Study of the B0–> K*e+e- decay with the LHCb detector and development of a novel concept of PID detector: the Focusing DIRC

Martino Borsato

LAL 15-254

Université Sorbonne Paris Cité

Thèse préparée à l'Université Paris Diderot

École doctorale STEP'UP - ED N°560

Laboratoire de l' Accélérateur Linéaire - UMR 8607

Study of the $B^0 \to K^{*0} e^+ e^-$ decay with the LHCb detector and development of a novel concept of PID detector: the Focusing DIRC

Étude de la désintégration $B^0 \to K^{*0} e^+ e^-$ avec le détecteur LHCb et développement d'un nouveau concept de détecteur PID: le FDIRC

par

Martino Borsato

Soutenue le 8 septembre 2015
devant un jury composé de:

Achille Stocchi Professeur (LAL)	Président du jury
Ulrik Egede Professor (Imperial College London)	Rappporteur
Marcella Bona Lecturer (Queen Mary University of London)	Rapporteur
Neville Harnew Professor (University of Oxford)	Membre
José Ocariz Professeur (Université Paris Diderot)	Membre
Damir Bečirević Directeur de recherche (Univ. Paris-Sud 11)	Membre
Marie-Hélène Schune Directeur de recherche (LAL)	Directrice de thèse
Nicolas Arnaud Chargé de recherche (LAL)	Co-encadrant de thèse

Resumé

Les courants neutres changeant la saveur de type $b \to s\gamma$ sont interdits au niveau le plus bas dans le Modele Standard (MS) et ils se produisent au premier ordre à travers des diagrammes radiatif à boucle. Par conséquent, ils sont sensibles à la Nouvelle Physique (NP), qui peut contribuer de façon significative. En outre, la chiralité de l'interaction faible dans le MS implique que le photon émis a une polarisation gauche. Cependant, toute une classe de théories de NP ne partagent pas cette caractéristique du MS et pourraient se manifester sans ambiguité comme une contribution droite à la polarization.
Cette thèse présente la première étude de la polarisation du photon provenant du processus $b \to s\gamma$ à travers une analyse angulaire du canal $B^0 \to K^{*0}e^+e^-$. Même si la desintégration $B^0 \to K^{*0}e^+e^-$ n'est pas une transition $b \to s$ de type radiatif, la contribution avec un photon virtuel qui se couple à la paire de leptons domine dans la région de bas masse invariante dileptonique (q). De plus, le canal avec des électrons plutôt que des muons permet de mieux isoler cette contribution dans la region de basses valeurs de q^2. La luminosité intégrée de $3\,\text{fb}^{-1}$ recueillie par LHCb au cours du Run 1 du LHC a permis pour la première fois de sélectionner un échantillon d'évenements $B^0 \to K^{*0}e^+e^-$ assez grand pour mesurer les asymétries transverses $\text{A}_\text{T}^{(2)} = -0.23 \pm 0.23 \pm 0.05$ et $\text{A}_\text{T}^{\text{Im}} = 0.14 \pm 0.22 \pm 0.05$ dans la région de bas q^2 entre $0.002\,\text{GeV}^2/c^4$ et $1\,\text{GeV}^2/c^4$. Ces mesures sont en accord avec les prédictions du MS et ils fournissent de nouvelles contraintes sur les contributions droities à la polarisation avec le même niveau de précision que celui fourni par la moyenne des mesures faites auprès des usines à B.

Les expériences de physique des saveurs nécessitent des détecteurs complexes dédiés à l'identification des différents types de hadrons chargés. Ces détecteurs sont pour la plupart basés sur la détection de la lumière Cherenkov émise par ces particules lorsqu'elles traversent un milieu diélectrique. Cette thèse presente également une activité de R&D sur un nouveau concept de détecteur basé sur le detecteur DIRC de l'exérience *BABAR* : le FDIRC. Son design est destiné à l'utilisation dans un environnement avec un bruit de fond deux ordres de grandeur plus grands que dans *BABAR* grâce à une camera d'imagerie plus rapide, de dimension réduite et composée de silice fondue résistant à l'irradiation. Le premier prototype à grande échelle a été testé au téléscope à rayons cosmiques de SLAC et a démontré la faisabilité du concept de détecteur.

Abstract

Flavour-changing neutral current processes of the type $b \to s\gamma$ are forbidden at the tree level in the Standard Model (SM) and occur at leading order through radiative loop diagrams. Therefore, they are sensitive to new physics (NP), which may contribute with competing diagrams. Furthermore, the chirality of the weak interaction in the SM implies that the photon emitted has left-handed polarisation. However, a whole class of NP theories do not share this SM feature and may manifest unambiguously as a right-handed contribution to the polarisation.

This thesis presents the first study of the $b \to s\gamma$ photon polarisation through an angular analysis of the $B^0 \to K^{*0}e^+e^-$ channel. Even though $B^0 \to K^{*0}e^+e^-$ is not a radiative $b \to s$ transition, the contribution from a virtual photon coupling to the lepton pair dominates in the low-q^2 region. Furthermore, the channel with electrons rather than muons allows to better isolate the virtual photon contribution at the low end of the q^2 spectrum.
The integrated luminosity of $3\,\text{fb}^{-1}$ collected by LHCb during LHC Run 1 allowed for the first time to select a sample of $B^0 \to K^{*0}e^+e^-$ events large enough to measure the transverse asymmetries $\text{A}_\text{T}^{(2)} = -0.23 \pm 0.23 \pm 0.05$ and $\text{A}_\text{T}^\text{Im} = 0.14 \pm 0.22 \pm 0.05$ in the low q^2 region between $0.002\,\text{GeV}^2/c^4$ and $1\,\text{GeV}^2/c^4$. These measurements are found to be consistent with SM predictions and provide new constraints on right handed contributions to the photon polarisation at the same level of precision as the one from the average of measurements involving radiative decays made at B-factories.

Flavour physics experiments require advanced detectors dedicated to the identification of the different flavours of charged hadrons. Most of them are based on the detection of the Cherenkov light emitted by these particles as they traverse a dielectric medium. This thesis presents also a R&D activity on a new detector concept based on the *BABAR* DIRC, the FDIRC. Its design is intended to operate the detector in an environment with a background two orders of magnitude larger than *BABAR* thanks to a faster and smaller imaging camera made of radiation-hard fused silica. The first full-scale prototype of the FDIRC was tested at the SLAC Cosmic Ray Telescope and demonstrated the feasibility of the detector concept.

Contents

Chapter 1

Theoretical overview

This chapter introduces the motivation for the study of the $B^0 \to K^{*0} e^+ e^-$ decay.

The Standard Model of particle physics and its shortcomings are introduced in Section 1.1. The study of $b \to s$ transitions as a powerful way to test this model is presented in Section 1.2 followed by an overview of the state of the art. In Section 1.3 more details are given about radiative $b \to s\gamma$ decays and their close relation to the physics that can be investigated with the study of $B^0 \to K^{*0} e^+ e^-$ at very low q^2.

1.1 The Standard Model of particle physics

The Standard Model (SM) [1–3] of particle physics is the theory best describing experimental data. The theoretical pillar of this model is local gauge invariance with respect to the group:

$$G_{\rm SM} \equiv SU(3)_C \times SU(2)_L \times U(1)_Y \ . \tag{1.1}$$

Strong interaction is described by the group $SU(3)_C$, while electroweak interaction is described by $SU(2)_L \times U(1)_Y$.

Gauge invariance determines the spin-1 particle content of the SM, the so-called gauge bosons. They consist in eight gluons associated with strong interaction, with coupling constant g_S, and four bosons associated with electroweak interaction W^i $(i = 1, 2, 3)$ and B, with coupling constants g and g', respectively.

Spin-1/2 particles (fermions) defined in the SM consist in three generations of quarks and three generations of leptons, whose transformation properties under $G_{\rm SM}$ are summarised by the quantum numbers shown in Figure 1.1 (top centre). Note that only quarks participate to strong interaction, only left-handed fermions participate to weak interaction and there is no right-handed neutrino.

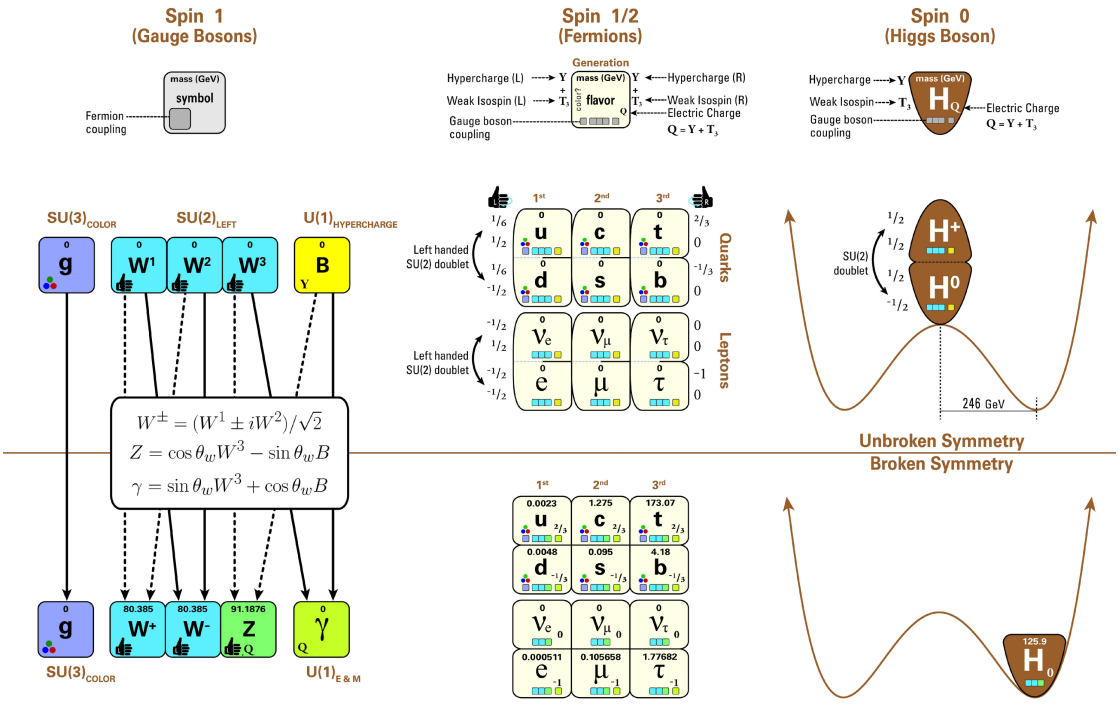


Fig. 1.1: The diagram[2] shows the elementary particles of the Standard Model: the gauge bosons, the three generations of quarks and leptons, and the Higgs bosons. Their names, masses, spins, charges, chiralities, and interactions with the strong, weak and electromagnetic forces are given. It also depicts the crucial role of the Higgs boson in electroweak symmetry breaking by showing how the properties of the various particles differ in the symmetric phase (top) and the broken-symmetry phase (bottom).

The last important ingredient of the SM is the spin-0 Higgs field, ϕ, the only known scalar field. It is defined as a complex $SU(2)$ doublet (see Figure 1.1) whose Lagrangian not only contains the gauge interaction term (coupling it to the electroweak gauge bosons) and the self interaction term (the Higgs potential V) but also the Yukawa interaction with fermions:

$$\mathcal{L}_{\mathcal{H}} = |D_\mu \phi|^2 - V(\phi) + \mathcal{L}_{\mathrm{Y}} \tag{1.2}$$

where the potential is

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2 \tag{1.3}$$

The potential is chosen such that $\mu^2 < 0$, which induces a degenerate family of

[2]The picture is based on *Standard Model Of Particle Physics, Most Complete Diagram.jpg* by Latham Boyle and is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

minima at a non-zero value of the field

$$\phi^\dagger\phi = -\frac{\mu^2}{\lambda} \equiv \frac{v^2}{2} \tag{1.4}$$

i.e. a non-vanishing vacuum expectation value v. This induces the spontaneous symmetry breaking (SSB) and choose a particular direction of the minima using the gauge freedom (the unitary gauge):

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix} . \tag{1.5}$$

The SSB of $SU(2)_L \times U(1)_Y$ is a fundamental feature of the SM. Through SSB, the Higgs gauge interaction gives a mass to the combinations of electroweak gauge bosons which are defined in Figure 1.1. These combinations are identified with the "physical" $W^\pm$ and Z bosons. The other orthogonal combination, γ, remains massless and is identified with the photon. Furthermore, SSB gives birth to the Higgs boson, H, a real field arising from fluctuations about the non-zero minimum of the Higgs potential.

SSB allows also to give mass to the fermions via the Yukawa couplings of the Higgs field ϕ (except for neutrinos, as they do not have a right-handed partner).

$$\mathcal{L}_\mathrm{Y} = -Y^d_{ij}\bar{Q}_{\mathrm{L}i}\phi d_{\mathrm{R}j} - Y^u_{ij}\bar{Q}_{\mathrm{L}i}\phi^* u_{\mathrm{R}j} - Y^\ell_{ij}\bar{L}_{\mathrm{L}i}\phi e_{\mathrm{R}j} + \mathrm{h.c.} \tag{1.6}$$

where $\bar{Q}_{\mathrm{L}i}$ and $\bar{L}_{\mathrm{L}i}$ are the left-handed quark and lepton doublets while $d_{\mathrm{R}j}$, $u_{\mathrm{R}j}$ and $e_{\mathrm{R}j}$ are the right handed singlets for up- and down-type quarks and leptons. Y^q_{ij} are the Yukawa couplings, where i and j run over the three generations.

Under SSB, the Higgs field acquires its vacuum expectation value, giving mass terms to quarks and charged leptons. Considering only quarks for simplicity:

$$\mathcal{L}_\mathrm{Y}^\mathrm{quarks} = -\bar{d}_{\mathrm{L}i}m^d_{ij}d_{\mathrm{R}j} - \bar{u}_{\mathrm{L}i}m^u_{ij}u_{\mathrm{R}j} + \mathrm{h.c.} \tag{1.7}$$

where the mass terms are related to Yukawa couplings $m^q_{ij} = \frac{v}{\sqrt{2}}Y^q_{ij}$ and are free to be non-diagonal among the three generations. These mass matrices are diagonalised by unitary rotations of the quark fields $\mathcal{U}^u_\mathrm{L}$, $\mathcal{U}^u_\mathrm{R}$, $\mathcal{U}^d_\mathrm{L}$ and $\mathcal{U}^d_\mathrm{R}$:

$$m^u_\alpha = (\mathcal{U}^{u\dagger}_\mathrm{L})_{i\alpha}m^u_{ij}(\mathcal{U}^u_\mathrm{R})_{\alpha j} \qquad\qquad m^d_\alpha = (\mathcal{U}^{d\dagger}_\mathrm{L})_{i\alpha}m^d_{ij}(\mathcal{U}^d_\mathrm{R})_{\alpha j} \tag{1.8}$$

where α indicates the quark generation. The values of quark masses m^u_α and m^d_α are arbitrary in the SM. The same is true for charged lepton masses m^ℓ_α. However, a hierarchy of the mass values among the three generations is found experimentally: $m_1 \ll m_2 \ll m_3$ for the three generations of quarks and charged leptons. Mass values in GeV/c^2 are reported in the centre bottom scheme of Figure 1.1.

The diagonalisation requires different transformations for the u_{L} and d_{L} quarks, which are part of the same $SU(2)$ doublet. As a consequence, charged weak currents (mediated by the $W^{\pm}$ boson) are not invariant under these rotations and give rise to quark flavour mixing. Namely, the charged current in the flavour basis reads:

$$\mathcal{L}_{\mathrm{CC}} = \frac{ig'}{\sqrt{2}} \left(W^{+}_{\mu} \bar{u}_{\mathrm{L}j} \gamma^{\mu} d_{\mathrm{L}j} + W^{-}_{\mu} \bar{d}_{\mathrm{L}j} \gamma^{\mu} u_{\mathrm{L}j} \right) \tag{1.9}$$

while, rotated in the mass basis gives:

$$\mathcal{L}_{\mathrm{CC}} = \frac{ig'}{\sqrt{2}} \left(W^{+}_{\mu} \bar{u}_{\mathrm{L}\alpha} (\mathcal{U}^{u}_{\mathrm{L}})_{\alpha j} (\mathcal{U}^{d\dagger}_{\mathrm{L}})_{j\beta} \gamma^{\mu} d_{\mathrm{L}\beta} + W^{-}_{\mu} \bar{d}_{\mathrm{L}\alpha} (\mathcal{U}^{d}_{\mathrm{L}})_{\alpha j} (\mathcal{U}^{u\dagger}_{\mathrm{L}})_{j\beta} \gamma^{\mu} u_{\mathrm{L}\beta} \right) \tag{1.10}$$

which includes the flavour mixing matrix $(\mathcal{U}^{u}_{\mathrm{L}})_{\alpha j} (\mathcal{U}^{d\dagger}_{\mathrm{L}})_{j\beta} \equiv V^{\mathrm{CKM}}_{\alpha\beta}$, known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix [4, 5]. This unitary complex 3×3 matrix can be parametrised by three rotation angles and a complex phase which is the only known source of CP violation. These four parameters are also arbitrary in the SM. Experimentally, it was found that the magnitude of the V^{CKM} elements have a hierarchical structure, which can be expressed [6] in terms of the expansion parameter $\lambda \simeq 0.24$:

$$V^{\mathrm{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1 & \lambda & 0 \\ -\lambda & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \mathcal{O}(\lambda^2) \,. \tag{1.11}$$

The weak neutral-current mediated by the Z boson on the other hand, is independent of the mixing due to the unitarity of the rotation matrices, *e.g.* for down-type quarks:

$$\begin{aligned} \mathcal{L}^{d}_{\mathrm{NC}} &= ig' \left(\bar{d}_{\mathrm{L}j} Z_{\mu} \gamma^{\mu} d_{\mathrm{L}j} \right) \\ &= ig' \left(\bar{d}_{\mathrm{L}\alpha} (\mathcal{U}^{d}_{\mathrm{L}})_{\alpha j} (\mathcal{U}^{d\dagger}_{\mathrm{L}})_{j\beta} Z_{\mu} \gamma^{\mu} d_{\mathrm{L}\beta} \right) \\ &= ig' \left(\bar{d}_{\mathrm{L}\alpha} \delta_{\alpha\beta} Z_{\mu} \gamma^{\mu} d_{\mathrm{L}\beta} \right) \,. \end{aligned} \tag{1.12}$$

Therefore, there are no flavour-changing neutral currents (FCNC) at the tree level in the SM. Hence, they can occur only through second order processes.

Shortcomings of the Standard Model

The SM is a very elegant theory which is having an unprecedented success in describing a wide variety of phenomena (see for example the extensive review by the Particle Data Group [7]). However, it presents problems, both in terms of unexplained observations and inconsistencies which ask for a theory beyond the standard model (BSM).

The SM is built with a large number of input parameters, namely 18 of which 13 from the Yukawa sector. Also, it does not explain the number of fermion generations nor their highly hierarchical structure in terms of mass.

Gravity cannot be included in the conceptual scheme of the SM since general relativity is incompatible with QFT. This makes it an effective theory which cannot be valid at the Planck energy scale.

The SM does not provide a candidate for cold dark matter (whose contribution to the mass content of the Universe is found to be about 5 times larger than ordinary matter). Furthermore, the amount of *CP* violation from the SM Yukawa sector is found to be 10 orders of magnitude smaller than that required to generate the matter-antimatter asymmetry observed in the Universe [8, 9]. A *CP* violating phase is allowed also in the strong sector, but so far it was found to be suppressed at the level of 10^{-9} with no SM explanation [10].

The simple formulation of the SM summarised in Section 1.1 does not give masses to neutrinos although the experimentally-proven oscillation between neutrino flavours [7] implies that they have non-zero masses.

Large quantum contributions to the square of the Higgs boson mass would inevitably make the mass huge, unless there is an incredible fine-tuning cancellation between the quadratic radiative corrections and the bare mass. This problem may be solved by the presence of physics beyond the SM at low mass scale (1 TeV), which would provide a more natural cancellation.

1.2 Rare $b \to s$ decays as probes of New Physics

1.2.1 Motivation and theoretical approach

Indirect searches for new physics (NP) through the analysis of SM processes have proven to be very powerful and complementary to direct searches. A famous example is the search for the suppressed decay $K_{\text{L}}^0 \to \mu^+\mu^-$ which led to the hypothesis of the GIM mechanism [11] and to the prediction of the charm quark. However, this approach requires not only precise experimental measurements but also a choice of NP-sensitive observables being cleanly predicted in the SM. One very promising way of pursuing such indirect searches is the study of B decays involving the $b \to s$ quarks transition. Different complementary reasons gave this process a place of honour in indirect searches for NP. Firstly, the b quark mass is much larger than the typical scale of the strong interaction ($m_b \gg \Lambda_{\text{QCD}}$), which means the otherwise troublesome long-distance contributions are generally less important than in processes involving lighter meson systems.
Secondly, given that the $b \to s$ transition involves a flavour-changing neutral-

current (FCNC), it can occur only at loop level in the SM, which makes it more sensitive to NP, which may contribute with new tree level diagram or by new particles entering loop contributions. Furthermore, the fact that the CKM matrix is approximately diagonal makes FCNC rare processes in the SM. Hence, the SM contribution is flavour suppressed and NP may show up very clearly if it does not share this very peculiar SM feature.

Another advantage of the $b \to s$ transition is that the b quark mass is much smaller than the masses of the electroweak bosons and of the top quark, which allows the construction of an effective low-energy theory describing the weak interactions of quarks, by separating the two mass scales. This approach is similar to Fermi's effective theory of weak decays. The theoretical framework of this effective approach is the Operator Product Expansion (OPE) [12] which allows to write:

$$\langle f|\mathcal{H}_{\text{eff}}|i\rangle = \sum_k \mathcal{C}_k(\mu)\langle f|\mathcal{O}_k(\mu)|i\rangle \tag{1.13}$$

and thus separate the decay amplitude in a low-energy part being encoded in local operators $\mathcal{O}_i$ with different Lorentz structures and in a high-energy part encoded in the so-called Wilson coefficients $\mathcal{C}_i$, which contain the short-distance physics. The effective Hamiltonian for $b\to s$ transitions reads:

$$\mathcal{H}_{\text{eff}} = -\frac{4\,G_F}{\sqrt{2}} V_{tb}V_{ts}^* \frac{\alpha_e}{4\pi} \sum_i (\mathcal{C}_i\mathcal{O}_i + \mathcal{C}_i'\mathcal{O}_i') + \text{h.c.} \tag{1.14}$$

where G_F is the Fermi constant , V_{ij} are CKM matrix elements and α_e is the fine structure constant. Operators $\mathcal{O}_i^{(\prime)}(\mu_s)$ and Wilson coefficients $\mathcal{C}_i^{(\prime)}(\mu_s)$ are evaluated at the renormalization scale μ_s. The most relevant operators to $b \to s$ studies are the electro-magnetic penguin, $\mathcal{O}_7^{(\prime)}$, and the semileptonic operators $\mathcal{O}_9^{(\prime)}$ and $\mathcal{O}_{10}^{(\prime)}$. Their expressions are:

$$\mathcal{O}_7 = \frac{m_b}{e}(\bar{s}\sigma_{\mu\nu}P_{\text{R}}b)F^{\mu\nu}, \qquad \mathcal{O}_7' = \frac{m_b}{e}(\bar{s}\sigma_{\mu\nu}P_{\text{L}}b)F^{\mu\nu}, \tag{1.15}$$

$$\mathcal{O}_9 = (\bar{s}\gamma_\mu P_{\text{L}}b)(\bar{\ell}\gamma^\mu\ell)\,, \qquad \mathcal{O}_9' = (\bar{s}\gamma_\mu P_{\text{R}}b)(\bar{\ell}\gamma^\mu\ell)\,, \tag{1.16}$$

$$\mathcal{O}_{10} = (\bar{s}\gamma_\mu P_{\text{L}}b)(\bar{\ell}\gamma^\mu\gamma_5\ell)\,, \qquad \mathcal{O}_{10}' = (\bar{s}\gamma_\mu P_{\text{R}}b)(\bar{\ell}\gamma^\mu\gamma_5\ell)\,, \tag{1.17}$$

where $P_{\text{R,L}} = \frac{1\pm\gamma_5}{2}$ denotes the right- and left-handed chiral projectors and $F^{\mu\nu}$ is the electro-magnetic tensor. In this scheme, Wilson coefficients are first calculated at the weak mass scale $\mu_s = m_W$ by matching the effective theory to the full SM theory. At this scale, QCD corrections are small and can be calculated precisely. Then, $\mathcal{C}^{(\prime)}(\mu_s)$ are evoluted down to $\mu_s = m_b$ using renormalization group equations, which allow to link the low and high energy scales.
BSM physics can modify the values of these coefficients as well as contribute to

others which are highly suppressed in the SM such as the ones corresponding to scalar and pseudoscalar operators $\mathcal{O}_{\rm S}$ and $\mathcal{O}_{\rm P}$ or to tensor operators $\mathcal{O}_T$ and $\mathcal{O}_{T5}$. Furthermore, the left-handedness of the weak interaction makes the primed Wilson coefficients $\mathcal{C}'_{7,9,10}$ to be suppressed in the SM, while they may get large values in NP models involving a different helicity structure.

1.2.2 Overview of experimental activity on $b \to s$

A wealth of different $b \to s$ processes has been studied and have contributed to put new very important constraints on extensions of the SM. Hereafter is a brief overview of the latest results in this field putting the accent on the ones related to the $B^0 \to K^{*0} e^+ e^-$ angular analysis, being the subject of this work. Rare $b \to s\ell\ell$ decays, very rare purely leptonic decays and also tests of lepton universality are presented hereafter while radiative $b \to s\gamma$ decays are discussed later in Section 1.3.

Rare $b \to s\ell\ell$ decays

B decays involving a $b \to s$ transition with the emission of a dilepton pair $\ell^+\ell^-$ are sensitive to a large number of Wilson coefficients, whose contributions vary as a function of q^2 as shown in Figure 1.2. Many channels have been investigated as well as many different observables ranging from branching ratios to asymmetries and angular coefficients. While inclusive branching ratios were studied only at B-factories [14, 15], many exclusive channels got more important contributions from hadron-collider experiments, which were able to collect much larger yields. Indeed, the high efficiency with muonic channels and the large production fraction of heavier b states at the LHC allowed the LHCb experiment to attain unprecedented precision in the branching ratios of $B \to K^{(*)}\mu^+\mu^-$ and $B_s^0 \to \phi\mu^+\mu^-$ (see Figure 1.3) as well as $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays [16–19] and also the ATLAS and CMS collaborations reported results on some of these branching ratios [20, 21]. These experimental results are now much more precise than the corresponding theoretical predictions since calculations are sensitive to hadronic uncertainties in the form factors. Hence, their sensitivity to NP is limited by the large theory errors rather than by the experimental reach. All these measurements are found to be compatible with the SM, although a trend towards smaller rates is found in several decays [22].

In order to increase sensitivity to BSM physics, it is useful to study observables in which the effects of form factor uncertainties are reduced. For example, CP- and isospin- asymmetries were studied in $B \to K^{(*)}\mu^+\mu^-$ decays by LHCb [16, 24]. They were both found to be compatible with SM across the whole q^2 spectrum.

Furthermore, angular analyses of $b \to s\ell\ell$ decays can be exploited to extract observables suffering less from hadronic uncertainties. Also, angular analyses are

more effective than branching ratio measurements in disentangling the different Wilson coefficients involved.
An angular analysis of charged and neutral $B \to K\mu^+\mu^-$ decays was performed at LHCb. The angular distribution of the angle defined by the dilepton decay with respect to the recoiling K in the B rest frame, θ_ℓ, was used to measure the forward-backward asymmetry $A_{\rm FB}$ as well as $F_{\rm H}$, a measure of the contribution from (pseudo)scalar and tensor amplitudes. Both these parameters are very suppressed in the SM across the whole q^2 range [25]. Measurements are found to be consistent with this prediction.

The angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay, with $K^{*0} \to K^+\pi^-$, is more complicated, but also richer in physics. As explained in Section 1.3.3, the angular decay rate (see Equation 1.28) is described by three angles, θ_ℓ, θ_K and ϕ, and can be used to measure the angular observables S_j and A_j (defined in Equation 1.29). The very large yield collected by LHCb in the full Run 1 dataset allows to fit all these parameters at the same time in fine q^2 bins and extract the correlation between them as well as the contamination from $K^+\pi^-$ in an S-wave configuration. The first measurement of the complete set of CP-averaged observables, S_j, was recently presented by the LHCb collaboration [26] using the whole Run 1 dataset (see Figure 1.4). The set of corresponding CP-asymmetries, A_j, is expected to be published soon. Good agreement with the SM predictions was found for all measured observables apart from S_5, which presents some tension

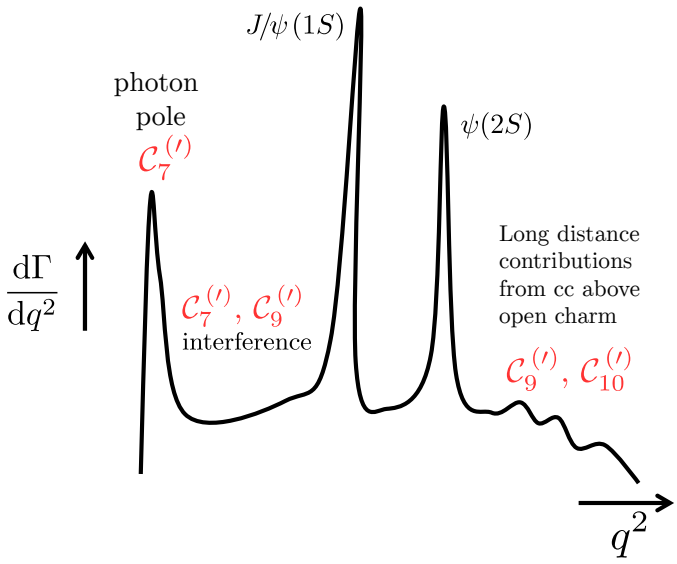


Fig. 1.2: Artistic sketch of the profile of the differential decay rate of $B^0 \to K^{*0}\ell^+\ell^-$ as a function of q^2 [13]. The main Wilson coefficients contributing to different q^2 regions are represented on top of the curve.

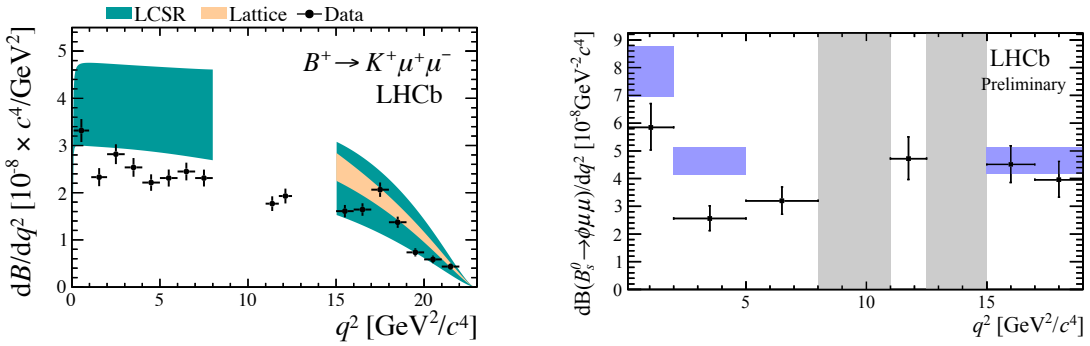


Fig. 1.3: Selected results on $b \to s\ell\ell$ differential branching ratios as a function of q^2. (left) $B^+ \to K^+\mu^+\mu^-$ as measured by LHCb [16] and compared to theory predictions. (right) $B_s^0 \to \phi\mu^+\mu^-$ from LHCb preliminary results [18]. The theory predictions are from [22, 23].

in the q^2 region below the J/ψ resonance (see Figure 1.5). This discrepancy was previously seen in the LHCb 1 fb^{-1} dataset on the related observable P_5' [27] and thus confirms an existing puzzle. In order to assess the significance of this deviation a fit taking into account all correlations both in data and in theory predictions is necessary. One such fit, including all the main $b \to s$ results, was performed and found that New Physics modifying the Wilson coefficient $\mathcal{C}_9$ is preferred over the Standard Model by 3.7 σ [22]. This very intriguing result raised large discussions in the theory community both on possible NP models that could accommodate this deviation [28–31] as well as on possible problems in the SM calculations [32–35].

Different subsets of these angular observables have also been measured by *BABAR*, Belle, CDF, ATLAS and CMS [20, 21, 36–38] and updates are expected from ATLAS and CMS with the full Run 1 dataset (25 fb^{-1}). Also, LHCb is going to re-analyse Run 1 data with a finer binning by taking advantage of the so-called "method of moments" [39] in place of a likelihood fit which suffers from convergence problems when dealing with a small number of events. Furthermore, another method is being set up to determine all observables as continuous distributions in q^2 rather than in bins [40], allowing to maximize the sensitivity (up to a gain equivalent to having a ∼70% larger dataset). This is the so-called "amplitude method" and consists in fitting directly the q^2-dependent $\overline{K}^{*0}$ spin amplitudes taking advantage of the symmetries of the angular distribution and exploiting a three-parameter ansatz for the q^2 dependence of the amplitudes (in the 1-6 GeV$^2/c^4$ q^2-region).

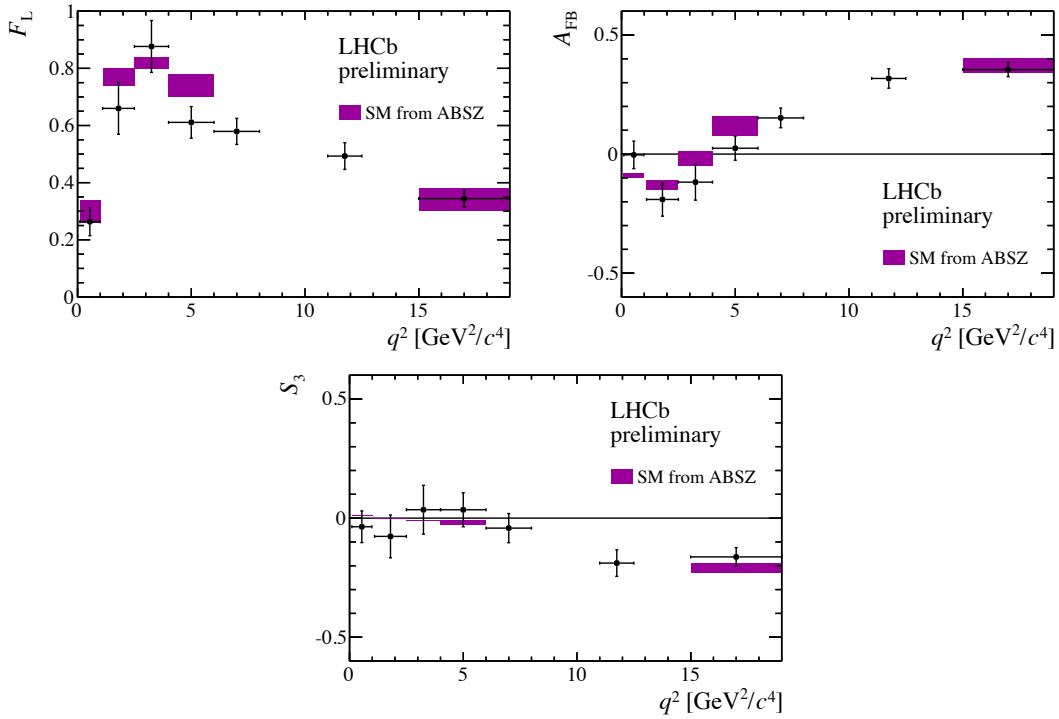


Fig. 1.4: Preliminary results of the angular analysis of $B^0 \to K^{*0}\mu^+\mu^-$ by LHCb using the whole Run 1 dataset ($3\,\text{fb}^{-1}$) [26] are compared to SM theory predictions from [23]. A selection of the set of CP-averaged angular observables measured is shown here: the longitudinal fraction F_L, the forward-backward asymmetry A_FB and the S_3 observable. As will be explained in Section 1.3.3, A_FB and S_3 are related to $\text{A}_\text{T}^\text{Re}$ and $\text{A}_\text{T}^{(2)}$, which are measured in the very low-q^2 region using the electronic channel $B^0 \to K^{*0}e^+e^-$.

Very rare decays

The $B_s^0 \to \mu^+\mu^-$ decay proceeds via loop diagrams similarly to $b \to s\ell\ell$ processes, but in addition it is helicity suppressed in the SM. Its time-integrated branching fraction is predicted in the SM to be as low as $\mathcal{B}(B_s^0 \to \mu^+\mu^-)_\text{SM} = (3.65 \pm 0.23) \times 10^{-9}$ [41]. A simultaneous analysis of this channel was performed by both the LHCb and CMS experiments with the full Run 1 dataset [42]. The combination of dataset allowed to observe this very rare process with a significance of 6 σ. The measured branching ratio was found to be in good agreement with the SM prediction:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-)_\text{exp} = (2.8\,^{+0.7}_{-0.6}) \times 10^{-9}\,. \tag{1.18}$$

This measurement is very sensitive to many NP models which may reduce the large helicity suppression of this decay. It provides also stringent constraints on the Wilson coefficient $\mathcal{C}_{10}^{(\prime)}$ as well as on possible scalar or pseudoscalar contributions

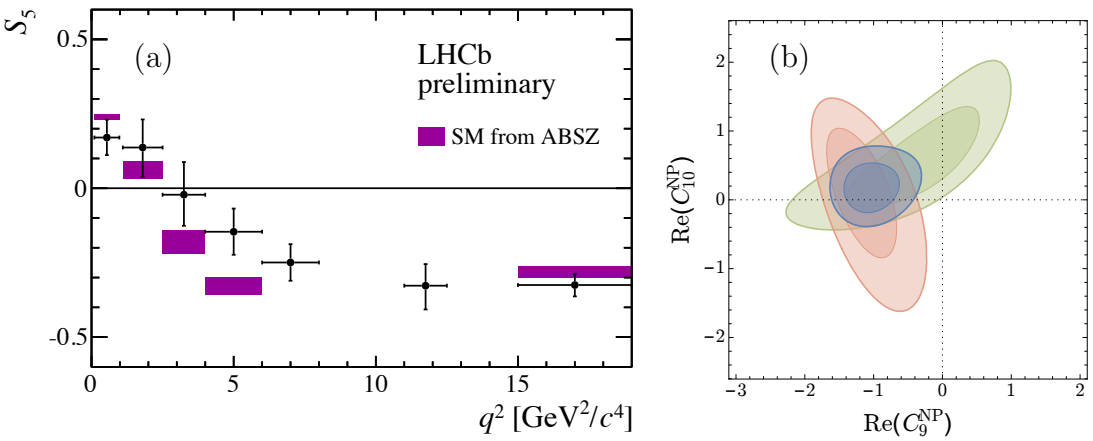


Fig. 1.5: (a) The angular observable S_5 as measured by LHCb [26] exhibits some tension in the q^2 region below the J/ψ resonance. (b) Allowed regions in the $\mathcal{R}e(\mathcal{C}_9^{(\mathrm{NP})}) - \mathcal{R}e(\mathcal{C}_{10}^{(\mathrm{NP})})$ plane obtained with a global fit of $b \to s$ data [22] including the latest result from LHCb [26]. The blue contours correspond to the 1 and 2 σ best fit regions from the global fit. The green and red contours correspond to the 1 and 2 σ regions if only branching ratio data or only data on $B^0 \to K^{*0}\mu^+\mu^-$ angular observables is taken into account

$\mathcal{C}_{\mathrm{S,P}}$.

Non-universal lepton couplings

Another interesting test of the SM using $b \to s$ decays is to look for non-universal couplings, *i.e.* to compare decays to different lepton flavours. Given that all SM particles couple equally to different lepton flavours (apart from the Higgs boson), these tests are sensitive to any BSM competing diagram not being lepton-flavour universal. LHCb performed one such test comparing $B^+ \to K^+\ell^+\ell^-$ decays involving electrons and muons in the 1-6 GeV$^2/c^4$ q^2 region [43] and found a deviation from the SM expectation [25] (very close to unity):

$$\mathcal{R}_K = 0.745^{+0.090}_{-0.074} \pm 0.036\,, \tag{1.19}$$

where the first uncertainty is statistical and the second is systematic. Despite this deviation being large, its significance is limited to 2.6 σ. Nonetheless, it attracted a lot of interest in the theory community as it can be related to the deviation in the $B^0 \to K^{*0}\mu^+\mu^-$ angular analysis [26] (as proposed by [44]) as well as to other deviations found in lepton universality tests with $B \to D^{(*)}\ell\nu$ decays [45–47] (as proposed by [48]). Recently, the possibility that some QED effects might have been underestimated in doing the SM calculation of $\mathcal{R}_K$ is

being investigated [49]. The $\mathcal{R}_K$ measurement is going to be followed by other lepton universality tests on $b \to s\ell\ell$ processes such as measurements of $\mathcal{R}_{K^{*0}}$ and $\mathcal{R}_\phi$ which may help to distinguish possible NP scenario thanks to their different spin-parity structure [50,51]. Also, an angular analysis of the $B^0 \to K^{*0}e^+e^-$ decay in the whole q^2 range and a direct comparison of the results to $B^0 \to K^{*0}\mu^+\mu^-$ can shed more light on this puzzle [22]. Both this analysis and a measurement of $\mathcal{R}_K$ at LHCb can take advantage of the experimental methods developed for the analysis of $B^0 \to K^{*0}e^+e^-$ described in this work.

1.3 Photon polarisation in $b \to s\gamma$ processes

The Wilson coefficients $\mathcal{C}_7^{(\prime)}$ corresponding to the electro-magnetic penguin operator $\mathcal{O}_7^{(\prime)}$ have been extensively studied in several $b \to s\gamma$ processes. Firstly, the focus has been on the $B \to X_s\gamma$ branching fraction. The world average value comes from the measurements made at the CLEO, *BABAR* and Belle experiments (see [52] and references therein). It is found to be in very good agreement with the latest SM calculation [53]:

$$\mathcal{B}(B \to X_s\gamma)_{\text{exp}} = (3.43 \pm 0.21 \pm 0.07) \times 10^{-4} \tag{1.20}$$

$$\mathcal{B}(B \to X_s\gamma)_{\text{theo}} = (3.36 \pm 0.23) \times 10^{-4} \tag{1.21}$$

Note that $\mathcal{B}(B \to X_s\gamma)$ is proportional at leading order to the sum of left- and right-handed electromagnetic Wilson coefficients squared, $|\mathcal{C}_7|^2 + |\mathcal{C}_7'|^2$. Hence, it is not directly sensitive to the handedness of the emitted photon.

In the SM the photon emitted in $b \to s\gamma$ transitions is predominantly left-handed in b decays (and right-handed in $\bar{b}$ decays). This is due to the fact that only left-handed quarks participate to the weak interaction. Therefore, an helicity flip is needed in one of the external quark-legs to create a right-handed contribution. This can be induced by the mass term of the quark fields (the Yukawa couplings), which results in a factor m_b for $b_R \to s_L\gamma_L$, and a factor m_s for $b_L \to s_R\gamma_R$. Therefore, the emission of right-handed photons in the SM is suppressed by a factor m_s/m_b which, in terms of Wilson coefficients, means that roughly $\mathcal{C}_7^{'}/\mathcal{C}_7 \simeq m_s/m_b \simeq 0.02$ (for real $\mathcal{C}_7^{(\prime)}$), although the exact level of suppression depends on the particular exclusive mode considered, due to different QCD effects [54, 55].
In contrast, in many extensions of the SM, NP may manifest as a large right handed current, hence a large $\mathcal{C}_7^{'}$ Wilson coefficient. These include for example Minimal Supersymmetric Standard Models (MSSM) [56–58], Left-Right Symmetric Models (LRSM) [59] as well as other models [60]. Various experimental methods have been proposed to measure the photon polarisation in $b \to s\gamma$. In the following, the most sensitive ones are briefly described.

1.3.1 Time-dependent rate of radiative decays

An indirect determination of the photon polarisation consists in studying the time-dependent rate of radiative decays of a B_q^0 meson to a *CP* eigenstate, f^{CP}, such as $B^0 \to K^{*0}(\to K_S^0\pi^0)\gamma$ or $B_s^0 \to \phi(\to K^+K^-)\gamma$. In general, the time dependent decay rate is given by:

$$\overset{(-)}{\mathcal{B}}(t) = \mathcal{B}_0 e^{-\Gamma t}\left(\cosh\frac{\Delta\Gamma\, t}{2} - \mathcal{A}^{\Delta}\sinh\frac{\Delta\Gamma\, t}{2} \pm \mathcal{C}\cos(\Delta m\, t) \mp \mathcal{S}\sin(\Delta m\, t)\right), \tag{1.22}$$

where $\mathcal{B}_0$ is the total decay rate, $\Delta\Gamma$ and Δm are the width and mass differences between the heavy and light mass eigenstates, $\mathcal{A}^\Delta$ is the CP-asymmetry induced by the width difference, $\mathcal{C}$ is the direct CP asymmetry and $\mathcal{S}$ is the CP asymmetry induced by the interference between mixing and decay. Both $\mathcal{A}^\Delta$ and $\mathcal{S}$ are sensitive to $\mathcal{C}_7^{(\prime)}$ as explained in the following.

Since a $\overline{B}{}_q^0$ (B_q^0) meson decays predominantly into a left- (right-) handed photon, the dominant amplitudes of $\overline{B}{}_q^0 \to f^{CP}\gamma_L$ and $\overline{B}{}_q^0 \to B_q^0 \to f^{CP}\gamma_R$ cannot interfere quantum-mechanically. Therefore, the CP-asymmetry generated by the $B - \overline{B}$ mixing, $\mathcal{A}_{CP}$, is expected to be zero up to $O(m_s/m_b)$. However, any NP contribution to the "wrong"-helicity amplitude could cause a deviation from zero. Indeed, $S_{f^{CP}\gamma}$ can be written in terms of $\mathcal{C}_7^{(\prime)}$ Wilson coefficients as [61] :

$$\mathcal{S} \approx \xi \frac{2\mathcal{I}m[e^{-i\phi_q}\,\mathcal{C}_7\mathcal{C}_7'|}{|\mathcal{C}_7|^2 + |\mathcal{C}_7'|^2}\,, \tag{1.23}$$

where $\xi(=\pm 1)$ is the CP-eigenvalue of f^{CP} and ϕ_q is the phase in the $B_q^0 - \overline{B}{}_q^0$ mixing, which in the SM is $\phi_d = 2\beta \simeq 43^\circ$ for the B^0 mixing ($\phi_s \simeq 0$ for the B_s^0). The coefficient $S_{K^*\gamma}$ has been measured by the *BABAR* and Belle experiments using $B^0 \to K^{*0}\gamma$ decays where the K^{*0} goes to the CP eigenstate $K_S^0\pi^0$, giving [52,62,63]

$$S_{K^*\gamma} = -0.16 \pm 0.22 \tag{1.24}$$

which is consistent with the SM prediction of -0.023 ± 0.016 [54].
Other measurements of $S_{f^{CP}\gamma}$ in different final states have not yet reached the same level of accuracy. Only results from $B^0 \to K_S^0\rho^0\gamma$ are coming close in sensitivity [64, 65]. LHCb is not competitive in this measurement as decays with K_S^0 and π^0 in the final state are difficult to reconstruct in the hadronic environment.

A measurement of $\mathcal{A}^\Delta$ using the B^0 meson is very difficult since the width difference is very small. However, for the B_s^0 meson, the large width difference $\Delta\Gamma_s$ allows to measure $\mathcal{A}^\Delta$ in B_s^0 decays by a measurement of the effective lifetime and without the need to tag the initial state. Indeed the CP-averaged decay rate reads:

$$\mathcal{B}(t) = \mathcal{B}_0 e^{-\Gamma_s t}\left(\cosh\frac{\Delta\Gamma_s\, t}{2} - \mathcal{A}^\Delta \sinh\frac{\Delta\Gamma_s\, t}{2}\right). \tag{1.25}$$

In the SM, the $\mathcal{A}^\Delta$ parameter is close to zero $\mathcal{A}^\Delta = 0.047^{+0.040}_{-0.025}$ [66] and any significant deviation from this value would be a sign of a right-handed contribution from NP. Indeed, in terms of $\mathcal{C}_7^{(\prime)}$ Wilson coefficients $\mathcal{A}^\Delta$ reads:

$$\mathcal{A}^\Delta \approx \xi \frac{2\mathcal{R}e[e^{-i\phi_s}\,\mathcal{C}_7\mathcal{C}_7'|}{|\mathcal{C}_7|^2 + |\mathcal{C}_7'|^2}\,. \tag{1.26}$$

where ϕ_s is close to 0 in the SM.

At B-factories only $B_{d,u}$ mesons are produced by the $\Upsilon(4S)$ decay, while at LHCb a large sample of B_s^0 mesons is produced by pp collisions, and the $B_s^0 \to \phi\gamma$ with $\phi \to K^+K^-$ can be reconstructed with a large yield [67]. Furthermore, an advantage of measuring $\mathcal{A}^\Delta$ rather than $\mathcal{S}$ is that the measurement does not require flavour tagging, which is challenging at LHCb. The main experimental difficulty is the modelling of the decay-time acceptance, which is required to be known to a very high level of accuracy [68]. Nonetheless, the $B^0 \to K^{*0}\gamma$ and $B_s^0 \to \phi J/\psi$ channels can be used as proxies to extract the acceptance.

1.3.2 Up-down asymmetry in $B^+ \to K^+\pi^-\pi^+\gamma$

Another way to access information on the photon polarisation is through the up-down asymmetry in $B^+ \to K^+\pi^-\pi^+\gamma$ decays [69,70]. Indeed, the photon direction with respect to the plane defined by the K^+ π^- π^+ system is related to its polarisation λ_γ, namely $\mathcal{A}_{\rm ud} = C\lambda_\gamma$. The constant of proportionality C, though, suffers from large hadronic uncertainties and depends on the intermediate states of the $K^+\pi^-\pi^+$ system as well as on their interference. This analysis has been performed by LHCb with the full Run I dataset [71]. $\mathcal{A}_{\rm ud}$ was measured in four intervals of $K^+\pi^-\pi^+$ invariant mass defined by the known resonances. By combining the significance of the four mass intervals, a non-zero value of $\mathcal{A}_{\rm ud}$ is observed at 5.2 σ, which means photon polarisation in $b \to s\gamma$ is directly observed for the first time. A deeper understanding of the structure of the $K^+\pi^-\pi^+$ system is needed to be able to extract information on $\mathcal{C}_7^{(\prime)}$ with this approach. Furthermore, the polarisation λ_γ depends on the difference of $\mathcal{C}_7$ and $\mathcal{C}_7'$ squared [69]

$$\lambda_g \simeq \frac{|\mathcal{C}_7'|^2 - |\mathcal{C}_7|^2}{|\mathcal{C}_7'|^2 + |\mathcal{C}_7|^2} \tag{1.27}$$

and therefore it is not sensitive to small NP contributions. Nonetheless, it contains complementary information with respect to the other measurements and could be useful to solve ambiguities in the determination of $\mathcal{C}_7^{(\prime)}$ [61].

1.3.3 Angular analysis of $B^0 \to K^{*0}e^+e^-$ as a probe to measure photon polarisation

The electromagnetic penguin operator corresponding to the Wilson coefficient $\mathcal{C}_7^{(\prime)}$ not only contributes to radiative B decays, but also to $B \to V\ell^+\ell^-$ (where V is a vector state) decays where it enters along with $\mathcal{C}_9^{(\prime)}$ and $\mathcal{C}_{10}^{(\prime)}$. As explained in Section 1.2.2 and sketched in Figure 1.2, these contributions vary as a function of the q^2 of the process. In particular, the electromagnetic penguin process is responsible for the pole at the low-q^2 end, where the branching ratio ramps up

as $1/q^2$. The $\mathcal{C}_7^{(\prime)}$ contribution to $b \to s\ell\ell$ can therefore be disentangled from $\mathcal{C}_9^{(\prime)}$ and $\mathcal{C}_{10}^{(\prime)}$ by an analysis restricted to the very-low q^2 region. There, the decay is almost equivalent to a radiative $b \to s\gamma$ decay with the real photon being replaced by a virtual one [73]. The main SM diagrams involved are shown in Figure 1.6, but in the very-low q^2 just the two penguin diagrams involving a photon rather than a Z^0 are dominating. Furthermore, an angular analysis of $B^0 \to K^{*0}\ell^+\ell^-$ with $K^{*0} \to K^+\pi^-$ is sensitive to the photon polarisation [73] which can influence the distribution of the angle ϕ between the planes defined by the K^{*0} decay and the dilepton in the B center-of-mass frame. In addition, in the region of low q^2, some theoretical uncertainties from long distance contributions are greatly reduced, thereby allowing more control over the SM prediction and increasing sensitivity to any NP effect [35, 74, 75].

The q^2 region above $0.1\,\mathrm{GeV}^2/c^4$ has been studied by LHCb through the analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay as outlined in Section 1.2.2. Experimentally, an analysis with muons rather than electrons in the final state produces a much higher yield at LHCb. This is primarily due to the distinctive signature that muons provide, which is efficiently exploited in the online selection, together with the better mass and energy resolutions and higher reconstruction efficiency of dimuon decays. However, as outlined in [76], the electronic channel $B^0 \to K^{*0}e^+e^-$

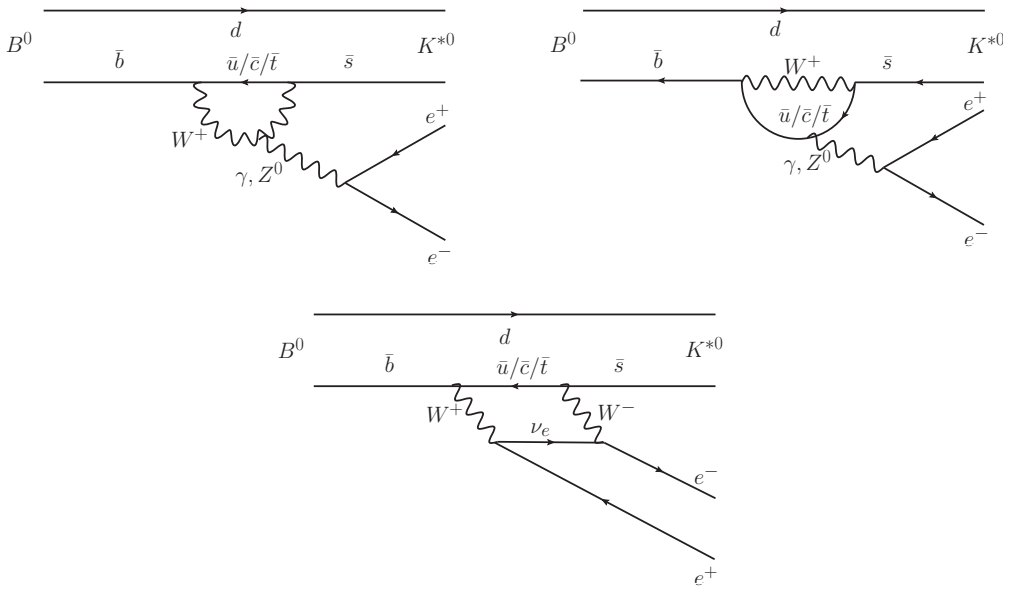


Fig. 1.6: Dominant Standard Model Feynman graphs involved in the $B^0 \to e^+e^-K^{*0}$ decay [72]. The two diagrams on the top are called "penguin diagrams" and may involve a photon or a Z^0. The one on the bottom is a "box diagram".

provides greater sensitivity to the photon polarisation by allowing to go much lower in q^2 thanks to the smaller threshold at $4m_e^2$. This is shown in Figure 1.7 with a comparison of the expected true q^2 spectra selected by LHCb for $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}\mu^+\mu^-$ channels. The larger photon pole contribution in electrons allows also to get a larger yield than the one would get by stopping at $0.1\,\text{GeV}^2/c^4$. Additionally, the formalism for the $B^0 \to K^{*0}e^+e^-$ decay is greatly simplified as the electron mass can be neglected even in the very low q^2 region. However, above a q^2 of $1\,\text{GeV}^2/c^4$, the muon mass terms become negligible and the electron and muon modes have essentially the same functional dependence on the Wilson coefficients (within the lepton flavour universality assumption).

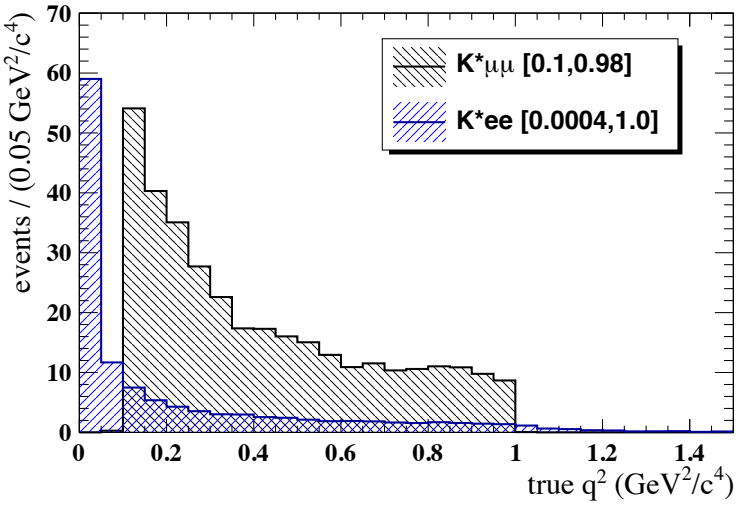


Fig. 1.7: Comparison of the expected *true*-q^2 spectra for selected $B^0 \to K^{*0}e^+e^-$ (blue) and $B^0 \to K^{*0}\mu^+\mu^-$ events from the LHCb MC. The requirements on the *reconstructed*-q^2 range are $[0.1, 0.98]\,\text{GeV}^2/c^4$ for muons, which is the first bin used for the angular analysis [26], and $[0.0004, 1.0]\,\text{GeV}^2/c^4$ for electrons. For illustration purposes, spectra are normalised to the signal yields obtained in data by the two LHCb analyses on the $3\,\text{fb}^{-1}$ [26, 72].

This work presents the first angular analysis of the $B^0 \to K^{*0}e^+e^-$ channel performed by LHCb in the very-low (reconstructed) q^2 region between $0.0004\,\text{GeV}^2/c^4$ and $1\,\text{GeV}^2/c^4$. The analysis is published in [72] and is detailed in this document in Chapters 3 and 4.

The partial decay width of the $B^0 \to K^{*0}e^+e^-$ decay can be described in terms of q^2 and three angles, θ_ℓ, θ_K and ϕ. The angle θ_ℓ is defined as the angle between the direction of the e^+ (e^-) and the direction opposite to that of the B^0 ($\overline{B}^0$) meson in the dielectron rest frame. The angle θ_K is defined as the angle between the direction of the kaon and the direction opposite to that of the B^0 ($\overline{B}^0$) meson

in the K^{*0} ($\overline{K}^{*0}$) rest frame. The angle ϕ is the angle between the plane containing the e^+ and e^- and the plane containing the kaon and pion from the K^{*0} ($\overline{K}^{*0}$) in the B^0 ($\overline{B}^0$) rest frame. The basis is designed such that the angular definition for the $\overline{B}^0$ decay is a CP transformation of that for the B^0 decay. A sketch of the three angles is given in Figure 1.8. These definitions are identical to those used for the $B^0 \to K^{*0}\mu^+\mu^-$ analysis [17] and are detailed in Appendix A.1.

Using the notation of Ref. [77], the decay distribution of the B^0 corresponds to

$$\begin{aligned}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\,\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\phi} = \frac{9}{32\pi}\Big[& I_1^s \sin^2\theta_K + I_1^c \cos^2\theta_K + \\ & I_2^s \sin^2\theta_K \cos 2\theta_\ell + I_2^c \cos^2\theta_K \cos 2\theta_\ell + \\ & I_3 \sin^2\theta_K \sin^2\theta_\ell \cos 2\phi + I_4 \sin 2\theta_K \sin 2\theta_\ell \cos\phi + \\ & I_5 \sin 2\theta_K \sin\theta_\ell \cos\phi + I_6 \sin^2\theta_K \cos\theta_\ell + \\ & I_7 \sin 2\theta_K \sin\theta_\ell \sin\phi + I_8 \sin 2\theta_K \sin 2\theta_\ell \sin\phi + \\ & I_9 \sin^2\theta_K \sin^2\theta_\ell \sin 2\phi \Big] ,\end{aligned} \tag{1.28}$$

where the angular coefficients I_j are only functions of q^2. The same equation holds for $\bar{\Gamma}$ with $\bar{I}_j$, no sign change is involved with the current definition of angles. Angular coefficients I_j can be expressed as bilinear combination of six K^{*0} transversity-amplitudes: four transverse, $A_\perp^{\mathrm{L,R}}$ and $A_{||}^{\mathrm{L,R}}$, and two longitudinal, $A_0^{\mathrm{L,R}}$ (the labels L and R refer to the left and right chirality of the dielectron system). All their expressions are reported in Appendix A.2. This is valid in the limit of massless leptons, which is a very good approximation for electrons, otherwise one would need one more amplitude of timelike type. Amplitudes encode the de-

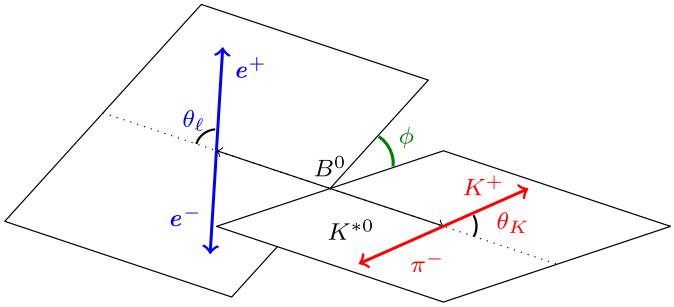


Fig. 1.8: A sketch of the definition of the three angles θ_ℓ, θ_K and ϕ for the $B^0 \to K^{*0}e^+e^-$ decay (adapted from [17]). Details are in the text below and in Appendix A.1

pendence to Wilson coefficients $\mathcal{C}^{(\prime)}_{7,9,10}$ and form factors. Their full expressions are reported in Appendix A.3.

Combining B^0 and $\overline{B}^0$ decays, it is possible to build angular observables as *CP*-averages or *CP*-asymmetries:

$$S_j = \left(I_j + \bar{I}_j\right) \Big/ \left(\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} + \frac{\mathrm{d}\bar{\Gamma}}{\mathrm{d}q^2}\right) \quad \text{or} \quad A_j = \left(I_j - \bar{I}_j\right) \Big/ \left(\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} + \frac{\mathrm{d}\bar{\Gamma}}{\mathrm{d}q^2}\right) \ . \tag{1.29}$$

The *CP*-averaged observables S_7, S_8 and S_9 are not interesting as they are suppressed by the small size of the strong phase difference between the decay amplitudes. They are consequently expected to be very close to zero not only in the SM but also in most extensions of it. However, the corresponding *CP*-asymmetries, A_7, A_8 and A_9, are not suppressed by the strong phases involved [78] and are sensitive to the effects of NP particles. Therefore, the sign of the ϕ angle is flipped for $\overline{B}^0$ decays so that in the angular distribution combining B^0 and $\overline{B}^0$ the *CP*-averages $S_{7,8,9}$ are replaced by *CP*-asymmetries $A_{7,8,9}$ (while S_{1-6} are kept as *CP*-averages). One more simplification allowed by the negligible electron mass is that the *CP* average of I_1^c, I_1^s, I_2^c and I_2^s (S_1^c, S_1^s, S_2^c and S_2^s) are related to F_L, the fraction of longitudinal polarisation of the K^{*0} meson,

$$S_1^c = -S_2^c = F_\mathrm{L} \qquad \text{and} \qquad \frac{4}{3} S_1^s = 4 S_2^s = 1 - F_\mathrm{L} \, . \tag{1.30}$$

In order to further simplify the expression of the angular distribution, the angle ϕ is transformed such that $\tilde{\phi} = \phi + \pi$ if $\phi < 0$, to cancel out the terms that have a $\sin\phi$ or $\cos\phi$ dependence, namely S_4, S_5, A_7 and A_8. Furthermore, the non-resonant contribution from the $B^0 \to K^+\pi^- e^+ e^-$ with the system $K^+\pi^-$ in an S-wave configuration is expected to be negligible in the low-q^2 region with the current sample size [79]. Indeed, using references [74, 79] it can be shown that the S-wave fraction, F_S, is proportional to the fraction of longitudinal polarisation of the K^{*0} in the $0 - 6\,\mathrm{GeV}^2/c^4$ q^2-range and since the decay is dominated by the transverse polarisation in the photon-pole region, F_S, is expected to be negligible. The remaining angular observables besides F_L are S_3, S_6 and A_9. The three of them can be reformulated as "transverse observables" [74] A_T^i by the ratio with the transverse fraction $(1 - F_\mathrm{L})$:

$$\begin{aligned} S_3 &= \frac{1}{2}\left(1 - F_\mathrm{L}\right) \mathrm{A}_\mathrm{T}^{(2)} \, , \\ S_6 &= \left(1 - F_\mathrm{L}\right) \mathrm{A}_\mathrm{T}^{\mathrm{Re}} \, , \\ A_9 &= \frac{1}{2}\left(1 - F_\mathrm{L}\right) \mathrm{A}_\mathrm{T}^{\mathrm{Im}} \, . \end{aligned} \tag{1.31}$$

The $\mathrm{A}_\mathrm{T}^{(2)}$, $\mathrm{A}_\mathrm{T}^{\mathrm{Re}}$ and $\mathrm{A}_\mathrm{T}^{\mathrm{Im}}$ are "optimised observables" for which the leading form-factor uncertainties cancel [74, 75], thus allowing more precise predictions. The

choice of this basis simplifies the comparison with theory, but it is worth saying that in principle the comparison is independent of the basis once the experimental correlations between observables as well as the theoretical ones are provided [22]. The same kind of reformulation is done to define P_i observables [80], which indeed are just a redefinition of transverse asymmetries: $\mathrm{A}_{\mathrm{T}}^{(2)} = P_1$, $\mathrm{A}_{\mathrm{T}}^{\mathrm{Re}} = 2P_2$ and $\mathrm{A}_{\mathrm{T}}^{\mathrm{Im}} = -2P_3^{CP}$. Finally, the $B^0 \to K^{*0} e^+ e^-$ angular distribution reads

$$\begin{aligned}\frac{1}{\mathrm{d}(\Gamma+\bar{\Gamma})/\mathrm{d}q^2}\,\frac{\mathrm{d}^4(\Gamma+\bar{\Gamma})}{\mathrm{d}q^2\,\mathrm{d}\cos\theta_\ell\,\mathrm{d}\cos\theta_K\,\mathrm{d}\tilde{\phi}} = \frac{9}{16\pi}\Bigg[\,&\frac{3}{4}(1-F_{\mathrm{L}})\sin^2\theta_K + F_{\mathrm{L}}\cos^2\theta_K + \\ &\left(\frac{1}{4}(1-F_{\mathrm{L}})\sin^2\theta_K - F_{\mathrm{L}}\cos^2\theta_K\right)\cos 2\theta_\ell \;+ \\ &\frac{1}{2}(1-F_{\mathrm{L}})\mathrm{A}_{\mathrm{T}}^{(2)}\sin^2\theta_K\sin^2\theta_\ell\cos 2\tilde{\phi} \;+ \\ &(1-F_{\mathrm{L}})\mathrm{A}_{\mathrm{T}}^{\mathrm{Re}}\sin^2\theta_K\cos\theta_\ell \;+ \\ &\frac{1}{2}(1-F_{\mathrm{L}})\mathrm{A}_{\mathrm{T}}^{\mathrm{Im}}\sin^2\theta_K\sin^2\theta_\ell\sin 2\tilde{\phi}\,\Bigg].\end{aligned} \tag{1.32}$$

The four angular observables F_{L}, $\mathrm{A}_{\mathrm{T}}^{(2)}$, $\mathrm{A}_{\mathrm{T}}^{\mathrm{Re}}$ and $\mathrm{A}_{\mathrm{T}}^{\mathrm{Im}}$ are related to the transversity amplitudes through [75]

$$\begin{aligned}F_{\mathrm{L}} &= \frac{|A_0|^2}{|A_0|^2 + |A_{||}|^2 + |A_\perp|^2}\,,\\ \mathrm{A}_{\mathrm{T}}^{(2)} &= \frac{|A_\perp|^2 - |A_{||}|^2}{|A_\perp|^2 + |A_{||}|^2}\,,\\ \mathrm{A}_{\mathrm{T}}^{\mathrm{Re}} &= \frac{2\mathcal{R}e(A_{||}^{\mathrm{L}}A_\perp^{*\mathrm{L}} + A_{||}^{\mathrm{R}}A_\perp^{*\mathrm{R}})}{|A_{||}|^2 + |A_\perp|^2}\,,\\ \mathrm{A}_{\mathrm{T}}^{\mathrm{Im}} &= \frac{2\mathcal{I}m(A_{||}^{\mathrm{L}}A_\perp^{*\mathrm{L}} + A_{||}^{\mathrm{R}}A_\perp^{*\mathrm{R}})}{|A_{||}|^2 + |A_\perp|^2}\,,\end{aligned} \tag{1.33}$$

where $|A_{\perp,||,0}|^2 = |A_{\perp,||,0}^{\mathrm{L}}|^2 + |A_{\perp,||,0}^{\mathrm{R}}|^2$. The three "transverse observables" $\mathrm{A}_{\mathrm{T}}^{(2)}$, $\mathrm{A}_{\mathrm{T}}^{\mathrm{Re}}$ and $\mathrm{A}_{\mathrm{T}}^{\mathrm{Im}}$ depend just on transverse amplitudes $A_\perp^{\mathrm{L,R}}$ and $A_{||}^{\mathrm{L,R}}$. The observable F_{L}, on the other hand, encodes the contribution from the longitudinal amplitude $A_0^{\mathrm{L,R}}$. Since it is related to the longitudinal polarisation of the K^{*0}, it is expected to be small at low q^2, as the virtual photon is then "quasi-real", hence transversely polarised. The observable $\mathrm{A}_{\mathrm{T}}^{\mathrm{Re}}$ is related to the forward-backward asymmetry A_{FB} by $\mathrm{A}_{\mathrm{T}}^{\mathrm{Re}} = \frac{4}{3}A_{\mathrm{FB}}/(1-F_{\mathrm{L}})$ [75]. Given the definition of $\tilde{\phi}$, the observable $\mathrm{A}_{\mathrm{T}}^{(2)}$ is averaged between B^0 and $\overline{B}^0$ decays, while $\mathrm{A}_{\mathrm{T}}^{\mathrm{Im}}$ corresponds to a CP asymmetry. These two observables bear the information on photon polarisation and indeed, in the limit $q^2 \to 0$, they can be expressed as the following simple functions of

$\mathcal{C}_7^{(\prime)}$ [75]

$$\mathrm{A}_{\mathrm{T}}^{(2)}(q^2 \to 0) = \frac{2\mathcal{R}e(\mathcal{C}_7\mathcal{C}_7'^*)}{|\mathcal{C}_7|^2 + |\mathcal{C}_7'|^2} \quad \text{and} \quad \mathrm{A}_{\mathrm{T}}^{\mathrm{Im}}(q^2 \to 0) = \frac{2\mathcal{I}m(\mathcal{C}_7\mathcal{C}_7'^*)}{|\mathcal{C}_7|^2 + |\mathcal{C}_7'|^2}. \tag{1.34}$$

The angular analysis of $B^0 \to K^{*0}e^+e^-$ therefore provides information on photon polarisation amplitudes through the terms involving the angle ϕ. The $B^0 \to K^{*0}e^+e^-$ sensitivity to $\mathcal{C}_7^{(\prime)}$ with the Run I data as well as a comparison to the other methods are provided in Section 4.6.

