

An illustration of a penguin standing on a sidewalk, looking up at a tall, ornate street lamp. The lamp has a glowing yellow light. The background is a blue sky with white clouds. The scene is set on a sidewalk next to a light-colored wall.

BEAUTY FROM EVERY ANGLE

Angular analysis of the
 $B^0 \rightarrow K^{*0} e^+ e^-$ decay
at LHCb

ALICE BIOLCHINI

BEAUTY FROM EVERY ANGLE

Angular analysis of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay channel
using Run1 and Run2 LHCb datasets

Alice Biolchini

2025

Copyright © Alice Biolchini, Amsterdam 2025, all rights reserved.

Beauty from every angle

*Angular analysis of the $B^0 \rightarrow K^{*0}e^+e^-$ decay channel
using Run1 and Run2 LHCb datasets*

Doctorial Thesis, VRIJE UNIVERSITEIT VAN AMSTERDAM

Cover artwork: Alice Biolchini

Description: The artwork is inspired by the surrealist style of René Magritte and realised digitally using the software Procreate.

On the **front cover**, a penguin gazes at the title of the thesis, written on the corner of a building. This scene symbolically recalls the rare penguin decays studied in this thesis. A streetlamp from the canals of Amsterdam is also depicted, representing the city where I completed my PhD.

The **back cover** features a geometric illustration of the LHCb detector. The collision between protons is surrealistically represented by a collision between tortellini, a reference to my Emilian roots.

Printed by Gildeprint

DOI: <https://doi.org/10.5463/thesis.1360>

ISBN: 978-94-6496-446-2



This publication is part of the research programme of the Netherlands Organisation for Scientific Research (NWO). The work is carried out at the National Institute of Subatomic Physics (Nikhef) in Amsterdam, The Netherlands.

VRIJE UNIVERSITEIT

BEAUTY FROM EVERY ANGLE

Angular analysis of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay channel using Run1 and Run2 LHCb datasets

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor of Philosophy aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. J.J.G. Geurts,
volgens besluit van de decaan
van de Faculteit der Bètawetenschappen
in het openbaar te verdedigen
op donderdag 18 september 2025 om 11.45 uur
in de universiteit

door

Alice Biolchini

geboren te Pavullo Nel Frignano, Italië

promotor:	prof.dr. H.G. Raven
copromotor:	dr. M. Senghi Soares
promotiecommissie:	prof.dr. W.D. Hulsbergen prof.dr. D. Boer prof.dr. P. Ferrari dr. M.H. Schune dr. E. Dall'Occo

*To my big, loud family,
including my grandmother,
la nonna Alma,
who passed away during my PhD.*

Contents

Introduction	vii
1 Theoretical framework and experimental overview	1
1.1 The Standard Model of particle physics	1
1.1.1 The SM Lagrangian	2
1.1.2 The Cabibbo–Kobayashi–Maskawa mechanism	4
1.1.3 Shortcomings of the SM	5
1.2 Flavour changing neutral currents	5
1.2.1 Effective field theory in flavour physics	6
1.2.2 q^2 structure in $b \rightarrow s\ell\ell$ decays	7
1.3 The $B^0 \rightarrow K^{*0}e^+e^-$ decay	8
1.3.1 The angular basis	9
1.3.2 Angular PDF	10
1.3.3 Angular observables reparametrisation	12
1.4 Experimental overview of $b \rightarrow s\ell\ell$ decays	13
1.4.1 Beauty production	13
1.4.2 State-of-the-Art measurements	14
1.5 Context and motivation of the current work	20
2 The LHCb experiment at the LHC	23
2.1 The Large Hadron Collider	23
2.2 LHCb detector	25
2.3 Tracking system	26
2.4 Particle identification	28
2.4.1 Particle identification system	28
2.4.2 Particle identification	30
2.5 LHCb Trigger	30
2.6 Electrons in LHCb	31
2.7 $B^0 \rightarrow K^{*0}e^+e^-$ nomenclature and trigger lines	33
2.7.1 <i>Brem</i> categories definition	34
3 Data samples and first stage selection	35
3.1 $B^0 \rightarrow K^{*0}e^+e^-$ candidates	35
3.2 Online selection: trigger lines	36
3.3 Offline selection	36
3.3.1 Stripping	36

3.3.2	Preselection	36
3.3.3	Backgrounds vetoes	38
3.3.4	K^{*0} selection	40
3.3.5	B^0 mass region	40
3.3.6	q^2 selection	40
3.4	Residual backgrounds	41
3.5	Samples of the analysis	43
3.5.1	MC simulations and corrections	43
3.5.2	Misidentified background sample	45
4	Selection optimization	51
4.1	Gradient Boosted Decision Trees (BDT)	51
4.2	BDT1	52
4.2.1	Input features	52
4.2.2	Training samples and BDT score	53
4.2.3	Data to MC agreement of the features	54
4.2.4	Figure of Merit (FoM) for signal-to-background optimisation	56
4.2.5	Signal efficiency and backgrounds rejection	56
4.3	BDT2 targeting backgrounds at high $\cos \theta_\ell$	58
4.3.1	Features	58
4.3.2	Data to MC agreement of the features	58
4.3.3	Training samples and BDT score	59
4.4	2D Optimisation of the cut on BDTs output	60
4.5	Signal efficiency and backgrounds rejection	61
4.6	Considerations	62
5	Analysis strategy	65
5.1	Angular acceptance	65
5.1.1	Acceptance and angular PDF	65
5.1.2	Acceptance parameterisation	68
5.2	Mass model	69
5.2.1	Signal modelling	69
5.2.2	Background modelling	70
5.2.3	Results	73
5.3	Angular fit	74
5.3.1	Angular model on signal MC simulation	76
6	Expected sensitivity and systematic uncertainties	81
6.1	Fit strategy validation using toy simulations	81
6.1.1	Limited sample size biases with signal-only toys	81
6.1.2	Realistic toys	83
6.2	Systematic uncertainties	87
6.2.1	Fit bias	88
6.2.2	$B^+ \rightarrow K^+ e^+ e^-$ background	89
6.2.3	Acceptance parameterisation	90
6.2.4	DSL and combinatorial backgrounds fitting strategy	91

6.2.5	Misidentified background fitting strategy	94
6.2.6	BDT	96
6.2.7	Robustness checks	97
6.2.8	Summary of the systematic uncertainties	98
6.3	Blinded fit to the collected data	100
7	Discussion and outlooks	107
7.1	Sensitivity comparison between analyses	107
7.2	Conclusions and Outlooks	119
Appendix A Misidentified electron background samples		121
Appendix B BDTs features		129
B.1	BDT1	129
B.2	BDT2	131
Appendix C Fit components mass shapes		133
C.1	Signal	133
C.2	Backgrounds	135
Appendix D Additional material on Chap. 6		139
D.1	Pulls distributions of signal-only toys	140
D.2	Pull distributions of realistic toys	143
D.3	Blinded 4D-fit	146
Appendix E Folding studies		147
Appendix F Alternative strategies		149
E1	Selection on $\cos \theta_\ell$	149
Appendix G Physical regions of the angular parameters		153
Bibliography		157
Summary		165
Samenvatting		171
Riassunto		177

Introduction

“There is something strange and unsettling about humanity’s relationship with the cosmos that is rarely mentioned: the fact that the universe is actually dark. There is no light, the stars do not shine, the Sun is not luminous, the Moon does not reflect its rays. Everything is black, frighteningly black. Why? Because light exists only if there are eyes and a brain capable of transforming electromagnetic waves into luminous signals, as the human brain does.” - Piero Angela, Superquark

These words are not just about the visible universe but also about knowledge itself. What we do not yet know remains in darkness - unseen, incomprehensible. But all it takes is a new way of observing to illuminate what was once hidden. To me, fundamental-particle detectors like LHCb represent exactly that: a new eye through which we can scrutinise the universe, lighting it up with new insights and unveiling secrets that would otherwise remain out of reach.

Dear reader, let me tell you right away: this thesis will not reveal anything extraordinary or mysterious. However, through technical details and nuances, it lifts a small part of the veil of Maya, offering a quantitative measure of how precisely and accurately we can grasp something new.

The protagonist of this dissertation is a flavoured particle known as the “Beauty” particle. This particle consists of a bottom quark, b , and a down quark d , in the specific case of this analysis¹. This particle decays spontaneously almost instantly, on average after 1.5 picoseconds. However, due to the high energy at which it is produced, the distance it travels before decaying is dilated, meaning it travels a few centimetres inside our detector after being produced. This allows us to detect its decay vertex within our detector, which is placed about 8 millimetres away from the beam collision point. This is no coincidence. The LHCb detector was specifically designed to be sensitive to Beauty particles to enable us to study its properties in detail. Once in a while - on average, once in every 6 million Beauty particles produced - it decays into a K^{*0} and two oppositely charged electrons, which is the decay that is the subject of this dissertation. The peculiarity of the decay $B^0 \rightarrow K^{*0} e^+ e^-$ is that the quark b , composing the B^0 and with electrical charge $-1/3$, goes into a quark s with the same electrical charge and two electrons of opposite charges. These transitions are called “flavour-changing neutral current”, as the quark changes its “flavour” but the net total exchanged charge is zero. According to the Standard Model, they occur as shown by the Feynman diagram in **Fig. 1**. Diagrams of this type are known as “Penguin” diagrams [1], and the loop at the centre of the diagram is characteristic of these decays and explains why they are so rare.

¹For the sake of completeness, this particle exists as matter and its antimatter counterpart: B^0 and $\bar{B}^0$ with a total electrical charge equal to zero. They respectively consist of an anti-bottom quark, $\bar{b}$ and a down quark, d , or a b and $\bar{d}$. This Thesis do not make a distinction between them and considers both.

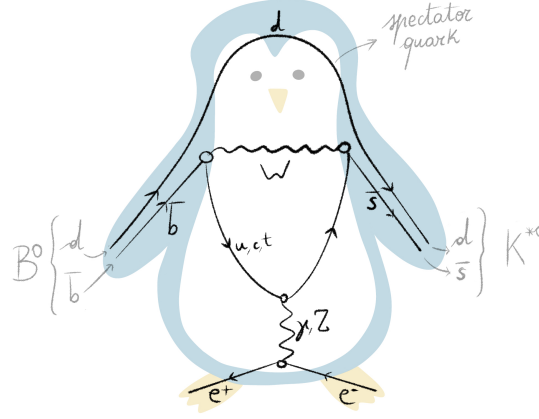


Figure 1: Feynman diagram for the $B^0 \rightarrow K^{*0} e^+ e^-$, mediated by flavour changing neutral current. This is also known as “Penguin” diagram describing the transition of a quark x to another quark y with the same charge. The intermediate quarks have charge differing from $Q(x) = Q(y)$ by one unit [1]. Normally the “penguin” diagram does not show the spectator quark, here is reported just for clarity purposes.

By analysing this decay as a function of the invariant mass squared of the di-lepton system, q^2 , we gain sensitivity to several Wilson Coefficients - parameters that encode effects which originate at higher energy scales. Put simply, these coefficients can hint at new, yet-unknown structures in physics, often referred to as “New Physics”, existing at energies beyond those currently accessible. Measurements like this are known as “indirect searches”: rather than searching for a direct signal of something new, they examine well-established phenomena and compare the results with theoretical predictions, searching for traces that could reveal the presence of unknown physics. If New Physics exists, it could manifest itself as deviations from these expectations. “Penguin” transitions are particularly suitable for indirect searches because of their rarity: even a tiny contribution from New Physics could lead to a noticeable deviation from the Standard Model prediction. The q^2 spectrum features several distinct regions of interest. This thesis focuses on the interval $[1.1, 6.0] \text{ GeV}^2/c^4$ and its separation in two smaller regions: $[1.10, 3.65] \text{ GeV}^2/c^4$ and $[3.65, 6.00] \text{ GeV}^2/c^4$. This analysis aims to measure the angular parameters describing the decay and compare them with Standard Model predictions. In fact, the $B^0 \rightarrow K^{*0} e^+ e^-$ decay is fully characterized by three angles ($\theta_K, \theta_\ell, \phi$)²: the K^{*0} particle decays into a $K\pi$ pair, defining θ_K , the electrons pair defines θ_ℓ , and the angle between the planes defined by the two pairs defines ϕ . The angular parameters are extracted by fitting the three joint angular distributions. At the time of writing, the analysis is still ongoing, and the angular distributions in data - as well as the corresponding angular parameters - remain blinded to prevent potential human biases. Therefore, the primary objective of this thesis is to assess the precision and accuracy of the angular parameters in the three aforementioned q^2 regions, two of which have never before been explored separately in this decay channel. *Chapter 1* provides a more detailed explanation of the theoretical foundations of this study, from the Standard Model of particle physics to the angular description of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay.

²The mathematical definition of the angles is given in **Sec. 1.3.1**

Chapter 2 explains the functioning of the LHCb experiment, the “eye and brain” that allows us to convert the information from pp collisions into data that we can analyse.

Chapter 3 presents the dataset used in the analysis and the tailored selection criteria applied to enhance the purity of the decay of interest over the background events present in the data. The chapter concludes with a detailed description of the residual background surviving the selection.

Chapter 4 describes the optimisation techniques employed to refine the selection, utilising boosted decision trees to further suppress backgrounds that persist after the initial selection.

Chapter 5 outlines the strategy adopted in this thesis to measure the angular parameters. In particular, it focuses on the acceptance function, which accounts for the fact that the detector cannot contain all the particles produced and distorts the angular distributions of the decay. Additionally, it presents the probability distribution functions used to model the reconstructed B^0 invariant mass and angular distributions, and the fit strategy to extract the angular parameters describing the decay.

Chapter 6 evaluates key systematic uncertainties related to strategy and design choices, necessary to provide a complete overview of the expected sensitivity, namely both systematic and statistical uncertainties of the adopted methodology in each q^2 region of interest in this decay channel.

Finally, *Chapter 7* concludes this work with final remarks and reflections from the author.

For readers interested in additional details, the appendices provide further insights into various technical aspects of the analysis.

1 | Theoretical framework and experimental overview

*This chapter provides a concise overview of the Standard Model (SM), with particular focus on Flavour Changing Neutral Currents (FCNCs) and the $b \rightarrow s \ell \ell$ decays, which are promising for probing potential New Physics (NP) beyond the SM. It presents the state-of-the-art measurements in this area and provides a comprehensive picture of the current experimental landscape. The chapter concludes with an in-depth discussion of the necessary concepts for the understanding of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay mode and its angular analysis.*

1.1 The Standard Model of particle physics

The Standard Model (SM) of particle physics is a quantum field theory that describes the electroweak and strong interactions between elementary particles. It is based on three key elements [2]:

- The symmetries of the Lagrangian,
- The representations of fermions and scalars,
- The pattern of spontaneous symmetry breaking.

The gauge symmetry of the SM is expressed as:

$$G_{SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \quad (1.1)$$

where $SU(3)_C$ represents the symmetry group of Quantum Chromodynamics (QCD), which governs the strong interactions, with the subscript C indicating the colour charge. The group $SU(2)_L \otimes U(1)_Y$ describes the symmetry of the electroweak interactions, as introduced by the Glashow-Weinberg-Salam theory, with L and Y referring to the weak interaction chirality and hypercharge, respectively.

The representations of the fundamental particles under these symmetries are [2]:

$$Q_{Li}^I(3, 2)_{+1/6}, \quad u_{Ri}^I(3, 1)_{+2/3}, \quad d_{Ri}^I(3, 1)_{-1/3}, \quad L_{Li}^I(1, 2)_{-1/2}, \quad \ell_{Ri}^I(1, 1)_{-1}, \quad (1.2)$$

where $i = 1, 2, 3$ denotes the generation of fermions, L (R) indicates left (right) chirality, and I represents the interaction eigenstates. Finally, there is a single scalar multiplet:

$$\phi(1, 2)_{+1/2}. \quad (1.3)$$

This notation is compact yet informative, clearly displaying the quantum numbers of the fields. For instance, the left-handed quarks Q_L^I form a triplet under $SU(3)_C$, a doublet under $SU(2)_L$, and have hypercharge $Y = +1/6$. The numbers in parentheses denote the dimensions of the representations of $SU(3)_C$ and $SU(2)_L$, while the subscript indicates the hypercharge. The $SU(2)_L$ doublets can be decomposed as follows:

$$Q_{Li}^I = \begin{pmatrix} u_{Li}^I \\ d_{Li}^I \end{pmatrix}, \quad L_{Li}^I = \begin{pmatrix} \nu_{Li}^I \\ \ell_{Li}^I \end{pmatrix}.$$

Usually, these particles are identified under two main groups:

- *quarks* ($Q_{Li}^I, u_{Ri}^I, d_{Ri}^I$) that are $SU(3)_C$ triplets and interact with the strong force;
- *leptons* (L_{Li}^I, ℓ_{Ri}^I) that are $SU(3)_C$ singlets.

In addition, it is possible to use a different classification:

- *left-handed* (Q_{Li}^I, L_{Li}^I) that are $SU(2)_L$ doublets and interact with W^i bosons;
- *right-handed* ($u_{Ri}^I, d_{Ri}^I, \ell_{Ri}^I$) that are $SU(2)_L$ singlets, therefore they do not interact with W^i bosons.

W^i (with $i = 1, 2, 3$) together with B are the gauge fields coming from the electroweak invariance under $SU(2)_L \otimes U(1)_Y$. A non-zero vacuum expectation value (VEV) of the additional scalar field ϕ leads to a spontaneous breaking of the $SU(2)_L \otimes U(1)_Y$ symmetry causing the W^3 and B bosons to mix, which leads to the physically observed Z^0 and A bosons. The expression of the VEV for the scalar field is:

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix}, \quad (1.4)$$

consequently, the gauge symmetry group is spontaneously broken:

$$G_{SM} \rightarrow SU(3)_C \otimes U(1)_{EM}, \quad (1.5)$$

where $U(1)_{EM}$ is the symmetry group of electromagnetism. The field ϕ is often parameterised as

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + H(x) \end{pmatrix} \quad (1.6)$$

where $H(x)$ is a neutral scalar field known as the Higgs boson. This symmetry breaking must occur in such a way that the $U(1)_{EM}$ symmetry remains unbroken, ensuring that the resulting gauge boson A - identified with the photon - remains massless, consistent with experimental observations.

1.1.1 The SM Lagrangian

The SM Lagrangian, consistent with the gauge symmetry group G_{SM} given in **Eq. 1.1**, is divided into three main components:

$$\mathcal{L}_{SM} = \mathcal{L}_{kin} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}. \quad (1.7)$$

The kinetic term, $\mathcal{L}_{kin}$, comprises the kinetic terms for fermions, represented by $i\bar{\psi}\gamma^\mu\partial_\mu\psi$. To preserve local gauge invariance, the standard derivative is replaced by a covariant derivative:

$$D^\mu = \partial^\mu + ig_s G_a^\mu L_a + ig W_b^\mu T_b + ig' B^\mu Y, \quad (1.8)$$

where

- L_a are the generators of the $SU(3)_C$ group: $\frac{1}{2}\lambda_a$ for triplets (with λ_a being the eight 3×3 Gell-Mann matrices) and 0 for singlets. G_a^μ denote the eight gluon fields;
- T_b are the generators of the $SU(2)_L$ group: $\frac{1}{2}\sigma_b$ for doublets (where σ_b are the three 2×2 Pauli matrices) and 0 for singlets. W_b^μ are the three weak interaction bosons;
- Y represents the $U(1)_Y$ charges and B^μ is the single hypercharge boson.

The Higgs potential, $\mathcal{L}_{Higgs}$, describes the scalar self-interactions of the Higgs field:

$$\mathcal{L}_{Higgs} = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \quad (1.9)$$

where λ is the self-coupling constant and $\mu^2 = -v^2\lambda$. If μ^2 is negative, there are infinitely many vacuum expectation values (VEVs) different from zero ($v \neq 0$). These infinitely many values are directly related to a $U(1)$ symmetry which guarantees the photon to be massless, as mentioned above.

The Yukawa term, $\mathcal{L}_{Yukawa}$, describes the interactions between fermions and the scalar field and is given by:

$$\mathcal{L}_{Yukawa} = -Y_{ij}^d \bar{Q}_{Li}^I \phi d_{Rj}^I - Y_{ij}^u \bar{Q}_{Li}^I \bar{\phi} u_{Rj}^I - Y_{ij}^\ell \bar{L}_{Li}^I \phi \ell_{Rj}^I + \text{h.c.}, \quad (1.10)$$

where $Y_{ij}^{u,d,\ell}$ are 3×3 complex matrices, $i, j = 1, 2, 3$ are generation indices, and $\bar{\phi} = i\sigma^2 \phi^\dagger$. The fermion mass terms emerge when ϕ acquires a vacuum expectation value, v (see **Eq. 1.4**). By substituting **Eq. 1.6** into **Eq. 1.10** and neglecting interaction terms, the mass terms can be written as:

$$\mathcal{L}_M = -(M_d)_{ij} \bar{d}_{Li}^I d_{Rj}^I - (M_u)_{ij} \bar{u}_{Li}^I u_{Rj}^I - (M_\ell)_{ij} \bar{\ell}_{Li}^I \ell_{Rj}^I, \quad (1.11)$$

where $M_{d,u,\ell} = \frac{v}{\sqrt{2}} Y_{d,u,\ell}$. The physical mass eigenstates are obtained by diagonalizing $M_{d,u,\ell}$:

$$V_{fL} M_f V_{fR}^\dagger = M_f^{diag} \quad \text{for } f = d, u, \ell \quad (1.12)$$

where V_{fL} and V_{fR} are unitary matrices and M_f^{diag} is diagonal and real. The mass eigenstates are thus identified as:

$$d_{Li} = (V_{dL})_{ij} d_{Lj}^I, \quad d_{Ri} = (V_{dR})_{ij} d_{Rj}^I, \quad (1.13)$$

$$u_{Li} = (V_{uL})_{ij} u_{Lj}^I, \quad u_{Ri} = (V_{uR})_{ij} u_{Rj}^I, \quad (1.14)$$

$$\ell_{Li} = (V_{\ell L})_{ij} \ell_{Lj}^I, \quad \ell_{Ri} = (V_{\ell R})_{ij} \ell_{Rj}^I, \quad (1.15)$$

$$\nu_{Li} = (V_{\nu L})_{ij} \nu_{Lj}^I. \quad (1.16)$$

1.1.2 The Cabibbo–Kobayashi–Maskawa mechanism

Within the SM, the CP symmetry - invariance under the combined action of the charge conjugation (C) and the parity (P) - is preserved in strong and electromagnetic interactions, but violated in weak interactions: in the SM Lagrangian the CP symmetry is broken by an irreducible complex physical phase in the Yukawa quark-term [3, 4]. From the $\mathcal{L}_{kin}$ term, the charge current interaction between the quarks and charged $SU(2)_L$ gauge bosons, $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp W_\mu^2)$, can be expressed as functions of the weak interaction eigenstates

$$\mathcal{L}_{W^\pm} = \frac{-g}{\sqrt{2}} \overline{Q}_{Li}^I \gamma^\mu W_\mu^\pm Q_{Li}^I + \text{h.c.} \quad (1.17)$$

CP asymmetry arises in this equation because only left-handed quarks participate, as the charged weak bosons $W^\pm$ preserve the $SU(2)_L$ symmetry and couple exclusively to left-handed particles, unlike the Z^0 boson. Expressing it as a function of the mass eigenstates (Eq. 1.13 – Eq. 1.16), the $\mathcal{L}_{W^\pm}$ becomes

$$\mathcal{L}_{W^\pm} = \frac{-g}{\sqrt{2}} \overline{u}_{Li} V_{uL} \gamma^\mu W_\mu^\pm V_{dL}^\dagger d_{Li}^I + \text{h.c.}, \quad (1.18)$$

where u_i stands for the up type quarks (u, c, t) and d_i stands for the down type, (d, s, b). Therefore, one can write:

$$\mathcal{L}_{W^\pm} = \frac{-g}{\sqrt{2}} (\overline{u}_L, \overline{c}_L, \overline{t}_L) \gamma^\mu W_\mu^\pm V_{CKM} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} + \text{h.c.}.$$

The V_{CKM} is the product between the unitary matrices V_{uL} and $V_{dL}^\dagger$, hence it is a 3×3 unitary matrix, known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix [3, 4]:

$$V_{CKM} \equiv V_{uL} V_{dL}^\dagger = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}.$$

The CKM matrix can be parameterised by three mixing angles and a complex phase, as reported below

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{-i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix},$$

where $s_{ij} = \sin \theta_{ij}$, $c_{ij} = \cos \theta_{ij}$ and δ is the phase, above mentioned, responsible for all CP -violating phenomena in flavour-changing processes in the SM [3].

The CKM matrix elements are fundamental parameters of the SM, their precise determination is important to make predictions and to put strong constraints on beyond SM theories. The magnitudes of the CKM matrix elements can be determined directly by measuring the decay rates of tree-level processes, *e.g.* by measuring branching ratios and lifetimes. In this way, it is possible to directly extract the value of $|V_{ij}|$. However, $|V_{td}|$ and $|V_{ts}|$ cannot be measured directly, as the processes in which they appear are loop-suppressed and occur at very low rates. In these cases, indirect measurements are performed using a higher-order process. Combining both the direct and indirect measurements, the up-to-date magnitudes of the CKM matrix elements are

the following [5]:

$$V_{CKM} = \begin{pmatrix} 0.97435 \pm 0.00016 & 0.22501 \pm 0.00068 & 0.003732^{+0.000090}_{-0.000085} \\ 0.22487 \pm 0.00068 & 0.97349 \pm 0.00016 & 0.04183^{+0.00079}_{-0.00069} \\ 0.00858^{+0.00019}_{-0.00017} & 0.04111^{+0.00077}_{-0.00068} & 0.999118^{+0.000029}_{-0.000034} \end{pmatrix}. \quad (1.19)$$

The GIM mechanism [6] ensures that direct transitions within the same type of quark (either up-type or down-type), namely quarks of the same charge but different flavor (such as $s \rightarrow d$ or $b \rightarrow s$), are forbidden at tree level, while transitions that change both flavor and charge of the quark are allowed. The illustration in **Fig. 1.1** gives a graphical representation of the CKM matrix elements' order of magnitude.

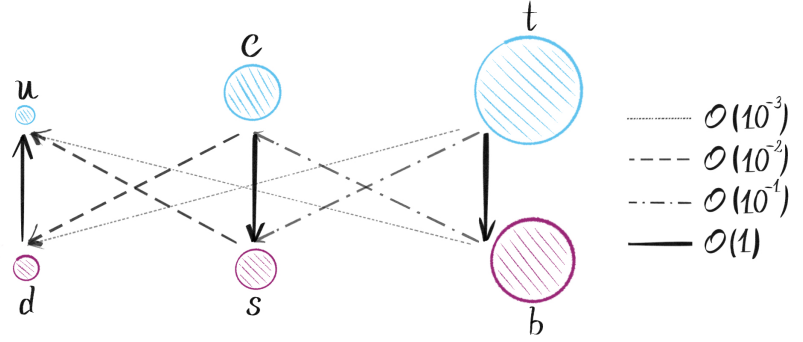


Figure 1.1: Illustration of the magnitudes of the CKM elements. Up-type quarks are represented in light blue, and down-type quarks are in purple.

1.1.3 Shortcomings of the SM

The SM is currently the most robust framework for describing the subatomic world, yet it remains incomplete. While the SM successfully integrates three of the four fundamental forces, it notably excludes gravity. The nature of dark matter - despite substantial evidence for its existence - is not accounted for within the SM. Additionally, neutrino masses, implied by the observation of neutrino oscillations, are not incorporated in the model. The underlying reasons for the existence of three generations of quarks and leptons, as well as the hierarchy of mass scales, remain unresolved mysteries. These limitations suggest the existence of a more comprehensive theory that encompasses the SM as a low-energy approximation.

1.2 Flavour changing neutral currents

Rare decays, which correspond to highly suppressed amplitudes within the SM, serve as excellent laboratories for testing the model and exploring potential New Physics (NP) beyond it. Generally they are mediated by loops of virtual particles, where heavy particles can contribute to the loop. Additionally, due to their rarity, even a small contribution from NP could result in a sizable deviation from the predictions of the SM. In particular, flavour-changing neutral current (FCNC) decays are considered golden channels for indirect NP searches. These processes involve transitions between two up-type quarks and two down-type quarks. As seen in **Sec. 1.1.2**, a direct transition between quarks of the same type is not allowed according to the SM [6]. Therefore, they

can only occur through loop or box diagrams, mediated by an electroweak boson. In addition to the loop suppression, these decays are also suppressed by the CKM matrix element involved, as shown by **Fig. 1.1**. As a result, these processes are extremely suitable for indirect searches.

The $b \rightarrow s\ell\ell$ transitions are a specific example of FCNCs, these SM Feynman diagrams are shown in **Fig. 1.2**. The $b \rightarrow s\ell\ell$ decays provide model-independent tests for NP models that alter the known properties of these decays, such as their decay rates or angular distributions.

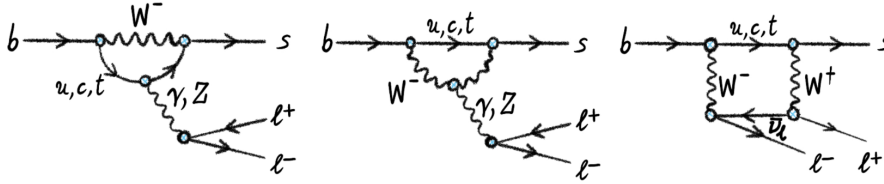


Figure 1.2: Illustration of the Feynman diagrams for the $b \rightarrow s\ell\ell$ decays allowed in the SM.

1.2.1 Effective field theory in flavour physics

Reliable SM predictions are essential to perform SM probes on these decays. However, obtaining accurate theoretical predictions for some observables can be highly challenging. The primary difficulty arises from the contributions of widely separated energy scales involved in these decays. The FCNC transition is governed by gauge boson interactions at the electroweak scale, $m_W \sim 80$ GeV. On the other hand, the energy scale associated with the beauty-hadron decay is approximately 5 GeV. Finally, the energy scale of strong interactions between hadrons is at the QCD scale, $\Lambda_{\text{QCD}} \sim 0.2$ GeV. These widely separated energy scales complicate the computation of the rare decay amplitudes because they involve the presence of large logarithms, making the theoretical calculations particularly difficult [7].

The multi-scale issue can be tackled with the help of Effective Field Theory (EFT) methods. They offer a model-independent way to calculate the amplitudes of the $b \rightarrow s\ell\ell$ decays. The EFT is a quantum field theory that is valid only in a limited energy range. The idea is that the b decay formalism can be expressed using the operator product expansion which integrates out all the degrees of freedom above a certain energy scale μ . If μ is taken large enough, the weak interactions can be approximated as point-like from the point of view of the hadronic scales m_b and Λ_{QCD} [7].

The effective Hamiltonian for a $b \rightarrow s\ell\ell$ transition is given by [8]

$$\mathcal{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}} \left(\lambda_t \mathcal{H}_{\text{eff}}^{(t)} + \lambda_u \mathcal{H}_{\text{eff}}^{(u)} \right) \quad (1.20)$$

where G_F is the Fermi constant, λ_i ($i = t, u$) are the CKM matrix elements, $V_{ib} V_{is}^\dagger$. The contribution $\mathcal{H}_{\text{eff}}^{(u)}$ is doubly Cabibbo-suppressed, therefore it can be neglected for simplicity. Then, it follows

that the B decay amplitude to a final state f , $A(B \rightarrow f)$, can be written schematically as:

$$A(B \rightarrow f) = \langle f | \mathcal{H}_{\text{eff}} | B \rangle = -\frac{4G_F}{\sqrt{2}} \frac{\alpha_e}{4\pi} V_{tb} V_{ts}^\dagger \sum_i C_i(\mu) \langle f | \mathcal{O}_i(\mu) | B \rangle, \quad (1.21)$$

where the operators $\mathcal{O}_i$, together with their chiral-flipped counterparts ($\mathcal{O}'_i$), are detailed below:

$$\mathcal{O}_7 = \frac{m_b}{e} \bar{s} \sigma^{\mu\nu} P_R b F_{\mu\nu}, \quad \mathcal{O}'_7 = \frac{m_b}{e} \bar{s} \sigma^{\mu\nu} P_L b F_{\mu\nu} \quad (1.22)$$

$$\mathcal{O}_9 = (\bar{s} \gamma_\mu P_L b) (\bar{\ell} \gamma^\mu \ell), \quad \mathcal{O}'_9 = (\bar{s} \gamma_\mu P_R b) (\bar{\ell} \gamma^\mu \ell), \quad (1.23)$$

$$\mathcal{O}_{10} = (\bar{s} \gamma_\mu P_L b) (\bar{\ell} \gamma^\mu \gamma_5 \ell), \quad \mathcal{O}'_{10} = (\bar{s} \gamma_\mu P_R b) (\bar{\ell} \gamma^\mu \gamma_5 \ell). \quad (1.24)$$

The semileptonic operators $\mathcal{O}_9$ and $\mathcal{O}_{10}$, along with the operator $\mathcal{O}_7$, govern the $b \rightarrow s \ell \ell$ transitions. This last operator, $\mathcal{O}_7$, represents the electromagnetic dipole operator, and it primarily governs the $b \rightarrow s \gamma$ transition but also contributes, albeit less dominantly, to $b \rightarrow s \ell \ell$ transitions [9].

Thanks to this effective notation, which separates short- and long-distance contributions, these calculations can be managed more easily, allowing to focus on different physics effects independently. The operators $\mathcal{O}_i$ represent the low-energy (long-distance) manifestation of the $b \rightarrow s \ell \ell$ decays, corresponding to interactions such as those depicted in **Fig. 1.3**. Meanwhile, the Wilson coefficients, C_i , encode the effects of the high-energy (short-distance) scale. The key point is that NP at energy scale $\Lambda_{\text{NP}} \gg m_W$ can be described by an EFT and its presence will manifest itself by the need to introduce new operators ($\mathcal{O}_{\text{NP}}$) or modified existing Wilson coefficients.

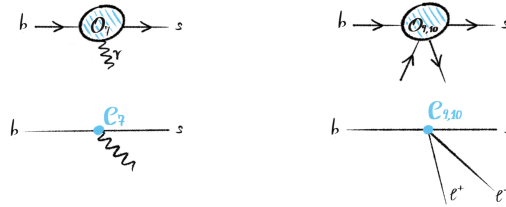


Figure 1.3: Illustration of the semileptonic electroweak penguin operators $\mathcal{O}_{9,10}$, and the radiative dipole operator $\mathcal{O}_7$. The diagrams on top highlight operators, while the diagrams at the bottom highlight the Wilson Coefficients $C_{7,9,10}$, namely the coupling constants of these interactions.

1.2.2 q^2 structure in $b \rightarrow s \ell \ell$ decays

The decay $b \rightarrow s \ell \ell$ can be analysed as a function of the dilepton invariant mass squared, denoted as q^2 . Different regions of q^2 are sensitive to specific Wilson coefficients (WCs) due to the various underlying decay mechanisms. The differential branching fraction as a function of q^2 is illustrated in **Fig. 1.4**, along with the WCs indicating the regions where they contribute (**Sec. 1.2.1**).

At low q^2 , the decay is influenced by the electromagnetic current, giving rise to the “photon pole”, a characteristic peak caused by the strong coupling between vector mesons (e.g., K^{*0} , with $J^P = 1^-$) and virtual photons. This peak is absent in decays involving pseudoscalar mesons (e.g., $K^\pm$, with $J^P = 0^-$).

Beyond the photon pole, a non-resonant region allows direct access to the rare $b \rightarrow s \ell \ell$ decay.

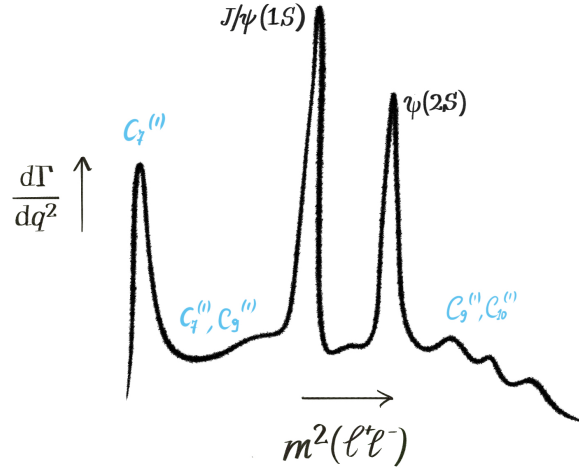


Figure 1.4: Illustration of the differential branching fraction for a $b \rightarrow s\ell\ell$ decay as a function of the dielectron invariant mass, a variable known as q^2 . The figure highlights which WCs dominate a certain q^2 range as well as the central charmonium resonances, J/ψ and $\Psi(2S)$.

At higher q^2 , two prominent charmonium resonances, J/ψ and $\psi(2S)$, appear due to tree-level decays into e^+e^- pairs, creating sharp peaks in the spectrum. The Feynman diagrams contributing to the $B^0 \rightarrow K^{*0}J/\psi$ ($\psi(2S)$) decays are shown in **Fig. 1.5**.

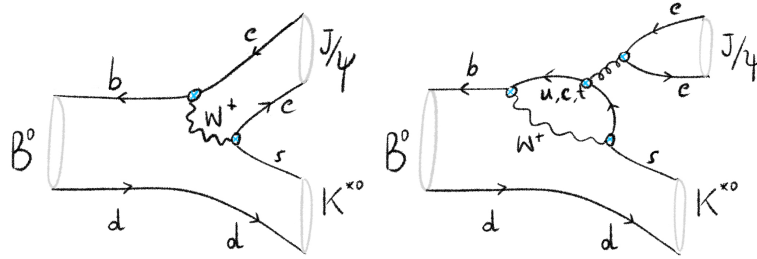


Figure 1.5: Tree-level (left) and loop-level(right) contributions to the charmonium resonances $B^0 \rightarrow K^{*0}J/\psi$ ($\psi(2S)$). The tree-level contribution is what makes these decays occur frequently, giving rise to prominent resonances. These illustrations also include the spectator quark, but the focus should remain on the flavour-changing currents.

The primary interest of this thesis lies in the “rare mode regions”, where the decay $b \rightarrow s\ell\ell$ occurs without resonance contributions. Analysing the decay as a function of q^2 is crucial, as it provides valuable information about different Wilson coefficients and helps probe the SM.

1.3 The $B^0 \rightarrow K^{*0}e^+e^-$ decay

The decay $B^0 \rightarrow K^{*0}e^+e^-$ is a $b \rightarrow s\ell^+\ell^-$ transition mediated by an FCNC. The $K^{*0}(892)$ is a resonance that promptly decays into a kaon and a pion, resulting in a final state with four particles. By analyzing this decay as a function of the dilepton invariant mass, q^2 , the measurement

becomes sensitive to several WCs. The q^2 spectrum, as discussed in **Sec. 1.2.2** and shown in **Fig. 1.4**, features several distinct regions of interest. A measurement in such a decay channel can probe a combination of WCs by focusing on these specific regions. In the low- q^2 region, WCs $C_7^{(\prime)}$ and $C_9^{(\prime)}$ are dominant, while in the high- q^2 region, the focus shifts to $C_9^{(\prime)}$ and $C_{10}^{(\prime)}$. Additionally, by studying the very-low q^2 region near the photon-pole, one can access the WC $C_7^{(\prime)}$.

This thesis focuses on the measurement of the angular parameters that characterise the $B^0 \rightarrow K^{*0} e^+ e^-$ decay. The angular description of the decay, together with the angular observables derived from this analysis, is detailed in this section.

1.3.1 The angular basis

The goal of an angular analysis is to measure the coefficients which describe the differential decay width as a function of the decay angles. The angular basis used in this analysis is illustrated in **Fig. 1.6**.

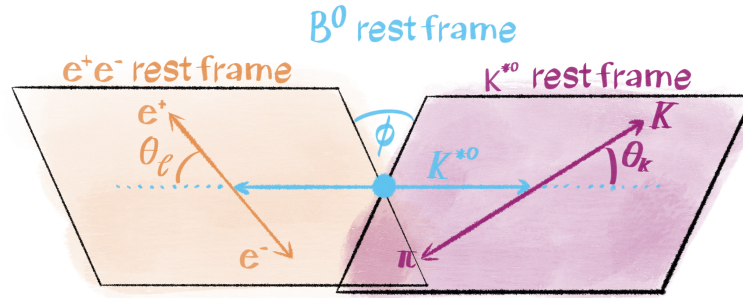


Figure 1.6: Illustration of the angular basis used to describe the decay.

The angle θ_ℓ is defined as the angle between the direction of the e^+ (e^-) in the dielectron rest frame and the direction of the dielectron in the B^0 ($\bar{B}^0$). For B^0 :

$$\cos \theta_\ell = \left(\hat{p}_{e^+}^{(e^+e^-)} \right) \cdot \left(\hat{p}_{e^+e^-}^{(B^0)} \right) = \left(\hat{p}_{e^+}^{(e^+e^-)} \right) \cdot \left(-\hat{p}_{B^0}^{(e^+e^-)} \right), \quad (1.25)$$

and for the $\bar{B}^0$:

$$\cos \theta_\ell = \left(\hat{p}_{e^-}^{(e^+e^-)} \right) \cdot \left(\hat{p}_{e^+e^-}^{(\bar{B}^0)} \right) = \left(\hat{p}_{e^-}^{(e^+e^-)} \right) \cdot \left(-\hat{p}_{\bar{B}^0}^{(e^+e^-)} \right), \quad (1.26)$$

The angle θ_K is defined as the angle between the direction of the kaon in the K^{*0} ($\bar{K}^{*0}$) rest frame and the direction of the K^{*0} ($\bar{K}^{*0}$) in the B^0 ($\bar{B}^0$) rest frame. For B^0 :

$$\cos \theta_K = \left(\hat{p}_{K^+}^{(K^{*0})} \right) \cdot \left(\hat{p}_{K^{*0}}^{(B^0)} \right) = \left(\hat{p}_{K^+}^{(K^{*0})} \right) \cdot \left(-\hat{p}_{B^0}^{(K^{*0})} \right) \quad (1.27)$$

and for the $\bar{B}^0$:

$$\cos \theta_K = \left(\hat{p}_{K^-}^{(\bar{K}^{*0})} \right) \cdot \left(\hat{p}_{\bar{K}^{*0}}^{(\bar{B}^0)} \right) = \left(\hat{p}_{K^-}^{(\bar{K}^{*0})} \right) \cdot \left(-\hat{p}_{\bar{B}^0}^{(\bar{K}^{*0})} \right) \quad (1.28)$$

The angle ϕ is the angle between the plane containing the e^+ and e^- and the plane containing the kaon and pion from the K^{*0} . The definition of the angle ϕ for the B^0 is given by:

$$\cos \phi = \left(\hat{p}_{e^+}^{(B^0)} \times \hat{p}_{e^-}^{(B^0)} \right) \cdot \left(\hat{p}_{K^+}^{(B^0)} \times \hat{p}_{\pi^-}^{(B^0)} \right), \quad (1.29)$$

$$\sin \phi = \left[\left(\hat{p}_{e^+}^{(B^0)} \times \hat{p}_{e^-}^{(B^0)} \right) \times \left(\hat{p}_{K^+}^{(B^0)} \times \hat{p}_{\pi^-}^{(B^0)} \right) \right] \cdot \hat{p}_{K^{*0}}^{(B^0)} \quad (1.30)$$

and for the $\bar{B}^0$:

$$\cos \phi = \left(\hat{p}_{e^-}^{(\bar{B}^0)} \times \hat{p}_{e^+}^{(\bar{B}^0)} \right) \cdot \left(\hat{p}_{K^-}^{(\bar{B}^0)} \times \hat{p}_{\pi^+}^{(\bar{B}^0)} \right), \quad (1.31)$$

$$\sin \phi = \left[\left(\hat{p}_{e^-}^{(\bar{B}^0)} \times \hat{p}_{e^+}^{(\bar{B}^0)} \right) \times \left(\hat{p}_{K^-}^{(\bar{B}^0)} \times \hat{p}_{\pi^+}^{(\bar{B}^0)} \right) \right] \cdot \hat{p}_{\bar{K}^{*0}}^{(\bar{B}^0)} \quad (1.32)$$

The $\hat{p}_X^{(Y)}$ are unit vectors describing the direction of a particle X in the rest frame of the system Y . The particle momenta are first boosted to the B^0 (or $\bar{B}^0$) rest frame. In this basis, the angular definition for the B^0 decay is a CP transformation of that for the $\bar{B}^0$ decay.

1.3.2 Angular PDF

The decay rate of $B^0 \rightarrow K^{*0} e^+ e^-$ is given by [10]:

$$\frac{d^4\Gamma(B^0 \rightarrow K^{*0} e^+ e^-)}{dq^2 d\vec{\Omega}} = \frac{9}{32\pi} \sum_i |I_i(q^2) f_i(\vec{\Omega})|^2, \quad (1.33)$$

and that of its charge conjugate $\bar{B}^0 \rightarrow \bar{K}^{*0} e^+ e^-$ by:

$$\frac{d^4\Gamma(\bar{B}^0 \rightarrow \bar{K}^{*0} e^+ e^-)}{dq^2 d\vec{\Omega}} = \frac{9}{32\pi} \sum_i |\bar{I}_i(q^2) f_i(\vec{\Omega})|^2. \quad (1.34)$$

In these expressions, the f_i terms are combinations of spherical harmonics that depend on $\vec{\Omega} = (\cos \theta_l, \cos \theta_K, \phi)$. The I_i and $\bar{I}_i$ terms are q^2 dependent angular observables that can be expressed as bilinear combinations of six complex decay amplitudes, $A_{\perp}^{L,R}$, $A_0^{L,R}$, and $A_{\parallel}^{L,R}$, which in turn can be expressed as functions of WCs ($\mathcal{C}_{7,9,10}$) and q^2 dependent hadronic form factors¹ [11, 12]. The decay amplitudes correspond to different transversality (parallel, perpendicular) states of the K^{*0} meson and chirality (left- and right-handed) states of the dilepton system. These six amplitudes correspond to resonant P-wave contribution from the vector meson K^{*0} decaying into the $K^+ \pi^-$ particles. However, the decay into $K^+ \pi^-$ can also receive contributions through an S-wave configuration [13]. This introduces additional complex amplitudes $\mathcal{A}_S^{R,L}$. All the angular coefficients $I_i(q^2)$ are listed in **Tab. 1.1** with the corresponding angular functions. The index $i \in [1 - 9]$ corresponds to P-wave, while $i \in [10 - 17]$ corresponds to S-wave contributions.

From these decay amplitudes, it is possible to define twelve CP averaged coefficients,

$$S_i = (I_i + \bar{I}_i) \left/ \frac{d(\Gamma + \bar{\Gamma})}{dq^2} \right. \quad (1.35)$$

¹Form factors are functions that describe the internal structure of composite particles during interactions with external fields or other particles, and they provide corrections to the simple point-like particle approximation. In the context of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay, form factors parameterise the hadronic matrix elements between the initial and final state hadrons, capturing how the internal structure of the hadrons influences the decay process.

i	$I_i(q^2)$	$f_i(\vec{\Omega})$
1s	$\frac{3}{4} [A_{\perp}^L ^2 + A_{\parallel}^L ^2 + A_{\perp}^R ^2 + A_{\parallel}^R ^2]$	$\sin^2 \theta_K$
1c	$[A_0^L ^2 + A_0^R ^2]$	$\cos^2 \theta_K$
2s	$\frac{1}{4} [A_{\perp}^L ^2 + A_{\parallel}^L ^2 + A_{\perp}^R ^2 + A_{\parallel}^R ^2]$	$\sin^2 \theta_K \cos 2\theta_{\ell}$
2c	$- [A_0^R ^2 + A_0^L ^2]$	$\cos^2 \theta_K \cos 2\theta_{\ell}$
3	$\frac{1}{2} [A_{\perp}^L ^2 - A_{\parallel}^L ^2 + A_{\perp}^R ^2 - A_{\parallel}^R ^2]$	$\sin^2 \theta_K \sin^2 \theta_{\ell} \cos 2\phi$
4	$\sqrt{\frac{1}{2}} \Re (A_0^L A_{\parallel}^{L*} + A_0^R A_{\parallel}^{R*})$	$\sin 2\theta_K \sin 2\theta_{\ell} \cos \phi$
5	$\sqrt{2} \Re (A_0^L A_{\perp}^{L*} - A_0^R A_{\perp}^{R*})$	$\sin 2\theta_K \sin \theta_{\ell} \cos \phi$
6s	$2 \Re (A_{\parallel}^L A_{\perp}^{L*} - A_{\parallel}^R A_{\perp}^{R*})$	$\sin^2 \theta_K \cos \theta_{\ell}$
7	$\sqrt{2} \Im (A_0^L A_{\parallel}^{L*} - A_0^R A_{\parallel}^{R*})$	$\sin 2\theta_K \sin \theta_{\ell} \sin \phi$
8	$\sqrt{\frac{1}{2}} \Im (A_0^L A_{\perp}^{L*} + A_0^R A_{\perp}^{R*})$	$\sin 2\theta_K \sin 2\theta_{\ell} \sin \phi$
9	$\Im (A_{\perp}^L A_{\parallel}^{L*} + A_{\perp}^R A_{\parallel}^{R*})$	$\sin^2 \theta_K \sin^2 \theta_{\ell} \sin 2\phi$
10	$\frac{1}{3} [A_S^L ^2 + A_S^R ^2]$	1
11	$\sqrt{\frac{4}{3}} \Re (A_S^L A_0^{L*} + A_S^R A_0^{R*})$	$\cos \theta_K$
12	$-\frac{1}{3} [A_S^L ^2 + A_S^R ^2]$	$\cos 2\theta_{\ell}$
13	$-\sqrt{\frac{4}{3}} \Re (A_S^L A_0^{L*} + A_S^R A_0^{R*})$	$\cos \theta_K \cos 2\theta_{\ell}$
14	$\sqrt{\frac{2}{3}} \Re (A_S^L A_{\parallel}^{L*} + A_S^R A_{\parallel}^{R*})$	$\sin \theta_K \sin 2\theta_{\ell} \cos \phi$
15	$\sqrt{\frac{8}{3}} \Re (A_S^L A_{\perp}^{L*} - A_S^R A_{\perp}^{R*})$	$\sin \theta_K \sin \theta_{\ell} \cos \phi$
16	$\sqrt{\frac{8}{3}} \Im (A_S^L A_{\parallel}^{L*} + A_S^R A_{\parallel}^{R*})$	$\sin \theta_K \sin \theta_{\ell} \sin \phi$
17	$\sqrt{\frac{2}{3}} \Im (A_S^L A_{\perp}^{L*} + A_S^R A_{\perp}^{R*})$	$\sin \theta_K \sin 2\theta_{\ell} \sin \phi$

Table 1.1: Dependence of the angular observables $I_i(q^2)$ on the transversity amplitudes $\mathcal{A}_{0,\parallel,\perp}^{L,R}$ and the corresponding angular terms $f_i(\cos \theta_{\ell}, \cos \theta_K, \phi)$ [12].

The normalisation condition implies $\frac{3}{4}(2S_{1s} + S_{1c}) - \frac{1}{4}(2S_{2s} + S_{2c}) = 1$. The electron rest mass of the final state electrons can be considered negligible for q^2 much larger than the electron rest mass. Therefore, in the limit of massless leptons, the CP -averaged observables are related by $S_{2c} = -S_{1c}$ and $S_{2s} = S_{1s}/3$. These conditions reduce the number of independent CP -averaged observables of the P-wave from eleven to eight. The observable S_{6s} is related to the forward-backward asymmetry, $A_{FB} = \frac{3}{4}S_{6s}$. Finally, the observable S_{1c} corresponds to the fraction of longitudinal polarisation of the K^{*0} meson. It is referred to as F_L , and it is defined in **Eq. 1.36**.

$$F_L = S_{1c} = \frac{|A_0^L|^2 + |A_0^R|^2}{|A_0^L|^2 + |A_0^R|^2 + |A_{||}^L|^2 + |A_{||}^R|^2 + |A_{\perp}^L|^2 + |A_{\perp}^R|^2}. \quad (1.36)$$

The CP -averaged angular distribution of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay can be written as [10]

$$\begin{aligned} \frac{1}{d^4(\Gamma + \bar{\Gamma})/dq^2} \frac{d^4(\Gamma + \bar{\Gamma})}{d\vec{\Omega}dq^2} &= \frac{9}{32\pi} \left[\frac{3}{4}(1 - F_L) \sin^2 \theta_K + F_L \cos^2 \theta_K \right. \\ &\quad + \frac{1}{4}(1 - F_L) \sin^2 \theta_K \cos 2\theta_\ell \\ &\quad - F_L \cos^2 \theta_K \cos 2\theta_\ell + S_3 \sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi \\ &\quad + S_4 \sin 2\theta_K \sin 2\theta_\ell \cos \phi + S_5 \sin 2\theta_K \sin \theta_\ell \cos \phi \\ &\quad + \frac{4}{3}A_{FB} \sin^2 \theta_K \cos \theta_\ell + S_7 \sin 2\theta_K \sin \theta_\ell \sin \phi \\ &\quad \left. + S_8 \sin 2\theta_K \sin 2\theta_\ell \sin \phi + S_9 \sin^2 \theta_K \sin^2 \theta_\ell \sin 2\phi \right]. \end{aligned} \quad (1.37)$$

If the S-wave is taken into account, six more terms need to be added, and the angular distribution is modified to [10]

$$\begin{aligned} \left. \frac{d\Gamma}{d\vec{\Omega}dq^2} \right|_{S+P} &= (1 - F_S) \left. \frac{d\vec{\Gamma}}{d\vec{\Omega}dq^2} \right|_P + \frac{d\Gamma}{dq^2} \frac{3}{16\pi} \times \\ &\quad \left[F_S \sin^2 \theta_\ell \right. \\ &\quad + S_{S1} \sin^2 \theta_\ell \cos \theta_K \\ &\quad + S_{S2} \sin 2\theta_\ell \sin \theta_K \cos \phi \\ &\quad + S_{S3} \sin \theta_\ell \sin \theta_K \cos \phi \\ &\quad + S_{S4} \sin \theta_\ell \sin \theta_K \sin \phi \\ &\quad \left. + S_{S5} \sin 2\theta_\ell \sin \theta_K \sin \phi \right], \end{aligned} \quad (1.38)$$

with F_S indicating the fraction of the S-wave:

$$F_S = \frac{|\mathcal{A}_S^L|^2 + |\mathcal{A}_S^R|^2}{|\mathcal{A}_S^L|^2 + |\mathcal{A}_S^R|^2 + |\mathcal{A}_0^L|^2 + |\mathcal{A}_0^R|^2 + |\mathcal{A}_{||}^L|^2 + |\mathcal{A}_{||}^R|^2 + |\mathcal{A}_{\perp}^L|^2 + |\mathcal{A}_{\perp}^R|^2} \quad (1.39)$$

1.3.3 Angular observables reparametrisation

The angular observables can be parameterised such that leading form-factor uncertainties cancel to first order. Although this set of reparameterised angular observables is not employed in this thesis, it is referenced in the context of other analyses. The basis of F_L , and the observables $P_i^{(\prime)}$

can be calculated from the observables S_i [14]:

$$\begin{aligned}
 P_1 &= 2 \frac{S_3}{1 - F_L}, & P_2 &= \frac{1}{2} \frac{S_6^s}{1 - F_L} = \frac{2}{3} \frac{A_{FB}}{1 - F_L}, \\
 P_3 &= -\frac{S_9}{1 - F_L}, & P'_4 &= \frac{S_4}{\sqrt{F_L(1 - F_L)}}, \\
 P'_5 &= \frac{S_5}{\sqrt{F_L(1 - F_L)}}, & P'_6 &= \frac{S_7}{\sqrt{F_L(1 - F_L)}}, \\
 P'_8 &= \frac{S_8}{\sqrt{F_L(1 - F_L)}}.
 \end{aligned} \tag{1.40}$$

1.4 Experimental overview of $b \rightarrow s\ell\ell$ decays

1.4.1 Beauty production

The studies of flavour physics through beauty-quark is currently performed in two main ways: through e^+e^- and pp collisions. These different colliders come with different production mechanisms and thus with largely varying production cross sections.

e^+e^- colliders Since the e^+e^- interaction occurs at a point-like level, the experimental environment is clean, with a well-defined interaction energy and a minimal number of interaction vertices per event. As a result, the background levels are very low, and the detection efficiency is extremely high, in comparison to a hadron collider. Colliders (e^+e^-) designed for flavour studies purposes are commonly referred to as **B-factories**. The B-factories operate at a centre-of-mass energy specifically tuned to produce the $\Upsilon(4S)$ resonance, which decays into quantum-correlated $B^0\bar{B}^0$ or B^+B^- pairs. The BaBar (PEP-II collider, SLAC, California) and Belle (KEKB collider, KEK, Japan) experiments are examples of B-factories. They have overcome the challenge of small cross-sections by achieving high integrated luminosities, approximately 500 fb^{-1} and 1000 fb^{-1} , respectively.

Hadron colliders At hadron colliders, the $b\bar{b}$ production cross-section is several orders of magnitude larger compared to that at e^+e^- colliders. However, the integrated luminosity is generally lower. In this case, the interaction occurs between the fundamental constituents of the protons, making it more difficult to accurately determine both the energy and the location of the interaction. A summary of the beauty production cross sections at LHCb over the years is provided in **Tab. 1.2**.

Experiment	Year	Beam	$\sqrt{s}$	$\sigma_{acc}(b\bar{b})$	$\int \mathcal{L}$	$\sim n(B^0)$
LHCb Run1	2011-2012	pp	7-8 TeV	$74 \mu\text{b}$ [15]	$1+2 \text{ fb}^{-1}$	2.2×10^{11}
LHCb Run2	2015-2018	pp	13 TeV	$144 \mu\text{b}$ [15]	5.9 fb^{-1}	8.5×10^{11}
LHCb Run3	2024	pp	13.6 TeV	$\sim 150 \mu\text{b}$	9.56 fb^{-1} [16]	1.4×10^{12}

Table 1.2: Summary of $b\bar{b}$ production cross-sections for different experiments inside the respective detector acceptances. The LHCb Run 3 production cross-section at $\sqrt{s} = 13.6 \text{ TeV}$ has been extrapolated from that at $\sqrt{s} = 13 \text{ TeV}$ assuming linear scaling with $\sqrt{s}$.

1.4.2 State-of-the-Art measurements

The study of $b \rightarrow s\ell\ell$ decays has garnered increasing attention over the past decade, fueled by a series of intriguing experimental results. Deviations from the SM predictions were observed across several observables, including the differential branching fractions, angular distributions, and the ratio of branching fractions of muons relative to electrons, denoted as R_X (where X represents the common hadronic particle in the final states of the two decays). While none of these deviations were statistically significant enough to constitute a discovery, they were compelling enough to attract substantial attention. The consistency of the discrepancies across different measurements suggested the possibility of NP contributions, potentially involving lepton flavour violation. In 2022, however, with the release of simultaneous analyses of $R_{K^{*0}}$ and R_K , the excitement surrounding these deviations began to subside. These measurements aligned closely with SM predictions within their experimental uncertainties, resulting in a re-evaluation of the so-called “flavour anomalies”. Although these results altered the overall landscape of the field, the broader implications of these findings are still being explored, and several anomalies remain unresolved.

This section aims to provide an overview of the observables and their measurements that exhibited patterns of deviation from the SM, extending to the latest measurements available at the time of writing (spring 2025).

Differential branching fraction measurements Measurements of differential branching fractions of $b \rightarrow s\ell\ell$ decays in bins of q^2 , $d\mathcal{B}/dq^2$, have been performed within several experiments such as LHCb, BaBar, Belle, and CMS. The LHCb collaboration performed the analysis of the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ [17], $B_s^0 \rightarrow \phi\mu^+\mu^-$ [18], $B^+ \rightarrow K^+\mu^+\mu^-$, $B^0 \rightarrow K^0\mu^+\mu^-$, and $B^+ \rightarrow K^{*+}\mu^+\mu^-$ [19] decays². The results were found to be systematically lower than SM predictions at different degrees of precision, as shown in **Fig. 1.7**. The plot shows only the results

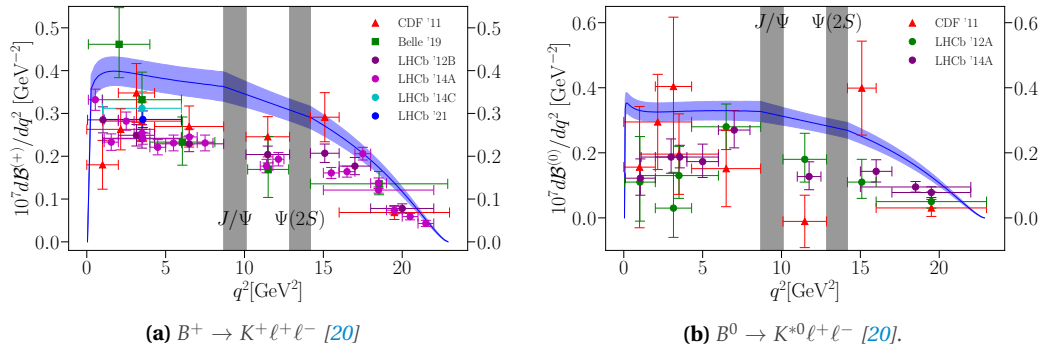


Figure 1.7: Figure (a) shows the differential branching fractions of the $B^+ \rightarrow K^+ \ell^+ \ell^-$ decays. For which Belle '19, and LHCb '14C and '21 have $\ell = e$, whilst otherwise $\ell = \mu$. Figure (b) shows the differential branching fraction of $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decay, where all experimental results take $\ell = \mu$. Both plots report measurements carried out in bins of q^2 from several experiments. They are compared with the theoretical prediction [20] - blue band. The grey regions correspond to regions dominated by the charmonium resonances (*i.e.* J/ψ and $\Psi(2S)$), the experimental measurement and the prediction avoid those regions.

of $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decays from various experiments and compares them to their latest theoretical prediction [20]. For completeness, LHCb results have been compared with

²The analyses presented in this chapter are CP -averaged unless stated otherwise. Therefore, the charge conjugate decays are implicitly included.

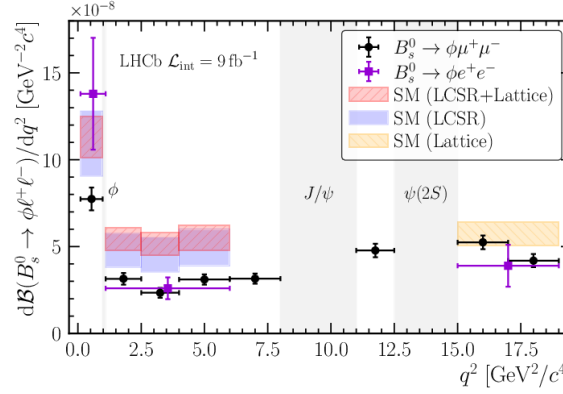


Figure 1.8: Differential branching fraction of the $B_s^0 \rightarrow \phi e^+ e^-$ decay in bins of q^2 , compared with measurements from $B_s^0 \rightarrow \phi \mu^+ \mu^-$ decays in Ref. [18] and SM predictions from FLAVIO [27] using form factors calculated with light-cone sum-rule [22] and lattice QCD techniques [28, 29]. Figure taken from appendix B of Ref. [26].

other theoretical predictions over time, notably in 2011 [21] and in 2014 [19]. These comparisons show that theoretical predictions (with larger uncertainties) and the experimental measurements tend to agree more closely in the q^2 region above 15 GeV^2/c^4 . A similar trend is also visible in the measurement of $B_s^0 \rightarrow \phi \mu^+ \mu^-$ of Ref. [18], comparing the results with SM predictions using Light Cone Sum Rules [22]. Other results from BaBar [23], Belle [24] and CMS [25] collaborations are available and could hint at the same trend, albeit with larger uncertainties. A new analysis was recently published (April 2025) measuring, for the first time, the differential branching fraction of the $B_s^0 \rightarrow \phi e^+ e^-$ decay channel at LHCb in three q^2 regions [26]. The results are shown in **Fig. 1.8**, in comparison with those from the same decay with muons in the final state [18], as well as with their SM predictions. The measurements are consistent with each other, except in the low- q^2 region ($[0.1-1.1], \text{GeV}^2/c^4$). They also agree with the SM predictions within uncertainties (although these uncertainties are generally larger than those in the muon channel), except in the central q^2 region where the results tend to lie below the SM expectations, similarly to the other analysis.

Angular analysis The LHCb collaboration performed several angular analyses in the muon sector ($b \rightarrow s\ell\ell$ decays where $\ell = \mu$) such as $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ [30], $B_s^0 \rightarrow \phi \mu^+ \mu^-$ [31] and $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ [32] modes. The ATLAS and CMS collaborations also analysed the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay [33, 34], and the Belle collaboration made a combined measurement of the $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ and $B^+ \rightarrow K^{*+} \ell^+ \ell^-$ modes [35], where $\ell = e, \mu$. The results of these analyses were generally in good agreement with the SM predictions, except for the observable P'_5 (Eq. 1.40), which exhibited a noticeable deviation in a few central q^2 bins. This variable is an angular observable, in the theoretically optimised base, as explained in **Sec. 1.3.3**. The plot in **Fig. 1.9a** shows the results of the angular observable P'_5 in different q^2 bins from several experiments which performed the angular analysis on the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay channel, compared to its SM prediction. This tension was consistently observed across experiments, except for the CMS measurement, using data from pp collisions at $\sqrt{s} = 8$ TeV [34]. During 2024, CMS updated these results using data from pp collisions at $\sqrt{s} = 13$ TeV. They are shown in **Fig. 1.9b** (black data points) in comparison to their old measurement [34] (purple) and to the LHCb measurement [30] (yellow). One can see that the updated and more precise measurement agrees in all the q^2 bins with the LHCb

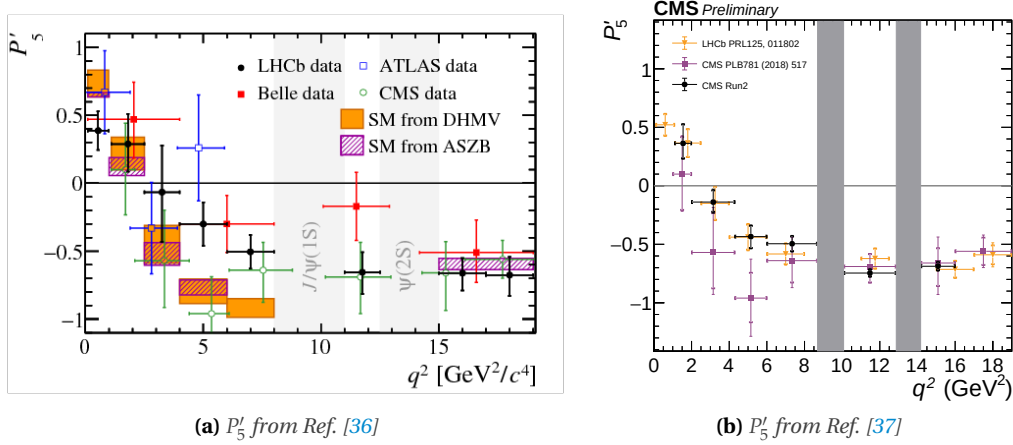


Figure 1.9: Figure (a) show the measurements of the angular observable P'_5 in bins of q^2 from the LHCb [30] (black), ATLAS [33] (blue), CMS [34] (green), and Belle [35] (red) collaborations. The values reported for the measurements performed by the Belle collaboration are the average between the electrons and muons modes. The SM predictions are from Refs. [11, 38]. The plot is taken from Ref. [36]. Figure (b) show the new CMS measurements [37] (black) of the angular observable P'_5 versus q^2 , in comparison to results from LHCb [30] and previous CMS publications [34]. The inner vertical bars represent the statistical uncertainties, while the outer vertical bars give the total uncertainties. The horizontal bars show the bin widths. The vertical shaded regions correspond to the J/ψ and $\Psi(2S)$ resonances. The definition of the q^2 bins may differ between the various measurements.

measurement. This represents an important cross-check between experiments performed on independent datasets.

To better understand the implications of such results, it is insightful to examine the likelihood profile of the Wilson coefficients fit, particularly that of the Wilson coefficient C_9 . The plots in **Fig. 1.10** show the likelihood profiles as a function of the shift in the real part of C_9 from the SM prediction [32]. The minimum for this likelihood is consistently shifted to the left, with a

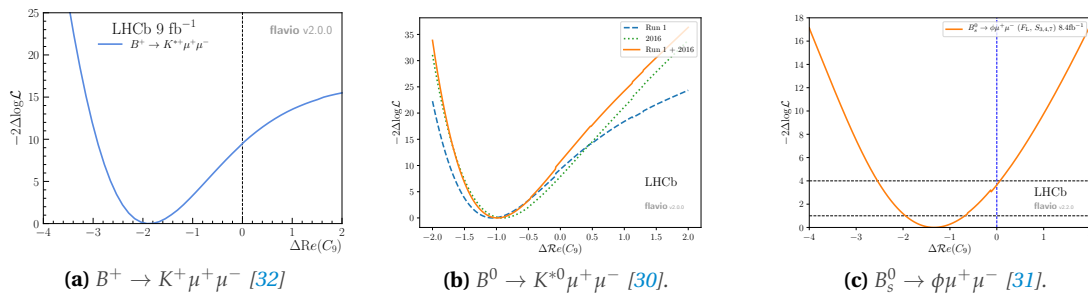


Figure 1.10: The FLAVIO [27] software package is used to fit the results obtained for the S_i CP -averaged angular observables by varying the real value of the vector coupling $\text{Re}(C_9)$. The plots show the likelihood scans of the shift of C_9 from its SM value. The focus is on the real part because the angular observables used in this fit are CP -averaged, making it impossible to access the imaginary part.

significance of $2\text{--}3\sigma$.

During early 2025, the LHCb collaboration published the first angular analysis in the electron sector, using the $B^0 \rightarrow K^{*0} e^+ e^-$ mode and the luminosity collected during LHC Run 1 and Run 2, 9fb^{-1} [39]. The analysis is performed in the central q^2 bin that spans from 1.1 to 6.0 GeV^2/c^4 . The full set of angular observables exhibits good agreement with SM predictions, as

shown in **Fig. 1.11a**. The plot in **Fig. 1.11b** shows the difference between the P_i (defined in **Sec. 1.3.3**) observables measured for the muon and the electron modes. This quantity is directly

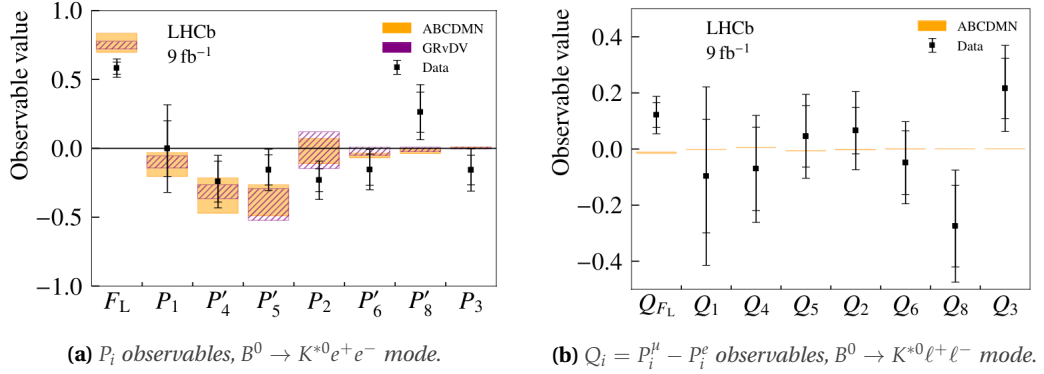


Figure 1.11: Figure (a) compares the CP -averaged angular observables resulting from the analysis of the decay $B^0 \rightarrow K^{*0} e^+ e^-$ with the SM predictions from Ref. [40] (yellow boxes) and Refs. [41, 42] (hatched boxes). Figure (b) shows the lepton flavour universality observables Q_i calculated using the P -basis angular observables of the muon and electron modes. The SM predictions (orange boxes) are based on Ref. [40]

related to the concept of “Lepton Flavour Universality” (LFU). The SM states that the coupling of gauge bosons to leptons is universal. The only difference arises from their Yukawa couplings, so from their mass differences. If LFU is conserved, then the angular observables describing $B^0 \rightarrow K^{*0} e^+ e^-$ and $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ are the same, leading to their difference being zero (when the mass difference is negligible with respect to the energy of the decay, which is the case for these analyses). This is the reason why variables such as Q_i are probes of LFU. In this case, no evidence of lepton flavour violation was observed. To enable a more precise comparison between the electron and muon sectors, it would be ideal to conduct the analysis using smaller bins, at least matching the bin size used in the corresponding muon analysis (Ref. [30]). However, the lower energy resolution and reduced reconstruction efficiency of electrons, compared to muons in LHCb, present significant challenges for such an analysis. In March 2025, constraints on the photon polarisation were published by performing an angular analysis of the $B_s^0 \rightarrow \phi e^+ e^-$ decay in the very low q^2 region between 0.0009 and 0.2615 GeV²/c⁴, using Run 1 and Run 2 datasets [43]. In this q^2 region, the measurement is mainly sensitive to the C_7 Wilson coefficient. The results are consistent with the SM predictions. Finally, the angular analysis of the same decay channel in q^2 regions [0.1, 1.1], [1.1, 6.0], and [15.0, 19.0] GeV²/c⁴ has been performed and presented at the “Moriond EW ‘25” conference [44], while the paper is still in preparation [45]. Also in this case, the results are found to be consistent with both the SM predictions and the more precise measurements on the same decay channel with muons in the final state [31], following the LFU hypothesis.

