Direct and Indirect Searches for New Physics in $b \to s\ell\ell$ Decays

---

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)
vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich

von
Andrea Mauri
aus
Italien

Promotionskommission

Prof. Dr. Nicola Serra
Dr. Rafael Silva Coutinho
PD Dr. Olaf Steinkamp
Dr. Katharina Müller

Zürich, 2019

ABSTRACT

The current era of particle physics is dominated by searches for physics beyond the Standard Model, which can be pursued with direct or indirect approaches. This dissertation investigates the features of rare semileptonic $b \to s\ell\ell$ transitions as probe of New Physics. Recent analyses of these decays have indicated an anomalous behaviour in measurements of angular distributions of the decay $B^0 \to K^{*0}\mu^+\mu^-$ and in lepton-flavour-universality observables. However, due to the limited understanding of hadronic uncertainties, the origin of some of these deviations is of difficult interpretation. These uncertainties, in fact, can mimic or camouflage New Physics effects limiting the discovery potential of measurements of $B^0 \to K^{*0}\mu^+\mu^-$ decays.

In order to solve this long-standing issue, a novel combined theory-experimental approach is presented. The proposed method is first applied to examine the prospects of an amplitude analysis of $B^0 \to K^{*0}\mu^+\mu^-$ decays and then extended to the channel $B^0 \to K^{*0}e^+e^-$ as a test of Lepton Flavour Universality. It is found that a separation between New Physics effects and hadronic uncertainties is possible and, if the current hints are confirmed, an early discovery of physics beyond the Standard Model can be achieved within the end of LHCb Run-II. Furthermore, the proposed method not only surpasses the sensitivity of previous approaches, but will also enable a deeper understanding of the nature of New Physics. Moreover, sensitivity studies, including all the experimental effects, have been performed investigating the expected significance of the proposed measurement with the current dataset collected by the LHCb experiment.

Finally, the study of $B^+ \to K^+\mu^+\mu^-$ decays is presented as an example of a direct search for New Physics. The Run-I dataset collected by the LHCb experiment, corresponding to 3 fb^{-1}, is analysed by looking for an hypothetical new scalar particle produced via $B^+ \to \chi K^+$ decays and subsequently decaying into a muon pair. The signal would manifest itself as an excess in the di-muon invariant mass distribution over the Standard Model background. No significant excess is observed in the accessible ranges of mass $250 < m_\chi < 4700$ MeV$/c^2$ and lifetime $0.1 < \tau_\chi < 1000$ ps. Therefore, upper limits on the branching fraction $\mathcal{B}(B^+ \to K^+\chi(\to \mu^+\mu^-))$ at 95% confidence level are set as a function of m_χ and τ_χ. These are the most stringent limits up-to-date on the process and constrain several inflaton models predicting the existence of a new light particle.

CONTENTS

INTRODUCTION

The Standard Model of particle physics (SM) aims to explain, in a unified theory, the elementary constituents of the Universe and their interactions. The SM has been successfully tested with extraordinary precision in the last few decades, however, a series of phenomena that can not be predicted within the SM have been observed in different fields. Therefore, the current theory is experimentally found to be incomplete.

Experiments designed to explore possible new phenomena are active in several fields, *e.g.* particle and astro-particle physics, aiming to detect signs of physics beyond the Standard Model. These researches can be performed based on two complementary strategies, *direct* or *indirect* approaches. The former intents to directly observe signs of hypothetical new particles, *e.g.* reconstructing their decay products or detecting their interaction with SM particles, while indirect searches aim to perform precise measurements and, in case of deviations from the SM predictions, to reveal possible contributions from new phenomena. This dissertation focuses on the physics case that can be explored by studying $b \to s\ell\ell$ transitions and investigates the discovery potential of such decays both in direct and indirect perspectives.

Despite no clear sign of New Physics (NP) has been observed so far, a series of non-standard measurements has been recently reported in flavour physics. These deviations, although non statistically significant when considered separately, show a coherent pattern that can be interpreted as a hint of NP. Several of these intriguing results consist of measurements of rare b-meson decays, including a large variety of $b \to s\ell\ell$ transitions. Nevertheless, the weak knowledge of theoretical uncertainties in these decays and the limited available statistics currently limit any claim of physics beyond the Standard Model.

The LHCb experiment [1] is an excellent example of precision measurements in flavour physics. Its designed characteristics, optimised for the study of b and c-hadron decays, together with the enormous collected datasets, allow the investigation of the flavour sector of the Standard Model with unprecedented precision.

In this thesis, a novel approach to unambiguously distinguish theoretical uncertainties from possible New Physics contributions in $B^0 \to K^{*0}\ell^+\ell^-$ decays is presented. This consists of an amplitude fit to the decay rate and is first applied to $B^0 \to K^{*0}\mu^+\mu^-$ decays [2]. The proposed method combines the state-of-the-art theoretical knowledge with the full experimental information that can be extracted from the kinematic distributions of signal decays. Afterwards, prospects for a simultaneous amplitude fit to $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$

decays are examined as advantageous solution to solve the interplay between hadronic uncertainties and NP effects [3].

Moreover, the results of a direct search for New Physics based on the analysis of $B^+ \to K^+\mu^+\mu^-$ decays performed on pp collision data corresponding to an integrated luminosity of 3 fb^{-1} collected by the LHCb experiment are presented. This search looks for the production of a hypothetical new scalar particle, χ, in $B^+ \to K^+\chi$ decays, with χ decaying into an opposite sign muon pair [4].

Furthermore, the performance of the Tracker Turicensis detector (TT) is examined by mean of a modified algorithm that allows to evaluate the detector resolution, efficiency and alignment in an unbiased way.

This dissertation is organised as follows; chapter 1 briefly introduces the properties of the Standard Model and illustrates possible NP scenarios, chapter 2 presents a detailed description of the theoretical framework employed to study $B^0 \to K^{*0}\ell^+\ell^-$ decays together with a review of the current experimental results, chapter 3 is dedicated to profile the hadronic uncertainties in $B^0 \to K^{*0}\ell^+\ell^-$ decays, while a general description of the LHCb detector is given in chapter 4, where particular attention is dedicated to the performance of the Tracker Turicensis, chapter 6 describes the results of the search for $B^+ \to K^+\chi(\to \mu^+\mu^-)$ decays, chapter 7 discusses the experimental sensitivity foreseen at LHCb for the amplitude analysis of $B^0 \to K^{*0}\mu^+\mu^-$ decays and, finally, conclusions to this work are given in chapter 8.

The sum of things is unlimited,
and they all change into one another.
The All includes the empty as well as the full.
The worlds are formed when atoms fall into the void
and are entangled with one another;
and from their motion as they increase in bulk,
arises the substance of the stars.

— Leucippus (fl. 450 B.C.)

1

THE STANDARD MODEL AND BEYOND

The Standard Model of particle physics stands out as one of the most successful theories of the last sixty years. It describes the behaviour of elementary particles (*i.e.* particles without internal substructure) and their interactions, incorporating in a single portrait strong, weak and electromagnetic forces. Particles are grouped based on their spin quantum number and are divided in *fermions*, half-integer spin particles, and *bosons*, integer spin particles and mediators of the various interactions. Fermions are the constituents of ordinary matter of the Universe and are in turn divided between *quarks* and *leptons*. While leptons are neutral with respect to the strong interaction, quarks appear in three different colour-charges: *red*, *blue* and *green*. Since only colourless objects can freely propagate in Nature, quarks are bound in mesons (quark-antiquark system) or baryons (three quarks objects).[1] Both groups of fermions appear in three generations, of two members each, and the resulting six types of fermions are known as *flavours*. The main components of the SM are summarised in Fig. 1.

The Standard Model has been experimentally verified with astonishing precision for many years,[2] with experimental measurements confirming its accuracy at different energy scales, from low-energy β-decay to high energy collider experiments. Nevertheless, many fundamental questions still remain open and suggest the necessity for an extension of the SM. These can be formulated as

i. *The flavour puzzle* - the flavour sector of the SM contains 20 free parameters[3] to describe masses and mixing of quarks and

---

[1] Bound states made of more exotic combinations of (anti-)quarks, known as *tetraquarks* and *pentaquarks*, have been recently discovered [5,6].

[2] One of most noticeable achievements of this theory has been the ability to predict the existence of particles that would have been only subsequently discovered, from the *charm* quark, whose discovery gave origin to *November revolution* (1974) [7,8], to the W and Z bosons in the '80s [9–12] and the third generation of quarks [13–15], until the most recent Higgs boson in 2012 [16,17].

[3] A total of 20 parameters is obtained considering a minimally extended SM (νMSM) to include neutrino masses and oscillations.

leptons, while only a total of 5 parameters is required to characterise gauge interactions and bosons masses; why there are three generations of fermions and what is the origin of their mass hierarchy make the flavour sector the most puzzling part of the SM;

ii. *The hierarchy problem* - the observed value of the Higgs boson mass leads to a non-trivial large fine-tuning mechanism in absence of New Physics at an energy scale lower than the Planck scale;[4]

iii. Why the gravitational force is not included in a common description with the other three fundamental interactions?

Beyond these intriguing, albeit purely speculative interrogations, a set of observations that cannot be explained within the Standard Model can be found in particle and astroparticle physics, as well as in cosmology. Examples are

i. Many experimental observations demonstrated that only a tiny fraction of the matter present in the Universe belongs to the components of the SM, while the vast majority corresponds to the so called *Dark Matter*, which interacts only gravitationally with the SM [19];[5]

ii. The CP violation[6] measured in weak interactions of quarks is by far too small to explain the matter-antimatter asymmetry observed in the Universe;

iii. The Standard Model predicts massless neutrinos, while oscillations between the three neutrino families are observed and unambiguously indicate non-zero neutrinos masses.

Therefore, it is commonly accepted that the Standard Model is an incomplete theory and a huge effort in the field is addressed to look for signs of New Physics.

1.1 THE STRUCTURE OF THE STANDARD MODEL

The Standard Model is formulated in terms of a quantum field theory, where bosons and fermions are described as excitations of local

---

[4] The Planck scale identifies the regime of space, time or energy beyond which quantum field theory and general relativity are no longer applicable and quantum effects of gravity are expected to dominate [18].

[5] In addition to Dark Matter, cosmological observations hypothesised the presence of Dark Energy, responsible for the accelerating expansion of the current Universe. The Dark Energy component is estimated to be approximately 70% of the total density of the Universe.

[6] The term CP violation refers to a breaking of the CP symmetry, that is the combination of the charge conjugation C and the parity transformation P. Perfect CP symmetry implies that the law of Physics are conserved if a particle is interchanged with its antiparticle while the spatial coordinates are inverted [20].

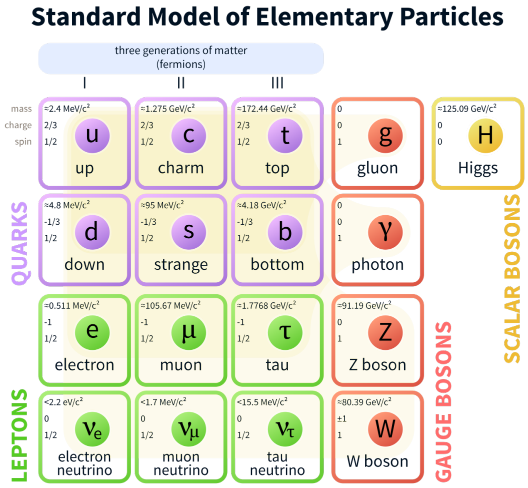


Figure 1: Elementary particles of the Standard Model [21].

quantum fields. The dynamics of the SM is completely determined by the *local* gauge symmetry $SU(3)_C \times SU(2)_L \times U(1)_\mathcal{Y}$, where the $SU(3)_C$ term describes the strong and the $SU(2)_L \times U(1)_\mathcal{Y}$ part the electroweak interaction, and by the fermion content, which consists of five fields with different quantum numbers under the gauge group. These are identified by $\psi = Q_L, d_R, u_R, L_L$ and ℓ_R, where the subscripts L and R indicate the handedness of the fermion and $Q(L)$ and $u/d(\ell)$ are respectively the doublets and singlets for quarks (leptons) under $SU(2)_L$. The five fermion fields appear in three identical replicas and the highly symmetric structure of the gauge sector gives rise to a large *global* flavour symmetry, where the fermions undergo the same gauge interactions. The kinetic term for the fermion fields can be expressed by

$$\mathcal{L}_{\text{kinetic}} = \sum_\psi \bar{\psi} i \gamma^\mu D_\mu \psi \,, \tag{1}$$

where γ^μ are Dirac matrices and the gauge invariance is preserved by the covariant derivative

$$D_\mu = \partial_\mu - i\frac{g_s}{2}\lambda_a G^a_\mu - i\frac{g}{2}\sigma_b W^b_\mu - i\frac{g'}{2}\mathcal{Y} B_\mu \,, \tag{2}$$

where G^a_μ, W^b_μ and B_μ are the $SU(3)_C$, $SU(2)_L$ and $U(1)_\mathcal{Y}$ gauge fields respectively, g_s, g and g' are the associated coupling constants and λ_a, σ_b and $\mathcal{Y}$ are the corresponding fundamental representations, identified by the Gell-Mann matrices, Pauli matrices and hypercharge, respectively.

The local symmetry is spontaneously broken by the introduction of a $SU(2)_L$ scalar doublet H, the Higgs boson, whose vacuum expectation value

$$\langle H \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{3}$$

gives mass to the fermion fields via the Yukawa interaction

$$\mathcal{L}_{\text{Yukawa}} = -Y^d_{ij} \bar{Q}^i_L H d^j_R - Y^u_{ij} \bar{Q}^i_L H^c u^j_R - Y^\ell_{ij} \bar{L}^i_L H \ell^j_R + h.c., \tag{4}$$

where i and j label the flavour generations, Y_{ij} are complex coupling matrices and H^c is the complex conjugated of the Higgs doublet $H^c = i\sigma_2 H^*$. Furthermore, the fact that the matrices $Y^{d,u,\ell}$ are not proportional to the identity gives rise to an explicit breaking of the global flavour symmetry.

In general, the diagonalisation of each Yukawa coupling requires two independent unitary matrices, $V_L Y V_R^\dagger = \lambda^{\text{diag}}$. In the lepton sector the invariance of the gauge Lagrangian under $SU(3)$ allows to freely choose the two matrices necessary to diagonalise Y^ℓ without breaking gauge invariance. This is not the case for the quark sector, where only three of the four needed rotation matrices to diagonalise both Y^d and Y^u can be freely chosen, leading to a non-identical mass and flavour-eigenstate bases. Choosing the basis where Y^d is diagonal,

$$Y^d = \lambda_d\,, \quad Y^u = V^\dagger \lambda_u\,, \tag{5}$$

a single non-trivial unitary mixing, V, is left and it represents the misalignment of the up- and down-quark mass bases. Within arbitrary phase-convention differences, the obtained 3×3 matrix results in the Cabibbo-Kobayashi-Maskawa (CKM) matrix [22]

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{6}$$

where each element V_{ij} represents the mixing between quark flavours. The CKM matrix shows a strongly hierarchical pattern that makes the Wolfenstein parametrisation [23] a convenient way to exhibit the mixing between generations of quarks in a more explicit way,

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{7}$$

where $\lambda = \sin\theta_c \sim 0.22$, with θ_c known as the Cabibbo angle describing the mixing between the first two generations of quarks [24], while A, ρ and η are real parameters of the order of unity. The two parameters ρ and η give origin to a single complex phase, this *weak* phase raises from the unitarity of the CKM matrix and is the only known source of CP violation in the Standard Model.

When moving from the flavour to the mass-eigenstate basis, the CKM matrix naturally appears in the interaction between quarks and $SU(2)$ gauge bosons described in Eq. 1 and 2. Therefore, the charged current mediated by the combination of gauge bosons

$$W^{\pm}_{\mu} = \frac{1}{\sqrt{2}}(W^1_{\mu} \mp iW^2_{\mu}), \tag{8}$$

reads as

$$\mathcal{L}_{CC} = -\frac{g}{\sqrt{2}}\bar{u}^i_L \gamma^{\mu} V_{ij} d^j_L W^+_{\mu} + h.c., \tag{9}$$

where $\mathcal{L}_{CC}$ has been limited to the quark sector. It is evident from the previous equation how the charged W boson rules a flavour changing transition between up-type and down-type quarks, with a coupling determined by the weak constant g and the relevant element of the CKM matrix V_{ij}.

On the other hand, the neutral Z boson originates from the combination of the remaining $SU(2)$ and $U(1)$ gauge fields

$$\begin{pmatrix} W^3_{\mu} \\ B_{\mu} \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} Z^0_{\mu} \\ A_{\mu} \end{pmatrix}, \tag{10}$$

where the vector field A_{μ} is identified as the photon, γ, and the weak mixing angle θ_W is defined as

$$\cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}} \quad \text{and} \quad \sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}} \tag{11}$$

and shows the unification of the weak and electromagnetic interactions. Both Z and γ are neutral bosons and couple exclusively to same flavour fermions with a structure given by

$$\begin{aligned} \mathcal{L}_{NC} = \bar{\psi}^i \gamma^{\mu} \Big[& A_{\mu}\left(g\frac{\sigma_3}{2}\sin\theta_W + g'\mathcal{Y}\cos\theta_W\right) \\ & + Z_{\mu}\left(g\frac{\sigma_3}{2}\cos\theta_W - g'\mathcal{Y}\sin\theta_W\right)\Big]\psi_i, \end{aligned} \tag{12}$$

where the first part identifies the electromagnetic interaction under the conditions

$$g\sin\theta_W = g'\cos\theta_W = e \quad \text{and} \quad \mathcal{Y} = Q - \frac{\sigma_3}{2}, \tag{13}$$

and Q denotes the electromagnetic charge operator.

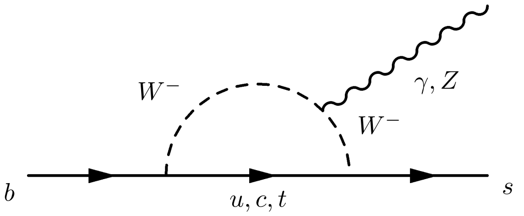


Figure 2: Feynman diagram of flavour changing neutral current $b \to s$ processes.

As a consequence, *Flavour Changing Neutral Current*[7] (FCNC) transitions are forbidden at the tree level in the Standard Model. These processes can only happen at the loop level (see Fig. 2) and are therefore suppressed with respect to tree level decays. This suppression makes FCNC transitions an ideal framework to look for New Physics. In fact, hypothetical new particles can either be directly detected or indirectly modify physical observables from their Standard Model predictions.

1.2 BEYOND THE STANDARD MODEL

Searches for physics beyond the Standard Model can be separated in two macroscopic areas: *direct* and *indirect* searches. While the former requires to reach the energy necessary to produce the given particle, the latter aims to perform precise measurements and detect possible deviations from SM predictions.

Since no evidence of New Physics has been observed yet, it is of fundamental importance to concurrently pursue both approaches. In fact, if the energy scale of NP is higher than what can be explored with current colliders, of the order of few TeVs, the only possible way for a discovery of NP is through indirect searches. On the other hand, if the scale of NP is low enough to be directly accessed by the current experiments, direct searches constitute the privileged channel.

1.2.1 *Direct searches in the dark sector*

A possible extension of the Standard Model consists in introducing a *dark sector* [25], which can contain a rich variety of distinct particles operating through forces that are hitherto unknown. Under this scenario, dark sector particles would be gauge singlet states with respect to the SM, and only be able to communicate with known particles via weakly interacting messengers through one of four portals: the vector, axion, Higgs, and neutrino portals.

---

[7] The term Flavour Changing Neutral Current refers to all the processes that modify the flavour of a particle but keeping unchanged its electric charge.

For the scope of this dissertation, only the Higgs portal is further discussed. In this scenario, the new scalar particle can mix with the Higgs boson and, therefore, can be explored at high-energy colliders. A prime example of such theories are so-called *inflaton* models [26, 27], where a hypothetical new scalar field was responsible for the exponential expansion of the Universe after the *big bang*.

Inflation theories have been formulated in the early '80s [28] as proposed solutions to the *horizon problem* and *flatness problem*. The former regards the homogeneous and isotropic aspect of the Universe. This condition of causally connected Universe cannot be reached assuming the evolution of the Universe being only dominated by radiation and matter.[8] Inflationary mechanisms postulate that currently disconnected regions were initially part of a single Hubble volume and underwent an exponential expansion in the first 10^{-32} seconds of the Universe that led them outside the observable horizon. Secondly, the flatness problem concerns the density of matter in the Universe that is found to be very close to the critical value required to have a flat Universe. Extrapolating this peculiarity to the early Universe, the obtained value is found to be identical to the critical density apart from one part of 10^{-60}. This incredible fine-tuned condition, without which the Universe would have collapsed or exploded much before the current epoch, is accommodated by the rapid expansion during inflation that is able to rapidly reduce $1 - \Omega$ to arbitrarily small values, where Ω is the ratio between the actual density and its critical value. Finally, these models can also help to solve the hierarchy problem and to explain the baryon asymmetry in the Universe [29, 30].

In conclusion, cosmological inflation predicts the existence of a new scalar inflaton field, χ, that is responsible for the thermal equilibrium of the early Universe. Moreover, some inflaton models allow the mixing of the inflaton with the SM Higgs boson. The Lagrangian of these extended models reads [26]

$$\mathcal{L} = \mathcal{L}_{\mathrm{SM}} + \frac{1}{2}\partial_\mu \chi \partial^\mu \chi + \frac{1}{2} m_\chi^2 \chi^2 - \frac{\beta}{4}\chi^4 - \lambda\left(H^\dagger H - \frac{\alpha}{\lambda}\chi^2\right)^2, \qquad (14)$$

where $\mathcal{L}_{\mathrm{SM}}$ is the SM Lagrangian without the Higgs potential and both quartic coupling, β, and Higgs-to-inflaton mixing, α, are very weak, while Higgs boson self-coupling, λ, is strong ($\lambda \gtrsim 0.1$). This new scalar potential has SM-like vacuum expectation values

$$\langle H \rangle = \frac{v}{\sqrt{2}} \quad \text{and} \quad \langle \chi \rangle = \sqrt{\frac{\lambda}{2\alpha}}\, v, \qquad (15)$$

---

[8] Two widely separated regions that move apart from each other faster than the speed of light never come into causal contact and, therefore, cannot reach thermal equilibrium. Assuming only a Universe dominated by matter/radiation, thermal equilibrium would be found only within $\sim 1.5°$, *i.e.* radiation coming from wider angles would be originated from disconnected regions.

with $v \approx 246$ GeV and gives origin to the scalar masses

$$m_H = \sqrt{2\lambda} v \quad \text{and} \quad m_\chi = m_H \sqrt{\frac{\beta}{2\alpha}}, \tag{16}$$

where the rotation between the mass and gauge eigenstates of the Higgs and the inflaton fields is described by the mixing angle

$$\theta = \sqrt{\frac{2\alpha}{\lambda}} = \frac{\sqrt{2\beta} v}{m_\chi}. \tag{17}$$

The size of this mixing angle establishes the magnitude of the mixing process, *i.e.* large values result in strong coupling of the inflaton to the SM sector and vice versa.

Under these assumptions the mass of the inflaton is only weakly bound. Constraints are imposed by the reheating temperature[9] of the Universe, that must be above the electroweak scale [31]; stronger mixing implies more efficient energy transfer from the inflaton to the SM particles and, consequently, higher reheat temperature in the early Universe. Hence, combining Eqs. 16 and 17, higher reheat temperature results in lighter inflaton. Secondly, quantum corrections originated from the mixing mechanism to the bare inflationary potential $\beta\chi^4/4$ should be small [26], this settles a upper bound on α and, as a consequence, a lower bound on the inflaton mass.

These conditions identify two possible mass ranges, an *heavy* inflaton, with $200 \lesssim m_\chi \lesssim 400$ GeV$/c^2$, and a *light* inflaton, with m_χ between $\mathcal{O}(10)$ MeV and $\mathcal{O}(10)$ GeV [26]. In the case of heavy inflaton the tiny coupling with the Standard Model disfavours any direct searches in the foreseeable future, since it would require an amount of statistics at high energy collisions that goes beyond the possibilities of current and planned future hadronic machines. On the other hand, searches for light inflaton appear promising given the large available datasets at low energy at current experiments.

Furthermore, two fundamental properties must be investigated in order to exploit direct searches opportunities: inflaton decay products and lifetime. The inflaton inherits through the Higgs mixing Yukawa-type couplings to the SM. Thus, its branching ratio coincides with that of a SM Higgs boson of mass m_χ [32], once taken into account the small mixing angle - see Fig. 3. Secondly, the lifetime of the inflaton depends on the accessible final states and is inversely proportional to the value of the quartic coupling β (or, equivalently, inversely proportional to the square of the mixing angle, θ^2). Figure 4 shows the obtained lifetime

---

[9] The term reheating refers to the end of the inflationary expansion, when the inflaton field decays into SM particles and the radiation-dominated phase of the Universe begins. The name originates from the thermalisation process subject to the production of the SM fields after the drop of temperature during inflation.

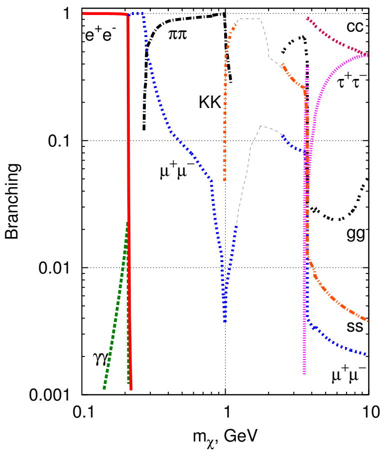


Figure 3: Inflaton branching ratios as function of the inflaton mass [33]. The region between $1.5 \lesssim m_\chi \lesssim 2.5$ GeV$/c^2$ suffers from significant QCD uncertainties, due to the opening of new hadronic channels, primarily $\chi \to \eta\eta$.

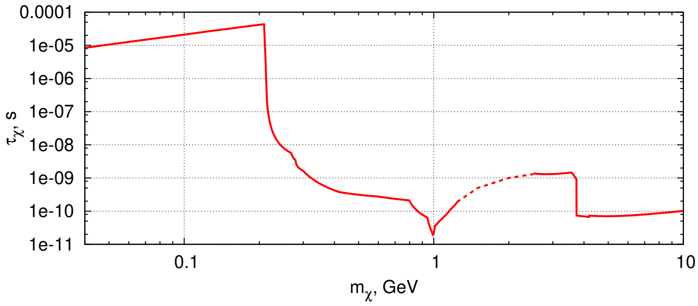


Figure 4: Inflaton lifetime as function of the inflaton mass for $\beta = \beta_0$, while for other values the lifetime is inversely proportional to the quartic coupling β (or, equivalently, inversely proportional to the square of the mixing angle, θ^2) [26].

for the choice $\beta = \beta_0 \equiv 1.5 \times 10^{-13}$.[10] The resulting lifetime can span several orders of magnitude, $\mathcal{O}(10^{-6} \div 10^{-12})$ s, and may become relevant for detection purpose at high energy beam experiments.

Rare B and K-meson decays are an ideal framework to look for light inflaton production. Figure 5 shows a possible diagram of FCNC processes that allows to directly access the inflaton via mixing with the SM Higgs boson. Previous searches for such phenomenon have been performed by the CHARM [34], NA48/2 [35] and LHCb [36] experiments. While the first one sets a lower limit on the inflaton mass

---

[10] The value of the quartic coupling $\beta_0 \equiv 1.5 \times 10^{-13}$ [26] assumes the quartic term $\beta\chi^4/4$ dominating the potential during the inflationary phase of the Universe and corresponds to a number of e-folding of approximately $N_e \simeq 62$, where N_e determines the timescale/magnitude of the exponential expansion of the early Universe $\sim e^{N_e}$.

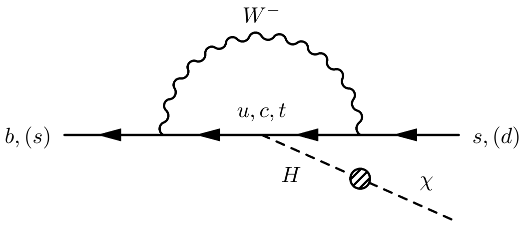


Figure 5: Feynman diagram of the production of the inflaton χ via flavour changing neutral current, where the inflaton interacts with the Standard Model by mixing with the Higgs boson.

$m_\chi \gtrsim 280\ \mathrm{MeV}/c^2$, the other two exclude a subregion of the parameter space (m_χ, θ^2). The NA48/2 collaboration analysed decays of $K^\pm \to \pi^\pm\mu^+\mu^-$, while LHCb published a search using the $B^0 \to K^{*0}\mu^+\mu^-$ decays, where in both cases the inflaton is assumed to decay into a muon pair.[11] However, the different quark transitions $K \to \pi$ and $B \to K$ have a strong impact on the sensitivity of the two decays. In fact, due to the large Yukawa coupling with the Higgs, the inflaton production is dominated by the contribution of the *top* quark in the loop diagram of Fig. 5. Thus, $s \to d$ transitions are suppressed compared to $b \to s$ transitions by the sub-leading terms of the CKM matrix V_{ts} and V_{td}.

In the following, the interest is restricted to B-meson decays to either a kaon (pseudo-scalar) or a K^{*0} (vector) in the final state. The branching ratio of $B^+ \to K^+\chi$ decays is predicted to be [37]

$$\mathcal{B}(B^+ \to K^+\chi) = 4 \times 10^{-7} \left(\frac{\theta}{10^{-3}}\right)^2 F_K(m_\chi)\sqrt{\lambda}\,, \tag{18}$$

where the dependence on the unknown inflaton mass resides in the form factors F and in the phase space suppression $\lambda = (1 - m_B^{-2}(m_\chi + m_K)^2) \times (1 - m_B^{-2}(m_\chi - m_K)^2)$. An analogous formalism can be used to express neutral $B^0 \to K^{*0}\chi$ decays where the only difference resides in the available phase space given by $\sqrt{\lambda^3}$ due to the presence of a vector in the final state [37].

Figure 6 shows a comparison of the branching ratios for the two channels. The $B^0 \to K^{*0}\chi$ mode is suppressed for large inflaton masses, $m_\chi \gtrsim 2\ \mathrm{GeV}/c^2$, due to angular momentum conservation. At low mass, where the available energy from the B decay compensates the angular momentum barrier factor, both channels have similar sensitivity.

A search based on $B^+ \to K^+\mu^+\mu^-$ decays, that is expected to have comparable or better sensitivity than the $B^0 \to K^{*0}\mu^+\mu^-$ search of Ref. [36], is presented in Chapter 6.

---

[11] The choice to restrict the search to a final state containing muons is experimentally motivated. Muons give a unique signature in the detector, typically surrounded by a heavy hadronic environment, and are measured with an excellent precision.

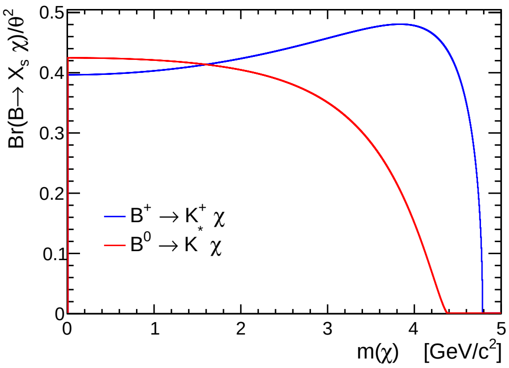


Figure 6: Branching ratio of the decays $B \to K^+\chi$ and $B \to K^*\chi$ as from [37].

1.2.2 *Indirect searches in flavour physics*

While results of direct searches of physics beyond the Standard Model are of straightforward interpretation, indirect searches can point to specific directions to understand the nature of New Physics. In addition, indirect measurements can be the only possible way to investigate new phenomena in case that the scale of NP is too high to be directly accessed at colliders. Precision measurements in the flavour sector are an excellent example of such indirect tests of the SM. A (non-exhaustive) list of representative measurements consists of

i. *CP violation and the unitarity triangle* - the unitarity of the CKM matrix leads to a set of relations between its elements, one of these,

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0\,, \tag{19}$$

is of particular interest since it involves the sum of terms all of the same order in λ and is historically known as the unitarity triangle. The measurement of the three angles of the unitarity triangle, accessible by several independent particle decays, represents a strong probe of the CP violation mechanism in the SM [38,39].

ii. *Rare hadron decays* - particle decays that are strongly suppressed in the Standard Model, *e.g.* loop processes/flavour changing neutral currents, are expected to be particularly sensitive to New Physics. In fact, virtual particles can contribute to the decays and enhance or suppress their branching ratio, modify angular distributions or introduce new sources of CP violation. Among these processes, purely leptonic decays as $B_{s(d)} \to \ell^+\ell^-$ are prime examples of indirect test of the SM. These FCNC decays undergo an additional helicity suppression by m_ℓ^2/M_B^2 and, within the SM, have branching ratio of the order of $\mathcal{O}(10^{-9})$ and $\mathcal{O}(10^{-10})$ for $B_s \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$, respectively.

A recent publication of the CMS and LHCb experiments [40] reported the first observation of $B_s \to \mu^+\mu^-$ decays and evidence of $B^0 \to \mu^+\mu^-$ decays, where both decay rates have been found to be compatible with the SM predictions. Despite less clean from a theoretical point of view, $b \to s\ell\ell$ transitions provide an analogously powerful investigation of NP and great attention is addressed in this dissertation to the study of $B^0 \to K^{*0}\ell^+\ell^-$ decays.

ii. *Forbidden Standard Model decays* - the Lagrangian of the SM is invariant under $U(1)_e \times U(1)_\mu \times U(1)_\tau$ rotations of the lepton fields. This discrete symmetry implies the conservation of the leptonic quantum number, commonly known as *Lepton Flavour Conservation*. On the contrary, the experimental evidence of non-zero mass neutrinos allows Lepton Flavour Violation (LFV) in the neutral lepton sector of the SM and, potentially, also in the charged sector. However, the predicted contribution from neutrino oscillations to charged LFV processes is extremely small, below 10^{-54}. Therefore, any sizeable contribution to decays that do not conserve lepton flavour would be a clear sign of NP.

These indirect searches provide important constraints on several New Physics scenarios. Popular examples are: Z' models [41–44], that introduce an additional $U(1)'$ gauge symmetry allowing neutral flavour-changing coupling at the tree level, *e.g.* $Z' \to bs$; and leptoquark models [45–47], that postulate the existence of a new boson that carries both quark and lepton quantum numbers. An interesting implication of these leptoquark models is the enhanced branching ratio of LFV decays (*e.g.* $B \to K\tau\mu$, $B_s \to \tau\mu$ [48,49]) to an extent that can be accessed at current colliders. Finally, further models that can be inspected by indirect approaches consist of composite Higgs [50,51] and extra-dimensions theories [52].

"E pur si move!"

— attributed[†] to Galileo Galilei, (1633)

2

THE $B^0 \to K^{*0}\ell^+\ell^-$ DECAY AND THE FLAVOUR ANOMALIES

This chapter discusses all the theoretical and experimental aspects relevant for the study of $B^0 \to K^{*0}\ell^+\ell^-$ decays.[1] These decays are characterised by $b \to s\ell\ell$ transitions and are found to be rare, with branching ratio of the order of $\mathcal{O}(10^{-6})$. In the first part, a complete theoretical description is presented, followed by a review of the current status of experimental measurements in the field.

2.1 THE STANDARD MODEL AS AN EFFECTIVE FIELD THEORY

It is common in flavour physics to consider the Standard Model as a low-energy Effective Field Theory (EFT) [53]. This is possible since decays of B, D and K mesons are ruled primarily by two different energy scales: the electroweak scale, characterised by the W boson mass, which determines the flavour-changing transition at the quark level, and the scale of strong interactions Λ_{QCD}, related to the hadron formation. Thus, the description of meson decays is simplified by integrating out the heavy SM fields (W and Z bosons, as well as the top quark) and rewriting the Lagrangian density in terms of local operators of dimension six and higher, containing only light SM fermions and suppressed by inverse powers of the W mass as

$$\mathcal{L}_{\text{eff}} = -4\frac{G_F}{\sqrt{2}}\sum \mathcal{C}_i(\mu)\mathcal{O}_i(\mu)\,, \tag{20}$$

where the sum runs over the complete basis of operators $\mathcal{O}_i$. In general, the effective couplings $\mathcal{C}_i(\mu)$, known as Wilson coefficients, depend on the renormalisation scale μ. This scale acts as a separator of short-distance effects, included in the $\mathcal{C}_i(\mu)$, and long-distance effects, encoded in the operators.

The Wilson coefficients are firstly evaluated at the electroweak scale ($\mu \approx m_W$) by matching the decay amplitudes of the full theory and of

---

[†] The origin of the phrase "And yet it moves", commonly attributed to the Italian mathematician, physicist and philosopher Galileo Galilei, has never been historically demonstrated, since it does not appear either in his trial's acts or in private manuscripts. The sentence is first reported by writer Giuseppe Baretti in his book *The Italian Library*, London (1757).

[1] Throughout this dissertation, the symbol K^{*0} refers to the neutral $K^{*0}(892)$ vector resonance.

the effective theory. Afterwards, renormalisation group equations [54, 55] are used to evolve the Wilson coefficients from the electroweak scale down to the energy scale of the physical process ($\mu \approx m_B$). Following this procedure, the presence of New Physics at high scale can modify the initial values of the Wilson coefficients, appearing as a shift with respect to the Standard Model predictions, *i.e.* $\mathcal{C}_i = \mathcal{C}_i^{\mathrm{SM}} + \mathcal{C}_i^{\mathrm{NP}}$.

In the case of $b \to s\ell\ell$ transitions, the effective weak Lagrangian is given by [56]

$$\begin{aligned}\mathcal{L}_{\mathrm{eff}}^{b\to s} = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \left\{ [\mathcal{C}_1 \mathcal{O}_1^c + \mathcal{C}_2 \mathcal{O}_2^c] + \frac{\alpha_e}{4\pi} [\mathcal{C}_7 \mathcal{O}_{7\gamma} + \mathcal{C}_9 \mathcal{O}_{9V} + \mathcal{C}_{10} \mathcal{O}_{10A}] \right\} \\ + \mathcal{O}\left(\frac{V_{ub} V_{us}^*}{V_{tb} V_{ts}^*}, \mathcal{C}_{3,\ldots,6}, \alpha_s \mathcal{C}_8 \right),\end{aligned} \tag{21}$$

where the relevant dimension-six operators are the $b \to sc\bar{c}$ four-quarks operators

$$\begin{aligned}\mathcal{O}_1^c &= \bar{b}_L^\alpha \gamma^\mu c_L^\alpha \bar{c}_L^\beta \gamma_\mu s_L^\beta, \\ \mathcal{O}_2^c &= \bar{b}_L^\alpha \gamma^\mu c_L^\beta \bar{c}_L^\beta \gamma_\mu s_L^\alpha,\end{aligned} \tag{22}$$

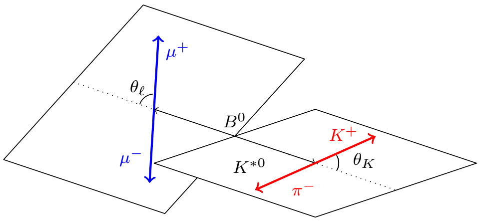
and radiative/semileptonic operators

$$\begin{aligned}\mathcal{O}_{7\gamma} &= \frac{e}{16\pi^2} m_b \bar{b}_R^\alpha \sigma^{\mu\nu} F_{\mu\nu} s_L^\alpha, \\ \mathcal{O}_{9V} &= \frac{1}{2} \bar{b}_L^\alpha \gamma^\mu s_L^\alpha \bar{\ell} \gamma_\mu \ell, \\ \mathcal{O}_{10A} &= \frac{1}{2} \bar{b}_L^\alpha \gamma^\mu s_L^\alpha \bar{\ell} \gamma_\mu \gamma_5 \ell,\end{aligned} \tag{23}$$

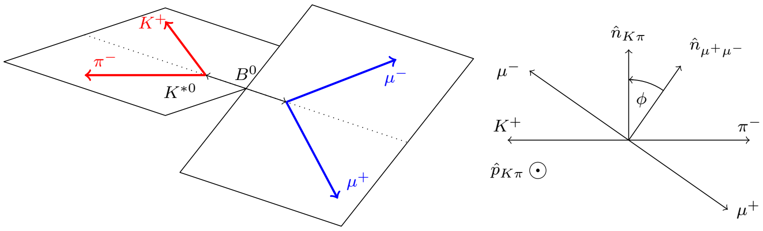
where b, s, c and ℓ are the SM fermionic fields, α and β denote colour indices and $F_{\mu\nu}$ is the electromagnetic tensor. The initial condition of $\mathcal{O}_{7\gamma}$, $\mathcal{O}_{9V}$ and $\mathcal{O}_{10A}$ are found to be particularly sensitive to NP. These operators correspond to the photon production $b \to s\gamma$ and to the *vector* and *axial* leptonic currents, respectively. Finally, by virtue of the left-handed nature of the weak interaction, right-handed current operators $\mathcal{O}_i'$, obtained by replacing $q_{L(R)} \to q_{R(L)}$ in Eq. 23, are suppressed in the Standard Model. Nevertheless, these chirality-flipped operators can receive non-negligible contributions from NP models involving a different helicity structure.

2.2 THE $B^0 \to K^{*0}\ell^+\ell^-$ DECAYS

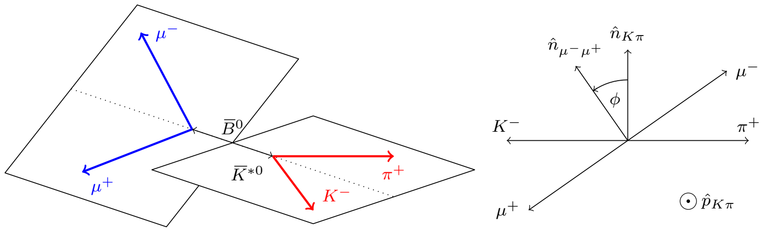
Rare $B^0 \to K^{*0}\ell^+\ell^-$ decays, where the K^{*0} is reconstructed as $K^{*0} \to K^+\pi^-$, provide a rich framework to investigate possible New Physics effects. In particular, the presence of a vector meson in the final state introduces additional degrees of freedom to the system, which can be parametrised in terms of the di-lepton invariant mass squared, q^2, and the angular distributions between the final-state particles.



(a) θ_K and θ_ℓ definitions for the B^0 decay



(b) ϕ definition for the B^0 decay



(c) ϕ definition for the $\overline{B}^0$ decay

Figure 7: Definition of the decay angles θ_ℓ, θ_K and ϕ in $B^0 \to K^{*0}\ell^+\ell^-$ decays.

Figure 7 shows the definition of the three decay angles, $\vec{\Omega} = (\theta_\ell, \theta_K, \phi)$, that completely characterise the topology of the decay. The angle θ_ℓ is the angle between the direction of the ℓ^+ (ℓ^-) and the direction opposite to that of the B^0 ($\overline{B}^0$) in the di-lepton reference frame. The angle θ_K is the angle between the direction of the K^+ (K^-) and the direction of the B^0 ($\overline{B}^0$) in the K^* ($\overline{K}^*$) reference frame. The angle ϕ is the angle between the plane of the di-lepton pair and the plane of the K and π originated from the K^* ($\overline{K}^*$) decay in the B^0 ($\overline{B}^0$) reference frame.

2.2.1 *Differential decay rate*

The differential decay rate of $B^0 \to K^*(\to K^+\pi^-)\ell^+\ell^-$ can, summing over the lepton spins, assuming an on-shell K^* of narrow width and integrating over the $K\pi$ invariant mass, be written as [57]

$$\begin{aligned} \frac{32\pi}{9}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\,\mathrm{d}\vec{\Omega}} &= (I_{1c} + I_{2c}\cos 2\theta_\ell)\cos^2\theta_K \\ &+ (I_{1s} + I_{2s}\cos 2\theta_\ell + I_{6s}\cos\theta_\ell)\sin^2\theta_K \\ &+ (I_3\cos 2\phi + I_9\sin 2\phi)\sin^2\theta_K\sin^2\theta_\ell \\ &+ (I_4\cos\phi + I_8\sin\phi)\sin 2\theta_K\sin 2\theta_\ell \\ &+ (I_5\cos\phi + I_7\sin\phi)\sin 2\theta_K\sin\theta_\ell\,, \end{aligned} \tag{24}$$

where $I_i = I_i(q^2)$ are q^2-dependent angular coefficients. These coefficients can be conveniently expressed in terms of transversity amplitudes, $\mathcal{A}^{L,R}_{0,\perp,\|}$ and $\mathcal{A}_t$, where $\lambda = 0, \perp, \|, t$ refers to the polarisation of the K^{*0} meson and the indices L and R denote the chirality of the lepton current [58]. Considering only the operators relevant in the Standard Model, the angular coefficients are

$$\begin{aligned} I_{1s} &= \frac{2+\beta_l^2}{4}\Big[|\mathcal{A}_\perp^L|^2 + |\mathcal{A}_\|^L|^2 + (L\to R)\Big] \\ &\quad + \frac{4m_l^2}{q^2}\,\mathcal{R}e\Big(\mathcal{A}_\perp^L\mathcal{A}_\perp^{R*} + \mathcal{A}_\|^L\mathcal{A}_\|^{R^*}\Big), \\ I_{1c} &= \Big[|\mathcal{A}_0^L|^2 + |\mathcal{A}_0^R|^2\Big] + \frac{4m_l^2}{q^2}\Big[|\mathcal{A}_t|^2 + 2\,\mathcal{R}e(\mathcal{A}_0^L\mathcal{A}_0^{R^*})\Big], \\ I_{2s} &= \frac{\beta_l^2}{4}\Big[|\mathcal{A}_\perp^L|^2 + |\mathcal{A}_\|^L|^2 + (L\to R)\Big], \\ I_{2c} &= -\beta_l^2\Big[|\mathcal{A}_0^L|^2 + |\mathcal{A}_0^R|^2\Big], \\ I_3 &= \frac{\beta_l^2}{2}\Big[|\mathcal{A}_\perp^L|^2 - |\mathcal{A}_\|^L|^2 + (L\to R)\Big], \\ I_4 &= \frac{\beta_l^2}{\sqrt{2}}\,\mathcal{R}e\Big[\mathcal{A}_0^L\mathcal{A}_\|^{L^*} + (L\to R)\Big], \\ I_5 &= \sqrt{2}\beta_l\,\mathcal{R}e\Big[\mathcal{A}_0^L\mathcal{A}_\perp^{L^*} - (L\to R)\Big], \\ I_{6s} &= 2\beta_l\,\mathcal{R}e\Big[\mathcal{A}_\|^L\mathcal{A}_\perp^{L^*} - (L\to R)\Big], \\ I_7 &= \sqrt{2}\beta_l\,\mathcal{I}m\Big[\mathcal{A}_0^L\mathcal{A}_\|^{L^*} - (L\to R)\Big], \\ I_8 &= \frac{\beta_l^2}{\sqrt{2}}\,\mathcal{I}m\Big[\mathcal{A}_0^L\mathcal{A}_\perp^{L^*} + (L\to R)\Big], \\ I_9 &= \beta_l^2\,\mathcal{I}m\Big[\mathcal{A}_\perp^L\mathcal{A}_\|^{L^*} + (L\to R)\Big], \end{aligned} \tag{25}$$

with $\beta_l = \sqrt{1 - 4m_l^2/q^2}$. These amplitudes are the physical quantities that describe the decay process[2] and can be expressed as [56, 59]

$$\begin{aligned}
\mathcal{A}_\perp^{L,R} &= N\Big\{\Big[(\mathcal{C}_9 + \mathcal{C}_9') \mp (\mathcal{C}_{10} + \mathcal{C}_{10}')\Big]\mathcal{F}_\perp(q^2) \\
&\qquad + \frac{2m_b M_B}{q^2}\Big[(\mathcal{C}_7 + \mathcal{C}_7')\mathcal{F}_\perp^T(q^2) - 16\pi^2 \frac{M_B}{m_b}\mathcal{H}_\perp(q^2)\Big]\Big\}\,, \\
\mathcal{A}_\parallel^{L,R} &= -N\Big\{\Big[(\mathcal{C}_9 - \mathcal{C}_9') \mp (\mathcal{C}_{10} - \mathcal{C}_{10}')\Big]\mathcal{F}_\parallel(q^2) \\
&\qquad + \frac{2m_b M_B}{q^2}\Big[(\mathcal{C}_7 - \mathcal{C}_7')\mathcal{F}_\parallel^T(q^2) - 16\pi^2 \frac{M_B}{m_b}\mathcal{H}_\parallel(q^2)\Big]\Big\}\,, \\
\mathcal{A}_0^{L,R} &= -N\Big\{\Big[(\mathcal{C}_9 - \mathcal{C}_9') \mp (\mathcal{C}_{10} - \mathcal{C}_{10}')\Big]\mathcal{F}_0(q^2) \\
&\qquad + \frac{2m_b M_B}{q^2}\Big[(\mathcal{C}_7 - \mathcal{C}_7')\mathcal{F}_0^T(q^2) - 16\pi^2 \frac{M_B}{m_b}\mathcal{H}_0(q^2)\Big]\Big\}\,, \\
\mathcal{A}_t &= -2N\,(\mathcal{C}_{10} - \mathcal{C}_{10}')\mathcal{F}_t(q^2)\,,
\end{aligned} \tag{26}$$

where the normalisation factor N is given by

$$N = G_F \alpha_e V_{tb} V_{ts}^* \sqrt{\frac{q^2 \beta_l \sqrt{\lambda}}{3 \cdot 2^{10}\pi^5 M_B}}\,. \tag{27}$$

The functions $\mathcal{F}_\lambda^{(T)}(q^2)$ express the *local* hadronic matrix elements, known as *form factors*,[3] and $\mathcal{H}_\lambda(q^2)$ parametrise the *non-local* hadronic matrix elements. While the form factors can be accessed either from first principles through Lattice QCD simulations [60, 61], or from quark-hadron-duality arguments through QCD Light-Cone Sum Rules [62, 63], the matrix elements of non-local operators result more difficult to determine and are discussed in the next section. The translation between the presented choice of form factor basis and the commonly-used in the literature (see Refs. [64, 65]) reads as

$$\begin{aligned}
\mathcal{F}_\perp &\mapsto \frac{\sqrt{2\lambda}}{M_B(M_B + M_{K^{*0}})} V\,, \\
\mathcal{F}_\parallel &\mapsto \frac{\sqrt{2}(M_B + M_{K^{*0}})}{M_B} A_1\,, \\
\mathcal{F}_0 &\mapsto \frac{(M_B^2 - q^2 - M_{K^{*0}}^2)(M_B + M_{K^{*0}})^2 A_1 - \lambda A_2}{2M_{K^{*0}} M_B (M_B + M_{K^{*0}})\sqrt{q^2}}\,, \\
\mathcal{F}_\perp^T &\mapsto \frac{\sqrt{2\lambda}}{M_B^2} T_1\,,
\end{aligned} \tag{28}$$

---

[2] In quantum mechanics, the transition between one quantum state to an other is determined by the invariant matrix element (or amplitude) between the initial and final states contracted with the relevant operators, whose modulus square gives the probability of the given process.

[3] The form factors describe the hadronic transition between the initial and final state $B^0 \to K^{*0}$.

$$\begin{aligned}
\mathcal{F}_{\parallel}^{T} &\mapsto \frac{\sqrt{2}(M_B^2 - M_{K^{*0}}^2)}{M_B^2} T_2\,,\\
\mathcal{F}_{0}^{T} &\mapsto \frac{\sqrt{q^2}(M_B^2 + 3M_{K^{*0}}^2 - q^2)}{2M_B^2 M_{K^{*0}}} T_2 - \frac{\sqrt{q^2}\lambda}{2M_B^2 M_{K^{*0}}(M_B^2 - M_{K^{*0}}^2)} T_3\,,\\
\mathcal{F}_{t} &\mapsto A_0\,,
\end{aligned}$$

where

$$\lambda = M_B^4 + M_{K^{*0}}^4 + (q^2)^2 - 2M_B^2(M_{K^{*0}}^2 + q^2) - 2q^2 M_{K^{*0}}^2\,. \tag{29}$$

2.2.1.1 *Non-local hadronic matrix elements form analyticity*

Matrix elements involving insertions of non-local operators, $\mathcal{O}_{bscc} \equiv \mathcal{C}_1\mathcal{O}_1 + \mathcal{C}_2\mathcal{O}_2$, are particularly difficult to determine or to assess reliably from first principles and have been the focus of much attention over the last two decades. The parametrisation of these hadronic matrix elements proposed in Ref. [59] is summarised in the following.

The non-local hadronic functions $\mathcal{H}_\lambda(q^2)$ can be expressed in terms of a "conformal" variable z [66,67]

$$z(q^2) \equiv \frac{\sqrt{t_+ - q^2} - \sqrt{t_+ - t_0}}{\sqrt{t_+ - q^2} + \sqrt{t_+ - t_0}}\,, \tag{30}$$

where $t_+ = 4M_D^2$ and $t_0 = t_+ - \sqrt{t_+(t_+ - M_{\psi(2S)}^2)}$. With this choice, the functions $\mathcal{H}_\lambda(z)$ are meromorphic[4] in $|z| < 1$, with two simple poles at $z_{J/\psi(1S)} \simeq 0.18$ and $z_{\psi(2S)} \simeq -0.44$. Therefore, after dividing out these singularities

$$\mathcal{H}_\lambda(z) = \frac{1 - zz_{J/\psi(1S)}^*}{z - z_{J/\psi(1S)}} \frac{1 - zz_{\psi(2S)}^*}{z - z_{\psi(2S)}} \hat{\mathcal{H}}_\lambda(z), \tag{31}$$

the remaining functions $\hat{\mathcal{H}}_\lambda(z)$ are analytical in $|z| < 1$ and can be Taylor-expanded around $z = 0$. In addition, in order to improve the convergence of the expansion, these non-local functions are assumed to inherit all the singularities of the form factors, resulting in

$$\hat{\mathcal{H}}_\lambda(z) = \Big[\sum_{k=0}^{K} \alpha_k^{(\lambda)} z^k\Big] \mathcal{F}_\lambda(z)\,, \tag{32}$$

where $\alpha_k^{(\lambda)}$ are complex coefficients, and the analytical expansion is truncated after the term z^K. This parametrisation is expected to converge reasonably well in the region of interest, $|z| < 0.52$,[5] how-

---

[4] In complex analysis, a meromorphic function on an open subset D of the complex plane is a function that is analytic on all of D except for a discrete set of isolated points, which are poles of the function.