1.3.4 Photon polarisation in $B^0 \to K^{*0}\gamma$ from the γ to e^+e^- conversion?

The idea of observing a high energy photon transverse polarisation by looking at the plane of the e^+e^- pair creation is not new and had been studied and published in the years around 1950 [81]. In LHCb Run 1 data a large yield of $B^0 \to K^{*0}\gamma$ events with the γ conversion to an e^+e^- pair was collected. In principle, one could use these events to extract information on the $b \to s\gamma$ polarisation using the plane identified by the e^+e^- conversion pair. This possibility is analysed in the following.

Pair creation by a polarised photon

A high energetic photon can create an e^+e^- pair by interacting with a nucleus via Coulomb forces (by the so-called Bethe-Heitler process [82]). The cross-section for pair production by a polarised photon (at lowest order) is given in [73].

$$\begin{aligned}
\frac{\mathrm{d}\sigma}{\mathrm{d}E_1\,\mathrm{d}\Omega_1\,\mathrm{d}\Omega_2} &= \frac{Z^2e^6}{16\pi^3}\frac{|\vec{p}_1||\vec{p}_2|}{E_\gamma^3 k^4}\left\{\frac{(k^2-4E_2^2)(\vec{e}\cdot\vec{p}_1)(\vec{e}^{\,*}\cdot\vec{p}_1)}{(E_1-|\vec{p}_1|\cos\theta_1)^2} + \frac{(k^2-4E_1^2)(\vec{e}\cdot\vec{p}_2)(\vec{e}^{\,*}\cdot\vec{p}_2)}{(E_2-|\vec{p}_2|\cos\theta_2)^2}\right. \\
&- \frac{k^2+4E_1E_2}{(E_1-|\vec{p}_1|\cos\theta_1)(E_2-|\vec{p}_2|\cos\theta_2)}\left[(\vec{e}\cdot\vec{p}_1)(\vec{e}^{\,*}\cdot\vec{p}_2)+(\vec{e}^{\,*}\cdot\vec{p}_1)(\vec{e}\cdot\vec{p}_2)\right] \\
&+ \left. E_\gamma^2\frac{\vec{p}_1{}^2\sin^2\theta_1 + \vec{p}_2{}^2\sin^2\theta_2 + 2|\vec{p}_1||\vec{p}_2|\sin\theta_1\sin\theta_2\cos(\phi_1-\phi_2)}{(E_1-|\vec{p}_1|\cos\theta_1)(E_2-|\vec{p}_2|\cos\theta_2)}\right\},
\end{aligned} \tag{1.35}$$

where Z and e are the charge of the nucleus and the elementary charge respectively while $\vec{e}$ is the photon polarisation vector and $p_{1,2} = (E_{1,2}, \vec{p}_{1,2})$ is the positron (electron) momentum in the laboratory reference frame. If $q = (E_\gamma, \vec{p}_\gamma)$ is the photon momentum, $k = p_1 + p_2 - q$ is the momentum transferred to the nucleus, which will be taken to be infinitely heavy. In this limit the nucleus does not carry away any energy and the photon energy is transferred entirely to the electron pair $E_\gamma = E_1 + E_2$. The polar and asimuthal angle of the positron (electron) with respect to the photon momentum direction are denoted with $\Omega_{1,2} = (\theta_{1,2}, \phi_{1,2})$.

Equation 1.35 contains many variables and is not easy to interpret. Variables more directly linked to the experimental observables (and similar to the ones used in the $B^0 \to K^{*0} e^+ e^-$ angular analysis) are the mass of the e^+e^- system $m(e^+e^-)$, its total energy $E(e^+e^-)$ which is equal to the incident gamma energy E_γ in the hypothesis of an infinitely heavy nucleus, θ_ℓ, the e^+e^- polar decay angle in the e^+e^- center of mass and ϕ, the angle between the plane of the dilepton pair and the plane containing the photon momentum $\vec{q}$ and the photon electric vector $\vec{e}$. There are then three other variables peaked at very small values corresponding to the momentum exchanged with the nucleus, $\vec{k}$. The component k_z is not independent since $k_z = (m(e^+e^-)^2)/(2E_\gamma)$ but k_x and k_y are independent and have to be integrated out as they are typically between a fraction of a keV/c to tens of keV/c which is unobservable compared to the tens of GeV/c of the $e^\pm$. It is the unbalance between these two "hidden" variables and the others which makes formula 1.35 difficult to use. In an article of 1951 [81] G.C.Wick, using the Weizsacker-Williams approximation [83], wrote Equation 1.35 in terms of variables closer to physical observables, obtaining:

$$\frac{\mathrm{d}\sigma}{\sin\theta_\ell \,\mathrm{d}\theta_\ell \,\mathrm{d}\phi} = \frac{\beta\, r_0^2\, m_e^2\, c^4}{2E_\gamma^2} \left(\frac{1}{1-\beta^2\cos^2\theta_\ell} - \frac{1}{2} + 2\beta^2 \frac{1-\beta^2}{(1-\beta^2\cos^2\theta_\ell)^2} \sin^2\theta_\ell \cos^2\phi \right), \tag{1.36}$$

where m_e is the electron mass, $r_0 = e^2/m_e c^2$ and E_γ is the photon energy; the other variables are in the dilepton centre-of-mass: β is the velocity of either the electron or the positron, θ_ℓ is the angle between the electron and the photon, while ϕ is the azimuth of the plane containing electron and photons measured from the plane of polarization of the photon. The term multiplying $\cos^2\phi$ gives the sensitivity to the plane of the photon polarisation.
Given that $m(e^+e^-) = 2m_e/\sqrt{1-\beta^2}$, Equation 1.36 means that for large $m(e^+e^-)/m_e$ the sensitivity to photon polarisation is smaller. Equation 1.36 can be further simplified by integrating over $\cos\theta_\ell$:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\phi} = A(1 + a\cos^2\phi) \tag{1.37}$$

where a gives the sensitivity to the photon polarisation. The coefficient a depends on $\cos\theta_\ell$ and $m(e^+e^-)$ as it is shown in Figures 1.9 and 1.10. From these plots one can conclude that the process of pair creation is sensitive to photon polarisation only when the mass of the pair is small, and that the sensitivity does not depend in a critical way on $\cos\theta_\ell$.

Maximon et al [84] did these calculations in a classical way and confirmed the validity of the Weizacker-Williams approximation. Their formulae are more complicated as they take into account the screening effect of the nucleus, but the conclusion is the same: the sensitivity to the photon polarisation decreases as

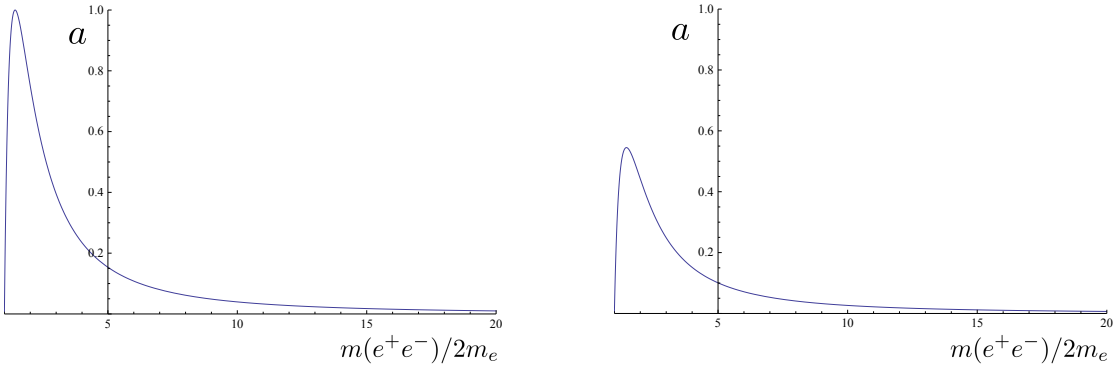


Fig. 1.9: The coefficient a of Equation 1.36 is plotted as a function of $m(e^+e^-)/(2m_e)$ at $\cos\theta_\ell = 0$ (left) and $\cos\theta_\ell = 0.7$ (right).

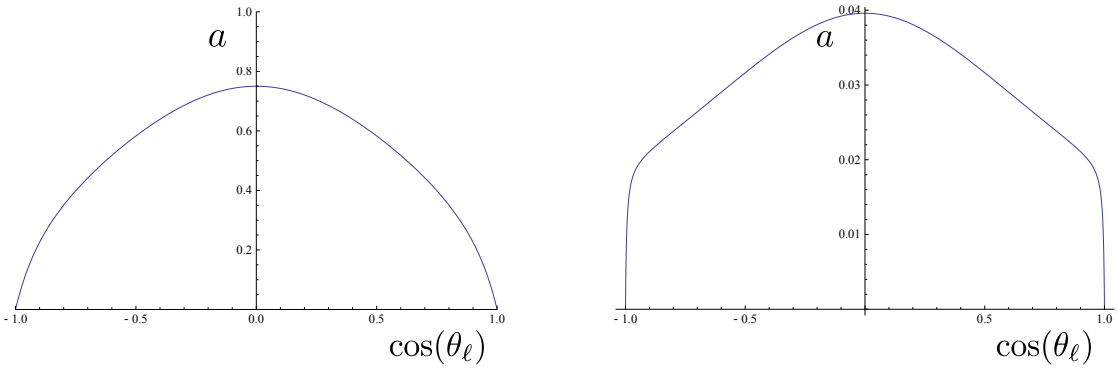


Fig. 1.10: The coefficient a of Equation 1.36 is plotted as a function of $\cos\theta_\ell$ at $m(e^+e^-) = 4m_e$ (left) and $m(e^+e^-) = 20m_e$ (right).

$1/m(e^+e^-)^2$. However, as explained in the following, in this low-mass region it is not possible to measure the plane of the e^+e^- pair using the LHCb detector.

Reconstruction of the e^+e^- pair

To access the polarisation of the photon, one needs to reconstruct the plane of the e^+e^- pair. For low masses, a first difficulty is that the e^+ and e^- form a very small angle and their tracks can be exactly superimposed in the detector, preventing the definition of a plane. This could be mitigated by using detectors at a large distance downstream of the origin of the pair. Nonetheless, another limiting factor is the multiple scattering of the $e^\pm$. The size of the effect can be evaluated in the following way: after conversion of the photons in a radiator, multiple scattering in this same radiator will give a transverse momentum $p_\perp$ to each lepton in one

plane with a gaussian width that can roughly be aproximated [85, 86] with

$$15\sqrt{X/X_0}\ \text{MeV}/c\,, \tag{1.38}$$

where X/X_0 is the thickness of the scaterring medium in terms of radiation lengths[3]. With a radiator of 10% X_0, $p_\perp$ is about $5\,\text{MeV}/c$. The e^+e^- mass measured downstream for pairs created with very small masses will be about $5\,\text{MeV}/c^2$ multiplied by $\sqrt{2}$ because of the two particles and another $\sqrt{2}$ for x and y plane. This would mean that for $m(e^+e^-) < 10\,\text{MeV}/c^2$ the measurement of the orientation of the dielectron plane is dominated by multiple scattering.
This is the case in LHCb as argumented in Section 3.1 and made clear by Figure 3.1 where the bias on the reconstructed ϕ angle is plotted against the dilepton invariant mass. Therefore, the $B^0 \to K^{*0}\gamma$ process cannot be used to measure the polarisation of the photon using the LHCb detector. What would be needed is a much much lighter detector. Indeed, gamma-ray telescope experiments are being proposed to measure gamma-ray polarisation via the pair production in a very light detector such as a TPC [87–89].

[3] A radiation legth is both the mean distance over which a high-energy electron loses all but $1/e$ of its energy by bremsstrahlung, and 7/9 of the mean free path for pair production by a high-energy photon (see [7] and references therein).

Chapter 2

The LHCb detector at the LHC

2.1 The Large Hadron Collider

The CERN Large Hadron Collider (LHC) [90] is the world's most powerful particle collider. It is housed in the 27 km circular tunnel built to house the Large Electron Positron (LEP) collider. It was designed to accelerate proton beams up to an energy of 7 TeV and thus to collide them with a centre-of-mass energy of 14 TeV. Protons are progressively accelerated through the chain of accelerators shown in Figure 2.1. At an energy of 450 GeV proton beams are injected in the LHC, which accelerates them up to the final energy using 16 radiofrequency (RF) cavities. Along the perimeter of the LHC, 1232 superconducting Niobium-Titanium dipole magnets are used to create a magnetic field up to 8.3 T in order to bend protons along the LHC ring. The two beam channels required to keep the proton beams in opposite orbits are accommodated in the same cryostat and share a common yoke which provides them opposite magnetic fields. The LHC also features 392 quadrupole magnets to focus the beams. Nominal proton beams are composed of bunches of 1.2-1.4 $\times 10^{11}$ protons separated by 25 ns (40 MHz).
The LHC was operated during 2011 and 2012 at collision energies of 7 and 8 TeV respectively and with bunches having 50 ns time spacing. Beams are collided in four interaction points (IPs) along the LHC ring. A total of seven detectors share these four interaction points to study the physics produced by the unprecedented multi-TeV scale particle collisions. Two IPs are surrounded by huge General-Purpose Detectors (GPDs), ATLAS [91] and CMS [92], which were designed to study collisions producing high transverse momentum (p_{T}) particles. The two experiments share a wide physics programme including general searches for signals from NP particles produced on shell, b and t quark physics and the detection of the Higgs boson which indeed they observed for the first time in 2012 [93, 94]. Now, a large part of their physics programme is devoted to measure precisely the properties of the recently discovered scalar boson, which is highly sensitive to contributions

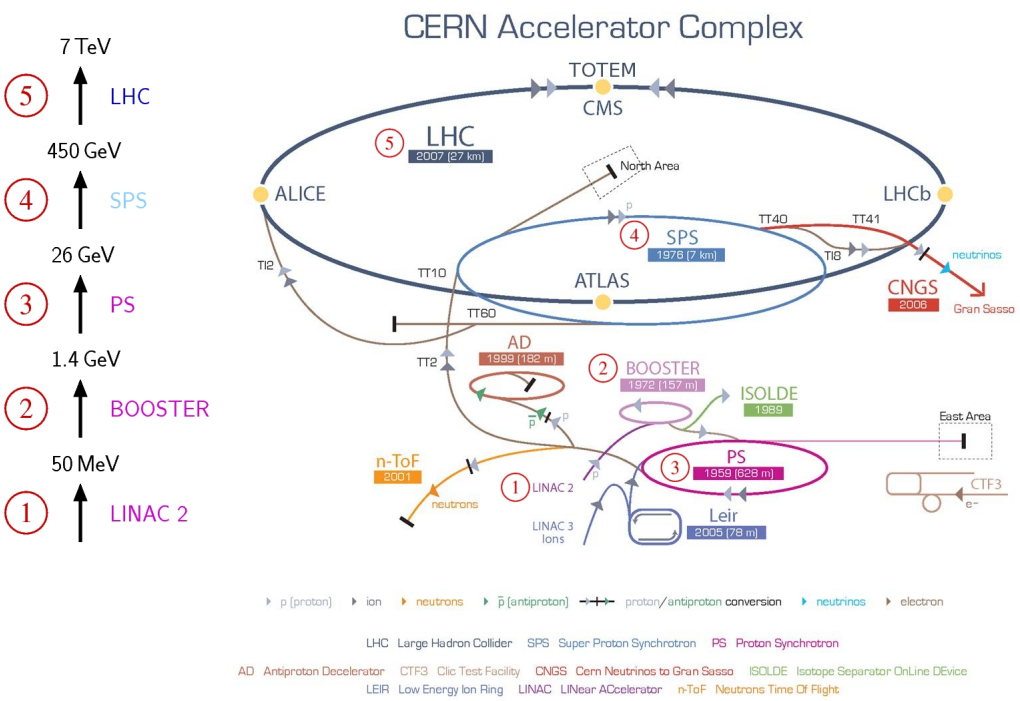


Fig. 2.1: Representation of the different steps, within the CERN accelerator complex, necessary to bring protons at 7 TeV. Protons are extracted from an H source and then progressively accelerated through the LINear ACCelerator 2 (LINAC 2) up to 50 MeV, the BOOSTER up to 1.4 GeV, the Proton Synchroton (PS) up to 26 GeV and the Super Proton Synchroton (SPS) up to 450 GeV. Then, they are injected in the LHC where they reach the final collision energy. (Figure adapted from CERN-DI-0606052 ©CERN Geneva)

from BSM physics.
The other experiments being operated at the LHC are: LHCb [95] (which is described in detail in Section 2.2), ALICE [96], TOTEM [97], LHCf [98] and MoEDAL [99]).

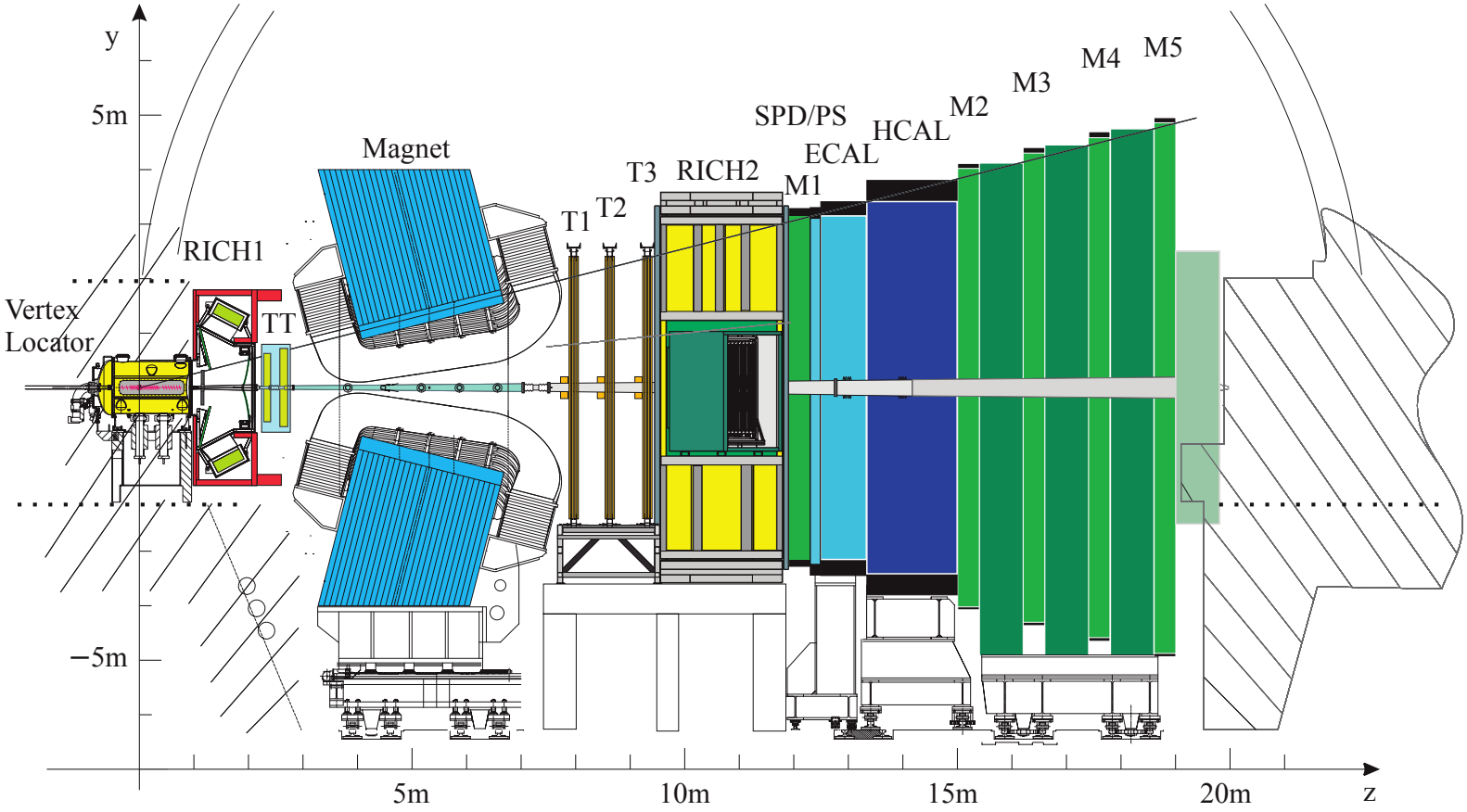


Fig. 2.2: Cross section of the LHCb detector [100]. The y and z axes of the default reference frame are defined in the figure. The x axis is perpendicular to the plane of the page oriented with the standard right-hand rule.

2.2 The LHCb experiment

LHCb is a dedicated heavy flavour physics experiment whose main goal is the indirect search of NP effects by studying *CP* violation and rare decays of beauty and charm hadrons. It exploits the large beauty and charm production cross-sections [101, 102] at the LHC to collect large samples of heavy flavour decays. The LHCb detector is a single-arm forward spectrometer designed to cover only the forward pseudorapidity[1] region $1.8 < \eta < 4.9$. Although this coverage corresponds to only the ∼4% of the solid angle, $b\bar{b}$ pairs are mainly produced at small angles, hence 25% of them are in the LHCb acceptance. Indeed, due to the large momentum differences involved in the parton-parton scattering, which can be up to few TeV in LHC pp collisions, and to the relatively small invariant mass necessary to produce a $b\bar{b}$, the pair is boosted in the forward or backward direction. For comparison, GPDs are covering pseudorapidities in the range $-2.4 < \eta < 2.4$ (more than 90% of the solid angle) and get ∼45% of the $b\bar{b}$ pairs (see the comparison in Figure 2.3). However instrumenting a smaller solid angle is more convenient and furthermore, the LHCb forward coverage selects b-mesons with larger boosts whose resulting displaced b-decay vertices provide a good signal signature and allow to precisely measure decay lifetimes. Last but not least, the forward detector layout allows to put dead material outside the acceptance, thus minimizing multiple scattering and improving the momentum resolution.

[1] $\eta = -\ln\tan\frac{\theta}{2}$, where θ is the polar angle with respect to the beam axis z.

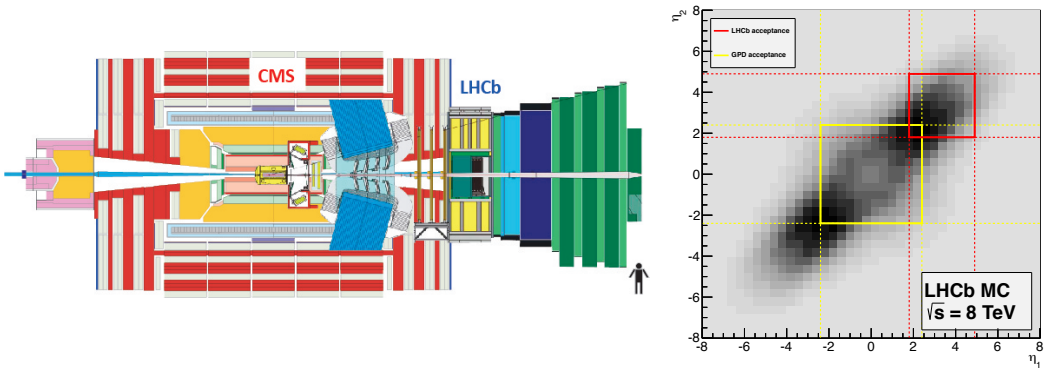


Fig. 2.3: Comparison of the angular acceptances of LHCb and GPDs. On the right the LHCb detector scheme is superimposed to the CMS one. ©CERN LHCb-PHO-GEN-2012-001. On the right the $b\bar{b}$ production rates as a function of the their pseudorapidities η_1 and η_2 is plotted for collision centre-of-mass energies of 8 TeV (the picture is very similar for 7 TeV). Red (yellow) dotted lines mark the LHCb (GPDs) acceptance limits. ©CERN bb_ProductionAngles

Compared to B-factories at e^+e^- colliders (*e.g.* *BABAR* [103] and Belle [104]), the LHCb experiment can count on higher cross-sections, a larger boost and the fact that all species of b hadrons are produced. However, the $b\bar{b}$ cross section in pp collisions is about 200 times smaller than the total inelastic cross section, implying a more polluted environment. Indeed, LHCb has worse b flavour tagging efficiency and more difficulties in the reconstruction of events with neutral or missing particles, meaning for example that inclusive branching ratio measurements are very challenging. Nonetheless, it has collected the world's largest sample of exclusively reconstructed charm and beauty decays and is now the leading beauty and charm physics experiment.

In order to optimise the LHCb performance in this high-multiplicity hadronic environment, LHC collisions for LHCb are delivered at a considerably lower luminosity than for ATLAS and CMS. Yet, this value is two times larger than the design luminosity, since, during the 2011 run, it was demonstrated that the physics output is not compromised by the corresponding pile-up μ_{vis} (number of visible interactions per beam-beam crossing). Peak luminosity and pile-up values are shown in Figure 2.4 (left). Furthermore, as shown in Figure 2.4 (right), the luminosity delivered to LHCb is levelled by adjusting the transversal beam overlap. This minimises the effects of luminosity decay during an LHC fill, allowing to maintain the same trigger configuration and to reduce systematic uncertainties. During 2011 and 2012, LHCb recorded an integrated luminosity of about $3\,\text{fb}^{-1}$.

The strategy to separate beauty and charm hadron decays from background is

[1] `LHCb-PHO-GEN-2012-001`, ©2012 CERN, for the benefit of the LHCb Collaboration

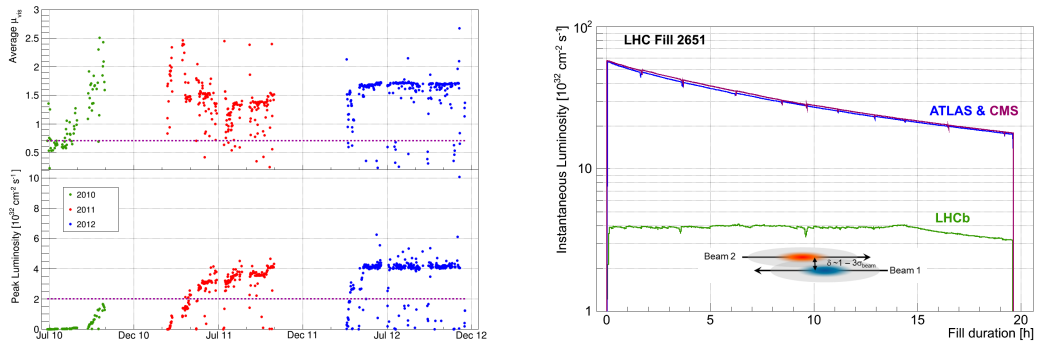


Fig. 2.4: (left) Peak luminosity and pile-up in LHCb as a function of time during LHC Run 1. (right) Instantaneous luminosity for ATLAS, CMS and LHCb during LHC fill 2651. In LHCb, after reaching the desired value of $4\times10^{32}\,\text{cm}^2\text{s}^{-1}$, the luminosity is kept stable by adjusting the transversal beam overlap [105]. The run shown is exceptionally long: after 15h beams in LHCb were already head-on and luminosity started to decay before the beam dump.

to exploit both the displaced vertex signature and the high transverse momenta of final state particles. In LHCb charged particle tracks and momenta are measured via a VErtex LOcator (VELO) detector positioned very close to the interaction point, a bending magnet with a bending power of about 4 Tm and tracking stations upstream and downstream the magnet as described in Section 2.2.1. Tracking provides also very good mass resolution, which together with good particle identification is a fundamental tool for a flavour physics experiment aiming to well separate a large number of decay channels and eventually isolate the rarest. Electrons and photons are identified by the silicon pad detector (SPD), the preshower (PS) and electromagnetic calorimeter (ECAL), while another calorimeter is dedicated to hadrons (HCAL) (see Section 2.2.3). Ring imaging Cherenkov (RICH) detectors are used to distinguish different types of hadrons as shown in Section 2.2.2. Downstream, muon stations composed of alternating layers of iron and multiwire proportional chambers provide good identification of traversing muons (Section 2.2.4). Finally, a fast hardware trigger is supplied by the calorimeters and the muon system, which is followed by a software based trigger running a simplified version of the offline event reconstruction. The trigger system is a fundamental tool in order to benefit from the high event rate at the LHC (Section 2.2.5).

2.2.1 Tracking system and vertex reconstruction

The VErtex LOcator (VELO)

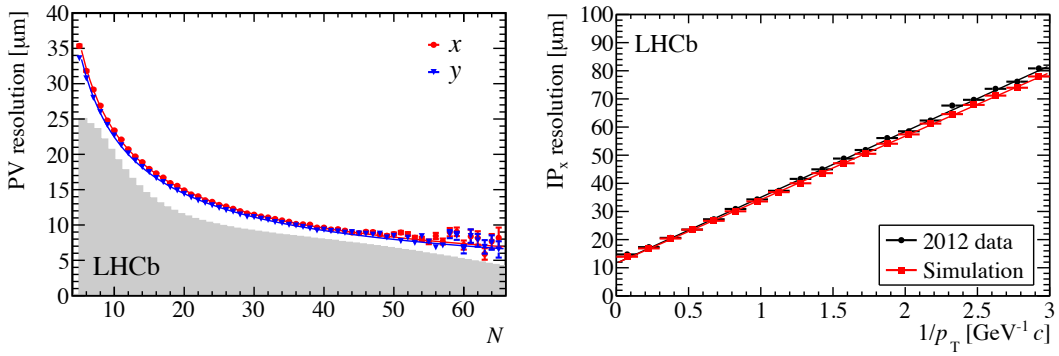


Fig. 2.5: Performance of the VELO detector. (left) Primary Vertex (PV) resolution as a function of the number of tracks composing the vertex. The x (red) and y (blue) resolutions are separately shown and the superimposed histogram (grey) shows the distribution of number of tracks per reconstructed primary vertex for events passing the high level trigger. (right) The impact parameter in x resolution as a function of $1/p_{\mathrm{T}}$. Both plots are made using LHCb data collected in 2012. [105, 106]

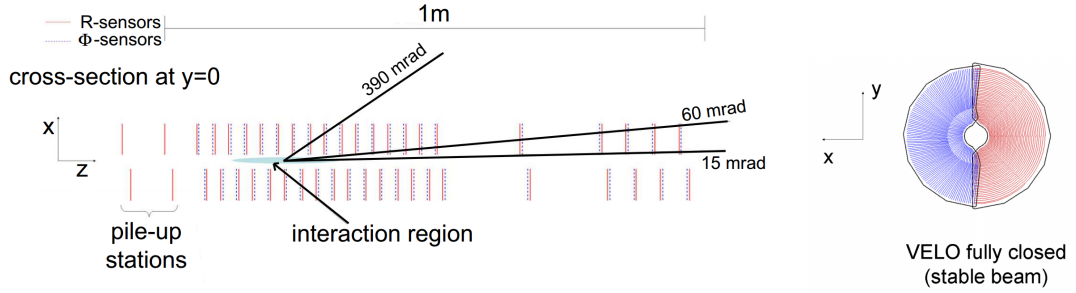


Fig. 2.6: (left) Schematic cross-section of the VELO detector. (right) Schematic front view of two VELO modules in closed position. [107]

The VELO detector provides very precise tracking of charged particles close to the interaction region allowing to measure vertices, which is important for both background rejection and precise measurements of lifetimes (it is required to resolve fast B_s^0 oscillations). The capability of assigning a track to the right vertex and of distinguishing secondary vertices is given by the resolution on the impact parameter (IP), *i.e.* the distance of closest approach of a track to a vertex. The VELO detector provides very good IP resolution (see Figure 2.5) thanks to fine silicon strips placed as close as 8 mm from the beam axis. This performance is crucial to fight against the combinatorial background coming from candidates in which one track is associated with the wrong decay vertex (this background is often the dominant one in the analysis of rare decay such as the one discussed in this work).
The detector is divided in two halves, each consisting of 21 modules mounted downstream of the interaction point and perpendicularly to the beam as shown in Figure 2.6. The number of modules is chosen such that tracks that are inside the acceptance of the rest of the tracking system (and originate up to 10 cm downstream of the interaction point) traverse at least 3 modules. Each module is equipped with silicon strips oriented in the r and ϕ directions to measure the azimuthal and radial coordinates of charged particles (this strip geometry is chosen to speed up pattern recognition). The pitch within a module varies from 38 µm at the inner radius of 8.2 mm, increasing linearly to 102 µm at the outer radius of 42 mm.
Two additional stations are placed upstream of the interaction point. They are used to aid the instantaneous measurement of luminosity. To protect the detector while LHC beams are not squeezed at the IP, the two VELO halves are retracted 35 mm from the beam axis.

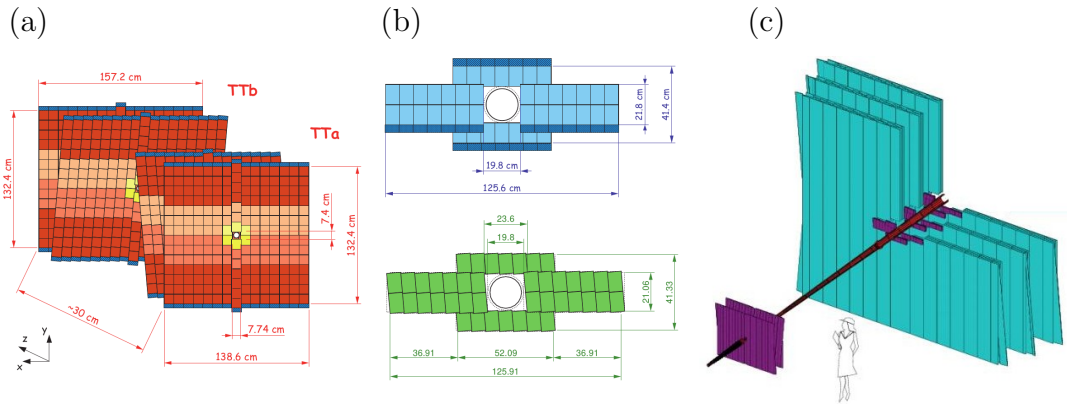


Fig. 2.7: Schematic views of the LHCb trackers: (a) the two Tracker Turicensis stations, (b) the Inner Tracker and (c) the Outer Tracker (together with the TT and IT which are shown in purple). [108, 109]

Tracking stations and magnet

The magnetic field required to measure momenta of highly energetic charged particles is provided by a large dipole magnet with a bending power of 4.2 Tm which is operated at room temperature. As shown in Figure 2.2, the magnet is placed about 5 m from the interaction point and is preceded by the Tracker Turicensis (TT) and followed by the T1, T2 and T3 tracking stations. While the TT is completely instrumented with silicon microstrips, each of the T1-3 stations is composed of two detectors with different technology: the region close to the beam pipe, the Inner Tracker (IT), is instrumented with silicon microstrips, while the outer region, the Outer Tracker (OT), is composed of straw tubes (see Figure 2.7).

The TT and IT detectors use silicon microstrip sensors with a strip pitch of about 200 µm. The Outer Tracker consists of approximately 200 straw-tube modules with drift-time read-out. Each module contains two layers of drift-tubes with an inner diameter of 4.9 mm. They provide a spatial resolution of 200 µm while keeping the drift time below 50 ns. Each of the tracking stations (TT, IT and OT) has four detection layers in an x-u-v-x arrangement with vertical strips in each of the two x layers, and strips rotated by ± 5 deg in the u and v layers, in order to get a stereo view of the particle trajectory. In LHC Run 1, the LHCb tracking system gave very good performances: a relative momentum resolution below 0.5% at low momenta and at 0.9% at 100 GeV/c and as a consequence a very good dimuon mass resolution of about 5 per mille all the way up to the Υ masses (see Figure 2.8).

Tracking in LHCb has the best performances in terms of momentum resolution for tracks leaving hits in the VELO and in the various tracking stations. These tracks are called “long tracks” and are the most used for physics analyses (included

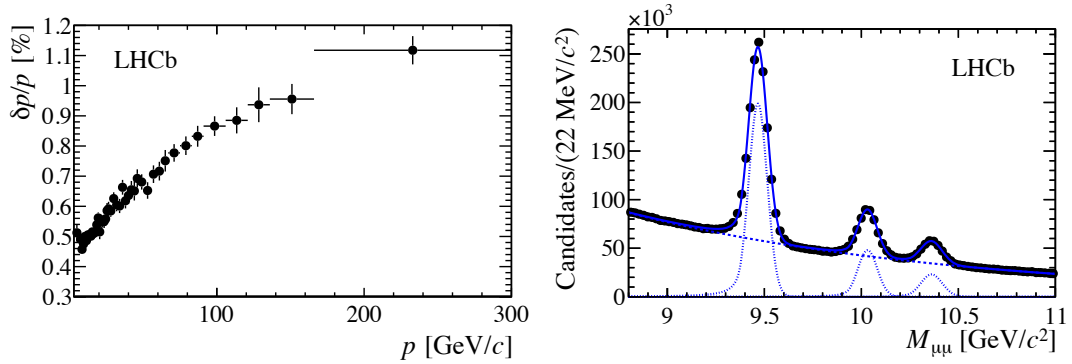


Fig. 2.8: (left) Relative momentum resolution versus momentum for long tracks from $J/\psi \to \mu^+\mu^-$ events. (right) Mass distribution for the Υ resonances $\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$ showing the excellent mass resolution of LHCb of about 5 per mille [105].

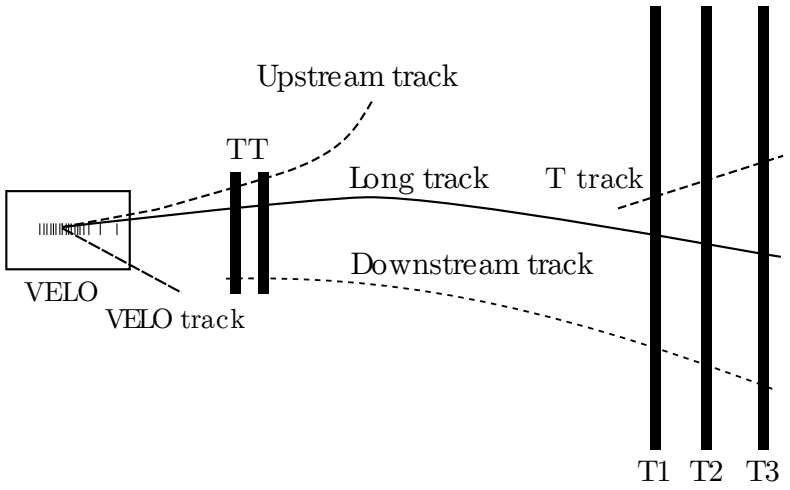


Fig. 2.9: A schematic illustration of the various track types [95].

the one discussed in this work). Nonetheless, as shown in Figure 2.9 other types of tracks are used, *e.g.* "downstream tracks" (not leaving hits in the VELO), which are important for the reconstruction of long lived particles such as K^0_S and Λ, or VELO tracks, which are useful for primary vertex reconstruction.

2.2.2 RICH detectors

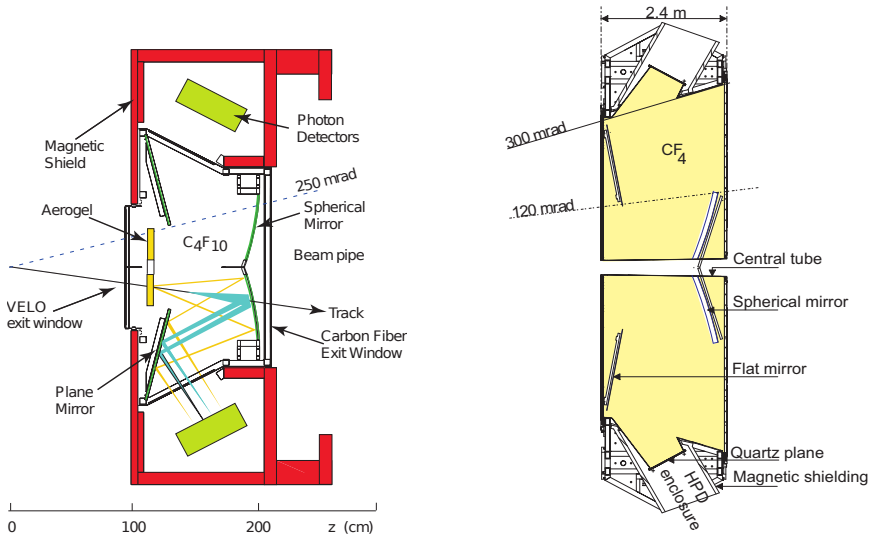


Fig. 2.10: Schematic cross section of the RICH1 (left) and RICH2 (right) detectors. [110]

Charged hadron identification in the momentum range from 2 to 100 GeV/c is achieved by two Ring Imaging Cherenkov detectors[2] (RICH1 and RICH2) one upstream and one downstream of the magnet, adjacent to the tracking stations. The upstream detector, RICH1, covers the momentum range from about 2 to 60 GeV/c and uses two radiators coupled to the same optics: a silica aerogel for low momenta (phase index of refraction $n = 1.03$) and C_4F_{10} gas for higher momenta ($n = 1.0014$). The downstream detector, RICH2, provides PID information for the high momentum range from about 15 GeV/c to 100 GeV/c, using a CF_4 gas ($n = 1.0005$) as radiator. While RICH1 covers the LHCb tracking acceptance, RICH2 has a reduced angular acceptance of ± 120 mrad (horizontal) and ± 100 mrad (vertical), as it is dedicated to the PID of particles with high momenta, which are mainly at small angles.
The complementarity of the three radiators in providing charged hadron identification in the entire momentum range is shown in Figure 2.11 (left). Both RICH1 and RICH2 use two mirrors to image the produced Cherenkov light on a plane of Hybrid Photon Detectors (HPDs) located outside the LHCb acceptance. Furthermore, the first mirror has a spherical curvature in order to focus Cherenkov light into rings on the HPDs plane. Focusing is necessary to compensate the fact that photons are emitted all along the track path in the gas (see for example the track in RICH1 in Figure 2.10).

[2]More information on the principle of PID with Cherenkov light and RICH detectors is given in Chapter 5.

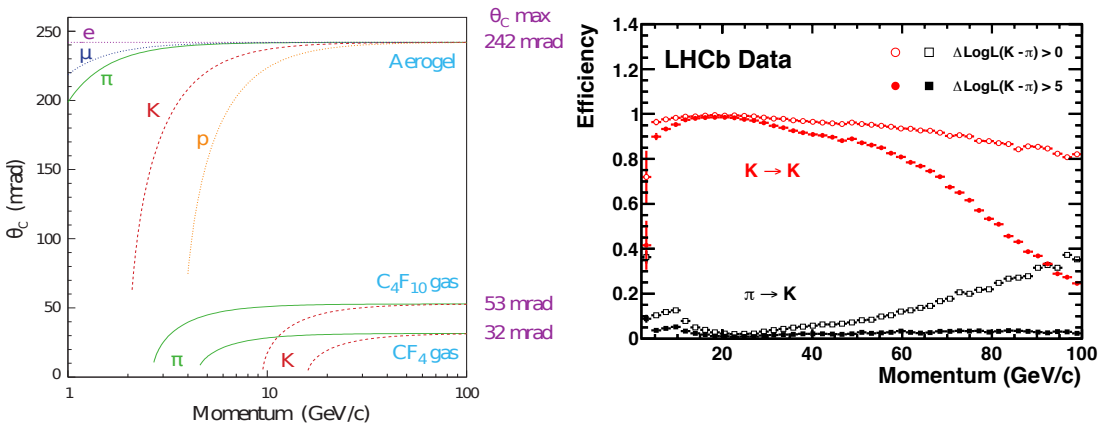


Fig. 2.11: (left) Cherenkov angle as a function of momentum for various final state particles and the three radiating media used in RICH1 and RICH2. The combination of different media allows to separate kaons and pions in a wide range of momenta as shown in the plot on the right where Kaon identification and pion misidentification probabilities are shown as a function of the charged track momentum. [111]

Being located between the VELO and the TT, RICH1 uses special lightweight spherical mirrors built from a carbon-fibre reinforced polymer rather than glass, in order to minimise the material budget. Input from the tracking system is used to compute the position of the expected Cherenkov ring. Its radius is compared to the one expected for the various particle hypotheses: π, K, p and likelihoods for each of them, $\mathcal{L}_i$, are calculated. A certain level of identification efficiency (*e.g.* kaon candidate is a kaon) and mis-identification probability (*e.g.* kaon candidate is a pion) can be chosen with a requirement on the difference of logarithms of likelihoods, $(-\Delta \log \mathcal{L})$ or `DLL`, as shown in Figure 2.11 (right). For example the performance of the RICH system for two different requirement on $(-\Delta \log \mathcal{L})$ is shown in Figure 2.11 (right) as a function of momentum. Another way of discriminating is based combining the detector information using a neural network trained on simulation events. This results in a discriminating variable, `ProbNN`, which is found to have slightly better PID performances than `DLL`.
The RICH system contributes also to the identification of electrons and muons, complementing the information from the electromagnetic calorimeter and the muon system.

2.2.3 Electromagnetic and hadronic calorimeters

The calorimeter system provides the identification of electrons, photons and hadrons and a measurement of their energies and positions. Also, high transverse energy (E_T) deposits from these particles are used in the first level of the trigger,

requiring the detector response to be as fast as possible. It is composed of a Scintillating Pad Detector (SPD), a Preshower (PS), an electromagnetic calorimeter (ECAL) and a hadronic calorimeter (HCAL). The SPD and the PS both consist of a plane of scintillator tiles, while the ECAL has shashlik-type construction [113], *i.e.* a stack of alternating slices of lead absorber and scintillator being penetrated perpendicularly by wavelength shifting fibers. The HCAL is a sampling device made from iron and scintillator tiles being orientated parallel to the beam axis. All four subdetectors use wavelength-shifting fibers to transmit scintillation light to photomultiplier tubes (PMT). As the hit density varies two orders of magnitude over the surface of the calorimeters, they are divided in regions with different granularity as detailed in Figure 2.12.

The PS and SPD consist of two identical planes of scintillator pads with a 15 mm thick lead plane in between. They are used to get information on the longitudinal development of the energy deposit in order to separate showers initiated by photons, electrons and hadrons, as shown in Figure 2.13 (left). Indeed, the lead thickness corresponds to 2.5 radiation lengths, but to only ∼6% hadronic interaction lengths, meaning that electrons and photons induce a shower much more often than hadrons. Therefore, the energy deposit in the PS allows to distinguish hadrons from electrons and photons. Furthermore, electrons are more likely than photons to produce signal in the SPD, giving it discriminating power. Additionally, the SPD hit multiplicity information is used in the hardware trigger as it is correlated to the multiplicity of the event.

The ECAL is composed of alternated layers of 2 mm thick lead and 4 mm thick scintillator tiles, for a total length of 84 cm. The total thickness is equivalent to 25 radiation lengths (X_0), allowing to fully contain electromagnetic showers. The ECAL provides a photon energy resolution that allows the study of radiative decays in LHCb. It results, for example, in a $B_s^0 \to \phi\gamma$ mass resolution of ∼90 MeV/c^2 [115] as shown in Figure 2.13.

The thickness of the HCAL is limited to 5.6 interaction lengths due to space

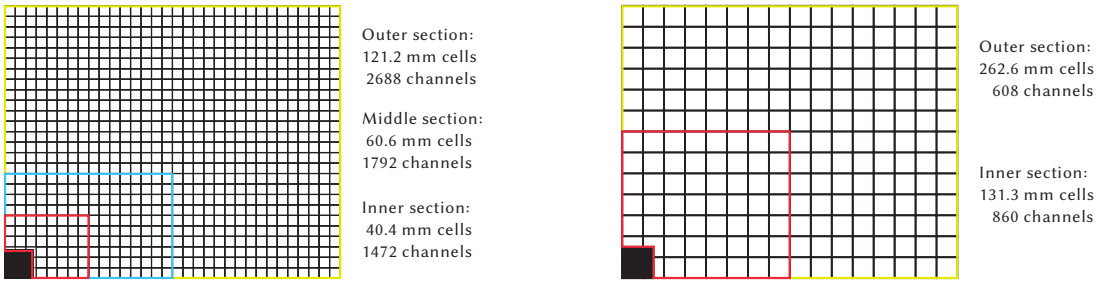


Fig. 2.12: (left) Segmentation of the SPD, PS and ECAL; (right) segmentation of the HCAL. [112]

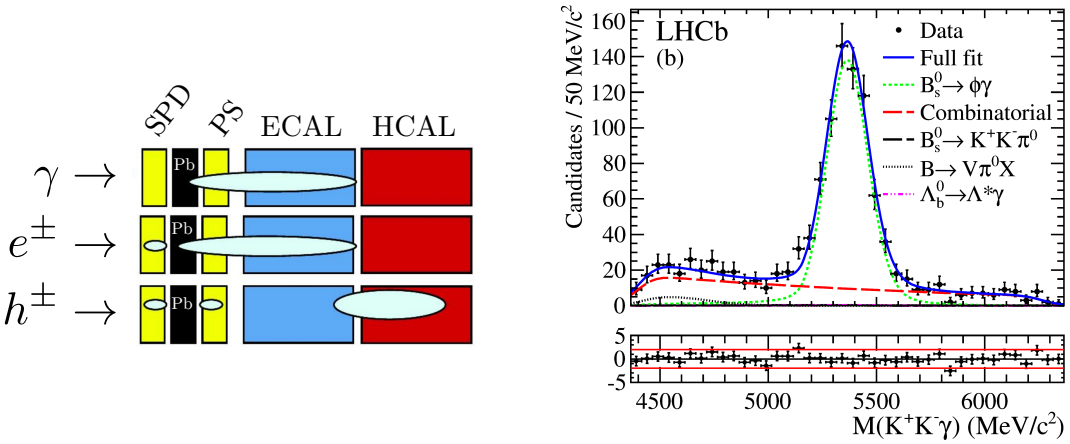


Fig. 2.13: (left) Schematic explanation of the principle of identification of photons, electrons and hadrons with the calorimeter system. Adapted from [114]. (right) Distribution of $m(K^+K^-\gamma)$ for a sample of $B_s^0 \rightarrow \phi\gamma$ events collected by LHCb with 2011 $1\,\text{fb}^{-1}$ data [115]. The mass resolution is dominated by the reconstructed photon energy by the ECAL.

constraints and less stringent requirements on the resolution. Indeed, its main purpose is to provide a trigger for charged hadrons having high transverse energy.

Electron identification and reconstruction

Electrons are identified by the energy deposit signature they leave in the calorimeter system, but also by looking for a track associated to the cluster and based on

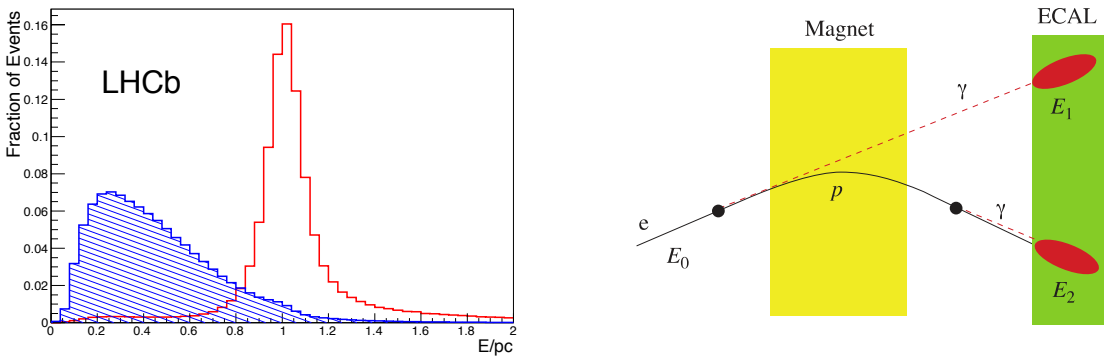


Fig. 2.14: (left) Distribution for the ECAL of E/pc for electrons (red) and hadrons (blue), as obtained from the first $340\,\text{pb}^{-1}$ recorded in 2011 [105]. (right) Schematic representation of bremsstrahlung photons emitted by an electron before and after the magnet [95].

the measurement of the ratio between energy and momentum E/pc provided by the various calorimeter devices. This is a very powerful discriminator, as shown in Figure 2.14 (left). Like for hadrons the identification probability is defined either with difference of log-likelihoods (`DLL`) by comparing the electron and pion hypotheses or with a neural network based on MC events.
The reconstruction of the electron momentum is challenging as most of the time electrons emit one or more bremsstrahlung photon upstream the magnet, which therefore impinges on the ECAL in a different cell (or is lost) as shown in Figure 2.14 (right). A lost bremsstrahlung photon shifts invariant masses involving electrons, such as the B^0 mass for $B^0 \to K^{*0}e^+e^-$, towards smaller values; however, it does not alter the ratio E/pc used for electron ID. To improve the momentum reconstruction, a dedicated bremsstrahlung recovery is used. Contributions from photon candidates, neutral clusters with transverse energy greater than 75 MeV, found within a region of the ECAL defined by the extrapolation of the electron track upstream of the magnet, are added to the measured electron momentum. Furthermore, if the same bremsstrahlung photon is associated with both the e^+ and the e^- of a dielectron pair candidate, its energy is added randomly to one of the tracks.

2.2.4 Muon system

The LHCb muon system is composed of five stations, M1 to M5, interspersed with 80 cm thick layers of iron absorbers for a total of 20 interaction lengths (meaning a muon must have a momentum of at least 6 GeV/c to traverse them all). They

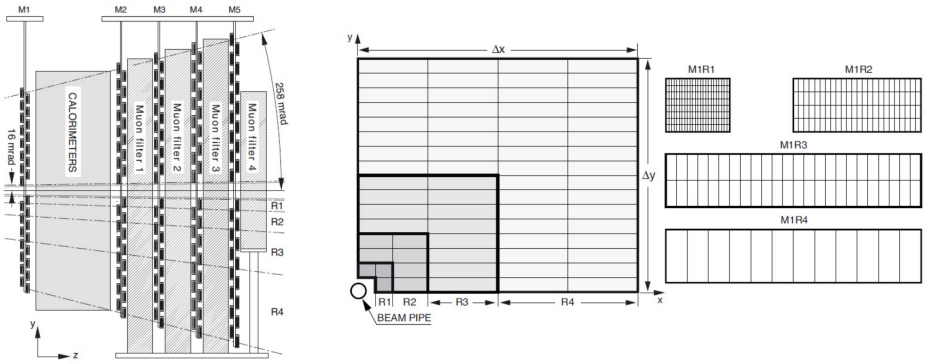


Fig. 2.15: (left) Schematic muon system cross section. (centre) A quadrant of one of the muon stations with the four separated region R1-4; each square represents a muon chamber. (right) Segmentation of the four types of muon chambers installed in the different regions R1-4. [116]

provide excellent muon identification and a fast detection of high p_{T} muons which is used for the first level of the trigger. M2-5 stations are located downstream of the calorimeter system, while M1 is upstream, in order to provide a standalone estimate of the p_{T} for the trigger. Stations are equipped with Multi Wire Proportional Chambers (MWPC) with four different regions (R1-4) of segmentation to accommodate for different occupancies with respect to the distance from the beam axis (see Figure 2.15). The highest rate region of M1 is the only one being instrumented with Gas Electron Multiplier (GEM) detectors rather than MWPC. To meet the requirement of more than 99% efficiency per muon station, each one has four active MWPC layers (except for M1 which has two).

2.2.5 Trigger system

At the 2012 luminosity and pile-up running conditions, the rate of visible beam crossings in LHCb is about 10 MHz. The purpose of the LHCb trigger system is to reduce the rate down to about 5 kHz (the highest that can be written to storage) while efficiently selecting events suitable for LHCb physics analyses [117].

The trigger consists of two levels: a Level 0 (L0) which is implemented in hardware, operates synchronously with the LHC nominal bunch crossing frequency of 40 MHz and reduces the rate down to 1 MHz, which is low enough to read out the whole detector information. The L0 trigger is based on information from the calorimeter and the muon systems and selects events containing particles with high transverse momenta as they are likely to come from particles with high mass such as B mesons. The calorimeter based L0 trigger computes the transverse energy deposited by incident particles in clusters of 2×2 cells and builds three types of L0 candidates based on the signatures explained in Section 2.2.3: `L0Hadron` for the highest-E_T HCAL cluster, `L0Electron` is the highest-E_T ECAL cluster with a PS and an SPS hit in front and `L0Photon` which is like `L0Electron` but with no SPS hit. If the E_T of one of the candidates is found to be above a certain threshold, the L0 trigger is fired. An additional condition is posed on the total number of SPD hits in order to veto events with too high multiplicity that would take too much processing time in the following trigger level.
The muon based L0 trigger looks for high p_{T} muon tracks using a fast standalone p_{T} measurement based only on the hits on the first two muon stations (25% resolution). To fire the L0 trigger, the highest p_{T} (or the product of the two highest) is required to be above a certain threshold.
E_T and p_{T} thresholds for 2011 and 2012 are summarised in Table 2.1. This output rate consists of approximately 400 kHz of muon triggers, 500 kHz of hadron triggers and 150 kHz of electron and photon triggers.

After the L0 trigger, the detector is read out and sent to the Event Filter Farm (EFF) of processors where they are analysed by the software High Level Trigger

	p_{T} or E_{T}		SPD
	2011	2012	2011 and 2012
single muon	1.48 GeV/c	1.76 GeV/c	600
dimuon $p_{\mathrm{T}1} \times p_{\mathrm{T}2}$	$(1.30\,\mathrm{GeV}/c)^2$	$(1.60\,\mathrm{GeV}/c)^2$	900
hadron	3.50 GeV	3.50-3.74 GeV	600
electron	2.50 GeV	2.72-2.96 GeV	600
photon	2.50 GeV	2.72-2.96 GeV	600

Tab. 2.1: Typical L0 thresholds used in Run I. The tightening of L0 thresholds in 2012 is a consequence of the increased luminosity and beam energy

(HLT). The HLT is composed of a first level (HLT1), where a fast selection is applied based on a partial reconstruction of the event and a second level (HLT2), where the full event is reconstructed. The same software as the offline analysis is used in the HLT. The filtering is organised in so-called trigger lines consisting of a sequence of reconstruction algorithms and selection requirements. The HLT1 performs a full 3D reconstruction of the VELO tracks and identifies primary vertices. It exploits the geometry of the VELO to do a first track reconstruction using only the VELO R-sensors. Track segments with large IP are extrapolated to the tracking stations to determine the momentum and apply minimum momentum criteria. For muon triggered events a fast muon identification is performed by matching tracks to muon chamber hits. In general, tracks are selected by requiring good track quality, large impact parameter from the Primary Vertex (PV) and large transverse momentum. The HLT2 takes as input a rate of about 80 kHz and performs a full events reconstruction. It is composed of a mixture of exclusive and inclusive lines such as the "topological lines", mainly designed to select partially reconstructed b-hadron decays, "muon lines" for events with one or two muons or "charm lines", optimised to select charm decays.

Reconstructed data are stored in "data summary tape" (dst) files for physics analysis. To allow the analysis of this huge amount of data, a first loose selection is applied, splitting it into different "stripping lines" conceived for specific analyses. Stripping selections can be revisited since new processing of dst files are run periodically.

2.2.6 The LHCb simulation

A full simulation of pp collisions and of the LHCb detector response is used. The pp collisions are generated using pythia [118] with a specific LHCb configuration [119]. Decays of hadronic particles are described by EvtGen [120], in which final-state radiation is generated using Photos [121].

The interaction of the generated particles with the LHCb detector, and its response, are implemented using the GEANT4 toolkit [122] as described in Ref. [123]. Also part of the response of the sub-detector electronics and of the hardware trigger is emulated. The HLT processing and the reconstruction are run on MC in the same way as on real data. Yet, some differences between MC and data remain, such as the average event multiplicity and the PID efficiency, which however can be easily kept under control in physics analyses.

Chapter 3

Selection of the $B^0 \to K^{*0} e^+ e^-$ decay at low q^2

The $B^0 \to K^{*0} e^+ e^-$ channel represents a privileged way to access the physics in the very low q^2 region of $b \to sll$ processes and thus to collect information on the photon polarisation in $b \to s\gamma$ as outlined in Section 1.3.3.
This chapter is about the selection of $B^0 \to K^{*0} e^+ e^-$ candidates from the LHCb data, while Chapter 4 explains the methods used for the angular analysis. The data sample used corresponding to an integrated luminosity of $3.0\,\text{fb}^{-1}$ that was collected by the LHCb experiment in pp collisions at center-of-mass energies of 7 and 8 TeV during 2011 and 2012. First of all, in Section 3.1 the analysis strategy is presented. Then, in Section 3.2 the online and offline selection of $B^0 \to K^{*0} e^+ e^-$ candidates is explained in detail. Finally, in Section 3.3 all possible backgrounds from specific channels are studied and vetoed when necessary. Then, a full mass fit of the $m(K^+\pi^- e^+ e^-)$ invariant mass distribution for the selected $B^0 \to K^{*0} e^+ e^-$ candidates is set up in Section 3.4.

3.1 q^2 range

As outlined in Section 1.3.3 the interest in using $B^0 \to K^{*0} \ell^+ \ell^-$ with electrons rather than muons is to go as low as possible in q^2 to isolate the photon pole. Therefore, the (reconstructed) dilepton-mass region chosen for this study is between $20\,\text{MeV}/c^2$ and $1\,\text{GeV}/c^2$, which corresponds to the q^2 range $[0.0004, 1]\,\text{GeV}^2/c^4$. Dilepton masses below $1\,\text{GeV}/c^2$ are chosen because at LHCb the analysis $B^0 \to K^{*0}\mu^+\mu^-$ is more sensitive than $B^0 \to K^{*0} e^+ e^-$ above $1\,\text{GeV}/c^2$ due to the larger yield. The lower bound is set at $20\,\text{MeV}/c^2$ following two experimental reasons: first, it is required to veto the background from $B^0 \to K^{*0}\gamma$ with the γ converting in the detector material to an e^+e^- pair, called hereafter

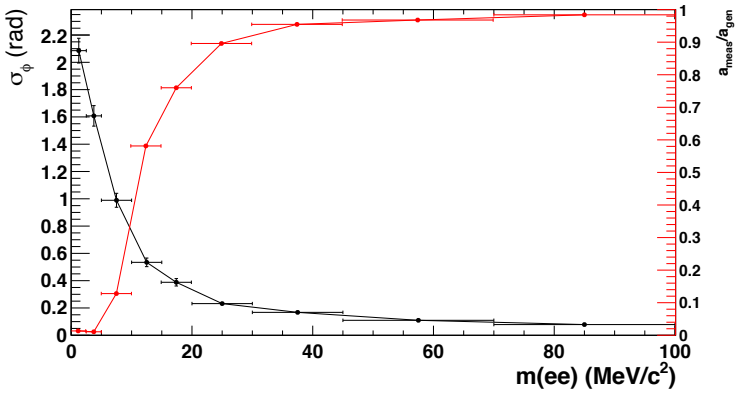


Fig. 3.1: In black the mean resolution on ϕ is plotted against the reconstructed mass of the di-electron pair for a sample of reconstructed $B^0 \to K^{*0}e^+e^-$ MC events. In red is the ratio between the measured a parameter, as defined in Equation 1.37, and the generated one in $B^0 \to K^{*0}e^+e^-$ toy MC. The measured value is always smaller as a consequence of the ϕ resolution, but the bias is negligible at high masses.

$B^0 \to K^{*0}\gamma_{e^+e^-}$; second, dielectrons with low mass yield a bad resolution in the ϕ angle between the dilepton plane and the plane defined by the decay of the K^{*0} (see definition in Appendix A.1). Indeed, a very low $m(e^+e^-)$ corresponds to a very small angle between the two electrons, which are thus quasi-collinear. Therefore, the precision in the orientation of the dielectron plane degrades at very low masses due to multiple scattering (see Figure 3.1). A bad resolution on ϕ is not suitable since it flattens all modulations of the ϕ distribution and thus it may hide effects of NP showing up as non-zero values of the angular observables $A_T^{(2)}$ and A_T^{Im}; indeed, these observables are related to modulations of type $\cos(2\phi)$ and $\sin(2\phi)$, respectively. In order to quantify the bias in the measurement of these observables that such a flattening would introduce, toy MC data are generated for the ϕ angle with a distribution of the type $1 + a\cos(2\phi)$ (corresponding to the same modulation that would be induced by a non-zero value of $A_T^{(2)}$). Then, they are smeared according to the expected ϕ resolution, which in turn is dependent on the dilepton mass. In Figure 3.1 (red curve) the bias $a_{\text{meas}}/a_{\text{gen}}$ is plotted as a function of the dilepton mass and therefore of the corresponding resolution. This plot shows clearly that below $20\,\text{MeV}/c^2$ the ϕ resolution is increasingly worse and the bias gets very large, *i.e.* $a_{\text{meas}}/a_{\text{gen}}$ becomes much smaller than 1.
For the same reason, a measurement of the photon polarisation using converted photons from $B^0 \to K^{*0}\gamma$ (see Section 1.3.4) is impossible in LHCb.
However, integrating the $B^0 \to K^{*0}e^+e^-$ differential decay rate between $20\,\text{MeV}/c^2$ and $1\,\text{GeV}/c^2$ in the dielectron mass the bias remains well below 1%, thus completely negligible with the current statistics. The mass cut is lowered down from

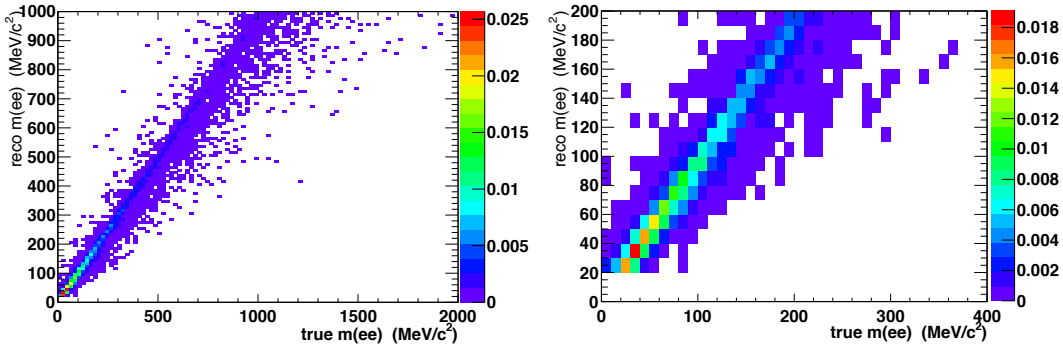


Fig. 3.2: $B^0 \to K^{*0}e^+e^-$ MC showing the reconstructed dilepton mass versus the true one. The cut in the reconstructed dilepton mass is applied between 20 MeV/c^2 and 1000 MeV/c^2. The z-axis represents the fraction of events in the bin. The right plot is zoomed version of the left one.

30 to 20 MeV/c^2 compared to the previous paper measuring the branching ratio with 1 fb^{-1} [124]. Figure 3.2 shows the reconstructed dilepton mass versus the true one for a simulated sample of $B^0 \to K^{*0}e^+e^-$ events. Dilepton mass migration from higher true- to lower reconstructed masses is due to the loss of one or more bremsstrahlung photons.

3.2 Selection

3.2.1 Online selection

Signal events are selected online by requiring that $B^0 \to K^{*0}e^+e^-$ candidates fulfill the L0 hardware electron or hadron triggers (see their definition in Section 2.2.5), or that the trigger is fulfilled independently of any of the decay products of the signal candidate. The latter case is called Trigger Independent of Signal (TIS) and is usually caused by the decay products of the other b hadron in the event. Given that the $m(K\pi e^+e^-)$ signal shape as well as the angular acceptance show a dependence on the L0 trigger, the entire $B^0 \to K^{*0}e^+e^-$ analysis is performed separately for three *trigger categories* defined to be mutually exclusive as follows (see Figure 3.3):

- **"L0Ele"**: the signal candidate has triggered the electron L0 line;
- **"L0Had"**: the signal candidate has triggered the hadron L0 line but not the electron one;
- **"L0TIS"** : the event has triggered the TIS L0 line and no candidate has triggered the electron or hadron lines.

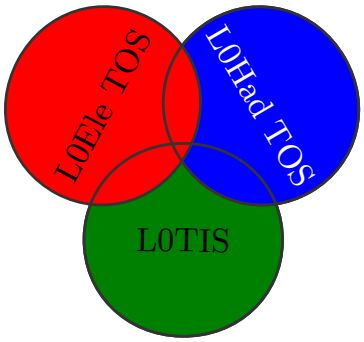


Fig. 3.3: Venn diagram for the definition of the three trigger categories.

Trigger On Signal (TOS) categories are preferred over the Trigger Independent of Signal category (TIS) in order to avoid relying on the MC description of generic b-hadron decays. The L0Ele category is preferred over L0Had as it has a clearer signal signature.
Concerning the software trigger, candidates are required to have triggered at first level (HLT1) a line requiring a track with transverse momentum $p_{\mathrm{T}} > 1.7\,\mathrm{GeV}/c$ and be inconsistent with originating from the primary vertex (PV). At second level (HLT2), the software trigger requires a two-, three- or four-tracks secondary vertex with a significant impact parameter (IP) with respect to pp primary vertices. Also, a multivariate algorithm is used for the identification of secondary vertices consistent with the decay of a b hadron [125].

