Search for $B_{(c)}^{+} \to \mu^{+}\nu_{\mu}\gamma$ decays with the LHCb experiment and upgrade of the electromagnetic calorimeter

Recherche des désintégrations $B_{(c)}^{+} \to \mu^{+}\nu_{\mu}\gamma$ avec l'expérience LHCb et la mise à niveau du calorimètre électromagnétique

Thèse de doctorat de l'université Paris-Saclay et de l'université de Heidelberg

École doctorale n°576, Particules, Hadrons, Énergie, Noyau, Instrumentation, Imagerie, Cosmos et Simulation (PHENIICS) Spécialité de doctorat: Physique des particules
Graduate School : Physique,
Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche **IJCLab** (Université Paris-Saclay, CNRS), sous la direction de **Patrick ROBBE**, Directeur de recherche Univerité Paris-Saclay,
et la co-direction **Stephanie HANSMANN-MENZEMER**, Professeure Universität Heidelberg

Thèse soutenue à Paris-Saclay, le 17 Decembre 2025, par

Fabian Christoph GLASER

Composition du jury

Membres du jury avec voix délibérative

Achille STOCCHI Directeur de recherche, IJCLab / Université Paris-Saclay	Président
Florian BERNLOCHNER Professeur, Universität Bonn	Rapporteur & Examinateur
Marta CALVI Professeure, Università degli Studi di Milano-Bicocca	Rapportrice & Examinatrice
Karim NOUI Professeur, IJCLab / Université Paris-Saclay	Examinateur
Philipp ROLOFF Chercheur, CERN	Examinateur
Keri VOS Professeure associée, Maastricht University	Examinatrice
Marie-Hélène SCHUNE Directrice de recherche, IJCLab / Université Paris-Saclay	Invitée

THESE DE DOCTORAT

NNT : 2020UPASA001

DISSERTATION

submitted to the
Combined Faculty of Mathematics, Engineering and Natural Sciences
of Heidelberg University, Germany
for the degree of
Doctor of Natural Sciences
as part of a joint doctorate in collaboration with
Université Paris-Saclay, France

Put forward by
M.Sc. Fabian Christoph Glaser
born in Spaichingen, Germany

Oral examination: 17.12.2025

Search for $B_{(c)}^{+} \to \mu^{+}\nu_{\mu}\gamma$ decays with the LHCb experiment and upgrade of the electromagnetic calorimeter

Referees:

Prof. Dr. Stephanie Hansmann-Menzemer

Prof. Dr. Florian Bernlochner

université
PARIS-SACLAY

ÉCOLE DOCTORALE
Particules, hadrons, énergie et noyau :
instrumentation, imagerie, cosmos
et simulation (PHENIICS)

Title: Search for $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays with the LHCb experiment and upgrade of the electromagnetic calorimeter

Keywords: Particle Physics, Calorimeter, LHCb, Radiative decays

Abstract: Knowledge of the internal structure of the B^+ meson is crucial for precise theory predictions of B^+ meson decays. In particular, the first inverse moment of the B^+-meson light-cone distribution amplitude (LCDA) is a key ingredient in QCD factorisation schemes. Inferring the value of the first inverse moment from theory models is very challenging and often limits the precision of theory predictions. Instead, it is proposed to infer its value from experimental data.

The radiative leptonic decay $B^+ \to \mu^+\nu_\mu\gamma$ is considered the golden mode to probe the internal structure of the B^+ meson, as its branching fraction is sensitive to the first inverse moment of the B^+-meson LCDA. Studies of radiative leptonic decays of B_c^+ mesons, on the other hand, allow to infer the dynamics in the bound B_c^+ meson system.

This thesis presents the simultaneous search for $B^+ \to \mu^+\nu_\mu\gamma$ and $B_c^+ \to \mu^+\nu_\mu\gamma$ decays using data recorded with the LHCb experiment in proton-proton collisions at a centre-of-mass energy of 13 TeV from 2016 to 2018, corresponding to an integrated luminosity of $5.4\,\text{fb}^{-1}$. The decays are reconstructed from photons converting to di-electron tracks in the detector material, which provide three tracks and a displaced decay vertex. The main challenge is to identify and suppress background from $\pi^0 \to \gamma\gamma$ decays, for which novel techniques are developed. The study aims to improve the current best limit on the branching fraction for $B^+ \to \mu^+\nu_\mu\gamma$ decays and set the first limit on the branching fraction for $B_c^+ \to \mu^+\nu_\mu\gamma$ decays. The expected limits on the branching fractions for radiative leptonic decays are of the order of 10^{-6} for decays of B^+ mesons and 10^{-4} for the decay of B_c^+ mesons.

With the LHCb experiment aiming to collect data at higher instantaneous luminosities in future runs, the detector is upgraded. This thesis reports on the organisation of the cabling of the electromagnetic calorimeter and presents performance studies of the particle identification system for Run 4 for different upgrade scenarios.

université
PARIS-SACLAY

ÉCOLE DOCTORALE
Particules, hadrons, énergie et noyau :
instrumentation, imagerie, cosmos
et simulation (PHENIICS)

Titre: Recherche des désintégrations $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ avec l'expérience LHCb et la mise à niveau du calorimètre électromagnétique

Mots clés: Physique des particles, Calorimètre, LHCb, Désintégrations radiatives

Résumé: La connaissance de la structure interne du méson B^+ est essentielle pour établir des prédictions théoriques précises sur les désintégrations du méson B^+. En particulier, le moment de premier ordre inverse de l'amplitude de distribution du cône de lumière (LCDA) du méson B^+ est un élément clé dans les schémas de factorisation de la QCD. Déduire la valeur du moment de premier ordre inverse à partir de modèles théoriques est très difficile et limite souvent la précision des prédictions théoriques. Il est donc proposé de déduire sa valeur à partir de données expérimentales.

La désintégration leptonique radiative $B^+ \to \mu^+\nu_\mu\gamma$ est considérée comme le mode idéal pour étudier la structure interne du méson B^+, car son taux de branchement est lié au moment de premier ordre inverse de la LCDA du méson B^+. D'autre part, les études des désintégrations leptoniques radiatives des mésons B_c^+ permettent de déduire la dynamique du système lié des mésons B_c^+.

Cette thèse présente la recherche simultanée des désintégrations $B^+ \to \mu^+\nu_\mu\gamma$ et $B_c^+ \to \mu^+\nu_\mu\gamma$ à partir des données enregistrées par l'expérience LHCb dans des collisions proton-proton à une énergie dans le centre de masse de 13 TeV entre 2016 et 2018, correspondant à une luminosité intégrée de 5,4 fb^{-1}. Les désintégrations sont reconstruites à partir de photons se convertissant en traces de deux électrons dans le matériau du détecteur, qui fournissent trois traces et un vertex de désintégration déplacé. Le principal défi consiste à identifier et à supprimer les événements de fond provenant des désintégrations $\pi^0 \to \gamma\gamma$, pour lesquels de nouvelles techniques ont été développées. L'étude vise à améliorer la meilleure limite actuelle sur la taux de branchement pour les désintégrations $B^+ \to \mu^+\nu_\mu\gamma$ et à établir la première limite sur la taux de branchement pour les désintégrations $B_c^+ \to \mu^+\nu_\mu\gamma$. Les limites attendues sur les taux de branchement pour les désintégrations leptoniques radiatives sont de l'ordre de 10^{-6} pour les désintégrations des mésons B^+ et de 10^{-4} pour la désintégration des mésons B_c^+. L'expérience LHCb visant à collecter des données à des luminosités instantanées plus élevées lors des prochains runs, le détecteur est en cours de mise à niveau. Cette thèse rend compte de l'organisation du câblage du calorimètre électromagnétique et présente des études de performance du système d'identification des particules pour le run 4 pour différents scénarios de mise à niveau.

université
PARIS-SACLAY

ÉCOLE DOCTORALE
Particules, hadrons, énergie et noyau :
instrumentation, imagerie, cosmos
et simulation (PHENIICS)

Titel: Suche nach $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ Zerfällen mit dem LHCb Experiment und Weiterentkicklung des elektromagnetischen Kalorimeters

Schlagwörter: Teilchenphysik, LHCb, Kalorimeter, Strahlende Teilchenzerfälle

Kurzfassung: Die Kenntnis der inneren Struktur des B^+-Mesons ist für präzise theoretische Vorhersagen der Zerfälle von B^+-Mesonen von entscheidender Bedeutung. Insbesondere das inverse Moment erster Ordnung der „light-cone distribution amplitude" (LCDA) des B^+-Mesons, λ_B, ist ein wichtiger Bestandteil von QCD-Faktorisierungsschemata. Den Wert des inversen Moments erster Ordnung aus theoretischen Modellen abzuleiten, ist schwierig und limitiert häufig die Genauigkeit theoretischer Vorhersagen. Alternativ kann der Wert auch aus experimentellen Daten abgeleitet werden.

Der leptonische Zerfall unter Aussendung eines Photons, $B^+ \to \mu^+\nu_\mu\gamma$, ist gut geeignet, um die innere Struktur des B^+-Mesons zu untersuchen, da sein Verzweigungsverhältnis direkt von λ_B abhängt. Die Studien zum analogen Zerfall von B_c^+-Mesonen ermöglichen vergleichbare Rückschlüsse auf die Dynamik im gebundenen B_c^+-Meson.

Diese Arbeit präsentiert die simultane Suche nach $B^+ \to \mu^+\nu_\mu\gamma$ und $B_c^+ \to \mu^+\nu_\mu\gamma$-Zerfällen in Daten, die mit dem LHCb-Experiment in Proton-Proton-Kollisionen bei einer Schwerpunktsenergie von 13 TeV von 2016 bis 2018 aufgezeichnet wurden, was einer integrierten Luminosität von 5,4 fb^{-1} entspricht. Für die Rekonstruktion des Zerfalls werden Photonen genutzt, die im Detektormaterial in ein Elektron-Positron-Paar umgewandelt werden. Daraus lässt sich der vom Primärvertex separierte Zerfallsvertex aus drei Spuren rekonstruieren. Die größte Herausforderung besteht darin, Hintergrundereignisse aus $\pi^0 \to \gamma\gamma$-Zerfällen zu identifizieren und zu reduzieren, wofür neue Techniken entwickelt wurden. Das Ziel der vorgelegten Arbeit ist, die derzeitige obere Schranke für das Verzweigungsverhältnis für $B^+ \to \mu^+\nu_\mu\gamma$-Zerfälle zu verbessern und die erste obere Schranke für das Verzweigungsverhältnis für $B_c^+ \to \mu^+\nu_\mu\gamma$-Zerfälle zu bestimmen. Die erwarteten Sensitivitäten für die Verzweigungsverhältnisse der untersuchten Zerfälle liegen in der Größenordnung von 10^{-6} für Zerfälle von B^+-Mesonen und 10^{-4} für Zerfälle von B_c^+-Mesonen.

Da das LHCb-Experiment in Zukunft Daten bei höheren instantanen Luminositäten aufzeichnen soll, wird der Detektor derzeit umgebaut. In dieser Arbeit werden die Planung zur Verkabelung des elektromagnetischen Kalorimeters und Studien zur Leistungsfähigkeit des Systems zur Teilchenidentifikation in Run 4 für verschiedene Szenarien vorgestellt.

Résumé

Le Modèle Standard (MS) de la physique des particules représente actuellement le meilleur modèle des particules fondamentales et de leurs interactions. Il fournit une description remarquablement précise des interactions fortes, faibles et électromagnétiques des particules fondamentales et a été largement testé. Cependant, il est connu pour être une théorie incomplète, car il ne fournit aucune explication à l'asymétrie matière-antimatière observée dans l'univers, à la présence de matière noire et aux masses non nulles des neutrinos, et il n'inclut pas de théorie quantique de la gravité. Cela suggère que le Modèle Standard doit être étendu afin de fournir une théorie complète de toutes les particules fondamentales et de toutes les interactions.

Les expériences modernes de physique des particules, comme par exemple l'expérience LHCb, sont conçues pour effectuer des mesures de précision et trouver des preuves de l'existence de nouvelles particules ou de nouvelles interactions au-delà du Modèle Standard. Ces extensions du Modèle Standard suggèrent souvent l'existence de particules lourdes supplémentaires, c'est pourquoi l'étude de la physique du b présente un intérêt particulier. Les systèmes comprenant un quark b (quark bottom), par exemple le méson B^+, font partie des états liés les plus lourds et se désintègrent uniquement par l'interaction faible avec une particule d'échange lourde comme médiateur. La présence d'une physique au-delà du Modèle Standard pourrait, par exemple, modifier le taux de branchement, soit par couplage direct à l'une des particules lourdes, soit par des effets d'interférence dans des processus d'ordre supérieur. Ainsi, l'étude de la physique du b constitue un moyen idéal pour tester le Modèle Standard et rechercher une nouvelle physique. Pour effectuer des tests de précision du Modèle Standard, il est essentiel de disposer de prédictions théoriques précises. Dans de nombreux cas, celles-ci sont limitées par la connaissance de l'interaction forte à basse énergie, qui ne peut être calculée de manière perturbative car la constante de couplage forte est de l'ordre de l'unité. Une approche pour surmonter ce problème consiste à utiliser la méthode de factorisation de la QCD afin de séparer le calcul des contributions à haute et à basse énergie. Un élément essentiel pour appliquer la factorisation aux désintégrations des mésons B^+ est la connaissance de l'amplitude de distribution du cône de lumière (LCDA) du méson B, et en particulier de son moment de premier ordre inverse, λ_B. Déduire ce paramètre à partir de calculs théoriques est très difficile et les incertitudes actuelles restent importantes. Il est donc proposé d'extraire λ_B à partir de données expérimentales via la désintégration $B^+ \to \ell^+\nu_\ell\gamma$. Cette désintégration n'a jamais été observée, la meilleure limite actuelle sur la taux de branchement rapportée par la collaboration Belle étant de $\mathcal{B}(B^+ \to \ell^+\nu_\ell\gamma) < 3.0 \times 10^{-6}$ avec un niveau de confiance de 90% [1].

Dans cette thèse, nous présentons la première recherche des désintégrations $B^+ \to \mu^+\nu_\mu\gamma$ à LHCb, en analysant les données collectées lors de collisions pp entre 2016 et 2018 correspondant à une luminosité intégrée de $5.4\,\text{fb}^{-1}$. Une recherche simultanée des désintégrations $B_c^+ \to \mu^+\nu_\mu\gamma$ est effectuée, ce qui marque la toute première recherche de ce type de désintégration. Les désintégrations sont reconstruites à partir des conversions de photons dans le matérial du détecteur afin de permettre la reconstruction du vertex de désintégration déplacé du B. À partir de la position du vertex, la direction de vol du B est

déduite, ce qui permet de récupérer partiellement l'impulsion du neutrino. En corrigeant le déséquilibre dans la désintégration d'impulsion perpendiculaire à la direction de vol du B, une masse corrigée est définie, qui est utilisée comme variable de discrimination entre les candidats signal et bruit de fond.

Le nombre de candidats signaux est extrait d'un ajustement par maximum de vraisemblance étendue à la distribution de masse corrigée dans les données. Dans cet ajustement, la source dominante de bruit de fond est identifiée comme provenant des désintégrations $\pi^0 \to \gamma\gamma$. Une nouvelle méthode a été développée pour modéliser ce type de bruit de fond, en utilisant une approche basée sur les données et complétée par des données supplémentaires issues de simulations, qui constituent le cœur de cette analyse. L'ajustement final de la distribution de masse corrigée dans les données reste confidentiel, car l'analyse fait actuellement l'objet d'un examen interne.

Le taux de branchement des désintégrations $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ est obtenu par rapport au canal de normalisation $B^0 \to K^{*0}\gamma$, qui est également reconstruit à partir des conversions de photons afin d'annuler de nombreuses incertitudes systématiques. Les incertitudes systématiques restantes, qui ne s'annulent pas dans le rapport entre le signal et le canal de normalisation, sont évaluées en détail afin d'obtenir le résultat suivant (la valeur centrale restant cachée)

$$\mathcal{B}(B^+ \to \mu^+\nu_\mu\gamma)_{E^*_\gamma > 1\,\text{GeV}} = (x.x \pm {}^{+1.5}_{-1.3}\,(\text{stat.}) \pm 1.0\,(\text{syst.})) \times 10^{-6},$$
$$\mathcal{B}(B_c^+ \to \mu^+\nu_\mu\gamma)_{E^*_\gamma > 1\,\text{GeV}} = (x.x \pm {}^{+2.9}_{-2.2}\,(\text{stat.}) \pm 1.4\,(\text{syst.})) \times 10^{-4}.$$

L'incertitude systématique est évaluée en supposant un taux de branchement du signal de 3.0×10^{-6} pour la désintégration $B^+ \to \mu^+\nu_\mu\gamma$, correspondant à la meilleure limite supérieure actuelle de Belle [1], et un taux de branchement de 8×10^{-5} pour la désintégration $B_c^+ \to \mu^+\nu_\mu\gamma$ selon la prédiction du modèle standard de la référence [2]. Dans le cas où aucun signal significatif n'est observé, les limites supérieures des taux de branchement sont fournies. En l'absence de signal, les limites supérieures attendues à un intervalle de confiance de 90% (95%) sont

$$\mathcal{B}_{exp}(B^+ \to \mu^+\nu_\mu\gamma)_{E^*_\gamma > 1\,\text{GeV}} < 2.3\,(2.7) \times 10^{-6}$$
$$\mathcal{B}_{exp}(B_c^+ \to \mu^+\nu_\mu\gamma)_{E^*_\gamma > 1\,\text{GeV}} < 6.3\,(7.8) \times 10^{-4}.$$

La deuxième analyse présentée dans cette thèse est la toute première recherche des désintégrations $B_c^+ \to J/\psi D^{*+}$. Cette désintégration est considérée comme le principal canal de normalisation pour les futures études des transitions $b \to d\ell\ell$ dans les désintégrations $B_c^+ \to D^{*+}\ell^+\ell^-$. Ces transitions $b \to d\ell\ell$ constituent un courant neutre à changement de saveur (FCNC), qui est interdit au niveau de l'arbre dans le Modèle Standard et ne peut être médié que par des processus d'ordre supérieur, ce qui se traduit par de faibles taux de branchement. La rareté des désintégrations FCNC en fait un outil idéal pour étudier la physique au-delà du Modèle Standard. Des particules ou interactions supplémentaires étendant le MS pourraient permettre aux FCNC d'être médiées au niveau de l'arbre ou conduire à des diagrammes supplémentaires au niveau de la boucle, ce qui modifierait les taux de branchement de ces désintégrations rares.

La recherche de la désintégration au niveau de l'arbre $B_c^+ \to J/\psi D^{*+}$ est effectuée sur les

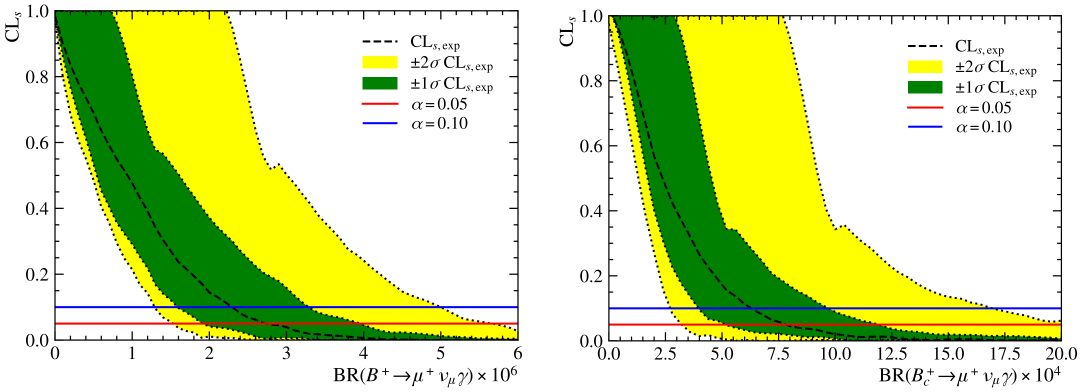


Figure 1: Contour des niveaux de confiance attendus pour la taux de branchement du signal $B^+ \to \mu^+\nu_\mu\gamma$ (à gauche) et $B_c^+ \to \mu^+\nu_\mu\gamma$ (à droite) dérivés de la méthode CL_s en l'absence de signal. Les bandes vertes et jaunes reflètent les bandes d'erreur sur le niveau de confiance attendu tracé en noir.

données enregistrées avec l'expérience LHCb lors des runs 1 et 2, correspondant à une luminosité intégrée de $9\,\text{fb}^{-1}$. Les désintégrations sont reconstruites à partir de $D^{*+} \to D^0\pi^+$ avec des désintégrations subséquentes $D^0 \to K^-\pi^+$ et $D^0 \to K^-\pi^+\pi^+\pi^-$ et à partir de $J/\psi \to \mu^+\mu^-$. Les candidats sont sélectionnés à l'aide d'une combinaison de critères de sélection unidimensionnels et d'une analyse multivariée, optimisée pour maximiser la significance attendue de l'observation des désintégrations $B_c^+ \to J/\psi D^{*+}$.
Le taux de branchement est mesuré par rapport à la désintégration de normalisation $B_c^+ \to J/\psi D_s^+$, reconstruite à partir de $J/\psi \to \mu^+\mu^-$ et $D_s^+ \to K^+K^-\pi^+$. À partir d'un ajustement maximal de vraisemblance étendue à la distribution de masse invariante reconstruite dans les données, un nombre total de 382 ± 20 candidats de normalisation est observé.
Le taux de branchement des désintégrations $B_c^+ \to J/\psi D^{*+}$ est extraite d'un ajustement simultané à la masse invariante des candidats reconstruits comme $D^0 \to K^-\pi^+$ et $D^0 \to K^-\pi^+\pi^+\pi^-$. L'analyse étant actuellement en cours de préparation pour un examen interne, la valeur centrale du taux de branchement reste masquée et le résultat est rapporté comme suit

$$\frac{\mathcal{B}(B_c^+ \to J/\psi D^{*+})}{\mathcal{B}(B_c^+ \to J/\psi D_s^+)} = x.xxx \pm 0.026\,(\text{stat}) \pm 0.007\,(\text{syst}), \tag{1}$$

où la première erreur est statistique et la seconde systématique. Avec une valeur attendue de $\mathcal{B}(B_c^+ \to J/\psi D^{*+})/\mathcal{B}(B_c^+ \to J/\psi D_s^+) = 0.098 \pm 0.011$, dérivée des arguments des principes fondamentaux du Modèle Standard, l'analyse présentée devrait permettre de trouver des preuves de la désintégration $B_c^+ \to J/\psi D^{*+}$ avec une significance supérieure à trois écart-types. Grâce à l'augmentation des statistiques enregistrées avec l'expérience LHCb lors du run 3, une première observation est tout à fait envisageable.

Dans la recherche des désintégrations $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$, la reconstruction des photons à partir des clusters neutres est d'une importance capitale pour la méthode de modélisation des processus de fond $\pi^0 \to \gamma\gamma$. Afin de garantir que l'expérience LHCb puisse exploiter pleinement son potentiel à l'avenir, le calorimètre électromagnétique (ECAL) sera amélioré avec une granularité plus fine.

Dans le cadre de ce travail, la cartographie des câbles de signal reliant les détecteurs du ECAL aux cartes frontales, ainsi que la répartition des cartes frontales dans les racks électroniques, ont été définies.

En outre, des études sur les performances du système d'identification des particules (PID) pour le run 4 de l'expérience LHCb ont été menées. Pour les futures mises à niveau du détecteur LHCb, il est prévu de remplacer le calorimètre hadronique (HCAL) par un écran absorbant afin d'améliorer les performances du système muonique. Une autre alternative consiste à conserver le HCAL en place et à l'utiliser comme écran. Pour décider de la marche à suivre, il est impératif de comprendre l'impact sur les performances du détecteur. L'étude présentée dans cette thèse se concentre sur le cas où le HCAL n'est pas remplacé et où la décision d'utiliser ou non le HCAL et d'en lire les données lors du run 4 doit être prise.

Dans ce contexte, l'identification des électrons et des muons, ainsi que la mauvaise identification respective des pions chargés et des kaons chargés, sont étudiées avec et sans les informations fournies par le calorimètre hadronique (HCAL). Les études de performance PID sont menées sur des données enregistrées avec l'expérience LHCb en 2024, dans des conditions similaires à celles prévues pour le run 4 d'exploitation. L'identification des leptons chargés est étudiée à l'aide d'une méthode de balisage et de sondage sur des candidats $B^+ \to K^+ J/\psi$, reconstruits à partir de $J/\psi \to \mu^+\mu^-$ et $J/\psi \to e^+e^-$. La probabilité que des pions et des kaons chargés soient identifiés à tort comme des leptons est évaluée sur des candidats $D^0 \to K^-\pi^+$ provenant de désintégrations $D^{*+} \to D^0\pi^+$.

Pour l'identification des muons, aucun changement significatif n'est constaté dans les performances. Pour les électrons, une légère dégradation des performances PID est observée pour les électrons énergétiques sans photon de freinage associé et avec des exigences PID strictes. L'effet global sur l'efficacité de la sélection est étudié à l'aide des désintégrations $B^+ \to K^+ J/\psi(\to e^+e^-)$ comme référence. En supposant un niveau identique de bruit de fond dû à des erreurs d'identification, la perte d'efficacité de sélection est d'environ 1%. Avec les mises à niveau prévues du système RICH et du ECAL, les performances globales du PID lors du run 4 d'exploitation devraient dépasser celles du run 3, même sans utiliser les informations PID provenant du HCAL.

Acknowledgements

I want to express my gratitude to my supervisors, Stephanie Hansmann-Menzemer and Patrick Robbe, for their guidance and support over the past three years and beyond. I am very grateful to you for initiating this PhD project and for the opportunity to do it jointly between Heidelberg and Orsay.

I want to thank Florian Bernlochner and Matra Calvi for agreeing to review my thesis and for the time and effort they invested. On the same note, I would like to thank Achille Stocchi, Karim Noui, Philipp Roloff and Keri Vos for being part of the thesis jury and accompanying me in taking the last hurdle.

Another big thank you goes to Marie-Hélène Schune for all your help and support during this project, answering all sorts of questions on physics, but in particular regarding all organisational matters. My sincere thanks to Martino Borsato for playing a central role in everything leading up to this thesis and for all your help and support, which stretches back to the beginning of my master's thesis.

Furthermore, I want to extend my gratitude to all current and former members of the LHCb groups in Heidelberg and at IJCLab. A special thanks goes to Gaëlle, Marie and Piera for the great company over the last two years and the fantastic spirit in the office.

Ein weiterer Dank gilt meinen Freunden, insbesondere Basti und Hannah für die langen Snookerabende, die vielen Runden Siedler und das gelegentlich gewährte Asyl auf eurer Couch; den besten Mitbewohnerinnen, Jana und Marie, für die schöne Zeit in Heidelberg; und Jakob für die viele Fahrradfachsimpelei.

Der letzte und wichtigste Dank gilt meiner Familie, insbesondere meiner Mutter, Elisabeth, und meinen beiden "großen" Schwestern, Daniela und Katja, die mich immer unterstützt haben meinen Weg zu gehen; meinen Nichten und Neffen Lauri, Finn, Santi und Emmi die mich bei Seifenblasen und LEGO alles Andere vergessen lassen; und meiner Großmutter, die diese Arbeit nicht mehr miterleben sollte, aber mit ihrem Satz "du gäisch a mol en Doktr" recht behalten sollte.

Contents

Introduction

The idea that the universe is composed of fundamental building blocks dates back nearly 2500 years to ancient Greece. Around 400 BC, Greek philosopher Democritus proposed that indivisible particles called *atoms* (from Greek *átomos*: indivisible) are the smallest units of nature, existing in infinite varieties. In the early 19th century, the term atom was adopted to refer to the smallest unit of chemical elements, which nowadays is known to be a composite object and, contrary to the etymology, not a fundamental building block.

As of today, 17 distinct fundamental particles[1] are known, contrasting Democritus' idea of infinitely many fundamental types. The particles and their interactions are described by the Standard Model (SM) of particle physics, which marks the current best knowledge of modern particle physics. It provides a remarkably accurate description of the fundamental interactions of the strong and electroweak force and has been tested extensively at a multitude of particle physics experiments. However, it is known to be an incomplete theory, as it lacks an explanation for the observed matter-antimatter asymmetry in the universe, the presence of dark matter, and non-zero neutrino masses, and it does not include a quantum theory of gravity. This suggests the Standard Model needs to be extended to provide a complete theory of all fundamental particles and interactions.
Modern particle physics experiments, like for example the LHCb experiment, are designed to perform precision measurements and find evidence for new particles or new interactions beyond the Standard Model. Such extensions of the Standard Model often suggest additional heavy particles, which is why the study of b-physics is of particular interest. Systems including a b-quark (bottom quark), for example the B^+ meson, are among the heaviest bound states and only decay through the weak interaction mediated by a heavy exchange particle. The presence of physics beyond the Standard Model might, for example, modify the decay rate, either by directly coupling to one of the heavy particles or through interference effects in higher-order processes. Thus, the study of b-physics provides a sensitive probe to test the Standard Model and search for New Physics.
To perform precision tests of the Standard Model, it is crucial to have precise theory predictions. In many cases, these are limited by the knowledge of the strong interaction at low energy, which cannot be calculated perturbatively as the strong coupling constant is of the order of unity. To overcome this issue, many different approaches have been proposed, one of which is the method of factorisation to separate the calculation of high- and low-energy contributions. One crucial ingredient for applying the factorisation approach to the decays of B^+ mesons is the first inverse moment of the B meson light-cone distribution amplitude, λ_B, which encodes the internal dynamics of the bound B^+ meson state. Inferring this parameter from theoretical calculations is very challenging and the current uncertainties remain large. Instead, it is suggested to extract λ_B from experimental data through the decay $B^+ \to \ell^+\nu_\ell\gamma$. The decay has never been observed, with the current best limit on the branching fraction reported by the Belle collaboration, with a limit of $\mathcal{B}(B^+ \to \ell^+\nu_\ell\gamma) < 3.0 \times 10^{-6}$ at 90% confidence level [1].

---

[1] Considering also the respective anti-particles and assuming neutrinos to be Dirac particles, a total of 29 particles are known.

In this thesis, the first search for $B^+ \to \mu^+\nu_\mu\gamma$ decays at LHCb is reported, analysing data collected in *pp* collisions from 2016 to 2018. A simultaneous search for $B_c^+ \to \mu^+\nu_\mu\gamma$ decays is performed, which has not been conducted previously. The aim is to observe both decays for the first time or improve the current best limit on the B^+ decay and set the first limit on the B_c^+ decay branching fraction. The analysis is entirely in place and currently under internal review.

Beyond the search for radiative leptonic decays, a separate analysis is presented, marking the first-ever search for $B_c^+ \to J/\psi D^{*+}$ decays using the full Run 1 and Run 2 datasets recorded with the LHCb experiment. The aim is to observe the decay for the first time or provide an upper limit on its branching fraction. This analysis is currently in preparation for internal review.

The thesis is structured as follows. In Chapter 1, an introduction to the Standard Model of particle physics and the QCD factorisation approach for $B^+ \to \ell^+\nu_\ell\gamma$ decays is presented. In Chapter 2, the experimental apparatus of the LHCb experiment is described. In Chapter 3, the author's contribution to the upgrade of the electromagnetic calorimeter at the LHCb experiment is presented.

In Chapter 4, an overview of the search for radiative leptonic $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$ decays at LHCb is given. The selection of candidates is detailed in Chapter 5. In Chapter 6, the fit to the corrected mass distribution is established to extract the number of observed signal candidates. At the core of this lies a novel approach to model background from $\pi^0 \to \gamma\gamma$ decays through data-driven methods. To reduce systematic uncertainties, the analysis is performed relative to the normalisation channel $B^0 \to K^{*0}\gamma$, which is analysed in Chapter 7. The statistical sensitivity is derived in Chapter 8 and the systematic uncertainties are discussed in Chapter 9. The blinded results and expected upper limits on the branching fractions are presented in Chapter 10.

The search for $B_c^+ \to J/\psi D^{*+}$ decays is introduced in Chapter 11. Details on the selection of candidates are given in Chapter 12 and the fit to the reconstructed invariant mass is explained in Chapter 13. The blinded results are presented in Chapter 14.

Part I

Theoretical foundation and the LHCb experiment

1

Theoretical foundation

This chapter provides a brief overview of the Standard Model of particle physics in Section 1.1 and introduces the concepts of QCD factorisation in Section 1.2. These concepts are further discussed in the context of radiative leptonic decays $B^+ \to \ell^+\nu_\ell\gamma$ and $B_c^+ \to \ell^+\nu_\ell\gamma$ in Sections 1.3 and 1.4, respectively.

1.1 The Standard Model of particle physics

The fundamental particles and their interactions are described in the Standard Model (SM) of particle physics, which is the most accurate description of modern particle physics. It is formulated as a renormalisable quantum field theory in which the fundamental particles manifest themselves as excitations of the underlying fields. Interactions between the particles arise from invariance under local gauge transformations of the symmetry $SU(3)_C \times SU(2)_L \times U(1)_Q$. They can be interpreted as the three fundamental forces: strong, weak and electromagnetic forces. The conjecture established by Noether in 1918 [3] implies the existence of three charges, which are conserved under the Standard Model gauge symmetry. The $SU(3)$ symmetry of the strong force conserves color charge, and the $SU(2)_L \times U(1)_Y$ symmetry of the weak and electromagnetic force conserves the electric charge and weak isospin charge.

1.1.1 Particle content

Based on the charges, or more generally on the *quantum numbers*, the fundamental particles of the Standard Model can be separated into different categories. Particles with half-integer spin[1] fall in the category of matter particles, denoted *fermions*. In contrast, particles with integer spin are denoted *bosons*, like the scalar Higgs boson or the vector bosons which mediate the fundamental forces.
Fermions that carry color charge and interact via the strong interactions are denoted *quarks*. Today, six different quarks are known, three up-type quarks with electric charge[2]

---

[1] The particles are assigned a spin based on the transformation of the underlying quantum field under Lorentz transformations.

[2] Throughout the document natural units are used setting $\hbar = c = 1$.

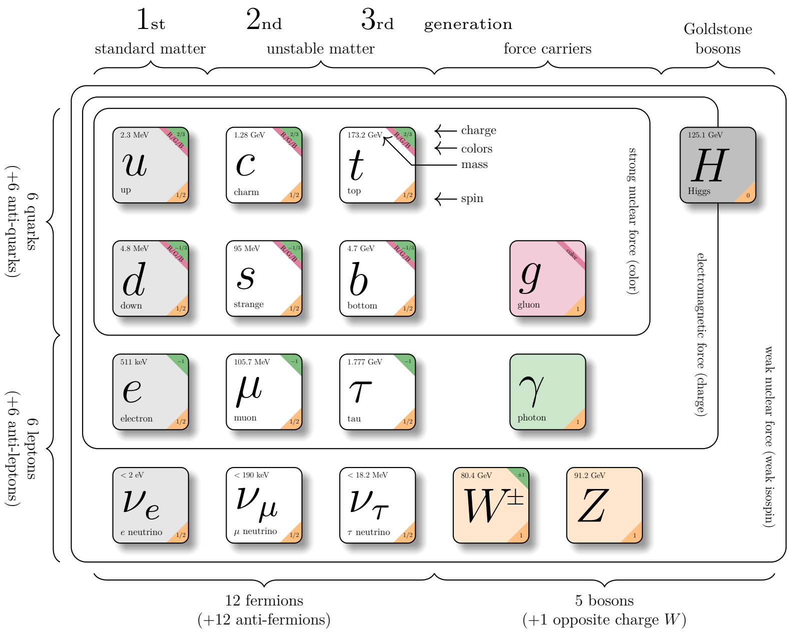


Figure 1.1: Particle content in the Standard Model of particle physics. Modified from Ref. [4] with particle properties from Ref. [5].

+2/3 named *up* (u), *charm* (c) and *top* (t), as well as three down-type quarks with electric charge $-1/3$ named *down* (d), *strange* (s) and *bottom* (b). All three up-type (down-type) quarks share the same quantum numbers and can be separated into three *generations* by their mass. The lightest pair of up- and down-type quarks (up and down) are quarks of the first generation; strange- and charm-quarks fall into the second generation, and the heaviest quarks (bottom and top) fall into the third generation.

Besides the quarks, the Standard Model contains six additional fermions, named *leptons*, which do not carry color charge. The *electron* (e) with an electric charge of -1 is the most prominent lepton in the Standard Model. Similarly to the quarks, also the leptons come in three different *flavours* of identical quantum numbers but different mass, which are named *muon* (μ) and *tauon* (τ). Each of the charged leptons has an associated electrically neutral lepton, named *neutrino*, which are labelled ν_e, ν_μ and ν_τ according to their associated charged lepton.

In addition to the twelve fermions described, the symmetry of the Standard Model ensures that each fermion has an associated antiparticle, which carries identical quantum numbers with an inverted sign, except for the spin. Since neutrinos are electrically neutral, they are the only fermions in the Standard Model that have the potential of being their own antiparticles. The question whether neutrinos exhibit this feature and are *Majorana* particles or are *Dirac* particles with a distinct anti-particle, like all other fermions, is subject to current research.

The gauge bosons of spin one mediate the fundamental interactions of particles. The electroweak interaction is mediated by the massive $W^{\pm}$ and Z^0 bosons as well as the massless photon, γ, while the strong interaction is mediated by massless gluons, g. The Higgs boson with spin zero is the only scalar particle in the Standard Model. It is the self-excitation of the underlying Higgs field which generates the masses of the fundamental particles.

1.1.2 Fundamental interactions

Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the modern theory of the strong interaction. It introduces an additional quantum number, the *color* charge with values $\alpha = 1, 2, 3$ (also referred to as *red*, *green* and *blue*), which is preserved under a local $SU(3)_C$ gauge symmetry acting on the color index. The gauge structure results in eight massless gluon fields, $G^i,\ i = 1...8$, which are the mediating bosons of the strong interaction. The gluons couple to all fields which carry color charge, which are the quarks and the gluons themselves, with a strong coupling constant g_s. The QCD Lagrangian density is

$$\mathcal{L}_{QCD} = \bar{\psi^{\alpha}}(i\gamma^{\mu}\partial_{\mu} - m)\psi_{\alpha} - \frac{g_s}{\sqrt{2}}(\bar{\psi^{\alpha}}\gamma^{\mu}\tau^{i}_{\alpha\beta}\psi^{\beta})G^{i}_{\mu} - \frac{1}{4}G^{i}_{\mu\nu}G^{i\mu\nu} + \frac{\theta_{QCD}}{32\pi^2}g_s G^{i}_{\mu\nu}\widetilde{G}^{i\mu\nu} \tag{1.1}$$

where ψ are the six different quark fields $\psi = u, d, s, c, b, t$ and γ^{μ} are the Dirac matrices. The elements τ^i are the eight generators of the $SU(3)$ Lie group represented by the Gell-Mann matrices and $G^{i}_{\mu\nu}$ is the gluon field strength tensor defined as

$$G^{i}_{\mu\nu} = \partial_{\mu}G^{i}_{\nu} - \partial_{\nu}G^{i}_{\mu} - g_s f_{ijk}G^{j}_{\mu}G^{k}_{\nu} \tag{1.2}$$

where the coefficients f_{ijk} are the (totally antisymmetric) structure constants of $SU(3)$. The first term in Eq. 1.1 is the kinetic term of the quarks and the second term describes the interaction of quark and gluon fields. The third term is the kinetic term of the gluon field itself and, due to the non-abelian structure of $SU(3)$, gives rise to gluon self-interactions. The last term in Eq. 1.1 introduces another dimensionless parameter, θ_{QCD}, and the dual field strength tensor, $\widetilde{G}^{i}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\rho\sigma}G^{i\rho\sigma}$. This represents the only term in QCD that violates CP symmetry[3] in the sector of strong interactions. Experimental limits on the neutron electric dipole moment suggest $|\theta_{QCD}| < 10^{-10}$ [6, 7], which is why for pure QCD typically the case $\theta_{QCD} = 0$ is considered.

Higher-order corrections to the gluon propagator from quark loops and gluon self-interactions can be absorbed in a running coupling constant, $\alpha_s(\mu^2) = g_s(\mu^2)/4\pi$, where μ is the energy scale of the process. This defines the typical energy scale of QCD, which is denoted Λ_{QCD}, and is of the order of 200 MeV. At large energies, $\mu \gg \Lambda_{QCD}$, the coupling constant, $\alpha_s(\mu)^2$, becomes small and quarks and gluons can be treated as quasi-free particles in perturbation theory. This behaviour is known as asymptotic freedom. In contrast to that, the strong coupling constant $\alpha_s(\mu^2) \sim \mathcal{O}(1)$ for energies of the order of $\mu \sim \Lambda_{QCD}$ and QCD becomes non-perturbative. This also implies colored objects, in particular isolated quarks, cannot exist freely but are *confined* in color-singlet bound states, e.g. in mesons ($q\bar{q}$) or baryons (qqq).

The running of the strong coupling constant makes the calculation of matrix elements involving low-energy QCD effects notoriously difficult. One approach to simplify calculations including the non-perturbative QCD regime is *QCD factorisation*, which is discussed in Section 1.2.

---

[3] CP symmetry refers to the combined symmetry under charge conjugation (C) and parity (P).

Electroweak interaction

For a fermion field, ψ, the left- and right-hand chiral projections can be defined as

$$\psi_L = P_L\psi = \frac{1-\gamma^5}{2}\psi, \quad \psi_R = P_R\psi = \frac{1+\gamma^5}{2}\psi, \tag{1.3}$$

with the chiral projection operators $P_{L,R}$ satisfying the relations $P_L + P_R = 1$ and $P^2_{L,R} = P_{L,R}$. The left- and right-hand chiral projections of the fermion field can be seen as independent degrees of freedom with $\psi = \psi_L + \psi_R$ and, in the limit of massless fermions, coincide with the negative and positive helicity states.
In the standard electroweak theory, the Lagrangian is invariant under local gauge transformations of the symmetry group $SU(2)_L \times U(1)_Y$. The three generators of $SU(2)$ give rise to three massless gauge fields, W^i, $i = 1, 2, 3$, which couple to the *weak isospin* component I_3 of left-handed chiral fermion fields with a gauge coupling g. In the Standard Model, the up-type quarks, $u_i = u, d, s$, as well as the neutrinos, ν_i, have weak isospin component $I_3 = +1/2$, while down-type quarks, $d_i = d, s, b$, and the charged leptons, $e_i = e, \mu, \tau$, have weak isospin component $I_3 = -1/2$. The left-handed chiral projections in each family of fermions are grouped into $SU(2)$ doublets, while the right-handed projections[4] are singlets with respect to $SU(2)$

$$q_{i,L} = \begin{pmatrix} u_i \\ d_i \end{pmatrix}_L, \quad \ell_{i,L} = \begin{pmatrix} \nu_i \\ e_i \end{pmatrix}_L, \quad (u_i)_R, \quad (d_i)_R, \quad (e_i)_R. \tag{1.4}$$

The additional $U(1)_Y$ symmetry acts on the *weak hypercharge* Y, which is defined as

$$Y = Q - I_3, \tag{1.5}$$

where Q is the electric charge. The symmetry under $U(1)_Y$ rotations gives rise to another gauge field, B, and gauge coupling g'. The electroweak Lagrangian then reads

$$\mathcal{L}_{EW} = i\bar{\psi}_{j,L}\gamma^\mu D_{\mu,L}\psi_{j,L} + i\bar{\psi}_{j,R}\gamma^\mu D_{\mu,R}\psi_{j,R} - \frac{1}{4}W^i_{\mu\nu}W^{\mu\nu i} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \tag{1.6}$$

with the field strength tensors for $SU(2)$ and $U(1)$, respectively

$$\begin{aligned} W^i_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\epsilon_{ijk}W^j_\mu W^k_\nu, \quad i, j, k = 1...3 \\ B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu. \end{aligned} \tag{1.7}$$

The covariant derivatives for the left- and right-handed chiral fermion currents are

$$\begin{aligned} D_{\mu,L} &= \partial_\mu - ig'\frac{Y}{2}B_\mu - ig\tau_k W^k_\mu, \\ D_{\mu,R} &= \partial_\mu - ig'\frac{Y}{2}B_\mu. \end{aligned} \tag{1.8}$$

In this simple model of electroweak interactions, the four gauge fields W^i and B remain massless as any additional mass term in the Lagrangian breaks gauge invariance. However, the physical bosons $W^\pm$ and Z, which are associated with the electroweak interaction, are found to be massive. Thus, the electroweak sector is extended by the introduction of an additional scalar field, ϕ.

---

[4] There are no right-handed chiral projections of neutrinos in the Standard Model.

Higgs mechanism

The Higgs mechanism postulated by Brout, Englert and Higgs [8,9] introduces a complex scalar field

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{1.9}$$

with hypercharge $Y = 1$ and a Lagrangian of the form

$$\mathcal{L}_\phi = (D^\mu \phi)^\dagger D_\mu \phi - V(\phi), \tag{1.10}$$

where $V(\phi)$ is a potential of the form

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \tag{1.11}$$

For $\mu^2 < 0$, the Higgs field acquires a non-vanishing vacuum expectation value (VEV) $v = \langle 0|\phi|0\rangle \neq 0$ and the symmetry of $SU(2)_L \times U(1)_Y$ is spontaneously broken. Without loss of generality, the Higgs field can be rotated in the four-dimensional space, such that the VEV reads

$$\phi \to v = \langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix}. \tag{1.12}$$

The vacuum of the Higgs field v carries no electric charge, resulting in the breaking of $SU(2)_L \times U(1)_Y \to U(1)_Q$, where the $U(1)_Q$ symmetry of electromagnetism is not broken. Around this minimum, the field ϕ can be expanded and quantised, which leads to the physics Higgs boson H in

$$\phi \to \phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + H \end{pmatrix}. \tag{1.13}$$

In this minimum, the covariant kinetic energy term reads

$$(D_\mu \phi)^\dagger (D^\mu \phi) = \frac{1}{2} (0\ \nu) \left(\frac{g}{2} \tau^i W_\mu^i + \frac{g'}{2} B_\mu \right)^2 \begin{pmatrix} 0 \\ \nu \end{pmatrix} + H \, \text{terms}, \tag{1.14}$$

from which the physical gauge bosons emerge as linear combinations of the gauge fields

$$\begin{aligned} W_\mu^\pm &= \frac{1}{2}(W_\mu^1 \mp W_\mu^2) \\ Z_\mu &= \frac{1}{\sqrt{g^2 + g'^2}} (g W_\mu^3 - g' B_\mu) \\ A_\mu &= \frac{1}{\sqrt{g^2 + g'^2}} (g W_\mu^3 + g' B_\mu) \end{aligned} \tag{1.15}$$

with masses $M_W = \frac{g}{2} v$, $M_Z = \frac{\sqrt{g^2+g'^2}}{2} v$ and $M_A = 0$. The remaining degree of freedom in the minimum of the Higgs field manifests itself as the massive Higgs boson with mass $m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda} v$.

Yukawa interaction

Following the spontaneous breaking of the electroweak symmetry, the masses of the gauge bosons $W^{\pm}$ and Z are generated, while the fermions remain massless. To generate the fermion masses, a Yukawa coupling between the Higgs doublet and the fermion fields is introduced as

$$\mathcal{L}_{Yuk} = y^{u}_{i,j}\bar{q}_{i,L}\widetilde{\phi}u_{j,R} + y^{d}_{i,j}\bar{q}_{i,L}\phi d_{j,R} + y^{\ell}_{i,j}\ell_{i,L}\phi e_{j,R} + h.c., \tag{1.16}$$

where $\widetilde{\phi} = i\tau^2\phi$ is the Higgs conjugate. Note that there are no right-handed neutrinos in the Standard Model and neutrinos remain massless fermions. The Yukawa coupling matrices y^f can be diagonalized by unitary transformations $A^{f\dagger}_{L} y^f A^{f}_{R} = y^{f}_{D}$. The eigenvalues of this matrix determine the masses of the quarks and leptons $m_f = \frac{v}{\sqrt{2}} y^{f}_{D,ii}$.

CKM mechanism

Using the unitary transformation matrices $A^{f}_{L,R}$, the physical mass eigenstates of the quarks can be defined as

$$u_{L,mass} = A^{u}_{L} u_{L,weak}, \quad d_{L,mass} = A^{d}_{L} d_{L,weak}. \tag{1.17}$$

It is important to note that the transformation matrices $A^{u}_{L,R}$ and $A^{d}_{L,R}$ are different for up- and down-type quarks, and both fermion fields cannot be diagonalised simultaneously. For charged current interactions of the form

$$\mathcal{L}_{CC} = -\frac{g}{\sqrt{2}}\left[\bar{d}_L\gamma^{\mu}A^{d,\dagger}_{L}A^{u}_{L}u_L W^{-}_{\mu} + \bar{u}_L\gamma^{\mu}A^{u,\dagger}_{L}A^{d}_{L}d_L W^{+}_{\mu}\right], \tag{1.18}$$

this introduces a fundamental difference between the weak eigenstates, which couple to the gauge fields and the mass eigenstates, which propagate as physical particles. The unitary transformation matrices can be absorbed into a single unitary matrix, the *Cabibbo-Kobayashi-Maskawa* (CKM) matrix

$$V_{CKM} = A^{u,\dagger}_{L}A^{d}_{L} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{1.19}$$

In general, any unitary $N \times N$ matrix has N^2 real parameters. However, not all of these parameters manifest physical observables as the quark fields have an unobservable phase. This allows to redefine the fields and remove $2N-1$ phase differences from the CKM matrix. The resulting $N^2-(2N-1)$ real parameters are $N(N-1)/2$ angles and $(N-1)(N-2)/2$ phases, which for the Standard Model with $N=3$ leaves three angles and one phase. This phase allows for violation of CP symmetry in weak charged current interactions. It is convenient to expand the CKM matrix in the Wolfenstein parametrization [10]

$$V_{CKM} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho-i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{1.20}$$

where it becomes apparent the Standard Model exhibits a strong hierarchical structure with $\lambda \approx 0.225$ [5].

1.1.3 Physics beyond the Standard Model

The Standard Model offers a remarkably powerful framework for describing the fundamental particles and their interactions with unprecedented accuracy. Despite its success, there is also striking experimental evidence of physics phenomena, which cannot be explained in the Standard Model.
Firstly, experimental observations of neutrino oscillations by the Sudbury Neutrino Observatories [11, 12] and Super-Kamiokande Observatory [13] suggest a non-trivial mixing of neutrino flavours, similar to the CKM mechanism. However, this requires the neutrino masses to be non-zero and non-degenerate as opposed to the massless neutrinos incorporated in the Standard Model.
Secondly, astronomical and cosmological observations indicate that the amount of ordinary baryonic matter in stars and galaxies only accounts for about 5% of the energy in the observable universe [14, 15]. About 25% of the energy is attributed to *Dark Matter*, which is clearly observable from its gravitational influence but does not couple to the electromagnetic interaction. The remaining 70% are attributed to *Dark Energy*, whose negative pressure allows to explain the accelerated expansion of the universe. Neither Dark Matter nor Dark Energy can be explained in the framework of the Standard Model. In addition, the Standard Model does currently not include a quantum theory of gravity.
Furthermore, the Standard Model falls short in explaining the magnitude of matter-antimatter asymmetry observed in the universe. One necessary condition to break matter-antimatter symmetry is the existence of a CP-violating process [16] at a time shortly after the Big Bang. The CKM mechanism in the Standard Model allows for such CP-violating processes; however, the degree to which CP is violated in the Standard Model is several orders of magnitude too small to explain the observed baryon asymmetry [17].
These shortcomings of the Standard Model suggest the Standard Model is the low-energy limit of a more general theory of particle physics that is yet to be discovered. Modern experiments aim at finding evidence of such physics beyond the Standard Model. A typical starting point is to probe properties of the Standard Model that do not arise from the principles of gauge invariance. For example, there is no fundamental motivation for why the Standard Model has exactly three generations of fermions, which exhibit a strict mass hierarchy, or why QCD is CP-conserving with vanishing θ_{QCD}.

1.2 QCD factorisation in non-leptonic B decays

As outlined in the previous section, the strong coupling constant, $\alpha_s(\mu)$, exhibits a running as a function of the energy scale μ and is of order unity for energies $\mu \sim \Lambda_{QCD}$. This makes the calculation of QCD effects notoriously difficult, as low-energy processes cannot be treated perturbatively.
In the framework of QCD factorisation, the contributions from low- and high-energy processes are *factorised* to approximate an exact QCD matrix element. The high-energy contributions are calculated in perturbation theory as an expansion in $\alpha_s(\mu)$, while low-energy effects are absorbed in an effective interaction, which is calculated non-perturbatively (e.g. from Lattice QCD) or derived from experimental data. This often results in a simpler structure than the exact matrix element or in a process-independent quantity. In the following, the concepts of QCD factorisation are introduced for non-leptonic B meson decays before being applied to the radiative leptonic decay $B^+ \to \ell^+ \nu_\ell \gamma$ in Section 1.3.

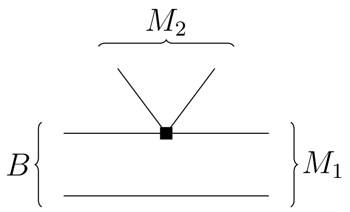


Figure 1.2: Leading order contribution in the naive factorisation approach of a $B \to M_1 M_2$ decay. The black square corresponds to an effective weak decay through a four-fermion operator.

A prime example for the application of QCD factorisation is the decay of a heavy B meson into two mesons M_1 and M_2. While QCD factorisation can be applied for a wide range of non-leptonic B meson decays, in the following, the mesons $M_{1,2}$ are assumed to be light pseudoscalar mesons, e.g. K^- and π^0, with masses, $m_{1,2}$, of the order of Λ_{QCD}. The matrix element for such a weak two-body decay $B \to M_1 M_2$ reads $\langle M_1 M_2 | \mathcal{H} | B \rangle$, where $\mathcal{H}$ is the Hamiltonian.

In a first step, the contribution of the weak current to the matrix element is addressed. The maximum energy of the weak current in a B-meson decay is set by the mass of the b-quark, m_b, which is much smaller than the weak energy scale set by the mass of the $W^\pm$ boson, m_W. This allows to integrate out the high-energy weak propagator and obtain an effective Hamiltonian

$$\mathcal{H}_{\text{eff}} = \frac{G_F}{\sqrt{2}} \lambda \sum_i C_i(\mu) \mathcal{O}_i, \tag{1.21}$$

where G_F is the Fermi constant, λ is the contraction of the CKM-matrix elements involved in the decay and $C_i(\mu)$ are the so-called *Wilson coefficients* at an energy scale μ multiplying local operators $\mathcal{O}_i$.

1.2.1 QCD factorisation at leading order in α_s

The simplest method to evaluate the resulting hadronic matrix elements $\langle M_1 M_2 | \mathcal{O}_i | B \rangle$ is through naive factorisation, in which the matrix element of the four-quark operator is approximated by the product of two separate matrix elements. Considering the leading order process with no gluon exchange between the mesons M_1 and M_2 or B and $M_{1,2}$, depicted in Figure 1.2, the matrix element simplifies to

$$\langle M_1 M_2 | \mathcal{O}_i | B \rangle \to \langle M_2 | \mathcal{O}_i | 0 \rangle \langle M_1 \mathcal{O}_i | B \rangle. \tag{1.22}$$

The matrix element involving M_2 simplifies to the process-independent decay constant, f_{M_2}, which can be calculated in Lattice QCD [18] or measured from experimental data. The second factor on the right side of Eq. 1.22 reduces to the form factors $F_j^{B \to M_1}(m^2)$, where m_2 is the momentum transfer of the weak current. The form factors encode the process-dependent non-perturbative QCD effects of the decay and can be calculated from Lattice QCD or extracted from experimental data. In the limit of $m_2 = 0$, the decay is characterised by a single form factor and the decay amplitude scales approximately as

$$\mathcal{A}(B \to M_1 M_2) \sim G_F m_b^2 f_{M_2} F^{B \to M_1}(0). \tag{1.23}$$

While this approach of naive factorisation provides a simplistic approximation of the hadronic matrix element, it suffers from neglecting processes that cannot be factorised

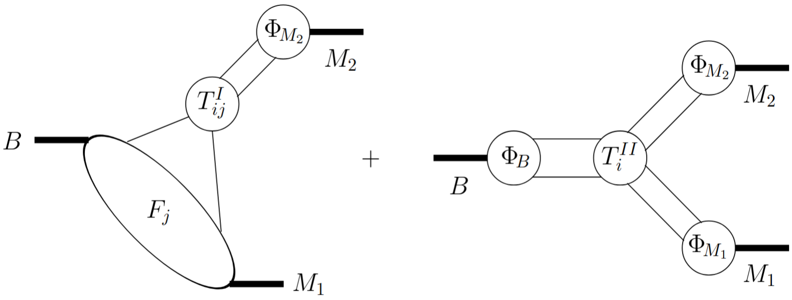


Figure 1.3: Graphical representation of the factorization formula in Eq. 1.24. Taken from [19]

into the product of the decay constant and the form factors, e.g. from gluon exchanges between the light spectator quark and M_2.

1.2.2 QCD factorisation at order α_s

The naive factorisation can be systematically improved upon by including perturbative QCD corrections arising from gluon exchanges. In general, these contributions can be organised according to their order in the strong coupling constant, α_s, and their suppression in powers of Λ_{QCD}/m_b. The process $B \to M_1 M_2$ at arbitrary order of α_s but leading power in Λ_{QCD}/m_b can be expressed through the factorisation formula, which is also graphically illustrated in Figure 1.3

$$\begin{aligned}\langle M_1 M_2|\mathcal{O}_i|B\rangle = \sum_j F_j^{B\to M_1}(m_2^2)\int_0^1 du\, T^I_{ij}(u)\phi_{M_2}(u)\\ + \int_0^\infty d\omega \int_0^1 du \int_0^1 dv\, T^{II}_i(\omega,u,v)\phi_B(\omega)\phi_{M_1}(v)\phi_{M_2}(u).\end{aligned} \tag{1.24}$$

Here, $F_j^{B\to M_1}$ are the process-dependent form factors, and the T^I_{ij} and T^{II}_i are hard-scattering kernels encoding corrections from hard-gluon exchanges up to power corrections in Λ_{QCD}/m_b. The functions ϕ denote the light-cone distribution amplitudes (LCDAs) and encode the non-perturbative substructure of the corresponding mesons.

For light mesons, the leading-twist LCDA, $\phi_M(u,\mu)$, represents the amplitude for finding a valence quark carrying a longitudinal momentum fraction u and the anti-quark carrying $1-u$, evaluated at a renormalisation scale, μ. The LCDA for the heavy B meson is introduced later in the discussion of structure-dependent corrections. Such amplitudes play a central role in encapsulating soft-QCD dynamics in QCD factorisation.

In the following discussion, the focus is restricted to the leading-power corrections arising from one-gluon exchanges at order α_s, which can be calculated in perturbation theory within the QCD factorisation framework.

Structure-independent corrections

The first term in Eq. 1.24, corresponding to the left-hand side in Figure 1.3, encapsulates all contributions, which are proportional to the form factor $F_j^{B\to M_1}$ and are independent of the structure of the B meson. This includes, for example, the leading-order process in the naive factorisation, as well as corrections from one-gluon exchanges.

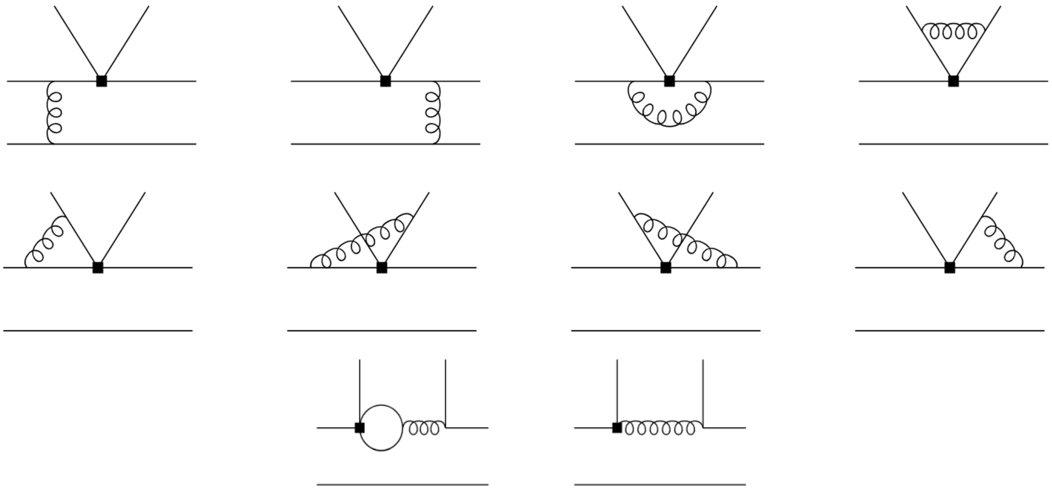


Figure 1.4: Structure-independent processes with one-gluon exchange. Factorisable vertex corrections (top). Non-factorisable vertex corrections (centre). Penguin diagram (bottom left) and diagram corresponding to the chromomagnetic dipole operator in the weak effective Hamiltonian (bottom right). Taken from Ref. [19]

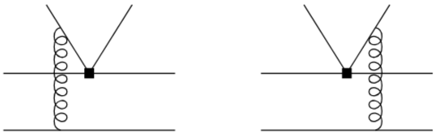


Figure 1.5: Hard one-gluon interaction with the spectator quark. Taken from [19]

The one-gluon exchanges illustrated in the top row of Figure 1.4 are the so-called *factorisable vertex corrections*, which are already absorbed in the definition of the physical form factors and meson decay constant.

By contrast, the *non-factorisable* vertex corrections and *penguin diagrams* shown in the second and third row of Figure 1.4 cannot be absorbed in the form factors or decay constants and are not accounted for in the naive factorisation approach. It can be shown that for these diagrams, the exchange of soft gluons is suppressed and only hard-gluon exchanges contribute[5]. They can be calculated perturbatively, leading to process-dependent correction terms to the hard scattering kernel T^I_{ij}. However, the hard-gluon exchanges with the constituents of M_2 make these corrections sensitive to the LCDA of the emitted meson, $\phi_{M_2}(u,\mu)$. The resulting contributions to the hadronic matrix element are computed as a convolution of ϕ_{M_2} with the hard scattering kernel T^I_{ij}, reflecting the interplay of the hadronic soft-QCD structure and the hard-gluon dynamics.

Structure-dependent corrections

The second term in Eq 1.24 corresponding to the right-hand side of Figure 1.3 arises from hard-gluon interactions between the spectator quark in the B meson and the quarks forming the meson M_2 as illustrated in Figure 1.5. Unlike the previously discussed corrections,

---

[5] The color-transparency argument, first proposed by Bjorken [20], states that soft gluons interact only with the color dipole moment of the color-singlet M_2. Such processes are suppressed by Λ_{QCD}/m_b and thus outside the scope of this introduction.

these interactions are also sensitive to the structure of the B meson. Consequently, they are described by a convolution of the hard scattering kernel T_i^{II} with the LCDAs of all three mesons ϕ_{M_1}, ϕ_{M_2} and ϕ_B.

In contrast to the LCDAs of light-mesons, the B meson LCDA is defined within the framework of *Heavy Quark Effective Theory* (HQET), which simplifies the description of the B meson in the limit $m_b \to \infty$. In this limit, a B meson is represented by a static heavy b-quark field located at the origin and a light spectator quark carrying momentum q in the B meson rest frame. This momentum is projected onto the light cone using a light-like vector $n = (1, 0, 0, -1)$, yielding the light-cone momentum of the spectator quark defined as $\omega = n \cdot q$. The leading-twist[6] B meson LCDA, $\phi_B^+(\omega, \mu)$, encodes the amplitude to find the light spectator quark in the B meson carrying light-cone momentum ω at the renormalisation scale μ. A possible model for the B meosn LCDA is shown in Figure 1.6.

In the hard spectator interaction, the gluon-propagator connecting the spectator quark and the meson M_2 is sensitive to the light-cone momentum of the spectator quark, such that the hard scattering kernel behaves approximately as $\sim 1/\omega$. This leads to a contribution to the hadronic matrix element of

$$\langle M_1 M_2 | \mathcal{O}_i | B \rangle \sim \int_0^\infty \frac{d\omega}{\omega} \phi_B^+(\omega) \int_0^1 du \int_0^1 dv T_i^{II}(u, v) \phi_{M_1}(u) \phi_{M_2}(v) + F_j^{B \to M_1} \text{ terms.} \tag{1.25}$$

The first integral defines the first inverse moment of the B meson LCDA,

$$\lambda_B^{-1}(\mu) \equiv \int_0^\infty \frac{d\omega}{\omega} \phi_B^+(\omega, \mu), \tag{1.26}$$

which is a non-perturbative, process-independent parameter encoding the B meson substructure within the factorised matrix elements.

Since the contribution from hard-spectator interactions can be numerically comparable to the form-factor terms, precise knowledge of the B meson LCDAs is crucial for reliable predictions within QCD factorisation. While the LCDAs for light mesons are accessible through Lattice QCD computations [21] and experimental measurements, for example from the process $\gamma\gamma^* \to \pi^0$ [22, 23], the B meson LCDA and its inverse moment are less well constrained. This uncertainty on λ_B currently drives the theoretical uncertainty in QCD factorisation calculations.

Conclusion

As illustrated by the example of an exclusive non-leptonic $B \to M_1 M_2$ decay into two light mesons, the QCD factorisation approach enables a systematic separation of short- and long-distance QCD contributions, thereby simplifying the calculation of hadronic matrix elements. Short-distance effects are isolated in perturbatively calculable hard-scattering kernels through expansion in powers of α_s, while the non-perturbative, process-dependent dynamics are absorbed in form factors and decay constants. By convoluting the hard-scattering kernel with the LCDAs of the participating mesons, the interplay between the hadronic substructure and the short-distance QCD dynamics is encapsulated.

---

[6] In addition to the leading-twist LCDA, $\phi_B^+(\omega, \mu)$, there exists a subleading-twist distribution amplitude $\phi_B^-(\omega, \mu)$, which encodes the transverse momentum and spin-structure of the light spectator quark inside the B meson. Contributions involving $\phi_B^-(\omega, \mu)$ are generally power suppressed by Λ_{QCD}/m_b but can play an important role in subleading-power corrections.

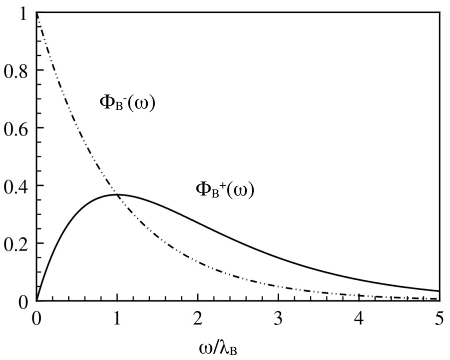


Figure 1.6: Model for the B meson LCDA. Adapted from Ref. [24].

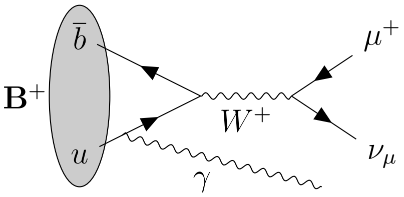


Figure 1.7: Leading order Feynman diagram for the decay $B^+ \to \mu^+\nu_\mu\gamma$ with photon emission from the light quark.

It is shown that precise knowledge of the B meson LCDA, and in particular its inverse moment, λ_B, is essential for reliable predictions of non-leptonic B meson decays within the QCD factorisation framework. In the following section, the decay $B^+ \to \ell^+\nu_\ell\gamma$ is studied in more detail, as it provides a sensitive probe to extract knowledge on the B meson LCDA from experimental data.

1.3 QCD factorisation in radiative leptonic *B* decays

Similarly to the purely leptonic decay $B^+ \to \ell^+\nu_\ell$, the decay $B^+ \to \ell^+\nu_\ell\gamma$ is mediated by an annihilation of the constituent quarks into a virtual W^+ boson decaying into a lepton and a neutrino. The leading order Feynman diagram is shown in Figure 1.7. While the emission of the additional photon is suppressed by one power of α_{em}, it also lifts the helicity suppression that is present in the purely leptonic decay. For muonic final states, this leads to branching fractions for the radiative decay that are an order of magnitude larger than their non-radiative counterpart [25].
The decay $B^+ \to \ell^+\nu_\ell\gamma$ has received a great deal of attention as it is considered the golden mode to probe the B meson substructure [25–27]. As was pointed out in the previous section, the limited knowledge of the leading-twist B meson LCDA $\phi_B^+(\omega)$ and its inverse moment, λ_B, is the leading uncertainty in QCD factorisation approaches. Inferring the value of λ_B from theoretical models is very challenging. The most advanced calculations from QCD sum-rules lead to a value of $\lambda_B = 440 \pm 110\,\text{MeV}$ [28]. From the phenomenology of non-leptonic B meson decays, a value of $\lambda_B \approx 220\,\text{MeV}$ is inferred [29], while constraints from lattice QCD point to a value of $\lambda_B = 338 \pm 68\,\text{MeV}$ [30]. To improve the knowledge on the value of λ_B, it has been suggested to measure it experimentally, for which the decay $B^+ \to \ell^+\nu_\ell\gamma$ is a prime candidate [27].