Ratio of branching fractions R_X These observables, as Q_i , are probes of LFU and they are given by the ratio between branching ratios of semileptonic decays that differ only in the lepton content, while the hadron (X) in the final state is the same:

$$R_X = \frac{\int_{q_{min}^2}^{q_{max}^2} \frac{d\Gamma(B \rightarrow X \mu^+ \mu^-)}{dq^2} dq^2}{\int_{q_{min}^2}^{q_{max}^2} \frac{d\Gamma(B \rightarrow X e^+ e^-)}{dq^2} dq^2}. \quad (1.41)$$

The fact that the hadron X is in common between the two decays allows for the cancellation of both form factor and non-local hadronic effects. Consequently the theoretical prediction of the value of R_X is exact at the level of $\mathcal{O}(1\%)$ in the q^2 region of $[1.0, 6.0] \text{ GeV}^2/c^4$ [46]. The variable R_X is then one of the most precise and powerful tests of LFU. LHCb has performed several measurements of this quantity over the years. Among those, the most up-to-date measurements at the time of writing are summarised in **Fig. 1.12**. The measurements show an overall agreement

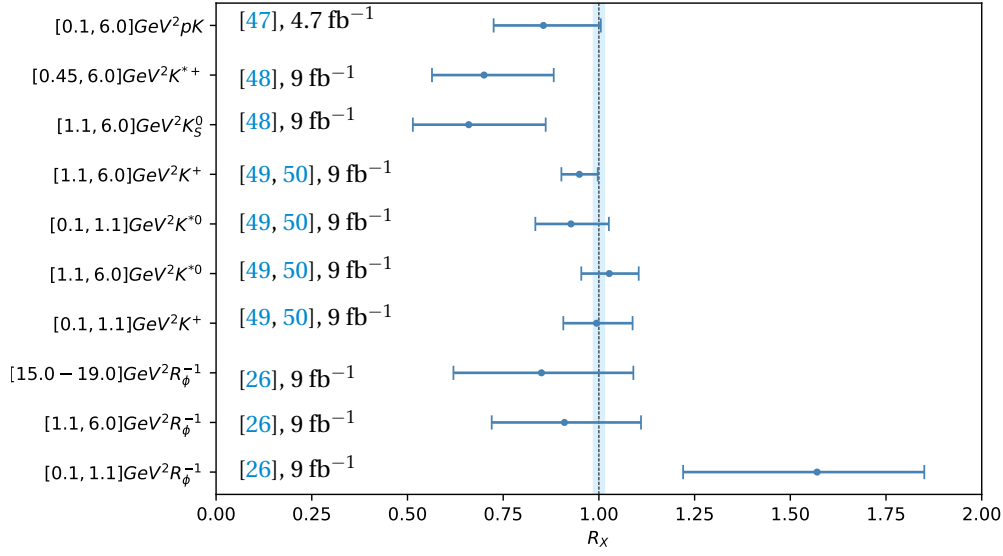


Figure 1.12: State-of-the-art R_X measurements performed by the LHCb collaboration. On the x -axis, the measured value of the R_X observable, while on the y -axis are listed several measurements specifying the q^2 bin where they are performed and the integrated luminosity utilised. The hadrons specified on the y -axis correspond to the X in R_X .

with the SM prediction, which is 1 as a consequence of the assumption of LFU³. Measurements of R_X type have also been performed by BaBar [23] and Belle [51, 52], and the results are in agreement with the SM, albeit with large uncertainties.

Wilson Coefficients fits The full range of $b \rightarrow s\ell\ell$ observables can be analysed simultaneously via “global fits” using the EFT framework introduced in **Sec. 1.2.1** to determine a possible common underlying feature. This can be done by parametrising the observables in terms of WCs, and performing maximum likelihood fits to determine potential NP contributions ($\mathcal{C}_i^{\text{NP}}$). Several research groups have conducted global fits involving various combinations of WC [40, 53, 54]. The plot in **Fig. 1.13a**, shows the comparison of the results of the global fits performed by different collaborations on the $\mathcal{C}_{9\mu}^{\text{NP}} - \mathcal{C}_{10\mu}^{\text{NP}}$ plane. The WCs in the plot are defined as

$$\mathcal{C}_i^{\text{NP}} = \mathcal{C}_i - \mathcal{C}_i^{\text{SM}}.$$

Despite the diverse methodologies pursued, a substantial level of agreement is evident when comparing the results within the $\mathcal{C}_{9\mu}^{\text{NP}} - \mathcal{C}_{10\mu}^{\text{NP}}$ plane. For completeness, although it is not shown

³The SM predictions for R_X measurements in low q^2 regions are slightly different from 1 due to the difference in mass between muons and electrons. However, these deviations have been ignored in the plot.

in the plot, one approach stands out for its significant disagreement with the others: the Phenomenological Data Driven (PDD) approach from the CFFPSV group. The PDD approach allows all parameters of their polynomial parametrisation to vary freely without constraint and fits them to the available data. Therefore, this discrepancy arises from the vast number of free parameters which allow it to absorb a large part of any potential NP effects [55]. This outlier underscores the complexities introduced by the presence of numerous parameters in the analysis and emphasises the importance of exercising caution when interpreting potential new physics (NP) effects. The reliability of these results is likely contingent on the fixed values of these parameters and the uncertainties associated with them, highlighting the need for a thorough evaluation of how parameter choices influence the overall conclusions.

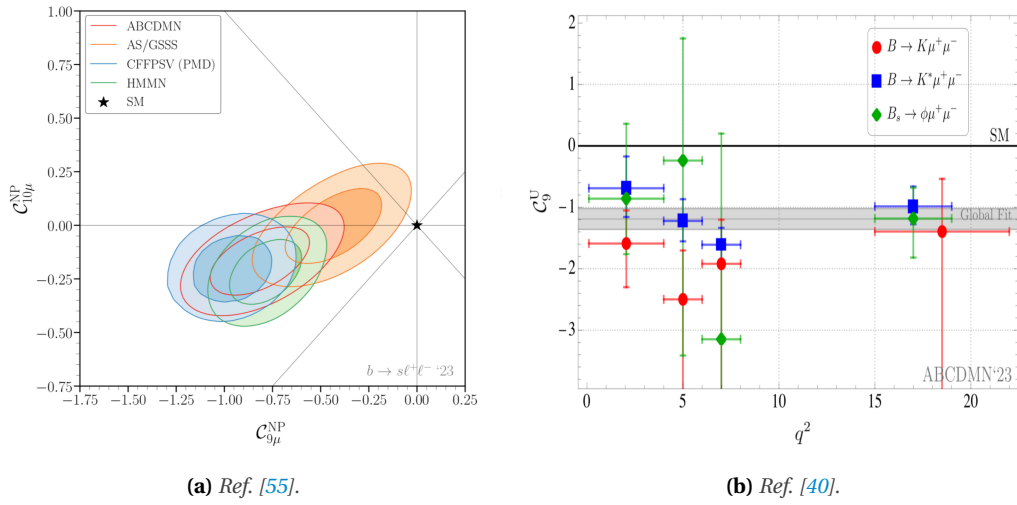


Figure 1.13: Figure (a) shows the comparison of global fits results of the global fits by different collaborations in the $C_{9\mu}^{NP} - C_{10\mu}^{NP}$ plane: ABCDMC [56], AS/GSSS [53], CFFPSV [54] (Phenomenological Model Driven approach), HMMN [57, 58]. Figure (b) shows the preferred value of C_9^U for different bins of q^2 . The data included in these fits correspond to the branching fractions for both the charged and neutral $B \rightarrow K^{(*)}\mu^+\mu^-$, the branching fractions for the $B_s^0 \rightarrow \phi\mu^+\mu^-$ mode, and the angular observables for both the $B \rightarrow K^*\mu^+\mu^-$ and $B_s^0 \rightarrow \phi\mu^+\mu^-$ modes. The grey band corresponds to the 1σ confidence interval for the global fit to C_9^U .

Among the various the NP scenarios, the most favoured is C_9^U scenario, representing a LFU-NP contribution [56, 40] ($C_{9\mu}^{NP} = C_{9e}^{NP} = C_9^U$). It has proven particularly effective in explaining the remaining tensions. The plot in **Fig. 1.13b** shows the best-fit points for C_9^U (assuming the SM value for the other Wilson coefficients) when performing fits on distinct data sets for the $B^\pm \rightarrow K^\pm\mu^+\mu^-$, the $B \rightarrow K^*\mu^+\mu^-$ and the $B_s^0 \rightarrow \phi\mu^+\mu^-$ modes in different q^2 bins [40].

In 2024, the LHCb collaboration published the first amplitude analysis of the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decay in the q^2 range of $[0.1, 18.0] \text{ GeV}^2/c^4$ [59]. The analysis uses a model of one- and two-particle non-local amplitudes to explicitly isolate the local and non-local contributions to the decay and capture their interference. This procedure allows for direct measurement of the $b \rightarrow s\mu\mu$ WCs. The result shows a 2.1σ deviation in the C_9 parameter. This set of results is consistent with those reported in recent global analyses of $b \rightarrow s\ell\ell$ decays [40], which favour lepton flavour universal NP contributions to the WC C_9 .

The global fit presented in this section does not include the latest results from analyses of type $b \rightarrow see$, as it was performed before the publication of these analyses and an updated plot has

not yet been published. However, the angular analysis of $B^0 \rightarrow K^{*0} e^+ e^-$ decays [39] published the C_9 WC coefficient fit using their results, and compared to the corresponding fit obtained using the angular parameters from the same decay channel with muons ($B^0 \rightarrow K^{*0} \mu^+ \mu^-$). The two likelihood profiles as a function of C_9 are shown in **Fig. 1.14a**. The likelihood profile can be expressed directly as a function of the $\Delta C_9 = C_{9\mu} - C_{9e}$ WC. **Fig. 1.14b** shows ΔC_9 compatible with zero within 1σ , confirming LFU and the shift in $C_{9\mu}$ with a significance of approximately 2σ . This is particularly remarkable since the systematic uncertainties of electrons and muons are quite different and uncorrelated. These results lend further support to the LFU-NP scenario.

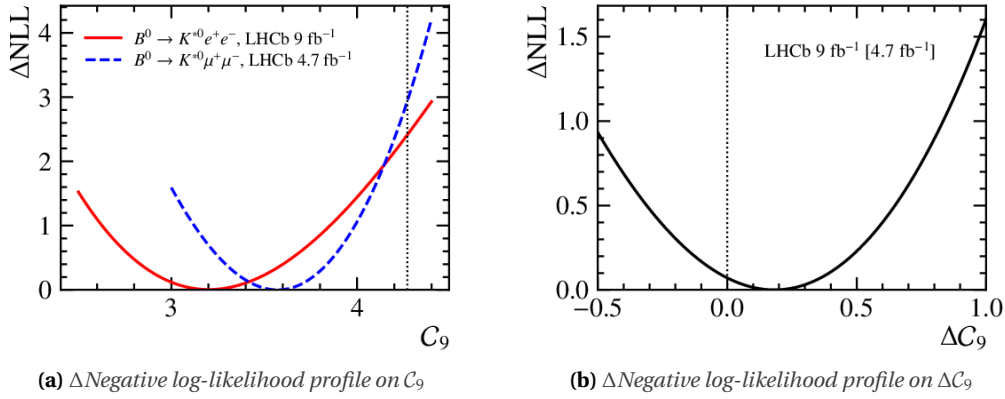


Figure 1.14: Figure (a) shows the negative log-likelihood scan of C_{9e} (red) and $C_{9\mu}$ (blue). Figure (b) shows the negative log-likelihood scan of ΔC_9 . The dotted vertical line corresponds to the SM prediction [60, 61]. Plots from Ref. [39]

1.5 Context and motivation of the current work

The analyses presented in **Sec. 1.4.2** highlight an intriguing puzzle, as several measurements show tensions with SM predictions, particularly in the muon sector. Among the existing NP scenarios that could explain such tensions, the most favoured is one which preserves LFU. The results from the latest angular analysis of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay in the central q^2 region [39] and the unbinned angular analysis of the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay [59] overall align with the SM prediction, although a few observables display small tensions leading to a WCs fit which seem to support the LFU-NP hypothesis. However, the theoretical predictions of these observables are very challenging, and unaccounted SM hadronic contributions could also generate an effect on C_9^U [62], mimicking the presence of a universal NP. Nonetheless, these tensions demand further interpretations, whether in terms of experimental effects, theoretical systematics or NP explanations. Therefore, additional analyses are necessary to have a clearer picture. In particular, $Q_5 = P_5^\mu - P_5^e$ is extremely important to break the degeneracy among the scenarios [62], as it would provide information on C_9^U . This quantity requires angular analyses to be performed also in the electron sector. This is the main motivation leading the analysis presented in this thesis, which will be performed using data collected by the LHCb experiment during LHC Run 1 and Run 2 data-taking periods.

While an angular analysis of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay channel has already been presented

using the same dataset, this new study represents a significant step forward. It builds upon the foundations of the previous measurement by refining analysis techniques and expanding the scope of the investigation. This thesis focuses on the central q^2 region and its two sub-bins. As discussed in **Sec. 1.2.2**, analysing the decay as a function of q^2 is essential to access information about different WCs. Furthermore, this approach enables a more precise comparison with the muon decay channel, allowing for fine-grained tests of LFU. Experimentally, this is achieved by performing the analysis in bins of q^2 . While this thesis focuses on the central q^2 region, the full analysis will also provide new results in the previously unexplored low q^2 region, $[0.1, 1.1] \text{ GeV}^2/c^4$.

2 | The LHCb experiment at the LHC

This chapter provides a brief overview of the LHCb detector, the experimental apparatus used to collect the data analysed in this dissertation. Finally, a section dedicated to electron reconstruction is also included, as it presents significant challenges for this analysis.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a superconducting circular hadron accelerator and, more specifically, a proton-proton (pp) collider operating at CERN (*Conseil Européen pour la Recherche Nucléaire*). It is housed in a 27 km tunnel that previously accommodated the Large Electron-Positron Collider (LEP), located approximately 100 m underground along the French-Swiss border near Geneva [63].

Protons are extracted from hydrogen gas, then pass through several sequential acceleration stages to increase their energy before being injected into the LHC. A schematic representation of the CERN accelerator complex is shown in **Fig. 2.1**. Hydrogen atoms are first extracted from a

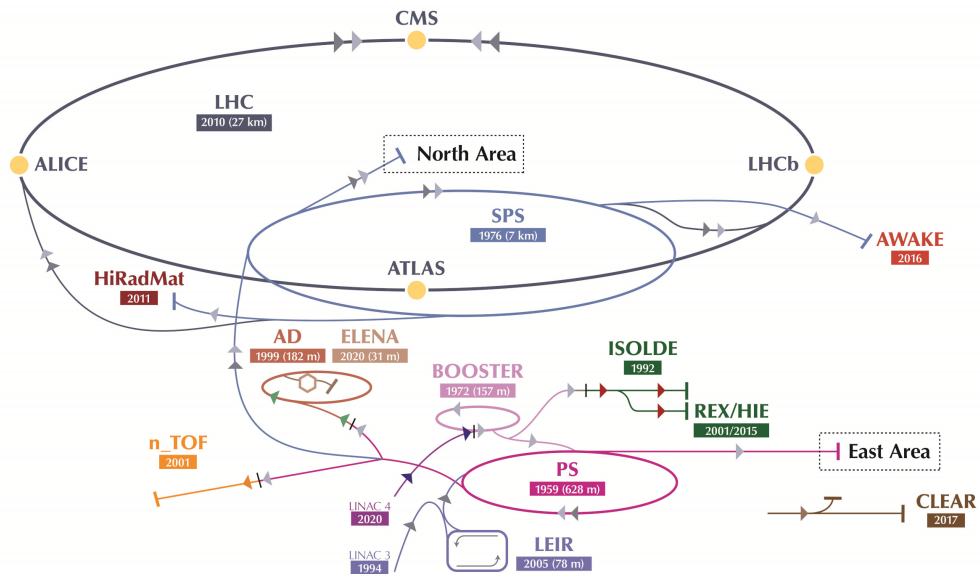


Figure 2.1: CERN accelerator complex [64].

hydrogen source and stripped of their electron, leaving protons. The protons are then accelerated in the LINAC 2. Afterwards, they are injected into the BOOSTER, a system of four synchrotron rings. At this stage, the protons reach an energy of 1.5 GeV. Next, they are transferred to the Proton Synchrotron (PS) and then to the Super Proton Synchrotron (SPS), where their energy is

increased to 450 GeV. Finally, the proton beams are injected into the LHC, where they circulate in opposite directions and are further accelerated by radio-frequency (RF) cavities. The beams are steered around the LHC's circumference using superconducting dipole magnets, which operate at an ultra-low temperature of 1.9 K, maintained by a liquid helium cooling system. These dipole magnets generate a magnetic field of 8.3 T. The beam pipes require an ultra-high vacuum ($1.013 \cdot 10^{-10}$ mbar, or 10^{-13} atmospheres) to prevent collisions between the particle beams and residual gas molecules inside the accelerator. Collisions occur at four interaction points along the LHC ring, where the detectors of the four main LHC experiments are located: ALICE, ATLAS, CMS, and LHCb.

During 2011, the centre-of-mass energy of the collisions was set to $\sqrt{s} = 7$ TeV, which was increased to 8 TeV in 2012. The data-taking period during 2011 and 2012 is referred to as Run 1. Following a two-year shutdown for upgrades and maintenance, particularly on the magnet systems, the LHC restarted operations in 2015 at a collision energy of 13 TeV, which was maintained throughout Run 2 (2015–2018).

Proton beams are organised into bunches, each containing approximately 10^{11} protons, with a time spacing of 25 ns between bunches (50 ns during Run 1). This spacing corresponds to a nominal bunch crossing rate of 40 MHz (20 MHz during Run 1). However, larger gaps are introduced between certain bunches to facilitate beam injection and extraction, resulting in an effective average crossing rate of approximately 30 MHz (15 MHz during Run 1).

Two important quantities related to the LHC machine can be defined:

Instantaneous luminosity Assuming the distribution of the protons in the bunches to be Gaussian, $\mathcal{L}$ can be defined according to the relation [65]:

$$\mathcal{L} = \frac{kfn_1n_2}{4\pi\sigma_x\sigma_y} \cdot \frac{1}{\sqrt{1 + \left(\frac{\sigma_z}{\sigma_y} \tan \frac{\phi}{2}\right)^2}} \quad (2.1)$$

where

- k is the number of bunches per beam (~ 2500 in Run 2).
- f is the revolution frequency of the bunches ($f \sim 11\text{kHz}$).
- n_1 and n_2 are the number of protons in the two colliding bunches: $n_1 = n_2 = 10^{11}$.
- $\sigma_x, \sigma_y, \sigma_z$ are the standard deviations in the three spatial directions of the Gaussian distributed bunch $\sigma_x = \sigma_y \sim 16 \mu\text{m}$ and $\sigma_z \gg \sigma_x, \sigma_y$, as z is the direction along which the beams are accelerated.
- ϕ is the crossing-angle between the two beams, which cross in the zy plane at LHCb. The LHCb coordinate system is described in **Sec. 2.2**.

Integrated luminosity The integrated luminosity, L , is defined as the time integral of the instantaneous luminosity:

$$L = \int \mathcal{L} dt.$$

This quantity is particularly useful because, when multiplied by the cross-section of a specific process, it directly yields the number of events produced for that process. The event rate for a

given process is expressed as:

$$\frac{dN}{dt} = \mathcal{L} \cdot \sigma(\sqrt{s}) \cdot \epsilon, \quad (2.2)$$

where $\sigma(\sqrt{s})$ is the cross-section at the centre-of-mass energy $\sqrt{s}$, and ϵ represents the efficiency for detecting the process.

While CMS and ATLAS aim to achieve a peak instantaneous luminosity of approximately $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$, LHCb was designed to operate at a much lower instantaneous luminosity. To maintain this, the luminosity was intentionally limited to $2 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ by introducing a transverse offset between the two colliding beams [66]. In 2012, this value increased to $4 \cdot 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$, a level that was maintained until the end of 2018.

2.2 LHCb detector

The Large Hadron Collider beauty (LHCb) experiment is the only LHC experiment primarily dedicated to heavy flavour physics. It is a single-arm spectrometer which covers the forward pseudorapidity range of $1.8 < \eta < 4.9^1$. The reason for this design choice must be sought in the main research aim of the experiment: in a high-energy pp collider, b and c quarks are mainly produced as $b\bar{b}$ and $c\bar{c}$ pairs by the strong interaction. As the masses of these pairs are small compared to the centre-of-mass energy, they are likely to be produced in forward and backward cones along the beam axis.

The LHCb experiment coordinate system is defined as follows:

- The *origin* of the coordinate system O is the nominal interaction vertex position of the pp collision.
- The z -axis is aligned to the nominal beam direction.
- the x -axis and y -axis are respectively the horizontal and vertical axis, perpendicular to each other and to the beam axis.

The data analysed in this dissertation were collected during LHC Run 1 and Run 2. The description of the detector corresponds to its apparatus during those years. In the following lines a schematic description of the LHCb sub-detectors and magnet is given, and a sketch of the detector is reported in **Fig. 2.2**:

- **VELO**: the VERTex LOcator is a silicon strip detector placed around the pp interaction vertex, which aims to reconstruct the position of decay vertices with high precision [67].
- **RICH1**: the first Ring Imaging CHERenkov detector is placed directly after the VELO. Its main function is to help in particle identification, covering a momentum range from 1 to 60 GeV/c [68].
- **TT**: the Tracker Turicensis is a silicon strip detector placed just before the magnet [69, 70].
- **Magnet**: a warm dipole magnet provides a bending power of 4 Tm for the measurement of charged particles momenta [71].

¹ $\eta = -\log(\tan(\frac{\theta}{2}))$ and θ is the polar angle to the beam direction

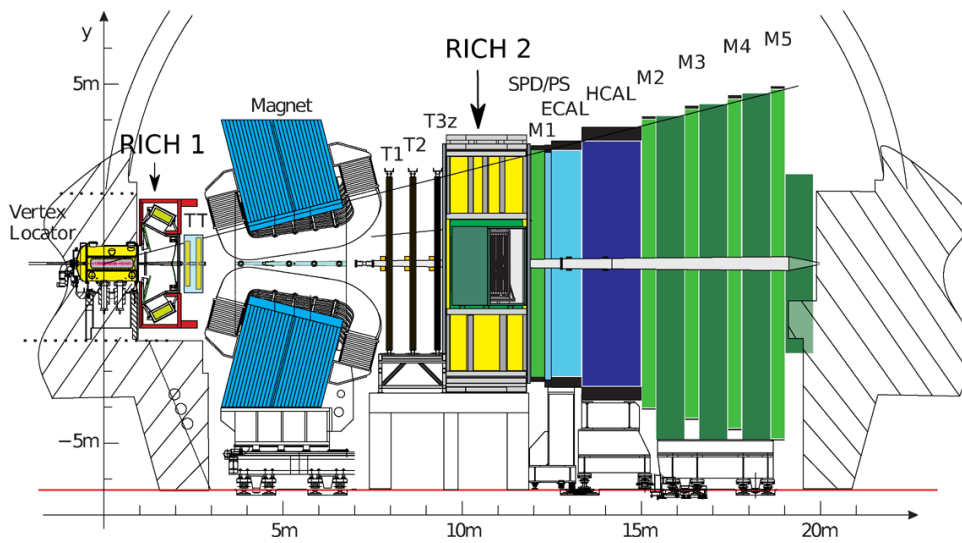


Figure 2.2: Side view of LHCb detector.

- **T1, T2, T3:** three stations, each consisting of two sub-detectors, the Inner Tracker (silicon strips) and Outer Tracker (straw tubes) [70, 72].
- **RICH2:** the second Ring Imaging Cherenkov covers a different momentum range with respect to the RICH1 (15-100 GeV/c) [68].
- **ECAL:** the Electromagnetic CALorimeter identifies electrons and photons and gives important information for the trigger system [73].
- **SPD/PS:** the Scintillating Pad and the Preshower Detectors improve the identification of electrons [73].
- **HCAL:** the hadronic calorimeter is used to identify and trigger hadrons [73].
- **M1-5:** finally five muon stations (multi-wire proportional chambers interspersed with iron layers) identify and trigger muons, the only particles capable to reach this region of the detector [74, 75, 76].

Utilising all these subsystems, LHCb measures all main observables of the particles produced in the collision: energy, momentum, velocity and position. These measurements allow the determination of their trajectories, charge, momenta, and, ultimately, enable particle identification.

2.3 Tracking system

The tracking system provides precise spatial measurements of charged particle tracks, enabling the determination of key quantities such as charge, momentum, and vertex positions. By combining information from different tracking sub-detectors, it is possible to reconstruct tracks, *i.e.*, the trajectories of charged particles. The first step in track reconstruction is pattern recognition, which involves identifying and grouping the detector hits produced by a single charged particle. The tracking system consists of the VELO, TT, and T-stations. The magnet introduces a deflection in the particles' trajectories, causing them to follow curved paths. This curvature allows for the measurement of their momenta.

VELO

The VELO is a silicon strip vertex detector that surrounds the interaction region at LHCb. The principal task of the VELO is to enable LHCb to trigger on and reconstruct displaced vertices, as well as to play a significant role in the tracking. It provides precise measurement of charged particles' trajectories in this region, so that it is possible to precisely:

- measure particles' momenta, by comparing the tracks' directions reconstructed in the VELO (before the magnet) with those reconstructed in the T stations (after the magnet);
- reconstruct pp interaction vertices (primary vertices) and displaced secondary vertices.

The presence of a displaced secondary vertex is a signature of b - or c -hadron decay, whose mean flight distance from their creation in the primary vertex to their decay is $\mathcal{O}(\text{cm})$.

TT

The TT is a silicon micro-strip detector placed just in front of the magnet, and has two purposes: first, to detect low-momentum particles that are bent out of the acceptance of the experiment by the magnet and thus do not reach stations T1-T3. Second, to reconstruct long-lived particles, such as K_S^0 and Λ , as their decay products may be produced after the VELO sub-detector.

The dipole magnet

The LHCb magnet is placed between the TT and the T-stations and bends the trajectories of charged particles, thereby enabling the measurement of their momentum. The magnet is formed by two saddle-shaped coils placed with a small angle with respect to the beam axis, to increase the opening window with z , in order to follow the acceptance of the LHCb detector. The field mainly develops in the y direction, hence particles are bent in the xz plane. The maximum intensity of the field is about 1.1 T, while its integral $\int Bdl = 4 \text{ Tm}$. To help determine and reduce any asymmetry in the detection of charged particles, an important feature is incorporated: the possibility of reversing the polarity of the magnet (up or down). The magnet polarity is inverted about every two weeks.

T stations

The T-stations (T1, T2, T3) are three tracking stations placed just downstream of the magnet to measure charged particles' momentum. They consist of two sub-detectors:

- a small **Inner Tracker** (IT), made of silicon micro-strips. It covers the high-occupancy region around the beam pipe.
- a much bigger **Outer Tracker** (OT), made of straw tubes. It is a gaseous ionisation detector that covers an area of about $6 \text{ m} \times 5 \text{ m}$, the whole LHCb geometrical acceptance, excluding the high-occupancy region covered by the IT.

LHCb type of tracks

The different types of tracks reconstructed at LHCb are shown in **Fig. 2.3**. Their classification depends on the sub-detector(s) traversed.

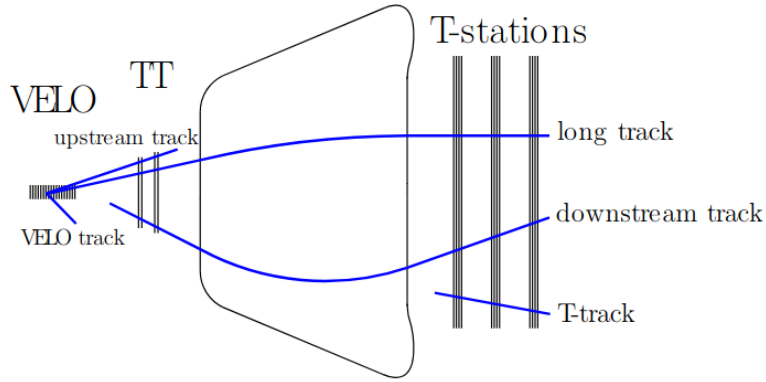


Figure 2.3: Tracks classification at LHCb depending on the crossed sub-detectors.

2.4 Particle identification

An excellent particle identification (PID) performance is vital for LHCb physics to reject backgrounds and for the classification of the final states. Cherenkov detectors can distinguish between charged kaons and pions, while calorimeter detectors allow identification of electrons, photons and hadrons. Muons are identified by their passage through the muon chambers.

2.4.1 Particle identification system

The Ring Cherenkov detectors The primary role of the RICH system is the identification of charged hadrons (π , K , p). One of the major requirements in a flavour-physics experiment is the reduction of combinatorial background [77]. In fact, many of the interesting decay modes of b - and c -flavoured hadrons involve hadronic multibody final states. The two Cherenkov detectors (RICH1 and RICH2) allow for the distinction of the particle type of charged hadrons in a range from 1 GeV/ c to 100 GeV/ c , exploiting the Cherenkov effect. Consider a particle that crosses a medium (refraction index n) with velocity v . If its velocity is higher than the speed of light in that medium ($v > c/n$), photons are emitted in a cone with angle θ_C , defined as $\cos \theta_C = 1/\beta n$, with $\beta = v/c$. It follows that Cherenkov light is emitted only if $\beta > 1/n$. In both RICH detectors, a complex system of spherical and planar mirrors reflects the emitted photons outside the LHCb acceptance, where they are collected by a lattice of hybrid photon detectors (HPDs). In this way, HPDs can be shielded from the magnetic field. Since the momentum p of the particle is known from the tracking system, it is possible to find its mass from the following relation:

$$n \cos(\theta_C) = \sqrt{1 + \left(\frac{mc}{p}\right)^2}.$$

In order to achieve the best separation power between pion and kaon mass hypotheses, the difference between the Cherenkov angles for these hypotheses must be maximized, which implies a small refractive index (approaching 1). However, the smaller the refractive index, the larger the threshold momenta below which Cherenkov light is emitted. Therefore, a compromise must be reached. To cover a large range of momenta, two RICH detectors are used at LHCb, based on two different refractive indices. The RICH1 is placed before the magnet (between VELO and TT) to identify the particles that are bent out from the LHCb acceptance by the magnetic field, and it

covers the low momenta range 1-60 GeV/c. The RICH2 is placed after T3 station and covers the high-momentum range 15-100 GeV/c.

Calorimeter detectors Calorimeter detectors provide fast information for the low-level trigger and offer identification of electrons, photons, and hadrons, together with a raw measurement of their energies and positions. The following sub-detectors form the calorimetric system:

- the **Scintillator Pad** (SPD) and **PreShower** (PS). The SPD is placed in front of the ECAL and the PS is placed after the SPD. In the middle, there is a 15 mm lead absorber, corresponding to ~ 2.5 radiation lengths for electrons. They consist of two planes of scintillating pads, used at the low-level electron trigger to reject background from charged and neutral pions and to improve electron identification. The SPD, just like a tracking detector, reveals only charged particles and allows electrons to be separated from photons [78]. The PS detector identifies electromagnetic particles.
- the **Electromagnetic Calorimeter** (ECAL). The ECAL consists of alternating 4 mm thick scintillator tiles and 2 mm thick lead plates. This thickness corresponds to about 25 radiation lengths, guaranteeing an almost complete electromagnetic shower containment and providing a good energy resolution of approximately $0.9\% \oplus (8\% \div 10\%) / \sqrt{E/\text{GeV}}$ [79].
- the **Hadron Calorimeter** (HCAL). The HCAL consists of alternating 4 mm thick scintillator tiles sandwiched between 16 mm iron sheets, corresponding to about 5.6 interaction lengths with a resulting resolution of $9\% \oplus 69\% / \sqrt{E/\text{GeV}}$ [79]. This value is due to a limited thickness, which does not allow the complete shower containment and hence energy measurement. For this reason, the HCAL is exploited mainly for trigger purposes.

A schematic representation of the energy deposit on the calorimeter system for different particles is shown in **Fig. 2.4**.

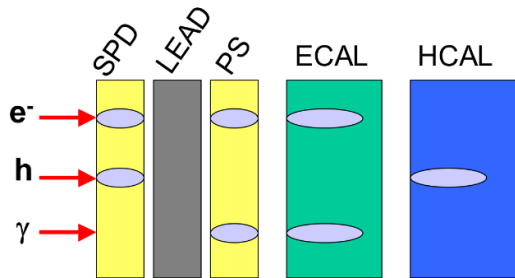


Figure 2.4: Signal deposited on the different parts of the calorimeter by an electron, a hadron, and a photon [78].

Muon detector The muon detector is composed of five stations (M1–M5) of rectangular shape, placed along the beam axis, designed to provide muon identification and fast measurement of their momenta for the low-level trigger [80]. The first station, M1, is placed before the calorimeter system, and it is used to improve the transverse momentum (p_T) measurement in the trigger (the calorimeter system, with its material budget, introduces additional uncertainties on position measurements downstream of it, due to multiple scattering). Stations M2 to M5 are placed downstream of the calorimeters and are interspersed with 80 cm thick iron absorbers to select

penetrating muons. In order to traverse the whole detector, the minimum momentum needed by a muon is about 6 GeV/ c , since the total absorber thickness (including calorimeters) corresponds to approximately 20 interaction lengths. In the trigger, muon reconstruction is performed by the stand-alone muon system, which achieves an average transverse momentum resolution of $\sim 20\%$.

2.4.2 Particle identification

Particle identification starts from the hits in the MUON stations that are searched in different regions as a function of the momentum of the muon candidate. Tracks with $p < 3$ GeV cannot be identified as muons: they would not be able to reach the MUON stations. Below a momentum of 6 GeV, the muon identification algorithm requires hits in the first two stations after the calorimeters, whereas, if $6 < p < 10$ GeV, an additional hit is required in one of the last two stations. Finally, above 10 GeV, hits are required in all four MUON stations. To identify photons, electrons and π^0 candidates, the calorimeter system is used. Neutral particles are distinguished from charged ones thanks to the presence or absence of tracks in front of the energy deposits. The discrimination between photons and π^0 mesons is performed by looking at the shape of the cluster. The PID information is obtained by combining all the information from the MUON stations, the RICH and the calorimeter system. Each sub-system gives a likelihood information for a certain particle hypothesis, $\mathcal{L}(X)$. Since pions are the most abundant species produced and detected at LHCb, the particle identification is usually calculated with respect to the pion hypothesis, $\mathcal{L}(\pi)$. The difference between the logarithms of two likelihoods (DLL) with the X and the π hypothesis is computed as $\text{PID}_X = \log \mathcal{L}(X) - \log \mathcal{L}(\pi) = \log \left[\frac{\mathcal{L}(X)}{\mathcal{L}(\pi)} \right]$. It immediately follows that the larger the PID_X value is, the more likely is that the particle belongs to the X species, rather than being a pion.

Another class of particle identification algorithms uses a Neural Network (NN) machine learning algorithm, which also combines information from the particle identification detectors and adds tracking system information [81]. The output of this algorithm for a particle y under the hypothesis that is of type x is called $y_{\text{ProbNN}x}$.

2.5 LHCb Trigger

The LHCb trigger is designed to efficiently select heavy flavour decays from the overwhelming QCD background produced in pp interactions [82]. At the nominal LHCb luminosity ($4 \cdot 10^{32} \text{cm}^{-2} \text{s}^{-1}$), over 200k $b\bar{b}$ pairs are produced every second. This means that, in order to sufficiently reduce the output rate of the trigger, not only must non- b backgrounds be efficiently rejected, but b -hadron candidates must also be classified, since the inclusive rate of b events alone already exceeds the available bandwidth [83]. The LHC bunch-crossing rate is 40 MHz, however, only a limited rate of events can be written on storage: 5 KHz during Run 1 and 12.5 KHz in Run 2. It is also true that only a small fraction of the total rate is made of interesting events. The reduction of the rate is performed in two trigger levels: Level-0 (L0), a custom hardware trigger intertwined with the read-out, and High Level Trigger (HLT), which is implemented in software and runs on a dedicated computer farm.

L0-Trigger: hardware trigger The read-out of some sub-detectors cannot transmit their complete data at a 40 MHz frequency, hence, the task of the L0 trigger is to reduce the events rate to 1 MHz. In order to do this, it exploits information from calorimeters and muon stations, as these are the only pieces of information available at such a high rate. This information is used to build a set of requirements which determines whether the detector information is passed on to the next stage. The Level-0 trigger consists of independent triggers that run in parallel, exploiting the information of these sub-detectors: L0Hadron, L0Electron and L0Photon use information from the calorimeter system to identify potential particles, and select events where the computed transverse energy of these candidates is above a set threshold. The L0Muon trigger line takes information from the muon stations to create candidate muons, and makes an estimate of their momenta. The trigger decision depends on the transverse momentum of one or two muon candidates.

High Level Trigger: software trigger The High Level Trigger (HLT) system [84] processes events selected by the L0 trigger in the Event Filter Farm (EFF) and consists of two stages: HLT1 and HLT2. HLT1 performs an inclusive selection by partially reconstructing charged particle tracks with $p_T > 500$ MeV and reconstructing primary vertices (PV) to reduce the input rate from 1 MHz to about 110 kHz. Events are then buffered for detector calibration and alignment before they are processed by HLT2. HLT2 conducts a full event reconstruction, including the reconstruction of charged and neutral particles, as well as PID. HLT2 achieves near-offline reconstruction quality by exploiting the full information from the tracking sub-detectors, performing additional steps of the pattern recognition which are not possible in HLT1 due to tight time constraints.

2.6 Electrons in LHCb

As electrons traverse the LHCb detector, due to their small mass, they lose energy through bremsstrahlung radiation, particularly in the high-material-budget regions before the calorimeter, complicating their reconstruction. The impact of bremsstrahlung depends on where it occurs in the detector. For instance, if an electron emits bremsstrahlung after passing through the dipole magnet, the photon and electron trajectories remain close and result in a single calorimeter cluster. However, if bremsstrahlung occurs upstream of the magnet, as represented in **Fig. 2.5a**, the photons' trajectory remains straight, while the electrons' trajectory bends due to the magnetic field. This results in separate calorimeter clusters and an underestimate of the electrons' (initial) momentum, leading to a degradation of the momentum resolution.

To mitigate this, LHCb employs a dedicated bremsstrahlung recovery algorithm [86]. The electron track upstream of the magnet is extrapolated to the calorimeter, and the energy deposits from neutral clusters with transverse energy (E_T) above 75 MeV are added back to the electron's momentum. In $B^0 \rightarrow K^{*0} e^+ e^-$ decays, approximately half of the events have one bremsstrahlung cluster added, a quarter have two clusters added, and the remaining quarter have none. These event categories are referred to as Brem1, Brem2, and Brem0, respectively. Recovering bremsstrahlung photons results in a less biased momentum determination, a more symmetric momentum resolution, and improved particle identification. However, this approach has limitations, such as the ECAL's poorer energy resolution compared to the tracking system and the possibility of bremsstrahlung photons falling outside the ECAL acceptance or being too soft to be detected. Moreover, electron signals suffer from significant background contamination, partic-

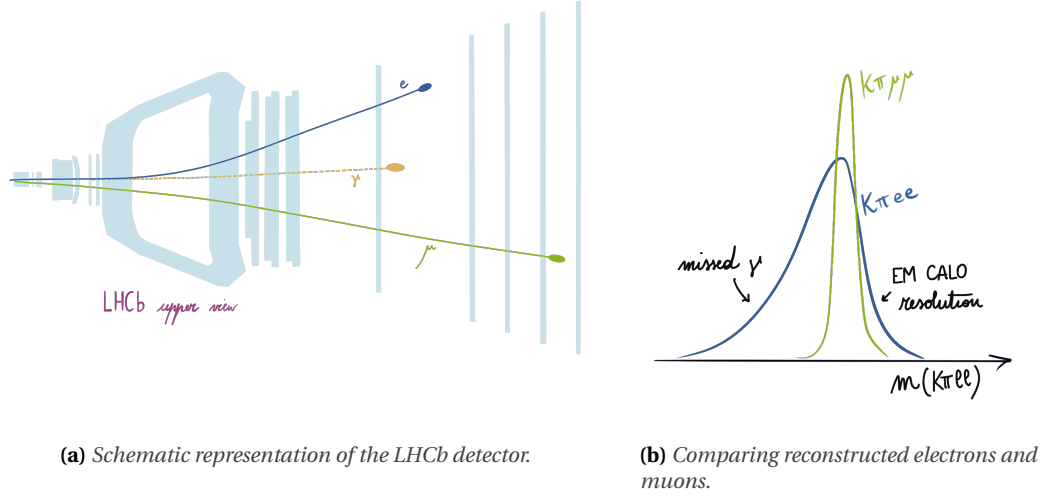


Figure 2.5: Figure (a) qualitatively shows the differences in trajectories between electrons, photons and muons during their path in the detector. If the brems photon (yellow) is emitted from the electron (blue) before the magnet, the photon cluster will be far from that of the electron. The muons (green) generally arrive undisturbed to the end of the detector. Figure (b) illustrates the difference in the reconstruction of the $m(K\pi\ell\ell)$ invariant mass between electrons and muons. [85].

ularly from hadrons and photons in the ECAL. To mitigate this, higher p_T thresholds are required for electrons, which reduces the efficiency of signal reconstruction [86, 87]. **Fig. 2.5b** illustrates the effect of reconstructing the $m(K\pi\ell\ell)$ invariant mass for the cases where ℓ corresponds to electrons or muons. One can observe larger tails in the electron channel, both on the left and on the right sides of the invariant mass distribution. The first arises from missed bremsstrahlung photons, while the right tail is a consequence of the ECAL's limited resolution, which is worse than the tracking resolution used for muons. Furthermore, the signal peak for the channel with electrons is also lower, due to the lower efficiency. These effects can be directly seen on real data: in **Fig. 2.6**, the invariant mass distributions of the final-state particles in $B^0 \rightarrow K^{*0}e^+e^-$ and $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decays highlight the electron channel's lower efficiency, poorer resolution, and higher background levels. These factors significantly reduce the signal-to-background ratio, posing major challenges for analyses involving electrons.

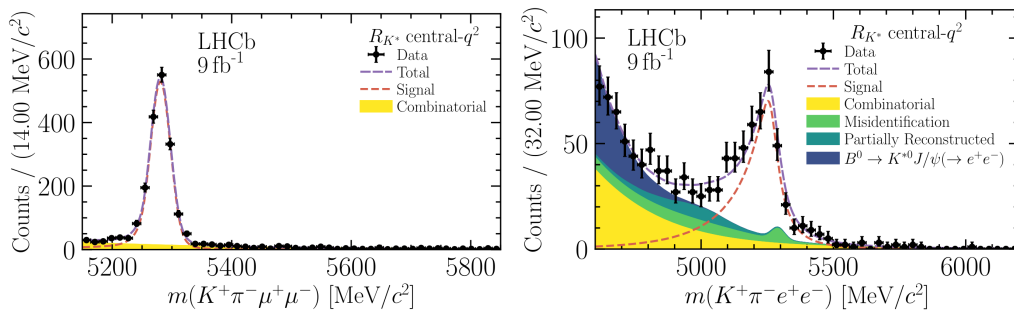


Figure 2.6: Invariant mass of the final state particles $K^+\pi^-\mu^+\mu^-$ and $K^+\pi^-e^+e^-$ on the left and right, respectively [49, 50].

2.7 $B^0 \rightarrow K^{*0}e^+e^-$ nomenclature and trigger lines

It is useful to define the relevant variables used in the present analysis.

Primary Vertex (PV) It is the position of the reconstructed primary pp interaction.

Decay Vertex (DV) It is the position of the reconstructed point where a certain particle X decayed.

Pseudorapidity (η) It is defined as $\eta = -\log(\tan \theta/2)$. It is another way to parametrise the θ angle with respect to the z -axis. It is widely used in collider physics because, in the limit of ultra-relativistic particles, the pseudorapidity approximates the rapidity, hence, it is Lorentz invariant for boosts along the beam direction.

N_{PV} It is the number of primary vertices for an event.

n_{SPD} It is the number of hits in the SPD (see **Sec. 2.4.1**).

θ_{DIRA} It is the direction angle, *i.e.* the angle between the momentum of the particle and the displacement vector, defined by the difference between the DV and PV of the particle.

Impact Parameter (IP) Distance of closest approach of a particle trajectory to a given point. Daughters of the B^0 normally have large impact parameters to the PV because of the displaced decay vertex of the B^0 meson. Instead, the IP of the B^0 meson itself to the PV tends to be closer to zero as the B^0 is produced in the initial collision corresponding to the PV.

Impact Parameter Chi square (χ_{IP}^2) Difference between the χ^2 of the primary vertex fit, obtained with and without considering the particle in the fit. If the particle does not come from the PV, the χ_{IP}^2 will generally be larger than the one obtained with a prompt particle.

Particle identification ($YPID_X$) It is computed through the difference between the logarithm of the likelihoods in the X and the π hypothesis. A high value of $YPID_X$ means a high probability for a particle Y to be of species X (see **Sec. 2.4.2**).

Track ghost probability ($\mathcal{P}_{ghost}$) It is the probability for a track to be a “ghost” track, *i.e.* a misidentified track, created from spurious combinations of hits. This quantity is calculated with a Neural Network [88], that combines information from different variables which describe the track reconstruction and global event properties to separate ghost tracks from real tracks.

Decay Tree Fitter (DTF) Algorithm which takes a complete decay chain, parametrises it in terms of vertex positions, decay lengths and momentum parameters. Then, it fits these parameters simultaneously, taking into account constraints such as measured parameters and 4-momentum conservation at each vertex [89].

Vertex fit (χ^2_{vtx}/ndf) It is the χ^2 value of the vertex fit, normalised to the degrees of freedom. The fitter takes as input a vector of particles, performs the fit, and updates the estimated decay vertex of the particle [90]. For signal, the average χ^2_{vtx}/ndf is expected to be close to one.

Track fit (χ^2_{trk}/ndf) It is the χ^2 value of the fit performed to the hits in the tracking detector normalised to the degrees of freedom. For more details on the track fit, please refer to Ref. [91]. For correctly reconstructed tracks, the average χ^2_{trk}/ndf is expected to be close to one.

Trigger On Signal (TOS) events that are explicitly triggered by the signal decay chain, *i.e.* one of the particles in the decay chain fires a certain line of the trigger.

Trigger Independent of Signal (TIS) events that are triggered independently of the presence of the signal, *i.e.* the trigger is fired by other particles in the events. As the trigger makes multiple decisions per event, an event can be both TIS and TOS simultaneously.

2.7.1 *Brem* categories definition

As discussed in **Sec. 2.6**, LHCb employs a dedicated algorithm to recover the energy loss of electrons (both positive and negative) due to bremsstrahlung emission. As a result it is possible to define three different categories:

- **Brem0:** this category includes candidates where no energy cluster is associated with the electrons.
- **Brem1:** this category includes candidates in which at least one energy cluster is recovered and associated with only one of the two electrons.
- **Brem2:** this category includes candidates in which both electrons have one or more energy clusters associated.

These category are graphically shown in **Fig. 2.7**

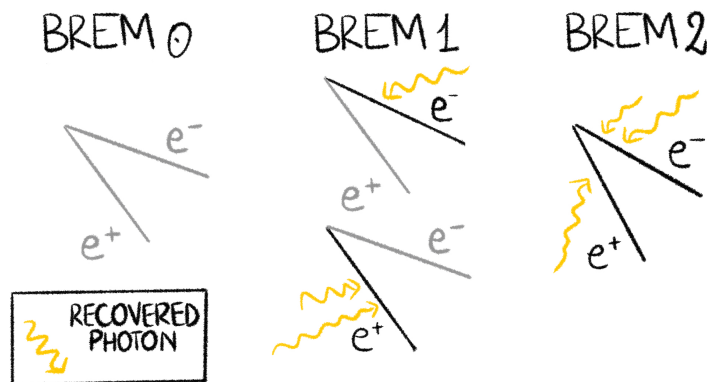


Figure 2.7: Illustration of the *brem* categories. The oscillating yellow arrows represent photons associated with the electron or positron by the bremsstrahlung recovery algorithm.

3 | Data samples and first stage selection

This chapter describes the first set of requirements applied to the data accepted by the trigger. Additionally, it details the major residual backgrounds that require further removal. The second part of the chapter focuses on the samples used in this analysis, which originate from the output of the LHCb data acquisition and trigger systems. The misidentified background sample has a dedicated subsection explaining its derivation.

3.1 $B^0 \rightarrow K^{*0} e^+ e^-$ candidates

The present study uses the data sample recorded by the LHCb experiment, from 2011 to 2018 (LHC Run 1 and Run 2) at centre-of-mass energies of 7, 8 and 13 TeV. The total integrated luminosity collected during this data-taking period is about $\sim 9 \text{ fb}^{-1}$. **Tab. 3.1** shows the luminosity collected at the LHCb experiment each year of the LHC run and the corresponding centre-of-mass energy ($\sqrt{s}$).

Year	Run 1		Run 2			
	2011	2012	2015	2016	2017	2018
$\sqrt{s}$ [TeV]	7	8		13		
$\int \mathcal{L}$ [fb^{-1}]	1.1	2.1	0.3	1.7	1.7	2.2

Table 3.1: centre-of-mass energy and inverse luminosity collected during the years at LHCb.

This work focuses on the $B^0 \rightarrow K^{*0} e^+ e^-$ decays and its charge conjugate $\bar{B}^0 \rightarrow \bar{K}^{*0} e^+ e^-$ ($\mathcal{BR}(B^0 \rightarrow K^{*0} e^+ e^-) \simeq 1.03 \times 10^{-6}$ [5]), where the K^{*0} ($\bar{K}^{*0}$) decays into a $K\pi$ pair about 100% of the time [5]. The latter decay occurs via strong interaction, it is instantaneous with respect to the vertex resolution, resulting in a final state with four particles: $B^0 \rightarrow K^+ \pi^- e^+ e^-$ or $\bar{B}^0 \rightarrow K^- \pi^+ e^+ e^-$ as explained in **Sec. 1.3**. For ease of reading, this decay and its charge conjugate are referred to as $B^0 \rightarrow K\pi e^+ e^-$ from now on¹. The tracks of the final state particles are reconstructed from the hits they left in the LHCb detector, as explained in **Chap. 2**. Finally, it is possible to infer the decay vertex of the B^0 particle by combining the four tracks.

However, multiple decays can produce the same final-state particles, making it essential to isolate the decay of interest from all others. An event that potentially contains the desired decay is referred to as a candidate. The candidate selection process is divided into online and offline selections. When the online selection is applied, discarded events are lost permanently. The offline selection, on the other hand, is performed later on and can be optimised as needed.

¹The inclusion of charge-conjugate will be implicitly assumed for all the decays throughout this document.

3.2 Online selection: trigger lines

Events generated in the pp collisions are selected using the trigger lines specifically designed to select decays with two electrons in the final state. For the nomenclature, please refer to **Sec. 2.5** and **Sec. 2.7**. Trigger lines chosen at L0 trigger are TOS on L0Electron together with events that fire *any* line among L0Hadron(TIS), L0Muon(TIS), and L0Electron(TIS). At HLT1, the trigger lines are of type *one track* (Run 1 and Run 2) and *two tracks* (Run 2). The first refers to lines with a decision based on a single reconstructed track, while the second refers to lines with a decision based on a two-body vertex [92]. HLT2 topological lines are designed to trigger efficiently on any B decay with at least 2 charged daughters. They are based on two-, three- and four-body combinations of displaced charged tracks [93, 86]. More details about the LHCb trigger lines can be found in Ref. [94].

3.3 Offline selection

The events passing the trigger selection are stored for further analysis and undergo the offline selection. The first step is called *stripping*. It aims at reducing the size of the sample, saving events relevant to a specific analysis. Subsequently, a series of cuts are applied, aiming at reducing backgrounds by selecting good-quality tracks and particles identified correctly. This set of requirements is referred to as *preselection*. Vetoes are also applied to reject specific backgrounds. Additionally, specific cuts are designed to select kaons and pions - from which the K^{*0} candidates are constructed - and B^0 candidates with reconstructed invariant mass around the B^0 pole mass ($\sim 5279 \text{ MeV}/c^2$ [5]). Finally, the kinematic region in q^2 where to perform the analysis needs to be selected. All these cuts are presented and motivated in this section.

3.3.1 Stripping

The requirements at this stage are loose and purely based on the reconstructed vertex quality and decay kinematics. The K^{*0} candidates are formed by combining kaons and pions of opposite charge. The e^+e^- system is formed by combining pairs of opposite sign electrons (corrected for bremsstrahlung emission, as explained in **Sec. 2.6**). Next, the B^0 candidate is reconstructed by combining the K^{*0} candidate and the e^+e^- pair. The list of the requirements of the *stripping* line used in this analysis can be found in Ref. [95].

3.3.2 Preselection

The requirements presented in this section ensure good-quality tracks and proper particle identification. If some requirements are shared with the *stripping* step, they are applied more tightly here. The cuts are listed in **Tab. 3.2**. All particles have a selection on their track fit χ^2 over degree of freedom and ghost probability (see **Sec. 2.7**), which should ideally be 1 and 0, respectively. Additionally, they are required to have associated clusters in the RICH sub-detector (hasRich). For electrons, they must be in the ECAL acceptance (InAccECAL). Furthermore, a fiducial cut is applied ($! (|x\text{Projection}_{\text{ECAL}}^{\text{LoCalo}}| < 363.6\text{mm} | y\text{Projection}_{\text{ECAL}}^{\text{LoCalo}}| < 282.6\text{mm})$). This removes a region of the ECAL that contains cells that are not properly read out [96], which can introduce discrepancies between data and simulated events. Kinematic requirements are also applied to the electron momentum ($p > 3000 \text{ MeV}/c$) and the transverse momentum ($p_T >$

500 MeV/c). Additional transverse momentum requirements are imposed on the K and π ($p_T > 250$ MeV/c). These kinematic requirements enhance the chance that these particles originate from a b -hadron decay. Particle identification cuts are applied using the $YPID_X$ and $yProbNNx$ variables, described in **Sec. 2.4.2**. Additional selection criteria are applied on the angle between two final-state tracks: if the angle is less than 0.5 mrad the candidate is rejected, ensuring the removal of *clone* particles, defined as tracks that share at least 70% of hits in the VELO and part of the T-stations [97]. Finally, a few cuts are dedicated to the rejection of backgrounds arising from electron-hadron and hadron-hadron misidentification. These are standard cuts for this decay, they were also applied in Refs. [49, 50] and [39].

		Type	Requirement
Quality		all tracks	$\chi^2_{\text{trk}}/\text{ndf} < 3$ $\mathcal{P}_{ghost} < 0.4$
			$\text{region}_{\text{ECAL}}^{\text{LOCalo}} \geq 0$
	e		$!(x\text{Projection}_{\text{ECAL}}^{\text{LOCalo}} < 363.6 \text{ mm})$
			$ y\text{Projection}_{\text{ECAL}}^{\text{LOCalo}} < 282.6 \text{ mm})$
			InAccECAL $\sqrt{p_{x,track}^2 + p_{y,track}^2} > 200$
Clone	K, π, e		$\theta(\pi, e) > 0.0005, \theta(K, e) > 0.0005, \theta(K, \pi) > 0.0005$
PID	all		hasRich
	e		hasCalo
	K, π		$p_T > 250 \text{ MeV}/c$
	e		$p_T > 500 \text{ MeV}/c, p > 3000 \text{ MeV}/c$
	K		$K\text{ProbNNK} \cdot (1 - K\text{ProbNNp}) > 0.05$
	π		$\pi\text{ProbNN}\pi \cdot (1 - \pi\text{ProbNNK}) \cdot (1 - \pi\text{ProbNNp}) > 0.1$
	e		$e\text{ProbNNe} > 0.2$
	K		$KPID_K > 0$
	e		$ePID_e > 2$
BKG	$c\bar{c} e - h \text{ swap}^*$		$!(m(K \rightarrow e \pi e \rightarrow K e)_{J/\psi, \psi(2S)} - m_{B^0}^{\text{PDG}} < 60 \text{ \& } e\text{ProbNNe} < 0.8 \text{ OR } m(K \pi \rightarrow e e e \rightarrow \pi)_{J/\psi, \psi(2S)} - m_{B^0}^{\text{PDG}} < 60 \text{ \& } e\text{ProbNNe} < 0.8)$
	$K - \pi \text{ swap}$		$KPID_K - \pi PID_K > 0$
	$B_s^0 \rightarrow \phi e^+ e^-$		$! m(K \pi \rightarrow K) < 1040 \text{ MeV}/c^2 \text{ \& } \pi\text{ProbNN}\pi < 0.8$

Table 3.2: List of offline requirements relative to the quality of the tracks and particle identification criteria. For nomenclature refer to **Sec. 2.7**.

*This selection concerns the $c\bar{c}$ resonances J/ψ and $\psi(2S)$. In particular, it acts to reduce background coming from the $e - h$ swap that could leak into the rare mode region due to their abundance. This cut is part of the standard cuts applied in such decay channels.

3.3.3 Backgrounds vetoes

A series of selection criteria is dedicated to rejecting specific misidentified semi-leptonic and hadronic background decays. The vetoes presented in this section build upon the results from Ref. [98, 99]. The vetoes are studied by examining the reconstructed masses of two or three-body final-state particles under the hypothesis that one electron has been misidentified. In practical terms, this can be done by assigning the mass of a π or a K to the misidentified electron when computing the invariant masses. Since hadrons emit a negligible amount of bremsstrahlung radiation, these requirements are only applied when there are no bremsstrahlung photons associated with the electrons (Brem0 category, see Sec. 2.7). Additionally, the resonances targeted by these vetoes occur only when the misidentified electron has the same charge as the hadron ($e_{\rightarrow h}^{SS}$), not the other way around² ($e_{\rightarrow h}^{WS}$), as the former involves Cabibbo-favoured processes (see Sec. 1.1.2).

The plots in Fig. 3.1 confirm these considerations, showing these effects for the $m(K e_{\rightarrow \pi}^{SS(WS)})$ reconstructed masses, where SS(WS) indicates when the misidentified electron used to reconstruct the two-body mass has the same(opposite) charge as(to) the π . All plots show the data sample at the *stripping* selection stage. Fig. 3.1a displays the peak of the $D^0 \rightarrow K\pi$ decay for $m(K e_{\rightarrow \pi}^{SS})$. In contrast, Fig. 3.1b shows that the D^0 peak disappears for $m(K e_{\rightarrow \pi}^{WS})$, as expected. Fig. 3.1c presents $m(K e_{\rightarrow \pi}^{SS})$ with an additional requirement that candidates belong to the Brem1 or Brem2 categories. Under these conditions, the hadronic peak is no longer visible. This pattern applies to all hadronic particles affected by these vetoes.

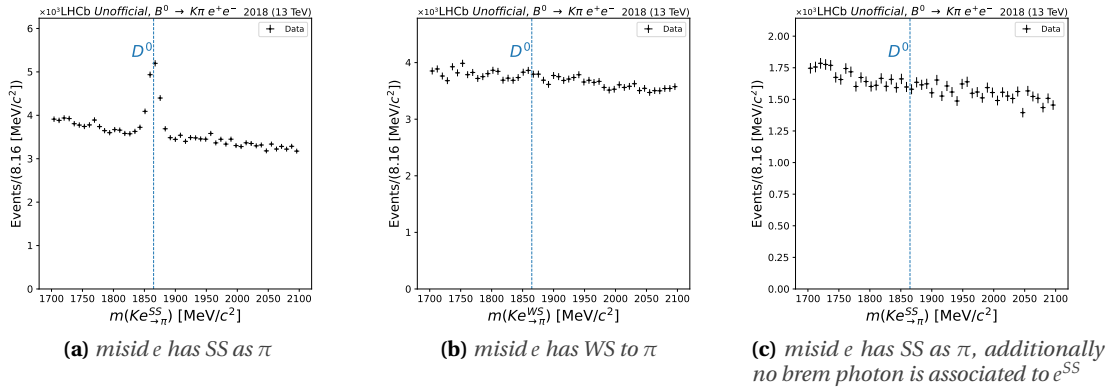


Figure 3.1: All the plots shows the $B^0 \rightarrow K\pi e^+ e^-$ data sample at the *stripping* selection stage. They show the two-body reconstructed mass $m(K e_{\rightarrow \pi})$ in different scenarios. The resonance shows up only when the electron has the same charge (SS) as the pion and when no associated bremsstrahlung photon is present.

As a result, the vetoes are applied only to electrons with the same charge as the pion and only when no associated bremsstrahlung photon is present. This ensures a very high signal efficiency for these cuts. The vetoes are listed in Tab. 3.3, where the first column indicates the vetoed background decay. The column $m(X)$ specifies the final-state particles invariant mass for which the requirement is applied, highlighting the assumed misidentification ($e \rightarrow h$). This denotes which mass assumption (m_h^{PDG}) has been assigned to the electron in order to compute the reconstructed invariant mass.

²Backgrounds involving an electron with the opposite charge to the hadron do exist, but they are Cabibbo-suppressed, making them much less likely to occur (see Sec. 1.1.2).

	Type	$m(X)$	Requirement
Vetoed	$\phi(\rightarrow KK)\pi$	$m(K\pi\rightarrow K)$	$ m_\phi^{PDG} - m(X) < 15 \text{ MeV}/c^2 \text{ \& } e\text{ProbNNK} > 0.7$
	$B^0 \rightarrow D^\pm(\rightarrow K\pi\pi^\pm)e^\mp\nu_e$	$m(K\pi e\rightarrow\pi)$	$ m_D^{PDG} - m(X) < 20 \text{ MeV}/c^2 \text{ \& } e\text{ProbNN}\pi > 0.95$
	$B^0 \rightarrow D^{*\pm}(\rightarrow K\pi\pi^\pm)e^\mp\nu_e$	$m(K\pi e\rightarrow\pi)$	$ m_{D^*}^{PDG} - m(X) < 30 \text{ MeV}/c^2 \text{ \& } e\text{ProbNN}\pi > 0.98$
	$D^0(\rightarrow K\pi)\pi^\pm e^\mp\nu_e$	$m(K\pi\rightarrow\pi)$	$ m_{D^0}^{PDG} - m(X) < 30 \text{ MeV}/c^2 \text{ \& } e\text{ProbNN}\pi > 0.8$
	$B_s^0 \rightarrow D_s^\pm(\rightarrow KK\pi)e^\mp\nu_e$	$m(K\pi e\rightarrow K)$	$ m_{D_s}^{PDG} - m(X) < 30 \text{ MeV}/c^2 \text{ \& } e\text{ProbNNK} > 0.7$
	$\Lambda_b^0 \rightarrow \Lambda_c^+(\rightarrow K\pi p)e^-\nu_e$	$m(K\pi e\rightarrow p)$	$ m(\Lambda_c^+) - m(X) < 30 \text{ MeV}/c^2 \text{ \& } e\text{ProbNN}p > 0.4$

Table 3.3: Offline requirements to veto misidentified semi-leptonic and hadronic backgrounds. Columns show the vetoed decay, the reconstructed mass - highlighting the misidentification -, and the applied requirements.

Fig. 3.2 illustrates the effect of these vetoes on the $B^0 \rightarrow K\pi e^+ e^-$ data sample, with each plot corresponding to an $m(X)$ value from **Tab. 3.3**. Black dots represent the data at the *stripping* selection stage, where each target resonance is visible, marked by a light blue dashed line and labelled with its name. The blue dots show the same dataset after applying all vetoes. Each plot includes a bin-by-bin ratio before and after the vetoes, demonstrating that they primarily remove hadronic peaks while keeping the rest of the distribution flat.

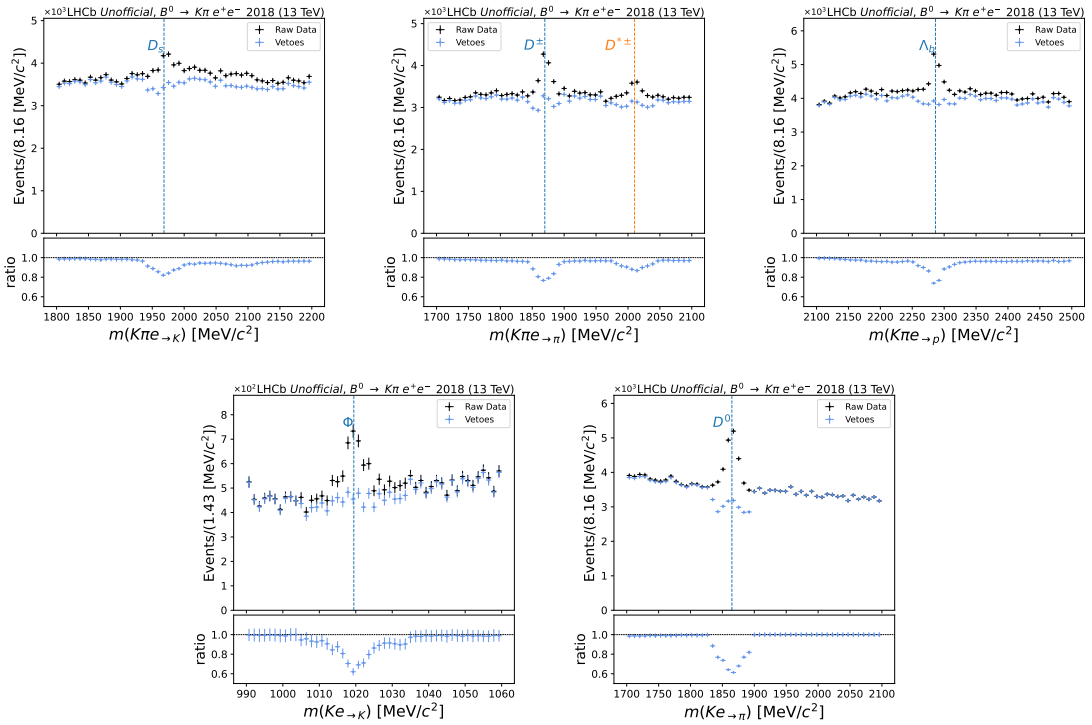


Figure 3.2: Effect of the vetoes listed in **Tab. 3.3** on the $B^0 \rightarrow K\pi e^+ e^-$ data sample at the *stripping* selection stage. Each plot corresponds to an invariant mass $m(X)$ from **Tab. 3.3**. The black dots represent the data sample, with the invariant masses of the targeted resonances marked by light blue dashed lines and labelled with their names. The blue dots show the same dataset after applying all vetoes.

3.3.4 K^{*0} selection

K^{*0} candidates are reconstructed by combining oppositely charged kaons and pions. However, not all combinations originate from actual K^{*0} candidates. The quality of a candidate is assessed using the invariant mass distribution of the kaon-pion system ($m(K\pi)$). High quality K^{*0} candidates cluster around the pole mass of the K^{*0} ($m_{K^{*0}}^{PDG} \sim 895 \text{ MeV}/c^2$ [5]). Therefore, a selection around this mass is applied, as shown in **Tab. 3.4**.

Requirement	
K^{*0}	$ m(K\pi) - m_{K^{*0}}^{PDG} < 100 \text{ MeV}/c^2$

Table 3.4: Criteria on the invariant mass $m(K\pi)$, to clean sample from K^{*0} mesons that do not resonate.

3.3.5 B^0 mass region

To improve signal purity, a region around the $m_{B^0}^{PDG} \sim 5279 \text{ MeV}/c^2$ [5] in the spectrum of the invariant mass of the final-state particles, $m(K\pi e^+ e^-)$, is selected. The requirement is shown in **Tab. 3.5**. The requirement aligns with Ref. [49, 50].

Requirement	
B^0	$4600 \text{ MeV}/c^2 < m(K\pi e^+ e^-) < 6200 \text{ MeV}/c^2$

Table 3.5: Selection requirement on the reconstructed invariant mass of the four final-state particles $m(K\pi e^+ e^-)$. This cut removes events that are inconsistent with the hypothesis that they originate from the decay of a B^0 meson.

3.3.6 q^2 selection

This analysis is performed separately in different q^2 regions, as mentioned in **Sec. 1.5**. The focus of this thesis is on the central region ($[1.1, 6.0] \text{ GeV}^2/c^4$), which is further divided into two sub-regions ($[1.10, 3.65] \text{ GeV}^2/c^4$ and $[3.65, 6.00] \text{ GeV}^2/c^4$), chosen to ensure equal signal content in both bins. The selection of q^2 is performed on the q^2 variable computed using the momenta of the final state particles obtained from the DTF algorithm, see **Sec. 2.7**. In particular, the algorithm performs a kinematic fit to determine these momenta, whilst constraining the B^0 origin vertex to the observed primary vertex position, and similarly, the momenta are constrained such that their invariant mass equals the nominal B^0 mass. This quantity is referred to as $q_{PV\&B}^2$. A rectangular cut on $q_{PV\&B}^2$ helps reducing leakage from $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ decays at small values of $m(K\pi e^+ e^-)$, though it introduces some combinatorial events at larger $m(K\pi e^+ e^-)$ values, as shown in **Fig. 3.3**. These combinatorial events can be further vetoed by cutting the invariant mass squared of the two electrons (q^2 without constraints), $q_{unconst}^2$. To summarise, the cuts for each kinematic region are reported in table **Tab. 3.6**.

q^2 region	Requirement
low q^2	$(0.1 < q_{PV\&B}^2 < 1.0) \text{ GeV}^2/c^4$
central q^2	$(1.1 < q_{PV\&B}^2 < 6.0 \text{ \& } q_{unconst}^2 < 7.0) \text{ GeV}^2/c^4$
central _{low} q^2	$(1.10 < q_{PV\&B}^2 < 3.65 \text{ \& } q_{unconst}^2 < 7.00) \text{ GeV}^2/c^4$
central _{high} q^2	$(3.65 < q_{PV\&B}^2 < 6.00 \text{ \& } q_{unconst}^2 < 7.00) \text{ GeV}^2/c^4$

Table 3.6: Selection requirements for each kinematic region. The quantity $q_{PV\&B}^2$ is reconstructed with the DTF algorithm as described in the text, while $q_{unconst}^2$ is the dielectron invariant mass squared.

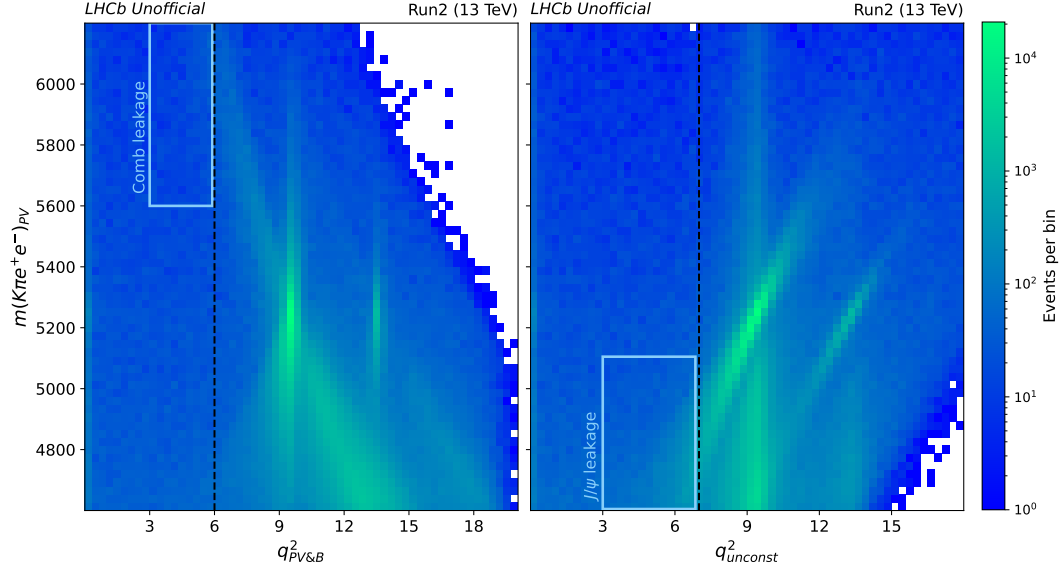


Figure 3.3: Distribution of $m(K\pi e^+ e^-)$ as a function of q^2 (right) and $q_{PV\&B}^2$ (left) using Run 2 data at 13 TeV, all the offline selection is applied. The regions corresponding to combinatorial background leakage (left) and $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ leakage (right) are highlighted with light blue squares. The cut on $q_{PV\&B}^2$ removes most of the $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ leakage, while the cut on the $q_{unconst}^2$ removes residual combinatorial background. The black dashed vertical lines at $q_{PV\&B}^2 = 6 \text{ GeV}^2/c^4$ and $q_{unconst}^2 = 7 \text{ GeV}^2/c^4$ mark the boundaries of the selection.

3.4 Residual backgrounds

This section details the residual backgrounds after applying the requirements presented in **Sec. 3.3**. The shape and position of these backgrounds relative to the signal are illustrated in **Fig. 3.4** for the $m(K\pi e^+ e^-)$ observable. This drawing is purely qualitative and not to scale, it is intended solely to aid in visualising the various backgrounds affecting the $m(K\pi e^+ e^-)$ spectrum while reading this section.

- The **combinatorial background** arises from pairing unrelated tracks that do not come from the decay of interest, but accidentally satisfy all the selection criteria. However, the probability of observing a specific invariant mass from random combinations decreases as the mass increases. This is because the phase space for a combination with a larger invariant mass is smaller in terms of the available energy and momentum.
- The **Partially Reconstructed (PR) background** comes from a decay of the B^0 featuring heavier kaon resonances, such as $B^+ \rightarrow K_1^+ (\rightarrow K^+ \pi^+ \pi^-) e^+ e^-$ and $B^+ \rightarrow K_2(1430)^+ (\rightarrow K^+ \pi^+ \pi^-) e^+ e^-$, which can be reconstructed as the signal with one or more pions missing.