3.2.2 Preselection

For the preselection of $B^0 \to K^{*0} e^+ e^-$ signal candidates a specific "stripping line" is developed. It consists of loose cuts on the transverse and absolute momenta of the final state particles, requirements to have good quality tracks, *i.e.* asking for small $\chi^2_{\mathrm{Track}}/\mathrm{ndof}$, and that tracks are coming from a secondary vertex by asking a large displacement from the primary vertex in terms of χ^2 of the impact parameter (χ^2_{IP}). Also some loose PID cuts are applied on both hadrons and electrons on the `DLL` variables defined in Sections 2.2.2 and 2.2.3. Moreover, a cut is put on the probability for a track to be a ghost, *i.e.* a duplicate track containing the same hits as another one or resulting from a cluster from one track being split in two. Good vertexing is required on the dilepton, K^{*0} and B^0 vertices by a cut on $\frac{\chi^2_{\mathrm{Vertex}}}{\mathrm{ndof}}$. Also, a small θ_{flight} is required, which means a small angle between the B momentum direction (calculated summing up final-state-particle momenta), and the direction of flight of the B, *i.e.* the direction from the PV to the B decay vertex. The e^+e^- invariant mass requirement is composed of a low-mass window below $1.2\,\mathrm{GeV}/c^2$

and of one between 2.2 and 4.2 GeV/c^2 to select $B^0 \to J/\psi(e^+e^-)K^{*0}$ events to be used as a control channel. In the low-mass window, no minimum cut is applied to the dilepton mass, and therefore a large number of $B^0 \to K^{*0}\gamma_{e^+e^-}$ are allowed to pass the stripping cut. These events are also used as a control channel for the $B^0 \to K^{*0}e^+e^-$ analysis. All cuts are listed in Table 3.1. These cuts being applied,

Particle	Condition
B^0	$\frac{\chi^2_{\text{Vertex}}}{\text{ndof}} < 16$, $\quad |m(B^0) - 5280| < 1000\,\text{MeV}/c^2$, $\theta_{\text{flight}} < 45\,\text{mrad}$
K^{*0}	$\frac{\chi^2_{\text{Vertex}}}{\text{ndof}} < 16$, $\quad |m(K^{*0}) - 892| < 150\,\text{MeV}/c^2$
e^+e^- (J/ψ)	$\frac{\chi^2_{\text{Vertex}}}{\text{ndof}} < 16$, $\quad m(e^+e^-) = [0, 1200]$ $([2200, 4200])$ MeV/c^2
K	$p_{\text{T}} > 400\,\text{MeV}/c$, $\quad p > 3\,\text{GeV}/c$, $\quad \frac{\chi^2_{\text{Track}}}{\text{ndof}} < 3$, $\quad \chi^2_{\text{IP}} > 4$, $\quad$ PIDK > -5
π	$p_{\text{T}} > 250\,\text{MeV}/c$, $\quad p > 2\,\text{GeV}/c$, $\quad \frac{\chi^2_{\text{Track}}}{\text{ndof}} < 3$, $\quad \chi^2_{\text{IP}} > 4$, $\quad$ PIDK < 10
$e^\pm$	$p_{\text{T}} > 200\,\text{MeV}/c$, $\quad \frac{\chi^2_{\text{Track}}}{\text{ndof}} < 3$, $\quad \chi^2_{\text{IP}} > 1$, $\quad$ PIDe > -2

Tab. 3.1: Cuts applied for the offline preselection.

a multivariate discriminator based on a Boosted Decision Tree [126,127] is trained with the input variables listed in Table 3.2. Together with p_{T}, χ^2_{IP} and θ_{flight} also the χ^2 of the flight distance (FD) of the dielectron, K^{*0} and B^0 vertices from the primary vertex are used as discrimant variables. The signal sample used consists of $B^0 \to K^{*0}e^+e^-$ Monte Carlo, while the background sample is composed of events from 2011 LHCb data having a B mass value larger than 5.6 GeV/c^2. This new BDT-based preselection is more efficient than a cut-based one that was developed previously [128] by a factor of about 1.4 [129].

Particle	BDT Variable			
B^0	p_{T},	χ^2_{IP},	χ^2_{FD},	θ_{flight}
K^{*0}	p_{T},	χ^2_{IP},	χ^2_{FD}	
e^+e^- (J/ψ)	p_{T},	χ^2_{IP},	χ^2_{FD}	
K, π, e^+, e^-	p_{T},	χ^2_{IP}		

Tab. 3.2: Variables used to train the Boosted Decision Tree for the stripping.

3.2.3 Simulation - data corrections

While there is in general good agreement between data and MC for the samples used in this analysis, some well-known differences are corrected a-posteriori assigning a weight to MC events.

Correction of the event occupancy

A significant difference consists in the event occupancy being largely underestimated in the MC. A way to accommodate this data-MC difference consists in correcting the SPD hits multiplicity distribution, as it is highly correlated to the event occupancy. Therefore, a re-weighting of the MC events is applied in order to match the SPD multiplicity distribution found in data.
The weights are calculated using the $B^0 \to J/\psi(e^+e^-)K^{*0}$ channel, since a very clean signal can be extracted from data (see Fig 3.4). Indeed, exploiting the narrow J/ψ resonance, the B^0 mass is re-calculated rescaling the electron momenta so that $m(e^+e^-)$ matches the world average value $m_{J/\psi} = 3096\,\mathrm{MeV}/c^2$ [130]. The resulting $m(K^{*0}e^+e^-)$ distribution has a very clean signal as shown in Figure 3.4 and is fitted with a double Crystal Ball[1] function [131] (CB) and an exponential, the two CBs sharing the same μ_B, α and n parameters. The fitted PDF is used to apply the *sPlot* technique [132] in order to extract the SPD multiplicity distribution for the signal only (see Figure 3.5 (left)). Then, weights are extracted from the ratio of data and MC distributions (see Figure 3.5(right)).

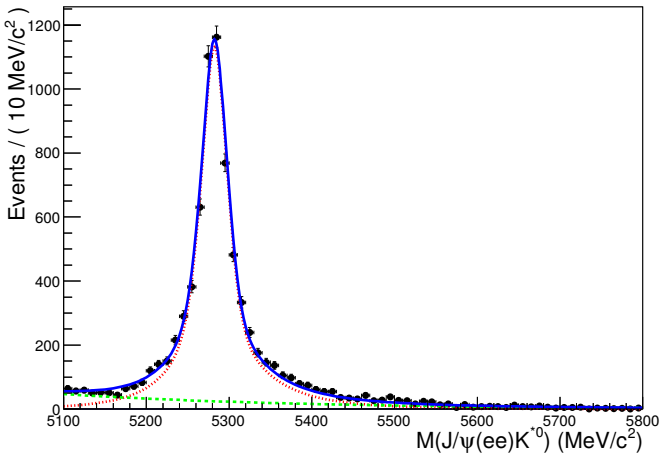


Fig. 3.4: The fit of the $B^0 \to J/\psi(e^+e^-)K^{*0}$ data that is used for the *sPlot* technique.

The occupancy is slightly correlated to the invariant mass shape of the signal because of the addition of cluster energies in the ECAL to correct the electron momenta for bremsstrahlung photons emitted before the magnet. However, the $m(K^+\pi^-e^+e^-)$ mass calculated with the constraint on $m(e^+e^-) = m_{J/\psi}$ is not correlated, allowing to apply the *sPlot* technique. The occupancy has also a sizeable effect on the matching of the *isolation variables* distributions which are used to discriminate signal from background. These variables are defined as the number

[1] More details on the Crystal Ball distribution are given in Section 3.4.1

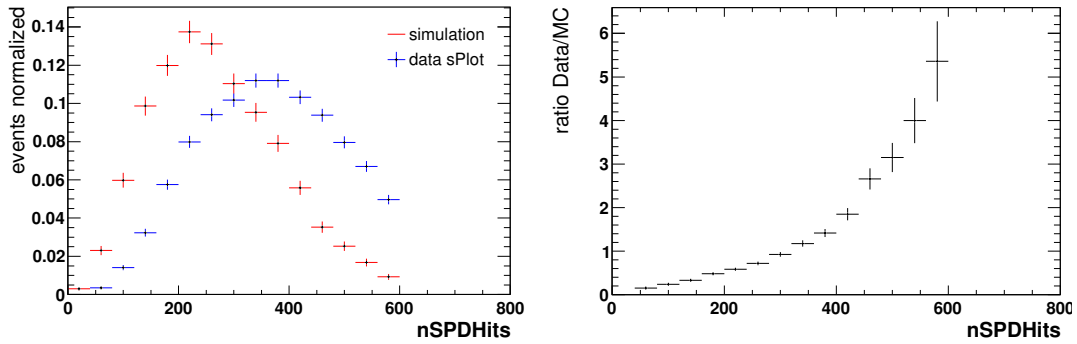


Fig. 3.5: (left) distribution of the number of hits in the SPD for $B^0 \to J/\psi(e^+e^-)K^{*0}$ MC and data (where the signal distribution is selected using the *sPlot* technique. (right) The ratio between MC and data distributions of the number of hits on the SPD.

of good two-track vertices one of the candidate signal tracks can make with any other track in the event [133], and are therefore correlated to the event occupancy. As an example one can see in Figure 3.6 the distribution of the K isolation variable before and after the re-weighting compared to an *sPlot* of data for the signal proxy $B^0 \to J/\psi(e^+e^-)K^{*0}$. The small, but significant disagreement with data goes away once the weighting is applied. The same happens for the isolation variable distributions of the other final-state particles.

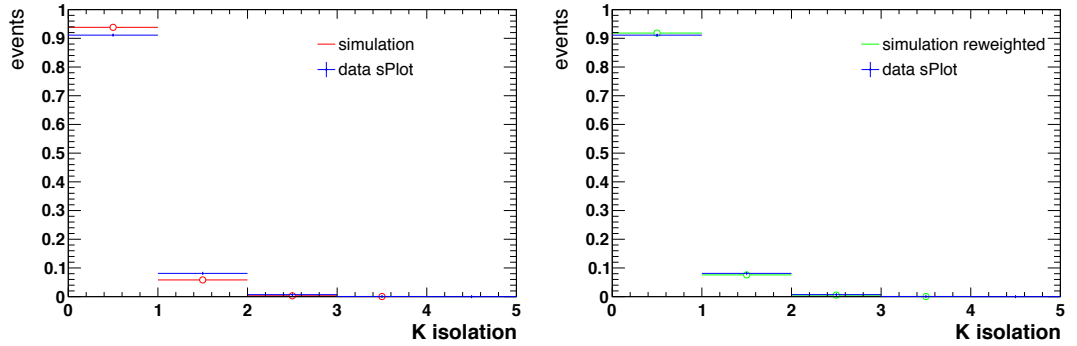


Fig. 3.6: Effect of the re-weighting of the MC on the distribution of the K isolation variable. $B^0 \to J/\psi(e^+e^-)K^{*0}$ MC is compared to data (plotted with the *sPlot* technique) before (left) and after (right) the re-weighting is applied.

Correction of the PID response

The PID response is not very well simulated in the LHCb MC. Therefore, rather than applying the cut on the PID variables from the MC, events are re-weighted

according to data-based PID efficiency tables[2]. Some golden modes that can be reconstructed without the use of information from the RICH detectors are used to provide samples of pure $\pi^{\pm}$, $K^{\pm}$ and $e^{\pm}$ candidates. Then, these samples are used to calculate the efficiency to select a certain particle with a given set of PID cuts in bins of particle momentum p, rapidity η and number of tracks in the event. Indeed, these three variables are the most correlated to the PID response in LHCb. Then, for each MC $B^0 \to K^{*0} e^+ e^-$ event, the expected efficiency to pass the given PID requirement is extracted. This efficiency is simply calculated as the product of the efficiencies obtained for each particle.

Correction of the L0 trigger response

The response of the electron and hadron L0 trigger lines is corrected using tables built from data samples. Data samples containing $K^{\pm}$, $\pi^{\pm}$ and $e^{\pm}$ candidates are selected from the unbiased L0TIS line and the probability of a certain particle track to have triggered a certain TOS line is estimated from the ratio:

$$\epsilon_{\text{TOS}} = \frac{\text{events TIS} \wedge \text{TOS}}{\text{events TIS}}. \tag{3.1}$$

This trigger efficiency is calculated in bins of the transverse energy the track deposits in the calorimeter (either the HCAL or the ECAL) and differentiating among calorimeter regions (see Section 2.2.3). Then, for each MC event the probability for the signal particles to have triggered the hadron L0 trigger line is calculated as $P_{\texttt{L0Had}} = \epsilon(\pi) + \epsilon(K) - \epsilon(\pi)\epsilon(K)$. On the other hand, the two electron candidates in the event are used to compute the probability of triggering the electron L0 line $P_{\texttt{L0Ele}} = \epsilon(e^+) + \epsilon(e^-) - \epsilon(e^+)\epsilon(e^-)$. Then, a weight is computed for each of the trigger categories defined in Section 3.2.1:

- "L0Ele" category: simply taking as a weight $P_{\texttt{L0Ele}}$.
- "L0Had" category: weight with $P_{\texttt{L0Had}} \times (1 - P_{\texttt{L0Ele}})$
- "L0TIS" category: weight the MC `L0TIS` line with $(1 - P_{\texttt{L0Had}}) \times (1 - P_{\texttt{L0Ele}})$.

No efficiency $\epsilon_{\texttt{TIS}}$ can be calculated using the technique described by Equation 3.1. Nonetheless, since these events were triggered by tracks not belonging to the signal decay, their efficiency is expected to be independent on the kinematic of the decay considered.

[2]For readers in LHCb, the `Urania/PIDCalib` package version `v2r3` was used

3.2.4 Multivariate Analysis

In order to optimise the selection, a multivariate analysis (MVA) is used rather than a simple selection based on rectangular cuts. This kind of approach provides better performance in the discrimination between signal and background as it is capable of combining the information of all discriminating variables together, rather than treat each of them independently. This way, it can fully exploit the information on the correlation between them and optimise the extraction of the information being useful to separate signal from background events.
It is based on a "supervised machine learning" algorithm which takes as input samples of events being previously labelled as signal or background and provide as output an optimal discriminator to classify unlabelled events. The algorithm used for this analysis is based on a Boosted Decision Tree (BDT)[3] [126, 127] with the "gradient" boosting technique, since it can exploit non-linear correlations between variables and it is known to have good performances with samples where the signal yield is orders of magnitude smaller than the background one. The samples used for the training are the following:

- *signal sample* from the $B^0 \to K^{*0}e^+e^-$ MC. It was not taken from data since the signal proxy $B^0 \to J/\psi(e^+e^-)K^{*0}$ is too different in terms of kinematics from the signal $B^0 \to K^{*0}e^+e^-$ events in the very low $m(e^+e^-)$. The MC is re-weighted for MC/Data differences as explained in Section 3.2.3.

- *background sample* from LHCb data (2011 and 2012). The main goal is to reject background of combinatorial type, therefore the sample is taken from the upper sideband above the B mass ($m(K^{*0}e^+e^-) > 5600$ MeV/c^2). The sample is highlighted in the distribution shown in Figure 3.7 and consists of $\sim 12 \times 10^3$ events.

BDT training

Applying the BDT classifier to the data it was trained on may result in a bias of the data sample. However, the very low yield of the $B^0 \to K^{*0}e^+e^-$ channel does not allow to use part of the data for the training of the BDT. Therefore, a "crossed-BDT" approach is used: first of all, the data sample is split randomly in two samples (labelled *sample A* and *sample B*). Then, two BDTs are built, the first (*BDT 1*) is trained, tested and optimised on *sample A* while the second, *BDT 2*, on *sample B*. Then, it is possible to avoid any danger of bias by applying *BDT 1* to *sample B* and *BDT 2* to *sample A*. Half of each sample is used for training,

[3] built within the TMVA4 package which is implemented in ROOT

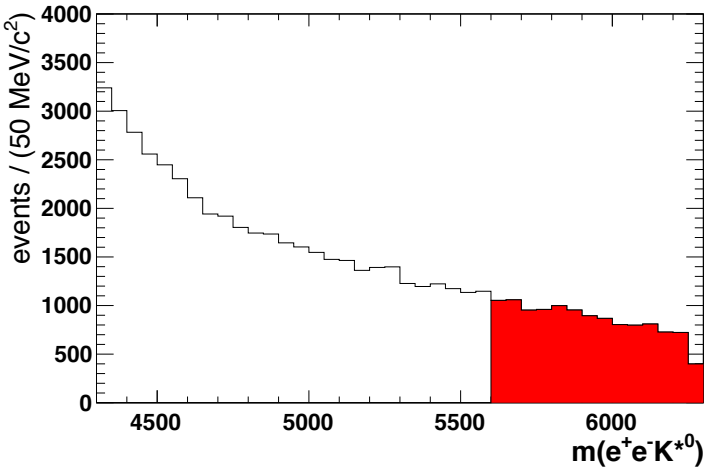


Fig. 3.7: The background sample that is used to train the BDT is obtained from the upper sideband (red) of $B^0 \to K^{*0}e^+e^-$ data.

while the other half is used for testing and optimizing. This results in training each BDT on $\sim$ 7000 signal events and $\sim$ 3000 background events.

The BDT is trained using kinematics-related variables, namely, the p_{T} and the χ^2_{IP} for all final state particles as well as for the dilepton pair, the K^{*0} and the B^0. Also, the quality of the e^+e^-, K^{*0} and B^0 vertices (χ^2_{Vertex}), the θ_{flight} angle and *isolation variables* related to each track totalling 25 degrees of freedom (listed in Table 3.3) As a conservative choice, no PID variable is given as input to the BDT training, since they are not well reproduced in the MC. The distributions of a choice of these discriminating variables is shown in Figure 3.8, while all distributions of the two samples are reported in Appendix B.1. In Figure 3.9 is the linear-correlation matrix for the discriminating variables in the signal and background samples used to train *BDT 1*. The BDT classifier output variable is normalised to take values in the range between -1, for events classified as background, and $+1$, for events classified as signal. Its distribution for the signal and background sample (and for the two BDTs trained) are shown in Figure 3.10. The BDT classifier distributions for the testing and training samples are in good agreement, meaning there is no significant over-training. The agreement is quantified by the *Kolmogorov-Smirnov test*, whose results are on Figure 3.10. The curves for the rejection of background versus the efficiency on signal calculated for the testing samples are in Figure 3.11 and show that the two BDTs behave very similarly. These curves do not correspond to the performances the BDT has on real data, as the proportion of signal and background for the training samples are quite different from data; however an optimisation is carefully carried out in Section 3.2.4.
As a cross-check, the BDT response on the signal proxy $B^0 \to J/\psi\,(e^+e^-)K^{*0}$

Particle	Variable				
B^0	p_{T},	$\log(\chi^2_{\mathrm{IP}})$,	$\log(\chi^2_{\mathrm{FD}})$,	$\log(\chi^2_{\mathrm{Vertex}})$,	θ_{flight}
K^{*0}	p_{T},	$\log(\chi^2_{\mathrm{IP}})$,	$\log(\chi^2_{\mathrm{FD}})$,	$\log(\chi^2_{\mathrm{Vertex}})$	
(e^+e^-)	p_{T},	$\log(\chi^2_{\mathrm{IP}})$,	$\log(\chi^2_{\mathrm{FD}})$,	$\log(\chi^2_{\mathrm{Vertex}})$	
K	p_{T},	$\log(\chi^2_{\mathrm{IP}})$,	Isolation		
π	p_{T},	$\log(\chi^2_{\mathrm{IP}})$,	Isolation		
e^+	p_{T},	$\log(\chi^2_{\mathrm{IP}})$,	Isolation		
e^-	p_{T},	$\log(\chi^2_{\mathrm{IP}})$,	Isolation		

Tab. 3.3: Variables used to train the Boosted Decision Tree.

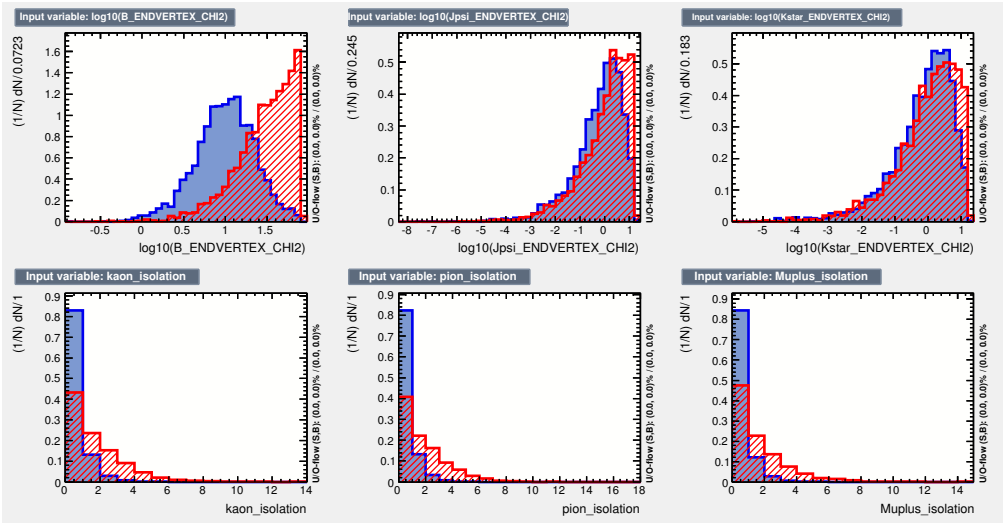


Fig. 3.8: The distributions of a choice of the variables used in BDT training. In red is the distribution of the background sample, in blue the signal sample.

is studied for both a MC sample and a data sample re-weighted with the *sPlot* technique. As shown in Figure 3.12, data and MC have similar responses in the region of interest, validating the training on a MC sample. The disagreement of few percent in the region of high BDT value is negligible for the optimisation procedure.

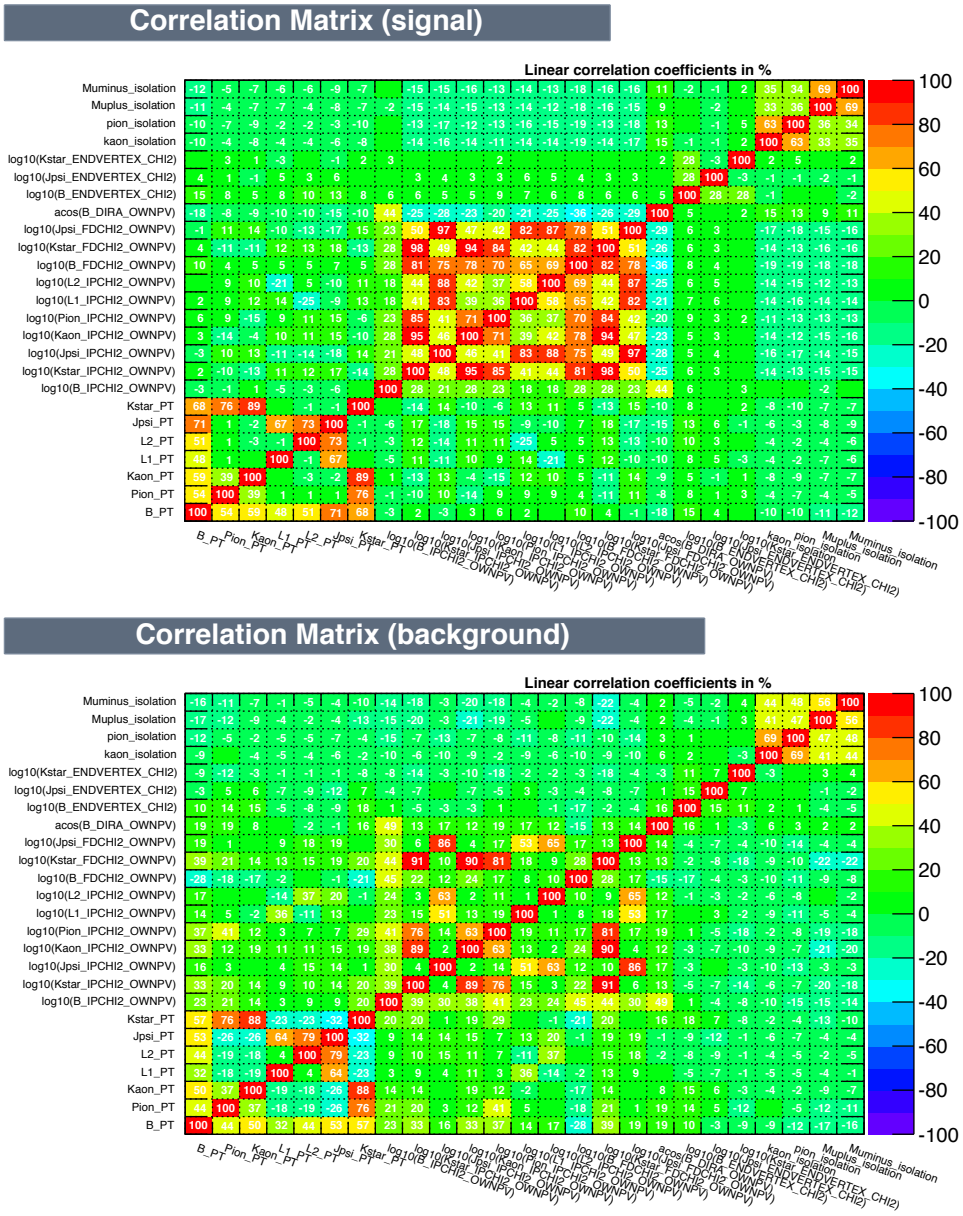


Fig. 3.9: Linear correlations between the discriminating variables in the signal (top) and background (bottom) samples used to train *BDT 1*.

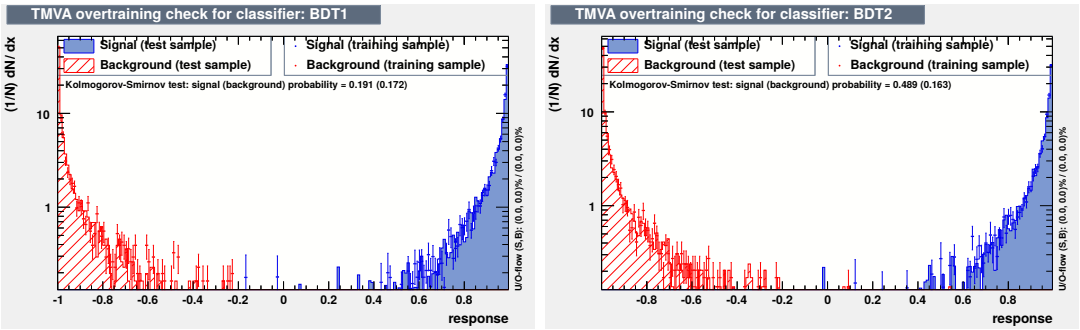


Fig. 3.10: Response of *BDT 1* (left) and *BDT 2* (right). Signal is in blue and background in red. The distributions for the testing sample (dots) is superimposed to the training sample (shaded) for both signal and background. The two samples are in good agreement for both signal and background, showing that the BDT is not over-trained.

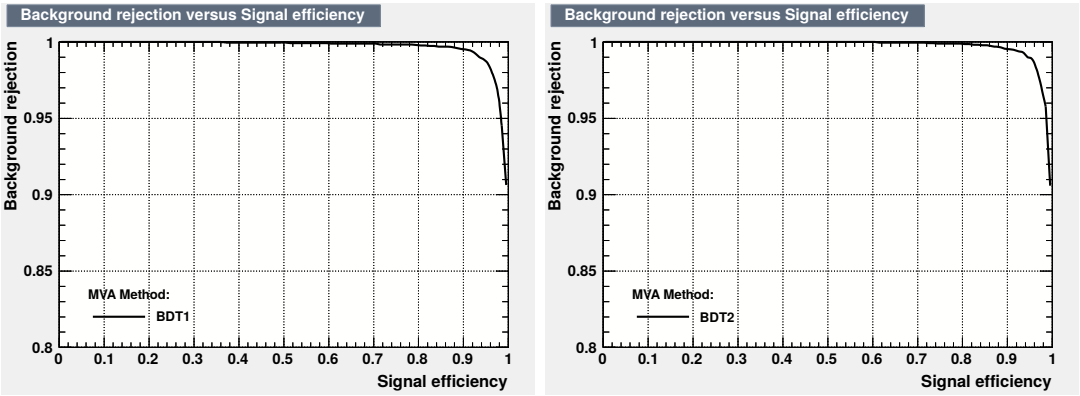


Fig. 3.11: Background rejection efficiency versus signal selection efficiency for the two BDTs trained.

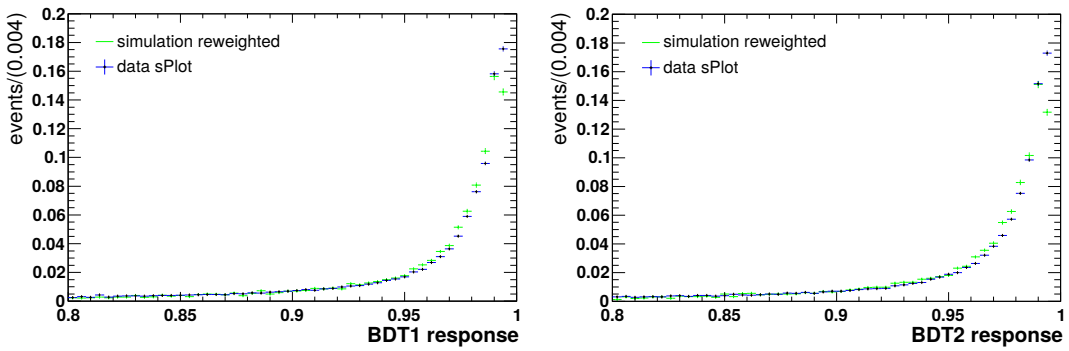


Fig. 3.12: BDT response for $B^0 \to J/\psi(e^+e^-)K^{*0}$ MC reweighted (green) and data using the *sPlot* technique (blue). Only the region of the signal (high BDT values) is plotted here. The two BDTs trained are shown: *BDT 1* on the left and *BDT 2* on the right. The last bin is empty (*i.e.* there are no BDT values being above 0.995) because of an artifact due to the normalization of the BDT value to be between -1 and 1. The differences between data and MC may be due to the presence of the S-wave component in the data.

Optimisation

A cut on the BDT classifier output variable (BDT cut) has to be chosen in order to reject efficiently background events while retaining signal events. The chosen optimisation metric is the so-called "signal significance" $S/\sqrt{S+B}$, where S (B) is the expected signal (background) yield.

The optimisation of the BDT cut is done together with that of the PID cuts for the $e^{\pm}$ and for the K. Indeed, their values may be correlated to the BDT cut in terms of background rejection efficiency.
The PID cut for the π is instead fixed to `ProbNNpi` > 0.2, since the efficiency does not depend strongly on it. The PID variable used for the K and the π are of the `ProbNN` type (see Section 2.2.2). These variables are found to be more efficient than `DLL` variables. However, for electrons `DLL` variables are used instead as the new `ProbNN` variables were not giving better performances (the neural network training was not optimised for electrons yet).
As for the whole analysis, the optimisation is performed separately for the three trigger categories.

Given that the $B^0 \to K^{*0}e^+e^-$ MC does not reproduce well the PID variables distributions, the number of expected signal events S is calculated using the number of signals found in data for the $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ channel and rescaled for the different selection efficiencies and branching ratios using the selection efficiencies extracted from MC:

$$S(B^0 \to K^{*0}e^+e^-) = S(B^0 \to J/\psi\, K^{*0}) \cdot \frac{\epsilon^{\text{sel}}_{\text{MC}}(B^0 \to K^{*0}e^+e^-)}{\epsilon^{\text{sel}}_{\text{MC}}(B^0 \to J/\psi\, K^{*0})} \cdot \frac{\mathcal{B}(B^0 \to K^{*0}e^+e^-)}{\mathcal{B}(B^0 \to J/\psi\, K^{*0})} \quad (3.2)$$

The number of $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ events in data is extracted with a fit to the $m(K^{*0}e^+e^-)$ invariant mass calculated with the e^+e^- pair mass fixed to the J/ψ PDG central value as it is done for the *sPlot* technique used in Section 3.2.3. An example is shown in Figure 3.13 (left).
Selection efficiencies are extracted from MC samples of $B^0 \to K^{*0}e^+e^-$ and $B^0 \to J/\psi\,(e^+e^-)K^{*0}$. $\mathcal{B}(B^0 \to J/\psi\,(e^+e^-)K^{*0})$ is taken from the PDG while $\mathcal{B}(B^0 \to K^{*0}e^+e^-)$ is taken from [134] and rescaled for the larger mass window of $20 - 1000$ MeV/c^2.

The background B is estimated from data mass-sidebands (under 4800 MeV/c^2 and above 5500 MeV/c^2), with the signal mass region blinded. An exponential distribution is fitted to the data, which describes the background in the sidebands to a first approximation (see Figure 3.13 (right) for an example). Then, the integral of the exponential is calculated in the signal region to estimate the background B.

The cut values for the BDT, the $e^{\pm}$ `PIDe` and the K `ProbNNk` are chosen

altogether (to take into account correlations) maximizing the significance value as shown in Figure 3.14. Each of the two BDTs is optimized separately on its optimisation sample (defined in Section 3.2.4). The best values for the cuts are chosen to be the same for the two BDTs as their response is found to be very similar. However, some differences among trigger categories are found for the BDT cut and the PID of the $e^{\pm}$. The error on the calculated value for the significance is of the order of $\sim 2\%$ for the L0Ele trigger category, $\sim 4\%$ for L0Had and $\sim 3\%$ for L0TIS. Large oscillations correspond to few unimportant cut values for which the background sideband fit did not converge correctly. The chosen working point for the three trigger categories are reported in Table 3.4. BDT cuts are applied in a crossed way to avoid any bias: the cut on *BDT 1* is applied to *sample B* and the one on *BDT 2* on *sample A*.

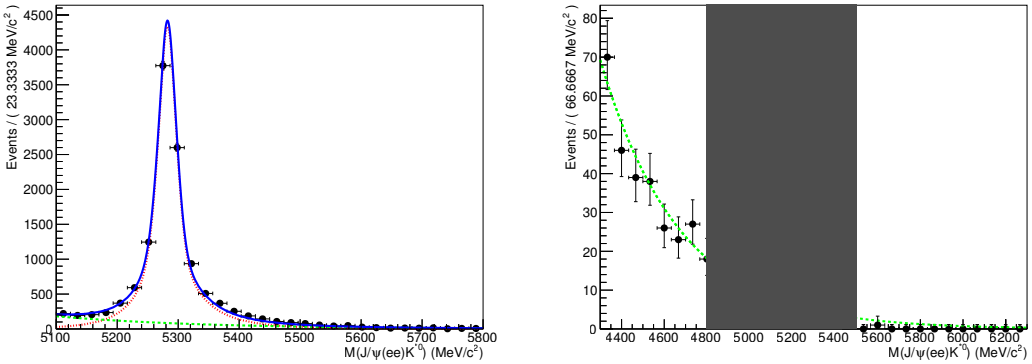


Fig. 3.13: Optimisation of the BDT cut. (left) Example of a fit for the extraction of the number of signal events $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ with a B mass fit to data for the trigger category L0Ele. The invariant mass of the e^+e^- pair is constrained to the J/ψ mass central value. (right) Example of a fit for the extraction of the expected background of $B^0 \to K^{*0}e^+e^-$ candidates in the signal region from an exponential fit of the data sidebands for the category L0Ele.

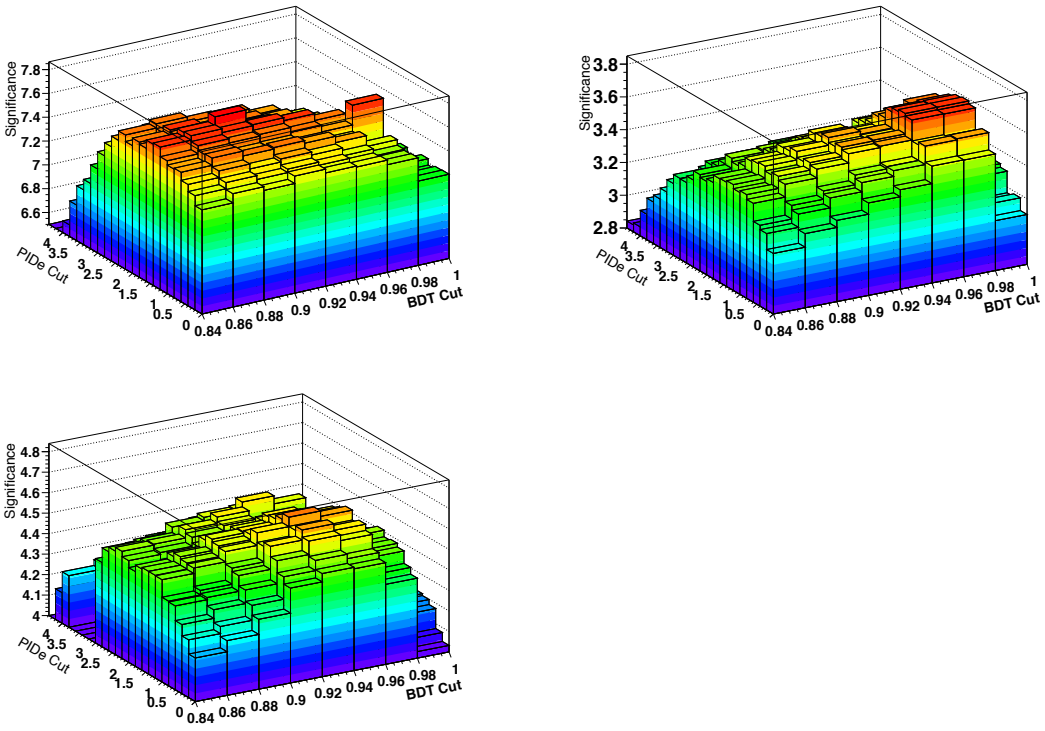


Fig. 3.14: Significance $S/\sqrt{(S+B)}$ for different combination of the cut on the *BDT 1* and the $e^{\pm}$ `PIDe` and for a K `ProbNNk` cut fixed to 0.05. Top left: L0Ele category; Top right: L0Had category; Bottom left: L0TIS category. Half of data *sample A* is used for this optimisation.

	BDT	$e^{\pm}$ `PIDe`	K `ProbNNk`	$S/\sqrt{S+B}$
L0Ele	> 0.88	> 1.2	> 0.05	7.5
L0Had	> 0.94	> 1.2	> 0.05	3.9
L0TIS	> 0.96	> 1.6	> 0.05	4.8

Tab. 3.4: Result of the optimisation of the cuts on the BDT value and the PID requirements of the $e^{\pm}$ and the K candidates. The PID requirement for the π candidate is fixed to `ProbNNpi` > 0.2

3.3 Specific backgrounds studies

Being $B^0 \to K^{*0}e^+e^-$ a rare decay channel it is sensitive to many sources of background. Some decay channels with properties similar to the signal are expected to be picked up as signal candidates. These background sources are studied in this chapter and, when possible, reduced by veto cuts based on their specific properties.

3.3.1 The $B^0 \to D^-e^+\nu$ decay

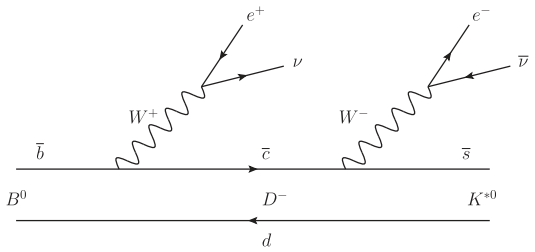


Fig. 3.15: Feynman diagram for the $B^0 \to D^-e^+\nu$ decay with $D^- \to K^{*0}e^-\bar{\nu}$

A specific background is coming from the $B^0 \to D^-e^+\nu$ decay in which the D^- in turn decays semileptonically to $K^{*0}e^-\bar{\nu}$. Its first-order Feynman diagram can be seen in Figure 3.15. This decay has a branching ratio four orders of magnitude larger than $B^0 \to K^{*0}e^+e^-$ and it may pass the selection cuts when the two neutrinos carry low momentum.
In the selection made for the branching ratio measurement of $B^0 \to K^{*0}e^+e^-$, a cut is applied to the events with an invariant mass of the K^{*0} and the e^- below 1900 MeV/c^2 ($m_{D^\pm} \simeq 1870\,\text{MeV}/c^2$) assuming a very low momentum neutrino. Unfortunately this cut biases the angular distribution of $\cos\theta_\ell$. Indeed, asking for $m(K^{*0}e^-) < 1900\,\text{MeV}/c^2$ one selects a e^+ with high energy and, as $\cos\theta_\ell$ is roughly equivalent to the energy asymmetry of the two electrons

$$\cos\theta_\ell \sim \frac{E_{e^+} - E_{e^-}}{E_{e^+} + E_{e^-}}, \tag{3.3}$$

this mass cut removes events having $\cos\theta_\ell$ close to +1. The lower mass sideband with $m(K^{*0}e^+e^-) < 4800$ MeV/c^2 (Figure 3.16 (left)) is dominated by this background and indeed it presents a clear peak close to $\cos\theta_\ell = +1$. Therefore, as an angular analysis aiming to use the $\cos\theta_\ell$ information is envisaged, this background is vetoed through a cut on $\cos\theta_\ell < 0.8$, which has almost the same effect as the $m(K^{*0}e) < 1900$ MeV/c^2 cut. In order not to introduce systematic

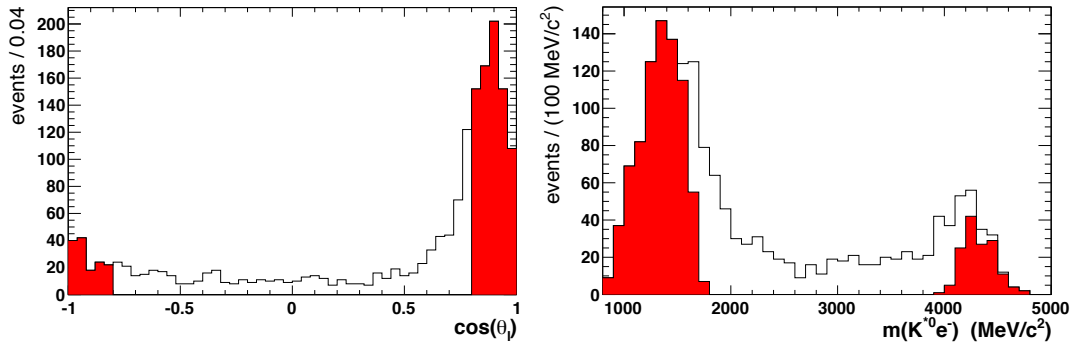


Fig. 3.16: Distribution of $\cos\theta_\ell$ (left) and of the invariant mass $m(K^{*0}e^-)$ (right) from data events with $m(K^{*0}e^+e^-) < 4800$ MeV/c^2. In this low-mass region the contribution from $B^0 \to D^- e^+ \nu$ is dominant. The red area corresponds to events cut by the symmetric requirement on $|\cos\theta_\ell| < 0.8$.

effects on the measurement of A_T^{Re} (which is related to the forward-backward asymmetry and therefore to an asymmetry in the $\cos\theta_\ell$ distribution), a cut is applied also at $\cos\theta_\ell > -0.8$ so that the range is symmetric. The cut on $\cos\theta_\ell$ is compared to the one on $m(K^{*0}e^-)$ in Figure 3.16, the area in red being the one that is removed by the cut on $\cos\theta_\ell$. The percentage of $B^0 \to K^{*0}e^+e^-$ signal events lost applying this cut is as small as 10% and not 20%, as one would expect for a flat distribution of $\cos\theta_\ell$. This is caused by the fall of the angular acceptance towards $\cos\theta_\ell$ values of ± 1, which corresponds to a low efficiency in selecting events with one of the two electrons carrying a very low energy. Furthermore, given that in the differential decay rate (Equation 1.32) terms sensitive to $A_T^{(2)}$ and A_T^{Im} are proportional to $\sin^2(\theta_l)$, the loss of events at $\cos\theta_\ell \sim \pm 1$ have a very small impact on the sensitivity to these observables.

3.3.2 The $B^0 \to K^{*0}\gamma$ decay followed by γ conversion

The branching fraction of $B^0 \to K^{*0}\gamma$ has been measured to be $\mathcal{B} = (4.33 \pm 0.15) \times 10^{-5}$ [7], *i.e.* two orders of magnitude larger than $B^0 \to K^{*0}e^+e^-$. When the photon converts in the detector material to a dielectron pair, the $B^0 \to K^{*0}\gamma_{e^+e^-}$ decay has the same final state as $B^0 \to K^{*0}e^+e^-$. In LHCb, around 40% of photons convert before the calorimeter, and although only a small fraction of these, of the order of 10%, convert in the VELO and are reconstructed as long tracks, the $m(K^{*0}e^+e^-)$ mass peaks at the B mass exactly as the signal. It is thus important to veto this signal-like background.

A very effective veto is the requirement on the reconstructed mass of the dielec-

tron to be higher than 20 MeV/c^2, which is required also for the angular analysis as explained in Section 3.1. Moreover, the e^+e^- pair from $B^0 \to K^{*0}\gamma_{e^+e^-}$ has a vertex at the point of conversion of the photon, but it may still be reconstructed as originating from the B^0 decay when the e^+e^- vertex is determined with a large error. Therefore, another effective veto is a requirement on the error on the reconstructed z-coordinate of the e^+e^- pair $\sigma_Z(e^+e^-)$ to be lower than 30 mm. The 2D distributions in $m(e^+e^-)$ and $\sigma_Z(e^+e^-)$ are shown in Figure **??** for $B^0 \to K^{*0}e^+e^-$ (left) and $B^0 \to K^{*0}\gamma_{e^+e^-}$ (right). The two variables are anticorrelated as a small

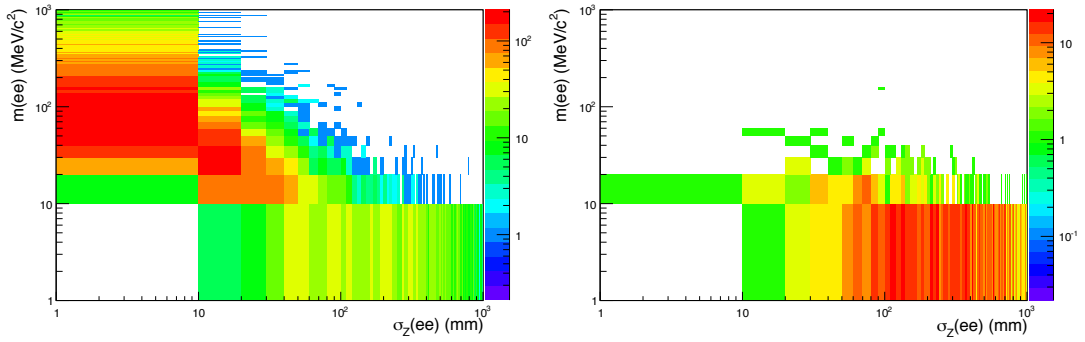


Fig. 3.17: The distribution of MC events as a function of the invariant mass $m(e^+e^-)$ and $\sigma_Z(e^+e^-)$ for $B^0 \to K^{*0}e^+e^-$ (left) and $B^0 \to K^{*0}\gamma_{e^+e^-}$ (right). Please note that all three axes are in logarithmic scale.

dielectron mass corresponds to a small opening angle which in turn causes a bad measurement of the z-coordinate of the dilepton vertex. They are not completely anticorrelated, though, therefore cutting on both gives a more efficient veto.
The $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ channel is used to check for MC-data differences in the distribution of $\sigma_Z(e^+e^-)$. As shown in Figure 3.18, data and MC agree reasonably well. The effect of adding the requirement on $\sigma_Z(e^+e^-)$ on top of the one on the reconstructed mass $m(e^+e^-)$ can be seen in the efficiency curve as a function of the true dielectron mass that is shown in Figure 3.19. The significant increase of efficiency in the very-low-dielectron-mass region before the $\sigma_Z(e^+e^-)$ cut is applied (Figure 3.18) is due to the fact that when the invariant mass is very close to threshold, the two electron tracks overlap in the VELO making a single segment with twice the charge, which is easier to reconstruct. This is confirmed by looking at the average charge deposited by one of the electron track for each VELO hit, which doubles when the two electrons overlap perfectly (Figure 3.20(left)). On the other hand, in the intermediate region (5-50 MeV/c^2) the dip in efficiency is attributed to the worsening in the tracking efficiency due to the 2 tracks being quite close but not overlapping anymore. This is demonstrated looking at the χ^2 of the VELO track as a function of the dilepton invariant mass (Fig 3.20(right)).

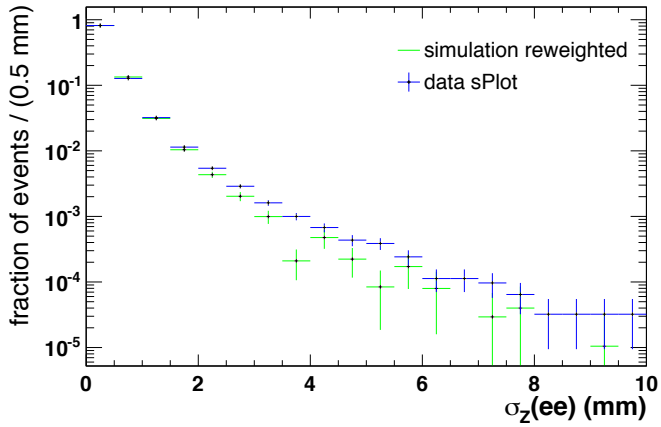


Fig. 3.18: The error on the determination of the z-coordinate of the e^+e^- vertex is compared between $B^0 \to J/\psi(e^+e^-)K^{*0}$ data and MC. The *sPlot* technique is used to extract the signal distribution from data.

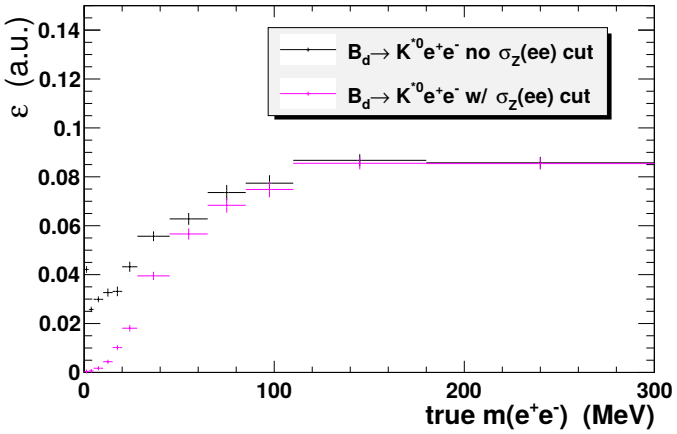


Fig. 3.19: Variation of the efficiency as a function of the true dielectron invariant mass with and without the $\sigma_Z(e^+e^-) < 30$ mm cut for $B^0 \to K^{*0}e^+e^-$ MC.

Once the veto is applied, very few event are left in the MC, corresponding to a $B^0 \to K^{*0}\gamma_{e^+e^-}$ pollution smaller than 1%. The reliability of this estimate is tested in the following. The *sPlot* technique is used with the mass fit described in Section 3.4.4 in order to estimate from data the dilepton invariant mass distribution before the $B^0 \to K^{*0}\gamma$ veto cut. This distribution represents the sum of $B^0 \to K^{*0}\gamma$ and $B^0 \to K^{*0}e^+e^-$ events and it is compared to MC to estimate the normalization to be applied to compute the $B^0 \to K^{*0}\gamma_{e^+e^-}$ pollution. The plot comparing data and MC is shown in Figure 3.21. The $B^0 \to K^{*0}e^+e^-$ MC has been normalised to the data in the region of dilepton mass between 100 MeV/c^2 and 1000 MeV/c^2,

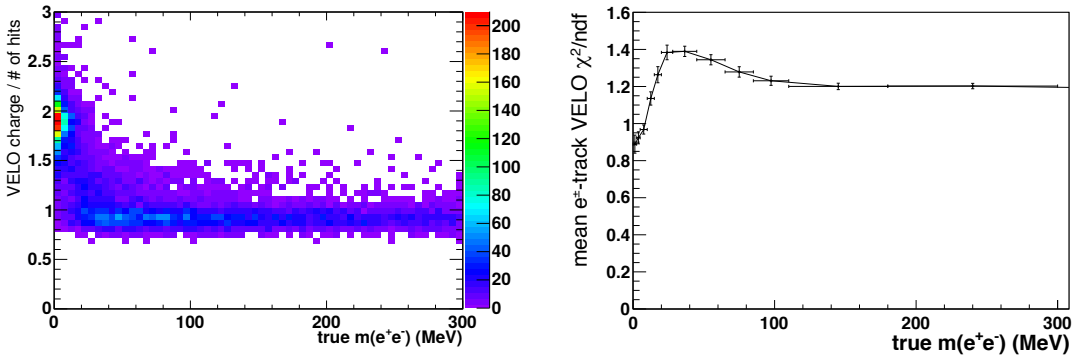


Fig. 3.20: (left) Mean charge deposited by one electron track for each VELO hit as a function of the true dilepton mass for $B^0 \to K^{*0}e^+e^-$ MC. (right) Variation of the mean value of the two electrons VELO χ^2 as a function of the true dielectron invariant mass for $B^0 \to K^{*0}e^+e^-$ MC.

which is free from $B^0 \to K^{*0}\gamma_{e^+e^-}$ background. On the other hand, the $B^0 \to K^{*0}\gamma_{e^+e^-}$ MC is normalised to the data in the region of dilepton mass between 0 MeV/c^2 and 5 MeV/c^2 (after having subtracted a small 3% contribution of $B^0 \to K^{*0}e^+e^-$ events as predicted by the MC). In Figure 3.22 the *sPlot* obtained with the $\sigma_Z(e^+e^-) < 30\,\text{mm}$ cut applied (based on the fit of Section 3.4.4) is compared to the corresponding MC distributions normalised with the factors extracted above. As expected the events with a dilepton invariant mass close to 0 are killed by this cut. Then, using this normalization of MC to data, the number of predicted $B^0 \to K^{*0}\gamma_{e^+e^-}$ events above $20\,\text{MeV}/c^2$ is $(1.9 \pm 1.0)\%$ of the signal yield.

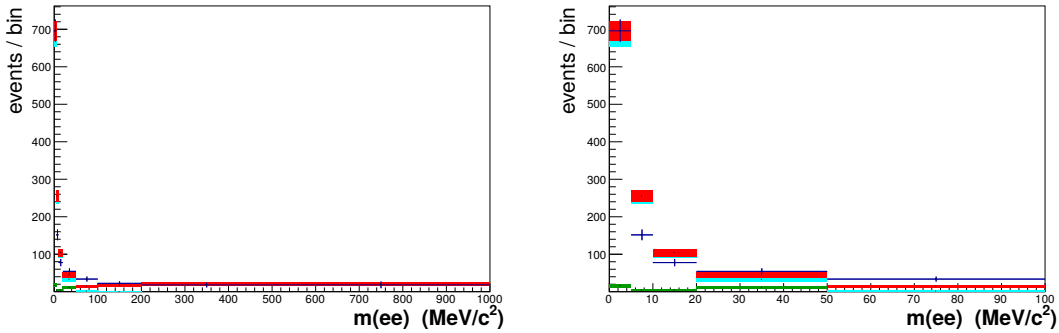


Fig. 3.21: dilepton invariant mass distributions for data (points) with the background removed by the *sPlot* technique and MC (histograms). The cyan distribution is the $B^0 \to K^{*0}\gamma_{e^+e^-}$ MC distribution, the green is the $B^0 \to K^{*0}e^+e^-$ MC distribution and the red is the sum. The plot on the right is a zoom of the overall plot.

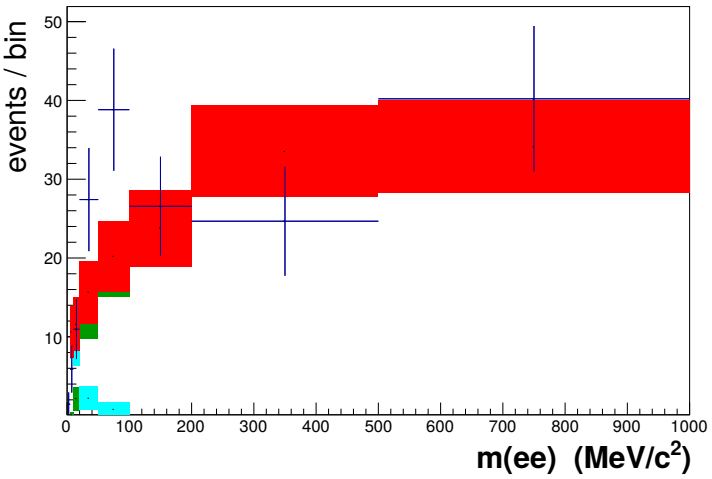


Fig. 3.22: dilepton invariant mass distributions for data (points) with the background removed by the *sPlot* technique and MC (histogram). The cyan distribution is the $B^0 \to K^{*0}\gamma_{e^+e^-}$ MC distribution, the green is the $B^0 \to K^{*0}e^+e^-$ MC distribution and the red is the sum. All selection cuts are applied including the $\sigma_Z(e^+e^-) < 30\,\text{mm}$ cut

However, one cannot fully rely on the MC to estimate the $B^0 \to K^{*0}\gamma_{e^+e^-}$ contamination due to an improper modelling of the dilepton invariant mass in the $B^0 \to K^{*0}\gamma_{e^+e^-}$. The generation of pair production by photons is described in the GEANT4 toolkit [122, 135] used for the generation (in the User Guide, section `PHYS211`). After generating the sharing of the photon energy between the electron and positron energy the polar angle (with respect to the photon direction) of the particle with the lowest energy is generated. However, for computing efficiency reasons the authors decided to model this angular distribution by the sum of two exponentials rather than using the exact formulae. This results in an e^+e^- mass distribution being approximately correct for low masses, but being incorrect at higher masses. Indeed, the mass distribution is expected to decrease approximately as $1/m^3_{e^+e^-}$ and not as an exponential. The exact mass distribution obtained from the Bethe Heitler formula is given in [136] and reproduced here for completeness

as a function of $x = \frac{m_{e^+e^-}}{2m_e}$:

$$\frac{\mathrm{d}P(\gamma \text{ conv})}{\mathrm{d}x} = \frac{1}{x^3}\left[\frac{1}{12}\sqrt{1-\frac{1}{x^2}}\left(28+\frac{17}{x^2}\right) - \frac{1}{6}\left(16-\frac{17}{x^4}+\frac{21}{x^2}\right)\cosh^{-1}x \right.$$
$$\left. + \left(-\sqrt{1-\frac{1}{x^2}}\left(1+\frac{1}{x^2}\right) + \left(2-\frac{1}{x^4}+\frac{2}{x^2}\right)\cosh^{-1}x\right)\ln\left(\frac{89.62x}{\sqrt{1+\left(89.62\frac{x^2}{2y}\right)^2}}\right)\right] \tag{3.4}$$

with $y = \frac{E_\gamma}{2m_e}\frac{1}{c^2}$. The logarithmic term in the equation is linked to the shielding functions of the target material. It is calculated for Silicon and an average E_γ of 40 GeV. This introduces a weak dependence of the mass distributions given in table 3.5. If E_γ is changed to 20 GeV, the fractions are only modified by about $5-10\%$. The resulting probabilities to get a conversion pair above a given value of e^+e^- invariant mass, are given in Table 3.5 for both this formula and the GEANT4 generator. Given the number of events generated at high mass

true $m_{e^+e^-} >$	GEANT4	Bethe Heitler
5 MeV/c^2	13.4%	15.9%
10 MeV/c^2	3.2 %	5.4 %
20 MeV/c^2	0.61%	1.71%
30 MeV/c^2	0.22%	0.82%
40 MeV/c^2	0.11%	0.48%
50 MeV/c^2	0.06%	0.31%

Tab. 3.5: Fraction of events above a given $m_{e^+e^-}$ cut. These numbers do not take into account detector effects such as efficiency and dielectron mass resolution.

in GEANT4 is very small (5 events are left after the $B^0 \to K^{*0}\gamma_{e^+e^-}$ veto cut is applied), it is impractical to correct the GEANT4 distribution by weights in order to reproduce the Bethe Heitler one. However, an estimate of the correction to the $B^0 \to K^{*0}\gamma_{e^+e^-}$ contamination due to the improper modelling of the dilepton masses in GEANT4 can be done using the $B^0 \to K^{*0}e^+e^-$ simulation and the knowledge of the expected $B^0 \to K^{*0}\gamma_{e^+e^-}$ distribution from Equation 3.4. Indeed, one can make the hypothesis that dielectron pairs from $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}\gamma_{e^+e^-}$ undergo approximately the same detector effects in terms of efficiency and mass resolution. Therefore, weights are applied to the $B^0 \to K^{*0}e^+e^-$ MC events to match their generator-level dielectron mass distribution to the two hypothetical $B^0 \to K^{*0}\gamma_{e^+e^-}$ distributions. Then, the expected $B^0 \to K^{*0}\gamma_{e^+e^-}$ having passed the whole selection (included the $B^0 \to K^{*0}\gamma_{e^+e^-}$ veto cut) can be

estimated for both Bethe-Heitler and GEANT4 and the ratio of the two is found to be $f_{\mathrm{BH/Geant}} = 2.0$. In figure 3.23 the true $m(e^+e^-)$ distributions of these two re-weighted samples are compared, showing much more large mass events pass the selection in the Bethe-Heitler case.

In summary, the $B^0 \to K^{*0}\gamma$ MC is normalised to data in the region 0-5 MeV/c^2 before the veto is applied. Then, the predicted contamination of events after the veto, (1.9±0.9)%, is corrected for the improper modelling of the conversion process in GEANT4 by the factor $f_{\mathrm{BH/Geant}} = 2.0$. This fact reduces the already small disagreement between data and MC that is visible in Figure 3.22 in the dilepton mass region 10-100 MeV/c^2. The $B^0 \to K^{*0}\gamma_{e^+e^-}$ contamination is thus estimated to contribute to $(3.8 \pm 1.9)\%$ of the signal yield.

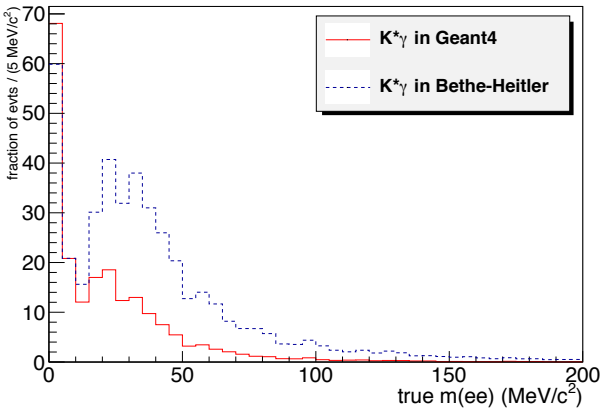


Fig. 3.23: $B^0 \to K^{*0}e^+e^-$ MC events re-weighted according to the true dilepton invariant mass to reproduce the $B^0 \to K^{*0}\gamma_{e^+e^-}$ $m(ee)$ distribution as predicted by the Bethe-Heitler formula 3.4 (dashed blue line) and as implemented in GEANT4 (solid red line). The whole selection included the cut to veto $B^0 \to K^{*0}\gamma_{e^+e^-}$ is applied to both samples.

3.3.3 $B^0 \to K^{*0}\eta$ and $B^0 \to K^{*0}\pi^0$

$B^0 \to K^{*0}\eta$ events can be a source of partially reconstructed background in case of a η decay to two photons followed by a γ conversion or of a η decay to a Dalitz pair ($\eta \to \gamma e^+e^-$). In the case of the conversion electrons, the two electron candidates may come from the same photon or from two different conversions. The first case is similar to $B^0 \to K^{*0}\gamma_{e^+e^-}$ and is thus vetoed by the same cuts. The expected pollution can therefore be calculated using the ratio of branching ratios from the

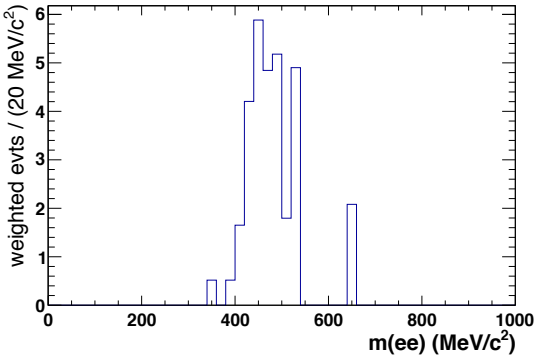


Fig. 3.24: Invariant mass distribution of the dielectron pair from a *double conversion* of the $B^0 \to K^{*0}\eta(\to \gamma\gamma)$ decay from MC.

PDG [130]:

$$\frac{\mathcal{B}(B^0 \to K^{*0}\eta) \times \mathcal{B}(\eta \to \gamma\gamma)}{\mathcal{B}(B^0 \to K^{*0}\gamma)} \simeq \frac{1.6 \times 10^{-5} \ \times \ 3.9 \times 10^{-1}}{4.3 \times 10^{-5}} \simeq 0.15$$

Then this fraction has to be multiplied by a factor 2 since the η decays in two photons. It gives therefore a contamination of 30% of the expected $B^0 \to K^{*0}\gamma_{e^+e^-}$ yield. This is completely negligible since only a small fraction of these events is expected to have a reconstructed $m(K^+\pi^-e^+e^-)$ in the signal region.

A slightly larger number of events comes from cases in which the e^- comes from one of the converted photons and the e^+ from the other (*double conversion*). Indeed, in this case the resulting dielectron mass is close to $m_\eta \simeq 548\,\mathrm{MeV}/c^2$ (see Figure 3.24) and thus passes the $B^0 \to K^{*0}\gamma_{e^+e^-}$ veto. This background is evaluated using a sample of $B^0 \to K^{*0}\eta(\to \gamma\gamma)$ MC. It is expected to contribute to about 2.9 events not peaking under the B mass and is therefore negligible.

To estimate the contribution from Dalitz pairs, one needs to know the fraction of them having a reconstructed mass above $20\,\mathrm{MeV}/c^2$. The theoretical mass distribution of electrons coming from a Dalitz decay of η is thus convoluted with the $m(e^+e^-)$ efficiency extracted from the $B^0 \to K^{*0}e^+e^-$ MC. In practice, this is done through a re-weighting of the $B^0 \to K^{*0}e^+e^-$ MC sample at generator level which allows to reproduce the reconstructed $m(e^+e^-)$ distribution of the electrons from the Dalitz decay of η, see Figure 3.25(left). The whole selection is applied to this reweighted sample, see Figure 3.25(right), and a contamination of 15.8% of the $B^0 \to K^{*0}e^+e^-$ signal yield is found. This is an upper limit, as it neglects the fact that the photon is not reconstructed, which would slightly reduce the efficiency to select Dalitz decay events compared to signal events (for a given dilepton mass). These events have a large flat distribution in $m(K^+\pi^-e^+e^-)$ with a fraction of just 0.18 of them being above $4.8\,\mathrm{GeV}/c^2$ and with no event above

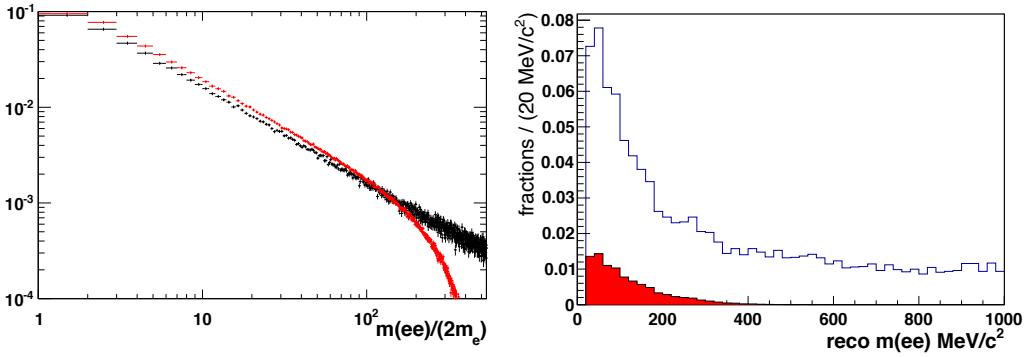


Fig. 3.25: (left) The generated dilepton-mass distribution from the $B^0 \to K^{*0}e^+e^-$ MC, in black, is reweighted to reproduce the expected distribution for the $B^0 \to K^{*0}\eta$ followed by a Dalitz decay of η, in red. The corresponding reconstructed dilepton-mass distribution (right) for the $B^0 \to K^{*0}e^+e^-$ decay, in blue, and for the $B^0 \to K^{*0}\eta(\to e^+e^-\gamma$ decay, in red.

5.25 GeV/c^2. This amounts to a contribution of 4.2 events above 4.8 GeV/c^2, not peaking in the B mass, thus negligible. Furthermore, this background being flat, it is mostly included in the partially reconstructed background fraction fitted to data (see Section 3.4.4).

Backgrounds from the decay $B^0 \to K^{*0}\pi^0$ are of the same type as the η, and therefore the same approach is applied to estimate their contribution. The probability of making a π^0 that then goes to a couple of photons is about two times smaller than for η:

$$\frac{\mathcal{B}(B^0 \to K^{*0}\pi^0) \times \mathcal{B}(\pi^0 \to \gamma\gamma)}{\mathcal{B}(B^0 \to K^{*0}\gamma)} \simeq \frac{3.3 \times 10^{-6} \times 0.99}{4.3 \times 10^{-5}} \simeq 0.08 \qquad (3.5)$$

Therefore, it gives a negligible contamination and so does the double conversion case.

The contamination coming from Dalitz decays of the π^0 is evaluated in the same way as for the η, see Figure 3.26, but taking into account the possibility of a merging of the photon energy-deposit in the ECAL to the deposit of one of the electron. This is likely to happen for π^0 Dalitz decays due to the small m_{π^0} and thus the small angle of aperture between its decay products. The effect of this merging on the expected m(e^+e^- K^{*0}) distribution has been evaluated from a MC sample of $B^0 \to K^{*0}\pi^0(\to \gamma\gamma)$ where one of the photons converted in the material. Indeed, looking at the $m(e^+e^-K^{*0})$ distribution for these events, Figure 3.27, one can see a small peak at the B^0 mass, while one would expect just a flat distribution below the B^0 mass if there was no merging of the photon with one of the electrons. Nonetheless the background remains negligible, yielding a

contamination of 1.0 expected events above a $m(e^+e^-K^{*0})$ of 4.8 GeV/c^2 (and 0.6 above 5.1 GeV/c^2).

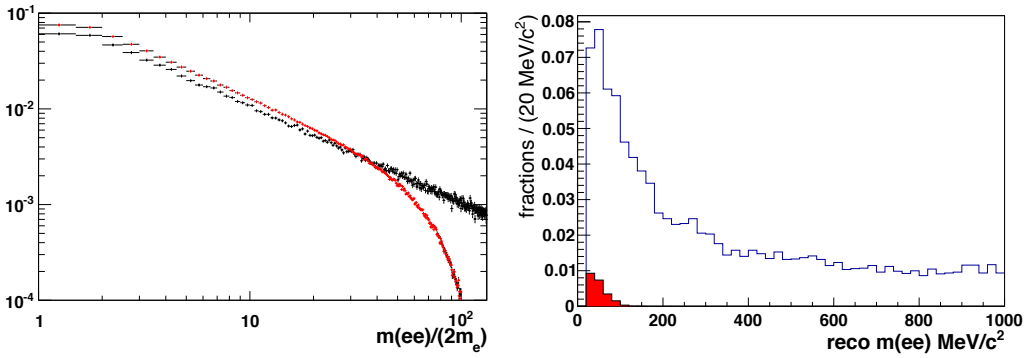


Fig. 3.26: (left) The generated dilepton-mass distribution from the $B^0 \to K^{*0}e^+e^-$ MC, in black, is reweighted to reproduce the expected distribution for the $B^0 \to K^{*0}\pi^0$ followed by a Dalitz decay of π^0, in red. The corresponding reconstructed dilepton-mass distribution (right) for the $B^0 \to K^{*0}e^+e^-$ decay, in blue, and for the $B^0 \to K^{*0}\pi^0(\to e^+e^-\gamma$ decay, in red.

