THESE DE DOCTORAT

NNT : 2024UPASP049

Study of rare beauty baryon decays with the LHCb detector

Étude des désintégrations rares des baryons beaux avec le détecteur LHCb

Thèse de doctorat de l'université Paris-Saclay et de l'université TU Dortmund

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 **Marie-Hélène SCHUNE**, Directrice de recherche - Université Paris-Saclay,
et la co-direction **Johannes ALBRECHT**, Professeur - TU Dortmund University

Thèse soutenue à Paris-Saclay, le 05 Juillet 2024, par

Janina NICOLINI

Composition du jury

Membres du jury avec voix délibérative

Fabien CAVALIER Directeur de recherche, IJCLab / Université Paris-Saclay	Président
Chris DELITZSCH Professeure associée (équiv. HDR), TU Dortmund University	Rapportrice & Examinatrice
Jonas RADEMACKER Professeur, University of Bristol	Rapporteur & Examinateur
Danny VAN DYK Professeur assistant, Durham University	Examinateur
Justine SERRANO Directrice de recherche, Centre de Physique des Particules de Marseille	Examinatrice
Timothy GERSHON Professeur, University of Warwick	Examinateur

technische universität dortmund

université PARIS-SACLAY

Study of rare beauty baryon decays with the LHCb detector

Janina Nicolini

2024

A document submitted in partial fulfillment of the requirements for the degree of

Doctor rerum naturalium
at
Fakultät Physik, Technische Universität Dortmund

as part of a joint doctorate in collaboration with
Université Paris-Saclay

Supervised by
Prof. Dr. Johannes Albrecht and Dr. Marie-Hélène Schune

Reviewed by
Prof. Dr. Johannes Albrecht and Dr. Chris Delitzsch

université PARIS-SACLAY

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

Title: Study of rare beauty baryon decays with the LHCb detector

Keywords: *Flavour Physics, LHCb, B physics, Angular analysis, Baryons, Rare Decays*

Abstract: Flavour-changing neutral currents, such as $b \to s\ell^+\ell^-$ transitions, are forbidden at tree level in the Standard Model (SM), and can only occur at loop level. Therefore, they are sensitive to small deviations from the presence of possible New Physics (NP) effects with respect to the SM predictions. While ordinary matter is made out of baryons, very few measurements have been performed with baryonic $b \to s\ell^+\ell^-$ transitions due to the production suppression of beauty baryons at collider experiments.

In this thesis, several analyses are presented to improve the knowledge of the four weakly decaying beauty baryons and respective $b \to s\ell^+\ell^-$ transitions.

First of all, the current status of the angular analysis of $\Lambda_b^0 \to \Lambda\ell^+\ell^-$ decays is presented for the first time using the full Run 1 and 2 datasets collected by the LHCb experiment. All ten angular coefficients are attempted to be measured in the dilepton invariant mass squared, q^2, range from 15 to 20 GeV$^2/c^4$ for the muon mode. The muon mode fit procedure has been validated with pseudo-experiments. The angular observables of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ are kept blind in the thesis.

Furthermore, the charmonium decay modes, $\Xi_b^- \to \Xi^- J/\psi$ and $\Omega_b^- \to \Omega^- J/\psi$, have been used to improve the knowledge of the Ω_b^- baryon properties. Similarly, the previously not observed $\Xi_b^0 \to \Xi^0 J/\psi$ and $\Xi_b^0 \to \Xi^0\psi(2S)$ decays have been studied. In addition, this thesis also presents the current status for the first search for the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay and the observation of the respective $\Xi_b^- \to \Xi^-\psi(2S)$ decay.

Besides the analysis work conducted in this thesis, the contributions to the Scintillating Fibre (SciFi) detector of the LHCb experiment are discussed.

université PARIS-SACLAY

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

Titre: Étude des désintégrations rares des baryons beaux avec le détecteur LHCb

Mots clés: *Physique de la saveur, LHCb, physique du B, analyse angulaire, baryons, désintégrations rares*

Résumé Les courants neutres avec changement de saveur, tels que les transitions $b \to s\ell^+\ell^-$, sont interdits au premier ordre dans le Modèle Standard (MS) et ne peuvent se produire qu'au niveau des boucles. Par conséquent, ils sont sensibles à la présence potentielle de Nouvelle Physique (NP) qui se manifestait par la présence de petites déviations par rapport aux prédictions du MS. Alors que la matière ordinaire est composée de baryons, très peu de mesures ont été réalisées avec des transitions baryoniques $b \to s\ell^+\ell^-$ en raison de la suppression de la production de baryons de beauté dans les expériences travaillant auprès de collisionneurs.

Dans cette thèse, plusieurs analyses sont présentées pour améliorer la connaissance des quatre baryons de beauté pour ce qui concerne leurs désintégrations via des transitions $b \to s\ell^+\ell^-$.

Tout d'abord, le statut actuel de l'analyse angulaire des désintégrations $\Lambda_b^0 \to \Lambda\ell^+\ell^-$ est réalisé pour la première fois en utilisant l'ensemble de la statistique collectée par l'expérience LHCb. Les dix coefficients angulaires sont mesurés dans la plage de masse invariante dileptonique au carré, q^2, de 15 à 20 GeV^2/c^4 pour le mode muonique. La procédure d'ajustement a été validée avec des pseudo-expériences. Les observables angulaires des désintégrations $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ ne sont pas dévoilées sur les données dans le document de thèse.

De plus, les modes de désintégration via des charmonia, $\Xi_b^- \to \Xi^- J/\psi$ et $\Omega_b^- \to \Omega^- J/\psi$, ont été utilisés pour améliorer la connaissance des propriétés du baryon Ω_b^-. De manière similaire, les désintégrations $\Xi_b^0 \to \Xi^0 J/\psi$ et $\Xi_b^0 \to \Xi^0\psi(2S)$ précédemment non observées ont été étudiées.Cette thèse présente également le statut actuel de la première recherche de la désintégration $\Xi_b^- \to \Xi^-\mu^+\mu^-$ et de l'observation de la désintégration $\Xi_b^- \to \Xi^-\psi(2S)$.

Outre le travail d'analyse mené dans cette thèse, la contributions au détecteur à fibre scintillante (SciFi) de l'expérience LHCb est discutée.

université PARIS-SACLAY

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

Titel: Untersuchung von seltenen Beauty-Baryonenzerfällen mit dem LHCb-Detektor
Schlagwröter: *Flavourphysik, LHCb, B-Physik, Winkelanalyse, Baryonen, Seltene Zerfälle*

Kurzfassung: Flavour-ändernde neutrale Ströme, wie $b \to s\ell^+\ell^-$ Übergänge, sind im Standardmodell (SM) auf Baumgraphen-Niveau verboten und können nur auf Schleifenebene auftreten. Daher sind sie sehr sensitiv gegenüber kleinsten Abweichungen in Bezug auf die SM-Vorhersagen erzeugt durch die mögliche Existenz der Neuen Physik (NP). Während gewöhnliche Materie aus Baryonen besteht, wurden bisher nur sehr wenige Messungen mit baryonischen $b \to s\ell^+\ell^-$-Übergängen durchgeführt, da die Produktion von Beauty-Baryonen bei Beschleunigerexperimenten unterdrückt wird.
In dieser Arbeit werden mehrere Analysen vorgestellt, um das Wissen über die vier schwach zerfallenden Beauty-Baryonen und die entsprechenden $b \to s\ell^+\ell^-$-Übergänge zu verbessern.
Zunächst wird der aktuelle Stand der Winkelanalyse von $\Lambda_b^0 \to \Lambda\ell^+\ell^-$-Zerfällen präsentiert, in der zum ersten Mal die vollständigen Datensätze vom Run 1 und 2 des LHCb-Experiments genutzt werden. Alle zehn Winkelkoeffizienten werden im Bereich der quadrierten invarianten Dileptonenmasse q^2 von 15 bis 20 GeV^2/c^4 für den myonischen Zerfallskanal gemessen. Das Verfahren zur Messung der Winkelkoeffizienten wurde mit Pseudo-Experimenten validiert. Der Signalbereich zur finalen Bestimmung der Werte der Winkelobservablen von dem $\Lambda_b^0 \to \Lambda\mu^+\mu^-$-Zerfall wird in dieser Arbeit ausgespart.
Darüber hinaus wurden die Charmonium-Zerfallskanäle $\Xi_b^- \to \Xi^- J/\psi$ und $\Omega_b^- \to \Omega^- J/\psi$ dazu verwendet, die Erkenntnis über die Eigenschaften des Ω_b^- Baryonens zu verbessern.
Ebenso wurden die bisher nicht beobachteten $\Xi_b^0 \to \Xi^0 J/\psi$ und $\Xi_b^0 \to \Xi^0\psi(2S)$ Zerfälle untersucht. Darüber hinaus wird in dieser Arbeit der aktuelle Stand der ersten Suche nach dem $\Xi_b^- \to \Xi^-\mu^+\mu^-$ Zerfall und die Beobachtung des $\Xi_b^- \to \Xi^-\psi(2S)$-Zerfalls dargestellt.
Neben den in dieser Arbeit durchgeführten Analysen werden auch die Beiträge zum SciFi-Detektor des LHCb Experiments diskutiert.

Contents

Synthèse

Les courants neutres changeant de saveur (FCNC) constituent un environnement particulièrement intéressant pour effectuer des recherches indirectes sensibles aux plus petites déviations par rapport aux prédictions du Modèle Standard (SM). Ce sont des transitions entre différentes saveurs de quarks de même charge électrique, qui ne peuvent apparaître qu'au niveau des boucles impliquant des particules de masse plus élevée. Dans le SM, elles sont fortement supprimées à la fois par la presence de particules massives dans la boucle et par le mécanisme Glashow-Iliopoulos-Maiani (GIM) [1]. La contribution du SM à de telles transitions est très faible, ce qui rend ces processus particulièrement sensibles aux moindres déviations dues à de la nouvelle physique (NP). Comme mentionné précédemment, la NP peut contribuer dans les diagrammes en boucle, mais de nouvelles contributions peuvent également apparaître en produisant des FCNC au niveau de l'arbre. Un exemple de ces FCNC fortement supprimés sont les transitions au niveau des quarks $b \to s\ell^+\ell^-$. En raison de la nature des quarks, ils ne peuvent être étudiés que dans des états liés de quarks, appelés hadrons. Au cours de la dernière décennie, un certain nombre de déviations par rapport aux prédictions du SM ont été observées dans les transitions $b \to s\ell^+\ell^-$ par des expériences de physique des hautes énergies [2–17]. Cet ensemble de déviations, appelé anomalies du B, peut indiquer deux directions opposées. Soit ces anomalies sont dues à un manque de compréhension des effets causés par l'interaction forte, qui est également la cause des états liés de quarks, soit elles constituent les premiers indices de physique au-delà du SM. Les anomalies peuvent être étudiées en analysant plus de données disponibles et en incluant davantage de modes de désintégration hadronique basés sur les transitions $b \to s\ell^+\ell^-$.

Alors que la matière ordinaire est constituée de hadrons formés de trois quarks, appelés baryons, la recherche actuelle se concentre principalement sur les transitions mésoniques $b \to s\ell^+\ell^-$ produites plus abondamment. Cependant, au Grand Collisionneur de Hadrons (LHC), tous les types d'hadrons b (mésons ou baryons) peuvent être produits en abondance. L'étude des transitions baryoniques $b \to s\ell^+\ell^-$ est particulièrement intéressante car elle offre non seulement une façon complémentaire de tester les anomalies observées dans les mésons B, mais donne également accès à des observables supplémentaires d'angle et de polarisation. Cela est dû à leur spin non nul, qui peut également être utilisé pour tester une éventuelle dépendance du spin à la nouvelle physique. L'expérience LHCb offre un environnement idéal pour étudier les désintégrations baryoniques.

Il existe plusieurs états de baryons qui contiennent un quark b et deux quarks légers (u, d, s). Cependant, seuls quatre d'entre eux se désintègrent par l'interaction faible, ce qui donne une durée de vie d'environ 1.5 ps. Ceux-ci sont: le Λ_b^0 (avec des quarks de

valence udb), le Ξ_b^- (dsb), le Ξ_b^0 (usb) et l'Ω_b^- (ssb). Cependant, seules les désintégrations du Λ_b^0 ont été étudiées en détail. Pour ce qui est des baryons Ξ_b et Ω_b^-, moins de dix modes de désintégration ont été observés pour chacun d'eux, la plupart utilisant des données de l'expérience LHCb.

Afin d'améliorer la connaissance des désintégrations rares des baryons beaux se désintégrant faiblement, cette thèse présente des études sur les canaux de désintégration de quatre d'entre eux. Des premières mises à jour sur les mesures de précision des propriétés du baryon Ω_b^- sont rapportées, ainsi que des études de plusieurs transitions baryoniques au niveau des boucles $b \to s\ell^+\ell^-$ et de leurs équivalents au niveau des arbres $b \to s(c\bar{c} \to \ell^+\ell^-)$. Ces études comprennent la première recherche du désintégration $\Xi_b^- \to \Xi^-\mu^+\mu^-$, l'observation des désintégrations $\Xi_b^- \to \Xi^-\psi(2S)$ et $\Xi_b^0 \to \Xi^0 J/\psi$, ainsi que les premiers indices prometteurs pour le désintégration $\Xi_b^0 \to \Xi^0\psi(2S)$ et l'analyse angulaire du désintégration $\Lambda_b^0 \to \Lambda\mu^+\mu^-$. À part la désintégration $\Lambda_b^0 \to \Lambda\mu^+\mu^-$, les trois autres modes de désintégration n'avaient pas été observés au début de cette thèse, et seul le $\Xi_b^- \to \Xi^-\psi(2S)$ avait été récemment observé pour la première fois par la collaboration CMS [18]. Toutes les analyses présentées sont réalisées avec des données collectées par l'expérience LHCb.
La première analyse présentée dans cette thèse comprend deux parties et a été publiée dans Ref. [19]. La première partie concerne la mesure de la masse du baryon Ω_b^- à l'aide de l'ensemble des données des Run 1 et Run 2 de l'expérience LHCb, correspondant à une luminosité intégrée de $9\,\text{fb}^{-1}$. Le fait que la différence de masse entre le baryon b initial et ses produits de désintégration (libération d'énergie, ou valeur Q) est assez petite, tant pour les désintégrations $\Omega_b^- \to \Omega^- J/\psi$ que $\Xi_b^0 \to \Xi^0 J/\psi$, fait de ces modes des candidats idéaux pour des mesures de masse. Pour améliorer davantage la sensibilité, la différence de masse entre les baryons Ω_b^- et Ξ_b^- est extraite, plutôt que la valeur absolue de la masse du Ω_b^-. Cette approche présente l'avantage d'annuler de nombreuses incertitudes systématiques en raison des topologies presque identiques des deux modes de désintégration. Ils ne diffèrent que par la présence d'un kaon contre un pion chargé dans l'état final. La sélection développée pour les deux modes de désintégration tire largement parti des topologies uniques qui ne peuvent être imitées par aucun mode de désintégration mésonique. La différence de masse est extraite à partir d'un ajustement simultané avec quatre projections, les ensembles de données du Run 1 et du Run 2 pour les désintégrations Ω_b^- et Ξ_b^-. La mesure de masse la plus précise de la masse du Ξ_b^- [20] à ce jour a ensuite été utilisée pour déterminer la masse du Ω_b^-:

$$\begin{aligned} m(\Omega_b^-) - m(\Xi_b^-) &= 248.54 \pm 0.51\,(\text{stat}) \pm 0.38\,(\text{syst})\,\text{MeV}/c^2\ , \\ m(\Omega_b^-) &= 6045.9 \pm 0.5\,(\text{stat}) \pm 0.6\,(\text{syst})\,\text{MeV}/c^2. \end{aligned}$$

Ce résultat est en accord avec la valeur moyenne mondiale, mais il est plus de deux fois plus précis : cela prouve que LHCb est le principal acteur dans la physique du baryon Ω_b^-. De plus, la moyenne de LHCb pour la masse du Ω_b^- $[m(\Omega_b^-)]_{\text{LHCb average}} = 6045.7 \pm 0.5\,(\text{stat}) \pm 0.6\,(\text{syst})\,\text{MeV}/c^2$ a été déterminée pour tenir compte correctement des corrélations entre différentes mesures de LHCb. Le résumé des mesures de masse du Ω_b^- réalisées par les expériences CDF et LHCb est montré dans la Fig. 1.
En utilisant les mêmes modes de désintégration, la première mesure de la fraction de production relative multipliée par la fraction de branchement du baryon Ω_b^- au LHC

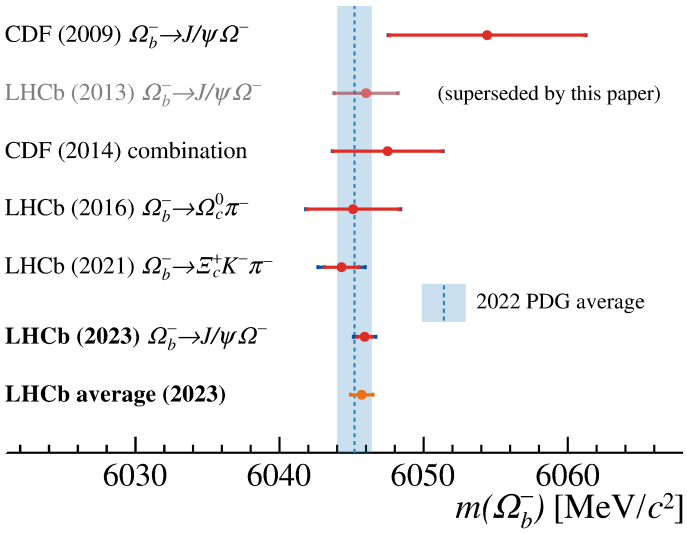


Figure 1: Résumé des mesures existantes de la masse du Ω_b^-. Le graphique exclut la mesure de l'expérience D0 [21] car elle n'est pas cohérente avec toutes les autres et se trouve en dehors de la plage de ce graphique.

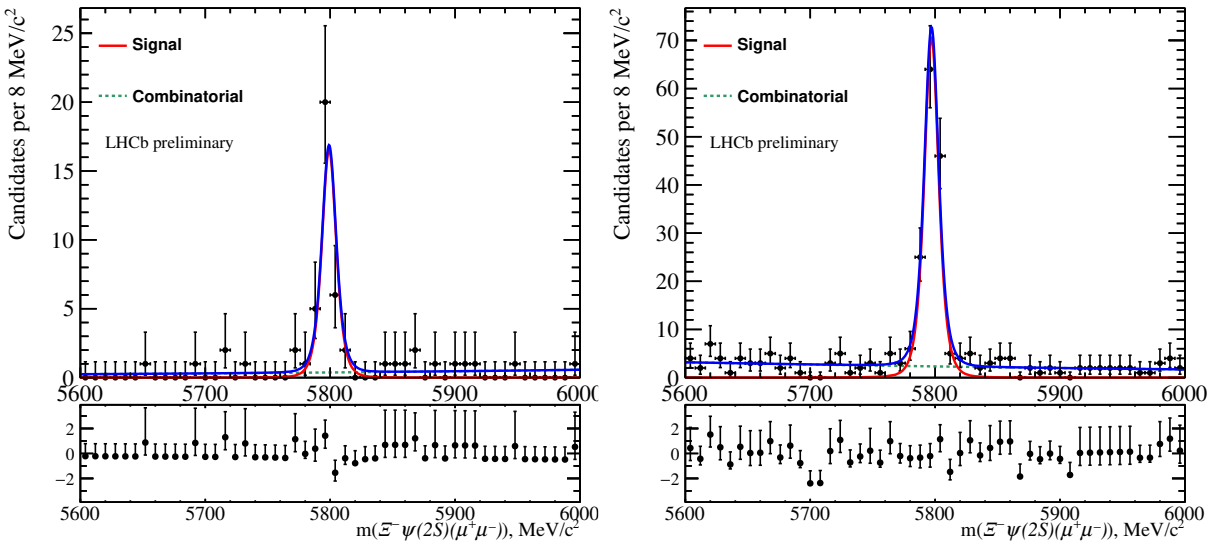
a été réalisée. En raison de la dépendance énergétique des fractions de production et des statistiques limitées de l'échantillon Ω_b^-, la mesure est effectuée uniquement avec l'ensemble des données du Run 2, correspondant à $6\,\text{fb}^{-1}$. Le résultat

$$\frac{f_{\Omega_b^-}}{f_{\Xi_b^-}} \times \frac{\mathcal{B}(\Omega_b^- \to J/\psi\Omega^-)}{\mathcal{B}(\Xi_b^- \to J/\psi\Xi^-)} = 0.120 \pm 0.008(\text{stat}) \pm 0.008(\text{syst}),$$

est obtenu à partir d'un ajustement simultané avec deux projections dans la région cinématique définie de $2 < \eta < 6$ et $p_t < 20$ GeV/c pour les baryons Ξ_b^- et Ω_b^-. Le plus grand défi de l'analyse était la connaissance limitée des propriétés des baryons étranges, ainsi que la grande fraction de traces provenant de désintégrations à longue distance. En raison de leur moindre efficacité de suivi et de vertexing, il est nécessaire de connaître précisément leur proportion dans chaque ensemble de données afin de garantir qu'aucun biais n'est introduit lors de l'étalonnage et de la fusion subséquente des différents ensembles de données pour extraire les résultats finaux. Une bonne stabilité du résultat a été atteinte à travers différents sous-ensembles de données, prouvant le succès de l'étalonnage.
Le chapitre suivant était consacré à l'étude des modes de désintégration rares du Ξ_b^-. La première partie était l'observation du mode de désintégration $\Xi_b^- \to \Xi^-\psi(2S)$ pour la première fois dans l'expérience LHCb, utilisant l'ensemble des données du Run 1 et du Run 2 correspondant à $9\,\text{fb}^{-1}$. Le résultat final est extrait d'un ajustement simultané avec quatre projections, le résultat des projections de l'ajustement du Run 1 pour $\Xi_b^- \to \Xi^-\psi(2S)$ étant montré à la Fig. 2.
La fraction de branchement du mode de désintégration par rapport à celle du $\Xi_b^- \to \Xi^- J/\psi$ a été évaluée et a donné

$$r_{BR}^{\psi}(\Xi_b^-) = 0.492 \pm 0.041(\text{stat}) \pm 0.021(\text{syst}).$$

Lorsqu'on compare ce résultat à la première observation réalisée par la collaboration CMS [18], une différence d'environ 1.5 écart-type est observée, mais le résultat de LHCb



(a) Run 1 (b) Run 2

Figure 2: Projections de l'ajustement simultané pour les ensembles de données $\Xi_b^- \to \Xi^-\psi(2S)$ dans Run 1 (à gauche) et Run 2 (à droite).

présenté dans cette thèse est cinq fois plus précis. De plus, le rapport r_{BR}^{ψ} diffère également des prédictions du modèle des quarks [22], tout comme les mesures des rapports présentés dans les Refs. [23, 24] pour les désintégrations du Λ_b^0 (6.3). Ce résultat souligne l'importance d'étudier également les modes de désintégration des charmonia et leur potentiel pour tester les prédictions du modèle des quarks.

La deuxième partie était consacrée à l'état actuel de la première recherche du mode de désintégration $\Xi_b^- \to \Xi^-\mu^+\mu^-$ avant le démasquage des données. En raison de l'absence de toute mesure absolue de fraction de branchement des désintégrations de Ξ_b^-, le mode de désintégration $\Xi_b^- \to \Xi^- J/\psi$ a été choisi comme norme pour annuler le maximum d'incertitudes systémiques possibles. Une estimation de la fraction de branchement de $\Xi_b^- \to \Xi^- J/\psi$ a ensuite été utilisée pour déterminer le ratio r_{BR} et les rendements de signaux attendus dans les trois différentes régions de q^2. Ces prédictions ont ensuite été utilisées pour tester l'ajustement simultané avec six projections, deux modes de désintégration et trois sous-ensembles de données, à l'aide de pseudo-expériences générées en fonction des taux prévus par le Modèle Standard. Enfin, la signification statistique attendue du signal de type SM a été évaluée et la procédure d'établissement des limites a été mise en place, préparant le démasquage des données. Les limites supérieures attendues sur r_{BR} selon l'hypothèse uniquement basée sur le bruit de fond sont fixées à 90 et 95% CL dans les trois régions de q^2 considérées dans cette analyse. Pour l'établissement des limites, le cadre GammaCombo [25, 26] est utilisé avec la méthode du niveau de confiance CL_s en appliquant une statistique de test unilatérale, où les valeurs p sont déterminées à l'aide de pseudo-expériences. Pour comparer les limites supérieures attendues aux prédictions du SM, la fraction de branchement attendue $\mathcal{B}(\Xi_b^- \to \Xi^-\mu^+\mu^-)$, discutée dans la section 6.2.2, est utilisée pour convertir les prédictions théoriques en prédictions sur r_{BR}. L'incertitude sur $\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi)$ est ignorée dans ce calcul pour le moment. Le résultat est montré dans la Fig. 3 et est compatible avec les attentes concernant la signification

du signal de la section précédente.

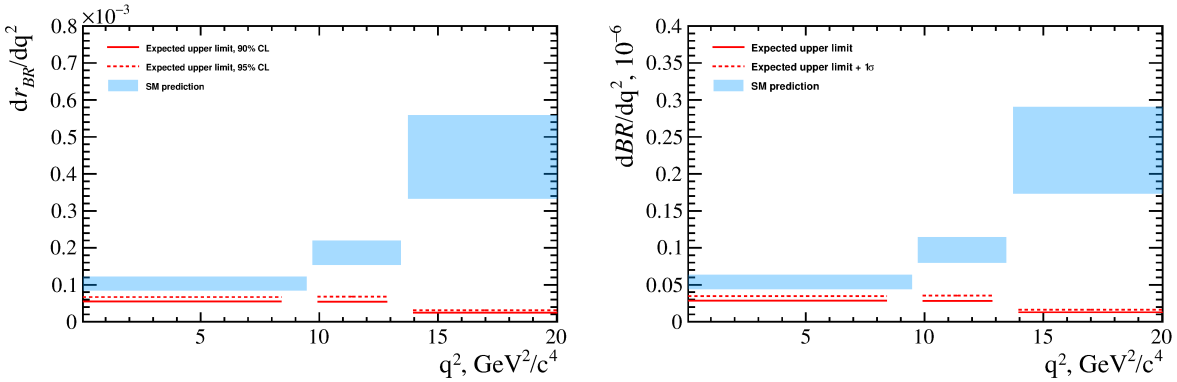


Figure 3: À gauche : Limite supérieure attendue à 90% (ligne rouge continue) et 95% (ligne rouge pointillée) CL sur dr_{BR}/dq^2, par rapport aux prédictions du SM de Ref. [27] (boîtes bleues). À droite : la même chose mais multipliée par notre estimation de $\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi)$, donnant $d\mathcal{B}/dq^2$. Seules les incertitudes statistiques sont prises en compte.

Ces limites supérieures soutiennent fortement une observation du désintégration $\Xi_b^- \to \Xi^-\mu^+\mu^-$ dans la région de q^2 la plus élevée. Les plus grands défis de ce processus résultent de la dépendance des efficacités à q^2 et de l'absence de prédictions des facteurs de forme pour la partie hadronique des modes de désintégration. Ce dernier point conduit à des échantillons de simulation de l'espace des phases qui ont été calibrés selon le modèle physique $\Lambda_b^0 \to \Lambda\mu^+\mu^-$, pour mieux tenir compte des angles de polarisation et de la distribution de q^2 du désintégration. La dépendance de l'efficacité à q^2 a été étudiée avec succès par la simulation et est incluse dans la procédure d'ajustement pour extraire la fraction de branchement.
L'analyse suivante décrit les premières étapes entreprises pour la recherche des modes de désintégration $\Xi_b^0 \to \Xi^0 J/\psi$ et $\Xi_b^0 \to \Xi^0\psi(2S)$. L'analyse est réalisée avec un ensemble de données partiel du Run 2, correspondant à $5.4\,\text{fb}^{-1}$. Le principal défi de cette analyse réside dans la combinaison de la faible section efficace de production du baryon Ξ_b^0 et la présence de particules neutres dans la chaîne de désintégration, notamment un π^0 dans l'état final. Comme le Ξ^0 se désintègre en particules neutres, son vertice de désintégration ne peut pas être reconstruit via des traces, donc le π^0 est supposé être issu du vertice primaire. Cela a plusieurs impacts : premièrement, le niveau de bruit de fond dû à la production abondante de mésons π^0 dans toutes sortes de processus est très élevé. Le deuxième impact est qu'il provoque un taux d'échec dépendant de la masse du fitter de l'arbre de désintégration, ce qui est crucial pour améliorer la précision des quantités reconstruites dans les désintégrations avec des particules neutres. Enfin, la présence d'un π^0 aggrave significativement la résolution en masse. Dans le travail présenté dans cette thèse, une première sélection basée sur des coupures a été développée et le mode de désintégration $\Xi_b^0 \to \Xi^0 J/\psi$ a été observé pour la première fois. Il est à noter qu'il s'agit peut-être de la première fois que le baryon Ξ^0 est reconstruit dans une expérience du LHC. Deux défis majeurs ont également été relevés : le développement d'une première stratégie de vérité (truthmatching) pour la simulation du signal, et la réduction significative du taux d'échec du DTF. Cette dernière a permis le développement du BDT pour la sélection, ce qui a

conduit à des indications fortes pour la première observation du mode de désintégration $\Xi_b^0 \to \Xi^0\psi(2S)$ également. Les ajustements de masse résultants pour les désintégrations $\Xi_b^0 \to \Xi^0 J/\psi$ et $\Xi_b^0 \to \Xi^0\psi(2S)$ utilisant la présélection et la sélection BDT sont montrés dans la Fig. 4.

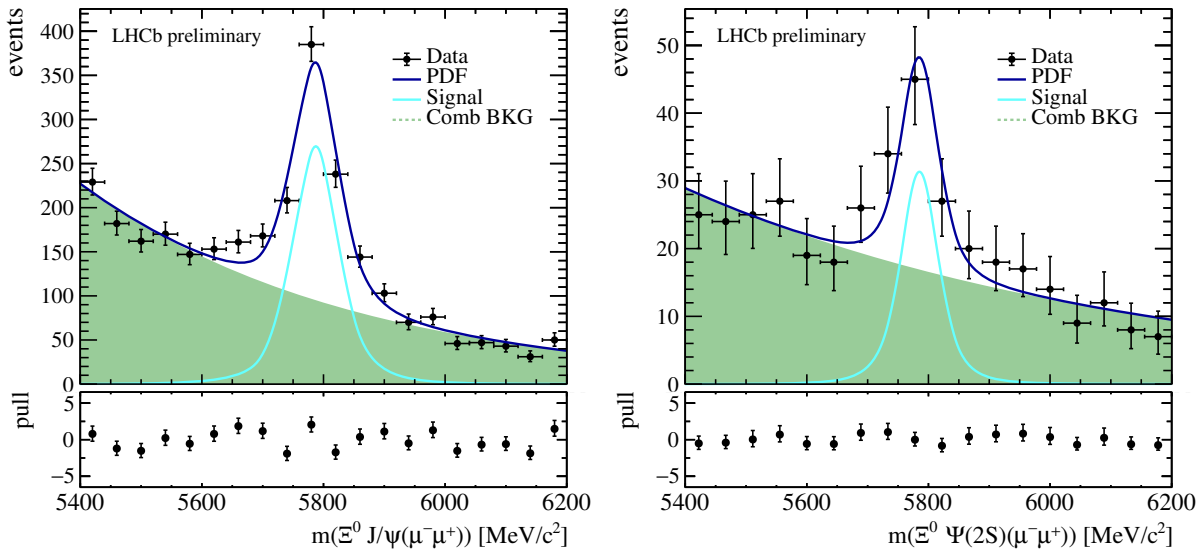


Figure 4: L'ensemble complet de données LHCb disponible (2016-18) est représenté par les points de données noirs. Ajustements de la masse invariante pour les désintégrations $\Xi_b^0 \to \Xi^0 J/\psi$ (à gauche) et $\Xi_b^0 \to \Xi^0\psi(2S)$ (à droite).

Cependant, il convient de noter que ces résultats sont seulement préliminaires et des études dédiées des défis introduits par la reconstruction du π^0 sont nécessaires. Parmi eux, des études de bruit de fond pour examiner la contribution des désintégrations partiellement reconstruites de Ξ_b^- combinées avec des mésons π^0 aléatoires, ainsi que l'étalonnage de la simulation pour garantir qu'aucun biais non pris en compte ne soit introduit. De plus, ces changements ouvrent la possibilité d'améliorer encore la sélection. Néanmoins, cette étude constitue un énorme pas en avant et prouve la capacité de reconstruire la désintégration $\Xi_b^0 \to \Xi^0\mu^+\mu^-$ à l'avenir avec des ensembles de données plus importants disponibles.
Dans le dernier chapitre d'analyse de cette thèse, l'état actuel de l'analyse angulaire du désintégration $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ utilisant l'ensemble complet des données des Run 1 et Run 2 est présenté. L'analyse est actuellement toujours masquée, mais toute la stratégie d'analyse a été évaluée afin de se préparer à la revue interne. L'analyse est réalisée dans le carré de la masse invariante du dilepton au-dessus de $q^2 > 15\,\mathrm{GeV}^2/c^4$, où la majorité du signal est attendue selon la mesure de la fraction de branchement différentielle du $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ [7]. L'analyse angulaire tente d'extraire pour la première fois les dix coefficients angulaires en utilisant un ajustement par maximum de vraisemblance. Un ensemble dédié de critères de sélection est développé afin d'obtenir des signaux propres du désintégration rare et de supprimer tous les modes de bruit de fond dominants. Cela comprend un classificateur multivarié dédié pour supprimer le bruit de fond constitué de combinaisons aléatoires de pistes. Un modèle d'ajustement de la masse a été développé, essentiel pour distinguer plus tard entre les événements de signal et de bruit de fond dans les composantes angulaires de l'ajustement par maximum de vraisemblance à quatre

dimensions. Des échantillons de simulation de l'espace des phases ont été utilisés pour extraire l'acceptation angulaire du détecteur LHCb, et un modèle d'ajustement a été développé pour extraire les coefficients angulaires. Ce dernier a été étudié en profondeur, notamment avec des pseudo-expériences, afin de garantir qu'aucun biais n'est introduit en raison du rendement de signal relativement faible par rapport à la complexité du modèle d'ajustement. En fait, un petit biais a été découvert, causé par la similarité de la distribution de $\cos(\theta_\Lambda)$ dans le signal et le bruit de fond due à la sélection des vrais hyperons Λ dans les deux cas, qui sera corrigé dans l'ajustement final des données. Un exemple aléatoire de pseudo-expérience réaliste avec les composantes de signal et de bruit de fond incluses peut être trouvé à la Fig. 5.

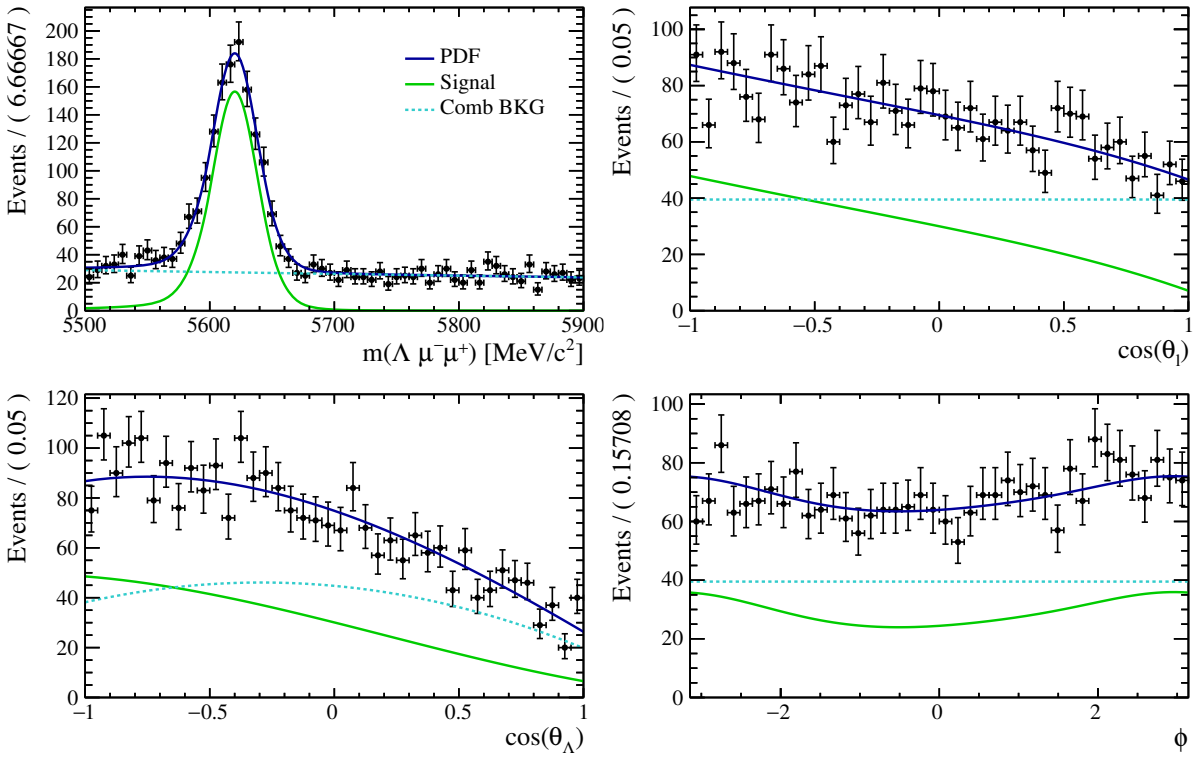


Figure 5: Un exemple aléatoire de pseudo-expérience réaliste utilisant les valeurs combinées du Run 1 et du Run 2 pour l'acceptation angulaire et les estimations de bruit de fond pour les formes angulaires. La distribution de masse invariante $m(\Lambda\mu^+\mu^-)$ est montrée en haut à gauche, celle de $\cos(\theta_\ell)$ en haut à droite, celle de $\cos(\theta_\Lambda)$ en bas à gauche et la distribution de ϕ en bas à droite.

Le point restant à finaliser est l'étalonnage de la simulation, car des différences résiduelles ont été découvertes. Ensuite, l'analyse complète sera mise à jour afin de commencer à effectuer les étapes de validation manquantes de la procédure d'ajustement reposant sur des données appropriées et à évaluer les incertitudes systémiques. Ces étapes sont actuellement effectuées en vue de la revue interne en attente.

Outre les mesures présentées dans cette thèse, j'ai également contribué à l'assemblage du détecteur SciFi et à une recherche de pannes lors de l'arrêt de fin d'année en 2023. Les

fibres optiques nécessaires pour une communication efficace et une lecture des données entre les cartes frontales SciFi et le centre de données ont dû subir une inspection optique approfondie pour s'assurer qu'aucune saleté ou rayure sur les fibres optiques ne réduise la qualité des données. Après avoir participé à l'inspection optique pendant le processus d'assemblage en 2021, j'ai pu appliquer les compétences acquises lors de l'enquête sur les pannes du détecteur SciFi en 2023. Pendant la collecte des données, plusieurs demi-ROBs et liens de données ont dû être exclus en raison d'instabilités réduisant la qualité des données. Lors de l'inspection optique, j'ai pu retracer l'instabilité à des émanations du VTRx bloquant le centre des fibres optiques. C'était le cas pour tous les demi-ROBs exclus et les liens de données, où aucune raison de l'instabilité n'avait été trouvée auparavant. Par conséquent, toutes les fibres optiques affectées et accessibles ont été nettoyées, conduisant à seulement deux liens de données exclus après l'intervention.

Introduction

The human curiosity and the idea of matter being composed of several different fundamental parts was first recorded about 600 B.C. [28]. Since then the belief that the atom (from the greek word *atemnein*, indivisible) is the smallest building block has been disproved. The first discovered fundamental particle was the electron [29], laying the ground for modern day particle physics.
The Standard Model (SM) is currently the best available theory aiming to describe all known fundamental particles and their interactions. It has been tested extensively and provides remarkable accuracy especially in the description of the electroweak interaction. However, it lacks, for example, an explanation for the existence of the three lepton and quark generations as well as their measured mass hierarchies. In combination with its inability to provide an explanation for certain cosmological observations like the existence of dark matter [30, 31], the SM is believed to only be an approximate and incomplete model of nature. Consequently, the field of particle physics is devoted to not only studying the SM in precision tests, but also directly searching for new particles at the energy frontier. Such new particles have been predicted in extensions to the SM, as well as in fundamentally different theoretical models. However, the reach of direct searches is limited by the reachable center-of-mass energies of accelerators. Therefore, indirect searches are an equally important tool to find answers to the open questions in the field of particle physics. As a direct consequence of the Heisenberg uncertainty relation, all possible interactions contribute to a process simultaneously. Therefore, heavy particles will contribute as virtual particles and influence the amplitude of a studied decay.
The goal of indirect searches is therefore to detect (tiny) deviations from the SM predictions in low-energy processes, due to effects of new fields (and consequently new interactions) by interpreting them with effective field theories. These theories allow extracting information on new physics beyond the Standard Model without reaching its natural energy scales to directly produce it. The same approach has been successfully used in the description of the nuclear beta-decay leading to the prediction, and the eventual discovery at collider experiments, of the $W^{\pm}$ boson [32]. Indirect searches can be performed either with large datasets of common processes enabling great precision and therefore sensitivity to even smallest deviations, or with processes that are heavily suppressed in the SM. In the latter case, even the smallest deviations have large impact on the measurements (in particular, due to interference between the new physics (NP) and SM processes) and can, therefore, be spotted.
One especially interesting environment to perform such indirect searches are flavour changing neutral currents (FCNC). They are transitions between different quark flavours with

the same electric charge and can only appear at loop level involving higher-mass particles. In the SM they are strongly suppressed by both the massive particles in the loop and the Glashow–Iliopoulos–Maiani (GIM) mechanism [1]. Therefore, the SM contribution of such transitions is very small and makes these processes especially sensitive to tiny deviations from NP. As mentioned before, NP can contribute in the loop, but the new contributions can also appear by producing FCNC at tree level.
One example of such suppressed FCNC are the $b \to s\ell^+\ell^-$ quark-level transitions. Due to the nature of the quarks, they can only be studied from experiments in quark bound states, called hadrons. In the past decade a number of curious deviations from the SM expectations have been observed in the $b \to s\ell^+\ell^-$ transitions by high-energy physics experiments [2–17]. The, in itself, consistent set of deviations, called B anomalies, can hint in two opposite directions. Either they emphasise a lack of understanding of effects caused by the strong interaction, which is also the cause of the quark bound states, or they are first hints for physics beyond the SM. The anomalies can be probed by analysing more available data and additionally including more hadronic decay modes based on the $b \to s\ell^+\ell^-$ transitions.
While ordinary matter consists of three-quark hadrons, called baryons, (and of electrons) current research focuses mostly on mesonic $b \to s\ell^+\ell^-$ transitions. This is partially driven by the fact that baryons cannot be produced at lepton colliders operated at the $\Upsilon(4S)$ resonance where most flavour-physics experiments are based. However, at the Large Hadron Collider (LHC) all types of b hadrons (mesons or baryons) can be produced abundantly. The study of baryonic $b \to s\ell^+\ell^-$ transitions is particularly interesting since they provide not only a complementary way to test the observed B-meson anomalies, but also give access to additional angular and polarisation observables. This is a result of their non-zero spin, which can also be utilised to test possible spin dependence of new physics. Since the hadronic environment at the hadron colliders is much busier than at lepton colliders, the experiments need to fulfil a dedicated set of requirements to perform precision measurements. These include a good background control and precise tracking, which the LHCb experiment provides and is therefore an ideal environment to study baryonic decays.
There are multiple baryon states that contain one b quark and two light (u, d, s) quarks. However, only four of them decay through the weak interaction, resulting in a lifetime of about 1.5 ps. These are: Λ_b^0 (valence quark content udb), Ξ_b^- (dsb), Ξ_b^0 (usb) and Ω_b^- (ssb). However, only the Λ_b^0 decays have been studied in reasonable detail [33]. For what concerns the Ξ_b and Ω_b^- baryons, less than ten decay modes have been observed for each of them, most of them using data from the LHCb experiment.
During the first two run periods of the LHC, the LHCb experiment has collected a large dataset of up to 10^{12} b and $\bar{b}$ quarks, out of which up to 25% hadronise into b baryons. Therefore, the collected data provide the ideal environment to study heavy-baryon properties and baryonic $b \to s\ell^+\ell^-$ transitions. The work presented in this thesis contributes to the study of baryonic decays.

This thesis starts with the most successful theory describing particle physics, the SM, and using it incompleteness to emphasise the importance of indirect and model-independent searches for NP. The current experimental knowledge about beauty baryons, in particular the four weakly-decaying ones important for this thesis, are discussed. After emphasising the importance of $b \to s\ell^+\ell^-$ transitions to test the SM and search for new physics, the

current experimental knowledge is discussed and finally a detailed introduction into the the angular distribution of the $\Lambda_b^0 \to \Lambda \ell^+ \ell^-$ decay is provided in Ch. 1. The following Ch. 2 describes the LHCb experiment, before the Upgrade of the LHCb detector for the Run 3 and 4 of the LHC is briefly described in the beginning of Ch. 3. The main part of the chapter is dedicated to the author's contributions to the new Scintillating Fibre (SciFi) detector. A tracking detector downstream of the dipol magnet of the LHCb experiment. Beside the author's contributions to the SciFi detector, I was lucky enough to work with an incredible team, that enabled me to contribute to several different analyses. Before describing each analysis in detail, Ch. 4 explains commonly used approaches. I performed, together with Dr. Vitalii Lisovskyi, a measurement providing significant improvement on the accuracy of the Ω_b^- mass. In the same analysis we measured the production fraction of Ω_b^- baryons relative to the Ξ_b^- baryons for the first time at the LHC. This analysis is described in Ch. 5 and the results of this analysis have been published in Ref. [19]. Beside our contributions to the improvements of knowledge about the b-baryon properties, we prepared an analysis for the first observation of the decay $\Xi_b^- \to \Xi^- \mu^+ \mu^-$, which is the first $b \to s \ell^+ \ell^-$ transition measured for a beauty baryon other than the Λ_b^0 baryon. In addition to this result, the measurement reports the observation of the $\Xi_b^- \to \Xi^- \psi(2S)$ decay. It has a much higher precision than the recently announced first observation of this decay mode by the CMS collaboration [18]. The analysis is presented in Ch. 6 and currently in the LHCb internal review process in preparation of the unblinding in the next months. Motivated by the Ξ_b^- measurements and the prospects of a possible isospin asymmetry measurement with the Ξ_b^0 baryon, Dr. Lisovskyi and I supervised M.Sc. Leandra Moeser performing the first search for the decay modes $\Xi_b^0 \to J/\psi \Xi^0$ and $\Xi_b^0 \to \psi(2S)\Xi^0$, which is described in Ch. 7. Finally, under the supervision of Dr. Marie-Hélène Schune and Prof. Dr. Johannes Albrecht, I performed an angular analysis of the $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ decay. The analysis is currently in preparation for the LHCb internal review process and described in detail in Ch. 8.
To summarise, this document reveals a general fascination for baryonic decays in the context of $b \to s \ell^+ \ell^-$ transitions and describes the contributions to the study of all four weakly-decaying beauty baryons.

Chapter 1

The theoretical framework and experimental status

This chapter covers the theoretical background and experimental status needed to study the work presented in this thesis. It includes a brief introduction to the Standard Model (SM) of particle physics (Sec. 1.1), its (baryon) bound states (Sec. 1.2), as well as a short introduction to (weak) effective field theories (Sec. 1.3) to describe the studied $b \to s\ell^+\ell^-$ transitions (Sec. 1.4). Finally, an overview of the current experimental status is given and the analyses performed during the duration of this PhD project are motivated.

1.1 The Standard Model of particle physics

The SM of particle physics is the most accurate theory describing all known fundamental particles and their interactions today. It is a relativistic quantum field theory that can be described completely by the Lagrangian density

$$\mathcal{L} = \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}, \tag{1.1}$$

which is gauge-invariant under local transformations of the $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry group. The fundamental particles can be grouped depending on their spin in fermions (half-integer spin, represented by fields ψ) and bosons (integer spin). The latter can be split in the gauge bosons with spin one, mediating the interactions, and the spinless Higgs boson.
The Quantum Chromodynamics (QCD) describes the strong interaction, which preserves the colour charge C due to its underlying symmetry $SU(3)_C$[1]. It is mediated by the eight massless gluons, g^a, coupling to the fermions carrying colour charge, called quarks [2]. Due to the non-abelian group structure of the QCD, the gluons are self-coupling and therefore

[1] There are three different colour charges, that are denoted as red (R), green (G) and blue (B).

[2] The other group of fermions is called leptons and will be discussed in more detail in the following. They are colour-neutral and do not interact through the strong interaction.

carry colour charge themselves. The QCD Lagrangian density function is given by

$$\mathcal{L}_{\mathrm{QCD}} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi + g_{\mathrm{s}}(\bar{\psi}\gamma^{\mu}T_a\psi)A_{\mu}^{a} - \frac{1}{4}G_{\mu\nu}^{a}G_{a}^{\mu\nu}. \tag{1.2}$$

with the field strength tensor of the gluon field A_{μ}^{a}

$$G_{\mu\nu}^{a} = \partial_{\mu}A_{\nu}^{a} - \partial_{\nu}^{a}A_{\mu}^{a} - g_{\mathrm{s}}f^{abc}A_{\mu,b}A_{\nu,c}. \tag{1.3}$$

where γ_{μ} presents the Dirac matrices, the coupling strength is defined by $g_s = \sqrt{4\pi\alpha_s}$ with the coupling of the strong interaction α_s, and f_{abc} the structure constants for different combinations of the colour charge configurations of the gluon a, b and c.
The remaining part of the SM symmetry group $SU(2)_L \times U(1)_Y$ relates to the non-abelian $SU(2)_L$ group preserving the left-hand chirality L, described by the weak isospin I_3 and the $U(1)_Y$ group that preserves the hypercharge Y, which differentiates between different fermions. The weak isospin is $1/2$ for the up-type quarks (u, c, t) and the electrical-neutral neutrinos (ν_i), while it is $-1/2$ for the down-type quarks (d, s, b) [3] and the charged leptons (e, μ, τ). The left-handed components of the fermion fields can be grouped in (three) isospin doublets with an up- and down-type quark $Q_L = (u_{i,L}, d_{i,L})^T$ and an electric uncharged and charged lepton $L_L = (\nu_{i,L}, e_{i,L})^T$, where i indicates the generation from the doublet. The right-handed components of the fermions have an isospin of $I_3 = 0$ and only exists for the up-, down-type quarks $(u_{i,R}, d_{i,R})$ and the charged leptons $(e_{i,R})$. Since neutrinos are massless particles in the SM, no right-handed isospin singlet exists for them. The $SU(2)_L$ symmetry group can be associated with three self-interacting massless gauge boson fields W_{μ}^{i} and the $U(1)_Y$ group with the gauge B_{μ} boson field, having the respective coupling constants g and g'. Together, the four gauge fields correspond to the electroweak symmetry,

$$\mathcal{L}_{\mathrm{EW}} = -\frac{1}{4}(W^{i\mu\nu}W_{\mu\nu}^{i} + B^{\mu\nu}B_{\mu\nu}) + \bar{\psi}_j i\gamma^{\mu}D_{\mu}\psi_j \tag{1.4}$$

with the left- and right-handed covariant derivative

$$D_{\mu,L} = \partial_{\mu} - ig'\frac{Y}{2}B_{\mu} - ig\tau^{k}W_{\mu}^{k}, \tag{1.5}$$

$$D_{\mu,R} = \partial_{\mu} - ig'\frac{Y}{2}B_{\mu}. \tag{1.6}$$

The $SU(2)_L \times U(1)_Y$ symmetry is spontaneously broken by the presence of the complex $SU(2)_L$ Higgs scalar doublet, $\phi = (\phi^{+}, \phi^{0})^T$, together with the Higgs scalar potential $V(\phi) = \mu^2\phi^{\dagger}\phi + \lambda(\phi^{\dagger}\phi)^2$. The parameters μ and λ are free parameters of the Higgs mass and four-vertex Higgs field coupling term. The spontaneous symmetry breaking (SSB) is caused by the non-vanishing vacuum expectation value of the Higgs field v. The remaining $U(1)_Q$ symmetry corresponds to the electromagnetic interaction, conserving the electric charge $Q = I_3 + \frac{1}{2}Y$, which is called the Gell-Mann-Nishijima relation [34,35]. The Higgs mechanism [36,37] can be described by the Higgs Lagrangian

$$\mathcal{L}_{\mathrm{Higgs}} = |D_{\mu,L/R}\,\phi|^2 - V(\phi). \tag{1.7}$$

[3] The six different quarks are also called flavours. Therefore, the study focusing on quarks is called flavour physics.

The introduction of the two complex Higgs scalar fields explains additionally the generation of boson masses in the SM. The SSB generates masses for the weak gauge bosons $W^{\pm}$, which are a linear combination $W^{\pm}_{\mu} = \frac{1}{\sqrt{2}}(W^1_{\mu} \mp iW^2_{\mu})$. In addition, the remaining vector fields W^3_{μ} and B_{μ} are mixed into the new massive vector field Z_{μ} (Z^0) and the massless photon vector field A_{μ} (γ):

$$Z_{\mu} = \cos\theta_{\mathrm{W}} W^3_{\mu} - \sin\theta_{\mathrm{W}} B_{\mu}, \tag{1.8}$$

$$A_{\mu} = \sin\theta_{\mathrm{W}} W^3_{\mu} + \cos\theta_{\mathrm{W}} B_{\mu}. \tag{1.9}$$

The Weinberg angle θ_W describing the mixing is defined by the initial electroweak coupling constants $\theta_W = \arctan(g'/g)$ and is related to the masses of the weak bosons $W^{\pm}$ and Z^0 with $\cos\theta_w = M_W/M_Z$. Glashow, Weinberg and Salam were the first to utilise the Higgs mechanism to prove the postulate electro-weak symmetry $SU(2)_L \times U(1)_Y$ [38–40]. While the masses of the three weak gauge bosons are degrees of freedom in the SM, that have been measured [33], they absorbed three degrees of freedom of the Higgs field. The last remaining degree of freedom is the massive Higgs boson H with a hypercharge of $Y = -1$. The Yukawa interaction

$$\mathcal{L}_{\mathrm{Yukawa}} = y^{\ell}_{ij}\left(\bar{\Psi}^{\ell i}_L \phi \Psi^{\ell j}_R\right) + y^{d}_{ij}\left(\bar{\Psi}^{qi}_L \phi \Psi^{dj}_R\right) + y^{u}_{ij}\left(\bar{\Psi}^{qi}_L \phi_c \Psi^{uj}_R\right) + h.c.\,. \tag{1.10}$$

couples left- and right-handed fermion fields with the Higgs field ϕ, where the indices i, j indicate the fermion generation. The three matrices y^f_{ij} of the Yukawa interaction (1.10) also lead to the fermion masses $m_f = (y^f_{ij} v)/\sqrt{2}$ for charged leptons, up- and down-type quarks. In the context of Yukawa couplings as the origin of fermion masses, it is important to consider their generation structure as 3×3 matrices. These matrices can be diagonalised by transforming the fermion fields to extract the Yukawa couplings introducing the fermion masses. However, there is no eigenbasis in which both quark mass matrices (up- and down-type) can be diagonalised simultaneously, which leads to the unitary *Cabibbo-Kobayashi-Maskawa* (CKM) matrix [41, 42]. It describes the transformation from the flavour eigenstates postulated by the SM (q) to the mass eigenstates (q'), which can be accessed in measurements. While neutrinos were initially considered massless in the SM, the observation of neutrino oscillations [43] can be described by the *Pontecorvo–Maki–Nakagawa–Sakata* (PNMS) matrix [44, 45] an equivalent to the CKM matrix for neutrinos. However, the weak interaction couples equally to the different lepton generations. This symmetry is called lepton flavour universality. The CKM matrix for the three quark generations is defined by:

$$\begin{aligned}\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} &= V_{\mathrm{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{i\delta_{13}} \\ -s_{23}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{13}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}\end{aligned} \tag{1.11}$$

where $s_{ij} = \sin(\theta_{ij})$ and $c_{ij} = \cos(\theta_{ij})$, θ_{ij} are the three CKM mixing angles and δ_{13} is the *CP*-violating phase. Together with the chiral structure of the weak interaction, the existence of the non-diagonal CKM matrix enables transitions between quark generations

transferred by the $W^{\pm}$ bosons, called charged current interactions. However, the small size of the off-diagonal elements imply the relative suppression of such transitions. It should be noted that the three CKM angles and the *CP*-violating phase are degrees of freedom in the SM. The same is true for the Yukawa couplings and therefore the fermion masses, leading to nine additional degrees of freedom. Together with the three gauge couplings, the vacuum expectation value of the Higgs potential and the Higgs mass, the SM has only 18 degrees of freedom. Consequently, the precise study of them will help to deepen the understanding of the universe and the fundamental particles that make it up. The discussed structure of the SM with the interactions of the respective fundamental particles is displayed in Fig. 1.1.

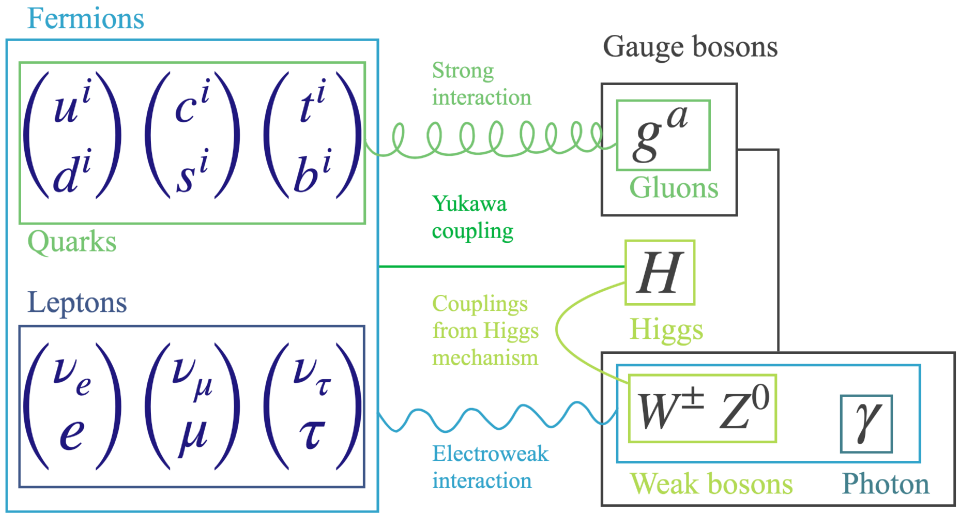


Figure 1.1: Graphical representation of the fundamental particles and their respective interactions with each other as described by the SM. Each of the quarks can have three possible colour charges that are indicated by the index i. Similarly, the index a highlights the eight possible colour charge configurations of the gluons mediating the strong interaction (QCD). The Yukawa coupling expresses the couplings of fermions to the Higgs field and how consequently the fermion masses are generated. The vector fields for the electroweak gauge bosons γ, $W^{\pm}$ and Z^0 are produced through the Higgs mechanism, which also generates their masses and couplings. Apation from Ref. [46].

The fundamental particles can be identified by their quantum numbers, like isospin, mass and electric charge. The current world averages of the quantities are listed in Ref. [33]. In addition to the twelve fermions, the Dirac equation for them leads to two equal energy solutions with different signs. Since only positive energies are physically meaningful, the existence of two solutions led to the postulation of antiparticles [47]. They are the counterpart for each SM fermion with inverted charge quantum numbers and equal mass. They almost immediately annihilate each other when produced together. It should be noted, that from all these fundamental particles only the up and down quark, the neutrinos and the electron are stable, meaning they have an infinite lifetime. All other heavier fermions decay through the described interactions into the stables fundamental particles.

Despite successful prediction of new particles, such as the Higgs boson [48, 49], and accurate description of interactions over the past decades, the SM of particle physics is known to be incomplete. This is in particular driven by its inability to explain certain phenomena. Among the most famous are astrophysical observations, such as that the theory surrounding the big bang [50] predict an initial equal production of matter and antimatter and therefore the need to a process that violates *CP* symmetry[4] in order to create a matter dominated universe. This theory is today expected to be the most accurate for the creation of the universe, after it successfully predicted the cosmic microwave background [51, 52]. However, the *CP*-violation in weak interaction of quarks is too small [33, 53] to explain the matter-antimatter-asymmetry. In addition, other observations such as the rotation curves of galaxies [31, 54] and the shape of gravitational lenses [30] imply a larger amount of matter than observed from measurements based on electromagnetic interactions. This implies the existence of *dark matter* which creates a gravitational pull, but does not interact through the electromagnetic interactions. To this day, no particle had been found that could be a candidate for dark matter. However, different theories predict different particles as possible explanations [55,56]. Furthermore, the observation of the accelerating expansion of the universe requires an extra source of energy, called *dark energy*. Another famous observation is the measurement of neutrino oscillation [43], which require neutrinos to carry mass contradicting the SM assumption of massless neutrinos. Therefore, it lacks an explanation of the origin of neutrinos masses. Beside these observations from astroparticle physics, the SM has no explanation for the mass hierarchies observed for the quark and lepton masses. The same is true for the origin of only three fermion generations or the left-handedness of the weak interaction.
Therefore, particle physics is still an evolving field searching for extensions of the current SM to provide the missing explanations or a more fundamental and complete theory. In particular, $b \to s\ell^+\ell^-$ transitions provide a clean probe of the SM that can help to answer open questions and deepen the understanding about particle physics. They will be discussed in more detail in Sec. 1.4.

1.2 Towards baryons – bound states of the SM

In nature, all fundamental particles creating matter (beside neutrinos) are only observed in bound states [5]. Reflecting upon the coupling constants of the individual interactions enables to understand the origin of these bound states. For the weak interaction, the large masses of the respective gauge bosons lead to a short range of the coupling constant g_W. Therefore, the weak interaction does not create specific bound states. The electromagnetic interaction has in principle an infinite range and is quantified by its coupling constant α_{EM}. It has a logarithmic dependence with increasing energy and therefore produces atoms as electromagnetic bound states. The last coupling constant α_s of the QCD is several orders of magnitudes stronger as the electromagnetic coupling constant α_{EM}. In addition, the coupling constant α_s has an asymptotic energy dependence[6]. At large energies the coupling constant becomes small, leading to colour-charged, elementary particles

[4] *CP* symmetry describes that the laws of physics are identical if a particle is exchanged with its antiparticle (charge (C) conjugation) and its spatial coordinates are inverted (parity (P) transformation).
[5] A detailed introduction and overview about possible configurations can be found in Ref. [47, 180].
[6] As a result QCD effects can only be calculated perturbatively for hard scattering processes.

being asymptotically free at short distances. But at low energies they almost immediately hadronise into colour-neutral bound states, so-called hadrons. This effect is referred to as colour-confinement. In particle physics hadrons are classified as mesons ($q\bar{q}$) and baryons (qqq) alongside with some evidence for tetraquarks ($qq\bar{q}\bar{q}$) [57, 58] and pentaquarks ($qqq\bar{q}\bar{q}$) [59, 60]. The only exception for QCD bound states of quarks is the top quark. Its high mass leads to a lifetime that is too short for the top quark to hadronise [33] before it decays. The number of baryons in each interaction needs to be preserved

$$\mathcal{B} = \frac{1}{3}(N_q - N_{\bar{q}}). \tag{1.12}$$

This accidental symmetry is referred to as baryon number conservation and crucial when studying decays with baryons.
Scattering experiments with hadrons proved [61], that the QCD bound states have a more complicated structure than only their initial quarks, called valence-quarks. They also include virtual quark pairs as well as virtual gluons, referred to as partons. These structures are studied at accelerators for protons and described with parton distribution functions. Regardless of the more complicated underlying structure, the quark-model [62] provides a sufficient theory to classify most hadrons by their quantum numbers only taking valence quarks into account.
In order to classify baryons, it is important to better understand their total wave functions:

$$\Psi = \Psi_{\text{Space}}\Psi_{\text{Colour}}\Psi_{\text{Spin}}\Psi_{\text{Flavour}}, \tag{1.13}$$

as it is directly linked to the allowed ground states. The wave function can be split into four sub-wave functions, the spatial Ψ_{Space}, the colour Ψ_{Colour}, the spin Ψ_{Spin} and the flavour wave function Ψ_{Flavour}. As baryons contain three quarks, they are fermionic particles and their total wave function needs to be antisymmetric under the exchange of two quarks due to the fermionic nature of the quarks. Baryons are classified with two angular momenta, one referring to two quarks and the second refers to the third one with respect to the combination of the first two. For ground states[7] both angular momenta L are zero, therefore the spatial wave function is symmetric due to the proportionality $\Psi_{\text{Space}} \propto (-1)^L$.
As a consequence of the exactness of the colour symmetry $SU(3)_C$ and the observation of colour confinement, it follows that every quark in a baryon needs to have a different colour in order to create a colour-singlet. The corresponding wave function is given by

$$\Psi_{\text{Colour}} = \frac{1}{\sqrt{6}}(RGB - RBG + BRG - BGR + GBR - GRB), \tag{1.14}$$

which is antisymmetric under the exchange of two arbitrary quarks.
Therefore, the remaining parts of the total wave function $\Psi_{\text{Spin}}\Psi_{\text{Flavour}}$, need to be symmetric. Each quark can occur in two spin configurations, spin up $\uparrow$ and spin down $\downarrow$. It follows that a baryon has symmetric spin wave function Ψ_{Spin} for the quadruplet of spin states with $j = \frac{3}{2}$. For spin states with $j = \frac{1}{2}$ a mixed symmetry of (anti-)symmetric and undefined symmetry parts is found. As a result, it is important to have a look at

[7] Since only ground states have been studied as part of this PhD project, excited states will not be discussed in this section.

the combination $\Psi_{\text{Spin}}\Psi_{\text{Flavour}}$ of the spin and flavour wave functions. Historically, the flavour symmetry was based on the isospin symmetry ($m(u) \approx m(d)$) corresponding to an $SU(2)$ symmetry. With the discovery of N further quark flavours this symmetry can be extended to $SU(N+2)$. But this symmetry is strongly broken due to the hierarchical mass structure of quarks, nevertheless it can be used to group hadrons in a first simple way.
When discussing baryons, there are three different cases that need to be discussed, the baryon containing only one quark flavour, two or three different ones. The first case is the easiest as the flavour wave function Ψ_{Flavour} is fully symmetry under the exchange of two quarks, the spin wave function Ψ_{Spin} also needs to be symmetric. Therefore, only ground states with spin $j = 3/2$ are possible. One example is the baryon $\Omega^- = |sss\rangle$. In the case of two different quark flavours, one gets again a case with states that are symmetric under the exchange of two quarks and some that have a mixed symmetry. When combining the symmetric flavour and spin wave function states, one obtains a ground state with spin 3/2, that is neither orbital nor radially excited. In addition, the combination of the mixed symmetry states allows to construct a spin 1/2 baryon ground state, that has again a symmetric wave function under the exchange of two quarks. The wave function is then given by

$$\begin{aligned}\frac{1}{\sqrt{18}}[|qqq^{'}\rangle(\uparrow\downarrow\uparrow + \downarrow\uparrow\uparrow - 2\uparrow\uparrow\downarrow) + |qq^{'}q\rangle(\uparrow\uparrow\downarrow + \downarrow\uparrow\uparrow - 2\uparrow\downarrow\uparrow) \\ + |q^{'}qq\rangle(\uparrow\downarrow\uparrow + \uparrow\uparrow\downarrow - 2\downarrow\uparrow\uparrow)].\end{aligned} \tag{1.15}$$

An example are the $|bss\rangle$ baryons, the Ω_b^- corresponding to spin $j = \frac{1}{2}$ state and Ω_b^{*-} to $j = \frac{3}{2}$. Under the same assumption, it can be concluded that there are three different ground states in the case of three different quark flavours. Two states have a spin 1/2 and one a 3/2. One example for the case of the $|udb\rangle$ states are the Λ_b^0 and Σ_b^0 baryon with spin 1/2 and Σ_b^{*0} with spin 3/2.
Initially Gell-Mann proposed the quark-model only containing the u, d and s quark in 1964. By making use of the approximate $SU(3)$ symmetry, baryons can be grouped by the electric charge, strangeness and projection of the isospin in multiplets distinguished by the spin. When extending this symmetry by including the b quark to a (strongly broken) $SU(4)$ symmetry, the representations of all possible baryon ground states in Fig. 1.2 is the result.