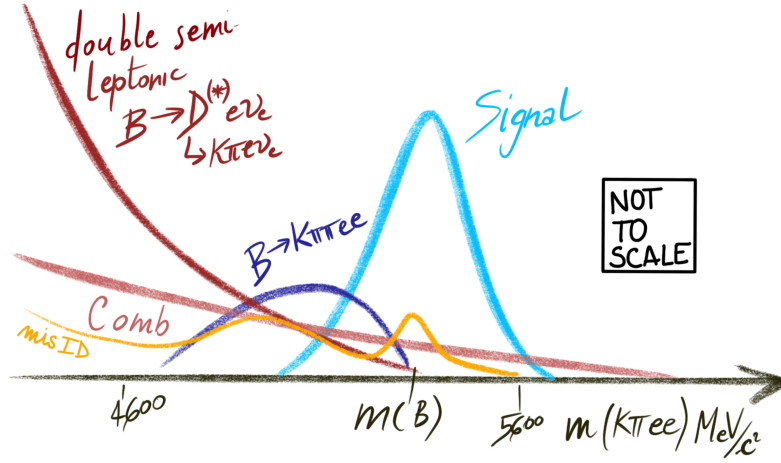


Figure 3.4: Illustration of the residual background components in the $m(K\pi e^+ e^-)$ spectrum. The drawing is not to scale and has the sole purpose of familiarising the reader with the different backgrounds.

This type of background generally shows up in the $m(K\pi e^+ e^-)$ distribution as a broad peak shifted towards values below the B^0 mass, due to the missing energy of particles that have not been included.

- The **Double Semi-Leptonic (DSL) background** consists of $B^0 \rightarrow D^+ (\rightarrow K^{*0} e^+) e^- \nu_e$ decays that have the same visible decay product as the $B^0 \rightarrow K^{*0} (\rightarrow K\pi) e^+ e^-$ signal decay. These decays are particularly problematic because they produce the same visible final state particles as the signal, pass all the PID and K^{*0} requirements, and have branching fractions of $\mathcal{O}(10^{-3})$, roughly a thousand times larger than the signal. Furthermore, this background exhibits a characteristic asymmetric $\cos \theta_\ell$ distribution, with a peak around one. This can be explained by two factors: first, θ_ℓ is defined as the angle between the B^0 ($\bar{B}^0$) and the positron (electron). Second, the positron (electron) from the B^0 ($\bar{B}^0$) typically has higher momentum than the one from the D meson. The higher momentum of the positron (electron) causes θ_ℓ to shift towards smaller values.
- The **hadronic misidentified background (misID)** consists of a mixture of decays in which a K or π is misidentified as an electron and it passes the selection criteria. This includes both single and double misidentification cases, where one or both electrons are actually hadrons. In the case of double misidentification, the background comes from the $B^0 \rightarrow K^{*0} h_1 h_2$ decays, where h represents either a K or a π . If h is a pion, the background peaks closer to the B^0 mass compared to the case where h are kaons, as the mass difference between electrons and pions is smaller than between electrons and kaons. For single misidentification, the background arises from a wide range of decays of the type $B^0 \rightarrow K^{*0} \pi^- (\pi^0, \gamma) X$, where X is any number of other final state particles [49, 50]. These backgrounds predominantly populate the lower $m(K\pi e^+ e^-)$ region. Due to the diverse nature of this background, which arises from multiple decay modes without a single dominant contribution, it cannot be easily vetoed. While each individual contribution is small, their combined effect is

significant.

- The J/ψ **leakage background**, as the name suggests, is the contamination from the $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ decays that has a branching ratio of $\mathcal{O}(10^{-3})$, three orders of magnitude larger than the decay of interest. Normally, by selecting q^2 regions below $6 \text{ GeV}^2/c^4$, most of these decays are removed as they resonate around $m(J/\psi)^2 \approx 9.6 \text{ GeV}^2/c^4$. However, a fraction of these events is reconstructed with a lower invariant mass than the true J/ψ resonance ($\sim 3096 \text{ MeV}/c^2$ [5]) due to detector resolution effects, leading to a low-mass tail that extends into the selected $m(K\pi e^+ e^-)$ region.

3.5 Samples of the analysis

To perform the analysis, the data sample collected at LHCb is not sufficient, it is necessary to use MC simulations of the signal decay $B^0 \rightarrow K^{*0} e^+ e^-$ and of backgrounds. Sometimes, data samples from other decays can be useful as well. This is the case for the lepton-flavour-violating $B^0 \rightarrow K^{*0} e^\pm \mu^\mp$ decay. This decay is forbidden by the SM, and the current upper limit on the branching ratio is 10^{-8} at 90% CL [5], which is negligible for the purposes of this analysis. This data sample mainly contains combinatorial and DSL decay of the type: $B^0 \rightarrow D^+ (\rightarrow K^{*0} e^+) \mu^- \nu_\mu$ and $B^0 \rightarrow D^+ (\rightarrow K^{*0} \mu^+) e^- \nu_e$. Therefore, it is suitable to study both combinatorial and DSL backgrounds. The list in **Tab. 3.7** shows all the samples used to perform this analysis. It specifies the sample as *data* to differentiate events collected from LHCb from simulated decays, labelled as *MC* samples. The last sample in the list (misID) differs from the rest, as it is a sample made by events estimated using a data-driven approach, introduced by Ref. [49, 50] to obtain models for misidentified hadronic decays. The methodology used to estimate it is detailed in **Sec. 3.5.2**.

Name	Type	Decay
$B^0 \rightarrow K\pi e^+ e^-$ data	<i>Data</i>	$K\pi e^+ e^-$
$B^0 \rightarrow K^{*0} e^+ e^-$	<i>MC</i>	$B^0 \rightarrow K^{*0} e^+ e^-$
$B^0 \rightarrow K^{*0} J/\psi$	<i>MC</i>	$B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$
$B \rightarrow K\pi\pi e^+ e^-$	<i>MC</i>	$B^+ \rightarrow K\pi \pi^+ e^+ e^-$
$B^0 \rightarrow K^{*0} e\mu$	<i>Data</i>	$B^0 \rightarrow K^{*0} e^\pm \mu^\mp$
misID	<i>Data driven*</i>	$B^0 \rightarrow K^{*0} \pi^- (\pi^0, \gamma) X, B^0 \rightarrow K^{*0} h_1 h_2$

Table 3.7: The table lists the samples used in this analysis, showing the reference name, the type of sample (data or MC), and the corresponding decays they describe. All these decays imply their charge-conjugate processes as well.

*The estimation of this dataset is detailed in **Sec. 3.5.2**.

3.5.1 MC simulations and corrections

Although this thesis does not focus on MC simulation corrections, it is important to highlight the key sources of discrepancies and how they have been mitigated. Given the extensive use of MC simulations in this analysis, it is crucial that they reliably model the data. In particular, the generated B meson kinematics and event multiplicity should match the data, the trigger must be properly emulated with its performance well described, the event reconstruction should be accurately represented, and the particle identification performance should be properly characterised.

Any misreconstruction or mismodelling effect that alters the shape of the angular distributions potentially introduces uncertainties in the final result.

MC simulations employ PYTHIA [100, 101] to generate pp collisions, using a configuration specifically tuned for LHCb [102]. The decays of unstable particles are modelled with EvtGen [103], while final-state radiation effects are handled by PHOTOS [104]. The interactions between the generated particles and the detector, as well as the detector response, are simulated using the Geant4 framework [105, 106, 107]. Despite these efforts to faithfully reproduce the experimental conditions, discrepancies remain due to potential mis-modelling of the pp collisions or the detector response. To mitigate these discrepancies, a correction procedure is applied. The primary approach is to first address discrepancies related to detector effects (such as tracking efficiency, trigger, and particle identification), followed by corrections for physics-related discrepancies (such as kinematics). Each correction is computed sequentially, with later corrections accounting for the effects of the previous ones. This approach ensures that possible correlations among all the effects are considered. Each correction results in a per-event weight (w), unless stated otherwise. They are performed in the following order:

1. w_{TCK} : correction for electrons and hadrons track reconstruction efficiency is computed using the *tag-and-probe* method [108, 87].
2. w_{MULT} : a second correction is used to align the event multiplicity between data and MC. For this purpose, the $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ data sample with background subtracted is used. This correction specifically reshapes the event multiplicity variable, rather than affecting other dataset variables. This approach is based on the fact that the observed discrepancy is not due to incorrect underlying physics processes, but rather the result of a poorly simulated detector response. This reshaped variable is then used to compute the subsequent MC correction for particle identification.
3. w_{PID} : another correction targets the particle identification response. The simulated PID performance deviates from that observed in data, primarily due to detector occupancy effects and variations in operational conditions across data-taking periods. This discrepancy is addressed through a common practice called PID calibration procedure that exploits high-purity control samples [109]. These include $D^{*+} \rightarrow D^0 \pi^+$, $\Lambda^0 \rightarrow p \pi^-$, and $\Lambda_c^+ \rightarrow p K^+ \pi^-$ for hadrons, $J/\psi \rightarrow \mu^+ \mu^-$ for muons, and $B^+ \rightarrow J/\psi (\rightarrow e^+ e^-) K^+$ for electrons.
4. $w_{\text{L0(HLT)}}$: a correction is applied to account for mismodelled trigger efficiencies, which can arise because simulation may not accurately reproduce trigger performances, especially due to varying detector occupancy. To address this, a data-driven approach is used, leveraging the abundant $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ decays to compute the L0 (HLT) trigger efficiency. The correction is applied by taking the ratio of trigger efficiencies extracted from data and from the MC simulation and applying this weight to the MC sample.
5. w_{KIN} : another key correction addresses discrepancies in the kinematics of B mesons. The most significant modelling discrepancies affect variables such as the transverse momentum (p_T) and pseudorapidity (η) of the B^0 candidate. To compensate for this, a bi-dimensional re-weighting procedure is implemented, to assign a weight that varies according to the B^0 p_T and η values.

The product of those weights provides the final correction to be applied to each MC candidate used in this analysis. An additional correction is applied to the $B \rightarrow K\pi\pi ee$ MC sample: since this is a five-body decay reconstructed as a four-body decay, it requires further kinematic corrections.

When a MC sample is used in this thesis, the application of these corrections is always implied, even when not explicitly stated.

3.5.2 Misidentified background sample

The methodology used to estimate the misidentified background in the $B^0 \rightarrow K\pi e^+ e^-$ dataset exploits the electrons that fail the particle identification criteria presented in **Sec. 3.3.2**. This practice is generally known as *method of sideband extrapolation*. The misidentified background is estimated per data-taking period (Run 1 and Run 2) - not per year -, due to the limited amount of electrons failing the PID criteria, as the cut $PID_e > 0$ is applied at the *stripping* level. The use of a distinct shape for Run 1 is based on the fact that there is no a priori reason to assume that the residual misidentified backgrounds would have the same shape in both runs, as particle identification algorithms have changed between Run 1 and Run 2. The final shape of the misidentified background strongly depends on the final selection applied to the data. The final selection is presented in the following chapter, where the optimisation of the current selection is discussed. Nevertheless, this section details the procedure to estimate the misidentified background composed of decays with single and double misidentified electrons, which can be outlined as follows:

1. Define the pass and fail criteria for the particle under study in terms of misidentification. In this case, these criteria correspond to the PID requirements for electrons seen in **Sec. 3.3.2**:
 - $\text{Pass}(Y): (Y_{\text{ProbNNe}} > 0.2 \ \& \ Y_{\text{PIDe}} > 2)$
 - $\text{Fail}(Y): \neg(Y_{\text{ProbNNe}} > 0.2 \ \& \ Y_{\text{PIDe}} > 2) \ \& \ Y_{\text{PIDe}} > 0$, where Y denotes the particle to which the requirement is applied. The $e_{\text{PIDe}} > 0$ is applied at the stripping level, as mentioned before.
2. Compute a scale function on a sample enriched with decays to real K and real π . For this study, the same calibration sample used for the particle identification algorithm PID-Calib [109] is exploited: a data sample of $D^{*-} \rightarrow D^0(\rightarrow K\pi)\pi^-$ candidates, with their background subtracted. The scale function is defined as $w_{\text{fake}}(\vec{X}) = \frac{N_{\text{Pass}}(\vec{X})}{N_{\text{Fail}}(\vec{X})}$. Where $N_{\text{Pass(Fail)}}(\vec{X})$ represents the number of events that pass (fail) the electron PID criteria defined earlier, and $\vec{X}$ represent all the dependencies of the scale functions. However, in this case, these criteria are applied to kaons and pions. This approach allows us to compute the efficiency of real kaons or pions passing the electron PID criteria over those failing them.
3. Finally, applying the scale function to electrons that fail the PID criteria in the $B^0 \rightarrow K\pi e^+ e^-$ dataset provides an estimate of the residual contribution of real kaons and pions that pass the electron PID criteria.

The following paragraph provides more details for each step presented above.

Scale Function

The scale function $w_{fake}(\vec{X})$ is computed from $D^{*-} \rightarrow D^0(\rightarrow K\pi)\pi^-$ calibration samples, with background subtracted. These decays are selected such that the kinematic and detector acceptance requirements are aligned as closely as possible to those presented in **Sec. 3.3**. Scale factors maps, w_{fake} , are computed as the ratio of the number of events passing the PID criteria over those failing them, in bins of $\log_{10} p_T$ and η . Since both pions and kaons misidentified as electrons are of interest, $w_{fake}(p_T, \eta)$ are divided into two categories: $w_{fake}^{\pi, K}(p_T, \eta)$ based on whether $K(\pi)\text{ProbNNk}$ is above (below) 0.1, respectively. To further align the calibration sample with the analysis samples, an additional categorisation is introduced based on the events trigger status, except for Run 1, as the trigger status is unavailable for this specific calibration sample. The cases where the triggered electron is part of the B^0 decay signal (or not) are separated and denoted here as LOElectronTOS ($!\text{LOElectronTOS}$) events. This formulation ensures a more precise estimation of the scale factors. In summary, the scale factors are computed in the following four categories and bins of $\log_{10} p_T$ and η :

1. Map for π -like LOElectronTOS

$$w_{fake}^{\pi, \text{TOS}} = \frac{\text{Pass}(\pi) \& \pi\text{ProbNNk} < 0.1 \& \text{LOElectronTOS}}{\text{Fail}(\pi) \& \pi\text{ProbNNk} < 0.1 \& \text{LOElectronTOS}} (\log_{10}(p_T), \eta)$$

2. Map for π -like $!\text{LOElectronTOS}$

$$w_{fake}^{\pi, !\text{TOS}} = \frac{\text{Pass}(\pi) \& \pi\text{ProbNNk} < 0.1 \& !\text{LOElectronTOS}}{\text{Fail}(\pi) \& \pi\text{ProbNNk} < 0.1 \& !\text{LOElectronTOS}} (\log_{10}(p_T), \eta)$$

3. Map for K -like LOElectronTOS

$$w_{fake}^{K, \text{TOS}} = \frac{\text{Pass}(K) \& K\text{ProbNNk} > 0.1 \& \text{LOElectronTOS}}{\text{Fail}(K) \& K\text{ProbNNk} > 0.1 \& \text{LOElectronTOS}} (\log_{10}(p_T), \eta)$$

4. Map for K -like $!\text{LOElectronTOS}$

$$w_{fake}^{K, !\text{TOS}} = \frac{\text{Pass}(K) \& K\text{ProbNNk} > 0.1 \& !\text{LOElectronTOS}}{\text{Fail}(K) \& K\text{ProbNNk} > 0.1 \& !\text{LOElectronTOS}} (\log_{10}(p_T), \eta)$$

The fact that the scale factors vary depending on the characteristics of each candidate justifies referring to them as scale function $w_{fake}(\vec{X})$, where $\vec{X}$ represent the characteristics: p_T, η, K -like, π -like, TOS , and $!\text{TOS}$ of each electron.

Derivation of misidentified background shape

Each B^0 candidate in the $B^0 \rightarrow K\pi e^+ e^-$ dataset falls in one of the following orthogonal scenarios:

- a) Both electrons satisfy the PID requirements.
- b) Opposite Sign (OS): the e which fails the PID cut has the opposite charge to the kaon, while the other passes the requirement, $B^0 \rightarrow K^+ \pi^- e_{pass}^+ e_{fail}^-$.
- c) Same Sign (SS): the e which fails the PID cut has the same charge to the kaon, while the other passes the requirement, $B^0 \rightarrow K^+ \pi^- e_{fail}^+ e_{pass}^-$.

- d) Both electrons fail the PID requirements. This scenario is referred to as ee , to indicate both electrons failed the criteria.

For each B^0 candidate, electrons failing the PID criteria are classified in the same 4 categories identified when computing the scale function:

1. π -like^{TOS}: $eProbNNk < 0.1 \ \& \ LOElectronTOS$
2. π -like^{!TOS}: $eProbNNk < 0.1 \ \& \ !LOElectronTOS$
3. K -like^{TOS}: $eProbNNk > 0.1 \ \& \ LOElectronTOS$
4. K -like^{!TOS}: $eProbNNk > 0.1 \ \& \ !LOElectronTOS$

To estimate the shape of the misidentified component, the idea is to apply the scale function to the failing electron of the corresponding category and kinematic bin. Candidates in data falling in scenarios b) and c) have just one of the two electrons undergoing the weighting procedure. The $w_{fake}^{ij}(p_T, \eta)_k$ depend on the nature of the electron track, namely its kinematic bin $(p_T, \eta)_k$, whether the track is pion or kaon-like ($i = \pi, K$) and whether or not $LOElectron$ is TOS ($j = TOS, !TOS$). For simplicity and to highlight the per candidate nature, weights belonging to scenarios b) and c) are defined as follows: $w_{fake,i}^{eOS(SS)}$, whether the electron failing has the opposite (OS) or same (SS) charge of the K , while i , this time, run over the B^0 candidates. Candidates belonging to the scenario c) are weighted according to the product of the scale factors evaluated on both tracks, $w_{fake}^{i_1,j_1}(p_T, \eta)_{k_1} w_{fake}^{i_2,j_2}(p_T, \eta)_{k_2} = w_{fake}^{ee}$. Finally, to find the shape of the misidentified background, the weights are combined: $+1 \cdot w_{fake,i}^{eOS} + 1 \cdot w_{fake,i}^{eSS} - 1 \cdot w_{fake,i}^{ee}$, where the last term is added to avoid double counting³. Following this lead, by summing all the weights, it is possible to estimate the number of misidentified events as in the following equation:

$$\overline{N} = \sum_i^{N_{OS}} w_{fake,i}^{eOS} + \sum_j^{N_{SS}} w_{fake,j}^{eSS} - \sum_k^{N_{ee}} w_{fake,k}^{ee} \quad (3.1)$$

and to compute the associated statistical uncertainty as follows:

$$\overline{\sigma}^2 = \sum_i^{N_{OS}} (w_{fake,i}^{eOS})^2 + \sum_j^{N_{SS}} (w_{fake,j}^{eSS})^2 + \sum_k^{N_{ee}} (w_{fake,k}^{ee})^2 \quad (3.2)$$

The shape of the expected misidentified background as a function of the reconstructed B^0 invariant mass at this stage of the computation is shown in the first column of **Fig. 3.5**. The solid blue histograms represent the data before reweighting, while the orange distributions show the data after applying the $w_{fake}(\vec{X})$ weights in each of the categories (SS, OS, and both electrons failing). The last plot of the first column shows the total distribution when combining the $w_{fake}(\vec{X})$ of each category, and the expected quantity of misidentified events is written on the plot under the name of $\sum(w_i)$. These shapes are obtained using the final selection, including the optimisation detailed in **Chap. 4**.

However, this strategy does not account for the contribution from genuine signal events that fall into the region failing the PID criteria. The following section discusses the correction for this effect. Currently, the only correction applied pertains to the real electron contamination

³Such a formulation is correct in the limit where the double misidentification ($FailFail$) factorizes between tracks, verified from Ref. [49, 50].

originating from the signal, as this is expected to be the dominant contribution. Nonetheless, future corrections should also consider the real electrons arising from background processes.

Correction for real electrons contamination

First of all, an initial estimation of the signal yield ($\mathcal{N}_{sig}$) in the $B^0 \rightarrow K\pi e^+ e^-$ dataset is computed. This is done by performing a fit to the $m(K\pi e^+ e^-)$ distribution without misidentified background component and with the full offline selection applied (**Chap. 3** and **Chap. 4**). Afterwards, the efficiencies for the scenarios b) (OS), c) (SS), and d) (ee) are computed from the $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation as follows:

$$\epsilon_{OS} = \frac{N(e_{OS}^{Fail} \& e_{SS}^{Pass})}{N(e_{OS}^{Pass} \& e_{SS}^{Pass})}, \quad \epsilon_{SS} = \frac{N(e_{SS}^{Fail} \& e_{OS}^{Pass})}{N(e_{OS}^{Pass} \& e_{SS}^{Pass})}, \quad \epsilon_{ee} = \frac{N(e_{OS}^{Fail} \& e_{SS}^{Fail})}{N(e_{OS}^{Pass} \& e_{SS}^{Pass})} \quad (3.3)$$

Additionally, the weighted average per run of each of the efficiencies above is computed using luminosity factors (L_y) as weights per year y . The luminosity factors are obtained by the product of the integrated luminosity collected for each year (see **Tab. 3.1**) and the $b\bar{b}$ cross-section normalised to that at $\sqrt{s} = 8 \text{ TeV}^4$. This procedure gives the following *effective* efficiencies per scenario:

$$\epsilon_{OS(SS, e^+ e^-)}^{eff} = \frac{\sum_y L_y \epsilon_{OS(SS, ee), y}}{\sum_y L_y} \quad (3.4)$$

Finally, the expected number of signal events in each scenario (OS, SS, $e^+ e^-$) is extracted according to the following equation:

$$\mathcal{N}_{sig, OS(SS, ee)} = \mathcal{N}_{sig} \epsilon_{OS(SS, ee)}^{eff} \quad (3.5)$$

It is now possible to integrate these quantities in (3.1). This is done as follows:

$$\begin{aligned} \bar{\mathcal{N}} = & \sum_i^{N_{OS}} w_{fake, i}^{e_{OS}} + \sum_j^{N_{SS}} w_{fake, j}^{e_{SS}} - \sum_k^{N_{ee}} w_{fake, k}^{ee} + \\ & - \left(\frac{\sum_l^{N'_{OS}} w_{fake, l}^{e_{OS}}}{\sum_l^{N'_{OS}} w_l} \mathcal{N}_{sig, OS} + \frac{\sum_m^{N'_{SS}} w_{fake, m}^{e_{SS}}}{\sum_m^{N'_{SS}} w_m} \mathcal{N}_{sig, SS} - \frac{\sum_n^{N'_{ee}} w_{fake, n}^{ee}}{\sum_n^{N'_{ee}} w_n} \mathcal{N}_{sig, ee} \right), \end{aligned} \quad (3.6)$$

where the indexes i, j and k , run over the number of B^0 candidates of the dataset per scenario (N_{OS}, N_{SS}, N_{ee}). The indices l, m , and n , run over the $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation entries per scenario ($N'_{OS}, N'_{SS}, N'_{ee}$). The weights $w_{l, m, n}$ are the simulation correction weights. The ratio between the sums of $w_{fake, l, m, n}$ and $w_{l, m, n}$ provides the rescaling factor needed to match the data. This ratio can be interpreted as the estimated efficiency for real electrons to fail the PID criteria. By multiplying this efficiency by the expected number of signal events in the data, one obtains the expected number of real electrons failing the PID selection. The effect of $w_{fake}(\vec{X})$ applied to the MC simulation is shown in the second column of **Fig. 3.5**, where the histogram in solid blue shows the simulation events before applying $w_{fake}(\vec{X})$, but with MC and luminosity corrections applied, while the orange distribution shows the simulation events after applying the $w_{fake}(\vec{X})$ weights in each of the categories (OS, SS, ee). The last plot in the second column shows the total

⁴The values for the normalised cross-sections have been provided by the Refs. [49, 50], via personal communication.

distribution when combining the $w_{fake}(\vec{X})$ of each category. Finally, the last column shows the misidentified background distribution after the real-electron contamination is subtracted, for each of the scenarios (OS, SS, ee). The plot in the last column and last row is the final expected misidentified electron contribution to the data, combining the three scenarios and subtracting the real electrons.

The plots in **Fig. 3.5** show the shape of this background component in $m(K\pi e^+e^-)$ for the dataset collected during Run 2. The corresponding plots for Run 1 $m(K\pi e^+e^-)$ are shown in **Fig. A.4**. Additionally, the background shapes for different angular variables are presented in **App. A**: $\cos\theta_\ell$ for Run 1 (**Fig. A.5**) and Run 2 (**Fig. A.1**), $\cos\theta_K$ for Run 1 (**Fig. A.6**) and Run 2 (**Fig. A.2**), and ϕ for Run 1 (**Fig. A.7**) and Run 2 (**Fig. A.3**). By examining the angular distributions in the appendix, one can observe that this background manifests prominently at high values of $\cos\theta_\ell$, similarly to the DSL background. A more detailed discussion of this shape is provided in **Chap. 5**.

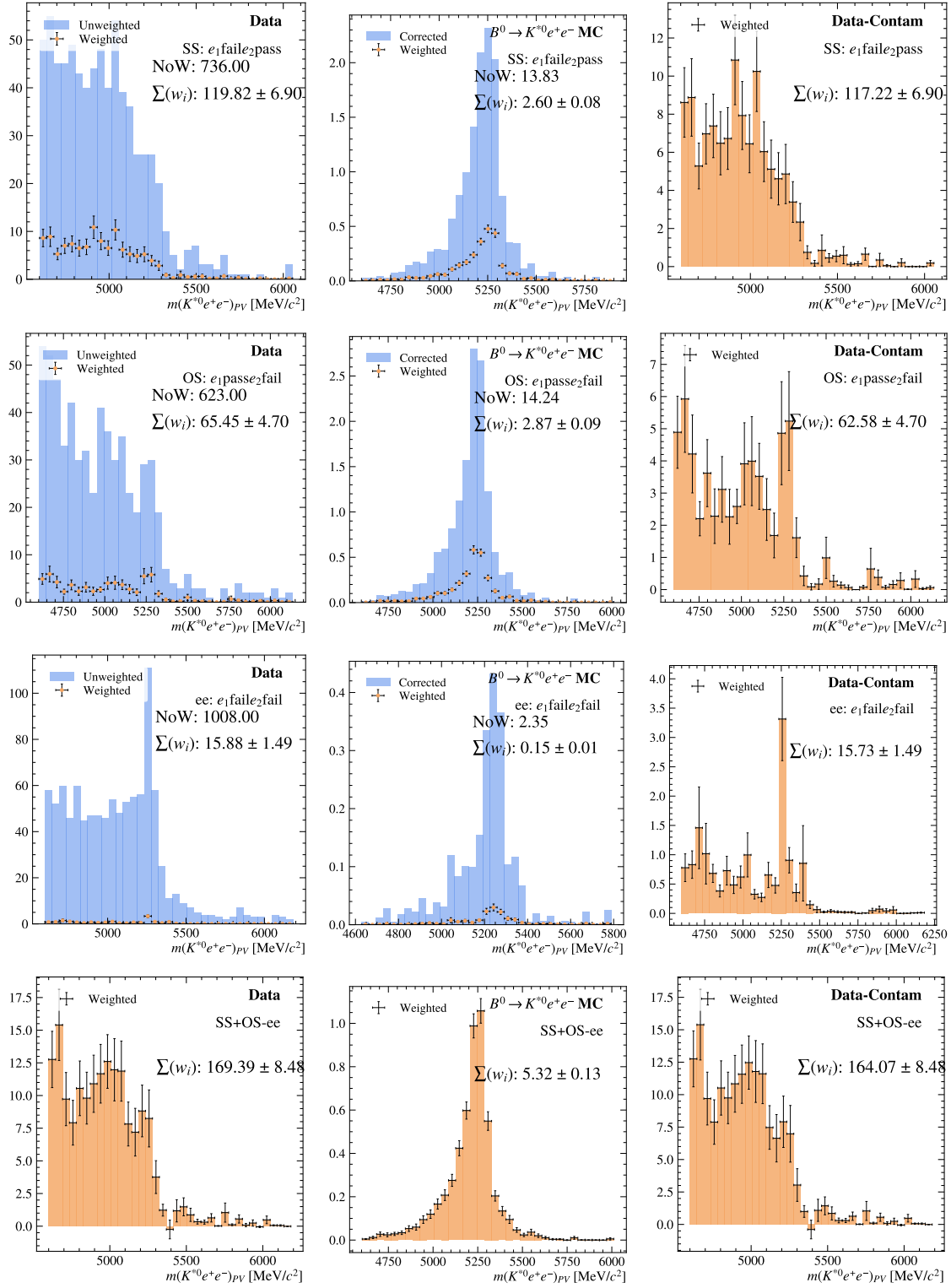


Figure 3.5: Distributions of $B^0 \rightarrow K\pi e^+ e^-$ candidates obtained merging trigger categories for the Opposite Sign(OS), Same Sign(SS), Fail Fail(ee) from the first to the third row, respectively. The first and the second columns show real data and $B^0 \rightarrow K^0 e^+ e^-$ MC simulation. The histograms in blue show the data (or MC simulation) without applying $w_{fake}(\vec{X})$, and the overlaid orange points show the same data after applying $w_{fake}(\vec{X})$. The last row shows the distributions obtained by combining the $w_{fake}(\vec{X})$ contributions shown above (SS+OS- ee). The last column is the overall combined $w_{fake}(\vec{X})$ weighted data–contamination in OS, SS, ee in the first three rows. The plot in the row and column shows the final dataset representing the expected misidentified electrons background content. “NoW” is the number of events belonging to the solid blue distribution before applying the scale function.

4 | Selection optimization

The offline selection can be further optimised to increase the signal-to-background ratio. This can be achieved using Machine Learning (ML) classifiers, particularly gradient-boosted decision trees (BDTs). This chapter introduces BDTs and presents two models trained specifically for this analysis.

4.1 Gradient Boosted Decision Trees (BDT)

The Boosted Decision Tree (BDT) is a supervised machine-learning technique used to classify events by assigning a score to each candidate, where a lower score indicates a higher probability of being background. Boosting is a method that builds a powerful classifier by combining an ensemble of weak learners [110]. In the case of a BDT, the weak learner is a *decision tree*, which recursively partitions the input feature space [111].

Training Phase For the purpose of training the model, the set of input features x is provided for both signal-like and background-like candidates. These candidates are labelled as signal ($y = 1$) and background ($y = 0$). The model is built iteratively by adding decision trees one at a time. The prediction score after N trees is then given by:

$$\hat{y}^{(N)} = \sum_{j=1}^N f_j(x) \quad (4.1)$$

where $f_j(x)$ is the prediction of the j -th tree. Each new tree is trained to minimise a loss function $L(\hat{y}, y)$, which measures the difference between the predicted ($\hat{y}$) and true values (y). This optimisation step, known as *gradient boosting*, adjusts each tree to correct the errors of the previous ones. For more details, see Ref. [112].

Evaluation Phase Once trained, the BDT can be applied to new data points. Given an unseen input x_{new} , the final prediction is obtained by summing the outputs of all trained trees:

$$\hat{y}_i^{(N)} = \sum_{j=1}^N f_j(x_i^{\text{new}}). \quad (4.2)$$

Unlike training, this step does not involve loss function minimisation or tree optimisation - it simply evaluates the trained trees and produces a classification score. This evaluation is performed on each data point i of the new dataset.

Two BDT models are trained in the analysis using XGBoost [113]. The score is distributed between 0 and 1, where lower values correspond to background-like candidates and higher values to signal-like candidates. The first BDT (BDT1) is used to effectively reduce the overall

background. The second BDT (BDT2) specifically targets residual backgrounds that peak at high values of $\cos \theta_\ell$.

4.2 BDT1

4.2.1 Input features

The input features of this first BDT are listed in **Tab. 4.1**. Besides standard geometric and kinematic variables (χ_{IP}^2 , χ_{vtx}^2 , p_T , η), two isolation variables have been introduced, as they proved to be able to improve the discrimination power of the classifier. The isolation variables measure the activity in terms of other particles in the proximity of the selected track or vertex. With more activity near the signal candidate, it becomes less likely that the candidate is a true $B^0 \rightarrow K^{*0} e^+ e^-$.

	Particle	Variable
1	$e^+ e^-$	Min χ_{IP}^2
2	$e^+ e^-$	Max p_T
3	J/ψ	χ_{vtx}^2
4	K^{*0}	χ_{vtx}^2
5	K^{*0}	χ_{IP}^2
6	K^{*0}	p_T
7	B^0	η
8	B^0	χ_{IP}^2
9	B^0	TOT χ^2
10	B^0	$ISO_{vtx}^{oneTrack} \Delta\chi^2$
11	B^0	$MASS(ISO_{vtx}^{oneTrack} \Delta\chi^2)$

Table 4.1: List of features that serve as input to the BDT1.

The isolation variables are generally effective in rejecting combinatorial-like backgrounds. The tracks building the K^{*0} or the B^0 vertex $\mathcal{V}$ are combined with a single other track from the event, and fitted together into a new vertex $\mathcal{V}^*$. The isolation variables used are defined as follows [114]:

- $ISO_{vtx}^{oneTrack} \Delta\chi^2$ the smallest difference in χ^2 between the $\mathcal{V}$ and the $\mathcal{V}^*$ vertex fits among all the other tracks in the event.
- $MASS(ISO_{vtx}^{oneTrack} \Delta\chi^2)$ the invariant mass of the tracks used to build the $\mathcal{V}^*$ vertex.

A few other variables have not been mentioned before - all the other variables have been defined in **Sec. 2.7**:

- $e^+ e^-$ Min χ_{IP}^2 : the smallest value of the χ_{IP}^2 between the two electrons.
- $e^+ e^-$ Max p_T : the largest value of the p_T between the two electrons.
- TOT χ^2 : the sum of the of B^0 χ_{IP}^2 , B^0 χ_{vtx}^2 , K^{*0} χ_{vtx}^2 , and J/ψ χ_{vtx}^2 .

The K^{*0} and B^0 variables target mainly combinatorial and partially reconstructed backgrounds, as both backgrounds are less likely to result in signal-like values for K^{*0} quantities and kinematic observables of a real $B^0 \rightarrow K^{*0} e^+ e^-$ decay. The J/ψ and $e^+ e^-$ variables target both combinatorial and possibly DSL backgrounds. They are expected to contribute less in rejecting partially reconstructed backgrounds.

4.2.2 Training samples and BDT score

The Run 2 $B^0 \rightarrow K\pi e^+ e^-$ dataset has been used to train the BDT labelled as background with the following selection applied:

- Full offline selection presented in **Chap. 3**, except for the q^2 and B^0 mass selection.
- $m(K\pi ee) > 5600 \text{ MeV}/c^2$: this cut selects a combinatorial enriched data sample.
- $(4600 < m(K\pi ee) < 4900) \text{ MeV}/c^2$: this cut selects the left sideband of the B^0 mass peak, where all the residual backgrounds discussed in **Sec. 3.4** are located, as shown by the illustration in **Fig. 3.4**. Events below this threshold are not selected because the $B^0 \rightarrow K^{*0} J/\psi$ leakage becomes abundant, and this background is expected to be very similar to the signal. As a consequence, adding those events to the training sample would harm the performance of the BDT. This selection is a background-enriched data sample.
- $(1 < q_{PV\&B}^2 < 7) \text{ GeV}^2/c^4$, releasing this cut and the q^2 veto, allows to have more events in the training sample for the background, as it is statistically limited.

The $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation has been used to train the BDT labelled as signal. The offline selection presented in **Chap. 3** is applied to the $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation, except for the q^2 cut which has been aligned to the background training sample. Additionally, the selection around the B^0 mass has been tightened to $4900 \text{ MeV}/c^2 < m(K^{*0} e^+ e^-) < 6200 \text{ MeV}/c^2$. An independent test sample has been set aside to evaluate its performance. The BDT classifier has been trained on 75% of the aforementioned dataset, and its performance was subsequently evaluated on the remaining 25%. This helps in assessing the true generalisation of the model. Additionally, k-fold cross-validation has been performed. If the average performance remains consistent, this generally ensures that no significant bias has been introduced by using this portion of the data for training the BDT.

The distribution of the features for the signal and background training samples and their correlation matrices are shown in **Fig. B.1** and **Fig. B.2**, respectively, in **App. B**.

The performance of the BDT together with the importance plot of the features are shown in **Fig. 4.1**. The left plot shows the classifier output for both signal (red) and background (blue), including training and test samples. The performance of the classifier is qualitatively indicated by its ability to separate signal from background and by the agreement between the test and training sample results. The Area Under the Curve (AUC) is also reported in the figure, it is a metric that quantifies the performance of a classifier by measuring the area under the Receiver Operating Characteristic (ROC) curve [115]. The AUC values range from 0 to 1, where 0.5 indicates random classification, and 1 represents a perfect classifier. Therefore, an AUC equal to ~ 0.97 signifies a good discrimination between signal and background. The right plot illustrates the order of importance of the input variables to distinguish signal from background.

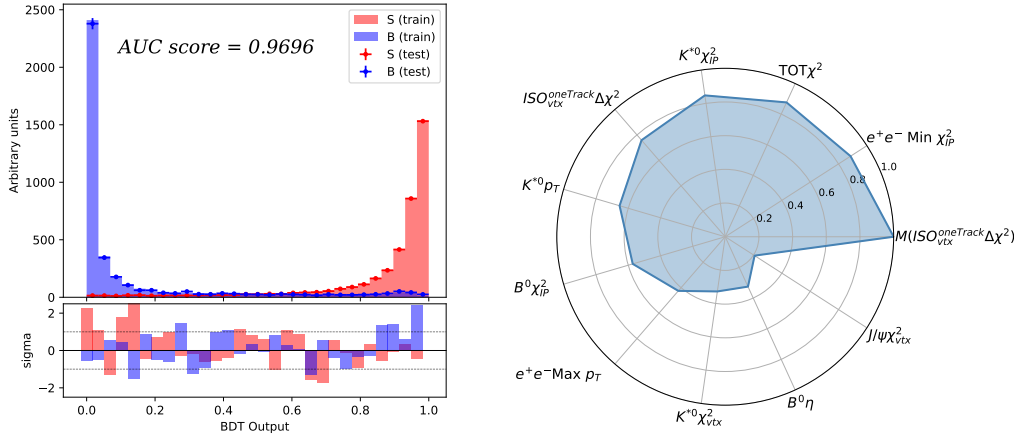


Figure 4.1: The left plot shows the output of the BDT1 classifier for signal (red) and background (blue), train and test samples. The right plot illustrates the importance of each variable in discriminating signal from background events. They are normalised to the variable with the highest score.

A 10-fold cross-validation is performed which provided an average AUC and accuracy of 0.97 and 91%, respectively. The accuracy is a metric that measures the proportion of correctly classified instances in a dataset. It is given by the formula:

$$\text{Accuracy} = \frac{TP + TN}{P + N}, \quad (4.3)$$

where P and N indicate all the signal (positive) and background (negative) events, while TP (True Positive) and TN (True Negative) indicate the events correctly identified as signal or background by the classifier.

4.2.3 Data to MC agreement of the features

Since the signal sample represents the expected signal in the data, it is crucial to verify the agreement between the data and MC for the relevant features. This ensures that any selection applied to the simulation affects the data in a consistent manner. The abundant charmonium resonance $B^0 \rightarrow K^{*0}J/\psi (\rightarrow e^+e^-)$ is used to check the agreement, in particular the background subtracted $B^0 \rightarrow K^{*0}J/\psi$ data sample has been used for this purpose¹. The $B^0 \rightarrow K^{*0}J/\psi$ MC simulation variables are normalised to those of the background-subtracted $B^0 \rightarrow K^{*0}J/\psi$ data sample, and the pulls among the two distributions are computed, as shown in **Fig. 4.2**. Looking at the pull distributions, two variables catch the attention as the disagreement is beyond 5σ in several bins: the total B^0 $TOT\chi^2$ and the B^0 $IP\chi^2$. Discarding these two variables as training features has been considered. However, they have been proven valuable as part of the input to the BDT: removing them would lead to a less effective BDT selection. Therefore, they are part

¹The $B^0 \rightarrow K\pi e e$ data sample is selected in the q^2 region of the $B^0 \rightarrow K^{*0}J/\psi$ resonance ($7.0 < q_{PV\&B}^2 < 11.0$) GeV^2/c^4 , see **Fig. 1.4**. The $B^0 \rightarrow K^{*0}J/\psi$ dataset is also affected by partially reconstructed background, with the difference that a single cut on the B^0 mass reconstructed using the DTF with PV and J/ψ mass constraints, removes all this background, $m(K\pi e^+e^-)_{PV\&J/\psi} > 5150 \text{ MeV}/c^2$. Additionally, besides the full offline selection, the sample has been reduced to a narrow region around the B^0 pole mass to reduce as much as possible the amount of combinatorial background, $(5150 < m(K\pi e^+e^-) < 5350) \text{ MeV}/c^2$. With this selection, the residual combinatorial background is subtracted with a fit to the data.

of the final features, and instead a systematic error is studied and assigned to account for this difference, see **Chap. 6**.

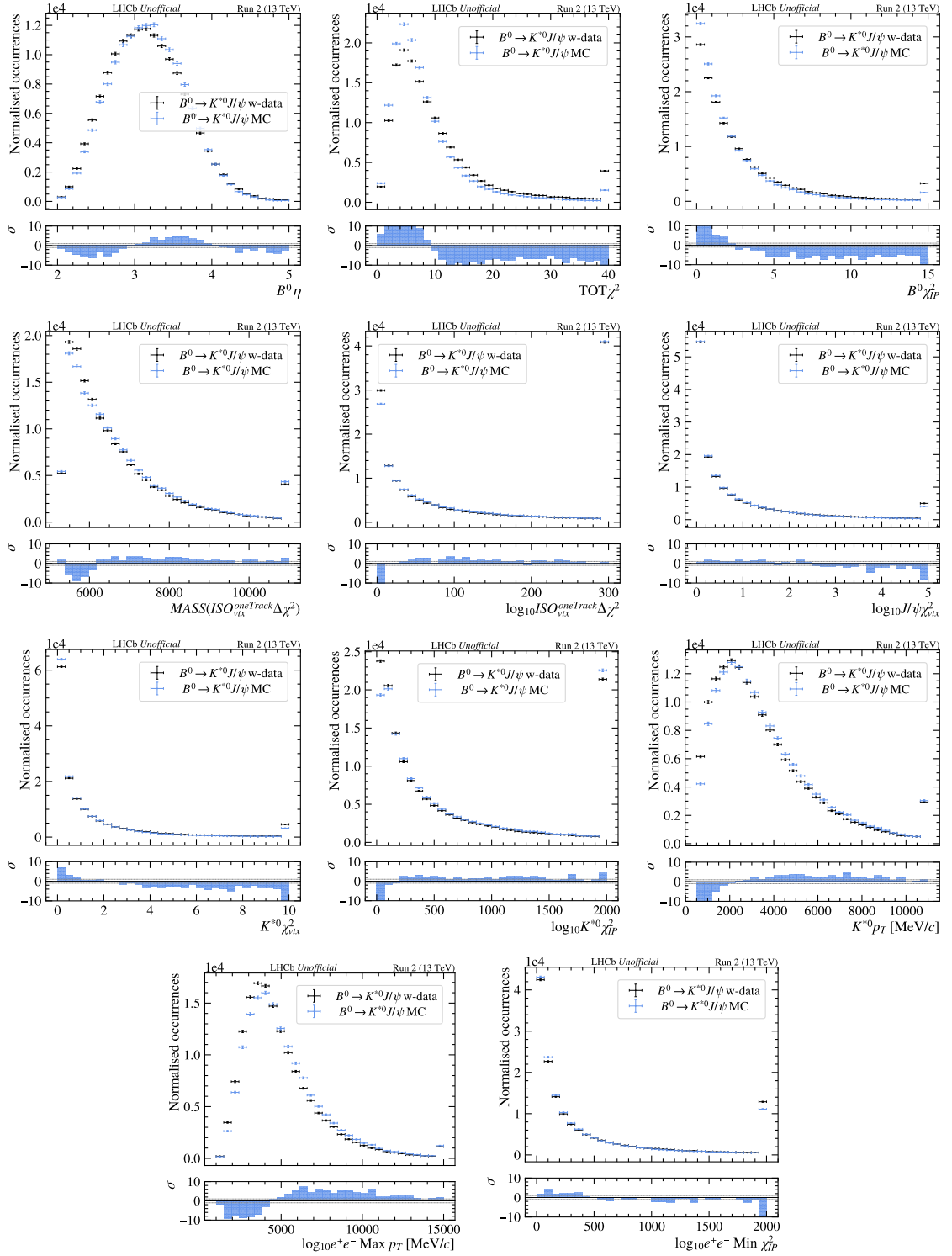


Figure 4.2: Comparison between background subtracted $B^0 \rightarrow K^{*0} J/\psi$ data sample and $B^0 \rightarrow K^{*0} J/\psi$ MC simulation of the features listed in **Tab. 4.1**. The $B^0 \rightarrow K^{*0} J/\psi$ corrected MC simulation has been normalised to the background-subtracted $B^0 \rightarrow K^{*0} J/\psi$ data sample. Each plot shows an under- and overflow bin.

4.2.4 Figure of Merit (FoM) for signal-to-background optimisation

In order to find the optimal cut on the BDT output score, the following FoM is used:

$$\text{FoM} = \frac{\epsilon_{sig}}{\sqrt{D}} \quad (4.4)$$

where ϵ_{sig} is the efficiency of the cut on the signal MC sample, while D is the yield of data candidates in the signal region around the nominal B^0 mass in the $m(K\pi e e)$ distribution of the $B^0 \rightarrow K\pi e^+ e^-$ dataset. This quantity relates the signal efficiency to the statistical uncertainty on data, which contains the information of the remaining signal and background. Ideally, the optimal cut minimises the statistical error on the signal while maximising the signal efficiency. Therefore, the cut that provides the larger value of this FoM is optimal. This FoM is computed by applying to both data and MC samples the same selection as in the signal training samples. Additionally, the cut around the B^0 pole mass is tightened to $[5150, 5350] \text{ MeV}/c^2$. This FoM has been evaluated as a function of the cut on the BDT output, on the central q^2 bin. The search for the optimal cut is shown in **Fig. 4.3** where the maximum of the curve is highlighted with a dotted line and the optimal BDT score is written. $BDT\ output > 0.96$ is the cut that optimises the amount of signal and background, providing the maximum background rejection while keeping the highest signal efficiency possible.

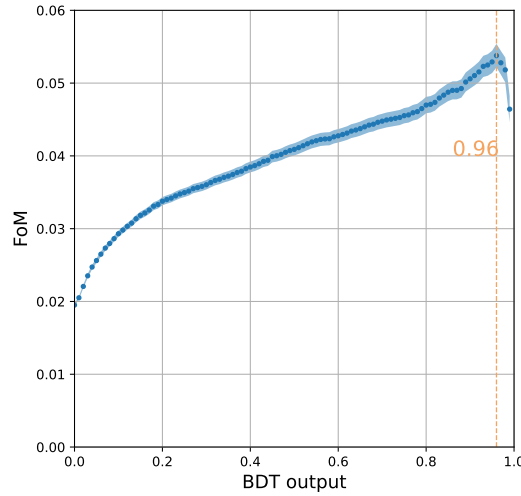


Figure 4.3: FoM computed using $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation and $B^0 \rightarrow K\pi e^+ e^-$ data sample by applying the cut $BDT\ output > X$, with X increasing of 0.01 at each step. The blue band indicates the statistical error associated with each FoM value. The orange dotted line highlights the maximum of the curve, corresponding the optimal BDT score.

4.2.5 Signal efficiency and backgrounds rejection

The effect of the BDT cut on the angles and mass distributions can be tested on the signal $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation. The plots in **Fig. 4.4** show the effect of the BDT cut on the sample with the full offline selection applied, specifically in the full central q^2 region. The blue distributions and ratios show the effect of the cut $BDT\ output > 0.96$, while the orange distributions show the effect of the inverted cut $BDT\ output < 0.96$, namely, what the former cut is removing. The ratio

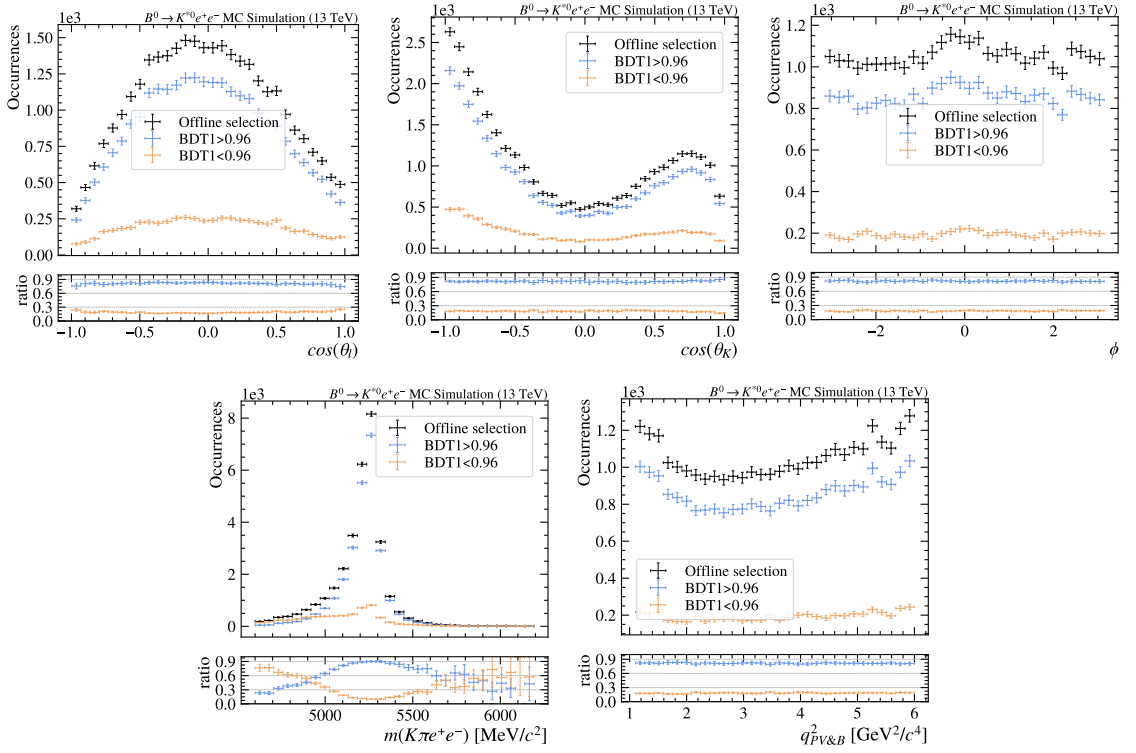


Figure 4.4: The figure shows the distribution of the B^0 mass, $q^2_{PV\&B}$, and three angular variables for the signal $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation at 13 TeV. The black distribution represents the baseline used to compute the ratio and the total signal efficiency, with the full offline selection applied in the central q^2 region. The distributions show the effect of the cut on the BDT score: > 0.96 . The orange distributions show the effect of the inverted cut: < 0.96 .

distributions show that the effect remains mostly flat for the angular variables and q^2 , ensuring that no additional dependence on the MC simulation is introduced. However, the selection modifies the mass distribution by preferentially removing events at lower masses. Since the mass distribution will be used as a variable in the final fit, accepting this effect assumes that the same modification occurs in the data. To account for any potential bias introduced by the BDT selection, a systematic uncertainty is assigned (**Sec. 6.2.6**). The total signal efficiency of the BDT cut is $(81.6 \pm 0.2)\%$, computed relative to the offline-selected sample.

The signal efficiency and background rejection percentages are listed in **Tab. 4.2** for both the low and central q^2 bins. These values are computed using either MC simulation at 13 TeV or Run 2 datasets, as specified in the table. The BDT effectively suppresses all residual backgrounds. However, the estimated amount of background peaking at high $\cos \theta_\ell$ remains significant. The $\cos \theta_\ell$ distributions of the peaking backgrounds are shown in **App. F** in **Fig. F.1** for the flavour-violating and misidentified electron samples. To further reduce these backgrounds while preserving as much signal as possible, an additional BDT has been studied. This would be particularly beneficial in minimizing systematic uncertainties related to background modeling.

Sample	Type	Background	Background rejection %		Years
			low q^2	central q^2	
$B^0 \rightarrow D(\rightarrow K^{*0} e \nu_e) e \nu_e$	MC	DSL	76.3 ± 2.6	77.3 ± 0.6	2016
$B^0 \rightarrow K^{*0} e \mu$	Data	DSL + Comb	97.33 ± 0.15	96.3 ± 0.1	Run 2
$B \rightarrow K \pi \pi e e$	MC	PR	61.7 ± 0.9	63.8 ± 0.4	@13 TeV
misID	Data driven	misID	94.4 ± 0.93	93.3 ± 0.5	
			Signal efficiency %		
			low q^2	central q^2	
$B^0 \rightarrow K^{*0} e e$	MC	-	83.3 ± 0.3	81.6 ± 0.2	@13 TeV

Table 4.2: Signal efficiency and background rejections of the optimal cut on the BDT output computed with respect to samples with the full offline selection applied in the q^2 region reported in the table.

4.3 BDT2 targeting backgrounds at high $\cos \theta_\ell$

Even if a substantial part of the background has been removed, an explicit effort has been made to target specifically those backgrounds which predominantly populate the high $\cos \theta_\ell$ region.

4.3.1 Features

The input features of the BDT are listed in **Tab. 4.3**. The first five variables are in common with the previous BDT, while additional electron variables are added. Kinematic and geometric variables target the DSL background, while PID variables target the misidentified background.

	Particle	Variable
1	$e^+ e^-$	Min χ_{IP}^2
2	$e^+ e^-$	Max p_T
3	J/ψ	χ_{vtx}^2
4	K^{*0}	χ_{IP}^2
5	B^0	$ISO_{vtx}^{oneTrack} \Delta \chi^2$
6	$e^+ e^-$	Max χ_{IP}^2
7	$e^+ e^-$	Min p_T
8	e_1	PIDe
9	e_2	PIDe
10	e_2	P

Table 4.3: List of features that serve as input to the second BDT (BDT2).

4.3.2 Data to MC agreement of the features

The same type of validation strategy described in **Sec. 4.2.3** is performed on the additional variables used by this BDT. The comparison between data and MC samples of the five additional variables is shown in **Fig. 4.5**.

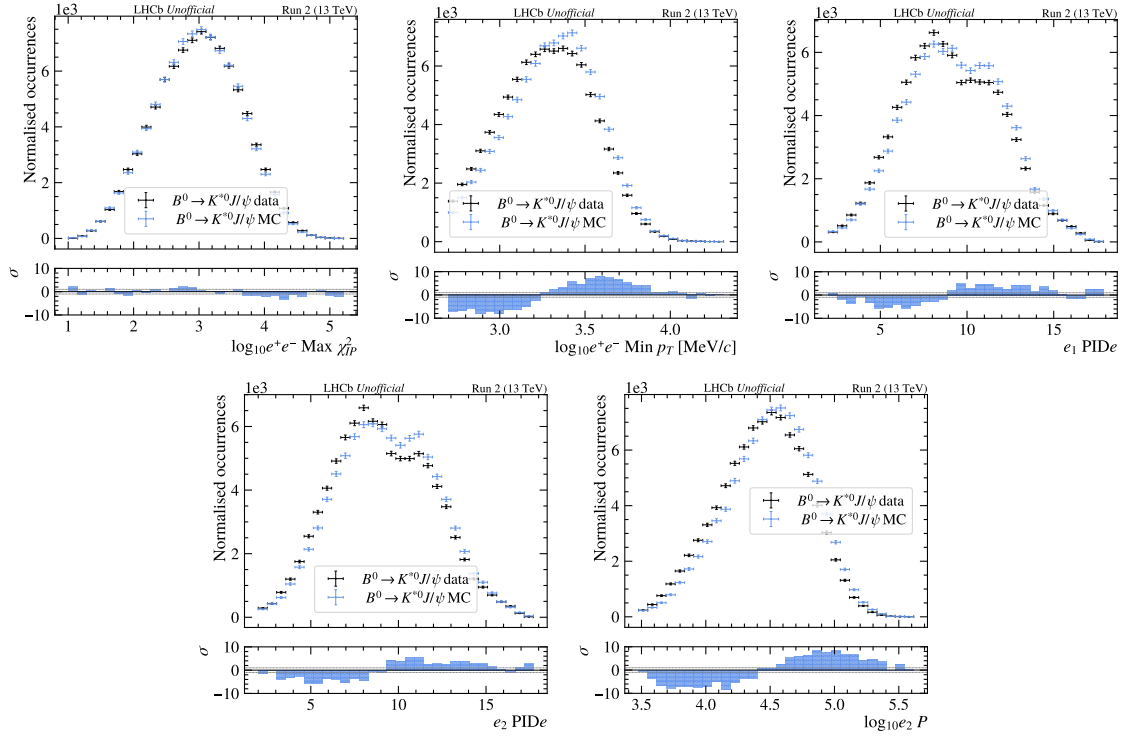


Figure 4.5: Comparison between background subtracted $B^0 \rightarrow K^{*0} J/\psi$ data sample and $B^0 \rightarrow K^{*0} J/\psi$ MC simulation of the features listed in **Tab. 4.3**. The $B^0 \rightarrow K^{*0} J/\psi$ corrected MC simulation has been normalised to the background subtracted $B^0 \rightarrow K^{*0} J/\psi$ data sample.

4.3.3 Training samples and BDT score

The full Run 2 $B^0 \rightarrow K\pi e^+ e^-$ dataset has been used to train the BDT labelled as background with the following selection applied:

- Full offline selection presented in **Chap. 3**, except for the q^2 and B^0 mass selection.
- $m(K\pi ee) < 4800 \text{ MeV}/c^2$: this cut selects the left sideband of the B^0 mass peak, where all the residual backgrounds discussed in **Sec. 3.4** are located. In this case, the full left sideband has been used to increase the number of events in the background training sample.
- $(0.1 < q_{PV\&B}^2 < 7) \text{ GeV}^2/c^4$: releasing the cut and the $q_{unconst}^2$ veto allows to have more events in the training sample for the background, as it is statistically limited. For the same reason, the full low and central q^2 regions are used.
- $BDT1 > 0.4$: A loose cut on the first BDT score is applied to remove part of the background. This cut is less stringent than the optimal one to preserve a larger training sample.
- $\cos \theta_\ell > 0.6$: this cut specifically targets those backgrounds that peak at $\cos \theta_\ell \sim 1$.

The $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation at 13 TeV has been used to train the BDT, labelled as signal. The same selection applied to the background training sample is used for the signal sample, except for the B^0 mass requirement, which in this case is defined as $(4900 < m(K^{*0} e^+ e^-) < 6200) \text{ MeV}/c^2$, and the cut on $\cos \theta_\ell$ is not applied. An independent test sample has been set aside to evaluate its performance. The BDT classifier has been trained on 75% of the aforementioned dataset, and its

performance was subsequently evaluated on the remaining 25%. The performance of the BDT together with the importance plot of the features is shown in **Fig. 4.6**. The AUC is estimated to be about 0.91, which proves the BDT to be effective in distinguishing signal from background. Additionally, a 10-fold cross-validation has been performed which provided an average AUC of 0.91 and an accuracy of 83%.

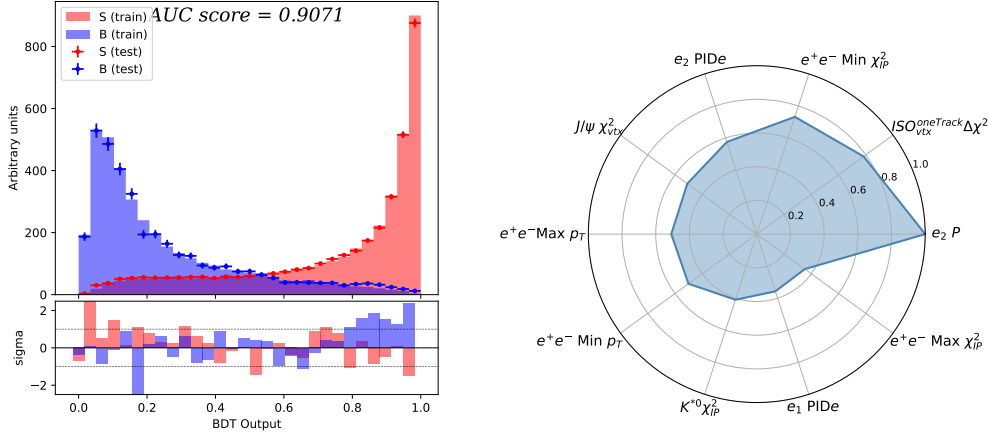


Figure 4.6: The left plot shows the output of the BDT2 classifier for signal (red) and background (blue), train and test samples. The right plot illustrates the importance of each variable in discriminating signal from background events. They are normalised to the variable with the highest score.

The distribution of the features for the signal and background training samples and their correlation matrices are shown in **Fig. B.3** and **Fig. B.4**, respectively, in **App. B**.

4.4 2D Optimisation of the cut on BDTs output

In order to find the optimal cut point, a 2-dimensional optimisation has been performed using both the BDT outputs presented in this chapter. The figure of merit that has been used is the following:

$$FoM = \frac{\epsilon_{sig}}{\sqrt{S+B}} \quad (4.5)$$

where ϵ_{sig} is the signal efficiency of a selection configuration and $\sqrt{S+B}$ is taken to be the error on the signal yield returned by a mass fit to the data. The 2D scan of the BDT scores is shown in **Fig. 4.7**, for each cell both the value of the FoM and the signal efficiency have been reported.

The largest FoM value corresponds to the cuts $BDT1 > 0.96$ and $BDT2 > 0.2$. These cuts are going to be applied in the rest of this dissertation, together with the offline selection, unless specified otherwise.

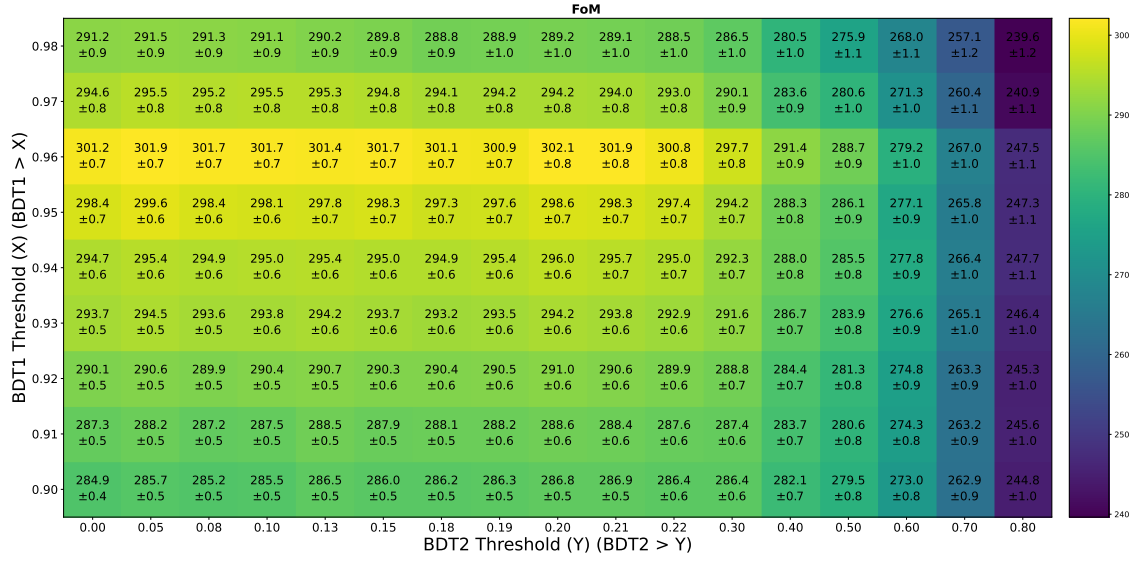


Figure 4.7: FoM computed using $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation and $B^0 \rightarrow K \pi e^+ e^-$ data sample by applying the cuts $BDT1 > X$, and $BDT2 > Y$ scanning the two-dimensional plane of the two BDT scores.

4.5 Signal efficiency and backgrounds rejection

The effect of these two cuts on the signal $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation at 13 TeV is shown in **Fig. 4.8** on the angular, $m(K^{*0} e^+ e^-)$ and $q_{PV\&B}^2$ variables. The blue distributions illustrate the effect of the cuts, while the orange distributions show the impact of the inverted cuts, highlighting the events removed by the selection. The ratio relative to the sample without the BDT cuts applied indicates that the selection on the new BDT slightly affects the signal at values of $\cos \theta_\ell$ close to one. This effect amounts to the fact that among the input features there were some correlating with $\cos \theta_\ell$ - such as the minimum and maximum electron p_T . This selection is applied to BDT2 output under the assumption that the same distortion also affects the data. This assumption is accounted for as a systematic uncertainty in the final measurement (**Sec. 6.2.6**). Including this systematic uncertainty remains preferable to applying a direct cut on $\cos \theta_\ell$. This choice is further discussed in **Chap. 7** and justified in **App. F**.

The signal efficiency and background rejection percentages for the two combined cuts are listed in Table **Tab. 4.4**, computed relative to the samples with full offline selection in both the low and central q^2 bins. These values are obtained using MC simulation at 13 TeV or Run 2 datasets, unless otherwise specified in the table.

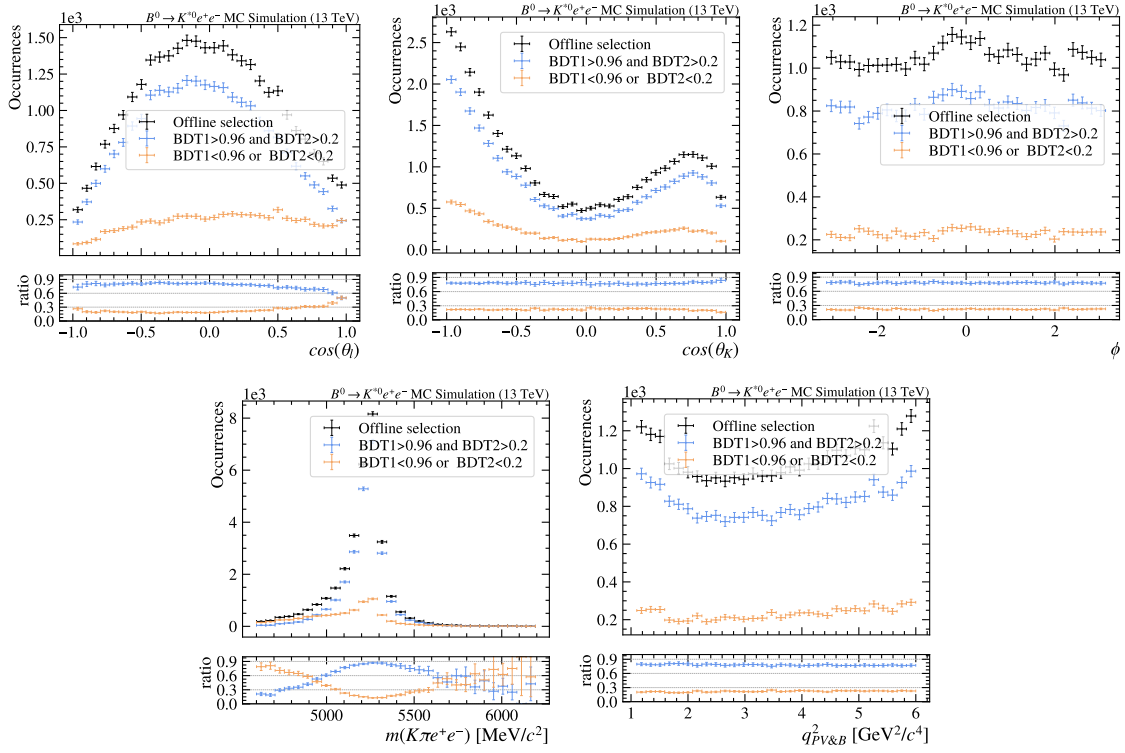


Figure 4.8: The figure shows the distribution of the B^0 mass, $q^2_{PV\&B}$, and three angular variables for the signal $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation at 13 TeV. The black distribution represents the baseline used to compute the ratio and the total signal efficiency, with the full offline selection applied in the central q^2 region. The blue distributions show the effect of the cuts $BDT1 > 0.96$ and $BDT2 > 0.2$. The orange distributions show the effect of the inverted cut: $BDT1 < 0.96$ and $BDT2 > 0.2$.

Sample	Type	Background	Background rejection %		Years
			low q^2	central q^2	
$B^0 \rightarrow D(\rightarrow K^{*0} e \nu_e) e \nu_e$	MC	DSL	96 ± 1	83.9 ± 0.4	2016
$B^0 \rightarrow K^{*0} e \mu$	Data	DSL + Comb	98.2 ± 0.1	97.09 ± 0.07	Run 2
$B \rightarrow K \pi \pi e e$	MC	PR	64.1 ± 0.9	65.8 ± 0.4	Run 2
misID	Data driven	misID	95.8 ± 0.8	95.5 ± 0.4	
Signal efficiency %					
			low q^2	central q^2	
$B^0 \rightarrow K^{*0} e e$	MC	-	79.0 ± 0.3	78.0 ± 0.2	Run 2

Table 4.4: Signal efficiency and background rejections for the two combined cuts $BDT1 > 0.96$ and $BDT2 > 0.2$.

4.6 Considerations

After applying this optimised selection, the residual backgrounds discussed in **Sec. 3.4** are still present. Comparing **Tab. 4.2** and **Tab. 4.4**, one can see that BDT2 improves significantly the DSL rejection, while reducing also other sources of backgrounds with a moderate penalisation of about 4% in signal efficiency.

Both the BDTs are trained using Run 2 samples. The same models are used to evaluate their

outputs also on the Run 1 sample. While this approach may be suboptimal, the low number of candidates available in the Run 1 dataset is insufficient to train a separate classifier.

This thesis focuses on the central q^2 region. However, the selection and optimisation of the two BDTs have also been proven effective in the low q^2 region. Therefore, the scores can be applied in low q^2 as well.

As previously mentioned, the systematic uncertainties associated with the selection on the BDT outputs are estimated in **Sec. 6.2.6**. These account for effects introduced by the selection, such as distortions in the distributions of variables to be fitted - specifically, the signal mass distribution from the first BDT and the $\cos \theta_\ell$ distribution from the second - as well as differences between data and MC simulation.

The final selection is defined by the criteria detailed in **Chap. 3**, along with the optimal cuts on the two BDT outputs determined in this chapter: $BDT1 > 0.96$ and $BDT2 > 0.2$.

5 | Analysis strategy

*This chapter outlines the analysis strategy adopted to extract the angular parameters describing the $B^0 \rightarrow K^{*0} e^+ e^-$ decays using the dataset collected during Run 1 and Run 2.*

5.1 Angular acceptance

The observation of any physics process is inherently shaped by the “eyes” through which it is viewed. The LHCb detector, much like human vision, captures only a partial perspective of reality. Due to its geometrical design, the detector covers only a portion of the solid angle, leaving some regions unmonitored. Additionally, even within its active regions, the detector does not always efficiently register all particles that pass through. These limitations - both geometric and instrumental - are most naturally represented by an efficiency: the ratio of observed processes to the total number of occurrences. In this study, the efficiency that accounts for the detector effects must be characterised as a function of the three relevant angles, defining an efficiency function referred to as acceptance. Another important detector-related effect concerns the momenta of the final-state particles, from which the angles are computed. Due to the detector's finite precision, these momenta are estimated with a non-perfect resolution. This effect must also be taken into account when evaluating the overall angular acceptance of the detector.

This section outlines how the acceptance is parameterised and incorporated in the analysis, with a specific paragraph dedicated to the resolution effect. Extensive studies have been conducted on this methodology using the abundant charmonium resonance $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ decays [95]. Here, only the fundamental concepts are summarised.

5.1.1 Acceptance and angular PDF

The acceptance function $\varepsilon(\cos \theta_\ell, \cos \theta_K, \phi)$ can be seen as the average probability of recording a candidate as a function of $\cos \theta_\ell$, $\cos \theta_K$ and ϕ , alias Ω , integrating over all other degrees of freedom. This analysis takes into account the detector acceptance including it directly into the probability distribution function (PDF, $\mathcal{P}$) describing the decay. Therefore the PDF of recorded candidates, $\mathcal{P}_{\text{obs}}(\Omega, \vec{S})$, is expressed as:

$$\mathcal{P}_{\text{obs}}(\Omega, \vec{S}) = \frac{\varepsilon(\Omega) \mathcal{P}_{\text{phys}}(\Omega, \vec{S})}{\int d\Omega \varepsilon(\Omega) \mathcal{P}_{\text{phys}}(\Omega, \vec{S})}, \quad (5.1)$$

where $\vec{S}$ are the parameters of interest of this analysis, the coefficients characterising $\mathcal{P}_{\text{phys}}(\Omega, \vec{S})$ which is the normalised physics PDF describing only the physics-related properties of the decay,

given in **Eq. 1.37**. This quantity can be written as a linear combination of angular functions:

$$\mathcal{P}_{\text{phys}}(\Omega, \vec{S}) = \sum_{k=0} C_k f_k(\Omega), \quad (5.2)$$

where f_k are base functions and C_k are angular coefficients, closely related to the parameters of interest $\vec{S}$. The base functions and the corresponding coefficients are listed in **Tab. 5.1**.

k	f_k	C_k in $\vec{S}$ base
0	1	0
1	$\sin^2 \theta_K (1 + \cos^2 \theta_\ell) / 2$	1
2	$(-1 + 5 \cos^2 \theta_K - \cos^2 \theta_\ell - 3 \cos^2 \theta_\ell \cos^2 \theta_K) / 2$	F_L
3	$\sin^2 \theta_K \sin^2 \theta_\ell \cos 2\phi$	S_3
4	$\sin 2\theta_K \sin 2\theta_\ell \cos \phi$	S_4
5	$\sin 2\theta_K \sin \theta_\ell \cos \phi$	S_5
6	$\sin^2 \theta_K \cos \theta_\ell$	$\frac{4}{3} A_{FB}$
7	$\sin 2\theta_K \sin \theta_\ell \sin \phi$	S_7
8	$\sin 2\theta_K \sin 2\theta_\ell \sin \phi$	S_8
9	$\sin^2 \theta_K \sin^2 \theta_\ell \sin 2\phi$	S_9
10	$\sin^2 \theta_\ell$	$F_S / (1 - F_S)$
11	$\sin^2 \theta_\ell \cos \theta_K$	$S_{S1} / (1 - F_S)$
12	$\sin 2\theta_\ell \sin \theta_K \cos \phi$	$S_{S2} / (1 - F_S)$
13	$\sin \theta_\ell \sin \theta_K \cos \phi$	$S_{S3} / (1 - F_S)$
14	$\sin \theta_\ell \sin \theta_K \sin \phi$	$S_{S4} / (1 - F_S)$
15	$\sin 2\theta_\ell \sin \theta_K \sin \phi$	$S_{S5} / (1 - F_S)$

Table 5.1: The angular functions $f_k(\cos \theta_\ell, \cos \theta_K, \phi)$, together with the coefficients with which they appear in the PDF.