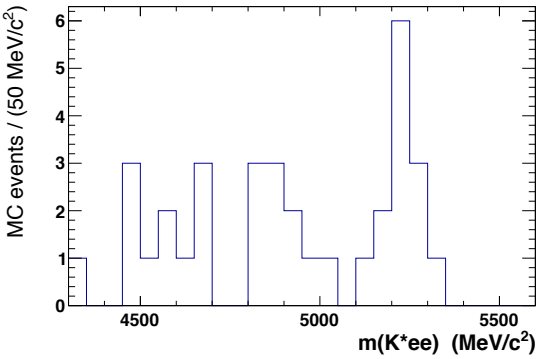


Fig. 3.27: Invariant mass distribution $m(e^+e^-K^{*0})$ from a MC sample of $B^0 \to K^{*0}\pi^0(\to \gamma\gamma)$, where one of the photons has converted to an e^+e^- pair.

3.3.4 The $B^0 \to K^{*0}V(\to e^+e^-)$ decays with $V = \rho, \omega, \phi$

To calculate the contribution of decays of the type $B^0 \to K^{*0}V(\to e^+e^-)$ with $V = \rho, \omega, \phi$, numbers are taken from the PDG [130] for the branching ratios of $B^0 \to K^{*0}V$ and $V \to e^+e^-$ processes (the latter are taken from storage rings data). All the branching ratios are reported in Table 3.6.

V	m_V (MeV/c^2)	$\mathcal{B}(B^0 \to K^{*0}\mathrm{V})$	$\mathcal{B}(V \to e^+e^-)$	$\mathcal{B}(B^0 \to K^{*0}V(\to e^+e^-))$
ρ	775.49 ± 0.34	$(3.9 \pm 1.3)10^{-6}$	$(4.71 \pm 0.05)10^{-5}$	$(1.8 \pm 0.6)10^{-10}$
ω	782.65 ± 0.12	$(2.0 \pm 0.5)10^{-6}$	$(7.16 \pm 0.12)10^{-5}$	$(1.4 \pm 0.6)10^{-10}$
ϕ	1019.455 ± 0.020	$(9.8 \pm 0.8)10^{-6}$	$(2.97 \pm 0.04)10^{-4}$	$(3.1 \pm 0.2)10^{-9}$

Tab. 3.6: Measured parameters for the three vector mesons considered [130]

These have to be compared to a $\mathcal{B}(B^0 \to K^{*0}e^+e^-)$ between 30 MeV/c^2 and 1000 MeV/c^2 of $(3.1^{+0.9}_{-0.8}\ ^{+0.2}_{-0.3} \pm 0.2) \times 10^{-7}$ [137][4]. This simple calculation already shows that the contribution from ρ and ω is negligible, while the one from ϕ is more sizeable. The argument holds even in the presence of interference effects with the most favourable phase between the background and the signal contributions [33, 138]. In order to evaluate the contribution of the ϕ resonance, the $B^0 \to K^{*0}\phi(\to \mu\mu)$ contribution as found in the $B^0 \to K^{*0}\mu\mu$ LHCb analysis with $3\,\mathrm{fb}^{-1}$ [139] is taken. The mass resolution being much better, the ϕ contribution gives a very clear and narrow peak in muons whose yield was extracted to be of 18 events. This number is normalised through the ratio of expected yields of $B^0 \to K^{*0}e^+e^-$ and $B^0 \to \mu^+\mu^-K^{*0}$ between $\sqrt{0.1}\,\mathrm{GeV}/c^2$ and $1.0\,\mathrm{GeV}/c^2$, giving a contribution of 2.6 ± 0.6 $B^0 \to K^{*0}\phi(\to ee)$ events. Note that in this q^2 range selection efficiencies of both analyses are quite flat. The ϕ naturally peaks very narrowly at $1020\,\mathrm{MeV}/c^2$, but the broad resolution effects in $B^0 \to K^{*0}e^+e^-$ make the 67% of ϕ events to get a reconstructed $m(e^+e^-)$ value below $1\,\mathrm{GeV}/c^2$ and therefore be selected. In conclusion this gives an expected contribution of 1.8 ± 0.4 $B^0 \to K^{*0}\phi(\to ee)$ events, thus corresponding to a negligible $(1.2 \pm 0.3)\%$ of the signal yield.

3.3.5 The $\Lambda_b^0 \to \Lambda^*(pK)e^+e^-$ decay

Decays of Λ_b^0 to pKe^+e^- can be a background to $B^0 \to K^{*0}e^+e^-$ either if the p is mis-identified as a π or if a double mis-identification $(p \to K)$ and $(K \to \pi)$ occurs. In order to investigate the contamination of this background, a MC sample of $\Lambda_b^0 \to$

[4]The branching ratio between 20 and 1000 MeV/c^2 will be even larger

$\Lambda^*(1520)^0\gamma$ with $\Lambda^*(1520)^0 \to pK$ is selected while releasing the veto for the $B^0 \to K^{*0}\gamma$ decay (see Section 3.3.2). The selection efficiency obtained is then compared to the one obtained for $B^0 \to K^{*0}\gamma$. This efficiency ratio is supposed to be the same among the decays $\Lambda_b^0 \to \Lambda^*(1520)^0 e^+e^-$ and $B^0 \to K^{*0}e^+e^-$, this assumption being motivated by the very low q^2 range considered in this analysis. The PID response, corrected for MC/data differences, gives a probability of $(p \to \pi)$ mis-identification of 30.0%. On the other hand, the one for double mis-identification $(p \to K)$ and $(K \to \pi)$ is 5.8%. Therefore, the contamination due to the $(p \to \pi)$ mis-identification is computed as:

$$\frac{\epsilon\left(\Lambda_b^0 \to \gamma\ \Lambda^*(1520)^0_{(p\to\pi)\text{mis-id}}\right)}{\epsilon(B^0 \to K^{*0}\gamma)} \times \frac{\mathcal{B}\left(\Lambda_b^0 \to \gamma\ \Lambda^*(1520)^0(\to pK)\right)}{\mathcal{B}(B^0 \to K^{*0}\gamma) \times \mathcal{B}(K^{*0} \to K\pi)} \times \frac{f_{\Lambda_b^0}}{f_d} = 1.77\% \quad (3.6)$$

This already small fraction of expected background events is further reduced by an explicit cut for the $(p \to \pi)$ mis-identification probability as `ProbNNpi*(1-ProbNNp)` > 0.2. This cut lowers the contamination by a factor of 2 making this background negligible, while having a 99.5% efficiency on the signal. The double mis-identification $(p \to K)$ and $(K \to \pi)$ is even lower and thus neglected.

3.3.6 K^{*0}-$\overline{K}^{*0}$ mis-identification and double-counted events

The double mis-identification of the K as a π and the π as a K can be a sizeable background for the angular analysis of $B^0 \to K^{*0}e^+e^-$. It leads to a $\cos\theta_\ell$ with opposite sign which can bias the measurement of A_T^{Re} to lower values. The measurement of $\cos\theta_K$ is also affected and could bias F_L. The measurement of A_T^{Im}, being a odd-*CP* observable, is also affected, while the $A_T^{(2)}$ observable is completely independent of the amount of K^{*0} mis-identified as $\overline{K}^{*0}$ ("swaps").
The MC sample was used to evaluate the number of "swaps": a $B^0 \to K^{*0}e^+e^-$ sample is selected removing all PID cuts (including stripping level ones). Then, the PID efficiency of these cuts is evaluated using with the same method used in Section 3.2.3. MC events are re-sampled according to these PID probabilities. The sample of candidates left has a $(2.6 \pm 0.1)\%$ of swaps, but a large part of them corresponds to events double counted both as K^{*0} and $\overline{K}^{*0}$ candidates. Indeed, the number of swaps reduces to $(1.6 \pm 0.1)\%$ after a removal of double counted events with a random choice of one of the two candidates. Furthermore, the choice between the two candidates can be improved: for each couple of double counted events having the K and π candidate tracks swapped, the one with the higher probability of being correctly identified is chosen, *i.e.* the one with the larger value of K`ProbNNk*`π`ProbNNpi`. This method lowers the contamination of "swaps" in

the MC down to:

$$\frac{(K^{*0} - \overline{K}^{*0}) \text{ mis-id candidates}}{\text{all signal candidates}} = (1.1 \pm 0.1)\%. \tag{3.7}$$

The same percentage is expected on data, making the contamination negligible in terms of the associated systematic error on the angular observables.

3.3.7 The $B_s^0 \to \phi e^+ e^-$ decay

The $B_s^0 \to \phi e^+ e^-$ decay with the ϕ going to K^+K^- was also considered and vetoed. It has a branching ratio of the same order of magnitude as the $B^0 \to K^{*0} e^+ e^-$ decay, very similar kinematics and if one of the K is mis-identified as a π the ϕ may be reconstructed as a K^{*0}. Therefore, a specific cut is applied on the invariant mass of the $m(K^+K^-)$ by assigning to the π candidate the K mass (see Figure 3.28). A cut on this invariant mass is set right above the mass of the ϕ at $m(K^+K^-) > 1040$ MeV/c^2. The efficiency of this cut on signal is 99.4%.

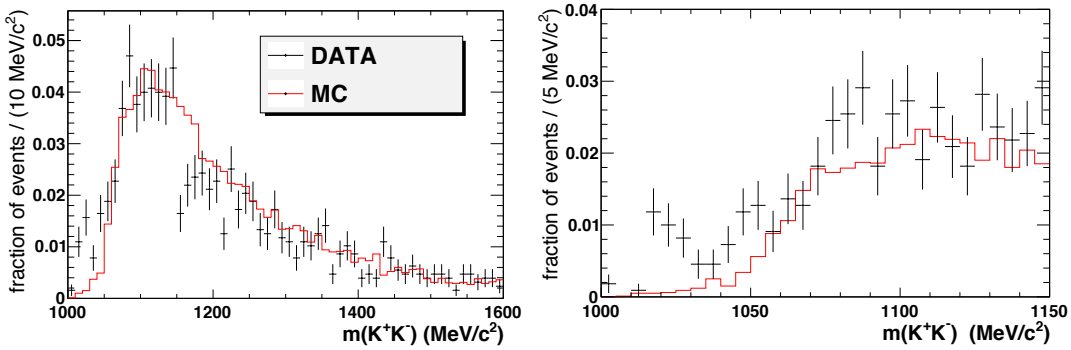


Fig. 3.28: Distribution of the invariant mass $m(K^+K^-)$ calculated under the assumption that the reconstructed π is mis-identified as a K. On the right is just a zoom of the left plot. The histogram shows the distribution from the $B^0 \to K^{*0} e^+ e^-$ MC sample to which all the selection cuts are applied but the $B_s^0 \to \phi e^+ e^-$ veto cut. In black is the same distribution from LHCb data, which has a different shape because it contains much more background. Nevertheless in data a clear peak can be seen around the ϕ invariant mass $m(\phi) = 1019\,\text{MeV}/c^2$, where the signal yield is instead very low.

3.3.8 The $B^+ \to K^+ e^+ e^-$ decay

A very small contribution from the $B^+ \to J/\psi\, K^+$ decay can be found in the signal proxy $B^0 \to J/\psi\, K^{*0}$, while the corresponding non resonant decay $B^+ \to K^+ e^+ e^-$ is highly suppressed compared to the $B^0 \to K^{*0} e^+ e^-$ channel for the dilepton masses

considered $(20 - 1000\,\mathrm{MeV}/c^2)$. This is due to the fact that the $B^+ \to K^+e^+e^-$ q^2 distribution does not have the photon pole, being thus ~ 6 times smaller in the $m(e^+e^-)$ range considered. It corresponds to events where one additional π is picked up from the combinatorial background to form a K^{*0} thus giving a reconstructed invariant mass $m(K\pi e^+e^-)$ around the B^0 mass. This background peaks in the reconstructed mass m(K e^+e^-) where the e^+e^- pair invariant mass is constrained to the J/ψ mass (see Figure 3.29) and is therefore removed with a cut at $m(Ke^+e^-) < 5200\,\mathrm{MeV}/c^2$. The relative fraction of this background in the $B^0 \to J/\psi(e^+e^-)K^{*0}$ channel is estimated to be 2%, thus quite low. It is nevertheless removed in order to avoid biases in the signal shape modelling when comparing data and MC (see Section 3.4.2).

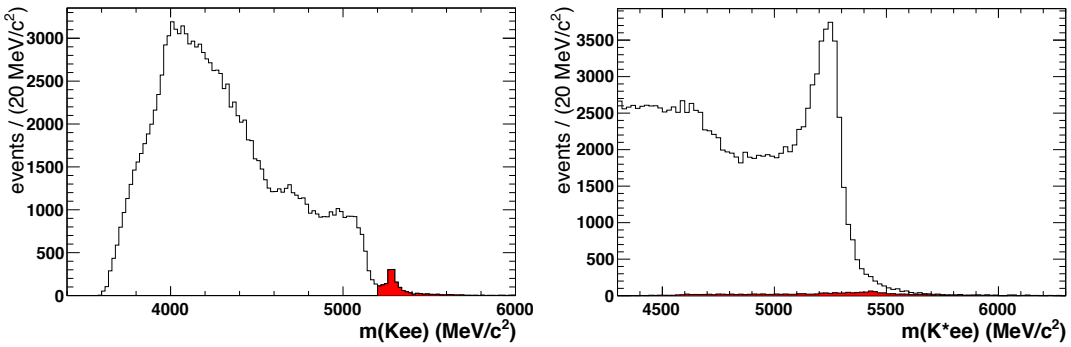


Fig. 3.29: On the left is the distribution of $m(Ke^+e^-)$ in data with the e^+e^- invariant mass constrained to the J/ψ mass. On the right the same data are used to plot the $m(K^{*0}e^+e^-)$ invariant mass (without the constrain on $m(e^+e^-)$). The events vetoed by the $m(Ke^+e^-) > 5200$ MeV/c^2 cut are shown in red in both distributions.

3.3.9 Summary of offline selection cuts

The overall set of selection cuts are summarised in Table 3.7. PID and BDT cuts are in Table 3.8. The same selection is applied to MC and to data. A summary of the estimated specific background yields is given in Table 3.9.

	$B^0 \to K^{*0} e^+ e^-$	$B^0 \to J/\psi\,(e^+e^-) K^{*0}$
$m(e^+e^-)$ window	$20 < m(e^+e^-) < 1000\,\text{MeV}/c^2$	$2200 < m(e^+e^-) < 3400\,\text{MeV}/c^2$
$m(K\pi)$ window	$\lvert m(K\pi) - 892\,\text{MeV}/c^2 \rvert < 100\,\text{MeV}/c^2$	
K^{*0} γ veto	$\sigma_Z(e^+e^-) < 30$ mm	
$B^0 \to D^- e^+ \nu$ veto	$\lvert \cos\theta_\ell \rvert < 0.8$	
$B_s^0 \to \phi e^+ e^-$ veto	$m(K^+K^-) > 1040$ MeV/c^2	
$\Lambda_b^0 \to e^+e^-$ $\Lambda^*(1520)^0$	πProbNNpi $*\,(1 - \pi$ProbNNp$) > 0.2$	
$B^+ \to K^+ e^+ e^-$ veto	no cut	$m^{J/\psi}(Ke^+e^-) < 5200$ MeV/c^2

Tab. 3.7: Summary of all selection cuts. They are the same among the three trigger categories.

	L0Ele	L0Had	L0TIS
BDT	> 0.88	> 0.94	> 0.96
$e^\pm$ min PIDe	> 1.2	> 1.2	> 1.6
K PID	ProbNNk > 0.05		
π PID	ProbNNpi > 0.2		

Tab. 3.8: Summary of all PID and BDT cuts. They are the same for the $B^0 \to K^{*0} e^+ e^-$ channel and for the control channels $B^0 \to J/\psi\,(e^+e^-) K^{*0}$ and $B^0 \to K^{*0}\gamma$

$B^0 \to K^{*0} \gamma_{e^+e^-}$	$(3.8 \pm 1.9)\%$	peaking
$B^0 \to K^{*0} \eta(\gamma\gamma)$	$\sim$3.2%	not peaking
$B^0 \to K^{*0} \eta(\gamma e^+ e^-)$	$\sim$3.4%	not peaking
$B^0 \to K^{*0} \pi^0(\gamma\gamma)$	$\sim$1.5%	not peaking
$B^0 \to K^{*0} \pi^0(\gamma e^+ e^-)$	$< 0.7\%$	peaking
$B^0 \to K^{*0} \phi(e^+e^-)$	$\sim$1.2%	peaking
$\Lambda_b \to \Lambda^* e^+ e^-$	$\sim$0.9%	not peaking
$B^+ \to K^+ e^+ e^-$	$\sim$0.3%	not peaking

Tab. 3.9: Summary of all specific background expected yields with respect to the signal. The partially reconstructed background involving higher K^{*0} resonances and the combinatorial background are not listed here as they are fitted to data.

3.4 Invariant mass fits

Before proceeding to the angular analysis of the $B^0 \to K^{*0}e^+e^-$ decay, a fit of the $m(K^+\pi^-e^+e^-)$ invariant mass distribution is set up in order to validate the understanding of the selected data sample and extract the signal yield. The $m(K^+\pi^-e^+e^-)$ invariant mass is then fitted together with the distributions of the three angles to measure the angular observables as described in Chapter 4.
The signal mass shape is modelled on MC (Section 3.4.1) and is then corrected for small MC/data differences in the mass shape which are extracted from the $B^0 \to J/\psi(e^+e^-)K^{*0}$ channel (Section 3.4.2). Then, Section 3.4.3 details the modelling of the $m(K^+\pi^-e^+e^-)$ distribution of partially reconstructed backgrounds arising from $B \to K^{*0}e^+e^-X$ decays, where one or more of the decay products (X) from the B is not reconstructed. The mass fits to the rare $B^0 \to K^{*0}e^+e^-$ channel and to the proxy $B^0 \to K^{*0}\gamma_{e^+e^-}$ are presented together in Section 3.4.4. Indeed, the $B^0 \to K^{*0}\gamma_{e^+e^-}$ proxy is used to validate the mass fit to the rare $B^0 \to K^{*0}e^+e^-$ decay and to extract the fraction of partially reconstructed background.

3.4.1 Invariant mass fit on $B^0 \to K^{*0}e^+e^-$ signal Monte-Carlo

The $B^0 \to K^{*0}e^+e^-$ signal $m(K^+\pi^-e^+e^-)$ invariant mass distribution is modelled on the MC. The signal shape is found to be different for the three trigger categories, thus they are modelled separately. Also, following [140], the Probability Distribution Function depends on the number of neutral clusters that are added to the dielectron candidate to correct for the effect of bremsstrahlung. Three " bremsstrahlung" categories are defined: 0γ, 1γ and 2γ corresponding to, respectively, 0, 1 and 2 or more neutral clusters added. The signal PDF is described by the sum of a Crystal Ball function [131] (CB) and a wide Gaussian function accounting for the cases where background neutral clusters were added. The centre of the Gaussian is fixed to the nominal B^0 mass value. No Gaussian is added for the 0γ category while for the other two categories, the Gaussian accounts to less than the 10% of the total signal PDF. The CB function is based on an empirical model developed for the radiative tail from $b\bar{b}$ resonances [131]. It consists of a Gaussian distribution with an exponential tail and has the following expression as a function of the invariant mass m:

$$\mathcal{CB}(m;\ M,\sigma,\alpha,n) = N \begin{cases} \exp\left(-\frac{(m-M)^2}{2\sigma^2}\right) & \text{for } \frac{m-M}{2\sigma} > -\alpha \\ \frac{\left(\frac{n}{\alpha}\right)^n \exp\left(-\frac{1}{2}\alpha^2\right)}{\left(\frac{m-M}{\sigma}+\frac{n}{\alpha}-\alpha\right)^n} & \text{for } \frac{m-M}{2\sigma} \leq -\alpha \end{cases} \tag{3.8}$$

where N is the signal normalisation, M is the position of the maximum, σ is the Gaussian width and the parameters n and α model the exponential tail. The n parameter is highly correlated with α, thus often one of the two is fixed in the fit. The signal PDF fitted to the MC mass distribution reads:

$$\mathcal{S}_{ij}^{\mathrm{MC}} = f_{ij}^{\mathrm{Core}}\ \mathcal{CB}(M_{ij}^{\mathrm{CB}}, \sigma_{ij}^{\mathrm{CB}}, \alpha_{ij}, n_{ij}) + (1. - f_{ij}^{\mathrm{Core}})\ \mathcal{G}(M_{ij}^{\mathrm{G}}, \sigma_{ij}^{\mathrm{G}}) \tag{3.9}$$

where the subscripts i =L0Ele, L0Had, L0TIS and j =0γ, 1γ, 2γ denote the trigger and bremsstrahlung cateogories[5], $\mathcal{G}(M^{\mathrm{G}}, \sigma^{\mathrm{G}})$ is the Gaussian function and f^{Core} is the fraction of signal PDF being described by the CB function. No parameter is shared between the 3×3 categories. The results of the fit of the MC sample are shown in Figure 3.30 and the values of the parameters are given in Table 3.10.

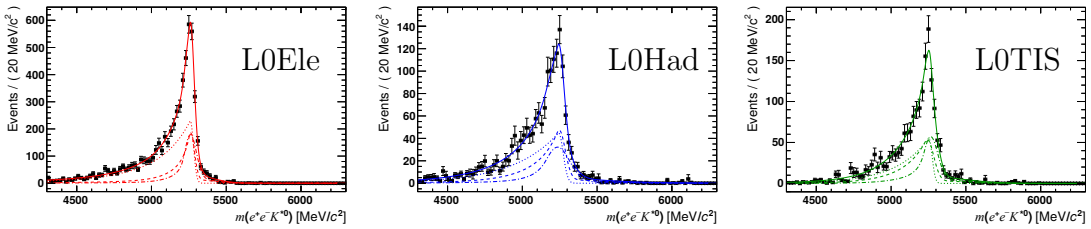


Fig. 3.30: Mass fit of the $B^0 \to K^{*0}e^+e^-$ decay from LHCb MC. The three trigger categories are shown separately on the three plots. The full line is the overall signal PDF, while the dotted, dashed and dashed-dotted lines are the signal PDFs for, respectively, the 0γ, 1γ and 2γ bremsstrahlung categories.

3.4.2 MC/data differences from $B^0 \to J/\psi\,(e^+e^-)K^{*0}$

A fit of the invariant mass shapes to the $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ MC is compared to a data sample of $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ with high signal purity in order to extract MC/data differences in the mass shapes. This information is then used to correct the $B^0 \to K^{*0}e^+e^-$ mass shape taken from MC (Section 3.4.1).

Invariant mass fit to $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ signal MC

Firstly, the signal shapes are extracted from $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ signal Monte-Carlo. This is done, as for the $B^0 \to K^{*0}e^+e^-$ channel (Section 3.4.1), by fitting separately the three trigger categories as well as the three bremsstrahlung categories. Also, the same functional shapes are used, a CB function plus a wide

[5]when assigned to a function, these subscripts indicate the parameters of the function are different for the various categories, not its functional form; the only exception is the signal PDF of the 0γ bremsstrahlung category which does not include the wide Gaussian component.

Parameter	Fitted value	Parameter	Fitted value	Parameter	Fitted value
$\alpha_{\text{L0Ele;0}\gamma}$	0.1	$\alpha_{\text{L0Ele;1}\gamma}$	0.4 ± 0.2	$\alpha_{\text{L0Ele;2}\gamma}$	1.2 ± 0.7
$\alpha_{\text{L0Had;0}\gamma}$	0.1	$\alpha_{\text{L0Had;1}\gamma}$	0.3	$\alpha_{\text{L0Had;2}\gamma}$	1
$\alpha_{\text{L0TIS;0}\gamma}$	0.1	$\alpha_{\text{L0TIS;1}\gamma}$	0.3	$\alpha_{\text{L0TIS;2}\gamma}$	1
		$f^{\text{Core}}_{\text{L0Ele;1}\gamma}$	0.85 ± 0.03	$f^{\text{Core}}_{\text{L0Ele;2}\gamma}$	0.79 ± 0.06
		$f^{\text{Core}}_{\text{L0Had;1}\gamma}$	0.84 ± 0.05	$f^{\text{Core}}_{\text{L0Had;2}\gamma}$	0.85 ± 0.07
		$f^{\text{Core}}_{\text{L0TIS;1}\gamma}$	0.86 ± 0.03	$f^{\text{Core}}_{\text{L0TIS;2}\gamma}$	0.76 ± 0.06
$M^{\text{CB}}_{\text{L0Ele;0}\gamma}$	$5260 \pm 5\,\text{MeV}/c^2$	$M^{\text{CB}}_{\text{L0Ele;1}\gamma}$	$5262 \pm 7\,\text{MeV}/c^2$	$M^{\text{CB}}_{\text{L0Ele;2}\gamma}$	$5261 \pm 8\,\text{MeV}/c^2$
$M^{\text{CB}}_{\text{L0Had;0}\gamma}$	$5243 \pm 4\,\text{MeV}/c^2$	$M^{\text{CB}}_{\text{L0Had;1}\gamma}$	$5247 \pm 8\,\text{MeV}/c^2$	$M^{\text{CB}}_{\text{L0Had;2}\gamma}$	$5236 \pm 7\,\text{MeV}/c^2$
$M^{\text{CB}}_{\text{L0TIS;0}\gamma}$	$5252 \pm 3\,\text{MeV}/c^2$	$M^{\text{CB}}_{\text{L0TIS;1}\gamma}$	$5242 \pm 6\,\text{MeV}/c^2$	$M^{\text{CB}}_{\text{L0TIS;2}\gamma}$	$5265 \pm 5\,\text{MeV}/c^2$
		$M^{\text{G}}_{\text{L0Ele;1}\gamma}$	$5279\,\text{MeV}/c^2$	$M^{\text{G}}_{\text{L0Ele;2}\gamma}$	$5279\,\text{MeV}/c^2$
		$M^{\text{G}}_{\text{L0Had;1}\gamma}$	$5279\,\text{MeV}/c^2$	$M^{\text{G}}_{\text{L0Had;2}\gamma}$	$5279\,\text{MeV}/c^2$
		$M^{\text{G}}_{\text{L0TIS;1}\gamma}$	$5279\,\text{MeV}/c^2$	$M^{\text{G}}_{\text{L0TIS;2}\gamma}$	$5279\,\text{MeV}/c^2$
$n_{\text{L0Ele;0}\gamma}$	5 ± 3	$n_{\text{L0Ele;1}\gamma}$	1.9 ± 0.4	$n_{\text{L0Ele;2}\gamma}$	1.5 ± 1
$n_{\text{L0Had;0}\gamma}$	100	$n_{\text{L0Had;1}\gamma}$	4 ± 3	$n_{\text{L0Had;2}\gamma}$	6
$n_{\text{L0TIS;0}\gamma}$	100	$n_{\text{L0TIS;1}\gamma}$	10 ± 4	$n_{\text{L0TIS;2}\gamma}$	2.6 ± 0.6
$\sigma^{\text{CB}}_{\text{L0Ele;0}\gamma}$	$21 \pm 3\,\text{MeV}/c^2$	$\sigma^{\text{CB}}_{\text{L0Ele;1}\gamma}$	$25 \pm 5\,\text{MeV}/c^2$	$\sigma^{\text{CB}}_{\text{L0Ele;2}\gamma}$	$39 \pm 10\,\text{MeV}/c^2$
$\sigma^{\text{CB}}_{\text{L0Had;0}\gamma}$	$30 \pm 2\,\text{MeV}/c^2$	$\sigma^{\text{CB}}_{\text{L0Had;1}\gamma}$	$40 \pm 13\,\text{MeV}/c^2$	$\sigma^{\text{CB}}_{\text{L0Had;2}\gamma}$	$80 \pm 9\,\text{MeV}/c^2$
$\sigma^{\text{CB}}_{\text{L0TIS;0}\gamma}$	$22 \pm 1\,\text{MeV}/c^2$	$\sigma^{\text{CB}}_{\text{L0TIS;1}\gamma}$	$42 \pm 5\,\text{MeV}/c^2$	$\sigma^{\text{CB}}_{\text{L0TIS;2}\gamma}$	$47 \pm 5\,\text{MeV}/c^2$
		$\sigma^{\text{G}}_{\text{L0Ele;1}\gamma}$	$134 \pm 17\,\text{MeV}/c^2$	$\sigma^{\text{G}}_{\text{L0Ele;2}\gamma}$	$125 \pm 25\,\text{MeV}/c^2$
		$\sigma^{\text{G}}_{\text{L0Had;1}\gamma}$	$253 \pm 33\,\text{MeV}/c^2$	$\sigma^{\text{G}}_{\text{L0Had;2}\gamma}$	$506 \pm 227\,\text{MeV}/c^2$
		$\sigma^{\text{G}}_{\text{L0TIS;1}\gamma}$	$247 \pm 41\,\text{MeV}/c^2$	$\sigma^{\text{G}}_{\text{L0TIS;2}\gamma}$	$208 \pm 39\,\text{MeV}/c^2$

Tab. 3.10: Fitted parameters as obtained from the invariant mass fit to $B^0 \to K^{*0}e^+e^-$ MC. When the error on the parameter is not given it means the parameter was fixed in the mass fit.

Gaussian. The results of the fit are shown in Figure 3.31 and the values of the parameters are given in Table 3.11.

Invariant mass fit to $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ data

A special selection is applied to $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ data to remove the partially reconstructed background and obtain a highly pure signal sample that can be easily compared to MC. The momenta of the electron candidates are rescaled so that the dielectron invariant mass matches the nominal J/ψ mass. These momenta are then used to compute the $m(K^+\pi^-e^+e^-)$ mass, allowing to eliminate the long

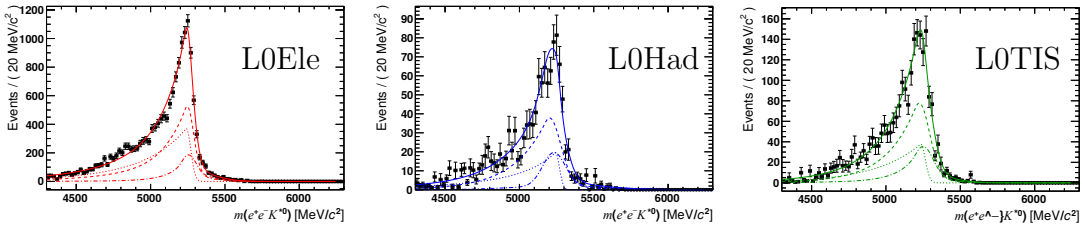


Fig. 3.31: Mass fit of the $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ decay from LHCb MC. The three trigger categories are shown separately on the three plots. The full line is the overall signal PDF, while the dotted, dashed and dashed-dotted lines are the signal PDFs for, respectively, the 0γ, 1γ and 2γ bremsstrahlung categories.

Parameter	Fitted value	Parameter	Fitted value	Parameter	Fitted value
$\alpha_{\mathrm{L0Ele};0\gamma}$	0.1	$\alpha_{\mathrm{L0Ele};1\gamma}$	0.246 ± 0.006	$\alpha_{\mathrm{L0Ele};2\gamma}$	0.82 ± 0.04
$\alpha_{\mathrm{L0Had};0\gamma}$	0.1	$\alpha_{\mathrm{L0Had};1\gamma}$	0.3	$\alpha_{\mathrm{L0Had};2\gamma}$	8 ± 4
$\alpha_{\mathrm{L0TIS};0\gamma}$	0.1	$\alpha_{\mathrm{L0TIS};1\gamma}$	0.3	$\alpha_{\mathrm{L0TIS};2\gamma}$	0.67 ± 0.06
		$f^{\mathrm{Core}}_{\mathrm{L0Ele};1\gamma}$	0.889 ± 0.009	$f^{\mathrm{Core}}_{\mathrm{L0Ele};2\gamma}$	0.76 ± 0.03
		$f^{\mathrm{Core}}_{\mathrm{L0Had};1\gamma}$	0.93 ± 0.02	$f^{\mathrm{Core}}_{\mathrm{L0Had};2\gamma}$	0.68 ± 0.04
		$f^{\mathrm{Core}}_{\mathrm{L0TIS};1\gamma}$	0.94 ± 0.02	$f^{\mathrm{Core}}_{\mathrm{L0TIS};2\gamma}$	0.85 ± 0.04
$M^{\mathrm{CB}}_{\mathrm{L0Ele};0\gamma}$	$5240 \pm 1\,\mathrm{MeV}/c^2$	$M^{\mathrm{CB}}_{\mathrm{L0Ele};1\gamma}$	$5245 \pm 2\,\mathrm{MeV}/c^2$	$M^{\mathrm{CB}}_{\mathrm{L0Ele};2\gamma}$	$5254 \pm 3\,\mathrm{MeV}/c^2$
$M^{\mathrm{CB}}_{\mathrm{L0Had};0\gamma}$	$5242 \pm 4\,\mathrm{MeV}/c^2$	$M^{\mathrm{CB}}_{\mathrm{L0Had};1\gamma}$	$5202 \pm 6\,\mathrm{MeV}/c^2$	$M^{\mathrm{CB}}_{\mathrm{L0Had};2\gamma}$	$5232 \pm 7\,\mathrm{MeV}/c^2$
$M^{\mathrm{CB}}_{\mathrm{L0TIS};0\gamma}$	$5234 \pm 4\,\mathrm{MeV}/c^2$	$M^{\mathrm{CB}}_{\mathrm{L0TIS};1\gamma}$	$5224 \pm 4\,\mathrm{MeV}/c^2$	$M^{\mathrm{CB}}_{\mathrm{L0TIS};2\gamma}$	$5250 \pm 6\,\mathrm{MeV}/c^2$
		$M^{\mathrm{G}}_{\mathrm{L0Ele};1\gamma}$	$5279\,\mathrm{MeV}/c^2$	$M^{\mathrm{G}}_{\mathrm{L0Ele};2\gamma}$	$5279\,\mathrm{MeV}/c^2$
		$M^{\mathrm{G}}_{\mathrm{L0Had};1\gamma}$	$5279\,\mathrm{MeV}/c^2$	$M^{\mathrm{G}}_{\mathrm{L0Had};2\gamma}$	$5279\,\mathrm{MeV}/c^2$
		$M^{\mathrm{G}}_{\mathrm{L0TIS};1\gamma}$	$5279\,\mathrm{MeV}/c^2$	$M^{\mathrm{G}}_{\mathrm{L0TIS};2\gamma}$	$5279\,\mathrm{MeV}/c^2$
$n_{\mathrm{L0Ele};0\gamma}$	11 ± 2	$n_{\mathrm{L0Ele};1\gamma}$	4.3 ± 0.3	$n_{\mathrm{L0Ele};2\gamma}$	2.3 ± 0.2
$n_{\mathrm{L0Had};0\gamma}$	3	$n_{\mathrm{L0Had};1\gamma}$	81	$n_{\mathrm{L0Had};2\gamma}$	43 ± 23
$n_{\mathrm{L0TIS};0\gamma}$	100	$n_{\mathrm{L0TIS};1\gamma}$	100	$n_{\mathrm{L0TIS};2\gamma}$	3.3 ± 0.5
$\sigma^{\mathrm{CB}}_{\mathrm{L0Ele};0\gamma}$	$29.9 \pm 0.5\,\mathrm{MeV}/c^2$	$\sigma^{\mathrm{CB}}_{\mathrm{L0Ele};1\gamma}$	$43.0 \pm 0.8\,\mathrm{MeV}/c^2$	$\sigma^{\mathrm{CB}}_{\mathrm{L0Ele};2\gamma}$	$55 \pm 2\,\mathrm{MeV}/c^2$
$\sigma^{\mathrm{CB}}_{\mathrm{L0Had};0\gamma}$	$26 \pm 2\,\mathrm{MeV}/c^2$	$\sigma^{\mathrm{CB}}_{\mathrm{L0Had};1\gamma}$	$76 \pm 4\,\mathrm{MeV}/c^2$	$\sigma^{\mathrm{CB}}_{\mathrm{L0Had};2\gamma}$	$76 \pm 6\,\mathrm{MeV}/c^2$
$\sigma^{\mathrm{CB}}_{\mathrm{L0TIS};0\gamma}$	$31 \pm 1\,\mathrm{MeV}/c^2$	$\sigma^{\mathrm{CB}}_{\mathrm{L0TIS};1\gamma}$	$70 \pm 2\,\mathrm{MeV}/c^2$	$\sigma^{\mathrm{CB}}_{\mathrm{L0TIS};2\gamma}$	$69 \pm 4\,\mathrm{MeV}/c^2$
		$\sigma^{\mathrm{G}}_{\mathrm{L0Ele};1\gamma}$	$140 \pm 5\,\mathrm{MeV}/c^2$	$\sigma^{\mathrm{G}}_{\mathrm{L0Ele};2\gamma}$	$158 \pm 8\,\mathrm{MeV}/c^2$
		$\sigma^{\mathrm{G}}_{\mathrm{L0Had};1\gamma}$	$184 \pm 30\,\mathrm{MeV}/c^2$	$\sigma^{\mathrm{G}}_{\mathrm{L0Had};2\gamma}$	$330 \pm 30\,\mathrm{MeV}/c^2$
		$\sigma^{\mathrm{G}}_{\mathrm{L0TIS};1\gamma}$	$166 \pm 22\,\mathrm{MeV}/c^2$	$\sigma^{\mathrm{G}}_{\mathrm{L0TIS};2\gamma}$	$165 \pm 23\,\mathrm{MeV}/c^2$

Tab. 3.11: Fitted parameters as obtained from the invariant mass fit to $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ signal MC. When the error on the parameter is not given it means the parameter was fixed in the mass fit.

radiative tail and efficiently separate $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ signal events from partially reconstructed background ones. Namely, the J/ψ constrained $m(K^+\pi^-e^+e^-)$ invariant mass is required to be larger than $5175\,\mathrm{MeV}/c^2$, as shown in Figure 3.32. The $m(K^+\pi^-e^+e^-)$ distribution without the J/ψ mass constraint is then fitted,

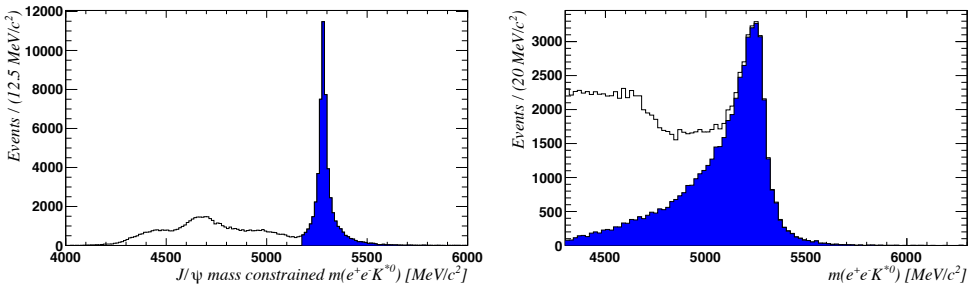


Fig. 3.32: On the left the is the distribution in data of the $m(K^+\pi^-e^+e^-)$ invariant mass calculated by requiring the dielectron mass to match the nominal J/ψ mass. This variable is used to separate the $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ signal events from partially reconstructed background. On the right is the $m(K^+\pi^-e^+e^-)$ distribution obtained without this constrain on the dielectron mass. The solid blue distributions correspond in both plots to the events selected by requiring the J/ψ constrained $m(K^+\pi^-e^+e^-)$ to be larger than $5175\,\mathrm{MeV}/c^2$.

for the three trigger categories, to the shapes determined in Section 3.4.2 by letting float only an additional scale factor (S) multiplying the resolutions of the CB and the Gaussian functions and a shift (δ) on the CB peak position. The S and δ parameters are fitted separately for the three trigger categories, but they are shared by the three bremsstrahlung categories:

$$\mathcal{S}_{ij}^{\mathrm{data}} = f_{ij}^{\mathrm{Core}}\ \mathcal{CB}(M_{ij}^{\mathrm{CB}} + \delta_i, \sigma_{ij}^{\mathrm{CB}}\ S_i, \alpha_{ij}, n_{ij}) + (1. - f_{ij}^{\mathrm{Core}})\ \mathcal{G}(M_{ij}^{\mathrm{G}}, \sigma_{ij}^{\mathrm{G}}\ S_i) \quad (3.10)$$

with i =L0Ele, L0Had or L0TIS and $j = 0\gamma$, 1γ or 2γ. The M_{ij}^{CB}, $\sigma_{ij}^{\mathrm{CB}}$, α_{ij} ,n_{ij} and f_{ij}^{Core} parameters are fixed to the ones obtained in MC (Table 3.11). The PDF is simplified by summing up bremsstrahlung categories:

$$\mathcal{S}_i^{\mathrm{data}} = \sum_{j=0\gamma}^{2\gamma} \Phi_j\ \mathcal{S}_{ij}^{\mathrm{data}} \quad (3.11)$$

where Φ_j are the fractions of bremsstrahlung categories and $\Phi_{0\gamma} + \Phi_{1\gamma} + \Phi_{2\gamma} = 1$ by construction. In the data fit these fractions are fixed to the ones found in MC. Indeed, very good agreement is found between the bremsstrahlung fraction in data and MC. This allows to get an simpler expression for the signal PDF and eases the combination with the background PDFs which are not modelled according to

bremsstrahlung categories.
The results for δ_i and S_i are reported in Table 3.12. An additional exponential component with slope parameter a is added to the signal PDF to model a small contribution of combinatorial background. The fits for the three trigger categories are shown in Figure 3.33. Fits are good and so are the pull distributions. The disagreement at very low $m(K^+\pi^-e^+e^-)$ is due to the requirement that the dielectron mass is larger than 2200 MeV/c^2 (Table 3.7). The obtained mass shifts δ are small and resolution scale factors S are all significantly larger than 1, but not by a large factor.

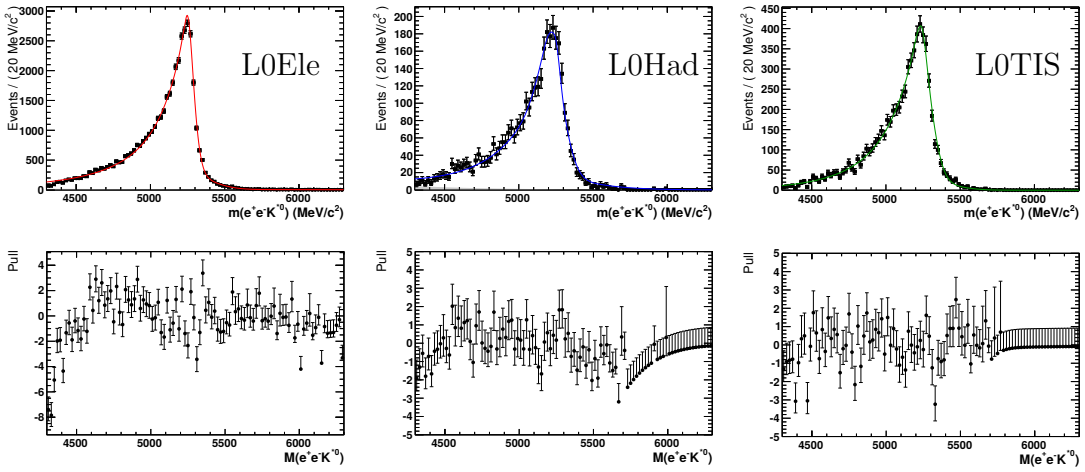


Fig. 3.33: Invariant mass distributions for the $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ decay mode for the three trigger categories as fitted on the data with the additional requirements that the $m(K^+\pi^-e^+e^-)$ mass computed with a J/ψ mass constraint is larger than 5175 MeV/c^2. The exponential function used to model the combinatorial background is in light grey, but its fraction is so small it is hardly visible on the plots. The three bottom plots display the pulls for the three trigger categories.

3.4.3 Modelling of the partially reconstructed background $m(K^+\pi^-e^+e^-)$ distribution.

Partially reconstructed background is expected to come from decays involving higher K^{*0} resonances of the type $B \to e^+e^-Y(\to K^+\pi^-X)$ where X is not reconstructed. Given the dominance of the photon pole contribution in the very low-q^2 region considered, this background is expected to be very similar to the partially reconstructed background affecting the $B^0 \to K^{*0}\gamma_{e^+e^-}$ channel, namely events of the type $B \to \gamma_{e^+e^-}Y(\to K^+\pi^-X)$. Therefore, MC events of this type

Parameter	L0Ele	L0Had	L0TIS
$a^{B^0\to J/\psi K^{*0}} \times 10^3$	-0.8 ± 0.1	-1.9 ± 0.3	-1.0 ± 0.5
S	1.043 ± 0.009	1.16 ± 0.03	1.09 ± 0.02
δ	(3.6 ± 0.6) MeV/c^2	(0.0 ± 2.5) MeV/c^2	(-0.4 ± 1.6) MeV/c^2
$N^{B^0\to J/\psi K^{*0}}$	43467 ± 223	3336 ± 63	6647 ± 82
$N^{\text{comb}}_{\text{L0ele}}$	983 ± 86	119 ± 27	30 ± 13

Tab. 3.12: Fitted parameters as obtained from the invariant mass fit to $B^0 \to J/\psi(e^+e^-)K^{*0}$ data with the additional requirements that the B mass computed with a J/ψ mass constraint is larger than 5175 MeV/c^2. The parameters δ are the shift of the CB peak position, S the scale factors multiplying the widths of the CB and gaussian functions, a are the exponential functions parameters and N are the yields.

channel	N_{gen}	N_{sel}	BR(PDG)
$B^+ \to \gamma K_1^+(1270)$	2×10^6	62	4.3×10^{-5}
$B^+ \to \gamma K_2^{+*}(1430)$	1×10^6	83	1.3×10^{-5}
$B^+ \to \gamma K_1^+(1400)$	1×10^6	107	$\sim 1.3 \times 10^{-5}$

Tab. 3.13: MC samples used to construct the expected partially reconstructed background coming from decays of the type $B \to \gamma Y(\to K\pi X)$.

were used to model the $m(K^+\pi^-e^+e^-)$ shape of the partially reconstructed background affecting $B^0 \to K^{*0}e^+e^-$. Namely, events for the K^* higher resonances listed in Table 3.13 were simulated. The three different samples were normalised according to the BR of each channel. Although measured BRs have large errors, the $m(K^+\pi^-e^+e^-)$ shapes of the various channel are similar, therefore the mass model does not depend strongly on the BR values. Actually, $\mathcal{B}(B^+ \to \gamma K_1^+(1400))$ has just an upper limit at 1.5×10^{-5}, but it was roughly taken equal to the one of $B^+ \to \gamma K_2^{+*}(1430)$. The $m(K^+\pi^-e^+e^-)$ mass shape was then modelled using a "RooKeysPdf" [141] function from the RooFit package in ROOT [142], which provides an un-binned non-parametric estimate of the probability density function. Given the low statistics available, the same shape was taken for the three trigger categories. The result is given in Figure 3.34. It was verified that the single event found above the nominal B mass is a genuine $B^+ \to K_1^+(1270)\gamma$ with $K_1^+(1270) \to \rho(\pi\pi)K$ where one of the two pions was not reconstructed.

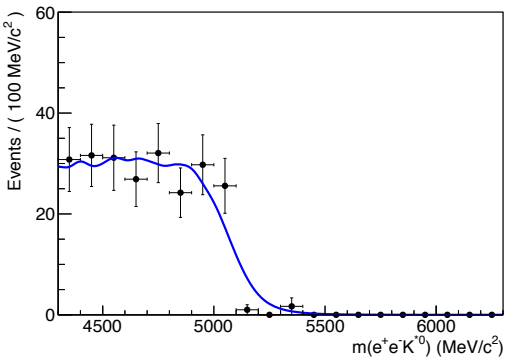


Fig. 3.34: $m(K^+\pi^-e^+e^-)$ distribution for partially reconstructed events of the type $B \to \gamma_{e^+e^-}Y(\to K^+\pi^-X)$ from MC. The higher K^* resonances simulated are listed in Table 3.13. The three trigger categories are summed up and the $m(K^+\pi^-e^+e^-)$ invariant mass distribution is fitted with a "RooKeysPdf" [141].

3.4.4 Invariant mass fit on $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}\gamma_{e^+e^-}$ data

While being a background to the low-q^2 $B^0 \to K^{*0}e^+e^-$ signal (as discussed in Section 3.3.2), the $B^0 \to K^{*0}\gamma_{e^+e^-}$ process is a good control sample having the same final state as the rare channel, very close $m(e^+e^-)$ invariant mass and a much larger yield once the veto cuts are released. The $B^0 \to K^{*0}\gamma_{e^+e^-}$ invariant mass distribution of the signal and of the related backgrounds are expected to be similar to low-q^2 $B^0 \to K^{*0}e^+e^-$, therefore this channel is used to validate the $m(K^+\pi^-e^+e^-)$ invariant mass fit by fitting it to the same functional shape.

As explained in Section 3.4.3, the partially reconstructed backgrounds related to the two channels are expected to be closely related due to the dominance of the photon pole in $B^0 \to K^{*0}e^+e^-$. The similarity holds not only for the $m(K^+\pi^-e^+e^-)$ shape, but also for the fraction of partially reconstructed background h^{PR} with respect to the signal yield. Hence, the information on this fraction is taken from the $B^0 \to K^{*0}\gamma_{e^+e^-}$ invariant mass fit and is then used as a constrain in the $B^0 \to K^{*0}e^+e^-$ mass fit to counterbalance the lack of statistics which does not allow a fit of this fraction in the $B^0 \to K^{*0}e^+e^-$.

Therefore, before fitting the $B^0 \to K^{*0}e^+e^-$ invariant mass distribution, a fit is performed on a data sample being selected with the same criteria with the exception of $B^0 \to K^{*0}\gamma_{e^+e^-}$ vetoes, namely the $\sigma_Z(e^+e^-) < 30\,\mathrm{mm}$ cut, and the $m(e^+e^-) > 20\,\mathrm{MeV}/c^2$ cut.
The functional shape used for both fits are the same. The signal PDF is the one used for the $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ data fit (Equation 3.11). Parameters of this PDF are taken from a fit of the related MC sample and are corrected with MC/data

differences from Table 3.12.To describe the background two other components are added:

- a `RooKeysPdf` to describe the partially reconstructed background PDF, $\mathcal{P}$, as detailed in Section 3.4.3;
- an exponential function to describe the "combinatorial" background PDF, $\mathcal{C}$, which actually includes low-mass background from the semileptonic cascade $B^0 \to D^-(\to K^{*0}e^-\bar{\nu})e^+\nu$.

In summary the overall PDF shape to describe the $m(K^+\pi^-e^+e^-)$ invariant mass fit of data is:

$$\mathcal{PDF}_i^{\text{data}} = (1 - h_i^{\text{comb}}) \left(h_i^{\text{PR}}\mathcal{P}_i + (1 - h_i^{\text{PR}})\, \mathcal{S}_i^{\text{data}}\right) + h_i^{\text{comb}}\, \mathcal{C}(a_i) \tag{3.12}$$

The free parameters for both the $B^0 \to K^{*0}\gamma_{e^+e^-}$ and the $B^0 \to K^{*0}e^+e^-$ fits are: the slope of the exponential function describing the combinatorial background a_i, the fraction of combinatorial background h^{comb} and the fraction of partially reconstructed background h^{PR}. In the $B^0 \to K^{*0}e^+e^-$ fit, a Gaussian constrain is added on the value of h^{PR} having width and central value taken from the $B^0 \to K^{*0}\gamma_{e^+e^-}$ fit (Table 3.14).

The $m(K^+\pi^-e^+e^-)$ invariant mass fit to $B^0 \to K^{*0}\gamma_{e^+e^-}$ is shown in Figure 3.35 and results are given in Table 3.14, while the one to $B^0 \to K^{*0}e^+e^-$ is Figure 3.36 and results are in Table 3.15. The total yield of $B^0 \to K^{*0}e^+e^-$ events fitted is 149 ± 16, which includes a $(3.8 \pm 1.9)\%$ contamination of $B^0 \to K^{*0}\gamma_{e^+e^-}$ events. The fits to the two channels are shown in Figure 3.37 with the three trigger categories added up.

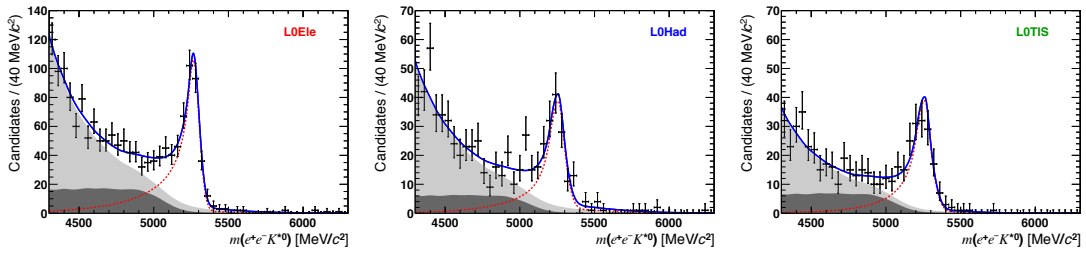


Fig. 3.35: Invariant mass distributions for the $B^0 \to K^{*0}\gamma_{e^+e^-}$ decay mode for the three trigger categories. The dashed line is the signal PDF, the light grey area corresponds to the combinatorial background, the dark grey area is the PR background. The solid line is the total PDF.

Parameter	L0Ele	L0Had	L0TIS
$a \times 10^3$	-3.7 ± 0.2	-3.4 ± 0.3	-3.5 ± 0.4
h^{PR}	0.42 ± 0.05	0.41 ± 0.08	0.43 ± 0.07
h^{comb}	0.58 ± 0.02	0.61 ± 0.03	0.48 ± 0.03
$N^{B^0 \to K^{*0}\gamma_{e^+e^-}}$	621 ± 30	234 ± 20	222 ± 17

Tab. 3.14: Parameters obtained from the invariant mass fit to $B^0 \to K^{*0}\gamma$ data. The total yield of $B^0 \to K^{*0}\gamma_{e^+e^-}$ events is 1077 ± 40 events, which includes a fraction of $B^0 \to K^{*0}e^+e^-$ events of the order of 20%.

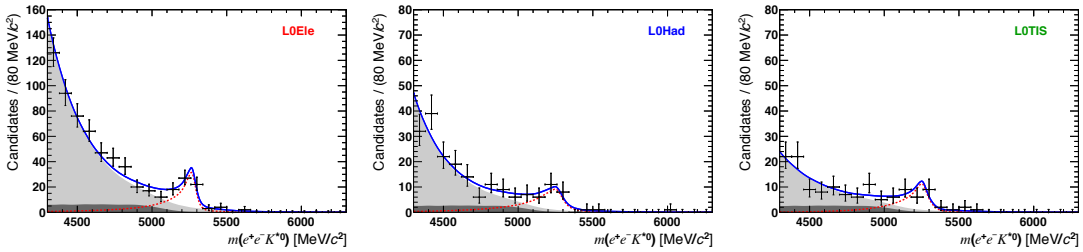


Fig. 3.36: Invariant mass distributions for the $B^0 \to K^{*0}e^+e^-$ decay mode for the three trigger categories. The dashed line is the signal PDF, the light grey area corresponds to the combinatorial background, the dark grey area is the PR background. The solid line is the total PDF.

Parameter	L0Ele	L0Had	L0TIS
$a \times 10^3$	-3.9 ± 0.3	-4.2 ± 0.7	-3.4 ± 0.8
h^{PR}	0.42 ± 0.05	0.40 ± 0.08	0.43 ± 0.07
h^{comb}	0.73 ± 0.04	0.68 ± 0.08	0.57 ± 0.11
$N^{B^0 \to K^{*0}e^+e^-}$	79 ± 12	37 ± 8	33 ± 8

Tab. 3.15: Above the line, the fitted parameters as obtained from the invariant mass fit on $B^0 \to K^{*0}e^+e^-$ data are given. Below the line the signal yields as computed from the parameters are also given.

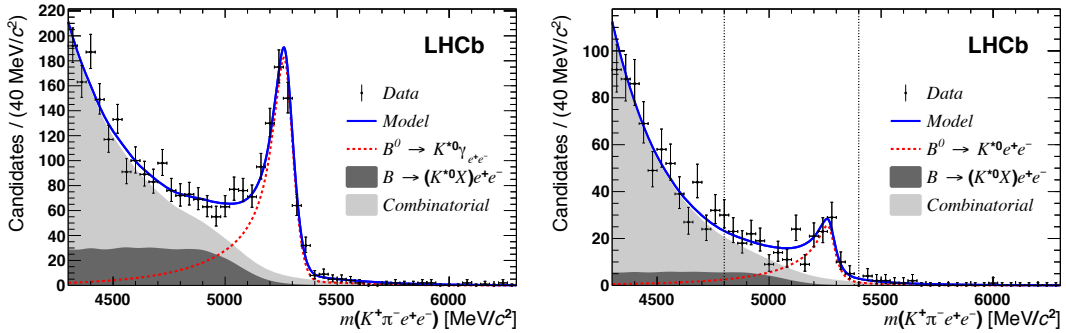


Fig. 3.37: The fits to $B^0 \to K^{*0}\gamma_{e^+e^-}$ and $B^0 \to K^{*0}e^+e^-$ are shown side by side with all trigger categories being summed up.

Chapter 4

Angular analysis of $B^0 \to K^{*0} e^+ e^-$ at low q^2

The selection described in Chapter 3 allows to extract a number of $B^0 \to K^{*0} e^+ e^-$ candidates large enough to perform a fit of the three angles describing the 4-body decay and get sensitivity to the polarisation of $b \to s\gamma$ processes as explained in Chapter 1.

Firstly, the strategy of the angular analysis is presented in Section 4.1. Then, the angular acceptance is extracted in Section 4.2 and background angular distributions are modelled in Section 4.3. The fit to measure angular observables is validated in Section 4.4 and the systematic uncertainties associated to this measurement are discussed in Section 4.5. Finally, results of the fit are presented in Section 4.6 and a discussion of their interpretation follows.

4.1 Angular fit strategy

The goal of this analysis is to fit $B^0 \to K^{*0} e^+ e^-$ signal candidates to the three-dimensional differential branching fraction described in Section 1.3.3, which is a function of $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$. Given the non-negligible background contamination of candidates, the angular distributions of the two main backgrounds have to be modelled and fitted together with the signal. Furthermore, in order to include precious information on the probability of a given candidate to be signal rather than background, the invariant mass $m(K^+\pi^- e^+ e^-)$ is fitted together with the angular variables.
Nonetheless, the angular fit is performed in a smaller $m(K^+\pi^- e^+ e^-)$ range to allow the use of candidates out of it for the modelling of the background angular-distributions. This *angular-fit window* is chosen between 4.8 and 5.4 GeV/c^2, while the one used for the signal yield extraction is between 4.3 and 6.3 GeV/c^2. The

boundaries of the mass window are chosen by maximizing the *significance* figure of merit being defined as $S/\sqrt{S+B}$, where S (B) stands for the signal (background) yield in the window. The yields of signal and background in this mass window are shown in Table 4.1. In this angular-fit mass-window the distributions of $\cos\theta_\ell$

Parameter	L0Ele	L0Had	L0TIS	Total
N^{sig}	64 ± 7	29 ± 5	30 ± 3	124 ± 9
N^{comb}	56 ± 10	14 ± 6	13 ± 3	83 ± 12
N^{PR}	20 ± 4	9 ± 3	9 ± 3	38 ± 5
S/B	0.85	1.30	1.40	1.03
$S/\sqrt{S+B}$	5.4	4.1	4.2	7.9

Tab. 4.1: Yields for the three trigger categories in the angular-fit mass-window $[4.8, 5.4]\,\text{GeV}/c^2$ for $B^0 \to K^{*0}e^+e^-$ data. The mass fit is described in Section 3.4.4.

$\cos\theta_K$ and $\tilde{\phi}$ are not looked at before the full fit is judged to be stable and has been validated on control samples from data and MC. This procedure of not looking at the final result of the fit during the whole analysis is widely used in experimental particle physics in order to prevent the analysts from biasing the results towards their own preconceptions. Nowadays, this is becoming more important as analysis methods are more complex and experiments are more difficult to reproduce.
The geometry of the LHCb detector as well as the processes of reconstruction and selection introduce a distortion of angular distributions which has to be modelled and corrected for. The modelling of this angular acceptance function is described in detail in Section 4.2. The strategy chosen to apply this correction is to multiply the angular part of the probability distribution function of the signal to the acceptance function.
Therefore, the probability density function (PDF) of the signal reads:

$$\begin{aligned} &\mathcal{S}_i\left(m_B, \cos\theta_\ell, \cos\theta_K, \tilde{\phi}\right) = \\ &\mathcal{S}_i^{\text{data}}(m_B)\ \ \mathcal{S}\left(\cos\theta_\ell, \cos\theta_K, \tilde{\phi}; F_{\text{L}}, \text{A}_{\text{T}}^{(2)}, \text{A}_{\text{T}}^{\text{Im}}, \text{A}_{\text{T}}^{\text{Re}}\right)\ \ \varepsilon_i\left(\cos\theta_\ell, \cos\theta_K, \tilde{\phi}\right) \end{aligned} \tag{4.1}$$

the subscript i standing for the trigger category and the variable m_B for $m(K^+\pi^-e^+e^-)$. The mass part $\mathcal{S}_i^{\text{data}}(m_B)$ is defined in Equation 3.11, the angular acceptance $\varepsilon_i(\cos\theta_\ell, \cos\theta_K, \tilde{\phi})$ is given in Section 4.2 and the angular part $\mathcal{S}\left(\cos\theta_\ell, \cos\theta_K, \tilde{\phi}; F_{\text{L}}, \text{A}_{\text{T}}^{(2)}, \text{A}_{\text{T}}^{\text{Im}}, \text{A}_{\text{T}}^{\text{Re}}\right)$ is in Equation 1.32. The acceptance function as well as the mass part depend on the trigger category of the candidate while the angular PDF, containing all the underlying physics information, is of course

independent of it and in general of the detector response.
Adding the PDF for the two types of background, the total PDF reads:

$$\begin{aligned}\mathcal{PDF}_i^{\text{Full}} = \quad & (1-h_i^{\text{comb}}) \quad \Big[(1-h_i^{\text{PR}})\ \mathcal{S}_i\left(m_B, \cos\theta_\ell, \cos\theta_K, \tilde{\phi}\right) \\ & \qquad +h_i^{\text{PR}}\ \mathcal{P}_i\left(m_B, \cos\theta_\ell, \cos\theta_K, \tilde{\phi}\right)\Big] \\ & +h_i^{\text{comb}} \qquad \mathcal{C}_i\left(m_B, \cos\theta_\ell, \cos\theta_K, \tilde{\phi}\right)\end{aligned} \tag{4.2}$$

where $\mathcal{S}_i$, $\mathcal{P}_i$ and $\mathcal{C}_i$ are the $4D$ functions for the signal, the partially reconstructed background and the combinatorial background respectively. The fractions h_i^{comb} are taken from the $B^0 \to K^{*0}e^+e^-$ mass fit in the larger mass range $[4.3, 6.3]\,\text{GeV}/c^2$ (Table 3.15) and recalculated in the smaller mass range $[4.8, 5.4]\,\text{GeV}/c^2$ (results are in Table 4.2). On the other hand, the fractions of partially reconstructed background h_i^{PR} are taken from the $B^0 \to K^{*0}\gamma_{e^+e^-}$ fit and recalculated in the smaller mass window (results are in Table 4.2). Both fractions are put as input to the fit as Gaussian constraints.

	L0Ele	L0Had	L0TIS
h_i^{PR}	0.24 ± 0.03	0.22 ± 0.06	0.25 ± 0.05
h_i^{comb}	0.38 ± 0.07	0.30 ± 0.13	0.24 ± 0.07

Tab. 4.2: Fractions of background recalculated in the signal mass window. Errors shown are coming from the limited size of the sample.

As for the signal, the PDFs for the two types of background are factorised in mass part and angular part, namely

$$\mathcal{P}_i\left(m_B, \cos\theta_\ell, \cos\theta_K, \tilde{\phi}\right) = \mathcal{P}_i(m_B)\ \mathcal{P}_i\left(\cos\theta_\ell, \cos\theta_K, \tilde{\phi}\right) \tag{4.3}$$

$$\mathcal{C}_i\left(m_B, \cos\theta_\ell, \cos\theta_K, \tilde{\phi}\right) = \mathcal{C}_i(m_B)\ \mathcal{C}_i\left(\cos\theta_\ell, \cos\theta_K, \tilde{\phi}\right) \tag{4.4}$$

where $P_i(m_B)$ and $\mathcal{C}(m_B)$ were defined in Section 3.4.4 while the angular part is described in Sections 4.3.2 and 4.3.1.

4.2 Angular Acceptance

The geometry of the LHCb detector and the processes of reconstruction and selection may introduce a distortion on the three angles considered. In order to isolate these acceptance effects, a phase-space MC sample of $B^0 \to K^{*0}e^+e^-$ events is

generated. The $B^0 \to K^{*0}e^+e^-$ decay process is generated with no underlying physics but momentum conservation which means the Dalitz plane $m^2(e^+e^-)$ vs. $m^2(K^{*0}e^+)$ is populated evenly. This results in generating flat uncorrelated distributions for $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$, which are very convenient to study the effect of acceptance. The whole $B^0 \to K^{*0}e^+e^-$ selection is applied (Sec. 3.2) and the resulting events are fitted to extract acceptance functions. The acceptance is assumed to be factorizable, namely:

$$\varepsilon(\cos\theta_\ell, \cos\theta_K, \tilde{\phi}) = \varepsilon(\cos\theta_\ell)\ \varepsilon(\cos\theta_K)\ \varepsilon(\tilde{\phi}). \tag{4.5}$$

This hypothesis is validated in Section 4.2.3.

4.2.1 Phase-space Monte Carlo weighting

The phase space MC generated has a very different q^2 distribution compared to the one generated with Standard Model physics, the main difference being the absence of the photon pole. Despite the q^2 range considered in this analysis being quite narrow, the acceptance has some correlation with the di-electron mass. For example, as shown in Figure 4.1, one of the parameters of the polynomial used to model the $\cos\theta_\ell$ acceptance has lower values for large $m(e^+e^-)$ The parameter here shown, c_2^ℓ, is the one varying the most with $m(e^+e^-)$. Therefore, modelling the

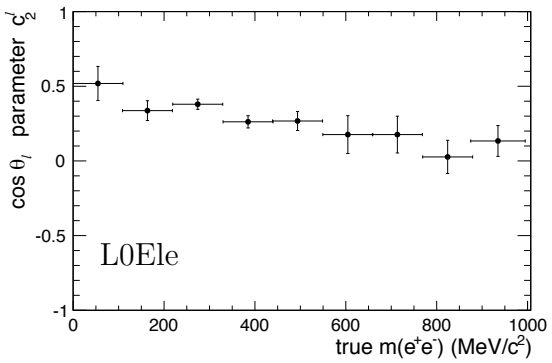


Fig. 4.1: The parameter c_2^ℓ used for the modelling of the $\cos\theta_\ell$ angular acceptance of the L0Ele trigger category (see Section 4.2.2) is extracted in bins of the $m(e^+e^-)$ invariant mass.

acceptance on a phase-space MC sample having a dramatically different $m(e^+e^-)$ distribution is incorrect. One solution to this issue could be to use the MC simulated with SM physics and to make the ratio between the angular distribution selected over the one generated. But this solution would be complicated for $\cos\theta_\ell$ and $\cos\theta_K$ since their SM distributions change as a function of q^2 (because of

the dependence on q^2 of the longitudinal fraction $F_{\rm L}$) and at the same time the selection efficiency changes as a function of q^2. Therefore, to compensate the lack of statistics at low q^2, a second sample of 2 millions phase space events with a requirement on $m(e^+e^-) < 440\,\mathrm{MeV}/c^2$ at generator level is simulated and added to the present one. This sample is then weighted in order to reproduce the expected SM q^2-distribution. Weights are calculated in bins of $\sim 100\,\mathrm{MeV}/c^2$. A finer binning is not needed as the acceptance does not change very fast with q^2. A coarser binning is preferred to avoid having events with very large weights. Figure 4.2 shows the phase space sample, enriched with events at $m(e^+e^-)$ lower than $450\,\mathrm{MeV}/c^2$, before and after weights are applied. The weighted distribution is in good agreement with the SM MC one.

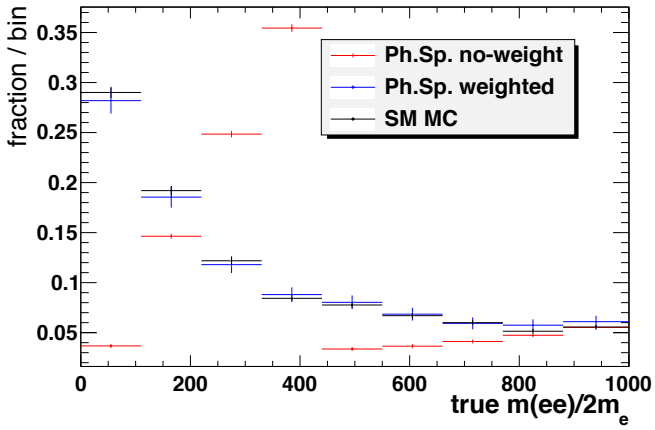


Fig. 4.2: In red is the $m(ee)/(2m_e)$ distribution of the sum of the two $B^0 \to K^{*0}e^+e^-$ phase space samples. The weighted distribution of this sample is in blue and is to be compared to the SM Monte Carlo distribution, which is in black.

4.2.2 Angular acceptance fit

Distributions for $\cos\theta_\ell$ and $\cos\theta_K$ of the weighted phase-space MC sample after the whole selection are fitted to Legendre polynomials to extract the acceptance functions. Polynomials are truncated at order 4 (Equation 4.6) as the statistics of the sample does not give sensitivity to higher degrees. Nonetheless, given the yield of $B^0 \to K^{*0}e^+e^-$ events in data to be fitted, this approximation is accurate enough.

$$1+\sum_{i=1}^{4} c_i P_i(x) = 1+c_1 x+c_2\frac{1}{2}(3x^2-1)+c_3\frac{1}{2}(5x^3-3x)+c_4\frac{1}{8}(35x^4-30x^2+3) \quad (4.6)$$

Choosing Legendre polynomials minimizes the correlation between the fitted parameters thus allowing a simpler treatment of the systematic uncertainties introduced by the modelling of the acceptance. This property is assured by the orthonormality of Legendre polynomials with respect to an integration between -1 and 1, namely $\int_{-1}^{1} P_n(x)P_m(x)\mathrm{d}x = 0$ if $n \neq m$. Note that Chebychev polynomials, another set of polynomials widely used to model acceptance functions, do not have this property, they are just orthogonal with respect to the weight $\frac{1}{\sqrt{1-x^2}}$.
Due to different cuts on the p_{T} of final state particles, the three trigger categories have significantly different acceptance curves, therefore they are modelled separately. All fits are based on an unbinned maximum likelihood.

