Analyse angulaire de désintégrations B $\to$K* e^+e avec le détecteur LHCb et upgrade de l'électronique des calorimètres

Fabrice Desse

université
PARIS-SACLAY

Angular analysis of $B^0 \to K^{*0} e^+ e^-$ decays with the LHCb detector and upgrade of the electronics of the calorimeters

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 576, Particules, Hadrons, Énergie, Noyau: Instrumentation, Imagerie, Cosmos et Simulation (PHENIICS)
Spécialité de doctorat: Physique des particules
Unité de recherche: Université Paris-Saclay, CNRS, IJCLab, 91405, Orsay, France.
Référent: Faculté des sciences d'Orsay

Thèse présentée et soutenue à Orsay, le 22 juillet 2020, par

Fabrice DESSE

Composition du jury:

Achille Stocchi Professeur, Université Paris-Saclay	Président
Tim Gershon Professeur, University of Warwick	Rapporteur & examinateur
Cristina Lazzeroni Professeure, University of Birmingham	Rapporteur & examinatrice
Aoife Bharucha Chargée de recherche, Université d'Aix-Marseille	Examinatrice
Florian Beaudette Directeur de recherche, École Polytechnique	Examinateur
Marie-Hélène Schune Directrice de recherche, Université Paris-Saclay	Directrice de thèse

Thèse de doctorat

NNT: 2020UPASS126

Contents

0 Résumé **1**

1 Theoretical and experimental context **13**
1.1 The Standard Model of particle physics 14
1.2 Probing New Physics in rare $b \to s\gamma$ transitions 20
1.3 Current status of the measurement of the photon polarization in $b \to s\gamma$ transitions 22
1.3.1 Inclusive branching fractions 22
1.3.2 Time dependent rate of radiative decays 22
1.3.3 Up-down asymmetry 24
1.3.4 Measuring the photon polarization in the baryonic sector . . 25
1.4 Measuring the photon polarization in the $B^0 \to K^{*0}e^+e^-$ decay mode 27
1.4.1 Differential decay width of the $B^0 \to K^{*0}e^+e^-$ decay 28
1.4.2 Choice of the q^2 range 33

2 The LHCb experiment at the LHC **39**
2.1 The Large Hadron Collider at CERN 39
2.1.1 Data samples 41
2.2 The LHCb detector during Run 1 and Run 2 41
2.2.1 General layout 43
2.2.2 Tracking and vertex reconstruction 45
2.2.3 Particle identification and energy measurement 48
2.2.4 Trigger system 57
2.2.5 The LHCb simulation 59
2.3 Upgrade I of the LHCb detector 59
2.3.1 The Front-End board of the upgraded LHCb calorimeters . . 63
2.3.2 The upgraded Low Level Trigger 66

Preamble **73**

3 Selection of $B^0 \to K^{*0} e^+ e^-$ events **75**
3.1 Trigger and pre-selection . . . 75
3.2 Corrections to simulation . . . 77
3.2.1 Particle identification . . . 78
3.2.2 L0 trigger . . . 78
3.2.3 B kinematics, multiplicity and reconstruction . . . 81
3.3 Specific backgrounds . . . 86
3.3.1 Semileptonic background . . . 86
3.3.2 $B^0 \to K^{*0}\gamma(\to e^+e^-)$ background . . . 88
3.3.3 $B^0 \to K^{*0}\eta$ and $B^0 \to K^{*0}\pi^0$ backgrounds . . . 91
3.3.4 $B^0 \to K^{*0}\eta(\to \gamma\gamma)$ and $B^0 \to K^{*0}\pi^0(\to \gamma\gamma)$ backgrounds . . . 94
3.3.5 $B^0_s \to \phi e^+ e^-$ and $B^+ \to K^+ e^+ e^-$ backgrounds . . . 95
3.3.6 $\Lambda^0_b \to pK^- e^+ e^-$ background . . . 97
3.3.7 $\pi \to e$ mis-identification . . . 98
3.3.8 K^{*0} and $\overline{K}^{*0}$ mis-identification and multiple candidates . . . 102
3.4 Combinatorial and partially reconstructed backgrounds . . . 103
3.4.1 Partially reconstructed background . . . 103
3.4.2 Combinatorial background . . . 105
3.4.3 BDT optimization . . . 106
3.5 Selection summary and efficiency . . . 116
3.6 Mass fits . . . 118
3.6.1 Mass shapes determination via fits to MC samples . . . 118
3.6.2 Partially reconstructed background . . . 121
3.6.3 $B^0 \to K^{*0}\pi^0/\eta$ background . . . 121
3.6.4 MC/data differences from $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ . . . 121
3.6.5 Invariant mass fit to $B^0 \to K^{*0} e^+ e^-$ data . . . 128

4 Angular analysis of the $B^0 \to K^{*0} e^+ e^-$ decay at very low q^2 **133**
4.1 Angular fit strategy . . . 133
4.2 Angular acceptance . . . 135
4.2.1 Phase-space Monte Carlo reweighting . . . 136
4.2.2 Acceptance fit . . . 136
4.3 Angular modeling of backgrounds . . . 140
4.3.1 Angular modeling of $B^0 \to K^{*0}\gamma(\to e^+e^-)$ background . . . 141
4.3.2 Angular modeling of combinatorial background . . . 141
4.3.3 Angular modeling of partially reconstructed background . . . 147
4.3.4 Angular modeling of $B^0 \to K^{*0}\eta$ and $B^0 \to K^{*0}\pi^0$ backgrounds . . . 149
4.4 Angular and mass fit validation . . . 151

4.4.1 Toys . . . 151
4.4.2 Angular fit to $B^0 \to K^{*0} e^+ e^-$ MC . . . 151
4.4.3 Angular fit to $B^0 \to K^{*0}\gamma(\to e^+e^-)$ MC and data . . . 154
4.5 Systematic uncertainties . . . 156
4.5.1 Angular acceptance . . . 158
4.5.2 Combinatorial background modeling . . . 158
4.5.3 $B^0 \to K^{*0}\eta/\pi^0(\to e^+e^-\gamma)$ backgrounds modeling . . . 158
4.5.4 Partially reconstructed background modeling . . . 159
4.5.5 Resolution of the $\tilde{\phi}$ angle . . . 159
4.5.6 K^{*0} and $\overline{K}^{*0}$ mis-identification . . . 160
4.5.7 Data/MC differences in mass shape . . . 161
4.6 Results . . . 161
4.6.1 Comparison to previous results . . . 162
4.6.2 Effective q^2 bin . . . 164
4.6.3 Constraints on the photon polarization . . . 166
4.7 Conclusion and perspectives for the future . . . 170

Appendices **172**

A Angular differential decay rate of the $B^0 \to K^{*0} e^+ e^-$ decay **173**
A.1 Definition of the angular basis . . . 173
A.2 Definition of the angular coefficients I_i . . . 176
A.3 Definition of the transversity amplitudes . . . 177

B Combinatorial background **178**
B.1 Variables used for the training of the BDT . . . 178
B.2 Training of the BDT . . . 182

C Corrections to simulation **185**
C.1 $B^0 \to K^{*0} e^+ e^-$ MC angular distributions . . . 185
C.2 Angular acceptance . . . 188

D Bremsstrahlung recovery **191**

E Angular and mass fit validation **193**

Bibliography **195**

Chapter 0

Résumé

Le Modèle Standard (MS) de la physique des particules décrit trois des quatre forces fondamentales: la force électromagnétique, l'interaction forte et l'interaction faible. Malgré son succès inégalé pour décrire les particules subatomiques et leurs interactions, le MS présente un certain nombre de limites. Premièrement, la gravité est absente de la théorie. De plus, le MS laisse un certain nombre de faits expérimentaux inexpliqués, comme par exemple les différences de masse importantes entre les générations de fermions ou encore la masse non nulle des neutrinos. Enfin, le MS est incapable d'expliquer certaines observations cosmologiques. En particulier, le MS n'offre aucun candidat à la matière noire et ne peut expliquer l'asymétrie matière/anti-matière observée dans l'Univers. Ces limites sont les principales raisons qui motivent la recherche d'une Nouvelle Physique (NP) au-delà du MS.

Les transitions d'un quark b vers un quark s (appelées transitions $b \to s$) sont un laboratoire privilégié de recherche de NP. En effet, les transitions $b \to s$ sont des courants neutres changeant la saveur qui n'apparaissent qu'au niveau des boucles dans le MS. Ces processus sont donc particulièrement sensibles à des effets de NP intervenant dans les boucles quantiques qui pourraient notamment se manifester via la modification de certaines observables angulaires. Les transitions $b \to s$ peuvent être décrites à l'aide de l'Expansion des Produits de l'Opérateur (OPE) [1] permettant de séparer les processus de haute et basse énergie, où l'Hamiltonien effectif prend la forme

$$\mathcal{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_j (\mathcal{C}_j \mathcal{O}_j + \mathcal{C}'_j \mathcal{O}'_j) \qquad , \tag{0.1}$$

où G_F est la constante de Fermi et V_{ij} les éléments de la matrice CKM. Les opérateurs locaux $\mathcal{O}_j$ décrivent les processus de basse énergie tandis que les co-

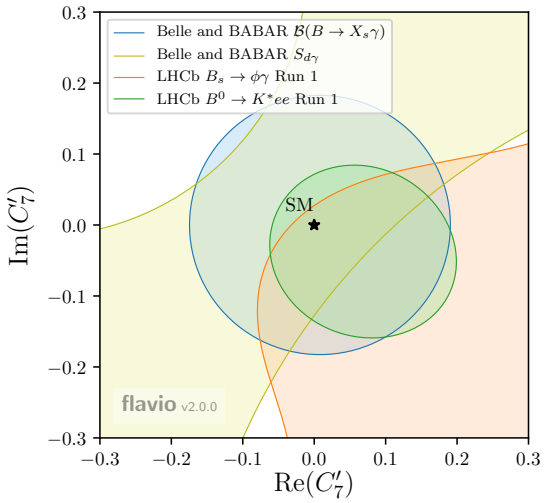


Figure 1: Contraintes (à 2σ) sur $\mathcal{C}_7'$ de mesures publiées. Les contraintes venant de la mesure de rapports d'embranchements inclusifs et de $S_{d\gamma}$ par les expériences Belle et BaBar sont représentées en bleu et jaune respectivement. Les contraintes des mesures de l'expérience LHCb utilisant les données du Run 1 ($3\,\mathrm{fb}^{-1}$) sont représentées en orange pour la désintégration $B_s^0 \to \phi\gamma$ et en vert pour l'analyse angulaire de la désintégration $B^0 \to K^{*0}e^+e^-$. La prédiction du MS est représentée par l'étoile noire.

efficients de Wilson $\mathcal{C}_j$ décrivent les processus de haute énergie. En particulier, les coefficients de Wilson $\mathcal{C}_7^{(\prime)}$ décrivent la polarisation du photon dans les transitions de type $b \to s\gamma$. Dans le MS, le photon est principalement émis avec une polarisation gauche, ce qui implique $\mathcal{C}_7'/\mathcal{C}_7 \simeq m_s/m_b \simeq 0.02$ [2] où m_i est la masse du quark i. Cependant plusieurs modèles de physique au-delà du MS (par exemple [3–5] ou [6]) autorisent une importante contribution des courants droits. La mesure de la polarisation du photon dans les processus de type $b \to s\gamma$ est donc une excellente manière d'explorer la NP.

Il existe plusieurs mesures de la polarisation du photon dans des transitions de type $b \to s\gamma$ [7–10] qui permettent de contraindre la valeur du coefficient de Wilson $\mathcal{C}_7'$ comme illustré par la figure 1. Les mesures de rapports d'embranchements inclusifs de type $B \to X_s\gamma$ où X_s est un état quelconque contenant un quark s imposent une contrainte circulaire proportionnelle à $|\mathcal{C}_7|^2 + |\mathcal{C}_7'|^2$. Dans les désintégrations $B^0 \to K_S^0\pi^0\gamma$ et $B_s^0 \to \phi\gamma$ l'asymétrie CP due à l'interférence entre l'oscillation et la désintégration du méson B est donnée par

$$S_{q\gamma} \approx \xi \frac{2\mathcal{I}m[e^{-i\phi_q}\mathcal{C}_7\mathcal{C}_7{}']}{|\mathcal{C}_7|^2 + |\mathcal{C}_7'|^2} \quad , \tag{0.2}$$

où ϕ_q est la phase de l'oscillation $B_q^0 - \bar{B}_q^0$ et $\xi = +1(-1)$ pour un méson B_q^0 ($\bar{B}_q^0$). Dans le cas du méson B_s^0 il est également possible de mesurer l'asymétrie CP due à la différence de largeur

$$\mathcal{A}_{s\gamma}^{\Delta} \approx \xi \frac{2\mathcal{R}e[e^{-i\phi_s}\mathcal{C}_7\mathcal{C}_7{}']}{|\mathcal{C}_7|^2 + |\mathcal{C}_7'|^2} \quad . \tag{0.3}$$

Les désintégrations $B^0 \to K^{*0}(\to K^+\pi^-)e^+e^-$ sont sensible quasi exclusivement au coefficients de Wilson $\mathcal{C}_7'$ à très bas q^2, où q est la masse de la paire d'électrons. La désintégration peut être décrite en fonction de q^2 et trois angles [2] comme

$$\begin{aligned}\left\langle \frac{d^4\Gamma}{dq^2 d\cos\theta_\ell d\cos\theta_K d\tilde{\phi}} \right\rangle_{CP} = \frac{9}{16\pi} \Big\{ & \frac{3}{4}(1 - F_{\rm L})\sin^2\theta_K + F_{\rm L}\cos^2\theta_K \\ & + \left[\frac{1}{4}(1 - F_{\rm L})\sin^2\theta_K - F_{\rm L}\cos^2\theta_K\right]\cos 2\theta_l \\ & + \frac{1}{2}(1 - F_{\rm L})A_T^{(2)}\sin^2\theta_K \sin^2\theta_l \cos 2\tilde{\phi} \\ & + (1 - F_{\rm L})A_T^{Re}\sin^2\theta_K \cos\theta_l \\ & + \frac{1}{2}(1 - F_{\rm L})A_T^{Im}\sin^2\theta_K \sin^2\theta_l \sin 2\tilde{\phi} \Big\} \quad ,\end{aligned} \tag{0.4}$$

où θ_l est l'angle entre la direction du e^+ (e^-) et la direction opposée à celle du B^0 ($\overline{B}^0$) dans le référentiel du diélectron, θ_K est l'angle entre la direction du kaon et la direction opposée à celle du B^0 ($\overline{B}^0$) dans le référentiel du K^{*0} ($\overline{K}^{*0}$), ϕ est l'angle entre le plan contenant les deux électrons et le plan contenant les deux hadrons de l'état final dans le référentiel du B^0 ($\overline{B}^0$) et $\tilde{\phi} = \phi + \pi$ si $\phi < 0$. Le paramètre F_L est la fraction de polarisation longitudinale du méson K^{*0} tandis que A_T^{Re} est lié à l'asymétrie *forward-backward*. Les deux paramètres sensibles à la polarisation du photon sont $A_T^{(2)}$ et A_T^{Im} qui sont directement proportionnels aux coefficients de Wilson $\mathcal{C}_7^{(\prime)}$ dans la limite $q^2 \to 0$

$$A_T^{(2)}(q^2 \to 0) = \frac{2\mathcal{R}e\left(\mathcal{C}_7\mathcal{C}_7'^*\right)}{|\mathcal{C}_7|^2 + |\mathcal{C}_7'|^2} \quad , \tag{0.5}$$

$$A_T^{Im}(q^2 \to 0) = \frac{2\mathcal{I}m\left(\mathcal{C}_7\mathcal{C}_7'^*\right)}{|\mathcal{C}_7|^2 + |\mathcal{C}_7'|^2} \quad . \tag{0.6}$$

L'analyse angulaire de la désintégration $B^0 \to K^{*0}e^+e^-$ présentée dans cette thèse est effectuée dans une unique région en q^2 $[0.0001, 0.25]\,\mathrm{GeV}^2/c^4$. Le choix de la limite inférieure est principalement motivé par la résolution de l'angle ϕ.

En effet, lorsque q^2 diminue, l'angle entre les deux électrons rétrécit, rendant la reconstruction du plan du diélectron plus difficile et diminuant ainsi la précision sur l'angle ϕ. Le choix de la limite supérieure est motivé par trois aspects. Premièrement, lorsque q^2 n'est pas strictement égal à zéro, $A_T^{(2)}$ et A_T^{Im} reçoivent également des contributions des coefficients de Wilson $\mathcal{C}_{9,10}$ affectant ainsi la proportionnalité directe avec les coefficients de Wilson $\mathcal{C}_7^{(\prime)}$. Par ailleurs, les termes proportionnels à $A_T^{(2)}$ et A_T^{Im} dans l'équation (0.4) sont multipliés par un facteur $(1 - F_{\mathrm{L}})$. Or, F_{L} est en première approximation proportionnel à q^2. Enfin la limite supérieure à $0.25\,\mathrm{GeV}^2/c^4$ permet d'éviter une contamination de désintégrations $B^0 \to K^{*0}\rho^0(\to e^+e^-)$.

La mesure présentée dans cette thèse utilise les $9\,\mathrm{fb}^{-1}$ de collisions proton-proton collectées par l'expérience LHCb au Large Hadron Collider (LHC) au CERN entre 2011 et 2018. Le détecteur LHCb [11] permet de sélectionner les désintégrations de mésons B grâce à une excellente résolution du vertex primaire ($\sim 150\,\mu\mathrm{m}$ selon l'axe du faisceau et $\sim 20\,\mu\mathrm{m}$ dans le plan transverse) et du paramètre d'impact ($\sim 50\,\mu\mathrm{m}$). Par ailleurs, le système de tracking permet de mesurer le moment des traces chargées avec une précision de $\delta p/p < 1\%$. Enfin, la combinaison d'un calorimètre électronique, d'un calorimètre hadronique et de deux détecteurs Cherenkov permet une excellente identification des particules de l'état final de la désintégration $B^0 \to K^{*0}e^+e^-$. Entre 2019 et 2021 le détecteur LHCb subit une importante mise à niveau afin de pouvoir faire fonctionner l'expérience avec une luminosité de $2 \times 10^{33}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$, soit un facteur 5 par rapport à la luminosité actuelle. En particulier, l'ensemble de l'électronique du détecteur doit être mise à niveau avec de nouvelles cartes Front-End (CFE) adaptées à une lecture à 40 MHz. Cette thèse présente également la mise à niveau du module *Low Level Trigger* de la CFE qui est notamment responsable de l'identification des clusters d'énergie transverse maximale dans les calorimètres pour chaque évènement. En particulier, une simulation complète du *Low Level Trigger* a été réalisée dans le cadre de cette thèse. Cette simulation est utilisée pour tester les CFE et sera utilisée dans le futur pour leur monitoring en ligne.

En complément de la région $q^2 \in [0.0001, 0.25]\,\mathrm{GeV}^2/c^4$ (dénommée très-bas-q^2), la région $q^2 \in [0, 0.0001]\,\mathrm{GeV}^2/c^4$ (dénommée gamma-q^2) est également étudiée. La région gamma-q^2 est dominée par la désintégration $B^0 \to K^{*0}\gamma$ où le photon se convertit en une paire d'électrons en interagissant avec le matériel du détecteur. Parce qu'il présente des caractéristiques très proches du canal $B^0 \to K^{*0}e^+e^-$ pour un rapport d'embranchement environ 100 fois supérieur, la désintégration $B^0 \to K^{*0}\gamma(\to e^+e^-)$ est un excellent canal de contrôle pour la présente analyse. Les candidats $B^0 \to K^{*0}e^+e^-$ sont tout d'abords sélectionnés dans les données.

Les bruits de fonds potentiels sont ensuite analysés et, si nécessaire, leur distribution en masse reconstruite $m(K^+\pi^-e^+e^-)$ et leur forme angulaire sont modélisées. Un premier ajustement de la masse $m(K^+\pi^-e^+e^-)$ est effectué sur les données dans une fenêtre relativement large entre 4500 et 6200 MeV/c^2 afin de modéliser correctement la queue radiative et le bruit de fond partiellement reconstruit. Puis, un second ajustement des données en 4 dimensions (masse + 3 angles) est effectué dans une fenêtre de masse restreinte entre 5000 et 5400 MeV/c^2 afin de faciliter la modélisation angulaire des bruits de fond et de réduire leur contribution.

Une pré-sélection est appliquée pour sélectionner les candidats $B^0 \to K^{*0}e^+e^-$. Les candidats de l'état final sont soumis à des critères sur leur moment transverse, la qualité des traces et la qualité de leur identification. Il est également requis que la masse reconstruite du système $K^+\pi^-$ soit proche de celle du méson K^{*0}. Enfin, toutes les traces doivent converger en un unique vertex, lequel doit être déplacé par rapport au vertex primaire de la collision proton-proton.
Le bruit de fond dominant après cette pré-sélection provient de désintégrations semi-leptoniques du type $B^0 \to D^-(\to K^{*0}(\to K^+\pi^-)e^-\overline{\nu}_e)e^+\nu_e$, dont le rapport d'embranchement est environ 10^4 fois supérieur à celui du signal. Cependant seuls les évènements dont les neutrinos sont peu énergétiques passent la pré-sélection. En d'autres termes, ces évènements présentent une forte asymétrie entre l'énergie du positron et de l'électron. Étant donné que $\cos\theta_\ell$ est en première approximation proportionnel à cette asymétrie, le bruit de fond semi-leptonique présente une distribution centrée vers les hautes valeurs de $\cos\theta_\ell$. Afin d'éviter tout biais du paramètre A_T^{Re} une coupure symétrique $|\cos\theta_\ell| < 0.8$ est appliquée, supprimant 97% du bruit de fond semi-leptonique résiduel pour une efficacité de 95% sur le signal.
Les désintégrations $B^0 \to K^{*0}\gamma$ où le photon se convertit en une paire d'électrons est également un bruit de fond important puisque le rapport d'embranchement de ce type de désintégration est environ 100 fois supérieur à celui du signal. Cependant, la coupure en q^2 diminue fortement la contamination de ce bruit de fond à environ 30% du signal. Afin de réduire la contamination à moins de 2% du signal, les évènements dont le vertex d'origine de la paire d'électrons est situé dans une région où la densité de matériel dans le détecteur est élevée sont rejetés. Cette coupure additionnelle a une efficacité sur le signal d'environ 98%.
Les désintégrations $B^0 \to K^{*0}\eta(\to e^+e^-\gamma)$ et $B^0 \to K^{*0}\pi^0(\to e^+e^-\gamma)$ sont également une source de bruit de fond, dans les cas où l'énergie du photon est faible ou si le photon est ajouté à tort comme un photon de Bremsstrahlung à l'un des deux électrons. Malgré des rapports d'embranchements un à deux ordre(s) de grandeur(s) inférieurs au signal, la contamination de ces bruits de fonds peut atteindre jusqu'à 5% dans la fenêtre de masse utilisée pour l'analyse angulaire. Ceci

est dû au fait que la quasi-totalité de l'espace de phase de ces deux désintégrations se situe dans la région de q^2 de l'analyse.
Le bruit de fond partiellement reconstruit provenant d'une résonance de kaon génériquement appelée K^{**}, de type $B^+ \to K^{**}(\to K^+\pi^-\pi^+)e^+e^-$ où l'un des pions n'est pas reconstruit, peut également polluer le signal. Dans la zone de q^2 considérée dans la présente analyse, le bruit de fond partiellement reconstruit est dominé par la désintégration $B^+ \to K_1(1270)(\to K^+\pi^-\pi^+)e^+e^-$. Il est possible de diminuer fortement la contamination de ce type de bruit de fond en utilisant la répartition cinématique du moment entre la partie hadronique et la partie électronique de la désintégration. Grâce à cette méthode, 70% du bruit de fond partiellement reconstruit est supprimé tout en gardant 90% des désintégrations $B^0 \to K^{*0}e^+e^-$.
Enfin le bruit de fond combinatoire provenant de l'association aléatoire des nombreux produits de désintégration de la collision proton-proton constitue un important bruit de fond. Un algorithme de *machine learning* utilisant un *Boosted Decision Tree* (BDT) est entraîné pour différencier le bruit de fond combinatoire du signal. Le BDT est entraîné avec un échantillon "signal" constitué de simulation Monte Carlo (MC) de désintégrations $B^0 \to K^{*0}e^+e^-$, et un échantillon "bruit de fond" constitué de données $B^0 \to K^{*0}e^+e^-$ reconstruits avec une masse $m(K^+\pi^-e^+e^-)$ supérieure à $5600\,\mathrm{MeV}/c^2$. Après l'application de la coupure BDT, 90% du bruit de fond combinatoire est rejeté pour une efficacité signal de 94%.
Les bruits de fond modélisés dans l'ajustement de la masse reconstruite $m(K^+\pi^-e^+e^-)$ et des trois angles sont donc les désintégrations $B^0 \to K^{*0}\gamma(\to e^+e^-)$, $B^0 \to K^{*0}\eta(\to e^+e^-\gamma)$, $B^0 \to K^{*0}\pi^0(\to e^+e^-\gamma)$, le bruit de fond partiellement reconstruit, les bruits de fonds combinatoire et semi-leptoniques. Ces deux derniers sont modélisés ensembles.
La modélisation des distributions en masse du signal et des différents bruits de fonds est effectuée à partir de simulations MC, à l'exception du bruit de fond combinatoire/semi-leptonique qui est modélisé par une fonction exponentielle. Les seuls paramètres libres de l'ajustement sont les fractions des différentes composantes ainsi que la pente de la fonction exponentielle modélisant le bruit de fond combinatoire/semi-leptonique. L'ajustement est réalisé simultanément dans les deux régions de q^2. Le résultat de l'ajustement aux données est donné dans la figure 2.

La reconstruction et la sélection des candidats $B^0 \to K^{*0}e^+e^-$ peut affecter la distribution angulaire. C'est pourquoi l'acceptance angulaire est modélisée, en supposant la factorisation des acceptances individuelles càd $\varepsilon(\cos\theta_\ell, \cos\theta_K, \phi) = \varepsilon(\cos\theta_\ell) \times \varepsilon(\cos\theta_K) \times \varepsilon(\phi)$. Tout d'abord, une simulation MC d'évènements $B^0 \to K^{*0}e^+e^-$ est générée avec une distribution angulaire plate. L'ensemble du processus de reconstruction et de sélection est ensuite appliqué à ces évène-

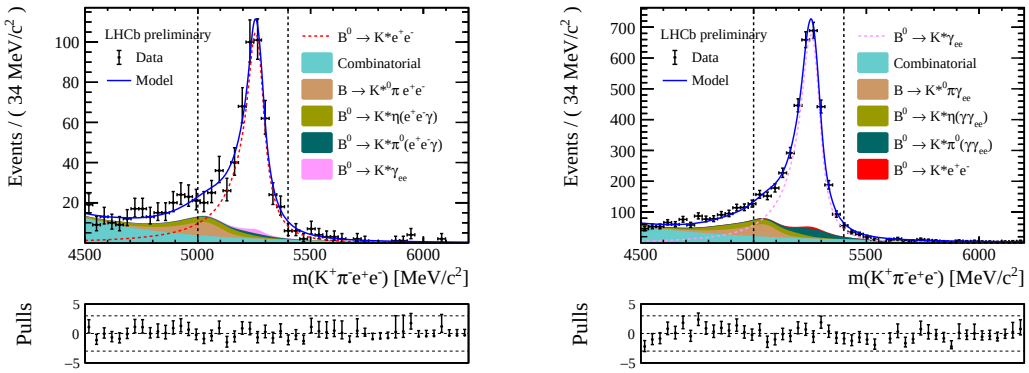


Figure 2: Ajustement de la masse reconstruite $m(K^+\pi^-e^+e^-)$ des données $B^0 \to K^{*0}e^+e^-$ dans la région très-bas-q^2 (gauche) et des données $B^0 \to K^{*0}\gamma(\to e^+e^-)$ dans la région gamma-q^2 (droite).

ments simulés. Puis, les distributions angulaires sont ajustées individuellement afin d'obtenir l'acceptance angulaire.
La forme angulaire des différents bruits de fonds est ensuite modélisée. Pour la plupart des bruits de fonds, la forme angulaire est modélisée avec des évènements simulés. Pour le bruit de fond combinatoire/semi-leptonique cependant, un échantillon de données reconstruits en $B^0 \to K^{*0}e^+\mu^-$ est utilisé. Étant donné que cette désintégration est interdite dans le MS, cet échantillon est constitué d'un mélange de bruit de fond combinatoire et semi-leptonique.

La plupart des erreurs systématiques sont évaluées à l'aide de pseudo-expériences. Deux types d'erreurs systématiques sont considérées. Les erreurs systématiques liées à la taille des échantillons utilisés pour les modélisations sont évaluées à l'aide d'une méthode appelée *bootstrapping*, qui consiste à générer un grand nombre de pseudo-échantillons en effectuant un tirage avec remise au sein de l'échantillon initial. Les erreurs systématiques liées à la modélisation des bruits de fonds et de l'acceptance angulaire sont évaluées en ajustant les données avec un modèle alternatif. Puis, un grand nombre de pseudo-échantillons sont générés à partir de cet ajustement. Enfin, chacun de ces pseudo-échantillons est ajusté à nouveau avec le modèle nominal. Le biais ainsi mis en évidence est assigné comme une erreur systématique. L'ensemble des erreurs systématiques est résumé dans le tableau 1.

L'ajustement en quatre dimensions ($m(K^+\pi^-e^+e^-)$, $\cos\theta_\ell$, $\cos\theta_K$, $\tilde{\phi}$) sur les

Source d'erreur systématique	$\sigma A_T^{(2)}$	σA_T^{Im}	σA_T^{Re}	$\sigma F_{\rm L}$
Acceptance: taille d'échantillon	< 0.007	< 0.007	< 0.007	< 0.003
Acceptance: modèle	0.004	0.001	0.008	0.001
Combinatoire/semi-lept.: taille d'échantillon	< 0.007	< 0.007	< 0.007	< 0.003
Combinatoire/semi-lept.: coupure $q2$	0.009	0.003	0.004	0.001
Combinatoire/semi-lept.: coupure BDT	0.007	0.003	0.003	0.005
Combinatoire/semi-lept.: fusion des Runs	0.002	0.003	0.004	0.001
Partiellement reconstruit: modèle	0.001	0.003	0.002	0.001
η/π^0: modèle	0.0004	0.0001	0.002	0.01
Résolution en ϕ	-0.004	-0.001	-	-
Échange $K^{*0} - \overline{K}^{*0}$	-	0.0002	0.0001	0.00004
Corrections MC	0.003	0.001	0.003	0.007
Données/MC diff. de forme en masse	0.002	0.002	0.004	0.001
Biais de l'ajustement	~ 0	~ 0	~ 0	-0.003
Total	$^{+0.016}_{-0.017}$	0.012	0.015	0.014
Erreur statistique	0.103	0.102	0.077	0.026

Table 1: Résumé des erreurs systématiques.

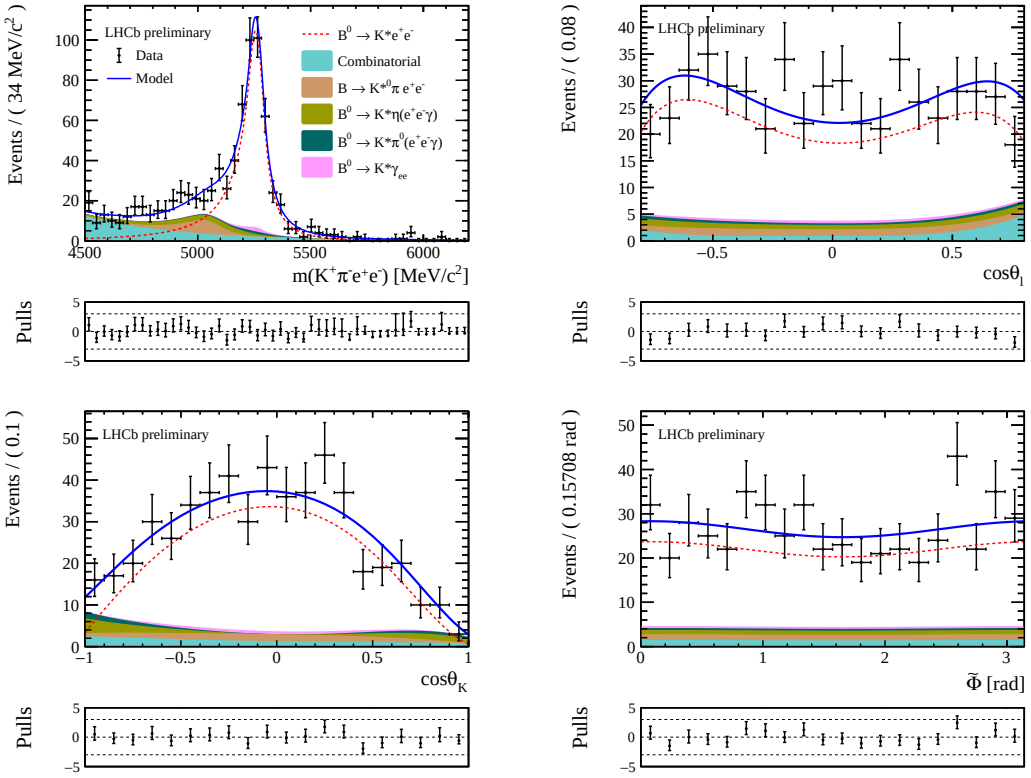


Figure 3: Ajustement simultané de la masse invariante $m(K^+\pi^-e^+e^-)$ et des trois angles $\cos\theta_\ell$, $\cos\theta_K$ et $\tilde{\phi}$ des données $B^0 \to K^{*0}e^+e^-$.

données $B^0 \to K^{*0}e^+e^-$ se trouve sur la figure 3 et a pour résultat

$$\begin{aligned} A_T^{(2)} &= 0.106 \pm 0.103\ ^{+0.016}_{-0.017} \\ A_T^{Im} &= 0.015 \pm 0.102 \pm 0.012 \\ A_T^{Re} &= -0.064 \pm 0.077 \pm 0.015 \\ F_{\rm L} &= 0.044 \pm 0.026 \pm 0.014 \qquad , \end{aligned} \tag{0.7}$$

où la première erreur est statistique tandis que la seconde est systématique. Afin de comparer ce résultat expérimental aux prédictions du MS, il est nécessaire de construire la région de q^2 effectif de l'analyse. En effet, en raison de l'efficacité qui varie en fonction de q^2 et de la résolution en q^2, la région de q^2 reconstruit est légèrement différente du vrai q^2, comme l'illustre la figure 4. Une région de q^2 effectif, tenant compte des effets de reconstruction, sélection et de résolution est ainsi déterminée à hauteur de $[0.0008 \pm 0.0004,\ 0.257 \pm 0.003]\ \text{GeV}^2/c^4$.

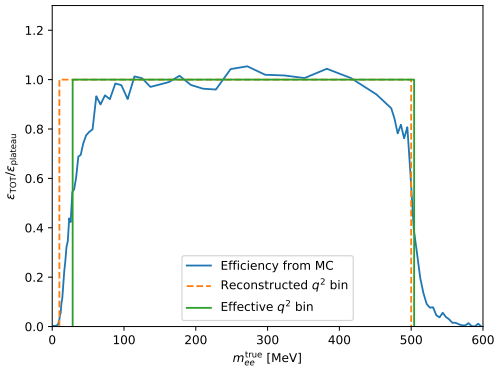


Figure 4: Efficacité de reconstruction et sélection des candidats $B^0 \to K^{*0}e^+e^-$ simulés (bleu). La région de q^2 reconstruit est tracée en pointillés orange, tandis que la région de q^2 effectif est tracée en vert.

Les prédictions du MS dans cette région de q^2 effectif sont comparées aux résultats expérimentaux dans le tableau 2. Les valeurs expérimentales sont compatibles avec les prédictions théoriques à hauteur de 0.5σ.

	Prédictions du MS [12]	Mesure LHCb
$A_T^{(2)}$	0.034 ± 0.021	$0.106 \pm 0.103\ ^{+0.016}_{-0.017}$
A_T^{Im}	$(-0.9 \pm 3.5) \times 10^{-4}$	$0.015 \pm 0.102 \pm 0.012$
A_T^{Re}	$(-6.3 \pm 0.5) \times 10^{-3}$	$-0.064 \pm 0.077 \pm 0.015$
F_{L}	0.051 ± 0.014	$0.044 \pm 0.026 \pm 0.014$

Table 2: Comparaison des prédictions du MS dans la région de q^2 effectif avec les mesures expérimentales présentées dans cette thèse.

La mesure présentée dans cette thèse représente actuellement la plus importante contrainte mondiale sur la polarisation droite du photon dans les transitions de type $b \to s\gamma$, comme l'illustre la figure 5.

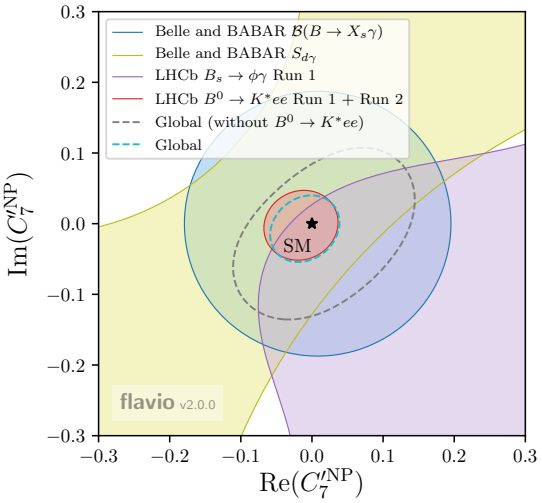


Figure 5: Contrainte actuelle de la contribution de NP sur la partie réelle et imaginaire du coefficient de Wilson $\mathcal{C}_7'$. La contrainte provenant de la mesure présentée dans cette thèse est tracée en rouge. La contrainte globale de toutes les mesures (excepté celle présentée dans cette thèse) est tracée en pointillés gris, tandis que la contrainte globale de toutes les mesures (incluant celle présentée dans cette thèse) est tracée en pointillés bleu clair.

Chapter 1

Theoretical and experimental context

Physics, from Ancient Greek $\varphi\upsilon\sigma\iota\kappa\acute{\eta}$ - literally "the knowledge of Nature" - is the science which aims at describing the Universe and its laws. Even though the laws of physics are generally assumed to be invariant in space and time - that is, the laws of Nature are assumed to be the same at any place and any time in the Universe - they are found to vary quite dramatically depending on the size or the speed of the physical system.
Quite naturally, physicists started to make experiments and build theories to describe phenomena which are close to human scale: objects of typically a few meters in size and relatively slow, *i.e.* much slower than the speed of light in vacuum ($\sim 3\times10^8$ m/s). It took until the end of the nineteenth century to get a global and coherent description of almost all known phenomena close to the human scale by the means of three complementary and entangled theories: Classical Mechanics to describe the motion of objects, Classical Electrodynamics to describe light, electricity and magnetism, and Classical Thermodynamics to describe heat exchanges and phase transformations (solid, liquid, gaseous). These three theories taken together are usually referred to as Classical Physics.
Although very effective and predictive, Classical Physics was unable to explain a few phenomena which were puzzling physicists at the end of the nineteenth century. One of the most famous failures of Classical Physics was the explanation of the results of the Michelson and Morley experiment. The latter was an interferometry experiment aiming at measuring the relative speed of the Earth with respect to the aether (the medium in which light was believed to travel). To overcome the problem, Albert Einstein formulated the theory of Special Relativity in 1905. While being in accordance with Classical Physics at low velocities, Special Relativity predicts the existence of additional effects which cannot be neglected at very high velocities - which is the case of light in the Michelson and Morley experiment.

Around the same period, another prediction of Classical Physics was challenged by experiment. Indeed, Classical Electrodynamics and Thermodynamics predict that an object with a non vanishing temperature emits radiation in all frequency ranges, emitting more energy as the frequency increases (the energy density is predicted to be proportional to the frequency squared). This would mean that any object would instantaneously radiate all its thermal energy, which is obviously in contradiction with observations. The solution has been found in the early twentieth century by Max Planck who introduced the fact that at the microscopic level, energy is quantized. This lead to the development of Quantum Mechanics, the theory that describes the laws of Nature at very small scales. At macroscopic scales, most quantum effects can be neglected, and indeed in the limit where the energy quanta goes to zero, many predictions of Quantum Mechanics match to the ones of Classical Physics.
Soon after his proposal of Special Relativity, Albert Einstein introduced the concept of dynamical space-time in his theory of General Relativity, where very massive objects like stars can induce a curvature in space-time. This theory is thus needed to properly describe the evolution of large scale structures like galaxies or the whole Universe itself, where the curvature of space-time cannot be neglected anymore. Indeed, one of the first historical test of General Relativity was to explain the anomalous (with respect to Classical Mechanics predictions) precession of Mercury, due to the curved space-time in the vicinity of the Sun.
To finish, to describe subatomic particles moving at high velocities, one has to combine Special Relativity and Quantum Mechanics. This has given birth to Quantum Field Theory which has been developed during the mid to end twentieth century and has lead to the Standard Model (SM) of particle physics, which is described in more details in the next section.
It is interesting to note that all these theories are not in direct contradiction with Classical Physics but have rather to be seen as extensions needed for specific regimes. Indeed, low space-time curvature General Relativity, Quantum Mechanics at large scale or Special Relativity at low speed all tend to Classical Physics. Quite remarkably however, there is no such smooth transition between Quantum Mechanics (or Quantum Field Theory) and General Relativity. Binding these two theories inside a unique description is one of the major challenges of contemporary physics.

1.1 The Standard Model of particle physics

The Standard Model (SM) of particle physics is the best known theory to describe subatomic particles and their interactions. It describes three of the four fundamental interactions, namely the electromagnetic, the strong and the weak interactions.

Atoms which form ordinary matter are bound states of a positively electrically charged nucleus and one or several negatively electrically charged electrons. The electrostatic force which binds them together is the low energy manifestation of the electromagnetic interaction. The nucleus itself is made of protons and neutrons which are bound together by the strong interaction. The third interaction described by the SM is the weak interaction which is for example responsible for the nuclear β-decay of radioactive isotopes, and plays an important role to initiate nuclear fusion which fuels the Sun. Only one fundamental interaction, gravity, is missing in the SM[1].
The three interactions are mediated by spin 1 particles which are called gauge vector bosons. The gluons mediate the strong force. There are eight types of gluons, each carrying one color and one anti-color. The color is the charge of the strong interaction (akin to the electric charge for the electromagnetic force). The Z and the W bosons are the mediators of the weak force, while the photon is the mediator of the electromagnetic force. The Z and the W bosons are massive and have three degrees of freedom (polarization), while the gluons and the photon are massless and have two degrees of freedom (polarization). The measurement of the polarization of the photon in decays involving a transition from a bottom to a strange quark (see below) is the main goal of this thesis.
In addition, there are twelve fermions (spin 1/2) divided into two groups (quarks and leptons) which further constitute the building blocks of the SM. Each fermion f also has its associated anti-particle noted $\bar{f}$. Each group of fermions is discussed in more details below.

- **Quarks** carry a color charge which makes them subject to the strong interaction. Since they also carry an electric charge, they are subject to the electromagnetic interaction as well. As all elementary fermions, quarks are subject to the weak interaction. Quarks are the constituents of mesons (bound state of a quark and an anti-quark) and baryons (bound state of three quarks)[2]. Among the remarkable baryons are the protons and the neutrons which are the constituents of the nuclei of all ordinary matter. The bound state of a down quark and a beauty anti-quark is called a B^0 meson whose rare decay to a K^{*0} meson (the bound state of a down quark and a strange anti-quark), an electron and an anti-electron (see below) is studied in detail in this thesis.

- **Leptons** are divided into charged leptons and neutrinos. They do not carry a color charge and are thus not subject to the strong interaction. Contrary to

[1]Practically, this is of little importance. Indeed, due to the very small masses of fundamental particles, the effects of gravity are completely negligible with respect to the three other fundamental interactions.

[2]There are hints of exotic bound states of more than three quarks, see for instance [13]

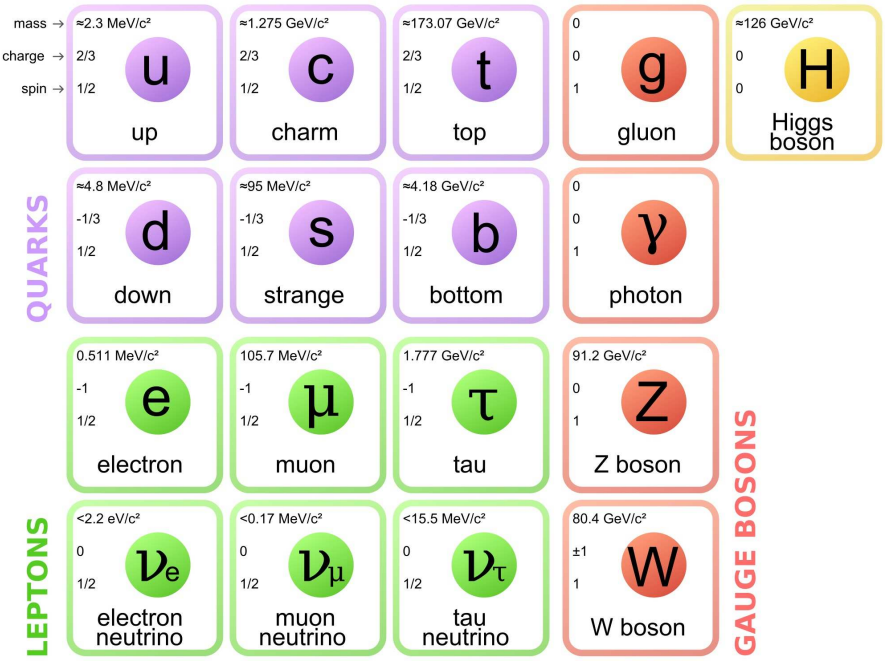


Figure 1.1: Elementary particles of the SM with their mass, electric charge and spin. Picture taken from PBS NOVA, Fermilab, Office of Science, United States Department of Energy, Particle Data Group.

charged leptons, neutrinos are electrically neutral and are thus only subject to the weak interaction. One of the most remarkable lepton is the electron which, when electrically bound to a nucleus, forms the atoms of all ordinary matter.

The last piece of the SM is the Higgs scalar boson which provides mass to the particles it interacts with. The Higgs potential is responsible for the spontaneous symmetry breaking of the electroweak group (see below). A summary of the full list of the elementary particles of the SM and their main properties is given in figure 1.1.

The SM is based on symmetries. Besides the Lorentz-Poincarré and CPT symmetries ensuring momentum and energy conservation, the SM is built on the local gauge invariance with respect to the group [14, 15]

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \quad , \tag{1.1}$$

where $SU(3)_C$ describes the strong interaction and $SU(2)_L \otimes U(1)_Y$ describes the electroweak interaction. The gauge Lagrangian density of the SM is thus given

by

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}G^a_{\mu\nu}G^{a,\mu\nu} - \frac{1}{4}W^b_{\mu\nu}W^{b,\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \qquad , \tag{1.2}$$

where $G^a_{\mu\nu}$ are the gluon fields with $a = \{1, 2, ..., 8\}$ while $W^b_{\mu\nu}$ and $B_{\mu\nu}$ are the primary electroweak fields with $b = \{1, 2, 3\}$. Adding the fermions and the Higgs scalar boson, the full Lagrangian density of the SM reads

$$\mathcal{L}_{\text{SM}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\bar{\psi}\not{D}\psi + h.c. + \psi_i y_{ij}\psi_j\phi + h.c. + |D_\mu\phi|^2 - V(\phi) \qquad , \tag{1.3}$$

where ψ are the fermion fields (one for each fermion), ϕ is the Higgs scalar field, $\not{D}$ is the scalar product of the gamma matrices γ^μ and the SM covariant derivative D_μ, and *h.c.* refers to the hermitian conjugate of the previous term. The first term is the scalar product of the field strength tensor $F_{\mu\nu}$ encoding the gauge vector boson fields and their interactions with each other. The second term describes how vector gauge bosons interact with fermions, *i.e.* how the strong, electromagnetic and weak forces interact with fermions. The third term describes how particles couple to the Higgs scalar boson and thereby how they acquire some mass, while the components of the Yukawa matrix y_{ij} are the coupling parameters to the Higgs field of each particle. The fourth term describes how the Z and W bosons interact with the Higgs field and thereby how they obtain their mass. To finish, the last term is the potential of the Higgs field.
One of the major features of the SM is the spontaneous symmetry breaking (SSB) of $SU(2)_L \otimes U(1)_Y$ induced by the Higgs field. Indeed, the Higgs potential has the form

$$V(\phi) = \lambda(\phi^\dagger\phi)^2 + \mu^2(\phi^\dagger\phi) \qquad , \tag{1.4}$$

where the potential is chosen such that $\mu^2 < 0$ and $\lambda > 0$. This leads to a potential with two degenerate non-zero minima given by

$$\phi^\dagger\phi = -\frac{\mu^2}{\lambda} \equiv \frac{v^2}{2} \qquad , \tag{1.5}$$

with a non vanishing vacuum expectation value v. In the process of the SSB, the primary electroweak fields $W^b_{\mu\nu}$ and $B_{\mu\nu}$ mix and form the physically observables states

$$\begin{cases} Z_\mu & = \cos\theta_{\text{W}} W^3_\mu - \sin\theta_{\text{W}} B_\mu \\ W^\pm_\mu & = \frac{W^1_\mu \mp W^2_\mu}{\sqrt{2}} \\ A_\mu & = \sin\theta_{\text{W}} W^3_\mu + \cos\theta_{\text{W}} B_\mu \qquad , \end{cases} \tag{1.6}$$

where θ_W is the Weinberg angle, A_μ is the photon field while Z_μ and $W_\mu^\pm$ are the Z and W bosons fields. The W and Z bosons acquire some mass through the process while the photon remains massless.
In addition, the SSB is also the process through which fermions (except neutrinos which are massless in the SM) acquire their mass via the Yukawa coupling to the Higgs field. The Yukawa part of the Lagrangian density is given by

$$\mathcal{L}_{\text{Yukawa}} = -y^d_{ij}\bar{Q}_{L,i}\phi d_{R,j} - y^u_{ij}\bar{Q}_{L,i}\tilde{\phi}u_{R,j} - y^e_{ij}\bar{L}_{L,i}\phi e_{R,j} + h.c. \quad , \tag{1.7}$$

where $\bar{Q}_{L,i}$ and $\bar{L}_{L,i}$ are the left handed quark and lepton doublet respectively while $d_{R,j}$, $u_{R,j}$ and $e_{R,j}$ are the right handed down quark, up quark and lepton singlets respectively. The indices i and j run over the three generations of fermions. Replacing ϕ by its vacuum expectation value and considering only quarks for simplicity the Yukawa part of the Lagrangian density now reads

$$\mathcal{L}^{\text{quarks}}_{\text{Yukawa}} = -\bar{u}_{R,i}\mathcal{M}^u_{ij}u_{L,j} - -\bar{d}_{R,i}\mathcal{M}^d_{ij}d_{L,j} \quad , \tag{1.8}$$

where the quark mass matrices in the flavor basis each containing nine complex numbers are given by

$$\begin{cases} \mathcal{M}^u_{ij} & = \frac{v}{\sqrt{2}}y^u_{ij} \\ \mathcal{M}^d_{ij} & = \frac{v}{\sqrt{2}}y^d_{ij} \end{cases} . \tag{1.9}$$

In order to find the quark mass eigenstates one has to diagonalize the $\mathcal{M}^u_{ij}$ and $\mathcal{M}^d_{ij}$ matrices which can be achieved by defining four unitary matrices U_L, U_R, D_L and D_R such that

$$\begin{aligned} U_R^{-1}\mathcal{M}^u U_L &= \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix} \\ \text{, and } \quad D_R^{-1}\mathcal{M}^d D_L &= \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix} \quad , \end{aligned} \tag{1.10}$$

where the masses of the quarks m_i are free parameters of the SM. Since the up-type quarks and the down type quarks have been rotated using different matrices U and D the $SU(2)$ symmetry of the left-handed quark doublets is broken. By the term broken, it is meant that the rotation by two different matrices induces quark mixing via charged weak current mediated by the W bosons of the form

$$J^{+,\mu} = (\bar{u}_1, \bar{u}_2, \bar{u}_3)_L \, \gamma^\mu \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}_L \quad , \tag{1.11}$$

and the analogous negatively charged current $J^{-,\mu}$, where the index $\{1,2,3\}$ represents the quark generation in the flavor basis. Rotating to the mass basis one gets

$$J^{+,\mu} = (\bar{u},\bar{c},\bar{t})_L\, U_L^\dagger\, \gamma^\mu D_L \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L = (\bar{u},\bar{c},\bar{t})_L\, \gamma^\mu V^{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L \quad , \tag{1.12}$$

where the Cabibbo-Kobayashi-Maskawa (CKM) matrix is defined as $V^{\mathrm{CKM}} \equiv U_L^\dagger D_L$. Being a 3×3 complex unitary matrix, the CKM matrix can be parameterized by three mixing angles and one complex phase (the origin of CP violation in the SM), which are free parameters of the SM. However, a strong hierarchy is found among the CKM matrix elements experimentally. The CKM matrix can be explicated as

$$\begin{aligned} V^{\mathrm{CKM}} &= \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \\ &= \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}\left(\lambda^4\right) \quad , \end{aligned} \tag{1.13}$$

where the second form is obtained through the Wolfenstein parameterization [16] with the expansion parameter $\lambda \simeq 0.24$, while the other parameters are given by $A \simeq 0.81$, $\rho \simeq 0.14$ and $\eta \simeq 0.35$. The neutral current interaction mediated by the Z boson and the photon however can be directly written in the flavor basis since the coupling of the Z boson and the photon to each fermion generation are universal. As a consequence, flavor changing neutral currents (FCNCs), such as $b \to s$ transitions studied in this thesis, are forbidden at tree level in the SM.

Despite its unprecedented success to predict a vast amount of observables verified by experiment at a very high precision, the SM presents some shortcomings. Some of the main limitations of the SM are briefly mentioned below.
First of all, the strong hierarchy experimentally observed in the fermion masses and the CKM matrix elements are unexplained by the SM (since they are free parameters of the theory). Moreover, the SM does not provide any mechanism to attribute a non-vanishing mass to neutrinos, implied by observed neutrino oscillations [17]. In addition, it is inconsistent with two major cosmological observations: it does not provide any Dark Matter candidate [18] nor enough CP violation to explain the matter/anti-matter asymmetry of the Universe [19]. To finish, the SM

does not include gravity due to serious incompatibilities between Quantum Field Theory and General Relativity.
This thus motivates the need for a theory Beyond the Standard Model (BSM) which could solve some if not all of the shortcomings of the SM.

1.2 Probing New Physics in rare $b \to s\gamma$ transitions

There are two main ways of searching for New Physics (NP). The most straightforward way is via direct searches by trying to produce the NP particles and study their properties. This is typically done in collider experiments, which deliver a large kinetic energy which can be transformed into mass energy at the collision point. The main limit of direct searches is the beam energy, which has to be greater than the mass of the NP particle times the speed of light squared for the particle to have any chance to be produced. The largest beam energy available in the world is delivered at the Large Hadron Collider (LHC) at CERN which has provided a center of mass energy up to 13 TeV. This thus allows to probe NP particles with masses up to the TeV scale. The most important constraints on NP searches via this direct method have been obtained by the ATLAS and CMS experiments (see for instance [20]). The second way of probing NP, also called indirect searches, is via the potential contributions of NP particles in loop processes which can affect specific observables (see below). Since the NP particles appearing in the quantum loop are virtual, the scale which can be probed is much larger than the beam energy. Such indirect approaches have been historically very successful, one of the most famous example being the interpretation of CP violation in kaon decays as the proof of the existence of a third quark family [21, 22] several years before quarks of the third family had been directly observed [23, 24]. However, the drawback of these indirect searches is that they only provide results in a 2D plane (coupling, NP scale).

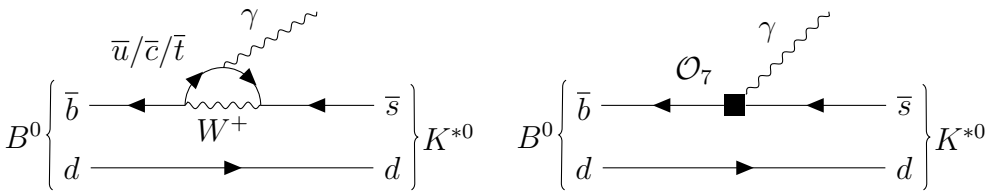


Figure 1.2: Example of a $b \to s\gamma$ transition with the $B^0 \to K^{*0}\gamma$ decay. The full Feynman diagram is shown on the left while the effective diagram is shown on the right.

Processes involving a transition from a bottom to a strange quark ($b \to s$) are FCNCs which are forbidden at tree level in the SM (see section 1.1). They can thus

only occur at loop level. This makes them sensitive to NP which may manifest itself via new tree level processes or new contributions in the loop. The absence of dominating tree level SM contribution makes $b \to s$ transitions a very popular laboratory for indirect searches for NP by looking at deviations with respect to SM predictions from specifically chosen sets of observables, in particular angular observables and ratios of branching fractions.
In addition, the mass of the b quark is much smaller than the electroweak (and top quark) scale. Thus, similarly to Fermi's four fermion interaction, one can build an effective low-energy theory to describe the quark interactions. The main idea of the Operator Product Expansion (OPE) [1] is to separate low-energy (large distance) from high energy (small distance) degrees of freedom by building an effective Hamiltonian of the form

$$\langle f|\mathcal{H}_{\text{eff}}|i\rangle = \sum_j \mathcal{C}_j \langle f|\mathcal{O}_j|i\rangle \qquad , \tag{1.14}$$

where the low-energy part is encoded in the local operators $\mathcal{O}_j$ while the high energy part is contained in the Wilson coefficients $\mathcal{C}_j$. In other words, the propagation over distances typically smaller than $1/\mu$ (where μ is the factorization scale) of the heavy degrees of freedom (the W, Z, Higgs bosons and top quark in the SM) are neglected and reduced to point-like interactions with an effective coupling given by the Wilson coefficients (see figure 1.2 for an illustration). For $b \to s$ transitions the effective Hamiltonian takes the form [25]

$$\mathcal{H}_{\text{eff}} = -\frac{4G_F}{\sqrt{2}} V_{tb}V_{ts}^* \sum_j (\mathcal{C}_j\mathcal{O}_j + \mathcal{C}_j'\mathcal{O}_j') \qquad , \tag{1.15}$$

where G_F is the Fermi constant and V_{ij} are the CKM matrix elements. The relevant operators for $b \to s\gamma$ transitions are

$$\begin{aligned} \mathcal{O}_7 &= \frac{e}{16\pi^2} m_b (\bar{s}\sigma_{\mu\nu} P_R b) F^{\mu\nu} \qquad , \\ \mathcal{O}_7' &= \frac{e}{16\pi^2} m_b (\bar{s}\sigma_{\mu\nu} P_L b) F^{\mu\nu} \qquad , \end{aligned} \tag{1.16}$$

where $\sigma_{\mu\nu} = i/2[\gamma_\mu, \gamma_\nu]$, e is the Quantum Electrodynamics (QED) coupling constant and $P_{L,R} = (1 \mp \gamma_5)/2$ are the projectors on the left- (right-) handed chirality.
The Wilson coefficients are first computed at the high energy scale $\mu_{EW} \sim \mathcal{O}(m_t, M_W)$ by matching the effective theory to the full SM. Then, the renormalization group equations, which are used to compute the Wilson coefficients at the low energy scale $\mu_b \sim \mathcal{O}(m_b)$, insure that all relevant Quantum Chromodynamics (QCD) and

electroweak corrections are taken into account. This framework can be easily extended to various NP scenarios, which will manifest themselves by changing the values of Wilson coefficients or by adding or enhancing the ones which are absent or strongly suppressed in the SM.
The Wilson coefficients $\mathcal{C}_7^{(\prime)}$ describe the helicity structure of $b \to s\gamma$ processes. Indeed, in the SM the emitted photon is predominantly left-handed (right-handed) in b ($\bar{b}$) decays. Due to the left-handedness of the weak interaction, the fraction of right-handed photons is suppressed by a factor m_s/m_b. In terms of Wilson coefficients this means that in the SM $\mathcal{C}_7'/\mathcal{C}_7 \simeq m_s/m_b \simeq 0.02$ (for real $\mathcal{C}_7'$) up to QCD corrections [2]. On the other hand, many BSM models - including Minimal Supersymmetric Standard Models (MSSM) [3–5] and Left-Right Symmetric Models (LRSM) [6] - allow large right-handed currents, *i.e.* large $\mathcal{C}_7'$. Measuring the photon polarization in $b \to s\gamma$ processes is thus an excellent way of probing NP.

1.3 Current status of the measurement of the photon polarization in $b \to s\gamma$ transitions

Several measurements have constrained the allowed values of the Wilson coefficients $\mathcal{C}_7^{(\prime)}$ associated to the electroweak penguin operators $\mathcal{O}_7^{(\prime)}$, as illustrated by figure 1.3 A short review is presented in this section.

1.3.1 Inclusive branching fractions

First, inclusive branching fractions of $B \to X_s\gamma$, where X_s is any state containing an s quark, are proportional to $|\mathcal{C}_7|^2+|\mathcal{C}_7'|^2$. Thus, even though they are not directly sensitive to the handedness of the photon polarization, they can put a circular constraint in the $(\mathcal{C}_7, \mathcal{C}_7')$ plane. The inclusive branching fraction $\mathcal{B}(B \to X_s\gamma)$ has been measured by the Belle, BaBar and CLEO experiments. The world average is ([7] and references therein) $\mathcal{B}(B \to X_s\gamma)_{exp}(E_\gamma > 1.6\,\text{GeV}) = (3.32 \pm 0.15) \times 10^{-4}$ which is in very good agreement with the SM prediction $\mathcal{B}(B \to X_s\gamma)_{th}(E_\gamma > 1.6\,\text{GeV}) = (3.36 \pm 0.23) \times 10^{-4}$ [26].

1.3.2 Time dependent rate of radiative decays

The time dependent decay rate at which a B_q^0 ($\bar{B}_q^0$) meson, where $q = \{d, s\}$, decays to a common final state containing a photon is proportional to [10]

$$\mathcal{P}(t) \propto e^{-\Gamma t}\left[\cosh\frac{\Delta\Gamma_q t}{2} - \mathcal{A}^\Delta \sinh\frac{\Delta\Gamma_q t}{2} + \xi C\cos(\Delta m_q t) - \xi S\sin(\Delta m_q t)\right] \quad , \tag{1.17}$$

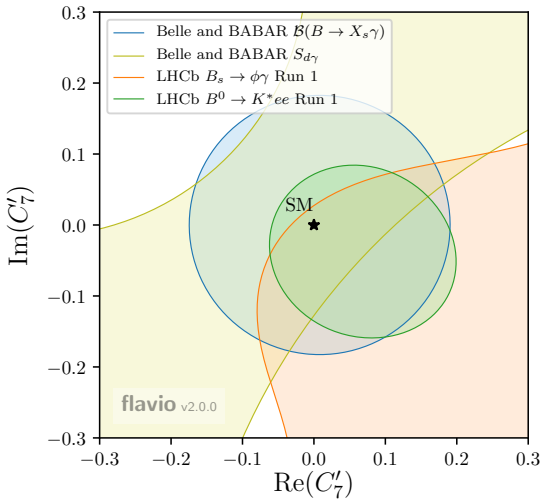


Figure 1.3: Constraints (at 2σ level) to $\mathcal{C}_7'$ from published measurements. The constraint from inclusive branching fractions and $S_{d\gamma}$ measurements by the Belle and BaBar experiments are shown in blue and yellow respectively. The constraints from LHCb measurements using the $3\,\mathrm{fb}^{-1}$ of Run 1 data are shown in orange for the $B_s^0 \to \phi\gamma$ decay observables and in green for the angular analysis of the $B^0 \to K^{*0}e^+e^-$ decay. The prediction of the SM is represented by the black star.

where $\Delta\Gamma_q$ and Δm_q are the width and mass differences between the B_q^0 mass eigenstates, Γ is the mean decay width between these eigenstates, $\xi = +1(-1)$ in the case of a B_q^0 ($\bar{B}_q^0$), C is the direct CP-asymmetry, $\mathcal{A}^\Delta$ is the CP-asymmetry related to the width difference and S is the CP-asymmetry induced by the mixing/decay interference. Both S and $\mathcal{A}^\Delta$ are sensitive to $\mathcal{C}_7^{(\prime)}$. Indeed, in the SM B_q^0 ($\bar{B}_q^0$) mesons decay predominantly into right-(left-)handed photons. As a consequence, the dominant amplitudes of $\bar{B}_q^0 \to f^{CP}\gamma_L$ and $\bar{B}_q^0 \to B_q^0 \to f^{CP}\gamma_R$, where f^{CP} is a CP eigenstate, cannot interfere quantum mechanically, since their final state is different. Thus, the CP-asymmetry due to the $B_q^0 - \bar{B}_q^0$ mixing is suppressed up to $\mathcal{O}(m_s/m_b)$ corrections. In terms of Wilson coefficients, the CP-asymmetries $S_{q\gamma}$ and $\mathcal{A}_{q\gamma}^\Delta$ can be expressed as [2]

$$\begin{aligned} S_{q\gamma} &\approx \xi \frac{2\mathcal{I}m[e^{-i\phi_q}\mathcal{C}_7\mathcal{C}_7{}']}{|\mathcal{C}_7|^2 + |\mathcal{C}_7'|^2} \quad , \\ \mathcal{A}_{q\gamma}^\Delta &\approx \xi \frac{2\mathcal{R}e[e^{-i\phi_q}\mathcal{C}_7\mathcal{C}_7{}']}{|\mathcal{C}_7|^2 + |\mathcal{C}_7'|^2} \quad , \end{aligned} \tag{1.18}$$

where ϕ_q is the phase of the $B_q^0 - \bar{B}_q^0$ mixing. In the standard model $\phi_d \simeq 43^\circ$ while $\phi_s \simeq -2^\circ$ for the B^0 and B_s^0 mixing respectively [10].

Experiment	Decay channel	$S_{d\gamma}$
BaBar and Belle average	$B^0 \to K^0_S\pi^0\gamma$ (including $B^0 \to K^{*0}(\to K^0_S\pi^0)\gamma$)	-0.15 ± 0.20
BaBar and Belle average	$B^0 \to K^0_S\rho\gamma$	-0.06 ± 0.23
BaBar	$B^0 \to K^0_S\eta\gamma$	$-0.18^{+0.49}_{-0.46} \pm 0.12$
Belle	$B^0 \to K^0_S\phi\gamma$	$0.74^{+0.72\ +0.10}_{-1.05\ -0.24}$
Theory	-	-0.023 ± 0.016

Table 1.1: Experimental measurements and theory prediction of the SM of the CP-asymmetry $S_{d\gamma}$ ([27] and references therein, [8, 9]). When given, the third term corresponds to systematic uncertainties.

The $S_{d\gamma}$ CP-asymmetry has been measured by the BaBar and Belle experiments in various decay channels. The most precise measurement has been obtained in $B^0 \to K^0_S\pi^0\gamma$ (including $B^0 \to K^{*0}(\to K^0_S\pi^0)\gamma$) decays. It is followed closely in precision by the measurement in the $B^0 \to K^0_S\rho\gamma$, while the measurements in the $B^0 \to K^0_S\eta\gamma$ and $B^0 \to K^0_S\phi\gamma$ are much less precise. All the experimental results are summarized in table 1.1 and are compatible with the theory prediction of the SM, given in the same table. The LHCb experiment has also provided a measurement of the $S_{s\gamma,exp} = 0.43 \pm 0.30 \pm 0.11$ CP-asymmetry in $B^0_s \to \phi\gamma$ decays, which is compatible with the SM prediction [10] $S_{th\gamma,exp} = 0 \pm 0.0002$. Although less precise than the most precise measurements of BaBar and Belle it is a unique and complementary measurement since B^0_s mesons are not produced by B-factories working at the $\Upsilon(4S)$ resonance.

Furthermore, one can exploit the large B^0_s width difference $\Delta\Gamma_s$ to measure $\mathcal{A}^{\Delta}_{s\gamma}$, with an important advantage being that the knowledge of the flavor of the B_s meson, which is experimentally challenging at LHCb, is not needed. The $\mathcal{A}^{\Delta}_{s\gamma}$ CP-asymmetry has been measured by LHCb in $B^0_s \to \phi\gamma$ decays [8] and yields $\mathcal{A}^{\Delta}_{s\gamma,\mathrm{exp}} = -0.67^{+0.37}_{-0.41} \pm 0.17$ which is compatible with the SM prediction [10] $\mathcal{A}^{\Delta}_{s\gamma,\mathrm{th}} = 0.047 \pm 0.025 \pm 0.15$.

1.3.3 Up-down asymmetry

The photon polarization can also be directly measured in three body radiative decays of $B \to K\pi\pi\gamma$. Indeed, the photon polarization parameter λ_γ is defined by [28]

$$\lambda_\gamma \equiv \frac{|\mathcal{A}(\bar{B} \to K^i_{\mathrm{res}}\gamma_R)|^2 - |\mathcal{A}(\bar{B} \to K^i_{\mathrm{res}}\gamma_L)|^2}{|\mathcal{A}(\bar{B} \to K^i_{\mathrm{res}}\gamma_R)|^2 + |\mathcal{A}(\bar{B} \to K^i_{\mathrm{res}}\gamma_L)|^2} \quad , \tag{1.19}$$

where $\mathcal{A}(\bar{B} \to K^i_{\rm res}\gamma_{R/L})$ are the weak radiative amplitudes of a $\bar{B}(b\bar{q})$ species to a K resonance i for right- and left-handed photons. In terms of Wilson coefficients, the photon polarization parameter is given by

$$\lambda_\gamma \simeq \frac{|\mathcal{C}_7'|^2 - |\mathcal{C}_7|^2}{|\mathcal{C}_7'|^2 + |\mathcal{C}_7|^2} \quad . \tag{1.20}$$

A convenient way of measuring λ_γ is by looking at the proportion of photons emitted above or below the normal plane of the three final state hadrons. This observable, which is called the up-down asymmetry $\mathcal{A}_{\rm ud}$ is given by

$$\mathcal{A}_{\rm ud} \equiv \frac{\int_0^1 d\cos\theta \frac{d\Gamma}{d\cos\theta} - \int_{-1}^0 d\cos\theta \frac{d\Gamma}{d\cos\theta}}{\int_{-1}^1 d\cos\theta \frac{d\Gamma}{d\cos\theta}} \quad , \tag{1.21}$$

where $d\Gamma$ is the differential decay width and θ is the angle between the z-axis and the vector orthogonal to the plane of the three final state hadrons. The up-down asymmetry $\mathcal{A}_{\rm ud}$ has been measured by LHCb [29] in the $B^+ \to K^+\pi^-\pi^+\gamma$ decay channel in four bins of $m(K^+\pi^-\pi^+)$ in order to cope with the major difficulty of this measurement which is the isolation of the different resonant contributions of the complex $K^+\pi^-\pi^+$ spectrum and their interferences. It is the first direct observation at a 5.2σ significance level of a non-zero parity violating photon polarization in $b \to s\gamma$ processes. Nevertheless, this measurement does not put any constraint on $\mathcal{C}_7^{(\prime)}$. Indeed the up-down asymmetry $\mathcal{A}_{\rm ud}$ is only proportional to the photon polarization parameter through $\mathcal{A}_{\rm ud} \propto \kappa\lambda_\gamma$. The term κ depends in particular on the various resonances of the $K^+\pi^-\pi^+$ system and is very challenging to estimate. Thus, up to now, the measurement of $\mathcal{A}_{\rm ud}$ only allows to perform a null test of the polarization of the photon. In addition, even under the hypothesis of a precise knowledge of the term κ, the photon polarization parameter λ_γ is only proportional to the squared Wilson coefficients. Thus it is not sensitive to small NP contributions and any uncertainty on the measurement of λ_γ is propagated quadradically to Wilson coefficients.

1.3.4 Measuring the photon polarization in the baryonic sector

The photon polarization can also be measured in baryonic $b \to s$ transitions such as $\Lambda_b^0 \to \Lambda\gamma$ [30]. Indeed, the differential decay width of this decay is proportional to [31]

$$\frac{d\Gamma}{d\cos\theta_\gamma} \propto 1 - \alpha_\gamma P_{\Lambda_b^0} \cos\theta_\gamma \quad , \tag{1.22}$$

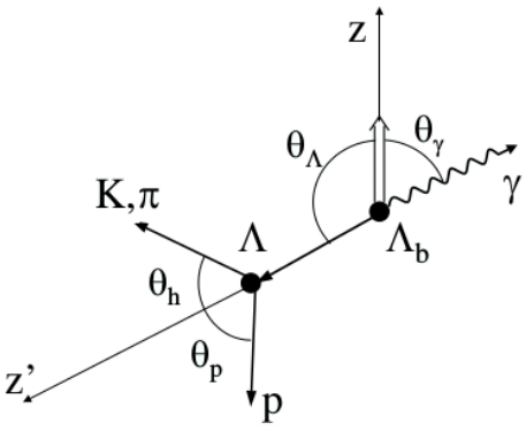


Figure 1.4: Sketch of the angle definition for the $\Lambda_b^0 \to \Lambda\gamma$. The angles θ_Λ and θ_γ are defined in the Λ_b^0 rest frame while the angles θ_h and θ_p are defined in the Λ rest frame. (from [31])

$$\frac{d\Gamma}{d\cos\theta_p} \propto 1 - \alpha_\gamma \alpha_{p,1/2} \cos\theta_p \qquad , \tag{1.23}$$

where θ_γ is the angle between the photon momentum and the Λ_b^0 spin direction (Z) in the Λ_b^0 rest frame and θ_p is the angle between the proton and the Λ momenta in the rest frame of the Λ (see figure 1.4). $P_{\Lambda_b^0}$ is the initial Λ_b^0 polarization and $\alpha_{p,1/2} = 0.642 \pm 0.013$ [32] is the weak $\Lambda \to p\pi^-$ decay parameter. The photon polarization is defined as

$$\alpha_\gamma \equiv \frac{P(\gamma_L) - P(\gamma_R)}{P(\gamma_L) + P(\gamma_R)} \qquad , \tag{1.24}$$

where $P(\gamma_{L/R})$ is the fraction of left/right-handed polarized photons in a $b \to s$ decay.