Factorisation of $B^+ \to \ell^+\nu_\ell\gamma$ decays has been studied in the heavy quark expansion and for sufficiently energetic photons $E_\gamma \gg \Lambda_{QCD}$ [26,31]. In this case, the strong interaction inside the B meson becomes perturbative and factorisation can be established. At first order, the branching ratio only depends on the leading-twist B meson LCDA $\phi_B^+(\omega)$ and is proportional to $\sim 1/\lambda_B^2$ [28,32]. Following [33], the decay amplitude can be written as

$$\mathcal{A}(B^+ \to \ell^+\nu_\ell\gamma) = \frac{G_F V_{ub}^*}{\sqrt{2}} \langle \bar{\ell}\nu_\ell\gamma | \ell\gamma^\mu(1-\gamma_5)\bar{\nu}_\ell \cdot u\gamma_\mu(1-\gamma_5)\bar{b}|B^+\rangle. \tag{1.27}$$

For photon emissions from the constituents of the B meson, the hadronic transition can be re-written in terms of two form factors F_V and F_A, which are denoted the vector and axial-vector form factors, respectively. These form factors can be parametrised in terms of the momentum transfer, q^2, or equivalently the photon energy, E_γ. The differential decay rate[7] is then given as

$$\frac{d\Gamma}{dE_\gamma} = \frac{\alpha_{\text{em}} G_F^2 \left|V_{ub}\right|^2}{6\pi^2} m_B E_\gamma^3 \left(1 - \frac{2E_\gamma}{m_B}\right) \left(|F_V|^2 + \left| F_A + \frac{Q_\ell f_B}{E_\gamma} \right|^2 \right), \tag{1.28}$$

where m_B is the mass of the B meson, Q_ℓ is the charge of the charged lepton and f_B is the B meson decay constant. For energetic photons, $E_\gamma \gg \Lambda_{QCD}$, the strong interaction inside the B meson becomes perturbative and the form factors can be expanded as a function of E_γ as

$$\begin{aligned} F_V\left(E_\gamma\right) &= \frac{Q_u f_B m_B}{2E_\gamma \lambda_B(\mu)} R\left(E_\gamma, \mu\right) + \xi\left(E_\gamma\right) + \Delta\xi\left(E_\gamma\right), \\ F_A\left(E_\gamma\right) &= \frac{Q_u f_B m_B}{2E_\gamma \lambda_B(\mu)} R\left(E_\gamma, \mu\right) + \xi\left(E_\gamma\right) - \Delta\xi\left(E_\gamma\right), \end{aligned} \tag{1.29}$$

where Q_u is the charge of the light quark. The first term is the leading power contribution in the heavy-quark expansion and originates from the emission of the photon from the B mesons's constituents. The term $R(E_\gamma, \mu)$ is a radiative correction factor of the leading power contribution and is unity at tree level. This results in the decay amplitude being inversely proportional to λ_B^2 for the leading-order process in which the photon is emitted from the light quark, making this decay a perfect probe for measuring the value of λ_B experimentally. The terms $\xi(E_\gamma)$ and $\Delta\xi(E_\gamma)$ in Eq. 1.29 are corrections power suppressed in $1/m_b$ and $1/2E_\gamma$ with a symmetry-preserving part common to both F_V and F_A and a symmetry-breaking part. These include higher order corrections for hard and soft photon emissions from the light quark, as well as corrections for photon emissions from the heavy b quark or final state lepton.

Note that due to the radiative correction term, a dependence on the shape of the LCDA beyond the first inverse moment is introduced. These can, in general, be parametrised by the logarithmic moments

$$\sigma_n = \int_0^\infty \frac{d\omega}{\omega} ln^n \left(\frac{\mu}{\omega}\right) \phi_B^+(\omega, \mu). \tag{1.30}$$

This makes it a more delicate operation to extract information on the parameter λ_B from a measurement of the branching ratio of $B^+ \to \ell^+\nu_\ell\gamma$ decays and introduces an inherent

---

[7] Charge conjugation is implied throughout the document unless stated otherwise.

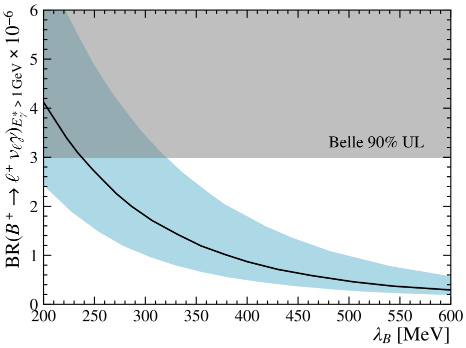


Figure 1.8: Integrated partial branching fraction $\mathcal{B}(B^+ \to \ell^+\nu_\ell\gamma)_{E_\gamma>1\,\mathrm{GeV}}$. The back line represents the exponential LCDA model and the colored band corresponds to LCDA models with σ_1 in the range of ± 0.69. Adapted from Ref. [33]. The region indicated in grey is the exclusion limit from Belle at 90% confidence level [1].

model dependence as shown in Figure 1.8. A commonly used model for the B meson LCDA is the *exponential model*, which only depends on λ_B and has $\sigma_n = 0$. This will serve as the baseline model in the analysis presented in Chapter 4. For λ_B in the region of $200\,\mathrm{MeV} < \lambda_B < 600\,\mathrm{MeV}$ the predicted branching ratio for photon energies $E_\gamma > 1\,\mathrm{GeV}$ ranges from 4×10^{-6} to 3×10^{-7}.

1.3.1 Experimental status on $B^+ \to \ell^+\nu_\ell\gamma$

While it has been shown that the decay $B^+ \to \ell^+\nu_\ell\gamma$ is a prime candidate to probe the B meson LCDA from experimental data, the decay has never been observed. There have been previous searches by the BaBar [34,35] and Belle [1,36] collaborations, the latter of which set the current best limit on the branching ratio for $E_\gamma > 1\,\mathrm{GeV}$, combining the searches with electronic and muonic final states at

$$\mathcal{B}(B^+ \to \ell^+\nu_\ell\gamma) < 3.0 \times 10^{-6} \text{ @90\%CL}, \quad \ell = e^+, \mu^+. \tag{1.31}$$

From the combined result, a lower limit on $\lambda_B > 238\,\mathrm{MeV}$ is set, which manifests the most stringent limit on λ_B from experiment.

1.4 The decay $B_c^+ \to \ell^+\nu_\ell\gamma$

The decay $B_c^+ \to \ell^+\nu_\ell\gamma$ can proceed through the same leading-order processes as its B^+ counterpart, as illustrated in Figure 1.9. Naively, one would expect the decay rate of $B_c^+ \to \ell^+\nu_\ell\gamma$ decays to be enhanced with respect to the respective B^+ decay by a factor $|V_{cb}|^2/|V_{ub}|^2$. A detailed study suggests there are a few more differences between the B^+ and B_c^+ decays linked to the mass of the charm quark. In particular, the charm quark can be assumed to move slowly in a non-relativistic bound state and the interline connecting the weak and electromagnetic decay vertices is of virtuality $m_c \gg \Lambda_{QCD}$. This allows the application of perturbative methods in non-relativistic frameworks and reduces the dependence of the decay rate on the internal structure of the B_c^+ meson [37].

The theoretical description of $B_c^+ \to \ell^+\nu_\ell\gamma$ decays has been established following the discovery of the B_c^+ meson at the CDF experiment [38]. Several different frameworks have

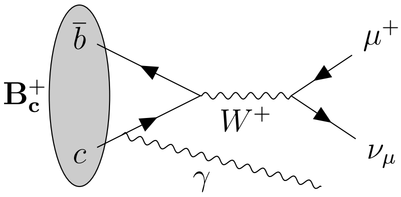


Figure 1.9: Leading order Feynman diagram for the decay $B_c^+ \to \mu^+\nu_\mu\gamma$ with photon emission from the charm quark.

Framework	$\mathcal{B}(B_c^+ \to \mu^+\nu_\mu\gamma)$	Ref.
Light cone sum rules	$\simeq 1 \times 10^{-5}$	[39]
Relativistic QCD	$\simeq 3 \times 10^{-5}$	[40]
Non-relativistic QCD	$\simeq 4.4 \times 10^{-5}$	[41]
Relativistic independent quark model	5.5×10^{-5}	[42]
Non-relativistic QCD (NLO)	8.2×10^{-5}	[2]

Table 1.1: Predicted branching fractions for the decay $B_c^+ \to \mu^+\nu_\mu\gamma$.

been applied to incorporate the internal dynamics of the B_c^+ meson in the prediction of $B_c^+ \to \ell^+\nu_\ell\gamma$ branching fractions. A selection of predictions using different frameworks is given in Table 1.1.

Measuring the branching fraction of $B_c^+ \to \ell^+\nu_\ell\gamma$ decays from experimental data allows to study the applicability of the various frameworks to the decay of B_c^+ mesons and resolve the tensions among the different predictions. In particular the ratio of $\mathcal{B}(B^+ \to \ell^+\nu\gamma)/\mathcal{B}(B_c^+ \to \ell^+\nu\gamma)$ probes the validity of the constituent quark model in B_c^+ decays [39, 40]. In the long term, the decay $B_c^+ \to \ell^+\nu_\ell\gamma$ can also be used as an independent probe of the B_c^+ meson decay constant [37, 41, 42] or the CKM matrix element $|V_{cb}|$ [37, 40–42].

As of today, the decay $B_c^+ \to \ell^+\nu_\ell\gamma$ has never been searched for as B_c^+ mesons are not produced at B-factories. At hadron colliders, the production of B_c^+ mesons is possible, albeit with a much smaller rate compared to B^+ mesons (see Chapter 2). In this context, the search for $B_c^+ \to \mu^+\nu_\mu\gamma$ decays at the LHCb experiment in Chapter 11 marks the first search for this decay.

2

The LHCb experiment at the LHC

In this chapter, a brief overview of the LHCb experiment at the Large Hadron Collider is given. Section 2.1 introduces the Large Hadron Collider with its accelerator complex. Section 2.2 gives an overview of the LHCb experiment, in the configuration as it was operated during Run 2 of the LHC from 2015 to 2018, which is the configuration in which the data for the search for $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays and the majority of the data for the search for $B_c^+ \to J/\psi D^{*+}$ decays are recorded. In Section 2.7, a brief overview of the upgrades of the LHCb experiment for current and future operation is presented.

2.1 The Large Hadron Collider

The Large Hadron Collider [43,44] (LHC) is the world's largest and most powerful particle accelerator to date. It is located at the *Conseil Européen pour la Recherche Nucléaire* (CERN) near Geneva, Switzerland. Inside a circular tunnel about 100 m underground and with a circumference of 27 km, two proton beams[1] are orbiting in opposite directions in two vacuum pipes. The beam pipes are surrounded by superconducting magnets to focus and steer the beams. They are brought to collisions at four distinct interaction points with collisions occurring at a rate of approximately[2] 30 MHz. At these interaction points, the four main experiments ALICE, ATLAS, CMS and LHCb are located.

Before the proton beams are injected into the main LHC accelerator, they pass through several accelerators to gradually increase their energy. A schematic of the CERN accelerator complex is shown in Figure 2.1. Isolated protons are produced by stripping electrons from hydrogen atoms and injected into the Linear Accelerator. In subsequent stages, the protons are injected into the Proton Synchrotron Booster, the Proton Synchrotron (PS) and Super Proton Synchrotron (SPS) and accelerated to energies of 450 GeV. From the SPS, the proton bunches are injected into the LHC, where they are further accelerated to the nominal energy. During Run 1 of the LHC, protons were accelerated to a centre-of-mass energy of $\sqrt{s} = 7$ TeV in 2011 and 8 TeV in 2012. For Run 2 from 2015 to 2018, the centre-of-mass energy was increased to 13 TeV. After the LHC has resumed operation for Run 3 in 2022 the centre-of-mass energy was further increased to 13.6 TeV.

---

[1] In a special configuration, also heavy ions can be injected.

[2] The bunch crossing rate is 40 MHz, but the proton beams also contain empty bunches which do not produce pp collisions.

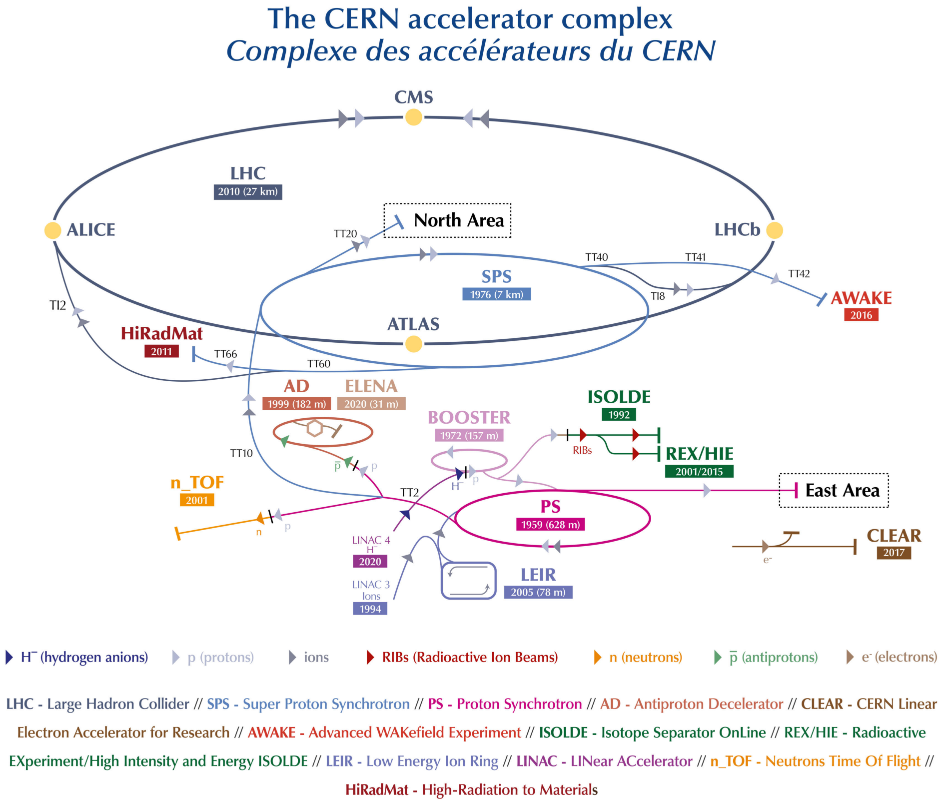


Figure 2.1: Schematic of the CERN accelerator complex [45].

2.2 The LHCb experiment

The *Large Hadron Collider beauty* (LHCb) experiment [46] is primarily designed to study the decays of *beauty* and *charm* hadrons. In the decays of such hadrons, precise measurements of CP violation in weak decays or searches for physics beyond the Standard Model in rare electroweak decays are conducted. To enable such precision studies, the LHCb experiment is operated with a lower instantaneous luminosity than ATLAS and CMS, thereby limiting the number of primary pp interactions. In Run 1 and 2, there was an average of one pp collision per bunch crossing. This is increased to approximately 5.3 pp collisions per bunch crossing for Run 3.

The LHCb detector is a single-arm forward spectrometer covering a region approximately from 10 mrad to 300 mrad (250 mrad) in the bending (non-bending) plane with respect to the beam axis. A schematic of the experiment is shown in Figure 2.2. The geometry is vastly different from other particle detectors and is closely related to the production of beauty hadrons at the LHC. They are formed from an initial $b\bar{b}$ quark anti-quark pair, which subsequently *hadronise* into bound beauty hadron states containing at least one (anti-) bottom quark. The energy required to create a $b\bar{b}$ pair is much smaller than the total available energy of the pp collision. Thus, they are predominantly created in gluon-gluon fusion processes $gg \to b\bar{b}$. The gluons of the colliding protons typically carry a small fraction of the proton energy and are likely to have asymmetric momenta. Consequently, the $b\bar{b}$ pairs produced at the LHC are boosted along the beam axis in forward

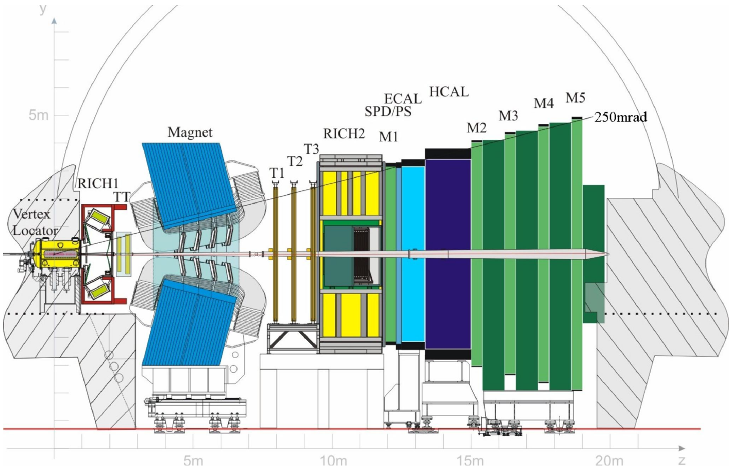


Figure 2.2: Schematic of the LHCb experiment in the configuration of Run 2. Taken from Ref. [46].

or backward direction. The LHCb detector is specifically designed to cover the forward region $2 < \eta < 5$, which is expressed in terms of the pseudorapidity, η, defined as

$$\eta = -log\left(tan\frac{\theta}{2}\right) \tag{2.1}$$

where θ is the polar angle with respect to the beam axis. The geometric acceptance of the LHCb detector covers only about 2.4% of the full solid angle, yet about 20% of all $b\bar{b}$ pairs produced in pp collisions are contained in the LHCb acceptance [47].
The production of beauty hadrons from the hadronisation of $b\bar{b}$ pairs also results in the fact that all possible beauty hadron species are created at the LHC. This includes beauty mesons with a light quark, such as $B^0(d\bar{b})$, $B^+(u\bar{b})$, $B^0_s(s\bar{b})$, beauty mesons with a heavy quark $B^+_c(c\bar{b})$ as well as beauty baryons, such as Λ_b, $\Sigma_b^{0,\pm}$ and $\Xi_b^{0,\pm}$. The relative rates at which the different beauty hadrons are produced are called the *fragmentation fractions*, f_i, which are normalised to unity

$$f_u + f_d + f_s + f_c + f_{\Lambda_b} + f_{\Sigma_b} + f_{\Xi_b} = 1. \tag{2.2}$$

Due to the non-perturbative nature of QCD, they cannot be calculated but are measured from experimental data. In this work, decays of the B^+ and B^+_c mesons are studied for which the relative fragmentation fractions in pp collisions at 13 TeV in the LHCb acceptance are measured to be [48]

$$\frac{f_c}{f_u + f_d} = (3.78 \pm 0.90) \times 10^{-3} \quad \text{and} \quad f_u = f_d = 0.346 \pm 0.007. \tag{2.3}$$

2.3 Tracking system

The trajectories of charged particles traversing the detector are reconstructed from hits in the different tracking stations of the LHCb detector. Together with the dipole magnet, they form the tracking system. From the curvature of the trajectory in the magnetic field, the momentum of charged particles is determined, and by combining multiple tracks, the

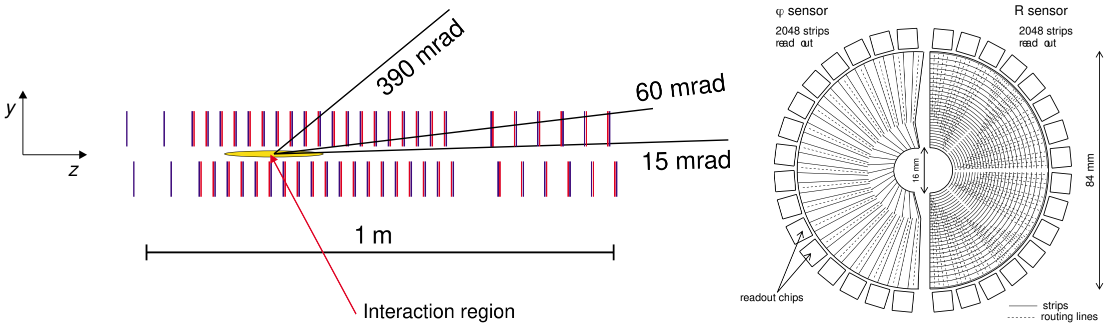


Figure 2.3: Schematic view of the position of VELO modules (left) and the R- and ϕ sensors of each module (right). Taken from Ref [49].

position of decay vertices can be derived. The following provides a brief overview of the subdetectors that form the tracking system.

2.3.1 Vertex Locator (VELO)

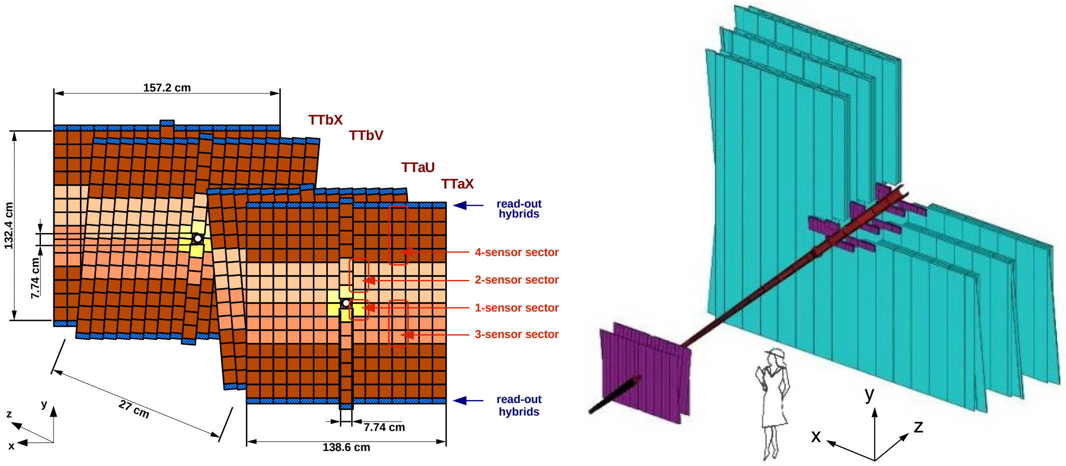
The LHCb Vertex Locator (VELO) [49, 50] is the innermost part of the LHCb detector. Its main purpose is to allow for a precise reconstruction of the position of the primary pp interaction vertices and the reconstruction of secondary B decay vertices. The VELO is composed of two halves, surrounding the region of the primary pp collision, and is separated from the LHC vacuum by a thin aluminium foil. Each half is equipped with 21 semi-circular silicon microstrip modules.

In stable p-beam conditions, the VELO modules are separated from the beam axis by just 7 mm. During injection of the beams, the two VELO halves are retracted to an aperture of about 60 mm to prevent damage to the modules from direct interactions with the uncollimated beams.

Each module has a thickness of about 300 µm and consists of two layers, one with radial (R sensors) and one with azimuthal (ϕ sensors) silicon strips. A schematic of the VELO halves and the arrangement of the R- and ϕ sensors is shown in Figure 2.3. The single hit resolution of a VELO sensor ranges from 4 µm to 30 µm depending on the angle of the track with respect to the sensor and the position in (R, ϕ) [50]. From the tracks of charged particles reconstructed with the VELO, the position of the PV can be determined. The experimental precision on the PV position depends strongly on the number of reconstructed VELO tracks. For a typical event with 35 reconstructed VELO tracks, the PV can be determined with a precision of 13 µm transverse to the beam axis and 71 µm along the beam axis [50].

2.3.2 Tracking stations

The LHCb detector has two additional planar tracking systems upstream and downstream of the dipole magnet. The Trigger Tracker (TT) [51,52] is located upstream of the magnet at a position $z = 2.5$ m from the primary interaction vertex. It is composed of four planar silicon microstrip layers measuring about 150 cm in width and 130 cm in height to cover the full LHCb acceptance. A schematic of the TT stations is shown in Figure 2.4a. To



(a) (b)

Figure 2.4: (a) Schematic of the TT stations. Taken from Ref [52]. (b) Schematic of the TT and T1-T3 tracking stations. Silicon microstrip detectors are shown in violet and the OT using drift tube technology is shown in blue. Taken from Ref [46].

achieve sensitivity to the y-position of hits from charged particles, two stereo layers are tilted by ±5° with respect to the vertical axis. This yields a spatial hit resolution of about 50 μm [51].

The main tracking stations (T1-T3) are located downstream of the magnet at a position about 8 m from the primary interaction vertex. A schematic of the tracking stations is shown in Figure 2.4b. Each of the downstream tracking stations is composed of four layers measuring about 6 m in width and 5 m in height covering the full LHCb acceptance. Similarly to the TT, two of the four layers are stereo layers tilted by ±5° with respect to the vertical axis, making them sensitive to the y-position of charged particle hits. In each of the layers, two different detector technologies are used for the inner and outer regions. The Inner Tracker (IT) [53] is the innermost part of the downstream tracking system, which uses the same silicon microstrips as the TT. Despite the IT covering only 1.2% of the detector acceptance, it is subject to 20% of all track hits.

The area of lower occupancy surrounding the IT is covered by the Outer Tracker (OT) [54–56]. It is constructed from 2.5 m long drift tubes with a diameter of 4.9 mm. The drift tubes are filled with a gas mixture of Ar, CO_2 and O_2. Charged particles traversing the drift tubes ionise the gas and the resulting charges are separated by an electrical field across the tube. From the drift time to the anode, the position of the charged particle hit is determined with a resolution of 200 μm.

2.3.3 Dipole magnet

The LHCb dipole magnet [57] is installed at a position $4\,\text{m} < z < 7\,\text{m}$ downstream of the primary interaction vertex. The coils enclose the full detector acceptance of 300 mrad horizontally and 250 mrad vertically and are surrounded by an iron yoke. The magnetic field is strongest in the volume enclosed by the coils, but it is not restricted to the volume of the magnet and stretches far beyond the coils. A schematic of the magnet geometry and the magnetic field strength as a function of the z-position is shown in Figure 2.5. For

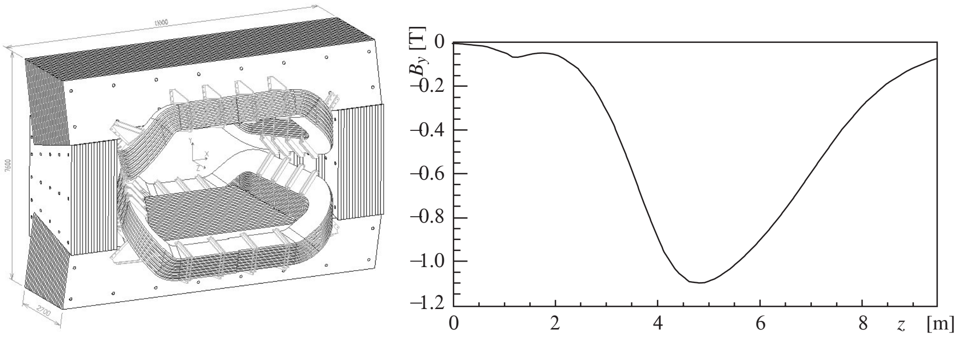


Figure 2.5: Left: Schematic of the LHCb dipole magnet [57]. Right: Magnetic field as a function of the z position [51].

particles traversing the detector up to $z = 10\,\text{m}$, the integrated magnetic field is about $4\,\text{Tm}$. From the track segments upstream (using VELO and TT) and downstream (using T1-T3) of the magnet, the curvature of a charged particle's trajectory is determined, which directly relates to its momentum.

2.3.4 Performance of the tracking system

The track reconstruction algorithms in LHCb are designed such that track segments are built in each tracking station (VELO, TT, T1-T3) before being merged into a single track with inputs from different tracking stations. Depending on which tracking systems are included to form a track, different *track types* are defined, which are also illustrated in Figure 2.6.

- Long tracks: Including VELO and T1-T3 and optionally TT
- Downstream tracks: Including TT and T1-T3 but not VELO
- Upstream tracks: Including VELO and TT but not T1-T3
- VELO tracks: Including only VELO (no TT or T1-T3)
- T-tracks: Including only T1-T3 (no VELO or TT)

For the analysis of particle decays at LHCb, three performance numbers related to the tracking system are key: the relative momentum resolution of a charged particle track, the vertex resolution combining multiple charged tracks and the efficiency of reconstructing a charged particle track. For the majority of physics analyses in LHCb, as well as for the ones presented in this thesis, only long tracks are considered, as they have superior momentum resolution and offer the best secondary vertex resolution. For long tracks, the momentum resolution $\Delta p/p$ depends on the absolute momentum and ranges from 0.5% at 20 GeV to 1% at 200 GeV [58]. The vertex resolution of secondary B-hadron decays in general depends on the topology of the decay. For the example of a $B_s^0 \to D_s^+\pi^-$ decay, a resolution of $24\,\mu\text{m}$ in transverse and $144\,\mu\text{m}$ in longitudinal direction is achieved [59]. The track reconstruction efficiency for muons reconstructed as long tracks is measured to be above 95% [60] while for electrons it is expected to be lower due to the emission of bremsstrahlung.

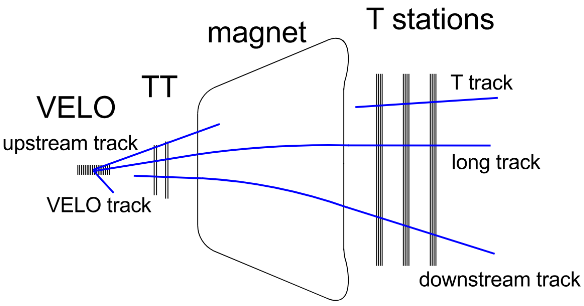


Figure 2.6: Illustration of the LHCb tracking detectors and the different track types reconstructed.

2.4 Particle identification system

In many physics analyses, the identification of charged particles is crucial for precision measurements. For this reason, LHCb is equipped with a particle identification system to separate different particle species traversing the detector. Furthermore, the calorimeter system allows for the reconstruction of neutral objects, such as neutral hadrons and photons. The following provides a brief overview of the particle identification (PID) system.

2.4.1 RICH

Two Ring Imaging Cherenkov (RICH) detectors [61, 62] are positioned upstream and downstream of the magnet. Charged particles traversing a medium with a velocity $v = \beta c$ higher than the specific speed of light in the medium emit Cherenkov radiation. The radiation is emitted in a cone around the flight direction with opening angle $\cos\theta_C = 1/\beta n$, where n is the refractive index of the medium. The RICH detectors in LHCb are designed to measure the opening angle of the Cherenkov cone, θ_C. Together with the momentum information from the tracking system, the velocity and thus the mass of the charged particle can be inferred. This allows to distinguish different particle species and assign a particle hypothesis to charged tracks, as is shown in Figure 2.7.

The Cherenkov light produced in the radiator medium is reflected outside of the detector acceptance using a number of spherical mirrors, where it is recorded by hybrid photon detectors as shown in Figure 2.8. In order to increase the sensitivity in a wide momentum

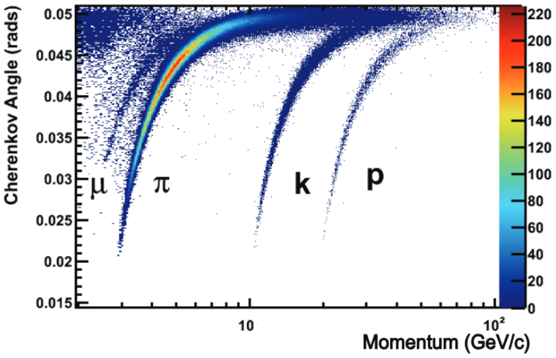


Figure 2.7: Reconstructed Cherenkov angle as a function of track momentum in the C_4F_{10} radiator of RICH1 [62].

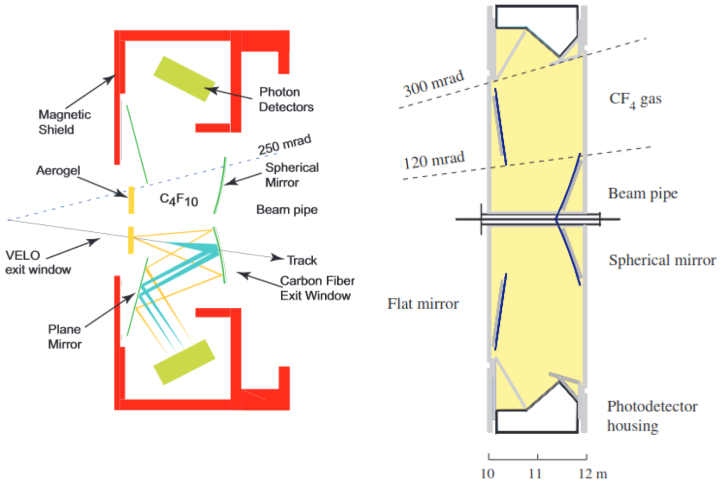


Figure 2.8: Schematic of the RICH1 (left) and RICH2 (right). Taken from Ref [61].

range, the two RICH detectors use different radiator mediums with different refractive indices. In RICH1 upstream of the magnet C_4F_{10} gas with $n = 1.0014$ is used as a radiator providing information on particle identification in a momentum range from 2 GeV to 40 GeV across the full LHCb acceptance. In contrast, RICH2 downstream of the magnet uses CF_4 with $n = 1.0005$ as a radiator to enhance sensitivity in the high-momentum region from 15 GeV to 100 GeV. Also, the acceptance of RICH2 is limited to 15-120 mrad as a result of high momentum particles being produced in the very forward region.

2.4.2 Calorimeter

The LHCb calorimeter system [63] is crucial for measuring the energy of neutral particles and improving the particle identification of both charged and neutral particles. The calorimeter system consists of the Scintillating Pad Detector, Preshower Detector, electromagnetic calorimeter, and hadronic calorimeter.

Scintillating Pad and Preshower Detector

The Scintillating Pad Detector (SPD) and the Preshower Detector (PS) are the first subdetectors of the calorimeter system. Two planar layers of rectangular scintillator pads are installed in front of the electromagnetic calorimeter. The two layers are separated by a 15 mm thick lead converter. The geometry and arrangement of the read-out channels are aligned with the geometry of the electromagnetic calorimeter discussed in the following paragraph.

The first layer of scintillator detector, installed before the converter material, forms the SPD detector. The separation of charged and neutral particles, particularly electrons and photons, is achieved as charged particles leave an energy deposit in the scintillator material, whereas neutral particles do not. Figure 2.9 shows an illustration of the energy deposition of different particle species in the calorimeter subsystems. Similarly, the PS installed after the lead converted enhances the separation between electrons and charged hadrons based on the energy deposit in the scintillator material.

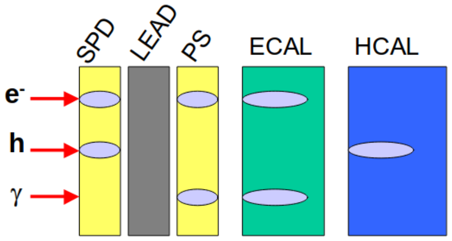


Figure 2.9: Illustration of the energy deposit of electrons, hadrons and photons in the different calorimeter subsystems.

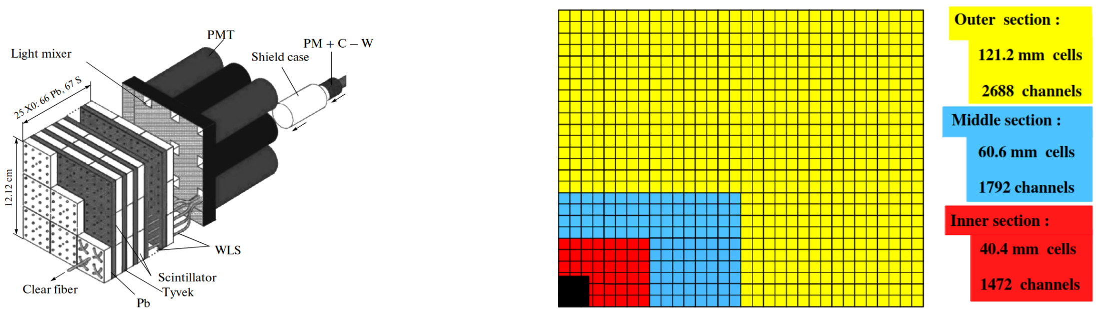
Electromagnetic calorimeter

The main purpose of the electromagnetic calorimeter (ECAL) is to measure the energy of electromagnetic showers induced primarily by photons and electrons. It covers the full LHCb acceptance across an area of 7.8 m × 6.3 m and is installed about 12.5 m downstream from the primary interaction vertex. It is constructed from alternating layers of absorber and scintillator plates, also known as *shashlik* technology, with a total thickness corresponding to 25 radiation lengths X_0. A schematic of a shashlik module is shown in Figure 2.10a.

The account for the different occupancy across the detector, the ECAL is divided into three separate regions, each with a different granularity of readout cells. In the innermost section with the highest occupancy, the cells measure 4 cm × 4 cm, in the middle section 6 cm × 6 cm and in the outer section with the lowest occupancy, they measure 12 cm × 12 cm. The separation of the ECAL in the different sections is shown in Figure 2.10b.

As a particle penetrates the ECAL modules, it is gradually stopped by the absorber material and generates an electromagnetic *shower* while interacting with the scintillator material. The light produced in the scintillator material is transported by wavelength-shifting fibres to photomultiplier tubes (PMT) installed on the downstream side of the ECAL. The light yield recorded by the PMT is then used to infer the energy deposit in a given cell.

To measure the energy of a particle, the energy deposition of several neighbouring cells has to be combined as electromagnetic showers extend beyond the geometry of a single



(a)

(b)

Figure 2.10: (a) Schematic of a shashlik module in the LHCb ECAL. Taken from Ref. [64]. (b) Lateral segmentation of ECAL modules in the different regions for one quadrant of the detector. Taken from Ref. [63].

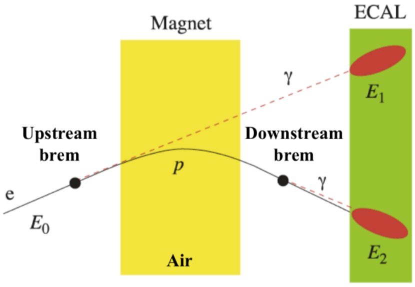


Figure 2.11: Illustration of bremsstrahlung emission of an electron upstream and downstream of the magnet. Taken from Ref. [46].

module or when particles strike near the edge of a cell. This is achieved by a clustering algorithm, which computes the total energy from a 3×3 cluster around the cell with a local maximum energy deposit. For $\pi^0 \to \gamma\gamma$ decays, special algorithms are used to take into account the overlap of the two photon clusters. The energy resolution which is achieved is [65]

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E[\,\mathrm{GeV}]}} \oplus 1\%. \tag{2.4}$$

Bremsstrahlung recovery

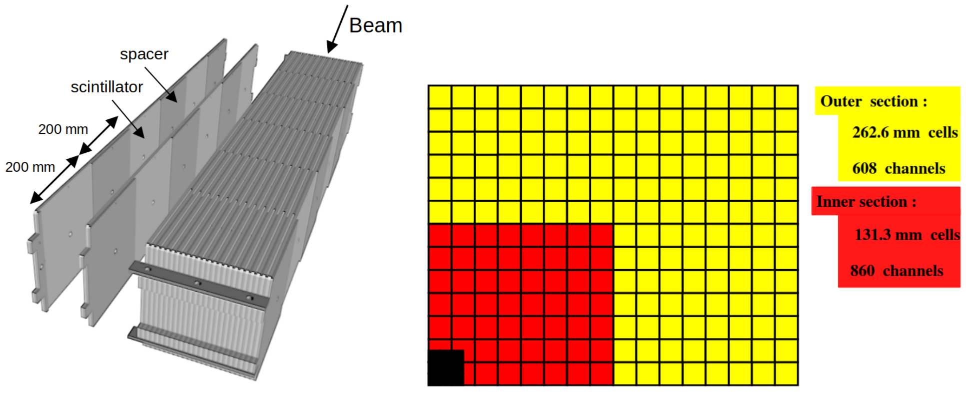
Electrons that traverse the detector are subject to emitting bremsstrahlung photons as they interact with the detector material. They are produced co-linearly to the electron flight direction. In the case, the bremsstrahlung photon is emitted downstream of the magnet, the momentum of the electron is still measured correctly from the tracking system. Also, the measurement of E/p using the ECAL is correct as the energies of the electron and the accompanying photon are contained in the same cluster.

In the case, the bremsstrahlung photon is emitted upstream of the magnet, the electron trajectory is bent by the magnet, whereas the photon continues in a straight line. Thus, the electron and the accompanying bremsstrahlung photon end up in different ECAL clusters. To recover these bremsstrahlung photons, the electron track upstream the magnet is extrapolated linearly to the ECAL and photon clusters with $p_T > 75\,\mathrm{MeV}$ are associated as bremsstrahlung photons to the electron track. The procedure is illustrated in Figure 2.11

Hadronic calorimeter

The hadronic calorimeter (HCAL) is located downstream of the ECAL and has an active area of $8.4\,\mathrm{m}x6.8\,\mathrm{m}$ covering the full LHCb acceptance. It is designed to measure the energy of hadron-induced showers using a shashlik technology similar to the ECAL. Contrary to the ECAL, the shashlik structure in the HCAL modules is oriented longitudinally to the beam axis as shown in Figure 2.12a.

Each module is composed of periodically stacked layers of iron plates interlaced by layers of six steel spacer plates and six polystyrene scintillator plates. This yields a 1.2 m long module weighing 9.5 tons corresponding to about 5.6 nuclear interaction lengths λ_I. Wavelength-shifting fibres along the edge of the modules transport the scintillation light to the PMTs. The energy resolution of the HCAL is significantly reduced with respect to



(a) (b)

Figure 2.12: (a): Schematic of an HCAL module. (b): Lateral segmentation of HCAL modules in the different regions for one quadrant of the detector. Adapted from Ref [63].

the ECAL and reads [65]

$$\frac{\sigma_E}{E} = \frac{69\%}{\sqrt{E[\,\mathrm{GeV}]}} \oplus 9\%. \tag{2.5}$$

Similar to the ECAL, the HCAL is also divided into different regions to account for the varying occupancy across the detector. In the outer region, the readout cells measure 262 mm × 262 mm while in the inner regions they measure 131 mm × 131 mm. The segmentation into the different regions is shown in Figure 2.12b.

2.4.3 Muon stations

The muon system [66, 67] is composed of five muon stations M1-M5, each equipped with 276 multi-wire proportional chambers. While the station M1 is located between RICH2 and the calorimeter system, the remaining stations M2-M5 are located downstream of the HCAL and are separated by 80 cm thick iron absorbers. Due to the muons' characteristic of being minimally ionising particles, they can traverse all five muon stations if their momentum is above $p > 6\,\mathrm{GeV}$. For muons above this threshold, hits in all five muon chambers can be found and their transverse momentum can be estimated solely from the muon stations with a relative momentum resolution of 25%.
The muon stations play an essential role in the operation of the LHCb experiment as they allow muons to be identified with high efficiency and are crucial for the LHCb trigger system in Run 1 and Run 2 as explained in Section 2.5.

2.4.4 PID performance

Identification of charged particles is achieved by constructing PID variables that combine information from the RICH detectors, the calorimeter system, the muon stations, and the bremsstrahlung recovery algorithm, and matching them to reconstructed particle tracks. The most commonly used PID variables are the $\mathrm{DLL}_{x\pi}$ variables in which a particle hypothesis x is tested with respect to the pion-hypothesis. The response of each PID subsystem is evaluated under the two different particle hypotheses to yield a system-wise difference in log-likelihood. The difference in log-likelihood is summed across the different

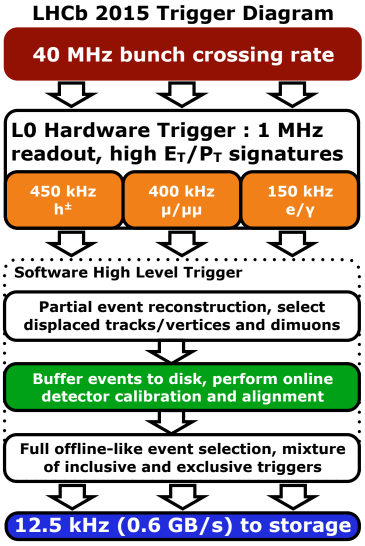


Figure 2.13: Overview of the LHCb trigger system in Run 2. Taken from Ref. [71].

PID systems to yield a single $\text{DLL}_{x\pi}$ variable. More details on the construction of the DLL variables are given in Section 3.3.1. For the identification of kaons, an efficiency of 95% is reached with a $\pi^+ \to K^+$ mis-ID efficiency of 5% [68]. For electrons, the PID efficiency is about 90% for a $\pi^+ \to e^+$ mis-ID efficiency of 6% [58].

Another commonly used set of PID variables are the `ProbNN` variables. These utilise a neural network, trained on clean samples of a given particle species, to construct a probability-like quantity indicating whether a particle is of a particular species.

For muon identification, a special binary PID variable, `isMuon`, is defined. It requires multiple subsequent hits in the muon chambers to be aligned within a certain window. This allows for muon identification with an efficiency of 97% at a $\pi^+/K^+ \to \mu^+$ mis-ID efficiency of 1-3% [69].

2.5 Trigger system

During Run 2, the LHCb experiment recorded *pp* collision data at the LHC's bunch-crossing rate of 40 MHz with an average instantaneous luminosity of $2 \times 10^{33}\,\text{cm}^{-2}\,\text{s}^{-1}$. To cope with about 3 Tb/s of raw data produced by the LHCb experiment, it is necessary to perform a fast selection of events to be saved that is compatible with the LHCb physics program of studying beauty- and charm-hadron decays. This is achieved by a multi-stage trigger system [70] consisting of a level zero (L0) hardware trigger and two high-level software triggers (HLT) to reduce the amount of data saved for physics analyses. An overview of the LHCb trigger system as it was used in Run 2 is shown in Figure 2.13.

2.5.1 L0 hardware trigger

The L0 trigger [72] is integrated in the front-end electronics of the detector subsystems and uses minimal information to reduce the event rate from 40 MHz to about 1 MHz. Events are selected based on large transverse momentum or energy, which is a typical signature of a heavy b-hadron decay. Depending on the particle and detector subsystem that triggers the L0 decision, several L0 trigger categories are defined. The L0 trigger categories, which are used in the presented searches for $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ and $B^+_c \to J/\psi D^{*+}$ decays, are the `L0Muon`, `L0Electron` and `L0Hadron` decisions. The former requires subsequent straight-line hits in all five muon stations with a sufficiently large transverse momentum. In contrast, the `L0Electron` and `L0Hadron` triggers require hits in the SPD and subsequent energy deposits in the calorimeter systems with sufficiently large transverse energy, defined as

$$E_T = E\, sin\theta, \tag{2.6}$$

where E is the sum of the energy deposits in the SPD+ECAL (+HCAL for `L0Hadron`) and θ is the polar angle of the energy deposit. The cluster energy, E, at the L0 stage is calculated combining the energies of 2×2 calorimeter cells to reduce the complexity of the L0 trigger in the calorimeter system.

On top of the kinematic selections at the L0 level, a requirement on the event multiplicity is imposed. For the L0 trigger categories mentioned above, the number of hits in the SPD detector is required to be less than 450 during Run 2. This is to limit the CPU time in the subsequent trigger stages by reducing events with a high detector occupancy.

2.5.2 HLT software trigger

For events that passed the L0 trigger stage, the detector is fully read out and the event information is passed to a CPU-based computing farm on which the high-level software trigger stages HLT1 and HLT2 [72] are run.

At the HLT1 stage, a partial event reconstruction is performed by reconstructing the tracks of charged particles with large transverse momentum ($p_T > 500\,\text{MeV}$), combining information from the different tracking detectors. Also, the primary interaction vertex is reconstructed from VELO tracks. Similarly to the L0 stage, a set of trigger *lines* is defined, which encapsulate typical event signatures for different physics programs. For example, in the b-hadron decays studied in this work, the HLT1 selection requires the presence of energetic particles which are displaced from the primary interaction vertex. Events that meet the requirements of any HLT1 trigger line are saved to a buffer.

Events selected with dedicated HLT1 lines are used to perform an automated detector alignment and calibration in real-time. This includes, in particular, a spatial alignment of the tracking system, calibration of the RICH gas refractive-index and photon detectors, as well as an energy calibration of the ECAL. With constantly updated alignment and calibration constants, optimal conditions for running the full event reconstruction at the HLT2 trigger stage are ensured.

At HLT2, the track reconstruction of HLT1 is repeated, followed by reconstruction of low-momentum tracks and clusters in the calorimeter system. Neutral particles are identified using the isolation of calorimeter clusters from charged tracks and possible bremsstrahlung photons are identified. Combining with information from the PS/SPD as well as the RICH system, full particle identification is provided at HLT2 level. Similarly to HLT1, a set of trigger lines is defined to select events that match the typical characteristics of decays studied in the various physics programs. Events passing the HLT2 selection are saved to storage, at which point the event rate has been reduced to about 12.5 kHz.

2.6 LHCb data flow and simulation

The data flow of the LHCb experiment is designed to provide high-quality data while allowing a high degree of flexibility for the different physics analyses. As explained in the previous section, data from the detector is filtered using a multi-stage trigger system. While the L0 stage is implemented in hardware, the HLT is a software trigger stage and is executed by MOORE [73]. Offline reconstruction from raw data is performed by the BRUNEL [74] application. In Run 2, the data is further filtered in a central event selection, known as *stripping*. This is achieved by a set of *stripping lines*, typically designed by analysts to select specific particle decays for physics analyses. This is controlled by the DAVINCI [74] application, which is also the application used to write the data to a format that can be accessed by analysts.

The same data flow is also used to process simulated events, also called *Monte Carlo data*, which allows to treat simulated events as closely to recorded data as possible. Generation of the initial pp collision is performed by PYTHIA [75] with a LHCb-specific configuration in GAUSS [76]. For the production of B_c^+ mesons, the special event generator BCVEGPY [77] generates the hard scattering process before the hadronisation is generated with PYTHIA. The decay and propagation of particles in the detector is handled by EVTGEN [78] and GEANT4 [79], respectively, while final-state radiation is generated using PHOTOS [80]. Lastly, the response of the electronics is emulated by BOOLE [74].

2.7 Upgrade I of the LHCb detector

After the 2018 data-taking was completed, the LHC entered the four-year phase of *Long Shutdown 2* (LS2) during which the LHCb experiment received a major upgrade, referred to as *Upgrade I*. This upgrade aims to allow the LHCb experiment to be operated at an instantaneous luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in Run 3 as opposed to $4 \times 10^{32}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ during Run 2 with an average of 5.3 visible pp interactions per bunch-crossing. A schematic of the upgraded LHCb detector for Run 3 is shown in Figure 2.14.

Operating the detector in higher pile-up conditions requires a redesign of the tracking stations with finer granularity and increased radiation hardness. A new VELO is designed [81] with pixel sensors arranged in L-shaped half-modules instead of the semi-circular modules equipped with R and ϕ silicon microstrip sensors. The changes of the module geometry allow for moving the VELO even closer to the beam line with a minimum aperture of 5 mm during stable beams.

The TT stations are completely replaced by the Upstream Tracker (UT) [82], which consists of four layers of silicon microstrip detectors, providing finer granularity and improved radiation hardness.

The downstream tracking stations T1-T3 are entirely replaced by the Scintillating Fibre-Tracker (SciFi) [82]. It is composed of three stations, each equipped with four layers of scintillating fibres that are read out by silicon photo-multipliers.

In the two RICH detectors, the hybrid photon detectors are replaced with faster multi-anode photomultiplier tubes and the optics of RICH1 are redesigned to deal with the increased occupancy in the central region [83]. The new photon detection chain is also equipped with new front-end electronics.

The PS and SPD of the calorimeter system are removed, as they cannot be operated in

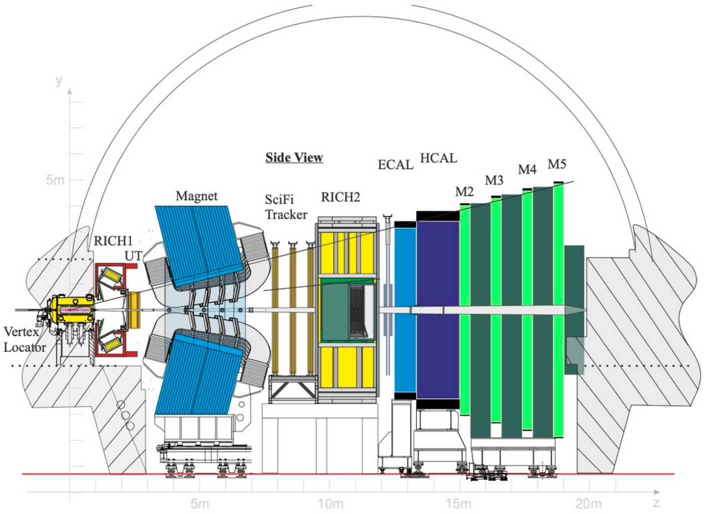


Figure 2.14: Schematic of the LHCb detector in Run 3. Taken from Ref. [85].

the high-occupancy environment of Run 3 [82]. For the same reason, the muon station M1 located in front of the calorimeter system is removed. While the remaining stations of the calorimeter system and muon system remain largely unchanged, they are all equipped with new front-end electronics.

In addition to the changes of the different subsystems, the trigger system is entirely reworked [84]. The higher average event multiplicity in Run 3 would lead to a saturation of the L0 hardware trigger. To overcome this issue, the L0 trigger is removed and a purely software-based trigger system is built. In Run 3, the detector is fully read out at the bunch-crossing rate of 40 MHz. The partial event reconstruction at HLT1 is moved to a GPU-based event filter, which performs the event selection in real-time. Events that pass the HLT1 software-trigger are stored in a buffer to perform automated real-time alignment and calibration. Once aligned and calibrated, the data are passed to the CPU-based HLT2 trigger, at which stage a full event reconstruction is performed.

The LHCb experiment resumed operation in 2022 for Run 3 of the LHC. After commissioning at the start of the run and operation at reduced luminosity, the LHCb experiment reached the nominal luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ in 2024. The current Run 3 is scheduled to continue until 2026, after which the LHC enters *Long Shutdown 3* (LS3). For the time beyond LS3, additional upgrades of the LHCb experiment are planned to increase the amount of recorded data and maximise the physics potential of the experiment. In the following Chapter 3, the focus is drawn to the future upgrades of the calorimeter system.

3 Towards an upgrade of the LHCb calorimeter

Following the conclusion of the currently ongoing Run 3 of the LHC, the accelerator's long-term schedule foresees two additional Run periods: Run 4, presently planned from 2030 to 2033, and Run 5, expected to last from 2036 to 2041. For these Run periods, the accelerator is upgraded to enter the era of the *High-Luminosity LHC* (HL-LHC), capable of delivering much higher instantaneous luminosities to the ATLAS and CMS experiments. For Run 5, these upgrades will also become effective for the ALICE and LHCb experiments. The long-term schedule of the LHC is illustrated in Figure 3.1.

During Long Shutdown 3 (LS3), the LHCb experiment undergoes an upgrade to ensure that high-quality data can be recorded in Run 4. The *LS3 enhancements* target primarily an upgrade of the PID system by improvements to the RICH and calorimeter systems [86]. Looking further ahead to Run 5, the LHCb experiments will need to be operational in an environment of approximately 30 *pp* collisions per bunch crossing when the HL-LHC upgrades come into effect. For comparison, in Runs 3 and 4, there are approximately five primary interactions per bunch crossing. To sustain this busy environment, the LHCb experiment is planned to undergo a major upgrade, named *Upgrade II* [86, 87]. In this upgrade, almost all subdetectors will be rebuilt or redesigned to withstand the higher radiation and occupancy. In addition, the subdetectors are equipped with timing capabilities to record the timestamp of an event with a precision of a few tens of picoseconds, which will allow for differentiation between different *pp* collisions within the same bunch crossing. As the time in LS4 is rather short, some installations are already prepared during LS3.

In the following sections, the upgrades of the LHCb calorimeter system for the LS3 enhancements and Upgrade II are discussed. The mapping of the signal cables presented in Section 3.2 and studies on the expected PID performance for Run 4 in Section 3.3 are the author's original work towards the upgrades of the calorimeter system.

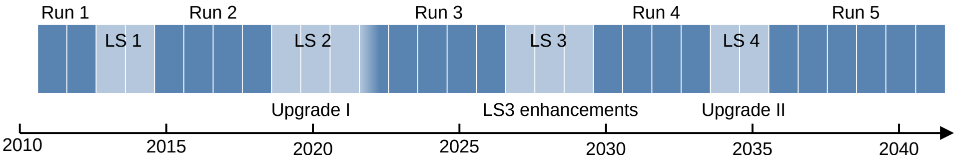


Figure 3.1: Long-term LHC schedule with the run periods relevant to the LHCb experiment indicated on the top. The upgrades of the LHCb experiment are indicated at the bottom.

3.1 Overview of the calorimeter upgrade

For Run 4, the ECAL will be entirely dismantled and rebuilt with finer granularity [86]. The layout of the ECAL for Run 4 is shown on the left in Figure 3.2. In the outer region, the currently installed Shashlik modules will be refurbished and equipped with new wavelength-shifting fibres and with new PMT readout. Some modules are equipped with smaller scintillator tiles to accommodate the increased number of shashlik modules with finer granularity. A schematic of a shashlik module is shown in Figure 3.3a.
In the inner regions, the ECAL is equipped with new *Spaghetti Calorimeter* (SpaCal) modules with finer granularity that can withstand the radiation damage in the high-occupancy region [88]. The SpaCal modules are built from lead (Pb) or tungsten (W) absorber material with scintillating fibres inserted in the longitudinal direction. A schematic of a SpaCal module is shown in Figure 3.3b. As the light is transported directly by the scintillating fibres, no wavelength-shifting fibres need to be inserted. The insertion of scintillating fibres into the absorber material results in a non-uniform material distribution. To compensate for this, the SpaCal cells are tilted by $\pm 3°$ in x and y with respect to the beam axis as shown in Figure 3.4. This results in a more uniform material distribution in the transverse plane and a better energy resolution, despite a small loss in acceptance near $x = 0$ and $y = 0$.
For the readout, the modules are equipped with new PMTs converting the scintillation light into an electric pulse proportional to the energy of the incoming particle. The energy is measured by the ICECAL ASICs, which integrate the charge of the electric signal within the 25 ns bunch-crossing window.
In Upgrade II, the granularity on the ECAL is increased once again [87,89] with the layout shown on the right in Figure 3.2. Additional SpaCal modules with dimensions 4 cm × 4 cm using lead absorber are added, and the innermost part is equipped with tungsten modules of 1.5 cm × 1.5 cm utilising crystal scintillating fibres. In addition, all modules are equipped with a double-sided readout. For SpaCal modules, a reflective layer is inserted in each SpaCal module to achieve longitudinal segmentation of the module.
The most striking modification to the ECAL in the Upgrade II period is the readout of the modules. New *Front-End Boards* (FEBs) are produced for all $\sim 30,000$ readout channels, which split the incoming analogue signal from the PMTs and process it with two different ASICs. The measurement of the integrated charge is provided by an upgraded version of

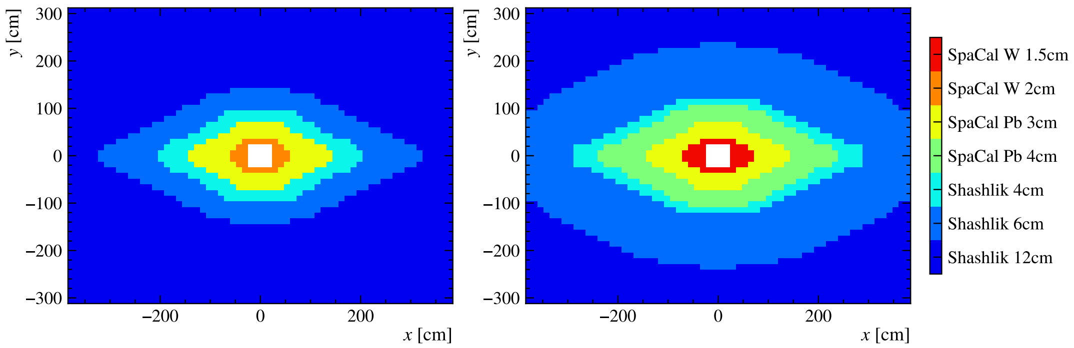


Figure 3.2: ECAL regions and cell sizes in the proposed configurations for Run 4 (left) and Run 5 (right). Adapted from Ref. [86] and [87].

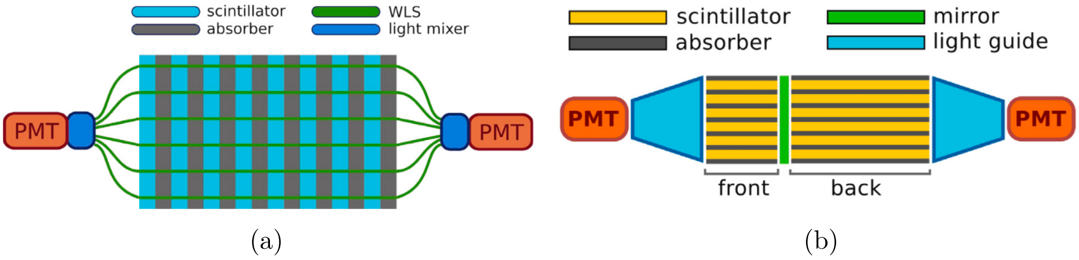


Figure 3.3: (a) Schematic of a SpaCal module and (b) shashlik module. Longitudinal segmentation and double-sided read-out are only introduced in Upgrade II. Taken from Ref. [90].

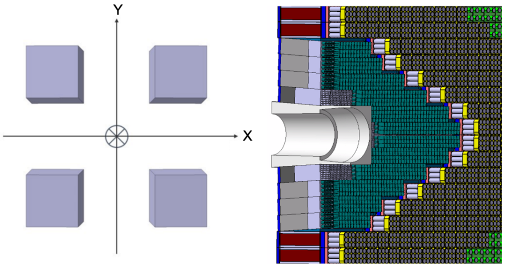


Figure 3.4: Rotation of the SpaCal modules in the inner part of the ECAL. Taken from Ref. [86].

the ICECAL ASIC, operated in a two-gain system and providing energy measurements in the dynamic range $E_T < 5\,\text{GeV}$ and $5\,\text{GeV} < E_T < 40\,\text{GeV}$. In addition, the new SPIDER ASIC is introduced, which uses waveform-sampling to provide timing information on the pulse with a resolution of $\mathcal{O}(10\,\text{ps})$. This allows to distinguish different pp collisions within the same bunch crossing and also offers the opportunity to study the pulse shape.

3.1.1 The Hadronic Calorimeter

The main purpose of the HCAL in Run 1 and Run 2 was to provide the L0 trigger decision on energetic hadrons, `L0Hadron` [72]. Since the removal of the L0 hardware trigger in the trigger architecture for Run 3 [84], the value of the HCAL is significantly diminished. Nonetheless, it is still operated in Run 3 to reconstruct the energy of hadronic showers, which is expected to improve the identification of charged particles, especially the separation of electrons from hadrons.

Looking ahead to Run 5, the higher expected occupancy would saturate the HCAL immediately. For this reason, an operation of the HCAL is not foreseen for Run 5. However, the HCAL, with its large amount of absorber material, could still passively benefit the LHCb muon system by reducing the occupancy in the muon stations. An alternative approach suggests that the performance of the muon system could be enhanced by removing the HCAL and introducing a dedicated shielding in front of the muon stations [87]. In this case, the disassembly of the HCAL and installation of the dedicated shielding requires a significant amount of work, which, due to time constraints during LS4, needs to be done already during LS3. At the time of writing this work, the decision whether the HCAL is kept in place or replaced instead is still pending.

3.2 Mapping of the signal cables

The signal cables connect the PMTs of each ECAL module to the FEBs in the electronics racks on the platform above the ECAL. A schematic showing the ECAL with the surrounding support structure and the electronics racks is shown in Figure 3.5. The signal cables are routed from the PMTs upwards to the top of the ECAL in vertical cable bands. From there, they are distributed among the different electronics racks, each of which is equipped with up to four crates that house the FEBs. The crates used for Run 4 can house up to 16 FEBs, while new crates are installed for Run 5, housing up to 21 FEBs.

The mapping of the signal cables is organised such that neighbouring PMTs are connected to the same FEB whenever possible. For cells with double-sided readout, the PMTs on the front and back sections of a cell are also mapped to the same FEB. For Upgrade II, this allows for potentially performing some pre-clustering on the front-end board, decreasing the data that needs to be sent to the event reconstruction farm.

To simplify the installation of the signal cables, cells in the same column (i.e. with the same horizontal position) are routed to the same electronics rack. The mapping also ensures that the cable length for SpaCal cells does not exceed 10 m, to avoid deteriorating the signal pulse.

To define the mapping of the signal cables, an algorithm is developed that groups signal channels connected to the same FEB simultaneously from both the inside and the outside until convergence is reached. Next, the grouped channels are distributed among the electronics racks according to their horizontal position, which ensures a simple installation of the cables. Lastly, the grouped signal channels are distributed among the crates within an electronics rack according to their vertical position on the ECAL wall. The mapping for the signal cables for the LS3 enhancements and Upgrade II is shown in Figure 3.6. The enclosed regions correspond to ECAL cells connected to the same FEB. For Upgrade II, mapping of the front- and back sections to the same FEB is implicit. The colored regions indicate in which electronics rack the respective FEB is housed and the three different shades correspond to the different crates. The distribution of the FEBs among the different crates is illustrated in Figure 3.7. For the LS3 enhancements, a total of 9,088 signal cables connect the ECAL cells to 284 FEBs in 18 crates. For Upgrade II, 30,368 signal cables connect the ECAL cells to 476 FEBs in 24 crates. An alternative projection of the signal cable mapping scheme overlaid with the different regions of the ECAL, denoting different cell sizes and technologies, is shown in Appendix A.1.

3.3 Run 3 performance studies without HCAL

To decide whether to replace the HCAL with a dedicated shield to reduce the flux in the muon stations or to keep the HCAL in place, the impact on the detector performance needs to be understood. In Run 3, the HCAL is operated in conditions similar to those expected for Run 4, and consequently, it could still be operated if kept in place. For Run 5, the decision not to operate the HCAL is definite due to the largely increased occupancy.

To understand the expected impact of not operating the HCAL in Run 4 on the detector performance, the performance is studied in the nominal conditions of Run 3. Of particular interest is the PID performance for electron and muon identification and their respective separation from hadrons, which is crucial for the flagship LHCb analyses of lepton flavor

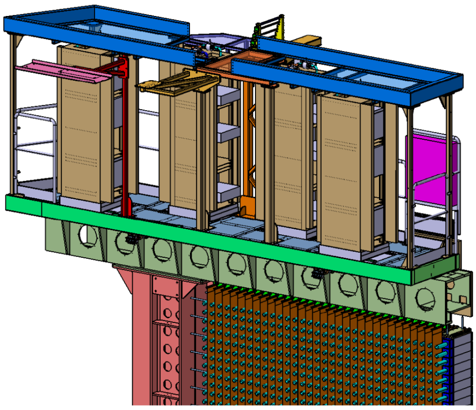


Figure 3.5: Schematic of the four electronics racks installed on top of one half of the ECAL. Taken from Ref. [91].

universality. In the past, studies on the impact of the HCAL on the PID performance in Run 4 have been conducted using simulated data. However, it is known that the PID response is not accurately reproduced in LHCb simulation. The study presented is using data recorded with the LHCb experiment in 2024 to get more reliable results on the PID performance. The data were collected in pp collisions at $\sqrt{s} = 13.6\,\text{TeV}$ with an average multiplicity of 5.3 primary interactions per bunch-crossing.

3.3.1 Charged particle PID

The most common variables for the identification of a particle species are the $\text{DLL}_{x\pi}$ variables. They are constructed as the difference in log-likelihood between two PID hypotheses on the reconstructed track, typically one of which is the pion hypothesis. The response of each subdetector of the PID system is evaluated under the two hypotheses from which a global PID response is generated.