1.2.1 Beauty baryons

There are fifteen beauty baryons, that contain only one heavy quark as displayed in the $1b$ plane of Fig. 1.2. Out of these fifteen b baryons, four are particular interesting as they are weakly decaying. The lightest one of them is the $\Lambda_b^0 = |udb\rangle$ baryon, which is the most abundantly produced one of the four weakly decaying b baryons and therefore the best studied one. Then there is the Ξ_b isospin doublet, $\Xi_b^0 = |usb\rangle$ and $\Xi_b^- = |dsb\rangle$, and lastly the heaviest one, the $\Omega_b^- = |ssb\rangle$. Since all four baryons are weakly decaying, they have a typical lifetime in the order of picoseconds. All other of the fifteen, neither radial nor orbital excited b baryon states can in principle decay strongly or electromagnetically into one of the four weakly decaying states. This is important as it can lead to backgrounds,

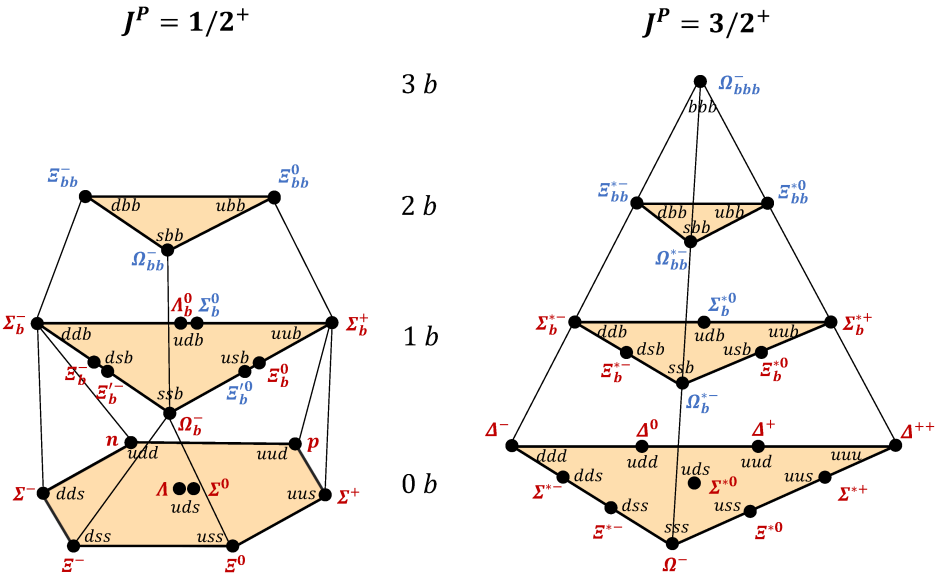


Figure 1.2: $SU(4)$ multiplets of baryons made of u, d, s and b quarks. Left: the multiplet with an $SU(3)$ octet on the lowest level. Right: the multiplet with an $SU(3)$ decuplet on the lowest level. Observed states are shown in red, while predicted unobserved states in blue. Taken from Ref. [63].

when studying the weak decays. However, due to the higher masses, they are less abundantly produced and some even have not yet been observed.
An overview of the current experimental status of the fifteen beauty baryons is given in Tab. 1.1.

Table 1.1: Properties of the lightest b-baryon states. Based on Ref. [33].

Name	Quark content	Mass, MeV/c^2	Lifetime or width	J^P	(I,I_3)	S
Λ_b^0	udb	5619.60 ± 0.17	1.471 ± 0.009 ps	$1/2^+$	$(0,0)$	0
Ξ_b^0	usb	5791.9 ± 0.5	1.480 ± 0.030 ps	$1/2^+$	$(1/2,1/2)$	-1
Ξ_b^-	dsb	5797.0 ± 0.6	1.572 ± 0.040 ps	$1/2^+$	$(1/2,-1/2)$	-1
Σ_b^+	uub	5810.56 ± 0.25	5.0 ± 0.5 MeV/c^2	$1/2^+$	$(1,1)$	0
Σ_b^0	udb	not yet observed		$1/2^+$	$(1,0)$	0
Σ_b^-	ddb	5815.64 ± 0.27	5.3 ± 0.5 MeV/c^2	$1/2^+$	$(1,-1)$	0
$\Xi_b'^0$	usb	not yet observed		$1/2^+$	$(1/2,1/2)$	-1
$\Xi_b'^-$	dsb	5935.02 ± 0.05	< 0.08 MeV/c^2	$1/2^+$	$(1/2,-1/2)$	-1
Ω_b^-	ssb	6046.1 ± 1.7	1.64 ± 0.18 ps	$1/2^+$	$(0,0)$	-2
Σ_b^{*+}	uub	5830.32 ± 0.27	9.4 ± 0.5 MeV/c^2	$3/2^+$	$(1,1)$	0
Σ_b^{*0}	udb	not yet observed		$3/2^+$	$(1,0)$	0
Σ_b^{*-}	ddb	5834.74 ± 0.30	10.4 ± 0.8 MeV/c^2	$3/2^+$	$(1,-1)$	0
Ξ_b^{*0}	usb	5952.3 ± 0.6	0.90 ± 0.18 MeV/c^2	$3/2^+$	$(1/2,1/2)$	-1
Ξ_b^{*-}	dsb	5955.33 ± 0.13	1.65 ± 0.33 MeV/c^2	$3/2^+$	$(1/2,-1/2)$	-1
Ω_b^{*-}	ssb	not yet observed		$3/2^+$	$(0,0)$	-2

In view of the work presented in this thesis, it is worthwhile to take a closer look at the current knowledge about the four weakly decaying beauty baryons.
The lightest and most frequently produced b baryon, the Λ_b^0, was discovered in 1991 by the UA1 experiment through the $\Lambda_b^0 \to \Lambda J/\psi$ decay [64]. Since the Λ_b^0 is so light, the SPS accelerator at CERN provided enough energy to produce the baryon. In the following decade the properties and many different decay modes of the Λ_b^0 have been studied at facilities including the LEP experiments, CDF at Fermilab as well as the LHCb experiment at the LHC at CERN. The full list of the currently 74 observed decay modes and measured properties can be found in the PDG [33]. Especially the large production fraction of Λ_b^0 baryons at the LHC has contributed to the extensive studies performed at the LHCb experiment. It also opened the possibility to study baryonic $b \to s\ell^+\ell^-$ transitions for the first time. More details about these will be given in Sec. 1.4.2.
The next b baryon relevant for this work, the Ξ_b^-, was directly observed for the first time in 2007 by the CDF collaboration [65], in the channel $\Xi_b^- \to \Xi^- J/\psi$. Since then, LHCb measured its mass, lifetime and relative production rate exploiting the same decay channel [66, 67]. In addition [20, 68, 69], it was also studied via $\Xi_b^- \to \Xi_c^{\,0}\pi^-$, $\Xi_b^- \to J/\psi\Lambda K^-$ and charmless $\Xi_b^- \to pK^-K^-$ decay modes. The latter decay was also used to perform the first search for *CP* violation in Ξ_b^- decays [70] in 2021. The measurement lead to the observation of the decay modes $\Sigma\ (1385)K^-$, $\Lambda\ (1405)K^-$, $\Lambda\ (1520)K^-$, $\Lambda\ (1670)K^-$, $\Sigma(1775)K^-$ and $\Sigma\ (1915)K^-$. The latest search was for the decay $\Xi_b^- \to \Xi^-\gamma$ [71]. So, with only a handful of decay channels established, Ξ_b^- is much less well studied as the Λ_b^0 baryon. However, with its decay into charged particles it is an excellent candidate to search for another baryonic $b \to s\ell^+\ell^-$ transition with the decay $\Xi_b^- \to \Xi^-\mu^+\mu^-$.
The isospin doublet partner of the Ξ_b^-, the Ξ_b^0, was observed for the first time by the CDF collaboration in 2011 using the decay $\Xi_b^0 \to \Lambda_c^+\pi^-$ [72] and in 2014 the same channel was used for a more precise determination of the Ξ_b^0 mass by the LHCb collaboration [73]. In addition, they confirmed the observation in 2013 using the decay channels $\Xi_b^0 \to pD^0\pi^-$ and $\Xi_b^0 \to \Lambda_c^+K^-$ [74]. Afterwards, the charmless four-body decay modes $\Xi_b^0 \to pK^-K^+K^-$, $\Xi_b^0 \to pK^-\pi^+\pi^-$ and $\Xi_b^0 \to pK^-\pi^+K^-$ were observed as well [75], and the latter two were exploited to search for *CP* violation [76]. Finally, first evidence for the decay modes $\Xi_b^0 \to J/\psi\Xi^0$ and $\Xi_b^0 \to J/\psi\Lambda$ are reported [77], but no limit was set. Motivated by this first evidence and the importance of the search for the rare decay mode $\Xi_b^0 \to \Xi^0\mu^+\mu^-$, the first dedicated search for the decays $\Xi_b^0 \to J/\psi\Xi^0$ and $\Xi_b^0 \to \psi(2S)\Xi^0$ was performed as part of this thesis. The details will be discussed in Ch. 7.
Even less knowledge is available for the Ω_b^- baryon, the last of the four weakly decaying beauty baryons. In 2008, the D0 collaboration claimed a discovery of the Ω_b^- candidate in the $J/\psi\Omega^-$ channel [21] with the mass of about 6165 MeV/c^2, being significantly heavier than theoretically expected. This was followed by the CDF observation of the Ω_b^- candidate in the same channel [78], but with the mass about 6054 MeV/c^2, much closer to the theoretical expectations and inconsistent with the D0 result. CDF has provided not only the mass and lifetime measurements, but also measured the relative production rate of the Ω_b^-, Ξ_b^- and Λ_b^0 baryons. The situation was clarified by LHCb in 2013 [67] with the confirmation of the state observed by the CDF collaboration using the same decay mode. It is therefore assumed that the D0 result likely was a statistical fluctuation. In 2014, CDF has observed the second decay mode, $\Omega_b^- \to \Omega_c^0\pi^-$ [79], confirmed by LHCb in 2016 [80]. Both collaborations have measured the mass and lifetime, but in both cases

the used dataset had very low yield of a few tens of signal events. In addition, the LHCb collaboration set limits on the charmless three-body decay modes $\Omega_b^- \to pK^-K^-K^-$, $\Omega_b^- \to p\pi^-\pi^-$ and $\Omega_b^- \to pK^-K^-$ in 2016 [69], which were updated in 2021 [70]. Recently, the LHCb collaboration also observed the decay mode $\Omega_b^- \to \Xi_c^+ K^-\pi^-$ [81].

1.2.2 Properties of beauty baryons

When discussing beauty baryons, it is important to notice that even for the lightest one, the Λ_b^0, it is not possible to produce a pair of them when operating electron colliders at the $\Upsilon(4S)$ resonance. Therefore, the current knowledge about b baryons is mostly driven by results obtained at the hadron colliders Tevatron and LHC. As a consequence, it is necessary to precisely know the production fraction of a certain type of b baryon f_i from the initial overall produced $b\bar{b}$ pairs at hadron colliders. This quantity is production environment dependent and needed to measure absolute branching fractions of decays.
The production probability depends on the mass of a hadron and therefore the number of heavy quarks. Here, it is sufficient to only consider b mesons and baryons with only one b quark. The production of baryons containing also a c quark in addition to the b quark is much smaller than 0.1 % and therefore, not considered in the following discussion. Since the direct production of the weakly decaying beauty baryons can not be distinguished from the production through the electromagnetic or strong decay of a heavier state, only the total production fractions of the $\Lambda_b^0, \Xi_b^0, \Xi_b^-$ and Ω_b^- baryons are taken into account. The sum over all production fractions needs to sum up to unity, giving

$$f_u + f_d + f_s + f_c + f_{\Lambda_b^0} + f_{\Xi_b^0} + f_{\Xi_b^-} + f_{\Omega_b^-} = 1. \tag{1.16}$$

In this notation f_i [8] denotes for the corresponding production fraction. The first four fractions give the probability of b quark to hadronise with a respective quark i into the specific B meson. Therefore, f_u is the hadronisation probability into a B^+ meson, f_d into a B^0 meson, f_s into a B_s^0 meson and f_c into a B_c^+ meson.
The production fraction of the lightest b-baryon, the Λ_b^0, was first measured in 2014 [82]. While the production fraction increases with the transverse momentum, for the calculations in this section, the value obtained when averaging over the kinematics is sufficient. The most recent update by the LHCb collaboration [83] measured the relative production fraction of the Λ_b^0 to:

$$\frac{f_{\Lambda_b^0}}{f_u + f_d} = 0.259 \pm 0.018 \tag{1.17}$$

at a center-of-mass energy of 13 TeV. In addition, the LHCb collaboration also measured the Ξ_b^- production fraction with respect to the Λ_b^0 baryon with the same Run 2 dataset:

$$\frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} = (8.2 \pm 0.7 \pm 0.6 \pm 2.5) \times 10^{-2}. \tag{1.18}$$

Similarly to the production fraction of the Λ_b^0 baryon, the LHCb collaboration also measured the production fractions f_s of the B_s^0 [83] and f_c of the B_c^+ meson [84] relatively to

[8] Note that the following, measured values are limited to the decays that are inside the acceptance of LHCb experiment.

the B^+ and B^0 mesons:

$$\frac{f_s}{f_u + f_d} = 0.122 \pm 0.006, \tag{1.19}$$

$$\frac{f_c}{f_u + f_d} = (2.61 \pm 0.03 \pm 0.062 \pm 0.06) \times 10^{-3} \tag{1.20}$$

at a center-of-mass energy of 13 TeV. The production fractions of the Ξ_b^0 and Ω_b^- baryon have not been measured yet, but can be estimated. For the Ξ_b^0 the isospin symmetry with respect to its doublet partner Ξ_b^- can be utilised to assume $f_{\Xi_b^0} \approx f_{\Xi_b^-}$. As previously discussed, for a baryon containing three different quark flavours, there are two spin 1/2 ground states with different symmetry configurations in the wave function. Therefore, the isospin doublet Ξ_b^0 and Ξ_b^- is paired with another doublet denoted as $\Xi_b^{0'}$ and $\Xi_b^{-'}$. Both particles could, in principle, decay electromagnetically into $\Xi_b^{0/-}\gamma$ or strongly into $\Xi_b^{-/0}\pi^{0/-}$. However, the not yet observed $\Xi_b^{0'}$ baryon is expected to only decay electromagnetically [85, 86], which breaks the isospin symmetry. Ref. [87, 88] gives a detailed estimation on the relative Ξ_b^0 production fractions by accounting for all known or predicted low-lying Ξ_b states and the phase-space factors in their decays to the ground states. The obtained relative production fraction of the Ξ_b^0 baryon is $f_{\Xi_b^0}/f_{\Xi_b^-} = 1.37 \pm 0.09$. Finally, for the Ω_b^- production fraction, there is no measurements at LHC. One estimate which is used in Ref. [89] is that the Ω_b^- contribution is about 15% of the Ξ_b (sum of the two isospin partners). This assumption is merely an educated guess and has never been tested at the LHC [9]. Combining the information from the measurements, neglecting the uncertainties, and including the discussed assumptions finally enables to determine $f_u + f_d$ from Eqn.(1.16)

$$(f_u + f_d) \times (1 + 0.122 + 0.0026 + 0.259 \times (1 + 0.082 \times (1 + 1.37) \times 1.15)) = 1, \tag{1.21}$$

where the whole term multiplied by 0.259 represents the contributions of b baryons – Λ_b^0, Ξ_b^-, Ξ_b^0 and Ω_b^-. This equation (1.21) leads to

$$f_u + f_d = 0.69 \tag{1.22}$$

and assuming isospin symmetry for the B mesons, the respective B^+ and B^0 production fractions are

$$f_u = f_d = 0.35. \tag{1.23}$$

This finally leads to the production fractions of all weakly decaying b hadrons in Tab. 1.2.

Table 1.2: Estimation of the production fractions of the weakly-decaying b hadrons at the LHC. Entries marked with $*$ are based on assumptions rather than measurements.

Fraction	f_u	f_d	f_s	f_c	$f_{\Lambda_b^0}$	$f_{\Xi_b^-}$	$f_{\Xi_b^0}$	$f_{\Omega_b^-}$
Value	0.35	0.35	0.085	0.002	0.18	0.015	0.020*	0.005*

[9] For simplicity, the measurement of the relative production fraction of the Ω_b^- baryon is not included in this discussion. It will be elaborated later in Ch. 5.

It should be noted that these assumptions do not take double and triple-heavy baryons into account, as well as bottomonia $b\bar{b}$. The contribution from bottomonia is challenging to assess due to the necessity of accounting for states unable to decay to b hadrons, i.e., those located below the $B\bar{B}$ threshold. Many of these states have high decay widths to the lowest-lying bottomonium states such as $\Upsilon(1S)$ or η_b, but also a wide variety of electromagnetic or strong decays to light particles, introducing a risk of double-counting. Production cross-sections have only been measured at the LHC for the Υ family, but not for the η_b or χ_b or h_b states. For instance, the LHCb collaboration has measured $\mathcal{B}(\Upsilon(1S) \to \mu^+\mu^-) \times \sigma(\Upsilon(1S)) = 4687 \pm 10 \pm 294\,\text{pb}$ [90], leading to a cross-section of about $0.2\,\mu\text{b}$, when using the know $\mathcal{B}(\Upsilon(1S))$ [33]. This can be compared to the B^+ production cross-section of about $86.6\,\mu\text{b}$ [91]. Consequently, the total bottomonia production fraction, $f_{b\bar{b}}$, can be up to about 0.5%. However, the current uncertainties on the measured b-hadron production fractions allow to neglect this contribution.
To improve the knowledge about b-hadron production fractions, measurement of the ones, $f_{\Xi_b^0}$ and $f_{\Omega_b^-}$, for the Ξ_b^0 and Ω_b^- baryon are needed. Therefore, as part of this thesis, the relative production fraction of the Ω_b^- baryon was measured for the first time at the LHC at a center-of-mass energy of 13 TeV. This is a crucial step towards enabling absolute branching fraction measurements of Ω_b^- decays, but also to cross-check the approximated production fraction of 0.5 %. In addition, this value is included in the calculation of the $b\bar{b}$ cross-section and therefore can influence all measurements involving b quarks. The measurement is described in more detail in Ch. 5.
While the quark-model provides an adequate way to predict hadrons, their most fundamental properties, their masses, need to include all QCD effects that are especially driving the binding energies:

$$m_{\text{hadr}} = \sum_i m^i_{\text{quark}} + E_{\text{bind}} + \sum \text{gluons} + \sum q\bar{q}. \tag{1.24}$$

Due to the non-perturbative nature of QCD, the theoretical prediction of hadrons masses is rather challenging. This is especially true for baryons due to their three-body nature when compared to mesons. There are different approaches for SM predictions like light-cone sum rule [92] or lattice QCD [93], that work with different assumptions. Therefore, the precise measurements of hadron masses, especially those of baryons, can help to exclude some approaches or contribute to refine them. This is also especially important in the context of predicting branching fractions as the same theories are used to predict form factors.
The described importance of mass measurements, together with the limited knowledge of the Ω_b^- baryon motivated to perform a mass measurement as part of this thesis. The measurement itself is described in more detail in Ch. 5.

1.3 (Weak) Effective field theory

Effective Field Theories (EFTs) are a theoretical framework, that enables to describe high energy NP phenomena and study their effect on SM precision observables without a priori knowledge of the underlying ultra violet theory [94,95]. This is possible, since EFTs treat several energy scales at the same time by fully separating them into the different scales. Therefore, they are completely model independent, making them a great tool for indirect

searches for NP.
A concrete example is the weak effective field theory used to describe the phenomenology of weak b decays. Due to the energy-scale separation, all degrees of freedom above the mass of the beauty quark m_b are integrated out. This is possible since the energy transfer of the dedicated particles, the $W^{\pm}, Z^0$ boson and the top quark, is much smaller compared to their masses $m_{W^{\pm}}, m_{Z^0}$ and m_t. As a result any weak transition of a b hadron can be described as a point-like four fermion transition at quark-level, which can be seen in the feynman diagrams in Fig. 1.3.

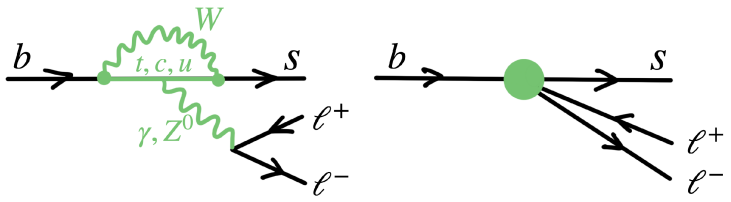


Figure 1.3: Example of the "integrating out" of the heavy degrees freedom for an FCNC process. The loop process indicated in green (left) turns into a point-like interaction (right) for the process.

The separation of the multiple-scale problems into the different scales, namely the (high-energy, or "short-distance") electroweak one and the (low-energy, or "long-distance") one of QCD effects, can be described simultaneously with Effective Hamiltonian formalism [96, 97]. The decay of a certain particle P into a final state F can therefore be fully described by the decay amplitude

$$A(P \to F) = \langle F|\mathcal{H}_{eff}|P\rangle \tag{1.25}$$

where H_{eff} is the relevant effective Hamiltonian. The most general effective Hamiltonian for b hadron decays can be written as:

$$H_{eff} \approx -\frac{4G_{\mathrm{F}}}{\sqrt{2}} V_{CKM} \sum_{i=1} \left(C_i(\mu)\mathcal{O}_i(\mu) + C_i'(\mu)\mathcal{O}_i'(\mu)\right). \tag{1.26}$$

In this equation G_{F} is the Fermi constant and V_{CKM} is the set of relevant CKM factors describing the quark transition of interest. The chosen energy scale $\mu = \Lambda_{\mathrm{QCD}}$ enables to describe electroweak and QCD effects at the same time. The left-handed $\mathcal{O}_i$ and right-handed $\mathcal{O}_i'$ [10] operators describe the low-energy physics (QCD) below the scale μ. The matrix elements of the QCD effects cannot be calculated perturbatively, but there are several methods allowing for predication of the from factors describing the non-perturbative QCD effects. On the other hand, the so-called Wilson coefficients $C_i(C_i')$ characterise the high-energy part, which is sensitive to NP effects and can be calculated fully perturbatively. NP can manifest itself as shifts to the Wilson coefficients or enhancements of new operators which are either absent or strongly suppressed in the SM. This procedure based on the factorisation of the short- and long-distance dynamics and is also referred to as Operator Product Expansion.

[10] Due to the chiral structure of the weak interaction, the right-handed operators are suppressed.

1.4 The magic of $b \to s\ell^+\ell^-$ transitions

As mentioned before, the suppression of $b \to s\ell^+\ell^-$ transitions in the SM makes them a particular powerful tool to search indirectly for NP effects. However, due to the colour-confinement, the transitions can only be studied in hadronised states. At hadron level, these quark transitions can be described by hadronic matrix elements, which can be parameterised in terms of form factors [11] multiplied by their respective independent Lorentz structures.
As described by the Heisenberg uncertainty principle, all possible decay modes contribute to a transition simultaneously. Therefore, the effective weak Hamiltonian is used to describe $b \to s\ell^+\ell^-$ transitions, includes both the SM and unknown NP contributions. Coming back to the effective Hamiltonian (1.26), ten operators are sufficient to fully describe the electroweak and QCD effects of the SM for $b \to s$ transitions. There are six hadronic four-quark operators describing $b \to sq\bar{q}$ transitions. They include the two $\mathcal{O}_{1-2}$ current-current tree-level $W^\pm$ exchange operators and $\mathcal{O}_{3-6}$ which are QCD penguins mediated by gluons. The remaining operators $\mathcal{O}_{7-10}$ are electroweak penguin operators, that can be distinguished into the electro- and chromo-magnetic penguins $\mathcal{O}_{7-8}$

$$\mathcal{O}_{7^{(\prime)}} = \frac{m_b}{e}[\bar{s}\sigma^{\mu\nu}P_{R(L)}b]F_{\mu\nu} \qquad \mathcal{O}_8^g = \frac{m_b}{g}[\bar{s}\sigma^{\mu\nu}T^aP_{R(L)}b]G^a_{\mu\nu} \tag{1.27}$$

and the two semileptonic operators $\mathcal{O}_{9-10}$

$$\mathcal{O}_9 = [\bar{s}\gamma^\mu P_L b][\bar{l}\gamma_\mu l] \qquad \mathcal{O}_{10} = [\bar{s}\gamma^\mu P_L b][\bar{l}\gamma_\mu\gamma_5 l] \tag{1.28}$$

with $P_{L,R} = 1/2(1 \mp \gamma_5)$ being the projectors for left-(right-)handed chirality.
When focusing on $b \to s\ell^+\ell^-$ transitions, it is sufficient to only consider the radiative operator $\mathcal{O}_7$ describing the $b \to s\gamma$ transition with an on-shell photon then decaying into the lepton pair and the vectorial and axial semileptonic operators $\mathcal{O}_9$ and $\mathcal{O}_{10}$. An example Feynman diagram for all three operators can be found in Fig. 1.4.

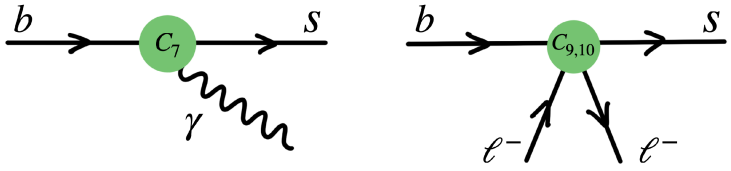


Figure 1.4: Sketch of the radiative dipole operator $\mathcal{O}_7$ (left) and the electroweak penguin operators $\mathcal{O}_{9,10}$ (right) in green.

It is necessary to examine the possible quarks in the loop of the $b \to s\ell^+\ell^-$ transition, in order to determine the V_{CKM} factor of the effective Hamiltonian. The CKM factor $V_{ub}V_{us}^*$ is only taken into account as a higher order correction, since the b to u quark transition is Cabibbo-suppressed. The remaining two factors $V_{tb}V_{ts}^*$ and $V_{cb}V_{cs}^*$ are fundamentally different, since processes involving the top quark are above the EFT energy scale μ and

[11] It should be noted, that the prediction of form factors is influenced by QCD effects between the quark transition of interest and the spectator quark. Since baryons have two spectator quarks, their form factor predictions are expected to be more challenging.

the ones with the charm quark below it. As a consequence the unitary relations of the CKM matrix are used to express $V_{cb}V_{cs}^*$ via the remaining CKM factor $V_{tb}V_{ts}^*$. Therefore, $V_{tb}V_{ts}^*$ is the only CKM factor needed to describe $b \to s\ell^+\ell^-$ transitions and the effective weak SM Hamiltonian can be written as followed:

$$\mathcal{H}_{eff} = -\frac{4G_F}{\sqrt{2}} V_{tb}V_{ts}^* \frac{\alpha_e}{4\pi} \sum_{i=7,9,10} \mathcal{C}_i(\mu)\mathcal{O}_i = N_1 \sum_{i=7,9,10} \mathcal{C}_i(\mu)\mathcal{O}_i. \tag{1.29}$$

Potential NP effects manifest in the modification of the Wilson coefficients of the SM operators and by feeding into new effective dimension-six operators:

$$\mathcal{H}_{eff}|^{\text{NP, naive}} = N_1 \sum_{i=9',10',S^{(\prime)},P^{(\prime)},T,T5} \mathcal{C}_i(\mu)\mathcal{O}_i. \tag{1.30}$$

These include the chirality-flipped counterparts of the semileptonic operators $\mathcal{O}_{9'}$ and $\mathcal{O}_{10'}$, that are suppressed due to the left-handedness of the weak interaction, as well as the scalar $\mathcal{O}_{S^{(\prime)}}$, pseudoscalar $\mathcal{O}_{P^{(\prime)}}$, tensor $\mathcal{O}_T$ and pseudotensor $\mathcal{O}_{T5}$ operators. All these operators are strongly suppressed and will not be considered further in this thesis.
In order to evaluate if a shift in the Wilson coefficients is indeed caused by the presence of NP, it is crucial to ensure all possible QCD contributions are sufficiently accounted for. A point of open debate is the contributions of so called "charm loops" as displayed in Fig. 1.5.

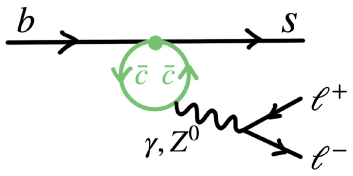


Figure 1.5: Sketch of a charm loop in green occurring at $b \to s$ vertex.

There are two energy scenarios to consider, where the energies of the charm quarks are above and below the beauty quark mass. While for larger energies, the effect is local and can therefore be absorbed in the current-current operators $\mathcal{O}_{1,2}$ and calculated perturbatively through form factors, the effect is non-local for smaller energies. As a result the charm quarks form hadrons, that can travel finite distances. These non-local effects cannot be calculated perturbatively and therefore their size is subject to debates. However, the most problematic part is that the non-local hadronic effects can create shifts in the C_7 and C_9 Wilson coefficients, that could falsely be interpreted as effects from NP contributions.

1.4.1 Towards measurements of $b \to s\ell^+\ell^-$ transitions

When classifying $b \to s\ell^+\ell^-$ transitions, it is important to choose a parameter that is equally important for theory and experiment in order to enable a good interpretation of the results for conclusions on NP scenarios. One such unit is the square of the transferred energy between the b and s quark, which is equal to the dilepton four momentum squared $q^2 = |p_{\ell^+} + p_{\ell^-}|^2$. It is a unit that can be experimentally easily accessed, while at the same time the form factors having a dependence on it making it very interesting for

theoretical interpretations. The most commonly chosen measure to show the importance of the transferred energy $\sqrt{q^2}$ is the differential decay rate as shown in Fig. 1.6. It is important to note that different transferred energy regions of the differential decay rate enable access to different Wilson coefficients and therefore test different NP scenarios. Therefore, it is always the most useful to explore the targeted observables binned in the transferred energy $\sqrt{q^2}$.

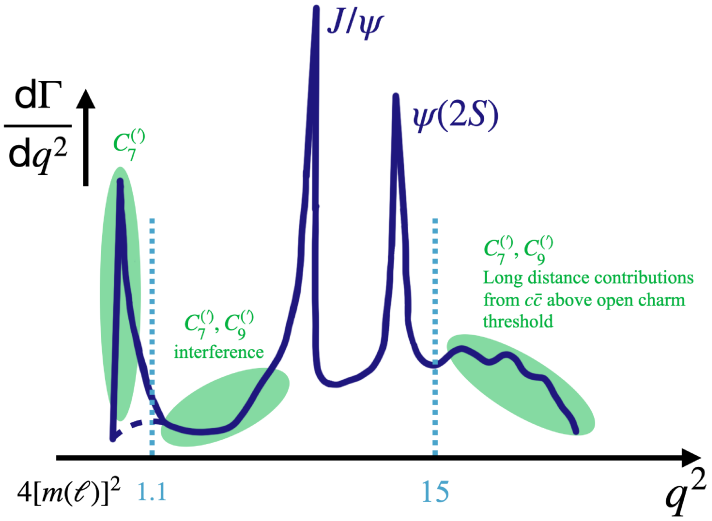


Figure 1.6: Sketch of the differential decay width as function of the squared dilepton invariant mass $q^2 = |p_{\ell^+} + p_{\ell^-}|^2$. The left peak represents the photon pole. The other two peaks are from the charmonium resonances J/ψ and $\psi(2S)$, which are $b \to c\bar{c}s$ tree-level transitions. They are included, as the final states are identical to the ones from $b \to s\ell^+\ell^-$ transitions. The areas sensitive to certain Wilson coefficients are marked by green ellipses.

The shape of the q^2 spectrum is highly depending on the form factors, available phase-space and allowed spin transitions, which will be elaborated in the following. The lower limit enabling $b \to s\ell^+\ell^-$ transition is, when the square of the transferred energy is at least $4m_\ell^2$ to produce the leptons at rest. At this q^2 value, the photon peak is visible below $q^2 \sim 1.1\,\mathrm{GeV}^2/c^4$, which corresponds to the decays, in which the massless γ radiation dominates over the virtual Z^0 boson. This region is particularly interesting, since it solely gives access to the photon penguin Wilson coefficient $C_7^{(\prime)}$. In addition, the decay of the photon to a pair of leptons gives direct access to the photon polarisation. The best lepton candidate for these measurements are electrons, since their much smaller mass leads to lower accessible areas and higher photon poles. However, this kind of measurement is not possible for decays into pseudoscalars such as $B^+ \to K^+\ell^+\ell^-$, since the radiative decay $B^+ \to K^+\gamma$ is spin-violating, resulting in absence of photon penguin enhancement (indicated by the dashed line below the photon pole in Fig. 1.6). The narrow $\phi(1020)$ resonance technically pollutes the region $q^2 = 1.05\,\mathrm{GeV}^2/c^4$, but the small branching fraction $\mathcal{B}(\phi(1020) \to \ell^+\ell^-) = 2.9 \times 10^{-4}$ [33] justifies neglecting it in the sketch. The region between the very-low q^2 and the $\phi(1020)$ resonance is called low-q^2 region. The higher-mass part of the spectrum is characterised by pollution from different charmonium resonances. They appear through the tree-level transition $b \to c\bar{c}s$ and the

two most prominent and narrow peaks are from the J/ψ and the $\psi(2S)$ resonances at $q^2 \sim 9.6\,\mathrm{GeV}^2/c^4$ and $q^2 \sim 13.6\,\mathrm{GeV}^2/c^4$. These two decay modes are especially important when studying $b \to s\ell^+\ell^-$ transitions, as they lead to the same final state particles. Therefore, they are quite powerful tools to cross-check analyses of the electroweak penguin transitions. The region between the $\phi(1020)$ and J/ψ is called the central-q^2 and is characterised by the interference of the Wilson coefficients C_7 and C_9. The region above $\psi(2S)$ is referred to as the high-q^2 (marked by the light blue dotted line), or the region of low hadronic recoil. In this region the Wilson coefficients C_9 and C_{10} interfere and long distance contributions from broad charmonia resonances above the open charm threshold contribute. However, they have rather small branching fractions for decays into a pair of leptons and can be neglected at first order. In the high-q^2 region, the energy of the hadron is smaller than the QCD energy scale Λ_{QCD}. This enables to perform theoretical predictions in the heavy quark limit ($m_b \to \infty$), which is particularly interesting since the relevant effective theories can exploit the spin-flavour symmetry of QCD [98, 99]. As a result, different form factors can be related and the number of unknown hadronic quantities is reduced. In addition, the calculations get an explicit m_b dependence due to the decoupling of the heavy-quark spin.

When comparing q^2-spectra of mesonic and baryonic $b \to s\ell^+\ell^-$ transitions, it is important that the q^2-dependence of the form factor shapes them such, that for mesons most of the signal is expected below the J/ψ resonance, while for baryons it is expected to be at high q^2 values above the $\psi(2S)$ peak (if the phase space allows so). This assumption is based on the only available form factor predictions for the $\Lambda_b^0 \to \Lambda$ transition, but e.g. the $\Xi_b^- \to \Xi^-$ transition can be related through $SU(3)$ symmetry to the $\Lambda_b^0 \to \Lambda$ one. A second order effect determining the shape of the q^2-spectrum is the available phase-space. For the $\Lambda_b^0 \to \Lambda(1520)$ transition, the phase-space is smaller compared to the ground state Λ transition due to the larger mass of the pK^- system. This results in most of the signal being situated in the region of the J/ψ and $\psi(2S)$ resonances [100].

Simply motivated by the higher available statistics due to the higher production fractions, most studies of $b \to s\ell^+\ell^-$ transitions have been performed with meson decays. The current status will be discussed in Sec. 1.4.2. However, it is especially interesting to study baryon transitions, since they probe possible spin dependence of NP. In addition, there are several promising NP models, that explicitly include a spin-dependence, that can only be tested with baryons [101]. In addition, the half-integer spin nature of the four weakly decaying beauty baryons gives unique access to the helicity structure of $b \to s\ell^+\ell^-$ transitions. The same information can not be accessed in the decay of B mesons as the information is lost in the hadronisation process.

As previously discussed, there are only four weakly-decaying ground states, namely the $\Lambda_b^0, \Xi_b^-, \Xi_b^0$ and Ω_b^- baryons. Their counterparts with spin $j = \frac{3}{2}$, decay all strongly or electromagnetically to another weakly decaying b-baryon together with a soft photon or pion. Due to the challenge of reconstructing such photons and pions, the analyses described in this thesis focus on the four directly weakly decaying baryons only.

When this PhD project was started, the only observed $b \to s\ell^+\ell^-$ transitions in baryons were those of Λ_b^0 decays. Therefore, the main focus of this PhD thesis was to examine the angular distributions of $\Lambda_b^0 \to \Lambda\ell^+\ell^-$ decays. The study of this particular decay has several benefits. It has not only the highest production fraction of the four weakly decaying b baryons, but also provides a rich angular structure due to the subsequent weak

decay of the Λ baryon, when compared to mesonic decay modes. In addition, the measurement of the differential branching fraction of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay [7], as shown in Fig. 1.7, has most of its signal in the high-q^2 region, where background levels from random combinations of tracks are lower.

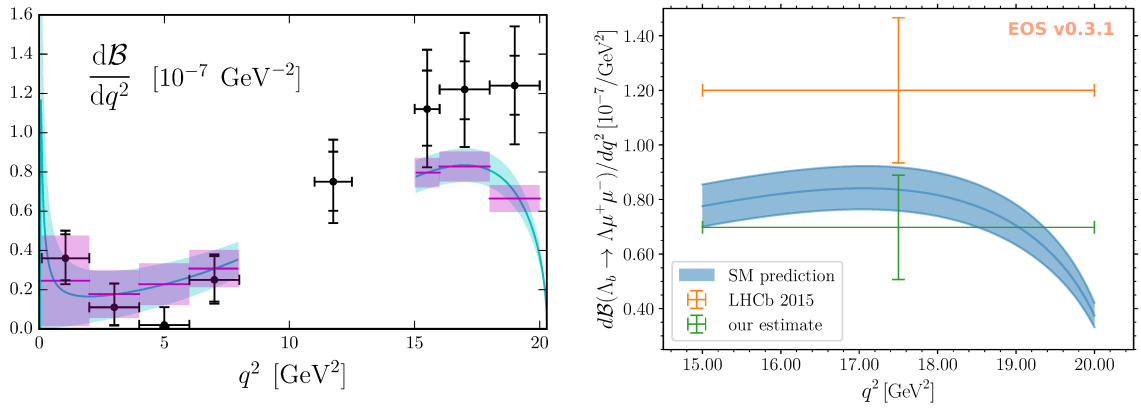


Figure 1.7: Differential rate as a function of q^2 in $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay, as measured by the LHCb collaboration [7] and compared to the SM predictions [102] (left). The dominant experimental uncertainty comes from the knowledge of the $\Lambda_b^0 \to \Lambda J/\psi$ absolute branching fraction and is correlated among all bins. Re-interpretation of the LHCb result in the bin $15 < q^2 < 20\,\text{GeV}^2/c^4$ with a more correct determination of the fragmentation fraction $f_{\Lambda_b^0}$ [103] (right). In both plots the uncertainty bands are included for the SM predictions.

Motivated by these advantages, it was decided to update the angular analysis of the muon decay mode, trying to extract all ten angular coefficients. Details about the current status of the experimental knowledge of $b \to s\ell^+\ell^-$ transitions of Λ_b^0 decays are given in Sec. 1.4.2. The angular basis, as well as the angular distributions analysed in this thesis are described in Sec. 8.1. The section ends with a short description, how angular analysis of the two different leptonic decay modes can be used to test the lepton flavour universality of the SM.

The $b \to s\ell^+\ell^-$ transition of the Ω_b^- baryon, $\Omega_b^- \to \Omega^-\ell^+\ell^-$, is particularly interesting, since the Ω_b^- baryon is decaying weakly into a strange baryon with spin $j = \frac{3}{2}$. This strange baryon is the only known weakly-decaying ground state with spin $j = \frac{3}{2}$. All other baryons with the same spin decay through the strong or electromagnetic interaction, making them more challenging to use for polarisation measurements. The tree-level decay $\Omega_b^- \to \Omega^- J/\psi(\to \mu^+\mu^-)$ was used to provide a mass measurement of the Ω_b^- baryon and its relative production fraction at the LHC. These measurements are crucial to update the limited knowledge about the properties of the Ω_b^- baryon. The details can be found in Ch. 5.

Finally, the $b \to s\ell^+\ell^-$ transitions of the Ξ_b isospin doublet are currently unobserved. Their observation is especially challenging due to their low production fractions. However, while the Ξ_b^- production fraction is very small, the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay is a decay chain of charged, weakly decaying hyperons, that cannot be mimicked by any mesonic decay. Therefore, although the signal yields were expected to be small, there was a high chance of observing this decay due to expected low background levels. As part of this thesis, the

first search for the decay $\Xi_b^- \to \Xi^-\mu^+\mu^-$ is performed and will be discussed in more detail in Ch. 6. For the Ξ^0 and the decay $\Xi_b^0 \to \Xi^0\mu^+\mu^-$ in particular, this is not true as it is a decay chain of neutral particles. They are much harder to correctly reconstruct, as they do not produce many tracks in a detector. So far, the Ξ^0 baryon has not been observed at the LHC[12]. Therefore, as a first step towards the observation of the rare decay mode, the search for the charmonia decays $\Xi_b^0 \to \Xi^0 J/\psi(\to \mu^+\mu^-)$ and $\Xi_b^0 \to \Xi^0\psi(2S)(\to \mu^+\mu^-)$ is carried out. This choice is motivated by much higher statistics for the tree-level decays into charmonia, as it can be seen in the sketch in Fig. 1.6, while at the same time leading to the same final state particles. Therefore, the charmonia decays enable to study crucial challenges and how to solve them, which can be implemented for the rare mode afterwards. The details of this analysis can be found in Ch. 7.

1.4.2 Experimental status of $b \to s\ell^+\ell^-$ transitions

The LHCb experiment has studied a variety of different decay modes of mesonic $b \to s\ell^+\ell^-$ transitions. Many different observables show tensions towards the SM predictions. While most of them remain below 3σ, individual bins show discrepancies of even above the 3σ limit. These deviations are referred to as B anomalies and the LHCb collaboration is keen to improve the understanding of these discrepancies with improved analysis techniques and larger datasets in order to validated their nature. This section includes a selection of the most recent publications in this area and discusses the importance of the different types of measurements.

Branching fractions

Branching fraction measurements are the natural first step, when observing a decay channel for the first time, and especially the differential branching fractions as function of q^2, $\frac{\mathrm{d}\mathcal{B}}{\mathrm{d}q^2}$, are interesting as they can be used to extract limits on Wilson coefficients. However, they suffer quite a bit from theoretical uncertainties on hadronic form factors. In addition, the size of the contribution from non-factorisable effects, such as quark loop contributions ($b \to sq\bar{q}$), in particular of charm loops, is subject of current debates. The differential branching fractions focus on muonic decay modes most of the time due to the much easier reconstruction compared to electrons[13] and have been measured for $B^+ \to K^+\mu^+\mu^-$, $B^0 \to K^0\mu^+\mu^-$, $B^0 \to K^{*0}\mu^+\mu^-$ and $B_s^0 \to \phi\mu^+\mu^-$ [14, 15, 17] by the LHCb collaboration. Fig. 1.8 shows the measurements of these branching fractions. They all show some deficit of events in the central-q^2 region with respect to the SM predictions. Beside the LHCb measurements, several other collaborations such as CDF, Belle, BaBar and CMS observed similar deviations [2, 4, 5, 8, 13].
For baryonic $b \to s\mu^+\mu^-$ transitions the differential branching fractions have only been measured for $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ [3, 7] and most recently for $\Lambda_b^0 \to \Lambda(1520)\mu^+\mu^-$ [100]. While for the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ a similar discrepancy in the central-q^2 bin can be seen as for the mesonic decays (see Fig. 1.7), the theory predictions for the $\Lambda_b^0 \to \Lambda(1520)\mu^+\mu^-$ disagree with each other making it impossible to draw a conclusion.

[12] The PDG claims the $\Xi_b^0 \to \Xi^0 J/\psi$ decay was seen, but the measurement in question only looks at Ξ_b^0 and Ξ_b^- baryons combined with partial reconstruction. Therefore, the Ξ^0 was not seen directly yet.
[13] The details will be discussed in Sec. H.

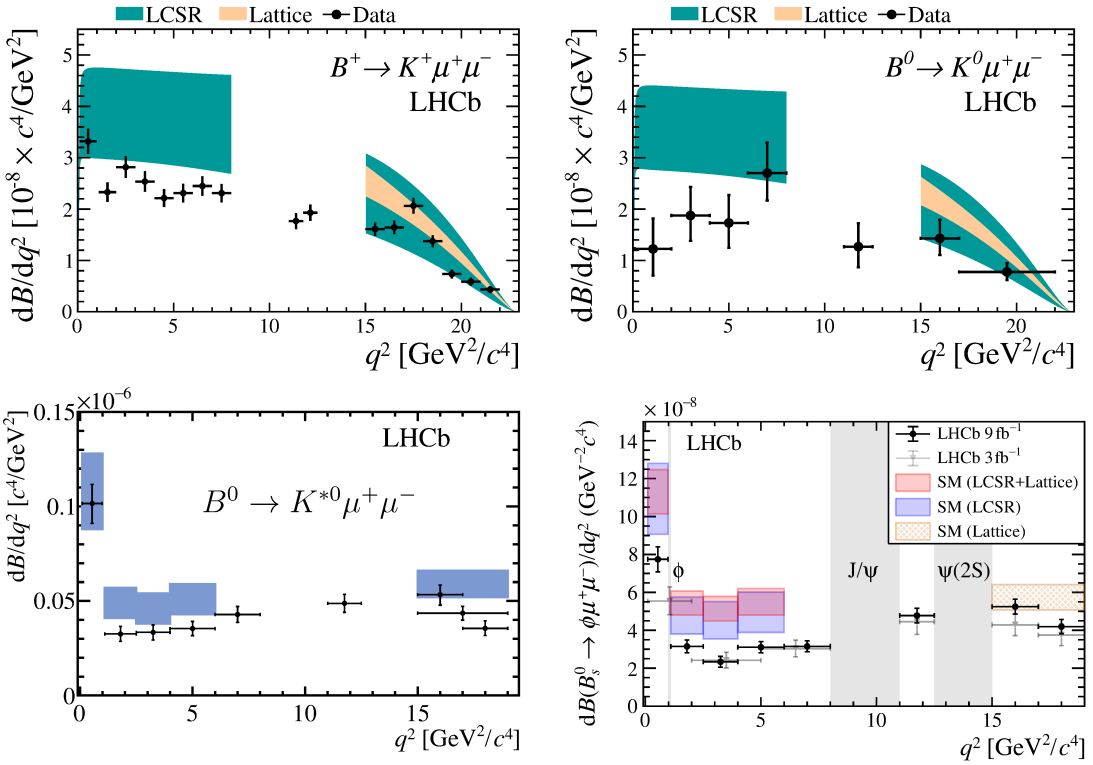


Figure 1.8: The LHCb measurements of the differential BFs of the $B^+ \to K^+\mu^+\mu^-$ (top left), $B^0 \to K^0_S\mu^+\mu^-$ (top right), $B^0 \to K^{*0}\mu^+\mu^-$ (bottom left), $B^0_s \to \phi\mu^+\mu^-$ (bottom right) are shown [6, 9, 16]. The black crosses are the data points from the measurements and the bands and boxes display SM predictions determined with lattice QCD and light-cone sum rule (LCSR) calculations.

Lepton Flavour Universality Ratios

The obstacles of large theoretical uncertainties in the branching fraction predictions can be overcome by taking ratios of decay modes with the same hadronic part, but comparing different leptons in the final states. As a consequence such ratios test the accidental lepton flavour symmetry of the SM by the ratios being unity within the limit of massless leptons. In addition, the uncertainties related to the form factor predictions completely cancel. Therefore, lepton flavour universality ratios are theoretically very clean. In the experimental measurements, the ratio is determined as double ratio with respect to the tree-level decay via J/ψ resonances as shown in Eqn. (1.31)

$$R = \frac{\mathcal{B}(H_b \to H_s\ \mu^+\mu^-)}{\mathcal{B}(H_b \to H_s\ J/\psi(\to \mu^+\mu^-))} \Big/ \frac{\mathcal{B}(H_b \to H_s\ e^+e^-)}{\mathcal{B}(H_b \to H_s\ J/\psi(\to e^+e^-))}, \tag{1.31}$$

where H_b represents the original beauty hadron decaying into the strange hadron H_s. The benefit of this choice is that the decay modes have the same final state particles and therefore of lot of the experimental uncertainties *e.g.* related to the reconstruction of the leptons cancel. Therefore these double ratios are experimentally and theoretically the cleanest observables. They are referred to as lepton flavour universality (LFU) ratios.

Several different collaborations have measured LFU ratios [2,4,13,104–106] being in agreement with unity but consistently showing deviations in different mesonic decay modes, that all measure fewer muon mode decays compared to the electron ones. The most recent lepton flavour universality measurement by the LHCb experiment [107,108] simultaneously determines the LFU ratios R_{K^+} and $R_{K^{*0}}$ in the low- and central-q^2 bin. The ratios are fully in agreement with SM predictions and discovered that previous deviations were mainly a result of underestimated mis-identification background of charmless decays. In addition, there was also one lepton flavour universality measurement performed with $\Lambda_b^0 \to pK^-\ell^+\ell^-$ [109]. The measurement was performed with the (pK^-) system first, since is has a higher reconstruction efficiency compared to the long-lived hyperon Λ and is in agreement with unity within one standard deviation.

Angular analyses

The last kind of possible measurements are angular analyses. They provide a large set of observables, that can be used to disentangles contributions from different Wilson coefficients. Since each angular coefficient depends on form factors, they can (sometimes) be organised in ratios of different angular coefficients that subsequently cancel the hadronic uncertainties. These new observables are called $P^{'}$ observables and have been studied for many different mesonic decay modes [10–12, 14, 15]. An examples is given by $P_5^{'} = S_5/\sqrt{F_L(1-F_L)}$, where S_5 is one of the angular coefficients and F_L the fraction of longitudinal polarisation of the strange meson. In the past there have been tensions observed in several of these theoretical clean prime observables. Examples are given in Fig. 1.9.

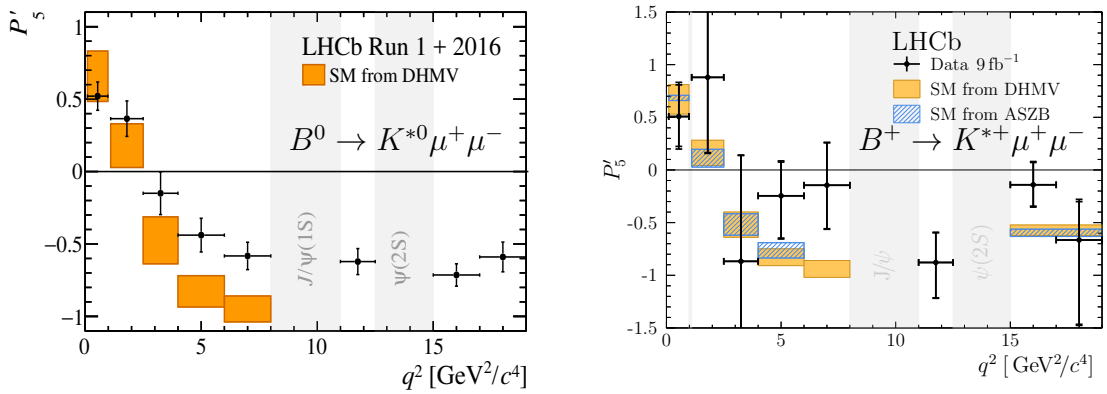


Figure 1.9: The LHCb measurements of the angular observables in $B^0 \to K^{*0}\mu^+\mu^-$ (left) and $B^+ \to K^{*+}\mu^+\mu^-$ (right). The black crosses are the data points and the blue and yellow boxes display SM predictions calculated by individual groups. Taken from Ref. [14, 15].

With the increasing precision of form factor predictions, the regular angular observables become more useful, as they enable to test the predictions of different models. Very recently, an unbinned amplitude analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay [110, 110] has been published. While it is in agreement with the SM predictions, it shows the same pattern

of deviations as results obtained from binned angular analysis.
As previously mentioned the very-low q^2 region of $b \to se^+e^-$ decay modes can be used to directly extract the photon polarisation. This has been done for the first time with the decay mode $B^0 \to K^{*0}e^+e^-$ [111] at the LHCb experiment.
In the context of this thesis and baryonic $b \to s\ell^+\ell^-$ transitions, there has been a likelihood angular analysis [7] of the decay $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ performed at the LHCb experiment. In addition, a method of moment analysis [112] was performed by the LHCb collaboration for the same decay channel including a larger dataset. A more detailed discussion about angular analyses follows in the next Ch. 8 focusing on the angular distribution of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$.

Chapter 2

The LHCb experiment at the LHC

In order to perform precision measurements in particle physics, it is crucial to collect large datasets with high selection and reconstruction efficiency for the final states of interest. When talking about beauty- or charm-quark physics, experiments such as the BaBar experiment [113] or the Belle and Belle II experiments [114, 115] are particularly interesting due to the cleanness of electron-positron collision events. Electron-positron colliders dedicated to the study of beauty hadrons are usually operated at a center-of-mass energy of $\sqrt{s} = 10.58\,\text{GeV}$, corresponding to the $\Upsilon(4S)$ resonance. This particular resonance has a mass just above the $B\bar{B}$ production threshold, and predominantly decays immediately into a pair of neutral or charged B mesons with relatively equal probability [1]. Due to the electromagnetic nature of the electron-positron collision and thanks to colliding the elementary particles, the underlying background is quite low. Furthermore, the $e^+e^- \to b\bar{b}$ cross-section is enhanced thanks to presence of the resonance. However, these colliders are limited in their available energy by their design choices. Indeed, the Belle experiment has collected some data at the $\Upsilon(5S)$ resonance that allows to study heavier B_s^0 meson pairs. However, unlike the $\Upsilon(4S)$ where the only allowed strong decay is that to $B\bar{B}$, the $\Upsilon(5S)$ meson can decay to a variety of final states, such as $B\bar{B}$, $B\bar{B}^*$, $B\bar{B}\pi$ or $B_s^0\bar{B}_s^0$ [33], which introduces complications. It has never been attempted to produce b-baryons, which are even heavier, at such "B-factory" electron colliders. However, b baryons were studied at the LEP experiments where they were predominantly produced in Z boson decays [116–118]. Operating dedicated flavour physics experiments at hadron colliders allows to collect much larger datasets compared to ones at B-factories due to the higher production cross-sections of $b\bar{b}$ and $c\bar{c}$ pairs with increased center-of-mass energy at the cost of worse purity.

Therefore, the design and successful operation of the LHCb (Large Hadron Collider beauty) experiment [119] at the Large Hadron Collider at CERN has ushered in a new era for precision measurements in flavour physics. This chapter is dedicated to proving this claim and explaining why the LHCb experiment offers the ideal environment to perform the measurements introduced in Ch. 1.

[1] The exact fraction can differ up to 5% due to the mass difference between the neutral and charged B meson and electromagnetic corrections.

Hence, the key concepts of the LHCb detector and how they enable meeting the experimental challenges of the showcased analyses are presented. It should be noted that the LHCb detector as described in this chapter was operated during the Run 1 and Run 2 of the LHC. It is also the version of the experiment used to collect the data analysed in the presented work in this thesis. The Upgrade detector will be described in a dedicated Ch. 3.

2.1 The Large Hadron Collider

The international particle physics laboratory CERN - *Conseil européen pour la recherche nucléaire* - is currently operating the world's most powerful particle collider, the Large Hadron Collider (LHC). The unique environment at CERN enables to reuse previously built smaller accelerators and connect them into a chain to reach a design energy of 7 TeV per proton beam and, therefore, a center-of-mass energy of 14 TeV during the collisions. However, so far due to limitations of the cryogenic system [120] only a peak energy of 13 TeV was reached at the end of the second run period of the LHC [2]. In the beginning the LHC was operated at a center-of-mass energy of 7 TeV which was increased to 8 TeV during Run 1 (2010-2012) and to 13 TeV for Run 2 (2015-2018). The LHC has been installed in a circular tunnel of 27 km circumference placed 100 m underground, initially built for the Large Electron-Positron collider (LEP). The full accelerator complex at CERN is displayed in Fig. 2.1.

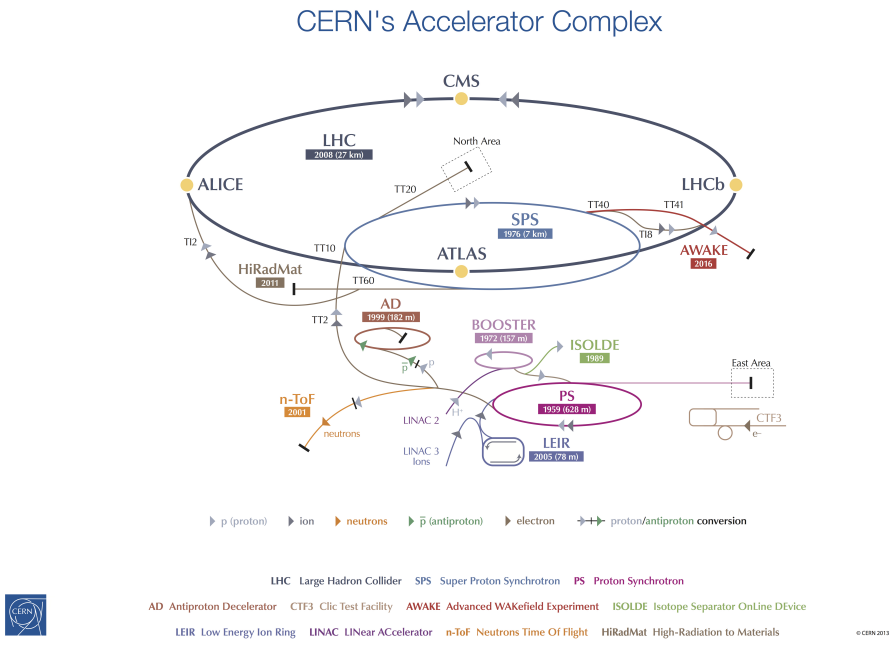


Figure 2.1: Overview of the accelerator complex at CERN taken from Ref. [121].

The protons for the proton beams are obtained by stripping hydrogen atoms of their

[2]Different upgrades, enable to reach a peak energy of 13.6 TeV for the third run period of the LHC, that started in 2022.

electrons and accelerating them in a linear accelerator called Linac2 [122]. The created proton bunches are then injected into a chain of different circular accelerators where the protons are accelerated up to 450 GeV. In a last step, the sixteen radio-frequency cavities of the LHC accelerate the beam to its final maximum energy. Due to precise control, the spacing between the bunches was decreased from 50 ns in Run 1 to 25 ns in most of Run 2, leading to a higher provided luminosity. The proton bunches are forced on their circular path along the ring by magnetic fields of up to 8.3 T, created by more than a thousand superconducting Nb-Ti dipole magnets. Two opposite magnetic fields are applied to the two proton beams moving in opposite directions. To ensure that the proton bunches in the beams are not spreading out, they are focused using quadrupole magnets located along the ring. Indicated with yellow dots on the dark blue LHC circle in the schematic drawing of the accelerator complex in Fig. 2.1 are the four experimental points. Here, the optics of the LHC can be used to collide two proton bunches and, therefore, the four main experiments of the LHC are placed there. One of them is the ALICE experiment [123], dedicated to the study of heavy ion collisions and quark-gluon plasma assumed to have existed in the early stages of our universe. The two largest experiments operating at LHC are the two general-purpose detectors ATLAS [124] and CMS [125], which are centered around their interaction points and cover almost the full solid angle. Their physics program is targeted towards high-transverse-momentum physics, precision tests of the Standard Model, top-quark physics, Higgs boson physics, as well as search for the phenomena beyond the SM. In addition, they are also able to perform flavour physics studies, but have several limitations.

Since quark production cross-sections in proton-proton collisions like $\sigma_{b\bar{b}}$ are significantly higher as the ones for top quarks or Higgs bosons[3], the proton beams do not need to collide head-on at the highest possible center-of-mass energy, in order to reach considerable high statistics for e.g. top or Higgs physics. This comes with the advantage of having a better control over the large contributions of backgrounds in hadron collider environments when performing flavour physics. In addition, the lower number of interactions per collision (pile-up), if the beams do not collide head-on, and the resulting decreased track multiplicity enable to lower the trigger thresholds while remaining with an amount of produced data at a processable level. Lastly, these detectors suffer from the absence of charged hadron identification.
Consequently, a dedicated flavour physics experiment was designed, the LHCb detector, tackling these disadvantages and more to push the boundaries of precision. The following sections detail the concept and data taking of the LHCb experiment.

2.2 The LHCb experiment

The LHCb experiment is a dedicated heavy-flavour physics experiment [127], whose main goal is the study of CP violation in weak interactions and indirect search for NP effects in rare decays of beauty and charm hadrons. It is a single-arm forward spectrometer covering only the forward pseudorapidity region $2 < \eta < 5$, corresponding to a coverage of about 4% of the solid angle.
In order to understand this particular design choice it is important to evaluate the $b\bar{b}$

[3]The dedicated plot displaying different cross-sections can be found in Ref. [126].

production at the LHC. As previously discussed, valence quarks (uud) are confined in bound states, such as the proton. With increasing energy in proton-proton collisions, an increasing fraction of the protons momentum is carried through sea quarks and even more by gluons. At the LHC, gluon-gluon fusion $gg \to b\bar{b}$ is the dominant $b\bar{b}$ production process and quark-antiquark annihilation from sea quarks $q\bar{q} \to b\bar{b}$ plays only a sub-leading role. This is particularly interesting, since there are usually large momentum asymmetries between the interacting partons from the two protons. Combined with the much larger center-of-mass energy of the LHC compared to the beauty quark mass, the produced $b\bar{b}$ pair is usually boosted in the forward or backward direction along the beam axis. The closeness to the beam axis is usually characterised in terms of the pseudorapidity defined as

$$\eta = -\ln \tan \frac{\theta}{2}$$

where θ is the polar angle with respect to the beam axis. The left plot in figure 2.2 displays the boost of the $b\bar{b}$ pairs, produced at the LHC. The right plot shows the pseudorapidity coverage of the LHCb experiment in red and a general purpose detector in yellow with $-2.4 < \eta < 2.4$. For the latter, this translates to a coverage of 90% of the solid angle and detecting about 45% of the produced $b\bar{b}$ pairs, compared to only 25% for LHCb. However, in higher pseudorapidity regions, the $b\bar{b}$ pairs are boosted more and therefore the b-hadron decay vertices are more displaced compared to the lower pseudorapidity regions of general purpose detectors. It enables to identify b-hadrons in the harsh environment of the hadronic collision. Thus, this is the first benefit of the unusual design of the LHCb experiment as a single-arm forward spectrometer[4]. Beside the displacement of the beauty decay vertices, the LHCb experiment has a second reason for being called a spectrometer due to its ring-imagine Cherenkov detectors (RICH) that are crucial to providing a reliable particle identification system. These systems are challenging to scale to the general-purpose detectors' size and shape. The combination of precise tracking and particle identification systems enables to focus the use of the calorimeter system to significantly improvements for particle identification for electron, photons and neutral pions and the measurement of energies for neutral particles.
Before going into more details about the actual detector design, it is useful to consider LHCb's special setup for the provided collision rate by the LHC. It is usually characterised by the instantaneous luminosity $\mathcal{L}(t)$

$$\int \mathcal{L}(t)dt = \frac{N}{\sigma},$$

which integrated over time is equal to the rate of a given process (number of events N) divided by its cross-section σ. Therefore, if a process has a large cross-section more events are produced at the same *integrated luminosity* compared to a process with a smaller cross-section. As a result due to the high production cross-sections for $b\bar{b}$ pairs, the LHC can provide collisions for LHCb at a significantly lower rate than for ATLAS and CMS. The

[4]In principle, it would be beneficial to increase the statistics by a factor of two, by simply mirroring the LHCb experiment with respect to the interaction point to double the available dataset. But beside the fact that no such space is available in the cavern, the precision only increases with a factor of square root of two, which does not support the doubling in costs.

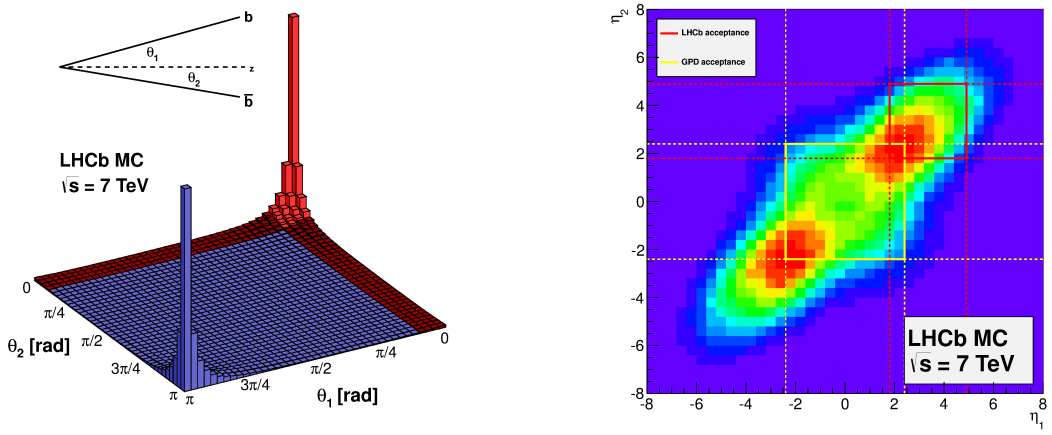


Figure 2.2: $b\bar{b}$ production rates as a function of their angles (left) and pseudorapidities (right). The regions fully covered by LHCb are presented in red and the ones of general purpose detectors in yellow (only for the right plot). Taken from Ref. [128].

used technique is called *luminosity levelling*. It describes a procedure, where the proton beams are deliberately slightly misaligned in the beginning of a fill and only over time, with the decrease of the beam quality, the beam overlap is increased until they collide head-on without an angle between the two beams. The left plot in Fig. 2.3 shows the instantaneous luminosities for ATLAS, CMS and LHCb for a random long LHC fill during Run 1. It can be clearly seen that the instantaneous luminosity remains constant for LHCb, till about 14 h. This is the point in time, when the two beams are colliding head-on also for LHCb. The luminosity levelling results in an optimised LHCb performance, since the integrated luminosity over a fill can be maximised, the trigger configuration does not need to be adjusted and related systematic uncertainties are reduced.

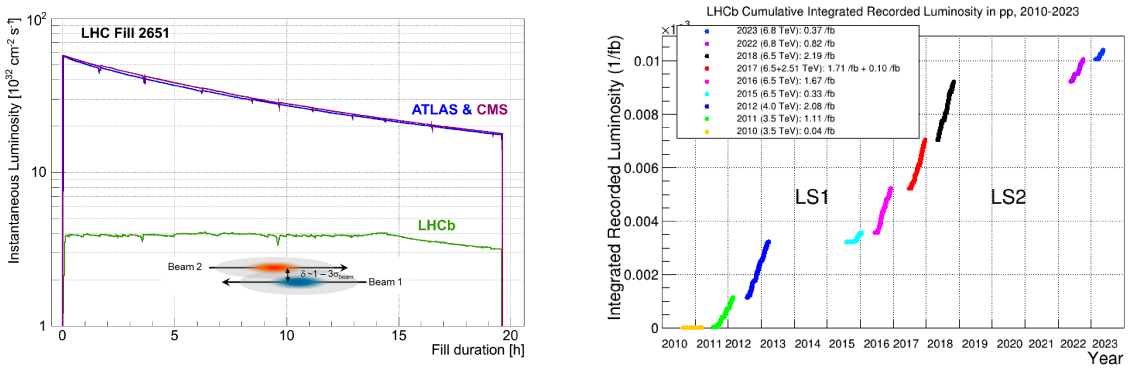


Figure 2.3: The left plot shows the development of the instantaneous luminosity for ATLAS (blue), CMS (pink) and LHCb (green) during one long LHC fill in Run 1. Taken from Ref. [129]. The right plot shows the cumulative integrated luminosity, recorded by the LHCb detector during each year of Run 1 and II. Taken from Ref. [130].

The right plot of Fig. 2.3 displays the recorded integrated luminosity collected by the LHCb experiment, split for each year of the data taking during Run 1 and Run 2. The total collected dataset during Run 1 and Run 2 combined, sums up to $9\,\text{fb}^{-1}$. The average data taking efficiency for the LHCb experiment is about 90%, and about 99% of the recorded data is usable for analysis [129]. The inefficiency during the data taking arise dominantly due to safety procedures, dead-time of some subdetector front-end systems, and short technical problems.
As a consequence of the above described concepts, the LHCb collaboration provides for most measurements a better precision in flavour physics measurements than those of ATLAS or CMS with much larger recorded luminosity.

2.2.1 Detector overview

A schematic representation of the LHCb detector and its coordinate system can be found in Fig. 2.4. The LHCb experiment uses a coordinate frame in which the z-axis is defined along the beam axis, with its positive direction pointing from the interaction point towards the muon system. The y-axis is vertical, oriented from the interaction point towards the surface and perpendicular to the LHC tunnel. Finally, the x-axis is defined in order to form a right-handed coordinate system.

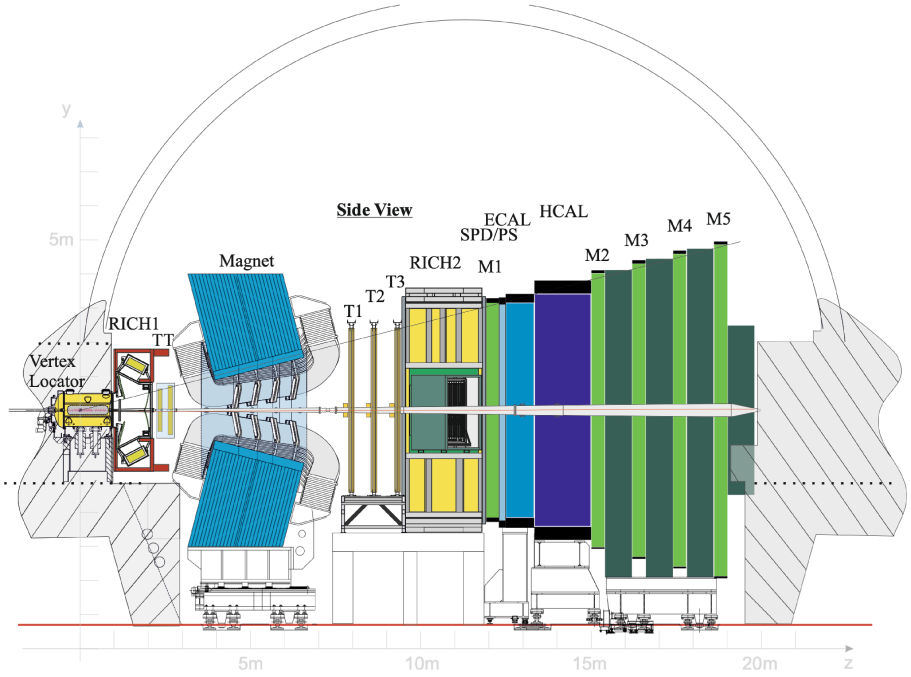


Figure 2.4: Schematic view of the LHCb detector as operated during Run 1 and Run 2. Taken from Ref. [131].