[5] The domain $|z| < 0.52$ includes q^2 values between -7 GeV^2 and $M_{\psi(2S)}^2$, as discussed in Sec. 3.2.2, where the functions $\mathcal{H}_\lambda$ will be extensively employed.

ever, the truncation of the series unavoidably introduces some model dependence.

2.2.2 *Observables in $B^0 \to K^{*0}\ell^+\ell^-$ decays*

Rare $B^0 \to K^{*0}\ell^+\ell^-$ decays are an excellent example of FCNC processes that are highly sensitive to physics beyond the SM. The presence of virtual particles, in fact, can significantly modify many of the numerous observables that can be studied analysing the decay.

2.2.2.1 *Branching ratio measurements*

One of the simplest observables of interest is the (differential) branching ratio, *i.e.* the rate of appearance of a given decay channel. However, the SM predictions for this kind of measurements suffer from large uncertainties, due to theoretical uncertainties on the form factors [68, 69]. These calculations involve non-perturbative QCD effects, that can not be exactly evaluated.

2.2.2.2 *Angular analyses*

The non-trivial angular distribution of the decay products determined by the differential decay rate given in Eq. 24 can be studied performing an angular analysis. The most natural set of observables can be built in terms of the *CP*-averaged and the *CP*-violating observables as

$$
\begin{aligned}
S_i &= \frac{I_i + \bar{I}_i}{(\mathrm{d}\Gamma + \mathrm{d}\bar{\Gamma})/\mathrm{d}q^2}, && (33) \\
A_i &= \frac{I_i - \bar{I}_i}{(\mathrm{d}\Gamma + \mathrm{d}\bar{\Gamma})/\mathrm{d}q^2}, && (34)
\end{aligned}
$$

where $\bar{I}_i$ and $\bar{\Gamma}$ refers to the *CP* conjugated decays.

Signal events can be analysed in bins of q^2, thereby effectively averaging the observables over the width of the bins. The resulting three-differential decay rate is given by

$$
\begin{aligned}
\frac{1}{\frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2}} \frac{\mathrm{d}(\Gamma+\bar{\Gamma})}{\mathrm{d}\vec{\Omega}} &= \frac{9}{32\pi}\Big[\frac{3}{4}(1-F_L)\sin^2\theta_K + F_L\cos^2\theta_K \qquad (35) \\
&+ \frac{1}{4}(1-F_L)\sin^2\theta_K\cos 2\theta_\ell - F_L\cos^2\theta_K\cos 2\theta_\ell \\
&+ S_3\sin^2\theta_K\sin^2\theta_\ell\cos 2\phi + S_4\sin 2\theta_K\sin 2\theta_\ell\cos\phi \\
&+ S_5\sin^2\theta_K\sin\theta_\ell\cos\phi + \frac{4}{3}A_{FB}\sin^2\theta_K\cos\theta_\ell \\
&+ S_7\sin^2\theta_K\sin\theta_\ell\sin\phi + S_4\sin 2\theta_K\sin 2\theta_\ell\sin\phi \\
&+ S_9\sin^2\theta_K\sin^2\theta_\ell\sin 2\phi\Big],
\end{aligned}
$$

where the mass of the leptons has been neglected, $F_L = (\frac{3}{4}I_{1c} - \frac{1}{4}I_{2c})/\Gamma \simeq S_{1c}$ (the last equality holds in the case of massless leptons when $S_{1c} = -S_{2c}$) is the longitudinal polarisation and $A_{FB} = 3/4\, S_{6s}$ is the lepton forward-backward asymmetry. The obtained differential decay rate can be used to extract the angular observables with a fit to data. Note that, the determination of the S_i observables, analogously to the branching ratio, suffers from sizeable uncertainties from form factors. Nevertheless, Refs. [70, 71] proposed an optimised set of observables where such uncertainties cancel to large extent. These are built from F_L and $S_{3,\ldots,9}$ and read as follows

$$\begin{aligned}
P_1 &= \frac{2\,S_3}{1-F_L} = A_T^{(2)}, \\
P_2 &= \frac{2}{3}\frac{A_{FB}}{1-F_L}, \\
P_3 &= \frac{-S_9}{1-F_L}, \\
P'_{4,5,8} &= \frac{S_{4,5,8}}{\sqrt{F_L(1-F_L)}}, \\
P'_6 &= \frac{S_7}{\sqrt{F_L(1-F_L)}}.
\end{aligned} \tag{36}$$

2.2.3 *Lepton Flavour Universality tests*

A complementary set of measurements that provides excellent probe of the Standard Model consists in *Lepton Flavour Universality* (LFU) tests. LFU is an intrinsic feature of the SM, originated by its flavour symmetry (*i.e.* apart from the Yukawa interaction, the Lagrangian of the SM is lepton-flavour universal). This implies that the electroweak couplings of the three generations of leptons are the same. An obvious consequence is that, once the difference of the lepton mass has been taken into account, amplitudes of processes involving different lepton flavours are predicted to be identical in the SM. This becomes of particular interest when considering rare B decays where the mass of the leptons $\ell = e, \mu$ can be neglected. Therefore, semileptonic decays involving $b \to s\ell\ell$ transitions are expected to behave identically when considering muon and electron final states. On the contrary, a non-trivial flavour structure beyond the SM can break Lepton Flavour Universality and be observed in the rich phenomenology of semileptonic decays, from branching ratios to angular observables.

The simplest example of such universality tests inspects branching ratios. It is common to define the ratio [72]

$$R_X = \frac{\int \frac{\mathrm{d}\Gamma(B\to X_s\mu^+\mu^-)}{\mathrm{d}q^2}\,\mathrm{d}q^2}{\int \frac{\mathrm{d}\Gamma(B\to X_s e^+e^-)}{\mathrm{d}q^2}\,\mathrm{d}q^2}, \tag{37}$$

where X_s stands for a hadron containing a *strange* quark, such as a K or a K^* meson, and the decay rate is integrated over a given q^2 range. The advantage of such formulation is that all the hadronic uncertainties (both form factors and non-local effects) cancel in the theoretical predictions, giving a Standard Model prediction close to unity with an accuracy at the percent level [73].

2.3 REVIEW OF THE FLAVOUR ANOMALIES

The term *flavour anomalies* refers to a set of measurements displaying tensions with the SM predictions that have been recently reported by several experiments in flavour physics. Despite each of these deviations has not exhibited singularly a discrepancy larger than $5\,\sigma$, these measurements globally show an intriguing anomalous pattern that attracted the attention of the entire community. These results can be grouped in two distinct areas: branching ratios and angular analyses of decays involving $b \to s\mu\mu$ transitions and tests of Lepton Flavour Universality in $b \to s\ell\ell$ decays.[6] In the following, a review of the current anomalous measurements in $b \to s\ell\ell$ decays is presented and their impact on the field is discussed.

2.3.1 *Branching ratios in $b \to s\mu\mu$ decays*

The LHCb experiment measured the branching ratios of several decay channels as $B^0 \to K^0\mu^+\mu^-$, $B^+ \to K^+\mu^+\mu^-$, $B^+ \to K^{*+}\mu^+\mu^-$, $B_s \to \phi\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ [75–77]. These channels are all governed by $b \to s\mu\mu$ transitions but differ by the choice of the spectator quark that gives origin to different mesons. In all these measurements the observed differential decay rate $\mathrm{d}\Gamma/\mathrm{d}q^2$ shows a deficit with respect to the Standard Model predictions, especially in the region $1 \lesssim q^2 \lesssim 8\ \mathrm{GeV}^2/c^4$. Figure 8 shows two of these examples. However, despite the revealed coherent behaviour, the interpretation of this set of measurements is limited by the form factor uncertainties in the Standard Model predictions.

2.3.2 *The P_5' anomaly*

The optimised set of angular observables defined in Eq. 36 has been widely used in the community as test of the Standard Model and, from an historical point of view, it is precisely in one of these observables - P_5' - that the first anomaly was observed in 2013 by the LHCb experiment. The first measurement of the optimised angular observables observed a local discrepancy in one of the q^2 bins of the P_5' variable

---

[6] Beside $b \to s\ell\ell$ transitions, a set of deviations from the SM has been measured also in $b \to c\ell\nu$ decays [74]. These processes are governed by tree-level charged current transitions and their discussion goes beyond the scope of this dissertation.

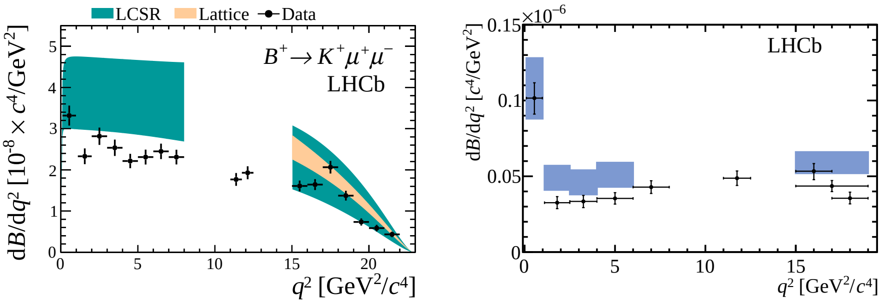


Figure 8: Differential branching ratio $d\Gamma/dq^2$ for (left) $B^+ \to K^+\mu^+\mu^-$ [75] and (right) $B^0 \to K^{*0}\mu^+\mu^-$ [77] decays. The shaded regions illustrate the theoretical predictions and their uncertainties.

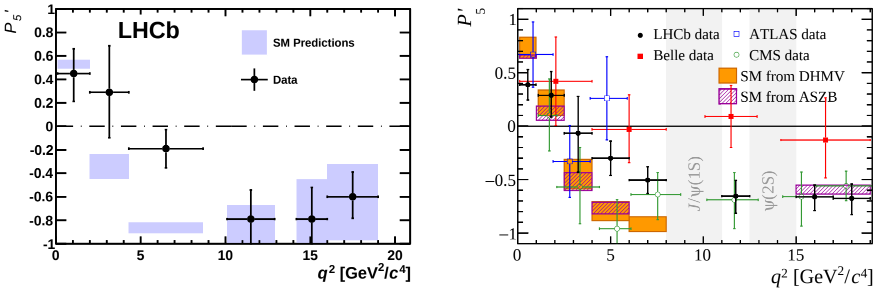


Figure 9: Left: first P_5' measurement by the LHCb experiment corresponding to a dataset of 1 fb^{-1} [78]. Right: Updated measurement of P_5' including LHCb 3 fb^{-1} [79], ATLAS [81], CMS [82] and Belle [80] results.

(see Fig. 9) [78]. Thereafter, several updates have been published in more recent years, both by LHCb (with three times more statistics) and by other experiments (Belle, ATLAS and CMS). Figure 9 (right) summarises the results obtained by the different experiments. While the updated publication of LHCb [3 fb^{-1}] [79] confirmed the initially-observed tension, analyses by Belle [80], ATLAS [81] and CMS [82] present large uncertainties that do not allow to draw any further conclusions. It is worth to note that, despite the limited sensitivity to this kind of measurements, the effort performed by general-purpose experiments like ATLAS and CMS proves the strong interest of the entire particle-physics community in the flavour anomalies.

2.3.3 *Charm loop or New Physics?*

Despite the long-standing P_5' deviation has been confirmed by two independent experiments, the physics community sees this anomaly with skepticism. In fact, if on one hand several models have been suggested in order to interpret this tension as a signature of New

Physics [41–47], on the other hand, the vector-like nature of this pattern could be also explained by non-perturbative QCD contributions from $\mathcal{O}_{bscc}$ operators (known as *charm loop*) that are able to either mimic or camouflage New Physics effects [83,84]. Presently, the largest systematic uncertainty that limits the discovery potential of $B^0 \to K^{*0}\mu^+\mu^-$ decays arises from the insufficient knowledge of these non-local hadronic matrix elements.

Some promising approaches suggest to extract these contributions from data-driven analyses [85,86] or by exploiting analytical properties of its structure [59], as discussed in Sec. 2.2.1.1. However, these models still have intrinsic limitations, in particular in the assumptions that enter in the parametrisation of the di-lepton invariant mass distribution. Chapter 3 investigates in details the potentialities and limitations of the analytical approach of Ref. [59], with the awareness that this is a crucial issue for the understanding of $B^0 \to K^{*0}\mu^+\mu^-$ anomalies.

2.3.4 *The R_K and R_{K^*} anomalies*

The LHCb experiment recently published two measurements, R_K [87] and R_{K^*} [88], of LFU tests involving $b \to s\ell^+\ell^-$ transitions. These measurements analyse $B^+ \to K^+\ell\ell$ and $B^0 \to K^{*0}\ell^+\ell^-$ decays, respectively. The observed values are

$$R_K = 0.745^{+0.090}_{-0.074} \pm 0.036\,, \tag{38}$$

measured in the range $1 < q^2 < 6\ \text{GeV}^2/c^4$, and

$$R_{K^*} = \begin{cases} 0.66^{+0.11}_{-0.07} \pm 0.03 & \text{for } 0.045 < q^2 < 1.1\ \text{GeV}^2/c^4\,, \\ 0.69^{+0.11}_{-0.07} \pm 0.05 & \text{for } 1.1 \quad\ < q^2 < 6.0\ \text{GeV}^2/c^4\,, \end{cases} \tag{39}$$

where the analysis has been performed in two distinct q^2 bins, namely low-q^2 and central-q^2 region. For all results, the first uncertainty is statistical and the second systematic.

These measurements are in tension with the Standard Model predictions at a level of 2.6 standard deviations (R_K) and 2.1 and 2.4 standard deviations (R_{K^*}) for the low- and central-q^2 bins respectively. Nevertheless, the interpretation of the observed value in the low-q^2 bin of the R_{K^*} measurement requires a separate discussion. In fact, the low-q^2 region of $B^0 \to K^{*0}\ell^+\ell^-$ decays is dominated by the $\mathcal{O}_7$ operator and analyses of $B \to K^{*0}\gamma$ decays [89–92] provide strong constraints on a SM-like value of $\mathcal{C}_7$. Hence, the small value of R_{K^*} at low-q^2 deviates both from the SM and from the most popular NP scenarios and only the analysis of larger datasets will elucidate such regime.

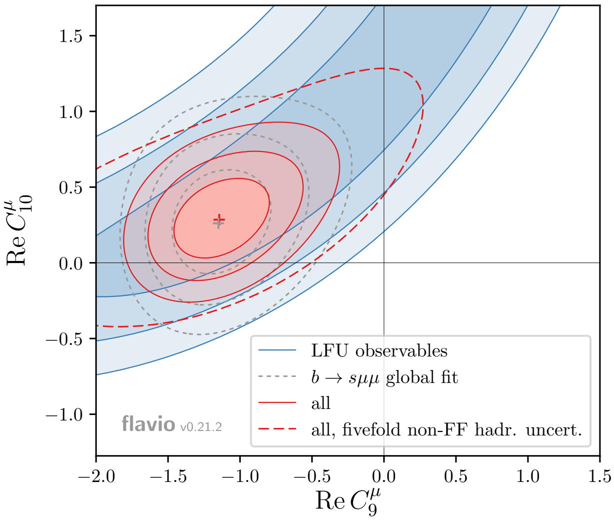


Figure 10: Global fit to $b \to s\ell^+\ell^-$ anomalies from Ref. [93].

2.4 GLOBAL FITS

Despite none of the current measurements exceeds the $5\,\sigma$ deviation required for a claim of New Physics, global fits seem to suggest a common explanation to these anomalies [93–95]. The most popular interpretations point to a shift in $\mathcal{C}_9$ or, alternatively, in $\mathcal{C}_9$ and $\mathcal{C}_{10}$ simultaneously.[7] Fig. 10 shows one example of global fits when varying $\mathcal{C}_9$ and $\mathcal{C}_{10}$ while keeping all the other Wilson coefficients to their Standard Model values. The results are also displayed when considering $b \to s\mu\mu$ measurements only, or LFU tests alone, resulting in a great compatibility of the favoured $\mathcal{C}_9$-$\mathcal{C}_{10}$ parameters space between the two sets of measurements. Furthermore, this set of anomalies suggests the presence of New Physics in the muon channel (*e.g.* lower branching ratios, modified angular distributions) while keeping the electron mode as in the Standard Model.

---

[7] In the hypothesis of a simultaneous NP contribution to both $\mathcal{C}_9$ and $\mathcal{C}_{10}$, an appealing scenario is provided by the case $\mathcal{C}_9^{\mathrm{NP}} = -\mathcal{C}_{10}^{\mathrm{NP}}$, which preserves the $V - A$ structure of the leptonic current of the Standard Model.

> *"[..] Considerate la vostra semenza:*
> *fatti non foste a viver come bruti,*
> *ma per seguir virtute e canoscenza.*
>
> *Li miei compagni fec'io sì aguti,*
> *con questa orazion picciola, al cammino,*
> *che a pena poscia li avrei ritenuti;*
>
> *e volta nostra poppa nel mattino,*
> *de' remi facemmo ali al folle volo,*
> *sempre acquistando dal lato mancino."*
>
> — Dante (XIV sec.) Inferno, Canto XXVI

3

TOWARDS A DISCOVERY OF NEW PHYSICS?

This chapter includes original work done during my PhD. Firstly, phenomenological studies are performed to understand the New Physics discovery potential of $B^0 \to K^{*0}\mu^+\mu^-$ decays, proposing a new method to advance the understanding of long-standing issues that so far limited the confidence of the flavour community in the anomalies observed in this decay [2]. Secondly, new observables that can lead to an early discovery of New Physics in Lepton Flavour Universality test of $B^0 \to K^{*0}\ell^+\ell^-$ decays (if the current hints are confirmed) are proposed [3].

3.1 PRELIMINARIES

In experiments, rare semileptonic decays are in general accessed only outside the regions corresponding to the $c\bar{c}$ resonant modes. In particular, the decays $B^0 \to J/\psi(1S)(\to \ell^+\ell^-)K^{*0}$ and $B^0 \to \psi(2S)(\to \ell^+\ell^-)K^{*0}$, which proceed at the tree level in the Standard Model, completely prevail over the $B^0 \to K^{*0}\ell^+\ell^-$ rare mode due to their branching ratio larger by three orders of magnitude.

In addition, bremsstrahlung effects limit the electron q^2 resolution at LHCb, resulting in a strong deterioration of the $J/\psi(1S)$ and $\psi(2S)$ resonant peaks. As a consequence, the following kinematic regions are investigated in these studies

$$\begin{aligned} Q_1^{(\mu)}: &\quad 1.1 \leq q^2 \leq 8.0 \text{ GeV}^2/c^4\,, \\ Q_2^{(\mu)}: &\quad 11.0 \leq q^2 \leq 12.5 \text{ GeV}^2/c^4\,, \\ Q^{(e)}: &\quad 1.1 \leq q^2 \leq 7.0 \text{ GeV}^2/c^4\,, \end{aligned} \tag{40}$$

Table 1: Number of expected signal candidates extrapolated to the different scenarios.

Number of events	Run-II	Upgrade [50 fb^{-1}]
$B^0 \to K^{*0}\mu^+\mu^-$	4,900	44,000
$B^0 \to K^{*0}\mu^+\mu^-$ $[Q^{(\mu)}_{\text{ext}}]$	6,900	62,000
$B^0 \to K^{*0}e^+e^-$	525	4,700

where two q^2 intervals are considered for the muon channel while the electron mode is limited to a single domain. The definition of these q^2 ranges is broadly consistent with published results and assumes improvements in the electron mode resolution [96]. Furthermore, a modified definition for the q^2 muon ranges is considered, namely

$$\begin{aligned} Q^{(\mu)}_{1,\text{ext}}: &\quad 1.1 \leq q^2 \leq 9.0 \text{ GeV}^2/c^4\,, \\ Q^{(\mu)}_{2,\text{ext}}: &\quad 10.0 \leq q^2 \leq 13.0 \text{ GeV}^2/c^4\,. \end{aligned} \tag{41}$$

These *extended* domains assume a veto of approximately $\pm\,0.5$ GeV^2 around the $J/\psi(1S)$ and $\psi(2S)$ invariant mass squared and may appear experimentally challenging. However, considering that the effect of the investigated non-local hadronic contributions is expected to rise in proximity of the charmonium resonances, the analysis of these extended q^2 regions is of high interest for this study.

Finally, the considered q^2 range includes neither the low-q^2 region, where several light resonances (*e.g.* ϕ, ρ, ω, etc.) must be considered,[1] nor the high-q^2 region, where the presence of broad charmonium resonances and new opening channels invalidates the current parametrisation of the non-local hadronic contributions.

Based on the selected q^2 intervals, the expected signal yields foreseen at LHCb Run-II [8 fb^{-1}] and Upgrade [50 fb^{-1}] [97] can be predicted. These are extrapolated from Refs. [79,88] by scaling the luminosities and the $b\bar{b}$ production cross section $\sigma_{b\bar{b}} \propto \sqrt{s}$, where $\sqrt{s}$ denotes the designed centre-of-mass energy of the b-quark pair. The expected number of signal candidates for the different scenarios is reported in Table 1.

The *Probability Density Function,*[2] *pdf*, for signal decays can be written as

$$pdf \equiv \frac{1}{\Gamma}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\,\mathrm{d}^3\Omega}\,, \qquad \text{with} \quad \Gamma \equiv \int_{q^2\in Q_i} \mathrm{d}q^2\,\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2}\,, \tag{42}$$

---

[1] An appropriate description of light resonances goes beyond the scope of this study.

[2] The Probability Density Function of a continuous variables is the normalised representation of its distribution, namely a function whose value at any given point returns the probability of the variable to equal that value (or to fall within a particular range of values, if integrated over that interval).

where $\mathrm{d}^4\Gamma$ is the differential decay rate of Eq. 24 and the q^2 domain Q_i depends on the lepton flavour under study. The signal *pdfs* contain the following list of parameters

i. the Wilson coefficients $\mathcal{C}_7^{(\ell)}$, $\mathcal{C}_9^{(\ell)}$ and $\mathcal{C}_{10}^{(\ell)}$ and, optionally, the right-handed counterparts $\mathcal{C}'^{(\ell)}_{7,9,10}$, where $\ell = e, \mu$ denotes the lepton flavour;[3]

ii. the CKM Wolfenstein parameters $\{\alpha_m^{(\mathrm{CKM})}\} \equiv \{\lambda, A, \bar{\rho}, \bar{\eta}\}$;

iii. the form factor parameters $\{\alpha_l^{(F)}\}$ for form factors $F = V, A_{0,\dots,2}, T_{1,\dots,3}$;

iv. the non-local hadronic parameters (real and imaginary parts) $\{\alpha_k^{(\lambda)}\}$ for each polarisation $\lambda = \perp, \|, 0$ and order of the analytical expansion z^k;

where the parameters of interest are the Wilson coefficients and all the remaining ones are treated as nuisance parameters.

In the following, several sensitivity studies are performed to investigate the sensitivity to New Physics in $B^0 \to K^{*0}\ell^+\ell^-$ decays. Ensembles of 1000 pseudoexperiments are generated based on the signal *pdf* of Eq. 42, where the values of the nuisance parameters are set as in the state-of-the-art literature[4]

i. $\{\alpha_m^{(\mathrm{CKM})}\}$, as obtained from a tree-level analysis of the unitarity triangle [99];

ii. $\{\alpha_l^{(F)}\}$, as determined from a combined fit to Light-Cone Sum Rules and lattice QCD results in Ref. [100];

iii. $\{\alpha_k^{(\lambda)}\}$, as in Ref. [59], with $\alpha_k^{(\lambda)} = 0$ for $k > 2$, unless stated otherwise.

Note that the evaluation of the non-local contributions proposed in Ref. [59] consists of a *genuine* Standard Model prediction since it does not make use of any information from $B^0 \to K^{*0}\mu^+\mu^-$ decays but rather theoretical calculations at $q^2 < 0$, where the current approaches for the large recoil region based on QCD-factorisation are reliable, and experimental data of $B^0 \to K^{*0}J/\psi(1S)$ and $B^0 \to K^{*0}\psi(2S)$ decays, which are independent on hypothetical New Physics in semileptonic operators. However, these predictions are affected by a non-quantifiable model dependence due to the truncation of the polynomial expansion at the order $K = 2$ assumed in Ref. [59].

---

[3] Within the Standard Model, Lepton Flavour Universality presupposes $\mathcal{C}_i^{(e)} = \mathcal{C}_i^{(\mu)}$, nonetheless, eventual New Physics contributions that violate LFU implies non equal $\mathcal{C}_i^{(\ell)}$ for the different lepton flavours.

[4] Some precautions must be taken into account when generating pseudoexperiments when external constraints are applied to some of the parameters. A detailed description of the correct procedure can be found in Ref. [98].

Finally, regarding the Wilson coefficients, different scenarios are considered, namely the Standard Model setup and two New Physics benchmark points (BMP),

i. `SM` - the Wilson coefficients are set to their SM values, namely $\mathcal{C}_9^{\mathrm{SM}} = 4.27$, $\mathcal{C}_{10}^{\mathrm{SM}} = -4.17$ and $\mathcal{C}_7^{\mathrm{SM}} = -0.34$ [101];[5]

ii. $\texttt{BMP}_{\mathcal{C}_9}$ - NP is inserted only in $\mathcal{C}_9^{(\mu)}$, with a shift of $\mathcal{C}_9^{\mathrm{NP}} = -1$;

iii. $\texttt{BMP}_{\mathcal{C}_{9,10}}$ - NP is inserted in both $\mathcal{C}_9^{(\mu)}$ and $\mathcal{C}_{10}^{(\mu)}$, with a simultaneous shift of $\mathcal{C}_9^{\mathrm{NP}} = -\mathcal{C}_{10}^{\mathrm{NP}} = -0.7$;

where for the two NP hypotheses, New Physics is inserted only in the muon channel, *i.e.* $\mathcal{C}_i^{(e)} = \mathcal{C}_i^{\mathrm{SM}}$. These two modified scenarios are favoured by several global fit analyses with similar significance [93–95].

An extended unbinned maximum likelihood fit [102] is then performed to these simulated samples, where the notation $\mathcal{H}_\lambda[z^K]$ refers to fits performed with non-local hadronic functions truncated at the order z^K. In general, in the case that prior knowledge on the true value of one or more fit parameters is available, this information can be incorporated into the fit procedure by multiplying the total likelihood by a (multi-)Gaussian function based on their central values, uncertainties and, eventually, correlation matrices. These Gaussian constraints are applied to $\{\alpha_m^{(\mathrm{CKM})}\}$, $\{\alpha_l^{(F)}\}$ and, optionally,[6] $\{\alpha_k^{(\lambda)}\}$, based on the above-mentioned references.[7] The total likelihood reads as

$$\mathcal{L}_{\mathrm{tot}} \equiv \mathcal{L}_{pdf} \times \mathcal{L}_{\mathcal{G}} \times \mathcal{L}_{\mathcal{B}}\,, \tag{43}$$

where $\mathcal{L}_{pdf} = \prod \mathcal{L}_{pdf_i}$ are unbinned likelihoods of N_i pseudoevents $x_{n,i} \sim pdf_i$,

$$\mathcal{L}_{pdf_i} = \prod_n^{N_i} pdf_i(x_{n,i})\,, \tag{44}$$

where the subscript i refers to the different q^2 regions and/or lepton decay channels, while $\mathcal{L}_{\mathcal{G}}$ includes the Gaussian constraints on the nuisance parameters and $\mathcal{L}_{\mathcal{B}} = \prod \mathcal{L}_{\mathcal{B}_i}$ incorporates Poissonian terms that take into account the integrated branching ratios of the decay in the kinematic regions of interest. These terms take into account the observed signal rate in the given regions of interest and provide complementary information with respect to the normalised four-dimensional *pdfs* of Eq. 42 and are expressed by

$$\mathcal{L}_{\mathcal{B}_i} = (n_{sig_i})^{N_i} \times e^{-n_{sig_i}} \tag{45}$$

---

[5] These values correspond to a fixed renormalisation scale of $\mu = 4.2\,\mathrm{GeV}/c^2$.

[6] Unless stated otherwise, the nominal configuration of the fit has no constraints on the non-local hadronic parameters $\{\alpha_k^{(\lambda)}\}$.

[7] In order to guarantee a good agreement between Light-Cone Sum Rules [62, 63] and Lattice results [60, 61], uncertainties on the form factors parameters $\{\alpha_l^{(F)}\}$ are doubled with respect to Ref. [100].

where the number of signal candidates is directly related to the branching ratio of the rare mode in a given kinematic region and is expressed as function of the nuisance parameters and Wilson coefficients

$$n_{sig_i} \propto \mathcal{B}_{K^{*0}\mu\mu,i} = \frac{\tau_B}{\hbar} \int_{q^2 \in Q_i} \frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} \mathrm{d}q^2 \,. \tag{46}$$

For the purpose of sensitivity studies, this is assumed to be perfectly known, while, in the case of a measurement, experimental uncertainties when extracting the branching ratio must be considered. A detailed discussion of all the experimental effects that can affect the amplitude fit to $B^0 \to K^{*0}\mu^+\mu^-$ decays is ignored at this stage and is postponed to chapter 7. This includes the uncertainty on the above-mentioned branching ratio measurement, detector acceptance/efficiency, background and S-wave contaminations, while a preliminary study of the finite q^2 resolution in $B^0 \to K^{*0}e^+e^-$ decays is discussed in these sensitivity studies, due to its major role in measurements involving electrons in the final state.

The stability of the model and the convergence to the global minimum is enforced by repeating the fit with randomised starting parameter values; the solution with smallest negative log-likelihood is taken as the default. In the following, all the quoted uncertainty intervals for the various parameters are determined in a frequentist way and correspond to 68% confidence level intervals from profile likelihoods.

In conclusion, this framework parametrises the dynamics of $B^0 \to K^{*0}\ell^+\ell^-$ in terms of an amplitude fit that includes angular information, branching ratio measurements, as well as the state-of-the-art knowledge on the nuisance parameters. In addition, due to its unbinned nature, this approach is expected to overcome the sensitivity of previous methods, as will be shown in Sec. 3.2.5.

3.2 DIAGNOSING NEW PHYSICS IN $B^0 \to K^{*0}\mu^+\mu^-$ DECAYS

This section investigates the prospects to disentangle problematic non-local hadronic effects described in Sec. 2.3.3 from possible NP contribution in $B^0 \to K^{*0}\mu^+\mu^-$ decays. Despite the fact that both the Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ can be significantly modified by the presence of New Physics, $\mathcal{C}_{10}$ is expected to be insensitive to the lack of knowledge on the non-local contributions [103], by virtue of the axial structure of $\mathcal{O}_{10A}$, which does not mix with any of the four-quark operators of Eq.22.

Therefore, primary attention is directed to the Wilson coefficient $\mathcal{C}_9$, *i.e.* the default configuration of the amplitude fit presented in this section assumes $\mathcal{C}_{10}$ to be fixed to its Standard Model value. Finally, $\mathcal{C}_7$ is always fixed to its SM value since it is strongly constrained by radiative decays and is unaffected by the most common NP models.

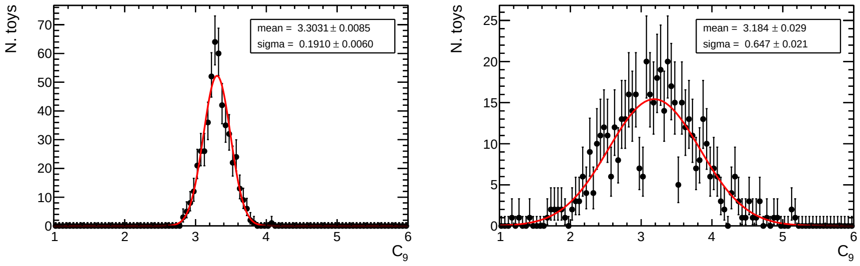


Figure 11: Fit results for the Wilson coefficient $\mathcal{C}_9$ for the expected statistics at the LHCb Run-II. Fits are performed with (left) $\mathcal{H}_\lambda[z^2]$ and (right) $\mathcal{H}_\lambda[z^3]$.

In the following, $B^0 \to K^{*0}\mu^+\mu^-$ decays are investigated in two kinematic regions $Q^{(\mu)}_{1,\,\text{ext}}$ and $Q^{(\mu)}_{2,\,\text{ext}}$ as defined in Eq. 40.

3.2.1 *On the truncation of the $\mathcal{H}_\lambda(z)$ series*

One of the main intents of this study is to inquire the model dependence introduced by the truncation of the polynomial expansion $\mathcal{H}_\lambda(z)$. This is explored by generating ensembles of pseudoexperiments corresponding to the $\texttt{BMP}_{\mathcal{C}_9}$ scenario and the expected LHCb Run-II statistics and fitting the simulated samples with $\mathcal{H}_\lambda[z^2]$ and $\mathcal{H}_\lambda[z^3]$. Figure 11 shows the fit results obtained for $\mathcal{C}_9$ for the two different model assumptions. The uncertainty on $\mathcal{C}_9$ more than triplicate between fits with $\mathcal{H}_\lambda[z^2]$ and $\mathcal{H}_\lambda[z^3]$. Hence, the model-dependence of the fit is large in the absence of any theory constraints on the parameters $\{\alpha_k^{(\lambda)}\}$.

In conclusion, a model-independent measurement can not be achieved purely analysing experimental data of semileptonic $B^0 \to K^{*0}\mu^+\mu^-$ decays, but requires additional information on the nonlocal hadronic functions. To overcome this problem, a novel combined theory-experimental analysis is pursued by extending the amplitude fit to include all the available (theoretical and experimental) information on the process.

3.2.2 *A novel theory-experimental approach*

This new combined approach merges the SM predictions of the nonlocal hadronic contributions of Ref. [59] and the amplitude fit to $B^0 \to K^{*0}\mu^+\mu^-$ decays into a single-step analysis. The previously described framework is then modified to include

i. theoretical predictions at negative q^2 for the ratios $\mathcal{H}_\lambda/\mathcal{F}_\lambda$, the values calculated in Ref. [59] for four points at $q^2 \in [-1,-3,-5,-7]\ \text{GeV}^2/c^4$ are reported in Table 2;

Table 2: Mean values μ_i (in units of 10^{-4}), and standard deviations σ_i (in units of 10^{-4}) of the theory constraints at negative q^2 (in units of GeV^2/c^4). Values are taken from Ref. [59].

	$\mathcal{R}e[\mathcal{H}_\perp]/\mathcal{F}_\perp$				$\mathcal{R}e[\mathcal{H}_\parallel]/\mathcal{F}_\parallel$				$\mathcal{R}e[\mathcal{H}_0]/\mathcal{F}_0$			
q^2	−7.0	−5.0	−3.0	−1.0	−7.0	−5.0	−3.0	−1.0	−7.0	−5.0	−3.0	−1.0
μ	6.656	4.878	4.076	3.750	6.033	4.384	3.728	3.586	−1.997	1.596	1.818	0.768
σ	2.553	1.048	0.621	0.561	2.446	0.971	0.575	0.538	4.077	1.368	0.472	0.125
	$\mathcal{I}m[\mathcal{H}_\perp]/\mathcal{F}_\perp$				$\mathcal{I}m[\mathcal{H}_\parallel]/\mathcal{F}_\parallel$				$\mathcal{I}m[\mathcal{H}_0]/\mathcal{F}_0$			
q^2	−7.0	−5.0	−3.0	−1.0	−7.0	−5.0	−3.0	−1.0	−7.0	−5.0	−3.0	-1.0
μ	1.581	1.294	1.291	1.380	1.517	1.246	1.257	1.366	6.328	1.970	0.583	0.136
σ	0.835	0.610	0.565	0.585	0.803	0.590	0.553	0.581	7.411	2.107	0.528	0.082

Table 3: Pseudo-observables from $B^0 \to K^{*0}\psi_n$ decays as obtained in Ref. [59].

Pseudo-observable	$J/\psi(1S)$	$\psi(2S)$
$\lvert r_\perp \rvert$	$(2.027 \pm 0.190) \cdot 10^{-3}$	$(1.06 \pm 0.21) \cdot 10^{-3}$
$\lvert r_\parallel \rvert$	$(1.713 \pm 0.260) \cdot 10^{-3}$	$(0.98 \pm 0.18) \cdot 10^{-3}$
$\lvert r_0 \rvert$	$(2.303 \pm 0.357) \cdot 10^{-3}$	$(1.40 \pm 0.36) \cdot 10^{-3}$
$\arg\{r_\perp r_0^*\}$	$+2.926 \pm 0.032$	$+2.799 \pm 0.314$
$\arg\{r_\parallel \vert r_0^*\}$	-2.944 ± 0.036	-2.815 ± 0.403

ii. experimental measurements of the hadronic decays $B^0 \to K^{*0}J/\psi(1S)$ and $B^0 \to K^{*0}\psi(2S)$. For each of these decays, a set of five pseudo-observables can be obtained from the corresponding amplitudes, these consist of three magnitudes and two relative phases, namely

$$|r_\perp^{\psi_n}|,\ |r_\parallel^{\psi_n}|,\ |r_0^{\psi_n}|,\ \arg\{r_\perp^{\psi_n} r_0^{\psi_n *}\},\ \arg\{r_\parallel^{\psi_n} r_0^{\psi_n *}\}, \tag{47}$$

where

$$r_\lambda^{\psi_n} \equiv \operatorname*{Res}_{q^2 \to M_{\psi_n}^2} \frac{\mathcal{H}_\lambda(q^2)}{\mathcal{F}_\lambda(q^2)}\,. \tag{48}$$

Constraints on these pseudo-observables are obtained in Ref. [59] from branching ratio measurements and angular analyses by Babar [104, 105], Belle [106–108] and LHCb [109] and are summarised in Table 3.

iii. The unbinned amplitude fit to $B^0 \to K^{*0}\mu^+\mu^-$ decays described in the previous section.

The conditions [i.] and [ii.] are included in the fit as multivariate Gaussian constraints on the relevant pseudo-observables.

To ensure the consistency of the model, these additional constraints are tuned accordingly to the values of the non-local hadronic func-

Table 4: Expected central value and uncertainties for the $\mathcal{C}_9$ Wilson coefficient obtained from fits with $\mathcal{H}_\lambda[z^2, ..., z^5]$ for the $\texttt{BMP}_{\mathcal{C}_9}$scenario.

	LHCb Run-II		LHCb Upgrade [50 fb^{-1}]	
	$\mathcal{C}_9^{\text{NP}}$ mean	$\mathcal{C}_9^{\text{NP}}$ sigma	$\mathcal{C}_9^{\text{NP}}$ mean	$\mathcal{C}_9^{\text{NP}}$ sigma
$\mathcal{H}_\lambda[z^2]$	-0.965 ± 0.004	0.120 ± 0.003	-0.996 ± 0.002	0.060 ± 0.001
$\mathcal{H}_\lambda[z^3]$	-0.991 ± 0.007	0.220 ± 0.005	-1.015 ± 0.004	0.124 ± 0.003
$\mathcal{H}_\lambda[z^4]$	-1.028 ± 0.008	0.235 ± 0.005	-1.012 ± 0.005	0.141 ± 0.003
$\mathcal{H}_\lambda[z^5]$	-0.945 ± 0.010	0.294 ± 0.007	-0.977 ± 0.005	0.162 ± 0.004

tions assumed in the production of the ensembles. The corresponding central values are shifted to match the predictions of each pseudo-observables given a certain set of parameters $\{\alpha_k^{(\lambda)}\}$, while their uncertainties are scaled to keep the relative error constant.

Within this setup, the study on the model-bias introduced by the truncation of the series is repeated. Table 4 reports the obtained sensitivity to $\mathcal{C}_9^{\text{NP}}$ for fits with $\mathcal{H}_\lambda[z^2, ..., z^5]$ for datasets corresponding to the LHCb Run-II and LHCb Upgrade expected statistics. From these results the following conclusions can be inferred

i. the uncertainty on $\mathcal{C}_9^{\text{NP}}$ roughly doubles moving from fits with $\mathcal{H}_\lambda[z^2]$ to $\mathcal{H}_\lambda[z^3]$, for both statistics under consideration;

ii. for the dataset corresponding to the expected statistics at the LHCb Upgrade, the uncertainty on $\mathcal{C}_9^{\text{NP}}$ only slightly increases for orders higher than $\mathcal{H}_\lambda[z^3]$;

iii. for the dataset corresponding to the expected statistics at the LHCb Run-II the saturation of the uncertainty is confined to the orders $\mathcal{H}_\lambda[z^3]$ and $\mathcal{H}_\lambda[z^4]$. In fact, for fits with $\mathcal{H}_\lambda[z^5]$ the uncertainty starts to increase again, pointing to a statistical limitation.

The observed distributions of the non-local hadronic functions $\mathcal{H}_\lambda(q^2)/\mathcal{F}_\lambda(q^2)$ for different orders in $\mathcal{H}_\lambda[z^k]$ are shown in Fig. 12, corresponding to the statistics expected at the LHCb Upgrade. In all cases, the uncertainty drastically increases for higher orders in the $\psi(2S)$ window. However, in the regions of interest of $B^0 \to K^{*0}\mu^+\mu^-$ decays, $Q^{(\mu)}_{1,\text{ext}}$ and $Q^{(\mu)}_{2,\text{ext}}$, the uncertainty of the real part of the hadronic functions reflects the behaviour observed for $\mathcal{C}_9^{\text{NP}}$, *i.e.* the uncertainty mildly increases for orders higher than $\mathcal{H}_\lambda[z^3]$.

3.2.2.1 *Residual model dependence*

This section investigates in details several aspects of the problematics concerning the truncation of the expansion of the non-local hadronic functions $\mathcal{H}_\lambda(z)$. The studies presented so far assumed a fixed set of values for the parameters $\{\alpha_{k\leq 2}^{(\lambda)}\}$ based on Ref. [59], while in the

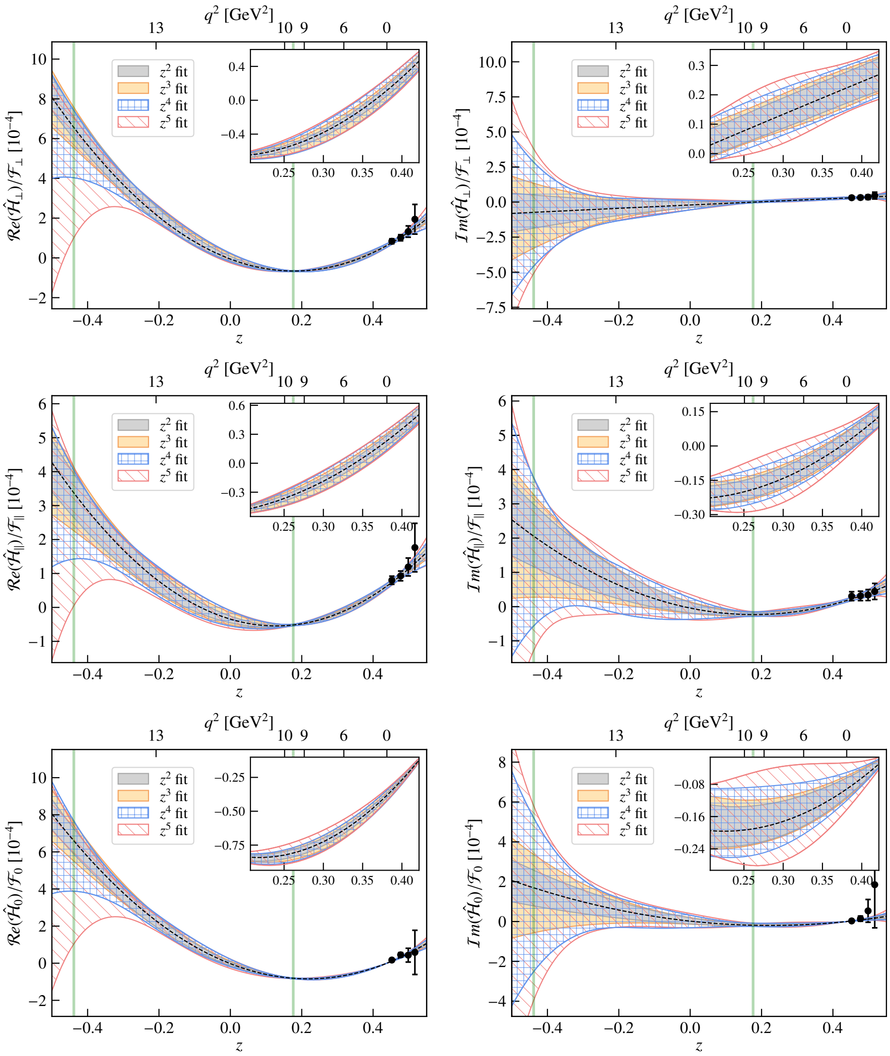


Figure 12: Results of the fits for the ratio $\hat{\mathcal{H}}_\lambda(z)/\mathcal{F}_\lambda(z)$ obtained with different orders of the expansion $\mathcal{H}_\lambda[z^k]$ for the $\mathtt{BMP}_{\mathcal{C}_9}$scenario and the expected statistics at LHCb Upgrade. The vertical bands correspond to the $J/\psi(1S)$ and $\psi(2S)$ regions and the points to the theoretical inputs at negative q^2. The top right box of each plot zooms in the q^2 range between 1.1 and 9.0 GeV$^2/c^4$.

following different possible choices for the hadronic parameters - including higher order of the expansion - are considered. Since *a-priori* there is no physics argument to assume which is the correct order to describe Nature, this is a crucial aspect in the determination of the model dependence of the proposed approach. In particular, the cases with non-zero coefficients for order of the expansion up to z^3 (*i.e.* $\alpha_{i\geq 4}^{(\lambda)} = 0$) and z^4 (*i.e.* $\alpha_{i\geq 5}^{(\lambda)} = 0$), labelled as $\mathcal{H}_\lambda\{\alpha_3\}$ and $\mathcal{H}_\lambda\{\alpha_4\}$ respectively, are examined.

The choice of the above-mentioned non-zero coefficients is based on the following criteria

i. compatibility with the theory predictions at negative q^2 and the pseudo-observables from the hadronic decays;

ii. attenuation of the tension with the P_5' anomaly, *i.e.* hadronic effects mimic the behaviour of NP.

The hadronic functions obtained from the resulting set of parameters are displayed in Fig. 13, while the corresponding values of the P_5' angular observable within the `SM` and `BMP`$_{\mathcal{C}_9}$scenarios are shown in Fig. 14. It is interesting to note that the predicted values of the non-local hadronic contributions of Ref. [59] shift the value of the P_5' observable downward, increasing the tension of the LHCb measurement with the SM hypothesis. However, the values obtained when allowing higher order of the expansion can produce the opposite effect and represent equally valid solutions. This fact emphasises once more the paramount importance of a complete understanding of the order of the series to be curtailed at to unambiguously discriminate between non-local hadronic effects and New Physics contribution.

A series of simulated ensembles produced with either $\mathcal{H}_\lambda\{\alpha_3\}$ or $\mathcal{H}_\lambda\{\alpha_4\}$ is fitted under different model assumptions, $\mathcal{H}_\lambda[z^2, ..., z^5]$. Table 5 shows the fit results for the $\mathcal{H}_\lambda\{\alpha_3\}$ configuration, similar results are obtained for the study of $\mathcal{H}_\lambda\{\alpha_4\}$. From these, the following conclusions can be deduced

i. fitting with $\mathcal{H}_\lambda[z^2]$ (*i.e.* with a lower order than what is used for the production of the ensembles) introduces a strong bias in the estimator for $\mathcal{C}_9$;

ii. when the order of the truncation in the fitting procedure reaches the one used for the production of the ensembles, the estimator for $\mathcal{C}_9$ is approximately unbiased (the extent of the validity of this statement is discussed below);

iii. the uncertainty on $\mathcal{C}_9$ varying the order of the fit follows the pattern observed in section 3.2.2, *i.e.* it strongly increases between $\mathcal{H}_\lambda[z^2]$ and $\mathcal{H}_\lambda[z^3]$ and smoothly grows for higher orders;

iv. the determination of the a-priori unknown description of non-local hadronic effects in Nature can be investigated by scanning the order of the truncation of the series until the central value of $\mathcal{C}_9$ stabilises.

It is worth to further elaborate on the sensitivity of $\mathcal{C}_9$ for the expected LHCb Run-II statistics. Figure 15 displays the central value and the statistical uncertainty of the estimator of $\mathcal{C}_9$, as obtained from fits with different orders $\mathcal{H}_\lambda[z^2, ..., z^5]$ performed on ensembles simulated with distinct conditions, $\mathcal{H}_\lambda\{\alpha_3\}$, $\mathcal{H}_\lambda\{\alpha_4\}$ as well as the default configuration, namely $\mathcal{H}_\lambda\{\alpha_2\}$. A bias in the estimation of $\mathcal{C}_9$ is noticeable, nevertheless the entity of such effect is mitigated for

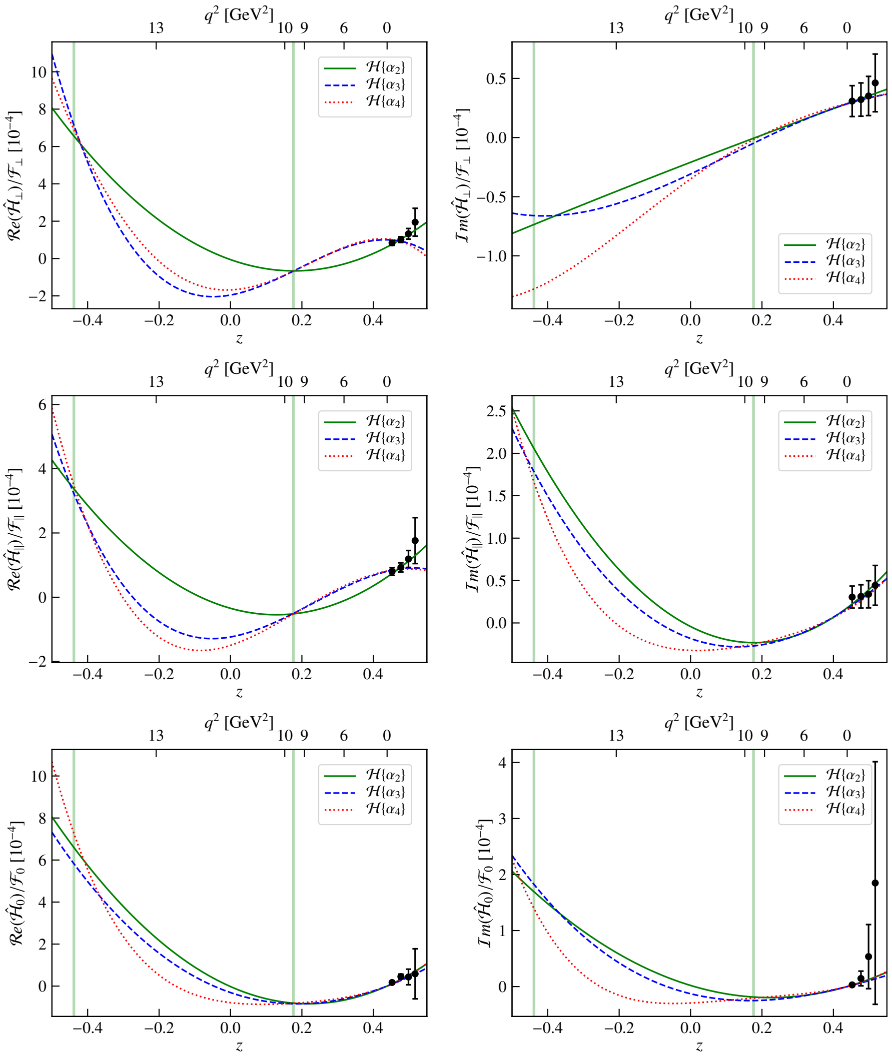


Figure 13: $\hat{\mathcal{H}}_\lambda(z)/\mathcal{F}_\lambda(z)$ corresponding to the set of parameter $\{\alpha_k^{(\lambda)}\}$ used for the production of the ensembles in the different hypotheses as described in the text. The vertical bands correspond to the $J/\psi(1S)$ and $\psi(2S)$ regions and the points to the theoretical inputs at negative q^2.

the expected statistics of the LHCb Upgrade. The study presented in appendix A.1.2 proves the statistical origin of this phenomenon analysing the results of the estimator of $\mathcal{C}_9$ as function of the number of events. However, the observed bias presents a non-trivial pattern as function of the increasing order of the $\mathcal{H}_\lambda(z)$ expansion in the fit and, therefore, requires to be accounted in any experimental measurements.