In the following, the identification of electrons and muons and their separation from hadrons is evaluated utilising the $\text{DLL}_{e\pi}$ and $\text{DLL}_{\mu\pi}$ variables, respectively. For the identification of electrons, the $\text{DLL}_{e\pi}$ variable is defined as

$$\text{DLL}_{e\pi} = \text{DLL}_{e\pi,RICH} + \text{DLL}_{e\pi,ECAL} + \text{DLL}_{e!e,BREM} + \text{DLL}_{e\pi,HCAL} \tag{3.1}$$

where the responses of the RICH, ECAL, bremsstrahlung recovery, and HCAL systems are summed up to yield the global PID response. As bremsstrahlung is only assigned for particles reconstructed under an electron hypothesis, the bremsstrahlung recovery system only allows for separating the electron hypothesis from a non-electron hypothesis, $!e$. All other PID subsystems are evaluated using the electron- and pion hypothesis, except the muon system, which is only separating muons μ from non-muons $!\mu$ and is thus not contributing to the separation of electrons from hadrons. However, for the identification of muons, the muon system is crucial and the $\text{DLL}_{\mu\pi}$ is defined as

$$\text{DLL}_{\mu\pi} = \text{DLL}_{\mu\pi,RICH} + \text{DLL}_{\mu\pi,ECAL} + \text{DLL}_{\mu\pi,HCAL} + \texttt{isMuon} \times \text{DLL}_{\mu!\mu,MUON}. \tag{3.2}$$

The muon system also relies on the binary `isMuon` identification, which requires a number of subsequent hits in the muon stations to be one and is zero otherwise. The alternative PID response without HCAL information is emulated offline by reevaluating the global

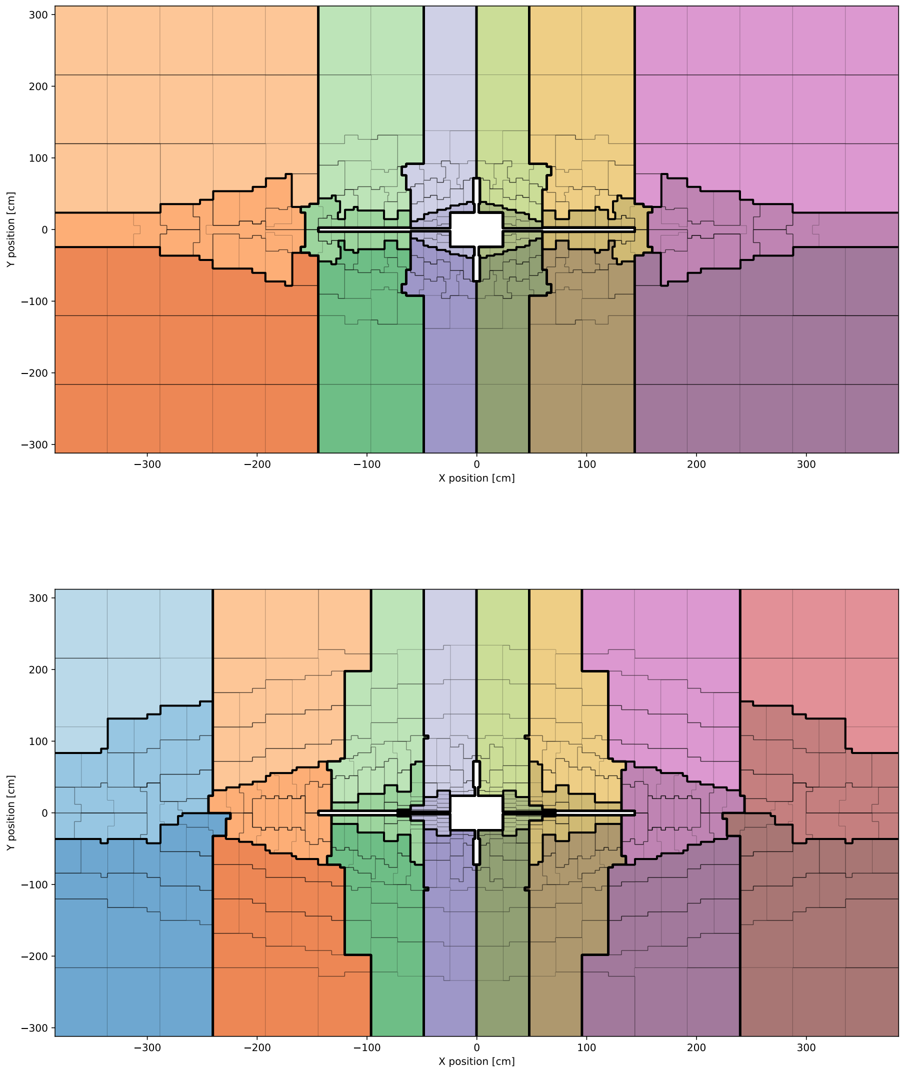
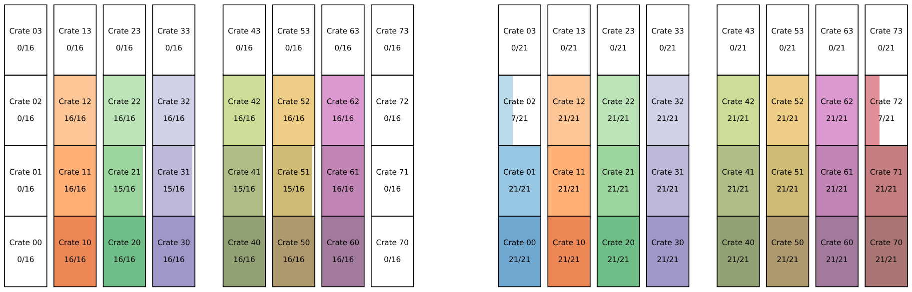


Figure 3.6: Mapping scheme of the signal cables the for LS3 enhancements (top) and Upgrade II (bottom). The regions enclosed by the black lines are mapped to the same FEB. The color and shade indicate the rack and crate the FEBs are housed in and correspond to the color scheme in Figure 3.7.



Crate 03 0/16	Crate 13 0/16	Crate 23 0/16	Crate 33 0/16	Crate 43 0/16	Crate 53 0/16	Crate 63 0/16	Crate 73 0/16
Crate 02 0/16	Crate 12 16/16	Crate 22 16/16	Crate 32 16/16	Crate 42 16/16	Crate 52 16/16	Crate 62 16/16	Crate 72 0/16
Crate 01 0/16	Crate 11 16/16	Crate 21 15/16	Crate 31 15/16	Crate 41 15/16	Crate 51 15/16	Crate 61 16/16	Crate 71 0/16
Crate 00 0/16	Crate 10 16/16	Crate 20 16/16	Crate 30 16/16	Crate 40 16/16	Crate 50 16/16	Crate 60 16/16	Crate 70 0/16

Crate 03 0/21	Crate 13 0/21	Crate 23 0/21	Crate 33 0/21	Crate 43 0/21	Crate 53 0/21	Crate 63 0/21	Crate 73 0/21
Crate 02 7/21	Crate 12 21/21	Crate 22 21/21	Crate 32 21/21	Crate 42 21/21	Crate 52 21/21	Crate 62 21/21	Crate 72 7/21
Crate 01 21/21	Crate 11 21/21	Crate 21 21/21	Crate 31 21/21	Crate 41 21/21	Crate 51 21/21	Crate 61 21/21	Crate 71 21/21
Crate 00 21/21	Crate 10 21/21	Crate 20 21/21	Crate 30 21/21	Crate 40 21/21	Crate 50 21/21	Crate 60 21/21	Crate 70 21/21

Figure 3.7: Illustration of the rack layout for the LS3 enhancements (left) and Upgrade II (right). The number of FEBs installed in each crate is indicated and the color coding matches the ECAL cells shown in Figure 3.6.

DLL, forcing DLL_{HCAL} to zero. In the following, the performance of the PID system under the nominal and alternative DLL is studied for the identification of electrons and muons.

3.3.2 Deriving PID efficiencies in data

The identification of electrons and muons is studied in $B^+ \to K^+ J/\psi$ decays with $J/\psi \to e^+e^-$ and $J/\psi \to \mu^+\mu^-$, respectively. A *tag-and-probe* method is used, in which one of the charged leptons is selected as *tag* with tight PID requirements imposed. The respective other charged lepton is used as the *probe*, selected without PID requirements. Imposing requirements on the DLL variables of the probe track calculated with and without HCAL information, the PID performance in both cases is studied.

The misidentification (misID) of charged hadrons as electrons and muons is studied using $D^{*+} \to D^0\pi^+$ decays with $D^0 \to K^-\pi^+$. Candidates are selected without PID requirements on one of the D^0 daughters, based on which the misidentification of $K^+ \to \ell^+$ and $\pi^+ \to \ell^+$ is studied for $\ell^+ = e^+, \mu^+$.

The PID efficiency is calculated in data by splitting the reconstructed candidates into a *pass category* with events passing the PID requirement on the probe track, and the *fail category* with candidates failing to pass the respective requirement. From the number of candidates in each category, $N_{pass,fail}$, the PID efficiency is determined as

$$\varepsilon = \frac{N_{pass}}{N_{pass} + N_{fail}}. \tag{3.3}$$

The number of reconstructed candidates in the pass- and fail category is determined from an extended maximum likelihood fit to the invariant mass distribution. The *Probability Density Function* (PDF) to model the invariant mass distribution in the two categories $\mathcal{P}_{pass,fail}$ is composed of a signal and one or more background PDFs multiplied by the respective number of candidates

$$\begin{aligned}\mathcal{P}_{pass} &= \quad\;\; \varepsilon \times N_{sig} \times \mathcal{P}_{sig} + \sum_{bkg} N_{bkg,pass} \times \mathcal{P}_{bkg},\\ \mathcal{P}_{fail} &= (1-\varepsilon) \times N_{sig} \times \mathcal{P}_{sig} + \sum_{bkg} N_{bkg,fail} \times \mathcal{P}_{bkg}.\end{aligned} \tag{3.4}$$

Performing a simultaneous fit to the two categories, the PID efficiency, ε, is directly inferred. As the impact of HCAL on the PID efficiency might show a momentum dependence, the PID efficiencies are determined in four bins of the probe track momentum $p_{probe} \in [0, 15, 30, 50, 300]\,\text{GeV}$.

Muon PID efficiency from $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$

The muon PID efficiency, $\varepsilon_{\mu\to\mu}$, is derived on $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$ candidates. The number of correctly reconstructed candidates is determined from a fit to the invariant mass, $m_{J/\psi}(K^+\ell^+\ell^-)$, where the di-electron is constrained to the known mass of the J/ψ resonance. Correctly identified $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$ candidates are modelled using a

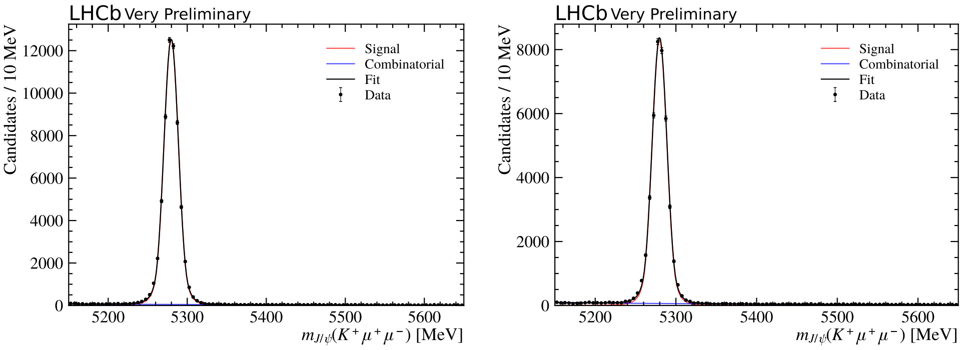


Figure 3.8: Fit to the invariant mass distribution for $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$ candidates in the pass (left) and fail (right) category for a probe muon with momentum of $p_{probe} \in [30, 50]$ GeV.

double-sided Crystal Ball function, defined as:

$$f(x;\mu,\sigma,\alpha_L,n_L,\alpha_R,n_R) = \begin{cases} A_L \cdot \left(B_L - \frac{x-\mu}{\sigma}\right)^{-n_L}, & \text{for } \frac{x-\mu}{\sigma} < -\alpha_L \\ \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), & \text{for } -\alpha_L \leqslant \frac{x-\mu}{\sigma} \leqslant \alpha_R \\ A_R \cdot \left(B_R + \frac{x-\mu}{\sigma}\right)^{-n_R}, & \text{for } \frac{x-\mu}{\sigma} > \alpha_R \end{cases} \tag{3.5}$$

with the tail parameters $\alpha_{L,R} > 0$ and $n_{L,R} > 0$ and the factors $A_{L,R}$ and $B_{L,R}$ are defined as

$$\begin{aligned} A_{L/R} &= \left(\frac{n_{L/R}}{\left|\alpha_{L/R}\right|}\right)^{n}_{L/R} \cdot \exp\left(-\frac{\left|\alpha_{L/R}\right|^2}{2}\right), \\ B_{L/R} &= \frac{n_{L/R}}{\left|\alpha_{L/R}\right|} - \left|\alpha_{L/R}\right|. \end{aligned} \tag{3.6}$$

The parameters $\alpha_{L,R}$ and $n_{L,R}$ are fixed from a fit to simulated $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$ candidates. *Combinatorial background* candidates, in which at least one track is not originating from the same B meson decay, are modelled with an exponential.
An example for the simultaneous fit to the pass and fail categories to derive the muon PID efficiency is shown in Figure 3.8.

Electron PID efficiency from $B^+ \to K^+ J/\psi(\to e^+e^-)$

The electron PID efficiency, $\varepsilon_{e\to e}$, is derived from $B^+ \to K^+ J/\psi(\to e^+e^-)$ candidates, separately for probe electrons with and without bremsstrahlung assigned. The fit setup largely follows the same strategy as for the decay to muons. However, an additional background component from partially reconstructed $B^0 \to K^+[\pi^-] J/\psi(\to e^+e^-)$ candidates needs to be considered. It is modelled by a non-parametric *Kernel Density Estimation* of simulated $B^0 \to K^{*0}(\to K^+\pi^-) J/\psi(\to e^+e^-)$ candidates. An example for the simultaneous fit to the pass and fail categories to derive the electron PID efficiency is shown in Figure 3.9

Misidentification from $D^0 \to K^-\pi^+$

The efficiencies for hadron to lepton misID, $h^+ \to \ell^+$, are derived from $D^0 \to K^-\pi^+$ candidates separately for $h^+ = K^+, \pi^+$ and $\ell^+ = e^+, \mu^+$. In the fit to the reconstructed

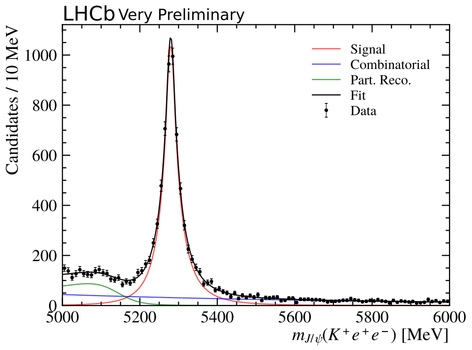


Figure 3.9: Fit to the invariant mass distribution for $B^+ \to K^+ J/\psi(\to e^+e^-)$ candidates in the pass (left) and fail (right) category for a probe electron without a bremsstrahlung photon attached and momentum of $p_{probe} \in [30, 50]$ GeV.

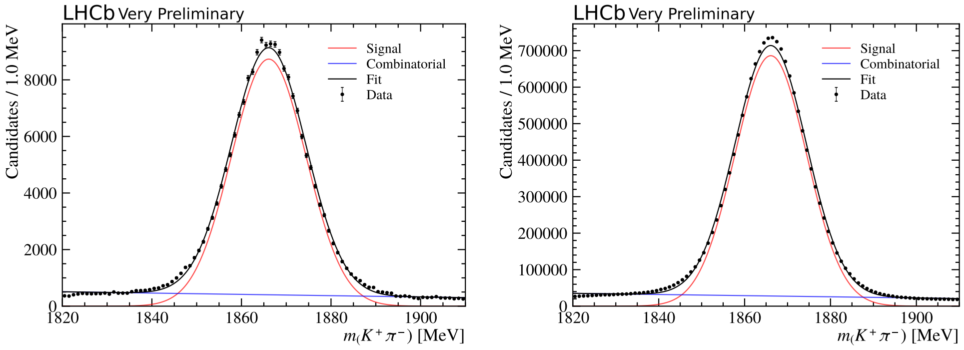


Figure 3.10: Fit to the invariant mass distribution for $D^0 \to K^-\pi^+$ candidates with the probe pion passing (left) and failing (right) the electron PID requirement without a bremsstrahlung photon attached and momentum of $p_{probe} \in [30, 50]$ GeV.

invariant mass, $m(K^+\pi^-)$, the component from $D^0 \to K^-\pi^+$ candidates is modelled by a modified Gaussian with an exponential tail on the left-hand side. The combinatorial background is modelled from a first-order polynomial. An example for the simultaneous fit to the pass and fail categories to derive the $\pi^+ \to e^+$ misID efficiency is shown in Figure 3.10.

3.3.3 PID performance without HCAL

The PID and misID efficiencies are evaluated for a wide range of cuts on the PID response, with and without information from the HCAL. The derived muon PID efficiencies are plotted against the respective $\pi^+ \to \mu^+$ misidentification efficiencies in Figure 3.11. For electrons, the respective efficiencies are shown in Figure 3.12 and 3.13 in the two bremsstrahlung categories. The comparisons against $K^+ \to \ell^+$ are shown in Appendix A.1.

For the identification of muons, the PID system shows no significant loss of performance when HCAL information is removed. For electrons, a slight degradation of the PID performance can be observed in the separation from hadrons for high-momentum tracks and tight PID requirements, especially when no bremsstrahlung is associated with the particle track.

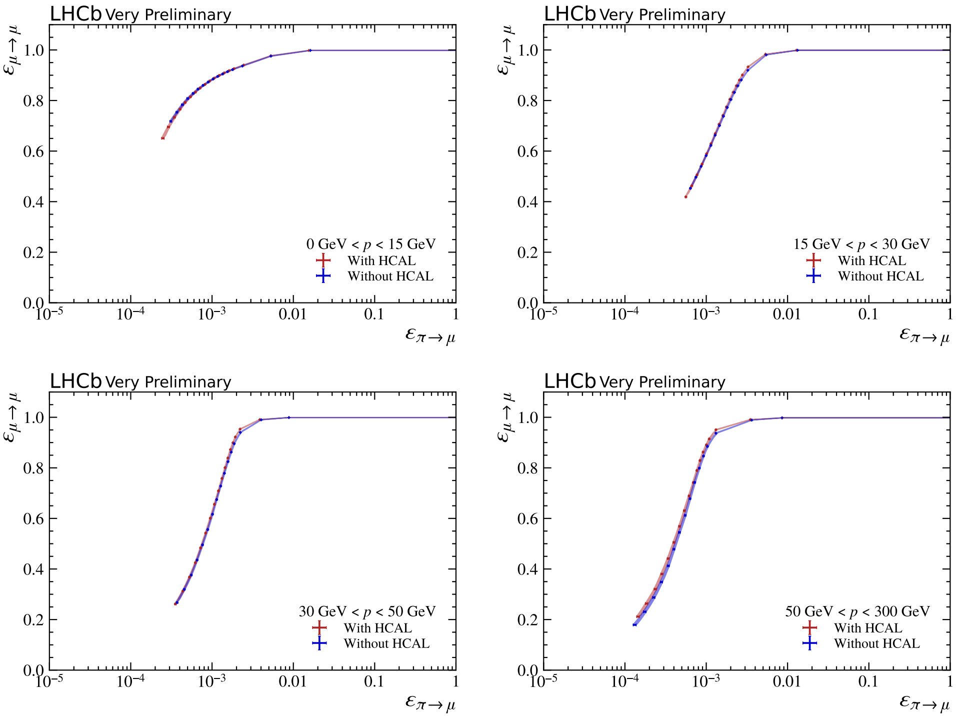


Figure 3.11: Performance of the PID system with HCAL (red) and without HCAL (blue) for muon identification and $\pi^+ \to \mu^+$ misidentification in four bins of momentum.

The presence of electron PID requirements in HLT2 lines to select beauty-hadron decays to electrons directly relates the loss in PID performance to a loss in selection efficiency. This effect is evaluated on $B^+ \to K^+ J/\psi(\to e^+e^-)$ candidates with a standard cut for the electron PID selection in Run 3 of $\mathrm{DLL}_{e\pi} > 2$. Assuming the trigger lines for Run 4 are tuned such that the data are selected with the same level of background from $h^+ \to e^+$ misidentification, the degraded PID performance results in a loss of selection efficiency of about 1%.

3.3.4 Concluding remarks

The expected PID performance for electrons and muons in Run 4 is studied using data recorded in 2024, in conditions similar to those anticipated for Run 4. Removing the readout of the HCAL for Run 4 is expected to have a negligible impact on muon identification. For the identification of electrons, a slight degradation of the PID performance is observed for high-momentum tracks and tight cuts on the PID response. The impact on the selection efficiency is evaluated to be of the order of 1% for $B^+ \to K^+ J/\psi(\to e^+e^-)$ candidates. This suggests a generally small impact on the selection of electrons from b-hadron decays selected with typical PID requirements. It is noted that these results do not account for the expected improvements of the PID system resulting from the LS3 enhancements, with which the overall PID performance is expected to improve for Run 4.

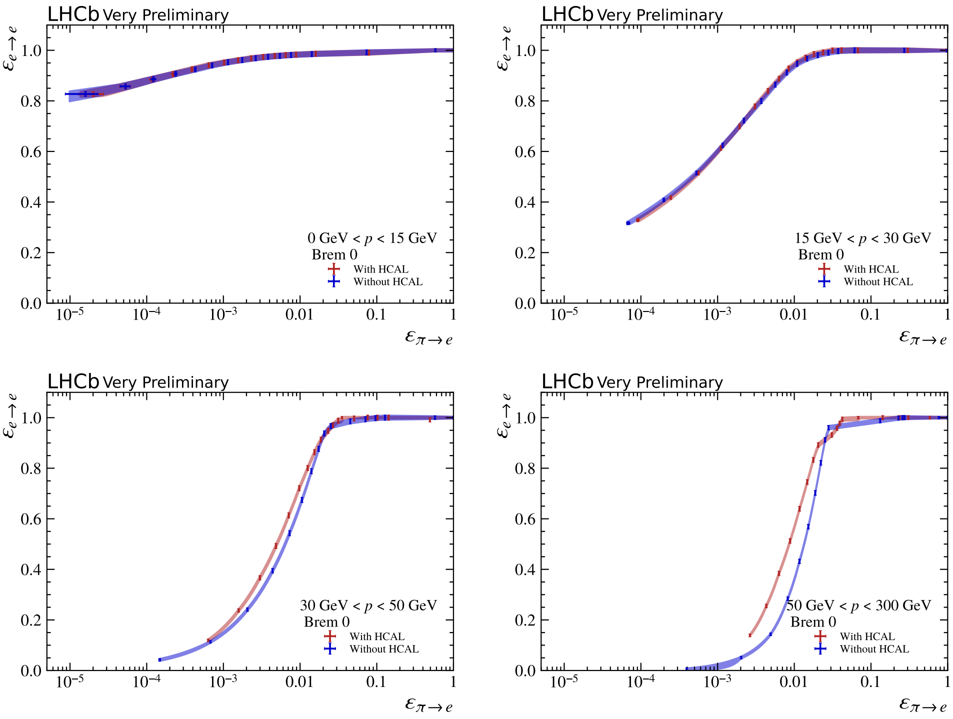


Figure 3.12: Performance of the PID system with HCAL (red) and without HCAL (blue) for electron identification and $\pi^+ \to e^+$ misidentification in four bins of momentum and without a bremsstrahlung photon attached to the probe track.

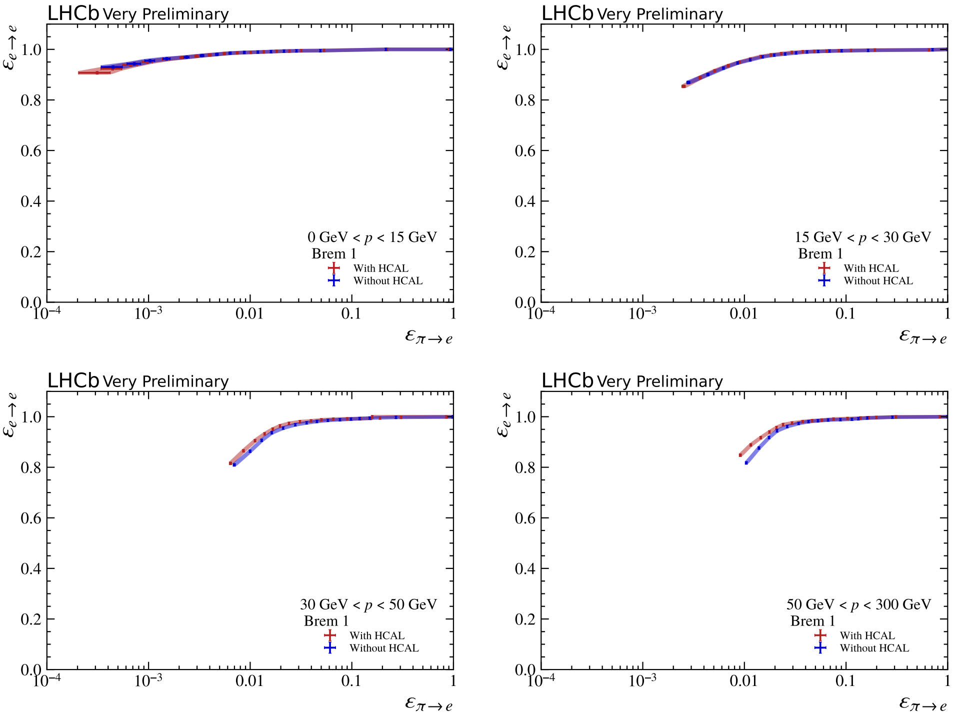


Figure 3.13: Performance of the PID system with HCAL (red) and without HCAL (blue) for electron identification and $\pi^+ \to e^+$ misidentification in four bins of momentum and with a bremsstrahlung photon attached to the probe track.

Part II

Search for radiative leptonic $B^{+}_{(c)} \to \mu^{+}\nu_{\mu}\gamma$ decays

4

Analysis strategy

In this part, the author's original work on the search for $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays is presented. As discussed in Section 1.3, the decay $B^+ \to \ell^+\nu_\ell\gamma$ is considered the golden mode to probe the B meson LCDA and measure its first inverse moment, λ_B, from experimental data. The decay $B^+ \to \ell^+\nu_\ell\gamma$ has never been observed with a current best upper limit on the branching fraction of $\mathcal{B}(B^+ \to \ell^+\nu_\ell\gamma) < 3.0 \times 10^{-6}$ from the Belle experiment, combining the results for $\ell^+ = e^+\mu^+$ [1]. The radiative leptonic decay of the B_c^+ meson, $B_c^+ \to \ell^+\nu_\ell\gamma$, has never been searched for. Consequently, the presented work manifests the first search for $B^+ \to \ell^+\nu_\ell\gamma$ decays at the LHCb experiment and the first ever search for $B_c^+ \to \ell^+\nu_\ell\gamma$ decays.

The analysis of $B^+ \to \ell^+\nu_\ell\gamma$ decays has been exclusively attributed to experiments at e^+e^- colliders and deemed impossible at the LHCb experiment. The reason for this is twofold. Firstly, the event selection for B physics at LHCb is specifically designed to select B candidates with displaced decay vertices. However, in a decay $B^+ \to \ell^+\nu_\ell\gamma$, only a single charged track is present, which makes it impossible to reconstruct a B decay vertex and perform a standard event selection. Secondly, the momentum of the neutrino cannot be reconstructed, nor can it be inferred from the kinematics with which the B mesons are produced, as they are unknown at a *pp* collider. The strategy outlined for the presented analysis allows to overcome both of these problems.

Firstly, candidates are reconstructed using $\gamma \to e^+e^-$ photon conversions in the material of the VELO. This results in multiple charged tracks pointing back to the displaced B decay vertex[1], allowing its position to be reconstructed. Photons reconstructed from conversions with two electron long tracks are denoted as γ_{ee} throughout the document.

Secondly, a fraction of the neutrino momentum can be recovered by correcting for the momentum imbalance in the decay perpendicular to the B flight direction. This allows to define a *corrected mass*, m_{corr}, as

$$m_{corr} = \sqrt{m(\ell^+\gamma_{ee})^2 + |p_\perp|^2} + |p_\perp|, \tag{4.1}$$

where $m(\ell^+\gamma_{ee})$ is the invariant visible mass of the charged lepton and the converted photon, and $|p_\perp|$ is the absolute momentum imbalance perpendicular to the B meson flight direction illustrated in Figure 4.1.

---

[1] While the photon might travel some distance before converting in the VELO, extrapolation of the electron tracks points back to the B decay vertex as they are produced with zero opening angle.

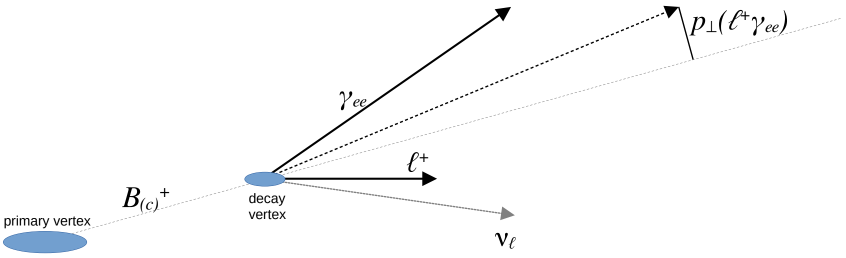


Figure 4.1: Illustration of the momentum $p_\perp$ in a semileptonic $B_{(c)}^+ \to \ell^+\nu_\ell\gamma_{ee}$ decay.

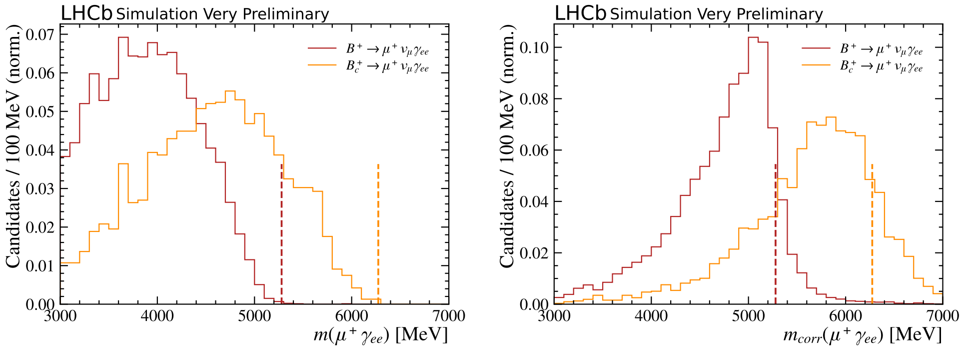


Figure 4.2: Reconstructed invariant mass (left) and corrected mass (right) for simulated $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$ candidates. The vertical dashed lines indicate the true mass of the B^+ and B_c^+ mesons, respectively.

Reconstructing the B decay vertex using photon conversions in the VELO material, the B flight direction is determined from the reconstructed positions of the primary vertex and the B decay vertex. The improvement on the mass resolution for $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$ decays using the corrected mass is shown in Figure 4.2.

With these considerations, the search for $B_{(c)}^+ \to \ell^+\nu_\ell\gamma$ decays becomes conceptually possible at LHCb. In the following, the decays are reconstructed solely in the muonic final state, which offers superior momentum resolution and PID performance compared to electrons. The number of observed $B^+ \to \mu^+\nu_\mu\gamma$ and $B_c^+ \to \mu^+\nu_\mu\gamma$ candidates is determined simultaneously from a single binned fit to the corrected mass distribution in data.

As discussed in Sections 1.2 and 1.3, the QCD factorisation approach for $B^+ \to \ell^+\nu_\ell\gamma$ decays is only valid for energetic photons[2]. For this reason, the analysis is performed for $E_\gamma^* > 1.0\,\text{GeV}$, which can be considered safe for factorisation to hold [33] and allows for direct comparison with the result from Belle [1]. Imposing a tighter requirement on E_γ^* could allow for more reliable theory predictions, but would also decrease the chances of observing the decay for the first time.

The quantity E_γ^* is calculated using an approximate B rest-frame, which is defined using the momentum of the reconstructed decay products and correcting for the missing neutrino momentum as introduced in [92]. More details can be found in Appendix B.1. Using the same reference rest frame, the muon energy in the (approximate) B meson rest-frame, E_μ^*, is defined.

Another quantity that is typically used in the discussion of the decay kinematics is the decay angle, θ_W. It is defined in the (approximate) $\mu^+\nu_\mu$ rest-frame as the angle between the muon momentum and the inverted B momentum and is illustrated in Figure 4.3.

---

[2] The photon needs to be energetic in the B rest-frame, which is marked by an asterisk.

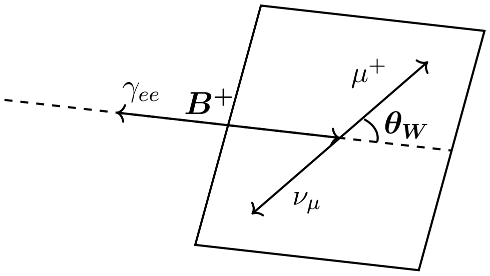


Figure 4.3: Illustration of the decay angle θ_W defined in a $B^+ \to \mu^+\nu_\mu\gamma_{ee}$ decay.

Due to uncertainties on the luminosity and especially the $\sigma_{pp\to b\bar{b}}$ cross-section as well as the fragmentation fractions f_i, measurements of absolute branching fractions suffer from large uncertainties at LHCb. Instead, the branching fractions are obtained relative to a *normalisation channel* with well-measured branching fraction. The presented search for $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays is performed using the normalisation channel $B^0 \to K^{*0}\gamma$ reconstructed from photon conversions. Besides a very clean signal selection, it has the advantage that systematic uncertainties related to the photon conversion cancel in the ratio. This includes, in particular, the uncertainties related to the description of the detector material for the photon conversion efficiency and uncertainties linked to the electron reconstruction and identification.

Details on the selection of candidates are provided in Chapter 5 along with a brief overview of the simulation samples used in the analysis and the corresponding corrections to simulated events. The fit of the reconstructed corrected mass distribution in data is established in Chapter 6. At the heart of this lies the modelling of background candidates involving the decay of light neutral mesons $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ in Section 6.3. A novel method is developed to derive the corrected mass templates for such backgrounds using a data-driven approach with additional inputs from simulation. The normalisation channel $B^0 \to K^{*0}\gamma$ is analysed in Chapter 7 and the number of observed $B^0 \to K^{*0}\gamma$ candidates is determined from a fit to the reconstructed invariant mass. In Chapter 8, the efficiencies for the selection of $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ and $B^0 \to K^{*0}\gamma$ candidates are calculated. Systematic uncertainties are studied in Chapter 9. The still *blinded* results are presented in Section 10 as the analysis is currently under internal review.

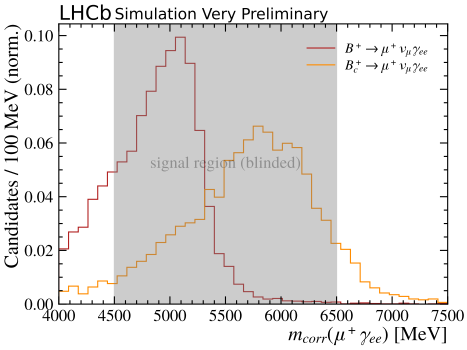


Figure 4.4: Reconstructed corrected mass of simulated $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ candidates with the blinded mass range from 4500 MeV to 6500 MeV shown in grey.

Blinding strategy

The presented analysis is performed blinded by excluding candidates from the mass range $4500\,\text{MeV} < m_{corr} < 6500\,\text{MeV}$ in the data. This is to prevent creating fake signals from statistical fluctuations in data. The blinded mass range is chosen to remove the signal components of both B^+ and B_c^+ as illustrated in Figure 4.4. For the B^+ (B_c^+) signal decay, about 72% (86%) of the simulated candidates passing the event selection fall into the blinded region. The majority of the remaining candidates populate the lower mass sideband $m_{corr} < 4500\,\text{MeV}$, which is expected to be dominated by background. The small fraction of $B_c^+ \to \mu^+\nu_\mu\gamma$ candidates populating the upper mass sideband above 6500 MeV is negligible, as the total number of expected signal candidates for this decay passing the selection is of the order of one.

5

Candidate selection

In Section 5.1, the data and simulation samples used in the search for $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays are discussed. The multi-stage selection of candidates is detailed in the following sections. First, several essential decay variables are defined in Section 5.2. The trigger selection is discussed in Section 5.3 and details on the central event filtering are provided in Section 5.4. In Section 5.5, the one-dimensional selection requirements are defined, followed by a multivariate analysis in Section 5.6. Data-driven corrections to the simulated samples are derived in Section 5.7.

5.1 Data and simulation samples

5.1.1 Data samples

The search for $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays is performed on data recorded with the LHCb experiment in *pp* collisions between 2016 and 2018 at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ corresponding to an integrated luminosity of $5.4\,\text{fb}^{-1}$. The integrated luminosities per year of data-taking are given in Table 5.1. Data recorded in earlier years is not considered, as the selection of $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ candidates in the central event filtering (see Section 5.4) is only available for the years 2016 to 2018.

5.1.2 Simulation samples

Throughout the analysis *Monte Carlo* (MC) simulation samples of various particle decays are used to optimise the selection of candidates, determine selection efficiencies and derive corrected mass templates. All simulation samples are generated using the data-flow described in Section 2.6. The simulation samples are generated separately for each year and magnet polarity to emulate the conditions during data-taking as closely as possible.

Year	2016	2017	2018
$\mathcal{L}_{int}$ [fb^{-1}]	1.6	1.7	2.1

Table 5.1: Integrated luminosity recorded with the LHCb experiment in *pp* collisions per year of data-taking.

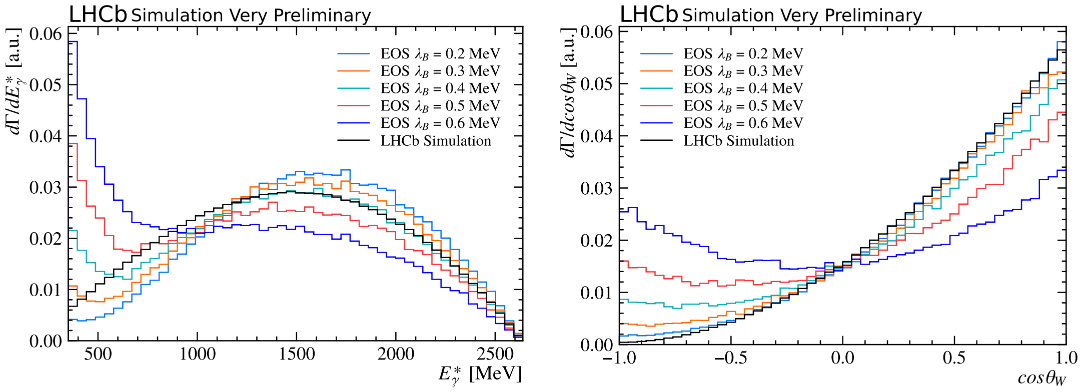


Figure 5.1: Comparison of the $B^+ \to \mu^+\nu_\mu\gamma$ model from Ref. [33] as implemented in EOS [93] for different values of λ_B and the leading-order model from Ref. [25].

For simulation samples involving photon conversions, an additional event filter is applied. After simulating the propagation and detector response for the signal decay, the signal photon is checked to have converted in the VELO material before simulating the detector response of the full event; otherwise, the event is discarded. This allows for the simulation of many more signal candidates with photon conversions using the same computing power. The impact of the event filter on the reconstruction efficiency is studied in Chapter 8.

Simulation of $B^+ \to \mu^+\nu_\mu\gamma$

Simulation samples of $B^+ \to \mu^+\nu_\mu\gamma$ decays are generated using an amplitude model based on the leading-order calculation from Ref. [25] that is implemented in the LHCb simulation framework. The implementation assumes an exponential model of the B meson LCDA with vanishing first logarithmic moment, $\sigma_1 = 0$, and the model parameters are set to $E^*_\gamma > 0.2\,\text{GeV}$, $\lambda_B = 333\,\text{MeV}$.
To check the validity of the leading-order model, it is compared to the state-of-the-art model from Ref. [33], which is implemented in the software EOS [93]. Small simulation samples are generated in EOS, setting the input parameters to the values in Table 1 of Ref. [33] and varying the value of λ_B within the loosely constrained region $200\,\text{MeV} < \lambda_B < 600\,\text{MeV}$. In Figure 5.1, the distributions in E^*_γ and $\cos\theta_W$ (see Figure 4.3 for the definition of the decay angle) for different values of λ_B are compared to the simulation from the leading-order model. Due to a non-uniform selection efficiency in E^*_γ and $\cos\theta_W$, this introduces an inherent model dependence. To quantify the model dependence in the study of the systematic uncertainties in Chapter 9, an additional simulation sample is generated using a phase-space model, which is flat in E^*_γ and $\cos\theta_W$.

Simulation of $B_c^+ \to \mu^+\nu_\mu\gamma$

For the simulation of $B_c^+ \to \mu^+\nu_\mu\gamma$ decays, the available predictions [2, 37, 40–42] are not implemented in the LHCb simulation framework. However, the leading-order calculation for the $B_c^+ \to \mu^+\nu_\mu\gamma$ decay from Ref. [40] predicts similar decay kinematics to the leading-order model for the B^+ decay from Ref. [25]. Hence, the same decay model is used for the B^+ and B_c^+ decays, except that the phase-space is enlarged for the respective B_c^+

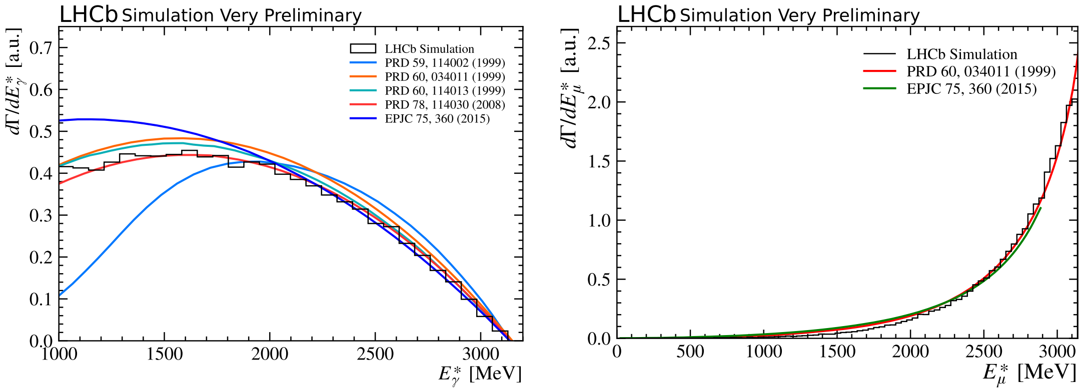


Figure 5.2: Predicted branching fraction of $B_c^+ \to \mu^+\nu_\mu\gamma$ decays for different theory models and the leading-order model for the decay $B^+ \to \mu^+\nu_\mu\gamma$ decay [25] applied to the decay of B_c^+ mesons differential in E_γ^* (left) and E_μ^* (right).

decay. A comparison of the decay kinematics for the various models for $B_c^+ \to \mu^+\nu_\mu\gamma$ decays and the leading-order model used in the simulation is shown in Figure 5.2. A separate simulation sample is generated using a phase-space model to evaluate the model dependence of the result.

Additional simulation samples

Additional simulation samples are generated for prominent background decays, such as $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ and $B^+ \to \mu^+\nu_\mu\eta(\to \gamma_{ee}\gamma)$ as well as an inclusive sample of $B^{+,0} \to \mu^+\nu_\mu X_u(\to X\gamma_{ee})$ decays in which the remaining resonant and non-resonant semileptonic $b \to u$ decays are simulated and X represents any number of additional particles. In addition simulation samples for the decays $B_{(c)}^+ \to \tau^+(\to \mu^+\nu_\mu\overline{\nu}_\tau)\nu_\tau\gamma$ are generated to study their contribution explicitly. For the analysis of the normalisation channel $B^0 \to K^{*0}\gamma$, the already existing simulation samples from the analysis in Ref. [94] of $B^0 \to K^{*0}(\to K^+\pi^-)\gamma_{ee}$ decays and some prominent background decays are used.

Truth matching

In all simulation samples, the reconstructed candidates are checked to be correctly identified. When simulating the detector response for a particle, the generated hits are flagged by an identification number (ID) linked to the respective particle, which is propagated throughout the event reconstruction. In all reconstructed simulation samples, this ID is checked for all particles used to build the B candidates to ensure the particles are correctly reconstructed and identified, and they match the simulated decay chain.

5.2 Definition of decay variables

The selection of candidates at LHCb is performed by imposing conditions on decay variables, also referred to as *cuts*. Many of these conditions are imposed upon geometric properties of recorded particle trajectories or reconstructed decays. For this purpose, several common decay variables characterising the decay geometry are defined in the following.

Polar coordinate system
From the rotational symmetry of the $b\bar{b}$-production at LHCb around the beam axis, it is convenient to define a polar coordinate system for the LHCb experiment. The reference point of this coordinate system is at $z = 0$ along the beam axis, which is the approximate position of the primary vertex. The polar angle, θ, is defined as the angle with respect to the beam axis, z, which is pointing downstream from the VELO towards the muon stations. The azimuthal angle, ϕ, is defined in the plane perpendicular to the beam axis and is zero along the positive y-axis.

Impact parameter
The *Impact Parameter* (IP) of a particle track is the distance of closest approach of the extrapolated particle track to the PV. Due to the uncertainty on the extrapolation of the particle track and the position of the PV, it is convenient to define the χ^2_{IP}. It is constructed from the difference in vertex fit quality between the PV reconstructed with and without the track under consideration. More intuitively, it behaves approximately as

$$\chi^2_{IP} = \left(\frac{IP}{\sigma_{IP}}\right)^2 , \tag{5.1}$$

where σ_{IP} is the experimental uncertainty on the IP.

Distance of closest approach (DOCA)
The *Distance Of Closest Approach* (DOCA) is defined similarly as the smallest geometric distance between the extrapolation of two particle tracks.

Flight distance
The spatial separation between a particle's origin vertex and its decay vertex is defined as the *Flight Distance* (FD). Due to the finite vertex resolution, it is subject to an experimental uncertainty, σ_{FD}, from which the significance of the vertex separation is defined as

$$\chi^2_{FD} = \left(\frac{FD}{\sigma_{FD}}\right)^2 . \tag{5.2}$$

Direction angle (DIRA)
For particles with a distinct origin vertex and decay vertex, for example a B meson, the flight direction can be inferred by connecting the two vertices. Then, the *Direction Angle* (DIRA) is defined as the cosine of the angle between the flight direction and the reconstructed momentum vector.

5.3 Trigger selection

The LHCb experiment used a multi-stage trigger system to reduce the event rate from 40 MHz to $\mathcal{O}(1\,\text{kHz})$ that is saved to disk for physics analysis as described in Section 2.5. At each trigger stage, a subset of trigger lines is selected from which the requirements of at least one line must be met. In this analysis, only events are retained in which a particle of the signal candidate fulfils the trigger requirements referred to as *Triggered On Signal* (TOS).

Based on the L0 trigger decision, the data is split into two mutually exclusive L0 trigger categories, defined as:

- **L0E**: At least one of the electrons belonging to the signal candidate passes the `L0Electron` decision, which requires an energy deposit in the calorimeter system with sufficiently large transverse energy. The threshold on the transverse energy is about 2.5 GeV, but the exact value varied throughout the data-taking period.
- **L0M**: None of the electrons belonging to the signal candidate passes the `L0Electron` decision and the muon track of the signal candidate passes the `L0Muon` decision. It requires a single muon track with large transverse momentum in the muon stations. The threshold on the transverse momentum is about 1.4 GeV, but the exact value varied throughout the data-taking period.

For the $B^0 \to K^{*0}\gamma$ normalisation channel, candidates are selected exclusively from the L0E category.
Candidates from the L0E category are required to pass the selections of the `Hlt1TrackMVA` or `Hlt1TwoTrackMVA` lines. The former requires a track with large transverse momentum originating from a displaced vertex. The latter uses multivariate analysis to identify two charged tracks originating from a common displaced decay vertex. Candidates in the L0M category may also fulfil the criteria of the `Hlt1TrackMuon` line, selecting muons with transverse momentum and significant displacement from the primary vertex.
At the HLT2 level, multivariate classifiers are used to select candidates based on a decay topology that is consistent with a two- or three-body decay from a displaced vertex. For candidates from the L0M category, additional classifiers are used to specifically select two- and three-body decay topologies, including a final state muon.

5.4 Central event filtering

Data that has passed the trigger selection is processed by a central event filtering, referred to as *stripping*, before being made available for physics analysis. A large number of selection lines have been developed for specific physics analyses, which select candidates based on kinematics, reconstruction quality and decay topology.
For this analysis, candidates are selected from a dedicated stripping line designed for selecting $B^+ \to \mu^+\nu_\mu e^+e^-$ decays. In the early stages of this analysis, it was found that the existing selection suffered from double-counting of bremsstrahlung photons when combining the two electrons. This was corrected for the 2023 re-stripping campaign, in which all data recorded with the LHCb experiment from 2016 to 2018 were reprocessed. The data used in this analysis were processed after these modifications were applied. The selection requirements are summarised in Table 5.2.
The B candidate is reconstructed from three long tracks with track segments in all tracking stations. Good quality tracks are selected by requiring a small χ^2 of the track fit divided by the degrees of freedom (ndof). In addition, the probability of the track to be a fake track (*ghost track*) is required to be below 0.35 and the extrapolation of the tracks must not point back too close to the PV.
Muon candidates must satisfy the `isMuon` requirement and the likelihood of the particle

Candidate	**Selection**
all tracks	`isLong` $\chi^2/\text{ndof} < 3$ GhostProb < 0.35
μ^+	$\chi^2_{IP} > 9$ $\text{DLL}_{\mu\pi} > 0$ (data only) $\text{DLL}_{\mu\text{K}} > 0$ (data only)
$e^\pm$	$\chi^2_{IP} > 25$ $p_T > 200\,\text{MeV}$ $\text{DLL}_{\text{e}\pi} > 2$ (data only) $\text{DLL}_{\text{eK}} > 0$ (data only)
B^+	$m_{corr}(\mu^+e^+e^-) \in [2500, 10000]\,\text{MeV}$ $m(\mu^+e^+e^-) < 7500\,\text{MeV}$ `DIRA` > 0.99 $p_T > 2000\,\text{MeV}$ $\chi^2_{FD} > 30$ $\chi^2_{\text{Vertex}}/\text{ndof} < 4$

Table 5.2: Stripping requirements for the selection of $B^+ \to \mu^+\nu_\mu e^+e^-$ decays with the modification for the processing of simulation data indicated.

being a muon must exceed the likelihood of being a pion or kaon. Electron candidates are required to have a transverse momentum exceeding 200 MeV and the likelihood of being an electron must be larger than the likelihood of being a kaon and significantly larger than being a pion.

To select three leptons originating from the same secondary B vertex, a good quality of the vertex fit and significant displacement from the PV are required. The transverse momentum of the resulting B candidate must exceed 2 GeV and a lower limit on the DIRA ensures the B momentum is well aligned with its flight direction. Lastly, loose cuts on the three-particle invariant mass and the corrected mass of the B are imposed.

Candidates of the normalisation channel $B^0 \to K^{*0}\gamma$ are selected from a stripping line developed specifically for the selection of $B \to X\gamma(\to e^+e^-)$. The selection criteria are listed in Appendix B.2. The selection of (di-)electrons and the requirements on the B candidate are similar to those in the selection of the signal channel. To ensure a better cancellation of systematic uncertainties linked to the reconstruction and selection of candidates between the signal and normalisation channel, the selection criteria are further aligned in the following selection steps.

Due to known imperfections of the PID response in the LHCb simulation, the PID requirements are removed for the processing of simulation samples and the efficiency of the PID selection is emulated from dedicated data control samples (see Section 5.7.2).

5.5 One-dimensional selection cuts

After candidates have passed the trigger and stripping selection, further selection requirements are imposed to suppress background candidates, while retaining a high signal efficiency. This stage of the selection is also used to align the selections of $B^+_{(c)} \to \mu^+\nu_\mu\gamma$

signal candidates and $B^0 \to K^{*0}\gamma$ normalisation candidates, in particular, the selection of the photon conversion.

Acceptance

Due to the high occupancy, the innermost cells of the ECAL are not equipped with readout electronics. To ensure the electrons originating from the photon conversion are found in the active region of the ECAL, a cut on the cluster position in the ECAL is imposed and the innermost cells of the ECAL are excluded.

Photon conversions

Photon conversions are selected from two oppositely charged electrons with a di-electron invariant mass below 10 MeV. Due to the small invariant mass and the absence of a magnetic field in the VELO, they are produced with vanishing opening angle. To ensure the electron tracks are overlapping in the VELO, ADC information in the form of the `VeloCharge` is used. This quantity is constructed from the charge deposit in the VELO modules and scaled according to the expected charge deposit for a minimally ionising particle. For a photon conversion, the two electron tracks are overlapping in the VELO, resulting in a charge deposit that is approximately twice as large as for a minimally ionising particle. A cut of `VeloCharge` > 1.25 is found to be sufficient to suppress electrons not originating from photon conversions.
Background from soft photons is reduced by requiring the di-electron track-momentum (i.e. without bremsstrahlung correction) to be larger than 20 GeV. Using the track momentum without bremsstrahlung correction improves the suppression of neutral pion decays $\pi^0 \to \gamma_{ee}\gamma$ for which the additional photon can be mistaken for bremsstrahlung.

Clone tracks

Two tracks are defined as clones if they share at least 70% of their total hits and are removed by a dedicated algorithm in the event reconstruction. However, there can be residual contributions from tracks which share the same VELO track segment but are matched to different track segments downstream of the magnet. While this is a desired feature of electrons from photon conversions, tracks other than the two electrons must not share the same VELO track segment. Such tracks are removed offline, requiring a minimum opening angle of 0.5 mrad between an electron and a muon (hadron) track for the signal (normalisation) selection. These requirements are found to have a negligible impact on the selection efficiencies.

Particle identification

To reduce background from mis-identification, additional PID requirements are imposed. For the electron identification, the PID selection in the stripping of the signal and normalisation channels is aligned.

One dimensional selection cuts		
Type		**Requirement**
Fiducial, Calibration	all tracks	$\texttt{hasRICH}$
	K^+, π^-	$\chi^2_{IP} > 9$ $2\,\text{GeV} < p < 200\,\text{GeV}$ $p_T > 400\,\text{MeV}$
	μ^+	$\chi^2_{IP} > 9$ $3\,\text{GeV} < p < 200\,\text{GeV}$ $p_T > 1200\,\text{MeV}$ $\texttt{isMuon}$
	$e^\pm$	$\chi^2_{IP} > 25$ $3\,\text{GeV} < p < 200\,\text{GeV}$ $p_T > 500\,\text{MeV}$ $\texttt{hasCalo}$
Acceptance	$e^\pm$	$\lvert x_{\text{ECAL}}\rvert > 363.6\,\text{mm}$ or $\lvert y_{\text{ECAL}}\rvert > 282.6\,\text{mm}$
Conversion	$e^\pm$	$\texttt{VeloCharge} > 1.25$
	γ_{ee}	$m(e^+e^-) < 10\,\text{MeV}$ $E^*_\gamma > 1.0\,\text{GeV}$ $p_{track} > 20\,\text{GeV}$ $\texttt{VeloCharge}(e^+) = \texttt{VeloCharge}(e^-)$
Clones	all tracks	$\theta(e^\pm, \mu^+) > 0.5\,\text{mrad}$ $\theta(e^\pm, K^+) > 0.5\,\text{mrad}$ $\theta(e^\pm, \pi^+) > 0.5\,\text{mrad}$
PID	K^+	$\text{DLL}_{\text{K}\pi} > 0$ $\texttt{ProbNNk} \cdot (1 - \texttt{ProbNNp}) > 0.05$
	π^-	$\texttt{ProbNNpi} \cdot (1 - \texttt{ProbNNk}) \cdot (1 - \texttt{ProbNNp}) > 0.1$
	μ^+	$\texttt{ProbNNmu} > 0.2$
	$e^\pm$	$\text{DLL}_{\text{e}\pi} > 3$ $\text{DLL}_{\text{eK}} > 0$
Signal channel	$B^+_{(c)}$	$\chi^2_{FD} > 100$ $\texttt{DIRA} > 0.998$ $1500\,\text{MeV} < m(\mu^+\gamma_{ee}) < 6500\,\text{MeV}$
Normalisation channel	K^{*0}	$m(K^+\pi^-_{\rightarrow K^-}) > 1040\,\text{MeV}$ $\lvert m - m_{PDG}\rvert < 100\,\text{MeV}$
	B^0	$m_{HOP} > 4900\,\text{MeV}$

Table 5.3: Summary of the one-dimensional selection requirements for the signal and normalisation channels applied on top of the stripping and trigger requirements.

The additional PID requirements on the muon track for signal candidates, as well as on the hadron tracks for normalisation candidates, are aligned with the selection criteria in previous LHCb analyses [94,95].
In Section 5.7.2, data-driven corrections to the PID response in simulation are derived. To use the standard methods of LHCb for these corrections, the cuts listed in Table 5.3 under "Fiducial, Calibration" align the data of this analysis with the phase-space of the standard calibration samples. Additional acceptance cuts are imposed on all tracks to have information from the RICH system, while electrons are further required to have information from the calorimeter system. For muons, the cut on the transverse momentum, $p_T > 1200\,\mathrm{MeV}$, is chosen tighter than the default requirement to suppress background candidates with soft muons.

Mass range

For the $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ signal candidates, the invariant visible mass is required to be within $1500\,\mathrm{MeV} < m(\mu^+\gamma_{ee}) < 6500\,\mathrm{MeV}$. The cuts on the `DIRA` and the significance of the B flight distance are tightened with respect to the stripping selection to suppress background and improve the corrected mass resolution.
Candidates of the normalisation channel $B^0 \to K^{*0}\gamma$ are required to contain a K^{*0} candidate within 100 MeV of the known K^{*0} mass, corresponding to about twice the measured width [5]. To suppress background from $\phi \to K^+K^-$, where one kaon is misidentified as a pion, the $(K^+\pi^-)$-invariant mass is required to be larger than 1040 MeV when assigning the kaon mass hypothesis to the pion. Partially reconstructed decays $H_b \to K^+\pi^-\gamma_{ee}X$, where X denotes any number of unreconstructed particles, are reduced by cutting on the HOP mass [96] similarly to [94]. It is constructed by attributing the momentum imbalance perpendicular to the B flight-direction to unaccounted bremsstrahlung emission and correcting the di-electron momentum according to

$$p_{HOP}(e^+e^-) = \frac{p_\perp(K^+\pi^-)}{p_\perp(e^+e^-)}p(e^+e^-) \tag{5.3}$$

from which a corrected HOP mass, $m_{HOP}(K^+\pi^-e^+e^-)$, is calculated. Selecting candidates with a HOP mass exceeding 4900 MeV removes a large fraction of background and has an efficiency of approximately 80% on $B^0 \to K^{*0}\gamma$ decays.

5.6 Multi-variate classification

To further suppress background candidates in data, a *multivariate analysis* (MVA) is performed. In MVA, machine learning algorithms are used to construct multi-dimensional decision trees from representative signal and background proxies. For each candidate, the resulting classifier maps the N input variables to the decision trees onto a single classifier response value, which allows for separation between signal and background.
A *decision tree* is a series of subsequent binary decisions splitting the candidates into multiple smaller subsets, called *branches*. The binary decisions are taken such that the separation between the representative signal and background proxies is maximal. The separation power is further extended by *boosting* the decision trees (BDT), which refers

to the iterative process of growing an ensemble of trees and applying weights for misclassified candidates to increase the separation power of subsequent trees. This significantly enhances the classifier's power to separate between signal and background candidates. However, boosting based on weighted decision trees can cause the classifier to suffer from *overtraining* and be susceptible to statistical fluctuations in the training samples. To avoid this issue, the method of *gradient boosting* is applied in this analysis. Thereby, the performance of a decision tree is encoded in a loss function and its derivative with respect to the training variables of the BDT is computed. This allows to iteratively grow new trees following the gradient of the loss function until convergence is reached, effectively performing a multi-dimensional optimisation of the classifier performance.

Another method of protecting the classifier from overtraining is the k-fold cross-validation technique. Thereby, k independent BDTs are trained, splitting the signal and background training samples into k statistically independent subsets of equal size. Of these, $k-1$ are used to train the BDT and one subset to evaluate the performance of the BDT, respectively.

Searching for $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays in candidates reconstructed from μ^+ and γ_{ee}, three classes of background are identified, which are suppressed using dedicated BDTs. These are

- Combinatorial background from random combinations of tracks
- Candidates in which the converted photon originates from the decay of $\pi^0 \to \gamma_{ee}\gamma$ or $\eta \to \gamma_{ee}\gamma$
- Partially reconstructed B decays with additional charged tracks

Note, there is significant overlap between the different categories, which is why the cut on the classifiers' response is determined in a three-dimensional optimisation process in Section 5.6.4.

5.6.1 Combinatorial BDT

Candidates in which at least one reconstructed final state particle does not originate from the signal B decay are referred to as *combinatorial background* candidates. To suppress this type of background, a dedicated BDT is trained using the XGBoost algorithm [97] for gradient boosting implemented in the SKLEARN python package [98]. To avoid overtraining, a k-fold cross-validation with $k = 5$ is applied using 4/5 of the training samples to train the BDT and 1/5 to evaluate its performance.

Training samples

The combinatorial BDT is trained to separate simulated $B^+ \to \mu^+\nu_\mu\gamma$ signal candidates from background candidates in the upper mass sideband $6500\,\text{MeV} < m_{corr} < 10000\,\text{MeV}$ of recorded data. This sideband is expected to be fully dominated by combinatorial background as the pollution from $B^+_c \to \mu^+\nu_\mu\gamma$ is expected to be negligible. Due to the random nature of this type of background, the properties of candidates in the upper mass sideband are expected to be representative of combinatorial background candidates in the signal region. The resulting training samples comprise approximately 5,200 background candidates and 7,700 signal candidates. For the signal proxy, all simulation correction

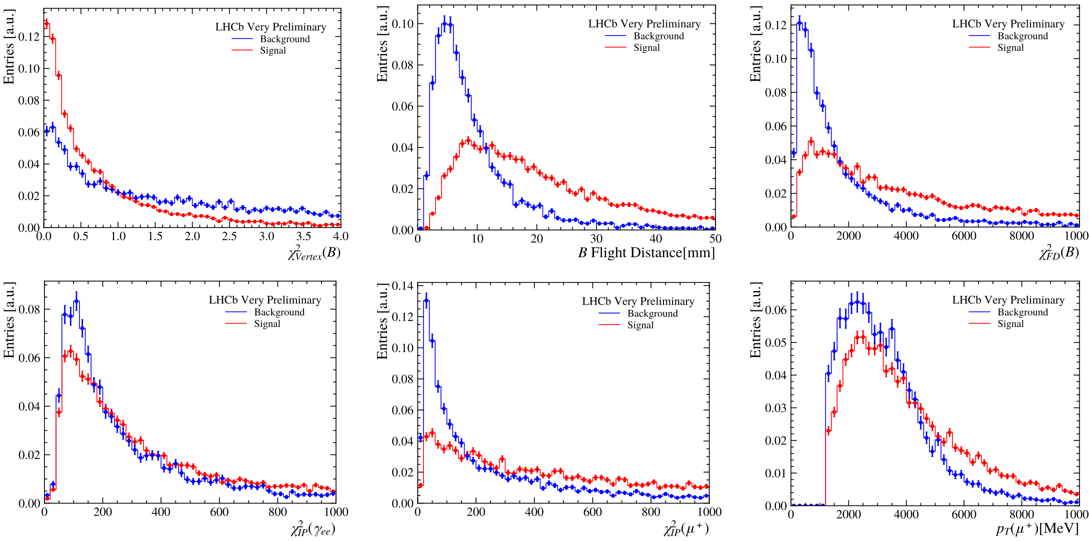


Figure 5.3: Distribution of signal (red) and background (blue) proxies in the variables used in the training of the combinatorial BDT.

weights are applied (see Section 5.7) to ensure correct distributions of the decay properties.

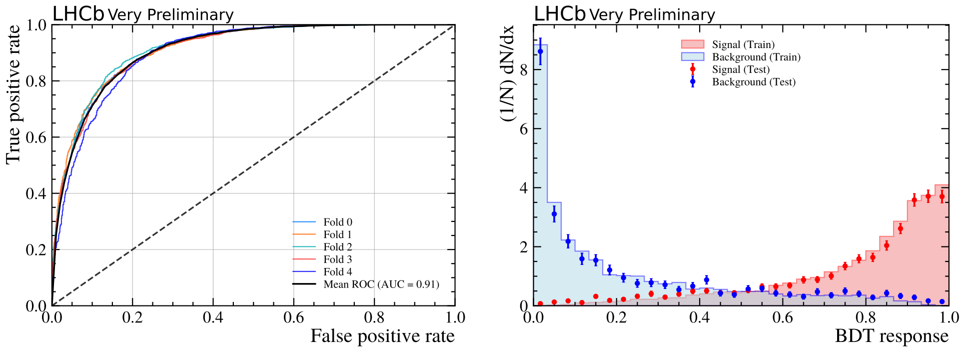
Training variables

For the training variables, a large set of decay properties is considered, including kinematic and topological quantities as well as features related to the reconstruction quality. Properties that show large correlations with each other are removed from the set of training variables. The same applies to variables which are correlated with $m_{corr}(\mu^+\gamma_{ee})$ as this would break the assumption that the upper mass sideband is a representative background proxy. Lastly, decay variables that offer no significant separation of signal and background candidates are removed. The resulting set of decay variables used in the training is the quality of the B decay vertex χ^2, the B flight-distance and its significance; the χ^2_{IP} with respect to the PV of the converted photon and the muon, as well as the transverse momentum of the muon. The distributions of the signal and background proxies in the training variables are shown in Figure 5.3.

BDT performance

The separation power of the combinatorial BDT is illustrated in Figure 5.4a for each classifier fold as well as for all five folds combined. The distribution of the classifier response for the signal and background candidates used in the training is shown in Figure 5.4b for a single fold. None of the folds show signs of overtraining, as the training and testing scores agree well.

To add physical meaning to the classifier response value, it is rescaled to be an approximate measure of the efficiency loss $1 - \varepsilon_{signal}$ on $B^+ \to \mu^+\nu_\mu\gamma$ signal candidates induced by the BDT. The resulting BDT response is denoted `CombBDT` with a cut of e.g. `CombBDT` > 0.4, rejecting approximately 40% of $B^+ \to \mu^+\nu_\mu\gamma$ signal candidates. The optimal BDT cut



(a) (b)

Figure 5.4: (a) ROC curves for the combinatorial BDT drawn for the five folds and the mean of all folds. (b) Check for overtraining of the BDT for one fold. The red and blue histograms represent the signal and background proxies used for training, corresponding to 4/5 of the total statistics. The red and blue points represent 1/5 of the signal and background proxies used for testing.

is determined by maximising the sensitivity to the branching ratio of $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays as explained in detail in Section 5.6.4 and is found to be `CombBDT` > 0.50. Due to the different decay characteristics of the normalisation candidates (e.g. the absence of a muon), the requirement on the `CombBDT` is not applied in the selection of normalisation candidates.

5.6.2 Neutral isolation BDT

Candidates in which the converted photon, γ_{ee}, originates from the decay of a neutral pion $\pi^0 \to \gamma_{ee}\gamma$ rather than from a heavy B meson are suppressed with a dedicated BDT. The idea is to search for an additional photon in the calorimeter to identify $\pi^0 \to \gamma_{ee}\gamma$ decays. The same XGBoost algorithm as for the combinatorial BDT is used with k-fold cross-validation of $k = 5$.

Training samples

To reduce potential biases from imperfections in the simulation, the classifier uses simulated candidates as a signal and background proxy. This ensures that any residual mis-modelling in the event simulation does not bias the classifier response. Simulated $B^+ \to \mu^+\nu_\mu\gamma$ candidates are used as signal proxy, and candidates from the inclusive simulation of $B^{+,0} \to \mu^+\nu_\mu X_u(\to X\pi^0)$ decays are used as background proxy. To enhance the available statistics in the training samples, the PID requirements on the electrons are removed. The resulting signal and background proxies comprise approximately 13,600 and 12,400 candidates, respectively.