The LHCb detector is designed for measurements that iteratively reconstruct beauty and charm hadrons from its decay products of stable or *quasi-stable* particles[5]. These following particles can be detected directly: charged pions ($\pi^{\pm}$), charged kaons ($K^{\pm}$), protons (p and $\bar{p}$), electrons and positrons ($e^{\pm}$), muons ($\mu^{\pm}$), photons (γ) and deuterons (d).

[5]These are particles, that live long enough to travel through the detector, e.g. muons.

In addition, neutral hadrons can be reconstructed via the calorimeter system. Unstable hadrons are reconstructed from their decay products using the four-momentum conservation.
Only neutrinos can not be reconstructed within LHCb, since the small solid angle coverage of the experiment does not enable reconstructions via missing transverse energy as done in general purpose detectors. A specific class of long-lived particles is also considered, which includes the weakly-decaying strange hadrons such as K^0_S, Λ, $\Xi^{0,\pm}$ or Ω^-. They can cross several subdetectors before decaying and will be called hyperons (except for K^0_S) in the following analysis.
As previously discussed, the LHCb experiment is optimised in a way to harvest most of the highly boosted b-hadrons and incorporate a dedicated particle identification system. Therefore, the LHCb detector incorporates several specialised sub-detectors, shown schematically in Fig. 2.4, which enable to precisely measure the discussed particles and therefore reconstruct properties of hadrons with high precision. The track reconstruction system includes a vertex detector, called the Vertex Locator (VELO) which is located very close to the collision point, and two systems with planar tracking stations: the Tracker Turicensis (TT) located upstream of the dipole magnet, and three tracking (T1-T3) stations located downstream of the magnet. The VELO is a very special subdetector at the LHCb, which can locate precisely the collision point and decay vertices of b-hadrons due to their forward boost. This is crucial in order to separate the b-hadron decays from the promptly produced tracks, to measure the b- and c-hadron lifetimes and to distinguish between prompt and displaced event topologies. The LHCb dipole magnet reaches up to 4 Tm and bends charged particles in the horizontal plane, enabling to determine the charge of the track, as well as estimate its momentum based on the curvature. The non-uniformity of the magnetic field was measured to be about 1% [132]. The magnet polarity can be reversed to reduce detection asymmetries, this is particularly important when measuring CP violation. The two possible polarity configurations are called *MagDown* and *MagUp*. Polarities are regularly reversed, so that the samples of both polarities are approximately equal for each data-taking year.
Charged particle identification (PID) is provided by two Ring Imaging Cherenkov detectors (RICH1 and RICH2), enabling to especially distinguish between different types of hadrons. In addition, muons are identified by the muon system and information from the calorimeter help to identify electrons, photons and neutral particles. The calorimeter system is composed of an electromagnetic calorimeter (ECAL) with its Preshower (PS) and Scintillator Pad Detector (SPD); and a hadronic calorimeter (HCAL). The muon system is composed by five muon chambers (M1-M5), all but M1 located behind the calorimeter system. Both the calorimeter and muon system are used as a first step in the trigger system used to reduce the obtained data to a processable level.

2.2.2 Subdetector of the LHCb experiment

The individual subdetectors are described in detail in the this section.

The Vertex Locator

The VELO detector [133, 134] encircling the interaction point at the LHCb experiment with two silicon-based sensor halves is a unique detector system to the whole LHC environment. Instead of a beam-pipe, the LHC beams are separated by a thin aluminium foil, commonly refereed to as *RF-foil*. The name arises from the fact that it not only separates the LHC beam vacuum from the primary vacuum in the VELO vacuum container protecting the sensors, but the foil also shields the latter from radio-frequency interference from the LHC beams. Due to the absence of a fixed beam pipe, the sensors can be positioned as close as 8 mm away from the beam center during the data taking. This position can pose risks during beam instability such as during injections. This leads to the second novelty of the VELO detector. To mitigate potential damage to the sensors, the two halves can move: the gap remains open during injection and is only closed when the beams are in a stable condition.

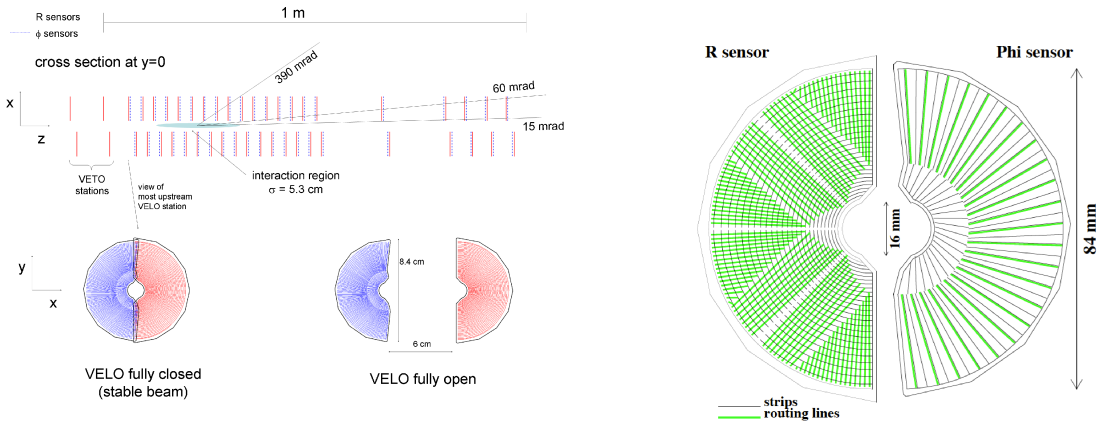


Figure 2.5: Overview of the VELO detector, showing spacing of modules along z-axis, and open and closed positions (left). Close up of the VELO R-φ silicon strip sensors (right). Taken from Ref. [135].

Each half of the VELO subdetector consists out of 21 sensor modules, that are designed to precisely determine the position of the primary vertex. The left plot of Fig. 2.5 shows the positions of the individual sensor positions with respect to the interaction region. In addition, the position the two halves are shown in the fully closed and fully open state depending on the beam condition. The right plot of Fig. 2.5 displays the structure of the R-φ silicon strip sensors. The R-type and φ-type sensors have their routing lines oriented perpendicularly and parallel to the silicon strips, respectively. This configuration facilitates the measurement of the radial distance to the beam axis and the azimuths angle, respectively.

TT and downstream tracking stations

The left plot of Fig. 2.6 shows the tracking detectors of the LHCb experiment. The silicon based detectors are shown in purple and the straw tube based detector in turquoise.

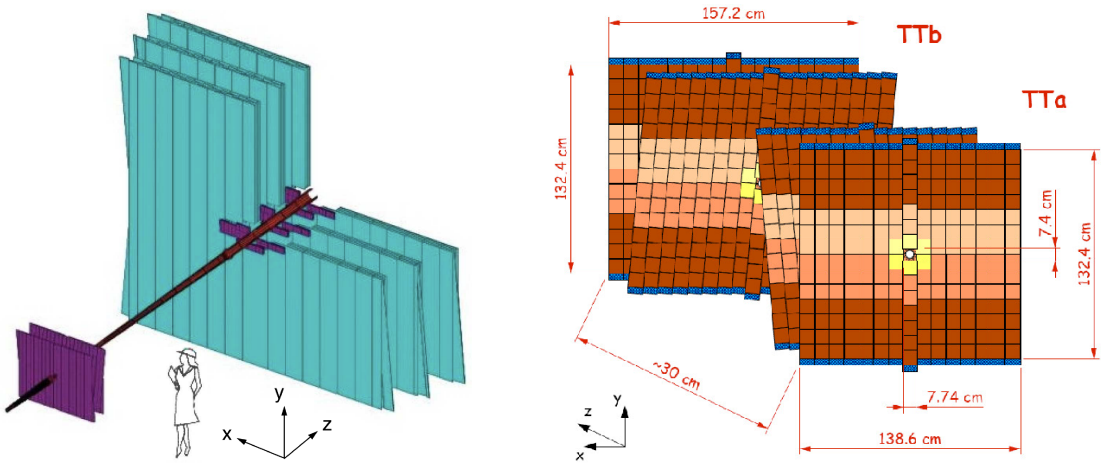


Figure 2.6: LHCb tracking detectors. The TT detector can be seen to the left, with three larger tracking stations (T1-T3) to the right. The silicon trackers are shown in purple for the TT and inner part (IT) of the tracking stations, and the straw tubes for the outer part (OT) – in turquoise (left). Schematic view of the TT stations (right). Taken from Refs [132, 136].

The first comprises two stations, each containing two layers, as displayed in the right plot of Fig. 2.6 and situated right before the magnet. Among these layers, the middle two are tilted by ±5 degrees relatively to the vertical axis, enabling three-dimensional track reconstruction. The TT enhances the momentum resolution of long tracks and is particularly crucial for forming downstream tracks and therefore the reconstruction of long-lived particles as studied in this thesis.
The tracking stations utilize two distinct detector technologies to optimise the tracking precision for a large surface detector, while staying cost efficient. Similarly to the TT, each tracking station consists of four layers, with two of them being tilted. The inner tracker (IT) [137] uses the same silicon-strip technology as the TT and covers the region proximal to the beam axis to provide precise tracking for strongly boosted particles. The straw tubes of the Outer Tracker (OT) [138–140] span over the outer regions of each tracking station. They are filled with a blend of argon, CO_2, and oxygen, enabling a drift time of ionisation clusters of less than 50 ns. The hit positions of particles passing through the OT are determined by measuring the drift time generated in the gas to the anode wire.

RICH detectors

For precision heavy-flavour measurements, it is crucial to have a powerful tool able to separate protons, kaons and pions in order to suppress backgrounds and correctly select signal candidates. The key detectors for this purpose are the two RICH detectors as displayed in Fig. 2.7.

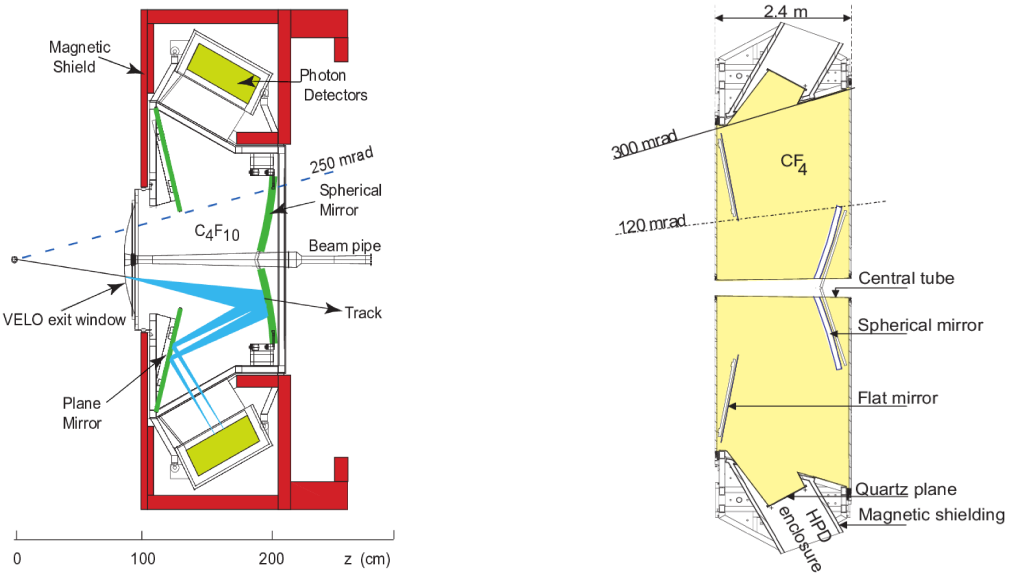


Figure 2.7: Design of the RICH 1 (left) and RICH 2 (right) detectors during Run 2. Taken from Ref. [141].

If a charged particle travels through a dielectric medium with a reflection index n faster than the phase velocity of light in this medium c/n, it will emit Cherenkov light. The emitted photons will form a cone with an opening angle $\cos\theta = \frac{1}{n\beta}$, where $\beta = v_{particle}/c$, pointing in the flight direction of the particle. By combining the Cherenkov angles with the track momentum information, the mass of the particle can be determined:

$$\cos\theta = \sqrt{1 + \left(\frac{mc}{p}\right)^2}. \tag{2.1}$$

In order to exploit a wider momentum range the two RICH detectors use different media, with different refraction indices. RICH1 uses the C_4F_{10} gas with $n = 1.0014$, and RICH2 – the CF_4 gas with $n = 1.0005$. RICH1 covers the full LHCb acceptance, while RICH2 covers only high-pseudorapidity range, where the majority of the targeted high-momentum particles are expected to end up. Cherenkov photons emitted inside radiators are reflected by the spherical mirrors and are directed to the Hybrid PhotoDetectors (HPD), that detect photons in the optical range (200 – 600 nm). During Run 1 RICH 1 possessed additional volumes filled with aerogel, which was eliminated before the start of Run 2 due to its performance decrease in high-multiplicity events [141].
Dedicated algorithms assign single photons to Cherenkov rings and compute the corresponding Cherenkov angle. Fig. 2.8 displays the reconstructed Cherenkov angle θ as a function of momentum of the tracks crossing the C_4F_{10} radiator of RICH1.
There are three clear curves identifying protons, kaons and pions, with the latter having the most produced Cherenkov light (and a fourth one from muons). This is direct consequence of the abundant production of pions during the proton-proton collisions and many decay processes as they are the lightest mesons. In addition, at the low momentum range there is another curve identifying muons. It is evident that there are kinematic regions where a reliable separation of different hadrons is not possible. At low momentum, kaons and protons do not have a sufficient velocity to create a Cherenkov ring. At the

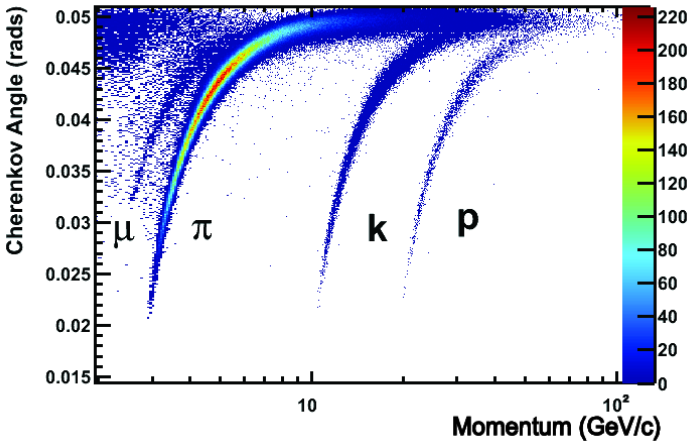


Figure 2.8: Reconstructed Cherenkov angle θ as a function of momentum of the tracks crossing the C_4F_{10} radiator, for several charged particles in 2011 LHCb data. Taken from Ref. [129].

same time, at high momentum, the Cherenkov angle is behaving in a similar manner for all charged particles, making their separation difficult and can be seen by the saturation of all particle types towards an Cherenkov angle of $\theta \sim 0.05\,\text{mrad}$.

Calorimeter system

The LHCb calorimeter [142] allows to measure particles energies and positions. It aims to measure the energy of the particles, by stopping and absorbing them. However, since the combined precision of the momentum resolution from the tracking system and the particles identification is so high, the energy of charged particles can be estimated much more precisely by utilising Einstein's relativistic energy

$$E = \sqrt{m^2c^4 + p^2c^2}.$$

Therefore, it is mainly used to measure the energy of neutral particles and more importantly provides crucial information for the separation of electrons, photons and hadrons. In addition, it plays a crucial role in the hardware trigger system, which will be discussed in detail in Sec. 2.2.5. For this multipurpose use of the calorimeter system it includes four different subdetectors – the SPD, PS, ECAL and HCAL as displayed in Fig 2.9.
SPD and PS consist of scintillator pads, with a lead plate located between the two subdetectors. Similarly to other experiments calorimeter systems, the ECAL and HCAL rely on an approach, where the absorber is used to initiate the shower and eventually stop the particles of interest, while the scintillator is used to measure the resulting energy deposits. For the ECAL, stacks of scintillator and lead absorber plates are alternated, which is a technology called *shashlik*. For the HCAL, an alternation of iron absorber and scintillator tiles is used. All four subdetectors use wavelength-shifting fibers crossing the internal structures to propagate the produced scintillation light to the readout photomultipliers. Comparably to the tracking stations, the four subdetectors increase the granularity with the hit density towards the beam pipe. The HCAL is only divided in two regions, as its energy resolution is too poor for proper energy measurements, it is primarily used to

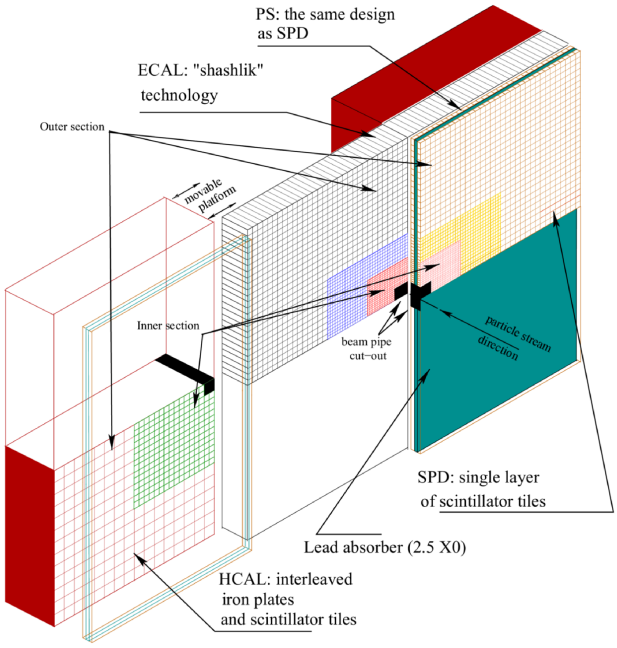


Figure 2.9: Schematic representation of the layout of the calorimeter system. Taken from Ref. [143].

identify high energy hadronic clusters for the hardware trigger stage [144,145]. The ECAL is crucial for the identification and reconstruction of electrons, photons and π^0, therefore it has larger granularity with a respective cell widths of 40.4, 60.6, and 121.2 mm in the inner, middle, and outer regions. The energy resolution of the ECAL is about

$$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} + 1\%, \tag{2.2}$$

where the energy is given in GeV [144,145].
The schematic representation of the calorimeter system in the left of Fig. 2.10 can be used to explain the principle of the particle identification with its different subdetectors.
The chosen thickness of the ECAL is about 25 radiation lengths, allowing to fully contain electromagnetic showers. Since only charged particles leave hits in the SPD, photons can be identified by the absence of such hits. The about 2.5 radiation lengths thick lead plate between the SPD and PS, forces the photons to initiate an electromagnetic shower, which is then detected by the PS and ECAL. In addition, they are identified by the fact that no tracks are pointing towards the created ECAL cluster. Due to the much larger hadronic interaction lengths compared to the electromagnetic ones, charged hadrons leave hits in the SPD, PS and only start to shower at the end of the ECAL, but are unaffected by the lead plate between the SPD and PS. They deposit most of the energy release in the HCAL, which has a thickness of about 5.6 hadronic interaction lengths. Since electrons release hits in the SPD, these can be used to distinguish them from photon showers that likewise start in the lead plate between the SPD and PS. In addition, the fraction of energy released in the ECAL, combined with tracking information is a powerful tool to distinguish charged hadrons and electrons. When a track is combined with an calorimeter

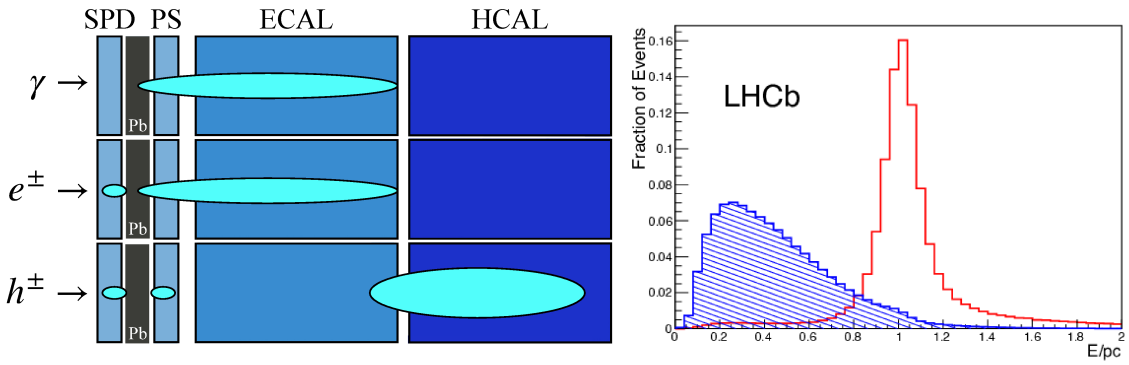


Figure 2.10: Schematic explanation of the principle of identification of photons, electrons and hadrons with the LHCb calorimeter system (left). Adaption from Ref. [146]. Distribution of E/p for electrons (red) and hadrons (blue), as obtained from a part of the 2011 dataset (right). Taken from Ref. [129].

cluster, the ratio of the deposited energy over the track momentum E/p can be determined. As can be seen in the left plot of Fig. 2.10 for electrons, this ratio is close to unity (modulo the resolution effects), while hadrons leave only a little fraction of their energy in the ECAL[6].

The remaining identification challenge of the calorimeter system is to distinguish between single photons and neutral π^0 mesons. Since π^0 decay in about 99% of the time into a pair of photons [33], they can leave the same signature if the opening angle is too small. Within LHCb, those pion decays are distinguished into two classes: *merged* when both pion photons end up in the same ECAL cluster, and *resolved*, when the two photons end up in separate clusters. The merged pions have usually a high transverse momentum and have the similar signature as a photon. The photon clusters from single photons and merged π^0 are distinguished by multivariate algorithm, called `isPhoton`. The cluster shape of a merged π^0 is usually a bit broader and has an elliptical shape, that can be exploited by the algorithm.

The correct identification of single photons is important as photons reconstructed in the ECAL are crucial to recover bremsstrahlung radiation emitted by electrons due to their interactions with the LHCb material. In addition, utilising the calorimeter energy deposits, as well as the positions of clusters is the only way to perform energy measurements for neutral particles.

Since no tracking information is available for neutral particles, the position of the ECAL cluster of a π^0 candidate is used to extrapolate the assumed path to the PV. This is a good assumption for neutral pions originating from b-hadron decays, as the typical flight distance of the b hadrons (about a centimeter) is negligible compared to the distance between the PV and the ECAL. However, if the targeted π^0 is assumed to originate from nearby the PV but actually arises from the decay of a long-lived hyperon like Ξ^0 (that can fly several meters), a large bias in the decay vertex position will be introduced. This would then lead to a bias in the estimated kinematic properties of π^0.

[6] A special feature of this ratio is that it is unaffected by bremsstrahlung losses. If a bremsstrahlung photon is not reconstructed, the effect is equally effecting the momentum and energy measurement and therefore cancelling. More details about bremsstrahlung can be found in App. H.

Muon stations

The last subsystem of the LHCb detector is the muon system [147–149] composed of five muon stations (M1-M5). They are used for the identification of muons and triggering purposes. The muon particle identification involves the use of data from the muon stations, the RICH and tracking systems, as well as the calorimeters [150]. The initial station, M1, is positioned upstream of the calorimeters to enhance the transverse momentum resolution in the muon trigger, by mitigating the impact of multiple scattering within the calorimeter material. The remaining four stations are situated after the calorimeter. Between the muon stations, absorbers consisting of 80 cm thick iron plates are placed. Combined with the stopping power of the calorimeters, this configuration effectively suppresses the majority of charged particles, except for muons. Only the M2 station occasionally sees some spillover from the calorimeter system. In order to traverse all muon chambers and absorbers, muons must possess a momentum exceeding 6 GeV/c.
Fig. 2.11 displays the layout of the muon system. Each of the muon stations is divided in four regions (R1-R4) of a different granularity to incorporate a finer segmentation in the regions close to the beam pipe, which experience a higher particle multiplicity. All regions in all the muon stations use Multiwire Proportional Chambers (MWPC). The only exception is the innermost region of M1, which uses the Gas Electron Multiplier (GEM) technology. This choice was made as the technology has a higher radiation tolerance [148].

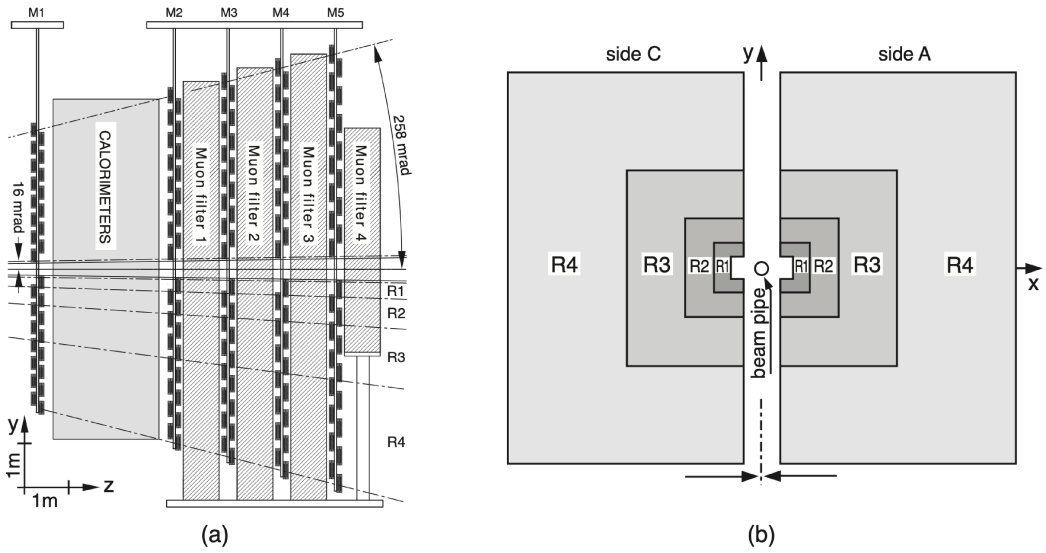


Figure 2.11: The layout of the LHCb muon system as a side view (left). The station layout featuring four regions of a different granularity (right). Taken from Ref. [149].

2.2.3 Vertexing and tracking

In order to understand the design choices made for the vertexing and tracking system, it is helpful to evaluate the constraints and the physics case. As previously discussed, beauty and charm hadrons are highly boosted, therefore the whole detector design is optimised in

the forward region. In order to reject backgrounds and improve the resolution on invariant mass distributions, that can be used for precision measurements such as the mass measurement presented in Chapter 5, a precise vertexing and a good momentum resolution is needed. These goals can be reached with a good spatial resolution, while simultaneously keeping a low material budget. This is important to reduce the multiple scattering and hadronic interactions of particles travelling through the detector. The tracking system is designed to reconstruct the trajectories of charged particles from the hits recorded by the various subdetectors.
Depending on the lifetime of a created particle and its resulting decay point, its mass and its momentum, several different types of particle tracks can be created. Their reconstruction can be subdivided into a track finding/pattern recognition part and a track fitting part which is done by a Kalman filter [151]. The possible types of tracks are displayed in Fig. 2.12.

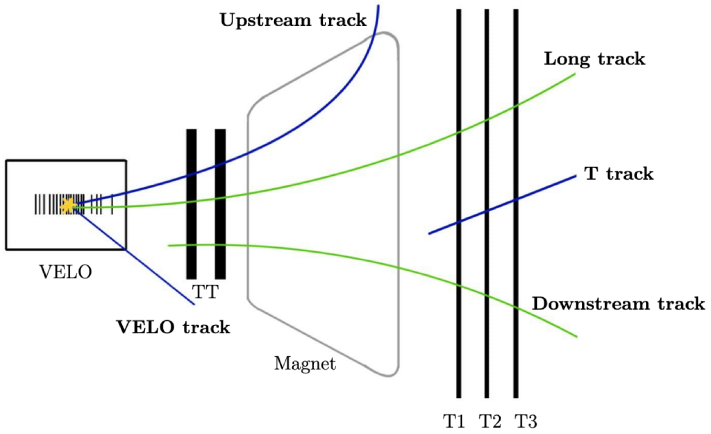


Figure 2.12: Schematic view of the different track types created within the LHCb detector, along with a schematic of the essential tracking detectors, showing the VELO detector, TT detector, magnet and the downstream tracking stations. The long and downstream track types relevant for the work presented in this thesis are highlighted in green, while the other types are displayed in blue. The yellow star in the VELO sketch indicates the proton-proton interaction point.

First basic track finding algorithms, called VELO tracking [152] and T seeding [153], are used to reconstruct straight VELO and T track candidates. These can be combined [154–156], forming the tracks with the most precise momentum estimate, that are called long tracks. In addition, if there are hits in TT on the trajectory of the long track [157], they are also added to further improve the momentum resolution and reduce the fake-track rate. In total, long tracks are reconstructed with about 95% efficiency [158, 159] and are therefore used in most analyses of LHCb. The analyses presented in this thesis study decays of long-lived hyperons, which often happen outside of the VELO, but the resulting charged decay products leave hits in the TT and the downstream tracking stations. In this case, a dedicated algorithm [160] matches T tracks with hits in the TT subdetector to reconstructed the so-called downstream tracks. All other track types have rather poor momentum resolution, which is a result of the fact that the charged particle trajectories are to first order only bent within the magnet volume [129, 142, 158],

but travel mostly in straight lines through the different tracking detectors. Consequently, tracks that have only information from tracking detector located before (VELO or Upstream tracks) or after the magnet (T tracks) have rather poor momentum resolution and (due to lack of information from certain subdetectors) unreliable identification. Nevertheless, these tracks are still important for various reconstruction and track isolation algorithms. For example, the VELO tracks, reconstructed using only VELO hits, enable the precise localisation of the primary vertex (PV). This arises as a result of the VELO being the only subdetector that has an acceptance in the backward direction. Sometimes a particle is reconstructed more than once, as different track types. In this case only the track with the larger number of crossed subdetectors and consequently the best momentum resolution is kept [158]. By combining as many subdetector hits as possible, fake tracks, i.e. tracks which are created as a combination of random unrelated hits in tracking detectors, can be rejected. These fake tracks are rejected by a neural network considering the information from all subdetectors [161].
In a second step, LHCb tracking algorithms rely on a fit based on the Kalman filter [151], allowing to improve the resolution on the track parameters and reject fake ('*ghost*') tracks. The fit improves the reconstructed particle trajectory by taking into account effects from multiple scattering and energy loss in the material interactions. The χ^2 per number of degrees of freedom, $\chi^2_{\text{track}}/\text{nDOF}$, of this fit, is a proxy for the quality of the track and therefore a good measure to select signal candidates.
The quantity of distinct tracks within an event, referred to as `nTracks`, is often utilised as a proxy for the event's multiplicity. Due to the high number of tracks produced during a proton-proton collision, further algorithms have been developed to reduce reconstruction errors. Next to the neural network against fake tracks [161], this includes the CloneKiller [162,163], removing duplicated tracks, called clones. If two tracks share more than 70% of their hits, only the track with the better quality is kept.

Tracking performance

The momentum resolution and efficiency of the tracking algorithms can be estimated with abundant decay modes. The $J/\psi \to \mu^+\mu^-$ decay is particularly well suited as the muons leave hits in the full tracking system and in the muon stations. The momentum resolution $\delta p/p$ can be approximated[7] as:

$$\left(\frac{\delta p}{p}\right)^2 = 2\left(\frac{\sigma_m}{m}\right)^2 - 2\left(\frac{p\sigma_\theta}{mc\theta}\right)^2,$$

where m is the invariant mass of the J/ψ candidate and σ_m is the Gaussian width obtained from a fit to the mass distribution. The second term is a correction for the opening angle θ between the two muons. The momentum resolution has been measured to vary between $\delta p/p = 0.5\%$ at low momentum to 1% at 200 GeV/c for Run 1 data [129].
The tracking efficiency can be determined with the *tag-and-probe* technique using again $J/\psi \to \mu^+\mu^-$ decays [158]. The total track reconstruction efficiency is found to be above 95% [158,159] for long tracks. A small reduction in the tracking efficiency is observed in 2015 due to the fact that the bunch spacing of the LHC was reduced from 50 to 25 ns,

[7] This approximation is only valid under the assumption that the muons mass can be neglected and the momenta of the muons are approximately equal

resulting in larger spillover effects in the OT detector whose readout window is larger than 25 ns [159].
While in Ref. [158, 160] the reconstruction of downstream tracks, which is used in Run 1 and Run 2, is described, no direct information about the downstream tracking efficiency have been published. An alternative approach for downstream tracks of long-lived particles [164] has been developed. They report a downstream tracking efficiency of about 76% for $\Lambda \to p\pi^-$ decays, which is in agreement with the old approach. The about 20% difference in the tracking efficiency between long and downstream tracks is a direct result of the missing information from the VELO for downstream tracks.

2.2.4 The particle identification

The production cross-section of beauty quarks in inelastic proton-proton collisions is two or three orders of magnitude smaller than the total inelastic cross section within the LHCb acceptance [89, 165, 166]. This leads to high track multiplicity and higher background levels (especially when compared to B-factories). Consequently, LHCb does not only need precise tracking of charged particles, but also has to identify their nature as well as identify neutral particles such as photons and neutral pions. The successful particle identification can either be based on the response of a specific subdetector, or on global likelihoods or neural networks which use information from all subdetectors to assign a particle type to a track.
As already mentioned, the particle identification for charged particles (beside electrons and muons) is mostly based on the response of the RICH detectors. However, in high-multiplicity collisions, identifying separate rings is challenging and the problem is tackeled using statistical tools. A specialised algorithm is employed to link reconstructed tracks to rings reconstructed within the RICH system and combined with information from the other subdetectors, assigning a likelihood to each track indicating its likelihood of being a particular particle type. The likelihood of each hypothesis is then compared to that of the pion (the most abundant particles at LHCb), and the difference in log-likelihood is denoted as `DLLX`. Typically, the differences in likelihoods between the kaon (proton) hypothesis and the pion one, known as `DLLK` (`DLLp`), serve as discriminating variables in the offline selection process[8]. Alternatively, a neural network, specifically a multilayer perceptron with one hidden layer, is trained independently for each charged particle species. This network integrates information from all subdetectors, including also the tracking system, and generates an output for each charged particle hypothesis (pion, kaon, proton, muon, electron, or deuteron), referred to as `ProbNNX` [167]. `ProbNNX` is a more potent discriminating variable compared to the difference in likelihoods as it uses more information. However, as discussed above, the separation between different hadron types is challenging in the low and high momentum regime, and even in the intermediate momentum range, the particle identification efficiency is not reaching 100%.
The muon identification involves several sequential steps [149]. Initially, a binary variable called `isMuon` is constructed, based on the information of a muon candidate penetrated

[8]The PID variables are known to be mismodelled in simulation and very few dedicated studies for the calibration of downstream tracks were available throughout this thesis. Therefore, to avoid introducing biases no such hadron PID variables were used as a majority of the decays studied in this thesis rely on downstream tracks.

through the calorimeters and iron plates. This variable exhibits high efficiency for muons with momenta exceeding 3 GeV/c (the threshold required to cross the calorimeter material), thereby reducing hadron misidentifications already to a percent level. However, in order to transverse all muon chambers the muons need to have a momenta of at least 6 GeV/c. In the subsequent step, the trajectories of charged particles from the tracking system are extrapolated to the muon stations with high precision. The pattern of hits in the muon stations around the extrapolation is then analysed to generate a discriminating variable, `muDLL`, representing the logarithm of the ratio between muon and non-muon hypotheses. Subsequently, `muDLL` is incorporated into a combined likelihood function alongside information from the RICH system and the calorimeter. This yields another discriminating likelihood-based variable, `PIDmu`, representing the logarithm of the ratio between muon and pion hypotheses, which is employed as a PID discriminator in the offline selection processes of analyses involving muons. Alternatively, the output of a neural network, `ProbNNmu`, can also be used.

2.2.5 Data processing chain

The LHC has a specialised bunch filling scheme to provide optimal conditions for each of the four experiments and ensure they can cope with the provided amount of data. The maximum bunch crossing rate at the LHC is 40 MHz, but was reduced to 15 MHz during Run 1 and increased to nearly 30 MHz for Run 2 at the LHCb experiment [168]. This corresponds to the production of 3 Tb/s of data. While not all of these data are interesting from a physics point of view and therefore storing all of them would be a huge waste of resources, there was no technology available to process such a big amount of data at such a high rate. Therefore, the recorded raw data pass through a dedicated data flow, called trigger, to optimise the data-taking efficiency and the purity of the recorded signals. At the end of this processing chain, the amount of data is reduced to a level that can be written to tape.
The trigger has two dedicated stages, a hardware stage (L0) reducing the data rate to about 1 MHz, the output of which can be processed by computing farms (event filter farm). The L0 is based on a very simple selection and uses information about the collected energy provided promptly by certain subdetectors. It is then followed by the two step software (high-level trigger, HLT) stage. In the first stage, HLT1, the events are partially reconstructed to apply selection on tracks. In the next step, HLT2, the full event reconstruction[9] is performed, enabling complicated and very specific selections. During Run 1 the computing resources limited the access to calibrations leading to an offline reconstruction to improve the output. Before the start of Run 2, several upgrades on the hardware and software, including the introduction of a buffer, enabled to run a fully automated detector calibration at the HLT stage, making the offline re-reconstruction obsolete in the default reconstruction. In addition, the trigger strategy was changed to account for the higher data rate with increase center-of-mass energy of the LHC. The Run 1 and Run 2 workflow of the trigger are displayed in Fig 2.13 and an overview of their performance can be found in the Refs. [159, 169, 170].

[9] This stage is also sometimes called online reconstruction.

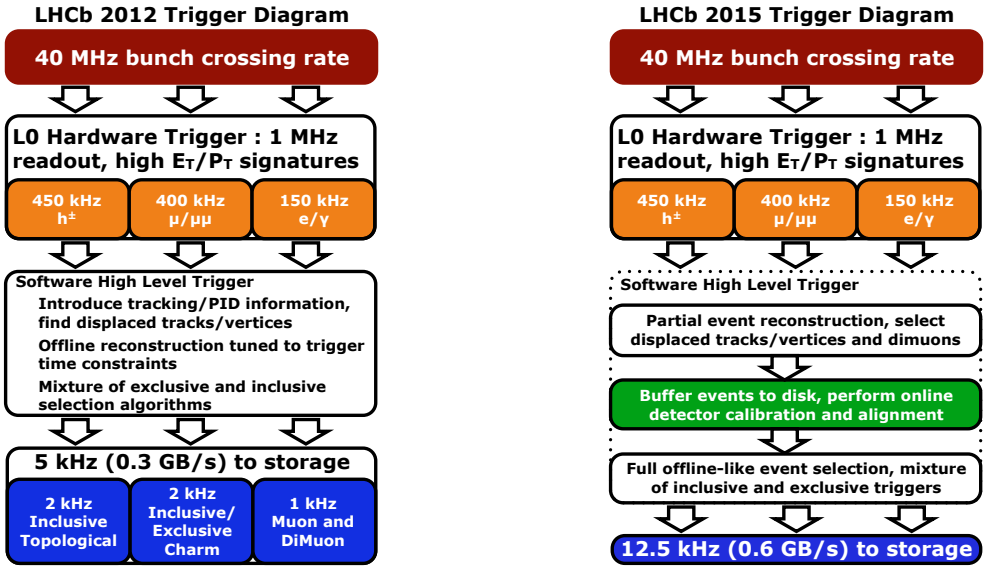


Figure 2.13: Trigger diagrams for the data flow at the LHCb experiment for Run 1 (left) and Run 2 (right). Taken from Ref. [171].

L0 - Hardware trigger

The L0 hardware trigger enables to reduce the data rate from the respective collision rate to about 1 MHz in a time window maximal 4 μs [169]. Due to the high mass of b-hadrons, the decay products have usually significantly high energies, in particular, large transverse (with respect to the beam direction) momenta and transverse energies.
There are four different types of L0 trigger categories[10], the `L0Muon`, `L0Hadron`, `L0Electron` and `L0Photon` triggers. For the first trigger category a straight track in all five muon stations is searched for [159]. Since the muon stations are normally only reached by muons (which have lower abundance compared to hadrons at the LHC), the trigger threshold on the p_T estimated from the track direction is lower compared to other L0 thresholds. Additional, there is also the `L0DiMuon` trigger line, which checks if the product of the two highest transverse momenta is above a certain threshold. The other three trigger selections rely on the calorimeter system [159]. Adding information from the SPD and PS subdetectors enables to distinguish between hadrons, electrons and photons, as previously discussed. For hadrons, photons or electrons, a significant amount of transverse energy needs to be deposited in the calorimeter cells in order to pass the hardware trigger. For example, the transverse energy for an electron is calculated by combining 2×2 ECAL cells into a cluster:

$$E_T = \sum_{i=1}^{4} E_i \sin\theta_i, \tag{2.3}$$

where E_i is the deposited energy in each cell i and θ_i the angle between the z-axis and a line from the cell centre to the average proton-proton interaction point. For the electron trigger an event is passed on to the next trigger stage, if at least one candidate is above the threshold [172]. However, for muons if several candidates pass the threshold only the

[10] There are a few additional trigger lines are less significant and therefore here not explicitly discussed.

one with the highest p_T is used for the trigger.
The trigger thresholds for hadrons are significantly tighter to ensure the rate of the L0 trigger stays within the 1 MHz range. It is need due to the significant QCD backgrounds produced at hadron colliders. For similar reasons the L0 trigger vetoes events with high multiplicity, which would exceed the HLT1 trigger decision time [173], by requiring the number of hits in the SPD to be inferior to 600 in Run 1 and 450 in Run 2 for most of the triggers [173].
The L0 reconstructed transverse momenta and energies are far from being precise. Therefore, after passing through the full reconstruction the L0 cuts are rather smeared out than a sharp cut in the respective kinematic distributions. In addition, the trigger threshold have been adapted throughout each data-taking period to optimise the trigger selection according to changes in the beam conditions and instantaneous luminosity.

HLT - High-level software trigger

After passing the L0 trigger, selected events are handed over to the software-based trigger. Its purpose is to select beauty and charm decays by searching for signatures of displaced vertices and tracks. Therefore, it partially reconstructs the final states by performing a track and primary vertex (PV) fit [152] and applies a selection on the track quality, particle momenta and PV displacement. Muons are already identified in the HLT 1 step, but events involving photons require calorimeter information that are not available at this step. The HLT 1 trigger reduces the rate to $\mathcal{O}(10)$ kHz.
In HLT2, the events are reconstructed using the full detector information including RICH and the calorimeter system. Furthermore, a more complete track reconstruction is performed. With the start of Run 2 it was possible to save events with positive HLT1 decisions to a buffer. Consequently, the HLT2 reconstruction and selection could be performed when the data-taking was on hold, *e.g.* in between the LHC fills or during technical shutdowns [159]. This optimisation allowed to lower the p_{T} thresholds on tracking algorithms and the implementation of a dedicated tracking algorithms for long-lived particles [170]. With the selection from HLT2 the data rate was further reduced to 5 kHz in Run 1 and 12.5 kHz in Run 2, which were the respective limits for writing to tape.
The HLT2 trigger selection uses two different strategies called *exclusive* and *inclusive*. The *exclusive* strategy is optimised for one specific final state, with the obsolete additional offline reconstruction as done in Run 1 these events where passed through the *Turbo stream*. This stream significantly reduced the amount of saved data by saving only the signal particles instead of the full event (except for events used for hadron spectroscopy). In addition, this stream does not save the raw detector information needed to reprocess the reconstruction. The other type of trigger selections uses an *inclusive* strategy by combining several decays with similar decay topologies and final states, the output is then saved in the *full stream* conserving all raw detector information. These inclusive triggers typically rely on signatures of b-hadron decays, such as a displacement of the decay vertex from the PV, high transverse momenta, or sometimes a presence of (di)leptons in the final state. In order to reduce the output from a certain inclusive trigger lines to a manageable level for user analysis, an additional selection step is applied, called *stripping*. The output data files are grouped in smaller *streams* depending on their final state and selection patterns. Therefore, the analyses presented in this thesis use inclusive trigger lines with additional stripping lines being applied on top.

Chapter 3

LHCb Upgrade I: An almost brand new detector

During the Runs 3 and 4 of the LHC, the instantaneous luminosity at the LHCb experiment is planned to be increased from $4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ to $2 \times 10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. This will increase the visible proton-proton interaction to about 5 and consequently lead to a higher pile-up.
In order to keep delivering high quality data, the LHCb experiment had to undergo a significant upgrade. Especially the radiation damage of the vertexing and tracking system started to affect the data quality. However, the bigger challenge is the handling of the produced amount of data with an increase in luminosity. Fig. 3.1 shows, that with an increased value of instantaneous luminosity, the relative trigger rate for hadronic decay channels starts to saturate. This limitation is caused by the tight calorimeter thresholds at the hardware trigger stage, required to reject the abundantly produced low-energy hadronic backgrounds.

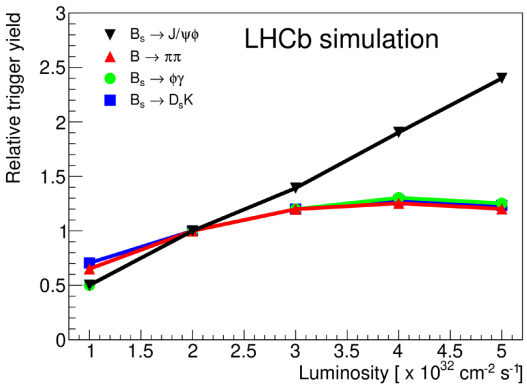


Figure 3.1: Relative trigger yields as a function of instantaneous luminosity. Taken from Ref. [174].

Consequently, the LHCb collaboration decided to fully remove the hardware stage of the trigger system and operate a fully software-based trigger during the next runs of the LHC

to keep up with the data acquisition rate. As a result, not only the complete vertexing and tracking system was replaced, but also the readout systems of all subdetectors in order to enable a readout at the nominal LHC bunch-crossing rate of 40 MHz. Together with other upgrades, over ninety percent of the LHCb detector components have been replaced, leading to an almost brand new detector.
In the following sections, a general overview of the LHCb Upgrade I, a more detailed introduction to the Scintillating Fiber Tracker (SciFi) and my contributions to its commissioning are presented.

3.1 Overview of the upgrade

The LHCb Upgrade I detector [119, 174] is displayed in Fig. 3.2.

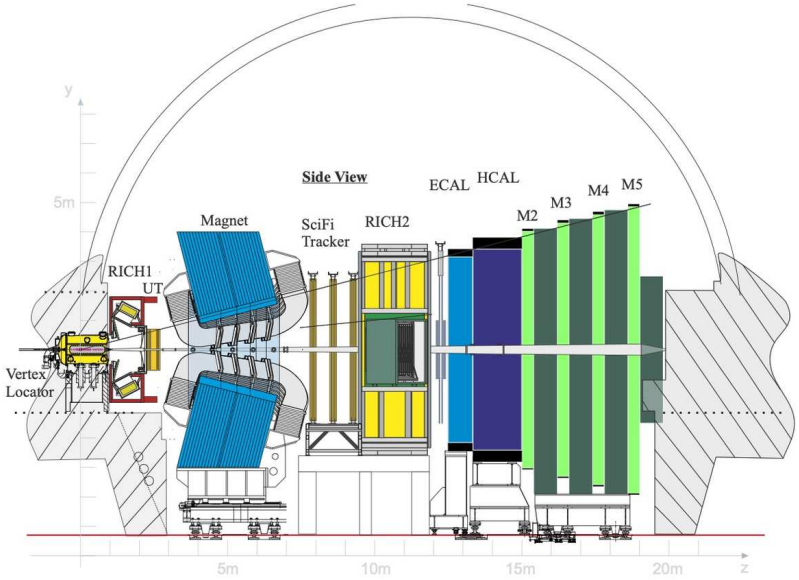


Figure 3.2: Schematic representation of the upgraded LHCb detector, called Upgrade I. Taken from Ref. [174].

During this upgrade, the first muon station M1 as well as the SPD and PS systems of the calorimeter have been removed. This decision was made as the occupancy in the scintillator used in the SPD and PS could not be read out in a meaningful matter. Similarly, the M1 station was removed since it was needed due to the new software-based trigger system. In their place a polyethylene shield was installed to shield the RICH2 and SciFi detectors from neutron backslashes from the calorimeters.
The optics of the RICH1 detector have been redesigned and exchanged to reduce the peak occupancy in the central region while enlarging the imaging area. The main upgrade of RICH1 and RICH2 was the exchange of the Hybrid PhotoDetectors by multi-anode photomultipliers to enable fast readout [174].
The biggest detector upgrade was done for the tracking system. The new vertex detector uses silicon pixels instead of silicon strips, resulting in a finer granularity. This change in technology enables to cope with the higher multiplicity and prevents an increase of

the ghost rate [174]. The modules are again split into A- and C-side, but constructed as L-shape instead of half circles. They can get as close as 5.1 mm to the beam when the detector is fully closed. The number of module layers remains at 26 and the position was reoptimised to enhance the impact parameter resolution, which is improved by 40% for tracks possessing a first hit at $z > 300\,\text{mm}$ [175]. In addition, the design tracking efficiency for long tracks with a momentum of at least $5\,\text{GeV}/c$ increases from 96% in Run 2 to 99% in Run 3 [175].
In addition, the TT stations are replaced by the Upstream Tracker (UT), aiming to reduce the ghost rate and to improve the momentum resolution [176]. With respect to the TT, a finer granularity of the silicon strips and a stronger radiation hardness is facilitated.
A schematic representation of the new VELO (top) and the UT detector (bottom) can be found in Fig. 3.3.

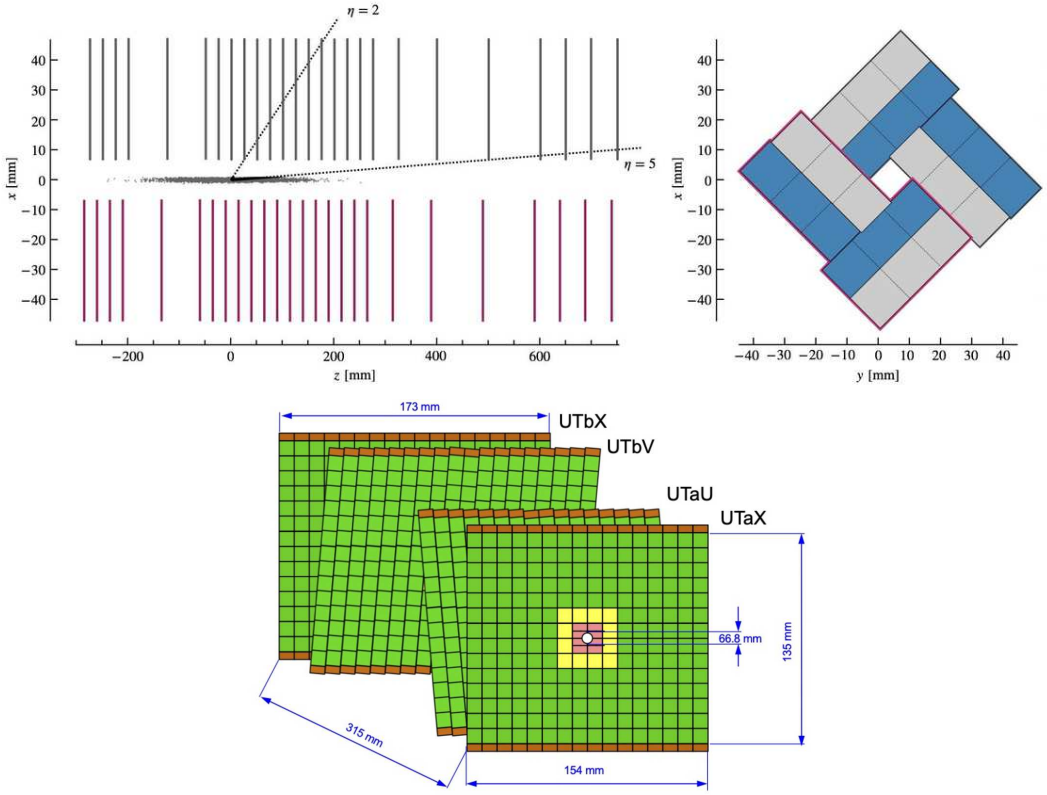


Figure 3.3: Schematic representation of the upgraded LHCb VELO and UT tracking detectors. Taken from Ref. [174].

The last new subdetector, the Scintillating Fiber Tracker (SciFi), replaces the IT and OT of the tracking stations. As the name suggested, the silicon strips and drift tubes have been replaced by scintillating fibres. The detector design will be discussed in detail in the next section.
As explained in the beginning of this chapter, it is necessary to remove the hardware stage and move to a fully software-based trigger system to achieve similar efficiencies for hadrons and leptons compared to Run 2. The new trigger strategy and offline analysis architecture for Run 3 is displayed in Fig. 3.4. The track reconstruction begins in a GPU-based HLT1 event filter, implemented in the Allen software [177], which reads out the

full detector. The HLT1 trigger decision is based on partially reconstructed events, this includes the attachment of bremsstrahlung photons, crucial for *e.g.* $b \to s\ell^+\ell^-$ analyses with electrons in the final state. After an event is selected, it is passed on to a buffer, where the real-time alignment and calibration is performed. The calibrated events are then passed to HLT2 trigger stage, operated with the Moore software. In this step, the full event reconstruction is carried out before the HLT2 trigger selection is applied. If the HLT2 trigger fires, the events are saved to tape. Similarly to Run 1 and Run 2, inclusive HLT2 lines are passed through an additional offline selection called sprucing. In order to run the trigger selection fully on software, the front-end electronics of all subdetectors have been upgraded to Field Programmable Gate Arrays (FPGAs), that can process the data at the full LHC bunch crossing rate.

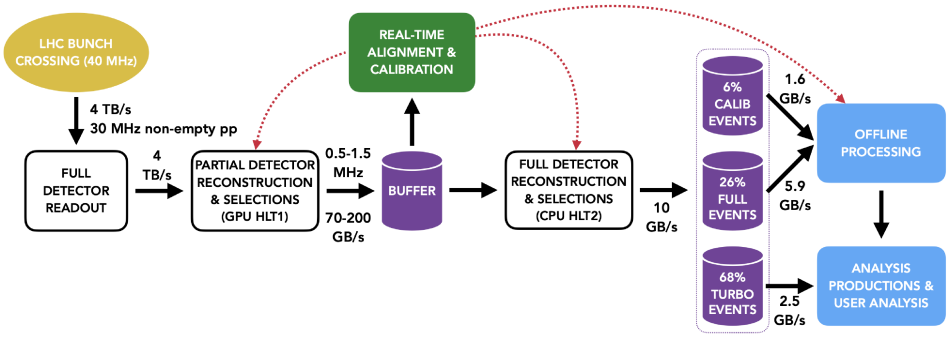


Figure 3.4: Schematic representation of the upgraded LHCb trigger system. Taken from Ref. [174].

3.2 Upgrade of the downstream tracking system

In order to understand the design choices of the upgrade of the downstream tracking system, it is useful to revisit the design of the Run 1 and Run 2 tracking stations in more detail.

Each of the three tracking stations consisted of four module layers, where the two middle layers of each station where tilted by $\pm 5°$ to enable a two-dimensional hit reconstruction. The outer part, the OT tracker, had a size of approximately $6 \times 5\,\text{m}$ and consisted out of 2.4 m long straw tubes filled with a mixture of argon, carbon dioxide and oxygen [138]. They were combined in modules of 4×64 tubes with two of the 2.4 m modules stuck onto each other to cover the full height of the tracking stations. The resulting average hit efficiency was estimated to be 99.2% and the position resolution was measured to be close to the design value of 200 µm [129]. During Run 2, utilising novel real-time calibration methods, as described in Ref. [140], the resolution could be further improved to 171 µm. The central region of approximately $11\,\text{m}^2$ was covered by the IT consisting of silicon microstrip sensors. Individual silicon sensors were 7.8 cm in width and 11 cm in height and located in detector boxes placed in a plus-shape around the beampipe. About 20 % of all charged particles were passing through the active area of the IT, that was only covering about 1.3 % of the geometrical acceptance. The hit efficiency was estimated to be 99.7 % and the resolution was measured to be 58 µm [129].

While it was believed that the lifetime of the tracking station could be extended, it was found that an integrated luminosity beyond $10^{33}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ would require a finer detector granularity in order to achieve the same momentum resolution and track efficiency for beauty and charm hadrons as in Run 2. Different scenarios including an increased area of the IT along with shorter OT modules in the central part were considered. But one of the main caveats was the fact that the drift times of almost 40 ns in the OT prevent a readout of such a design at the full LHC bunch crossing rate of 40 MHz. In the end, the decision favoured a scintillating fibre tracker as a cost-effective, light and uniform detector covering the full detector area of $360\,\mathrm{m}^2$[1].
The goal was to build a detector with a hit efficiency of 99% while keeping the noise rate down to 10% and providing a single hit position resolution in the xz plane below 100 µm [174,176]. In addition to the needed granularity and a radiation hardness sufficient for at least $50\,\mathrm{fb}^{-1}$, the material budget needed to be low enough to limit further multiple scattering and secondary particle production. A SciFi tracker using scintillating fibre technology with silicon photomultiplier (SiPM) readout fulfils these requirements. While the scintillating fibre technology is used in a wide range of applications, this is among the first times it is used in an high-energy physics environment for tracking after the D0 detector [178]. Schematic front and side views of the detector can be found in Fig. 3.5.

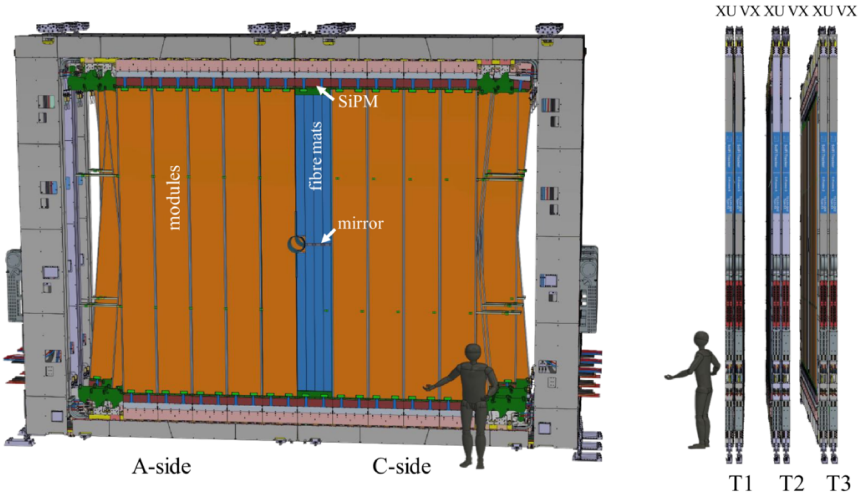


Figure 3.5: Schematic front and side views of the SciFi Tracker. Taken from Ref. [174].

The general concept of the tracking stations with three stations and twelve layers, including the tilt of the inner ones by by ±5°, is preserved. Each station consists of four C-frames carrying two layers on each side of the beam pipe. The first two stations hold two times five modules in each layer, while the last station has two times six modules to increase the detector's coverage. The modules are constructed in a honeycomb with six layers of fibres and protected by a carbon-fibre outer casing. Each module consists of two about 2.4 m long fibre mats that are separated by mirrors in the middle and read out with SiPMs at the end. The scintillating fibres are made out of polystyrene doped with p-Terphenyl dye to speed up the polystrene scintillating process in a feasable range for

[1] In the end the SciFi detector has a slightly reduced detector area of $320\,\mathrm{m}^2$.

the LHC bunch crossing rate. In addition, a tetraphenylbutadiene wavelength shifter is added improving the photon detection efficiency. Lastly, the core is surrounded by two claddings with lower refraction indices to enable total reflection and therefore ensuring any photon independent of its production point can reach the SiPMs. In the end, a total of 524 288 SiPMs channels need to be read out. The SiPMs are read out with frontend boards[2] (FEB) connected to the backend for the data acquisition. The center piece of backend electronics is the PCIe40 card, which was designed for the LHCb experiment to enable readout at the 40 MHz bunch-crossing rate of the LHC [179]. It uses an FPGA with three different firmwares, the first one is TELL40 dedicated to receiving the data. The second one is SOL40 helping to perform the communication from and to the FEBs and therefore the control of them. The is the SODIN firmware, receiving the 40 MHz LHC clock and distributing it to the other cards together with control commands it generates. The half-ROBs themselves consist of three types boards with dedicated functions. The PACIFIC board reads, amplifies, and finally digitises the data obtained from the SiPMs. Afterwards, the four Clusterisation boards perform the zero-suppression and group the signals in clusters. The four Master board provide the control to the other boards, as well as the power and clock. In addition, they are responsible for serialising and sending the data and are used for monitoring of the detector [174].
The single hit position resolution of the modules have been found to be about 64 µm during a test beam campaign[3] [174].

3.3 Commissioning of the SciFi detector

For the successful commissioning of the SciFi detector, it is not only crucial to have a strict quality control during the production process and assembly of the different parts, but also during the final assembly of components and the installation in the LHCb cavern. Optical fibres are not only used to collect the signals from the charged particles passing through the SciFi detector, but also to transport the signal from the detector in the cavern up into the data center.

3.3.1 Commissioning during the long shutdown 2

It is important to check these fibre cables carefully and ensure no dirt or scratches on the connectors of them reduce the quality or even completely stop the signal. In the path between the FEBs of the SciFi and the data center are three places, where connectors bridge two different optical fibres. The first ones are the Lucent Connector (LC) of single fibres going into the FEBs. These fibres are routed within the C-frames and end at the red patch panels of the C-frames. They have been combined into one fibre just below the FEBs and end in Multi-Fiber Push-On (MPO) connectors. From here an intermediate 27 m fibre cable is routed to the *bunker* patch panels below the SciFi detector, that allow a connection to the MPO connectors of the data and control fibres going all the way up to

[2] Alternatively the frontend boards are also called readout boards (ROB). They are split in two units called half-ROBs (H0, H1), each managing eight data fibres and a recieving (RX) and transmitting (TX) control link.

[3] As of the writing of this thesis, no independent performance paper has been published yet, quoting the single hit resolution during the Run 3 data taking.

the surface and into the data center, where the collected data will be processed. The data center itself has also an LC patch panel, which will not be covered in this thesis. Pictures of the optical fibres connected to the bunker patch panel (left), the outside view of the C-frame patch panels with the cable routers in red and yellow in the cavern (middle) and during the connection to the FEBs in the assembly hall (right) can be found in Fig. 3.6.

Figure 3.6: Pictures of the bunker patch panel of the C-side (left), the red patch panel directly on the C-Frames (middle) and optical fibres connected to FEBs in the assembly hall (right).

Especially, dirt or scratches on the center of a connector of an optical fibre can reduce its light transmitting efficiency, leading to a loss of data. In addition, during the installation processes the delicate optical fibres can be bent too much and therefore break inside. While this is not visible from outside, it effectively leads to a complete loss of transmission of signals through this fibre. Therefore, each fibre is handled with extreme caution and all connectors are checked when installed in the C-frames, and later again the MPO connectors at the C-frame patch panel before they are connected with the cables in the cavern. They have been checked with an optical microscope and luckily only very rarely scratches have been found on the connectors. Very few of the checked connectors had a scratch directly through the center, where the light is transmitted. In that case the fibre was exchanged. If dirt was found, the connector was carefully cleaned until all the dirt was removed. Fig. 3.7 shows an example of a scratched fibre (left) and dirt as it was found on some connectors (right). While for the latter the dirt was also not in the center it can pass onto the optical center, when the detector has to be moved.

During the construction process in the assembly hall, it was possible to remove individual fibres from the C-frames, if they showed damaged. However, since this is not possible in the cavern after the installation, each half-ROB is in principle reached by two spare cables in the full chain from the data centre to the FEB boards in the C-frame, to use as an replacement if a data or control link fibre is broken[4].
After a successful installation during the long shutdown 2 the commissioning of the SciFi relying on collision data started with the start of Run 3.

[4]It should be noted that the fibres connecting from the bunker area below the SciFi detector to the data center are shared with the ECAL, reducing the number of available spares.

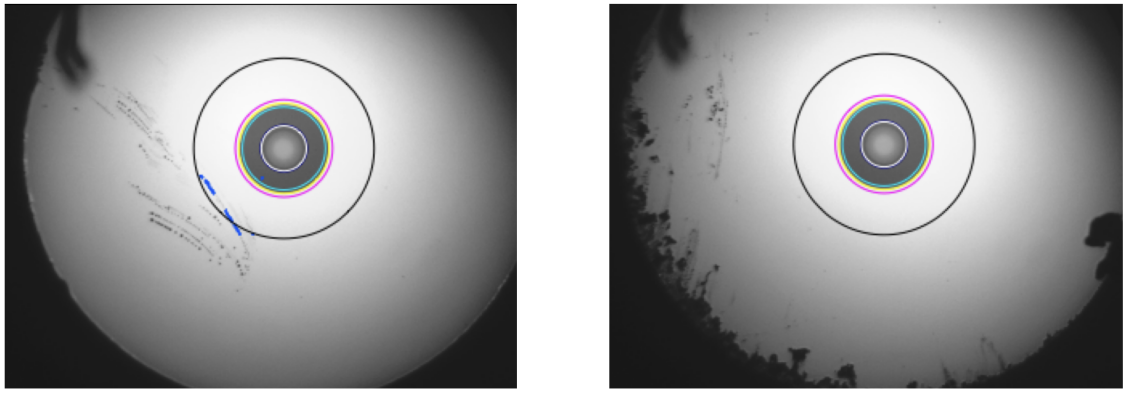


Figure 3.7: Pictures taken during the optical inspection of the optical fibres during the long shutdown 2. The left picture displays an LC connector with scratches, while the right pictures shows an LC connector with dirt at the outside of the fibre. The grey filled circle in the middle of the connector shows the active center of the optical fibres. The grey, pink and green circular lines are from the analysis software of the microscope.

3.3.2 Failure investigation during the YETS 2023

While the beam time in 2022 was mostly used to commission the LHCb Upgrade 1, especially during the data taking in 2023, several issues occurred with individual data links and whole half-ROBs. One example of an observed technical issue across all tracking detectors was, that some data links were suddenly sending corrupt data, that showed up as decoding errors in the timing and fast control (TFC) counters of the TELL40 [180]. These errors could lead to complete communication loss, which is problematic for the data quality. A first solution was found, by restarting the run, but the errors were occurring too frequently to make this a sustainable solution. Therefore, a deeper investigation was performed and revealed that these errors were linked to control links with high Forward Error Correction (FEC) counts caused by infrequent bit signals. This problem was solved by sending additional individual bits to overcome the infrequency. While this issue was solved, some of them were so severe and needed deeper investigations, that the connected half-ROBs and data links have been excluded from the nominal data taking for a better data taking stability. The excluded half-ROBs and data links are displayed in Fig. 3.8. In order to better understand the observed issues, a dedicated intervention was planned for the end of the year shutdown (YETS) in 2023.

For a better orientation, Fig. 3.9 displays the naming convention for the SciFi detector. It will be used in the following, when referring to individual ROBs.

Before discussing the dedicated intervention on all the excluded data links and half-ROBs, it is important to discuss a technical issue affecting all LHC experiments, including the SciFi detector at the LHCb experiment. During the long shutdown 2 it was discovered, that the versatile transceivers (VTRx)[5], developed at CERN, showed temperature-dependent problematic behaviours leading as far as the total loss of the communication. It was discovered, that due to the increase in temperature the epoxy glue of

[5]They are commonly used by all LHC experiments for the communication of upgraded frontend electronics due to their radiation and magnetic field tolerance.

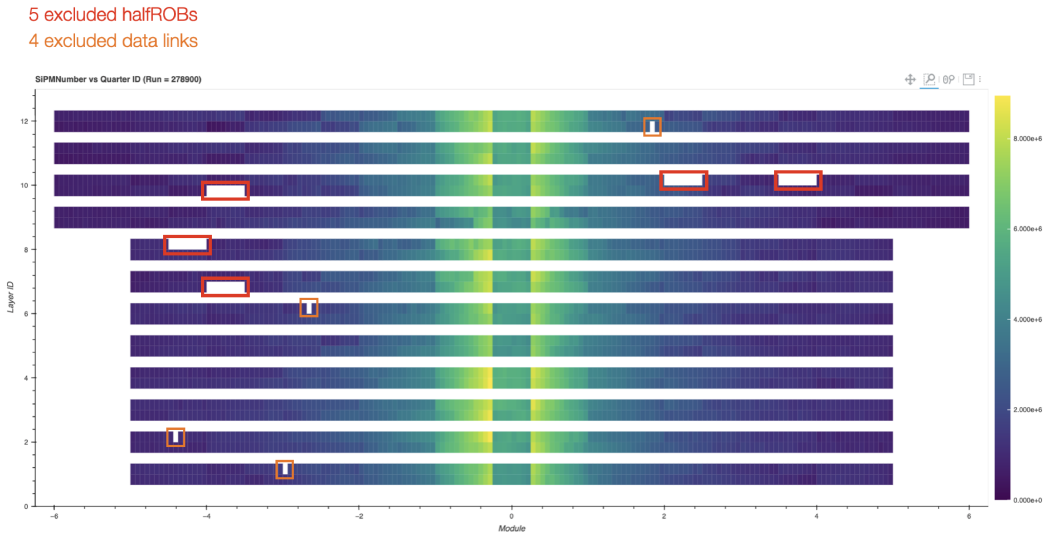


Figure 3.8: SciFi hitmap reference with the excluded data links (orange) and half-ROBs (red) extracted from the LHCb online monitoring in November 2023 [181]. The hitmap displays the number of clusters per SiPM. The zy-axis is the module number, with negative numbers corresponding to those on the C-side; the y-axis is the layer index. The horizontal band in the middle is caused by the beam pipe.