The curves chosen to model the $\cos\theta_\ell$ acceptance are taken symmetric by default (i.e. parameters c_1^ℓ and c_3^ℓ are fixed to 0), since no difference is expected in the detector response to electrons and positrons. Nonetheless, a second fit is done including a linear term to check for a potential bias that could affect the measurement of $\mathrm{A_T^{Re}}$. No significant effect is found, but the sensitivity on the asymmetry is included in the estimate of systematic uncertainties (Section 4.5). The $\cos\theta_\ell$ angular distributions for the three trigger categories are shown in Figure 4.3 and results are in Table 4.3. As expected the $\cos\theta_\ell$ acceptance for the L0Ele trigger category is different from the other two due to the threshold on the electron p_T to trigger the line.

Acceptance shapes for $\cos\theta_K$ are naturally asymmetric due to the mass difference between K and π which results in a momentum imbalance in the laboratory reference frame. The $\cos\theta_K$ angular distributions are in Figure 4.4 and the fit results in Table 4.4. The threshold on the p_T of one of the hadrons results in a $\cos\theta_K$ acceptance of the L0Had category having a shape clearly different from the other two.

For the angle $\tilde{\phi}$ the MC generated following SM physics rather than the phase-space one is used for the extraction of the acceptance. This choice is driven by the possibility to have a smaller statistical uncertainty on the acceptance modelling and is possible thanks to the fact that in such a low q^2 range there is no modulation in $\tilde{\phi}$ at generator level. Indeed, in the SM the physical parameters $\mathrm{A_T^{(2)}}$ and $\mathrm{A_T^{Im}}$ are very close to 0 for q^2 values below $1\,\mathrm{GeV}^2/c^4$. Furthermore, $\mathrm{A_T^{(2)}} = 0$ (and $\mathrm{A_T^{Im}} = 0$) means no correlation of the angle $\tilde{\phi}$ with $\cos\theta_\ell$ and $\cos\theta_K$ at generator level. A fit to $\tilde{\phi}$ at generator level is presented in Section 4.4.1 (Table 4.10 and Figure 4.21) and shows that it is compatible with being flat. The $\tilde{\phi}$ acceptance is expected to be flat since it is quite difficult to find an experimental effect which would bias it. One way would be a cut on an invariant mass made with one of the electrons and one of the hadrons, but no such cut is used in the offline selection. Another way would be a much higher efficiency for tracks being on the same plane in the laboratory frame which may translate to a better acceptance for small and large

$\tilde{\phi}$ angles. No significant effect is found. The most dangerous effects would be in $\cos(2\tilde{\phi})$ or in $\sin(2\tilde{\phi})$ since such modulations may be confused with physics coming from non-zero values of $A_T^{(2)}$ or A_T^{Im}. Therefore, effects of this type are checked by fitting with curves of type $1 + c_c \cos(2\tilde{\phi})$ and $1 + c_s \sin(2\tilde{\phi})$. No significant effect is found, i.e. the fitted parameter c_c, c_s are compatible with 0 for all three trigger categories. The error on the parameters c_c, c_s will be included in the systematic errors of the full data fit to take into account the degree of knowledge on the flatness of the $\tilde{\phi}$ acceptance. The $\tilde{\phi}$ angular distributions for the three trigger categories are shown in Figure 4.5 and the results are in Table 4.5.

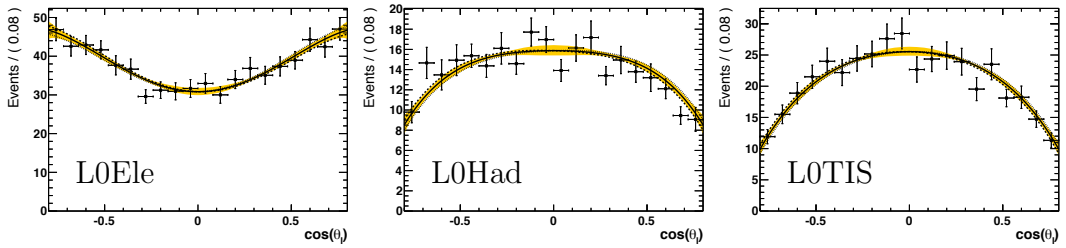


Fig. 4.3: Angular acceptance curves for $\cos\theta_\ell$ for the three trigger categories: L0Ele, L0Had and L0TIS. Polynomials are fitted to distributions taken from Phase space MC re-weighted as explained in Sec. 4.2.1. The default acceptance curves are represented by a solid line. The orange region shows the excursion used in the evaluation of systematics uncertainties. In addition, alternative shapes with a linear term in $\cos\theta_\ell$ were tested and are represented by dotted and dashed curves.

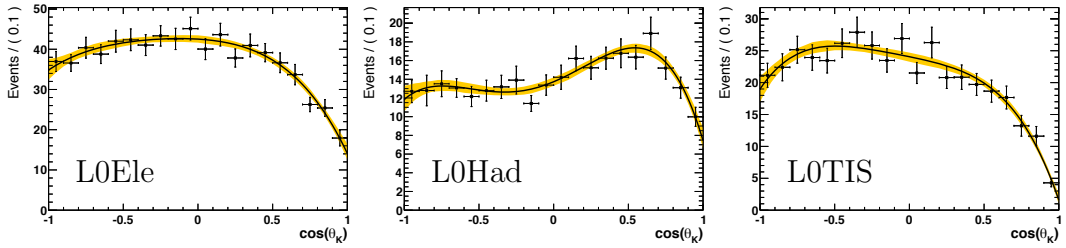


Fig. 4.4: Angular acceptance curves for $\cos\theta_K$ for the three trigger categories: L0Ele, L0Had and L0TIS. Polynomials are fitted to distributions taken from Phase space MC re-weighted as explained in Sec. 4.2.1. The default acceptance curves are represented by a solid line. The orange region shows the excursion used in the evaluation of systematic uncertainties.

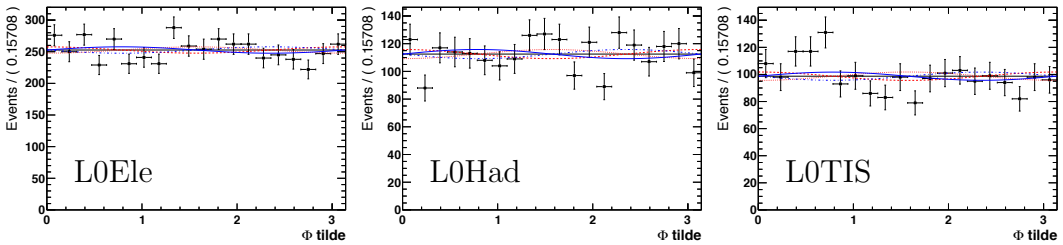


Fig. 4.5: Angular acceptance curves for $\tilde{\phi}$ for the three trigger categories: L0Ele, L0Had and L0TIS. Acceptance is taken from MC generated following the Standard Model. The default acceptance curves used are flat (represented in black). Alternative shapes tested for the evaluation of systematics are shown in red and blue. They have a non-zero term in either $\cos 2\tilde{\phi}$ (red, dashed and dotted lines) or $\sin 2\tilde{\phi}$ (blue, solid and dashed-dotted).

parameter	L0Ele	L0Had	L0TIS
c_2^ℓ	$+0.309 \pm 0.034$	-0.317 ± 0.042	-0.471 ± 0.041
c_4^ℓ	-0.065 ± 0.046	-0.076 ± 0.059	-0.058 ± 0.052
c_1^ℓ	$+0.008 \pm 0.027$	-0.058 ± 0.036	-0.068 ± 0.033

Tab. 4.3: Parameters of the polynomial functions of the type $1 + \sum_{i=1} c_i^\ell P_i(x)$ used to model the acceptance in $\cos\theta_\ell$. A first degree parameter is added to check if there is any significant asymmetry.

parameter	L0Ele	L0Had	L0TIS
c_1^K	-0.190 ± 0.024	$+0.096 \pm 0.035$	-0.339 ± 0.032
c_2^K	-0.301 ± 0.031	-0.108 ± 0.043	-0.382 ± 0.042
c_3^K	-0.086 ± 0.037	-0.253 ± 0.053	-0.073 ± 0.049
c_4^K	-0.047 ± 0.042	-0.205 ± 0.060	-0.128 ± 0.051

Tab. 4.4: Parameters of the polynomial functions of the type $1 + \sum_{i=1} c_i^K P_i(x)$ used to model the acceptance in $\cos\theta_K$.

parameter	L0Ele	L0Had	L0TIS
c_c	-0.014 ± 0.020	-0.023 ± 0.029	$+0.059 \pm 0.031$
c_s	$+0.010 \pm 0.020$	-0.012 ± 0.029	$+0.056 \pm 0.031$

Tab. 4.5: Results of the fit of the $\tilde{\phi}$ acceptance to possible dangerous effects of the type $1 + c_c \cos(2\tilde{\phi})$ and $1 + c_s \cos(2\tilde{\phi})$. All fitted a_i $(i = c, s)$ parameters are found to be compatible with 0, therefore the acceptance is taken flat by default.

4.2.3 Acceptance factorization

The hypothesis that the acceptance can be factorized in the three variables $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ is validated in the following.

Testing the factorization $\varepsilon(\cos\theta_\ell, \cos\theta_K) = \varepsilon(\cos\theta_\ell)\ \varepsilon(\cos\theta_K)$

The hypothesis of factorization between $\cos\theta_\ell$ and $\cos\theta_K$ is tested by trying to recover the generator level distributions using a weight $(\varepsilon(\cos\theta_\ell)\ \varepsilon(\cos\theta_K))^{-1}$. The weighted distribution is therefore expected to be flat for both $\cos\theta_\ell$ and $\cos\theta_K$. To assess the degree of flatness the weighted distributions in $\cos\theta_\ell$ and $\cos\theta_K$ are fitted to a Legendre polynomial of order 4. In principle any correlation effect should show up as a non-zero polynomial parameter, but no significant deviation from 0 is found. All distributions and results are reported in Appendix B.3. For each angle and each trigger category, distributions weighted with $(\varepsilon(\cos\theta_\ell)\ \varepsilon(\cos\theta_K))^{-1}$ are compared with flat-by-construction distributions weighted with either $\varepsilon^{-1}(\cos\theta_\ell)$ or $\varepsilon^{-1}(\cos\theta_K)$. An example for $\cos\theta_\ell$ in the L0Ele trigger category is shown in Figure 4.6.

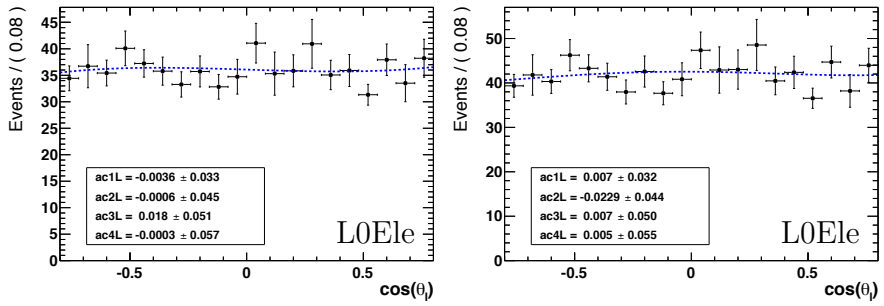


Fig. 4.6: Phase space MC for L0Ele trigger category. The distribution on the right is weighted with $1/\varepsilon(\cos\theta_\ell)$ (it is therefore flat by construction), the one on the left is weighted with $1/(\varepsilon(\cos\theta_\ell)\varepsilon(\cos\theta_K))$. Both are fitted to a Legendre polynomial of order 4.

Testing the factorization of the $\tilde{\phi}$ acceptance with respect to $\cos\theta_\ell$ and $\cos\theta_K$.

In order to test if the acceptance in $\tilde{\phi}$ is correlated to the acceptance in $\cos\theta_\ell$ or to the one in $\cos\theta_K$, the Standard Model MC (which is used to model the $\tilde{\phi}$ acceptance) is weighted with either $\varepsilon^{-1}(\cos\theta_\ell)$ or $\varepsilon^{-1}(\cos\theta_K)$. The weighted $\tilde{\phi}$ distributions are found to remain flat, the "amount of flatness" being evaluated once again by fitting with $1+c_c\cos(2\tilde{\phi})$ and $1+c_s\sin(2\tilde{\phi})$. In Figure 4.7 it is shown

	c_c		c_s	
	$\varepsilon^{-1}(\cos\theta_\ell)$	$\varepsilon^{-1}(\cos\theta_K)$	$\varepsilon^{-1}(\cos\theta_\ell)$	$\varepsilon^{-1}(\cos\theta_K)$
L0Ele	-0.011 ± 0.020	-0.010 ± 0.020	0.006 ± 0.020	0.017 ± 0.021
L0Had	-0.026 ± 0.030	-0.015 ± 0.029	-0.008 ± 0.030	-0.010 ± 0.030
L0TIS	0.059 ± 0.033	0.045 ± 0.029	0.049 ± 0.033	0.060 ± 0.033

Tab. 4.6: Results of the fit to the $\tilde{\phi}$ distribution weighted with the inverse of the acceptances in $\cos\theta_\ell$ and $\cos\theta_K$ to test a possible correlation.

the fit for the L0Ele trigger category while all the others are in Appendix B.4. All fitted parameters c_c and c_s for the weightings $\varepsilon^{-1}(\cos\theta_\ell)$ and $\varepsilon^{-1}(\cos\theta_K)$ are reported in Table 4.6.

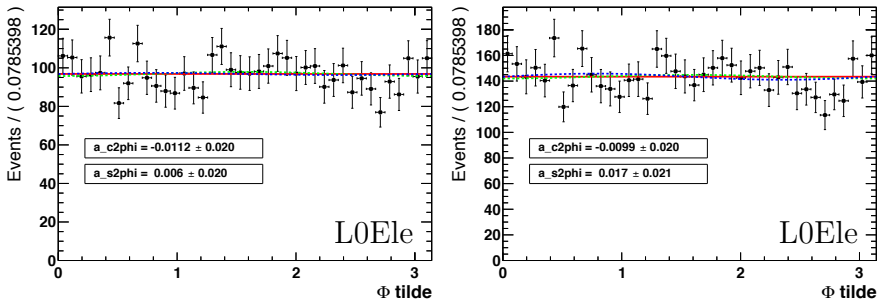


Fig. 4.7: Standard Model MC for L0Ele trigger category. The distribution on the right is weighted with $1/\varepsilon(\cos\theta_\ell)$, the one on the left is weighted with $1/\varepsilon(\cos\theta_K)$. Both are fitted to $1 + c_c\cos(2\tilde{\phi})$ (dashed green line) and $1 + c_s\sin(2\tilde{\phi})$ (dashed blue line).

4.2.4 Validation of the angular acceptance on data

In order to validate the angular acceptance modelling (and indirectly the MC/Data corrections) the angular distributions for MC and data are compared for the signal proxy $B^0 \to J/\psi\,(e^+e^-)K^{*0}$. The *sPlot* technique is used on data to extract the signal distribution (as it is done in Section 3.2.3). As shown in Figures 4.8 and 4.9, the distributions for $\cos\theta_\ell$ and $\tilde{\phi}$ are in good agreement. However, the same comparison cannot be used for $\cos\theta_K$ because the $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ MC does not include the S-wave contribution.

Nonetheless, one can check the modelling of the $\cos\theta_K$ acceptance on the other proxy: $B^0 \to K^{*0}\gamma_{e^+e^-}$. The $B^0 \to K^{*0}\gamma_{e^+e^-}$ decay is expected to have $F_{\rm L} = 0$ due to the polarization of the real photon being purely transversal. Also, a

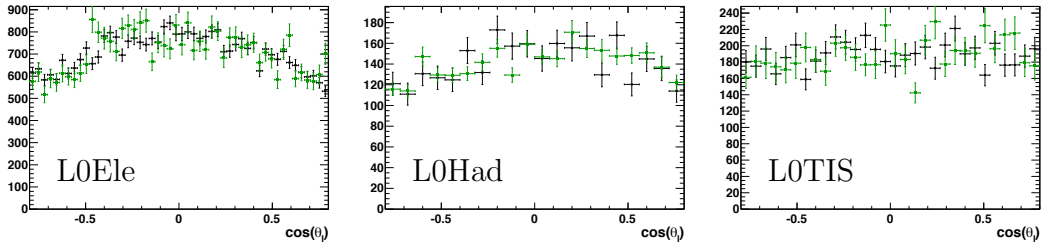


Fig. 4.8: $\cos\theta_\ell$ distributions for $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ for the three trigger categories. In black is an *sPlot* of data while in green is the MC with all MC/Data corrections applied.

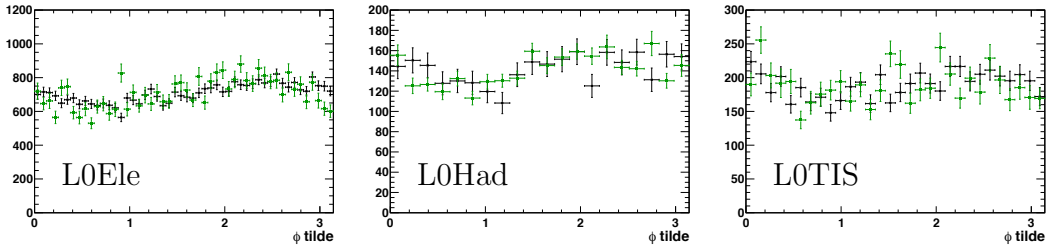


Fig. 4.9: $\tilde{\phi}$ distributions for $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ for the three trigger categories: from left to right L0Ele, L0Had, L0TIS. In black is an *sPlot* of data while in green is the MC with all MC/Data corrections applied.

data sample containing mainly $B^0 \to K^{*0}\gamma$ events can be selected asking the diletpon mass to be lower than $20\,\mathrm{MeV}/c^2$. It will contain also a small amount of $B^0 \to K^{*0}e^+e^-$ events, which however are also expected to have $F_\mathrm{L} = 0$ in this dielectron region. Therefore, a 2D fit of $m(K^+\pi^-e^+e^-)$ together with the $\cos\theta_K$ angle is done to extract F_L. The backgrounds are modelled in the same way as for $B^0 \to K^{*0}e^+e^-$ with the caveat that F_LPR for the modelling of the partially reconstructed background is taken to be 0.17 as it is fitted on $B^0 \to K^{**}\gamma$ MC in Section 4.3.2.

The $\cos\theta_K$ acceptance used for this test is taken from a $B^0 \to K^{*0}e^+e^-$ Standard Model MC sample requiring the dilepton mass to be below $20\,\mathrm{MeV}/c^2$. Due to the very small q^2 window and the fact that F_L is close to 0 at very-low q^2 in the SM, $\cos\theta_K$ and $\cos\theta_\ell$ are decorrelated and the acceptance can be extracted directly from the SM MC rather than from the phase space MC which lacks of statistic at very low $m(e^+e^-)$.

The result of the fit is $F_\mathrm{L} = 0.019 \pm 0.022$, thus well compatible with 0 to a precision of about 2%, much better than the one expected for the $B^0 \to K^{*0}e^+e^-$ full fit.

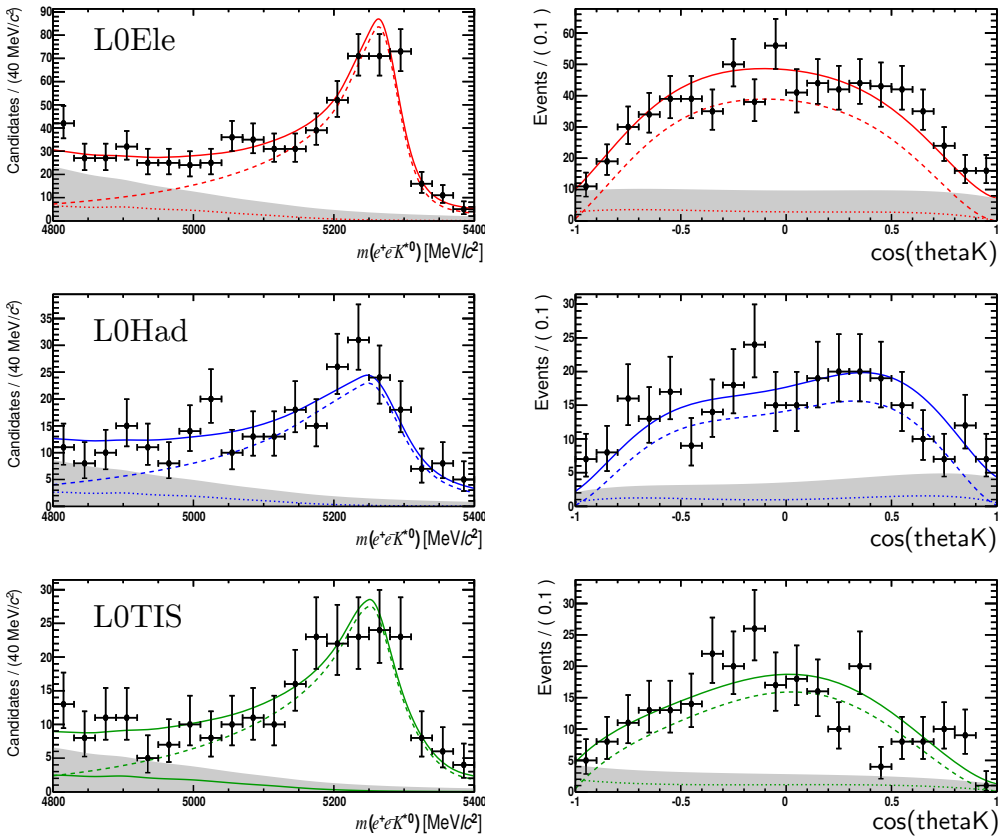
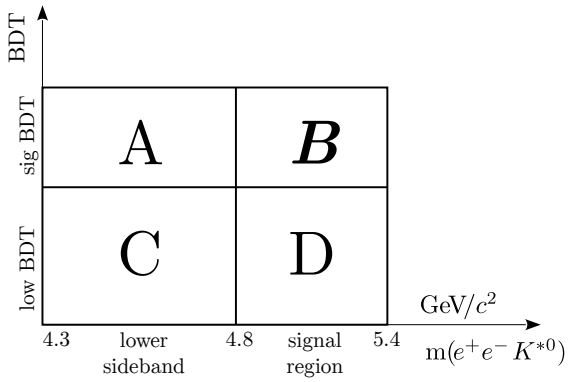


Fig. 4.10: 2D fit to $B^0 \to K^{*0}\gamma_{e^+e^-}$ to extract the value of $F_{\rm L}$ and validate the $\cos\theta_K$ angular acceptance. The solid line is the fit to the total PDF, the dotted line is the signal PDF, the dashed line the partially reconstructed background and the light grey area is the combinatorial background. The result of the fit is $F_{\rm L} = 0.019 \pm 0.022$, compatible with 0 to the 2% level.

4.3 Angular modelling of background

4.3.1 Angular modelling of the combinatorial background

Combinatorial background largely dominates at low mass: between 4.3 GeV/c^2 and 4.8 GeV/c^2 (below the signal window) the data mass fit (Sec. 3.4.4) predicts a fraction of 7% of partially reconstructed background and 3% of signal. This lower side-band sample can be used to infer the angular distribution of background in the signal window, but one has to take care of the possible correlation of the angular distributions with the reconstructed mass $m(K^{*0}e^+e^-)$. In particular, a correlation is found in the $\cos\theta_\ell$ distribution, which is more asymmetric at lower $m(K^{*0}e^+e^-)$. The effect of this correlation on the changing of the angular distributions between the lower sideband and the signal window is extracted from the ratio of distributions of the same two mass windows at lower BDT response values, namely between -0.5 and the signal BDT cut. In this region of reduced BDT response the combinatorial background largely dominates also in the angular-fit mass window. So, the BDT-value vs $m(K^+\pi^-e^+e^-)$ space is divided in four regions labelled A, B, C and D as described by the following scheme:



Then, one can estimate, for example, the $\cos\theta_\ell$ distribution in the signal region $B(\cos\theta_\ell)$ taking the distributions of the lower sideband sample $A(\cos\theta_\ell)$, and multiplying and dividing it by the distributions at lower BDT to account for the changing in mass window, namely by doing:

$$B(\cos\theta_\ell) = A(\cos\theta_\ell)\frac{D(\cos\theta_\ell)}{C(\cos\theta_\ell)} \tag{4.7}$$

These four distributions are shown in Figure 4.11 for the L0Ele trigger category. The same procedure is used to estimate the distributions for $\cos\theta_K$ and all results for $\cos\theta_\ell$ and $\cos\theta_K$ and for the three trigger categories are in Appendix B.5.

Distributions for $\cos\theta_\ell$ are modelled with a polynomial of type

$$P^{\rm exp}(\cos\theta_\ell) = 1 + a_3^\ell \cos^3\theta_\ell + a_4^\ell \cos^4\theta_\ell. \tag{4.8}$$

This choice of parametrization is arbitrary and is a good enough description of the data. These two parameters, being even and odd, are not correlated, thus the evaluation of systematic errors is simple and there is no need to use orthonormal Legendre polynomials. For $\cos\theta_K$ just a linear term is used:

$$P^{\rm exp}(\cos\theta_K) = 1 + a_1^K \cos\theta_K \tag{4.9}$$

All results for the modelling of the $\cos\theta_\ell$ and $\cos\theta_K$ distributions are reported in Table 4.7. The $\tilde{\phi}$ distributions are fitted to the function

$$P^{\rm exp}(\tilde{\phi}) = 1 + a_c \cos(2\tilde{\phi}) + a_s \sin(2\tilde{\phi}) \ , \tag{4.10}$$

but no significant deviation from flatness is found when fitting the trigger categories separately. The three trigger categories are therefore added up and the degree of flatness is measured in the summed up distribution (see Figure 4.12). The results of the fit are also shown in Table 4.7. They are compatible with a flat distribution. The angular PDF for the combinatorial background then reads:

$$P_i^{\rm exp}(\cos\theta_\ell, \cos\theta_K, \tilde{\phi}) = P_i^{\rm exp}(\cos\theta_\ell)\ P_i^{\rm exp}(\cos\theta_K)\ P^{\rm exp}(\tilde{\phi}) \tag{4.11}$$

where i =L0Ele, L0Had, L0TIS.

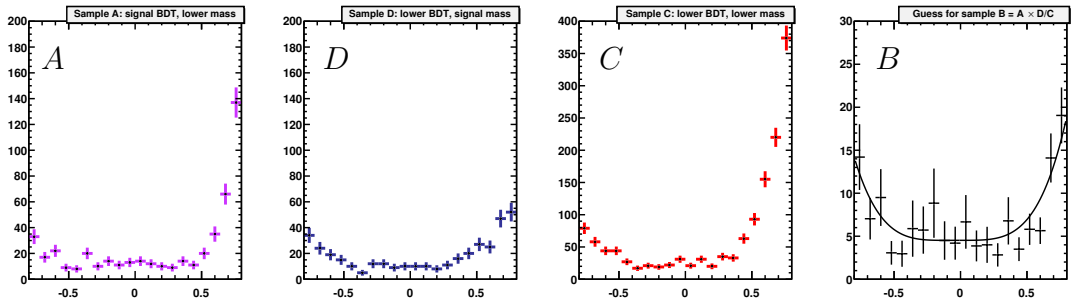


Fig. 4.11: Distributions of the $\cos\theta_\ell$ angle for the 3 samples A,D and C defined above in the L0Ele category. On the right is the guess for the distribution of the sample B obtained as $A\ D/C$.

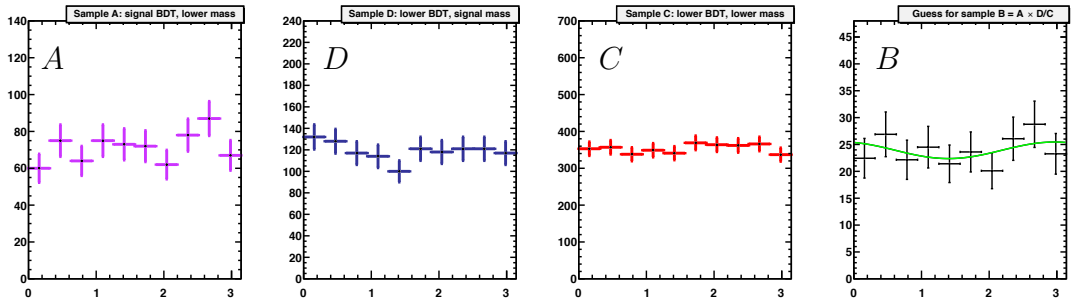


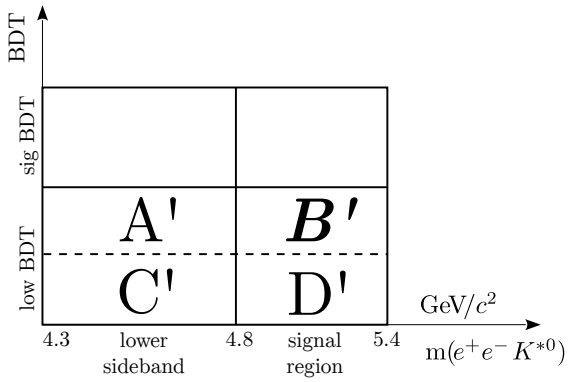
Fig. 4.12: Distributions of the $\tilde{\phi}$ angle for the 3 samples A,D and C defined above for all trigger categories summed up. On the right is the guess for the distribution of the sample B obtained as $A\ D/C$.

	L0Ele	L0Had	L0TIS
a_3^ℓ	1.03 ± 0.74	0.63 ± 0.66	0.12 ± 0.78
a_4^ℓ	6.5 ± 2.1	1.6 ± 1.4	1.3 ± 1.7
a_1^K	0.04 ± 0.13	0.43 ± 0.17	-0.36 ± 0.21
a_c	0.06 ± 0.07		
a_s	-0.02 ± 0.07		

Tab. 4.7: Results of the fit to the $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ distributions of the combinatorial background.

Method validation

The method is validated splitting again in four sub-samples the events having a BDT response value below the signal cut:



These four samples, dubbed A′, B′, C′ and D′, are all dominated by combinatorial background, therefore the guessed $B' = A'D'/C'$ distribution can be checked against the real one. The lower BDT window is between -0.5 and 0.2, while the higher one is between 0.2 and the signal cut (see Table 3.7). The test is done both separately for the three trigger cateogories and also summing them up to gather more statistics. No significant effect is found, so the method is judged to be accurate enough. This means there is no strong correlation between the BDT value and the mass, or at least the effect of these correlations is small compared to the statistical errors on the parameters which are included in the estimate of the systematic error. All the results are shown in Table 4.8. An example of the plot of the 4 subsamples and the guess of the sample at high mass and high BDT is shown in Figure 4.13 for the $\cos\theta_\ell$ angle in the L0Ele trigger category. The plots for all angles and all trigger categories are in Appendix B.6.

	L0Ele	L0Had	L0TIS	All
a_3^ℓ data	3.9 ± 1.2	1.5 ± 0.5	1.3 ± 0.7	2.0 ± 0.4
a_3^ℓ guess	1.0 ± 1.3	0.0 ± 0.8	0.5 ± 0.7	0.6 ± 0.5
a_4^ℓ data	12.2 ± 3.1	1.1 ± 0.9	4.5 ± 1.5	4.8 ± 0.9
a_4^ℓ guess	11.2 ± 4.2	2.4 ± 1.9	3.2 ± 1.7	5.2 ± 1.2
a_1^K data	-0.10 ± 0.14	0.41 ± 0.12	-0.14 ± 0.13	0.06 ± 0.08
a_1^K guess	0.07 ± 0.16	0.46 ± 0.15	-0.20 ± 0.17	0.09 ± 0.09

Tab. 4.8: Results of the fit for the validation of the method used to estimate the angular distribution of the combinatorial background in the signal region.

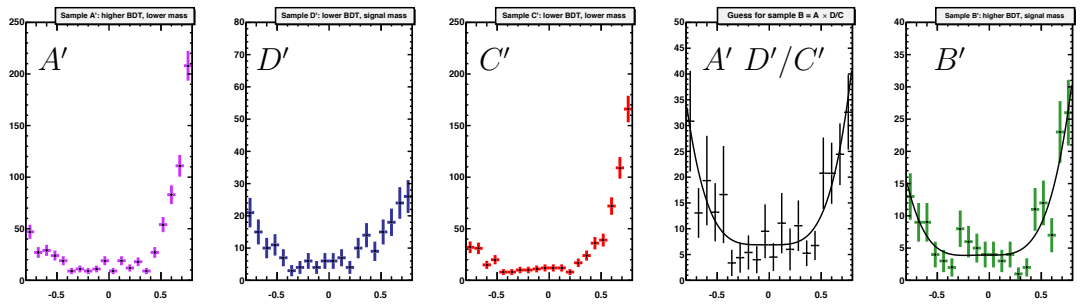


Fig. 4.13: Distributions of the $\cos\theta_\ell$ angle used to validate the method for the extraction of the angular distribution of the combinatorial background. From left to right the first three plots are the 3 samples A', D' and C' defined above for the L0Ele trigger category. Then, the fourth is the guess for the distribution of the sample B' obtained as $A'\ D'/C'$ and the next is the real distribution of data in B'.

4.3.2 Angular modelling of the partially reconstructed background

The partially reconstructed background accounts for about 15% of the events in the angular fit mass window. These events cannot be treated in the same way as the combinatorial ones. Since only one or two particles are not reconstructed, the observed angular distributions retain some of the features induced by the dynamics of the decay. Hence, they are modelled using the same functional shapes as the signal, but with independent physics parameters, $F_{\mathrm{L,PR}}$, $A^{(2)}_{\mathrm{T,PR}}$, $A^{\mathrm{Im}}_{\mathrm{T,PR}}$ and $A^{\mathrm{Re}}_{\mathrm{T,PR}}$.

The loss of one or more final-state hadrons leads to an apparent polarisation of the K^{*0} closer to the random polarization value, corresponding to $F_{\mathrm{L}} = 1/3$. Indeed, while on $B^0 \to K^{*0}\gamma_{e^+e^-}$ simulated events the F_{L} parameter is found to be 0 (as expected, since the real γ has no longitudinal component), it reaches $(17 \pm 4)\%$ for simulated $B \to \gamma K^{**}(\to K\pi X)$ events. For the signal decay $B^0 \to$

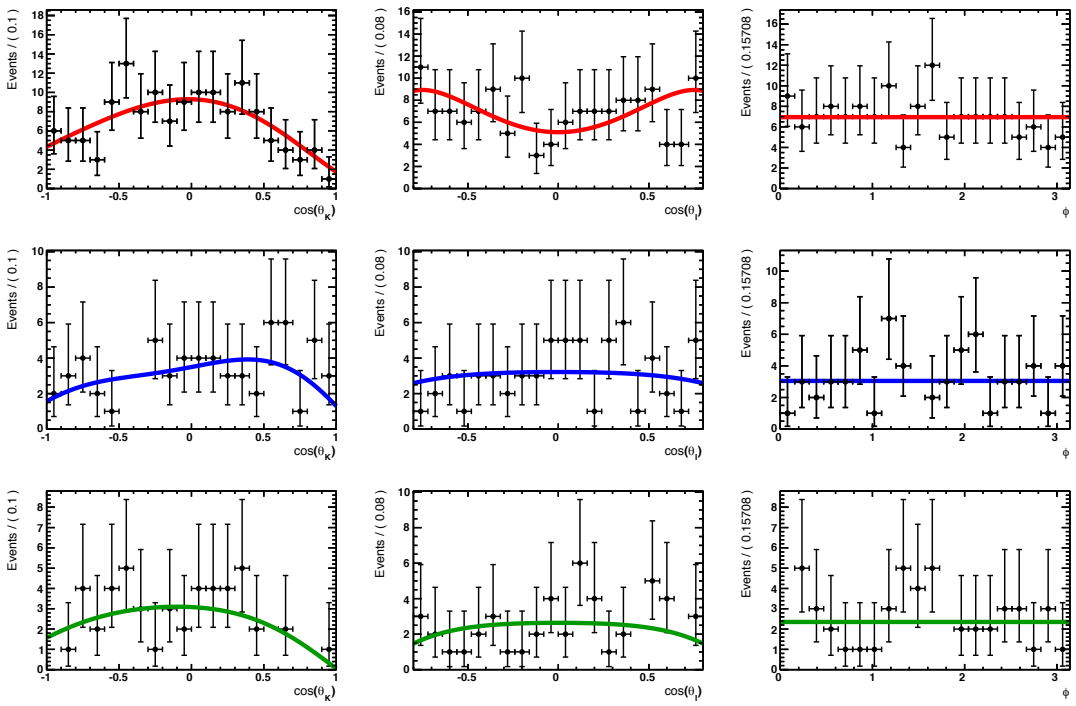


Fig. 4.14: Angular fit to $B^0 \to K^{**}\gamma(ee)$ MC. On the left is the $\cos\theta_K$ angle, on the center $\cos\theta_\ell$ and on the right $\tilde{\phi}$. On the top row is the L0Ele trigger category, on the central one is L0Had and on the bottom is L0TIS.

$K^{*0}e^+e^-$, in the very-low q^2 bin considered, one expects an F_{L} value of the order of 10%. Therefore, the F_{L} value for the partially reconstructed background, $F_{\mathrm{L,PR}}$, is assumed to be equal to 1/3, which is equivalent to no polarisation. This parameter

is varied between 17% and 50% to assess the size of the systematic uncertainty associated with this hypothesis. This value is validated also on a sample of $B \to K_1(1270)e^+e^-$ MC at generator level. Firstly, some cuts on the final state particles p_T are applied as well as the nominal window around the K^{*0} mass, the q^2 range and the cut on $|\cos\theta_\ell| < 0.8$. The same cuts are applied on a sample of $B^0 \to K^{*0}e^+e^-$ MC at generator level in order to compare. Then, the $\cos\theta_K$ distributions of these two samples are fit to extract $F_{\rm L}$ (they are shown in Figure 4.15). The results are $F_{{\rm L}K^{*0}ee} = 0.140 \pm 0.002$ and $F_{{\rm L}K_1e^+e^-} = 0.34 \pm 0.04$, therefore validating the hypothesis.

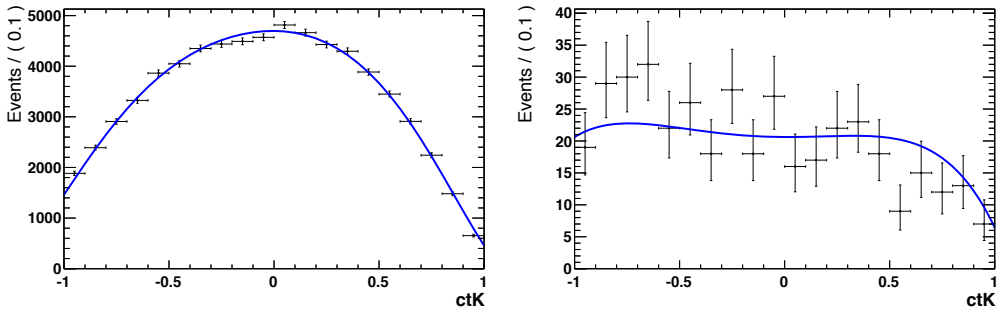


Fig. 4.15: Angular fit of $\cos\theta_K$ for generator level MC samples of $B^0 \to K^{*0}e^+e^-$ (left) and $B^+ \to K_1(1270)e^+e^-$ (right).

Similarly, the loss of information due to the non-reconstructed particles leads to a damping of the transverse asymmetries of the partially reconstructed background, $A^{(2)}_{\rm T,PR}$, $A^{\rm Im}_{\rm T,PR}$ and $A^{\rm Re}_{\rm T,PR}$, compared to those of the signal. The signal transverse asymmetries are expected to be small in the SM, therefore their values are set to zero to describe the angular shape of the PR background. For $A^{(2)}_{\rm T,PR}$ and $A^{\rm Im}_{\rm T,PR}$ the validity of this assumption is tested by comparing angular fits to $B \to J/\psi K^{**}(\to K\pi X)$ and $B^0 \to J/\psi K^{*0}$ simulated events. The angular fits are done without using the information on the mass $m(K^{*0}e^+e^-)$ and taking an angular acceptance correction modelled on a phase space MC sample with a q^2 around the J/ψ resonance. The two fits to the $B^0 \to J/\psi(e^+e^-)K^{*0}$ signal and partially reconstructed background are shown in Figures 4.16 and 4.17 and results are summed up in Table 4.9. They show that the large $\sin(2\tilde{\phi})$ effect (proportional to $(1-F_{\rm L}){\rm A}^{\rm Im}_{\rm T}$) measured on the $B^0 \to J/\psi(e^+e^-)K^{*0}$ signal is damped by a factor

$$D = \frac{(1-F_{\rm L}^{\rm sig}){\rm A}_{\rm T}^{\rm Imsig}}{(1-F_{\rm L}^{\rm PR}){\rm A}_{\rm T}^{\rm Im,part}} = 0.2 \pm 0.3 \qquad (4.12)$$

in the sample of partially reconstructed background.

The systematic uncertainty associated with this assumption is estimated by varying $A^{(2)}_{\rm T,PR}$ and $A^{\rm Im}_{\rm T,PR}$ up to half of the fitted signal values of ${\rm A}^{(2)}_{\rm T}$ and ${\rm A}^{\rm Im}_{\rm T}$, *i.e.*

MC sample	$F_{\rm L}$	$\mathrm{A}_{\rm T}^{(2)}$	$\mathrm{A}_{\rm T}^{\rm Im}$	$\mathrm{A}_{\rm T}^{\rm Re}$
$K^{*0}J/\psi\,(ee)$	0.600 ± 0.006	-0.121 ± 0.032	-0.451 ± 0.032	$(-5.7 \pm 10.3)10^{-3}$
$K^{**}J/\psi\,(ee)$	0.370 ± 0.021	-0.096 ± 0.083	-0.054 ± 0.081	$(-4.8 \pm 30.0)10^{-3}$

Tab. 4.9: Results of the angular fit to the $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ MC and to the $B^0 \to J/\psi\,(ee)K^{**}$ MC

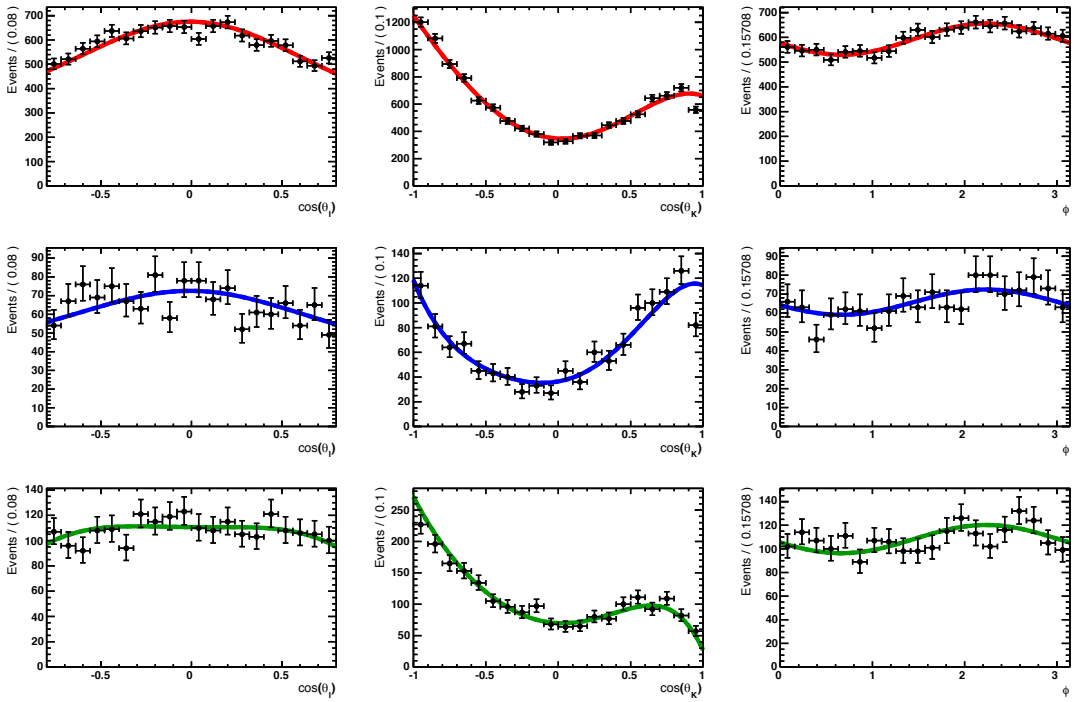


Fig. 4.16: Angular Fit to $B^0 \to J/\psi\,(e^+e^-)K^{*0}$ MC for the three trigger categories. On the left is the $\cos\theta_\ell$ angle, on the center $\cos\theta_K$ and on the right $\tilde{\phi}$. On the top row is the L0Ele trigger category, on the central one is L0Had and on the bottom is L0TIS.

assuming a damping factor of 0.5. For the $A_{\rm T,PR}^{\rm Re}$ parameter, however, one cannot estimate a damping factor with the same method since in the $B \to J/\psi\,K^{*0}$ decay the value of $\mathrm{A}_{\rm T}^{\rm Re}$ is zero. The systematic uncertainty is evaluated by allowing the $A_{\rm T,PR}^{\rm Re}$ parameter to be as high as the $\mathrm{A}_{\rm T}^{\rm Re}$ value obtained from the $B^0 \to K^{*0}e^+e^-$ angular fit. The results in Table 4.9 confirm also the trend of the partially reconstructed background to get values of $F_{\rm L}$ closer to the 1/3.

In summary the angular PDF for the partially reconstructed background reads:

$$P_i^{\rm PR}(\cos\theta_\ell, \cos\theta_K, \tilde{\phi}) = P^{B^0\to K^{*0}e^+e^-}(\cos\theta_\ell, \cos\theta_K, \tilde{\phi}; F_{\rm L,PR}, A^{(2)}_{\rm T,PR}, A^{\rm Im}_{\rm T,PR}, A^{\rm Re}_{\rm T,PR}) \\ \epsilon_i(\cos\theta_\ell, \cos\theta_K, \tilde{\phi}) \quad (4.13)$$

where $P^{B^0\to K^{*0}e^+e^-}$ is in Equation 1.32 and the acceptance function (trigger category dependent) is described in Section 4.2. The value of $F_{\rm L,PR}$ is fixed to 1/3 and the values of transverse asymmetries $A^{(2)}_{\rm T,PR}$, $A^{\rm Im}_{\rm T,PR}$ and $A^{\rm Re}_{\rm T,PR}$ are fixed to 0.

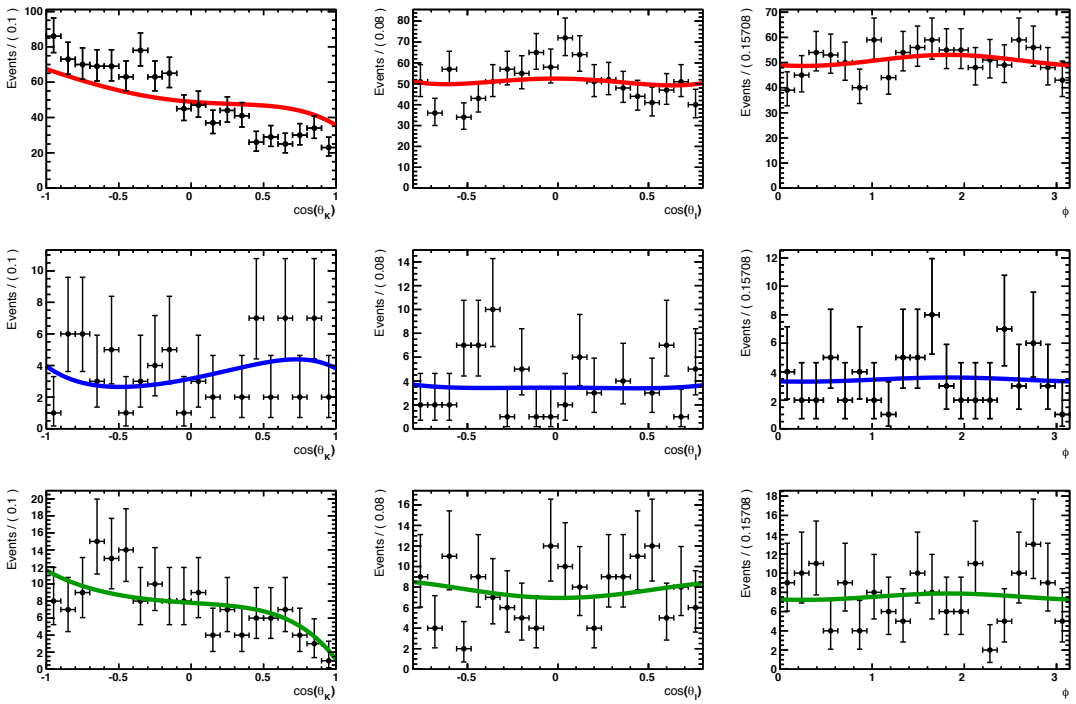


Fig. 4.17: Angular fit to $B^0 \to K^{**}J/\psi(ee)$ MC. On the left is the $\cos\theta_K$ angle, on the center $\cos\theta_\ell$ and on the right $\tilde{\phi}$. On the top row is the L0Ele trigger category, on the central one is L0Had and on the bottom is L0TIS.

4.4 Fit validation

4.4.1 Test of the angular fit on the LHCb MC

In order to further test the angular fit and the modelling of the acceptance the following test is done. A sample is created summing up $B^0 \to K^{*0}e^+e^-$ LHCb MC after the whole selection and toy-MC events for the two types of background. The fractions for the two types of background with respect to the signal yield are fixed to the ones found on data. Then, this sample is fitted with the nominal PDF (as described in Section 4.1) with the only caveat that no MC/Data correction is applied. The four physical parameters $F_{\rm L}$, $\mathrm{A}_{\rm T}^{(2)}$, $\mathrm{A}_{\rm T}^{\rm Im}$ and $\mathrm{A}_{\rm T}^{\rm Re}$ are thus extracted. The results are shown in Table 4.10 and the fit are shown in Figures 4.18, 4.19 and 4.20.
The results are compared to a fit to a large sample of events at generator level of the signal angular PDF (Equation 1.32) not multiplied by the acceptance correction for the three angles. Nonetheless, the q^2 distribution at generator level is corrected with a weight as to reproduce the q^2 distribution after selection. This way physical observables depending on q^2, like $F_{\rm L}$, are integrated correctly and these numbers can be directly compared to the results of the fit to MC events after the whole selection.The generator-level fit is shown in Figure 4.21 and the results are in Table 4.10.
Fitted values after the whole selection are well compatible with the ones fitted at generator level. This proves that the fitting PDF does not introduce any bias on the values of the physical parameters extracted. Also, this test may be affected by the non-factorization of the acceptance correction. It is therefore one more proof that the acceptance can be factorized in the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$.

	$F_{\rm L}$	$\mathrm{A}_{\rm T}^{(2)}$	$\mathrm{A}_{\rm T}^{\rm Im}$	$\mathrm{A}_{\rm T}^{\rm Re}$
MC gen-level	0.151 ± 0.002	0.007 ± 0.008	0.013 ± 0.008	-0.056 ± 0.006
MC after reco/sel	0.149 ± 0.008	0.028 ± 0.030	0.020 ± 0.030	-0.053 ± 0.021

Tab. 4.10: Results of the fit to the MC at generator level and after the whole selection. Even with a yield of $\sim 10^4$ signal events the reconstruction and selection processes do not introduce significant biases (in the MC).

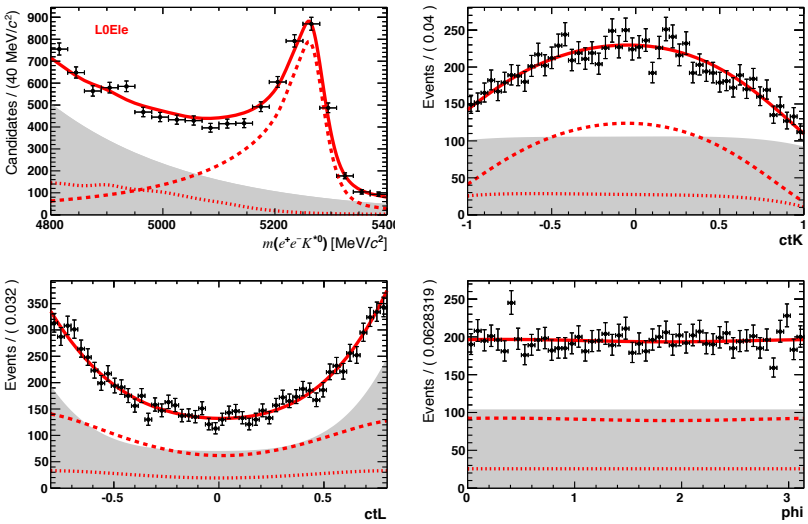


Fig. 4.18: L0Ele trigger category. LHCb MC for the signal is summed to distributions for the two types of background generated with toy-MC, the fractions being fixed to the ones fitted to data. The solid line is the fit to the total angular PDF, the dashed line is the signal PDF, the dotted line the partially reconstructed background and the light grey area is the combinatorial background.

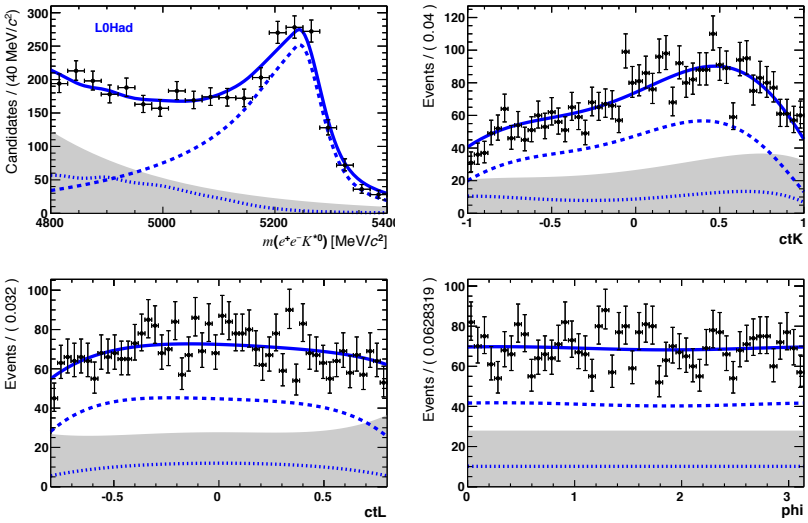


Fig. 4.19: L0Had trigger category. LHCb MC for the signal is summed to distributions for the two types of background generated with toy-MC, the fractions being fixed to the ones fitted to data. The solid line is the fit to the total angular PDF, the dashed line is the signal PDF, the dotted line the partially reconstructed background and the light grey area is the combinatorial background).

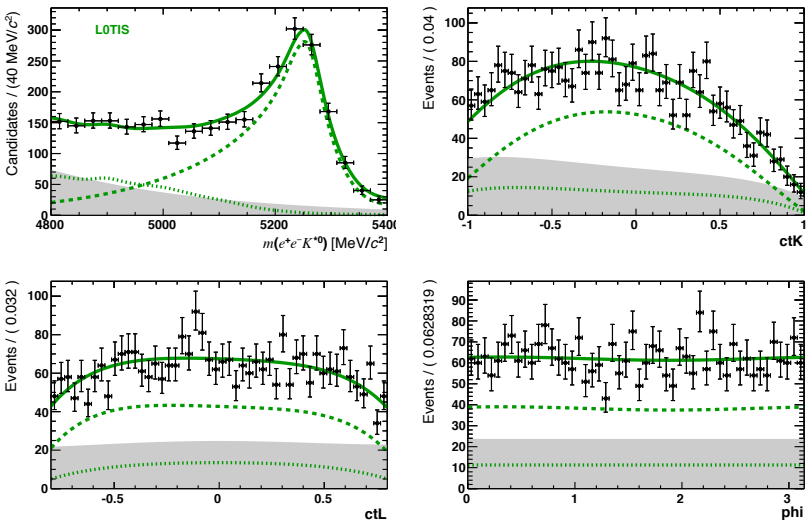


Fig. 4.20: L0TIS trigger category. LHCb MC for the signal is summed to distributions for the two types of background generated with toy-MC, the fractions being fixed to the ones fitted to data. The solid line is the fit to the total angular PDF, the dashed line is the signal PDF, the dotted line the partially reconstructed background and the light grey area is the combinatorial background.

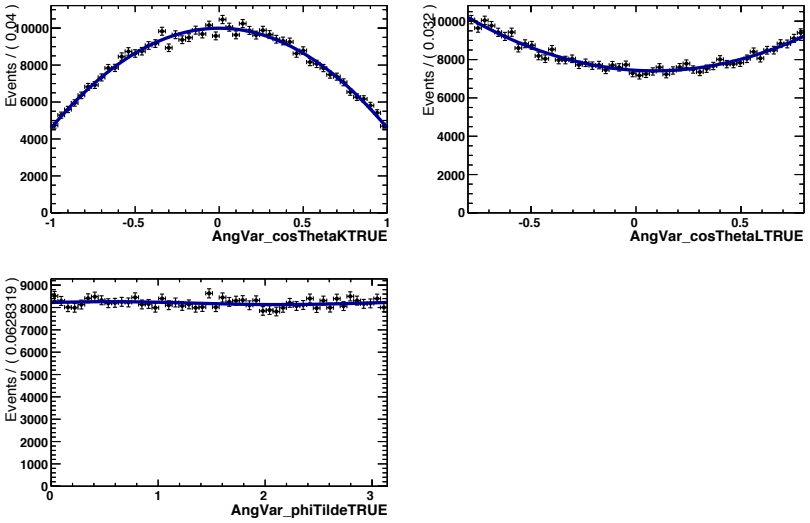


Fig. 4.21: Generator level distributions for the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$. The q^2 distribution is weighted with the acceptance, while the angular distributions are not.

4.4.2 Fit stability with toy-MC

The full fit was checked using toy Monte-Carlo experiments. Signal and background events are generated according to the mass and angular PDFs described in Sections 3.4 and 4.1. Each toy experiment is generated with the same statistics and background fractions as data (in Table 4.1). The signal is generated with SM-inspired values for the four observables: $F_{\mathrm{L}} = 0.18$, $\mathrm{A}_{\mathrm{T}}^{\mathrm{Re}} = -0.06$ and $\mathrm{A}_{\mathrm{T}}^{(2)} = \mathrm{A}_{\mathrm{T}}^{\mathrm{Im}} = 0$. Two examples of the likelihood of two different toy MC are shown in Figure 4.22. The corresponding results are :

$$\begin{aligned} \mathrm{A}_{\mathrm{T}}^{(2)} &= 0.12 \pm 0.22 & \mathrm{A}_{\mathrm{T}}^{(2)} &= -0.31 \pm 0.31 \\ \mathrm{A}_{\mathrm{T}}^{\mathrm{Im}} &= 0.16 \pm 0.21 & \mathrm{A}_{\mathrm{T}}^{\mathrm{Im}} &= 0.19 \pm 0.28 \\ \mathrm{A}_{\mathrm{T}}^{\mathrm{Re}} &= -0.31 \pm 0.15 & \mathrm{A}_{\mathrm{T}}^{\mathrm{Re}} &= 0.09 \pm 0.20 \\ F_{\mathrm{L}} &= 0.05 \pm 0.06 & F_{\mathrm{L}} &= 0.31 \pm 0.07 \end{aligned}$$

Then, five thousands toy experiments are generated and each of them is fitted to the nominal PDF to measure the four observables and check if they correspond to the generated ones. The distributions of the their uncertainties as calculated from the fit are shown in Figure 4.23. The central values for the statistical uncertainties are in Table 4.11.

	F_{L}	$\mathrm{A}_{\mathrm{T}}^{(2)}$	$\mathrm{A}_{\mathrm{T}}^{\mathrm{Im}}$	$\mathrm{A}_{\mathrm{T}}^{\mathrm{Re}}$
σ^{stat}	0.07	0.25	0.25	0.17

Tab. 4.11: Statistical uncertainties on the four physical observables.

The proper behaviour of the fit is checked using the pull distributions. They are fitted to gaussian functions, giving σ-width values in agreement with one (fits and results are in Figure 4.24). The central values of the gaussians are generally either in agreement with zero, or reasonably close to this value. The main shifts are observed on the F_{L} and $\mathrm{A}_{\mathrm{T}}^{\mathrm{Re}}$ parameters which are slightly biased towards smaller values. This effect is thus further investigated. First of all, it appears to be caused by a small left tail of the F_{L} pull distribution rather than a general shift of it. Furthermore, it is found that this effect disappears in toys generated with much larger samples. Therefore, this effect appears to be caused by an asymmetric likelihood function, the asymmetry showing up just at low statistics as the large uncertainty let the parameters have a broader spectrum of values. This effect is very small compared to the statistical uncertainties but it is nevertheless included in the systematic error under the label "fit procedure".

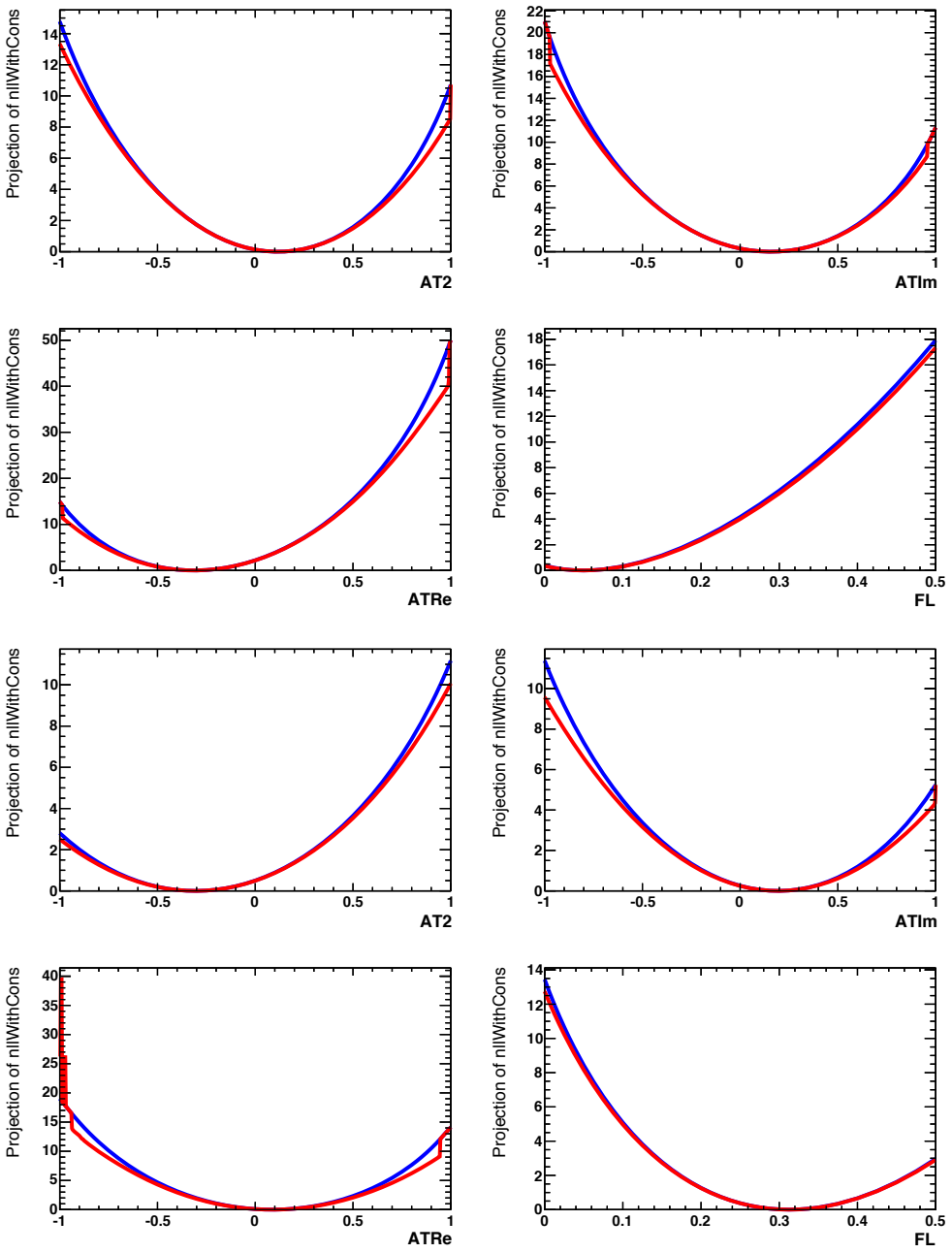


Fig. 4.22: Scan of the likelihood (in blue) and profile likelihood (in red) for the four observables respectively $A_T^{(2)}$, A_T^{Im} and A_T^{Re} and F_L. The first four plots correspond to the first set of results and the four last ones to the second set of results

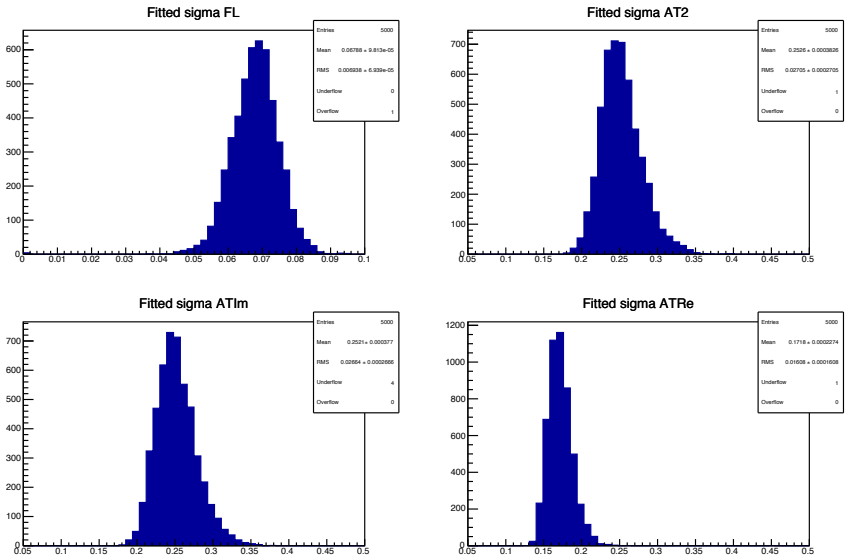


Fig. 4.23: The distributions of the uncertainties for the four parameters of interest : F_{L}, $\mathrm{A}_{\mathrm{T}}^{(2)}$ $\mathrm{A}_{\mathrm{T}}^{\mathrm{Im}}$ and $\mathrm{A}_{\mathrm{T}}^{\mathrm{Re}}$. The generated values are respectively $F_{\mathrm{L}} = 0.18$, $\mathrm{A}_{\mathrm{T}}^{(2)} = 0$, $\mathrm{A}_{\mathrm{T}}^{\mathrm{Im}} = 0$ and $\mathrm{A}_{\mathrm{T}}^{\mathrm{Re}} = -0.06$.

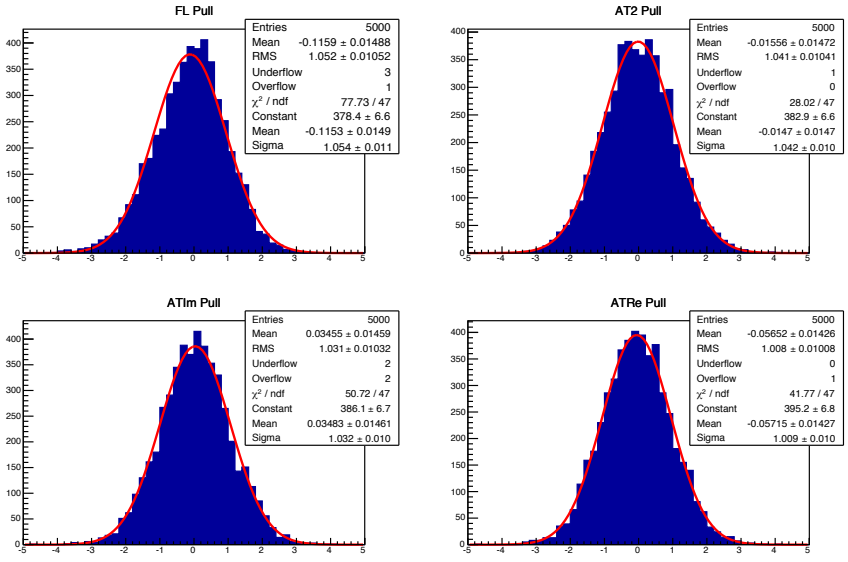


Fig. 4.24: Pull distributions of a toy-MC generated with SM values of the physical parameters of interest and nominal modelling of acceptance and backgrounds. The four distributions are fitted to a gaussian whose results are shown in the legend.

4.4.3 Check for various physics parameters values

In the toys Monte-Carlo the values of the generated physics parameters are changed and the fitted parameters can be compared with the generated ones. There is no sign of deviation from the expected values, as can be seen in Figure 4.25

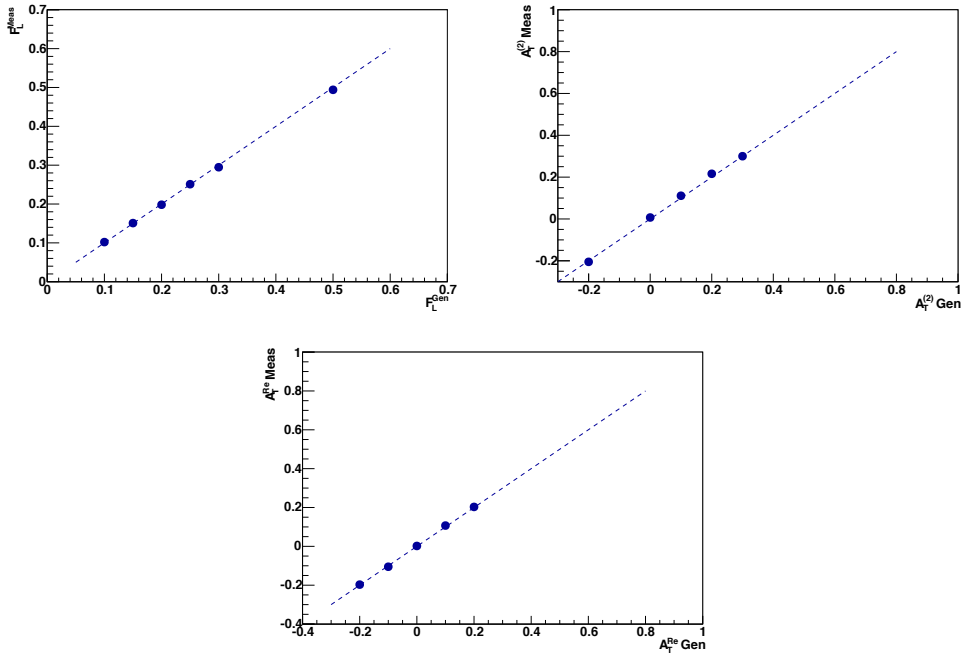


Fig. 4.25: Comparison of the fitted and generated values for various physics parameters. The mean and uncertainty on the fitted value are obtained from 1000 toys experiments and are of the order of 0.002 for F_{L}, 0.008 for $\mathrm{A}_{\mathrm{T}}^{(2)}$ and 0.006 for $\mathrm{A}_{\mathrm{T}}^{\mathrm{Re}}$.