Training variables

Candidates of neutral pion decays are built by combining the converted photon, γ_{ee}, with an additional photon reconstructed from a neutral cluster in the ECAL, γ_{calo}. In searching for the additional photon, the momentum of the converted photon is extrapolated to the

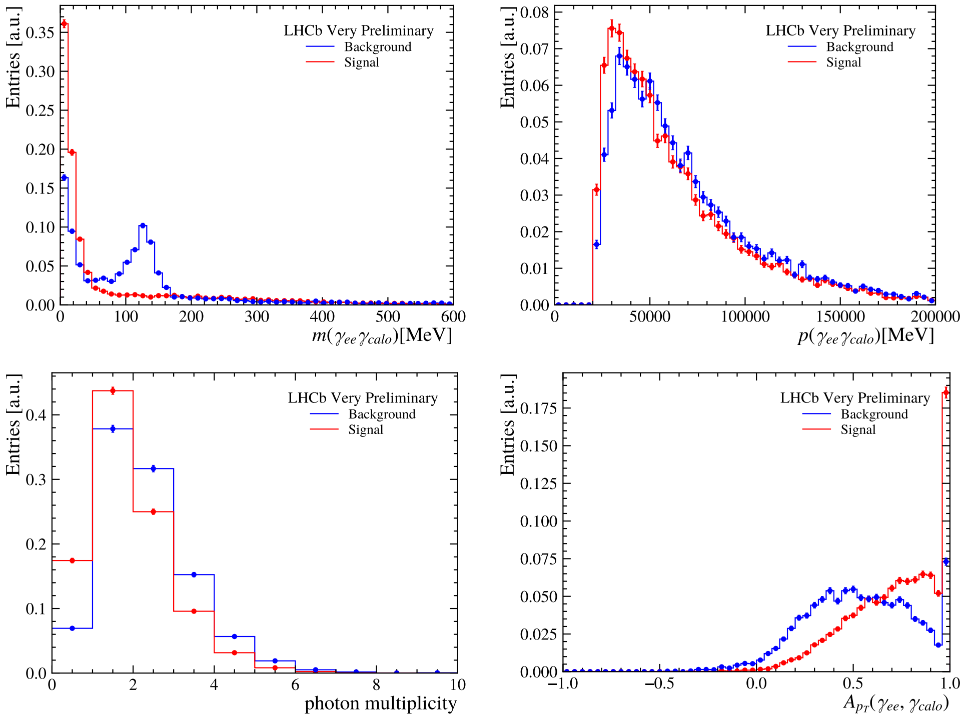


Figure 5.5: Distribution of signal (red) and background (blue) proxies in the training variables.

ECAL and a search window of radius R around this axis is defined. The radius R is given in terms of the pseudorapidity, η, and azimuthal angle, ϕ, as

$$R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{5.4}$$

Within this radius, the reconstructed photon with the maximum transverse momentum is selected as γ_{calo}. The radius is set to $R = 0.3$ to use the same definition of the calorimeter photon as in the modelling of the corrected mass distribution, detailed in Section 6.3. Using this definition of γ_{calo}, the training variables for the neutral isolation BDT are defined as: the invariant mass $m(\gamma_{ee}\gamma_{calo})$ and momentum $p(\gamma_{ee}\gamma_{calo})$ of the π^0 candidate, the number of reconstructed photons in the search window and the p_T asymmetry, A_{p_T}, of the γ_{calo} and γ_{ee} defined as

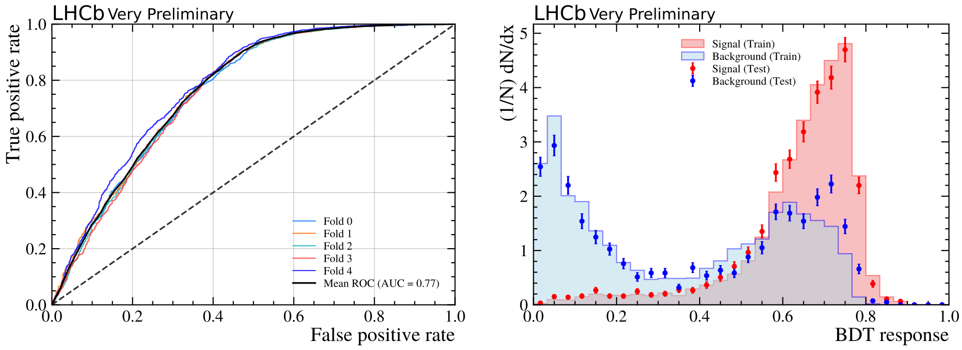
$$A_{p_T} = \frac{p_T(\gamma_{ee}) - p_T(\gamma_{calo})}{p_T(\gamma_{ee}) + p_T(\gamma_{calo})}. \tag{5.5}$$

The distributions of the training variables for signal and background proxies are shown in Figure 5.5.

BDT performance

Figure 5.6a shows the separation power of the classifier per fold as well as for all folds combined. The classifier response of the neutral isolation BDT is shown in Figure 5.6b for the training and validation samples for a single fold. No significant overtraining is observed for the signal response across the five folds.

The final classifier response `PiBDT` is again rescaled to reflect the efficiency loss $1 - \varepsilon_{signal}$



(a) (b)

Figure 5.6: (a) ROC curves for the neutral isolation BDT drawn for the five folds and the mean of all folds. (b) Check for overtraining of the BDT for one fold. The red and blue histograms represent the signal and background proxies used for training, corresponding to 4/5 of the total statistics. The red and blue points represent 1/5 of the signal and background proxies used for testing.

on $B^+ \to \mu^+\nu_\mu\gamma$ signal candidates induced by a requirement on `PiBDT`. The optimal cut is found to be `PiBDT` > 0.50, which maximises the sensitivity towards the branching ratios of $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays as explained in Section 5.6.4. The same requirement is also applied to $B^0 \to K^{*0}\gamma$ normalisation candidates to achieve better cancellation of systematic uncertainties.

Veto of $\eta \to \gamma_{ee}\gamma$ decays

Similarly to background candidates in which the converted photon originates from the decay of a $\pi^0 \to \gamma_{ee}\gamma$, decays of $\eta \to \gamma_{ee}\gamma$ form a similar type of background. These decays can be suppressed using a similar method as for the suppression of $\pi^0 \to \gamma_{ee}\gamma$ decays by searching for an additional calorimeter photon in a certain region of the ECAL. To reconstruct $\eta \to \gamma_{ee}\gamma_{calo}$ decays, the same definition of γ_{calo} is used as in the reconstruction of neutral pion decays. Instead of suppressing $\eta \to \gamma_{ee}\gamma$ decays with a dedicated BDT, a simple veto of $m(\gamma_{ee}\gamma_{calo}) \notin [450, 600]$ MeV is imposed as the peak in $m(\gamma_{ee}\gamma_{calo})$ is well separated from the region close to zero populated by $B^+ \to \mu^+\nu_\mu\gamma$ signal candidates. Figure 5.7 shows the distribution of $m(\gamma_{ee}\gamma_{calo})$ for $B^+ \to \mu^+\nu_\mu\gamma$ and $B^+ \to \mu^+\nu_\mu\eta(\to \gamma_{ee}\gamma)$ along with the imposed veto.

For modelling the residual background from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ decays in Section 6.3, it is convenient to define a single quantity to suppress both sources of background simultaneously. This is achieved by redefining the score of `PiBDT` as

$$\texttt{PiBDT} \to \begin{cases} \texttt{PiBDT} & \text{if } m(\gamma_{ee}\gamma_{calo}) \notin [450, 600]\,\text{MeV} \\ 0 & \text{if } m(\gamma_{ee}\gamma_{calo}) \in [450, 600]\,\text{MeV}. \end{cases} \tag{5.6}$$

5.6.3 Charged isolation BDT

The third prominent source of background is from partially reconstructed decays, in which at least one charged track is not reconstructed as part of the signal candidate. The overlap with background from $\pi^0 \to \gamma_{ee}\gamma$ decays is large, e.g. in $B^+ \to \overline{D}^0(\to K^+\pi^-\pi^0)\mu^+\nu_\mu$

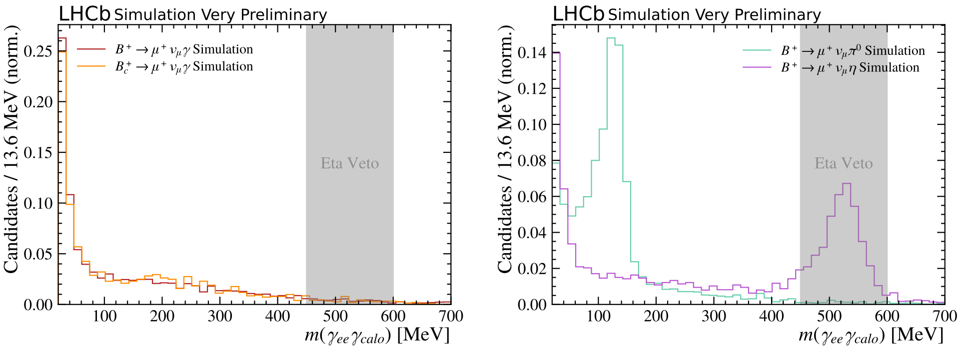


Figure 5.7: Simulation of $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ signal candidates (left) and $B^+ \to \mu^+\nu_\mu\pi^0$ and $B^+ \to \mu^+\nu_\mu\eta$ background candidates (right) with the veto of $\eta \to \gamma_{ee}\gamma_{calo}$ decays shown in grey. The population at low invariant mass present in both signal and background simulation plots is due to the selected γ_{calo} being a bremsstrahlung photon of one of the electrons.

decays. Previous LHCb analyses [92, 99] have established a powerful tool in which candidates are classified as background if an additional track is found in the event that forms a good vertex with a given signal track.

In this analysis, the classification trained in the analysis of $B^0_s \to K^-\mu^+\nu_\mu$ decays [99] is applied. Due to the similarity of the decay topology between $B^0_s \to K^-\mu^+\nu_\mu$ and $B^+ \to \mu^+\nu_\mu\gamma$ decays[1], the exact classifier of Ref. [99] is applied. Additional tracks from the underlying event are combined with a signal track (μ^+ and $e^\pm$), and the compatibility of the two tracks originating from a common vertex is evaluated using the BDT. The response of the charged isolation BDT on the signal tracks ranges from -1 (classified as a non-isolated track) to 1 (classified as an isolated track). For the selection of signal (normalisation) candidates, the minimum BDT response among the three (four) final state tracks is used and denoted `IsoBDT`. The distribution of the `IsoBDT` response in data and simulation is shown in Figure 5.8. The difference in the BDT performance for the B^+ and B_c^+ signal candidates is due to the longer lifetime of the B^+ meson, which causes a better separation from other charged tracks.

Following the optimisation process in Section 5.6.4, a requirement of `IsoBDT` > -0.70 is imposed on signal candidates. The same cut on the `IsoBDT` is also applied in the selection of normalisation candidates.

5.6.4 BDT optimisation

Strategy

The selection on the BDT response is optimised by maximising the sensitivity towards the branching fractions of $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays. Due to the different expected levels of background in the regions of the $B^+ \to \mu^+\nu_\mu\gamma$ and $B_c^+ \to \mu^+\nu_\mu\gamma$ decays, the set of BDT selections that optimises the individual branching ratios is potentially different. To still be able to use a single set of BDT cuts, the primary aim is to maximise the sensitivity towards the branching ratio of the B^+ decay. If multiple sets of BDT selections yield a similar sensitivity, the secondary aim is to maximise the sensitivity towards the B_c^+ decay.

---

[1] Due to the tracks from the converted photon overlapping in the VELO, both decays exhibit an effective two-track + neutrino decay topology.

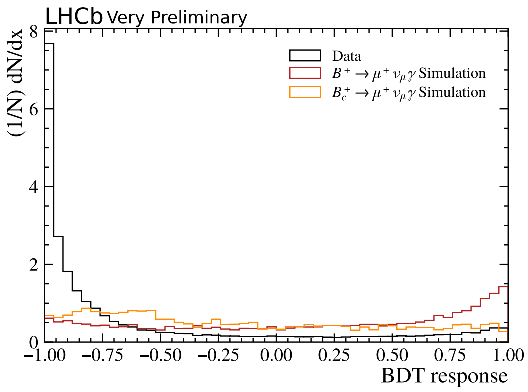


Figure 5.8: Distribution of the `IsoBDT` output for data (black) and simulation of the signal components $B^{+} \to \mu^{+}\nu_{\mu}\gamma$ in red and $B_{c}^{+} \to \mu^{+}\nu_{\mu}\gamma$ in orange.

Furthermore, it is computationally very expensive to perform a three-dimensional optimisation, scanning cuts on the `PiBDT`, `CombBDT` and `IsoBDT` simultaneously. Hence, the first step of the optimisation is to choose the requirements on the `PiBDT` and `CombBDT` in a two-dimensional scan of the sensitivity for a loose cut on the charged isolation BDT of `IsoBDT` > -0.9. In a second step, the `IsoBDT` is optimised in a one-dimensional scan of the sensitivity for fixed cuts of the `PiBDT` and `CombBDT`.

Figure of Merit

To estimate the sensitivity on the $B_{(c)}^{+} \to \mu^{+}\nu_{\mu}\gamma$ branching fractions, a number of *pseudo-experiments* are conducted. For each BDT selection, background-only *pseudo-data* are generated (see Section 6.6) and fitted with the nominal fit strategy established in Section 6.5 with freely floating signal strengths for the B^{+} and B_{c}^{+} signal components. From the profiling of the log-likelihood, the approximate coverage at 90% confidence level is estimated. The median across 300 pseudo-experiments is denoted μ_{90}.
To relate the values of μ_{90} to a limit on the branching fractions, knowledge of the signal efficiencies is required. However, at this stage of the analysis, the absolute scale of the sensitivity can be omitted. Instead, the optimisation of the BDT selection is performed using a figure of merit (FoM), defined as

$$FoM = \frac{\mu_{90}}{\varepsilon_{BDT}} \tag{5.7}$$

with ε_{BDT} being the efficiency of just the BDT selection determined on simulation data.

PiBDT and CombBDT optimisation

The FoM is scanned simultaneously in `PiBDT` and `CombBDT` separately for the two signal components $B^{+} \to \mu^{+}\nu_{\mu}\gamma$ and $B_{c}^{+} \to \mu^{+}\nu_{\mu}\gamma$ in increments of 0.10. Figure 5.9 shows the FoM for the two signal channels relative to the best FoM across the two-dimensional scan. In $B^{+} \to \mu^{+}\nu_{\mu}\gamma$, the FoM shows a wide minimum with a large number of BDT selections maximising the sensitivity. The selection `PiBDT` > 0.50 and `CombBDT` > 0.50 is chosen as it also fulfils the secondary requirement of maximising the sensitivity for $B_{c}^{+} \to \mu^{+}\nu_{\mu}\gamma$.
The stability of the BDT selection in `PiBDT` and `CombBDT` is ensured by performing the

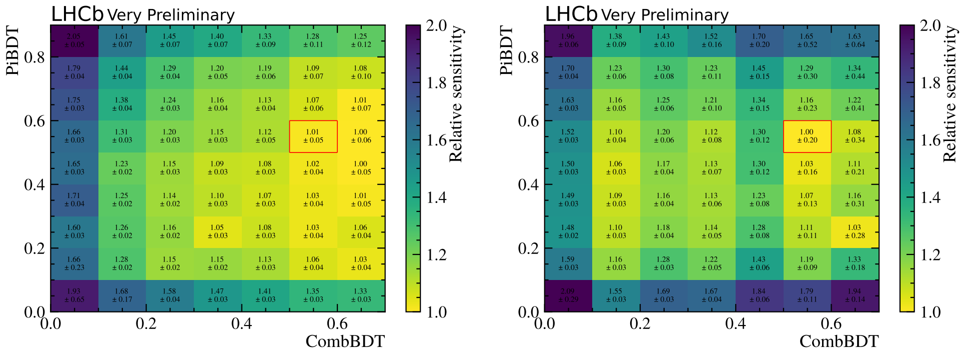


Figure 5.9: Two-dimensional scan of the FoM in `PiBDT` and `CombBDT` relative to the lowest FoM for the $B^+ \to \mu^+\nu_\mu\gamma$ (left) and $B_c^+ \to \mu^+\nu_\mu\gamma$ (right) signal channels. The FoM is determined for a fixed cut on `IsoBDT` > -0.90. The position indicated in red is the chosen nominal BDT selection `CombBDT` > 0.50 and `PiBDT` > 0.50.

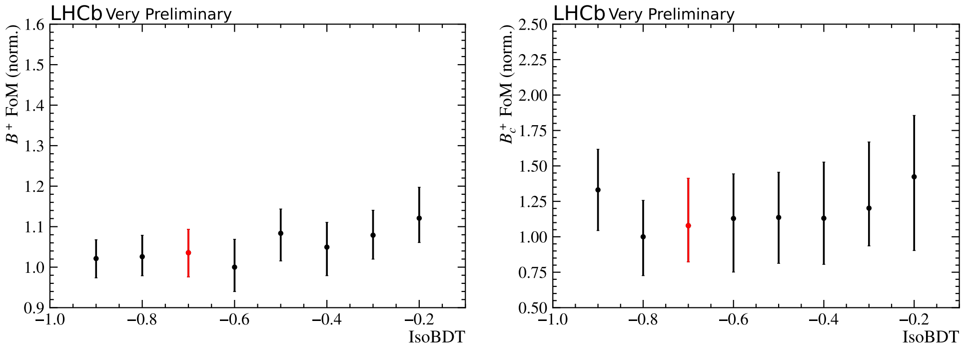


Figure 5.10: One-dimensional scan of the FoM in `IsoBDT` relative to the lowest FoM for $B^+ \to \mu^+\nu_\mu\gamma$ (left) and $B_c^+ \to \mu^+\nu_\mu\gamma$ (right) signal channels. The FoM is determined for the nominal BDT requirement of `CombBDT` > 0.50 and `PiBDT` > 0.50. The point indicated in red is the nominal BDT cut `IsoBDT` > -0.70.

same two-dimensional scan for alternative selections on the `IsoBDT`. No significant change in the choice of the optimal `PiBDT` and `CombBDT` selection is found.

IsoBDT optimisation

After optimisation of the `PiBDT` and `CombBDT` selection, the cut on `IsoBDT` is optimised in a one-dimensional scan of the FoM for both $B^+ \to \mu^+\nu_\mu\gamma$ and $B_c^+ \to \mu^+\nu_\mu\gamma$. Figure 5.10 shows the FoM relative to the lowest FoM for the two signal channels. Again, a range of selections on the `IsoBDT` is found to maximise the sensitivity towards the $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$ branching fraction from which the nominal selection is chosen as `IsoBDT` > -0.70.

5.7 Corrections to simulation

The analysis relies on simulation data in the training of multivariate classifiers (see Sections 5.6.1 and 5.6.2), study of background components (see Section 6.3), modelling mass shapes (see Section 6.5) and estimation of the selection efficiencies (see Chapter 8). It is known that some properties are not modelled accurately in the LHCb simulation. These

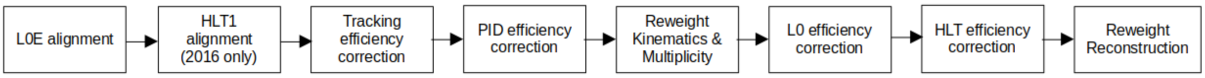


Figure 5.11: Set of subsequent corrections to simulation.

include, in particular, the initial B kinematics and event multiplicity, the trigger response, the reconstruction of electron tracks and particle identification. Simulated events are corrected for the effects mentioned above to represent the event properties in the data accurately.

The strategy of correcting simulation data is adopted from a previous LHCb analysis in [94]. Still, due to differences in the candidate selection, all corrections are derived explicitly for this analysis. The correction chain is illustrated in Figure 5.11 with corrections applied for all subsequent steps.

5.7.1 Trigger alignment

Due to irradiation, the performance of the ECAL degrades with time, causing the meaning of the recorded transverse energy, E_T, for the `L0Electron` trigger decision to vary with time. Despite frequent recalibrations, this is not fully reflected in the simulation. In addition, different E_T-thresholds have been used throughout the data-taking, especially in 2016 and 2017; however, the trigger response in simulation is generated from a single trigger setting. To better align the trigger decision in data and simulation, an additional cut on the electron E_T is applied that is tighter than any of the E_T thresholds used in data-taking

$$\texttt{L0Electron} \to \texttt{L0Electron}\,\&\, E_T^{\mathrm{L0}} > \begin{cases} 2700\,\mathrm{MeV} & \text{for 2016,2017} \\ 2400\,\mathrm{MeV} & \text{for 2018,} \end{cases} \tag{5.8}$$

where E_T^{L0} is the transverse energy as reconstructed at the L0 level. Alignment of the L0E trigger requirement on $B^+ \to K^+ J/\psi(\to e^+e^-)$ candidates results in a loss of statistics of about 6% for 2016, 3% for 2017 and 0.2% for 2018.

Also, the thresholds of the HLT selection have changed during data taking. Especially, the requirements of the `Hlt1TrackMVA` decision are varied a lot during the 2016 data-taking, which is not reflected in the trigger emulation of simulated events. To better align the `Hlt1TrackMVA` decision for the 2016 simulation, the trigger decision is recomputed offline for each simulated event using the correct constants employed during the 2016 data-taking.

5.7.2 Particle Identification corrections

The LHCb simulation is known not to produce the correct PID response reliably. To account for this, collaboration-wide calibration tools are used to correct the PID efficiency in simulation. In this analysis, the PID efficiencies of muons and hadrons are obtained from the PIDCalib2 [100] package. High statistics data calibration samples are used to derive the efficiency for a given PID requirement in data using a tag-and-probe method (see Section 3.3.2 for an explanation of the method). The muon PID efficiency is derived from a calibration sample of $J/\psi \to \mu^+\mu^-$ decays, while for charged pions and kaons, a sample of $D^{*+} \to D^0(\to K^+\pi^-)\pi^+$ decays is used. For electron PID, a calibration sample of $B^+ \to K^+J/\psi(\to e^+e^-)$ decays is used.

Tight kinematic and PID requirements are applied on the tag track and residual background is statistically subtracted using the ${}_s\mathcal{P}lot$ technique [101]. The PID efficiency is determined in bins of `nTracks` to represent the event multiplicity, particle momentum and pseudorapidity, separately for each year, polarity and particle species.

For the calibration of electron PID, the background subtraction in the calibration samples is known to suffer from limited statistics and correlations between the invariant mass and the momentum and pseudorapidity of the probe track. Instead, the PID efficiency for electrons is determined from a simultaneous fit to the reconstructed invariant mass of $B^+ \to K^+J/\psi(\to e^+e^-)$ candidates in the PIDCalib2 control sample using the pass-fail method as in Section 3.3.2. The PID efficiency of electrons is determined in bins of `nTracks`, transverse momentum, p_T, and pseudorapidity, η, separately for each year of data-taking and whether or not a bremsstrahlung photon is associated with the probe electron. One of the many mass fits from the pass-fail method is shown in Figure 5.12.

The simulation is corrected by multiplying the derived PID efficiency for each track and assigning it as a per-event-weight, e.g. $w_{\text{PID}} = w_{\text{PID}}(\mu^+) \times w_{\text{PID}}(e^+) \times w_{\text{PID}}(e^-)$ for $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ candidates.

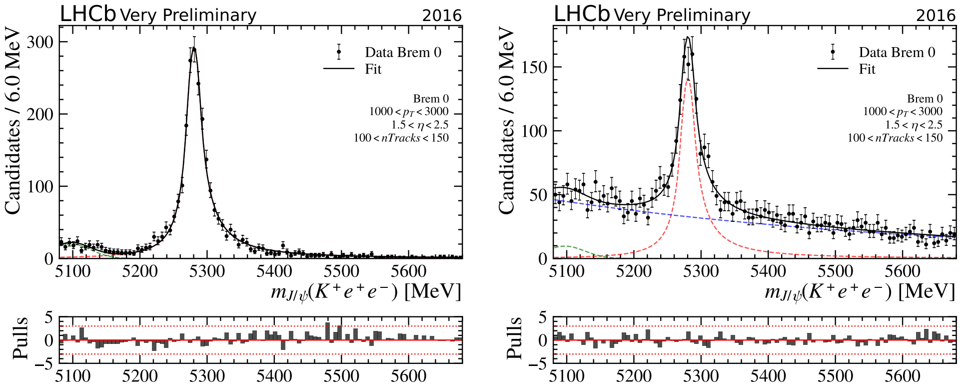


Figure 5.12: Exemplary fit to $B^+ \to K^+J/\psi(\to e^+e^-)$ calibration data for candidates passing (left) and not passing (right) the electron PID requirement $DLL_{e\pi} > 3$ on the probe electron. The calibration data shown is from 2016 without bremsstrahlung associated with the probe electron and for `nTracks` $\in [100, 150]$, $p_T \in [1000, 3000]$ MeV and $\eta \in [1.5, 2.5]$.

5.7.3 Tracking corrections

The efficiency of reconstructing a long track in simulation is susceptible to uncertainties due to modelling particle interactions with the detector material. For electrons, the tracking efficiency is corrected from the results of Ref. [102] in which the electron tracking efficiency in data is derived from $B^+ \to K^+J/\psi(\to e^+e^-)$ decays using a tag-and-probe method. For muon tracks, correction weights are taken from the TrackCalib package developed in the LHCb collaboration from a similar tag-and-probe method using $J/\psi \to \mu^+\mu^-$ decays [60]. In both cases, the correction weights are provided in bins of track momentum and pseudorapidity for each year of data-taking.

For the tracking efficiency of hadrons, the same corrections are applied as for muons. Due

to hadronic interactions with the detector material, an additional systematic uncertainty is assigned for these tracks linked to the uncertainty on the material description in simulation (see Chapter 9).
The final tracking efficiency correction weight is taken as the product of the weights for all reconstructed final state particles, e.g. $w_{\text{Trk}} = w_{\text{Trk}}(\mu^+) \times w_{\text{Trk}}(e^+) \times w_{\text{Trk}}(e^-)$ for $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ candidates.

5.7.4 Multiplicity and kinematics corrections

The initial kinematics of simulated B meson decays exhibit some discrepancies with respect to the kinematics observed in data, particularly in the transverse momentum, p_T, and pseudorapidity, η, of the B meson. In addition, the event multiplicity in simulated events is not representative of the event multiplicity in recorded data. Both event multiplicity and initial B kinematics are corrected simultaneously using a GBReweighter implemented in the HEP_ML python package [103]. The underlying method uses BDTs trained on the event multiplicity and B kinematics in simulation and data to predict the correction weights for simulated events, accounting for possible correlations.
The correction weights are derived from data and simulation of $B^+ \to K^+J/\psi(\to \mu^+\mu^-)$ decays as a function of the B momentum and transverse momentum, the B pseudorapidity and the event multiplicity, `nTracks`. Due to a lack of simulation events generated with high event multiplicity, a cut on `nTracks` < 330 is applied to prevent the weights from becoming excessively large. Figure 5.13 shows the improved agreement of data and simulation following the correction of the event multiplicity and initial B kinematics on $B^+ \to K^+J/\psi(\to e^+e^-)$ candidates.
For corrections of the initial decay kinematics and event multiplicity in the simulation of B_c^+ decays, a dedicated GBReweighter is trained. The same strategy as for B^+ decays is applied using simulation and background-subtracted data of $B_c^+ \to J/\psi(\to \mu^+\mu^-)\pi^+$ decays.

5.7.5 L0 trigger calibration

The modification to the `L0Electron` trigger decision is already increasing the alignment between data and simulation. Nonetheless, there is still some residual disagreement between the trigger performance of recorded data and simulated events. The efficiency of the L0 trigger selection is corrected using the TISTOS method [104]. The trigger efficiency is computed separately in data and simulation on an unbiased dataset that is *triggered independently of signal* (TIS), for example, by the respective other beauty hadron originating from the same $b\bar{b}$ pair. In this dataset, the efficiency of trigger decisions on the signal candidate (TOS) is computed from

$$\varepsilon_{L0}^{Data,MC} = \left(\frac{N_{TIS\&TOS}}{N_{TIS}}\right)^{Data,MC}, \tag{5.9}$$

where N_{TIS} is the number of events that are TIS and $N_{TIS\&TOS}$ denotes the number of events that are TIS and TOS. From this, simulation correction weights are derived as

$$w_{L0} = \frac{\varepsilon_{L0}^{Data}}{\varepsilon_{L0}^{MC}}. \tag{5.10}$$

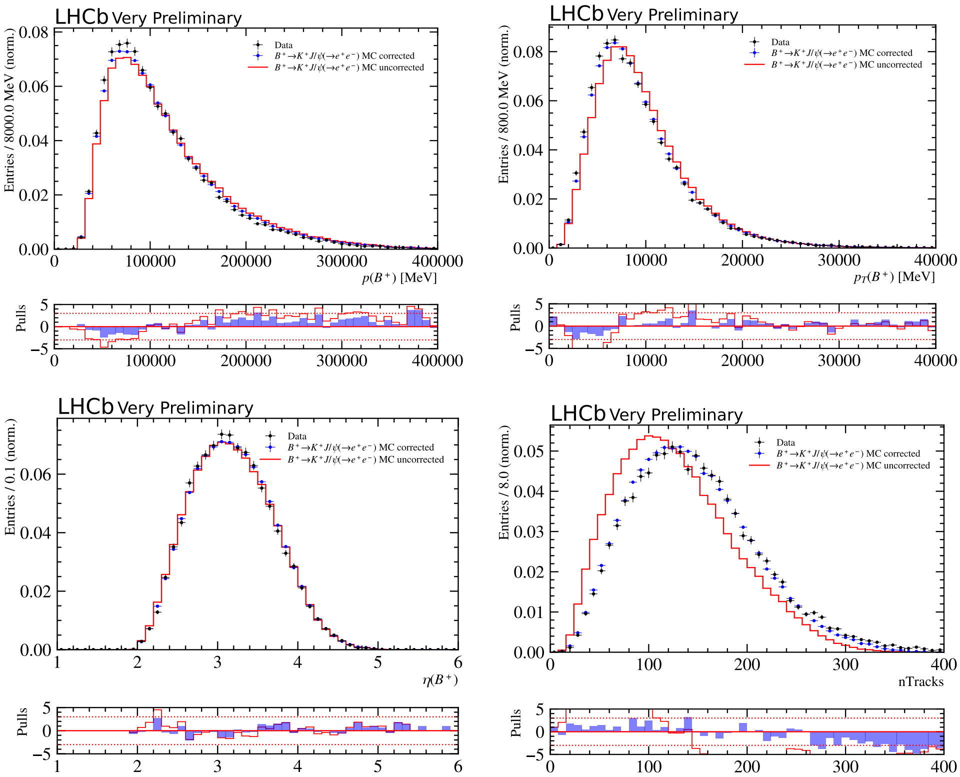


Figure 5.13: Event multiplicity and B kinematics in 2016 $B^+ \to K^+ J/\psi(\to e^+e^-)$ data (black) and simulation with (blue) and without (red) correction weights applied.

First, the `L0Electron` and `L0Muon` trigger efficiencies for a single electron and muon are calculated. In a second step, the overlap between the different trigger decisions is taken into account.

Calibration of L0E

The `L0Electron` trigger efficiency is derived on $B^+ \to K^+ J/\psi(\to e^+e^-)$ candidates in data and simulation as a function of the electron's transverse energy, E_T. The ratio of the obtained efficiencies, $w = \varepsilon^{Data}_{L0Electron}/\varepsilon^{MC}_{L0Electron}$, defines the correction weight for a single electron to pass the `L0Electron` decision, which is shown in Figure 5.14.
As the L0E category is defined as the logical *or* of the `L0Electron` TOS decision on the e^+ and e^-, the overlap between the two needs to be taken into account

$$
\begin{aligned}
\varepsilon^{Data,MC}_{L0E} &= \varepsilon^{Data,MC}_{L0Electron}(e^+) + \varepsilon^{Data,MC}_{L0Electron}(e^-) - \varepsilon^{Data,MC}_{L0Electron}(e^+) \cdot \varepsilon^{Data,MC}_{L0Electron}(e^-) \\
&= 1 - \left(1 - \varepsilon^{Data,MC}_{L0Electron}(e^+)\right) \cdot \left(1 - \varepsilon^{Data,MC}_{L0Electron}(e^-)\right) .
\end{aligned}
\tag{5.11}
$$

The final L0E correction weights are then given as

$$
w_{L0E} = \frac{\varepsilon^{data}_{L0E}}{\varepsilon^{MC}_{L0E}} = \frac{1 - \left(1 - \varepsilon^{data}_{L0Electron}(e^+)\right) \cdot \left(1 - \varepsilon^{data}_{L0Electron}(e^-)\right)}{1 - (1 - \varepsilon^{MC}_{L0Electron}(e^+)) \cdot (1 - \varepsilon^{MC}_{L0Electron}(e^-))} . \tag{5.12}
$$

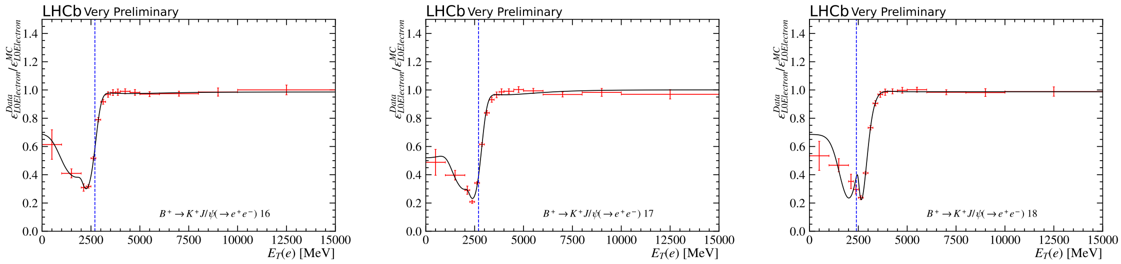


Figure 5.14: Ratio of efficiencies of the `L0Electron` TOS decision in data and simulation as a function of the transverse energy of the electron for the years 2016 to 2018 (left to right). The dashed blue line indicates the E_T threshold used to align the `L0Electron` trigger alignment.

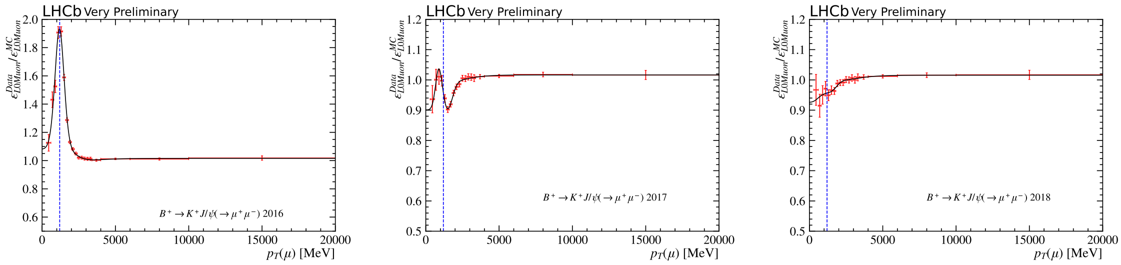


Figure 5.15: Ratio of efficiencies of the `L0Muon` TOS decision in data and simulation as a function of the transverse momentum of the muon for the years 2016-2018 (left to right). The dashed blue line indicates the cut $p_T > 1200\,\text{MeV}$ on muon tracks applied in the nominal selection.

Calibration of L0M

Analogously, the efficiency of the `L0Muon` decision is evaluated on $B^+ \to K^+J/\psi(\to \mu^+\mu^-)$ candidates in data and simulation as a function of the muon p_T. The derived correction weights are shown in Figure 5.15.
To obtain the L0 correction weights for $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ candidates in the exclusive L0M trigger category, the overlap with L0E needs to be considered. Similar to the overlap between the two electrons in the L0E category, the efficiency for the exclusive L0M category is derived from

$$\varepsilon^{Data,MC}_{L0M} = \varepsilon^{Data,MC}_{L0Muon}(\mu^+)\cdot\Big(1-\varepsilon^{Data,MC}_{L0Electron}(e^+)\Big)\cdot\Big(1-\varepsilon^{Data,MC}_{L0Electron}(e^-)\Big)\,. \tag{5.13}$$

The correction weight for the exclusive L0M decision then reads

$$w_{L0M} = \frac{\varepsilon^{data}_{L0M}}{\varepsilon^{MC}_{L0M}} = \frac{\varepsilon^{data}_{L0Muon}(\mu^+)}{\varepsilon^{MC}_{L0Muon}(\mu^+)}\cdot\frac{\Big(1-\varepsilon^{data}_{L0Electron}(e^+)\Big)\cdot\Big(1-\varepsilon^{data}_{L0Electron}(e^-)\Big)}{(1-\varepsilon^{MC}_{L0Electron}(e^+))\cdot(1-\varepsilon^{MC}_{L0Electron}(e^-))} \tag{5.14}$$

5.7.6 HLT calibration

Corrections to the efficiency of the HLT selection are derived using the same *TISTOS* method as for the correction of the L0 trigger. The efficiencies of the combined HLT1 and HLT2 selections are determined on $B^+ \to K^+J/\psi(\to e^+e^-)$ and $B^+ \to K^+J/\psi(\to \mu^+\mu^-)$

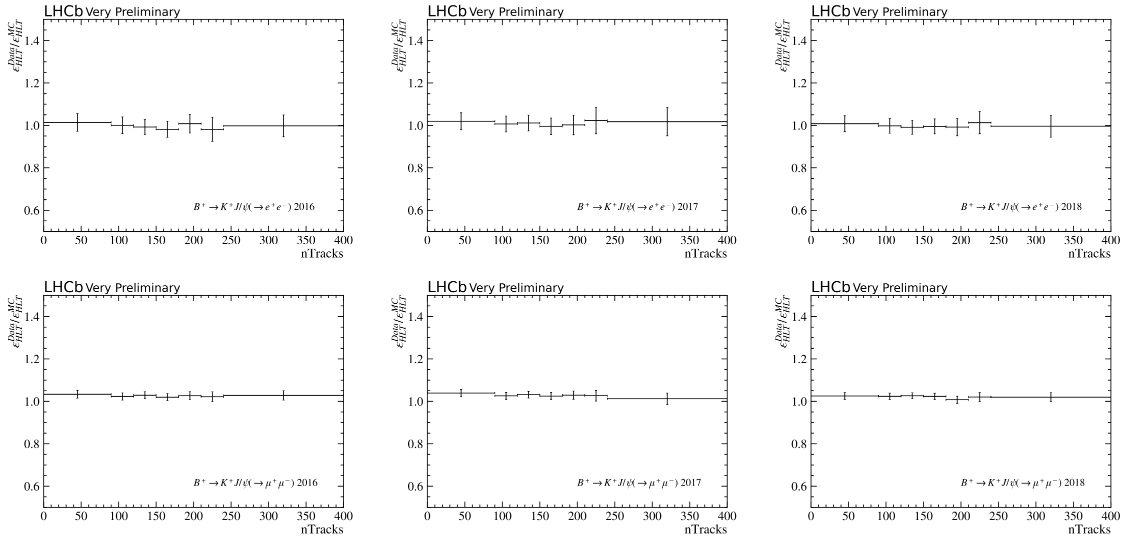


Figure 5.16: Ratio of efficiencies of the HLT selection in data and simulation in bins of `nTracks` for the years 2016-2018 (left to right). The efficiencies are derived from $B^+ \to K^+J/\psi(\to e^+e^-)$ decays triggered on L0E (top) and $B^+ \to K^+J/\psi(\to \mu^+\mu^-)$ decays triggered on L0M (bottom).

candidates in data and simulation, triggered on L0E and L0M, respectively. From the ratio of the efficiency in data and simulation, the HLT simulation correction weights, $w_{HLT} = \varepsilon^{Data}_{HLT}/\varepsilon^{MC}_{HLT}$, are derived. Figure 5.16 shows the weights for the L0E and L0M categories in bins of `nTracks`.

5.7.7 Correction of event reconstruction

Differences between data and simulation in the event reconstruction, especially the vertex resolutions, are corrected for using $B^+ \to K^+J/\psi(\to e^+e^-)$ candidates in data and simulation. A GBReweighter is trained on the quality of the B vertex fit, $\chi^2_{Vertex}(B^+)$, and the significance of the B flight distance, $\chi^2_{FD}(B^+)$. To account for correlation with the B kinematics, also $p_T(B^+)$ and $\eta(B^+)$ are considered. The improved agreement between data and simulation for $\chi^2_{Vertex}(B^+)$ and $\chi^2_{FD}(B^+)$ is shown in Figure 5.17.

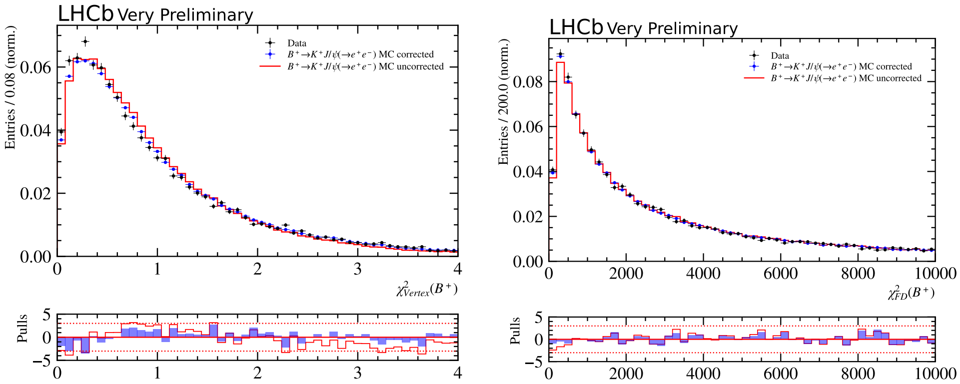


Figure 5.17: B vertex quality (left) and B flight distance (right) for $B^+ \to K^+J/\psi(\to e^+e^-)$ candidates in 2016 data (black) and simulation with (blue) and without (red).

6

Fitting corrected mass distribution

The number of observed $B^+ \to \mu^+\nu_\mu\gamma$ and $B_c^+ \to \mu^+\nu_\mu\gamma$ candidates is extracted from a single fit to the corrected mass distribution, combining the recorded data from 2016 to 2018. In Section 6.1, the fundamentals of the likelihood fit are discussed. The templates modelling the corrected mass distribution of $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$ signal candidates are derived in Section 6.2. The corrected mass distribution for background from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ decays is modelled in Section 6.3. Additional sources of background are discussed in Section 6.4. The fit to the reconstructed corrected mass is established in Section 6.5 and validated in Section 6.6.

6.1 Fundamentals

Assuming a set of observed data points x_i and a probability density function (PDF) $\mathcal{P}(x_i|\boldsymbol{\theta})$ with model parameters $\boldsymbol{\theta} = (\theta_1, \theta_2, ...)$, a maximum likelihood estimation can be used to find the most probable parameter values, $\hat{\boldsymbol{\theta}}$, based on the observed data points. For N independent measurements, the likelihood function is defined as

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i=1}^{N} \mathcal{P}(x_i|\boldsymbol{\theta}). \tag{6.1}$$

The objective is to maximise the likelihood function to determine the most probable model parameters. In practice, this is achieved by defining the negative log-likelihood function

$$-\log \mathcal{L}(\boldsymbol{\theta}) = -\sum_{i=1}^{N} \log \mathcal{P}(x_i|\boldsymbol{\theta}) \tag{6.2}$$

and minimise it with respect to the model parameters. The likelihood function can also be applied to data that is binned in an observable. In a counting experiment, the number of observed counts, n_i, in each bin, i, is expected to follow a Poisson distribution

$$\mathcal{P}_{Poisson}(n_i|\mu_i) = \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!}, \tag{6.3}$$

where μ_i is the expected number of counts in bin i. The resulting log-likelihood reads

$$-\log \mathcal{L} = -\sum_i \log \left(\frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!} \right) = \sum_i (-n_i \log \mu_i + \mu_i + \log n_i!), \tag{6.4}$$

where the last term is constant and can be omitted for minimisation. In typical applications, the PDF describing the data is the sum of several PDFs, $\mathcal{P}_j$, modelling different processes, j. For example, in the fit to derive the observed number of $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ candidates, the expected bin-count is the sum of all signal and background components

$$\mu_i = \sum_j \mu_{ij} = N_{B^+}\mathcal{P}_{i,B^+\to\mu^+\nu_\mu\gamma} + N_{B_c^+}\mathcal{P}_{i,B_c^+\to\mu^+\nu_\mu\gamma} + \sum_{\text{background } b} N_b\mathcal{P}_{i,b}, \tag{6.5}$$

where N_j are the model parameters quantifying the observed number of candidates of the respective process j.

In this analysis, the PDFs to model signal and background processes are derived from simulation and data control samples with finite statistics. This introduces an uncertainty on each process-specific PDF and consequently the most probable model parameters, $\hat{\boldsymbol{\theta}}$, which is generally difficult to quantify. For this reason, a binned maximum likelihood fit is performed in which the observed data is modelled from templates rather than from analytic functions. This allows to account for the finite statistics of the simulation and data control samples by bin-wise uncertainties on the respective templates.

Beeston and Barlow have developed a method to incorporate the template uncertainties in the likelihood function [105] and account for them in the minimisation. For each bin, i, and each template, j, a nuisance parameter, θ_{ij}, is introduced, which modifies the bin count as $\mu_{ij} \to \theta_{ij}\mu_{ij}$. Poisson constraints on the nuisance parameters ensure that the bin counts are modified in a range that reflects the respective template uncertainty. While this method offers a statistically exact treatment of the template uncertainties, the large number of nuisance parameters makes it computationally very expensive.

Conway proposed a solution [106], in which the nuisance parameters are instead constrained by a Gaussian penalty term. The negative log-likelihood in a particular bin then reads

$$-\log\mathcal{L}_i = -n_i \log\left(\sum_j \theta_{ij}\mu_{ij}\right) + \sum_j \theta_{ij}\mu_{ij} + \sum_j \frac{(\theta_{ij}-1)^2}{2\sigma_{ij}^2}, \tag{6.6}$$

where σ_{ij} reflects the template uncertainty. Instead of fitting the nuisance parameters, they are solved analytically by setting the derivative of the negative log-likelihood with respect to the θ_{ij} to zero. Expanding the logarithm by a Taylor series to second order, this results in a quadratic equation for the θ_{ij} that is easily solved. This approximation drastically reduces the computational expenses in minimising the negative log-likelihood to find the remaining model parameters, N_j, that best match the observed data.

In this analysis, the observed numbers of $B^+ \to \mu^+\nu_\mu\gamma$ and $B_c^+ \to \mu^+\nu_\mu\gamma$ signal candidates are extracted from an extended binned maximum likelihood fit in the IMINUIT software package [107]. The templates for the corrected mass distributions of the signal candidates are derived from simulation in Section 6.2. Background templates are derived from data control samples in Section 6.3 and Section 6.4. Template uncertainties are incorporated in the likelihood function using the approximation from Conway implemented in the TEMPLATE cost function in IMINUIT.

6.1.1 Choice of observable

The corrected mass, $m_{corr}(\mu^+\gamma_{ee})$, computed without bremsstrahlung correction, is used to discriminate between signal and background candidates. This improves the separation

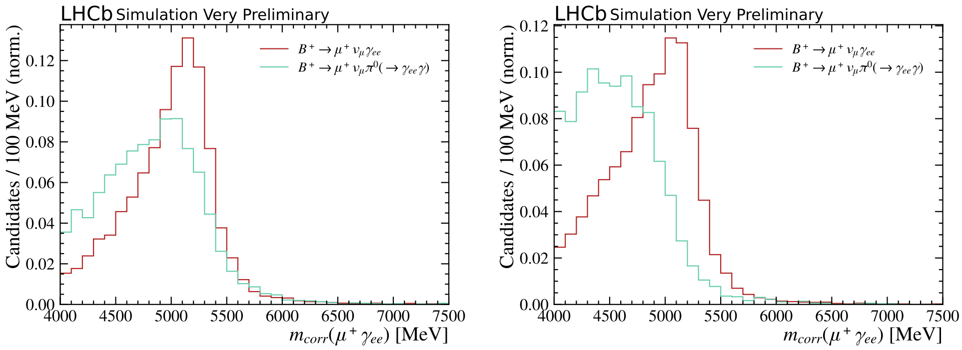


Figure 6.1: Normalised corrected mass distribution of simulated $B^+ \to \mu^+\nu_\mu\gamma$ signal candidates (red) and simulated $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ background candidates (green) with (left) and without (right) bremsstrahlung in the computation of m_{corr}.

of the $B^+ \to \mu^+\nu_\mu\gamma$ signal mass distribution from the background of $\pi^0 \to \gamma_{ee}\gamma$ decays. In Figure 6.1, the corrected mass distributions with and without bremsstrahlung for simulated $B^+ \to \mu^+\nu_\mu\gamma$ and $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ candidates are compared. The separation between signal and background can be significantly improved using the corrected mass without bremsstrahlung. This is important, considering that the branching fraction of $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ is at least an order of magnitude larger than for $B^+ \to \mu^+\nu_\mu\gamma$ decays. In the following, the symbol m_{corr} always denotes the corrected mass computed without bremsstrahlung correction.

The fit range is chosen between $4000\,\text{MeV} < m_{corr}(\mu^+\gamma_{ee}) < 7500\,\text{MeV}$. The data are binned in 35 uniform bins with a width of 100 MeV. While the analysis is still blinded in the signal region of $4500\,\text{MeV} < m_{corr}(\mu^+\gamma_{ee}) < 6500\,\text{MeV}$, this allows the upper and lower corrected mass sidebands to be used in a blinded fit as described in Section 6.5.2.

6.2 Signal modelling

The corrected mass templates for $B^+ \to \mu^+\nu_\mu\gamma$ and $B_c^+ \to \mu^+\nu_\mu\gamma$ signal candidates are taken from the respective simulation samples. For both components, the simulation from the leading-order calculation for $B^+ \to \mu^+\nu_\mu\gamma$ decays in Ref. [25] is used as the nominal model. The simulation is corrected for data/simulation differences according to the method discussed in Section 5.7.

Figure 6.2 shows the nominal corrected mass template for $B^+ \to \mu^+\nu_\mu\gamma$ candidates. The statistical uncertainty of the template is computed using a bootstrapping method, where the template is resampled 100 times.

In Figure 6.3, the corrected mass template for $B_c^+ \to \mu^+\nu_\mu\gamma$ candidates is shown. In the lower mass region, significant fluctuations in the template are observed. Such fluctuations with a scale of below 100 MeV are expected to be unphysical and can be attributed to the limited statistics of the simulation sample. A Gaussian filter is applied on the derived templates by convoluting the template with a Gaussian kernel to remove some of the unphysical fluctuations. The width of the Gaussian kernel is chosen at $\sigma = 100\,\text{MeV}$ to keep the bias to the region of the mass peak minimal. For the template of $B^+ \to \mu^+\nu_\mu\gamma$ decays, the Gaussian filter is not applied due to the larger simulation statistics available.

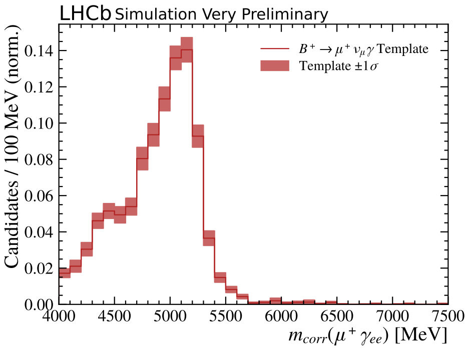


Figure 6.2: Nominal signal templates for $B^+ \to \mu^+\nu_\mu\gamma$ candidates.

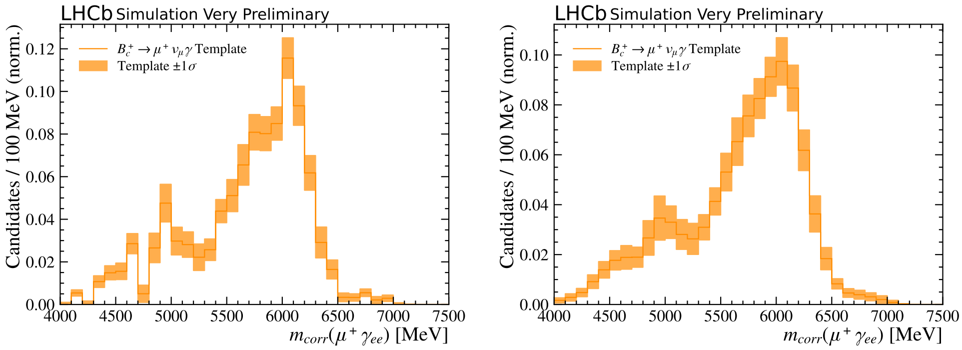


Figure 6.3: Corrected mass template for $B_c^+ \to \mu^+\nu_\mu\gamma$ candidates. For the plot on the right, a Gaussian filter is applied to remove some of the unphysical fluctuations in the template.

6.3 Background from $\pi^0/\eta \to \gamma_{ee}\gamma$ decays

The dominant source of background in this analysis arises from photon conversions not directly originating from a heavy B meson decay. Instead, they are produced in the decays of light neutral mesons, such as $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$. The additional photon can be missed or may be attached to one of the electrons as a bremsstrahlung photon. The large number of such background processes leading to a $\mu^+\gamma_{ee}$ final state can be categorised into:

- Semileptonic $b \to u$ decays, e.g. $B^+ \to \mu^+\nu_\mu\pi^0$ and $B^+ \to \mu^+\nu_\mu\eta$
- Semileptonic $b \to u$ cascade decays, e.g. $B^+ \to \mu^+\nu_\mu\rho^-(\to \pi^-\pi^0)$
- Semileptonic $b \to c \to u$ cascade decays, e.g. $B^0 \to D^-(\to \pi^0 X)\mu^+\nu_\mu$
- Combinatorial background, i.e. random association of π^0 and μ^+ not originating from the same B decay

The first category of semileptonic $b \to u$ decays has a signature particularly close to that of the signal decay and differs only by an additional photon that is not reconstructed. This type of background is expected to populate the region just below the B^+ signal peak. For the B_c^+ meson, there is no equivalent of this decay.

The second and third categories of semileptonic cascade decays are another type of partially reconstructed decays with additional charged or neutral particles in the final state. For semileptonic $B^{0,+}$ cascade decays, it is expected to populate a region in m_{corr} below the B^+ signal peak. However, partially reconstructed decays of heavier b-hadrons, such as B_s^0, B_c^+ or Λ_b^0, may also populate the region of the B^+ signal peak.
For combinatorial background candidates, the distribution of the corrected mass is generally not restricted, although it is expected to follow an approximate exponential distribution.

A large fraction of these background candidates is removed by means of the neutral isolation BDT and veto of $\eta \to \gamma_{ee}\gamma$ decays. Nonetheless, a sizeable fraction of background candidates is expected to pass the selection, which needs to be modelled for the fit of the corrected mass distribution in data. Simulating all sources of background from $\pi^0 \to \gamma_{ee}\gamma$ or $\eta \to \gamma_{ee}\gamma$ decays is computationally very expensive and considered not feasible. Instead, a data-driven approach is developed to obtain the corrected mass templates, which ensures all possible processes leading to background from π^0 and η decays are included by construction. The method is outlined below and further explained in the following sections.
The idea is to select clean samples of $\mu^+\pi^0(\to \gamma_{ee}\gamma)$ and $\mu^+\eta(\to \gamma_{ee}\gamma)$ candidates in data by explicitly reconstructing the $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ decays from an additional calorimeter photon, γ_{calo}, as explained in Section 6.3.1. To increase the available statistics in these samples, the requirement on the `PiBDT` is removed. Correctly reconstructed $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ candidates are identified by fitting the reconstructed diphoton invariant mass, $m(\gamma_{ee}\gamma_{calo})$, as detailed in Section 6.3.2. Using the ${}_s\mathcal{P}lot$ technique, background-subtracted samples of $\mu^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ and $\mu^+\eta(\to \gamma_{ee}\gamma_{calo})$ candidates are obtained.
These background-subtracted samples are used to derive the corrected mass templates for background from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ decays. As the efficiency of reconstructing the additional photon and resolving the π^0 and η decays potentially affects the corrected mass distribution, this effect is corrected for in Sections 6.3.3 and 6.3.4 using additional input from simulation. The efficiency of the `PiBDT`, which is removed in the selection of $\mu^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ and $\mu^+\eta(\to \gamma_{ee}\gamma_{calo})$ candidates, is accounted for in the same step, such that the derived corrected mass templates derived in Section 6.3.5 are representative of all background candidates passing the full selection. A thorough validation of the method is presented in Section 6.3.6.
With this data-driven approach, all sources of background from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ decays are included by construction. In addition, correcting for the efficiency of reconstructing the γ_{calo} and resolving the $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ decay and the efficiency of the `PiBDT`, the derived templates are also scaled to the number of expected background candidates in the final fit of the corrected mass distribution in data.

6.3.1 Reconstructing $\pi^0/\eta \to \gamma_{ee}\gamma_{calo}$ in data

Candidates of $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ decays are reconstructed from an additional calorimeter photon, γ_{calo}, as already introduced in the development of the neutral isolation BDT in Section 5.6.2. The γ_{calo} is selected from a neutral cluster in the ECAL within a cone around the γ_{ee} flight direction with radius, R, defined as

$$R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{6.7}$$

Within this radius, the photon with the maximum transverse momentum and $p_T > 75\,\text{MeV}$ is selected. Henceforth, this photon is referred to as *the* calorimeter photon or just γ_{calo}.

Choice of cone size $\boldsymbol{R}$

The choice of the cone size affects the efficiency with which $\pi^0 \to \gamma_{ee}\gamma_{calo}$ or $\eta \to \gamma_{ee}\gamma_{calo}$ decays are resolved. In the case the cone size is chosen too narrow, the additional photon may fall outside of the search window, resulting in a lower efficiency. In the case the cone size is chosen too wide, the chance of selecting a random photon with large p_T increases, again reducing the efficiency of selecting the correct photon.
Since the efficiency of resolving $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ decays from the correct photon is corrected for in the later steps, the choice of the cone size R is not critical for the validity of the background modelling. Nonetheless, a simple optimisation of the choice of R is performed on simulated $B^+ \to \mu^+\nu_\mu\pi^0$ and $B^+ \to \mu^+\nu_\mu\eta$ candidates. Based on the distribution in $m(\gamma_{ee}\gamma_{calo})$, a cone of size $R = 0.3$ is chosen to resolve a large number of $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ decays, while keeping the level of background from random photons low.

6.3.2 Fitting $m(\gamma_{ee}\gamma_{calo})$

Correctly resolved $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ candidates are identified by fitting the reconstructed $m(\gamma_{ee}\gamma_{calo})$. An unbinned maximum likelihood fit is performed using the ZFIT [108] software package. The fit range is chosen from 20 MeV to 700 MeV to cover the region populated by the π^0 and η with masses of $m_{\pi^0} = 135\,\text{MeV}$ and $m_\eta = 548\,\text{MeV}$, respectively [5]. To account for the varying mass resolution, the fit is performed in four bins of di-photon momentum, $p(\gamma_{ee}\gamma_{calo})$. The momentum is calculated without bremsstrahlung correction on the converted photon to avoid double counting of momentum in the case the selected γ_{calo} is assigned as a bremsstrahlung photon to one of the electrons. The four bins in $p(\gamma_{ee}\gamma_{calo})$ are chosen to be approximately equally populated with bin edges set to [20, 40, 60, 80, 300] GeV. The fit model comprises the following four components:

1. Combinatorial background, i.e. a random combination of γ_{ee} and γ_{calo}
2. Events in which the selected γ_{calo} is the bremsstrahlung photon of one of the electrons
3. True $\pi^0 \to \gamma_{ee}\gamma_{calo}$ decays
4. True $\eta \to \gamma_{ee}\gamma_{calo}$ decays

The fit models for all four components are obtained from simulation. A brief overview is given here, which is expanded on in the following paragraphs. In a first step, the di-photon invariant mass of simulated $B^+ \to \mu^+\nu_\mu\gamma$ candidates is fitted. In this sample, the additional γ_{calo} is either a bremsstrahlung photon or of combinatorial nature. Consequently,

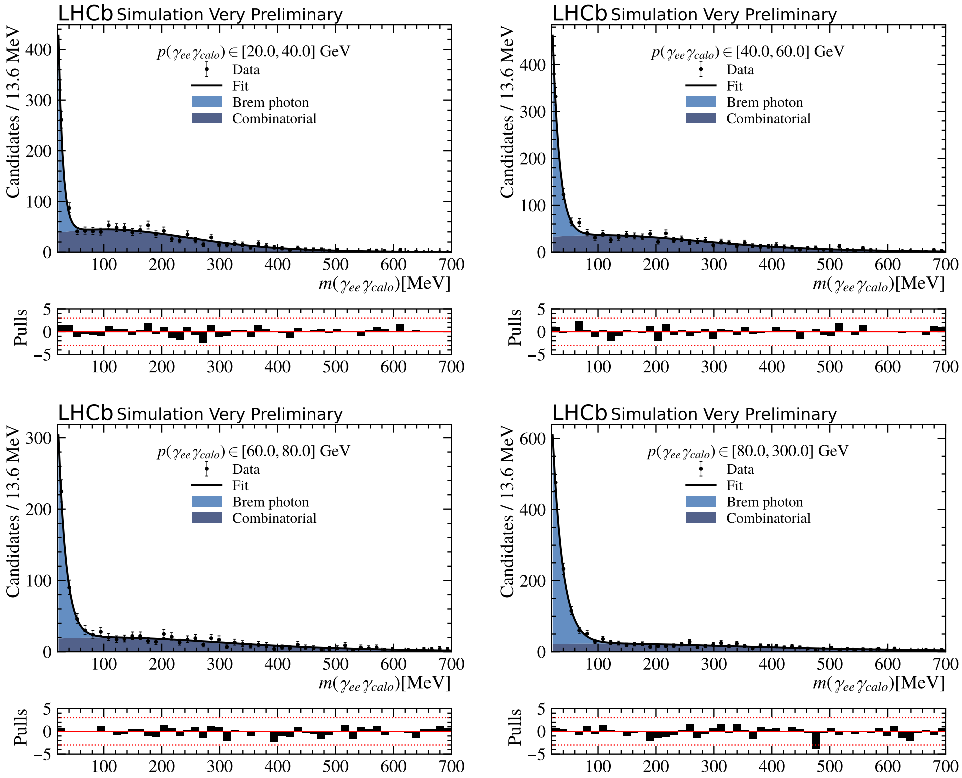


Figure 6.4: Fit in $m(\gamma_{ee}\gamma_{calo})$ of simulated $B^+ \to \mu^+\nu_\mu\gamma$ candidates in four bins of $p(\gamma_{ee}\gamma_{calo})$.

the fit is used to derive the model for both background components. In a second step, the simulation samples of $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ and $B^+ \to \mu^+\nu_\mu\eta(\to \gamma_{ee}\gamma)$ are used to derive the model for $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ candidates, respectively. Using these signal and background models, the fit of reconstructed candidates in data is performed.

Modelling of bremsstrahlung and combinatorial backgrounds

For simulated $B^+ \to \mu^+\nu_\mu\gamma$ candidates, the additional γ_{calo} is expected to either be a bremsstrahlung photon or a random photon. Thus, a fit of the $m(\gamma_{ee}\gamma_{calo})$ distribution is used to simultaneously extract the mass distributions for background from bremsstrahlung photons and combinatorial background.
In the case that the selected γ_{calo} is the true bremsstrahlung photon of one of the electrons, the di-photon invariant mass is close to zero and is modelled by an exponential function with slope λ_{Brem}.
For combinatorial background, the typical model of an exponential with slope λ_{comb} is modified by a threshold function due to the restricted phase-space. The threshold function is parametrised by a generalised logistic function, defined as

$$g(x) = \frac{1}{(1+e^{-\frac{x-\mu}{\sigma}})^{\nu}}, \tag{6.8}$$

which is the minimal extension of the logistic function that accurately describes the observed threshold effect. The parameters μ and σ set the central position and growth rate of the logistic function, with the parameter ν defining at which asymptote maximum growth occurs. The model for the combinatorial background then reads

$$f(x) = \frac{e^{-\lambda_{comb}x}}{(1+e^{-\frac{x-\mu}{\sigma}})^{-2\sigma\lambda_{comb}}}, \tag{6.9}$$

where the parameter is set to $\nu = -2\sigma\lambda_{comb}$ to reduce the number of free parameters in the fit by one and fix the maximum at $x = \mu$ irrespective of the slope λ_{comb}. Comparing the distribution of combinatorial background in data and simulation, the position of the maximum is found to agree well, while the slope λ_{comb} shows some discrepancy. Hence, the parametrisation of the model is chosen such that λ_{comb} can be left floating in the fit to the data, while the remaining model parameters are fixed based on the simulation.
In each of the four bins in $p(\gamma_{ee}\gamma_{calo})$, the two slopes λ_{Brem} and λ_{comb} are fitted independently. The parameters μ and σ of the threshold function are shared across all four momentum bins. The fits to simulated $B^+ \to \mu^+\nu_\mu\gamma$ candidates in all four momentum bins are shown in Figure 6.4.

Modelling of $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$

The di-photon invariant mass for $\pi^0 \to \gamma_{ee}\gamma_{calo}$ candidates is modelled with a double-sided Crystal Ball function (see Eq. 3.5). In contrast, $\eta \to \gamma_{ee}\gamma_{calo}$ candidates are modelled with a single-sided Crystal Ball function with a power-law tail only on the left-hand side. The model parameters are extracted from a simultaneous fit to the $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ and $B^+ \to \mu^+\nu_\mu\eta(\to \gamma_{ee}\gamma)$ simulation samples, which is shown in Figure 6.5. The left-hand tail parameters, α_L and n_L, in each bin of $p(\gamma_{ee}\gamma_{calo})$ are shared between the shapes for π^0- and η decays. The components of bremsstrahlung and combinatorial background are fixed from the fit to the simulation of $B^+ \to \mu^+\nu_\mu\gamma$ decays.

Fit to Data

The di-photon invariant mass in data is fitted using an extended unbinned maximum likelihood fit. The components for $\pi^0 \to \gamma_{ee}\gamma_{calo}$ candidates, $\eta \to \gamma_{ee}\gamma_{calo}$ candidates, background from bremsstrahlung and combinatorial background are included with the respective models derived from simulation. All model parameters are fixed from the fit to simulation, except for the slope λ_{comb} and the right-hand tail parameter α_R, of the π^0 mass peak. In addition, a common shift, $\Delta\mu$, and scaling of the width, $\Delta\sigma$, are applied to the Crystal Ball functions modelling the π^0 and η mass peaks to account for residual differences in the photon reconstruction between data and simulation. The resulting fit of $m(\gamma_{ee}\gamma_{calo})$ in data is shown in Figure 6.6.

Unfolding event distributions using ${}_s\mathcal{P}lot$

The ${}_s\mathcal{P}lot$ technique [101] allows for unfolding two overlapping distributions based on the separation in an uncorrelated discriminating variable. In this case, it is applied to obtain distributions of background-subtracted $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ candidates in other decay variables. Based on the maximum likelihood fit in the discriminating variable $m(\gamma_{ee}\gamma_{calo})$, per-event weights, so-called *sWeights*, are calculated using the HEPSTATS library of the SCIKIT HEP project [109]. Applying these sWeights to data allows to

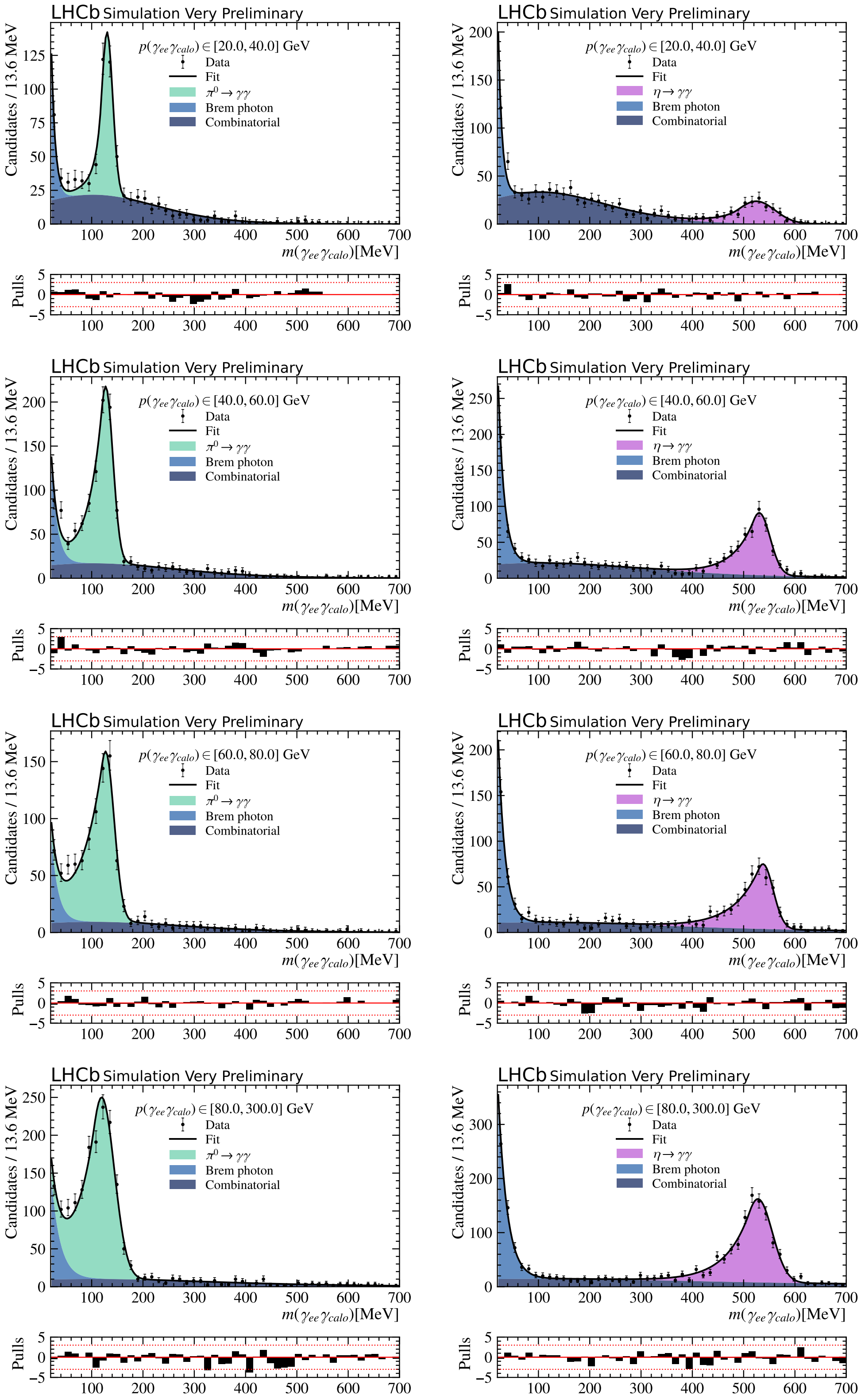


Figure 6.5: Simultaneous fit of $m(\gamma_{ee}\gamma_{calo})$ in simulated $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ (left) and $B^+ \to \mu^+\nu_\mu\eta(\to \gamma_{ee}\gamma)$ (right) candidates in four bins of $p(\gamma_{ee}\gamma_{calo})$ (top to bottom).