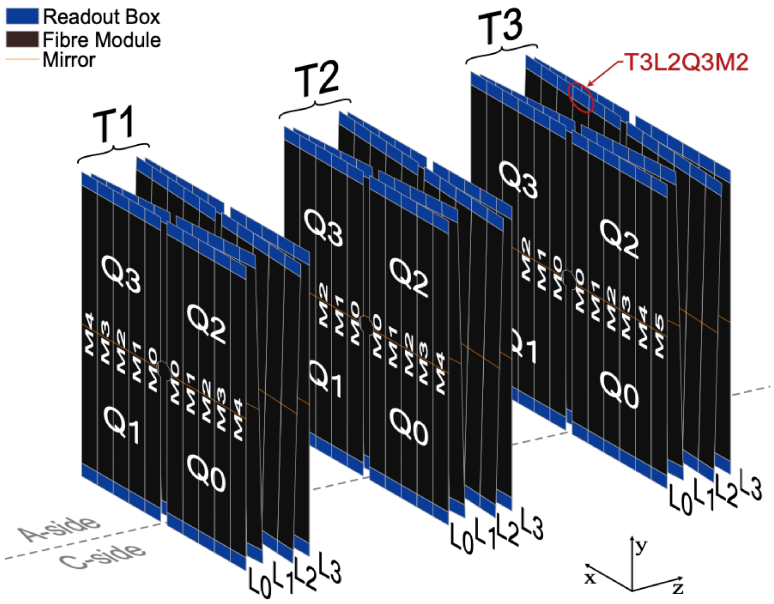


Figure 3.9: Schematic view of the SciFi Tracker with the naming convention overlaid. An example ROB and its location identifier is highlighted in red. Distances between the layers (L) and stations (T) are exaggerated for better visibility. The inner layers per station (L1 & L2) are tilted by $\pm 5°$ to obtain the x-coordinate of detected hits. Taken from Ref. [182].

the VTRx started to outgas, leading to condensation on the communication fibres.

The problem was communicated and investigated by the CMS HCAL team [183]. They were able to prove that active cooling of the VTRx can mostly stop the outgassing. It was later confirmed that the issue was caused by an insufficient curing time by the production company of the VTRx supplying CERN. Since the issue was discovered during a late stage and the impact of the Covid-19 pandemic had caused severe delays in the installation schedule, it was decided to install cooling where possible to already installed detector components [184]. The SciFi detector uses two VTRx for the communication with the Master boards of each FEB. When the issue with the VTRx was discovered the C-side of the SciFi was already installed. The detector is in principle operated water-cooled at room temperature, however to reduce noise the SiPM are actively cooled to $-50°C$. Therefore, it was assumed that the VTRx might not get too warm during the operation of the detector. For the A-side of the SciFi, it was possible to anneal the VTRxs to cure the epoxy glue before the installation in the cavern.
This is most likely the source of the dirt found on the optical fibres dedicated to the control of the FEBs and therefore in direct contact with the VTRx. It should be noted, that the unannealed C-side of the SciFi showed more problematic half-ROBs and data links as it can be seen in the left side of Fig. 3.8.
The first excluded half-ROB was `T2L2Q0M3H0` due to an DC-DC converter failure. In order to repair the converter the FEB had to be removed and reinstalled. Before reconnecting the data and control fibres, they underwent an optical inspection revealing VTRx outgasing. An example for light (left) and strong (right) VTRx outgasing covering the LC connectors of the optical fibres can be found in Fig 3.10.

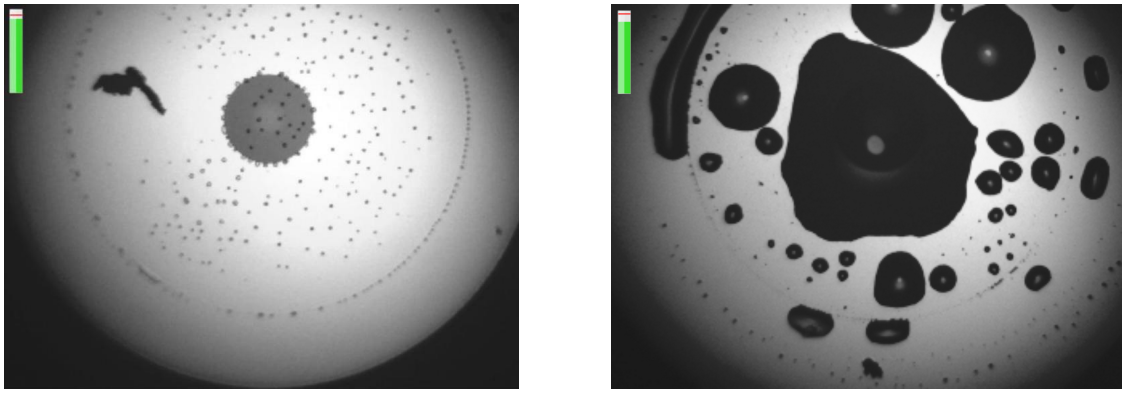


Figure 3.10: Pictures taken during the optical inspection of the FEB `T3L3Q1M1`. The light grey shows the body of the optical fibre and the smaller dark grey area in the middle is the active center of the fibre transmitting the light. On the RX control link of H0 half-ROB signs of light outgasing of the VTRx was found together with some other dirt. For the RX control link of the H1 half-ROB heavy outgasing both found. Both fibres have been cleaned before reinstallation.

Since this was the first time the outgasing was observed for an SciFi FEB, a thorough investigation was triggered. All FEBs in the same quadrant have been tested and for two more FEBs signs of outgasing were found, while no failure was found during the data taking. The next investigation was performed for the data link `T3L3Q1M1H0D2`, which had frequently sent errors with the start of a run, a reset or configuration change of the ROB.

Therefore, it was decided to also exchange the corresponding FEB. After the exchange the optical inspection revealed heavy outgasing on the H0 half-ROB, while the H1 only showed mild signs of outgasing. The half-ROB T3L1Q3M2H0 was excluded, since no optical power was received at the SOL40 card, leading to a loss of control of the half-ROB. A dedicated optical investigation with an optical source was performed to investigate possible broken fibres as a source of the communication loss. It was found that the fibre trunk between the data center and the bunker patch panel in the cavern is broken and it was replaced with a spare. Similarly, for the half-ROB T3L1Q3M3H1 an unstable control link was found. The same optical investigation did not reveal a broken fibre. However, during the optical inspection some smaller dirt patches were found that have been removed. Afterwards, the control links between the H0 and H1 half-ROB have been swapped for some tests. After the tests they were swapped back and a stable communication with the half-ROB was observed. The half-ROB T2L3Q2M4H1 was continuously getting into an error state when the ROB was reset, consequently it was excluded. An inspection of the optical power and the fibre connectors was performed. However, during this test the receiving control links could not be tested due to some other on-going work. Afterwards the controls were cleaned and swapped between H0 and H1, leading to stable performance. After they were swapped back severe signs of outgasing were found. This is expected to have caused the problems observed during the data taking. The last half-ROB under investigation was T3L1Q0M3H1, the previously discussed issue with high FEC counts could not be fixed by sending individual bits for this half-ROB. Therefore, it was excluded due to the control link instability. Optical inspection of the fibres for the connectors and with a source was performed with the same caveat of not being able to access the receiving control link as for T2L3Q2M4H1. The swap of the H0 and H1 led to stable performance and again after swapping them back signs of outgasing were found on the optical fibre. After cleaning the fibres, the half-ROBs are performing stably. The data link T1L0Q2M4H1D7 was constantly sending decoding errors, that vanished after cleaning the optical fibres at the FEB. For the last two data links only a partial success was obtained. The data link T1L1Q2M4H1D1 showed zero optical power at the TELL40 card, therefore the RX cable was inspected and showing signs of VTRx outgasing. The swap of the control links showed an improvement of the stability. However, after the reverse swap the data link remains slightly unstable. Therefore, it is excluded in the beginning of the data taking till further investigations lead to a complete stability. The last data link T2L1Q2M2H0D5 was excluded since no optical power was observed due to a problem with the FEB. Unfortunately, it was inaccessible during the YETS and therefore not exchanged.
Therefore, the data link remains excluded for the data taking in 2024. Beside these last two data links, it was expected that after the intervention all other excluded data-links and half-ROBs can be used in the data taking of 2024. The top plot in Fig. 3.11 shows the hitmap of the SciFi detector during a test run for the FASER experiment [185] on the 30th of march. It can be seen that the two remaining excluded data links T1L1Q2M4H1D1 and T2L1Q2M2H0D5 are shown up in white on the left side of the plot. During a previous noise run also a third data link had to be excluded due to instabilities. In addition, the exchanged FEBs, T2L2Q0M3H0 and T3L3Q1M1H0, can be identified by the two yellow areas in the middle left (close to the excluded T2L1Q2M2H0D5 data link) and middle top right. This is a result of the calibration still needed to be updated after the FEB exchange.

Finally, the bottom plot of Fig. 3.11 shows the hitmap of the SciFi detector after the missing calibration during one of the first stable beam runs in beginning of April 2024.

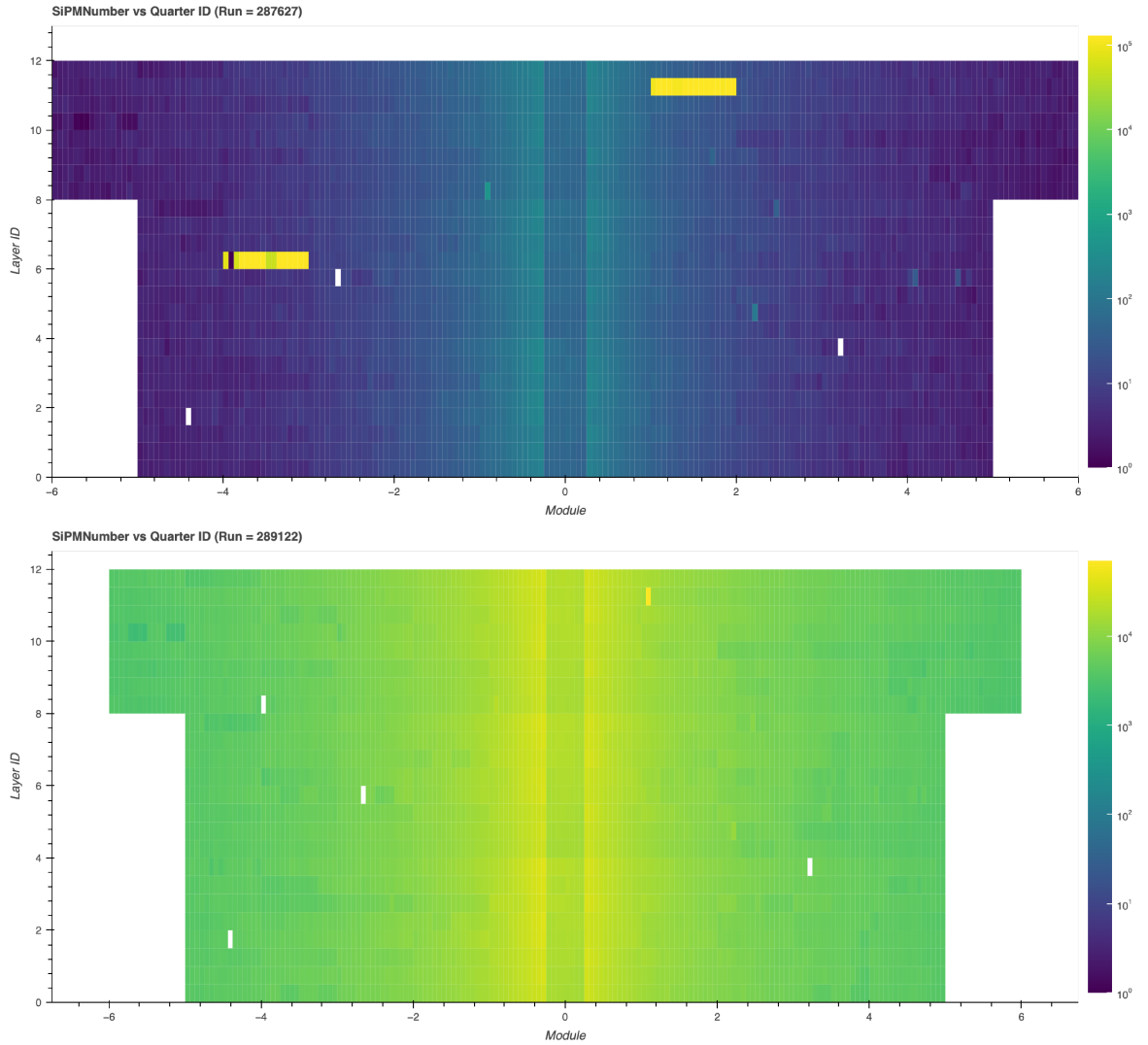


Figure 3.11: SciFi hitmap reference extracted from the LHCb online monitoring during a FASER test Run 1n March 2024 [186] after the performed intervention (top). SciFi hitmap during the first stable beams runs in beginning of April, when all FEBs have been properly calibrated (bottom).

To summarise, the intervention successfully helped to recover most of the excluded data-links and half-ROBs. It also showed, that the VTRx within the SciFi frontend boards outgas. Consequently, during the long-shutdown 3 a removal of all FEB that have not previously been annealed is needed. On a more positive side, it was also found that some fibres showed signs of the VTRx outgasing but were still operating as expected. Therefore, it is expected, that the SciFi detector can be operated without mayor losses due to the continued outgasing of the VTRx in Run 3.

Chapter 4

Key concepts for the presented analyses

The following section summarises information relevant for all four presented analyses in the subsequent chapters. In order to minimise the repetition of information, this section covers different topics. This includes the discussion of the datasets used throughout this thesis, calibrations of the simulation that is equivalent and the approach used for the signal probability distribution function (PDF).

4.1 Data samples

The LHCb experiment collected the Run 1 dataset (2011-2012), corresponding to about 3 fb^{-1} at center-of-mass energies of 7 and 8 TeV and the Run 2 dataset (2015-2018), corresponding to about 6 fb^{-1} taken at 13 TeV center-of-mass energy, where the b quark production cross-section is higher [89]. The center-of-mass energy and the collected integrated luminosities for each individual data-taking year are given in Tab. 4.1.

Table 4.1: Integrated luminosity collected in each year at the corresponding center-of-mass energy $\sqrt{s}$. Relative luminosity weight enables to combine the partial Run 1 and Run 2 datasets to correctly determine the efficiencies using simulation samples.

Year	$\sqrt{s}$, [TeV]	Luminosity, [fb^{-1}]	Relative weight
2011	7	1.	0.5
2012	8	2.	1
2015	13	0.33	0.15
2016	13	1.67	0.76
2017	13	1.71	0.78
2018	13	2.19	1

The data-taking conditions change over time: both due to ageing of the detector and varying accelerator conditions. Therefore, when multiple datasets are combined, knowledge of their relative luminosity values is used to obtain a reliable estimate of the detection efficiency. Thanks to the experience gained during Run 1 and the continuous improvement of

algorithms, it was possible to enhance the reconstruction, trigger and stripping selections. As a result, some of the cross-checks in the presented analyses have been performed with the Run 1 and Run 2 datasets separately to ensure full compatibility when extracting the final results.
Unless stated otherwise, the analyses presented in this analysis are performed by using the full Run 1 and Run 2 datasets to improve the statistical precision of the measurements.

4.2 Simulation samples

Monte-Carlo simulation of the decays of interest has multiple use cases in LHCb. It can be used to determine the properties of a certain background distribution in order to correctly incorporate it in the fits of the invariant mass distributions in data. The more important instance for the use of simulated samples is to simulate the properties and the invariant-mass shape of the signal decays. This is particularly important for blinded analysis, where key information like the invariant mass of the studied decay are concealed during the development of the analysis strategy to prevent bias and ensure objectivity in the results. Throughout this thesis, the four most important use cases of signal simulation are as signal proxy in multivariate analyses, to determine tail parameters of invariant-mass distributions for the decays of interest, and to determine reconstruction efficiencies and detector acceptance effects.
In order to correctly determine these quantities, it is important that the statistics of the simulation samples scale with the luminosity of the respective collision data samples. This is particularly the case, when several years of data-taking are merged and *e.g.* efficiencies are determined for the combined sample. In this case the luminosity values can be transformed into a weight (see Tab. 4.1) to ensure the individual simulation samples have the same relative importance as the individual collision data samples for the final result.

4.2.1 Truthmatching

When simulation samples are produced, the processes of the proton-proton collision, production of the b hadron and its decay are generated with different MC generators. The proton-proton collisions, the hard scattering processes and hadronisation process are produced through the *Pythia8* package [187], while the decay process is usually handled by the *EvtGen* package [188]. At this step the nature of each particle and their *true* values *e.g.* for the momentum are known. This stage is called *generator level*. In order to understand how the detector effects and the reconstruction software affect true physics distributions, it is important to pass the simulation samples at generator level through the model of the detector and the same software as real collision data would. At the end, *reconstruction level* simulated datasets are obtained, that technically look exactly the same way as real collision data. However, during the reconstruction process certain mistakes can be made, mis-reconstructing signal to make it look like background, or vice versa. In order to ensure that only correctly reconstructed simulated samples are used for the above described tasks, a dedicated procedure called *truthmatching* is performed. There are two different ways to perform the truthmatching in LHCb.
The first approach uses the *true* information from the generator level. Each particle is

assigned a *TRUEID*, which indicates their particle type, as well as a *mother* and *grandmother TRUEID* according to their origin. By requesting that a reconstructed particle has matching *TRUEID* to its generator-level counterpart, the simulation is automatically truthmatched. This approach can also be expanded to matching the mother and grandmother *TRUEID*s.
The problem is that the *TRUEID* truthmatching does not easily account for partially reconstructed signal candidates, when extra particles are present in the decay, which are not reconstructed. Therefore, a new dedicated algorithm was developed, called *background category* (BKGCAT). It categorises each particle as signal or a certain type of background by testing each candidate against a specific background hypothesis. The candidate is tested iteratively against lower and lower-priority background hypothesis, til one could not be rejected. Only if all the alternative hypotheses have been rejected, the process is classified as signal (category 0). Similarly to the reconstruction of real physics data, some of the background categories obtain poorly reconstructed signal events and are therefore also added to the truthmatching strategy. These is in particular the low-mass background category 50, where the reconstructed mass is slightly lower due to final-state radiation. In addition, sometimes there is remaining signal in the category 60. This is the so-called ghost category and includes poorly reconstructed tracks, where the reconstructed shared hits do not match with any generated hits. Therefore, these tracks are artifacts of the reconstruction software. As a result, this category was not used in the default truthmatching and only included for systematic studies. The truthmatching with the categories 0, 10 and 50 was used for the analyses presented in this thesis, unless otherwise stated.

Besides the truthmatching, a second procedure needs to be performed in order to ensure the simulation looks exactly like data. MC generators are based on theoretical predictions for the kinematic distributions in production and decay, as well as models of the detector. Both of these items can have imperfections, like missing material in the detector simulation or offsets in the used particle properties due to the precision of the available data at the time. Furthermore, a detailed simulation of the proton-proton collision is complicated and mismodelling can easily occur. In order to ensure that these and other mismodelling sources do not bias measurements, simulated samples need to be calibrated. Each chapter will have a dedicated explanation for their respective calibration procedure. However, the PID response and the hardware trigger response is calibrated in the same way in all presented analyses. Therefore it is included already in this chapter.

4.2.2 Calibration of the particle identification response

The response of multiple subdetectors is combined to provide high-level quantities enabling particle identification (PID), typically referred to as *PID variables*. Due to complexity of the relevant subdetectors, the PID variables are known to be mismodelled in simulation and therefore need a dedicated calibration. In all analyses presented in this thesis, no hadron PID variables are used. This choice was made as the selection of long-lived hadrons is most efficient without any PID requirements. In addition, the calibration of the PID from downstream particles (decay products of long-lived hadrons) is more complicated compared to typical tracks, because the response from the VELO and even

RICH1 is often not available. Therefore, the utilisation of hadron PID variables was dispensed and PID selections have only been applied to leptons. In general, there are three different types of lepton PID variables that need to be considered:

1. The binary variables like `isMuon` used in the stripping selection,
2. the likelihood-based `DLLX` variables testing against a pion hypothesis,
3. and the neural-network based `ProbNN` variables.

There are three different approaches available to calibrate the PID response.

1. The first one is the `MakePerfHistsRunRange` method of the `PIDCalib` package described in Ref. [189]. This method allows to compute the efficiency of a certain PID cut applied to a given track (this cut can be itself composed of several requirements), based on a tag and probe approach, and applying this efficiency as a per-event weight to the simulation samples;
2. The `PidGen` method, allows to resample the actual distribution of given PID variables in the simulation, by creating a new variable which shows a better data-simulation agreement. A weak aspect of this method is that it does not preserve the correlation between different variables for a given track;
3. The `PidCorr` method, is based on `PidGen` and uses dedicated simulation samples, provided with the `PIDCalib` package, in order to also account for the correlations between different PID variables for a single track, and therefore it preserves this correlation after the resampling. The methods 2 and 3 are described in details in Ref. [190].

All of these methods rely on the dedicated calibration samples recorded in parallel to the general data taking. In Run 2, this selection was performed at the high-level trigger stage directly, which allowed to minimise possible biases due to PID requirements present in physics triggers. Different samples are provided to cover different kinematic regions for each particle type, to ensure the best performance. The first important step before correcting the PID response in the simulation is the choice of the appropriate calibration sample. The calibration samples rely on abundant decay modes that can be selected without imposing PID requirements on at least one particle, referred to as the "probe" particle. The kinematic coverage of the calibration samples and the ones of the decays of interest should overlap sufficiently. The PID calibration methods assume that the PID response can be parametrised as a function of three variables: track momentum, track pseudorapidity, and event multiplicity. The latter can either be represented by the number of tracks reconstructed in the event, `nTracks`, or by the number of hits in the SPD subsystem, `nSPDHits`[1]. Both event multiplicity variables are known to be mismodelled, as they depend heavily on modelling of the *pp* collision itself, as well as secondary particles produced in material interactions. However, in previous analysis usually no significant

[1]It should be noted, that L0 triggers apply cuts to the `nSPDHits` variable, which also effects the `nTracks` variable. However, due to the non-trivial correlation between both variables, the `nSPDHits` selection does not create a sharp cut in the `nTracks` distribution.

trend was seen for PID efficiency of muons as a function of event multiplicity, and only a slight one for electrons. This is explained by the fact that the muon identification relies mainly on the muon stations, located at the very end of the detector, where the particle multiplicity is the lowest.
In order to perform the PID calibration, the dedicated calibration samples are statistically cleaned from the background contamination using the ${}_s\mathcal{P}lot$ method [191]. This relies on the assumption of the absence of correlation between the invariant mass, the PID observables, and the kinematic variables used to parametrise the efficiency: a systematic uncertainty is typically assigned for deviations from these assumptions. Afterwards, the PID cut efficiency can be determined with the tag-and-probe approach, allowing to extract the efficiency maps. The `PIDCalib` package provides a dedicated binning scheme optimiser to ensure the binning of each variable is fine enough to describe sufficiently the efficiency variations, and at the same time the statistics in each bin is sufficient to avoid large fluctuations.
One interesting fact is, that the `PidGen` method builds the distribution of the resampled PID variables based on the three variables discussed above. The `PidGen` method provides an option to automatically account for the mismodelling of the `nTracks` distribution by rescaling it in simulation by a certain factor.
It should be noted that during this thesis a new PID calibration package, called `PIDCalib2`, became available. It has not been used for the mass and production fraction analysis, as well as the analysis related to the charged Ξ_b^- baryon, as both analyses were already in review. In addition, so far also no attempt on the calibration on the electron PID variables has been performed. However, it is worth mentioning that the correlations between the invariant mass and the kinematic variables used for the calibration, as well as the different mass shapes for different bremsstrahlung categories[2] affect the performance of the ${}_s\mathcal{P}lot$ technique. It is important as the PID variables for electrons rely on bremsstrahlung information, as it is an distinct feature of electrons. The issues with the ${}_s\mathcal{P}lot$ technique can lead to an overestimation of the electron PID efficiencies. Therefore, the calibration of the electron variables needs to be checked carefully. The details on the PID calibration strategy are given for each analysis separately in the respective chapter.

4.2.3 Trigger calibration

Besides the PID calibration, the trigger response needs to be corrected for mismodelling in simulation. This mostly concerns the hardware trigger, where the thresholds were changed frequently to account for changes in the LHC performance, which is not properly reflected in simulation. During the time of writing this thesis, no calibration of the electron mode has been done and the L0 trigger strategy has been identical for all muon-based analyses in this thesis. Therefore, this section describes the calibration of the trigger response for all decay modes with muons. It can be performed by using the *TISTOS* method [192].
The trigger efficiency can be estimated from the fraction of events in a dataset triggered independently of signal (*TIS*). In such a dataset, there are also events that passed a

[2]Electrons radiate bremsstrahlung when passing the detector. A dedicated algorithm, called BremAdder, reconstruct these photons and attaches them back to the electron to improve its variables resolution. The algorithm has three classes: no bremsstrahlung photon added, one added or more than two added. More details can be found in Sec. H.

trigger on signal (*TIS&TOS*) selection, allowing to estimate the trigger efficiency for the signal. This assumption has been thoroughly tested and, therefore, the efficiency can be estimated with:

$$\epsilon_{TOS} = \frac{N_{TIS\&TOS}}{N_{TIS}} \tag{4.1}$$

where the denominator, N_{TIS}, is the number of events in the sample passing the TIS selection and the numerator, $N_{TIS\&TOS}$, is the number of those events that also pass the TOS decision of interest. The quantity ϵ_{TOS} needs to be computed on simulation and data in order to determine the final correction weight $w = \epsilon^{\text{data}}_{TOS}/\epsilon^{\text{sim}}_{TOS}$.
It is crucial that the same selection is applied to the data and simulation sample. In addition, the background has been statistically subtracted from the collision data sample by using the ${}_s\mathcal{P}lot$ technique [191].
Due to the statistical limitation of the decay modes studied in this thesis, the calibration has been performed with a much more abundant produced B-meson decay mode, $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$. The decay mode has been selected with the `Bu2LLK_mmLine` stripping line and the `L0Global_TIS` trigger requirement on the B^+ meson. In addition, only a J/ψ mass window ($3000 < m(\mu^+\mu^-) < 3150$ MeV/c^2) and a loose `ProbNNmu>0.1` cut has been applied to reject fake muons.
For the hardware trigger stage, two trigger categories need to be calibrated:

1. `L0Muon`: where either of the two muon candidates need to pass a certain p_T threshold that depends on the data-taking period,

2. `L0DiMuon`: where a certain threshold needs to be passed based on the product of the p_T of the two muons.

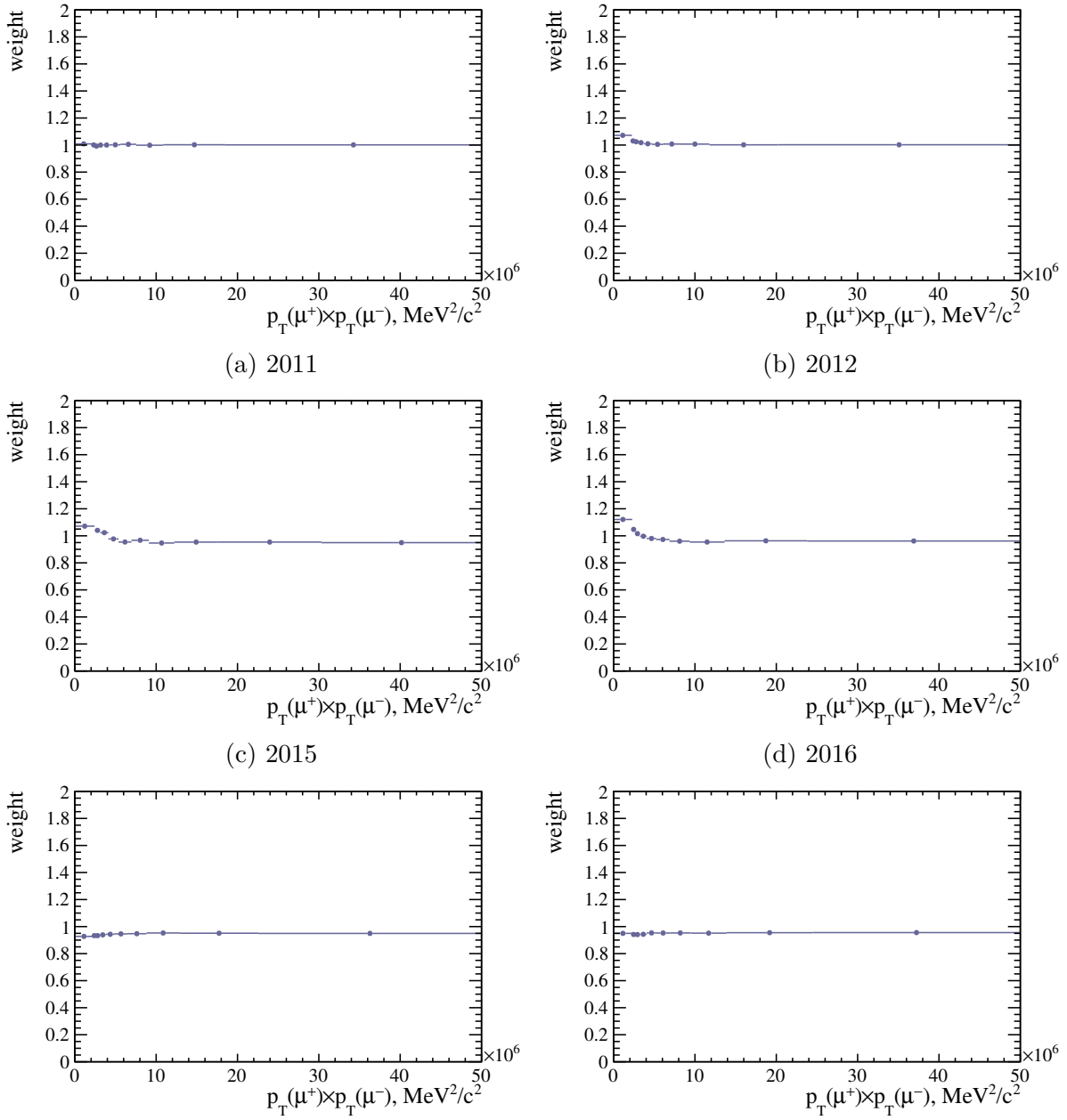
In addition, the trigger categories always include a cut on the number of hits in the SPD detector, `nSPDHits`, as it is a proxy of the event multiplicity. This allows to reject very high-multiplicity events that would slow down the subsequent software trigger. This `nSPDHits` variable is known to be mismodelled in simulation. Since in Run 2 the `nSPDHits` cut in the `L0Muon` trigger selection was tighter as on the `L0Global_TIS` trigger, the correction weight is expected to differ from unity, effectively accounting for the `nSPDHits` cut efficiency.
The TISTOS trigger sample has been defined by the fraction of events from the B^+ `L0Global_TIS` sample that also pass the `L0Muon` or `L0DiMuon` trigger selection respectively. While commonly the respective trigger calibration weight is determined as function of the variable used for the cut in the trigger selection, the ratio $w = \epsilon^{\text{data}}_{TOS}/\epsilon^{\text{sim}}_{TOS}$ is studied as a function of the product of transverse momenta of the two muons for both categories. This choice simplifies a strategy where the calibration weights can be applied simultaneously to events that pass only one trigger selections and those that pass both. The correction maps for the different years of data taking are shown in Fig. 4.1 for `L0Muon` and in Fig. 4.2 for `L0DiMuon`. The resulting reweighted per-event L0 decision is computed as follows, accounting for the overlap between the two:

$$\begin{aligned}\texttt{L0Decision}_{\mu\mu} = {} & \texttt{L0MuonDecision}_{\mu\mu} \times w_{\texttt{L0Muon}} + \texttt{L0DiMuonDecision}_{\mu\mu} \times w_{\texttt{L0DiMuon}} \\ & - \texttt{L0MuonDecision}_{\mu\mu} \times w_{\texttt{L0Muon}} \times \texttt{L0DiMuonDecision}_{\mu\mu} \times w_{\texttt{L0DiMuon}}\end{aligned} \tag{4.2}$$

This approach does not take into consideration the correlation between the two L0 categories, however the residual difference compared to an individual weight for each of them is found to be in third per mille level and cancel in the ratio. Therefore, this simplified approach has been selected. While the correction maps in the low p_T region show a significant difference from unity, they are not expected to affect significantly the final results of the analyses which are either performed in the charmonia regions or have most of their signal in the high-q^2 region where muons are more energetic. In the case of normalised measurements the effect of the calibration is small when the kinematics of the two datasets is similar, as they largely cancel. However, at low q^2, where muons are softer, the effect is more significant. In addition, for the angular analysis of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$, in the targeted high-$q^2$ region the transverse momenta of the muons are expected to be large and therefore also less affected by the differences in the lower p_T region.

HLT calibration

The HLT trigger response is assumed to be well modelled in simulation, as it is known to be the case from other analyses with dimuon final states [7,107,108]. Since, especially in high-q^2 regions, where most of the signal lies, no significant trend in kinematic variables has been observed, no HLT corrections are applied in the baseline of the presented analyses. Typically, a dedicated systematic uncertainty is assigned, which will be discussed in the respective analysis chapter.

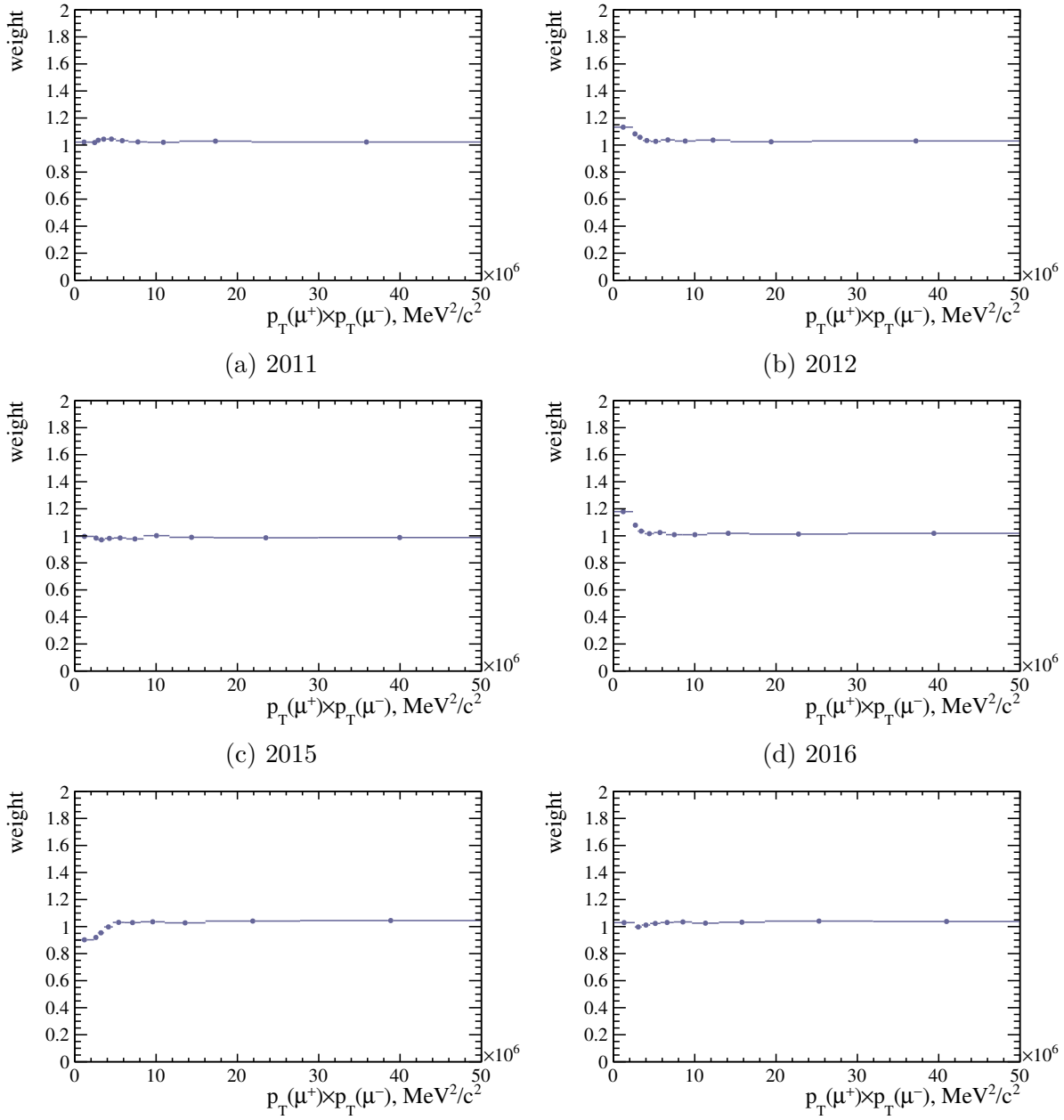


(a) 2011 (b) 2012

(c) 2015 (d) 2016

(e) 2017 (f) 2018

Figure 4.1: `L0Muon` weights as a function of product of transverse momenta of the two muons.



(a) 2011

(b) 2012

(c) 2015

(d) 2016

(e) 2017

(f) 2018

Figure 4.2: `L0DiMuon` weights as a function of product of transverse momenta of the two muons.

4.3 The Decay Tree Fitter algorithm

Particle decays, where an unstable particle decays into a certain unstable intermediate state, before this state finally decays into stable final-state particles, are referred to as decay chains or decay trees. The latter names refers to the fact, that from the initial particle at least two new ones per decay branch off. In the standard LHCb reconstruction procedure, these decay chains are reconstructed through a bottom-up approach, where the information from the stable final states is combined to reconstruct the point of origin of these stable particles and therefore reconstruct the decay vertex of the intermediate particle. A least-square fit to the properties of the decay products is then used to reconstruct the parameters of the intermediate state. It is known, that constrains such as from the primary vertex can help to improve the accuracy of a reconstruction. However, since the least-square fit is a bottom-up approach starting from the reconstructed final state particles, such information is inaccessible.
The DecayTreeFitter (DTF) algorithm [193] utilising a Kalman filter [194] while performing a least square fit was first developed for the BaBar experiment as a solution to the disadvantage of the bottom-up fitting approach. By performing an iterative vertex fit which extracts all parameters simultaneously, information from particles above the currently calculated point in the decay chain can be considered when determining the properties of a targeted intermediate state. The χ^2_{DTF} distribution is defined by

$$\chi^2_{\text{DTF}} = \sum_{\text{track } i} (q_i - h(x, p_i))^T V_i^{-1} (q_i - h(x, p_i)),$$

testing the measured track parameters q_i against a model of it, h, which relies on the vertex position x and the momentum p_i. The uncertainties are included by multiplying the expression with the covariance matrix of the tracks V. The integration of the Kalman filter enables to compare the current model prediction with the measurements and therefore recursively improve them, before the whole χ^2_{DTF} distribution is minimised. Consequently, this approach enables the simultaneous extraction of decay time, position and momentum parameters and their uncertainties and correlations for all particles in a decay chain. Since the successful deployment for the BaBar experiment, the DTF has been become a common reconstruction approach in flavour physics and is currently also used by the LHCb collaboration.
It is particularly important when reconstructing decay chains with long-lived neutral particles leaving no tracks in the detector, such as $K^0_S \to \pi^0\pi^0$ or $\Xi^0 \to \Lambda\pi^0$. For both of these cases, the decay vertex can only be determined with the use of the DTF. It should be noted that a DTF fit can also fail, in this case a very large negative value of the DTF output (-350×10^{-36}) is assigned to ensure that failed events do not interfere with any correctly returned values.

4.4 Description of the mass distribution

Measurements in particle physics are probability measurements. In particular, due to limited detector resolution, the measured invariant mass of a particle can be defined as a random variable distributed around its true value. Therefore, the invariant mass

distribution should be described by a continuous probability distribution, called Gaussian distribution, able to parameterise the detector resolution. However, the final state particles, used to reconstruct the invariant mass, can radiate photons while travelling through the detector, leading to an asymmetric tail on the left side. Furthermore, each instance of the particle crossing the detector material is followed by a corresponding energy loss. While a correction is applied for the average energy loss, the probabilistic nature of this process leads to possible under- or over-estimation of the invariant mass, creating additional effects on both tails of the resolution function. This led to the development of the Crystal-Ball function with a Gaussian core and a power-law tail describing the left tail due to photon radiation, and its double-sided version with asymmetric power-law tails on both sides:

$$f(x;\mu,\sigma,\alpha_L,n_L,\alpha_R,n_R) = \begin{cases} A_L \cdot \left(B_L - \frac{x-\mu}{\sigma}\right)^{-n}, & \text{if } \frac{x-\mu}{\sigma} < -\alpha_L \\ \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), & \text{if } -\alpha_L \leq \frac{x-\mu}{\sigma} < \alpha_R \\ A_R \cdot \left(B_R - \frac{x-\mu}{\sigma}\right)^{-n}, & \text{if } \frac{x-\mu}{\sigma} > \alpha_2 \end{cases} \tag{4.3}$$

with

$$A_{L,R} = \left(\frac{n_{L,R}}{|\alpha_{L,R}|}\right)^{n}_{L,R} \cdot \exp\left(-\frac{|\alpha_{L,R}|^2}{2}\right)$$
$$B_{L,R} = \frac{n_{L,R}}{|\alpha_{L,R}|} - |\alpha_{L,R}|.$$

In this definition σ is the standard deviation of the Gaussian core, defining the resolution of the distribution, μ is the mean of the Gaussian core, defines the extracted mass values from the invariant mass fit, $\alpha_{L,R}$ are the left and right tail parameters and finally $n_{L,R}$ are the left and right power-law parameters.
Measured events can have different uncertainties with respect to the extracted mass. This causes a distortion of the Crystal-Ball function, making a Gaussian assumption non-feasible. The per-event uncertainties are usually a continuous function and can be described approximately by a (non-continuous) sum of delta functions. This principle can be utilised by performing a linear combination of two double-sided Crystal-Ball functions to describe the invariant mass distribution.
However, approximating an infinite sum with only two delta functions has limited precision. Therefore, an alternative approach was developed and is discussed in detail in Ref. [195]. The developed function, called Hypatia, has a Crystal-Ball-like powerlaw tail and a generalised hyperbolic core using Bessel functions of the third kind to deal precisely with per-event uncertainties. In the same way as for the double-sided Crystal-Ball function, a double-sided Hypatia function was created. This function is often called Hypatia2 and has a total of 9 parameters, which are the mean and standard deviation of the core, two parameters defining the hyperbolic nature, the order of the Bessel functions and four tail parameters.
Both approaches, the double-sided Hypatia function and the sum of two double-sided Crystal-Ball function, have been used for analyses presented in this thesis. While the Hypatia function is more precise, the large number of parameters are non-trivial to control in order to have a good fit stability when fitting invariant masses of rare baryonic decay modes. As a consequence, the double-sided Hypatia approach was only used in

the analysis searching for the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ and $\Xi_b^- \to \Xi^-\psi(2S)$, while for all other mass distributions the linear combination of two double-sided Crystal-Ball function with a shared mean was used. This choice was made after dealing with a challenging control of the fit stability for the multi-dimensional simultaneous fit in case of the Ξ_b^- analysis. In the first searches, as well as the Ω_b^- mass and relative production fraction measurement, the mass fits were performed simultaneously to the signal and normalisation modes, as well as in different data taking periods if necessary.
For the angular analysis of the Λ_b^0 decays, the invariant mass distribution is part of a simultaneous fit together with three angles in order to extract the angular coefficients. The mass fit is a crucial component as it enables to determine the signal fraction in the data, needed to discriminate the signal and background components in the angular distributions.
For all invariant mass fits, the tail parameters of the signal functions have been fixed to the values obtained from fits to the truthmatched, corrected and fully selected simulation samples of each respective decay mode. This step is crucial to improve the fit stability by decreasing the number of free parameters in the fit.
The only significant observed background, unless discussed otherwise, is the combinatorial background in all presented work. It can typically be described by an exponential shape, as a collection of random tracks is more likely to produce a low-mass combination rather than a high-mass one. However, the shape of this background is known to be affected by selection cuts applied, especially in kinematic edge regions such as the high-q^2 region. Therefore, it has been carefully checked whether the remaining background in data is sufficiently well described by an exponential function.
More details for the individual invariant mass fits can be found in the respective analysis chapters.

Chapter 5

Mass and production fraction measurement of the Ω_b^- baryon

This chapter is dedicated to the first analysis out of four covered in this thesis. As discussed in the subsections about beauty baryons and their properties from Sec. 1.2.1, the knowledge about Ω_b^- baryon properties is restricted to measurements from the CDF and LHCb collaborations, all with very limited statistical power. Therefore, it is important to measure its properties with larger datasets, as presented in this chapter.
This chapter is dedicated to the measurement of the mass of the Ω_b^- baryon and its relative production fraction with respect to the Ξ_b^- baryon. It was equally motivated by the limited knowledge about the Ω_b^- baryon and the cleanness of the presented decay modes in this chapter. The measurement was performed together with Dr. Vitalii Lisovskyi and has been published in Ref. [19].

5.1 Analysis strategy

The work presented in this chapter utilises the tree-level Cabibbo-favoured decay modes $\Omega_b^- \to \Omega^- J/\psi$ and $\Xi_b^- \to \Xi^- J/\psi$. Combined with the unique topology of two long-lived, weakly decaying hyperons, that cannot be mimicked by any mesonic decay mode, both decays allow to obtain exceptionally clean Ω_b^- and Ξ_b^- invariant mass distributions. Therefore, they are ideal to perform property measurements for the Ω_b^- baryon.
Both decay modes have very similar topologies leading to only one differing final state particle. The Ξ^- hyperon is reconstructed via its decay to the $\Lambda(\to p\pi^-)$ baryon and a charged pion [1], while the Ω^- hyperon is reconstructed via its decay to the $\Lambda(\to p\pi^-)$ baryon and a charged kaon. Consequently, a lot of the systematic uncertainties cancel when extracting results for the Ω_b^- baryon w.r.t. the Ξ_b^- baryon. Another minor difference arises from the about two times larger lifetime of the Ξ^- ($(1.639 \pm 0.015) \times 10^{-10}$ s) [33] compared to that of the Ω^- hyperon ($(8.21 \pm 0.11) \times 10^{-11}$ s) [33], leading to a considerably different detector acceptance of the two decay channels.

[1] To distinguish between the two same-charge pions in the Ξ^- cascade, they are denoted as π_Ξ and π_Λ, respectively.

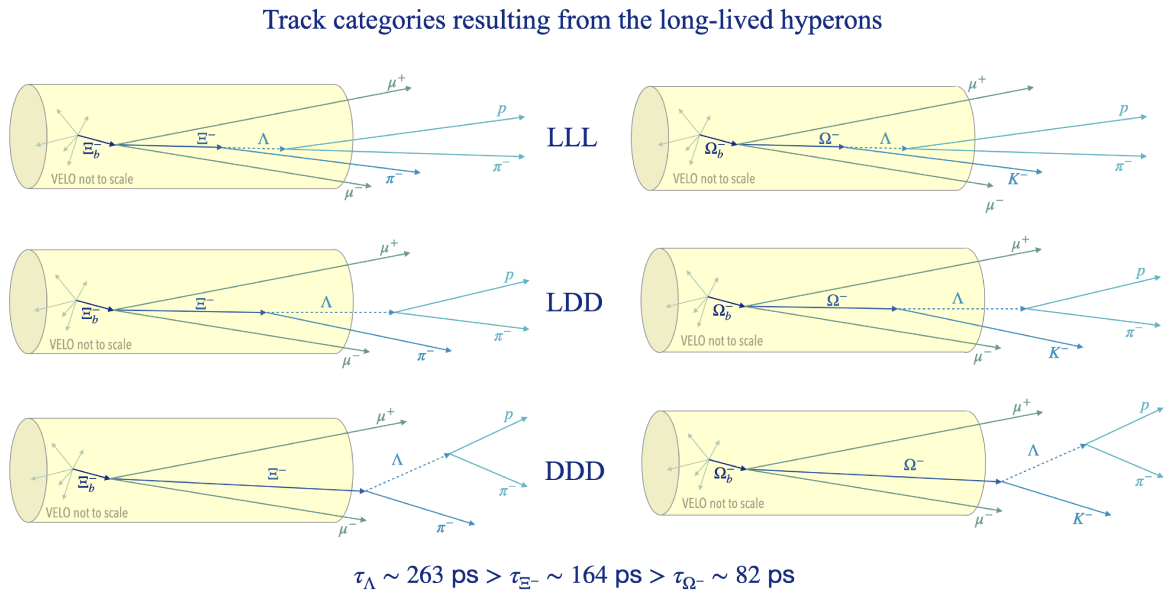


Figure 5.1: Schematic representation of the $\Xi_b^- \to \Xi^- J/\psi(\to \mu^+\mu^-)$ and $\Omega_b^- \to \Omega^- J/\psi(\to \mu^+\mu^-)$ decay chain in the three track categories LLL, LDD, DDD depending on the origin of the final state particles from the hyperon decays with respect to the VELO, indicated by the yellow cylinder.

Since the Ω^- (Ξ^-) and Λ baryons can decay within or outside of the VELO detector, three different track type categories have to be considered in this analysis. Depending on the origin of the final-state particles $K^-(\pi_\Xi), p, \pi_\Lambda$, the beauty baryon decays can be classified into LLL, LDD and DDD track type categories. They are visualised in Fig. 5.1 for the $\Xi_b^- \to \Xi^- J/\psi$ (left) and $\Omega_b^- \to \Omega^- J/\psi$ (right) decay.
The relative proportions of each track category combination have been estimated with simulation and are about 10% for LLL, 43% for LDD and 47% for DDD in case of the $\Xi_b^- \to \Xi^- J/\psi$ decay. For the $\Omega_b^- \to \Omega^- J/\psi$ decay the fractions are about 15% for LLL, 45% for LDD and 40% for DDD, highlighting the slight difference due to the smaller Ω^- lifetime. The LLL track category is characterised by somewhat harsher background conditions when compared the other two categories, due to more (unrelated) decays happening close to the PV.
Both the mass and production fraction measurements have been performed in a well-defined phase-space in the kinematic region $2 < \eta < 6$ and $p_T < 20$ GeV/c for the Ξ_b^- and Ω_b^- baryons. This is especially important for interpretation of the production-ratio measurement, as its value can depend on the kinematics, as seen previously for the Λ_b^0 baryon [83]. In addition, the momentum scale calibration has been applied to all data samples. The details about the procedure can be found in App. A.1.

5.1.1 Obtaining the Ω_b^- mass through mass differences

The measurement of the Ω_b^- mass is obtained by extracting the mass difference between the Ω_b^- and Ξ_b^- baryon from a four-dimensional simultaneous maximum likelihood fit to the $\Omega_b^- \to \Omega^- J/\psi$ and $\Xi_b^- \to \Xi^- J/\psi$ decays, and adding it to the most precise meas-

urement of the Ξ_b^- mass [20] [2]. This strategy is motivated by two technical facts. The resolution of an invariant mass highly depends on the momentum resolution of each final-state track. The momentum resolution is typically better for low-momentum tracks, as larger curvatures are easier to quantify precisely in the tracking system. This is directly related to the Q-value of a decay, which is defined by the mass difference between an initial hadron and its decay products. The larger the masses of the decay products, the smaller energy is left to be distributed among their momenta. Therefore, decay modes with smaller Q-values have better mass resolution. Consequently, the relatively high masses of the J/ψ resonance and the Ω^- and Ξ^- hyperons, lead to relatively small Q-values for both studied Ω_b^- and Ξ_b^- decay modes. Secondly, the dominant uncertainty in mass measurements is the driven by the knowledge of the absolute momentum scale within an experiment. However, if two decay modes have a very similar topology and Q-value, the related uncertainty largely cancels in the mass difference. Ideally, the final state particles of both decay modes would be identical, but within the expected statistical uncertainty of the measurement the difference from the kaon from the Ω_b^- versus the charged pion for the Ξ_b^- decay can be neglected.
The four projections of the fit are the Run 1 and Run 2 dataset respectively for both decay modes. Therefore, the full available LHCb dataset collected at 7, 8 and 13 TeV is used. The Run 1 and Run 2 datasets are fitted separately, as the invariant-mass resolution can be slightly different between the two.

5.1.2 Relative Ω_b^- production fractions at the LHC

The observed yield of b-hadron decays at the LHC depends both on the production rate of a given b hadron as well as its branching fraction to a given final state. To measure each of the two quantities individually, theoretical assumptions or measurements from independent experiments are required. In particular, for B mesons, absolute branching fractions are measured at so-called B-factories, where the collision energy is tuned in a way that all b quarks hadronise only into B mesons. The production fractions of beauty baryons are very difficult to extract since no absolute branching fraction measurements can be obtained from B factories. For the Λ_b^0 baryon, the production rate was measured relatively to the B mesons by exploiting the theoretical relation between their semileptonic decay widths within the heavy-quark limit [83]. For the Ω_b^- baryon, semileptonic decays have not been properly studied. Furthermore, the dominant semileptonic decay of the Ω_b^- would be that to the Ω_c^0 baryon and a lepton-neutrino pair. However, the Ω_c^0 baryon is also poorly studied and its absolute branching fractions have not been measured. Therefore, different approaches need to be explored for the moment. The Ω_b^- production fraction has been measured with respect to the Λ_b^0 and Ξ_b^- baryons by the CDF collaboration [78] for the first time using the respective J/ψ decay modes. However, both measurements report the ratio of production fractions times the ratio of branching fractions, since the ratio of the respective branching fractions is not known. The reported ratios can also not be disentangled by utilising $SU(3)$ symmetry like previously discussed for the partially reconstructed background from the Ξ_b decays in the Λ_b^0 angular analysis. The symmetry assumptions can only be applied if the parent beauty baryons as well as their resulting

[2] This choice was made to obtain the most precise possible result and reduce correlations due to overlapping Ξ_b^- datasets, that could be used in the world average [33]

hyperons are part of the same baryon multiplet (see Fig. 1.2). However, the Ω^- hyperon is a member of the baryon decuplet while the Ξ^- (Λ) hyperons are members of the baryon octet, consequently, no simple symmetry-based relations can be used. In addition, the predictions on the rate of $\mathcal{B}(\Omega_b^- \to \Omega^- J/\psi)$ differ in various theoretical models [196]. The reported estimates on the Ω_b^- production fraction $f_{\Omega_b^-}$, based on the CDF result, have large uncertainties, but are broadly consistent with the $f_{\Omega_b^-}/(f_{\Xi_b^-} + f_{\Xi_b^0}) = 15\%$ assumption introduced in Sec. 1.2.
However, since the production fractions depend on the production environment, it is important to measure the relative production fractions times the ratio of branching fractions for the first time at the LHC as well. In the meantime, more reliable theoretical estimates for the ratios of branching fractions might become available. Therefore, the ratio

$$r_{BR} \equiv \frac{\mathcal{B}(\Omega_b^- \to \Omega^- J/\psi)}{\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi)} \times \frac{f_{\Omega_b^-}}{f_{\Xi_b^-}} \tag{5.1}$$

has been extracted using a dataset of $6\,\text{fb}^{-1}$ collected at a center-of-mass energy of 13 TeV. It is known that the LHCb simulation inadequately models the downstream tracking efficiency in Run 1. This, together with the low statistical power for the $\Omega_b^- \to \Omega^- J/\psi$ signal events in Run 1 would lead to a large systematic uncertainty, if this effect is corrected using data. Unfortunately, the uncertainty is not expected to fully cancel due to the difference in the relative fraction of downstream tracks in the Ω_b^- and Ξ_b^- decay modes. In addition, it is not possible to extract the result combined for the Run 1 and Run 2 datasets, since the production cross-sections are energy-dependent and are therefore not equal in the two run periods. Therefore, combined with the low statistical power of the Run 1 dataset for the Ω_b^- decay mode, the result is only obtained for the Run 2 dataset. A simultaneous unbinned maximum-likelihood fit to the $\Omega_b^- \to \Omega^- J/\psi$ and $\Xi_b^- \to \Xi^- J/\psi$ datasets is performed.

5.1.3 Simulation

The simulation samples are used to determine the invariant mass shapes of the signal and to extract the efficiencies needed to measure the relative production fractions of the Ω_b^- baryon. Both the $\Omega_b^- \to \Omega^- J/\psi$ and $\Xi_b^- \to \Xi^- J/\psi$ samples are simulated assuming the phase-space model, in absence of better models. One peculiarity is, that the Ω_b^- decay has been produced through the Ξ_b^- baryon with altered mass and lifetime matching the world-average values from Ref. [197]. This is done in order to reduce the used computing resources otherwise necessary due to the small Ω_b^- production rate in the Monte Carlo generator Pythia [187]. This approach comes with the caveat of potentially incorrectly generated Ω_b^- kinematics. However, since they are generally known to mismodelled for all beauty hadrons in simulation, they will anyway be corrected. Similarly, the samples were filtered with the used stripping line to reduce consumption of computing resources, since only events that passed the stripping selection are fully processed. The subsequent Ω^-, Ξ_b^- and Λ decays have been simulated with a helicity amplitude model using the parity-valuating parameters reported in Ref. [197] to improve the kinematic properties of the decays.

5.2 Background studies and signal selection

The very distinct topologies of a cascade of two weakly decaying hyperons reduces the amount of possible significant backgrounds drastically for both the Ω_b^- and Ξ_b^- decay mode. In both cases the dominant background source is the combinatorial background, which can be suppressed by applying mass windows not only around the J/ψ resonance, but also the known hyperon masses for the Ξ^-, Ω^- and Λ. This section discusses possible background sources and how the applied selection suppresses them.

5.2.1 Trigger and stripping selection

The trigger selection in this analysis largely overlaps with the previously discussed strategy. The samples are again selected via the `L0Muon` and `L0DiMuon` trigger `TOS` decisions on the hardware stage. Since the trigger efficiency of downstream tracks in LHCb is not very high, most of the trigger power is driven by the two muons from the J/ψ decay. Therefore, the HLT1 trigger stage is driven by the `MVADecision` trigger using a loose cut on high transverse momenta of displaced tracks or pairs of tracks, together with a MVA classifier to increase the number of signal candidates and the `DiMuonHighMass` line. The latter selects events with dimuon vertices, whose mass is near or larger than the J/ψ mass. For the HLT2 trigger selection, the same strategy relying on the dimuon information is used. In addition, the two-body topology trigger was used as it was found to have non-zero contributions. For Run 2 the HLT2 trigger selection was simplified by using a trigger line specialised for the the J/ψ selection.
The dedicated data and simulation samples have been selected with the `B2XMuMu` stripping line selecting inclusive b hadrons decays with two muons in the final state and the details of the line can be found in App. C.

5.2.2 Background sources

As already mentioned, the combinatorial background is the main source of background in this analysis. Due to the tight mass window applied around the J/ψ mass, the background predominantly arises from the combinations of a true J/ψ meson with unrelated hadrons. The background is continuously distributed across the full invariant mass range probed in this measurement. While lepton misidentification backgrounds are already suppressed by the J/ψ mass window, an additional `ProbNNmu`> 0.1 cut was applied to both muon candidates [3].

Partially reconstructed

In typical measurements of b-hadron decays, an important source of background comes from partially reconstructed decays, when some final state particles were not reconstructed or bent out of the detector acceptances. As a consequence, the remaining topology can be very similar to that of the signal, and these backgrounds may even created a broadly peaking-like structure nearby to the signal invariant-mass region.

[3] Similarly to the previous discussion, there are also loose muon ID cuts already applied in the stripping line.

Typically, the main source of partially-reconstructed backgrounds are the decays where one of the final-state particle is produced in its excited state, with a de-excitation to the state of interest with an accompanying set of pions or photons. Partially reconstructed backgrounds from excited charmonium states that decay to $J/\psi X$ can safely be neglected, since the large mass difference between the J/ψ and the excited charmonia (and, therefore, the large missing energy in case of partial reconstruction) moves the reconstructed invariant mass outside of the considered range. An additional suppression comes from the pointing requirements applied in this analysis (the requirement of b-hadron momentum vector being aligned with its measured direction of flight), which is not fulfilled in case of a large missing energy.
For this reason, the decays with excited strange hadrons can be of a larger importance, as typically seen in B-meson decays [108]. The first baryonic decay mode to consider is ${\Xi_b}^{0,-} \to \Xi^*(\to \Xi^-\pi^{+,0})J/\psi$, where the excited $\Xi(1530)$ state is the lowest lying excited Ξ state. The $\Xi(1530)$ resonance always decays to $\Xi^-\pi$ and can therefore create a partially reconstructed background to the $\Xi_b^- \to \Xi^- J/\psi$ signal, if the pion has a very low momentum. The invariant-mass shape of such background would create a bump with the upper kinematic endpoint at $m(\Xi_b) - m(\pi)$. However, the wave function of the $\Xi(1530)$ state is totally symmetric in flavour, while the wave function of the parent Ξ_b^0 baryon is antisymmetric in flavour [198]. This creates a suppression of the transition $\Xi_b \to \Xi(1530)$ and, therefore, the decay rate of ${\Xi_b}^{0,-} \to \Xi^*(\to \Xi^-\pi^{+,0})J/\psi$. The same situation holds for the channel $\Xi_b^0 \to \Xi^*\gamma$, and so, by consequence, also for the J/ψ decay mode and the background is therefore negligible. Indeed, the rate of this process is consistent with being very small, as shown in an unpublished evidence for the decay in a LHCb thesis [199]. All known higher mass Ξ^* states [4] predominantly decay to ΛK rather than to $\Xi\pi$. Furthermore, with them having high masses, the missing energy is expected to be large and therefore these backgrounds are also expected to be negligible for the $\Xi_b^- \to \Xi^- J/\psi$ decay mode.
In case of the Ω_b^- baryon, decays of all excited Ω^- resonances to the $\Omega^-\pi^0$ final state violate isospin symmetry and therefore are expected to have small rates, this way reducing any $\Omega_b^- \to \Omega^-\pi^0 J/\psi$ background to a negligible level. In addition, all four currently known excited Ω^- states decay to $\Xi K(\pi)$ final states [33]. Furthermore, recent theoretical predictions of the Ω^- spectrum do not show evidence for a possible yet-unknown state below the strong-decay threshold [201], which could only decay to $\Omega^-\gamma$.
Consequently, the channels $\Xi_b^- \to \Xi^- J/\psi$ and $\Omega_b^- \to \Omega^- J/\psi$ are clean from partially-reconstructed background in vicinity of the signal region, underlining their usefulness for precise mass measurements.

Hadron misidentification background

Similarly to the partially reconstructed background, the misidentification of hadrons can move backgrounds into the signal regions, but they are expected to be negligible due to the striking topology of the hyperon decay chain.
As already discussed when dealing with the Λ_b^0 angular analysis, the K_S^0 meson may be reconstructed as a Λ candidate, if one of the pions is identified as a proton. However, the requirement that the Λ candidate originates from a decay of another long-lived particle,

[4] Some of these resonances are seen in the $\Xi_b^- \to J/\psi\Lambda K^-$ analysis [200].

virtually eliminates such background: this topology is specific to hyperons. Indeed, after the complete preselection is applied, no K^0_S peak can be found in the dataset used in this analysis. More dangerous are misidentification backgrounds from the $\Xi_b^- \to \Xi^- J/\psi$ decay as the $\Omega_b^- \to \Omega^- J/\psi$ candidate and vice versa, due to their nearly identical topology. However, the Ω^- misidentification as the Ξ^- is already suppressed by the low Ω_b^- production rate and, after being combined with the mass window around the Ξ^- and the Ξ_b^- invariant mass fit range, it is found to be negligible.
However, due to the much higher production rate of the Ξ_b^-, a clear Ξ^- misidentification background can be found in the Ω^- signal, as shown by the black histogram in Fig. 5.2.

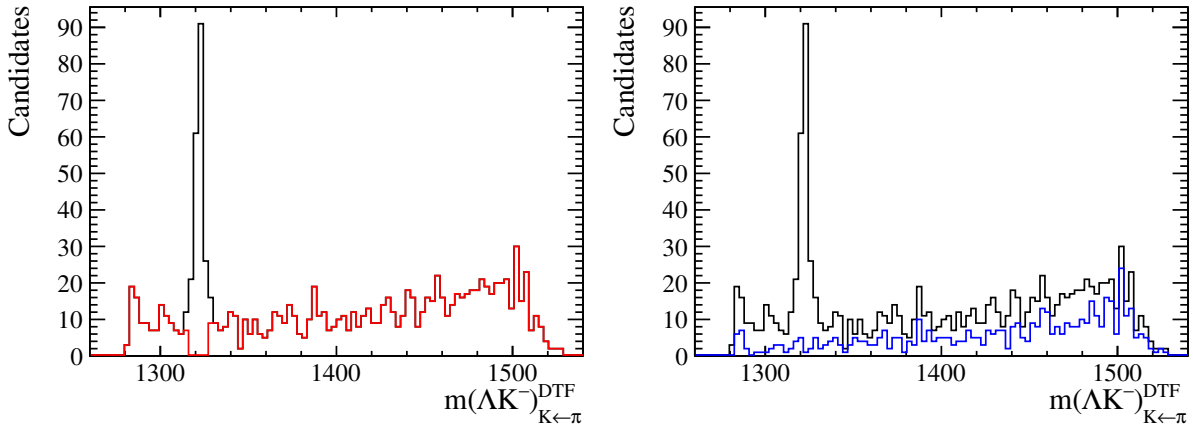


Figure 5.2: ΛK^- invariant mass with the kaon assigned a pion mass hypothesis, with the Λ mass constraints applied. Black: all selected events, showing a prominent Ξ^- peak (PDG mass 1321.7 MeV/c^2). Red: Same but with a 7 MeV/c^2 veto around the Ξ^- peak. Blue: same but requiring the kaon PIDK$>$ 5 instead of the mass veto.

The background can be removed with a PID cut on the kaon of `PIDK`$>$ 5. This cut suppresses simultaneously a part of the combinatorial background. However, the calibration of the PID cut efficiency for downstream tracks is significantly limited by the available calibration datasets and has not been thoroughly validated in previous LHCb measurements. To avoid any biases due to knowledge of the efficiencies needed for the extraction of the relative Ω_b^- production fraction, the PID cut is replaced with a mass veto for the latter.

Over-reconstructed background

Any real decay involving a Λ hyperon may be combined with a random charged pion or kaon and produce an "over-reconstructed" background in the signal region of the decay modes of interest, mimicking a Ξ^- or Ω^- candidate. The most likely source of this effect could be the abundant $\Lambda_b^0 \to \Lambda J/\psi$ decay.
Due to the heavier mass of the kaon, the background of a Λ_b^0 decay with a random K^- is larger than the Ω_b^- mass and therefore only present in the upper sideband. Within the obtained resolution, the over-reconstructed background has practically no overlap with the Ω_b^- signal peak. However, for the Ξ_b^-, due to the much lighter mass of the charged pion, the over-reconstructed Λ_b^0 background can peak right below the signal. This is because

$m(\Lambda_b^0) + m(\pi^-)$ is slightly lower than $m(\Xi_b^-)$. When applying only a loose selection [5], the two-dimensional plot of the $J/\psi\Xi^-$ invariant mass with the $J/\psi\Lambda$ invariant mass shows a clear Λ_b^0 contribution, as can be seen in Fig. 5.3.

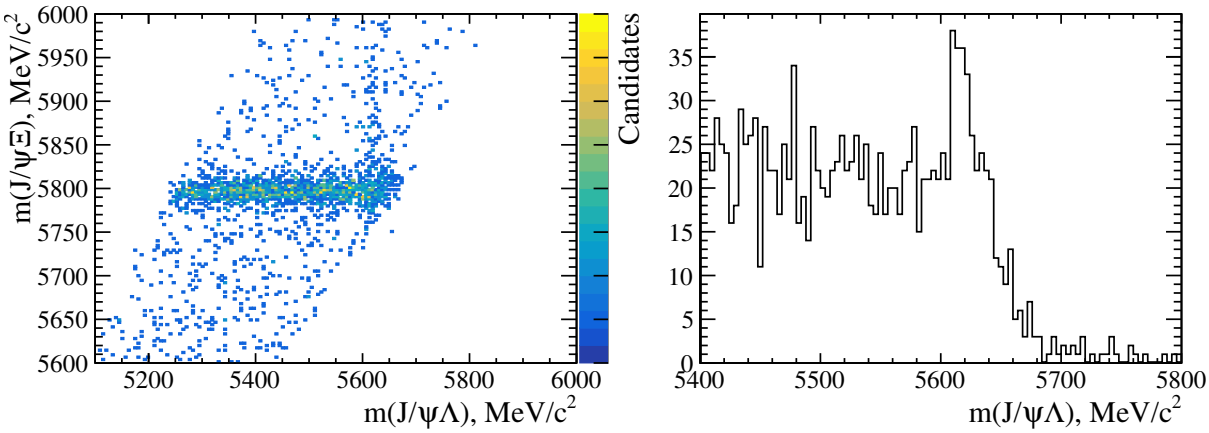


Figure 5.3: Left: Horizontal band is the $\Xi_b^- \to \Xi^- J/\psi$ signal, vertical band corresponds to combinations of $\Lambda_b^0 \to \Lambda J/\psi$ with a random pion. Right: $\Lambda J/\psi$ invariant mass with a visible Λ_b^0 peak. The plots are prepared using the Run 2 dataset.