3.2.3 *Simultaneous fit to $\mathcal{C}_9$ and $\mathcal{C}_{10}$*

In the previous sections the attention was entirely pointed to establish whether is possible to disentangle non-local hadronic effects from New

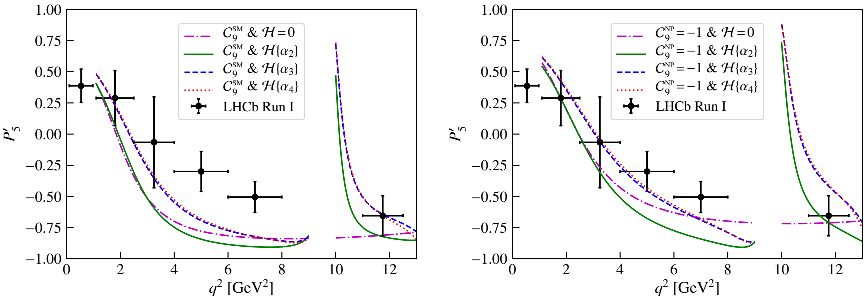


Figure 14: Projections on the P_5' angular observable of the different hypotheses for the non-local hadronic effects in the (left) `SM` and (right) `BMP`$_{\mathcal{C}_9}$ scenario. The results of the LHCb Run-I analysis [79] are overlaid as reference.

Table 5: Expected central value and uncertainties for the $\mathcal{C}_9$ Wilson coefficient obtained from fits with $\mathcal{H}_\lambda[z^2,...,z^5]$ for the `BMP`$_{\mathcal{C}_9}$scenario when produced with $\mathcal{H}_\lambda\{\alpha_3\}$ as described in the text.

	LHCb Run-II		LHCb Upgrade	
	$\mathcal{C}_9^{\rm NP}$ mean	$\mathcal{C}_9^{\rm NP}$ sigma	$\mathcal{C}_9^{\rm NP}$ mean	$\mathcal{C}_9^{\rm NP}$ sigma
$\mathcal{H}_\lambda[z^2]$	-1.708 ± 0.004	0.134 ± 0.003	-1.722 ± 0.002	0.057 ± 0.002
$\mathcal{H}_\lambda[z^3]$	-1.001 ± 0.006	0.200 ± 0.005	-1.017 ± 0.003	0.107 ± 0.002
$\mathcal{H}_\lambda[z^4]$	-1.045 ± 0.007	0.212 ± 0.005	-1.013 ± 0.004	0.122 ± 0.003
$\mathcal{H}_\lambda[z^5]$	-0.950 ± 0.009	0.267 ± 0.006	-0.983 ± 0.005	0.146 ± 0.003

Physics contribution in $\mathcal{C}_9$. However, some of the most popular beyond the SM scenarios indicate that NP can simultaneously affect $\mathcal{C}_9$ and $\mathcal{C}_{10}$ [48,49]. Furthermore, the determination of $\mathcal{C}_{10}$ is commonly known to be unaffected by the limited knowledge on the non-local hadronic contributions [103]. As a consequence, possible measurements of $\mathcal{C}_{10}$ can be even more appealing for the unambiguous claim of physics beyond the SM.

A direct measurement of $\mathcal{C}_{10}$ from experimental data can be investigated by floating simultaneously the Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ in the amplitude fit to $B^0 \to K^{*0}\mu^+\mu^-$ decays. Figure 16 shows the obtained sensitivity to the parameter space $\mathcal{C}_9$ vs $\mathcal{C}_{10}$ for fits with $\mathcal{H}_\lambda[z^2,...,z^5]$ assuming the `BMP`$_{\mathcal{C}_9}$scenario and for datasets corresponding to the LHCb Run-II and LHCb Upgrade expected statistics. The single projections are also reported in Tables 6. Several conclusion can be driven by these results

i. the uncertainty on $\mathcal{C}_9$ varying the order of the fit follows the pattern observed in section 3.2.2;

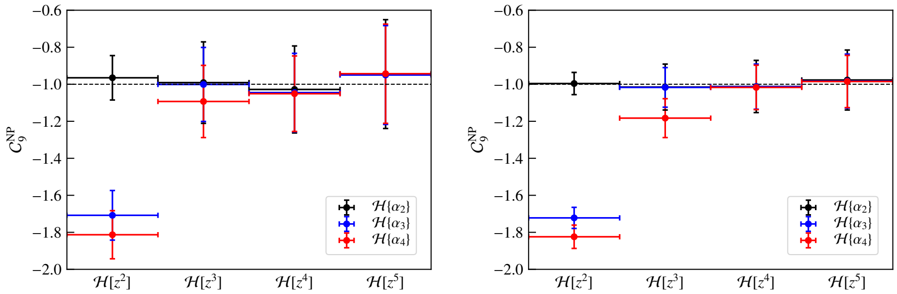


Figure 15: Results for the determination of $\mathcal{C}_9$ from the fit for the (left) LHCb Run-II and (right) LHCb Upgrade expected statistics. The expected central value and statistical uncertainties are shown for fits with $\mathcal{H}_\lambda[z^2, ..., z^5]$ (the four horizontal points of the graph) for the $\texttt{BMP}_{\mathcal{C}_9}$scenario when produced with $\mathcal{H}_\lambda\{\alpha_2, \alpha_3, \alpha_4\}$ (in different colours).

ii. due to the correlation between $\mathcal{C}_9$ and $\mathcal{C}_{10}$ the uncertainty on $\mathcal{C}_9$ alone is larger compared to the case of section 3.2.2 with $\mathcal{C}_{10}$ fixed in the fit;

iii. a precise determination of $\mathcal{C}_{10}$ is limited by the uncertainties on the form factors, which already saturates with the statistics expected to be collected after LHCb Run-II.[8] This conclusion is deduced from the approximately unmodified uncertainties between the LHCb Run-II and LHCb Upgrade expected datasets. In addition, an intrinsic bias is observed (even in the case of large datasets) for higher order of the fit, which, analogously to the determination of $\mathcal{C}_9$, must be considered in experimental measurements.

Figure 17 shows the obtained sensitivity to the parameter space $\mathcal{C}_9$ vs $\mathcal{C}_{10}$ the two considered NP scenarios $\texttt{BMP}_{\mathcal{C}_9}$and $\texttt{BMP}_{\mathcal{C}_{9,10}}$. For simplicity, only results obtained from fits with $\mathcal{H}_\lambda[z^3]$ are displayed. The expected sensitivity on $\mathcal{C}_{10}$ is found to be of great interest, already for the statistics expected at LHCb Run-II, leading to possible observation of New Physics from models with strong contribution to $\mathcal{O}_{10A}$ [48, 49]. In addition, the large dataset expected to be collected during the LHCb Upgrade will allow a deeper understanding of the nature of New Physics, by discriminating hypothetical purely vectorial from a vector-axial structures.

---

[8] The precision on the determination of $\mathcal{C}_{10}$ can be substantially improved by including constraints from $B_s \to \mu^+\mu^-$ decays, however the inclusion of this additional channel, despite intriguing, goes beyond the scope of these studies.

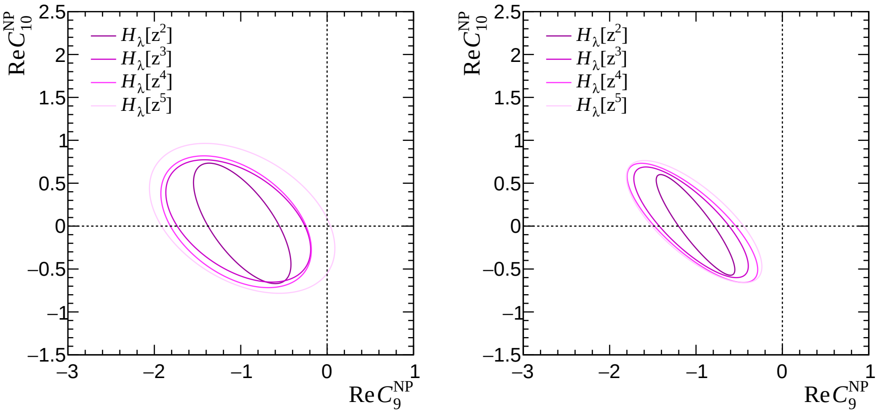


Figure 16: Two-dimensional sensitivity scans for the pair of Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ for different non-local hadronic parametrisation models $\mathcal{H}_\lambda[z^2, ..., z^5]$. The contours correspond to $3\,\sigma$ statistical-only uncertainty bands evaluated with the expected statistics after LHCb Run-II (left) and LHCb Upgrade (right).

Table 6: Fit results for the pair of Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ obtained from fits with $\mathcal{H}_\lambda[z^2, ..., z^5]$ for the $\texttt{BMP}_{\mathcal{C}_9}$scenario with the corresponding statistics expected at the LHCb Run-II and LHCb Upgrade.

	$\mathcal{C}_9^{\mathrm{NP}}$ mean	$\mathcal{C}_9^{\mathrm{NP}}$ sigma	$\mathcal{C}_{10}^{\mathrm{NP}}$ mean	$\mathcal{C}_{10}^{\mathrm{NP}}$ sigma	correlation $\mathcal{C}_9^{\mathrm{NP}}$- $\mathcal{C}_{10}^{\mathrm{NP}}$
			LHCb Run-II		
$\mathcal{H}_\lambda[z^2]$	-0.982 ± 0.008	0.164 ± 0.006	0.032 ± 0.010	0.204 ± 0.007	-0.68 ± 0.03
$\mathcal{H}_\lambda[z^3]$	-1.029 ± 0.012	0.244 ± 0.009	0.060 ± 0.010	0.207 ± 0.007	-0.47 ± 0.04
$\mathcal{H}_\lambda[z^4]$	-1.053 ± 0.013	0.253 ± 0.009	0.051 ± 0.011	0.223 ± 0.008	-0.43 ± 0.04
$\mathcal{H}_\lambda[z^5]$	-0.983 ± 0.017	0.312 ± 0.012	0.091 ± 0.013	0.254 ± 0.009	-0.40 ± 0.04
			LHCb Upgrade		
$\mathcal{H}_\lambda[z^2]$	-1.005 ± 0.007	0.132 ± 0.005	0.014 ± 0.008	0.171 ± 0.006	-0.89 ± 0.01
$\mathcal{H}_\lambda[z^3]$	-1.057 ± 0.010	0.193 ± 0.007	0.044 ± 0.009	0.188 ± 0.007	-0.79 ± 0.02
$\mathcal{H}_\lambda[z^4]$	-1.041 ± 0.011	0.220 ± 0.008	0.037 ± 0.010	0.202 ± 0.007	-0.78 ± 0.02
$\mathcal{H}_\lambda[z^5]$	-1.021 ± 0.011	0.228 ± 0.008	0.051 ± 0.010	0.207 ± 0.007	-0.73 ± 0.02

3.2.4 *Unbinned determination of angular observables*

One of the benefits of the proposed approach is that it takes advantage of the full unbinned description of the decay, which enables the reproducibility of the commonly used angular observables defined in Eq. 33 and 36.

In order to examine the statistical uncertainty expected for the angular observables, the results of the fits conduced with $\mathcal{H}_\lambda[z^2, ..., z^5]$ are analysed. The projections of the fit results on the angular observables are found to be independent on the truncation of the series expansion. This guarantees that the confidence intervals obtained for the angular

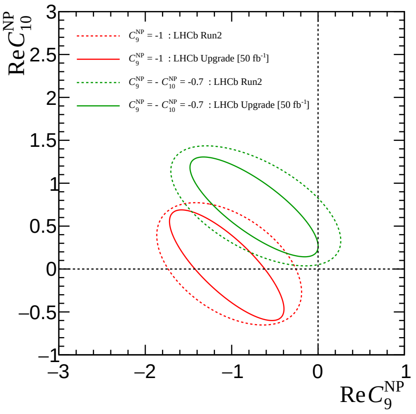


Figure 17: Two-dimensional sensitivity scans for the pair of Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ for the two considered NP scenarios (red) $\texttt{BMP}_{\mathcal{C}_9}$and (green) $\texttt{BMP}_{\mathcal{C}_{9,10}}$. The contours correspond to $3\,\sigma$ statistical-only uncertainty bands obtained with $\mathcal{H}_\lambda[z^3]$ fits for the (dotted) LHCb Run-II and (solid) LHCb Upgrade expected statistics.

observables are free from systematic uncertainties on the non-local hadronic parametrisation.

It is interesting to compare the statistical uncertainty on the angular observables obtained by the unbinned amplitude fit with respect to the classic binned approach. A q^2 binned angular fit, built on the signal *pdf* defined in Eq. 35, is performed on the same ensembles generated above, splitting in 1 GeV$^2/c^4$ q^2 bins. Figure 18 shows the result of the angular observables obtained by the unbinned amplitude fit compared to a binned angular analysis.

The derivation of the angular observables from the amplitude fit represents a result of extreme interest that, thanks to its independence on the non-local hadronic parametrisation, can play an important role in the validation of the fit results.

3.2.5 *Unbinned vs binned sensitivity to New Physics*

Despite the obvious improvement in the uncertainty on the the angular observables, it should be emphasised that the total gain in sensitivity to New Physics cannot be directly read from Fig. 18, since the points of the binned likelihood fit are all uncorrelated. To fully access the comparison of the two approaches, a binned fit to the angular observables with the same physics model should be performed. Figure 19 shows the comparison in the determination of the Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ between the unbinned amplitude fit and the binned fit to the angular observables for the $\mathcal{H}_\lambda[z^2]$ and $\mathcal{H}_\lambda[z^3]$ non-local hadronic parametrisations. Both approaches are applied to the same ensembles

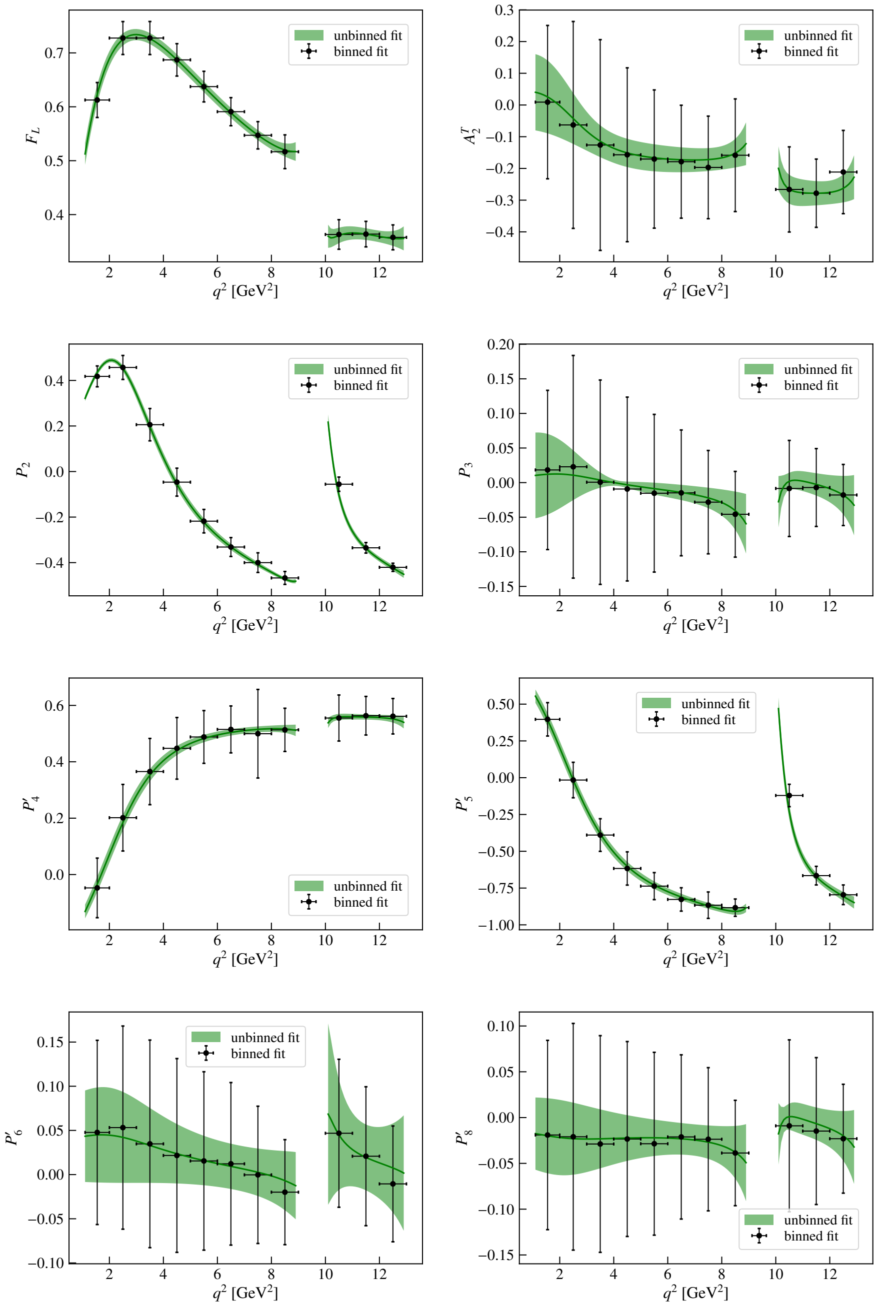


Figure 18: Optimised angular observables obtained from the unbinned amplitude fit compared to the binned angular analysis. Both approaches analyse the same set of ensembles generated with the $\texttt{BMP}_{\mathcal{C}_9}$scenario and the expected statistics at LHCb Run-II.

of pseudoexperiments produced with the $\texttt{BMP}_{\mathcal{C}_9}$scenario and the expected statistics at LHCb Run-II. For the binned fit, a χ^2 minimisation is performed using the binned angular observables and their correlations as input, together with the theoretical constraints of Sec. 3.1 and 3.2.2. The large difference in sensitivity between the two methods can

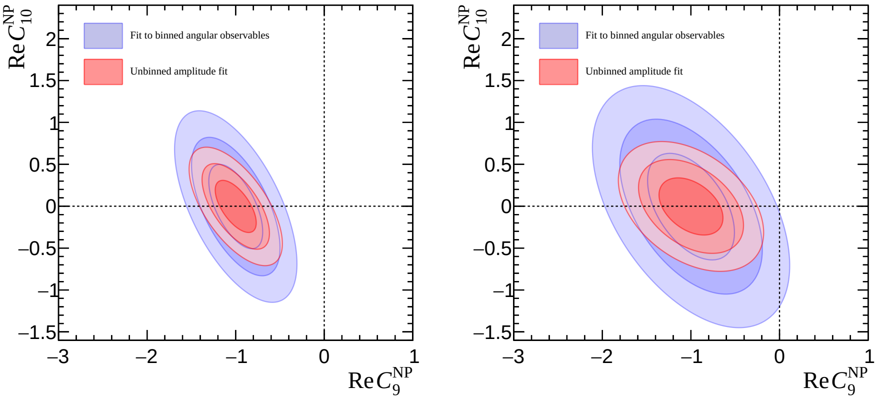


Figure 19: Two-dimensional sensitivity scans for the pair of Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ as obtained from the unbinned amplitude fit and from fits to the full set of angular observables for (left) $\mathcal{H}_\lambda[z^2]$ and (right) $\mathcal{H}_\lambda[z^3]$ non-local hadronic parametrisation models. The contours correspond to 1, 2, 3 σ statistical uncertainty evaluated with the expected statistics at LHCb Run-II for the $\mathtt{BMP}_{\mathcal{C}_9}$scenario.

be quantified in 5.9 (8.1) σ and 3.1 (4.2) σ for the binned (unbinned) and for the $\mathcal{H}_\lambda[z^2]$ and $\mathcal{H}_\lambda[z^3]$ non-local hadronic models, respectively.

This result strongly illustrates the advantage of the proposed amplitude fit with respect to the conventional binned approach.

3.3 LEPTON FLAVOUR UNIVERSALITY TEST IN $B^0 \to K^{*0}\ell^+\ell^-$

The second part of this chapter aims to extend the proposed framework to perform a simultaneous amplitude fit to $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ decays as a test of Lepton Flavour Universality. Currently, published LFU tests consist of measurements of the ratios R_X, that simply take into account the number of events observed in the two channels, and differences in the binned angular observables of $B^0 \to K^{*0}\ell^+\ell^-$ decays, in particular $\Delta P_4'$ and $\Delta P_5'$ published by Belle [80]. Performing a simultaneous amplitude fit to the full differential decay rate benefits from the full description of the decay and combines the discovery potential of the standalone ratio R_{K^*} and the flavour-dependent angular analysis. Therefore, the proposed approach is expected to enhance the sensitivity of previous LFU measurements, while preserving the beneficial property of a clean measurement free from most of the theoretical and experimental uncertainties.

3.3.1 *Simultaneous amplitude fit*

The $B^0 \to K^{*0}e^+e^-$ decay mode is included in the amplitude fit by introducing an additional $pdf^{(e)}$ based on the formalism presented in section 3.1. The key feature of this strategy is to realise that all local (form-factor) and non-local (charm-loop) hadronic matrix elements are known to be lepton-flavour universal and intervene identically in the two channels. Therefore, these are shared in the fit between both semileptonic decays resulting in the convenient cancellation of all the hadronic uncertainties, similarly to the case of the ratio measurements R_X. This approach benefits from the large statistics available for $B^0 \to K^{*0}\mu^+\mu^-$ decays that is sufficient to enable the determination of these multi-space parameters. Note that an amplitude analysis of the electron mode only has been previously disregarded in terms of current probe, given the limited dataset in either LHCb or Belle experiments. In this common framework, the hadronic contributions are treated as nuisance parameters, while only the Wilson coefficients $\mathcal{C}_9^{(\mu,e)}$ and $\mathcal{C}_{10}^{(\mu,e)}$ are kept separately for the two channels. For consistency the Wilson coefficients C_7 is also shared in the fit and fixed to its Standard Model value.

Commonly to most of the LFU tests performed at hadron colliders, the limiting factor for this measurements is represented by the limited statistics of $B^0 \to K^{*0}e^+e^-$ decays, due to the lower electron reconstruction efficiency (which includes triggering, reconstruction and identification) compared to the muon case. In the following, the kinematic regions $Q_1^{(\mu)}$, $Q_2^{(\mu)}$ and $Q^{(e)}$ defined in Eq. 40 are used as domain for the simultaneous amplitude fit.

Similarly to the study of Sec. 3.2.1, the model-(in)dependence of the framework is investigated by modifying the non-local hadronic parametrisation. In addition to the baseline model $\mathcal{H}_\lambda[z^K]$ (*i.e.* free-floating analytical functions truncated at different orders, independent on any theoretical inputs) two complementary assumptions are considered

i. the Standard Model predictions of the the non-local contributions of Ref. [59] are included as a multivariate Gaussian constraints on $\{\alpha_k^{(\lambda)}\}$ (this assumption is limited to $\mathcal{H}_\lambda[z^2]$);

ii. the non-local hadronic functions are re-parametrisation as proposed in Ref. [86], *i.e.* using instead an expansion in terms of $1 + a_\lambda + b_\lambda(q^2/6\,[\mathrm{GeV}^2/c^4])$.

Figure 20 shows the fit results for several alternative parametrisations of the non-local hadronic matrix elements for the $\mathtt{BMP}_{\mathcal{C}_9}$hypothesis, with signal yields corresponding to LHCb Run-II scenario. The sensitivity to $\mathcal{C}_9^{(\mu,e)}$ is proven to be strongly dependent on the model assumption, as seen before. For this reason, in the following - where no theory information is used - the usual Wilson coefficients $\mathcal{C}_i^{(\mu,e)}$ are renamed as $\widetilde{\mathcal{C}}_i^{(\mu,e)}$, since an accurate disentanglement between the physical meaning of $\mathcal{C}_i^{(\mu,e)}$ and the above-mentioned hadronic pollu-

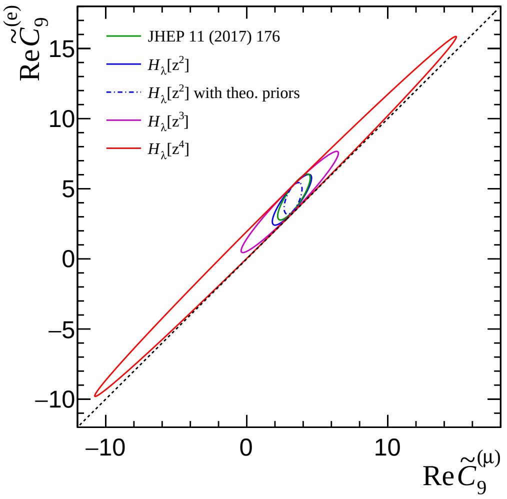


Figure 20: Two-dimensional sensitivity scans for the pair of Wilson coefficients $\widetilde{\mathcal{C}}_9^{(\mu)}$ and $\widetilde{\mathcal{C}}_9^{(e)}$ for different non-local hadronic parametrisation models evaluated at $\texttt{BMP}_{\mathcal{C}_9}$ and with the expected statistics after LHCb Run-II. The contours correspond to $3\,\sigma$ statistical-only uncertainty bands and the dotted black line indicates the LFU hypothesis.

tion cannot be achieved without additional external information, as evidenced in Sec. 3.2.

Nonetheless, it is noticeable that the high correlation of the $\widetilde{\mathcal{C}}_9^{(\mu)}$ and $\widetilde{\mathcal{C}}_9^{(e)}$ coefficients is sufficient to preserve the true underlying physics at any order of the series expansion $\mathcal{H}_\lambda[z^K]$ and without any parametric theoretical input, *i.e.* the two-dimensional pull estimator with respect to the LFU hypothesis is unbiased.

This remarkable feature leads to the definition of new parameters of interest, *i.e.* the difference of Wilson coefficients between muons and electrons

$$\Delta\mathcal{C}_i = \widetilde{\mathcal{C}}_i^{(\mu)} - \widetilde{\mathcal{C}}_i^{(e)}\,, \tag{49}$$

particularly convenient in the investigation of LFU-breaking hypothesis. Figure 21 demonstrates this principle with a series of signal-only sensitivity studies with the two $\texttt{BMP}_{\mathcal{C}_9}$and $\texttt{BMP}_{\mathcal{C}_{9,10}}$scenarios using the proposed parameters $\Delta\mathcal{C}_i$. The obtained sensitivity is found to be completely independent on the parametrisation of the non-local hadronic contributions. The maximal expected significance with respect to the Standard Model can be quantified as 4.6 and $5.3\,\sigma$ for $\texttt{BMP}_{\mathcal{C}_9}$and $\texttt{BMP}_{\mathcal{C}_{9,10}}$, respectively. Realistic experimental effects are necessary to determine the exact sensitivity achievable. Nevertheless, these results suggest that a first observation (with a single measurement) of Lepton Flavour Uni-

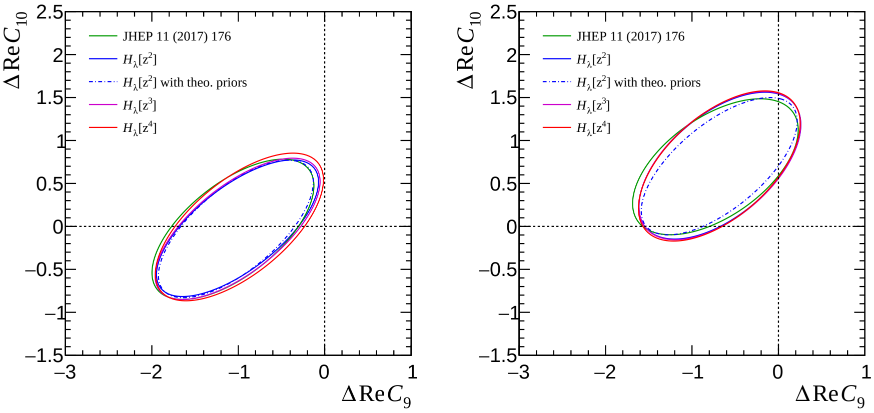


Figure 21: Two-dimensional sensitivity scans for the proposed parameters $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ for different non-local hadronic parametrisation models evaluated at (left) $\mathtt{BMP}_{\mathcal{C}_9}$ and (right) $\mathtt{BMP}_{\mathcal{C}_{9,10}}$, and with the expected statistics after LHCb Run-II. The contours correspond to $3\,\sigma$ statistical-only uncertainty bands.

versality breaking appears to be feasible with the expected recorded statistics by the end of LHCb Run-II.[9]

One interesting characteristic of the proposed parameters $\Delta\mathcal{C}_i$ is that these are not only unaffected by the non-local hadronic pollution but also independent on form factors uncertainties. Thus, it is appealing to examine the prospects for the upcoming LHCb Upgrade. Figure 22 summarises the two-dimensional statistical-only significances for the designed LHCb Run-II and Upgrade luminosities. The unprecedented dataset expected to be collected after the LHCb Upgrade will not only yields insight on this phenomena but also enables a deeper and clean understanding of the Nature of New Physics - independently on any hadronic uncertainties. In the proposed approach, in fact, the complete cancellation of hadronic uncertainties occurs also in presence of New Physics. This unique characteristic is a natural consequence of the amplitude fit, which allows direct inspection of the physical amplitudes that, in turn, depend linearly on NP effects.[10]

3.3.2 *Proving the cancellation of hadronic elements*

Figure 21 demonstrates the independence of the parameters $\Delta\mathcal{C}_i$ on the parametrisation of the non-local hadronic contributions. However, it is interesting to investigate to which extension this assumption is

---

[9] The reader must note that the two NP scenarios considered here correspond to a very conservative value of R_{K^*} for the central q^2 bin, $R_{K^*} \simeq 0.85\,(0.75)$ for $\mathtt{BMP}_{\mathcal{C}_9}$and $\mathtt{BMP}_{\mathcal{C}_{9,10}}$, respectively.

[10] Note that "classic" LFU observables (*i.e.* R_X, $\Delta P_i'$) are quadratic in the amplitudes, therefore, in presence of non-LFU NP, the cancellation of hadronic uncertainties is only partial due to the contribution of mix terms $\mathcal{C}_9 \cdot \mathcal{H}_\lambda$.

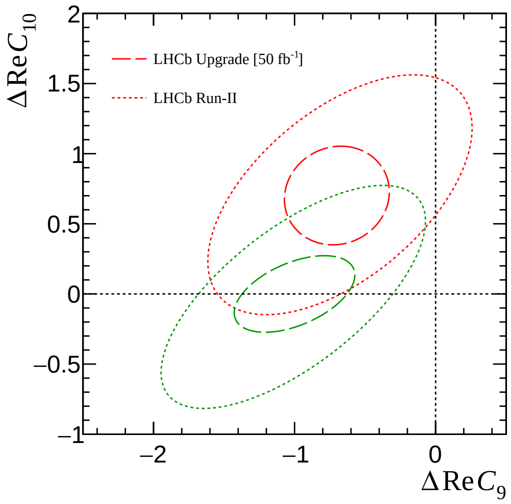


Figure 22: Two-dimensional sensitivity scans for the proposed parameters $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ for the two considered NP scenarios: (green) $\mathtt{BMP}_{\mathcal{C}_9}$ and (red) $\mathtt{BMP}_{\mathcal{C}_{9,10}}$. The contours correspond to $3\,\sigma$ statistical-only uncertainty bands expected for the (dotted) LHCb Run-II and (dashed) LHCb Upgrade statistics.

Table 7: Fit results for the parameters $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ obtained from fits with $\mathcal{H}_\lambda[z^2, z^3, z^4]$ for the $\mathtt{BMP}_{\mathcal{C}_9}$scenario when produced with $\mathcal{H}_\lambda\{\alpha_4\}$ as described in the text.

	LHCb Run-II				
	$\Delta\mathcal{C}_9$ mean	$\Delta\mathcal{C}_9$ sigma	$\Delta\mathcal{C}_{10}$ mean	$\Delta\mathcal{C}_{10}$ sigma	correlation $\Delta\mathcal{C}_9$ - $\Delta\mathcal{C}_{10}$
$\mathcal{H}_\lambda[z^2]$	-1.021 ± 0.013	0.265 ± 0.009	0.020 ± 0.010	0.192 ± 0.007	0.51 ± 0.04
$\mathcal{H}_\lambda[z^3]$	-0.997 ± 0.014	0.279 ± 0.010	0.004 ± 0.010	0.202 ± 0.007	0.41 ± 0.04
$\mathcal{H}_\lambda[z^4]$	-0.986 ± 0.014	0.278 ± 0.010	-0.007 ± 0.010	0.202 ± 0.007	0.45 ± 0.04

satisfied. A stringent probe to provide further confidence in the robustness of the proposed approach corresponds to the case of ensembles produced with high orders of the non-local hadronic expansion $\mathcal{H}_\lambda(z)$ and fitted with a lower *"wrong"* order. Table 7 shows the fit results for the pair of parameters $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ for ensembles generated with $\mathcal{H}_\lambda\{\alpha_4\}$ as described in section 3.2.2.1 and fitted with different orders of the series $\mathcal{H}_\lambda[z^2, z^3, z^4]$. The obtained expected values are found to be unbiased independently on the order of the fitted parametrisation, confirming the cancellation of the hadronic contributions that results in a measurement free from any theoretical uncertainties.[11]

---

[11] Despite only analytical functions are considered in this study, possible non-analytical contributions are not expected to break this mechanism.

3.3.3 *Preliminary experimental considerations*

Experimental resolution and detector acceptance/efficiency effects are not considered in this chapter, as these would require further information from current (non-public for a phenomenology study) or planned B-physics experiments. Nevertheless, it is important to highlight that the finite q^2 resolution is often the limiting factor for measurements involving electrons. Thus, preliminary studies to investigate the impact of q^2 resolution effects on the proposed method are of extreme interest to establish the feasibility of the measurement.

A q^2-constant asymmetric smearing of the di-lepton invariant mass is assumed for the electron mode; the size and asymmetry of such smearing are described by a double Gaussian

$$\begin{aligned} dG &= 0.7 \times \mathcal{G}(\mu = -30\,\mathrm{MeV}/c^2,\, \sigma = 70\,\mathrm{MeV}/c^2) \\ &+ 0.3 \times \mathcal{G}(\mu = -120\,\mathrm{MeV}/c^2,\, \sigma = 150\,\mathrm{MeV}/c^2) \end{aligned} \tag{50}$$

which values are naively chosen to reproduce the mass fits of Ref. [88]. The obtained fit results are shown in Fig. 23 (left) and Table 8. Remarkably, sizeable resolution effects typical of hadronic machines are found not to affect the determination of $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$, whose estimators remains unbiased, even if no-correction is applied. This result further highlights the great stability of the proposed approach, which benefits from the large dataset available for $B^0 \to K^{*0}\mu^+\mu^-$ decays to determine all the nuisance parameters, while only $\widetilde{\mathcal{C}}_9^{(e)}$ and $\widetilde{\mathcal{C}}_{10}^{(e)}$ are bound to the electron sample. This interpretation can be tested by "inverting " the statistics of the two channels, in which case most of the parameters are determined by the (mis-reconstructed) electron sample and the estimator of $\Delta\mathcal{C}_9$ results to be strongly biased.

As a consequence, the proposed approach can influence the design of future experiments and upgrades, which are mostly focused on how to improve the reconstruction and resolution for electron channels. These studies shows that, although such development is desirable, muon channels can still be the key target for hadronic machines.

Moreover, the default q^2 range $Q^{(e)}$ defined in Eq. 40 assumes an improvement on the selection of $B^0 \to K^{*0}e^+e^-$ decays at LHCb. All publications related to the electron channel so far define the central q^2 region up to 6 GeV^2/c^4. Figure 23 (right) shows the improvement in the significance to the $\mathtt{BMP}_{\mathcal{C}_9}$scenario accomplished by the extension of the electron domain up to 7 GeV^2/c^4, as proven to be possible in [96]. Since the limiting factor is driven by the electron statistics, the uncertainty on the measurement greatly benefits from this experimental development and the enhancement in its sensitivity can be quantified from 3.6 to 4.6$\,\sigma$ for the $\mathtt{BMP}_{\mathcal{C}_9}$scenario.

Finally, the differential decay width can receive additional complex amplitudes from signal-like backgrounds, *e.g.* $K\pi$ S-wave from a non-

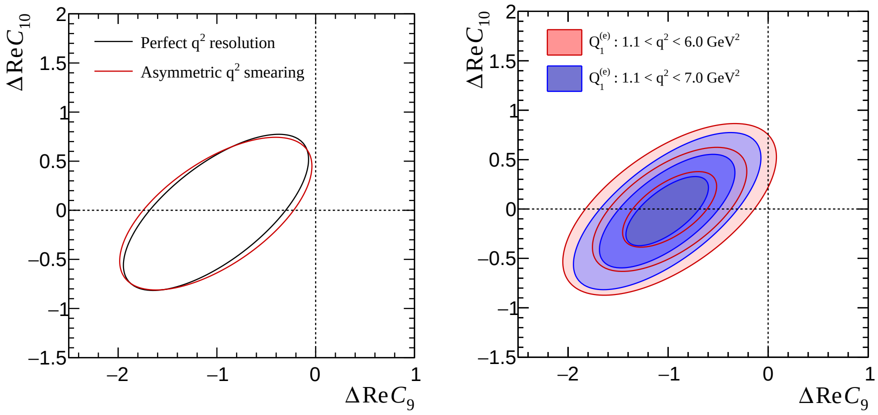


Figure 23: Two-dimensional sensitivity scans for the proposed parameters $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$. The nominal fit is compared with two modified configuration for the electron mode: (left) an asymmetric q^2 smearing is applied to the di-lepton invariant mass; (right) $B^0 \to K^{*0}e^+e^-$ decays are selected in a reduced q^2 range. The contours correspond to $3\,\sigma$ (or 1, 2, $3\,\sigma$) statistical-only uncertainty bands for the $\mathtt{BMP}_{\mathcal{C}_9}$scenario with the expected statistics after LHCb Run-II.

Table 8: Fit results for the parameters $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ obtained from the fit for the $\mathtt{BMP}_{\mathcal{C}_9}$scenario assuming perfect resolution or an asymmetric q^2 smearing.

	LHCb Run-II				
	$\Delta\mathcal{C}_9$ mean	$\Delta\mathcal{C}_9$ sigma	$\Delta\mathcal{C}_{10}$ mean	$\Delta\mathcal{C}_{10}$ sigma	correlation $\Delta\mathcal{C}_9$ - $\Delta\mathcal{C}_{10}$
Perfect q^2 resolution	-1.002 ± 0.014	0.267 ± 0.009	-0.018 ± 0.012	0.236 ± 0.008	0.68 ± 0.03
Asymmetric q^2 smearing	-1.011 ± 0.014	0.283 ± 0.010	-0.034 ± 0.011	0.226 ± 0.008	0.62 ± 0.03

resonant decay and/or a scalar resonance (see detailed discussion in chapter 7). These contributions are determined to be small [77, 79], and, in the proposed formalism, benefit from the same description between the muon and electron mode. Therefore, in this constrained framework these effects are even further suppressed and can then be neglected for the purpose of these sensitivity studies.

3.3.4 *Unbinned vs binned approach*

Similarly to the results obtained in section 3.2.4, the classic angular observables can be derived from the results of the amplitude fit. It is interesting to compare the gain in sensitivity obtained with the simultaneous amplitude fit with respect to the traditional binned angular analysis, in particular in the perspective of LFU tests. Figure 24 shows the values obtained for the angular observable P_5' for the two channels, assuming the $\texttt{BMP}_{\mathcal{C}_9}$scenario and the expected statistics at LHCb Run-II. As a comparison, the results obtained by the binned angular analyses of the two decay modes are overlaid in the figure, where 1 GeV^2/c^4 and 2 GeV^2/c^4 q^2 bins have been considered for the muon and electron channels, respectively. To compensate the low statistics expected for the electron channel, the binned angular analysis of $B^0 \to K^{*0}e^+e^-$ decays assumes the folding technique described in Ref. [78] to reduce the number of degrees of freedom of the angular *pdf*, nevertheless the uncertainty associated to the electron channel is found to be large and limits the sensitivity to New Physics of classic angular analyses. In addition, in order to fairly compare the expected sensitivity to NP, a fit to the Wilson coefficients based on the two methods is performed on ensembles produced with the expected statistics at LHCb Run-II and the $\texttt{BMP}_{\mathcal{C}_9}$scenario.

The results of the amplitude fit of Fig. 21 (left) is compared to a binned fit to the angular observables, where different foldings have been applied to the electron channel in order to separately extract $P_4'^{(e)}$ $P_5'^{(e)}$, $P_6'^{(e)}$ and $P_8'^{(e)}$, while the full set of angular observables is employed for the muon mode. The sensitivity to the differences of Wilson coefficients results to be extremely poor, *i.e.* the uncertainty on $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ is found to be $\mathcal{O}(1)$, and makes impossible any claim of New Physics in the near future purely based on the (difference) of angular observables. Note that the binned observables $\Delta P_i'$ are based on separate analyses of the muon and electron decay modes and, despite their clean SM predictions, their experimental determination unavoidably ignores the correlations between the two channels. Oppositely, the amplitude fit benefits from the sharing of all the nuisance parameters and enables an accurate determination of the eventual different physics underlying the two decays. In summary, the crucial advantage of the proposed amplitude fit stands in the simultaneous treatment of the muon and electron channels.

Finally, the great extent and versatility of the amplitude fit allow to also calculate the unbinned value for the R_{K^*} observable. Figure 25 shows the obtained results as function of q^2 for the different considered scenarios. Notice that, due to the presence of light resonances, the present description of the amplitude fit does not allow to examine the low-q^2 region, where the classic binned analysis remains the only current possible approach.

3.4 CONCLUSIONS

This chapter described a novel approach to investigate the New Physics discovery potential of $B^0 \to K^{*0}\ell^+\ell^-$ decays. It consists of an unbinned amplitude fit that allows a direct determination of the Wilson coefficients from data, combining all the available information in the decay, from angular analysis to branching ratio measurements.

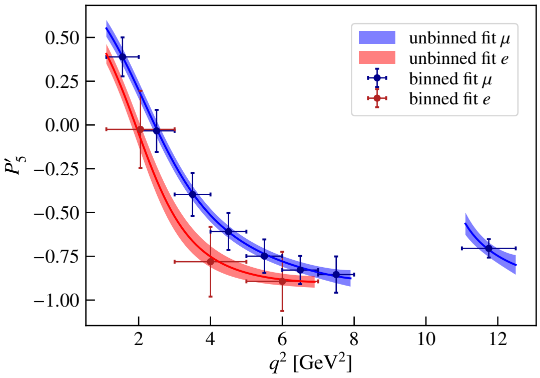


Figure 24: F_L and P_5' angular observables obtained for the muon and electron channel from the unbinned amplitude fit compared to binned angular analyses separately of the two decay modes. Both approaches analyse the same set of ensembles generated with the $\texttt{BMP}_{\mathcal{C}_9}$ scenario and the expected statistics at LHCb Run-II.

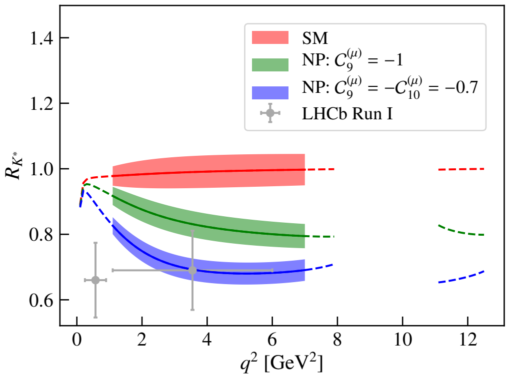


Figure 25: Values of $R_{K^{*0}}$ obtained from the unbinned amplitude fit for the SM and the two NP scenarios $\texttt{BMP}_{\mathcal{C}_9}$ and $\texttt{BMP}_{\mathcal{C}_{9,10}}$ and for the expected statistics after the LHCb Run-II. The results are shown in the q^2 region given by $Q^{(e)} = [1.1, 7.0]$ GeV$^2/c^4$ and are extrapolated for illustration to the low-q^2 region and the kinematic regions $Q_1^{(\mu)}$ and $Q_2^{(\mu)}$. The Run-I LHCb measurement [88] is overlaid as a reference.

The proposed method is firstly employed to analyse $B^0 \to K^{*0}\mu^+\mu^-$ decays alone, where large effort is dedicated to the determination of the model-dependence introduced by the truncation of the analytical expansion of the non-local hadronic matrix elements. The state-of-the-art knowledge on these non-local hadronic contributions (charm-loop), as well as on the local matrix elements (form factors) and CKM parameters is included in the framework creating a propitious combined theory-experimental synergy to unravel the long-standing issue that limited the confidence in the $b \to s\mu\mu$ anomalies. For the first time, a quantitative evaluation of the model-dependence associated to the non-local hadronic contributions is accessed by studying the impact of the order of the z-expansion on the parameters of interest. Following the proposed approach, an adequate estimation of systematic uncertainties in direct measurements of Wilson coefficients can be achieved.

Secondly, the described amplitude fit is extended to $B^0 \to K^{*0}e^+e^-$ decays in the perspective of Lepton Flavour Universality test. This simultaneous approach results in a clean and model-independent method for a precise determination of LFU-breaking differences of Wilson coefficients, *i.e.* $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$. This relies on a shared parametrisation of the local and non-local hadronic matrix elements between the muonic and electronic channels, which in turn enables the determination of the parameters of interest free from any theoretical uncertainty. In addition, this simultaneous analysis is found to be robust against experimental effects such as mismodeling of the detector resolution, since most parameters are effectively determined from the muon mode. This would be an important benefit for LHCb where the electron resolution is significantly worse than that of muons. Figure 26 illustrates the usefulness of the newly-proposed parameters by combining the different information of the decay, from angular analysis to branching ratio measurements. Due to the inclusiveness of the approach, the expected sensitivity surpasses any of the projections for the foreseen measurements of *e.g.* R_{K^*} or P_5' alone - given the benchmark points. Therefore, this novel formalism can be the most immediate method to observe unambiguously New Physics in $B^0 \to K^{*0}\ell^+\ell^-$ decays.

Finally, a promising feature of this framework is the possibility to extend this analysis to include other decay channels involving flavour changing neutral currents. For instance, the charged decay $B^+ \to K^{*+}\ell^+\ell^-$ undergoes the same physics and is easily accessible at the B factories, while other rare semi-leptonic decays such as $B^+ \to K^+\ell^+\ell^-$ and $\Lambda_b \to \Lambda^{(*)}\ell^+\ell^-$ have different phenomenologies but access the same New Physics information in terms of Wilson coefficients description. Thus, an unbinned global simultaneous fit to all data involving $b \to s\ell\ell$ transitions is a natural and appealing extension of this work. Moreover, the parameter space of the investigated Wilson coefficients can be broadened to incorporate direct determination of the right-handed $\mathcal{C}_i'$, as described in appendix A.1.4, and hypothetical non-trivial imaginary part $\mathcal{I}m\,\mathcal{C}_i$, which can be accessed by exploiting the CP asymmetries between $B^0 \to K^{*0}\ell^+\ell^-$ and $\bar{B} \to \bar{K}^{*0}\ell^+\ell^-$ decays.

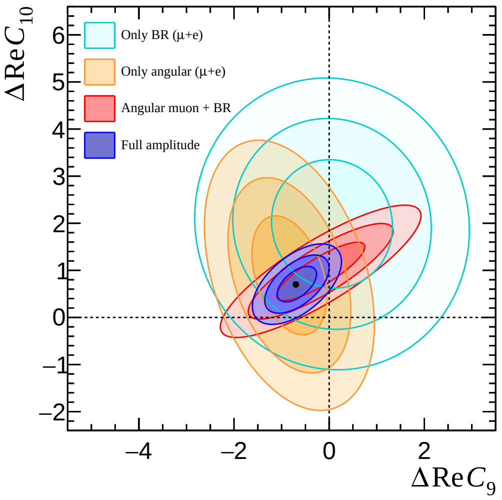


Figure 26: Sensitivity to $\texttt{BMP}_{\mathcal{C}_{9,10}}$ scenario for the expected statistics after the LHCb Run-II. The relative contribution (1, 2, $3\,\sigma$ contours) of each step of the analysis is shown in different colours, together with the result of full amplitude method proposed in this chapter.

"Beim Menschen ist es wie beim
Velo. Nur wenn er faehrt, kann
er bequem die Balance halten."

— Albert Einstein, *Letter to his son Eduard*, February 5, 1930

4

LHC AND THE LHCb DETECTOR

From a theoretical point of view, decays of elementary particles are very well defined phenomena that can be described by their interactions at all the stages of the process. In particle physics experiments, particle decays can only be accessed through the reconstruction of the final state products. In general, the detection these particles is based on their interaction with matter, which can be different for charged or neutral particles and dependent on the energy domain. These principles rule the geometry and materials used to build particle detectors, optimised for the reconstruction of momentum, mass/energy and trajectory of the final state particles.

This chapter gives a brief review of the LHC accelerator complex and the LHCb experiment, with a more detailed description of the performance of the TT silicon detectors.

4.1 THE LARGE HADRON COLLIDER

The Large Hadron Collider (LHC) is a particle accelerator and collider located at the *Conseil Européenne pour la Recherche Nucléaire* (CERN). The LHC is placed in a 27 Km long tunnel, approximately 100 meters under the ground, that hosted the former Large Electron Positron Collider (LEP). Other accelerators that had been originally built at CERN for previous experiments work as an injection chain for the LHC, as shown in Fig. 27. The proton beam starts from LINAC, a small linear accelerator, where its energy firstly reaches 50 MeV. Afterwards, the beam passes through a booster, where is split in four bunches, the Proton Synchrotron (PS), where is accelerated up to 25 GeV, and the Super Proton Synchrotron (SPS) up to 450 GeV. Finally, proton beams are injected in the LHC ring, where they circulate in opposite trajectories. Proton beams are bent around the circumference of the LHC using a total number of about 9300 superconducting dipole magnets, operating at a temperature of 1.9 K and producing a magnetic field of 8.3 T, while the accelerating power is given by radiofrequency cavities working at 400.79 MHz. The energy gain per revolution is 485 keV, in spite of the 7 keV loss per turn due to synchrotron radiation.

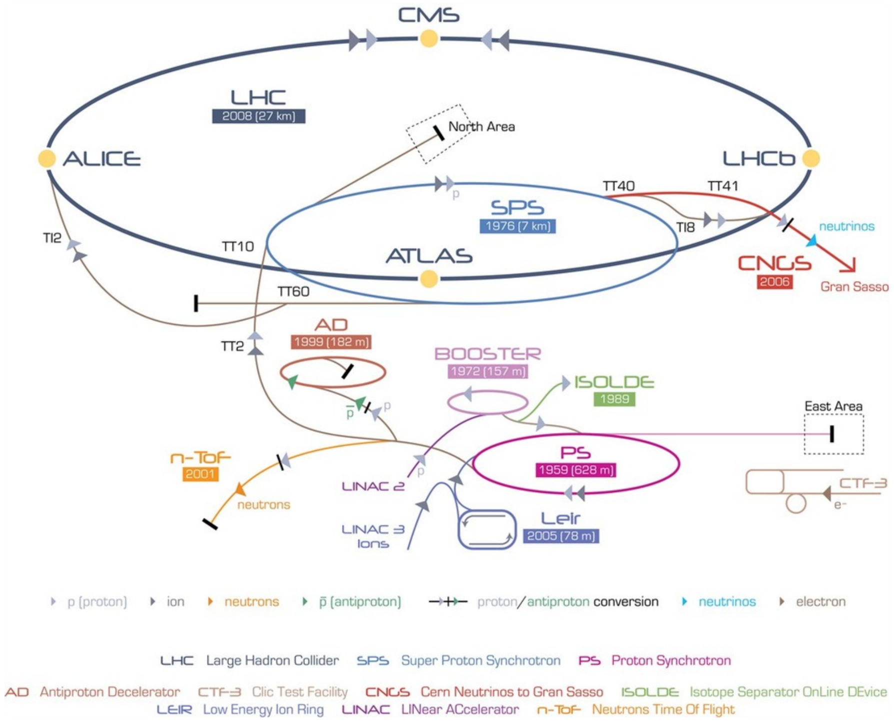


Figure 27: Schematic view of the accelerator complex at CERN [110].

The expected rate of a process produced during the collisions is given by the product of its cross section and the instantaneous luminosity. The latter can be expressed as

$$\mathcal{L} = \frac{f \gamma N_p^2 n_b}{4\pi \epsilon_n \beta^*} F, \tag{51}$$

where N_p and n_b are respectively the number of particles per bunch and bunches per beam, f the revolution frequency, γ the relativistic factor, ϵ_n the normalised transverse beam emittance, β^* quantifies the amount of focusing of the beam and F is a geometrical reduction factor due to the beams crossing angle.

The high luminosity of the LHC, designed to reach up to an instantaneous luminosity of $10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, is obtained with a high frequency bunch crossing, a high density of protons per bunch and a focusing system that reduces the transverse dimension of the beam at the interaction points. Each proton beam at full intensity consists of 2808 bunches, and each bunch contains 1.15×10^{11} protons.

During the Run-I of the LHC program, proton bunches were colliding at a proton-proton collision energy of $\sqrt{s} = 7$ and 8 TeV spaced by 50 ns, while during the ongoing Run-II the energy has been increased to 13 TeV and the collision rate has been doubled.

The proton beams interact in four points of the LHC ring, where the main experiments are located: ATLAS [111], CMS [112], LHCb [1]

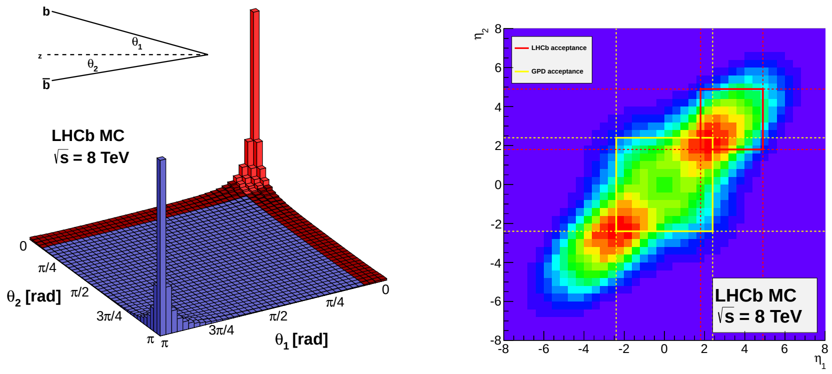


Figure 28: The correlation of the $b\bar{b}$ pair angle is shown in terms of pseudorapidity (left) and polar angle (right) together with a comparison between the LHCb [1] and ATLAS/CMS acceptances [111, 112].

and ALICE [113]. ATLAS and CMS are general purpose detectors, ideally suited for direct searches for New Physics, while LHCb is designed to investigate heavy flavour physics phenomena and ALICE focuses on the study of quark-gluon plasma processes through heavy ion collisions.