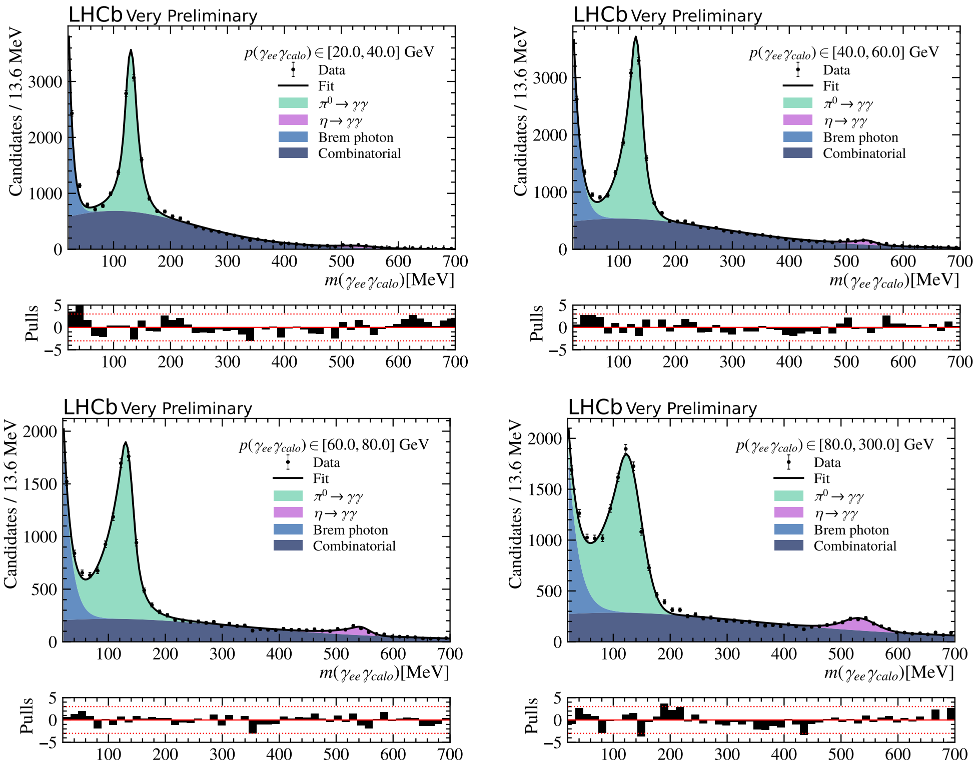


Figure 6.6: Fit to $m(\gamma_{ee}\gamma_{calo})$ in data in four bins of $p(\gamma_{ee}\gamma_{calo})$.

statistically recover the distributions of $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ candidates in other variables, given that the variable of interest and the discriminating variable, $m(\gamma_{ee}\gamma_{calo})$, are uncorrelated.

6.3.3 Efficiency of resolving $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ in data

Naively, the corrected mass templates for background from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ decays could be obtained from the corrected mass distribution in the background-subtracted $\mu^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ and $\mu^+\eta(\to \gamma_{ee}\gamma_{calo})$ data samples. However, reconstructing the additional calorimeter photon and resolving the $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ decays is subject to an additional reconstruction efficiency, which potentially distorts the corrected mass distribution. To compensate for this effect, a novel strategy is developed in which the efficiency of reconstructing the γ_{calo} and resolving the π^0 and η decays[1] is corrected for.

The strategy, illustrated in Figure 6.7, is briefly outlined here and explained in greater detail in the following paragraphs. The photon reconstruction efficiency depends strongly on the transverse momentum, $p_T(\gamma_{calo})$. Therefore, the background-subtracted samples (1) are corrected for the photon reconstruction efficiency as a function of $p_T(\gamma_{calo})$ by applying per-event weights (2) to yield the original $p_T(\gamma_{calo})$ distribution in data (3). The cor-

---

[1] For simplicity, the combined efficiency of reconstructing an additional calorimeter photon from a $\pi^0 \to \gamma_{ee}\gamma_{calo}$ or $\eta \to \gamma_{ee}\gamma_{calo}$ decay and resolving the respective decays in $m(\gamma_{ee}\gamma_{calo})$ is referred to as *photon reconstruction efficiency.*

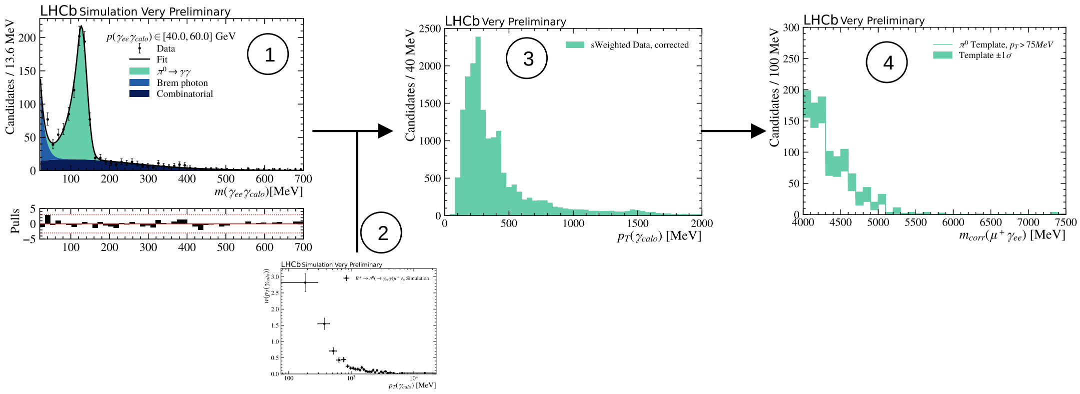


Figure 6.7: Illustration of the data-driven method to obtain the corrected mass templates for background from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ decays.

rected mass templates are then derived from the distribution in $m_{corr}(\mu^+\gamma_{ee})$ of efficiency-corrected $\mu^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ and $\mu^+\eta(\to \gamma_{ee}\gamma_{calo})$ candidates in data (4).

Correcting the photon reconstruction efficiency

In this paragraph, the correction of the photon reconstruction efficiency as a function of $p_T(\gamma_{calo})$ is established (steps 1 → 2 → 3 in Figure 6.7). The strong dependence on $p_T(\gamma_{calo})$ is induced by the energy resolution of the ECAL as well as the selection of γ_{calo} as the photon with maximum p_T in a certain region of the ECAL.

The photon reconstruction efficiency is derived separately for $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ from simulation of $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ and $B^+ \to \mu^+\nu_\mu\eta(\to \gamma_{ee}\gamma)$ candidates, respectively. The same reconstruction of the additional calorimeter photon is used for simulated candidates and the ${}_s\mathcal{P}lot$ technique is applied based on the fit to $m(\gamma_{ee}\gamma_{calo})$, shown in Figure 6.5. From the background-subtracted simulation samples of correctly resolved $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ decays, the photon reconstruction efficiency, $\varepsilon_{\gamma_{calo}}$, is derived in bins of $p_T(\gamma_{calo})$ as the fraction of resolved candidates in simulation

$$\varepsilon_{\gamma_{calo}} = \frac{N^{MC}_{resolved}}{N^{MC}}. \tag{6.10}$$

As mentioned in the beginning, the idea is to use large samples of background-subtracted $\mu^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ and $\mu^+\eta(\to \gamma_{ee}\gamma_{calo})$ candidates, selected without `PiBDT` requirement, to model the corrected mass distribution of candidates passing the `PiBDT` selection. The efficiency $\varepsilon_{\gamma_{calo}\&\texttt{PiBDT}}$ takes into account the additional efficiency of the `PiBDT` selection and is defined as

$$\varepsilon_{\gamma_{calo}\&\texttt{PiBDT}} = \frac{N^{MC}_{resolved}}{N^{MC}(\texttt{PiBDT} > 0.50)}. \tag{6.11}$$

The inverse of this efficiency defines the efficiency correction weight, $w_{\gamma_{calo}}$, and is computed in 30 uniformly populated bins of $p_T(\gamma_{calo})$ as shown in Figure 6.8 for the processes $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$. With these weights applied, the distributions of decay variables from resolved and background-subtracted $\mu^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ and $\mu^+\eta(\to \gamma_{ee}\gamma_{calo})$ reflect the distributions of all π^0 and η background processes, not just those resolved from an additional γ_{calo}, passing the `PiBDT` selection. The effect of the efficiency correction on the $p_T(\gamma_{calo})$ distribution in background-subtracted $\mu^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ candidates is illustrated in Figure 6.9.

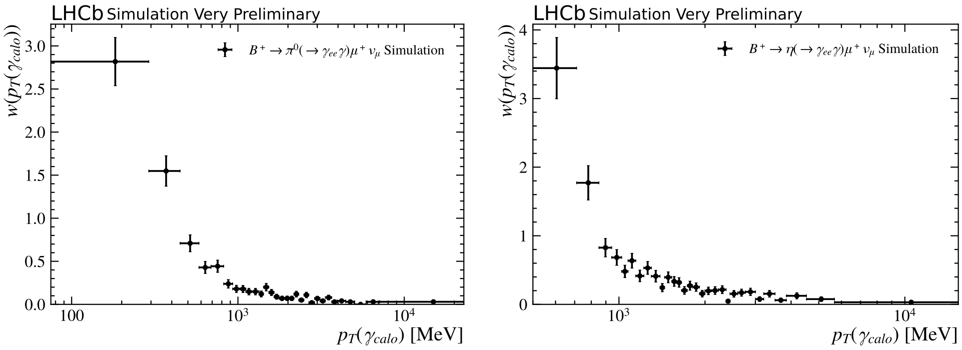


Figure 6.8: Correction weights for reconstructing and resolving $\pi^0 \to \gamma_{ee}\gamma_{calo}$ (left) and $\eta \to \gamma_{ee}\gamma_{calo}$ (right), including the efficiency of the `PiBDT` selection.

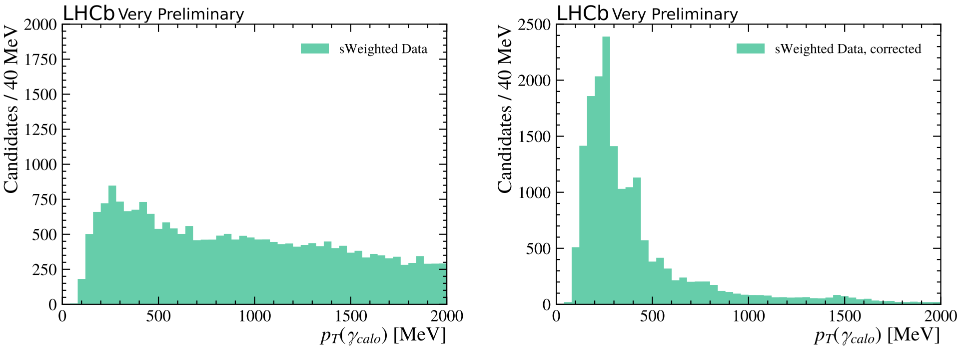


Figure 6.9: Reconstructed distribution of $p_T(\gamma_{calo})$ in background subtracted $\pi^0 \to \gamma_{ee}\gamma_{calo}$ candidates in data (left). Applying the efficiency correction weights shown in Figure 6.8, the expected distribution shown on the right can be recovered.

Constructing the corrected mass templates

The final step (step 4 in Figure 6.7) is the construction of the corrected mass templates for background candidates from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ decays. They are derived from the distribution of background-subtracted and efficiency-corrected samples of resolved $\mu^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ and $\mu^+\eta(\to \gamma_{ee}\gamma_{calo})$ candidates in data. The resulting templates are shown in Figure 6.10. By construction, the templates are scaled to the expected number of background candidates in data, which are 920 ± 50 candidates from $\pi^0 \to \gamma_{ee}\gamma$ decays and 21 ± 7 candidates from $\eta \to \gamma_{ee}\gamma$ decays. The statistical uncertainty on the templates is derived from a bootstrapping technique, where the underlying data samples are resampled one hundred times.

Limitations of the photon efficiency correction

The method of obtaining the corrected mass templates by applying efficiency correction weights to resolved $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ candidates has one fundamental limitation. As shown in Figure 6.8, the weights $w_{\gamma_{calo}}$ are not defined for very low transverse momentum. This is caused by the limitation of the LHCb experiment to reconstruct and resolve $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ decays with very soft photons. In the case of $\pi^0 \to \gamma_{ee}\gamma_{calo}$, the limit at a transverse momentum of 75 MeV is given by the minimum p_T with which neutral clusters in the ECAL are reconstructed. For $\eta \to \gamma_{ee}\gamma_{calo}$

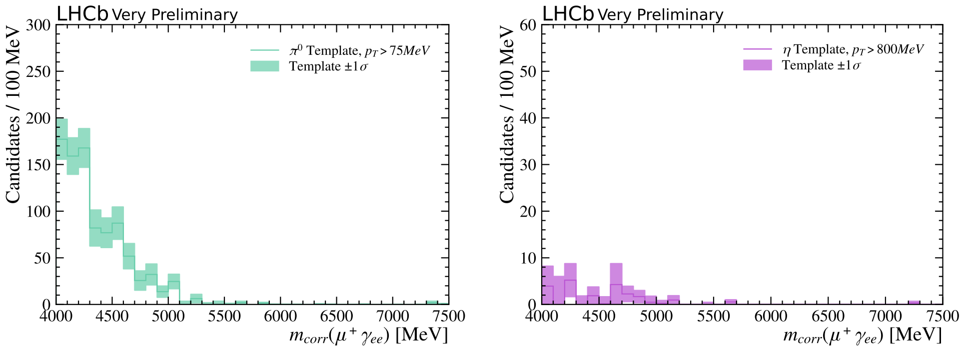


Figure 6.10: Corrected mass templates for backgrounds from $\pi^0 \to \gamma_{ee}\gamma$ (left) and $\eta \to \gamma_{ee}\gamma$ (right) with efficiency correction weights applied.

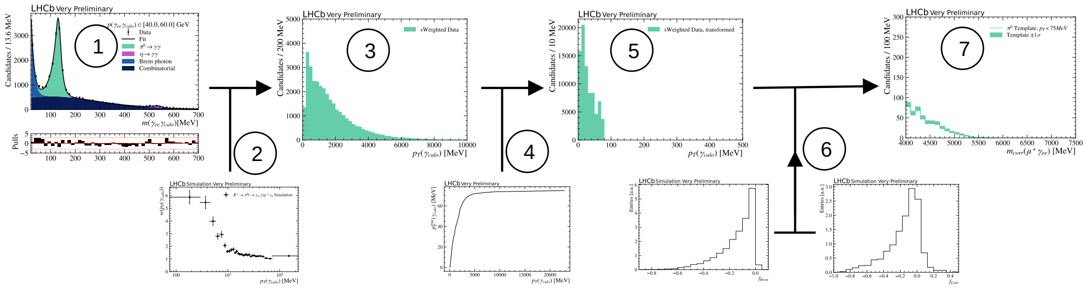


Figure 6.11: Illustration of the method developed to obtain the corrected mass templates for $\pi^0 \to \gamma_{ee}\gamma$ background candidates for γ_{calo} with low transverse momentum.

decays, the weights $w_{\gamma_{calo}}$ are only defined for $p_T(\gamma_{calo}) > 800\,\text{MeV}$, as the decay is not resolved for softer photons. Consequently, the derived templates are only representative of background candidates from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ decays in which the additional photon has sufficient transverse momentum. Background candidates in which the additional photon from a $\pi^0 \to \gamma_{ee}\gamma$ or $\eta \to \gamma_{ee}\gamma$ decay is too soft for the method to apply are referred to as *soft photon background* candidates and are modelled using a different approach in the following.

6.3.4 Extrapolation for soft photon background candidates

Soft photon background candidates from $\pi^0 \to \gamma_{ee}\gamma$ or $\eta \to \gamma_{ee}\gamma$ need to be treated differently, as one cannot select a representative sample of these decays in data simply because the photon has too low transverse momentum to be reconstructed. Instead, $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ decays reconstructed from calorimeter photons with higher transverse momentum are utilised to model soft photon background candidates. To achieve this, a novel method is developed to extrapolate the decay kinematics to lower $p_T(\gamma_{calo})$, resulting in a data-driven model of the corrected mass for soft photon backgrounds. The resulting model is then combined with the templates derived in the previous section to yield a complete model of the corrected mass for background from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ decays of the entire $p_T(\gamma_{calo})$ range.

The following presents a brief overview of the methodology, which is expanded upon in the subsequent sections. To simplify the discussion, the method is explained for $\pi^0 \to \gamma_{ee}\gamma$ decays with $p_T(\gamma_{calo}) < 75\,\text{MeV}$, but it works analogously for $\eta \to \gamma_{ee}\gamma$ decays in the

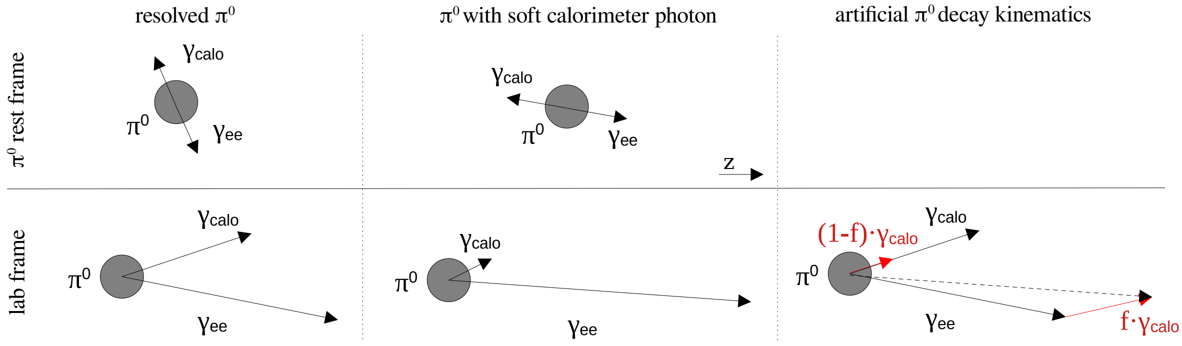


Figure 6.12: Illustration of a $\pi^0 \to \gamma_{ee}\gamma_{calo}$ decay in the centre-of-mass frame (top) and the lab frame boosted along the z-axis (bottom) with a resolved π^0 from an energetic γ_{calo} (left) and unresolved π^0 with soft γ_{calo} (centre). The plot on the right illustrates how resolved $\pi^0 \to \gamma_{ee}\gamma_{calo}$ decays are used to model background candidates with soft γ_{calo} transferring a fraction of the reconstructed γ_{calo} momentum to the γ_{ee}.

range $p_T(\gamma_{calo}) < 800\,\text{MeV}$ unless otherwise noted. Key results are always presented for both $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ decays.
The method illustrated in Figure 6.11 is based on the same background-subtracted data samples of $\mu^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ candidates from the previous section, reconstructed with larger $p_T(\gamma_{calo})$ (1). The previously derived weights to correct the photon efficiency are applied (2) to yield the original $p_T(\gamma_{calo})$ distribution of $\pi^0 \to \gamma_{ee}\gamma_{calo}$ decays in data for $p_T(\gamma_{calo}) > 75\,\text{MeV}$ (3). Using input from simulation, a transfer function is defined (4), which transforms the original $p_T(\gamma_{calo})$ distribution in data to the expected distribution $p_T(\gamma_{calo}) < 75\,\text{MeV}$ (5). This is achieved by creating pseudo-candidates with artificial γ_{ee} and γ_{calo} kinematics that reflect the anticipated kinematics of background candidates with soft photons, taking into account resolution effects (6). The corrected mass, $m_{corr}(\mu^+\gamma_{ee})$, is recomputed with the modified γ_{ee} four-momentum to yield the corrected mass templates for soft photon background candidates from $\pi^0 \to \gamma_{ee}\gamma$ decays (7).

Extrapolation of decay kinematics to low $p_T(\gamma_{calo})$

In the selection of candidates (see Chapter 5), a cut on the reconstructed momentum of the converted photon $p_{track}(\gamma_{ee}) > 20\,\text{GeV}$ is imposed. This directly implies that any π^0 that contributes as a source of background must be at least of the same momentum. Therefore, background candidates with soft additional photons involve the decay of an energetic π^0 in which the γ_{ee} carries a large fraction of the momentum and the γ_{calo} carries a small fraction of the π^0 momentum[2] as illustrated in the centre column of Figure 6.12. In case the calorimeter photon carries a slightly larger fraction of the π^0 momentum in the lab frame, the γ_{calo} can be reconstructed and the $\pi^0 \to \gamma_{ee}\gamma_{calo}$ candidate can be resolved as illustrated in the left column of Figure 6.12. This motivates emulating the decay kinematics of π^0 decays with soft photons from resolved $\pi^0 \to \gamma_{ee}\gamma_{calo}$ decays, reconstructed from a more energetic γ_{calo}. The idea is to create a pseudo-candidate with artificial γ_{ee} kinematics by transferring a fraction, f, of the reconstructed γ_{calo} four-

---

[2] In the π^0 centre-of-mass frame, the two photons have the same absolute momentum. In the lab frame, the momenta are very different if the two photons are produced along the axis of the π^0 boost.

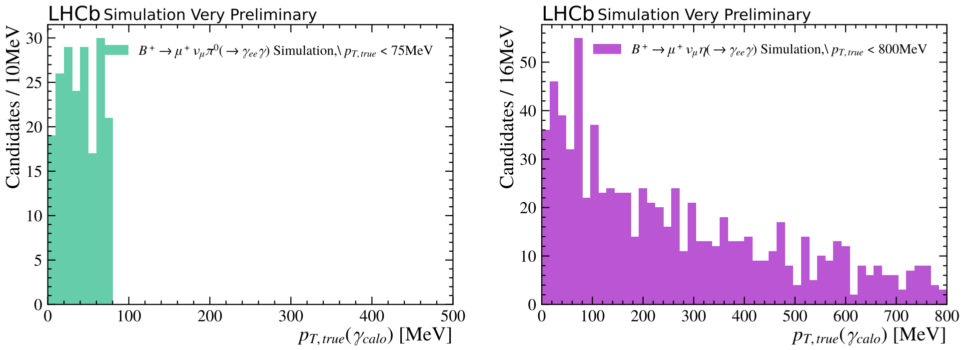


Figure 6.13: True p_T-distribution of soft photons in simulation of $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ candidates (left) and $B^+ \to \mu^+\nu_\mu\eta(\to \gamma_{ee}\gamma)$ candidates (right). Only the region of soft photon backgrounds for the respective decay is shown.

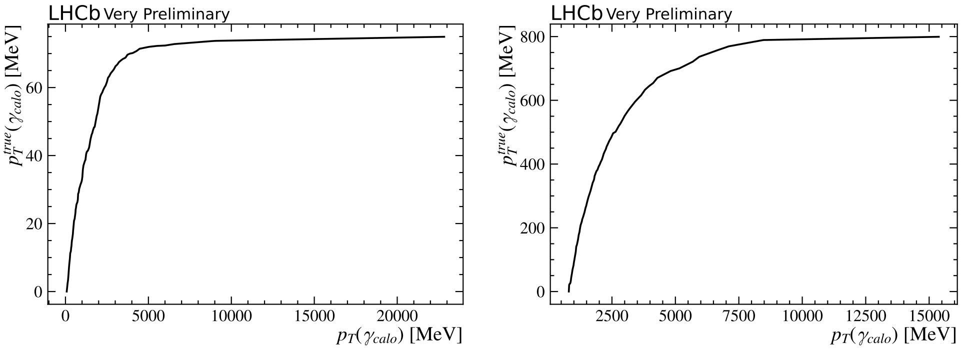


Figure 6.14: Transfer function relating the reconstructed $p_T(\gamma_{calo})$ of efficiency-corrected and background-subtracted $\pi^0 \to \gamma_{ee}\gamma_{calo}$ candidates to the $p_T^{true}(\gamma_{calo})$ in the region of soft photon backgrounds, established on $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ (left) and $B^+ \to \mu^+\nu_\mu\eta(\to \gamma_{ee}\gamma)$ (right).

momentum to the reconstructed γ_{ee} four-momentum as illustrated on the right-hand side of Figure 6.12. The fraction, f, is chosen such that the modified transverse momentum of the calorimeter photon $(1-f)p_T(\gamma_{calo})$ matches the expected distribution in data below 75 MeV.

Transforming the $p_T(\gamma_{calo})$ distribution

The momentum distribution of soft photons in the region $p_T(\gamma_{calo}) < 75$ MeV is a priori unknown and cannot be inferred from data. However, in simulated candidates of $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ and $B^+ \to \mu^+\nu_\mu\eta(\to \gamma_{ee}\gamma)$ decays, the *true* event information is available. In particular, the true p_T distribution of the additional photon in the region below 75 MeV (800 MeV for $\eta \to \gamma_{ee}\gamma$) can be accessed, which is shown in Figure 6.13.
Assuming the same p_T distribution of soft photon backgrounds in data and simulation risks a dependence on the simulated π^0 kinematics. Especially, as the simulated π^0 kinematics are only representative of the kinematics in $B^+ \to \mu^+\nu_\mu\pi^0$ decays; they are not necessarily representative of the kinematics of all possible $\pi^0 \to \gamma_{ee}\gamma$ decays contributing as a source of background.
A weaker assumption is to expect the same relation to hold between the p_T distributions above and below 75 MeV in data and simulation. In simulation, such a relation is established between the *true* p_T distribution below 75 MeV and the reconstructed p_T distribu-

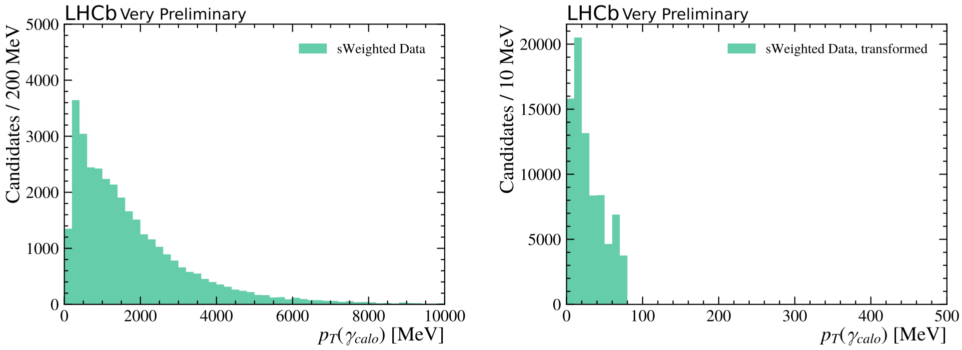


Figure 6.15: Reconstructed distribution of $p_T(\gamma_{calo})$ in efficiency-corrected and background-subtracted $\pi^0 \to \gamma_{ee}\gamma_{calo}$ decays (left). Applying the transfer function, the expected p_T distribution of soft photon background candidates in data can be estimated (right).

tion of resolved candidates in efficiency-corrected and background-subtracted simulation. This relation defines the *transfer function* shown in Figure 6.14. Evaluating the transfer function on efficiency-corrected and background-subtracted $\pi^0/\eta \to \gamma_{ee}\gamma_{calo}$ candidates in data, each candidate is assigned a value of $p_{T,modified}$ in the region below 75 MeV. The distribution of these assigned $p_{T,modified}$ reflects the expected true $p_T(\gamma_{calo})$ distribution of soft photon background candidates in data and is shown in Figure 6.15.

Creating artificial $\pi^0 \to \gamma_{ee}\gamma$ decay kinematics

From the expected $\pi^0 \to \gamma_{ee}\gamma$ decay kinematics for $p_T(\gamma_{calo}) < 75$ MeV, the kinematics of the resolved $\pi^0 \to \gamma_{ee}\gamma_{calo}$ candidates in data are modified to create artificial pseudo-candidates that match the expected kinematics of soft photon background candidates. As illustrated on the right side of Figure 6.12, the idea is to transfer a fraction, f, of the reconstructed γ_{calo} momentum to the converted photon, to resemble the expected kinematics of decays with soft photons. The fraction f is naively defined as

$$f = \frac{p_T(\gamma_{calo}) - p_{T,modified}(\gamma_{calo})}{p_T(\gamma_{calo})}. \tag{6.12}$$

While this definition is a good starting point for creating artificial decay kinematics, two additional resolution effects need to be taken into account.

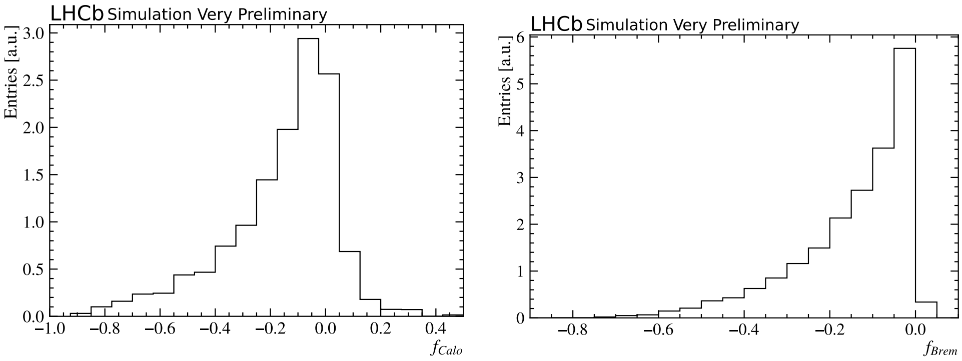
Firstly, the transverse momentum of the calorimeter photon for revolved $\pi^0 \to \gamma_{ee}\gamma_{calo}$ decays has the tendency to be overestimated. Transferring the fraction f of the reconstructed γ_{calo} momentum to the converted photon would create an artificial decay kinematic with too energetic γ_{ee}. This is quantified in the calorimeter resolution f_{Calo} defined as

$$f_{Calo} = \frac{p_{T,true}(\gamma_{calo}) - p_{T,reco}(\gamma_{calo})}{p_{T,reco}(\gamma_{calo})}. \tag{6.13}$$

The calorimeter resolution is obtained from simulated $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ candidates and is shown in Figure 6.16a.

Secondly, the two electrons from the photon conversion are subject to energy loss due to bremsstrahlung. When increasing the momentum of the converted photon, the fractional momentum loss due to bremsstrahlung, f_{Brem}, needs to be considered. It is defined as

$$f_{Brem} = \frac{p_{T,track}(\gamma_{ee}) - p_{T,true}(\gamma_{ee})}{p_{T,true}(\gamma_{ee})}, \tag{6.14}$$



(a) (b)

Figure 6.16: Calorimeter resolution f_{Calo} (a) and fractional energy loss due to bremsstrahlung of converted photons (b) derived from simulation of $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$.

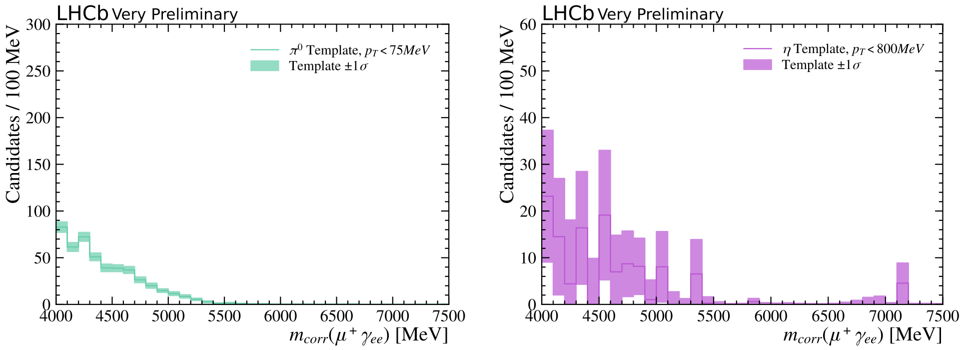


Figure 6.17: Corrected mass templates for backgrounds from $\pi^0 \to \gamma_{ee}\gamma$ (left) and $\eta \to \gamma_{ee}\gamma$ (right) decays with soft photons.

where $p_{T,true}$ denotes true photon momentum and $p_{T,track}$ is the momentum reconstructed from the tracking system, i.e. without bremsstrahlung correction. The fractional momentum loss is obtained from simulation of $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ decays and is shown in Figure 6.16b.
Taking into account these resolution effects, the four-momentum, p^α, of the converted photon is modified according to

$$p^\alpha_{track}(\gamma_{ee}) \to f \times (1 + f_{Calo}) \times (1 + f_{Brem}) \times p^\alpha(\gamma_{calo}) + p^\alpha_{track}(\gamma_{ee}). \tag{6.15}$$

Pseudo-candidates with modified γ_{ee} kinematics are created for all $\mu^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ candidates resolved in data. For these candidates, the corrected mass, $m_{corr}(\mu^+\gamma_{ee})$, is recalculated using the modified γ_{ee} kinematics. The resulting corrected mass distribution of pseudo-candidates is representative of all background candidates from $\pi^0 \to \gamma_{ee}\gamma$ decays with additional photons too soft to be reconstructed. The corrected mass templates are shown in Figure 6.17. The uncertainties are derived from a bootstrapping technique, where the sample of resolved $\mu^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ candidates is resampled one hundred times.

Expected number of soft photon background candidates

To scale the derived templates to the expected number of background candidates with soft additional photons, the expected yields in data are estimated. The assumption is to expect the same relation between the number of $\pi^0 \to \gamma_{ee}\gamma$ events with $p_T(\gamma)$ above and below 75 MeV in data and simulation. Special care needs to be taken to include the selection on the `PiBDT` in the expected number of events, as it is not applied when selecting the resolved $\mu^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ candidates in data. The relation between the number of $\pi^0 \to \gamma_{ee}\gamma$ candidates with $p_T(\gamma_{calo})$ below 75 MeV passing the `PiBDT` selection and the number above the respective p_T threshold is defined in simulation as

$$\begin{aligned} f_\pi &= \frac{N_{MC}(p_{T,true}(\gamma_{calo}) < 75\,\text{MeV}; \texttt{PiBDT} > 0.50)}{N_{MC}(p_T(\gamma_{calo}) > 75\,\text{MeV})}, \\ f_\eta &= \frac{N_{MC}(p_{T,true}(\gamma_{calo}) < 800\,\text{MeV}; \texttt{PiBDT} > 0.50)}{N_{MC}(p_T(\gamma_{calo}) > 800\,\text{MeV})}. \end{aligned} \tag{6.16}$$

The relative fractions are evaluated using the simulation of $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ and $B^+ \to \mu^+\nu_\mu\eta(\to \gamma_{ee}\gamma)$ decays, respectively. The expected number of soft photon background candidates in data is then computed by multiplying the relative fractions $f_{\pi/\eta}$ with the number of resolved $\mu^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ and $\mu^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ candidates in data. This results in an expected number of 470 ± 20 soft photon background candidates from $\pi^0 \to \gamma_{ee}\gamma$ and 120 ± 20 from $\eta \to \gamma_{ee}\gamma$ decays. The corrected mass templates shown in Figure 6.17 are scaled to the respective number of background candidates.

6.3.5 Templates for background from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta\gamma_{ee}\gamma$

The corrected mass templates for background candidates from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ decays derived in Sections 6.3.3 and 6.3.4 are orthogonal to one another as they describe backgrounds from two exclusive regions in $p_T(\gamma_{calo})$. As the templates in both regions are scaled to the expected number of background candidates in data, they can be combined for a complete model of all sources of background from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ decays, respectively. They are shown in Figure 6.18 and the number of background candidates are predicted to be 1390 ± 60 and 145 ± 25, respectively.

In the background templates, in particular the one to model $\eta \to \gamma_{ee}\gamma$ backgrounds, significant fluctuations are observed. Such fluctuations with a scale below 100 MeV are unphysical due to the absence of beauty hadron decays of the type $H_b \to \mu^+\pi^0(\to \gamma_{ee}\gamma)$. In fact, the narrowest structures in the background template are expected from $B^+ \to \mu^+\nu_\mu\pi^0$ decays, which have the best expected corrected mass resolution of all background processes. Hence, the background templates should not show any structures with a resolution much below the corrected mass resolution of $B^+ \to \mu^+\nu_\mu\pi^0$ decays. This is incorporated by applying a Gaussian filter to the background templates with $\sigma = 100\,\text{MeV}$, which removes unphysical fluctuations. The width of the Gaussian kernel is chosen such that the corrected mass distribution of $B^+ \to \mu^+\nu_\mu\pi^0$ decays remains unchanged.

The final corrected mass templates for background from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ are shown in Figure 6.19. The smoothed templates with the Gaussian filter applied offer a much more reliable way of modelling the background components in data and are the nominal templates to be used in the fit to the corrected mass distribution. The systematic uncertainty linked to this smoothing is discussed in Chapter 9.

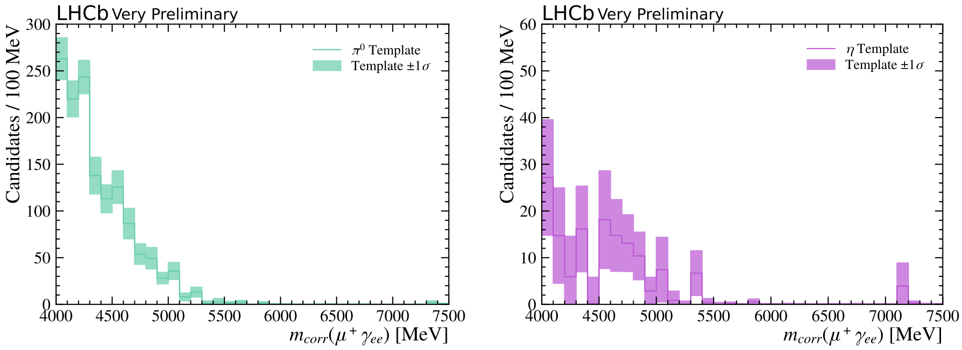


Figure 6.18: Corrected mass templates for background from $\pi^0 \to \gamma_{ee}\gamma$ (left) and $\eta \to \gamma_{ee}\gamma$ (right) scaled to the expected yield.

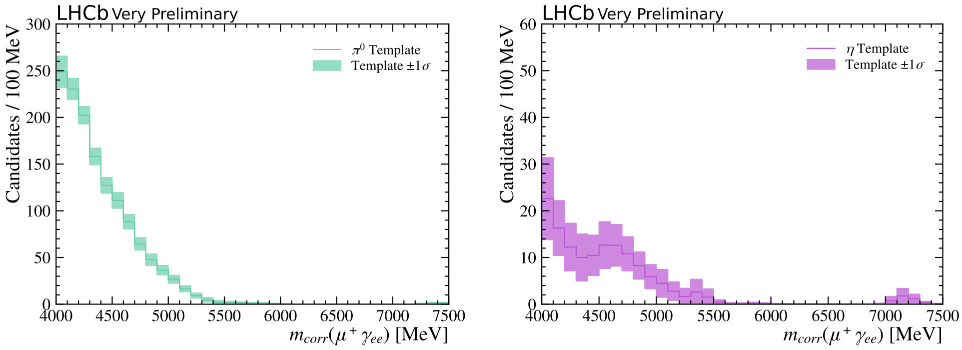


Figure 6.19: Corrected mass templates for background from $\pi^0 \to \gamma_{ee}\gamma$ (left) and $\eta \to \gamma_{ee}\gamma$ (right) scaled to the expected yield with a Gaussian filter applied to remove unphysical fluctuations.

6.3.6 Validation for background from $\pi^0 \to \gamma_{ee}\gamma$ decays

Given that the background from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ decays is the dominating source of background in the search for $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays, it is crucial to validate the method of modelling the corrected mass distribution of this type of background. For this, $D^{*+} \to D^0\pi^+$ candidates are selected in data, reconstructed as $D^0 \to K^-\pi^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ with and without the additional calorimeter photon. Applying the method developed in Sections 6.3.3 and 6.3.4, it should be possible to model the mass distribution, $m(K^-\pi^+\gamma_{ee})$, of all partially reconstructed $D^0 \to K^-\pi^+\pi^0(\to \gamma_{ee}[\gamma])$ candidates without additional γ_{calo} from the data sample reconstructed with an additional γ_{calo}. Due to the superior mass resolution of $D^0 \to K^-\pi^+\pi^0$ candidates in $m(K^-\pi^+\gamma_{ee})$ compared to corrected mass resolution in $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ candidates, this exercise demands a much higher precision and constitutes a rigorous validation of the method.

Template for $D^0 \to K^-\pi^+\pi^0$ in $m(K^-\pi^+\gamma_{ee})$

First, the decay $\pi^0 \to \gamma_{ee}\gamma_{calo}$ is resolved in data with a calorimeter photon, which is selected with the same requirements as defined in Section 6.3.1. Resolved $\pi^0 \to \gamma_{ee}\gamma_{calo}$ candidates are identified by a fit to $m(\gamma_{ee}\gamma_{calo})$, as shown in Figure 6.20.
Secondly, the photon efficiency correction is applied using the same weights as derived on simulation of $B^+ \to \mu^+\nu_\mu\pi^0(\to \gamma_{ee}\gamma)$ decays in Section 6.3.3. The resulting template for resolved and efficiency-corrected $\pi^0 \to \gamma_{ee}\gamma_{calo}$ candidates is shown in Figure 6.21a.

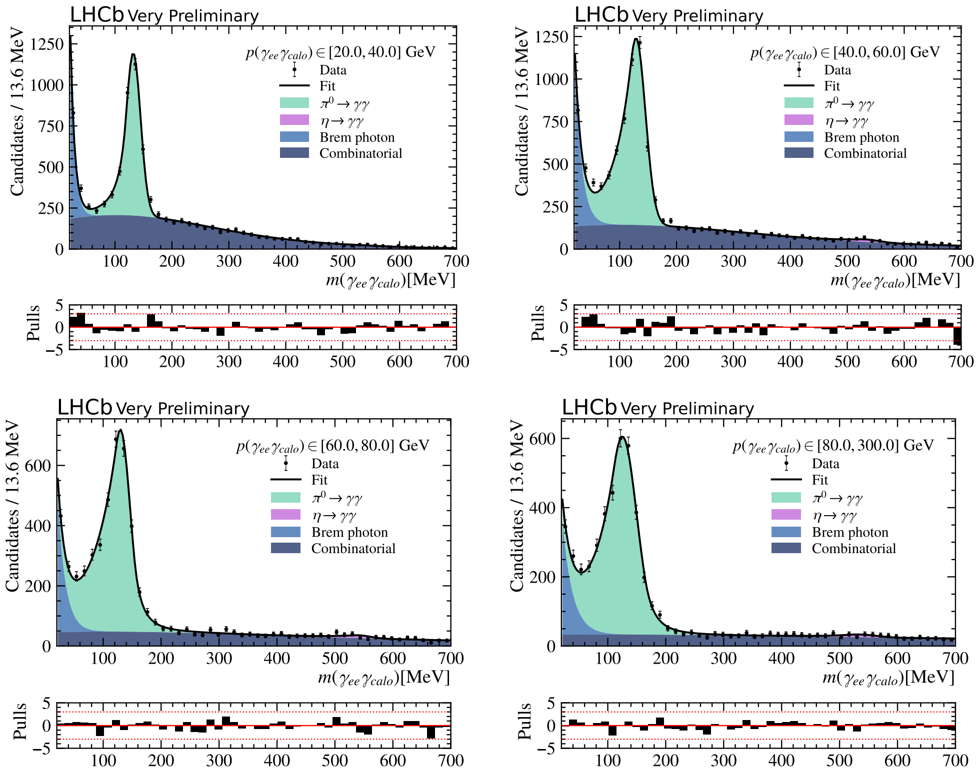
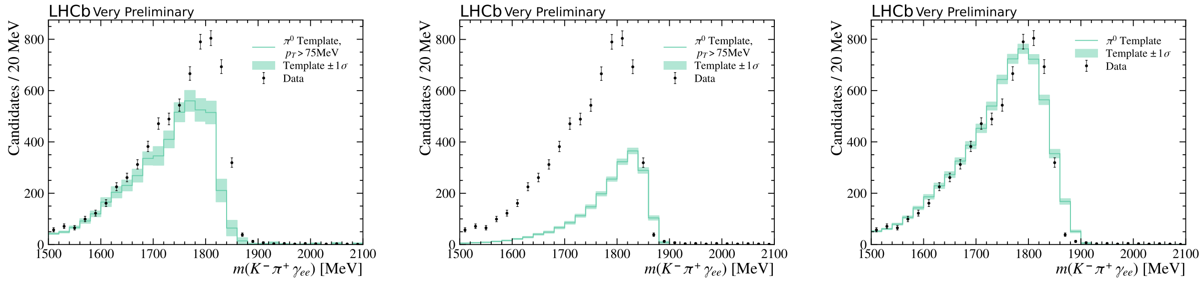


Figure 6.20: Fit to $D^0 \to K^-\pi^+\pi^0(\to \gamma_{ee}\gamma_{calo})$ candidates reconstructed in data in four bins of $p(\gamma_{ee}\gamma_{calo})$.

Next, the method to extrapolate the decay kinematics to the region $p_T(\gamma_{calo}) < 75\,\mathrm{MeV}$, developed in Section 6.3.4, is applied. The resulting template in $m(K^-\pi^+\gamma_{ee})$ is shown in Figure 6.21b.
Comparing the derived templates for the two kinematic regions of the additional photon in Figure 6.21a and 6.21b, the correlation between the kinematics of the γ_{calo} and the reconstructed mass $m(K^-\pi^+\gamma_{ee})$ becomes apparent. This underlines the necessity of accounting for events with soft γ_{calo} with the dedicated method developed in Section 6.3.4. The two templates are combined in Figure 6.21c. Good agreement is observed with the distribution of partially reconstructed $D^0 \to K^-\pi^+\pi^0(\to \gamma_{ee}[\gamma])$ candidates in $m(K^-\pi^+\gamma_{ee})$ both in terms of the mass distribution as well as in the number of expected candidates. This demonstrates that the developed method allows for the accurate modelling of the mass distribution of all $\pi^0 \to \gamma_{ee}\gamma$ decays from a sample of resolved $\pi^0 \to \gamma_{ee}\gamma_{calo}$ candidates. The residual differences between the data and the template in Figure 6.21c at the right-hand edge of the mass peak are expected to be absorbed by the corrected mass resolution when applied to the search for $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$ decays. This is confirmed in a second validation step, which is explained in the following.

Template for $D^0 \to K^-\pi^+\pi^0$ in $m_{corr}(\pi^+\gamma_{ee})$

In a similar validation, the templates are derived in a corrected D^0 mass, which is defined with a comparable relative mass resolution, $\Delta m/m$, as for $B^+ \to \mu^+\nu_\mu\gamma$ decays. The



(a) Template for $p_T(\gamma_{calo}) > 75\,\text{MeV}$ (b) Template for $p_T(\gamma_{calo}) < 75\,\text{MeV}$ (c) Combined template

Figure 6.21: Template (green) for $D^0 \to K^-\pi^+\pi^0$ candidates in $m(K^-\pi^+\gamma_{ee})$. The respective templates in (a) and (b) are scaled according to the expected yield and added up to obtain the template in (c). The distribution of partially reconstructed $D^0 \to K^-\pi^+\pi^0(\to \gamma_{ee}[\gamma])$ candidates (black) is overlaid in all plots.

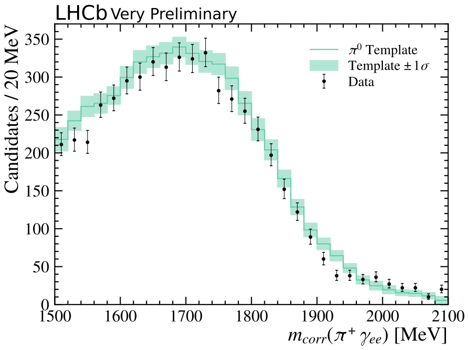


Figure 6.22: Template (green) and observed data (black) for partially reconstructed $D^0 \to K^-\pi^+\pi^0(\to \gamma_{ee}[\gamma])$ candidates in $m_{corr}(\pi^+\gamma_{ee})$.

corrected D^0 mass is defined without momentum information on the kaon[3], treating it as a (massive) neutrino, correcting for the momentum imbalance in the decay perpendicular to the D^0 flight direction. To account for the mass of the kaon, m_K, in the definition of the corrected mass, Eq. 4.1 is modified for

$$m_{corr}(\pi^+\gamma_{ee}) = \sqrt{m^2_{\pi\gamma_{ee}} + |p_\perp|^2} + \sqrt{m^2_K + |p_\perp|^2}. \tag{6.17}$$

Figure 6.22 shows the comparison of the derived template and $D^0 \to K^-\pi^+\pi^0$ candidates in data using the corrected D^0 mass. Excellent agreement is found, confirming the residual discrepancies observed for the templates in $m(K^-\pi^+\gamma_{ee})$ shown in Figure 6.21c are negligible.

---

[3] The kaon is still used to reconstruct the D^0 decay vertex. This counterbalances the effect of the significantly shorter flight distance of D^0 mesons with respect to B mesons on the corrected mass resolution.

6.4 Additional sources of background

6.4.1 Background from muon misidentification

Background candidates in which a final-state particle is misidentified are referred to as *misID background*. To derive the amount and model the mass shape of misID background after the full selection is applied, a data-driven approach is taken, utilising dedicated data control samples.

Data control regions

A dedicated data control sample is selected from a stripping selection with all PID requirements on the muon candidate removed. To reduce the retention rate of this selection, it is pre-scaled by a factor of 0.05.
The expected misID background is expected to be composed of π^+ and K^+ misidentified as muons. Misidentification of electrons can be neglected in this case, as they do not reach the muon stations and therefore do not pass the `isMuon` requirement. The data control sample is further divided into two PID regions, enriched in pions and kaons, respectively. The set of PID selections for the two regions, along with the nominal PID cut for the muon candidate in the signal selection, are given in Table 6.1.

Particle	PID requirement
μ^+	`isMuon`==1, $\text{DLL}_{\mu\pi} > 0$, $\text{DLL}_{\mu K} > 0$, $\texttt{ProbNN}_{\mu} > 0.2$
π^+	`isMuon`==0, $\text{DLL}_{\mu\pi} < 0$, $\text{DLL}_{K\pi} < 0$
K^+	`isMuon`==0, $\text{DLL}_{\mu\pi} < 0$, $\text{DLL}_{K\pi} > 0$

Table 6.1: PID selection for hadron-enriched regions of the misID data sample as well as the nominal PID selection for muons.

PID efficiencies using PIDCalib2

The efficiencies of π^+ and K^+ passing the PID cut of the respective control region, as well as the efficiency for passing the nominal muon PID requirement, are derived using the PIDCalib2 package [100]. The efficiencies are obtained in bins of the track momentum, the track pseudorapidity, and the event multiplicity, `nTracks`. The ratio of the two efficiencies is assigned as a per-event weight to candidates in the data-control samples to determine the contamination of misID background that passes the nominal PID requirement. The prescale of the stripping line is accounted for by scaling each derived weight by a factor of 20.

Corrected mass template

The corrected mass templates for the $h^+ \to \mu^+$ misID backgrounds are derived separately by applying the weights mentioned above to the respective misID data control regions. Figure 6.23 shows the corrected mass templates scaled to the expected level for the different species of $h^+ \to \mu^+$ misID background in the signal region.

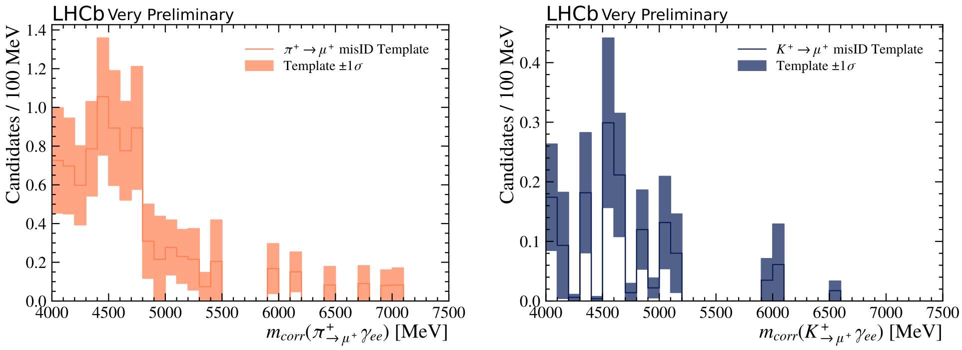


Figure 6.23: Corrected mass templates for background from $\pi^+ \to \mu^+$ (left) and $K^+ \to \mu^+$ (right) misID.

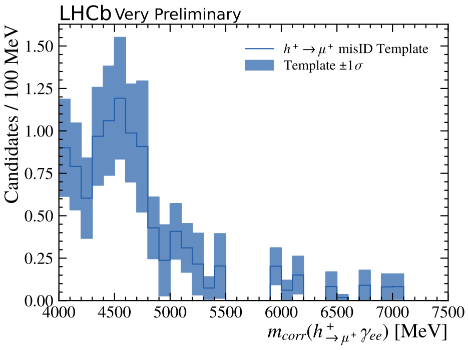


Figure 6.24: Corrected mass template for the combined $h^+ \to \mu^+$ misID background.

For $\pi^+ \to \mu^+$, a total of 8.6 ± 1.0 background candidates are expected, for $K^+ \to \mu^+$, the expectation is 1.5 ± 0.3 background candidates. The corrected mass template for the combined misID backgrounds is constructed by adding the contributions from kaon and pion misID. A total number of 10.1 ± 1.0 background candidates is expected and the corrected mass template is shown in Figure 6.24.

6.4.2 Background from electron misidentification

In principle, hadron to electron misidentification can be treated the same way as for muons in the previous section, using a data control sample without a PID requirement on one of the electron tracks selected from a dedicated stripping line. However, the electrons from conversions are selected using tight requirements on the di-electron invariant mass and on the VELO tracks of the two electrons overlapping. This drastically reduces the available phase space for background from electron misidentification, resulting in a negligible contribution.

6.4.3 Background from $B^+_{(c)} \to \tau^+\nu_\tau\gamma_{ee}$

Radiative leptonic decays $B^+_{(c)} \to \ell^+\nu_\ell\gamma$ can also proceed to a final state with tau leptons as $B^+_{(c)} \to \tau^+\nu_\tau\gamma$. With subsequent $\tau^+ \to \mu^+\nu_\mu\bar{\nu}_\tau$ decays, this yields an irreducible source of partially reconstructed background that differs only by the two additional neutrinos. The contamination from such tauonic decays that pass the full event selection is studied using dedicated simulation samples of $B^+_{(c)} \to \tau^+\nu_\tau\gamma_{ee}$ decays. The relative contamination, f_τ, is derived separately for decays of B^+ and B^+_c mesons as

$$f_\tau^{B^+_{(c)}} = \frac{N^{\text{expected}}_{B^+_{(c)}\to\tau^+\nu_\tau\gamma}}{N^{\text{expected}}_{B^+_{(c)}\to\mu^+\nu_\mu\gamma}} = \frac{\varepsilon^\tau_{tot}}{\varepsilon^\mu_{tot}} \frac{\mathcal{B}(B^+_{(c)} \to \tau^+\nu_\tau\gamma) \times \mathcal{B}(\tau^+ \to \mu^+\nu_\mu\bar{\nu}_\tau)}{\mathcal{B}(B^+_{(c)} \to \mu^+\nu_\mu\gamma)}, \tag{6.18}$$

where N^{expected} is the number of expected signal candidates and $\varepsilon^{\mu/\tau}_{tot}$ is the total selection efficiency of muonic/tauonic decays. Using the predicted branching fractions of B^+ decays to the muonic and tauonic final states for $E^*_\gamma > 1\,\text{GeV}$ from Ref. [110] and the measured branching fraction of $\tau^+ \to \mu^+\nu_\mu\bar{\nu}_\tau$ from the PDG [5], a ratio of $f_\tau^{B^+} \approx 0.01$ is computed. For the respective decay of B^+_c mesons, the authors of [42] predict the branching fractions of the muonic and tauonic decays for $E^*_\gamma > 1\,\text{GeV}$ to be of similar size. With the slightly lower selection efficiency, this also leads to $f_\tau^{B^+_c} \approx 0.01$. As the expected contamination is approximately two orders of magnitude smaller than the respective signal, background from $B^+_{(c)} \to \tau^+\nu_\tau\gamma$ decays is neglected.

6.5 Fitting the corrected mass

6.5.1 Nominal fit strategy

The number of $B^+ \to \mu^+\nu_\mu\gamma$ and $B^+_c \to \mu^+\nu_\mu\gamma$ candidates in data is obtained simultaneously from a single fit to the corrected mass distribution in data. The fit is using the templates derived for the $B^+ \to \mu^+\nu_\mu\gamma$ and $B^+_c \to \mu^+\nu_\mu\gamma$ signal components as well as the π^0 and η backgrounds and the background from muon misID. Each of the fit components is extended by a separate yield parameter, N_j. By default, the signal yields are allowed to be negative to improve the statistical behaviour of the likelihood function. The statistical methods to obtain the branching fractions in the physically allowed region are discussed in Section 8.7. On the number of background candidates from muon misID, a Gaussian constraint is imposed with the central value and width of the Gaussian set to the predicted yield and its uncertainty, respectively. For background from $\pi^0 \to \gamma_{ee}\gamma$, the validation in Section 6.3.6 shows a slight discrepancy in the predicted and observed number of background candidates. To allow for a variation of the number of background candidates, the yields N_π and N_η are allowed to float in the fit, but a constraint on the relative fit fraction N_π/N_η is imposed. The central value and width of the Gaussian constraint are set according to the predicted number of background candidates from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ decays, accounting for the respective uncertainties.

6.5.2 Blinded mass fit

As the analysis is blinded at this stage, the signal region $4500\,\text{MeV} < m_{corr} < 6500\,\text{MeV}$ is masked and excluded from the fit and the signal yields N_{B^+} and $N_{B^+_c}$ are forced to zero. Figure 6.25 shows the blinded fit to data with the fit parameters listed in Table 6.2. For reference, approximately 56 $B^+ \to \mu^+\nu_\mu\gamma$ signal candidates are expected based on the upper limit on the branching fraction from Belle [1] and about one $B^+_c \to \mu^+\nu_\mu\gamma$ candidate based on the Standard Model prediction from Ref. [2]. The corresponding leakage

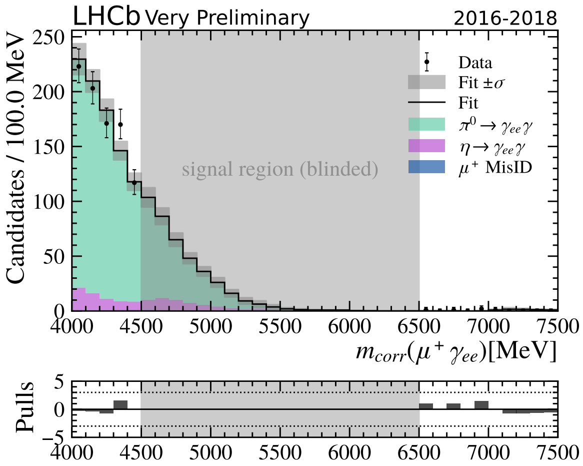


Figure 6.25: Blinded fit to the corrected mass distribution combining the data recorded in the years 2016 to 2018.

Parameter	**Fitted value**	**Predicted value**
N_π	1170 ± 60	1390 ± 60
N_η	135 ± 29	150 ± 40
N_{misID}	10.1 ± 1.0	10.1 ± 1.0 (constraint)
N_η / N_π	0.116 ± 0.026	0.105 ± 0.028 (constraint)

Table 6.2: Fit result for the blinded mass fit shown in Figure 6.25.

of signal outside of the blinded region is therefore completely negligible.
The observed number of π^0 background candidates is 2.5σ lower than the prediction, which could be related to the different kinematics of π^0 in $B^+ \to \mu^+\nu_\mu\pi^0$ simulation and the actual π^0 decays contributing in data. It is worth noting that the uncertainty on the predicted yield of π^0 background candidates quoted in Table 6.2 does not take into account systematic uncertainties linked to the method with which the background yield is predicted. However, the fact that the observed level of background is found to be slightly smaller than the expectation suggests that no significant background component is missing in the fit. The observed level of background from misidentification is consistent with the prediction by construction since a constraint is applied on N_{misID}.

6.6 Fit validation

The performance of the fit to the corrected mass is validated in a series of *pseudo-experiments*. From the derived signal and background templates, samples of *pseudo-data* are generated with known compositions that are representative of the expected corrected mass distribution in data. Studies of pseudo-experiments typically involve fitting the corrected mass distribution with different signal strengths injected to evaluate potential biases and correlations among the fit results.

Generating pseudo-data

The pseudo-data are generated from the nominal templates for the signal and background components derived in the previous sections. To account for the template uncertainties, bin-wise Gaussian modifiers are applied to vary the expected mass distribution within the respective uncertainties for each fit component. Bins in which the resulting PDF is negative are forced to zero. For each component, events are sampled randomly from the resulting corrected mass distribution, which generates the expected statistical fluctuations in the pseudo-data. The number of events sampled from the background templates is taken from the result of the blinded fit to the data and varied within their respective uncertainties. The number of injected $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ signal is specified separately in each study.

Pull studies

The performance of the fit is validated in a series of pseudo-experiments, each of which is analysed in terms of the *pull*. It is defined as the difference between the fitted number of signal candidates, μ, and the number of signal candidates injected in the pseudo-data, μ_0, divided by the respective uncertainty on the fit result, σ, as

$$pull = \frac{\mu - \mu_0}{\sigma}. \tag{6.19}$$

According to the central limit theorem, the pull distributions for a series of pseudo-experiments should be Gaussian distributed and centred at zero with a standard deviation of one.

To validate the performance of the fit to the corrected mass, three pull studies are performed with different levels of injected signal. The first study is conducted for the background-only hypothesis in the absence of $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ signal. A second study is conducted for a realistic signal hypothesis in which the injected signal for $B^+ \to \mu^+\nu_\mu\gamma$ decays corresponds to Belle's upper limit on the branching fraction of 3.0×10^{-6}, which is equivalent to 56 signal candidates. For $B^+_c \to \mu^+\nu_\mu\gamma$, the signal is injected according to the Standard Model branching fraction of 8×10^{-5} [2], corresponding to one signal candidate. The final validation study is performed on pseudo-data, which is generated with a signal strength that would lead to an observation with a significance of 5σ. From the results on the statistical sensitivity in Section 8.7.1 and extrapolation to the 5σ level, this is expected for approximately 130 $B^+ \to \mu^+\nu_\mu\gamma$ signal candidates and 25 $B^+_c \to \mu^+\nu_\mu\gamma$ signal candidates. The fit for one such a pseudo-experiment is shown in Figure 6.26.

For all three studies, a total of one thousand pseudo-experiments are conducted. The resulting pull distributions on the extracted $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ signal yields are shown in Figure 6.27. The pull distributions of the extracted $B^+ \to \mu^+\nu_\mu\gamma$ signal show a small shift of the central position towards lower values. This fit bias is found to be consistent with the number of generated events and is corrected in the following section. The width of the respective distributions is consistent with unity for all three studies. The pull distributions of the extracted $B^+_c \to \mu^+\nu_\mu\gamma$ signal are slightly skewed due to the proximity to the physical boundary of the fit being negative in some bins. In Section 8.7, the correct confidence intervals of the extracted signals are recovered based on pseudo-experiments, ensuring correct coverage even in the case that the pull distributions are not following a Gaussian profile. For the extraction of the $B^+_c \to \mu^+\nu_\mu\gamma$ signal, the study with a realistic

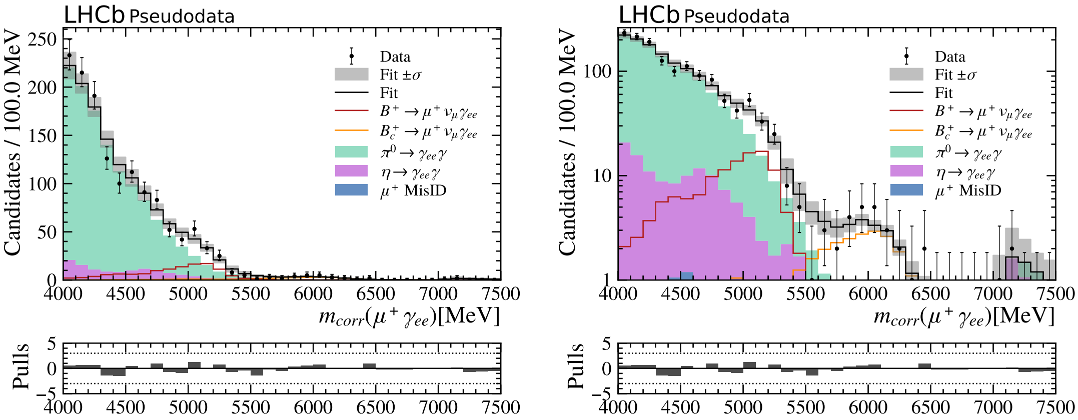


Figure 6.26: Fit to pseudo-data generated with $N_{B+} = 130$ and $N_{B_c^+} = 25$ in linear (left) and logarithmic (right) scale.

signal injected shows a small negative bias, which is also corrected to yield an unbiased central value for the final result.

Correcting the fit bias for $B^+ \to \mu^+\nu_\mu\gamma$

In the extraction of the $B^+ \to \mu^+\nu_\mu\gamma$ signal, a slight negative bias is observed, which is consistent as a function of the number of injected signal candidates. This suggests correcting for the observed bias and recovering the correct central value. The average fit bias is quantified by conducting a series of pseudo-experiments with different number of injected $B^+ \to \mu^+\nu_\mu\gamma$ signal. The observed bias on the extracted number of signal candidates as a function of the number of injected signal candidates is shown on the left-hand side in Figure 6.28. A mean bias on the number of fitted $B^+ \to \mu^+\nu_\mu\gamma$ signal candidates of -5.8 is found, which is used to correct the number of extracted signal candidates for the final result.

The same strategy is applied for the number of extracted $B_c^+ \to \mu^+\nu_\mu\gamma$ decays, which also shows a slight negative bias. The bias is evaluated as a function of the induced number of $B_c^+ \to \mu^+\nu_\mu\gamma$ signal events up to ten times the Standard Model expectation as shown on the right-hand side of Figure 6.28. A mean bias of -0.8 is found in the extracted number of $B_c^+ \to \mu^+\nu_\mu\gamma$ signal candidates, which is used to correct the final result. The systematic uncertainty induced by this correction is studied in Chapter 9.