In order to suppress the Λ_b^0 background, a displacement cut is applied to the Ξ^- pion, requiring it to be significantly displaced from the interaction point. This choice is made, since explicit mass vetoes may shape the combinatorial background and introduce a bias on the final mass measurement. Fig. 5.4 shows events passing the cut, and the size of the Λ_b^0 peak is significantly reduced. The same cut is applied on the kaon in the Ω_b^- dataset to ensure that the selections align as much as possible between the two datasets.

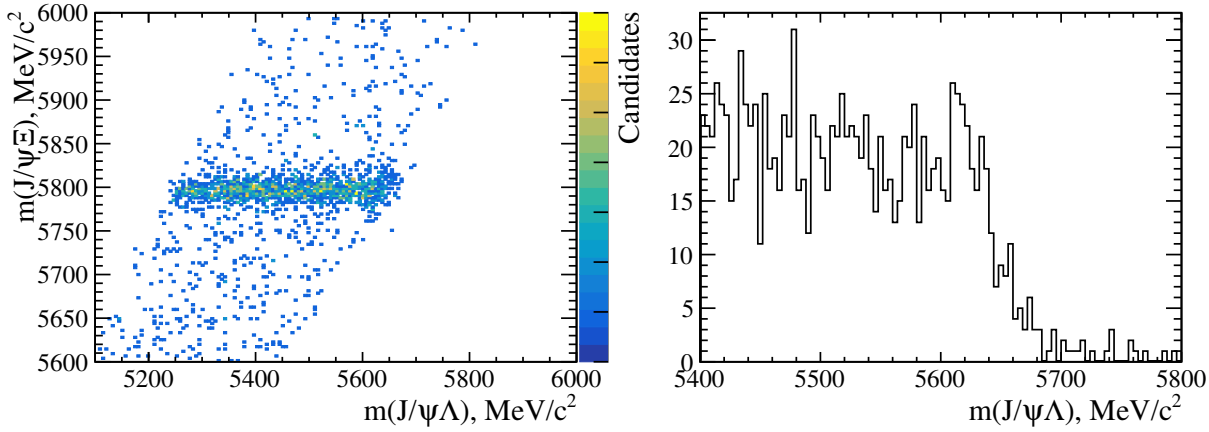


Figure 5.4: Same as Fig. 5.3, but only events with $\chi^2_{\text{IP}}(\pi_\Xi) > 20$ are selected. It can be seen that the Λ_b^0 background becomes smaller.

[5] This includes the trigger selection, the hyperon and J/ψ mass windows as well as the muon PID and some ordinary acceptance cuts.

Hadron combination backgrounds

The last bigger group of background considered, is connected to combination of the correct final-state hadrons from decay modes other than the signal ones. While the direct decay $\Xi_b^- \to J/\psi \Lambda \pi^-$, where Λ and π^- do not originate from the Ξ^- hyperon, has not been observed to date, the same final state can arise from the Cabibbo-suppressed transitions $\Xi_b^- \to J/\psi \Sigma(1385)^- (\to \Lambda \pi^-)$. Additionally, the tail of $\Sigma(1385)^-$ resonance could reach into the Ξ^- mass window, but the resonance decays promptly and is therefore suppressed by the displacement requirements on the π_Ξ, and cannot be present in the sample where the pion forms a downstream track. The background is expected to be negligible due to the displacement requirement and the Cabibbo suppression. Since a similar displacement cut is applied for the Ω^-, the equivalent Cabibbo-suppressed process $\Omega_b^- \to J/\psi \Xi^-(1690) (\to \Lambda K^-)$ is also negligible.
The never observed Cabibbo-suppressed decay $\Omega_b^- \to \Xi^- J/\psi$ is removed in the $\Xi_b^- \to \Xi^- J/\psi$ channel by the choice of the fit range that excludes the large Ω_b^- mass, as well as due to its small production rate. Similarly, the $\Xi_b^- \to J/\psi \Lambda K^-$ decays are removed by the choice of the fit range used for the $\Omega_b^- \to \Omega^- J/\psi$ channel.
The partially reconstructed decay $\Omega_b^- \to \Omega^- J/\psi$ with $\Omega^- \to \Xi^- \pi^0$, where the pion is not reconstructed, is negligible. It is not only suppressed by the Ω_b^- production rate and the small branching fraction of the $\Omega^- \to \Xi^- \pi^0$ decay, but also the topology requirements: the Ξ^- hyperons from Ω^- decays would be significantly displaced from the dimuon vertex, due to the long lifetime of Ω^-.

The charged pions and clone track background

As discussed in the detector chapter, the LHCb reconstruction software runs a dedicated algorithm to remove tracks that share at least 70% of their hits ("clone tracks"). However, since the $\Xi_b^- \to \Xi^- J/\psi$ decay mode has two same-charge pions, it has a higher risk for the presence of clone tracks even after the algorithm was run.
In the LLL track category, where both pions are long tracks, some clone tracks were found. The other two track categories as well as the Ω_b^- samples did not show any presence of clone tracks. The angle between two clones is very close to zero, because they typically share several hits in the VELO detector, therefore they can be removed with the cut $\theta(\pi_\Xi, \pi_\Lambda) > 0.5$ mrad in the laboratory frame, while preserving the entire signal. The same cut, but between pion and kaon, was also applied to the Ω_b^- sample for consistency in the selections.

5.2.3 Final selection

The final selection can be found in Tab. 5.1. For both muon candidates the boolean selection that there are within the acceptance of the muon chambers (`InAccMuon`) and have information from the rich detectors (`hadRich`) is applied. This selection is applied to enable good quality muon candidates. The only additional selection is the removal of multiple candidates. In data, only one random candidate is kept if multiple are present in the same event, this affected only 0.8% of the data events after the full selection was applied. In simulation, the candidate with the lower background category is kept. In case several belong to the same category, one is randomly kept. It should be noted that the

fraction of multiple candidates is very low and their removal does not affect significantly any of the results in this measurement.

Table 5.1: Offline selection requirements for the mass and relative production fraction measurement of the Ω_b^- baryon. The exchangeable respective Ξ_b^- background vetoes are marked in bold for the mass and relative production fraction measurement.

Type	Particle(s)	Requirement
Quality and acceptance	Ξ_b^-, Ω_b^-	$2 < \eta < 6$
	Ξ_b^-, Ω_b^-	$p_T < 20$ GeV/c
	μ	`InAccMuon, hasRich`
PID	μ	$3 < p < 150\,\mathrm{GeV}/c$
	μ	`ProbNNmu>0.1`
	K	`PIDK>5` / **for mass measurement**
Mass windows	dilepton	$3000 < \mathrm{m}(\mu^+\mu^-) < 3150$ MeV/c^2
	Λ	$\lvert m(p\pi^-) - 1115.68\rvert < 8$ MeV/c^2
	Ξ^-	$\lvert m(\Lambda\pi^-) - 1321.71 - m(\Lambda) + 1115.68\rvert < 10$ MeV/c^2
	Ω^-	$\lvert m(\Lambda K^-) - 1672.45 - m(\Lambda) + 1115.68\rvert < 10$ MeV/c^2
Background rejection		$\chi^2_{\mathrm{IP}}(\pi_\Xi) > 20$
		$\chi^2_{\mathrm{IP}}(K^-) > 20$
	clone tracks	$\theta(\pi_\Xi, \pi_\Lambda) > 0.5$ mrad
	clone tracks	$\theta(K, \pi_\Lambda) > 0.5$ mrad
	Ξ^-, Ω^-	$p_T > 1$ GeV/c
	Ξ^- veto	$\lvert m(\Lambda K^-_{\leftarrow\pi^-})_{DTF} - 1322\rvert > 8$ MeV/c^2 / **for PF meas.**

5.3 Measurement of the Ω_b^- and Ξ_b^- mass difference

As discussed in the analysis strategy, the mass difference between the Ω_b^- and Ξ_b^- baryons, rather than the actual Ω_b^- mass, is extracted from the simultaneous invariant-mass fit to the $\Omega_b^- \to \Omega^- J/\psi$ and $\Xi_b^- \to \Xi^- J/\psi$ decay modes. The fitted variables are the invariant masses of the $\Xi^- J/\psi$ and $\Omega^- J/\psi$ systems using the PV constraint as well as the mass constraints on the J/ψ, Ξ^- (Ω^-) and Λ hadrons.
As discussed in the Sec. 4.4, the signal peaks of the Ω_b^- and Ξ_b^- decays are each described by a linear combination of two double-sided Crystal Ball functions with their fraction being fixed to 0.5 in order to improve the fit stability. In addition, the fraction of the widths between the two Crystal Ball functions is shared between the Ξ_b^- and Ω_b^- and between run periods to further improve the fit stability. The maximum-likelihood fit is simultaneously performed to the two decay modes and two data taking periods, Run 1 and Run 2. The mass difference between the two beauty baryons is directly extract from the fit by expressing the Ω_b^- mass values obtained from the fit in terms of the mass difference. For what concerns the fit procedure, the fit is first performed with the respective simulation samples, where all parameters are kept floating. The fits to simulation can be found in Fig D.1 in App. D and the obtained mass difference is $m(\Omega_b^-) - m(\Xi_b^-) = 251.16 \pm 0.02\,\mathrm{MeV}/c^2$. When compared to the actual simulated values of the Ω_b^- and Ξ_b^- baryon

masses, the fit returns a bias of $+0.06 \pm 0.02$ MeV/c^2 on the $m(\Omega_b^-) - m(\Xi_b^-)$ measurement. Since the expected statistical uncertainty is expected to be an order of magnitude larger than the bias, the result extracted from data will be corrected for it and its full value is assigned as a systematic uncertainty. The uncertainty from this bias arises from the size of the simulation sample. The origin of this bias might stem from the choice of the signal distributions, from the final-state radiation, other reconstruction effects or is simply a statistical effect. However, at the present level of precision, no need for a detailed investigation was perceived.
After the fit to simulation, the tail parameters and the fractions of the widths of the signal functions are fixed to the obtained values. The background is modelled by an exponential and its slope is kept floating in the final fit. The resulting plots from the simultaneous fit to data can be seen in Fig. 5.5.

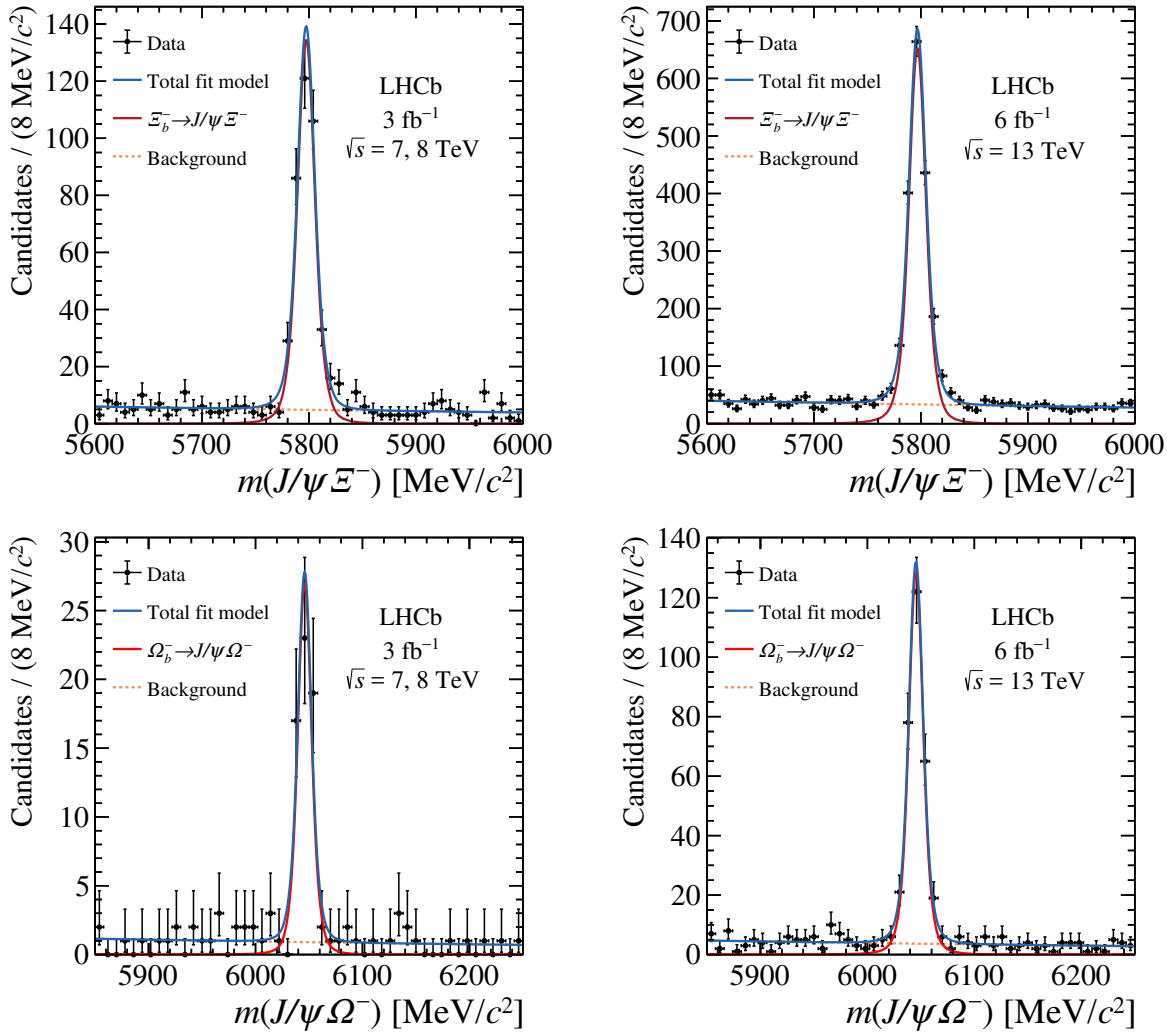


Figure 5.5: Invariant-mass distributions in the (top row) $\Xi_b^- \to \Xi^- J/\psi$ and (bottom row) $\Omega_b^- \to \Omega^- J/\psi$ datasets, in the (left column) Run 1 and (right column) Run 2 data-taking periods. The results of the simultaneous fit used for the mass-difference measurement are overlaid.

The values of the extracted fit parameters are given in Tab. 5.2. It can be seen that the

resolution of $\Omega_b^- \to \Omega^- J/\psi$ decay is slightly better compared to $\Xi_b^- \to \Xi^- J/\psi$ decay. This is expected due to the larger Q-value in the Ξ_b^- decay. The obtained value for the mass difference

$$m(\Omega_b^-) - m(\Xi_b^-) = 248.54 \pm 0.51\,(\text{stat})\,\text{MeV}/c^2$$

shows only the statistical uncertainty after correcting for the fit bias (discussed above) and will be referred to as nominal result in the following.

Table 5.2: Fitted values of all free parameters obtained from the fit to Run 1 and Run 2 collision dataset.

Parameter	Value
$m(\Omega_b^-) - m(\Xi_b^-)$	248.60 ± 0.51 MeV/c^2
$m(\Xi_b^-)$ Run1	5797.37 ± 0.50 MeV/c^2
$m(\Xi_b^-)$ Run2	5796.68 ± 0.25 MeV/c^2
width Ω_b^- Run1	5.40 ± 0.73 MeV/c^2
width Ω_b^- Run2	5.66 ± 0.35 MeV/c^2
width Ξ_b^- Run1	6.83 ± 0.40 MeV/c^2
width Ξ_b^- Run2	6.65 ± 0.19 MeV/c^2
signal yield Ω_b^- Run1	60 ± 8
signal yield Ω_b^- Run2	300 ± 18
signal yield Ξ_b^- Run1	377 ± 21
signal yield Ξ_b^- Run2	1790 ± 47
bkg yield Ω_b^- Run1	45 ± 7
bkg yield Ω_b^- Run2	184 ± 15
bkg yield Ξ_b^- Run1	244 ± 17
bkg yield Ξ_b^- Run2	1677 ± 46
bkg slope Ω_b^- Run1	$(-1.2 \pm 1.3) \times 10^{-3}$
bkg slope Ω_b^- Run2	$(-1.3 \pm 0.6) \times 10^{-3}$
bkg slope Ξ_b^- Run1	$(-8.8 \pm 5.4) \times 10^{-4}$
bkg slope Ξ_b^- Run2	$(-8.8 \pm 2.1) \times 10^{-4}$

To ensure that the result is not biased due to the applied hyperon mass constraints, as well as merging of significant portions of the dataset, several cross-checks were performed. First, the fit is performed with only the J/ψ constraint applied. The returned mass difference $m(\Omega_b^-) - m(\Xi_b^-) = 248.66 \pm 0.56$ MeV/c^2 is in excellent agreement with the nominal result. Afterwards, the fit is performed with different data-taking periods (Run 1, Run 2, as well as 2015+16, 2017, 2018 separately), both magnet polarities, particles and antiparticles and the different track-type categories (LLL, LDD, DDD). An overall good stability of the mass difference is found.

5.3.1 Systematic uncertainty

In order to extract the final result for the mass difference and determine the value of the Ω_b^- mass, the non-cancelling sources of the systematic uncertainties need to be determined. The full set of systematic uncertainties is summarised in Tab. 5.3.
As already mentioned, the absolute value of the observed fit bias in the signal-only study is assigned as a systematic uncertainty additionally to the correction of the fit value for it. Additionally, the fit can be biased when using the full fit model, *i.e.* including the

Table 5.3: Systematic uncertainties on the mass-difference measurement. The total systematic uncertainty is obtained by summing all sources in quadrature.

Source	Uncertainty [MeV/c^2]
Fit bias	0.06
Full fit model	0.01
$\Lambda_b^0 \to \Lambda J/\psi$ background	0.10
Momentum scale	0.09
$\mathrm{d}E/\mathrm{d}x$ correction	0.01
Hyperon mass	0.35
Total	0.38

background model. This bias was studied with pseudoexperiments with realistic signal and background yields and no significant effect was found. Therefore only the signal fit bias is treated as a source of systematic uncertainty.
Several checks have been performed to estimate the systematic uncertainty related to the choice of the fit model. First, pseudoexperiments have been performed, where datasets generated with a linear background model are fitted with the nominal exponential: no significant deviation from the zero has been seen for the pull on the mass difference extracted from these experiments. Similarly, the signal fit model has been replaced by a single Crystal-Ball distribution and the obtained difference on the mass difference is $0.01\,\mathrm{MeV}/c^2$. Additionally, the resolution effects of the fits were tested by fixing the ratio of the Ξ_b^- and Ω_b^- resolution to that from simulation. No significant biases were observed, therefore the combined systematic uncertainty of $0.01\,\mathrm{MeV}/c^2$ is assigned for the fit model choice.
As discussed in the background section, a residual contribution of an over-reconstructed $\Lambda_b^0 \to \Lambda J/\psi$ background might remain in the Ξ_b^- samples after the full selection is applied. In order to access the related systematic uncertainty, four alternative approaches and their effect on the mass difference have been studied. The first three approaches aim at suppressing the background to a negligible level by including further cuts. In particular, these are a dedicated Λ_b^0 mass veto, requiring a large flight-distance of the Ξ^- candidate utilising its large lifetime, or rejecting good-quality $J/\psi\Lambda$ vertices that indicate the J/ψ and Λ candidate originate from the same point in space. The last approach includes a dedicated fit component for the over-reconstructed Λ_b^0 background. Its shape is accessed with a kernel density estimator [202] using simulation that were passed through the same selection as the Ξ_b^- samples and the yield is kept floating in the fit. The final uncertainty is obtained by averaging over all four shifts from the nominal result of the mass difference and is found to be $0.10\,\mathrm{MeV}/c^2$.
The momentum scale calibration has a relative uncertainty of $\pm 0.03\%$ [203]. Therefore, the momentum scale in simulation has been varied according to this value and a systematic uncertainty of $0.09\,\mathrm{MeV}/c^2$ on the mass difference was obtained.
In the case of equal number of final-state tracks between two decay modes used for measurements of mass difference, the systematic uncertainty on the amount of material assumed in the track reconstruction for the energy loss ($\mathrm{d}E/\mathrm{d}x$) has been found to cancel up to $0.01\,\mathrm{MeV}/c^2$ [67].
The last uncertainty is related to the limited knowledge of the Ξ^- and Ω^- hyperon masses used to constrain the reconstructed masses. In order to estimate the resulting bias on

the hyperon-mass-constrained b-baryon mass, the central values of the mass constraints are varied in the simulated samples according to the uncertainties on the known Ω^- and Ξ^- hyperon masses, respectively. The average bias on the measured b-baryon mass is evaluated for both cases and the two uncertainties are then summed in quadrature. No such procedure is performed for the Λ hyperon and the J/ψ mass, since their uncertainties are significantly smaller and the resulting effects are expected to cancel in the mass difference. The obtained systematic uncertainty of $0.35\,\text{MeV}/c^2$ is dominated largely by the uncertainty of the Ω^- mass with $m_{\text{PDG}}(\Omega^-) = 1672.45 \pm 0.29\,\text{MeV}/c^2$ [33].
Finally, the mass difference is determined to be

$$m(\Omega_b^-) - m(\Xi_b^-) = 248.54 \pm 0.51\,(\text{stat}) \pm 0.38\,(\text{syst})\,\text{MeV}/c^2, \tag{5.2}$$

taking into account the fit bias observed in simulation and accounting for all sources of the systematic uncertainties. The result is in agreement with the previous measurement, $m(\Omega_b^-) - m(\Xi_b^-) = 247.4 \pm 3.2\,(\text{stat}) \pm 0.5\,(\text{syst})\,\text{MeV}/c^2$, by LHCb [80] using the Run 1 dataset and fully hadronic decay modes, but significantly more precise.
Finally, the Ω_b^- mass value, $m(\Omega_b^-)$, is obtained using the statistically independent [6] and most precise measurement of the Ξ_b^- baryon mass to date: $m(\Xi_b^-) = 5797.33 \pm 0.24 \pm 0.29\,\text{MeV}/c^2$ [20]. The systematic uncertainties related to the momentum-scale calibration and the energy-loss correction are assumed to be fully correlated, while all remaining ones are expected to be uncorrelated. Therefore, the final mass value for the Ω_b^- mass is

$$m(\Omega_b^-) = 6045.9 \pm 0.5\,(\text{stat}) \pm 0.6\,(\text{syst})\,\text{MeV}/c^2,$$

rounded to uncertainties of 0.1 MeV/c^2.
This result is the most precise determination of the Ω_b^- mass and is consistent with the current world average of $6045.2 \pm 1.2\,\text{MeV}/c^2$ [33].
It should be noted, that this result supersedes the value reported in Ref. [67] as the latter uses the smaller subset of data and the same final states as presented in this measurement.

LHCb Ω_b^- mass average

Finally, the LHCb average of the Ω_b^- mass measurements is also evaluated, to correctly account for correlations due to overlaps in the data samples and selected decay modes, as well as correlated systematic uncertainties.
In order to proceed with determining the LHCb average of the Ω_b^- mass, the average of the two LHCb determinations of the mass difference $m(\Omega_b^-) - m(\Xi_b^-)$ is performed first. Besides the value discussed in this chapter $248.54 \pm 0.51\,(\text{stat}) \pm 0.38\,(\text{syst})\,\text{MeV}/c^2$, the mass difference has been determined in Ref. [80] to be $247.4 \pm 3.2\,(\text{stat}) \pm 0.5\,(\text{syst})\,\text{MeV}/c^2$ using $\Omega_b^- \to \Omega_c^0(\to pK^-K^-\pi^+)\pi^-$ and $\Xi_b^- \to \Xi_c^0(\to pK^-K^-\pi^+)\pi^-$ decays. Consequently, the two measurements are statistically uncorrelated and only correlated systematic uncertainties from the momentum scale calibration and the energy loss due to material interactions need to be treated accordingly. While the latter systematic uncertainty is not quoted in Ref. [80], it is assumed to be zero. All other systematic uncertainties are assumed to be uncorrelated. A weighted average based on the statistical precision of the measurements is performed and the resulting value is found to be $248.50 \pm 0.51\,(\text{stat}) \pm 0.37\,(\text{syst})\,\text{MeV}/c^2$

[6] Since this measurement uses fully hadronic $\Xi_b^- \to \Xi_c^0\pi^-$ decays.

accounting for all correlated and uncorrelated uncertainties.
This result can be used in combination with the statistically independent Ξ_b^- mass measurement from Ref. [20] to determine the Ω_b^- mass value from measurements through the mass difference $m(\Omega_b^-) - m(\Xi_b^-)$. Accounting for correlated systematic uncertainties in the same manner as for the mass difference, the Ω_b^- mass is found to be

$$m(\Omega_b^-) = 6045.9 \pm 0.5\,(\text{stat}) \pm 0.6\,(\text{syst})\,\text{MeV}/c^2. \tag{5.3}$$

Finally, three different LHCb measurement of the Ω_b^- mass from different final states enter the PDG average of it [33]. These are the partial Run 1 dataset analysis [67] using the $\Omega_b^- \to \Omega^- J/\psi$ decay, the full Run 1 analysis [80] with $\Omega_b^- \to \Omega_c^0 \pi^-$ decays and the full Run 1 + Run 2 analysis of $\Omega_b^- \to \Xi_c^+ K^- \pi^-$ decays [81]. While the first one is superseded by the result presented in this chapter and published in Ref. [19], the other two are used to determine an LHCb average of the Ω_b^- mass. When correctly accounting for correlations or systematic uncertainties, the LHCb average of the Ω_b^- baryon mass is found to be

$$m(\Omega_b^-)_{\text{LHCb}} = 6045.7 \pm 0.5\,(\text{stat}) \pm 0.6\,(\text{syst})\,\text{MeV}/c^2. \tag{5.4}$$

This is the most precise determination of Ω_b^- mass to date by a single experiment, and has a factor two improvement in precision with respect to the world average as of 2022 [33].

Summary of the mass measurement

To summarise the mass measurements discussed in this chapter and published in Ref. [19], reports the most precise determination of the Ω_b^- baryon mass, being consistent with the current world average of $6045.2 \pm 1.2\,\text{MeV}/c^2$ [33], but reducing the uncertainty by a factor two. It should be noted that this measurement supersedes the value measured in Ref. [67] which uses a subset of the data and the same final state as the presented analysis. The summary of the Ω_b^- mass measurements from the CDF and LHCb experiments is shown in Fig. 5.6.

5.4 Measurement of the relative Ω_b^- production x BF ratio

The ratio of the relative production rates times the ratio of the branching fractions of the Ω_b^- and Ξ_b^- decays in Eq. (5.1)

$$r_{BR} = \frac{N(\Omega_b^- \to \Omega^- J/\psi)}{N(\Xi_b^- \to \Xi^- J/\psi)} \times \frac{\varepsilon(\Xi_b^- \to \Xi^- J/\psi)}{\varepsilon(\Omega_b^- \to \Omega^- J/\psi)} \times \frac{\mathcal{B}(\Xi^- \to \Lambda\pi^-)}{\mathcal{B}(\Omega^- \to \Lambda K^-)} \tag{5.5}$$

can be accessed by determining the respective efficiencies ϵ of the decay modes, their signal yields N from invariant mass fits and correcting by the branching fractions of the subsequent Ξ^- and Ω^- hyperon decays. Consequently, it is crucial to calibrate the simulation to avoid any mismodelling in order to obtain correct efficiencies.

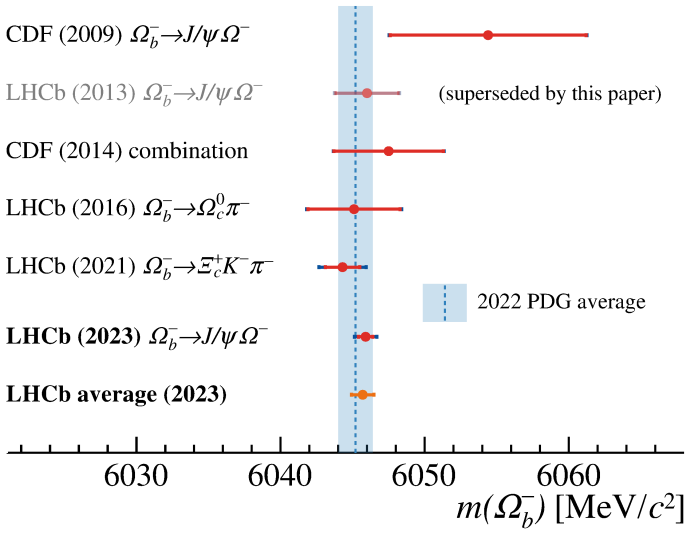


Figure 5.6: Summary of the existing measurements of the Ω_b^- mass. The plot excludes the measurement by the D0 experiment [21] as it is inconsistent with all the others and lies outside the range of this plot.

5.4.1 Correction to the simulations

For certain calibration steps, centralised LHCb calibration samples can be used. In order to calibrate the mismodelling of certain quantities (such as production kinematics) that cannot be performed with these calibration samples, background-subtracted Run 2 datasets have been created for both the $\Omega_b^- \to \Omega^- J/\psi$ and $\Xi_b^- \to \Xi^- J/\psi$ decay modes via the ${}_s\mathcal{P}lot$ technique. The details can be found in the App. D.

Particle identification and event multiplicity

While the `B2XMuMu` stripping line contains a loose PID selection requirement on the muon candidates, its efficiency is not corrected in the nominal calibration and only considered as source of a systematic uncertainty. This decision was made, since the simulation samples have been filtered with the stripping line during the production process, while the size of this mismodelling is known to be small. As a consequence, the stripping cut cannot be removed from the samples in order to calibrate them. However, the offline applied `ProbNNmu` PID cut is stricter and the variable resampled with the `PIDGen` toolkit discussed in Sec. 4.2.2. The mismodelling of the event multiplicity has been corrected by scaling its distribution by a factor 1.3 during the calibration process. This factor has been determined by comparing the mismodelling of the `nTracks` distribution, which can be found in Fig. D.3 in the App. D. The same calibration procedure was performed for both muon candidates in the Ω_b^- and Ξ_b^- simulation samples.

Trigger response

The L0 trigger response is calibrated in both Ξ_b^- and Ω_b^- datasets according to the discussion in Sec. 4.2.3. While the HLT response is expected to be modelled correctly,

a dedicated cross-check is performed with the $\Xi_b^- \to \Xi^- J/\psi$ data sample confirming this assumption.

Properties of baryons and their decays

After applying the PID calibration, the kinematics of the beauty baryons and properties of the decays in the decay chain need to be calibrated.
Before determining any weights related to the kinematics, the correct modelling of the fraction of long and downstream tracks was checked in the simulation. This was done, since the downstream tracking efficiency depends on the track momentum and therefore can be affect by miscalibrations of the fractions. It was found that the modelling of this effect in simulation agrees within the uncertainties with the background-subtracted datasets for both the Ξ_b^- and Ω_b^- decay modes.
No lifetime calibrations are needed for the Ω_b^- and Ξ_b^- baryons, as the values used in the generator of the simulation are equivalent with the latest world-averages [33].
However, the generated beauty hadron kinematics are known to be mismodelled and expected to possibly be additionally affected by the generation through the modified Ξ_b^- baryon in case of the Ω_b^- sample, as discussed above. In particular, the transverse momentum p_T dependence of the b-baryon production in the LHCb experiment is known to be mismodelled. Therefore, a dedicated calibration of both the Ξ_b^- and Ω_b^- transverse momentum distributions is performed by calculating a ratio between the normalised reconstruction-level distribution in the sWeighted data and simulation, with identical selection between the two. Six isopopulated bins in baryon p_T are used to compute the correction map, to provide a sufficient granularity while using a reasonable amount of data per bin. The comparison between the background-subtracted data and the simulation, as well as the obtained weights can be found in Fig. D.4 and Fig. D.5 in the App. D. When applying the p_T weights to the simulation, they are interpolated linearly to the corresponding value of transverse momentum from the weight map. Kinematic weights are attached to both reconstructed- and generator-level simulation samples.
No other kinematic variables are calibrated due to the limited statistics after the background subtraction on data.
In phase-space simulation the helicity decays are all generated flat. Since for both the Ξ_b^- and Ω_b^- decays no reliable form-factor predictions are available, no dedicated model is available for the generation of the simulation. In an attempt to improve the modelling of the Ξ^-, Ω^- and Λ hyperon decays, they have been produced with a helicity amplitude model which includes the latest parity-violating parameter values [197]. However, it was found that a disagreement for the helicity angles remains, which can be visualised and corrected through the momentum asymmetries

$$\frac{p(A) - p(B)}{p(A) + p(B)} \tag{5.6}$$

between the baryon decay products A and B. The reason for the mismodelling is likely to be related to the helicity structure of the top-level beauty baryon decay not being properly modelled by the phase-space simulation. These correction maps are calculated and applied in the same manner as the transverse momentum correction weights. The weights have been iteratively calculated following the order of the decays from the collision to the

final state particles. The disagreements between data and simulation for the momentum asymmetries of the Ξ^- decay products (left) and Λ ones (right) are shown in Fig. 5.7.

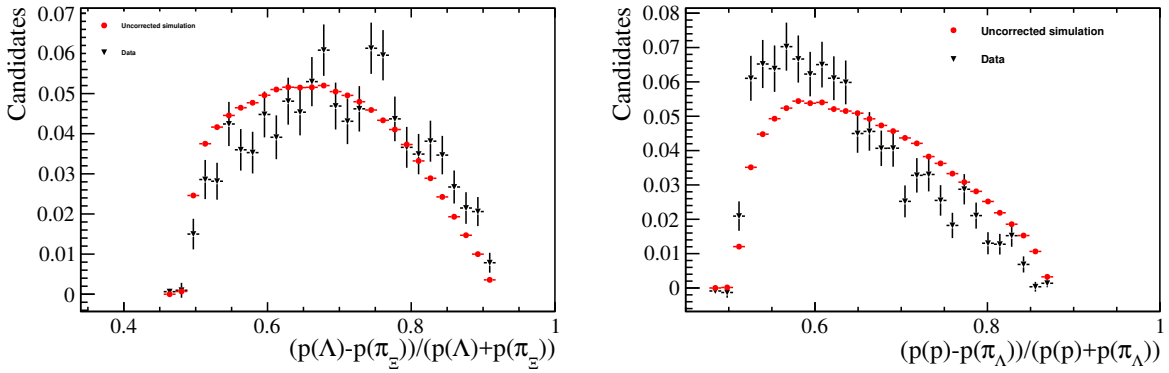


Figure 5.7: Momentum asymmetry distributions in hyperon decays from the Ξ_b^- decay chain (left: Ξ^-, right: Λ) for Run 2 data and simulation, normalised to unit area.

Similarly, the momentum asymmetry of the Ω^- and Λ decay for the Ω_b^- decay chain have been studied. Due to the statistical limitation, fewer bins have been chosen. Within the available statistics, asymmetry in the Ω^- decay chain is in agreement between the data and simulation, as can be seen on the left plot in Fig. 5.8. Therefore, it was decided to not apply a correction on this distribution. However, there is a slight disagreement in the Λ decay chain (right plot).

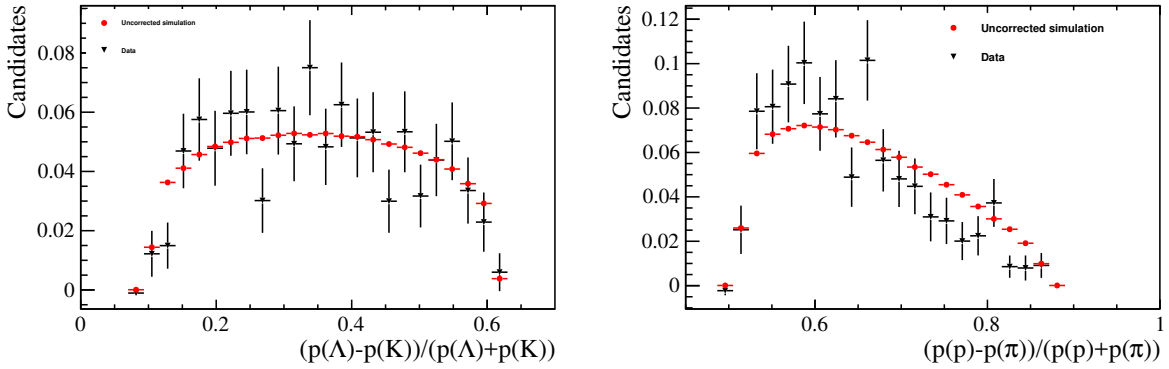


Figure 5.8: Momentum asymmetry distributions in hyperon decays from the Ω_b^- decay chain (left: Ω^-, right: Λ) for Run 2 data and simulation, normalised to unit area.

The correction weight maps for the Ξ_b^-, Ω^- and respective Λ momentum asymmetries can be found in Fig. D.6 and Fig. D.7 in the App. D.

Remaining calibration

After applying the aforementioned corrections, a residual disagreement between data and simulation is found for quality of the decay-tree-fit, χ^2_{DTF} per degree of freedom. Although this variable is not used in the selection chain of the current analysis, it can be a sign of

mismodelling of other quantities related to the vertex quality. Therefore, it is only applied to the Ξ_b^- sample and dedicated systematic uncertainty is assigned, as the cancellation between the Ξ_b^- and Ω_b^- channels cannot be guaranteed. The corresponding comparison plots and weights can be found in Fig. D.8 (Ξ_b^-) and Fig. D.9 (Ω_b^-) in the App. D.
The resulting comparisons between the data and corrected simulation show a good agreement is seen on the kinematic variables of the Ξ_b^- and Ω_b^-. The notable exceptions are the track ghost probability variables, their effect is however expected to cancel in the ratio of the two decay modes (a dedicated systematic uncertainty is considered in Sec. 5.4.3).
Lastly, several kinematic and vertexing variables are studied, splitting the dataset in equally populated halves, as well as the three track categories, to ensure that the proportion of long and downstream tracks between data and simulation agrees in different parts of the phase space. No significant deviations have been found.

5.4.2 Efficiency calculation

Estimation of the signal selection efficiency with simulated datasets is a commonly used tool to unfold true physics distribution from the measured distributions in an experiment. The signal efficiency describes the proportion of signal candidates that are produced within the acceptance of the LHCb detector, and have passed the full reconstruction, trigger and selection requirements. In LHCb, the total selection efficiency can be factorised as the product of the following efficiencies:

$$\epsilon_{tot} = \epsilon_{gen} \times \epsilon_{filt} \times \epsilon_{sel}, \tag{5.7}$$

where the first term represents the generator-level efficiency, the second one is the filtering efficiency, and the last one – the trigger and offline selection efficiency. The following sections describe each of the efficiencies independently and reports the obtained values. Each mentioned efficiency is computed iteratively on top of the previous ones.
To summarise, the efficiencies for both the $\Xi_b^- \to \Xi^- J/\psi$ and $\Omega_b^- \to \Omega^- J/\psi$ decay modes have to be estimated in order to measure the relative branching fractions times the production rates from the observed signal yields in the fits to the invariant mass distributions.

Generator-level efficiencies

The generator-level efficiency ϵ_{gen} accounts for the geometric acceptance of the LHCb detector. The name "generator-level" accounts for the fact that in the LHCb simulation software, only decays that end up in LHCb acceptance are saved after generation, to preserve the computing resources. In this analysis, a set of special requirements was applied during the production process of the simulation samples, due to nontrivial topology of the decay that involves long-lived hyperons. In order to reduce the consumption of the computing power, a defined set of geometrical requirements was set to ensure that only signal particles reconstructible as long or downstream tracks are further processed. This selection set is also accounted for by the ϵ_{gen} generator level efficiency. The LHCb simulation group provides official tables with this efficiency split for each data taking period and magnet polarity. A summary of these efficiencies for the different datasets is provided in Table 5.4, where the values have been averaged between the magnet polarities.

Table 5.4: Generator-level efficiencies, %.

Dataset	$\Xi_b^- \to \Xi^- J/\psi$	$\Omega_b^- \to \Omega^- J/\psi$
2015	7.63 ± 0.02	10.28 ± 0.03
2016	7.66 ± 0.02	10.21 ± 0.03
2017	7.66 ± 0.02	10.24 ± 0.03
2018	7.64 ± 0.02	10.25 ± 0.03

It should be noted that the higher generator-level efficiency of the $\Omega_b^- \to \Omega^- J/\psi$ decays is related to the Ω^- hyperons twice shorter lifetime compared to that of the Ξ^- hyperon. Consequently, the Ω_b^- decay products end up more frequently in the LHCb detector acceptance.
It should be noted that the efficiencies determined in this section use the full acceptance of the LHCb detector, while the final measurement is performed in a slightly narrower defined phase-space region. However, the respective fiducial cuts are only applied to the beauty baryons Ω_b^- and Ξ_b^-, which have very similar kinematics. Therefore, the effects from the phase-space selection on generator level will cancel. The same argumentation is true for the filtering efficiencies as both efficiencies are obtained from the generator level efficiencies.

Filtering efficiencies

The filtering efficiency ϵ_{filt} accounts for the fact that the simulation samples were produced using the so-called filtering script, which selects only events that passed selection of the `B2XMuMu` stripping line. This is done to save the disk space by focusing on the events that are actually useful for the analysis. A dedicated script from the LHCb simulation group enables to determine the filtering efficiency from the number of generated events and the one of those that passed the filtering. The summary of the efficiencies is again provided for both decay channels and all data-taking periods, while averaging the efficiencies of both magnet polarities, and displayed in the Tab. 5.5.

Table 5.5: Filtering efficiencies, %

Channel	$\Xi_b^- \to \Xi^- J/\psi$	$\Omega_b^- \to \Omega^- J/\psi$
2015	15.41 ± 0.02	13.58 ± 0.02
2016	15.56 ± 0.01	13.82 ± 0.01
2017	12.95 ± 0.01	13.89 ± 0.01
2018	12.93 ± 0.01	13.87 ± 0.01

While the filtering efficiency is stable for all data-taking periods for the Ω_b^-, the Ξ_b^- ones differ between 2015-16 and 2017-18. This is a result of small changes in the stripping selection for the Ξ_b^-, that can be found in Tab. C.2 and C.3 provided in the App. C. It should be noted, that the stripping line selects several decay modes, therefore the filtering efficiency does not solely account for the $\Omega_b^- \to \Omega^- J/\psi$ and $\Xi_b^- \to \Xi^- J/\psi$ decays.

Selection efficiencies

The final part of the total efficiency, the selection efficiency, is defined on simulation samples as a ratio between the number of reconstructed and selected events and the number of generated events that pass the filtering. The number of selected events is defined by passing the reconstruction and stripping algorithms, the truthmatching selection and the full offline signal selection. Therefore, the selection efficiency accounts by construction the reconstruction, trigger and the truthmatching efficiencies and counter-balances the inclusiveness of the filtering efficiency. All while simultaneously accounting for the offline selection requirements. The last term combines a larger amount of in principle independent efficiencies, since they may be difficult to separate. This particularly true for the reconstruction and offline selection efficiencies, when trying to properly account for multiple candidates and calibrations of the simulation. In order to obtain the most accurate selection efficiency, all corrections are applied when computing the efficiency. It is defined as the ratio of the fully calibrated number of events passing the selection requirements in the reconstructed-level simulated sample, and the weighted (only for the kinematics and decay model [7]) number of events in the generator-level sample:

$$\epsilon_{sel} = \frac{N_{reco-level}^{\text{filtered, reweighted, fully selected}}(\text{truthmatched})}{N_{gen-level}^{\text{filtered, reweighted}}} \tag{5.8}$$

In this step the phase-space acceptance cuts are applied both to generator and reconstruction level simulation. The obtained total selection efficiencies are given in Tab. 5.6. Similarly to the generator efficiency, the reconstruction efficiency of the $\Omega_b^- \to \Omega^- J/\psi$

Table 5.6: Total nominal Run 2 selection efficiencies, %

Dataset	$\Xi_b^- \to \Xi^- J/\psi$	$\Omega_b^- \to \Omega^- J/\psi$
2015	6.82±0.05	12.56±0.06
2016	8.19±0.03	14.64±0.04
2017	9.78±0.04	15.58±0.04
2018	8.80±0.03	14.04±0.04

decay is higher due to the shorter Ω^- lifetime.

Total efficiencies

By computing the product of the above reported generation, filtering, and selection efficiencies, the total efficiencies, needed for the ratio of the production and branching fractions of the Ω_b^- and Ξ_b^-, can be accessed. The total efficiencies have been calculated averaged for all four years using the luminosity weighted simulation. For overview, the Tab. 5.7 does not only contain the total efficiencies for the Ξ_b^- and Ω_b^- independently, but also their ratio as used in the actual measurement. The provided uncertainties are of statistical nature, account only for the size of the simulated samples.

[7] This is a consequence of the fact, that all other variables needed for the calibration are not available at the generator level.

Table 5.7: Total nominal Run 2 efficiencies (including geometric, filtering, reconstruction, selection).

Channel	Value
$\Xi_b^- \to \Xi^- J/\psi$	$(0.923 \pm 0.002) \times 10^{-3}$
$\Omega_b^- \to \Omega^- J/\psi$	$(2.065 \pm 0.004) \times 10^{-3}$
ratio Ω_b^- / Ξ_b^-	2.236 ± 0.007

In addition, the total efficiencies have also been calculated split per track category. These values are needed to perform dedicated cross-checks validating that the result does not depend on the track category. They are calculated by adding the track type selection to the reconstruction level MC only and therefore reflects as fractions. Consequently, the total efficiency of each channel can be reobtained by summing over the individual track type efficiencies $\varepsilon_{total} = \varepsilon_{DDD} + \varepsilon_{LDD} + \varepsilon_{LLL}$. The obtained values are listed in Tab. 5.8.

Table 5.8: Total Run 2 efficiencies (including geometric, filtering, reconstruction, selection), for different track type configurations.

Channel	DDD	LDD	LLL
$\Xi_b^- \to \Xi^- J/\psi$	0.403×10^{-3}	0.432×10^{-3}	0.089×10^{-3}
$\Omega_b^- \to \Omega^- J/\psi$	0.737×10^{-3}	1.071×10^{-3}	0.258×10^{-3}
ratio Ω_b^- / Ξ_b^-	1.828	2.479	2.907

The efficiencies are the lowest for the LLL track category since due to the long lifetimes of the hyperons, most decay chains include at least two downstream tracks. In addition, it can be seen that the efficiency ratios get larger the more long tracks are involved, consistent with the expectations from the ratio of Ω^- and Ξ^- lifetimes.

5.4.3 Invariant mass fit and systematic uncertainties

After the successful calibration and the determination of the efficiencies, the invariant mass fits can be performed. The same fit setup as for the mass measurement described in Sec. 5.3 is used, however, two crucial changes were made. As discussed above, the simultaneous fit is reduced to only two projections, the Run 2 samples for the $\Omega_b^- \to \Omega^- J/\psi$ and $\Xi_b^- \to \Xi^- J/\psi$ decay. A side effect of this is that it reduces the number of degrees of freedom in the fit and improves the fit stability. In addition, the signal yield of the Ω_b^- sample is expressed by the ratio (5.4) to directly extract the statistical uncertainty on the ratio from the fit. The efficiencies are taken from Tab. 5.7 and for the branching fractions of the hyperons the world-averages [33], $\mathcal{B}(\Xi^- \to \Lambda\pi^-) = 0.9989$ and $\mathcal{B}(\Omega^- \to \Lambda K^-) = 0.678$, are used. The uncertainties on the hyperon branching fractions and the ratio of efficiencies are not considered in the fit, but are added as systematic uncertainties. In the final fit to data, the background yields and slopes, the Ξ_b^- yield and width, the mass difference between the beauty baryons and the ratio of the resolution parameters, as well as the ratio (5.4) are kept floating. The projections of the simultaneous fit can be seen in Fig. 5.9. The fitted values of the floating parameters are shown in Table 5.9 and the obtained value will again be called nominal result.

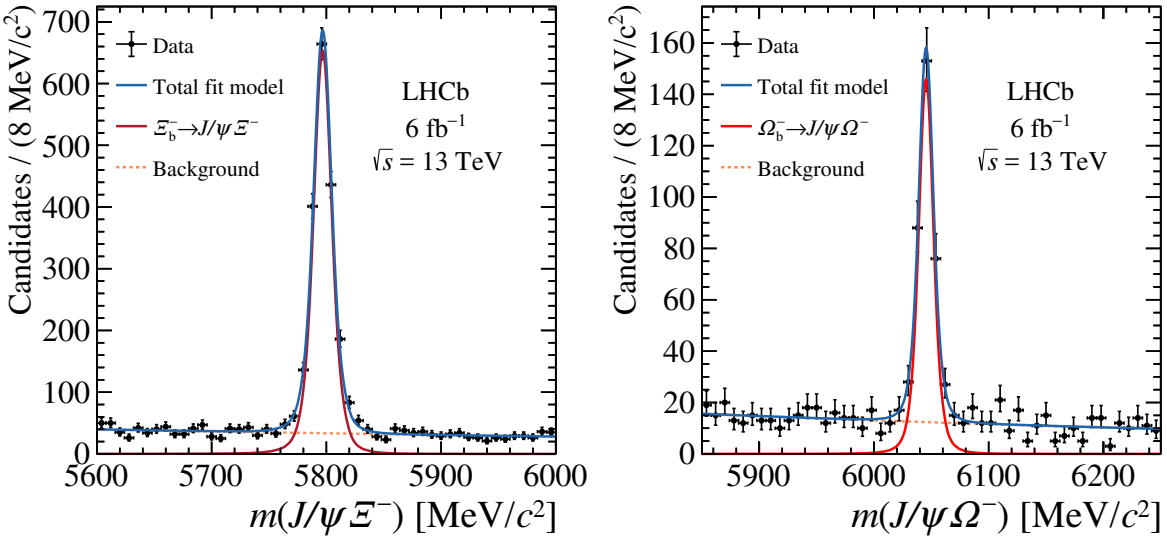


Figure 5.9: Invariant mass distributions in the $\Xi_b^- \to \Xi^- J/\psi$ and $\Omega_b^- \to \Omega^- J/\psi$ datasets, with the results of the simultaneous fit overlaid.

Table 5.9: Fitted values of all free parameters in the fit to data

Parameter	Value
r_{BR}	0.120 ± 0.008
$m(\Omega_b^-) - m(\Xi_b^-)$	248.66 ± 0.55 MeV/c^2
$m(\Xi_b^-)$	5796.68 ± 0.25 MeV/c^2
width ratio Ω_b^-/Ξ_b^-	0.812 ± 0.064 MeV/c^2
width Ξ_b^-	6.64 ± 0.19 MeV/c^2
signal yield Ξ_b^-	1790 ± 47
bkg yield Ω_b^-	622 ± 27
bkg yield Ξ_b^-	1677 ± 46
bkg slope Ω_b^-	$(-1.2 \pm 0.4) \times 10^{-3}$
bkg slope Ξ_b^-	$(-8.6 \pm 2.1) \times 10^{-4}$

Before evaluating the systematic uncertainties and consequently obtaining the final result, the fit and therefore stability of the ratio has been tested by splitting the datasets according to the data-taking periods, the magnet polarities, particles and antiparticles, as well as the three track-type categories (LLL, LDD, DDD). No significant deviations was seen, supporting the stability of the fit procedure. In addition, pull studies have been performed to validate the fit procedure and no bias has been found.

Systematic uncertainties

The systematic uncertainties discussed in this paragraph are determined relative to the nominal results, therefore only the total systematic uncertainty is converted to an absolute value in the end. Several possible different sources of systematic uncertainty are studied in detail and summarised in Table 5.10.
The relative uncertainty on the efficiency ratio due to the size of the available simulation samples is found to be 0.3%. The largest relative uncertainty arises due to the calibration of the simulation and is mostly driven by the smallness of the background-subtracted

dataset, especially for the Ω_b^- decay. The systematic uncertainty associated to this limited amount of data is propagated using a bootstrapping technique [204]. For this procedure, the data candidates used for the determination of the calibration weights are assigned a Poisson-distributed weight of unity mean and the calibration chain is re-evaluated. The procedure is repeated 100 times, resulting in 100 alternative values of the efficiency-ratio. The square root of the mean square of the efficiency ratio is then determined to access the relative systematic uncertainty of 3.9%. As an example, Fig. 5.10 shows the obtained correction weight maps for the transverse momentum of the Ξ_b^- and the Ω_b^-.

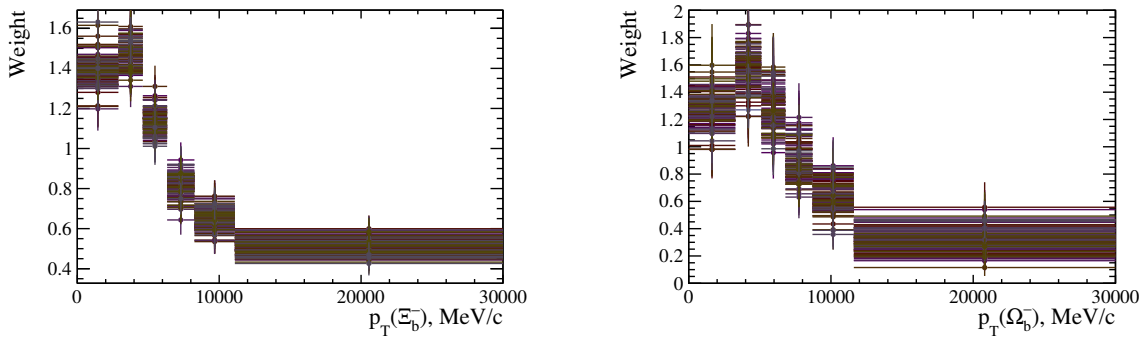


Figure 5.10: Example of the bootstrapping method: 100 bootstrapped correction maps for the $p_T(\Xi_b^-)$ (left), or for the $p_T(\Omega_b^-)$ (right). Each of the 100 maps is painted in a slightly different colour.

Furthermore, alternative binning schemes are tested for the calibration weights on the simulation. Therefore, the original weights before the interpolation are used and the obtained uncertainty is 0.7%. In addition, the kinematics of the Ξ_b^- baryon is alternatively calibrated with a two dimensions weight map, $p_{\mathrm{T}}(\Xi_b^-) : \eta(\Xi_b^-)$, rather than the default calibration of $p_{\mathrm{T}}(\Xi_b^-)$ alone. For the Ω_b^- baryon, the available amount of data is too small to perform the same two-dimensional calibration. But the effect is expected to be similar to the one in case of the Ξ_b^-. The resulting uncertainty is 1.0%.
Accounting for the residual mismodelling of the quality χ^2 of the decay tree fit leads to an relative uncertainty of 1.7%. Similarly, the L0 weights have been removed and no shift in the result had been seen. Therefore, no dedicated systematic uncertainty is needed for the trigger calibration.
The standard calibration procedure used for the calibration of the `ProbNNmu` variables is accounted for by a relative uncertainty of 0.2%, following the recommendations of the

Table 5.10: Relative systematic uncertainties on R.

Source	Uncertainty [%]
Size of simulated samples	0.3
Calibration of simulation	5.5
Selection criteria	0.1
Fit model	0.8
Lifetimes of b baryons	3.1
Material interactions	0.7
External input ($\mathcal{B}$)	1.0
Total	6.5

LHCb PID group. Since the option of using a factor to correct the mismodelling of the event multiplicity in the resampling was used, the alternative approach is utilised to access the related systematic uncertainty. In this case, the event multiplicity is corrected with a dedicated weight and the resampling is run without a scaling factor. The resulting uncertainty due to the treatment of the event multiplicity is 0.5%.
As previously mentioned, only the additional PID cuts applied offline for the two muon candidates have been calibrated, since the simulation samples have been filtered by the Stripping selection. The calibration of the muon ID cuts present in the stripping have been produced with privately produced, unfiltered simulation samples of the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay covering also the J/ψ phase-space using the efficiency maps accessed with `PIDCalib` [189]. It is found that the PID efficiency is flat in q^2 and slightly overestimated in simulation, however this effect is expected to approximately cancel in the ratio between the two decay modes. Nevertheless, as the kinematics can differ slightly, an uncertainty of 0.6% is assigned which is half of the q^2-averaged efficiency difference.
The unfiltered simulation samples of the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ and $\Omega_b^- \to \Omega^-\mu^+\mu^-$ decays, from the rare mode searches, can be used similarly to estimate the effect of the kinematic calibration on the generator and filtering efficiency. The ratios of the weighted and unweighted values of the efficiencies can then be used to determined the ratio between the Ξ_b^- and Ω_b^- for both cases. The combined relative uncertainty from both checks is 0.87%.
A further source of uncertainty originates from imprecise modelling of the downstream tracking efficiency in simulation. While the effect is expected to be small, due to the proper modelling of the track type fractions and hyperon momenta, the stability of the result is tested as a function of the Ξ^- (Ω^-) momentum. This approach enables to test for a momentum-dependent bias and the observed variation of 2.8% from the nominal result is assigned as a systematic uncertainty.
Beside the calibration of the simulation, the efficiency ratio can also be effected by the truthmatching strategy. Including the background category 60, which accounts for unmatched tracks *e.g.* in case of a soft pion, for the Ξ_b^- samples, leads to an uncertainty of 0.2%. Similarly, the ghost probability of a track can be mismodelled. Since the effect can be asymmetric between long and downstream tracks and their proportion differs between the Ω_b^- and Ξ_b^- samples, a dedicated systematic uncertainty is estimated. The requirement suppressing fake tracks is tightened for all five tracks with respect to the default value in the stripping selection for data and simulation. Afterwards, the efficiencies as well as the ratio of fitted yields are recalculated and a systematic uncertainty of 1.25% is determined.
Combining all these individual uncertainties related to the calibration and efficiencies, the overall relative uncertainty is 5.5%.
For an estimation of the systematic uncertainty related to the selection, the effect of the Ξ^- veto in the Ω_b^- dataset is studied. While the mass veto can shape the background, the more concerning effect is that the efficiency of the Ξ^- mass veto can be incorrectly estimated, if the invariant mass resolution is not well modelled in simulation. This effect is probed by widening and narrowing the vetoed mass region and determining the shift on the obtained ratio of the production and branching fractions. The associated relative uncertainty is estimated to be 0.1%.
Several sources of systematic uncertainty related to the fit model are considered. A fit bias of 0.1% is obtained from 20 000 pseudoexperiments generated and fitted with the

nominal strategy. In addition, the biases with respect to the chosen signal and background models have been evaluated. The bias on the signal model has been accessed by generating 10 000 pseudoexperiments according to an alternative signal model based on a single double-sided Crystal-Ball function. The resulting bias when fitting the generated dataset with the default fit model is found to be 0.1% as well. For the background model, two checks where performed, first pseudoexperiments were generated according to a linear background model and the resulting relative uncertainty is found to be 0.1%. The more crucial second check includes a dedicated $\Lambda_b^0 \to \Lambda J/\psi$ background component in the $\Xi_b^- \to \Xi^- J/\psi$ fit model. While the nominal result is not changed, when the $\Lambda_b^0 \to \Lambda J/\psi$ background fit component is included, the respective bias due to the statisical limitation is accessed with pseudoexperiments. It is found to amount to 0.8% and the total relative systematic uncertainty related to the fit model is therefore 0.8%.
Since the simulation only uses the nominal value of the beauty baryon lifetimes [33], the final result can be biased due to the imprecise knowledge of their values. In order to access the related systematic uncertainties, the uncertainties on the Ω_b^- and Ξ_b^- baryon lifetimes are propagated to the final result by shifting the baryon decay-time distributions with a dedicated exponential weight. The obtained relative uncertainty on the production ratio of 3.1% is driven by the relatively poor knowledge of the Ω_b^- baryon lifetime.
Long-lived hyperons can interact with the detector material prior to their decay, this can affect the measurement in two different manners. First, events where the decay process is altered due to the hadronic interaction are not taken into account in the generator level simulation. Although the respective lost events can be recovered, their fraction is not equal for the Ξ_b^- and Ω_b^- decay. Therefore, a dedicated uncertainty of 0.5% is assigned. The second more crucial uncertainty arises from mismodelling of the material interactions in simulation. This is a result of the poor knowledge of the cross-sections of hyperons. A dedicated systematic uncertainty is assigned by comparing the hyperon material-interaction probability estimated with two different versions of the GEANT4 simulation toolkit [8] [205]. They use different models for hadronic interaction cross-sections and the relative uncertainty is found to be 0.5% as well and, therefore, the total relative uncertainty due to material interactions is 0.7%.
Lastly, the uncertainties on the branching fractions of the $\Omega^- \to \Lambda K^-$ and $\Xi^- \to \Lambda\pi^-$ decays are propagated to the final result by summing these two relative uncertainties in quadrature, leading to a relative uncertainty of 1.0%. The uncertainty is mostly driven by the uncertainty on the Ω^- branching fraction $\mathcal{B}(\Omega^- \to \Lambda K^-) = (67.8 \pm 0.7)\%$ compared to the relatively small uncertainty for the Ξ^- baryon one, $\mathcal{B}(\Xi^- \to \Lambda\pi^-) = (99.887 \pm 0.0035)\%$. Combining all these individual systematic uncertainties leads to a total uncertainty of 6.5%, which translate to a shift of 0.008. Taking into account the central value of our nominal result, the following result is obtained for the ratio of production fractions times branching fractions

$$r_{BR} = 0.120 \pm 0.008(\text{stat}) \pm 0.008(\text{syst}), \tag{5.9}$$

in the kinematic region $2 < \eta < 6$ and $p_{\mathrm{T}} < 20$ GeV/c for the Ξ_b^- and Ω_b^- baryons. It is the first time this quantity was evaluated at the LHC at a center-of-mass energy of 13 TeV.
To summarise, the obtained result is lower than the CDF measurement [78] by about

[8] The GEANT4 versions used are `v9r6p4` and `v10r6p2`.

1.3 standard deviations. Since the CDF measurement [78] was used to estimate the production fractions of the Ω_b^- baryon at the LHC, the resulting overestimation can affect the measurement of the total $b\bar{b}$ production cross-section which makes assumptions on the fractions of different species [89]. However, due to the very small production rate of the Ω_b^- baryon, the effect on the total $b\bar{b}$ cross-section is most likely negligible.
Lastly, since the theoretical predictions of the branching fractions of the $\Omega_b^- \to \Omega^- J/\psi$ decay vary and no direct measurements of the $\Omega_b^- \to \Omega^- J/\psi$ and $\Xi_b^- \to \Xi^- J/\psi$ branching fraction exists, no attempt on disentangling the ratio of production fractions from the one of branching fractions was made.

Chapter 6

Towards observation of rare Ξ_b^- decays

The study of $b \to s\ell^+\ell^-$ transitions with other weakly decaying beauty baryons than the Λ_b^0 offers large potential to increase the knowledge with respect to the B anomalies. Out of the three remaining beauty baryons, the Ξ_b isospin doublet has larger production fractions when compared to the Ω_b^- baryon. The Ξ_b^- baryon additionally creates mostly charged particles in its decay chain, improving its reconstruction efficiencies significantly when compared to the neutral Ξ_b^0 baryon. When combined with the unique topology, equally to the one of the tree-level decay $\Xi_b^- \to \Xi^- J/\psi(\to \mu^+\mu^-)$ in Fig. 5.1 in Ch. 5, very low background levels are expected, motivating the search for the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay. Therefore, this chapter is dedicated to the first measurement of the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay and measuring the $\Xi_b^- \to \Xi^-\psi(2S)$ decay mode with the same final state particles. This chapter will only discuss additional background sources and calibration steps due to the large overlap with the $\Xi_b^- \to \Xi^- J/\psi$ decay in Ch. 5. As of the point of writing this thesis, the analysis is still blind and therefore, no unblinded results will be shown for the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay. This chapter will cover the analysis strategy, specific backgrounds and the dedicated selection strategy, as well as the fit models and the efficiency determination. Lastly, systematic uncertainties and some pre-unblinding checks will be presented. The described analysis has again been performed together with Dr. Vitalii Lisovskyi.

6.1 Current theoretical and experimental knowledge

Before going into details about the analysis strategy, it is helpful to reflect a bit more on the current knowledge of rare Ξ_b^- decays. During the duration of this PhD, the measurement searching for the complementary radiative decay mode $\Xi_b^- \to \Xi^-\gamma$ has been published [71,206]. This decay mode is interesting, since baryonic $b\to s\gamma$ transitions enable to measure the photon polarisation [198,207] as the helicity structure is preserved in weak decays. No signal was found in the Run 2 dataset and therefore an upper limit was set. But the measurement only uses signal candidates, where the Ξ^- and Λ hyperon decayed within the VELO detector due to reconstruction constraints, which significantly reduces the amount of signal. In addition, decay modes with two muons in the final states are easier to trigger and improve the reconstruction of the Ξ_b^- decay vertex significantly.

This can be seen by the discussed $\Xi_b^- \to \Xi^- J/\psi$ decay mode in the previous section, but also for more suppressed decay modes like $\Xi_b^- \to \Xi^- \psi(2S)$, which was just recently observed for the first time by the CMS collaboration [18][1].
The $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ branching fraction was first predicted using the light-cone QCD sum-rule in Ref. [208]. When integrating the prediction of Ref. [27], nearly the same result is obtained, but the latter is more precise. This is mostly driven due to the effect that the knowledge on the required input has been improved over the decade between both predictions. Another prediction of the differential branching fraction is given in Ref. [209]. However, this prediction relies on the LHCb measurement of $\mathcal{B}(\Lambda_b^0 \to \Lambda \mu^+ \mu^-)$, which can be affected by long-distance or NP effects. Therefore, the latter is not a pure SM prediction. It is nevertheless listed with the other predictions in Tab. 6.1. All these predictions expect a branching fraction of about 2×10^{-6} for the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay, which is about a factor two larger than all other known $b \to s \ell^+ \ell^-$ transitions.

Table 6.1: Theoretical predictions for $\mathcal{B}(\Xi_b^- \to \Xi^- \mu^+ \mu^-)$. The first two predictions are based on the light-cone QCD sum rule (LCQCD) technique. The third prediction is not an independent SM prediction as it relies on experimental $\mathcal{B}(\Lambda_b^0 \to \Lambda \mu^+ \mu^-)$.

Approach	Value
LCQCD [208]	$(2.23 \pm 0.67) \times 10^{-6}$
LCQCD [27]	$(2.33 \pm 0.38) \times 10^{-6}$
$SU(3)$ [209]	$(1.87^{+0.56}_{-0.54}) \times 10^{-6}$

From a theoretical viewpoint the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ is also interesting when compared to the $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ decay. Both decay modes proceed through a spin 1/2 to spin 1/2 hadronic transition, therefore the decay modes have only two differences. These are the respective form factors and one spectator quark ($u \to s$) that differs and consequently affects the wave functions of the transition due to the two strange quarks for the Ξ^- baryon. However, the effects are minor and the form factor predictions for the Ξ_b^- decay [210] also predict most of the signal to be located in the high-q^2-region similar to the $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ case. A more recent binned in q^2 prediction [27] through the light-cone QCD sum rules also determines that most of the signal is in the high-q^2-region. Both predictions [27, 210] are displayed in Fig. 6.1. It should be noted, that for mesons the form factor predictions lead to most of the signal being located in the central q^2-region just below the J/ψ resonance. Therefore, due to the similarity between the predictions for the $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ and $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decays, the latter are a powerful tool to cross-check if this might be a general behaviour for baryonic $b \to s \ell^+ \ell^-$ transitions.

6.2 Analysis strategy

The main target of this measurement is the first observation of the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay using the full Run 1 and Run 2 datasets of the LHCb experiment. All three track type

[1] It should be noted that the unique topology of a cascade of hyperon decays leads to an absence of hadron PID requirements. Therefore, the decay mode is less challenging to reconstruct for general purpose detectors with no dedicated PID detectors.

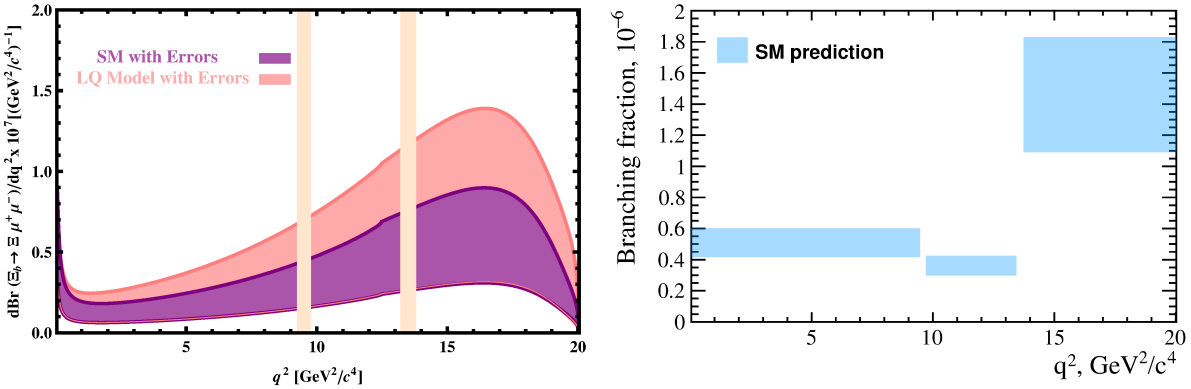


Figure 6.1: Left: The light pink band shows the predicted SM differential decay rate of the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay from Ref. [210] (taken from the paper). Right: binned values of absolute SM branching fraction of the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay from Ref. [27] (our illustration from Tab.4 of their paper). Same as in the $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ case, most of signal is expected at high q^2 in both theoretical papers.

categories, as discussed in Sec. 5.1, will be utilised to maximise the available statistics. In order to minimise the systematic uncertainties on the measurement of interest, the branching fraction of the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay is measured with respect to the $\Xi_b^- \to \Xi^- J/\psi$ branching fraction

$$r_{BR} = \frac{BR(\Xi_b^- \to \Xi^- \mu^+ \mu^-)}{BR(\Xi_b^- \to \Xi^- J/\psi(\to \mu^+ \mu^-))}. \tag{6.1}$$

Since both decay modes have the same initial and final states, not only do a lot of systematic uncertainties cancel, but also the Ξ_b^- production fraction and all hyperon decay branching fractions do so. Since the $\Xi_b^- \to \Xi^- J/\psi$ transition is a tree-level CKM-favoured decay and therefore has a branching fraction several orders of magnitude larger than the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay, it can additionally be used to calibrate the simulation and study all background sources that are not specific to the rare decay mode. No other decay mode with a higher production and/or branching fraction, like the $\Lambda_b^0 \to \Lambda J/\psi$ decay, offers the same benefits.