However, the initial Λ_b^0 polarization has been measured by LHCb [33] and is as small as $P_{\Lambda_b^0} = 0.06 \pm 0.07 \pm 0.021$. Thus the sensitivity to the photon polarization in Eq. (1.22) is suppressed by this factor. Nevertheless, Eq. (1.23) can still be used. The main advantage of measuring the photon polarization in $\Lambda_b^0 \to \Lambda\gamma$ is that at leading order (LO) the photon polarization is not dependent on form factors and can thus be computed with high accuracy in the SM, providing a clean observable. However, next-to-leading-order (NLO) corrections can affect the photon polarization. These corrections are not well known, but have been estimated to potentially reach a level of $\mathcal{O}(1 - 30\%)$ and depend on the values of the Wilson coefficients, in particular $\mathcal{C}_8$ [34]. Thus some progress is needed in the computation of NLO corrections for the measurement of the photon polarization in

$\Lambda_b^0 \to \Lambda\gamma$ decays to be sensitive to NP. In addition, the study of $\Lambda_b^0 \to \Lambda\gamma$ at LHCb (the only experiment in the world which produces large amounts of Λ_b^0 baryons to date) comes with some additional experimental challenges with respect to B decays. First, $b\bar{b}$ pairs produced in the proton-proton collisions at LHCb produce roughly twice as less Λ_b^0 baryons than B^0/B^+ mesons. In addition, the Λ baryon is a long lived particle which most of the time decays far from the interaction point. Thus typically only $\sim 20\%$ of reconstructed $\Lambda_b^0 \to \Lambda\gamma$ candidates at LHCb have long tracks[3]. To finish, most of the other Λ_b^0 decays also involve a Λ baryon in the decay chain, meaning potentially challenging sources of backgrounds for $\Lambda_b^0 \to \Lambda\gamma$ decays. Nevertheless, the $\Lambda_b^0 \to \Lambda\gamma$ decay has been observed for the first time at LHCb and its branching fraction has been measured [35], opening the way towards a measurement of the photon polarization in baryonic decays.

1.4 Measuring the photon polarization in the $B^0 \to K^{*0}e^+e^-$ decay mode

Instead of measuring the polarization of real photons in $b \to s\gamma$ processes, it is also possible to measure the polarization of (quasi-real) virtual photons in $b \to s\ell^+\ell^-$ processes at very low dilepton mass squared q^2, such as $B^0 \to K^{*0}\ell^+\ell^-$. Indeed, as sketched in figure 1.5, $B^0 \to K^{*0}\ell^+\ell^-$ decays are sensitive to different Wilson coefficients depending on the invariant mass of the final state leptons. In particular, at the very low end of the q^2 spectrum, the contribution of $\mathcal{C}_7^{(\prime)}$ is largely dominant and can be almost completely isolated from the contribution from the other Wilson coefficients. The very-low-q^2 region however is not accessible to muons and taus which have a q^2 threshold at $4m_\mu^2 \simeq 0.05\,\text{GeV}/c^2$ and $4m_\tau^2 \simeq 12.6\,\text{GeV}/c^2$, while for electrons the threshold $4m_e^2 \simeq 10^{-6}\,\text{GeV}/c^2$ is very close to the $q^2 = 0$ limit. As illustrated in figure 1.6, the predominance of the $\mathcal{C}_7^{(\prime)}$ Wilson coefficients at very-low-q^2 is explained by the fact that in this q^2 region the dominant Feynman diagrams of the $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}\gamma$ decays are exactly the same, up to the fact that the photon is virtual in the case of the $B^0 \to K^{*0}e^+e^-$ decay. The closer to $q^2 \to 0$ the dielectron mass is, the *less virtual* the photon is.

It is therefore possible to measure the photon polarization in $B^0 \to K^{*0}e^+e^-$ decays at very-low-q^2 by studying the angular distribution of the final state particles of the decay. In addition, due to the photon pole, one benefits from a sizable enhancement of the branching fraction in the very-low-q^2 region. Furthermore, some of the hadronic uncertainties are considerably reduced at very-low-q^2 allowing more precise predictions from the SM [37–39] and thus a greater sensitivity to

[3] A long track is defined by a track which leaves hits in all tracking subdetectors. More details are given in section 2.2.2

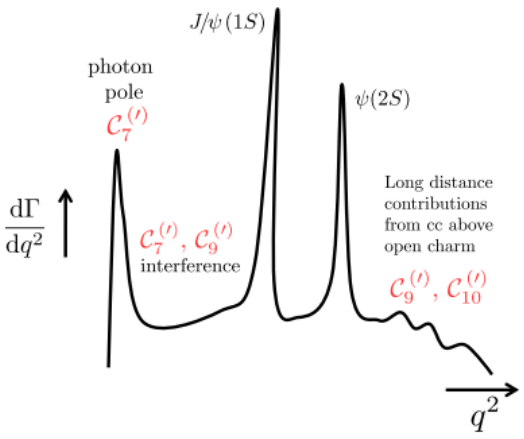


Figure 1.5: Artistic view of the $B^0 \to K^{*0}\ell^+\ell^-$ spectrum showing the dominant Wilson coefficients as a function of q^2 [36].

NP.

1.4.1 Differential decay width of the $B^0 \to K^{*0}e^+e^-$ decay

The differential decay width of the $B^0 \to K^{*0}e^+e^-$ decay can be described as a function of three angles $(\cos\theta_\ell, \cos\theta_K, \phi)$ and q^2. The angular basis is defined such that the $\overline{B}^0$ angular definition is a CP transformation of that of the B^0. The angle θ_l is defined as the angle between the direction of the e^+ (e^-) and the direction opposite to the B^0 $\left(\overline{B}^0\right)$ direction 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 the B^0 $\left(\overline{B}^0\right)$ direction in the K^{*0} $\left(\overline{K}^{*0}\right)$ rest frame. The angle ϕ is defined as the angle between the plane containing the two leptons and the plane containing the two hadrons of the final state in the B^0 $\left(\overline{B}^0\right)$ rest frame. A detailed description of the angular basis is given in Appendix A.1. Using the notations of [40] the

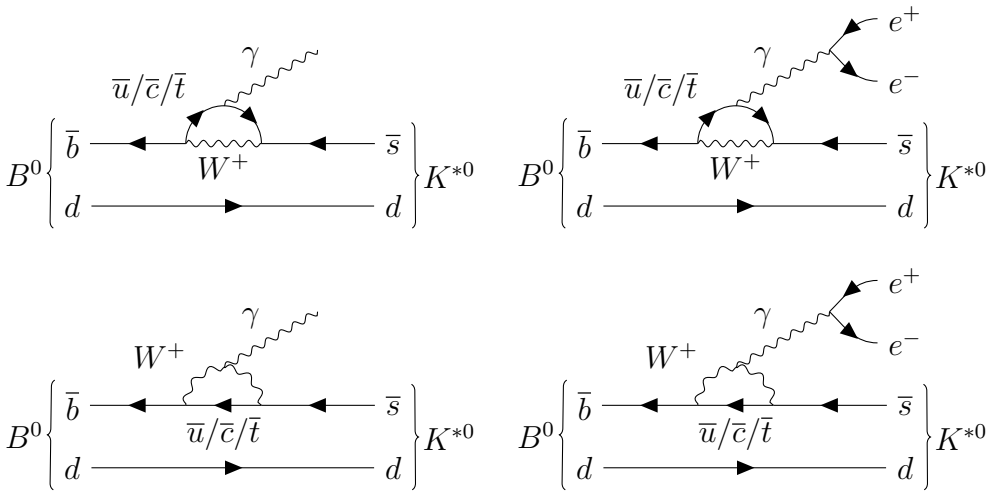


Figure 1.6: The two dominant Feynman diagrams in $B^0 \to K^{*0}\gamma$ decays (left) and $B^0 \to K^{*0}e^+e^-$ decays at very-low-q^2 (right).

differential decay width of the $\overline{B}^0 \to \overline{K}^{*0}e^+e^-$ decay reads [2]

$$\begin{aligned}
\frac{d^4\Gamma(\overline{B}^0 \to \overline{K}^{*0}e^+e^-)}{dq^2 d\cos\theta_\ell d\cos\theta_K d\phi} = \frac{9}{32\pi} \Big\{ & I_1^S(q^2)\sin^2\theta_K + I_1^C(q^2)\cos^2\theta_K \\
& + \left[I_2^S(q^2)\sin^2\theta_K + I_2^c(q^2)\cos^2\theta_K\right]\cos 2\theta_l \\
& + I_3(q^2)\sin^2\theta_K \sin^2\theta_l \cos 2\phi \\
& + I_4(q^2)\sin 2\theta_K \sin 2\theta_l \cos\phi \\
& + I_5(q^2)\sin 2\theta_K \sin\theta_l \cos\phi \\
& + I_6^S(q^2)\sin^2\theta_K \cos\theta_l \\
& + I_7(q^2)\sin 2\theta_K \sin\theta_l \sin\phi \\
& + I_8(q^2)\sin 2\theta_K \sin 2\theta_l \sin\phi \\
& + I_9(q^2)\sin^2\theta_K \sin^2\theta_l \sin 2\phi \Big\} \qquad ,
\end{aligned} \tag{1.25}$$

where the angular coefficients I_i are functions of q^2, two transverse $A_{\perp,\parallel}(q^2)^{L,R}$ and one longitudinal $A_0^{L,R}(q^2)$ amplitudes. These amplitudes are related to the spin state of the K^{*0} meson. In principle, the angular coefficients I_i are also function of two additional amplitudes, namely A_S related to the S-wave contribution to the $K^+\pi^-$ system and A_t related to the spin state of the *off-shell* virtual gauge boson decaying into the lepton pair and. However, the ratio of the S-wave fraction with respect to the fraction of longitudinal polarization of the K^{*0} is constant as a function of q^2 in the range $[0, 6]\,\mathrm{GeV}^2/c^4$ [38, 41]. Thus at very-low-q^2 the

fraction of S-wave is negligible. Moreover, in the limit of massless electrons, the terms proportional to A_t vanish. Thus, in the case of the $B^0 \to K^{*0} e^+ e^-$ decay at very-low-q^2 the angular coefficients I_i can be expressed in terms of q^2 and the six transverse and longitudinal amplitudes $A^{L,R}_{\perp,\parallel,0}$ only. These amplitudes are functions of q^2, Wilson coefficients and hadronic form factors. The complete expressions of the angular coefficients I_i are given in Appendix A.2 while the full expressions of the transversity amplitudes are given in Appendix A.3. Note that Eq. (1.25) also holds for the CP conjugate width $\bar{\Gamma}$ by replacing

$$I_{1-6} \to \bar{I}_{1-6} \quad , \qquad I_{7-9} \to -\bar{I}_{7-9} \qquad , \tag{1.26}$$

where $\bar{I}_i$ equals I_i with all weak phases conjugated. The fact that the I_{7-9} angular coefficients transform with a minus sign is bound to the definition of the angles. Indeed, with this convention, the CP transformation implies $\phi \to -\phi$.
Since at LHCb it is experimentally challenging to measure the flavor of the B meson, it is convenient to define the CP-averages S_i and CP-asymmetries A_i

$$S_i = \frac{I_i + \bar{I}_i}{\frac{d\Gamma}{dq^2} + \frac{d\bar{\Gamma}}{dq^2}} \quad , \qquad A_i = \frac{I_i - \bar{I}_i}{\frac{d\Gamma}{dq^2} + \frac{d\bar{\Gamma}}{dq^2}} \qquad . \tag{1.27}$$

Thus the CP-averaged differential decay width reads

$$\begin{aligned}
\left\langle \frac{d^4\Gamma}{dq^2 d\cos\theta_\ell d\cos\theta_K d\phi} \right\rangle_{CP} = \frac{9}{32\pi} \Big\{ & S_1^S \sin^2\theta_K + S_1^C \cos^2\theta_K \\
& + \left[S_2^S \sin^2\theta_K + S_2^C \cos^2\theta_K \right] \cos 2\theta_l \\
& + S_3 \sin^2\theta_K \sin^2\theta_l \cos 2\phi \\
& + S_4 \sin 2\theta_K \sin 2\theta_l \cos\phi \\
& + S_5 \sin 2\theta_K \sin\theta_l \cos\phi \\
& + S_6 \sin^2\theta_K \cos\theta_l \\
& + A_7 \sin 2\theta_K \sin\theta_l \sin\phi \\
& + A_8 \sin 2\theta_K \sin 2\theta_l \sin\phi \\
& + A_9 \sin^2\theta_K \sin^2\theta_l \sin 2\phi \Big\} \qquad .
\end{aligned} \tag{1.28}$$

The expression in Eq. (1.28) can be simplified without any loss to the sensitivity of the photon polarization [42] by folding the angle ϕ by

$$\tilde{\phi} \equiv \begin{cases} \phi \text{ if } \phi \geq 0 \\ \phi + \pi \text{ if } \phi < 0 \end{cases} \qquad . \tag{1.29}$$

The folded CP-averaged differential decay width is therefore given by

$$\begin{aligned}\left\langle \frac{d^4\Gamma}{dq^2 d\cos\theta_\ell d\cos\theta_K d\tilde{\phi}} \right\rangle_{CP} &= \frac{9}{32\pi}\left\{ S_1^S \sin^2\theta_K + S_1^C \cos^2\theta_K \right. \\ &+ \left[S_2^S \sin^2\theta_K + S_2^C \cos^2\theta_K \right] \cos 2\theta_l \\ &+ S_3 \sin^2\theta_K \sin^2\theta_l \cos 2\tilde{\phi} \\ &+ S_6 \sin^2\theta_K \cos\theta_l \\ &+ \left. A_9 \sin^2\theta_K \sin^2\theta_l \sin 2\tilde{\phi} \right\} \quad . \end{aligned} \tag{1.30}$$

Moreover, in the limit of massless lepton, two free parameters are dropped due to the relations [40]

$$S_1^S = 3S_2^S \quad , \qquad S_1^C = -S_2^C \qquad , \tag{1.31}$$

where the S_2 CP-averages are related to the fraction of longitudinal polarization of the K^{*0} meson $F_{\rm L}$ by

$$F_{\rm L} = -S_2^C \quad , \qquad 1 - F_{\rm L} = 4S_2^S \qquad . \tag{1.32}$$

The remaining CP-averages S_3 and S_6 and the CP-asymmetry A_9 can be expressed in terms of transverse observables as [38]

$$\begin{aligned} S_3 &= \frac{1}{2}(1 - F_{\rm L}) A_T^{(2)} \qquad , \\ S_6 &= (1 - F_{\rm L}) A_T^{Re} \qquad , \\ A_9 &= \frac{1}{2}(1 - F_{\rm L}) A_T^{Im} \qquad , \end{aligned} \tag{1.33}$$

where the transverse observables A_T^i are optimized observables for which the leading form-factor uncertainties cancel [38, 39] allowing for more precise theoretical predictions. An equivalent set of optimized observables P_i [43] can also be used. The relation between the A_T^i and the P_i observables is given by

$$\begin{aligned} A_T^{(2)} &= P_1 \qquad , \\ A_T^{Re} &= 2P_2 \qquad , \\ A_T^{Im} &= -2P_3^{CP} \qquad . \end{aligned} \tag{1.34}$$

The $B^0 \to K^{*0} e^+ e^-$ differential decay width can thus be expressed as a function of the three angles $\cos\theta_\ell$, $\cos\theta_K$, $\tilde{\phi}$ and the four angular observables $F_{\rm L}$, $A_T^{(2)}$, A_T^{Im} and A_T^{Re} as

$$\begin{aligned}\left\langle \frac{d^4\Gamma}{dq^2 d\cos\theta_\ell d\cos\theta_K d\tilde{\phi}} \right\rangle_{CP} &= \frac{9}{16\pi}\left\{ \frac{3}{4}(1-F_{\rm L})\sin^2\theta_K + F_{\rm L}\cos^2\theta_K \right. \\ &+ \left[\frac{1}{4}(1-F_{\rm L})\sin^2\theta_K - F_{\rm L}\cos^2\theta_K\right]\cos 2\theta_l \\ &+ \frac{1}{2}(1-F_{\rm L})A_T^{(2)}\sin^2\theta_K \sin^2\theta_l \cos 2\tilde{\phi} \\ &+ (1-F_{\rm L})A_T^{Re}\sin^2\theta_K\cos\theta_l \\ &+ \left.\frac{1}{2}(1-F_{\rm L})A_T^{Im}\sin^2\theta_K\sin^2\theta_l\sin 2\tilde{\phi}\right\} \quad ,\end{aligned} \tag{1.35}$$

where the four angular observables can be expressed as functions of the transversity amplitudes (see Appendix A.3) as [39]

$$\begin{aligned} F_{\rm L} &= \frac{|A_0|^2}{|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2} \quad , \\ A_T^{(2)} &= \frac{|A_\perp|^2 - |A_\parallel|^2}{|A_\perp|^2 + |A_\parallel|^2} \quad , \\ A_T^{Re} &= \frac{2\mathcal{R}e(A_\parallel^L A_\perp^{L*} + A_\parallel^R A_\perp^{R*})}{|A_\perp|^2 + |A_\parallel|^2} \quad , \\ A_T^{Im} &= \frac{2\mathcal{I}m(A_\parallel^L A_\perp^{L*} + A_\parallel^R A_\perp^{R*})}{|A_\perp|^2 + |A_\parallel|^2} \quad . \end{aligned} \tag{1.36}$$

A_T^{Re} is related to the forward-backward asymmetry by $A_T^{Re} = \frac{4}{3}A_{FB}/(1-F_{\rm L})$, while $A_T^{(2)}$ and A_T^{Im} are sensitive to the photon polarization and thus to $\mathcal{C}_7^{(\prime)}$ at very-low-q^2. Indeed in the limit $q^2 \to 0$ they can be expressed as [39]

$$A_T^{(2)}(q^2 \to 0) = \frac{2\mathcal{R}e\left(\mathcal{C}_7\mathcal{C}_7^{\prime *}\right)}{|\mathcal{C}_7|^2 + |\mathcal{C}_7^\prime|^2} \quad , \tag{1.37}$$

$$A_T^{Im}(q^2 \to 0) = \frac{2\mathcal{I}m\left(\mathcal{C}_7\mathcal{C}_7^{\prime *}\right)}{|\mathcal{C}_7|^2 + |\mathcal{C}_7^\prime|^2} \quad . \tag{1.38}$$

The measurement of the photon polarization in $b \to s\gamma$ processes is the main work presented in this thesis. The angular analysis of the $B^0 \to K^{*0}(\to K^+\pi^-)e^+e^-$ decay is performed in the q^2 bin $[0.0001, 0.25]\,{\rm GeV}/c^2$ using the full Run 1 and Run

2 data collected at the LHCb experiment representing an integrated luminosity of $\sim 9\,\mathrm{fb}^{-1}$.

1.4.2 Choice of the q^2 range

Since the main motivation for this analysis is to measure the photon polarization, the q^2 range is chosen as low as possible in order to isolate the photon pole. The dielectron mass range is chosen as $[10, 500]\,\mathrm{MeV}/c^2$ corresponding to $q^2 \in [0.0001, 0.25]\,\mathrm{GeV}^2/c^4$. The choice of the lower limit is mainly driven by the resolution on the ϕ angle and the converted photons background, while the upper boundary is a trade-off between statistics and sensitivity to $\mathcal{C}_7^{(\prime)}$ plus the avoidance of unwanted resonances.

On the usage of the B^0 and vertex constrained dielectron mass

In order to be closer to the true dielectron mass m_{ee}^{true}, the reconstructed dielectron mass is computed with several constraints when determining the range of the analysis. The four vectors of the electrons are recomputed via a kinematic fit with the constraints that the four final state particles tracks point to the same vertex and that the B^0 candidate points to the primary vertex of the proton-proton interaction. An additional constraint is added to the fit, namely that the reconstructed $m(K^+\pi^-e^+e^-)$ is equal to the nominal B^0 mass. This constrained reconstructed dielectron mass is called $m_{ee}^{B^0}$. As shown in figure 1.7, since the constrained mass is closer to the true dielectron mass, one gets less leakage when cutting on $m_{ee}^{B^0}$ to define the q^2 bin of the analysis. This allows to reduce the contamination from converted photons coming from $B^0 \to K^{*0}\gamma(\to e^+e^-)$ and eases the comparison with theoretical predictions since the experimental dielectron mass range is closer to the true one.

Lower boundary

In principle, one could go down to the threshold $m_{ee}^{\mathrm{true}} = 2m_e$. However, when the dielectron mass gets lower, the angle between the two electrons gets smaller and thus the plane containing the electron pair is determined with a worse precision. This leads to a worse measurement of ϕ, the angle between the plane containing the electron pair and the plane containing the hadrons. This mainly affects the measurement of $A_T^{(2)}$ and A_T^{Im}. Indeed, after integrating eq. (1.35) over $\cos\theta_\ell$ and $\cos\theta_K$, the CP averaged differential decay rate of $B^0 \to K^{*0}e^+e^-$ is given by

$$\left\langle \frac{d\Gamma}{dq^2 d\tilde{\phi}} \right\rangle = \frac{1}{\pi}\Big[1 + \frac{(1-F_L)A_T^{(2)} cos(2\tilde{\phi})}{2(\alpha+\beta F_L)} + \frac{(1-F_L)A_T^{\mathrm{Im}} sin(2\tilde{\phi})}{2(\alpha+\beta F_L)}\Big] \quad , \tag{1.39}$$

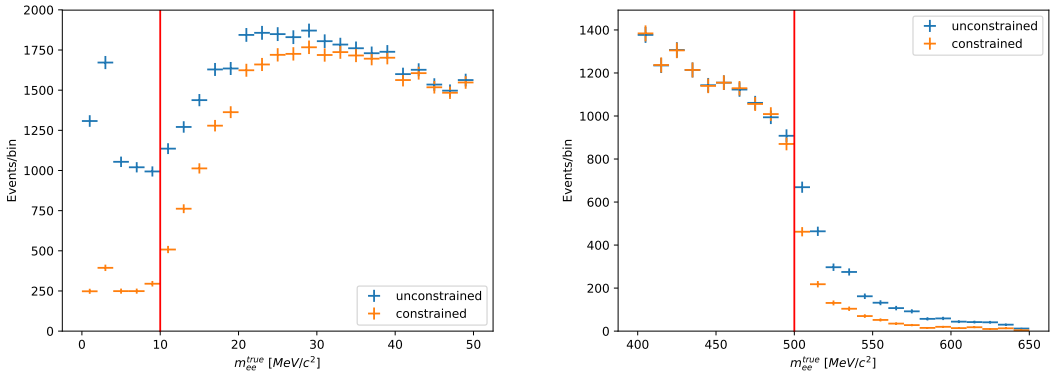


Figure 1.7: Leakage of events in simulated $B^0 \to K^{*0}e^+e^-$ from lower m_{ee}^{true} (left) and higher m_{ee}^{true} (right) when cutting between 10 and 500 MeV/c^2 m_{ee} (blue) or $m_{ee}^{B^0}$ (orange). The red lines indicate the cut imposed on the reconstructed $m_{ee}^{B^0}$ dielectron mass.

where the α and β parameters depend on the angular acceptance of the detector[4]. Since the photon polarization is mostly extracted from $A_T^{(2)}$, let us ignore the second term proportional to A_T^{Im} for this discussion. Moreover, the value of F_L is highly constrained by the fit on $\cos\theta_K$ of $< d\Gamma/dq^2 d\cos\theta_K >$. Hence, let us assume that at first order F_L is fixed in eq. (1.39). Fitting the data on $\tilde{\phi}$ to extract $A_T^{(2)}$ therefore effectively reduces to the extraction of the amplitude of a sinusoidal function. However, if $\tilde{\phi}$ is poorly measured, the distribution will be flattened and the extracted amplitude will be lower, leading to a bias towards a lower measured $A_T^{(2)}$.
Down to 10 MeV/c^2, the error on ϕ is below 500 mrad but reaches up to more than 1 rad below 10 MeV/c^2 (see figure 1.8). Thus a lower limit of 10 MeV/c^2 is chosen for the analysis range.

The bias to the measurement of $A_T^{(2)}$ is estimated with toys and is as low as $\sim$ 4%. The fitted value on data is therefore corrected accordingly while this correction is assigned as a systematic error (see section 4.5.5). Another issue at very low q^2 is the contamination from $B^0 \to K^{*0}\gamma$ where the photon converts to a dielectron pair in the material of the detector. Nevertheless, in the bin $[10, 500]$ MeV/c^2 and after applying the relevant veto cut, the contamination from converted photons is as low as $\sim 2\%$ (see section 3.3.2).

[4]In the (unrealistic) case where the acceptance is flat, $\alpha = 1$ and $\beta = 0$.

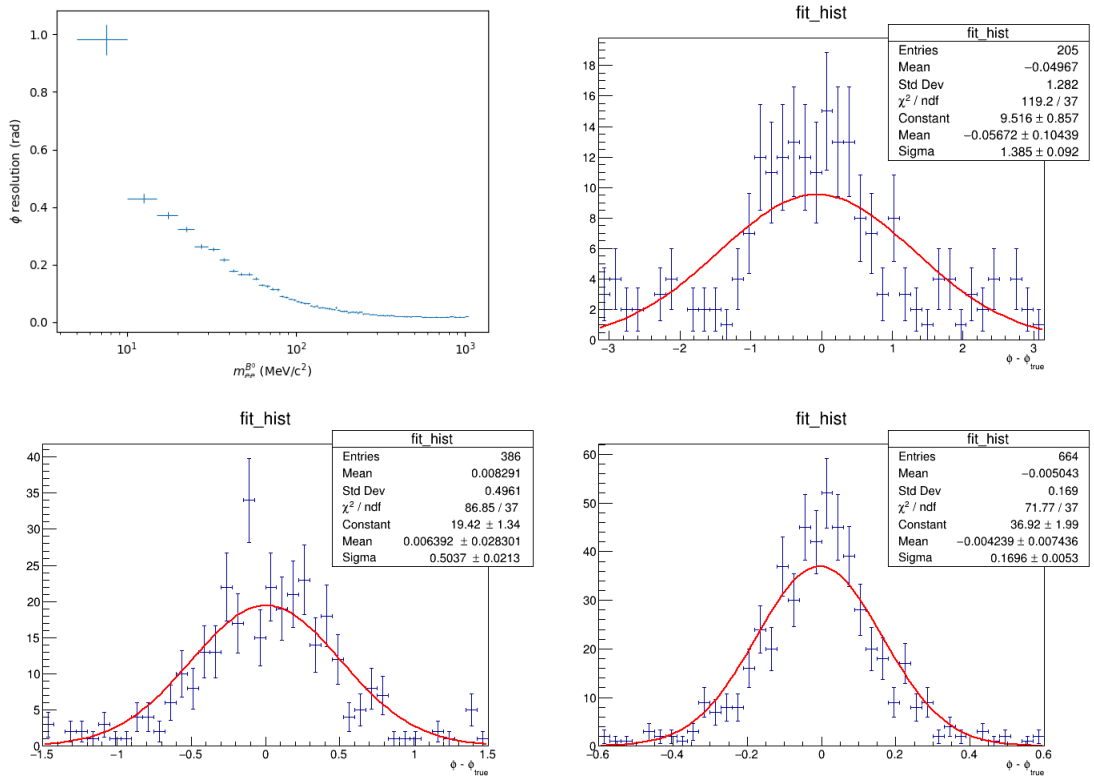


Figure 1.8: Resolution of the angle ϕ as a function of $m^{B^0}_{ee}$ (top left), extracted by a Gaussian fit to $\phi - \phi_{\text{true}}$ simulated $B^0 \to K^{*0}e^+e^-$ events. Three examples of such fits are shown, for $m^{B^0}_{ee} \in [5, 10]\,\text{MeV}/c^2$ (top right), $m^{B^0}_{ee} \in [10, 15]\,\text{MeV}/c^2$ (bottom left) and $m^{B^0}_{ee} \in [50, 55]\,\text{MeV}/c^2$ (bottom right)

.

Upper boundary

Extending the upper limit of the q^2 bin to higher q^2_{true} values decreases the sensitivity to $\mathcal{C}_7^{(\prime)}$ when measuring $A_T^{(2)}$. Indeed, in the limit of small lepton mass, $A_T^{(2)}$ is given by [38]

$$A_T^{(2)} = 2\frac{\mathcal{R}e\left(\mathcal{C}_7\mathcal{C}_7'^*\right)}{|\mathcal{C}_7|^2 + |\mathcal{C}_7'|^2}\left(1 + \alpha\frac{\mathcal{C}_9}{\mathcal{C}_7}\right)\frac{1}{\left(1 + 2\alpha\frac{\mathcal{C}_7\mathcal{C}_9}{|\mathcal{C}_7|^2+|\mathcal{C}_7'|^2} + \alpha^2\frac{\mathcal{C}_9^2+\mathcal{C}_{10}^2}{|\mathcal{C}_7|^2+|\mathcal{C}_7'|^2}\right)} \quad , \qquad (1.40)$$

where $\alpha = 0.0226 q^2_{\text{true}}$. Thus, for small α, *i.e.* small m^{true}_{ee}, $A_T^{(2)}$ is at first order only related to the ratio $\mathcal{C}_7'/\mathcal{C}_7$. But for higher m^{true}_{ee}, one has also to take into account the terms proportional to $\mathcal{C}_9$ and $\mathcal{C}_{10}$, thus losing the disentanglement of

$\mathcal{C}_7^{(\prime)}$ and $\mathcal{C}_{9,10}$. In the SM, eq. (1.40) reduces to

$$A_T^{(2)} = 2\frac{\mathcal{C}_7'}{\mathcal{C}_7}\frac{1 - 0.352 q_{\text{true}}^2}{1 - 0.704 q_{\text{true}}^2 + 0.216 q_{\text{true}}^4} \quad , \tag{1.41}$$

using $|\mathcal{C}_7|^2 + |\mathcal{C}_7'|^2 \simeq |\mathcal{C}_7|^2$, $\mathcal{C}_7 = -0.304$, $\mathcal{C}_9 = 4.73$ and $\mathcal{C}_{10} = -4.1$.
Figure 1.9 shows the evolution of the q^2-dependent term of eq. (1.41) as a function of the average true dielectron mass. This term can be interpreted as a correction to the sensitivity to $\mathcal{C}_7^{(\prime)}$ when measuring $A_T^{(2)}$ in the SM. In the range of the analysis $[10, 500]\,\text{MeV}/c^2$ (see below), the average dielectron mass taking into account the efficiency is around $200\,\text{MeV}/c^2$ giving rise to a correction to $A_T^{(2)}$ below 2%.

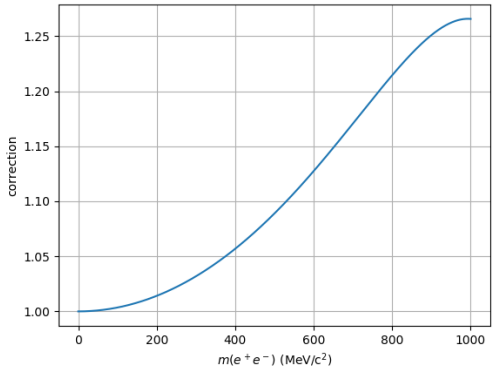


Figure 1.9: Correction to the sensitivity to $\mathcal{C}_7^{(\prime)}$ when measuring $A_T^{(2)}$ in the SM, as a function of the average m_{ee}^{true}.

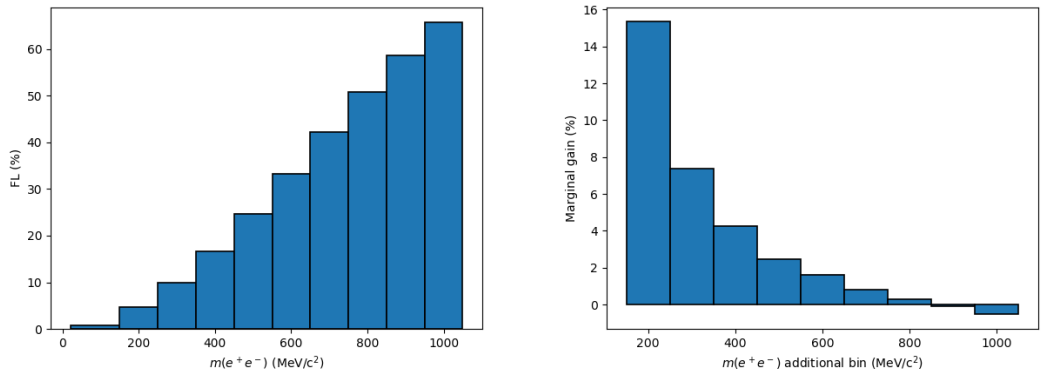
The statistical error on $A_T^{(2)}$ goes as $1/\sqrt{N}$, where N is the number of signal events. However, due to the fact that the first term in eq. (1.39) is proportional to $(1 - F_L)A_T^{(2)}$, events with higher values of F_L are less useful to determine $A_T^{(2)}$. This can be translated in terms of an effective number of events defined as

$$N_{\text{eff}} = N(1 - < F_L >)^2 \quad , \tag{1.42}$$

where, in a given bin, $< F_L >$ is the average F_L value and N (computed from eq. 4.12) is the number of signal events.
For a given bin, one can compute the reduction in the statistical error of $A_T^{(2)}$ if an additional bin is added to the first bin. This marginal error gain is given by

$$\sigma_{\text{marg}} = \frac{\sigma_{\text{bin1}} - \sigma_{\text{bin(1+2)}}}{\sigma_{\text{bin(1+2)}}} = \sqrt{\frac{N_{\text{eff}}\,(\text{bin(1+2)})}{N_{\text{eff}}(\text{bin1})}} - 1 \quad , \tag{1.43}$$

where $\sigma_{\text{bin}(1+2)}$, $N_{\text{eff}}\,(\text{bin}(1+2))$ are the statistical error and effective yield of the two bins merged together into a single one, while $\sigma_{\text{bin}1}$, $N_{\text{eff}}\,(\text{bin}1)$ are the corresponding values for the lower bin alone.



(a) (b)

Figure 1.10: Left: Mean SM F_{L} value in bins of dielectron mass. Right: Relative gain on the statistical error of $A_T^{(2)}$ by adding a bin to the sum of all previous bins.

Figure 1.10 shows that above 850 MeV/c^2, the marginal gain becomes slightly negative, meaning that adding these events slightly increases the statistical error on $A_T^{(2)}$. This is due to the fact that q^2 is integrated over the whole range of the analysis. Thus, adding more events of higher q^2 degrades the sensitivity to $A_T^{(2)}$ of all the events, including the ones with a low q^2, because the average F_{L} increases. Above 850 MeV/c^2, this negative effect is not compensated anymore by the additional statistics gained when increasing the upper limit. Besides, the resonant $B^0 \to K^{*0}\rho^0(\to e^+e^-)$ decay pollutes the $B^0 \to K^{*0}e^+e^-$ decay in the vicinity of the ρ^0 resonance at 770 MeV/c^2. To be safely away from the ρ^0 resonance and taking into account its large width (149 MeV/c^2), the upper boundary for the analysis is chosen to be 500 MeV/c^2.

Chapter 2

The LHCb experiment at the LHC

2.1 The Large Hadron Collider at CERN

The Large Hadron Collider (LHC) [44] is the world's largest and most powerful particle accelerator. It is located at CERN, near Geneva and has been in operation since September 2008. The main purpose of the LHC is to accelerate two proton beams at nearly the speed of light.
The LHC is actually the latest stage of an accelerator complex. Protons are first extracted from a hydrogen gas source, using an electric field to get rid of the electron bound to the proton. The protons are then accelerated to the energy of 50 MeV by LINAC2, a linear accelerator. In LINAC2, the protons pass through alternatively positively or negatively charged cylindrical conductors. Then the protons are injected into four superimposed synchrotron rings called the Proton Synchrotron Booster (PSB) which accelerates the protons to 1.4 GeV. The next stages are the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) which are circular accelerators of 628 m and 7 km circumferences respectively and which accelerate protons to 25 GeV and 240 GeV respectively. Finally, the proton beams are injected in the LHC in two beam pipes: one that circulates clockwise and the other one that circulates anticlockwise. The full chain of the CERN accelerator complex is sketched in figure 2.1.
The LHC is a 27 km ring located 100 m underground. The beams are guided into the ring by superconducting electromagnets which need to be cooled down at $-271.3°$C. This is achieved by a distribution system of liquid helium. To avoid any unwanted collision, an ultrahigh vacuum of 1.013×10^{-10} mbar (this is lower than the interstellar void) is maintained inside the beam pipes where the protons are circulating. The proton beams are accelerated up to 6.5 TeV using radio-frequency cavities. When the beams have reached the wanted energy, they are deviated such that the two beams cross each other in several collision points where

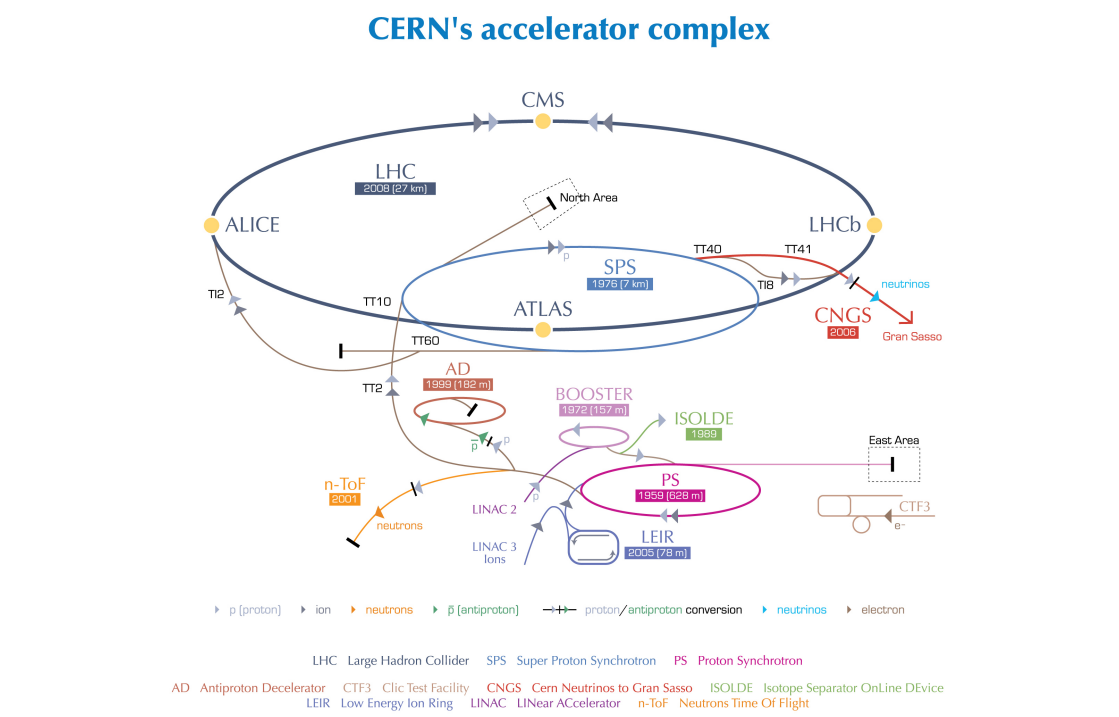


Figure 2.1: Sketch of the CERN accelerator complex. The date below each accelerator denotes the year when the accelerator has been launched for the first time. Besides producing the inputs for the LHC, several lower stage accelerators also provide beams for lower energy experiments (picture from CERN-DI-0812015).

the center-of-mass collision energy is equal to the sum of the energy of the two beams. The energy at which the proton beams are accelerated has been increased gradually over time: 3.5 TeV in 2011, 4 TeV in 2012 and 6.5 TeV in 2015, 2016, 2017 and 2018 (corresponding to 7 TeV, 8 TeV and 13 TeV center-of-mass energy at the collision points respectively). Since protons are composite objects made of quarks and gluons which mediate the strong interaction binding them together, only a fraction of the total kinetic energy of the proton is involved in the collision. At the LHC, gluon-gluon fusion and quark-antiquark annihilation are the main processes which produce pairs of $b\bar{b}$ quarks which in turn may hadronize to $B^0/\overline{B}^0$ mesons studied in this thesis.

At each collision point, a fraction of the protons collide and transform their kinetic energy into mass. Experiments are thus built around these collision points to detect and study the particles created by the collisions. There are four major experiments at the LHC: ATLAS [45], CMS [46], ALICE [47] and LHCb [48]. The

ATLAS and CMS experiments are general purpose detectors designed to study collisions producing high transverse momentum particles. They have a wide physics program with a focus on direct searches for NP particles as well as the detection and study of the properties of the Higgs boson. The ALICE experiment is a heavy-ion detector designed to study the physics of strongly interacting matter at extreme energy densities. The LHCb experiment is described in details in the section below.

2.1.1 Data samples

The analysis presented in this thesis uses the full Run 1 (R1) and Run 2 (R2) proton-proton collision data collected at LHCb during 2011 and 2012 (R1), and 2015, 2016, 2017 and 2018 (R2), corresponding to a total integrated luminosity of $9.1\,\text{fb}^{-1}$. The center-of-mass energy as well as the integrated luminosity corresponding to each year is given in table 2.1.

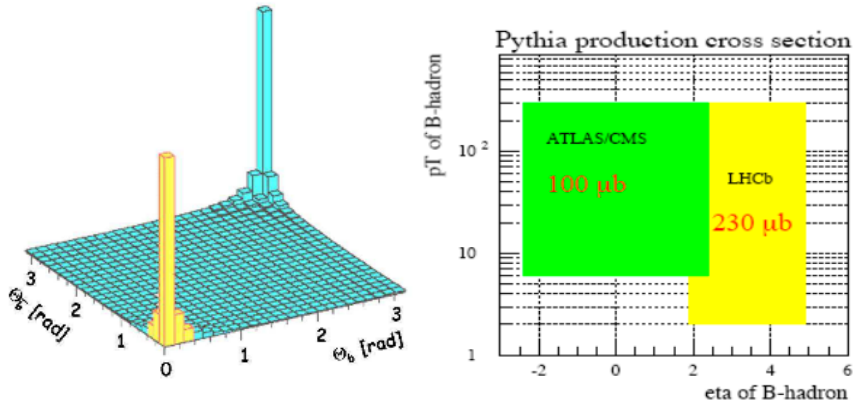
Year	$\mathcal{L}[\,\text{fb}^{-1}]$	$\sqrt{s}[\,\text{TeV}]$
2011	1.1	3.5
2012	2.1	4.0
2015	0.3	6.5
2016	1.7	6.5
2017	1.7	6.5
2018	2.2	6.5

Table 2.1: Summary table of the data samples used in the analysis, where $\mathcal{L}$ is the integrated recorded luminosity and $\sqrt{s}$ is the center-of-mass energy of the collision.

2.2 The LHCb detector during Run 1 and Run 2

The LHC is currently the world's largest B meson source. Out of the four major experiments at the LHC, LHCb is the only experiment which is specifically designed to study b (and c) hadrons. LHCb is a single-arm spectrometer with a narrow forward angular coverage from $\sim$ 10 mrad to $\sim$ 300 (250) mrad in the bending (non-bending) plane. The choice of the detector geometry is motivated by the fact that at the high energies provided by the LHC, the production of a $b\bar{b}$ pair occurs mostly in a small forward or backward cone, as shown in figure 2.2a. Indeed, the main $b\bar{b}$ production process at the LHC is gluon fusion in which the momenta of the partons[1] are very asymmetric. Therefore, the $b\bar{b}$ pair is boosted

[1]The partons are the constituents of hadrons, here quarks and gluons.



(a) (b)

Figure 2.2: Left: Simulation of the $b\bar{b}$ production cross section as a function of the polar angles of the b and $\bar{b}$ quarks with respect to the beam direction (from [11]). Right: Coverage of the LHCb and ATLAS/CMS experiments as a function of the transverse momentum p_T and the pseudo-rapidity η of simulated b hadrons. An estimation of the effective $b\bar{b}$ cross-section is given in red (from [49])

in the forward or backward direction. As a consequence, even if LHCb covers only $\sim 4\%$ of the 4π solid angle it covers $\sim 25\%$ of the $b\bar{b}$ production cross section (for comparison, the CMS detector covers $\sim 95\%$ of the solid angle which corresponds to $\sim 45\%$ of the $b\bar{b}$ production cross section). Thus, even though a large amount of B mesons are also produced at ATLAS and CMS, their detection efficiencies of B meson decays are much lower than the one of LHCb. This is mostly due to the fact that ATLAS and CMS trigger on events with higher transverse momentum and have a lower coverage at high pseudo-rapidity[2], as shown in figure 2.2b.

The second largest source of B mesons are B-factories at e^+e^- colliders working at the $\Upsilon(4S)$ resonance, like the Belle (II) [50, 51] and BaBar [52] experiments. Producing a very large amount of B mesons is crucial when studying rare decays like $B \to K^* e^+ e^-$ which typically have branching fractions of the order of $\sim 10^{-7}$. The total number of $B \to K^* e^+ e^-$ events produced are given by

$$N = \sigma_{B\bar{B}} \times L \times 2 \times \mathcal{B}(B \to K^* e^+ e^-) \times \varepsilon \quad , \tag{2.1}$$

where $\sigma_{B\bar{B}}$ is the $B\bar{B}$ cross section, L is the integrated luminosity and ε is the production efficiency (the latter is made explicit in each case below).
At Belle and BaBar, $\sigma_{B\bar{B}} \sim 1$ nb, while the total recorded luminosity over ten

[2]The pseudo-rapidity η is defined as $\eta \equiv -ln \tan \frac{\theta}{2}$, where θ is the polar angle with respect to the beam axis. Thus a track very close to the beam ($\theta \to 0$) has a pseudo-rapidity $\eta \to +\infty$, while a track perpendicular to the beam ($\theta \to \pi/2$) has a pseudo-rapidity $\eta \to 0$.

years has been $\sim 1000\,\text{fb}^{-1}$. In addition, the production efficiency $\varepsilon \sim 1$ at Belle and BaBar. Indeed, the detectors cover the full solid angle, the branching fraction of $\Upsilon(4S)$ to $B\bar{B}$ is almost 100% and the precise knowledge of the center-of-mass energy of the collision as well as the clean environment of e^+e^- collisions allows to reconstruct B decays to all final states. The total amount of $B \to K^*e^+e^-$ produced over ten years is thus of $\mathcal{O}(100)$ per experiment. At LHCb the $b\bar{b}$ cross section at 13 TeV center-of-mass energy is about 0.5 mb. However, in proton-proton collisions, not only B mesons are produced but also B baryons[3], like the Λ_b^0. In addition, B decays decaying to neutral final state particles (*e.g.* $B^0 \to K^{*0}(K_S^0\pi^0)e^+e^-$ or $B^+ \to K^{*+}(K^+\pi^0)e^+e^-$) are very challenging to reconstruct at LHCb due to the busy environment of proton-proton collisions. Thus at first order, only $B^0 \to K^{*0}e^+e^-$ decays can be studied at LHCb among all possible $B \to K^*e^+e^-$ decays. Taking into account the limited angular coverage of LHCb, the effective $\sigma_{B\bar{B}}$ cross section is at the order $\sim 10\,\mu\text{b}$. During the full Run 1 and Run 2, LHCb recorded about $\sim 9\,\text{fb}^{-1}$ integrated luminosity (see figure 2.3) which thus corresponds to about $\sim$ 20k $B^0 \to K^{*0}e^+e^-$ events, *i.e.* two orders of magnitude more than B-factories[4]. However, due to an inelastic cross section typically higher by a factor 200 with respect to the $b\bar{b}$ cross section in a proton-proton collision, LHCb suffers from a much higher background level than B-factories. In addition, since electrons are elementary particles, the energy of the e^+e^- collision at B-factories is well known. At LHCb however, only the energy of the protons is known but the fraction carried by the colliding partons is unknown. In particular, this makes challenging the flavor tagging[5] and the reconstruction of decays with neutral particles (such as π^0 mesons or photons) or neutrinos in the final state at LHCb.

2.2.1 General layout

The layout of the LHCb detector is shown in figure 2.4. Due to the busy environment of proton-proton collisions, the 10^{34} $\text{cm}^{-2}\text{s}^{-1}$ luminosity of the LHC is intentionally reduced to 4×10^{32} $\text{cm}^{-2}\text{s}^{-1}$ by regularly adjusting a small displacement of the two proton beams in the LHCb collision point [53]. This is mainly

[3]Even if this is a drawback when studying B meson decays, the ability to produce large amounts of bottom and charmed baryons is a unique feature of LHCb and opens the horizon to a largely unexplored sector of flavor physics.

[4]These numbers are only orders of magnitudes. In addition, they do not take into account the selection efficiencies which can be very different depending on the experiment and the particular analysis which is considered.

[5]Flavor tagging refers to the identification of the flavor of neutral B meson, for instance $B^0/\bar{B}^0$ mesons.

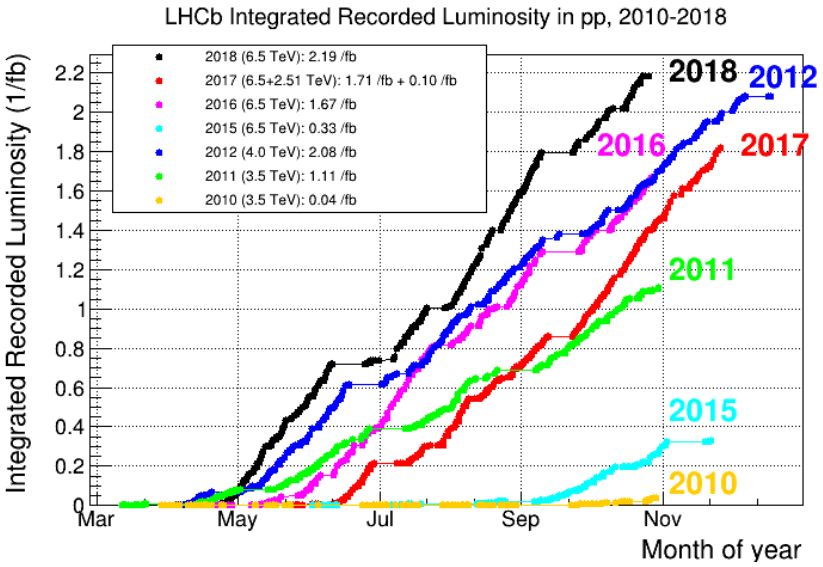


Figure 2.3: Integrated recorded luminosity of proton-proton collisions at LHCb during Run 1 and Run 2 (from the LHCb public website http://lhcb-public.web.cern.ch/lhcb-public/).

done to reduce the pile up[6] and thus be able to associate the primary vertex of the proton-proton collision to the displaced vertex of the studied b or c hadron. The proton-proton collision happens inside the VErtex LOcator ($VELO$), a high precision tracking detector used to measure the coordinates of vertices near the interaction point and to determine the impact parameter[7] (IP) of tracks. A magnet creates a magnetic field which bends charged particles trajectories. A series of trackers measure the track of charged particles enabling the measurement of their momentum. A calorimeter system made of an Electromagnetic CALorimeter (ECAL) and a Hadronic CALorimeter (HCAL) provides the identification of electrons, photons and hadrons as well as the measurement of their energies. The different types of hadrons, in particular kaons and pions, are identified with two Ring Imaging CHerenkov (RICH) detectors. At the downstream end of the detector, muon chambers provide muon identification. To finish, a two staged trigger system is implemented at LHCb. A fast hardware Level 0 (L0) trigger using information from the calorimeters and the muon chambers selects candidates with a high transverse energy or momentum. The L0 trigger is followed by a software

[6]The pile up is the superposition of several collisions at the same time in the detector. This can happen if the decay products of a second collision interact with the detector while the latter is still busy with the interactions of a first collision.

[7]The impact parameter is the transverse distance of closest approach between a particle's trajectory and a vertex, here the primary proton-proton interaction vertex.

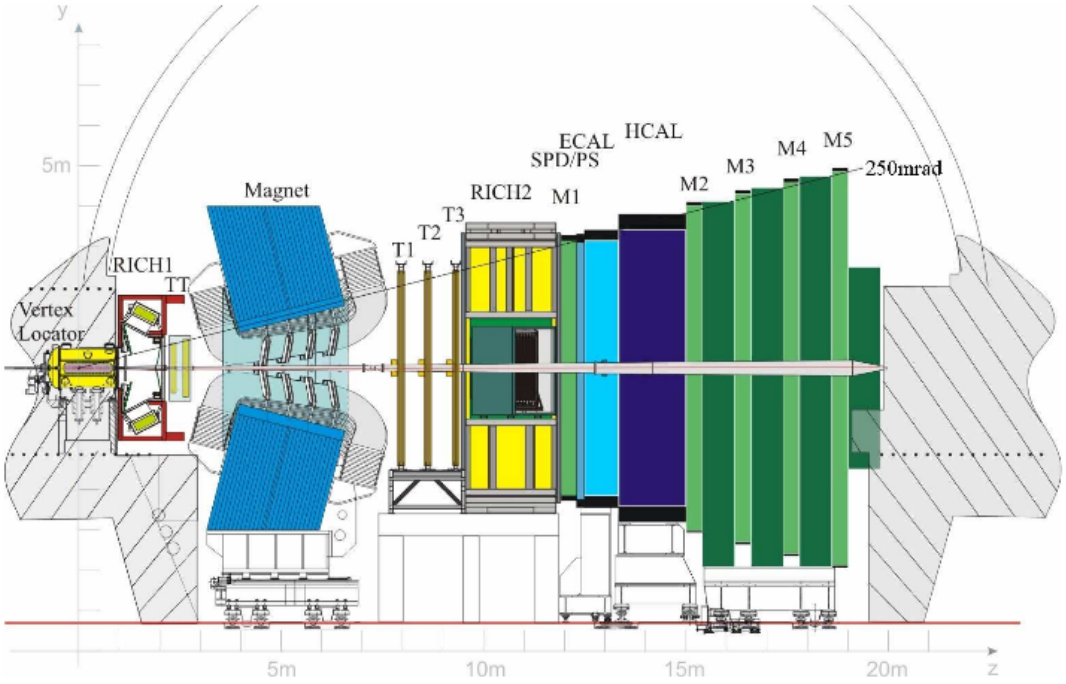


Figure 2.4: Layout of the LHCb detector [11]. A right-handed coordinate system is defined with z along the beam axis, y vertical and x horizontal.

High Level Trigger (HLT) to select interesting candidates out of the pre-selected one by the L0 trigger. Each subdetector is discussed in details below.

2.2.2 Tracking and vertex reconstruction

The main purpose of the tracking and vertexing system is to reconstruct vertices close to the proton-proton interaction point and to measure the momentum of charged particles by reconstructing their tracks.

Vertex reconstruction

Hadrons containing a b or a c quark (*e.g.* the B^0 meson studied in this thesis) are longed lived particles. Their daughter particles thus present the distinctive feature to originate from displaced (typically a few millimeters) secondary vertices with respect to the primary vertex (PV) of the proton-proton interaction. The $VELO$ thus provides the crucially needed precise measurements of track coordinates near the interaction point to be able to select b or c-hadrons.

The $VELO$ is made of a series of semi-disk of silicium disposed along the beam

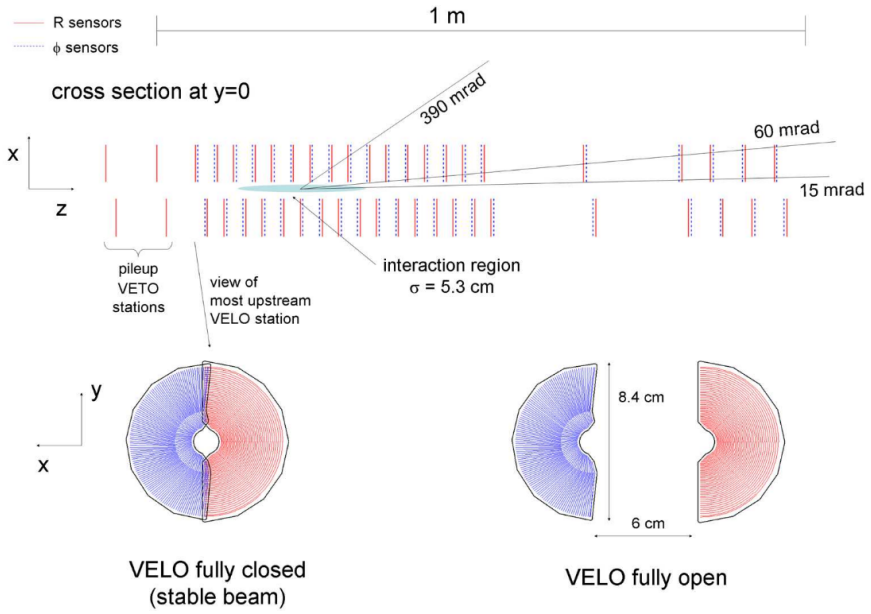


Figure 2.5: (Top) Cross-section of the $VELO$ showing the disposition of the sensors. (Bottom) Disposition of the $VELO$ semi-disk stations in closed (left) and open (right) position (from [48]).

axis. Each sensor gives a measurement of the radial distance r and the azimuth angle ϕ with respect to the beam axis. The disposition of the sensors, shown in figure 2.5, is chosen such that a track within the LHCb acceptance must cross at least three stations. In addition, there are two stations placed upstream of the interaction point whose main purpose is to contribute to the instantaneous luminosity measurement. Since the aperture of the $VELO$ is smaller than the beam radius during injection, the $VELO$ modules are retractable. As shown in figure 2.5, the $VELO$ modules are only closed in stable beam condition. To reduce the interaction before a particle crosses the sensors, the $VELO$ is kept in a vessel that maintains vacuum around the sensors. The $VELO$ is separated from the beam vacuum by an aluminum foil. The latter is called RF-foil, because it also protects the electronics of the $VELO$ from radio frequency (RF) interferences with the beam. The disposition of the RF-foil around the beam as well as the material budget in the $VELO$ as a function of the azimuth angle ϕ and the pseudo-rapidity η are shown in figure 2.7. The material budget has a crucial impact on photon conversions into e^+e^- pairs inside the $VELO$, which is an important background for $B^0 \to K^{*0}e^+e^-$ decays (see section 3.3.2).

The ability to distinguish secondary from primary vertices is given by the impact parameter (IP) and the PV resolutions. As shown in figure 2.6 the $VELO$

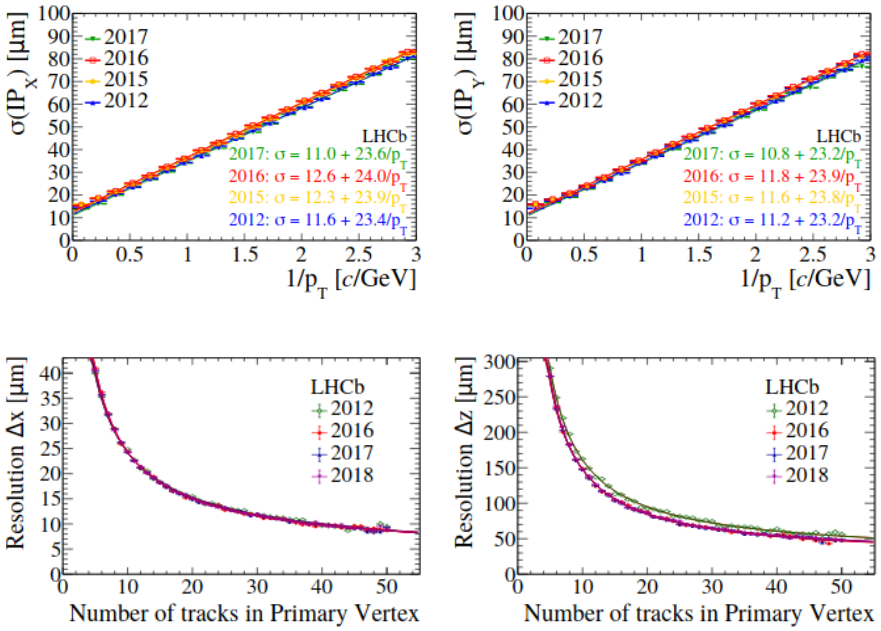


Figure 2.6: (Top) Impact parameter resolution along the x (left) and the y (right) axes, as a function of the inverse transverse momentum of the track for different years of data taking. (Bottom) PV resolution along the x (left) and the z axes, as a function of the number of tracks used to reconstruct the vertex for different years of data taking (from [54]).

provides excellent IP and PV resolutions. In rare decays like $B^0 \to K^{*0}e^+e^-$, one of the major background comes from all other particles surrounding the signal event which can be randomly combined to form a fake signal event. Being able to assign a specific primary and secondary vertex to each track greatly reduces the combinatorics of such random associations.

Track reconstruction

A dipole magnet operating at room temperature with an integrated magnetic field of about 4 Tm bends charged particles trajectories. The magnetic field along the beam axis is shown in figure 2.8. Charged particles leave hits in a series of trackers arranged upstream and downstream of the magnet. The four stations of the Tracker Turicensis (TT) are installed upstream of the magnet, while the three downstream stations (T1-3) are divided in two parts: the Inner Trackers (IT) arranged close to the beam axis and the Outer Tracker (OT) farther away from the beam axis (see figure 2.8). The TT and the IT share the same technology and are made of silicium microstrips while the OT are made of straw chambers.

Since the magnetic field is known to a high precision (the relative uncertainty $\delta B/B$ is at the level $\sim 4 \times 10^{-4}$) the momentum of charged tracks can be measured via the measurement of the bending radius of the tracks. The best tracking performances are obtained when combining information from the $VELO$ and the trackers. These tracks are called long tracks (see figure 2.9) and are the most used in physics analyses [8]. In particular, the analysis presented in this thesis only uses long tracks. The tracking system provides an excellent measurement of charged particles momentum with a relative momentum resolution $\delta p/p$ of $0.4\% - 1\%$ and a tracking efficiency of about 95% (see figure 2.10) for long tracks.

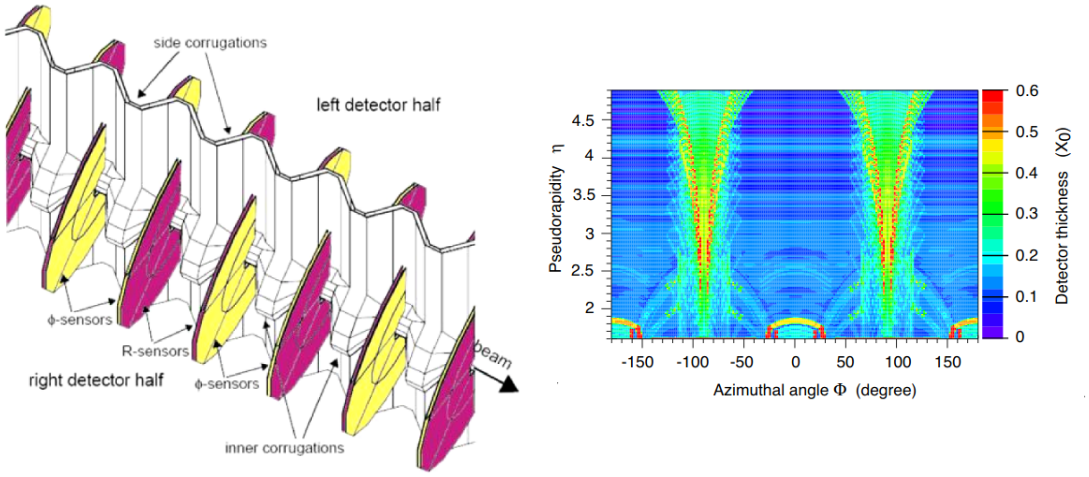


Figure 2.7: (left) Disposition of the RF-foil around the beam. (right) Average radiation length as a function of the azimuth angle ϕ and the pseudo-rapidity η (from [48]).

2.2.3 Particle identification and energy measurement

The tracking system having provided the measurement of the momentum of charged particles, two major tasks remain for the detector: the measurement of the energy and the identification of the particle specie of each candidate. For charged particles, these are actually two aspects of a single problem, due to the relativistic energy-momentum relation

$$E = \sqrt{p^2c^2 + m^2c^4} \quad , \tag{2.2}$$

where E is the energy, p the momentum and c the speed of light. Since the momentum is measured by the tracking system, knowing the specie of the candidate

[8] Some specific analyses also use other types of tracks. For instance, downstream tracks are used for the reconstruction of long lived particles such as K^0_S or Λ.

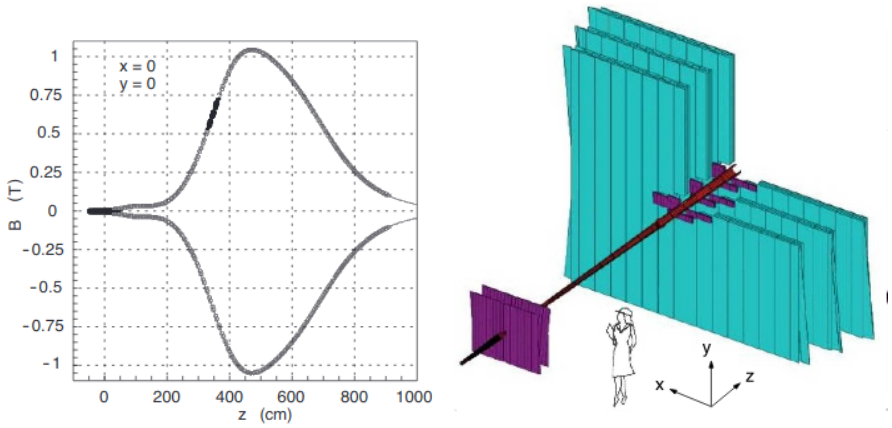


Figure 2.8: (left) Magnetic field along the beam axis. (right) Tracking system. The TT and IT are shown in purple while the OT is shown in cyan (from [48]).

(*i.e.* its mass) gives access to its energy. The core of the problem thus consists in identifying the specie of each track candidate. This is achieved by the combination of several subdetectors. A system of calorimeters (see section 2.2.3) allows to distinguish between photons, electrons and hadrons. In addition, two RICHs detectors (see section 2.2.3) considerably improve the particle identification (PID) of the LHCb detector and in particular allow to separate kaons from pions and from protons. To finish, muons chambers (see section 2.2.3) provide muon identification. Two particular types of particles require a special treatment, namely photons and electrons (see section 2.2.3). Indeed, since photons are neutral, they do not leave hits in the trackers and thus do not have associated tracks (*i.e.* their momentum is not measured). Therefore, the energy of photon candidates (which gives at the same time access to their momentum since Eq. (2.2) for a massless particle is just $E = pc$) is measured by the ECAL. Moreover, electrons require a special care due to Bremsstrahlung photon emission.

The calorimeter system

The calorimeter system [48] performs two major tasks. It identifies electrons, photons and hadrons, and measures their energies. It also plays an essential role in the L0 trigger by selecting candidates with a high transverse momentum (see section 2.2.4). The calorimeter system is made of two blocks: the electromagnetic block and the hadronic block. The electromagnetic block is composed of the PreShower (PS) detector, the Scintillator Pad Detector (SPD) and the ECAL, while the hadronic block is made of the HCAL. The SPD and the PS are two planes of scintillator pads. They are placed upstream of the ECAL and are separated by

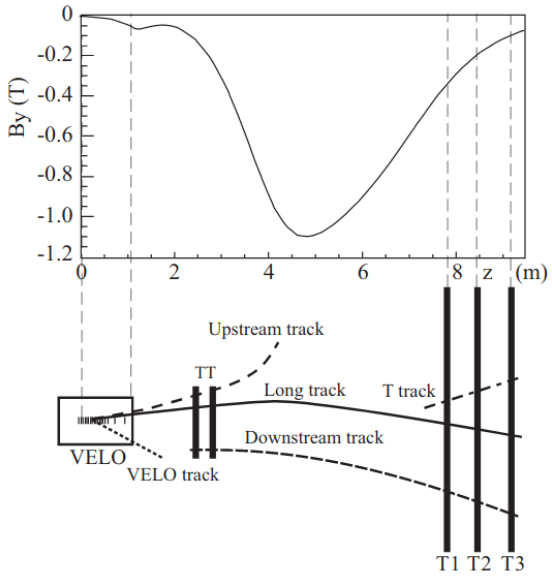


Figure 2.9: Sketch of the various track types at LHCb. The corresponding main component of the magnetic field (B_y) is also shown on the top of each region of the detector (from [48]).

a 15 mm thick lead plate. The latter corresponds to 2.5 radiation lengths while corresponding to only ~ 0.06 hadronic interaction length. Thus electrons and photons typically start to initiate an electromagnetic shower in the lead plate while hadrons do not. Nevertheless, all charged tracks leave hits in the SPD. In addition, the PS and the SPD are followed by the ECAL, a shashlik calorimeter made of alternating scintillator and lead absorber plates, arranged perpendicular to the beam axis. The ECAL is about 25 radiation length long and thus absorbs the full electromagnetic shower produced by electrons and photons, while absorbing only a small fraction of the shower produced by hadrons.

The ECAL is followed by the HCAL which is made of iron and scintillator tiles, oriented along the beam axis. The HCAL has an important role in the L0 trigger by selecting highly energetic hadrons (see section 2.2.4), but it is also used for particle identification. Both the ECAL and the HCAL have a variable granularity, to adapt to the higher hit density closer to the beam (see figure 2.11). When particles interact with the scintillators, they produce light which is collected by optical fibers and carried to photomultipliers (PMT). The variable granularity is obtained by varying the number of optical fibers which deposit their light to the same PMT. The HCAL has a thickness of 5.6 hadronic interaction lengths. Thus, while hadrons only have a small deposit in the ECAL, most of their energy is deposited in the HCAL. An illustration of the separation power between electrons

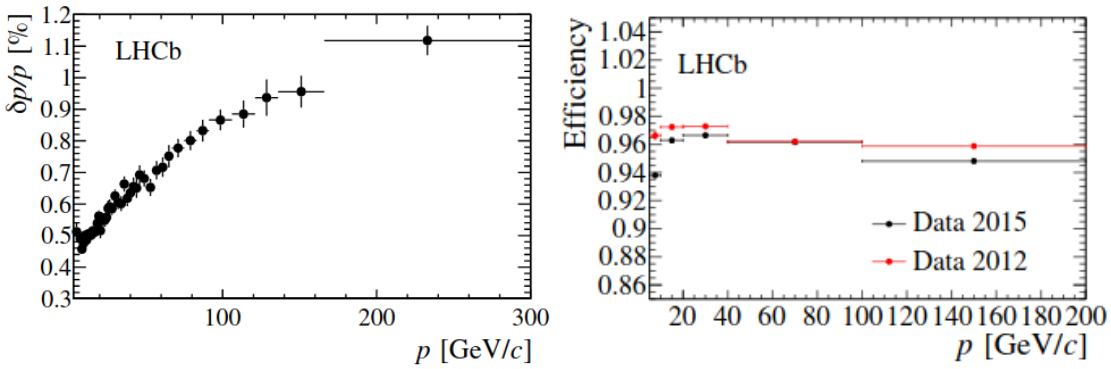


Figure 2.10: (left) Relative momentum resolution and (left) tracking efficiency as a function of momentum (from [54, 55]).

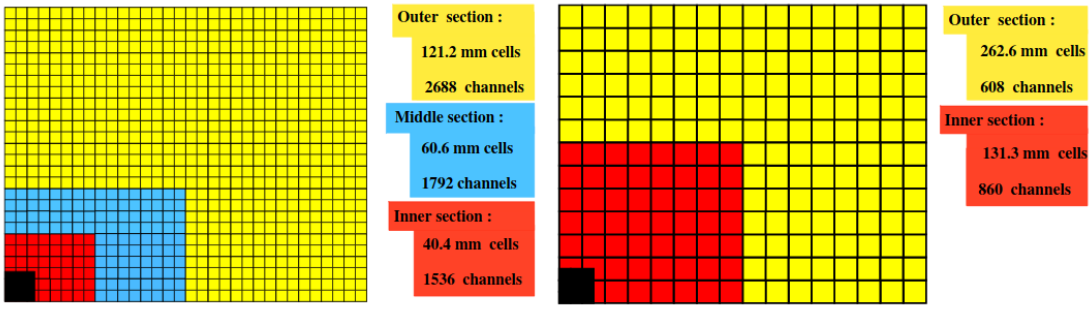


Figure 2.11: Lateral segmentation of one quarter of the detector front face of the SPD/PS and ECAL (left) and of the HCAL (right). In the left figure the cell dimensions are given for the ECAL (from [48]).

and hadrons is given in figure 2.13b where the ratio of the energy deposited in the ECAL (E^{ECAL}) divided by pc is shown. Since electrons deposit all their energy in the ECAL and their mass is negligible with respect to their momentum, this ratio is centered around 1. Hadrons, on the contrary, deposit only a small fraction of their energy in the ECAL and can have much higher masses, so their E^{ECAL}/pc is expected to be lower than 1. In order to distinguish photons from electrons, two main features are used: the fact that a photon is a neutral particle and thus has no track associated to it and the fact that photons have a low probability to leave hits in the SPD. Thus a photon (an electron) is identified as an energy cluster in the ECAL with no (a) track pointing to it and no (some) hits in the SPD. A dedicated algorithm based on the topology of the neutral clusters (*i.e.* trackless clusters) is used to separate photons from highly boosted π^0 mesons. The separation power of the calorimeter system to distinguish electrons from photons and from hadrons

is summarized in figure 2.13a.

Ring Imaging Cherenkov detectors

Two RICHs detectors [48] (RICH1 and RICH2) allow to identify the different hadrons and further improve the PID of electrons and muons. When a charged particle travels faster than light in a given medium, it produces a cone of Cherenkov light with an angle θ which depends on the velocity as

$$\cos\theta = \frac{1}{n\beta} \quad , \tag{2.3}$$

where n is the refractive index of the medium and β is the ratio of the velocity of the particle with respect to the speed of light. Combined with the measurement of the momentum by the trackers, the measurement of the velocity of the particle gives access to its mass through

$$m = \frac{p}{c\beta\gamma} \quad , \tag{2.4}$$

where $\gamma = (1-\beta^2)^{-1/2}$ is the relativistic Lorentz factor. RICH1 is located upstream of the magnet and covers low momentum charged particles range $\sim 1-60\,\mathrm{GeV}/c$. It is composed of two radiators combined to a single optical system: a silica aerogel ($n = 1.03$) for the lowest momenta and C_4F_{10} gas ($n = 1.0014$) for middle range momenta. RICH2 is located downstream of the magnet and covers high momentum charged particles range from $\sim 15\,\mathrm{GeV}/c$ up to and beyond $100\,\mathrm{GeV}/c$ using a CF_4 ($n = 1.0005$) radiator. The layout of RICH1 and RICH2, as well as the separating power of the different particle species as a function of their momentum are shown in figure 2.12. Being able to distinguish pions from kaons is crucial for the angular analysis of the $B^0 \to K^{*0}e^+e^-$ decay since it allows to identify the flavor of the K^{*0} meson (and thus the B^0 meson), which is needed in the definition of the ϕ angle (see section 1.4). A good separation of protons to kaons and pions also provides useful vetoes against backgrounds from b baryons like $\Lambda_b^0 \to pK^-e^+e^-$ (see section 3.3.6).

Muon system

Muon identification is provided by five muon stations (M1-5). The M2-5 stations are placed at the downstream end of the detector and are separated by 80 cm thick iron absorber to select penetrating muons. The stations are equipped with Multi Wire Proportional Chambers (MWPC) segmented into four regions (see figure 2.14) to accommodate with the occupancy variations as one gets further away from the beam. The full LHCb detector accounts for ~ 20 interaction length, meaning

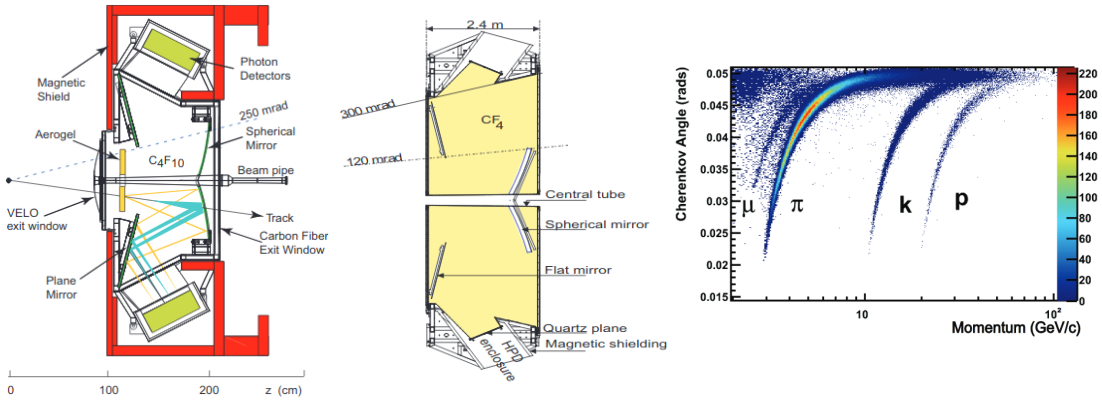


Figure 2.12: Schematic side view of the RICH1 detector (left) and top view of the RICH2 detector (middle), from [48]. Right: Measured Cherenkov angle as a function of the momentum of the track in the C_4F_{10} radiator in 2011 LHCb data (from [56]).