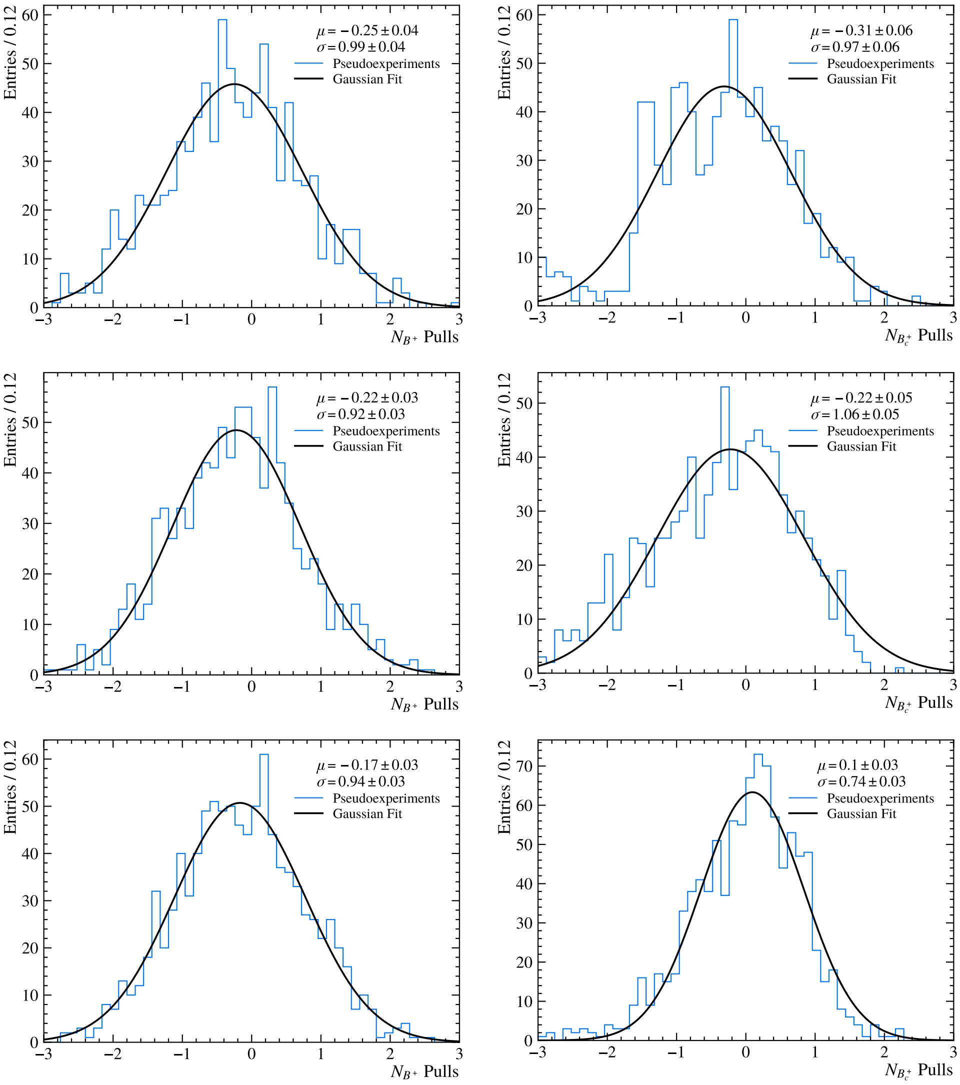


Figure 6.27: Pull distribution for the extracted $B^+ \to \mu^+\nu_\mu\gamma$ (left) and $B_c^+ \to \mu^+\nu_\mu\gamma$ (right) signal from one thousand pseudo-experiments. The pseudo-data is generated in the absence of signal (top), for a realistic signal hypothesis (centre) and for an expected observation with a significance of 5σ (bottom).

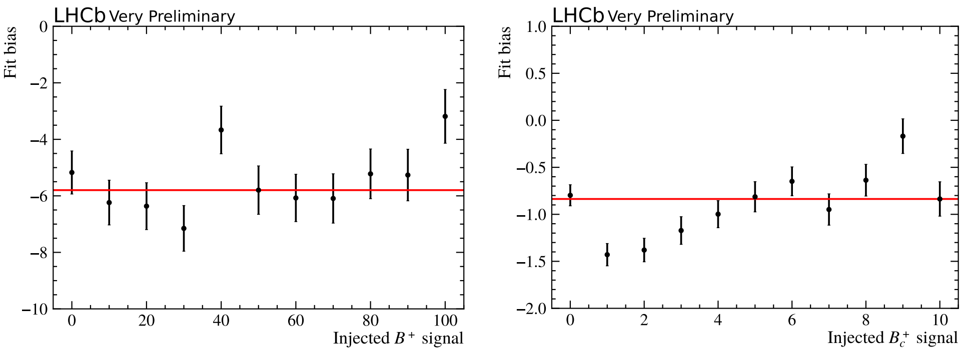


Figure 6.28: Observed fit bias on the number of extracted $B^+ \to \mu^+\nu_\mu\gamma$ signal candidates as a function of the number of injected signal candidates. The red line marks the mean observed mean bias.

7

Normalisation channel $B^0 \to K^{*0}\gamma$

Since there are large uncertainties on the B meson production rate at LHCb, an absolute measurement of the branching fraction of $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays is prone to large uncertainties. Instead, its branching fraction is measured relative to a *normalisation channel* with well-measured branching fraction. Many uncertainties related to the B meson production rate, including luminosity, $b\bar{b}$ cross section, fragmentation fractions (only for $B^+ \to \mu^+\nu_\mu\gamma$) cancel in the ratio of branching fractions, along with other systematic uncertainties linked to the reconstruction and selection of candidates.
In this analysis, the normalisation channel $B^0 \to K^{*0}\gamma$ is used with $\gamma \to e^+e^-$ conversions and electrons reconstructed from long tracks. It allows a cancellation of the B meson production rate due to identical fragmentation fractions $f_u = f_d$, offers a clean signal selection and has been studied extensively in previous LHCb analyses [94, 95]. Its branching ratio is measured as [5]

$$\mathcal{B}(B^0 \to K^{*0}\gamma) \times \mathcal{B}(K^{*0} \to K^+\pi^-) = (2.79 \pm 0.17) \times 10^{-5} \tag{7.1}$$

with a relative uncertainty of 6% and about an order of magnitude larger than the expected branching ratio for $B^+ \to \mu^+\nu_\mu\gamma$ decays. Crucially, the efficiency of the photon conversion and its reconstruction from a pair of e^+e^- long tracks cancels in the ratio. Due to the presence of two hadrons instead of a muon in the final state, systematic uncertainties related to hadron-muon differences do not cancel in the ratio. In Chapter 9 this is studied and a systematic uncertainty is assigned to the final result.
For the decay $B^+_c \to \mu^+\nu_\mu\gamma$ the same normalisation channel is used. In contrast to the $B^+ \to \mu^+\nu_\mu\gamma$ decay, the B meson production rate does not cancel in the ratio and a dependence on f_d/f_c remains for the final result on the branching fraction.

Selection alignment

The selection of the candidates for the normalisation channel is described in detail in Chapter 5. The selection of the hadronic part is aligned with the previous LHCb analyses of $B^0 \to K^{*0}(\to K^+\pi^-)e^+e^-$ decays at very low q^2 [95] in which the decay $B^0 \to K^{*0}\gamma$ is studied as a control channel. The selection of the photon conversion is fully aligned between the $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ signal and $B^0 \to K^{*0}\gamma$ normalisation channel. Details on the selection requirements can be found in Table 5.3. Note that only the charged and neutral isolation BDTs are applied, while the combinatorial BDT cannot be applied in $B^0 \to K^{*0}\gamma$ due to the absence of a final state muon.

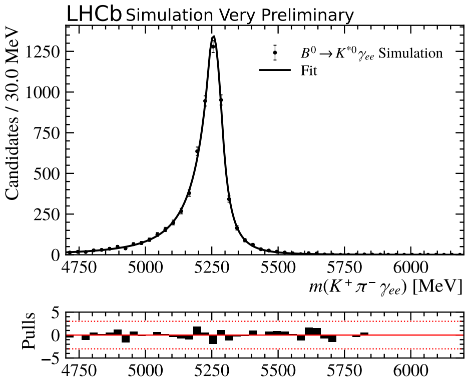


Figure 7.1: Fit to the $m(K^{+}\pi^{-}\gamma_{ee})$ invariant mass of $B^0 \to K^{*0}\gamma$ candidates in simulation.

7.1 Fit setup

The number of $B^0 \to K^{*0}\gamma$ candidates is extracted from an extended unbinned maximum likelihood fit to the $m(K^{+}\pi^{-}\gamma_{ee})$ invariant mass spectrum using the python library ZFIT [108]. The model for the $B^0 \to K^{*0}\gamma$ signal component is derived from simulation. Prominent partially reconstructed backgrounds are modelled from dedicated simulation samples, while the combinatorial background is modelled using an exponential shape.

Signal shape

The $B^0 \to K^{*0}\gamma$ signal shape is modelled using a double-sided Crystal-Ball PDF. The model parameters are determined from a fit to the reconstructed invariant mass in simulated $B^0 \to K^{*0}(\to K^{+}\pi^{-})\gamma_{ee}$ candidates. The fit projection is shown in Figure 7.1.

Partially reconstructed background with missing hadron

Partially reconstructed background candidates originate predominantly from the decay $B^{+} \to K_1(1270)^{+}\gamma$ with $K_1(1270)^{+} \to K^{+}\pi^{-}X$ where X is not reconstructed. The invariant mass shape of this background is modelled from $B^{+} \to K_1(1270)^{+}\gamma(\to e^{+}e^{-})$ simulation, where the $K_1(1270)^{+}$ may decay to $K^{*0}(\to K^{+}\pi^{-})\pi^{+}$, $\rho^0(\to \pi^{+}\pi^{-}X)K^{+}$, $\omega(782)(\to \pi^{+}\pi^{-})K^{+}$ or non-resonantly to $K^{+}\pi^{+}\pi^{-}$. The relative branching fractions of the $K_1(1270)^{+}$ decays are fixed according to their measured branching fractions [5]. The invariant mass shape of simulated candidates passing the full selection is described using a double-sided Crystal-Ball PDF with the corresponding fit shown in Figure 7.2.

Background from $\mathbf{B^0 \to K^{*0}\pi^0}$ and $\mathbf{B^0 \to K^{*0}\eta}$

Background candidates in which the converted photon does not originate from the B meson decay but rather from the decay of a light $\pi^0/\eta \to \gamma_{ee}\gamma$ decay are modelled from respective simulation samples. The invariant mass shapes of $B^0 \to K^{*0}\pi^0$ and $B^0 \to K^{*0}\eta$ candidates are modelled separately using a non-parametric Kernel Density Estimation (KDE) from the ZFIT python package. The respective simulation data along with the derived KDE shape are shown in Figure 7.3.

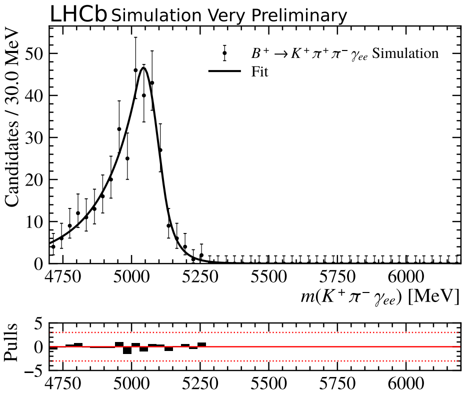


Figure 7.2: Fit to the $m(K^+\pi^-\gamma_{ee})$ invariant mass spectrum of $B^+ \to K_1(1270)^+(\to K^+\pi^-[\pi^+])\gamma_{ee}$ simulation.

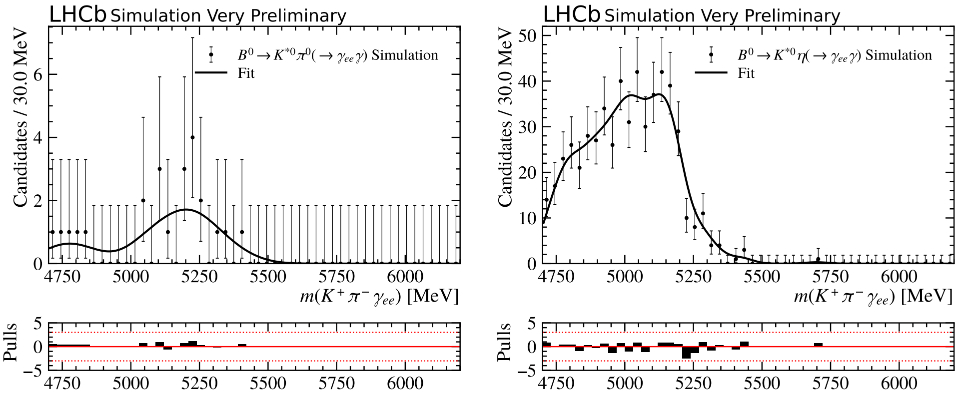


Figure 7.3: Distribution to the $m(K^+\pi^-\gamma_{ee})$ invariant mass of $B^+ \to K^{*0}\pi^0(\to \gamma_{ee}\gamma)$ (left) and $B^0 \to K^{*0}\eta(\to \gamma_{ee}\gamma)$ simulation overlaid with the respective KDE.

Combinatorial background

The combinatorial background is composed of candidates in which at least one reconstructed final state particle originates from a process other than the signal B meson decay. The analysis of $B^0 \to K^{*0}\gamma$ decays in [95] shows it can be described by a falling exponential shape. The same method is used in this analysis and an exponential shape with a floating slope is used to model the combinatorial background in the fit to data.

7.2 Invariant mass fit to $B^0 \to K^{*0}\gamma$ data

The invariant mass fit to data is performed in a single extended unbinned maximum likelihood fit combining all three years of data-taking. The shapes for the signal and partially reconstructed background components are fixed from the invariant mass fits to the simulation. To account for differences between data and simulation in the signal mass shape, a shift of the central position, $\Delta\mu$, is applied while the width of the Gaussian core is scaled by $(1+\Delta\sigma)$. The slope of the combinatorial background shape is free to vary in the fit to data. The combined PDF is given by the sum of the signal and background components

$$\begin{aligned}\mathcal{P}_{tot}(x;\Delta\mu,\Delta\sigma,\lambda) = {} & N_{sig}\mathcal{P}_{sig}(\mu_{MC}+\Delta\mu,\sigma_{MC}(1+\Delta\sigma)) \\ & + N_{PR}\mathcal{P}_{PR} + N_\pi\mathcal{P}_\pi + N_\eta\mathcal{P}_\eta \\ & + N_{comb}e^{-\lambda x},\end{aligned} \tag{7.2}$$

where N_i and $\mathcal{P}_i$ are the respective yields and PDFs of the individual signal and background components.

To improve the fit stability, the yields of the $B^0 \to K^{*0}\pi^0$ and $B^0 \to K^{*0}\eta$ background components are constrained relative to the signal yield. Using the measured branching fractions of these decays from [5] along with the signal branching fraction, the yields N_π and N_η are substituted for

$$\begin{aligned} N_\pi &= f_\pi N_{sig} = \frac{\mathcal{B}(B^0 \to K^{*0}\pi^0(\to\gamma\gamma))}{\mathcal{B}(B^0 \to K^{*0}\gamma)}\frac{\varepsilon^{MC}_{B^0\to K^{*0}\pi^0}}{\varepsilon^{MC}_{B^0\to K^{*0}\gamma}}N_{sig} \\ N_\eta &= f_\eta N_{sig} = \frac{\mathcal{B}(B^0 \to K^{*0}\eta(\to\gamma\gamma))}{\mathcal{B}(B^0 \to K^{*0}\gamma)}\frac{\varepsilon^{MC}_{B^0\to K^{*0}\eta}}{\varepsilon^{MC}_{B^0\to K^{*0}\gamma}}N_{sig}\end{aligned} \tag{7.3}$$

where $\varepsilon^{MC}_{B\to f}$ is the total selection efficiency of the respective B meson decay to a final state f evaluated on simulation. A Gaussian constraint is imposed upon the relative fit fractions $f_\pi = 0.008 \pm 0.016$ and $f_\eta = 0.022 \pm 0.003$, where uncertainties are derived from the external uncertainties on the branching fractions and the uncertainties on the selection efficiencies due to limited statistics in the simulation samples. The fit to data is shown in Figure 7.4 and yields 692 ± 31 reconstructed $B^0 \to K^{*0}\gamma$ candidates in data.

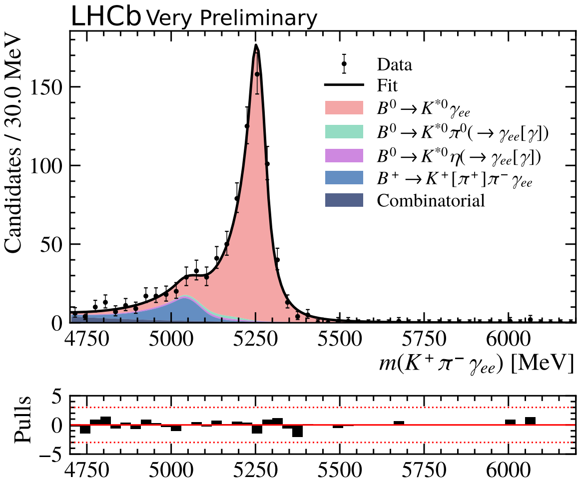


Figure 7.4: Invariant mass fit of $B^0 \to K^{*0}\gamma$ candidates reconstructed in data. The fit components are stacked on top of each other.

8 Statistical sensitivity

In this chapter, the expected statistical sensitivity on the branching ratios for $B^+ \to \mu^+\nu_\mu\gamma$ and $B_c^+ \to \mu^+\nu_\mu\gamma$ decays is obtained. First, the efficiencies of reconstructing and selecting $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$ and $B^0 \to K^{*0}\gamma$ decays are derived. They are calculated on the respective event simulation and factorised into

$$\varepsilon_{tot} = \varepsilon_{geo} \times \varepsilon_{filter} \times \varepsilon_{reco\&sel}, \tag{8.1}$$

where ε_{geo} is the geometrical detector acceptance efficiency (see Section 8.1), ε_{filter} is the efficiency of the conversion filter in the event generation (see Section 8.2) and $\varepsilon_{reco\&sel}$ is the remaining reconstruction and selection efficiency (see Section 8.3). All efficiencies are calculated separately per year of data taking.
In Section 8.5, the number of observed $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$ candidates is related to the respective branching fraction, defining the single event sensitivity. The expected statistical sensitivity on the branching ratio of $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$ decays is studied in Section 8.7.

8.1 Geometric efficiency ε_{geo}

The detector acceptance is defined as the fraction of events for which all decay products are produced in the instrumented forward region of the LHCb experiment. Charged particles are required to be produced in the forward detector acceptance defined by the polar angle $\theta \in [10, 400]$ mrad. Neutral particles are required to be in the region [0,400] mrad. Both these regions are chosen larger than the actual LHCb acceptance to allow for particles re-entering the acceptance due to the magnetic field or scattering. For neutrinos, no requirement on the polar angle is imposed. The geometrical acceptance efficiencies for centrally produced simulations are provided by simulation experts and are listed in Table 8.1.

8.2 Conversion filter efficiency ε_{filter}

In the simulation of $B_c^+ \to \mu^+\nu_\mu\gamma$ and $B^0 \to K^{*0}\gamma$ decays, a conversion filter is applied in order to reduce the computational expenses in the simulation process (see Section 5.1.2). It requires the photon in the decay of interest to convert within a given volume that

Decay	ε_{geo} $[10^{-2}]$		
	2016	2017	2018
$B^+ \to \mu^+\nu_\mu\gamma$	26.88 ± 0.07	26.83 ± 0.06	26.91 ± 0.07
$B_c^+ \to \mu^+\nu_\mu\gamma$	25.48 ± 0.03	25.50 ± 0.03	25.46 ± 0.03
$B^0 \to K^{*0}\gamma$	24.98 ± 0.04	24.94 ± 0.03	25.02 ± 0.03

Table 8.1: Geometrical detector acceptance efficiencies per year averaged across the two magnet polarities.

Decay	Filter z_{max}	ε_{filter} $[10^{-2}]$		
		2016	2017	2018
$B^+ \to \mu^+\nu_\mu\gamma$	$z < 500\,\text{mm}$	4.76 ± 0.04	4.81 ± 0.04	4.68 ± 0.04
$B_c^+ \to \mu^+\nu_\mu\gamma$	$z < 500\,\text{mm}$	4.74 ± 0.03	4.73 ± 0.03	4.65 ± 0.03
$B^0 \to K^{*0}\gamma$	$z < 2500\,\text{mm}$	14.97 ± 0.04	14.97 ± 0.04	14.84 ± 0.04

Table 8.2: Filtering efficiencies of the conversion filter used in the production of the simulation samples.

is defined by the z-coordinate of the true $\gamma \to e^+e^-$ vertex. The efficiencies for the conversion filter are derived from the central event generation and are listed in Table 8.2.

8.3 Reconstruction and selection efficiency $\varepsilon_{reco\&sel}$

The combined reconstruction and selection efficiency, $\varepsilon_{reco\&sel}$, is determined using simulation samples with all corrections to simulation applied. Because the weights generated by the GBReweighter to correct the initial B kinematics and multiplicity, and reconstruction effects, are not normalised, they need to be normalised before computing the selection efficiency.

Normalising $w_{\text{Mult\&Kin}}$

The weights $w_{\text{Mult\&Kin}}$ used to correct the initial B kinematics in simulation are normalised using the *generator level simulation*, which holds the true event information of all simulated decays prior to the interaction with the detector. The normalisation constant for the weights $w_{\text{Mult\&Kin}}$ is defined as the ratio between the number of generated events, N_{Gen}, and the sum of the correction weights $w_{\text{Mult\&Kin}}$ on the generator level simulation. The correction weight $w_{\text{Mult\&Kin}}$ is re-defined as:

$$w_{\text{Mult\&Kin}} \to \frac{N_{Gen}}{\sum_{i=1}^{N_{Gen}} w_{i,\text{Mult\&Kin}}} w_{\text{Mult\&Kin}}. \tag{8.2}$$

Normalising w_{Reco}

Also, the weights w_{Reco} used to correct for discrepancies in the reconstruction between data and simulation need to be normalised. The normalisation constant is derived from simulation samples with and without the correction weight w_{Reco} applied. For the number of all reconstructed simulated candidates, N_{sel}, that pass the selection, the remaining

Channel	$\varepsilon_{reco\&sel}$ $[10^{-4}]$		
	2016	2017	2018
$B^+ \to \mu^+\nu_\mu\gamma$	6.45 ± 0.27	5.86 ± 0.25	5.62 ± 0.23
$B_c^+ \to \mu^+\nu_\mu\gamma$	0.66 ± 0.08	0.67 ± 0.08	0.66 ± 0.08
$B^0 \to K^{*0}\gamma$	2.93 ± 0.11	2.77 ± 0.12	2.44 ± 0.11

Table 8.3: Reconstruction and selection efficiencies of the signal and normalisation channels for each year of data-taking.

simulation correction weights are applied. The correction weight w_{Reco} is then re-defined as:

$$w_{\mathrm{Reco}} \to \frac{\sum_{i=1}^{N_{Sel}} \left(w_{i,\mathrm{PID}} \times w_{i,\mathrm{Trk}} \times w_{i,\mathrm{L0}} \times w_{i,\mathrm{HLT}} \times w_{i,\mathrm{Mult\&Kin}}\right)}{\sum_{i=1}^{N_{Sel}} \left(w_{i,\mathrm{PID}} \times w_{\mathrm{i,Trk}} \times w_{i,\mathrm{L0}} \times w_{i,\mathrm{HLT}} \times w_{i,\mathrm{Mult\&Kin}} \times w_{i,\mathrm{Reco}}\right)} w_{\mathrm{Reco}}. \tag{8.3}$$

Deriving the reconstruction and selection efficiency

The efficiency of reconstructing and selecting a decay of interest is defined as the fraction of simulated candidates that pass the full selection and accounting for all simulation correction weights. For the normalisation channel $B^0 \to K^{*0}\gamma$, the selection listed in Table 5.3 is applied as well as the nominal cut on the charged (`IsoBDT` > -0.70) and neutral isolation (`PiBDT` > 0.50) BDTs. For the signal channels $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$, the same BDT selection is applied on top of the preselection listed in Table 5.3 as well as the cut on the combinatorial BDT with the nominal cut of `CombBDT` > 0.50.
For the normalisation channel $B^0 \to K^{*0}\gamma$, the efficiency of reconstruction and selection is derived simply using the number of generated decays in the simulation sample, N_{gen}, as the denominator for the efficiency. For the signal channels $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$, the search is performed for a given cut on the photon energy in the B meson rest frame, E_γ^*. Thus, the efficiency of selection and reconstruction is based on the number of generated events *given* a cut of $E_\gamma^* > 1\,\mathrm{GeV}$:

$$\begin{aligned}\varepsilon_{reco\&sel}^{sig} &= \frac{\sum_{i=1}^{N_{sel}} \left(w_{i,\mathrm{PID}} \times w_{i,\mathrm{TRK}} \times w_{i,\mathrm{L0}} \times w_{i,\mathrm{HLT}} \times w_{i,\mathrm{Mult\&Kin}} \times w_{i,\mathrm{Reco}}\right)}{N_{gen}(E_\gamma^* > 1\,\mathrm{GeV})}, \\ \varepsilon_{reco\&sel}^{norm} &= \frac{\sum_{i=1}^{N_{sel}} \left(w_{i,\mathrm{PID}} \times w_{i,\mathrm{TRK}} \times w_{i,\mathrm{L0}} \times w_{i,\mathrm{HLT}} \times w_{i,\mathrm{Mult\&Kin}} \times w_{i,\mathrm{Reco}}\right)}{N_{gen}}.\end{aligned} \tag{8.4}$$

The reconstruction and selection efficiency $\varepsilon_{reco\&sel}$ for the signal and normalisation channels are given in Table 8.3.

8.4 Total efficiency

The total efficiency is computed as the product of $\varepsilon_{tot} = \varepsilon_{geo} \times \varepsilon_{filter} \times \varepsilon_{reco\&sel}$. It is averaged across all three years of data-taking from a weighted average accounting for the integrated luminosity, $\mathcal{L}_y$, recorded with the LHCb experiment per year, y:

$$\varepsilon_{tot} = \frac{\sum_{y\in[16,17,18]} \mathcal{L}_y \times \varepsilon_{tot,y}}{\sum_{y\in[16,17,18]} \mathcal{L}_y}. \tag{8.5}$$

Channel	ε_{total} $[10^{-6}]$			
	2016 ($1.6\,\text{fb}^{-1}$)	2017 ($1.7\,\text{fb}^{-1}$)	2018 ($2.1\,\text{fb}^{-1}$)	Averaged
$B^+ \to \mu^+\nu_\mu\gamma$	8.3 ± 0.4	7.55 ± 0.33	7.07 ± 0.30	7.57 ± 0.19
$B_c^+ \to \mu^+\nu_\mu\gamma$	0.80 ± 0.09	0.81 ± 0.10	0.78 ± 0.09	0.79 ± 0.05
$B^0 \to K^{*0}\gamma$	11.0 ± 0.4	10.4 ± 0.4	9.0 ± 0.4	10.03 ± 0.24

Table 8.4: Total efficiencies of the signal and normalisation channels for each year of data-taking and the weighted average across all years of data-taking.

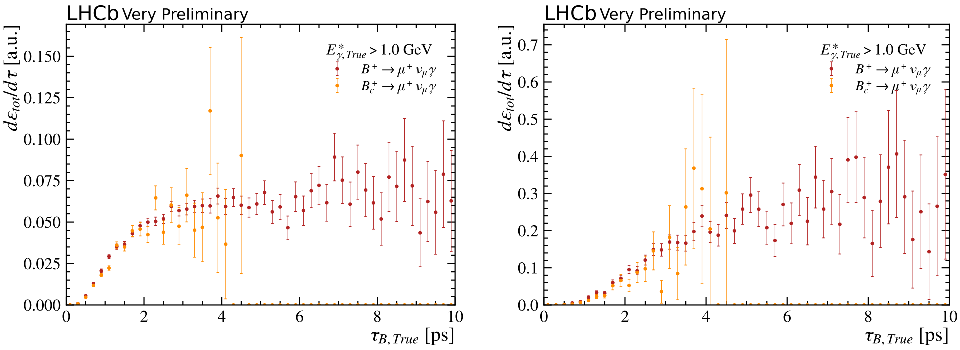


Figure 8.1: Signal selection efficiency as a function of the true B decay time without BDT selection (left) and with the BDT selection applied (right).

The total efficiencies for the signal and normalisation channels are listed in Table 8.4. They are given per year of data-taking as the product of the efficiencies listed in Tables 8.1, 8.2 and 8.3, as well as the weighted average across all years of data-taking.
The efficiency of selecting $B_c^+ \to \mu^+\nu_\mu\gamma$ decays is about an order of magnitude smaller than for the corresponding decay of the B^+. This is expected from the significantly shorter life-time of the B_c^+ meson ($\tau_{B_c^+} = 0.510\,\text{ps}$) compared to the B^+ ($\tau_{B^+} = 1.638\,\text{ps}$) [5]. Figure 8.1 shows the signal selection efficiencies as a function of the true B decay time, which agrees well between the B^+ and B_c^+ channels for short decay times.

8.5 Single event sensitivities

In general, the branching fraction, $\mathcal{B}$, for a B meson decaying to a final state f can be measured from the number of observed signal candidates, $N_{B\to f}$, through

$$\mathcal{B}(B \to f) = N_{B\to f} \times \mathcal{L}_{int} \times \sigma_{pp\to b\bar{b}} \times f_i \times \varepsilon_{B\to f}, \tag{8.6}$$

where $\mathcal{L}_{int}$ is the total integrated luminosity, $\sigma_{pp\to b\bar{b}}$ is the $b\bar{b}$ cross-section, f_i is the fragmentation fraction and $\varepsilon_{B\to f}$ is the total efficiency of the event selection. Measuring the branching fractions of $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays relative to $B^0 \to K^{*0}\gamma$, this simplifies to

$$\begin{aligned}\mathcal{B}_{\text{sig}} &= N_{sig} \times \alpha_{sig} \\ &= \frac{N_{\text{sig}}}{N_{B^0\to K^{*0}\gamma}} \times \frac{f_{\text{d}}}{f_{\text{sig}}} \times \mathcal{B}\left(B^0 \to K^{*0}\gamma\right) \times \mathcal{B}\left(K^{*0} \to K^+\pi^-\right) \times \frac{\varepsilon_{B^0\to K^{*0}\gamma}}{\varepsilon_{\text{sig}}}\end{aligned} \tag{8.7}$$

Parameter	**Value**
$N_{B^0\to K^{*0}\gamma}$	692 ± 31 (from Sec. 7.2)
$\frac{f_d}{f_c}$	132 ± 31 [48]
$\mathcal{B}(B^0 \to K^{*0}\gamma) \times \mathcal{B}(K^{*0} \to K^+\pi^-)$	$(2.79 \pm 0.17) \times 10^{-5}$ [5]

Table 8.5: Inputs used in the calculation of the single event sensitivity.

where "sig" denotes either $B^+ \to \mu^+\nu_\mu\gamma$ or $B_c^+ \to \mu^+\nu_\mu\gamma$ decays. The term α_{sig} defines the single-event sensitivity, which relates the number of observed signal candidates to the respective branching ratio. The number of observed candidates in data, N_{sig} and $N_{B^0\to K^{*0}\gamma}$, are obtained from the procedures described in Sections 6.5 and 7.2, respectively and the ratios of efficiencies are calculated above. The external input on $\mathcal{B}(B^0 \to K^{*0}\gamma) \times \mathcal{B}(K^{*0} \to K^+\pi^-)$ is taken from [5] and amounts to $(2.79 \pm 0.17) \times 10^{-5}$. With the inputs summarised in Table 8.5 the single event sensitivities can be calculated to

$$\begin{aligned}
\alpha_{B^+} &= (5.3 \pm 0.3 \pm 0.3) \times 10^{-8} \\
\alpha_{B_c^+} &= (5.1 \pm 0.4 \pm 0.3) \times 10^{-7} \times \frac{f_d}{f_c} = (6.7 \pm 0.6 \pm 0.4 \pm 1.6) \times 10^{-5},
\end{aligned} \tag{8.8}$$

where the first uncertainty is the systematic uncertainty due to the finite size of the simulation samples and the uncertainty on the observed number of $B^0 \to K^{*0}\gamma$ events in data. The second uncertainty is due to the external uncertainty on the measured branching fraction of the $B^0 \to K^{*0}\gamma$ decay. The third uncertainty in the last term of $\alpha_{B_c^+}$ is due to the external uncertainty on the ratio of fragmentation fractions $\frac{f_d}{f_c}$.

8.6 Model dependence

The total selection efficiencies are obtained from the simulation based on the leading-order calculation from Ref. [25]. However, the latest prediction from Ref. [33] shows a significant dependence of distributions in E_γ^* and $cos\theta_W$ on the choice of LCDA model, in particular the value of λ_B (see Figure 5.1) and is discussed in more detail in Section 5.1. As a result of the non-uniform selection efficiency in E_γ^* and $cos\theta_W$, shown in Figure 8.2, this introduces an inherent model dependence on the derived selection efficiency and thus on the derived branching fractions.

To resolve this model dependence, the two-dimensional efficiency maps as a function of E_γ^* and $cos\theta_W$ are provided together with the final result on the branching fraction. They allow for re-interpretation of the result on the signal branching fractions obtained in the leading-order LCDA model for an arbitrary choice of the LCDA model. The efficiency maps shown in Figure 8.3 are calculated from the signal simulation samples generated with a phase-space model to ensure good coverage across the entire range of E_γ^* and $cos\theta_W$.

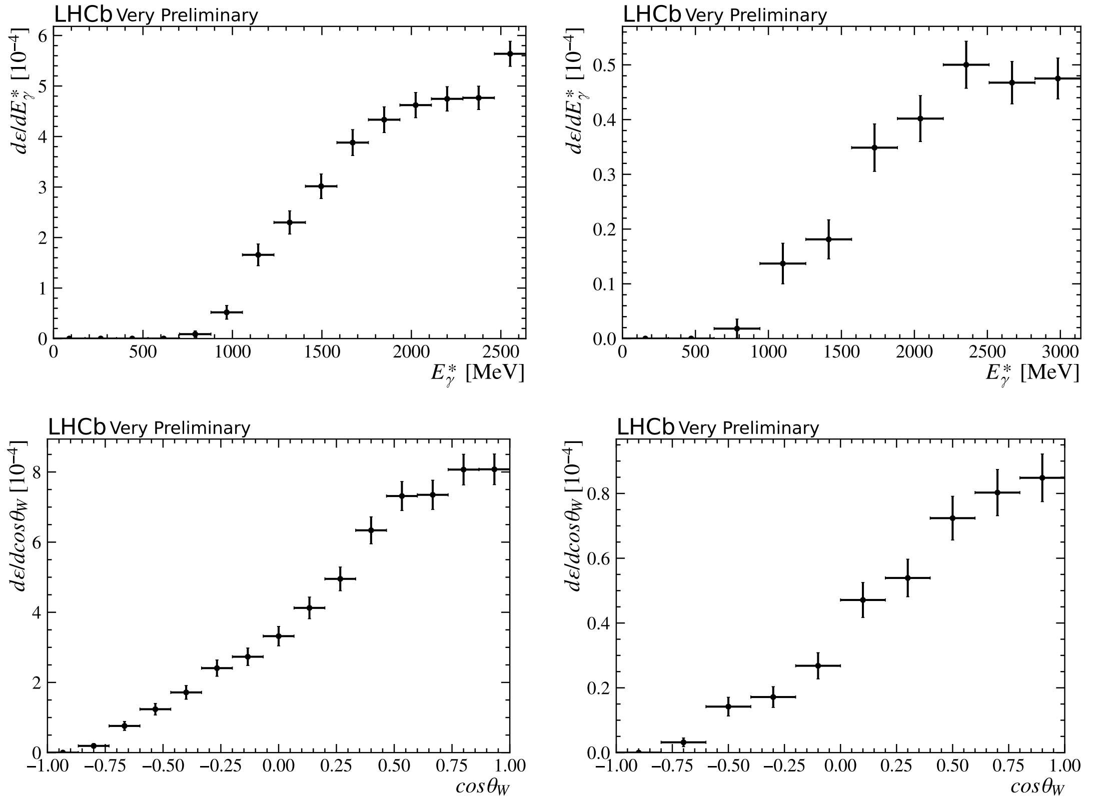


Figure 8.2: Reconstruction and selection efficiency for $B^+ \to \mu^+\nu_\mu\gamma$ candidates (left) and $B^+_c \to \mu^+\nu_\mu\gamma$ candidates (right) as a function of the true E^*_γ (top) and the true decay angle $cos\theta_W$ (right). The efficiencies are obtained from $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ simulation samples generated with a phase-space model.

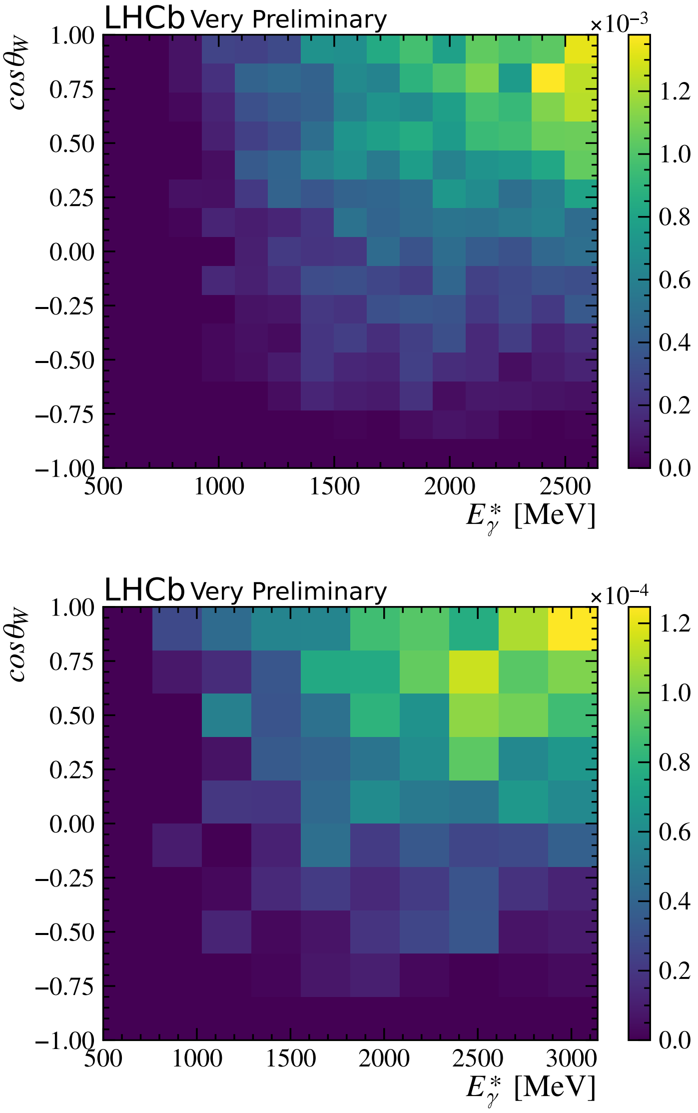


Figure 8.3: Two-dimensional map of the reconstruction and selection efficiency for $B^+ \to \mu^+\nu_\mu\gamma$ candidates (top) and $B_c^+ \to \mu^+\nu_\mu\gamma$ candidates (bottom) as a function of the true E^*_γ and the true decay angle $cos\theta_W$. The efficiencies are obtained from $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ simulation samples generated with a phase-space model.

8.7 Expected statistical sensitivity

In this section, the expected statistical sensitivities on the signal branching fractions are determined. First, the statistical uncertainties are computed, followed by a discussion of the statistical method to obtain the significance of observing a positive signal.

Statistical uncertainty

Typically, the confidence intervals in maximum likelihood fits are derived from profiling the log-likelihood function, which is approximately χ^2-distributed according to Wilk's theorem [111]. The introduction of additional nuisance parameters with Gaussian constraints in the method from Conway (see Section 6.1) breaks the assumption of Wilk's theorem, resulting in a log-likelihood that is no longer approximately χ^2-distributed. In addition, the proximity of the total PDF to being negative in some bins can cause the log-likelihood to be asymmetric. For this reason, the confidence intervals on the measured branching fractions are determined from Feldman-Cousins scans [112]. The method constructs confidence intervals around the measured value, $\hat{\mu}$, by scanning the coverage probability in steps of the signal parameter, μ, using a likelihood ratio evaluated on data and a series of pseudo-experiments.

While this method yields the correct coverage, it is computationally quite expensive. It will be used to obtain the confidence intervals on the measured signal branching fractions once the data is unblinded. However, since the central values are still blinded, the expected statistical uncertainty is approximated using a much simpler approach. A series of pseudo-experiments is conducted with a realistic number of signal candidates injected. For the decay $B^+ \to \mu^+\nu_\mu\gamma$, this corresponds to the upper limit on the branching fraction of 3.0×10^{-6} from Belle [1] and for $B_c^+ \to \mu^+\nu_\mu\gamma$ to the Standard Model prediction from Ref. [2] of 8×10^{-5}. The expected statistical uncertainty is derived as the interval which covers 68% of the fitted signal branching fractions around the median value. The values obtained are $\Delta\mathcal{B}(B^+ \to \mu^+\nu_\mu\gamma) = {}^{+1.49}_{-1.33} \times 10^{-6}$ and $\Delta\mathcal{B}(B_c^+ \to \mu^+\nu_\mu\gamma) = {}^{+2.90}_{-2.16} \times 10^{-4}$.

8.7.1 Strategy and performance for the observation of signal

In case a positive signal is observed, the correct significance of the observation is obtained from a one-sided hypothesis test on an ensemble of pseudo-experiments, testing the signal+background hypothesis against the background-only hypothesis. The preference of the signal+background hypothesis is quantified by means of the *test-statistic* q_0 defined in Eq. 12 of Ref. [113]

$$q_0 = \begin{cases} -2\log\frac{\mathcal{L}(0,\hat{\hat{\boldsymbol{\theta}}}(0))}{\mathcal{L}(\hat{\mu},\hat{\boldsymbol{\theta}})} & \hat{\mu} \geq 0 \\ 0 & \hat{\mu} < 0 \end{cases} \tag{8.9}$$

The test statistic utilises the ratio of the likelihood, $\mathcal{L}$, of an unconstrained fit with signal strength $\hat{\mu}$ and a fit under the background-only hypothesis, $\mu = 0$. In the case the unconstrained fit yields a negative signal parameter, $\hat{\mu} < 0$, the test statistic is set to zero to indicate perfect compatibility with the hypothesis, $\mu = 0$. This definition reflects the boundary condition that any physical signal must be non-negative. The parameters $\hat{\boldsymbol{\theta}}$ represent the nuisance parameters of the unconstrained fit and $\hat{\hat{\boldsymbol{\theta}}}(\mu)$ are the conditional nuisance parameters under a certain signal hypothesis μ.

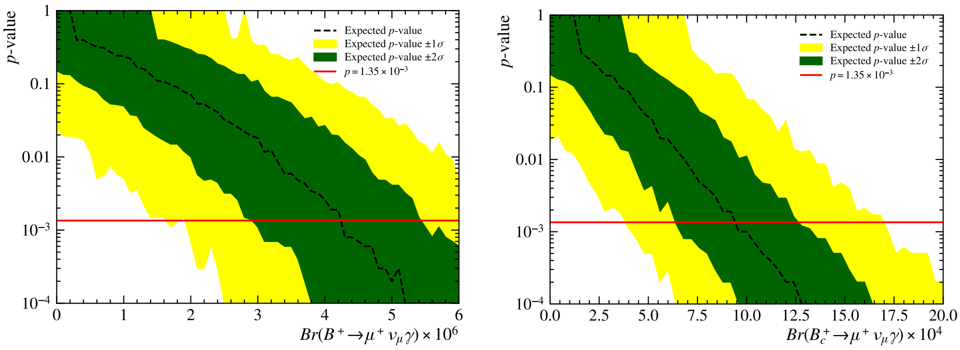


Figure 8.4: Expected p-value for an observation of $B^+ \to \mu^+\nu_\mu\gamma$ (left) and $B_c^+ \to \mu^+\nu_\mu\gamma$ (right) decays as a function of the branching fraction derived from pseudo-experiments. The red line indicates the p-value corresponding to an observation with a significance of three standard deviations. Only the statistical uncertainty is considered.

Let $q_{0,data}$ be the value of the test-statistic evaluated on data and $f(q_0|0)$ the distribution of the test statistic evaluated on background-only pseudo-experiments, normalised to unity. Then the p-value of observing a positive signal in data is defined as the fraction of background-only pseudo-experiments in which the compatibility with the signal+background hypothesis exceeds the compatibility found in data

$$p = \int_{q_{0,data}}^{\infty} dq_0 \, f(q_0|0). \tag{8.10}$$

While the analysis is still blinded, the test statistic cannot be evaluated on data. However, it can be evaluated on pseudo-data generated with different levels of signal injected to study the *expected* p-value of an observation. Figure 8.4 shows the expected p-value for different branching fraction hypotheses, evaluated on 1000 signal-induced and 10,000 background-only pseudo-experiments. To claim the observation of a positive signal with a significance of three standard deviations, the p-value must fall below 1.35×10^{-3}. For the decay $B^+ \to \mu^+\nu_\mu\gamma$ this is expected for a branching fraction of $\mathcal{B}(B^+ \to \mu^+\nu_\mu\gamma) \gtrsim 3.6 \times 10^{-6}$ or 72 signal candidates, which is just above the limit set by the Belle experiment at 3.0×10^{-6} [1]. For the measurement of the branching fraction of $B_c^+ \to \mu^+\nu_\mu\gamma$, the observation of a positive signal with a significance exceeding 3σ is expected for a branching fraction of $\mathcal{B}(B_c^+ \to \mu^+\nu_\mu\gamma) \gtrsim 1.2 \times 10^{-3}$ or 20 signal candidates, which is about an order of magnitude larger than most Standard Model predictions. Therefore, it is expected that the statistical sensitivity of the analysis is insufficient to find evidence of the decays $B^+ \to \mu^+\nu_\mu\gamma$ or $B_c^+ \to \mu^+\nu_\mu\gamma$ and instead an upper limit on the branching fraction is provided. The expected upper limit is calculated in Chapter 10, including all sources of systematic uncertainties, which are discussed in the following.

9

Systematic uncertainties

The search for $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$ decays is expected to be dominated by the statistical uncertainty. Nonetheless, the systematic uncertainties affecting the measured branching fractions are thoroughly evaluated. The sources of systematic uncertainties are divided into two categories: In Section 9.1, systematic uncertainties related to the single event sensitivity are discussed. In Section 9.2, the systematic uncertainties associated with the fit of the corrected mass are derived.

9.1 Systematic uncertainties on the efficiency ratio

In this section, the systematic uncertainties linked to the single event sensitivity are discussed. As such, they are directly proportional to the number of observed signal candidates and are derived as a relative uncertainty on the respective branching fraction. They are summarised in Table 9.1. The systematic uncertainty on the number of observed $B^0 \to K^{*0}\gamma$ candidates is directly taken from the fit of the normalisation channel in Section 7.2. The external uncertainty on the normalisation branching fraction is taken from the PDG [5] and the uncertainty on the ratio of fragmentation fractions f_d/f_c from Ref. [48]. The remaining systematic uncertainties linked to the efficiency calculation for the signal and normalisation channels are detailed in the following.

Tracking efficiency

The corrections of the tracking efficiencies applied in Section 5.7.3 have a standard systematic uncertainty of 0.8% assigned per track, due to the finite precision with which the tracking efficiency is measured in data. Due to the non-cancellation of the tracking efficiency between the signal (three tracks) and the normalisation channels (four tracks), a systematic uncertainty of 0.8% remains.

Hadronic interactions

The corrections of the tracking efficiency from the collaboration-wide TrackCalib package are derived from $J/\psi \to \mu^+\mu^-$ decays and consequently apply to the tracking efficiency of muons. While these can also be applied to tracking efficiencies for hadrons, there is

Source	$\frac{\Delta\mathcal{B}}{\mathcal{B}}(B^+ \to \mu^+\nu_\mu\gamma)$ [%]	$\frac{\Delta\mathcal{B}}{\mathcal{B}}(B_c^+ \to \mu^+\nu_\mu\gamma)$ [%]
Tracking efficiency	0.8	0.8
Hadronic interactions	1.8	1.8
PID efficiency	1.8	1.8
L0 efficiency	2.0	2.0
m_{HOP}	1.0	1.0
Simulation statistics	3.5	7.6
Normalisation yield	4.5	4.5
$\mathcal{B}(B^0 \to K^{*0}\gamma)$	6.0	6.0
f_d/f_c	-	23.8
Total	9.0	26.3

Table 9.1: Summary of the systematic uncertainties linked to the single event sensitivity. The relative uncertainties listed are directly proportional to the respective branching fractions and are given in percent.

an additional systematic uncertainty due to hadronic interactions with the detector material, which are not accounted for in the tracking efficiency tables. From simulations of $B^0 \to K^{*0}(\to K^+\pi^-)J/\psi$, it is found that about 11% of kaons and 14% of pions cannot be reconstructed as long tracks due to hadronic interactions before the last tracking station [60]. With a standard uncertainty of 10% on the material budget, this leads to a systematic uncertainty of 1.1% on the kaon and 1.4% on the pion tracking efficiency. The resulting systematic uncertainty due to hadronic interactions calculated from

$$\left(\frac{\Delta\varepsilon}{\varepsilon}\right)_{Hadr.Int.} = \sqrt{0.011^2 + 0.014^2} = 0.018. \tag{9.1}$$

PID Efficiency

The finite statistics of the data control samples used to derive the PID efficiencies introduce a systematic uncertainty. It is evaluated by propagating the uncertainty on the PID efficiency in the simulation of the signal and normalisation channels. The resulting systematic uncertainty on the signal branching fractions is calculated to be 1.8%.

L0 efficiency

For signal and normalisation candidates in the L0E trigger category, the L0 efficiency of the electrons is assumed to factorise into the `L0Electron` efficiencies for the two electrons. This assumption introduces a systematic uncertainty on the L0E trigger efficiency, which is quantified by generating an alternative correction as a function of the maximum E_T of the two electrons. The resulting change in the total selection efficiency of 2.0% is assigned as systematic uncertainty on the signal branching fractions.

Efficiency due to selection on m_{HOP}

$B^0 \to K^{*0}\gamma$ normalisation candidates are selected with a cut of $m_{HOP} > 4900\,\text{MeV}$ before correcting for differences between data and simulation. This introduces a systematic uncertainty on the selection efficiency of the normalisation channel. In previous studies of the decays $B^+ \to K^+ J/\psi$ and $B^0 \to K^{*0} J/\psi$ [94], the size of the systematic uncertainty is found to be consistent between the two decays and independent of the di-electron invariant mass. The systematic uncertainty on the selection efficiency of $B^0 \to K^{*0}\gamma$ decays due to the selection on m_{HOP} is therefore evaluated using $B^+ \to K^+ J/\psi(\to e^+e^-)$ decays. $B^+ \to K^+ J/\psi(\to e^+e^-)$ candidates are selected in data and simulation with and without the requirement $m_{HOP} > 4900\,\text{MeV}$. From the selected number of candidates, the efficiency of the cut on m_{HOP} is derived separately in data and simulation. The discrepancy between the efficiency in data and simulation is found to be 1.0% and is assigned as a systematic uncertainty on the efficiency ratio.

Simulation statistics

The finite size of the simulation samples for the $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ and $B^0 \to K^{*0}\gamma$ decays introduces an uncertainty on the selection efficiency. This is reflected in the uncertainty on the total efficiency listed in Table 8.4. The systematic uncertainty on the efficiency ratio takes into account the uncertainty from the finite statistics of the $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ and $B^0 \to K^{*0}\gamma$ simulation samples and equates to 3.5% for $B^+ \to \mu^+\nu_\mu\gamma$ and 7.6% for $B^+_c \to \mu^+\nu_\mu\gamma$.

Signal physics model

As was briefly discussed in Section 5.1.2, the B^+ decay kinematics show a significant dependence on the parameter λ_B. As the value of λ_B is only loosely constrained, this leaves a wide range of possible decay kinematics in E^*_γ and $\cos\theta_W$. Together with an observed non-uniformity of the $B^+ \to \mu^+\nu_\mu\gamma$ selection efficiency discussed in Section 8.6, this introduces a systematic uncertainty on the selection efficiency.

Instead of absorbing the dependence of the measured branching fractions on the LCDA model into a systematic uncertainty, the model dependence on the selection efficiency is resolved by providing the two-dimensional selection efficiency in $(E^*_\gamma,\ \cos\theta_W)$ together with the final result. They are shown in Figure 8.3 and allow for reinterpretation of the observed number of signal candidates for an arbitrary LCDA model. Consequently, no systematic uncertainty is assigned to the model dependence of the selection efficiency.

Pollution from $B^+_{(c)} \to \mu^+\nu_\mu e^+e^-$ and $B^0 \to K^{*0}e^+e^-$

The decay $B^0 \to K^{*0}\gamma$ is reconstructed from the conversion of a real photon into two oppositely charged electrons. However, the decay can also proceed through a $b \to s\gamma^*$ transition with a virtual photon $\gamma^* \to e^+e^-$. Due to the experimental resolution on the invariant di-electron mass, the two processes cannot be distinguished if the virtuality of the photon is close to zero. A previous LHCb analysis of $B^0 \to K^{*0}\gamma$ decays reconstructed from photon conversions with electron long-tracks has found the pollution from $b \to s\gamma^*$ processes to be of the order of a few percent [95].

However, the decays $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ are subject to the same type of pollution from $B^+_{(c)} \to \mu^+\nu_\mu e^+e^-$ proceeding through a $\gamma^* \to e^+e^-$ transition. Consequently, the corrections to the number of observed signal and normalisation events cancel in the ratio and no additional systematic uncertainty needs to be assigned to the derived branching fractions for $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays.

Source	$\Delta\mathcal{B}(B^+ \to \mu^+\nu_\mu\gamma)$ $[10^{-7}]$	$\Delta\mathcal{B}(B_c^+ \to \mu^+\nu_\mu\gamma)$ $[10^{-5}]$
$B^+ \to \mu^+\nu_\mu\gamma$ model	7.2 (51%)	6.5 (26%)
$B_c^+ \to \mu^+\nu_\mu\gamma$ model	0.9 (7.0%)	8.1 (31%)
B_c^+ template smoothing	0.6 (4.3%)	4.6 (18%)
Background mass shape	5.6 (40%)	8.3 (33%)
MisID yield	0.2 (1.8%)	0.1 (0.4%)
Fit bias	0.6 (4.2%)	2.2 (9.1%)
Table 9.1	2.7 (19%)	2.1 (8.3%)
Total	10.0 (79%)	14.5 (58%)

Table 9.2: Summary of the systematic uncertainties evaluated on pseudo-data generated with $\mathcal{B}(B^+ \to \mu^+\nu_\mu\gamma) = 3.0 \times 10^{-6}$ corresponding to the upper limit from Belle [1] and $\mathcal{B}(B_c^+ \to \mu^+\nu_\mu\gamma) = 8 \times 10^{-5}$ corresponding to the Standard Model prediction from Ref. [2]. The systematic uncertainties from Table 9.1 are converted for the same branching fractions. The numbers in brackets reflect the relative size of the systematic uncertainty with respect to the expected statistical uncertainty.

9.2 Systematic uncertainties on the signal extraction

Systematic uncertainties linked to the extraction of the observed signal yield are determined from studies of pseudo-experiments. In these studies, the pseudo-data are generated from an alternative model, where the corrected mass distributions or background yields are varied, and then fitted with the nominal fit model. The size of the systematic uncertainty is derived from the shift of the central value of the fitted signal branching fractions across 1000 pseudo-experiments. While the analysis is still blinded, the pseudo-experiments are generated with a true signal injected according to the Belle upper limit on the $B^+ \to \mu^+\nu_\mu\gamma$ branching fraction of 3.0×10^{-6} [1], corresponding to 56 signal candidates, and the Standard Model prediction of 8×10^{-5} [2] for the $B_c^+ \to \mu^+\nu_\mu\gamma$ branching fraction, corresponding to one signal candidate. The derived systematic uncertainties are summarised in Table 9.2. After unblinding, the systematic uncertainties discussed in this section will be re-evaluated based on the number of observed signal and background candidates in the data.

9.2.1 Signal physics model

The model dependence on the LCDA model and the value of λ_B may not only affect the selection efficiency but also the corrected mass distribution of $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ candidates. This introduces a model dependence on the extracted number of observed signal candidates in the fit to the corrected mass distribution in data that cannot be resolved.

The systematic uncertainty is quantified using simulation samples of $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays, which are generated from a phase-space model. Figure 5.1 suggests that the nominal LHCb simulation based on the leading-order physics model from Ref. [25] and the phase-space model mark the extremes of the expected model dependence. A comparison of the corrected mass distributions for simulated $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decays generated from the leading-order physics model and a phase-space model is shown in Figure 9.1. Separate pseudo-experiments are conducted in which the $B^+ \to \mu^+\nu_\mu\gamma$ and $B_c^+ \to \mu^+\nu_\mu\gamma$ signal is

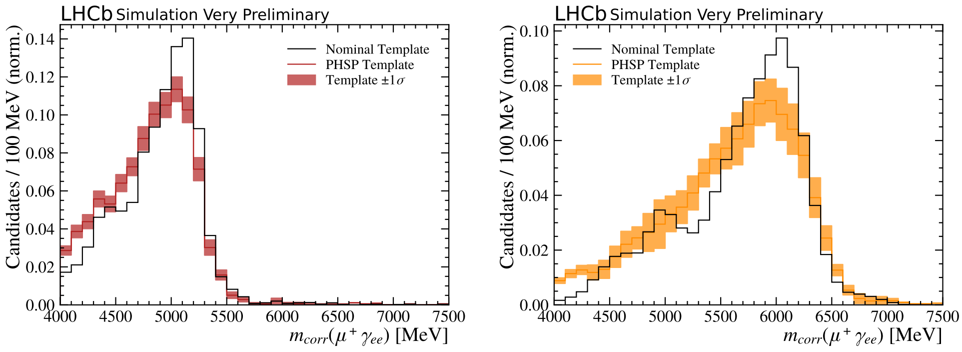


Figure 9.1: Comparison of the nominal corrected mass template for $B^+ \to \mu^+\nu_\mu\gamma$ decays (left) and $B_c^+ \to \mu^+\nu_\mu\gamma$ decays (right) derived from simulation with the leading-order physics model of [25] and a phase-space model. Only the uncertainty of the phase-space template is shown.

injected according to the phase-space model, respectively. The systematic uncertainties on the $B^+ \to \mu^+\nu_\mu\gamma$ branching fraction are found to be 7.2×10^{-7} and 0.9×10^{-7} when using the alternative B^+ and B_c^+ signal model, respectively. For the branching fraction of $B_c^+ \to \mu^+\nu_\mu\gamma$, the systematic uncertainties are 6.6×10^{-5} and 8.1×10^{-5}, respectively. Note that these uncertainties are expected to scale approximately linearly with the actual signal yield. Especially for the B^+ signal, the uncertainty is conservative as a result of the pseudo-data being generated for a branching ratio of 3.0×10^{-6} and is likely to change for the unblinded result.

9.2.2 Signal mass smoothing

The systematic uncertainty linked to the smoothing of the signal mass distribution for $B_c^+ \to \mu^+\nu_\mu\gamma$ candidates is evaluated by generating pseudo-data from the template without the smoothing applied. The resulting systematic uncertainty is assessed to be 0.6×10^{-7} for the B^+ and 4.6×10^{-5} for the B_c^+ signal.

9.2.3 Background mass shape

For the systematic uncertainty linked to the background model, two separate studies are conducted. In a first step, the effect of the smoothing of the $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ background templates is addressed. Alternative background models are generated in which the Gaussian kernel in the smoothing is reduced from 100 MeV to 50 MeV. The pseudo-data are then generated from a linear interpolation between the nominal model, $\mathcal{P}_0$, and the alternative model, $\mathcal{P}_{alt}$, according to $\xi \times \mathcal{P}_0 + (1-\xi) \times \mathcal{P}_{alt}$, where $\xi \in [0, 1]$. From 1000 pseudo-experiments, the systematic uncertainty on the signal branching fractions is evaluated to be 5.3×10^{-7} for the B^+ and 1.6×10^{-5} for the B_c^+ signal.
In a second step, the uncertainty on the $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ background model introduced by the method to extrapolate event information to low $p_T(\gamma_{calo})$ (see Section 6.3.4) is assessed. Instead of using the extrapolation only for soft photons, the method is extended across the full $p_T(\gamma_{calo})$ range and alternative templates are derived solely from the extrapolation method. Pseudo-data are generated using linear interpolation between the nominal and alternative background models. The resulting systematic uncertainties on the branching fractions are 1.8×10^{-7} for the B^+ and 8.1×10^{-5} for the B_c^+ signal.

The total systematic uncertainty linked to the background templates is obtained by adding the two uncertainties, assuming them to be uncorrelated. This yields a total systematic uncertainty of 5.6×10^{-7} on the B^+ and 8.3×10^{-5} on the B_c^+ signal branching fractions.

9.2.4 MisID yield

The expected number of background candidates from $h^+ \to \mu^+$ misidentification is evaluated from PIDCalib2, which is known to underestimate the contribution from π^+ or K^+ decaying in flight. To account for this, pseudo-data are generated in which the background from misidentification is enlarged by a conservative factor of two. The systematic uncertainty is evaluated to be 0.2×10^{-7} for the B^+ and 1.0×10^{-4} for the B_c^+ signal branching fractions.

9.2.5 Fit bias

The stability of the fit bias is studied as a function of the injected signal across a wide range of signal strengths in Section 6.6. For the extracted number of signal candidates, a bias of -5.8 ± 1.1 is found for the $B^+ \to \mu^+\nu_\mu\gamma$ signal and -0.9 ± 0.5 for the $B_c^+ \to \mu^+\nu_\mu\gamma$ signal. The mean bias is corrected for the final result and the uncertainty is assigned as a systematic uncertainty. This yields an uncertainty of 0.6×10^{-7} on the branching fraction for $B^+ \to \mu^+\nu_\mu\gamma$ decays and 2.2×10^{-5} for $B_c^+ \to \mu^+\nu_\mu\gamma$ decays .

10 Results and concluding remarks

In Section 10.1, the blinded results on the $B^+ \to \mu^+\nu_\mu\gamma$ and $B_c^+ \to \mu^+\nu_\mu\gamma$ branching fractions are presented. The expected upper limits on the signal branching fractions are calculated from the CL_s method and presented in Section 10.2. In Section 10.3, some concluding remarks are provided and future prospects of the analysis at LHCb are discussed.

10.1 Blinded results

Combining the expected statistical and systematic uncertainties on the branching fractions, the following blinded results are obtained:

$$\begin{aligned} \mathcal{B}(B^+ \to \mu^+\nu_\mu\gamma)_{E_\gamma^*>1\,\mathrm{GeV}} &= (x.x \pm {}^{+1.5}_{-1.3}\,(\mathrm{stat}) \pm 1.0\,(\mathrm{syst})) \times 10^{-6}, \\ \mathcal{B}(B_c^+ \to \mu^+\nu_\mu\gamma)_{E_\gamma^*>1\,\mathrm{GeV}} &= (x.x \pm {}^{+2.9}_{-2.2}\,(\mathrm{stat}) \pm 1.4\,(\mathrm{syst})) \times 10^{-4}. \end{aligned} \tag{10.1}$$

The systematic uncertainties are evaluated assuming a signal branching fraction on the $B^+ \to \mu^+\nu_\mu\gamma$ decay of 3.0×10^{-6}, corresponding to the current best upper limit from Belle [1], and a branching fraction of 8×10^{-5} for the decay $B_c^+ \to \mu^+\nu_\mu\gamma$ according to the Standard Model prediction from Ref. [2]. Once the central values are unblinded, the statistical coverage will be evaluated from Feldman-Cousins scans to ensure correct coverage. The systematic uncertainties linked to the signal extraction will then be re-evaluated on pseudo-experiments generated with the observed number of signal candidates in data. It is noted that the systematic uncertainties quoted in the blinded results above are evaluated for conservative choices of the signal branching fractions. For example, the value of $\lambda_B = 338 \pm 68\,\mathrm{MeV}$ derived from lattice QCD [30] would lead to a branching fraction around 1.2×10^{-6} in the latest prediction from Ref. [33] using an exponential LCDA model. Similarly, the branching fraction of 8×10^{-5} for the decay $B_c^+ \to \mu^+\nu_\mu\gamma$ is at the upper end of the range of Standard Model predictions [2, 39–42]. In both cases, evaluating the systematic uncertainties on pseudo-experiments generated for smaller signal branching fractions would reduce the sizable uncertainties on the signal physics models. Consequently, the systematic uncertainties will potentially be reduced once the central values are unblinded.

10.2 Expected upper limits

The sensitivities on the signal branching fractions suggest no significant observation of a positive signal and the upper limits on the signal branching fractions are provided. They are derived using the CL_s method [114], which is a modification of the previously used frequentist approach, designed to produce conservative upper limits and avoid excluding signal models where the experiment has little sensitivity. Two separate p-values are derived under the signal+background hypothesis and the background-only hypothesis. From this, the CL_s confidence level is defined as

$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{p_{s+b}}{1-p_b}, \tag{10.2}$$

where p_{s+b} is the probability of preferring the signal+background hypothesis and $1-p_b$ is the probability of rejecting the background-only hypothesis. By construction, the CL_s prevents the exclusion of a signal model that does not yield a significantly better description of the data than the background alone.

For the construction of the upper limit of a non-negative signal, the $\widetilde{q}_\mu$ test statistic defined in Eq. 16 of [113] is utilised. It reads

$$\widetilde{q}_\mu = \begin{cases} -2\log\frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(0,\hat{\hat{\boldsymbol{\theta}}}(0))} & \hat{\mu} < 0 \\ -2\log\frac{\mathcal{L}(\mu,\hat{\hat{\boldsymbol{\theta}}}(\mu))}{\mathcal{L}(\hat{\mu},\hat{\boldsymbol{\theta}})} & 0 \le \hat{\mu} \le \mu \\ 0 & \hat{\mu} > \mu \end{cases} \tag{10.3}$$

where μ is the tested branching fraction hypothesis, $\hat{\mu}$ is the preferred signal strength from an unconstrained fit and $\hat{\boldsymbol{\theta}}$ ($\hat{\hat{\boldsymbol{\theta}}}(\mu)$) are the (conditional) nuisance parameters. The distributions of $\widetilde{q}_\mu$ under the signal+background and background-only hypotheses are evaluated on one thousand pseudo-experiments with and without induced signal in the event generation and for each signal hypothesis, μ. Since the analysis is blinded, the *expected* upper limits are computed from pseudo-experiments. Figure 10.1 shows the expected CL_s confidence levels on the signal branching fractions. The expected upper limits at 90% (95%) confidence level are defined as the branching fraction, for which the CL_s falls below 0.10 (0.05). The are evaluated as

$$\begin{aligned} \mathcal{B}_{exp}(B^+ \to \mu^+\nu_\mu\gamma)_{E^*_\gamma>1\,\text{GeV}} &< 2.3\,(2.7)\times 10^{-6}, \\ \mathcal{B}_{exp}(B_c^+ \to \mu^+\nu_\mu\gamma)_{E^*_\gamma>1\,\text{GeV}} &< 6.3\,(7.8)\times 10^{-4}. \end{aligned} \tag{10.4}$$

10.3 Concluding remarks and prospects

The presented analysis marks the first search for $B^+ \to \mu^+\nu_\mu\gamma$ decays at a hadron collider and the first ever search for $B_c^+ \to \mu^+\nu_\mu\gamma$ decays. Signal candidates are selected from photon conversions in the VELO detector material to allow for the reconstruction of the displaced B decay vertex and the computation of the corrected mass. A selection of signal candidates is established using a combination of one-dimensional selection cuts and multivariate analysis. Two dedicated BDTs are trained to help suppress the combinatorial background and background from $\pi^0 \to \gamma_{ee}\gamma$ decays. The number of observed

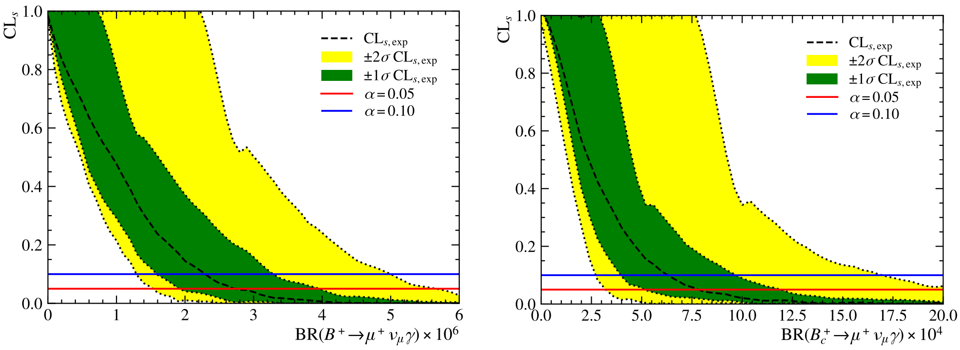


Figure 10.1: Contour of the expected confidence levels on the $B^+ \to \mu^+\nu_\mu\gamma$ (left) and $B_c^+ \to \mu^+\nu_\mu\gamma$ (right) signal branching fraction derived from the CL_s method in the absence of signal. The green and yellow bands reflect the error bands on the expected confidence level drawn in black.

$B^+ \to \mu^+\nu_\mu\gamma$ and $B_c^+ \to \mu^+\nu_\mu\gamma$ signal candidates is extracted from a single binned extended maximum likelihood fit to the corrected mass distribution combining the data recorded from 2016 to 2018.