The coefficient C_0 is zero for the PDF in **Eq. 1.37**. The function f_0 is included when the simulation of the signal decay is generated using flat distributions for all the angles, the so-called “flat MC”. In this case, $S_0 = 1$ and all other coefficients are zero. Using **Eq. 5.2**, the normalisation factor of **Eq. 5.1** becomes:

$$\int d\Omega \varepsilon(\Omega) \mathcal{P}(\Omega, \vec{S}) = \int d\Omega \varepsilon(\Omega) \sum_k C_k f_k(\Omega) = \sum_k C_k \left[\int d\Omega \varepsilon(\Omega) f_k(\Omega) \right], \quad (5.3)$$

where the factors $\int d\Omega \varepsilon(\Omega) f_k(\Omega)$ define the acceptance moments a_k in the f_k basis:

$$a_k \equiv \int d\Omega \varepsilon(\Omega) f_k(\Omega). \quad (5.4)$$

With this definition, the PDF of the recorded candidates defined in **Eq. 5.1**, can be written as:

$$\mathcal{P}_{\text{obs}}(\Omega, \vec{S}) = \frac{\varepsilon(\Omega) \mathcal{P}_{\text{phys}}(\Omega, \vec{S})}{\sum_k C_k a_k}. \quad (5.5)$$

In the minimisation process of such a PDF, when defining the negative log-likelihood (NLL), the product involving the acceptance function in the PDF transforms into a sum due to the logarithm. Since the factor $\varepsilon(\Omega)$ is independent of $\vec{S}$, its derivative vanishes, meaning the parameter estimation is independent of the acceptance. Consequently, a parameterisation of the acceptance as

a function of Ω is only needed for visualisation, while detector effects are fully encoded in the acceptance moments, which are solely required for normalisation.

Unfortunately, this reasoning is only valid for signal-only angular fits. When fitting a model that includes background, an explicit representation of $\varepsilon(\Omega)$ is required since each background component appears as an additional term in the PDF:

$$\mathcal{P}_{\text{obs}}(\Omega, \vec{S}) = \frac{\varepsilon(\Omega) \mathcal{P}_{\text{phys}}(\Omega, \vec{S}) + \mathcal{P}_{\text{bkg1}} + \mathcal{P}_{\text{bkg2}} + \dots + \mathcal{P}_{\text{bkg}_n}}{\int d\Omega (\varepsilon(\Omega) \mathcal{P}_{\text{phys}}(\Omega, \vec{S}) + \mathcal{P}_{\text{bkg1}} + \mathcal{P}_{\text{bkg2}} + \dots + \mathcal{P}_{\text{bkg}_n})}. \quad (5.6)$$

In this case, the acceptance function no longer factorises during minimisation, making its explicit parameterisation necessary. The parameterisation function is detailed in **Sec. 5.1.2**.

Acceptance moments The acceptance moments a_k can be estimated numerically by MC integration [116]. Assuming a uniformly distributed Ω -e.g., when the integration is performed on a “flat MC” sample - **Eq. 5.4** can be approximated with the following sum:

$$a_k = \frac{1}{N} \sum_i^N \varepsilon(\Omega) f_k(\Omega_i), \quad (5.7)$$

where N is the total number of candidates in the MC simulation. However, for the approximation of the integral by a summation, $\varepsilon(\Omega)$ takes only two values: 1 for candidates within the detector acceptance and 0 for others. Consequently, the acceptance moments can be computed as:

$$a_k = \frac{1}{\sum_i w_i} \sum_i w_i f_k(\Omega_i), \quad (5.8)$$

where i runs only over the candidates within the detector acceptance, and w_i are the weights accounting for differences between the simulation sample and the data (see **Sec. 3.5**). A detailed derivation of **Eq. 5.8** is available in Ref. [95]. When Ω is not uniformly distributed, the angular functions must be normalised by the probability density function (PDF) of Ω . The computation introduces a systematic uncertainty due to the weights w_i and a statistical uncertainty related to the number of candidates in the MC sample. Acceptance moments computed using this approach depend solely on the generated angular values, capturing only the geometrical effects and detection efficiency of the detector while neglecting other factors, such as energy resolution, that appear as a smearing of the generated angular values. This particular effect is outlined in the following section.

Acceptance moments including angular resolution effects The resolution effect is incorporated into the acceptance using the angular values after detector effects have been simulated. Within a MC simulation, “generated” and “reconstructed” angles, Ω_{gen} and Ω_{reco} , are available. The formers correspond to the angular values generated before interacting with the detector, while the latter represents the value of those angles after reconstruction. Reconstruction simulates the passage of particles through the detector and includes the effects of the detector resolution and inefficiencies. The PDF $\mathcal{P}_{\text{obs}}(\Omega_{\text{reco}} | \vec{S}_{\text{phys}})$, defined in **Eq. 5.1**, can incorporate both acceptance and resolution effects through a convolution, as follows:

$$\mathcal{P}_{\text{obs}}(\Omega_{\text{reco}} | \vec{S}_{\text{phys}}) = \frac{1}{N} \int d\Omega_{\text{true}} R(\Omega_{\text{reco}} | \Omega_{\text{true}}) \varepsilon(\Omega_{\text{true}}) \mathcal{P}_{\text{phys}}(\Omega_{\text{true}} | \vec{S}_{\text{phys}}), \quad (5.9)$$

where N is the normalisation constant, and $R(\Omega_{\text{reco}}|\Omega_{\text{true}})$ is the resolution function. Ω_{true} are the real angular distributions of the decay, and they correspond to Ω_{gen} for the $B^0 \rightarrow K^{*0}e^+e^-$ MC samples. In practice, the finite angular resolution of the detector manifests as a distortion of the observed angular distributions, and it can be absorbed into a modified efficiency function, $\bar{\varepsilon}(\Omega_{\text{reco}})$. Consequently, the same PDF can also be expressed in the more familiar form:

$$\mathcal{P}_{\text{obs}}(\Omega_{\text{reco}}|\vec{S}_{\text{phys}}) = \frac{1}{N} \mathcal{P}_{\text{phys}}(\Omega_{\text{reco}}|\vec{S}_{\text{phys}}) \bar{\varepsilon}(\Omega_{\text{reco}}). \quad (5.10)$$

By equating **Eq. 5.9** and **Eq. 5.10**, one can express $\bar{\varepsilon}(\Omega_{\text{reco}})$ as:

$$\bar{\varepsilon}(\Omega_{\text{reco}}) = \frac{\int d\Omega_{\text{true}} R(\Omega_{\text{reco}}|\Omega_{\text{true}}) \varepsilon(\Omega_{\text{true}}) \mathcal{P}_{\text{phys}}(\Omega_{\text{true}}|\vec{S}_{\text{phys}})}{\mathcal{P}_{\text{phys}}(\Omega_{\text{reco}}|\vec{S}_{\text{phys}})}. \quad (5.11)$$

This equation shows how the resolution introduces a dependency of the acceptance on the real values of the angular parameters, $\vec{S}_{\text{phys}}$. Given these assumptions, one can derive the acceptance moments as follows:

$$a_k = \frac{1}{\sum w_i} \sum_i w_i \frac{f_k(\Omega_{\text{reco},i})}{\mathcal{P}(\Omega_{\text{reco},i}|\vec{S}_{\text{phys}})} \frac{\mathcal{P}(\Omega_{\text{gen}}|\vec{S}_{\text{phys}})}{\mathcal{P}(\Omega_{\text{gen}}|\vec{S}_{\text{gen}})}. \quad (5.12)$$

The full derivation is provided in Ref. [95]. Since the true values of the angular parameters are unknown (they are the parameters of interest of this analysis) and only predicted, the first ratio in **Eq. 5.12** quantifies how they can be computed, while the second ratio serves as a correction for this lack of knowledge. In fact, the first ratio would provide an exact estimation if $\vec{S}_{\text{gen}} \equiv \vec{S}_{\text{phys}}$. The uncertainty associated with this assumption can be evaluated using a recursive maximum likelihood fit, in which the values of $\vec{S}_{\text{gen}}$ are updated at each step with $\hat{\vec{S}}_{\text{phys}}$ estimated from the fit. This recursive approach has been adopted in Ref. [95], and shown to introduce a small uncertainty on the parameters estimation. This systematic uncertainty is evaluated in this thesis in **Sec. 6.2.3**.

In summary, this analysis uses the acceptance moments for the normalisation of the angular PDF describing the $B^0 \rightarrow K^{*0}e^+e^-$ decay, computed under the assumption $\vec{S}_{\text{gen}} \equiv \vec{S}_{\text{phys}}$:

$$a_k = \frac{1}{N_{\text{in}}} \sum_i^{N_{\text{in}}} w_i \frac{f_k(\Omega_{\text{reco},i})}{\mathcal{P}(\Omega_{\text{reco},i}|\vec{S}_{\text{gen}})}, \quad (5.13)$$

where $\mathcal{P}(\Omega_{\text{reco},i}|\vec{S}_{\text{gen}})$ is the PDF used to generate the Ω_{gen} in the MC simulation. In this analysis, the MC simulation used to compute and parametrise the acceptance is the so-called ‘‘Physics MC’’ where the $\vec{S}_{\text{gen}}$ are expected to be close to $\vec{S}_{\text{phys}}$.

5.1.2 Acceptance parameterisation

As mentioned earlier, a parameterisation of the acceptance is needed in case of angular fit with background components as well as for visualisation. In this analysis, the acceptance is parameterised with a combination of Legendre polynomials $L(x)$ and Fourier $F(x)$ series terms as:

$$\varepsilon(\cos \theta_\ell, \cos \theta_K, \phi) = \sum_{klm} c_{klm} L_k(\cos \theta_\ell) L_l(\cos \theta_K) F_m(\phi), \quad (5.14)$$

where c_{klm} are the acceptance coefficients and l and m denotes the polynomial order. Fourier series were chosen to parametrise the azimuthal angle ϕ , to ensure the continuity of the function at $\pm\pi$:

$$F_m(\phi) = \begin{cases} \cos\left(\frac{m\phi}{2}\right) & \text{if } m \text{ even} \\ \sin\left(\frac{(m+1)\phi}{2}\right) & \text{if } m \text{ odd} \end{cases} . \quad (5.15)$$

The parameters c_{klm} are calculated via Monte Carlo integration method [116]:

$$c_{klm} = \frac{1}{\sum_{i=1}^N w_i} \sum_{i=1}^N \frac{w_i}{\mathcal{P}(\Omega_{\text{reco},i}|\vec{S}_{\text{gen}})} \left[\left(\frac{2}{2k+1}\right) \left(\frac{2}{2l+1}\right) \left(\frac{1}{p_{mm'}}\right) L_k(\cos\theta_\ell) L_l(\cos\theta_K) F_m(\phi) \right]_i , \quad (5.16)$$

by exploiting the orthogonality of the Polynomial functions:

$$\int_{-1}^{+1} L(x,k) L(x,k') dx = \frac{2}{2k+1} \delta_{kk'} , \quad (5.17)$$

and that of the Fourier terms:

$$\int_{-1}^{+1} F(x,m) F(x,m') dx = p_{mm'} , \quad (5.18)$$

where

$$p_{mm'} = \begin{cases} 2\pi & \text{if } m = m' = 0 \\ \pi & \text{if } m = m' \neq 0 \\ 0 & \text{if } m \neq m' \end{cases} . \quad (5.19)$$

The PDF used to generate the ‘‘Physics MC’’ simulation, $\mathcal{P}(\Omega_{\text{reco},i}|\vec{S}_{\text{gen}})$, is evaluated on the i^{th} -candidate using the reconstructed angles Ω_{reco} and the generator level angular parameters $\vec{S}_{\text{gen}}$. This is also the same PDF that is used to evaluate the acceptance moments for the normalisation, as they must remain consistent between the two methods.

5.2 Mass model

5.2.1 Signal modelling

The invariant mass distribution of the signal candidates has a clear asymmetric tail that results from a combination of effects, such as final state radiation, bremsstrahlung (over-) corrections and stochastic tracking imperfections. The model chosen to parametrise $B^0 \rightarrow K^{*0} e^+ e^-$ decays consists of a sum of two Crystal Ball (CB) functions, referred to as Double Crystal Ball (DCB), that share common values for the peak position and width, but have independent power law tails on opposite sides of the peak. The DCB function is defined as [117]:

$$\mathcal{P}_{\text{DCB}}(m|\mu, \sigma, \alpha_L, \alpha_R, n_L, n_R) = \begin{cases} A_L \cdot \left(B_L - \frac{(m-\mu)}{\sigma}\right)^{-n_L} & \text{if } \frac{(m-\mu)}{\sigma} < -\alpha_L \\ \exp\left(-\frac{(m-\mu)^2}{2\sigma^2}\right) & \text{if } -\alpha_L < \frac{(m-\mu)}{\sigma} < \alpha_R \\ A_R \cdot \left(B_R + \frac{(m-\mu)}{\sigma}\right)^{-n_R} & \text{if } \frac{(m-\mu)}{\sigma} > \alpha_R \end{cases} \quad (5.20)$$

with

$$A_{L,R} = \left(\frac{n_{L,R}}{|\alpha_{L,R}|} \right)^{n_{L,R}} \cdot \exp \left(-\frac{1}{2} |\alpha_{L,R}|^2 \right), \quad (5.21)$$

$$B_{L,R} = \frac{n_{L,R}}{|\alpha_{L,R}|} - |\alpha_{L,R}|, \quad (5.22)$$

where μ and σ are the mean value and resolution of the Gaussian part of the function. The $\alpha_{R/L}$ parameters define the points - expressed in units of σ - at which the Gaussian turns into a power-law tail on the right or left side, respectively. The parameters $n_{R/L}$ control the shape of these power-law tails, determining how rapidly the function falls off beyond the transition point.

The bremsstrahlung recovery procedure significantly affects the shape of the signal mass distribution. As a result, the three categories identified in Section 2.7.1 are parameterised separately. To accommodate for the observed distribution of the data in the *Brem1* and *Brem2* categories a Gaussian function is used in addition to the DCB. The fits to the Run 2 MC simulation of the signal decay in each *brem* category of the central q^2 region are shown in **Fig. 5.1**. The values of the parameters are listed in **Tab. C.1** for Run 1 and Run 2 datasets, and all the q^2 regions under analysis. The fits for the two smaller q^2 regions are shown in **Fig. C.1**.

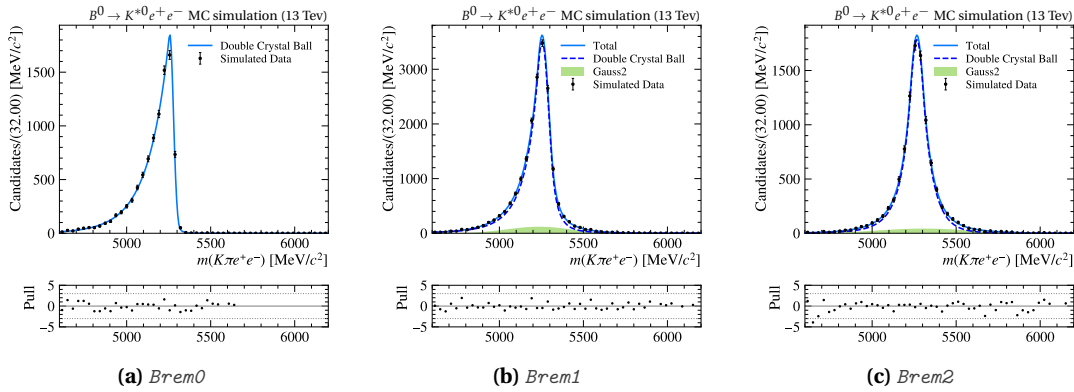


Figure 5.1: Fit to the $m(K\pi e^+ e^-)$ reconstructed mass of the $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation at 13 TeV. Selected q^2 region: $[1.1, 6.0] \text{ GeV}^2/c^4$.

5.2.2 Background modelling

Combinatorial and DSL

These two backgrounds are modelled together using a single exponential function. The slope is obtained by fitting the flavour-violating data sample, $B^0 \rightarrow K^{*0} e \mu$. The fit in the central q^2 region is shown in **Fig. 5.2**. In the fit to the $B^0 \rightarrow K \pi e^+ e^-$ data, the slope of this background is fixed to the value obtained by fitting the $B^0 \rightarrow K^{*0} e \mu$ sample, and the yield is left free. **App. C** shows the fits (**Fig. C.2**) and their corresponding results (**Tab. C.3**) in the two smaller q^2 regions under analysis.

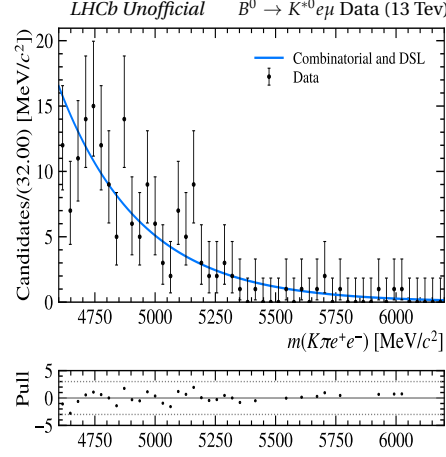


Figure 5.2: Fit to the $m(K\pi e^+ e^-)$ reconstructed mass of the $B^0 \rightarrow K^{*0} e \mu$ using Run 2 samples at 13 TeV. Selected q^2 region: $[1.1, 6.0] \text{ GeV}^2/c^4$.

Misidentified electrons background

The misidentified background is modelled using the sample described in **Sec. 3.5.2**. The functional form for this background in mass follows the choice made in Refs. [50, 49]:

$$\mathcal{P}(m|\mu_{\pi\pi}, \sigma_{\pi\pi}, \mu_{KK}, \sigma_{KK}, \lambda, f_1, f_2) = f_1 \text{Exp}(m|\lambda) + \quad (5.23)$$

$$+ f_2 \text{Gauss}(m|\mu_{KK}, \sigma_{KK}) + (1 - f_1 - f_2) \text{Gauss}(m|\mu_{\pi\pi}, \sigma_{\pi\pi}) \quad (5.24)$$

where the two Gaussian components account for cases where both electrons are misidentified and an exponential function accounts for the case where just one electron is misidentified. When the two electrons are pions ($\text{Gauss}(m|\mu_{\pi\pi}, \sigma_{\pi\pi})$), the missing energy due to the mass difference between electrons and pions is sufficient to shift the peak away from that of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay of about 20 MeV/c^2 below the B^0 pole mass. When both electrons are kaons, the larger mass difference between the kaon and electron results in a larger shift of the peak to lower mass, approximately 230 MeV/c^2 below the B^0 mass peak. Using this information, the means of the two Gaussian components are fixed to the expected values of the corresponding peaks, while the widths (σ) are left floating to find an optimal value that works for both Run 1 and Run 2. The exponential component accounts for all other single misidentified decays and varies among samples corresponding to the two runs. A feature of this background, in common with the DSL background, is a strong correlation between $m(K\pi e^+ e^-)$ and $\cos \theta_\ell$. This effect is visible in **Fig. 5.3a**, where the two-dimensional plot of the misidentified background is shown. This suggests that most of the decays contributing to this background are secondary B^0 decays proceeding through intermediate heavy particles, such as a D meson. For example, residual $B^0 \rightarrow D^- (\rightarrow K^{*0} \pi^+) \pi^-$ decays can pass the K^{*0} selection criteria and have a branching ratio of $\mathcal{O}(10^{-5})$, which is an order of magnitude larger than that of the signal decay. The fit to the misidentified dataset estimated for Run 2 in the central q^2 region is shown in **Fig. 5.3b**. **App. C** shows the fits (**Fig. C.2**) and their corresponding results (**Tab. C.2**) in the two smaller q^2 regions under analysis.

In the final fit to the data, the shape of this background is fully fixed and the yield is Gaussian constrained to the expected amount of misidentified events in data, computed as in **Eq. 3.6** in

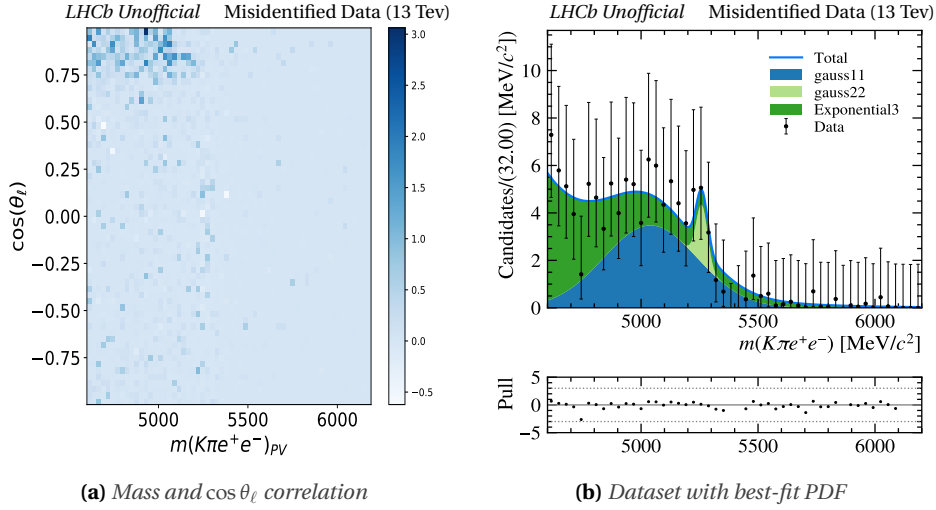


Figure 5.3: Both plots make use of the misidentified background enriched sample, obtained using Run 2 dataset in the central q^2 region. The first highlights the correlation between invariant mass and angular distribution by showing the distribution on the two-dimensional plane, $m(K\pi e^+ e^-) - \cos \theta_\ell$. The second shows the best-fit PDF on the invariant mass distribution.

Sec. 3.5.2.

Partially Reconstructed

The modelling strategy for this background involves using a Kernel Density Estimator (KDE), a non-parametric method for approximating the probability density function of a given dataset. In this case, the $B \rightarrow K\pi\pi e^+ e^-$ MC sample is used for modelling. In the final fit the shape is fixed, while the yield is left free to vary to account for uncertainties. The shape of the KDE in the central q^2 region for Run 2 is shown in **Fig. 5.4a**, superposed to the MC events. The fits are shown in **Fig. C.3** for the two smaller q^2 regions, and for both Run 1 and Run 2 samples.

J/ψ leakage

The last background to be considered is the leakage of the abundant $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ decays. This background is modelled with a KDE in the q^2 region of interest, using the $B^0 \rightarrow K^{*0} J/\psi$ MC simulation. The shape of the KDE in the central q^2 region for Run 2 is shown in **Fig. 5.4b**, superposed to the MC events. The expected amount of leakage is estimated as follows:

$$N_{q^2 \text{ RareMode}, \text{data}}^{J/\psi} = N_{q^2 J/\psi, \text{data}}^{J/\psi} \frac{N_{q^2 \text{ RareMode}, \text{MC}}^{J/\psi}}{N_{q^2 J/\psi, \text{MC}}^{J/\psi}}, \quad (5.25)$$

where $N_{q^2 J/\psi, \text{data}}^{J/\psi}$ is estimated by performing a fit to the $m(K\pi e^+ e^-)$ distribution in data selecting the q^2 region from $7 \text{ GeV}^2/c^4$ to $11 \text{ GeV}^2/c^4$. The scaling ratio is computed using the MC simulation: $N_{q^2 \text{ RareMode}, \text{MC}}^{J/\psi}$ is the number of MC events in the q^2 region of interest, $N_{q^2 J/\psi, \text{MC}}^{J/\psi}$ is the total number of J/ψ events in the MC simulation in the q^2 region from 7 to 11 GeV^2/c^4 . In the final fit, the yield of this background is constrained using a Gaussian distribution with a mean of $N_{q^2 \text{ RareMode}, \text{data}}^{J/\psi}$ and width of $\sqrt{\sum_i w_i^2}$, with w_i being the MC correction weights. The fits are shown in **Fig. C.3** for the two smaller q^2 regions, and both Run 1 and Run 2 samples.

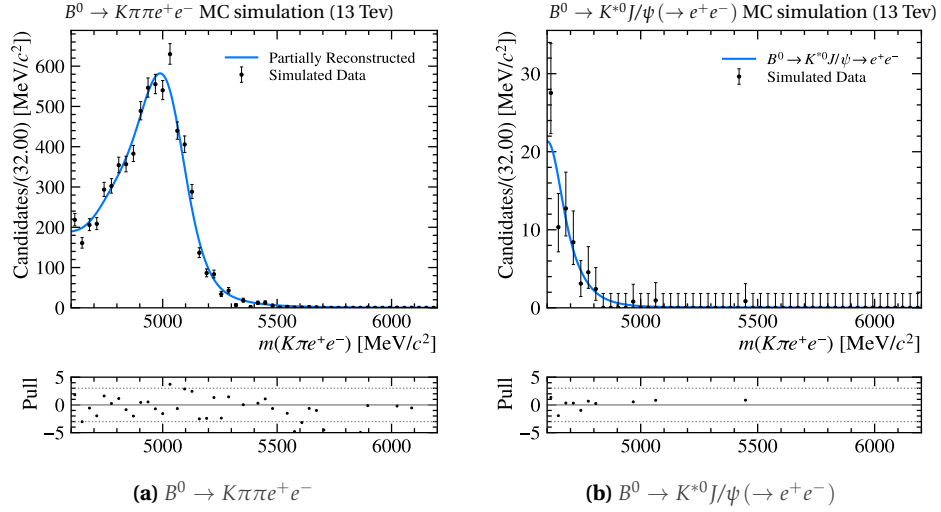


Figure 5.4: Fit to the $m(K\pi e^+e^-)$ reconstructed mass of the $B^0 \rightarrow K\pi\pi e^+e^-$ MC samples at 13 TeV (left). Fit to the $m(K\pi e^+e^-)$ reconstructed mass of the $B^0 \rightarrow K^{*0}J/\psi$ MC samples at 13 TeV (right). Selected q^2 region: $[1.1, 6.0] \text{ GeV}^2/c^4$.

5.2.3 Results

The fit to the reconstructed $m(K\pi e^+e^-)$ data is performed as described in the sections above and the yields obtained in all the cases are listed in **Tab. 5.2**. One can observe that for certain datasets, the J/ψ leakage and the partially reconstructed backgrounds are not reported. In the case of J/ψ leakage, the fit is performed neglecting this background component, as its expected contribution is zero. For the partially reconstructed background, the yield is not shown when an initial fit - allowing all yields to float - returns a negative value. In such cases, the table reports the results of an alternative fit in which this background is neglected.

The total background is also reported in the table as the sum of all components, with the error computed by adding all uncertainties in quadrature.

Yields	central _{low}		central _{high}		central	
	Run 2	Run 1	Run 2	Run 1	Run 2	Run 1
Signal	154.62 ± 16.12	37.67 ± 8.28	185.76 ± 16.98	38.24 ± 8.96	341.07 ± 23.44	78.50 ± 11.92
DSL-Comb	73.74 ± 21.00	26.20 ± 7.64	82.05 ± 25.84	20.51 ± 15.51	154.67 ± 33.05	48.25 ± 13.25
J/ψ	-	-	32.10 ± 4.20	5.15 ± 1.21	33.15 ± 4.24	5.15 ± 1.2
Misid	51.35 ± 7.17	28.68 ± 5.24	56.16 ± 7.53	41.73 ± 6.66	110.04 ± 10.48	73.21 ± 8.49
PR	92.80 ± 23.01	-	68.33 ± 24.57	9.09 ± 14.87	154.44 ± 33.16	-
TOT Bkg	217.89 ± 31.96	54.88 ± 9.26	238.64 ± 36.68	76.48 ± 22.52	452.30 ± 48.16	126.61 ± 49.00

Table 5.2: Yield of the various fit components extracted by separate fits to each of the datasets.

An example of fit is shown in **Fig. 5.5**, for the central q^2 region from 1.1 to 6.0 GeV^2/c^4 , with the full selection described in **Chap. 3** and **Chap. 4**.

Using the yields extracted from this fit, a second fit is performed in which the relative fractions of the background components are fixed, while only the total background and signal yields are left floating. This approach allows for the extraction of signal and background weights, and the results of this second fit are listed in **Tab. 5.3**.

The main reason why the background weights are not extracted from the first fit - where all yields were left floating - is due to the Gaussian-constrained yields. The presence of two partially

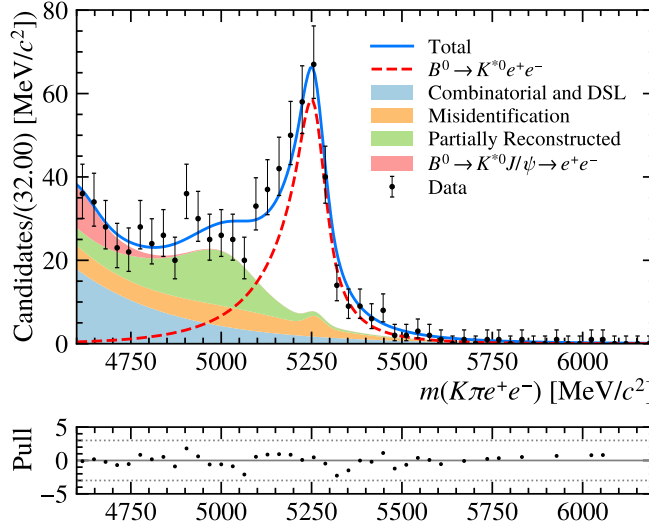


Figure 5.5: Fit to Run 2 $B^0 \rightarrow K\pi e^+ e^-$ candidates with final selection applied and selecting the full central q^2 region from 1.1 to 6.0 GeV^2/c^4 .

constrained background components caused instabilities in the weight computation. Therefore, performing a second fit with fixed background fractions serves as a workaround to this problem.

These background weights play an important role in the final fit strategy, which is detailed in the following section.

Yield	central_low		central_high		central	
	Run 2	Run 1	Run 2	Run 1	Run 2	Run 1
sig	152.94 ± 16.28	37.53 ± 8.66	184.98 ± 17.00	41.21 ± 8.96	341.18 ± 23.43	78.87 ± 12.45
bkg	218.89 ± 18.19	53.56 ± 9.55	236.00 ± 18.43	71.79 ± 10.52	451.72 ± 25.69	124.77 ± 14.18

Table 5.3: Signal and background yields resulting from separate fit to reported datasets when fixing the relative fractions of the background to the values in **Tab. 5.2**.

5.3 Angular fit

The values of the angular observables are extracted from an unbinned maximum likelihood fit to the reconstructed $m(K\pi e^+ e^-)$ mass, and the three decay angles $\Omega = (\cos \theta_\ell, \cos \theta_K, \phi)$. Both signal and background components are described by analytically normalised PDFs, which depend on observables of interest, the angular parameters, S_i , describing the signal angular PDF on **Eq. 1.37**, and nuisance parameters. The total PDF is of the form:

$$\mathcal{P}_{\text{tot}}(m_{K\pi ee}, \Omega | \vec{S}) = f_S \mathcal{P}_S(m_{K\pi ee}, \Omega | \vec{S}) + \sum_i^{n-1} f_{B,i} \mathcal{P}_{B,i}(m_{K\pi ee}, \Omega) + \left(1 - f_S - \sum_i^{n-1} f_{B,i}\right) \mathcal{P}_{B,n}(m_{K\pi ee}, \Omega) \quad (5.26)$$

where f_S is the signal fraction and $f_{B,i}$ are the fractions of the background components. The PDF of all the components are assumed to factorize between the mass and the angular parts:

$$\mathcal{P}_S(m_{K\pi ee}, \Omega | \vec{S}) = \mathcal{P}_S(m_{K\pi ee}) \mathcal{P}_{phys}(\Omega | \vec{S}) \quad (5.27)$$

$$\mathcal{P}_{B,i}(m_{K\pi ee}, \vec{\Omega}) = \mathcal{P}_{B,i}(m_{K\pi ee}) \mathcal{P}_B(\Omega) \quad (5.28)$$

where $\mathcal{P}_{phys}(\Omega | \vec{S})$ is defined in **Eq. 1.37**, $\mathcal{P}_S(m_{K\pi ee})$ and $\mathcal{P}_{B,i}(m_{K\pi ee})$ are the mass PDFs for the signal and background components defined in the previous section. The mass shape of each background is fixed in the fit, and the corresponding parameters are treated as nuisance parameters. The parameterisation of the angular PDF of all backgrounds, $\mathcal{P}_B(\Omega)$, assumes factorisation among the angular variables, and it is written as follows:

$$\mathcal{P}_{B,n}(\Omega | \alpha_k, \beta_i, \gamma_m) = L_k^n(\cos \theta_\ell | \alpha_k) L_i^n(\cos \theta_K | \beta_i) F_m(\phi | \gamma_m), \quad (5.29)$$

where L is a Legendre polynomial of order k, i for $\cos \theta_\ell$ and $\cos \theta_K$ respectively. F_m is a Fourier series, defined in **Eq. 5.15**. This factorising function is an approximation, as it is known to break down for the partially reconstructed background due to correlations between $\cos \theta_\ell$ and $\cos \theta_K$ arising from the energy of the missing particle. A more detailed discussion about this assumption and potential improvements is provided in **Chap. 7**. The order of the polynomial functions and the initial values of their coefficients are determined by fitting the background-weighted data. The background weights are extracted from the fit to $m(K\pi e^+ e^-)$ described in **Sec. 5.2.3**, exploiting the “sPlot” strategy [118]. To summarise, the angular function for the backgrounds is determined in three steps:

- Retrieve the relative fraction of each background component from a fit to the $m(K\pi e^+ e^-)$ invariant mass.
- Retrieve background weights by performing a second fit to $m(K\pi e^+ e^-)$, this time fixing the relative yields of the background components and floating only the signal and total background yields.
- Fit the background-weighted data to determine the best-fit polynomial orders. The resulting coefficients can be used as initial values in the final fit.

Once the angular shape of the backgrounds is determined, the final fit to the mass and angular distributions can be performed. The free parameters in this fit are:

- S_i : angular parameters of interest,
- $B^0 \rightarrow K^{*0} e^+ e^-$ yield,
- all the background yields - although the leakage originating from the $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ decay and the misidentified background yields are Gaussian constrained,
- $\alpha_k, \beta_i, \gamma_m$: angular coefficients of the background introduced in **Eq. 5.29**.

The fit is done simultaneously between Run 1 and Run 2 using the full integrated luminosity collected from 2011 to 2018, 9 fb^{-1} . The common parameters of the simultaneous fit are only the S_i angular parameters. The acceptance function is estimated separately for the two runs.

5.3.1 Angular model on signal MC simulation

The angular model for the signal can be validated using the $B^0 \rightarrow K^{*0}e^+e^-$ MC simulated data. Specifically, this is the same dataset and angular distributions used to compute the acceptance moments and acceptance parametrisation (see **Sec. 5.1**). As explained in **Sec. 5.1.1**, the generator-level angular parameters used to generate the MC samples are required for computing the acceptance moments. These parameters can be obtained by fitting the generator-level angles of the $B^0 \rightarrow K^{*0}e^+e^-$ MC simulation, which do not include any detector effects.

Once these values are obtained, they can be used to estimate the acceptance moments required for normalising the signal probability density function (PDF), as discussed in **Sec. 5.1.1**. The efficiency function has been parametrised using Legendre polynomials of order five for $\cos \theta_\ell$, order six for $\cos \theta_K$, and Fourier series terms up to order five for ϕ (see **Sec. 5.1.2**).

The parameters resulting from the fit to the reconstructed angles in the $B^0 \rightarrow K^{*0}e^+e^-$ MC samples are compared to the angular parameters obtained from fitting the generator-level angular distributions. These comparisons are shown in **Tab. 5.4** for central q^2 , in **Tab. 5.5** for central_{low} q^2 , and in **Tab. 5.6** for central_{high} q^2 . The values extracted from fits to the reconstructed angles using Run 1 and Run 2 samples are listed under the labels *Run 1* and *Run 2*, respectively, while the angular parameters obtained from generator-level distributions are listed under the label *Gen.*.

The values obtained by fitting the reconstructed angles are compatible with the generator values because the same sample of simulated data was used both to compute the acceptance moments and to perform the fit itself. The statistical uncertainties from this fit reflect the limited size of the MC samples used to compute the acceptance, and are therefore treated as systematic uncertainties in **Sec. 6.2.3**.

Fig. 5.6, **Fig. 5.7**, and **Fig. 5.8** also show the angular distributions obtained using the generated angles, labeled *Gen. Angles*, as well as those reconstructed from signal MC simulations that emulate the Run 1 and Run 2 datasets, labelled *Reco. Angles Run 1* and *Reco. Angles Run 2*, respectively. The corresponding best-fit models are overlaid on each angular distribution. Each figure allows for a direct comparison between generator-level and reconstructed-level simulations. This comparison highlights the impact of detector effects, which are especially pronounced in the $\cos \theta_K$ distribution.

Central q^2 [1.1,6.0] GeV^2/c^4			
Param	Gen.	Run 1	Run 2
N. events	135 212	18 947	44 069
A_{FB}	-0.0431 ± 0.0016	-0.0422 ± 0.0044	-0.0422 ± 0.0029
F_L	0.7262 ± 0.0017	0.7277 ± 0.0047	0.7279 ± 0.0031
S_3	-0.0065 ± 0.0023	-0.0058 ± 0.0056	-0.0075 ± 0.0036
S_4	-0.1270 ± 0.0028	-0.1279 ± 0.0079	-0.1259 ± 0.0052
S_5	-0.1486 ± 0.0028	-0.1484 ± 0.0072	-0.1484 ± 0.0047
S_7	0.0095 ± 0.0029	0.0067 ± 0.0076	0.0095 ± 0.0050
S_8	0.0052 ± 0.0029	0.0037 ± 0.0083	0.0049 ± 0.0054
S_9	0.0060 ± 0.0023	0.0068 ± 0.0056	0.0055 ± 0.0036

Table 5.4: Fit results to the Generator level MC sample at 13 TeV, and Run 1 (Run 2) signal reconstructed $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulated data used to compute the acceptance functions in the central q^2 region.

Central _{low} q^2 [1.10,3.65] GeV^2/c^4			
Param	Gen.	Run 1	Run 2
N. events	68 860	9 705	22 137
A_{FB}	-0.1504 ± 0.0020	-0.1502 ± 0.0060	-0.1497 ± 0.0041
F_L	0.7378 ± 0.0024	0.7391 ± 0.0067	0.7380 ± 0.0045
S_3	-0.0045 ± 0.0030	-0.0031 ± 0.0073	-0.0053 ± 0.0048
S_4	-0.0561 ± 0.0039	-0.0559 ± 0.0115	-0.0546 ± 0.0076
S_5	0.0086 ± 0.0041	0.0065 ± 0.0103	0.0095 ± 0.0068
S_7	0.0121 ± 0.0041	0.0110 ± 0.0104	0.0110 ± 0.0070
S_8	0.0066 ± 0.0040	0.0063 ± 0.0114	0.0054 ± 0.0077
S_9	0.0018 ± 0.0031	0.0058 ± 0.0072	0.0016 ± 0.0048

Table 5.5: Fit results to the Generator level MC sample at 13 TeV, and Run 1 (Run 2) signal reconstructed $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulated data used to compute the acceptance functions in the central_{low} q^2 region.

Central _{high} q^2 [3.65,6.00] GeV^2/c^4			
Param	Gen.	Run 1	Run 2
N. events	66 352	9 242	21 932
A_{FB}	0.0681 ± 0.0022	0.0695 ± 0.0060	0.0687 ± 0.0038
F_L	0.7144 ± 0.0024	0.7160 ± 0.0064	0.7172 ± 0.0042
S_3	-0.0092 ± 0.0032	-0.0107 ± 0.0079	-0.0103 ± 0.0049
S_4	-0.2028 ± 0.0034	-0.2044 ± 0.0095	-0.2018 ± 0.0064
S_5	-0.3135 ± 0.0036	-0.3114 ± 0.0094	-0.3136 ± 0.0061
S_7	0.0065 ± 0.0041	0.0023 ± 0.0108	0.0072 ± 0.0069
S_8	0.0029 ± 0.0040	0.0015 ± 0.0112	0.0037 ± 0.0072
S_9	0.0095 ± 0.0032	0.0072 ± 0.0078	0.0083 ± 0.0051

Table 5.6: Fit results to the Generator level MC sample at 13 TeV, and Run 1 (Run 2) signal reconstructed $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulated data used to compute the acceptance functions in the central_{high} q^2 region.

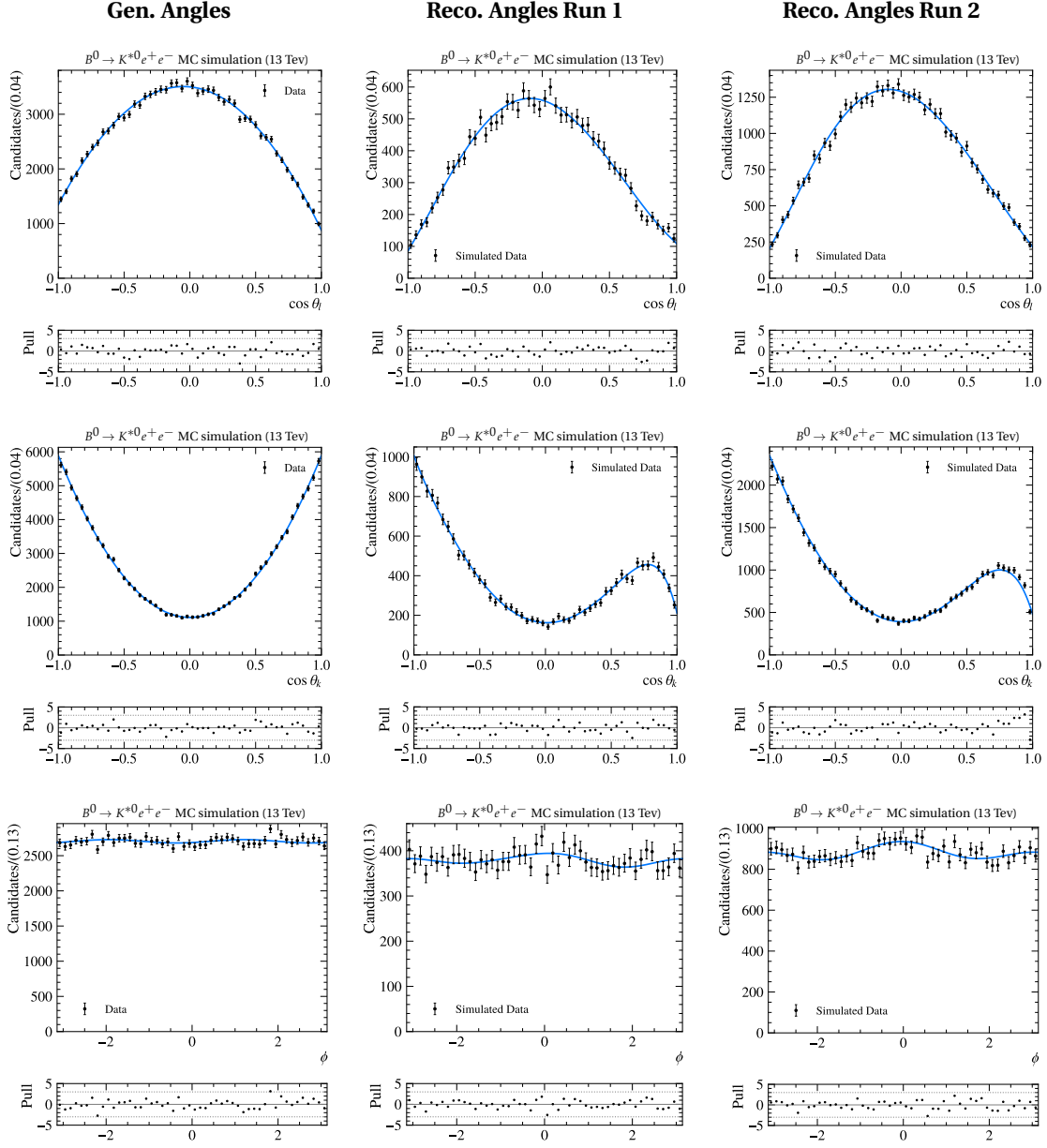


Figure 5.6: Angular distributions for signal $B^0 \rightarrow K^{*0} e^+ e^-$ simulated data, at generator level (left), reconstructed Run 1 (centre) and Run 2 (right), in the central q^2 ($[1.10, 6.0]$ GeV^2/c^4).

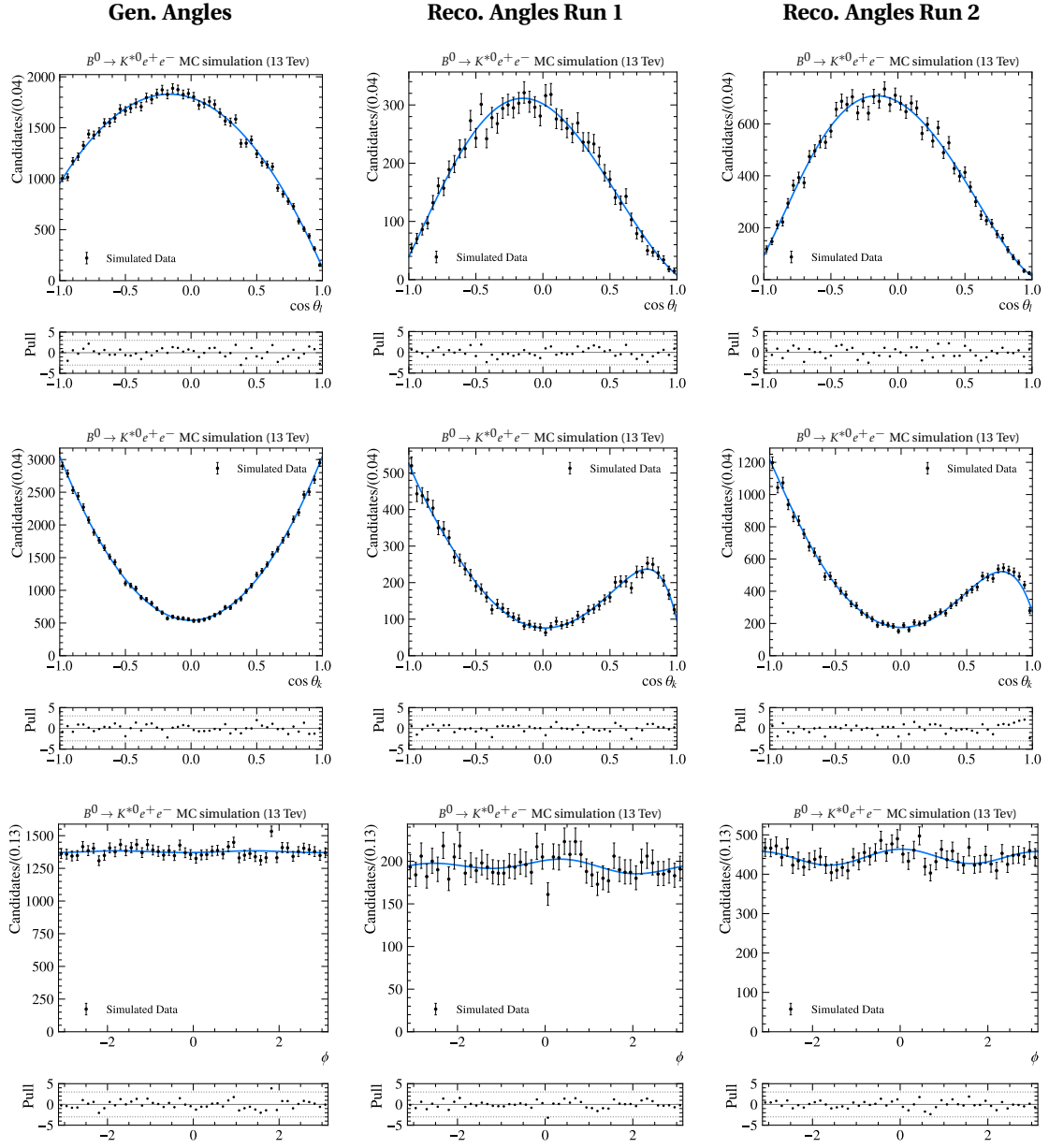


Figure 5.7: Angular distributions for signal $B^0 \rightarrow K^{*0} e^+ e^-$ simulated data, at generator level (left), reconstructed Run 1 (centre) and Run 2 (right), in the central q^2 ($[1.10, 3.65] \text{ GeV}^2/c^4$).

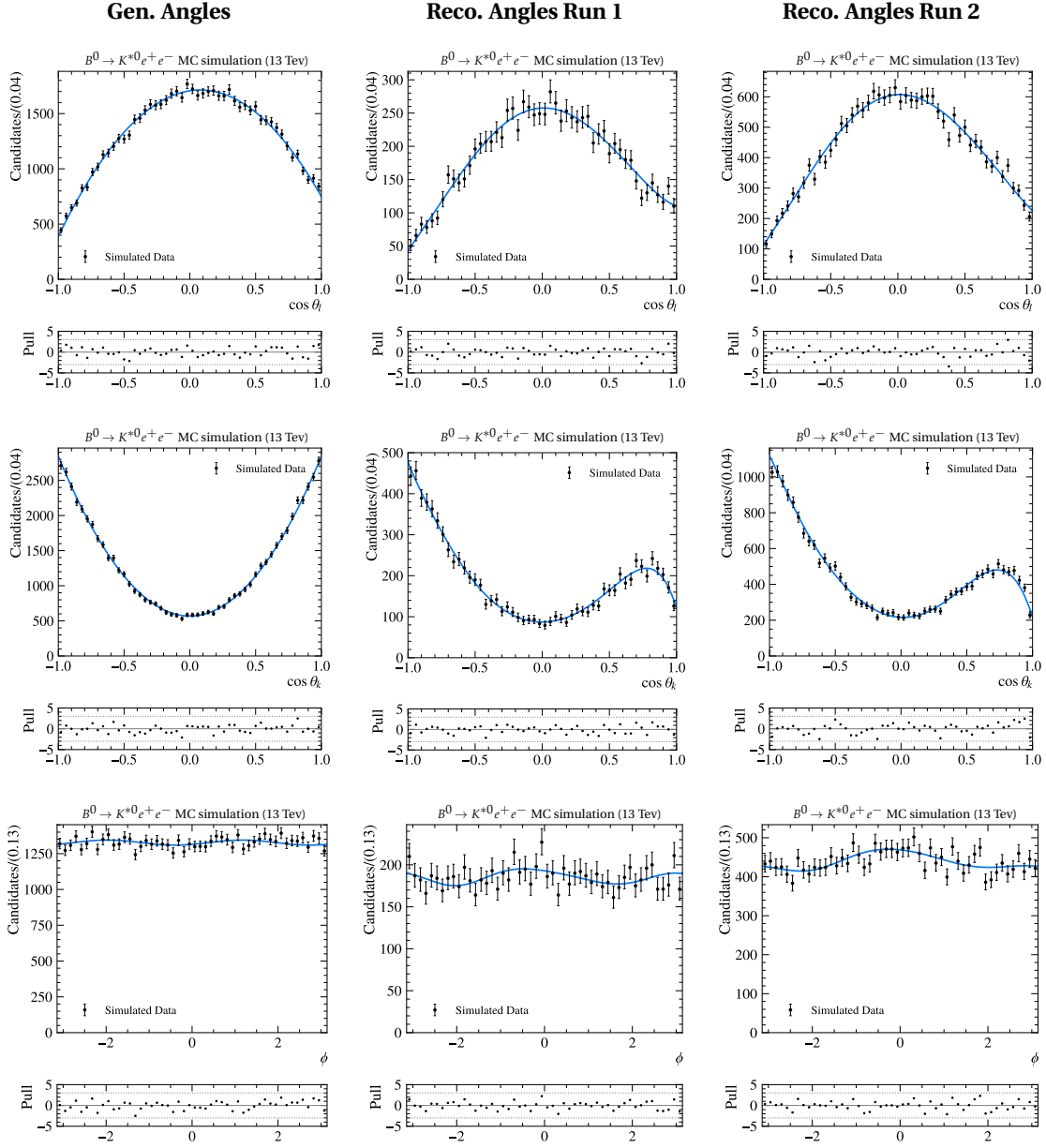


Figure 5.8: Angular distributions for signal $B^0 \rightarrow K^{*0} e^+ e^-$ simulated data, at generator level (left), reconstructed Run 1 (centre) and Run 2 (right), in the central q^2 ($[3.65, 6.00] \text{ GeV}^2/c^4$).

6 | Expected sensitivity and systematic uncertainties

This chapter presents the validation of the fit strategy using toy simulations. Initially, only signal events are generated based on the signal PDF. Subsequently, realistic toy datasets are produced, incorporating both signal and background contributions. This approach validates the fit strategy adopted and the estimation of several systematic uncertainties. Finally, a blinded fit is conducted on the collected dataset. The results are presented for the full central q^2 region, as well as for the two smaller central q^2 bins.

6.1 Fit strategy validation using toy simulations

The validation of the fit strategy described in **Chap. 5** is performed using toy simulations and comes in two steps. The first is to assess biases due to the limited number of signal events available in the collected dataset. The second is to generate realistic toys incorporating both signal and background contributions. From each toy, the pull value of each parameter of interest is computed as follows:

$$Pull(\vec{\alpha})_i = \frac{\vec{\alpha}_i^{fit} - \vec{\alpha}_i^{gen}}{\vec{\sigma}_i^{fit}}, \quad (6.1)$$

where $\vec{\alpha} = (A_{FB}, F_L, S_3, \dots, S_9)$ are all the angular parameters and $\vec{\alpha}^{gen}$ are their generated values, listed under the column named **Tab. 5.4**, **Tab. 5.5**, and **Tab. 5.6** for the central, central_{low}, and central_{high} q^2 regions, respectively. $\vec{\alpha}_i^{fit}$ and $\vec{\sigma}_i^{fit}$ are the parameter values and their statistical errors from the maximum likelihood Estimators (MLEs). The index $i = 1, \dots, N$ runs over the number of toy samples generated (N). Thanks to the central limit theorem, the pulls are normally distributed for sufficiently large samples. This property allows to investigate for the presence of bias, and allows the verification of the error coverage [119].

6.1.1 Limited sample size biases with signal-only toys

The first step to validate the fit strategy is to assess its behaviour in a signal-only scenario, particularly when dealing with a limited number of signal events, as it is known that the MLE is only asymptotically unbiased [120]. This is achieved using signal-only toy simulations. Signal events are generated based on the signal PDF including acceptance effects, see **Eq. 5.5**. The number of signal events for each toy sample is randomly sampled from a Poisson distribution with a mean equal to the signal yield expected in each q^2 region, **Tab. 5.2**. An unbinned maximum likelihood fit is then performed on each toy sample using the very same PDF used to generate

them. **Fig. 6.1** shows an overview of the mean (μ) and the variance (σ) of the Gaussian function used to fit the pull distributions of ~ 1000 toys, in each q^2 bin. The pull distribution for each of the angular parameters is shown for completeness on **App. D.3, Fig. D.1**, for the central q^2 , **Fig. D.2** for the central_{low} q^2 , and **Fig. D.3** for central_{high} q^2 regions. While the full central

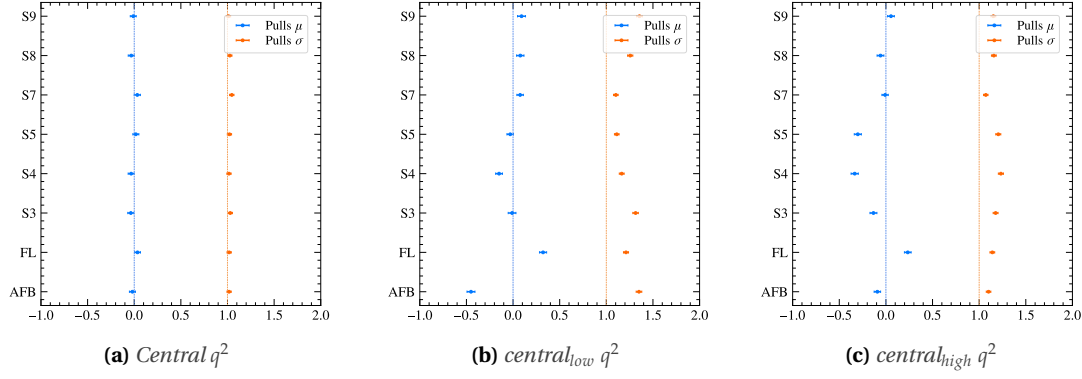


Figure 6.1: Signal-only toys results. Mean and variance of the pull distributions from 1000 toys with realistic stat. Number of events per toy randomly extracted from poissonian with mean at 154 or 185 for central -low or -high q^2 respectively.

q^2 region, $[1.1, 6.0] \text{ GeV}^2/c^4$, presents pull distributions which are very well aligned around zero (means) and one (sigmas), that is not the case for the central -low or -high q^2 regions. This effect

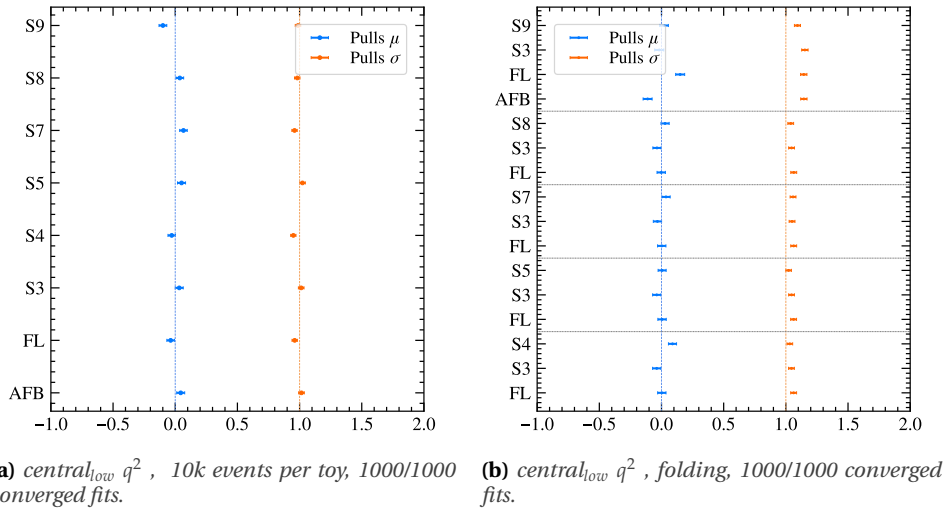


Figure 6.2: Signal-only toys results for central_{low} q^2 region. Mean and variance of the pulls distribution from 1000 toys, showing how increasing the number of events per toy (a) or fixing some of the parameters to zero and performing multiple fits (b) cures most of the biases.

is shown to stem from the limited number of signal events and seems to be aggravated when a fitted parameter is close to the border of its allowed physics region (shown in **App. G**), as it will be discussed in **Sec. 7.1**. The pull distributions are well centred when generating 1000 toy samples with approximately 10 000 events each, as shown in **Fig. 6.2a**. One way to address this behaviour at low statistics is to apply the so-called “folding” strategy, already used within the LHCb collaboration, *e.g.* in Ref. [32]. This technique exploits symmetries of the differential decay rate in the angles to remove some observables, thereby reducing the number of free parameters in

the fit. To fully cover all angular parameters, five foldings are required. This procedure is detailed in Ref. [121], and **App. E** shows that it essentially amounts to fixing the removed parameters to zero. The same toy samples generated with a realistic number of events for the central_{low} q^2 region have been fitted fixing to zero the removed parameters of each folding. An overview of the mean and standard deviation of the pulls is shown in **Fig. 6.2b**, where the parameters extracted from the five fits are grouped by the grey lines. While fixing to zero a few parameters is not the preferred strategy for this analysis, it remains a viable option for obtaining unbiased estimates of the angular parameters without loss in precision. This thesis instead presents “unfolded” fit and systematic studies.

6.1.2 Realistic toys

To fully validate the fit strategy, toy simulations are generated mimicking data in terms of sample size, signal and background components. These toy simulations are referred to in this thesis as realistic toys. In the realistic toys, the signal mass and angular distributions are sampled from the signal PDF as reported in **Eq. 5.27**. The background components are generated as follows:

- **Partially Reconstructed:** the mass and the three angular distributions are bootstrapped from the $B^0 \rightarrow K\pi\pi e^+e^-$ MC sample accounting for the MC simulation correction weights (**Sec. 3.5.1**). The yield is drawn from a Poisson distribution with mean fixed to the value obtained from the $m(K\pi e^+e^-)$ fit, as reported in **Tab. 5.2**.
- **Misidentified background:** due to the limited size of the available sample, this background is modelled in mass and angles assuming a fully factorised PDF:

$$\mathcal{P}_{misid}(m, \cos \theta_\ell, \cos \theta_K, \phi) = \mathcal{P}(m)\mathcal{P}(\cos \theta_\ell)\mathcal{P}(\cos \theta_K)\mathcal{P}(\phi) \quad (6.2)$$

$$= \mathcal{P}_{misid}(m)L_k(\cos \theta_\ell|\alpha_k)L_i(\cos \theta_K|\beta_i)F_m(\phi|\gamma_m), \quad (6.3)$$

where $\mathcal{P}_{misid}(m)$ is defined in **Eq. 5.23** and depends on the q^2 region and on the data-taking run. L is a Legendre polynomial of order k, i for $\cos \theta_\ell$ and $\cos \theta_K$ respectively. F_m is a Fourier series, defined in **Eq. 5.15**. Events are generated according to the resulting four-dimensional PDF, $\mathcal{P}_{misid}(m, \cos \theta_\ell, \cos \theta_K, \phi)$. The number of generated events fluctuates according to a Poisson distribution with the mean value fixed to the expected amount reported in **Tab. 5.2**.

- **DSL and Combinatorial:** the mass and the angular distributions of these backgrounds are simulated by bootstrapping the lepton-flavour-violating data sample $B^0 \rightarrow K^{*0}e\mu$. The yield is drawn from a Poisson distribution with mean fixed to the value “DSL-comb” obtained from the $m(K\pi e^+e^-)$ fit, as reported in **Tab. 5.2**.
- **$B^0 \rightarrow K^{*0}J/\psi$ leakage:** the mass and angular distributions are bootstrapped from the $B^0 \rightarrow K^{*0}J/\psi$ MC simulation accounting for the MC simulation correction weights (**Sec. 3.5.1**). The number of events are drawn from a Poisson distribution with fixed mean on the expected amount reported in **Tab. 5.2**.

Each component is then summed up to create a toy sample emulating the size and composition of the real data.

The realistic toys are generated 1000 times emulating the full central q^2 , central_{low} q^2 , central_{high} q^2 , Run 1, and Run 2 datasets. Each realistic toy is compared to data and a χ^2

test is performed on the mass and the angular binned distributions. Of course, the value depends on the binning (30 bins in each variable) and varies as each toy is generated using Poissonian fluctuations. The χ^2 distributions for the mass and angular variables are reported in **Fig. 6.3** for the Run 1 and Run 2 datasets and in each of the q^2 region under analysis.

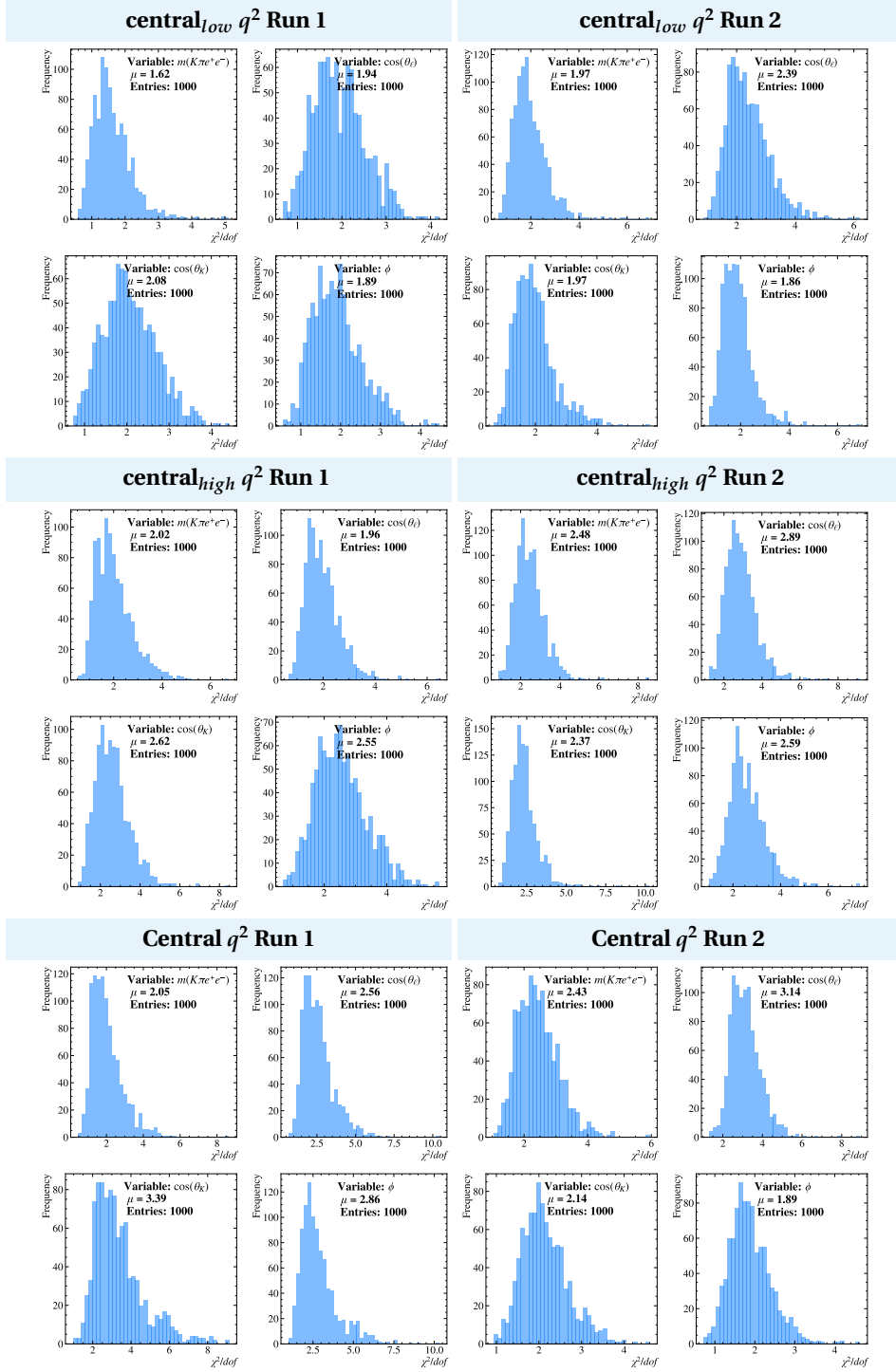


Figure 6.3: The plots displays the χ^2/dof statistic computed between data and the realistic toys emulating the Run 1 and the Run 2 datasets, on the $m(K\pi^+e^-)$, $\cos\theta_\ell$, $\cos\theta_K$, and ϕ binned distributions (30 bins). This is done in the three q^2 regions under analysis.

The plots show reasonably distributed $\chi^2/ndof$, where $ndof = N_{bins} - 1$ with a peak around 2 indicating moderate deviations. It is important to note that this test is not intended as a hypothesis test, since the distributions are known a priori to not to be exactly the same. Rather, it serves as a way to quantify the overall level of agreement between toys and data, providing a summary metric that is more practical than visually inspecting 1000 toy experiments individually. In light of these considerations, this level of agreement is deemed acceptable.

The four-dimensional fit on mass and angular distributions is performed simultaneously on the Run 1 and Run 2 toy samples in each of the q^2 regions. The fit procedure adopted for each toy, it follows the method outlined in **Chap. 5**, with one key difference: Gaussian-constrained yields are assigned values randomly sampled from a Gaussian distribution centered at the expected quantity, with a standard deviation equal to the uncertainty on the expected value [122]. Fluctuating constrained yields by generating them with a Poisson distribution whose mean is itself drawn from the Gaussian constraint would be incorrect [119], as this procedure implicitly alters the original likelihood model.

As described in **Chap. 5**, background weights are extracted from a fit to the $m(K\pi e^+e^-)$ distribution fixing the relative fractions of each background. The b-weighted angular distribution is used to set an angular function for the sum of the backgrounds. As shown in **Fig. 6.4**, the projections of the b-weighted distribution onto the three angular variables closely match those of the generated background samples. It should be noted that only the signal and misidentified components of the toy samples are generated under the assumption of factorisation between the mass and angular distributions. In contrast, the DSL, combinatorial, partially reconstructed, and J/ψ leakage components are bootstrapped from samples that inherently contain information about potential correlations. Therefore, the agreement observed in **Fig. 6.4** is not trivial. While this level of agreement is evident in the 1D projections, it is not guaranteed to hold in the full 3D angular space. Nevertheless, it is considered sufficient to justify the use of this approach. Once

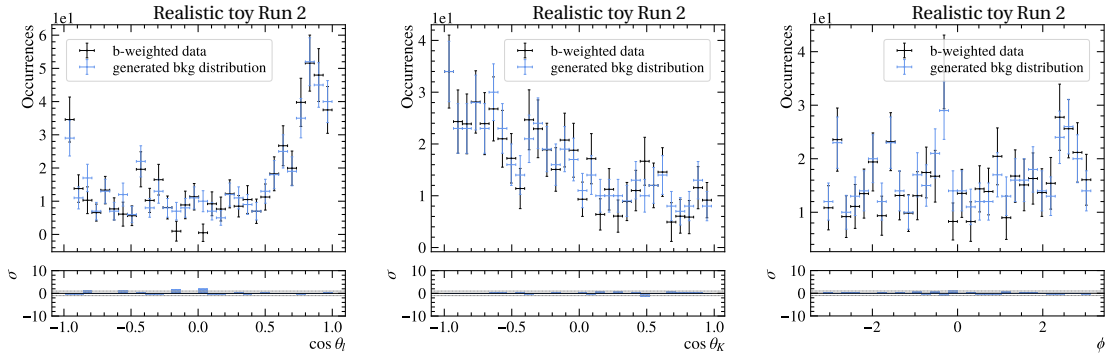


Figure 6.4: Comparison of the b-weighted angular distributions with the sum of the generated background samples. This is an example of a realistic toy emulating the Run 2 dataset in the full central q^2 region.

the angular background shape is determined, the four-dimensional fit is performed. An example of four-dimensional fit to the toy sample emulating central q^2 region for Run 1 and Run 2 datasets is shown in **Fig. 6.5**.

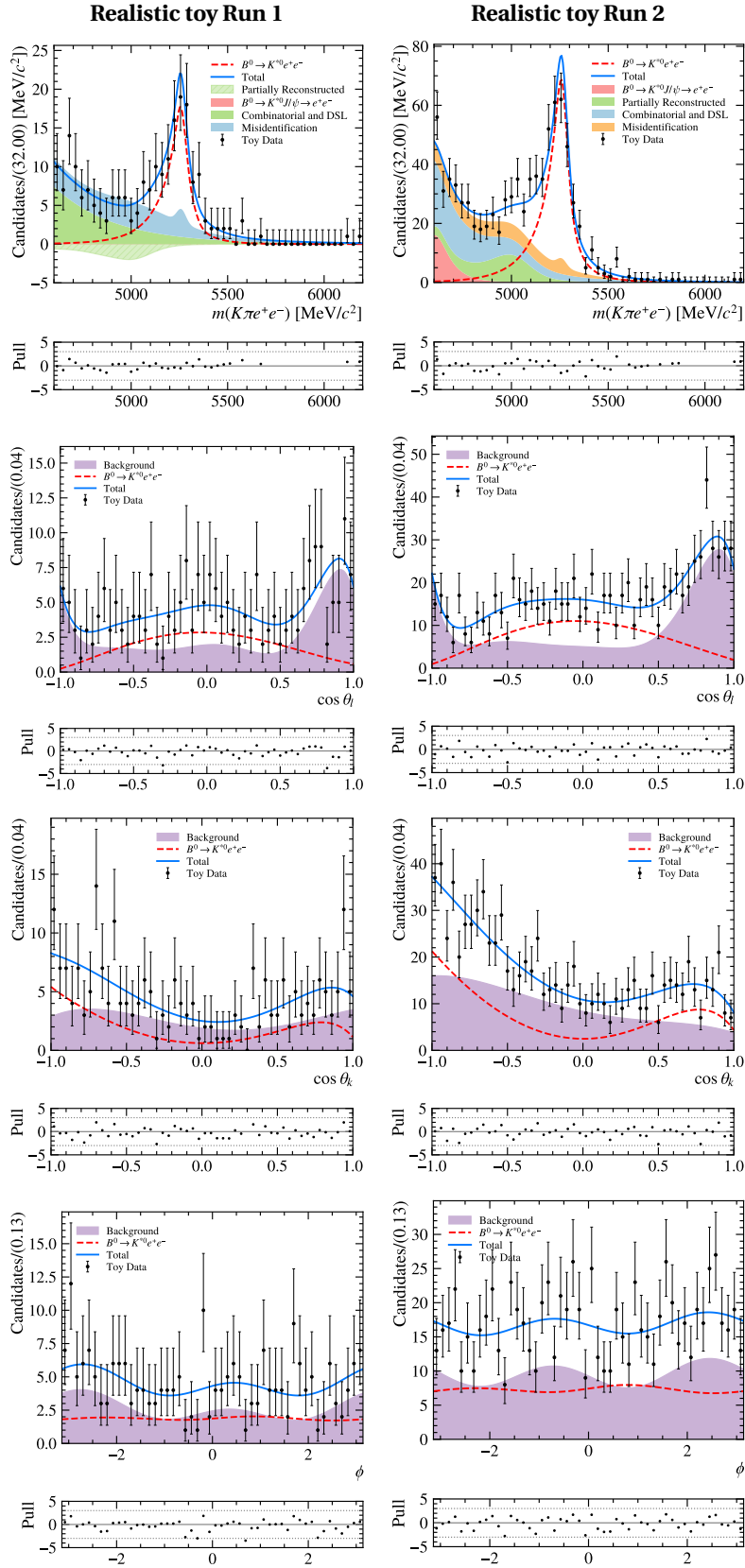


Figure 6.5: Realistic toy simulating the $B^0 \rightarrow K\pi e^+ e^-$ Run 1 (left column) and Run 2 (right column) datasets in the central q^2 region. From the top to the bottom, the distribution of the invariant mass, $\cos \theta_\ell$, $\cos \theta_K$, and ϕ are displayed. The blue lines overlaid on the data represent the best-fit PDFs.