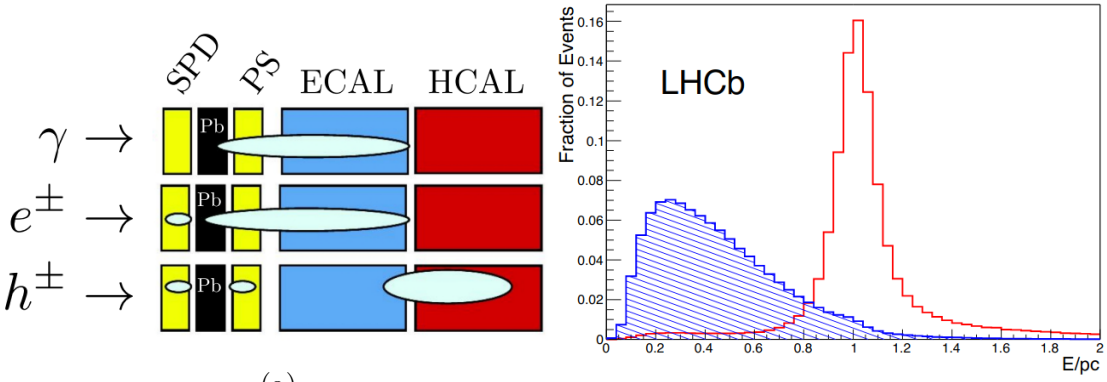


Figure 2.13: (left) Sketch of the typical energy deposit of photons, electrons and hadrons in the calorimeter system (from [36]). (right) Distribution of E^{ECAL}/pc for electrons (red) and hadrons (blue) in 2011 data (from [55]).

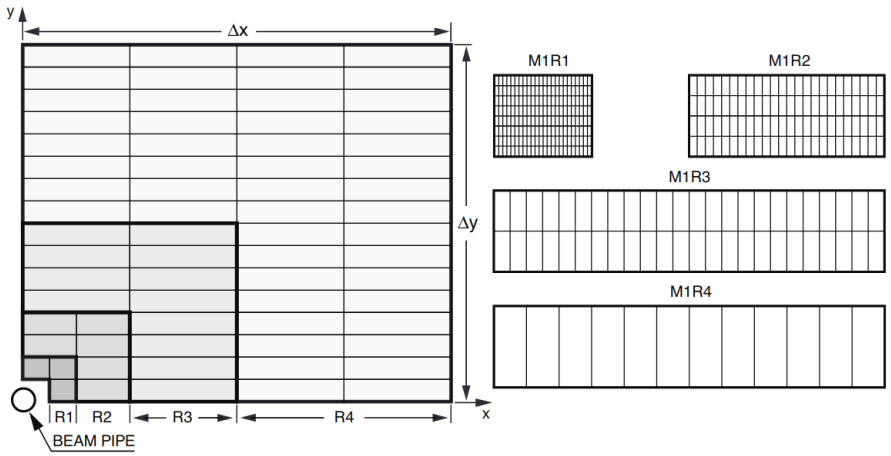


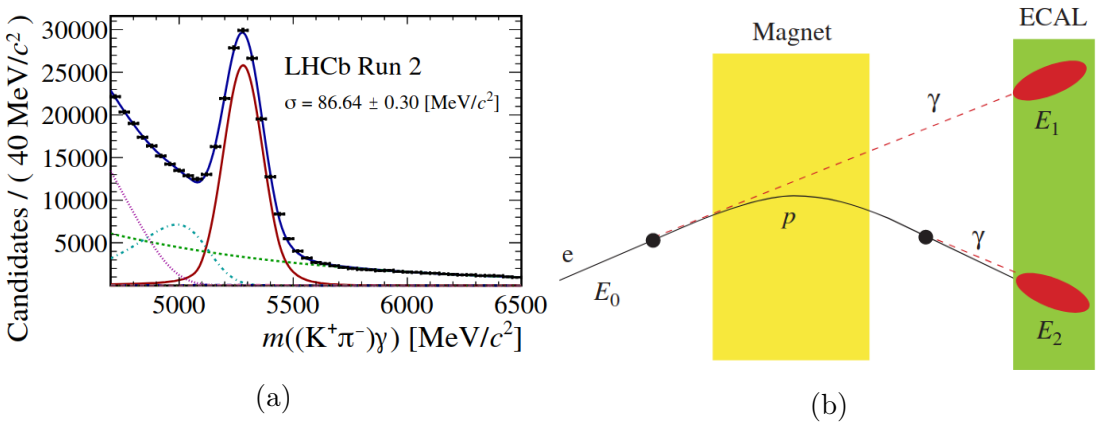
Figure 2.14: Layout of the muon stations. (left) Front view of a quarter of one muon station. Each square represents one chamber. (right) Segmentation of the four types of muon chambers installed in each of the four regions (from [48]).

that to cross all five muon chambers a muon needs a minimum momentum of about 6 GeV/c. The M1 station is placed in front of the calorimeters. It is also instrumented with MWPC except for the innermost part which is equipped with Gas Electron Multiplier (GEM) detectors. The main purpose of the M1 station is to select muons with high transverse momentum for the L0 trigger (see section 2.2.4).

Summary of the particle identification

Two types of algorithms combine the information from the various subdetectors to create PID variables which can be used in the offline selection.
Using the information from the tracking system and the RICHs, an algorithm computes a likelihood $\texttt{DLL}_X$ of being a given particle specie X. The difference between the likelihood of being a pion and that of being of type X is defined as $\texttt{PIDX} = \texttt{DLL}_\pi - \texttt{DLL}_X$.
An alternative approach uses the information from all subdetectors, including tracking, calorimeters, RICH detectors and muon stations to train a neural network[9] and build $\texttt{ProbNN\_}X$ variables, which can be interpreted as the probability to be of type X. Although $\texttt{ProbNN\_}X$ usually tend to perform slightly better than PIDX variables [57], they often offer some complementarity.

[9] A neural network is a type of machine learning algorithm which is trained to perform a specific task.



(a) (b)

Figure 2.15: (left) Invariant mass fit of $B^0 \to K^{*0}\gamma$ candidates in LHCb Run 2 data. The signal is shown in red while the various backgrounds are shown in green and purple. The sum of the signal and backgrounds is shown in blue (from [54]). (right) Sketch of the two main types of Bremstrahlung photon emissions upstream or downstream of the magnet (from [58]).

Photon and electron reconstruction

Since photons are neutral particles, they do not leave hits in the trackers. Thus one has to fully rely on the ECAL to measure their energy. The relative energy resolution of the ECAL is at the level of $\sigma_E/E = 10\%/\sqrt{E} \oplus 1\%$ (where the energy E is given in GeV) [48] resulting in a B mass resolution of about $90\,\mathrm{MeV}/c^2$ for the $B^0 \to K^{*0}\gamma$ decay[10], as shown in figure 2.15a.

Electrons also require a special care. Indeed, due to their small mass, electrons emit colinear Bremstrahlung photons when interacting with the material of the detector. As sketched in figure 2.15b, two main cases can occur: the Bremstrahlung photon can be emitted upstream of the magnet or downstream of the magnet. If the Bremstrahlung photon is emitted downstream of the magnet, there are no consequences. Indeed, the charged track is not bent anymore thus the energy loss has no impact on the measurement of the momentum. In addition, since the photon is emitted parallel to the electron trajectory, its energy will end up in the same cluster and the the energy measurement in the ECAL will be correct as well. However, if the Bremstrahlung photon is emitted upstream of the magnet (mainly by interacting with the $VELO$, RICH1, TT or the beam pipe), the measured momentum is biased towards lower values due to the momentum taken away by the emitted

[10]The resolution obtained for $B^0 \to K^{*0}\gamma$ candidates in this thesis is better (see section 3.3.2). Indeed, only events where the photon has converted into an e^+e^- pair are selected. Thus the tracking information as well as the calorimeter information can be used.

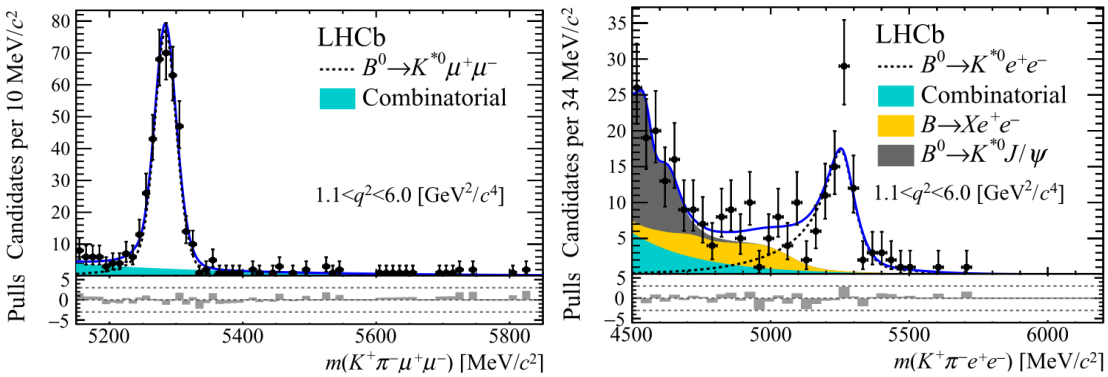


Figure 2.16: Invariant mass fits of $B^0 \to K^{*0}\mu^+\mu^-$ candidates (left) and $B^0 \to K^{*0}e^+e^-$ candidates (right) in the invariant dilepton mass squared region $[1.1, 6]\,\text{GeV}^2/c^4$ in Run 1 LHCb data (from [59]). Note the different scales in the x axis.

photon. In addition, since the photon is neutral, its trajectory is not bent by the magnet and its energy cluster ends up in a different place of the ECAL. In order to recover the lost energy, an algorithm searches for neutral clusters (clusters with no associated charged track) in the ECAL with an energy above 75 MeV in a region extrapolated from the track of the electron upstream of the magnet. If such a cluster is found, its energy is added to the electron candidate. If the same neutral cluster is associated to two electron candidates of different charges, its energy is added randomly to one of the candidates.
Nonetheless, the Bremstrahlung emission does not degrade the particle identification of electrons, since in all the cases (upstream or downstream emission, Bremstrahlung photon(s) recovered or not) the E/p ratio stays correct.

This results in a much worse resolution of electrons than muons at LHCb (an example is given in figure 2.16). Compared to the muon channel, the mass shape in the electron channel presents several special features. First, it has a long tail towards lower masses. This is due to unrecovered Bremstrahlung photons. This typically happens if the cluster created by the Bremstrahlung emission is lower than the 75 MeV threshold, if a random charged track from the event is pointing to the cluster formed by the Bremstrahlung photon(s) or in the rare (because of the lower material budget in this region) cases when the Bremstrahlung emission occurs inside the magnet. In addition, the mass shape of the electron channel also presents a small tail towards higher masses. This is due to wrongly added Bremstrahlung photons. This can occur for instance if a random photon or π^0 meson creates a neutral cluster in the region of the ECAL where the Bremstrahlung algorithm is searching. To finish, the bulk of the mass shape in the electron channel is wider than the one in the muon channel. This happens because, even if the

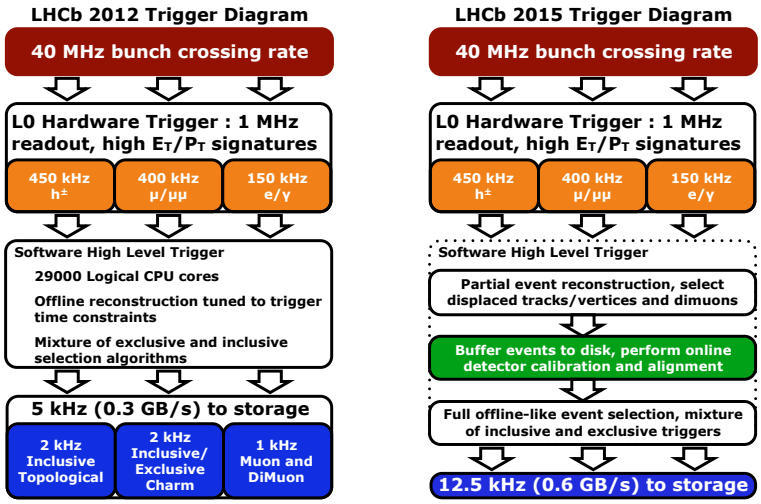


Figure 2.17: Sketch of the trigger chain during Run 1 (left) and Run 2 (right) at LHCb (from [60]).

Bremstrahlung emission is fully recovered, the energy resolution of the ECAL is worse by about a factor 10 with respect to the tracking resolution.
One of the major consequences of what is discussed above for the angular analysis of the $B^0 \to K^{*0} e^+ e^-$ decay channel presented in this thesis is the pollution coming from partially reconstructed backgrounds (see section 3.4.1) which mainly pollute the radiative tail at lower masses.

2.2.4 Trigger system

The proton-proton bunch crossing frequency at LHCb is at the level of 40 MHz. However, since most of the collisions do not contain interesting events, saving them all would be an enormous waste of disk space. The rate is thus reduced by a three steps trigger system which selects only interesting events to be saved on disk. A hardware Level 0 (L0) trigger uses basic information from the calorimeters, the muon chambers and the $VELO$ to reduce the rate from 40 MHz to 1 MHz. Then a two staged software High Level Trigger (HLT1 and HLT2) reduces the rate to a few kHz. The full trigger chain is summarized in figure 2.17.

Level 0 trigger

Due to the large mass of the B meson, its decay products have a high probability to have high transverse momentum or energy. Thus the L0 trigger is a fast hardware trigger aiming at reconstructing the highest transverse energy hadron, electron and

photon clusters in the calorimeters as well as the two highest transverse momentum muons in the muon chambers. Hadron, electron and photon candidates are built based on their signature in the calorimeter only (as explained in section 2.2.3). Then, the maximum 2×2 cell energy cluster is taken to build the `L0Hadron`, `L0Electron` and the `L0Photon` candidate. In addition, the transverse momentum (p_T) of muon candidates is reconstructed using only the muon chambers (the p_T resolution is $\sim 20\%$). The two candidates with the highest p_T form the `L0Muon` candidates. If the energy of the `L0Hadron`, `L0Electron` or `L0Photon` candidates or the momentum of the `L0Muon` candidates is above a predefined threshold, the corresponding trigger line is fired and the event is kept. The L0 trigger decision has to be made within 2 µs. Thus, an estimate of the number of charged tracks is done by the $VELO$ to reject events with too many tracks which would take more than 2 µs to be computed by the L0 trigger.

High Level Trigger

The High Level Trigger (HLT) is a software trigger divided in two parts, HLT1 and HLT2. The first stage HLT1 only uses part of the full event data. The main purpose of HLT1 is to select beauty or charmed decays. Thus, using only information from the trackers and the $VELO$, the HLT1 software performs a partial reconstruction of the tracks. Events are then selected by making some requirements on track quality, momentum and displaced vertices. The rate of events is thus reduced to a few tens of kHz.
The events which pass HLT1 then go to HLT2. This second step performs a full reconstruction using the information from all subdetectors. During Run 1, for computational reasons, a simplified reconstruction algorithm has been used in HLT2 called *online* reconstruction. After the data that had passed HLT2 had been saved to disk, a second reconstruction (called *offline* reconstruction) was performed with more complex algorithms. During Run 2, to avoid problems related to having a different *online* and *offline* reconstruction, the HLT software has been optimized and the computer farms have been upgraded such that the *offline* algorithms could also be applied *online*. To still allocate more computing time to HLT2, instead of directly injecting the output of HLT1 to HLT2, the HLT1 output was saved to a buffer disk. This allowed HLT2 to have a margin and to compute the data saved on the buffer disk when no data taking was performed (for instance during the fill of the LHC or during short technical shutdowns).
Once the event is fully reconstructed, HLT2 fires when an event matches the requirements of one of the HLT2 *lines*. A *line* is a set of lose requirements aiming at selecting a specific type of decay. There are inclusive lines which aim at selecting events with a certain topology and exclusive lines which are optimized to select a specific type of decay (for instance, a B meson decaying to two muons of opposite

charge plus anything else). In the analysis presented in this thesis, only topological HLT lines are used.
The events which have fired the L0 trigger in the first place and then one or several HLT lines are then saved to disk. However, since these trigger lines are very general, each physics analysis is only interested in some specific decays and thus only uses a small fraction of the data collected at LHCb. To prevent unnecessary disk usage (and costs) the datasets are further split in several streams called *stripping lines*. Each *stripping line* corresponds to a lose set of cuts to select events of a given type. For instance there is a *stripping line* dedicated to $b \to s\ell^+\ell^-$ decays, which is the one used for the angular analysis of the $B^0 \to K^{*0}e^+e^-$ decay presented in this analysis (see section 3.1).

2.2.5 The LHCb simulation

Monte Carlo (MC) simulation is used to study various properties of the collisions and their decay products as well as their response in the LHCb detector. The proton-proton collisions are generated by the PYTHIA [61] software with a specific LHCb configuration [62]. The decay of the particles is simulated by the EVTGEN [63] software while the final state radiation is generated with the PHOTOS [64] software. The full LHCb detector and its interaction with the simulated particles (including the hardware L0 trigger) is simulated with the GEANT4 software [65, 66]. Using the information from the simulated detector, the simulated events are reconstructed using the same reconstruction software (including the HLT software trigger) as in data. This is done to have simulated events as close as possible to the real data collected at LHCb.
Throughout this thesis, generator level MC (sometimes also referred to as TRUE variables) refers to particles as they have been generated by the LHCb simulation software before any interaction with the simulated detector. On the contrary, reconstructed level MC refers to the measured properties of the particles, *i.e.* the variables reconstructed out of the interaction with the simulated detector.

2.3 Upgrade I of the LHCb detector

Most physics measurements at LHCb, as the one presented in this thesis, are limited by the statistical uncertainty. Given that the latter scales with the square root of the number of events, another $30\,\text{fb}^{-1}$ of data would be needed to reduce the statistical uncertainty by a factor two with respect to the full Run 1 and Run 2 data collected so far by LHCb. With the instantaneous luminosity of Run 2 of $4 \times 10^{32}\ \text{cm}^{-2}\text{s}^{-1}$ this would require another fifteen years of data taking. Thus, the LHCb detector is undertaking an important upgrade during the LHC long

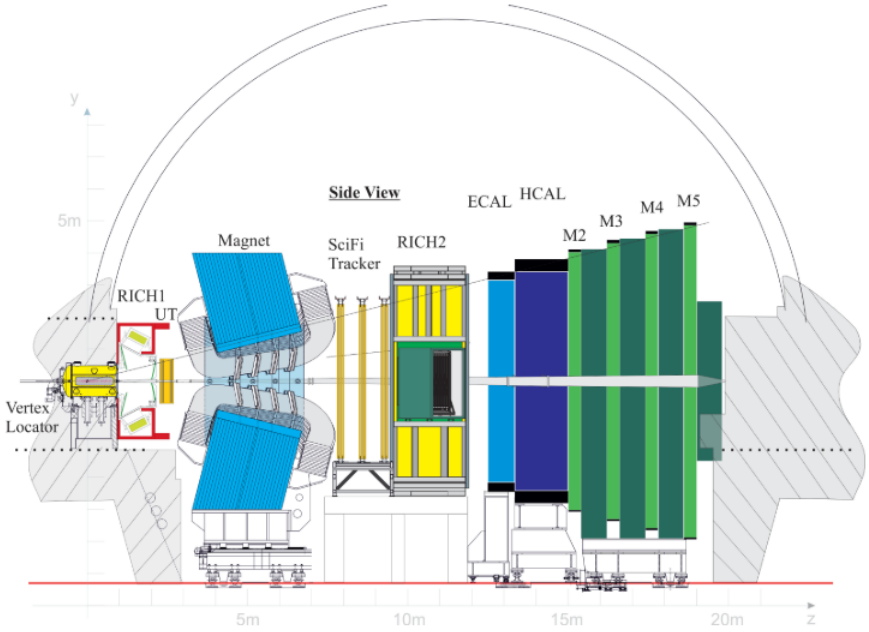


Figure 2.18: The Upgrade I LHCb detector (from [67])

shut down LS2 in 2019/2020 to be able to run the experiment at an instantaneous luminosity of 2×10^{33} cm^{-2}s^{-1} corresponding to a factor five increase.
Before the Upgrade I, the rate of data taking has been intentionally reduced through two means. First the nominal luminosity of the LHC has been reduced by about two orders of magnitude to 4×10^{32} cm^{-2}s^{-1} (see section 2.2). Moreover, the maximum readout of the electronics of the full detector before the Upgrade I was 1 MHz, while the LHC bunch crossing rate is as high as 40 MHz. To reduce the rate by a factor forty, only basic information from the calorimeters and the muon chambers have been recorded at 40 MHz. This information has then been used by a hardware Level 0 trigger (L0) to lower the rate to 1 MHz (see section 2.2.4). To fully exploit the luminosity delivered by the LHC, the L0 trigger will be removed after Upgrade I and replaced by a fully software trigger.

In order to reach the goal of running the LHCb experiment at an instantaneous luminosity of 2×10^{33} cm^{-2}s^{-1}, most subdetectors have to be upgraded to increase their granularity and radiation hardness. In addition, the front-end electronics and the data acquisition system has to be replaced to be able to read out and record events at 40 MHz [68]. The layout of the Upgrade I LHCb detector is given in figure 2.18. The various upgrades of all the subdetectors are summarized below.

- The upgraded Vertex detector ($VELO$) [69] will be made of 26 planar stations of hybrid silicon pixel detectors. Each pixel will have a square shape of

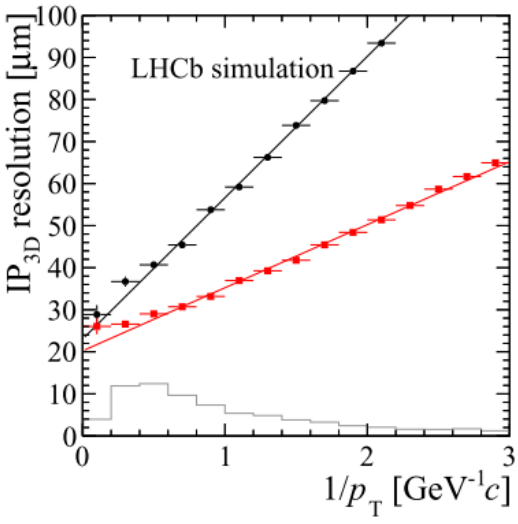


Figure 2.19: Estimation of the impact parameter resolution as a function of the inverse transverse momentum of a given track. The resolution of the original $VELO$ is shown in red while the resolution of the upgraded $VELO$ is shown in black. The inverse transverse momentum distribution of particles produced in b decays is shown in light gray (from [71]).

55×55 μm. In total there will be 41 million pixels read out by a new custom electronics [70]. Moreover, the thickness of the RF foil will be reduced from 300 μm to $200 - 250$ μm. The combination of a reduced material budget and smaller extrapolation distance will significantly enhance the resolution of the impact parameter as shown in figure 2.19.

- The Upstream Tracker (UT) [67] will replace the Tracker Turicensis (TT). It will consist in four layers of silicon microstrip detectors. The granularity is divided in four regions: in the outer region, 10 cm long strips with a pitch of 190 μm are used, in the middle region the pitch is reduced to 95 μm and 5 cm long strips are used in the inner region with the same pitch as in the middle region. To improve the acceptance of the detector at small polar angles, the innermost sensors present a quarter circle cutout. A new readout electronics, the SALT [72], which is needed to cope with the 40 MHz readout, is also developed. The number of fake tracks (ghosts) reconstructed per events with or without the UT is shown in figure 2.20.

- The SciFi downstream tracker is going to replace the original Inner Tracker (IT) and Outer Tracker (OT). The SciFi is made of three stations of scintillating fibers with four planar detection layers each. The 250 μm diameter

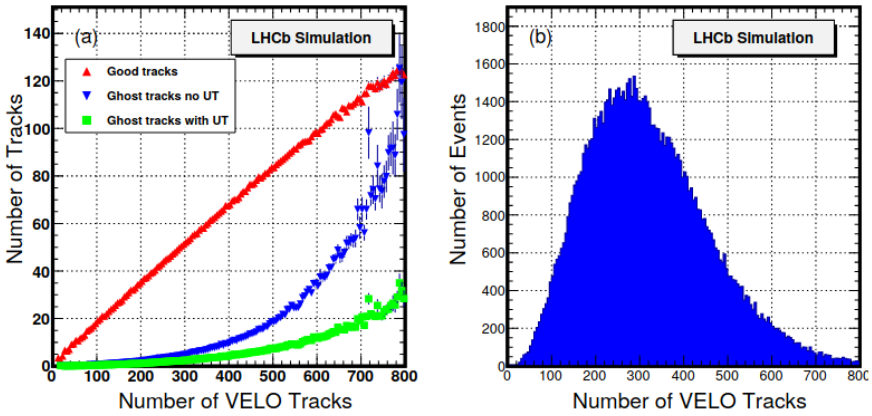


Figure 2.20: (left) Estimation of the number of fake tracks (ghosts) per event as a function of the number of $VELO$ tracks with or without UT. (right) Distribution of the number of $VELO$ tracks (from [67]). In both plots, inclusive b-hadron decays at 14 TeV center-of-mass energy have been generated (from [67]).

thick and 2.5 m long fibers will be readout by silicon photo-multipliers (operated at -40 ° C to reduce the radiation damage) at the top and bottom of each layer. As shown in figure 2.21 the track reconstruction efficiency of the upgraded downstream tracker is slightly lower than the original downstream tracker for a low number of primary vertices (PV). This is due to the lower granularity in the inner region of the SciFi. However, the tracking efficiency improves significantly for higher number of PVs and, as shown in figure 2.21 the upgraded downstream tracker performs better over the full pseudo-rapidity range.

- The system of light collection of the two RICH detectors will be changed to Multi-anode Photo Multiplier Tubes (MaPMT) with a granularity of 2.9×2.9 mm^2 [73, 74]. The optical design of the mirror system of RICH1 will also be updated to adapt to the higher occupancy. A new front-end electronics is also developed to cope with the 40 MHz readout.

- Out of the five original muon stations (M1-5) only four will remain. Indeed, the M1 station, which has been essentially used for the L0 trigger, will be removed. In addition, the inner region of the M2 station will be equipped with Triple-GEM detectors [75] to handle the higher particle density. The readout electronics will also be changed to cope with the 40 MHz readout.

- The PS and SPD detectors of the original calorimeter system will be removed, since their main purpose was to provide inputs to the L0 trigger. In addition,

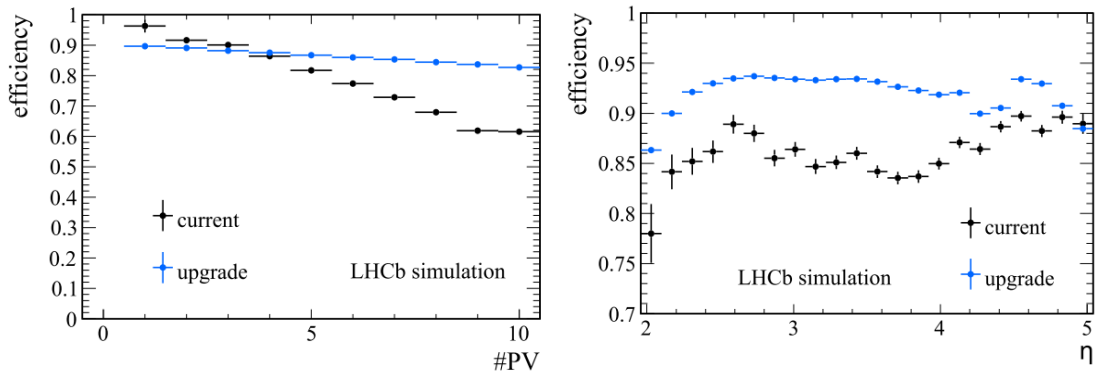


Figure 2.21: Estimation of the tracking efficiency as a function of the number of primary vertices (left) and as a function of the pseudo-rapidity (right), from [67].

the readout electronics of both the ECAL and the HCAL will be updated to cope with the 40 MHz readout.

- The L0 trigger will be removed and replaced by a fully software trigger [76]. The latter will only reduce the rate to 30 MHz instead of 1 MHz. The software trigger is decomposed into two parts: the Low Level Trigger (LLT) and the High Level Trigger (HLT). The software LLT selects events containing clusters with high transverse energy in the calorimeters or tracks with high transverse momentum in the muon detector. Thus, the software LLT is conceptually very similar to the hardware L0 trigger. The processing time is estimated at 20 ms per event and the output bandwidth is fixed to 20 kHz $\times$ 100 kBytes = 2 GBytes/s. Nevertheless, the calorimeter part of the L0 hardware trigger will be kept in the upgraded detector. Its main purposes will be to serve as a backup in case of a temporary failure of the software LLT and to provide higher level information to the software trigger in order to speed up the computing time of each event. The upgrade of the electronics of the hardware L0 trigger of the calorimeters has been performed in this thesis and is presented in details in section 2.3.2.

2.3.1 The Front-End board of the upgraded LHCb calorimeters

The data of the two calorimeters (ECAL and HCAL) is collected by front-end boards (FEB) [77]. Both the ECAL and the HCAL use the same FEBs. Each FEB is connected to 32 photomultiplier (PMT) outputs of the calorimeter corresponding to one module which is a rectangle of 4×8 cells. Figure 2.22 shows the rectangles

of cells associated to the FEBs for the ECAL and the HCAL.

One FEB provides the measurement of the 32 transverse energy (E_T) of the corresponding 32 calorimeter cells to the software trigger. In addition, each FEB also provides the Level 0 Low Level Trigger (L0LLT) information which consists in the maximum transverse energy cluster and its location as well as the total transverse energy and the multiplicity of the event (see details in section 2.3.2). As in the pre-upgrade L0 trigger (see section 2.2.4) the maximum transverse energy cluster aims at selecting events coming from a B meson while the multiplicity measurement may be used for global event cuts, to remove very busy events which would require too much computing time for the HLT farms. The multiplicity of the L0LLT is akin to the SPD hits measurement of the pre-upgrade L0 trigger. The information is stored as ADC counts. The ADC bandwidth of each FEB is $12 \times 32 = 384$ bits (12 bits to store the transverse energy per channel times 32 channels), plus 32 bits for the L0LLT information. Each FEB operates at 40 MHz in accordance with the LHC bunch crossing rate.

As before the upgrade, the FEBs will be gathered in 18 crates (14 for the ECAL and 4 for the HCAL) themselves gathered in racks and installed in platforms on top of the detector in the LHCb cavern. In each crate, a control board is located in the middle of the crate. The control board is, in particular, used to manage the clocks of the FEBs and to configure the FEBs. An illustration of the layout of a crate is given in figure 2.23.

As shown in figure 2.24, the digital processing in the FEB is done by radiation-hard Field-Programmable Gate Arrays (FPGA) which are organized in several blocks. The main ones are the front-end blocks (four identical blocks), the trigger block and the (de-)serializers block. The front-end block is responsible for the E_T measurement while the trigger block computes the L0LLT information. In addition, since the L0LLT needs inputs from the neighbor FEBs (see section 2.3.2), the (de-)serializers block is needed for the exchange of information between the FEBs. The transverse energy of each cell is computed by integrating the charge and transforming it into a 12 bit ADC digital word. However, even in the absence of signal, the ADC count will not be zero, due to electronic noise. A procedure is thus implemented to subtract the noise. There are two methods implemented (the choice of the method can be chosen via a software switch). The simplest method subtracts the smallest ADC count of the two previous measurements. Since the occupancy of a single channel is low, the probability that there had been three consecutive signal events in the same channel is negligible. The second method subtracts the lowest of the value of the previous noise subtraction (plus a small constant of typically a few ADC counts) and the previous measurement. This avoids large fluctuations in the (rare) cases where there had been several consecutive signal events in the same channel. To avoid ending up with negative

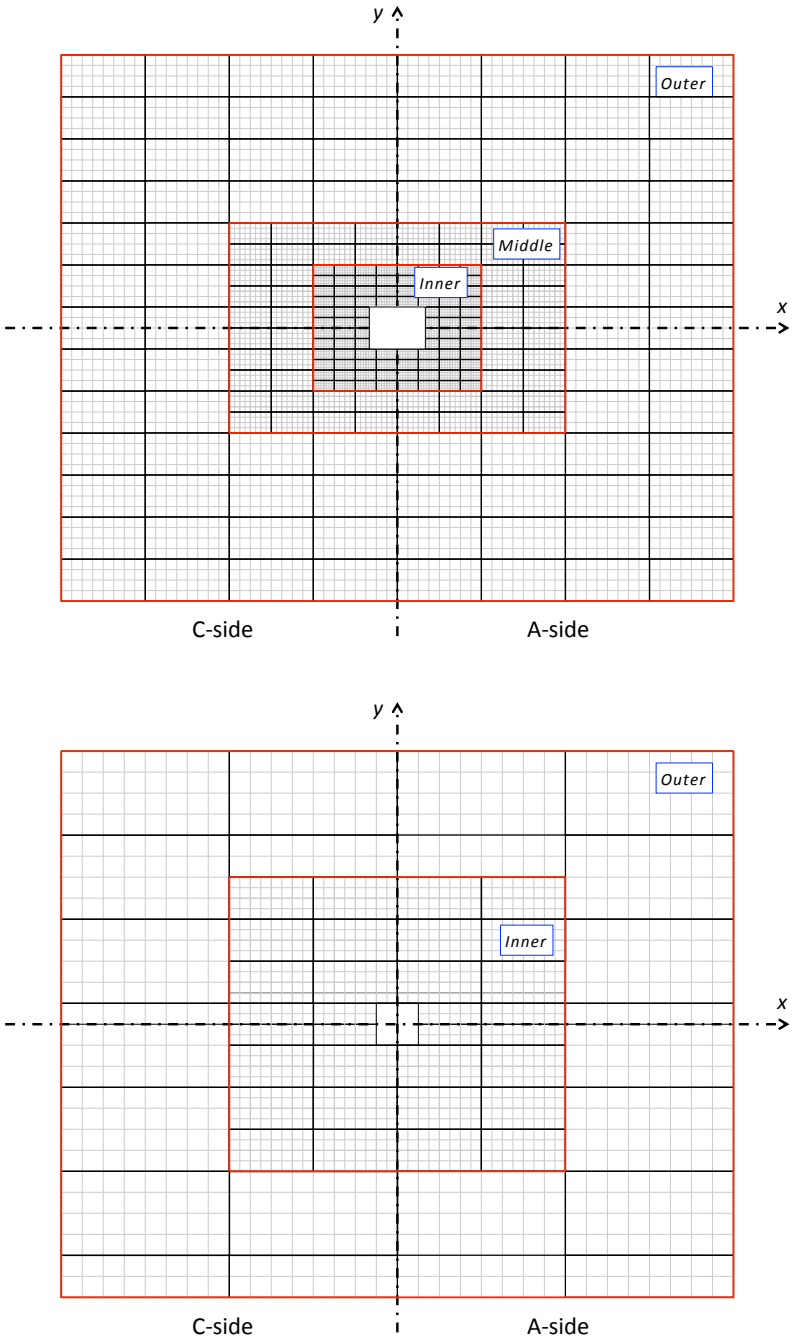


Figure 2.22: Sketch of the ECAL (top) and HCAL (bottom) cells associated to each FEB. The small squares represent cells while the black rectangles represent the 4×8 cells associated to each FEB.

Figure 2.23: Picture of a crate in a test bench at IJCLab. The control board is located in the middle of the crate, while four FEBs are plugged in various slots of the crate. Similar crates will be installed in the calorimeters, where all empty slots of the crate will be filled.

ADC count values (in the case the noise in the current event is smaller than the noise in the previous event) a pedestal is added to all ADC counts. That is, a value N_{ped} (typically around 256) is added to the ADC count, before the noise is subtracted. Thus, the noise-corrected ADC count for the transverse energy measurement is given by

$$N_{ADC} = N_{ADC}^{raw} + N_{ped} - N_{noise} \quad , \tag{2.5}$$

where N_{ADC}^{raw} is the raw 12 bit ADC count of the integrated charge and N_{noise} is the ADC count noise subtraction. The minimum value is saturated at zero, while the maximum value is saturated at 4095.

The FEBs show good performances. The average noise per channel is below 2 ADC counts, as shown in figure 2.25a. Moreover, a $\sim 2\%$ linearity is observed between the size of the pulse and the ADC count (see figure 2.25b). In addition, the cross-talk of one channel to another is below a few percent (see figure 2.25c). To finish, the time stability is better than 1% over ± 2 ns on the integration pulse.

2.3.2 The upgraded Low Level Trigger

The upgrade of the electronics of the hardware Low Level Trigger (L0LLT) is part of this thesis. The main purpose of the L0LLT is to build electron and hadron candidates based on calorimeter information only. The information computed by

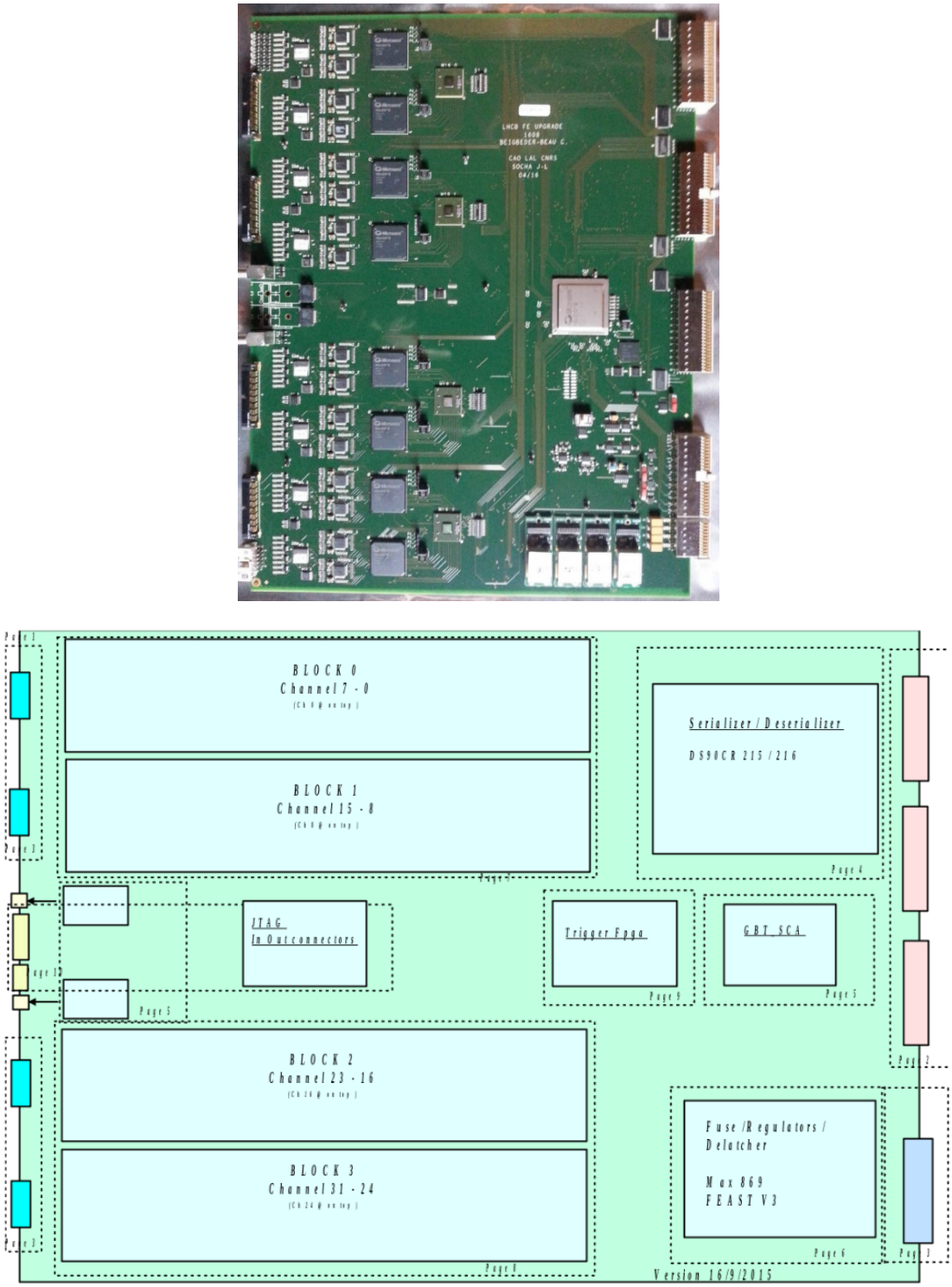
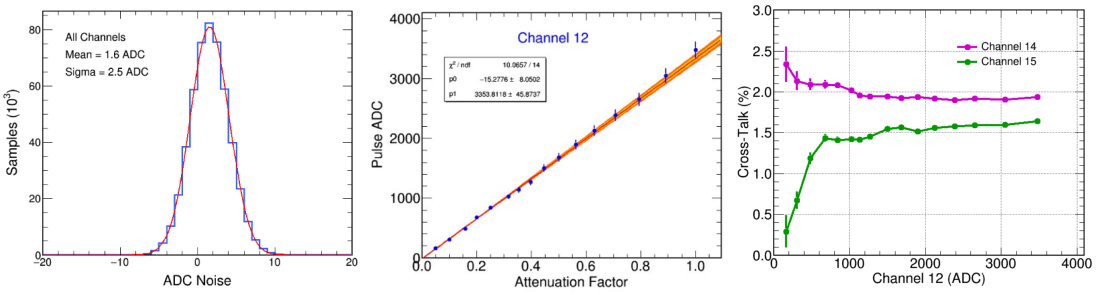


Figure 2.24: (top) Picture of a FEB prototype. (bottom) Sketch of the blocks of the FEB (from [77]).



(a) (b) (c)

Figure 2.25: (left) Average ADC noise over the 32 channels. (middle) ADC count as a function of the injected charge. In order to get a continuous variation, the same large pulse is injected through an attenuator. (right) Cross-feed ADC count received by the neighbor channels as a function of the ADC count in the channel where the signal is injected (from [77]).

the L0LLT is added to the raw data sent to the software trigger. It is mainly used to speed up the software LLT by providing higher level information or can be used as electron, photon or hadron seeds in the first stages of the HLT sequence. In addition, it could be used as a backup in case of a temporary drop of the computing farms capabilities of the software trigger. In this scenario, the information provided by the L0LLT could be used directly (*i.e.* without further reprocessing nor adding any additional information) by the software LLT to considerably reduce the computing resources it needs. However, since the PS/SPD will be removed in the upgraded LHCb detector, photon and electron candidates cannot be distinguished at the L0LLT level anymore. Thus contrary to the pre-upgrade L0 trigger, only two types of candidates can be built: hadron and generic electromagnetic candidates.

The L0LLT module of the FEB computes, for each bunch crossing, the total transverse energy (E_T^{tot}) deposited in the 32 cells associated to the FEB, the transverse energy of the maximum 2×2 cluster (E_T^{max}), the location of the maximum 2×2 cluster, and the multiplicity, defined as the number of cells having a transverse energy deposit higher than a certain threshold.

Due to bandwidth limitations, and since the L0LLT does not require a high precision (its purpose is not a measurement of the energy, but to trigger on clusters above a certain threshold), the 12 bits ADC count of the E_T measurement are converted into 10 bits ADC count. The aim of the conversion is to obtain a 10 bits word such that a well-calibrated channel with 12 bits saturation at 20 GeV transverse energy produces a 10 bits calibrated trigger signal saturated at 10 GeV. It is not necessary to resolve transverse energies above this value since the purpose

of the L0LLT is only to trigger on events above a certain E_T threshold, and this threshold is not foreseen to be ever above 10 GeV. The calibration is tuned via the mean of a 10 bits calibration constant N_c which is unique for each channel (*i.e.* for each cell of the calorimeter which corresponds to one PMT). The 10 bits trigger ADC count is thus given by

$$N_T = (N_{ADC} - N_{ped}) \times \frac{N_c}{512} \times \frac{1024}{4095} \quad , \tag{2.6}$$

where the last term corresponds to the 12 to 10 bits conversion. Since the calibration constant has a value between 0 and 1023, it allows to increase the gain of the PMT up to a factor 2 or to decrease it down to zero. A value of $N_c = 0$ can be used to completely mask a given channel which has been identified as defective. Due to the limited computing resources of the electronics of the L0LLT FPGA, the conversion is done as a series of additions instead of a multiplication. Let ABCDEFGHIJKL be the initial 12 bits word, where each letter equals 0 or 1. Similarly, let the letters ZYXWVUTSRQ be a 10 bits word defined by $N_c/512$. First, ten 22 bits words are defined as the product of the shifted initial ABCDEFGHIJKL 12 bits word and individual bits of the calibration word as

$$\begin{aligned}
L_1 &= 0000000000ABCDEFGHIJKL \times Q \quad , \\
L_2 &= 000000000ABCDEFGHIJKL0 \times R \quad , \\
L_3 &= 00000000ABCDEFGHIJKL00 \times S \quad , \\
L_4 &= 0000000ABCDEFGHIJKL000 \times T \quad , \\
L_5 &= 000000ABCDEFGHIJKL0000 \times U \quad , \\
L_6 &= 00000ABCDEFGHIJKL00000 \times V \quad , \\
L_7 &= 0000ABCDEFGHIJKL000000 \times W \quad , \\
L_8 &= 000ABCDEFGHIJKL0000000 \times X \quad , \\
L_9 &= 00ABCDEFGHIJKL00000000 \times Y \quad , \\
L_{10} &= 0ABCDEFGHIJKL000000000 \times Z \quad ,
\end{aligned} \tag{2.7}$$

where each product is just a multiplication by 1 or 0, so not resource intensive. Then, the computation of the 10 bits E_T is performed in several steps. First, the lines are added two by two as

$$S_{i,i+1} = L_i + L_{i+1} \quad , \; i = \{1, 3, 5, 7, 9\} \quad . \tag{2.8}$$

Then, the two first $S_{i,i+1}$ are added two by two as

$$S_{j,j+1,j+2,j+3} = S_{j,j+1} + S_{j+2,j+3} \quad , \; j = \{1, 5\} \quad , \tag{2.9}$$

and then added together as

$$S_{12345678} = S_{1234} + S_{5678} \quad . \tag{2.10}$$

The final result is obtained by adding the last two components

$$abcdefghijklmnopqrstuv = S_{12345678} + S_{910} \quad , \tag{2.11}$$

where only $abcdefghijklmno$ is needed to compute the final result. The final result is obtained by rounding the $abcdefghijklmno$ word in the following way:

- if no is 10 or 11, the $abcdefghijklmno$ word is incremented
- else the $abcdefghijklmno$ word is kept as it is

The final 10 bits word is given by the 10 bits in the center of the word, namely $cdefghijklm$, *i.e.* with a saturation at 1023. In addition, an alternative mode of 12 to 10 bits conversion is also implemented. In this alternative mode, the 12 bits are shifted by 2 bits, before the calibration. This corresponds to increasing the gain of the trigger calibration by a factor four. This mode of computation can be used in case the PMT is operated at reduced gain due to an excess in the anode current or instability with high voltage.

Once the 12 to 10 bits conversion has been computed for each channel, the total transverse energy as well as the multiplicity is trivial to compute. The maximum energy cluster however, is more complex, because it requires information from neighbor FEBs. Indeed, the maximum energy cluster might be located at the border the rectangle associated to each FEB. That is why each FEB receives, in addition to the 32 E_T of its own associated cells, the E_T of the left cell column of the right neighbor FEB, the upper row of the bottom neighbor FEB and of the upper-left corner cell of the bottom right neighbor FEB, as sketched in figure 2.26. In addition, each FEB sends the information of its left column, top row and upper-left cell to the relevant neighbor FEB. The clusters are then built by scanning 2×2 cells and summing the E_T of the four cells. There are thus 32 clusters to be computed, as shown in figure 2.26. Note that a given cell may contribute to several clusters. The cluster with the highest transverse energy defines E_T^{max}. To finish, each cell has an associated address (0 to 31 as shown in figure 2.26). The address of each cluster is defined as the address of the top-left cell of this cluster. The L0LLT then returns a 32 bits word which is encoded as follows:

- bit 0-4 (5 bits): address of the maximum energy cluster
- bit 5-14 (10 bits): E_T^{max}
- bit 15-26 (11 bits): E_T^{tot}

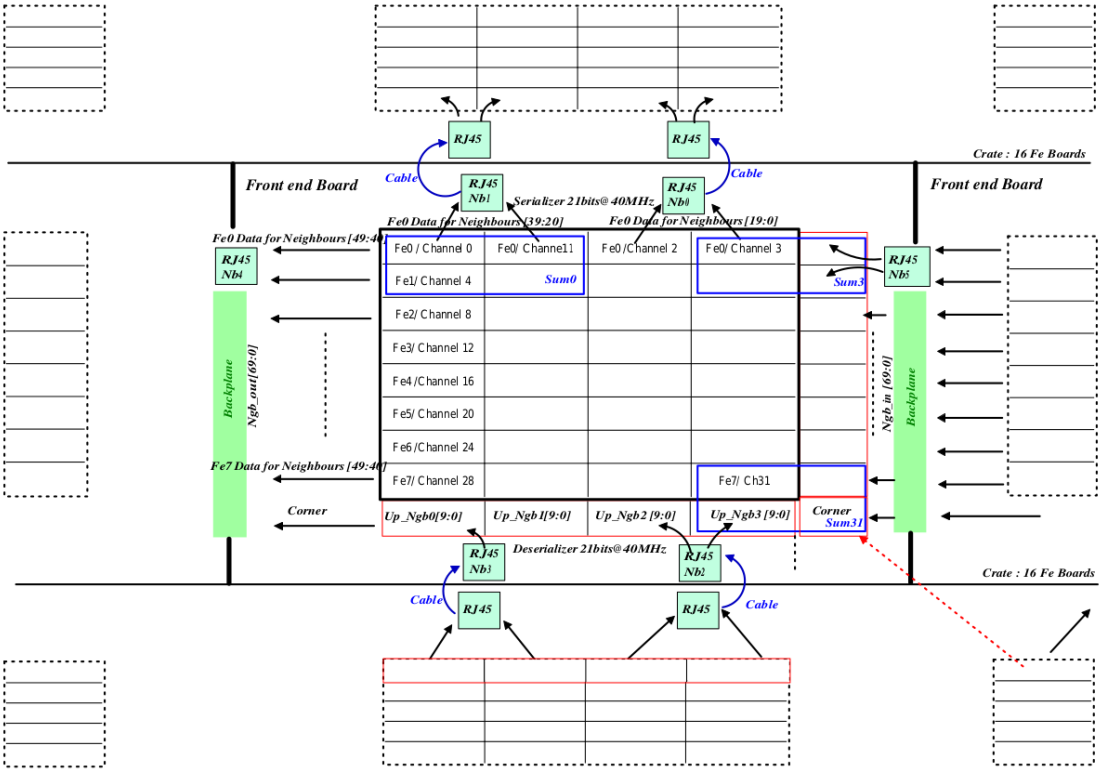


Figure 2.26: Sketch of the maximum transverse energy cluster computation. This is a zoom of figure 2.22 in the vicinity of one random FEB associated to 32 calorimeter cells. Within one crate, each board sends its left column cells data to its left FEB neighbor. Each FEB also sends its top row information to the FEB at the same position in the upper crate. To finish, each FEB sends its top-left cell information to the cell located in the upper crate on its left. Some examples of energy clusters are represented by blue rectangles (from [77]).

1	2	3	4	5	6	1’
3’	4’	5’	6’	1’	2’	3’

Figure 2.27: Layout of the test bench. Each number represents a FEB, and each row represents a crate. The crate on the top is the real one in which the FEBs (number 1 to 6) are plugged in. The crate on the bottom is a fake one created by copies of the FEBs (the primed numbers) which send their information via RJ45 cables to mimic missing neighbors. The numbering is shifted such that no card receives its own information.

- bit 26-31 (6 bits): multiplicity (number of cells with E_T above a certain threshold)

The FEBs are produced by an external company. Thus, before the FEBs are installed in the platforms of the calorimeters, they have to be tested in the laboratory. A specific test of the L0LLT module of the FEB has been designed and will be added to the full test procedure of each FEB. The development of a full simulation of the L0LLT has been one of the major tasks related to the upgrade performed in this thesis. Since the software L0LLT reproduces all the computations of the hardware L0LLT, it is the main testing tool of the L0LLT module of the FEBs. The FEBs will be tested on two identical test benches. Each test bench consists in one crate filled with one control board and six FEBs. In addition, a device called MiniDAQ will be connected to the boards. The main purpose of the MiniDAQ is to send inputs via optical fibers to the boards to simulate energy deposited in the cells of the calorimeters, as well as to record the outputs of the FEBs, namely the E_T of each channel and the 32 bits L0LLT word. Since each FEB needs three neighbors (including from neighbors in a different crate) to compute the L0LLT word, fake neighbors are created by sending the information of some FEBs via RJ45 cables to the neighbor-less boards, as sketched in figure 2.27. Two types of patterns will be injected to the FEBs via the MiniDAQ. The pattern A consist in a series of seven 0s followed by one 1. This pattern has been found to maximize the (de-)serializers consumption and is thus used to test hardware related issues. The pattern B is made of random values and aims at testing the global functioning of the L0LLT, including software issues related to the firmware[11]. In both cases (pattern A and B), the same input is sent to the hardware and the software L0LLT. If the output of the hardware L0LLT matches that of the software one, the card passes the test and is ready to be installed in the calorimeter platforms.

[11]The firmware is the internal software of the electronics components of the FEBs.

Preamble

In order to ease the reading, the angular analysis of the $B^0 \to K^{*0}e^+e^-$ decay channel at very low q^2 presented in this thesis has been separated in two chapters. The selection of $B^0 \to K^{*0}(\to K^+\pi^-)e^+e^-$ candidates is explained in Chapter 3 while the angular analysis is detailed in Chapter 4.

A second q^2 bin $[0, 0.0001]\,\mathrm{GeV}^2/c^4$, denoted as gamma-$q^2$ bin, is considered alongside the q^2 bin of the analysis (see section 1.4.2) $[0.0001, 0.25]\,\mathrm{GeV}^2/c^4$, denoted as very-low-$q^2$ bin. Due to the large branching fraction (about two orders of magnitudes higher than the $B^0 \to K^{*0}e^+e^-$ one) of the $B^0 \to K^{*0}\gamma$ decay, the reconstructed $B^0 \to K^{*0}e^+e^-$ candidates in the gamma-q^2 bin are largely dominated by $B^0 \to K^{*0}\gamma(\to e^+e^-)$ events where the photon converted into an e^+e^- pair in the material of the detector (see section 3.3.2). The physics as well as the detector response of $B^0 \to K^{*0}\gamma(\to e^+e^-)$ decays are very similar to $B^0 \to K^{*0}e^+e^-$ decays at low q^2, while the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ decay channel offers much higher yields thanks to its higher branching fraction. This makes the gamma-q^2 bin an excellent crosscheck bin for the angular and mass fits.
Due to the kinematic differences coming from the different center-of-mass energies between Run 1 (R1) and Run 2 (R2), the two datasets are treated separately, akin to the two trigger categories (see section 3.1) and the two q^2 bins. The analysis is thus separated in eight categories (2 years x 2 triggers x 2 q^2 bins) fitted simultaneously and sharing in particular the four physical observables F_L, $A_T^{(2)}$, A_T^{Im} and A_T^{Re}. First, the main backgrounds are studied and their contamination is estimated in section 3.3. If possible, veto cuts are applied to reduce the background contamination. In case the background contribution remains non negligible, the $m(K^+\pi^-e^+e^-)$ (see section 3.6.3) and angular (see section 4.3) shapes are modeled and added to the mass and angular fits. Then, the $m(K^+\pi^-e^+e^-)$ invariant mass is fitted to data in section 3.6.5 in a wide range $[4500, 6200]\,\mathrm{MeV}/c^2$ to properly model the radiative tail and the partially reconstructed background.

Next, the distortion to the angular distributions caused by the reconstruction and the selection of $B^0 \to K^{*0} e^+ e^-$ candidates is studied and modeled in section 4.2. In order to make the angular modeling of backgrounds easier and to reduce the background contamination, the angular fit is performed in a reduced mass window $[5000, 5400]\,\mathrm{MeV}/c^2$ (see section 4.1). A four dimensional fit to the $m(K^+\pi^- e^+ e^-)$ invariant mass and the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ is finally performed in section 4.6 to extract the four physical observables F_L, $A_T^{(2)}$, A_T^{Im} and A_T^{Re}.

Comparing the experimental results to the expected (or desired) results during the process of the analysis could lead to potential biases. Indeed, one could tune (even unconsciously) the analysis to adjust towards the targeted results. To avoid such biases, the angular fits to $B^0 \to K^{*0} e^+ e^-$ data are blinded. That is, the angular distributions in $B^0 \to K^{*0} e^+ e^-$ data are not looked at and the angular fit to data is not performed until the analysis is fully finalized. The $m(K^+\pi^- e^+ e^-)$ distribution however is not blinded since it does not reveal any information about the four physical observables F_L, $A_T^{(2)}$, A_T^{Im} and A_T^{Re}.

Chapter 3

Selection of $B^0 \to K^{*0}e^+e^-$ events

This Chapter explains how $B^0 \to K^{*0}(\to K^+\pi^-)e^+e^-$ candidates are selected in the LHCb data. The main backgrounds are studied in sections 3.3 and 3.4. After the full selection (summarized in section 3.5) is applied, the $m(K^+\pi^-e^+e^-)$ invariant mass shape of $B^0 \to K^{*0}e^+e^-$ events as well as the backgrounds which are found to be non negligible are modeled in section 3.6. A fit to the $m(K^+\pi^-e^+e^-)$ invariant mass is finally performed on the selected $B^0 \to K^{*0}e^+e^-$ candidates in LHCb data in section 3.6.5.

3.1 Trigger and pre-selection

As explained in section 2.2.4, only events which have passed the LHCb L0 hardware trigger are saved. In this analysis, $B^0 \to K^{*0}(\to K^+\pi^-)e^+e^-$ candidates are selected requiring that either

- At least one of the two electron candidates has fired the electron L0 trigger. These types of events are referred to as L0Lepton (L0L).
- The event has been triggered by a decay product which is not part of the signal candidate. Since b hadrons are produced in pairs at LHCb, this typically happens when the event has been triggered by a decay product of the other b hadron in the event. These types of events are referred to as L0Independant (L0I).

Given that the shape of the reconstructed $m(K^+\pi^-e^+e^-)$ mass, the background contributions as well as the angular acceptance depends on the L0 trigger, the entire analysis is performed separately for the two different trigger categories described above. As sketched in figure 3.1a, it may happen that an event fulfills the two requirements, *i.e.* that one (or both) electron(s) has fired the electron L0

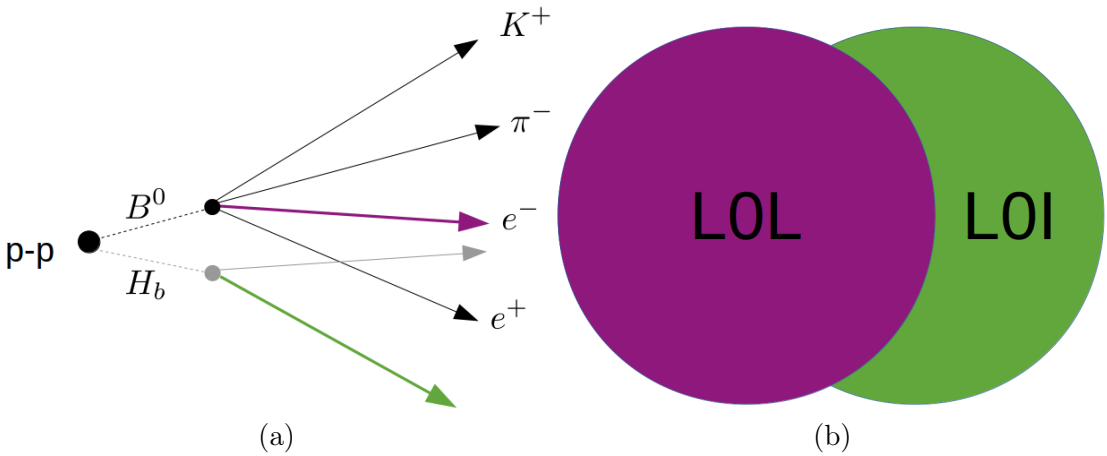


Figure 3.1: (left) Sketch of an event which is both L0L and L0I, where only the decay products associated to the two B mesons produced in the collision are drawn. In this example, the purple electron of the signal candidate has fired the electron L0 line, while the green kaon of the other B meson produced in the event has fired the hadron L0 line. (right) Illustration of the exclusive definition of the trigger categories. Events which are in principle both L0L and L0I (like the example in the left figure) end up in the primary trigger category L0L.

trigger and another decay product of the event which is not part of the signal candidate has fired a trigger line as well. In order for each candidate to be associated to a single trigger category, the latter are defined to be mutually exclusive. As illustrated in figure 3.1b, the L0L trigger category is taken as primary category while the L0I trigger category is defined as events triggered by a decay product independent of the signal and not triggered by at least one electron of the signal candidate. Unless explicitly stated otherwise, the term L0I throughout this thesis refers to the exclusive L0I category.

On top of the L0 hardware trigger, $B^0 \to K^{*0}e^+e^-$ events are also required to fire some HLT software trigger lines (see section 2.2.4). The HLT1 lines are based on generic track information and are fired by rather basic requirements (such as transverse momentum for instance). The HLT2 lines rely on multivariate algorithms trained to select events of a given topology, here $B \to$ 2-, 3-body decays involving electrons.

On top of the trigger requirements, a pre-selection (including the *stripping* selection, see section 2.2.4) is applied to $B^0 \to K^{*0}e^+e^-$ candidates. All final state candidates are required to have high transverse momentum and good track quality. The latter is insured by requiring small χ^2/ndf, which is the χ^2 of the fit to a given decay vertex divided by the number of degrees of freedom of the fit. Besides, the

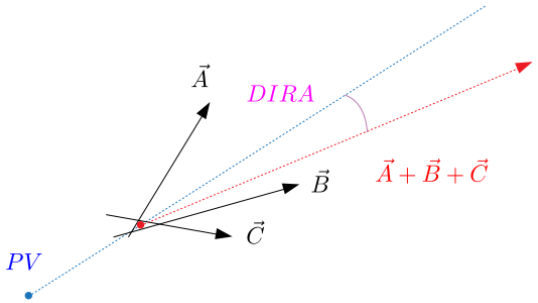


Figure 3.2: Sketch of a B meson decaying to three particles A, B and C, whose reconstructed momenta are drawn in black. The direction of the sum of the three final state particles is drawn in red, while the line between the primary and the reconstructed decay vertex is drawn in blue. The angle between the two (shown in purple) is defined as the `DIRA` angle.

probability[1] of all tracks to be a fake track coming from a random combination of hits in the tracking system (called a ghost track) is required to be low.
In addition, good vertexing is required. All final state particles are required to have a large displacement from the primary vertex. This is obtained by requiring a large χ^2_{IP}, which is the difference of χ^2 if the primary vertex is computed with a given candidate candidate or without it. Besides, the angle (called `DIRA`) between the direction of the sum of all final state particles momenta and a line drawn from the PV to the decay vertex (see figure 3.2 for illustration) is required to be small. Moreover, PID requirements are applied to all final state candidates, with a combination of `PIDX` and `ProbNN` (see section 2.2.3) variables.
Furthermore, all final state candidates are required to have hits in the $RICH$ detectors and electrons are required to be within the $ECAL$ acceptance.
To finish, the reconstructed mass of the $K^+\pi^-$ system is required to be within $100\,\text{MeV}/c^2$ of the nominal mass of the K^{*0} meson. The full selection is summarized in tables 3.3 and 3.4 of section 3.5.

3.2 Corrections to simulation

The analysis relies on Monte Carlo (MC) simulation to study and estimate background contaminations (see section 3.3), to compute efficiencies (see section 3.5), to train a multivariate classifier to reduce the combinatorial background (see section 3.4.2), to model the mass shapes (see section 3.6.1) as well as the angular shape of some background components (see section 4.3) and to model the angular

[1]The probability to be a ghost track is computed by a dedicated neural-network based algorithm [78] using tracking information.

acceptance (see section 4.2). However, the LHCb MC is known to not perfectly reproduce the data. Thus, the simulation samples are corrected for the PID response, the L0 trigger response, the generated event multiplicity and B kinematics as well as the event reconstruction. This part of the work relies on a general approach developed by the LHCb collaboration for Lepton Universality tests ([59] as well as its R2 update to be published). The weights are computed step by step and applied on top of each other in the order mentioned above.
For the PID corrections, data-based PID efficiency tables are used (see below). All the other corrections are obtained via data/MC ratios, or via training a dedicated multivariate classifier using simulation and data of the resonant $B^+ \to K^+ J/\psi\,(\to e^+e^-)$ decay mode. This decay is used because a very clean signal can be extracted from data, using a specific cut to get rid of most of the background. Indeed, one can recompute the B mass using the constraint that the dielectron mass equals the nominal mass of the J/ψ. This way, all the events in the radiative tail due to large (unrecovered) bremsstrahlung emissions are pushed back under the peak. By requiring the J/ψ constrained reconstructed B mass to be above $5150\,\mathrm{MeV}/c^2$ one gets rid of most of the partially reconstructed and combinatorial backgrounds. The J/ψ-q^2 bin is defined as $[6, 11]\,\mathrm{GeV}^2/c^4$. The $B^+ \to K^+ J/\psi\,(\to e^+e^-)$ decay mode is preferred to the $B^0 \to K^{*0} J/\psi\,(\to e^+e^-)$ one because of its larger statistics. It is checked that the corrections obtained from the B^+ channel are compatible with the ones computed with the B^0 channel.

3.2.1 Particle identification

The PID response in MC samples is estimated using PID efficiency tables. These tables are produced centrally for the whole LHCb collaboration [79] and are called `PIDCalib` tables. They are produced separately for each year of data taking and particle species. Instead of cutting on the PID variables like in data, the efficiency of the PID cuts related to a given particle species (the definition of the PID cuts is given in table 3.4) is estimated for each particle species from the `PIDCalib` tables as a function of the momentum p and the pseudo-rapidity η. For electrons, due to different background compositions, two separate maps are computed for events with no bremsstrahlung and bremsstrahlung added. To illustrate the results, the PID efficiencies of pions, kaons and electrons for the year 2012 are shown in figures 3.3 and 3.4. These efficiencies are then applied per track.

3.2.2 L0 trigger

The analysis is performed in two trigger categories L0L and L0I (see section 3.1). Thus, it is important to properly evaluate their relative efficiencies.
The weights to correct for the L0 efficiencies are obtained by data/MC efficiency

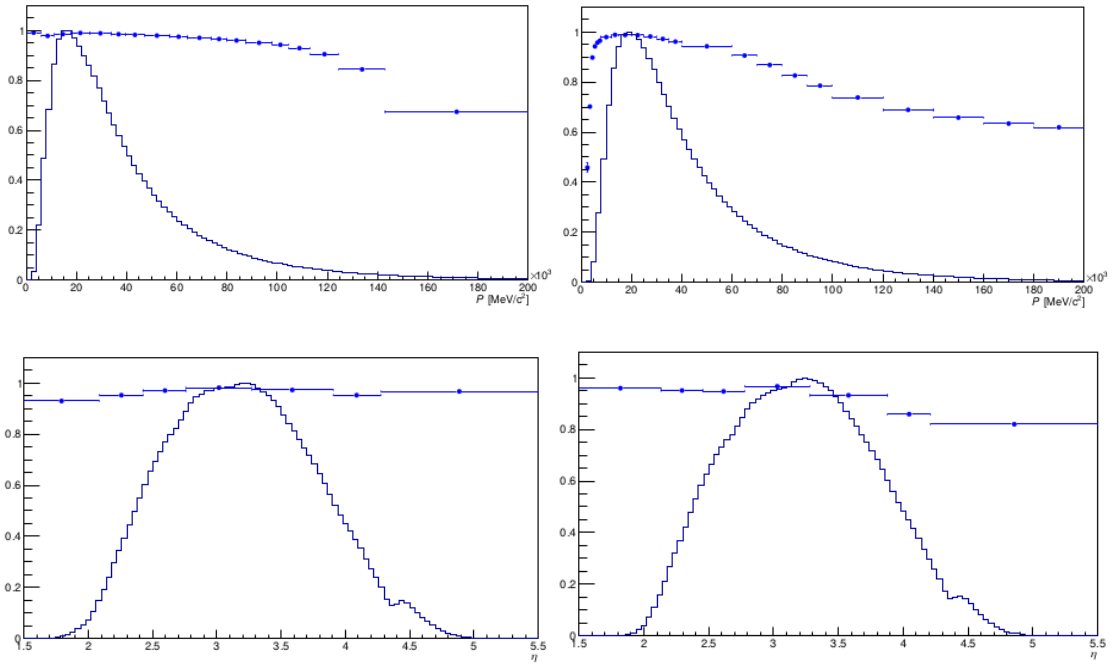


Figure 3.3: Efficiency (blue dots) of pions (left) and kaons (right) to pass the pion and kaon PID requirements respectively, as a function of the particle momentum (top) and pseudo-rapidity (bottom), for the data taking year 2012. The blue histogram represents the distribution of the corresponding variable in the `PIDCalib` sample used to produce the efficiency table.

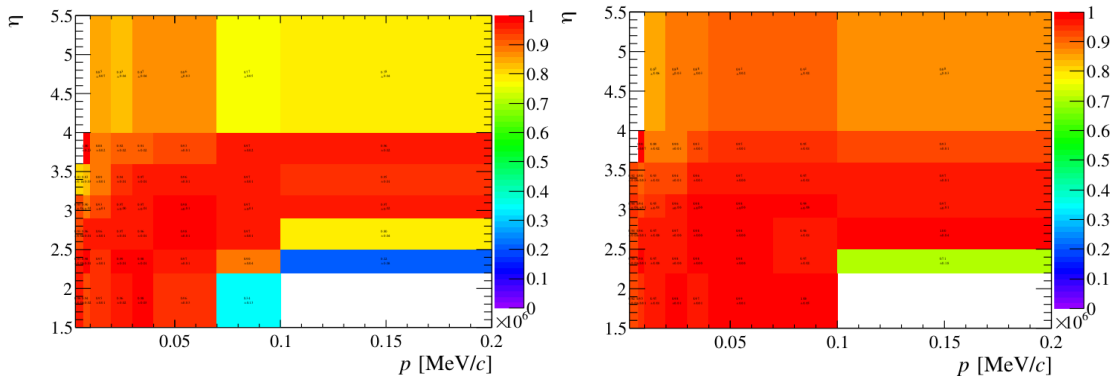


Figure 3.4: Efficiency to pass the electron PID requirements, as a function of the particle momentum and pseudo-rapidity, for the data taking year 2012. The efficiency for electrons without (with) bremsstrahlung added is shown on the left (right).

ratios, separately for each year of data taking, to reflect the different trigger configurations of each year. The full pre-selection (see section 3.1) is applied to both $B^+ \to K^+ J/\psi (\to e^+e^-)$ data and MC samples. For the MC sample, no PID cut is applied but the PID efficiency weights discussed in section 3.2.1 are applied instead. In addition to the J/ψ constrained reconstructed B mass cut discussed above, the remaining background is further reduced by applying a tight cut around the reconstructed mass $|m(K^+e^+e^-) - m^{\mathrm{PDG}}_{B^+}| < 60\,\mathrm{MeV}/c^2$, where m^{PDG} is the world average measured mass from the Particle Data Group (PDG) [80].
Since all events saved on disk have fired some trigger, there is no trigger-less data sample to evaluate efficiencies from. The weights to correct for the L0 efficiencies are therefore computed using the `Trigger Independent of Signal / Trigger On Signal` (`TISTOS`) [81] method. The main idea of the `TISTOS` method is to evaluate the number of events passing a `TOS` requirement within a `TIS` population which is as independent as possible to the `TOS` requirement which is probed. The efficiencies are computed as

$$\varepsilon^s_t = \left(\frac{N_{\texttt{TIS\&TOS}}}{N_{\texttt{TIS}}} \right)^s \quad , \tag{3.1}$$

where $s = \{\mathrm{data}, \mathrm{MC}\}$ denotes the sample, $t = \{\mathrm{L0L}, \mathrm{L0I}\}$ denotes the inclusive trigger category, and `TIS` and `TOS` are defined as

- L0L category:
  - `TIS` = event triggered by any track which is not part of the signal candidate
  - `TOS` = event triggered by at least one electron of the signal candidate
- L0I category:
  - `TIS` = event triggered by at least one electron or one hadron of the signal candidate
  - `TOS` = event triggered by any track which is not part of the signal candidate

The efficiencies for the L0L category are computed in bins of electron transverse energy and $ECAL$ region (inner, middle or outer) while the efficiencies for the L0I category are computed in bins of B transverse momentum (p_T). To reduce the dependency on the binning scheme, the distributions of ε^s_{L0L} and ε^s_{L0I} are fitted

respectively with ad hoc functions. The weights for the inclusive trigger categories are obtained by the ratio of the fitted efficiencies as

$$w_t = \frac{\varepsilon_t^{data}}{\varepsilon_t^{MC}} \quad . \tag{3.2}$$

Since the L0I trigger category in this analysis is defined as an exclusive category, the weights obtained from eq. (3.2) cannot be used directly. The weights for the exclusive L0I category are given by

$$w_{L0I!} = w_{L0I} \cdot \frac{1-\varepsilon_{L0L}^{data}}{1-\varepsilon_{L0L}^{MC}} = \frac{\varepsilon_{L0I}^{data}}{\varepsilon_{L0I}^{MC}} \cdot \frac{1-\varepsilon_{L0L}^{data}}{1-\varepsilon_{L0L}^{MC}} \quad . \tag{3.3}$$

To illustrate the results, the L0 trigger corrections for the data taking year 2012 are shown in figure 3.5 for the L0L trigger category and in figure 3.6 for the L0I trigger category. As shown in figure 3.7, the L0I data/MC efficiencies ratios are found to be similar for $B^+ \to K^+ J/\psi (\to e^+e^-)$ and $B^+ \to K^+ J/\psi (\to \mu^+\mu^-)$ samples. Therefore, the muon efficiencies are used for the L0I efficiency corrections to take advantage of the higher statistics in the muon channel.