The dominant source of background in this analysis originates from $\pi^0 \to \gamma_{ee}\gamma$ and $\eta \to \gamma_{ee}\gamma$ decays. To model these backgrounds, a novel method is developed that uses π^0 and η candidates reconstructed from candidates with an additional calorimeter photon. Using additional input from simulation, the efficiency of resolving the $\pi^0 \to \gamma_{ee}\gamma_{calo}$ and $\eta \to \gamma_{ee}\gamma_{calo}$ decays is corrected in a multi-step process to obtain the corrected mass templates that model the distribution of background candidates. Utilising the normalisation channel $B^0 \to K^{*0}\gamma$, the $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ branching fractions are extracted. At the time of writing this thesis, the analysis remains blinded and is currently undergoing review in the collaboration. The expected sensitivity on the branching fraction of $B^+ \to \mu^+\nu_\mu\gamma$ decays indicates a similar sensitivity as the results from the Belle collaboration [1], which provide the current best upper limit of $\mathcal{B}(B^+ \to \ell^+\nu_\ell\gamma) < 3.0 \times 10^{-6}$ at 90% confidence level [1]. For the observation of $B_c^+ \to \mu^+\nu_\mu\gamma$ decays, the current sensitivity is about an order of magnitude looser than most Standard Model predictions [2, 39–42].

While the expected sensitivity in the presented analysis might not be sufficient for a first observation of the $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ decay, there are excellent prospects of observing the decay for the first time using the dataset recorded with the LHCb experiment in Run 3. The aim is to record up to 25 fb^{-1} of pp collisions, which corresponds to a fivefold increase in integrated luminosity compared to the dataset used in this analysis. In addition, the removal of the L0 hardware trigger in Run 3 is expected to lead to a sizeable increase in the signal selection efficiency. Furthermore, the author has contributed to the development of dedicated HLT2 trigger lines for the selection of $B^+_{(c)} \to \ell^+\nu_\ell\gamma_{ee}$ decays in Run 3 with $\ell^+ = e^+, \mu^+$, leading to an additional increase in the selection efficiency and allowing to include the electronic mode in the analysis of Run 3 data. In summary, there are bright prospects for observing the decay $B^+ \to \ell^+\nu_\ell\gamma$ in future analyses at LHCb and significantly improving the limit in the branching fraction of $B_c^+ \to \ell^+\nu_\ell\gamma$ decays.

Part III

Search for $B_c^+ \to J/\psi D^{*+}$ decays

11

Introduction

In the Standard Model, flavour-changing neutral currents (FCNC) are forbidden at tree level and can only be mediated through higher-order loop-level Feynman diagrams. The decay $B^0 \to K^{*0}\ell^+\ell^-$ is a prime example for a *rare decay*, which is mediated by a $b \to s\ell\ell$ transition as shown in Figure 11.1a and has a branching fraction of the order of 10^{-6} [5]. Such rare decays are often used as a probe for *New Physics* (NP), in which new particles and/or interactions are postulated. Some of these models allow for FCNC at tree-level or result in additional loop-level diagrams, which would modify the branching fractions of rare decays, making decays involving $b \to s\ell\ell$ transitions sensitive to New Physics effects. Taking it one step further, the study of $b \to d\ell\ell$ transitions is an even more sensitive probe to New Physics effects due to the additional CKM suppression.

At LHCb, numerous analyses of $b \to s\ell\ell$ transitions have been performed [94, 115–122]. Due to the experimental challenges, the studies are typically performed relative to a normalisation channel. For the aforementioned decay of $B^0 \to K^{*0}\ell^+\ell^-$, the tree-level decay $B^0 \to K^{*0}J/\psi(\to \ell^+\ell^-)$ is typically used as a normalisation channel as it decays to the same final state and offers excellent control over the experimental efficiencies.

Similarly, the decay $B_c^+ \to J/\psi D^{*+}$ is the prime normalisation channel for future studies of $b \to d\ell\ell$ transitions in $B_c^+ \to D^{*+}\ell^+\ell^-$ decays, illustrated in Figure 11.1b. However, the decay $B_c^+ \to J/\psi D^{*+}$ has never been observed and the presented work marks the first search for this particular decay. While predictions of its branching fraction suffer from large uncertainties linked to the nature of the B_c^+ meson, its branching fraction can be estimated relative to the decay $B_c^+ \to J/\psi D_s^+$. Due to the similarities between the two decays, which are illustrated by the leading-order Feynman diagrams shown in

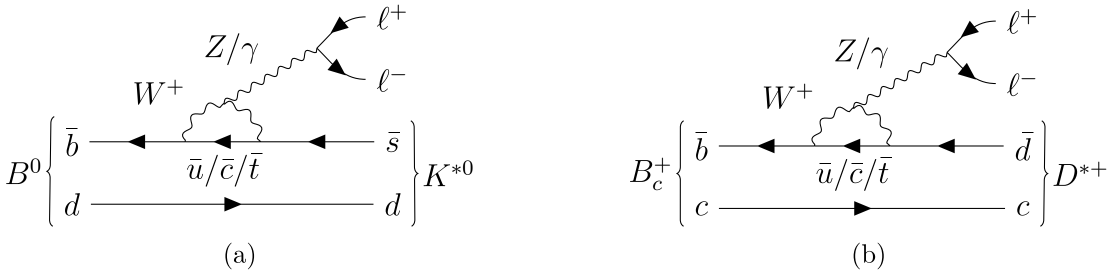


Figure 11.1: Standard Model Feynman diagram for a rare $b \to s\ell\ell$ transition in $B^0 \to K^{*0}\ell^+\ell^-$ (a) and a rare $b \to d\ell\ell$ transition in $B_c^+ \to D^{*+}\ell^+\ell^-$ (b).

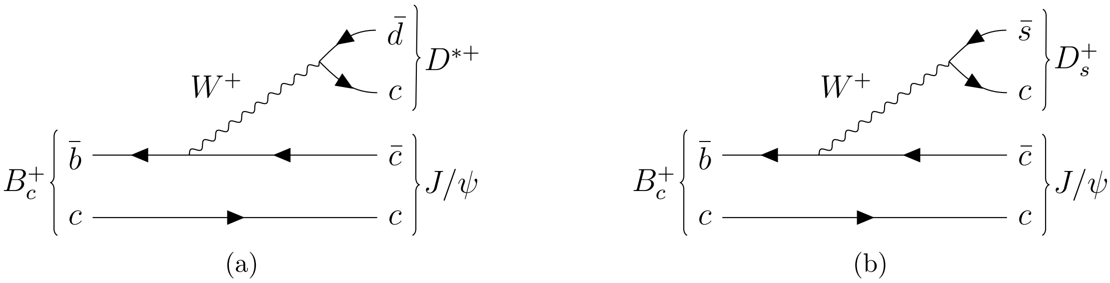


Figure 11.2: Leading order Feynman diagrams for the decay $B_c^+ \to J/\psi D^{*+}$ (a) and $B_c^+ \to J/\psi D_s^+$ (b). Only the colour-favoured tree-level diagrams are shown.

Figure 11.2, two first-principle arguments can be employed. Firstly, it is assumed that the ratio of branching fractions between decays to the vector and pseudoscalar charm mesons is identical in the $D^{(*)+}$ and $D_s^{(*)+}$ system

$$\frac{\mathcal{B}(B_c^+ \to J/\psi D^{*+})}{\mathcal{B}(B_c^+ \to J/\psi D^{+})} = \frac{\mathcal{B}(B_c^+ \to J/\psi D_s^{*+})}{\mathcal{B}(B_c^+ \to J/\psi D_s^{+})} = 1.91 \pm 0.20 \pm 0.07, \tag{11.1}$$

which is measured experimentally in the $D_s^{(*)+}$ system [123]. Secondly, is is assumed, the decays $B_c^+ \to J/\psi D^{(*)+}$ and $B_c^+ \to J/\psi D_s^{(*)+}$ differ only by the involved CKM-matrix element

$$\frac{\mathcal{B}(B_c^+ \to J/\psi D^{*+})}{\mathcal{B}(B_c^+ \to J/\psi D_s^{+})} = \frac{\mathcal{B}(B_c^+ \to J/\psi D_s^{*+})}{\mathcal{B}(B_c^+ \to J/\psi D_s^{+})} \frac{|V_{cd}|^2}{|V_{cs}|^2} = 0.098 \pm 0.011 \tag{11.2}$$

using the values from [123] and [5]. This assumption is supported by the similar masses of the charm mesons of $m_{D^0} = 1864.8$ and $m_{D_s^+} = 1968.4$ [5], resulting in a similar available phase-space for the B_c^+ decay. The expected value for the ratio of branching fractions is used to estimate the number of expected signal candidates in the LHCb dataset and to assess the expected significance for observing $B_c^+ \to J/\psi D^{*+}$ decays.

The decay $B_c^+ \to J/\psi D^{*+}$ is reconstructed at LHCb from $J/\psi \to \mu^+\mu^-$ and $D^{*+} \to D^0\pi^+$ with $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$. The full Run 1 and Run 2 datasets are exploited to maximise the statistical power. The measurement of the branching fraction is performed relative to the decay $B_c^+ \to J/\psi D_s^+$ reconstructed as $D_s^+ \to K^+K^-\pi^+$, which has been studied in a previous LHCb analysis [124]. In Section 12, the event selection is discussed. The extraction of the $B_c^+ \to J/\psi D^{*+}$ branching fraction from a fit to the reconstructed invariant mass in data is established in Section 13 relative to $B_c^+ \to J/\psi D_s^+$ decays. The still blinded results are presented in Section 14.

The event selection, particularly the selection of $B_c^+ \to J/\psi D_s^+$ normalisation candidates and the architecture of the BDT, was developed in collaboration with an internship student at IJCLab, under the supervision of the author.

12 Candidate selection

12.1 Data and simulation samples

The search for $B_c^+ \to J/\psi D^{*+}$ decays is performed on data recorded with the LHCb experiment in pp collisions in 2011 and 2012 at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ and $8\,\text{TeV}$, respectively, and during Run 2 from 2015 to 2018 at $13\,\text{TeV}$. The combined datasets correspond to an integrated luminosity of $9\,\text{fb}^{-1}$.
Monte Carlo simulation samples of the decay $B_c^+ \to J/\psi D^{*+}$ are produced separately for $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ using a phase-space model for the decay of the D^0. In Section 12.6, the simulation sample for $D^0 \to K^-\pi^+\pi^+\pi^-$ is reweighted to a decay model including intermediate resonances. Simulation samples for the normalisation channel $B_c^+ \to J/\psi D_s^+$ with subsequent $D_s^+ \to K^+K^-\pi^+$ decays and for $B_c^+ \to J/\psi\pi^+$ decays were already generated in a previous LHCb analysis [124] using a proper physics amplitude model implemented in EvtGen [125]. All samples are generated separately for each year and magnet polarity.

12.2 Trigger selection

At the L0 level, candidates are required to contain a muon as part of the signal candidate, satisfying the requirements of the `L0Muon` decision or the two signal muons passing the `L0DiMuon` decision. Both decisions require the muon candidates to have subsequent hits in the muon stations with a significant transverse momentum. For the `L0DiMuon` trigger, the requirement on the transverse momentum is imposed on the product of the measured muon transverse momenta.
In the subsequent HLT1 selection, additional requirements are imposed on the muon track or di-muon candidate. In both cases, the (di-)muon candidate is required to have a significant transverse momentum and to originate from a displaced decay vertex. In Run 2, additional MVA-based trigger decisions are used to identify (di-)muon candidates originating from displaced decay vertices, which were not implemented during Run 1.
At the HLT2 level, candidates are required to pass dedicated trigger lines to select $J/\psi \to \mu^+\mu^-$ decays. These require a good quality vertex between the two muon tracks and a resulting di-muon candidate with large transverse momentum and invariant mass that is within $120\,\text{MeV}$ of the known mass of the J/ψ resonance.

12.3 Central event filtering

The requirements of the central event filtering are only applied to the muon tracks and J/ψ-candidates and are listed in Appendix C.1. Good-quality muon tracks pointing back not too closely to the PV are required with a transverse momentum exceeding 500 MeV. In addition, muons must have subsequent hits in the muon stations to satisfy the `isMuon` requirement and the likelihood of being a muon must exceed the likelihood of being a pion.

The combination of two oppositely charged muons must form a good-quality vertex that is significantly displaced from the PV and the di-muon invariant mass must be within 100 MeV of the known mass of the J/ψ-resonance.

12.4 One-dimensional cut selection

Candidates passing the trigger and stripping selections are further selected by imposing one-dimensional selection requirements. The requirements are summarised here and listed in Table 12.1.

Good-quality tracks are required for all final state tracks. Muon candidates must originate from a vertex other than the primary vertex and satisfy a loose momentum selection. Combining two oppositely charged muons, the resulting di-muon candidate must have a good quality vertex and invariant mass within 30 MeV of the known mass of the J/ψ resonance.

Also, kaon and pion candidates must pass loose momentum and PID selections and form a good-quality vertex that is significantly displaced from the PV when combined into a D^0/D_s^+ candidate. For D^0 and D_s^+ candidates reconstructed from multi-hadron final states $K^-\pi^+\pi^+\pi^-$ and $K^+K^-\pi^+$, respectively, the sum of the track χ^2_{IP} must exceed 3σ to suppress background candidates in which one or more tracks originate from the PV. The resulting D^0/D_s^+ candidates are required to have an invariant mass within 25 MeV of the respective known meson masses.

Intermediate $D^{*+} \to D^0\pi^+_{soft}$ decays are reconstructed from an additional *soft pion*, π_{soft}, without the aforementioned loose momentum cuts nor PID cuts applied. To reduce background, a tight cut on the mass difference $\Delta m = m_{D^{*+}} - m_{D^0}$ is imposed.

Lastly, the B_c^+ candidates are selected combining the D^{*+}/D_s^+ candidates with the J/ψ candidates requiring a good quality vertex and a good alignment of the reconstructed B_c^+ candidate momentum and its flight direction.

Figure 12.1 illustrates the selection of the J/ψ and D^0 resonances, as well as the selection in Δm in data and simulation of $B_c^+ \to J/\psi D^{*+}$ candidates reconstructed as $D^0 \to K^-\pi^+$.

12.5 BDT selection

At this point of the selection, the data are still dominated by combinatorial background candidates in which at least one of the reconstructed final state tracks originates from a process other than the signal B_c^+ decay. These background candidates are suppressed with dedicated BDTs, trained separately for the three decay modes $B_c^+ \to J/\psi D_s^+$ and $B_c^+ \to J/\psi D^{*+}$ with $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$. The same XGBoost algo-

Candidate	**Selection**
all tracks	$\chi^2/\mathrm{ndof} < 3$ `GhostProb` < 0.4 `hasRICH`
$\mu^\pm$	$\chi^2_{IP} > 3$ $3\,\mathrm{GeV} < p < 200\,\mathrm{GeV}$ $p_T > 500\,\mathrm{MeV}$ `isMuon`
J/ψ	$\chi^2_{Vertex} < 9$ $\lvert m_{\mu^+\mu^-} - m^{PDG}_{J/\psi}\rvert < 30\,\mathrm{MeV}$
D^0/D^+_s daughters	$2\,\mathrm{GeV} < p < 200\,\mathrm{GeV}$ $p_T > 200\,\mathrm{MeV}$ $\mathtt{ProbNN}_{K,\pi} > 0.1$
D^0/D^+_s	$\chi^2_{Vertex}/\mathrm{ndof} < 16$ PV vertex separation $> 3\sigma$ $\sum_{h^\pm} \chi^2_{IP} > 9$ (not $D^0 \to K^-\pi^+$) $p_T > 1500\,\mathrm{MeV}$ $\lvert m - m^{PDG}_{D^0,D^+_s}\rvert < 25\,\mathrm{MeV}$
D^{*+}	$\chi^2_{Vertex}/\mathrm{ndof} < 9$ $p_T > 2\,\mathrm{GeV}$ $143.5\,\mathrm{MeV} < \Delta m < 147.5\,\mathrm{MeV}$
B^+_c	$\chi^2_{Vertex}/\mathrm{ndof} < 9$ $\chi^2_{IP} < 25$ `DIRA` > 0.99

Table 12.1: Summary of the one-dimensional selection requirements for the selection of $B^+_c \to J/\psi D^{*+}$ and $B^+_c \to J/\psi D^+_s$.

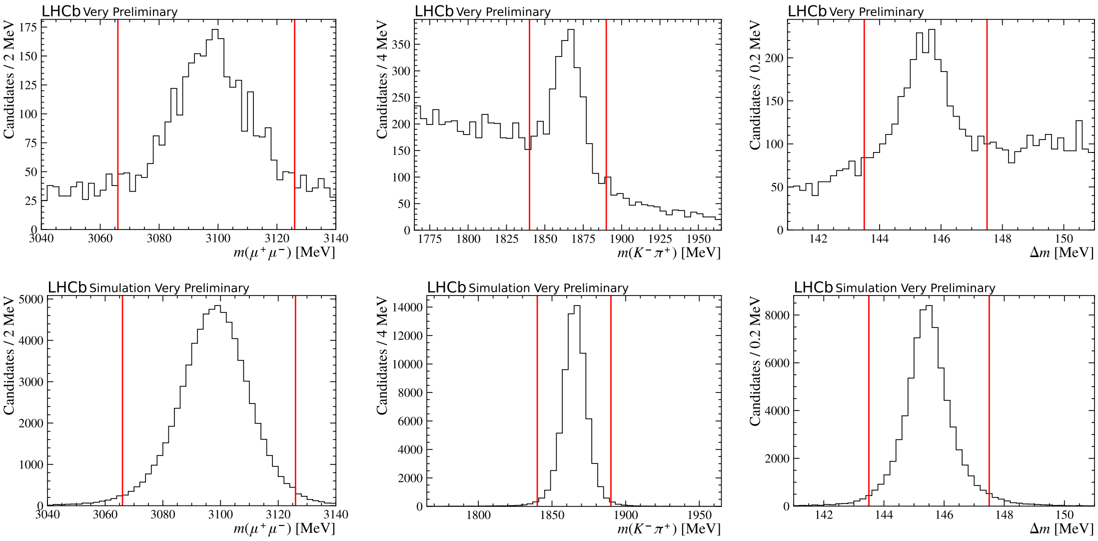


Figure 12.1: Selection of the J/ψ (left) and D^0 (centre) resonances, as well as the selection in Δm (right) of $B_c^+ \to J/\psi D^{*+}$ candidates reconstructed as $D^0 \to K^-\pi^+$ in data (top) and simulation (bottom). The two vertical lines indicate the selection window.

rithm from the SKLEARN software package with a k-fold cross validation of $k = 5$ is used as in the analysis of $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$ detailed in Section 5.6.

Training samples

The classifiers for all three decays are trained using candidates from simulation of the respective decay as signal proxy. The background proxies are taken from candidates in the upper mass sideband $6400\,\text{MeV} < m(J/\psi D_{(s)}^{(*)+}) < 7000\,\text{MeV}$ of data, where the mass of the J/ψ and D^0/D_s^+ candidates are constrained to the known masses. For the simulated samples, all simulation correction weights (see Section 12.6) are applied to ensure correct distributions in the training variables.

Training variables

The three classifiers are trained on similar sets of decay variables that provide significant separation between signal and background candidates and are not strongly correlated with each other. For the BDTs trained for the decay $B_c^+ \to J/\psi D^{*+}$ with $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$, the training variables are: the χ^2_{IP} of all D^0-daughter tracks as well as the sum of the χ^2_{IP} among all D^0-daughter tracks, the minimum transverse momentum among the D^0-daughters and the minimum transverse momentum among the two muons, the `DIRA` angle of the B_c^+ candidate and the significance of its decay vertex being displaced from the PV. An additional training variable is constructed by refitting the reconstructed decay with the mass of the J/ψ and D^0 candidates constrained to their known masses and using the goodness of this fit, $\chi^2_{J/\psi,D^0}/\text{ndof}$, in the training of the BDT.

For the BDT used to isolate the decay $B_c^+ \to J/\psi D_s^+$, the same training variables are used, replacing D^0 by D_s^+ in the definition of the training variables. In addition, the

`DIRA` angle of the B_c^+ candidate is replaced by the significance of D_s^+ decay vertex being displaced from the B_c^+ decay vertex.
The distributions of the training variables in the signal and background proxies for all three classifiers can be found in Appendix C.2

BDT performance

The classifier performance is shown in Figure 12.2 for all three classifiers and all five classifier folds. They all provide excellent separation between signal and background candidates and show no signs of overtraining.

12.5.1 BDT optimisation

Optimisation for the normalisation decay $B_c^+ \to J/\psi D_s^+$

The selection on the BDT response for the selection of $B_c^+ \to J/\psi D_s^+$ decays is optimised by maximising the expected significance of observing a signal, $S/\sqrt{S+B}$, where S is the number of expected $B_c^+ \to J/\psi D_s^+$ normalisation candidates in data and B is the expected number of background candidates in the signal region in data. An extended unbinned maximum likelihood fit is performed to the $m(J/\psi D_s^+)$ distribution of normalisation candidates in data, where the masses of the J/ψ and D_s^+ candidates are constrained to the respective known masses. The fit projection is shown in Figure 12.3a. The number of background candidates, B, for each cut on the BDT response is derived by integrating the background PDF in the signal region $6250\,\text{MeV} < m(J/\psi D_s^+) < 6300\,\text{MeV}$. The expected number of normalisation candidates, S, is instead extrapolated from the fitted number of $B_c^+ \to J/\psi D_s^+$ normalisation candidates for a loose cut on the BDT response of > 0.1 and extrapolated to tighter cuts on the BDT response by the efficiency evaluated on simulation. This ensures the derived significance is not subject to statistical fluctuations in data. The resulting expected significance of $B_c^+ \to J/\psi D_s^+$ in data is shown in Figure 12.3b. The nominal cut on the BDT response is chosen at $BDT > 0.996$, which maximises the expected significance.

Optimisation for the signal decay $B_c^+ \to J/\psi D^{*+}$

The optimisation of the cut on the BDT response for selecting $B_c^+ \to J/\psi D^{*+}$ candidates is determined using a similar strategy, maximising the expected significance, $S/\sqrt{S+B}$. However, as the analysis is blinded, the expected number of signal candidates cannot be determined from a fit to the reconstructed invariant mass of signal candidates in data. Instead, the predicted signal branching fraction from Eq. 11.2 is used to compute the expected number of $B_c^+ \to J/\psi D^{*+}$ signal candidates, N^f_{exp}, reconstructed from a specific final state, $D^0 \to f$, relative to the number of observed $B_c^+ \to J/\psi D_s^+$ decays, $N_{D_s^+}$, as

$$N^f_{exp} = N_{D_s^+} \frac{\mathcal{B}(B_c^+ \to J/\psi D_s^{*+})}{\mathcal{B}(B_c^+ \to J/\psi D_s^+)} \frac{|V_{cd}|^2}{|V_{cs}|^2} \frac{\mathcal{B}(D^{*+} \to D^0\pi^+)\mathcal{B}(D^0 \to f)}{\mathcal{B}(D_s^+ \to K^+K^-\pi^+)} \frac{\varepsilon_f^{MC}}{\varepsilon_{D_s^+}^{MC}}. \tag{12.1}$$

Hereby, $\varepsilon_{D_s^+}^{MC}$ is the total efficiency of selecting $B_c^+ \to J/\psi(\to \mu^+\mu^-)D_s^+(\to K^+K^-\pi^+)$ decays evaluated on the respective simulation samples and ε_f^{MC} is the respective efficiency

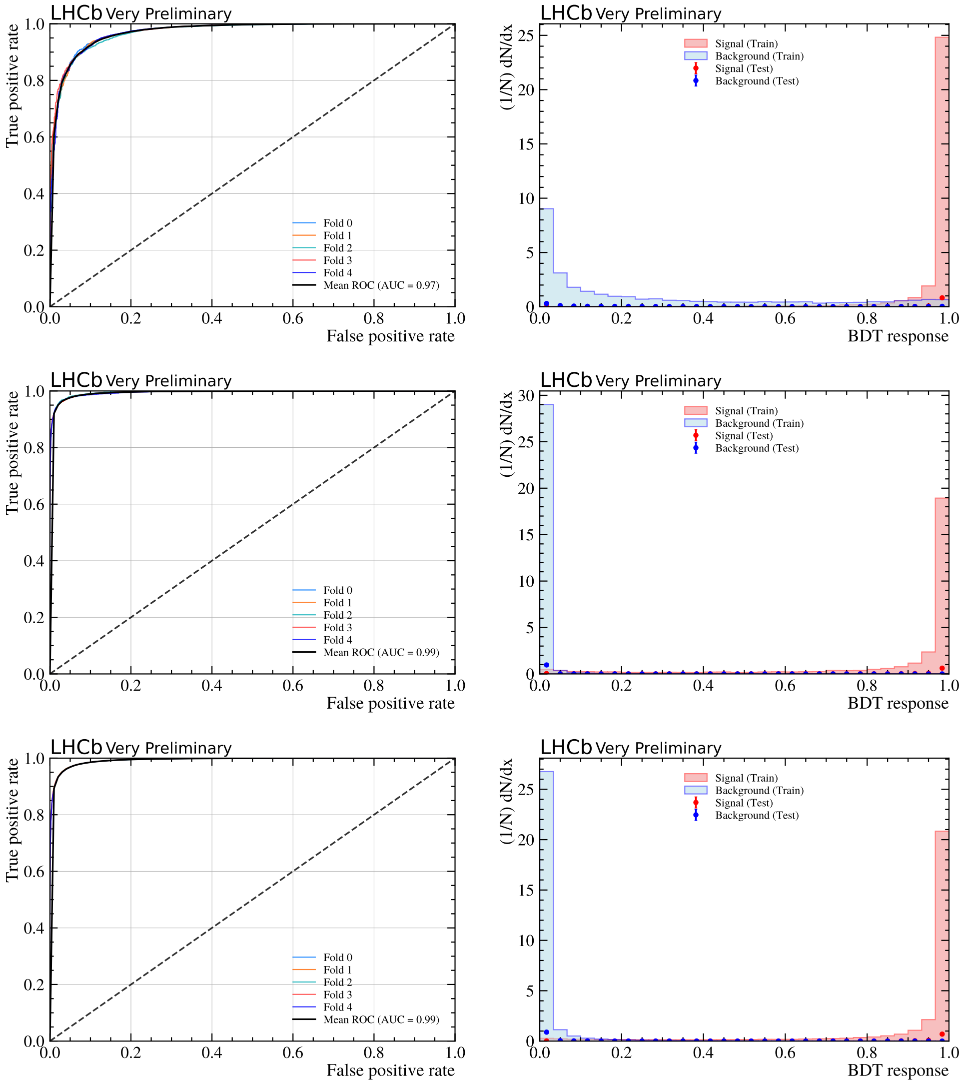
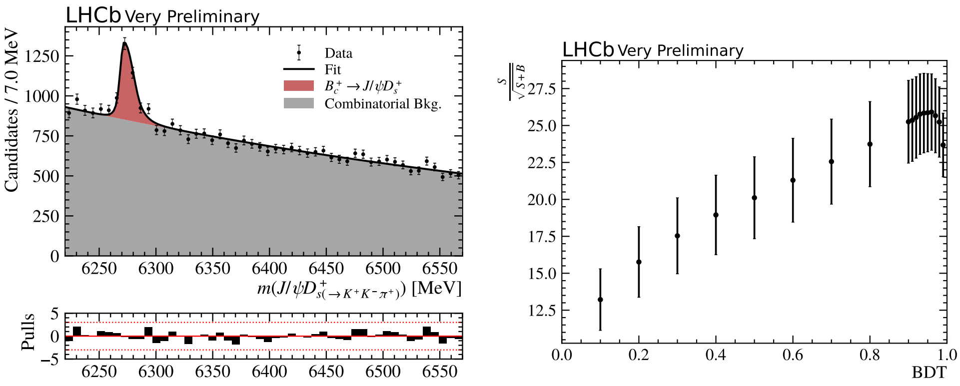


Figure 12.2: ROC curves for the BDT (left) drawn for the five folds and the mean of all folds. Check for overtraining of the BDT for one fold (right). The BDTs for $B_c^+ \to J/\psi D^{*+}$ with $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ are shown on the top and centre and the BDT for $B_c^+ \to J/\psi D_s^+$ is shown on the bottom. The red and blue histograms represent the signal and background proxies used for training, corresponding to 4/5 of the total statistics. The red and blue points represent 1/5 of the training proxies used for testing.



(a) (b)

Figure 12.3: (a) Mass fit to $m(J/\psi D_s^+)$ in data for a cut of BDT > 0.1 and (b) the expected significance for observing $B_c^+ \to J/\psi D_s^+$ decays.

for $B_c^+ \to J/\psi(\to \mu^+\mu^-)D^{*+}(\to D^0\pi^+)$ decays reconstructed in a specific $D^0 \to f$ decay. Using $N_{D_s^+} = 405{\pm}21$ as derived in Section 13.1 and the total efficiencies from Section 12.7, the number of expected signal candidates before the BDT selection is $N_{exp}^{K^-\pi^+} = 36 \pm 5$ and $N_{exp}^{K^-\pi^+\pi^+\pi^-} = 21 \pm 3$. Using the efficiency of the BDT selection on the simulation samples, the number of expected signal candidates is extrapolated to arbitrary cuts on the BDT response.

The number of background candidates, B, for a given cut on the BDT response is determined from a fit of the upper mass sideband $6350\,\text{MeV} < m(J/\psi D^0\pi^+_{soft}) < 6550\,\text{MeV}$ in data and extrapolation to the signal region ±20 MeV around the known B_c^+ mass. The resulting expected significance for $B_c^+ \to J/\psi D^{*+}$ reconstructed in $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ along with an exemplary mass fit is shown in Figure 12.4. The nominal cut on the BDT response is chosen at $BDT > 0.98$ and 0.997, respectively.

12.6 Corrections to simulation

Similar to the analysis of $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$, the search for $B_c^+ \to J/\psi D^{*+}$ decays relies on simulation data for which corrections to simulated candidates need to be applied. Simulation of $D^0 \to K^-\pi^+\pi^+\pi^-$ generated with a phase-space model is corrected in Section 12.6.1 to match a more realistic physics model with intermediate resonances. In Sections 12.6.2 to 12.6.6, similar corrections to the simulation as in the analysis of $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$ decays are applied to correct for the PID efficiency, tracking efficiency, decay kinematics and event multiplicity, as well as the trigger efficiency.

12.6.1 Decay model for $D^0 \to K^-\pi^+\pi^+\pi^-$

The simulation of $B_c^+ \to J/\psi D^{*+}$ decays with $D^0 \to K^-\pi^+\pi^+\pi^-$ is generated using a phase-space model for the decay of the D^0, which does not reflect the vast number of intermediate processes leading to the same final state. As this might affect the selection efficiency, per-event weights are assigned to the phase-space generated simulation to match a more realistic amplitude model.

Using the AmpGen software package [126], a sample of one million $D^0 \to K^-\pi^+\pi^+\pi^-$

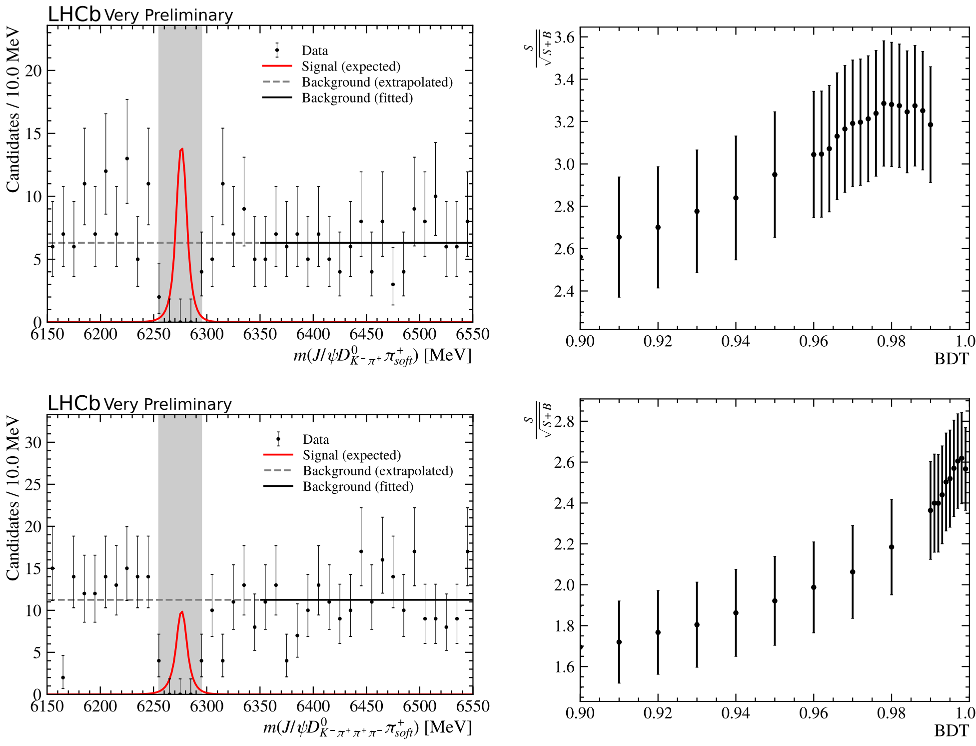


Figure 12.4: Blinded mass fit to the upper mass sideband $m(J/\psi D^0\pi^+_{soft}) > 6350\,\text{MeV}$ in data for a cut on the BDT of > 0.95 (left) and the expected significance for $B_c^+ \to J/\psi D^{*+}$ decays (right). Candidates are reconstructed from $D^0 \to K^-\pi^+$ (top) and $D^0 \to K^-\pi^+\pi^+\pi^-$ (bottom). The solid line represents the fitted background, and the dashed line indicates the extrapolation. The blinded signal region is shown in grey, and the expected signal is drawn in red.

decays is generated according to the physics model from Ref. [127], which is implemented in AmpGen as DtoKpipipi amplitude model. Correction weights to the phase-space simulation are generated by training a GBReweighter using five linearly independent and Lorentz invariant quantities, defined as $m_{K^-\pi_0^+}$, $m_{K^-\pi_1^+}$, $m_{K^-\pi^-}$, $m_{\pi_0^+\pi_1^+}$ and $m_{K^-\pi_0^+\pi_1^+}$, where the two pions of identical charge are denoted π_0^+ and π_1^+, respectively.

The correction weights are applied as per-event weights to the phase-space generated simulation based on the *generated* decay kinematics. The modification of the decay model is illustrated in Figure 12.5.

12.6.2 PID corrections

For muons, the PID efficiency is not corrected in this analysis, as any residual differences cancel in the ratio of the signal and normalisation decay. However, the PID response of hadrons is not as well modelled and does not cancel in the ratio. In the processing of simulated candidates, all PID requirements on the hadrons are removed and replaced by per-event PID efficiency weights generated with the PIDCalib2 package [100] analogously to Section 5.7.2. The PID efficiency weights for hadrons are generated in bins of `nTracks`, track momentum and pseudorapidity, separately for kaons and pions. The resulting per-event weight is constructed by multiplying the derived PID efficiency for each hadron track $w_{\text{PID}} = \prod_{h^\pm} w_{\text{PID}}(h^\pm)$. Note that due to the absence of PID requirements on the soft pion in $D^{*+} \to D^0\pi^+_{soft}$ decays, no correction weight is applied for π^+_{soft} tracks.

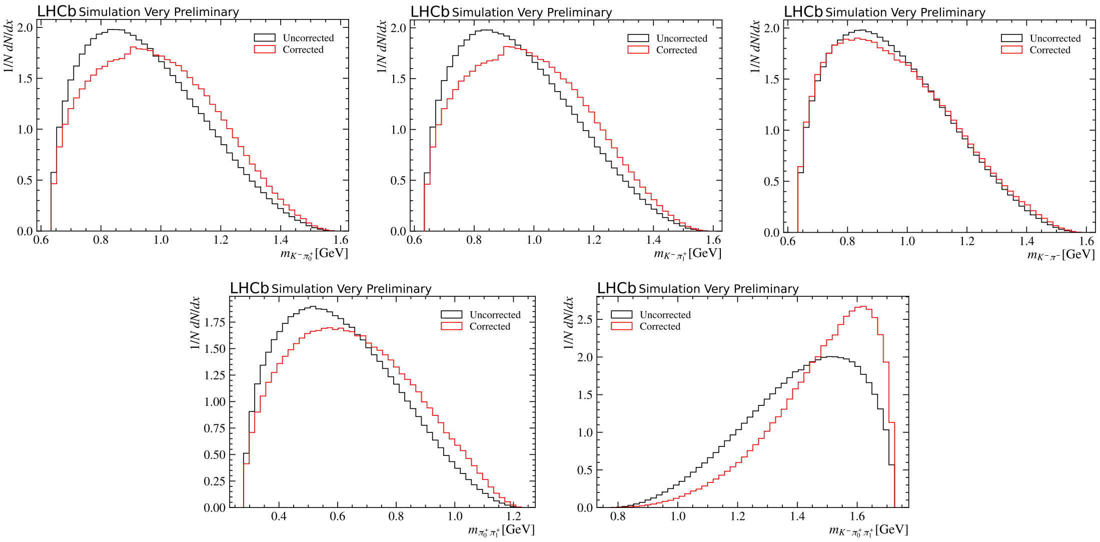


Figure 12.5: Simulated candidates of $D^0 \to K^-\pi^+\pi^+\pi^-$ decays generated with a phase-space decay model (black) and corrected to match the amplitude model from [128] in the invariant masses $m_{K^-\pi_0^+}$, $m_{K^-\pi_1^+}$, $m_{K^-\pi^-}$, $m_{\pi_0^+\pi_1^+}$ and $m_{K^-\pi_0^+\pi_1^+}$.

12.6.3 Tracking corrections

Corrections to the tracking efficiency in simulation are derived using standard correction tables provided in the TrackCalib package as a function of track transverse momentum, pseudorapidity and azimuthal angle as discussed in Section 5.7.3. The resulting per-event weight is constructed as the product of the correction weights for all final state tracks, e.g. $w_{\text{Trk}} = w_{\text{Trk}}(\mu^+) \times w_{\text{Trk}}(\mu^-) \times \prod_{h^\pm} w_{\text{Trk}}(h^\pm)$.

12.6.4 Kinematics and multiplicity

The generated B_c^+ kinematics and event multiplicity are corrected using a strategy similar to the one discussed in Section 5.7.4. A GBReweighter is trained on $B_c^+ \to J/\psi\pi^+$ candidates in simulation and background-subtracted data using the momentum, transverse momentum and pseudorapidity of the B_c^+ candidate, and the event multiplicity, `nTracks`, as training variables. The improved agreement of data and simulation following this correction is shown in Figure 12.6 for $B_c^+ \to J/\psi\pi^+$ candidates.

12.6.5 L0 trigger corrections

Corrections to the L0 trigger efficiency in simulation are derived from the high statistics control channel $B^+ \to K^+J/\psi(\to \mu^+\mu^-)$ using the *TISTOS* method as described in Section 5.7.5. Signal and normalisation candidates must either pass the `L0DiMuon` or the `L0Muon` trigger decisions. The efficiency of both trigger decisions is derived separately in data and simulation as a function of the product of the two muon transverse momenta, $p_{T,\mu\mu} = p_T(\mu^+) \times p_T(\mu^-)$. The resulting ratios of the `L0DiMuon` and `L0Muon` trigger efficiencies between data and simulation can be found in Appendix C.3. To account for the overlap of the two trigger decisions, the correction weight assigned to simulated

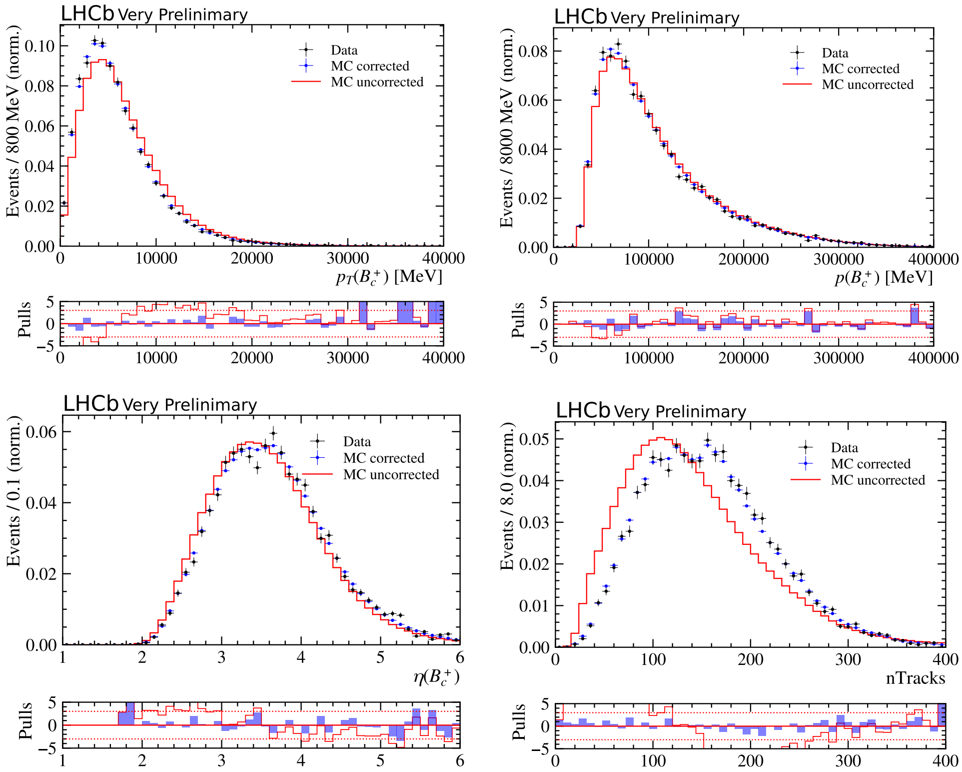


Figure 12.6: Comparison of background-subtracted data (black) and simulation of $B_c^+ \to J/\psi\pi^+$ candidates with (blue) and without (red) corrections to the B_c^+ kinematics and event multiplicity applied.

candidates is defined as

$$\begin{aligned} w_{\text{L0}} =& \texttt{L0DiMuon} \times w_{\texttt{L0DiMuon}}(p_{T,\mu\mu}) + \texttt{L0Muon} \times w_{\texttt{L0Muon}}(p_{T,\mu\mu}) \\ &- \texttt{L0DiMuon} \times \texttt{L0Muon} \times w_{\texttt{L0DiMuon}}(p_{T,\mu\mu}) \times w_{\texttt{L0Muon}}(p_{T,\mu\mu}) \end{aligned} \tag{12.2}$$

12.6.6 HLT trigger corrections

Corrections to the combined HLT1 and HLT2 trigger efficiency in simulation are derived using the same *TISTOS* method on data and simulation of $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$ decays. The trigger efficiencies in data and simulation are derived in bins of the event multiplicity, `nTracks`. From the ratio of efficiencies, the correction weight to simulation is derived as $w_{\text{HLT}} = \varepsilon_{\text{HLT}}^{Data}/\varepsilon_{\text{HLT}}^{MC}$. The derived correction weights can be found in Appendix C.3

12.7 Selection efficiency

The total efficiency for the selection and reconstruction of a signal or normalisation candidate is evaluated on the respective simulation samples. Due to a filter on the geometric detector acceptance applied in the event generation, the efficiency factorises into $\varepsilon_{tot} = \varepsilon_{geo} \times \varepsilon_{reco\&sel}$. All efficiencies are derived separately for each decay and each year of data-taking.

12.7.1 Geometric efficiency ε_{geo}

The simulation used in this analysis is produced, requiring all charged particles to be within the approximate detector acceptance of $\theta \in [10, 400]$ mrad, where θ is the polar angle with respect to the beam axis. The geometric efficiency is determined from the information provided with the centrally produced simulation samples, separately for each year of data taking. A table of the geometric efficiencies is provided in Appendix C.4.

12.7.2 Reconstruction and selection efficiency $\varepsilon_{reco\&sel}$

The combined reconstruction and selection efficiency, $\varepsilon_{reco\&sel}$, is determined from the respective simulation samples with all correction weights derived in Section 12.6 applied. As already discussed in the analysis of $B^+_{(c)} \to \mu^+\nu_\mu\gamma$ in Section 8.3, the weights to correct the initial kinematics of the B_c^+ and event multiplicity generated by a GBReweighter are not normalised. To simplify the calculation of the selection efficiency, these weights are first normalised according to Eq. 8.2. With the normalised weights, the reconstruction and selection efficiency is determined from the number of reconstructed candidates in simulation passing the event selection, N_{sel}, accounting for all simulation correction weights and the number of generated simulation events, N_{gen}

$$\varepsilon_{reco\&sel} = \frac{\sum_{i=1}^{N_{sel}} (w_{\text{PID}} \times w_{\text{Trk}} \times w_{\text{Mult\&Kin}} \times w_{\text{L0}} \times w_{\text{HLT}})}{N_{gen}}. \tag{12.3}$$

For the signal decay $B_c^+ \to J/\psi D^{*+}$ with $D^0 \to K^-\pi^+\pi^+\pi^-$, the weights to match the event simulation to the physical D^0 amplitude model, w_{D^0}, (see Section 12.6.1) need to be taken into account explicitly to ensure a proper normalisation of the efficiency

$$\varepsilon_{reco\&sel} = \frac{\sum_{i=1}^{N_{sel}} (w_{D^0} \times w_{\text{PID}} \times w_{\text{Trk}} \times w_{\text{Mult\&Kin}} \times w_{\text{L0}} \times w_{\text{HLT}})}{\sum_{i=1}^{N_{gen}} w_{D^0}}. \tag{12.4}$$

The resulting reconstruction and selection efficiencies for the signal and normalisation channels are listed in Appendix C.4.

12.7.3 Total efficiency ε_{tot}

The total efficiency is computed from the product of the geometric efficiency and the reconstruction and selection efficiency. As the number of signal candidates in data is determined from a single fit combining all years of data-taking, the combined efficiency across all years is calculated, accounting for the integrated luminosity recorded with the LHCb experiment per year of data-taking (see Eq. 8.5). The total efficiencies for the signal and normalisation channels are listed in Table 12.2.

Year	ε_{tot} $[10^{-3}]$		
	$B_c^+ \to J/\psi D^{*+}$ $D^0 \to K^-\pi^+$	$B_c^+ \to J/\psi D^{*+}$ $D^0 \to K^-\pi^+\pi^+\pi^-$	$B_c^+ \to J/\psi D_s^+$ $D_s^+ \to K^+K^-\pi^+$
2011	2.74 ± 0.02	0.79 ± 0.02	3.26 ± 0.02
2012	2.85 ± 0.03	0.82 ± 0.02	3.39 ± 0.02
2015	3.24 ± 0.03	0.93 ± 0.02	3.83 ± 0.02
2016	3.24 ± 0.03	0.93 ± 0.02	3.83 ± 0.02
2017	3.24 ± 0.03	0.93 ± 0.02	3.83 ± 0.02
2018	3.25 ± 0.03	0.93 ± 0.02	3.83 ± 0.02
Average	3.09 ± 0.03	0.89 ± 0.02	3.66 ± 0.02

Table 12.2: Total efficiencies for the signal and normalisation decays for each year of data-taking and averaged across all years, weighted by the respective recorded integrated luminosity.

13 Extraction of the branching fraction

The number of observed $B_c^+ \to J/\psi D^{*+}$ signal and $B_c^+ \to J/\psi D_s^+$ normalisation candidates is extracted from fitting the respective invariant mass distributions of reconstructed candidates in data passing the event selection. In both cases, an extended unbinned maximum likelihood fit is performed using the ZFIT python package. The strategy for the fit to the normalisation channel $B_c^+ \to J/\psi D_s^+$ is detailed in Section 13.1 and the fit to the signal channel $B_c^+ \to J/\psi D^{*+}$ is discussed in Section 13.2.

13.1 Fit to the normalisation channel $B_c^+ \to J/\psi D_s^+$

The number of $B_c^+ \to J/\psi D_s^+$ normalisation candidates is extracted from a fit to the $m(J/\psi D_s^+)$ distribution of normalisation candidates, where the J/ψ and D_s^+ candidates are constrained to the respective known masses. An extended unbinned maximum likelihood fit is performed, combining the data recorded during Run 1 and Run 2 into a single dataset. The PDF describing the invariant mass distribution is composed of a signal component derived from simulation of $B_c^+ \to J/\psi D_s^+$ candidates and a PDF to model the combinatorial background in data.

13.1.1 Signal model

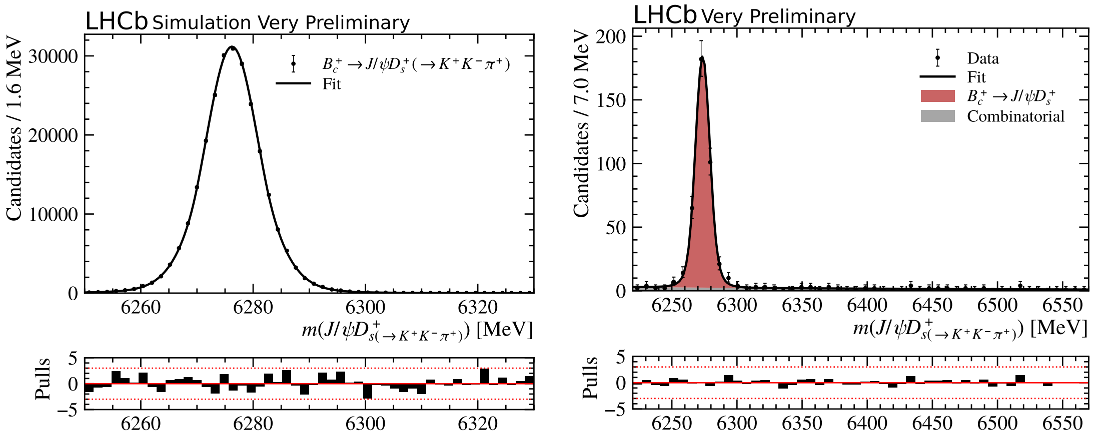
The invariant mass distribution of $B_c^+ \to J/\psi D_s^+$ candidates is modelled using the sum of a double-sided Crystal Ball function (defined in Eq. 3.5) and a Gaussian with shared mean. The model parameters are fixed from a fit to simulated candidates of $B_c^+ \to J/\psi D_s^+$ decays with the fit projection shown in Figure 13.1a.

13.1.2 Combinatorial background

Combinatorial background in data is modelled using an exponential PDF with a freely floating slope parameter λ_{Comb}.

13.1.3 Invariant mass fit

A single invariant mass fit is performed to extract the number of reconstructed candidates in the combined Run 1 and Run 2 datasets. For the signal shape, a common shift of the



(a) (b)

Figure 13.1: Fit to the reconstructed invariant mass of $B_c^+ \to J/\psi D_s^+$ candidates reconstructed from $D_s^+ \to K^+K^-\pi^+$ (a) in simulation and (b) in data combing the data recorded during Run 1 and 2.

mean position, $\Delta\mu$, and scaling of the width by $(1 + \Delta\sigma)$ are introduced, which are freely floating in the fit to data. The number of reconstructed normalisation candidates, $N_{D_s^+}$, and the number of background candidates are freely floating as well. The fit to the reconstructed invariant mass $m(J/\psi D_s^+)$ in data is shown in Figure 13.1b. The number of reconstructed candidates in the normalisation channel is found to be $N_{D_s^+} = 382 \pm 20$.

13.2 Fit to the signal channel $B_c^+ \to J/\psi D^{*+}$

The number of $B_c^+ \to J/\psi D^{*+}$ signal candidates passing the selection is extracted from a fit of the $m(J/\psi D^0\pi^+_{soft})$ distribution of reconstructed signal candidates in data, where the J/ψ and the D^0 candidates are constrained to the respective known masses. For the signal extraction, all years of data-taking are combined into a single dataset. The PDF to model the recorded mass distribution is composed of a signal PDF and a model for combinatorial background and is defined separately for decays reconstructed from $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ decays.

13.2.1 Signal model

Similar to the fit of the normalisation decay, the invariant mass distribution of signal candidates is modelled using the sum of a double-sided Crystal Ball function and a Gaussian with shared mean. The model parameters are determined from a fit to simulated $B_c^+ \to J/\psi D^{*+}$ candidates with $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$, respectively. The fit projections are shown in Figure 13.2.

13.2.2 Combinatorial background

Combinatorial background is modelled using an exponential PDF with a freely floating slope parameter, λ^f_{Comb}, separately for the two final states of the D^0 decay.

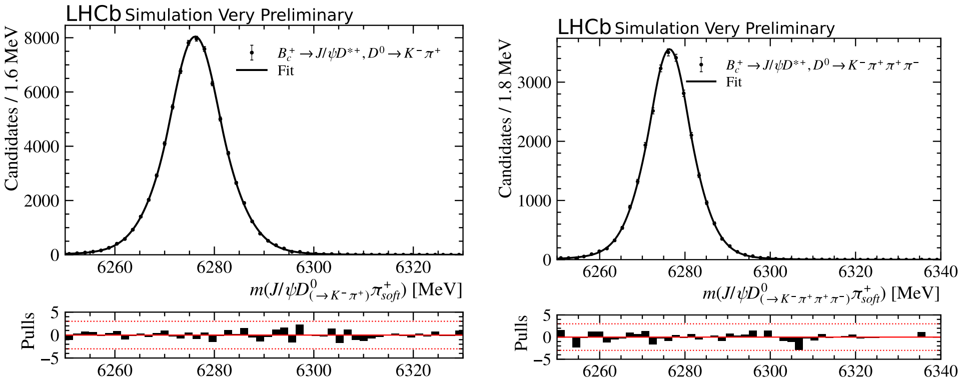


Figure 13.2: Invariant mass fit to simulated $B_c^+ \to J/\psi D^{*+}$ candidates reconstructed as $D^0 \to K^-\pi^+$ (left) and $D^0 \to K^-\pi^+\pi^+\pi^-$ (right).

13.2.3 Simultaneous fit setup

The fit to the reconstructed invariant mass in data is performed simultaneously for decays reconstructed from $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$. The signal shape parameters are fixed from a fit to the invariant mass of simulated candidates and a common shift of the mean position, $\Delta\mu$, and scaling of the width by $(1+\Delta\sigma)$ for both final states are introduced. The values of $\Delta\mu$ and $\Delta\sigma$ are Gaussian constrained to the values and uncertainties obtained in the invariant mass fit to the $B_c^+ \to J/\psi D_s^+$ normalisation channel, detailed in Section 13.1.

The simultaneous fit setup is used to directly extract the ratio of branching fractions from data, by expressing the number of reconstructed signal candidates, N_f, reconstructed in the two final states, $D^0 \to f$, with $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ as

$$N_f = \underbrace{\frac{\mathcal{B}(B_c^+ \to J/\psi D^{*+})}{\mathcal{B}(B_c^+ \to J/\psi D_s^+)}}_{\text{Fit parameter}} \times \underbrace{N_{D_s}}_{\text{Sec. 13.1}} \times \underbrace{\frac{\varepsilon_f^{MC}}{\varepsilon_{D_s}^{MC}}}_{\text{Sec. 12.7}} \times \underbrace{\frac{\mathcal{B}(D^{*+} \to D^0\pi^+_{soft}) \times \mathcal{B}(D^0 \to f)}{\mathcal{B}(D_s^+ \to K^+K^-\pi^+)}}_{\text{External input}} . \quad (13.1)$$

The yield of the normalisation channel, N_{D_s} is extracted from the fit to the reconstructed invariant mass in Section 13.1 and the ratio of efficiencies is derived from simulation in Section 12.7. The external inputs on the branching fractions for the $D^{(*)+}$ and D_s^+ decays are listed in Table 13.1. This leaves a single parameter, the ratio of branching fractions, to parametrise the signal strength in data reconstructed as $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$. Note, the nominal fit is performed using the central values of the parameters in Eq. 13.1. The systematic uncertainty on the ratio of branching fractions introduced by the uncertainties of the respective parameters in Eq. 13.1 are studied in Section 13.3.

13.2.4 Blinded invariant mass fit

As the analysis is still blinded, the fit to the invariant mass of candidates reconstructed in data is performed using the nominal fit setup, but the data in the signal region $m_{B_c^+} \pm 20\,\text{MeV}$ is not shown and the fitted ratio of branching fractions remains blinded. The blinded fit projections of the simultaneous invariant mass fit to data in the two different final states are shown in Figure 13.3.

Decay	$\mathcal{B}$ [10^{-2}] [5]
$D^{*+} \to D^0\pi^+_{soft}$	67.7 ± 0.5
$D^0 \to K^-\pi^+$	3.945 ± 0.030
$D^0 \to K^-\pi^+\pi^+\pi^-$	8.22 ± 0.14
$D_s^+ \to K^+K^-\pi^+$	5.54 ± 0.08

Table 13.1: External inputs to extract the ratio of branching fractions for $\frac{\mathcal{B}(B_c^+ \to J/\psi D^{*+})}{\mathcal{B}(B_c^+ \to J/\psi D_s^+)}$. Values are taken from the PDG [5].

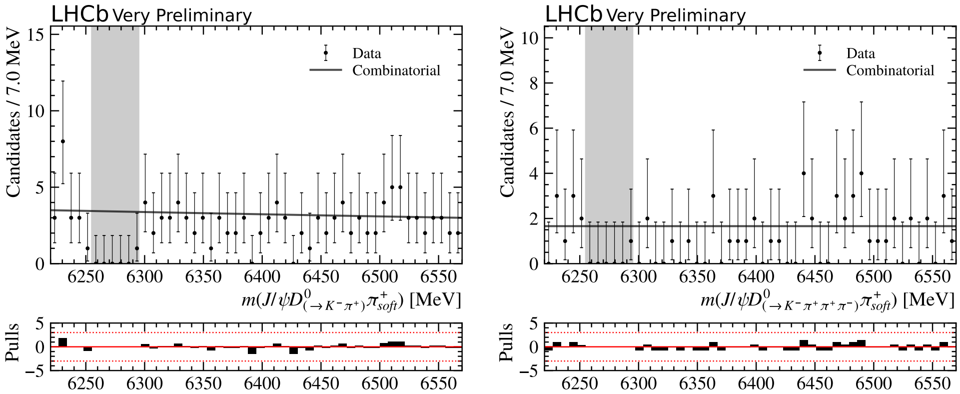


Figure 13.3: Blinded fit to the invariant mass of $B_c^+ \to J/\psi D^{*+}$ candidates reconstructed from $D^0 \to K^-\pi^+$ (left) and $D^0 \to K^-\pi^+\pi^+\pi^-$ (right). The grey region is the blinded signal region around the B_c^+ mass.

13.2.5 Validation of the simultaneous fit

The simultaneous fit is validated by performing a series of pseudo-experiments. Pseudo-data are generated from the derived signal and background PDFs according to the number of expected signal candidates following the predicted ratio of branching fractions in Eq. 11.2 and the number of background candidates derived from the blinded fit to data. The generated pseudo-data are fitted using the nominal fit strategy explained in the previous paragraphs in a total of 300 pseudo-experiments. The fit projections of one such pseudo-experiment are shown in Figure 13.4. The results on the fitted branching fractions across the ensemble of pseudo-experiments are evaluated in terms of the pull distribution, $\frac{\mu_{BR}-\mu_0}{\sigma_{BR}}$, where $\mu_{BR} \pm \sigma_{BR}$ is the fitted ratio of branching fractions with the respective uncertainty and μ_0 is the value used in the generation of the pseudo-data. The pull distribution and the uncertainty on the ratio of branching fractions, σ_{BR}, are shown in Figure 13.5. The pull distribution follows a Gaussian profile centred at -0.09 ± 0.04, which indicates a small negative bias, approximately an order of magnitude smaller than the expected statistical uncertainty. Nonetheless, a systematic uncertainty is added to cover the observed negative bias. The width of the pull distribution is found to be consistent with unity, indicating correct coverage of the fit. From the distribution of uncertainties, σ_{BR}, the expected statistical uncertainty on the ratio of branching fractions is taken as ± 0.028.

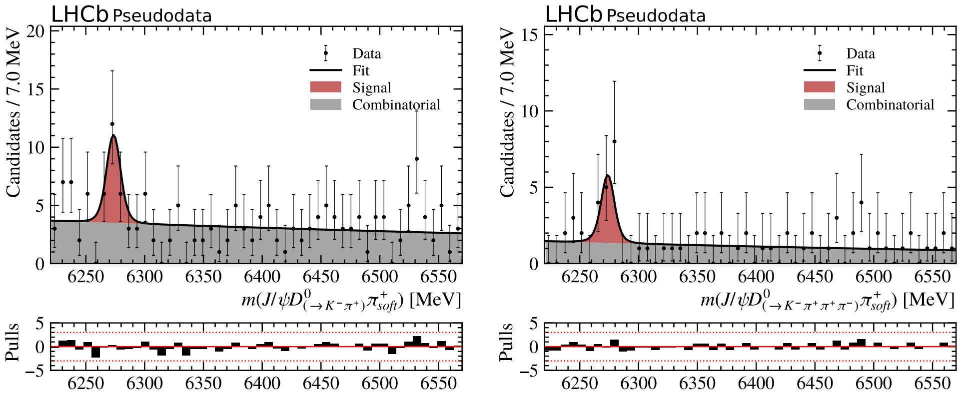


Figure 13.4: Fit projections of a simultaneous fit to $B_c^+ \to J/\psi D^{*+}$ pseudo-data reconstructed as $D^0 \to K^-\pi^+$ (left) and $D^0 \to K^-\pi^+\pi^+\pi^-$ (right).

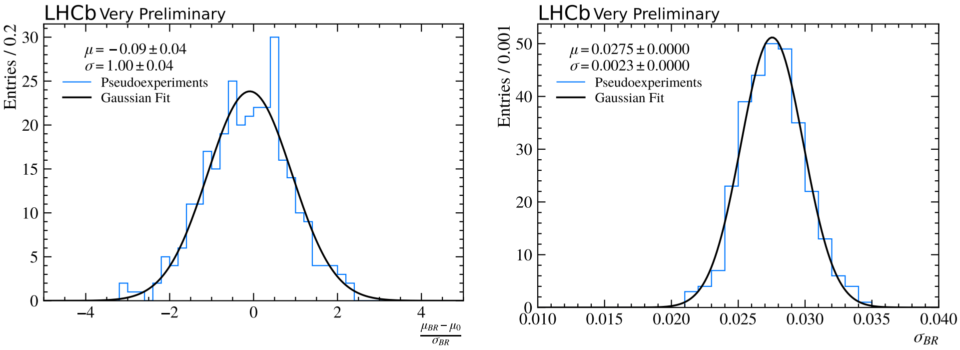


Figure 13.5: Distribution of pulls (left) and fit uncertainties (right) on the ratio of branching fractions $\frac{\mathcal{B}(B_c^+ \to J/\psi D^{*+})}{\mathcal{B}(B_c^+ \to J/\psi D_s^+)}$ for 300 pseudo-experiments.

Systematic uncertainty	$\Delta\frac{\mathcal{B}(B_c^+\to J/\psi D^{*+})}{\mathcal{B}(B_c^+\to J/\psi D_s^+)}[10^{-3}]$
Finite simulation samples	3.1
Tracking efficiencies	1.3
PID efficiency	0.2
Trigger efficiency	1.6
Signal fit bias	1.8
Choice of fit model	1.0
Normalisation yield	5.1
External uncertainties	2.6
Total systematic uncertainty	7.1

Table 13.2: Summary of the systematic uncertainties evaluated for a central value on the ratio of branching fractions of 0.098 as derived in Eq. 11.2.

13.3 Systematic uncertainties

In this section, the systematic uncertainties on the measured ratio of branching fractions $\mathcal{B}(B_c^+ \to J/\psi D^{*+})/\mathcal{B}(B_c^+ \to J/\psi D_s^+)$ are evaluated. Although the analysis is expected to be dominated by the statistical uncertainty of the signal channel $B_c^+ \to J/\psi D^{*+}$, a thorough evaluation of all sources of systematic uncertainties is performed. A summary of the systematic uncertainties considered in this analysis is provided in Table 13.2.

Finite simulation samples

The finite size of the simulation samples for the signal and normalisation decays introduces a systematic uncertainty on the total reconstruction and selection efficiency. The uncertainties linked to the finite size of the simulation samples averaged across all years of data-taking are taken from Table 12.2. The impact on the ratio of branching fractions is evaluated by conducting a series of pseudo-experiments in which the pseudo-data are generated by varying the efficiencies within the respective uncertainties. From the difference of the central value, $\Delta\mu$, and the difference in the spread around the central value, $\Delta\sigma$, with respect to the study using the central values in Section 13.2.5, the systematic uncertainty is evaluated as

$$\sigma_{syst} = \sqrt{\Delta\mu^2 + \Delta\sigma^2}. \tag{13.2}$$

From 300 pseudo-experiments the values $\Delta\mu = 8.5 \times 10^{-4}$ and $\Delta\sigma = 3.0 \times 10^{-3}$ are derived. This yields a systematic uncertainty of $\pm 3.1 \times 10^{-3}$ on the ratio of branching fractions

Tracking efficiencies

The tracking efficiency tables provided in the collaboration-wide TRACKCALIB package and used to correct the tracking efficiency in simulation are subject to a systematic uncertainty. For signal candidates reconstructed as $D^{*+} \to D^0\pi^+_{soft}$ and $D^0 \to K^-\pi^+$, this cancels in the ratio with the normalisation channel due to the same number of charged tracks in the final state. However, signal candidates reconstructed as $D^0 \to K^-\pi^+\pi^+\pi^-$

have two additional charged tracks, for which the systematic uncertainty of the tracking efficiency does not cancel. Hence, the standard uncertainties per track of 0.4% in Run 1 and 0.8% in Run 2 are assigned for the two additional tracks. Averaging across all years of data-taking and accounting for the respective recorded luminosity, this yields a total systematic uncertainty of 1.3%. Although there is no systematic uncertainty linked to the tracking efficiency in the $D^0 \to K^-\pi^+$ signal channel, the uncertainty of 1.3% is applied conservatively on the ratio of branching fractions. Using the predicted central value from 11.2, this corresponds to an uncertainty of 1.3×10^{-3} on the ratio of branching fractions.

PID efficiencies

The finite size of the calibration samples used to correct the hadron PID efficiency introduces a systematic uncertainty on the selection efficiency of signal and normalisation candidates. This results in a 0.1% relative uncertainty on the selection efficiency of signal and normalisation candidates, respectively, for which a conservative relative systematic uncertainty of 0.2% is assigned. This relates to an absolute uncertainty of 0.2×10^{-3} on the ratio of branching fractions.

Trigger efficiencies

Similarly, the uncertainties on the derived L0 and HLT trigger efficiencies are propagated to yield the uncertainty on the derived selection efficiency. The relative uncertainty is evaluated to be 1.1% for both the signal and the normalisation channels. Assuming the extreme case of uncorrelated uncertainties between the signal and normalisation channels, this leads to a relative systematic uncertainty of 1.6%. With the predicted central value from Eq. 11.2, this corresponds to an uncertainty of 1.6×10^{-3} on the ratio of branching fractions.

Signal fit bias

In Section 13.2.5, a slight negative bias on the extracted ratio of branching fractions is found. A systematic uncertainty is applied to the ratio of branching fraction, which accounts for the observed bias of -1.8×10^{-3}.

Choice of fit model

The systematic uncertainty due to the choice of fit model is evaluated by performing pseudo-experiments with the pseudo-data generated from an alternative model and fitted with the nominal model. The alternative model is chosen as a generalised Crystal Ball function for the signal and a second-order polynomial for the background. The definition of the generalised Crystal Ball function and the blinded fit to the data using the alternative fit model is provided in Appendix C.5. The systematic uncertainty is evaluated on 300 pseudo-experiments, which yield a shift of the central value of $\Delta\mu = 6 \times 10^{-4}$ and a difference in spread of $\Delta\sigma = 7 \times 10^{-4}$. Combining the effects using Eq. 13.2, this yields a total systematic uncertainty of 1.0×10^{-3} on the ratio of branching fractions.
A similar alternative fit model is chosen to perform the fit to the normalisation channel

$B_c^+ \to J/\psi D_s^+$. No change in the derived number of normalisation candidates is found and no additional systematic uncertainty is assigned.

Normalisation yield

The uncertainty on the extracted number of $B_c^+ \to J/\psi D_s^+$ normalisation candidates introduces a systematic uncertainty on the ratio of branching fractions. The size of the systematic uncertainty is taken as the relative uncertainty on the normalisation yield and evaluated to be 5.2%, based on the results from Section 13.1. This relates to an absolute uncertainty on the ratio of branching fractions of 5.1×10^{-3}.

External uncertainties

The external inputs on the branching fractions for the decay of the intermediate D^{*+}, D^0 and D_s^+ resonances listed in Table 13.1 constitute external sources of systematic uncertainties. The impact on the ratio of branching fractions is evaluated by conducting a series of 300 pseudo-experiments in which the pseudo-data are generated by varying the external branching fractions within the respective uncertainties. The systematic uncertainty is calculated from Eq. 13.2 using the obtained values $\Delta\mu = 2.4 \times 10^{-3}$ and $\Delta\sigma = 1.2 \times 10^{-4}$. This yields a total systematic uncertainty of 2.6×10^{-3} on the ratio of branching fractions.