The overview of the pull distributions obtained in the three q^2 region is shown in **Fig. 6.6**. The pull distribution for each of the angular parameters is shown for completeness on **App. D.3**, **Fig. D.4**, for the central q^2 , **Fig. D.5** for the central_{low} q^2 , and **Fig. D.6** for central_{high} q^2 regions. The values obtained for each angular parameter and their associated uncertainties are listed in **Tab. 6.1**, under the columns labelled *Sim. Run 1*, *Run 2*, which refer to a simultaneous fit to Run 1 and Run 2. The table also lists the values obtained from the fit to the Run 2 toys only for the central_{low} and central_{high} q^2 regions, as these serve as the reference for the evaluation of systematic uncertainties in smaller bins, as detailed in the next section.

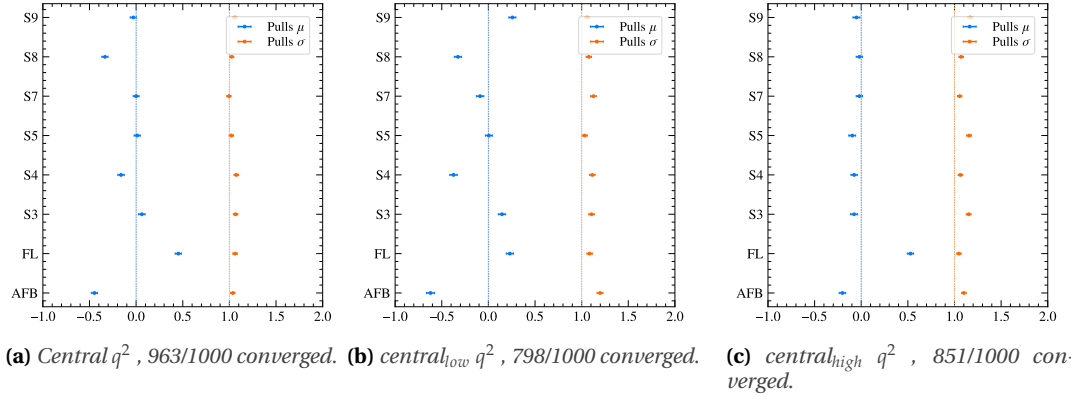


Figure 6.6: Mean and variance of the pulls distribution from 1000 toys with realistic stat.

Comparing these pulls to those with only the signal component (**Fig. 6.1**), one can observe that in the central_{low} q^2 region, the bias of approximately 0.5σ in A_{FB} remains visible, along with the bias in F_L . Additionally, new biases appear in S_8 , S_9 , and S_4 . In contrast, for the central_{high} q^2 region, most biases seem to be corrected, except for those in F_L , A_{FB} . Although the values of the Gaussian sigmas improve compared to the signal-only case, they systematically exceed one, indicating an underestimation of the uncertainties. Concerning the full central q^2 region, the realistic toys behave similarly to the central_{low} q^2 region, with visible biases in F_L , and A_{FB} , and to a lesser extent in S_4 and S_8 . Given that the signal-only toys were able to retrieve the generated values without biases, it is reasonable to attribute the biases arising when using realistic toys to imperfections of the fit strategy. More discussion on this in **Sec. 7.1**.

6.2 Systematic uncertainties

The measurement of the angular parameters is subject to both statistical uncertainties from the fit and systematic uncertainties arising from analysis design and data selection strategies. Key sources of systematic uncertainty include background modeling, acceptance parametrization, the impact of neglecting residual backgrounds and the BDT selection. In this dissertation, systematic uncertainties are primarily evaluated using toy sample generation, unless stated otherwise. Two main approaches are used to assess systematic effects:

1. Re-fitting with an alternative configuration:

- Generate a large number of toy samples, typically 1000, using the standard procedure (S).

Realistic toys fit values					
Param	central	central _{low}		central _{high}	
Fit	Sim. Run 1, Run 2	Sim. Run 1, Run 2	Run 2	Sim. Run 1, Run 2	Run 2
A_{FB}	-0.062 ± 0.041	-0.194 ± 0.068	-0.194 ± 0.078	0.055 ± 0.051	0.057 ± 0.055
F_L	0.744 ± 0.040	0.753 ± 0.066	0.764 ± 0.075	0.736 ± 0.052	0.743 ± 0.056
S_3	-0.004 ± 0.043	0.006 ± 0.070	-0.015 ± 0.080	-0.021 ± 0.057	-0.015 ± 0.063
S_4	-0.138 ± 0.067	-0.096 ± 0.107	-0.078 ± 0.121	-0.212 ± 0.085	-0.210 ± 0.093
S_5	-0.147 ± 0.055	0.009 ± 0.086	0.002 ± 0.097	-0.317 ± 0.071	-0.320 ± 0.078
S_7	0.009 ± 0.057	0.005 ± 0.086	-0.004 ± 0.099	0.011 ± 0.076	0.005 ± 0.083
S_8	-0.017 ± 0.068	-0.028 ± 0.104	-0.050 ± 0.120	-0.004 ± 0.088	0.001 ± 0.096
S_9	0.005 ± 0.043	0.020 ± 0.069	0.026 ± 0.079	0.005 ± 0.058	0.005 ± 0.063

Table 6.1: Average values of the angular parameters and their associated uncertainties obtained from the fits to 1000 realistic toys. For the two smaller q^2 regions, results are shown for both the simultaneous fit to the Run 1 and Run 2 toys, and for the fit to the Run 2 toys only.

- Fit each sample twice: once with the standard procedure (S), and a second time with the alternative configuration (A).

2. Generating toys with different configurations:

- Generate one set of toy samples using the standard configuration (S), and another using the alternative configuration (A). Usually, the difference between configurations affects either the bootstrapped background sample or the function used for generation. The components of the toys that do not change are taken from the standard configuration to isolate systematic uncertainties from statistical fluctuations.
- Both samples (S) and (A) are fitted using the standard configuration (S).
- Repeat this process 1000 times.

Each approach produces two sets of results: either (SS, SA) or (SS, AS). To isolate the effect of the alternative strategy from statistical fluctuations, the two results are subtracted: SS - SA and SS - AS. This difference is the deviation between the pull mean of each parameter, and it represents the systematic bias introduced when using A instead of S.

A detailed discussion of the main sources of systematic uncertainty and the methods used to quantify them is provided in the following subsections. Each subsection includes a table where uncertainties may accompany the reported values. When present, these uncertainties arise from the finite number of toy experiments that successfully converged. If no uncertainty is reported for a given systematic effect, it is because the estimate was obtained by performing two fits to real data and taking the difference between the resulting angular parameters.

The systematic uncertainties for the central q^2 region are computed using simultaneous fits to the Run 1 and Run 2 realistic toys, while the systematic uncertainties for the smaller bins are computed using Run 2 only.

6.2.1 Fit bias

The fit bias entries represent the biases observed during the validation of the fit strategy using pseudo-experiments (**Sec. 6.1**). In this thesis, these are treated as systematic uncertainties. The values are reported directly in the summary tables: **Tab. 6.9** for the central, **Tab. 6.10** for the central_{low}, and **Tab. 6.11** for central_{high} q^2 regions.

6.2.2 $B^+ \rightarrow K^+ e^+ e^-$ background

The $B^+ \rightarrow K^+ e^+ e^-$ can be a source of background in case a soft pion is associated with the kaon to form a K^{*0} , or if the kaon is misidentified as a pion and a soft kaon is associated with it. Both these cases reconstruct a fake K^{*0} candidate and consequently the invariant mass of the subsequent B^0 candidates peaks at larger values than the B^0 reconstructed mass. In the previous analysis [39] on this decay channel, the contamination from this background was efficiently removed requiring the maximum value between the three-body invariant masses $m(K e^+ e^-)$ and $m(\pi \rightarrow K e^+ e^-)$ to be smaller than $5100 \text{ MeV}/c^2$. However, this strategy has not been pursued as the cut introduces an artificial correlation between the variables $m(K \pi e^+ e^-)$ and $\cos \theta_K$, which breaks the assumption utilised in the fit that the mass and angular distributions factorise. The effect is shown in **Fig. 6.7**. The systematic associated to the B^+ veto was found to be larger than 50% of the statistical uncertainty on F_L and between 30% and 50% of the statistical uncertainty on P'_5 and P_2 [39]. The

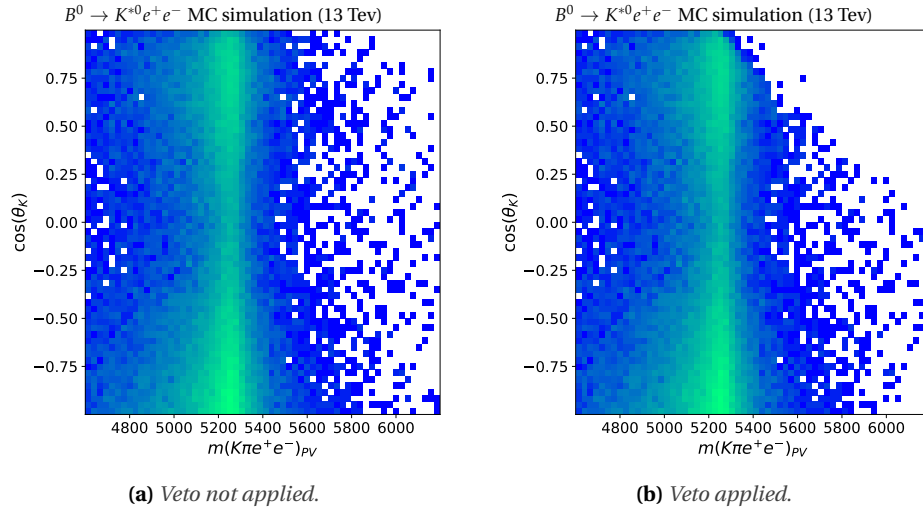


Figure 6.7: Effect of the B^+ veto ($\text{Max}[m(K e^+ e^-), m(\pi \rightarrow K e^+ e^-)] < 5100 \text{ MeV}/c^2$) on the signal $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation. This cut is not applied anymore.

choice of not applying this veto still comes with a systematic uncertainty, due to the fact that a background is neglected in the fit model. The expected number of $B^+ \rightarrow K^+ e^+ e^-$ in the total central q^2 regions of interest is estimated to be compatible with zero [123]. A thorough study has been performed on this systematic in the full central q^2 region, detailed in Ref. [123]. The approach in Ref. [123] is to generate toys with different configurations, standard and alternative and then fit both with the standard configuration. The *standard* (S) configuration consists in bootstrapping the reconstructed angular distributions from the $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation, while in the *alternative* configuration, the MC sample of $B^+ \rightarrow K^+ e^+ e^-$ reconstructed with a soft pion associated has been bootstrapped to simulate the contamination of B^+ backgrounds events and added to the simulated signal sample in the final fit region. Finally, the two types of toy, with and without the B^+ contamination, are both fitted with the standard configuration (S), namely using a signal-only PDF. This procedure is repeated for different B^+ contamination fractions. It has been found that with higher fraction of B^+ the parameter affected the most is F_L . This same angular parameter was the most affected by the veto in the previous analysis. Nevertheless, the systematic uncertainty associated to not applying a veto leads to an uncertainty of less than 50% of the statistical error even when the effect is overestimated. The highest fraction used in this

study is 10 B^+ events over 380 signal events. The choice of 10 events is based on the worst-case scenario obtained assuming that all events in the $m(Ke^+e^-)$ distribution in data located where the B^+ peak should appear are actually B^+ . This is clearly an upper limit for this systematic uncertainty. A more reasonable fraction is 0.1% which results in 0 expected events for 380 signal events, but could have positive statistical fluctuations. The value of this uncertainty for this analysis has been estimated using the same approach, but using realistic toys with 10 times the yield of each component (signal and backgrounds) and injecting 0.1% of the signal yield as B^+ background events. The same procedure is repeated in all three q^2 regions. The results are listed in **Tab. 6.2** compared also to those obtained in Ref. [123] with the same fraction. The values are always compatible with zero within the error due to the number of converged toys, therefore the systematics due to neglecting the B^+ backgrounds are reported as zero in the final tables: **Tab. 6.9**, **Tab. 6.10**, and **Tab. 6.11**.

Systematic uncertainties due to $B^+ \rightarrow K^+e^+e^-$ background ($\% \sigma$)					
Param. $\downarrow$	q^2 bin $\rightarrow$	Ref. [123]		This thesis	
		central	central	central _{low}	central _{high}
A_{FB}		4.5 ± 4.5	0.1 ± 4.8	0.1 ± 4.6	0.1 ± 4.6
F_L		2.1 ± 4.4	2.8 ± 4.8	1.7 ± 5.1	1.7 ± 5.1
S_3		1.7 ± 4.4	0.2 ± 4.6	0.0 ± 4.7	0.0 ± 4.7
S_4		0.0 ± 4.5	0.4 ± 4.5	0.1 ± 4.5	0.1 ± 4.5
S_5		1.5 ± 4.4	0.4 ± 4.5	0.4 ± 4.5	0.4 ± 4.5
S_7		0.2 ± 4.6	0.1 ± 4.4	0.1 ± 4.5	0.1 ± 4.5
S_8		0.7 ± 4.5	0.3 ± 4.5	0.2 ± 4.6	0.2 ± 4.6
S_9		0.8 ± 4.6	0.3 ± 4.6	0.0 ± 4.6	0.0 ± 4.6

Table 6.2: Systematic uncertainties associated with neglecting the $B^+ \rightarrow K^+e^+e^-$ background are expressed as a percentage of the fit uncertainty obtained from toy simulations. The values presented in this thesis are compared to those reported in Ref. [123], demonstrating consistency between the two results. Each value is quoted with an uncertainty, calculated as the square root of the sum of the squared errors on the means of the pull distributions.

6.2.3 Acceptance parameterisation

The fit to the simulated data used to compute the acceptance moments and parameterisation for Run 1 and Run 2 is presented in **Sec. 5.3.1**. The uncertainties from this fit are treated as systematic uncertainties on the angular parameters, reflecting the limited size of the MC samples used for acceptance determination. In the full central q^2 bin, it is computed as a weighted average between the Run 1 and Run 2 uncertainties obtained from the fit, using the expected yields of $B^0 \rightarrow K^{*0}J/\psi$ candidates for Run 1 and Run 2, taken from Ref. [95]. In the two smaller q^2 regions, the same systematic uncertainty is evaluated using Run 2 only, as for all the other systematic uncertainties concerning these bins. The values are reported in **Tab. 6.3** under the columns named *Size*. All values are expressed as fractions of the statistical error from the respective fit to the data: the simultaneous fit to the Run 1 and Run 2 datasets for the full central q^2 region, and the Run 2 dataset fit only for the smaller q^2 regions.

Another source of uncertainty arising from the acceptance parameterisation is due to the choice of the polynomial order used by the acceptance function. This systematic is quantified

using signal-only pseudo-experiments by increasing all polynomial orders by three, as it was also done in Ref. [39]. The results are listed in **Tab. 6.3** under the columns named *Poly order*.

Finally, the systematic uncertainty associated with the assumption of $\vec{S}_{\text{gen}} = \vec{S}_{\text{phys}}$ detailed in **Sec. 5.1.1**, is computed by repeating the fit changing the value of the physics parameters to those estimated from the fit itself, $\vec{S}_{\text{phys}} = \hat{\vec{S}}_{\text{phys}}$. Therefore, the acceptance moments are computed using **Eq. 5.12**, including the correction factor. The results are listed in **Tab. 6.3** under the columns named $\hat{S}_{\text{phys}}$.

The total systematic uncertainty associated with the acceptance function is computed as the square root of the quadratic sum of the three systematic uncertainties evaluated in this section, and listed in **Tab. 6.9**, **Tab. 6.10**, and **Tab. 6.11** for central, central_{low}, and central_{high} q^2 , respectively.

Systematic uncertainties due to the acceptance (% σ)									
q^2 bin	central			central _{low}			central _{high}		
Params	Size	Poly order	$\hat{S}_{\text{phys}}$	Size	Poly order	$\hat{S}_{\text{phys}}$	Size	Poly order	$\hat{S}_{\text{phys}}$
A_{FB}	7.7	1.2 ± 4.6	0.0	4.7	4.4 ± 6.2	10.8	7.0	1.5 ± 5.0	0.0
F_L	8.6	0.6 ± 4.5	0.0	6.2	2.7 ± 5.4	5.0	8.4	5.9 ± 5.0	1.3
S_3	9.4	0.3 ± 4.6	0.0	8.2	6.8 ± 6.0	2.4	10.3	6.4 ± 5.2	0.0
S_4	8.8	2.5 ± 4.6	1.0	5.8	5.1 ± 5.4	0.0	6.9	6.7 ± 5.5	1.4
S_5	9.8	1.0 ± 4.6	1.3	7.7	5.7 ± 5.0	0.0	8.4	3.1 ± 5.4	0.0
S_7	9.8	0.6 ± 4.7	0.0	7.6	9.9 ± 5.0	0.0	9.1	1.7 ± 4.9	0.9
S_8	8.9	0.5 ± 4.7	3.1	6.0	12.8 ± 5.5	6.9	8.1	4.7 ± 5.0	1.5
S_9	9.4	2.0 ± 4.6	0.0	6.5	11.7 ± 6.0	1.0	9.0	2.6 ± 5.1	0.0

Table 6.3: Systematic uncertainties associated with acceptance parameterisation. They are due to the sample size used to parametrise it, the polynomial order used for the acceptance function, and the angular parameter values used in the acceptance moments estimation.

6.2.4 DSL and combinatorial backgrounds fitting strategy

The fitting strategy adopted for these two backgrounds is to use one single exponential for both these background components, although it is known that the DSL background is expected to have a steeper slope, as discussed in **Sec. 3.4**. To assess a systematic uncertainty due to this strategy, the lepton-flavour-violating Run 2 data sample has been manipulated to force the distinction of the two components, following a strategy close to the one documented in Ref. [39].

First, the selection threshold of the first BDT is tightened from 0.96 to 0.99, enriching the sample in DSL candidates. The mass and angular distributions are then fitted using a four-dimensional PDF that assumes factorisation among all variables. Specifically, the fit employs an exponential function for the $m(K\pi e^+ e^-)$ distribution, Legendre polynomials of order 9 for $\cos \theta_\ell$, 4 for $\cos \theta_K$, and a Fourier series of order 4 for ϕ .

A combinatorial enriched sample is obtained by applying a cut $\cos \theta_\ell < 0.5$, removing most of the DSL peaking at large $\cos \theta_\ell$. This sample is then fitted using a four-dimensional PDF, under the assumption of factorisation. The mass distribution is modelled with an exponential function, a linear function with zero y-intercept describes $\cos \theta_\ell$, a Legendre polynomial of order 4 models $\cos \theta_K$, and a Fourier series of order 4 models ϕ .

Finally, these functional forms are used to fit the full dataset, yielding a composition of 67%

DSL and 33% of Combinatorial backgrounds, as reported in **Tab. 6.4**. The fits to the DSL-enriched, combinatorial-enriched, and total lepton-flavour-violating datasets are shown in **Fig. 6.8**.

Variable	Param	Value combinatorial	Value dsl
-	<i>fraction</i>	0.334 ± 0.014	-
$m(K\pi e\mu)$	λ	-0.0020 ± 0.0001	-0.0048 ± 0.0002
$\cos \theta_K$	$a1$	-0.292 ± 0.066	-0.349 ± 0.047
	$a2$	0.057 ± 0.085	-0.014 ± 0.061
	$a3$	-0.124 ± 0.104	-0.007 ± 0.073
	$a4$	-0.097 ± 0.114	-0.082 ± 0.083
ϕ	$c2$	0.002 ± 0.057	0.139 ± 0.040
	$c1$	0.041 ± 0.054	-0.000 ± 0.040
	$s2$	-0.025 ± 0.053	0.003 ± 0.039
	$s1$	-0.024 ± 0.055	-0.012 ± 0.038
$\cos \theta_\ell$	$a1$	-0.052 ± 0.171	2.618 ± 0.078
	$a2$	-	3.000 ± 0.002
	$a3$	-	1.651 ± 0.069
	$a4$	-	-0.057 ± 0.107
	$a5$	-	-1.041 ± 0.124
	$a6$	-	-1.604 ± 0.128
	$a7$	-	-1.180 ± 0.132
	$a8$	-	-0.639 ± 0.111
	$a9$	-	-0.408 ± 0.100

Table 6.4: Fit results to the Run 2 $B^0 \rightarrow K^{*0} e\mu$ data sample.

The systematic uncertainty is assessed by fitting twice the same toy sample with an alternative configuration. The toy samples are generated as realistic toys (**Sec. 6.1.2**) except that the combinatorial and DSL backgrounds are generated from the best-fit PDF detailed in **Tab. 6.4**. Each toy sample is then fitted twice using the following configurations:

- Alternative: fit using the same two exponential functions in mass that were used to generate the sample. Because the fitting strategy on the angular variables has only one PDF for all backgrounds, this affects mainly the mass distribution.
- Standard: fit using just one exponential component in mass, fixed to the value obtained by fitting the $m(K\pi e\mu)$ distribution of the $B^0 \rightarrow K^{*0} e\mu$ data sample with a single exponential function, as explained in **Chap. 5**.

The systematic uncertainties arising with the choice of fitting two components in the mass with one exponential only are shown in **Tab. 6.5**, under the name of *Fit strategy*. They are computed by performing the difference between the standard and alternative configurations of the pull means of 1000 toys on each angular parameter.

A second systematic is considered, arising from the fact that the DSL background breaks the mass-angles factorisation hypothesis. A way to test this assumption is to compare the standard configuration to an alternative where the DSL is generated assuming factorisation among the mass and the angles.

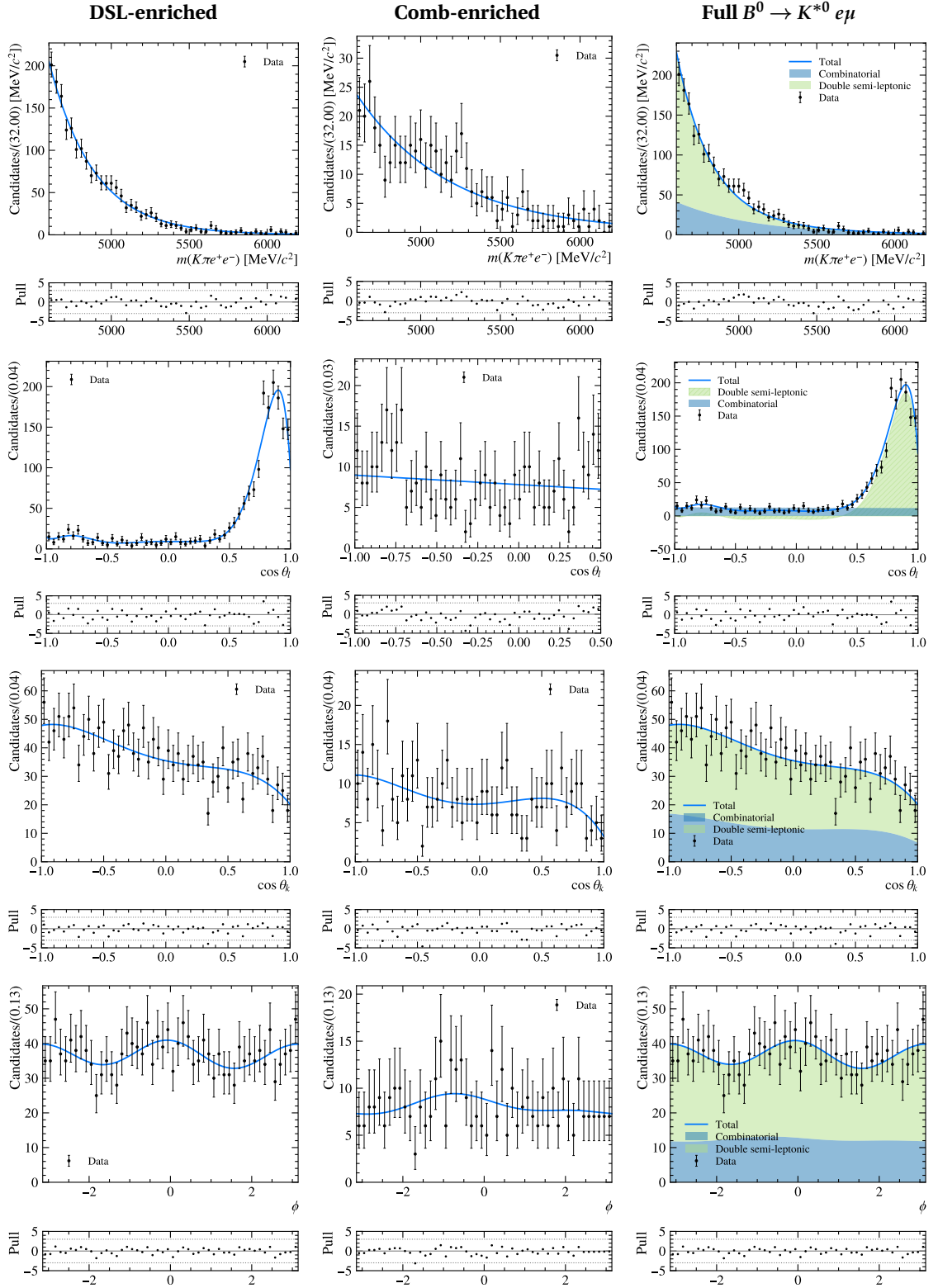


Figure 6.8: Starting from the left column: the DSL-enriched sample is shown with the best-fit PDF overlaid. The middle column presents the combinatorial-enriched sample, also with the best-fit PDF overlaid. Finally, the right column displays the full lepton-flavour-violating dataset, fitted with both components. Each row shows the projections for the different variables. The samples are the final selected Run 2 datasets with the additional discussed selection applied.

- **Standard:** Generate and fit the toy samples in the standard way. Therefore, generate realistic toys as explained in **Sec. 6.1.2**: the DSL and combinatorial background components are bootstrapped by the lepton-flavour-violating sample. The samples are then fitted as explained in **Chap. 5**: using one exponential function in the mass for both DSL and combinatorial, fixing the slope to the best fit value coming from the same sample that was bootstrapped.
- **Alternative:** Generate the DSL and Combinatorial components separately from the best-fit PDF listed in **Tab. 6.4**, hence assuming factorisation among all the variables, mass and 3 angles. The sample is then fitted in the standard way.

It is important to note that this systematic uncertainty is not independent of the one studied previously, as it also includes the systematic effect due to the fitting strategy (which is found to be negligible for all the coefficients and q^2 regions). Therefore, in the systematic uncertainties summary tables, only this last uncertainty is reported for all angular parameters (**Tab. 6.9**, **Tab. 6.10**, and **Tab. 6.11**). The result is shown in **Tab. 6.5** under the name of *Correlation* in comparison to the previous. These values have been computed by performing a simultaneous fit to Run 1 and Run 2 toys for the central q^2 region. For the Run 1 toys, the same functional form used for Run 2 was employed to generate the combinatorial and DSL background components. This choice is justified by the fact that only a few events survived the selection in the Run 1 lepton-flavour-violating data sample used to extract the DSL and combinatorial PDFs. Although affected by much larger uncertainties, the fit with the two components yields fractions for the DSL and combinatorial backgrounds that are consistent with those obtained in Run 2. Additionally, the functional forms used to generate the combinatorial and DSL background components have been kept the same across all three bins. This is because, in order to obtain a larger number of events to fit in the lepton-flavour-violating data sample, the full central q^2 region was considered.

Systematic uncertainties due to DSL and Combinatorial fitting strategy ($\% \sigma$)					
q^2 bin	central	central _{low}		central _{high}	
Params	Correlation	Fit strategy	Correlation	Fit strategy	Correlation
A_{FB}	11.9 ± 4.7	5.4 ± 5.2	5.2 ± 5.2	2.4 ± 4.9	22.7 ± 4.9
F_L	11.9 ± 4.8	8.9 ± 5.2	17.8 ± 5.1	5.1 ± 4.6	18.0 ± 4.7
S_3	20.0 ± 4.8	1.6 ± 5.4	25.5 ± 5.4	0.0 ± 5.1	9.2 ± 5.1
S_4	24.7 ± 4.9	1.3 ± 4.9	15.1 ± 5.0	1.8 ± 4.8	12.1 ± 4.8
S_5	11.9 ± 4.7	0.1 ± 4.8	6.2 ± 4.9	2.5 ± 5.1	8.2 ± 5.2
S_7	0.5 ± 4.5	1.8 ± 4.9	1.3 ± 5.0	0.8 ± 4.8	4.4 ± 4.7
S_8	17.1 ± 4.7	1.0 ± 4.9	30.9 ± 4.9	0.0 ± 4.8	1.5 ± 4.8
S_9	2.4 ± 4.9	1.1 ± 5.3	5.0 ± 5.3	0.1 ± 5.3	0.1 ± 5.3

Table 6.5: Systematic uncertainties associated to the choice of modeling two backgrounds, combinatorial and double semi-leptonic background, using just one exponential function for the invariant mass distribution, and no correlation between mass and angles within the current fit strategy. They are listed for each angular parameter and expressed in percentage of statistical error.

6.2.5 Misidentified background fitting strategy

The largest systematic uncertainty for the misidentified background arises from the choice of the functional form for the invariant mass distribution used in the fit. To evaluate this effect, two

strategies have been adopted:

- Generate toy samples as described in **Sec. 6.1.2**. Fit each toy sample with two different configurations:
 - Standard: fit the toy sample in the standard way described in **Chap. 5**.
 - Alternative: fit the toy sample in the standard way. However, all the parameters concerning the misidentified background mass PDF (**Eq. 5.23**) are left free in the four-dimensional fit, except for the means of the two Gaussian functions. This choice is physically motivated, as the mean positions are determined by the underlying processes.

The systematic uncertainty is then obtained as the difference between the means of the resulting pulls. The final result is reported in **Tab. 6.6** under the label *Float* for both the q^2 regions under study.

- Generate toy samples with alternative functional forms. In this case, the toys change the PDF used to generate the misidentified background, while the fitting function is kept the same. In particular the following scenarios have been investigated for the central and central_{low} q^2 regions where the PDF used to generate and fit in the standard scenario (S) is the sum of one exponential (E) and two Gaussian functions (G1 which is the Gaussian with lowest energy, and G2 which is the Gaussian with mean closer to the B^0 mass peak):
 - Alternative (A1) is the case where the misidentified background is generated using only the exponential (E) and the Gaussian at lower energy (G1). Each toy sample is then fitted with the standard function.
 - Alternative (A2) is the case where the misidentified background is generated using only the exponential (E) and the Gaussian at higher energy (G2). Each toy sample is then fitted with the standard function.

The scenarios are different in central_{high} q^2 region where the PDF used to generate and fit in the standard scenario (S) is the sum of one exponential (E) and just one Gaussian at lower energy (G1). In this case, just one alternative scenario (A) is taken into account:

- the misidentified background is generated as in the standard configuration adding a small fraction of Gaussian at higher energy, to cover for the possibility of an overlooked peaking background. The fraction has been chosen to be 5% of the total misidentified background.

The systematic error for all these scenarios is computed as the difference of the mean bias between the configurations S and A. The results are reported in **Tab. 6.6**.

The fact that the error on the *Float* systematic is larger than those on other systematics is due to the convergence rate of toys, which was 50%, when normally this varies between 90-100%. Due to the low convergence rate of the toys for the *Float* systematic, and since all these values relate to the shape of the misidentified background, only the following scenarios are considered as the total systematic uncertainty associated with the misidentified background shape: A1 for the central and central_{low} q^2 regions, and A for the central_{high} q^2 region. These values are mainly related to the peaking component of the background, which is considered the most dangerous

if ignored or significantly over- or under-estimated. The total results are reported in **Tab. 6.9**, **Tab. 6.10**, and **Tab. 6.11**.

Misidentified background systematics uncertainties ($\% \sigma$)								
q^2 bin	central			central _{low}			central _{high}	
Sys	Float	A1	A2	Float	A1	A2	Float	A
A_{FB}	19.5 ± 7.3	11.9 ± 4.7	19.3 ± 4.7	15.5 ± 6.2	23.9 ± 5.3	11.0 ± 5.2	16.1 ± 5.3	2.1 ± 4.9
F_L	35.0 ± 6.8	14.5 ± 4.9	34.5 ± 4.8	20.1 ± 6.5	21.9 ± 5.2	6.5 ± 5.2	32.7 ± 5.0	1.8 ± 4.7
S_3	1.0 ± 8.1	0.9 ± 4.8	1.6 ± 4.8	8.5 ± 6.7	2.4 ± 5.3	1.0 ± 5.3	1.6 ± 5.5	1.2 ± 5.1
S_4	3.2 ± 7.9	2.5 ± 4.9	9.8 ± 4.9	2.5 ± 6.5	3.7 ± 5.2	3.0 ± 5.2	10.0 ± 5.1	5.1 ± 4.8
S_5	3.3 ± 7.5	4.0 ± 4.7	11.2 ± 4.6	5.2 ± 6.3	0.3 ± 5.0	2.2 ± 5.0	11.8 ± 5.5	2.3 ± 5.2
S_7	5.5 ± 7.2	1.1 ± 4.6	0.8 ± 4.5	0.1 ± 6.3	1.6 ± 5.0	2.9 ± 5.0	0.7 ± 5.1	2.0 ± 4.7
S_8	3.0 ± 7.3	0.1 ± 4.7	3.7 ± 4.7	6.6 ± 6.3	0.5 ± 5.0	3.3 ± 4.9	0.6 ± 5.2	2.1 ± 4.8
S_9	2.8 ± 7.8	2.4 ± 4.8	4.0 ± 4.8	1.4 ± 6.7	2.7 ± 5.3	4.2 ± 5.3	1.2 ± 5.6	2.0 ± 5.3

Table 6.6: Systematic uncertainties associated with the modelling choice of the misidentified background, within the current fit strategy. They are listed for each angular parameter and expressed in percentage of statistical error. The associated error is purely related to the number of converging toys.

6.2.6 BDT

The systematic uncertainty associated with the BDT selection arises from the use of MC simulations to model fit components, particularly the signal, the partially reconstructed background, and the $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ leakage background. As discussed in **Chap. 4**, some variables show discrepancies between data and MC, requiring a corresponding systematic uncertainty. Furthermore, the selection on the BDTs on the signal MC samples shows local variations that alter the mass and angular distributions' shape (**Fig. 4.4** and **Fig. 4.8**). By using MC simulation to model signal and backgrounds, the underlying assumption is that the selection on the BDTs affects simulation and data in the same way. The impact of unaccounted discrepancies between data and MC simulations arising from the BDT selection is evaluated separately for the signal, the partially reconstructed MC simulation, and the $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ MC simulation in the three q^2 regions under analysis.

- **Signal:** The systematic is evaluated using realistic toys as follows:
 - Standard: toy samples are generated using the standard approach, with the signal component bootstrapped from the signal MC sample, applying the full selection. Each toy is then fitted using the standard approach.
 - Alternative: toy samples are generated using the standard approach, except the signal component is bootstrapped from the signal MC simulation sample without selection on the two BDT outputs. Each toy is then fitted using the standard approach.

The signal component should not be significantly affected by the BDT selection. This evaluation aims to estimate the impact of a slight difference in the signal shape between data and simulation.

- **Partially Reconstructed (PR) and $B^0 \rightarrow K^{*0} J/\psi$ leakage:** This particular systematic has been evaluated directly in data, while keeping the result blind (**Sec. 6.3**) by computing the difference between the angular parameters in two scenarios:

- Standard: The fit to the data is performed in the standard way.
- Alternative: The fit to the data is performed by changing the model of the partially reconstructed ($B^0 \rightarrow K^{*0} J/\psi$ leakage) used to fit. The samples are modified by removing the two BDT cuts, and their $m(K\pi e^+ e^-)$ distributions are modelled using a KDE derived directly from the modified MC simulation samples.

The systematic uncertainty for these two components is evaluated directly in the data to allow the fit to adjust to the new background shape. This approach is preferred because the background shape is expected to change with the selection. Using toy samples would enforce an incorrect configuration and could strongly overestimate the effect.

The results are listed in **Tab. 6.7**, highlighting the effect of each component. The systematic uncertainties are summed in quadrature and used as the total systematic uncertainty associated with the BDT selection (**Tab. 6.9**, **Tab. 6.10**, and **Tab. 6.11**).

Systematic uncertainties due to BDTs selection ($\% \sigma$)								
q^2 bin	central			central_{low}		central_{high}		
Parameters	Signal	PR	J/ψ	Signal	PR	Signal	PR	J/ψ
A_{FB}	20.3 ± 4.7	16.8	3.0	17.3 ± 5.5	23.6	24.5 ± 4.9	6.1	3.8
F_L	3.1 ± 4.7	13.2	3.4	14.8 ± 5.3	9.7	10.7 ± 5.0	10.4	4.0
S_3	8.9 ± 4.8	0.2	0.9	1.0 ± 5.3	3.3	2.5 ± 5.2	1.3	0.0
S_4	16.7 ± 4.8	0.5	2.2	2.1 ± 5.0	0.8	8.5 ± 5.0	5.7	1.1
S_5	5.2 ± 4.7	2.1	1.1	1.5 ± 4.9	4.5	13.3 ± 5.1	2.5	2.1
S_7	1.2 ± 4.6	2.7	1.0	5.3 ± 5.1	0.6	0.1 ± 4.9	0.42	1.7
S_8	13.7 ± 4.7	0.3	1.8	6.7 ± 5.1	0.3	1.4 ± 5.0	0.20	0.0
S_9	12.4 ± 4.7	4.3	1.2	9.9 ± 5.3	3.8	8.9 ± 5.2	1.7	0.6

Table 6.7: Systematic uncertainties associated to the selection on the BDTs output. They are listed for each angular parameter and expressed in percentage of statistical error. The associated error is purely related to the number of converging toys.

6.2.7 Robustness checks

Simulation corrections

MC simulations are generated as described in **Sec. 3.5.1** and corrected for possible discrepancies with the collected data. The uncertainty arising from the correction weights is estimated computing difference of the mean biases among these two scenarios:

- Standard: Generate toys as explained in **Sec. 6.1.2**, and fit accordingly to **Chap. 5**.
- Alternative: Generate toys as explained in **Sec. 6.1.2**, bootstrapping all the MC samples (also the signal), omitting the correction weights. Fit accordingly to **Chap. 5**.

This uncertainty combines the effects of individual correction weights. However, it is not included in the final systematic uncertainties tables, as it cannot be interpreted as an upper limit. Omitting the MC corrections weights does not guarantee the maximum possible variation. Instead, the results presented in **Tab. 6.8** serve as a robustness check, demonstrating that the angular parameters are not significantly affected by the MC weights. The final systematic uncertainty

may still be different from the values reported here and will be properly assessed in the future steps of this analysis.

Robustness check concerning MC corrections		
	central _{low} q^2	central _{high} q^2
A_{FB}	5.9 ± 5.5	6.7 ± 4.9
F_L	9.5 ± 5.3	4.6 ± 4.9
S_3	1.7 ± 5.5	4.1 ± 5.2
S_4	1.3 ± 5.1	1.9 ± 5.0
S_5	3.6 ± 4.9	5.0 ± 5.0
S_7	3.6 ± 5.1	7.0 ± 4.7
S_8	0.0 ± 5.1	0.1 ± 4.9
S_9	10.4 ± 5.4	11.0 ± 5.0

Table 6.8: Systematic check in the impact of the MC simulation corrections on the angular parameters. They are listed for each angular parameter and expressed in percentage of statistical error. The error associated is purely related to the number of toys thrown.

6.2.8 Summary of the systematic uncertainties

The results of these studies are summarised in **Tab. 6.9** for the central, **Tab. 6.10** for the central_{low}, and **Tab. 6.11** for central_{high} q^2 regions. The values are expressed as a percentage of their relative statistical uncertainties. The average statistical uncertainties to which the systematic values are normalised are provided in **Tab. 6.1**. For the central_{low} and central_{high} q^2 regions, the reference statistical values are those listed under the column labelled *Run 2*. A value of 0.0 is written in the tables if the magnitude of the systematic uncertainty is smaller than its associated statistical uncertainty.

Systematic uncertainties central q^2 [1.1, 6.0] GeV ² /c ⁴ (% σ)							
	Fit bias	B^+	BDT	acc	misid	DSL+Comb	TOT
A_{FB}	44.7	0.0	26.5	7.7	11.9	11.9	55.2
F_L	45.2	0.0	13.6	8.6	14.5	11.9	51.5
S_3	6.2	0.0	8.9	9.4	0.0	20.0	24.6
S_4	16.2	0.0	16.8	8.9	0.0	24.7	35.1
S_5	0.0	0.0	5.7	9.9	0.0	11.9	16.5
S_7	0.0	0.0	2.9	9.8	0.0	0.0	10.2
S_8	33.4	0.0	13.8	9.4	0.0	17.1	41.0
S_9	0.0	0.0	13.2	9.4	0.0	0.0	16.2

Table 6.9: Preliminary systematic uncertainties for the central q^2 bin, listed by source and by affected angular parameter. The total systematic uncertainty is obtained by summing in quadrature the contributions from all sources, assuming they are uncorrelated.

Systematic uncertainties central_{low} q^2 [1.10, 3.65] GeV ² /c ⁴ (% σ)							
	Fit bias	B^+	BDT	acc	misid	DSL+Comb	TOT
A_{FB}	54.5	0.0	29.3	11.8	23.9	0.0	67.4
F_L	32.5	0.0	17.7	7.9	21.9	17.8	47.2
S_3	13.7	0.0	3.3	10.9	0.0	25.5	31.1
S_4	18.3	0.0	0.8	5.8	0.0	15.1	24.4
S_5	5.4	0.0	4.5	9.6	0.0	6.2	13.4
S_7	17.0	0.0	5.3	12.5	0.0	0.0	21.8
S_8	47.0	0.0	6.7	15.7	0.0	30.9	58.8
S_9	31.1	0.0	10.6	13.4	0.0	0.0	35.5

Table 6.10: Preliminary systematic uncertainties for the central_{low} q^2 bin, listed by source and by affected angular parameter. The total systematic uncertainty is obtained by summing in quadrature the contributions from all sources, assuming they are uncorrelated.

Systematic uncertainties central_{high} q^2 [3.65, 6.00] GeV ² /c ⁴ (% σ)							
	Fit bias	B^+	BDT	acc	misid	DSL+Comb	TOT
A_{FB}	20.2	0.0	25.5	7.0	0.0	22.7	40.3
F_L	52.3	0.0	15.4	10.3	0.0	18.0	58.3
S_3	7.7	0.0	1.3	12.1	0.0	9.2	17.1
S_4	7.6	0.0	10.3	9.7	5.1	12.1	20.7
S_5	9.7	0.0	13.7	8.4	0.0	8.2	20.5
S_7	0.0	0.0	1.7	9.1	0.0	0.0	9.2
S_8	0.0	0.0	0.2	8.2	0.0	0.0	8.2
S_9	5.1	0.0	9.1	9.0	0.0	0.0	13.8

Table 6.11: Preliminary systematic uncertainties for the central_{high} q^2 bin, listed by source and by affected angular parameter. The total systematic uncertainty is obtained by summing in quadrature the contributions from all sources, assuming they are uncorrelated.

6.3 Blinded fit to the collected data

The PDF parameters S_i are extracted using a maximum likelihood fit to the invariant mass and decay angles with the strategy described in **Chap. 5**, simultaneously between the Run 1 and Run 2 datasets. A blinding procedure is applied to the angular parameters S_i to prevent unintended biases when analysing data, which consists of applying an unknown offset to the parameters of interest. The value of this offset is derived from a character string, ensuring that the blinding offset is consistent but unknown. The blinded values are reported as XX. The statistical uncertainties on the angular parameters given by the fit to the data are listed in **Tab. 6.12** under the name *Stat. Err.*. However, for the two smaller central q^2 bins these values need to be corrected, as the pseudo-experiments used to validate the fit show pull distributions with widths larger than one (**Sec. 6.1.2, Fig. 6.6**), which is an indicator of the fit underestimating the statistical uncertainties. The width values are listed in **Tab. 6.12** under the name *Corr.(σ)*. To correct for this effect, the statistical uncertainties are multiplied by the σ correction. The final statistical uncertainties are listed in **Tab. 6.12** under the name *Corr. Err.*, which stands for “corrected error”. The final statistical uncertainties for each q^2 region are highlighted with a bold font. The complete list of free parameters in the fit is shown in **Tab. D.1** for all the q^2 regions under analysis.

q^2 bin $\rightarrow$	central		central_{low}		central_{high}		
Parameters	Stat. Err	Stat.Err	Corr.(σ)	Corr. Err	Stat.Err	Corr.(σ)	Corr. Err
A_{FB}	XX$\pm$0.043	0.079	1.20	XX$\pm$0.095	0.059	1.10	XX$\pm$0.065
F_L	XX$\pm$0.039	0.066	1.08	XX$\pm$0.072	0.054	1.10	XX$\pm$0.060
S_3	XX$\pm$0.037	0.058	1.11	XX$\pm$0.064	0.051	1.07	XX$\pm$0.055
S_4	XX$\pm$0.071	0.115	1.11	XX$\pm$0.128	0.101	1.13	XX$\pm$0.114
S_5	XX$\pm$0.053	0.079	1.03	XX$\pm$0.081	0.079	1.11	XX$\pm$0.088
S_7	XX$\pm$0.055	0.084	1.13	XX$\pm$0.094	0.081	1.04	XX$\pm$0.085
S_8	XX$\pm$0.068	0.113	1.08	XX$\pm$0.122	0.094	1.05	XX$\pm$0.099
S_9	XX$\pm$0.042	0.068	1.06	XX$\pm$0.072	0.061	1.08	XX$\pm$0.067

Table 6.12: Columns named *Stat. Err.* show the statistical uncertainties of the blinded simultaneous fit. The columns named *Corr.(σ)* show the pull distributions width of the toys in **Sec. 6.1.2**. Columns named *Corr. Err.* show the corrected uncertainties. Columns in bold report the final statistical uncertainties associated with each parameter in the three q^2 regions under analysis. For the sake of the reading, only the uncertainty of the angular parameter is reported. The table with all the parameters minimised in the fit is reported in **App. D.3, Tab. D.1**

The plots in **Fig. 6.9**, **Fig. 6.10**, and **Fig. 6.11** show the unblinded mass distribution in data overlaid with the best-fit PDF and the residuals among the data and the best-fit PDF for the blinded angular distributions, for central, central_{low}, and central_{high} q^2 regions, respectively.

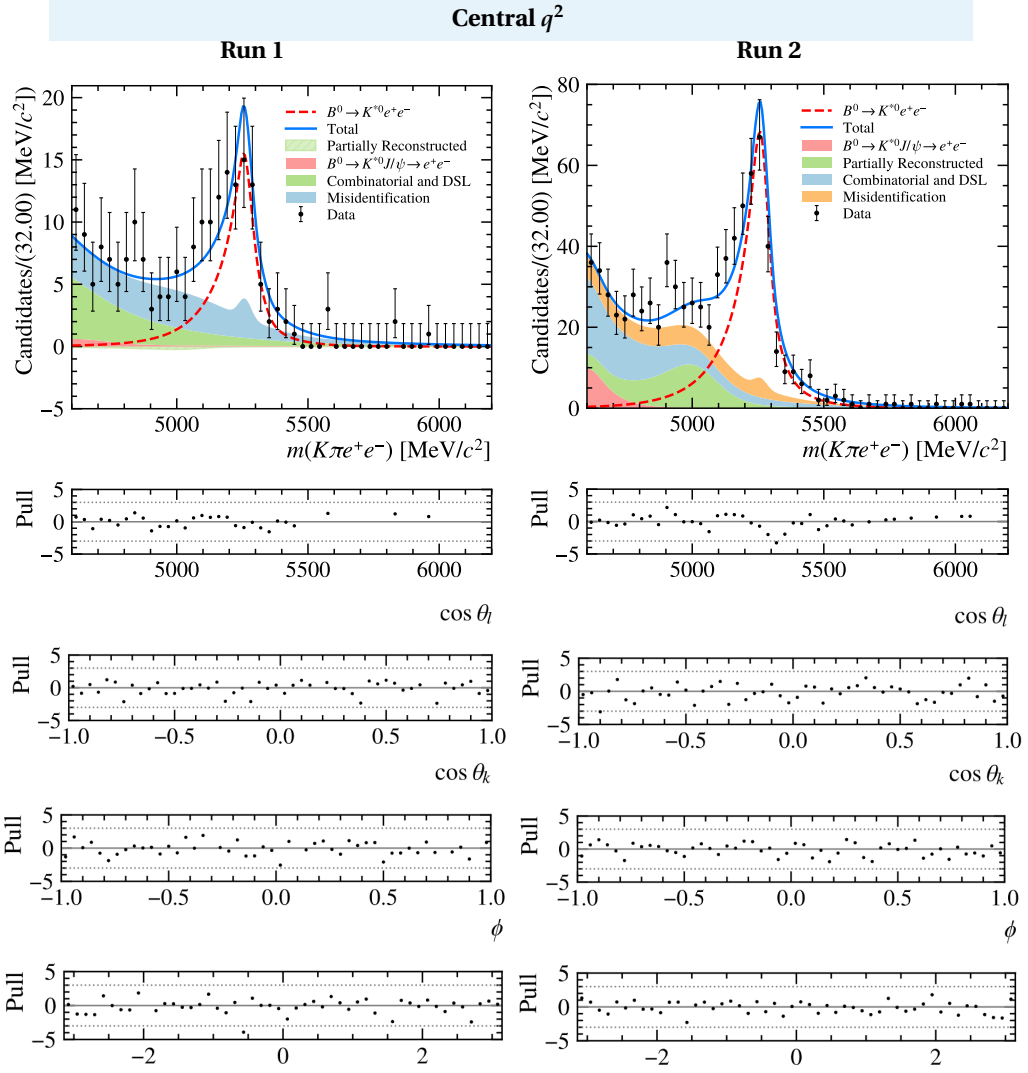


Figure 6.9: The top row displays the invariant mass dataset and total PDF with parameters set at best fit values. The remaining rows show only the residuals among the data and the total PDF. The left column corresponds to Run 1, while the right column corresponds to Run 2. The results pertain to the central q^2 region.

The plots in **Fig. 6.12**, **Fig. 6.13**, and **Fig. 6.14** show in black the 1D projections onto the angular variables of the b-weighted data for the central, central_{low}, and central_{high} q^2 regions, respectively. The background weights are derived from the second fit to the invariant mass, $m(K\pi e^+ e^-)$, where the relative fractions of the various background components are fixed, as explained in **Sec. 5.2**. The blue curve represents the best-fit background PDF. The first row refers to the Run 1 dataset, while the second row to the Run 2.

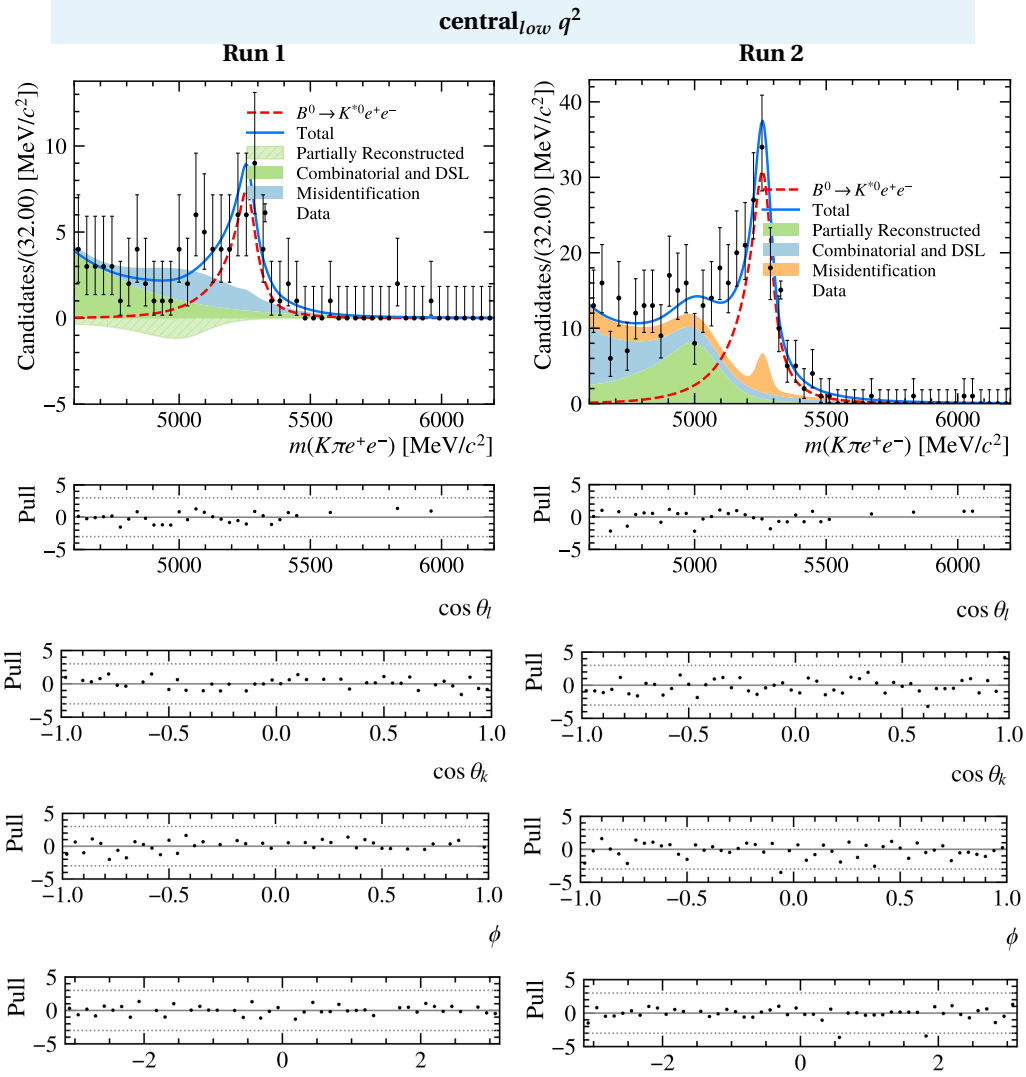


Figure 6.10: The top row displays the invariant mass dataset and total PDF with parameters set at best fit values. The remaining rows show only the residuals among the data and the total PDF. The left column corresponds to Run 1, while the right column corresponds to Run 2. The results pertain to the central_{low} q^2 region.

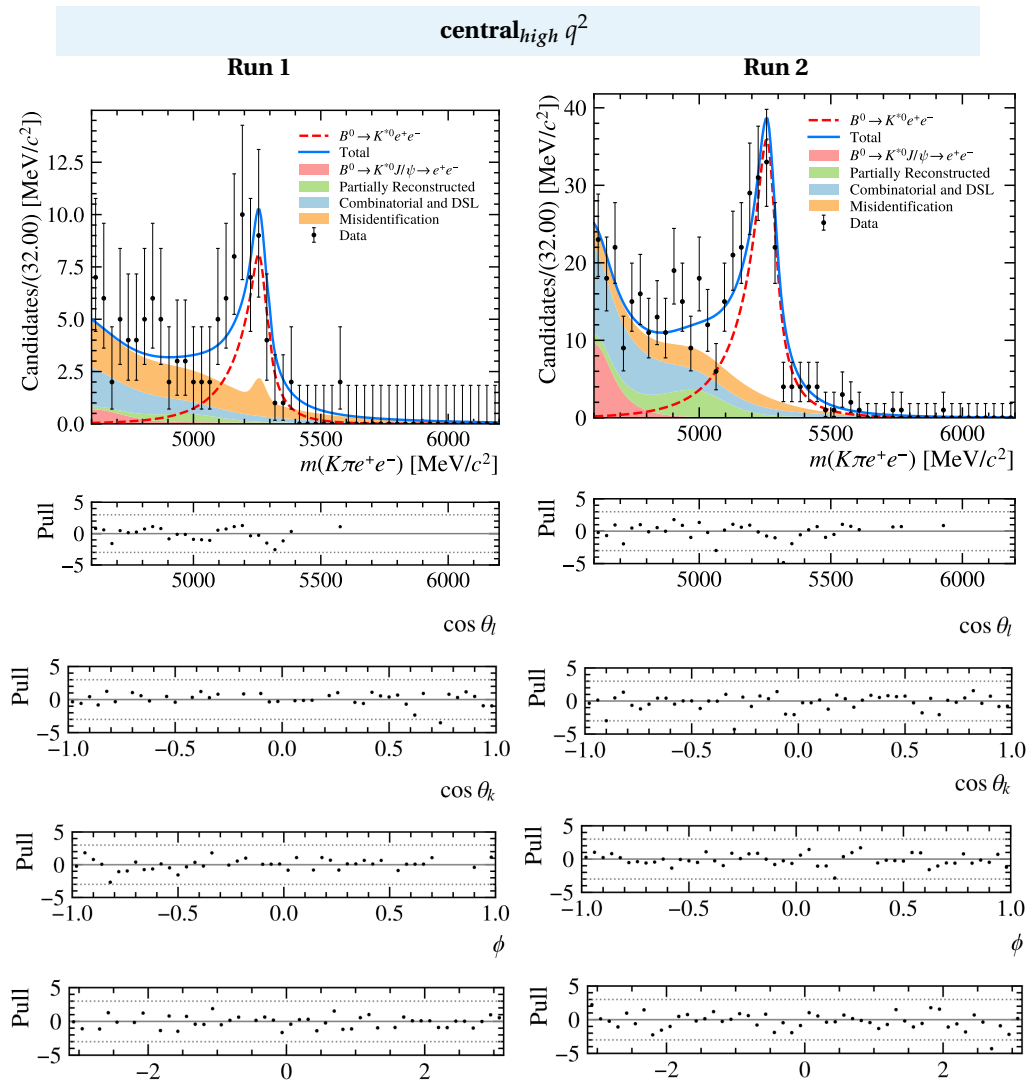


Figure 6.11: The top row displays the invariant mass dataset and total PDF with parameters set at best-fit values. The remaining rows show only the residuals among the data and the total PDF. The left column corresponds to Run 1, while the right column corresponds to Run 2. The results pertain to the central_{high} q^2 region.

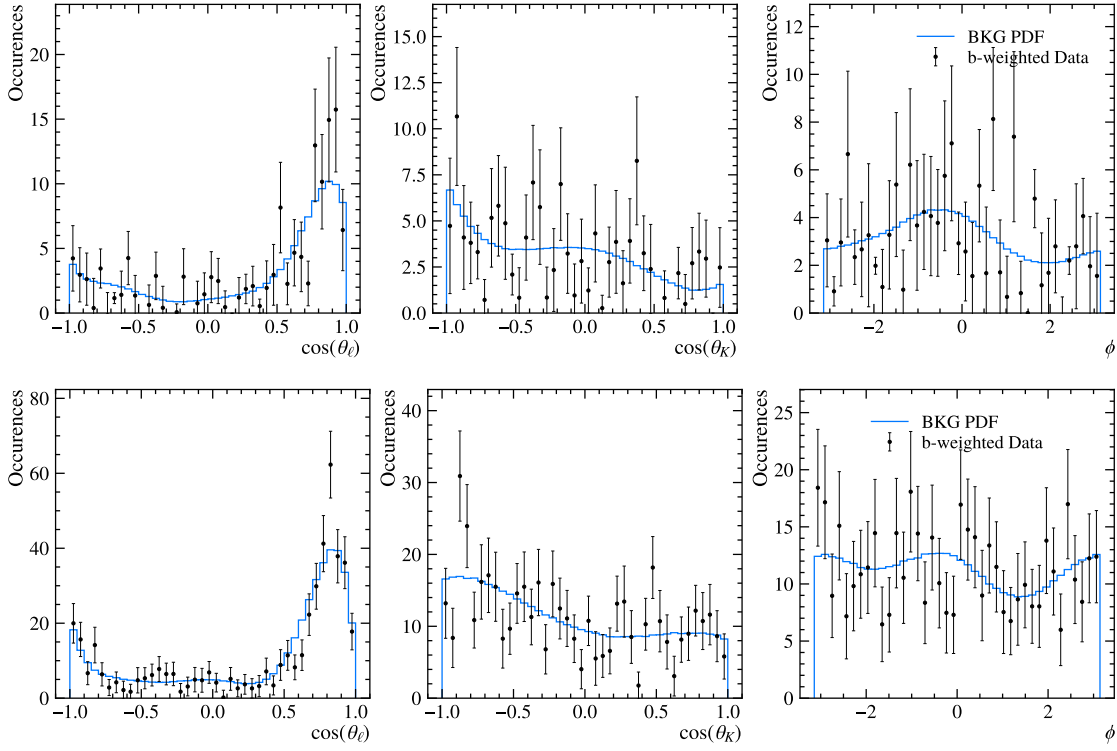


Figure 6.12: The b-weights for Run 1 (above) and Run 2 (below) extracted from the $m(K\pi e^+e^-)$ fit to the data are compared with best fit background PDF in central q^2 bin.

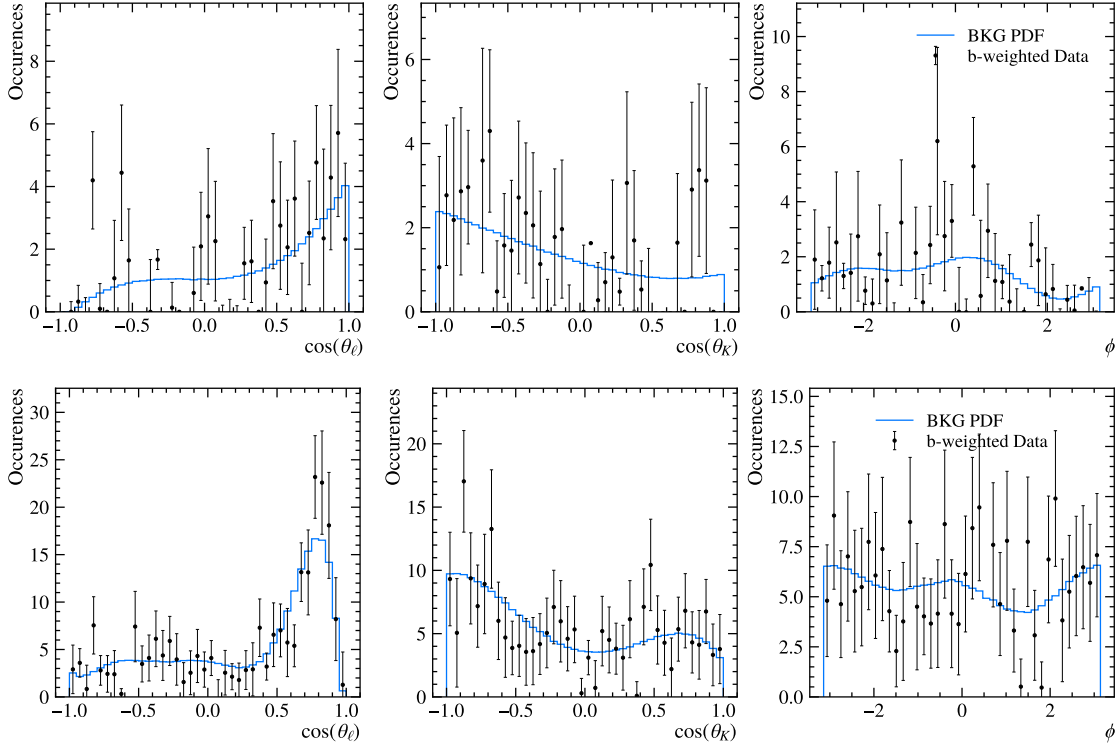


Figure 6.13: The b-weights for Run 1 (above) and Run 2 (below) extracted from the $m(K\pi e^+e^-)$ fit to the data are compared with best fit background PDF in central_{low} q^2 bin.

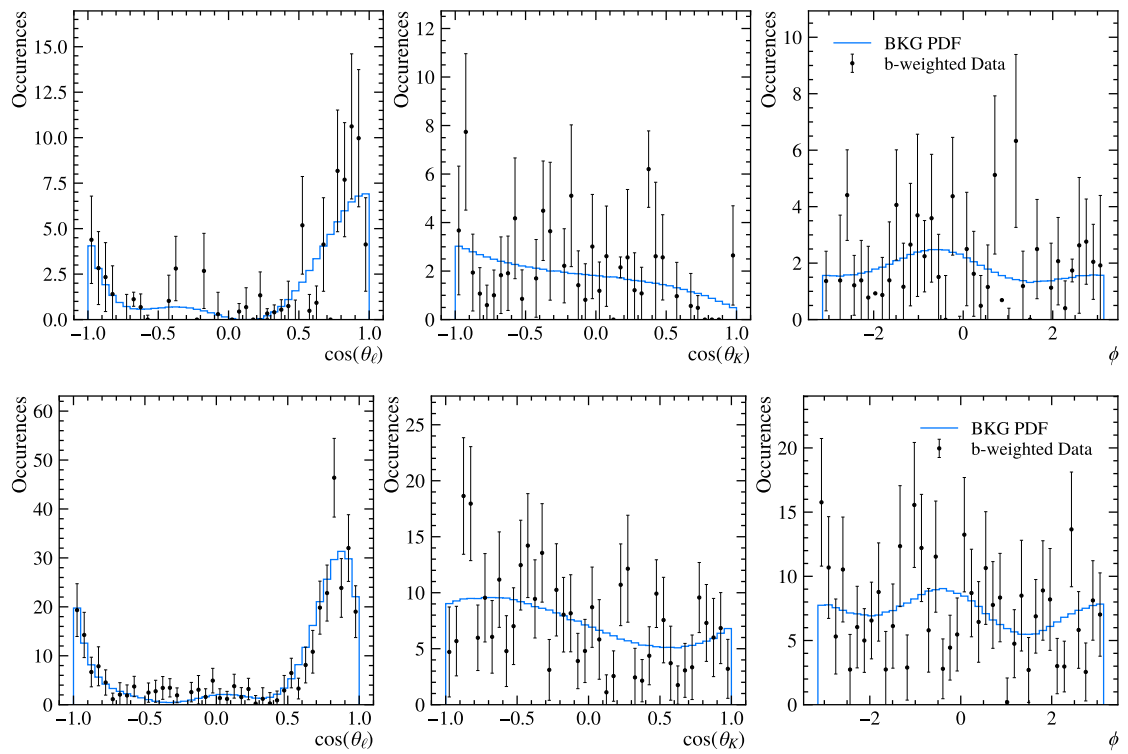


Figure 6.14: The b-weights for Run 1(above) and Run 2(below) extracted from the $m(K\pi e^+e^-)$ fit to the data are compared with best fit background PDF in central_{high} q^2 bin.

7 | Discussion and outlooks

7.1 Sensitivity comparison between analyses

This thesis described the $B^0 \rightarrow K^{*0}e^+e^-$ angular analysis, focusing on the selection applied to reduce the backgrounds, and on the analysis strategy to measure the angular parameters describing the decay in three kinematic regions: a central q^2 region, spanning $[1.1, 6.0] \text{ GeV}^2/c^4$, a central_{low}, spanning $[1.10, 3.65] \text{ GeV}^2/c^4$, and a central_{high}, spanning $[3.65, 6.00] \text{ GeV}^2/c^4$.

This analysis differs from the previous one [39] in several respects. In particular, some selection cuts have been re-optimised to retain a higher signal efficiency.

- The vetoes introduced in **Sec. 3.3.3** on background arising from decays involving Λ baryons, D , and ϕ mesons are applied only to candidates in which the electron has opposite charge to the kaon and is not associated with any photon cluster. This strategy results in a signal retention gain of approximately $\sim 3\%$.
- The veto on the $B^+ \rightarrow K^+e^+e^-$ background has been removed, resulting in a 2% gain in signal efficiency. In addition to the gain in efficiency, this choice avoids a correlation between the variables $\cos \theta_K$ and $m(K\pi e^+e^-)$ caused by the veto, as shown in **Fig. 6.7**. A dedicated systematic uncertainty study, presented in **Sec. 6.2.2**, shows that the uncertainty from not applying the veto is smaller than the one associated with keeping it.
- The invariant mass range, $m(K\pi e^+e^-)$, used for the final fit has been chosen to be broader. This allows for improved background modelling and aligns the selection with the analyses in Refs. [50, 49].

In total, these changes result in an approximate 8.5% gain in signal efficiency. These efficiencies have been estimated using signal MC simulations by comparing the sample before and after each individual selection step.

Furthermore, the BDTs used in this thesis are newly developed and trained to target a broader range of backgrounds by exploiting the lower B^0 mass sideband in data, rather than just combinatorial background. In particular, the second BDT is specifically designed to reduce systematic uncertainties arising from backgrounds that peak at $\cos \theta_\ell$ values close to one. While in the previous analysis, a cut $\cos \theta_\ell < 0.9$ is applied, in this thesis the peaking background in $\cos \theta_\ell$ is reduced by the selection on the second BDT, see **Sec. 4.3**. This provides an additional handle for background suppression and eliminates the need for a stringent cut on $\cos \theta_\ell$, which would otherwise reduce sensitivity to certain angular observables, as discussed in **App. F**.

Another substantial difference lies in the treatment of the detector acceptance, $\epsilon(\cos \theta_\ell, \cos \theta_K, \phi)$. In the previous analysis, the data were corrected for acceptance effects by applying a per-event

weight of $1/\epsilon(\cos\theta_\ell, \cos\theta_K, \phi)$. As a result, the PDF used to fit the weighted angular distributions was simply the physics PDF given in **Eq. 1.37**. In contrast, this analysis incorporates the acceptance function directly into the likelihood. Both approaches are valid. However, the latter was chosen to avoid the use of excessively large weights in regions of high $\cos\theta_\ell$ and $\cos\theta_K$, where the acceptance approaches zero. Furthermore, the introduction of the acceptance moments strategy allows for a systematic-free modelling of the acceptance in the case of a signal-only fit. Initially, a different approach was explored, in which signal weights were computed from a fit to the invariant mass distribution, $m(K\pi e^+ e^-)$, following the same procedure used in this analysis to extract background weights - namely, using the sPlot technique [118]. The angular fit would then be performed on the weighted data, effectively resulting in a signal-only fit. The advantage of this strategy was that it avoided the need to parametrize the background in the angular variables, which led to large systematic uncertainties in the previous analysis. However, this came at the cost of increased statistical uncertainty compared to a full four-dimensional fit to mass and angular distributions. In practice, this strategy proved to be unviable due to the large amount of background present. The resulting signal weights were unreliable, causing the angular fit to have an extremely low convergence rate on toys, around 10%. For this reason, the sPlot-based strategy was abandoned in favour of the method detailed in this dissertation. Interestingly, what was a weakness in the signal-weighting approach - namely, the large background - became a strength here, as it allowed for the extraction of reliable background weights that accurately represent the true background shape in data, as seen in **Fig. 6.4**.

Therefore, even though the fitting strategy remains a four-dimensional fit to the mass and angular distributions, as in the previous analysis, the background modelling and acceptance correction approaches are significantly different. While the underlying assumptions are strong, this dissertation demonstrates that the method yields correct results within (or better than) half a sigma, providing a key advantage: the background angular shape is modelled without relying on simulations, which is a major strength of this approach.

Statistical uncertainties

Tab. 7.1 presents a comparison of the statistical uncertainties in the full central q^2 bin, spanning from 1.1 to 6.0 GeV^2/c^4 , between the previous and the current analysis. The table highlights the improvements in statistical sensitivity resulting from the methodological differences described above.

Systematic uncertainties

The following tables (**Tab. 7.2**, **Tab. 7.3**, **Tab. 7.4**, **Tab. 7.5**) compare the systematic uncertainties between the strategy adopted in this thesis and that of the previous analysis. Some of the systematic sources from the earlier analysis are either no longer applicable or have changed in nature due to the updated methodology.

The tables are grouped by two angular parameters, with systematics from the previous analysis listed under a column labeled “[39]”¹, and those from this thesis reported for the three q^2 regions: central (C), central_{low} (C_{low}), and central_{high} (C_{high}).

All systematic uncertainties are expressed as fractions of the corresponding statistical uncertainty. This choice is motivated by the fact that the central_{low} and central_{high} q^2 regions are less

¹The values listed under Ref. [39] are not available in the public paper; they were provided via private communication.