4.2 THE LHCb EXPERIMENT

The LHC is an enormous source of b hadrons, including B^+, B^0, B^0_s, B^+_c mesons and b baryons (*e.g.* Λ^0_b and Ξ^0_b). At 7 TeV centre-of-mass energy the $b\bar{b}$ cross section is found to be [114]

$$\sigma(pp \to b\bar{b}X) = (284 \pm 20 \pm 49)\,\mu\text{b}, \tag{52}$$

and is expected to increase approximately linearly by raising the energy from 7 TeV to 13 TeV [115], resulting in roughly 10^5 $b\bar{b}$ pairs produced every second.

Quark-antiquark pairs are produced in the collisions by interactions between the constituents of the protons; these partons typically have asymmetric momenta and, since the mass of the $b\bar{b}$ is small compared to the energy of the collision, are strongly boosted in the forward (or backward) direction, resulting in the angular distributions shown in Fig. 28.

Furthermore, b hadrons have a relatively long lifetime, between 0.4 (B^+_c) and 1.6 ps (B^+) [116]. This results, combined with the relativistic boost, in an average flight distance of few millimetres in the detector. Thus, the identification of secondary vertices (SV) at short distance from the interaction point is of primary importance to distinguish

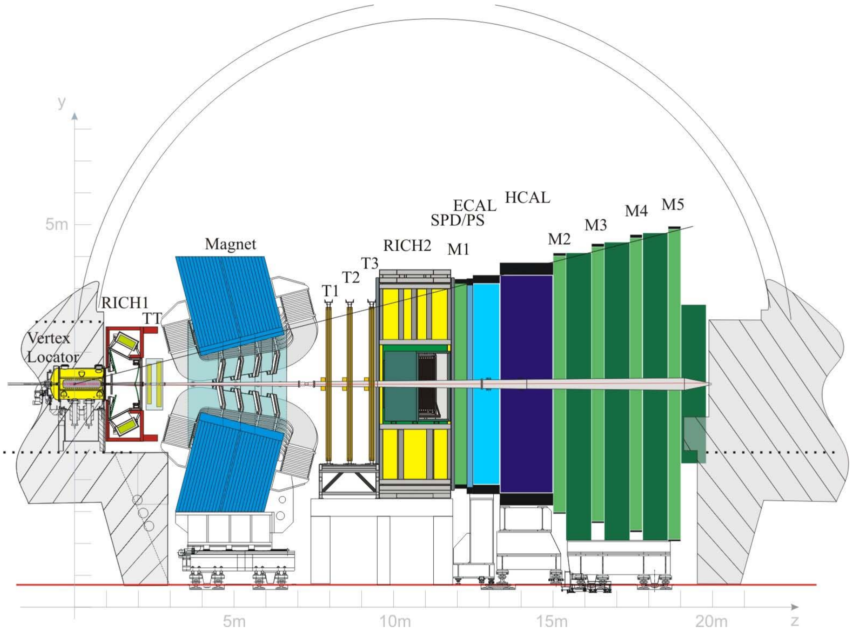


Figure 29: Schematic view of the LHCb detector [1]. The coordinate system is defined with the z-axis along the beam and the y-axis along the vertical.

b-hadron decays from charged particles originated directly from the pp collision.

The design of the LHCb detector [1] is optimised for precise measurements in the heavy quark sector. It consists of a single-arm forward dipole spectrometer with a forward angular coverage from 15 mrad to 300 (250) mrad in the bending (non-bending) plane, corresponding to a pseudorapidity[1] range of $1.8 < \eta < 4.9$. A schematic view of the LHCb detector is shown in Fig. 29 with highlighted the different sub-detectors.

In order to guarantee the optimal detector performance, the instantaneous luminosity provided by LHC is levelled down to 2×10^{32} cm^{-2}s^{-1}. This solution is adopted to reduce the pile-up (*i.e.* the number of pp collisions in a single bunch crossing) and the particle occupancy in the detector (extremely severe in the forward region) and is obtained by a misalignment of the two colliding proton beams at the interaction point at LHCb. In addition, the LHCb luminosity control ensures constant operations at the detector's optimal instantaneous luminosity, in order to stabilise the data taking conditions and maximising the recorded integrated luminosity. This procedure is often referred to as *luminosity levelling*.

---

[1] Pseudorapidity is defined as $\eta = -\ln(\tan\theta/2)$, where θ is the polar angle with respect to the beam axis.

In general, the LHCb operation system can be described by three components: tracking, particle identification (PID) and trigger systems, which are discussed in the following.

4.2.1 *Tracking system*

The trajectories of charged particles are reconstructed using information from different sub-detectors dedicated to the tracking: the Vertex Locator (VELO), the Tracker Turicensis (TT) and the Inner and Outer Tracker (IT and OT). Between the TT and the subsequent tracking stations is placed a dipole magnet with an integrated bending power of 4 Tm. The presence of the magnetic field allows, from the curvature of reconstructed tracks, the determination of the momentum of charged particles with a resolution of $\delta p/p \simeq 0.5\%$ for momenta up to 200 GeV/c. The polarity of the magnet is periodically inverted to control possible detection asymmetries.

4.2.1.1 *The Vertex Locator*

The VELO [117] is the closest sub-detector to the interaction point, surrounding the beam at only 8 mm from its axis. It is composed of two semi-circular halves that are retracted during the beam injection and energy ramping, to avoid possible damage to the sensors due to beam instabilities.

Each VELO half consists of 21 modules of pair of silicon microstrip sensors that provide radial and azimuthal coordinates, by mean of different strip directions, as shown in Fig. 30. Both R and ϕ sensors cover the region from $r = 8.2$ to $r = 42$ mm, while the strip pitch increases linearly for the ϕ sensors from 38 μm at the innermost radius to 102 μm at the outermost radius.

The VELO has the fundamental task to reconstruct primary vertices (PVs) and b-hadrons decays vertices, essential for all the analyses in LHCb. A PV resolution of 13 μm in the transverse plane and 71 μm along the beam axis is achieved for vertices with 25 tracks (typical LHCb event). An impact parameter (IP) resolution better than 35 μm is found for particles with transverse momentum greater than 1 GeV/c. The overall performance of the VELO is described in detail in Ref. [117].

4.2.1.2 *The Tracker Turicensis*

The Tracker Turicensis (TT) [118] is the tracking station placed upstream of the magnet. It consists of four layers of silicon microstrip sensors and it is designed to cope with the high track occupancy around the beam pipe. Each layer covers a 150 cm wide and 130 cm high area, while the orientation of the silicon strips follows a "*x-u-v-x*" layout, where the x layers are aligned vertically and measure the x

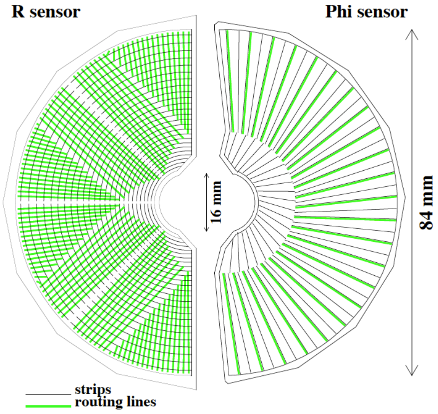


Figure 30: Schematic representation of an R-type and a ϕ-type VELO sensor [1].

coordinate while the u and v layers are rotated of $+5^\circ$ and -5° with respect to the vertical in order to allow a coarse determination of the y coordinate. The layout of the TT is sketched in Fig. 31.

Each TT layer is constituted of half-modules with an overlap of few millimeters to avoid insensitive areas in the acceptance. Each half-module is built of seven silicon sensors, grouped in two or three read-out sectors. The three central modules close to the beam pipe have three sectors containing one, two and four sensors from the inner to the outer region, respectively. The other modules are organised in three- and four-sensors sectors. The silicon sensors are 9.44 cm long, 9.64 cm wide and 500 μm thick and carry 512 read-out strips, corresponding to a strip pitch of 183 μm.

4.2.1.3 *The Inner Tracker*

The Inner Tracker (IT) [119] refers to the silicon detector placed in the three tracking stations downstream of the magnet, covering nearly 4 m^2 of active area around the beam pipe (120 cm wide and 40 cm high). Each IT station consists of four layers that follows the same x-u-v-x layout of the TT. As shown in Fig. 32, each IT layer is made of one (top and bottom units) or two-sensors (side units) modules. The silicon sensors of the IT are similar to the ones of the TT but with the slightly modified geometry of 7.6 cm of width, 11 cm of length and 384 strips with a pitch of 197 μm.

4.2.1.4 *The Outer Tracker*

The Outer Tracker (OT) [120] forms the outer part of the downstream tracking stations, where the particle flux is reduced compared to the innermost region. For this reason, a coarser granularity is tolerated and the lower-priced drift-tube technology is used to cover the large

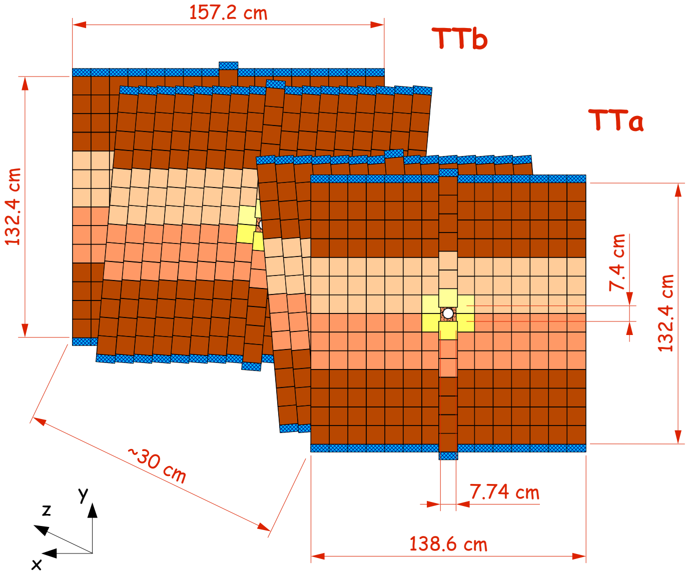


Figure 31: Schematic view of the four TT layers [1].

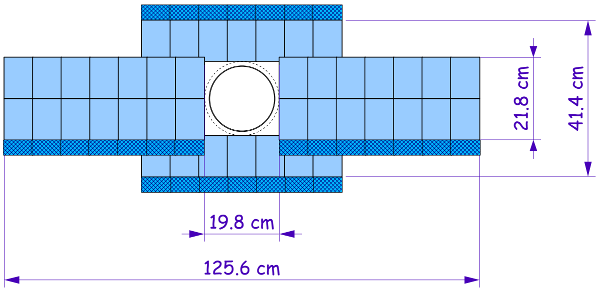


Figure 32: Schematic view of one layer of the IT [1].

active area of 597 cm × 485 cm. The OT naturally extends the IT stations to fully cover the LHCb acceptance, and preserves the same definition of the stereo angles described for the TT/IT.

The total twelve (four per station) detector layers consist of arrays of gas straw-tubes 2.4 m long and 4.9 mm in diameter. The gas is composed of a mixture of Ar (70%) and CO_2 (30%) that ensures a drift-time across the tubes below 50 ns, corresponding to the performance required by the tracking algorithm.

4.2.1.5 *Track reconstruction*

The LHCb tracking reconstruction combines information obtained from all sub-detectors to recreate the trajectory of charged particles. The reconstruction algorithm is based on a two-stage procedure; hits

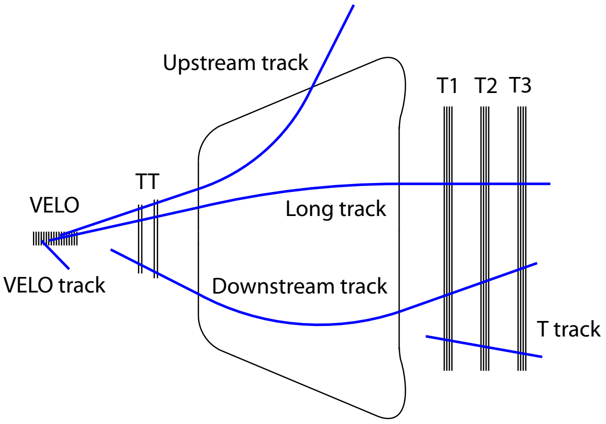


Figure 33: Illustration of the various LHCb track definitions.

are first associated to a single track by the pattern recognition software [121] and subsequently re-fitted (after removing possible *clones*, *i.e.* tracks sharing a significant subset of hits) to obtain their final reconstructed trajectory and momentum with the corresponding uncertainties.

Different pattern recognition algorithms exist in LHCb, depending on the possible track types. Generally, they are initialised by searching for straight segments connecting hits in the same sub-detector, referred to as VELO-seed or T-seed, then the track seeds are extrapolated to the other sub-detectors to look for the best compatibility with other clusters. For a detailed description of the track reconstruction algorithms see Refs. [121, 122].

Depending on their trajectory and on the parts of the detector involved in the reconstruction, tracks are divided in several categories (see Fig. 33)

i. VELO tracks - formed only with hits in the VELO detector; no momentum information is available, though an accurate precision on the primary vertex can be achieved;

ii. Upstream tracks - formed only with hits in the VELO and in the TT sub-detectors, these are typically low momentum particles that are bent out of the acceptance by the magnetic field;

iii. Long tracks - crossing the entire LHCb tracking system, these tracks provide the most precise determination of the particles' momentum and are the most useful for physics analysis;

iv. Downstream tracks - formed only with hits in the TT and in the T stations, these are typically due to long-living particles (*e.g.* K^0_S) that decay outside the VELO, hence they have worse momentum and impact parameter resolution compared to Long tracks;

v. T tracks - reconstructed only in the T stations.

4.2.2 *Performance of the Silicon Tracker*

The study presented in this section is entirely original work done during my PhD and consists of a novel approach to examine the Tracker Turicensis detector efficiency and resolution performance in an unbiased way. In order to avoid intrinsic bias in using reconstructed data events, a *masked* tracking approach is designed as follows

i. the TT layer under study is masked, *i.e.* hits information are stored but not used for the track reconstruction;

ii. track reconstruction algorithms are run on the hits collected by the remaining three TT layers and by the other tracking stations;

iii. the expected track coordinates are extrapolated from the track fit to the position of the masked layer;

iv. a permutation of the masked layer is executed in order to obtain the desired unbiased track reconstruction for all the four layers of the TT.

In order to ensure the high quality of the tracks and to reduce the contamination from "fake" tracks, as clones or *ghosts*,[2] the following selection is applied in Long tracks

- $p > 10000$ MeV$/c$,
- $p_{\rm T} > 500$ MeV$/c$,
- $\chi^2_{\rm track}/{\rm ndf} < 2$,
- $\mathcal{P}_{\rm ghost} < 0.1$,

where the track χ^2 per degrees of freedom, $\chi^2_{\rm track}/{\rm ndf}$, quantifies the quality of a track fit and the ghost probability, $\mathcal{P}_{\rm ghost}$, is obtained from a multivariate classifier that separates ghosts from good tracks combining information from different stages of the track reconstruction and from global event properties [123].

The distance between the observed hit and the extrapolated track position, defined as *unbiased residuals*, reveals information on the resolution and alignment of each silicon sensor. Figure 34 shows the obtained distributions for the three years of the Run-II, grouped by read-out configuration, *i.e.* one-, two-, three- and four-sensors, respectively. The single hit resolution, identified as the width of the residual

---

[2] Ghosts are tracks formed by combinations of hits that do not correspond to the passage of particles.

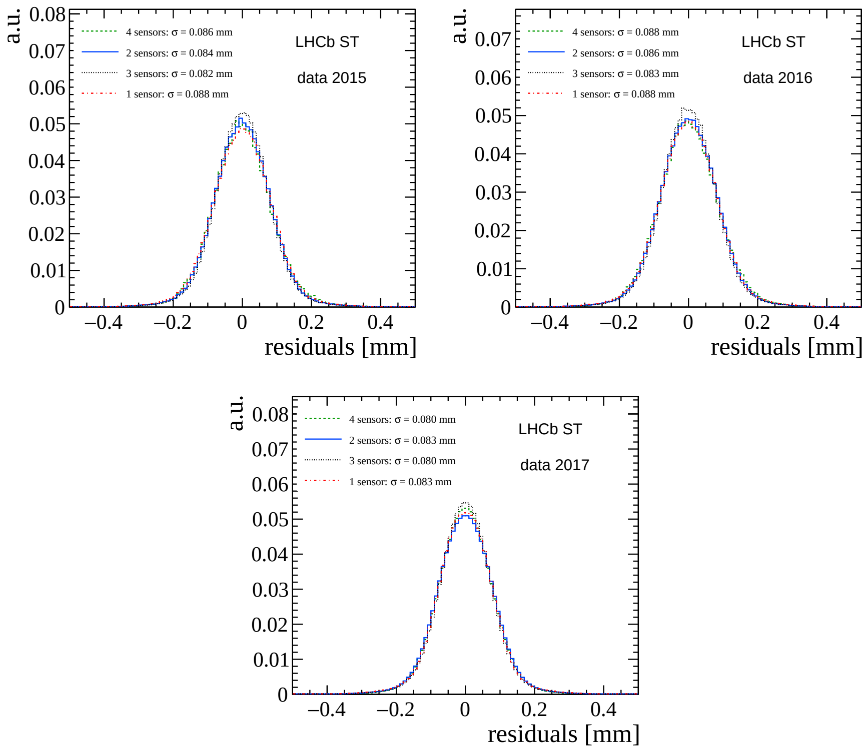


Figure 34: Residual distributions for the three years of the Run-II. Different colours correspond to the different TT read-out configurations.

distribution, results in approximately 80 μm and it is found to be stable during the years. Figure 35 also shows the width of the obtained residual distribution for each sector of the four layers of the TT. An improvement of the resolution is noticeable moving towards the outer region, due to the presence of tracks that are more inclined with respect to the x-axis, resulting in clusters that are typically made of a higher number of strips. Larger clusters allow a better resolution by averaging the information on the fraction of total charge deposited on each strip. The distribution of the cluster size, defined as the number of strips that received a deposit of charge by the passage of the particle, is shown in Fig. 36 separately for the different read-out systems. Sectors with one and two sensors, situated in the central region of the detector, have a higher probability to have a single-strip cluster. Figure 36 also shows the measured signal-to-noise distributions, defined as the ratio between the measured signal amplitude and the rms of the noise distribution in that sector, which are found to be the highest for the single-sensor sectors around the beam pipe.

Furthermore, a measurement of the misalignment of the detector can be obtained from the bias of the residual distributions, defined as the shift of the observed mean from zero. Figure 37 shows the value of this bias obtained for each TT sector. The precision of the alignment

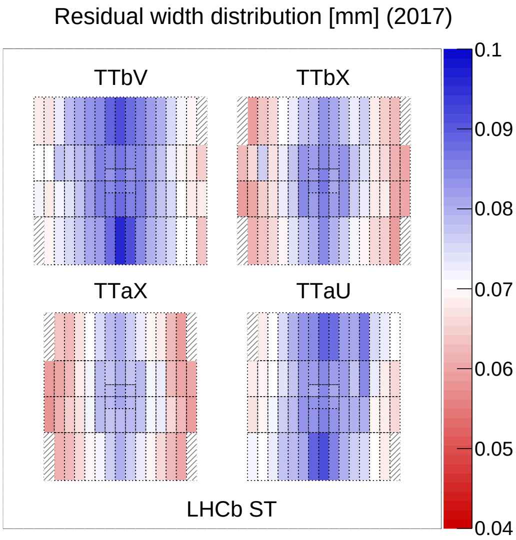


Figure 35: Width of the residual distribution obtained for each TT sector for 2017 pp collision data. Shaded sectors have limited statistics. Note that the masked algorithm must run on raw data and is computationally highly demanding.

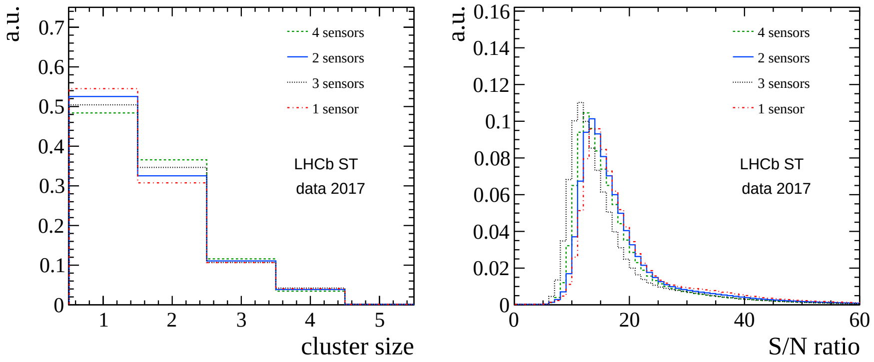


Figure 36: Cluster size (left) and the signal/noise (right) distributions for the different TT read-out configurations.

procedure is given by the RMS of these distributions and is found to be around 7 μm.

Finally, an evaluation of the efficiency of the TT can be obtained by studying the presence of hits in a determined window around the position of the reconstructed track trajectory. Figure 38 shows the map of the efficiency for the four layers of the TT corresponding to a window search of 240 μm, equivalent to approximately three times the hit resolution. The obtained efficiency exceed the 97% for all the TT sectors, resulting in an overall average of around 99.5%.

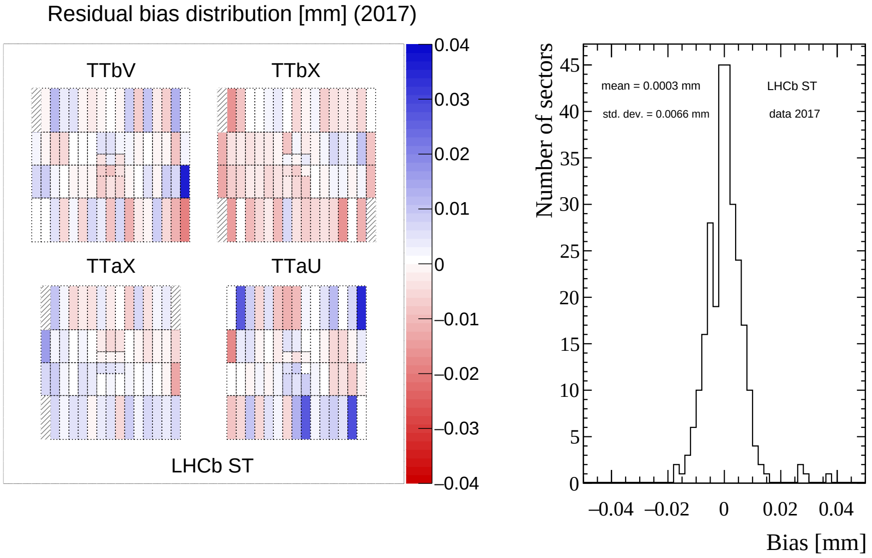


Figure 37: Mean of the residual distribution shown per TT sector (left) and summarised in a single histogram (right) for 2017 pp collision data.

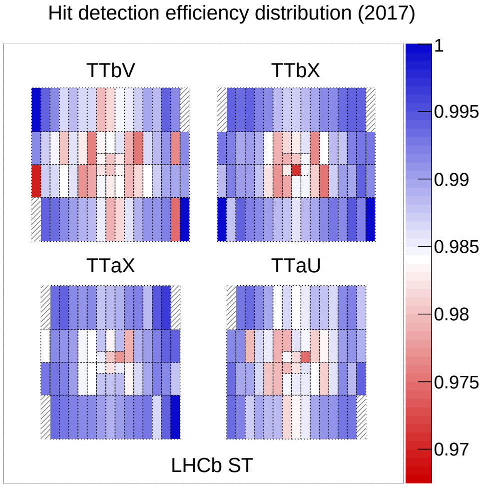


Figure 38: TT efficiency divided per sector for 2017 pp collision data.

4.2.3 *Particle identification*

Particle identification plays a fundamental role in the study of heavy flavour decays. In fact, rejecting backgrounds at the level required for the precision of the LHCb physics program often relies on the ability to distinguish the different species of particles produced in the collision.

An accurate separation between the variety of charged and neutral particles in their wide momentum range is accomplished by combining information from different sub-detectors: Ring Imaging Cherenkov (RICH) detectors are able to separate the different charged hadrons (*i.e.* K, π, p), the calorimeter system allows the identification of electrons, photons and hadrons, while the muon stations are dedicated to the detection of muons.

4.2.3.1 *Ring Imaging Cherenkov detectors*

A charged particle travelling faster than the velocity of light in a medium emits a cone of light, known as Cherenkov radiation. Ring Imaging Cherenkov (RICH) detectors use the characteristic angle of the emitted radiation

$$\cos\theta_C = \frac{1}{n\beta}, \tag{53}$$

where n is the refractive index of the medium and β is the ratio between the particle velocity and the speed of light in vacuum, to detect and identify charged particles.

Two RICH detectors are present in LHCb [124]; RICH-1 is placed between the VELO and the TT and makes use of two radiators, an aerogel[3] (SO_2) with $n = 1.03$ and C_4F_{10} with $n = 1.0014$, while the RICH-2 stands downstream of the T-stations and operates with CF_4 ($n = 1.0005$). The different refractive indices guarantees an extended momentum operational range, between 2 and 60 GeV/c for RICH-1 and between 15 and 100 GeV/c for RICH-2. The layout and the discriminating power of the RICH-1 is shown in Fig. 39.

Finally, combining information on the momentum of the particle (from the track reconstruction) and its velocity β (from the Cherenkov angle) the rest-mass of the particle can be inferred. The possible mass hypotheses are compared to the measurements and different probabilities are assigned to the reconstructed particle to be identified as each given type (*e.g.* pion, kaon, proton, etc.).

4.2.3.2 *Calorimeters*

The calorimeter system is dedicated to obtain information on the energy and position of electrons, photons and hadrons and plays an important role in the hardware stage of the trigger. The basic principle of calorimetry stands in the conversion of the energy-loss of particles in the scintillator to light. An adequate number of radiation (interaction) lengths guarantees that most of the energy of the incident particles is released giving origin to a showering process, *i.e.* production of secondary particles in cascade. The energy deposit of the entire shower is then reconstructed by the detection of the emitted photons by photomultipliers.

---

[3] The aerogel was used only during Run-I and has been remove in Long-Shutdown-I.

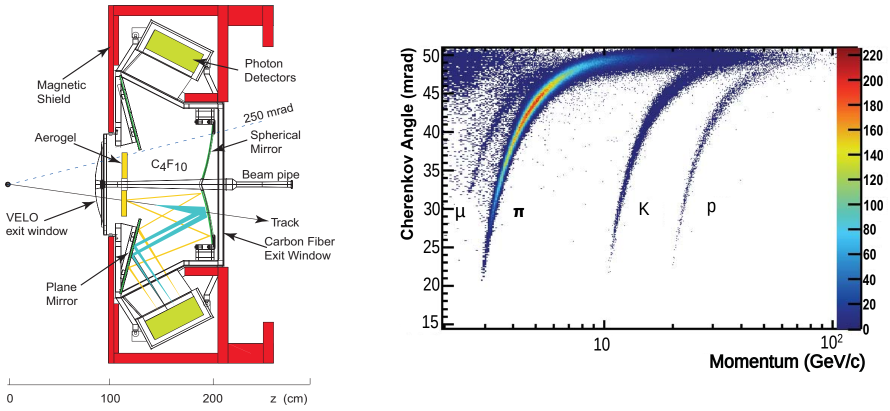


Figure 39: Left: side view schematic layout of the RICH-1 detector. Right: Reconstructed Cherenkov angle for isolated tracks, as a function of track momentum; the Cherenkov bands for muons, pions, kaons and protons are clearly visible [124].

The LHCb calorimeter system consists of three parts

i. The Scintillating Pad (SPD) and Preshower (PS) detectors [125] - the former assists the identification of charged particles from neutral ones (mainly electron from photons) before entering the calorimeters, the latter completes electron/hadron separation thanks to the different shower lengths of electrons and pions. The SPD/PS combined consist of a 15 mm thick absorber, corresponding to ~ 2.5 radiation lengths.

ii. The Electromagnetic Calorimeter (ECAL) [126] - designed as a *sampling* calorimeter,[4] is composed of 66 layers of 2 mm of lead alternated by planes of scintillating pads of 4 mm thickness, orientated perpendicularly to the beam direction. The resulting total of 25 radiation lengths allows to fully contain the electromagnetic cascade of photons and electrons, providing a measurements of their energy. The granularity of the calorimeter cells are adapted to the non-uniform detector occupancy, with increasing cell size for the inner, middle and outer regions.

iii. The Hadronic Calorimeter (HCAL) [127] - consisting of a sampling calorimeter but, differently from the ECAL, the 26 layers of iron and scintillating tiles are oriented parallel to the beam axis. With a total length of 19.7 cm, corresponding to 5.6 interaction lengths, the HCAL provides a measurement of the energy of the incoming hadrons.

---

[4] The name sampling calorimeter refers to a calorimetry technique in which the material that induces the particle shower is distinct, and typically alternated, to the scintillating material. This allows to produce a cascade that evolves quickly in a limited space, though, a disadvantage is that part of the energy is lost in the absorber material and the total energy of the cascade must be estimated.

The obtained energy resolution is found to be [128]

$$\frac{\sigma_E}{E} \simeq \begin{cases} \frac{9\%}{\sqrt{E(\,\mathrm{GeV})}} \oplus 0.8\% & \text{for the ECAL and} \\ \frac{69\%}{\sqrt{E(\,\mathrm{GeV})}} \oplus 10\% & \text{for the HCAL}\,, \end{cases} \tag{54}$$

where the first term is related to stochastic effect in shower development and the second one is due to instrumental resolution.

4.2.3.3 *The muons system*

Decays of B mesons with a muon pair in the final state play a leading role in the entirety of this dissertation, therefore, accurate muon reconstruction and identification are of fundamental importance. Energetic muons are able to pass through several meters of material loosing only a small fraction of their energy. Particle physics experiments take advantage of this unique property to design the layout of sub-detectors devoted to muon identification. Typically, muon detectors are placed as further point of the experiment where only muons can survive.[5]

The muon system [129] consists of five stations (M1-M5) placed along the beam axis as shown in Fig. 40. The inner and outer angular acceptances of the muon system are 20 (16) mrad and 306 (258) mrad in the bending (non-bending) plane, similar to that of the tracking system. The first station is positioned upstream of the calorimeters in order to avoid multiple scattering distortion, while the remaining four are placed after the HCAL each interleaved by 80 cm of iron, acting as absorber. Each station is equipped with 276 multi-wire proportional chambers with the exception of the inner part of M1, subject to the highest radiation, which is equipped with 12 gas electron multiplier detectors [130]. Each station is divided into four regions with increasing distance from the beam axis; the linear dimension and segmentation of these regions scale with the ratio $1 : 2 : 4 : 8$ (see Fig. 40).

In order to traverse the whole LHCb detector, muons must have a minimum momentum of approximately 6 GeV$/c$. For momentum regimes above 3 GeV$/c$, average muon identification efficiencies of 98% can be achieved for pion and kaon misidentification rates below 1% [131]. Moreover, including information from other PID subdetectors, hadrons misidentification probabilities can be reduced to below 0.6%, preserving a total muon efficiency above 93%. In this regime, about 60% of the misidentified pions and kaons are due to decays in flight.

Besides particle identification, the muons system provides fast-response transverse momentum measurement used for trigger purpose. Standalone muon track reconstruction can be achieved by re-

---

[5] Neutrinos are also able to travel through material but they are hardly detected due to their weak-interaction with matter.

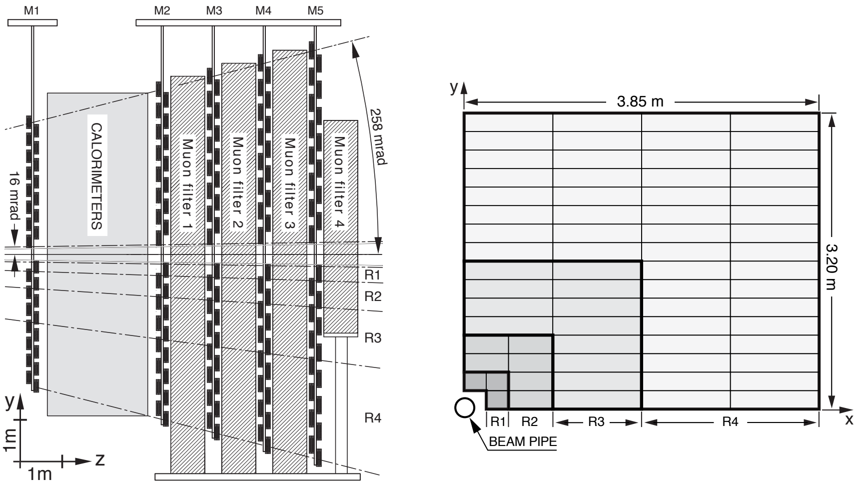


Figure 40: Left: side view of the LHCb Muon Detector [129]. Right: geometry of a quadrant, with each rectangle representing one chamber [129].

quiring to find hits in all the five stations, from this and the position of the interaction point the transverse momentum of the candidate can be estimated. Each muon station is designed to perform with an efficiency above 99% in a 20 ns time window, which guarantees five-fold coincidences among all the stations with an efficiency above 95% [129]. Stations M1, M2 and M3 provide the largest contribution to the determination of the p_T of the muon candidate, therefore, these have a higher x granularity with respect to stations M4 and M5, whose main purpose is the identification of penetrating particles.

4.2.4 *Trigger system*

At the nominal LHC conditions, proton bunches collide at a frequency up to 40 MHz, which corresponds to nearly 10 MHz where at least one visible primary pp interaction occurs at LHCb. The rate of $b\bar{b}$ pairs produced is approximately 100 kHz, of which only 15% decay within the angular acceptance of LHCb and only a tiny fraction of these contains interesting b-hadron decays. Therefore, considering the enormous collision rate and the limited capability to store events for off-line data analysis, the *trigger* system is required to identify events of interest and reduce the rate to manageable levels to be stored on disk.

The LHCb trigger is organised in a two-stage system [132]

i. The hardware-based Level-0 trigger (L0) - using only partial information from the detector, it is designed to reduce the event rate to 1.1 MHz, which is the maximum rate at which the full detector can be read out. The decision is based on the kinematics of b-hadron decays, in which decay products are typically produced with large momentum and transverse energy (E_T). Two sub-detectors contribute to the L0 trigger; the calorimeters select high E_T clusters that, combined with information on the number of hits in the PS and SPD, define three types of trigger candidates: photon, electron and hadrons. The muon stations provide a coarse estimate of the transverse momentum of the muon candidate, selecting high transverse momentum single-/di-muon events.

ii. The software-based High-Level Trigger (Hlt) - it consists of several trigger selections designed to collect specific events based on more complex kinematic, topological and particle identification requirements. There are two levels of software trigger; the first stage Hlt1 makes use of a partial event reconstruction and, thanks to the track information in the VELO, includes in the selection criteria primary vertex and impact parameter determination and qualities. During the second stage Hlt2, a complete event reconstruction is possible and more refined methods are available, such as secondary vertices, decay length and mass reconstruction, resulting in a great number of exclusive selections. The Hlt reduces the event rate up to 12.5 kHz, at this rate the selected events are permanently stored for physics analysis.

5

$B \rightarrow K^{+}(K^{*})\mu^{+}\mu^{-}$ GENERALITIES

In the remaining part of this dissertation two studies performed at the LHCb experiment are presented. Both works analyse $b \rightarrow s\mu\mu$ transitions, but in two different perspectives; a direct search for a light di-muon resonance in $B^{\pm} \rightarrow K^{\pm}\mu^{+}\mu^{-}$ decays and the measurement of the Wilson coefficients in $B^{0} \rightarrow K^{*0}\mu^{+}\mu^{-}$ decays as indirect search for New Physics. Both decay channels are ruled by the same flavour changing neutral current transition, where the only difference stands in the spectator quark, as shown in Fig. 41. Most of the experimental work to select and analyse the two modes is common; with final states either reconstructed with three or four charged tracks. Note that the neutral meson K^{*0} is reconstructed via the decay $K^{*0} \rightarrow K^{+}\pi^{-}$.

This chapter describes the main experimental techniques used to analyse $B^{\pm}(B^{0}) \rightarrow K^{\pm}(K^{*0})\mu^{+}\mu^{-}$ decays, while chapters 6 and 7 present the distinct strategies and results of the two studies.

5.1 DATASETS

At the time of this dissertation, the LHC has almost completed the Run-II of its program, delivering pp collisions at different centre-of-mass energies. The datasets collected by the LHCb experiment and the corresponding luminosities are reported in Table 9. During the Long-Shutdown between Run-I and Run-II, LHCb modified part of its software (*e.g.* reconstruction, trigger, variable definitions) to be

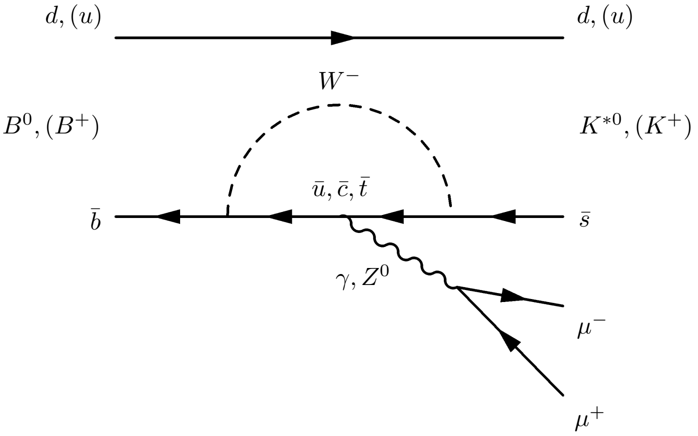


Figure 41: Feynman diagram of the flavour changing neutral current processes $B^{0}(B^{\pm}) \rightarrow K^{*0}(K^{\pm})\mu^{+}\mu^{-}$.

Table 9: Luminosity recorded by the LHCb experiment per year.

	Year	centre-of-mass energy [TeV]	Luminosity [fb^{-1}]
Run-I	2011	7	1.01
	2012	8	1.99
Run-II	2015	13	0.33
	2016	13	1.67
	2017	13	1.71

adapted to the upgraded collision conditions. In the following, when relevant, the differences between the treatment of Run-I and Run-II data are highlighted.

5.2 TRIGGER

The presence of (two) muons in the final state rules the trigger selection. In order to ensure an accurate modelling of the signal process in the simulation, only candidates triggered by tracks associated to signal decay products are retained for the analysis. The different beam energy conditions between Run-I and Run-II require distinguish trigger requirements for the two datasets.

The trigger lines used for the selection of the $B^0 \to K^{*0}\mu^+\mu^-$ candidates during the two data-taking periods are summarised in Table 10. At the hardware stage, the `L0Muon` and `L0DiMuon` lines require a minimum value for the transverse momentum and for the product of the two highest transverse momenta of the muon candidate(s), respectively. In addition, a requirement on the number of hits in the SPD is also applied to remove events with high charged-particle multiplicity, which would occupy a excessive fraction of the HLT processing time. At the Hlt1 level, the `Hlt1TrackAllL0` and `Hlt1TrackMuon` lines require at least one good-quality track with high momentum, transverse momentum and impact parameter χ^2 [1] [133], where these requirements are relaxed [134] in the case the track is associated to a muon candidate; for the Run-II the `Hlt1TrackAllL0` line has been replaced by the `Hlt1TrackMVA` and `Hlt1TwoTrackMVA` lines, where the transverse momentum and impact parameter information is exploited using a multivariate approach to select events with at least one and two good-quality tracks, respectively. Moreover, the `Hlt1DiMuonLowMass` line applies a requirement on the impact parameter of the two muon

---

[1] The impact parameter χ^2 represents the minimum χ^2 distance of a particle's trajectory from PVs. An accurate definition of the variables employed in the selection of signal candidates is given in the next section.

Table 10: Trigger requirements for the analysis of $B^0 \to K^{*0}\mu^+\mu^-$ decays.

Level	Run-I	Run-II
L0	`L0Muon`	`L0Muon` or `L0DiMuon`
Hlt1	`Hlt1TrackMuon` or `Hlt1TrackAllL0`	`Hlt1TrackMuon` or `Hlt1(Two)TrackMVA`
Hlt2	`Hlt2SingleMuon` or `Hlt2DiMuonDetached` or `Hlt2Topo(Mu)[2,3,4]BodyBBDT`	`Hlt2SingleMuon` or `Hlt2Topo(Mu)[2,3,4]Body`

candidates. Finally, at the Hlt2 level, the `Hlt2SingleMuon` line selects a clean muon candidate[2] with very good track quality, high p_T and very large impact parameter and impact parameter χ^2 [134]. The `Hlt2DiMuonDetached` line selects di-muon candidates with invariant mass below the J/ψ, significant impact parameter χ^2 and primary vertex separation [134], while the `Hlt2Topo(Mu)[2,3,4]BodyBBDT` lines require a different number of good-quality tracks (and/or muon candidates) with significant impact parameter χ^2 and compatible with the hypothesis of n-body b-hadron decays ($n = 2, 3, 4$) [135]. These topological lines pursue an inclusive selection, aiming to collect events in which only a subset of the decay products is reconstructed, and are built based on a multivariate approach (Bonsai Boosted Decision Tree [136]) which has been optimised separately for the Run-I and Run-II conditions [137].

With respect to Table 10, small differences are considered in the treatment of $B^+ \to K^+\mu^+\mu^-$ decays during Run-I: candidates that fire the `L0DiMuon` and `Hlt1DiMuonLowMass` lines are also retained, while at the Hlt2 level only candidates that fire the `Hlt2TopoMuBodyBBDT` line are used for the analysis. All the remaining topological `Hlt2` trigger lines provide a negligible contribution to the signal efficiency while increasing the combinatorial background of 4% and are therefore excluded.

The exact requirements of the Hlt trigger lines can slightly change depending on the data-taking conditions, even within the same Run/year, and are detailed in Ref. [138].

5.3 PRESELECTION

The selection of signal events occurs though several steps. Firstly, the size of the stored data is reduced by the introduction of a *stripping* framework. In this environment a set of requirements is applied

---

[2] Due to the enormous rate of single muon events, a prescaling factor of 2 is applied to the `Hlt2SingleMuon` line.

Table 11: List of requirements made on the candidates at stripping level for the `B2KX2MuMuDarkBoson` and `B2XMuMu` lines.

Particle	Variable	`B2KX2MuMuDarkBoson`	`B2XMuMu`
B meson	m	[4800, 5800] MeV/c^2	[4600, 7000] MeV/c^2
	p_{T}	$>$ 1000 MeV/c	-
	τ	$>$ 0.2 ps	-
	χ^2_{FD}	-	$>$ 121
	$\cos\theta_{\mathrm{dir}}$	$>$ 0	-
	θ_{dir}	-	$<$ 14 mrad
	$\chi^2_{\mathrm{vtx}}/\mathrm{ndf}$	$<$ 25	$<$ 8
	χ^2_{IP}	$<$ 50	$<$ 16
tracks	$\mathcal{P}_{\mathrm{ghost}}$	$<$ 0.3	$<$ 0.4
	χ^2_{IP}	$>$ 9	$>$ 9
	$\chi^2_{\mathrm{track}}/\mathrm{ndf}$	$<$ 3	-
K^+	p	$>$ 2000 MeV/c	-
	p_{T}	$>$ 250 MeV/c	-
	ProbNN_K	$>$ 0.1	-
K^{*0}	$m(K^+\pi^-)$	-	$<$ 6200 MeV/c^2
	$\chi^2_{\mathrm{vtx}}/\mathrm{ndf}$	-	$<$ 9
	χ^2_{FD} (wrt. PV)	-	$>$ 9
di-μ	m	$<$ 5000 MeV/c^2	$<$ 7100 MeV/c^2
	p_{T}	$>$ 250 MeV/c	-
	`doca`	$<$ 0.2 mm	-
	χ^2_{doca}	$<$ 25	-
	$\cos\theta_{\mathrm{dir}}$	$>$ 0	-
	χ^2_{FD} (wrt. PV)	$>$ 25	-
	$\chi^2_{\mathrm{vtx}}/\mathrm{ndf}$	$<$ 10	$<$ 9
μ	`isMuon`	`True`	`True`
	p_{T}	$>$ 100 MeV/c	-
	DLL_μ	> -5	> -3

prior to the data manipulation by the user. These algorithms, know as Stripping lines, are tuned according to the type of decays studied by each analysis. Two lines have been used for the decays of interest in this dissertation, the so-called `B2KX2MuMuDarkBoson` line for the selection of $B^+ \to K^+\chi(\to \mu^+\mu^-)$ candidates, and the `B2XMuMu` line for $B^0 \to K^{*0}\mu^+\mu^-$ decays. Besides the natural differences between the two stripping lines due to the selection of the K^+ meson in the former and the $K^{*0} \to K^+\pi^-$ system in the latter, in the `B2KX2MuMuDarkBoson` line particular attention is dedicated to the di-muon system. This stripping line, in fact, has been specifically designed for searches of new particles produced in B-meson decays and decaying into muon pairs;

requirements on the impact parameter and B-decay vertex reconstruction have been relaxed in order to select more efficiently events with a displaced di-muon vertex and, consequently, lower B-decay vertex reconstruction quality. Oppositely, the `B2XMuMu` line aims to select generic b-hadron decays with a well reconstructed decay vertex formed by a hadron and a pair of muons in the final state.

Since the stripping selection is performed off-line, the complete set of variables from the fully reconstructed event can be used. Table 11 presents the list of requirements included in the two considered stripping selections. These make use of variables involving topological criteria and fit-quality requirements defined as follows

i. Direction angle (θ_{dir}) - the angle between the momentum of a particle (the B meson in this case) and the direction of flight from the best PV to the candidate's decay vertex. A minimum requirement on $\cos\theta_{\mathrm{dir}}$ ensures the consistency of the event reconstruction.

ii. Flight distance χ^2 (χ^2_{FD}) - χ^2 separation of a particle's decay vertex from PVs. This variable is defined as the increase in χ^2 of the PV vertex fit when particles from the SV are added into the PV. A minimum requirement on the B candidate and/or subsystems of the decay (*e.g.* K^{*0}, di-μ) ensures the reconstructed event to be consistent with a B meson decay.

iii. Impact parameter χ^2 (χ^2_{IP}) - the minimum χ^2 distance of a particle's trajectory from PVs. This variable is defined as the increase in χ^2 of the PV vertex fit when the particle is added into the vertex. Similarly to χ^2_{FD}, a minimum requirement is typically applied to final states particles to reject tracks originating from PVs; oppositely, when applied to B candidates, a maximum cut ensures B mesons to be produced in the collision.

iv. Vertex χ^2 per degrees of freedom ($\chi^2_{\mathrm{vtx}}/\mathrm{ndf}$) - the χ^2 of the hypothesis of the tracks originating from the same vertex, divided by the number of degrees of freedom of the vertex. A maximum requirement on the B meson and/or subsystems of the decay guarantees high-quality vertex reconstruction candidates.

v. Distance of closest approach (`doca`) - a maximum requirement is typically applied to form a 2-body object.

vi. Distance of closest approach χ^2 (χ^2_{doca}) - `doca` significance cut, it is defined as the χ^2 of the hypothesis of the two tracks originating from the same vertex.

Furthermore, final state particles are required to be compatible with their correct PID hypothesis. The following particle-identification variables are commonly used

i. `isMuon` - this condition requires the track to have at least one hit in a subset of the muon stations;

ii. DLL_a - corresponds to the difference of log-likelihood between the hypothesis to be a particle of type a and a pion. This variable combines information from the RICH detectors, calorimeters and muon system.

iii. ProbNN - is the output of a multivariate classifier (Artificial Neural Network) that combines the information from the tracking and different particle identification systems resulting in a single probability for a given PID hypothesis.

5.4 EVENTS SIMULATIONS

Large samples of simulated events are often required in several aspects of the analysis. Simulations are used to determine selection efficiencies, to model events distributions and to investigate possible sources of background. Therefore, it is of great importance for simulated events to well reproduce the true underlying physics and detector response.

In the simulation, pp collisions are generated using Pythia [139] with a specific LHCb configuration [140]. Decays of hadronic particles are described by EvtGen [141], in which final-state radiation is generated using Photos [142]. The interaction of the generated particles with the detector, and its response, are implemented using the Geant4 toolkit [143] as described in Ref. [144].

5.4.1 *Simulations corrections*

Within all the LHCb analyses, a big effort is made to validate and to ensure the reliability of the simulated samples. The level of consistency between data and simulations is studied using calibration channels, for which the distributions of the variables of interest are extracted from background-subtracted data[3] and, in case of disagreements, simulations are corrected with data-driven approaches.

Two typologies of discrepancies are identified, originating from a mismodelling of the kinematics of the event, as well as from a mis-modelling of the detector response. Regarding the first category, three kinematic variables are found not to be described well in simulations, the detector track occupancy, the transverse momentum of the B meson and the B-meson decay vertex $\chi^2_{\mathrm{vtx}}/\mathrm{ndf}$. These distributions are reweighted in all simulated datasets accordingly to the distributions observed in the calibration sample. In the following,

---

[3] The *sPlot* technique [145] is one of the most popular tool to statistically to resolve different fit components. This approach is based on a discriminating variable, in this case the reconstructed invariant mass of the B candidate, being uncorrelated with the set of variables of interest and used to unfolding signal/background structures.

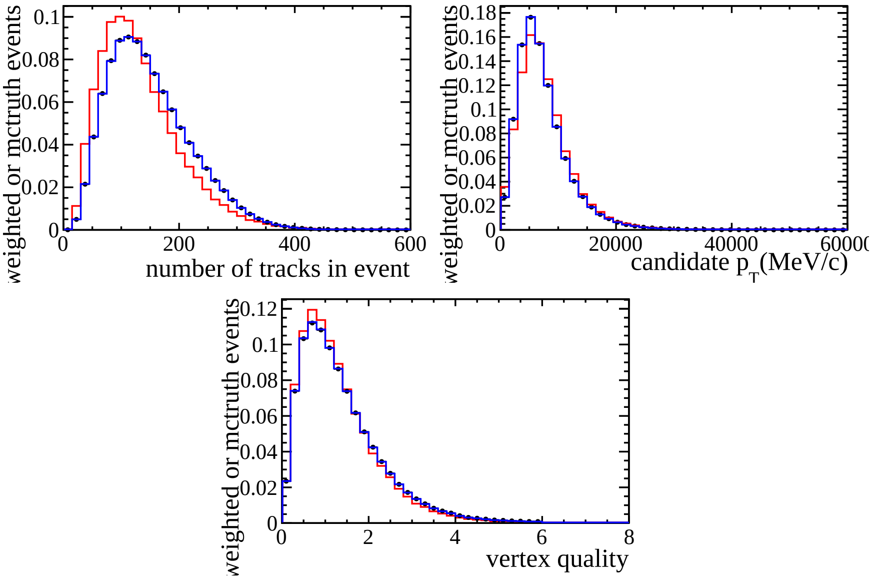


Figure 42: Distribution of the (left) number of tracks, (right) p_T and (centre) $\chi^2_{\text{vtx}}/\text{ndf}$, before (red) and after (blue) the reweighting procedure for simulated $B^0 \to J/\psi(1S)K^{*0}$ decays together with background-subtracted data (black points) [146].

decays of $B^+(B^0) \to J/\psi(1S)K^+(K^{*0})$, with the $J/\psi(1S)$ decaying to a pair of muons, are used as control channel. Event-by-event weights are derived sequentially due to potential small correlations between variables, with the $n-1$ weight being applied before deriving the n^{th} weight. Figure 42 shows the distributions of the three considered variables before and after the reweighting procedure and their consistency with the background-subtracted data after the described method.