The $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ dataset is kept blind in a $\pm 60\,\mathrm{MeV}/c^2$ mass window around the known Ξ_b^- mass, and has not been unblinded before this thesis was handed in. Therefore, the unblinding strategy will be discussed at the end of this section. The unique topology and presence of downstream tracks create a non-trivial q^2-dependence of the detector efficiency, which will need to be taken into account in order to provide an integrated and differential measurement of the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ branching fraction. Additionally, it is expected that the integrated result will have a larger uncertainty due to the poor knowledge of the $\Xi_b^- \to \Xi^- J/\psi$ decay. The differential result will be provided in three bins of q^2: below the J/ψ resonance (region 1), between the J/ψ and $\psi(2S)$ (region 2) and above the $\psi(2S)$ resonance (region 3). Beside the expected higher precision, it enables to cross-check the shape of the differential distribution of the $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ decay. The exact details of the different q^2 regions are given in Tab. 6.2. The window between 2900 and 3000 MeV/c^2 is deliberately excluded, as it is dominated by the radiative tail of the J/ψ resonance.

Table 6.2: Dimuon invariant mass windows considered in the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ analysis.

Range	Mass requirement
Region 1	$m(\mu^+\mu^-) < 2900\,\text{MeV}/c^2$
Region 1a	$0.1 < q^2 < 8\,\text{GeV}^2/c^4$
J/ψ window	$3000 < m(\mu^+\mu^-) < 3150$ MeV/c^2
Region 2	$3150 < m(\mu^+\mu^-) < 3586$ MeV/c^2
Region 2a	$11 < q^2 < 12.5\,\text{GeV}^2/c^4$
$\psi(2S)$ window	$3586 < m(\mu^+\mu^-) < 3736$ MeV/c^2
Region 3	$m(\mu^+\mu^-) > 3736$ MeV/c^2
Region 3a	$15 < q^2 < 20\,\text{GeV}^2/c^4$

In order to enable an easier comparison with the $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ branching fraction in the future, an alternative binning scheme denoted by "region 1a", "region 2a" and "region 3a" is provided. It should be noted that the difference between the nominal and Λ_b^0 binning scheme in the high-q^2-region is only minor since the difference in the masses of the Ξ_b^- and Ξ^- baryons, and therefore the upper kinematic cut-off, is only $q^2 = 20.03$ GeV$^2/c^4$. Therefore, it is not covered in this thesis and only the default binning scheme is discussed.

6.2.1 Prospects of the $\psi(2S)$ channel

When plotting the invariant mass of $\Xi^- \mu^+ \mu^-$ as a function of dimuon invariant mass[2]; as displayed in Fig. 6.2, not only the J/ψ resonance is visible as vertical band but also the $\psi(2S)$ one. The white horizontal band corresponds to the blinded signal window for the $\Xi^- \mu^+ \mu^-$ invariant mass.

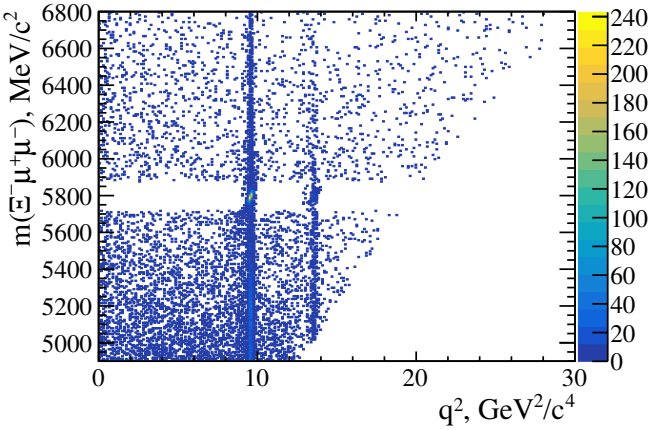


Figure 6.2: Invariant mass of $\Xi^- \mu^+ \mu^-$ candidates as a function of dilepton invariant mass squared (q^2), in the data. The full Run 2 dataset is exploited to prepare this plot, with preselection and a loose BDT cut applied. The signal window of the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay is blinded.

This motivates to also study the tree-level transition, $\Xi_b^- \to \Xi^- \psi(2S)(\to \mu^+ \mu^-)$, which

[2]This plot has a preselection applied, including a majority of the cuts entering the final selection described in Sec. 6.3.2.

will also be measured relatively to the J/ψ decay mode

$$r^{\psi}_{BR} = \frac{\mathcal{B}(\Xi_b^- \to \Xi^- \psi(2S)(\to \mu^+\mu^-))}{\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi(\to \mu^+\mu^-))}. \tag{6.2}$$

During the process of preparing this measurement the decay mode $\Xi_b^- \to \Xi^- \psi(2S)(\to \mu^+\mu^-)$ has not been observed, therefore the ratio (6.2) will be published as a standalone measurement as well. While the CMS collaboration [18] just published the first observation of this decay mode, the obtained result in this measurement will be significantly more precise, as it will be shown in the following sections.
Beside its observation, the ratio of the charmonia decays of the Ξ_b^- baryon offers a powerful tool to cross-check deviations observed with the ratio of the respective Λ_b^0 decay modes

$$r^{\psi}_{BR}(\Lambda_b^0) = \frac{\mathcal{B}(\Lambda_b^0 \to \Lambda\psi(2S))}{\mathcal{B}(\Lambda_b^0 \to \Lambda J/\psi)}. \tag{6.3}$$

The measurements [23, 24] of the Λ_b^0 ratio (6.3) disagree with the covariant quark model prediction[3], motivating the standalone measurement of the $\Xi_b^- \to \Xi^-\psi(2S)$ decay. This is particularly true, since the charmonia decays of Ξ_b^- and Λ_b^0 differ only by the flavour of one spectator quark. This also reflects in the fact that the predictions [22] of the ratios for the charmonia decay modes between the Λ_b^0 decays, $\mathcal{B}(\Lambda_b^0 \to \Lambda\psi(2S))/\mathcal{B}(\Lambda_b^0 \to \Lambda J/\psi) \sim 0.82$, and Ξ_b^- decays, $\mathcal{B}(\Xi_b^- \to \Xi^-\psi(2S))/\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi) \sim 0.84$, are very close. Therefore, the measurement of the ratio (6.2) will provide valuable input to understanding the baryon form-factors, which is crucial for predictions of baryonic $b \to s\ell^+\ell^-$ transitions.

6.2.2 Towards absolute branching fractions

Theoretical interpretation of the reported branching fractions might require the absolute branching fractions rather than the experimentally cleaner ratios. In order to obtain, an absolute branching fraction, either the absolute branching fraction of the chosen normalisation mode $\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi)$ needs to be known, or if a different one like $\Lambda_b^0 \to \Lambda J/\psi$ is chosen, the absolute production rate of the Ξ_b^- baryon, $f_{\Xi_b^-}$. However, no absolute branching fraction of any Ξ_b^- decay nor its production rate is known. The latter could in principle be determined using the semileptonic decay mode $\Xi_b^- \to \Xi_c^+ \mu^- \bar{\nu}_\mu$, since the semileptonic decay widths are predicted to be nearly equal for all b-hadrons [211]. In order to utilise this fact, one would need to measure the decay mode $\Xi_b^- \to \Xi_c^+ \mu^- \bar{\nu}_\mu$ and its counterpart $\Xi_b^- \to \Xi_c^{+*} \mu^- \bar{\nu}_\mu$ through excited charm-hadron states. Afterwards, an extrapolation can be applied to account for the suppressed decay channels based on $b \to u$ transitions, that are also challenging to reconstruct. Beside the fact that this determination of the absolute Ξ_b^- production rate, $f_{\Xi_b^-}$, would require a fully independent measurement, its precision will greatly suffer from the poor knowledge of the absolute branching fractions of charm baryons. Therefore, it is equally precise to determine the absolute branching fraction $\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi)$ through $SU(3)$ symmetry relations [212–214]

[3] For detailed covariant quark level discussion see 7.1

with the $\mathcal{B}(\Lambda_b^0 \to \Lambda J/\psi)$ decay:

$$\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi) = \frac{3}{2} \times \frac{\tau_{\Xi_b^-}}{\tau_{\Lambda_b^0}} \times \mathcal{B}(\Lambda_b^0 \to \Lambda J/\psi) \approx 1.6 \times \mathcal{B}(\Lambda_b^0 \to \Lambda J/\psi), \tag{6.4}$$

where $\tau_{\Xi_b^-}/\tau_{\Lambda_b^0}$ is the ratio of the average lifetimes [197] of the Ξ_b^- and Λ_b^0, $1.57 \pm 0.04\,\text{ps}$ and $1.47 \pm 0.01\,\text{ps}$, respectively. Unfortunately, the value of $\mathcal{B}(\Lambda_b^0 \to \Lambda J/\psi)$ is not known precisely either, therefore the value has been computed in the same way as in the LHCb analysis on $\Xi_b \to \Xi\gamma$ search [71] but with more recent PDG values. In order to extract the value of $\mathcal{B}(\Lambda_b^0 \to \Lambda J/\psi)$, the ratio of branching fractions and production fractions with respect to the $B^0 \to K_S^0 J/\psi$ can be used. The ratio has been measured by the D0 experiment [215]:

$$\frac{f_{\Lambda_b^0} \times \mathcal{B}(\Lambda_b^0 \to \Lambda J/\psi)}{f_d \times \mathcal{B}(B^0 \to K_S^0 J/\psi)} = 0.345 \pm 0.047. \tag{6.5}$$

In addition, the *heavy flavour averaging group* (HFLAV) updated their 2018 value of the production fractions at Tevatron using the CDF result and incorporating the current knowledge of branching fractions of charm hadrons [216], leading to:

$$\frac{f_{\Lambda_b^0}}{f_u + f_d} = 0.229 \pm 0.062. \tag{6.6}$$

This can be translated to $f_{\Lambda_b^0}/f_d$ by using the isospin symmetry $f_u = f_d$. Similarly, the $\mathcal{B}(B^0 \to K^0 J/\psi) = (8.68 \pm 0.30) \times 10^{-4}$ from the PDG2020 values [33] can be translated to the branching fraction $\mathcal{B}(B^0 \to K_S^0 J/\psi) = (4.34 \pm 0.15) \times 10^{-4}$ by assuming an equal production of K_L and K_S eigenstates in weak decays. Now combining the Eqs. (6.5)-(6.6), finally leads to the branching fraction of the $\Lambda_b^0 \to \Lambda J/\psi$ decay

$$\mathcal{B}(\Lambda_b^0 \to \Lambda J/\psi) = (3.27 \pm 1.00) \times 10^{-4}. \tag{6.7}$$

Finally, with Eq. (6.4) gives the branching fraction of the $\Xi_b^- \to \Xi^- J/\psi$ decay

$$\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi) = (5.2 \pm 1.6 \pm 1.6 \pm 0.1) \times 10^{-4} = (5.2 \pm 2.3) \times 10^{-4}, \tag{6.8}$$

where the first uncertainty comes from $\mathcal{B}(\Lambda_b^0 \to \Lambda J/\psi)$, the second – from $SU(3)$ breaking and the third from the lifetime uncertainties; while the last expression combines all uncertainties.

With this result (6.8), the SM prediction of the branching fraction $\mathcal{B}(\Xi_b^- \to \Xi^- \mu^+ \mu^-)$, discussed in Sec. 6.1, can be converted into the units of r_{BR} to obtain a reference value for the observables. Only the most recent result from Ref. [27] will be discussed. The central value integrated over q^2, $(2.33 \pm 0.38) \times 10^{-6}$ can be translated into the SM expectation of $r_{BR}(\text{SM}) \approx 0.0045 \pm 0.0007$ where the uncertainty arises from the SM prediction, while a very asymmetric uncertainty from the normalisation branching fraction has to be added separately.

The predicted branching fraction in Ref. [27] is also provided in three q^2-regions that can be translated into the r_{BR} values: $r_{BR}(\text{region I}) \approx 0.0010 \pm 0.0002$, $r_{BR}(\text{region II}) \approx 0.0007 \pm 0.0001$, and $r_{BR}(\text{region III}) \approx 0.0028 \pm 0.0007$. These numbers are used as expectations on the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ signal yield. However, it should be noted that the theoretical and experimental q^2-regions do not fully align, since the latter needs to take experimental resolutions from the charmonia resonances into account.

6.2.3 Blinding and unblinding strategy

In this measurement only the signal of the unobserved $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay mode is kept blind. The efficiencies for all decays of interest, as well as the mass projections of the charmonia decay modes, $\Xi_b^- \to \Xi^- J/\psi$ and $\Xi_b^- \to \Xi^-\psi(2S)$, are not kept blind to enable cross-checks.
The blinding range of $5737 - 5857$ MeV/c^2 in the $\Xi^-\mu^+\mu^-$ invariant mass is expected to contain above 96% of the signal in the nominal Ξ_b^- fit range ($5600 - 6000$ MeV/c^2). Therefore, the blinding window is also described as ($\sim 3\sigma$) window around the nominal Ξ_b^- mass.
A dedicated unblinding strategy is developed, to account for the attempt to measure the differential $\Xi_b^- \to \Xi^-\mu^+\mu^-$ branching fraction and not only integrated over q^2 one. After obtaining the authorisation to unblind from the collaboration, the statistical significance of the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ signal obtained from the merged dataset (all three q^2 regions combined) will be determined and checked if it exceeds three standard deviations. Afterwards, the statistical significance of the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ signal will be quantified in each of the three q^2 regions separately.
Under the assumption of observing the signal with at least three standard deviations in all three q^2 regions, the measurement of the differential branching fraction in all three bins will be pursued. If no significant signal is found, an upper limit on $\Xi_b^- \to \Xi^-\mu^+\mu^-$ branching fraction will be set in each bin individually and in the integrated q^2 region. The integrated BF will be provided as a sum of differential BFs in the three q^2 regions, corrected by the vetoed charmonium regions. The corrections for the excluded charmonium regions will be estimated using the physical model from the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ simulation. This step is done since, contrary to the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay, no form factor predictions are available yet for the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay.

6.3 Specifics of the rare Ξ_b^- decay channel

As already mentioned, the background sources and calibration strategy largely overlaps with that for the $\Xi_b^- \to \Xi^- J/\psi$ decay mode discussed in Ch. 5. Therefore, this chapter will only be used to elaborate on the background sources that are specific to the rare mode and treated differently. Additionally, the selection in this analysis is based on a BDT, therefore, its specifics will be discussed. Similarly, the calibration discussion will only give a brief overview and discuss specifics for the calibration strategy of the Run 1 simulation samples and application to the rare decay mode.

6.3.1 Additional background sources

The dominant background source is, similarly to the charmonia decays, of combinatorial nature. However, in the non-resonant q^2 windows, muons can also form random combinations. In particular, products of semileptonic b-hadron decays can be combined with one random muon. Therefore, the combinatorial background is significantly larger and is suppressed with a dedicated multivariate classifier described later in Sec. 6.3.2. It should be noted, that in the LLL track type category the relative level of combinatorial background is larger since more random combinations are formed by the prompt-like tracks created at

the PV or in decays of short-lived hadrons. With further displacement, this background gets smaller, explaining the lower combinatorial background level in the LDD and DDD track type categories.
Similarly to the $\Xi_b^- \to \Xi^- J/\psi$ decay channels, lepton misidentification backgrounds are expected to be negligible in all q^2 regions and are further suppressed by the muon ID requirement `ProbNNmu>0.1`. However, dedicated checks for backgrounds from double pion misidentification as muons are performed. The charmless decay $\Xi_b^- \to \Xi^- \pi^+ \pi^-$ mode is expected to have a small branching fraction and has never been observed, additionally similar charmless backgrounds have been found to be negligible in other $b \to s\ell^+\ell^-$ processes. Therefore, the same assumption is made for the $\Xi_b^- \to \Xi^- \pi^+ \pi^-$ decay and two dedicated post-unblinding checks will be performed to confirm it. First, the tighter PID selection will be applied and, second, checking the invariant mass distribution when a pion mass hypothesis is applied to both muons. The most prominent misidentified decay would be $\Xi_b^- \to \Xi_c^0 (\to \Xi^- \pi^+) \pi^-$, but no sign of a Ξ_c^0 mass peak ($m_{PDG} = 2470$ MeV/c^2) is seen in $\Xi^- \mu^+$ invariant mass when a pion mass hypothesis is applied on the muon, as can be seen in the left plot of Fig. 6.3.

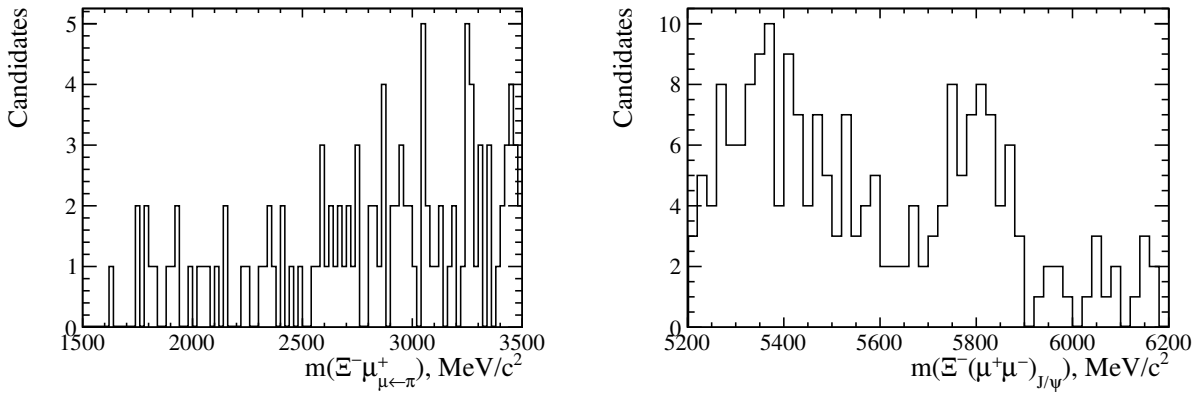


Figure 6.3: $\Xi^- \mu^+$ invariant mass with pion mass hypothesis on the muon, in the selected data of the q^2 region below the J/ψ resonance (left). No sign of a Ξ_c mass peak is found. The J/ψ-constrained $\Xi^- \mu^+ \mu^-$ invariant mass in the region $2900 < m(\mu^+\mu^-) < 3000$ MeV/c^2 (right). A small peak from the radiative tail of the $\Xi_b^- \to \Xi^- J/\psi$ decay is visible.

While backgrounds from higher charmonia resonances are negligible in the full q^2 region, the radiative tail of the $\Xi_b^- \to \Xi^- J/\psi$ decay is a source of background in the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay as it can be seen in the right plot in Fig. 6.3. The source of the tail is the emission of soft photons and energy losses in interactions with the detector material and removed by excluding the region $2900 < m(\mu^+\mu^-) < 3000$ MeV/c^2.
Another source of background could be posed by the currently unobserved $\Xi_b^0 \to \Xi^- \pi^+ J/\psi$ decay, that has an predicted low branching fraction due to symmetry constraint of the wave function. A dedicated check showed that it can be safely neglected.
Equivalent to the J/ψ decay mode, the most significant physical background source arises from the over-reconstructed $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ (including $\Lambda_b^0 \to \Lambda J/\psi$) background with a random charged pion. However, this background is mostly rejected by the selection and the remaining background is studied as a dedicated source for a systematic uncertainty by including the relevant component in the mass fits. the same $\chi^2_{\rm IP}(\pi_\Xi) > 20$ cut as in the

$\Xi_b^- \to \Xi^- J/\psi$ channel is applied. Additionally, the remaining background is studied as a dedicated source for a systematic uncertainty by including the relevant component in the mass fits.

6.3.2 Signal selection

This section is dedicated to the signal selection needed to observe the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ and $\Xi_b^- \to \Xi^- \psi(2S)$ decays. The selection has been developed based on the experience gained from the background studies and optimised to reach the best sensitivity for the observables of interest.
As first step of the selection, a trigger and stripping selection are applied. Equivalent to the previous analysis, most of the selection power for the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay arises from the muons, therefore the decay has been triggered on signal through the `L0Muon` or `L0DiMuon` trigger selection for the hardware stage, and the HLT trigger selection also utilises muonic trigger lines and additional topological triggers in the HLT2 stage to increase the trigger efficiencies. The details of the HLT trigger selection can be found in App. B and for the stripping selection in App. C. Since the selection requirements largely overlap with those discussed in Sec. 5.2.3, only the details of the multivariate analysis will be described in the following. The MVA selection is performed with a gradient boosted decision tree algorithm (BDT) (for general informartion see App. A.2), and optimised independently for the $\Xi_b^- \to \Xi^- \psi(2S)$ and $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay mode. In order to ensure the cancellation of related systematic uncertainties the respective BDT cut is also applied to the $\Xi_b^- \to \Xi^- J/\psi$ normalisation mode. A summary of the offline selection requirements applied throughout this analysis, is presented in the Tab. 6.3 for Run 1 and Tab. 6.4 for Run 2.

Table 6.3: Offline selection requirements for Run 1.

Type	Particle(s)	Requirement	Sample
Quality and acceptance	μ	`InAccMuon, hasRich`	all
	all L tracks	`TRACK_GhostProb<0.3`	all
	p, π	`TRACK_Type == 3 || 5`	all
PID	μ	$p > 3\,\text{GeV}/c$, $p_T > 800$ MeV/c	all
	μ	$p < 150$ GeV/c	all
	μ	`ProbNNmu>0.1`	all
Mass windows	dilepton	$3000 < \text{m}(\mu^+\mu^-) < 3150$ MeV/c^2	$J/\psi(\mu^+\mu^-)$
		$3586 < \text{m}(\mu^+\mu^-) < 3736$ MeV/c^2	$\psi(\mu^+\mu^-)$
		all except $2900 - 3150$ and $3586 - 3736$	rare
	Λ	$\lvert m(p\pi^-) - 1116\rvert < 8$ MeV/c^2	all
	Ξ^-	$\lvert m(\Lambda\pi^-)_{\Lambda\,\text{constr}} - 1322\rvert < 10$ MeV/c^2	all
Background rejection		$\chi^2_{\text{IP}}(\pi_\Xi) > 20$	all
	clones	$\theta(\pi_\Xi, \pi_\Lambda) > 0.5$ mrad	all
Combinatorial background	–	`BDT>0.35`	$\psi(\mu^+\mu^-)$ & norm.
	–	`BDT>0.94; BDT>0.85 (Region 3)`	rare & norm.

Table 6.4: Offline selection requirements for Run 2.

Type	Particle(s)	Requirement	Sample
Quality and acceptance	μ	`InAccMuon, hasRich`	all
	all tracks	`TRACK_GhostProb<0.3`	all
	p, π	`TRACK_Type == 3 || 5`	all
PID	μ	$p > 3\,\text{GeV}/c$, $p_T > 800$ MeV/c	all
	μ	$p < 150$ GeV/c	all
	μ	`ProbNNmu>0.1`	all
Mass windows	dilepton	$3000 < \text{m}(\mu^+\mu^-) < 3150$ MeV/c^2	$J/\psi(\mu^+\mu^-)$
		$3586 < \text{m}(\mu^+\mu^-) < 3736$ MeV/c^2	$\psi(\mu^+\mu^-)$
		all except $2900 - 3150$ and $3586 - 3736$	rare
	Λ	$\lvert m(p\pi^-) - 1116\rvert < 8$ MeV/c^2	all
	Ξ^-	$\lvert m(\Lambda\pi^-)_{\Lambda\,\text{constr}} - 1322\rvert < 10$ MeV/c^2	all
Background rejection		$\chi^2_{\text{IP}}(\pi_\Xi) > 20$	all
	clones	$\theta(\pi_\Xi, \pi_\Lambda) > 0.5$ mrad	all
Combinatorial background	–	`BDT>0.35`	$\psi(\mu^+\mu^-)$ & norm.
	–	`BDT>0.97; BDT>0.85 (Region 3)`	rare & norm.

Training of the multivariate classifier

The BDT has been trained on Run 2 samples, using truthmatched[4] $\Xi_b^- \to \Xi^-\mu^+\mu^-$ simulation samples as signal proxy and the upper and lower signal sidebands of the collision data (excluding the regions dominated by the charmonia resonances) as a background proxy. The lower sideband is defined as $5000 < m(\Xi^-\mu^+\mu^-) < 5600$ MeV/c^2; and the upper one as $5900 < m(\Xi^-\mu^+\mu^-) < 7000$ MeV/c^2. Since little partially reconstructed background is expected, the lower sideband is included in the training to gain statistical power and therefore a more reliable training. The same BDT classifier is applied to the different data taking periods and q^2 regions, although independent BDT selection cut values are determined for the $\psi(2S)$ and rare mode datasets. This choice is made to ensure that the BDT efficiencies largely cancel in the final normalised branching fractions and avoids introducing any bias due to different BDTs. Since the main objective is the observation of the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay mode, the q^2 regions with the J/ψ ($\psi(2S)$) resonances are excluded by vetoes of 150 (50) MeV around their known world-average masses. This removes the risk of biases due to different background compositions between the rare mode and the charmonia decay modes. Beside other reasons, this motivated to study the step-by-step efficiencies also for all track type categories independently. No significant effect is found and the BDT cut efficiency is similar between all categories. The described preselection is applied, and the entire calibration chain (Sec. 6.3.3), except for the trigger corrections, is applied to the simulation samples in order to correct for the data-simulation discrepancies prior to the training. A dedicated grid search showed the best performance for a training setup with 250 trees, a maximal depth of trees of four and a learning rate of 0.11. A large number of input variables has been studied and reduced to a final set of 15 variables describing the kinematics and geometry of the decay.
The classifier has been checked for overtraining by utilising the Kolmogorov-Smirnov (KS)

[4]The details of the truthmatching will be discussed as part of the calibration Sec. 6.3.3.

test [217] for a test and training dataset as shown in the left plot of Fig. 6.4. No sign for overtraining is found. Additionally, a 10-fold cross-validation is performed. The resulting ten ROC curves are shown in the right plot of Fig. 6.4 and show a very good agreement between all ten folds.

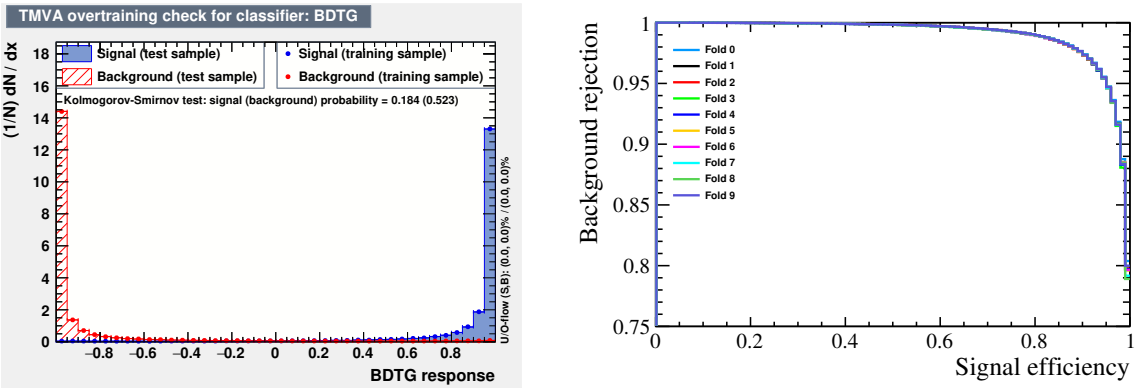


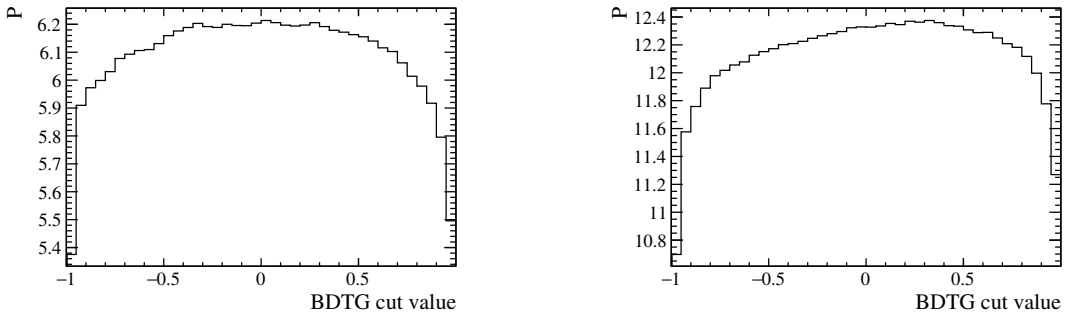
Figure 6.4: Left: BDT response distributions for signal and background proxy samples, compared between the training and testing samples. Right: ROC curves for the ten folds.

Additional information such as the BDT input variables, their distributions for the signal and background sample, their correlations and their importance for the BDT classifier can be found in App. E.2.

Cut optimisation for the $\Xi_b^- \to \Xi^-\psi(2S)(\to \mu^+\mu^-)$ channel

The individual BDT cut optimisation with figures of merit for each studied decay channel enables to achieve the optimal sensitivity for them. Since the resonant $\Xi_b^- \to \Xi^-\psi(2S)$ channel is not blinded and a signal peak is clearly visible even without the BDT cut, the chosen figure of merit (FOM) is defined as $P = \frac{S}{\sqrt{S+B}}$, where S is defined by $S = \epsilon_{\text{BDT}} \cdot S_0$. S_0 is the signal yield obtained from the fit to data prior to any BDT cut[5] and the efficiency of each BDT cut ϵ_{BDT} is obtained from simulation at each BDT cut value. Finally, B is the background yield in the signal region obtained from a fit at a given BDT cut. Since the final result will be obtained from a simultaneous fit to the Run 1 and Run 2 samples, the BDT cut value is also determined for both datasets independently to correctly reflect the global shift in kinematics due to change in center-of-mass collision energy. The obtained distributions of the FOM value P as a function of the BDT cut value, are shown in Fig. 6.5 for Run 1 and Run 2.
For the Run 2 dataset, the BDT cut of > 0.35 is chosen, since it provides the highest FOM values and simultaneously keeps a high signal efficiency. While the Run 1 FOM distributions suggests a looser cut value, the fact that the FOM shape is roughly flat around a large range, it was decided to align the BDT cut with Run 2 and also use the value > 0.35.

[5]The choice of obtaining the S_0 value from data is made to avoid a potential bias due to the poor knowledge of Ξ_b^- production rate, which can bias the estimated signal yield when using branching fractions.



(a) Run 1 (b) Run 2

Figure 6.5: BDT cut value optimisation for the $\Xi_b^- \to \Xi^- \psi(2S)(\to \mu^+\mu^-)$ channel by plotting the FOM values as a function of the BDT cut value.

Cut optimisation for the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ channel in the full q^2 range

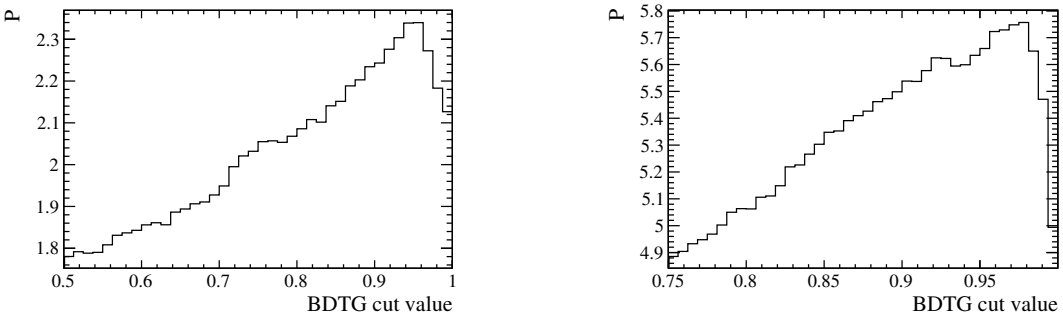
For the BDT cut optimisation of the rare $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay, the two FOM approaches are studied to ensure an optimal chosen value. In both approaches the background yield is obtained from a fit to data at a given BDT cut, however the signal yield cannot be estimated with fits to data, since the signal region is still blinded. Therefore, the expected signal yield S_0 for the FOM $\frac{S}{S+B}$, as already used for the $\Xi_b^- \to \Xi^- \psi(2S)$ decay channel, is defined by

$$\begin{aligned} S_0 = \mathcal{L} \times \sigma_{b\bar{b}} \times 2 \times f_{\Xi_b^-} \times \mathcal{B}(\Xi_b^- \to \Xi^- \mu^+ \mu^-) \times \varepsilon_0 \\ \times \mathcal{B}(\Xi^- \to \Lambda \pi^-) \times \mathcal{B}(\Lambda \to p\pi^-); \end{aligned} \tag{6.9}$$

where $\mathcal{L}$ is the integrated luminosity in the given dataset; $\sigma_{b\bar{b}}$ is the cross-section of creation the $b\bar{b}$ pair at a given center-of-mass energy; the factor 2 accounts for both b quark and anti-quark which can hadronise into a b hadron; $f_{\Xi_b^-}$ is the hadronisation fraction to Ξ_b^- (see Tab. 1.2); $\mathcal{B}(\Xi_b^- \to \Xi^- \mu^+ \mu^-)$ is assumed to be of about 10^{-6} similar to all other $b \to s\ell^+\ell^-$ transitions; and finally ε_0 is the total efficiency without any BDT cut applied. In order to avoid introducing a q^2 dependence trough the efficiency ε_0, the $\Xi_b^- \to \Xi^- J/\psi$ decay channel efficiency is used in the calculation of the expected signal yield S_0. Finally $\mathcal{B}(\Xi^- \to \Lambda\pi^-) \times \mathcal{B}(\Lambda \to p\pi^-)$ are the branching fraction of the subsequent hyperon decays. This estimation provides the expected signal yield of about 48 (8) events in Run 2 (Run 1), however it should be noted that the uncertainties on these expected signal yields are large, *e.g.* the theoretical prediction of the $\mathcal{B}(\Xi_b^- \to \Xi^- \mu^+ \mu^-)$ would require to include an additional factor two.
The distribution of FOM for Run 1 (left) and Run 2 (right) is presented in Fig. 6.6. A BDT cut value of 0.94 is chosen for Run 1 and 0.97 for Run 2.

Cut optimisation for the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ channel at high q^2

Finally, the study of the background level in the blinded invariant mass plots, as shown in Sec. 6.5.3, reveals that the background level drastically differs across the q^2. In particular, in the high-q^2 region (region 3) the background level is significantly lower compared to other q^2 windows. Therefore, it is decided to loosen the BDT cut value for the high-q^2



(a) Run 1 $S/\sqrt{S+B}$ (b) Run 2 $S/\sqrt{S+B}$

Figure 6.6: BDT cut value optimisation for the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ channel by plotting the FOM values as a function of the BDT cut value for the Run 1 and Run 2 dataset.

region in order to maximise the signal yield. Indeed, when comparing FOM curve in the high-q^2 region with the corresponding FOM in Fig. 6.6, where the full q^2 is considered, it is found that looser BDT cut values are preferred. The maximum of the FOMs are found around $0.85-0.9$. Similarly to Run 2, the maximum is reached at about $0.86-0.9$ for the Run 1 dataset. However, it should be noted, that only $\mathcal{O}(1-5)$ background events survive at tight BDT cut values leading to statistical fluctuations in the FOM distributions. In this measurement, the BDT cut value of 0.85 is chosen for the high q^2 region in both Run 1 and Run 2.
Finally, the chosen BDT cut values are summarised in the full selection summaries in Tab. 6.3 for Run 1 and Tab. 6.4 for Run 2.

Multiple candidates

As a last step of the selection, a search for candidates sharing the same event and run numbers is performed. In data, when such multiple candidates are found, only one is retained randomly for the analysis, and all the others are removed. This corresponds to a removal of about 0.8% of candidates in the fully selected dataset. In the simulation, only the candidate with the lowest background category is kept if multiple candidates are found. If there are several candidates sharing the same category, the one to be kept is selected randomly in the same manner as for data. These multiple candidates arise due to the not-fully matched simulated events when including the background category 60 in the truthmatching of the simulation.

6.3.3 Calibration of the simulation

Before discussing the calibration for the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ simulation samples, a quick recap on the truthmatching strategy is needed. It was found that in simulation the decay products of long-lived hyperons are sometimes not properly matched reflecting in signal events in the background category 60. Therefore, it is decided to truthmatch the Ξ_b^- through the background categories 0, 10[6], 50 or 60 to ensure that no bias from unmatched

[6]The background category 10 is the quasi-signal category, where all final particles are correctly reconstructed and connected to the correct initial hadron, but one part in the decay chain *e.g.* a intermediate

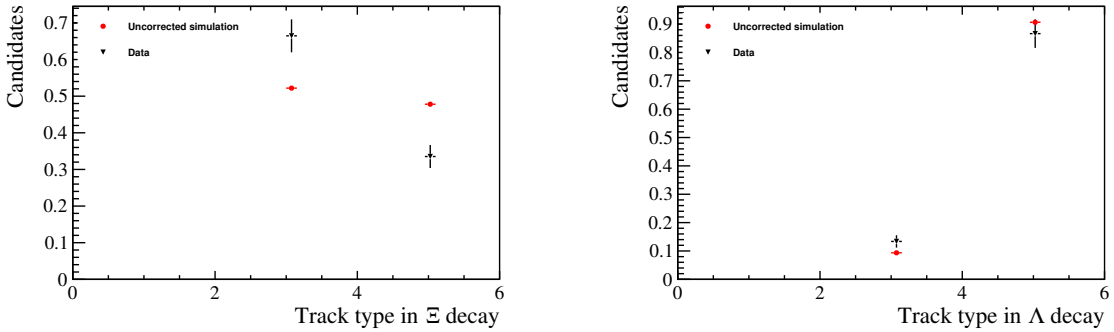


Figure 6.7: Relative proportions of long (type=3) and downstream (5) tracks in (left) Ξ^- decay and (right) Λ decay in Run 1 data and simulation samples.

pions is produced. Finally, the truthmatching through category 0, 50 or 60 is applied to the Ξ^- and Λ hyperons to reject events with swapped pions. Both latter contributions create a shape broader than the signal, and their potential effect on the signal yield is covered by the dedicated systematic uncertainty on the matching procedure discussed later. Since the calibration of the simulation is based on the $\Xi_b^- \to \Xi^- J/\psi$ channel and extrapolated to the other decay modes, the strategy largely overlaps with that discussed in Sec. 5.4.1. The same selection as discussed in App. D is used to obtain the background-subtracted data. For the PID calibration of the Run 1 data sample, the same resampling strategy through the `PidGen` method of the `PIDCalib` package was used. The next calibration needed for the Run 1 dataset is the fraction of the long and downstream tracks, since the kinematic calibration is performed with all track categories merged. Therefore, any mismodelling of the fraction could cause a bias in the calibrations of the kinematics since the vertexing efficiency is lower for downstream tracks. There are significant discrepancies for the Ξ^- children tracks in the Run 1 samples, as can be seen in Fig. 6.7. This is explained by a significant data-simulation mismodelling of the track ghost probability variable in Run 1 for downstream tracks. A cut is placed in the stripping on the pion from the Ξ^- decay (`GhostProb` < 0.5) which introduces this additional disagreement. The resulting improved agreement between data and simulation after the calibration can be seen in Fig. 6.8. Moreover, the effect of the correction is expected to nearly cancel in the ratio of the branching fractions and a dedicated systematic uncertainty is applied to Run 1 efficiency determination accounting for this assumption.
In order to avoid any potential biases due to remaining mismodelling of the Run 1 track type proportions and the ghost probability response, the kinematic corrections are evaluated using the Run 2 datasets and then ported to the Run 1 dataset. The same approach is used to correct the decay models of the Ξ^- and Λ hyperons with the momentum asymmetries of the daughter particles. Therefore, the correction strategy for these two calibrations is equivalent to the one explained in Sec. 5.4.1. Similarly, the trigger response is found to only need calibration of the L0 stage as described in Sec. 4.2.3. Residual mismodelling for the HLT response will be determined through a dedicated systematic uncertainty. One of the most powerful variables used in the training of the multivariate classifier in this analysis is the quality of the decay-tree-fit, χ^2_{DTF} per degree of freedom (see App. E.2.1). The benefit of this variable is that it has information about the whole decay chain and

resonance cannot be matched.

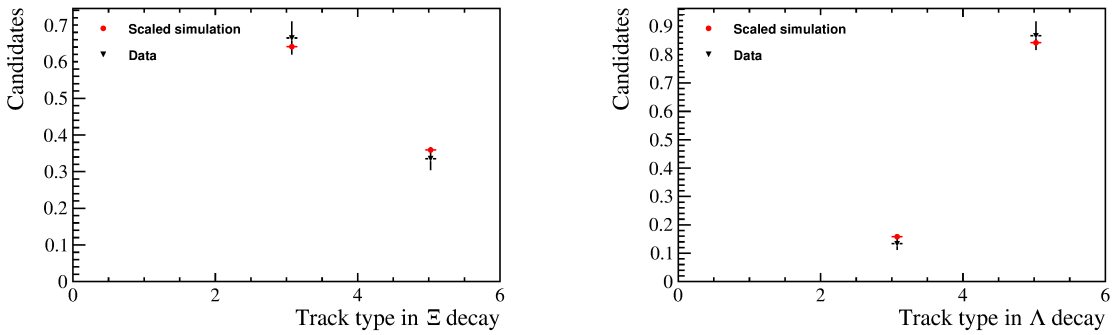


Figure 6.8: Relative proportions of long (type=3) and downstream (5) tracks in (left) Ξ^- decay and (right) Λ decay in Run 1 data and simulation samples after the scaling is applied to simulation.

has only a PV constraint. Consequently, it does not create any mass-dependent bias in the MVA training. However, after the full calibration is applied residual mismodelling was found as it can be seen in Fig. D.8 in App. D.2. Therefore, the DTF calibration as described in Sec. 5.4.1 is applied as part of the default calibration.

Decay model of the rare $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay

The final calibration step, is crucial for the correct branching fraction measurement of the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay mode. Since no form factor predictions are available, the simulation has been produced flat in phase-space. However, it is expected that the q^2 distribution has a similar shape to the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay. Additionally, the phase-space simulation has flat helicity angles for the muons, that need to be calibrated[7]. The helicity angles, assuming an unpolarised Ξ_b^- baryon, are displayed in Fig. 6.9.

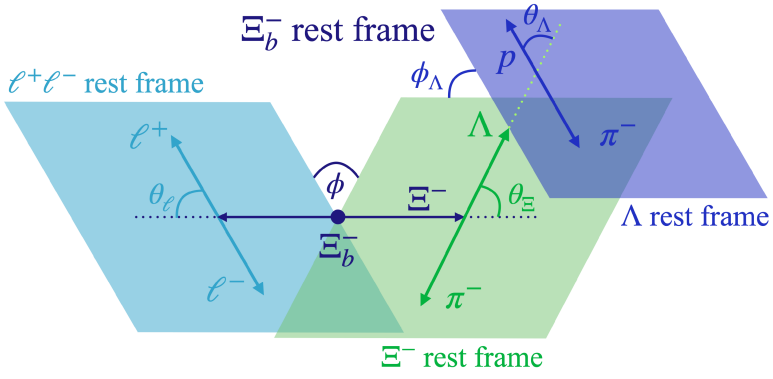


Figure 6.9: Angular distributions of the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay, in the helicity basis assuming an unpolarised Ξ_b^- baryon.

Under the assumption of $SU(3)$ symmetry, the mismodelling of the q^2 and muon helicity angle distributions can be calibrated with $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ simulation samples. The latter can be produced with a physics model based on lattice QCD predictions of the form

[7]The other helicity angles are better modelled due to the physics model used for the hyperon decays. Additional, residual mismodelling is corrected due to the momentum asymmetry weights.

factors in Ref. [102]. The same approach cannot be used for the hadronic side, due to the extra intermediate step with the Ξ^- hyperon in the Ξ_b^- decay chain and differing in the parity-violating asymmetry parameters in the Ξ^- and Λ decays. Consequently, the shape of the distributions of θ_Ξ in the Ξ_b decay and θ_Λ in the Λ_b^0 decay are expected to differ. However, the angle ϕ between the dimuon and first hyperon plane is expected to be similar as well. Therefore, the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ angular distributions in θ_ℓ and ϕ, as well as the q^2 distribution are calibrated at generator level to match the ones of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ simulation. After ensuring that all phase-space requirements align between both samples, a *GBReweighter* algorithm from the *hep-ml* package [218] is trained to reweight the Ξ_b distributions to the Λ_b^0 ones in three dimensions, as shown in Fig. 6.10(bottom).

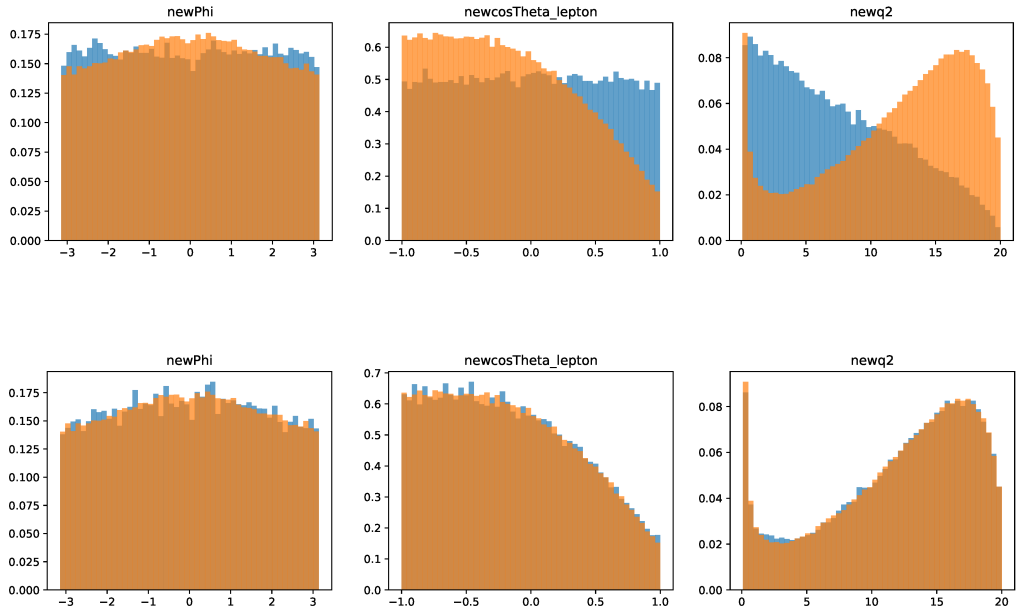


Figure 6.10: Top: distributions of ϕ, $\cos\theta_\ell$ and q^2 in (orange) $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ with the physics model and (blue) $\Xi_b^- \to \Xi^-\mu^+\mu^-$ with the phase-space model. Bottom: result of a multivariate reweighting of the Ξ_b distributions to match the Λ_b^0 ones.

Validation of the calibration

After applying the full correction chain described above, a check is performed for the data-simulation agreement in various variables. Most importantly, the data-simulation agreement is studied for the BDT output, and is found to be reasonable, as shown in Fig. 6.11 and 6.12.

No concerning deviations between the background substracted and calibrated data are found.

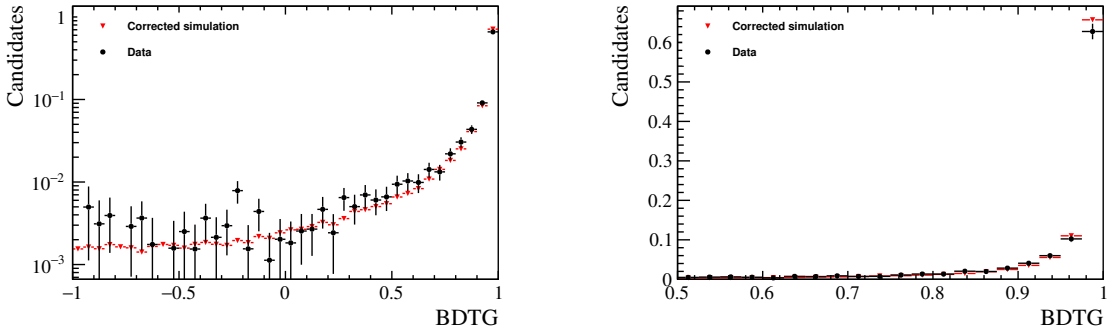


Figure 6.11: Comparison of the BDT output variable, between the $\Xi_b^- \to \Xi^- J/\psi$ data (black circles) and simulation (red triangles), for Run 2. A linear-scale zoom to the region BDT>0.5 is shown to the right.

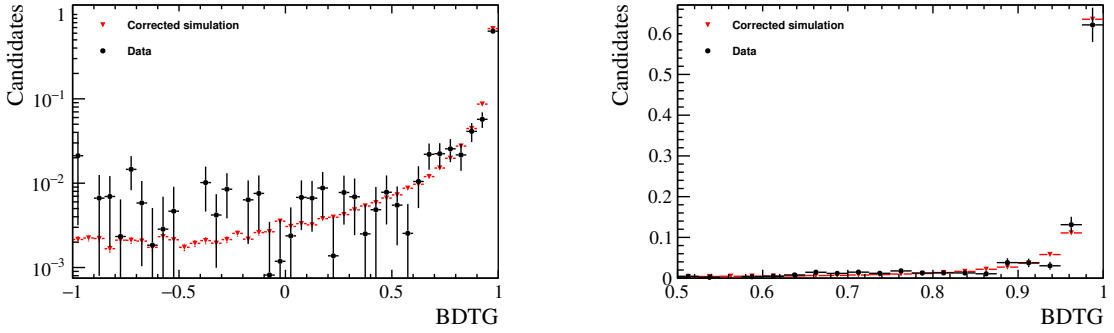
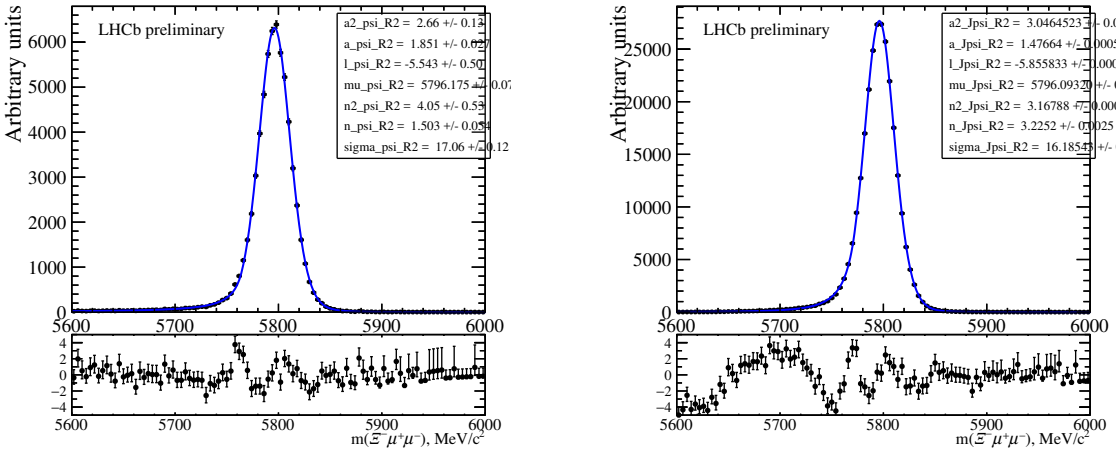


Figure 6.12: Comparison of the BDT output variable between the $\Xi_b^- \to \Xi^- J/\psi$ data (black circles) and simulation (red triangles), for Run 1. A linear-scale zoom to the region BDT>0.5 is shown to the right.

6.4 Fit model for the three decay modes

In each of the three decay modes, the signal is parameterised by a double-sided Hypatia function and the remaining background is modelled with an exponential. As in all previous analyses, the tail parameters are fixed to the values obtained from simulation. However, there are two important caveats that need to be considered. First, the $\Xi_b^- \to \Xi^- J/\psi$ and $\Xi_b^- \to \Xi^- \psi(2S)$ decay modes have mass constraints applied not only to the Ξ^- and Λ hyperons, but also to the respective charmonia. In the rare mode, the latter constraint is not possible. Therefore, it is important to study the effect of the mass constraints and how removing them affects the fits. This is done, since the charmonia mass constraints significantly improve the mass resolution on the Ξ_b mass. The constraint on the hyperon masses is not so useful, because the momenta of hyperon decay products (and so, their uncertainties) are much lower compared to muon momenta. Second, if the dataset is very small, it might be useful to fix the signal resolution as well to improve the fit stability. This is especially true if a multidimensional fit is needed. Therefore, a dedicated study is performed to compare the mass resolution between fits to simulation and data. The study of the effect of the charmonia mass constraint revealed, that the signal width gets wider without the constraint and the signal yield is changing slightly. A dedicated systematic uncertainty is therefore evaluated. The fits are performed with Run 2 data and both the

$\Xi_b^- \to \Xi^- J/\psi$ and $\Xi_b^- \to \Xi^- \psi(2S)$ decay modes. It is concluded that no extra treatment is needed when removing the J/ψ mass constraint for the normalisation fit in the rare mode. However, the comparison of the signal width of fits to data and simulation revealed that the resolution is in general under-estimated in simulation by about 10%. Therefore, the width obtained from simulation is multiplied by a factor 1.1, when fixed in fits to data. Additionally, it was found that the signal yield for both charmonia decay modes remains the same when the signal width is fixed and kept floating. Further details to undermine this discussion can be found in App. E.4.
In order to confirm the observation of the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay, all available datasets and q^2 regions are merged. The corresponding fits to $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ and $\Xi_b^- \to \Xi^- J/\psi$ signal simulation samples, where all years are merged, are shown in Fig. 6.13. Similarly, Fig. E.12 shows the fits to the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ simulation for the three different q^2 regions in App. E.4. This is important since the fit for final measurement of the branching fraction (or setting an upper limit on it) will be performed independently for the three regions.



(a) $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ (b) $\Xi_b^- \to \Xi^- J/\psi$

Figure 6.13: Fits to the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ and $\Xi_b^- \to \Xi^- J/\psi$ signal simulation, for the merged Run 1 and Run 2 simulation.

6.4.1 Efficiency calculations

The efficiency calculation follows the exact same scheme as described in Eqn. 5.7 in the previous Ch. 5 in Sec. 5.4.2. Therefore, this section only discusses significant deviations from the strategy in the previous chapter. It should be noted, that if two or more years are merged, the efficiencies of each year are weighted with a respective luminosity weight to mimic the same importance of a simulation sample as the respective data sample has.

Generator-level efficiency

The (average over q^2) generator-level efficiency is provided as a single number for each data taking period and magnet polarity in Tab. E.3 in App. E.5. The efficiencies for two magnet polarities are averaged. Since, the generator-level efficiency is a priori a function of q^2, a dedicated study is performed to obtain the q^2 trend of the generator-efficiency

and the details are discussed in App. E.5. The resulting generator-efficiency maps as a function of q^2 normalised to $\Xi_b^- \to \Xi^- J/\psi$ channel are shown in Fig. 6.14 for Run 1 (left) and Run 2 (right).

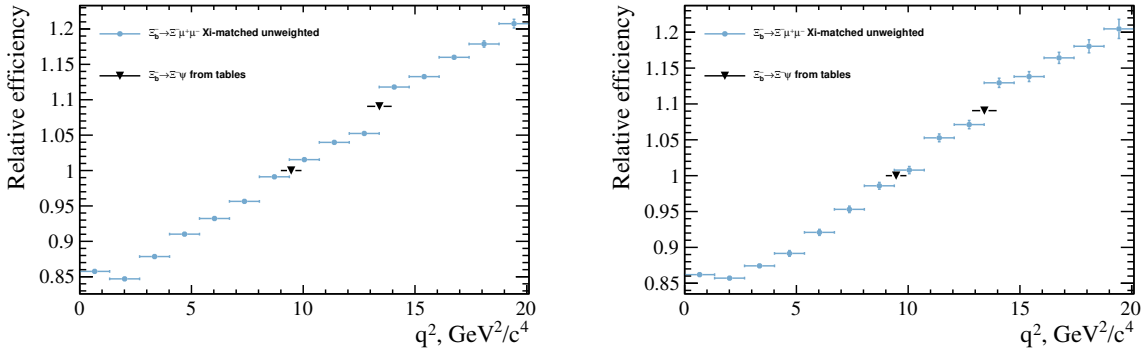


Figure 6.14: Trend of generator-level $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ efficiency as a function of q^2 normalised to the efficiency of the $\Xi_b^- \to \Xi^- J/\psi$ channel for Run 1 (left) and Run 2 (right). Uncertainties arise from the limited size of the produced simulated sample. For completeness, the values of generator-level efficiency for $\Xi_b^- \to \Xi^- J/\psi$ and $\Xi_b^- \to \Xi^- \psi(2S)$ channels, taken from Tab. E.4, are also shown in black.

The trend of the efficiency with q^2 arises from requirements on the daughters of long-lived particles to fall into the acceptance of LHCb. On average, the Ξ^- has lower momentum at high q^2, which allows it to more often decay before the magnet and be reconstructed, leading to higher efficiency at high q^2.

Filtering efficiency

The filtering efficiency ϵ_{filt} accounts for the fact that the simulation samples were produced using the filtering script, selecting only events passing the `B2XMuMu` stripping line. The summary of the (averaged over q^2) filtering efficiencies for all year and channels, merged for the magnet polarities can be found in Tab. E.5 in App. E.5 and the normalised to the $\Xi_b^- \to \Xi^- J/\psi$ filtering efficiency ones in Tab. E.6. Once again, a significant difference is observed between the two charmonium decay modes, initiating the study of the filtering efficiency as a function of q^2.
The maps of filtering efficiency as a function of q^2 and normalised to the $\Xi_b^- \to \Xi^- J/\psi$ mode efficiency maps are again obtained with a dedicated study described in App. E.5 and are shown in Fig. 6.15 for 2011-2012 (top left), 2015-2016 (top right) and 2017-2018 (bottom). The 2011-2012 map is also studied separately from 2015-2016, to account for possible differences in reconstruction in Run 1. Additionally, on the right side of the same figure the are shown. It can be noted that the difference between 2011-2012 and 2015-2016 is negligible in the ratio.

Selection efficiency

Since the stripping line is inclusive and selects several decay channels, this reflects on the filtering efficiency. However, the selection efficiency is constructed in a way that

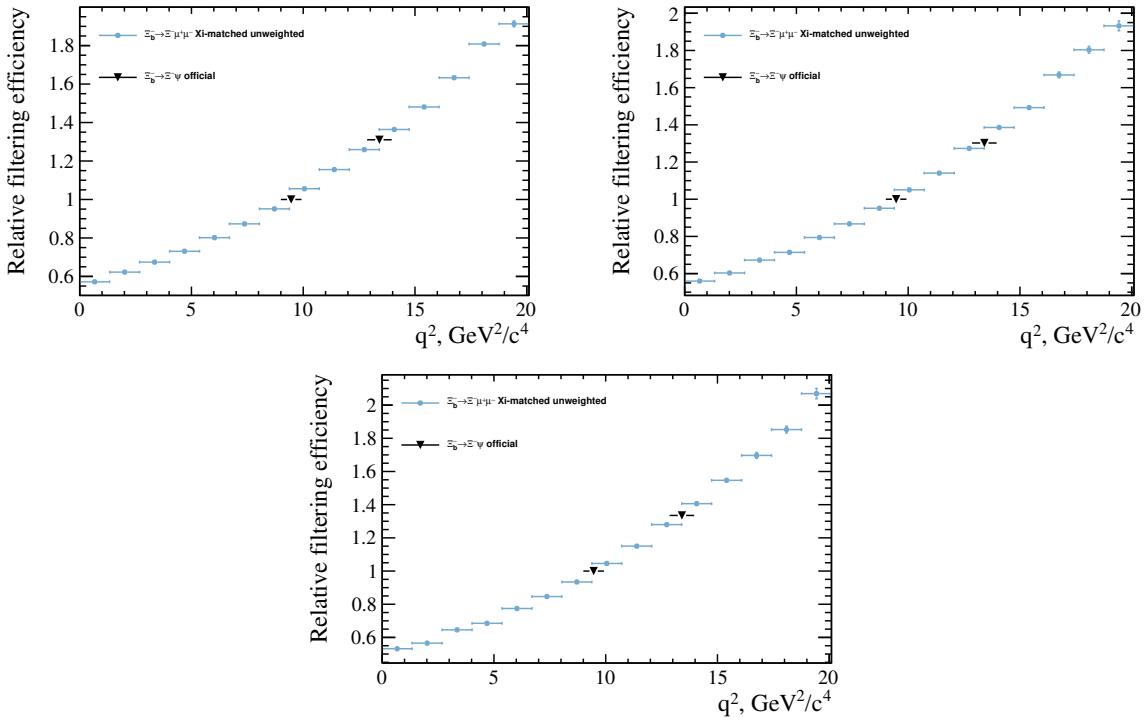


Figure 6.15: Trend of filtering $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ efficiency as a function of q^2 normalised to the efficiency of the $\Xi_b^- \to \Xi^- J/\psi$ channel in (top left) 2011-2012, (top right) 2015-2016 and (bottom) 2017-2018. Uncertainties arise from the limited size of the privately produced simulated sample. For completeness, the values of similar efficiency (ratios) for $\Xi_b^- \to \Xi^- J/\psi$ (defined as 1) and $\Xi_b^- \to \Xi^- \psi(2S)$ channels are also shown in black, taken from Tab. E.6.

includes the reconstruction and truthmatching efficiency for the final state of interest, which counterbalances the inclusiveness of the filtering efficiency.
The total selection efficiency as a function of q^2 for the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay mode is displayed in Fig. 6.16 in absolute units on the left, and normalised to the $\Xi_b^- \to \Xi^- J/\psi$ efficiency on the right.

The overall difference between 2015-2016 and 2017-2018 is largely caused by the equivalent difference in filtering efficiency (so, stripping selection). Further difference at low q^2 arises from the less efficient HLT2 selection in 2015-16 compared to 2017-18. The jump in efficiency around $q^2 = 14\,\mathrm{GeV}^2/c^4$ is due to the looser BDT cut in region 3. The *normalised* efficiency in region 3 is lower in Run 1 compared to Run 2, which is mostly driven by the different efficiency of the BDT cut used for the normalisation mode. Further differences arise from the trigger efficiencies, however, after normalised to the selection efficiency of the $\Xi_b^- \to \Xi^- J/\psi$ channel, the 2015-2016 and 2017-2018 datasets have very similar efficiencies at higher values of q^2. The remaining difference at low q^2 is driven by the thresholds of the L0 triggers.

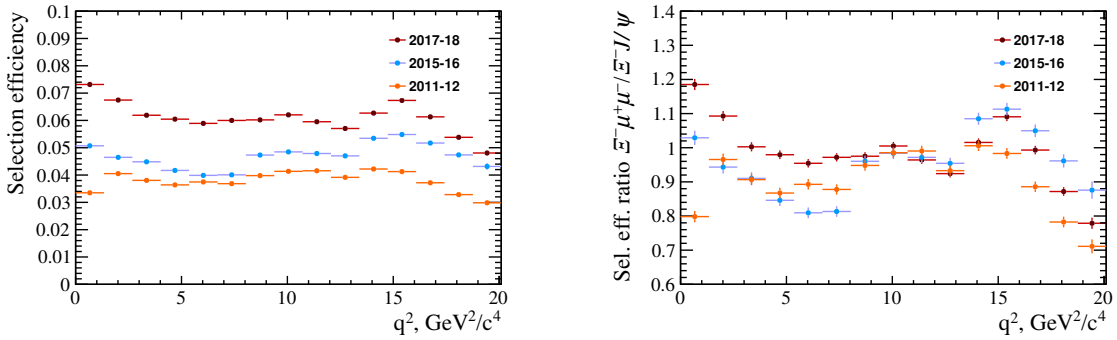


Figure 6.16: Trend of the selection $\Xi_b^- \to \Xi^-\mu^+\mu^-$ efficiency as a function of q^2 in (orange) 2011-2012, (blue) 2015-2016 and (red) 2017-2018, (left) absolute and (right) normalised to the efficiency of the $\Xi_b^- \to \Xi^- J/\psi$ channel. Uncertainties arise from the limited size of the simulated samples.

Total efficiency

Finally, the total efficiency is computed as a product of the generator, filtering, and, as one step, reconstruction & selection efficiencies. Tab. 6.5 lists the total efficiencies for the $\Xi_b^- \to \Xi^- J/\psi$ and $\Xi_b^- \to \Xi^-\psi(2S)$ modes, and their ratio. The efficiencies are lower in Run 1 due to suboptimal stripping (mostly due to tight fake track probability cuts) and trigger selections. However, the ratio of efficiencies between the two decay modes is very similar between Run 1 and Run 2. Additionally, the efficiencies split per track type category, used for cross-checks, are given in Tab. E.7 in App. E.5.1. Similarly, by

Table 6.5: Total efficiencies for the $\Xi_b^- \to \Xi^- J/\psi$ and $\Xi_b^- \to \Xi^-\psi(2S)$ modes, and their ratio. Uncertainties shown account for the size of simulated samples only.

Channel	Run 2	Run 1
$\Xi_b^- \to \Xi^- J/\psi$	$(0.876 \pm 0.002) \times 10^{-3}$	$(0.591 \pm 0.002) \times 10^{-3}$
$\Xi_b^- \to \Xi^-\psi(2S)$	$(1.157 \pm 0.004) \times 10^{-3}$	$(0.782 \pm 0.004) \times 10^{-3}$
ratio $\psi(2S)/J/\psi$	1.321 ± 0.005	1.323 ± 0.005

multiplying the three efficiency maps (geometric, filtering and selection), each already normalised to the $\Xi_b^- \to \Xi^- J/\psi$ channel, the final map of relative efficiency as a function of q^2 is obtained. It is shown in Fig. 6.17 and only a slight difference between 2015-2016 and 2017-2018 datasets is seen. For Run 1, the difference is larger due to difference in the BDT cuts for the normalisation channel as well as in trigger selection.
The map of relative efficiency, presented in Fig. 6.17, needs to be averaged over each q^2 bin. This cannot be done reliably with the samples produced using the phase-space model. However, it is expected that the actual q^2 distribution of the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay is very similar to the one of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay. Therefore, the ratio of the generator-level q^2 distributions in the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ and $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decays displayed in the left plot of Fig. E.13 in App. E.5 are taken to calculate the q^2 weight shown in the right plot of Fig. E.13 to correct for the q^2 shape. This q^2 weight is then added to the generator-level sample of $\Xi_b^- \to \Xi^-\mu^+\mu^-$ without any acceptance cuts applied to obtain the total efficiency-ratio averages over each of three given q^2 bins.

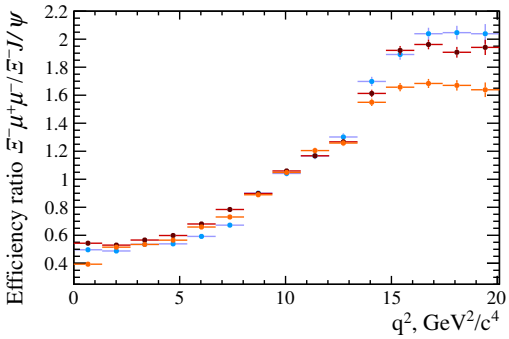


Figure 6.17: Trend of the total $\Xi_b^- \to \Xi^-\mu^+\mu^-$ efficiency as a function of q^2 in (orange) 2011-2012, (blue) 2015-2016 and (red) 2017-2018 periods, normalised to the efficiency of the $\Xi_b^- \to \Xi^- J/\psi$ channel. Uncertainties arise from the limited size of the simulated samples.

Fig. 6.18 displays the obtained efficiency ratios for the q^2 bins with the q^2 weight applied in the left plot. Additionally, the right plot shows the result obtained from the default phase-space simulations. Only minor difference can be spotted, that are below the expected statistical sensitivity. In both cases the statistical uncertainties from the independent efficiency maps have been fully propagated to the q^2 averages and the obtained values are quoted in Tab. 6.6.