Parameters	Stat errors Ref. [39]	Stat errors this thesis	Stat decrease
A_{FB}	0.052	0.043	-17.3%
F_L	0.045	0.039	-13.3%
S_3	0.042	0.037	-11.9%
S_4	0.073	0.071	-2.7%
S_5	0.054	0.053	-1.8%
S_7	0.056	0.055	-1.8%
S_8	0.072	0.068	-5.6%
S_9	0.045	0.042	-6.7%

Table 7.1: Comparison between the statistical uncertainties from the fit to the central q^2 regions performed by the analysis in Ref. [39] and by this thesis analysis strategy. These same uncertainties were listed in **Tab. 6.12**. The last column highlights the percentage decrease of the statistical uncertainties with respect the previous analysis.

populated than the full central q^2 region. Therefore, expressing them as fractions of the statistical uncertainties allows for a meaningful comparison between different regions. However, when comparing the same q^2 region, it should be noted that the value quoted in “[39]” corresponds to a larger absolute systematic uncertainty than the one quoted in “this thesis”, because the current analysis yields smaller statistical uncertainties than Ref. [39] (**Tab. 7.1**). Entries marked as “N/A” (Not Applicable) indicate that the systematic does not apply to the strategy used in this thesis. A dash (“ - ”) indicates that the systematic is relevant but has not yet been evaluated.

The row labelled “Effective total” corresponds to the square root of the quadratic sum of the values shown in black, allowing for a fair comparison based only on the systematics that are available for both strategies. The row “Total” shows the square root of the quadratic sum of all listed values, including those shown in gray, which are available only for the previous analysis. Values marked in red are those with systematic uncertainty larger than the previous analysis. The values marked in blue have a larger systematic uncertainty in units of sigma, but a smaller absolute value, when considering the statistical uncertainties given in **Tab. 7.1**.

An explanation of the systematic uncertainties included in the tables follows, taken from the internal analysis note of Ref. [39].

- **$K\pi e\mu$ sample size:** this systematic uncertainty is associated with the limited number of candidates populating the lepton-flavour-violating sample $B^0 \rightarrow K\pi e\mu$ used to evaluate the DSL and combinatorial backgrounds shape and to generate realistic toys.
- **DSL comb. model:** this systematic uncertainty is related to the modelling choice of these backgrounds. This uncertainty applies to this analysis as well, as it is evaluated by changing the functional form used to describe the background.
- **DSL and Combinatorial strategy:** this systematic uncertainty is related to the fitting strategy used in the previous analysis, where the two background components were separated: DSL and Combinatorial. This systematic uncertainty does not apply to the current strategy, since they are modelled separately in the previous analysis. Instead, this analysis estimates a systematic uncertainty due to the modelling strategy, as described in the next item.
- **DSL comb. factorisation:** this systematic uncertainty arises from the fact that factorisation is assumed between the reconstructed mass B^0 and $\cos\theta_\ell$. The value reported under

the columns *this thesis* estimates both the effect of the factorisation assumption and the fitting strategy (modelling DSL and combinatorial under a unique exponential function), as explained in **Sec. 6.2.4**.

- **Part. reco. model** is a systematic uncertainty related to the modelling choice in mass and angular distributions of the partially reconstructed background. In this analysis, the quoted values are those derived from the BDT systematic, as described in **Sec. 6.2.6**, specifically for the case where the part. reco. background is involved. Since an angular modelling for each specific background is not used in this analysis, this systematic is related only to the mass modelling.
- **Part. reco. strategy**: this systematic uncertainty is associated with the fact that the $B^+ \rightarrow K\pi\pi e^+e^-$ MC simulation has been used to model this background. However, alternative decays could have been used instead, such as $B^+ \rightarrow K_1^{*+}(1270)e^+e^-$ or $B^+ \rightarrow K_2^{*+}(1430)e^+e^-$ or even a cocktail of these partially reconstructed decays.
- **Part. reco. factorisation**: the partially reconstructed background is modelled under the assumption that its angular and mass distributions factorise. However, due to the nature of partial reconstruction, residual correlations from the underlying decay dynamics can be expected. Therefore, this assumption introduces a source of systematic uncertainty.
- **Misid bkg sample size**: this systematic uncertainty is associated with the limited number of candidates in the PID fail region in data, which is mainly due to the $ePIDe > 0$ cut applied at the stripping level. This uncertainty could be reduced by re-stripping the data without that cut, thus improving the estimate of the number of expected misidentified events.
- **Misid bkg model**: this systematic uncertainty arises from the modelling choice for the misidentified background. In the previous analysis, a kernel density estimate (KDE) was used. The values reported in this analysis are those detailed in **Sec. 6.2.5**.
- **Misid fail region**: this systematic uncertainty is related to how the fail region is defined when computing the expected number of misidentified events in data, see **Sec. 3.5.2**.
- **Misid map parameterisation**: this systematic uncertainty is connected to the dependencies of the transfer function $w_{\text{fake}}(\vec{X})$. The nominal approach parametrises it as a function of P_T and η , but other possible dependencies may have been neglected.
- **Higher orders**: this systematic uncertainty is related to the acceptance parameterisation. It is evaluated in the same way between Ref. [39] and this thesis: the fit is repeated increasing all polynomial orders describing the acceptance parameterisation by 3. The values computed in this thesis are described in **Sec. 6.2.3**.
- **Coefficient uncertainties**: this systematic uncertainty is also related to the acceptance. In particular, it arises from the limited number of events available in the MC simulations used to compute the acceptance parameterisation and the acceptance moments. The values computed in this thesis are described in **Sec. 6.2.3**.
- **Resolution correction**: the acceptance function, which accounts for both geometrical and resolution effects, introduces a dependency on the true physics angular parameters, as explained in **Sec. 6.2.3**. The nominal acceptance function is obtained by parametrisation

rare mode simulations based on the Standard Model prediction. However, the true physics underlying the data may deviate from the SM. In this analysis, the systematic uncertainty is estimated using a recursive fit that updates the angular parameters used to compute the acceptance with those obtained directly from the fit. Even though the acceptance is handled differently in this analysis, a similar systematic uncertainty was also considered in the previous analysis, although it was computed using a different methodology.

- **Simulation corrections:** this systematic uncertainty is related to the use of an alternative simulation correction chain, intended to test the robustness of the results against different simulation correction strategies. Specifically, it is compared to the approach adopted in Refs. [50, 49]. In the present analysis, the robustness checks discussed in **Sec. 6.2.7** already yield significantly lower values. While these cannot be interpreted as upper limits, it is nevertheless expected that computing the systematic uncertainty using the same approach as in the previous analysis would result in smaller values, since the correction strategy used in this analysis is more closely aligned with that of Refs. [50, 49].
- **PIDCalib sample size:** this systematic uncertainty is closely related to the method used to evaluate the previous on the simulation corrections. Some elements of the simulation correction procedure are shared between the nominal and alternative strategies described in the previous point. As a result, simply switching between them does not fully account for the associated uncertainties. The most relevant shared component is the statistical limitation of the PIDCalib sample, which is particularly significant in the electron mode. This systematic uncertainty is also applicable to this analysis, regardless of whether the simulation correction uncertainties are computed in the same way, as the PIDCalib sample is still used to correct the PID variables in the MC simulations.
- **Signal mass model:** this systematic uncertainty is related to the shape used to model the signal in the mass distribution. In this analysis, the values used are those computed in **Sec. 6.2.6**, specifically from the study of the BDT effect on the signal mass shape.
- **Shift and scale:** the control mode mass fit is used to extract the shift and scale parameters applied to the signal mass distribution. These two parameters regulate the position and width of the signal peak, respectively, and are fixed to the values obtained from a fit to the $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ data sample. Introducing these parameters helps account for residual differences between the simulation and the data.
- **Shift and scale (Strategy):** this systematic uncertainty arises due to the different mass functions used in the Rare and Control mode fits.
- **Signal mass model (sim-data):** the mismodelling observed in the control mode mass distribution - likely caused by remaining differences between simulation and data - suggests that similar effects may be present in the rare mode. However, these are difficult to detect directly due to limited statistics. This systematic uncertainty is designed to quantify their potential impact.
- **Signal factorisation:** the nominal fit strategy assumes full factorisation between the signal mass and angular distributions. However, this assumption is only approximately valid, as effects such as electron energy loss through photon emission and bremsstrahlung correction can introduce correlations, particularly between the mass and $\cos \theta_\ell$.

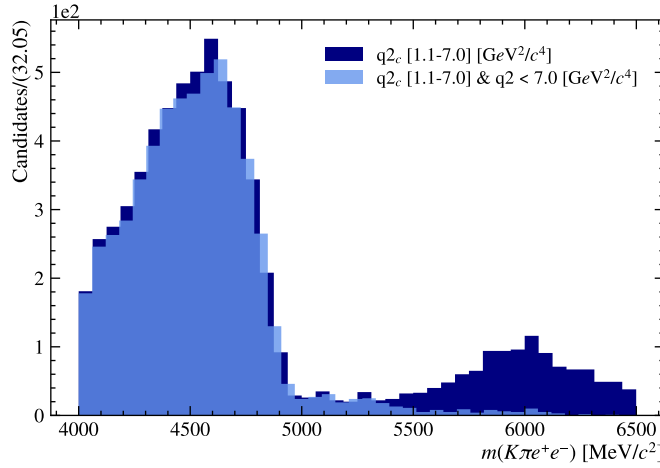


Figure 7.1: The plot shows the distributions of the reconstructed B^0 mass using the $B^0 \rightarrow K^{*0} J/\psi$ MC simulation with the *preselection* applied, without BDT cuts and selecting the $q^2_{PV\&B}$ region $[1.1, 7.0] \text{ GeV}^2/c^4$. The light blue distribution has the additional q^2 veto $q^2_{unconst} < 7 \text{ GeV}^2/c^4$.

- **Control leak:** in the previous analysis, this systematic was related to the fact that a smaller region around the B^0 mass ($[4900, 6200] \text{ MeV}/c^2$) was used in the final fit, and the $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ leakage was neglected. In this analysis, the region around the B^0 pole mass is larger ($[4600, 6200] \text{ MeV}/c^2$), meaning that the leakage cannot be neglected. Therefore, the systematic uncertainty is instead related to the shape of the modelled $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ leakage. The quoted values are taken from **Sec. 6.2.6**, specifically when the shape of the $B^0 \rightarrow K^{*0} J/\psi$ MC was adjusted.
- **J/ψ comb.:** this systematic uncertainty is related to the presence of combinatorial events at higher B^0 reconstructed mass leaking from the $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ control mode. However, this systematic uncertainty is not applicable in this analysis, as those events are vetoed with the cut $q^2_{unconst} < 7 \text{ GeV}^2/c^4$, as shown in **Fig. 7.1**. The plot shows the B^0 mass distribution using the $B^0 \rightarrow K^{*0} J/\psi$ MC simulation with the *preselection* applied in the rare mode region ($[1.1, 7.0] \text{ GeV}^2/c^4$). The BDT cuts are not applied here to clearly illustrate the effect of the q^2 veto, since they would otherwise strongly suppress the bump at high mass values. The excess of combinatorial events, visible as a bump in the darker distribution, disappears when the additional cut $q^2_{unconst} < 7 \text{ GeV}^2/c^4$ is applied.
- **Neglecting S-wave:** due to the limited sensitivity, the angular PDF used in the nominal fit neither includes the S-wave component nor the S-P interference terms.
- **B^+ background:** in the previous analysis, the systematic uncertainty was associated with distortions in the mass and $\cos \theta_K$ distributions introduced by the B^+ veto. In this analysis, the systematic error instead reflects the effect of neglecting residual B^+ background events, as discussed in **Sec. 6.2.2**.
- **Fit Bias:** this refers to the biases observed in the fit validation using pseudo-experiments. Although in this analysis the biases are larger than the average systematic uncertainties, they are included within this thesis to allow for a comparison with the previous analysis. However, the angular fit strategy will be refined and improved to reduce these biases and bring them within the statistical uncertainties on the pull distributions of the realistic toys.

Comparing the available systematic uncertainties - specifically the *Effective total* - in the full central q^2 region between this analysis and the previous one, one can observe that the current analysis strategy generally leads to smaller systematic uncertainties. This improvement is primarily driven by enhancements in DSL and combinatorial background modelling, a refined acceptance strategy, the removal of the B^+ veto, and the introduction of the q^2 veto. In the current strategy, the dominant source of systematic uncertainty typically originates from fit biases. Even when these fit biases are included, the overall comparison still shows a reduction in systematics.

As a side note, the values in the *Effective total* rows correspond to the total systematic uncertainties listed in **Tab. 6.9**, **Tab. 6.10**, and **Tab. 6.10**. The differences are due to rounding at the per-mille level.

Future improvements The current background description can be improved by adopting a more detailed modelling. In the current toy studies, the background samples - Partially Reconstructed, DSL, and Combinatorial - were generated without assuming angular factorisation (by bootstrapping MC simulations or data samples including those backgrounds), yet the fit model assumes factorisation among the mass and the angular variables. Addressing known angular correlations - particularly between $\cos \theta_\ell$ and $\cos \theta_K$ in the partially reconstructed background - for instance by adopting or incorporating a non-factorised PDF is expected to help mitigate this issue.

Additionally, including a second set of lower-order polynomial functions, grouping backgrounds that do and do not have peaks in $\cos \theta_\ell$ would improve further the description of backgrounds.

Another improvement implemented in the analysis - although not finalised at the time this thesis was written - is the computation of a new mass variable in the MC samples that inherently accounts for residual differences between data and MC, in addition to the standard MC corrections. This variable is computed by smearing the momenta of the final-state particles, with a function calibrated using the $B^0 \rightarrow K\pi e^+ e^-$ data sample selecting the q^2 region around the $J/\psi \rightarrow e^+ e^-$ resonance ($[7, 11] \text{ GeV}^2/c^4$). Currently, these residual differences are handled using two additional parameters: a shift of the signal mean and a scaling factor for the signal width. These parameters are fixed to the values obtained from a fit to the reconstructed B^0 invariant mass in the $B^0 \rightarrow K\pi J/\psi \rightarrow e^+ e^-$ data sample.

Among the systematic uncertainties listed, *Shift and scale (bkg)*, *Shift and scale (same strat.)*, and *Signal mass model (sim-data)* arise from the assumption that the shift and scale parameters are identical between the q^2 regions around the J/ψ resonance ($[7, 11] \text{ GeV}^2/c^4$) and the rare mode regions between $1 \text{ GeV}^2/c^4$ and $6 \text{ GeV}^2/c^4$ in the $B^0 \rightarrow K\pi e^+ e^-$ collected data. With the proposed strategy, the associated systematic uncertainties are expected to be both fewer and smaller.

Physical boundaries The results of the realistic pseudo-experiments show that the most significant fit biases are those in the parameters A_{FB} and F_L . It is important to note that both A_{FB} and F_L lie near the boundary of their allowed physical region - *i.e.*, the edge of the parameter space where the angular signal PDF remains positive-definite. The allowed regions for the q^2 bins studied in this thesis are reported in **App. G**. This proximity to the boundary can manifest as fit instability and biases, and it can be the source of the large biases observed in the parameters A_{FB} and F_L in the central q^2 region, as well as those observed for the central_{low} and the central_{high} q^2 regions. In fact, the limited number of candidates available in these regions may enhance the

bias, as the parameters are less constrained by the available statistics and more likely to fluctuate into unphysical regions. This hypothesis seems to hold, since even the signal-only low-statistics toys are subject to biases. Such biases can be minimised by applying appropriate corrections or by including them as asymmetric systematic uncertainties.

Angular parameters		A_{FB}				F_L			
Reference		[39]	this thesis			[39]	this thesis		
Source↓	q^2 bin →	C	C	C_{low}	C_{high}	C	C	C_{low}	C_{high}
$K\pi e\mu$ sample size		0.18	-	-	-	0.26	-	-	-
DSL comb. model		0.12	-	-	-	0.29	-	-	-
DSL strategy		0.32	N/A	N/A	N/A	0.42	N/A	N/A	N/A
Combinatorial strategy		0.15	N/A	N/A	N/A	0.30	N/A	N/A	N/A
DSL comb. factorisation		0.16	0.12	0.00	0.23	0.23	0.12	0.18	0.18
Part. reco. model		0.03	0.18	0.24	0.06	0.07	0.13	0.10	0.10
Part. reco. strategy		0.08	-	-	-	0.20	-	-	-
Part. reco. factorisation		0.01	-	-	-	0.02	-	-	-
Misid bkg sample sizes		0.10	-	-	-		-	-	-
Misid bkg model		0.08	0.12	0.24	0.00	0.17	0.14	0.22	0.00
Misid fail region		0.06	-	-	-	0.27	-	-	-
Misid map parameterisation		0.06	-	-	-	0.10	-	-	-
Higher orders		0.17	0.00	0.00	0.00	0.19	0.00	0.00	0.06
Coefficient uncertainties		0.06	0.08	0.05	0.07	0.07	0.09	0.06	0.08
Resolution correction		0.07	0.01	0.11	0.00	0.08	0.01	0.05	0.01
Simulation correction		0.23	-	-	-	0.26	-	-	-
PIDCalib sample sizes		0.17	-	-	-	0.19	-	-	-
Signal mass model		0.01	0.20	0.17	0.24	0.03	0.00	0.15	0.11
Shift and scale (bkg)		0.04	-	-	-	0.17	-	-	-
Shift and scale (same strat.)		0.01	-	-	-	0.03	-	-	-
Signal mass model (sim-data)		0.03	-	-	-	0.07	-	-	-
Signal factorisation		0.11	-	-	-	0.18	-	-	-
Control leak		0.02	0.02	N/A	0.04	0.01	0.03	N/A	0.04
J/ψ comb.		0.13	N/A	N/A	N/A	0.12	N/A	N/A	N/A
Neglecting S-wave		0.10	-	-	-	0.35	-	-	-
B^+ veto		0.24	0.00	0.00	0.00	0.50	0.00	0.00	0.00
Fit bias		0.05	0.45	0.54	0.20	0.01	0.45	0.32	0.52
Effective total		0.52	0.56	0.67	0.40	0.82	0.51	0.47	0.58
Total		0.68				1.13			

Table 7.2: Systematic uncertainties for the angular parameters A_{FB} and F_L . The row “Effective total” corresponds to the quadratic sum of the values shown in black, the uncertainties estimated in this thesis as well. The uncertainty on A_{FB} for the central q^2 region in this thesis yields a larger fraction of the statistical error, compared to Ref. [39]. However, since the statistical uncertainty itself is smaller, the total systematic uncertainty is still smaller in the former case in absolute terms. Full explanation in the text.

Angular parameters		S_3				S_4			
Reference		[39]	this thesis			[39]	this thesis		
Source↓	q^2 bin →	C	C	C_{low}	C_{high}	C	C	C_{low}	C_{high}
$K\pi e\mu$ sample size		0.09	-	-	-	0.08	-	-	-
DSL comb. model		0.10	-	-	-	0.12	-	-	-
DSL strategy		0.14	N/A	N/A	N/A	0.22	N/A	N/A	N/A
Combinatorial strategy		0.25	N/A	N/A	N/A	0.11	N/A	N/A	N/A
DSL comb. factorisation		0.20	0.20	0.25	0.09	0.08	0.25	0.15	0.12
Part. reco. model		0.04	0.00	0.03	0.01	0.02	0.00	0.01	0.06
Part. reco. strategy		0.04	-	-	-	0.07	-	-	-
Part. reco. factorisation		0.00	-	-	-	0.03	-	-	-
Misid bkg sample sizes		0.12	-	-	-	0.07	-	-	-
Misid bkg model		0.06	0.00	0.00	0.00	0.06	0.00	0.00	0.05
Misid fail region		0.21	-	-	-	0.05	-	-	-
Misid map parameterisation		0.03	-	-	-	0.03	-	-	-
Higher orders		0.15	0.00	0.07	0.06	0.24	0.00	0.00	0.07
Coefficient uncertainties		0.06	0.009	0.08	0.10	0.08	0.10	0.06	0.07
Resolution correction		0.05	0.00	0.02	0.00	0.08	0.01	0.00	0.01
Simulation correction		0.12	-	-	-	0.20	-	-	-
PIDCalib sample sizes		0.17	-	-	-	0.29	-	-	-
Signal mass model		0.02	0.09	0.00	0.00	0.01	0.17	0.00	0.08
Shift and scale (bkg)		0.04	-	-	-	0.05	-	-	-
Shift and scale (same strat.)		0.02	-	-	-	0.02	-	-	-
Signal mass model (sim-data)		0.03	-	-	-	0.03	-	-	-
Signal factorisation		0.03	-	-	-	0.04	-	-	-
Control leak		0.00	0.01	N/A	0.00	0.00	0.02	N/A	0.01
J/ψ comb.		0.04	N/A	N/A	N/A	0.04	N/A	N/A	N/A
Neglecting S-wave		0.09	-	-	-	0.13	-	-	-
B^+ veto		0.13	0.00	0.00	0.00	0.15	0.00	0.00	0.00
Fit bias		0.01	0.06	0.14	0.08	0.03	0.16	0.18	0.08
Effective total		0.42	0.24	0.31	0.17	0.41	0.35	0.24	0.21
Total		0.55				0.59			

Table 7.3: Systematic uncertainties for the angular parameters S_3 and S_4 . The row “Effective total” corresponds to the quadratic sum of the values shown in black, therefore it includes only those systematic already estimated in this thesis. Full explanation in the text.

Angular parameters	S_5					S_7			
Reference	[39]	this thesis				[39]	this thesis		
Source↓	q^2 bin →	C	C	C_{low}	C_{high}	C	C	C_{low}	C_{high}
$K\pi e\mu$ sample size		0.12	-	-	-	0.08	-	-	-
DSL comb. model		0.19	-	-	-	0.08	-	-	-
DSL strategy		0.13	N/A	N/A	N/A	0.09	N/A	N/A	N/A
Combinatorial strategy		0.14	N/A	N/A	N/A	0.09	N/A	N/A	N/A
DSL comb. factorisation		0.12	0.12	0.06	0.08	0.07	0.00	0.00	0.00
Part. reco. model		0.03	0.02	0.05	0.02	0.01	0.02	0.01	0.00
Part. reco. strategy		0.10	-	-	-	0.03	-	-	-
Part. reco. factorisation		0.02	-	-	-	0.00	-	-	-
Misid bkg sample sizes		0.14	-	-	-	0.15	-	-	-
Misid bkg model		0.07	0.00	0.00	0.00	0.12	0.00	0.00	0.00
Misid fail region		0.07	-	-	-	0.27	-	-	-
Misid map parameterisation		0.05	-	-	-	0.05	-	-	-
Higher orders		0.21	0.00	0.06	0.00	0.21	0.00	0.10	0.00
Coefficient uncertainties		0.08	0.10	0.08	0.08	0.08	0.10	0.08	0.09
Resolution correction		0.06	0.01	0.00	0.00	0.06	0.00	0.00	0.00
Simulation correction		0.20	-	-	-	0.21	-	-	-
PIDCalib sample sizes		0.30	-	-	-	0.43	-	-	-
Signal mass model		0.02	0.05	0.00	0.13	0.02	0.00	0.05	0.00
Shift and scale (bkg)		0.07	-	-	-	0.03	-	-	-
Shift and scale (same strat.)		0.02	-	-	-	0.01	-	-	-
Signal mass model (sim-data)		0.04	-	-	-	0.02	-	-	-
Signal factorisation		0.07	-	-	-	0.01	-	-	-
Control leak		0.00	0.01	N/A	0.02	0.00	0.01	N/A	0.02
J/ψ comb.		0.07	N/A	N/A	N/A	0.03	N/A	N/A	N/A
Neglecting S-wave		0.08	-	-	-	0.16	-	-	-
B^+ veto		0.18	0.00	0.00	0.00	0.16	0.00	0.00	0.00
Fit bias		0.03	0.00	0.05	0.10	0.00	0.00	0.17	0.00
Effective total		0.39	0.17	0.14	0.20	0.34	0.10	0.22	0.09
Total		0.62				0.70			

Table 7.4: Systematic uncertainties for the angular parameters S_5 and S_7 . The row “Effective total” corresponds to the quadratic sum of the values shown in black, therefore it includes only those systematic already estimated in this thesis. Full explanation in the text.

Angular parameters		S_8				S_9			
Reference		[39]	this thesis			[39]	this thesis		
Source↓	q^2 bin →	C	C	C_{low}	C_{high}	C	C	C_{low}	C_{high}
$K\pi e\mu$ sample size		0.07	-	-	-	0.09	-	-	-
DSL comb. model		0.07	-	-	-	0.14	-	-	-
DSL strategy		0.10	N/A	N/A	N/A	0.10	N/A	N/A	N/A
Combinatorial strategy		0.07	N/A	N/A	N/A	0.21	N/A	N/A	N/A
DSL comb. factorisation		0.65	0.17	0.31	0.00	0.08	0.00	0.00	0.00
Part. reco. model		0.01	0.00	0.00	0.00	0.02	0.04	0.04	0.02
Part. reco. strategy		0.03	-	-	-	0.04	-	-	-
Part. reco. factorisation		0.03	-	-	-	0.03	-	-	-
Misid bkg sample sizes		0.06	-	-	-	0.11	-	-	-
Misid bkg model		0.06	0.00	0.00	0.00	0.06	0.00	0.00	0.00
Misid fail region		0.09	-	-	-	0.06	-	-	-
Misid map parameterisation		0.03	-	-	-	0.03	-	-	-
Higher orders		0.24	0.00	0.13	0.00	0.13	0.00	0.12	0.00
Coefficient uncertainties		0.09	0.09	0.06	0.08	0.05	0.09	0.07	0.09
Resolution correction		0.07	0.03	0.07	0.01	0.04	0.00	0.01	0.00
Simulation correction		0.26	-	-	-	0.12	-	-	-
PIDCalib sample sizes		0.21	-	-	-	0.12	-	-	-
Signal mass model		0.02	0.14	0.07	0.00	0.02	0.13	0.10	0.09
Shift and scale (bkg)		0.03	-	-	-	0.03	-	-	-
Shift and scale (same strat.)		0.01	-	-	-	0.01	-	-	-
Signal mass model (sim-data)		0.03	-	-	-	0.03	-	-	-
Signal factorisation		0.01	-	-	-	0.04	-	-	-
Control leak		0.00	0.02	N/A	0.00	0.00	0.01	N/A	0.01
J/ψ comb.		0.03	N/A	N/A	N/A	0.05	N/A	N/A	N/A
Neglecting S-wave		0.01	-	-	-	0.10	-	-	-
B^+ veto		0.15	0.00	0.00	0.00	0.12	0.00	0.00	0.00
Fit bias		0.01	0.33	0.47	0.00	0.03	0.00	0.31	0.05
Effective total		0.73	0.41	0.59	0.08	0.32	0.16	0.36	0.14
Total		0.82				0.44			

Table 7.5: Systematic uncertainties for the angular parameters S_8 and S_9 . The row “Effective total” corresponds to the quadratic sum of the values shown in black, therefore it includes only those systematic already estimated in this thesis. Full explanation in the text.

7.2 Conclusions and Outlooks

To conclude, the current analysis strategy improves both the statistical precision and the control of systematic uncertainties compared to the previous approach. This analysis can supersede the existing results using the Run 1 and Run 2 datasets of the $B^0 \rightarrow K^{*0}e^+e^-$ decay channel in the full central q^2 region, $[1.1, 6.0] \text{ GeV}^2/c^4$. Additionally, it provides more detailed information for the smaller q^2 regions such as $[1.10, 3.65] \text{ GeV}^2/c^4$ and $[3.65, 6.00] \text{ GeV}^2/c^4$. Although beyond the scope of this thesis, this analysis also yields results in the previously unexplored low q^2 kinematic region $[0.1, 1.0] \text{ GeV}^2/c^4$.

To make the results public, the analysis will be unblinded following the official LHCb procedure: the internal analysis note undergoes peer review by designated experts who validate the analysis strategy. Once the strategy is deemed correct and robust, the results are unblinded and prepared for publication.

The development of this analysis and the expertise acquired in this decay channel have laid the groundwork for future studies using Run 3 data collected during 2024. The integrated luminosity in Run 3 alone already exceeds the total collected during Run 1 and Run 2, with 9.56 fb^{-1} [16] recorded to date. Furthermore, the LHCb Upgrade and the removal of the hardware trigger (L0) have significantly improved the trigger efficiency, as illustrated in **Fig. 7.2** from Ref. [124].

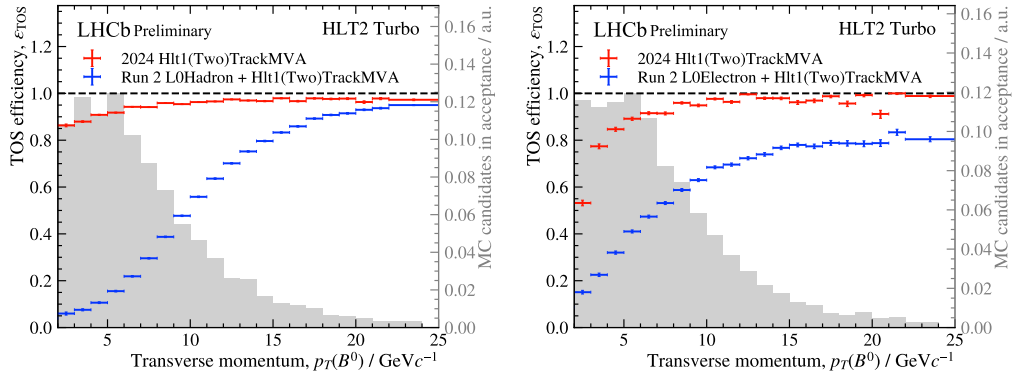


Figure 7.2: Combined TOS in B^0 decays to light hadrons, computed in bins of B^0 transverse momentum in 2024 data, shown in red. The combined TOS efficiencies calculated in Run 2 are shown in blue for comparison. Generator-level distributions for MC events within the LHCb acceptance are shown in grey. The plot on the left has been produced using $B^0 \rightarrow D\pi$ decays, while the plot on the right uses $B \rightarrow K^*e^+e^-$. Ref. [124]

Given these improvements, it is now possible to estimate the statistical uncertainty that can be achieved using the Run 3 dataset in the full central q^2 region. To do so, it is first necessary to estimate the total efficiency for detecting $B^0 \rightarrow K^{*0}e^+e^-$ decays, given by the production cross-sections at the relevant centre-of-mass energies ($\sigma(\sqrt{s}_{\text{Run1}})$ and $\sigma(\sqrt{s}_{\text{Run2}})$), the integrated luminosities for each run period (L_{Run1} , L_{Run2}), as reported in **Tab. 1.2**, the branching ratio of the $B^0 \rightarrow K^{*0}e^+e^-$ decay channel ($\mathcal{BR}(B^0 \rightarrow K^{*0}e^+e^-)$) [5], and the total number of observed signal events for Run 1 and Run 2, $N(B^0 \rightarrow K^{*0}e^+e^-)$, shown in **Tab. 5.3**. By substituting these quantities into **Eq. 2.2**, one can express the efficiency as:

$$\epsilon_{\text{Run1,2}}^{\text{tot}} = \frac{N(B^0 \rightarrow K^{*0}e^+e^-)}{(\sigma(\sqrt{s}_{\text{Run1}})L_{\text{Run1}} + \sigma(\sqrt{s}_{\text{Run2}})L_{\text{Run2}}) \cdot \mathcal{BR}(B^0 \rightarrow K^{*0}e^+e^-)} \simeq 0.39 \times 10^{-3}. \quad (7.1)$$

Using the same equation, it is now possible to estimate the expected number of $B^0 \rightarrow K^{*0}e^+e^-$ candidates for the integrated luminosity collected during Run 3, which amounts to 9.56 fb^{-1} [16]. The production cross-section at $\sqrt{s} = 13.6 \text{ TeV}$ is reported in **Tab. 1.2**. It is extrapolated from the values reported at $\sqrt{s} = 13.0 \text{ TeV}$, assuming a linear dependence on the centre-of-mass energy. This yields an estimated cross-section of approximately $150 \mu\text{b}$. Substituting these values into **Eq. 2.2**, the expected signal yield can be expressed as:

$$N(B^0 \rightarrow K^{*0}e^+e^-)_{\text{Run3}} = \varepsilon_{\text{Run1,2}}^{\text{tot}}(\sqrt{s}_{\text{Run3}})L_{\text{Run3}}\mathcal{BR}(B^0 \rightarrow K^{*0}e^+e^-) \simeq 559 \text{ candidates.} \quad (7.2)$$

This estimate assumes the same detection and trigger efficiency as in previous runs. However, for a B^0 meson with an average momentum of approximately 10 GeV , the efficiency is expected to improve significantly, from around 65% in Run 2 to about 90%, as shown in **Fig. 7.2**. This corresponds to a relative increase in signal yield of approximately 50%. Taking this into account, the number of expected $B^0 \rightarrow K^{*0}e^+e^-$ candidates in the central q^2 region from the full Run 3 dataset increases to roughly 1118 events. When combined with the existing Run 1 and Run 2 data, the total number of signal candidates would amount to approximately 1538. Since the statistical uncertainty scales as $1/\sqrt{N}$, this increase in statistics would reduce the uncertainty to about 0.025, representing a $\sim 49\%$ reduction relative to the previous value. This level of precision is now comparable to the current statistical uncertainties achieved in the corresponding muon channel [30] in the central q^2 region.

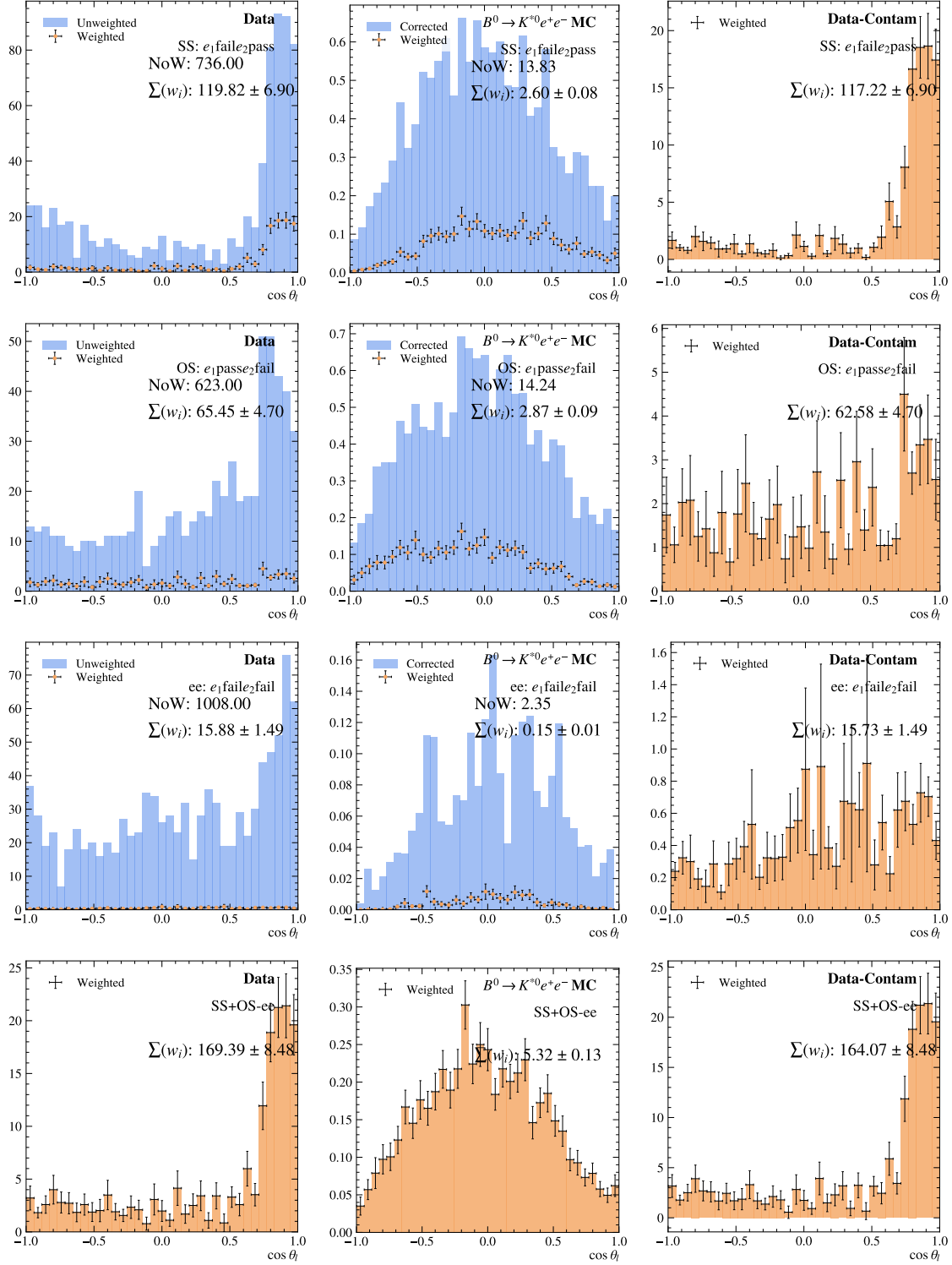
A | Misidentified electron background samples

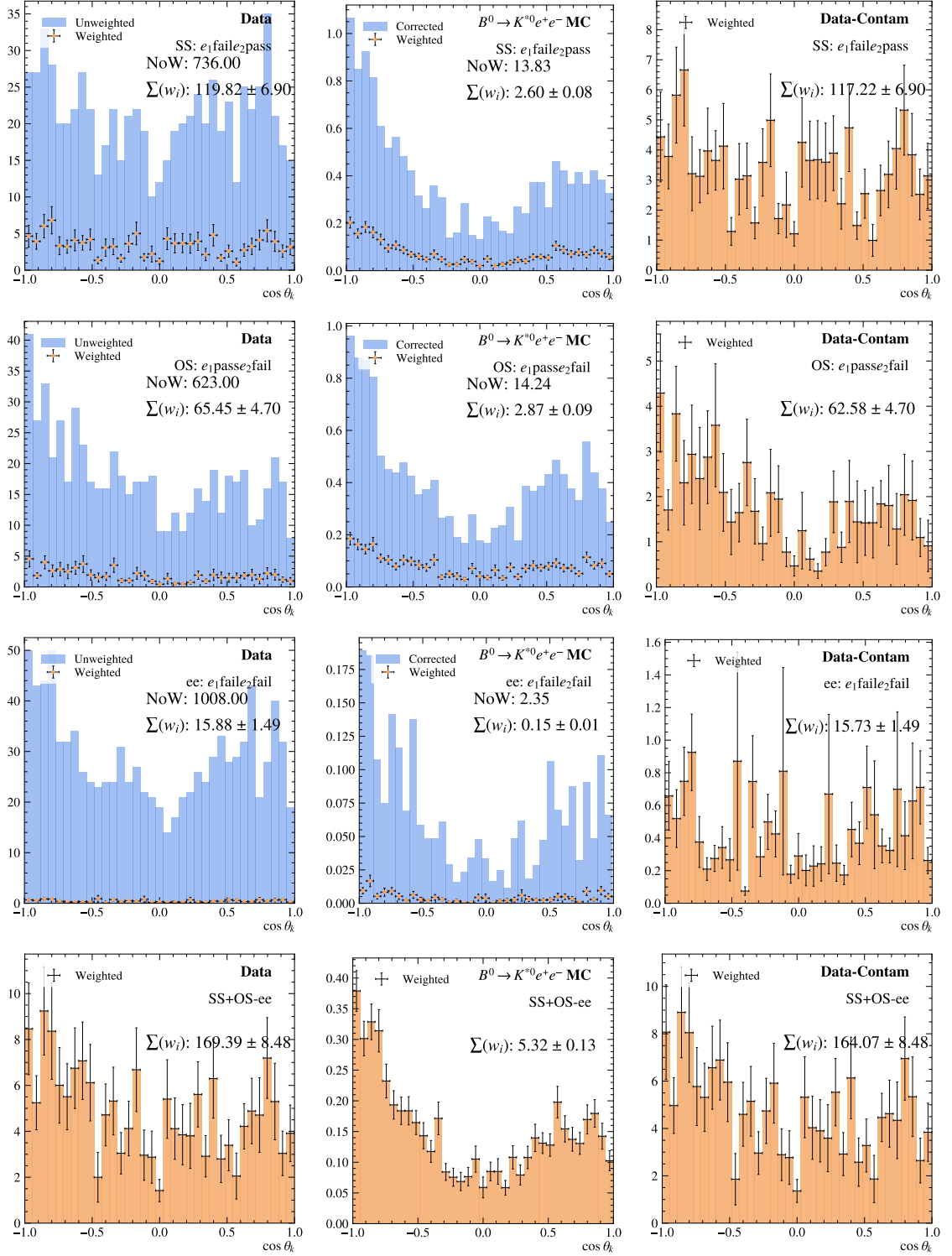
This appendix is a follow-up of **Chap. 3**, where the expected misidentified distributions have been shown only for the Run 2 dataset and in particular just for the $m(K\pi e^+ e^-)$. This appendix completes the chapter by showing, not only the $m(K\pi e^+ e^-)$, but also the $\cos\theta_\ell$, $\cos\theta_K$, and ϕ distributions for both the Run 1 and Run 2 datasets.

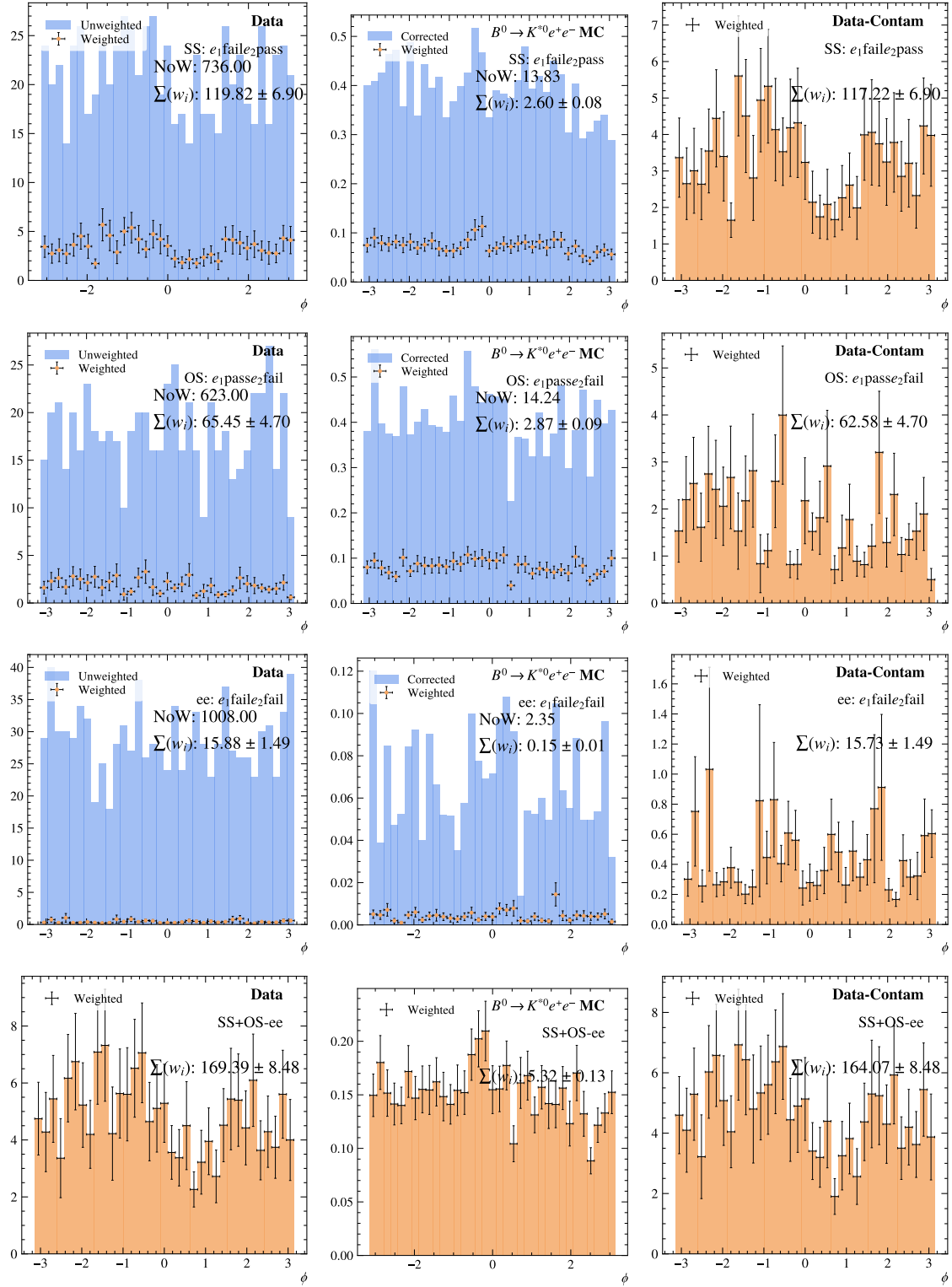
The plots are grouped by variables: $m(K\pi e^+ e^-)$, $\cos\theta_\ell$, $\cos\theta_K$, ϕ . Each group of figures shows the distributions of $B^0 \rightarrow K\pi e^+ e^-$ candidates obtained for the Same Sign (SS), Opposite Sign (OS), Fail-Fail (ee) from the first to third row, respectively. The first and the second columns show $B^0 \rightarrow K\pi e^+ e^-$ data and $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation, respectively. The solid blue distributions show the samples before applying the transfer function $w_{fake}(\vec{X})$, while the overlaid orange points show the same data after applying the $w_{fake}(\vec{X})$ weights. The last row shows the distributions obtained by combining the $w_{fake}(\vec{X})$ of the distributions above. The last column is the overall combined $w_{fake}(\vec{X})$ weighted data-contamination. The first three rows show OS, SS, ee , while the last row in the last column displays the final dataset representing the expected misidentified electrons background content. “NoW” shows the number of events belonging to the solid blue distribution before applying the scale function, while “ $\sum_i w_i$ ” gives the expected number of misidentified background events.

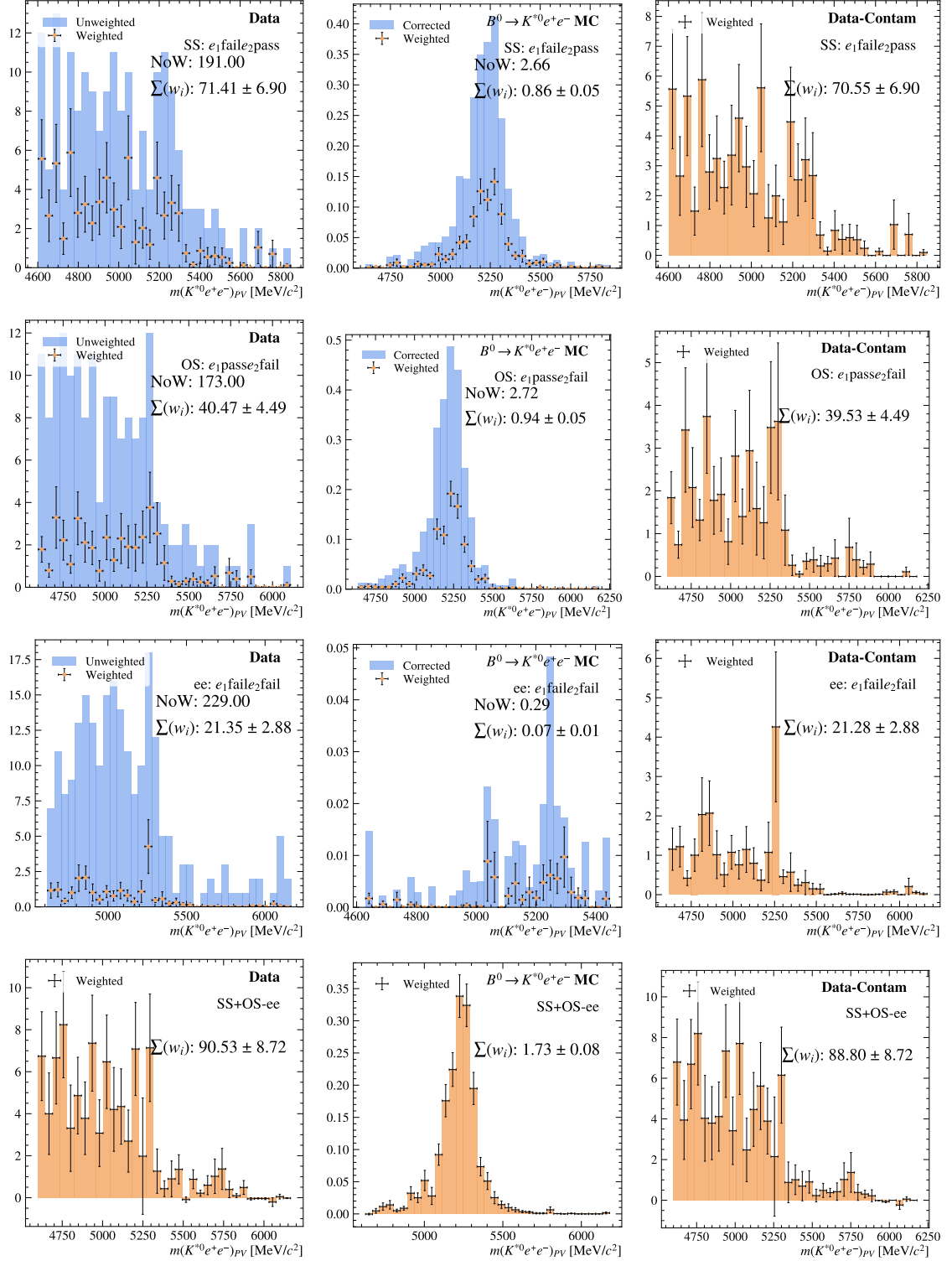
For clarity, the definitions of the OS, SS, and ee categories are recalled below:

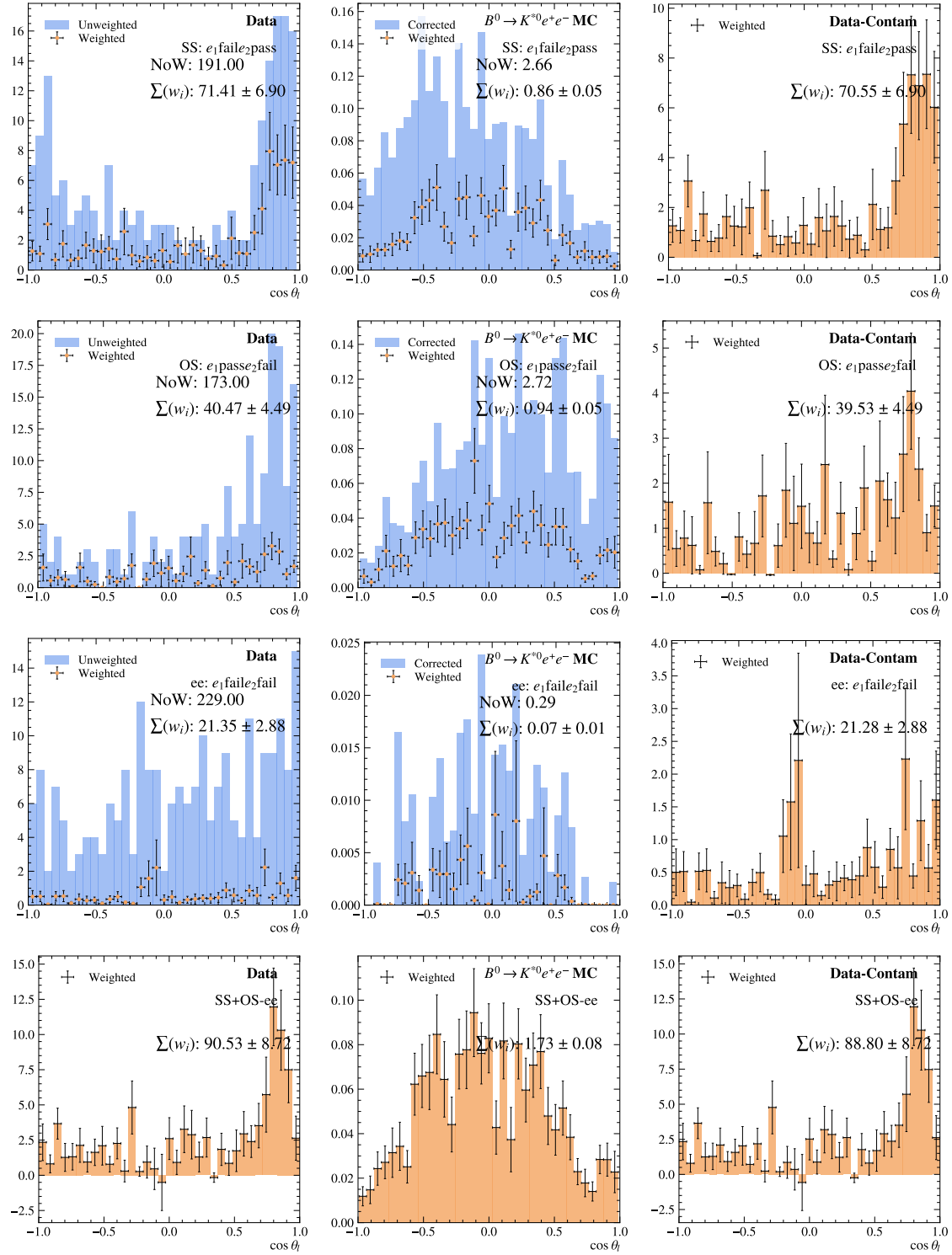
- OS: the e which fails the PID cut has the opposite charge to the kaon, while the other passes the requirement, $B^0 \rightarrow K^+ \pi^- e_{pass}^+ e_{fail}^-$.
- SS: the e which fails the PID cut has the same charge to the kaon, while the other passes the requirement, $B^0 \rightarrow K^+ \pi^- e_{fail}^+ e_{pass}^-$.
- ee : Both electrons fail the PID requirements.

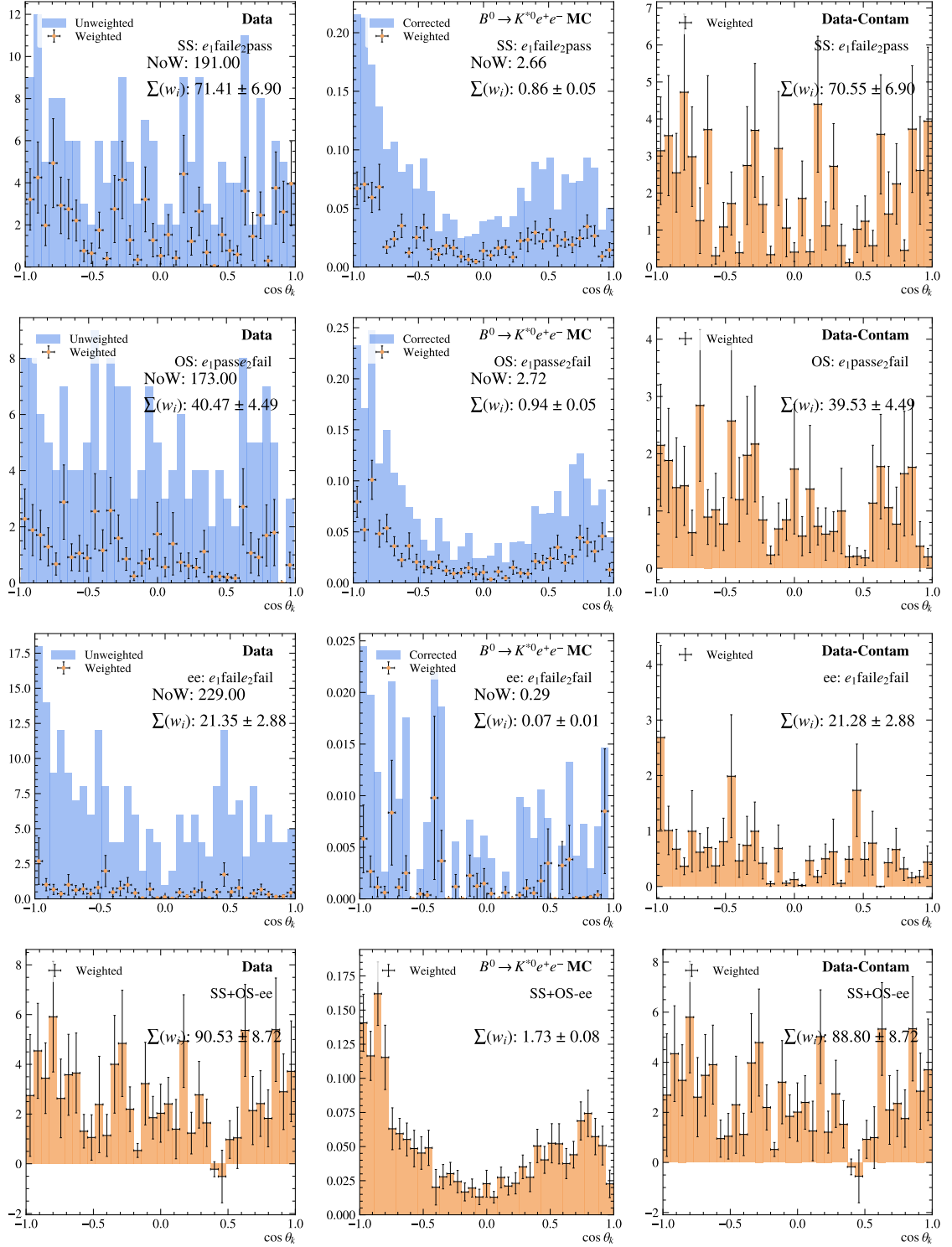
Figure A.1: Central q^2 , Run 2, $\cos \theta_\ell$.

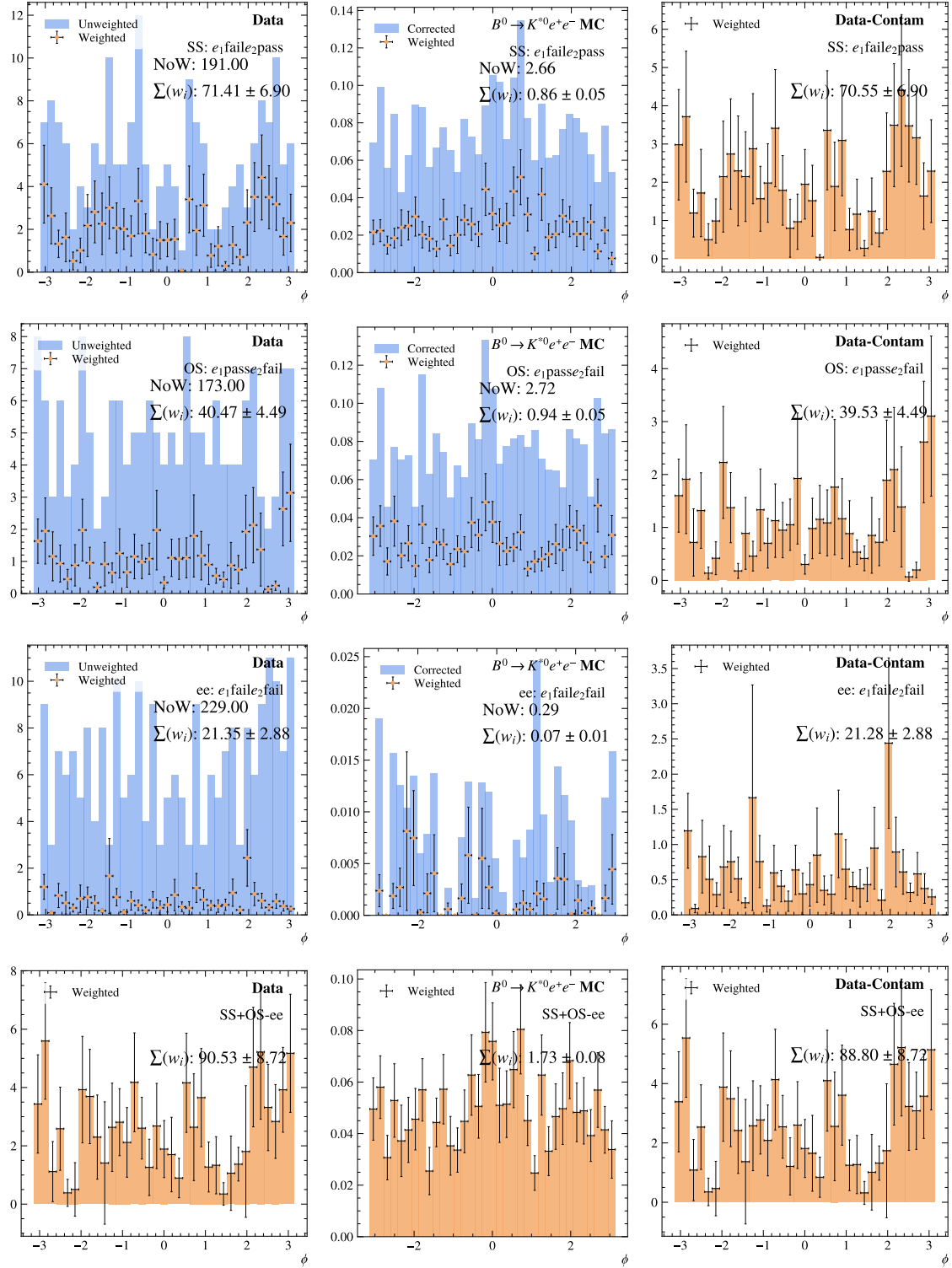
Figure A.2: Central q^2 , Run 2, $\cos \theta_K$.

Figure A.3: Central q^2 , Run 2, ϕ .

Figure A.4: Central q^2 , Run 1, $m(K^0 \pi^+ e^-)$.

Figure A.5: Central q^2 , Run 1, $\cos \theta_\ell$.

Figure A.6: Central q^2 , Run 1, $\cos \theta_K$.

Figure A.7: Central q^2 , Run 1, ϕ .

B | BDTs features

B.1 BDT1

This section concerns the BDT described in **Sec. 4.2**. **Fig. B.1** shows the distribution of the input features for the training samples (SIG and BKG).

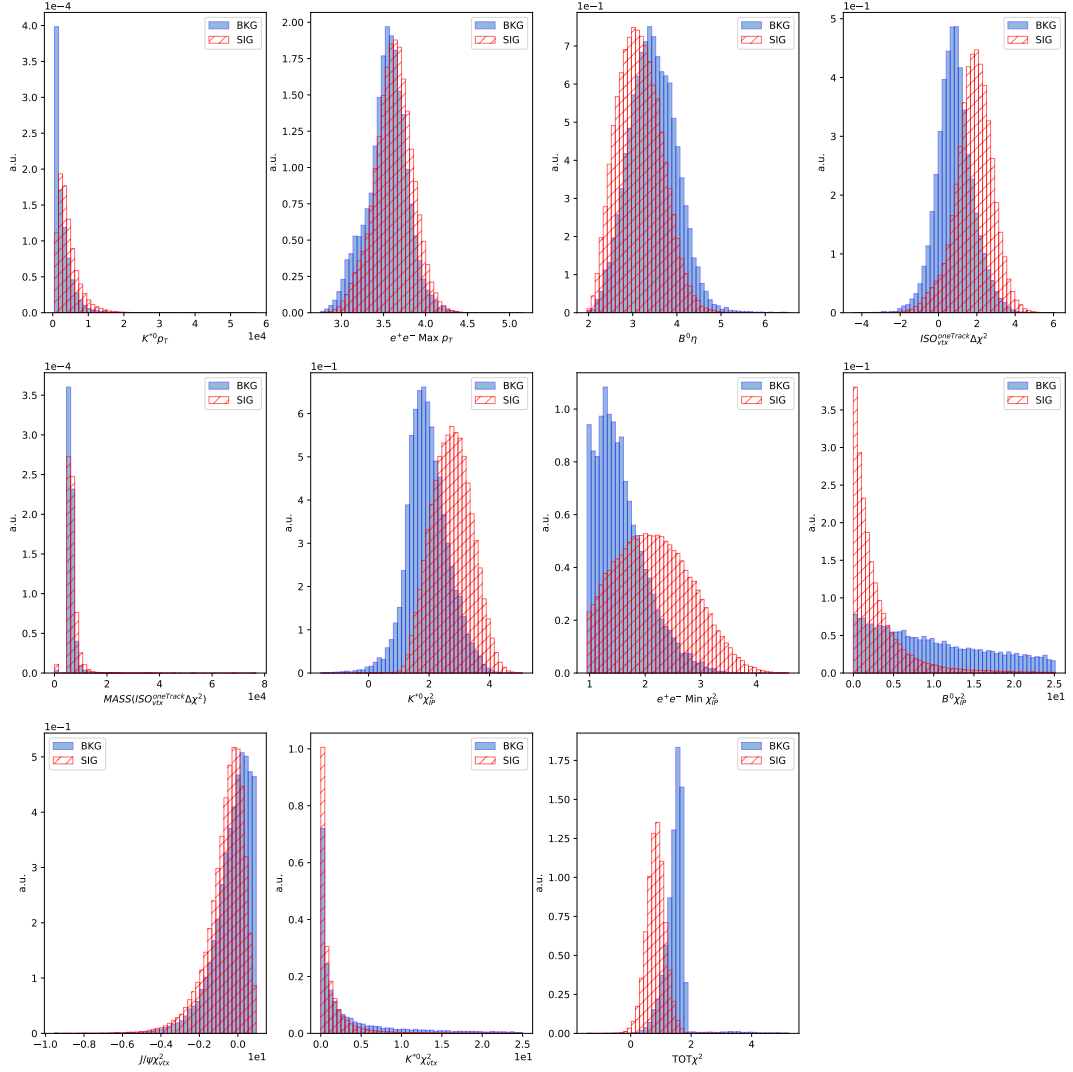


Figure B.1: Distributions of individual features for signal and background training samples used in the BDT training.

Fig. B.2 shows the correlation matrices of the background (above) and the signal (below) training samples.

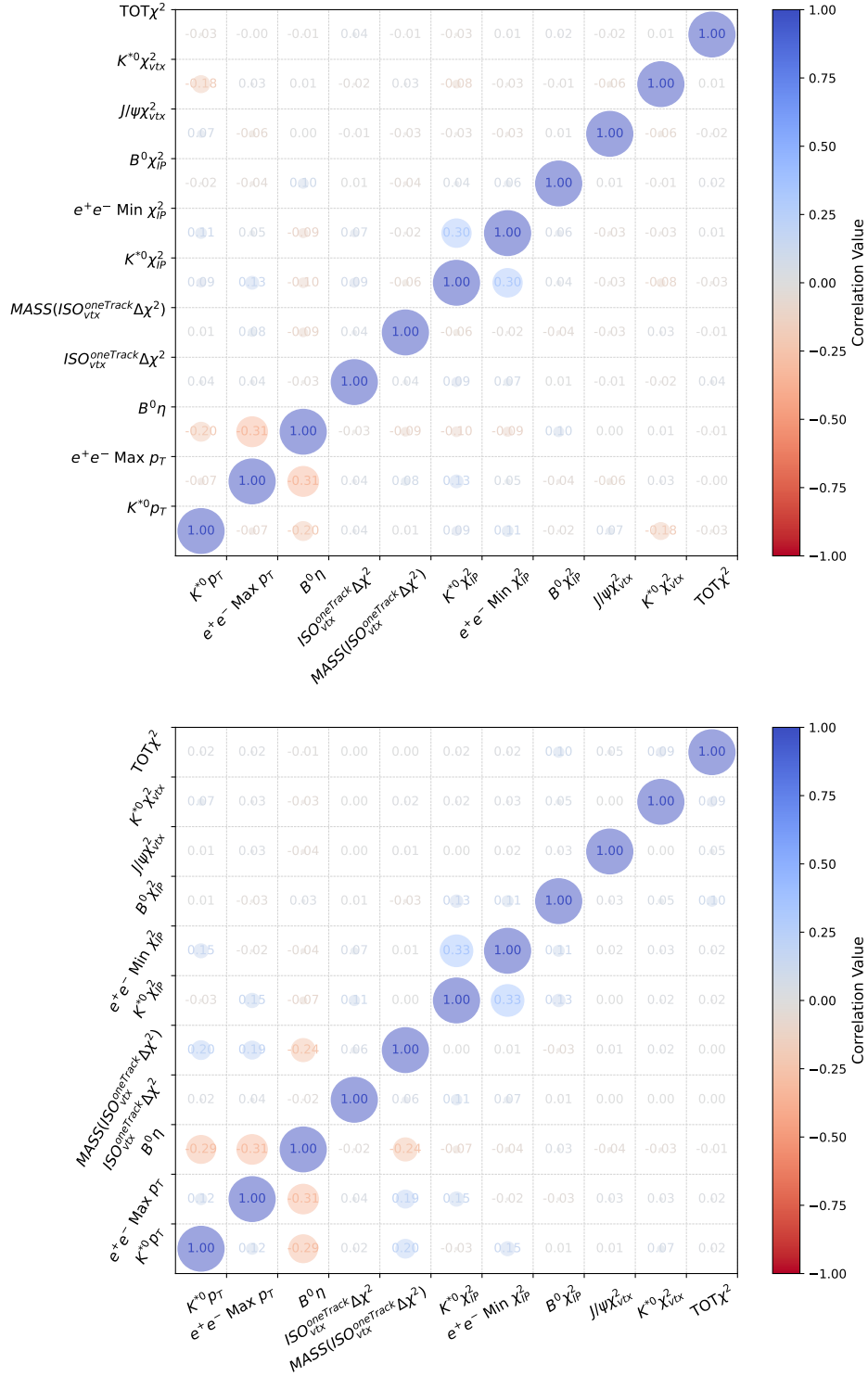


Figure B.2: Features correlation matrices for training samples. Signal (above) and background (below).

B.2 BDT2

This section concerns the BDT described in **Sec. 4.3**. **Fig. B.1** shows the distribution of the input features for the training samples (SIG and BKG).

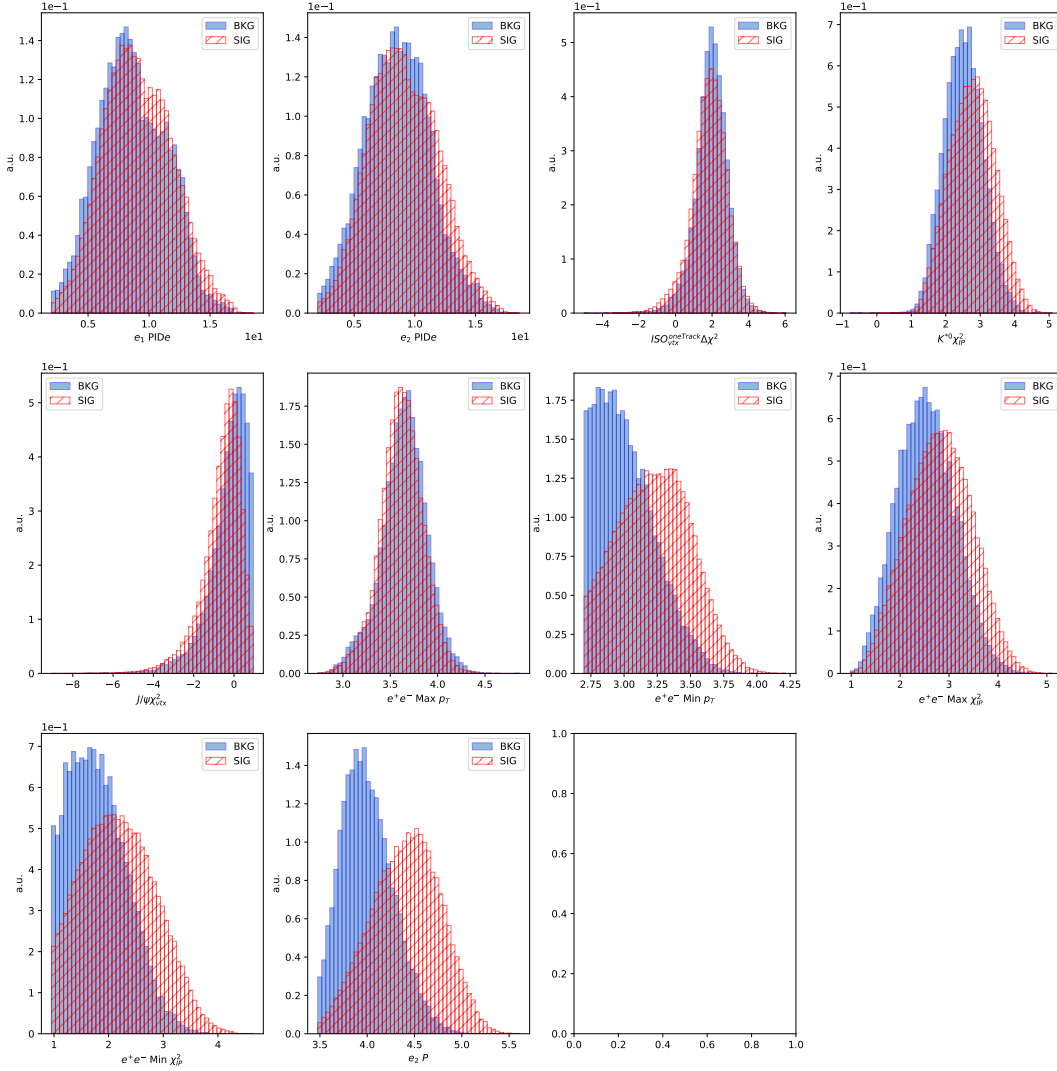


Figure B.3: Distribution in each of the features for the signal and background training samples.

Fig. B.2 shows the correlation matrices of the background (above) and the signal (below) training samples.