4.5 Systematic uncertainties

Systematic uncertainties are estimated using toy-MCs (unless explicitly stated). The physics values are SM ones : $F_{\rm L}$ = 0.18 , $A_{\rm T}^{\rm Re}$ = -0.06 , $A_{\rm T}^{(2)}$ = 0 and $A_{\rm T}^{\rm Im}$ = 0 . For this evaluation a given parameter (for example one of the parameter describing the acceptance) is moved one sigma up and down. Given the smallness of most of these effects, a too large number of toys having the same signal yield as real data should be generated to measure accurately enough their effect, i.e. to make statistical fluctuations negligible compared to real effects. Therefore, for each of these variations one single toy with a statistics 10^4 times larger than the data sample is generated. Then, each of the toy is fitted to the nominal PDF and shifts on the fitted parameters of interest are extracted. Up or down variations of each parameter type are done at the same time for the three trigger categories to be conservative about the possible presence of category-independent effects. Moving a parameter one sigma up and down resulted always in symmetric shifts (i.e. all asymmetries are found to be not significant). Then, each up and down shift is associated to an uncertainty and they are summed up quadratically to get the total uncertainties corresponding to each source of systematics. Correlations between parameters of the background modelling are neglected. Correlations between parameters modelling the angular acceptance (Appendix B.2) are in general quite small and furthermore, the larger correlations found are all negative. Therefore, they are neglected too, doing a safe pessimistic choice.

Note that the systematic uncertainty on $F_{\rm L}$ is mainly driven by uncertainties in symmetric terms in $\cos\theta_K$ and in $\cos\theta_\ell$ (in both the acceptance and the combinatorial background modelling). Also the uncertainty in the partially reconstructed background has a significant impact on it. The uncertainties on $A_{\rm T}^{(2)}$ and $A_{\rm T}^{\rm Im}$ on the other hand are almost independent of the modelling of $\cos\theta_K$ and $\cos\theta_\ell$ distributions, their uncertainty being driven only by the $\tilde{\phi}$ modelling. Furthermore, as expected, the uncertainty on $A_{\rm T}^{\rm Re}$ depends mainly on asymmetric terms in $\cos\theta_\ell$.

4.5.1 Estimation of the systematics due to the acceptance

For $\cos\theta_\ell$ and $\cos\theta_K$ the acceptances have been fitted on phase space Monte-Carlo reweighed to match the measured q^2 spectrum. The statistical uncertainties on the fitted parameters are used. For $\cos\theta_\ell$ a problem would be the presence of an asymmetric term in the acceptance function which would bias the extraction of $A_{\rm T}^{\rm Re}$. In principle there is no strong reason why such a term should exist. Nevertheless the $\cos\theta_\ell$ acceptance is also modelled allowing for a linear term in $\cos\theta_\ell$ (see Section 4.2). No deviation from 0 is observed, but the uncertainty on

this term is used to assess the systematic error.
For $\tilde{\phi}$ it is hard to find a reason for a dangerous $\cos 2\tilde{\phi}$ or $\sin 2\tilde{\phi}$ modulation and indeed no such modulations are observed in the fits of the three trigger categories. Therefore, to estimate the systematic uncertainties, the three trigger categories are summed up and a fit is performed allowing only for $\cos 2\tilde{\phi}$ or $\sin 2\tilde{\phi}$ terms. The different variations tested and the resulting systematics uncertainties on the four angular observables are given in Table 4.12.

	$\delta F_{\mathrm{L}} \times 10^{-2}$	$\delta \mathrm{A}_{\mathrm{T}}^{(2)} \times 10^{-2}$	$\delta \mathrm{A}_{\mathrm{T}}^{\mathrm{Im}} \times 10^{-2}$	$\delta \mathrm{A}_{\mathrm{T}}^{\mathrm{Re}} \times 10^{-2}$
$c_1^K - \sigma$	0.07 ± 0.07	-0.20 ± 0.25	0.00 ± 0.25	-0.19 ± 0.17
$c_1^K + \sigma$	-0.03 ± 0.07	0.46 ± 0.25	-0.23 ± 0.25	-0.23 ± 0.17
$c_2^K - \sigma$	-1.21 ± 0.07	0.13 ± 0.24	0.14 ± 0.24	-0.25 ± 0.16
$c_2^K + \sigma$	1.24 ± 0.07	-0.08 ± 0.25	0.24 ± 0.25	-0.12 ± 0.17
$c_3^K - \sigma$	0.02 ± 0.07	-0.02 ± 0.25	0.06 ± 0.25	-0.19 ± 0.17
$c_3^K + \sigma$	-0.10 ± 0.07	-0.10 ± 0.25	0.18 ± 0.25	-0.12 ± 0.17
$c_4^K - \sigma$	0.03 ± 0.07	-0.31 ± 0.25	0.02 ± 0.25	0.20 ± 0.17
$c_4^K + \sigma$	-0.15 ± 0.07	-0.37 ± 0.25	-0.04 ± 0.25	0.10 ± 0.17
$c_2^\ell - \sigma$	0.40 ± 0.07	0.16 ± 0.25	0.02 ± 0.25	0.24 ± 0.17
$c_2^\ell + \sigma$	-0.34 ± 0.07	0.22 ± 0.25	0.28 ± 0.25	-0.01 ± 0.17
$c_4^\ell - \sigma$	-0.10 ± 0.07	0.16 ± 0.25	-0.19 ± 0.25	-0.02 ± 0.17
$c_4^\ell + \sigma$	-0.06 ± 0.07	-0.60 ± 0.25	0.06 ± 0.25	-0.04 ± 0.17
$0 - \sigma(c_1^\ell)$	-0.07 ± 0.07	0.45 ± 0.25	0.02 ± 0.25	-3.23 ± 0.17
$0 + \sigma(c_1^\ell)$	0.03 ± 0.07	-0.19 ± 0.25	0.01 ± 0.25	2.84 ± 0.17
$0 - \sigma_{c_c}$	0.02 ± 0.07	-3.63 ± 0.25	0.41 ± 0.25	0.13 ± 0.17
$0 + \sigma_{c_c}$	-0.10 ± 0.07	3.89 ± 0.25	0.51 ± 0.25	0.01 ± 0.17
$0 - \sigma_{c_s}$	-0.03 ± 0.07	0.41 ± 0.25	-3.51 ± 0.25	0.03 ± 0.17
$0 + \sigma_{c_s}$	0.04 ± 0.07	0.02 ± 0.25	3.39 ± 0.25	-0.25 ± 0.17
Total	1.29	3.84	3.49	3.06

Tab. 4.12: Systematic uncertainties as extracted from MC toys generated with different shapes of the angular acceptance.

4.5.2 Estimation of the systematics due to the combinatorial background

The uncertainties on the parameters fitted in Section 4.3 are used to estimate the systematic uncertainties due to non-precise knowledge of the shape of the combinatorial background. The values of the parameters tested as well as the resulting systematics uncertainties are given in Table 4.13. Since the fractions are obtained directly from data, no systematic uncertainty is assigned to the fraction.

	$\delta F_{\mathrm{L}} \times 10^{-2}$	$\delta \mathrm{A}_{\mathrm{T}}^{(2)} \times 10^{-2}$	$\delta \mathrm{A}_{\mathrm{T}}^{\mathrm{Im}} \times 10^{-2}$	$\delta \mathrm{A}_{\mathrm{T}}^{\mathrm{Re}} \times 10^{-2}$
$a_1^K - \sigma$	0.18 ± 0.07	-0.70 ± 0.25	0.42 ± 0.25	0.00 ± 0.17
$a_1^K + \sigma$	-0.01 ± 0.07	-0.08 ± 0.25	-0.05 ± 0.25	-0.35 ± 0.17
$a_3^\ell - \sigma$	0.05 ± 0.07	0.06 ± 0.25	0.19 ± 0.25	-3.48 ± 0.17
$a_3^\ell + \sigma$	0.12 ± 0.07	-0.52 ± 0.25	0.14 ± 0.25	3.83 ± 0.17
$a_4^\ell - \sigma$	0.72 ± 0.07	-0.17 ± 0.24	0.41 ± 0.24	0.91 ± 0.17
$a_4^\ell + \sigma$	-0.33 ± 0.07	-0.33 ± 0.25	0.40 ± 0.25	-1.06 ± 0.17
$-\sigma_{a_c}$	0.10 ± 0.07	-3.23 ± 0.25	0.25 ± 0.25	0.07 ± 0.17
$+\sigma_{a_c}$	0.12 ± 0.07	2.64 ± 0.25	0.12 ± 0.25	0.04 ± 0.17
$-\sigma_{a_s}$	0.08 ± 0.07	-0.19 ± 0.25	-3.01 ± 0.25	0.15 ± 0.17
$+\sigma_{a_s}$	0.04 ± 0.07	-0.26 ± 0.25	2.75 ± 0.25	-0.13 ± 0.17
Total	0.56	2.99	2.93	3.79

Tab. 4.13: Systematic uncertainties as extracted from MC toys generated with different shapes of the combinatorial background angular distributions.

4.5.3 Estimation of the systematics due to the partially reconstructed background

The angular distribution of this background is described by a physics function identical to the one used for the signal but with different physics parameters (Equation 4.13). The systematic error related to $F_{\mathrm{L,PR}}$ is evaluated by varying its value between $\min F_{\mathrm{L,PR}} = 0.17$ and $\max F_{\mathrm{L,PR}} = 0.5$ as explained in Section 4.3.2. The values of the parameters tested as well as the resulting systematics uncertainties are given in Table 4.14. As expected changing $F_{\mathrm{L,PR}}$ has no effect on $\mathrm{A}_{\mathrm{T}}^{(2)}$, $\mathrm{A}_{\mathrm{T}}^{\mathrm{Im}}$

and A_T^{Re}.
Taking $A_{T,PR}^{(2)}$ and $A_{T,PR}^{Im}$ values by damping the signal $A_T^{(2)}$ and A_T^{Im} values by a factor $D = 0.2 \pm 0.3$ results roughly in a bias of $(2.5 \pm 2.5)\%$ of the fitted values of $A_T^{(2)}$ and A_T^{Im}, thus very small. To be conservative the related systematic uncertainty is taken as 5% of the fitted values of $A_T^{(2)}$ and A_T^{Im} (and it is symmetrized). The value of $A_{T,PR}^{Re}$ is also expected to be damped with respect to the fitted value A_T^{Re}, but the damping factor is unknown. To be conservative in the evaluation of the related systematic error, $A_{T,PR}^{Re}$ is thus fixed to the fitted value of A_T^{Re} in order to evaluate the systematic. The size of the systematic uncertainties are summarized in Table 4.14.

	$\delta F_L \times 10^{-2}$	$\delta A_T^{(2)} \times 10^{-2}$	$\delta A_T^{Im} \times 10^{-2}$	$\delta A_T^{Re} \times 10^{-2}$
$F_{L,PR} = 0.50$	-1.77 ± 0.07	0.03 ± 0.24	0.26 ± 0.24	-0.26 ± 0.16
$F_{L,PR} = 0.17$	2.02 ± 0.07	-0.12 ± 0.25	-0.16 ± 0.25	-0.02 ± 0.17
Total	1.90	0.08	0.21	0.14

Tab. 4.14: Systematic uncertainties as extracted from MC toys generated with different shapes of the partially reconstructed background angular distributions.

4.5.4 Systematic effect related to the $B^0 \to K^{*0}\gamma_{e^+e^-}$ contamination

The contamination of $B^0 \to K^{*0}\gamma_{e^+e^-}$ events is expected to be $(3.8 \pm 1.9)\%$ of the signal yield (see Section 3.3.2). It peaks in the B mass, but its angular distribution is expected to be flat in $\tilde{\phi}$ because the angular information on the polarization gets lost in the creation of the di-electron pair for $m(ee)$ masses above $20\,\mathrm{MeV}/c^2$, as discussed in Section 1.3.4. Therefore, to correct the fitted values of $A_T^{(2)}$ and A_T^{Im}, they have to be multiplied by a small factor $(1 - 0.038)^{-1}$. The same applies to A_T^{Re} and F_L as they are also 0 in $B^0 \to K^{*0}\gamma_{e^+e^-}$. Then, the systematic uncertainty on the corrected value corresponds to the uncertainty on the $B^0 \to K^{*0}\gamma_{e^+e^-}$ contamination. It is therefore $\sim 2\%$ of the fitted value (very small).

4.5.5 Systematic uncertainty related to interference effects with $B^0 \to K^{*0} V(\to e^+e^-)$ decays with $V = \rho, \omega, \phi$

One should also take into account possible interference effects of the vector resonances with the continuum as in principle, they contribute to the angular observables $F_{\rm L}$, $A_{\rm T}^{\rm Re}$, $A_{\rm T}^{(2)}$ and $A_{\rm T}^{\rm Im}$. Interference terms are taken into account by [33,138] when calculating the values angular observables as a function of q^2. When integrated over q^2 the effect completely vanishes for ρ and ω. This is not the case for the ϕ resonance: its interference effect on the angular observables does not get evenly integrated out as the efficiency is smaller for q^2 values above the resonance. Nevertheless, the effect is found to be completely negligible for all the four angular observables when compared to systematic uncertainties.

4.5.6 Fit statistics

In order to assess any effect due to the fit and taking into account the current statistics a sample of 5000 toy-MCs with the current statistics has been generated and fitted with the default PDFs. The pulls distributions in Section 4.4.2 show that the $F_{\rm L}$ and $A_{\rm T}^{\rm Re}$ parameters are exhibiting a small bias ($\sim 10\times$ smaller than the statistical error). The fit results are therefore corrected by $F_{\rm L} = F_{\rm L}^{\rm fit} + 0.008$ and $A_{\rm T}^{\rm Re} = A_{\rm T}^{\rm Re,fit} + 0.010$. These shifts are taken as a systematic error to be conservative and are shown in the line "Fit procedure" of Table 4.15.

4.5.7 Summary of systematic uncertainties

Results are summarized in Table 4.15. All systematic errors are small compared to the expected statistical errors.

Source	$\sigma(F_{\mathrm{L}})$	$\sigma(\mathrm{A}_{\mathrm{T}}^{(2)})$	$\sigma(\mathrm{A}_{\mathrm{T}}^{\mathrm{Im}})$	$\sigma(\mathrm{A}_{\mathrm{T}}^{\mathrm{Re}})$
Acceptance modelling	0.013	0.038	0.035	0.031
Combinatorial background	0.006	0.030	0.029	0.038
Partially reconstructed	0.019	$0.05 \times \mathrm{A}_{\mathrm{T}}^{(2)\,\mathrm{fit}}$	$0.05 \times \mathrm{A}_{\mathrm{T}}^{\mathrm{Im,fit}}$	$0.10 \times \mathrm{A}_{\mathrm{T}}^{\mathrm{Re,fit}}$
$B^0 \to K^{*0}\gamma_{e^+e^-}$ contamination	$0.02 \times F_{\mathrm{L}}^{\mathrm{fit}}$	$0.02 \times \mathrm{A}_{\mathrm{T}}^{(2)\,\mathrm{fit}}$	$0.02 \times \mathrm{A}_{\mathrm{T}}^{\mathrm{Im,fit}}$	$0.02 \times \mathrm{A}_{\mathrm{T}}^{\mathrm{Re,fit}}$
Fit procedure	0.008	~ 0	~ 0	0.010
total σ_{syst}	0.025	0.05	0.05	0.05
expected σ_{stat}	0.068	0.25	0.25	0.17

Tab. 4.15: Summary of systematic uncertainties classified by their source. The total systematic error is calculated here assuming SM values for the F_{L}, $\mathrm{A}_{\mathrm{T}}^{(2)}$, $\mathrm{A}_{\mathrm{T}}^{\mathrm{Im}}$, $\mathrm{A}_{\mathrm{T}}^{\mathrm{Re}}$ (which are needed in order to compute the effect of the $B^0 \to K^{*0}\gamma_{e^+e^-}$ contamination). In the last line is the expected statistical error (from Table 4.11). The comparison shows that the measurement is mainly limited by statistics.

4.6 Results of the full fit on $B^0 \to K^{*0}e^+e^-$ data

The results of the fit to $B^0 \to K^{*0}e^+e^-$ data is shown in Figure 4.26. Fitted values are corrected using the known contamination of $B^0 \to K^{*0}\gamma_{e^+e^-}$ events (a very small $\sim 4\%$ correction). Also, they are used to compute the values of the systematic errors related to the partially reconstructed background and the $B^0 \to K^{*0}\gamma_{e^+e^-}$ contamination as shown in Table 4.15. The corrected values and their associated errors (statistical placed first and systematic placed second) are in Equation 4.14:

$$\begin{aligned} F_{\rm L} &= 0.16 \pm 0.06 \pm 0.03 \\ {\rm A}_{\rm T}^{(2)} &= -0.23 \pm 0.23 \pm 0.05 \\ {\rm A}_{\rm T}^{\rm Im} &= 0.14 \pm 0.22 \pm 0.05 \\ {\rm A}_{\rm T}^{\rm Re} &= 0.10 \pm 0.18 \pm 0.05 \end{aligned} \tag{4.14}$$

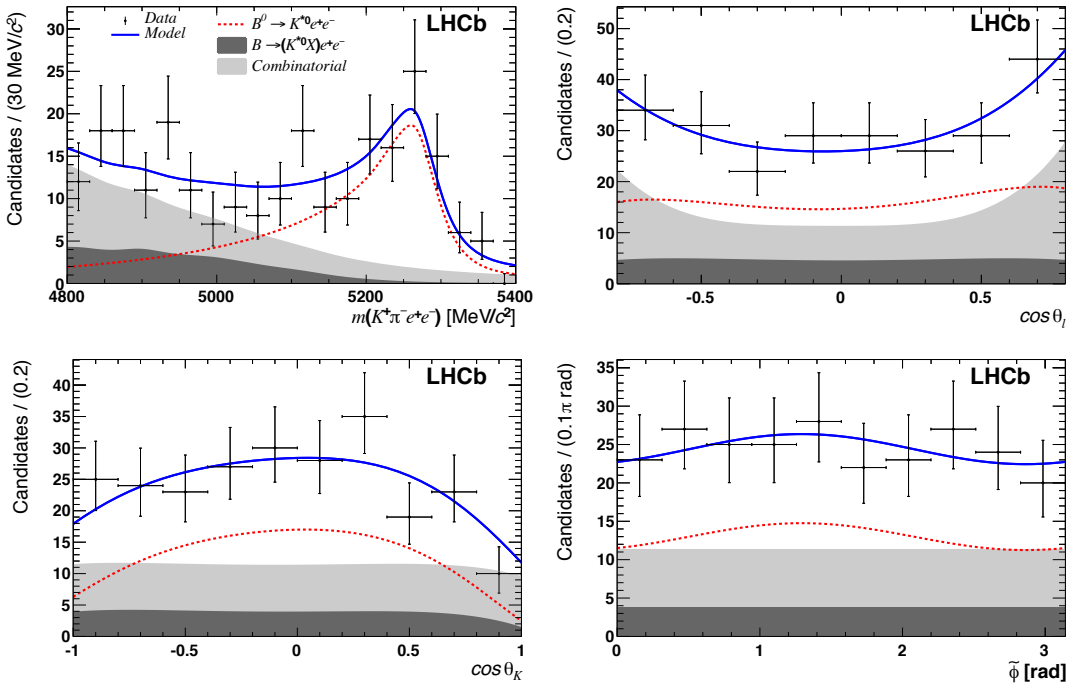


Fig. 4.26: Full fit of $B^0 \to K^{*0}e^+e^-$ data. The fit is unbinned and simultaneous on $m(e^+e^-K^{*0})$ and the three angles $\cos\theta_\ell$, $\cos\theta_K$, $\tilde{\phi}$. The results of the fit are in Equation 4.14.

4.6.1 q^2 region as input to theory

In order to compare the results obtained to predictions from theory one cannot rely solely on the reconstructed q^2 window considered ($[0.0004, 1]\,\mathrm{GeV}^2/c^4$) since the efficiency changes considerably as a function of q^2 and the effect of bremsstrahlung on the dilepton mass resolution is significant. Two different inputs are thus given in the following to do theory predictions in the correct q^2 range: the average q^2 value as extracted from data and the limits of an effective q^2 bin that can be used with the hypothesis of flat q^2 efficiency to integrate q^2 predictions for the observables.

Average q^2 value

The average reconstructed q^2 value is extracted from an *sPlot* of data. Then, it is corrected for reconstruction effects to recover the expected true average q^2.
In the MC, selected events have an average reconstructed $< q^2 >_{\mathrm{reco}} = 0.190\,\mathrm{GeV}^2/c^4$ while the corresponding average true $< q^2 >_{\mathrm{true}}$ is $0.238\,\mathrm{GeV}^2/c^4$. Therefore, the average q^2 found on data is roughly corrected by the ratio $0.238/0.190 = 1.25$ (this ratio is very similar among the three trigger categories: 1.24 in L0Ele, 1.25 in L0Had and 1.26 in L0TIS). This correction takes into account both efficiency and dilepton mass migration effects.
In order to extract the average $< q^2 >_{\mathrm{reco}}$ on $B^0 \to K^{*0}e^+e^-$ data the *sPlot* technique is used. A mass fit to the $B^0 \to K^{*0}e^+e^-$ data like the one described in Section 3.4.4 is done leaving the yields free but constraining all signal and

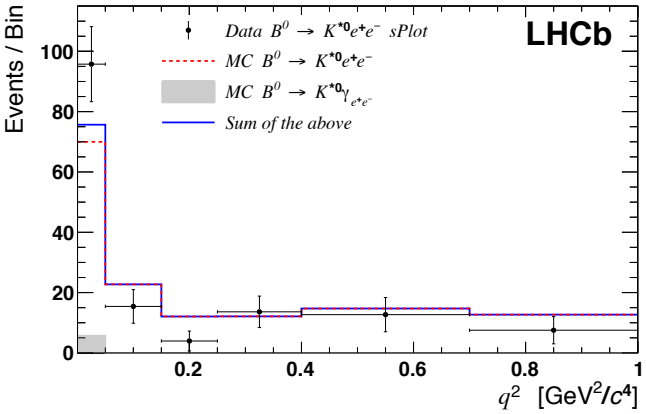


Fig. 4.27: Distribution of reconstructed $q^2 = m^2(e^+e^-)$ from an *sPlot* of data (black points). Overlaid is the MC distributions of $B^0 \to K^{*0}e^+e^-$ (dashed red line), the 3.8% $B^0 \to K^{*0}\gamma_{e^+e^-}$ contamination (light grey area) and the sum of the two (solid blue line).

background PDFs. The resulting reconstructed-q^2 distribution is shown in Figure 4.27. Data points are overlaid by the MC distributions for $B^0 \to K^{*0} e^+ e^-$ and $B^0 \to K^{*0} \gamma_{e^+e^-}$. The two curves are in good agreement with a $\chi^2/\text{ndof} = 0.76$. The photon pole at the low-q^2 end is clearly visible in data. The resulting average q^2 value is then computed from data and corrected by a factor $(1 - 0.038)^{-1}$ to account for an expected $B^0 \to K^{*0} \gamma_{e^+e^-}$ contamination of 3.8%. The result is $< q^2 >_{\text{reco, sPlot}} = (0.14 \pm 0.03)\,\text{GeV}^2/c^4$ corresponding to an expected value of the true average q^2 of:

$$< q^2 >= (0.17 \pm 0.04)\,\text{GeV}^2/c^4. \tag{4.15}$$

Effective q^2 bin

The average q^2-value is a good input for a theoretical estimation of observables as long as their behaviour can be approximated as being linear with respect to q^2 in the range considered, which is not always the case. Therefore, the boundaries of an effective q^2_{true}-bin are evaluated. This range may be used for integrating observables over q^2 while neglecting acceptance effects. This approach is a good approximation as the efficiency is rather flat in a large domain, except close to the reconstructed-q^2 limits, 0.0004 and 1 GeV^2/c^4.
The efficiency is thus assumed to be constant in the q^2_{true} region between 0.1 and 0.8 GeV^2/c^4, see Figure 4.28. Then, an effective lower q^2_{true}-limit is set so that the fraction of events that would be selected in the region below $0.1\,\text{GeV}^2/c^4$ assuming a flat q^2_{true}-acceptance is the same as the events selected in the MC having the detailed acceptance shape. Given the acceptance is falling down at low q^2_{true} the effective lower q^2_{true}-limit is set to a higher value (at $0.0020\,\text{GeV}^2/c^4$), compared to the reconstructed one (which is at $0.0004\,\text{GeV}^2/c^4$).
The same method is applied to set the effective top-q^2_{true}-limit. Due to bremsstrahlung, events with q^2_{true} higher than $1\,\text{GeV}^2/c^4$ are selected and this results in an effective top-q^2_{true} limit higher than $1\,\text{GeV}^2/c^4$ (namely at $1.12\,\text{GeV}^2/c^4$). As a crosscheck, it was verified that a sample of $B^0 \to K^{*0} e^+ e^-$ events at generator level being in this effective q^2_{true}-bin has a very similar average q^2_{true} value as the sample of selected events in the MC having the detailed q^2_{true} acceptance shape.
The uncertainties on the values set for the effective q^2 limits are set empirically to half of the correction applied. The values for the effective q^2_{true} limits thus read:

$$\begin{aligned} q^2_{\text{min}} &= 0.0020 \pm 0.0008\,\text{GeV}^2/c^4 \\ q^2_{\text{max}} &= 1.12 \pm 0.06\,\text{GeV}^2/c^4 \end{aligned} \tag{4.16}$$

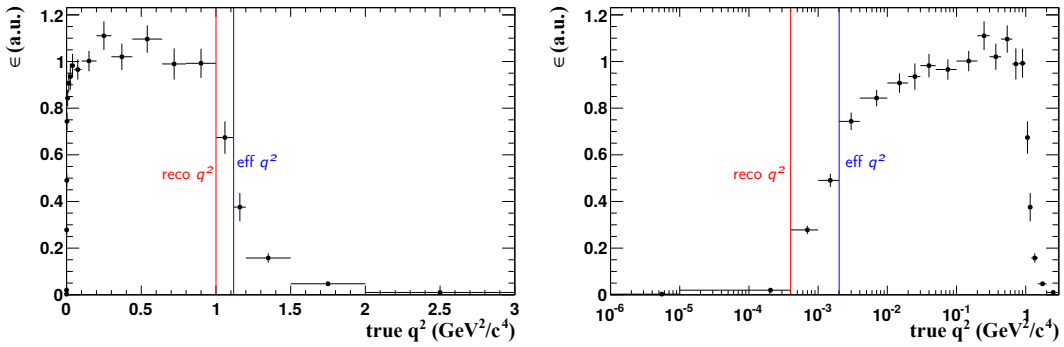


Fig. 4.28: Efficiency for the selection of $B^0 \to K^{*0}e^+e^-$ as a function of q^2_{true}. The same curve is shown in linear scale on the left and in logarithmic scale on the right. The reconstructed-q^2 is required to lie between $0.0004\,\text{GeV}^2/c^4$ and $1\,\text{GeV}^2/c^4$ (boundaries are shown in red). Due to the experimental q^2 resolution, this requirement corresponds to effective-q^2 limits between $0.0020\,\text{GeV}^2/c^4$ and $1.12\,\text{GeV}^2/c^4$ (boundaries shown in blue).

4.6.2 Implications

The effective-q^2 range defined in Section 4.6.1 allowed theorists to make predictions for the four angular observables.

	adapted from [35]	Straub p.c. [23]	LHCb measurement [72]
F_{L}	$0.10^{+0.11}_{-0.05}$	0.175 ± 0.032	$0.16 \pm 0.06 \pm 0.03$
$\text{A}_{\text{T}}^{(2)}$	$0.03^{+0.05}_{-0.04}$	0.018 ± 0.002	$-0.23 \pm 0.23 \pm 0.05$
$\text{A}_{\text{T}}^{\text{Im}}$	$(-0.2^{+1.2}_{-1.2}) \times 10^{-4}$		$0.14 \pm 0.22 \pm 0.05$
$\text{A}_{\text{T}}^{\text{Re}}$	$-0.15^{+0.04}_{-0.03}$		$0.10 \pm 0.18 \pm 0.05$

Tab. 4.16: Predictions of the $B^0 \to K^{*0}e^+e^-$ angular observables from [35] and [23] are compared to the LHCb measurement [72] presented in this work. Good agreement with SM predictions is found for all the measured observables. From Straub private communication a prediction for A_{FB} rather than $\text{A}_{\text{T}}^{\text{Re}}$ is provided and no prediction for $\text{A}_{\text{T}}^{\text{Im}}$, which however is expected to be very suppressed in the SM as can be seen in the prediction provided by [35].

The errors associated to the predictions by [35] include a systematic uncertainty associated to the definition of the effective-q^2 range. However, this uncertainty was found to be small. The longitudinal fraction F_{L} is found to be small, as expected in this low q^2 range. The measured value of $\text{A}_{\text{T}}^{\text{Re}}$ is compatible at the level of 1.3 σ. The theoretical predictions for $\text{A}_{\text{T}}^{(2)}$ and $\text{A}_{\text{T}}^{\text{Im}}$ are very precise and both close to 0.

They are also well in agreement with the experimental values. As explained in Section 1.3.3 their measurement can be used to place new stringent constraints on the $\mathcal{C}_7$ and $\mathcal{C}_7^{'}$ Wilson coefficients. In particular, in the limit of $q^2 \to 0$ the relation between $A_T^{(2)}$, A_T^{Im} and $\mathcal{C}_7$, $\mathcal{C}_7^{'}$ is very simple as it can be seen in Equation 1.34. For the q^2 range considered in this analysis this approximation is valid at the level of 5% if SM values for the ratios $\mathcal{C}_9/\mathcal{C}_7$ and $\mathcal{C}_{10}/\mathcal{C}_7$ are considered. Then, constraints can be extracted on possible NP contributions to $\mathcal{C}_7$ and $\mathcal{C}_7^{'}$, hereafter called $\mathcal{C}_7^{(NP)}$ and $\mathcal{C}_7^{'(NP)}$ In a model-independent approach these parameters are allowed to be complex, making a total of four degrees of freedom. Therefore, the problem is simplified further by assuming a particular scenario in which NP appears only in the right handed operator $\mathcal{C}_7^{'(NP)}$. Then, constraints can be placed in the complex plane of $\mathcal{C}_7^{'(NP)}$ by assuming the SM for $\mathcal{C}_7$. In this plane, the SM is represented by the point in the center at $(0, 0)$. Constraints from the new measurements of $A_T^{(2)}$ and A_T^{Im} on $B^0 \to K^{*0}e^+e^-$ can then be compared and combined to the other existing constraints coming from radiative decays which are summarized in Section 1.3. Namely, constraints considered here are taken from:

- the measurement of the inclusive branching ratio of $B \to X_s\gamma$, where the experimental combination is taken from [52] and the SM prediction from [53];
- the time-dependent measurement of CP asymmetry in $B^0 \to K^{*0}(\to K_S^0\pi^0)\gamma$ decays which was measured at B-factories [62, 63] and is suppressed in the SM.

The constraint from $\mathcal{B}(B \to X_s\gamma)$ is concentric as $\mathcal{C}_7^{'}$ contributes to it with its modulus only. On the other hand, the constraint from $B^0 \to K^{*0}(\to K_S^0\pi^0)\gamma$ is oblique, the angle being given by the $B^0 - \overline{B}^0$ mixing phase as shown in Equation 1.23. The observables for different NP values of $\mathcal{C}_7^{'}$ are computed following equations in [61]. The compatibility of predictions with measurements are computed assuming Gaussian probability distributions and summing up in quadrature statistical and systematic uncertainties.
The constraints from the measurements of the transverse asymmetries $A_T^{(2)}$ and A_T^{Im} in $B^0 \to K^{*0}e^+e^-$ are shown in Figure 4.30. The measurement of $A_T^{(2)}$ constrains mainly the real part of $\mathcal{C}_7^{'}$, and A_T^{Im} the imaginary part. Constraints from $B^0 \to K^{*0}e^+e^-$ are found to be more stringent than the time dependent CP asymmetry from $B^0 \to K^{*0}(\to K_S^0\pi^0)\gamma$ which was measured at B-factories. The combination of all these constraints in Figure 4.31 shows very good agreement with the SM.

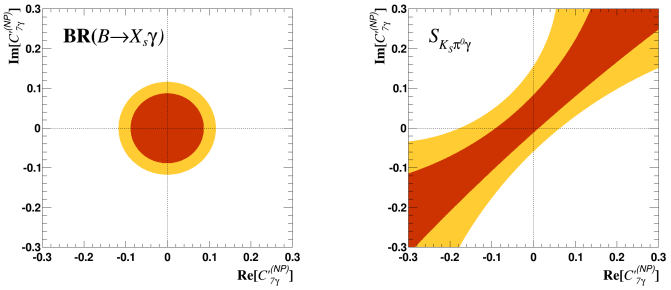


Fig. 4.29: Constraints from older $b \to s\gamma$ measurements on possible NP in $\mathcal{C}_7'$ are shown on the $\mathcal{C}_7'^{(\mathrm{NP})}$ complex plane. The orange (yellow) bound represents the 1σ (2σ) constraint. On the left is the constraint from $\mathcal{B}(B \to X_s\gamma)$ [52] while on the right is the one from the time-dependent CP asymmetry in $B^0 \to K^{*0}(\to K^0_S\pi^0)\gamma$ [62, 63].

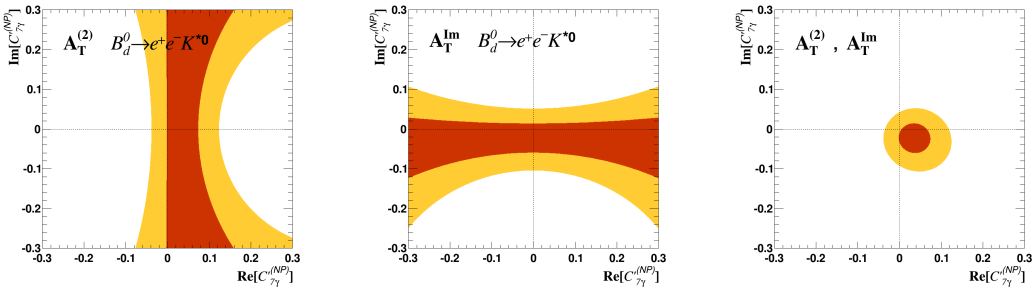


Fig. 4.30: Constraints from the angular analysis of $B^0 \to K^{*0}e^+e^-$ on possible NP in $\mathcal{C}_7'$ are shown on the $\mathcal{C}_7'^{(\mathrm{NP})}$ complex plane. Orange (yellow) bound represents the 1σ (2σ) constraint. On the left is the constraint from $\mathrm{A}_\mathrm{T}^{(2)}$ and in the center the one from $\mathrm{A}_\mathrm{T}^{\mathrm{Im}}$. On the right is the combination of the two, i.e. all the information coming from the angular analysis of $B^0 \to K^{*0}e^+e^-$ only.

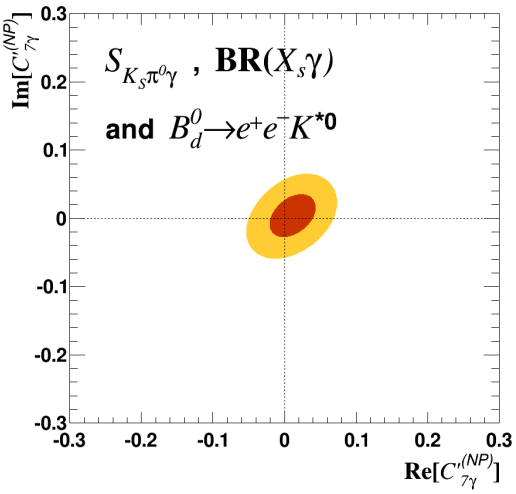


Fig. 4.31: Constraints on possible NP in $\mathcal{C}_7'$ are shown on the $\mathcal{C}_7'^{(\mathrm{NP})}$ complex plane coming from the inclusive branching ratio $\mathcal{B}(B \to X_s\gamma)$ [52], the time-dependent CP asymmetry in $B^0 \to K^{*0}(\to K^0_S\pi^0)\gamma$ [62,63] as well as the ones coming from the angular analysis of $B^0 \to K^{*0}e^+e^-$. The orange (yellow) bound represents the 1σ (2σ) constraint. Very good agreement with the SM is found.

4.7 Conclusions and perspectives

The data collected by the LHCb experiment in pp collisions at centre-of-mass energies of 7 and 8 TeV during 2011 and 2012 were used to select a sample of the rare FCNC decay $B^0 \to K^{*0} e^+ e^-$ in the q^2 region dominated by the photon pole below $1\,\mathrm{GeV}^2/c^4$. A full angular analysis of this sample was performed to extract four angular observables: F_L and $\mathrm{A}_\mathrm{T}^\mathrm{Re}$, which are related, respectively, to the longitudinal polarization fraction and to the lepton forward-backward asymmetry, and the transverse asymmetries $\mathrm{A}_\mathrm{T}^{(2)}$ and $\mathrm{A}_\mathrm{T}^\mathrm{Im}$, which in this very low q^2 region are sensitive to the right-handed component of the $\mathcal{C}_7$ Wilson coefficient, $\mathcal{C}_7^\prime$.

The result presented in this thesis demonstrates the power of the $B^0 \to K^{*0} e^+ e^-$ angular analysis at very-low q^2 in constraining possible BSM right-handed contributions to the $b \to s\gamma$ transition. This method is competitive with the ones relying on radiative decays: with the $3\,\mathrm{fb}^{-1}$ luminosity collected by the LHCb experiment in Run 1, this analysis provides at least the same level of precision on $\mathcal{C}_7^\prime$ as the combination of measurements done with radiative decays at B-factories. Moreover, the LHCb measurement is currently limited by the low statistics of the sample and the errors on the theoretical predictions are very small, therefore more data collected in the future will surely enhance the sensitivity.

With the data LHCb is going to collect in Run 2 (between 2015 and 2018) the total $B^0 \to K^{*0} e^+ e^-$ statistics is roughly expected to be 4 times larger thanks to both a higher integrated luminosity and a larger $b\bar{b}$ production cross-section (due to the higher pp collision centre-of-mass energy of 13-14 TeV). Therefore, statistical errors on $\mathrm{A}_\mathrm{T}^{(2)}$ and $\mathrm{A}_\mathrm{T}^\mathrm{Im}$ are expected to halve. Moreover, after the upgrade happening in 2018-2019 [143], LHCb will collect $5\,\mathrm{fb}^{-1}$ per year. The current plan is to run the detector in these conditions for 10 years, collecting a statistic which would allow to divide the current statistical error by roughly a factor 6. The Belle II experiment should start taking data in 2018 and is expected to collect about $50\,\mathrm{ab}^{-1}$ by 2024. While LHCb will continue to have the best statistics for $B^0 \to K^{*0} \mu^+ \mu^-$ decays, Belle II will be competitive in the electronic channel $B^0 \to K^{*0} e^+ e^-$ thanks to the $\Upsilon(4S)$ $e^+ e^-$ collider environment.

With more statistics available the analysis approach could benefit from moving the upper q^2 limit from $1\,\mathrm{GeV}^2/c^4$ down to a smaller value, *e.g.* $0.1\,\mathrm{GeV}^2/c^4$. According to the MC this cut would retain just 60% of the $B^0 \to K^{*0} e^+ e^-$ events, something which is not desirable with the statistics currently available as the angular fit could have stability issues. Nonetheless this approach could be considered when having additional data. Indeed, $B^0 \to K^{*0} e^+ e^-$ decays below $0.1\,\mathrm{GeV}^2/c^4$ are expected to have a lower value of F_L and thus more sensitivity to $\mathcal{C}_7^\prime$ with respect to events above $0.1\,\mathrm{GeV}^2/c^4$. Namely, in the MC the value of F_L changes from 0.16 to 0.03, resulting in 20% better sensitivity on the transverse asymmetries (for the same statistics). One more advantage is that the combinatorial background is

phase-space suppressed at very-low q^2 compared to the photon pole rise which is characteristic of the signal and of the partially reconstructed background. Hence, the systematic uncertainty associated to its angular modelling, which is one of the main contributors to the total systematic error (see Table 4.15), is expected to be reduced in the region of $q^2 < 0.1\,\mathrm{GeV}^2/c^4$. Moreover, the shape of partially reconstructed background could be better validated on data if the combinatorial component is suppressed and the statistics is high enough. If needed, also the systematic uncertainty related to the acceptance modelling could be reduced with a larger MC sample and a more accurate crosscheck on the $B^0 \to K^{*0}\gamma_{e^+e^-}$ control channel.

Moreover, with the current data other $b \to s\ell\ell$ angular observables such as the ones already measured in the muonic channel [26] can be measured in $B^0 \to K^{*0}e^+e^-$ in the very-low q^2 region that cannot be reached by muons. Although the $\mathrm{A}_\mathrm{T}^{(2)}$ and $\mathrm{A}_\mathrm{T}^{\mathrm{Im}}$ observables are the most sensitive to $\mathcal{C}_7'$, more orthogonal information on the photon polarization is in principle contained in other observables such as the P_i' ones. Furthermore, the $B^0 \to K^{*0}e^+e^-$ analysis methods presented in this document will contribute to the test of lepton universality of the $b \to s\ell\ell$ transitions. In particular, the measurement of $\mathcal{R}_{K^{*0}}$ which the LHCb collaboration is now preparing, will be an important followup of the $\mathcal{R}_K$ result as discussed in Section 1.2.2. Finally, with more statistic available the $B^0 \to K^{*0}e^+e^-$ angular analysis could be extended to higher q^2 values and ratios of angular observables with the muonic channel could give interesting lepton universality tests.

Chapter 5

The Focusing DIRC: a novel PID detector concept

This chapter presents my contributions to the R&D effort towards a new PID detector based on the *BABAR* DIRC concept, the Focusing DIRC (FDIRC). The work was carried out at the beginning of my PhD, between 2012 and 2013.
As an introduction, a brief overview of particle identification in particle physics experiments is made in Section 5.1. Then, the design of the *BABAR* DIRC is presented (Section 5.2) followed by a discussion of the reasons leading to the novel design of the Focusing DIRC detector (Section 5.3). The first full-scale prototype of the FDIRC was built and operated at the SLAC Cosmic Ray Telescope. The construction and commissioning of the prototype as well as the reconstruction procedure, the Geant4 simulation and the application of the so-called "chromatic correction" are detailed in Section 5.4. Then, in Section 5.5, the main results of this study are summarised and the benefit of the FDIRC R&D to similar detectors which are being developed for other experiments is overviewed.

5.1 Particle identification in particle physics

In most particle physics experiments, stable[1] Particle IDentification (PID) is of crucial importance, at the same level as tracking and calorimetry. The classification of particles is done either by analysing the way they interact with the different parts of the detector or by an indirect measurement of their mass. The first method is already incorporated in the design of most particle physics detectors, which are generally composed of a tracking system, electromagnetic and hadronic calorimeters and a muon system. This method allows to classify electrons, pho-

[1] In this context, particles are defined as "stable" if they have a lifetime long enough to pass through the detector material.

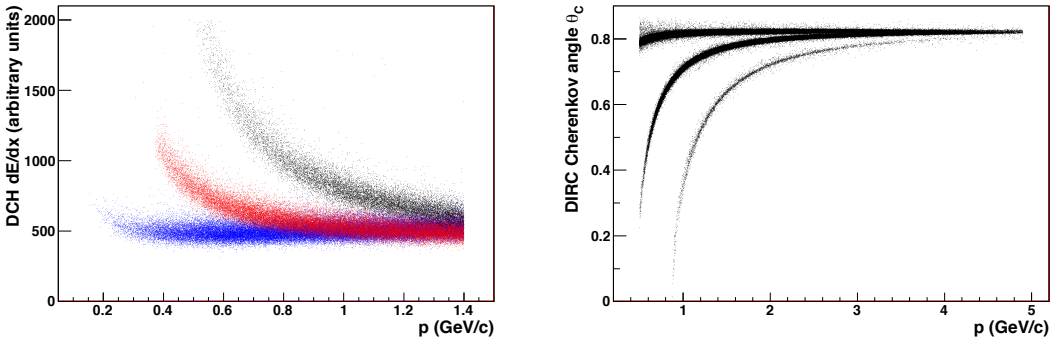


Fig. 5.1: (left) The measured $\mathrm{d}E/\mathrm{d}x$ in the *BABAR* drift chamber as a function of track momentum for three final state particles: pions (blue, lower band), kaons (red, middle band) and protons (black, upper band) [144]. (right) The measured Cherenkov angle in the *BABAR* DIRC as a function of track momentum for pions (upper band) kaons (middle band) and protons (lower band) [144].

tons, hadrons and muons. However, charged hadrons (π, K and p), have mostly the same behaviour in these detectors (charge deposit in the tracking system and hadronic shower in the calorimeters) and are therefore very difficult to identify by this method. Nonetheless, their identification is of major importance for any experiment willing to study hadronic decays and a way to access their mass difference becomes necessary. Since stable particle masses cannot be measured directly, they are normally deduced from the measurement of the momentum p and the velocity $c\beta$ through the basic equation:

$$m = \frac{p}{c\beta\gamma} \tag{5.1}$$

where c is speed of light and γ is the relativistic Lorentz factor $\gamma = (1-\beta^2)^{-1/2}$. The particle momentum is usually measured precisely by the track curvature in a magnetic field, while a measurement of the particle velocity is in general more difficult as most of the time it is very close to the light speed. Various methods can be employed: a time-of-flight (TOF) measurement, the detection of the Cherenkov radiation emission angle (see later for details), the detection of transition radiation or a measurement of the energy deposit due to ionization ($\mathrm{d}E/\mathrm{d}x$). The identification power of each method is restricted to certain momentum ranges. For example, the measurements of energy deposit by ionization and Cherenkov angle by *BABAR* are shown in Figure 5.1 as a function of p. However, the choice of the PID method to employ depends also on requirements such as size, event rate, luminosity, material budget and geometrical coverage.

Most experiments requiring good charged hadron identification, such as flavour-physics experiments, have a dedicated detector for PID which relies on the mea-

surement of the Cherenkov radiation emission angle. Cherenkov radiation is a shock wave resulting from a charged particle moving through a material faster than the velocity of light in that material. Its discovery and interpretation earned the Nobel prize to P. Čerenkov, I. Y. Tamm and I. M. Frank in 1958. Cherenkov light is emitted with a polar angle θ_C with respect to the particle velocity given by [145, 146]:

$$\cos\theta_C = \frac{1}{\beta n_{\mathrm{p}}(\lambda)}, \tag{5.2}$$

and thus depending only on the particle velocity β and on the phase refractive index of the material traversed $n_{\mathrm{p}}(\lambda)$, which in turn depends on the wavelength of the emitted light λ. Cherenkov detectors are normally composed of a transparent dielectric medium through which charged particles pass and of some optic elements to image the emitted light into photodetectors in order to measure θ_C and thus access β. These are known as Ring Imaging CHerenkov (RICH) detectors [147]. Furthermore, since $|\cos\theta_C| \leq 1$, Cherenkov radiation is emitted only above a threshold velocity $\beta_t = 1/n_{\mathrm{p}}(\lambda)$. Some simple detectors called "threshold counters" [148] use this principle to distinguish particles in a narrow momentum range by the fact that they emit or not Cherenkov light.
Cherenkov radiation is emitted promptly as a charged particle passes through the medium and thus provides very accurate timing information, which can be used, for example, to measure precisely the TOF of the particle. The light emitted by a single particle corresponds to a small number of observed photo-electrons $N_{\mathrm{p.e.}}$ which can be approximated by the equation:

$$N_{\mathrm{p.e.}} \simeq N_0 L \sin^2\theta_C \tag{5.3}$$

where L is the path length of the particle through the radiator and the factor N_0 contains the light transmission, collection and detection efficiencies and has typical values between 30 and 180 cm^{-1}.

Since the first RICH detector, designed and built by A. Roberts in 1960 [147], many different types of RICH detectors were developed. Various radiative media and imaging techniques have been used. The choice of the radiative medium is driven by the average momentum of hadrons that need to be classified. For high momentum, β approaches 1 and therefore one needs a medium with low refractive index n (*i.e.* close to 1) in order to have sensitivity on the velocity of the particle. Indeed, common refractive media span from solid fused silica ($n_{\mathrm{p}} \simeq 1.4$) or water ($n_{\mathrm{p}} \simeq 1.3$) for momenta of few GeV/c [149] to aerogels ($n_{\mathrm{p}} \simeq 1.03$) or gases such as CF_4 ($n_{\mathrm{p}} \simeq 1.0005$) for higher momenta [111].
The imaging technique used to measure θ_{C} is quite straightforward if the radiative material is thin and thus the photon emission region is small compared to the imaging camera size: the emitted Cherenkov light forms a ring whose radius is

proportional to θ_C. This is normally the case for dense solid/liquid radiators, as a small thickness is enough to cause the emission of a sufficient amount of photons. However focusing mirrors are required when the path length of the charged particle in the detector is long, as for lighter media such as aerogel and gas.

5.2 The *BABAR* DIRC

The DIRC (Detector of Internally Reflected Cherenkov light) [150, 151] is a PID detector concept that was developed for the *BABAR* experiment [144, 152] and has been crucial to its performance [149]. The *BABAR* detector (shown in Figure 5.2 operated from 1999 to 2008 at the PEP-II asymmetric e^+e^- collider at the SLAC National Laboratory. The collider was run mainly at the $\Upsilon(4S)$ centre-of-mass energy (10.58 GeV), a resonance that decays exclusively to $B^0\overline{B}^0$ and B^+B^- pairs. The asymmetric energies of the two colliding beams, 9.0 and 3.1 GeV for the electron and positron beam respectively, resulted in boosted $B\bar{B}$ pairs ($\beta\gamma = 0.56$). The boost allowed to distinguish the B decay vertices, determine their relative decay time and measure time-dependent decay rates. The PEP-II collider has been run at a luminosity of up to 10^{34} cm^{-2}s^{-1}. The *BABAR* flavor physics program required a detector being able to separate kaon and pion tracks with momenta up to 4 GeV/c. Below 0.7 GeV/c, PID relied on the dE/dx measurement in the drift chamber and silicon vertex tracker, but for higher momenta a dedicated PID device based on the Cherenkov angle measurement was necessary. This sub-detector was required to cover the whole barrel acceptance, to account for less than 10% X_0 (radiation lengths) and to be as thin as possible to leave space for tracker and calorimeter (the cost of the calorimeter system scales roughly like the inner radius squared). This requirements led to the development of a novel Cherenkov detector concept, the DIRC, which allows to have the radiative material in the barrel and the Cherenkov-ring imaging outside of it.

In fact, instead of imaging directly the Cherenkov cone on a plane in front of it, photons are transported to a camera located outside the barrel via total internal reflection inside the same dielectric material they originate from, as shown in Figure 5.3. Indeed, the radiative material is shaped in long bars with rectangular cross section, which work both as radiator and light guide. Light angular information is preserved while propagating along bars, as it is designed with almost perfectly parallel and flat sides, namely with surface roughness below 5 Å (rms) and squared to better than 0.3 mrad.
The entire *BABAR* DIRC has 144 bars of pure synthetic fused silica, each 4.9 m long and with cross section 1.7×3.5 cm^2. Each bar is actually composed of four 1.225 m shorter bars optically glued together with an epoxy glue. Bars are grouped 12 by 12 in a light and tight support structure, the bar box. Bars are set along the beam

line and cover the whole azimuthal range for polar angles between −51.4 and +64.5 degrees (corresponding to 83% of the polar angle coverage in the centre-of-mass frame), as shown in Figure 5.4. Cherenkov photons are transported to the back end of the bars (towards negative z in the reference frame defined in Figure 5.2) where they exit into a pinhole camera consisting of a large volume of ultra-pure water. This medium was chosen because it is inexpensive, transparent, easy to clean and with average index of refraction and relative chromatic dispersion sufficiently close to fused silica corresponding properties. The camera hosts 10,752 densely packed Photo Multiplier Tubes (PMTs) displayed uniformly in terms of angle coverage and about 1.2 m away from the quartz bar exit window. A wedge of fused silica is optically glued to the bar exit to "bend" photons coming out at large angles relative to the bar axis. It reduces the size of the required detection surface and recovers photons that would otherwise be lost due to internal reflection at the fused-silica-to-water interface.
The reconstruction of the Cherenkov angle is done combining information from the tracking system together with the positions of the PMT hits in the DIRC. As shown in Figure 5.3 (right), the Cherenkov ring imaged on the PMT plane is usually doubled due to the fact that each photon can do its last side-reflection

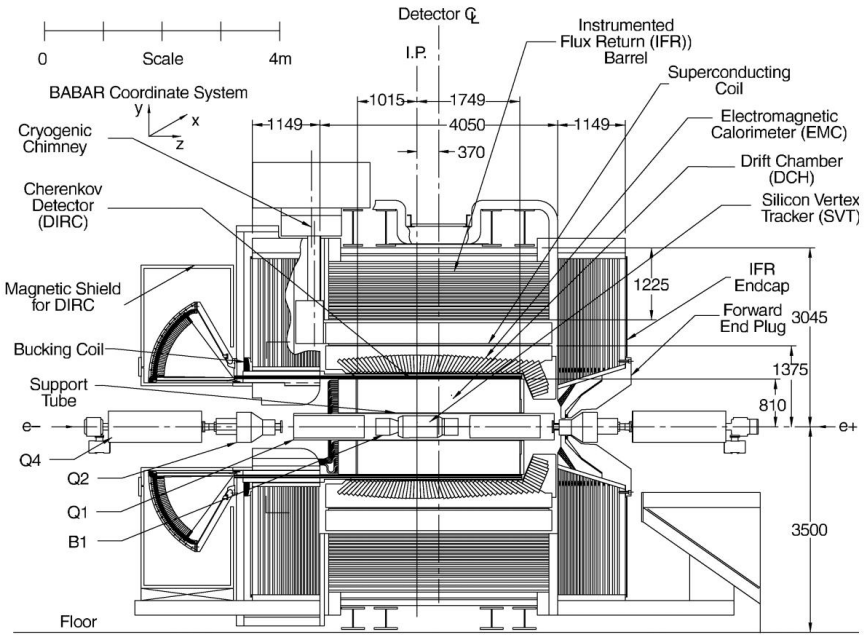


Fig. 5.2: Schematic of the *BABAR* detector [149]. Radially, the DIRC sub-detector occupies a thin layer (∼ 1.7 cm) after the silicon vertex tracker and the drift chamber and right before the electromagnetic calorimeter.

on either side of the bar *i.e.* in the direction perpendicular to the bar side-view in Figure 5.3 (left). Moreover, the ring image is affected by the wedge, that can move part of it to another region of the PMT imaging plane. Photon hits were measured with a time resolution of 1.7 ns, mainly driven by the PMT Transit Time Spread (TTS) of 1.5 ns. The information on the photon time of arrival was used to reject background hits, as shown in Figure 5.3 (right), as well as to resolve some of the ambiguities arising from the different possible photon paths from the charged track to a PMT. The dominant contributor to the overall detection efficiency is the quantum efficiency of the PMT (which has a strong dependence on wavelength). However, the main low cut-off on wavelength is due to the transmission through the epoxy glue used to glue the fused-silica short-bars together, which cuts sharply ultraviolet photons with $\lambda < 290$ nm. Furthermore, a factor playing an important role is the minimum angle of incidence for total internal reflection (the so-called critical angle) which is given by Snell's law:

$$\sin \alpha_{\mathrm{ref}}^{\mathrm{min}} = \frac{1}{n_{\mathrm{p}}(\lambda)} \tag{5.4}$$

For the interface fused-silica / air and an average 410 nm photon, the minimum angle of incidence is thus $\alpha_{\mathrm{ref}}^{\mathrm{min}} \simeq 42.9°$. This requirement makes a strict selection of photons at their first hit on a bar surface which largely depends on the track

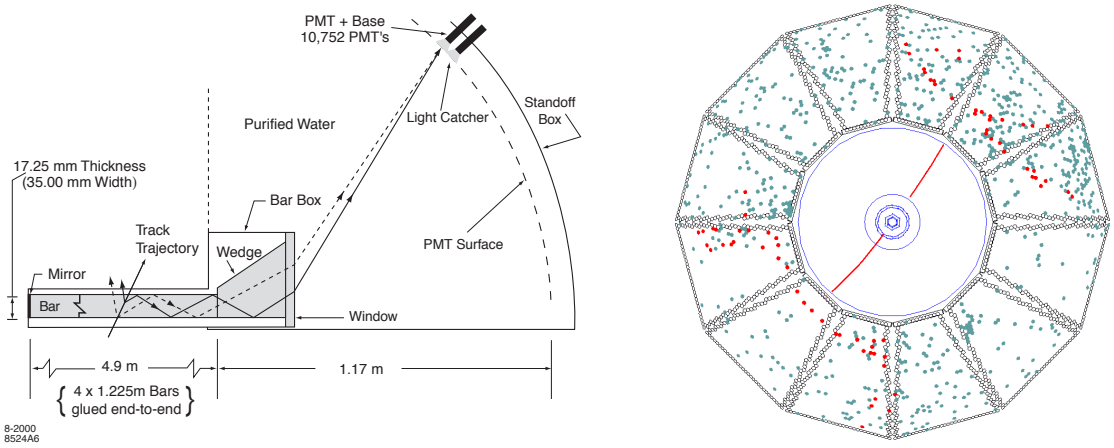


Fig. 5.3: (left) Schematic of the DIRC fused silica radiator bar and the imaging region [149]. The bars active region covers polar angles between −51.4 and +64.5 degrees. (right) Display of the PMT hits (black and grey dots) in the *BABAR* DIRC camera originating from a dimuon event [149]. Green dots are background, while red dots are signals from the muon being selected by requiring hits to be within 8 ns of the expected Cherenkov photon arrival time.

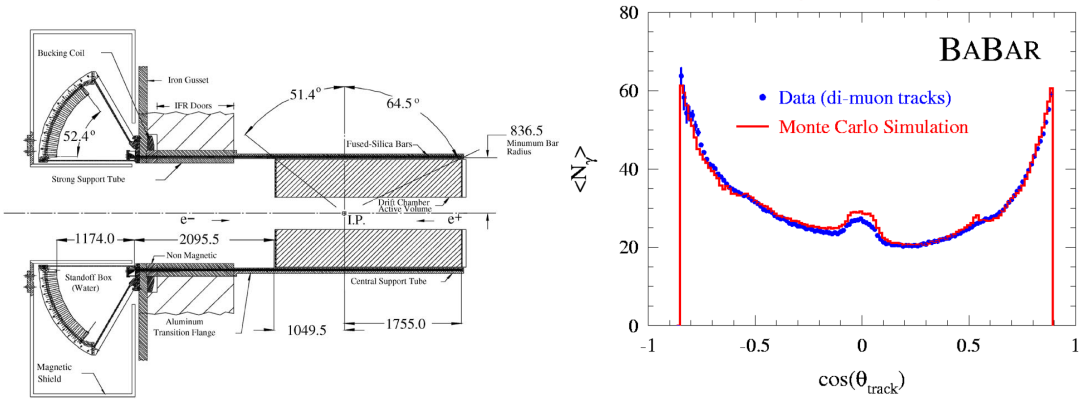


Fig. 5.4: (left) Elevation view of the nominal DIRC system geometry [149]. All dimensions are given in millimeters. (right) Number of detected photo-electrons as a function of the polar angle of the track for di-muon events [149].

incidence angle, see Figure 5.4 (right). A larger part of the ring is lost for tracks with a small dip angle, giving a number of photo-electrons as low as 20, while up to 60 photo-electrons are detected for tracks with larger incidence angle. This feature is a strong point of the DIRC design as it allows to get a better Cherenkov angle measurement ($\sigma_{\theta_C} \propto 1/\sqrt{N_{\text{p.e.}}}$) for tracks with large polar angles, which in general are also the ones with higher momenta. Photons can undergo hundreds of reflections while propagating along the bars, therefore a coefficient for internal reflection of ~ 0.9997 is required[2].

The overall single photon Cherenkov angle resolution receives its largest contribution from the pinhole imaging method, *i.e.* from the fact that the size of the bar exit window (as well as that of PMT photo-cathodes, which is about 2.5 cm in diameter) is not completely negligible compared to the distance between the end of the bars and the PMTs. For the *BABAR* DIRC, this accounts to about 7 mrad, slightly larger than the resolution contribution coming from photon production and dispersion during transmission through the optical elements. The latter accounts to 5.4 mrad and depends on the chromatic dispersion of the fused silica as well as on the accepted range of wavelengths. The overall single photon resolution is thus 9.6 mrad.

The *BABAR* DIRC performed reliably and efficiently during the whole *BABAR* data taking period. Its main performance parameters are the following:

- time resolution of about 1.7 ns, close to the PMT TTS of 1.5 ns;

[2]This corresponds to a $(0.9997)^{100} \simeq 97\%$ probability for a photon to undergo 100 internal reflections.

- single photon Cherenkov angle resolution of 9.6 mrad for dimuon events;
- Cherenkov angle resolution per track of 2.5 mrad in dimuon events;
- $K - \pi$ separation above 2.5 σ from the pion Cherenkov threshold up to momenta of 4.2 GeV/c.

5.3 The Focusing DIRC concept

Based on the success of the DIRC detector, an R&D program has been pursued to develop a compact and fast detector for future PID systems [154, 155]. The large water tank of the *BABAR* DIRC was sensitive to backgrounds resulting mainly from neutrons interacting with the H_2O molecules in the huge imaging camera (6000 litres). Moreover, it was a permanent concern as water could leak in the bar boxes containing the DIRC quartz bars and from there reach other parts of the *BABAR* detector, causing serious and permanent damage to the apparatus. A new design of the detector was thus needed for it to operate in experiments with higher

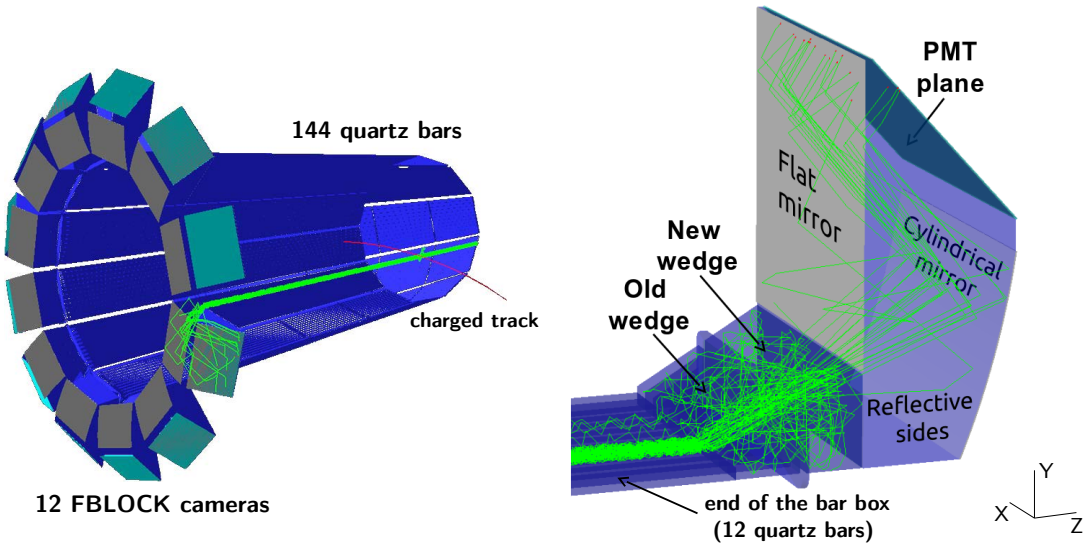


Fig. 5.5: (left) A 3-D picture of the full FDIRC design with 12 modular cameras and (right) of one of the photon camera from the Geant4 simulation. The mechanical support is not shown. The paths of Cherenkov photons from a charged muon track are represented in green. Few of them got reflected on the interface between the different quartz pieces. These photons may result on a PMT hit, but they are significantly delayed in time and can easily be rejected. [153]

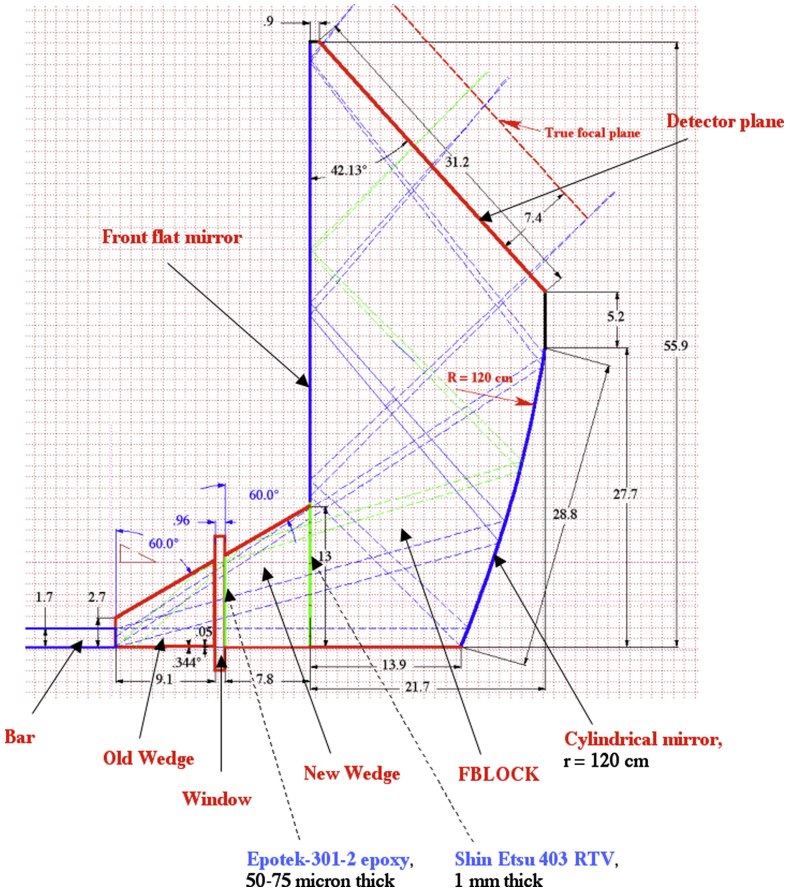


Fig. 5.6: FDIRC optical design (dimensions are in cm). Dashed lines show photon paths from the bar exit to the detector plane where they are focused onto. All paths include reflections on the two mirrors. The green photon path also includes a reflection on the top of the wedge. [158]

background rates. Initially, this R&D effort was mainly intended to build a DIRC of new generation to cover the barrel region of the *SuperB* detector. The *SuperB* project [156], consisted in a B-factory with a luminosity two orders of magnitude higher than previous B-factories, PEP-II for *BABAR* and KEK-B for Belle. The main background was expected to grow proportionally to the luminosity and was not bearable with the *BABAR* DIRC design. Unfortunately, the *SuperB* project was cancelled at the end of 2012 due to lack of funding. A similar Super B-Factory is being built in Japan by the Belle-II collaboration [157].

This new detector concept, the Focusing DIRC (FDIRC) [159], consists in a new compact and fast imaging camera made of fused silica and coupled to quartz bars of the same material. Indeed, the main strategy to reduce background is to make the camera faster and much smaller, thanks to fine pixelated photo-detectors and

focusing optics. While the imaging camera is completely redesigned, the invaluable *BABAR* DIRC quartz bars are reused to limit costs, requiring the new camera to be compatible with them. The new camera design is modular: it consists of 12 solid blocks of fused silica in place of the single large water-filled DIRC camera (see Figure 5.5). These blocks are machined from radiation-hard pieces of fused silica, the same optic material the bars are built of. The major design constraints for the new camera are the following:

- it must be compatible with the existing *BABAR* bar box design;
- it requires very fine photon detector pixelation (because of the smaller dimensions of the camera)
- it requires very fast photon detectors (∼0.2 ns) for background rejection and *chromatic correction*[3].

Imaging is provided by two mirrors focusing photons onto a plane containing highly pixelated photomultiplier tubes. The reduced volume of the new camera as well as the use of fused silica for coupling to the bar boxes, is expected to reduce the sensitivity to background by about one order of magnitude. The very fast timing of the new PMTs provides separation of ambiguous solutions in the folded optical system and roughly another order of magnitude improvement in background rejection. In addition to being useful for background rejection, the FDIRC timing allows to correct the Cherenkov angle measurement for the chromatic dispersion contribution which accounts for about half of the total resolution ($\sigma_{\text{Chromatic}} \sim 4.5\,\text{mrad}$). Nonetheless, the FDIRC detector has been designed to measure Cherenkov angles with very good precision using the geometrical imaging only, which improves the robustness of the design in high multiplicity environments.

Figures 5.5 (right) and 5.6 show the new FDIRC photon camera design (see Ref. [160] for more details). It consists of two parts: a focusing block (FBLOCK) with a cylindrical mirror and a flat mirror, and a new wedge. The wedge at the end of the bar "rotates" rays with large transverse angles in the vertical plane (see the green pattern in Figure 5.6). The old *BABAR* wedge which is attached to the end of the bars is too short to ensure that all rays strike the FDIRC cylindrical mirror, therefore it is prolonged with a new one. As shown in Fig. 5.6, the cylindrical mirror focuses in the radial y-direction, while just "pinhole focusing" is used in the direction out of the plane of the schematic (the x-direction). Geometrically, focusing in the x-direction in the same way as in y is impossible as the mirror is shared by the 12 bars of each bar box. Alternative solutions are being investigated by recent R&D for other DIRC-like detectors. For example, the PANDA collaboration is considering a DIRC design with spherical lenses instead of a mirror to

[3]See Section 5.4.6 for more details.

focus on both x and y directions simultaneously [161]), while the LHCb TORCH detector camera, designed in order to adapt to *BABAR* bars, has a focusing mirror sliced and curved in the x-direction as shown in Figure 5.24. Both options would have been difficult to implement in the FDIRC and it would not have strongly benefitted from focusing in x since the most interesting angular information is in the y-direction as just a portion of the Cherenkov ring is detected (Figure 5.7. Thus, the FDIRC optics has a worsening of the angular resolution for photons in the peripheral "wings" of the partial Cherenkov ring detected [160].