Furthermore, the simulation of the detector response fails to reproduce particle identification variables, particularly due to disagreements in the detector occupancy and non-trivial operation conditions not modelled in the simulation. The `PIDCalib` package is the most widely used tool within LHCb to achieve the necessary PID corrections. This package makes use of clean and high-statistics calibration samples such as $D^{*+} \to D^0\pi^+$, $\Lambda \to p\pi^-$ and $\Lambda_c^+ \to pK^+\pi^-$ for hadrons and $J/\psi(1S) \to \mu^+\mu^-$ for muons and provides two alternative approaches. The "classic" `MCReweighting` calculates per-event weights, based on the kinematic information of the analysed tracks (typically momentum, pseudorapidity and the event track multiplicity), to be applied to the simulated samples. Therefore, the estimated efficiency for a given PID requirement is calculated as the average of per-event probabilities to satisfy such condition. This approach is completely data-driven and has proved to be efficient in cases when simple requirements ("rectangular cuts") are placed on PID variables. On the other hand, this procedure is bound to the kinematic coverage of the calibration sample that may differ from the one of the analysed

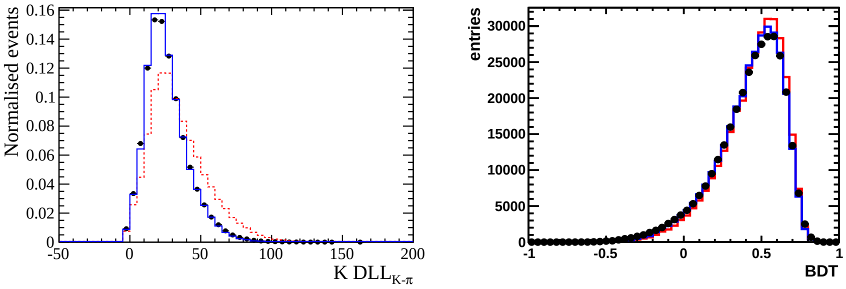


Figure 43: Comparison between simulated samples and background-subtracted data (black points) from $B^0 \to J/\psi(1S)K^{*0}$ decays [146]. Left: the red dashed line corresponds to the original $\mathrm{DLL}_K(K)$ distribution while the blue line to the resampled one. Right: destribution of the multivariate classifier for (red) simulated events with only the PID resampling applied and (blue) both the kinematic reweighting and the PID resampling.

channel. This is the case for the muons obtained from the default `PIDCalib` calibration sample, which include a kinematic cut given by $p_T > 800$ MeV/c that does not allow to inspect low di-muon invariant masses. In order to overcome this limitation, a second calibration channel is selected using $D_s \to \phi(\mu^+\mu^-)\pi$ decays and a data-driven PID efficiency determination is achieved with a `cut&fit` approach, *i.e.* the efficiency is estimated from the signal yields extracted from data with and without a given PID requirement. The analysis of $B^+ \to K^+\mu^+\mu^-$ decays makes use of both the above methods to evaluate the correct PID efficiencies.

An alternative approach is provided by the `MCResampling` tool. This method builds a *pdf* from the PID distributions of calibration data and generates random PID variables based on those *pdfs*. As before, in order to account for dependence of PID distributions on the kinematics of tracks, the calibration sample is binned in p, η and event track multiplicity. This approach, however, only preserves the correlations of the PID response with the sampled variables while breaking all the others.[4] This may have an irreducible effect on the data-simulation agreement of the multivariate response. Figure 43 shows the result of the resampling procedure and the comparison between data and simulated events for the $\mathrm{DLL}_K(K)$ variable and for the multivariate classifier obtained for $B^0 \to J/\psi(1S)K^{*0}$ candidates as discussed in the following. A good overall agreement is observed, proving the validity of the `MCResampling` approach.

---

[4] A more recent development employs multidimensional transformation of variables rather than complete resampling of the PID response [147]. In this way the correlations with other PID variables and, in general, other variables of the event are preserved.

5.5 MULTIVARIATE SELECTION

Random combinations of tracks originated from the collision can be a source of background if their reconstructed invariant mass lies in the region of interest. A large fraction of this *combinatorial background* is removed by mean of a multivariate selection.

Machine learning algorithms are built to distinguish events from different samples based on dissimilarities in the multi-dimensional distribution of a set of discriminating variables. These algorithms require a training process on two input samples, labelled as *signal* and *background* samples, while the obtained classifier response indicates the probability to be assigned to the signal or background category. The ideal working point defining the optimal background rejection and signal efficiency is then determined based on the necessity of each analysis. Typically, the performance of the classifier is studied on a sub-part of the signal (background) sample, previously excluded from the training process, referred as *test* samples.

One common example of these multivariate techniques is represented by Decision Trees algorithms [148] . Decision trees are classification methods based on a branch-like segmentation of the events. Signal and background datasets are optimally separated based on a given input variable, this process is repeated for all discriminants until a certain node has reached either a minimum number of events or a minimum/maximum signal purity. The performance of these supervised learning processes is improved by assigning increasing weights - *boosting* [149] - to misclassified events, *i.e.* iteratively modifying the sample distribution accordingly to the difficult instances. Numerous loss functions are available among the machine learning community, examples are adaptive boosting (BDT) [150] and gradient boosting (BDTG) [151], where the former is used the select $B^0 \to K^{*0}\mu^+\mu^-$ candidates and the latter is used for the analysis of $B^+ \to K^+\chi(\to \mu^+\mu^-)$ decays.

For both channels, the background sample is selected from data, using the right sideband of the B invariant mass distribution.[5] This region provides a data sample of pure combinatorial background, while the left-hand sideband is removed in order to avoid potential contamination from partially reconstructed backgrounds. In order to exploit at most the information contained in the dataset, a *k-folding* technique is applied [152]. The procedure can be summarised in the following steps

i. the considered sample is split in k parts;

---

[5] The exact definition of the sideband mass window depends on the specific channel and is defined as [5450, 5800] MeV/c^2 and [5350, 5800] MeV/c^2 for $B^+ \to K^+\mu^+\mu^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ decays, respectively.

Table 12: Input variables for the multivariate selection of $B^+(B^0) \rightarrow K^+(K^{*0})\mu^+\mu^-$ candidates.

Training variables	
$B^+ \rightarrow K^+\chi(\rightarrow \mu^+\mu^-)$	$B^0 \rightarrow K^{*0}\mu^+\mu^-$
decay-time (B^+)	decay-time (B^0)
p_{T} (B^+)	$p(B^0)$
$\cos\theta_{\mathrm{dir}}$ (B^+)	p_{T} (B^0)
χ^2_{IP} (B^+)	$\cos\theta_{\mathrm{dir}}$ (B^0)
χ^2_{vtx} $(K^+\chi)$	χ^2_{vtx} $(K\pi\mu\mu)$
χ^2_{IP} (K^+)	`BDT-iso2` (μ)
χ^2_{FD} (χ)	`track-iso` (K)
χ^2_{IP} (μ)	`track-iso` (π)
$\mu^+\mu^-$ `doca`	$\mathrm{DLL}_{\mu}(\mu)$
`BDT-iso1` (χ)	$\mathrm{DLL}_{K}(K)$
`track-iso` (μ)	$\mathrm{DLL}_{K}(\pi)$

ii. $k-1$ of these subsamples are used for the training and testing stages, while the obtained classifier is applied to the remaining part;

iii. a cyclical permutation is done in order to apply the classifier to all the subsamples.

This strategy allows to include in the training process all the events of the dataset while keeping an unbiased sample to test and to optimise the performance of the obtained classifier. The optimisation procedure aims to identify the selection to be applied on the multivariate classifier that provides the best sensitivity for each considered analysis. This operation is performed on the $k-1$ subsamples of point [ii.] above and guarantees an unbiased (*i.e.* independent on the performance on the dataset that will be employed for the analysis) optimisation procedure.

Finally, for the study of $B^0 \rightarrow K^{*0}\mu^+\mu^-$ decays, the signal sample required for the training process is selected from data and obtained from the control channel $B^0 \rightarrow J/\psi(1S)(\rightarrow \mu^+\mu^-)K^{*0}$, while for the analysis of $B^+ \rightarrow K^+\mu^+\mu^-$ decays simulated signal samples are employed.

Table 12 summarises the set of variables used as input of the multivariate classifier for the two analyses. For $B^+ \rightarrow K^+\mu^+\mu^-$ decays, these consist of kinematic properties of the B meson, topological variables of the reconstructed decay and two isolation variables. The `BDT-iso1` is the result of a multivariate classifier trained on $B^0 \rightarrow D_s^{*-}\mu^+\nu_\mu$ decays, as described in Ref. [153]. This algorithm looks at each track in the event and returns a value based on the probability for the

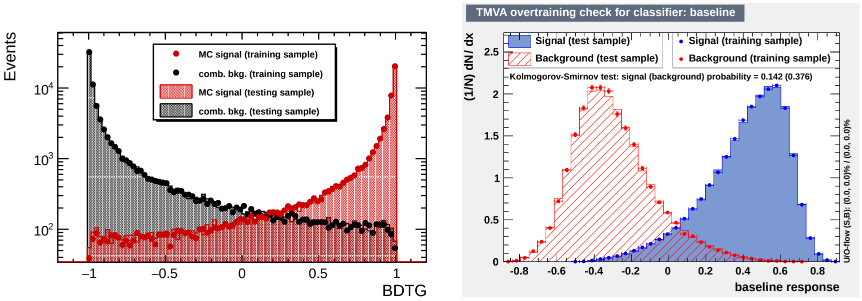


Figure 44: Multivariate classifier distributions for (left) $B^+ \to K^+\chi$ decays and (right) $B^0 \to K^{*0}\mu^+\mu^-$ decays [146]. The training and testing samples are also shown as overtraining cross-checks. Distributions correspond to one of the k folds of the training process, all the sub-samples responses are found to be statistically compatible.

track to match the reconstructed B candidate. Higher values correspond to tracks which are more likely to come from the B candidate. Typically, the track with the highest isolation value is sufficient to achieve a combinatorial background rejection equivalent to the case when the isolation information of the entire event is used. Hence, only the isolation of most discriminating track is included in the training process. The track isolation `track-iso` [154] is applied to each muon candidate. This variable identifies the number of tracks that can create a good vertex with the considered muon, after excluding the other muon track from the event. In general, both isolation variables are designed to identify tracks that are close to the signal candidate and can form a vertex with it. These are expected to be more abundant for combinatorial background than for signal events.

For the selection of $B^0 \to K^{*0}\mu^+\mu^-$ decays, in addition to the time of decay, p, p_T and $\cos\theta_{\rm dir}$ of the B^0 candidate and $K^+\pi^-\mu^+\mu^-$ vertex χ^2, the isolation of the four final state particles is included. For the hadrons, the track isolation described above is employed, while for the two muons an isolation variable based on a multivariate approach is used. The `BDT-iso2` variable[6] extends the use of the information from the reconstructed tracks (*e.g.* p_T, IP, $\chi^2_{\rm IP}$, $\chi^2_{\rm track}/{\rm ndf}$, `doca`, etc.) and employs a BDT classifier trained on a simulated signal sample and di-muon combinatorial data [146]. Finally, the following PID variables are used: DLL$_K$ of the kaon and pion and DLL$_\mu$ of the muons.

Figure 44 shows the obtained distributions of the classifier for the signal and background samples in the two described configurations. At the optimal working point, chosen to maximise the $S/\sqrt{S+B}$ figure of merit, the multivariate classifier is found to provide a background rejection of 97% and a signal efficiency of 85% for $B^0 \to K^{*0}\mu^+\mu^-$

---

[6] The "2" in the name of the `BDT-iso2` variable is used to be distinguished from the multivariate isolation variable `BDT-iso1` previously introduced for the selection of $B^+ \to K^+\chi(\to \mu^+\mu^-)$ candidates.

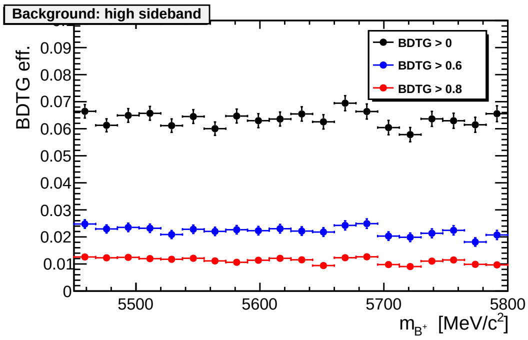


Figure 45: Combinatorial background efficiency for several selection cuts as function of the reconstructed B-meson invariant mass. Results are obtained by using the right B invariant mass sideband.

decays [79], while the optimisation of the multivariate selection for the search of $B^+ \to K^+\chi$ decays is detailed in Sec. 6.5.

Finally, in order not to create artificial distortions when applied to the signal region, the multivariate response is required to be flat as function of the reconstructed B invariant mass. Figure 45 shows the flatness of the classifier output as function of the $K^+\mu^+\mu^-$ invariant mass tested in the right-hand sideband of $B^+ \to K^+\mu^+\mu^-$ candidates. Several possible selections are shown as reference and the background rejection is found to be constant with very good accuracy.

"Verily at the first Chaos came to be, but next
wide-bosomed Gaia, the ever-sure foundations of all
the deathless ones who hold the peaks of snowy Olympus,
and dim Tartarus in the depth of the wide-pathed Earth,
and Eros, fairest among the deathless gods,
who unnerves the limbs and overcomes the mind
and wise counsels of all gods and all men within them."

— Hesiod, *Theogony*, 116-122† (ca. VIII – VII century BC)

6

DIRECT SEARCH FOR LOW MASS DI-MUON RESONANCES

This chapter presents a search for low-mass di-muon resonances in $B^{\pm} \to K^{\pm}\mu^{+}\mu^{-}$ decays and describes in detail the result published in Ref. [4]. The analysis looks for the production of a hypothetical new scalar particle, χ, via the decay $B^{+} \to K^{+}\chi$[1] and its subsequent decay into a pair of opposite charged muons $\chi \to \mu^{+}\mu^{-}$. The presented results are based on the dataset collected during the LHC Run-I, corresponding to an integrated luminosity of 3.0 fb^{-1}.

6.1 STRATEGY

A model-independent approach to search for a new particle of unknown mass and lifetime has been designed for this analysis. The method consists in performing scans on the di-muon invariant mass spectrum searching for an excess of events, and the result is presented as a 95% confidence level excluded limit, as function of the mass, m_{χ}, and lifetime, τ_{χ}, of the hypothetical particle.

6.1.1 *How to set a limit: the CLs method*

Any search for previously unobserved modes is susceptible to the absence of a significant signal. When interpreting the result of an experiment, a scientist must answer to the following questions: did I or did I not establish a discovery? And, in case of positive answer, how well does my alternative model describes this discovery? Oppositely, in the "unlucky" case of no discovery, one must wonder: to which level my alternative model is excluded? This last interrogation is equivalent

---

† Translation by Hugh G. Evelyn-White (1914).

[1] In the following, the inclusion of the charge-conjugate decay $B^{-} \to K^{-}\chi$ is always implied.

to establish an upper limit on the strength of the new phenomenon with a certain confidence.

In order to answer to these various points, it is necessary to apply a test statistic to the sample of data under the different hypotheses. It is common to identify the test hypothesis as the *signal-plus-background* scenario, whose likelihood is given by $\mathcal{L}_{s+b}$, and the null hypothesis as the *background-only* scenario, whose likelihood is given by $\mathcal{L}_b$. Thus, the likelihood ratio

$$Q = \frac{\mathcal{L}_{s+b}}{\mathcal{L}_b} \tag{55}$$

can be used to build the desired test statistic, conveniently expressed in terms of $x = -2 \ln Q$, where x follows a χ^2 distribution with one degree of freedom. This formulation allows to define the quantity

$$p = \int_{-\infty}^{x_{obs}} g(x|\mathcal{H}_0) \mathrm{d}x \,, \tag{56}$$

where $g(x|\mathcal{H}_0)$ is the distribution of the test statistic x for the null hypothesis $\mathcal{H}_0$ and x_{obs} is the value of the test statistic obtained in the experiment. This quantity, known as *p-value*, corresponds to the probability to get a result which is equally or less compatible with the null hypothesis than the measured value, *i.e.* the null hypothesis is rejected with a probability of $1-p$.

In the event of a discovery, the p-value is useful to determine the strength (or significance) of the newly observed phenomenon. On the other hand, when setting a limit on a certain process, it is convenient to define the *confidence levels* CL_b and CL_{s+b} of the background-only and signal-plus-background hypotheses, respectively. The confidence level is defined as

$$CL = \int_{x_{obs}}^{\infty} g(x|\mathcal{H}) \mathrm{d}x \,, \tag{57}$$

which, in the case of the null-hypothesis, is equivalent to $1-p$. Finally, the ratio

$$CL_s = \frac{CL_{s+b}}{CL_b} \tag{58}$$

is frequently used to generalise the investigation of the two hypotheses and to approximate the confidence in the *signal-only* hypothesis [155]. In summary, in the absence of an observation, any signal hypothesis that results in $CL_s < \alpha$ is excluded at $(1-\alpha)$ confidence level.

In order to increase the sensitivity of a search, a common technique consists in splitting the sample in categories with different signal/background ratios. In this case, the likelihood for the signal-plus-background hypothesis can be defined as

$$\mathcal{L}_{s+b} = \prod_i \mathcal{P}(n^i_{obs}|s_i + b_i) \,, \tag{59}$$

where the index i runs on the different categories, s_i and b_i are the expected number of signal and background events, respectively, and $\mathcal{P}(n_i|\lambda_i)$ is the Poissonian probability of observing n_i events in a distribution with mean λ_i. On the other hand, for the background only hypothesis the likelihood is reduced to

$$\mathcal{L}_b = \prod_i \mathcal{P}(n^i_{obs}|b_i)\,. \tag{60}$$

6.1.2 *Searching in the lifetime dimension*

Since the lifetime of the new particle can be long compared to the detector resolution, the di-muon vertex is allowed, but not required, to be displaced from the B-meson decay vertex. In order to increase the sensitivity for different lifetimes, three bins of decay-time of the χ candidate are considered. The division in the three bins reflects the following criteria

i. The first bin corresponds to small decay-times and contains events that can be identified as *prompt* decays. All the irreducible $B^+ \to K^+\mu^+\mu^-$ Standard Model decays, which are the main background of this analysis, are meant to be contained in the first bin. This region is characterised by a high signal efficiency but also by a large background contamination.

ii. The second bin contains *displaced* di-muon decays and is supposed to be most sensitive for long lifetimes; this region suffers from lower reconstruction efficiency compared to the first bin but benefits of a low background contamination.

iii. The third bin contains *very displaced* di-muon candidates and is chosen to determine a "zero background" search region.

6.1.3 *Searching in the mass dimension*

For each of the three above-mentioned decay-time bins, a search in the mass dimension is performed by scanning the di-muon invariant mass distribution in steps of $\frac{1}{2}\sigma(m)$, where $\sigma(m)$ is the local di-muon mass resolution. For each tested mass m_t, a search region is defined as

$$|m - m_t| < 2\sigma(m) \tag{61}$$

and a background region is defined as the sidebands of the search region according to

$$3\sigma(m) < |m - m_t| < (2k+3)\sigma(m); \tag{62}$$

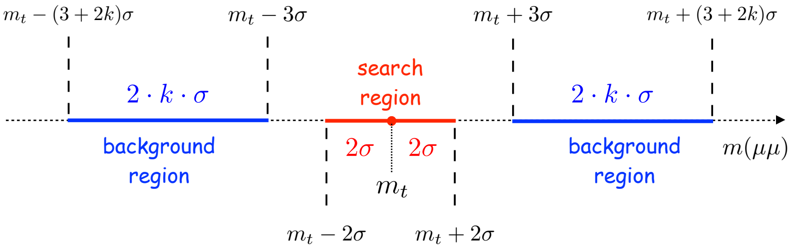


Figure 46: Illustration of the signal and background region as defined in Eqs. 61 and 62.

therefore, the parameter k is defined as the ratio between the size of the background and search regions. The definition of the search and background regions is illustrated in Fig. 46.

For each m_t a test statistic is performed; the background-plus-signal and the background-only hypotheses are compared using the CL_s method, where the information from the three bins is combined into a single likelihood. For each bin the expected number of signal events is obtained from the simulation and normalised to the observed yield in the control channel, while the expected number of background events is extracted directly from the data with a linear fit to the background region (di-muon sidebands).[2] Note that the SM di-muon resonances $\phi(1020)$, $J/\psi(1S)$, $\psi(2S)$, $\psi(3770)$ and $\psi(4160)$ would appear as signal in this approach and are therefore vetoed from the search. The choice of the parameter k is a trade between a large value of k, that would give a more precise background estimation, and the approximation of local linearity,[3] whose validity is studied with pseudoexperiments and discussed in Sec. 6.4.

6.1.4 *Mass-lifetime parameter space*

One of the critical aspects of this analysis is the coverage of the explored m_χ-τ_χ parameter space. The mass domain is naturally bounded by the kinematic threshold for di-muon production, $2m_\mu$, and by the difference between the B and K meson masses, $m_{B^+} - m_{K^+}$. Furthermore, some of the crucial features of the analysis (*e.g.* signal efficiency, resolution, etc.) are found to be dependent on the mass and lifetime of the hypothetical new particle. Thus, the study of the signal properties by mean of the simulations must carefully take into account these dependences. This is obtained by generating signal simulations with

---

[2] An exception to this approach for the background prediction is applied in the third bin, where some prior assumptions are required (see section 6.5.3).

[3] In order to be able to estimate the expected background yield from the sidebands of the search region as in Eq. 62, the background distribution must be approximately linear in the region of interest $|m_t - m| < (2k+3)\sigma(m)$.

Table 13: Signal simulated samples with different mass and lifetime hypotheses. Numbers correspond to the number of generated events per sample (in millions).

	Mass (MeV/c^2)												
	250	500	800	1000	1500	1700	1775	2000	2500	3000	3500	4000	4500
$\tau_\chi = 1$ ps				1.5									
$\tau_\chi = 10$ ps				1.6									
$\tau_\chi = 100$ ps	1.5	1.5	1.5	1.6	1.5	0.2	0.2	1.5	1.5	1.5	1.5	1.5	1.5
$\tau_\chi = 1000$ ps				1.5									

several mass and lifetime hypotheses, which configurations aim to ideally cover the explored parameter space. The optimal set of values is reported in Table 13 and largely consists of samples produced with $\tau_\chi = 100$ ps. This decision balances the reduced signal efficiency observed for long lifetimes with the robustness of the reweighting approach used to parametrise all the lifetimes of interest from a single sample produce with a given τ_χ, for a detailed discussion see Sec. 6.6.

6.2 SELECTION FOR UNBIASED m_χ-τ_χ PARAMETER SPACE

The search is performed on candidates that have the reconstructed $K^+\mu^+\mu^-$ invariant mass in a window of $\pm 50\ \mathrm{MeV}/c^2$ around the nominal B^+ mass [156], denoted as signal region. This definition correspond to a range of almost $\pm 3\sigma$ around the B-meson mass. In order to avoid to introduce any bias during the selection procedure, the analysis is performed in a *blind* way, *i.e.* all the aspects of the analysis are fixed before looking at the data in the region of interest.

In general, χ candidates can be either reconstructed as Long or Downstream tracks; the latter is typical of long decay-times, when the χ candidate decays outside the VELO. Unfortunately, no dedicated trigger lines for the selection of χ candidates built from Downstream tracks has been implemented in Run-I, as discussed in Sec. 5.2. Therefore, this analysis is limited to events formed by Long tracks.

The multivariate algorithms discussed in Sec. 5.5 relies on the use of simulated signal datasets. The set of variables used in the training process has been examined for different signal hypotheses as given in Table 13. Noticeable differences are observed predominantly with $m_\chi = 250$, 1000, 2500 and 4000 MeV/c^2 where $\tau_\chi = 100$ ps and $\tau_\chi = 1$, 10 and 100 ps where $m_\chi = 2500\ \mathrm{MeV}/c^2$. For simplicity, the classifier training process has been limited to this sub-sample of simulated signal configurations. Figure 47 shows the signal efficiency as function of the signal mass and lifetime for different considered training configurations. All efficiencies are shown for a fixed background rejection rate of 0.9. Configurations that were trained with a long

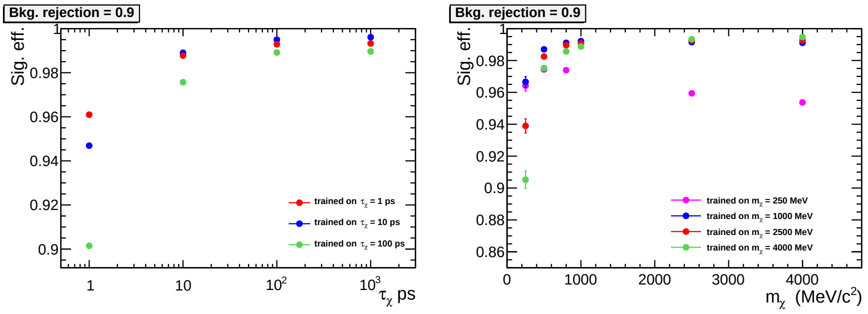


Figure 47: Signal efficiency for different (left) lifetimes and (right) masses hypotheses. The efficiencies correspond to a background rejection of 0.9 obtained with several distinct training configurations.

signal lifetime show a considerable loss of efficiency for short lifetime samples, while the configuration trained with the signal sample with $\tau_\chi = 1$ ps gives the best overall performance and is chosen as default training configuration. Similarly, training configurations with high mass show an efficiency drop at low mass, vice versa, training on low mass reduces the performance at high mass. The performance obtained with $m_\chi = 1000$ and 2500 MeV$/c^2$ is found to be very similar over most of the mass spectrum. For convenience, the sample with $m_\chi = 2500$ MeV$/c^2$ is selected as baseline classifier, since the minimal gain observed at very low mass does not justify the computationally expensive production of complete set of simulated samples.

6.3 BACKGROUND

Several sources of background can pollute the search for $B^+ \to K^+\chi$ decays and must be carefully treated. Among these, the so called *peaking backgrounds*, decays of B mesons or other hadrons that due to particle misidentification or misreconstruction of the decay can appear as a peak in the signal region, are of extreme importance. In this section, all potential sources of backgrounds are briefly discussed. These are processed with the same selection as the signal mode and their contribution is estimated using data-corrected simulations. Table 14 summarises all the cuts and vetoes applied to reject the different sources of background discussed in the following.

The main categories of background are identified as

i. Decays of B mesons to SM di-muon resonances, *i.e.* $\phi(1020)$, $J/\psi(1S)$, $\psi(2S)$, $\psi(3770)$ and $\psi(4160)$, with a large branching ratio, which decay to the same final state of interest. Events are removed applying a veto on the $\mu^+\mu^-$ mass in the ranges $985 < m_{\mu^+\mu^-} < 1055$, $2946 < m_{\mu^+\mu^-} < 3176$, $3586 < m_{\mu^+\mu^-} < 3850$ and $4103 < m_{\mu^+\mu^-} < 4270$ MeV$/c^2$. The upper edge of the $\psi(2S)$ veto falls around the $\psi(3770)$ mass, so they are included into a single vetoed region.

Table 14: List of all vetoes and cuts applied to reduce possible background contaminations.

Background	Variable	Mass region (MeV/c^2)	Cut
$\phi \to \mu^+\mu^-$	$m_{\mu^+\mu^-}$	[985, 1055]	vetoed-1st bin
$J/\psi \to \mu^+\mu^-$	$m_{\mu^+\mu^-}$	[2946, 3176]	vetoed
$\psi(2S)$, $\psi(3770) \to \mu^+\mu^-$	$m_{\mu^+\mu^-}$	[3586, 3850]	vetoed
$\psi(4160) \to \mu^+\mu^-$	$m_{\mu^+\mu^-}$	[4103, 4270]	vetoed-1st bin
$J/\psi \to \mu^+\mu^-$ with $K^+ \leftrightarrow \mu^+$	$m_{(K\mu\leftrightarrow\mu\mu)}$	[3000, 3200]	$\texttt{isMuon}(K^+) = \texttt{false}$
$B^+ \to D^0(\to K\pi)X$	$m_{(\mu\mu\leftrightarrow K\pi)}$	[1840, 1890]	$\text{ProbNN}_\mu\ (\mu) > 0.4$
$K^0_S \to \pi^+\pi^-$	$m_{(\mu\mu\leftrightarrow\pi\pi)}$	[443, 523]	vetoed
$\Lambda^0 \to p\pi$	$m_{(\mu\mu\leftrightarrow p\pi)}$	[1090, 1120]	vetoed

All these resonances are originated directly from B decays. In particular, the expected yields of $B^+ \to K^+\phi(1020)$ and $B^+ \to \psi(4160)K^+$ decays in the displaced region of the analysis (2nd and 3rd decay-time bins) are found to be negligible. In addition, no peaks corresponding to the $\phi(1020)$ and $\psi(4160)$ resonances are observed in the combinatorial background sample from the right B invariant mass sideband. Thus, the corresponding vetoes can be relaxed and they are applied only in the first bin of the search. The $B^+ \to J/\psi(1S)K^+$ and $B^+ \to \psi(2S)K^+$ decays have very large branching ratio and their tail can still populate the displaced region.

ii. Decays of $B^+ \to J/\psi K^+$ can escape the J/ψ veto if the kaon is misidentified with the same-sign muon and vice versa. This background is rejected by computing the $K^+\mu^-$ invariant mass under the $\mu^+\mu^-$ hypothesis, $m_{(K\mu\leftrightarrow\mu\mu)}$, and requiring the kaon to fail the `isMuon` flag for candidates with a mass $m_{(K\mu\leftrightarrow\mu\mu)}$ within the range $3000 < m_{(K\mu\leftrightarrow\mu\mu)} < 3200$ MeV/c^2.

iii. Possible three body hadron decays, *i.e.* $B^+ \to h^+h^-h^+$ with $h = \{K, \pi\}$, is a class of background if pions or kaons are misidentified as muons, though most of these events are reduced after particle identification criteria.

iv. The decays of $B^+ \to D^0\pi$ and $B^+ \to D^0K$, with $D^0 \to K^+\pi^-$, $D^0 \to K^+K^-$ or $D^0 \to \pi^+\pi^-$, can be particularly dangerous for the presence of a displaced secondary vertex given by the decay of the D^0 meson. The decays $B^+ \to D^0(\to KK)h$ and $B^+ \to D^0(\to \pi\pi)h$, with $h = \{K, \pi\}$, are Cabibbo suppressed and are found to be negligible, while a small contribution of $B^+ \to D^0(\to K\pi)\pi$ and $B^+ \to D^0(\to K\pi)K$ are expected in data. These backgrounds are rejected by the introduction of a more stringent muon PID requirement for candidates with invariant mass, after the correct mass hypothesis assignment $m_{(\mu\mu\leftrightarrow K\pi)}$, in a window around the D^0 mass. A cut at ProbNN$_\mu$ $(\mu) > 0.4$ is sufficient to remove these sources of background and is applied for events with $m_{(\mu\mu\leftrightarrow K\pi)} \in [1840, 1890]$ MeV$/c^2$.

v. Peaks corresponding to $K^0_S \to \pi^+\pi^-$ and $\Lambda^0 \to p\pi$ decays are observed in the right B invariant mass sideband of the dataset and removed by a veto around the K^0_S and Λ^0 mass after the correct mass hypothesis assignment. Due to the small difference in mass between a muon and a pion, the veto around the K^0_S mass results in a significant loss of efficiency for signal candidates with m_χ close to the K^0_S mass.

vi. Signal events from $B^+ \to K^+\mu^+\mu^-$ decays are an irreducible source of background in this measurement. The di-muon pair is expected to come from a prompt decay of the B meson and these events require a dedicated rejection strategy (in particular for the first bin) as discussed in Sec. 6.5.1. Unlike this dominant decay mode, other decays that contain two opposite sign muons and a pseudo-scalar or a vector meson are not expected to contribute. While for $B^+ \to \pi^+\mu^+\mu^-$ decays the low branching fraction combined with the misidentification efficiency reduces this background to a negligible level, partially reconstructed $B^0 \to K^{*0}(K^+\pi^-)\mu^+\mu^-$ events are either rejected by the B mass signal region or reduced due to isolation criteria. Finally, possible contributions from B_s and Λ_b decays are found to be negligible.

vii. Combinatorial background is the only source of background that can populate the displaced region of the analysis. Figure 48 shows the distribution of the decay-time of the di-muon system for combinatorial background events observed in the right B invariant mass sideband of data. The tail of the distribution extends up to tens of ps. Events with time of decay $t < -1$ ps are rejected, since these are not physical. The value of the cut takes into account the time resolution of the detector and it is found to reject more than 30% of the combinatorial background in the first bin of the search.

The combinatorial background yield is extrapolated from data with a fit to the right B invariant mass sideband, as shown in Fig. 49. The expected yields after the selection optimisation described in section 6.5.2 are found to be approximately 980, 170 and 2.6 events for first, second and third bins of the analysis, respectively.

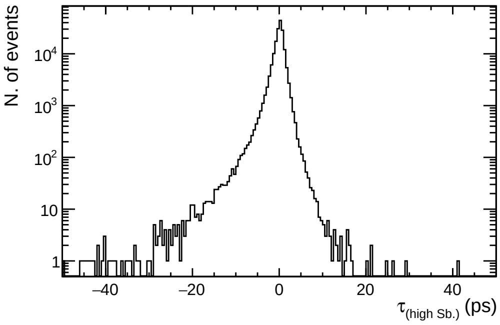


Figure 48: Distribution of time of decay for combinatorial background, events are taken from the right B invariant mass sideband.

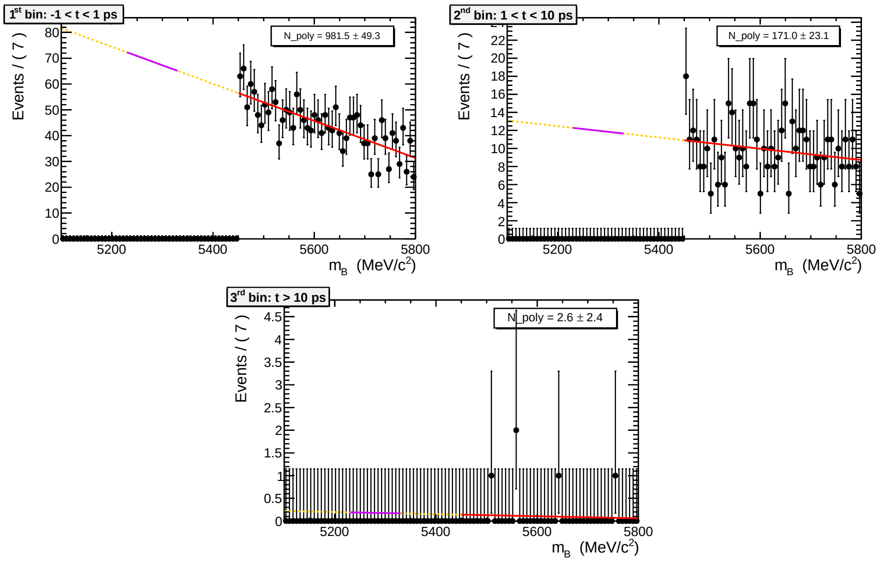


Figure 49: Fit of the right B invariant mass sideband to obtain the expected combinatorial background yield in the first (top left), second (top right) and third bin (bottom) of the analysis. The distribution is fitted with a linear function, the range marked with the solid violet line corresponds to the signal window.

6.4 STRATEGY VALIDATION

The strategy of the search introduced in section 6.1 is validated with a series of pseudoexperiments including the surviving sources of background, *i.e.* $B^+ \to K^+\mu^+\mu^-$ decays and combinatorial background. The di-muon invariant mass distribution of $B^+ \to K^+\mu^+\mu^-$ decays and combinatorial background is derived from simulations and from the right B invariant mass sideband, respectively. Figure 50 shows the obtained distributions for the two background sources; a 7^{th} order polynomial is used to parametrise the $B^+ \to K^+\mu^+\mu^-$ component, while the combinatorial background is described with a kernel density estimator. The generated ensembles are produced from these two

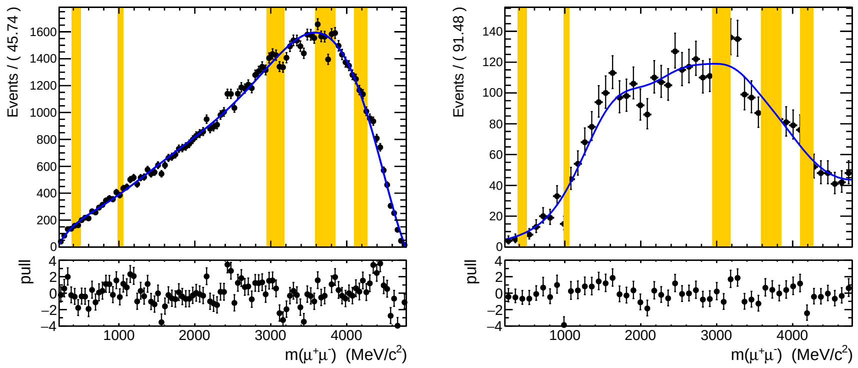


Figure 50: Di-muon distribution used to generate pseudoexperiments for (left) simulated events from $B^+ \to K^+\mu^+\mu^-$ decays and (right) data combinatorial background. The $B^+ \to K^+\mu^+\mu^-$ di-muon spectrum is fitted with a 7^{th} order polynomial, while the combinatorial background is described with a kernel density estimator. The yellow bands correspond to vetoed regions in the search.

background *pdfs* and scaled by the corresponding number of expected events.

The value of the parameter k of Eq. 62 must satisfy the approximation of local linearity in the entire mass range. The di-muon invariant mass distribution obtained in these pseudoexperiments is scanned as described in section 6.1.3 and, for each tested mass, the pull of the deviation between the observed and the expected number of background events is built. The approximation of local-linearity is better satisfied as the pulls become closer to a standard Gaussian. Among the tested hypotheses, values between $k = 5$ and 10 provide very similar performances, while higher values badly satisfies the local linearity approximation. Oppositely, smaller values are found to be too sensitive to statistical fluctuation. Thus, the choice $k = 10$ guarantees at the same time the validity to the local linearity approximation and gives more stability in the fit. For this value, the mean of the pulls obtained from pseudoexperiments are found to differ from zero for less than 0.08 standard deviations for all the tested masses. A systematic uncertainty is then assigned to the background prediction in order to take this discrepancy into account.

6.5 OPTIMISATION OF THE SEARCH

The search is optimised in order to reach the best sensitivity in the entire m_χ-τ_χ plane. The five parameters that complete the selection are given by the two decay-time bin boundaries, named t_1 and t_2 and defined as

- 1^{st} bin: $|t| < t_1$,
- 2^{nd} bin: $t_1 < t < t_2$,

  - 3rd bin: $t > t_2$,

and the selection cuts on the BDTG response in the three bins.

The optimisation proceeds through a sequential approach, based on the following steps. Firstly, a simplified *only-two-bins* scenario is considered and the value of t_1 is defined dividing the *prompt* from the *displaced* region. The selection in the 1st and 2nd bins is optimised under these conditions. Thereafter, the *displaced* region is split in two bins by determining the value of t_2 and, ultimately, the selection in the 3rd bin is established. This sequential approach is possible since the optimisation of each bin is mostly independent on the other two.

6.5.1 *Prompt-displaced regions*

The bin boundary t_1 has the fundamental role to separate the *prompt* from the *displaced* regions of the analysis and the corresponding background regimes. The optimal value is chosen based on the criterion that all the irreducible background from $B^+ \to K^+\mu^+\mu^-$ decays must be included in the first bin. Table 15 shows the expected number of events for different cuts on the time of decay of the di-muon system. It is found that only $\mathcal{O}(1)$ event is expected to have $t > 1$ ps, therefore, the split between the *prompt* and *displaced* regions is set to $t_1 = 1$ ps. In order to take into account the limited decay-time resolution of the detector, the 1st bin is then symmetrically defined as $|t| < 1$ ps.

6.5.2 *Selection optimisation*

Several figures of merit are commonly used in data analysis to identify the ideal threshold for the selection of a given signal. These typically approximate the signal significance by comparing the signal and background yields or efficiencies. In the following, an accurate optimisation is obtained by investigating the best expected excluded limit through the minimisation of the CLs.

Each bin of the classifier response is scanned in step of 0.01 in its domain $[-1, 1]$. For each value, the expected limit is evaluated for several signal hypotheses: four masses $m_\chi = 500$, 1000, 2500 and 4500 MeV$/c^2$, each with lifetimes $\tau_\chi = 0.1$, 1, 10 and 100 ps. In order to compute the CLs for a given background and signal-plus-background hypotheses, the expected signal and background yields are required for each configuration.

The expected signal yield is obtained from the simulation and normalised in terms of the observed number of events of the control channel. For the entire optimisation procedure, a branching ratio of 1×10^{-9} is assumed for the signal mode.[4]

---

[4] It has been shown that the optimisation is independent on this choice [157].

Table 15: Contribution of the irreducible background $B^+ \to K^+\mu^+\mu^-$. The number of expected events is shown for different cuts on the decay-time of the di-muon system. The separation at $t = 1$ ps illustrates the optimised division between first and second bins of the analysis.

$B^+ \to K^+\mu^+\mu^-$	Expected N. of events
-	7586 ± 455
$t > 0$ ps	3754 ± 225
$t > 0.1$ ps	779 ± 47
$t > 0.2$ ps	215 ± 13
$t > 0.3$ ps	86 ± 6
$t > 0.4$ ps	38 ± 3
$t > 0.5$ ps	18.0 ± 1.6
$t > 0.6$ ps	10.5 ± 1.1
$t > 0.7$ ps	6.3 ± 0.8
$t > 0.8$ ps	3.1 ± 0.5
$t > 0.9$ ps	2.2 ± 0.4
$t > 1.0$ ps	1.2 ± 0.3
$t > 1.1$ ps	1.0 ± 0.3
$t > 1.2$ ps	0.9 ± 0.3
$t > 1.5$ ps	0.6 ± 0.3
$t > 2$ ps	0.12 ± 0.08

For the expected background yield the following procedure is adopted

i. For each bin and for each selection cut, the total number of background events is determined. The expected $B^+ \to K^+\mu^+\mu^-$ decay yield is derived from the simulation, while the combinatorial background is obtained with a linear fit to the right B invariant mass sideband of data;[5]

ii. A large number of pseudoexperiments is generated with the obtained total background yields and accordingly to the background distributions;

iii. For each pseudoexperiment and for each tested mass, the background prediction in the di-muon search window is obtained with the linear fit to the di-muon sidebands;

---

[5] In order not to introduce any bias in the optimisation procedure, the sample receives the BDTG classifier from one of the five testing parts defined in the folding procedure and not from the default one used in the final selection.

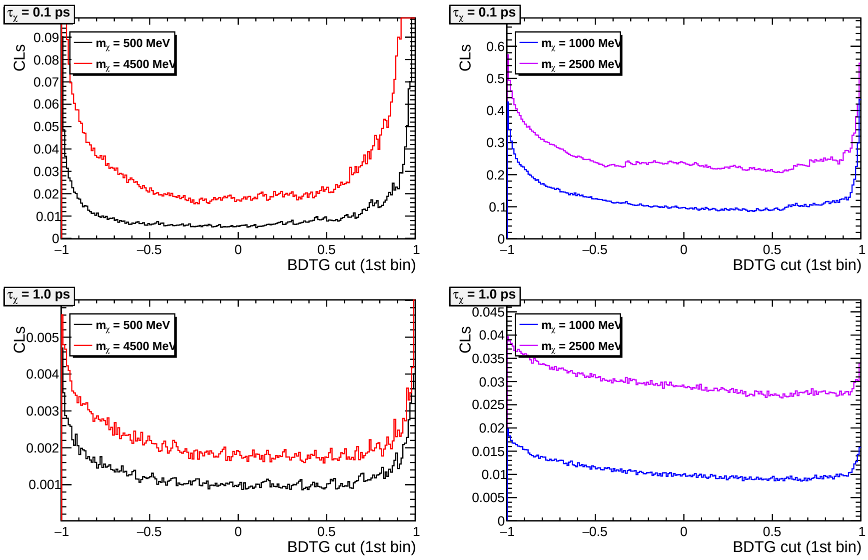


Figure 51: Expected CLs as function of the BDTG selection in the first bin for signal hypotheses with (top) $\tau_\chi = 0.1$, (bottom) $\tau_\chi = 1$ ps, (left) $m_\chi = 500$, 4500 MeV$/c^2$ and (right) $m_\chi = 1000$, 2500 MeV$/c^2$.

iv. Finally, from the distribution of the obtained values from the various pseudoexperiments, the expected background yield and its uncertainty is obtained for each tested mass.

6.5.2.1 *First bin optimisation*

Figure 51 shows the CLs as function of the BDTG selection in the 1^{st} bin for different signal mass and lifetime hypotheses. Only lifetimes $\tau_\chi \leq t_1 = 1$ ps are relevant for the first bin optimisation and are considered at this stage. No strong dependence on the mass is observed in the minimisation of the CLs, therefore, a common BDTG selection is defined in the first bin. The final cut is set to BDTG > 0.6, this choice privileges the high purity of the selected sample, while higher values would result in a reduced sensitivity at very low and very high masses.

6.5.2.2 *Second bin optimisation*

The background contamination of the *displaced* region is strongly suppressed. Despite the obvious advantage in terms of signal sensitivity, there are limiting factors in the optimisation procedure. In fact, a tight selection may limit the number of background events for the di-muon sideband fit, that is the key feature of this search. Thus, studies with pseudoexperiments are required in order to guarantee a good and reliable background extrapolation and to satisfy the following conditions

i. the di-muon sidebands must contain at least four events in total;

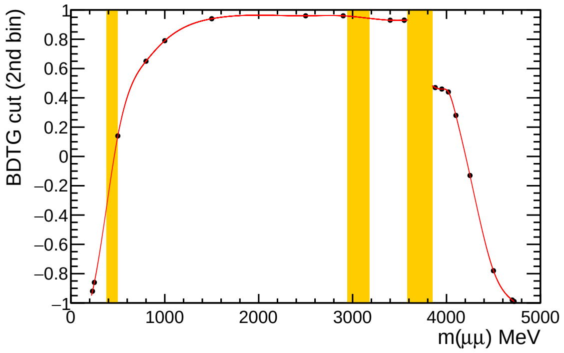


Figure 52: Value of the BDTG selection as function of the mass in the second bin. The function is interpolated between points where the requirements described in the text were tested with pseudoexperiments. Yellow bands correspond to vetoed regions in the search.

ii. each sideband must contain at least one event;

iii. the probability that both of the above points are satisfied must be $> 99\%$.

The selection optimisation proceeds similarly to the first bin. It is found that, for all the tested signal hypotheses, the largest sensitivity is always obtained for selection corresponding to the tightest possible BDTG cut compatible with the above-mentioned requirements. Figure 52 displays the resulting mass-dependent BDTG selection applied in the second bin. The value of the obtained cut is interpolated between each mass point and allowed to be discontinuous on each sides of vetoed regions. After this optimised selection, the expected number of background events in the displaced region of the analysis (and in the corresponding di-muon window search) is found to be of the order of $\mathcal{O}(1)$ for all the considered mass hypotheses.

6.5.3 *The third bin: the "zero background" region*

This section introduces the third bin of the analysis, that consists of very displaced events and is intended as a "zero background" search. This region makes an exception to the sidebands background extrapolation approach. In fact, it is based on the criterion to have the smallest possible number of background events in the full di-muon mass range. However, the sensitivity of the search is found to be mostly independent on the choice of the bin boundary t_2, with multiple configurations of t_2 and BDTG cut in the 3^{rd} bin that give identical performances. In the following, the third bin is defined as $t > t_2 = 10$ ps and the relative selection is tightened to reduce the total expected number of background events to $\mathcal{O}(1)$ in the entire B

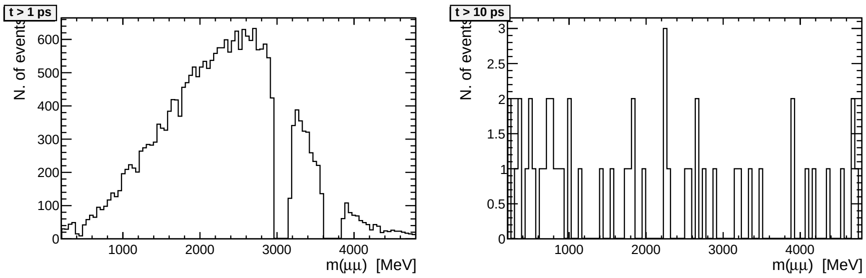


Figure 53: Di-muon invariant mass distribution for events in the right B mass sideband for $t > t_1 = 1$ ps (left) and $t > t_2 = 10$ ps (right) before the multivariate selection.

signal region. This requirement corresponds to a selection of BDTG > 0.8.

Due to the limited population of events in this region, any assumptions on the di-muon mass distribution are not reliable. Figure 53 shows the combinatorial background distribution of the di-muon invariant mass for the second and third bins of the analysis, $t > 1$ and $t > 10$ ps, respectively. In order to overcome the uncertainty on the background prediction for the di-muon search in the third bin, two hypotheses are tested: a flat lineshape and a *pdf* built from the di-muon distribution observed in the second bin. A large systematic uncertainty is added to the background prediction, assigned as three times the difference between the two obtained *pdfs*. Despite this conservative approach and potentially large systematics, no difference is found in the expected sensitivity for the two tested configurations, with and without the systematics.

6.5.3.1 *Sensitivity of an hypothetical observation with injected signal*

The third bin is mainly introduced in this analysis in order to capitalise the discovery potential in the case of an observation compatible with a long lifetime signal. However, its impact on the upper limit procedure is found to be marginal, even in the case of very long lifetimes, for which most of the sensitivity is already achieved by the second bin alone. Figure 54 shows the comparison of the expected 95% confidence level excluded limit when including/removing the third bin, for different signal mass and lifetime hypotheses. It is observed that the introduction of the third bin slightly improves the excluded limit for signal with long lifetimes.

The sensitivity to a hypothetical signal observation is tested by injecting an artificial signal in the framework. This assumption is studied with signal of $m_\chi = 2500\ \mathrm{MeV}/c^2$, $\tau_\chi = 100$ ps and $\mathcal{B} = 1 \times 10^{-9}$. Table 16 shows the p-value of the artificially created observation under the two-bins and three-bins scenarios, together with the obtained signal and background yields. The enormous difference in the sensitivity, estimated to increase from a p-value of $\mathcal{O}(10^{-2})$, in the simplified two

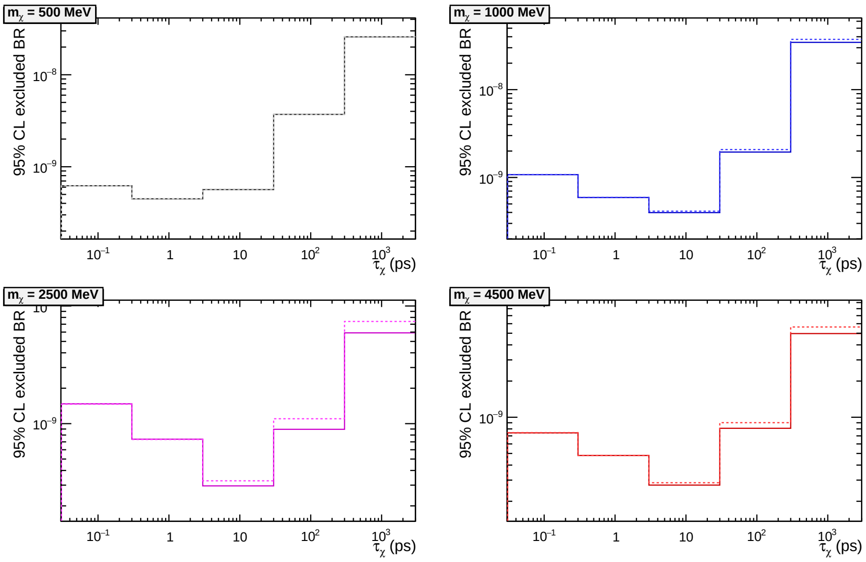


Figure 54: Expected excluded branching ratios for different masses and lifetimes at 95% confidence level. The excluded limits are shown both in the three bins (solid line) and two bins (dashed line) scenario for the four tested mass: $m_\chi = 500$ (top left), $m_\chi = 1000$ (top right), $m_\chi = 2500$ (bottom left) and $m_\chi = 4500$ MeV$/c^2$ (bottom right).

Table 16: Expected signal and background yields and observed number of events obtained with an artificially-injecteded signal with $m_\chi = 2500$ MeV$/c^2$, $\tau_\chi = 100$ ps and $\mathcal{B} = 1 \times 10^{-9}$. The two columns correspond to the two and three bins scenarios, respectively.

	2 bins		3 bins		
n_{sig}	(0.14,	3.5)	(0.14,	1.3,	2.5)
n_{bkg}	(64,	1.4)	(64,	1.4,	0.003)
n_{obs}	(64,	5)	(64,	2,	3)
p-value	0.008		$< 10^{-7}$		

bins scenario, to less than $\mathcal{O}(10^{-7})$,[6] for the complete analysis, proves the fundamental role of the third bin in the case of a long lifetime signal observation.

6.5.4 *Summary of the optimised selection*

After the optimisation procedure the binning scheme and corresponding multivariate selections can be listed as

---

[6] The computing resources required to generate an appropriate number of pseudoexperiments to exactly quantify the deviation from the background-only hypothesis exceeds the availability for the scope of this test, however, the presented boundary p-value $\lessapprox \mathcal{O}(10^{-7})$ adequately illustrates the result of this study.

- 1^{st} bin: $|t| < 1$ ps, BDTG > 0.6,
- 2^{nd} bin: $1 < t < 10$ ps, BDTG $> f(m)$, where $f(m)$ is plotted in Fig. 52,
- 3^{rd} bin: $t > 10$ ps, BDTG > 0.8,

where the combinatorial background rejection rate is found to be approximately 0.98 for the first and second bins, and around 0.9 for the third bin.

6.6 MODELLING THE SIGNAL

One of the challenges of the analysis is to correctly parametrise the signal in the entire m_χ-τ_χ parameter space. In the following, all the aspects relevant for the signal description are discussed. These are studied with simulations, validated on control channels and, if required, corrected for possible discrepancies.

6.6.1 *Signal mass resolution*

The size of the di-muon search window is defined in terms of the local mass resolution, which varies across the whole mass range. The mass resolution dependence as a function of m_χ is studied with simulated samples. The reconstruction of candidates in the B mass signal region is repeated with the constraints to originate from the PV and to have the invariant mass fixed to the nominal mass of the B^+ meson [156]. This is found to significantly improve the mass resolution, especially for large signal masses.