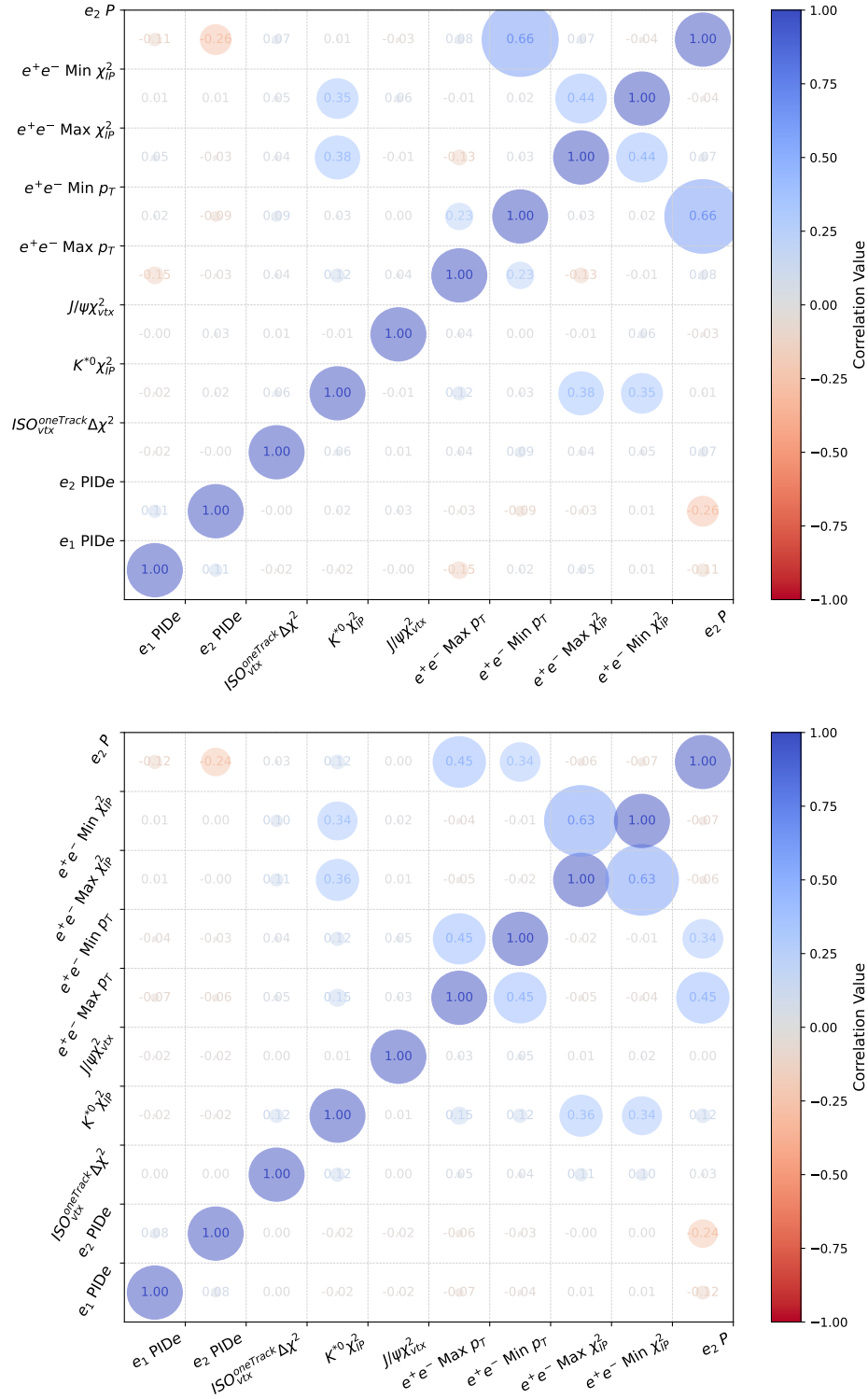


Figure B.4: Features correlation matrices for training samples. Signal (above) and background (below).

C | Fit components mass shapes

C.1 Signal

This section shows the tables with the best-fit values of the signal mass PDF presented in **Sec. 5.2.1**, using the $B^0 \rightarrow K^{*0} e^+ e^-$ MC samples, in the three q^2 regions under analysis, Run 1, and Run 2. **Fig. C.1** shows the signal MC samples, grouped according to the run period they simulate and the corresponding q^2 region, with the best-fit signal PDF overlaid. The plots are shown for the two smaller bins only, as the plots for the central region are already shown in **Chap. 5**.

$m(K\pi e^+ e^-)$ fit to the $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation						
q^2 region $\rightarrow$	central		central _{low}		central _{high}	
Parameter $\downarrow$	Run 2	Run 1	Run 2	Run 1	Run 2	Run 1
a_0 Brem0	0.17	0.19	0.19	0.21	0.16	0.17
a_0 Brem1	0.54	0.43	0.60	0.40	0.49	0.44
a_0 Brem2	0.80	1.07	0.75	1.14	0.80	0.70
a_1 Brem0	2.46	2.67	2.46	2.74	2.39	2.33
a_1 Brem1	1.38	1.19	1.41	1.12	1.37	1.22
a_1 Brem2	0.79	1.58	0.78	1.71	0.77	0.72
frac Brem1	0.91	0.97	0.89	0.99	0.93	0.97
frac Brem2	0.91	0.76	0.87	0.71	0.87	0.99
μ_G Brem1	5234.76	5328.04	5240.26	5431.48	5232.93	5329.66
μ_G Brem2	5306.95	5272.89	5280.89	5257.27	5256.30	5299.53
μ_{DCB} Brem0	5256.80	5251.31	5256.67	5250.09	5256.97	5253.84
μ_{DCB} Brem1	5254.15	5256.30	5254.80	5258.40	5253.26	5254.41
μ_{DCB} Brem2	5266.43	5270.00	5268.74	5271.05	5264.22	5266.74
n_0 Brem0	36.26	15.11	24.33	13.23	104.93	19.96
n_0 Brem1	6.05	9.39	5.09	14.29	7.47	7.33
n_0 Brem2	5.76	3.79	11.99	4.80	9.43	6.15
n_1 Brem0	1.94	1.86	2.61	1.70	1.76	2.65
n_1 Brem1	3.44	4.10	3.28	3.95	3.52	4.35
n_1 Brem2	15.68	2.71	31.41	2.12	58.54	11.48
σ_G Brem1	187.75	182.39	180.00	140.74	193.10	182.83
σ_G Brem2	270.34	144.19	282.33	146.80	257.21	5.96
σ_{DCB} Brem0	21.57	22.60	21.84	23.23	21.22	20.91
σ_{DCB} Brem1	41.42	37.37	41.19	36.24	41.92	38.06
σ_{DCB} Brem2	47.23	52.50	44.63	51.84	48.59	45.93

Table C.1: Result of the fit to the $m(K\pi e^+ e^-)$ distribution in the $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation sample. The table lists the fit parameters for the central, central_{low}, and central_{high} q^2 regions, separately for Run 1 and Run 2.

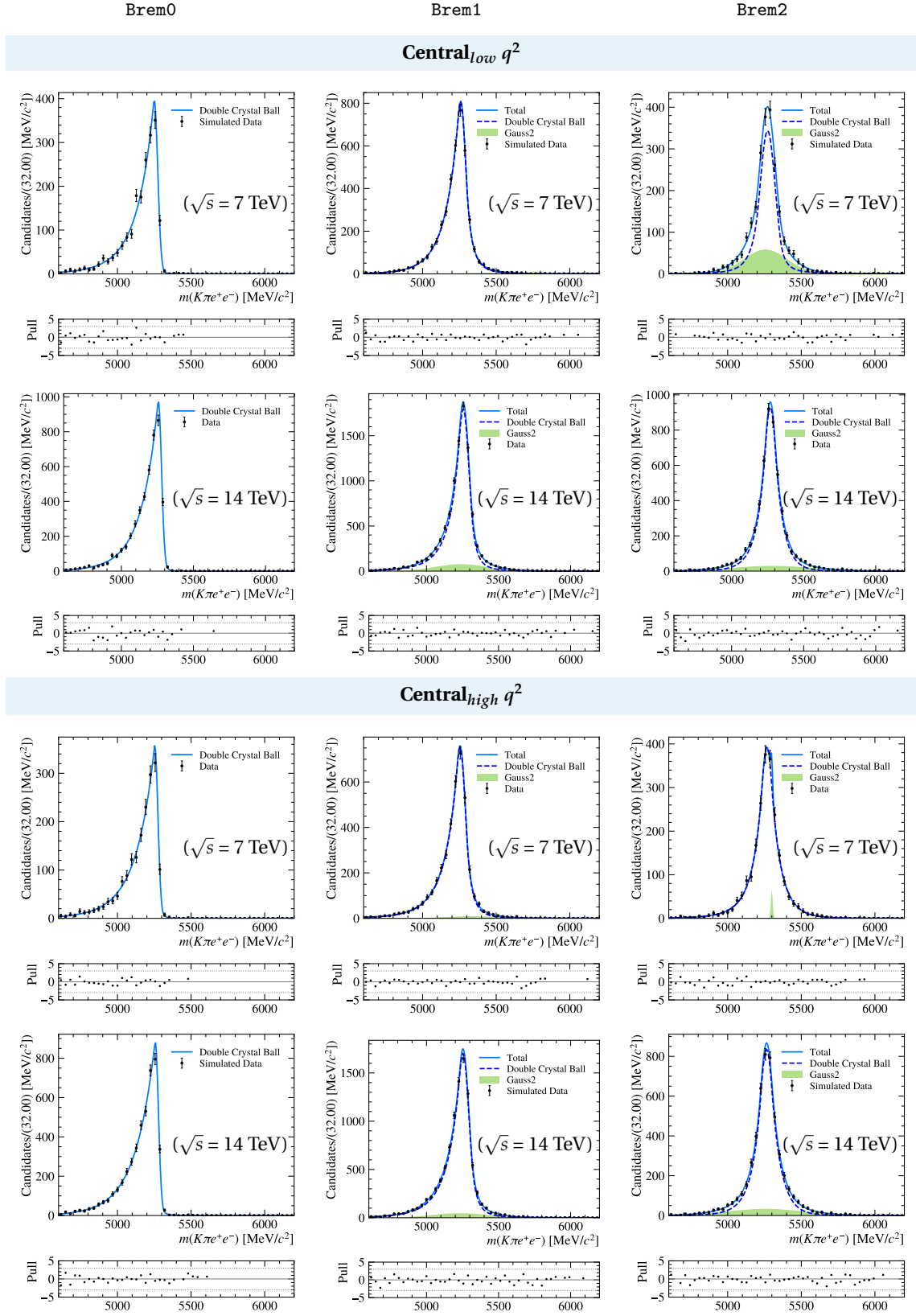


Figure C.1: Fit to the $m(K\pi^+e^-)$ reconstructed mass of the $B^0 \rightarrow K^{*0} e^+ e^-$ MC simulation.

C.2 Backgrounds

This section shows the tables with the best-fit values of the mass PDFs presented in **Sec. 5.2.2** of the electron misidentified background (**Tab. C.2**) and for the lepton flavour-violating sample $B^0 \rightarrow K^{*0}e\mu$ (**Tab. C.3**), which is used to model DSL and combinatorial backgrounds. The results are shown separately for three q^2 regions, Run 1, and Run 2. **Fig. C.2** shows electron misidentified samples - estimated as explained in **Sec. 3.5.2** - and the $B^0 \rightarrow K\pi e\mu$ data samples, grouped by run period and q^2 region, with the best-fit signal PDF overlaid. **Fig. C.3** shows the $B^0 \rightarrow K\pi\pi^+e^-$ and $B^0 \rightarrow K^{*0}J/\psi$ MC samples, grouped according to the run period they simulate and the corresponding q^2 region, with the best-fit PDF overlaid. The plots are shown for the two smaller bins only as the plot for the central region are already shown in **Chap. 5**.

$m(K\pi e^+e^-)$ fit to the expected misidentified distribution in data						
q^2 region $\rightarrow$	central		central _{low}		central _{high}	
Parameter $\downarrow$	Run 2	Run 1	Run 2	Run 1	Run 2	Run 1
frac _{G1}	0.49	0.40	0.30	0.99	0.57	0.25
frac _{G2}	0.042	0.040	0.16	1-0.99	-	0.046
λ	-0.0032	-0.0024	-0.0031	-	-0.0036	-0.0024

Table C.2: Result of the fit to the $m(K\pi e^+e^-)$ distribution for the expected misidentified sample. The table lists the fit parameters for the central, central_{low}, and central_{high} q^2 regions, separately for Run 1 and Run 2.

$m(K\pi e^+e^-)$ fit to the $B^0 \rightarrow K^{*0}e\mu$ data sample						
q^2 region $\rightarrow$	central		central _{low}		central _{high}	
Parameter $\downarrow$	Run 2	Run 1	Run 2	Run 1	Run 2	Run 1
λ	-0.0036	-0.0031	-0.0034	-0.0034	-0.0038	-0.0029

Table C.3: Result of the fit to the $m(K\pi e^+e^-)$ distribution in the $B^0 \rightarrow K^{*0}e\mu$ data sample. The table lists the fit parameters for the central, central_{low}, and central_{high} q^2 regions, separately for Run 1 and Run 2.

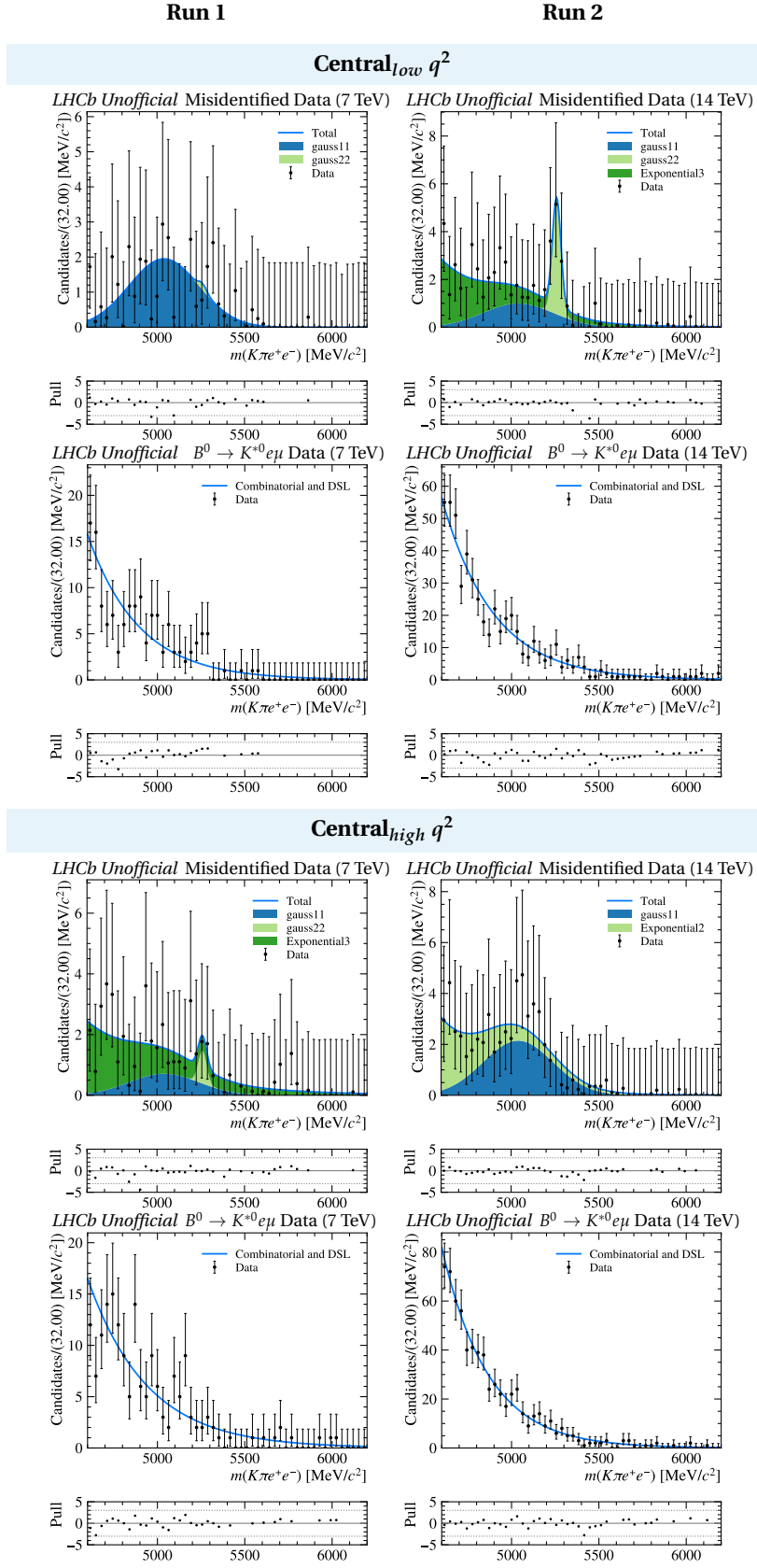


Figure C.2: Estimated misidentified samples and $B^0 \rightarrow K\pi e\mu$ data samples divided by run period and q^2 regions overlaid by their best-fit PDFs.

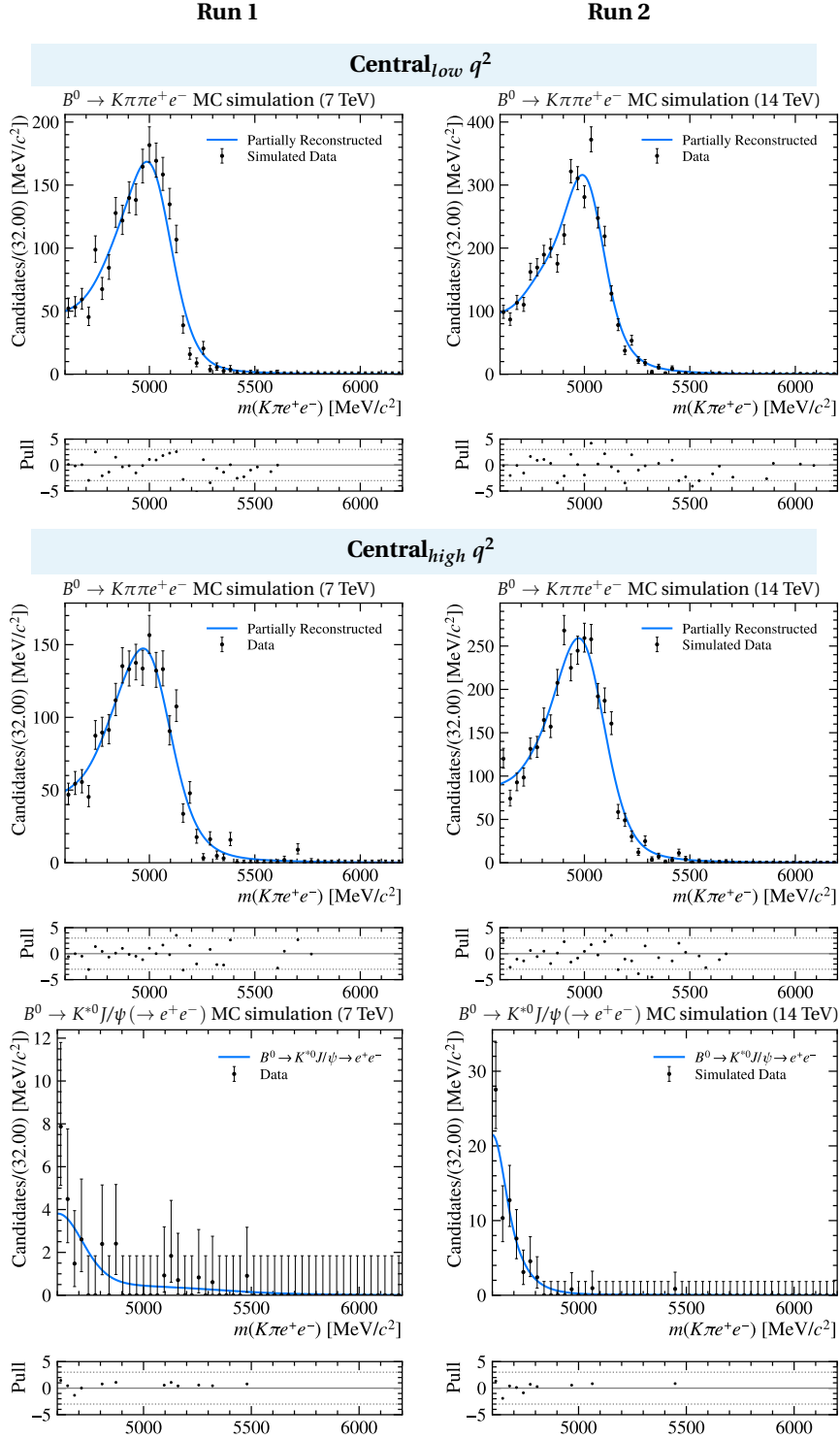


Figure C.3: $B^0 \rightarrow K\pi\pi^+e^-$ (top rows) and $B^0 \rightarrow K^{*0}J/\psi$ (bottom row) MC samples, grouped according to the run period they simulate and the corresponding q^2 region, with their best-fit PDF overlaid.

D | Additional material on Chap. 6

Sec. D.1 and **Sec. D.2** report the pull distributions (Eq. 6.1) for the angular parameters obtained by fitting 1 000 signal-only and realistic toys simulations, respectively. Each plot shows the mean and the sigma of the Gaussian function used to fit the distribution, together with the number of toys that converged, N_{toys} . The means and the widths are summarised in **Fig. 6.1** and **Fig. 6.6** for signal-only and realistic toys, respectively.

Sec. D.3 reports a table listing the values of the free parameters and their associated uncertainties, as determined by fitting simultaneously the Run 1 and Run 2 data across all the q^2 regions considered in this thesis. The values of the angular parameters remain blinded.

D.1 Pulls distributions of signal-only toys

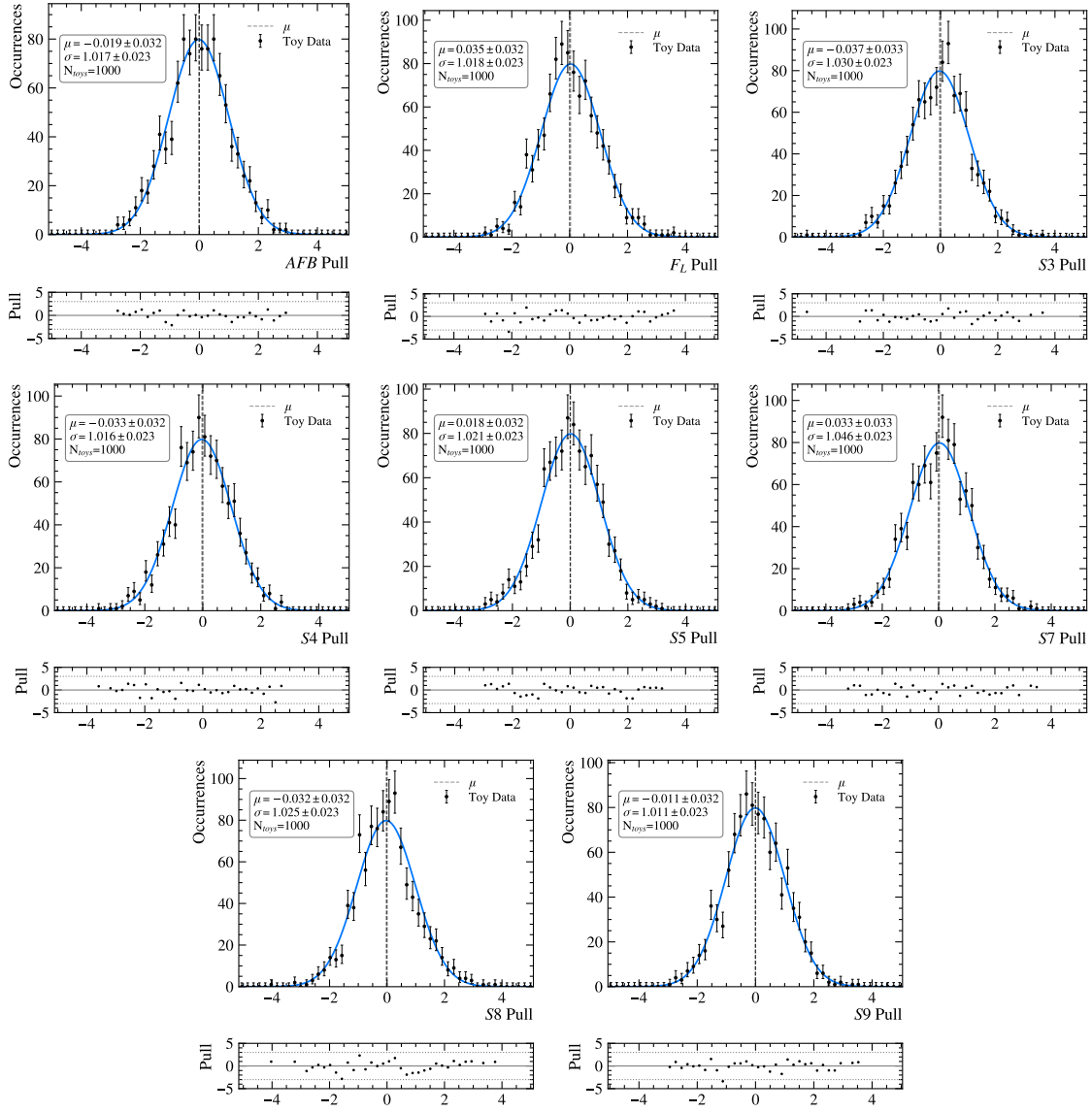


Figure D.1: Central q^2 : results of 1000 signal-only toys simulating the realistic number of $B^0 \rightarrow K^{*0} e^+ e^-$ events expected in the central q^2 region ($[1.1 - 6.0] \text{ GeV}^2/c^4$). The number of events per toy is sampled from a Poisson with a mean corresponding to the value reported in **Tab. 5.2**.

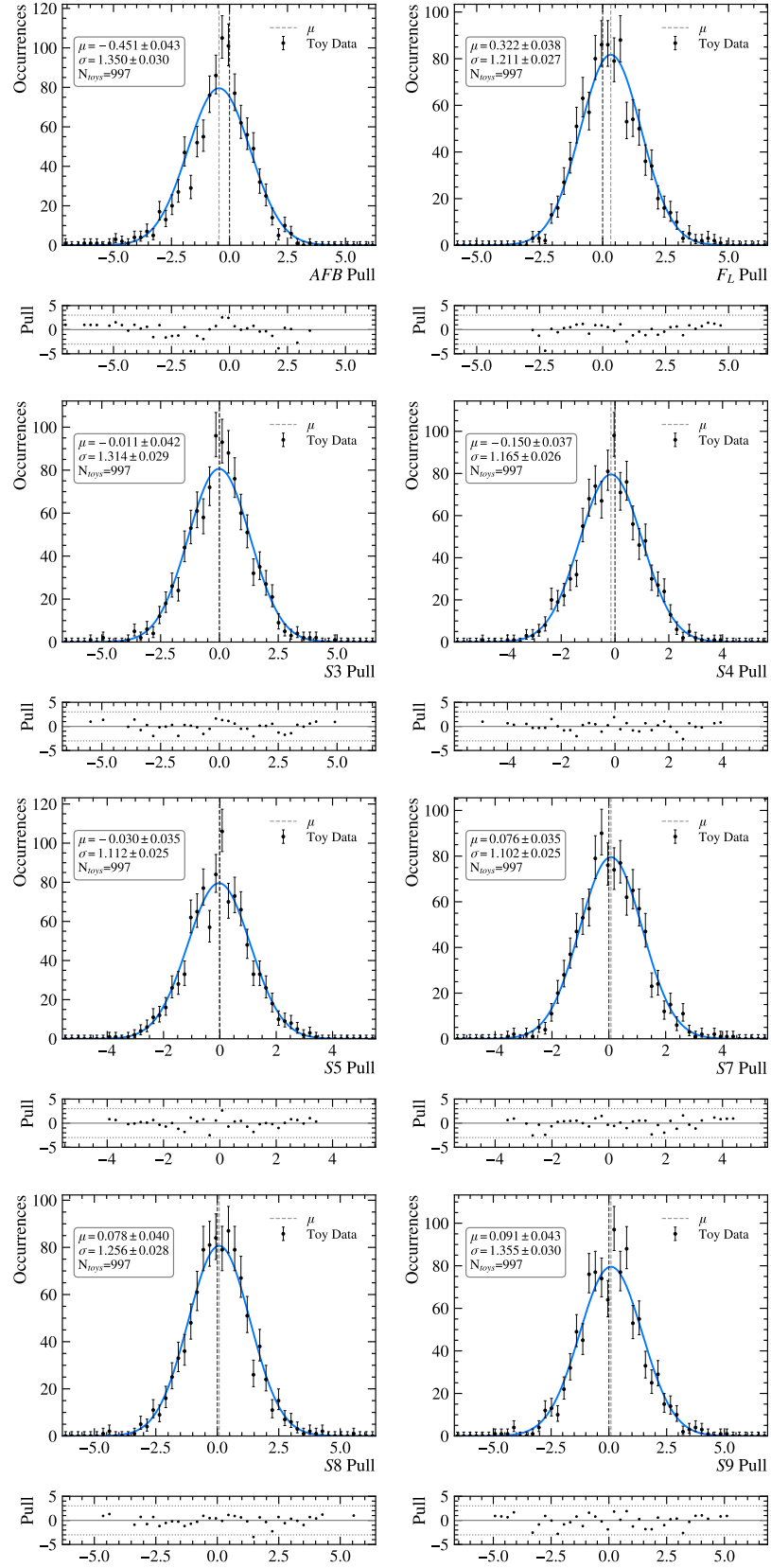


Figure D.2: Central_{low} q^2 : results of 1000 signal-only toys simulating the realistic number of $B^0 \rightarrow K^{*0} e^+ e^-$ events expected in the central q^2 region ($[1.10 - 3.65] \text{ GeV}^2/c^4$). The number of events per toy is sampled from a Poisson with a mean corresponding to the value reported in **Tab. 5.2**.

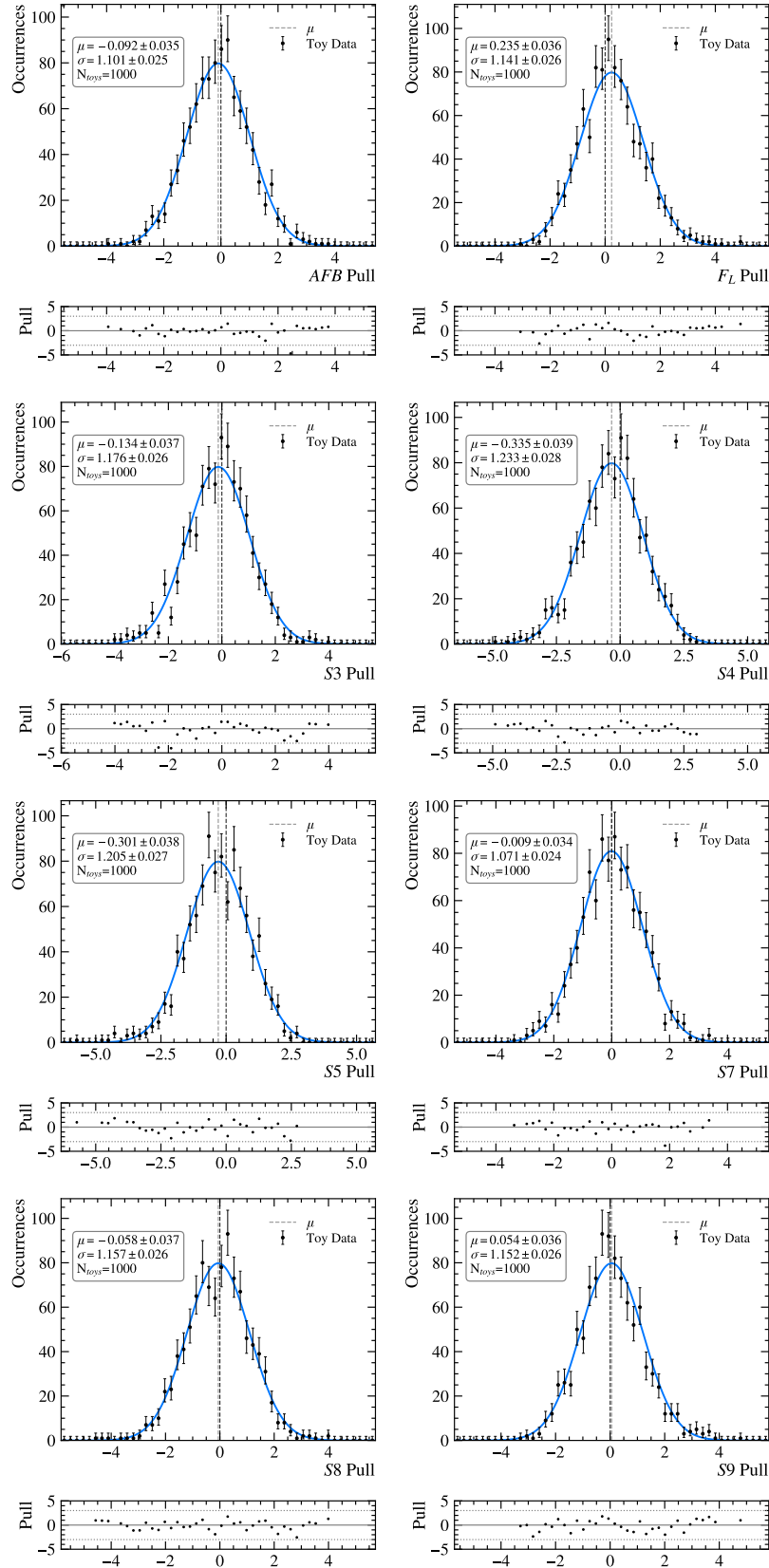


Figure D.3: Central_{high} q^2 : results of 1000 signal-only toys simulating the realistic number of $B^0 \rightarrow K^{*0} e^+ e^-$ events expected in the central q^2 region ($[3.65 - 6.00] \text{ GeV}^2/c^4$). The number of events per toy is sampled from a Poisson with a mean corresponding to the value reported in **Tab. 5.2**.

D.2 Pull distributions of realistic toys

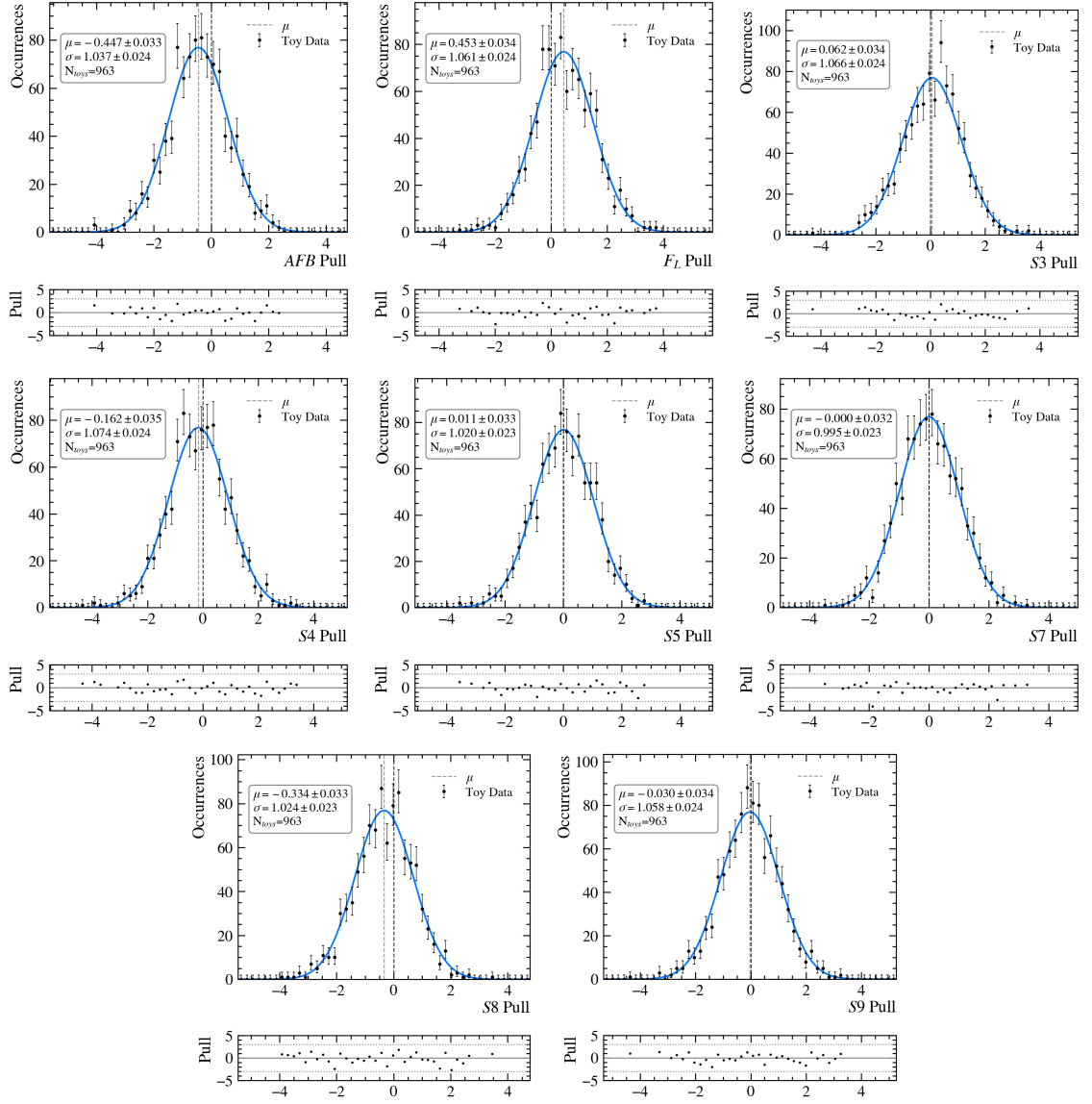


Figure D.4: Central q^2 : results of 1000 realistic toys simulating the number of events expected in data in the central q^2 region ($[1.1 - 6.0] \text{ GeV}^2/c^4$). For each toy, the number of events per component is sampled from a Poisson distribution with a mean equal to the corresponding yield reported in **Tab. 5.2**.

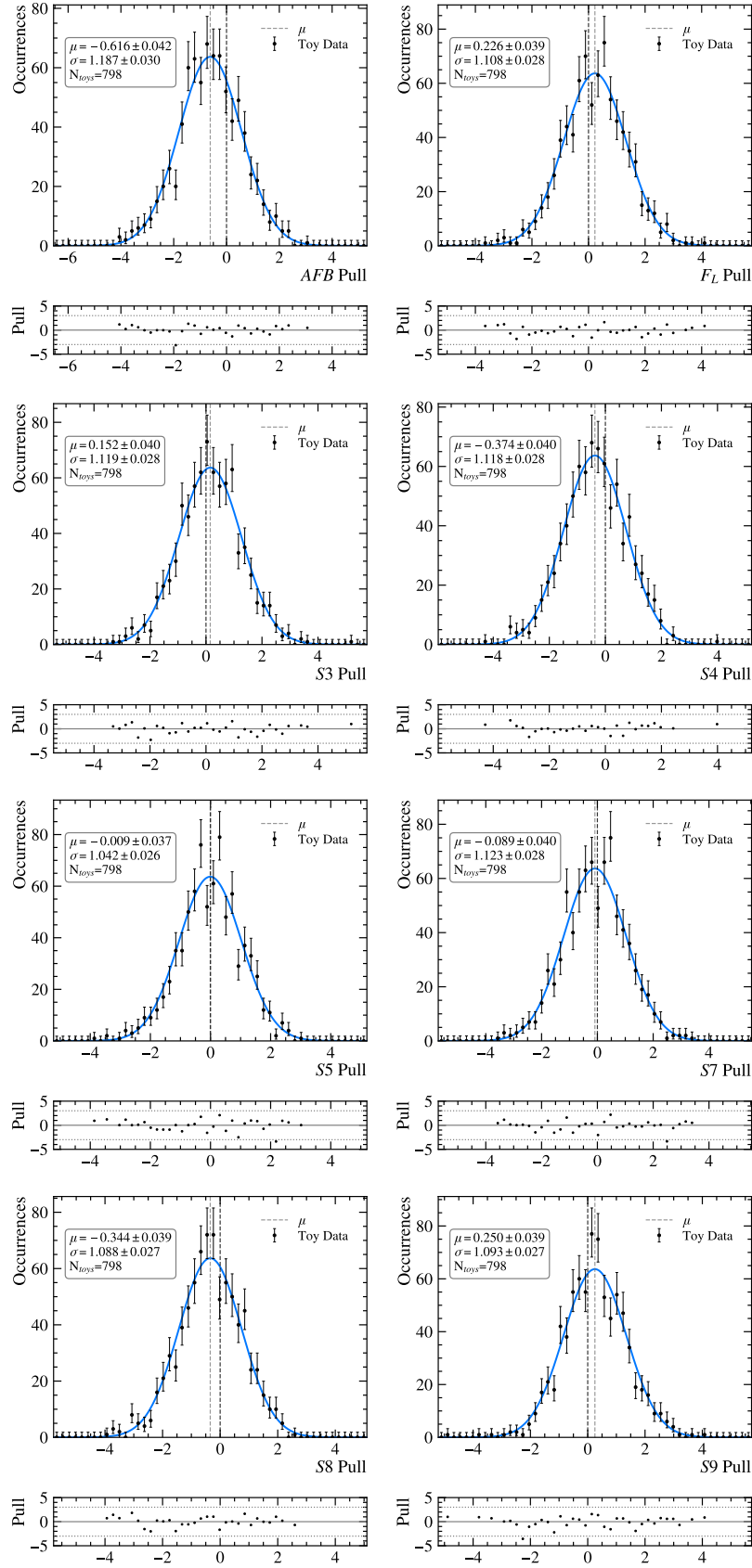


Figure D.5: Central_{low} q^2 : results of 1000 realistic toys simulating the number of events expected in data in the central q^2 region ($[1.10 - 3.65] \text{ GeV}^2/c^4$). For each toy, the number of events per component is sampled from a Poisson distribution with a mean equal to the corresponding yield reported in **Tab. 5.2**.

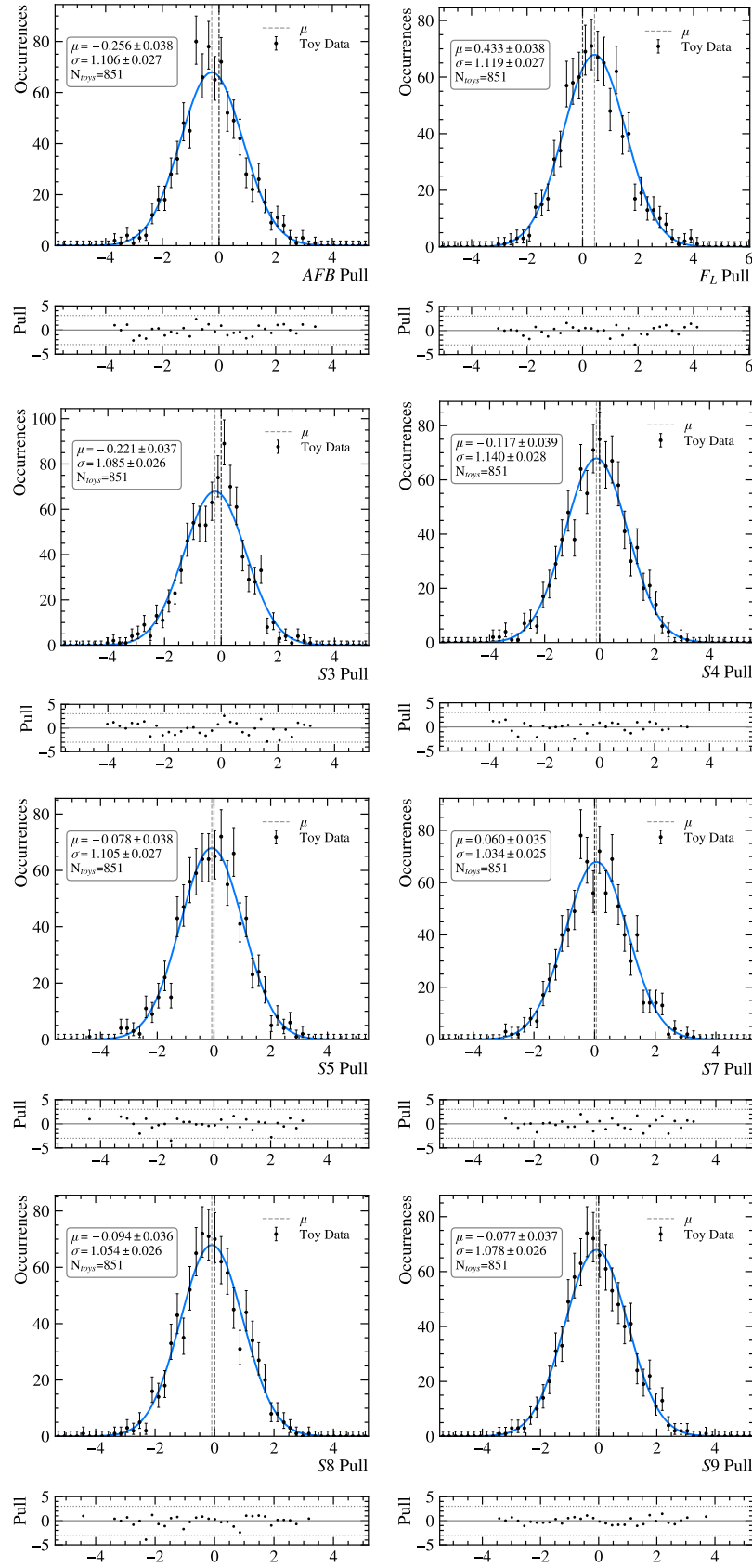


Figure D.6: Central_{high} q^2 : results of 1000 realistic toys simulating the number of events expected in data in the central q^2 region ($[3.65 - 6.00] \text{ GeV}^2/c^4$). For each toy, the number of events per component is sampled from a Poisson distribution with a mean equal to the corresponding yield reported in **Tab. 5.2**.

D.3 Blinded 4D-fit

Parameter	central	central _{low}	central _{high}
A_{FB}	$XX \pm 0.04$	$XX \pm 0.08$	$XX \pm 0.06$
F_L	$XX \pm 0.04$	$XX \pm 0.07$	$XX \pm 0.05$
S_3	$XX \pm 0.04$	$XX \pm 0.06$	$XX \pm 0.05$
S_4	$XX \pm 0.07$	$XX \pm 0.11$	$XX \pm 0.10$
S_5	$XX \pm 0.05$	$XX \pm 0.08$	$XX \pm 0.08$
S_7	$XX \pm 0.06$	$XX \pm 0.08$	$XX \pm 0.08$
S_8	$XX \pm 0.07$	$XX \pm 0.11$	$XX \pm 0.09$
S_9	$XX \pm 0.04$	$XX \pm 0.07$	$XX \pm 0.06$
$\alpha_{1,bkg} \cos \theta_\ell$ Run 1	0.99 ± 0.20	1.07 ± 0.29	1.12 ± 0.27
$\alpha_{2,bkg} \cos \theta_\ell$ Run 1	1.43 ± 0.25	0.45 ± 0.43	2.05 ± 0.31
$\alpha_{3,bkg} \cos \theta_\ell$ Run 1	0.30 ± 0.29	0.69 ± 0.55	0.35 ± 0.45
$\alpha_{4,bkg} \cos \theta_\ell$ Run 1	-0.18 ± 0.34	-	0.21 ± 0.49
$\alpha_{5,bkg} \cos \theta_\ell$ Run 1	-0.46 ± 0.39	-	-0.88 ± 0.44
$\alpha_{6,bkg} \cos \theta_\ell$ Run 1	-0.39 ± 0.38	-	-
$\alpha_{7,bkg} \cos \theta_\ell$ Run 1	-0.35 ± 0.38	-	-
$\alpha_{1,bkg} \cos \theta_\ell$ Run 2	1.01 ± 0.10	0.81 ± 0.13	1.15 ± 0.15
$\alpha_{2,bkg} \cos \theta_\ell$ Run 2	1.41 ± 0.13	0.41 ± 0.15	2.20 ± 0.17
$\alpha_{3,bkg} \cos \theta_\ell$ Run 2	0.09 ± 0.14	-0.16 ± 0.17	0.18 ± 0.23
$\alpha_{4,bkg} \cos \theta_\ell$ Run 2	-0.26 ± 0.17	-1.10 ± 0.21	0.34 ± 0.27
$\alpha_{5,bkg} \cos \theta_\ell$ Run 2	-1.15 ± 0.17	-1.36 ± 0.23	-1.15 ± 0.28
$\alpha_{6,bkg} \cos \theta_\ell$ Run 2	-0.76 ± 0.18	-0.99 ± 0.27	-0.78 ± 0.25
$\alpha_{7,bkg} \cos \theta_\ell$ Run 2	-0.50 ± 0.20	-0.52 ± 0.31	-0.81 ± 0.28
$\beta_{1,bkg} \cos \theta_K$ Run 1	-0.63 ± 0.16	-0.64 ± 0.27	-0.57 ± 0.23
$\beta_{2,bkg} \cos \theta_K$ Run 1	0.02 ± 0.24	0.24 ± 0.37	-0.02 ± 0.26
$\beta_{3,bkg} \cos \theta_K$ Run 1	-0.23 ± 0.26	0.08 ± 0.45	-0.17 ± 0.28
$\beta_{4,bkg} \cos \theta_K$ Run 1	0.38 ± 0.29	-	-
$\beta_{1,bkg} \cos \theta_K$ Run 2	-0.43 ± 0.09	-0.50 ± 0.14	-0.36 ± 0.13
$\beta_{2,bkg} \cos \theta_K$ Run 2	0.23 ± 0.12	0.45 ± 0.18	0.08 ± 0.17
$\beta_{3,bkg} \cos \theta_K$ Run 2	0.05 ± 0.14	-0.14 ± 0.21	0.24 ± 0.20
$\beta_{4,bkg} \cos \theta_K$ Run 2	-0.15 ± 0.17	-0.32 ± 0.23	0.00 ± 0.24
$\gamma_{1,bkg} \phi$ Run 1	0.08 ± 0.14	0.08 ± 0.22	0.06 ± 0.19
$\gamma_{2,bkg} \phi$ Run 1	0.24 ± 0.14	0.36 ± 0.23	0.19 ± 0.19
$\gamma_{3,bkg} \phi$ Run 1	-0.04 ± 0.14	0.25 ± 0.23	-0.11 ± 0.18
$\gamma_{4,bkg} \phi$ Run 1	-0.22 ± 0.14	-0.22 ± 0.22	-0.19 ± 0.18
$\gamma_{1,bkg} \phi$ Run 2	0.10 ± 0.08	0.12 ± 0.11	0.13 ± 0.10
$\gamma_{2,bkg} \phi$ Run 2	-0.01 ± 0.08	-0.08 ± 0.11	0.05 ± 0.11
$\gamma_{3,bkg} \phi$ Run 2	-0.04 ± 0.08	-0.02 ± 0.11	-0.07 ± 0.11
$\gamma_{4,bkg} \phi$ Run 2	-0.11 ± 0.07	-0.09 ± 0.10	-0.12 ± 0.10
DSL _{comb.} Run 1	50.16 ± 17.46	36.44 ± 9.74	20.29 ± 14.96
DSL _{comb.} Run 2	165.89 ± 32.09	71.67 ± 20.68	92.56 ± 24.37
J/ψ leak Run 1	5.01 ± 1.00	-	5.04 ± 1.00
J/ψ leak Run 2	33.02 ± 3.97	-	31.75 ± 3.95
Misid Run 1	72.11 ± 7.92	28.39 ± 5.24	41.51 ± 5.94
Misid Run 2	107.98 ± 7.02	50.15 ± 5.02	57.26 ± 4.91
Part. reco. Run 1	-3.30 ± 16.76	-11.48 ± 9.94	5.14 ± 14.50
Part. reco. Run 2	125.95 ± 31.75	94.04 ± 22.43	42.83 ± 22.46
Signal Run 1	82.75 ± 11.79	39.63 ± 7.97	43.80 ± 8.50
Signal Run 2	364.58 ± 23.01	159.66 ± 15.67	199.31 ± 16.62

Table D.1: Result of the simultaneous fit in the three q^2 regions under analysis.

E | Folding studies

This study has been performed to investigate the behaviour of the parameter that are removed when using the folded PDF. A likelihood scan has been carried out under the following scenarios:

- **Folding:** The likelihood profile is computed using the folded PDF.
- **Fix2Zero Params:** The full PDF is used to compute the likelihood profile, with the removed parameters fixed to zero.
- **Fix2GenVal Params:** The full PDF is used to compute the likelihood profile, with the removed parameters fixed to their generated values.
- **Free Params:** The full PDF is used to compute the likelihood profile.

The plots below compare the likelihood profiles of the angular parameters evaluated across the five folding schemes and for each of the scenarios considered. The results show that, for all parameters and folding types, not only the minimum of the likelihood but also the full likelihood profile is identical between the “folding” and “fixing to zero” scenarios. This indicates that, in these two scenarios, parameters with non-zero generated values would be subject to bias.

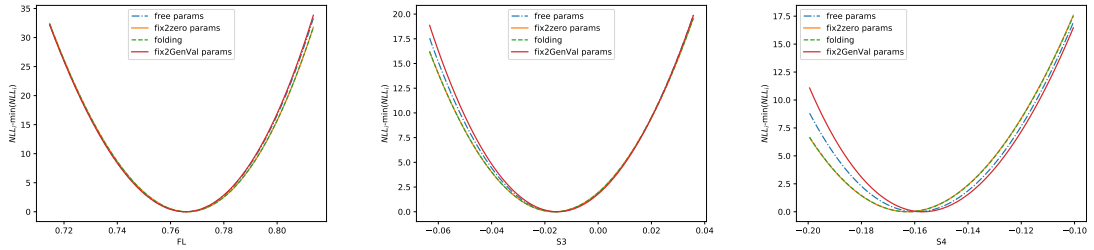


Figure E.1: S_4 folding, $[1,6] \text{ GeV}^2/c^4$

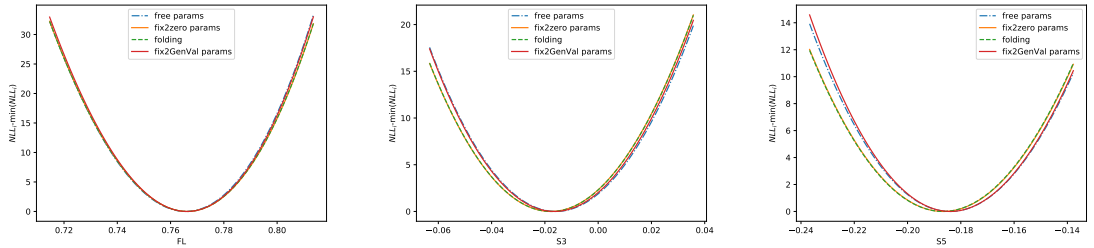
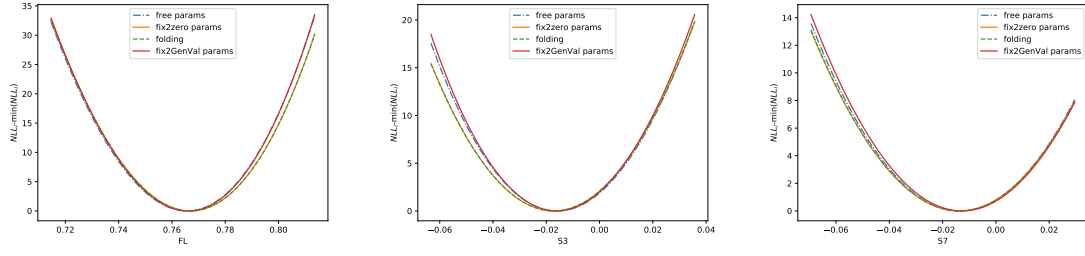
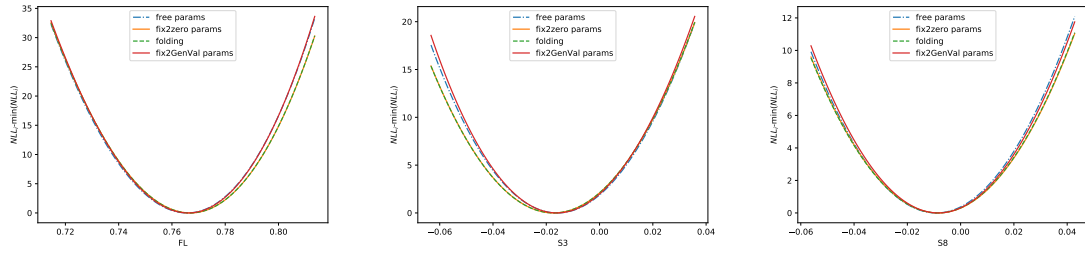
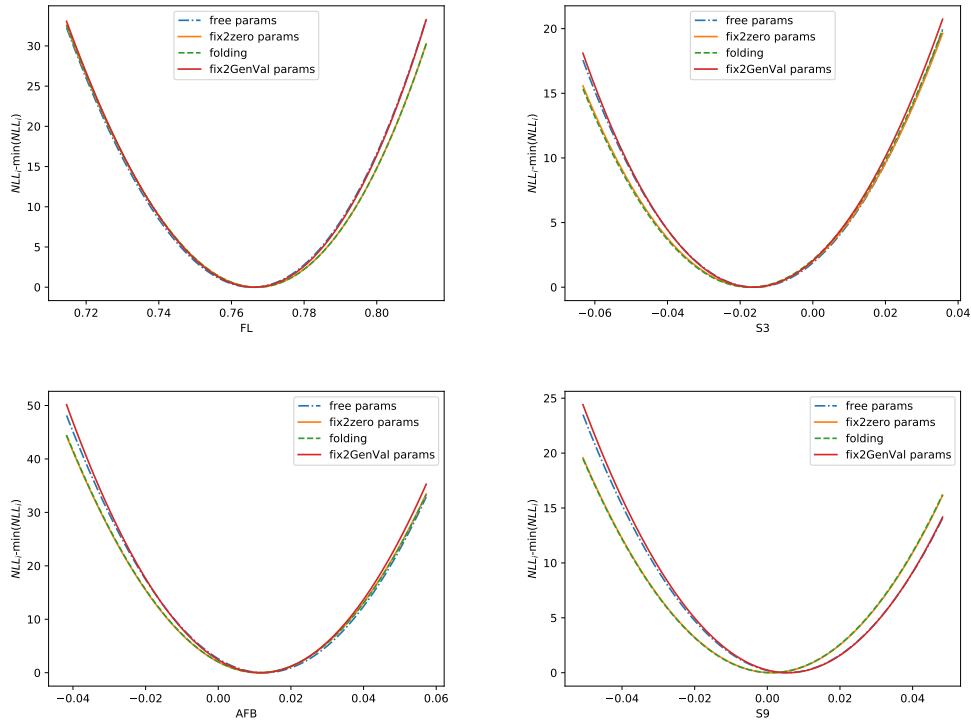


Figure E.2: S_5 folding, $[1,6] \text{ GeV}^2/c^4$

Figure E.3: S_7 folding, $[1,6] \text{ GeV}^2/c^4$ Figure E.4: S_8 folding, $[1,6] \text{ GeV}^2/c^4$ Figure E.5: S_9, A_{FB} folding, $[1,6] \text{ GeV}^2/c^4$

F | Alternative strategies

F.1 Selection on $\cos \theta_\ell$

The main objective of the second BDT is to target backgrounds that peak at large values of $\cos \theta_\ell$, a characteristic feature of both DSL and misidentified electron backgrounds, as illustrated in **Fig. F.1**. A straightforward strategy to suppress this peaking component would be to apply a direct cut on $\cos \theta_\ell$. This section explores the impact of such a cut, specifically $|\cos \theta_\ell| < 0.8$, which is highlighted in orange in **Fig. F.1**, and explains why the second BDT was ultimately preferred over this simpler approach. As a first step, the figure of merit defined in **Eq. 4.5** was used to verify

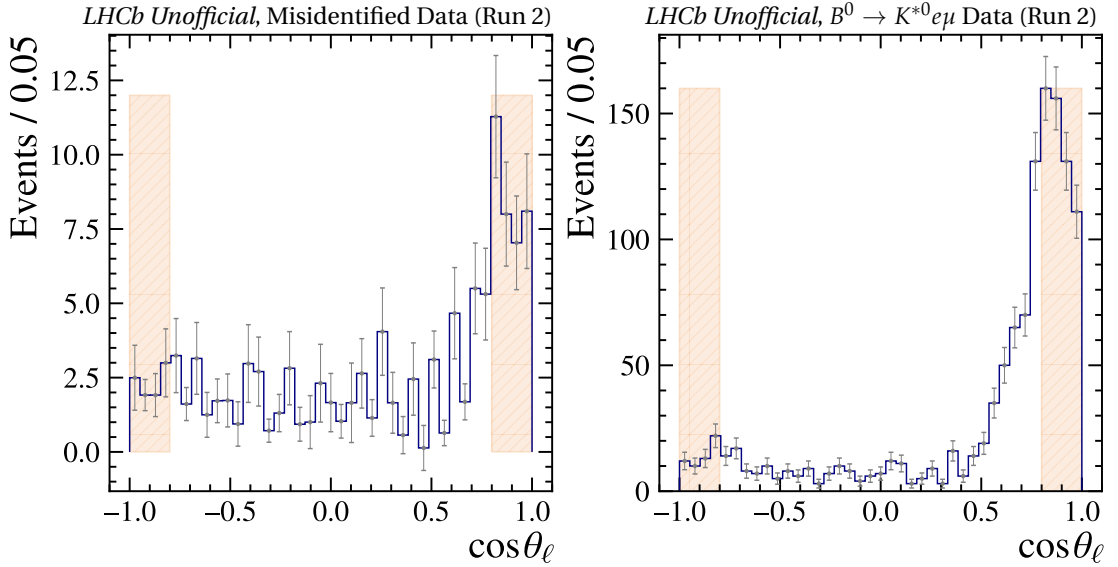


Figure F.1: $\cos \theta_\ell$ distribution of two final selected samples: expected misidentified background on the left and lepton-flavour-violating $B^0 \rightarrow K^{*0} e \mu$ data on the right. Orange areas indicate the regions removed by the $|\cos \theta_\ell| < 0.8$ cut.

whether the optimal BDT1 threshold changed after applying the $\cos \theta_\ell$ cut. The optimal value was found to remain at $BDT1 > 0.96$. Based on this, the alternative configuration was compared to the standard one as follows:

- Standard (S): $BDT1 > 0.96$ and $BDT2 > 0.20$ with figure-of merit $FoM = 302.1 \pm 0.8$ and signal efficiency $\epsilon_{sig} = 76\%$.
- Alternative (A): $BDT1 > 0.96$ and $|\cos \theta_\ell| < 0.8$ with figure-of-merit $FoM = 295.1 \pm 0.8$ and

signal efficiency $\epsilon_{sig} = 72\%$.

To compare these configurations, pseudoexperiments were performed. Specifically, 1000 toys were generated using configuration S and fitted using the standard fitting strategy described in **Chap. 5**. Another 1000 toys were generated with only the $BDT1 > 0.96$ requirement, after which the $|\cos \theta_\ell| < 0.8$ cut was applied. These were then fitted following the same strategy, with shapes derived from samples selected with both $BDT1 > 0.96$ and $|\cos \theta_\ell| < 0.8$. These toys simulate the Run 2 data only. The outcomes of this comparison are presented in **Tab. F.1** and **Tab. F.2** for the $central_{low} q^2$ and $central_{high} q^2$ regions, respectively. Each table reports the pulls mean (Bias), the mean statistical uncertainties from the fit (Fit error), the bias increase (in units of sigma), and the relative increase in statistical uncertainty of the alternative configuration compared to the standard. The bias increase is calculated as the difference between the absolute values of the pulls obtained in the alternative and standard configurations. The error associated is the square root of the sum in quadrature of the uncertainties on the means from the two biases. The relative increase in statistical uncertainty is computed as the difference between the statistical errors from the fit for the alternative and standard configurations, divided by the statistical error in the alternative configuration.

The results indicate a consistent trend of increased biases and larger statistical uncertainties. The latter increase was expected due to the slight loss in signal efficiency of the alternative configuration and because $\cos \theta_\ell$ is a fit variable — imposing a cut on it naturally reduces the sensitivity to angular parameters.

Based on these considerations, the standard configuration has been adopted throughout this dissertation.

central _{low} q^2						
Param	Standard		Alternative		Alternative vs Standard	
	Bias	Fit Error	Bias	Fit Error	Bias increase	Error increase
A_{FB}	-0.55 ± 0.04	0.08	-0.61 ± 0.04	0.11	$+0.07 \pm 0.05$	+28%
F_L	0.33 ± 0.04	0.08	0.93 ± 0.03	0.10	$+0.60 \pm 0.05$	+22%
S_3	-0.14 ± 0.04	0.08	-0.22 ± 0.04	0.10	$+0.09 \pm 0.05$	+18%
S_4	-0.18 ± 0.04	0.12	-0.19 ± 0.04	0.15	$+0.01 \pm 0.05$	+21%
S_5	-0.05 ± 0.04	0.10	-0.10 ± 0.04	0.12	$+0.05 \pm 0.05$	+16%
S_7	-0.17 ± 0.04	0.10	-0.01 ± 0.04	0.12	-0.16 ± 0.05	+17%
S_8	-0.47 ± 0.04	0.12	-0.26 ± 0.04	0.15	-0.11 ± 0.05	+21%
S_9	0.31 ± 0.04	0.8	-0.00 ± 0.04	0.10	-0.31 ± 0.05	+20%

Table F.1: Pulls mean (Bias), the mean statistical uncertainties from the fit (Fit error), the bias increase, and the relative increase in statistical uncertainty (Error increase) of the alternative configuration (A) compared to the standard (S). For their computation refer to the text. The errors associated to the Biases are the errors on the mean of the Gauss function used to fit the pull distributions. Results obtained from 1000 toys simulating Run 2 central_{low} q^2 .

Param	central _{low} q^2					
	Standard		Alternative		Alternative vs Standard	
	Bias	Fit Error	Bias	Fit Error	Bias increase	Error increase
	A	S	A	S		
A_{FB}	-0.20 ± 0.04	0.06	-0.39 ± 0.04	0.08	$+0.19 \pm 0.05$	+34%
F_L	0.53 ± 0.03	0.06	0.90 ± 0.03	0.08	$+0.37 \pm 0.05$	+26%
S_3	-0.08 ± 0.04	0.06	-0.29 ± 0.04	0.12	$+0.22 \pm 0.05$	+19%
S_4	-0.08 ± 0.04	0.09	-0.31 ± 0.04	0.09	$+0.24 \pm 0.05$	+23%
S_5	-0.10 ± 0.04	0.08	-0.41 ± 0.04	0.10	$+0.32 \pm 0.05$	+15%
S_7	-0.02 ± 0.04	0.08	0.04 ± 0.04	0.00	$+0.02 \pm 0.05$	+16%
S_8	-0.02 ± 0.04	0.10	-0.06 ± 0.04	0.12	$+0.04 \pm 0.05$	+23%
S_9	-0.05 ± 0.04	0.06	0.04 ± 0.04	0.08	-0.01 ± 0.06	+21%

Table F.2: Pulls mean (Bias), the mean statistical uncertainties from the fit (Fit error), the bias increase, and the relative increase in statistical uncertainty (Error increase) of the alternative configuration (A) compared to the standard (S). For their computation refer to the text. The errors associated to the Biases are the errors on the mean of the Gauss function used to fit the pull distributions. Results obtained from 1000 toys simulating Run 2 central_{high} q^2 .

G | Physical regions of the angular parameters

This appendix presents the 2-dimensional projections of the physically allowed regions for the angular parameters. The blue areas indicate regions where the angular PDF is positive - *i.e.*, where the negative log - likelihood (NLL) is defined. The red dot corresponds to the Standard Model prediction for the angular parameters.

These regions are obtained using the SM predictions provided by the FLAVIO [27] package. All angular parameters are fixed to their SM values, except for those being scanned. The scanned parameters vary from -1 to 1 , except for F_L , which varies from 0 to 1 .

This procedure is repeated for the central_{low} and central_{high} q^2 bins separately, as well as for the full central q^2 bin. The projections reveal that the SM prediction lies at the edge of the physical region when F_L and A_{FB} are involved. This proximity to the boundary can induce a fit bias due to physical constraints, as the results from pseudo-experiments tend to be pulled toward the allowed region.

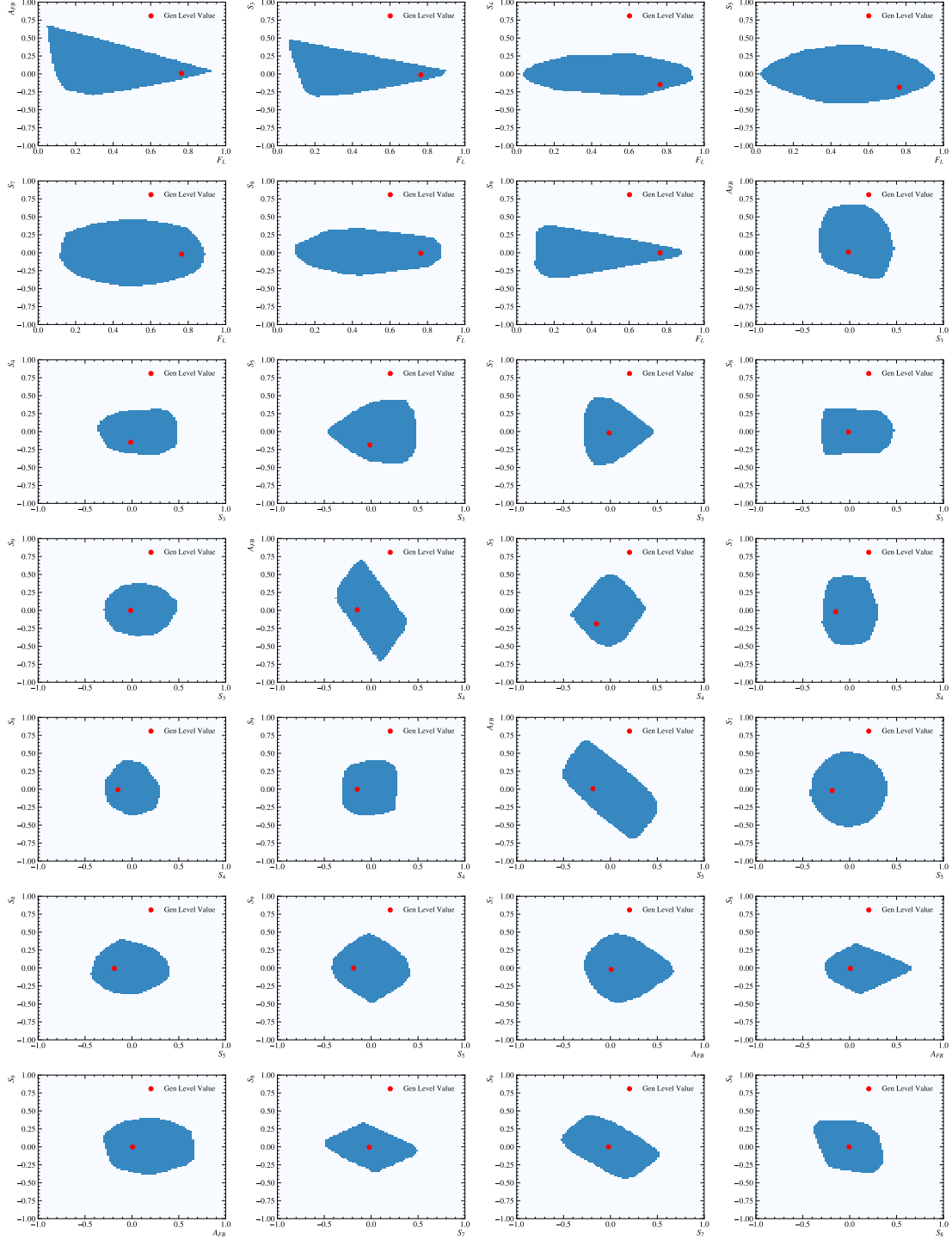


Figure G.1: Physical regions of the angular parameter in the full central q^2 region, $[1.1, 6.0] \text{ GeV}^2/c^4$.

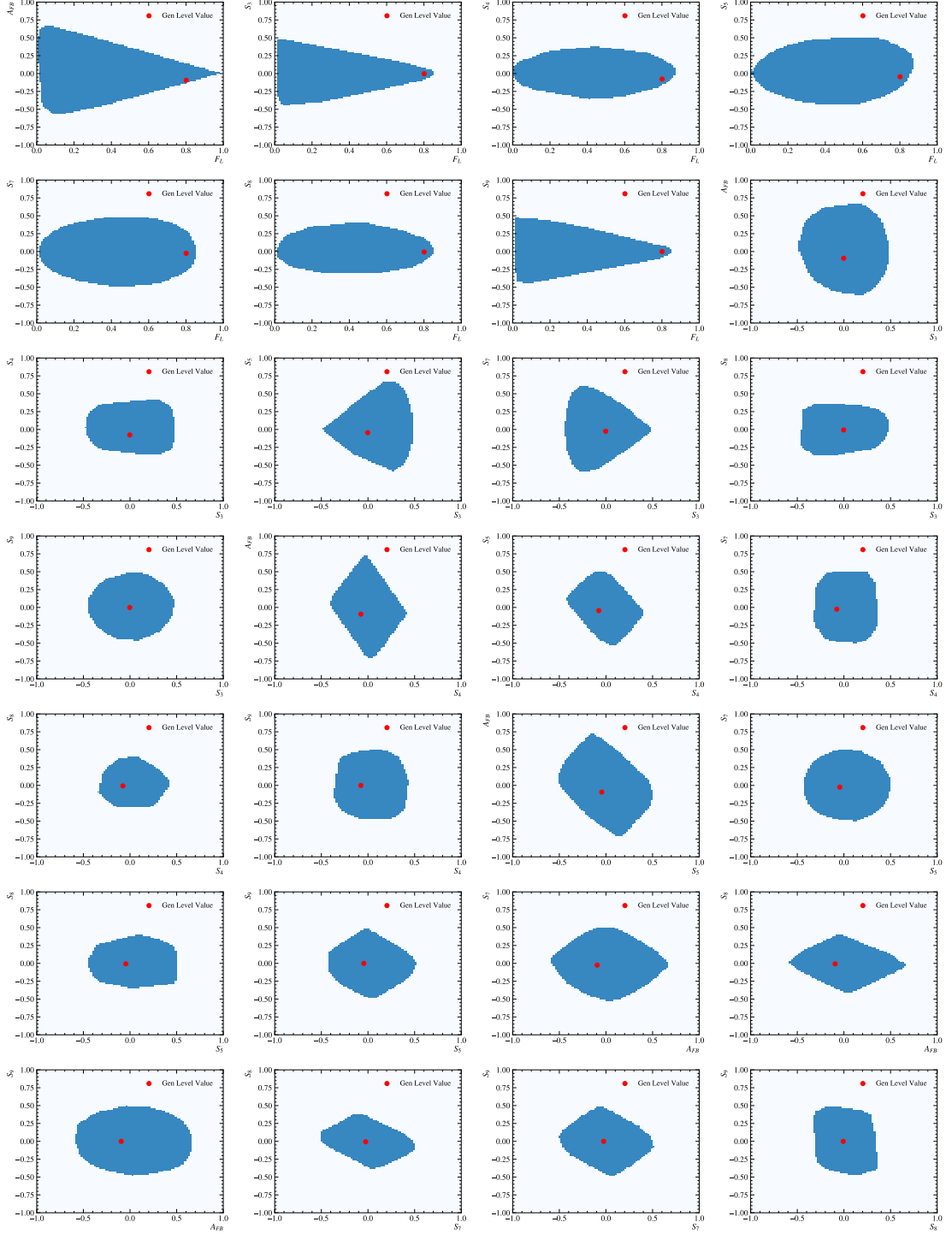


Figure G.2: Physical regions of the angular parameter in the central_{low} q^2 region, $[1.10, 3.65] \text{ GeV}^2/c^4$.