Photons that enter the FBLOCK at large x-angles reflect from the parallel sides, leading to more complexity in the ring pattern, which is folded on the PMT plane as shown in Figure 5.7. This leads to more complexity in the reconstruction procedure (more details in Section 5.4.5). However, the folded design makes the optical piece small, and places the photon detectors in an accessible location, improving both the mechanics and decreasing the sensitivity to background. A flat mirror reflects rays onto a plane instrumented with photon detectors in order to further reduce the volume of the camera. The optics is designed so that photons impinge on the instrumented plane with an incidence angle of almost 90° in the plane of Figure 5.6, thus minimizing the possibility of reflection and maximizing

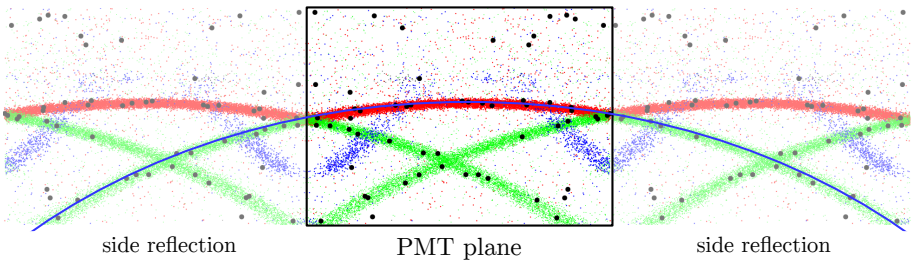


Fig. 5.7: Illustration of the effect of the new folded optics on a Cherenkov ring pattern imaged by the FDIRC. The expected image produced by a pion entering one a quartz bar with perpendicular angle of incidence is represented in red, green and blue, the different colours representing 0, 1 and 2 reflections on the sides of the FBLOCK (actually, the probability of doing 2 reflections is very small). The expected ring pattern is overlaid by black circles representing PMT hits from a single track. The black square in the centre frames the picture in the PMT plane. For illustration purposes this image is mirrored on the sides following possible side reflections. Putting side by side all mirrored pictures allows to follow by eye the Cherenkov ring pattern, which is indicated by a blue line. The ring pattern is quite symmetrical because Cherenkov light is produced in one of the bar in the middle of the barbox. Furthermore, given that the track is perpendicular to the bar, the forward- and backward-going parts of the ring are one on top of the other, simplifying further the picture.

the collection efficiency. The instrumented plane is located in a slightly under-focused position to reduce the FBLOCK size and therefore its weight. Precise focusing is unnecessary, as the finite pixel size would not take advantage of it. The total weight of the solid fused silica FBLOCK is about 80 kg, requiring good mechanical support.

In summary, there are several important advantages gained in moving from the *BABAR* DIRC pinhole focused design with water coupling to the FDIRC compact focusing design made of solid fused silica:

- the design is modular (12 independent modules), therefore easier to maintain
- the sensitivity to background, especially to neutrons, is significantly reduced
- the total number of photomultipliers is reduced by about one half compared to a non-focusing design with equivalent performance
- there is no risk of water leaks, and no time-consuming maintenance of a water system, as was required to operate *BABAR* safely.

5.4 The first full-size FDIRC sector prototype

In order to check the feasibility of the Focusing DIRC concept, a first prototype was built at SLAC almost a decade ago. It was composed of a single fused silica bar and the produced Cherenkov light was imaged by a focusing device made of a spherical mirror (instead of the cylindrical one which has been chosen for the final FDIRC design), an oil-filled photon camera, and highly-pixelated photon detectors. It was constructed and operated in a test beam in 2005, 2006, and 2007 [162, 163], and later tested in the SLAC cosmic ray telescope (CRT) [164] using cosmic muons [165]. These initial tests concentrated on proving the principle of focusing optics, learning how to operate highly pixelated fast detectors, and showing that the chromatic broadening can be corrected by timing [162,163]. Then, a full-scale prototype of $1/12^{\text{th}}$ of the FDIRC was built and tested with multi-directional muon tracks at the SLAC Cosmic Ray Telescope in 2011-2013. In the following sections, the construction, commissioning and testing of this prototype are described in detail.

5.4.1 Building of the new compact optical camera

One spare *BABAR* DIRC bar box was reused, containing 12 bars with (old) wedges at the bar exit. The imaging camera, on the other hand, was built from scratch following the design described in Section 5.3. The focusing block (FBLOCK) and the new wedge were machined from radiation-hard Corning 7980 fused silica. The required surface-polishing quality was only 30 Å (rms) because each photon only bounces 3−5 times inside the photon camera. The two FBLOCK reflecting surfaces were coated with aluminium and with a SiO_2 overcoat to protect it; Figure 5.8(a) shows a picture of the optics. The new wedge was glued to the bar box with a $50 - 75\,\mu$m thick layer of optical epoxy (Epotek 301-2). The FBLOCK was enclosed in an aluminium box, see Figure 5.8(b), and optically coupled to the wedge with 1 mm thick Shin-Etsu 403 Room Temperature Vulcanizing (RTV) silicone, as shown in Figure 5.8(c). This type of coupling allows a possible separation of the two pieces in case of problems.

5.4.2 Photon detectors and electronics

The photon detector for the FDIRC detector needs to have pixels of few mm and timing capabilities well below the ns. No stringent magnetic field tolerance is required as the photo detectors can stay outside the magnetic field thanks to the long bars penetrating the magnet iron. The choice was directed to the Hamamatsu H-8500 multi-anode PMT (MaPMT) shown in Figure 5.9(a). This model is meeting all requirements and comes at a small price since it is largely used by the medical

Fig. 5.8: (a) The FBLOCK made of radiation hard fused silica (Corning 7980). (b) The FBLOCK enclosed in its aluminium box. (c) Coupling of the new wedge to the FBLOCK.

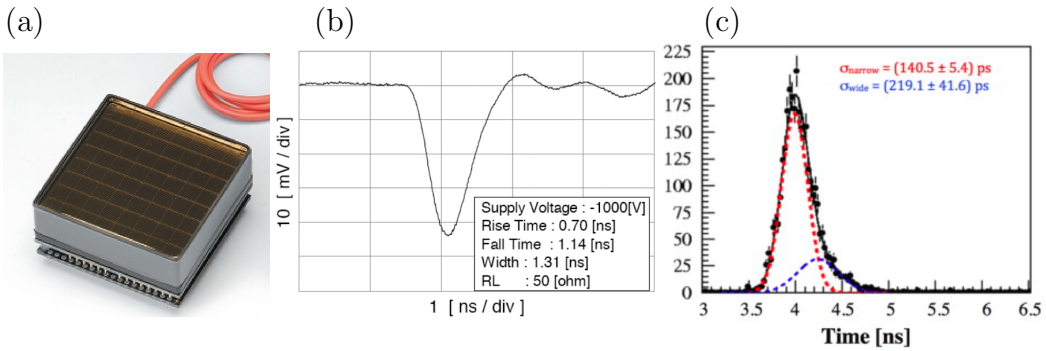


Fig. 5.9: (a) Picture of the Hamamatsu H-8500 Multi-anode PMT. (b) A typical shape of a signal from an H-8500 without an amplifier. The rise time is about 0.7 ns. (c) The H-8500 MaPMT single photoelectron transit time spread is $\sigma_{TTS} \sim 140$ ps [155].

community. This MaPMT has 64 pixels approximately 6×6 mm^2, a rise time of $\sim$0.7 ns, and a total pulse width of $\sim$1.3 ns as shown in Figure 5.9(b). It also features good Quantum Efficiency (Q.E.) around 24% for an average wavelength photon of 410 nm. Figure 5.10 is an illustration of the main factors affecting the wavelength spectrum of the collected light: the photocathode Q.E. plays the leading role, but the cutoff at short wavelentgh is due to the epotek used to glue the short bars together and the reflection coefficient also depends strongly on wavelength. Due to funding constraints, the FDIRC prototype camera was populated with 12 H-8500 MaPMTs. This only partially covers the focal plane since a com-

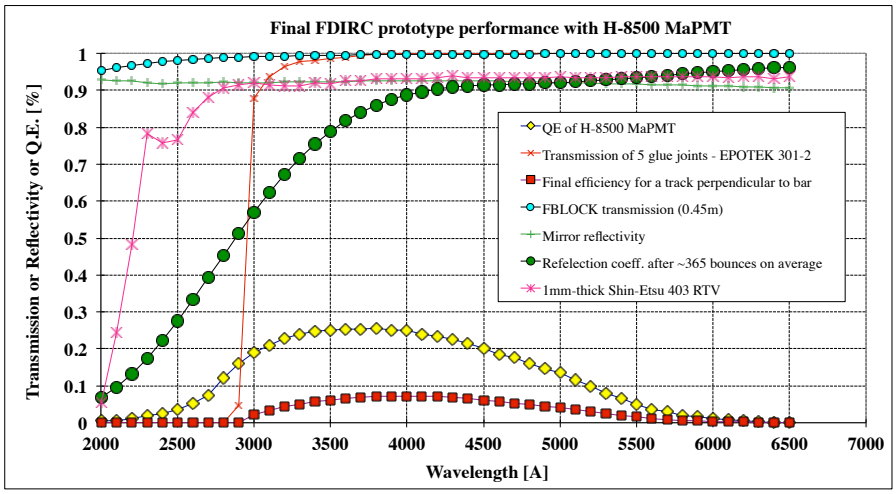


Fig. 5.10: Various contributions to the photon detection efficiency as a function of wavelentgh. The main constrain comes from the Quantum Efficiency (Q.E.) of the H-8500 MaPMT photo-cathode. Nonetheless, the cut-off at short wavelength is due to the epotek used to glue the short bars together. The expected total efficiency depends on several other factors, those shown here are the ones having the strongest dependence on wavelength.

plete coverage would require 48 such PMTs. The location of the 12 MaPMTs was chosen to optimise the coverage for the expected photon hit locations from close-to-vertical cosmic-ray muons hitting the bars in the SLAC CRT. To benefit from the speed of the tube, a pre-amplifier with a modest gain of a factor 40 was used, the same as the one that was successfully used on the initial focusing DIRC prototype [154, 155]. Using this amplifier and a simple constant fraction discriminator (CFD), a single photo-electron Transit Time Spread (σ_{TTS}) of $\sim$140 ps [154, 155] was recorded. The corresponding timing distribution is shown in Figure 5.9(c).

After passing through the pre-amplifier, PMT signals are sent into the IRS2 application-specific integrated circuit (ASIC) [166]. The proposed *SuperB* design was based on CFD-on-a-chip electronics [167] but that board was not available before the end of the data taking, hence the IRS2 ASIC was used instead. This is a fast waveform sampler configured to sample at a frequency of $\sim$2.7 GSa/s. Digitised data are sent to a back-end CompactPCI® (cPCI) crate via gigabit fiber links; a linux-based data acquisition (DAQ) system running on a cPCI CPU collects the digitised waveforms and stores them to disk for off-line analysis. Raw IRS2 waveforms require a substantial amount of calibration and processing in order to extract the photon time of arrival. An example of a waveform coming from a PMT signal is shown in Figure 5.11 (right). The time-of-arrival of each pulse is calculated based on a software CFD method [168]; this method was found to

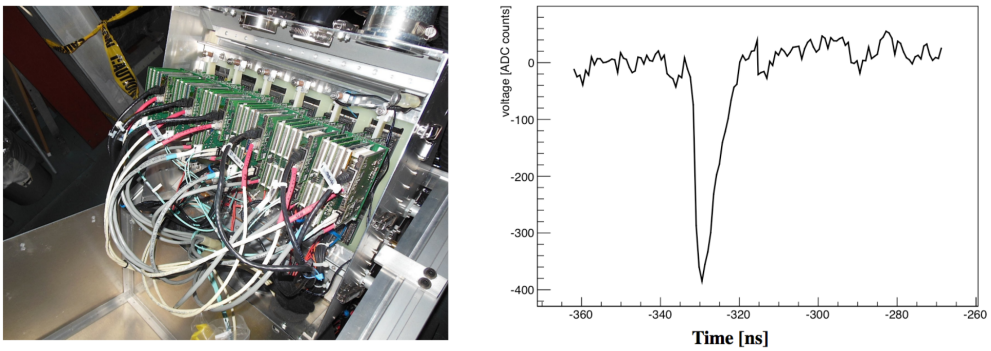


Fig. 5.11: (left) The 12 H8500 MaPMTs with pre-amplifiers and IRS2 digitizing electronics mounted on the FDIRC prototype. (right) An example of a pedestal-subtracted waveform from a PMT signal.

be competitive with other timing techniques [169]. Linear interpolation is used between points to obtain sub-sample precision on the time of arrival.

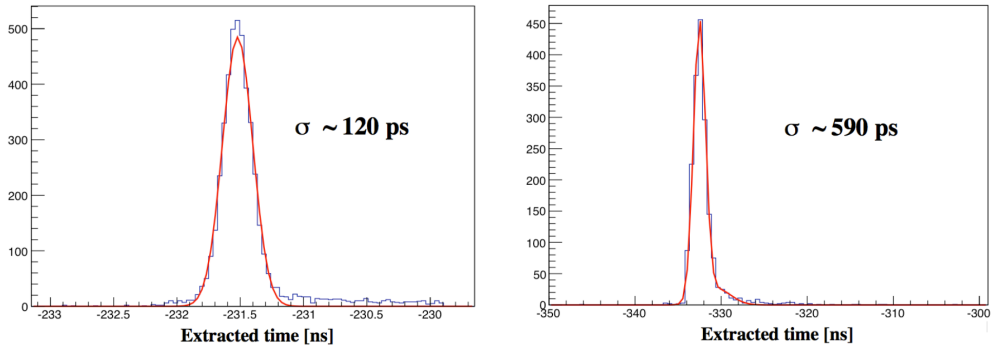


Fig. 5.12: (left) Timing distribution obtained with fixed-amplitude signal injected in the last dynode of a MaPMT. (right) Timing distribution obtained with laser-induced single photoelectrons.

The electronics timing performance is evaluated using two methods: first, test pulses of fixed amplitude are injected into the signal chain through the last dynode of a H-8500 MaPMT. The timing of these test pulses is measured using a simple threshold crossing technique on the digitised data. This test allowed to measure the timing resolution of the system, excluding most waveform processing limitations. Resolutions obtained are in the range of $\sim 100 - 200$ ps, as shown in Figure 5.12 (left). This indicates that such resolutions are in principle achievable. The second method used to evaluate the timing performance consists in producing laser-induced single photoelectrons. As PMT signals have significant variation in pulse height, the timing of these pulses is calculated with a software

CFD technique, similar to the one used for Cherenkov photons data. As shown in Figure 5.12 (right), a significant degradation of the timing resolution is observed. Typical timing resolutions for these signals are around ∼600 ps, suggesting that variation in the IRS2 amplitude response is the primary cause of the degradation. Such amplitude-dependent variations arise due to instabilities in operating the IRS2. In particular, a large number of control signals with tight margins between them is required to operate the ASIC and small misalignments of these signals can cause complicated systematic amplitude variations. The IRS2 also has no on-chip method to lock the time-base, so these amplitude variations drift with environmental conditions, making it extremely difficult to remove such effects in the full FDIRC-prototype dataset. A number of revisions have been undertaken to resolve these and other issues with the IRS2 architecture, such as the implementation of an on-ASIC delay locked loop to stabilize the time base, and an internal timing generator that allows very fine internal adjustment of the signals. These improvements have culminated in the IRSX ASIC[4], which will be used to read out the TOP (Time-Of-Propagation) counter of the Belle-II experiment.

5.4.3 Installation at the SLAC Cosmic Ray Telescope

The full-size FDIRC sector prototype was installed at the SLAC Cosmic Ray Telescope (CRT) as shown in Fig. 5.13. This facility [164] is well-suited for testing this kind of detector as it features:

- two planes of scintillator hodoscopes that provide tracking of cosmic ray muons with an angular resolution of ∼1.5 mrad and track position resolution of $4 - 5$ mm;
- a fused silica start counter consisting of two fused silica bars coupled to a 4-pixel MCP-PMT which provides a start time resolution better than 70 ps;
- a stack of four iron absorbers (and one lead absorber) instrumented with scintillators which is used to select muons with energy greater than ∼2 GeV.

The scintillator hodoscopes have active areas of 51 cm × 107 cm and the angular acceptance shown in Figure 5.14 (left) in terms of the X and Z components of the track direction $\hat{n}$ (the reference frame is defined in Figure 5.13). The large angular acceptance allowed to study the response of the FDIRC to multi-directional tracks, something that would be difficult using a test beam. The bar box is located within the CRT such that selected cosmic muons pass near the midpoint along the length of the fused silica radiators. Most of muon tracks cross central bars as shown in Figure 5.14.

[4]Information on the IRSX ASIC is expected to be published in the near future.

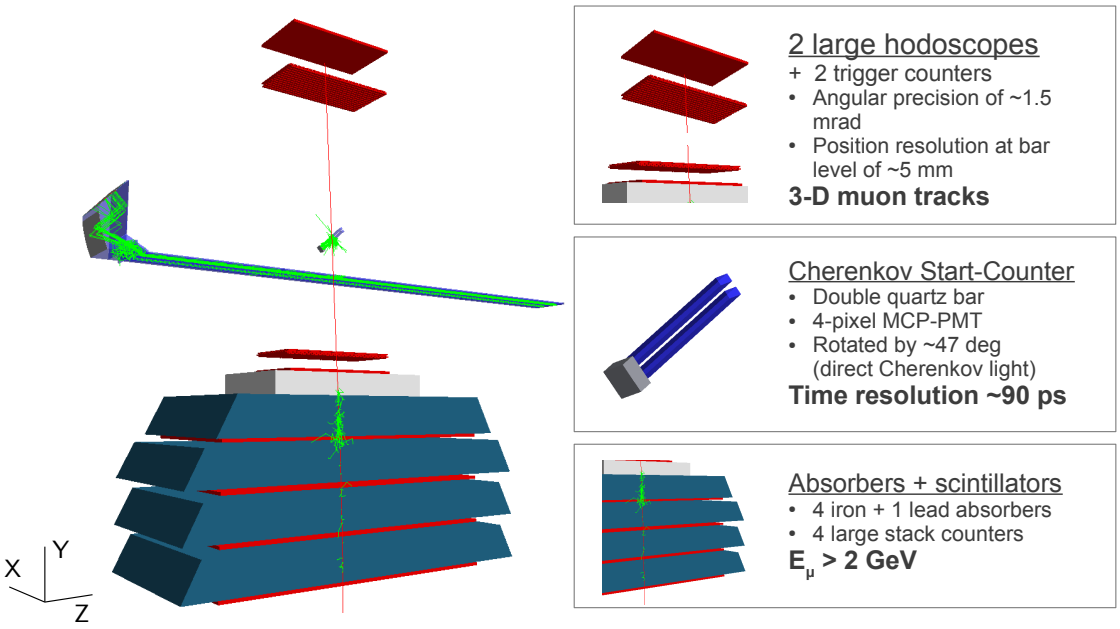


Fig. 5.13: Picture of the FDIRC prototype installed in the CRT from the Geant4 simulation. High energetic cosmic muons (red track in the figure) are used as a probe to test the detector. The simulation is as realistic as possible: it includes all the sub-detectors composing the CRT.

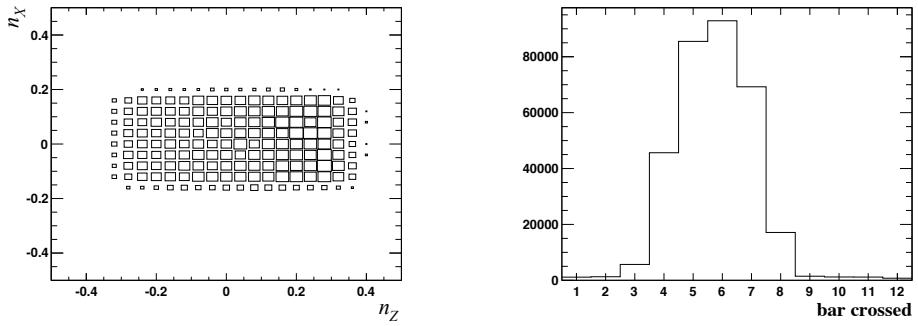


Fig. 5.14: (left) CRT angular acceptance in terms of the track angular components $\hat{n}_X$ and $\hat{n}_Z$ from the Geant4 simulation. The asymmetry in the Z direction is due to the Cherenkov start counter. (right) Distribution of the number identifying the first quartz bar crossed by muon tracks in the CRT acceptance. Central bars are hit the most.

The timing resolution of the start counter was obtained in a test beam, the result is shown in Figure 5.15 (left). However, in order to get the exact time the track passes through a bar, the start time has to be corrected for the Cherenkov photon TOP within the fused silica bar, t_1, as well as for the muon TOF between

the start counter and the bar, t_2. Both t_1 and t_2 depend on the tracking as shown in Figure 5.15 (right). The correction is carried out analytically and was cross-checked with the Geant4 simulation. The resulting time correction is between 250 and 850 ps, therefore it is not negligible if one wants to do a chromatic correction based on timing information.

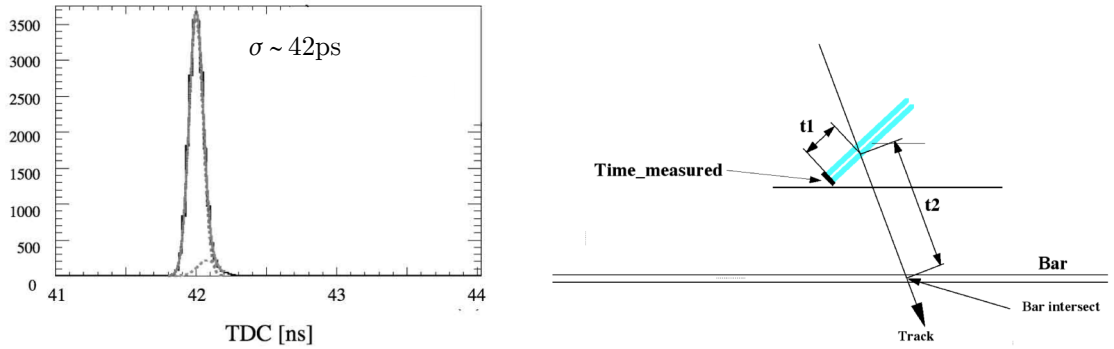


Fig. 5.15: (left) Timing resolution of the start counter as obtained from a test-beam. (right) Scheme of the time corrections needed to get the time at which the muon track passes through a bar.

Time calibration with laser

To calibrate the time offsets of each pixel, the FBLOCK was instrumented with a calibration laser. In particular, a Picosecond Injection Laser (PiLas)[5] was coupled to a diffuser[6] and located at the bottom of the FBLOCK as shown in Figure 5.16. The device provides a relatively uniform spray of 406 nm photons in the detector plane. The intensity is adjusted so that one deals with single photon hits per pixel, and a time offset with respect to the CRT trigger is set so that the laser hits do not overlap with signal from cosmic ray events. This enabled continuous real-time calibration, which was very useful to check the stability of the system over many months of running. The presence of the two reflective sides and the two mirrors make available multiple paths from the diffuser to a pixel which have slightly different time of arrival. Nonetheless, with the present time resolution of ∼600 ps different photon paths are clustered in the same peak in the timing distribution of each pixel. Furthermore, an implementation of the laser in the Geant4 simulation demonstrated that the time spread between mean values of timing distributions of the different pixels is less than 200 ps in the small portion of focal plane instrumented in the FDIRC prototype (see Figure 5.16). Therefore,

[5]PiLas made by Advanced Laser Diode Systems, D-12489 Berlin, Germany
[6]Opal diffuser, 5 mm diameter, P/N K46-162, Edmund Scientific

a simple alignment of the mean values was used to calibrate relative time offsets between pixels. Then, an global time offset is applied to align the expected TOP of Cherenkov photons with the measured one.

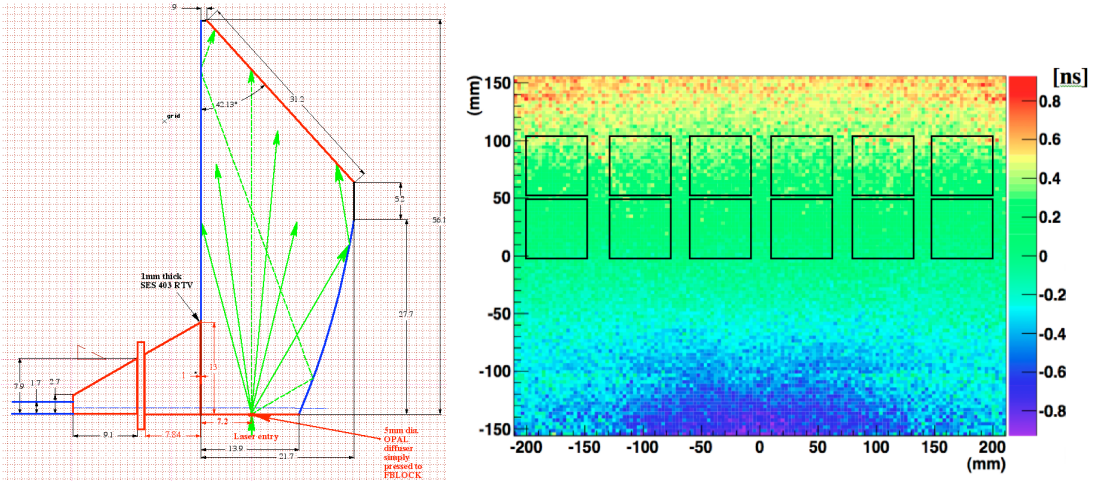


Fig. 5.16: (left) The calibration laser is positioned in the bottom of the FBLOCK and coupled to a diffuser. (right) Mean time of arrival of laser photons as a function of the position in the focal plane. An offset corresponding to the mean time in the PMT area was subtracted. PMT positions are represented by black squares.

5.4.4 Geant4 simulation of the FDIRC detector

The initial design of the FDIRC camera was verified using ray tracing within the Mathematica framework [160]; soon after a Geant4-based [170] standalone simulation of the FDIRC detector was developed [171] to verify its performance. The simulation models all optical elements: bars, glue joints, mirrors, quartz wedge, quartz window and focusing block. Cherenkov photons are simulated with the "optical photon" Geant4 object, which reproduces very carefully the most important optical processes such as bulk absorption, Rayleigh scattering and refraction. Either the full barrel detector (all 12 sectors) or as single sector (the FDIRC prototype) can be simulated. Furthermore, all CRT elements were modelled in the simulation (see Figure 5.13) so that a full simulation of the FDIRC prototype installed at the CRT can be used to make detailed studies. Cosmic muons are fired from the top with the expected distribution in terms of angle and energy. More detailed tuning of the FDIRC prototype include: the modelling of the PMT Quantum Efficiency (QE) as a function of wavelength; the implementation of pixel-by-pixel efficiencies taken from a scan of the 12 PMTs with a laser in a test-bench; a precise modelling of the geometry using "as-built" dimensions rather

than nominal ones. There is no simulation of the electronics response in terms of digitization and waveform processing, but a time jitter is applied to reproduce the timing resolution measured in data.

5.4.5 FDIRC reconstruction

The first task of the FDIRC reconstruction algorithm is to extract the direction of emission $\hat{\mathbf{k}}$ of a Cherenkov photon based on its PMT hit and compute the corresponding Cherenkov angle using:

$$\cos(\theta_C) = \hat{\mathbf{k}} \cdot \hat{\mathbf{n}} \tag{5.5}$$

where $\hat{\mathbf{n}}$ is the direction of the charged track provided by the tracking algorithm.

The Geant4 simulation is essential to the data-analysis since it is used to map PMT pixel positions to the original photon directions $\hat{\mathbf{k}}$ [158]. Lookup tables are filled by firing millions of photons in each bar with random directions and collecting the corresponding PMT hits. For each pixel, different photon paths are formed by clustering hits due to photons coming from close enough directions $\hat{\mathbf{k}}_i$.

Given that the exact number of bounces inside the bar is unknown, the direction components of the photons at bar exit correspond to the original one up to unknown signs $(\pm\hat{k}_X, \pm\hat{k}_Y, \pm\hat{k}_Z)$. Once at bar exit, the complex folded optics of the FDIRC results in more ambiguities in the possible photon directions for a given pixel hit. In particular, from a given bar exit-window, a photon can either go "directly" to the PMT plane or bounce on one of the FBLOCK sides. On top of that, it can either bounce on the top surface of the wedge or not, making a total of $3 \times 2 = 6$ paths per bar per pixel. This is a simplified picture though: for some pixels more or less than 6 paths are available. A way of visualizing and understanding the complex folding is shown in Fig. 5.18. Camera ambiguities are disentangled from sign ambiguities identifying them by the absolute value of k_Y *at bar exit* and the signed value of k_X *at the end of the old wedge*. Indeed, the old wedge is split in 12 pieces in the $X - Z$ plane, just like the bars. However, in the $Y - Z$ plane, the wedge is different from the bar as it has an inclined top reflective surface (Fig. 5.6 right). This is understood when looking at the projections in Figure 5.17. For illustration, all lookup tables entries corresponding to bar number 6 are plotted on the $(k_X, |k_Y|)$ plane in Figure 5.18.

Furthermore, the expected TOP depends on the photon path length and is also subject to ambiguities. Hence, the path length in the camera is stored in the lookup table for each solution. The one in the bar, instead, is calculated analytically and depends only on the component $|k_Z| = \sqrt{k_X^2 + k_Y^2}$, on its sign

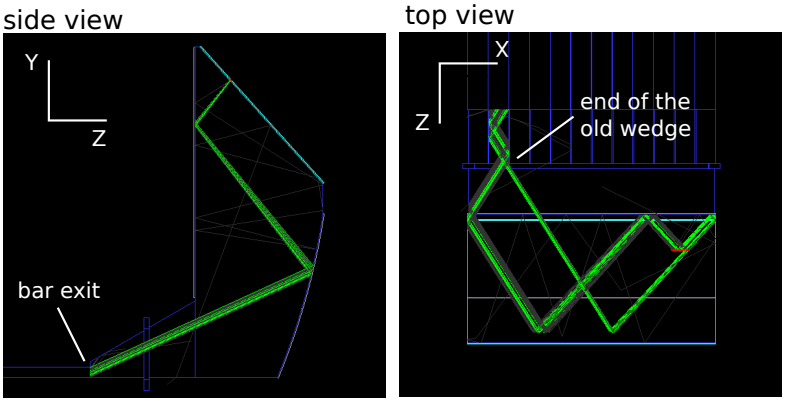


Fig. 5.17: Two projections of the FDIRC imaging camera are shown in blue. Photon paths from a bar exit to a pixel are shown in green. Different paths are defined by the k_Y value *at bar exit* (left) and the k_X value *at the end of the old wedge* (right).

$(\pm)_Z$ and on the distance between the track and the bar exit Δz:

$$d_{\text{bar}} = \frac{\Delta z}{|k_Z|} \qquad \text{for } (\pm)_Z = - \quad \text{“forward photons”} \tag{5.6}$$

$$d_{\text{bar}} = \frac{2L_{\text{bar}} - \Delta z}{|k_Z|} \qquad \text{for } (\pm)_Z = + \quad \text{“backward photons”} \tag{5.7}$$

Therefore, the FDIRC design produces as many as $6 \times 8 = 48$ ambiguities for each pixel hit. Still, it is possible to choose the right solution for most hits. A very powerful method is to compare the expected TOP for each path to the measured one. Some ambiguities, though, have just the opposite sign of k_X or k_Y and therefore the same expected TOP; yet, they normally give non-physical values for θ_C, so they can be rejected.

For the FDIRC prototype test at the CRT, only performances for single photons were evaluated, but actually, a significant ambiguity discrimination power would come from correlating hits from the same track. In practice, this is done by a reconstruction algorithm based on an unbinned maximum likelihood as the one used for the *BABAR* DIRC [149]. Such an algorithm takes as input all photon hits in an event with all ambiguities and returns a likelihood value for each of the five stable charged particle types (e, μ, π, K and p) and for each track. The algorithm also makes an estimation of the number of signal and background photons. In a fully instrumented sector, the number of hits from a Cherenkov event is expected to be between 20 and 60, like in the *BABAR* DIRC, and therefore their combination is expected to contribute significantly to the disentangling of ambiguities.

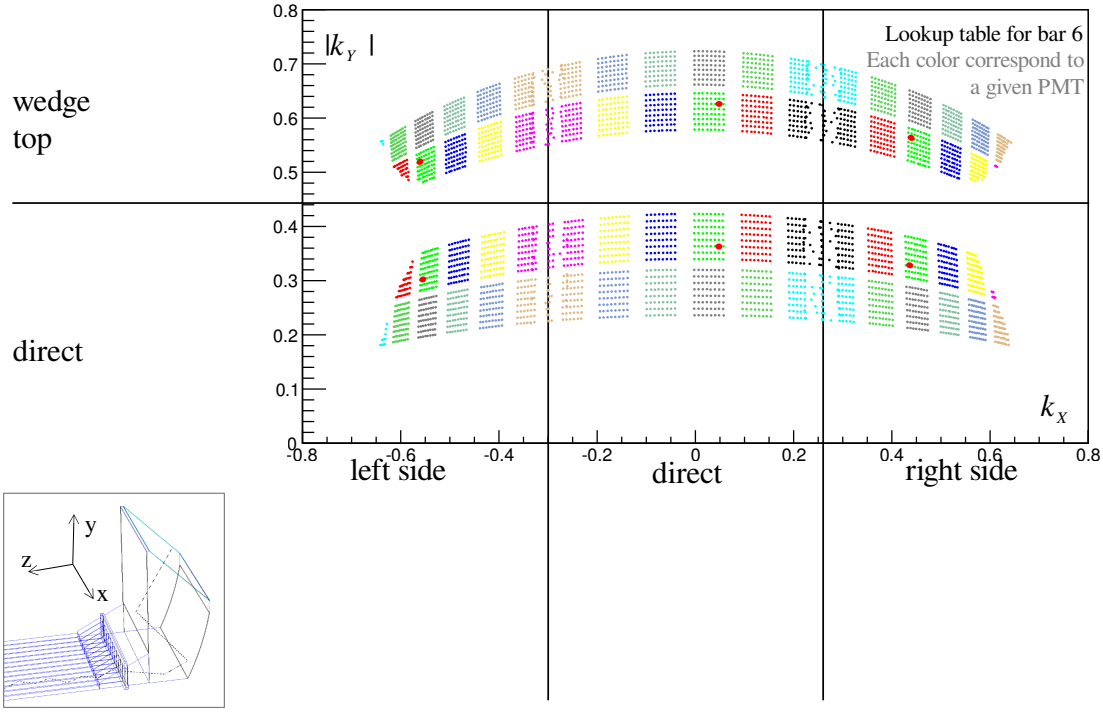


Fig. 5.18: In the centre, a plot of all the possible directions $\hat{\mathbf{k}}$ from the exit of bar number 6 (in the middle of the bar box) to each of the 12×64 pixels. The directions of pixels belonging to different PMTs are represented with dots of different colours. The corresponding photon paths in the FBLOCK are illustrated by the two FBLOCK pictures on the left and the three on the bottom, where photon paths corresponding to the pixel highlighted in red are shown as an example. The bottom left plot is just an illustration of the reference frame used. In the $(k_X, |k_Y|)$ plane, the image of the instrumented plane (composed of 12 PMTs) is mirrored on the left and on the right, because of side reflections. Furthermore, the whole image is reflected on the top resulting in a total of $3 \times 2 = 6$ images of the PMT plane. Ambiguous paths to pixels very close to side mirrors which are or not bouncing on the mirror are clustered together. Indeed, they correspond to very similar directions since there is no focusing in X.

5.4.6 Results from the test of the FDIRC prototype

The FDIRC prototype recorded Cherenkov light from cosmic muons characterised by the SLAC CRT during several months. I participated to the commissioning phase both on site at SLAC and from LAL. Preliminary results came out in Summer 2013 [153].

The performances of the prototype were evaluated thereafter using a large sample of 3×10^5 cosmic muon events [172,173]. Hard muons with energies larger than $\sim$2 GeV are selected by requiring hits in the four scintillators of the iron stack. At this energy, muons have $\beta \simeq 1$ and therefore produce Cherenkov light with the maximal angle for quartz, namely $\theta_C = 822$ mrad for the average detected wavelength. All the information from CRT detectors is combined with the corresponding FDIRC information coming from the processing of IRS2 waveforms. The same data format is produced by the Geant4 simulation so that data and MC can be analysed by the same code.

The main goal of the FDIRC prototype was to validate the new camera design and measure the single-photon Cherenkov angle resolution. In order to minimise the problem of ambiguities, muon tracks are required to have crossed a single bar and the measured TOP is required to correspond to the expected one within a narrow time window. Their difference, $\mathrm{dTOP} = \mathrm{TOP}_{\mathrm{measured}} - \mathrm{TOP}_{\mathrm{expected}}$, is shown in Figure 5.19 separately for forward and backward photons. To obtain a clean data sample to plot dTOP, the Cherenkov angle (at least one of the ambiguities) is required to be within 30 mrad from the expected value for hard muons. The dTOP resolution comes from the following contributions:

$$\sigma^2_{\mathrm{dTOP}} \simeq \sigma^2_{\mathrm{TTS}} + \sigma^2_{\mathrm{elec}} + \sigma^2_{\mathrm{chrom}} + \sigma^2_{\mathrm{geom}} + \sigma^2_{\mathrm{tracking}} \ , \tag{5.8}$$

where the last two contributions are coming from the FDIRC optics and from the CRT tracking respectively and are negligible with respect to the others. As explained above, the photon detection resolution is driven by the IRS2 electronics, $\sigma_{\mathrm{elec}} \sim 600$ ps rather than the Transit Time Spread of the MaPMTs, $\sigma_{\mathrm{TTS}} \sim 140$ ps. The chromatic contribution depends on the photon path length in the quartz, $\sigma_{\mathrm{chrom}} \sim 65$ ps/m. Indeed, as shown in Figure 5.19, the obtained dTOP distribution for backward photons is broader than for forward ones as they propagate for about 10-15 m in the quartz, while forward photons do 2-3 m as shown in Figure 5.20. The same effect is found in MC, as shown in Figure 5.19. These distributions are fitted by the sum of two Gaussian functions in order to extract the width parameter of the narrower one and to roughly model the background from photon path ambiguities with the larger one. Good agreement between data and MC is found for the dTOP distribution width of forward photons, validating the time jitter added to MC hit to simulate the PMT resolution. However, the data distributions show some features in the tails, which may be due to residual imperfect time alignment of pixels. For

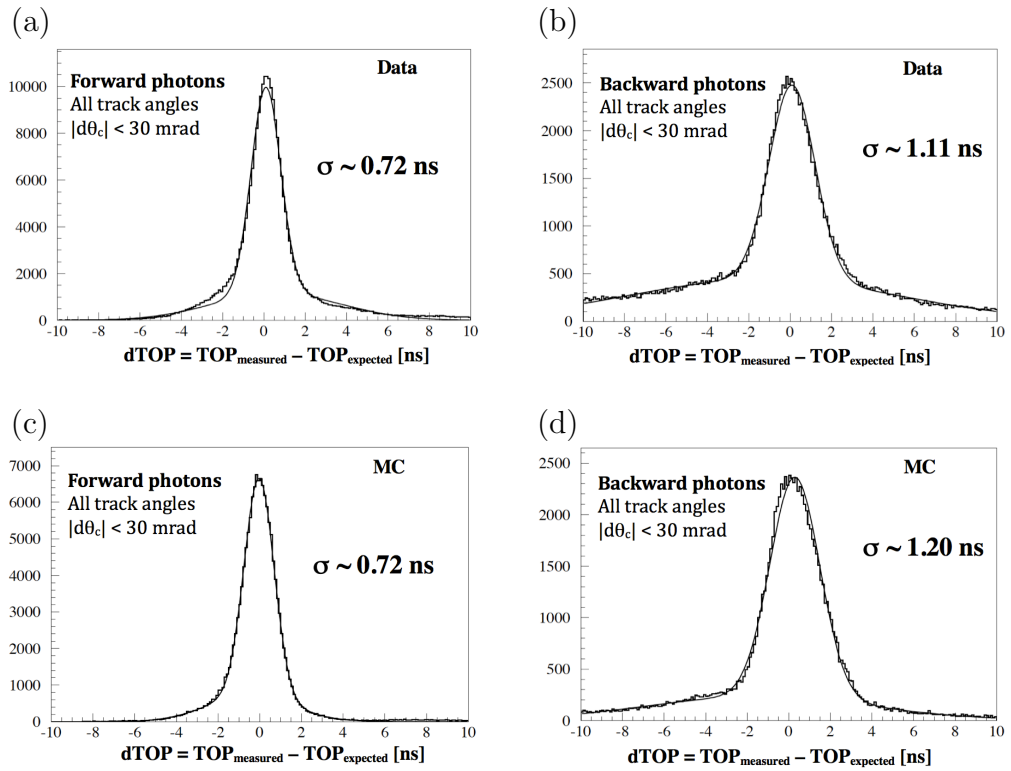


Fig. 5.19: dTOP distributions obtained in data for forward (a) and backward (b) photons. Good photon paths are selected by requiring the Cherenkov angle to be within 30 mrad from the expected value for hard muons. The same distributions coming from the MC are shown for forward (c) and backward (d) photons. Both data and MC show a broadening of the distribution for backward photons due to the chromatic effect.

backward photons, the distribution is more asymmetric and not well modelled by the fit. Nonetheless, the dTOP distribution width is found to be slightly narrower in data, an effect that probably points to an imperfect modelling of the collection efficiency as a function of wavelength. Indeed, various effects may contribute to the detected bandwidth of photons, as shown in Figure 5.10.

Based on these distributions, the window around dTOP is taken as: $|\text{dTOP}| < 2\,\text{ns}$ for forward photons and $|\text{dTOP}| < 2.5\,\text{ns}$ for backward photons [173]. The width of the background photons dTOP distribution is ∼50% larger, but the window is enlarged by only 25% in order to cope with the larger fraction of background. This simple method of dealing with ambiguities allows to get a clean enough and unbiased Cherenkov peak distribution which can be used to estimate the FDIRC single-photon angular resolution. The obtained Cherenkov angle distribution is shown in Figure 5.21. More ambiguities exist for the Cherenkov angle than for the TOP as the first depends also on the sign of the $\hat{k}_X$ and $\hat{k}_Y$ components, while

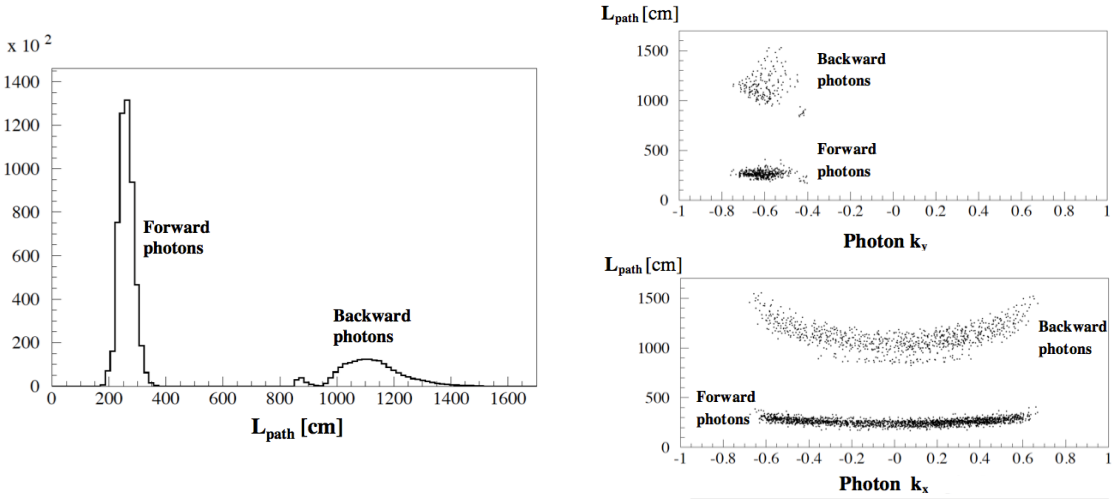


Fig. 5.20: (left) Length of the photon path in the quartz between emission and arrival to a PMT for the CRT data sample. Forward photons are well separated from backward photons. (right) Photon path lengths are plotted as a function of k_X and k_Y.

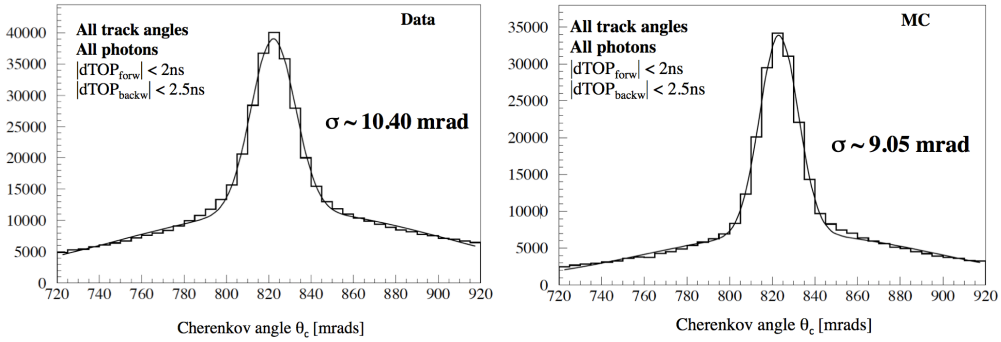


Fig. 5.21: Single photon Cherenkov angle resolution obtained for CRT data (left) and the corresponding Geant4 simulation (right).

the expected TOP depends only on $\hat{k}_Z = \pm\sqrt{\hat{k}_X^2 + \hat{k}_Y^2}$. Actually, since tracks in the CRT acceptance have dip angles in the range $[-350, 350]$ mrad, only the $\hat{k}_Y$ solution with a negative sign gives a physical Cherenkov angle, as shown in Figure 5.20. Furthermore, the k_Y acceptance has a large gap between photon paths bouncing on the wedge top and direct ones (Figure 5.18), making the separation of these two types of path easy by employing the TOP measurement. Actually, most Cherenkov photons from tracks selected by the CRT bounce on the wedge top (large $|\hat{k}_Y|$ region), as shown in Figure 5.20. The θ_C distribution is fitted by the sum of two Gaussian functions, one for the signal and one for the background

coming from ambiguous solutions. The measured resolution is 10.4 mrad, which is somewhat worse than the prediction from MC. The difference is probably due to several small imperfections of the prototype not simulated in the MC. For example, the bars used are expected to have some surface imperfections and the chromatic response, both in terms of wavelength efficiency and chromatic dispersion, may be not perfectly tuned in the MC. Howevever, the Cherenkov angle resolution measured validates the new compact focusing camera design.

Chromatic correction

The chromatic dispersion in the Cherenkov angle of emission is due to the wavelength dependence of the phase refractive index of the radiative medium - see Equation 5.2. The refractive index, n_{p} of fused silica changes by few percents within the FDIRC sensitive range, *i.e.* between 290 and 650 nm. It is larger for shorter wavelength, meaning that "blue" photons are emitted at a smaller Cherenkov angle than "red" photons, giving a chromatic aberration term of ∼5.4 mrads. This aberration is not irreducible: the photon group velocity depends on its wavelength through:

$$v_{\mathrm{g}} = \frac{c}{n_{\mathrm{g}}(\lambda)} \tag{5.9}$$

where c is the speed of light in vacuum, while $n_{\mathrm{g}}(\lambda)$ is the group index of refraction of the medium. Indeed, it is the relation between the phase and group indices of refraction, n_{p} and n_{g}, which correlates θ_C and v_{g}. At first order they are related by the dispersion relation:

$$n_{\mathrm{g}} = n_{\mathrm{p}} - \lambda \frac{\partial n_{\mathrm{p}}}{\partial \lambda} \tag{5.10}$$

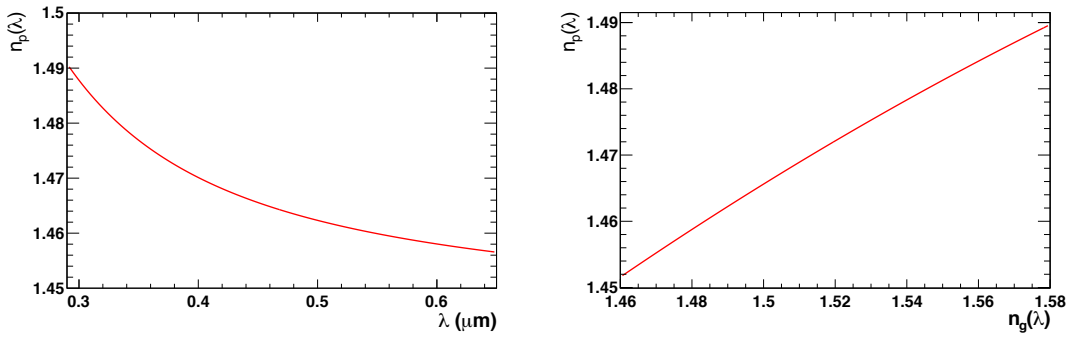


Fig. 5.22: (left) Dependence on the wavelength of the phase refractive index of fused silica. (right) Relation between group and phase refractive indices of fused silica.

Taking the parametrization of $n_{\mathrm{p}}(\lambda)$ for fused silica from the Melles-Griot® catalog[7] shown in Figure 5.22 (left), and computing Equation 5.10, one can obtain the relation between n_{p} and n_{g}, shown in Figure 5.22 (right). In the wavelength range of interest, the relation is close to be linear; n_{p} varies by ∼2.5%, corresponding to a chromatic contribution to the θ_C resolution of 5.4 mrads, and n_{g} by ∼6.5%. A similar linear relation holds between the variations of the Cherenkov angle and of the group velocity:

$$\Delta\theta_C(\lambda) \propto \Delta v_{\mathrm{g}}(\lambda) \propto \frac{\mathrm{dTOP}(\lambda)}{\mathrm{TOP}}. \tag{5.11}$$

Ultimately, this means that to correct chromaticity one needs a measurement of the photon TOP to a precision of at least $6.5\%/\sqrt{12} \simeq 2\%$.
The FDIRC detector can take advantage of the long path Cherenkov photons travel before reaching PMTs, which results in TOPs as long as 20 ns for forward-going photons and 60 ns for backward-going ones (for muon tracks in the CRT acceptance). Therefore, in order to benefit from the chromatic correction, the timing resolution is required to be below 0.4 ns for forward photons and 1.2 ns for backward photons. Indeed, the *BABAR* DIRC could not benefit from this correction as it had a timing resolution of 1.6 ns.

The feasibility of the chromatic correction was proved by the first oil-filled Focusing DIRC prototype [163, 165], which reached a time resolution of 220 ps. The present FDIRC full-scale prototype was tested with an electronics that reached only a 600 ps timing resolution and therefore is capable of correcting significantly the Cherenkov angle only for photon doing long paths. Therefore, in Figure 5.23 the effect of the chromatic correction is shown only for backward photons. Furthermore, for this test only the central part of the ring is selected as it has con-

[7] CVI Melles-Griot, chapter 4: material properties

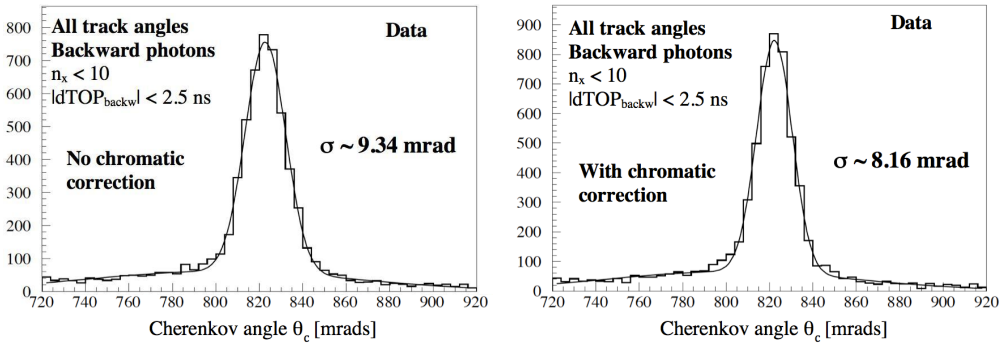


Fig. 5.23: Effect of the chromatic correction on a data sample of backward-going Cherenkov photon belonging to the central part of the ring ($n_X < 10$ requirement). A significant improvement of the θ_C resolution is observed.

siderably better θ_C resolution and is less affected by background from ambiguous photon path solutions. This is done by requiring photon paths with less than 10 bounces on the vertical side surfaces of the bar, $n_X < 10$. In this data sample, the chromatic correction allows the Cherenkov angle resolution to be improved from 9.34 mrad to 8.16 mrad, demonstrating the feasibility of the principle.

5.5 Conclusion and future

The R&D work on the FDIRC detector was presented in this Chapter. A new, fast and compact focusing-camera made of radiation hard fused silica was designed to allow operating the detector with backgrounds two order of magnitude larger than in the *BABAR* experiment. A full scale FDIRC prototype was successfully constructed and tested with cosmic muons with energies greater than 2 GeV selected and characterised by the SLAC CRT. The prototype obtained a single photon Cherenkov angle resolution of 10.40 mrad, which meets the expected performace and validates the design and construction of the new camera.
The focal plane was partly covered with 12 H-8500 MaPMTs having 64 pixels each and fast rise time. Signals were read out with fast digitizing IRS2 electronics. The expected timing performace of the order of 0.2 ns was not achieved due to a variation in the IRS2 amplitude response drifting with the environmental conditions. This test led to the development of a new chip, the IRSX ASIC, which will be used for the TOP counter of the Belle-II experiment.
Nonetheless, the achieved timing resolution of ∼0.6 ns allowed to see a sizeable effect of the chromatic correction on the Cherenkov angle resolution of the prototype. This advanced technique consists in reducing the effect of chromatic broadening of the Chernekov angle using the timing information of single photon and a precise estimation of their path in the fused silica.
Several R&D projects to desing new detectors similar to the FDIRC are ongoing. For example, the GlueX collaboration plans to build a barrel DIRC detector [174] with a design very similar to the FDIRC as shown in Figure 5.24 (left). Thanks to the availability of *BABAR* quartz bars, GlueX plans to reuse one third of them and build fused silica focusing photon cameras to image photons. Also the PANDA collaboration is considering to build a barrel DIRC [175,176] to provide good PID performance to the PANDA experiment at the FAIR facility. Among other options they are considering long fused silica bars coupled to modular cameras built out of the same material.
Furthermore, Time Of Flight (TOF) devices are being developed with a DIRC-like design, such as the TOP counter for Belle-II and the TORCH [177] detector proposed for the upgraded LHCb experiment. The TORCH detector might as well reuse *BABAR* DIRC bars [178] and couple them to focusing cameras as the

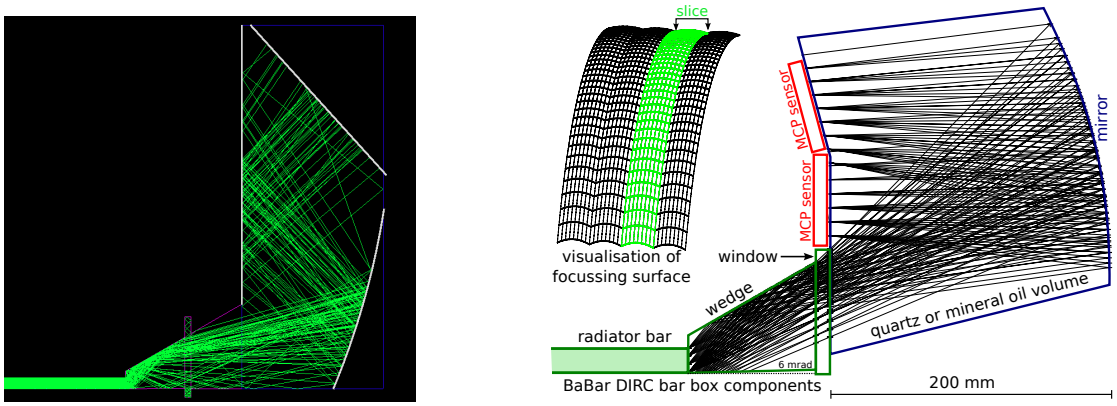


Fig. 5.24: Similar focusing camera design conceived for the GlueX FDIRC (left) and the LHCb TORCH (right) detectors. Both designs shown consider to reuse *BABAR* bars which come with the small wedge attached at bar exit. While the GlueX FDIRC camera design is nearly identical to the FDIRC one, the one for the TORCH detector has only one mirror. Furthermoe, in the design shown, this mirror is splitted in vertical slices with cylindrical section in order to get focusing in the horizontal direction.

one shown in Figure 5.24 (right). Indeed, TORCH plans to measure particle hit time with a resolution of 15 ps employing the prompt Cherenkov light produced in the bar which gives about 30 photo-electrons detected. Therefore, it requires an imaging camera in order to trace back the path of Cherenkov photons in the quartz as well as to correct for the chromatic contribution to the TOP by measuring the Cherenkov angle.

Bibliography

[1] S. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl.Phys.* **22** (1961) 579–588. (Cited on page 9.)

[2] S. Weinberg, *A Model of Leptons*, *Phys.Rev.Lett.* **19** (1967) 1264–1266. (Cited on page 9.)

[3] A. Salam, *Weak and Electromagnetic Interactions*, *Conf.Proc.* **C680519** (1968) 367–377. (Cited on page 9.)

[4] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, *Phys. Rev. Lett.* **10** (1963) 531–533. (Cited on page 12.)

[5] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652–657. (Cited on page 12.)

[6] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, *Phys. Rev. Lett.* **51** (1983) 1945. (Cited on page 12.)

[7] Particle Data Group, K. A. Olive et al., *Review of particle physics*, *Chin. Phys.* **C38** (2014) 090001. (Cited on pages 12, 13, 32, and 69.)

[8] A. Sakharov, *Violation of CP Invariance, c Asymmetry, and Baryon Asymmetry of the Universe*, *Pisma Zh.Eksp.Teor.Fiz.* **5** (1967) 32–35. (Cited on page 13.)

[9] Planck collaboration, P. Ade et al., *Planck 2013 results. XVI. Cosmological parameters*, *Astron.Astrophys.* **571** (2014) A16, [arXiv:1303.5076]. (Cited on page 13.)

[10] M. Dine, W. Fischler, and M. Srednicki, *A Simple Solution to the Strong CP Problem with a Harmless Axion*, *Phys. Lett.* **B104** (1981) 199. (Cited on page 13.)

[11] S. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, *Phys.Rev.* **D2** (1970) 1285–1292. (Cited on page 13.)

[12] K. G. Wilson, *Nonlagrangian models of current algebra*, *Phys.Rev.* **179** (1969) 1499–1512. (Cited on page 14.)

[13] Adapted from a presentation of Thomas Blake, 2014. (Cited on page 16.)

[14] BaBar collaboration, J. Lees et al., *Measurement of the $B \to X_s \ell^+ \ell^-$ branching fraction and search for direct CP violation from a sum of exclusive final states*, *Phys.Rev.Lett.* **112** (2014) 211802, [arXiv:1312.5364]. (Cited on page 15.)

[15] Belle collaboration, Y. Sato et al., *Measurement of the Lepton Forward-Backward Asymmetry in $B \to X_s \ell^+ \ell^-$ Decays with a Sum of Exclusive Modes*, arXiv:1402.7134. (Cited on page 15.)

[16] 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]. (Cited on pages 15 and 17.)

[17] LHCb collaboration, R. Aaij et al., *Differential branching fraction and angular analysis of the decay $B^0 \to K^{*0} \mu^+ \mu^-$*, *JHEP* **08** (2013) 131, [arXiv:1304.6325]. (Cited on pages 15 and 26.)

[18] LHCb collaboration, C. Linn, *Electroweak penguin decays: $b \to s\ell\ell$, talk presented at Flavor Physics & CP violation 2015 (Nagoya, Japan, 25-29 May 2015)* (2015). (Cited on pages 15 and 17.)

[19] LHCb collaboration, R. Aaij et al., *Differential branching fraction and angular analysis of $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ decays*, *JHEP* **1506** (2015) 115, [arXiv:1503.07138]. (Cited on page 15.)

[20] CMS collaboration, S. Chatrchyan et al., *Angular analysis and branching fraction measurement of the decay $B^0 \to K^{*0} \mu^+ \mu^-$*, *Phys.Lett.* **B727** (2013) 77–100, [arXiv:1308.3409]. (Cited on pages 15 and 17.)

[21] ATLAS collaboration, *Angular Analysis of $B_d \to K^{*0} \mu^+ \mu^-$ with the ATLAS Experiment*, Tech. Rep. ATLAS-CONF-2013-038, CERN, Geneva, Apr, 2013. (Cited on pages 15 and 17.)

[22] W. Altmannshofer and D. M. Straub, *New physics in $b \to s$ transitions after LHC run 1*, arXiv:1411.3161. (Cited on pages 15, 17, 19, 20, and 28.)

[23] A. Bharucha, D. M. Straub, and R. Zwicky, *$B \to V \ell^+ \ell^-$ in the Standard Model from Light-Cone Sum Rules*, arXiv:1503.05534. (Cited on pages 17, 18, and 134.)

[24] LHCb collaboration, R. Aaij et al., *Measurement of CP asymmetries in the decays $B^0 \to K^{*0}\mu^+\mu^-$ and $B^+ \to K^+\mu^+\mu^-$*, *JHEP* **09** (2014) 177, [arXiv:1408.0978]. (Cited on page 15.)

[25] C. Bobeth, G. Hiller, and G. Piranishvili, *Angular distributions of $\bar{B} \to K\bar{\ell}\ell$ decays*, *JHEP* **0712** (2007) 040, [arXiv:0709.4174]. (Cited on pages 16 and 19.)

[26] LHCb collaboration, *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay*, LHCb-CONF-2015-002. (Cited on pages 16, 18, 19, 25, and 139.)

[27] 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]. (Cited on page 17.)

[28] A. J. Buras, F. De Fazio, and J. Girrbach, *331 models facing new $b \to s\mu^+\mu^-$ data*, *JHEP* **1402** (2014) 112, [arXiv:1311.6729]. (Cited on page 17.)

[29] W. Altmannshofer, S. Gori, M. Pospelov, and I. Yavin, *Quark flavor transitions in $L_\mu - L_\tau$ models*, *Phys.Rev.* **D89** (2014) 095033, [arXiv:1403.1269]. (Cited on page 17.)

[30] 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]. (Cited on page 17.)

[31] A. Crivellin, G. DAmbrosio, and J. Heeck, *Explaining $h \to \mu^\pm\tau^\mp$, $B \to K^*\mu^+\mu^-$ and $B \to K\mu^+\mu^-/B \to Ke^+e^-$ in a two-Higgs-doublet model with gauged $L_\mu - L_\tau$*, *Phys.Rev.Lett.* **114** (2015) 151801, [arXiv:1501.00993]. (Cited on page 17.)

[32] A. Khodjamirian, T. Mannel, A. Pivovarov, and Y.-M. Wang, *Charm-loop effect in $B \to K^{(*)}\ell^+\ell^-$ and $B \to K^*\gamma$*, *JHEP* **1009** (2010) 089, [arXiv:1006.4945]. (Cited on page 17.)

[33] S. Jager and J. Martin Camalich, *On $B \to V\ell\ell$ at small dilepton invariant mass, power corrections, and new physics*, *JHEP* **1305** (2013) 043, [arXiv:1212.2263]. (Cited on pages 17, 79, and 129.)

[34] J. Lyon and R. Zwicky, *Resonances gone topsy turvy - the charm of QCD or new physics in $b \to s\ell^+\ell^-$?*, arXiv:1406.0566. (Cited on page 17.)

[35] 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*, arXiv:1412.3183. (Cited on pages 17, 24, and 134.)

[36] BaBar collaboration, B. Aubert et al., *Measurements of branching fractions, rate asymmetries, and angular distributions in the rare decays $B \to K\ell^+\ell^-$ and $B \to K^*\ell^+\ell^-$*, *Phys.Rev.* **D73** (2006) 092001, [arXiv:hep-ex/0604007]. (Cited on page 17.)

[37] Belle collaboration, J.-T. Wei et al., *Measurement of the Differential Branching Fraction and Forward-Backword Asymmetry for B —¿ K(*)l+l-*, *Phys.Rev.Lett.* **103** (2009) 171801, [arXiv:0904.0770]. (Cited on page 17.)

[38] CDF collaboration, T. Aaltonen et al., *Measurements of the Angular Distributions in the Decays $B \to K^{(*)}\mu^+\mu^-$ at CDF*, *Phys.Rev.Lett.* **108** (2012) 081807, [arXiv:1108.0695]. (Cited on page 17.)

[39] F. Beaujean, M. Chrzszcz, N. Serra, and D. van Dyk, *Extracting Angular Observables without a Likelihood and Applications to Rare Decays*, arXiv:1503.04100. (Cited on page 17.)

[40] U. Egede, M. Patel, and K. A. Petridis, *Method for an unbinned measurement of the q^2 dependent decay amplitudes of $\bar{B}^0 \to \bar{K}^{*0}\mu^+\mu^-$ decays*, arXiv:1504.00574. (Cited on page 17.)

[41] C. Bobeth et al., *$B_{s,d} \to \ell^+\ell^-$ in the Standard Model with Reduced Theoretical Uncertainty*, *Phys.Rev.Lett.* **112** (2014) 101801, [arXiv:1311.0903]. (Cited on page 18.)

[42] 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*, arXiv:1411.4413. (Cited on page 18.)

[43] 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]. (Cited on page 19.)

[44] S. L. Glashow, D. Guadagnoli, and K. Lane, *Lepton Flavor Violation in B Decays?*, *Phys.Rev.Lett.* **114** (2015) 091801, [arXiv:1411.0565]. (Cited on page 19.)

[45] BaBar collaboration, J. Lees et al., *Measurement of an Excess of $\bar{B} \to D^{(*)}\tau^-\bar{\nu}_\tau$ Decays and Implications for Charged Higgs Bosons*, *Phys.Rev.* **D88** (2013), no. 7 072012, [arXiv:1303.0571]. (Cited on page 19.)

[46] Belle collaboration, T. Kuhr, *New $B \to D^{(*)}\tau\nu$ Result from Belle*, *talk presented at Flavor Physics & CP violation 2015 (Nagoya, Japan, 25-29 May 2015)* (2015). (Cited on page 19.)

[47] LHCb collaboration, G. Ciezarek, *Semileptonic B baryon decays and $B \to D^*\tau\nu$, talk presented at Flavor Physics & CP violation 2015 (Nagoya, Japan, 25-29 May 2015)* (2015). (Cited on page 19.)

[48] B. Bhattacharya, A. Datta, D. London, and S. Shivashankara, *Simultaneous Explanation of the R_K and $R(D^{(*)})$ Puzzles*, *Phys.Lett.* **B742** (2015) 370–374, [arXiv:1412.7164]. (Cited on page 19.)

[49] J. Gratrex, M. Hopfer, and R. Zwicky, *Generalised helicity formalism, higher moments and the $B \to K_{J_K}(\to K\pi)\bar{\ell}_1\ell_2$ angular distributions*, arXiv:1506.03970. (Cited on page 19.)

[50] G. Hiller and M. Schmaltz, *Diagnosing lepton-nonuniversality in $b \to s\ell\ell$*, *JHEP* **1502** (2015) 055, [arXiv:1411.4773]. (Cited on page 20.)

[51] Bečirević, Damir and Fajfer, Svjetlana and Košnik, Nejc, *Lepton flavor nonuniversality in $b \to s\ell^+\ell^-$ processes*, arXiv:1503.09024. (Cited on page 20.)

[52] Heavy Flavor Averaging Group (HFAG), Y. Amhis et al., *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2014*, arXiv:1412.7515. (Cited on pages 21, 22, 135, 136, and 137.)

[53] M. Misiak et al., *Updated NNLO QCD predictions for the weak radiative B-meson decays*, *Phys.Rev.Lett.* **114** (2015), no. 22 221801, [arXiv:1503.01789]. (Cited on pages 21 and 135.)

[54] P. Ball, G. W. Jones, and R. Zwicky, *$B \to V\gamma$ beyond QCD factorisation*, *Phys.Rev.* **D75** (2007) 054004, [arXiv:hep-ph/0612081]. (Cited on pages 21 and 22.)

[55] B. Grinstein, Y. Grossman, Z. Ligeti, and D. Pirjol, *Photon polarization in $B \to X\gamma$ in the standard model*, *Phys.Rev.* **D71** (2005) 011504, [arXiv:hep-ph/0412019]. (Cited on page 21.)

[56] L. L. Everett, G. L. Kane, S. Rigolin, L.-T. Wang, and T. T. Wang, *Alternative approach to $b \to s\gamma$ in the uMSSM*, *JHEP* **01** (2002) 022, [arXiv:hep-ph/0112126]. (Cited on page 21.)

[57] J. Foster, K.-i. Okumura, and L. Roszkowski, *New constraints on SUSY flavour mixing in light of recent measurements at the Tevatron*, *Phys.Lett.* **B641** (2006) 452–460, [arXiv:hep-ph/0604121]. (Cited on page 21.)

[58] E. Lunghi and J. Matias, *Huge right-handed current effects in $B \to K^{*0}(K\pi)\ell^+\ell^-$ in supersymmetry*, *JHEP* **04** (2007) 058, [arXiv:hep-ph/0612166]. (Cited on page 21.)

[59] E. Kou, C.-D. L, and F.-S. Yu, *Photon Polarization in the $b \to s\gamma$ processes in the Left-Right Symmetric Model*, *JHEP* **1312** (2013) 102, [arXiv:1305.3173]. (Cited on page 21.)

[60] T. Goto, Y. Okada, T. Shindou, and M. Tanaka, *Patterns of flavor signals in supersymmetric models*, *Phys.Rev.* **D77** (2008) 095010, [arXiv:0711.2935]. (Cited on page 21.)

[61] Bečirević, Damir and Kou, Emi and Le Yaouanc, Alain and Tayduganov, Andrey, *Future prospects for the determination of the Wilson coefficient $C'_{7\gamma}$*, *JHEP* **1208** (2012) 090, [arXiv:1206.1502]. (Cited on pages 22, 23, and 135.)

[62] BaBar collaboration, B. Aubert et al., *Measurement of time-dependent CP asymmetry in $B^0 \to K^0_S\pi^0\gamma$ decays*, *Phys.Rev.* **D78** (2008) 071102, [arXiv:0807.3103]. (Cited on pages 22, 135, 136, and 137.)

[63] Belle collaboration, Y. Ushiroda et al., *Time-dependent CP asymmetries in $B^0 \to K^0_S\pi^0\gamma$ transitions*, *Phys.Rev.* **D74** (2006) 111104, [arXiv:hep-ex/0608017]. (Cited on pages 22, 135, 136, and 137.)

[64] Belle collaboration, J. Li et al., *Time-dependent CP Asymmetries in $B^0 \to K^0_S\rho^0\gamma$ Decays*, *Phys.Rev.Lett.* **101** (2008) 251601, [arXiv:0806.1980]. (Cited on page 22.)

[65] BaBar collaboration, S. Akar, *Penguin and rare decays in BABAR*, *J.Phys.Conf.Ser.* **556** (2014), no. 1 012047, [arXiv:1410.6897]. (Cited on page 22.)