The mass distribution of the signal simulated samples is fitted with the sum of two Gaussian distributions with the same mean. No significant dependence on the signal lifetime is found, hence the mass resolution is assumed to be constant for all the considered lifetimes. Simulations are found to underestimate the mass resolution observed in data, therefore a scaling factor evaluated using $J/\psi(1S) \to \mu^+\mu^-$ decays of $\sigma^{DATA}_{J/\psi(1S)}/\sigma^{MC}_{J/\psi(1S)} = 1.167$ is applied to the values obtained from simulated datasets.

Figure 55 shows the resulting di-muon mass resolution as a function of m_χ, interpolated across the whole mass range and varying between 2 and 9 MeV/c^2. The mass resolution obtained from data for $\phi(1020) \to \mu^+\mu^-$ and $\psi(2S) \to \mu^+\mu^-$ decays is overlaid as reference. These are found to be compatible within two standard deviations, equivalent to 6% of the scale factor. This discrepancy is assigned as uncertainty on the scale factor and is added as a source of systematics, as described in section 6.7.2.

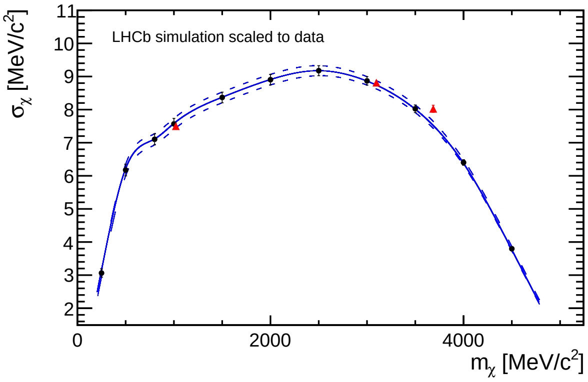


Figure 55: Signal mass resolution as a function of m_χ. The black points show the resolution obtained from the simulation and scaled to the $J/\psi(1S) \to \mu^+\mu^-$ resolution observed in data; the red triangles show the resolution observed in data for $\phi(1020) \to \mu^+\mu^-$, $J/\psi(1S) \to \mu^+\mu^-$ and $\psi(2S) \to \mu^+\mu^-$ decays as reference.

6.6.2 *Signal efficiency*

The overall signal efficiency is factorised in several components: acceptance, reconstruction and selection efficiencies. The last one, in particular, contains all the requirements of the analysis, namely stripping, preselection, particle identification, multivariate selection and finally the probability to lie within the signal region defined around the B meson mass and within the di-muon search window around each tested mass. The efficiency related to the di-muon search window is obtained by integrating the fitted di-muon mass distribution in the range between $\pm 2\sigma_m$, and its uncertainty is evaluated by varying the resolution σ_m within its error. Particle identification efficiency includes the PID requirements contained in the stripping and is obtained with the `PIDCalib` package. All the other efficiency components and corresponding uncertainties are evaluated with simulations.[7] Figure 56 shows the dependence of the total signal efficiency on the signal mass of the simulated datasets.

6.6.2.1 *MC signal reweighted lifetimes*

A clear decrease of the reconstruction efficiency is observed for long lifetimes, since χ candidates start to decay after the VELO. Each simulated signal sample is then reweighted in order to obtain the

---

[7] Trigger efficiencies are often inaccurate in the simulation. However, Ref. [79] proved, using a data-driven approach, that muon trigger lines are well modelled and do not require additional corrections.

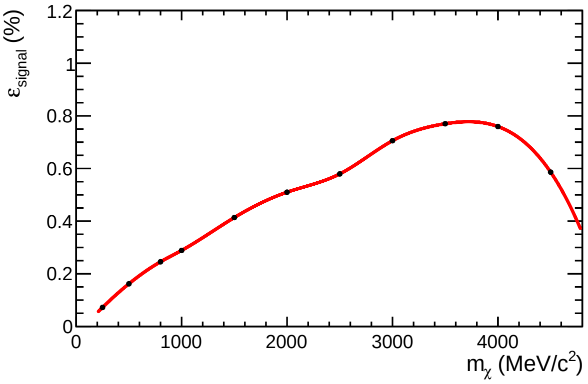


Figure 56: Total signal efficiency as function of the mass for simulated samples generated with $\tau_\chi = 100$ ps.

signal efficiency as function of the lifetime for each simulated mass. The weight to be assigned to simulated events is given by

$$w_i = \frac{1}{\tau_2} e^{-t_i/\tau_2} \bigg/ \frac{1}{\tau_1} e^{-t_i/\tau_1}\,, \tag{63}$$

where τ_1 is the lifetime of the generated sample, τ_2 is the new lifetime to be rescaled and t_i is the time of decay of the single event.

Firstly, the procedure is validated on simulated datasets with $\tau_\chi =$ 1, 10, 100 and 1000 ps and $m_\chi = 2500$ MeV$/c^2$. The sample with $\tau_\chi =$ 100 ps is used to obtain the reweighted efficiency for all the others. While a good overall agreement between the reweighted approach and the values obtained from the original samples is seen, a discrepancy of approximately three standard deviations is observed for the signal sample with $\tau_\chi = 1000$ ps. A systematic uncertainty is added to take this effect into account for signal with very long lifetime, as described in section 6.7.1.

The obtained signal efficiency is then interpolated across the entire m_χ-τ_χ parameter space. Figure 57 shows the resulting two-dimensional efficiency complete of the full list of the background rejection cuts and vetoes. Note that the structure around the invariant masses 1.8 GeV$/c^2$ corresponds to the effect of the additional ProbNN$_\mu > 0.4$ applied to reject D^0 backgrounds (that results to be more than 90% efficient on signal), while the inefficiencies around the invariant masses 1 GeV$/c^2$ and 4.2 GeV$/c^2$ correspond to the $\phi(1020)$ and $\psi(4160)$ regions where only the second and third decay-time bins are used for the search.

6.7 SYSTEMATIC UNCERTAINTIES

The exclusion limit resulting from this search is expected to be dominated by the limited statistics of the available dataset rather than by systematic effects. Table 17 summarises the systematic uncertainties re-

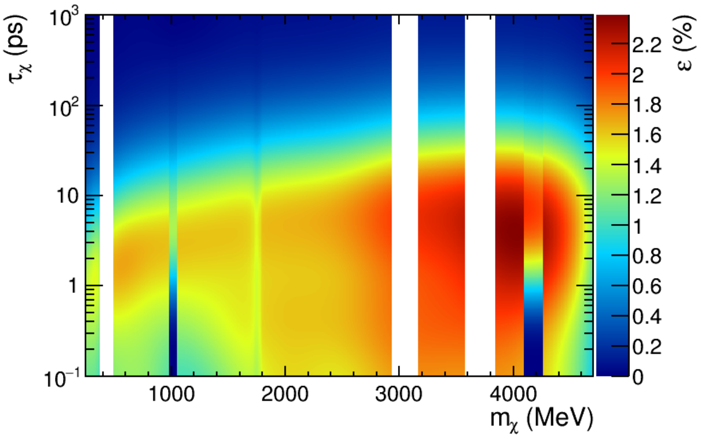


Figure 57: Two-dimensional signal efficiency as function of m_χ and τ_χ. Background rejection cuts and the vetoes are also applied and appear as empty or discontinuous regions in the figure.

Table 17: List of systematic uncertainties.

Source	Uncertainty
Signal resolution	$(1.5 \div 2)\%$
MC size	$(2 \div 6)\%$
MC lifetime reweighting	$(0 \div {}^{+0}_{-20})\%$
Background mass shape mismodelling	0.08× stat. err.
Normalisation branching ratio	3 %

lated to this measurement. Note that the impact of these uncertainties on the excluded limit is found to be minimal, on average the excluded limit is increased by only 2% in the considered m_χ-τ_χ plane. In the following, additional systematic uncertainties besides the mismodelling of the background mass shape (introduced in Sec. 6.4) and the uncertainty on the branching ratio of the normalisation channel [156] are discussed.

6.7.1 *Modelling of the signal efficiency*

A systematic uncertainty is introduced to take into account the finite size of the simulated samples used to estimate the signal efficiency. This affects especially the two boundaries of the examined lifetimes, $\tau_\chi = \mathcal{O}(0.1)$ and $\mathcal{O}(1000)$ ps, where large weights are assigned to a limited number of events. For these two domains the uncertainty associated to the signal efficiency results in approximately 6%, while for all the intermediate hypotheses, $\tau_\chi = \mathcal{O}(1 \div 100)$ ps, is reduced to

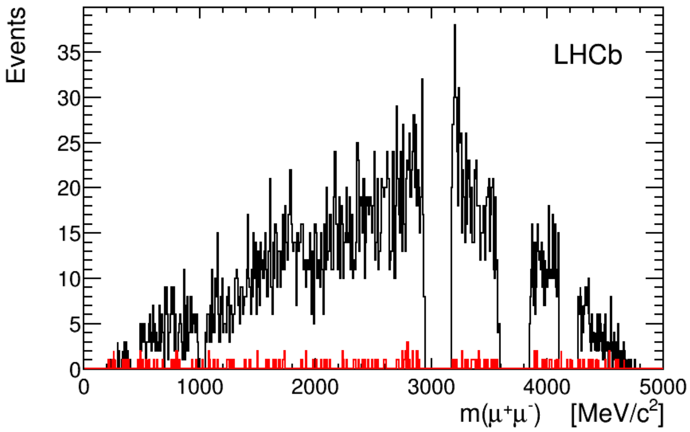


Figure 58: Di-muon invariant mass distribution for events in the B mass signal region in the first (black) and second (red) bin of the search. The binning scheme corresponds to one sigma of the mass-dependent mass-resolution model. No events are observed in the third bin of the analysis.

2%. Furthermore, the reweighting procedure is found to overestimate of approximately 20% the signal efficiency for very long lifetimes. An asymmetric and linearly varying value between ${}^{+0\%}_{-0\%}$ and ${}^{+0\%}_{-20\%}$ uncertainty is introduced from $\tau_\chi = 100$ to 1000 ps, respectively.

6.7.2 *Modelling of the signal mass resolution*

The systematic uncertainty associated to the mass resolution is given by two contributions: the uncertainty on the resolution obtained from the fit to the simulation and the relative error of 6% assigned to the scaling factor. The latter is found to be the dominant and its impact is studied by varying the mass resolution within its uncertainty and by considering the difference on the resulting efficiencies. This results between 1.5% and 2% depending on the signal mass.

6.8 RESULTS

Figure 58 shows the di-muon invariant mass distribution for events in the B mass signal region in the first and second bins of the analysis. No events are observed in the third bin. Since no significant excess from the background-only hypothesis is found, an upper limit on the branching ratio $\mathcal{B}(B^+ \to K^+\chi(\to \mu^+\mu^-))$ is set at 95% confidence level.

The excluded limit is shown in Fig. 59 as function of m_χ and τ_χ. Figure 60 shows the same result in one dimensional plots for several mass and lifetime hypotheses. The obtained limits vary between 2×10^{-10} and 10^{-7} and are most stringent in the region corresponding

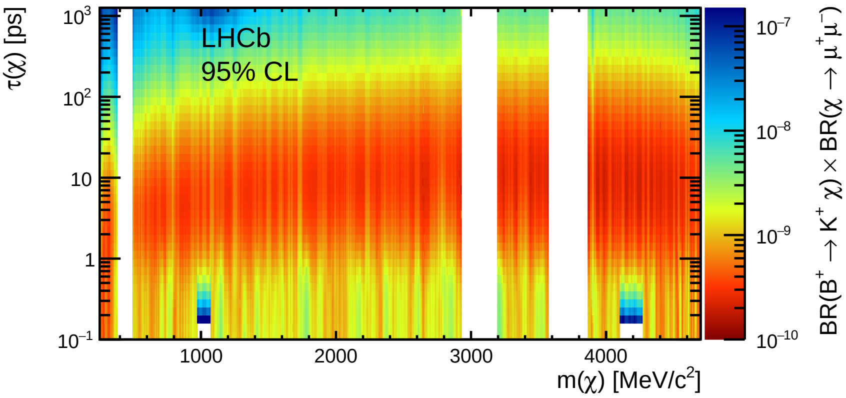


Figure 59: Excluded branching ratio for the $B^+ \to K^+\chi(\to \mu^+\mu^-)$ decay as a function of m_χ and τ_χ at 95% confidence level. Regions corresponding to the fully-vetoed K^0_S, $J/\psi(1S)$, $\psi(2S)$ and $\psi(3770)$ and to the partially-vetoed $\phi(1020)$ and $\psi(4160)$ resonances are excluded from the figure.

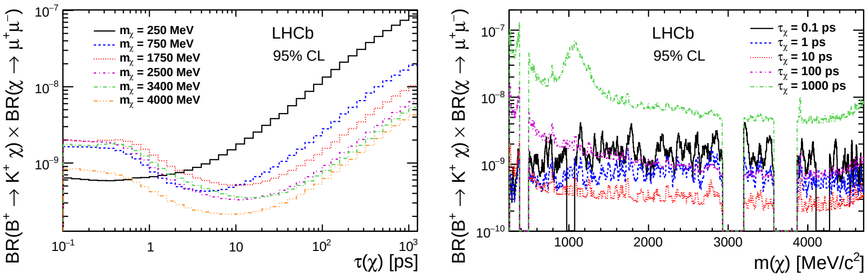


Figure 60: Excluded branching fraction for the $B^+ \to K^+\chi(\to \mu^+\mu^-)$ decay at the 95% confidence level for (left) several mass and (right) lifetime hypotheses. Regions corresponding to the fully-vetoed K^0_S, $J/\psi(1S)$, $\psi(2S)$ and $\psi(3770)$ and to the partially-vetoed $\phi(1020)$ and $\psi(4160)$ are excluded from the figure.

to $\tau_\chi = \mathcal{O}(10)$ ps. Shorter lifetimes are affected by the larger background of the first decay-time bin, while for longer lifetimes the limit becomes weaker as the probability for the χ to decay within the VELO decreases. Nevertheless, the presented analysis sets the most stringent upper limits on the studied decay to date, improving previous results by up to one order of magnitude, with the long lifetimes region of the explored parameter space as the one that benefits the most from the improvement.

This result implies stringent constraints on theories that predict the existence of new light scalar particles. Figure 61 shows the excluded region at 95% confidence level of the parameter space of the inflaton model presented in section 1.2.1. Constraints are placed on the square of the mixing angle, θ^2, as function of the inflaton mass. The translation between the excluded branching fraction for a given lifetime and the derived constraint on the inflaton mixing angle squared is ruled

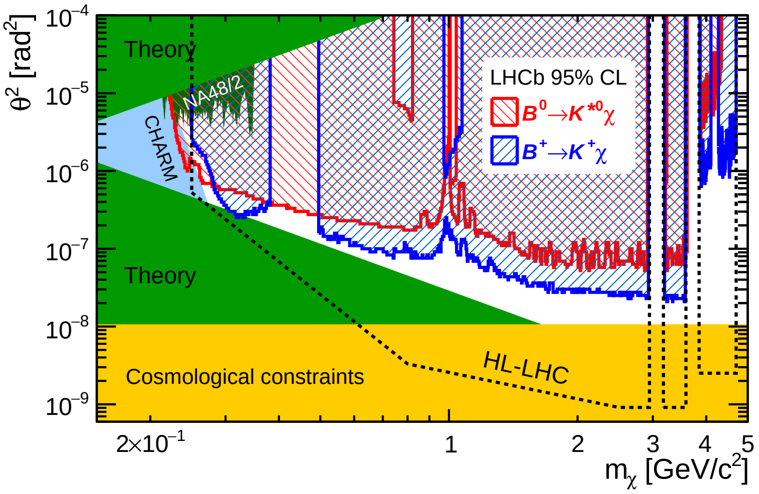


Figure 61: Parameter space of the inflaton model described in Refs. [26,33,37]. The region excluded at 95% confidence level by this analysis is shown by the blue hatched area. The region excluded by the search with $B^0 \to K^{*0}\chi(\to \mu^+\mu^-)$ decays [36] is indicated by the red hatched area. Direct experimental constraints set by the CHARM [34] and NA48/2 [35] experiments are also shown, as well as regions forbidden by theory or cosmological constraints [33]. Finally, the dashed line corresponds to the projected sensitivity for the presented measurement with the dataset expected to be collected after the LHCb Phase-2 Upgrade [300 fb^{-1}].

by Eq. 18 and the information included in Figs. 4 and 3. Theoretical and cosmological constraints on the model [27,33] are also included together with experimental limits previously set by CHARM [34], NA48/2 [35] and LHCb [36] experiments. This measurement excludes a large fraction of the theoretically allowed parameter space confirming the great interest in the study of B-meson decays as possible direct searches of New Physics. Finally, the large statistics foreseen to be collected during the LHCb Run-II and future upgrades will allow to completely rule out these light inflaton models (within the accessible mass range) and are expected to improve to upper limit on θ^2 by up to more than two orders of magnitudes - see Fig. 61.

7

SENSITIVITY OF A DIRECT FIT TO WILSON COEFFICIENTS

This chapter discusses the experimental features involved in the amplitude fit presented in chapter 3 for $B^0 \to K^{*}\mu^{+}\mu^{-}$ decays and the corresponding expected sensitivity. At the time of this dissertation, a binned angular analysis of $B^0 \to K^{*}\mu^{+}\mu^{-}$ decays is ongoing within the LHCb collaboration, aiming to update the published results of Ref. [79] with the dataset collected during the years 2015 and 2016 of Run-II. The strong link between the "classic" angular observables and the new proposed method leads to the natural choice of using the same dataset and analysis strategy in both measurements. This allows a synergy between the two approaches, especially for the validation of the results of the amplitude fit obtained through the projections into angular observables.

In summary, the sensitivity studies reported in this chapter assume the analysis of the datasets collected by the LHCb experiment during Run-I and the first two years of Run-II, corresponding to a total integrated luminosity of approximately 5 fb^{-1}.

7.1 TENSORFLOW FRAMEWORK FOR AMPLITUDE ANALYSIS

This analysis is performed within a new framework built on the `TensorFlow` python library. `TensorFlow` is a machine learning interface developed by Google [158] that can be used to express a wide variety of algorithms. Despite its extremely young age, the first release only dates back to 2015, this tool has seen a remarkable success in the machine learning community, leading to dozens of applications in computer science and numerous other fields.[1] The great versatility of the system and the natural capability in processing large amount of data makes `TensorFlow` an optimal framework for data analysis in particle physics as well.

The essence of this network stands in the construction of computational *graphs*, which are composed of *nodes*. The graph represents a dataflow computation, while each node represents the instance of an operation and can have zero or more inputs/outputs. Figure 62 shows the structure of a graph for a simple operation with one variable x and two parameters, b and W. For illustration, the graph corresponding to

---

[1] `TensorFlow` is successfully applied in several machine learning applications in different fields, including speech recognition, computer vision, robotics, information retrieval, natural language processing, geographic information extraction and computational drug discovery.

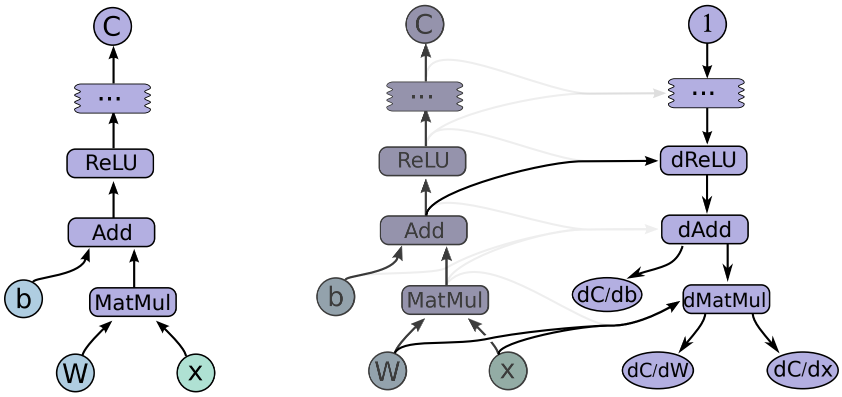


Figure 62: Examples of `TensorFlow` graphs for (left) a simple operation and (right) its gradient, where b and W are parameters (`tf.variables`), x is the input of data (`tf.placeholder`) and C is the function to be calculated, while the rest of the graph shows the required operations classified as nodes [158].

the gradient of the same operation is also reported.[2] The user accesses the `TensorFlow` graph by the creation of a *session*, while the execution of the calculation occurs only when executing (*run*) the session.

Following these guidelines, an innovative framework that conjugates the machine learning expertise with the typical necessities of physics analyses (*e.g.* generation and fitting of a given dataset, estimation of systematic uncertainty, etc.) is implemented. While the generation of events from a given *pdf* is of straightforward implementation, an accept/reject technique based the given function returns the required simulated sample, the fitting procedure requires more attention. The act of performing a fit can be reduced to the problem of maximising the likelihood (or, more commonly, minimising the negative log-likelihood, *nll*). Therefore, it is sufficient to explicitly write the *nll*

```
def unbinned_nll(pdf):
    return -tf.reduce_sum(tf.log(pdf))
```

where `pdf` is an one-dimensional *tensor* corresponding to the evaluation of the *pdf* on the analysed dataset. The great flexibility of the framework allows to easily modify the *nll* to include all the necessary information. For example, in the case of an extended maximum likelihood fit, the *nll* is modified to include a Poissonian term

---

[2] Note that in deep learning algorithms the training of neural networks is based on the minimisation of the *cost* function, therefore, the calculation of the gradient of a given function assumes a central role in the entire framework.

```
def extended_unbinned_nll(pdf,n_sig,n_obs):
    nll = -tf.reduce_sum(tf.log(pdf))
    nll += -tf.reduce_sum(-n_sig+n_obs*tf.log(n_sig))
    return nll
```

where n_{sig} is the signal yield (parameter of the fit) and n_{obs} is the number of observed events. Moreover, in the case of prior knowledge on one or more of the fit parameters, such constraint can be included directly in the likelihood as

```
def ext_unbinned_nll_gauss(pdf,n_sig,n_obs,p,mu,s):
    nll = -tf.reduce_sum(tf.log(pdf))
    nll += -tf.reduce_sum(-n_sig+n_obs*tf.log(n_sig))
    nll += tf.reduce_sum(tf.square(p-mu)/(2*tf.square(s)))
    return nll
```

where p is a *tensor* of fit parameters and μ and s are the corresponding central values and uncertainties. These examples clearly illustrated the great versatility of the presented framework that can be advantageously adapted to all the specific requirements of each analysis.

Once defined the expression of the likelihood for the desired fit model, the second necessary step consists in the minimisation of the *nll*. This point particularly benefits from the legacy of the machine-learning oriented development of the `TensorFlow` package, since numerous minimisers are available (*e.g.* GradientDescentOptimizer, AdamOptimizer, RMSPropOptimizer, etc.). In addition to these built-in methods, Minuit [159], the minimiser embedded in the ROOT software commonly used by the vast majority of the particle physics community, can be easily accessed.[3]

In summary, this analysis is performed within a new framework implemented specifically for the application of the `TensorFlow` interface to amplitude fits. The developed package aims to provide a tool for physics analysis independent on RooFit, a public release of the software is foreseen in the near future along with a peer-reviewed publication. The framework takes advantage of the data-flow system employed by `TensorFlow` to optimally elaborate large amount of data. In addition, thanks to the wide variety of different hardware platforms allowed by `TensorFlow`, running on GPU cards guarantees optimal computational performances speeding up the computing time up to a factor of 40, compared to a single CPU process.

---

[3] Note that, if Minuit automatically provides the calculation of parameter uncertainties and covariance matrices, this is not true for other `TensorFlow` minimisers that lack of a "physical" interpretation and require more manipulation by the user.

7.2 EXPERIMENTAL ASPECTS

Most of the experimental aspects involved in the study of $B^0 \to K^{*}\mu^+\mu^-$ decays are in common with the binned angular analysis [79] and are briefly summarised in this section. The main difference stands in the unbinned treatment of q^2, largely discussed in chapter 3, which is investigated in two kinematic regions, $q^2 \in [1.1, 8.0]$ GeV$^2/c^4$ and $q^2 \in [11.0, 12.5]$ GeV$^2/c^4$, respectively.

Moreover, despite the natural differences between the Run-I and Run-II beam energy conditions, which introduce separate treatments of the datasets, the off-line selection of events is designed as close as possible to the Run-I publication [79]. Further details on these requirements are discussed in chapter 5.

7.2.1 *Invariant B mass fit*

Candidates are required to have $K^+\pi^-\mu^+\mu^-$ reconstructed invariant mass between 5170 and 5700 MeV$/c^2$. An appropriate modelling of the invariant mass is essential to separate the signal events from the combinatorial background surviving the full selection. The signal mass shape is described with a double Crystal Ball [160], consisting in the sum of two Crystal Ball functions with common mean, μ, and tail parameters, α and n, but different widths, σ_1 and σ_2,

$$\mathcal{P}_{sig}(m) = f \times \mathcal{P}_{CB}(m|\mu,\sigma_1,n,\alpha) + (1-f) \times \mathcal{P}_{CB}(m|\mu,\sigma_2,n,\alpha)\,. \tag{64}$$

Each Crystal Ball function is defined by a Gaussian core and an asymmetric power law tail as

$$\mathcal{P}_{CB}(m|\mu,\sigma_1,n,\alpha) = \begin{cases} e^{-\frac{1}{2}(\frac{m-\mu}{\sigma})^2} & \text{for } \frac{m-\mu}{\sigma} > \alpha\,, \\ \frac{a}{(b-\frac{m-\mu}{\sigma})^n} & \text{for } \frac{m-\mu}{\sigma} < \alpha\,, \end{cases} \tag{65}$$

where

$$\begin{aligned} a &= \left(\frac{n}{|\alpha|}\right)^n e^{-\frac{1}{2}\alpha}\,, \\ b &= \frac{n}{|\alpha|} + \alpha\,. \end{aligned} \tag{66}$$

The large branching fraction associated to the tree-level $B^0 \to J/\psi(1S)K^{*0}$ decays is used as a proxy for several constraints and corrections. For the mass fit to $B^0 \to J/\psi(1S)K^{*0}$ decays, a second signal component is included to take into account the suppressed decay $B^0_s \to J/\psi(1S)K^{*0}$ [161], this is expressed by the same signal parametrisation with a shift on the mean by $\Delta m = m(B^0_s) - m(B^0)$. However, the $B^0_s \to K^{*0}\mu^+\mu^-$ decay is neglected when fitting $B^0 \to K^{*0}\mu^+\mu^-$ candidates. Finally, the combinatorial background is parametrised with an exponential function.

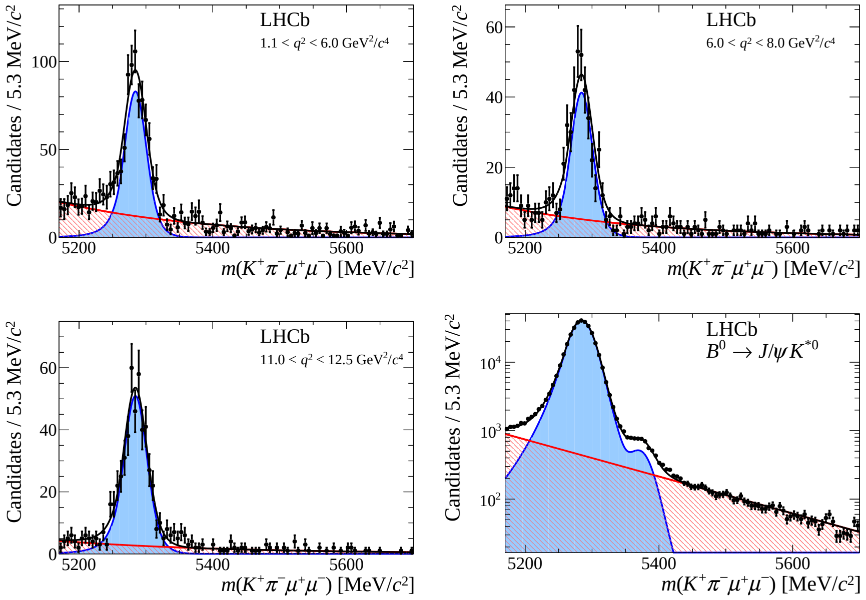


Figure 63: Distribution of the $K^+\pi^-\mu^+\mu^-$ invariant mass for the signal decay $B^0 \to K^{*0}\mu^+\mu^-$ selected in three q^2 regions and for the control channel $B^0 \to J/\psi(1S)(\to \mu^+\mu^-)K^{*0}$, obtained from the LHCb Run-I dataset [79]. The total fitted distribution (black line) as well as the signal (blue shaded area) and background (red hatched area) components are overlaid to data.

Figure 63 shows the results of the $K^+\pi^-\mu^+\mu^-$ invariant mass fit obtained in Ref. [79] for the signal decay $B^0 \to K^{*0}\mu^+\mu^-$, split in three bins of q^2, and for the control channel $B^0 \to J/\psi(1S)(\to \mu^+\mu^-)K^{*0}$. The obtained signal and background yields are reported in Table 18 for the rare mode,[4] and in Table 19 for the control channel together with the values of the mass fit parameters.

7.2.2 *Detector acceptance*

The reconstruction and selection criteria, generally referred to as *acceptance* effects, distort the distributions of the decay angles, as well as the q^2 distribution. The signal acceptance is accurately studied with

---

[4] Note that Ref. [79] made use of several complementary q^2 binning definitions, namely a 1 GeV^2/c^4 and 2 GeV^2/c^4 binning schemes and a wider region $[1.1, 6.0]\ \text{GeV}^2/c^4$. The fit results of Fig. 63 and Table 18 are reported for illustration purpose, since the proposed method combines the region $1.1 < q^2 < 8.0\ \text{GeV}^2/c^4$ in a unique kinematic domain.

Table 18: Signal and background yields in $B^0 \to K^{*0}\mu^+\mu^-$ decays in several bins of q^2 as obtained in Ref. [79].

$q^2[\mathrm{GeV}^2/c^4]$	N_{sig}	N_{bkg}
[1.1, 2.5]	179.7 ± 15.4	124.4 ± 13.5
[2.5, 4.0]	165.4 ± 15.9	206.6 ± 17.1
[4.0, 6.0]	279.5 ± 20.2	300.4 ± 20.7
[6.0, 8.0]	344.3 ± 22.1	344.8 ± 22.1
[11.0, 12.5]	329.8 ± 21.0	212.1 ± 18.0

Table 19: Mass model parameters determined from the fit of the control channel $B^0 \to J/\psi(\to \mu^+\mu^-)K^{*0}$ decay as obtained in Ref. [79].

Parameter	Value
α	1.533 ± 0.033
n	4.23 ± 0.6
σ_1	15.36 ± 0.19
σ_2	25.85 ± 0.82
f	0.704 ± 0.031
μ	5284.339 ± 0.043
Δm	87.21 ± 0.83
exp. slope	−0.006319 ± 0.0001
N_{sig}	343763 ± 822
$N_{B_s^0}$	4199 ± 162
N_{bkg}	26877 ± 649

the simulations and it is parametrised in four dimensions, according to

$$\varepsilon(\cos\theta_L, \cos\theta_K, \phi, q^2) = \sum_{ijmn} c_{ijmn} L_i(\cos\theta_L) L_j(\cos\theta_K) L_m(\phi) L_n(q^2)\,, \tag{67}$$

where the terms $L_h(x)$ denote Legendre polynomials of order h.[5] The coefficients c_{ijmn} are determined with a moment analysis of simulated $B^0 \to K^{*0}\mu^+\mu^-$ phase-space decays, where the three decay angles, $\cos\theta_L$, $\cos\theta_K$ and ϕ are generated flat and a per-event weight is assigned to correct for the non-flat phase-space distribution of q^2. Following this approach, the distribution of simulated events that pass the reconstruction and selection requirements provides a four-

---

[5] Note that the variables must be rescaled in the range $-1 < x < 1$ when evaluating the acceptance to preserve the orthogonality of Legendre polynomials.

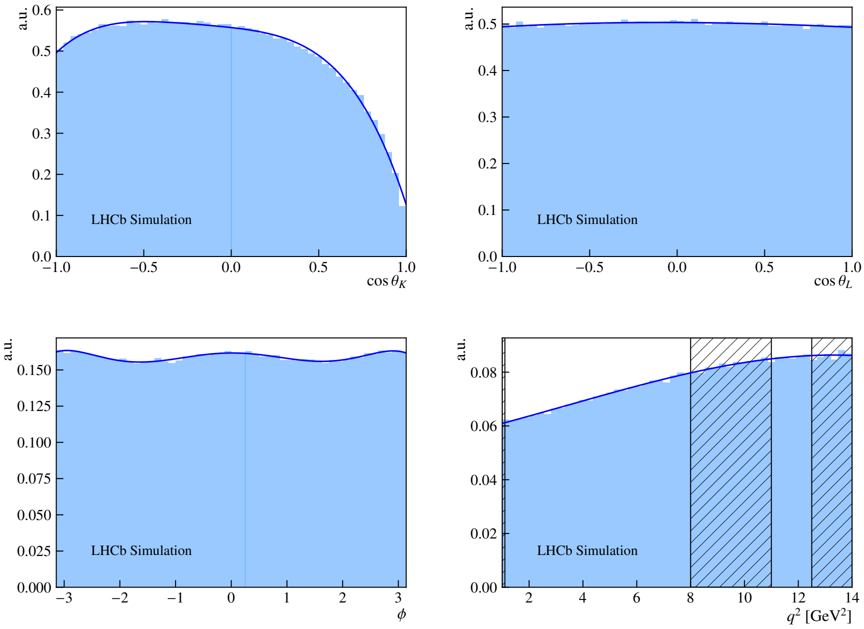


Figure 64: Projections of the four-dimensional acceptance parametrisation in $\cos\theta_L$, $\cos\theta_K$, ϕ and q^2. The filled histograms indicate the simulated $B^0 \to K^{*0}\mu^+\mu^-$ phase-space events, while the overlaid function (solid line) shows the signal efficiency after integrating over the other three variables. For illustration, the kinematic regions excluded from the analysis are shaded in the q^2 projection.

dimensional model-independent[6] description of the signal acceptance, also referred to as *efficiency map* in the following.

Unlike Ref. [79], the q^2 domain considered for the definition of the acceptance is limited to the interval $1.0 < q^2 < 14.0\ \text{GeV}^2/c^4$, accordingly to the kinematic regions studied in this analysis. This allows to reduce the number of polynomial functions required for the parametrisation of q^2 to polynomials of order three and lower, while the parametrisation of $\cos\theta_L$, $\cos\theta_K$ and ϕ comprises $L_h(x)$ up to third, fourth and sixth order, respectively. This results in a total of 560 coefficients that are determined using approximately 1.4 million of reconstructed and selected events simulated with the Run-I conditions. The projections of the obtained acceptance are shown in Fig. 64 for the four different variables.

7.2.3 *$K\pi$ mass system*

The $B^0 \to K^{*0}\mu^+\mu^-$ decay mode extensively analysed in this dissertation always implies the presence of the vector meson $K^{*0}(892)$,

---

[6] Thanks to the parametrisation in terms of all of the relevant kinematic variables needed to describe the decay, the obtained signal acceptance does not depend on the model used in the simulation.

commonly known as P-wave state and reconstructed through its decay products $K^{*0} \to K^+\pi^-$. However, the reconstructed $K^+\pi^-$ system can also originate either from a non-resonant decay or from the decay of scalar resonances, generically denominated as S-wave configuration. The contribution of the S-wave component has been measured to be small, below 10% in the range $796 < m_{K\pi} < 996$ MeV/c^2 [77]. Nevertheless, it can dilute the sensitivity to $B^0 \to K^{*0}\mu^+\mu^-$ decays and pollute the observables of interest [162, 163]. In order to inspect the S-wave contribution in the $K^+\pi^-$ mass region selected for this analysis, the reconstructed invariant mass of the $K^+\pi^-$ system is included in the fit.

The most appropriate and/or convenient description of a narrow isolated resonance such as $K^{*0}(892)$ consists of a relativistic Breit-Wigner lineshape as

$$f_{BW}(m^2_{K\pi}) = \left(\frac{k}{k_{892}}\right) \frac{B'_1(k, k_{892}, d)}{m^2_{K\pi} - m^2_{892} - i m_{892}\Gamma_{892}(m_{K\pi})}, \tag{68}$$

where m_{892} is the pole-mass of the resonance, k and k_{892} are the momentum of the K^+ in the rest frame of the K^{*0} resonance evaluated at a given $m_{K\pi}$ and at the mass of the resonance, respectively, while the orbital angular momentum barrier factors, $(k^L B'_L)$, account for spin-dependent effects in the conservation of the angular momentum,[7] where L is the orbital angular momentum between the decay products of the resonance. Finally, the running width $\Gamma_{892}(m_{K\pi})$ is given by

$$\Gamma_{892}(m_{K\pi}) = \Gamma_{892} B'^2_1(k, k_{892}, d) \left(\frac{k}{k_{892}}\right)^3 \left(\frac{m_{892}}{m_{K\pi}}\right), \tag{69}$$

where Γ_{892} is the resonance width. A more detailed description of the parametrisation of a resonance through the relativistic Breit-Wigner lineshape is given in appendix B.1.

In the event that there is more than one overlapping resonance in the same partial wave or a significant interference with a non-resonant component, the described Breit-Wigner parametrisation is not valid since the sum of multiple contributions violates unitarity. This is the case of the S-wave, where a spin-0 resonance, $K^*_0(1430)^0$, interferes strongly with a non-resonant (or "slowly varying") term. The so-called LASS lineshape [164] has been suggested[8] to model this

---

[7] Low momentum particles have difficulties in generating sufficient angular momentum to conserve the spin of the resonance.

[8] The parametrisation of the $K\pi$ scalar component is not well established. Alternative descriptions are common in the literature, such as an extremely broad Breit-Wigner $\kappa(800)$ resonance at low $m(K\pi)$ or through a dispersive analysis of K-π scattering [165].

scalar amplitude as a non-resonant effective form together with the $K_0^*(1430)^0$ resonance. It reads as

$$f_{LASS}(m_{K\pi}^2) = B_1'(p, p_{1430}, d)\left(\frac{p}{p_{1430}}\right)\left(\frac{1}{\cot\delta_B - i} + e^{2i\delta_B}\frac{1}{\cot\delta_R - i}\right), \tag{70}$$

where $p(p_{1430})$ is the momentum of the $K\pi$ system in the rest frame of the B^0, evaluated at a given $m_{K\pi}$ (at the pole-mass of the $K_0^*(1430)$ resonance) and the two terms $\cot\delta_B$ and $\cot\delta_R$ are defined as

$$\cot\delta_B = \frac{1}{ak} + \frac{rk}{2}, \tag{71}$$

where a and r are empirical parameters,[9] and

$$\cot\delta_R = \frac{m_{1430}^2 - m_{K\pi}^2}{m_{1430}\Gamma_{1430}(m_{K\pi})}. \tag{72}$$

The second term of Eq. 70 corresponds to a relativistic Breit-Wigner for the $K_0^*(1430)^0$ resonance, with

$$\Gamma_{1430}(m_{K\pi}) = \Gamma_{1430}\frac{k}{k_0}\frac{m_{1430}}{m_{K\pi}}. \tag{73}$$

7.2.4 *Background parametrisation*

Several sources of background can potentially mimic the signal decay if they are mis-reconstructed in the detector. These are rejected from the analysis with an appropriate set of cuts, as detailed in Ref. [79] and listed in Table 20. After these requirements, only combinatorial background remains in the selected datasets and particular attention is addressed to the description of the angular distributions. These are assumed to be uncorrelated and are parametrised with a polynomial expansion based on Chebyshev polynomials up to the second order. This choice follows the assumptions of Ref. [79], where the angular parametrisation is tested on the right B invariant mass sideband and is proven to have good compatibility with the observed distributions. In addition, a linear function in $m_{K\pi}$ is assumed to described the smooth variation of combinatorial background in the considered $K\pi$ mass window and is included in the background template. Table 21 lists the values of the background coefficients obtained from the angular fit to the control channel $B^0 \to J/\psi K^{*0}$, where the three angles are fitted simultaneously with the B and $K\pi$ invariant masses.

The above-mentioned parametrisation is aligned to the $B^0 \to K^{*0}\mu^+\mu^-$ binned angular analysis. However, the proposed amplitude fit also requires a parametrisation of the background q^2 distribution.

---

[9] These empirical parameters are fixed to the values $a = 3.83\,[\mathrm{GeV}/c]^{-1}$ and $r = 2.86\,[\mathrm{GeV}/c]^{-1}$ as from Ref. [166].

Table 20: List of requirements applied to reject all possible sources of peaking backgrounds. Mass units are in MeV/c^2.

Background	Rejected if:
$\Lambda_b^0 \to pK^-\mu^+\mu^-$	$(5575 < m_{(\pi\to p)K\mu\mu} < 5665)$ & $\mathrm{DLL}_p(\pi) > 0$ $(5575 < m_{(\pi\to K)(K\to p)\mu\mu} < 5665)$ & $\mathrm{DLL}_K(\pi) > 0$
$B^0 \to K^{*0}\mu^+\mu^-$ with $K \leftrightarrow \pi$	$(795 < m_{(K\leftrightarrow\pi)} < 995)$ & $(\mathrm{DLL}_K(K) + 10 < \mathrm{DLL}_K(\pi))$ & $(\mathrm{DLL}_K(K) - \mathrm{DLL}_K(\pi) > 0)$
$B^0 \to J/\psi K^{*0}$ with $h \leftrightarrow \mu$	$(3036 < m_{(h\to\mu)\mu} < 3156)$ & (`isMuon`(h) = `true` || $\mathrm{DLL}_\mu(h) > 5$)
$B_s^0 \to \phi(\to KK)\mu^+\mu^-$	$(5321 < m_{(\pi\to K)K\mu\mu} < 5411)$ $(1010 < m_{(\pi\to K)K} < 1030)$ || $\mathrm{DLL}_K(\pi) > -10)$ || $(1030 < m_{(\pi\to K)K} < 1075)$ || $\mathrm{DLL}_K(\pi) > 10)$
$B^+ \to K^+\mu^+\mu^-$	$m_{K\pi\mu\mu} > 5380$ & $5220 < m_{K\mu\mu} < 5340$

Table 21: Values of the coefficients of the Chebyshev polynomials used to parametrise the combinatorial background obtained from the angular fit to the control channel $B^0 \to J/\psi K^{*0}$.

Parameter	Value
$c^1_{bkg}(\cos\theta_L)$	0.003 ± 0.025
$c^2_{bkg}(\cos\theta_L)$	-0.301 ± 0.026
$c^1_{bkg}(\cos\theta_K)$	0.153 ± 0.026
$c^2_{bkg}(\cos\theta_K)$	-0.121 ± 0.026
$c^1_{bkg}(\phi)$	-0.023 ± 0.024
$c^2_{bkg}(\phi)$	0.016 ± 0.023
$c^1_{bkg}(m_{K\pi})$	-0.412 ± 0.050

This is obtained from a fit to the right B invariant mass sideband, $5420 < m_{K\pi\mu\mu} < 5800$ MeV/c^2, of the Run-I dataset, separately for the two considered q^2 regions. Chebyshev polynomials of second order and lower are found to describe well enough the observed q^2 distributions, as shown in Fig. 65, resulting in the following set of coefficients

$$\begin{aligned} c^1_{bkg}(q^2) &= 0.30 \pm 0.10\ (0.10 \pm 0.20)\,, \\ c^2_{bkg}(q^2) &= -0.15 \pm 0.10\ (0.20 \pm 0.18)\,, \end{aligned} \tag{74}$$

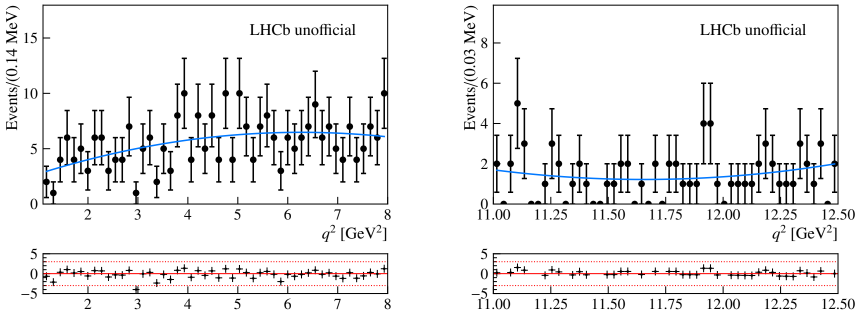


Figure 65: Distribution of q^2 in the two kinematic regions of the analysis for events obtained from the right B invariant mass sideband of the LHCb Run-I dataset. The result of the fit with a second order polynomial function is overlaid to data.

for the kinematic regions $1.1 < q^2 < 8.0$ GeV$^2/c^4$ ($11.0 < q^2 < 12.5$ GeV$^2/c^4$), respectively.

In conclusion, the background *pdf* is described by a six-dimensional function given by

$$pdf_{bkg} = \prod_x \left(\sum c_i T_i(x) \right) \times e^{-\lambda m_{K\pi\mu\mu}} , \qquad (75)$$

where $x \equiv \{\cos\theta_L, \cos\theta_K, \phi, q^2, m^2_{K\pi}\}$, $T_i(x)$ denote Chebyshev polynomials of order i and the last term is the exponential function used to parametrise the $m_{K\pi\mu\mu}$ distribution.

7.3 EXTENSION OF THE MODEL

This section presents the extension of the framework discussed in chapter 3, necessary in order to include all the above-mentioned experimental aspects.

7.3.1 *S-wave description*

Events with the $K\pi$ system in the S-wave configuration are known to modify the distributions of the decay angles and q^2. To appropriately take this contribution into account, two additional complex amplitudes $A^{L,R}_{S0,St}$ must be included in the model. These are taken from Ref. [167] and modified accordingly to Eq. 26 of chapter 2 to coherently include non-local hadronic effects in the S-wave as well. Thus, these scalar amplitudes read as

$$\begin{aligned}\mathcal{A}_{S0}^{L,R} &= -N_0 \frac{\sqrt{\lambda_{K_0^*}}}{M_B\sqrt{q^2}} \Big\{ \Big[(\mathcal{C}_9 - \mathcal{C}_9') \mp (\mathcal{C}_{10} - \mathcal{C}_{10}')\Big] f_+(q^2) \\ &\quad + \frac{2m_b M_B}{q^2}\Big[(\mathcal{C}_7 - \mathcal{C}_7') f_T(q^2) - 16\pi^2 \frac{M_B}{m_b}\mathcal{H}_{S0}(q^2)\Big]\Big\}, \\ \mathcal{A}_{St} &= -2N_0 \frac{M_B^2 - M_{K_0^*}^2}{M_B\sqrt{q^2}}(\mathcal{C}_{10} - \mathcal{C}_{10}') f_0(q^2), \end{aligned} \tag{76}$$

where the symbol K_0^* refers to the scalar $K_0^*(1430)^0$ resonance and the normalisation factor N_0 is given by

$$N_0 = G_F \alpha_e V_{tb} V_{ts}^* \sqrt{\frac{q^2 \beta_l \sqrt{\lambda_{K_0^*}}}{3 \cdot 2^{10}\pi^5 M_B}}, \tag{77}$$

and $\lambda_{K_0^*}$ is the same of Eq. 29 after the substitution $M_{K^*} \to M_{K_0^*}$. Three new form factors are required to parametrise the scalar $B \to K_0^*$ transition matrix elements, namely f_+, f_T and f_0, whose definitions slightly differ from Ref. [167] and are detailed in appendix B.2. The non-local hadronic functions $\mathcal{H}_{S0}(q^2)$ follow Eq. 31 of chapter 2, while Eq. 32 is replaced by

$$\hat{\mathcal{H}}_{S0}(z) = \Big[\sum_{k=0}^{K} \alpha_k^{S0} z^k\Big] f_+(q^2)\,. \tag{78}$$

In addition to the form factors, the described parametrisation introduces a new set of complex parameters $\{\alpha_k^{S0}\}$ to characterise the non-local hadronic contributions to the S-wave and guarantees a uniform formalism between P- and S-wave decays.

7.3.2 *Including the S-wave in the signal model*

In order to incorporate both the P- and S-wave components in the amplitude fit to $B^0 \to K^+\pi^-\mu^+\mu^-$ decays, the differential decay rate of Eq. 24 is modified by the inclusion of the following additional angular terms [167]

$$
\begin{aligned}
I_{1c}^{S} &= \frac{1}{3}\Big\{\Big[|\mathcal{A}_{S0}^{L}|^2+|\mathcal{A}_{S0}^{R}|^2\Big]+\frac{4m_l^2}{q^2}\Big[|\mathcal{A}_{St}|^2+2\,\mathcal{R}e(\mathcal{A}_{S0}^{L}\mathcal{A}_{S0}^{R\,*})\Big]\Big\},\\
I_{2c}^{S} &= -\frac{1}{3}\beta_l^2\Big[|\mathcal{A}_{S0}^{L}|^2+|\mathcal{A}_{S0}^{R}|^2\Big],\\
\tilde{I}_{1c} &= \frac{2}{\sqrt{3}}\,\mathcal{R}e\Big[\mathcal{A}_{S0}^{L}\mathcal{A}_{0}^{L*}+\mathcal{A}_{S0}^{R}\mathcal{A}_{0}^{R*}+\frac{4m_l^2}{q^2}\Big(\mathcal{A}_{S0}^{L}\mathcal{A}_{0}^{R*}+\mathcal{A}_{0}^{L}\mathcal{A}_{S0}^{R\,*}+\mathcal{A}_{St}\mathcal{A}_{t}^{*}\Big)\Big],\\
\tilde{I}_{2c} &= -\frac{2}{\sqrt{3}}\beta_l^2\,\mathcal{R}e\Big[\mathcal{A}_{S0}^{L}\mathcal{A}_{0}^{L*}+\mathcal{A}_{S0}^{R}\mathcal{A}_{0}^{R*}\Big],\\
\tilde{I}_{4} &= \sqrt{\frac{2}{3}}\beta_l^2\,\mathcal{R}e\Big[\mathcal{A}_{S0}^{L}\mathcal{A}_{\parallel}^{L*}+(L\to R)\Big],\\
\tilde{I}_{5} &= \sqrt{\frac{8}{3}}\beta_l^2\,\mathcal{R}e\Big[\mathcal{A}_{S0}^{L}\mathcal{A}_{\perp}^{L\,*}-(L\to R)\Big],\\
\tilde{I}_{7} &= \sqrt{\frac{8}{3}}\beta_l^2\,\mathcal{I}m\Big[\mathcal{A}_{S0}^{L}\mathcal{A}_{\parallel}^{L*}-(L\to R)\Big],\\
\tilde{I}_{8} &= \sqrt{\frac{2}{3}}\beta_l^2\,\mathcal{I}m\Big[\mathcal{A}_{S0}^{L}\mathcal{A}_{\perp}^{L\,*}+(L\to R)\Big],
\end{aligned}
\tag{79}
$$

where I_i^S are pure S-wave contributions and $\tilde{I}_i$ denotes interference terms. The total list of angular terms is summarised in Table 22 and results in

$$
\begin{aligned}
\frac{32\pi}{9}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\,\mathrm{d}\Omega} &= \frac{32\pi}{9}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\,\mathrm{d}\Omega}\bigg|_{\mathrm{P-wave}}\\
&+ \big(I_{1c}^{S}+I_{2c}^{S}\cos 2\theta_l\big)\\
&+ \big(\tilde{I}_{1c}+\tilde{I}_{2c}\cos 2\theta_l\big)\cos\theta_K\\
&+ \big(\tilde{I}_{4}\sin 2\theta_l+\tilde{I}_{5}\sin\theta_l\big)\sin\theta_K\cos\phi\\
&+ \big(\tilde{I}_{7}\sin\theta_l+\tilde{I}_{8}\sin 2\theta_l\big)\sin\theta_K\sin\phi\,.
\end{aligned}
\tag{80}
$$

Furthermore, to better separate the contribution of the S-wave from the P-wave, a simultaneous fit to the $K\pi$ invariant mass is included in the model. The mass lineshapes of Eqs. 68 and 70 contain the angular momentum barrier factors that, in turn, depend on the momentum of the decay products. This results in a non trivial dependence between $m_{K\pi}^2$ and q^2. However, the effect of those terms is found to be negligible when restricting to the kinematic region of interest.[10] The different amplitudes are then modified as follows

$$
\begin{aligned}
A_{0,\perp,\parallel,t}^{L,R}(q^2,m_{K\pi}^2) &= A_{0,\perp,\parallel,t}^{L,R}(q^2)\times f_{892}(m_{K\pi}^2)\,,\\
A_{S0,St}^{L,R}(q^2,m_{K\pi}^2) &= A_{S0,St}^{L,R}(q^2)\times f_{K_0^*}(m_{K\pi}^2)\,,
\end{aligned}
\tag{81}
$$

---

[10] The angular momentum barrier factors preserve momentum conservation and become relevant only in proximity of the border of the available $m_{K\pi}^2$-q^2 phase space.