3.2.3 B kinematics, multiplicity and reconstruction

To correct for the B kinematics, one has to chose the loosest trigger possible, to reduce the trigger bias to its minimum. Since at LHCb the muon trigger is much loser than the electron one, the muon channel $B^+ \to K^+ J/\psi (\to \mu^+\mu^-)$ is used. Moreover, one benefits from the higher muon statistics available at LHCb. In addition, since the occupancy of the event does not depend on the flavor of the final state leptons, the muon channel is also used to compute the occupancy corrections. It is checked that using the electron channel instead of the muon channel, and that using the $B^0 \to K^{*0} J/\psi (\to \mu^+\mu^-)$ decay channel instead of the $B^+ \to K^+ J/\psi (\to \mu^+\mu^-)$ one gives similar results. The proxy variables used for the B kinematics and multiplicity are the following:

- B kinematics: B transverse momentum and pseudo-rapidity.
- Multiplicity: number of primary vertices and number of tracks in the $VELO$.

In order to take into account the correlations between these proxy variables, a multivariate classifier is used: the GBreweighter from the hep_ml library [82]. The GBreweighter is trained using $B^+ \to K^+ J/\psi (\mu^+\mu^-)$ `L0_Muon`[2] MC and ${}_s\mathcal{P}$lotted[3]

[2]Similarly as for electrons, `L0_Muon` requires that at least one of the muons of the signal candidate has triggered the L0 muon line

[3]${}_s\mathcal{P}$lot is a statistical method to separate signal from background events [83].

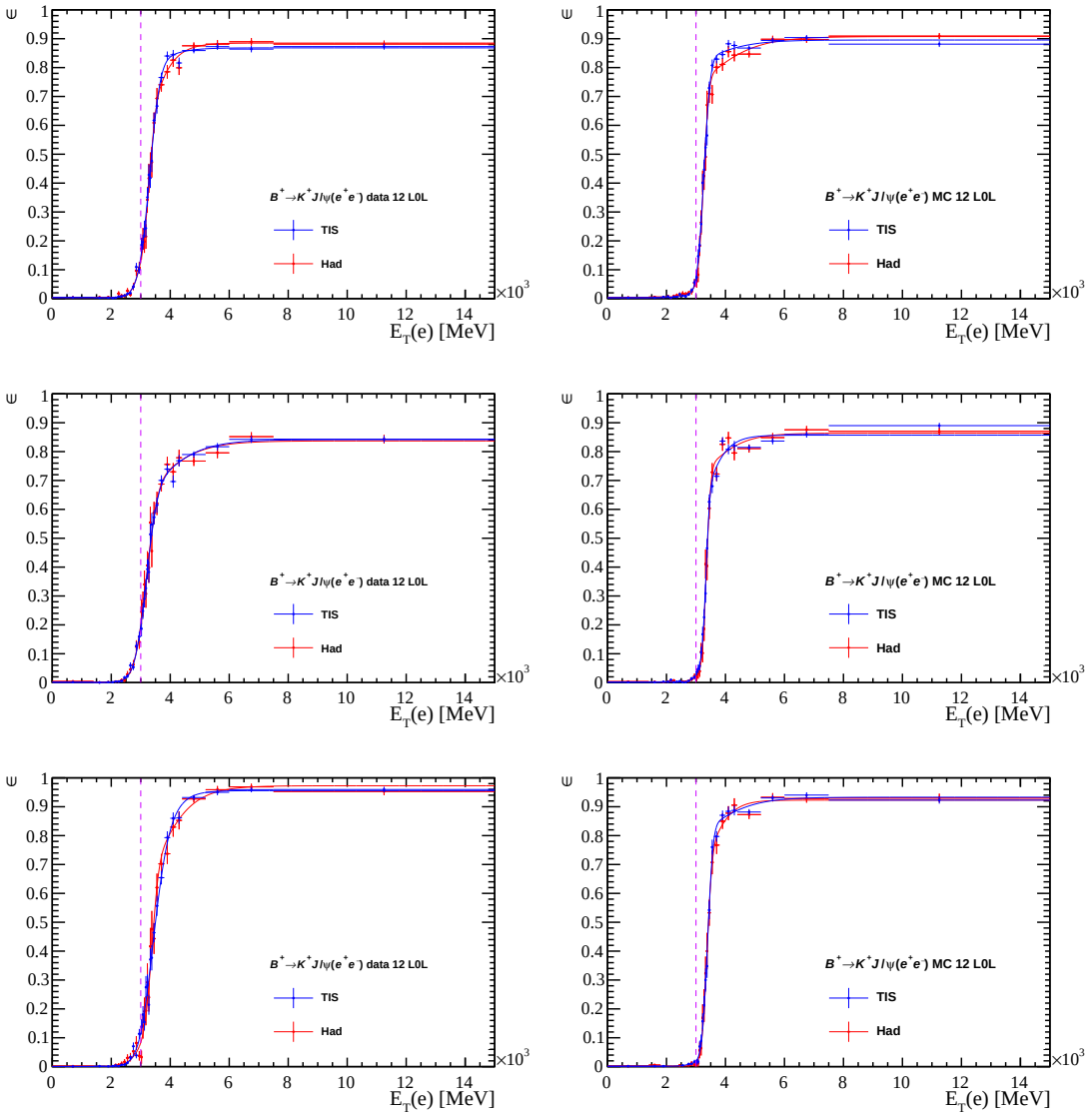


Figure 3.5: Efficiencies of the L0L trigger requirement as a function of the transverse energy of the electron in the three $ECAL$ regions (inner, middle, outer) from top to bottom. The left (right) plot shows 2012 $B^+ \to K^+ J/\psi (\to e^+e^-)$ data (simulation) events. The nominal data points and fitted functions are shown in blue, while the red points and curve show a crosscheck with an alternative `TIS` requirement based on hadrons only.

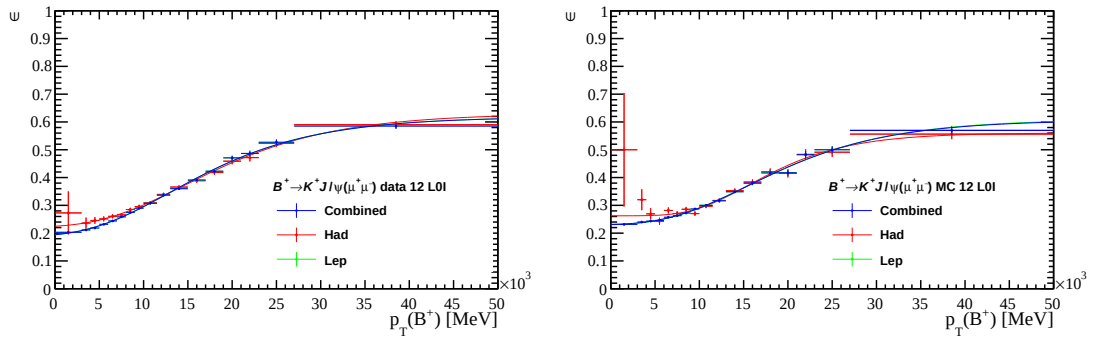


Figure 3.6: Efficiencies of the L0I trigger requirement as a function of the transverse momentum of the B meson. The left (right) plot shows 2012 $B^+ \to K^+ J/\psi (\to \mu^+\mu^-)$ data (simulation) events. The nominal data points and fitted functions are shown in blue. The green (red) points and curves show a crosscheck using an alternative `TOS` sample based on electrons (hadrons) only instead of the combination of the two.

data. The PID and L0 corrections discussed in sections 3.2.1 and 3.2.2 are already applied to the MC sample used for the training. Since the aim at this stage is to correct for generator level kinematics and not for reconstruction effects, the correction is applied on the true momentum and true pseudo-rapidity of the B meson in the MC samples.
To correct MC/data differences due to reconstruction effects, four proxy variables are used: the reconstructed B transverse momentum and pseudo-rapidity, the χ^2 of the vertex fit and the χ^2 of the impact parameter of the B meson. Since electrons and muons are known to perform differently in the reconstruction process, a second reweighter step is trained using $B^+ \to K^+ J/\psi (e^+e^-)$ MC and data. For the MC training sample, all corrections, including the B kinematics and multiplicity ones, are applied.

Figure 3.8 shows the distribution of some key variables of ${}_s\mathcal{P}$lotted $B^+ \to K^+ J/\psi (\to e^+e^-)$ data, compared to the bare unweighted MC or the same MC with all corrections applied. Before applying the corrections, some discrepancies are observed between the data and the bare MC, while after applying the corrections, the weighted MC matches the data. Nevertheless, these corrections do not affect much the angular distributions of $\cos\theta_K$, $\cos\theta_\ell$ and $\tilde{\phi}$ in the $B^0 \to K^{*0} e^+e^-$ MC as shown in Appendix C.1.

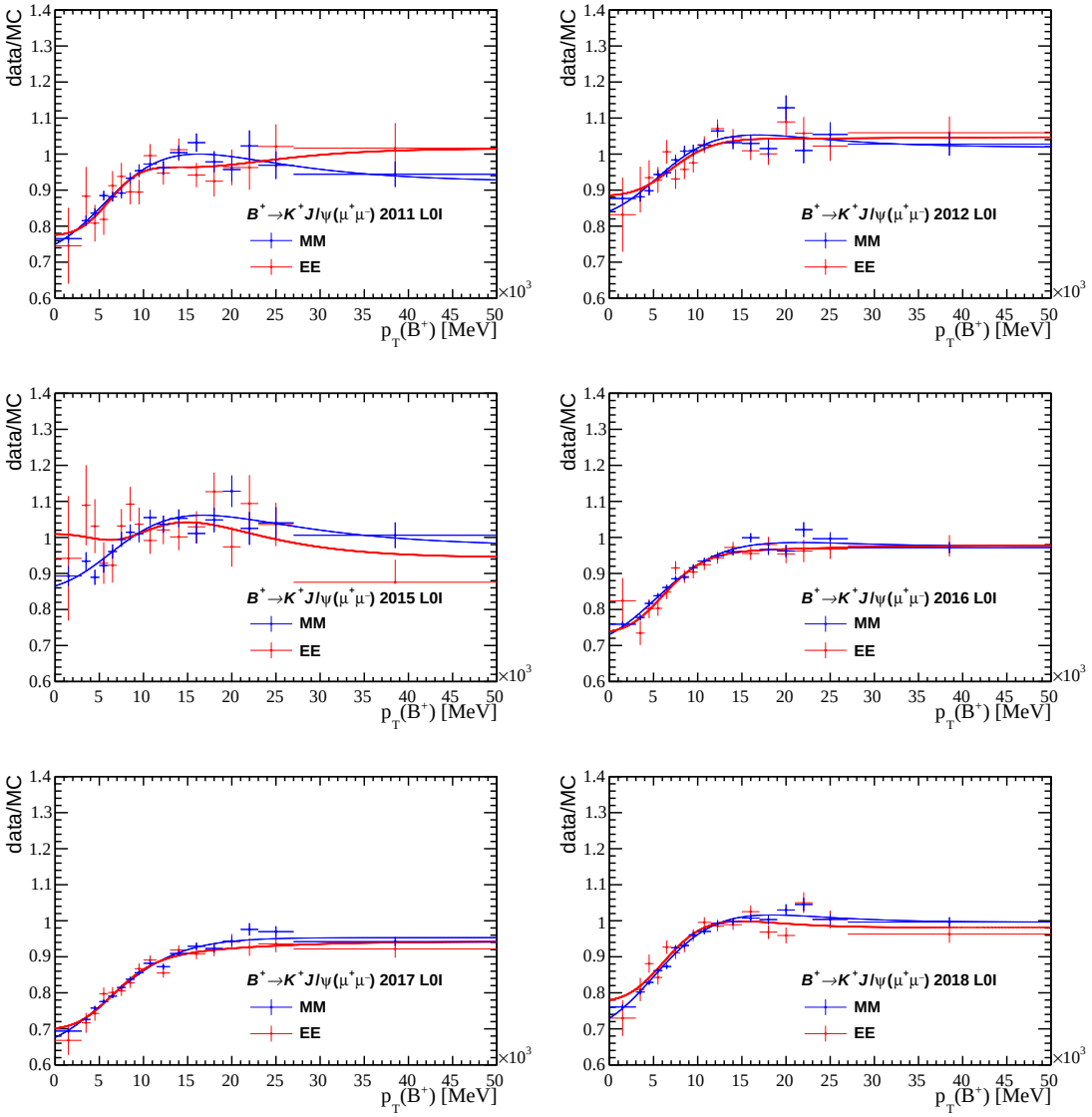


Figure 3.7: Data/MC efficiency ratios for the L0I trigger requirement as a function of the transverse momentum of the B meson. The ratios obtained from $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$ simulation and data are shown in blue while the ones obtained from $B^+ \to K^+ J/\psi(\to e^+e^-)$ simulation and data are shown in red.

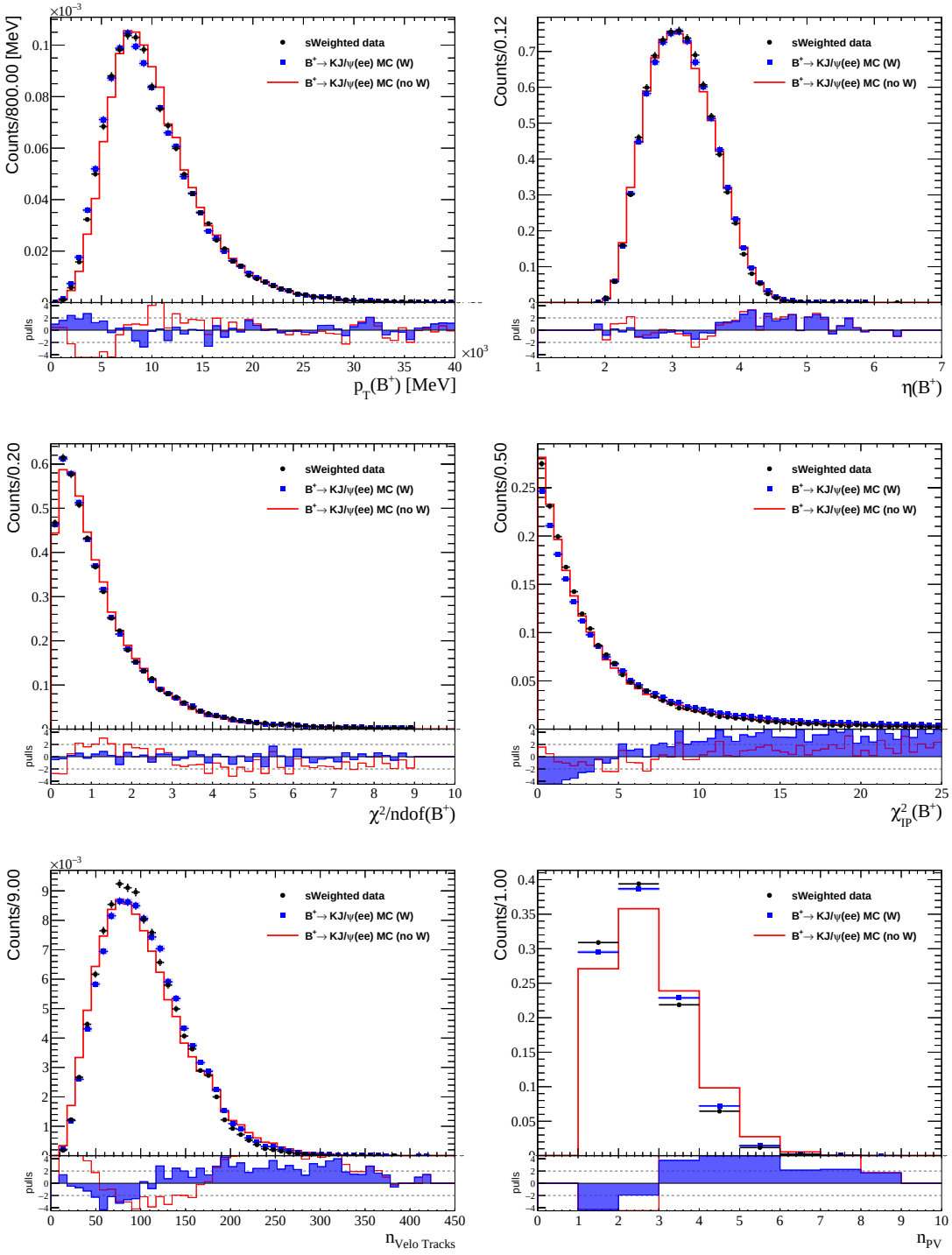


Figure 3.8: MC/data comparison plots for 2012 L0L $B^+ \to K^+ J/\psi (\to e^+e^-)$ samples. The ${}_s\mathcal{P}$lotted data is shown in black, the bare (without corrections) MC is shown in red and the reweighted MC is shown in blue.

3.3 Specific backgrounds

The decay $B^0 \to K^{*0}e^+e^-$ being rare, it is subject to many specific backgrounds which can pass the stripping and pre-selection. The main backgrounds which can pollute the $B^0 \to K^{*0}e^+e^-$ are studied and their contribution is estimated. When necessary, specific vetoes are developed to reduce the amount of background.

3.3.1 Semileptonic background

Semileptonic $B^0 \to D^-(\to K^{*0}(\to K^+\pi^-)e^-\overline{\nu}_e)e^+\nu_e$ decays are an important background for the $B^0 \to K^{*0}e^+e^-$ decay. Indeed, this decay has the same final state as the signal $B^0 \to K^{*0}e^+e^-$ decay (since neutrinos are not measured at LHCb) and has a branching ratio four orders of magnitude higher than the signal. If the lost energy carried by the neutrinos is small, then this event can be reconstructed as $B^0 \to K^{*0}e^+e^-$ and pass the selection. Since the $K^*e^-\overline{\nu}_e$ come from a D^- meson and the energy of the neutrino is lost, the mass of the K^*e^- is expected to be below the mass of the D^- meson ($\sim 1870\,\mathrm{MeV}/c^2$). Looking at the problem from another point of view, this means that such events present an asymmetry in the energy distribution between the two electrons, having a preference for high energy e^+ and low energy e^-. This is reflected in the θ_l angle which goes as

$$\cos\theta_\ell \sim \frac{E_{e^+} - E_{e^-}}{E_{e^+} + E_{e^-}} \quad \text{for } B^0 \text{ and} \quad \cos\theta_\ell \sim \frac{E_{e^-} - E_{e^+}}{E_{e^+} + E_{e^-}} \quad \text{for } \overline{B}^0. \tag{3.4}$$

Thus, $B^0 \to D^-(\to K^{*0}(\to K^+\pi^-)e^-\overline{\nu}_e)e^+\nu_e$ decays peak at high values of $\cos\theta_\ell$ as shown in figure 3.9. However, A_T^{Re} is related to the forward-backward asymmetry (see section 1.4), therefore it is roughly speaking a measure of the asymmetry between positive and negative $\cos\theta_\ell$ events. For this reason, a symmetric veto cut $|\cos\theta_\ell| < 0.8$ is applied. The $B^0 \to K^{*0}e^+e^-$ fraction lost with this cut is as low as $\sim 5\%$. Moreover, the terms sensitive to the photon polarization in eq. (1.35) involving $A_T^{(2)}$ and A_T^{Im} are proportional to $\sin^2\theta_l$. Thus these lost events were bearing a very small sensitivity to the photon polarization anyway. To finish, the veto cut $|\cos\theta_\ell| < 0.8$ has another nice feature. Indeed, $B^0 \to K^{*0}\gamma$ decays where the photon converts into an e^+e^- pair can leak into the signal q^2 region, if the dielectron angle is large or if the e^+e^- pair is very asymmetric (this background is studied in details in section 3.3.2). As shown by the relation in eq. (3.4), the veto cut $|\cos\theta_\ell| < 0.8$ already does half of the job by killing the asymmetric pairs.

The contamination from semileptonic decays with respect to signal $B^0 \to$

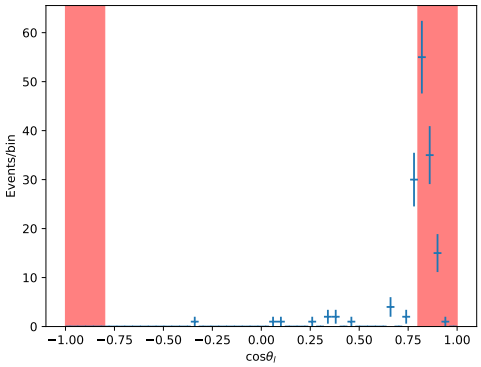


Figure 3.9: Distribution of $\cos\theta_\ell$ in the $B^0 \to D^-(\to K^{*0}(\to K^+\pi^-)e^-\overline{\nu}_e)e^+\nu_e$ MC after the full selection (except the veto cut and the BDT cut). The events removed by the veto cut are highlighted in red.

$K^{*0}e^+e^-$ is estimated by

$$C_{\mathrm{SL}} = \frac{\mathcal{B}\left(\mathrm{SL}\right) \times \varepsilon\left(\mathrm{SL}\right)}{\int_{q^2_{true,min}}^{q^2_{true,max}} \frac{d\mathcal{B}(B^0 \to K^{*0}e^+e^-)}{dq^2} dq^2 \times \varepsilon\left(B^0 \to K^{*0}e^+e^-\right)\left(q^2_{true,min}, q^2_{true,max}\right)} \quad , \tag{3.5}$$

where SL $\equiv B^0 \to D^-(\to K^{*0}(\to K^+\pi^-)e^-\overline{\nu}_e)e^+\nu_e$. The efficiencies are obtained from simulation. The integrated differential branching fraction of $B^0 \to K^{*0}e^+e^-$ is computed using `Flavio` [12] and yields $3.89 \pm 0.90 \times 10^{-7}$. A wide range of $q^2_{true,min} = 0.000009\,\mathrm{GeV}^2/c^4$ and $q^2_{true,max} = 0.64\,\mathrm{GeV}^2/c^4$ is chosen in order to take into account possible leakage at the border of the q^2 range of the analysis. The $B^0 \to D^-(\to K^{*0}(\to K^+\pi^-)e^-\overline{\nu}_e)e^+\nu_e$ branching fraction is taken from the PDG [80]. Before applying the veto cut, C_{SL} is as high as $220 \pm 70\%$, while after the veto cut is applied it goes down to $5.6 \pm 1.1\%$. This is in good agreement with what is found in the mass fit to data (see section 3.6.5), where the semileptonic background is expected to account for some sizable fraction of what is quoted as combinatorial background in the fit.
Since it is very hard to disentangle the semileptonic from the combinatorial backgrounds, these two backgrounds are modeled together by a single shape in the mass and angular fits (see sections 3.6 and 4.3.2).

3.3.2 $B^0 \to K^{*0}\gamma(\to e^+e^-)$ background

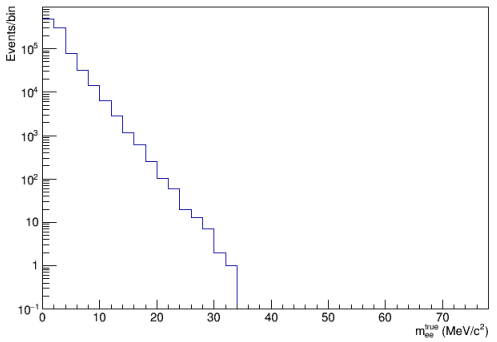
The radiative decay $B^0 \to K^{*0}\gamma$, where the real photon converts into an e^+e^- pair in the material of the detector is an important background to the $B^0 \to K^{*0}e^+e^-$ decay. Indeed, it has the exact same final state and its branching fraction is two orders of magnitude higher than the $B^0 \to K^{*0}e^+e^-$ decay. Nevertheless, since the dielectron pair comes from a real photon, it only acquires some mass via the momentum transfer to the nucleus of the material. For this reason, the dielectron mass of $B^0 \to K^{*0}\gamma(\to e^+e^-)$ events is usually very small. Therefore, removing events below 10 MeV/c^2 which is required by the q^2 range of the analysis (see section 1.4.2) already removes most of the converted photons. Nonetheless, without a specific veto, the contamination to $B^0 \to K^{*0}e^+e^-$ reaches $\sim 25\%$ in the analysis range (see figure 3.13).

Corrections to the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ MC

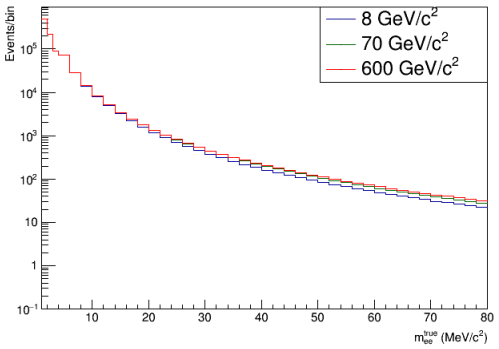
The full LHCb MC uses GEANT4 to simulate the interaction of particles with the detector. For computational reasons, the formula modeling the conversion of photons into an e^+e^- pair in the material of the detector is simplified. This leads to an underestimation of converted photons for high dielectron masses, as shown in figure 3.10.

In order to match the distribution of m_{ee}^{true} in the MC to the theoretical distribution [84], one can apply weights. However, after applying the cut $|\cos\theta_\ell| < 0.8$, there are very few events above 20 MeV/c^2 and none above 40 MeV/c^2 in the MC (see figure 3.10a), which would lead to very large or even infinite weights. To deal with this issue, a new sample (denoted as rescaled MC) is generated using the initial MC. In the rescaled MC, the angle between the dielectron pair is increased by a factor four in order to increase m_{ee}^{true}. The two samples are then merged together in the following way: for $m_{ee}^{\text{true}} < 10$ MeV/c^2, the events are taken from the unrescaled sample, while for m_{ee}^{true} rescaled above 10 MeV/c^2 the events are taken from the rescaled sample. After the rescaling and reweighting procedure, the obtained distribution at generator level in the MC is shown in figure 3.10c.

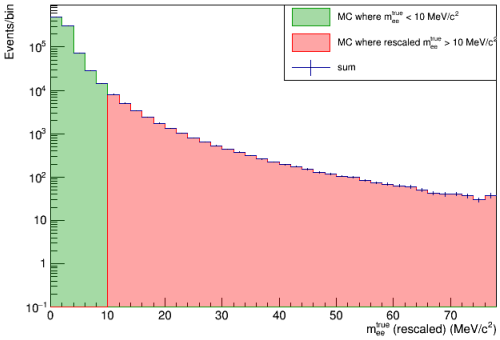
A similar procedure is then applied at reconstructed level. Some special care is needed when rescaling at reconstructed level. In particular one has to propagate the rescaling of the angles and momenta to the reconstructed helicity angle θ_l and the reconstructed dielectron mass, which require some additional corrections. The detailed procedure of rescaling and reweighting the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ both at generator and reconstructed level is explained in [85]. The dielectron mass distribution at reconstructed level obtained after all the weights and corrections is shown in figure 3.11



(a)



(b)



(c)

Figure 3.10: Dielectron mass at generator level for $|\cos\theta_\ell| < 0.8$ in the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ MC (3.10a) and the theoretical distribution for different values of the photon energy (3.10b). After applying the rescaling and reweighting procedure, the distribution in the MC matches the theory [84] (3.10c). The MC where $m_{ee}^{\text{true}} < 10\,\text{MeV}/c^2$ is shown in green, the MC where the rescaled $m_{ee}^{\text{true}} > 10\,\text{MeV}/c^2$ is shown in pink while the blue dots indicate their sum.

Veto cut and estimation of the remaining contamination

Now that the MC has the correct distribution, one can study how to veto converted photons and estimate the remaining contamination to $B^0 \to K^{*0}e^+e^-$.

Most converted photons are produced when the photon interacts with the nucleus of the material of the $VELO$. According to the MC, $\sim 70\%$ of the conversions happen in the RF foil, while the remaining $\sim 30\%$ happen in the modules (see figure 3.12). Thus, one can veto converted photons by rejecting events where the dielectron origin vertex position is compatible with the region where there is material in the detector. Using a precise mapping of the VELO one defines the distance

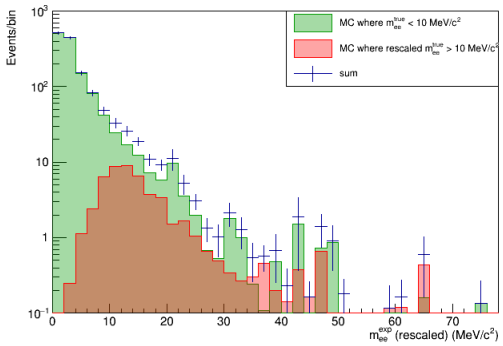


Figure 3.11: Dielectron mass at reconstructed level of the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ MC, after all weights and corrections, including the cut on the (rescaled) helicity angle $|\cos\theta_\ell| < 0.8$. In green the contribution from events with $m_{ee}^{\text{true}} < 10\,\text{MeV}/c^2$, in pink the contribution from rescaled events with a rescaled $m_{ee}^{\text{true}} > 10\,\text{MeV}/c^2$, in blue their sum.

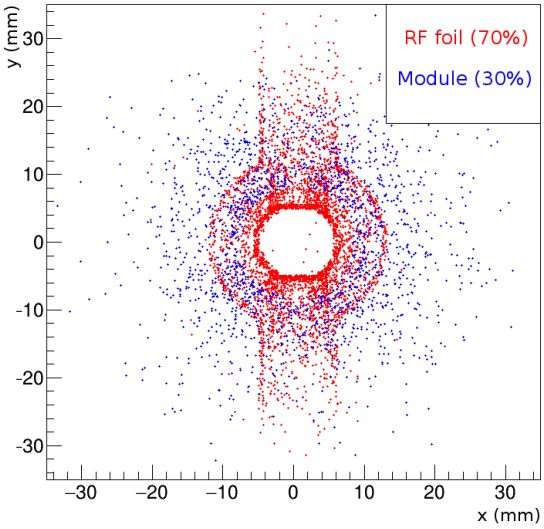


Figure 3.12: Conversion point in the (x, y) plane of $B^0 \to K^{*0}\gamma(\to e^+e^-)$ simulated events converting in the VELO. In red (blue) the conversion happened in the RF foil (modules) of the VELO.

to the nearest VELO modules in uncertainty space as [86]

$$\sigma_{\text{mod}} = \sqrt{\left(\frac{\delta_x}{\sigma_x}\right)^2 + \left(\frac{\delta_y}{\sigma_y}\right)^2 + \left(\frac{\delta_z}{\sigma_z}\right)^2} \quad , \tag{3.6}$$

where $\delta_{x,y,z}$ are the distances (in real space) between the reconstructed di-electron vertex and the nearest point in the VELO modules, while $\sigma_{x,y,z}$ are the

uncertainties of the vertex position. Since the uncertainty on the z coordinate of the dielectron vertex is larger than the typical distance to the foil for vertices in the foil region, the distance to the nearest point in the foil has little discrimination power and using it results in lowering the signal efficiency (see [85]). Thus, only the distance to the nearest $VELO$ module is used.
To find the optimal cut, the converted photons contamination to $B^0 \to K^{*0}e^+e^-$ is estimated for different cut values. First, a simultaneous fit is performed to data with no cut on σ_{mod}, in the ranges $m_{ee}^{B^0} \in [0, 10]\,\text{MeV}/c^2$ dominated by $B^0 \to K^{*0}\gamma(\to e^+e^-)$ events and $m_{ee}^{B^0} \in [10, 500]\,\text{MeV}/c^2$ dominated by $B^0 \to K^{*0}e^+e^-$ events. The details of the fit are given in section 3.6. For the bin $[0, 10]\,\text{MeV}/c^2$ the $B^0 \to K^{*0}e^+e^-$ leakage is computed using the yield returned by the fit in the bin $[10, 500]\,\text{MeV}/c^2$ constrained by the MC,

$$N_{ee}^{\text{data},t}(0, 10) = \frac{N_{ee}^{\text{MC},t}(0, 10)}{N_{ee}^{\text{MC},t}(10, 500)} \cdot N_{ee}^{\text{data},t}(10, 500) \qquad , \tag{3.7}$$

where $t = \{\text{L0L, L0I}\}$ denotes the trigger category. Similarly, the leakage of $B^0 \to K^{*0}\gamma(\to e^+e^-)$ in the bin $[10, 500]\,\text{MeV}/c^2$ is defined as

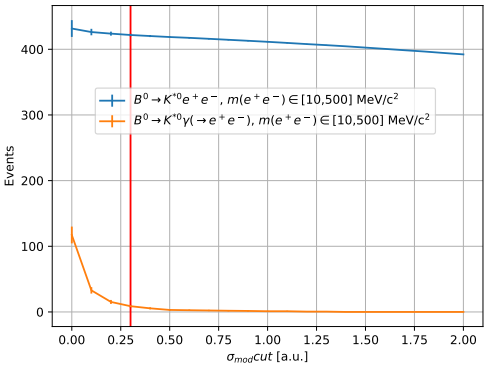
$$N_{\gamma ee}^{\text{data},t}(10, 500) = \frac{N_{\gamma ee}^{\text{MC},t}(10, 500)}{N_{\gamma ee}^{\text{MC},t}(0, 10)} \cdot N_{\gamma ee}^{\text{data},t}(0, 10) \qquad . \tag{3.8}$$

Then, for each σ_{mod} cut, the efficiency of the cut is estimated from simulation. The estimated $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}\gamma(\to e^+e^-)$ yields are then computed by the product of the yield with no cut and the efficiency of the cut.

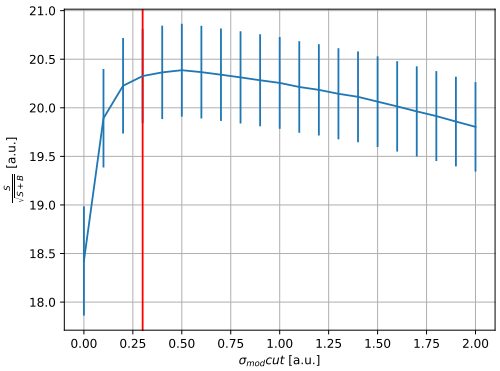
As the veto cut is tightened, the leakage of converted photons decreases, but the $B^0 \to K^{*0}e^+e^-$ efficiency is also slightly decreasing (see figure 3.13a). Thus $N_{ee}/\sqrt{N_{ee} + N_{\gamma ee}}$ is computed for each σ_{mod} cut value and reported in figure 3.13b. Knowing that the veto cut mostly removes $B^0 \to K^{*0}e^+e^-$ events at very low m_{ee} (because they have a larger vertex uncertainty due to their smaller dielectron angle) which are the most sensitive to the photon polarization, the optimal point is taken at the start of the plateau of $N_{ee}/\sqrt{N_{ee} + N_{\gamma ee}}$, at $\sigma_{\text{mod}} > 0.3$. After this veto cut has been applied, the converted photons contamination to $B^0 \to K^{*0}e^+e^-$ is as low as ~ 2 % (see figure 3.13c) while the signal efficiency of this cut is above 99%.

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

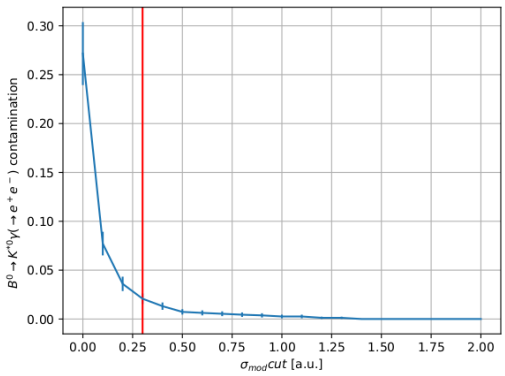
Another specific background for the $B^0 \to K^{*0}e^+e^-$ decay are the $B^0 \to K^{*0}\eta$ and $B^0 \to K^{*0}\pi^0$ decays where the η (π^0) decays to $e^+e^-\gamma$. There are two ways such Dalitz decays can end up being selected as $B^0 \to K^{*0}e^+e^-$: either the photon has been missed and was soft enough such that $m(K^+\pi^-e^+e^-)$ still peaks at the B^0



(a)



(b)



(c)

Figure 3.13: Top: Yields of $B^0 \to K^{*0} e^+ e^-$ (blue) and $B^0 \to K^{*0}\gamma(\to e^+e^-)$ (orange) in the bin $[10, 500]\,\mathrm{MeV}/c^2$ as a function of the cut on σ_{mod}. The red bar shows the chosen veto cut. Bottom left: $N_{ee}/\sqrt{N_{ee} + N_{\gamma ee}}$ as a function of the veto cut. Bottom right: Contamination of $B^0 \to K^{*0}\gamma(\to e^+e^-)$ to $B^0 \to K^{*0} e^+ e^-$ in the bin $[10, 500]\,\mathrm{MeV}/c^2$ as function of the veto cut.

mass, or the photon has been recovered and wrongly added as a bremsstrahlung photon. The branching fractions of these two decays are quite small with respect to the signal $B^0 \to K^{*0} e^+ e^-$ one (about two orders of magnitude lower for the π^0 and one order of magnitude for the η) but almost the whole phase space is located in the q^2 region of the analysis. Thus, the contamination of these backgrounds has to be carefully evaluated. In the case of the $B^0 \to K^{*0}\pi^0(\to \gamma e^+ e^-)$ decay, since the mass of the π^0 is smaller than the mass of the η, the opening angle between the dielectron and the bachelor photon is smaller. Therefore, it is more likely to recover the bachelor photon as a bremsstrahlung photon. This explains the fact

Signal contamination	R1		R2	
	L0L	L0I	L0L	L0I
C_{π^0}	$1.39 \pm 0.39\%$	$3.32 \pm 1.45\%$	$1.37 \pm 0.38\%$	$3.95 \pm 1.06\%$
C_{η}	$6.67 \pm 1.45\%$	$9.60 \pm 2.15\%$	$7.10 \pm 1.54\%$	$11.32 \pm 2.38\%$

Table 3.1: $B^0 \to K^{*0}\pi^0(\to \gamma e^+e^-)$ and $B^0 \to K^{*0}\eta(\to \gamma e^+e^-)$ contamination with respect to the signal $B^0 \to K^{*0}e^+e^-$ decay, estimated from simulation.

that the mass shape of the π^0 decay is more peaking than the one of the η decay, as shown in figure 3.36. The contamination of the $B^0 \to K^{*0}\pi^0(\to \gamma e^+e^-)$ after the full selection is applied is estimated by

$$C_{\pi^0} = \frac{\mathcal{B}\left(B^0 \to K^{*0}\pi^0\left(\to e^+e^-\gamma\right)\right) \times \varepsilon\left(B^0 \to K^{*0}\pi^0\left(\to e^+e^-\gamma\right)\right)}{\int_{q^2_{true,min}}^{q^2_{true,max}} \frac{d\mathcal{B}(B^0 \to K^{*0}e^+e^-)}{dq^2} dq^2 \times \varepsilon\left(B^0 \to K^{*0}e^+e^-\right)\left(q^2_{true,min}, q^2_{true,max}\right)} \quad , \tag{3.9}$$

where the efficiencies are obtained from simulation. The integrated differential branching fraction of $B^0 \to K^{*0}e^+e^-$ is computed using `Flavio` [12] and yields $3.89 \pm 0.90 \times 10^{-7}$. A wide range of $q^2_{true,min} = 0.000009\,\mathrm{GeV}^2/c^4$ and $q^2_{true,max} = 0.64\,\mathrm{GeV}^2/c^4$ is chosen in order to take into account possible leakage at the border of the q^2 range of the analysis. The $B^0 \to K^{*0}\pi^0\left(\to e^+e^-\gamma\right)$ branching fraction is taken from the PDG [80]. The $B^0 \to K^{*0}\pi^0(\to \gamma e^+e^-)$ contamination estimations for the different years and trigger categories are given in the first row of table 3.1.

For the $B^0 \to K^{*0}\eta(\to e^+e^-\gamma)$ decay, a similar procedure is applied. However, in the software used for the simulation of the B meson decay (EvtGen), there is no model for the $\eta \to e^+e^-\gamma$ decay. The MC is thus generated without any physics model for the η decay except for momentum conservation, resulting in a drastically wrong dielectron mass distribution with respect to the theoretical one [87,88] as shown in figure 3.14. To properly evaluate the contamination of this decay, the MC is reweighted at generator level such that the m^{true}_{ee} distribution matches the theoretical one. To avoid very large weights at low dielectron mass, an additional sample of $B^0 \to K^{*0}\eta(\to e^+e^-\gamma)$ is generated requiring $m^{\mathrm{true}}_{ee} < 60\,\mathrm{MeV}/c^2$ and added to the nominal sample. The reweighted m^{true}_{ee} distribution is shown in figure 3.14. The contamination from $B^0 \to K^{*0}\eta(\to e^+e^-\gamma)$ is then computed in a similar way as in eq. (3.9), but using reweigthed MC to compute the efficiency. The $B^0 \to K^{*0}\eta(\to \gamma e^+e^-)$ contamination estimations for the different years and trigger categories are given in the second row of table 3.1. Even if the reconstruction efficiency for the η decay is lower than the π^0 decay, its contamination is higher. This is mainly due to two effects. First, the η decay

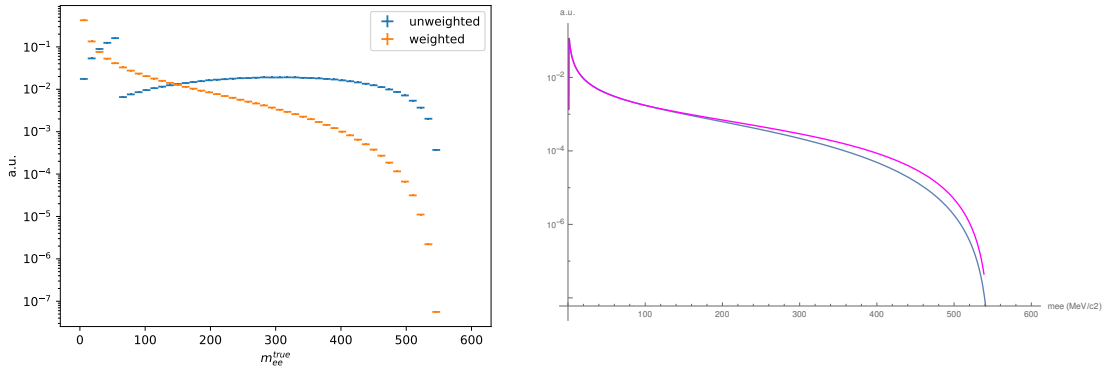


Figure 3.14: True dielectron mass distribution of the MC $B^0 \to K^{*0}\eta(\to \gamma e^+e^-)$ at generator level (left), where an additional sample with a generator level cut requiring $m_{ee}^{\text{true}} < 60\,\text{MeV}/c^2$ is added to the nominal sample. Due to the lack of physics model for the η decay in EvtGen, the m_{ee}^{true} distribution (left, blue points) is very different from the theoretical distribution (right). On the right plot, the pink curve takes into account the ρ^0 contribution, while the light blue curve does not. After reweighting (left, orange points), the true dielectron mass distribution of the MC at generator level matches the theory.

branching fraction is larger than the π^0 one by a factor ~ 3. Moreover, due to the larger mass of the η, the peak of its q^2 distribution is shifted towards higher dielectron masses. Thus, a fraction about twice as large of the η decay ends up in the q^2 bin of the analysis. Nonetheless, it is less dangerous than the π^0 background for the angular analysis. Indeed, due to its broader mass shape, it is easier to disentangle from the signal. On top of that, less than half of it is within the angular fit window $[5000, 5400]\,\text{MeV}/c^2$ (see section 4.1). The angular shape of this background is modeled (see section 4.3.4) and included in the angular fit to data.

3.3.4 $B^0 \to K^{*0}\eta(\to \gamma\gamma)$ and $B^0 \to K^{*0}\pi^0(\to \gamma\gamma)$ backgrounds

In the gamma-q^2 bin ($m_{ee}^{B^0} \in [0, 10]\,\text{MeV}/c^2$), $B^0 \to K^{*0}\pi^0(\to \gamma(\to e^+e^-)\gamma)$ and $B^0 \to K^{*0}\eta(\to \gamma(\to e^+e^-)\gamma)$ decays, where one of the photons converts to an e^+e^- pair, can be a source of background for the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ decay. Similarly to Dalitz decays (see section 3.3.3), due to the lower mass of the π^0 with respect to the η, it is much more likely to recover the second unconverted photon as a bremsstrahlung photon, as can be shown by the mass distributions in simulation in figure 3.15.

The efficiencies of these backgrounds are extracted from simulation. The rel-

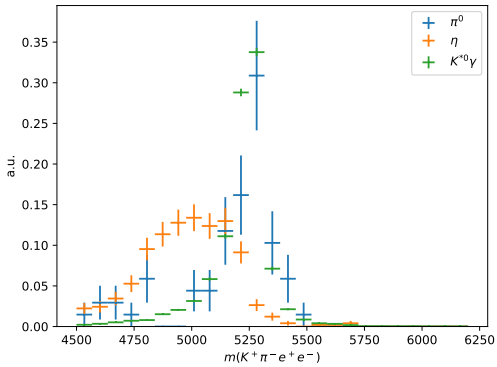


Figure 3.15: Distributions of $m(K^+\pi^-e^+e^-)$ for $B^0 \to K^{*0}\gamma(\to e^+e^-)$ (green), $B^0 \to K^{*0}\eta(\to \gamma(\to e^+e^-)\gamma)$ (orange) and $B^0 \to K^{*0}\pi^0(\to \gamma(\to e^+e^-)\gamma)$ (blue) MCs, for all Runs and trigger categories.

ative efficiencies with respect to $B^0 \to K^{*0}\gamma(\to e^+e^-)$ are ~ 1 for the π^0 decay while ~ 0.5 for the η decay. The contamination of these backgrounds with respect to the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ decay in the gamma q^2 bin are then extracted through the ratio of branching fractions (taken from the PDG [80]) and efficiencies. Due to the limited statistics available in simulation and since no difference is expected between the two runs, R1 and R2 simulated data are merged together. The contamination with respect to the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ decay yields $5.6 \pm 1.6\%$ ($11.9 \pm 3.4\%$) for the π^0 decay and $5.1 \pm 0.6\%$ ($8.7 \pm 1.2\%$) for the η decay for the L0L (L0I) trigger categories respectively. While this background is negligible in the very-low-q^2 bin, its mass and angular shapes are modeled in the gamma-q^2 bin and included in the mass and angular fits (see sections 3.6 and 4.4.3).

3.3.5 $B_s^0 \to \phi e^+e^-$ and $B^+ \to K^+e^+e^-$ backgrounds

Other species of B meson can also pollute the $B^0 \to K^{*0}e^+e^-$ candidates. This is the case of the $B_s^0 \to \phi(\to K^+K^-)e^+e^-$. If one of the kaon coming from the decay of the ϕ meson is misidentified as a pion, the ϕ can be reconstructed as a K^{*0}, with a relatively low $m(K^+\pi^-)$, but still large enough in some cases to pass the selection. Indeed, when looking at the mass of the $(K^+\pi^-)$ system in data, where the kaon mass hypothesis is assigned to the pion (see figure 3.16) a peaking structure is spotted around the ϕ mass (1019 MeV/c^2). To remove these events, a veto cut $m(K^+(\pi^- \to K^-)) > 1040\,\text{MeV}/c^2$ is applied. The efficiency of this cut on the signal $B^0 \to K^{*0}e^+e^-$ estimated from MC is above 99 %.

In addition, another species of B meson can also contaminate the signal, namely $B^+ \to K^+e^+e^-$. Indeed, this decay can mimic the signal if a random pion from

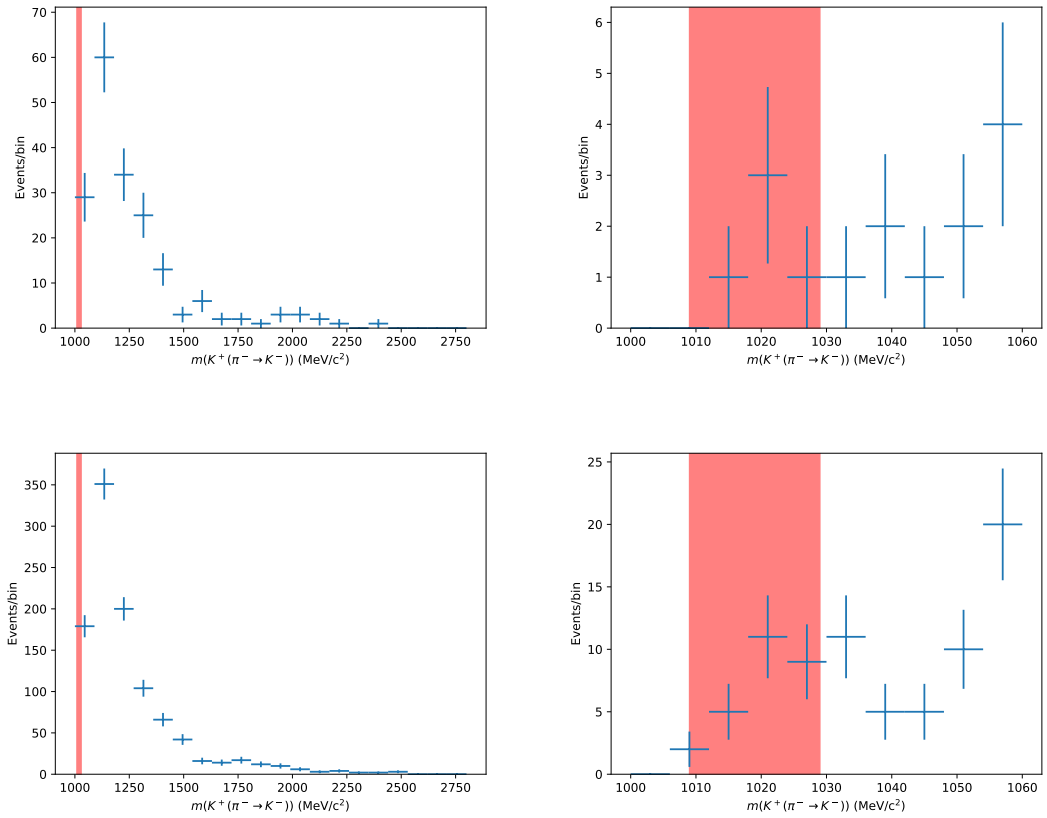


Figure 3.16: Mass of the $(K^+\pi^-)$ system in $B^0 \to K^{*0}e^+e^-$ data R1 (top) and R2 (bottom), where the kaon mass hypothesis is assigned to the pion. The right plot is a zoom near the ϕ peak at 1019 MeV/c^2.

the event is picked up and associated to the kaon or if the kaon is misidentified as a pion and a random kaon from the event is picked up. Since the dielectron pair does not come from a virtual photon, the $B^+ \to K^+e^+e^-$ decay does not have a photon pole at very low q^2. Therefore, this background is not expected to be present at very low q^2 and indeed the contamination from this decay is estimated to be below 0.03% from MC. In addition, the reconstructed $m(K^+\pi^-e^+e^-)$ of this decay peaks at a higher mass than the B^0 mass. This background is therefore neglected in the very-low-q^2 bin.

In the J/ψ-q^2 bin however, the background of the type $B^+ \to K^+J/\psi\,(\to e^+e^-)$ is present. As shown in figure 3.17, this background is peaking in the $m(K^+e^+e^-)$ mass where the dielectron mass is constrained to the J/ψ mass. To avoid any biases when computing corrections to the MC using the resonant mode (see section 3.2), a dedicated veto cut is applied to remove this background requiring that $m(K^+e^+e^-)$

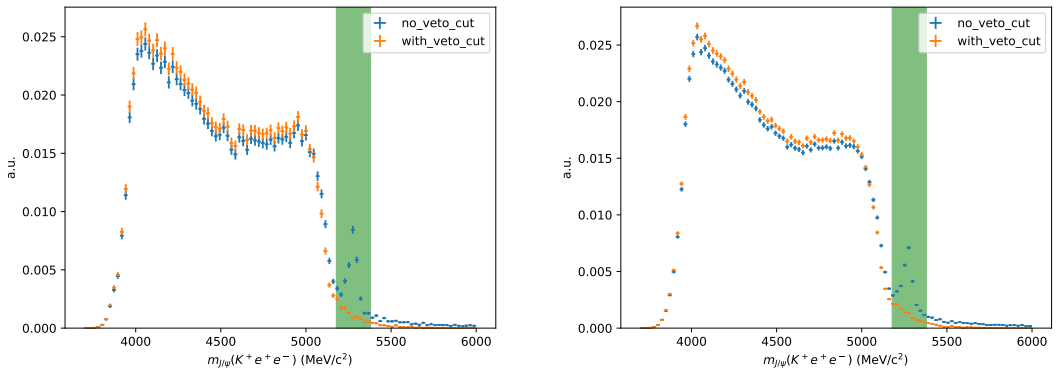


Figure 3.17: Distribution of the $m(K^+e^+e^-)$ mass where the dielectron mass is constrained to the J/ψ mass in $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ data for R1 (left) and R2 (right). The full selection is applied without the B^+ veto cut (blue) or with the veto cut (orange).

and $m((\pi \to K)e^+e^-)$ are lower than 5100 MeV/c^2. The efficiency of this cut on signal $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ is estimated from MC and is above 99 %.

3.3.6 $\Lambda_b^0 \to pK^-e^+e^-$ background

In the case of the mis-identification $\pi \to p$, or a double mis-identification $\pi \to K$ and $K \to p$, $\Lambda_b^0 \to pK^-e^+e^-$ events can pass the selection. Since the branching fraction of this decay as well as the structure of the pK spectrum are not well known, some assumptions have to be done to estimate the contamination of this background. First, it is assumed that the contamination in the $B^0 \to K^{*0}\gamma$ channel is the same as in the $B^0 \to K^{*0}e^+e^-$ channel at very low q^2, *i.e.*

$$\frac{\mathcal{B}(\Lambda_b^0 \to pK^-\gamma)\cdot\varepsilon(\Lambda_b^0 \to pK^-\gamma)}{\mathcal{B}(B^0 \to K^{*0}\gamma)\cdot\varepsilon(B^0 \to K^{*0}\gamma)} = \left[\frac{\mathcal{B}(\Lambda_b^0 \to pK^-e^+e^-)}{\mathcal{B}(B^0 \to K^{*0}e^+e^-)} \times \frac{\varepsilon(\Lambda_b^0 \to pK^-e^+e^-)}{\varepsilon(B^0 \to K^{*0}e^+e^-)}\right]\Bigg|_{m_{ee}\in[10,500]\,\mathrm{MeV}/c^2} . \tag{3.10}$$

The branching fraction of the decay $\Lambda_b^0 \to \Lambda^{*0}(\to pK^-)\gamma$ has been measured in [89] and is given by

$$\frac{f_d}{f_{\Lambda_b}}\frac{\mathcal{B}(B^0 \to K^{*0}\gamma)}{\mathcal{B}(\Lambda_b^0 \to \Lambda^{*0}(\to pK^-)\gamma)} = 3.348 \pm 0.071 \pm 0.154 \quad , \tag{3.11}$$

where Λ^{*0} denotes all resonances between 1400 and 2600 MeV/c^2 and f_d (f_{Λ_b}) is the B^0 (Λ_b^0) production fractions at LHCb. As shown in figure 3.18, the theoretical pK^- spectrum is expected to be dominated by the $\Lambda^*(1520)$ resonance. The $\Lambda_b^0 \to pK^-\gamma$ efficiency is computed using $\Lambda_b^0 \to \Lambda^*(1520)(\to pK^-)\gamma$ MC. The efficiency is then corrected to take into account the efficiency of the $m(K^+\pi^-) = m_{K^{*0}} \pm 100\,\text{MeV}/c^2$ cut performed in the selection, which is different for higher Λ^{*0} resonances. This effect has been studied in [35] and a correction factor of 0.575 ± 0.01 has to be applied to the efficiency. The contamination of $\Lambda_b^0 \to pK^-e^+e^-$ decays to $B^0 \to K^{*0}e^+e^-$ decays is estimated to be $0.81 \pm 0.07\%$ ($0.47 \pm 0.05\%$) for R1 (R2). The difference between R1 and R2 is attributed to the better training in R2 of the PID neural network used for `ProbNN` variables, which play a major role in the suppression of this background. Since the contamination is found to be below the percent level, this background is neglected in the mass and angular fits.

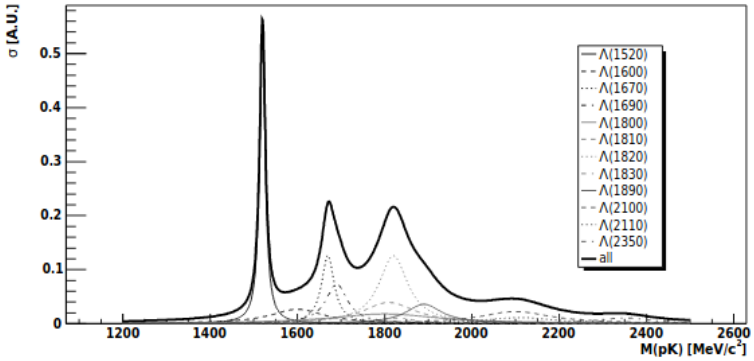


Figure 3.18: Theoretical pK spectrum in $\Lambda_b^0 \to \Lambda^{*0}(\to pK^-)\gamma$ decays (from [90]).

3.3.7 $\pi \to e$ mis-identification

The misidentification of a pion as an electron can lead to a contamination from semileptonic cascades of the type $B^0 \to D^-(K^{*0} \to (K^+\pi^-)\pi^-)e^+\nu$, $B^0 \to \overline{D}^0(\to K^+\pi^-)\pi^-e^+\nu$ or $B^0 \to D^{*-}(\to D^0(\to K^+\pi^-)\pi^-)e^+\nu$. In order to check for the presence of these types of backgrounds the $m(K^+(e^- \to \pi^-))$ and $m(K^+\pi^-(e^- \to \pi^-))$ are scanned in data after full selection. As shown in figures 3.19 and 3.20 no peak is found around the D^- / D^0 masses. Moreover, due to the missing neutrinos, this background is non peaking and, if present at all, would only pollute the lower mass sideband. For these reasons, these backgrounds are neglected.

In the case of a double $\pi \to e$ mis-identification, $B^0 \to K^{*0}\pi^+\pi^-$ decays can contaminate the $B^0 \to K^{*0}e^+e^-$ signal. The $m(e^+e^-)^{\pi\to e}$ spectrum of $B^0 \to$

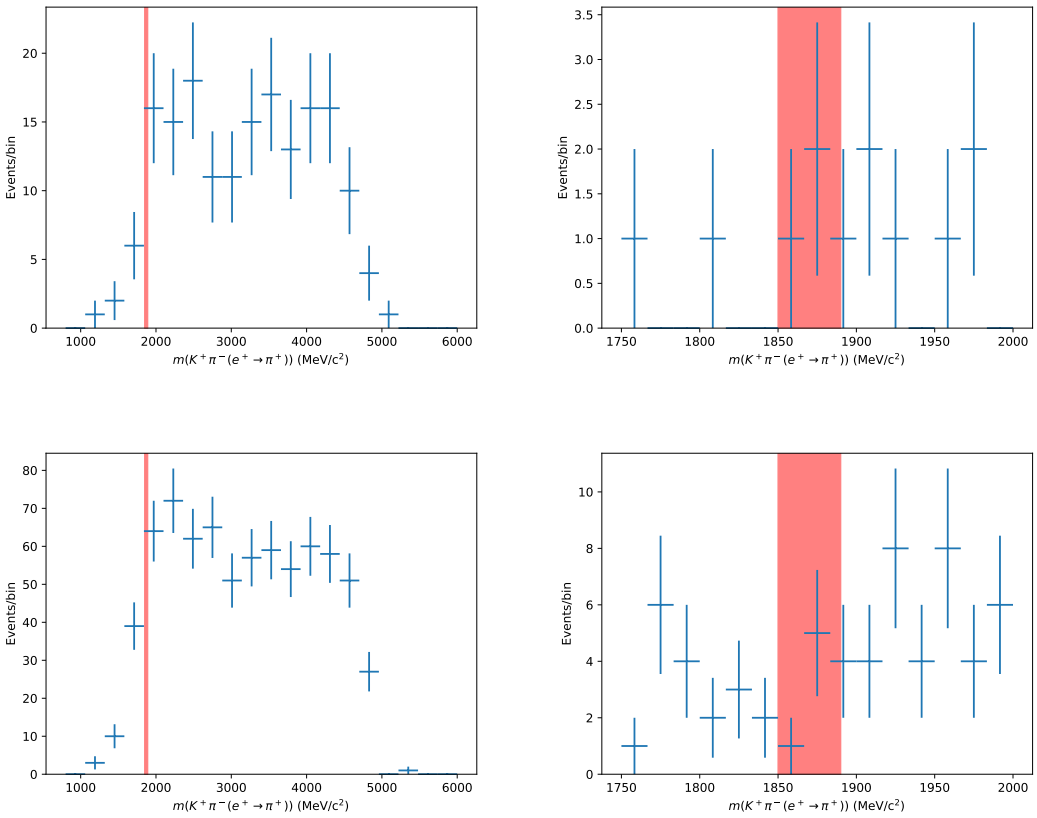


Figure 3.19: Distribution of $m(K^+\pi^-(e^-\to\pi^-))$ for R1 (top) and R2 (bottom) data after applying the full selection. The plot on the right is a zoom of the plot on the left. The red band represents the region around the D^-/D^0 mass. No peak is spotted.

$K^{*0}\pi^+\pi^-$ ${}_s\mathcal{P}$lotted data (from [91,92]) is shown in figure 3.21a. Even if the major part of the spectrum lies at higher $m(e^+e^-)^{\pi\to e}$, some events can still leak inside the q^2 range of the analysis.

Given the very good resolution of $B^0\to K^{*0}\pi^+\pi^-$ (see figure 3.21b) events with respect to $B^0\to K^{*0}e^+e^-$ events (largely due to the absence of bremsstrahlung emission), the fraction of $B^0\to K^{*0}\pi^+\pi^-$ events reconstructed as $B^0\to K^{*0}e^+e^-$ events can be largely enhanced by restricting to high m_{ee} and a tight cut around the reconstructed B^0 mass. The contribution of this background is therefore estimated by defining two regions:

- ee region: $m(K^+\pi^-e^+e^-)\in[5000,5400]\,\mathrm{MeV}/c^2$ (the mass range of the angular analysis, see section 4.1) and $\mathrm{m}_{ee}^{B^0}\in[10,500]\,\mathrm{MeV}/c^2$
- $\pi\pi$ region: $m(K^+\pi^-e^+e^-)\in m_{B^0}^{PDG}\pm40\,\mathrm{MeV}/c^2$ and $m_{ee}\in[500,1000]\,\mathrm{MeV}/c^2$

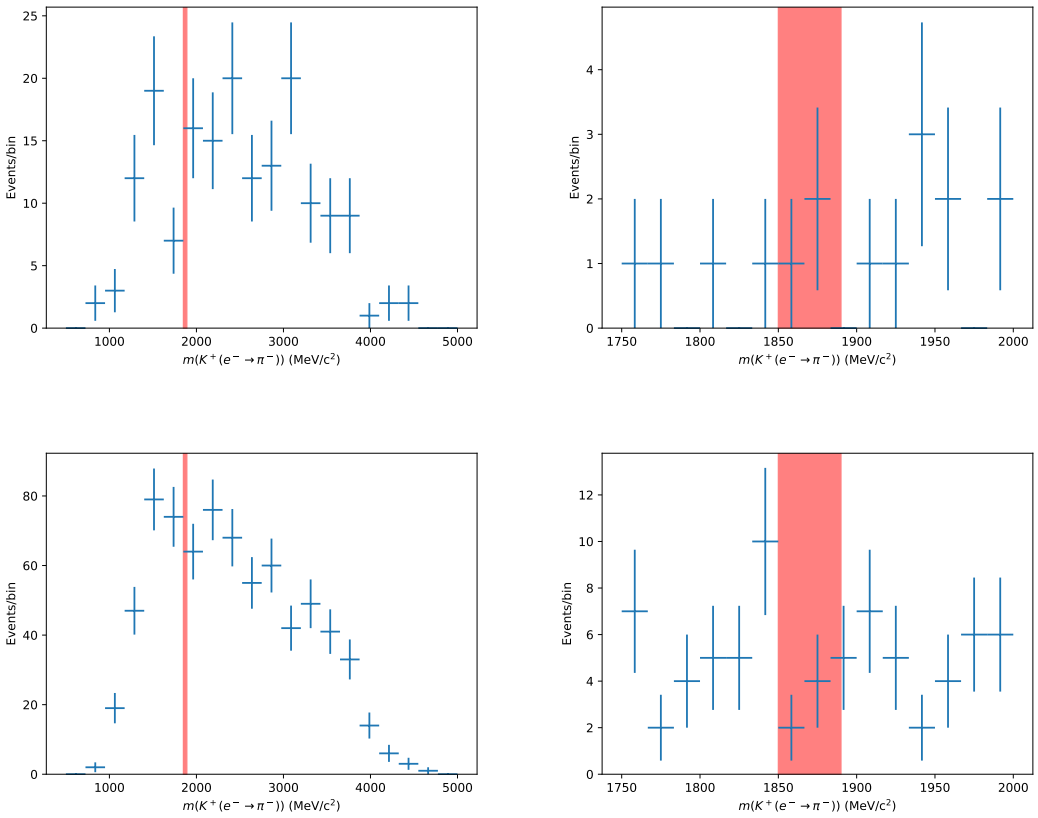
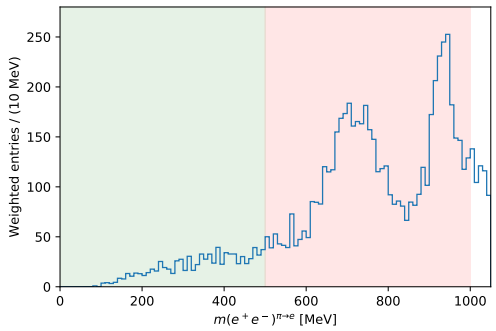


Figure 3.20: Distribution of $m(K^+(e^-\to\pi^-))$ for R1 (top) and R2 (bottom) data after applying the full selection. The plot on the right is a zoom of the plot on the left. The red band represents the region around the D^0 mass. No peak is spotted.

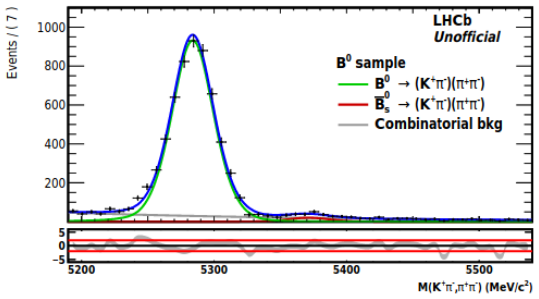
These two regions are defined on top of the full selection (see section 3.5) except the $\mathrm{m}_{ee}^{B^0}$ cut. The $B^0 \to K^{*0}e^+e^-$ data is fitted (see details of the fit procedure in section 3.6) both in the $\mathrm{m}_{ee}^{B^0} \in [10, 500]\,\mathrm{MeV}/c^2$ and $\mathrm{m}_{ee}^{B^0} \in [500, 1000]\,\mathrm{MeV}/c^2$ regions and shown in figure 3.22. In each of these two regions, there are three types of populations: $B^0 \to K^{*0}e^+e^-$ events, $B^0 \to K^{*0}\pi^+\pi^-$ events and other types of backgrounds given by

$$\begin{cases} N_{tot}^{data}(ee \text{ region}) & = N_{ee}^{data}(ee \text{ region}) + N_{\pi\pi}^{data}(ee \text{ region}) + N_{bkg}^{data}(ee \text{ region}) \\ N_{tot}^{data}(\pi\pi \text{ region}) & = N_{ee}^{data}(\pi\pi \text{ region}) + N_{\pi\pi}^{data}(\pi\pi \text{ region}) + N_{bkg}^{data}(\pi\pi \text{ region}) \end{cases}, \tag{3.12}$$

where N_{bkg}^{data} in each region is computed by integrating the sum of all background PDFs in the angular window $m(K^+\pi^-e^+e^-) \in [5000, 5400]\,\mathrm{MeV}/c^2$. In addition,



(a)



(b)

Figure 3.21: Left: $m(e^+e^-)^{\pi\to e}$ distribution of ${}_s\mathcal{P}lotted$ $B^0 \to K^{*0}\pi^+\pi^-$ events. In particular, the $m(K^+\pi^-)$ cut of the analysis (see section 3.5) is applied. The green (pink) region corresponds to the ee ($\pi\pi$) region. Right: reconstructed B^0 mass of $B^0 \to K^{*0}\pi^+\pi^-$ data, from [91, 92].

the number of $B^0 \to K^{*0}e^+e^-$ ($B^0 \to K^{*0}\pi^+\pi^-$) in the ee and $\pi\pi$ region can be related through

$$\begin{cases} N_{ee}^{data}(\pi\pi \text{ region}) &= \frac{N_{ee}^{MC}(\pi\pi \text{ region})}{N_{ee}^{MC}(ee \text{ region})} N_{ee}^{data}(ee \text{ region}) \\ N_{\pi\pi}^{data}(\pi\pi \text{ region}) &= \frac{N_{\pi\pi}^{s\mathcal{P}lotted}(\pi\pi \text{ region})}{N_{\pi\pi}^{s\mathcal{P}lotted}(ee \text{ region})} N_{\pi\pi}^{data}(ee \text{ region}) \end{cases} , \tag{3.13}$$

where N_{ee}^{MC} are the yields in the $B^0 \to K^{*0}e^+e^-$ MC and $N_{\pi\pi}^{s\mathcal{P}lotted}$ are the yields obtained with a sample of ${}_s\mathcal{P}lotted$ $B^0 \to K^{*0}\pi^+\pi^-$ events (from [91, 92]). Combining eq. (3.13) with eq. (3.12) the system of four equations with four unknowns is solvable. In particular, the quantities of interest $N_{\pi\pi}^{data}(ee \text{ region})$ and $N_{ee}^{data}(ee \text{ region})$ can be computed to extract the contamination of $B^0 \to K^{*0}\pi^+\pi^-$ to $B^0 \to K^{*0}e^+e^-$ events in the angular analysis region given by

$$C_{\pi\pi} = \frac{N_{\pi\pi}^{data}(ee \text{ region})}{N_{ee}^{data}(ee \text{ region})} = 0.88 \pm 0.32\% \quad . \tag{3.14}$$

The uncertainty ($\sim 10\%$) of the SM $B^0 \to K^{*0}e^+e^-$ branching fraction predictions in these low q^2 regions is taken into account. In addition, since the LHCb MC is generated with SM values, the possibility of NP is also considered. The ratio of branching fractions varies by $\sim 5\%$ at most (*e.g.* in the extreme case where $\mathcal{C}_9 = -1$). Nevertheless, the uncertainty related to possible NP effects is carefully taken into account.
The contamination $C_{\pi\pi}$ is found to be negligible and therefore this background is not modeled in the mass and angular fits.

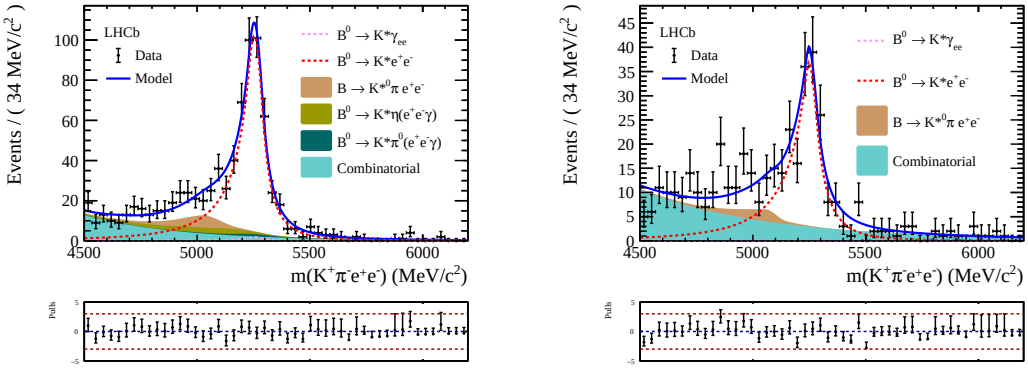


Figure 3.22: Fits to the $m(K^+\pi^-e^+e^-)$ invariant mass in the $\mathrm{m}_{ee}^{B^0} \in [10, 500]\,\mathrm{MeV}/c^2$ (left) and $\mathrm{m}_{ee}^{B^0} \in [500, 1000]\,\mathrm{MeV}/c^2$ (right) regions, corresponding to the ee and $\pi\pi$ regions respectively. The signal $B^0 \to K^{*0}e^+e^-$ is shown in red dotted line, the partially reconstructed background in brown and the combinatorial background in cyan. Two additional components (in light and dark green) are also present in the $\mathrm{m}_{ee}^{B^0} \in [10, 500]\,\mathrm{MeV}/c^2$ (see section 3.6) namely $B^0 \to K^{*0}\eta(\to \gamma e^+e^-)$ and $B^0 \to K^{*0}\pi^0(\to \gamma e^+e^-)$.

3.3.8 K^{*0} and $\overline{K}^{*0}$ mis-identification and multiple candidates

The case of a K^{*0} misidentified as a $\overline{K}^{*0}$, *i.e.* a double misidentification of the kaon as a pion and the pion as a kaon can lead to important biases to angular observables. Indeed, the measured flavor of the K^{*0} defines the measured flavor of the B^0. Since the definition of the angles has a sign flip depending on the flavor of the B^0 meson, a flavor flip leads to a sign flip in the angles. This introduces a bias in the measurement of F_{L}, A_T^{Re} and A_T^{Im}, while $A_T^{(2)}$ is unaffected.
To evaluate the K^{*0} misidentification rate, the PID maps discussed in section 3.2.1 are used. The K^{*0} misidentification rate is found to be $0.74 \pm 0.01\%$. However, most of the double misidentified events are multiple candidates, *i.e.* the reconstruction algorithms have found multiple ways of forming a $B^0 \to K^{*0}e^+e^-$ candidate with a single event (in this precise case, there is typically one $B^0 \to K^{*0}e^+e^-$ candidate where the pion and the kaon tracks are associated to the pion and the kaon candidates respectively and another $B^0 \to K^{*0}e^+e^-$ candidate where the pion and the kaon are swapped). After the full selection is applied, a very small amount ($\mathcal{O}(1\%)$) multiple candidates remain. If a candidate is chosen randomly among the multiple candidates, the double kaon-pion misidentification rate goes down to $0.38 \pm 0.01\%$. To reduce it even further, instead of a random choice, the candidate with the highest `K_ProbNNk` $\times$ `Pi_ProbNNpi` is kept, lowering the double misidentification rate to the negligible amount of $0.17 \pm 0.01\%$.

3.4 Combinatorial and partially reconstructed backgrounds

Now that all specific backgrounds have been treated, two types of background remain: combinatorial and partially reconstructed backgrounds. When the reconstruction algorithms look for $B^0 \to K^{*0}e^+e^-$ candidates among the many decay products produced in a proton-proton collision, it is common that a candidate is formed by combining random tracks coming from various decays. This type of background is called combinatorial background. In addition, it may also happen that a decay involving two electrons, and kaon and a pion plus some additional particle(s) X is reconstructed as $B^0 \to K^{*0}e^+e^-$ if the particle(s) X is (are) missed. This is called partially reconstructed background.
To lower the amount of partially reconstructed background, a specific veto cut (called HOP) is applied and explained in section 3.4.1. To reduce the amount of combinatorial background, a Multivariate Analysis (MVA) using a Boosted Decision Tree (BDT) is used. The samples used to train the BDT are detailed in section 3.4.2, the training itself is discussed in section 3.4.2, while the BDT cut optimization is explained in section 3.4.3.