[66] F. Muheim, Y. Xie, and R. Zwicky, *Exploiting the width difference in $B_s \to \phi\gamma$*, *Phys.Lett.* **B664** (2008) 174–179, [arXiv:0802.0876]. (Cited on page 22.)

[67] LHCb collaboration, R. Aaij et al., *Measurement of the ratio of branching fractions $\mathcal{B}(B^0 \to K^{*0}\gamma)/\mathcal{B}(B^0_s \to \phi\gamma)$ and the direct CP asymmetry in $B^0 \to K^{*0}\gamma$*, *Nucl. Phys.* **B867** (2013) 1–18, [arXiv:1209.0313]. (Cited on page 23.)

[68] F. Soomro, *Radiative decays of B mesons at LHCb*, CERN-THESIS-2011-035. (Cited on page 23.)

[69] M. Gronau, Y. Grossman, D. Pirjol, and A. Ryd, *Measuring the photon polarization in $B \to K\pi\pi\gamma$*, *Phys.Rev.Lett.* **88** (2002) 051802, [arXiv:hep-ph/0107254]. (Cited on page 23.)

[70] E. Kou, A. Le Yaouanc, and A. Tayduganov, *Determining the photon polarization of the $b \to s\gamma$ using the $B \to K_1(1270)\gamma \to (K\pi\pi)\gamma$ decay*, *Phys.Rev.* **D83** (2011) 094007, [arXiv:1011.6593]. (Cited on page 23.)

[71] LHCb collaboration, R. Aaij et al., *Observation of photon polarization in the $b \to s\gamma$ transition*, *Phys. Rev. Lett.* **112** (2014) 161801, [arXiv:1402.6852]. (Cited on page 23.)

[72] LHCb collaboration, R. Aaij et al., *Angular analysis of the $B^0 \to K^{*0}e^+e$ decay in the low-q^2 region*, *JHEP* **1504** (2015) 064, [arXiv:1501.03038]. (Cited on pages 24, 25, and 134.)

[73] Y. Grossman and D. Pirjol, *Extracting and using photon polarization information in radiative B decays*, *JHEP* **0006** (2000) 029, [arXiv:hep-ph/0005069]. (Cited on pages 24 and 29.)

[74] F. Kruger and J. Matias, *Probing new physics via the transverse amplitudes of $B^0 \to K^{*0}(\to K^+\pi^-)\ell^+\ell^-$ at large recoil*, *Phys.Rev.* **D71** (2005) 094009, [arXiv:hep-ph/0502060]. (Cited on pages 24 and 27.)

[75] Bečirević, Damir and Schneider, Elia, *On transverse asymmetries in $B \to K^*\ell^+\ell^-$*, *Nucl. Phys.* **B854** (2012) 321–339, [arXiv:1106.3283]. (Cited on pages 24, 27, 28, and 29.)

[76] J. Lefrancois and M. H. Schune, *Measuring the photon polarization in $b \to s\gamma$ using the $B^0 \to e^+e^-K^{*0}$ decay channel*, Tech. Rep. LHCb-PUB-2009-008. CERN-LHCb-PUB-2009-008. LHCb-INT-2009-007, CERN, Geneva, Jun, 2009. (Cited on page 24.)

[77] W. Altmannshofer et al., *Symmetries and Asymmetries of $B \to K^*\mu^+\mu^-$ Decays in the Standard Model and Beyond*, *JHEP* **0901** (2009) 019, [arXiv:0811.1214]. (Cited on page 26.)

[78] C. Bobeth, G. Hiller, and G. Piranishvili, *CP asymmetries in $\bar{B} \to \bar{K}^*(\to \bar{K}\pi)\bar{\ell}\ell$ and untagged $\bar{B}_s$, $B_s \to \phi(\to K^+K^-)\bar{\ell}\ell$ decays at NLO*, *JHEP* **07** (2008) 106, [arXiv:0805.2525]. (Cited on page 27.)

[79] C.-D. Lu and W. Wang, *Analysis of $B \to K_J^*(\to K\pi)\mu^+\mu^-$ in the higher kaon resonance region*, *Phys.Rev.* **D85** (2012) 034014, [arXiv:1111.1513]. (Cited on page 27.)

[80] 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* **1301** (2013) 048, [arXiv:1207.2753]. (Cited on page 28.)

[81] G. C. Wick, *Detection of gamma-ray polarization by pair production*, *Phys. Rev.* **81** (Feb, 1951) 467–468. (Cited on pages 29 and 30.)

[82] H. Bethe and W. Heitler, *On the Stopping of fast particles and on the creation of positive electrons*, *Proc.Roy.Soc.Lond.* **A146** (1934) 83–112. (Cited on page 29.)

[83] C. Von Weizsacker, *Radiation emitted in collisions of very fast electrons*, *Z. Phys* **88** (1934), no. 612 95; E. Williams, *Kgl. danske videnskab. selskab mat.-fys. medd. 13, no. 4 (1935); e. fermi*, *Z. Phys* **29** (1924) 315–327. (Cited on page 30.)

[84] L. C. Maximon and H. Olsen, *Measurement of linear photon polarization by pair production*, *Phys. Rev.* **126** (Apr, 1962) 310–319. (Cited on page 30.)

[85] B. Rossi and K. Greisen, *Cosmic-ray theory*, *Rev. Mod. Phys.* **13** (1941) 240–309. (Cited on page 32.)

[86] G. R. Lynch and O. I. Dahl, *Approximations to multiple Coulomb scattering*, *Nucl. Instrum. Meth.* **B58** (1991) 6–10. (Cited on page 32.)

[87] D. Bernard, *Polarimetry of cosmic gamma-ray sources above e^+e^- pair creation threshold*, *Nucl.Instrum.Meth.* **A729** (2013) 765–780, [arXiv:1307.3892]. (Cited on page 32.)

[88] S. D. Hunter et al., *A Pair Production Telescope for Medium-Energy Gamma-Ray Polarimetry*, *Astropart.Phys.* **59** (2014) 18–28, [arXiv:1311.2059]. (Cited on page 32.)

[89] D. Bernard et al., *HARPO: a TPC as a gamma-ray telescope and polarimeter*, *Proc.SPIE Int.Soc.Opt.Eng.* **9144** (2014) 91441M, [arXiv:1406.4830]. (Cited on page 32.)

[90] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001. (Cited on page 33.)

[91] ATLAS collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003. (Cited on page 33.)

[92] CMS collaboration, S. Chatrchyan et al., *The CMS experiment at the CERN LHC*, *JINST* **3** (2008) S08004. (Cited on page 33.)

[93] ATLAS collaboration, 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) 1–29, [arXiv:1207.7214]. (Cited on page 33.)

[94] CMS collaboration, 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) 30–61, [arXiv:1207.7235]. (Cited on page 33.)

[95] LHCb collaboration, A. A. Alves Jr. et al., *The LHCb detector at the LHC*, *JINST* **3** (2008) S08005. (Cited on pages 34, 41, and 45.)

[96] ALICE collaboration, K. Aamodt et al., *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008) S08002. (Cited on page 34.)

[97] TOTEM collaboration, G. Anelli et al., *The TOTEM experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08007. (Cited on page 34.)

[98] LHCf collaboration, O. Adriani et al., *The LHCf detector at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08006. (Cited on page 34.)

[99] MoEDAL collaboration, J. Pinfold et al., *Technical Design Report of the MoEDAL Experiment*, . (Cited on page 34.)

[100] R. Lindner, *LHCb layout*, LHCb Collection., Feb, 2008. (Cited on page 35.)

[101] LHCb collaboration, R. Aaij et al., *Prompt charm production in pp collisions at* $\sqrt{s} = 7\,\text{TeV}$, *Nucl.Phys.* **B871** (2013) 1–20, [arXiv:1302.2864]. (Cited on page 36.)

[102] 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) 209–216, [arXiv:1009.2731]. (Cited on page 36.)

[103] BaBar collaboration, B. et al., *The BaBar detector*, *Nucl. Instrum. Meth. A* **479** (2002), no. 1 1–116. (Cited on page 37.)

[104] A. Abashian et al., *The Belle Detector*, *Nucl.Instrum.Meth.* **A479** (2002) 117–232. (Cited on page 37.)

[105] LHCb collaboration, R. Aaij et al., *LHCb Detector Performance*, *Int.J.Mod.Phys.* **A30** (2015), no. 07 1530022, [arXiv:1412.6352]. (Cited on pages 37, 38, 41, and 45.)

[106] R. Aaij et al., *Performance of the LHCb Vertex Locator*, *JINST* **9** (2014) 09007, [arXiv:1405.7808]. (Cited on page 38.)

[107] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011. LHCB-TDR-005. (Cited on page 39.)

[108] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029. LHCB-TDR-008. (Cited on page 40.)

[109] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024. LHCB-TDR-006. (Cited on page 40.)

[110] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037. LHCB-TDR-003. (Cited on page 42.)

[111] M. Adinolfi et al., *Performance of the LHCb RICH detector at the LHC*, *Eur. Phys. J.* **C73** (2013) 2431, [arXiv:1211.6759]. (Cited on pages 43 and 143.)

[112] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036. LHCB-TDR-002. (Cited on page 44.)

[113] G. David et al., *Performance of the PHENIX EM calorimeter*, *IEEE Trans.Nucl.Sci.* **43** (1996) 1491–1495; J. Badier et al., *Shashlik calorimeter: Beam test results*, *Nucl.Instrum.Meth.* **A348** (1994) 74–86. (Cited on page 44.)

[114] A. Martin Sanchez, *CP violation studies on the $B^0 \to DK^{*0}$ decays and hadronic trigger performance with the LHCb detector at CERN*, CERN-THESIS-2013-311. (Cited on page 45.)

[115] LHCb collaboration, R. Aaij et al., *Measurement of the ratio of branching fractions $\mathcal{B}(B^0 \to K^{*0}\gamma)/\mathcal{B}(B^0_s \to \phi\gamma)$ and the direct CP asymmetry in $B^0 \to K^{*0}\gamma$*, *Nucl.Phys.* **B867** (2013) 1–18, [arXiv:1209.0313]. (Cited on pages 44 and 45.)

[116] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010. LHCB-TDR-004. (Cited on page 46.)

[117] LHCb collaboration, *LHCb trigger system: Technical Design Report*, CERN-LHCC-2003-031. LHCB-TDR-010. (Cited on page 47.)

[118] T. Sjöstrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 physics and manual*, *JHEP* **05** (2006) 026, [arXiv:hep-ph/0603175]; T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, *Comput.Phys.Commun.* **178** (2008) 852–867, [arXiv:0710.3820]. (Cited on page 48.)

[119] I. Belyaev et al., *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, *Nuclear Science Symposium Conference Record (NSS/MIC)* **IEEE** (2010) 1155. (Cited on page 48.)

[120] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth.* **A462** (2001) 152–155. (Cited on page 48.)

[121] P. Golonka and Z. Was, *PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays*, *Eur.Phys.J.* **C45** (2006) 97–107, [arXiv:hep-ph/0506026]. (Cited on page 48.)

[122] 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.* (2003) 250. (Cited on pages 49 and 73.)

[123] M. Clemencic et al., *The LHCb simulation application, Gauss: Design, evolution and experience*, *J. Phys. Conf. Ser.* **331** (2011) 032023. (Cited on page 49.)

[124] LHCb collaboration, R. Aaij et al., *Measurement of the $B^0 \to K^{*0}e^+e^-$ branching fraction at low dilepton mass*, *JHEP* **05** (2013) 159, [arXiv:1304.3035]. (Cited on page 53.)

[125] 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]. (Cited on page 54.)

[126] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and regression trees*. Wadsworth international group, Belmont, California, USA, 1984. (Cited on pages 55 and 59.)

[127] R. E. Schapire and Y. Freund, *A decision-theoretic generalization of on-line learning and an application to boosting*, *Jour. Comp. and Syst. Sc.* **55** (1997) 119. (Cited on pages 55 and 59.)

[128] M. Nicol, *Analysis of the rare decay $B^0 \to K^{*0}e^+e^-$ at LHCb*, CERN-THESIS-2012-289. (Cited on page 55.)

[129] C. Prouve, *Analysis of the $B^0 \to K^{*0}e^+e^-$ decay at LHCb*, Master Thesis, July, 2012. (Cited on page 55.)

[130] Particle Data Group, J. Beringer et al., *Review of Particle Physics (RPP)*, *Phys.Rev.* **D86** (2012) 010001. (Cited on pages 56, 76, and 79.)

[131] T. Skwarnicki, *A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances*. PhD thesis, Institute of Nuclear Physics, Krakow, 1986. DESY-F31-86-02. (Cited on pages 56 and 85.)

[132] M. Pivk and F. L. Diberder, *: A statistical tool to unfold data distributions*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **555** (2005), no. 12 356 – 369. (Cited on page 56.)

[133] CDF collaboration, A. Abulencia et al., *Search for $B_s \to \mu^+\mu^-$ and $B_d \to \mu^+\mu^-$ decays in $p\bar{p}$ collisions with CDF II*, *Phys.Rev.Lett.* **95** (2005) 221805, [arXiv:hep-ex/0508036]. (Cited on page 57.)

[134] LHCb collaboration, R. Aaij et al., *Measurement of the $B^0 \to K^{*0}e^+e^-$ branching fraction at low dilepton mass*, *JHEP* **1305** (2013) 159, [arXiv:1304.3035]. (Cited on page 65.)

[135] G. Patrick and L. Urbán, *GEANT: detector description and simulation tool, long writeup W5013 - PHYS211*. CERN Program Library. CERN, Geneva, 1993. (Cited on page 73.)

[136] A. Borsellino, *Momentum Transfer and Angle of Divergence of Pairs Produced by Photons*, *Phys.Rev.* **89** (1953) 1023–1025. (Cited on page 73.)

[137] J. He, J. Lefrançois, M. Nicol, and M.-H. Schune, *Measurement of the $B^0 \to K^{*0}e^+e^-$ branching fraction* , . Linked to LHCb-PAPER-2013-005. (Cited on page 79.)

[138] A. Y. Korchin and V. A. Kovalchuk, *Contribution of low-lying vector resonances to polarization observables in $B^0 \to K^{*0}e^+e^-$ decay*, *Phys.Rev.* **D82** (2010) 034013, [arXiv:1004.3647]. (Cited on pages 79 and 129.)

[139] Polci F. and Coquereau S. Private communication, 2014. (Cited on page 79.)

[140] J. Albrecht et al., *Measurement of the ratio $R_K = \mathcal{B}(B^+ \to K^+\mu^+\mu^-)/\mathcal{B}(B^+ \to K^+e^+e^-)$*, Tech. Rep. LHCb-ANA-2011-112, CERN, Geneva, January, 2014. (Cited on page 85.)

[141] K. S. Cranmer, *Kernel estimation in high-energy physics*, *Comput.Phys.Commun.* **136** (2001) 198–207, [arXiv:hep-ex/0011057]. (Cited on pages 91 and 92.)

[142] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl.Instrum.Meth.* **A389** (1997) 81–86. (Cited on page 91.)

[143] LHCb collaboration, *Framework TDR for the LHCb Upgrade: Technical Design Report*, CERN-LHCC-2012-007. LHCB-TDR-012. (Cited on page 138.)

[144] BaBar collaboration, B. Aubert et al., *The BaBar Detector: Upgrades, Operation and Performance*, *Nucl.Instrum.Meth.* **A729** (2013) 615–701, [arXiv:1305.3560]. (Cited on pages 142 and 144.)

[145] P. Čerenkov, *Visible radiation produced by electrons moving in a medium with velocities exceeding that of light*, *Physical Review* **52** (1937), no. 4 378. (Cited on page 143.)

[146] I. Frank and I. Tamm, *Coherent visible radiation of fast electrons passing through matter*, in *Dokl. Akad. Nauk SSSR*, vol. 14, pp. 109–114, 1937. (Cited on page 143.)

[147] A. Roberts, *A new type of Cerenkov detector for the accurate measurement of particle velocity and direction*, *Nucl.Instrum.Meth.* **9** (1960) 55–66. (Cited on page 143.)

[148] J. V. Jelley, *Detection of μ-mesons and other fast charged particles in cosmic radiation, by the cerenkov effect in distilled water*, *Proceedings of the Physical Society. Section A* **64** (1951), no. 1 82. (Cited on page 143.)

[149] BaBar DIRC group, I. Adam et al., *The DIRC particle identification system for the BaBar experiment*, *Nucl.Instrum.Meth.* **A538** (2005) 281–357. (Cited on pages 143, 144, 145, 146, 147, and 162.)

[150] B. Ratcliff, *The B factory detector for PEP-II: A Status report*, *AIP Conf.Proc.* **272** (1993) 1889–1896. (Cited on page 144.)

[151] B. Ratcliff, *Imaging rings in ring imaging Cherenkov counters*, *Nucl.Instrum.Meth.* **A502** (2003) 211–221. (Cited on page 144.)

[152] BaBar collaboration, B. Aubert et al., *The BaBar detector*, *Nucl.Instrum.Meth.* **A479** (2002) 1–116, [arXiv:hep-ex/0105044]. (Cited on page 144.)

[153] M. Borsato et al., *First results of the Focusing DIRC prototype, an innovative detector for charged particle identication*, *PoS* **EPS-HEP2013** (2013) 489. (Cited on pages 148 and 164.)

[154] C. Field et al., *Novel photon detectors for focusing DIRC prototype*, *Nucl.Instrum.Meth.* **A518** (2004), no. 1-2 565–568. (Cited on pages 148 and 155.)

[155] C. Field et al., *Development of photon detectors for a fast focusing DIRC*, *Nucl.Instrum.Meth.* **A553** (2005), no. 1 96–106. (Cited on pages 148, 154, and 155.)

[156] SuperB collaboration, M. Baszczyk et al., *SuperB Technical Design Report*, arXiv:1306.5655. (Cited on page 149.)

[157] Belle-II collaboration, T. Abe et al., *Belle II Technical Design Report*, arXiv:1011.0352. (Cited on page 149.)

[158] M. Borsato et al., *The focusing DIRC: An innovative PID detector*, *Nucl.Instrum.Meth.* **A732** (2013) 333–337. (Cited on pages 149 and 161.)

[159] J. Va'vra, D. Roberts, and B. Ratcliff, *FDIRC design for SuperB*, *Nucl.Instrum.Meth.* **A639** (2011) 282–286. (Cited on page 149.)

[160] J. Va'vra, *Simulation of the focusing DIRC optics with mathematica*, in *Nuclear Science Symposium Conference Record, 2008. NSS'08. IEEE*, pp. 2408–2412, IEEE, 2008. (Cited on pages 150, 151, and 160.)

[161] M. Hoek et al., *The PANDA Barrel DIRC detector*, *Nucl.Instrum.Meth.* **A766** (2014) 9–13. (Cited on page 151.)

[162] J. Va'vra et al., *The Focusing DIRC - the first RICH detector to correct the chromatic error by timing, and the development of a new TOF detector concept*, *Nucl.Instrum.Meth.* (2007). (Cited on page 153.)

[163] J. Benitez et al., *Status of the Fast Focusing DIRC (fDIRC)*, *Nucl.Instrum.Meth.* **A595** (2008) 104–107. (Cited on pages 153 and 168.)

[164] J. Va'vra, *Slac cosmic ray telescope facility*, January, 2010. SLAC-PUB-13873. (Cited on pages 153 and 157.)

[165] K. Nishimura et al., *A Detailed Study of FDIRC Prototype with Waveform Digitizing Electronics in Cosmic Ray Telescope Using 3D Tracks*, *Nucl.Instrum.Meth.* **A701** (2013) 115–126. (Cited on pages 153 and 168.)

[166] J. Va'vra et al., *The focusing DIRC with waveform digitizing electronics*, *Proceedings of the TIPP09 International Conference, 2009.* (Cited on page 155.)

[167] C. Beigbeder et al., *Front-end electronics for the SuperB charged particle identification detectors*, *Nucl.Instrum.Meth.* **A718** (2013) 186–188. (Cited on page 155.)

[168] G. Varner, L. Ruckman, and A. Wong, *The first version buffered large analog bandwidth (blab1) asic for high luminosity collider and extensive radio neutrino detectors*, *Nucl.Instrum.Meth.* **A591** (2008), no. 3 534–545. (Cited on page 155.)

[169] D. Breton et al., *High resolution photon timing with mcp-pmts: A comparison of a commercial constant fraction discriminator (cfd) with the asic-based waveform digitizers target and wavecatcher*, *Nucl.Instrum.Meth.* **A629** (2010) 123–132. (Cited on page 155.)

[170] GEANT4 collaboration, S. Agostinelli et al., *GEANT4: A simulation toolkit*, *Nucl. Instrum. Meth.* **A506** (2003) 250–303. (Cited on page 160.)

[171] D. Roberts, *Geant 4 model of FDIRC*, in *SuperB workshop at SLAC*, October, 2009. (Cited on page 160.)

[172] D. Roberts et al., *Results from the FDIRC prototype*, *Nucl.Instrum.Meth.* **A766** (2014) 114–117. (Cited on page 164.)

[173] B. Dey et al., *Design and performance of the Focusing DIRC detector*, *Nucl.Instrum.Meth.* **A775** (2014) 112–131, [arXiv:1410.0075]. (Cited on pages 164 and 165.)

[174] GlueX collaboration, M. Dugger et al., *A study of decays to strange final states with GlueX in Hall D using components of the BaBar DIRC*, arXiv:1408.0215. (Cited on page 169.)

[175] PANDA Cherenkov collaboration, J. Schwiening, *The barrel DIRC detector for the PANDA experiment at FAIR*, *Nucl.Instrum.Meth.* **A639** (2011) 315–318. (Cited on page 169.)

[176] A. Lehmann, *The DIRC Detectors at the PANDA Experiment*, *PoS* **TIPP2014** (2015) 112. (Cited on page 169.)

[177] LHCb collaboration, M. Charles and R. Forty, *TORCH: Time of Flight Identification with Cherenkov Radiation*, *Nucl.Instrum.Meth.* **A639** (2011) 173–176, [arXiv:1009.3793]. (Cited on page 169.)

[178] K. Fohl, *TORCH an Innovative High-Precision Time-of-Flight PID Detector for the LHCb Upgrade*, Tech. Rep. LHCb-PROC-2015-001. CERN-LHCb-PROC-2015-001, CERN, Geneva, Jan, 2015. (Cited on page 169.)

Ω

Appendix A

$B^0 \to K^{*0} e^+ e^-$ angular differential decay rate

A.1 Angular basis definition

The decay is completely described by four independent kinematic variables: the q^2 and the three angles θ_ℓ, θ_K and $\tilde{\phi}$ as defined in Figure A.1.

The angle θ_ℓ is defined as the angle between the direction of the e^+ (e^-) in the dimuon rest frame and the direction of the dimuon in the B^0 ($\overline{B}^0$) rest frame. The angle θ_K is defined as the angle between the direction of the kaon in the K^{*0} ($\overline{K}^{*0}$) rest frame and the direction of the K^{*0} ($\overline{K}^{*0}$) in the B^0 ($\overline{B}^0$) rest frame. The angle $\tilde{\phi}$ is the angle between the plane containing the e^+ and e^- and the plane containing the kaon and pion from the K^{*0}. Explicitly, $\cos\theta_\ell$ and $\cos\theta_K$ are defined as

$$\cos\theta_\ell = \left(\hat{p}_{e^+}^{(e^+e^-)}\right) \cdot \left(\hat{p}_{e^+e^-}^{(B^0)}\right) = \left(\hat{p}_{e^+}^{(e^+e^-)}\right) \cdot \left(-\hat{p}_{B^0}^{(e^+e^-)}\right) \ , \tag{A.1}$$

$$\cos\theta_K = \left(\hat{p}_{K^+}^{(K^{*0})}\right) \cdot \left(\hat{p}_{K^{*0}}^{(B^0)}\right) = \left(\hat{p}_{K^+}^{(K^{*0})}\right) \cdot \left(-\hat{p}_{B^0}^{(K^{*0})}\right) \tag{A.2}$$

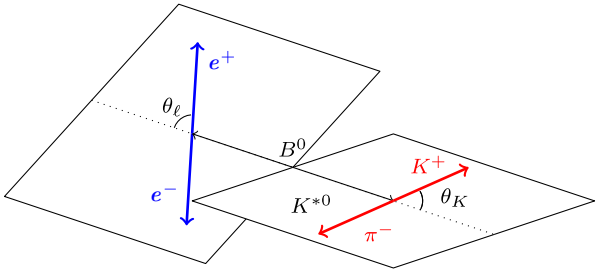
for the B^0 and

$$\cos\theta_\ell = \left(\hat{p}_{e^-}^{(e^+e^-)}\right) \cdot \left(\hat{p}_{e^+e^-}^{(\bar{B}^0)}\right) = \left(\hat{p}_{e^-}^{(e^+e^-)}\right) \cdot \left(-\hat{p}_{\bar{B}^0}^{(e^+e^-)}\right) \ , \tag{A.3}$$

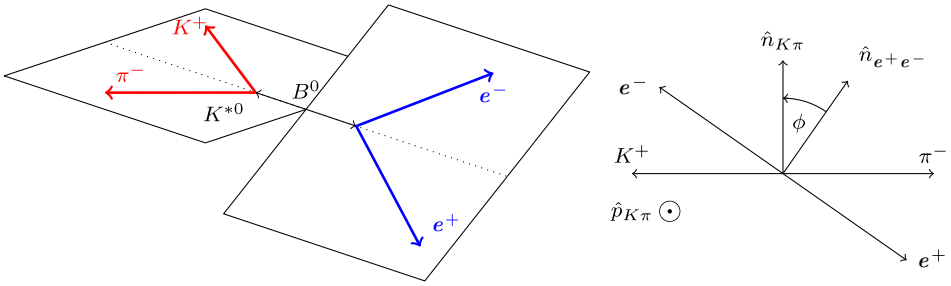
$$\cos\theta_K = \left(\hat{p}_{K^-}^{(K^{*0})}\right) \cdot \left(\hat{p}_{K^{*0}}^{(\bar{B}^0)}\right) = \left(\hat{p}_{K^-}^{(K^{*0})}\right) \cdot \left(-\hat{p}_{\bar{B}^0}^{(K^{*0})}\right) \tag{A.4}$$

for the $\overline{B}^0$ decay. The definition of the angle $\tilde{\phi}$ is given by

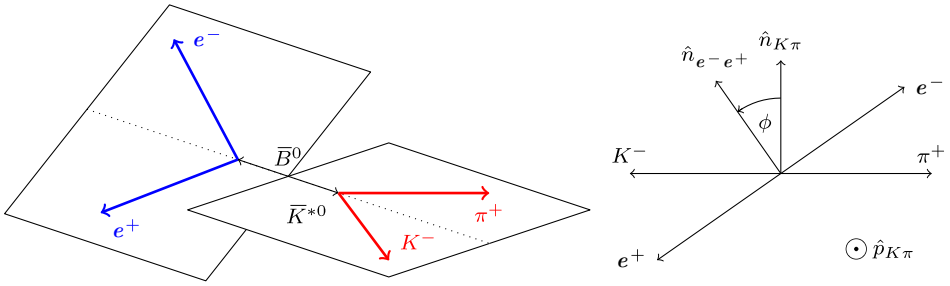
$$\cos\tilde{\phi} = \left(\hat{p}_{e^+}^{(B^0)} \times \hat{p}_{e^-}^{(B^0)}\right) \cdot \left(\hat{p}_{K^+}^{(B^0)} \times \hat{p}_{\pi^-}^{(B^0)}\right) \ , \tag{A.5}$$



(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

Fig. A.1: Graphical representation of the angular basis used for $B^0 \to K^{*0}e^+e^-$ and $\overline{B}^0 \to \overline{K}^{*0}e^+e^-$ decays. The notation $\hat{n}_{ab}$ is used to represent the normal to the plane containing particles a and b in the B^0 (or $\overline{B}^0$) rest frame. An explicit description of the angular basis is given in the text.

$$\sin\tilde{\phi} = \left[\left(\hat{p}_{e^+}^{(B^0)} \times \hat{p}_{e^-}^{(B^0)}\right) \times \left(\hat{p}_{K^+}^{(B^0)} \times \hat{p}_{\pi^-}^{(B^0)}\right)\right] \cdot \hat{p}_{K^{*0}}^{(B^0)} \tag{A.6}$$

for the B^0 and

$$\cos\tilde{\phi} = \left(\hat{p}_{e^-}^{(\overline{B}^0)} \times \hat{p}_{e^+}^{(\overline{B}^0)}\right) \cdot \left(\hat{p}_{K^-}^{(\overline{B}^0)} \times \hat{p}_{\pi^+}^{(\overline{B}^0)}\right) \ , \tag{A.7}$$

$$\sin\tilde{\phi} = -\left[\left(\hat{p}_{e^-}^{(\overline{B}^0)} \times \hat{p}_{e^+}^{(\overline{B}^0)}\right) \times \left(\hat{p}_{K^-}^{(\overline{B}^0)} \times \hat{p}_{\pi^+}^{(\overline{B}^0)}\right)\right] \cdot \hat{p}_{\overline{K}^{*0}}^{(\overline{B}^0)} \tag{A.8}$$

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

A.2 Angular observables

Tab. A.1: Angular observables I_j in terms of transversity amplitudes $A^{\mathrm{L,R}}_{\perp,\parallel,0}$ and their corresponding angular terms in the hypothesis of negligible lepton masses. Terms arising in the presence of an S-wave are neglected.

j	I_j	f_j
$1s$	$\frac{3}{4}\left[\lvert\mathcal{A}^{\mathrm{L}}_{\parallel}\rvert^2+\lvert\mathcal{A}^{\mathrm{L}}_{\perp}\rvert^2+\lvert\mathcal{A}^{\mathrm{R}}_{\parallel}\rvert^2+\lvert\mathcal{A}^{\mathrm{R}}_{\perp}\rvert^2\right]$	$\sin^2\theta_K$
$1c$	$\lvert\mathcal{A}^{\mathrm{L}}_{0}\rvert^2+\lvert\mathcal{A}^{\mathrm{R}}_{0}\rvert^2$	$\cos^2\theta_K$
$2s$	$\frac{1}{4}\left[\lvert\mathcal{A}^{\mathrm{L}}_{\parallel}\rvert^2+\lvert\mathcal{A}^{\mathrm{L}}_{\perp}\rvert^2+\lvert\mathcal{A}^{\mathrm{R}}_{\parallel}\rvert^2+\lvert\mathcal{A}^{\mathrm{R}}_{\perp}\rvert^2\right]$	$\sin^2\theta_K\cos 2\theta_\ell$
$2c$	$-\lvert\mathcal{A}^{\mathrm{L}}_{0}\rvert^2-\lvert\mathcal{A}^{\mathrm{R}}_{0}\rvert^2$	$\cos^2\theta_K\cos 2\theta_\ell$
3	$\frac{1}{2}\left[\lvert\mathcal{A}^{\mathrm{L}}_{\perp}\rvert^2-\lvert\mathcal{A}^{\mathrm{L}}_{\parallel}\rvert^2+\lvert\mathcal{A}^{\mathrm{R}}_{\perp}\rvert^2-\lvert\mathcal{A}^{\mathrm{R}}_{\parallel}\rvert^2\right]$	$\sin^2\theta_K\sin^2\theta_\ell\cos 2\phi$
4	$\sqrt{\frac{1}{2}}\mathrm{Re}(\mathcal{A}^{\mathrm{L}}_{0}\mathcal{A}^{\mathrm{L}*}_{\parallel}+\mathcal{A}^{\mathrm{R}}_{0}\mathcal{A}^{\mathrm{R}*}_{\parallel})$	$\sin 2\theta_K\sin 2\theta_\ell\cos\phi$
5	$\sqrt{2}\mathrm{Re}(\mathcal{A}^{\mathrm{L}}_{0}\mathcal{A}^{\mathrm{L}*}_{\perp}-\mathcal{A}^{\mathrm{R}}_{0}\mathcal{A}^{\mathrm{R}*}_{\perp})$	$\sin 2\theta_K\sin\theta_\ell\cos\phi$
$6\mathrm{s}$	$2\mathrm{Re}(\mathcal{A}^{\mathrm{L}}_{\parallel}\mathcal{A}^{\mathrm{L}*}_{\perp}-\mathcal{A}^{\mathrm{R}}_{\parallel}\mathcal{A}^{\mathrm{R}*}_{\perp})$	$\sin^2\theta_K\cos\theta_\ell$
7	$\sqrt{2}\mathrm{Im}(\mathcal{A}^{\mathrm{L}}_{0}\mathcal{A}^{\mathrm{L}*}_{\parallel}-\mathcal{A}^{\mathrm{R}}_{0}\mathcal{A}^{\mathrm{R}*}_{\parallel})$	$\sin 2\theta_K\sin\theta_\ell\sin\phi$
8	$\sqrt{\frac{1}{2}}\mathrm{Im}(\mathcal{A}^{\mathrm{L}}_{0}\mathcal{A}^{\mathrm{L}*}_{\perp}+\mathcal{A}^{\mathrm{R}}_{0}\mathcal{A}^{\mathrm{R}*}_{\perp})$	$\sin 2\theta_K\sin 2\theta_\ell\sin\phi$
9	$\mathrm{Im}(\mathcal{A}^{\mathrm{L}*}_{\parallel}\mathcal{A}^{\mathrm{L}}_{\perp}+\mathcal{A}^{\mathrm{R}*}_{\parallel}\mathcal{A}^{\mathrm{R}}_{\perp})$	$\sin^2\theta_K\sin^2\theta_\ell\sin 2\phi$

A.3 Full expressions of transversity amplitudes

The full expressions of the transversity amplitudes $A_{\perp \mathrm{L,R}}$, $A_{||\mathrm{L,R}}$ and $A_{0\mathrm{L,R}}$. Their dependence on Wilson coefficients $\mathcal{C}_{7,9,10}^{(\prime)}$ and form factors is made explicit.

$$A_{\perp L,R} = N\sqrt{2}\lambda^{1/2}\Bigg[[(C_9 + C_9') \mp (C_{10} + C_{10}')] \frac{V(q^2)}{m_B + m_{K^*}} + \frac{2m_b}{q^2}(C_7 + C_7')T_1(q^2)\Bigg], \tag{A.9}$$

$$A_{\| L,R} = -N\sqrt{2}(m_B^2 - m_{K^*}^2)\Bigg[[(C_9 - C_9') \mp (C_{10} - C_{10}')] \frac{A_1(q^2)}{m_B - m_{K^*}} + \frac{2m_b}{q^2}(C_7 - C_7')T_2(q^2)\Bigg], \tag{A.10}$$

$$A_{0L,R} = -\frac{N}{2m_{K^*}\sqrt{q^2}}\Bigg\{ [(C_9 - C_9') \mp (C_{10} - C_{10}')] \times \Bigg[(m_B^2 - m_{K^*}^2 - q^2)(m_B + m_{K^*})A_1(q^2) - \lambda\frac{A_2(q^2)}{m_B + m_{K^*}}\Bigg] + 2m_b(C_7 - C_7')\Bigg[(m_B^2 + 3m_{K^*}^2 - q^2)T_2(q^2) - \frac{\lambda}{m_B^2 - m_{K^*}^2}T_3(q^2)\Bigg]\Bigg\}, \tag{A.11}$$

where

$$N = V_{tb}V_{ts}^*\left[\frac{G_F^2\alpha^2}{3\cdot 2^{10}\pi^5 m_B^3}q^2\lambda^{1/2}\beta_\mu\right]^{1/2}, \tag{A.12}$$

with $\lambda = m_B^4 + m_{K^*}^4 + q^4 - 2(m_B^2 m_{K^*}^2 + m_{K^*}^2 q^2 + m_B^2 q^2)$ and $\beta_\mu = \sqrt{1 - 4m_\mu^2/q^2}$.

Appendix B

$B^0 \to K^{*0} e^+ e^-$ selection and angular analysis

B.1 Supplementary information on the BDT training

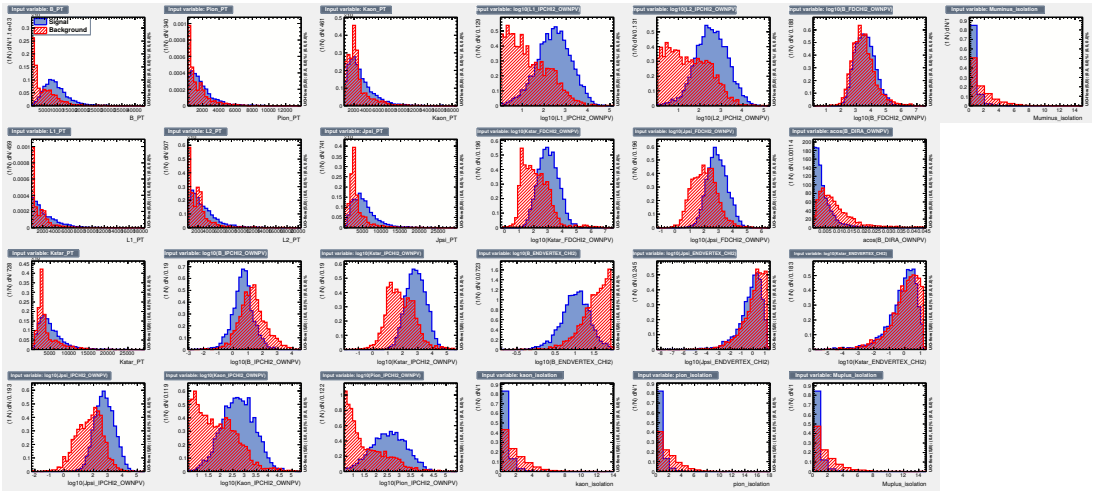


Fig. B.1: Distributions of discriminating variables of the signal (blue) and background (red) samples used to train *BDT 1*

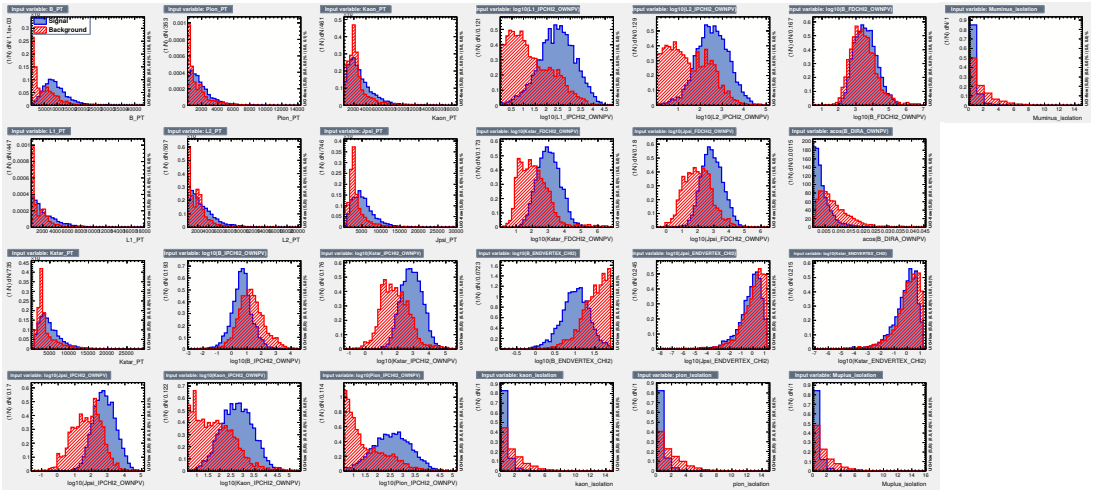


Fig. B.2: Distributions of discriminating variables of the signal (blue) and background (red) samples used to train *BDT 2*

B.2 Correlation between parameters of the polynomials fitting the angular acceptances

	(c_2^ℓ, c_4^ℓ)	(c_1^ℓ, c_2^ℓ)	(c_1^ℓ, c_4^ℓ)
L0Ele	0.11	0.01	-0.03
L0Had	-0.29	0.02	-0.06
L0TIS	-0.49	-0.32	0.06

Tab. B.1: Correlation between coefficients of the fit to the acceptance in $\cos\theta_\ell$. The curve fit is shown in Figure 4.3 and resulting parameters are in Table 4.3

	(c_1^K, c_2^K)	(c_1^K, c_3^K)	(c_1^K, c_4^K)	(c_2^K, c_3^K)	(c_2^K, c_4^K)	(c_3^K, c_4^K)
L0Ele	-0.12	-0.15	0.00	-0.09	-0.14	-0.11
L0Had	0.01	-0.07	-0.05	-0.02	-0.05	-0.05
L0TIS	-0.31	-0.34	0.01	-0.27	-0.41	-0.36

Tab. B.2: Correlation between coefficients of the fit to the acceptance in $\cos\theta_K$. The curve fit is shown in Figure 4.4 and resulting parameters are in Table 4.4

B.3 Correlation between $\cos\theta_\ell$ and $\cos\theta_K$ in the acceptance modelling

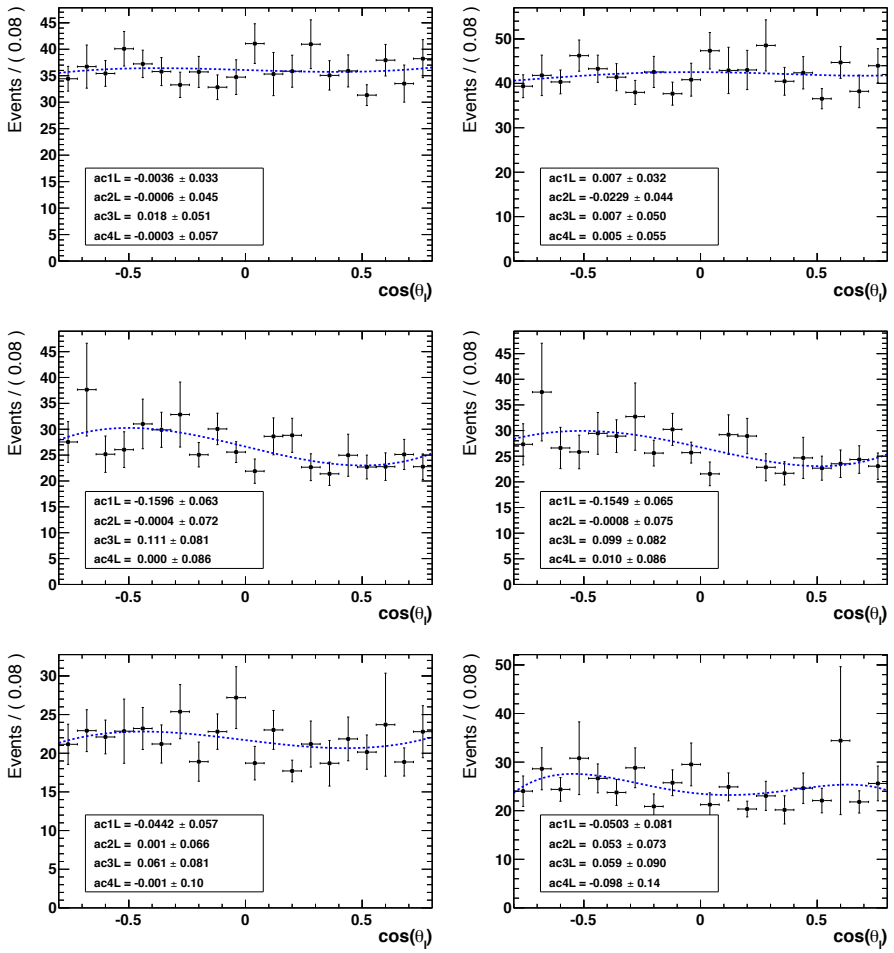


Fig. B.3: Phase space MC for L0Ele (top row), L0Had (middle row) and L0TIS (bottom row) trigger categories. The distribution on the right is weighted with $1/\varepsilon(\cos\theta_\ell)$, the one on the left is weighted with $1/(\varepsilon(\cos\theta_\ell)\varepsilon(\cos\theta_K))$. Both are fitted to a Legendre polynomial of order 4.

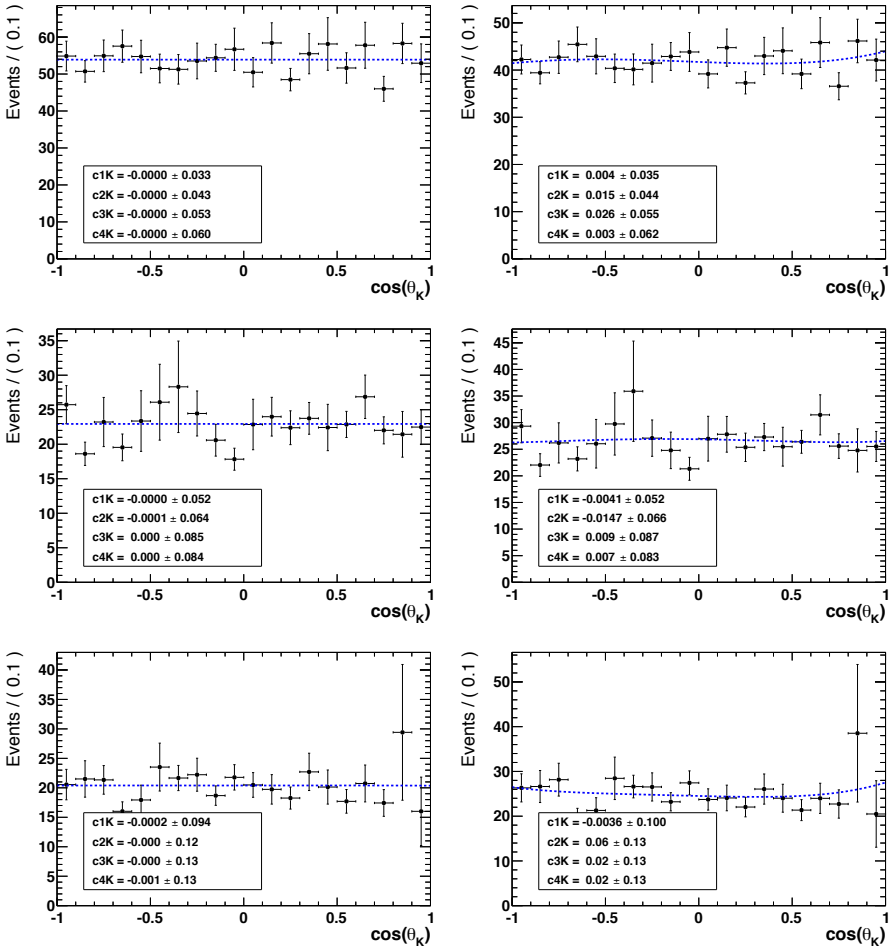


Fig. B.4: Phase space MC for L0Ele (top row), L0Had (middle row) and L0TIS (bottom row) trigger categories. The distribution on the right is weighted with $1/\varepsilon(\cos\theta_K)$, the one on the left is weighted with $1/(\varepsilon(\cos\theta_\ell)\varepsilon(\cos\theta_K))$. Both are fitted to a Legendre polynomial of order 4.

B.4 Correlation of $\tilde{\phi}$ with $\cos\theta_\ell$ and $\cos\theta_K$ in the acceptance modelling

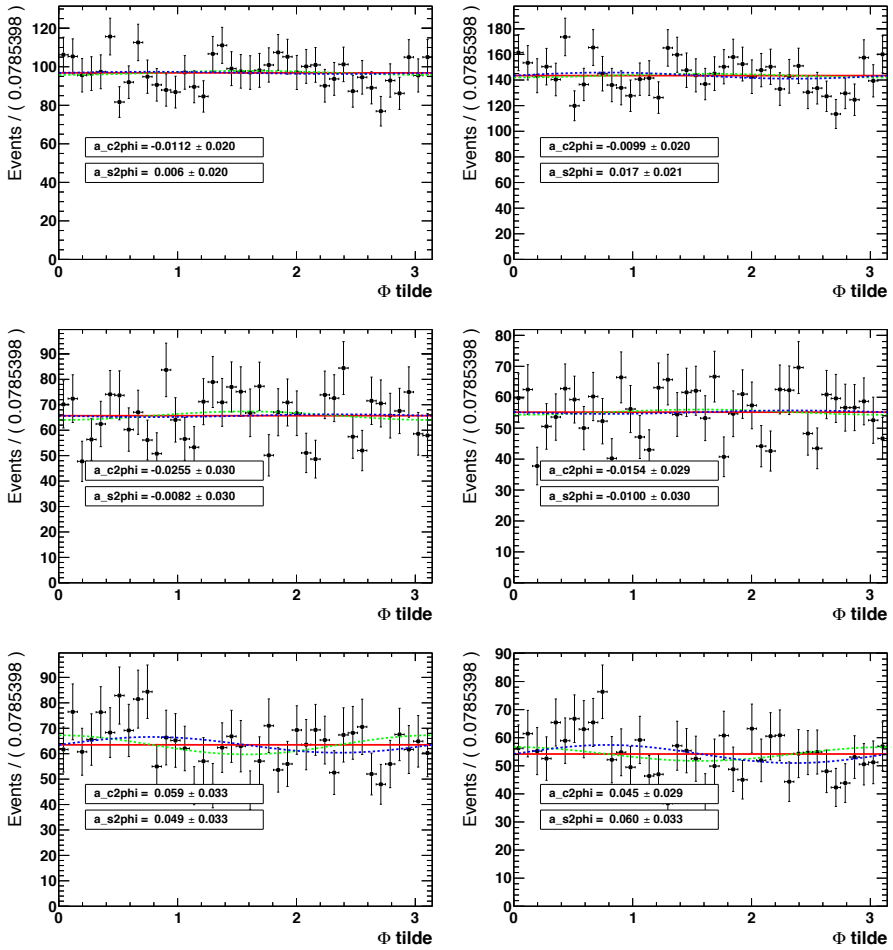


Fig. B.5: Standard Model MC for the L0Ele (top row), L0Had (middle row) and L0TIS (bottom row) trigger categories. The distribution on the right is weighted with $1/\varepsilon(\cos\theta_\ell)$, the one on the left is weighted with $1/\varepsilon(\cos\theta_K)$. Both are are fitted to $1 + a_c\cos(2\tilde{\phi})$ and $1 + a_s\sin(2\tilde{\phi})$.

B.5 Angular fit for the modelling of the combinatorial background

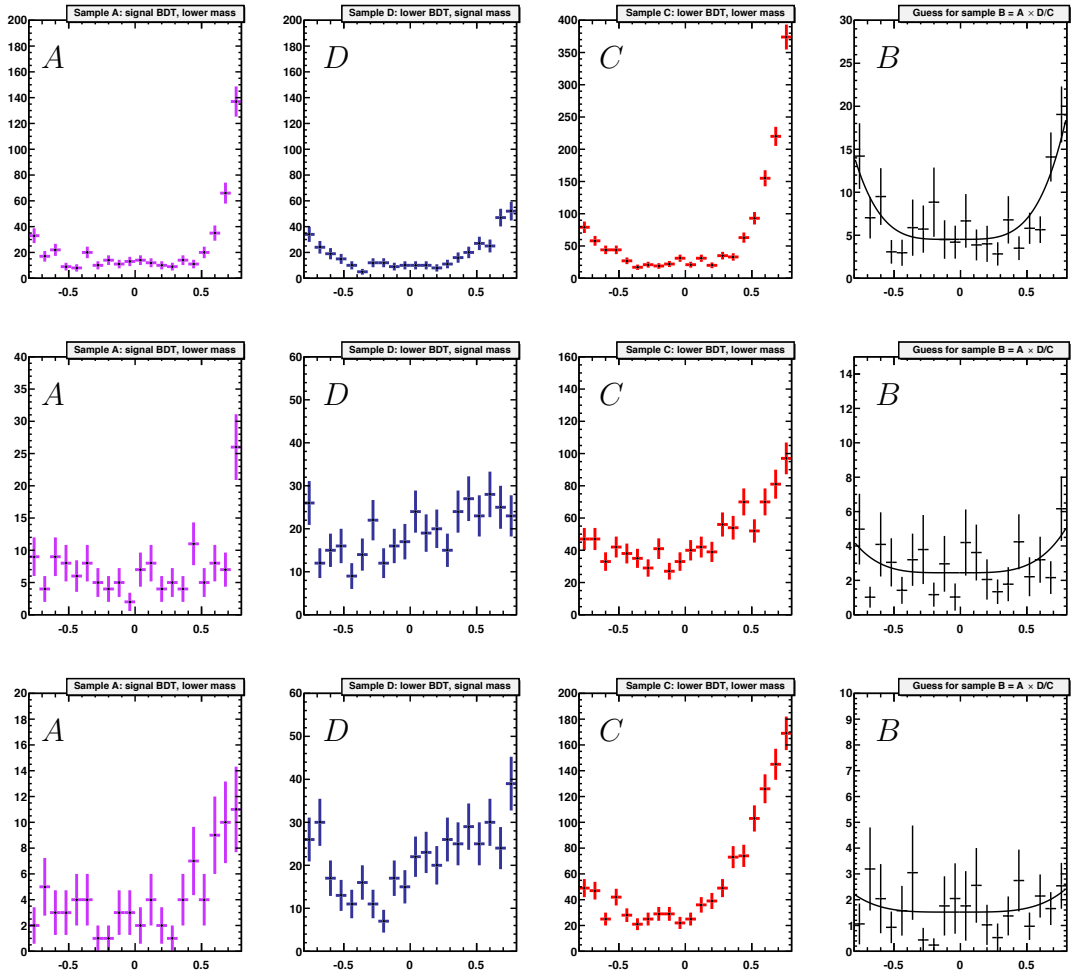


Fig. B.6: Distributions of $\cos\theta_\ell$ used to model the angular shape of the combinatorial background for the three trigger categories L0Ele, L0Had and L0TIS in the rows from top to bottom. From left to right are the three samples A, D and C defined in Section 4.3.1 and on the right is the guess for the distribution of the sample B obtained as $A\ D/C$.

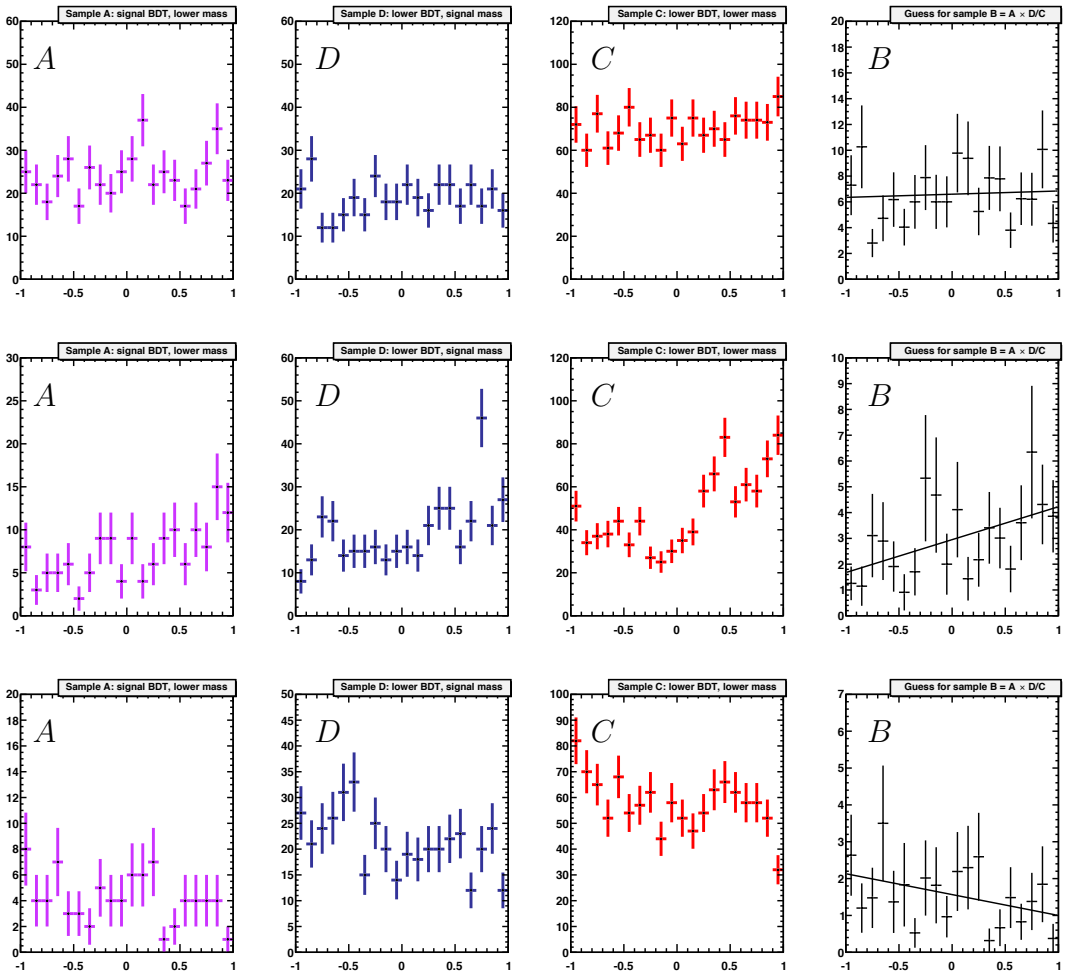


Fig. B.7: Distributions of $\cos\theta_K$ used to model the angular shape of the combinatorial background for the three trigger categories L0Ele, L0Had and L0TIS in the rows from top to bottom. From left to right are the three samples A, D and C defined in Section 4.3.1 and on the right is the guess for the distribution of the sample B obtained as $A\ D/C$.

B.6 Validation of the method for the extraction of the angular shape of the combinatorial background

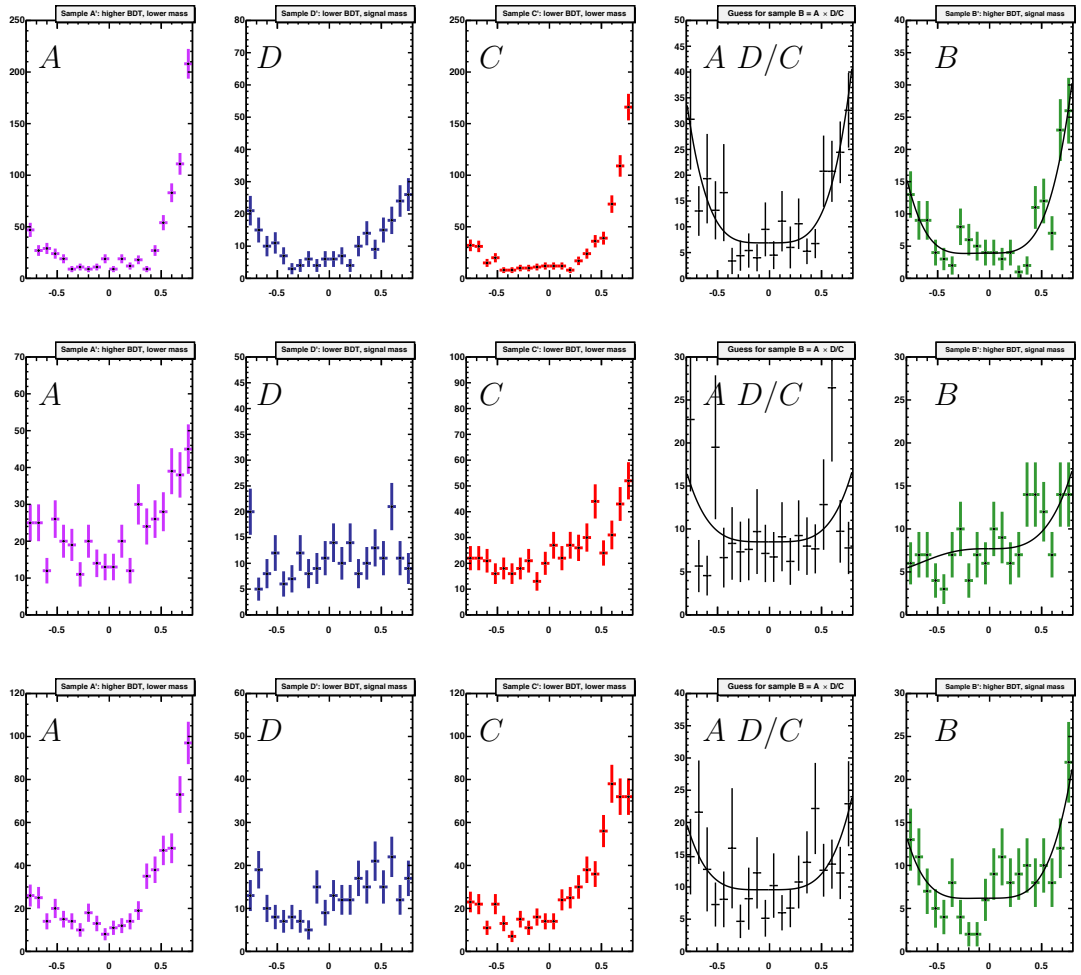


Fig. B.8: Distributions of the $\cos\theta_\ell$ angle used to validate the method for the extraction of the angular distribution of the combinatorial background. From left to right the first three plots are the 3 samples A, D and C defined in Section 4.3.1 for the L0Ele (top row), L0Had (middle row) and L0TIS (bottom row) trigger categories. Then, the fourth is the guess for the distribution of the sample B obtained as $A\ D/C$ and the next is the real distribution of data in B.

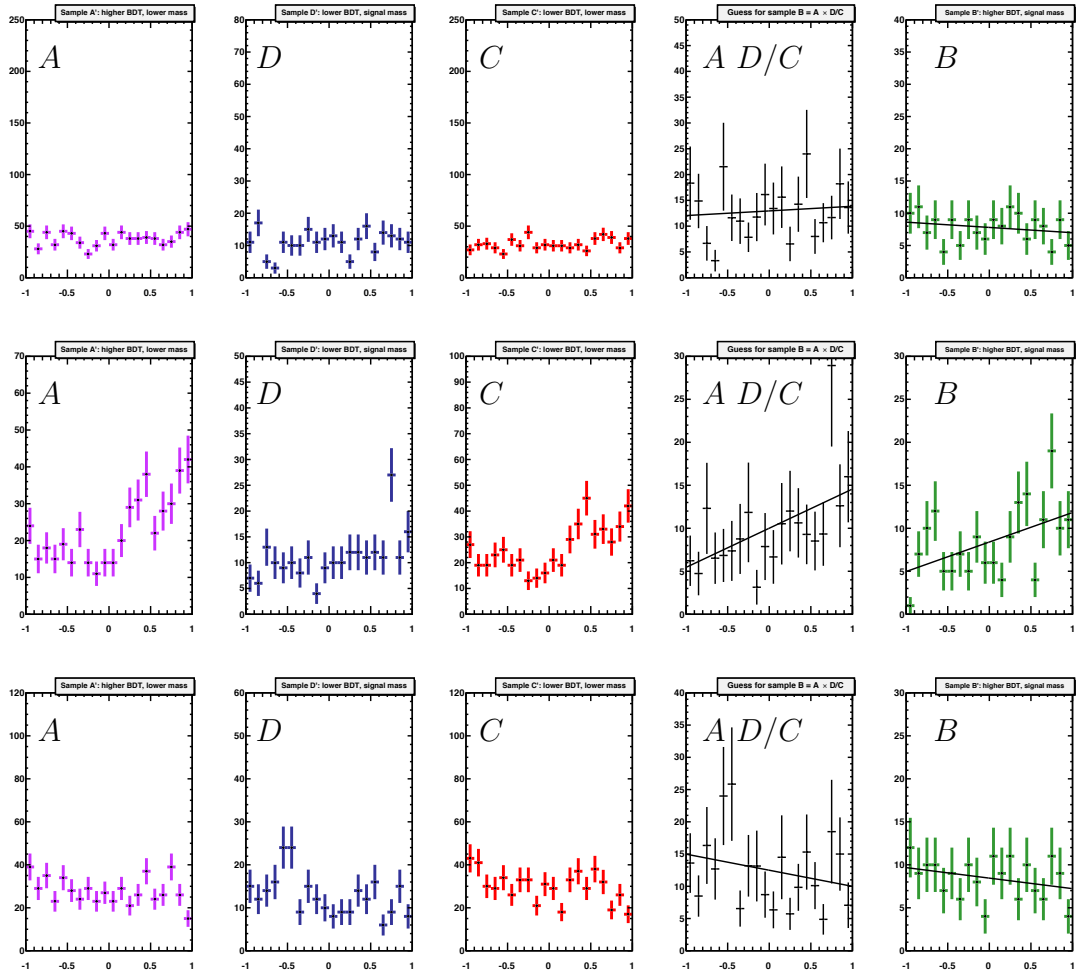


Fig. B.9: Distributions of the $\cos\theta_K$ angle used to validate the method for the extraction of the angular distribution of the combinatorial background. From left to right the first three plots are the 3 samples A, D and C defined in Section 4.3.1 for the L0Ele (top row), L0Had (middle row) and L0TIS (bottom row) trigger categories. Then, the fourth is the guess for the distribution of the sample B obtained as $A\ D/C$ and the next is the real distribution of data in B.

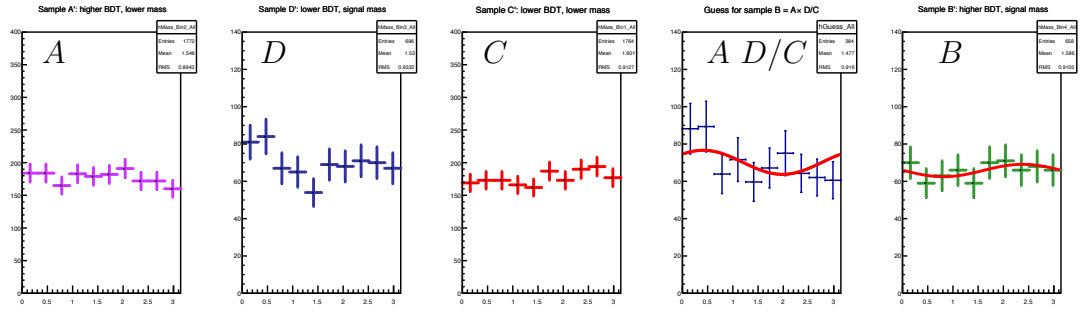


Fig. B.10: Distributions of the $\tilde{\phi}$ angle used to validate the method for the extraction of the angular distribution of the combinatorial background. From left to right the first three plots are the 3 samples A, D and C defined in Section 4.3.1 for all trigger categories summed up. Then, the fourth is the guess for the distribution of the sample B obtained as $A\ D/C$ and the next is the real distribution of data in B.

Aknowledgements

First of all I want to thank the director of the Laboratoire de l'Accélérateur Linéaire, Achille Stocchi, for letting me do my PhD thesis at LAL as well as for giving me the opportunity to follow the NPAC Master. And last but not least for his inspiring flavour physics classes.

I owe my gratitude to Marcella Bona and Ulrik Egede, who kindly accepted to be referees of my thesis. I really appreciated your careful reading and useful comments. Thanks a lot to Damir Becirevic, Neville Harnew, José Ocariz and, again, Achille Stocchi for being part of my jury.

Mon remerciement plus grand va à Marie-Hélène Schune, ma directrice de thèse, qui m'a adopté et m'a appri jour par jour comment devenir un scientifique. Merci pour ton inestimable disponibilité, pour tes conseils de vie, ton humanité et sincerité.

Un enorme merci à Nicolas Arnaud, qui a cru en moi dès le premier jour et qui a fait décoller ma thèse. Merci pour ton aide lors de mon installation en France et de ta grande confiance en moi en tout moment.

Un remerciement spécial à Jacques Lefrançois, mon idéal de physicien. J'ai été honoré de travailler à coté de toi. Sans toi cette thèse ne serait pas été possible, merci pour tes idées, tes conseils et ta ferveur scientifique contagieuse.

Pendant ma thèse j'ai eu la chance de travailler avec le groupe LHCb du LAL qui a été une grande famille pour moi. Je garderai un magnifique souvenir des discussions scientifique aux meetings, mais aussi de celles semi-scientifique au CESFO, à la cafét et lors de nos pique-nique sur la pelouse devant le 208. Merci à Frédéric Machefert qui a été le meilleur collegue de bureau possible malgré sa concurrence déloyale avec ses plantes d'avocat. Merci à Patrick Robbe pour son aide informatique et son bon humeur. Merci à Benoit Viaud pour ses blagues et à Yasmine Amhis pour ses magnifiques sourires. Et puis merci à Guy Wormser et Sergey Barsuk pour les conversations toujours interessants and thanks to Yiming, Victor, Laure and Giulia as well, it's a pity you came and I leave.

Thaks to all LAL PhD's of present and past, a Marta in primis (perché sei la presidente, non per altro, gne gne!), a Renato-l'instancabile per le discussioni sulla

fisica teorica e per quelle sull'amore platonico. Et puis merci aux beaux gosses du 208, Alexis et Camille, j'espère que notre amitié coninuera après le jour J. Thanks to Claire Prouve for her witty humour and her fluffy animals. Merci aux malheureux doctorants du batiment 200 et 209 pour les pauses café, les séminaire étudiant et les bières au Gramophone. Merci aux vieux gens de NPAC aussi: Cyril, Sam, Aurélie, Alexis et tous les autres.

Non posso dimenticare gli amici di Treviso, parte della mia anima. In ordine di apparizione: La Michela, imbattibile a Machiavelli. Maxi e il libro della natura, che un po' precursore di questo. La positività della Chiaretta. E poi i "Fioi", che la diaspora non ci fa nessuna paura. E in particolare: Enrico, per primo solo perch oggi il suo compleanno, Niccol, Caio e Checco per la loro prova d'amore, Nicola e l'Ogger i Grossi assenti. LaSara, si vabb troppo comodo! E poi in ordine sparso i Raccoons Marco e Tommasino, LaBea, LaVale, la SaraMamma, la GeGe, Albe, Racioppi, la Mària, La Gianna, VaninA, la ValeChan, Poppi, la Sbolzi, il Mammo, AndreaDeMeo, la Elsi, Vitto, Stefanone, ...

Poi ci sono gli amici fisici di Padova: Giovanni e Dario (i miei insegnanti), Giulia Cusin e Tommaso Brotto i grandi teorici. Ci sono i ritals del martedì: Michele, Andrea, Camilla, Cate, Silvia, Lapo, Anna, Giorgio, ...E poi in ordine sparso tante belle persone conosciute in questi ultimi tempi: Jessica, il dottor Gonnella, Luca Mastro, Ivo, Xavier, Lorenzo Caccianiga, Fabio Ferrari, Lucia Perrin, Paolo Agnes, Davide Poletti. Chiedo ufficialmente scusa a chi ho dimenticato, ditemelo, vi aggiungo a penna all'inizio della lista!

Un remerciement special à ma famille de Gentilly: Loeïza, Matthieu, Mattia, Simon mais aussi Vincent, Ana Lucia, Léa et Pierre...qu'est-ce que c'est beaux notre loft!

Grazie a Stefania per essere una donna d'amore e avermi insegnato cosa significa. Questa tesi è in gran parte tua.

Per finire l'ennesimo grazie alla mia famiglia, Papà, Mamma e Sorella, per l'eterna fiducia, il sostegno, la motivazione e l'amore incommensurabile.