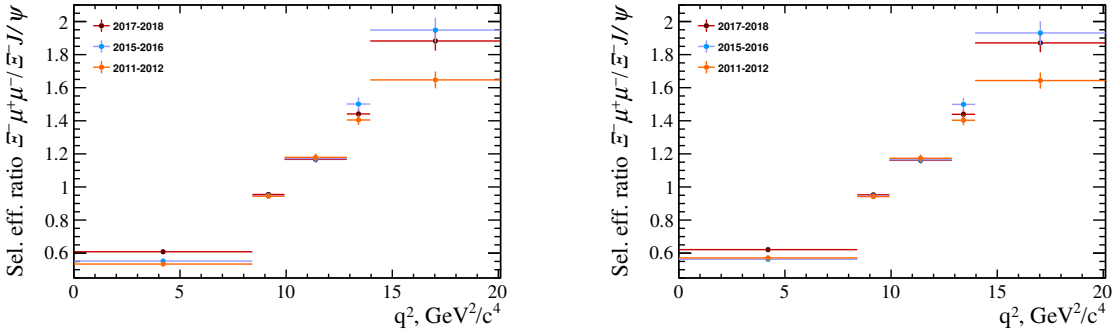


Figure 6.18: Efficiency ratio $\Xi_b^- \to \Xi^-\mu^+\mu^-/\Xi_b^- \to \Xi^- J/\psi$, averaged over the bins used in the search for the rare mode (the two narrow bins are vetoed and can be ignored). Left: our default values, with the q^2 weighting. Right: the values obtained from the PHSP model.

Table 6.6: Values of the ratio of total efficiencies between the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ and $\Xi_b^- \to \Xi^- J/\psi$ modes, averaged over the q^2 regions of interest, in 2011-2012, 2015-2016 and 2017-2018.

Region	Value 11-12	Value 15-16	Value 17-18
1	0.534	0.552	0.608
2	1.179	1.172	1.167
3	1.647	1.948	1.883

6.5 Branching fractions and systematic studies

This section is dedicated to the determination of the ratio of branching fractions and the related systematic uncertainties. First, the not-blinded ratio of the branching fractions of the charmonia decay modes will be discussed. Afterwards, the strategy towards the observation of the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay is discussed. This includes blinded fits, the strategy for the upper-limit setting, the estimation of the expected yields and a short discussion of the uncertainties.

6.5.1 Ratio $r^{\psi}_{BR} = \mathcal{B}(\Xi_b^- \to \Xi^-\psi(2S))/\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi)$

The ratio r^{ψ}_{BR}, defined in Eq. 6.2, can be expressed with the fitted yields N and efficiencies ε from Tab. 6.5, as follows:

$$r^{\psi}_{BR} = \frac{N(\Xi_b^- \to \Xi^-\psi(2S))}{N(\Xi_b^- \to \Xi^- J/\psi)} \times \frac{\varepsilon(\Xi_b^- \to \Xi^- J/\psi)}{\varepsilon(\Xi_b^- \to \Xi^-\psi(2S))} \times \frac{\mathcal{B}(J/\psi \to \mu^+\mu^-)}{\mathcal{B}(\psi(2S) \to \mu^+\mu^-)}. \tag{6.10}$$

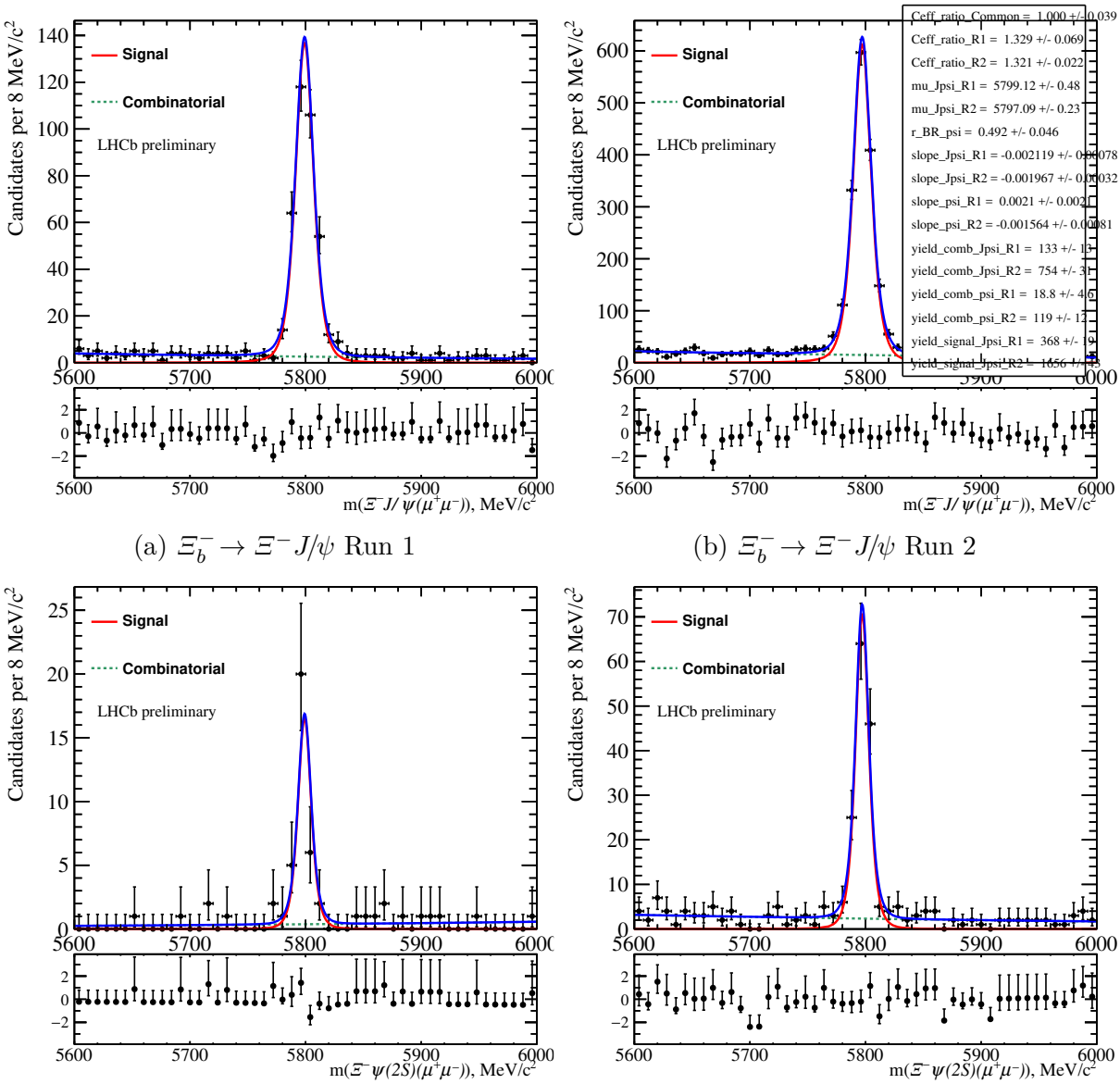
The latest world averages of the branching fractions of the charmonia decaying to two muons are $\mathcal{B}(\psi(2S) \to \mu^+\mu^-) = 0.79 \pm 0.02\%$ and $\mathcal{B}(J/\psi \to \mu^+\mu^-) = 5.96 \pm 0.03\%$ [33]. The final result on r^{ψ}_{BR} is obtained from a four-dimensional simultaneous fit to the $\Xi_b^- \to \Xi^- J/\psi$ and $\Xi_b^- \to \Xi^-\psi(2S)$ datasets in Run 1 and Run 2. The mass constraints on Ξ^-, Λ and charmonia are applied, and the tail parameters of the signal distributions are fixed to the values obtained from simulation. The means of the signal distributions are assumed to be shared between the $\Xi_b^- \to \Xi^- J/\psi$ and $\Xi_b^- \to \Xi^-\psi(2S)$ datasets, but can be different in Run 1 and Run 2. This difference is expected since the momentum scale has not been calibrated in this analysis. Additionally, the widths are fixed to 1.1 times the values from simulation, following studies with $\Xi_b^- \to \Xi^- J/\psi$ data. The background model consists only of combinatorial background, described in each dataset by an exponential shape with a floating slope parameter and a floating yield. In order to directly extract the result for r^{ψ}_{BR}, the yield of the $\Xi_b^- \to \Xi^-\psi(2S)$ is expressed via the yield of $\Xi_b^- \to \Xi^- J/\psi$ and r^{ψ}_{BR} and the relevant efficiencies:

$$N_{\Xi_b^- \to \Xi^-\psi(2S)} = N_{\Xi_b^- \to \Xi^- J/\psi} \times r^{\psi}_{BR} \times \frac{\varepsilon(\Xi_b^- \to \Xi^-\psi(2S))}{\varepsilon(\Xi_b^- \to \Xi^- J/\psi)} \times \frac{\mathcal{B}(\psi(2S) \to \mu^+\mu^-)}{\mathcal{B}(J/\psi \to \mu^+\mu^-)}. \tag{6.11}$$

Consequently, the fit has the following 13 free parameters: means of the Ξ_b mass peak in Run 1 and Run 2, r^{ψ}_{BR} (shared between Run 1 and Run 2), yield of $\Xi_b^- \to \Xi^- J/\psi$ signal in Run 1 and Run 2, yields and slopes of the combinatorial background in the $\Xi_b^- \to \Xi^- J/\psi$ and $\Xi_b^- \to \Xi^-\psi(2S)$ for Run 1 and Run 2 datasets. Additionally, the systematic uncertainties are included and the dedicated values of them are described in Sec. 6.5.2. There are two types of uncertainties: those which are considered uncorrelated between Run 2 and Run 1 datasets, and those which are common for the two (correlated). The uncorrelated uncertainties can be directly plugged into the fit as Gaussian constraints on the relevant ratios of efficiencies, with Run 1 and Run 2 having each one constraint. In order to reflect the correlated systematic uncertainties, a third constraint with a mean of unity and width of the size of the correlated systematic uncertainty is added to both the Run 1 and Run 2 ratios of efficiencies. This way, the fit fully accounts for all systematic uncertainties, and there is no need for further correction to the fit result. The final fit

result is shown in Fig. 6.19. And the extracted final result for the ratio of the charmonia decay mode branching fraction is

$$r_{BR}^{\psi} = 0.492 \pm 0.041(\text{stat}) \pm 0.021(\text{syst}). \tag{6.12}$$



(a) $\Xi_b^- \to \Xi^- J/\psi$ Run 1

(b) $\Xi_b^- \to \Xi^- J/\psi$ Run 2

(c) $\Xi_b^- \to \Xi^- \psi(2S)$ Run 1

(d) $\Xi_b^- \to \Xi^- \psi(2S)$ Run 2

Figure 6.19: Result of the simultaneous fit to the $\Xi_b^- \to \Xi^- J/\psi$ (top) and $\Xi_b^- \to \Xi^- \psi(2S)$ (bottom) datasets in Run 1 (left) and Run 2 (right), with systematic uncertainties included as Gaussian constraints. In the legend, the converged values of the three constraints are named as Ceff_ratio_R1 for Run 1, Ceff_ratio_R2 for Run 2 and Ceff_ratio_Common for the shared uncertainties between both Runs.

In order to test the stability of the result, the dataset has been split into different sub-datasets. The results are shown in App. E.6, no substantial deviations are found.

6.5.2 $\psi(2S)$: systematic uncertainties

The approaches to estimate the systematic uncertainties in this section, largely overlap with those discussed for the relative production fraction of the Ω_b^- baryon in Sec. 5.4.3. Therefore, this section is kept brief and only new sources of uncertainty are discussed in depth. It should be noted that usually two systematic uncertainties are reported, since they, especially those related to the calibration, may differ between the Run 1 and Run 2 datasets. For some other uncertainties like the external input for branching fractions, that are independent of the dataset, only one value is reported. Several possible different sources of systematic uncertainty are studied in detail and summarised in Table 6.7. The uncorrelated (per Run period) and correlated (relevant for both periods) contributions are determined to convert them to the final total uncertainty for Run 1 and Run 2[8].

Table 6.7: Relative systematic uncertainties on r_{BR}^{ψ}, %.

Type	Run 1	Run 2
Size of simulated samples	0.6	0.4
Calibration of simulation	5.4	1.6
Lifetimes of Ξ_b^- baryon	0.1	0.1
Fit model	3.1	
Material interaction	0.8	
External input (BR)	2.2	
Total uncorrelated	5.4	1.7
Total correlated	4.1	
Total	6.9	4.4

The relative uncertainty on the efficiency ratio due to the size of the available simulation samples is found to be 0.6% (0.4%). The systematic uncertainty due to the small amount of available statistics when producing the background-subtracted data samples is propagated using a bootstrapping technique [204]. The relative systematic uncertainty is found to be 0.2% (0.2%). Additionally, alternative binning schemes are tested for the calibration weights on the simulation and found to lead to shifts of 0.1% (0.2%) in the final result. In addition, the kinematics of the Ξ_b^- baryon is alternatively calibrated with a two-dimensional weight map, $p_{\mathrm{T}}(\Xi_b^-) : \eta(\Xi_b^-)$, rather than the default calibration of $p_{\mathrm{T}}(\Xi_b^-)$ alone. The resulting uncertainty is 0.1% (0.1%). In order to access the relative systematic uncertainty due to the porting of the correction weights from Run 2 to the Run 1 samples, alternative weights for Ξ_b^- kinematics and DTF χ^2 correction are determined using the Run 1 samples. The resulting shift on the result is found to be 0.3%. An important calibration step is the reweighting of the helicity angles to those of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ simulation samples and the momentum asymmetries of the Ξ^- and Λ decay products. The absences of those weights leads to an systematic uncertainty of 0.6% (0.9%).
Similarly, the L0 calibration weights have been removed and the resulting shift leads to the dedicated systematic uncertainty of 0.2% (0.1%). A dedicated cross-check showed, that any correction for minimal mismodelling of the HLT trigger response is expected to cancel in the ratio. Therefore, no dedicated uncertainty is assigned. The limitations in

[8]In the following, if two numbers are given, the one in brackets corresponds to the Run 2 number.

the standard calibration procedure used for the calibration of the ProbNNmu variables are accounted for by a relative uncertainty of 0.11% for both Run periods. The relative uncertainty when reweighting the event multiplicity and resampling the ProbNNmu without the scaling factor is 0.7% (0.6%). As previously mentioned, only the additional PID cuts applied offline for the two muon candidates have been calibrated, since the simulation samples have been filtered by the stripping selection. With the same procedure as described in Sec. 5.4.3, the effect of the calibration of the muon ID cuts in the stripping is determined using Run 2. The corresponding systematic uncertainty is found to be 0.55%. Due to the absence of similar unfiltered samples for Run 1, the same uncertainty is applied for both Run periods. A further source of uncertainty originates from imprecise modelling of the downstream tracking and vertexing efficiencies in simulation. While the effect is expected to be small in Run 2, the effect is expected to be large in Run 1 due to the significant mismodelling of the track types. A dedicated systematic uncertainty of 5% (1%) is assigned. Beside the calibration of the simulation, the efficiency ratio can also be affected by the truthmatching strategy. Including the background category 60 or removing it, leads to an uncertainty of 0.1% (0.1%). Similarly, the ghost probability of a track can be mismodelled and the related systematic uncertainty is determined to be 1.6% (0.3%). Combining all these individual uncertainties related to the calibration and efficiencies, the overall relative uncertainty is 5.4% for Run 1 and 1.6% for Run 2.
Since the simulation only uses the nominal value of the beauty baryon lifetimes [33], the final result can be biased due to the imprecise knowledge of their values. The obtained relative uncertainty on the ratio r^{ψ}_{BR} is found to be 0.1% (0.1%).
Several sources of systematic uncertainty related to the fit model are considered. The mean of the Ξ_b^- mass is allowed to float in the simultaneous fit to data. The related systematic uncertainty is addressed by fixing the mean to the world average of the Ξ_b^- mass [33] and determined to be 0.2%. Similarly, the requirement that the mean is shared between all four planes (Run 1 and 2, $\Xi_b^- \to \Xi^- J/\psi$ and $\Xi_b^- \to \Xi^- \psi(2S)$) is removed and the resulting uncertainty is 0.2%. Combining both uncertainties leads to a total systematic uncertainty of 0.3% related to the Ξ_b^- mass treatment. When removing all hyperon and charmonia mass constraints, the obtained result of the fit is shifted by 2.6%. Letting the signal width float in the mass fits, leads to a systematic uncertainty of 0.3%. The fit bias due to the choice of the signal model has been tested with 10 000 pseudoexperiments generated with an alternative signal model based on one double-sided Crystal-Ball function, and fitting with the nominal model. The obtained uncertainty is 0.65%. Similarly, 10 000 pseudoexperiments are produced modelling the combinatorial background with a second-order Chebyshev polynomial and fitting with the exponential. The resulting uncertainty is 1.01%. The last important check is related to possible contaminations from the over-reconstructed $\Lambda_b^0 \to \Lambda J/\psi$ ($\Lambda_b^0 \to \Lambda \psi(2S)$) decay combined with a random charged pion. While the effect due to the contamination is expected to nearly cancel in the ratio[9], 10 000 pseudo-experiments are generated with a dedicated Λ_b^0 background component and fitted with the default fit model. The resulting systematic uncertainty is determined to be 1.2%, leading to a total of systematic uncertainty of 3.1% due to the fit model choices.

[9]The ratio of partial width between any Ξ_b^- and Λ_b^0 decay is always 3/2 and cancels, however, the kinematics and resolution slightly differ between the $\psi(2S)$ and J/ψ decay modes, which could result in a not full cancellation of the efficicency ratios.

As discussed above, the generator-level datasets can be accessed, when fully matching the subsequent hyperon decays or not. The difference in the obtained efficiency arises due to material interactions, which are not considered when the subsequent decays after the Ξ^- are not considered. While the bias due to hadronic interactions should largely cancel in the ratio of the $\psi(2S)$ and J/ψ modes, comparing the ratio of efficiencies calculated with the two methods without applying the decay model correction[10] enables to obtain the related systematic uncertainty of 0.8%.
The uncertainties from the external input of the branching fractions $\mathcal{B}(\psi(2S) \to \mu^+\mu^-) = 0.79 \pm 0.02\%$ and $\mathcal{B}(J/\psi \to \mu^+\mu^-) = 5.96 \pm 0.03\%$ are propagated to the final result by summing these two relative uncertainties in quadrature. The resulting systematic uncertainty is 2.2%.
Combining all these individual systematic uncertainties and correctly accounting for correlations leads to a total uncertainty of 6.9% for the Run 1 and 4.4% for Run 2.

6.5.3 Towards the observation of the rare mode

In order to observe the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay, a simultaneous fit to the $\Xi_b^- \to \Xi^- J/\psi(\to \mu^+\mu^-)$ and $\Xi_b^- \to \Xi^-\mu^+\mu^-$ datasets is performed, where the only difference in the selection of both decay modes is the respective q^2 requirement. The BDT cut value applied to the normalisation channel is the same as that used in the regions 1 and 2 of the rare mode, this leads to smaller but cleaner $\Xi_b^- \to \Xi^- J/\psi(\to \mu^+\mu^-)$ signal than in the fits used for measurement of r^{ψ}_{BR}.
The signal models used for this fit are explained in Sec. 6.4 for the rare and $\Xi_b^- \to \Xi^- J/\psi$ modes, where no J/ψ mass constraint is applied to the latter. The tail parameters of the signal PDFs are all fixed to the values from simulation and the signal resolution is additionally fixed after rescaling the value obtained from simulation for both channels by the a factor 1.1. The mean of the Ξ_b^- mass peak is shared between the rare and J/ψ datasets as well as later for each data taking period, and is kept floating in the fit to data. The combinatorial background is modelled by an exponential in each plane of the simultaneous fit and the slope as well as the yield are kept floating. The resulting blind fit is shown in Fig. 6.20 using the fully merged dataset. The summary of resulting values of non-blinded fit parameters is provided in Tab. 6.8.

Table 6.8: Non-blinded fit parameters of the blinded fit

Parameter	Value
comb. yield (J/ψ)	183.0 ± 17.4
comb. slope (J/ψ)	$(-3.3 \pm 0.7) * 10^{-3}$
comb. yield (rare)	106.9 ± 11.9
comb. slope (rare)	$(-0.5 \pm 0.9) * 10^{-3}$
peak mean	5798.4 ± 0.5
sig. yield (J/ψ)	1373.1 ± 38.6

The same simultaneous fit setup will be used to confirm if the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay can be observed. Once the fit is unblinded, the hypothesis that the observed $\Xi_b^- \to \Xi^-\mu^+\mu^-$ signal peak is originating from a statistical fluctuation has to be tested. If the signal

[10] The effect is covered by the systematic uncertainty related to the helicity corrections.

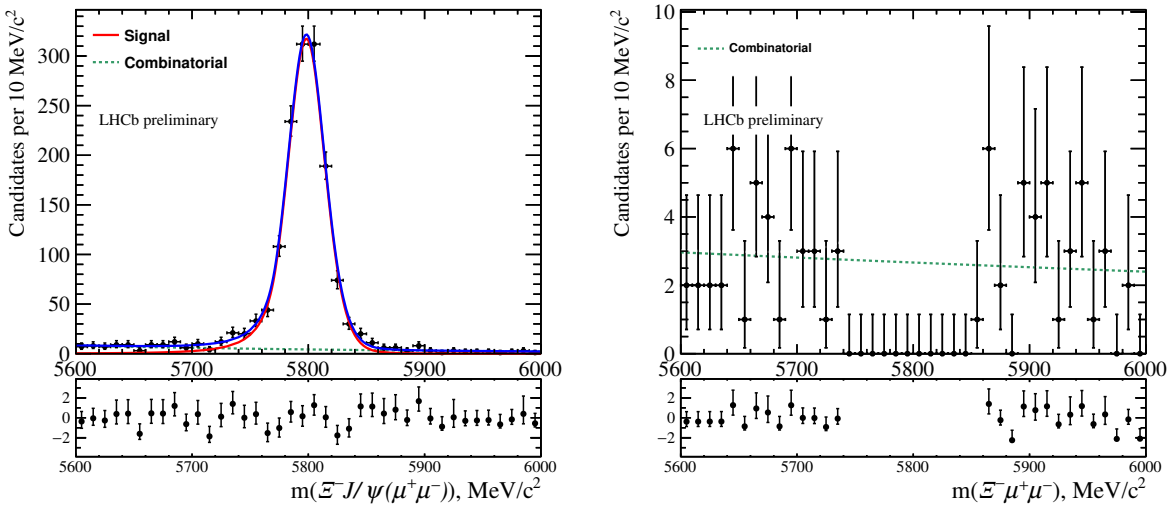


Figure 6.20: Simultaneous fit to data: left is normalisation $\Xi_b^- \to \Xi^- J/\psi$ channel; right is blinded $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ data.

hypothesis is confirmed, a simultaneous fit is performed to six projections: rare and normalisation modes in 2011-12, 2015-16, 2017-18. This choice is made due to the differences of the efficiency trends for each of these three datasets. The fit is performed independently for all three of the q^2 bins, enabling to measure the differential branching fraction, or set an upper limit on it, in each q^2 bin. In these fits, the efficiencies from Tab. 6.6 are used to convert the signal yield into the branching fraction into the fit, constrained in a Gaussian manner according to respective systematic uncertainties. This means, the yield of the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay mode is expressed via the ratio of branching fractions r_{BR} to directly extract the value of differential branching fraction in each bin.

$$r_{BR} = \frac{N(\Xi_b^- \to \Xi^- \mu^+ \mu^-)}{N(\Xi_b^- \to \Xi^- J/\psi)} \times \frac{\varepsilon(\Xi_b^- \to \Xi^- J/\psi)}{\varepsilon(\Xi_b^- \to \Xi^- \mu^+ \mu^-)} \times \mathcal{B}(J/\psi \to \mu^+ \mu^-) \tag{6.13}$$

Expected signal yield and significance

The predicted values of r_{BR} in the three q^2 regions in Sec. 6.2.2 can be combined with the ratios of the efficiencies in Tab. 6.6 to estimate the expected SM signal yields based on Eq. 6.13. It should be noted, that these predicted values of r_{BR} are based on the assumption of the $\Xi_b^- \to \Xi^- J/\psi$ branching fraction, which has large uncertainties as discussed in Sec. 6.2.2. Therefore, uncertainties are not quoted here due to their large size, where the dominant one comes from the SM-predicted rate to r_{BR}. Tab. 6.9 summarises all calculated yield predictions in the three q^2 regions, combined and split between the three data taking periods, as well as all three q^2 bins combined per data-taking period and lastly combining everything. According to the SM prediction, about 150 signal events should be found in the whole dataset.

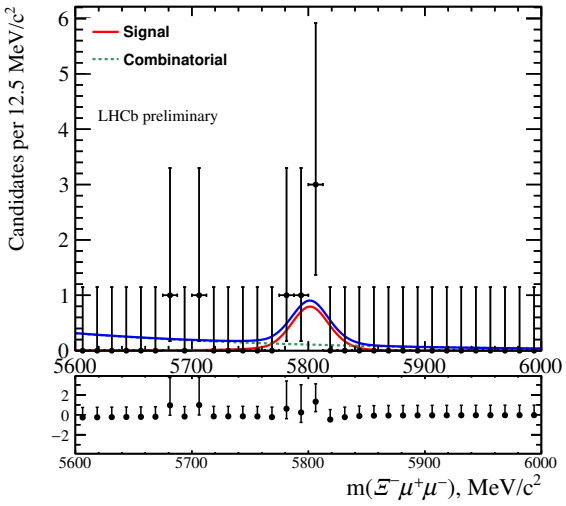
These values can be used to study the fit setup with pseudo-experiments, where the signal yield is set to the predicted values in Tab. 6.9 and the systematic uncertainties discussed in Sec. 6.5.4 are included. As an example Fig. 6.21 shows the projections of the fit for one of the 5 000 pseudoexperiments. The resulting pulls of r_{BR} and of the $\Xi_b^- \to \Xi^- J/\psi$ yield are shown in Fig. 6.22 for region 1. It should be noted, that the regions 1 and 2 have

Table 6.9: Expected rare signal yields under SM assumption based on the calculation of the $\Xi_b^- \to \Xi^- J/\psi$ branching fraction with large uncertainties.

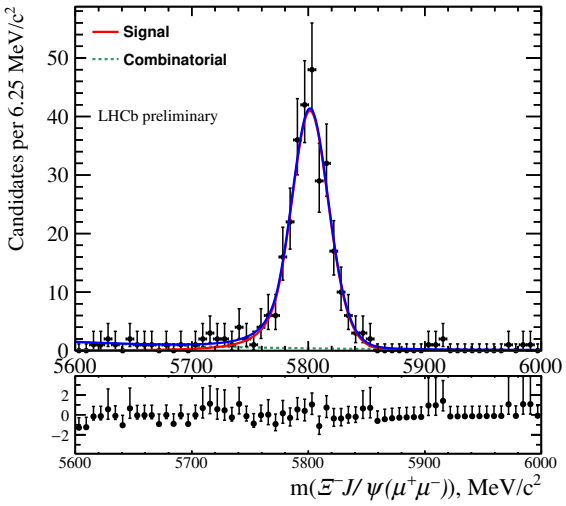
years	r_{BR}	norm. yield	eff. ratio	rare yield
		Region 1		
17-18		746	0.608	7.6
15-16	0.0010	349	0.552	3.2
11-12		279	0.534	2.5
all				15.2
		Region 2		
17-18		746	1.167	10.2
15-16	0.0007	350	1.172	4.8
11-12		278	1.179	3.9
all				20.4
		Region 3		
17-18		750	1.882	66.4
15-16	0.0028	351	1.948	32.2
11-12		272	1.647	21.1
all				125
		Total		
17-18				84.2
15-16				40.2
11-12				27.4
All				**151.8**

very low predicted signal yields and due to the tight BDT cut also a very low background level. Small yields can lead to biases seen in the pseudo-experiment studies, that need to be evaluated properly after the unblinding (with the actually observed yields) to ensure no bias is introduced on the r_{BR} measurement. Similarly, the low statistics in the BDT cut optimisation procedure can introduce an additional bias, that should be studied as well.

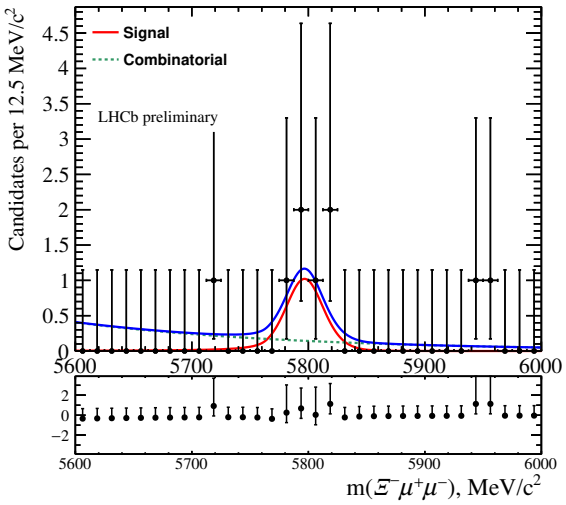
As a last step, the predicted SM signal yields can be used to determine the expected statistical significance. Therefore, two fits are performed: one with the signal yield kept floating and one set to zero, where all sub-datasets are merged and only the three q^2 regions are distinguished. For each of the two fits the log-likelihoods are recorded and the statistical significance is defined as $\sqrt{-2\Delta LL}$, where ΔLL is the difference of likelihoods of the two hypotheses. In order to estimate the expected significance while taking into account statistical behavior of signal and background contributions, 50 pseudoexperiments have been generated according to the predicted SM yields and varied following Poisson statistics and fitted both with the default fit model and the zero signal hypotheses (fixed signal yield). It should be noted that the SM estimates ignore the systematic uncertainty on the $\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi)$ which can easily shift the expected signal yield by up to 40%. The theoretical uncertainty itself is also ignored here. The obtained means of the $\sqrt{-2\Delta LL}$ distributions for each of the three q^2 regions are about 3.7, 5.4 and 16 standard deviations, respectively. Therefore, if the SM predictions are correct, it might be possible to find evidence for the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay mode in the first q^2 region and observe the signal in the other two regions, where the region 3, with $15 < q^2 < 20\,\mathrm{GeV}^2/c^4$, is most likely to be observed due to its large statistical significance. The corresponding plots can be found in Fig. E.19 in App. E.7.



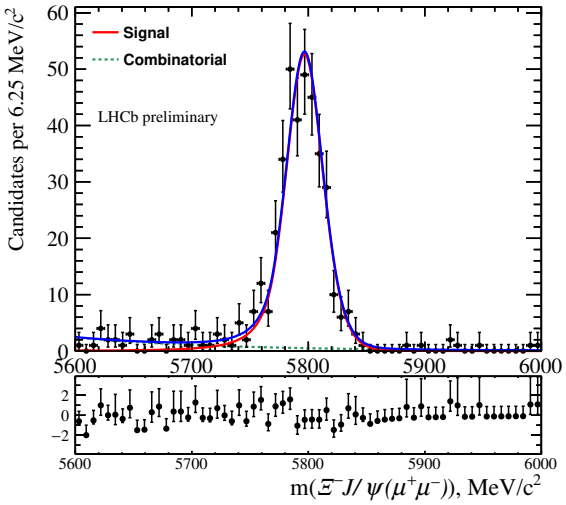
(a) 2011-12 rare



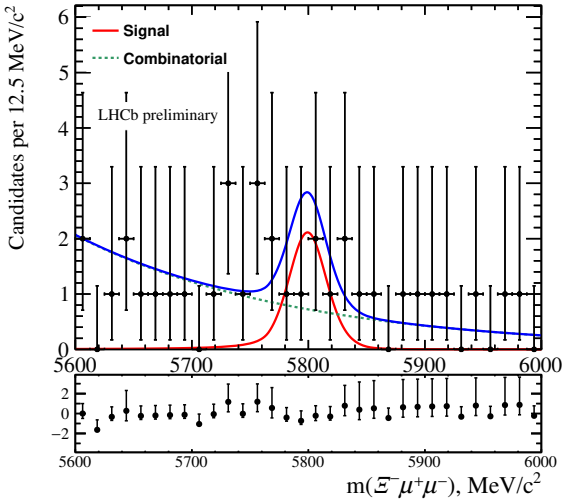
(b) 2011-12 J/ψ



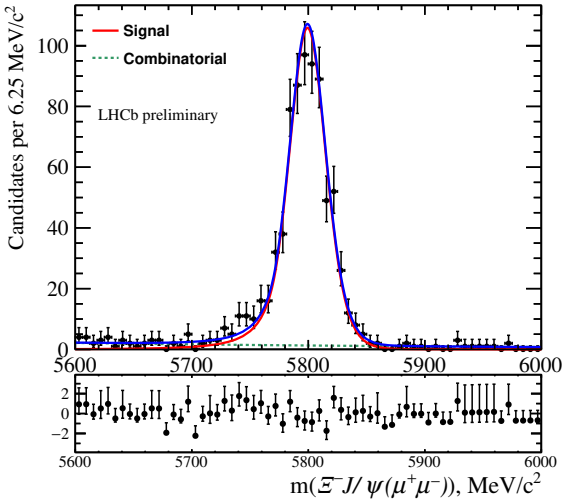
(c) 2015-16 rare



(d) 2015-16 J/ψ

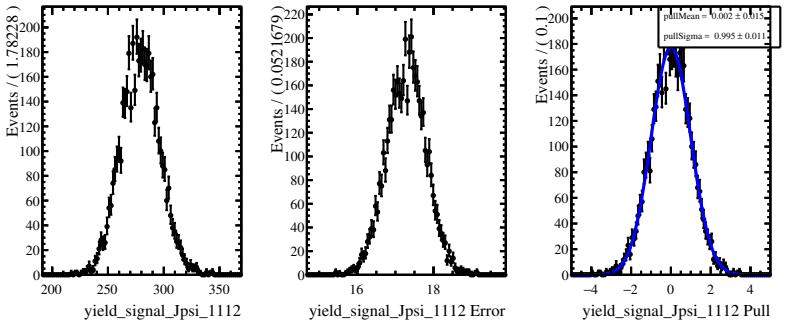


(e) 2017-18 rare

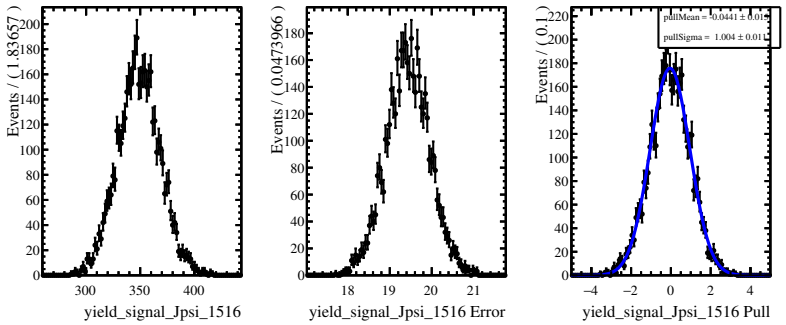


(f) 2017-18 J/ψ

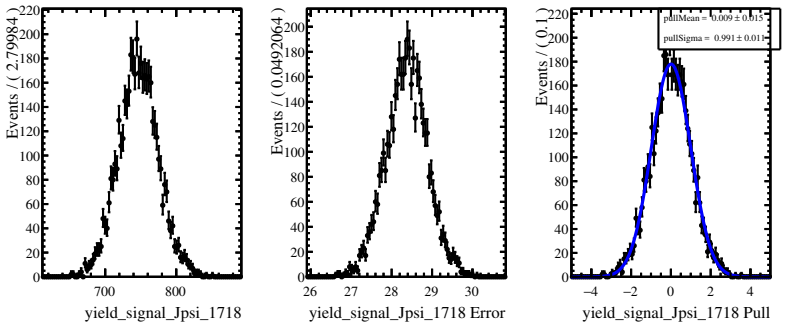
Figure 6.21: Results of SM-like pseudo-experiments in the q^2 region of $m(\mu^+\mu^-) < 2900\,\mathrm{MeV}/c^2$.



(a) $\Xi_b^- \to \Xi^- J/\psi$ yield in 2011-12: pull mean= -0.02 ± 0.02, pull width = 1.00 ± 0.01



(b) $\Xi_b^- \to \Xi^- J/\psi$ yield in 2015-16: pull mean= -0.04 ± 0.05, pull width = 1.00 ± 0.01



(c) $\Xi_b^- \to \Xi^- J/\psi$ yield in 2017-18: pull mean= 0.01 ± 0.02, pull width = 0.99 ± 0.01

Figure 6.22: Pulls on free fit parameters from 5 000 SM-like pseudo-experiments in the q^2 of $m(\mu^+\mu^-) < 2900\,\mathrm{MeV}/c^2$.

Limit setting

Finally, the procedure has been put into place to estimate the expected upper limits on the branching fraction, in case no signal is observed. As it has been shown above, there is a chance that the signal significance will not be sufficient to reach the threshold of an evidence ($> 3\sigma$) in at least one of the q^2 regions. In that case, an upper limit on the r_{BR} will be set in that region. Due to the large uncertainties on the SM predicted signal yields, the projected statistical significance can also be overestimated, therefore the limit setting framework is put into place for all three q^2 regions. Since the data is currently blinded, the expected upper limits on r_{BR} under the background-only hypothesis are set at 90 and 95% CL in each region of q^2. For the limit setting, the GammaCombo framework [25, 26]

is used with the confidence level CL_s method [219] applying a one-sided test statistic, where the p-values are determined using pseudo-experiments. The obtained expected upper limits in the three regions of q^2 are presented in Tab. 6.10. The values of the systematic uncertainties used for these results are presented in Sec. 6.5.4. As expected, including the systematic uncertainties in the setup returns a slightly worse expected upper limit.

Table 6.10: Values of the expected upper limits on r_{BR} in the q^2 regions of interest, with and without taking into account the systematic uncertainties.

Region	No syst (90% CL)	w/syst (90% CL)	No syst (95% CL)	w/syst (95% CL)
1	$< 0.46 \times 10^{-3}$	$< 0.48 \times 10^{-3}$	$< 0.56 \times 10^{-3}$	$< 0.59 \times 10^{-3}$
2	$< 0.16 \times 10^{-3}$	$< 0.17 \times 10^{-3}$	$< 0.20 \times 10^{-3}$	$< 0.21 \times 10^{-3}$
3	$< 0.15 \times 10^{-3}$	$< 0.16 \times 10^{-3}$	$< 0.19 \times 10^{-3}$	$< 0.20 \times 10^{-3}$

In order to compare the expected upper limits to the SM predictions, the expected branching fraction $\mathcal{B}(\Xi_b^- \to \Xi^- \mu^+ \mu^-)$, discussed in Sec. 6.2.2, is used to convert the theoretical predictions into predictions on r_{BR}. The uncertainty on $\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi)$ is ignored in this computation at the moment. The result is shown in Fig. 6.23 and is compatible with expectations on signal significance from the previous section.

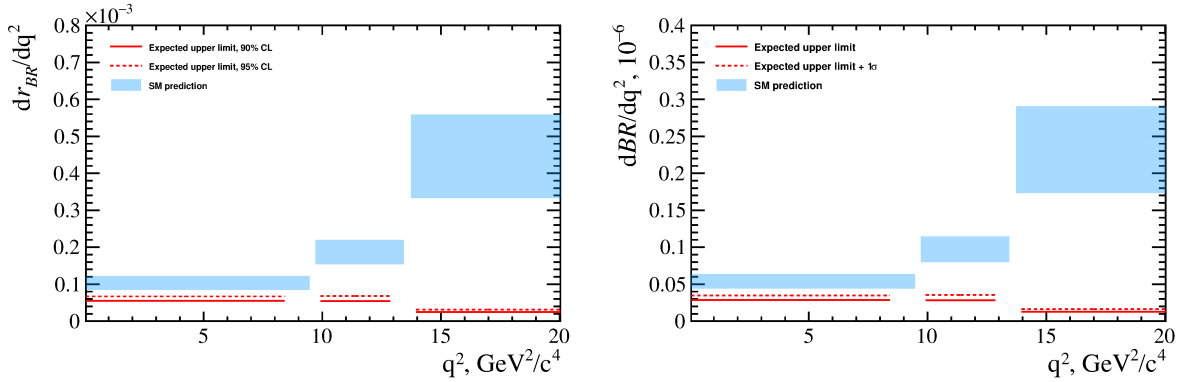


Figure 6.23: Left: Expected upper limit at 90% (red solid line) and 95% (red dashed line) CL on dr_{BR}/dq^2, compared to SM predictions from Ref. [27] (blue boxes). Right: the same but multiplied by our estimate of $\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi)$, yielding $d\mathcal{B}/dq^2$. Only statistical uncertainties are taken into account.

6.5.4 Rare mode systematic uncertainties

The estimates of the systematic uncertainties in this section are preliminary and will possibly be updated in the future. The sources of systematic uncertainty are mostly the same as in the previous section for the $\psi(2S)$ channel in Sec. 6.5.2. Thus, only the averaged uncertainties are reported and not the full details are discussed. But additional sources of uncertainty arise due to the little knowledge about the properties of the expected signal. Only the systematic uncertainties related to the decay model and the size of the simulated samples are expected to differ significantly between the three q^2 regions. Therefore, for

all other sources only one value is reported as they are expected to have roughly the same impact on all the three q^2 regions, which is furthermore small compared to the expected statistical uncertainty. The uncertainties are always estimated for the entire Run 1 and Run 2 dataset. A summary of the values of each source of uncertainty is provided in Table 6.11.

Table 6.11: Systematic uncertainties on r_{BR} in %. A value per q^2 region is given for uncorrelated uncertainties, while a single value is given for the common ones.

Type	Region 1	Region 2	Region 3
Decay model	3.7	0.7	1.7
Size of simulated samples	0.4	1.2	1.9
Calibration of simulation		2.1	
Lifetime of Ξ_b^-		0.3	
Fit model		1.7	
Material interaction		0.8	
External input (BR)		0.6	
Total	4.7	3.2	3.9

As it was discussed in Sec. 6.4.1, the default efficiency-ratio values in the three regions of q^2 are determined after the q^2-weighting of the rare-mode simulation. Since this weighting is not perfect, and the exact model cannot be known until the signal is experimentally observed, the obtained result is compared to the one with the efficiency-ratios obtained with PHSP simulation, described in Fig. 6.18. The resulting relative difference in efficiency, measured with 2017-18 simulation, is 3.7% in region 1, 0.7% in region 2 and 1.7% in region 3. The smallness of the shift in the region 2 is due to narrowness of this region, while in region 3, it is caused by the fact that the relative efficiency is almost flat there, consequently the q^2 weights have little impact.
Since the statistics of simulated samples affects the uncertainty on signal efficiencies, the related uncertainty is propagated to the final result. The resulting systematic uncertainty is 0.4% in region below $m(\mu^+\mu^-) < 2900\,\text{MeV}/c^2$, 1.2% in region of $3150 < m(\mu^+\mu^-) < 3586\ \text{MeV}/c^2$ and 1.9% in region of $15 < q^2 < 20\,\text{GeV}^2/c^4$. This uncertainty is larger compared to the computation of r_{BR}^{ψ}, since the statistically smaller, unfiltered samples used to evaluate the q^2 trend of the efficiencies are utilised.
The total systematic uncertainty related to the calibration of the simulation is found to be 2.1%. The used approaches are identical to the ones for r_{BR}^{ψ}. However, it should be noted, that the approach of accessing the L0 uncertainty by just removing the corresponding weights is expected to be not portable to the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ samples, since the effect of mis-calibration will be larger at low q^2. Therefore, a dedicated approach should be used to estimate the systematic uncertainty from the calibration through the TISTOS method. In order to do that, as a first step, the "TIS" tag definition is modified. Instead of a global TIS, a TOS selection is required on the second muon from the decay and the efficiency ratio is recalculated. Additionally, the effect of the binning scheme for the calibration weight maps should be evaluated. This can be done in two steps: first, instead of interpolating between the different bins, the binned weight can be applied to each event and, second, the binning scheme can be modified. The final systematic uncertainty can then be accessed by summing the relative size of observed deviations in each such test in quadrature. The total uncertainty due to the fit model choice is 1.7%, however there

are several differences. While the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay mode is still blind, the shift in the $\Xi_b^- \to \Xi^- J/\psi$ yields with the J/ψ-constrained vs unconstrained fits is determined, leading to an uncertainty of at most 0.5%. Similarly, so far no approach was performed to access uncertainties related to the sharing and fixing of the Ξ_b^- mass. The effect of the mass-resolution constraint is expected to be the largest in the broadest q^2 region, region 1. The systematic uncertainty obtained through pseudoexperiments is 1.1%, but needs to be re-evaluated after the unblinding. As an alternative model for the combinatorial background, a first-order Chebyshev polynomial is chosen due to the statistical limitations. The uncertainty due to the alternative fit model with the Λ_b^0 background is expected to be not larger as the r_{BR}^{ψ} one, which is therefore taken as current estimate. The effect of the Ξ_b^- lifetime correction is 0.3%. Finally, the external uncertainty on the $J/\psi \to \mu^+\mu^-$ branching fraction is propagated to the final result and the obtained uncertainty is 0.6%. Combining all these individual systematic uncertainties and correctly accounting for correlations leads to a total uncertainty of 4.7% in region 1, 3.2% in region 2 and 3.9% in region 3.
Finally, to obtain a result for the absolute branching fractions via the calculated value $\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi) = (5.2 \pm 1.6 \pm 1.6 \pm 0.1) \times 10^{-4}$, (Eq. 6.8),where the first uncertainty comes from $\mathcal{B}(\Lambda_b^0 \to \Lambda J/\psi)$, the second – from $SU(3)$ breaking and the third from the lifetime uncertainties, the uncertainties need to be propagated to the final results.

6.6 Summary

In summary, this chapter presented the observation of the $\Xi_b^- \to \Xi^-\psi(2S)$ decay mode for the first time at the LHCb experiment using the full Run 1 and Run 2 dataset corresponding to $9\,\text{fb}^{-1}$. The branching fraction of the decay mode with respect to the $\Xi_b^- \to \Xi^- J/\psi$ one has been evaluated and found to be

$$r_{BR}^{\psi} = 0.492 \pm 0.041(\text{stat}) \pm 0.021(\text{syst}).$$

When the result is compared to the first observation performed by the CMS collaboration [18]

$$r_{BR}^{\psi}(CMS) = 0.84 \pm 0.21\,(\text{stat}) \pm 0.10\,(\text{syst}) \pm 0.02(\mathcal{B}),$$

they differ by 1.5 standard deviation, but the LHCb result presented in this thesis is a factor 5 more precise. Additionally, the ratio r_{BR}^{ψ} is very similar to the measurements [23, 24] of the Λ_b^0 ratio (6.3) supporting the observed deviation from the quark model predictions [22], which are about $r_{BR}^{\psi}(\text{Theo}) \sim 0.84$.
Additionally, for the current status of the first search for the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay mode pre-unblinding is presented. The SM predictions of the ratio r_{BR} in the three signal regions have been used to estimate the expected signal yields, taking into account the q^2-dependent efficiencies. The fit model has been evaluated with blinded fits to data as well as fits to pseudoexperiments where the signal is injected according to the predicted SM yields. Additionally, the expected statistical significance of the SM-like signal has been evaluated and the limit setting procedure has been put into place. The studies support that the signal will be observed with the full Run 1 and Run 2 dataset of the LHCb experiment.

Chapter 7

First studies of unobserved Ξ_b^0 decays

The next analysis covered as part of this thesis is the first search for the tree-level decays of $\Xi_b^0 \to \Xi^0\psi$ with $\psi \in [J/\psi, \psi(2S)]$. As already briefly mentioned in Sec. 1.4.1, these decay modes have the same final state particles as the rare decay mode $\Xi_b^0 \to \Xi^0\mu^+\mu^-$ and an identical topology. Therefore, they enable to study the challenges arising due to the unknown, but expected to be small, production fraction of Ξ_b^0 baryons, and the presence of neutral particles in the decay chain.
The following sections describe the analysis strategy and the challenges that have to be overcome to observe the $\Xi_b^0 \to \Xi^0\psi$ decays and pave the way towards the observation of the rare decay mode $\Xi_b^0 \to \Xi^0\mu^+\mu^-$. Dominant background sources will be discussed and the resulting selection. In addition, the use of the DecayTreeFitter and required adaptions will be covered in order to increase the number of observed signal candidates. Lastly, the open challenges of this analysis will be discussed. This first study was performed as part of M.Sc. Leandra Moesers master thesis [220] and has been supervised by Dr. Vitalii Lisovskyi and myself.

7.1 Analysis strategy

Currently, predictions for b-baryons decays other than of Λ_b^0 decays rely on the covariant quark model as no QCD predictions of the form factors are available. Consequently, it is important to ensure that the predictions are confirmed by experimental data. One possible way to test the covariant quark model is by determining the ratio of branching fractions for $\psi(2S)$ over J/ψ decay modes for the same b-hadron. While the ratio has been proven to be in agreement with the prediction for B meson decays [33], the ATLAS collaboration measured the ratio $\mathcal{B}(\Lambda_b^0 \to \Lambda\psi(2S))/\mathcal{B}(\Lambda_b^0 \to \Lambda J/\psi)$ [23] to differ by 2.8σ from the covariant quark model predictions [221, 222]. The LHCb measurement [24] of the same quantity is compatible within one standard deviation with the measurement from the ATLAS collaboration [23], but has a better precision and therefore confirms the discrepancy with the covariant model predictions[1].

[1]Even when including different form factor predictions [223, 224] these discrepancies remain as the resulting values 0.65 to 1.14 are above the measured central values of 0.51 [23, 24].

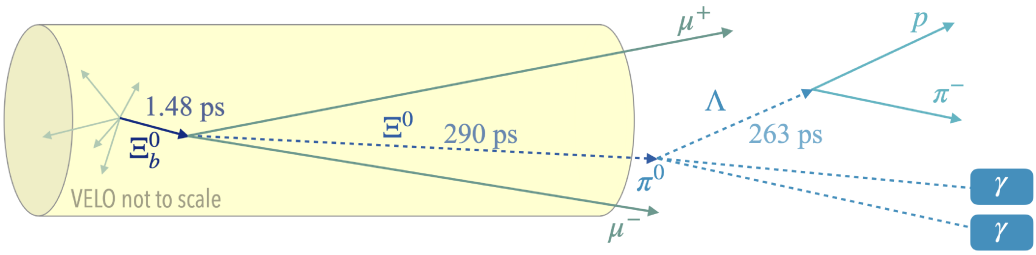


Figure 7.1: Schematic representation of the $\Xi_b^0 \to \Xi^0\psi$ decay chain. Neutral particles are indicated by dashed lines, while the tracks of charged particles are displayed with solid lines. The boxes around the photons are indicating calorimeter clusters.

Consequently, it is important to investigate if the same discrepancy is visible for other b-baryons. Therefore, under the assumption that both charmonia decay modes of the Ξ_b^0 will be observed, the ratio of the branching fractions of them

$$R = \frac{\mathcal{B}(\Xi_b^0 \to \Xi^0\psi(2S))}{\mathcal{B}(\Xi_b^0 \to \Xi^0 J/\psi)} \tag{7.1}$$

is the targeted result of the presented analysis.
Since the J/ψ and $\psi(2S)$ decay modes only differ by their respective charmonia resonances, not only the Ξ_b^0 production fraction $f_{\Xi_b^0}$ cancels in the ratio, but also the branching fractions and reconstruction efficiencies of the subsequent cascade decays, $\Xi^0 \to \Lambda\pi^0$, $\Lambda \to p\pi^-$ and $\pi^0 \to \gamma\gamma$, as well as systematic uncertainties that are not related to the specific charmonia. Consequently, the ratio is, regardless of all experimental challenges, a rather clean observable.
The analysis will be performed using datasets collected between 2016 and 2018 and corresponding to an integrated luminosity of $5.4\,\text{fb}^{-1}$. This choice was enforced by the dedicated stripping line selecting these decay modes being only available for these three years.
The following subsection elaborates the specific experimental challenges that need to be overcome in order to provide a first observation of both decay modes.

7.1.1 Challenges

Choosing the charmonia decay modes to test the covariant quark model predictions for Ξ_b^0 decays has several advantages. When charmonia are reconstructed through muonic decay modes $\psi \to \mu^+\mu^-$, they benefit from very high trigger efficiencies and can be used to reconstruct the Ξ_b^0 decay vertex. This enables to reduce the background due to displacement cuts between the proton-proton interaction point and the Ξ_b^0 decay vertex. In addition, applying charmonia mass constraints to the targeted invariant mass $m(\Xi^0\psi(\to \mu^+\mu^-))$ helps to improve the mass resolutions significantly due to their high masses and reconstruction efficiency. This is particularly important as the rest of the decay chain includes several neutral particles smearing the mass resolution. Fig. 7.1 displays a schematic representation of the full decay chain.
The Ξ_b^0 decays to the charmonia and a neutral Ξ^0 baryon, which then decays to a neutral Λ baryon and a neutral π^0 meson. Since both decay products of the Ξ^0 have no electrical

charge they can not be used directly to reconstruct the Ξ^0 decay vertex. The Λ baryon then finally decays into two charged particles, a proton and a π^- meson. These enable to determine the Λ decay vertex and improve the momentum information of the Λ baryon. However, since both the Ξ^0 and Λ are long-lived hyperons, the Λ is expected to decay in 99.5% of the time (according to uncalibrated simulation samples) after the VELO. Therefore, the Λ candidates are reconstructed from downstream tracks and consequently have a worse momentum resolution. However, similarly to the charged Ξ_b^- decay modes, it is expected that the long lifetimes of the hyperons help to suppress the backgrounds, as no similar topology exists for mesonic decay modes. But even if the precise knowledge of the Λ momentum helps to improve the knowledge on that of the Ξ^0 baryon, the latter suffers significantly from the challenging π^0 reconstruction. Since the π^0 decays into a pair of photons, no charged tracks are left in the detector to help with the reconstruction of the Ξ^0 decay vertex.
In addition, the reconstruction of π^0 mesons brings its own set of challenges. For this analysis, the π^0 momentum is expected to be relatively low due to the kinematics of the two-body Ξ^0 decay, given the much higher Λ mass. As a consequence, the π^0 candidates are reconstructed from two resolved photons. As discussed in Sec. 2.2.4, the estimated momentum of the π^0 is biased as it can only be reconstructed from the photon cluster positions and an extrapolation to the PV, rather than to the correct Ξ^0 decay position. In addition, there are many other sources of photons produced in pp collisions. As a result, the occupancy of the calorimeter is high, which can further bias the π^0 reconstruction. The two dominant effects are assigned wrong photons or assigning a part of a cluster of a π^0 photon to a different ECAL cluster, leading to an over- or underestimation of the energy assigned to the π^0 candidate. Therefore, the requirements for the π^0 candidate need to be rather loose to not lose signal. At the same time, the π^0 is the lightest meson and, therefore, is abundantly produced in proton-proton collisions and physics processes unrelated to the targeted result. Therefore, the multiplicity of possible π^0 candidates in the same event drives the amount of background in this analysis.
There are three technical challenges that have to be addressed to ensure a correct and efficient reconstruction of the targeted decay modes $\Xi_b^0 \to \Xi^0\psi$.
The first two are related to available simulation samples. Since no form factor predictions are available the produced MC samples are phase space samples produced flat in the helicity angles. The simulation samples are used to train a BDT and determine the shapes of the tail parameters of the function used to fit the signal peak. Therefore, it is important that the simulation samples are correctly truthmatched and calibrated to ensure no biases from mismodelling in MC are introduced.

Truthmatching

The background category (BKGCAT) truthmatching tool cannot be used alone for the truthmatching of the $\Xi_b^0 \to \Xi^0\psi$ decays. The problem is that the tool has no category that considers a background from combining one real signal photon and one random photon into a π^0 candidate. This can lead to different treatment of simulation in data, where rejection of such candidates is not trivial. Therefore, it was crucial to study the possibilities with a truthmatching through the background category and TRUEID approaches as well as a combination of both.

MC Calibration

Due to the large amount of neutral particles in the decay chain and the challenges they pose in their reconstruction, the disagreement between data and simulation is expected to exceed the previously discussed and corrected mismodelling. Especially the description of π^0 candidates is expected to be mismodelled in simulation. In order to ensure a correct modelling of the simulation with respect to the collision data, the samples need to be carefully checked for disagreements[2].

DecayTreeFitter failure rate

The effect of the DecayTreeFitter (DTF) on the accuracy of the reconstructed variables for the $\Xi_b^0 \to \Xi^0\psi$ decays is expected to be significantly larger compared to the previously discussed analyses. Since the π^0 is assumed to fly from the PV, the first order estimation of quantities like the momentum are expected to be largely biased, which also effects the Ξ^0 quantaties.
Unfortunately, the limited knowledge of the Ξ^0 decay vertex and therefore the origin of the π^0 meson is expected to cause a higher and mass-dependent failure rate of DTF, making the produced variables unreliable. Therefore, the last big challenge is to reduce the DTF failure rate in order to increase the available statistics and ensure the produced variables can be used in the signal selection. This is in particular important with variables related to the Ξ^0 hyperon, as its decay vertex can only be estimated by the DTF.

7.2 Background studies and signal selection

Similarly to the previous analyses, the dominant background source in this study is the combinatorial background due to random π^0 and π^0 candidates formed from random photons. This is a result of the abundant production of both photons and π^0 mesons in the initial collisions and underlying decays in processes that are not related to the signal. This background also influences the truthmatching of the simulation, which will therefore be discussed first.

7.2.1 MC truthmatching

Due to the missing check for π^0 candidates from one real and one random photon, the TRUEID truthmatching approach was studied first. When matching the TRUEID for the Ξ_b^0, Ξ^0 and Λ candidate, there remains a large asymmetric tail in the Ξ^0 invariant mass $m(\Lambda\pi^0)$. This can be seen in the dark blue distribution in the left plot of Fig. 7.2. The first assumption is, that this asymmetry arises due to the challenging reconstruction of the π^0 candidates. Therefore, on top of the Ξ_b^0, Ξ^0 and Λ TRUEIDs, the ones of the π^0 and the two γ candidates are also matched. While the tail is not fully gone, the light blue distribution in the left plot of Fig. 7.2 shows that the tail is significantly reduced. The most likely reason for the asymmetric tail is that the algorithm used to ensure that a cell is not double counted, when the two γ ECAL cluster are overlapping, was never trained

[2]The proper calibration of the simulation samples lies outside of the scope of this thesis.

on actual data. Therefore, it is likely that the π^0 energy might be overestimated due to double counting and therefore leading to a tail on the right side of the Ξ^0 invariant mass.

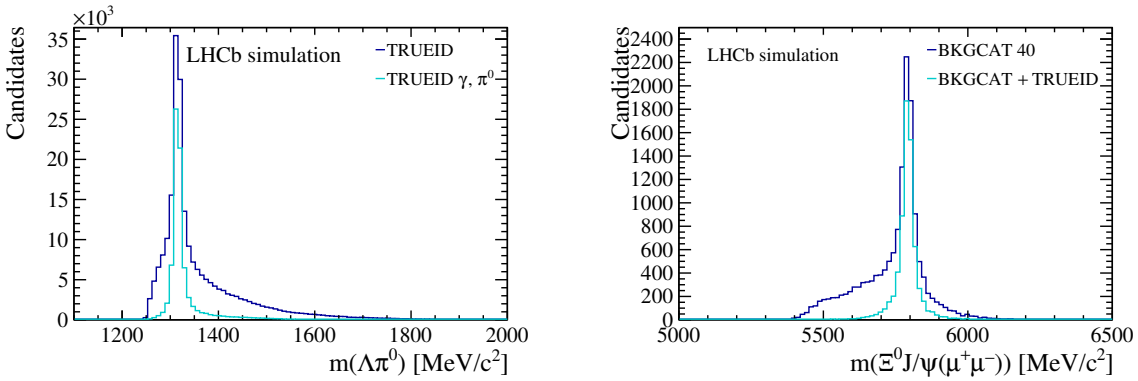


Figure 7.2: Invariant mass $m(\Lambda\pi^0)$ with TRUEID truthmatching for the Ξ_b^0, Ξ^0 and Λ candidates in dark blue and additional TRUEID requirements on the photons and π^0 in light blue (left). Invariant mass $m(\Xi^0 J/\psi)$ for background category 40 truthmatching in dark blue and with additional Ξ_b^0, Ξ^0 and Λ TRUEID requirements in light blue showing a clear signal peak (right). Both plots have been prepared with the full $\Xi_b^- \to \Xi^- J/\psi$ simulation.

As part of the truthmatching approach with the background categories, significant amount of signal found in the partially reconstructed background category (BKGCAT 40). This signal can be seen in the right plot of Fig. 7.2, which does not disappear when applying the Ξ_b^0, Ξ^0 and Λ TRUEID matching (light blue).
The final truthmatching strategy is based on the TRUEID applied to all particles beside the photons and ensures the mother TRUEID matches for the source of the Ξ^0, Λ and the π^0 candidate. Some events with a background category of 60 or above remain, classifying them as being background. Therefore, it is necessary to apply an additional truthmatching requirement that the Ξ^0 background category is below 60. Fig. 7.3 shows the comparison of the TRUEID truthmatching only applied to Ξ_b^0, Ξ^0 and Λ (dark blue), the truthmatching through the background category being smaller than 0 or 50 for Ξ_b^0, Ξ^0 and Λ candidates (dark green), and the final selection (ligh blue) for the $m(\Xi^0 J/\psi)$ (left) and $m(\Lambda\pi^0)$ (right) invariant mass distributions. It can be seen that the asymmetric tail in the Ξ^0 mass distribution is nearly fully removed, while the small asymmetry in the Ξ_b^0 mass distribution is small and remains rather unaffected by the different truthmatching strategies.

7.2.2 Cut-based offline selection

The data studied in this analysis are selected with the `RareBaryonicMuMuLine` stripping line designed to search for Ξ_b^0 decay modes such as $\Xi_b^0 \to \Xi^0\mu^+\mu^-$, $\Xi_b^0 \to \Xi^0 J/\psi$ and $\Xi_b^0 \to \Xi^0\psi(2S)$. The details of the stripping line can be found in the App. C, but it should be noted that the only applied cut for the π^0 candidate is that the transverse momentum should be above 100 MeV/c. This leads to the introduction of significant background levels due to random pions. The signal candidates are triggering on the signal (TOS) based on the muons as it has been done in all other analyses presented in this thesis. For the

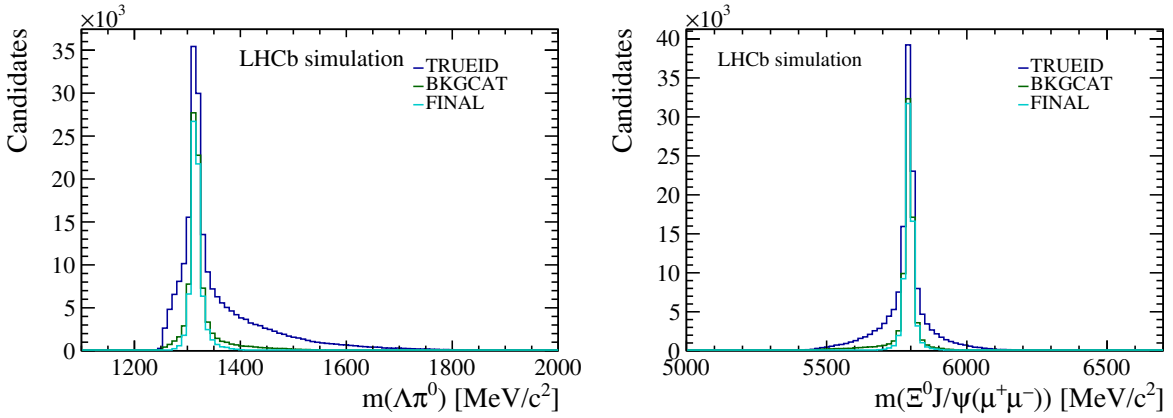


Figure 7.3: Invariant masses $m(\Lambda\pi^0)$ (left) and $m(\Xi^0 J/\psi)$ (right) comparing the TRUEID (Ξ_b^0, Ξ^0 andΛ), background category (0 or 50 for Ξ_b^0, Ξ^0 and Λ) and final truthmatching strategy. Both plots have been prepared with the full $\Xi_b^- \to \Xi^- J/\psi$ simulation.

hardware trigger, the event needs to either pass the `L0Muon` or `L0DiMuon` trigger selection. For the HLT1 trigger selection the event must either be selected by an MVA line or events a mass close to or above the J/ψ mass. The HLT2 selection is passed by events with dimuon vertices with a significant displacement from the primary vertex.

Background sources

The long-lived times of both the Ξ^0 and Λ baryon lead to the Λ nearly always decaying outside of the VELO. Therefore, requiring that both the track type of the proton and pion are downstream tracks, together with the selection of real Λ candidates, as it will be explained in Sec. 8.2.1, it can be used as a preselection to obtain manageable data samples.
In addition, tight cuts are applied around the known J/ψ and $\psi(2S)$ mass respectively to select real charmonia. Together with the selection of the Λ baryon, this explains additionally why the combinatorial background is driven by the random π^0 and γ candidates. Besides, the impact on the source of the combinatorial background, the applied mass windows sufficiently suppress any background arising from a hadron misidentified as lepton. To further suppress this lepton misidentification background, PID cuts of `ProbNNmu` > 0.1 are applied to both muon candidates. In principle the J/ψ channel can be polluted by partially reconstructed background from an excited charmonia $\Xi_b^0 \to \Xi^0 Y(\to J/\psi X)$. However due to the large mass difference between the J/ψ and the excited states, this background is not affecting the signal region. This type of background is negligible in the $\psi(2S)$ region.
As previously mentioned and already discussed for the charged Ξ_b baryon, the topology with two long-lived hyperons can also not be mimicked by any mesonic decay. Therefore, only other baryonic decay channels are the main source of physics backgrounds. For the completeness, hadronic misidentification backgrounds such as $\pi - p$ could move a B meson decay of the kind $B \to K_S^0(\to \pi^+\pi^-)\psi$ into the signal region, when a random π^0 candidate is additionally picked up. However, since only the K_S^0 is a long-lived particle, selecting requirements for two long-lived particles will lead to a negligible level of this background.

The most important physics backgrounds arise from Λ_b^0 baryon decay modes due to its much higher production fraction compared to the one of the Ξ_b^0. In addition, the masses of the Λ_b^0 and Ξ_b^0 baryon only differ by only about 170 MeV/c^2. The decay $\Lambda_b^0 \to pK^-(\to \pi^-\pi^0)J/\psi$ leads to the exact same final state, but at lower masses. In addition, the requirement of the Λ that the proton and charged pion candidate originate from the same vertex and form a narrow invariant-mass window allows to suppress this background, as proton and kaon from the decay $\Lambda_b^0 \to pK^-J/\psi$ do not originate from the same point in space. When a $\Lambda_b^0 \to \Lambda J/\psi$ decay is combined with a random π^0, this overreconstructed background peaks just below the signal. Therefore, it is necessary to remove it with a dedicated mass veto. Fig. 7.4 displays the Λ_b^0 mass veto in green and the inverted veto in dark blue, showing a clear Λ_b^0 peak at the known Λ_b^0 mass [33], highlighted in light blue. In principle, the same is true for the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay, however the branching fraction of this decay mode is several orders of magnitudes smaller and therefore in combination with the already discussed selections negligible.

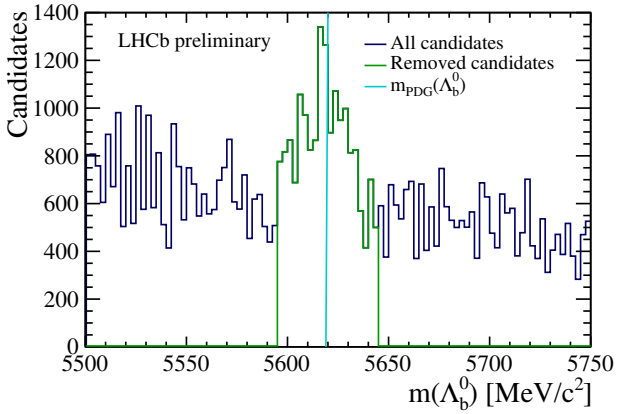


Figure 7.4: Visualisation of the Λ_b^0 signal peak from the over-reconstructed $\Lambda_b^0 \to \Lambda J/\psi$ background removed by the mass veto in green, the inverted mass veto in dark-blue and the known Λ_b^0 mass [33] highlighted in light blue. The plots are prepared with the 2018 dataset.

Since the production fraction of the Ξ_b^- baryon is roughly the same as the one of the Ξ_b^0 baryon due to isospin symmetry, it can create non-negligible backgrounds. The partially reconstructed decay mode $\Xi_b^- \to \Xi^-(\to \Lambda\pi^-)J/\psi$, where one charged pion is not reconstructed and a random π^0 candidate is added, can peak in the signal region. In order to correctly estimate the contribution of this decay mode and if it can be neglected or needs to be included in the invariant mass fit, a dedicated study of a $\Xi_b^- \to \Xi^-J/\psi$ simulation sample is needed. However, this study lays outside of the scope of this thesis and for simplicity the background is assumed to be negligible. Similarly, the decay mode $\Xi_b^- \to \Lambda(\to \pi^-p)K^-(\to \pi^-\pi^0)J/\psi$ leads to the same final state and due to the small mass of the π^- would not be moved outside of the signal window. In principle, semileptonic decay modes like $\Xi_b \to \Xi_c(\to \Xi\mu\nu_\mu)\mu\nu_\mu$ can also lead to the same final state particles, however they are suppressed by the applied ψ windows. The last source of backgrounds arise from Ω_b^- baryon decays. However, they are expected to be negligible due to the much smaller production fraction compared to that of the Ξ_b baryons. Nevertheless, two decays, $\Omega_b^- \to \Omega^-(\to \Lambda K^-(\to \pi^-\pi^0))J/\psi$ and $\Omega_b^- \to \Omega^-(\to \Xi^0\pi^-)J/\psi$, could contribute to par-

tially reconstructed background when one charged pion is not reconstructed. In general it is expected, that all decay modes relying on the decay $K^-(\to \pi^-\pi^0)$ are suppressed by its branching fraction of about 21% [33]. In addition, the long lifetime of the kaon would result in its decay happening far behind the magnet. Therefore, the $\Lambda_b^0 \to pK^-(\to \pi^-\pi^0)J/\psi$, $\Xi_b^- \to \Lambda K^-(\to \pi^-\pi^0)J/\psi$ and $\Omega_b^- \to \Omega^-(\to \Lambda K^-(\to \pi^-\pi^0))J/\psi$ decays are all expected to be negligible.
In general, pollution backgrounds producing a real Λ and/or J/ψ candidates can be moved in the signal region when wrongfully combined with a random π^0. The DTF is a powerful tool to handle this kind of background. In addition, the dominant expected source of background, the combinatorial background is expected to be suppressed sufficiently by a multivariate analysis (MVA).