3.4.1 Partially reconstructed background

To reduce the amount of partially reconstructed background (and, to a lesser extent, combinatorial background), a cut on the HOP mass is used [93]. The idea of HOP is to exploit the kinematic balance of the decay products of the B^0 meson along the plane transverse to the B^0 flight direction, as sketched in figure 3.23. The ratio of the transverse momentum of the hadronic part over the leptonic part is given by

$$\alpha_{\mathrm{HOP}} = \frac{p_T\left(K^{*0}\right)}{p_T\left(e^+e^-\right)} \quad . \tag{3.15}$$

If this ratio is different from unity, it means that some energy is missing in the final state. For signal $B^0 \to K^{*0}e^+e^-$ events, the missing energy is most likely lost through the emission of bremsstrahlung photons. Since bremsstrahlung photons are emitted in the same direction as the electron, the fraction of transverse momentum which is lost is the same as the longitudinal one. Therefore α_{HOP} can be used to recover the full momentum using

$$p_{corr}\left(e^+e^-\right) = \alpha_{\mathrm{HOP}} \cdot p\left(e^+e^-\right) \quad . \tag{3.16}$$

One can thus recompute the invariant mass of the B^0 candidate using the corrected momenta, $m_{\mathrm{HOP}}(K^+\pi^-e^+e^-)$.

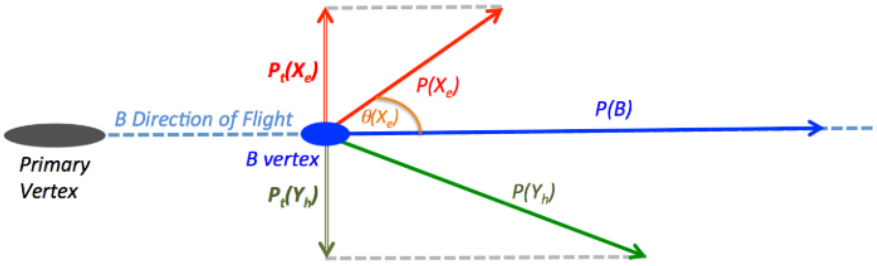


Figure 3.23: Sketch of the kinematics of a $B \to Y_h X_e$ decay, where Y_h is the hadronic part (here the K^{*0}) and X_e the leptonic part (here the dielectron pair).

In the case of partially reconstructed background however, the missing particles have no reason to fly in the same direction as the electron. The same argument holds for combinatorial backgrounds. Thus, the $m_{\text{HOP}}(K^+\pi^-e^+e^-)$ is a good discriminating variable to reject partially reconstructed and combinatorial background. Still, the $m_{\text{HOP}}(K^+\pi^-e^+e^-)$ resolution depends on the B^0 lifetime. Hence, three types of cuts are considered:

- A simple cut $m_{\text{HOP}}(K^+\pi^-e^+e^-) > a_{\text{HOP}}$
- A double-sided cut $b_{\text{HOP}} > m_{\text{HOP}}(K^+\pi^-e^+e^-) > a_{\text{HOP}}$
- A diagonal cut $m_{\text{HOP}}(K^+\pi^-e^+e^-) > a_{\text{HOP}} + b_{\text{HOP}} \cdot \log\left(B^0\chi^2_{FD}\right)$

The three types of cuts are illustrated in figure 3.24. The lower red line illustrates the simple cut, the upper red line combined with the lower one illustrates the double-sided cut and the green line illustrates the diagonal cut.

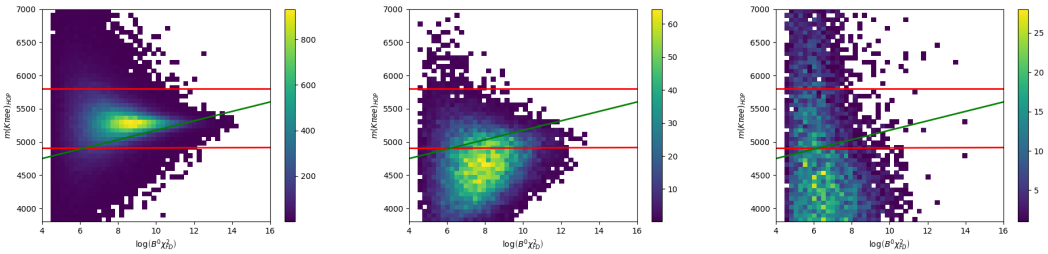


Figure 3.24: $m_{\text{HOP}}(K^+\pi^-e^+e^-)$ versus $\log\left(B^0\chi^2_{FD}\right)$ distributions for $B^0 \to K^{*0}e^+e^-$ MC (left), $B^+ \to K_1(1270)(\to K^+\pi^-X)e^+e^-$ MC (middle) and $B^0 \to K^{*0}e^+\mu^-$ data (right), which is a good proxy for combinatorial background (see section 4.3.2). The red lines illustrate the simple and double sided cuts, while the green line illustrates the diagonal cut.

The double-sided cut is more effective to reject combinatorial background, but is useless to reject additional partially reconstructed background and significantly reduces the signal efficiency. The diagonal cut yields similar performances as the simple cut. Thus, for simplicity, the simple cut on $m_{\mathrm{HOP}}(K^+\pi^-e^+e^-)$ is chosen as the partially reconstructed background veto. In order not to sculpt the signal mass shape, a rather loose cut at $m_{\mathrm{HOP}}(K^+\pi^-e^+e^-) > 4900\,\mathrm{MeV}/c^2$ is chosen. The efficiency of this cut on $B^+ \to K_1(1270)(\to K^+\pi^-X)e^+e^-$ MC [4] is lower than 30% while on signal $B^0 \to K^{*0}e^+e^-$ MC it is higher than 90 %.

3.4.2 Combinatorial background

Since the distributions of kinematics variables used to train the BDT are different between R1 and R2 due to the center-of-mass energy difference, two independent BDTs are trained separately for R1 and R2. The two BDTs are trained using the AdaBoostClasifier class from the Scikit-learn package [94]. To avoid biases, a k-folding approach using 10 folds is used: 10 independent BDTs are trained using 9/10 of the samples and are applied to the remaining 1/10. For each fold, half of the events are used to train the BDT while the other half is used to test it.

Training samples and selection

Simulated $B^0 \to K^{*0}e^+e^-$ events are used for the signal training sample, while data from the upper mass sideband ($m(K^+\pi^-e^+e^-) > 5600\,\mathrm{MeV}/c^2$) is used as a background training sample. The full selection, including all the specific vetoes discussed in the previous sections except the HOP cut are applied to both samples. Due to limited statistics, for both samples the upper limit of the q^2 range is increased to $5\,\mathrm{GeV}^2/c^4$. In addition, the cut on $m(K^+\pi^-)$ is relaxed to $|m(K^+\pi^-) - m^{PDG}_{K^{*0}}| < 200\,\mathrm{MeV}/c^2$. On the background sample, only a loose cut `ProbNNe` > 0.05 is performed on the electron PID, while on the signal sample the full PID weights (as well as all corrections discussed in section 3.2) are applied. Again, to enlarge the available statistics, both trigger categories are merged together for the training. It is checked that loosening the selection for the training with respect to the full selection does not change the distributions of the variables used for the training of the BDT. In figures 3.25, 3.26, 3.27 and 3.28 the distributions of the training variables are shown with the full cut or the relaxed cut, for both the signal and background samples. No large difference is spotted.

[4] Here X denotes the missing particle, which is usually a pion. The MC used is a mixture of $K_1(1270) \to K^+\pi^-\pi^+$ decays as well as $K_1(1270) \to \omega(\to \pi^+\pi^-)K^+$, $K_1(1270) \to \rho(\pi^+\pi^-)K^+$ and $K_1(1270) \to K^{*0}(K^+\pi^-)\pi^+$. The fraction of each decay is generated according to the known branching fractions.

Particle	Variables
B^0	p_T, $\chi^2_{FD\_OWNPV}$, χ^2_{DTF}/ndf
K^{*0}	$\log \chi^2_{IP\_OWNPV}$
h	$\log \chi^2_{IP\_OWNPV}$
$e^{\pm}$	$min(\log \chi^2_{IP\_OWNPV,e^+}, \log \chi^2_{IP\_OWNPV,e^-})$
e^+e^-	$\log \chi^2_{IP\_OWNPV}$

Table 3.2: Variables used to train the BDT: p_T is the transverse momentum, $\chi^2_{FD\_OWNPV}$ is the flight distance with respect to the primary vertex divided by its uncertainty, χ^2_{DTF}/ndf is the χ^2 of a kinematic fit to the full decay chain divided by the number of degrees of freedom of the fit and $\chi^2_{IP\_OWNPV}$ is the difference of χ^2 if the primary vertex is computed with our without a given track. In the table, h refers to the hadrons, *i.e.* the pion or the kaon.

The lower mass sideband is not used for the training since it contains a non negligible fraction of partially reconstructed background. For R1 (R2), the size of the signal training sample is $\sim$ 55k ($\sim$ 130k) events while the size of the background training sample is $\sim$ 5k ($\sim$ 45k) events.

BDT training

The BDT uses 8 kinematics and decay topology variables, listed in table 3.2, which are chosen because their distribution is different in the signal and background samples. The distributions of these variables for both the signal and background samples, as well as the correlation between the variables are shown in appendix B.1.

The most discriminating variables are the χ^2 of the kinematic fit to the full decay chain, the p_T of the B^0 and the χ^2 of the impact parameter of the $e^{\pm}$. The output of the BDT for both the training and the test samples for a random fold is shown in figure 3.29. Similar plots for all folds are shown in appendix B.2. No over-training is spotted since the distribution of the test sample is similar to the training sample, but with different fluctuations.

The effect of the BDT cut on the angular distributions of the signal are also checked. As shown in figures 3.30 and 3.31 no large variation is spotted. This is essentially due to the $\sim$ 94% efficiency of the optimal BDT cut (see section 3.4.3) on $B^0 \to K^{*0}e^+e^-$ signal.

3.4.3 BDT optimization

Given the fact that the combinatorial background level is expected to be higher in the L0I trigger category than in the L0L category and that the BDT outputs from

R1 and R2 are not exactly the same (see figure 3.29) a different BDT cut is considered for each run and trigger category. The choice of the best set of {BDT R1 L0L cut, BDT R1 L0I cut, BDT R2 L0L cut, BDT R2 L0I cut } is determined through the maximization of $S/\sqrt{S+B}$, where S and B are estimations of the signal and background yields in the angular fit window $m(K^+\pi^-e^+e^-) \in [5000, 5400]\,\mathrm{MeV}/c^2$ (see section 4.1) after the full selection. For each set of cut, $S^{t,y}$ and $B^{t,y}$, where $t = \{L0L, L0I\}$ and $y = \{R1, R2\}$ are computed separately and then added together to compute the total S and B.
To estimate the signal yield for a given set of cuts, a simultaneous fit is performed in the two trigger categories to the data in the J/ψ-q^2 bin (see figure 3.40 of section 3.6.4) with the full selection, except the HOP and BDT cuts. Then it is rescaled by the ratio of branching fractions and efficiencies extracted from MC samples such that

$$S^{t,y} = N^{t,y}_{J/\psi} \cdot \frac{\mathcal{B}\left(B^0 \to K^{*0}e^+e^-\right) \cdot \varepsilon^{t,y}_{cut}\left(B^0 \to K^{*0}e^+e^-\right)}{\mathcal{B}\left(B^0 \to K^{*0}J/\psi(\to e^+e^-)\right) \cdot \varepsilon^{t,y}\left(B^0 \to K^{*0}J/\psi(\to e^+e^-)\right)} \quad , \tag{3.17}$$

where $N^{t,y}_{J/\psi}$ is the signal yield of the fit to $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ data, $\varepsilon^{t,y}_{cut}\left(B^0 \to K^{*0}e^+e^-\right)$ is the $B^0 \to K^{*0}e^+e^-$ efficiency (including the HOP and angular fit window cuts) of the full selection including a set of BDT cuts, $\varepsilon^{t,y}(B^0 \to K^{*0}J/\psi(\to e^+e^-))$ is the $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ efficiency of the full selection with no BDT and HOP cuts. The fit follows a similar strategy to the one discussed in section 3.6.4. The branching fractions are computed with `Flavio` v.2.0.0 [12]. To account for possible leakage, the $B^0 \to K^{*0}e^+e^-$ branching fraction is computed in a wide range $[3, 800]\,\mathrm{MeV}/c^2$ and yields $3.89 \pm 0.90 \times 10^{-7}$. The $B^0 \to K^{*0}e^+e^-$ efficiency is thus also computed with the same cut $[3, 800]\,\mathrm{MeV}/c^2$ on m^{true}_{ee}. The $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ branching fraction is taken from the PDG [80] and yields $1.27 \pm 0.05 \times 10^{-3}$.

To estimate the background yield, a simultaneous fit is performed in the gamma- and very-low-q^2 bins in the two trigger categories. The fit procedure follows the same strategy as the one explained in section 3.6. The background PDFs are then integrated in the angular fit window between 5000 and 5400 MeV/c^2 to extract $B^{t,y}$.
The S and B used to compute the figure of merit $S/\sqrt{S+B}$ are obtained by summing over all triggers and all years. As shown in figure 3.32 the set of cuts that maximizes $S/\sqrt{S+B}$ is {BDT R1 L0L > 0.2, BDT R1 L0I > 0.4, BDT R2 L0L > 0.3, BDT R2 L0I > 0.4}. For the chosen set of cuts, the signal efficiency is $\sim 94\%$ while the background rejection is $\sim 90\%$.

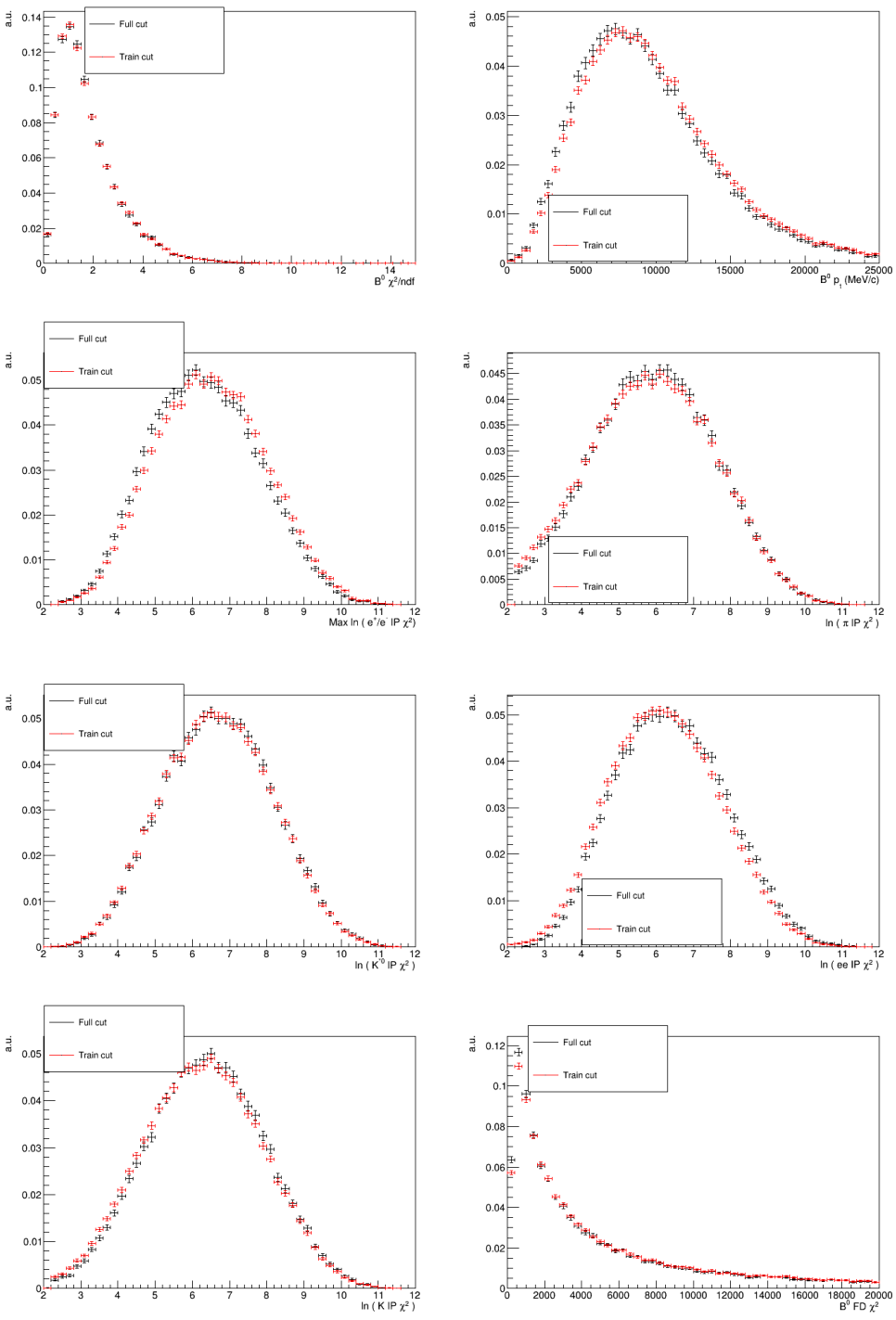


Figure 3.25: Comparison of the distributions of the training variables with the full selection applied (black) or the relaxed training selection applied (red), for the R1 signal sample.

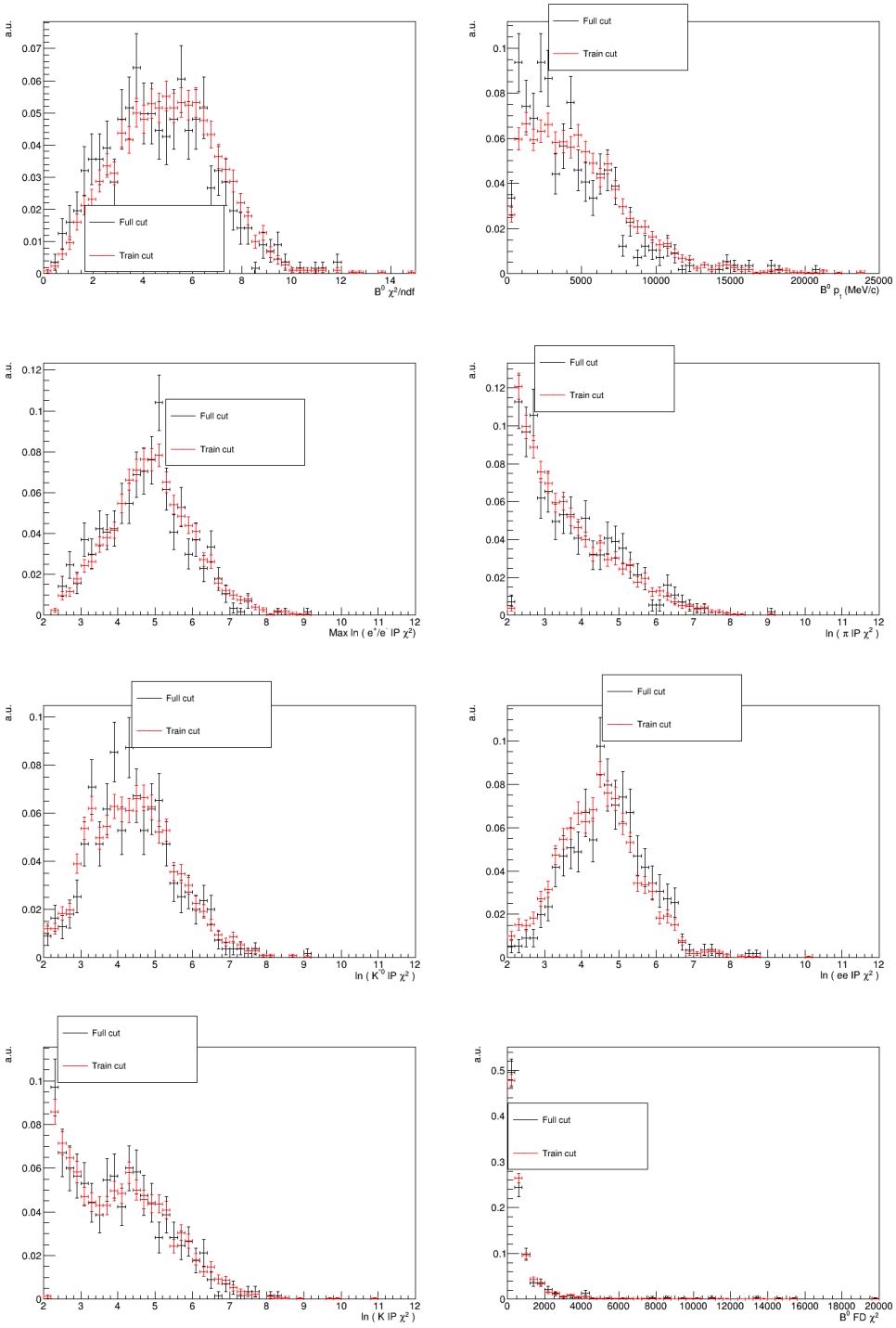


Figure 3.26: Comparison of the distributions of the training variables with the full selection applied (black) or the relaxed training selection applied (red), for the R1 background sample.

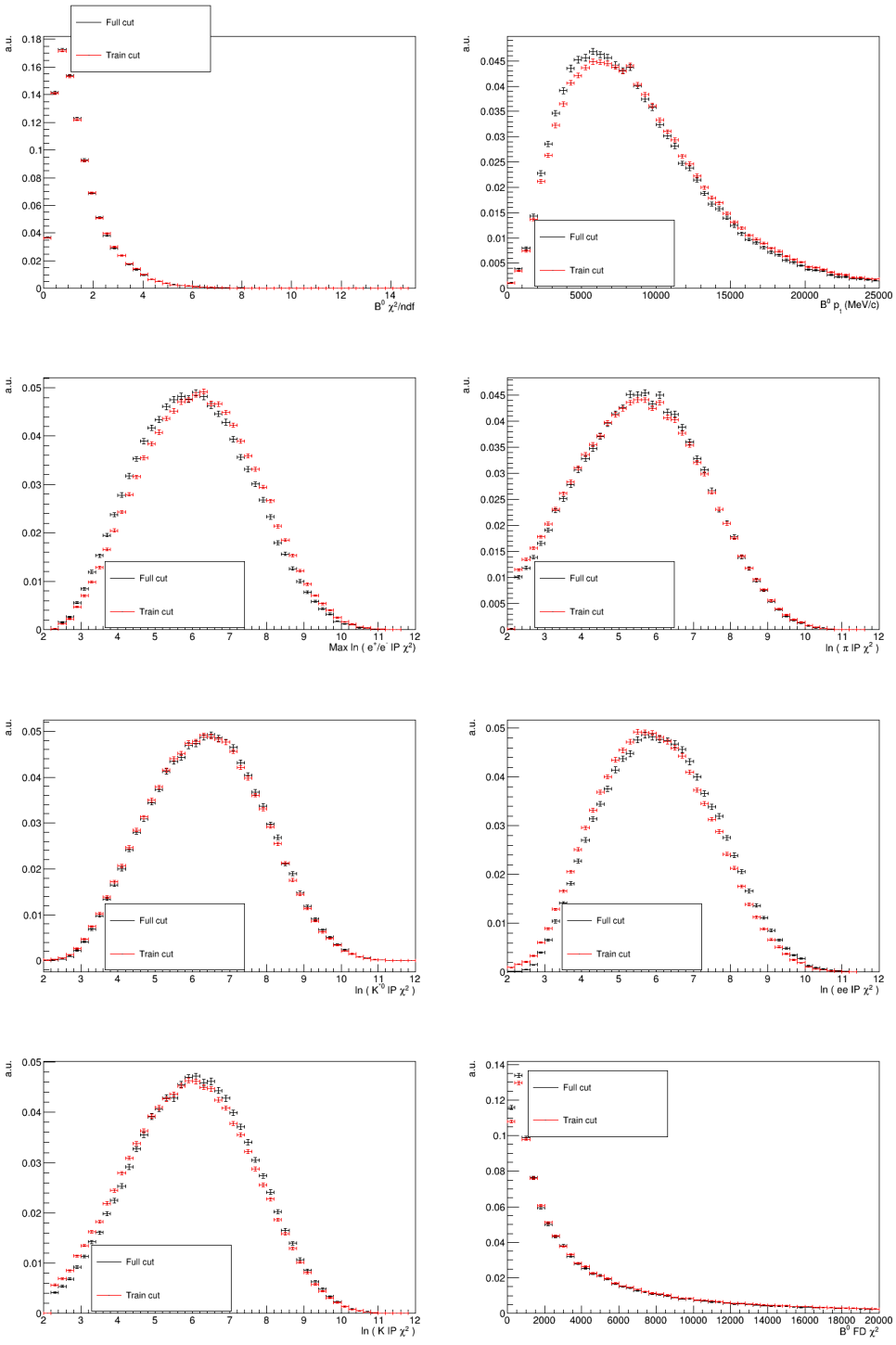


Figure 3.27: Comparison of the distributions of the training variables with the full selection applied (black) or the relaxed training selection applied (red), for the R2 signal sample.

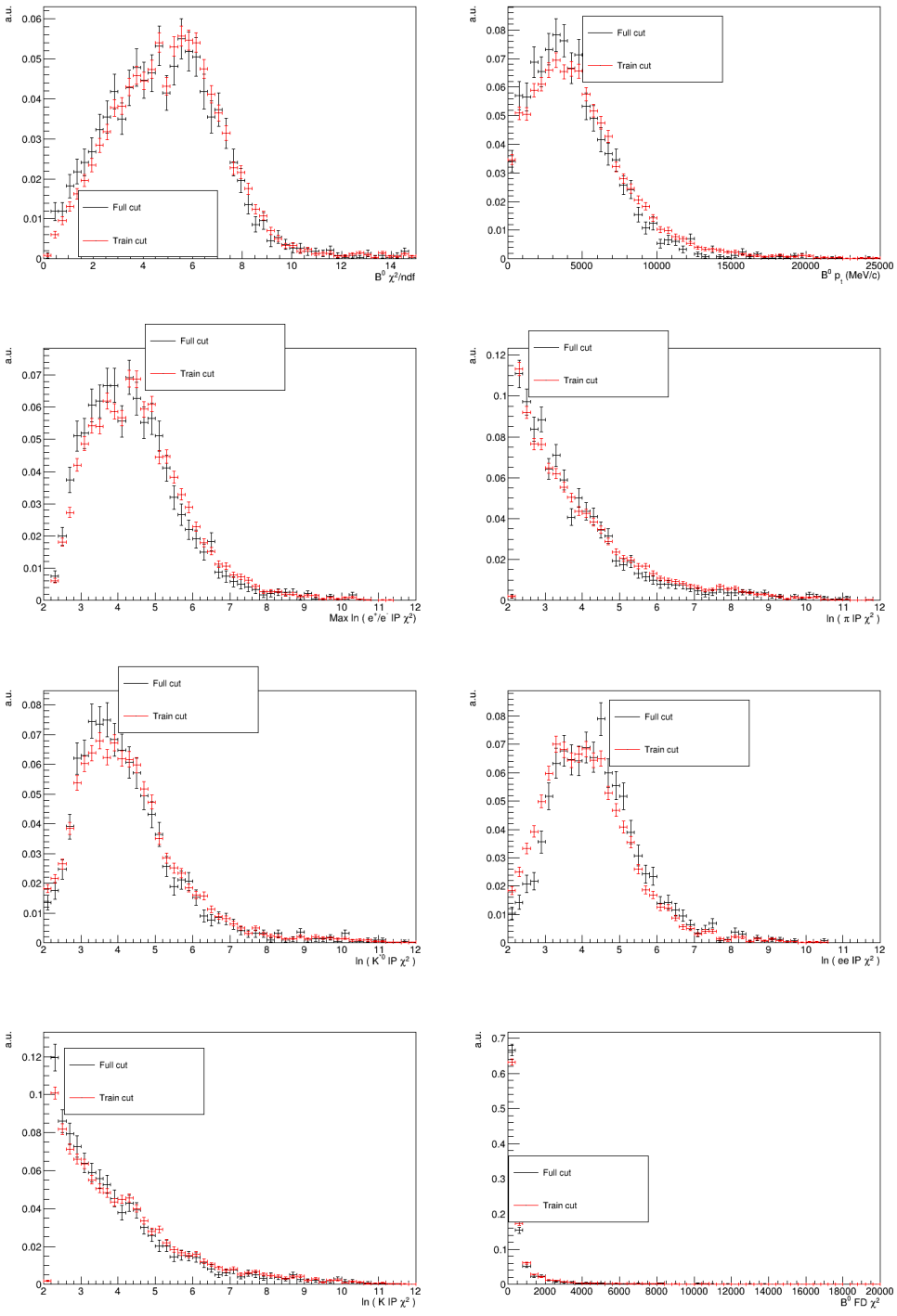


Figure 3.28: Comparison of the distributions of the training variables with the full selection applied (black) or the relaxed training selection applied (red), for the R2 background sample.

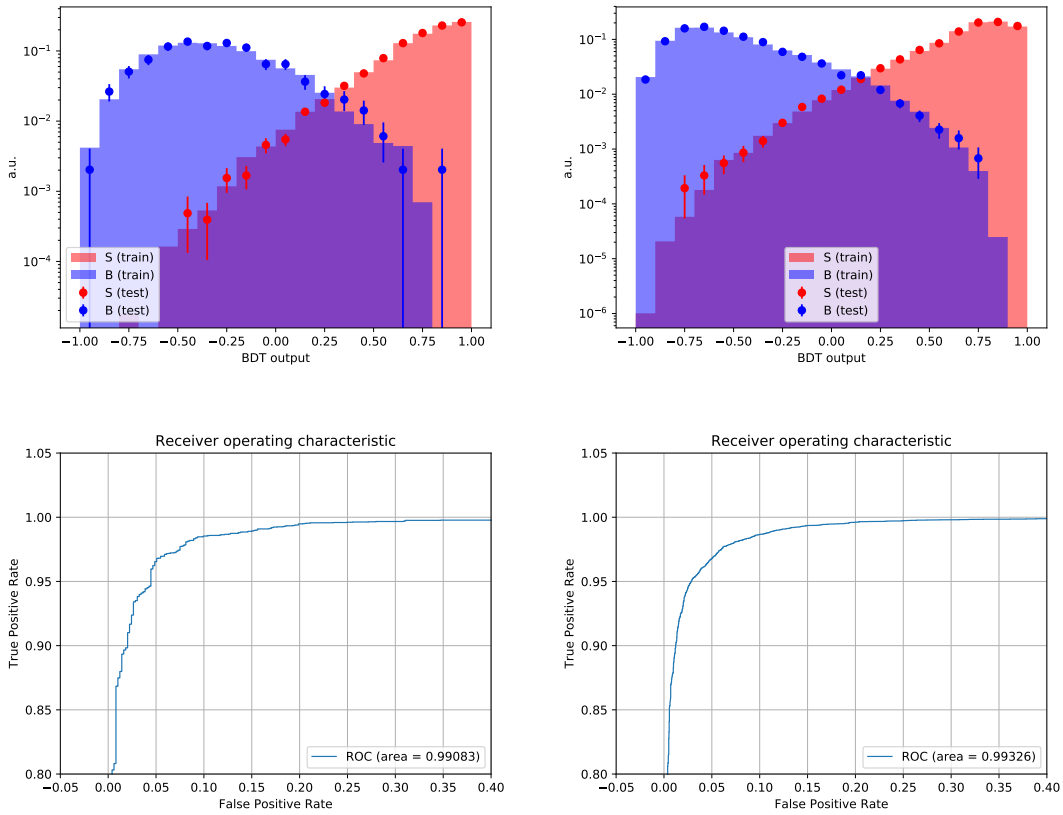


Figure 3.29: BDT output (top) of the first fold for R1 (left) and R2 (right) and their corresponding ROC curves (bottom). The test samples (dot) of the signal sample (red) and the background sample (blue) are shown on top of the training samples (histogram).

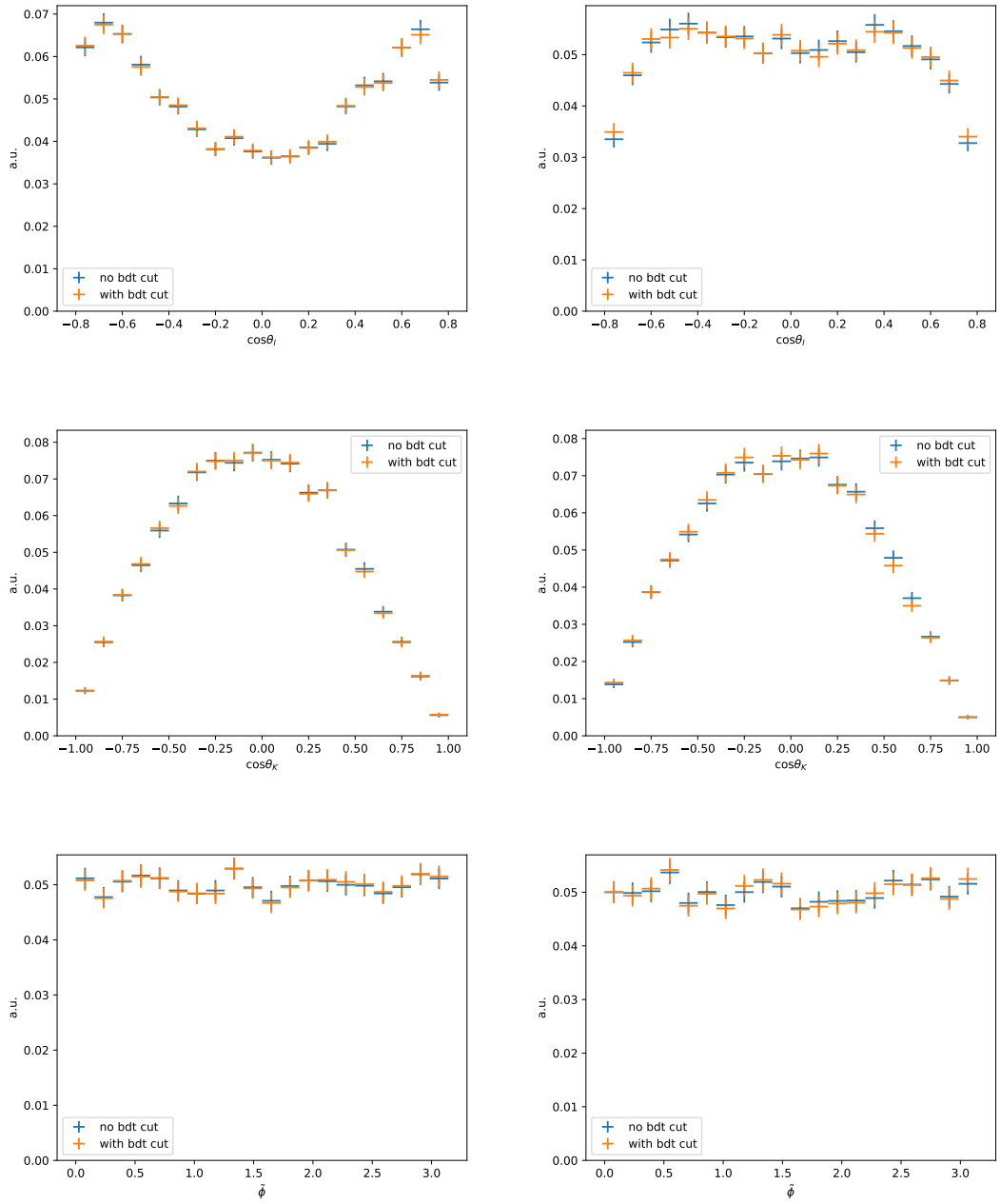


Figure 3.30: Distributions of the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ in the $B^0 \to K^{*0}e^+e^-$ MC with or without the optimal BDT cut for R1 for the L0L (left) and L0I (right) trigger category.

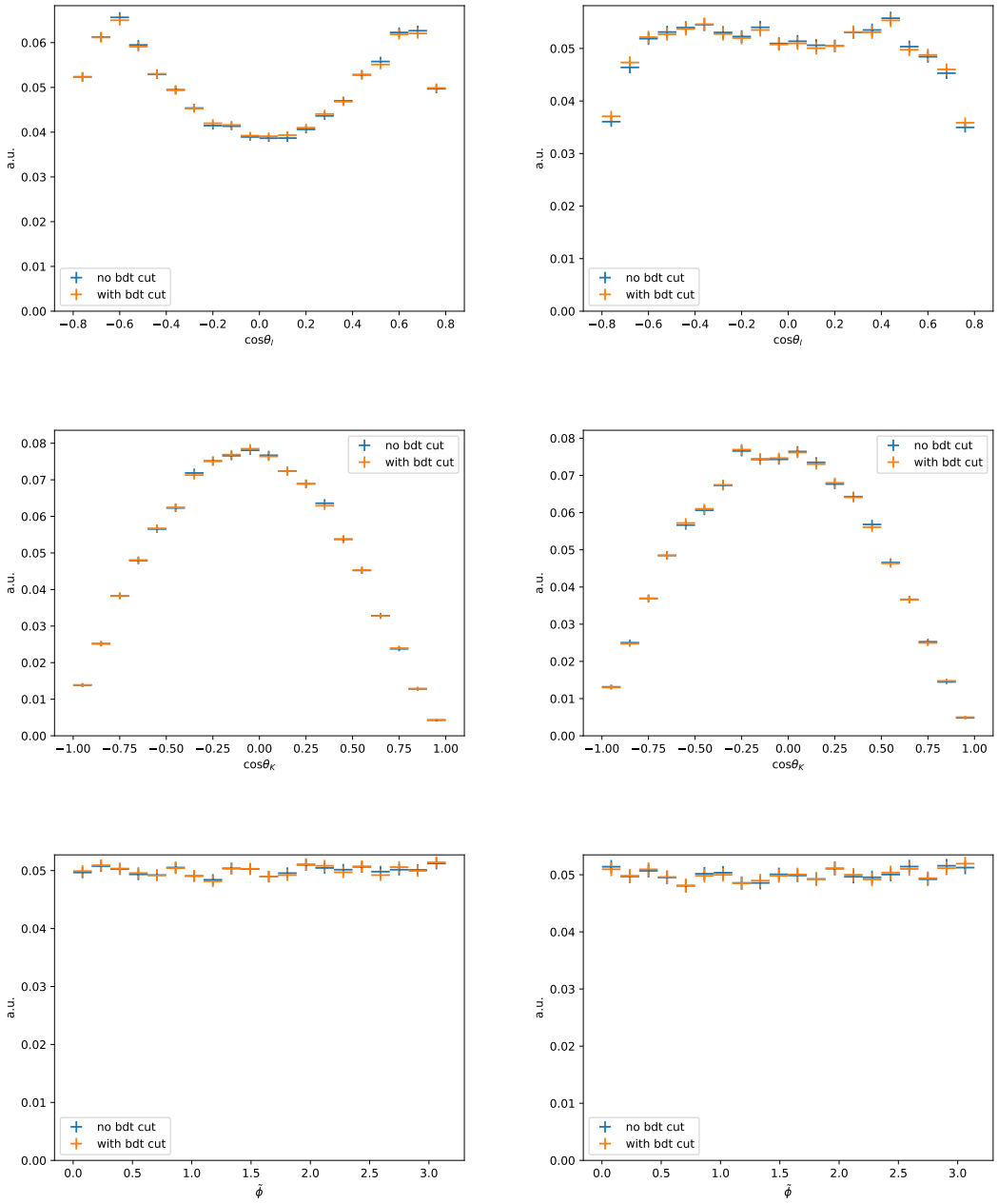


Figure 3.31: Distributions of the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ in the $B^0 \to K^{*0}e^+e^-$ MC with or without the optimal BDT cut for R2 for the L0L (left) and L0I (right) trigger category.

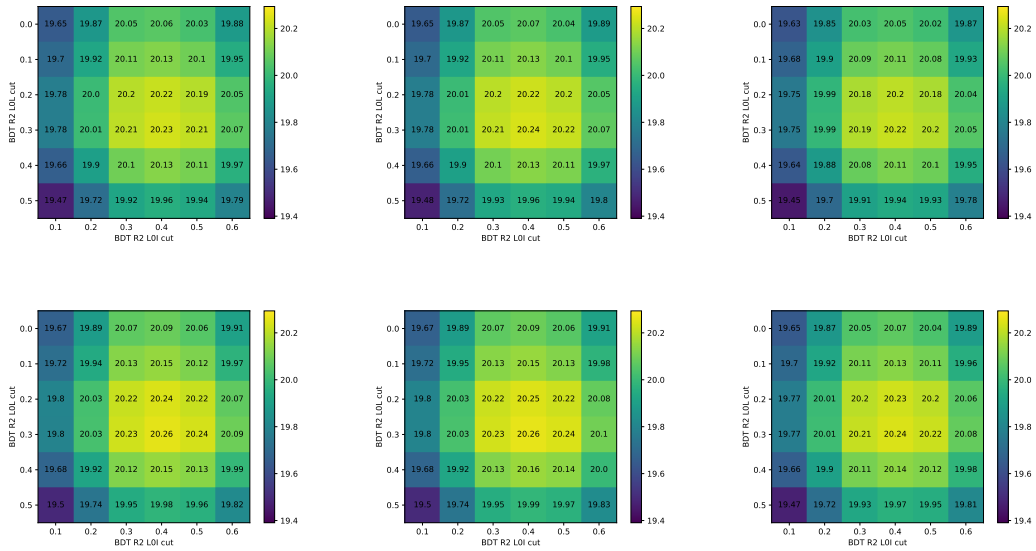


Figure 3.32: $S/\sqrt{S+B}$ for different sets of cuts {BDT R1 L0L cut, BDT R1 L0I cut, BDT R2 L0L cut, BDT R2 L0I cut }.

3.5 Selection summary and efficiency

The full selection of the analysis, applied on top of the trigger requirements (see section 3.1) is summarized in tables 3.3 and 3.4.

Type	**Requirement**
Global	`nSPDHits` $< 600(450)$ Run1(Run2)
B	$\lvert m - m_B^{\mathrm{PDG}}\rvert < 1500\,\mathrm{MeV}/c^2$
	`DIRA` > 0.9995
	$\chi^2_{\mathrm{IP}}(primary) < 25$
	end vertex $\chi^2/\mathrm{ndf} < 9$
	primary vertex χ^2 separation > 100
K^{*0}	$\lvert m - m_{K^{*0}}^{PDG}\rvert < 300\,\mathrm{MeV}/c^2$
	$p_{\mathrm{T}} > 500\,\mathrm{MeV}/c$
	origin vertex $\chi^2/\mathrm{ndf} < 25$
K	$\mathtt{DLL}_{K\pi} > -5$
	$\chi^2_{\mathrm{IP}}(primary) > 9(4)$
π	$\chi^2_{\mathrm{IP}}(primary) > 9$
$\ell\ell$	$m < 5500\,\mathrm{MeV}/c^2$
	end vertex $\chi^2/\mathrm{ndf} < 9$
	origin vertex χ^2 separation > 16
e	$\mathtt{DLL}_{e\pi} > 0$
	$p_{\mathrm{T}} > 300\,\mathrm{MeV}/c$
	$\chi^2_{\mathrm{IP}}(primary) > 9$

Table 3.3: Summary of the stripping requirements.

The efficiency of the full selection as a function of m_{ee}^{true} estimated from $B^0 \to K^{*0}e^+e^-$ MC is shown in figure 3.33. For R1 and 2015/2016 data taking periods, the simulated samples are stripping filtered (*i.e.* only the events passing the stripping requirements are saved to disk), thus the efficiency shown is the efficiency given that the event has already passed the stripping requirements (listed in table 3.3). For the 2017/2018 data taking period, the simulation is unfiltered thus the efficiency shown is the efficiency given that the final state particles of the event are in the acceptance of the LHCb detector. The drop of efficiency at very low m_{ee}^{true}

Type		Requirement
Quality	all tracks	$\chi^2/ndf < 3$
		`GhostProb` < 0.4
	e	Inside $ECAL$ acceptance
ID	K^{*0}	$\lvert m(K\pi) - m^{PDG}_{K^{*0}}\rvert < 100$ MeV/c^2
PID	all	Has hits in the $RICH$s detectors
	e	Has deposited a signal inside the $ECAL$
	K, π	$p_T > 250$ MeV/c
	e	$p_T > 500$ MeV/c
	K	`ProbNNk` $\cdot (1 -$ `ProbNNp`$) > 0.05$
		& `K_PIDK > 0`
	π	`ProbNNpi` $\times (1 -$ `ProbNNk`$)$
		$\times (1 -$ `ProbNNp`$) > 0.1$
	e	`V3ProbNNe` > 0.2 & `E_PIDe > 2`
BKG	Semileptonic	$\lvert\cos\theta_l\rvert < 0.8$
	$B^0 \to K^{*0}\gamma(\to e^+e^-)$	$\sigma_{\text{mod}} > 0.3$
	$B^0_s \to \phi e^+e^-$	$m(K(\pi \to K)) > 1040\,\text{MeV}/c^2$
	Partially reconstructed	$m_{HOP}(K^+\pi^-e^+e^-) > 4900$
	Combinatorial Run1	$L0L : BDT > 0.2$, $L0I : BDT > 0.4$
	Combinatorial Run2	$L0L : BDT > 0.3$, $L0I : BDT > 0.4$
Mass ranges	$m^{B^0}_{ee}$	$m^{B^0}_{ee} \in [10, 500]$ MeV/c^2
	$m(K^+\pi^-e^+e^-)$	$m(K^+\pi^-e^+e^-) \in [4500, 6200]$ MeV/c^2

Table 3.4: Summary of the selection cuts applied on top of the stripping and trigger requirements.

is mainly due to the converted photon veto cut (see section 3.3.2).

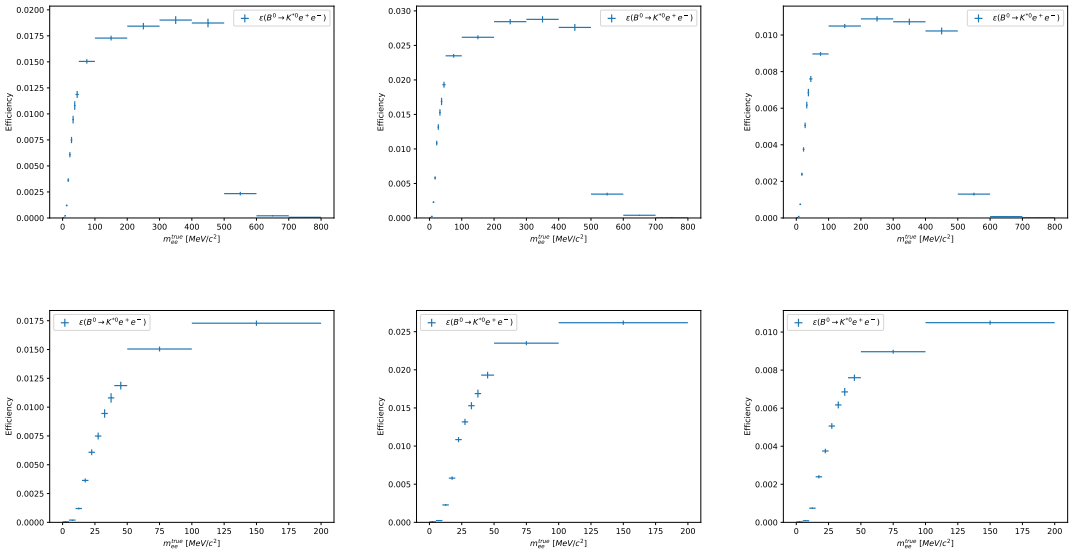


Figure 3.33: Efficiency of the full selection as a function of m_{ee}^{true} estimated from $B^0 \to K^{*0}e^+e^-$ MC for R1 (left), 2015/2016 (middle) and 2017/2018 (right) data taking periods. The bottom plots are zooms of the upper plots at low m_{ee}^{true}.

3.6 Mass fits

A simultaneous unbinned maximum likelihood fit to the $m(K^+\pi^-e^+e^-)$ invariant mass is performed in eight categories (two q^2 bins $\times$ two trigger categories $\times$ two Runs). The two q^2 bins, hereafter referred to as gamma-q^2 (very-low-q^2) are defined as $m_{ee}^{B^0} \in [0, 10]\,\text{MeV}/c^2$ ($m_{ee}^{B^0} \in [10, 500]\,\text{MeV}/c^2$). The two trigger categories L0L and L0I are defined in section 2.2.4. The two data taking periods are Run 1 (R1) and Run 2 (R2).
The mass shapes are fixed from MC and their modeling is detailed in section 3.6.1. To account for MC/data differences, the signal shapes are allowed to be modified by a scale and shift factor, which are fixed from a fit to the resonant mode $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ (see section 3.6.4). A wide range of $m(K^+\pi^-e^+e^-) \in [4500, 6200]\,\text{MeV}/c^2$ is chosen to properly model the radiative tail and the partially reconstructed background.

3.6.1 Mass shapes determination via fits to MC samples

The mass shapes of the signal components and of specific backgrounds is determined by a fit to simulated events. All simulation samples used for the mass shape determination are reweighted to correct for the PID response, the L0 trigger

response, the kinematics and the reconstruction (as described in section 3.2).

Signal PDFs

The signal Probability Density Functions (PDFs), $B^0 \to K^{*0}\gamma(\to e^+e^-)$ for the gamma-q^2 bin and $B^0 \to K^{*0}e^+e^-$ for the very-low-q^2 bin, are modeled by a Double Crystal Ball [5] (DCB). A fit to the MC with all bremsstrahlung categories included (*i.e.* events with no photon recovered, one or more photon(s) recovered are fitted together) is performed for each trigger category. It is checked (see Appendix D) that the fraction of events in each bremsstrahlung category is well reproduced in simulation. The fits to the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ and $B^0 \to K^{*0}e^+e^-$ MCs are shown in figure 3.34.

[5] A (double) crystal ball is a probability density function made of a Gaussian bulk with a (two) power law tail(s).

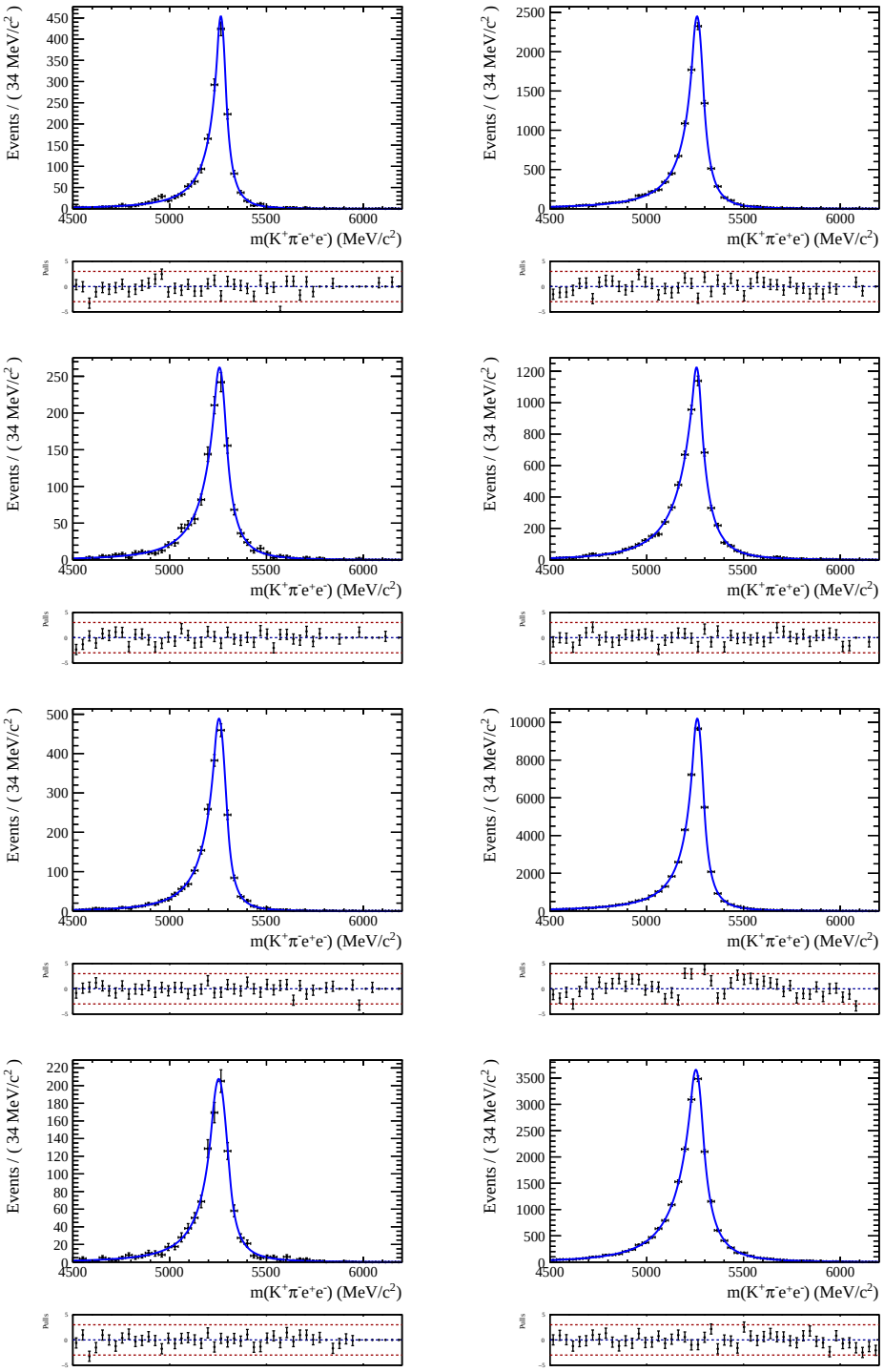


Figure 3.34: Left: Invariant mass fits to the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ MC. Right: Invariant mass fits to the $B^0 \to K^{*0}e^+e^-$ MC in the very-low-q^2 bin. From top to bottom: R1 L0L, R1 L0I, R2 L0L, R2 L0I.

3.6.2 Partially reconstructed background

In the m_{ee} range below 500 MeV/c^2, the partially reconstructed background mostly comes from $B^+ \to K_1(1270)(\to K^+\pi^- X)e^+e^-$ and $B^+ \to K_1(1270)(\to K^+\pi^- X)\gamma(\to e^+e^-)$, where X is not reconstructed. The partially reconstructed background mass shape is thus modeled by fits to $B^+ \to K_1(1270)(\to K^+\pi^- X)e^+e^-$ simulated events, using an ad hoc model of a sum of a Crystal Ball (CB) and a Gaussian, as shown in figure 3.35.

3.6.3 $B^0 \to K^{*0}\pi^0/\eta$ background

In the gamma-q^2 bin, the dominant π^0/η decays come from $B^0 \to K^{*0}\pi^0(\to \gamma(\to e^+e^-)\gamma)$ and $B^0 \to K^{*0}\eta(\to \gamma(\to e^+e^-)\gamma)$. The mass shapes of these backgrounds are thus obtained by fits to the corresponding MC samples with a `RooKeysPdf`[6] [95] from the `ROOT` software package [96]. Due to limited statistics, all years and trigger categories are merged together to extract the shape of these backgrounds. The fits to the $B^0 \to K^{*0}\pi^0(\to \gamma(\to e^+e^-)\gamma)$ and $B^0 \to K^{*0}\eta(\to \gamma(\to e^+e^-)\gamma)$ MCs are shown in figure 3.37.

In the very-low-q^2 bin, the dominant π^0/η decays come from $B^0 \to K^{*0}\pi^0(\to e^+e^-\gamma)$ and $B^0 \to K^{*0}\eta(\to e^+e^-\gamma)$. Similarly as above, the mass shapes of these backgrounds are obtained by fits to the corresponding MC samples with a `RooKeysPdf`. The fits to the $B^0 \to K^{*0}\pi^0/\eta(\to e^+e^-\gamma)$ MC are shown in figure 3.36.

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

In order to account for MC/data differences, a shift to the mean and a scale factor to the width of the DCB used to model the signal mass shapes are introduced when fitting to data. These factors are extracted from an invariant mass fit to the data of the resonant mode $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$.
The fit is performed in a similar fashion as the fit to the rare mode. The signal shape is first modeled by a DCB and is fixed by a fit to the $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$ MC (see figure 3.38). Then, the shift and the scale are introduced, such that the signal PDF fitted to the resonant data becomes

$$\begin{aligned}\mathcal{P}^{t,y}_{J/\psi} = DCB(&\mu^{t,y}_{J/\psi,\mathrm{MC}} + m^{y}_{J/\psi}, \sigma^{t,y}_{J/\psi,\mathrm{MC}} \cdot s^{y}_{J/\psi},\\ &\alpha^{t,y}_{1,J/\psi,\mathrm{MC}}, \alpha^{t,y}_{2,J/\psi,\mathrm{MC}}, n^{t,y}_{1,J/\psi,\mathrm{MC}}, n^{t,y}_{2,\gamma,\mathrm{MC}}) \quad ,\end{aligned} \tag{3.18}$$

[6] A `RooKeysPdf` is a non parametric probability density function. Its shape is obtained by summing many Gaussian functions.

where the upper script $t = \{\text{L0L, L0I}\}$ denotes the trigger category while $y = \{\text{R1,R2}\}$ denotes the years of data taking. $\mu_{J/\psi,\text{MC}}$ is the width, $\sigma_{J/\psi,\text{MC}}$ is the mean and $\alpha_{J/\psi,\text{MC}}$, $n_{J/\psi,\text{MC}}$ the parameters of the two tails of the DCB, all fixed from the fit to MC.

The following backgrounds are considered:

- $B_s^0 \to K^{*0}J/\psi(\to e^+e^-)$: modeled using the same mass shape as $B^0 \to K^{*0}J/\psi(\to e^+e^-)$, but shifting the mean by the difference of the nominal mass of the two mesons $m_{B^0} - m_{B_s^0}$.
- $\Lambda_b^0 \to pK^-J/\psi(\to e^+e^-)$: determined by an invariant mass fit using a `RooKeysPdf` to the $\Lambda_b^0 \to pK^-J/\psi(\to e^+e^-)$ MC (see figure 3.39), reweighted for the pK Dalitz plot [97].
- Combinatorial background: modeled by an exponential function.

When fitting to data, both fractions of $B_s^0 \to K^{*0}J/\psi(\to e^+e^-)$ and $\Lambda_b^0 \to pK^-J/\psi(\to e^+e^-)$ with respect to the signal $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ are fixed to

$$f_{B_s^0}^y = \frac{\mathcal{B}(B_s^0 \to K^{*0}J/\psi)}{\mathcal{B}(B^0 \to K^{*0}J/\psi)} \cdot \left(\frac{f_s}{f_d}\right)^y \quad , \tag{3.19}$$

and

$$f_{\Lambda_b^0}^{t,y} = \frac{\mathcal{B}(\Lambda_b^0 \to pK^-J/\psi)}{\mathcal{B}(B^0 \to K^{*0}J/\psi)} \cdot \left(\frac{f_{\Lambda_b}}{f_d}\right)^y \cdot \frac{\varepsilon^{t,y}(\Lambda_b^0 \to pK^-J/\psi(\to e^+e^-))}{\varepsilon^{t,y}(B^0 \to K^{*0}J/\psi(\to e^+e^-))} \quad , \tag{3.20}$$

where $\frac{f_s}{f_d}$ ($\frac{f_{\Lambda_b}}{f_d}$) is the B_s^0 (Λ_b^0) over B^0 production at LHCb [98] [99]. Since $\frac{f_{\Lambda_b}}{f_d}$ is p_T dependent, the ratio for the average p_T (estimated from the MC) is taken.

In summary, the full PDF for the fit to the resonant mode is given by

$$\begin{aligned}\mathcal{P}_{tot,J/\psi}^{t,y}(m_{J/\psi}^y, s_{J/\psi}^y, \lambda^{t,y}) = {} & f_{J/\psi}^{t,y} \cdot \mathcal{P}_{J/\psi}^{t,y}(m_{J/\psi}^y, s_{J/\psi}^y) \\ & + f_{B_s^0}^{t,y} \cdot \mathcal{P}_{B_s^0}^{t,y}(m_{J/\psi}^y, s_{J/\psi}^y) \\ & + f_{\Lambda_b^0}^{t,y} \cdot \mathcal{P}_{\Lambda_b^0}^{t,y} \\ & + (1 - f_{J/\psi}^{t,y} - f_{B_s^0}^{t,y} - f_{\Lambda_b^0}^{t,y} - f_{comb}^{t,y}) \cdot \mathcal{P}_{comb}^{t,y}(\lambda^{t,y}) \quad ,\end{aligned} \tag{3.21}$$

where $\lambda^{t,y}$ is the slope of the exponential function. Since the shift and scale factor $m_{J/\psi}$ and $s_{J/\psi}$ are found to be compatible for the two trigger categories, they are shared. A simultaneous fit to the two trigger categories is thus performed to the data in the q^2 bin $[6, 11]\,\text{GeV}^2/c^4$ using the PDF of eq. (3.21). The fit results are shown in figures 3.40. The values of $m_{J/\psi}$ and $s_{J/\psi}$ returned by the fits to R1 and R2 resonant data are then fixed and used for the fits to the rare modes.

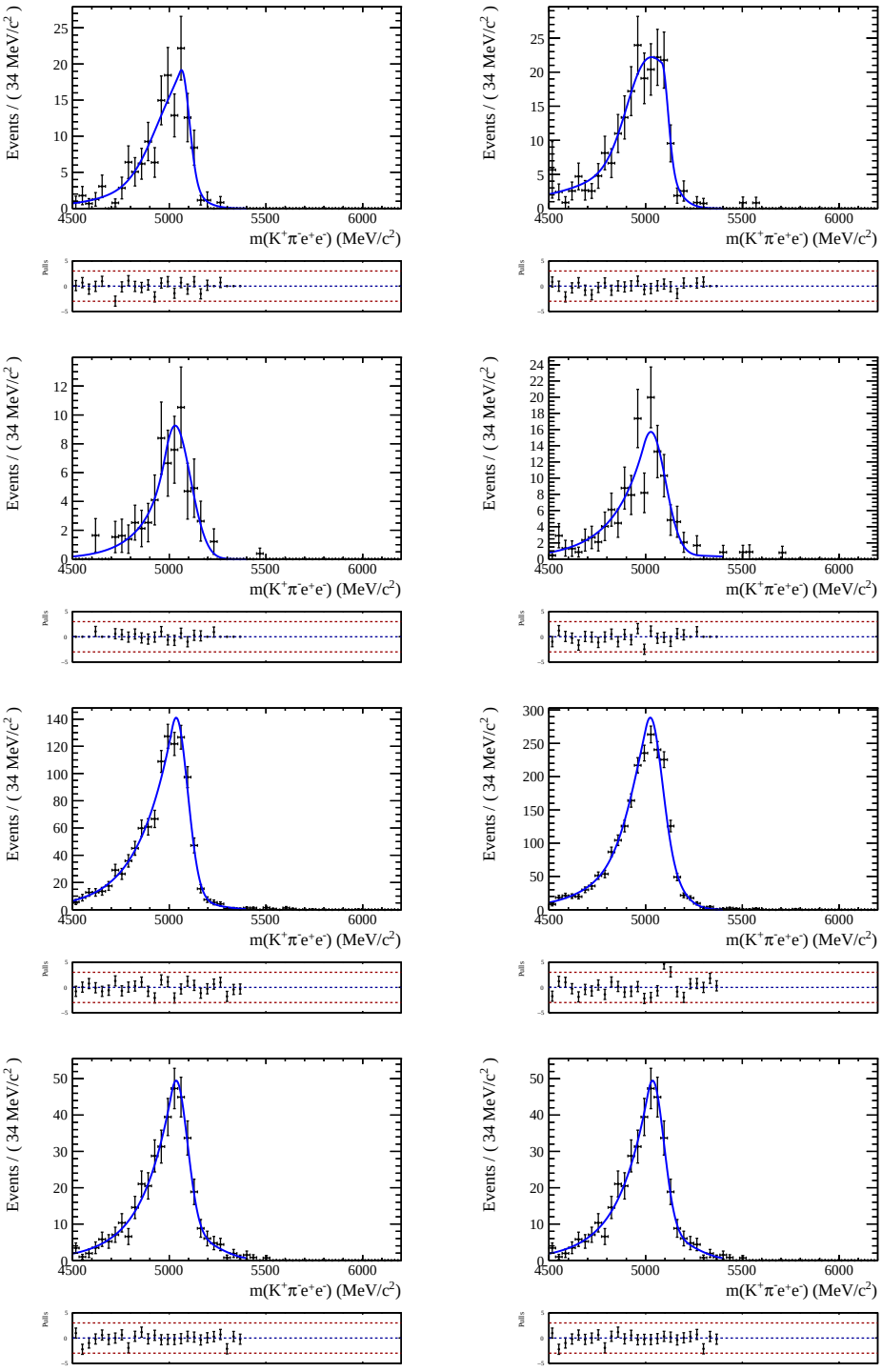


Figure 3.35: Invariant mass fits to the $B^+ \to K_1(1270)(\to K^+\pi^- X)e^+e^-$ MC in the gamma-q^2 (left) and very-low-q^2 (right) bins. From top to bottom: R1 L0L, R1 L0I, R2 L0L, R2 L0I.

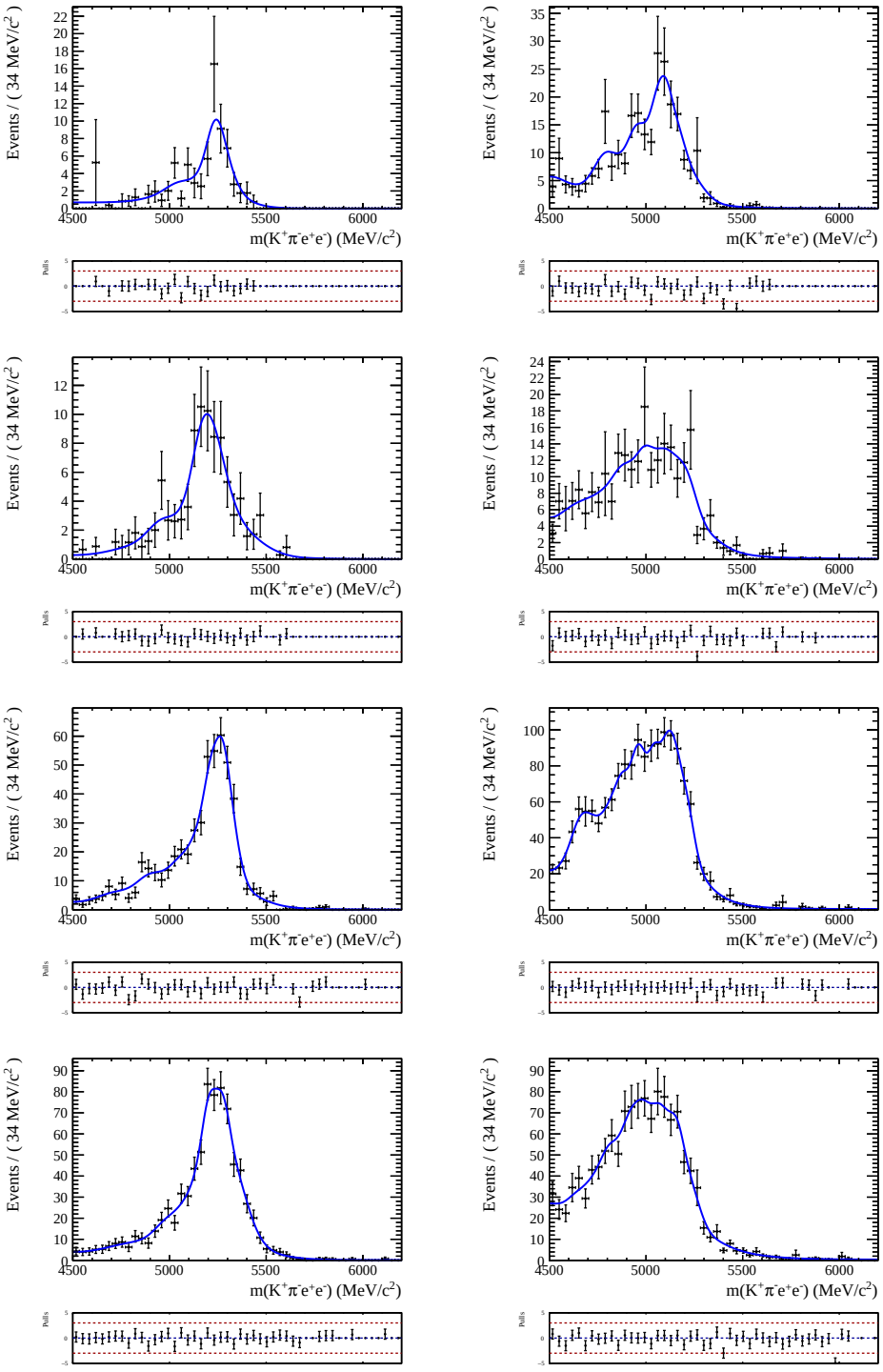


Figure 3.36: Invariant mass fits to the $B^0 \to K^{*0}\pi^0(\to e^+e^-\gamma)$ (left) and $B^0 \to K^{*0}\eta(\to e^+e^-\gamma)$ (right) MCs in the very-low-q^2 bin. From top to bottom: R1 L0L, R1 L0I, R2 L0L, R2 L0I.

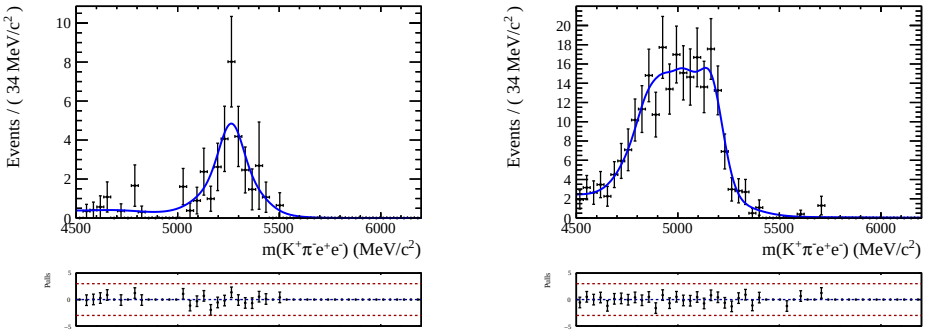


Figure 3.37: Invariant mass fits to the $B^0 \to K^{*0}\pi^0(\to \gamma(\to e^+e^-)\gamma)$ (left) and $B^0 \to K^{*0}\eta(\to \gamma(\to e^+e^-)\gamma)$ (right) MC in the gamma-q^2 bin, for all years and trigger categories merged together.

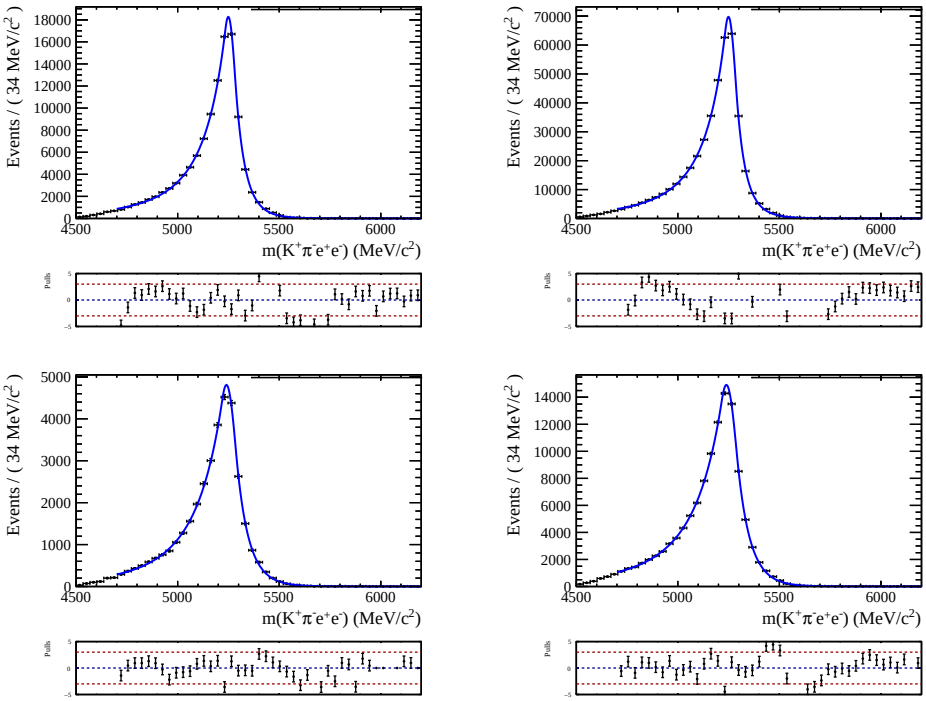


Figure 3.38: Invariant mass fits to the $B^0 \to K^{*0}J/\psi(\to e^+e^-)$ MC, in the trigger category L0L (top) and L0I (bottom), for R1 (left) and R2 (right).

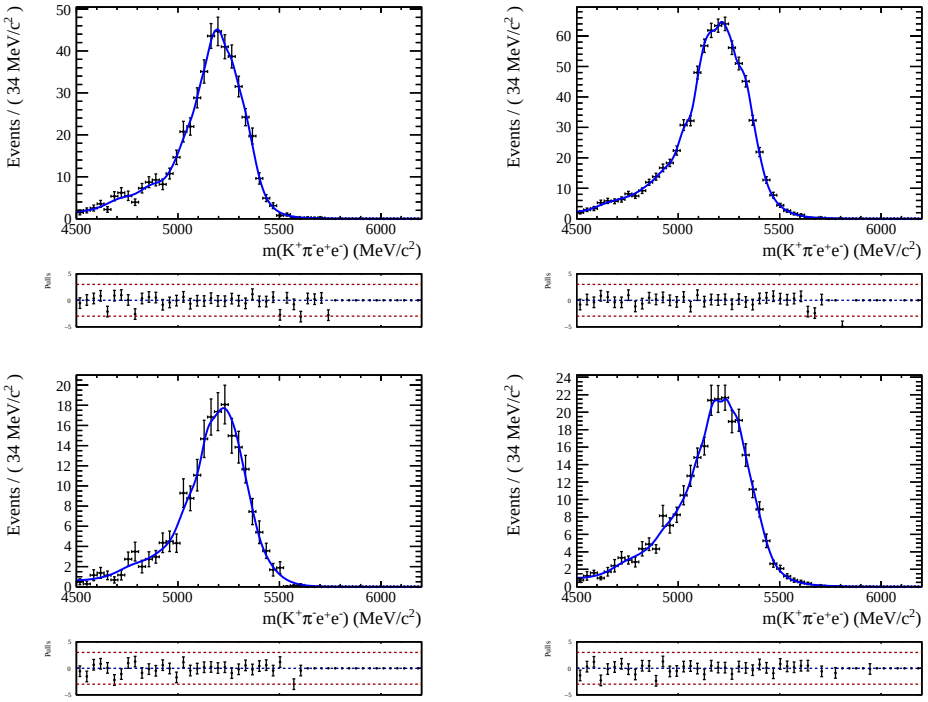


Figure 3.39: Invariant mass fits to the $\Lambda_b^0 \to pK^- J/\psi(\to e^+e^-)$ MC, in the trigger category L0L (top) and L0I (bottom), for R1 (left) and R2 (right).