Table 22: Dependence of the angular coefficients $I_i(q^2)$ on the transversity amplitudes and the corresponding angular terms $f_i(\cos\theta_l, \cos\theta_K, \phi)$.

i	$I_i(q^2)$	$f_i(\vec{\Omega})$
$1s$	$\frac{2+\beta_l^2}{4}\left[\lvert\mathcal{A}_\perp^L\rvert^2 + \lvert\mathcal{A}_\parallel^L\rvert^2 + (L \to R)\right] + \frac{4m_l^2}{q^2}\,\mathcal{R}e\left(\mathcal{A}_\perp^L\mathcal{A}_\perp^{R*} + \mathcal{A}_\parallel^L\mathcal{A}_\parallel^{R*}\right)$	$\sin^2\theta_K$
$1c$	$\left[\lvert\mathcal{A}_0^L\rvert^2 + \lvert\mathcal{A}_0^R\rvert^2\right] + \frac{4m_l^2}{q^2}\left[\lvert\mathcal{A}_t\rvert^2 + 2\,\mathcal{R}e(\mathcal{A}_0^L\mathcal{A}_0^{R*})\right]$	$\cos^2\theta_K$
$2s$	$\frac{\beta_l^2}{4}\left[\lvert\mathcal{A}_\perp^L\rvert^2 + \lvert\mathcal{A}_\parallel^L\rvert^2 + (L \to R)\right]$	$\sin^2\theta_K\cos 2\theta_l$
$2c$	$-\beta_l^2\left[\lvert\mathcal{A}_0^L\rvert^2 + \lvert\mathcal{A}_0^R\rvert^2\right]$	$\cos^2\theta_K\cos 2\theta_l$
3	$\frac{\beta_l^2}{2}\left[\lvert\mathcal{A}_\perp^L\rvert^2 - \lvert\mathcal{A}_\parallel^L\rvert^2 + (L \to R)\right]$	$\sin^2\theta_K\sin^2\theta_l\cos 2\phi$
4	$\frac{\beta_l^2}{\sqrt{2}}\,\mathcal{R}e\left[\mathcal{A}_0^L\mathcal{A}_\parallel^{L*} + (L \to R)\right]$	$\sin 2\theta_K\sin 2\theta_l\cos\phi$
5	$\sqrt{2}\beta_l\,\mathcal{R}e\left[\mathcal{A}_0^L\mathcal{A}_\perp^{L}{}^* - (L \to R)\right]$	$\sin 2\theta_K\sin\theta_l\cos\phi$
$6s$	$2\beta_l\,\mathcal{R}e\left[\mathcal{A}_\parallel^L\mathcal{A}_\perp^{L}{}^* - (L \to R)\right]$	$\sin 2\theta_K\cos\theta_l$
7	$\sqrt{2}\beta_l\,\mathcal{I}m\left[\mathcal{A}_0^L\mathcal{A}_\parallel^{L*} - (L \to R)\right]$	$\sin 2\theta_K\sin\theta_l\sin\phi$
8	$\frac{\beta_l^2}{\sqrt{2}}\,\mathcal{I}m\left[\mathcal{A}_0^L\mathcal{A}_\perp^{L}{}^* + (L \to R)\right]$	$\sin 2\theta_K\sin 2\theta_l\sin\phi$
9	$\beta_l^2\,\mathcal{I}m\left[\mathcal{A}_\perp^L\mathcal{A}_\parallel^{L*} + (L \to R)\right]$	$\sin^2\theta_K\sin^2\theta_l\sin 2\phi$
$1c^S$	$\frac{1}{3}\left\{\left[\lvert\mathcal{A}_{S0}^L\rvert^2 + \lvert\mathcal{A}_{S0}^R\rvert^2\right] + \frac{4m_l^2}{q^2}\left[\lvert\mathcal{A}_{St}\rvert^2 + 2\,\mathcal{R}e(\mathcal{A}_{S0}^L\mathcal{A}_{S0}^{R}{}^*)\right]\right\}$	1
$2c^S$	$-\frac{1}{3}\beta_l^2\left[\lvert\mathcal{A}_{S0}^L\rvert^2 + \lvert\mathcal{A}_{S0}^R\rvert^2\right]$	$\cos 2\theta_l$
$\tilde{1}c$	$\frac{2}{\sqrt{3}}\,\mathcal{R}e\left[\mathcal{A}_{S0}^L\mathcal{A}_0^{L*} + \mathcal{A}_{S0}^R\mathcal{A}_0^{R*} + \frac{4m_l^2}{q^2}\left(\mathcal{A}_{S0}^L\mathcal{A}_0^{R*} + \mathcal{A}_0^L\mathcal{A}_{S0}^{R}{}^* + \mathcal{A}_{St}\mathcal{A}_t{}^*\right)\right]$	$\cos\theta_K$
$\tilde{2}c$	$-\frac{2}{\sqrt{3}}\beta_l^2\,\mathcal{R}e\left[\mathcal{A}_{S0}^L\mathcal{A}_0^{L*} + \mathcal{A}_{S0}^R\mathcal{A}_0^{R*}\right]$	$\cos\theta_K\cos 2\theta_l$
$\tilde{4}$	$\sqrt{\frac{2}{3}}\beta_l^2\,\mathcal{R}e\left[\mathcal{A}_{S0}^L\mathcal{A}_\parallel^{L*} + (L \to R)\right]$	$\sin\theta_K\sin 2\theta_l\cos\phi$
$\tilde{5}$	$\sqrt{\frac{8}{3}}\beta_l^2\,\mathcal{R}e\left[\mathcal{A}_{S0}^L\mathcal{A}_\perp^{L}{}^* - (L \to R)\right]$	$\sin\theta_K\sin\theta_l\cos\phi$
$\tilde{7}$	$\sqrt{\frac{8}{3}}\beta_l^2\,\mathcal{I}m\left[\mathcal{A}_{S0}^L\mathcal{A}_\parallel^{L*} - (L \to R)\right]$	$\sin\theta_K\sin\theta_l\sin\phi$
$\tilde{8}$	$\sqrt{\frac{2}{3}}\beta_l^2\,\mathcal{I}m\left[\mathcal{A}_{S0}^L\mathcal{A}_\perp^{L}{}^* + (L \to R)\right]$	$\sin\theta_K\sin 2\theta_l\sin\phi$

where

$$\begin{aligned} f_{892}(m_{K\pi}^2) &= \mathcal{N}_{BW}\cdot f_{BW}(m_{K\pi}^2)\,, \\ f_{K_0^*}(m_{K\pi}^2) &= \mathcal{N}_{LASS}\cdot\lvert g_S\rvert e^{i\delta_S}\cdot f_{LASS}(m_{K\pi}^2)\,, \end{aligned} \tag{82}$$

and $\mathcal{N}_{BW}$ and $\mathcal{N}_{LASS}$ are normalisation factors, such that[11,12]

$$\begin{aligned} \int_0^\infty \left\lvert\mathcal{N}_{BW}\cdot f_{BW}(m_{K\pi}^2)\right\rvert^2 \mathrm{d}m_{K\pi}^2 &= 1\,, \\ \int_0^\infty \left\lvert\mathcal{N}_{LASS}\cdot f_{LASS}(m_{K\pi}^2)\right\rvert^2 \mathrm{d}m_{K\pi}^2 &= 1\,, \end{aligned} \tag{83}$$

---

[11] Due to the empirical nature of the LASS parametrisation, its normalisation does not have a strong physical meaning, however, despite not strictly necessary, it is introduced in analogy with the Breit-Wigner description of the K^* resonance. The latter, in fact, requires the normalisation condition of Eq. 82 in order to satisfy the consistency of the P-wave differential decay rate when integrating over the entire physical $m_{K\pi}^2$ region.

[12] Note that the LASS parametrisation is curtailed at 1.8 GeV/c^2, above which no data from the LASS experiment is available.

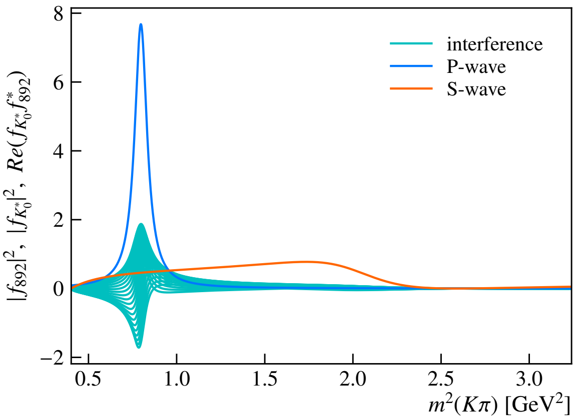


Figure 66: $k\pi$ invariant mass squared for the pure P-wave, pure S-wave and interference term when varying the relative phase δ_S between 0 and π.

while the coefficients g_S and δ_S determine the relative magnitude and phase between the two contributions. Figure 66 shows the $m^2_{K\pi}$ distribution of the P-wave and S-wave contributions, as well as their interference depending on the relative phase.

Within this formalism, the differential decay rate, after integrating in the angles, is given by

$$\frac{\mathrm{d}^2\Gamma}{\mathrm{d}q^2\,\mathrm{d}m^2_{K\pi}} = \left|f_{892}(m^2_{K\pi})\right|^2 \cdot \frac{\mathrm{d}\Gamma^{\mathrm{P}}}{\mathrm{d}q^2} + \left|f_{K_0^*}(m^2_{K\pi})\right|^2 \cdot \frac{\mathrm{d}\Gamma^{\mathrm{S}}}{\mathrm{d}q^2}\,, \tag{84}$$

with

$$\frac{\mathrm{d}\Gamma^{\mathrm{P}}}{\mathrm{d}q^2} = \frac{3}{4}(2I_{1s} + J_{1c}) - \frac{1}{4}(2J_{2s} + J_{2c}) \tag{85}$$

and

$$\frac{\mathrm{d}\Gamma^{\mathrm{S}}}{\mathrm{d}q^2} = 3J^S_{1c} - J^S_{2c}\,. \tag{86}$$

In conclusion, the five-dimensional differential decay rate is given by Eq. 80 when including the $m^2_{K\pi}$ dependence as from Eq. 81. This extended formalism completes the description of $B^0 \to K^*(\to K^+\pi^-)\mu^+\mu^-$ decays presented in chapter 2. For illustration, Figs. 67 and 68 show the projections of the differential branching ratio on the decay angles, q^2 and $m^2_{K\pi}$ for the two kinematic regions of the analysis. The relative magnitude between S- and P-wave is set to $g_S = 0.93$, which corresponds to an averaged S-wave fraction[13] of $F_S \sim 7\%$ and

---

[13] Note that, differently than the binned angular analysis, in the proposed formalism F_S is not a parameter of the fit, but rather an unbinned function $F_S(q^2, m_{K\pi}) = \left|f_{K_0^*}(m_{K\pi})\right|^2 \cdot \frac{\mathrm{d}\Gamma^{\mathrm{S}}}{\mathrm{d}q^2}(q^2) \Big/ \left(\left|f_{892}(m_{K\pi})\right|^2 \cdot \frac{\mathrm{d}\Gamma^{\mathrm{P}}}{\mathrm{d}q^2}(q^2) + \left|f_{K_0^*}(m_{K\pi})\right|^2 \cdot \frac{\mathrm{d}\Gamma^{\mathrm{S}}}{\mathrm{d}q^2}(q^2)\right)$ whose numerator and denominator must be integrated individually to obtain the averaged S-wave fraction in the region of interest.

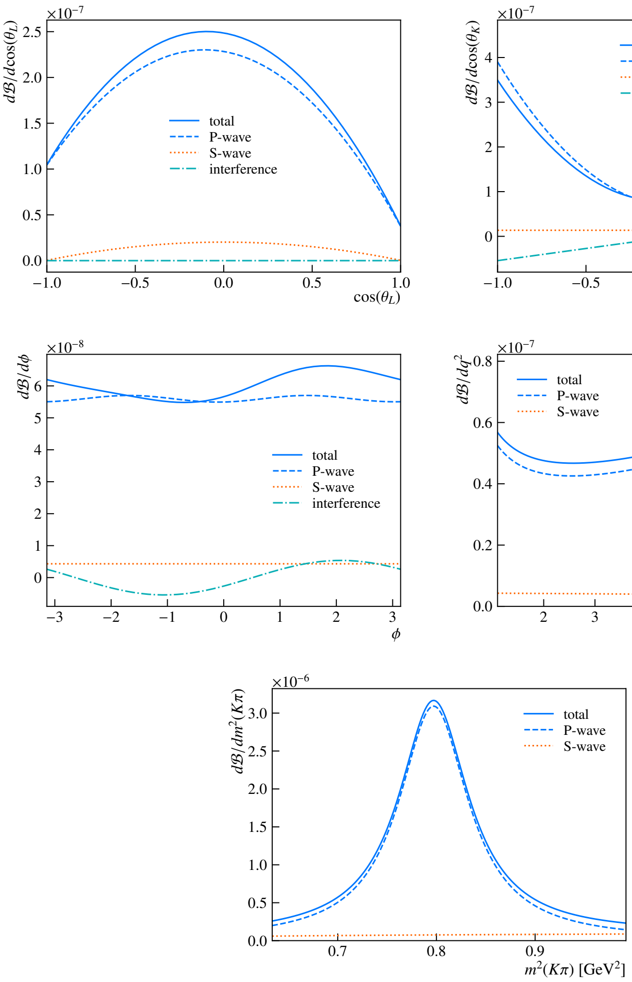


Figure 67: Differential $B^0 \to K^+\pi^-\mu^+\mu^-$ branching ratio as function of the three decay angles $\cos\theta_L$, $\cos\theta_K$ and ϕ, q^2 and $m^2_{K\pi}$ in the in the region $1.1 < q^2 < 8.0\ \mathrm{GeV}^2/c^4$.

2% in the range $1.1 < q^2 < 8.0\ \mathrm{GeV}^2/c^4$ and $11 < q^2 < 12.5\ \mathrm{GeV}^2/c^4$, respectively. This setting is consistent with the corresponding measured values of Refs. [77,79], while the relative phase is arbitrarily set to $\delta_S = \pi/2$.

7.3.3 *Full amplitude fit*

All the experimental effects described in the previous sections must be included in the fit model. This is done by modifying the signal *pdf* as follows

$$pdf_{sig} \equiv \mathcal{N}_{sig} \times \mathcal{P}_{sig}(m_{K\pi\mu\mu}) \times \varepsilon(q^2,\Omega) \times \frac{\mathrm{d}^5\Gamma(B^0 \to K^+\pi^-\mu^+\mu^-)}{\mathrm{d}q^2\,\mathrm{d}\Omega\,\mathrm{d}m^2_{K\pi}}, \tag{87}$$

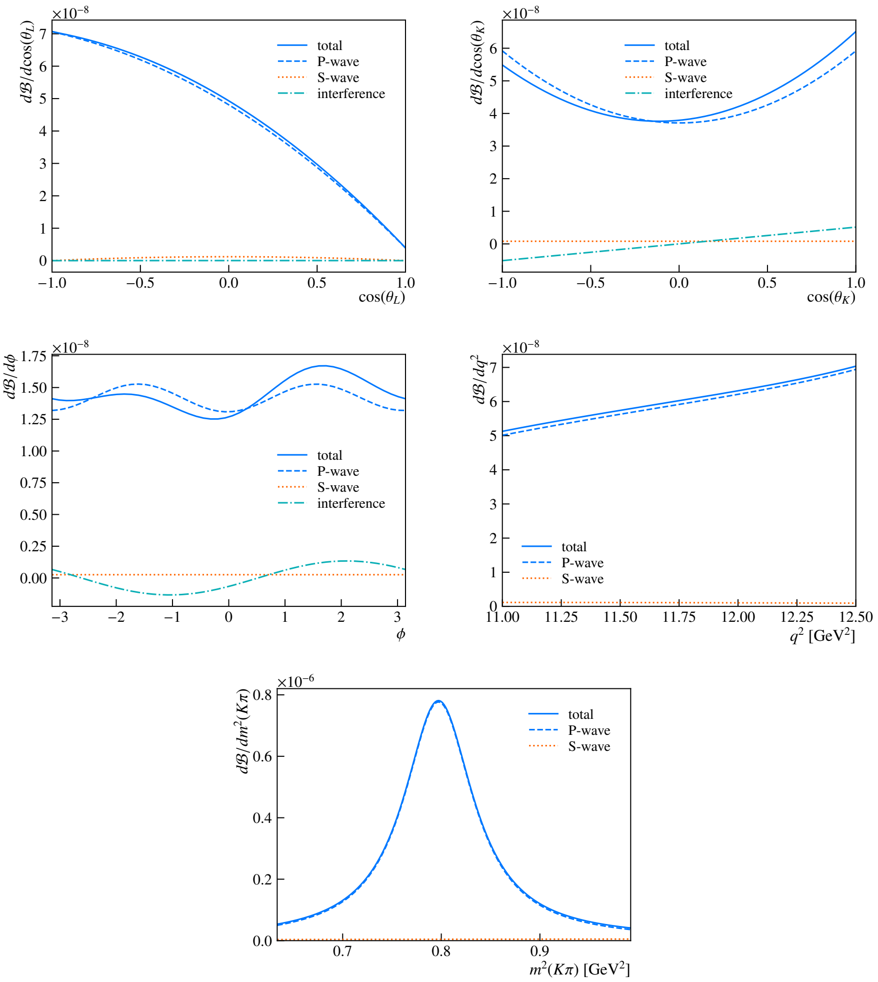


Figure 68: Differential $B^0 \to K^+\pi^-\mu^+\mu^-$ branching ratio as function of the three decay angles $\cos\theta_L$, $\cos\theta_K$ and ϕ, q^2 and $m^2_{K\pi}$ in the in the region $11.0 < q^2 < 12.5\ \text{GeV}^2/c^4$.

where $\mathcal{N}_{sig}$ is a numerical factor to normalise the *pdf* to unity, $\varepsilon(q^2, \Omega)$ is the four-dimensional efficiency map of Eq. 67 and $\mathcal{P}_{sig}(m_{K\pi\mu\mu})$ is the double Crystal Ball of Eq. 64 used to model the reconstructed invariant B mass. The resulting signal *pdf* is a six-dimensional function that takes into account the detector acceptance on an event-per-event basis. Finally, the total *pdf* is built from the signal and background components.

This modified formulation replaces the *signal-only* description presented in chapter 3, in reference to Eqs. 43 and 44, while Eq. 45 is adjusted with the introduction of the background yield, n_{bkg}, as additional fit parameter

$$\mathcal{L}_{\mathcal{B}} = (n_{sig} + n_{bkg})^{N_{obs}} \times e^{-(n_{sig}+n_{bkg})} \,. \tag{88}$$

7.3.4 *Branching ratio measurement*

In the presented model the measurement of the $B^0 \to K^{*0}\mu^+\mu^-$ branching ratio is strongly embedded in the framework. In fact, the signal yield directly enters in the likelihood through Eq. 88 and can be expressed as

$$n_{sig} = N_{J/\psi(1S)} \times \frac{\mathcal{B}_{K^{*0}\mu\mu}}{\mathcal{B}_{J/\psi(1S)}} \times \frac{\varepsilon_{K^{*0}\mu\mu}}{\varepsilon_{J/\psi(1S)}}, \tag{89}$$

where $N_{J/\psi(1S)}$ is the observed yield of the control channel, $\mathcal{B}_{J/\psi(1S)} = \mathcal{B}(B \to J/\psi(1S)K^{*0}) \times \mathcal{B}(J/\psi(1S) \to \mu^+\mu^-)$ is its branching ratio [156],

$$\mathcal{B}_{K^{*0}\mu\mu} = \frac{\tau_B}{\hbar} \int_{m^2_{K\pi}} \int_{q^2 \in Q_i} \frac{\mathrm{d}^2\Gamma}{\mathrm{d}q^2 \mathrm{d}m^2_{K\pi}} \mathrm{d}q^2 \mathrm{d}m^2_{K\pi} \tag{90}$$

is the integrated signal branching ratio in the region of interest and the last term of Eq. 89 is the relative efficiency between the signal mode and the control channel.

However, the efficiency ratio requires a more accurate knowledge of the signal efficiency compared to the differential decay rate fit, where the angular dependence described in Sec. 7.2.2 is sufficient. In fact, the ratio $\varepsilon_{K^{*0}\mu\mu}/\varepsilon_{J/\psi(1S)}$ is typically one of the main systematic uncertainties for branching ratio measurements [77]. Two possible approaches can be pursued to parametrise this term

i. The *classic* approach, which calculates the efficiency from simulation, separately for the two channels, and combines the obtained results afterwards. The value of the efficiency for each channel is simply given by the ratio of the number of events observed in a given region of interest over the number of events generated in the same region.

ii. An *innovative* approach, which aims to exploit the properties of the parametrisation of the signal acceptance of Sec. 7.2.2. The desired relative efficiency can be defined as

$$\frac{\varepsilon_{K^{*0}\mu\mu}}{\varepsilon_{J/\psi(1S)}} \equiv \frac{\int \varepsilon^{4D}\, pdf_{K^{*0}\mu\mu}\, d\Omega\, dq^2}{\int \varepsilon^{4D}_{(q^2=m^2_{J/\psi})}\, pdf_{J/\psi(1S)}\, d\Omega\, \Delta q^2}, \tag{91}$$

where ε^{4D} is the four-dimensional efficiency map defined in Eq 67, $pdf_{K^{*0}\mu\mu}$ is the *pure* signal *pdf* after integrating in the $m^2_{K\pi}$ dimension and normalised to unity, and $pdf_{J/\psi(1S)}$ is the angular *pdf* of the control channel [79, 109]. This approach relies on the use of the same efficiency map (sliced at the mass of the $J/\psi(1S)$ for the denominator) to preserve the required efficiency scale between signal and normalisation channel. Although this implies the extension of the framework to include the angular analysis of $B^0 \to J/\psi(1S)K^{*0}$ decays as a simultaneous component of the fit,[14] two major advantages can be foreseen

i. the efficiency calculation is model-independent - does not rely on the physics model included in the simulation, *e.g.* LHCb simulations do not account for non-local hadronic effects;[15]

ii. homogeneous propagation of systematic uncertainties - the same efficiency map and signal *pdf* concurrently contribute to the amplitude fit and to the branching ratio measurement, this allows to reduce the sources of systematic uncertainties and to uniformly propagate them within the measurement.

7.4 EXPECTED SENSITIVITY

The extended formalism described in the previous section is used to study the expected sensitivity, considering all the experimental effects. The different beam conditions between Run-I and Run-II require distinguished treatments of the two datasets, including separate acceptances, mass fits and background parametrisations. However, differences in these samples are not expected to significantly modify the sensitivity of the measurement, hence, the studies in the following assume homogeneous conditions for the entire dataset, where the Run-I setting is used as reference and the observed yields are scaled to the total analysed dataset.[16]

7.4.1 *Embedded simulations*

The generation of pseudoexperiments follows the guidelines of Sec. 3.1, with the following conditions

i. the S-wave relative magnitude and phase, g_S and δ_S, are set as in Sec. 7.3.2;

i. the S-wave form factors f_+, f_T and f_0, which are currently poorly known due to the limited knowledge on $B \to K_0^*$ transition matrix elements, are assumed to be the same as for $B^+ \to K^+$ decays [168], which parametrisation is further discussed in appendix B.2.1;

---

[14] Note that for a correct description of $B^0 \to J/\psi(1S)K^{*0}$ decays, exotic charmonium states as $Z_c^+(4200)$ and $Z_c^+(4430)$ must be considered [108].

[15] An evaluation of this effect has been studied for $B^0 \to K^{*0}e^+e^-$ decays and is found to be up to 2%. This estimate is obtained with the current framework by turning on/off the non-local hadronic parameters and studying the migration into the region of interest due to the finite q^2 resolution.

[16] The expected yields for the 2015 and 2016 detasets are scaled from Ref. [79] by luminosity and $\sigma_{b\bar{b}} \propto \sqrt{s}$.

iii. the values of the S-wave non-local hadronic parameters $\{\alpha_k^{S0}\}$ are set identically to the ones of the longitudinal P-wave polarisation, $\{\alpha_k^0\}$.[17]

iv. the invariant B mass fit signal parameters μ, σ_1, σ_2, f_{CB}, α and n are defined separately for the two q^2 regions and their values are set as in Table 19;

v. all the background parameters (see Eq. 75) are defined separately for the two q^2 regions, the slope of the exponential function for the $m_{K\pi\mu\mu}$ mass fit is set as in Table 19, while the remaining polynomial coefficients are set as in Table 21 and Eq. 74;[18]

This extended version of the framework introduces 48 new parameters[19] with respect to the *signal-only* configuration of chapter 3.

An extended maximum likelihood fit is then performed to these simulated samples. The truncation of the S-wave non-local hadronic functions $\mathcal{H}_{S0}$ is aligned with the corresponding P-wave components, where the analytical expansion is curtailed at order K for a given fit model assumption $\mathcal{H}_\lambda[z^K]$. In addition to the Gaussian constraints applied on the nuisance parameters detailed in Sec. 3.1 and on the theoretical and experimental pseudo-observables discussed in Sec.3.2.2, further constraints are imposed on the scalar form factors. These are Gaussian constrained based on the results of Ref. [168] with the uncertainties enlarged by a factor of 3 in order to take into account differences between $B^0 \to K_0^*$ and $B^+ \to K^+$ dynamics. This conservative choice is found to have a negligible impact on the fit results, due to the small S-wave component that contributes to the decay. This assumption has been tested by enlarging the considered $K\pi$ mass window to [644, 1200] MeV$/c^2$, in which case the higher relative contribution of the S-wave induces strong instability to the fit. Note that a wider $K\pi$ mass region implies also a more abundant background contamination. All the remaining S-wave parameters are free-floated in the fit. Moreover, concerning the signal parametrisation of the B invariant mass, the mean and the two widths of the double Crystal Ball are floated while the tail parameters and the relative fraction f_{CB} are kept fixed. A good stability in the determination of the tail parameters of the double Crystal Ball is only possible with a large

---

[17] Note that this condition is employed only for the production of the ensembles, no assumptions are implied on the S-wave non-local hadronic parameters $\{\alpha_k^{S0}\}$ during the fitting procedure.

[18] Note that the values of the angular coefficients of Table 21 result from the angular fit to the control mode $B^0 \to J/\psi(1S)K^{*0}$. In general, the background parametrisation is expected to vary for the different kinematic domains, however, the values of these coefficients are used exclusively as baseline for the production of the ensembles and do not undermine the validity of this study.

[19] The number of parameters required to describe the effect of the S-wave contribution depends on the truncation of the analytical expansion of the non-local hadronic matrix elements. The count of 48 parameters reported in the text assumes the baseline configuration of $\mathcal{H}_\lambda[z^2]$.

number of events. For this reason f_{CB}, α and n are bounded to the result of the fit to the control channel. Finally, all the background parameters are free-floated.

7.4.2 *Experimental sensitivity to $\mathcal{C}_9$ and $\mathcal{C}_{10}$*

Ensembles of pseudoexperiments are generated with the described configuration and analysed with the $\mathcal{H}_\lambda[z^2]$ fit model assumption. In order to examine the impact of each of the experimental effects, the expected sensitivity has been studied step by step. First, the effect of the S-wave and background contributions has been investigated separately, afterwards the sensitivity of the full amplitude fit has been examined.

For illustration, Figs. 69 and 70 show the distributions of the events for one pseudoexperiment of reference, including acceptance, background and S-wave contributions, together with the projections of the fit results in the $1.1 < q^2 < 8.0\ \text{GeV}^2/c^4$ and $11.0 < q^2 < 12.5\ \text{GeV}^2/c^4$ regions, respectively.

Table 23 shows the sensitivity to the pair of Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ obtained for the $\texttt{BMP}_{\mathcal{C}_9}$scenario. The fit results are reported for different stages of the amplitude fit together with the signal-only scenario of chapter 3 as a comparison. A non-negligible bias is observed in the determination of both $\mathcal{C}_9$ and $\mathcal{C}_{10}$, though, its size is found to be of the same order for all the tested configuration, proving a valuable stability of the method against all the experimental effects. This intrinsic model bias is found to disappear for large datasets (within the expected statistics at the LHCb Upgrade), even including all the experimental effects, proving its statistical origin.[20] Secondly, it is interesting to notice that the statistical sensitivity to the Wilson coefficients obtained performing a "realistic" experimental analysis is diluted of only approximately 10 % with respect to the signal-only scenario studied in chapter 3.

Finally, the sensitivity expected for different NP scenarios is investigated. Table 24 shows the fit results obtained for the $\texttt{BMP}_{\mathcal{C}_9}$, $\texttt{BMP}_{\mathcal{C}_{9,10}}$and $\texttt{SM}$ benchmark points. Interestingly, the statistical uncertainty is found to scale approximately linearly with the generated values of the Wilson coefficients, *e.g.* $\sigma(\mathcal{C}_9^{\texttt{SM}})/\mathcal{C}_9^{\texttt{SM}} \simeq \sigma(\mathcal{C}_9^{\texttt{BMP}_{\mathcal{C}_9}})/\mathcal{C}_9^{\texttt{BMP}_{\mathcal{C}_9}} \simeq 7.3\%$, resulting in a more precise measurement in case of a stronger NP contribution. Figure 71 shows the two-dimensional $\mathcal{C}_9$-$\mathcal{C}_{10}$ parameter space for the considered scenarios. The fit results have been corrected for the observed bias and a systematic uncertainty has been added in quadrature

---

[20] A detailed study on the dependence of the bias as function of the number of events is given in appendix A.1.2 for the signal-only scenario.

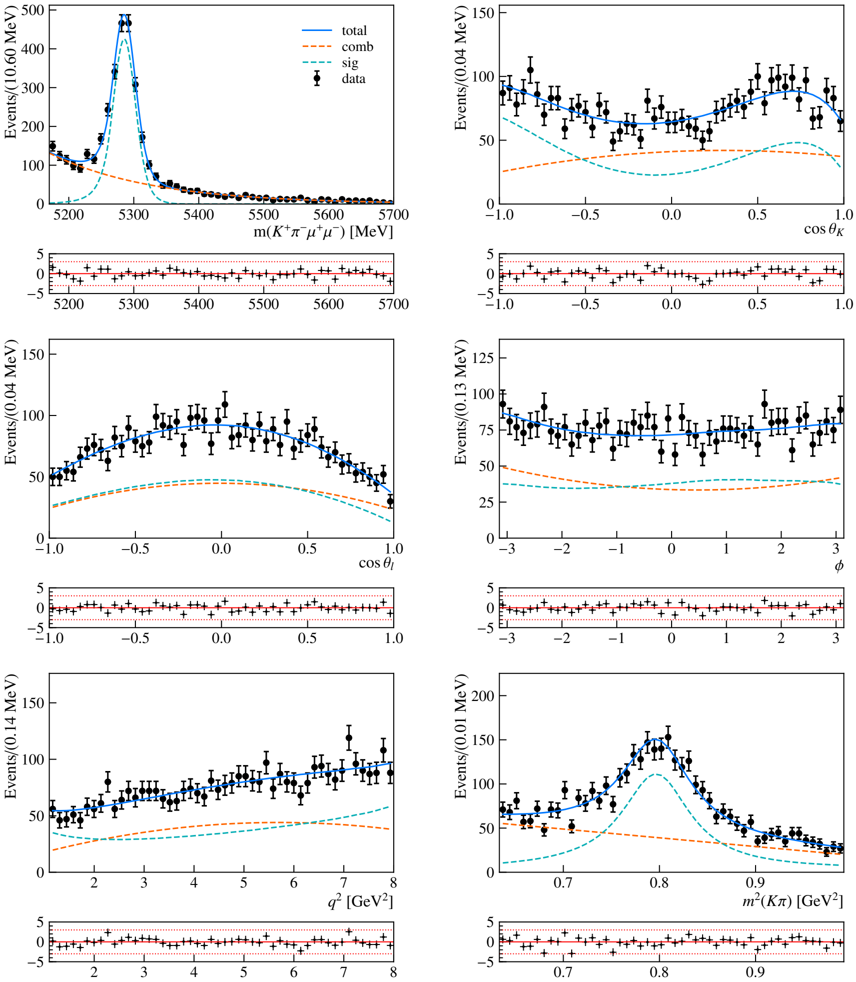


Figure 69: Distributions of events and projections of the fit result in the region $1.1 < q^2 < 8.0$ GeV$^2/c^4$ for one generated pseudoexperiment.

to account for this effect.[21] The expected sensitivity varies between 2.5 and 3.5 σ for the $\mathtt{BMP}_{\mathcal{C}_{9,10}}$ and $\mathtt{BMP}_{\mathcal{C}_9}$ scenarios, respectively. Note that a clear observation of NP that exceeds 5 σ in a single measurement can be achieved for NP scenarios that match the central values of $R_{K(K^*)}$, *e.g.* $\mathcal{C}_9^{\mathrm{NP}} \sim -1.7$ [94].

7.4.3 *Systematic uncertainties*

Two major sources of systematic uncertainties can be foreseen

---

[21] The systematic uncertainty that accounts for the observed bias in the estimators of the parameters has been applied to the results of Fig. 71 for illustration purpose, in order to visualise the limited impact of such effect on the sensitivity of the measurement. The use of the Feldman-Cousins method [169] is foreseen for an appropriate determination of the central values and uncertainties of the parameters of interest.

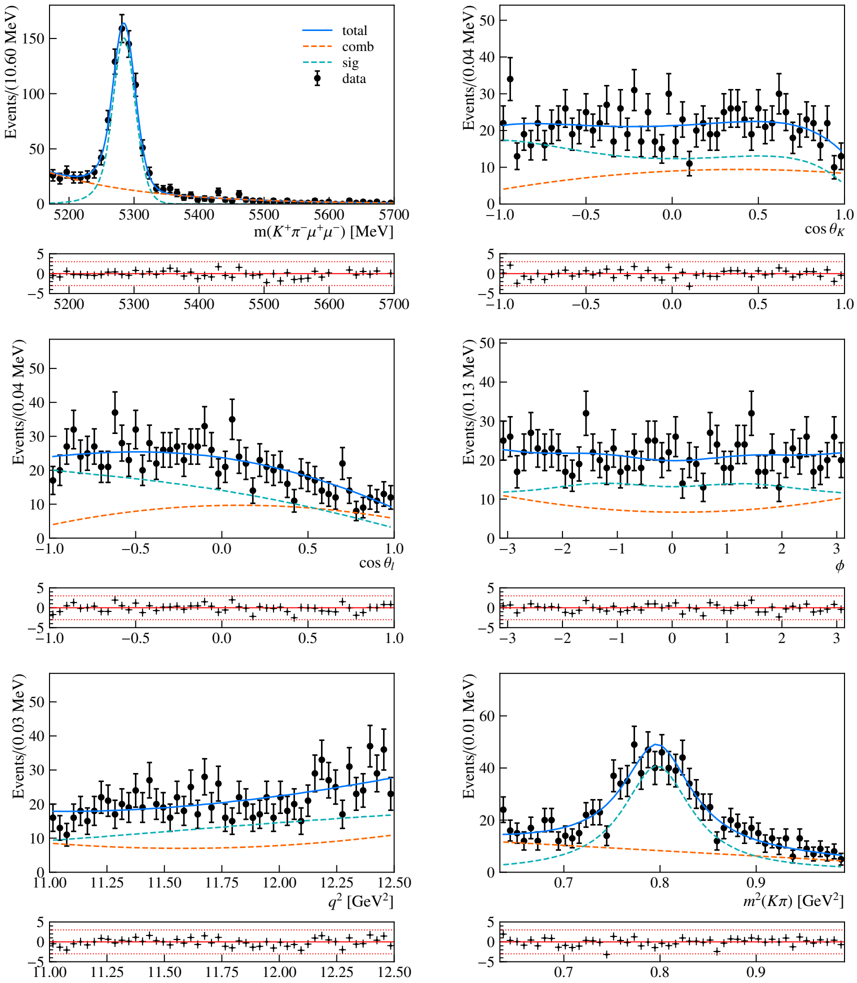


Figure 70: Distributions of events and projections of the fit result in the region $11.0 < q^2 < 12.5\ \mathrm{GeV}^2/c^4$ for one generated pseudoexperiment.

i. the entity of the bias observed in the estimators of $\mathcal{C}_9$ and $\mathcal{C}_{10}$ is non-negligible, between 25 and 30% of the statistical uncertainty, however, its statistical nature allows to apply the Feldman-Cousins method [169] in order to correct for the observed effect;

ii. a model dependence is introduced by the truncation of the analytical expansion of the non-local hadronic contributions. A statistical approach can be used to determine at which order to curtail the functions $\mathcal{H}_\lambda$; the order of the expansion is progressively increased until the improvement of the likelihood is statistically significant,[22] this determines the highest order at which the fit is sensitive to, given the analysed dataset. Once defined that, the next order can be used to assign a systematic considering eventual difference in the fitted central value of the

---

[22] Tipically, the threshold of $2\Delta\mathrm{NLL} > 9$ is used as discriminant.

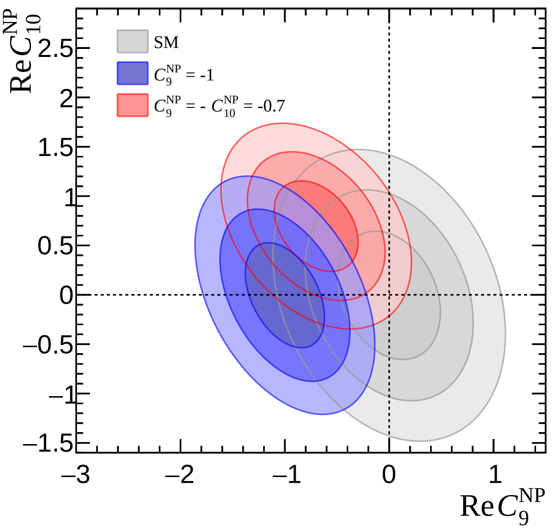


Figure 71: Two-dimensional sensitivity scans for the pair of Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ for the $\mathtt{BMP}_{\mathcal{C}_9}$, $\mathtt{BMP}_{\mathcal{C}_{9,10}}$ and `SM` scenarios obtained with fits with $\mathcal{H}_\lambda[z^2]$ and including all the experimental effects. The contours correspond to 1, 2, 3 σ contours obtained combining the statistical and bias-related systematic uncertainties.

Table 23: Fit results for the pair of Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ obtained from fits with $\mathcal{H}_\lambda[z^2]$ for the $\mathtt{BMP}_{\mathcal{C}_9}$ scenario.

$\mathcal{C}_9^{\mathrm{NP}}$ mean	$\mathcal{C}_9^{\mathrm{NP}}$ sigma	$\mathcal{C}_{10}^{\mathrm{NP}}$ mean	$\mathcal{C}_{10}^{\mathrm{NP}}$ sigma	correlation $\mathcal{C}_9^{\mathrm{NP}}$- $\mathcal{C}_{10}^{\mathrm{NP}}$
Signal only (P-wave)				
-0.96 ± 0.01	0.22 ± 0.01	0.05 ± 0.01	0.29 ± 0.01	-0.52 ± 0.03
P + S-waves				
-0.94 ± 0.01	0.23 ± 0.01	0.07 ± 0.02	0.31 ± 0.01	-0.54 ± 0.04
P-wave + Acc. + Bkg.				
-0.96 ± 0.01	0.23 ± 0.01	0.09 ± 0.02	0.33 ± 0.01	-0.45 ± 0.05
Full fit				
-0.93 ± 0.01	0.24 ± 0.01	0.08 ± 0.01	0.34 ± 0.01	-0.44 ± 0.03

parameters of interest. However, further studies are required to quantitatively estimate this systematic uncertainty.

Secondly, the theoretical uncertainties on the form factors play a determinant role in the proposed measurement. These have been proven to strongly impact the sensitivity to the Wilson coefficients, however, these uncertainties can be naturally investigated in the described framework by varying the corresponding parameters accordingly to their covariance matrix.

Table 24: Fit results for the pair of Wilson coefficients $\mathcal{C}_9$ and $\mathcal{C}_{10}$ obtained from fits with $\mathcal{H}_\lambda[z^2]$ to different physical scenarios.

	$\mathcal{C}_9^{\mathrm{NP}}$ mean	$\mathcal{C}_9^{\mathrm{NP}}$ sigma	$\mathcal{C}_{10}^{\mathrm{NP}}$ mean	$\mathcal{C}_{10}^{\mathrm{NP}}$ sigma	correlation $\mathcal{C}_9^{\mathrm{NP}}$- $\mathcal{C}_{10}^{\mathrm{NP}}$
`SM`	0.08 ± 0.01	0.31 ± 0.01	0.09 ± 0.02	0.41 ± 0.01	-0.26 ± 0.04
$\mathtt{BMP}_{\mathcal{C}_9}$	-0.93 ± 0.01	0.24 ± 0.01	0.08 ± 0.01	0.34 ± 0.01	-0.44 ± 0.03
$\mathtt{BMP}_{\mathcal{C}_{9,10}}$	-0.63 ± 0.01	0.26 ± 0.01	0.76 ± 0.01	0.30 ± 0.01	-0.35 ± 0.04

Finally, a longer list of minor systematic uncertainties must be considered. These are in common with the binned angular analysis [79] and consist of all the effects that can alter the mass and/or angular distributions of either the signal or background candidates. These contributions are expected to play a marginal role and can be summarised as

i. acceptance parametrisation - associated to the parametrisation (orders of the Legendre polynomials) that is used to describe the acceptance function;

ii. data-simulation differences - associated to the correction applied to the simulation sample from which the acceptance is determined;

iii. acceptance statistical uncertainty - resulting from the limited size of the simulation sample from which it is determined;

iv. $m_{K\pi}$ model - associated to the choice of the LASS parametrisation in order to describe the S-wave contribution;

v. background model - associated to the choice of the background parametrisation (orders of the Chebyshev polynomials);

vi. peaking backgrounds - due to possible sources of peaking background surviving the selection that can modify the angular distributions;

vii. $m_{K\pi\mu\mu}$ model - associated to the tail parameters of the double Crystal Ball that are determined by the mass fit to the control channel;

viii. the uncertainty on the branching fraction of the calibration sample $B^0 \to J/\psi(1S)K^{*0}$.

All these sources can be investigated using pseudoexperiments in which one or more parameters are varied, in order to render statistical fluctuations negligible large statistical samples are employed. The parameters of interest are determined from these simulated ensembles using either the nominal or the systematically varied models and eventual differences are assigned as systematic uncertainties.

7.4.4 *Conclusion and prospects*

In conclusion, the experimental aspects required for an amplitude fit to $B^0 \to K^{*0}\mu^+\mu^-$ decays have been studied. These include the extension of the formalism to the S-wave component, the parametrisation of the signal acceptance and the introduction of combinatorial background. The proposed method is found to be robust against all these experimental effects, which mildly dilute the expected sensitivity of the measurement.

Nevertheless, a sizeable bias in the determination of the Wilson coefficients is observed for the currently analysed dataset and, together with the model-dependence introduced by the truncation of the non-local hadronic functions, represent the largest sources of systematic uncertainties of this measurement.

Finally, the expected sensitivity foreseen for the analysis of the 5 fb^{-1} dataset collected by the LHCb experiment during Run-I and part of Run-II is estimated to be approximately 2.5 and $3.5\,\sigma$ for the two benchmark-like NP scenarios. These results complete the studies presented in chapter 3 and provide a "realistic" sensitivity of the proposed measurement.

The presented analysis is currently being performed at LHCb. Given the high profile of the measurement, a composite procedure for the validation of the results is foreseen. Thanks to the large amount of information that can be extracted from the amplitude fit, an unbinned determination of the angular observables and branching ratio can be achieved and represent extremely valuable information independent on the main systematic uncertainties of the analysis. Furthermore, one of the interesting aspects of this measurement is the prospect for a comparison with alternative approaches that use an empirical modelling of the non-local hadronic contribution based on the coherent sum of resonant amplitudes [85]. The comparison between these two complementary approaches is determinant for the comprehension of the q^2 dependence of the decay amplitudes and will provide a further understanding of the dynamics of the decay.

8

CONCLUSIONS

This dissertation has widely analysed $b \to s\ell^+\ell^-$ decays as probe of New Physics. Direct and indirect searches can exploit the rich phenomenology of flavour changing neutral current processes and strongly constrain NP models. The LHCb experiment provides an ideal environment to perform such precise measurements and to test the flavour structure of the Standard Model.

In this thesis, the study of the TT detector performance is presented, based on a *masked* algorithm that allow to extract the hit resolution and detector alignment and efficiency in an unbiased way. The misalignment of each TT sector is found to be below $7\,\mu$m, compared to the obtained resolution of approximately $80\,\mu$m, while the overall efficiency is found to be above 99%.

Secondly, a search for a new hypothetical scalar particle through the decay channel $B^+ \to K^+\chi(\to \mu^+\mu^-)$ is perform with the dataset collected by the LCHb experiment during Run-I. The sensitivity of the analysis is maximised splitting signal candidates in bins of decay-time, this allows to identify separate regions that are optimised in accordance with the different background regimes. No evidence of the signal is found and excluded limits are placed on the branching ratio of the process as function of the mass and lifetime of the new particle. These results set stringent constraints on the parameter space of inflaton models.

Furthermore, phenomenological studies to investigate the prospects for a novel theory-experimental approach to disentangle possible NP contributions from problematic non-local hadronic effects in $B^0 \to K^{*0}\mu^+\mu^-$ decays have been presented. Based on the proposed method, an adequate evaluation of the systematic uncertainties associated to the *charm-loop* is for the first time possible. In addition, the unbinned amplitude fit can be extended to the electron channel. In this case, thanks to the shared parametrisation of all the hadronic matrix elements, the determination of the differences of Wilson coefficients is found to be a clean and model-independent test of LFU, which combines in a single measurement all the benefits of previous analyses. Moreover, the parameters $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ are found to be extremely robust against poor resolution effects typical of experiments at hadronic machines. This feature is beneficial for the feasibility of the proposed measurement at the LHCb experiment.

Finally, sensitivity studies considering all experimental aspects of the analysis of $B^0 \to K^{*0}\mu^+\mu^-$ decays are performed to establish the "realistic" sensitivity of the proposed amplitude fit. Detector acceptance, S-wave and background contributions have been taken into account and are found not to significantly dilute the sensitivity to the parameters $\mathcal{C}_9$ and $\mathcal{C}_{10}$. Oppositely, the largest sources of systematic uncertainties are identified in the intrinsic model bias observed already in the signal-only configuration. Nevertheless, this measurement will provide further insight on the origin of the flavour anomalies and will be of paramount importance in understanding the nature of New Physics.

APPENDIX

A

ADDITIONAL MATERIAL FOR THE AMPLITUDE FIT TO $B^0 \to K^{*0}\ell\ell$

This appendix contains additional material relative to the proposed amplitude fit model described in chapter 3.

A.1 ADDITIONAL STUDIES ON $B^0 \to K^{*0}\mu\mu$ DECAYS

A.1.1 *On the determination of non-local parameters from data*

Despite most of the attention is directed to extract the Wilson coefficients as parameters of interest in the amplitude fit to $B^0 \to K^{*0}\mu^+\mu^-$ decays, genuine interest can be addressed to establish to which extent information concerning the non-local matrix elements can be inferred from experimental data. This question can be inspected in two possible ways

i. assuming the Standard Model predictions of the non-local hadronic contributions obtained in Ref. [59];

ii. abstaining from using any prior knowledge on those parameters and relying purely on the information extracted from $B^0 \to K^{*0}\mu^+\mu^-$ decays.

Under these hypotheses, a series of simulated pseudoexperiments is produced corresponding to the $\texttt{BMP}_{\mathcal{C}_9}$ scenario and the expected LHCb Run-II statistics. For the former condition, a multivariate Gaussian constraint on the non-local hadronic parameters $\{\alpha_k^{(\lambda)}\}$ is added to the fit, while these are freely floated when considering the second hypothesis. Both analyses have been performed with a truncation at the order of $\mathcal{H}_\lambda[z^2]$ on the modelling of the hadronic functions.

Figure 72 shows the fit results obtained for the non-local hadronic parameters under the two considered hypotheses. From these the following conclusion can be driven

i. The analysis with Gaussian constraints on the parameters $\{\alpha_k^{(\lambda)}\}$ is able to extract additional information on the non-local hadronic effects from the fit, *i.e.*, the obtained uncertainties are smaller than the corresponding Gaussian constraints. The uncertainties on the hadronic parameters scale by a factor between 0.5 and 0.8 for the expected statistics of the LHCb Run-II.

ii. When removing the constraints, the fit still converges and a disentangling of the non-local hadronic effects from $\mathcal{C}_9$ is still

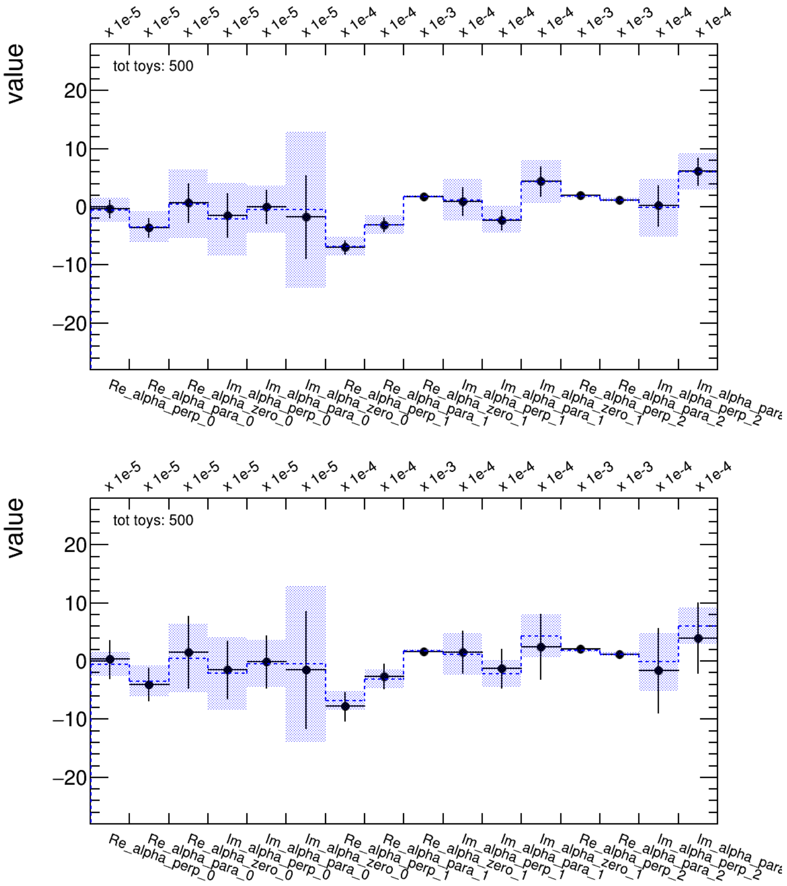


Figure 72: Fit results for the hadronic parameters $\{\alpha_k^{(\lambda)}\}$ (top) with and (bottom) without Gaussian constraints for the expected statistics of the LHCb Run-II. Points exhibit the mean and width of the fit results to the pseudoexperiments. The dashed blue line shows the parameter values used in the generation of the ensembles. The blue rectangular box represents the uncertainty on the prior prediction to each parameter and is overlaid as reference.

possible. For the expected statistics of the LHCb Run-II the uncertainties on the parameters $\{\alpha_k^{(\lambda)}\}$ obtained exclusively from the fit to $B^0 \to K^{*0}\mu^+\mu^-$ data are found to be comparable to the ones from the prior predictions.

A.1.2 *Study on the bias of the $\mathcal{C}_9$ estimator*

The capability of the model to disentangle non-local hadronic effects from $\mathcal{C}_9$ exclusively from the fit to $B^0 \to K^{*0}\mu^+\mu^-$ data, *i.e.* without any prior knowledge on the hadronic parameters, is remarkable. However, in order to completely assert the separation between $\mathcal{C}_9$ and the non-local hadronic effects, it is necessary to inquire eventual bias

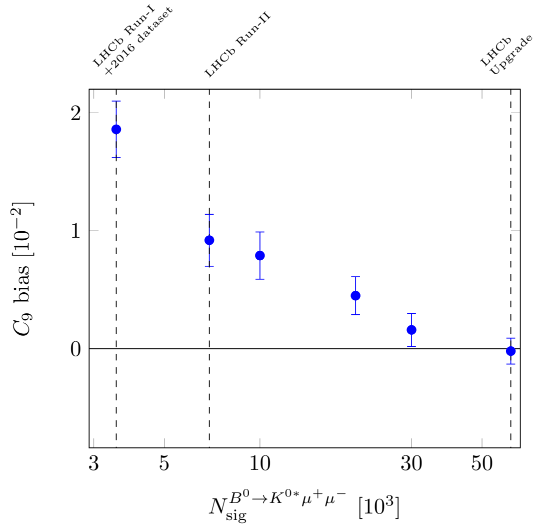


Figure 73: Bias of the estimator for $\mathcal{C}_9$ as function of the number of signal events for $\mathcal{H}_\lambda[z^2]$ fits.

in the estimator of the parameters of interest. Figure 73 shows the estimator of $\mathcal{C}_9$ as function of the number of signal events. A small positive bias is observed for datasets containing up to $\mathcal{O}(10^4)$ events and in found to disappear for larger datasets.

A.1.3 *Exploring the impact of the inputs from theory and hadronic decays*

An interesting aspect of the combined theory-experimental approach, consists in investigating the impact of additional pseudo-observables introduced in the fit. In particular, the following points are investigated

i. is it possible to perform a purely experimental analysis, *i.e.*, excluding the theoretical points at negative q^2 and relying only on semileptonic and hadronic decays?

ii. The pseudo-observables obtained for the hadronic decays currently constrain only two relative phase between the three polarisations. What is the impact of a hypothetical theory determination of the absolute phase of the hadronic decays and/or an increased precision of the relative phases?