14

Results and concluding remarks

The presented analysis marks the first search for $B_c^+ \to J/\psi D^{*+}$ decays using the full Run 1 and Run 2 datasets recorded with the LHCb experiment. Signal candidates are reconstructed from $D^{*+} \to D^0\pi^+_{soft}$ with $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ decays. The signal selection is established using a combination of one-dimensional cuts and multivariate analysis, maximising the expected significance of observing $B_c^+ \to J/\psi D^{*+}$ decays.
The branching fraction of the signal decay is measured relative to the branching fraction of the normalisation decay $B_c^+ \to J/\psi D_s^+$, reconstructed from $D_s^+ \to K^+K^-\pi^+$ decays. From an invariant mass fit to the reconstructed invariant mass, a total of 382 ± 20 normalisation candidates are found.
Using a simultaneous fit to the reconstructed invariant mass of signal candidates reconstructed from $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$ decays, the ratio of branching fractions is extracted from data. As the central value of the ratio of branching fractions is currently blinded, the result is reported as

$$\frac{\mathcal{B}(B_c^+ \to J/\psi D^{*+})}{\mathcal{B}(B_c^+ \to J/\psi D_s^+)} = x.xxx \pm 0.026\,(\text{stat}) \pm 0.007\,(\text{syst}), \tag{14.1}$$

where the first uncertainty marks the expected statistical uncertainty and the second marks the systematic uncertainty. Assuming the central value of the unblinded result on $\mathcal{B}(B_c^+ \to J/\psi D^{*+})/\mathcal{B}(B_c^+ \to J/\psi D_s^+)$ is of the order of the predicted value 0.098 ± 0.011, the presented analysis is expected to lead to evidence for the decay $B_c^+ \to J/\psi D^{*+}$ with a significance above three standard deviations. Currently, the analysis is being finalised and prepared for the internal review process before unblinding the central value on the ratio of branching fractions.

15

Conclusion

The precision of theoretical predictions for B meson decays is often limited by the knowledge of low-energy QCD contributions, which cannot be calculated perturbatively. In QCD factorisation approaches, the calculations of the high-energy and low-energy contributions are separated to reduce the uncertainties in the theoretical treatment of B meson decays. One crucial ingredient for the application of QCD factorisation is the knowledge of the leading-order B meson light-cone distribution amplitude and, in particular, its first inverse moment, λ_B. Inferring the parameter λ_B from theoretical calculations is very challenging and results in currently large uncertainties.

Radiative leptonic decays of $B^+ \to \ell^+ \nu_\ell \gamma$ are considered a golden mode to probe the B meson LCDA and measure λ_B experimentally. Such decays have never been observed with the current best limit on their branching fraction reported by the Belle collaboration at $\mathcal{B}(B^+ \to \ell^+ \nu_\ell \gamma) < 3.0 \times 10^{-6}$ [1]. This work presents the first search for radiative leptonic decays $B^+ \to \mu^+ \nu_\mu \gamma$ at a hadron collider and the first ever search for $B_c^+ \to \mu^+ \nu_\mu \gamma$ decays. The analysis is performed on data recorded with the LHCb experiment from 2016 to 2018, corresponding to an integrated luminosity of $5.4\,\mathrm{fb}^{-1}$. Signal candidates are reconstructed using photon conversions in the material of the VELO to allow for the displaced B vertex to be resolved and a standard event selection to be performed. The number of observed $B^+ \to \mu^+ \nu_\mu \gamma$ and $B_c^+ \to \mu^+ \nu_\mu \gamma$ signal decays is extracted from a fit to the corrected mass distribution in data. For the modelling of background from $\pi^0 \to \gamma_{ee} \gamma$ and $\eta \to \gamma_{ee} \gamma$ decays, a new powerful data-driven method is developed using additional inputs from simulation. From the observed number of signal candidates, the respective branching fractions are derived relative to the normalisation channel $B^0 \to K^{*0} \gamma$, which is also reconstructed from photon conversions to cancel systematic uncertainties. As the analysis is currently under review in the collaboration, the central values of the signal branching fractions remain blinded.

Assuming a branching fraction for the decay $B^+ \to \mu^+ \nu_\mu \gamma$ of 3.0×10^{-6}, corresponding to the current best upper limit from the Belle collaboration, and a branching fraction of 8×10^{-5} for the decay $B_c^+ \to \mu^+ \nu_\mu \gamma$ according to the prediction from [2], the blinded results are:

$$\begin{aligned}
\mathcal{B}(B^+ \to \mu^+ \nu_\mu \gamma)_{E^*_\gamma > 1\,\mathrm{GeV}} &= (x.x \pm {}^{+1.5}_{-1.3}\,(\mathrm{stat.}) \pm 1.0\,(\mathrm{syst.})) \times 10^{-6}, \\
\mathcal{B}(B_c^+ \to \mu^+ \nu_\mu \gamma)_{E^*_\gamma > 1\,\mathrm{GeV}} &= (x.x \pm {}^{+2.9}_{-2.2}\,(\mathrm{stat.}) \pm 1.4\,(\mathrm{syst.})) \times 10^{-4}.
\end{aligned} \tag{15.1}$$

The first uncertainty corresponds to the statistical uncertainty and the second uncertainty marks the systematic uncertainty. After unblinding of the central values, the systematic uncertainties will be recomputed using the observed number of signal candidates, which is likely to reduce the leading systematic uncertainty linked to the corrected mass model for $B^+ \to \mu^+\nu_\mu\gamma$ decays.

In the case that no significant signal is observed, the upper limits on the branching fractions are provided. They are evaluated from the CL_s method on pseudo-experiments. The expected upper limits in the absence of a true signal, including all statistical and systematic uncertainties, at 90% (95%) confidence level are

$$\begin{aligned} \mathcal{B}_{exp}(B^+ \to \mu^+\nu_\mu\gamma)_{E^*_\gamma > 1\,\mathrm{GeV}} &< 2.3\,(2.7) \times 10^{-6} \\ \mathcal{B}_{exp}(B_c^+ \to \mu^+\nu_\mu\gamma)_{E^*_\gamma > 1\,\mathrm{GeV}} &< 6.3\,(7.8) \times 10^{-4}. \end{aligned} \tag{15.2}$$

The expected sensitivity on the branching fraction for $B^+ \to \mu^+\nu_\mu\gamma$ decays is very close to the one reported by the Belle collaboration, possibly leading to an improvement of the upper limit on the branching fraction. For the decay $B_c^+ \to \mu^+\nu_\mu\gamma$, the expected sensitivity is an order of magnitude looser than most Standard Model predictions and no observation is expected. Instead, the first upper limit on the branching fraction will be provided. Despite the search for $B_{(c)}^+ \to \mu^+\nu_\mu\gamma$ decays being deemed impossible at hadron colliders, the presented work demonstrates that the analysis is feasible, pushing the limitations of the LHCb experiment and extending its physics program to areas the LHCb experiment was originally not designed for.

As the presented results are dominated by the statistical uncertainty, an update of the analysis using data recorded with Run 3 of the LHCb experiment is proposed. With the significant increase in statistics, the decay $B^+ \to \ell^+\nu_\ell\gamma$ can potentially be observed in the future and the limit on the branching fraction of $B_c^+ \to \ell^+\nu_\ell\gamma$ can be significantly improved.

In addition, the first search for $B_c^+ \to J/\psi D^{*+}$ decays is presented using the full Run 1 and Run 2 datasets recorded with the LHCb experiment, corresponding to an integrated luminosity of $9\,\mathrm{fb}^{-1}$. This decay can serve as a prime normalisation channel for future searches of $b \to d\ell\ell$ transition in decays of B_c^+ mesons. The decay is reconstructed from $D^{*0} \to D^0\pi^+_{soft}$ with $D^0 \to K^-\pi^+$ and $D^0 \to K^-\pi^+\pi^+\pi^-$. The number of observed signal candidates is extracted from a simultaneous fit to the reconstructed invariant mass in the two final states. Using the normalisation channel $B_c^+ \to J/\psi D_s^+$, the ratio of branching fractions is extracted as

$$\frac{\mathcal{B}(B_c^+ \to J/\psi D^{*+})}{\mathcal{B}(B_c^+ \to J/\psi D_s^+)} = x.xxx \pm 0.026\,(\mathrm{stat.}) \pm 0.007\,(\mathrm{syst.}), \tag{15.3}$$

where the first uncertainty is the statistical uncertainty and the second one is the systematic uncertainty. The central value remains blinded as the analysis is currently being prepared for a collaboration-wide review. With an expectation derived from first principles arguments of $\mathcal{B}(B_c^+ \to J/\psi D^{*+})/\mathcal{B}(B_c^+ \to J/\psi D_s^+) = 0.098 \pm 0.011$, the presented analysis is expected to find evidence for $B_c^+ \to J/\psi D^{*+}$ decays with a significance exceeding three standard deviations. Using the increased statistics recorded with the LHCb experiment in Run 3, a first observation is within reach in future analyses of this decay.

In light of the future upgrades of the LHCb detector, the author has contributed to the upgrade of LHCb's electromagnetic calorimeter. In the search for $B_{(c)}^{+} \to \mu^{+}\nu_{\mu}\gamma$, it is shown that the reconstruction of photons from neutral clusters in the ECAL is of paramount importance. To ensure the LHCb experiment can fully exploit its potential in the future, the ECAL is upgraded during LS3 with finer granularity and again in Upgrade II to enable operation in the high pile-up conditions expected for Run 5.
In this work, the mapping of the signal cables that connect the PMTs to the front-end boards and the distribution of the front-end boards across the electronics racks are defined. The signal cables for the LS3 enhancement are currently being produced and will be installed during LS3.
Furthermore, studies on the performance of the PID system for Run 4 have been conducted if the HCAL is not replaced during LS3. In particular, the identification of electrons and muons, as well as the respective misidentification of hadrons, is studied with and without using PID information from the HCAL. No significant change in PID performance is found for muons. For electrons, a slight degradation of the PID performance is found for energetic electrons without associated bremsstrahlung photon and tight PID requirements. The overall effect on the selection efficiency is studied using $B^{+} \to K^{+}J/\psi(\to e^{+}e^{-})$ decays as a benchmark. Assuming the same level of background from misidentification, the loss in selection efficiency is approximately 1%. With the planned LS3 enhancements to the RICH system and ECAL, the total PID performance in Run 4 is likely to surpass that in Run 3, even without utilising PID information from the HCAL.

A

Supplementary material on the upgrade of the calorimeter

A.1 Calorimeter signal cable mapping

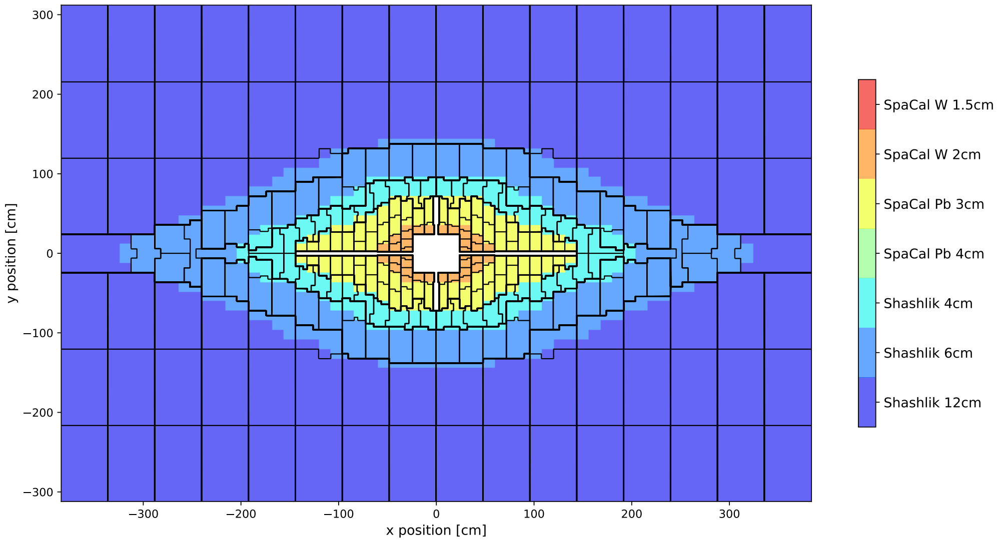


Figure A.1: Mapping scheme of the signal cables for the LS3 enhancements. Cells in the regions enclosed by a black line are mapped to the same FEB. The colour indicates the cell size and module technology.

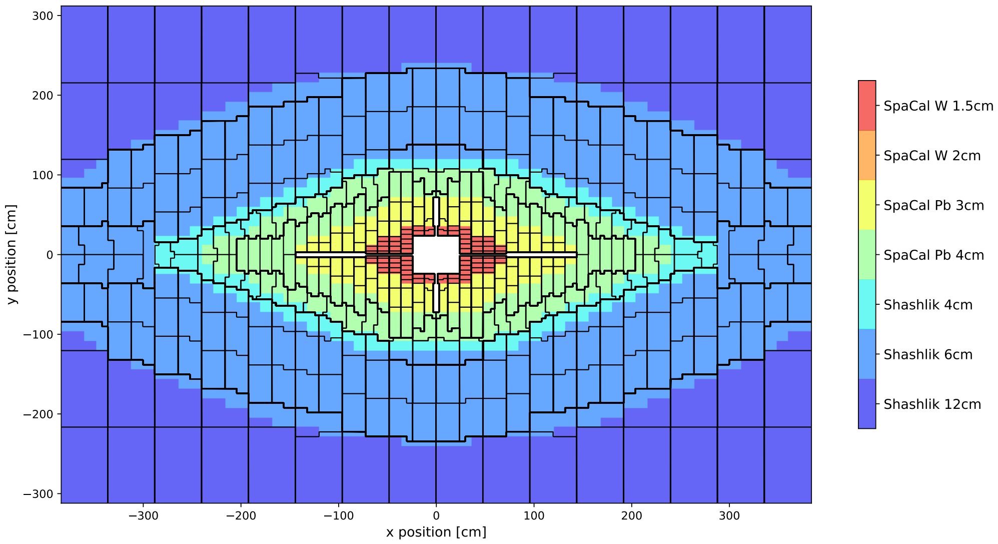


Figure A.2: Mapping scheme of the signal cables for Upgrade II (bottom). Cells in the regions enclosed by a black line are mapped to the same FEB. The colour indicates the cell size and module technology.

A.2 Charged PID performance with and without HCAL

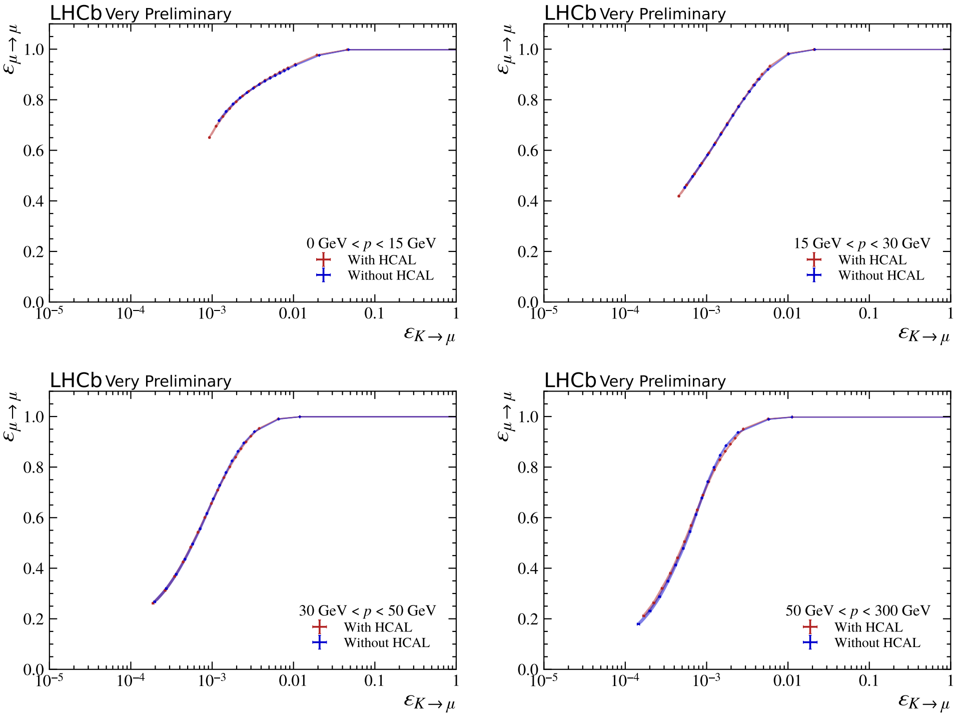


Figure A.3: Performance of the PID system with HCAL (red) and without HCAL (blue) for muon identification and $K^+ \to \mu^+$ misidentification in four bins of momentum.

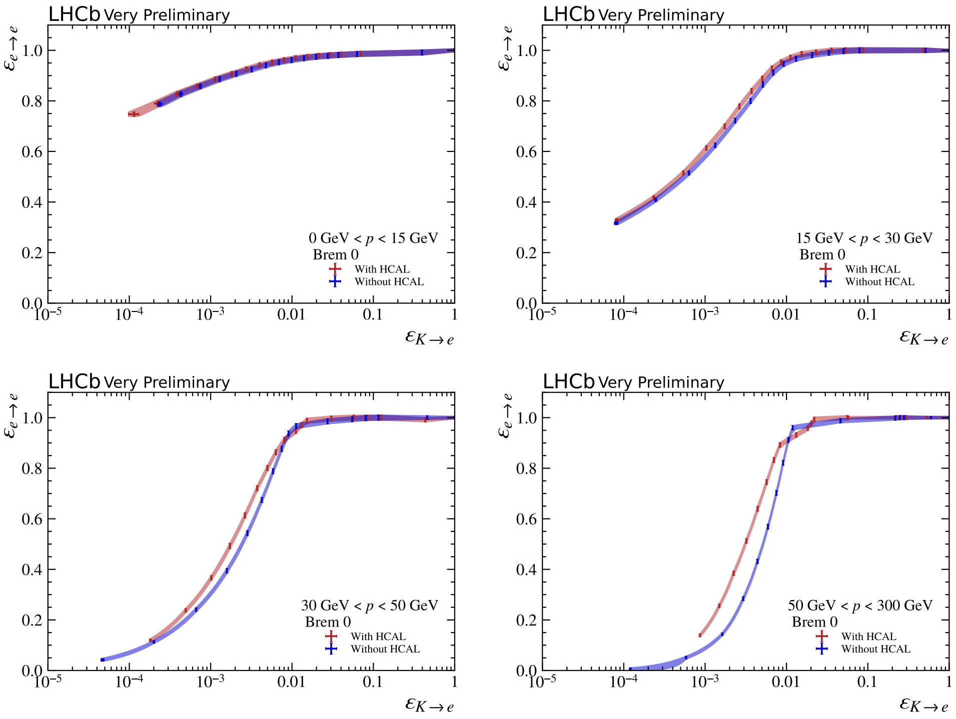


Figure A.4: Performance of the PID system with HCAL (red) and without HCAL (blue) for electron identification and $K^+ \to e^+$ misidentification in four bins of momentum and without a bremsstrahlung photon attached to the probe track.

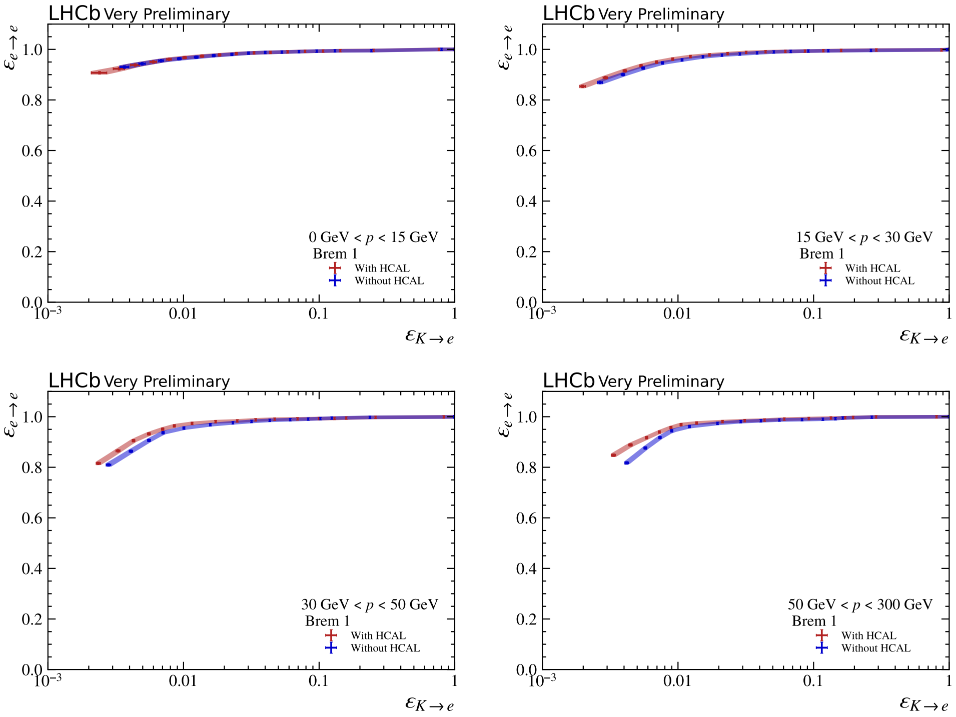


Figure A.5: Performance of the PID system with HCAL (red) and without HCAL (blue) for electron identification and $K^+ \to e^+$ misidentification in four bins of momentum and with bremsstrahlung photon attached to the probe track.

B Supplementary material on the search for $B_{(c)}^{+} \to \mu^{+}\nu_{\mu}\gamma$ decays

B.1 Approximating the $\boldsymbol{B}$ rest frame

The photon energy is calculated in the B rest frame, which is approximated using the visible decay products as introduced in [92]. The proper velocity, $\beta\gamma$, of the visible decay products (denoted *vis*) along the z axis is set equal to that of the B meson along the same axis. This results in a simple expression for the approximate B momentum along the z axis

$$(p_B)_z = \frac{m_B}{m_{vis}}(p_{vis})_z. \tag{B.1}$$

Using the B flight direction (which can be determined from the position of the PV and the B decay vertex), the remaining components are scaled by the same fraction to yield

$$|p_B| = \frac{m_B}{m_{vis}}(p_{vis})_z\sqrt{1+\tan^2\theta} \tag{B.2}$$

with θ being the angle between the B flight direction and the z axis.

B.2 Stripping requirements for the normalisation channel $B^0 \to K^{*0}\gamma$

Candidate	Selection
all tracks	`isLong` $\chi^2/\text{ndof} < 3$ GhostProb < 0.35
K^+	$\chi^2_{IP} > 9$ $p_T > 400\,\text{MeV}$ $\text{DLL}_{\text{K}\pi} > -5$
π^-	$\chi^2_{IP} > 9$ $p_T > 400\,\text{MeV}$
K^{*0}	$\lvert m - m_{PDG}\rvert < 300\,\text{MeV}$ $p_T > 500\,\text{MeV}$ $\chi^2_{FD} < 25$
$e^\pm$	$\chi^2_{IP} > 25$ $p_T > 500\,\text{MeV}$ $\text{DLL}_{\text{e}\pi} > 3$ (data only) VELO ADC > 50
e^+e^-	$m < 12\,\text{MeV}$ $p_T > 500\,\text{MeV}$ $\chi^2_{\text{Vertex}} < 4$ $\chi^2_{FD} > 16$
B^0	$\lvert m - m_{\text{PDG}}\rvert < 5000\,\text{MeV}$ `DIRA` > 0.9995 $\chi^2_{IP} < 25$ $\chi^2_{FD} > 100$ $\chi^2_{\text{Vertex}}/\text{ndof} < 9$

Table B.1: Stripping requirements for the selection of $B^0 \to K^{*0}\gamma$ candidates with the modifications for the processing of simulation data indicated.

C

Supplementary material on the search for $B_c^+ \to J/\psi D^{*+}$ decays

C.1 Requirements of the central event filtering

Candidate	**Selection**
$\mu^{\pm}$	track $\chi^2/\text{ndof} < 3$ `isMuon` $\text{DLL}_{\mu\pi} > 0$ $p_T > 500\,\text{MeV}$ $\chi^2_{\text{IP}} > 4$
$\mu^+\mu^-$	$m(\mu^+\mu^-) \in m_{J/\psi}^{\text{PDG}} \pm 100\,\text{MeV}$ DOCA $\chi^2 < 30$ $\chi^2_{Vertex}/\text{ndof} < 20$ Vertex separation from PV $> 3\sigma$

Table C.1: Stripping requirements for the selection of $B_c^+ \to J/\psi(\to \mu^+\mu^-)X$ candidates.

C.2 BDT training features

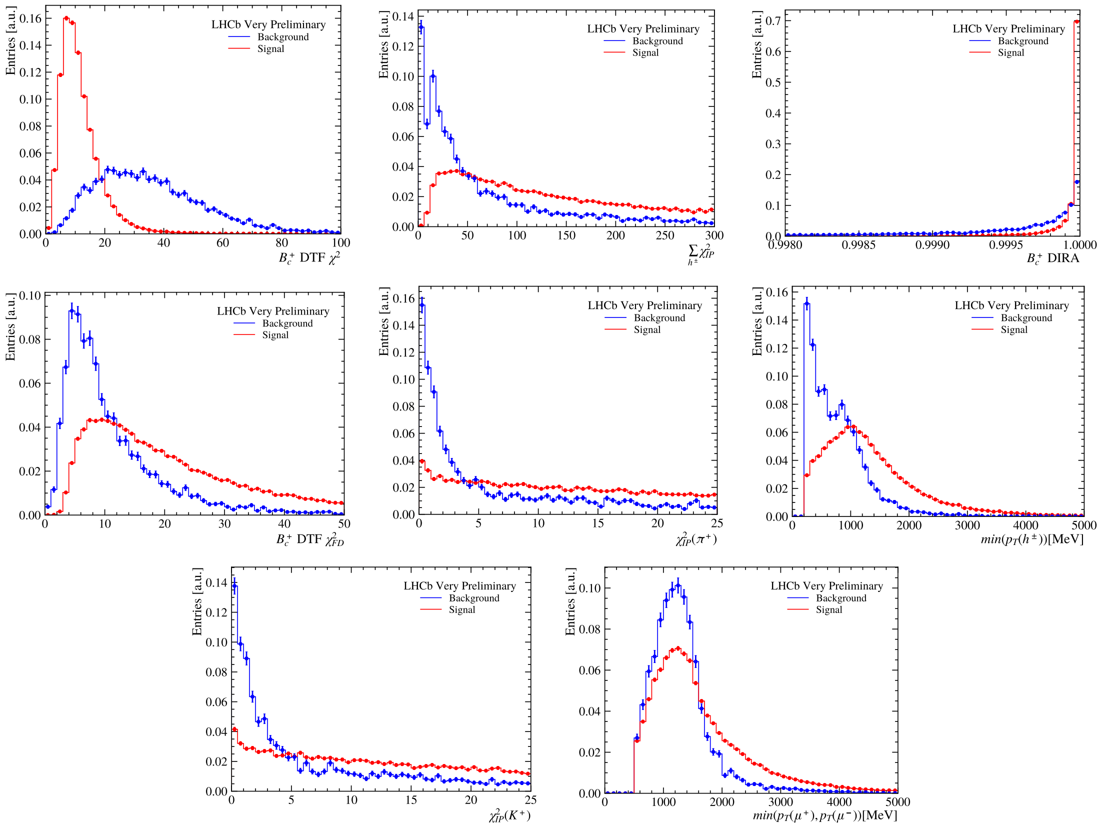


Figure C.1: Distributions of the signal proxy (red) and background proxy (blue) in the decay variables used in the training of the BDT for the $B_c^+ \to J/\psi D^{*+}$ signal decay reconstructed from $D^0 \to K^-\pi^+$.

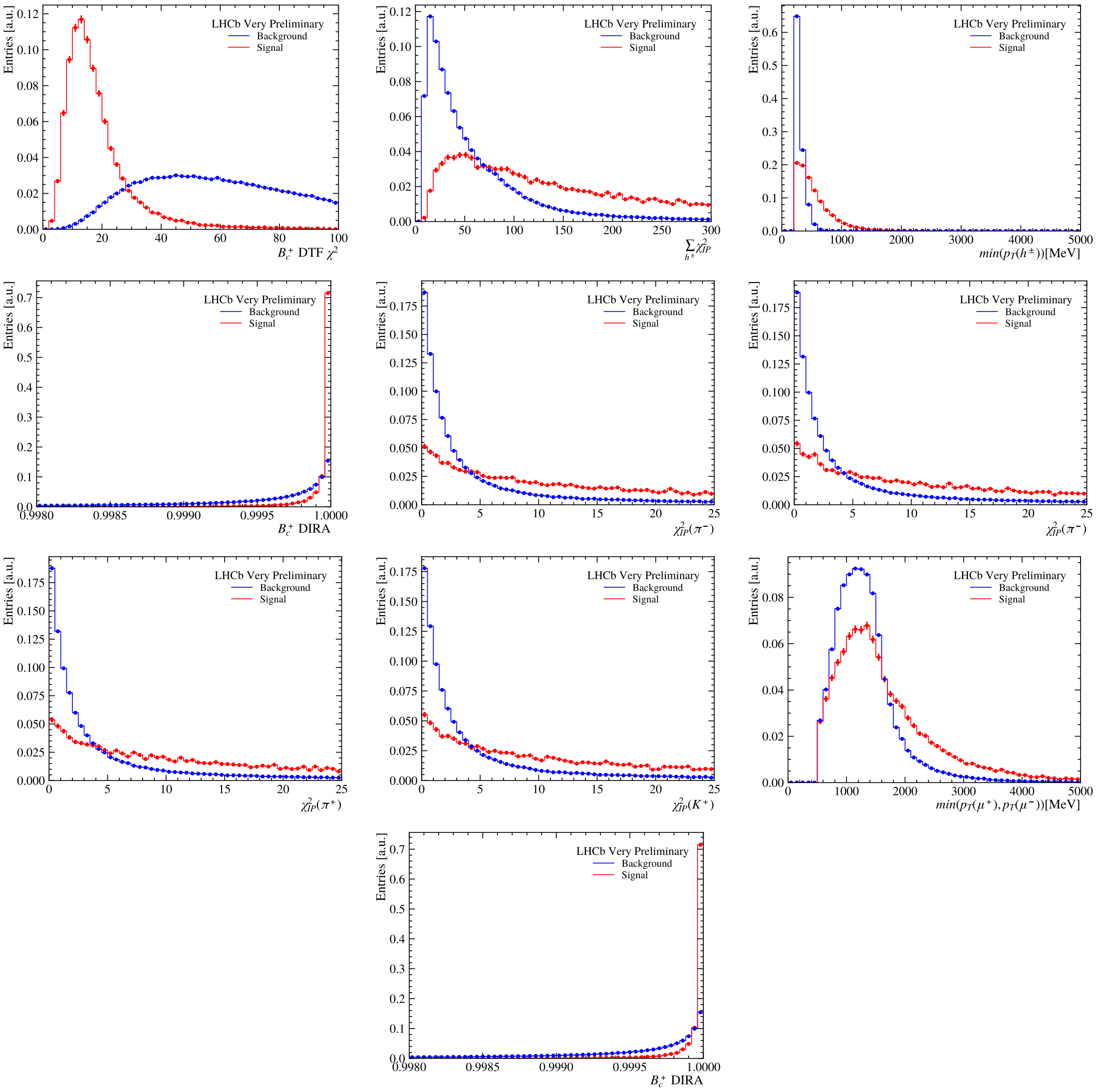


Figure C.2: Distributions of the signal proxy (red) and background proxy (blue) in the decay variables used in the training of the BDT for the $B_c^+ \to J/\psi D^{*+}$ signal decay reconstructed from $D^0 \to K^-\pi^+\pi^+\pi^-$.

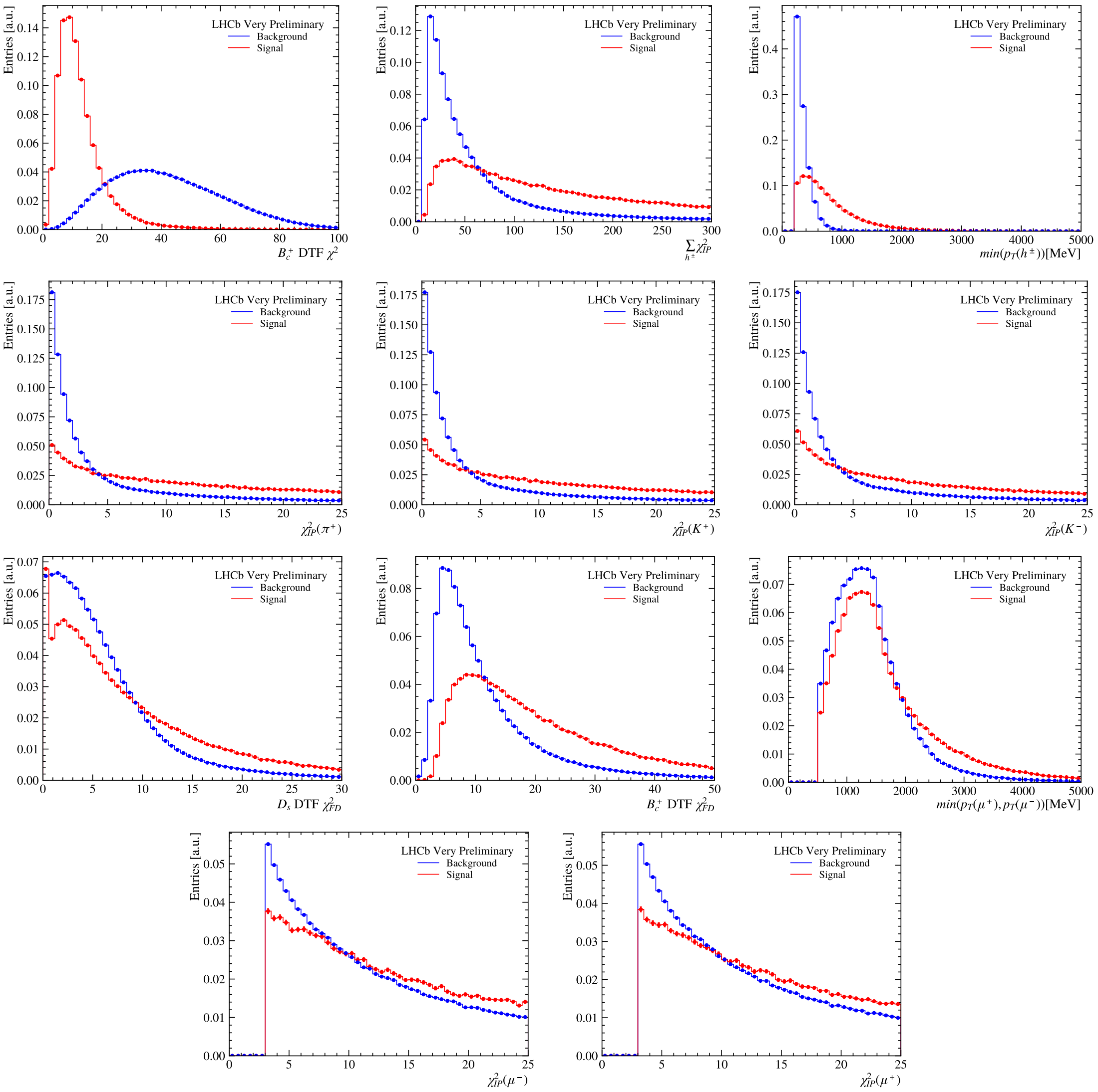


Figure C.3: Distributions of the signal proxy (red) and background proxy (blue) in the decay variables used in the training of the BDT for the $B_c^+ \to J/\psi D_s^+$ normalisation channel.

C.3 Trigger efficiency corrections

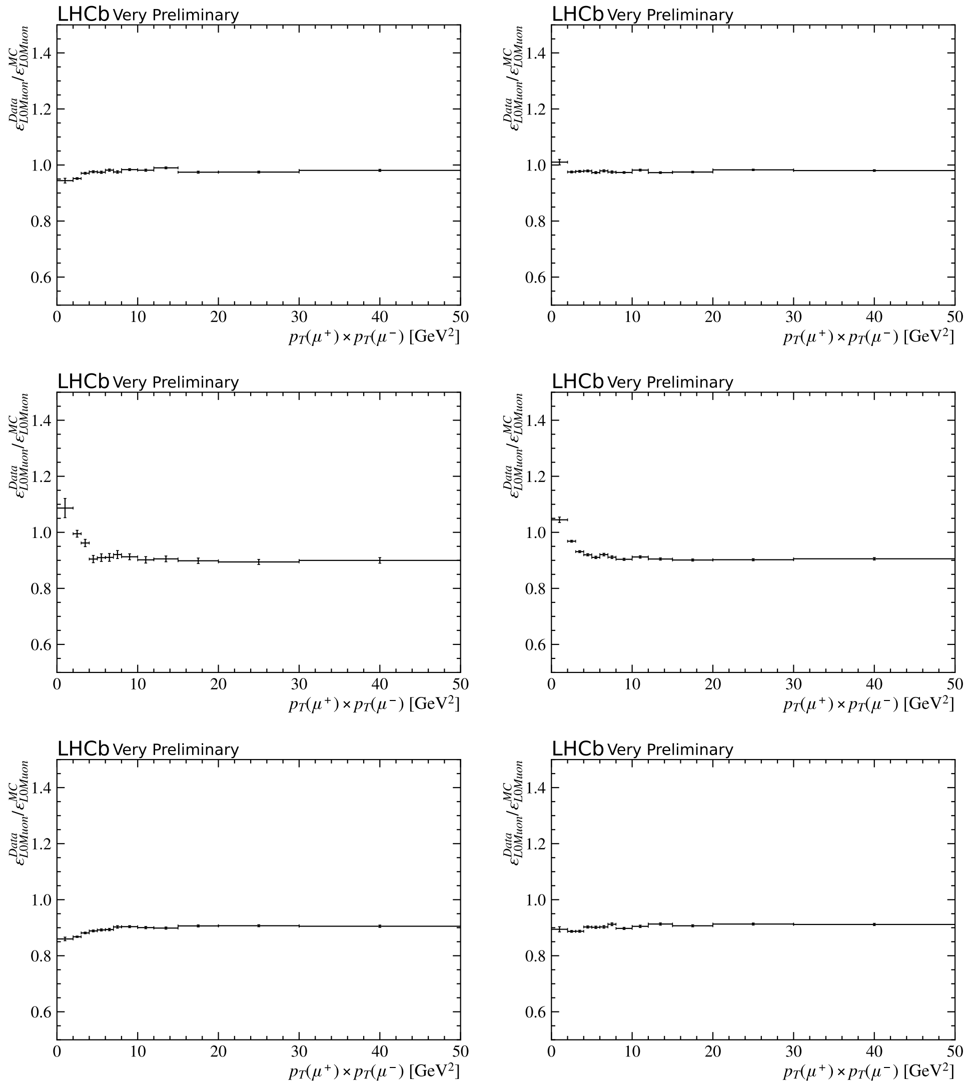


Figure C.4: Ratio of the `L0Muon` trigger efficiencies in data and simulation of $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$ decays for the years 2011 and 2012 (top), 2015 and 2016 (centre) and 2017 and 2018 (bottom).

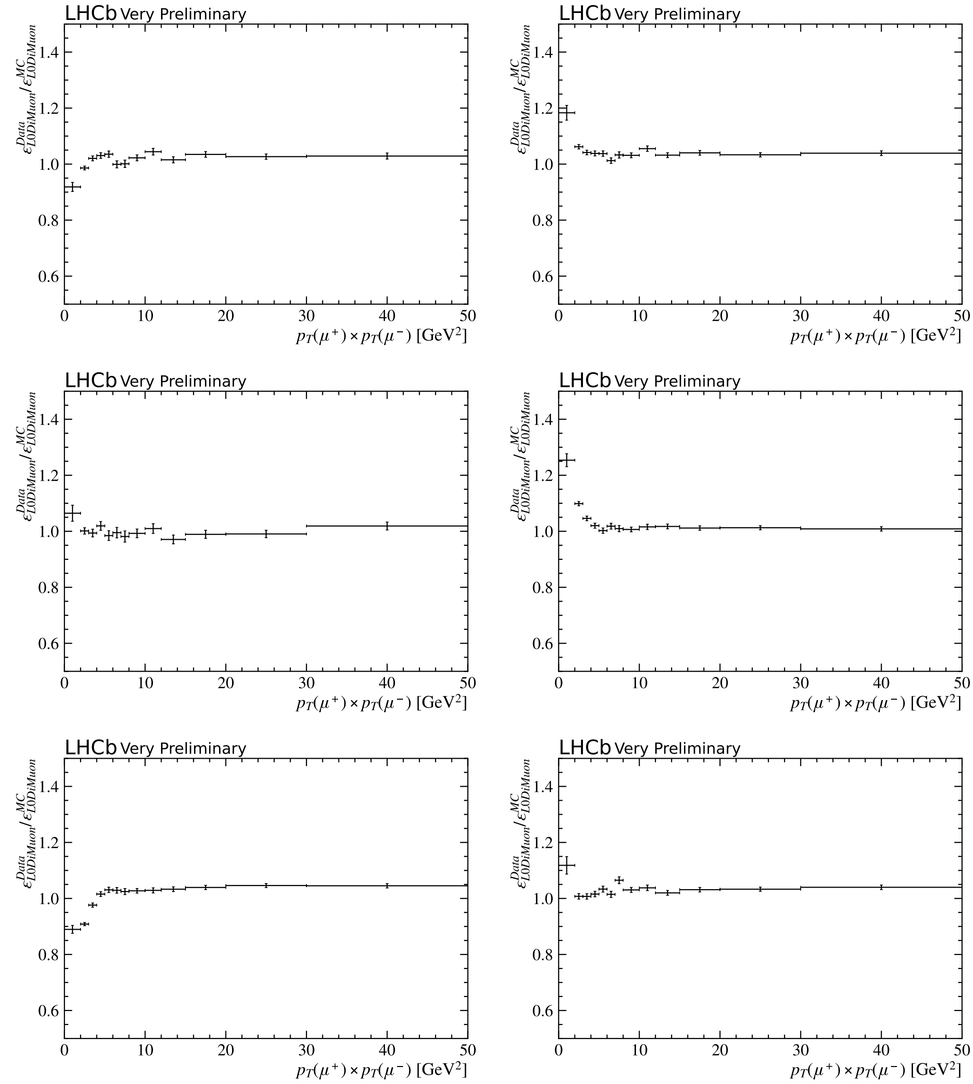


Figure C.5: Ratio of the `L0DiMuon` trigger efficiencies in data and simulation of $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$ decays for the years 2011 and 2012 (top), 2015 and 2016 (centre) and 2017 and 2018 (bottom).

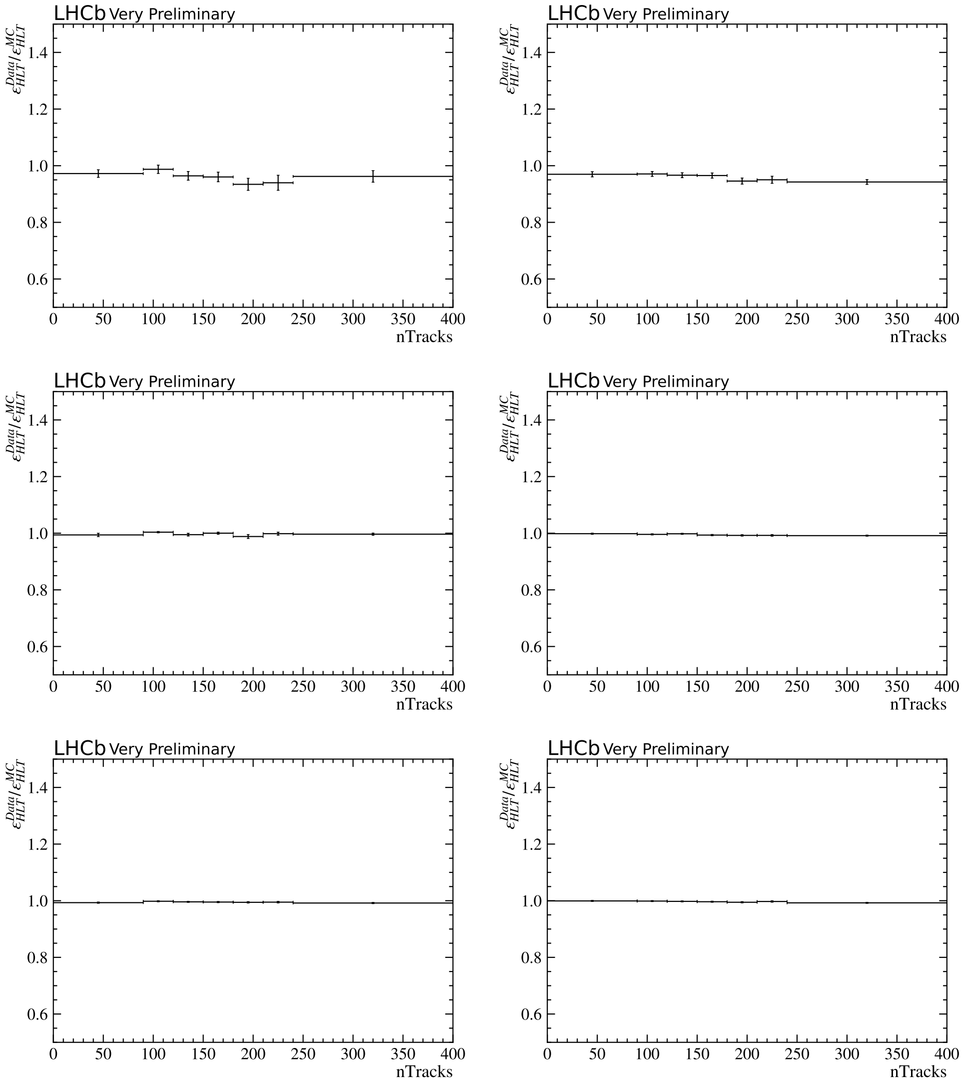


Figure C.6: Ratio of the HLT trigger efficiencies in data and simulation of $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$ decays for the years 2011 and 2012 (top), 2015 and 2016 (centre) and 2017 and 2018 (bottom).

C.4 Breakdown of efficiencies

Year	ε_{geo} $[10^{-2}]$		
	$B_c^+ \to J/\psi D^{*+}$ $D^0 \to K^-\pi^+$	$B_c^+ \to J/\psi D^{*+}$ $D^0 \to K^-\pi^+\pi^+\pi^-$	$B_c^+ \to J/\psi D_s^+$ $D_s^+ \to K^+K^-\pi^+$
2011	12.07 ± 0.02	11.54 ± 0.02	12.34 ± 0.01
2012	12.53 ± 0.02	11.97 ± 0.02	12.84 ± 0.01
2015	14.25 ± 0.03	13.54 ± 0.03	14.51 ± 0.02
2016	14.25 ± 0.02	13.56 ± 0.02	14.52 ± 0.02
2017	14.26 ± 0.02	13.55 ± 0.02	14.51 ± 0.02
2018	14.28 ± 0.02	13.56 ± 0.02	14.50 ± 0.02

Table C.2: Geometrical detector acceptance efficiencies per year averaged across the two magnet polarities.

Year	$\varepsilon_{reco\&sel}$ $[10^{-2}]$		
	$B_c^+ \to J/\psi D^{*+}$ $D^0 \to K^-\pi^+$	$B_c^+ \to J/\psi D^{*+}$ $D^0 \to K^-\pi^+\pi^+\pi^-$	$B_c^+ \to J/\psi D_s^+$ $D_s^+ \to K^+K^-\pi^+$
2011	1.53 ± 0.03	0.39 ± 0.02	1.97 ± 0.01
2012	1.61 ± 0.02	0.43 ± 0.01	2.03 ± 0.01
2015	2.24 ± 0.03	0.62 ± 0.02	2.73 ± 0.02
2016	2.19 ± 0.02	0.62 ± 0.01	2.51 ± 0.01
2017	2.27 ± 0.02	0.68 ± 0.01	2.64 ± 0.02
2018	2.10 ± 0.02	0.60 ± 0.01	2.51 ± 0.01

Table C.3: Reconstruction and selection efficiencies per year averaged across the two magnet polarities.

C.5 Systematic uncertainties

Alternative signal fit

An alternative signal model is taken from a generalised asymmetric double-sided Crystal Ball function, defined as

$$f\left(x;\mu,\sigma_L,\alpha_L,n_L,\sigma_R,\alpha_R,n_R\right)=\begin{cases}A_L\cdot\left(B_L-\frac{x-\mu}{\sigma_L}\right)^{-n_L}, & \text{for } \frac{x-\mu}{\sigma_L}<-\alpha_L\\ \exp\left(-\frac{(x-\mu)^2}{2\sigma_L^2}\right), & \text{for } -\alpha_L\leqslant\frac{x-\mu}{\sigma_L}\leqslant 0\\ \exp\left(-\frac{(x-\mu)^2}{2\sigma_R^2}\right), & \text{for } 0<\frac{x-\mu}{\sigma_R}\leqslant\alpha_R\\ A_R\cdot\left(B_R+\frac{x-\mu}{\sigma_R}\right)^{-n_R}, & \text{for } \frac{x-\mu}{\sigma_R}>\alpha_R\end{cases} \tag{C.1}$$

with $A_{L,R}$ and $B_{L,R}$ following the definition in Eq. 3.6.

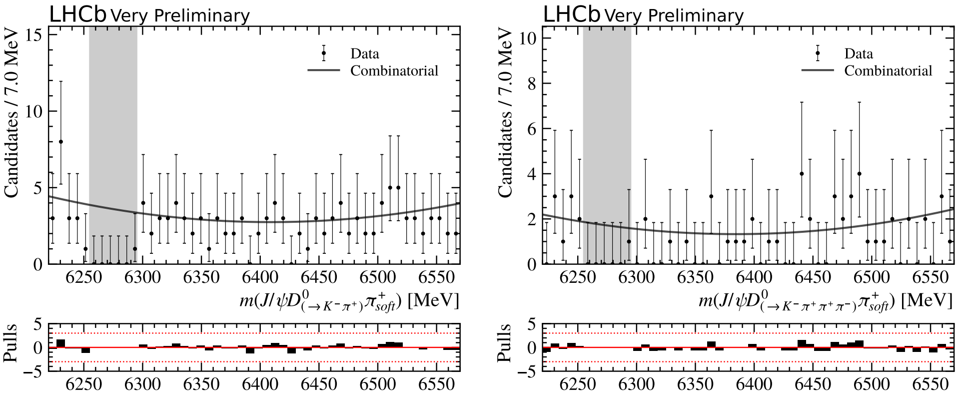


Figure C.7: Blinded fit to the reconstructed invariant mass of $B_c^+ \to J/\psi D^{*+}$ candidates reconstructed from $D^0 \to K^-\pi^+$ (left) and $D^0 \to K^-\pi^+\pi^+\pi^-$ (right) using an alternative fit model. The signal is modelled from a generalised Crystal Ball function. The background is modelled from a second order polynomial. The gray region is the blinded signal region around the B_c^+ mass.

Bibliography

[1] Belle collaboration, M. Gelb *et al.*, *Search for the rare decay of* $B^+ \to \ell^+\nu_\ell\gamma$ *with improved hadronic tagging*, Phys. Rev. D **98** (2018) 112016, `arXiv:1810.12976`.

[2] W. Wang and R.-L. Zhu, *Radiative leptonic* $B_c \to \gamma\ell\bar{\nu}$ *decay in effective field theory beyond leading order*, The European Physical Journal C **75** (2015) .

[3] E. Noether, *Invariant variation problems*, Transport Theory and Statistical Physics **1** (1971) 186–207.

[4] C. Burghard and D. Galbraith, *A standard diagram of the current standard model of physics.*, 2016. `http://www.texample.net/tikz/examples/model-physics/` (Accessed: 23 June 2025).

[5] Particle Data Group, R. L. Workman *et al.*, *Review of particle physics*, Prog. Theor. Exp. Phys. **2022** (2022) 083C01.

[6] C. Abel *et al.*, *Measurement of the permanent electric dipole moment of the neutron*, Physical Review Letters **124** (2020) .

[7] M. Pospelov and A. Ritz, *Theta vacua, QCD sum rules, and the neutron electric dipole moment*, Nuclear Physics B **573** (2000) 177–200.

[8] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[9] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[10] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51** (1983) 1945.

[11] Q. R. Ahmad *et al.*, *Measurement of the rate of* $\nu_e + d \to p + p + e^-$ *interactions produced by* ^8B *solar neutrinos at the Sudbury Neutrino Observatory*, Physical Review Letters **87** (2001) .

[12] Q. R. Ahmad *et al.*, *Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory*, Physical Review Letters **89** (2002) .

[13] S. Fukuda *et al.*, *Solar* ^8B *and hep Neutrino Measurements from 1258 Days of Super-Kamiokande Data*, Physical Review Letters **86** (2001) 5651–5655.

[14] N. Aghanim *et al.*, *Planck2018 results: VI. cosmological parameters*, Astronomy & Astrophysics **641** (2020) A6.

[15] DES Collaboration, T. M. C. Abbott *et al.*, *Dark energy survey year 3 results: Cosmological constraints from cluster abundances, weak lensing, and galaxy clustering*, `arXiv:2503.13632`.

[16] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[17] P. Huet and E. Sather, *Electroweak baryogenesis and standard model CP violation*, Phys. Rev. **D51** (1995) 379, `arXiv:hep-ph/9404302`.

[18] S. Aoki *et al.*, *Review of lattice results concerning low-energy particle physics*, `arXiv:1607.00299`.

[19] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, *QCD factorization for exclusive non-leptonic B-meson decays: general arguments and the case of heavy–light final states*, Nuclear Physics B **591** (2000) 313–418.

[20] J. D. Bjorken, *Topics in b-physics*, Nuclear Physics B - Proceedings Supplements **11** (1989) 325.

[21] RQCD collaboration, G. S. Bali *et al.*, *Light-cone distribution amplitudes of pseudoscalar mesons from lattice QCD*, JHEP **08** (2019) 065, `arXiv:1903.08038`, [Addendum: JHEP 11, 037 (2020)].

[22] CLEO Collaboration, J. Gronberg *et al.*, *Measurements of the meson-photon transition form factors of light pseudoscalar mesons at large momentum transfer*, Phys. Rev. D **57** (1998) 33.

[23] BABAR Collaboration, B. Aubert *et al.*, *Measurement of the $\gamma\gamma^* \to \pi^0$ transition form factor*, Phys. Rev. D **80** (2009) 052002.

[24] A. G. Grozin and M. Neubert, *Asymptotics of heavy-meson form factors*, Physical Review D **55** (1997) 272–290.

[25] G. P. Korchemsky, D. Pirjol, and T.-M. Yan, *Radiative leptonic decays of B mesons in QCD*, Phys. Rev. D **61** (2000) 114510, `arXiv:hep-ph/9911427`.

[26] E. Lunghi, D. Pirjol, and D. Wyler, *Factorization in leptonic radiative $B \to \gamma e\nu$ decays*, Nucl. Phys. B **649** (2003) 349, `arXiv:hep-ph/0210091`.

[27] M. Beneke and J. Rohrwild, *B meson distribution amplitude from $B \to \gamma\ell\nu$*, The European Physical Journal C **71** (2011) .

[28] V. M. Braun, D. Y. Ivanov, and G. P. Korchemsky, *B-meson distribution amplitude in QCD*, Physical Review D **69** (2004) .

[29] M. Beneke, T. Huber, and X.-Q. Li, *NNLO vertex corrections to non-leptonic B decays: Tree amplitudes*, Nuclear Physics B **832** (2010) 109–151.

[30] R. Mandal, S. Nandi, and I. Ray, *Constraining inverse moment of B-meson distribution amplitude using Lattice QCD data*, Physics Letters B **848** (2024) 138345.

[31] S. W. Bosch, R. J. Hill, B. O. Lange, and M. Neubert, *Factorization and Sudakov resummation in leptonic radiative B decay*, Physical Review D **67** (2003) .

[32] M. Beneke and T. Feldmann, *Symmetry-breaking corrections to heavy-to-light B meson form factors at large recoil*, Nuclear Physics B **592** (2001) 3–34.

[33] M. Beneke, V. M. Braun, Y. Ji, and Y.-B. Wei, *Radiative leptonic decay* $B \to \gamma \ell \nu_\ell$ *with subleading power corrections*, Journal of High Energy Physics **2018** (2018) .

[34] BaBar collaboration, B. Aubert *et al.*, *Search for the radiative leptonic decay* $B^+ \to \gamma \ell^+ \nu_l$, `arXiv:0704.1478`.

[35] BABAR Collaboration, B. Aubert *et al.*, *Model-independent search for the decay* $B^+ \to l^+ \nu_l \gamma$, Phys. Rev. D **80** (2009) 111105.

[36] Belle collaboration, A. Heller *et al.*, *Search for* $B^+ \to \ell^+ \nu_\ell \gamma$ *decays with hadronic tagging using the full Belle data sample*, Phys. Rev. D **91** (2015) 112009, `arXiv:1504.05831`.

[37] N. Barik, S. Naimuddin, P. C. Dash, and S. Kar, *Radiative leptonic* B_c *decay in the relativistic independent quark model*, Phys. Rev. D **78** (2008) 114030.

[38] CDF collaboration, P. P. Singh, *Discovery of the* B_c *meson*, in *12th Les Rencontre de Physique de la Vallee d'Aoste: Results and Perspectives in Particle Physics*, 1998, `arXiv:hep-ex/9807022`.

[39] T. M. Aliev and M. Savcı, $B_c \to \ell \nu \gamma$ *decay in light cone QCD*, Physics Letters B **434** (1998) 358–364.

[40] P. Colangelo and F. D. Fazio, *Radiative leptonic* B_c *decays*, Modern Physics Letters A **14** (1999) 2303–2312.

[41] G. Chiladze, A. F. Falk, and A. A. Petrov, *Radiative leptonic* B_c *decays in effective field theory*, Phys. Rev. D **60** (1999) 034011.

[42] C.-H. Chang, C.-D. Lü, G.-L. Wang, and H.-S. Zong, *Pure leptonic decays of the* B_c *meson and their radiative corrections*, Phys. Rev. D **60** (1999) 114013.

[43] L. R. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001, This report is an abridged version of the LHC Design Report (CERN-2004-003).

[44] O. S. Brüning *et al.*, *LHC Design Report*, CERN Yellow Reports: Monographs, CERN, Geneva, 2004.

[45] E. Mobs, *The CERN accelerator complex in 2019*, 2019. General Photo. 2019. https://cds.cern.ch/record/2684277.

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

[47] R. Aaij *et al.*, *Measurement of the b-quark production cross section in 7 and 13 TeV collisions*, Physical Review Letters **118** (2017) .

[48] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the B_c^- meson production fraction and asymmetry in 7 and 13 TeV pp collisions*, Phys. Rev. D **100** (2019) 112006.

[49] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011, 2001.

[50] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, `arXiv:1405.7808`.

[51] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030, 2003.

[52] C. Abellan Beteta *et al.*, *Monitoring radiation damage in the LHCb Tracker Turicensis*, `arXiv:1809.05063`.

[53] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029, 2002.

[54] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024, 2001.

[55] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, `arXiv:1311.3893`.

[56] P. d'Argent *et al.*, *Improved performance of the LHCb Outer Tracker in LHC Run 2*, JINST **12** (2017) P11016, `arXiv:1708.00819`.

[57] LHCb collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007, 2000.

[58] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, `arXiv:1412.6352`.

[59] S. Klous, *The LHCb vertex detector*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **549** (2005) 55, VERTEX 2003.

[60] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, `arXiv:1408.1251`.

[61] LHCb collaboration, *LHCb RICH: Technical Design Report*, CERN-LHCC-2000-037, 2000.

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

[63] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036, 2000.

[64] Y. N. Kharzheev, *Scintillation counters in modern high-energy physics experiments (review)*, Phys. Part. Nuclei **46** (2015) .

[65] C. Abellan Beteta *et al.*, *Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC*, `arXiv:2008.11556`, submitted to JINST.

[66] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010, 2001.

[67] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, `arXiv:1211.1346`.

[68] R. Calabrese *et al.*, *Performance of the LHCb RICH detectors during LHC Run 2*, JINST **17** (2022) P07013, `arXiv:2205.13400`.

[69] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, `arXiv:1306.0249`.

[70] LHCb collaboration, *LHCb trigger system: Technical Design Report*, CERN-LHCC-2003-031, 2003.

[71] LHCb collaboration, *Trigger schemes*, http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html (accessed 18. May 2025).

[72] 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`.

[73] G. Corti *et al.*, *Software for the lhcb experiment*, Nuclear Science, IEEE Transactions on **53** (2006) 1323 .

[74] LHCb collaboration, *LHCb computing: Technical Design Report*, CERN-LHCC-2005-019, 2005.

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

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

[77] C.-H. Chang, C. Driouichi, P. Eerola, and X.-G. Wu, *BCVEGPY: an event generator for hadronic production of the Bc meson*, Computer Physics Communications **159** (2004) 192–224.

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

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

[80] N. Davidson, T. Przedzinski, and Z. Was, *PHOTOS interface in C++: Technical and physics documentation*, Comp. Phys. Comm. **199** (2016) 86, `arXiv:1011.0937`.

[81] LHCb collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021, 2013.

[82] LHCb collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001, 2014.

[83] LHCb collaboration, *LHCb PID Upgrade Technical Design Report*, CERN-LHCC-2013-022, 2013.

[84] LHCb collaboration, *LHCb Trigger and Online Upgrade Technical Design Report*, CERN-LHCC-2014-016, 2014.

[85] LHCb collaboration, R. Aaij *et al.*, *The LHCb Upgrade I*, JINST **19** (2024) P05065, `arXiv:2305.10515`.

[86] LHCb collaboration, *LHCb Particle Identification Enhancement Technical Design Report*, CERN-LHCC-2023-005, 2023.

[87] LHCb collaboration, *LHCb Upgrade II Scoping Document*, CERN-LHCC-2024-010, 2024.

[88] L. An *et al.*, *Performance of a spaghetti calorimeter prototype with tungsten absorber and garnet crystal fibres*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **1045** (2023) 167629.

[89] LHCb collaboration, *LHCb Framework TDR for the LHCb Upgrade II Opportunities in flavour physics, and beyond, in the HL-LHC era*, CERN-LHCC-2021-012, 2022.

[90] LHCb ECAL UpgradeII R&D group, A. Bordelius, *The LHCb PicoCal*, Nucl. Instrum. Methods Phys. Res. , A **1079** (2025) 170608.

[91] LHCb collaboration, F. Machefert, *A calorimeter for the LHCb upgrade II*, Calorimeter upgrade workshop (IN2P3). 2021. (accessed 6. Aug 2025).

[92] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions* $\mathcal{B}(\overline{B}^0 \to D^{*+}\tau^-\overline{\nu}_\tau)/\mathcal{B}(\overline{B}^0 \to D^{*+}\mu^-\overline{\nu}_\mu)$, Phys. Rev. Lett. **115** (2015) 111803, Publisher's Note ibid. **115** (2015) 159901, `arXiv:1506.08614`.

[93] D. van Dyk *et al.*, *EOS - A Software for Flavor Physics Phenomenology*, `arXiv:2111.15428`.

[94] LHCb collaboration, R. Aaij *et al.*, *Measurement of lepton universality parameters in* $B^+ \to K^+\ell^+\ell^-$ *and* $B^0 \to K^{*0}\ell^+\ell^-$ *decays*, Phys. Rev. **D108** (2023) 032002, `arXiv:2212.09153`.

[95] LHCb collaboration, R. Aaij *et al.*, *Strong constraints on the* $b \to s\gamma$ *photon polarisation from* $B^0 \to K^{*0}e^+e^-$ *decays*, JHEP **12** (2020) 081, `arXiv:2010.06011`.

[96] M.-H. Schune, F. Polci, and M. Borsato, *HOP an additional tool for decays involving electrons*, LHCb-INT-2015-037, 2015.

[97] T. Chen and C. Guestrin, *XGBoost: A scalable tree boosting system*, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, 785–794, ACM, 2016.

[98] 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/`.

[99] LHCb collaboration, R. Aaij *et al.*, *First observation of the decay $B_s^0 \to K^- \mu^+ \nu_\mu$ and measurement of $|V_{ub}|/|V_{cb}|$*, Phys. Rev. Lett. **126** (2021) 081804, `arXiv:2012.05143`.

[100] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021, 2016.

[101] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005) 356, `arXiv:physics/0402083`.

[102] LHCb collaboration, R. Aaij *et al.*, *Measurement of the electron reconstruction efficiency at LHCb*, JINST **14** (2019) P11023, `arXiv:1909.02957`.

[103] J. Li and H. Sun, *HEP ML Lab: An end-to-end framework for applying machine learning into phenomenology studies*, `arXiv:2405.02888`.

[104] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, `arXiv:1211.3055`.

[105] R. J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. **77** (1993) 219.

[106] J. S. Conway, *Incorporating nuisance parameters in likelihoods for multisource spectra*, `arXiv:1103.0354`.

[107] H. Dembinski *et al.*, *scikit-hep/iminuit*, 2025. doi: 10.5281/zenodo.15157028.

[108] J. Eschle, A. Puig Navarro, R. Silva Coutinho, and N. Serra, *zfit: Scalable pythonic fitting*, SoftwareX **11** (2020) 100508.

[109] E. Rodrigues *et al.*, *The Scikit HEP project overview and prospects*, EPJ Web of Conferences **245** (2020) 06028.

[110] M. Rowe and R. Zwicky, *Structure-dependent QED in $B^- \to \ell^- \bar{\nu}(\gamma)$*, `arXiv:2404.07648`.

[111] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Annals Math. Statist. **9** (1938) 60.

[112] G. J. Feldman and R. D. Cousins, *Unified approach to the classical statistical analysis of small signals*, Phys. Rev. D **57** (1998) 3873.

[113] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, The European Physical Journal C **71** (2011) .

[114] A. L. Read, *Presentation of search results: The CL_s technique*, J. Phys. **G28** (2002) 2693.

[115] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the decay* $B_s^0 \to \phi e^+e^-$, JHEP **07** (2025) 069, `arXiv:2504.06346`.

[116] LHCb collaboration, R. Aaij *et al.*, *Test of lepton flavor universality with* $B^+ \to K^+\pi^+\pi^-\ell^+\ell^-$ *decays*, Phys. Rev. Lett. **134** (2025) 181803, `arXiv:2412.11645`.

[117] LHCb collaboration, R. Aaij *et al.*, *Test of lepton flavour universality with* $B_s^0 \to \phi\ell^+\ell^-$ *decays*, Phys. Rev. Lett. **134** (2025) 121803, `arXiv:2410.13748`.

[118] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of* $B^0 \to K^{*0}e^+e^-$ *decays*, JHEP **06** (2025) 140, `arXiv:2502.10291`.

[119] LHCb collaboration, R. Aaij *et al.*, *Comprehensive analysis of local and nonlocal amplitudes in the* $B^0 \to K^{*0}\mu^+\mu^-$ *decay*, JHEP **09** (2024) 026, Erratum ibid. **05** (2025) 208, `arXiv:2405.17347`.

[120] LHCb collaboration, R. Aaij *et al.*, *Determination of short- and long-distance contributions in* $B^0 \to K^{*0}\mu^+\mu^-$ *decays*, Phys. Rev. **D109** (2024) 052009, `arXiv:2312.09102`.

[121] LHCb collaboration, R. Aaij *et al.*, *Tests of lepton universality using* $B^0 \to K^0_S\ell^+\ell^-$ *and* $B^+ \to K^{*+}\ell^+\ell^-$ *decays*, Phys. Rev. Lett. **128** (2022) 191802, `arXiv:2110.09501`.

[122] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the* $B^+ \to K^{*+}\mu^+\mu^-$ *decay*, Phys. Rev. Lett. **126** (2021) 161802, `arXiv:2012.13241`.

[123] LHCb collaboration, R. Aaij *et al.*, *A search for rare* $B \to D\mu^+\mu^-$ *decays*, JHEP **02** (2024) 032, `arXiv:2308.06162`.

[124] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the* B_c^+ *meson mass*, JHEP **07** (2020) 123, `arXiv:2004.08163`.

[125] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152.

[126] T. Evans et. al. , *AmpGen software*, https://github.com/GooFit/AmpGen/ (accessed 25. August 2025).

[127] LHCb collaboration, R. Aaij *et al.*, *Studies of the resonance structure in* $D^0 \to K^{\mp}\pi^{\pm}\pi^+\pi^-$ *decays*, Eur. Phys. J. **C78** (2018) 443, `arXiv:1712.08609`.

[128] LHCb collaboration, R. Aaij *et al.*, *Search for* B_c^+ *decays to two charm mesons*, Nucl. Phys. **B930** (2018) 563, `arXiv:1712.04702`.