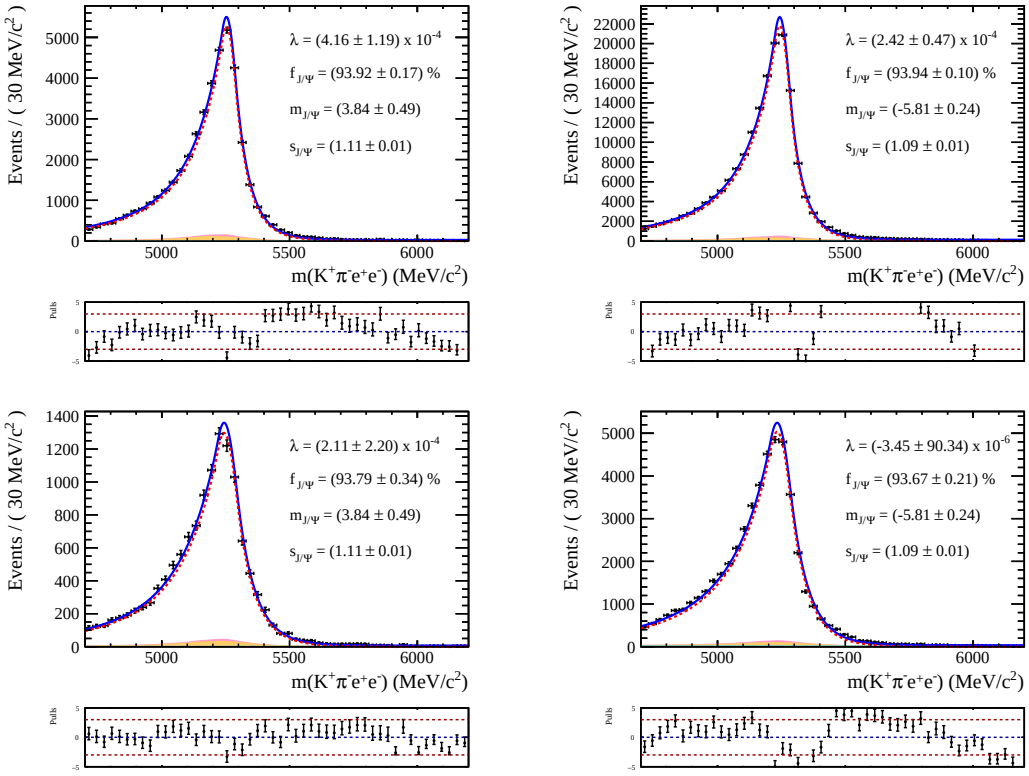


Figure 3.40: Invariant mass fits to the R1 (left) and R2 (right) data in the J/ψ-q^2 bin, in the trigger category L0L (top) and L0I (bottom). The signal is shown in red dotted line, the $B_s^0 \to K^{*0} J/\psi (\to e^+e^-)$ in pink, the $\Lambda_b^0 \to pK^- J/\psi (\to e^+e^-)$ in orange and the combinatorial background in cyan.

3.6.5 Invariant mass fit to $B^0 \to K^{*0} e^+ e^-$ data

The invariant mass fit to data is performed simultaneously in the two q^2 bins (gamma- and very-low-q^2), the two trigger categories (L0L and L0I) and the two data taking periods (R1 and R2). All the shapes are fixed from the fits to MC (see section 3.6.1). To account for data/MC differences, the signal DCB mean is shifted by a value $m_{J/\psi}$ while the width is scaled by a factor $s_{J/\psi}$, both fixed from the fit to data in the J/ψ-q^2 bin (see section 3.6.4). In the sections below the following notations are used to denote the years $y = \{\text{R1, R2}\}$, the trigger categories $t = \{\text{L0L, L0I}\}$ and the q^2 bins $q = \{g, l\}$ where g means the gamma- and l the very-low-q^2 bin. The different components of the mass fit to data are the following:

- Signal and leakage PDFs: $B^0 \to K^{*0}\gamma(\to e^+e^-)$ and $B^0 \to K^{*0}e^+e^-$, defined by

$$\begin{aligned}\mathcal{P}_s^{y,t,q} = DCB(&\mu_{s,\text{MC}}^{y,t,q} + m_{J/\psi}^{y}, \sigma_{s,\text{MC}}^{y,t,q} \cdot s_{J/\psi}^{y}, \\ &\alpha_{1,s,\text{MC}}^{y,t,q}, \alpha_{2,s,\text{MC}}^{y,t,q}, n_{1,s,\text{MC}}^{y,t,q}, n_{2,s,\text{MC}}^{y,t,q}) \quad .\end{aligned} \tag{3.22}$$

  The fraction $f_s^{y,t,q}$ associated to this PDF is defined as the sum of the fraction of $B^0 \to K^{*0}\gamma(\to e^+e^-)$ and $B^0 \to K^{*0}e^+e^-$ as

$$f_s^{y,t,q} = f_\gamma^{y,t,q} + f_{ee}^{y,t,q} \quad . \tag{3.23}$$

  The leakage from $B^0 \to K^{*0}e^+e^-$ events in the gamma-q^2 bin is constrained by the MC such that the fraction of the leakage is defined by

$$f_{ee}^{y,t,g} = \mathcal{G}_{ee}^{y,t} \left(\frac{N_{ee,\text{MC}}^{y,t,g}}{N_{ee,\text{MC}}^{y,t,l}} \right) \frac{N_{\text{data}}^{y,t,l}}{N_{\text{data}}^{y,t,g}} f_{ee}^{y,t,l} \quad , \tag{3.24}$$

  where $f_{ee}^{y,t,q}$ is the $B^0 \to K^{*0}e^+e^-$ fraction returned by the fit in each category while $N_{\text{data}}^{y,t,q}$ is the yield of data in each category. $\mathcal{G}$ is a Gaussian constraint parameterized by the ratio of yields extracted from simulation. Similarly, the leakage from $B^0 \to K^{*0}\gamma(\to e^+e^-)$ events in the very-low-q^2 bin is given by

$$f_\gamma^{y,t,l} = \mathcal{G}_\gamma^{y,t} \left(\frac{N_{\gamma,\text{MC}}^{y,t,l}}{N_{\gamma,\text{MC}}^{y,t,g}} \right) \frac{N_{\text{data}}^{y,t,g}}{N_{\text{data}}^{y,t,l}} f_\gamma^{y,t,g} \quad . \tag{3.25}$$

- Partially reconstructed background PDF: defined by

$$\begin{aligned}\mathcal{P}_{PR}^{y,t,q} = f_{CB}^{y,t,q} \cdot CB(\mu_{PR,\mathrm{MC}}^{y,t,q}, \sigma_{PR,\mathrm{MC}}^{y,t,q}, \alpha_{PR,\mathrm{MC}}^{y,t,q}, n_{PR,\mathrm{MC}}^{y,t,q}) \\ + (1 - f_{CB}^{y,t,q}) \cdot Gauss(\tilde{\mu}_{PR,\mathrm{MC}}^{y,t,q}, \tilde{\sigma}_{PR,\mathrm{MC}}^{y,t,q}) \quad .\end{aligned} \tag{3.26}$$

Since the underlying physics is similar between a real and a (quasi real) virtual photon at very low q^2, one expects the equality of the ratios

$$\frac{\mathcal{B}(B^+ \to K_1(1270)\gamma)}{\mathcal{B}(B^0 \to K^{*0}\gamma)} = \frac{\mathcal{B}(B^+ \to K_1(1270)e^+e^-)}{\mathcal{B}(B^0 \to K^{*0}e^+e^-)} \quad . \tag{3.27}$$

Thus, the fraction of partially reconstructed background with respect to the signal (including leakage) is shared between the two q^2 bins, imposing the additional constraint

$$f_{PR}^{y,t,q} = C_{PR}^{y,t} \cdot f_s^{y,t,q} \quad , \tag{3.28}$$

where $C_{PR}^{y,t}$ is a free parameter shared among the two q^2 bins.

- π^0/η background PDFs: defined by

$$\mathcal{P}_i^{y,t,q} \quad , \; i = \{\pi^0, \eta\} \quad , \tag{3.29}$$

which are non parametric shapes obtained from the fits to the corresponding MCs. The fractions of π^0/η backgrounds are Gaussian constrained from the MC such that

$$\begin{cases} f_i^{y,t,g} = \mathcal{G}_i^{y,t,g} \left(\frac{N_{i,\mathrm{MC}}^{y,t,g}}{N_{\gamma,\mathrm{MC}}^{y,t,g}} \right) f_\gamma^{y,t,g} \\ f_i^{y,t,l} = \mathcal{G}_i^{y,t,l} \left(\frac{N_{i,\mathrm{MC}}^{y,t,l}}{N_{ee,\mathrm{MC}}^{y,t,l}} \right) f_{ee}^{y,t,l} \end{cases} , \; i = \{\pi^0, \eta\} \quad . \tag{3.30}$$

- Combinatorial background PDF: defined by

$$\mathcal{P}_{comb}^{y,t,q} = \exp(\lambda^{y,t,q}) \quad . \tag{3.31}$$

In summary, the total PDF is given by

$$\begin{aligned}\mathcal{P}_{tot}^{y,t,q}(f_s^{y,t,q}, C_{PR}^{y,t}, \lambda^{y,t,q}, \mathcal{G}_{ee}^{y,t}, \mathcal{G}_{\gamma}^{y,t}, \mathcal{G}_{\eta}^{y,t,q}, \mathcal{G}_{\pi^0}^{y,t,q}) = \sum_y \sum_t \sum_q & f_{ee}^{y,t,q}(f_s^{y,t,q}, \mathcal{G}_{ee}^{y,t}) \cdot \mathcal{P}_s^{y,t,q} \\ & + f_{\gamma}^{y,t,q}(f_s^{y,t,q}, \mathcal{G}_{\gamma}^{y,t}) \cdot \mathcal{P}_s^{y,t,q} \\ & + f_{PR}^{y,t,q}(C_{PR}^{y,t}, f_s^{y,t,q}) \cdot \mathcal{P}_{PR}^{y,t,q} \\ & + f_{\eta}^{y,t,q}(\mathcal{G}_{\eta}^{y,t,q}) \cdot \mathcal{P}_{\eta}^{y,t,q} \\ & + f_{\pi^0}^{y,t,q}(\mathcal{G}_{\pi^0}^{y,t,q}) \cdot \mathcal{P}_{\pi^0}^{y,t,q} \\ & + \left(1 - \sum_{i \in \{s,PR,\eta,\pi^0\}} f_i^{y,t,q}\right) \cdot \mathcal{P}_{comb}^{y,t,q}\left(\lambda^{y,t,q}\right) \ , \end{aligned} \tag{3.32}$$

where the only free parameters are the signal plus leakage fractions $f_s^{y,t,g}$, the partially reconstructed fraction relative to the signal plus leakage shared between the two q^2 bins $C_{PR}^{y,t}$, the slopes of the exponential of the combinatorial background $\lambda^{y,t,q}$ and the Gaussian constraints on the $B^0 \to K^{*0}e^+e^-/B^0 \to K^{*0}\gamma(\to e^+e^-)$ leakages as well as the Gaussian constraints on the η/π^0 contaminations.

The fits to data are shown in figure 3.41. The fitted parameters are reported in the plots. A summary of all signal fractions is given in tables 3.5 and 3.6.

Sample	R1		R2	
	L0L	L0I	L0L	L0I
f_{ee}	5.4 ± 0.5	4.2 ± 0.6	6.3 ± 0.2	3.6 ± 0.3
f_{γ}	66.4 ± 2.4	59.2 ± 3.0	66.1 ± 1.6	57.1 ± 2.0
f_{PR}	5.5 ± 1.2	11.0 ± 2.0	5.6 ± 1.2	10.5 ± 1.9
f_{η}	3.9 ± 0.5	3.4 ± 0.5	4.7 ± 0.6	4.3 ± 0.5
f_{π^0}	3.9 ± 1.4	4.0 ± 1.2	5.4 ± 1.2	4.5 ± 1.1
f_{comb}	14.9 ± 2.4	18.2 ± 3.4	11.9 ± 1.3	20.0 ± 2.1
N_{data}	703	463	2556	1366

Table 3.5: Summary of the signal fractions (in %) in the gamma-q^2 bin obtained from the invariant mass fit. The uncertainties are only statistical.

Sample	R1		R2	
	L0L	L0I	L0L	L0I
f_{ee}	69.6 ± 6.0	56.2 ± 7.8	74.3 ± 2.8	53.8 ± 4.3
f_{γ}	1.7 ± 0.5	1.3 ± 0.4	0.9 ± 0.4	0.7 ± 0.3
f_{PR}	5.5 ± 1.3	10.0 ± 2.2	5.8 ± 1.2	9.5 ± 1.8
f_{η}	4.6 ± 0.6	5.3 ± 0.9	5.3 ± 0.5	6.1 ± 0.7
f_{π^0}	1.0 ± 0.2	1.9 ± 0.5	1.0 ± 0.2	2.1 ± 0.4
f_{comb}	17.6 ± 6.9	25.2 ± 10.0	12.7 ± 3.1	27.8 ± 5.6
N_{data}	88	66	413	241

Table 3.6: Summary of the signal fractions (in %) in the very-low-q^2 bin obtained from the invariant mass fit. The uncertainties are only statistical.

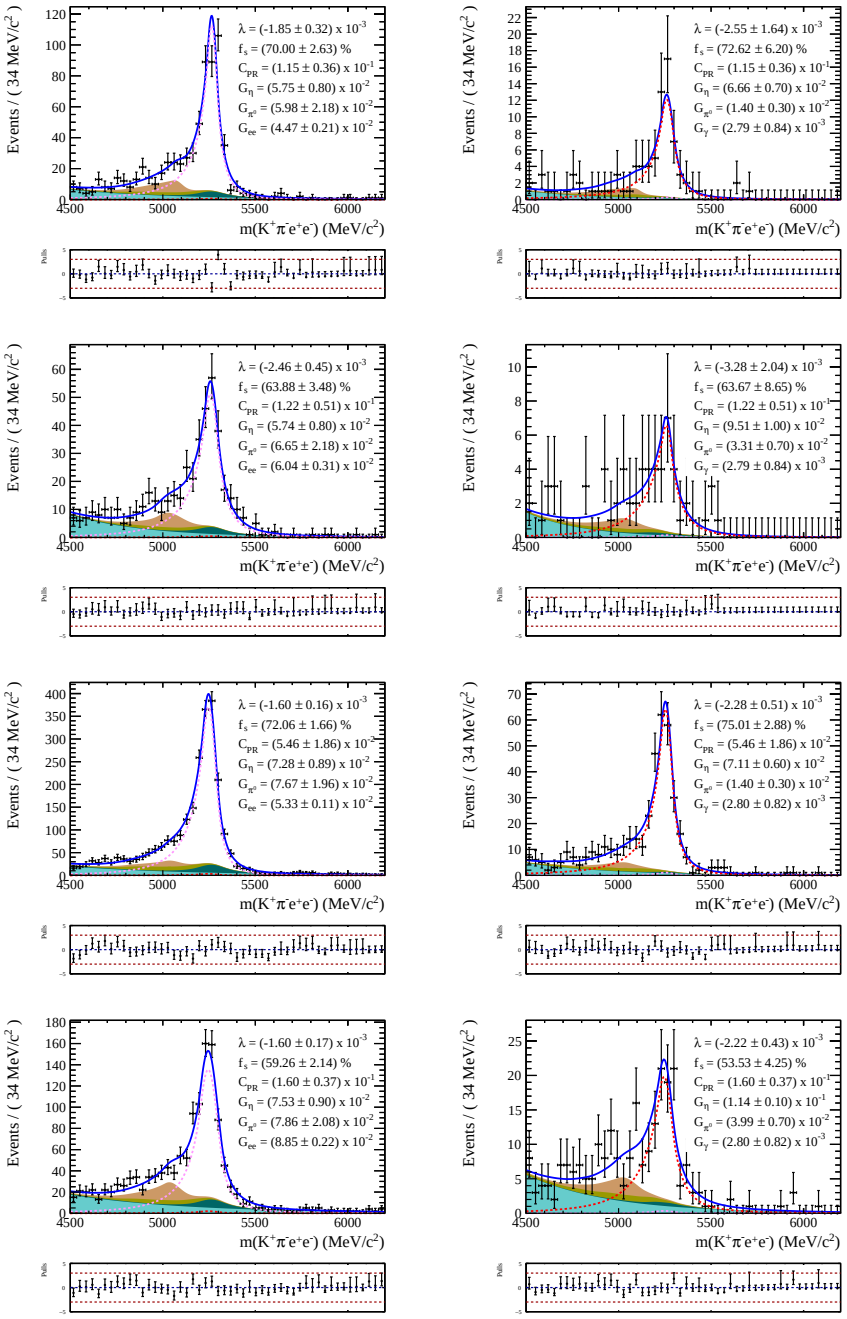


Figure 3.41: Invariant mass fits to data in the gamma-q^2 (left) and very-low-q^2 (right) bins. From top to bottom: R1 L0L, R1 L0I, R2 L0L, R2 L0I. The $B^0 \to K^{*0}\gamma(\to e^+e^-)$ contribution is shown in pink dotted line, the $B^0 \to K^{*0}e^+e^-$ one in red dotted line, the $B^0 \to K^{*0}\pi^0(\to \gamma(\to e^+e^-)\gamma)$ (left) and $B^0 \to K^{*0}\pi^0(\to \gamma e^+e^-)$ (right) backgrounds in dark green, the $B^0 \to K^{*0}\eta(\to \gamma(\to e^+e^-)\gamma)$ (left) and $B^0 \to K^{*0}\eta(\to \gamma e^+e^-)$ (right) backgrounds in light green, the partially reconstructed background in brown, and the combinatorial background in cyan.

Chapter 4

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

This chapter presents the measurement of the photon polarization in $b \to s\gamma$ processes through the angular analysis of the $B^0 \to K^{*0} e^+ e^-$ decay channel in the q^2 range $[0.0001, 0.25]\,\mathrm{GeV}^2/c^4$. The analysis uses the full Run 1 and Run 2 datasets collected at the LHCb experiment corresponding to $9.1\,\mathrm{fb}^{-1}$ integrated luminosity (see section 2.1.1), selected as explained in Chapter 3.
The angular acceptance induced by the geometry of the detector as well as the reconstruction and selection of $B^0 \to K^{*0} e^+ e^-$ candidates is modeled in section 4.2. The angular shape of the main backgrounds are modeled in section 4.3. The fit to measure the four angular observables $A_T^{(2)}$, A_T^{Im}, A_T^{Re} and F_L is validated in section 4.4. The systematic uncertainties of the measurement are detailed in section 4.5. Finally, the four dimensional fit to $B^0 \to K^{*0} e^+ e^-$ data to the $m(K^+\pi^- e^+ e^-)$ invariant mass and the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ is then performed. The results of the fits as well as their implications are presented in section 4.6.

4.1 Angular fit strategy

As explained in section 3.6, the invariant mass fit is performed in a wide mass window $[4500, 6200]\,\mathrm{MeV}/c^2$. However, for the angular fit, the mass window is reduced to decrease the background contamination and to make the modeling of the angular shape of the backgrounds easier. To get the signal and background yields in the reduced mass windows, the PDFs obtained after the full mass fit are integrated over the reduced region. The mass window for the angular fit is $[5000, 5400]\,\mathrm{MeV}/c^2$, which is chosen by maximizing $S/\sqrt{S+B}$ (where S is the signal yield while B is the sum of background yields). The B/S ratios in the

angular fit mass window are presented in table 4.1.

	R1		R2		All
	L0L	L0I	L0L	L0I	All
B/S	0.15 ± 0.03	0.24 ± 0.05	0.15 ± 0.02	0.34 ± 0.04	0.20

Table 4.1: Fraction of background over signal in the $[5000, 5400]\,\mathrm{MeV}/c^2$ angular mass window in the very-low-q^2 bin.

The aim of the angular analysis is to measure the four observables F_{L}, $A_T^{(2)}$, A_T^{Re} and A_T^{Im} by performing a four dimensional fit to the $B^0 \to K^{*0}e^+e^-$ signal candidates to the differential decay width given in eq. (1.35) and the $m(K^+\pi^-e^+e^-)$ invariant mass. The fit is performed simultaneously in the two trigger categories and the two data samples R1 and R2, with shared parameters F_{L}, $A_T^{(2)}$, A_T^{Re} and A_T^{Im}.
The reconstruction and selection of signal candidates induces a distortion in the angular distributions. To take this into account, the signal PDF is multiplied by an angular acceptance PDF (whose modeling and determination is discussed in details in section 4.2). The signal angular PDF thus reads

$$\begin{aligned}\mathcal{A}_{ee}^{t,y}\left(\cos\theta_\ell, \cos\theta_K, \tilde{\phi}\right) = \mathcal{A}_{ee}^{phy}\left(\cos\theta_\ell, \cos\theta_K, \tilde{\phi}; F_{\mathrm{L}}, A_{\mathrm{T}}^2, A_T^{Im}, A_T^{Re}\right) \\ \times \varepsilon_{\mathcal{A}}^{t,y}\left(\cos\theta_\ell, \cos\theta_K, \tilde{\phi}\right) \quad ,\end{aligned} \tag{4.1}$$

where $t = \{\mathrm{L0L,L0I}\}$ denotes the trigger category, $y = \{\mathrm{R1,R2}\}$ denotes the data sample, $\mathcal{A}_{ee}^{phy}$ is the physical differential decay width given in eq. (1.35) and $\varepsilon_{\mathcal{A}}^{t,y}$ is the trigger and Run dependent angular acceptance.
In addition, due to the non negligible background contamination, the angular distributions of the main backgrounds have to be modeled as well and added to

the total angular PDF, which reads

$$\begin{aligned}\mathcal{A}^{t,y}_{tot}\left(\cos\theta_\ell,\cos\theta_K,\tilde{\phi}\right) = &\left(1-p^{t,y}_{PR}-p^{t,y}_{comb}-p^{t,y}_{\pi^0}-p^{t,y}_{\eta}-p^{t,y}_{\gamma}\right)\cdot\mathcal{A}^{t,y}_{ee}\left(\cos\theta_\ell,\cos\theta_K,\tilde{\phi}\right)\\ &+p^{t,y}_{PR}\cdot\mathcal{A}^{t,y}_{PR}\left(\cos\theta_\ell,\cos\theta_K,\tilde{\phi}\right)\\ &+p^{t,y}_{comb}\cdot\mathcal{A}^{t,y}_{comb}\left(\cos\theta_\ell,\cos\theta_K,\tilde{\phi}\right)\\ &+p^{t,y}_{\pi^0}\cdot\mathcal{A}^{t,y}_{\pi^0}\left(\cos\theta_\ell,\cos\theta_K,\tilde{\phi}\right)\\ &+p^{t,y}_{\eta}\cdot\mathcal{A}^{t,y}_{\eta}\left(\cos\theta_\ell,\cos\theta_K,\tilde{\phi}\right)\\ &+p^{t,y}_{\gamma}\cdot\mathcal{A}^{t,y}_{\gamma}\left(\cos\theta_\ell,\cos\theta_K,\tilde{\phi}\right) \quad ,\end{aligned} \tag{4.2}$$

where $\mathcal{A}^{t,y}_{PR}$, $\mathcal{A}^{t,y}_{comb}$, $\mathcal{A}^{t,y}_{\pi^0}$, $\mathcal{A}^{t,y}_{\eta}$ and $\mathcal{A}^{t,y}_{\gamma}$ are the angular PDFs of the partially reconstructed, the combinatorial, the $B^0\to K^{*0}\pi^0(\to e^+e^-\gamma)$, the $B^0\to K^{*0}\eta(\to e^+e^-\gamma)$ and the $B^0\to K^{*0}\gamma(\to e^+e^-)$ backgrounds respectively. The modeling of the background angular PDFs is discussed in details in section 4.3. The fractions associated to each background, denoted $p^{t,y}_i$ in eq. (4.2) are Gaussian constrained from the mass fit (see section 3.6) result. The full four dimensional PDF thus reads

$$\begin{aligned}\mathcal{F}^{t,y}_{tot}\left(m(K^+\pi^-e^+e^-),\cos\theta_\ell,\cos\theta_K,\tilde{\phi}\right) = &\,\mathcal{P}^{t,y,l}_{tot}\left(m(K^+\pi^-e^+e^-)\right)\\ &\times\mathcal{A}^{t,y}_{tot}\left(\cos\theta_\ell,\cos\theta_K,\tilde{\phi}\right) \quad ,\end{aligned} \tag{4.3}$$

where the mass part of the PDF $\mathcal{P}^{t,y,l}_{tot}$ is fitted over the wide mass window $[4500,6200]\,\mathrm{MeV}/c^2$ while the angular part $\mathcal{A}^{t,y}_{tot}$ is fitted over the narrow mass window $[5000,5400]\,\mathrm{MeV}/c^2$.

4.2 Angular acceptance

As mentioned above the geometry of the LHCb detector, the reconstruction and the selection of signal candidates can induce a distortion in the three angular distributions. To study these effects, $B^0\to K^{*0}e^+e^-$ simulated events are generated with no underlying physics but momentum conservation. Hence the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ are generated with a flat uncorrelated distribution. This sample is referred to as $B^0\to K^{*0}e^+e^-$ phase-space MC. The full reconstruction algorithm as well as the full selection is then applied to the phase-space MC. The resulting

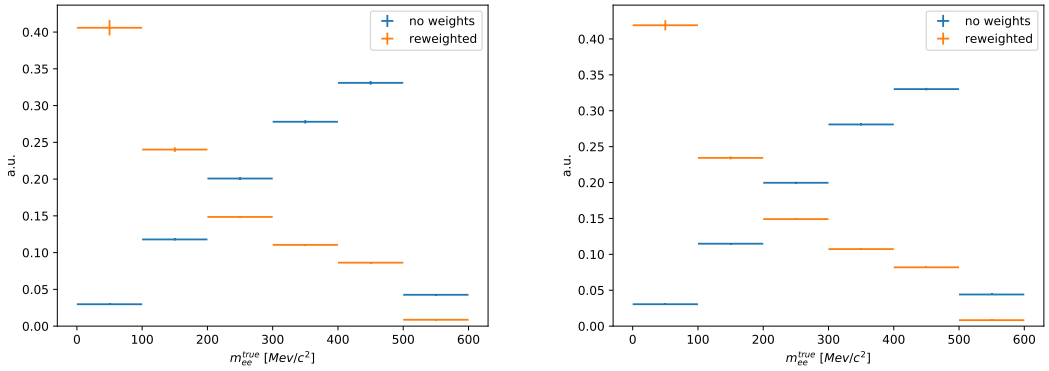


Figure 4.1: True dielectron mass distribution of the phase space $B^0 \to K^{*0}e^+e^-$ MC before and after reweighting for R1 (left) and R2 (right).

distribution of $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ thus give access to the angular acceptance of the three angles.

4.2.1 Phase-space Monte Carlo reweighting

The phase-space MC is generated with a flat q^2 distribution resulting in a dramatically different dielectron mass distribution than the SM physics. Even though the q^2 bin of the analysis is very narrow, the angular acceptance is expected to have some correlation with the dielectron mass, especially for the angle θ_l which directly involves the two leptons. To cope with this issue, an additional phase-space $B^0 \to K^{*0}e^+e^-$ MC sample is generated requiring $m_{ee}^{\text{true}} < 600\,\text{MeV}/c^2$ and added to the nominal one (this is done to avoid very large weights in the following reweighting procedure). Then, this enriched sample of phase-space $B^0 \to K^{*0}e^+e^-$ MC is reweighted in bins of m_{ee}^{true} to match the SM $B^0 \to K^{*0}e^+e^-$ MC. A rather coarse binning (100 MeV/c^2 width) is chosen to avoid large weights. This procedure is allowed since the angular acceptance changes slowly with m_{ee}. The true dielectron mass distribution of the phase-space MC before and after reweighting is shown in figure 4.1.

4.2.2 Acceptance fit

To extract the angular acceptance for each of the three angles $\cos\theta_\ell$, $\cos\theta_K$ and ϕ an unbinned maximum likelihood fit to the phase-space $B^0 \to K^{*0}e^+e^-$ MC after full reconstruction and selection is performed. All corrections discussed in section 3.2 as well as the q^2 reweighting discussed in section 4.2.1 are also applied. Note

however, that these corrections have a small impact on the distribution of the angular acceptances, as shown in Appendix C.2. Since the angular acceptance is expected to be trigger and Run dependent, the fit is performed separately for the two trigger categories and the two data samples (R1 and R2).
For the $\cos\theta_\ell$ and $\cos\theta_K$ angles, the acceptance is modeled by Legendre polynomials truncated at order four, given by

$$\begin{aligned} 1 + \sum_{i=1}^{4} c_i \mathcal{P}_i(x) = 1 + c_1 x + c_2 \frac{1}{2}(3x^2 - 1) \\ + c_3 \frac{1}{2}(5x^3 - 3x) + c_4 \frac{1}{8}(35x^4 - 30x^2 + 3) \quad . \end{aligned} \tag{4.4}$$

This choice is motivated by the fact that Legendre polynomials of different order are orthonormal to each other with respect to an integration between -1 and 1. Thus, this minimizes the correlation between the fitted coefficients c_i of eq. (4.4).
Since $\cos\theta_\ell$ is roughly equal to the energy asymmetry between the two leptons and since the detector and reconstruction effects as well as the selection are not expected to introduce any bias with respect to a particular lepton charge, the $\cos\theta_\ell$ acceptance is expected to be symmetric. Therefore only the lowest order odd coefficient of eq. (4.4) is kept. As can be seen in figure 4.2, this remaining odd coefficient is indeed compatible with zero. The PDF used for the fit of the angular acceptance of $\cos\theta_\ell$ thus reads

$$\begin{aligned} \varepsilon^{t,y}_{\mathcal{A},\cos\theta_\ell}(\cos\theta_\ell; c^{t,y}_{l2}, c^{t,y}_{l4}) = 1 + c^{t,y}_{l1}\cos\theta_\ell + c^{t,y}_{l2}\frac{1}{2}\left(3\cos{\theta_\ell}^2 - 1\right) \\ + c^{t,y}_{l4}\frac{1}{8}\left(35\cos{\theta_\ell}^4 - 30\cos{\theta_\ell}^2 + 3\right) \quad . \end{aligned} \tag{4.5}$$

The angular acceptance fits to $\cos\theta_\ell$ as well as the fit parameters are shown in figure 4.2.

The angular acceptance for $\cos\theta_K$ however is expected to be asymmetric, due to the mass difference between the pion and the kaon. The PDF used for the fit of the angular acceptance of $\cos\theta_K$ hence reads

$$\begin{aligned} \varepsilon^{t,y}_{\mathcal{A},\cos\theta_K}(\cos\theta_K; c^{t,y}_{K1}, c^{t,y}_{K3}) = 1 + c^{t,y}_{K1}\cos\theta_K \\ + c^{t,y}_{K2}\frac{1}{2}(3\cos{\theta_K}^2 - 1) \\ + c^{t,y}_{K3}\frac{1}{2}(5\cos{\theta_K}^3 - 3\cos\theta_K) \\ + c^{y,t}_{K4}\frac{1}{8}(35\cos{\theta_K}^4 - 30\cos{\theta_K}^2 + 3) \quad . \end{aligned} \tag{4.6}$$

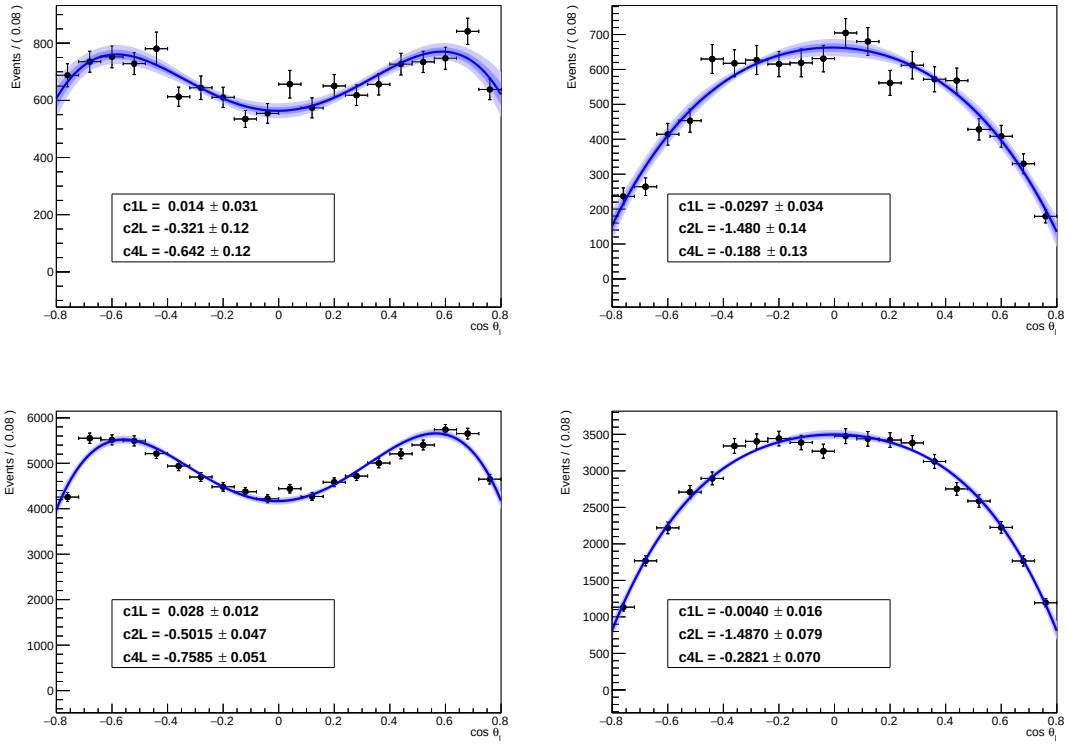


Figure 4.2: Fits to the phase space $B^0 \to K^{*0} e^+ e^-$ MC of the $\cos\theta_\ell$ distribution for R1 (top) and R2 (bottom) in the L0L (left) and L0I (right) trigger categories. The fitted function from Eq. (4.5) is shown in blue with its associated 1 σ and 2 σ error bands.

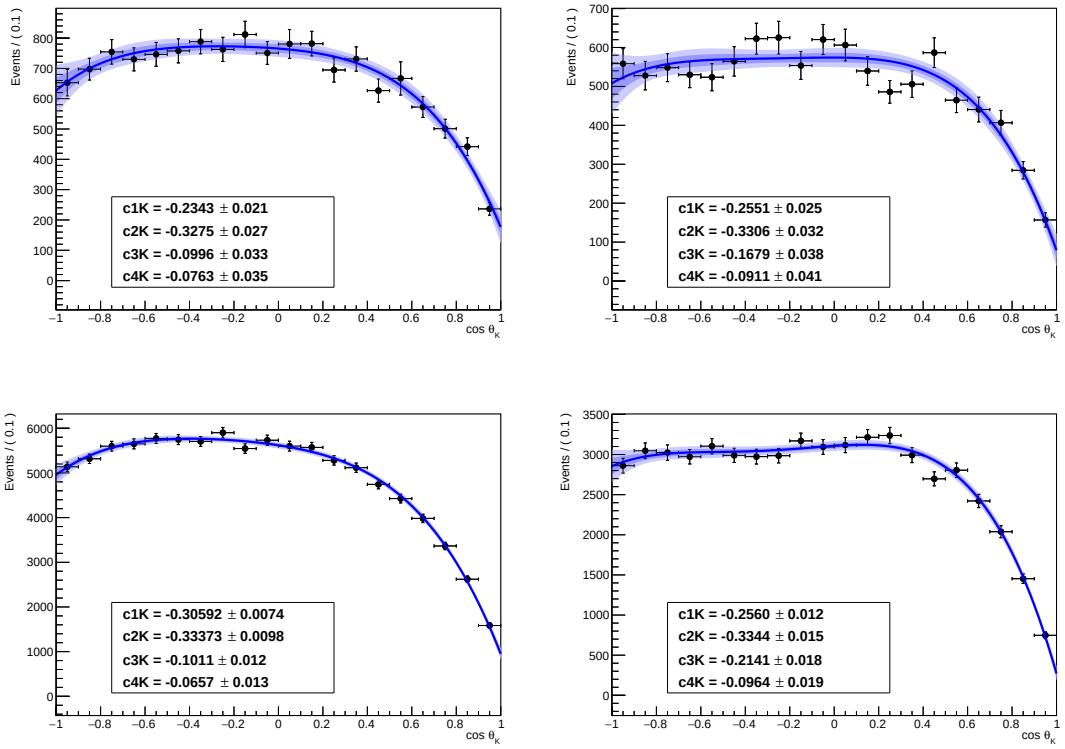


Figure 4.3: Fits to the phase space $B^0 \to K^{*0} e^+ e^-$ MC of the $\cos\theta_K$ distribution for R1 (top) and R2 (bottom) in the L0L (left) and L0I (right) trigger categories. The fitted function from Eq. (4.4) is shown in blue with its associated 1 σ and 2 σ error bands.

The angular acceptance fits to $\cos\theta_K$ as well as the fit parameters are shown in figure 4.3.

The ϕ angle is found to be unaffected by the reconstruction and selection effects. Nevertheless, the flatness of the ϕ angular acceptance is tested against $\cos 2\phi$, $\sin 2\phi$, $\cos\phi$ and $\sin\phi$ effects. These are the most dangerous effects because they could mimic NP with non-zero values of $A_T^{(2)}$ or A_T^{Im} (although any $\sin\phi$ or $\cos\phi$ effect would vanish due to the folding of the ϕ angle). Since $\cos 2\phi$, $\sin 2\phi$, $\sin\phi$ and $\cos\phi$ are orthogonal functions, they can be checked together. The ϕ angular acceptance PDF is thus given by

$$\varepsilon^{t,y}_{\mathcal{A},\phi}(\phi; c^{t,y}_{c1}, c^{t,y}_{s1}, c^{t,y}_{c}, c^{t,y}_{s}) = 1 + c^{t,y}_{c}\cos 2\phi + c^{t,y}_{s}\sin 2\phi + c_{c1}\sin\phi + c_{s1}\cos\phi \quad . \quad (4.7)$$

The angular acceptance fits to ϕ as well as the fit parameters are shown in figure 4.4. All parameters $c^{t,y}_{c(1)}$ and $c^{t,y}_{s(1)}$ are found to be compatible with zero within one

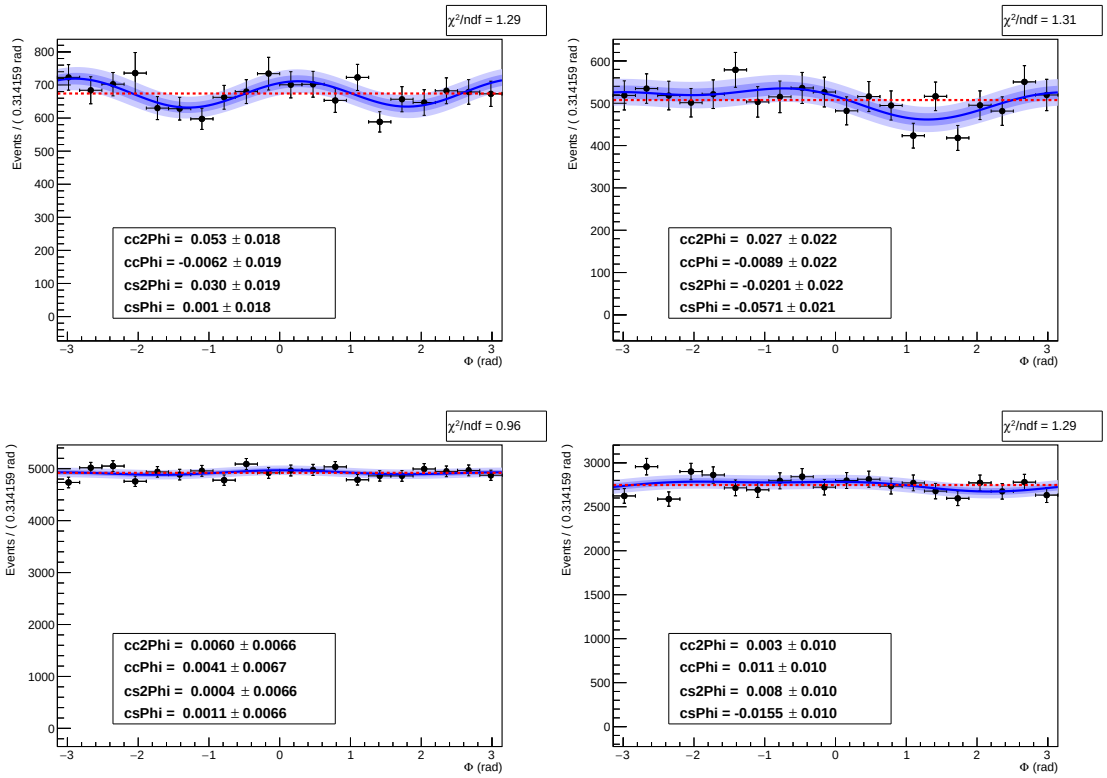


Figure 4.4: Fits to the phase space $B^0 \to K^{*0}e^+e^-$ MC of the ϕ distribution for R1 (top) and R2 (bottom) in the L0L (left) and L0I (right) trigger categories. The fitted function from Eq. (4.7) is shown in blue with its associated 1 σ and 2 σ error bands. The χ^2/ndf with respect to a flat distribution (shown in dotted red) is given in the top right corner.

to two σ. The χ^2/ndf with respect to a flat distribution is also computed and is found to be close to unity for all categories.

4.3 Angular modeling of backgrounds

The angular shape of the backgrounds has to be modeled to perform the four dimensional fit to the $m(K^+\pi^-e^+e^-)$ invariant mass and the three angles $\cos\theta_K$, $\cos\theta_\ell$ and $\tilde{\phi}$. The angular shapes of backgrounds are determined by various methods described in this section. Once they have been modeled, the angular shapes of backgrounds are all fixed when fitting to data (only the fractions of the various backgrounds are allowed to float).

4.3.1 Angular modeling of $B^0 \to K^{*0}\gamma(\to e^+e^-)$ background

To model the angular shape of the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ background, a dedicated ad hoc model is developed. Since in the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ decay the presence of the electrons is only due to the interaction of the real photon with a nucleus of the detector material, the $\cos\theta_\ell$ and $\tilde{\phi}$ distributions are not related to the physics of the $B^0 \to K^{*0}\gamma$ decay. The $\cos\theta_K$ dependent part of the full signal angular PDF is obtained by integrating over the $\cos\theta_\ell$ and $\tilde{\phi}$ angles, taking into account the angular acceptances. The obtained PDF is then fitted to the $\cos\theta_K$ distribution of the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ MC. The resulting fit is shown in figure 4.5. The fitted F_{L} is equal to $0.32 \pm 0.33\%$ which is compatible with zero, the expected value for a real photon.

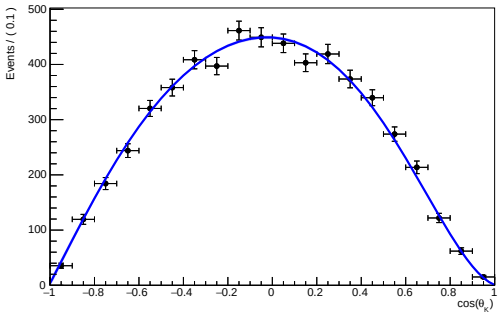


Figure 4.5: Fit to the $\cos\theta_K$ distribution of the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ MC for all years (R1+R2) and all trigger categories (L0L+L0I).

As shown in figure 4.6, it is checked that there is no remaining correlation between the $\cos\theta_\ell$ ($\tilde{\phi}$) and $\cos\theta_K$ angles. Thus the $\cos\theta_\ell$ and $\tilde{\phi}$ parts of the PDF are modeled with polynomials similar to the ones used to model the acceptance given by

$$\begin{aligned} p_l^{t,y}(a_1^{t,y}, a_2^{t,y}, a_4^{t,y}, \cos\theta_\ell) &= 1 + a_1^{t,y}\cos\theta_\ell + a_2^{t,y}\frac{1}{2}\left(3\cos{\theta_\ell}^2 - 1\right) \\ &\quad + a_4^{t,y}\frac{1}{8}\left(35\cos{\theta_\ell}^4 - 30\cos{\theta_\ell}^2 + 3\right) \quad , \\ p_{\tilde{\phi}}^{t,y}(C_s^{t,y}, C_c^{t,y}, \tilde{\phi}) &= 1 + C_s^{t,y}\sin(2\tilde{\phi}) + C_c^{t,y}\cos(2\tilde{\phi}) \quad . \end{aligned} \tag{4.8}$$

The resulting fits are shown in figure 4.7.

4.3.2 Angular modeling of combinatorial background

In order to model the angular shape of the combinatorial background, the upper mass sideband cannot be used directly. Indeed, what is called *combinatorial*

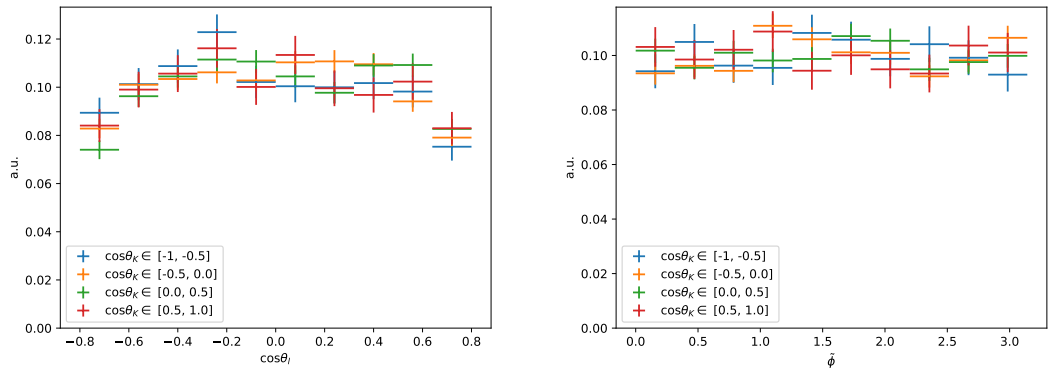


Figure 4.6: Distribution of the $\cos\theta_\ell$ (left) and $\tilde{\phi}$ (right) angles in several bins of $\cos\theta_K$ in the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ MC, for all years and trigger categories merged together.

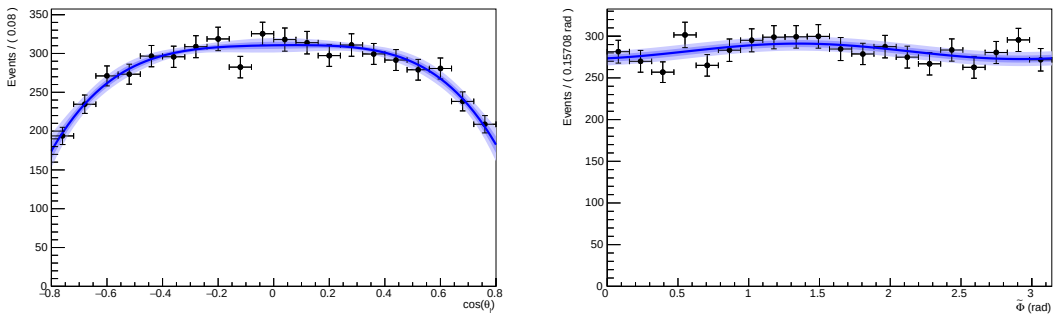


Figure 4.7: Fits of the ad hoc PDF to the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ for all years (R1+R2) and all trigger categories (L0L+L0I). The dark and light bands represent the 1σ and 2σ error respectively.

throughout this note is actually a mixture of semileptonic background and combinatorial background. These two types of background are expected to have different angular shapes. Since the semileptonic background is missing some particles, it is mostly present at low masses and thus absent from the upper mass sideband.
To cope with this, $B^0 \to K^{*0}e^+\mu^-$ data is used to model the combinatorial/semileptonic background. Indeed, due to lepton flavor conservation, this decay is forbidden in the SM. Therefore, when asking to reconstruct $B^0 \to K^{*0}e^+\mu^-$, all candidates found must have selected a random track to form the $e^+\mu^-$ pair. But since the same mass window as for the signal can be kept, the $B^0 \to K^{*0}e^+\mu^-$ data also contains the semileptonic background.

To validate the usage of $B^0 \to K^{*0}e^+\mu^-$ data to model the combinatorial /

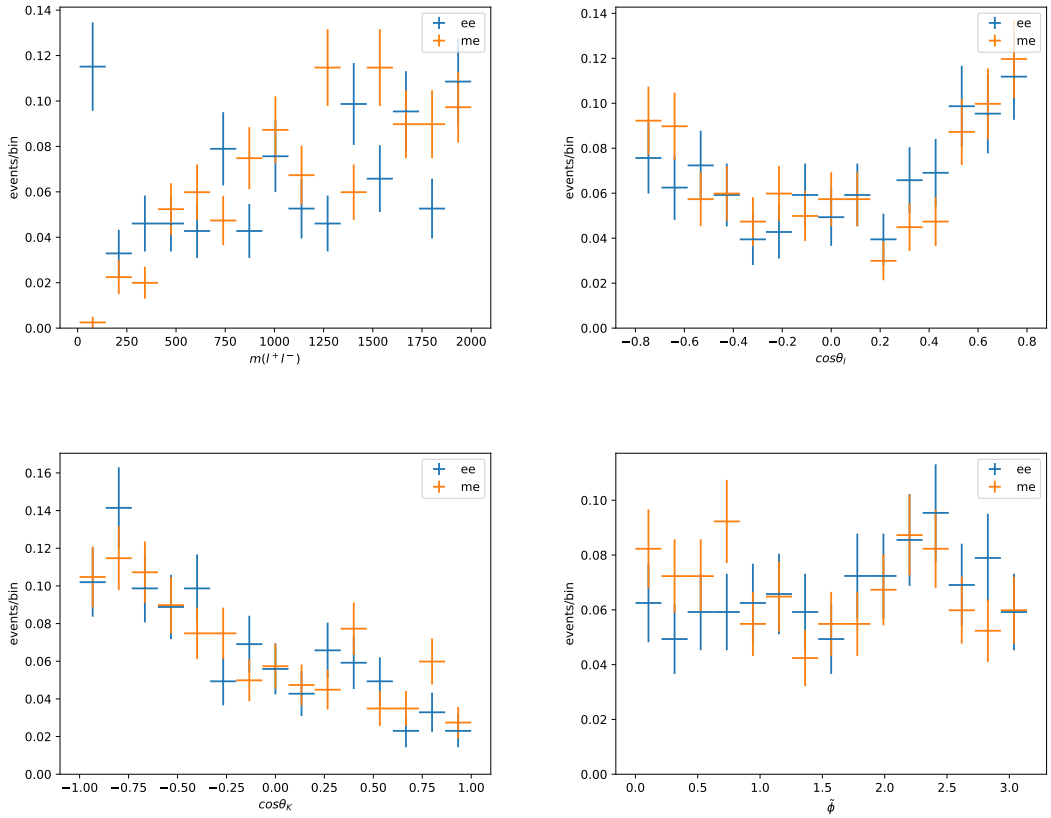


Figure 4.8: Comparison plots between $B^0 \to K^{*0}e^+\mu^-$ (blue) and $B^0 \to K^{*0}e^+e^-$ (orange) data (R1+R2) in the L0L trigger category, in the upper mass sideband $m(K^+\pi^-\ell^+\ell^-) > 5400\,\mathrm{MeV}/c^2$ and positive BDT output. From top left to bottom right, $m_{\ell^+\ell^-}$, $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$. The distributions are normalized to unity.

semileptonic background, its dilepton invariant mass, $m(K^+\pi^-\ell^+\ell^-)$ and angular distributions are compared to $B^0 \to K^{*0}e^+e^-$ combinatorial/semileptonic background in the corner of phase space where the latter can be isolated. For the $B^0 \to K^{*0}e^+\mu^-$ data, the dilepton mass $m_{\ell^+\ell^-}$ is defined as the dilepton invariant mass where the electron mass hypothesis is applied to the muon candidate. Moreover, since in $B^0 \to K^{*0}e^+\mu^-$ data one of the candidates is a muon, the L0L category is redefined as events where the electron has triggered the L0 electron line (but not the muon). The full selection (see section 3.5) is applied to both samples, but to gain statistics, the HOP cut is not applied and the $m_{\ell^+\ell^-}$ mass window is enlarged to $[10, 2000]\,\mathrm{MeV}/c^2$. It is checked that the HOP cut does not affect the angular distributions in $B^0 \to K^{*0}e^+\mu^-$ data.
In figure 4.8, the $m_{\ell^+\ell^-}$ and angular distributions of $B^0 \to K^{*0}e^+e^-$ and $B^0 \to$

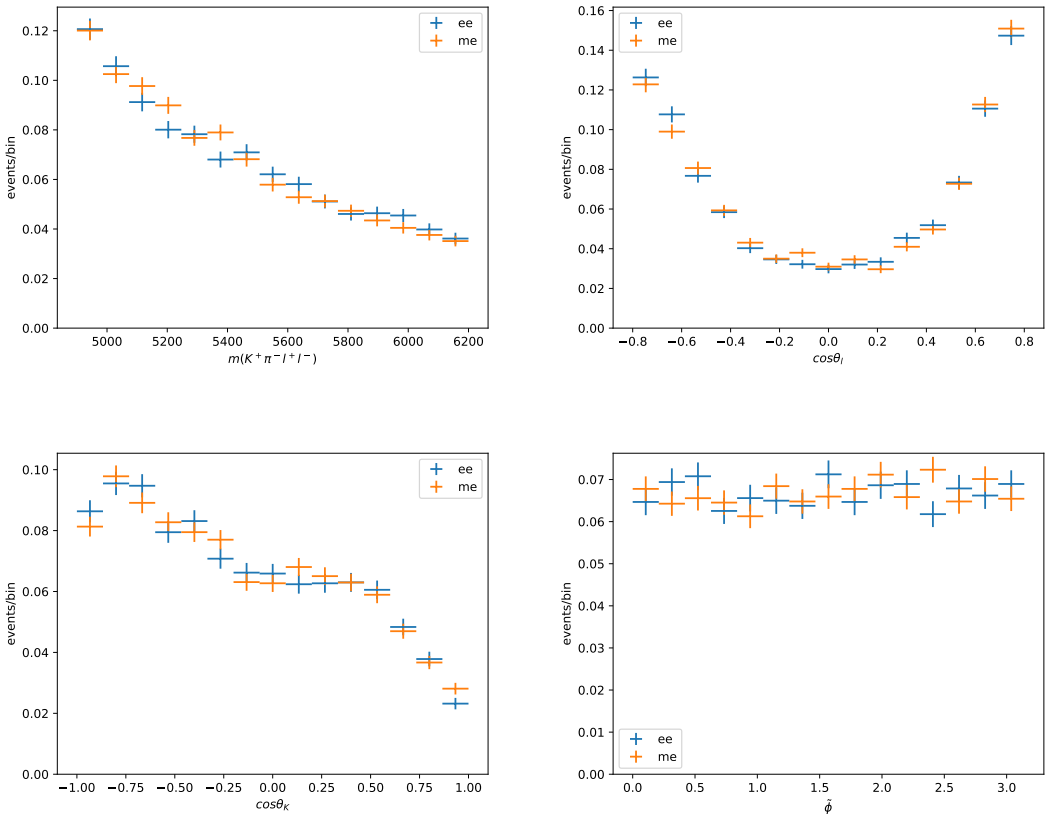


Figure 4.9: Comparison plots between $B^0 \to K^{*0}e^+\mu^-$ (blue) and $B^0 \to K^{*0}e^+e^-$ (orange) data (R1+R2) in the L0L trigger category, in the angular window $m(K^+\pi^-\ell^+\ell^-) \in [5000, 5400]\,\mathrm{MeV}/c^2$ and negative BDT output region. From top left to bottom right, $m(K^+\pi^-e^+e^-)$, $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$. The distributions are normalized to unity.

$K^{*0}e^+\mu^-$ data in the L0L trigger category are compared for $m(K^+\pi^-\ell^+\ell^-) > 5400\,\mathrm{MeV}/c^2$ and positive BDT output. The discrepancy in the first $m_{\ell^+\ell^-}$ bins is attributed to the absence of photon pole for $B^0 \to K^{*0}e^+\mu^-$. In figure 4.9, the invariant mass $m(K^+\pi^-e^+e^-)$ and the angular distributions of $B^0 \to K^{*0}e^+e^-$ and $B^0 \to K^{*0}e^+\mu^-$ data in the L0L trigger category are compared for $m(K^+\pi^-\ell^+\ell^-) \in [5000, 5400]\,\mathrm{MeV}/c^2$ and negative BDT output values. The $B^0 \to K^{*0}e^+\mu^-$ sample reproduces well the exponential mass shape of the $B^0 \to K^{*0}e^+e^-$ combinatorial/semileptonic background. In addition, both when comparing $B^0 \to K^{*0}e^+\mu^-$ and $B^0 \to K^{*0}e^+e^-$ data in the upper mass sideband and in the negative BDT output regions, the angular distributions are similar. In order to quantify their compatibility, the $\cos\theta_\ell$ and $\cos\theta_K$ distributions are fitted with an order 4 Legendre polynomial while the $\tilde{\phi}$ distributions are fitted with a $1 + c_c^{t,y}\cos 2\tilde{\phi} + c_s^{t,y}\sin 2\tilde{\phi}$

function. The fitted parameters obtained in the L0L trigger category are shown in figure 4.10. While the $\cos\theta_K$ and $\tilde{\phi}$ parameters are compatible, some small discrepancies are found for some of the parameters in the case of $\cos\theta_\ell$. Thus, in general, $B^0 \to K^{*0}e^+\mu^-$ is a good proxy to model the combinatorial/semileptonic background for the $\cos\theta_K$ and $\tilde{\phi}$ angles, while although very similar it does not perfectly reproduce the $\cos\theta_\ell$ shape. This is reflected in the systematic error associated to the combinatorial background angular shape modeling (see section 4.5.2).

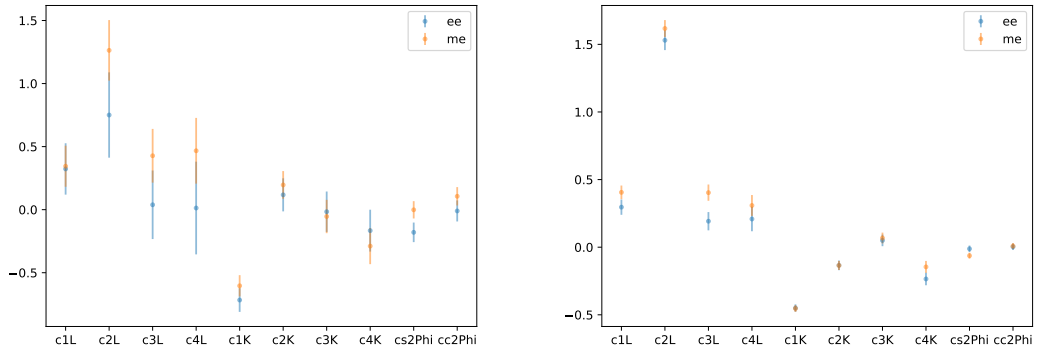


Figure 4.10: Fit parameters obtained on $B^0 \to K^{*0}e^+\mu^-$ (blue) and $B^0 \to K^{*0}e^+e^-$ (orange) data (R1+R2) in the L0L trigger category, in the angular window $m(K^+\pi^-\ell^+\ell^-) \in [5000, 5400]\,\text{MeV}/c^2$ and negative BDT output region (left) and in the upper mass sideband $m(K^+\pi^-\ell^+\ell^-) > 5400\,\text{MeV}/c^2$ and positive BDT output (right). The $\cos\theta_\ell$ and $\cos\theta_K$ distributions are fitted with order 4 Legendre polynomials while the $\tilde{\phi}$ distributions are fitted with a $1 + c_c^{t,y}\cos 2\tilde{\phi} + c_s^{t,y}\sin 2\tilde{\phi}$ function.

Ideally, to model the angular shape of the combinatorial/semileptonic background, one would like to use exactly the same selection as the one applied to $B^0 \to K^{*0}e^+e^-$ data. However, doing so the statistics is very limited and would result in very large uncertainties in the angular shape determination. To cope with this issue, since the angular distributions in $B^0 \to K^{*0}e^+\mu^-$ are similar between R1 and R2, the two data samples are merged together. In addition, it is considered to relax three cuts: the dilepton mass, the $m(K^+\pi^-\ell^+\ell^-)$ and the BDT cuts. For each angle, the following sets of cuts are considered:

- $m_{\ell^+\ell^-}$: $[10, 500]$, $[500, 1000]$ and $[1000, 2000]$ MeV/c^2
- $m(K^+\pi^-\ell^+\ell^-)$: $[4900, 5400]$ and $[5400, 6200]$ MeV/c^2
- BDT cut: BDT < 0, BDT > 0 and nominal BDT cut

For each set of cuts, the angular distributions are fitted similarly as above and the fitted parameters are compared. The fitted parameters obtained in each case are shown in figures 4.11, 4.12 and 4.13 for the L0L trigger category. The best trade-off of statistics versus compatibility of the parameters is obtained with the region $m_{\ell^+\ell^-} \in [10, 1000]$ MeV/c^2, $m(K^+\pi^-\ell^+\ell^-) \in [4900, 5400]$ MeV/c^2 and BDT > 0.

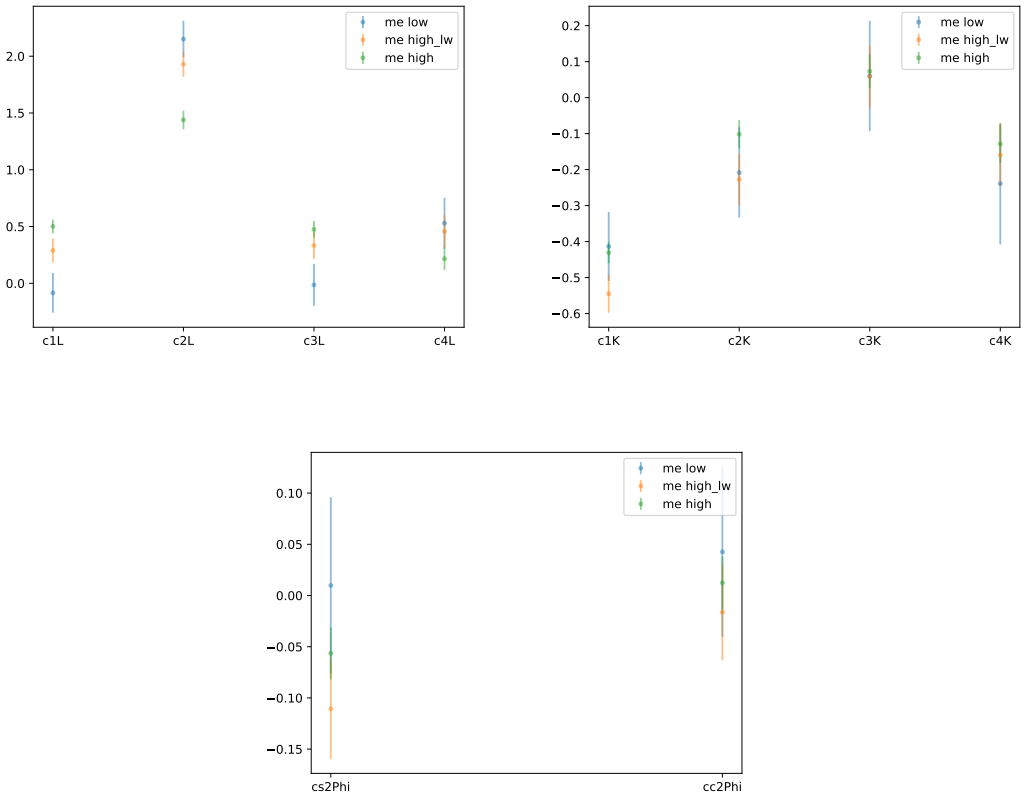


Figure 4.11: Fitted parameters obtained on $B^0 \to K^{*0}e^+\mu^-$ data (R1+R2) for the three $m_{\ell^+\ell^-}$ regions $[10, 500]$ (blue), $[500, 1000]$ (orange) and $[1000, 2000]$ (green) for the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ (from left to right).

Each of the three angles is thus fitted in this region and shown in figure 4.14. The obtained parameters define the angular shapes of the combinatorial/semileptonic background for the angular fit.

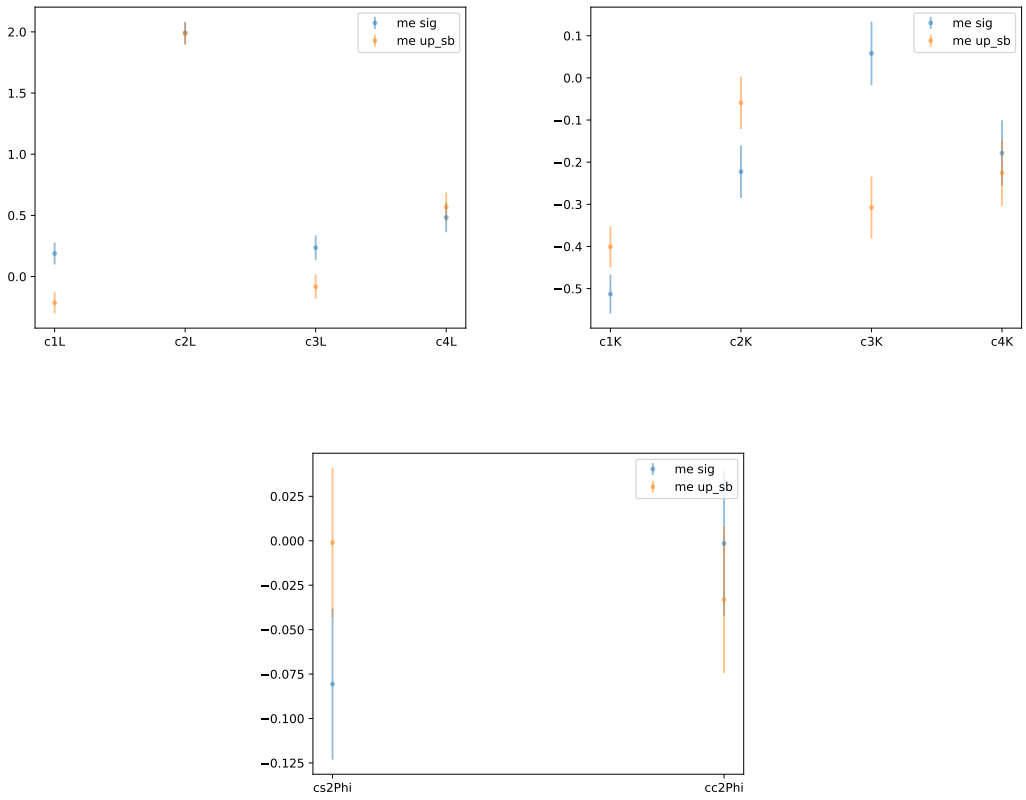


Figure 4.12: Fitted parameters obtained on $B^0 \to K^{*0} e^+ \mu^-$ data (R1+R2) for the three $m(K^+\pi^-\ell^+\ell^-)$ regions $[4900, 5400]$ (blue) and $[5400, 6200]$ MeV/c^2 (orange) for the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ (from left to right).

4.3.3 Angular modeling of partially reconstructed background

As for the mass shape, the angular shape of partially reconstructed background is modeled using $B^+ \to K_1(1270)(\to K^+\pi^- X)e^+e^-$ MC. The full selection is applied, including the corrections discussed in section 3.2. The full angular signal PDF (including angular acceptances) is fitted to the $B^+ \to K_1(1270)(\to K^+\pi^- X)e^+e^-$ MC, in a simultaneous fit to both Runs and trigger categories. The fits are shown in figure 4.15 while the fit result parameters are listed in table 4.2. The obtained shapes are added to the full angular PDF.

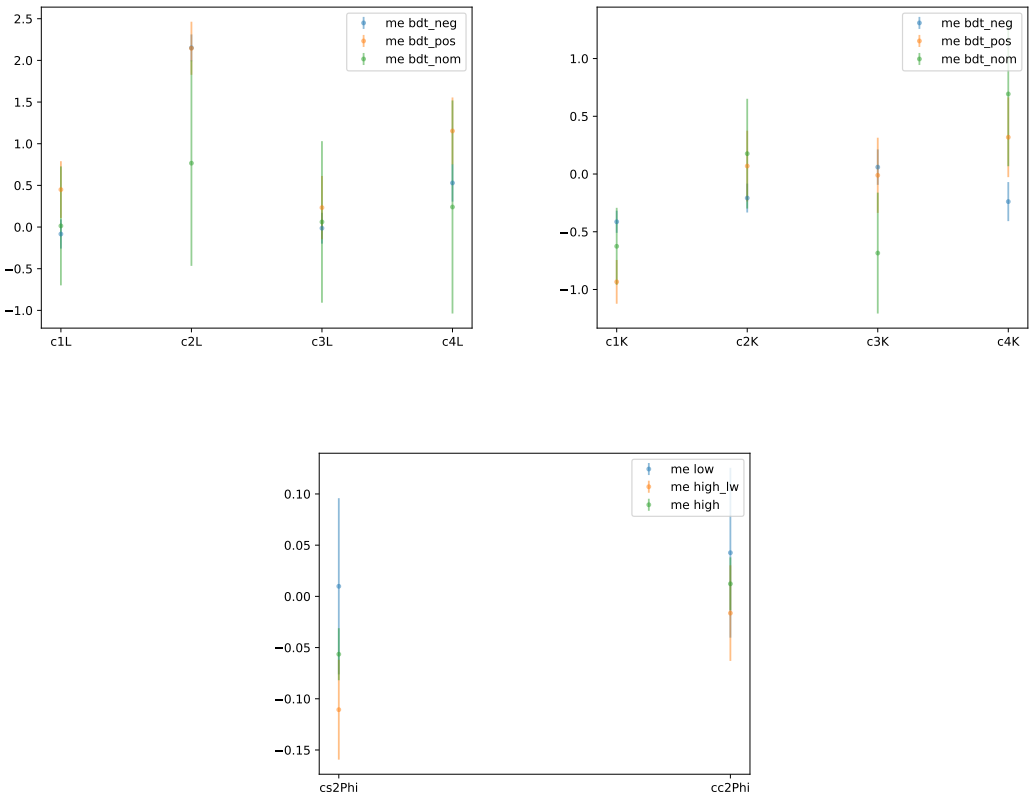


Figure 4.13: Fitted parameters obtained on $B^0 \to K^{*0} e^+ \mu^-$ data (R1+R2) for the three BDT output regions negative BDT output (blue), positive BDT output (orange) and nominal BDT cuts (green) for the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ (from left to right).

Parameter	Value
$A_T^{(2)}$	$(1.09 \pm 5.08) \times 10^{-2}$
A_T^{Im}	$(2.98 \pm 5.16) \times 10^{-2}$
A_T^{Re}	$(5.21 \pm 3.70) \times 10^{-2}$
F_{L}	$(16.78 \pm 1.39) \times 10^{-2}$

Table 4.2: Fitted parameters on the full $B^+ \to K_1(1270)(\to K^+\pi^- X) e^+ e^-$ MC dataset.

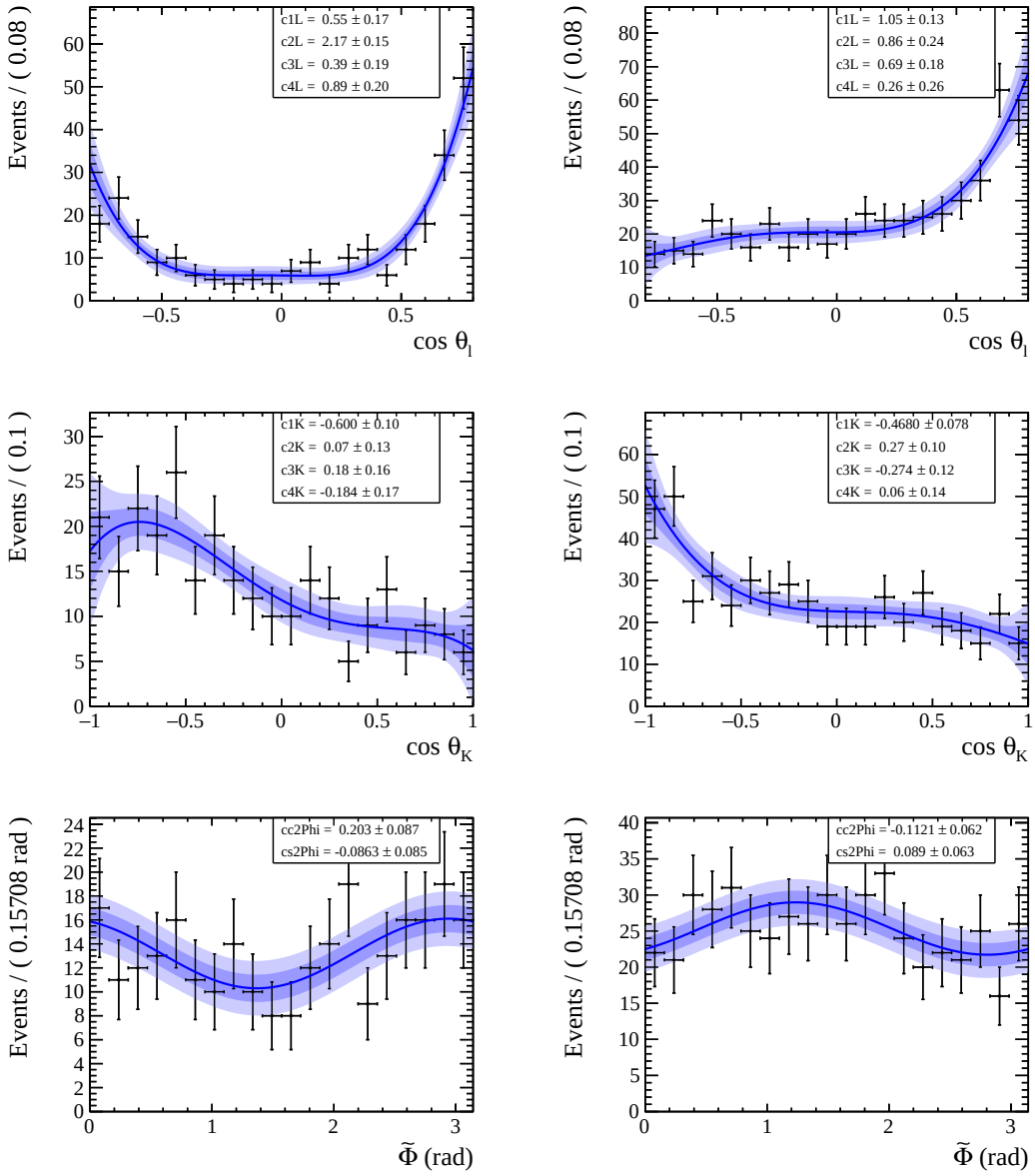


Figure 4.14: Fits to $B^0 \to K^{*0}e^+\mu^-$ data (R1+R2) in the L0L (left) and L0I (right) trigger categories. The angular distributions of the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ (left to right) are fitted. The dark and light bands represent the 1σ and 2σ error respectively.