To address the first question the fit is repeated removing the constraints on the non-local hadronic functions introduced by the theoretical calculation at negative q^2, and the stability of the fit is examined with different orders of the $\mathcal{H}_\lambda(z)$ expansion. Fig. 74 shows the result of the fit assuming $\mathcal{H}_\lambda[z^4]$, with and without the input from the theory points for the expected statistics at LHCb Upgrade and the $\mathtt{BMP}_{\mathcal{C}_9}$scenario. The results prove a strong model-bias, similarly to what is presented in section 3.2.1. Therefore, a purely experimental analysis

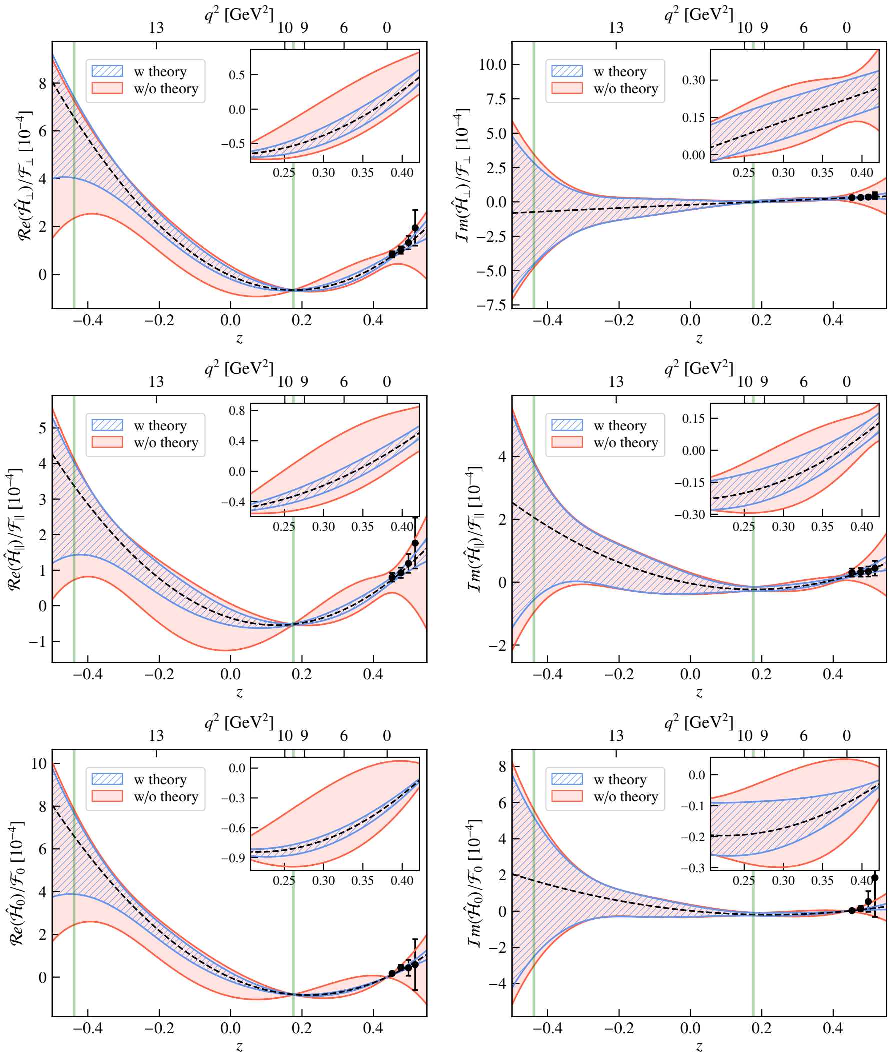


Figure 74: Results of the fits for the ratios $\hat{\mathcal{H}}_\lambda(z)/\mathcal{F}_\lambda(z)$ obtained with and without the theoretical points calculated at negative q^2. Fits correspond to the $\texttt{BMP}_{\mathcal{C}_9}$scenario, the expected statistics at LHCb Upgrade and $\mathcal{H}_\lambda[z^4]$ truncation of the series expansion. The vertical bands correspond to the $J/\psi(1S)$ and $\psi(2S)$ regions and the points to the theoretical inputs at negative q^2. The top right box of each plot zooms in the q^2 range between 1.1 and 9.0 GeV$^2/c^4$.

that combines information from semileptonic decays and hadronic $B^0 \to K^{*0}\{J/\psi(1S), \psi(2S)\}$ decays is currently not possible. The desired disentangling of the hadronic effect from possible New Physics contributions crucially relies on the theory inputs at negative q^2.

Possible benefits from hypothetical improvements to the constraints based on the hadronic decays are also inquired. First of all, the uncertainty on the non-local hadronic functions evaluated at the $J/\psi(1S)$ is extremely small - see Fig. 74. This is an intrinsic property of the analytical parametrisation, given by the change of one unit of π in

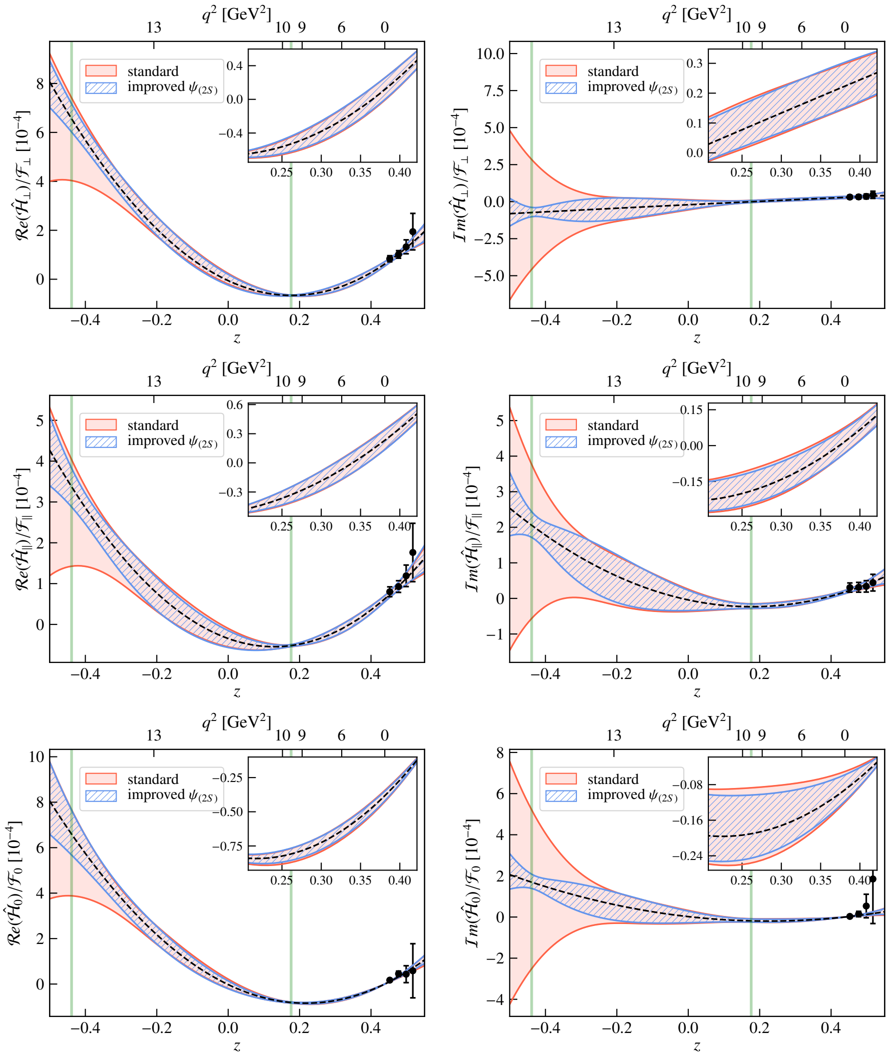


Figure 75: Results of the fits for the ratios $\hat{\mathcal{H}}_\lambda(z)/\mathcal{F}_\lambda(z)$ obtained with the current status of the theoretical and experimental knowledge, labelled "standard" in the legend, and assuming future improvements on the $B^0 \to K^{*0}\psi(2S)$ measurements as described in the text, labelled "improved". Fits correspond to the $\texttt{BMP}_{\mathcal{C}_9}$scenario, the expected statistics at LHCb Upgrade and $\mathcal{H}_\lambda[z^4]$ truncation of the series expansion. The vertical bands correspond to the $J/\psi(1S)$ and $\psi(2S)$ regions and the points to the theoretical inputs at negative q^2. The top right box of each plot zooms in the q^2 range between 1.1 and 9.0 GeV$^2/c^4$.

the phase of the amplitude when crossing the resonance pole.[1] We find that, for datasets corresponding to the expected statistics at LHCb Run-II, the impact of the pseudo-observables of the $J/\psi(1S)$ improves the determination of the $\mathcal{C}_9$ by approximately 15%, while it plays a minor role when increasing the z-expansion to higher orders. Further-

---

[1] This behaviour is typically represented with a circular motion in the Argand plane [170].

more, the fit is able to select the absolute phase of the $J/\psi(1S)$ with the same precision as the two relative phases.

The impact of the $\psi(2S)$ pseudo-observables on the combined fit is of great interest since these are weakly constrained, in particular the two relative phases [59]. This is inspected by including a hypothetical constraint on the absolute phase of the $\psi(2S)$ and repeating the analysis with two configurations

i. the relative uncertainty of the absolute phase of the $\psi(2S)$ is assumed to be of the same order as the one of the relative phases;

ii. the uncertainties of all phases of the $\psi(2S)$ are reduced to reflect the uncertainties of the $J/\psi(1S)$.

In both cases the central value of the absolute phase of the $\psi(2S)$ is set to the prediction obtained from the set of parameters $\{\alpha_k^{(\lambda)}\}$ used for the production of the ensembles.

Fit results obtained assuming condition [i.] are found to be comparable with respect to the standard configuration, disproving the benefit of an absolute determination of the $\psi(2S)$ phase. On the other hand, Fig. 75 shows the results obtained for the non-local hadronic functions with the current available information and assuming the future improvements on the $\psi(2S)$ determination as in condition [ii.]. Both analyses are performed with datasets corresponding to the LHCb Upgrade expected statistics and assume $\mathcal{H}_\lambda[z^4]$ truncation. The benefits produced by the foreseen future improvements on the $\psi(2S)$ pseudo-observables are limited to the region of the $\psi(2S)$, while their impact on the determination of the Wilson coefficient $\mathcal{C}_9$ is negligible.

A.1.4 *Extended New Physics models: the primed Wilson coefficients*

Despite the most popular global fits to the B-anomalies prefer to insert New Physics in the $\mathcal{C}_9$ and $\mathcal{C}_{10}$ Wilson coefficients, it is also possible to have beyond the Standard Model contributions in the chirality-flipped $\mathcal{C}_9'$ and $\mathcal{C}_{10}'$. The primed Wilson coefficients, zero in the SM, are only weakly constrained by current measurements [94]. Thus, it is interesting to study the feasibility to directly extract also those parameters from a fit to data.

For this purpose, the parameter space is further extended to include $\mathcal{C}_9'$ and $\mathcal{C}_{10}'$ in the amplitude fit and all the studies described in Sec. 3.2 are repeated. Firstly, the dependence on the non-local hadronic effects is tested also for the primed Wilson coefficients. The fit is repeated with different orders of the analytical parametrisation. Figure 76 (left) shows the fit results, obtained purely from amplitude fit to $B^0 \to K^{*0}\mu^+\mu^-$ decays (without the inputs from theory and hadronic decays of section 3.2.2). The first interesting conclusion is that both primed Wilson coefficients are not affected by non-local hadronic effects, since

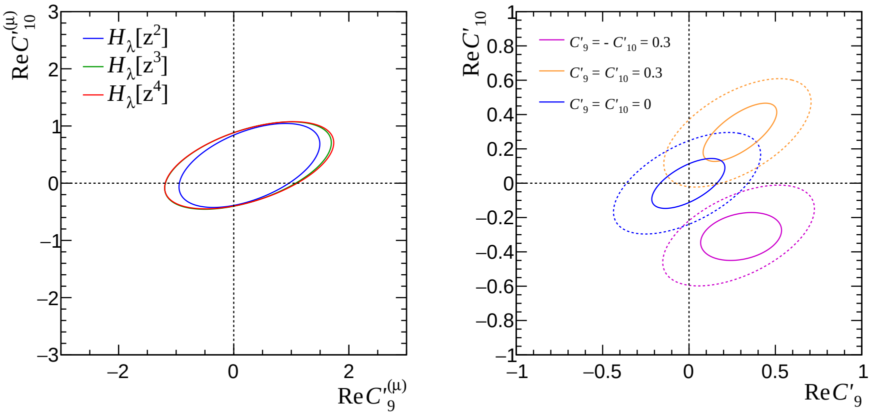


Figure 76: Two-dimensional sensitivity scans for the pair of Wilson coefficients $\mathcal{C}'_9$ and $\mathcal{C}'_{10}$. Left: different non-local hadronic parametrisation models are shown for a New Physics scenario with $\mathcal{C}'_9 = \mathcal{C}'_{10} = 0.3$ and the expected statistics after LHCb Run-II. Right: three different New Physics scenarios are considered for the LHCb Upgrade (dotted) $50\,\mathrm{fb}^{-1}$ and (solid) $300\,\mathrm{fb}^{-1}$ expected statistics. The contours correspond to $3\,\sigma$ statistical-only uncertainty bands.

their uncertainties clearly saturates with the order $\mathcal{H}_\lambda[z^3]$, *i.e.* the uncertainty resulting from fits with $\mathcal{H}_\lambda[z^3]$ and $\mathcal{H}_\lambda[z^4]$ is the same. With this baseline setup, it is possible to investigate the sensitivity for current and future experiments. As can be seen in Fig. 76, the sensitivity expected for the LHCb Run-II statistics does not allow any conclusive measurements on the right-handed $\mathcal{C}'_i$. On the other hand, convincing measurements of $\mathcal{C}'_9$ and $\mathcal{C}'_{10}$ will be possible starting from the LHCb Upgrades [$50\,\mathrm{fb}^{-1}$ - $300\,\mathrm{fb}^{-1}$] and will allow to extract the full physical information encoded in these decays.

A.2 ADDITIONAL STUDIES ON THE LFU TEST IN $B^0 \to K^{*0}\ell\ell$ DECAYS

A.2.1 *Exploring additional experimental scenarios*

It is interesting to extend the sensitivity studies based on the simultaneous fit to $B^0 \to K^{*0}\ell^+\ell^-$ decays presented in Sec. 3.3 to further explore future experimental scenarios. In particular, prospects for measurements of the newly proposed parameters $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ with the Belle-II experiment are investigated in the following. By virtue of the clean experimental environment and the advantageous electron resolution with respect to hadronic machines, experiments at B factories allow the investigation of the region between the $J/\psi(1S)$ and

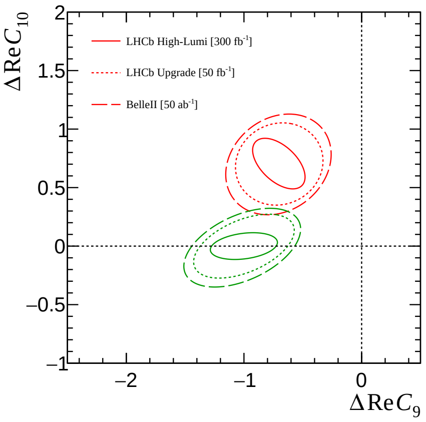


Figure 77: Two-dimensional sensitivity scans for the proposed parameters $\Delta\mathcal{C}_9$ and $\Delta\mathcal{C}_{10}$ for the two considered NP scenarios: (green) $\mathtt{BMP}_{\mathcal{C}_9}$ and (red) $\mathtt{BMP}_{\mathcal{C}_{9,10}}$. The contours correspond to $3\,\sigma$ statistical-only uncertainty bands expected for the (dashed) Belle II 50 ab^{-1} and LHCb Upgrade (dotted) $50\,\mathrm{fb}^{-1}$ and (solid) $300\,\mathrm{fb}^{-1}$ statistics.

the $\psi(2S)$ also for the electron channel [80]. Therefore, an additional kinematic region defined as

$$Q_2^{(e)}: \quad 11.0 \leq q^2 \leq 12.5\ \mathrm{GeV}^2/c^4 \qquad (92)$$

is included in the framework. The expected signal yields are extrapolated from Ref. [80] by scaling the luminosity expected for Belle-II [50 ab^{-1}]. Finally, the LHCb Upgrade [300 fb^{-1}] scenario foreseen for the LHCb high luminosity Phase-II program [97] is also examined.

Figure 77 shows the two-dimensional statistical-only significances for the LHCb/Belle upgrades given the designed luminosities. Both LHCb Upgrade [$50\,\mathrm{fb}^{-1}$] and Belle II expected sensitivities have comparable significance (within $8.0 - 10\,\sigma$), while LHCb High-Lumi results in an overwhelming significance. These exceptional datasets will give birth to a new era of precision measurements of Wilson coefficients.

B

S-WAVE PARAMETRISATION

B.1 MASS LINESHAPES

A relativistic Breit-Wigner form is often the most appropriate/convenient mass lineshape to describes isolated, non-overlapping resonances far from the threshold of additional decay channels. For a single resonance $R_J(\to K\pi)$, with orbital angular momentum J, the relativistic Breit-Wigner reads as

$$\begin{aligned} f_{BW}(m^2_{K\pi}) =& B'_{L_{K\pi}}(k,k_0,d)\cdot\left(\frac{k}{k_0}\right)^{L_{K\pi}}\cdot B'_{L_B}(p,p_0,d)\cdot\left(\frac{p}{p_0}\right)^{L_B} \\ &\times\frac{1}{m^2_{K\pi}-m^2_0-im_0\Gamma_J(m_{K\pi})} \end{aligned} \tag{93}$$

where m_0 is the pole-mass of the resonance, $k(p)$ is the momentum of the $K^+(R)$ in the rest frame of the $R(B)$ evaluated at a given $m_{K\pi}$ and the parameters k_0 and p_0 are the values of k and p evaluated at the pole-mass of the resonance. The orbital angular momentum barrier factors, $(k^L B'_L)$, must be taken into account for both the $B\to R_J X$ and $R_J\to K\pi$ parts of the decay. These involve the Blatt-Weisskopf functions [171] and account for spin-dependent effects in the conservation of the angular momentum, with L the necessary orbital angular momentum, which depends on the momentum of the decay products $k(p)$ - in the rest frame of the decaying particle - and on the size of the decaying particle given by the d constant, "meson radius". Thus, the orbital angular momentum between the resonance R and the bachelor particle(s) X in the B decay is denoted L_B, while $L_{K\pi}$ is the orbital angular momentum between the K^+ and π^- mesons in the decay of the resonance R. The relevant Blatt-Weisskopf barrier factors are [156]:

$$B'_{L=0}(k,k_0,d) = 1, \tag{94}$$

$$B'_{L=1}(k,k_0,d) = \sqrt{\frac{1+(k_0 d)^2}{1+(kd)^2}}. \tag{95}$$

Finally, the running width $\Gamma_J(m_{K\pi})$ is given by

$$\Gamma_J(m_{K\pi}) = \Gamma_0 B'^2_{L_{K\pi}}(k,k_0,d)\left(\frac{k}{k_0}\right)^{2L_{K\pi}+1}\left(\frac{m_0}{m_{K\pi}}\right), \tag{96}$$

where Γ_0 is resonance width.

For all $K\pi$ states of $K^*_J\to K\pi$, the daughters are spinless so $L_{K\pi}=J$ of the resonance/state. On the other hand, for $B^0\to[K^*_0][\mu^+\mu^-]$ the di-muon object may be in a vector state and therefore $L_B = 1$ in this case, while for $B^0 \to [K_1^*(892)^0][\mu^+\mu^-]$ there are three possible states, with $L_B = 0$ being the dominant.

In conclusion, for the P-wave component, $L_{K\pi} = 1$ and $L_B = 0$, and the $K\pi$ mass lineshape simplifies to the expression in Eq. 68, where the meson radius d is set to 1.6 $[\mathrm{GeV}/c]^{-1}$ [166]. On the other hand, for the S-wave component, the different spin configuration of the $K\pi$ system results in $L_B = 1$ and $L_{K\pi} = 0$ and only the angular momentum barrier factor $(p/p_{1430})B'_{L_B}$ survives in Eq. 70.

B.2 SCALAR FORM FACTORS

The scalar form factors f_+, f_T and f_0 introduced in Sec. 7.3.1 are determined by the following matrix elements,

$$\begin{aligned}\langle K_0^*(k)|\bar{s}\gamma_\mu\gamma_5 b|B(p)\rangle &= \left[(p+k)_\mu - \frac{M_B^2 - M_{K_0^*}^2}{q^2} q_\mu\right] f_+(q^2) \\ &+ \frac{M_B^2 - M_{K_0^*}^2}{q^2} q_\mu f_0(q^2) \qquad (97)\end{aligned}$$

and

$$\langle K_0^*(k)|\bar{s}\sigma_{\mu\nu}\gamma_5 q^\nu b|B(p)\rangle = i\left[(M_B^2 - M_{K_0^*}^2)q_\mu - q^2(p+k)_\mu\right]\frac{M_B}{q^2} f_T(q^2)\,. \qquad (98)$$

These correspond to Eq.10 of Ref. [167] except for the transformation

$$f_T \mapsto \frac{q^2}{M_B(M_B + M_{K_0^*})} f'_T\,, \qquad (99)$$

where f'_T is the one defined in Ref. [167] and commonly used in the literature.

B.2.1 *Scalar form factors parametrisation*

The form factors definition above is valid for a generic scalar → scalar transition, however, the exact evaluation of the f_+, f_T and f_0 functions depend on the kind of mesons involved in the decay. Unfortunately, the current theoretical knowledge on the $B^0 \to K_0^*$ form factors is very limited. For this reason, the scalar form factors required for the

Table 25: Results of z-expansion fits of the $B \to K$ form factors as from Ref. [168].

Coefficient	Value
b_0^+	$0.466(14)$
b_1^+	$-0.89(13)$
b_2^+	$-0.21(55)$
b_0^0	$0.292(10)$
b_1^0	$0.28(12)$
b_2^0	$0.15(44)$
b_0^T	$0.460(19)$
b_1^T	$-1.09(24)$
b_2^T	$-1.11(97)$

insertion of the S-wave in the amplitude fit are taken from $B^+ \to K^+$ transition and follow the parametrisation below [168]

$$f_{+,T'}(q^2) = \frac{1}{1 - q^2/M^2_{B_s^*}} \sum_{m=0}^{K-1} b_m^{+,T} \left[z^m - (-1)^{m-K} \frac{m}{K} z^K \right] \quad (100)$$

$$f_0(q^2) = \frac{1}{1 - q^2/M^2_{B_{s0}^*}} \sum_{m=0}^{K-1} b_m^0 z^m \quad (101)$$

with $M_{B_s^*} = 5415.4$ MeV and $M_{B_{s0}^*} = 5711$ MeV. As in Ref. [172] $K = 3$ is chosen for the expansion, while the values of the coefficients b_i are taken from Ref. [168] and reported in Table 25.

BIBLIOGRAPHY

[1] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[2] M. Chrzaszcz *et al.*, *Prospects for disentangling long- and short-distance effects in the decays $B \to K^* \mu^+ \mu^-$*, `arXiv:1805.06378`.

[3] A. Mauri, N. Serra, and R. Silva Coutinho, *Towards establishing Lepton Flavour Universality violation in $\bar{B} \to \bar{K}^* \ell^+ \ell^-$ decays*, `arXiv:1805.06401`.

[4] LHCb collaboration, R. Aaij *et al.*, *Search for long-lived scalar particles in $B^+ \to K^+ \chi(\mu^+ \mu^-)$ decays*, Phys. Rev. **D95** (2017) 071101, `arXiv:1612.07818`.

[5] Belle, K. Abe *et al.*, *Observation of a new narrow charmonium state in exclusive B+- —> K+- pi+ pi- J / psi decays*, in *Proceedings, 21st International Symposium on Lepton and Photon Interactions at High Energies (LP 03): Batavia, ILL, August 11-16, 2003*, 2003, `arXiv:hep-ex/0308029`.

[6] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi\, p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi\, p K^-$ decays*, Phys. Rev. Lett. **115** (2015) 072001, `arXiv:1507.03414`.

[7] J. J. Aubert *et al.*, *Experimental observation of a heavy particle j*, Phys. Rev. Lett. 33 (1974) 1404.

[8] J.-E. Augustin *et al.*, *Discovery of a narrow resonance in e^+e^- annihilation*, Phys. Rev. Lett. 33 (1974) 1406.

[9] UA1, G. Arnison *et al.*, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV*, Phys. Lett. **B122** (1983) 103, [,611(1983)].

[10] UA2, M. Banner *et al.*, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*, Phys. Lett. **B122** (1983) 476, [,7.45(1983)].

[11] UA1, G. Arnison *et al.*, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c**2 at the CERN SPS Collider*, Phys. Lett. **B126** (1983) 398, [,7.55(1983)].

[12] UA2, P. Bagnaia *et al.*, *Evidence for Zo —> e+ e- at the CERN anti-p p Collider*, Phys. Lett. **B129** (1983) 130, [,7.69(1983)].

[13] S. W. Herb *et al.*, *Observation of a dimuon resonance at 9.5 gev in 400-gev proton-nucleus collisions*, Phys. Rev. Lett. **39** (1977) 252.

[14] CDF Collaboration, F. Abe *et al.*, *Observation of top quark production in $\overline{p}p$ collisions with the collider detector at fermilab*, Phys. Rev. Lett. **74** (1995) 2626.

[15] D0 Collaboration, S. Abachi *et al.*, *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) 2632.

[16] ATLAS, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1, `arXiv:1207.7214`.

[17] CMS, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30, `arXiv:1207.7235`.

[18] L. J. Garay, *Quantum gravity and minimum length*, Int. J. Mod. Phys. **A10** (1995) 145, `arXiv:gr-qc/9403008`.

[19] J. Conrad, *Indirect Detection of WIMP Dark Matter: a compact review*, in *Interplay between Particle and Astroparticle physics London, United Kingdom, August 18-22, 2014*, 2014, `arXiv:1411.1925`.

[20] I. I. Bigi and A. I. Sanda, *CP violation*, , [Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.9,1(2009)].

[21] Wikipedia contributors, *Standard model — Wikipedia, the free encyclopedia*, `https://en.wikipedia.org/wiki/Standard_Model`, 2004. [Online; accessed 22-Oct-2018].

[22] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652.

[23] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[24] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531, [,648(1963)].

[25] R. Essig *et al.*, *Working Group Report: New Light Weakly Coupled Particles*, in *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*, 2013, `arXiv:1311.0029`.

[26] F. Bezrukov and D. Gorbunov, *Light inflaton Hunter's Guide*, JHEP **1005** (2010) 010, `arXiv:0912.0390`.

[27] F. Bezrukov and D. Gorbunov, *Relic Gravity Waves and 7 keV Dark Matter from a GeV scale inflaton*, Phys. Lett. **B736** (2014) 494, `arXiv:1403.4638`.

[28] A. H. Guth, *The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems*, Phys. Rev. **D23** (1981) 347, [Adv. Ser. Astrophys. Cosmol.3,139(1987)].

[29] M. P. Hertzberg and J. Karouby, *Generating the Observed Baryon Asymmetry from the Inflaton Field*, Phys. Rev. **D89** (2014) 063523, `arXiv:1309.0010`.

[30] M. P. Hertzberg and J. Karouby, *Baryogenesis from the Inflaton Field*, Phys. Lett. **B737** (2014) 34, `arXiv:1309.0007`.

[31] V. A. Rubakov and M. E. Shaposhnikov, *Electroweak baryon number nonconservation in the early universe and in high-energy collisions*, Usp. Fiz. Nauk **166** (1996) 493, `arXiv:hep-ph/9603208`, [Phys. Usp.39,461(1996)].

[32] J. R. Ellis, M. K. Gaillard, and D. V. Nanopoulos, *A Phenomenological Profile of the Higgs Boson*, Nucl. Phys. **B106** (1976) 292.

[33] F. Bezrukov and D. Gorbunov, *Light inflaton after LHC8 and WMAP9 results*, JHEP **07** (2013) 140, `arXiv:1303.4395`.

[34] F. Bergsma *et al.*, *Search for axion-like particle production in 400 gev proton-copper interactions*, Physics Letters B **157** (1985) 458 .

[35] NA48/2, J. R. Batley *et al.*, *Searches for lepton number violation and resonances in $K^{\pm} \to \pi\mu\mu$ decays*, Phys. Lett. **B769** (2017) 67, `arXiv:1612.04723`.

[36] LHCb collaboration, R. Aaij *et al.*, *Search for hidden-sector bosons in $B^0 \to K^{*0}\mu^+\mu^-$ decays*, Phys. Rev. Lett. **115** (2015) 161802, `arXiv:1508.04094`.

[37] B. Batell, M. Pospelov, and A. Ritz, *Multi-lepton Signatures of a Hidden Sector in Rare B Decays*, Phys. Rev. **D83** (2011) 054005, `arXiv:0911.4938`.

[38] CKMfitter Group, J. Charles *et al.*, *CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*, Eur. Phys. J. **C41** (2005) 1, `arXiv:hep-ph/0406184`, and updates at http://www.slac.stanford.edu/xorg/ckmfitter/.

[39] UTfit, M. Bona *et al.*, *Model-independent constraints on $\Delta F = 2$ operators and the scale of new physics*, JHEP **03** (2008) 049, `arXiv:0707.0636`, and updates at http://www.utfit.org/.

[40] CMS and LHCb collaborations, V. Khachatryan *et al.*, *Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data*, Nature **522** (2015) 68, `arXiv:1411.4413`.

[41] R. Gauld, F. Goertz, and U. Haisch, *An explicit Z'-boson explanation of the $B \to K^*\mu^+\mu^-$ anomaly*, JHEP **01** (2014) 069, `arXiv:1310.1082`.

[42] A. J. Buras and J. Girrbach, *Left-handed Z' and Z FCNC quark couplings facing new $b \to s\mu^+\mu^-$ data*, JHEP **12** (2013) 009, `arXiv:1309.2466`.

[43] W. Altmannshofer and D. M. Straub, *New Physics in $B \to K^*\mu\mu$?*, Eur. Phys. J. **C73** (2013) 2646, `arXiv:1308.1501`.

[44] A. Crivellin *et al.*, *Lepton-flavour violating B decays in generic Z' models*, Phys. Rev. **D92** (2015) 054013, `arXiv:1504.07928`.

[45] G. Hiller and M. Schmaltz, *R_K and future $b \to s\ell\ell$ physics beyond the standard model opportunities*, Phys. Rev. **D90** (2014) 054014, `arXiv:1408.1627`.

[46] S. Biswas, D. Chowdhury, S. Han, and S. J. Lee, *Explaining the lepton non-universality at the LHCb and CMS within a unified framework*, JHEP **02** (2015) 142, `arXiv:1409.0882`.

[47] B. Gripaios, M. Nardecchia, and S. A. Renner, *Composite leptoquarks and anomalies in B-meson decays*, JHEP **05** (2015) 006, `arXiv:1412.1791`.

[48] M. Bordone, C. Cornella, J. Fuentes-Martin, and G. Isidori, *A three-site gauge model for flavor hierarchies and flavor anomalies*, Phys. Lett. **B779** (2018) 317, `arXiv:1712.01368`.

[49] M. Bordone, C. Cornella, J. Fuentes-Martín, and G. Isidori, *Low-energy signatures of the* PS^3 *model: from B-physics anomalies to LFV*, `arXiv:1805.09328`.

[50] A. Azatov *et al.*, *Anatomy of $b \to c\tau\nu$ anomalies*, `arXiv:1805.03209`.

[51] D. Marzocca, *Addressing the B-physics anomalies in a fundamental Composite Higgs Model*, JHEP **07** (2018) 121, `arXiv:1803.10972`.

[52] E. Megias, M. Quiros, and L. Salas, *Lepton-flavor universality limits in warped space*, Phys. Rev. **D96** (2017) 075030, `arXiv:1707.08014`.

[53] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, Rev. Mod. Phys. **68** (1996) 1125, `arXiv:hep-ph/9512380`.

[54] A. Petermann, *La normalisation des constantes dans la théorie des quantaNormalization of constants in the quanta theory*, Helv. Phys. Acta **26** (1953) 499.

[55] M. Gell-Mann and F. E. Low, *Quantum electrodynamics at small distances*, Phys. Rev. **95** (1954) 1300.

[56] C. Bobeth, G. Hiller, and D. van Dyk, *General analysis of $\bar{B} \to \bar{K}^{(*)}\ell^+\ell^-$ decays at low recoil*, Phys. Rev. **D87** (2013) 034016, `arXiv:1212.2321`, [Phys. Rev.D87,034016(2013)].

[57] U. Egede *et al.*, *New observables in the decay mode $\bar{B}_d \to \bar{K}^{*0}l^+l^-$*, JHEP **11** (2008) 032, `arXiv:0807.2589`.

[58] F. Kruger and J. Matias, *Probing new physics via the transverse amplitudes of $B^0 \to K^{*0}(\to K^-\pi^+)l^+l^-$ at large recoil*, Phys. Rev. **D71** (2005) 094009, `arXiv:hep-ph/0502060`.

[59] C. Bobeth, M. Chrzaszcz, D. van Dyk, and J. Virto, *Long-distance effects in $B \to K^*\ell\ell$ from Analyticity*, `arXiv:1707.07305`.

[60] D. Becirevic, V. Lubicz, and F. Mescia, *An Estimate of the B —> K* gamma form factor*, Nucl. Phys. **B769** (2007) 31, `arXiv:hep-ph/0611295`.

[61] R. R. Horgan, Z. Liu, S. Meinel, and M. Wingate, *Lattice QCD calculation of form factors describing the rare decays $B \to K^*\ell^+\ell^-$ and $B_s \to \phi\ell^+\ell^-$*, Phys. Rev. **D89** (2014) 094501, `arXiv:1310.3722`.

[62] P. Ball and V. M. Braun, *Exclusive semileptonic and rare B meson decays in QCD*, Phys. Rev. **D58** (1998) 094016, `arXiv:hep-ph/9805422`.

[63] A. Khodjamirian, T. Mannel, and N. Offen, *Form-factors from light-cone sum rules with B-meson distribution amplitudes*, Phys. Rev. **D75** (2007) 054013, `arXiv:hep-ph/0611193`.

[64] C. Bobeth, G. Hiller, and D. van Dyk, *The Benefits of $\bar{B} \to \bar{K}^*l^+l^-$ Decays at Low Recoil*, JHEP **07** (2010) 098, `arXiv:1006.5013`.

[65] P. Ball and R. Zwicky, *$B_{d,s} \to \rho, \omega, K^*, \phi$ decay form-factors from light-cone sum rules revisited*, Phys. Rev. **D71** (2005) 014029, `arXiv:hep-ph/0412079`.

[66] C. G. Boyd, B. Grinstein, and R. F. Lebed, *Model independent extraction of |V(cb)| using dispersion relations*, Phys. Lett. **B353** (1995) 306, `arXiv:hep-ph/9504235`.

[67] C. Bourrely, I. Caprini, and L. Lellouch, *Model-independent description of B —> pi l nu decays and a determination of |V(ub)|*, Phys. Rev. **D79** (2009) 013008, `arXiv:0807.2722`, [Erratum: Phys. Rev.D82,099902(2010)].

[68] C. Bobeth, G. Hiller, and D. van Dyk, *More Benefits of Semileptonic Rare B Decays at Low Recoil: CP Violation*, JHEP **07** (2011) 067, `arXiv:1105.0376`.

[69] C. Bobeth, G. Hiller, D. van Dyk, and C. Wacker, *The Decay $B \to K\ell^+\ell^-$ at Low Hadronic Recoil and Model-Independent $\Delta B = 1$ Constraints*, JHEP **01** (2012) 107, `arXiv:1111.2558`.

[70] S. Descotes-Genon, J. Matias, M. Ramon, and J. Virto, *Implications from clean observables for the binned analysis of $B \to K^*\mu^+\mu^-$ at large recoil*, JHEP **01** (2013) 048, `arXiv:1207.2753`.

[71] J. Matias, F. Mescia, M. Ramon, and J. Virto, *Complete Anatomy of $\bar{B}_d \to \bar{K}^{*0}(\to K\pi)l^+l^-$ and its angular distribution*, JHEP **04** (2012) 104, `arXiv:1202.4266`.

[72] G. Hiller and F. Kruger, *More model-independent analysis of $b \to s$ processes*, Phys. Rev. **D69** (2004) 074020, `arXiv:hep-ph/0310219`.

[73] M. Bordone, G. Isidori, and A. Pattori, *On the Standard Model predictions for R_K and R_{K^*}*, Eur. Phys. J. **C76** (2016) 440, `arXiv:1605.07633`.

[74] G. Ciezarek *et al.*, *A Challenge to Lepton Universality in B Meson Decays*, Nature **546** (2017) 227, `arXiv:1703.01766`.

[75] LHCb collaboration, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetries of $B \to K^*\mu^+\mu^-$ decays*, JHEP **06** (2014) 133, `arXiv:1403.8044`.

[76] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of the decay $B_s^0 \to \phi\mu^+\mu^-$*, JHEP **07** (2013) 084, `arXiv:1305.2168`.

[77] LHCb collaboration, R. Aaij *et al.*, *Measurement of the S-wave fraction in $B^0 \to K^+\pi^-\mu^+\mu^-$ decays and the $B^0 \to K^*(892)^0\mu^+\mu^-$ differential branching fraction*, JHEP **11** (2016) 047, Erratum ibid. **04** (2017) 142, `arXiv:1606.04731`.

[78] LHCb collaboration, R. Aaij *et al.*, *Measurement of form-factor-independent observables in the decay $B^0 \to K^{*0}\mu^+\mu^-$*, Phys. Rev. Lett. **111** (2013) 191801, `arXiv:1308.1707`.

[79] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using $3fb^{-1}$ of integrated luminosity*, JHEP **02** (2016) 104, `arXiv:1512.04442`.

[80] Belle, S. Wehle *et al.*, *Lepton-Flavor-Dependent Angular Analysis of $B \to K^*\ell^+\ell^-$*, Phys. Rev. Lett. **118** (2017) 111801, `arXiv:1612.05014`.

[81] ATLAS Collaboration, A. Collaboration, *Angular analysis of* $B_d^0 \to K^* \mu^+ \mu^-$ *decays in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2017-023, CERN, Geneva, Apr, 2017.

[82] CMS, C. Collaboration, *Measurement of the* P_1 *and* P_5' *angular parameters of the decay* $\mathrm{B}^0 \to \mathrm{K}^{*0} \mu^+ \mu^-$ *in proton-proton collisions at* $\sqrt{s} = 8$ TeV, .

[83] S. Jäger and J. Martin Camalich, *On* $B \to V \ell \ell$ *at small dilepton invariant mass, power corrections, and new physics*, JHEP **05** (2013) 043, `arXiv:1212.2263`.

[84] S. Jäger and J. Martin Camalich, *Reassessing the discovery potential of the* $B \to K^* \ell^+ \ell^-$ *decays in the large-recoil region: SM challenges and BSM opportunities*, Phys. Rev. **D93** (2016) 014028, `arXiv:1412.3183`.

[85] T. Blake *et al.*, *An empirical model of the long-distance contributions to* $\bar{B}^0 \to \bar{K}^{*0} \mu^+ \mu^-$ *transitions*, `arXiv:1709.03921`.

[86] T. Hurth, C. Langenbruch, and F. Mahmoudi, *Direct determination of Wilson coefficients using* $B^0 \to K^{*0} \mu^+ \mu^-$ *decays*, JHEP **11** (2017) 176, `arXiv:1708.04474`.

[87] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using* $B^+ \to K^+ \ell^+ \ell^-$ *decays*, Phys. Rev. Lett. **113** (2014) 151601, `arXiv:1406.6482`.

[88] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality with* $B^0 \to K^{*0} \ell^+ \ell^-$ *decays*, JHEP **08** (2017) 055, `arXiv:1705.05802`.

[89] CLEO, T. E. Coan *et al.*, *Study of exclusive radiative B meson decays*, Phys. Rev. Lett. **84** (2000) 5283, `arXiv:hep-ex/9912057`.

[90] Belle, M. Nakao *et al.*, *Measurement of the* $B \to K^* \gamma$ *branching fractions and asymmetries*, Phys. Rev. **D69** (2004) 112001, `arXiv:hep-ex/0402042`.

[91] BaBar, B. Aubert *et al.*, *Measurement of Branching Fractions and CP and Isospin Asymmetries in* $B \to K^*(892)\gamma$ *Decays*, Phys. Rev. Lett. **103** (2009) 211802, `arXiv:0906.2177`.

[92] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions* $\mathcal{B}(B^0 \to K^{*0}\gamma)/\mathcal{B}(B_s^0 \to \phi\gamma)$ *and the direct CP asymmetry in* $B^0 \to K^{*0}\gamma$, Nucl. Phys. **B867** (2013) 1, `arXiv:1209.0313`.

[93] W. Altmannshofer, P. Stangl, and D. M. Straub, *Interpreting Hints for Lepton Flavor Universality Violation*, Phys. Rev. **D96** (2017) 055008, `arXiv:1704.05435`.

[94] B. Capdevila *et al.*, *Patterns of New Physics in $b \to s\ell^+\ell^-$ transitions in the light of recent data*, JHEP **01** (2018) 093, `arXiv:1704.05340`.

[95] T. Hurth, F. Mahmoudi, D. Martinez Santos, and S. Neshatpour, *Lepton nonuniversality in exclusive $b{\to}s\ell\ell$ decays*, Phys. Rev. **D96** (2017) 095034, `arXiv:1705.06274`.

[96] F. Lionetto, N. Serra, O. Steinkamp, and K. Mueller, *Measurements of Angular Observables of $B^0 \to K^{*0}\mu^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ Decays and the Upgrade of LHCb*, `https://cds.cern.ch/record/2624938`, Feb, 2018. Presented 22 Mar 2018.

[97] LHCb Collaboration, R. Aaij *et al.*, *Expression of Interest for a Phase-II LHCb Upgrade: Opportunities in flavour physics, and beyond, in the HL-LHC era*, Tech. Rep. CERN-LHCC-2017-003, CERN, Geneva, Feb, 2017.

[98] T. M. Karbach and M. Schlupp, *Constraints on Yield Parameters in Extended Maximum Likelihood Fits*, `arXiv:1210.7141`.

[99] UTfit Collaboration, M. Bona *et al.*, *The Unitarity Triangle Fit in the Standard Model and Hadronic Parameters from Lattice QCD: A Reappraisal after the Measurements of Δm_s and BR($B \to \tau\nu_\tau$)*, JHEP **0610** (2006) 081, `arXiv:hep-ph/0606167`, We use the updated data from Winter 2013 (pre-Moriond 13).

[100] A. Bharucha, D. M. Straub, and R. Zwicky, *$B \to V\ell^+\ell^-$ in the Standard Model from light-cone sum rules*, JHEP **08** (2016) 098, `arXiv:1503.05534`.

[101] EOS, *A hep program for flavor observables*, `https://github.com/eos/eos`.

[102] R. J. Barlow, *Extended maximum likelihood*, Nucl. Instrum. Meth. **A297** (1990) 496.

[103] B. Capdevila, S. Descotes-Genon, L. Hofer, and J. Matias, *Hadronic uncertainties in $B \to K^*\mu^+\mu^-$: a state-of-the-art analysis*, JHEP **04** (2017) 016, `arXiv:1701.08672`.

[104] BaBar, B. Aubert *et al.*, *Measurement of branching fractions and charge asymmetries for exclusive B decays to charmonium*, Phys. Rev. Lett. **94** (2005) 141801, `arXiv:hep-ex/0412062`.

[105] BaBar, B. Aubert *et al.*, *Measurement of decay amplitudes of $B \to J/\psi K^*, \psi(2S)K^*$, and $\chi_{c1}K^*$ with an angular analysis*, Phys. Rev. **D76** (2007) 031102, `arXiv:0704.0522`.

[106] Belle, R. Itoh *et al.*, *Studies of CP violation in B —> J/psi K* decays*, Phys. Rev. Lett. **95** (2005) 091601, `arXiv:hep-ex/0504030`.

[107] Belle, K. Chilikin *et al.*, *Experimental constraints on the spin and parity of the Z(4430)$^+$*, Phys. Rev. **D88** (2013) 074026, `arXiv:1306.4894`.

[108] Belle, K. Chilikin *et al.*, *Observation of a new charged charmoniumlike state in $\bar{B}^0 \to J/\psi K^- \pi^+$ decays*, Phys. Rev. **D90** (2014) 112009, `arXiv:1408.6457`.

[109] LHCb collaboration, R. Aaij *et al.*, *Measurement of the polarization amplitudes in $B^0 \to J/\psi K^*(892)^0$ decays*, Phys. Rev. **D88** (2013) 052002, `arXiv:1307.2782`.

[110] C. Lefèvre, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, Dec, 2008.

[111] ATLAS, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[112] CMS, S. Chatrchyan *et al.*, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.

[113] ALICE, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[114] LHCb collaboration, R. Aaij *et al.*, *Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} = 7$ TeV in the forward region*, Phys. Lett. **B694** (2010) 209, `arXiv:1009.2731`.

[115] LHCb collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*, Phys. Rev. Lett. **118** (2017) 052002, Erratum ibid. **119** (2017) 169901, `arXiv:1612.05140`.

[116] LHCb collaboration, R. Aaij *et al.*, *Measurements of the B^+, B^0, B^0_s meson and Λ^0_b baryon lifetimes*, JHEP **04** (2014) 114, `arXiv:1402.2554`.

[117] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, `arXiv:1405.7808`.

[118] M. Tobin, *The lhcb silicon tracker*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **831** (2016) 174 , Proceedings of the 10th International "Hiroshima" Symposium on the Development and Application of Semiconductor Tracking Detectors.

[119] LHCb collaboration, P. R. e. a. Barbosa-Marinho, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029, LHCb-TDR-008.

[120] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024, LHCb-TDR-006.

[121] J. Van Tilburg and M. Merk, *Track simulation and reconstruction in LHCb*, 2005. Presented on 01 Sep 2005.

[122] M. De Cian, U. Straumann, O. Steinkamp, and N. Serra, *Track Reconstruction Efficiency and Analysis of* $B^0 \to K^{*0}\mu^+\mu^-$ *at the LHCb Experiment*, Sep, 2013. Presented 14 Mar 2013.

[123] J. Brehmer, J. Albrecht, and P. Seyfert, *Ghost probability: an efficient tool to remove background tracks*, Tech. Rep. CERN-LHCb-INT-2012-025, CERN, Geneva, Sep, 2012.

[124] M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, `arXiv:1211.6759`.

[125] S. Filippov *et al.*, *Experimental performance of SPD/PS detector prototypes*, .

[126] LHCb, I. Machikhiliyan, *The LHCb electromagnetic calorimeter*, J. Phys. Conf. Ser. **160** (2009) 012047.

[127] LHCb, R. I. Dzhelyadin, *The LHCb hadron calorimeter*, Nucl. Instrum. Meth. **A494** (2002) 332.

[128] I. Machikhiliyan and the LHCb calorimeter group, *Current status and performance of the lhcb electromagnetic and hadron calorimeters*, Journal of Physics: Conference Series **293** (2011) 012052.

[129] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, `arXiv:1211.1346`.

[130] G. Bencivenni *et al.*, *Advances in triple-GEM detector operation for high-rate particle triggering*, Nucl. Instrum. Meth. **A513** (2003) 264.

[131] F. Archilli *et al.*, *Performance of the Muon Identification at LHCb*, JINST **8** (2013) P10020, `arXiv:1306.0249`.

[132] R. Aaij *et al.*, *The LHCb Trigger and its Performance in 2011*, JINST **8** (2013) P04022, `arXiv:1211.3055`.

[133] V. V. Gligorov, *A single track HLT1 trigger*, LHCb-PUB-2011-003.

[134] R. Aaij and J. Albrecht, *Muon triggers in the High Level Trigger of LHCb*, Tech. Rep. LHCb-PUB-2011-017. CERN-LHCb-PUB-2011-017, CERN, Geneva, Sep, 2011.

[135] M. Williams *et al.*, *The HLT2 Topological Lines*, Tech. Rep. LHCb-PUB-2011-002. CERN-LHCb-PUB-2011-002, CERN, Geneva, Jan, 2011.

[136] V. V. Gligorov and M. Williams, *Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree*, JINST **8** (2013) P02013, `arXiv:1210.6861`.

[137] T. Likhomanenko *et al.*, *LHCb topological trigger reoptimization*, J. Phys. Conf. Ser. **664** (2015) 082025.

[138] A. Puig, *The LHCb trigger in 2011 and 2012*, LHCb-PUB-2014-046.

[139] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, `arXiv:0710.3820`.

[140] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.

[141] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

[142] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays*, Eur. Phys. J. **C45** (2006) 97, `arXiv:hep-ph/0506026`.

[143] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270; Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[144] M. Clemencic *et al.*, *The LHCb simulation application, Gauss: Design, evolution and experience*, J. Phys. Conf. Ser. **331** (2011) 032023.

[145] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, `arXiv:physics/0402083`.

[146] S. Coquereau *et al.*, *The $B^0 \to K^{*0}\mu^+\mu^-$ selection for 3fb^{-1} of LHCb data*, Tech. Rep. LHCb-INT-2013-058. CERN-LHCb-INT-2013-058, CERN, Geneva, Nov, 2013.

[147] A. Poluektov, *Kernel density estimation of a multidimensional efficiency profile*, JINST **10** (2015) P02011, `arXiv:1411.5528`.

[148] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*, Wadsworth international group, Belmont, California, USA, 1984.

[149] R. Schapire, *The boosting approach to machine learning: An overview*, Lecture Notes In Statistics-New York-Springer Verlag (2003) 1149.

[150] B. P. Roe *et al.*, *Boosted decision trees, an alternative to artificial neural networks*, Nucl. Instrum. Meth. **A543** (2005) 577, `arXiv:physics/0408124`.

[151] J. H. Friedman, *Greedy function approximation: A gradient boosting machine*, The Annals of Statistics **29** (2001) 1189.

[152] A. Blum, A. Kalai, and J. Langford, *Beating the hold-out: Bounds for k-fold and progressive cross-validation*, in *Proceedings of the Twelfth Annual Conference on Computational Learning Theory*, COLT '99, (New York, NY, USA), pp. 203–208, ACM, 1999. doi: 10.1145/307400.307439.

[153] G. Ciezarek, *Searches for Lepton Number Violation, and Flavour Violation beyond the Yukawa Couplings at LHCb*, PhD thesis, Imperial Coll., London, 2014-05-01.

[154] C. Adrover *et al.*, *Search for the* $B_S^0 \to \mu^+\mu^-$ *and* $B^0 \to \mu^+\mu^-$ *decays with 3* fb^{-1} *at LHCb*, LHCb-ANA-2013-032 (2013).

[155] A. L. Read, *Presentation of search results: the cls technique*, Journal of Physics G: Nuclear and Particle Physics **28** (2002) 2693.

[156] Particle Data Group, C. Patrignani *et al.*, *Review of particle physics*, Chin. Phys. **C40** (2016) 100001, and 2017 update.

[157] M. Perrin-Terrin and G. Mancinelli, *Optimisation of the binning of the discriminating variables used in the computation of* $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ *upper limits with the modified frequentist approach*, Tech. Rep. LHCb-INT-2012-003. CERN-LHCb-INT-2012-003, CERN, Geneva, Feb, 2012.

[158] M. Abadi *et al.*, *Tensorflow: Large-scale machine learning on heterogeneous distributed systems*, CoRR **abs/1603.04467** (2016) `arXiv:1603.04467`.

[159] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975) 343.

[160] LHCb collaboration, R. Aaij *et al.*, *Observation of* J/ψ*-pair production in* pp *collisions at* $\sqrt{s} = 7$ *TeV*, Phys. Lett. **B707** (2012) 52, `arXiv:1109.0963`.

[161] LHCb collaboration, R. Aaij *et al.*, *Measurement of the* $B_s^0 \to J/\psi \bar{K}^{*0}$ *branching fraction and angular amplitudes*, Phys. Rev. **D86** (2012) 071102(R), `arXiv:1208.0738`.

[162] J. Matias, *On the S-wave pollution of* $B \to K^*\ell^+\ell^-$ *observables*, Phys. Rev. **D86** (2012) 094024, `arXiv:1209.1525`.

[163] T. Blake, U. Egede, and A. Shires, *The effect of S-wave interference on the* $B^0 \to K^{*0}\ell^+\ell^-$ *angular observables*, JHEP **03** (2013) 027, `arXiv:1210.5279`.

[164] D. Aston *et al.*, *A Study of K- pi+ Scattering in the Reaction K- p —> K- pi+ n at 11-GeV/c*, Nucl. Phys. **B296** (1988) 493.

[165] J. R. Pelaez and A. Rodas, $\pi\pi \to K\bar{K}$ *scattering up to 1.47 GeV with hyperbolic dispersion relations*, `arXiv:1807.04543`.

[166] LHCb collaboration, R. Aaij *et al.*, *Observation of the resonant character of the* $Z(4430)^-$ *state*, Phys. Rev. Lett. **112** (2014) 222002, `arXiv:1404.1903`.

[167] D. Becirevic and A. Tayduganov, *Impact of* $B \to K_0^*\ell^+\ell^-$ *on the New Physics search in* $B \to K^*\ell^+\ell^-$ *decay*, Nucl. Phys. **B868** (2013) 368, `arXiv:1207.4004`.

[168] J. A. Bailey *et al.*, $B \to Kl^+l^-$ *decay form factors from three-flavor lattice QCD*, Phys. Rev. **D93** (2016) 025026, `arXiv:1509.06235`.

[169] G. J. Feldman and R. D. Cousins, *A Unified approach to the classical statistical analysis of small signals*, Phys. Rev. **D57** (1998) 3873, `arXiv:physics/9711021`.

[170] F. Halzen and P. Minkowski, *On resonance interpretation in argand diagrams*, Nuclear Physics B **14** (1969) 522 .

[171] J. M. Blatt and V. F. Weisskopf, *Theoretical nuclear physics*, John Wiley & Sons, New York (1952).

[172] LHCb collaboration, R. Aaij *et al.*, *Measurement of the phase difference between the short- and long-distance amplitudes in the* $B^+ \to K^+\mu^+\mu^-$ *decay*, Eur. Phys. J. **C77** (2017) 161, `arXiv:1612.06764`.