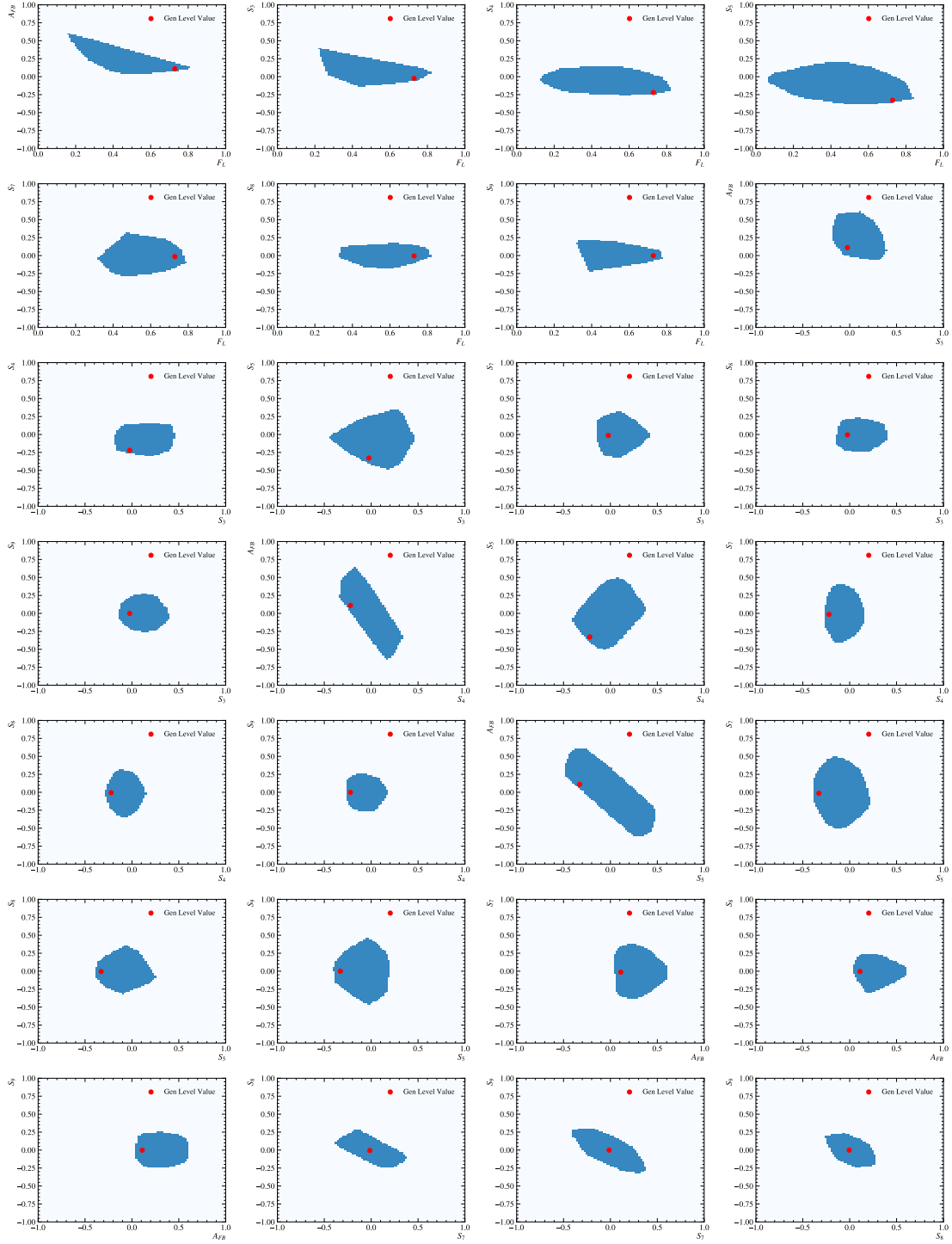


Figure G.3: Physical regions of the angular parameter in the central_{high} q^2 region, $[3.65, 6.00] \text{ GeV}^2/c^4$.

Bibliography

- [1] J. R. Ellis, M. K. Gaillard, D. V. Nanopoulos, and S. Rudaz, *The Phenomenology of the Next Left-Handed Quarks*, [Nucl. Phys. B **131** \(1977\) 285](#), [Erratum: Nucl.Phys.B 132, 541 (1978)].
- [2] Y. Nir, *CP violation in and beyond the standard model*, in *27th SLAC Summer Institute on Particle Physics: CP Violation in and Beyond the Standard Model*, pp. 165–243, 7, 1999. [arXiv:hep-ph/9911321](#).
- [3] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, [Phys. Rev. Lett. **10** \(1963\) 531](#).
- [4] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, [Prog. Theor. Phys. **49** \(1973\) 652](#).
- [5] Particle Data Group, S. Navas *et al.*, *Review of particle physics*, [Phys. Rev. D **110** \(2024\), no. 3 030001](#).
- [6] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, [Phys. Rev. D **2** \(1970\) 1285](#).
- [7] T. Blake, G. Lanfranchi, and D. M. Straub, *Rare B Decays as Tests of the Standard Model*, [Prog. Part. Nucl. Phys. **92** \(2017\) 50](#), [arXiv:1606.00916](#).
- [8] W. Altmannshofer *et al.*, *Symmetries and Asymmetries of $B \rightarrow K^* \mu^+ \mu^-$ Decays in the Standard Model and Beyond*, [JHEP **01** \(2009\) 019](#), [arXiv:0811.1214](#).
- [9] F. Beaujean, C. Bobeth, D. van Dyk, and C. Wacker, *Bayesian Fit of Exclusive $b \rightarrow s \bar{\ell} \ell$ Decays: The Standard Model Operator Basis*, [JHEP **08** \(2012\) 030](#), [arXiv:1205.1838](#).
- [10] LHCb, R. Aaij *et al.*, *Angular analysis of the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decay using 3 fb^{-1} of integrated luminosity*, [JHEP **02** \(2016\) 104](#), [arXiv:1512.04442](#).
- [11] S. Descotes-Genon, L. Hofer, J. Matias, and J. Virto, *On the impact of power corrections in the prediction of $B \rightarrow K^* \mu^+ \mu^-$ observables*, [JHEP **12** \(2014\) 125](#), [arXiv:1407.8526](#).
- [12] T. Hurth, C. Langenbruch, and F. Mahmoudi, *Direct determination of Wilson coefficients using $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ decays*, [JHEP **11** \(2017\) 176](#), [arXiv:1708.04474](#).
- [13] M. Algueró *et al.*, *A complete description of P- and S-wave contributions to the $B^0 \rightarrow K^+ \pi^- \ell^+ \ell^-$ decay*, [JHEP **12** \(2021\) 085](#), [arXiv:2107.05301](#).
- [14] S. Descotes-Genon, T. Hurth, J. Matias, and J. Virto, *Optimizing the basis of $B \rightarrow K^* \ell \ell$ observables in the full kinematic range*, [JHEP **05** \(2013\) 137](#), [arXiv:1303.5794](#).

- [15] LHCb, R. Aaij *et al.*, *Measurement of the b -quark production cross-section in 7 and 13 TeV pp collisions*, *Phys. Rev. Lett.* **118** (2017), no. 5 052002, [arXiv:1612.05140](#), [Erratum: *Phys.Rev.Lett.* 119, 169901 (2017)].
- [16] [LHCb Operations Plots Webpage](#).
- [17] LHCb, R. Aaij *et al.*, *Measurements of the S -wave fraction in $B^0 \rightarrow K^+ \pi^- \mu^+ \mu^-$ decays and the $B^0 \rightarrow K^*(892)^0 \mu^+ \mu^-$ differential branching fraction*, *JHEP* **11** (2016) 047, [arXiv:1606.04731](#), [Erratum: *JHEP* 04, 142 (2017)].
- [18] LHCb, 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), no. 15 151801, [arXiv:2105.14007](#).
- [19] LHCb, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetries of $B \rightarrow K^{(*)} \mu^+ \mu^-$ decays*, *JHEP* **06** (2014) 133, [arXiv:1403.8044](#).
- [20] HPQCD, W. G. Parrott, C. Bouchard, and C. T. H. Davies, *Standard Model predictions for $B \rightarrow K \ell^+ \ell^-$, $B \rightarrow K \ell^+ \ell^-$ and $B \rightarrow K \nu \nu$ using form factors from $N_f=2+1+1$ lattice QCD*, *Phys. Rev. D* **107** (2023), no. 1 014511, [arXiv:2207.13371](#), [Erratum: *Phys.Rev.D* 107, 119903 (2023)].
- [21] LHCb, R. Aaij *et al.*, *Measurement of the isospin asymmetry in $B \rightarrow K^{(*)} \mu^+ \mu^-$ decays*, *JHEP* **07** (2012) 133, [arXiv:1205.3422](#).
- [22] 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](#).
- [23] BaBar, J. P. Lees *et al.*, *Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays $B \rightarrow K^{(*)} l^+ l^-$* , *Phys. Rev. D* **86** (2012) 032012, [arXiv:1204.3933](#).
- [24] Belle, J.-T. Wei *et al.*, *Measurement of the Differential Branching Fraction and Forward-Backward Asymmetry for $B \rightarrow K^{(*)} \ell^+ \ell^-$* , *Phys. Rev. Lett.* **103** (2009) 171801, [arXiv:0904.0770](#).
- [25] CMS, V. Khachatryan *et al.*, *Angular analysis of the decay $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ from pp collisions at $\sqrt{s} = 8$ TeV*, *Phys. Lett. B* **753** (2016) 424, [arXiv:1507.08126](#).
- [26] LHCb, R. Aaij *et al.*, *Test of Lepton Flavor Universality with $B_s^0 \rightarrow \phi \ell^+ \ell^-$ Decays*, *Phys. Rev. Lett.* **134** (2025), no. 12 121803, [arXiv:2410.13748](#).
- [27] D. M. Straub, *flavio: a Python package for flavour and precision phenomenology in the Standard Model and beyond*, [arXiv:1810.08132](#).
- [28] 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](#).
- [29] 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](#).
- [30] LHCb, R. Aaij *et al.*, *Measurement of CP-Averaged Observables in the $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ Decay*, *Phys. Rev. Lett.* **125** (2020), no. 1 011802, [arXiv:2003.04831](#).

- [31] LHCb, R. Aaij *et al.*, *Angular analysis of the rare decay $B_s^0 \rightarrow \phi \mu^+ \mu^-$* , *JHEP* **11** (2021) 043, [arXiv:2107.13428](#).
- [32] LHCb, R. Aaij *et al.*, *Angular Analysis of the $B^+ \rightarrow K^{*+} \mu^+ \mu^-$ Decay*, *Phys. Rev. Lett.* **126** (2021), no. 16 161802, [arXiv:2012.13241](#).
- [33] ATLAS, 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](#).
- [34] CMS, *Measurement of the P_1 and P_5' angular parameters of the decay $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ in proton-proton collisions at $\sqrt{s} = 8$ TeV*, .
- [35] Belle, S. Wehle *et al.*, *Lepton-Flavor-Dependent Angular Analysis of $B \rightarrow K^* \ell^+ \ell^-$* , *Phys. Rev. Lett.* **118** (2017), no. 11 111801, [arXiv:1612.05014](#).
- [36] LHCb, F. Dettori, *Search for new physics in $b \rightarrow q \ell \ell$ decays*, in *53rd Rencontres de Moriond on Electroweak Interactions and Unified Theories*, pp. 47–52, 2018. [arXiv:1805.05073](#).
- [37] CMS, 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. B* **864** (2025) 139406, [arXiv:2411.11820](#).
- [38] A. Khodjamirian, T. Mannel, A. A. Pivovarov, and Y.-M. Wang, *Charm-loop effect in $B \rightarrow K^{(*)} \ell^+ \ell^-$ and $B \rightarrow K^* \gamma$* , *JHEP* **09** (2010) 089, [arXiv:1006.4945](#).
- [39] LHCb, R. Aaij *et al.*, *Angular analysis of $B^0 \rightarrow K^{*0} e^+ e^-$ decays*, [arXiv:2502.10291](#).
- [40] M. Algueró *et al.*, *To (b)e or not to (b)e: no electrons at LHCb*, *Eur. Phys. J. C* **83** (2023), no. 7 648, [arXiv:2304.07330](#).
- [41] N. Gubernari, M. Reboud, D. van Dyk, and J. Virto, *Improved theory predictions and global analysis of exclusive $b \rightarrow s \mu^+ \mu^-$ processes*, *JHEP* **09** (2022) 133, [arXiv:2206.03797](#).
- [42] EOS Authors, D. van Dyk *et al.*, *EOS: a software for flavor physics phenomenology*, *Eur. Phys. J. C* **82** (2022), no. 6 569, [arXiv:2111.15428](#).
- [43] LHCb, R. Aaij *et al.*, *Constraints on the photon polarisation in $b \rightarrow s \gamma$ transitions using $B_s^0 \rightarrow \phi e^+ e^-$ decays*, *JHEP* **03** (2025) 047, [arXiv:2411.10219](#).
- [44] [Conference Talk at “Moriond EW ’25”, L. Paolucci](#).
- [45] LHCb, LHCb-PAPER-2025-006, in preparation.
- [46] M. Bordone, G. Isidori, and A. Pattori, *On the Standard Model predictions for R_K and R_{K^*}* , *Eur. Phys. J. C* **76** (2016), no. 8 440, [arXiv:1605.07633](#).
- [47] LHCb, R. Aaij *et al.*, *Test of lepton universality with $\Lambda_b^0 \rightarrow p K^- \ell^+ \ell^-$ decays*, *JHEP* **05** (2020) 040, [arXiv:1912.08139](#).
- [48] LHCb, R. Aaij *et al.*, *Tests of lepton universality using $B^0 \rightarrow K_S^0 \ell^+ \ell^-$ and $B^+ \rightarrow K^{*+} \ell^+ \ell^-$ decays*, *Phys. Rev. Lett.* **128** (2022), no. 19 191802, [arXiv:2110.09501](#).
- [49] LHCb, 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. D* **108** (2023), no. 3 032002, [arXiv:2212.09153](#).

- [50] LHCb, R. Aaij *et al.*, *Test of lepton universality in $b \rightarrow s\ell^+\ell^-$ decays*, [Phys. Rev. Lett. **131** \(2023\), no. 5 051803](#), [arXiv:2212.09152](#).
- [51] BELLE, S. Choudhury *et al.*, *Test of lepton flavor universality and search for lepton flavor violation in $B \rightarrow K\ell\ell$ decays*, [JHEP **03** \(2021\) 105](#), [arXiv:1908.01848](#).
- [52] Belle, A. Abdesselam *et al.*, *Test of Lepton-Flavor Universality in $B \rightarrow K^*\ell^+\ell^-$ Decays at Belle*, [Phys. Rev. Lett. **126** \(2021\), no. 16 161801](#), [arXiv:1904.02440](#).
- [53] A. Greljo, J. Salko, A. Smolkovič, and P. Stangl, *Rare b decays meet high-mass Drell-Yan*, [JHEP **05** \(2023\) 087](#), [arXiv:2212.10497](#).
- [54] M. Ciuchini *et al.*, *Constraints on lepton universality violation from rare B decays*, [Phys. Rev. D **107** \(2023\), no. 5 055036](#), [arXiv:2212.10516](#).
- [55] B. Capdevila, A. Crivellin, and J. Matias, *Review of semileptonic B anomalies*, [Eur. Phys. J. ST **1** \(2023\) 20](#), [arXiv:2309.01311](#).
- [56] M. Algueró *et al.*, *Are we overlooking lepton flavour universal new physics in $b \rightarrow s\ell\ell$?*, [Phys. Rev. D **99** \(2019\), no. 7 075017](#), [arXiv:1809.08447](#).
- [57] T. Hurth, F. Mahmoudi, and S. Neshatpour, *B anomalies in the post $R_{K^{(*)}}$ era*, [Phys. Rev. D **108** \(2023\), no. 11 115037](#), [arXiv:2310.05585](#).
- [58] T. Hurth, F. Mahmoudi, D. M. Santos, and S. Neshatpour, *More Indications for Lepton Nonuniversality in $b \rightarrow s\ell^+\ell^-$* , [Phys. Lett. B **824** \(2022\) 136838](#), [arXiv:2104.10058](#).
- [59] LHCb, R. Aaij *et al.*, *Comprehensive analysis of local and nonlocal amplitudes in the $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decay*, [JHEP **09** \(2022\)](#), [arXiv:2405.17347](#).
- [60] C. Bobeth, M. Misiak, and J. Urban, *Photonic penguins at two loops and m_t dependence of $BR[B \rightarrow X_s l^+ l^-]$* , [Nucl. Phys. B **574** \(2000\) 291](#), [arXiv:hep-ph/9910220](#).
- [61] M. Gorbahn and U. Haisch, *Effective Hamiltonian for non-leptonic $|\Delta F| = 1$ decays at NNLO in QCD*, [Nucl. Phys. B **713** \(2005\) 291](#), [arXiv:hep-ph/0411071](#).
- [62] M. Algueró, J. Matias, B. Capdevila, and A. Crivellin, *Disentangling lepton flavor universal and lepton flavor universality violating effects in $b \rightarrow s\ell^+\ell^-$ transitions*, [Phys. Rev. D **105** \(2022\), no. 11 113007](#), [arXiv:2205.15212](#).
- [63] *LHC Machine*, [JINST **3** \(2008\) S08001](#).
- [64] Picture from CERN Accelerating science, home.cern/science/accelerators/accelerator-complex.
- [65] W. Herr and B. Muratori, *Concept of luminosity*, in *CERN Accelerator School and DESY Zeuthen: Accelerator Physics*, pp. 361–377, 9, 2003.
- [66] B. Muratori and T. Pieloni, *Luminosity levelling techniques for the LHC*, in *ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders*, pp. 177–181, 2014. [arXiv:1410.5646](#). doi: [10.5170/CERN-2014-004.177](https://doi.org/10.5170/CERN-2014-004.177).
- [67] LHCb, P. R. Barbosa-Marinho *et al.*, *LHCb VELO (Vertex LOCator): Technical Design Report*, Tech. Rep. CERN-LHCC-2001-011, CERN, Geneva, 2001.

- [68] LHCb, S. Amato *et al.*, *LHCb RICH: Technical Design Report*, Tech. Rep. CERN-LHCC-2000-037, CERN, Geneva, 2000.
- [69] LHCb, R. Antunes-Nobrega *et al.*, *LHCb reoptimized detector design and performance: Technical Design Report*, Tech. Rep. CERN-LHCC-2003-030, CERN, Geneva, 2003.
- [70] LHCb, P. R. Barbosa-Marinho *et al.*, *LHCb: Inner tracker technical design report*, Tech. Rep. CERN-LHCC-2002-029, CERN, 11, 2002.
- [71] LHCb, S. Amato *et al.*, *LHCb magnet: Technical Design Report*, Tech. Rep. CERN-LHCC-2000-007, CERN, Geneva, 2000.
- [72] LHCb Outer Tracker Group, R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, [JINST 9 \(2014\), no. 01 P01002](#), [arXiv:1311.3893](#).
- [73] LHCb, S. Amato *et al.*, *LHCb calorimeters: Technical Design Report*, Tech. Rep. CERN-LHCC-2000-036, CERN, Geneva, 2000.
- [74] LHCb, P. R. Barbosa-Marinho *et al.*, *LHCb muon system: Technical Design Report*, Tech. Rep. CERN-LHCC-2001-010, CERN, Geneva, 2001.
- [75] LHCb, *LHCb muon system: addendum to the Technical Design Report*, Tech. Rep. CERN-LHCC-2003-002, CERN, Geneva, 2003.
- [76] LHCb, *LHCb muon system: second addendum to the Technical Design Report*, Tech. Rep. CERN-LHCC-2005-012, CERN, Geneva, 2005. Submitted on 9 Apr 2005.
- [77] LHCb RICH Group, M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, [Eur. Phys. J. C 73 \(2013\) 2431](#), [arXiv:1211.6759](#).
- [78] LHCb, E. Picatoste Olloqui, *LHCb preshower(PS) and scintillating pad detector (SPD): Commissioning, calibration, and monitoring*, [J. Phys. Conf. Ser. 160 \(2009\) 012046](#).
- [79] I. Machikhiliyan *et al.*, *Current status and performance of the lhcb electromagnetic and hadron calorimeters*, in *Journal of Physics: Conference Series*, vol. 293, p. 012052, IOP Publishing, 2011.
- [80] J. Alves, A.Ã. *et al.*, *Performance of the LHCb muon system*, [JINST 8 \(2013\) P02022](#), [arXiv:1211.1346](#).
- [81] A. Hoecker *et al.*, *Tmva - toolkit for multivariate data analysis*, 2009.
- [82] T. Head, *The LHCb trigger system*, [JINST 9 \(2014\) C09015](#).
- [83] S. Benson, V. Gligorov, M. A. Vesterinen, and J. M. Williams, *The lhcb turbo stream*, in *J. Phys. Conf. Ser.*, vol. 664, p. 082004, 2015.
- [84] LHCb, R. Aaij *et al.*, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, [JINST 14 \(2019\), no. 04 P04013](#), [arXiv:1812.10790](#).
- [85] M. Borsato, *Electron reconstruction at lhcb and belle ii*, [URL](#).

- [86] M. van Veghel, *Pursuing forbidden beauty: Search for the lepton-flavour violating decays $B^0 \rightarrow e^\pm \mu^\mp$ and $B_s^0 \rightarrow e^\pm \mu^\mp$ and study of electron-reconstruction performance at LHCb*, PhD thesis, Groningen University, 2020, [URL](#).
- [87] LHCb, R. Aaij *et al.*, *Measurement of the electron reconstruction efficiency at LHCb*, [JINST](#) **14** (2019) P11023, [arXiv:1909.02957](#).
- [88] T. Likhomanenko *et al.*, *LHCb Topological Trigger Reoptimization*, [J. Phys. Conf. Ser.](#) **664** (2015), no. 8 082025, [arXiv:1510.00572](#).
- [89] <https://twiki.cern.ch/twiki/bin/view/LHCb/DecayTreeFitter>.
- [90] <https://twiki.cern.ch/twiki/bin/view/LHCb/VertexFitters>.
- [91] J. A. N. van Tilburg, *Track simulation and reconstruction in LHCb*, PhD thesis, Vrije Universiteit, 2005, [URL](#).
- [92] LHCb, *LHCb Upgrade GPU High Level Trigger Technical Design Report*, .
- [93] M. Williams *et al.*, *The HLT2 Topological Lines*, .
- [94] LHCb, 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), no. 04 P04013, [arXiv:1812.10790](#).
- [95] A. Snoch, *Angles and B-mons: Angular analysis of $B^0 \rightarrow K^{*0} J/\psi (\rightarrow e^+ e^-)$ decays with LHCb data*, PhD thesis, University of Groningen, 2024, [URL](#).
- [96] LHCb, R. Aaij *et al.*, *Test of lepton universality with $B^0 \rightarrow K^{*0} \ell^+ \ell^-$ decays*, [JHEP](#) **08** (2017) 055, [arXiv:1705.05802](#).
- [97] E. Rodrigues, *Dealing with clones in the tracking*, [CERN-LHCb-2006-057](#).
- [98] N. Oskam, *Hadronic backgrounds in $B^0 \rightarrow K^{*0} e^+ e^-$ decays*, Bachelor Thesis, Universiteit Van Amsterdam, Netherlands, 2023, [URL](#).
- [99] M. Geijssen, *Measurement of the angular decay rate of $B^0 \rightarrow K^{*0} e^+ e^-$ using the full available LHCb data sample (Runs 1 and 2)*, Master Thesis, Universiteit Van Amsterdam, Netherlands, 2024.
- [100] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, [JHEP](#) **05** (2006) 026, [arXiv:hep-ph/0603175](#).
- [101] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, [Comput. Phys. Commun.](#) **178** (2008) 852, [arXiv:0710.3820](#).
- [102] LHCb, I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, in *2010 IEEE Nuclear Science Symposium, Medical Imaging Conference, and 17th Room Temperature Semiconductor Detectors Workshop*, pp. 1155–1161, 2010. doi: [10.1109/NSSMIC.2010.5873949](#).
- [103] D. J. Lange, *The EvtGen particle decay simulation package*, [Nucl. Instrum. Meth. A](#) **462** (2001) 152.

- [104] N. Davidson, T. Przedzinski, and Z. Was, *PHOTOS interface in C++: Technical and Physics Documentation*, *Comput. Phys. Commun.* **199** (2016) 86, [arXiv:1011.0937](#).
- [105] J. Allison *et al.*, *Geant4 developments and applications*, *IEEE Trans. Nucl. Sci.* **53** (2006) 270.
- [106] GEANT4, S. Agostinelli *et al.*, *GEANT4 - A Simulation Toolkit*, *Nucl. Instrum. Meth. A* **506** (2003) 250.
- [107] LHCb, M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, *J. Phys. Conf. Ser.* **331** (2011) 032023.
- [108] LHCb, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, *JINST* **10** (2015), no. 02 P02007, [arXiv:1408.1251](#).
- [109] R. Aaij *et al.*, *Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2*, *EPJ Tech. Instrum.* **6** (2019), no. 1 1, [arXiv:1803.00824](#).
- [110] J. H. Friedman, *Stochastic gradient boosting*, *Comput. Stat. Data Anal.* **38** (2002) 367.
- [111] K. P. Murphy, *Machine learning: a probabilistic perspective*, MIT press, 2012.
- [112] [XGBoost Tutorial: Introduction to Boosted Trees](#).
- [113] [XGBoost Python: Using the Scikit-Learn Estimator Interface](#).
- [114] J. Cerasoli, *Search for the rare $B^0 \rightarrow K^{*0} \tau^+ \tau^-$ decay at the LHCb experiment*, PhD thesis, Aix Marseille Univ, CNRS/IN2P3, CPPM, Marseille, France, 2022, [URL](#).
- [115] T. Fawcett, *An introduction to roc analysis*, *Pattern recognition letters* **27** (2006), no. 8 861.
- [116] W. H. Press, *Numerical recipes 3rd edition: The art of scientific computing*, Cambridge university press, 2007.
- [117] [zfit documentation v0.18.1: DoubleCB](#).
- [118] M. Pivk and F. R. Le Diberder, *SPlot: A Statistical tool to unfold data distributions*, *Nucl. Instrum. Meth. A* **555** (2005) 356, [arXiv:physics/0402083](#).
- [119] L. Demortier and L. Lyons, *Everything you always wanted to know about pulls*, tech. rep., Technical Report CDF/ANAL/PUBLIC/5776, CDF, 2002.
- [120] G. Cowan, *Statistical data analysis*, 1998.
- [121] M. De Cian, *Track Reconstruction Efficiency and Analysis of $B^0 \rightarrow K^{*0} \mu^+ \mu^-$ at the LHCb Experiment*, PhD thesis, Zurich U., 2013, doi: [10.5167/uzh-85745](#).
- [122] T. M. Karbach and M. Schlupp, *Constraints on Yield Parameters in Extended Maximum Likelihood Fits*, [arXiv:1210.7141](#).
- [123] M. Yildiz, *An analysis of systematic uncertainties in an angular research in $B^0 \rightarrow K^{*0} e^+ e^-$* , Master Thesis, Universiteit Van Amsterdam, Netherlands, 2024 (in preparation).
- [124] LHCb, *HLT1 trigger efficiencies in 2024 data*, [LHCB-FIGURE-2024-030](#).

Summary

“There is something strange and unsettling about the relationship between humans and the cosmos that is never mentioned: the universe is actually dark, there is no light, the stars do not shine, the Sun is not luminous, the Moon does not reflect its rays. Everything is black, frighteningly black. Because light exists only if there are eyes and a brain capable of transforming electromagnetic waves into luminous signals, as the human brain does.” - Piero Angela, Superquark

These words speak not only of the visible universe, but also of knowledge itself. What we do not know is dark to our eyes -unseen, incomprehensible. Yet, a new way of observing is enough to illuminate what was once hidden. For me, fundamental particle detectors like LHCb represent precisely that: a new eye through which to explore the universe, lighting it up with new insights and revealing otherwise inaccessible secrets.

There is something extraordinary in these instruments: they not only show us the present, they also take us back in time. With the extreme energies achieved in particle accelerators, we can recreate conditions similar to those during the first moments after the Big Bang. In that instant, for just a brief moment, particles that have not existed for billions of years come to life. They are unstable, ephemeral: they decay almost immediately, but not before leaving a trace. And it is precisely from those traces that we build our understanding of the universe's past.

In the last seventy years, we have discovered an incredible number of particles - enough to fill an encyclopedia. Among all of them, only a few are fundamental: the ultimate building blocks of matter, those that - as far as we know today - cannot be divided any further.

These particles are shown in **Fig. I**. Some of them are responsible for transmitting the fundamental forces of nature and are called bosons. Others are the true constituents of matter: the fermions. With these ingredients, nature builds everything we know, from stars to our own bodies, including every atom in the universe.

All interactions among these particles are described by a theory that, to this day, is our best guide for navigating the infinitely small: the Standard Model. An elegant, precise theory, confirmed by decades of experiments. And yet, incomplete.

The Standard Model works, but still leaves many questions unanswered. For example, it does not include gravity - the force that keeps the planets in orbit and holds us to the ground. It also does not explain why the universe today is made almost entirely of matter and not of antimatter, even though our current theories suggest they should have been created in equal amounts.

And then there is the issue of galaxies: when we observe them, we see they rotate more quickly than can be explained by gravity and the visible matter they contain. To explain this discrepancy, scientists have hypothesised the existence of a form of invisible matter, aptly named dark matter.

The way to continue exploring this invisible and unknown universe is still the same as always:

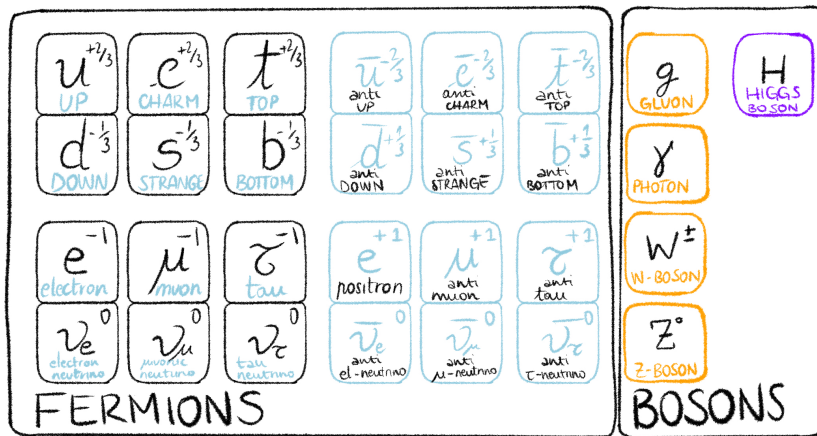


Figure I: Representation of the fundamental particles of the Standard Model. The number at the top right corner of each tile indicates the electrical charge of the particle.

the scientific method. We follow Popper's philosophy of falsifiability, constantly testing our theories. We collect data, analyse it, and let it speak for itself. If observations are consistent with the theoretical model, then the model survives a bit longer. But if something does not add up, if there is an anomaly the theory cannot explain, then a crack opens through which new light - or new physical laws *beyond the Standard Model* - might emerge.

Dear reader, let me tell you right away: this thesis does not reveal anything extraordinary or mysterious. But, among technicalities and details, it lifts a small part of the veil of Maya, offering a quantitative measure of how precisely we can grasp something new.

The protagonist of this story is a particle called "Beauty". It is composed of two quarks: one down and one bottom (also known as beauty, and referred to as b quark in this thesis). These two quarks have opposite electric charge and together form a neutral particle called the B^0 meson. On average after 1.5 picoseconds, this particle decays spontaneously. A picosecond is an incredibly short time: to make it comparable to a second, we'd have to imagine that one second lasts about 30,000 years! Yet, despite its very short lifetime, the Beauty particle manages to leave a trace. This is because it is created at such high energies that it moves almost at the speed of light. As a result, time "slows down" for it, allowing it to travel a few centimetres before decaying.

Just 8 millimetres from the point where the protons collide inside the Large Hadron Collider lies the first element of the LHCb (Large Hadron Collider beauty) detector, designed specifically to be sensitive to these particles. LHCb is the eye and the brain that transforms information, from the proton collision into data ready to be analysed.

Once in a while - more precisely, once every 6 million Beauty particles produced - it decays into a K^{*0} and two oppositely charged electrons. This specific decay, $B^0 \rightarrow K^{*0} e^+ e^-$, is the subject of this thesis. Its peculiarity lies in the fact that the b quark, a constituent of the B^0 meson with electric charge $-1/3$, transforms into an s quark with the same electric charge, emitting two oppositely charged electrons. These transitions are known as "flavour-changing neutral currents", because the quark changes its "flavour" while the net electric charge exchanged remains zero. According to the Standard Model, they occur as illustrated in the Feynman diagram in **Fig. II**.

The loop at the centre of the diagram is a distinctive feature of these decays and explains why they are so rare.

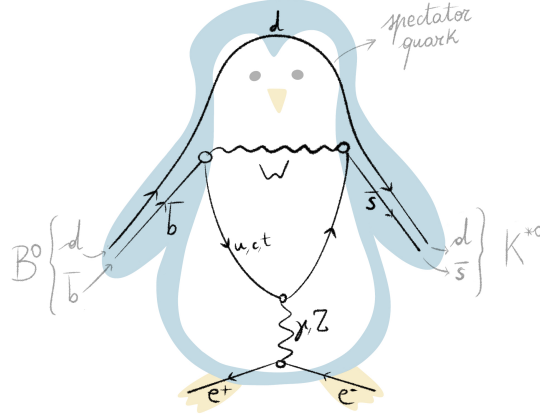


Figure II: Feynman diagram for the decay $B^0 \rightarrow K^{*0} e^+ e^-$, mediated by a flavour-changing neutral current. This type of diagram is also known as a “penguin” diagram and describes the transition of a quark x into a quark y with the same electric charge. The intermediate quarks differ in charge from x and y by one unit [1]. Typically, the spectator quark is not shown in the penguin diagram, but it is included here for clarity.

By analysing this decay as a function of the squared invariant mass of the di-electron system, called q^2 , one gains sensitivity to different Wilson Coefficients - parameters that encode underlying physics phenomena at higher energy scales. Simply put, these coefficients may hint at the existence of new physical structures that lie at energy scales not yet experimentally accessible. These new physical structures would imply that new fundamental particles, not listed in **Fig. II**, would exist, requiring our current theoretical knowledge to be expanded *beyond the Standard Model*. Measurements of this kind are known as “indirect searches”: instead of looking for direct signs of new particles, well-known phenomena are studied and compared with theoretical predictions, in search of discrepancies that might suggest the presence of physics beyond the Standard Model. Penguin-type decays are especially well suited for such indirect searches precisely because of their rarity: even the smallest contribution from something new could produce measurable effects.

The q^2 spectrum features several regions of particular interest. This thesis focuses on the interval $[1.10, 6.00] \text{ GeV}^2/c^4$ and two smaller sub-intervals within it: $[1.10, 3.65] \text{ GeV}^2/c^4$ and $[3.65, 6.00] \text{ GeV}^2/c^4$. The main goal of the analysis is to measure the angular parameters that describe the decay. These parameters, too, have Standard Model predictions that can be tested using the LHCb data. The decay $B^0 \rightarrow K^{*0} e^+ e^-$ can be described by three angles $(\theta_K, \theta_\ell, \phi)$ ¹: the K^{*0} meson decays into a $K\pi$ pair, which defines θ_K , the electron pair defines θ_ℓ , and the angle between the planes spanned by these two pairs defines ϕ . The parameters we aim to measure - the aforementioned angular parameters - are extracted by comparing the data distribution as a function of the three angles $(\theta_K, \theta_\ell, \text{ and } \phi)$ to the functional form predicted by the Standard Model, using a minimisation (fitting) procedure.

An example of the projections of this distribution is shown in **Fig. III**: the black points represent the experimental data, while the blue line is the function that best describes the observations.

¹The mathematical definition of the angles is provided in **Sec. 1.3.1**.

This is the best fit to the data, from which the angular parameters are extracted, within an uncertainty interval determined by the amount of data available.

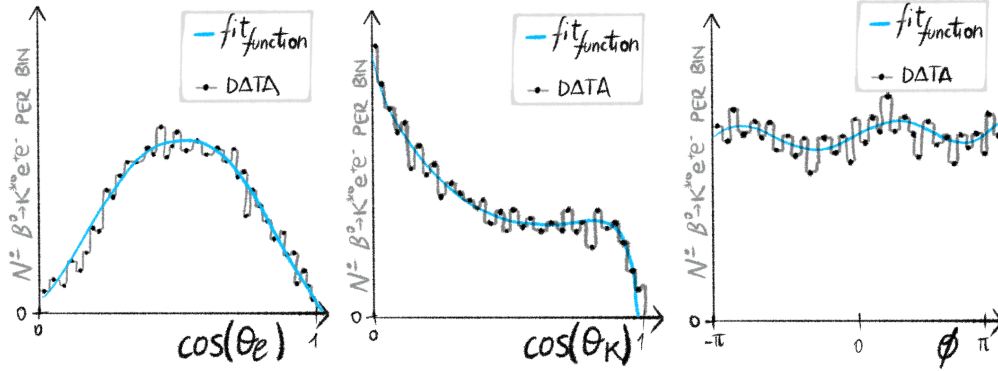


Figure III: Illustration not to scale that shows the $\cos \theta_e$, $\cos \theta_K$, and ϕ histograms for the $B^0 \rightarrow K^{*0} e^+ e^-$ decays.

At the time of writing this thesis, the analysis is still ongoing and the angular distributions - as well as the corresponding angular parameters - remain blinded. This means the analysts do not have access to the actual numerical values of the parameters, to avoid human biases - our natural tendency to confirm initial hypotheses. The main objective of this work is therefore to evaluate the precision and accuracy with which these parameters can be measured in the two selected q^2 regions, never before explored separately in this decay channel.

The first part of the thesis describes in detail the data selection process. In fact, the data collected by the experiment are not immediately ready for fitting. In rare decays like $B^0 \rightarrow K^{*0} e^+ e^-$, the signals are often submerged by numerous events that mimic the decay but are not truly the same, sharing only some features. These events are known as background. Therefore, the first step of the analysis is to filter out the background and isolate true $B^0 \rightarrow K^{*0} e^+ e^-$ candidates, namely events showing all expected features of the decay. However, even with very strict selection criteria, some background events may remain indistinguishable from real decays: these are known as irreducible backgrounds and must be modelled together with the signal in the final fit to the data.

At this point, the second part of the analysis can begin where the filtered data - containing both signal and background - are fitted with theoretical functions capable of describing them separately to extract the angular parameters describing the signal decay, as mentioned earlier.

The final step of the analysis involves assessing the impact of the approximations and assumptions made throughout the measurement. For instance, some background contributions are considered negligible - but is that assumption truly justified? Similarly, the functional forms used to model signal and background events are used because they empirically describe the dataset well. However, they are not the only plausible choices. How would the results change if alternative, yet reasonable, functions were used instead? Asking such questions and seeking quantitative answers is a fundamental part of the analysis. It allows us to quantify how much the residual ignorance impacts the final results, thereby assessing the robustness of the measurement.

Given the current state of knowledge in particle physics, progress stems from collective effort: it is thanks to the small contributions of many that we can explore fascinating realities otherwise beyond our reach. This work represents one small brick in the ongoing search for new

fundamental laws. Precisely measuring the angular parameters of the $B^0 \rightarrow K^{*0} e^+ e^-$ decay helps test the limits of the Standard Model, in the hope of opening new windows toward the unknown.

Samenvatting

“Er is iets vreemds en verontrustends aan de relatie tussen de mens en het heelal waar nooit over gesproken wordt: het universum is eigenlijk donker, er is geen licht, de sterren schijnen niet, de zon is niet lichtgevend, de maan weerkaatst haar stralen niet. Alles is zwart, beangstigend zwart. Want licht bestaat alleen als er ogen en een brein zijn die elektromagnetische golven kunnen omzetten in lichtsignalen, zoals het menselijk brein dat doet.” - Piero Angela, Superquark, presentator van een Italiaans populairwetenschappelijk televisieprogramma

Deze woorden gaan niet alleen over het zichtbare universum, maar ook over kennis zelf. Wat we niet weten is voor onze ogen donker – onzichtbaar, onbegrijpelijk. En toch is een nieuwe manier van observeren genoeg om dat wat verborgen was te verlichten. Voor mij zijn fundamentele-deeltjesdetectoren zoals LHCb precies dat: een nieuw oog om het universum te verkennen, het te verlichten met nieuwe inzichten en geheimen te onthullen die anders ontoegankelijk zouden blijven.

Er is iets buitengewoons aan het feit dat deze instrumenten ons niet alleen het heden tonen. Ze brengen ons terug in de tijd. Dankzij de extreme energieën die bereikt worden in deeltjesversnellers, kunnen we omstandigheden nabootsen die lijken op die van de eerste momenten na de oerknal. Op dat moment, heel even, komen deeltjes tot leven die al miljarden jaren niet meer bestaan. Ze zijn onstabiel, vluchtig: ze vervallen bijna onmiddellijk, maar laten een spoor achter. En precies uit die sporen bouwen we ons begrip van het verleden van het universum op.

In de afgelopen zeventig jaar hebben we een ongelooflijk aantal deeltjes ontdekt – genoeg om een encyclopedie te vullen. Maar van al die deeltjes zijn er slechts enkele fundamenteel: de uiteindelijke bouwstenen van materie, die – voor zover we nu weten – niet verder deelbaar zijn.

Deze deeltjes worden weergegeven in **Fig. I**. Sommige zijn verantwoordelijk voor het overbrengen van de fundamentele natuurkrachten en worden bosonen genoemd. Andere zijn de echte bestanddelen van materie: de fermionen. Met deze ingrediënten bouwt de natuur alles wat we kennen: van sterren tot onze eigen lichamen, inclusief elk atoom in het universum.

Alle interacties tussen deze deeltjes worden beschreven door een theorie die tot op heden onze beste gids is voor het navigeren in het oneindig kleine: het Standaardmodel. Een elegante, precieze theorie, bevestigd door tientallen jaren van experimenten. En toch, onvolledig.

Het Standaardmodel werkt, maar laat nog veel vragen onbeantwoord. Zo bevat het geen zwaartekracht – de kracht die planeten in een baan houdt en ons op de grond. Ook verklaart het niet waarom het universum vandaag bijna volledig uit materie bestaat en niet uit antimaterie, terwijl onze basis-theorieën suggereren dat ze in gelijke hoeveelheden zouden moeten zijn ontstaan.

En dan is er het probleem van de sterrenstelsels: wanneer we ze observeren, zien we dat ze te

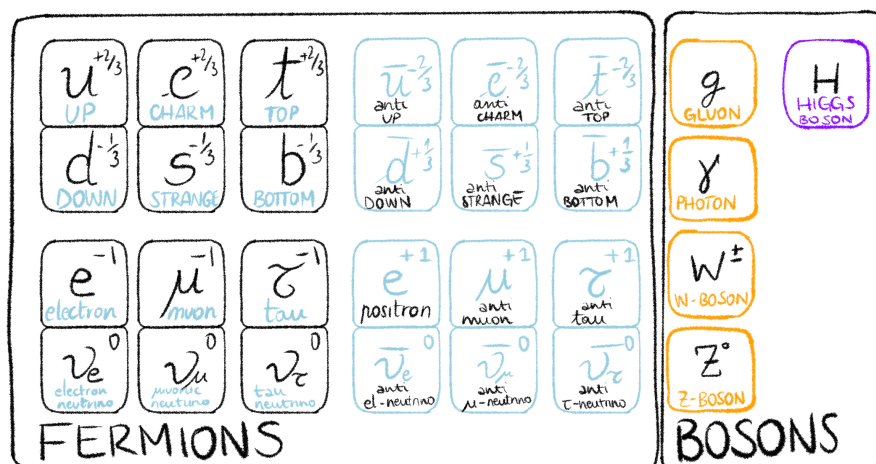


Figure I: Weergave van de fundamentele deeltjes van het Standaardmodel.

snel roteren in verhouding tot de zichtbare materie die ze bevatten. Om dit verschil te verklaren, hebben wetenschappers het bestaan van een nieuwe vorm van onzichtbare materie voorgesteld, toepasselijk donkere materie genoemd.

De manier om dit onzichtbare en onbekende universum verder te verkennen blijft dezelfde als altijd: de wetenschappelijke methode. We volgen de filosofie van falsifieerbaarheid van Popper, en testen onze theorieën voortdurend. We verzamelen data, analyseren die, en laten de gegevens voor zichzelf spreken. Als de observaties overeenkomen met het theoretische model, dan overleeft het model nog wat langer. Maar als er iets niet klopt, als er een anomalie is die we niet kunnen verklaren, dan opent zich een barst waarlangs nieuw licht – of nieuwe natuurwetten voorbij het Standaardmodel – zou kunnen verschijnen.

Beste lezer, laat me u meteen waarschuwen: deze dissertatie onthult niets buitengewoons of mysterieus. Maar tussen technische details en complexiteit, licht het een klein stukje van de sluier van Maya op, door een kwantitatieve maat te bieden van hoe precies we iets nieuws te weten kunnen komen.

De hoofdrolspeler van dit verhaal is een deeltje genaamd “Beauty”. Het is samengesteld uit twee quarks: één down en één bottom (ook wel beauty genoemd). Deze twee quarks hebben tegengestelde elektrische ladingen en vormen samen een neutraal deeltje, het B^0 -meson. Na ongeveer 1,5 picoseconde vervalt dit deeltje spontaan. Een picoseconde is een ongelooflijk korte tijd: om het vergelijkbaar te maken met één seconde, moeten we ons voorstellen dat één seconde ongeveer 30.000 jaar duurt!

Toch slaagt het Beauty-deeltje erin een spoor achter te laten, ondanks zijn korte levensduur. Dat komt omdat het wordt gecreëerd met zulke hoge energieën dat het zich bijna met de lichtsnelheid voortbeweegt. Daardoor “vertraagt” de tijd voor het deeltje, waardoor het enkele centimeters kan afleggen voordat het vervalt.

Op slechts 8 millimeter van het punt waar de protonen botsen in de Large Hadron Collider ligt het eerste onderdeel van de LHCb-detector (Large Hadron Collider beauty), speciaal ontworpen om gevoelig te zijn voor deze deeltjes. LHCb is het oog en het brein dat informatie omzet van de protonbotsing naar gegevens die geanalyseerd kunnen worden.

Heel af en toe – meer bepaald, eenmaal per 6 miljoen geproduceerde Beauty-deeltjes –

vervalt het in een K^{*0} -meson en twee tegengesteld geladen elektronen. Dit specifieke verval, $B^0 \rightarrow K^{*0} e^+ e^-$, is het onderwerp van deze dissertatie. De bijzonderheid ervan ligt in het feit dat de b -quark, een component van het B^0 -meson met elektrische lading $-1/3$, verandert in een s -quark met dezelfde elektrische lading, waarbij twee tegengesteld geladen elektronen worden uitgezonden. Deze overgangen worden “flavour-changing neutral currents” genoemd, omdat de quark zijn “flavour” verandert terwijl de netto uitgewisselde elektrische lading nul blijft. Volgens het Standaardmodel komen ze voor zoals geïllustreerd in het Feynman-diagram in **Fig. II**.

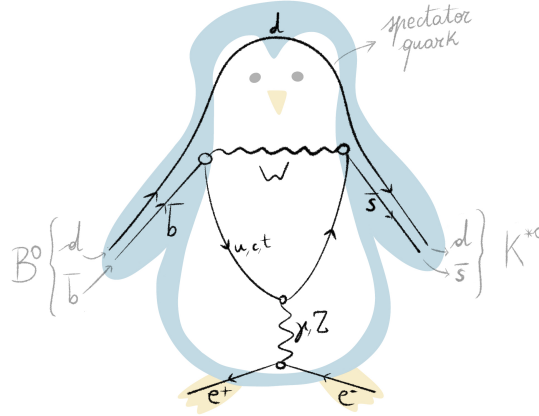


Figure II: Feynman-diagram voor het verval $B^0 \rightarrow K^{*0} e^+ e^-$, gemedieerd door een flavour-changing neutral current. Dit type diagram staat ook bekend als een “penguin”-diagram en beschrijft de overgang van een quark x naar een quark y met dezelfde elektrische lading. De intermediaire quarks verschillen in lading van x en y met één eenheid [1]. Gewoonlijk wordt de spectator-quark niet weergegeven in het penguin-diagram, maar is hier ter verduidelijking wel opgenomen.

De lus in het midden van het diagram is een kenmerkend element van deze vervallen en verklaart waarom ze zo zeldzaam zijn.

Door dit verval te analyseren als functie van de kwadratische invariante massa van het elektronensysteem, q^2 genoemd, krijgt men gevoeligheid voor verschillende Wilsoncoëfficiënten – parameters die de effecten van hogere energieschalen coderen. Simpel gezegd: deze coëfficiënten kunnen wijzen op het bestaan van nieuwe fysische structuren die zich op energieschalen bevinden die nog niet experimenteel toegankelijk zijn. Metingen van dit type worden “indirecte metingen” genoemd: in plaats van direct naar nieuwe deeltjes te zoeken, bestuderen we bekende fenomenen en vergelijken we ze met theoretische voorspellingen, op zoek naar afwijkingen die kunnen wijzen op fysica voorbij het SM. Penguin-vervallen zijn bijzonder geschikt voor zulke indirecte metingen, juist vanwege hun zeldzaamheid: zelfs de kleinste bijdrage van iets nieuws kan meetbare effecten veroorzaken.

Het q^2 -spectrum bevat verschillende gebieden van bijzonder belang. Deze dissertatie richt zich op drie intervallen: het centrale interval $[1.10, 6.00] \text{ GeV}^2/c^4$, en twee kleinere subintervallen daarin: $[1.10, 3.65] \text{ GeV}^2/c^4$ en $[3.65, 6.00] \text{ GeV}^2/c^4$. Het hoofddoel van de analyse is het meten van de hoekcoëfficiënten die het verval beschrijven. Ook voor deze coëfficiënten bestaan voorspellingen binnen het Standaardmodel die getest kunnen worden met LHCb-data. Het verval $B^0 \rightarrow K^{*0} e^+ e^-$ kan beschreven worden door drie hoeken ($\theta_K, \theta_\ell, \phi$): het K^{*0} -meson vervalt in

een $K\pi$ -paar, wat θ_K definieert, het elektronenpaar definieert θ_ℓ , en de hoek tussen de vlakken gespannen door deze twee paren definieert ϕ^2 . De parameters die we willen meten – de genoemde hoekcoëfficiënten – worden verkregen door het modelleren van de verdelingen van de drie hoeken: θ_K , θ_ℓ en ϕ .

Een voorbeeld van deze verdelingen is weergegeven in **Fig. III**: de zwarte punten stellen de experimentele data voor, terwijl de blauwe lijn de functie is die de observaties het best beschrijft – dat wil zeggen, de kromme die de afstand tot elk punt minimaliseert. Dit is de best passende functie bij de data, waaruit de hoekcoëfficiënten worden gehaald, binnen een onzekerheidsinterval dat bepaald wordt door de hoeveelheid beschikbare data.

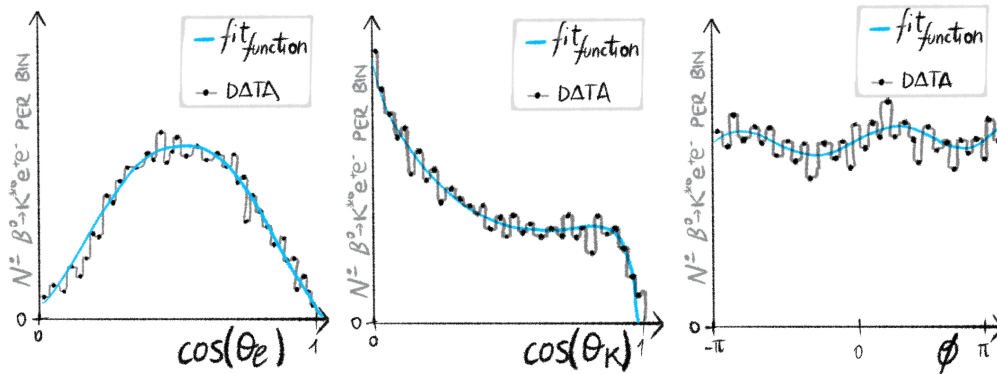


Figure III: Niet-op-schaal illustratie van de $\cos \theta_\ell$, $\cos \theta_K$ en ϕ histogrammen voor het $B^0 \rightarrow K^{*0} e^+ e^-$ -verval.

Op het moment van schrijven van deze dissertatie is de analyse nog aan de gang en zijn de hoekverdelingen – evenals de overeenkomstige hoekcoëfficiënten – nog geblindeerd. Dat betekent dat de onderzoekers geen toegang hebben tot de echte waarden van de coëfficiënten, om experimentele bias – onze natuurlijke neiging om eerste hypothesen te bevestigen – te vermijden. Het voornaamste doel van dit werk is dus de precisie en nauwkeurigheid te evalueren waarmee deze coëfficiënten gemeten kunnen worden in de twee geselecteerde q^2 -gebieden, die nooit eerder afzonderlijk in dit vervalkanaal zijn onderzocht.

De eerste helft van de dissertatie beschrijft gedetailleerd het dataselectieproces. De gegevens die het experiment verzamelt, zijn immers niet onmiddellijk klaar om gefit te worden. In zeldzame gevallen zoals $B^0 \rightarrow K^{*0} e^+ e^-$ is het signaal vaak omgeven door talloze event die het verval nabootsen, maar niet hetzelfde zijn, en enkel bepaalde kenmerken delen. Deze event worden achtergrond genoemd. De eerste stap van de analyse is daarom het filteren van de achtergrond en het isoleren van echte $B^0 \rightarrow K^{*0} e^+ e^-$ -kandidaten, dus event die alle verwachte kenmerken van het verval vertonen. Maar zelfs met zeer strenge selectiecriteria kunnen sommige achtergrond-event niet te onderscheiden zijn van echte vervallen: dit zijn de irreduceerbare achtergronden, die samen met het signaal moeten worden gemodelleerd in de uiteindelijke datafit.

Op dat punt kan het tweede deel van de analyse beginnen, waarbij de gefilterde data – met zowel signaal als achtergrond – worden gefit met theoretische functies die ze afzonderlijk kunnen beschrijven, om zo de eerder genoemde hoekcoëfficiënten te extraheren.

De laatste stap van de analyse bestaat erin de impact te evalueren van de aannames en benaderingen die tijdens de meting gemaakt zijn. Zo worden bepaalde achtergrondbijdragen als

²De wiskundige definitie van de hoeken wordt gegeven in **Sec. 1.3.1**

verwaarloosbaar beschouwd – maar is dat wel echt terecht? Evenzo zijn de functies die gebruikt worden om signaal- en achtergrond-event te modelleren gekozen omdat ze het dataprofiel empirisch goed beschrijven. Maar het zijn niet de enige plausibele keuzes. Hoe zouden de resultaten veranderen als we andere, eveneens redelijke functies gebruikten? Deze vragen stellen, en zoeken naar kwantitatieve antwoorden, is een fundamenteel onderdeel van de analyse. Het stelt ons in staat te kwantificeren in welke mate de resterende onwetendheid invloed heeft op de eindresultaten, en daarmee om de robuustheid van de meting te beoordelen.

Gezien de huidige stand van kennis in de deeltjesfysica komt vooruitgang voort uit collectieve inspanning: het is dankzij de kleine bijdragen van velen dat we fascinerende werkelijkheden kunnen verkennen die anders buiten ons bereik zouden liggen. Dit werk vertegenwoordigt één kleine bouwsteen in de voortdurende zoektocht naar nieuwe fundamentele wetten. Het nauwkeurig meten van de hoekcoëfficiënten van het $B^0 \rightarrow K^{*0} e^+ e^-$ -verval helpt om de grenzen van het Standaardmodel te testen, in de hoop nieuwe vensters te openen richting onbekende fysica.

Riassunto

"C'è una cosa strana e sconcertante che riguarda il rapporto dell'uomo con il cosmo e che non viene mai detta: è il fatto che l'universo in realtà è buio, non c'è alcuna luce, le stelle non brillano, il Sole non è luminoso, la Luna non riflette i suoi raggi. Tutto è nero, spaventosamente nero. Perché la luce esiste solo se ci sono degli occhi ed un cervello capace di trasformare delle onde elettromagnetiche in segnali luminosi, come fa appunto il cervello umano. " - Piero Angela, Superquark

Queste parole non parlano solo dell'universo visibile, ma anche della conoscenza. Ciò che non sappiamo è buio ai nostri occhi, invisibile, incomprensibile. Basta però un nuovo modo di osservare per illuminare ciò che prima era nascosto. Per me, i rivelatori di particelle fondamentali come LHCb rappresentano proprio questo: un nuovo occhio con cui scrutare l'universo, accendendolo di nuove luci e rivelandone segreti altrimenti inaccessibili.

C'è qualcosa di straordinario nel fatto che questi strumenti non ci mostrino solo il presente, ma ci riportano indietro nel tempo. Con le energie estreme che raggiungiamo negli acceleratori di particelle, riusciamo a ricreare condizioni simili a quelle dei primi istanti dopo il Big Bang. In quel momento, solo per qualche istante, prendono vita particelle che non esistono più da miliardi di anni. Sono instabili, effimere: si disintegrano quasi subito, ma non prima di lasciare una traccia. Ed è proprio da quelle tracce che costruiamo la nostra comprensione del passato dell'universo.

Negli ultimi settant'anni abbiamo scoperto un'incredibile quantità di particelle, abbastanza da riempirci un'enciclopedia. Ma tra tutte, solo alcune sono fondamentali: i mattoncini ultimi della materia, quelli che – per quanto ne sappiamo oggi – non si possono dividere ulteriormente.

Queste particelle sono rappresentate in **Fig. I**. Alcune di loro sono responsabili della trasmissione delle forze fondamentali della natura e vengono chiamate bosoni. Altre, invece, sono i costituenti della materia vera e propria: i fermioni. Con questi ingredienti, la natura costruisce tutto ciò che conosciamo, dalle stelle ai nostri corpi, passando per ogni atomo dell'universo.

Tutte le interazioni tra queste particelle sono descritte da una teoria che, ad oggi, è la nostra migliore bussola per orientarci nell'infinitamente piccolo: il Modello Standard. Una teoria elegante, precisa, confermata da decenni di esperimenti. Eppure, incompleta.

Il Modello Standard funziona, ma lascia ancora molte domande senza risposta. Ad esempio, non include la forza di gravità – quella che tiene i pianeti in orbita, che ci tiene ancorati a terra. E non spiega perché l'universo, oggi, sia fatto quasi esclusivamente di materia e non di antimateria, anche se dalle nostre teorie iniziali dovrebbero essersi create in quantità uguali.

E poi c'è la questione delle galassie: osservandole, vediamo che ruotano più velocemente rispetto a quanto predetto considerando la materia visibile che contengono. Per spiegare questa discrepanza, gli scienziati hanno ipotizzato l'esistenza di una nuova forma di materia, invisibile,

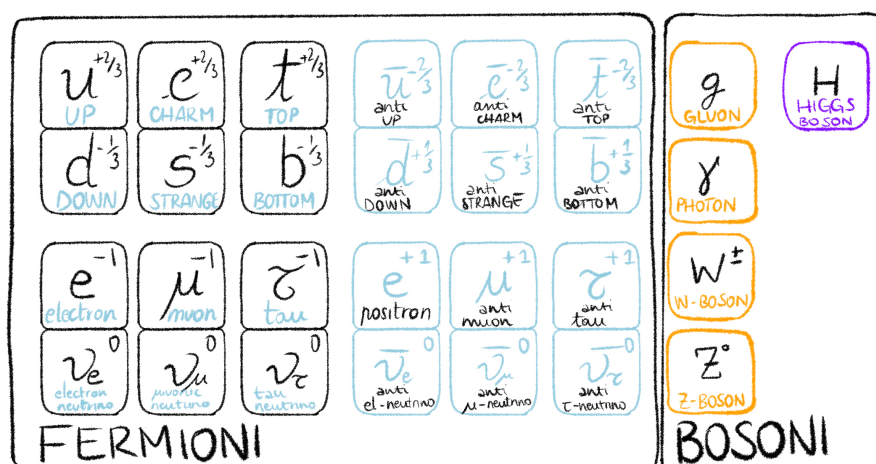


Figure I: Rappresentazione delle particelle fondamentali del Modello Standard. I numeri in alto a destra di ciascuna casella indicano la carica elettrica di ogni fermione.

che per questo viene chiamata materia oscura.

Il modo per continuare a esplorare questo universo invisibile e sconosciuto è ancora quello di sempre: il metodo scientifico. Seguiamo la filosofia della falsificabilità di Popper, mettendo continuamente alla prova le nostre teorie. Raccogliamo dati, li analizziamo, e lasciamo che siano loro a parlare. Se le osservazioni sono coerenti con il modello teorico, allora il modello sopravvive un po' più a lungo. Se invece qualcosa non torna, se c'è un'anomalia che non possiamo spiegare, allora si apre una crepa dalla quale potrebbe entrare una nuova luce o leggi fisiche *oltre* il Modello Standard.

Caro lettore, ti avverto subito: questa tesi non svelerà qualcosa di straordinario o oscuro. Ma, tra tecnicismi e dettagli, solleva una piccola parte del velo di Maya, offrendo una misura quantitativa di quanto precisamente ed accuratamente possiamo conoscere qualcosa di nuovo.

La protagonista di questa storia si chiama particella "Beauty". È composta da due quark: uno down, d , e uno bottom, b (anche chiamato beauty). Questi due quark hanno carica elettrica opposta e, insieme, formano una particella neutra, chiamata mesone B^{03} . Dopo circa 1,5 picosecondi, questa particella decade spontaneamente. Un picosecondo è un tempo incredibilmente breve: se volessimo renderlo paragonabile a un secondo, dovremmo immaginare che un secondo duri circa 30 000 anni!

Eppure, nonostante la sua brevissima vita, la particella Beauty riesce a lasciare una traccia. Questo perché viene creata a energie così elevate da muoversi quasi alla velocità della luce. A causa di questo, il tempo per lei "rallenta", e riesce a percorrere qualche centimetro prima di decadere.

A soli 8 millimetri dal punto in cui avviene la collisione tra i protoni che circolano dentro il Large Hadron Collider, si trova il primo elemento del rivelatore LHCb (Large Hadron Collider beauty), progettato appositamente per essere sensibile a queste particelle. LHCb è l'occhio ed il cervello che converte le informazioni, dalla collisione tra protoni fino ai dati pronti per essere essere analizzati.

³Per completezza, questa particella esiste come materia, B^0 , e come la sua controparte di anti-materia, $\bar{B}^0$. Esse consistono di un anti-bottom quark, $\bar{b}$, e di un down quark, d , oppure di un bottom quark, b , ed un anti-down quark, $\bar{d}$. Questa tesi non fa distinzione tra i due e li considera entrambi.

Una volta ogni tanto - più precisamente, una volta ogni 6 milioni di particelle Beauty prodotte - essa decade in un mesone K^{*0} e due elettroni di carica opposta. Questo particolare decadimento, $B^0 \rightarrow K^{*0} e^+ e^-$, è l'oggetto di studio di questa tesi. La peculiarità di questo decadimento sta nel fatto che il quark b , costituente del mesone B^0 e con carica elettrica $-1/3$, si trasforma in un quark s con la stessa carica elettrica, emettendo due elettroni di carica opposta. Queste transizioni sono dette “flavour-changing neutral current”, ovvero correnti neutre con cambiamento di sapore, perché il quark cambia “sapore”, cioè tipo di quark, ma la carica elettrica totale scambiata rimane nulla. Secondo il Modello Standard, esse avvengono come illustrato nel diagramma di Feynman in **Fig. II**.

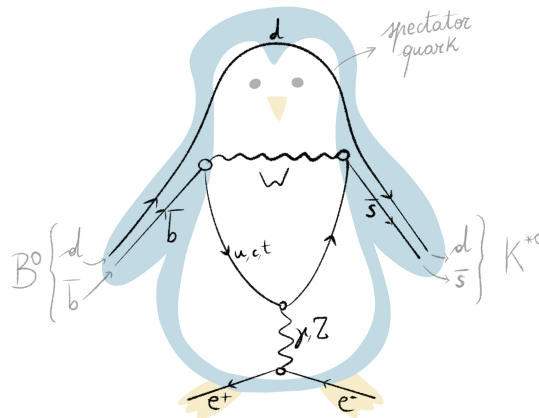


Figure II: Diagramma di Feynman per il decadimento $B^0 \rightarrow K^{*0} e^+ e^-$, mediato da una corrente neutra con cambiamento di sapore (flavour changing neutral current). Questo tipo di diagramma è anche noto come “diagramma a pinguino” (Penguin) e descrive la transizione di un quark x in un quark y con la stessa carica elettrica. I quark intermedi hanno carica che differisce da quella di x e y di un’unità [1]. Normalmente il diagramma a pinguino non mostra il quark spettatore, che qui è riportato per maggiore chiarezza.

I decadimenti che seguono diagrammi di questo tipo sono conosciuti come decadimenti a “pinguino”; il loop al centro del diagramma è una loro caratteristica distintiva, e spiega perché siano così rari.

Analizzando questo decadimento in funzione della massa invariante al quadrato del sistema di elettroni, q^2 , si ottiene sensibilità a diversi Coefficienti di Wilson - parametri che codificano gli effetti di scale di energia più elevate. In parole semplici, questi coefficienti possono suggerire l’esistenza di strutture fisiche ancora sconosciute che esisterebbero a energie non ancora accessibili sperimentalmente. Misure di questo tipo prendono il nome di “ricerche indirette”: invece di cercare segnali diretti di nuove particelle, si studiano fenomeni ben noti confrontandoli con le previsioni teoriche, alla ricerca di eventuali discrepanze che possano suggerire l’esistenza di nuovi fenomeni fisici. I decadimenti di particelle di tipo “pinguino” sono particolarmente adatte a questo tipo di ricerca indiretta proprio per la loro rarità: anche un piccolissimo contributo da parte di qualcosa di nuovo potrebbe produrre effetti misurabili.

Lo spettro in q^2 presenta diverse regioni di particolare interesse. Questa tesi si concentra sull’intervallo $[1.10, 6.00] \text{ GeV}^2/c^4$, e due suoi sottointervalli più piccoli: $[1.10, 3.65] \text{ GeV}^2/c^4$ e $[3.65, 6.00] \text{ GeV}^2/c^4$. L’obiettivo principale dell’analisi è la misura dei parametri angolari che

descrivono il decadimento. Infatti, anche per tali parametri esiste una predizione da parte del Modello Standard che può essere testata analizzando i dati raccolti da LHCb. Il decadimento $B^0 \rightarrow K^{*0} e^+ e^-$ è infatti descrivibile da tre angoli ($\theta_K, \theta_\ell, \phi$): la particella K^{*0} decade in una coppia $K\pi$ che definisce θ_K , la coppia di elettroni definisce θ_ℓ , e l'angolo tra i piani individuati da queste due coppie definisce ϕ^4 . I parametri che vogliamo misurare, quindi i parametri angolari menzionati prima, si estraggono tramite un fit tridimensionale alla distribuzione dei tre angoli: $\theta_K, \theta_\ell, \phi$.

Un esempio di queste distribuzioni è mostrato in **Fig. III**: i punti neri rappresentano i dati sperimentali raccolti, mentre la linea blu è la funzione che meglio descrive le osservazioni, ossia la curva che minimizza la distanza da ciascun punto. Si tratta del miglior fit ai dati, da cui si ricavano i parametri angolari, entro un intervallo di incertezza determinato dalla quantità di dati disponibili.

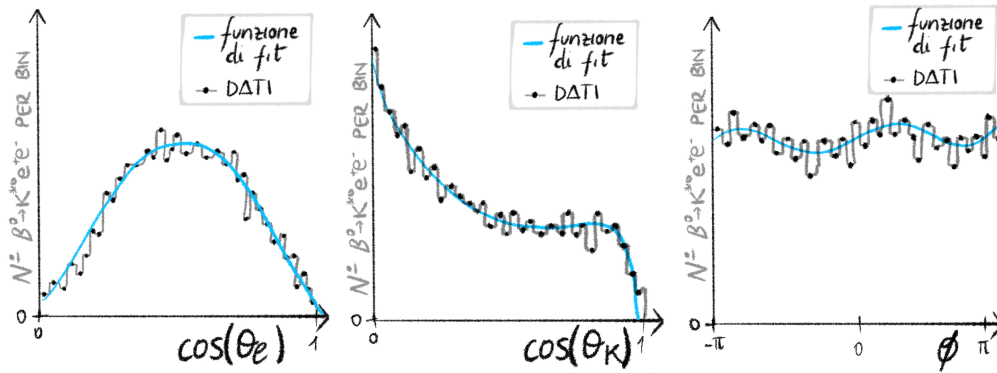


Figure III: Illustrazione non in scala delle distribuzioni angolari $\cos \theta_\ell$, $\cos \theta_K$, and ϕ del decadimento sotto analisi $B^0 \rightarrow K^{*0} e^+ e^-$.

Al momento della stesura di questa tesi, l'analisi è ancora in corso e le distribuzioni angolari - così come i parametri angolari corrispondenti - restano blinded, ovvero "accecati". Questo significa che chi effettua l'analisi non ha accesso ai valori reali dei parametri, per evitare bias sperimentali, cioè quella naturale tendenza umana a confermare le ipotesi iniziali. L'obiettivo principale di questo lavoro è quindi valutare la precisione e l'accuratezza con cui tali parametri possono essere misurati nelle due regioni di q^2 selezionate, mai esplorate separatamente in questo canale di decadimento.

La prima parte della tesi descrive in dettaglio il processo di selezione dei dati. Infatti, i dati raccolti dall'esperimento non sono immediatamente pronti per essere analizzati tramite fit. Nel caso di decadimenti rari come il $B^0 \rightarrow K^{*0} e^+ e^-$, i segnali sono spesso sommersi da numerosi eventi che imitano il decadimento ma non lo sono realmente, condividendo solo alcune caratteristiche. Questi eventi sono noti come fondo. Il primo passo dell'analisi è dunque filtrare il fondo per isolare i veri candidati $B^0 \rightarrow K^{*0} e^+ e^-$, cioè quegli eventi che mostrano tutte le caratteristiche attese dal decadimento. Tuttavia, anche imponendo criteri di selezione molto stringenti, possono rimanere eventi indistinguibili dai veri decadimenti: si tratta dei cosiddetti fondi irriducibili, che devono essere modellati insieme al segnale nel fit finale ai dati.

A questo punto si può procedere alla seconda parte dell'analisi, in cui i dati filtrati - contenenti sia segnale che fondo - vengono fittati con funzioni teoriche capaci di descriverli, al fine di estrarre

⁴La definizione matematica degli angoli è fornita in **Sec. 1.3.1**

i parametri angolari menzionati in precedenza.

L'ultima fase dell'analisi consiste nel valutare l'impatto delle approssimazioni e delle assunzioni effettuate durante tutta la misura. Ad esempio, alcuni contributi di fondo vengono considerati trascurabili - ma questa ipotesi è davvero giustificata? Analogamente, le forme funzionali utilizzate per modellare gli eventi di segnale e di fondo sono scelte perché descrivono bene il dataset da un punto di vista empirico. Tuttavia, non sono le uniche opzioni plausibili. Come cambierebbero i risultati se si utilizzassero funzioni alternative, ma comunque ragionevoli? Porsi queste domande e cercare risposte quantitative rappresenta un aspetto fondamentale dell'analisi. Ci permette di quantificare quanto l'ignoranza residua influisca sui risultati finali, e quindi di valutare la robustezza della misura.

Allo stato attuale della conoscenza in fisica delle particelle, il progresso nasce dal lavoro collettivo: è grazie ai piccoli contributi di molti che possiamo esplorare affascinanti realtà altrimenti inaccessibili. Questo lavoro rappresenta un piccolo mattone nella continua ricerca di nuove leggi fondamentali. Misurare con precisione i parametri angolari del decadimento $B^0 \rightarrow K^{*0} e^+ e^-$ contribuisce a testare i limiti del Modello Standard nella speranza di aprire spiragli verso una fisica ancora sconosciuta.

BEAUTY FROM EVERY ANGLE



Alice Biolchini