4.3.4 Angular modeling of $B^0 \to K^{*0}\eta$ and $B^0 \to K^{*0}\pi^0$ backgrounds

The angular shape of the $B^0 \to K^{*0}\pi^0/\eta(\to e^+e^-\gamma)$ backgrounds are modeled using $B^0 \to K^{*0}\pi^0(\to e^+e^-\gamma)$ MC. The $\cos\theta_\ell$ and $\tilde{\phi}$ angles are modeled using

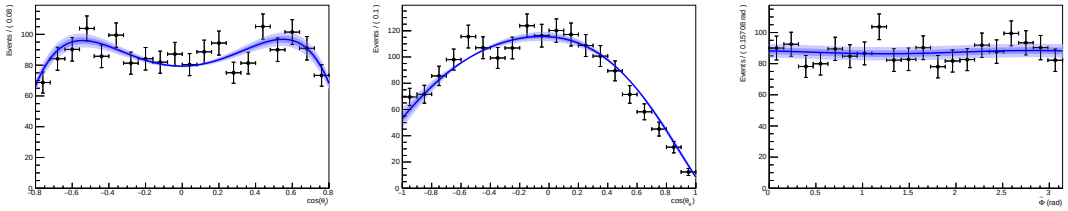


Figure 4.15: Fits of full angular signal PDF to the $B^+ \to K_1(1270)(\to K^+\pi^- X)e^+e^-$ for all years (R1+R2) and all trigger categories (L0L+L0I). The dark and light bands represent the 1σ and 2σ error respectively.

similar polynomials as the angular acceptance and the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ background given in eq. (4.8). The $\cos\theta_K$ angle is modeled using Chebychev polynomials of order five. This type of polynomial is used because it allows to ensure a positive value of the PDF. Indeed, since this background has $F_{\mathrm{L}} = 1$ due to the π^0 spin, the shape of the background goes roughly as x^2 in the vicinity of $\cos\theta_K \to 0$. To ensure that the PDF is equal to zero at $\cos\theta_K = 0$ the order four coefficient is constrained such that

$$d_k = b_k - 1 \qquad , \tag{4.9}$$

where d_k and b_k are the order four and two coefficients of the Chebychev polynomial respectively. The fits to the $B^0 \to K^{*0}\pi^0(\to e^+e^-\gamma)$ MC are shown in figure 4.16.

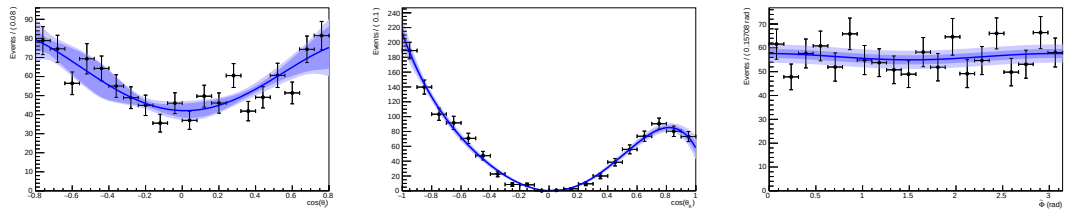


Figure 4.16: Fits to the $B^0 \to K^{*0}\pi^0(\to e^+e^-\gamma)$ for all years (R1+R2) and all trigger categories (L0L+L0I). The dark and light bands represent the 1σ and 2σ error respectively.

For the $B^0 \to K^{*0}\eta(\to e^+e^-\gamma)$ decay however, since there is no EvtGen model for this decay, the generated angular distributions in the MC are non physical. The $B^0 \to K^{*0}\eta(\to e^+e^-\gamma)$ MC can thus not be used to model the angular shape of this decay. Nevertheless, since the η and the π^0 mesons have the same spin and

parity, their angular distributions are expected to be very similar. Therefore, the angular shape extracted from the π^0 MC is also used for the η Dalitz decay.

4.4 Angular and mass fit validation

4.4.1 Toys

The stability of the fit is tested with toy datasets. Ten thousand datasets of pseudo-data are generated with the signal and background fractions given by the fit to data. For the various backgrounds, the angular PDFs are generated with the shapes discussed in section 4.3, while the signal is generated with the (blinded) central values of the fit to data. A second set of toys is generated with SM values computed with Flavio [12]: $F_{\rm L} = 0.051$, $A_T^{(2)} = 0.032$, $A_T^{Re} = -0.024$ and $A_T^{Im} = -0.002$. Each toy dataset is generated with the actual number of events in real LHCb data (873 in the very low q^2 bin and 6129 in the gamma q^2 bin).
Then the exact same fit procedure as the one carried out on real data is applied to each one of these toy datasets. First, the $m(K^+\pi^-\ell^+\ell^-)$ is fitted over the large mass window $[4500, 6200]$ MeV/c^2. Then, the obtained fractions for the different components of the mass fit are Gaussian constrained. To finish, the four dimensional PDF from Eq. (4.3) is fitted to the mass and the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ simultaneously. The fitted value, error and pull distributions for each of the four free parameters of the fit to the toy datasets generated with the fitted values in data are shown in figure 4.17, and fitted with a Gaussian function. The plots for the toys with SM values are given in Appendix E. No large bias is observed and the width of the pulls is compatible with unity, showing correct error estimation. For $F_{\rm L}$ however, a small bias of 0.003 is observed, which is due to the proximity to the unphysical $F_{\rm L} < 0$ region. The fitted $F_{\rm L}$ value in data is therefore corrected accordingly and a 100% systeamtic uncertainty is attributed (see section 4.5). The size of this systematic is nevertheless much smaller ($< 10\%$) than the statistical error on the $F_{\rm L}$ parameter.

4.4.2 Angular fit to $B^0 \to K^{*0}e^+e^-$ MC

In order to validate the angular fit procedure, the angular analysis is performed on $B^0 \to K^{*0}e^+e^-$ MC. The full angular PDF given in Eq.(4.2), including the angular acceptance, is fitted to this sample after full reconstruction and selection to extract the four observables $F_{\rm L}$, $A_T^{(2)}$, A_T^{Re} and A_T^{Im}. The fits are shown in figure 4.18.

The values of the four physical observables extracted from the fit to the $B^0 \to K^{*0}e^+e^-$ MC are then compared to the generated values of the MC. The latter are

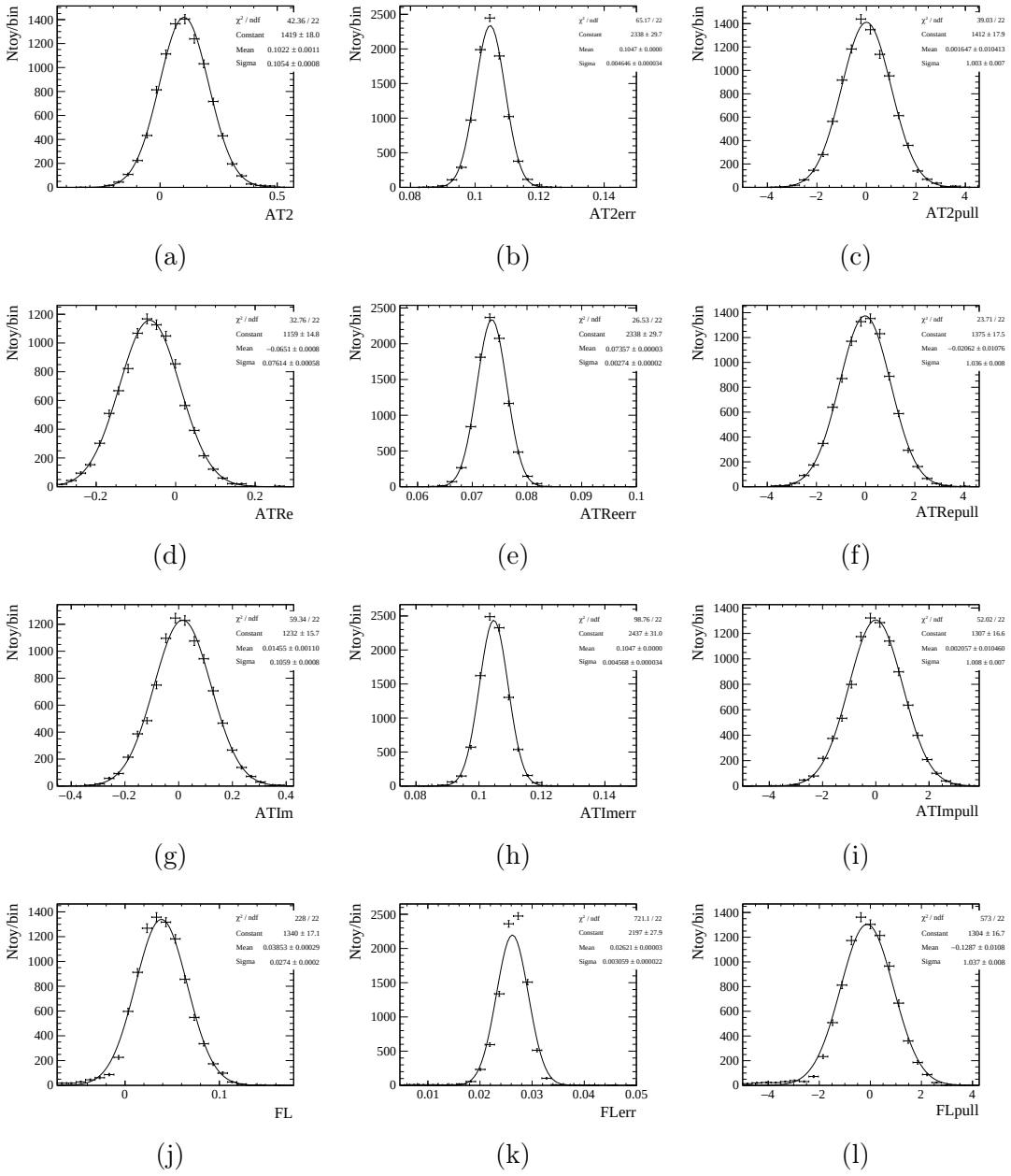


Figure 4.17: Fitted values (left), errors (middle) and pulls (right) of the 10^4 toy datasets for the four parameters $A_T^{(2)}$, A_T^{Re}, A_T^{Im} and $F_{\rm L}$ (from top to bottom). The four parameters ahve been generated with the (blinded) central values of the fit to data.

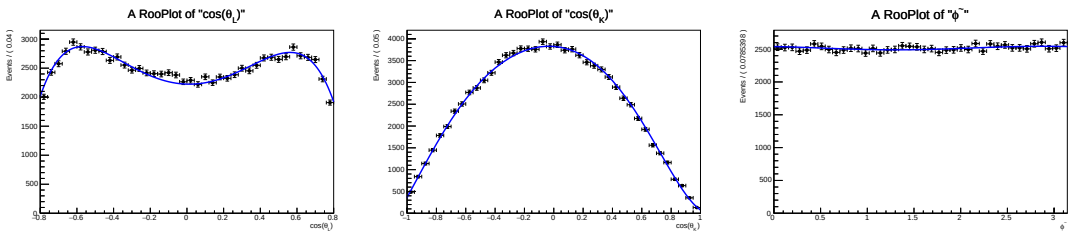


Figure 4.18: Fits to the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ to $B^0 \to K^{*0}e^+e^-$ MC after full reconstruction and selection, taking into account the angular acceptance.

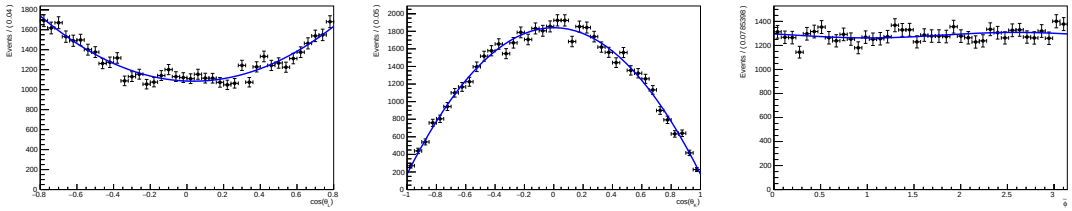


Figure 4.19: Fits to the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ to $B^0 \to K^{*0}e^+e^-$ generator level MC, reweighted to match the q^2 distribution of the fully reconstructed and selected $B^0 \to K^{*0}e^+e^-$ MC.

extracted from a fit to generator level $B^0 \to K^{*0}e^+e^-$ MC without any cut (see figure 4.19). However the differential decay width is integrated over q^2 in the full angular PDF, while the generator level MC is generated over the full phase space. Given the narrow q^2 range of the analysis, this results in drastically different q^2 distributions which in turns affects the values of the measured parameters, especially $F_{\rm L}$ which is very q^2 dependent. In order to be able to compare the fit to generator level MC with the one to reconstructed and selected MC, the generator level MC is reweighted such that its $m_{ee}^{\rm true}$ distribution matches that of the reconstructed MC after the full selection has been applied (see figure 4.20). Moreover, for consistency, none of the MC corrections discussed in section 3.2 is applied for this crosscheck.

Both fit results are compared in table 4.3. All four parameters are compatible within less than two sigmas, which validates the angular fit procedure, including the angular acceptance modeling. In particular, this test has a 0.015 precision level on $A_T^{(2)}$ which is seven times better than the precision expected in data.

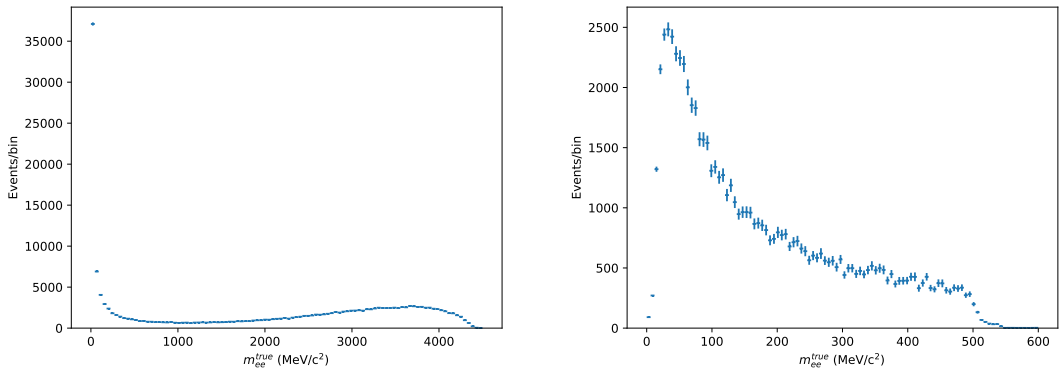


Figure 4.20: True dielectron mass distribution in the $B^0 \to K^{*0}e^+e^-$ generator level MC before reweighting (left) and after reweighting to match the distribution of the fully reconstructed and selected $B^0 \to K^{*0}e^+e^-$ MC.

	F_{L}	A_T^{Re}	$A_T^{(2)}$	A_T^{Im}
gen-level fit	$(3.55 \pm 0.24)\%$	-0.032 ± 0.009	0.012 ± 0.014	-0.026 ± 0.014
reco-level fit	$(3.95 \pm 0.14)\%$	-0.019 ± 0.005	-0.003 ± 0.006	-0.022 ± 0.006

Table 4.3: Comparison of the fit parameters extracted from the fit to generator level with respect to fully reconstructed and selected $B^0 \to K^{*0}e^+e^-$ MC.

4.4.3 Angular fit to $B^0 \to K^{*0}\gamma(\to e^+e^-)$ MC and data

An additional important validation of the fit and the angular acceptance modeling is made on $B^0 \to K^{*0}\gamma(\to e^+e^-)$ MC and data. Since in the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ decay the presence of the electrons is only due to the interaction of the real photon with a nucleus of the detector material, the $\cos\theta_\ell$ and $\tilde{\phi}$ distributions are not related to the physics of the $B^0 \to K^{*0}\gamma$ decay. Therefore, the full signal angular PDF, which includes the angular acceptance, is integrated over $\cos\theta_\ell$ and $\tilde{\phi}$. The resulting PDF is thus only sensitive to F_{L}. Due to the very low statistics in the phase space $B^0 \to K^{*0}e^+e^-$ MC in the region $[0, 10]$ MeV/c^2, the same $\cos\theta_K$ angular acceptance as in the bin $[10, 500]$ MeV/c^2 is taken.
First, the fit is performed on $B^0 \to K^{*0}\gamma(\to e^+e^-)$ MC (see section 4.3.1). The result of the fit yields $F_{\mathrm{L}} = 0.32 \pm 0.33\%$, which is compatible with the expected value of $F_{\mathrm{L}} = 0$ in the case of a real photon.

Then, the angular analysis is performed on data in the gamma-q^2 bin, which predominantly contains $B^0 \to K^{*0}\gamma(\to e^+e^-)$ events. The angular shapes of the backgrounds are modeled as follows:

- $B^0 \to K^{*0}\eta(\to \gamma(\to e^+e^-)\gamma)$ and $B^0 \to K^{*0}\pi^0(\to \gamma(\to e^+e^-)\gamma)$: the same ad hoc model as for the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ background in the very-low-q^2 bin (see section 4.3.1) is used. The PDF is then fitted to the $B^0 \to K^{*0}\eta(\to \gamma e^+e^-)$ and $B^0 \to K^{*0}\pi^0(\to \gamma e^+e^-)$ MCs to extract their respective angular shapes.

- Partially reconstructed background: the same procedure as for the signal $B^0 \to K^{*0}\gamma(\to e^+e^-)$ is applied. The resulting PDF is then fitted to the $\cos\theta_K$ distribution of the $B^+ \to K_1(1270)(\to K^+\pi^-X)e^+e^-$ MC in the gamma-q^2 bin.

- Combinatorial background: since the $\cos\theta_K$ distribution of the $B^0 \to K^{*0}e^+\mu^-$ data does not vary much with m_{ee} (see section 4.3.2) and due to the very low amount of $B^0 \to K^{*0}e^+\mu^-$ events below 10 MeV/c^2, the same angular shape than the one of the very-low-q^2 bin is used.

Up to the fact that the full angular PDF is integrated over the angles $\cos\theta_\ell$ and $\tilde{\phi}$, the exact same procedure as for the fit to $B^0 \to K^{*0}e^+e^-$ data in the very-low-q^2 is applied. The mass and angular fits are shown in figure 4.21. The result of the fit yields $F_{\rm L} = 0^{+0.66}_{-0}\%$.

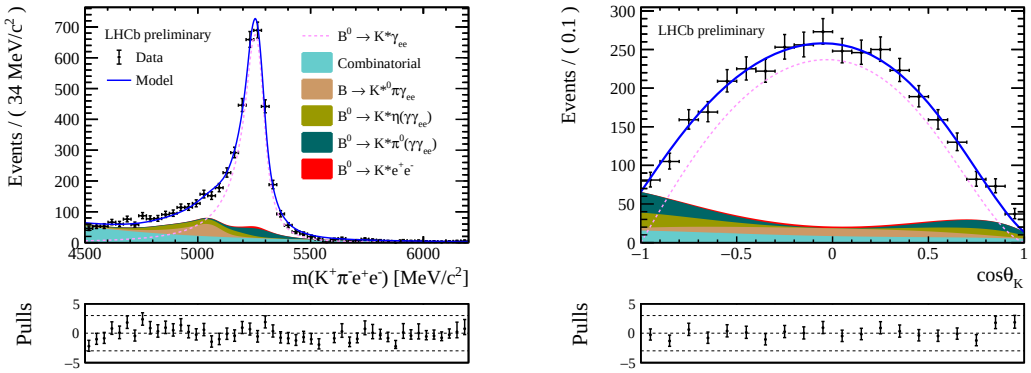


Figure 4.21: Fits to the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ data, projected on $m(K^+\pi^-e^+e^-)$ (left) and $\cos\theta_K$ (right) for all years (R1+R2) and trigger categories (L0L+L0I). The various components of the fits are given in the legend in the left plots.

In order to validate this result, toys are generated fixing the signal $B^0 \to K^{*0}\gamma(\to e^+e^-)$ $F_{\rm L}$ value to zero. About 95 % of the toys return a fitted $F_{\rm L}$ below 1.5%, while $\sim$ 60% hit the physical boundary at $F_{\rm L} = 0$. The result found on $B^0 \to K^{*0}\gamma(\to e^+e^-)$ is thus compatible with a true $F_{\rm L}$ value equal to zero.

4.5 Systematic uncertainties

Two main types of systematic uncertainties are considered: the uncertainties due to the statistics of the samples used for the modeling of angular shapes and angular acceptances, and the uncertainties due to the modeling itself. In both cases, unless stated otherwise, the systematic uncertainties are estimated using toys.
To evaluate systematic uncertainties related to the size of the samples used for the modeling, a bootstrapping method is used. This method consists in creating 1000 re-sampled datasets. Each re-sampled dataset is created by replacing each event by a randomly chosen event from the same sample, where the same event can be selected more than once. Then, the full fit is performed 1000 times, but using each time the re-sampled dataset for the modeling of the angular shape or acceptance. To avoid any fluctuation related to the available statistics in data, the fits are performed on toy datasets generated with the fitted values of the nominal fit to data but with 10 times more statistics. The pull distribution of the four parameters of interest $F_{\rm L}$, $A_T^{(2)}$, A_T^{Re} and A_T^{Im} for the 1000 toys is then fitted by a Gaussian. Let the systematic error be expressed as a fraction of the statistical error as

$$\sigma_{sys} = x \cdot \sigma_{stat} \quad . \tag{4.10}$$

The width of the Gaussian fit to the pull distribution is equal to

$$\sigma_{tot} = \sqrt{1 + x^2} \quad . \tag{4.11}$$

Given that having generated 1000 toys σ_{tot} is estimated with a ~ 0.026 error and that the toys have been generated with 10 times the nominal statistics, the bootstrapping method is sensitive to systematic errors larger than $\sim 7\%$ of the statistical error.

To evaluate systematic errors related to the modeling of angular shapes or angular acceptances, alternative models are considered. Similarly as above, toys are generated with the fitted values of the nominal fit to data but with 10 times more statistics. The full analysis is then performed on these toys generated with the alternative models, but fitting the toy datasets with the nominal model. A Gaussian fit is then performed to the distributions of the four parameters of interest $F_{\rm L}$, $A_T^{(2)}$, A_T^{Re} and A_T^{Im} to extract the deviations from the nominal values.
The different systematic uncertainties considered are detailed below and are summarized in table 4.4. The total systematic error is well below the statistical error, especially for the two parameters sensitive to the photon polarization $A_T^{(2)}$ and A_T^{Im} where the systematic error is about 10% of the statistical error.

Source of systematic	$\sigma A_T^{(2)}$	σA_T^{Im}	σA_T^{Re}	σF_{L}
Acceptance sample size	< 0.007	< 0.007	< 0.007	< 0.003
Acceptance model	0.004	0.001	0.008	0.001
Comb/SL sample size	< 0.007	< 0.007	< 0.007	< 0.003
Comb/SL $q2$ cut	0.009	0.003	0.004	0.001
Comb/SL BDT cut	0.007	0.003	0.003	0.005
Comb/SL merging Runs	0.002	0.003	0.004	0.001
Part. reco. model	0.001	0.003	0.002	0.001
η/π^0 model	0.0004	0.0001	0.002	0.01
ϕ resolution	-0.004	-0.001	-	-
$K^{*0} - \overline{K}^{*0}$ mis-ID	-	0.0002	0.0001	0.00004
MC corrections	0.003	0.001	0.003	0.007
Data/MC diff. in mass shape	0.002	0.002	0.004	0.001
Fit bias	~ 0	~ 0	~ 0	-0.003
Total	$^{+0.016}_{-0.017}$	0.012	0.015	0.014
Statistical error	0.103	0.102	0.077	0.026

Table 4.4: Summary of all systematic uncertainties.

4.5.1 Angular acceptance

The angular acceptance is evaluated using phase space $B^0 \to K^{*0}e^+e^-$ MC (see section 4.2). Three types of systematic uncertainties are considered for the angular acceptance: the uncertainty related to the size of the MC samples, the uncertainty related to the various corrections applied to the MC and the uncertainty related to the function used to model the acceptance.
To evaluate the uncertainty related to the MC corrections applied to the phase space $B^0 \to K^{*0}e^+e^-$ MC sample, an alternative modeling is considered using bare (uncorrected) MC.
To evaluate the uncertainty related to the function used to model the angular acceptance, an alternative modeling using Legendre polynomial of order 6 instead of 4 is considered.

4.5.2 Combinatorial background modeling

The angular shape of the combinatorial/semileptonic background is modeled using $B^0 \to K^{*0}e^+\mu^-$ data (see section 4.3.2). Two types of systematic uncertainties are considered for the angular modeling of combinatorial/semileptonic background: the uncertainty related to the size of the data samples and the uncertainty related to the modeling itself.
To evaluate the uncertainty related to the modeling itself, three alternative models are considered. The first alternative model is linked to the usage of a relaxed BDT cut due to statistical issues in the $B^0 \to K^{*0}e^+\mu^-$ sample. Alternative angular shapes are extracted using the nominal BDT cut of the analysis. Since the BDT mostly kills combinatorial background but is not very effective at killing semileptonic background, this alternative model is testing the effect of the mis-modeling of the ratio of combinatorial background with respect to semileptonic background. The second alternative model is linked to the usage of an enlarged q^2 bin due to statistical issues in the $B^0 \to K^{*0}e^+\mu^-$ sample. Similarly as above, alternative angular shapes are extracted using the nominal q^2 bin of the analysis. To finish, a third alternative model is considered, where the angular shape is fitted separately for R1 and R2. Thus, this alternative model is testing the effect of merging the two Runs together in the nominal modeling.

4.5.3 $B^0 \to K^{*0}\eta/\pi^0(\to e^+e^-\gamma)$ backgrounds modeling

The angular shape of the $B^0 \to K^{*0}\eta/\pi^0(\to e^+e^-\gamma)$ are modeled from MC (see section 4.3.4). Since the size of the MC samples is at least two orders of magnitude higher than the yields fitted in data, no systematic uncertainty related to the size of the MC sample is assigned.

To evaluate the uncertainty related to the modeling of this background, an alternative modeling is considered, using the ad hoc model used for the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ background (see section 4.3.1).

4.5.4 Partially reconstructed background modeling

The angular shape of the partially reconstructed background is modeled using $B^+ \to K_1(1270)(\to K^+\pi^-X)e^+e^-$ MC (see section 4.3.3). Since the size of the MC samples is at least two orders of magnitude higher than the yields fitted in data, no systematic uncertainty related to the size of the MC sample is assigned.

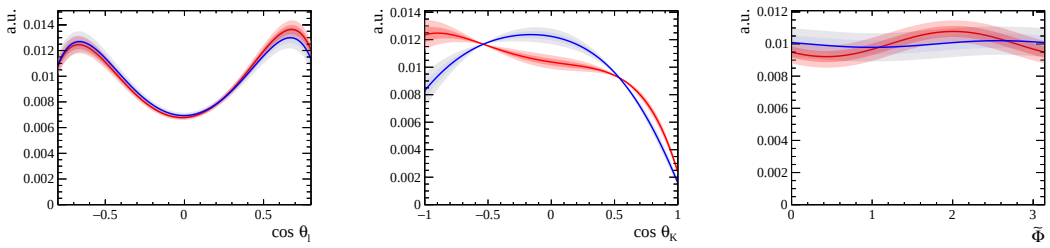


Figure 4.22: Angular modeling of the partially reconstructed background in the R2 L0L category. The nominal model is shown in blue while the alternative model is shown in red. The 1σ and 2σ errors are represented by the darker and lighter bands respectively.

The uncertainty due to the modeling is evaluated by considering an alternative model. To take into account possible variations in the angular shape due to other K resonances than the $K_1(1270)$, a dedicated sample is considered. A sample of phase space $B^+ \to K^+\pi^-\pi^+e^+e^-$ MC is used. This sample is then reweighted in the $K^+\pi^-\pi^+$ Dalitz plane to match the distribution in $B^+ \to J/\psi(\to e^+e^-)K^{res}(\to K^+\pi^-\pi^+)$ data, where K^{res} is any K resonance. The angular distributions of the reweighted phase space $B^+ \to K^+\pi^-\pi^+e^+e^-$ MC are then fitted using the full signal PDF, as in the nominal fit. Figure 4.22 illustrates the difference between the nominal and the alternative models with the example of the R2 L0L category.

4.5.5 Resolution of the $\tilde{\phi}$ angle

As explained in section 1.4.2, fitting the data on $\tilde{\phi}$ to extract $A_T^{(2)}$ and A_T^{Im} effectively reduces to the extraction of the amplitude of a sinusoidal function. Therefore, the poor $\tilde{\phi}$ resolution at the lower end of the q^2 bin flattens the distribution and the extracted amplitude is lower, leading to a bias towards lower measured $A_T^{(2)}$ and A_T^{Im}. In order to estimate the bias to the measurement of $A_T^{(2)}$ (and A_T^{Im}), toy data samples of the three angles ($\cos\theta_\ell$, $\cos\theta_K$, $\tilde{\phi}$) are generated from the full

signal PDF of eq. (1.35) bin per bin of dielectron mass. Assuming that $m_{ee}^{B^0}$ is roughly equal to m_{ee}^{true}, the signal yield in a bin of dielectron mass $[m_1, m_2]$ is given by

$$N(m_1, m_2) = \int_{m_1^2}^{m_2^2} \frac{d\mathcal{B}(B^0 \to K^{*0} e^+ e^-)}{dq^2_{\text{true}}} dq^2_{\text{true}} \times \epsilon(m_1, m_2) \times L \times 2\sigma_{b\bar{b}} \times f_d \quad , \tag{4.12}$$

where L is the integrated luminosity, $\sigma_{b\bar{b}}$ is the $b\bar{b}$ cross section and f_d is the B^0 production fraction at LHCb. The integrated differential branching fraction is computed with `Flavio` [12]. The efficiency of the full selection in the bin $\epsilon(m_1, m_2)$ is computed from $B^0 \to K^{*0} e^+ e^-$ simulation. To avoid statistical fluctuations, the toys are generated with 10 times the statistics in data. The number of generated events as a function of the diectron mass is shown in figure 4.23.

For each m_{ee} bin, $N(m_1, m_2)$ events are generated with the (averaged over the bin) SM value of F_L, A_T^{Im} and A_T^{Re}. $A_T^{(2)}$ is generated with an (unrealistically high) value of 0.5. This is done to prevent any underestimation of the effect in case the true value of $A_T^{(2)}$ is higher than the prediction of the SM. Then, to mimic the worse $\tilde{\phi}$ resolution, the generated $\tilde{\phi}$ is randomly smeared according to a Gaussian distribution with a width given by the Gaussian fit to the $B^0 \to K^{*0} e^+ e^-$ MC performed in each $m_{ee}^{B^0}$ bin to the $(\tilde{\phi}_{\text{true}} - \tilde{\phi})$ distribution. Each smeared sample is then fitted with the full PDF and the generated $A_T^{(2)}$ is compared with the measured $A_T^{(2)}$ in the pseudo experiments. As expected the measured $A_T^{(2)}$ is slightly undervalued. However, in the range $[10, 500]\,\text{MeV}/c^2$ the bias is as low as $\sim 4\%$ (see figure 4.23). The fitted value on data is therefore corrected accordingly while this correction is assigned as a systematic error.

4.5.6 K^{*0} and $\overline{K}^{*0}$ mis-identification

The amount of K^{*0} mis-identified as $\overline{K}^{*0}$, *i.e.* a double mis-identication of the kaon as a pion and the pion as a kaon has been evaluated to be equal to $0.17 \pm 0.01\%$ (see section 3.3.8).
The potentially affected parameters are F_{L}, A_T^{Re} and A_T^{Im}. The signal angular PDF is first fitted to the $B^0 \to K^{*0} e^+ e^-$ MC. Then, a second sample is constructed. The latter is made of $B^0 \to K^{*0} e^+ e^-$ MC combined with 0.17% of $B^0 \to K^{*0} e^+ e^-$ MC with a $K^{*0}/\overline{K}^{*0}$ swap. The difference of the fitted parameters in the two cases is assigned as a systematic error.

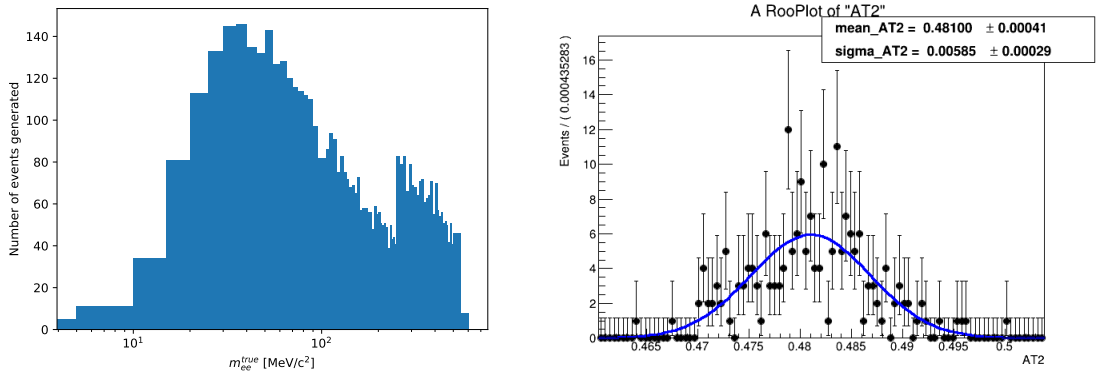


Figure 4.23: (left) Number of generated events as a function of the dielectron mass, as given by eq. (4.12). The bin width is 5 between 0 and 250 MeV/c^2, 10 between 250 and 500 MeV/c^2 and 50 above 500 MeV/c^2. (right) Measured $A_T^{(2)}$ for 1000 toys generated with $A_T^{(2)} = 0.5$. The bias on the measurement of $A_T^{(2)}$ in the range $m_{ee}^{B^0} \in [10, 500]\,\text{MeV}/c^2$ is $\sim 4\%$.

4.5.7 Data/MC differences in mass shape

Data/MC differences in the mass shape of $B^0 \to K^{*0}e^+e^-$ signal (and $B^0 \to K^{*0}\gamma(\to e^+e^-)$ in the gamma-q^2 bin) are accounted for by a shift $m_{J/\psi}$ to the mean and a scale factor $s_{J/\psi}$ to the width of the DCB. These parameters are obtained by a fit to the $B^0 \to K^{*0}J/\psi\,(\to e^+e^-)$ resonant decay mode (see section 3.6.4). To evaluate the uncertainty related to the data/MC differences, an alternative model is considered where the shift and scale factors are obtained from a fit to $B^0 \to K^{*0}\gamma(\to e^+e^-)$ data in the gamma-q^2 bin.

4.6 Results

The fit is performed on $B^0 \to K^{*0}e^+e^-$ data following the strategy described in section 4.1, and is shown in figure 4.24. The results are

$$\begin{aligned} A_T^{(2)} &= 0.106 \pm 0.103\ ^{+0.016}_{-0.017} \\ A_T^{Im} &= 0.015 \pm 0.102 \pm 0.012 \\ A_T^{Re} &= -0.064 \pm 0.077 \pm 0.015 \\ F_\text{L} &= 0.044 \pm 0.026 \pm 0.014 \qquad , \end{aligned} \tag{4.13}$$

where the first error is the statistical error while the second error is the systematic error.

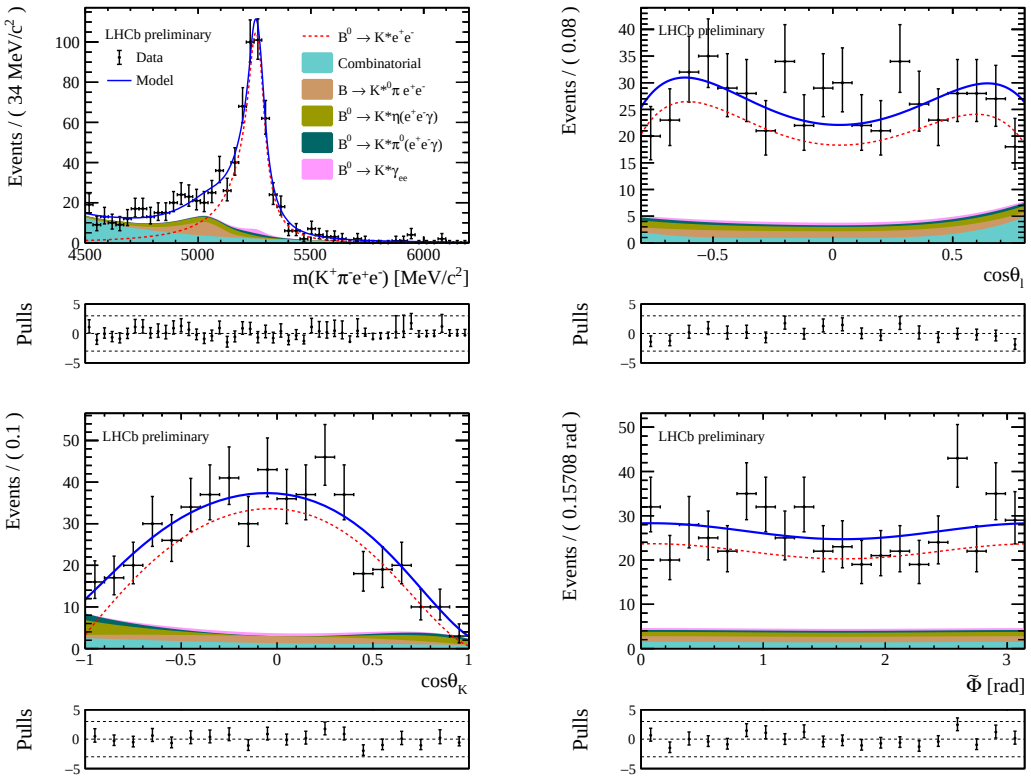


Figure 4.24: Fit to $B^0 \to K^{*0}e^+e^-$ data for all years (R1+R2) and trigger categories (L0L+L0I). The fit is unbinned and simultaneous in the four categories (two years and two trigger), and on the $m(K^+\pi^-e^+e^-)$ invariant mass and the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$.

4.6.1 Comparison to previous results

The four parameters $F_{\rm L}$, $A_T^{(2)}$, A_T^{Im} and A_T^{Re} have also been measured in a previous angular analysis [100] of the $B^0 \to K^{*0}e^+e^-$ decay in the q^2 region $[0.0004, 1]\,{\rm GeV}^2/c^4$ using the $3\,{\rm fb}^{-1}$ data collected by LHCb during R1. On top of the statistical gain due to the additional data collected during R2, the analysis presented in this thesis features several improvements with respect to the former R1 analysis.
First, the sensitivity to the photon polarization, in particular the disentanglement of the $\mathcal{C}_7^{(\prime)}$ from the $\mathcal{C}_{9,10}$ contributions is improved by lowering the effective true-q^2 range of the analysis from $[0.002 \pm 0.0008, 1.12 \pm 0.06]\,{\rm GeV}^2/c^4$ down to $[0.0008 \pm 0.0004, 0.257 \pm 0.003]\,{\rm GeV}^2/c^4$ (see section 4.6.3). This is largely due to the fact that the reconstructed dielectron mass range has been lowered from $[20, 1000]\,{\rm MeV}/c^2$ down to $[10, 500]\,{\rm MeV}/c^2$. In addition, the usage of the vertex and mass constrained dielectron mass $m_{ee}^{B^0}$ allows the reconstructed q^2 range to be

much closer to the true-q^2 range, while the development of a new veto cut against photon conversions (see section 3.3.2) allows to increase the signal efficiency of the selection by $\sim 20\%$ at the lower end of the q^2 range.
Moreover, the selection has been improved, lowering the overall background over signal ratio from $\sim 100\%$ to $\sim 20\%$. This improvement is illustrated in figure 4.25 where the mass fits of the two analyses are compared. This is in particular due to the introduction of a new veto cut to remove partially reconstructed and combinatorial backgrounds (see section 3.4.1), combined to the fact that the combinatorial and partially reconstructed backgrounds are phase space suppressed at lower q^2 while the signal benefits from the photon pole.
To finish, the mass and angular modeling of backgrounds has been improved. New backgrounds coming from η and π^0 Dalitz decays which were previously neglected are included to the model (see sections 3.6.1 and 4.3.4). In addition, a new data-based approach has been developed to model the combinatorial and semileptonic backgrounds (see section 4.3.2), which are the most challenging backgrounds to model in this analysis. Combined with the lower amount of these backgrounds, this is reflected by a decrease of about one order of magnitude of the systematic error associated to the angular modeling of these backgrounds.

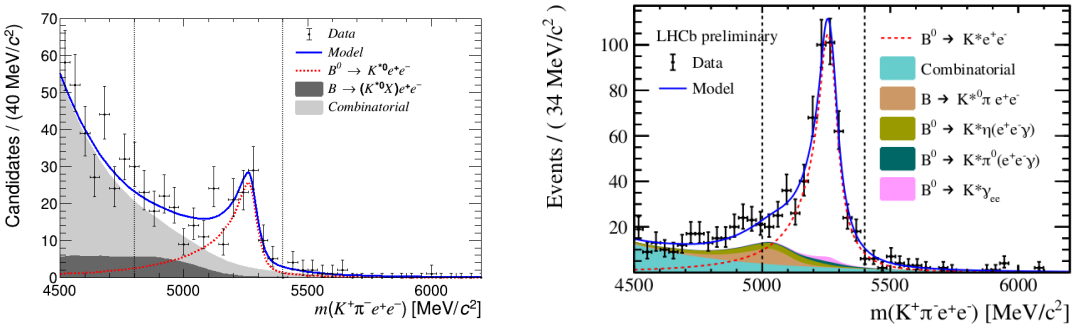


Figure 4.25: Comparison of the mass fits to the full datasets in the R1 analysis (left) and the present analysis (right). In both plots, the signal is shown in red dotted line while the sum of all components is shown in blue solid line. On the left plot, the partially reconstructed background is shown in dark grey and the combinatorial/semileptonic background in light grey. On the right plot, the $B^0 \to K^{*0}\gamma(\to e^+e^-)$ background is shown in pink dotted line, the partially reconstructed background in brown, the combinatorial/semileptonic background in cyan, the $B^0 \to K^{*0}\eta(\to \gamma e^+e^-)$ background in light green and the $B^0 \to K^{*0}\pi^0(\to \gamma e^+e^-)$ background in dark green.

The overall improvement to the measurement of the photon polarization is illustrated in figure 4.29b where the improvement of the constraint on the right-handed current of the photon is shown.

4.6.2 Effective q^2 bin

The signal acceptance as a function of the true q^2, obtained from the corrected MC, is uniform in a large domain except close to the limits of the reconstructed q^2, 0.0001 and 0.25 GeV^2/c^4 (10 to 500 MeV/c^2 in $m_{ee}^{B^0}$). For this reason, the R1 analysis [100] results were accompanied by the definition of an effective q^2 bin, to be used by theorist as an approximation of the acceptance as a function of true q^2. The q^2 resolution has been significantly improved in the present analysis using the B^0 mass constrained $m_{ee}^{B^0}$ (as explained in section 1.4.2), but the resolution effect has still to be taken into account.
In the same way as for the R1 analysis, the values of the effective q^2 boundaries are obtained by requiring that in the low- and high-q^2 regions the same number of events are obtained in a uniform acceptance model and in the MC. In other words, in the effective q^2 bin, the fraction of reconstructed events that have a true-q^2 below the effective lower limit is compensated by assuming a larger efficiency for the events just above the lower effective limit (and similarly for the higher effective limit).
The true-q^2 effective region is thus determined to be between 0.0008 ± 0.0004 and 0.257 ± 0.003 GeV^2/c^4. An uncertainty on the q^2 effective limits is assigned as half of the q^2 limit modification. The m_{ee} distribution of reconstructed events is compared to the one of the generator level events falling in the effective q^2 bin in figure 4.26.

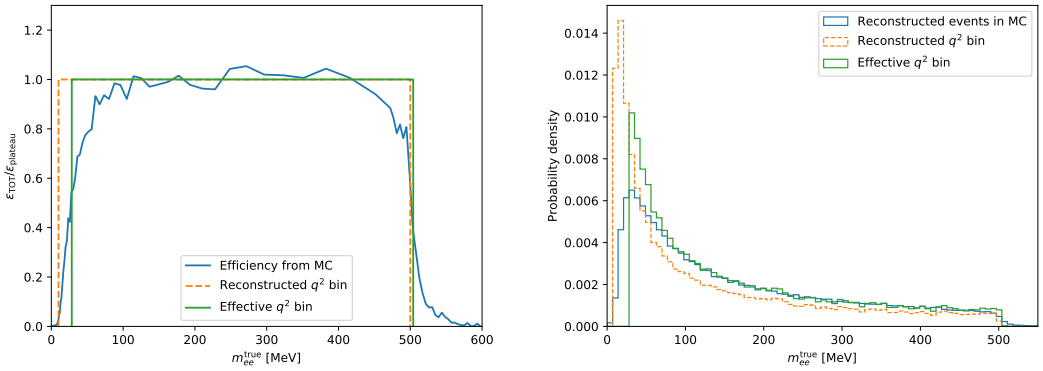


Figure 4.26: Left: Reconstruction efficiency from the MC compared to the effective-q^2-bin approximation as a function of $m_{ee}^{B^0}$. Right: The true-m_{ee} distribution of reconstructed MC events is compared to the one of generator-level events within the effective q^2 bin. The orange dashed curve shows the generator level distribution within the boundaries of the reconstructed-q^2 bin (the incorrect distribution one would naively assume if not using the effective q^2 bin).

It is checked that the average values of the true-q^2 and of the angular observables evaluated with a uniform acceptance in the effective q^2 region are in agreement with those obtained using the true-q^2 acceptance from the MC.

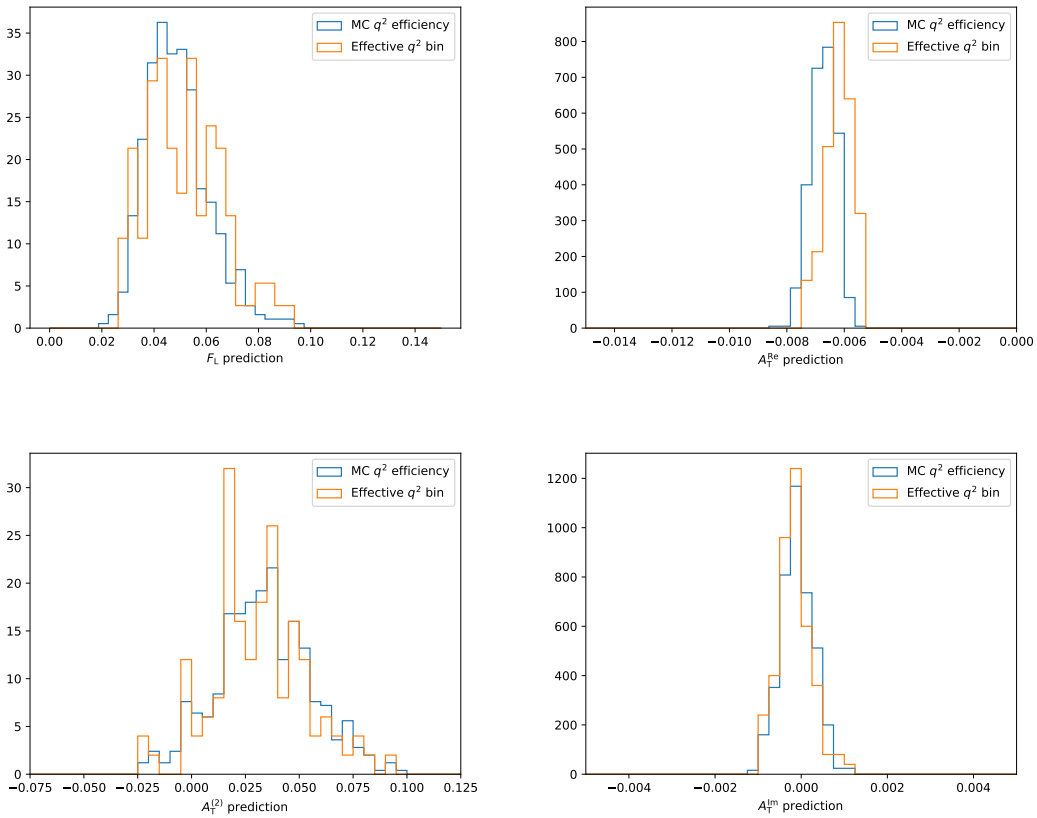


Figure 4.27: Theoretical predictions for the angular observables obtained by `flavio` and floated according to the theoretical inputs. Predictions using the full q^2 efficiency are compared to the ones obtained using the effective q^2 bin (taking into account its uncertainty).

Theoretical predictions for the angular observables are potentially affected by the effective q^2 bin approximation. This has been checked using `flavio` 2.0.0 [12]. The theoretical inputs to `flavio` predictions (mainly the form factors) are floated according to their errors with the same routine that is used by `flavio` to estimate theoretical uncertainties. For each set of theoretical inputs the observables are calculated using the full q^2 efficiency and the approximated effective q^2 bin. The uncertainty on the effective q^2 bin boundaries is taken into account by floating them. Results are shown in figure 4.27. The bias induced by the effective-q^2-bin approximation is completely negligible with respect to the theoretical uncertainties

for all observables but A_T^{Re}. However, the theoretical uncertainty on A_T^{Re} is 100 times smaller than the experimental sensitivity so this small bias is completely negligible in that respect.

The accuracy of the theoretical predictions in the effective q^2 bin is studied in the presence of New Physics in the Wilson coefficients $\mathcal{C}_7'$ and $\mathcal{C}_9^{\rm NP}$. Results are shown in figure 4.28. Once again, the effective q^2 definition proves to be solid as biases in the predictions are negligible.

4.6.3 Constraints on the photon polarization

Using the procedure described above, theoretical predictions for the four observables $F_{\rm L}$, $A_T^{(2)}$, A_T^{Im} and A_T^{Re} are computed and compared to the measured values in table 4.5. The uncertainty in the theoretical predictions takes into account both the error on the effective q^2 bin as well as the theoretical errors. Each of the four measured parameter is compatible with the SM predictions within one σ.

	SM predictions from [12]	LHCb measurement
$A_T^{(2)}$	0.034 ± 0.021	$0.106 \pm 0.103\ ^{+0.016}_{-0.017}$
A_T^{Im}	$(-0.9 \pm 3.5) \times 10^{-4}$	$0.015 \pm 0.102 \pm 0.012$
A_T^{Re}	$(-6.3 \pm 0.5) \times 10^{-3}$	$-0.064 \pm 0.077 \pm 0.015$
$F_{\rm L}$	0.051 ± 0.014	$0.044 \pm 0.026 \pm 0.014$

Table 4.5: Comparison of SM predictions in the effective q^2 bin of the analysis to the LHCb measurement presented in this thesis.

As given in eqs. (1.37) and (1.38), in the limit where $q^2 \to 0$, $A_T^{(2)}$ and A_T^{Im} are directly proportional to the $\mathcal{C}_7^{(\prime)}$ Wilson coefficients. As shown in section 1.4.2, in the q^2 bin of this analysis, the corrections to this proportionality relation coming from $\mathcal{C}_{9,10}$ contributions are smaller than 2% (assuming SM values for $\mathcal{C}_9$ and $\mathcal{C}_{10}$). Therefore, the measurement presented in this thesis allows to put constraints to the helicity structure of the photon polarization in $b \to s\gamma$ processes. In particular, the scenario where NP only appears in the right-handed current, *i.e.* $\mathcal{C}_7'$, is considered. The possible NP contributions to $\mathcal{C}_7'$ is denoted $\mathcal{C}_7'^{\rm NP}$. Since $\mathcal{C}_7'$ is a complex number, the constraints are expressed in the complex plane of $\mathcal{C}_7'^{\rm NP}$, where by definition the SM value in this plane is $(0, 0)$. As shown in figure 4.29a, the measurement of $A_T^{(2)}$ mainly constrains the real part while the measurement of A_T^{Im} mainly constrains the imaginary part of $\mathcal{C}_7'$. The constraints from the present measurement is compared to the constraints given by the previous measurements discussed in section 1.3: inclusive branching fractions of $B \to X_s\gamma$ decays and $S_{d\gamma}$ CP-asymmetry measurements at the Belle and BaBar experiments, the mea-

surement of the $\mathcal{A}_{s\gamma}^{\Delta}$ and $S_{s\gamma}$ CP-asymmetries in $B_s^0 \to \phi\gamma$ decays by the LHCb experiment and the previous $B^0 \to K^{*0}e^+e^-$ angular analysis using the full R1 LHCb data. As shown in figure 4.29b, the measurement presented in this thesis is by far the most constraining one. In addition, contrary to CP observables, angular observables in the $B^0 \to K^{*0}e^+e^-$ decay channel are not affected by potential NP in meson-antimeson mixing [101]. Moreover, $A_T^{(2)}$ and A_T^{Im} depend on the difference of the weak phases of $\mathcal{C}_7$ and $\mathcal{C}_7^{(\prime)}$ while the $\mathcal{A}_{s\gamma}^{\Delta}$ and $S_{s\gamma}$ CP-asymmetries depend on their sum. Therefore, while the present constraining power of $\mathcal{A}_{s\gamma}^{\Delta}$ and $S_{s\gamma}$ in the $\mathcal{C}_7^{\prime}$ complex plane does not yet reach the level of that of the measurement presented in this thesis, the two measurements are complementary in scenarios with complex contributions to both $\mathcal{C}_7$ and $\mathcal{C}_7^{\prime}$ Wilson coefficients. Their combination, alongside with B-factories measurements, allows to perform global fits in the $\mathcal{C}_7^{(\prime)}$ complex plane such as [101]. A similar global fit including the measurement presented in this thesis is shown in figure 4.29c.

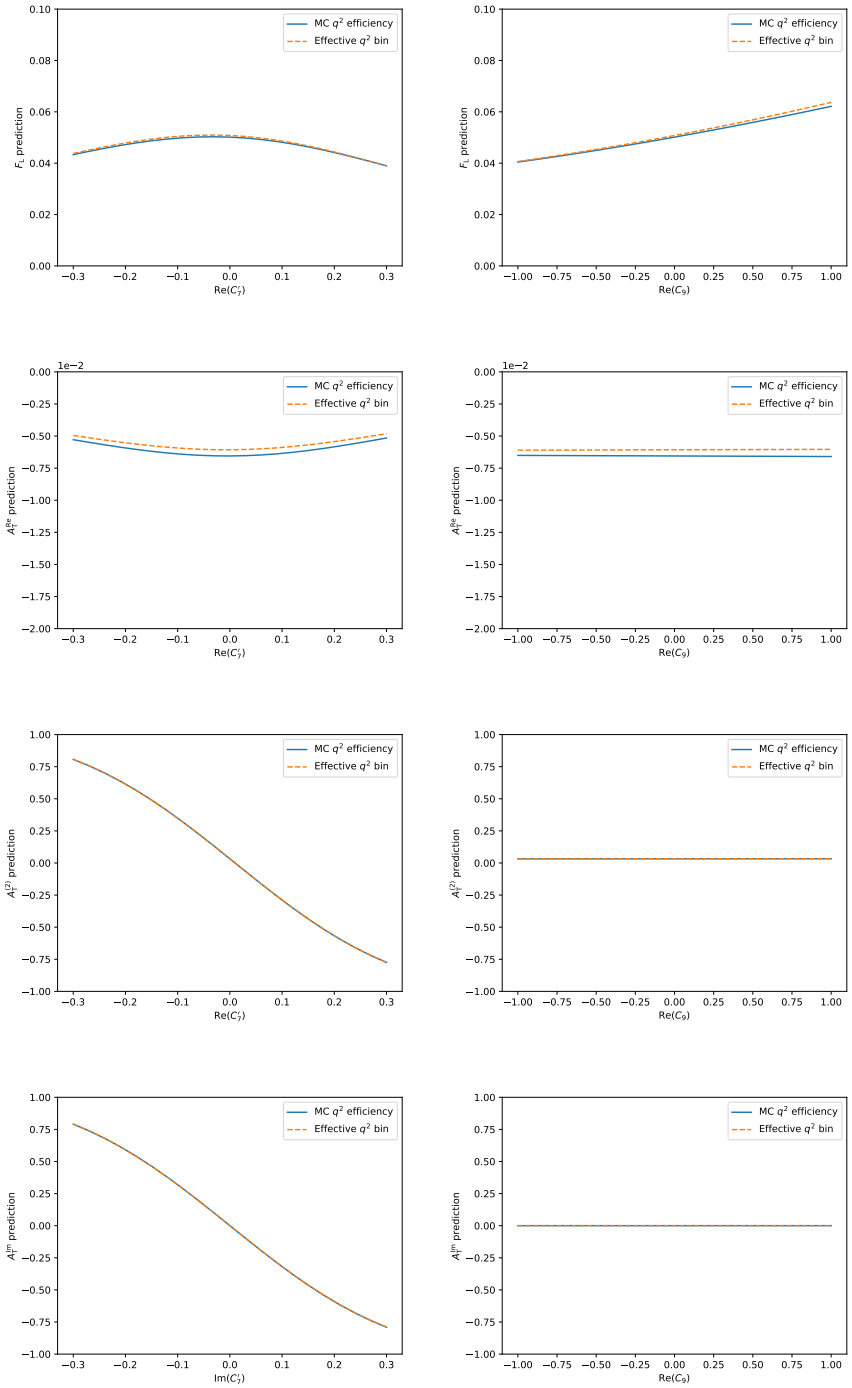


Figure 4.28: Theoretical predictions for the angular observables obtained by `flavio` as a function of the NP Wilson coefficients $\mathcal{C}_7'$ and $\mathcal{C}_9^{\text{NP}}$. Predictions using the full q^2 efficiency from MC (blue) are compared to the ones obtained using the effective q^2 bin (orange).

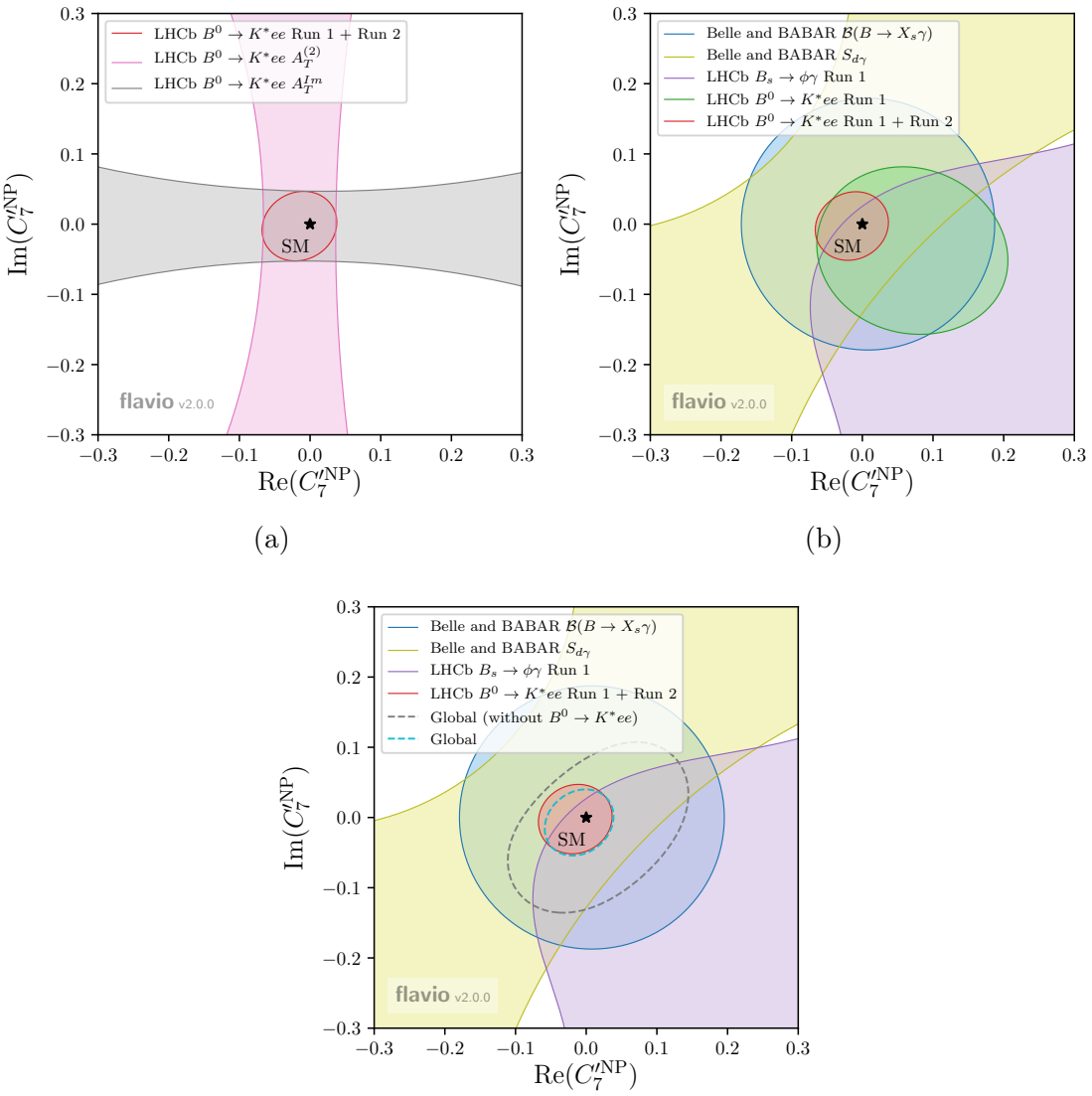
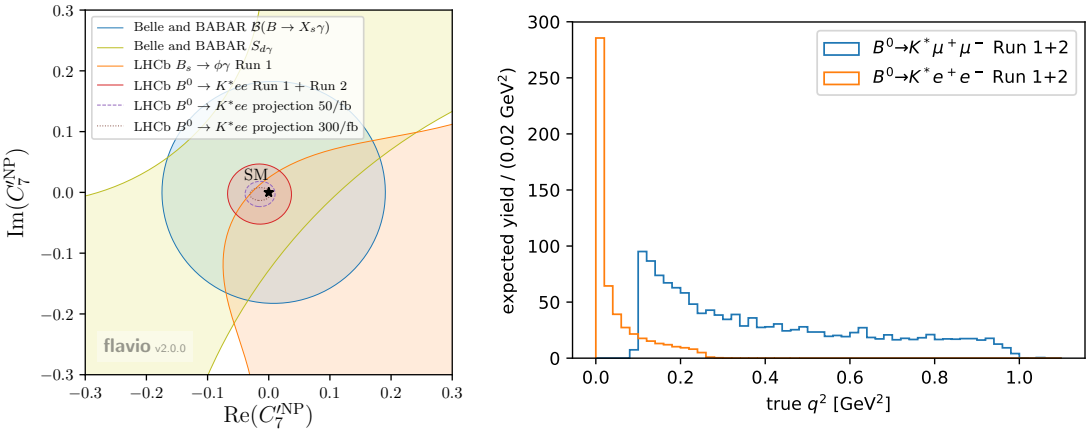


Figure 4.29: Top left: Constraints (at 2σ level) to the real and imaginary parts of $\mathcal{C}_7^{\prime\mathrm{NP}}$ from the measurement of $A_T^{(2)}$ (pink) and A_T^{Im} (grey) in the present angular analysis of the $B^0 \to K^{*0}e^+e^-$ decay. The combination of the two measurements is shown in red. Top right: Comparison of the constraints from past measurements to the measurement presented in this thesis (red). The inclusive branching fraction and the $S_{d\gamma}$ measurements by the Belle and BaBar experiments are shown in blue and yellow respectively, the $B_s^0 \to \phi\gamma$ measurements at LHCb in purple and the previous angular analysis of $B^0 \to K^{*0}e^+e^-$ using LHCb R1 data only in green. Bottom: Current constraints on the $\mathcal{C}_7^{\prime}$ Wilson coefficient. The global fit without the measurement presented in this thesis is shown in grey dashed lines while the global fit including this measurement is shown in cyan dashed lines. In all plots, the SM prediction is represented by a black star.

4.7 Conclusion and perspectives for the future

An angular analysis of the Flavor Changing Neutral Current $B^0 \to K^{*0} e^+ e^-$ decay has been performed using the full Run 1 and Run 2 LHCb data collected between 2011 and 2018 at 7 TeV (2011), 8 TeV (2012) and 13 TeV (2015-2018) centre-of-mass energies. A total amount of about 450 $B^0 \to K^{*0} e^+ e^-$ candidates are selected in the q^2 range $[0.0001, 0.25]\,\mathrm{GeV}^2/c^4$, with an estimated background over signal ratio of $\sim 20\%$. A simultaneous fit to the reconstructed $m(K^+\pi^- e^+ e^-)$ mass as well as the three angles $\cos\theta_\ell$, $\cos\theta_K$ and $\tilde{\phi}$ is performed to extract the four angular observables F_L, A_T^{Re}, $A_T^{(2)}$ and A_T^{Im}. While F_L is related to the longitudinal polarization of the K^{*0} meson and A_T^{Re} is related to the lepton forward-backward asymmetry, $A_T^{(2)}$ and A_T^{Im} in this low q^2 region are sensitive to the photon polarization in $b \to s\gamma$ transitions. The results presented in this thesis therefore allow to constrain possible Beyond the Standard Model right-handed currents embedded by the Wilson coefficient $\mathcal{C}_7^{'}$. The measurement is compatible with the Standard Model (SM) at 0.5σ in the $\mathcal{C}_7^{'}$ complex plane, while each of the four angular observables are compatible with the SM predictions within one σ. It is currently the world's best constraint on right-handed contributions to the photon polarization in $b \to s\gamma$ transitions.
The current measurement is dominated by statistical uncertainties on $A_T^{(2)}$ and A_T^{Im} which are at least one order of magnitude larger than the current theoretical errors. Therefore, additional data which will be collected by the LHCb experiment in the future will allow to reduce the statistical uncertainties and shrink the constraints. The LHCb detector is currently going trough a major upgrade to increase the luminosity at which the experiment can be run. A contribution to this upgrade regarding the electronics of the calorimeters is presented in this thesis. The goal at the end of Run 4 in 2029 is to have collected $50\,\mathrm{fb}^{-1}$ proton-proton collisions at LHCb. Then, the LHCb detector will undergo a second major upgrade aiming at having collected a total integrated luminosity of $300\,\mathrm{fb}^{-1}$ data at the end of 2040. Taking only into consideration the statistical gain (*i.e.* assuming the same method, q^2 bin, efficiency and systematic errors than the ones presented in the current analysis), the projected constraints to the $\mathcal{C}_7^{'}$ Wilson coefficient are shown in figure 4.30a.

Beyond the gain of statistics, the $B^0 \to K^{*0} e^+ e^-$ angular analysis could be further improved in the future. First, the large data samples available in the future could allow to reduce the upper limit of the q^2 bin to further improve the disentanglement of the $\mathcal{C}_7^{(')}$ from the $\mathcal{C}_{9,10}$ Wilson coefficients. The lower limit could also be reduced while including the ϕ resolution effects to the model. The contamination coming from converted photons would however be challenging. In addition, instead of integrating out q^2 as it is done in the present analysis, one



(a) (b)

Figure 4.30: Left: Projected constraints to the $\mathcal{C}_7'$ Wilson coefficient with the 50 fb^{-1} (purple dashed) and 300 fb^{-1} (brown dotted) LHCb datasets, compared to the current (red) $B^0 \to K^{*0}e^+e^-$ measurements as well as previous measurements from various decays and experiments (blue, yellow and orange). Right: true-q^2 distribution of the analysis presented in this thesis (orange) and the $B^0 \to K^{*0}\mu^+\mu^-$ angular analysis [102] (blue).

could perform a full amplitude analysis (in a similar fashion to what is proposed in [103]). This would in particular allow for a significant gain in the statistical error of F_{L}, although the gain for $A_T^{(2)}$ and A_T^{Im}, *i.e.* to the photon polarization, would be quite limited given the very narrow q^2 bin of the analysis. Furthermore, such an amplitude analysis involves major challenges including the q^2-dependent modeling of backgrounds as well as q^2 resolution effects.

In addition, the $B_s^0 \to \phi e^+e^-$ decay also allows to measure the photon polarization in $b \to s\gamma$ processes. Indeed, the differential decay width of the $B_s^0 \to \phi e^+e^-$ decay can be written in a similar fashion as the $B^0 \to K^{*0}e^+e^-$ one (akin to the one given in [104]). As in the B^0 case, the B_s^0 differential decay width has a similar term proportional to $A_T^{(2)}$. However, since the flavor of the B_s^0 meson is not accessible experimentally, the term proportional to A_T^{Im} vanishes. Since it gives access to additional CP-symmetries and CP-asymmetries, namely S_7 and $A_{5,6}$, while contrary to the B^0 channel it is not sensitive to $S_{5,6}$ and A_7, the B_s channel is complementary to the B^0 one. Although it suffers from a lower production rate $f_s/f_d \sim 1/4$ at LHCb, the narrow ϕ resonance allows to select a clean $B_s^0 \to \phi e^+e^-$ sample, in particular free from partially reconstructed background.

To finish, even if the photon polarization in $b \to s$ transitions is the main motivation for the angular analysis of the $B^0 \to K^{*0}e^+e^-$ decay at very low q^2,

the current measurement could be extended to more observables with the current dataset. Indeed, the current measurement is done with a folding on the ϕ angle. While it does not decrease the sensitivity to $A_T^{(2)}$ and A_T^{Im}, the folding removes the observables S_4, S_5, A_7 and A_8 from the full decay width in eq. (1.28). As shown in figure 4.30b, the angular analysis of the $B^0 \to K^{*0}e^+e^-$ decay channel probes a different q^2 region than the corresponding angular analysis in the muon channel $B^0 \to K^{*0}\mu^+\mu^-$ [102]. Doing a similar measurement than the one presented in this thesis without folding the ϕ angle and extending the upper q^2 region up to $1.1\,\mathrm{GeV}^2/c^4$ (the lower q^2 boundary of the muon analysis) would allow to probe $\mathcal{C}_{7,9,10}$ interferences in a q^2 region inaccessible to the muon channel. The unfolding, however, makes the analysis more complex and comes with additional challenges, in particular regarding the stability of the fit. In addition, in the scenario where one would increase the upper q^2 limit, one would have to deal with the pollution coming from the ρ^0 resonance. Moreover, the current dataset of selected $B^0 \to K^{*0}e^+e^-$ in the very-low-q^2 bin could be used to measure the $B^0 \to K^{*0}e^+e^-$ branching fraction in this q^2 region with an expected precision of $\sim 5\%$. While the current SM prediction suffers from a rather high uncertainty (at the level of $\sim 20\%$) in this low q^2 region, one can shrink the theoretical error to the sub-percent level by building the observable $(1 - F_\mathrm{L}) \cdot \mathcal{B}(B^0 \to K^{*0}e^+e^-)/\mathcal{B}(B^0 \to K^{*0}\gamma)$. This measurement could be especially useful for Lepton Flavor Universality tests (such as [59]), where the q^2 dependence of the efficiency plays an important role.

Appendix A

Angular differential decay rate of the $B^0 \to K^{*0} e^+ e^-$ decay

A.1 Definition of the angular basis

The full differential decay width of the $B^0 \to K^{*0} e^+ e^-$ decay can be described as a function of three angles $(\cos\theta_\ell, \cos\theta_K, \phi)$ and q^2, the invariant mass squared of the dielectron system. The three angles are sketched in figure A.1.

The angular basis is defined such that the $\overline{B}^0$ angular definition is a *CP* transformation of that of the B^0. The angle θ_l is defined as the angle between the direction of the e^+ (e^-) and the direction opposite to the B^0 $\left(\overline{B}^0\right)$ direction 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 the B^0 $\left(\overline{B}^0\right)$ direction in the K^{*0} $\left(\overline{K}^{*0}\right)$ rest frame. The angle ϕ is defined as the angle between the plane containing the two leptons and the plane containing the two hadrons of the final state in the B^0 $\left(\overline{B}^0\right)$ rest frame. Explicitly, $\cos\theta_\ell$ is given by

$$\cos\theta_\ell = \left(\hat{p}_{e^+}^{(e^+e^-)}\right) \cdot \left(\hat{p}_{e^+e^-}^{(B^0)}\right) = \left(\hat{p}_{e^+}^{(e^+e^-)}\right) \cdot \left(-\hat{p}_{B^0}^{(e^+e^-)}\right) \tag{A.1}$$

for the B^0 and

$$\cos\theta_\ell = \left(\hat{p}_{e^-}^{(e^+e^-)}\right) \cdot \left(\hat{p}_{e^+e^-}^{(\overline{B}^0)}\right) = \left(\hat{p}_{e^-}^{(e^+e^-)}\right) \cdot \left(-\hat{p}_{\overline{B}^0}^{(e^+e^-)}\right) \tag{A.2}$$

for the $\overline{B}^0$. The explicit definition of $\cos\theta_K$ is given by

$$\cos\theta_K = \left(\hat{p}_{K^+}^{(K^{*0})}\right) \cdot \left(\hat{p}_{K^{*0}}^{(B^0)}\right) = \left(\hat{p}_{K^+}^{(K^{*0})}\right) \cdot \left(-\hat{p}_{B^0}^{(K^{*0})}\right) \tag{A.3}$$

for the B^0 and

$$\cos\theta_K = \left(\hat{p}_{K^-}^{(\overline{K}^{*0})}\right) \cdot \left(\hat{p}_{\overline{K}^{*0}}^{(\overline{B}^0)}\right) = \left(\hat{p}_{K^-}^{(\overline{K}^{*0})}\right) \cdot \left(-\hat{p}_{\overline{B}^0}^{(\overline{K}^{*0})}\right) \tag{A.4}$$

for the $\overline{B}^0$. The explicit definition of ϕ is given by

$$\begin{aligned} \cos\phi &= \left(\hat{p}_{e^+}^{(B^0)} \times \hat{p}_{e^-}^{(B^0)}\right) \cdot \left(\hat{p}_{K^+}^{(B^0)} \times \hat{p}_{\pi^-}^{(B^0)}\right) \quad , \\ \sin\phi &= \left[\left(\hat{p}_{e^+}^{(B^0)} \times \hat{p}_{e^-}^{(B^0)}\right) \times \left(\hat{p}_{K^+}^{(B^0)} \times \hat{p}_{\pi^-}^{(B^0)}\right)\right] \cdot \hat{p}_{K^{*0}}^{(B^0)} \end{aligned} \tag{A.5}$$

for the B^0 and by

$$\begin{aligned} \cos\phi &= \left(\hat{p}_{e^-}^{(\overline{B}^0)} \times \hat{p}_{e^+}^{(\overline{B}^0)}\right) \cdot \left(\hat{p}_{K^-}^{(\overline{B}^0)} \times \hat{p}_{\pi^+}^{(\overline{B}^0)}\right) \quad , \\ \sin\phi &= -\left[\left(\hat{p}_{e^-}^{(\overline{B}^0)} \times \hat{p}_{e^+}^{(\overline{B}^0)}\right) \times \left(\hat{p}_{K^-}^{(\overline{B}^0)} \times \hat{p}_{\pi^+}^{(\overline{B}^0)}\right)\right] \cdot \hat{p}_{\overline{K}^{*0}}^{(\overline{B}^0)} \end{aligned} \tag{A.6}$$

for the $\overline{B}^0$, where $\hat{p}_X^{(Y)}$ 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 ($\overline{B}^0$) rest frame.

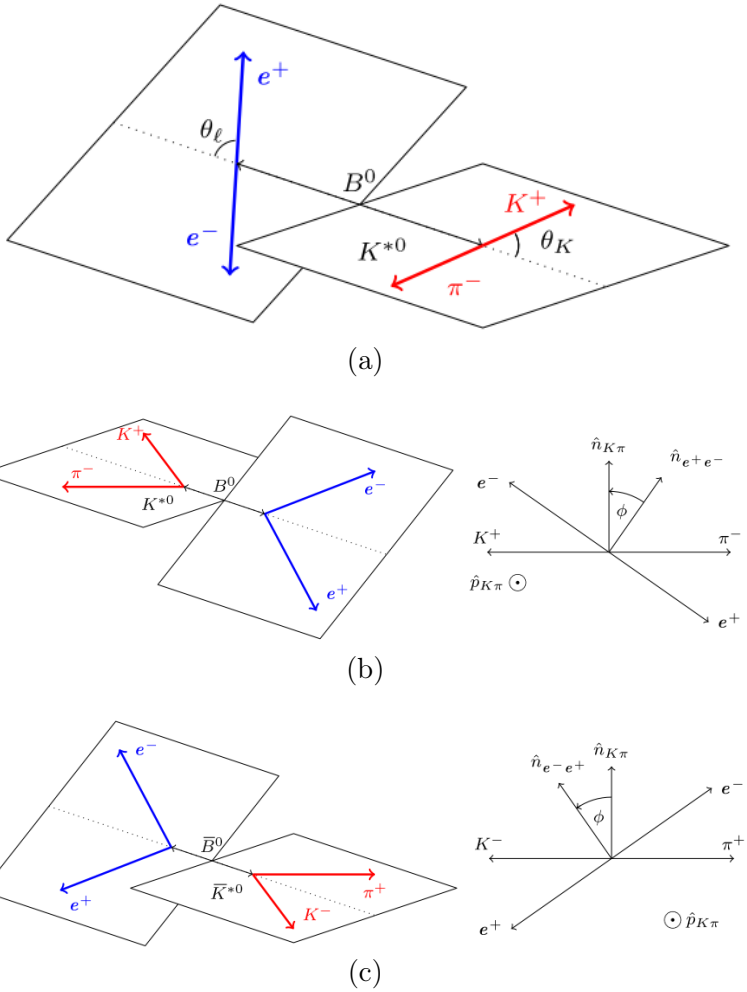


Figure A.1: Sketch of the θ_l and θ_K angles for the B^0 (A.1a), the ϕ angle for the B^0 (A.1b) and the ϕ angle for the $\overline{B}^0$ (A.1c).

A.2 Definition of the angular coefficients I_i

The expression of the angular coefficients of Eq. (1.25) in the limit of massless lepton and vanishing S-wave are given by [40]

$$
\begin{aligned}
I_1^S &= \frac{3}{4}\left[|A_\perp^L|^2 + |A_\parallel^L|^2 + (L \to R)\right] \quad , \\
I_1^C &= |A_0^L|^2 + |A_0^R|^2 \quad , \\
I_2^S &= \frac{1}{4}\left[|A_\perp^L|^2 + |A_\parallel^L|^2 + (L \to R)\right] \quad , \\
I_2^C &= -\left[|A_0^L|^2 + (L \to R)\right] \quad , \\
I_3 &= \frac{1}{2}\left[|A_\perp^L|^2 - |A_\parallel^L|^2 + (L \to R)\right] \quad , \\
I_4 &= \frac{1}{\sqrt{2}}\left[\mathcal{R}e\left(A_0^L A_\parallel^{L*}\right) + (L \to R)\right] \quad , \\
I_5 &= \sqrt{2}\left[\mathcal{R}e\left(A_0^L A_\perp^{L*}\right) - (L \to R)\right] \quad , \\
I_6^S &= 2\left[\mathcal{R}e\left(A_\parallel^L A_\perp^{L*}\right) - (L \to R)\right] \quad , \\
I_7 &= \sqrt{2}\left[\mathcal{I}m\left(A_0^L A_\parallel^{L*}\right) - (L \to R)\right] \quad , \\
I_8 &= \frac{1}{\sqrt{2}}\left[\mathcal{I}m\left(A_0^L A_\perp^{L*}\right) + (L \to R)\right] \quad , \\
I_9 &= \left[\mathcal{I}m\left(A_\parallel^{L*} A_\perp^L\right) + (L \to R)\right] \quad .
\end{aligned}
\tag{A.7}
$$

A.3 Definition of the transversity amplitudes

The transversity amplitudes of Eq. (A.7) can be expressed as [40]

$$
\begin{aligned}
A_{\perp}^{L,R}(q^2) &= N(q^2)\sqrt{2\lambda(q^2)}\left\{\frac{2m_b}{q2}\left(\mathcal{C}_7+\mathcal{C}_7^{'}\right)T_1(q^2)\right. \\
&\left.+\left[\left(\mathcal{C}_9+\mathcal{C}_9^{'}\right)\mp\left(\mathcal{C}_{10}+\mathcal{C}_{10}^{'}\right)\right]\frac{V(q^2)}{m_B+m_{K^{*0}}}\right\} \quad , \\
A_{\parallel}^{L,R}(q^2) &= -N(q^2)\sqrt{2}\left(m_B^2-m_{K^{*0}}^2\right)\left\{\frac{2m_b}{q^2}\left(\mathcal{C}_7-\mathcal{C}_7^{'}\right)T_2(q^2)\right. \\
&\left.+\left[\left(\mathcal{C}_9-\mathcal{C}_9^{'}\right)\mp\left(\mathcal{C}_{10}-\mathcal{C}_{10}^{'}\right)\right]\frac{A_1(q^2)}{m_B-m_{K^{*0}}}\right\} \quad , \\
A_0^{L,R}(q^2) &= -\frac{N(q^2)}{2m_{K^{*0}}\sqrt{q^2}}\left\{\left[\left(\mathcal{C}_9-\mathcal{C}_9^{'}\right)\mp\left(\mathcal{C}_{10}-\mathcal{C}_{10}^{'}\right)\right]\right. \\
&\times\left[\left(m_B^2-m_{K^{*0}}^2-q^2\right)\left(m_B+m_{K^{*0}}\right)A_1(q^2)-\lambda(q^2)\frac{A_2(q^2)}{m_B^2-m_{K^{*0}}^2}\right] \\
&\left.+2m_b\left(\mathcal{C}_7-\mathcal{C}_7^{'}\right)\left[\left(m_B^2+3m_{K^{*0}}^2-q^2\right)T_2(q^2)-\frac{\lambda(q^2)}{m_B^2-m_{K^{*0}}^2}T_3(q^2)\right]\right\} \quad ,
\end{aligned}
\tag{A.8}
$$

where

$$
\begin{aligned}
N(q^2) &= V_{tb}V_{ts}^*\left[\frac{G_F^2\alpha^2}{2^{10}\pi^5 m_B^3}\frac{\beta_e(q^2)}{3}q^2\sqrt{\lambda(q^2)}\right]^{1/2} \quad , \\
\beta_l(q^2) &= \sqrt{1-\frac{4m_e^2}{q^2}} \quad , \\
\lambda(q^2) &= \left[q^2-(m_B+m_{K^{*0}})^2\right]\left[q^2-(m_B-m_{K^{*0}})^2\right] \quad ,
\end{aligned}
\tag{A.9}
$$

and $V(q^2)$, $A_{1,2}(q^2)$ and $T_{1,2,3}(q^2)$ are hadronic form factors.

Appendix B

Combinatorial background

B.1 Variables used for the training of the BDT

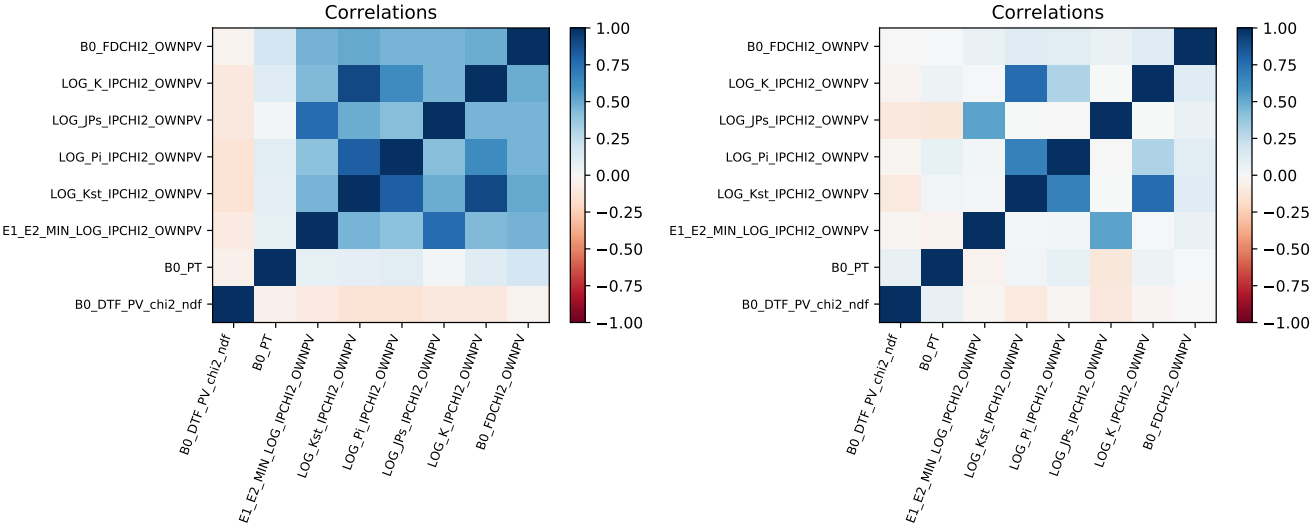


Figure B.1: Correlations between the variables used for the training of the BDT, in the R1 signal training sample (left) and in the R1 background training sample (right)

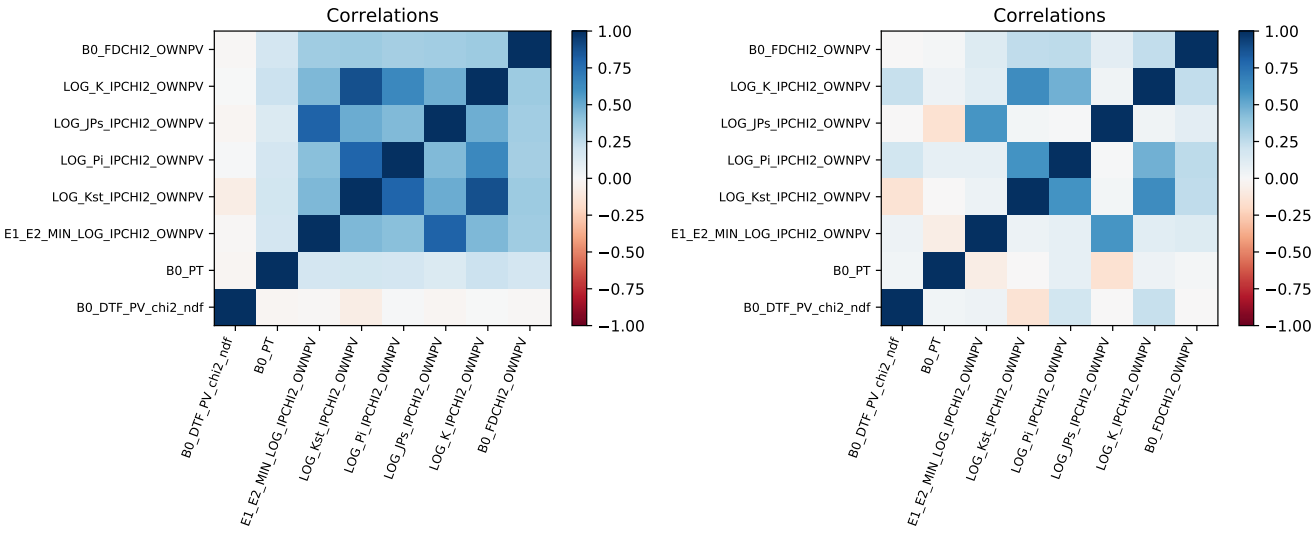


Figure B.2: Correlations between the variables used for the training of the BDT, in the R2 signal training sample (left) and in the R2 background training sample (right)

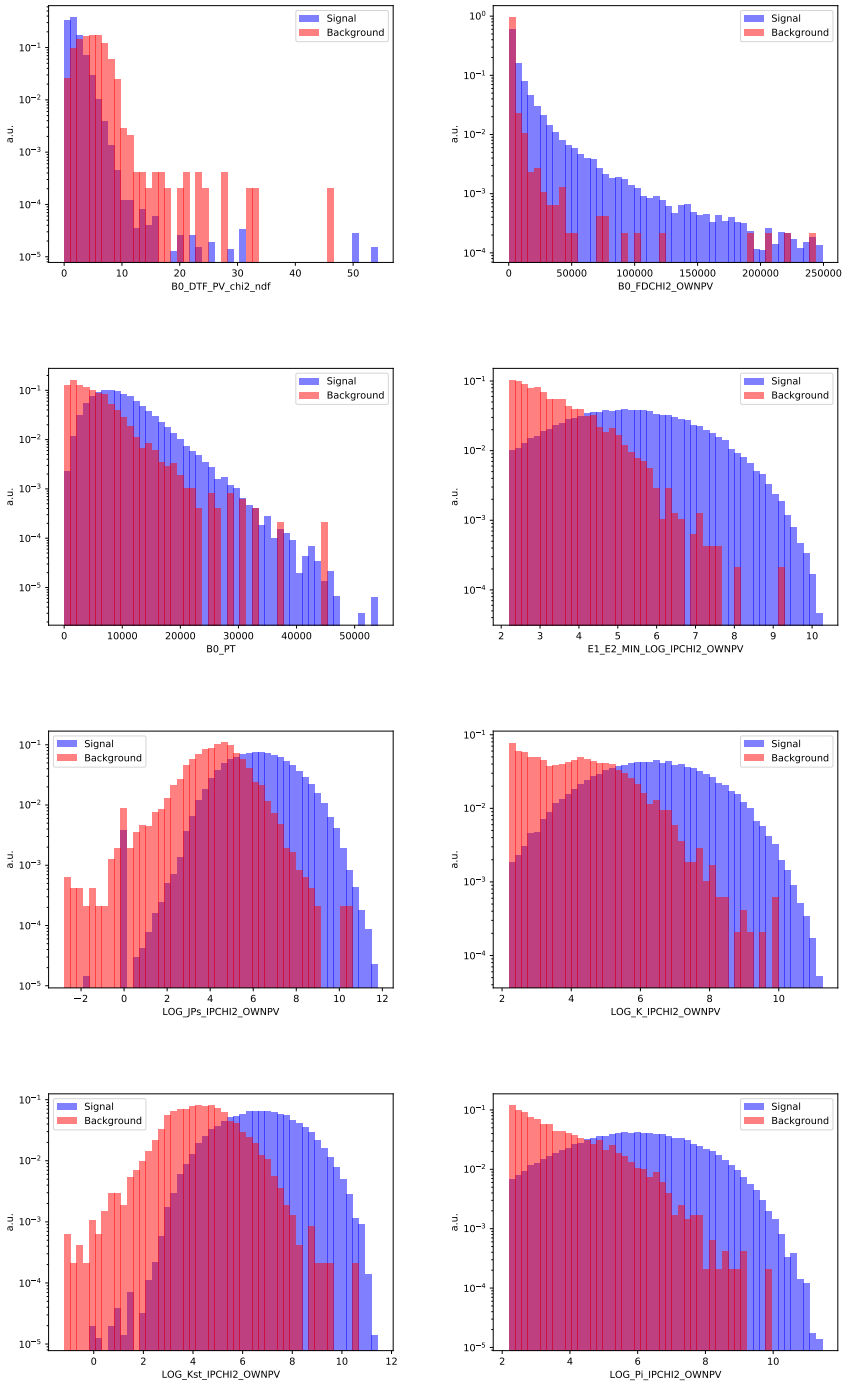


Figure B.3: Distributions of the variables used to train the R1 BDT. The signal training sample is shown in blue, the background training sample is shown in red.

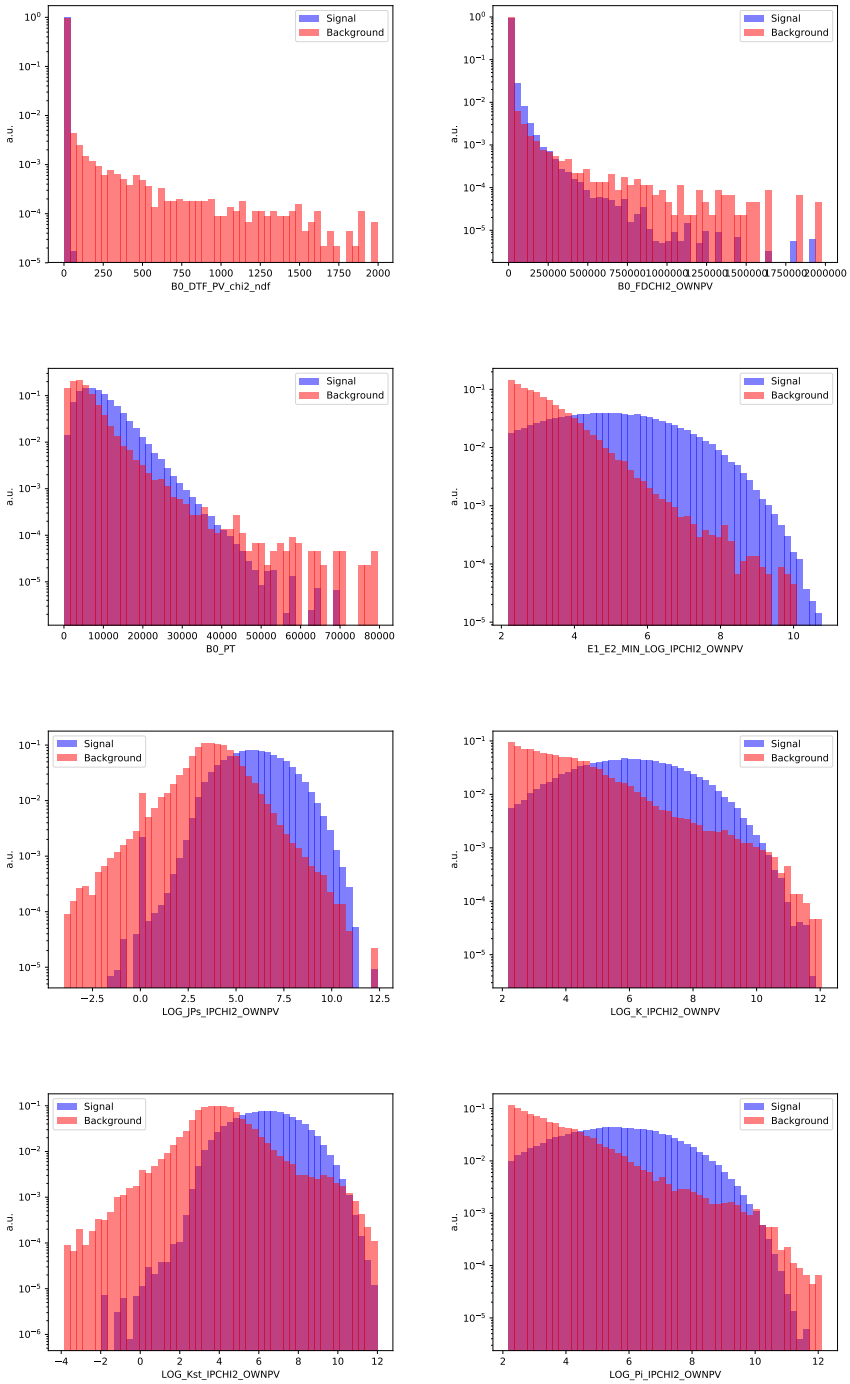


Figure B.4: Distributions of the variables used to train the R2 BDT. The signal training sample is shown in blue, the background training sample is shown in red.

B.2 Training of the BDT

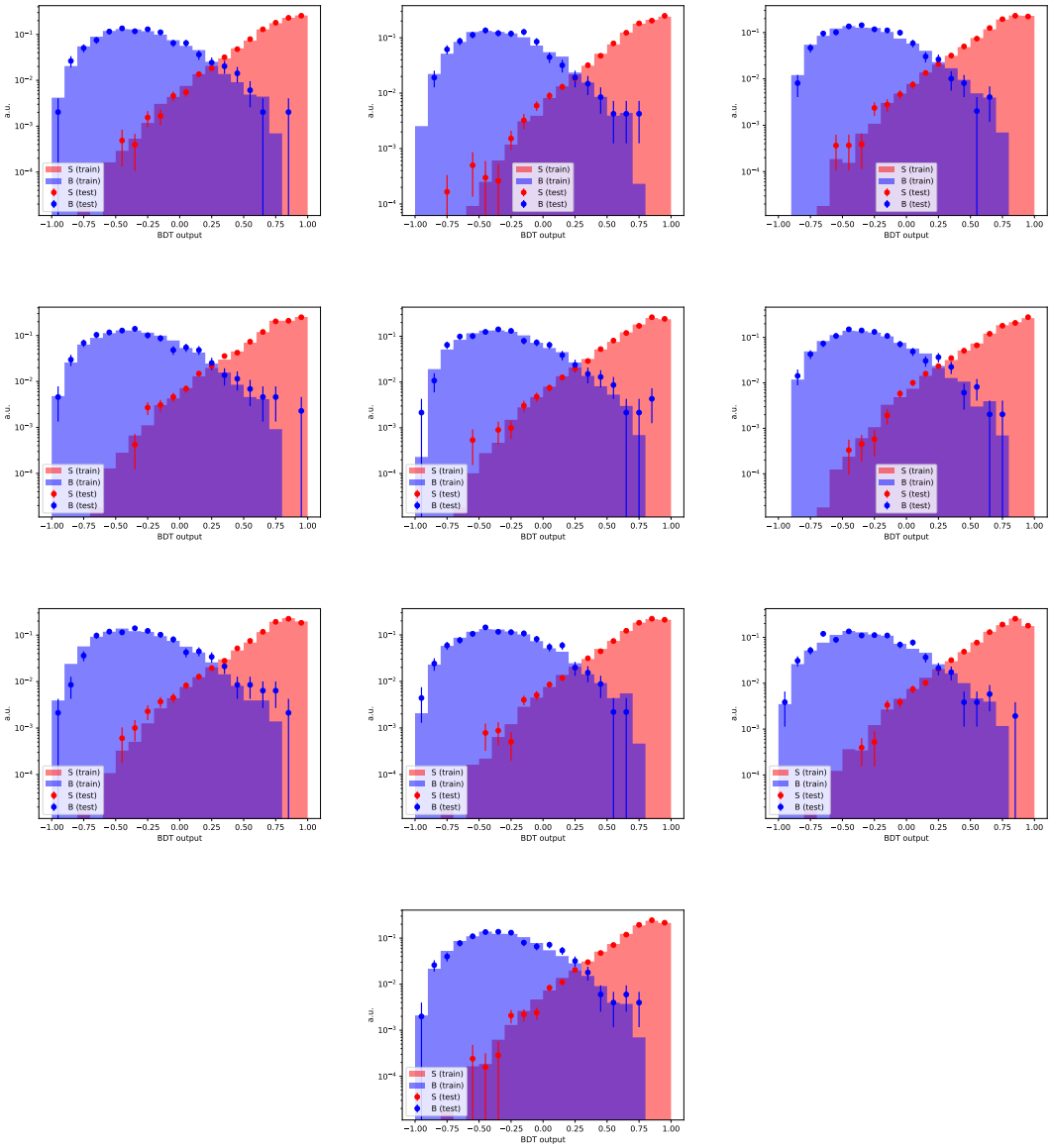


Figure B.5: BDT output of the ten different folds for R1. The test samples (dot) of the signal sample (red) and the background sample (blue) are shown on top of the training samples (histograms).

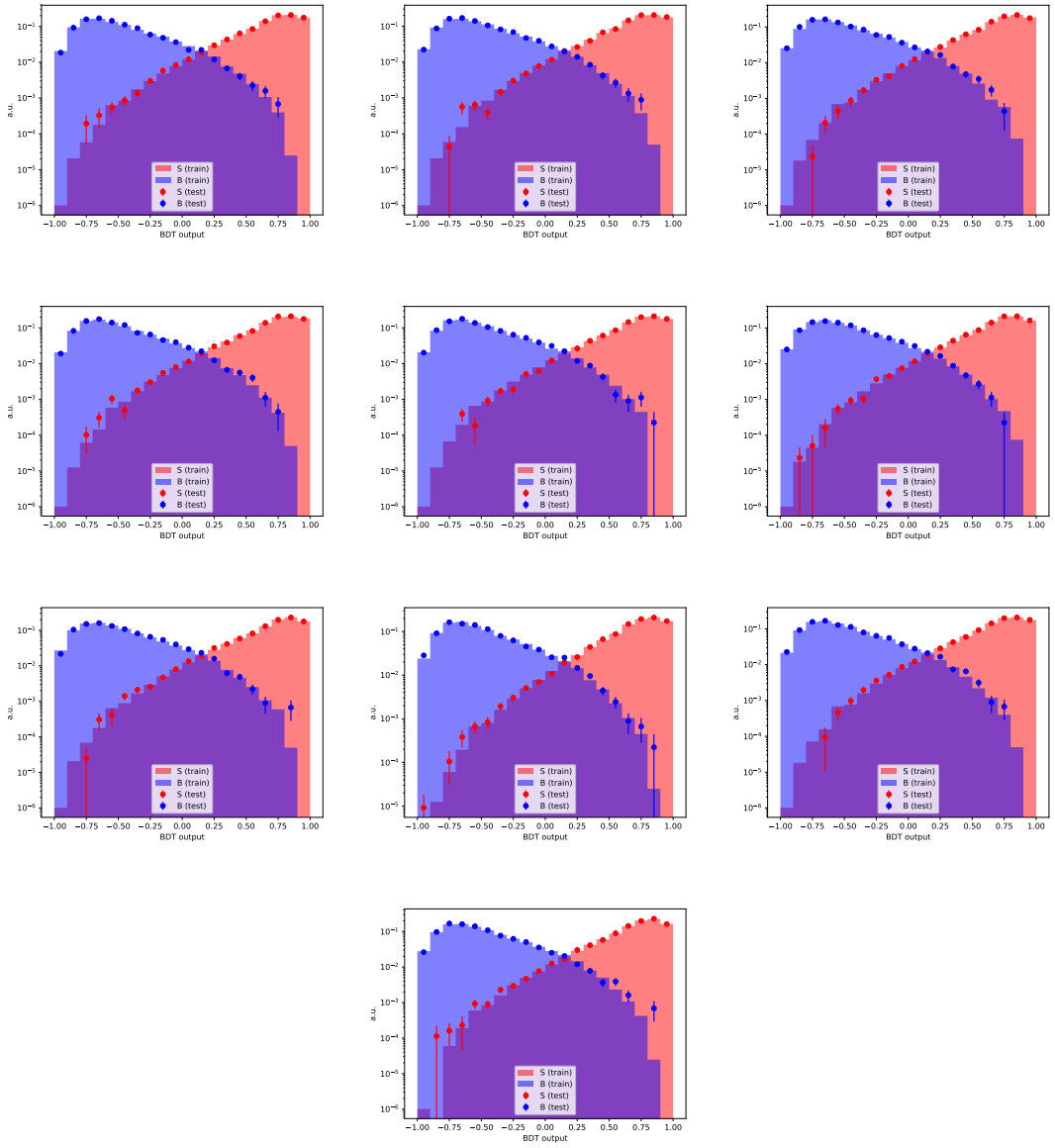


Figure B.6: BDT output of the ten different folds for R2. The test samples (dot) of the signal sample (red) and the background sample (blue) are shown on top of the training samples (histograms).

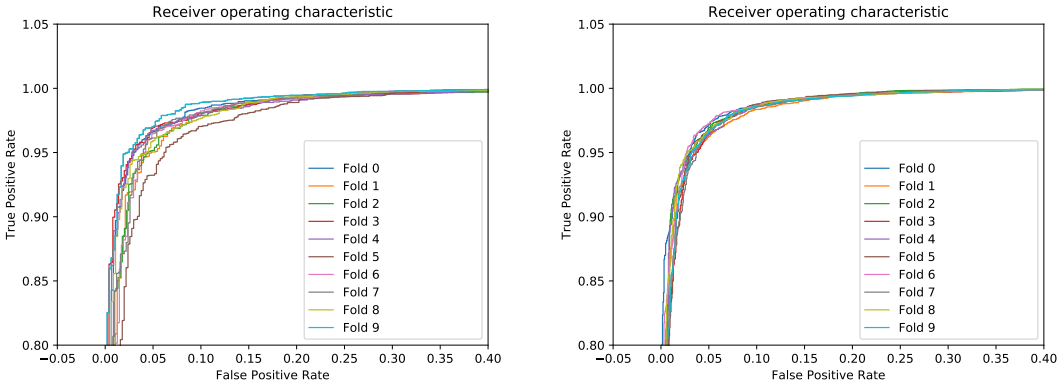


Figure B.7: ROC curves of all folds for R1 (left) and R2 (right).

Appendix C

Corrections to simulation

C.1 $B^0 \to K^{*0} e^+ e^-$ MC angular distributions

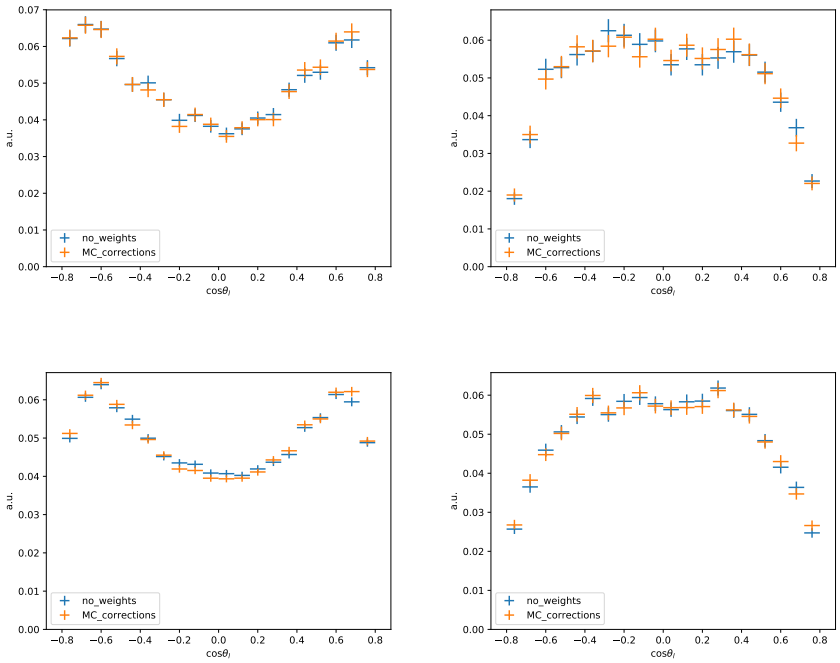


Figure C.1: Distribution of the $\cos\theta_\ell$ angle in the $B^0 \to K^{*0} e^+ e^-$ MC for R1 (R2) on the top (bottom) and in the L0L (L0I) trigger category on the left (right). The bare MC without any reweighting is shown in blue while the distribution including the MC/data corrections is shown in orange.

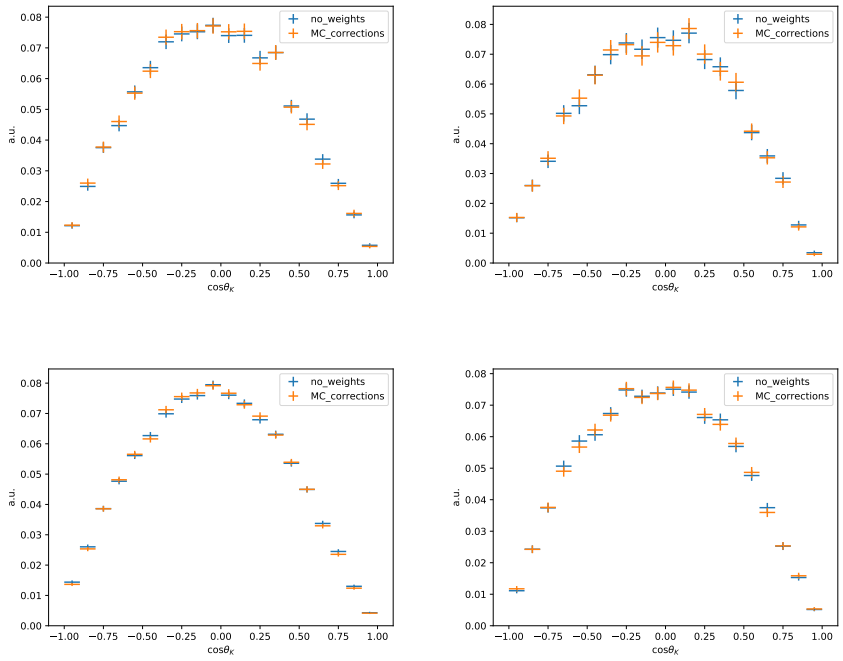


Figure C.2: Distribution of the $\cos\theta_K$ angle in the $B^0 \to K^{*0}e^+e^-$ MC for R1 (R2) on the top (bottom) and in the L0L (L0I) trigger category on the left (right). The bare MC without any reweighting is shown in blue while the distribution including the MC/data corrections is shown in orange.

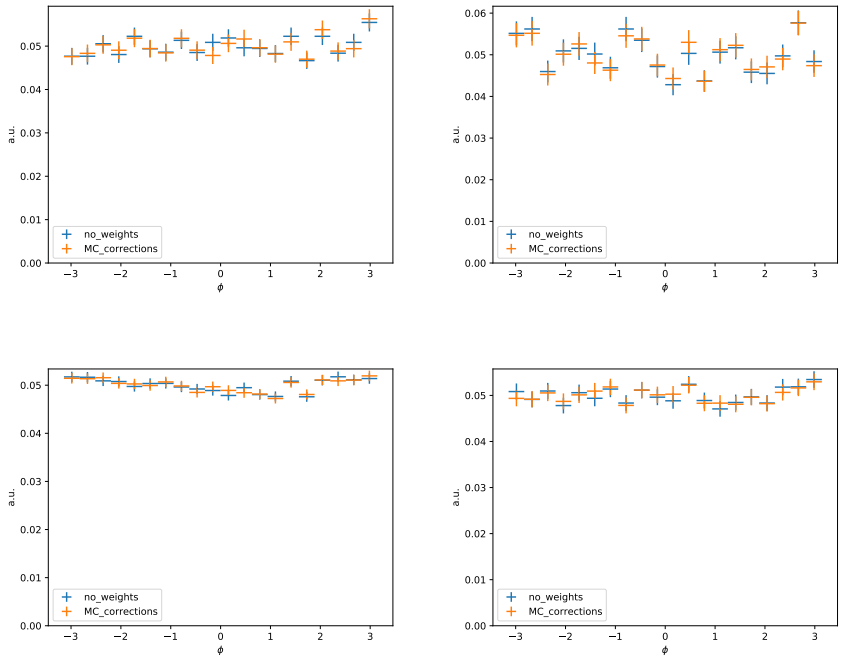


Figure C.3: Distribution of the ϕ angle in the $B^0 \to K^{*0}e^+e^-$ MC for R1 (R2) on the top (bottom) and in the L0L (L0I) trigger category on the left (right). The bare MC without any reweighting is shown in blue while the distribution including the MC/data corrections is shown in orange.

C.2 Angular acceptance

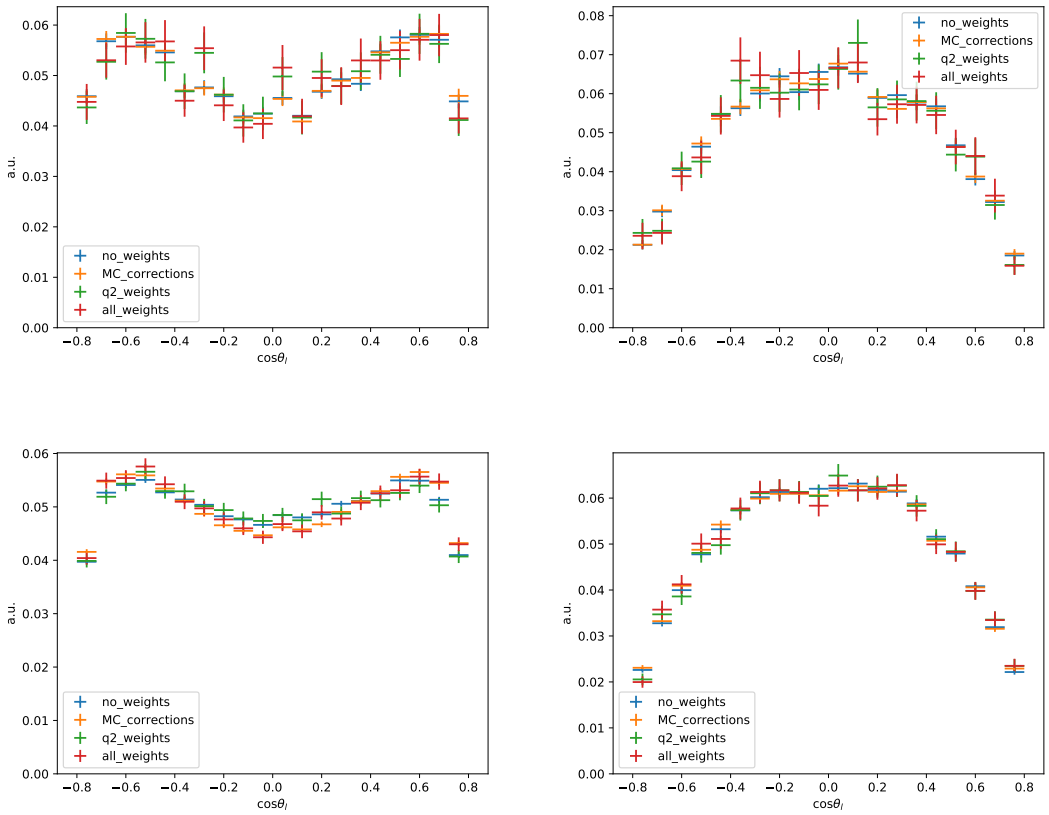


Figure C.4: Distribution of the $\cos\theta_\ell$ angle in the phase space $B^0 \to K^{*0}e^+e^-$ MC for R1 (R2) on the top (bottom) and in the L0L (L0I) trigger category on the left (right). The bare MC without any reweighting is shown in blue, the distribution including the MC/data corrections is shown in orange, the distribution including the q^2 reweighting is shown in green and the distribution including both MC/data and q^2 corrections is shown in red.

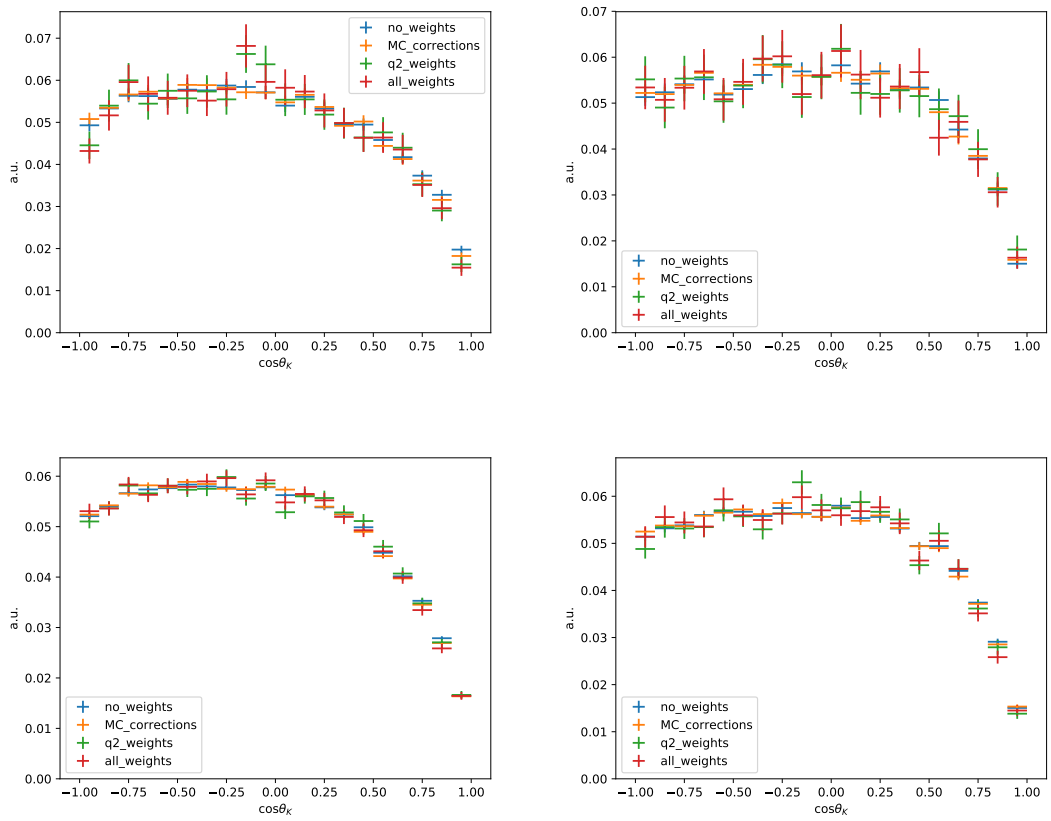


Figure C.5: Distribution of the $\cos\theta_K$ angle in the phase space $B^0 \to K^{*0}e^+e^-$ MC for R1 (R2) on the top (bottom) and in the L0L (L0I) trigger category on the left (right). The bare MC without any reweighting is shown in blue, the distribution including the MC/data corrections is shown in orange, the distribution including the q^2 reweighting is shown in green and the distribution including both MC/data and q^2 corrections is shown in red.

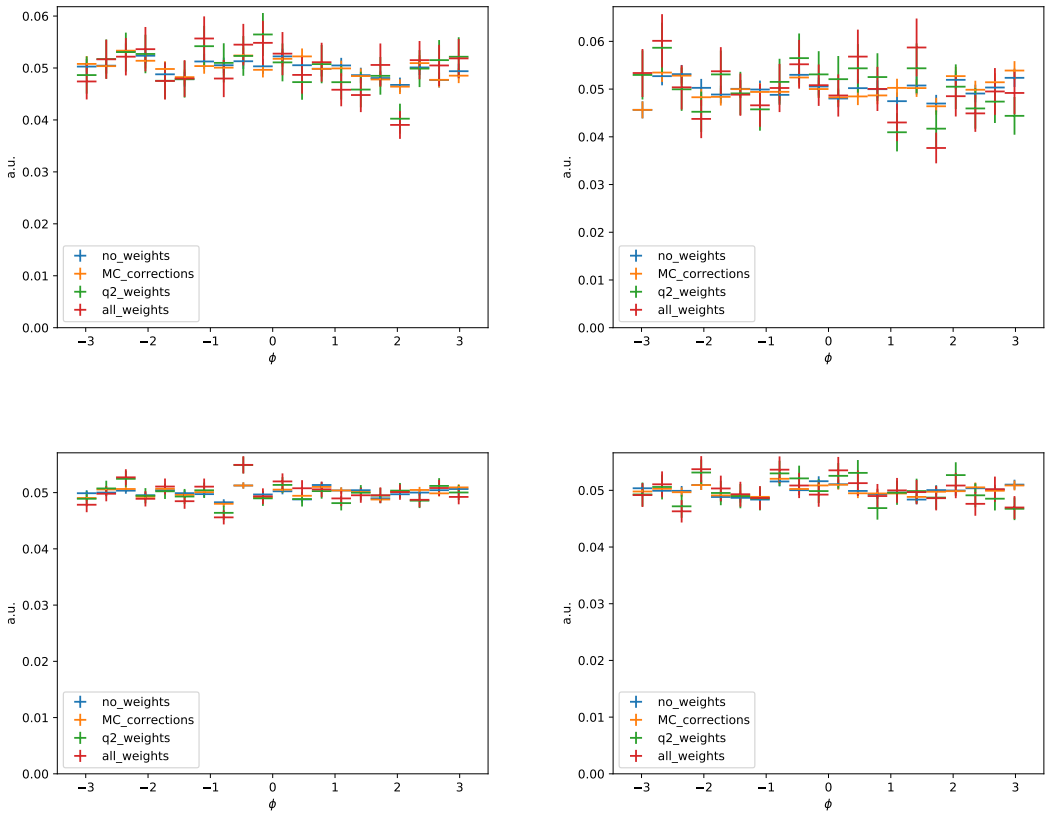


Figure C.6: Distribution of the ϕ angle in the phase space $B^0 \to K^{*0}e^+e^-$ MC for R1 (R2) on the top (bottom) and in the L0L (L0I) trigger category on the left (right). The bare MC without any reweighting is shown in blue, the distribution including the MC/data corrections is shown in orange, the distribution including the q^2 reweighting is shown in green and the distribution including both MC/data and q^2 corrections is shown in red.

Appendix D

Bremsstrahlung recovery

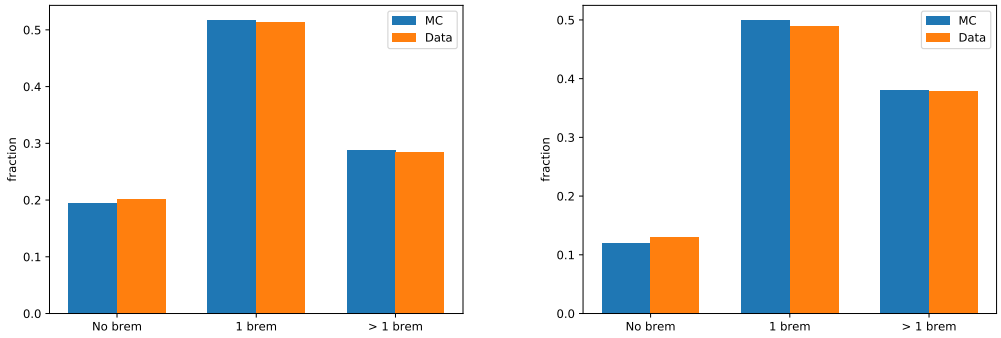


Figure D.1: Fraction of events in each bremstrahlung category for $B^0 \to K^{*0} J/\psi (\to e^+e^-)$ MC (blue) and data (orange): no bremstrahlung recovered, one cluster or more than one cluster recovered. The full selection is applied on both samples. To remove the residual background a cut $\pm 60\,\mathrm{MeV}/c^2$ on $m(K^+\pi^-e^+e^-)$ is performed around the nominal B^0 mass. The left plot shows the L0L trigger category while the right plot shows the L0I trigger category.

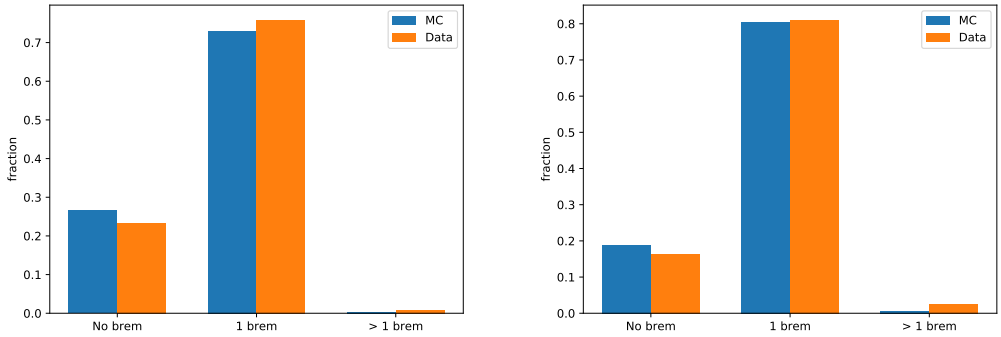


Figure D.2: Fraction of events in each bremstrahlung category for $B^0 \to K^{*0}\gamma(\to e^+e^-)$ MC (blue) and data (orange): no bremstrahlung recovered, one cluster or more than one cluster recovered. The full selection is applied on both samples. To remove the residual background a cut $\pm 60\,\mathrm{MeV}/c^2$ on $m(K^+\pi^-e^+e^-)$ is performed around the nominal B^0 mass. The left plot shows the L0L trigger category while the right plot shows the L0I trigger category.

Appendix E

Angular and mass fit validation

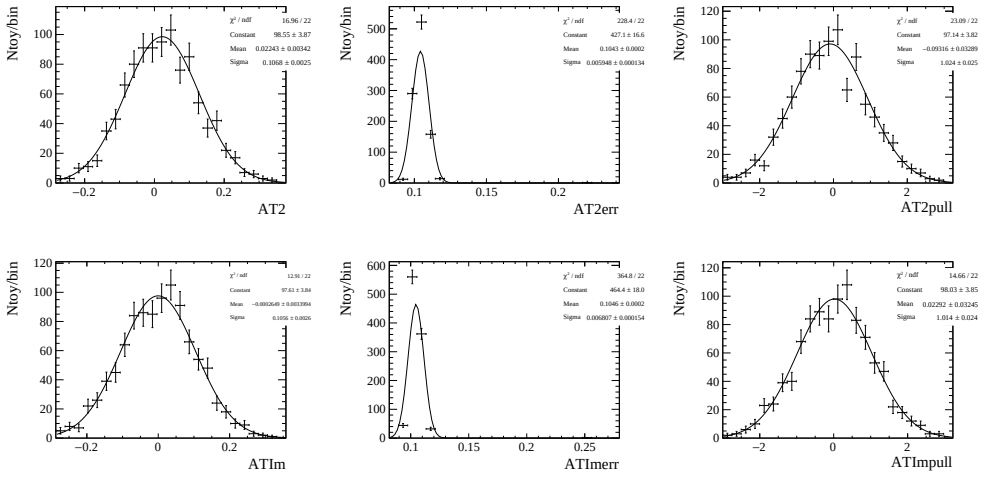


Figure E.1: Fitted values (left), errors (middle) and pulls (right) of the 1000 toy datasets for the parameters $A_T^{(2)}$ (top) and A_T^{Im} (bottom). All parameters have been generated with SM values.

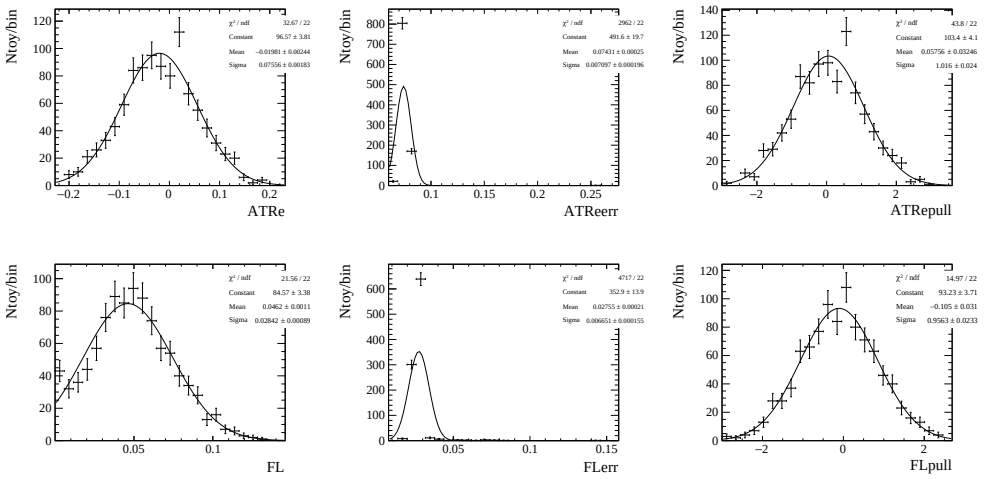


Figure E.2: Fitted values (left), errors (middle) and pulls (right) of the 1000 toy datasets for the parameters A_T^{Re} (top) and F_{L} (bottom). All parameters have been generated with SM values.

Bibliography

[1] K. G. Wilson, *Nonlagrangian models of current algebra*, Phys. Rev. **179** (1969) 1499.

[2] D. Becirevic, E. Kou, A. Le Yaouanc, and A. Tayduganov, *Future prospects for the determination of the Wilson coefficient $C'_{7\gamma}$*, JHEP **08** (2012) 090, arXiv:1206.1502.

[3] L. L. Everett *et al.*, *Alternative approach to $b \to s\gamma$ in the uMSSM*, JHEP **01** (2002) 022, arXiv:hep-ph/0112126.

[4] 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, arXiv:hep-ph/0604121.

[5] E. Lunghi and J. Matias, *Huge right-handed current effects in $B \to K^*(K\pi)l^+l^-$ in supersymmetry*, JHEP **04** (2007) 058, arXiv:hep-ph/0612166.

[6] E. Kou, C.-D. Lü, and F.-S. Yu, *Photon Polarization in the $b \to s\gamma$ processes in the Left-Right Symmetric Model*, JHEP **12** (2013) 102, arXiv:1305.3173.

[7] HFLAV, Y. S. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2018*, arXiv:1909.12524.

[8] LHCb, R. Aaij *et al.*, *Measurement of CP-violating and mixing-induced observables in $B_s^0 \to \phi\gamma$ decays*, Phys. Rev. Lett. **123** (2019) 081802, arXiv:1905.06284.

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

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

[11] LHCb, E. Jans, *The LHCb detector*, in *Particles and fields. Proceedings, Meeting of the Division of the American Physical Society, DPF 2009, Detroit, USA, July 26-31, 2009*, 2009, arXiv:0910.1740.

[12] D. Straub, *flavio: a Python package for flavour and precision phenomenology in the Standard Model and beyond*, arXiv:1810.08132.

[13] LHCb, R. Aaij *et al.*, *Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$*, Phys. Rev. Lett. **122** (2019) 222001, arXiv:1904.03947.

[14] M. Thomson, *Modern Particle Physics*, Cambridge University Press, 2013.

[15] D. M. Schwartz, *Quantum Field Theory and the Standard Model*, Cambridge University Press, 2014.

[16] L. Wolfenstein, *Parametrization of the kobayashi-maskawa matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[17] Super-Kamiokande, Y. Fukuda *et al.*, *Evidence for oscillation of atmospheric neutrinos*, Phys. Rev. Lett. **81** (1998) 1562, arXiv:hep-ex/9807003.

[18] Planck, Y. Akrami *et al.*, *Planck 2018 results. I. Overview and the cosmological legacy of Planck*, arXiv:1807.06205.

[19] L. Canetti, M. Drewes, and M. Shaposhnikov, *Matter and Antimatter in the Universe*, New J. Phys. **14** (2012) 095012, arXiv:1204.4186.

[20] A. Canepa, *Searches for Supersymmetry at the Large Hadron Collider*, Rev. Phys. **4** (2019) 100033.

[21] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the 2π Decay of the K_2^0 Meson*, Phys. Rev. Lett. **13** (1964) 138.

[22] M. Kobayashi and T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.

[23] M. L. Perl *et al.*, *Evidence for Anomalous Lepton Production in e+ - e- Annihilation*, Phys. Rev. Lett. **35** (1975) 1489, [,193(1975); ,193(1975)].

[24] S. W. Herb *et al.*, *Observation of a Dimuon Resonance at 9.5-GeV in 400-GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (1977) 252.

[25] S. Bifani, S. Descotes-Genon, A. Romero Vidal, and M.-H. Schune, *Review of Lepton Universality tests in B decays*, J. Phys. **G46** (2019) 023001, arXiv:1809.06229.

[26] M. Misiak *et al.*, *Updated NNLO QCD predictions for the weak radiative B-meson decays*, Phys. Rev. Lett. **114** (2015) 221801, arXiv:1503.01789.

[27] HFLAV, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016*, Eur. Phys. J. **C77** (2017) 895, arXiv:1612.07233.

[28] M. Gronau and D. Pirjol, *Photon polarization in radiative B decays*, Phys. Rev. **D66** (2002) 054008, arXiv:hep-ph/0205065.

[29] LHCb, R. Aaij *et al.*, *Observation of Photon Polarization in the b->sγ Transition*, Phys. Rev. Lett. **112** (2014) 161801, arXiv:1402.6852.

[30] C. Marin Benito, *Pushing the boundaries of the LHCb rare decays program: search for the $\Lambda_b \to \Lambda\gamma$ decay*, CERN-THESIS-2018-401, 2018, Presented 04 Apr 2018.

[31] F. Legger and T. Schietinger, *Polarized radiative Lambda/b decays at LHCb*, CERN-LHCB-2006-013, 2006.

[32] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. D **98** (2018) 030001.

[33] LHCb, R. Aaij *et al.*, *Measurements of the $\Lambda_b^0 \to J/\psi\Lambda$ decay amplitudes and the Λ_b^0 polarisation in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. **B724** (2013) 27, arXiv:1302.5578.

[34] G. Hiller and A. Kagan, *Probing for new physics in polarized Λ_b decays at the Z*, Phys. Rev. **D65** (2002) 074038, arXiv:hep-ph/0108074.

[35] LHCb, R. Aaij *et al.*, *First Observation of the Radiative Decay $\Lambda_b^0 \to \Lambda\gamma$*, Phys. Rev. Lett. **123** (2019) 031801, arXiv:1904.06697.

[36] M. Borsato, *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*, CERN-THESIS-2015-219, 2015.

[37] S. Jäger and J. Martin Camalich, *Reassessing the discovery potential of the $B \to K^*\ell^+\ell^-$ decays in the large-recoil region: SM challenges and BSM opportunities*, Phys. Rev. **D93** (2016) 014028, arXiv:1412.3183.

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

[39] D. Becirevic and E. Schneider, *On transverse asymmetries in $B \to K^* l^+ l^-$*, Nucl. Phys. **B854** (2012) 321, arXiv:1106.3283.

[40] W. Altmannshofer *et al.*, *Symmetries and Asymmetries of $B \to K^* \mu^+ \mu^-$ Decays in the Standard Model and Beyond*, JHEP **01** (2009) 019, arXiv:0811.1214.

[41] 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.

[42] J. Lefrançois and M. H. Schune, *Measuring the photon polarization in b->s gamma using the B-> K*e+e- decay channel*, LHCb-PUB-2009-008, 2009.

[43] Belle, A. Abdesselam *et al.*, *Angular analysis of $B^0 \to K^*(892)^0 \ell^+ \ell^-$*, in *Proceedings, LHCSki 2016 - A First Discussion of 13 TeV Results: Obergurgl, Austria, April 10-15, 2016*, 2016, arXiv:1604.04042.

[44] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[45] ATLAS, G. Aad *et al.*, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[46] CMS, S. Chatrchyan *et al.*, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.

[47] ALICE, K. Aamodt *et al.*, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[48] LHCb, A. A. Alves, Jr. *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[49] M. Pepe Altarelli and F. Teubert, *B Physics at LHCb*, Int. J. Mod. Phys. A **23** (2008) 5117, arXiv:0802.1901.

[50] A. Abashian *et al.*, *The Belle Detector*, Nucl. Instrum. Meth. **A479** (2002) 117.

[51] Belle-II, T. Abe *et al.*, *Belle II Technical Design Report*, arXiv:1011.0352.

[52] BaBar, B. Aubert *et al.*, *The BaBar detector*, Nucl. Instrum. Meth. **A479** (2002) 1, arXiv:hep-ex/0105044.

[53] F. Follin and D. Jacquet, *Implementation and experience with luminosity levelling with offset beam*, in *Proceedings, ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders (BB2013): CERN, Geneva, Switzerland, March 18-22 2013*, 183–187, 2014, arXiv:1410.3667. [,183(2014)].

[54] LHCb, R. Aaij *et al.*, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019) P04013, arXiv:1812.10790.

[55] LHCb, R. Aaij *et al.*, *LHCb Detector Performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[56] LHCb RICH Group, M. Adinolfi *et al.*, *Performance of the LHCb RICH detector at the LHC*, Eur. Phys. J. **C73** (2013) 2431, arXiv:1211.6759.

[57] LHCb, D. Derkach *et al.*, *Machine-Learning-based global particle-identification algorithms at the LHCb experiment*, J. Phys. Conf. Ser. **1085** (2018) 042038.

[58] LHCb, *LHCb technical design report: Reoptimized detector design and performance*, CERN-LHCC-2003-030, 2003.

[59] LHCb, R. Aaij *et al.*, *Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell^-$ decays*, JHEP **08** (2017) 055, arXiv:1705.05802.

[60] LHCb, R. Aaij *et al.*, *Trigger plots and diagrams for conferences*, 2020.

[61] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.

[62] I. Belyaev *et al.*, *Handling of the generation of primary events in Gauss, the LHCb simulation framework*, J. Phys. Conf. Ser. **331** (2011) 032047.

[63] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152.

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

[65] Geant4 collaboration, S. Agostinelli *et al.*, *Geant4: A simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[66] Geant4 collaboration, J. Allison *et al.*, *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53** (2006) 270.

[67] LHCb, I. Bediaga *et al.*, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001, 2014.

[68] LHCb, O. Steinkamp, *LHCb Upgrades*, J. Phys. Conf. Ser. **1271** (2019) 012010.

[69] LHCb, E. Lemos Cid and P. Vázquez Regueiro, *The LHCb Vertex Locator Upgrade*, PoS **Vertex 2017** (2018) 002.

[70] T. Poikela *et al.*, *The VeloPix ASIC*, JINST **12** (2017) C01070.

[71] LHCb, I. Bediaga *et al.*, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021, 2013.

[72] S. Bugiel *et al.*, *SALT, a dedicated readout chip for high precision tracking silicon strip detectors at the LHCb Upgrade*, JINST **11** (2016) C02028.

[73] LHCb RICH, S. Easo, *Overview of LHCb-RICH upgrade*, Nucl. Instrum. Meth. **A876** (2017) 160.

[74] LHCb RICH, L. Cassina, *Photodetectors and front-end electronics for the LHCb RICH upgrade*, Nucl. Instrum. Meth. **A876** (2017) 217, arXiv:1611.00406.

[75] D. Pinci and M. Santimaria, *Study of the GEM chamber for the upgrade of the LHCb muon system*, Nuovo Cim. **C39** (2016) 265.

[76] LHCb Collaboration, R. Aaij *et al.*, *LHCb Trigger and Online Upgrade Technical Design Report*, CERN-LHCC-2014-016, 2014.

[77] Y. S. Amhis *et al.*, *The Front-End board of the upgraded LHCb Calorimeter*, Tech. Rep. LHCb-INT-2019-004. CERN-LHCb-INT-2019-004, CERN, Geneva, 2019.

[78] M. De Cian, S. Farry, P. Seyfert, and S. Stahl, *Fast neural-net based fake track rejection in the LHCb reconstruction*, LHCb-PUB-2017-011, 2017.

[79] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021, 2016.

[80] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.

[81] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, LHCb-PUB-2014-039, 2014.

[82] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, J. Phys. Conf. Ser. **762** (2016) , arXiv:1608.05806, https://github.com/arogozhnikov/hep_ml.

[83] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, arXiv:physics/0402083.

[84] Y.-S. Tsai, *Pair Production and Bremsstrahlung of Charged Leptons*, Rev. Mod. Phys. **46** (1974) 815, [Erratum: Rev.Mod.Phys. 49, 521–423 (1977)].

[85] LHCb collaboration, F. Desse, J. Lefrancois, and M.-H. Schune, *On the photon conversion in Geant4*, LHCb-INT-2019-001, 2019.

[86] M. Alexander *et al.*, *Mapping the material in the LHCb vertex locator using secondary hadronic interactions*, JINST **13** (2018) P06008, arXiv:1803.07466.

[87] N. Q. et al, *Event generators for $\eta/\prime\eta$ decays ar BESIII*, Chinese Phys. **42 013001** (2018).

[88] R. H. Dalitz, *On an alternative decay process for the neutral pi-meson, Letters to the Editor*, Proc. Phys. Soc. **A64** (1951) 667.

[89] V. J. Rives Molina, *Study of b-hadron decays into two hadrons and a photon at LHCb and first observation of b-baryon radiative decays*, CERN-THESIS-2016-158, 2016-07-07.

[90] F. Legger and T. Schietinger, *Photon helicity in $\Lambda_b \to pK\gamma$ decays*, Phys. Lett. **B645** (2007) 204, arXiv:hep-ph/0605245, [Erratum: Phys. Lett.B647,527(2007)].

[91] LHCb, R. Aaij *et al.*, *Study of the $B^0 \to \rho(770)^0 K^*(892)^0$ decay with an amplitude analysis of $B^0 \to (\pi^+\pi^-)(K^+\pi^-)$ decays*, JHEP **05** (2019) 026, arXiv:1812.07008.

[92] A. A. Gallas Torreira *et al.*, *Study of the $B^0 \to \rho^0(770)K^*(892)^0$ mode and amplitude analysis of $B^0 \to (\pi^\pm\pi^\mp)(K^+\pi^-)$ decays*, LHCb-ANA-2016-076, 2016.

[93] LHCb collaboration, M. Borsato, S. M.-H. and F. Polci, *HOP: an additional tool for decays involving electrons*, LHCb-INT-2015-037, 2015.

[94] F. Pedregosa *et al.*, *Scikit-learn: Machine learning in Python*, J. Machine Learning Res. **12** (2011) 2825, arXiv:1201.0490, and online at http://scikit-learn.org/stable/.

[95] K. S. Cranmer, *Kernel estimation in high-energy physics*, Comput. Phys. Commun. **136** (2001) 198, arXiv:hep-ex/0011057.

[96] R. Brun and F. Rademakers, *ROOT: An objected oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997) 81.

[97] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi p K^-$ decays*, Phys. Rev. Lett. **115** (2015) 072001, arXiv:1507.03414.

[98] LHCb collaboration, R. Aaij *et al.*, *Measurement of b hadron production fractions in 7 TeV pp collisions*, Phys. Rev. **D85** (2012) 032008, arXiv:1111.2357.

[99] LHCb collaboration, R. Aaij *et al.*, *Measurement of b-hadron fractions in 13 TeV pp collisions*, arXiv:1902.06794, submitted to Phys. Rev. Lett.

[100] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the $B^0 \to K^* e^+ e^-$ decay in the low-q^2 region*, JHEP **04** (2015) 064, arXiv:1501.03038.

[101] A. Paul and D. M. Straub, *Constraints on new physics from radiative B decays*, JHEP **04** (2017) 027, arXiv:1608.02556.

[102] LHCb, R. Aaij *et al.*, *Angular analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay using 3 fb^{-1} of integrated luminosity*, JHEP **02** (2016) 104, arXiv:1512.04442.

[103] M. Chrzaszcz *et al.*, *Prospects for disentangling long- and short-distance effects in the decays $B \to K^*\mu^+\mu^-$*, JHEP **10** (2019) 236, arXiv:1805.06378.

[104] LHCb, R. Aaij *et al.*, *Angular analysis and differential branching fraction of the decay $B_s^0 \to \phi\mu^+\mu^-$*, JHEP **09** (2015) 179, arXiv:1506.08777.

Remerciements

I would like to thank Cristina Lazzeroni and Tim Gershon who kindly accepted to be rapporteurs of my thesis, as well as Aoife Bharucha, Florian Beaudette and Achille Stocchi who accepted to be part of my jury.
Un grand merci à Marie-Hélène Schune, ma directrice de thèse. Merci de m'avoir pris sous ton aile durant ces 3 ans, j'ai beaucoup appris à tes côtés. Tu as su trouver l'équilibre idéal en me laissant jouir d'une grande autonomie à laquelle je tiens tant tout en étant toujours disponible quand j'avais besoin de tes précieux conseils. Au-delà du côté scientifique, merci pour ces moments de convivialité, ton humanité et spontaniéité.
Je voudrais également remercier toute l'équipe qui a participé à cette analyse: Jacques Lefrançois, mon consultant de luxe, ce fut un réel plaisir et un honneur d'échanger avec toi. J'ai rarement croisé dans ma vie une personne avec un ratio skill/ego aussi élevé. Martino Borsato, Dr. de l'Univers et mon prof de statistique préféré. Ton arrivée pour la dernière année fut la meilleure surprise de ma thèse, aussi bien sur le plan scientifique qu'humain. Und natürlich Jonathan Brandt, vielen Dank für deine Hilfe, es hat wirklich spass gemacht, zusammen zu arbeiten.
Je souhaiterais également remercier le groupe LHCb du LAL-IJCLab au sein duquel j'ai eu la chance de travailler. Merci en particulier à Frédéric Machefert et Alexis Vallier pour m'avoir fait découvrir la physique des particules et LHCb, c'est sans doute un peu de votre faute si cette thèse existe. Merci à toute l'équipe des pingouins: Yasmine Amhis toujours de bonne humeur (mais comment fais-tu ??), Carla Marin Benito dont on ressort du bureau toujours la bouche pleine de turron, Vitalii Lisovskyi who's jokes are almost as unexpected as Elisabeth's ones, Anja Beck my special advisor in spherical harmonics and D functions, and Felicia Volle for our funny discussions at the delicious CESFO. Merci également à Patrick Robbe pour ton aide sur le LLT, pour avoir été un super prof de physique des particules en M2 et pour ton humour pince-sans-rire. Merci à Sergey Barsuk pour ta sympathie et tes légendaires chocolats ukrainiens. Merci à Michael Winn,

l'homme au bateau sur le toit, pour ces longues discussions de physique mais aussi et surtout politiques et philosophiques que nous avons eu dans notre magnifique bureau flambant neuf. Merci à Fransceso Bossu, qui venait régulièrement se joindre aux débats, sans doute aussi pour profiter de la vue de notre bureau. Et bien entendu, merci à tous les autres (anciens et présents) doctorants du groupe: the mysterious Andrii Usachov and his weird fun facts, Victor Daussy-Renaudin le conseillé France Culture du laboratoire et Elisabeth Niel dont la qualité des blagues me rassure quotidiennement sur le fait que je suis loin de toucher le fond avec les miennes (et ça c'est très précieux). Merci également à tous les participants des movie nights, que j'ai lâchement abandonnées quand je suis devenu papa (meilleure excuse du monde), je garde un très bon souvenir de ces soirées entre doctorants du LAL et d'ailleurs.
Last but not least, un grand merci à Marie-Claire Lefevre pour m'avoir soutenu et supporté durant ces trois années de thèse, avec ses hauts et ses bas.

université PARIS-SACLAY

ÉCOLE DOCTORALE
Particules, hadrons, énergie et noyau : instrumentation, imagerie, cosmos et simulation (PHENIICS)

Titre: Analyse angulaire de désintégrations $B^0 \to K^{*0}e^+e^-$ avec le détecteur LHCb et upgrade de l'électronique des calorimètres

Mots clés: Physique de la saveur, LHCb, physique du B, polarisation du photon, électrons, désintégrations rares

Résumé: Les courants neutres changeant la saveur de type $b \to s\gamma$ ne sont autorisés qu'au niveau des boucles dans le Modèle Standard (MS). Ils sont donc potentiellement sensibles à des effets de Nouvelle Physique (NP) intervenant dans les boucles quantiques qui pourraient se manifester via la modification de certaines observables angulaires. Dans le MS, le photon est principalement émis avec une polarisation gauche. Cependant, plusieurs théories de NP autorisent une importante contribution des courants droits. L'analyse angulaire des désintégrations $B^0 \to K^{*0}e^+e^-$ à très bas q^2 (la masse invariante de la paire de dielectron au carré) permet d'étudier la structure de l'hélicité des transitions $b \to s\gamma$ grâce à la contribution dominante du photon virtuel couplé à la paire de dielectron à très bas q^2.

Cette thèse présente l'analyse angulaire de désintégrations $B^0 \to K^{*0}e^+e^-$ utilisant l'ensemble des collisions proton-proton enregistrées par l'expérience LHCb durant les Run 1 et 2 entre 2011 et 2018, représentant une luminosité intégrée de $9.1\,\text{fb}^{-1}$ aux énergies dans le centre de masse allant de 7 à 13 TeV. Les asymétries transverses dans la région de q^2 $[0.0001, 0.25]\,\text{GeV}^2/c^4$ sont mesurées à hauteur de $A_T^{(2)} = 0.106 \pm 0.103^{+0.016}_{-0.017}$, $A_T^{Im} = 0.015 \pm 0.102 \pm 0.012$, $A_T^{Re} = -0.064 \pm 0.077 \pm 0.015$ et $F_\text{L} = 0.044 \pm 0.026 \pm 0.014$. Ces mesures représentent actuellement la plus importante contrainte mondiale sur la polarisation droite du photon dans les transitions $b \to s\gamma$.

Le détecteur LHCb subit une importante mise à jour en 2019/2020 afin de faire fonctionner l'expérience à une luminosité instantanée de $2 \times 10^{33}\ \text{cm}^{-2}\text{s}^{-1}$. L'électronique des calorimètres doit notamment être mise à niveau avec de nouvelles Cartes Front-End (CFE) adaptées à une lecture à 40 MHz. Cette thèse présente également l'upgrade du module Low Level Trigger de la CFE qui est notamment responsable de l'identification des cluster d'énergie transverse maximale pour chaque évènement.

Title: Angular analysis of $B^0 \to K^{*0} e^+ e^-$ decays with the LHCb detector and upgrade of the electronics of the calorimeters

Keywords: Flavor physics, LHCb, B physics, photon polarization, electrons, rare decay

Abstract: Flavor changing neutral current processes of the type $b \to s\gamma$ are forbidden at tree level in the Standard Model (SM). They are thus sensitive to potential New Physics (NP) effects occurring via loop processes which may manifest themselves through the modification of angular observables. In the SM, the photon is predominantly left handed. However several NP theories allow large right-handed currents. The angular analysis of $B^0 \to K^{*0} e^+ e^-$ decays at very low q^2 (the dielectron pair invariant mass squared) allows to study the helicity structure of $b \to s\gamma$ transitions thanks to the dominant contribution from the virtual photon coupling to the dielectron pair at very low q^2.

This thesis presents the angular analysis of $B^0 \to K^{*0} e^+ e^-$ decays using the full Run 1 and Run 2 proton-proton collision datasets collected between 2011 and 2018 at the LHCb experiment, corresponding to an integrated luminosity of $9.1\,\text{fb}^{-1}$ at centre-of-mass energies ranging from 7 to 13 TeV. The transverse asymmetries in the q^2 range $[0.0001, 0.25]\,\text{GeV}^2/c^4$ are measured as $A_T^{(2)} = 0.106 \pm 0.103^{+0.016}_{-0.017}$, $A_T^{Im} = 0.015 \pm 0.102 \pm 0.012$, $A_T^{Re} = -0.064 \pm 0.077 \pm 0.015$ and $F_\text{L} = 0.044 \pm 0.026 \pm 0.014$. These measurements provide the current world's best constraint on right handed contributions to the photon polarization in $b \to s\gamma$ transitions.

The LHCb detector is undertaking a major upgrade in 2019/2020 to be able to run the experiment at an instantaneous luminosity of 2×10^{33} $\text{cm}^{-2}\text{s}^{-1}$. In particular, the electronics of the calorimeters has to be upgraded with new Front-End boards (FEB) to cope with the 40 MHz readout. This thesis also presents the upgrade of the Low Level Trigger module of the FEB, which is in particular responsible for identifying the maximum transverse energy cluster of each event.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France