Mass fits

Before attempting to fix the DTF failure rate and therefore improve the performance of a possible trained BDT for the MVA, a simple cut-based selection was attempted to observe the Ξ_b^0 signal peak in the J/ψ mode. This is possible since the chosen fit range automatically excludes all events where the DTF fit has failed.
In order to lower the background to a manageable level, additional rectangular cuts on top of the ones discussed in the previous section are needed. One of the most powerful cuts is expected to be the decay time of the Ξ^0 candidate: it profits from the long lifetime of the Ξ^0 hyperon. As the Ξ^0 decay vertex cannot be built in a conventional manner (it decays to neutral particles), this variable can only be obtained using the DTF. In order to reduce the possible loss of signal due to the DTF failure, the cut on the Ξ^0 decay time is the only one applied to a DTF variable. All other cuts are based on standard variables returned by the reconstruction software. The first of these cuts is placed on the χ^2 of the reconstructed decay vertex of the Ξ_b^0. The decay-time cut on the Λ candidate is tightened and a cut on the transverse momentum of the Ξ^0 candidate is applied. In addition, the masses of the π^0, Λ and the Ξ^0 candidates are constrained to their world-average values. In order to reduce the background related to the random photons or hadrons in the calorimeter, requirements based on the product of transverse momenta, and product of the PID response (CL) for the two photons are applied. Lastly, a opening angle cut is placed on the Ξ^0 and the π^0 to reduce the number of selected random π^0 candidates. The final set of cuts, as of the status during the thesis writing, are summarised in Tab. 7.1.
The signal peak of the invariant mass distribution $m(\Xi^0 J/\psi(\to \mu^+\mu^-))$ is modelled by a sum of two double-sided Crystal-Ball functions, where the means and resolution parameters are shared. To increase the fit stability in fits to data, the fraction between the two functions, as well as their tail parameters are fixed to the values obtained from fits to truthmatched and fully selected simulation samples. The combinatorial background is modelled by an exponential with a free slope. The obtained invariant mass fits for the $\Xi_b^0 \to \Xi^0 J/\psi$ (left) and $\Xi_b^0 \to \Xi^0 \psi(2S)$ (right) decays using the 2018 datasets can be seen in Fig 7.5. The full cut-based selection is applied and multiple candidates are removed.
While there is a clear mass peak visible in the J/ψ channel, for the $\psi(2S)$ there is too much background to draw any conclusion. However, to clarify the nature of the visible enhancement, more data and a more sophisticated selection is needed.

Table 7.1: Preliminary offline selection requirements for the search for the $\Xi_b^0 \to \Xi^0 J/\psi$ and $\Xi_b^0 \to \Xi^0 \psi(2S)$ decays.

Type	Particle(s)	Requirement
Quality & Acc	μ	`InAccMuon, hasRich`
PID	μ	`ProbNNmu>0.1`
Mass windows	J/ψ	$\lvert m(\mu^+\mu^-) - 3096\rvert < 100$ MeV/c^2
	$\psi(2S)$	$\lvert m(\mu^+\mu^-) - 3686\rvert < 100$ MeV/c^2
	Λ	$\lvert m(p\pi^-) - 1116\rvert < 8$ MeV/c^2
Background rejection	Λ_b^0 veto	$\lvert m(\Lambda\psi) - 5620 - m(\psi) + m_{PDG}(\psi)\rvert > 25$ MeV/c^2
	real Λ	$0 < \tau(\Lambda) < 2$ ns, `DIRA`> 0, $\chi^2_{\rm FD} > 0$, $0 < {\rm VTX}_z < 2320$ mm
	track type	track type $(p) = 5$ (downstream)
	track type	track type $(\pi^-) = 5$ (downstream)
Cut-based selection (later replaced with BDT)	Ξ_b^0	χ^2 EndVTX < 30
	Ξ^0 mass	$\lvert m(\Lambda\pi^0) - 1315 - m(p\pi^-) + 1115 - m(\pi^0) + 135\rvert < 20$ MeV/c^2
	Ξ^0 decay time	$\tau(\Xi^0) > 3$ ps
	Ξ^0	$p_{\rm T} > 1000$ MeV/c
	γ PID	$CL(\gamma_1) \cdot CL(\gamma_2) > 0.15$
	γ	$p_{\rm T}(\gamma_1) \cdot p_{\rm T}(\gamma_2) > 35000$ MeV^2/c^2
	opening angle	$\cos([(\eta_{\Xi^0} - \eta_{\pi^0})^2 + (\phi_{\Xi^0} - \phi_{\pi^0})^2]^{1/2}) > 0.8$
	π^0 mass	$\lvert m(\pi^0) - 135\rvert < 20$ MeV/c^2

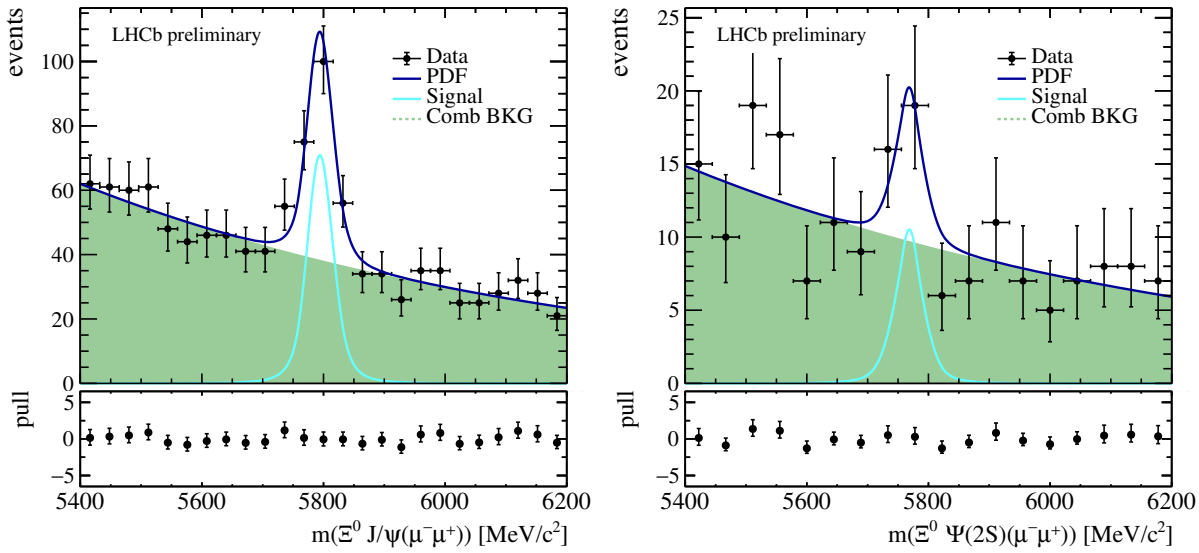


Figure 7.5: Invariant mass fits for the $\Xi_b^- \to \Xi^- J/\psi$ (left) and $\Xi_b^0 \to \Xi^0 \psi(2S)$ (right) decay using the 2018 dataset.

7.3 With DTF towards signal peaks

As shown in the measurements presented in the previous chapters, the selection through an MVA significantly improves the selection efficiencies. In addition, the DTF tool is crucial to calculate the correct parameters of a track, especially for decays involving neutral particles. Therefore, the DTF variables used in an MVA selection are expected to significantly help with the signal selection for the challenging Ξ_b^0 decays.
It should be noted that the failure rate of the DTF is quite high for the studied decays. It is assumed that this is driven by two effects. The first being from backgrounds combining a part of the signal decay and a random photon to form a π^0 candidate, and the second is that the DTF fails more often when the reconstructed invariant masses differ significantly from the PDG values used by the mass constraint. Both effects would lead to a mass-dependence of the failure rate and could potentially shape the background. In fact, after applying only the preselection, the DTF fails more often than it converges for both data and simulation, as can be seen in the top left plot for simulation and bottom for data in Fig. 7.6. In addition, it was checked if the fraction of failed DTF fits is reduced by applying truthmatching to the simulated signal. This indeed leads to a significant reduction of the failure rate, but does not completely remove it, as can be seen in the right plot with light blue bins from the truthmatched simulation in Fig. 7.6. Therefore, it is expected that the failed DTF events in data include not only background events but also signal events. Consequently, it is decided to fix the DTF failure, in order to recover the otherwise potentially lost signal candidates and remove the risk of introducing a bias. This is particularly important due to the expected low signal yields, however, the fix will simultaneously lead to an increase in the background level.

The DTF as described in Sec. 4.3 is a tool that is called when reading the stripping output via functions that enable to calculate the respective variables. In order to fix or at least reduce the DTF failure rate, the implementation of the tool itself needs to be modified. While there are many ways to alter the DTF setup, it was decided to only change two parameter values: the maximum number of iteration in the DTF fit and the number of iterations in which the difference in fit χ^2 between two fit iterations, $\Delta\chi^2_{\mathrm{DTF}}$, is allowed to increase. The maximum number of iterations allowed to the DTF fit was increased from 10 to 1000 iterations, while $\Delta\chi^2_{\mathrm{DTF}}$ increased from 3 to 15 attempts. This choice was made, as the changes do not alter any calculations performed in the DTF implementation itself but simply give the fit more time for convergence. The default values of these parameters in the DTF implementation have been optimised as a trade-off between having the fit converging for complicated fits and required computing time. Therefore, the presented changes will improve the fit convergence while simultaneously increasing the computing time. In addition, the DTF tool is forced to calculate the DTF variables also for non-converged fits, rather than simply returning a default value. This enables to use the DTF variables as input parameters in the BDT training without excluding events.
The described changes in the DTF implementation have two additional effects. First, the DTF fit is performed for all PV hypotheses in cases of multiple ones present in one event. The different PV fits are sorted according to the one that gives better χ^2. While in this study only the "best" PV was considered, the study of the different PVs could enable further input on the source of the DTF failure rate. Second, the DTF fit status is returned for each fits: 0 for converged fits, 1 for failed fits and 3 for non-converged ones.

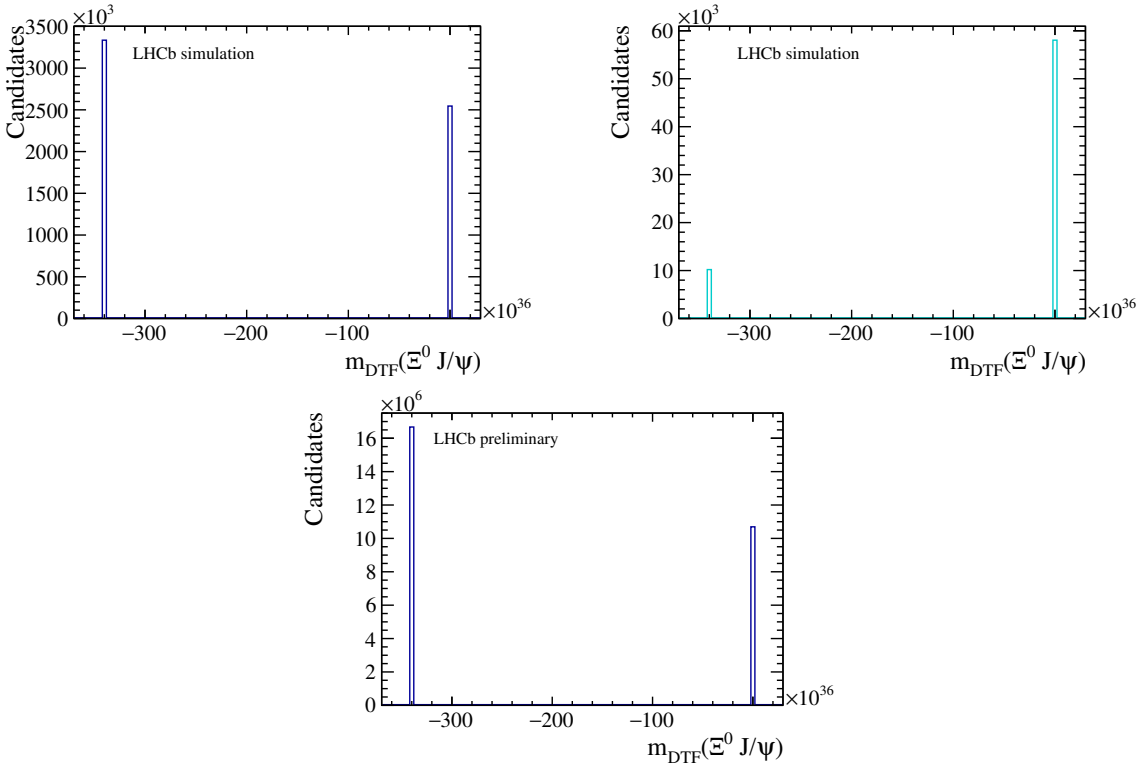


Figure 7.6: Visualisation of the DTF with the invariant mass $m_{\mathrm{DTF}(\Xi J/\psi)}$ for full $\Xi_b^0 \to \Xi^0 J/\psi$ simulation before (top left) and after (top right) truthmatching, as well as for the 2018 collision data (bottom). All three plots are from samples obtained through the default LHCb implementation of the DTF tool.

The classification of the three categories is based on the comparison of the $\Delta\chi^2_{\mathrm{DTF}}$ value and two additional parameters $\Delta\chi^2_{\mathrm{DTF,\,conv}}$ and $\Delta\chi^2_{\mathrm{DTF,\,quit}}$ that are set to defined values:

1. *converged*: $\Delta\chi^2_{\mathrm{DTF}} < \Delta\chi^2_{\mathrm{DTF,\,conv}}$,
2. *failed*: $\Delta\chi^2_{\mathrm{DTF}} > \Delta\chi^2_{\mathrm{DTF,\,quit}}$
3. and *non-converged*, if the converged and failed requirement are not met within the 1000 DTF fits and the 15 iterations $\Delta\chi^2_{\mathrm{DTF}}$ can increase.

As it can be seen in Fig. 7.7, only a small fraction of the fits on signal simulation did not converge. For the not truthmatched simulation more of the fits failed than converged as it can be seen in the left plot. However, after applying the truthmatching (right plot) only a small fraction of fits failed, further supporting the hypothesis that the DTF failure rate is much higher for background events.
This motivates to reject failed or non-converged fits to reduce background level in data. However, a dedicated study determining whether the DTF failure rate for signal is well modelled in simulation is needed in order to not introduce a possible bias. Additionally, the rejection of these two DTF status categories could sculpt variables such as the invariant mass of the Ξ^0 candidate. A more straightforward study would be to determine the impact of adjusting the values of the $\Delta\chi^2_{\mathrm{DTF,\,conv}}$ and $\Delta\chi^2_{\mathrm{DTF,\,quit}}$ to further reduce the rate of

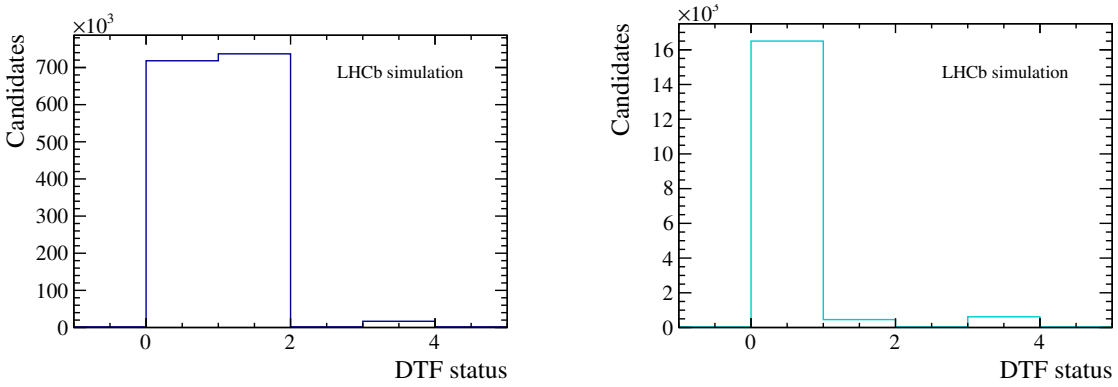


Figure 7.7: DTF status for the invariant mass $m_{\mathrm{DTF}(\Xi J/\psi)}$ as returned by the adapted DTF implementation for $\Xi_b^0 \to \Xi^0 J/\psi$ simulation before (left) and after (right) truthmatching.

failed and non-converged DTF fits. Both of these studies have not yet been performed and therefore are subject of future studies in order to converge on a final DTF strategy for the first measurement of the branching fractions of the $\Xi_b^0 \to \Xi^0 J/\psi$ and $\Xi_b^0 \to \Xi^0 \psi(2S)$ decays.

7.3.1 Preliminary BDT selection

The BDT presented in this section is the outcome from several previous studies, that were used to evaluate the effect of the DTF improvements. However, it is expected that the presented setup can be further optimised. This is in particular driven by the fact that the simulation is expected to be mismodelled significantly and has not been calibrated yet. In addition, as just discussed, there is still room for further DTF investigations. Consequently, this BDT is labelled as preliminary but nevertheless helps to understand the impact of a MVA selection on the reconstruction of the challenging Ξ_b^0 decay modes.

The BDT is trained utilising truthmatched $\Xi_b^0 \to \Xi^0 J/\psi$ simulation samples as a signal proxy and the lower and upper data sidebands, excluding the Ξ_b^0 signal region, as a background proxy. Both the signal and background sample are 2018 samples and are preselected with the trigger selection and the requirements in Tab. 7.1, excluding the cut-based selection. A large set of variables is studied and reduced to the following seven input variables:

- $\chi^2_{\mathrm{DTF}}(\Xi_b^0)$
- $\texttt{DIRA}(\Xi_b^0)$
- $\tau_{\mathrm{DTF}}(\Xi^0)$
- $m(\Xi^0)$
- $\chi^2_{EV}(\Lambda)$
- $p_{\mathrm{T}}\ (\Lambda)$
- $\chi^2_{OVTX}(J/\psi)$

The BDT response has been checked for signs of overtraining by comparing the training and testing sample. The left plot in Fig. 7.8 shows the BDT response in logarithmic scale. While there is a tiny hint of false classification of background as signal, there is no clear sign of overtraining. Additionally, the same setup was evaluated with a k-fold cross-validation of ten folds. The obtained ROC curves are displayed in the right plot of Fig. 7.8. All folds provide a uniform ROC curve with the same AUC score of 0.996.

Consequently, the BDT performance is expected to not depend on the given dataset. Therefore, it is not only used for the J/ψ decay mode, but also applied to the $\psi(2S)$ sample.

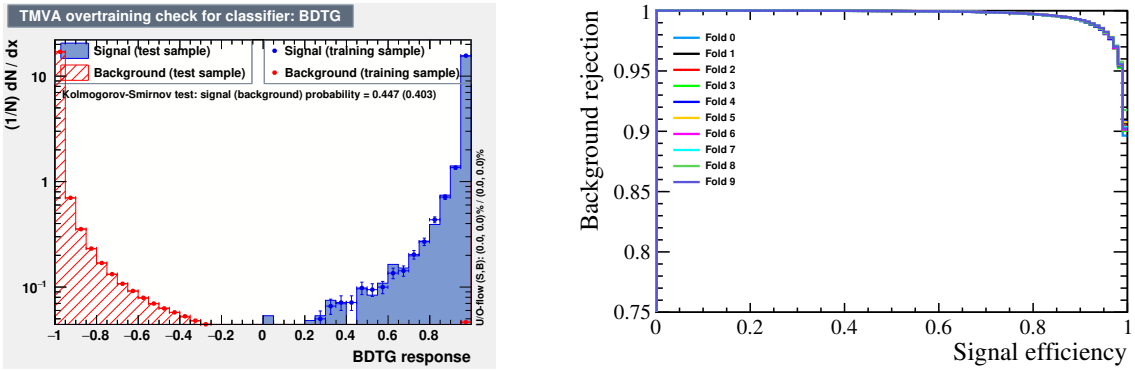


Figure 7.8: Evaluation the BDT response utilising a test and training sample (left). ROC curves for the ten folds to evaluate the statistical independence of the BDT (right). Each of the folds is displayed in a different colour. Both plots are prepared using the 2018 samples. Taken from Ref. [220].

The optimal BDT cut value is determined using a Punzi FOM, where the endeavoured significance a is chosen to be $a = 5$, equally to the strategy discussed in Sec. 6.3.2. The maximum of the Punzi FOM is determined independently for both decay modes, using the respective $\Xi_b^0 \to \Xi^0 J/\psi$ and $\Xi_b^0 \to \Xi^0 \psi(2S)$ samples. The best BDT cut for the J/ψ sample is found at 0.9925 and for the $\psi(2S)$ sample at 0.990. Both BDT cut values are very tight, this is a result of the effect of the input variables for the BDT training. Since the simulation was not calibrated yet, it was decided to not use variables related to the photons or π^0 candidate in the training. However, most of the background originates from random π^0 candidates. Therefore, the BDT has only information through the Ξ^0 available to veto this background. More details about the BDT setup, including the variable distributions, correlation matrices and the FOM plots can be found in the App. F.1.
After applying the preselection discussed in Tab. 7.1, excluding the cut-based selection, but applying the respective BDT cut, multiple candidates are removed as a last step of the selection. The obtained dataset is then fitted with the same fit model again. The resulting mass fits are shown in Fig. 7.9. Due to the missing calibration of the simulation and the overall early stage of this analysis, only absolute and rounded numbers are quoted. The signal yield for the $\Xi_b^0 \to \Xi^0 J/\psi$ decay is about 700 events with a resolution of about $24\,\mathrm{MeV}/c^2$ and the $\Xi_b^0 \to \Xi^0 \psi(2S)$ signal yield is about 60 events with a resolution of about $25\,\mathrm{MeV}/c^2$. The signal yield and resolution of the $\Xi_b^- \to \Xi^- J/\psi$ decay are shown in Fig. E.7 and in Fig. E.9 for the $\Xi_b^- \to \Xi^- \psi(2S)$ decay. When comparing both decay modes, it is found that the signal yield is a factor 2.3 larger for the charged Ξ_b^- decays and the resolution is a factor 1.3 worse for the J/ψ decays and even 2.9 for the $\psi(2S)$ decays. Both effects are caused by the challenges in the π^0 reconstruction.
While there is still room to improve the selection and background suppression, the selection through the MVA approach using the full available dataset of $5.4\,\mathrm{fb}^{-1}$ might lead to not only observing the $\Xi_b^0 \to \Xi^0 J/\psi$ but also the $\Xi_b^0 \to \Xi^0 \psi(2S)$ decay for the first time. However, before claiming the first observation of the $\Xi_b^0 \to \Xi^0 \psi(2S)$ decay, further work

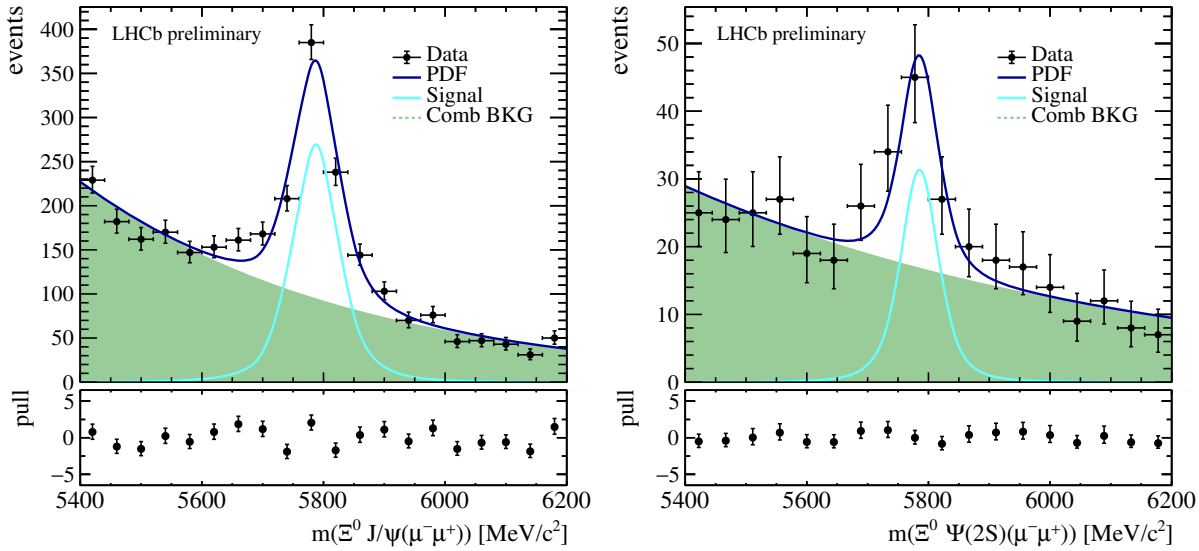


Figure 7.9: The full available LHCb dataset (2016-18) shown by the black data points. Invariant mass fits for the $\Xi_b^0 \to \Xi^0 J/\psi$ (left) and $\Xi_b^0 \to \Xi^0 \psi(2S)$ (right) decays.

is needed to rule out possible background contributions, as this is only a very preliminary result. Notable are the imperfections in the fit model description of the data, that might hint for the need of the study of the possible contributions. One example could be the partially reconstructed backgrounds arising from decays of the charged Ξ_b^- baryon, combined with a random π^0. As discussed in Sec. 7.2, the detailed investigation of this background source is laying outside of the progress discussed in this thesis. Additionally, the mass constraints on the hyperons and especially on the charmonia are known to help with the signal resolution. In order to verify that the constraints do not create an artificial peak by sculpting the background, it is also important to study the impact of the constraints themselves. The last but the most important bit is that the differences between data and simulation, especially in relation to the π^0 properties, needs to be understood in order to ensure the robustness of the studies and selection.

7.4 Summary and open challenges

The preliminary results shown in this Chapter strongly suggest the first observation of the $\Xi_b^0 \to \Xi^0 J/\psi$ decay and shows strong evidence that this might also be true for the $\Xi_b^0 \to \Xi^0 \psi(2S)$ decay. Consequently, this marks also the first time the Ξ^0 hyperon has been successfully reconstructed at the LHC. Considering that the quark model predictions of the ratio of branching fractions of the $\psi(2S)$ and J/ψ decay modes are not only challenged for the Λ_b^0 baryon [23, 24], but also for the charged Ξ_b^- baryon as discussed in Ch. 6, the measurement of the same ratio R (7.1) for the Ξ_b^0 baryon is even more interesting.
However, before measuring the ratio of branching fractions R, the composition of the $\Xi_b^0 \to \Xi^0 J/\psi$ and $\Xi_b^0 \to \Xi^0 \psi(2S)$ samples needs to be thoroughly investigated. Under the assumption that the peak is only caused by the decay of interest in absence of peaking

backgrounds, the yield of observed $\Xi_b^0 \to \Xi^0\psi(2S)$ decays can be used to estimate an upper limit on the rare mode $\Xi_b^0 \to \Xi^0\mu^+\mu^-$ signal yield. Since studies of other decay modes [7, 24] suggest that the signal yield of the respective $b \to s\ell^+\ell^-$ transition based decays are a factor 2–4 lower than their corresponding $\psi(2S)$ decay mode[3], most likely, only an upper limit on the branching fraction of the $\Xi_b^0 \to \Xi^0\mu^+\mu^-$ decay could be set. However, if the background level could be lowered significantly while at the same time retaining the current (assumed) signal level, chances of an observation could be improved. Still, these assumptions are still rather of speculative nature and significant further work is needed to obtain reasonable conclusions both on the composition of the peak in the $\Xi_b^0 \to \Xi^0\psi(2S)$ sample and the possibility of observing the $\Xi_b^0 \to \Xi^0\mu^+\mu^-$ decay with the partial Run 2 sample.
As discussed in Sec. 7.3, the impact of the changes in the DTF implementation and possible room for further improvements need to be studied in the future in order to ensure the obtained results are reliable. Additionally, the impact of possible partially reconstructed backgrounds from the Ξ_b^- baryon need to be studied. However, the most important open challenges are the calibration of the simulation and the understanding of the π^0 reconstruction. Since the latter is the dominant source of creating backgrounds and lowering the resolution of the signal, a bias can easily be introduced. For the calibration of the simulation, it is expected that the Ξ_b^0 baryon kinematics and the PID response needs to be calibrated for the muons and photons, if the latter will be used in the selection. Fig. F.8 in App. F.2 shows the comparison of the BDT input variables between the uncorrected simulation and the background-subtracted data. It can be seen that some of the variables show discrepancies between the two samples, highlighting the need for a proper calibration of the simulation to ensure the BDT does not introduce a bias.
To conclude, this chapter showed first promising hints for the observation of the $\Xi_b^0 \to \Xi^0 J/\psi$ and $\Xi_b^0 \to \Xi^0\psi(2S)$ decay, but further work is needed to confirm this interpretation.

[3]For example in Ref. [63], the signal yield for the decay mode $\Lambda_b^0 \to pK^-\psi(2S)(\to \mu^+\mu^-)$ is 421 (Fig. 3.67 p.161) and for the corresponding rare decay $\Lambda_b^0 \to pK^-\mu^+\mu^-$ is 198 (Fig. 3.70 on p.164) using the 2016 dataset. Consequently, the factor between both yields is about 2.1.

Chapter 8

Angular analysis of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ at high q^2

This chapter describes the current status of the angular analysis of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay. The analysis is performed in the high q^2 region above $15\,\mathrm{GeV}^2/c^4$ and aims to extract all ten angular coefficients with a maximum-likelihood fit for the first time using the full Run 1 and Run 2 dataset of the LHCb experiment, corresponding to an integrated luminosity of $9\,\mathrm{fb}^{-1}$. The described analysis has been performed under the supervision of Dr. Marie-Hélène Schune and Prof. Dr. Johannes Albrecht.
In the beginning an introduction to the used angular basis and the angular distribution is given, before the signal selection and background suppression are discussed. Afterwards, the analysis strategy and construction of the full fit model is elaborated, before the validation of that model is described. The next section is then dedicated to the discussion of systematic uncertainties that need to be studied to extract the final result. Finlly, the chapter concludes with open steps before the analysis can be reviewed and unblinded to extract the final result from the fit to data.

8.1 The foundation of the angular analysis of $\Lambda_b^0 \to \Lambda\ell^+\ell^-$

Before discussing the angular analysis of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay, it is useful to revisit how angular analyses help to improve our understanding of possible NP. As discussed in Sec. 1.4.1, $b \to s\ell^+\ell^-$ transitions are a powerful tool to search for NP effects by extracting limits on Wilson coefficients. While branching fractions are the most straight forward measurement to perform, it is extremely difficult to disentangle different Wilson coefficients from them. In contrast, disentangling the coefficients is much easier with angular analyses, because they have many more angular observables. The differential decay width is defined by the sum of the square value of the transversity amplitudes A_λ

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} = \sum_\lambda |A_\lambda|^2. \tag{8.1}$$

The transversity amplitudes $A_\lambda(C_i, H_i(FF))$ depend on Wilson coefficients C_i as well as on helicity amplitudes H_i, which themselves depend on form factors. Therefore, angular

analyses not only enable to extract limits on Wilson coefficients, but to test form factor predictions. With enough available statistics, it is important to perform a q^2-binned analysis as the non-local form factor contribution introduce a q^2 dependence to the amplitudes.
This chapter is dedicated to the angular analysis of the decay $\Lambda_b^0 \to \Lambda \ell^+ \ell^-$ developed for the muon decay mode, focusing on the high-q^2 region as most of the signal is to be found in that region. A schematic representation of the angles studied in this analysis and describing the decay can be found in Fig. 8.1.

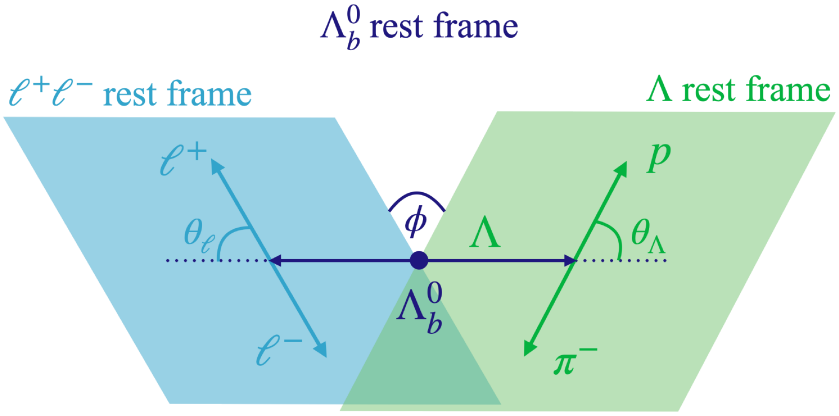


Figure 8.1: Sketch displaying the definition of the angles describing the $\Lambda_b^0 \to \Lambda \ell^+ \ell^-$ decay.

8.1.1 Angular basis

The angular basis used in this analysis, follows the commonly used basis within the LHCb collaboration to be consistent with previous angular analyses of $b \to s\ell^+\ell^-$ transitions. This enables an easier interpretation of the different results and to extract global limits on the Wilson coefficients.
The differential decay width of the $\Lambda_b^0 \to \Lambda \ell^+ \ell^-$ decay can be described as a function of three angles[1] ($\cos\theta_\ell$, $\cos\theta_\Lambda$, ϕ) and q^2, the invariant mass squared of the dilepton system. The angle θ_ℓ is defined between the positive charged lepton $\ell^+(\ell^-)$ in the dilepton rest frame and the direction of the dileptons in the Λ_b^0 ($\overline{\Lambda}_b^0$) rest frame. The angle θ_Λ is defined between the $p(\bar{p})$ in the Λ ($\overline{\Lambda}$) rest frame and the direction of the Λ ($\overline{\Lambda}$) in the Λ_b^0 ($\overline{\Lambda}_b^0$) rest frame. And finally, the angle ϕ is defined between the dilepton and the Λ ($\overline{\Lambda}$) plane.

[1]This is only true under the assumption of unpolarised Λ_b^0 decays. This choice was made since the LHCb experiment, ATLAS and the CMS collaborations providing a measurement of the Λ_b^0 polarisation consistent with zero [225–227].

The definition of the angles for the Λ_b^0 decay are given by

$$\begin{aligned}
\cos\theta_\ell &= \left(\hat{p}_{\ell^+}^{(\ell\ell)}\right)\cdot\left(\hat{p}_{\ell\ell}^{(\Lambda_b^0)}\right)\\
\cos\theta_\Lambda &= \left(\hat{p}_{p}^{(\Lambda)}\right)\cdot\left(\hat{p}_{\Lambda}^{(\Lambda_b^0)}\right)\\
\cos\phi &= \left(\hat{p}_{\ell^+}^{(\Lambda_b^0)}\times\hat{p}_{\ell^-}^{(\Lambda_b^0)}\right)\cdot\left(\hat{p}_{p}^{(\Lambda_b^0)}\times\hat{p}_{\pi^-}^{(\Lambda_b^0)}\right)\\
\sin\phi &= \left[\left(\hat{p}_{\ell^+}^{(\Lambda_b^0)}\times\hat{p}_{\ell^-}^{(\Lambda_b^0)}\right)\cdot\left(\hat{p}_{p}^{(\Lambda_b^0)}\times\hat{p}_{\pi^-}^{(\Lambda_b^0)}\right)\right]\cdot\hat{p}_{\Lambda}^{(\Lambda_b^0)}
\end{aligned} \tag{8.2}$$

utilising the normalised to unity momentum vector $\hat{p}_i^{(j)}$ of particle i in the rest frame j. Under the *CP* transformation the three angles transform as followed

$$\begin{gathered}
\theta_\ell \to \theta_\ell,\\
\theta_\Lambda \to \theta_\Lambda,\\
\text{and}\ \ \phi \to 2\pi - \phi.
\end{gathered}$$

Therefore, the angles for the $\overline{\Lambda}_b^0$ decay are defined as

$$\begin{aligned}
\cos\theta_\ell &= \left(\hat{p}_{\ell^-}^{(\ell\ell)}\right)\cdot\left(\hat{p}_{\ell\ell}^{(\bar{\Lambda}_b^0)}\right)\\
\cos\theta_\Lambda &= \left(\hat{p}_{\bar{p}}^{(\bar{\Lambda})}\right)\cdot\left(\hat{p}_{\bar{\Lambda}}^{(\bar{\Lambda}_b^0)}\right)\\
\cos\phi &= \left(\hat{p}_{\ell^-}^{(\bar{\Lambda}_b^0)}\times\hat{p}_{\ell^+}^{(\bar{\Lambda}_b^0)}\right)\cdot\left(\hat{p}_{\bar{p}}^{(\bar{\Lambda}_b^0)}\times\hat{p}_{\pi^+}^{(\bar{\Lambda}_b^0)}\right)\\
\sin\phi &= -\left[\left(\hat{p}_{\ell^-}^{(\bar{\Lambda}_b^0)}\times\hat{p}_{\ell^+}^{(\bar{\Lambda}_b^0)}\right)\cdot\left(\hat{p}_{\bar{p}}^{(\bar{\Lambda}_b^0)}\times\hat{p}_{\pi^+}^{(\bar{\Lambda}_b^0)}\right)\right]\cdot\hat{p}_{\bar{\Lambda}}^{(\bar{\Lambda}_b^0)}
\end{aligned} \tag{8.3}$$

8.1.2 Angular distribution

Using the described angular basis, the full angular distribution of the $\Lambda_b^0 \to \Lambda\ell^+\ell^-$ decay is given by the 4-fold differential decay width taken from Ref. [228]:

$$\begin{aligned}
K(q^2,\phi,\cos\theta_\ell,\cos\theta_\Lambda) &= \frac{8\pi}{3}\frac{\mathrm{d}^4\Gamma}{\mathrm{d}q^2\mathrm{d}\phi\mathrm{d}\cos\theta_\ell\mathrm{d}\cos\theta_\Lambda}\\
&= \big[(K_{1ss}\sin^2\theta_\ell + K_{1cc}\cos^2\theta_\ell + K_{1c}\cos\theta_\ell)\\
&+ (K_{2ss}\sin^2\theta_\ell + K_{2cc}\cos^2\theta_\ell + K_{2c}\cos\theta_\ell)\cos\theta_\Lambda\\
&+ (K_{3sc}\sin\theta_\ell\cos\theta_\ell + K_{3s}\sin\theta_\ell)\sin\theta_\Lambda\sin\phi\\
&+ (K_{4sc}\sin\theta_\ell\cos\theta_\ell + K_{4s}\sin\theta_\ell)\sin\theta_\Lambda\cos\phi\big].
\end{aligned} \tag{8.4}$$

The angular coefficients K_i depend on the dilepton invariant mass squared, q^2, the form factors and the Wilson coefficients. Therefore, they are sensitive to NP effects as previously discussed. They can be expressed in terms of the transversity amplitudes $A^{R,L}{}_{\perp_{1,0},\parallel_{1,0}}$, which depend on the helicity amplitudes of the $\Lambda_b^0 \to \Lambda$ transition and the decay parameter α in $\Lambda \to N\pi$. The derivation and exact definition of the transversity

amplitudes $A^{R,L}{}_{\perp_{1,0},\parallel_{1,0}}$ can be found in Ref. [228]. While helicity amplitudes are in general easier to obtain, they have mixed *CP* contributions. Therefore, they are transformed into the transversity base, which leads to *CP* even and odd amplitudes. This step is needed, since the analysis is performed as a *CP* averaged measurement.
The full definition of the angular observables is given in (8.5).

$$\begin{aligned}
K_{1ss}(q^2) &= \frac{1}{4}\left[|A^R_{\perp_1}|^2 + |A^R_{\parallel_1}|^2 + 2|A^R_{\perp_0}|^2 + 2|A^R_{\parallel_0}|^2 + (R \leftrightarrow L)\right] \\
K_{1cc}(q^2) &= \frac{1}{2}\left[|A^R_{\perp_1}|^2 + |A^R_{\parallel_1}|^2 + (R \leftrightarrow L)\right] \\
K_{1c}(q^2) &= -\mathrm{Re}\left\{A^R_{\perp_1} A^{*R}_{\parallel_1} - (R \leftrightarrow L)\right\} \\
K_{2ss}(q^2) &= +\frac{\alpha}{2}\mathrm{Re}\left\{A^R_{\perp_1} A^{*R}_{\parallel_1} + 2A^R_{\perp_0} A^{*R}_{\parallel_0} + (R \leftrightarrow L)\right\} \\
K_{2cc}(q^2) &= +\alpha\mathrm{Re}\left\{A^R_{\perp_1} A^{*R}_{\parallel_1} + (R \leftrightarrow L)\right\} \\
K_{2c}(q^2) &= -\frac{\alpha}{2}\left[|A^R_{\perp_1}|^2 + |A^R_{\parallel_1}|^2 - (R \leftrightarrow L)\right] \\
K_{3sc}(q^2) &= +\frac{\alpha}{\sqrt{2}}\mathrm{Im}\left\{A^R_{\perp_1} A^{*R}_{\perp_0} - A^R_{\parallel_1} A^{*R}_{\parallel_0} + (R \leftrightarrow L)\right\} \\
K_{3s}(q^2) &= +\frac{\alpha}{\sqrt{2}}\mathrm{Im}\left\{A^R_{\perp_1} A^{*R}_{\parallel_0} - A^R_{\parallel_1} A^{*R}_{\perp_0} - (R \leftrightarrow L)\right\} \\
K_{4sc}(q^2) &= +\frac{\alpha}{\sqrt{2}}\mathrm{Re}\left\{A^R_{\perp_1} A^{*R}_{\parallel_0} - A^R_{\parallel_1} A^{*R}_{\perp_0} + (R \leftrightarrow L)\right\} \\
K_{4s}(q^2) &= -\frac{\alpha}{\sqrt{2}}\mathrm{Re}\left\{A^R_{\perp_1} A^{*R}_{\perp_0} - A^R_{\parallel_1} A^{*R}_{\parallel_0} - (R \leftrightarrow L)\right\}
\end{aligned} \tag{8.5}$$

There are four commonly known observables, which are linear combinations and ratios of the initial angular observables in (8.5). All of these observables are normalised to the differential decay width in q^2

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} = 2K_{1ss} + K_{1cc} = 1 \tag{8.6}$$

to reduce the degree of freedom of the angular fit by one. The fraction of the transverse polarisation of the dilepton system

$$F_1 = \frac{2K_{1cc}}{2K_{1ss} + K_{1cc}} \tag{8.7}$$

is one of them and directly linked with the fraction of longitudinal polarisation

$$F_0 = 1 - F_1 = \frac{2K_{1ss} - K_{1cc}}{2K_{1ss} + K_{1cc}}. \tag{8.8}$$

The other three observables are the forward-backward asymmetries with respect to the leptonic scattering A^ℓ_{FB}, to the baryonic scattering angle A^Λ_{FB} and a combination of both asymmetries $A^{\ell\Lambda}_{\mathrm{FB}}$:

$$\begin{aligned}
A^\ell_{\mathrm{FB}} &= \frac{3}{2}\frac{K_{1c}}{2K_{1ss} + K_{1cc}}, \\
A^\Lambda_{\mathrm{FB}} &= \frac{1}{2}\frac{2K_{2ss} + K_{2cc}}{2K_{1ss} + K_{1cc}}, \\
A^{\ell\Lambda}_{\mathrm{FB}} &= \frac{3}{4}\frac{K_{2c}}{2K_{1ss} + K_{1cc}}.
\end{aligned} \tag{8.9}$$

These four observables can be calculated from the final fit in order to test the compatibility with the Run 1 result reported in Ref. [7]. Next to statistical limitations, one of the largest uncertainty on angular observables arises from uncertainties on the theoretical predictions of the form factors. Therefore, it is important to find clean observables, that allow to use ratios of the angular coefficients to cancel the form factor contributions. In the case of the decay $\Lambda_b^0 \to \Lambda \ell^+ \ell^-$, there is one clean observable

$$X_1 = \frac{K_{1c}}{K_{2cc}}. \tag{8.10}$$

In the absence of chirality-flipped operators (including $\mathcal{O}_{7'} \to 0$), there is one other clean observable

$$X_2 = \frac{K_{1cc}}{K_{2c}} \tag{8.11}$$

and the angular coefficient is $K_{3s} \to 0$.
The measurement is performed for Λ_b^0 and $\overline{\Lambda}_b^0$ simultaneously, therefore the CP averaged angular distribution of (8.4) needs to be used:

$$\begin{aligned}
&\frac{1}{\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} + \frac{\mathrm{d}\bar{\Gamma}}{\mathrm{d}q^2}} \left\langle \frac{\mathrm{d}^4(\Gamma + \bar{\Gamma})}{\mathrm{d}q^2 \mathrm{d}\phi \mathrm{d}\cos\theta_\ell \mathrm{d}\cos\theta_\Lambda} \right\rangle_{CP\text{ averaged}} \\
&\qquad = \frac{3}{8\pi}\big[(S_{1ss}\sin^2\theta_\ell + (1 - 2S_{1ss})\cos^2\theta_\ell + S_{1c}\cos\theta_\ell) \\
&\qquad + (S_{2ss}\sin^2\theta_\ell + S_{2cc}\cos^2\theta_\ell + S_{2c}\cos\theta_\ell)\cos\theta_\Lambda \\
&\qquad + (S_{3sc}\sin\theta_\ell\cos\theta_\ell + S_{3s}\sin\theta_\ell)\sin\theta_\Lambda\sin\phi \\
&\qquad + (S_{4sc}\sin\theta_\ell\cos\theta_\ell + S_{4s}\sin\theta_\ell)\sin\theta_\Lambda\cos\phi\big].
\end{aligned} \tag{8.12}$$

The new normalised and CP even coefficients are defined by

$$S_i = \frac{K_i + \bar{K}_i}{\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} + \frac{\mathrm{d}\bar{\Gamma}}{\mathrm{d}q^2}}. \tag{8.13}$$

As the measurement will be performed in one q^2 bin only, the full angular distribution (8.12) will be called three-dimensional in the following. Due to the statistical limitations of the available data samples, it might be useful to only perform a two- or even one-dimensional likelihood fit to extract the angular observables.
The SM predictions calculated using `EOS` [229] can be found in Ref. [228, 230]:

$$\begin{aligned}
K_{1ss} &= 0.353 \pm 0.00295 \\
K_{1cc} &= 0.295 \pm 0.00591 \\
K_{1c} &= -0.234 \pm 0.00887 \\
K_{2ss} &= -0.227 \pm 0.00631 \\
K_{2cc} &= -0.177 \pm 0.00667 \\
K_{2c} &= 0.218 \pm 0.00455 \\
K_{3sc} &= (6.45 \times 10^{-5}) \pm 0.000153 \\
K_{3s} &= -0.000351 \pm 0.000428 \\
K_{4sc} &= -0.0232 \pm 0.00602 \\
K_{4s} &= -0.121 \pm 0.0109.
\end{aligned} \tag{8.14}$$

In addition, it is possible to express the CP averaged angular distribution (8.12) in terms of CP asymmetries:

$$A_i = \frac{K_i - \bar{K}_i}{\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} + \frac{\mathrm{d}\bar{\Gamma}}{\mathrm{d}q^2}}. \tag{8.15}$$

To access angular distributions with CP asymmetries the following transformations have to be performed:

$$\theta_\Lambda \to \theta_\Lambda + \pi \text{ and} \tag{8.16}$$

$$\theta_\ell \to \pi - \theta_\ell \tag{8.17}$$

$$\begin{aligned}
&\frac{1}{\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} + \frac{\mathrm{d}\bar{\Gamma}}{\mathrm{d}q^2}} \left\langle \frac{\mathrm{d}^4(\Gamma + \bar{\Gamma})}{\mathrm{d}q^2 \mathrm{d}\phi \mathrm{d}\cos\theta_\ell \mathrm{d}\cos\theta_\Lambda} \right\rangle_{CP \text{ averaged}} \\
&\qquad = \frac{3}{8\pi}\big[(S_{1ss}\sin^2\theta_\ell + (1 - 2S_{1ss})\cos^2\theta_\ell + A_{1c}\cos\theta_\ell) \\
&\qquad + (A_{2ss}\sin^2\theta_\ell + A_{2cc}\cos^2\theta_\ell + A_{2c}\cos\theta_\ell)\cos\theta_\Lambda \\
&\qquad + (A_{3sc}\sin\theta_\ell\cos\theta_\ell + A_{3s}\sin\theta_\ell)\sin\theta_\Lambda\sin\phi \\
&\qquad + (A_{4sc}\sin\theta_\ell\cos\theta_\ell + A_{4s}\sin\theta_\ell)\sin\theta_\Lambda\cos\phi\big].
\end{aligned} \tag{8.18}$$

It is important to note, that the transformations do not give access to the CP odd angular observables A_{1ss} and A_{1cc}. To access all CP even and odd angular coefficients the analysis needs to be performed independently for the Λ_b^0 and $\bar{\Lambda}_b^0$ decay. In general, this is possible as the decay of the Λ baryon allows to tag the flavour of the initial b baryon.

8.2 Signal selection for the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay

Before discussing the angular analysis, the section covers the signal selection for the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay. It concludes with the mass fit, which is needed in the angular fit in order to distinguish signal and background candidates. It should be noted that the Λ_b^0 decay studied in this analysis has a subsequent long-lived hyperon decay, $\Lambda \to p\pi^-$, with the lifetime of the Λ being $(2.632 \pm 0.020) \times 10^{-10}$ ps. As a result, the Λ baryon can decay either inside or outside the VELO, depending on its decay time (flight distance) and be reconstructed via its decay products as either *long* (L) tracks or *downstream* (D) tracks. The two track type categories, displayed in Fig. 8.2, will be merged for the extraction of the final results in order to increase the fit stability, but dedicated cross-checks are performed.

8.2.1 Preselection

The inclusive stripping line `Bu2LLK_mmLine` has been used for this measurement utilising full Run 1 and Run 2 dataset of the LHCb experiment and the specific cuts are reviewed in the Tab. C.1 in App. C. The trigger selection profits from presence of muons in the final state, and either `L0Muon` or `L0DiMuon` trigger decisions, applied to the dimuon pair

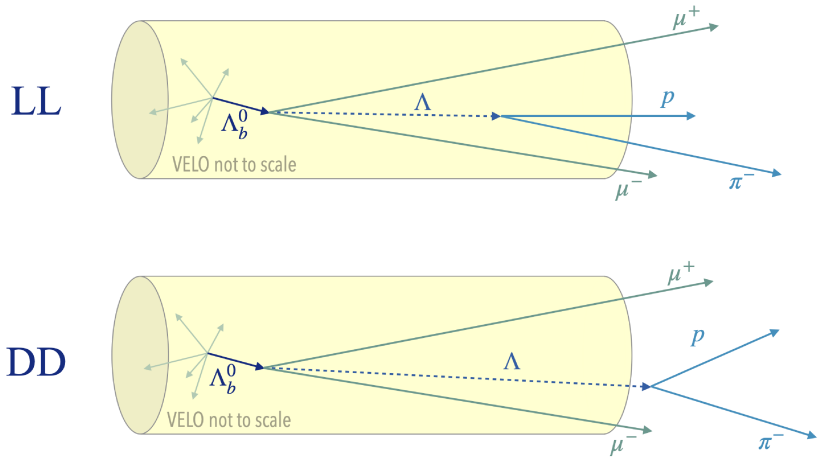


Figure 8.2: Schematic representation of the track type categories for the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay.

of the signal candidate, have to be passed. The HLT selection is similarly mostly relying on muonic trigger lines, in addition, generic topological b-hadron-decay trigger lines have been added to increase the trigger efficiency. The list of HLT lines used in this analysis are summarised in Tab. B.2 in the App. B.

The LHCb measurement of the differential branching fraction $\mathcal{B}(\Lambda_b^0 \to \Lambda\mu^+\mu^-)$ [7] (displayed in Fig. 1.7) showed that most of the signal can be found in the high-q^2 region above the $\psi(2S)$ resonance. Therefore, the targeted region to perform the angular analysis is $15.00 < q^2 < 20.28$ GeV$^2/c^4$[2]. The dimuon mass region around the J/ψ resonance, $3000 < m(\mu^+\mu^-) < 3150$ MeV/c^2 is used to select the $\Lambda_b^0 \to \Lambda J/\psi(\to \mu^+\mu^-)$ channel. The latter is used for certain calibration steps and for cross-checks, as it has the same final state as the targeted $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay.

The left plot of Fig. 8.3 features the invariant mass of $\Lambda\mu^+\mu^-$ candidates in data as a function of their dimuon invariant mass. The vertical bands corresponding to J/ψ and $\psi(2S)$ resonances combined with a Λ candidate (not necessarily from Λ_b^0 decays of interest), are well visible. The horizontal band, corresponding to the known Λ_b^0 mass value, represents the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay. The intersections of the vertical and horizontal bands correspond to the tree-level $\Lambda_b^0 \to \Lambda J/\psi(\to \mu^+\mu^-)$ and $\Lambda_b^0 \to \Lambda\psi(2S)(\to \mu^+\mu^-)$ decays. It is very well visible, that most of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ signal outside of the charmonia tree-level decay region is indeed found above the $\psi(2S)$ peak in the q^2 range. Finally, the uniformly distributed data points across the phase space correspond to backgrounds, mostly that of combinatorial nature. The diagonal cutoff in the bottom right corner stems from the fact that the $\Lambda\mu^+\mu^-$ invariant mass cannot be smaller than the sum of the $\mu^+\mu^-$ invariant mass and the known Λ mass. The slightly visible diagonal lines where the horizontal $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay line crosses the vertical charmonia lines arise from muon final-state radiation.

As a part of the previous analysis of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay [7], a cut-based preselection was developed to reject backgrounds without a genuine Λ baryon. These requirements have been revisited and applied in this analysis with a small modification. A mass window from 1108 MeV/c^2 to 1124 MeV/c^2 is placed around the Λ baryon mass [33] and only events

[2]The upper limit is defined by the difference in masses of the Λ_b^0 and Λ baryons squared.

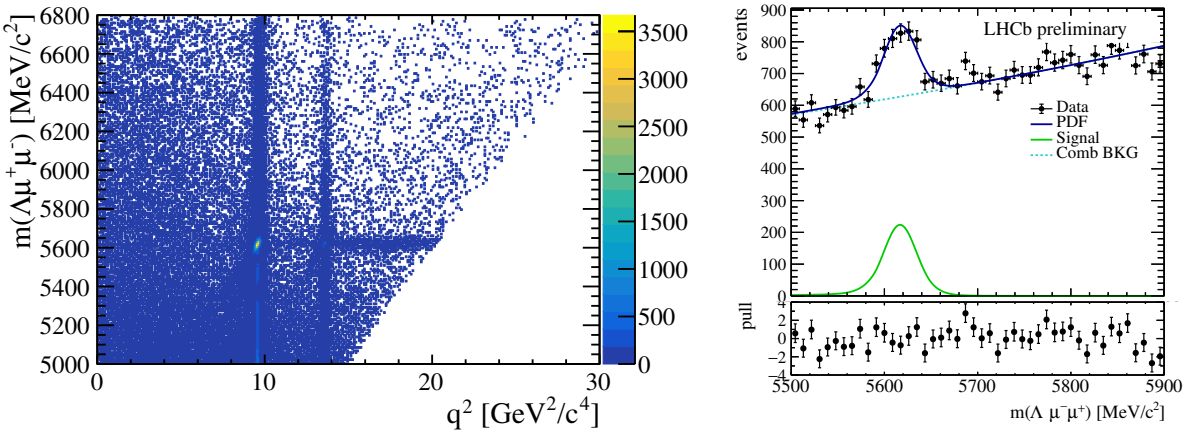


Figure 8.3: Left: Invariant mass of $\Lambda\mu^+\mu^-$ candidates as a function of dilepton invariant mass squared (q^2), in the data. The full Run 2 dataset is exploited to prepare this plot, with preselection and a tighter boosted decision tree (BDT) cut applied. Right: Invariant mass of $\Lambda\mu^+\mu^-$ candidates with only the preselection applied to the full Run 2 dataset.

with a Λ decay-time above 0 ps and below 2 ps are considered. A significant amount of Λ baryons are produced in material interactions around $z \approx 2350$ mm which corresponds to the boundary between the RICH1 exit window and the beam pipe. Therefore, all events with a decay vertex positions above $z > 2350$ mm have been excluded. Additionally, candidates with a flight distance goodness-of-fit from the primary vertex (PV) with values less than zero, and those with the cosine of the measured pointing angle of the Λ hyperon from the PV larger than zero. The remaining preselection cuts ensure to select good quality muons. They are shown in the Tab. 8.2 displaying the final selection.
The importance of the preselection can visualised by fitting[3] the invariant mass distribution $m(\Lambda\mu^+\mu^-)$ with only the preselection applied. The right plot of Fig. 8.3 shows that there is already a peak visible in the position of the Λ_b^0 mass. However, more dedicated background studies and additional selection steps are needed to ensure the peak does not include physical background, as well as to reduce the combinatorial background.

8.2.2 Simulation samples and their calibration

Two different types of simulation are utilised in this analysis in order to correctly model the signal and determine the angular acceptance influencing the extracted angles for the angular measurement. The latter is performed with phase-space simulation, where all helicity angles are produced with flat distributions. Therefore, any distortion from a straight line is a result of the detector effects and applied selection requirements. The angular acceptance is discussed in detail in Sec. 8.3.2. The $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ signal is studied with physics simulation samples based on the lattice QCD predictions [102] for the relevant form factors in the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay. These samples are used to determine the shape parameters of the signal distribution in the invariant-mass fit, and as a signal proxy in the training of a gradient boosted decision tree (BDT) used to suppress the combinatorial

[3]The fit procedure will be discussed in Sec. 8.2.6.

background described in Sec. 8.2.4. Since the simulation and collision data samples are merged in order to obtain the final result in this analysis, the requested statistics for the individual simulation samples has been scaled according to the luminosity for each respective year, as presented in Tab. 8.1.

Table 8.1: List of the generated amounts of simulated data for each year, with the two magnet polarities merged.

Year	2011	2012	2015	2016	2017	2018
Generated yield	1M	2M	0.5M	2.5M	2.5M	3.5M

Additional simulation samples have been used to study the contributions from different background sources. They are discussed in Sec. 8.2.3.
In order to select signal candidates in the reconstructed simulated samples, the samples are truthmatched using the background category tool. The truth-matching of signal candidates is done by demanding the background category of the Λ_b^0 to be equal to either 0, 10 or 50.
Before using the simulation samples as part of this analysis, they need to be calibrated, to account for known mismodelling effects with respect to the collision data that could create any biases. This is done by applying a combination of different dedicated weights, also called corrections, to the simulated distributions, in order to match them to those in data. The full correction chain is in the end applied to all used simulation samples. As a part of this thesis, three different calibrations steps are performed: the hardware trigger response, the Λ_b^0 kinematics and the Λ_b^0 lifetime value used in the production of the samples, are calibrated.

Calibration steps

The L0 trigger response is calibrated for the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ dataset according to the method introduced in Sec. 4.2.3. A large sample of $B^+ \to J/\psi K^+$ decays is used to perform this calibration, which is therefore statistically independent of all other correction steps. It has been processed with the `Bu2LLK_mmLine` stripping line as well, resulting in align muon ID requirements and kinematic selections crucial to port the calibration maps.
Checking the data-simulation agreement for the generated kinematics of the Λ_b^0 baryon revealed a significant discrepancy. This is caused partially by the fact that the Λ_b^0 baryon production in LHCb has a strong p_{T} dependence [83], which is known to be mismodelled in the simulation. As a consequence, the kinematic calibration should not be the same as for B-meson decays. Therefore, the correction is calculated using samples of the $\Lambda_b^0 \to \Lambda J/\psi$ tree-level decay, which has higher statistical power compared to the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay. In order to do that, the ${}_s\mathcal{P}lot$ technique [191] is used to access the background-subtracted data samples to be compared to the simulated samples. This can be done since the invariant mass does not exhibit as significant correlation with the kinematic variables of interest. The selection steps (trigger and preselection) between both samples have been aligned before determining the sweights and later the correction weights.
The transverse momentum p_{T} and the pseudorapidity η of the Λ_b^0 baryon are used to parametrise its kinematics. Since these two variables are somewhat correlated, the correction is computed in two dimensions simultaneously. The number of bins has been optimised in

order to catch the disagreement sufficiently but avoid introducing a bias due to statistical fluctuations in case of too fine binning. In order to overcome statistical limitations, it was necessary to calculate the maps for a combined Run 1 and a combined Run 2 samples, instead of each year individually. Indeed, while the Λ_b^0 kinematics depend on the collision energy, it should remain stable within different data-taking years at the same energy. In case of the Run 1 samples, 200 iso-populated two-dimensional bins are used to calculated the correction map, for the Run 2 sample 250 bins are used. The resulting correction maps for both samples are shown in Fig. 8.4.

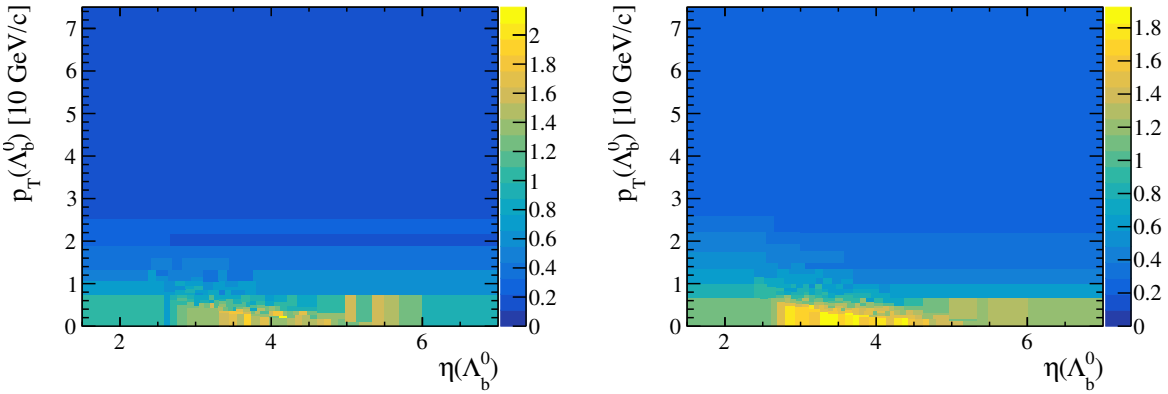


Figure 8.4: Kinematic weight maps for the Run 1 (left) and Run 2 (right) simulation samples.

Since the corrections are determined with the L0 trigger requirements applied to the $\Lambda_b^0 \to \Lambda J/\psi$ samples, the mismodelling of the trigger response can cause a bias in the kinematic corrections. Therefore, a dedicated systematic uncertainty will be determined. The Λ_b^0 lifetime value used in the productions of the simulation samples is $\tau_{\text{tag}} = 1.451$ ps, which slightly differs from the current average value of $\tau_{\text{average}} = (1.471 \pm 0.009)$ ps from HFLAV [231]. To correct for this, an analytical reweighting $w_{\tau_{\Lambda_b^0}}$ is applied to the Λ_b^0 decay-time, based on the ratio of two exponential distributions

$$w_{\tau_{\Lambda_b^0}} = \exp\left(-t/\tau_{\text{average}}\right) / \exp\left(-t/\tau_{\text{tag}}\right),$$

where t is the true decay-time of the candidate. The effect of this reweighting is rather small, but still included to improve the data-simulation agreement.
The resulting calibrated distributions have been compared to the background-subtracted data to ensure the success of the calibration steps. As an example the Λ_b^0 transverse momentum p_T, the pseudo-rapidity η and decay-time τ, displayed in Fig. 8.5, show very good agreement.

Additional calibration steps

In order to ensure, that the calibration does not create any unexpected shapes in the simulation of the angles used in the determination of the angular acceptance, the distributions between the uncalibrated and calibrated samples of each angle are compared. No significant difference are observed.
It should be noted that one crucial step of the correction of the simulation samples, the

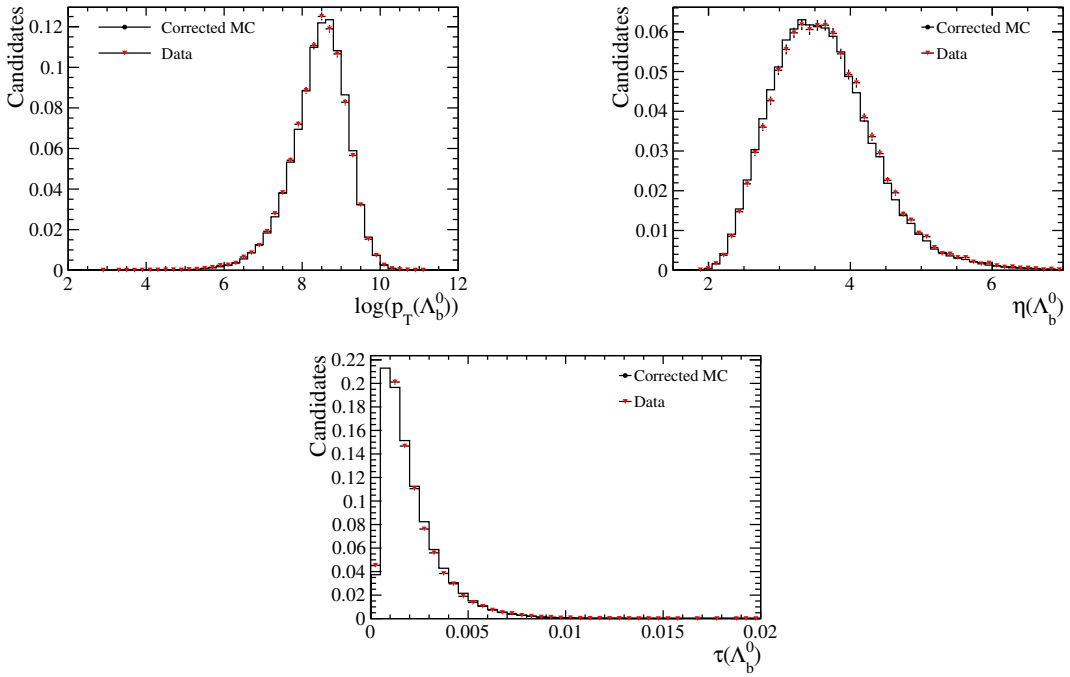


Figure 8.5: Comparison of the Run 2 $\Lambda_b^0 \to \Lambda J/\psi$ samples between background-subtracted data (red triangles) and calibrated simulation (black histogram).

calibration of the PID requirements used in the selection of the analysed samples, is still missing. Since the calibration of the PID variables can influence the other steps of the calibration it is done as a first step. Unfortunately, the calibration of the muon ID variables was only ongoing during the process of writing this thesis. Therefore, this step has not been included in the discussion and shown results. Once the PID calibration is done, the calibration steps discussed in this section will be redone to ensure a complete and unbiased correction of any mismodelling. Nevertheless, the disagreement between the simulated samples and the background subtracted data for the `ProbNNmu` variables are accessed after the discussed calibration and only a small difference remains, that is only visible on the logarithmic scale as shown in Fig. 8.6. Therefore, it is expected that the calibration will only have a small impact on the analysis, and the discussed results in this thesis should not change drastically once the analysis is rerun will the full calibration. Additional residual mismodelling can be found in track multiplicity distribution, however due to the small effect of this correction on other variables, it is decided to include it only as source of a systematic uncertainty. Indeed, while the PID efficiency in general depends on the track multiplicity, the effect is minimal for the muon ID which mostly relies on the muon system that has relatively low occupancy. Similarly, the fraction of track types is known to be mismodelled in Run 1, while the Run 2 samples show a reasonable agreement after the calibration. Therefore, the effect of a possible additional calibration will only be included as systematic uncertainty.

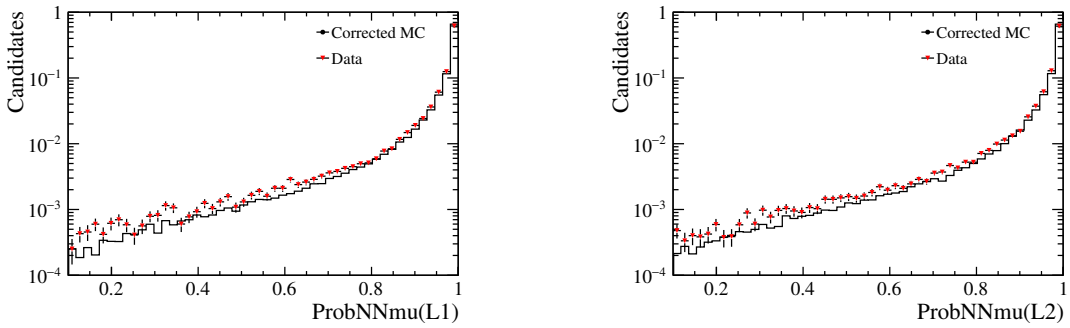


Figure 8.6: Comparison of the Run 2 $\Lambda_b^0 \to \Lambda J/\psi$ samples between background-subtracted data (red triangles) and calibrated simulation (black histogram) for the `ProbNNmu` PID variable for both muon candidates.

8.2.3 Background sources

In order to optimise the selection further, it is important to understand the background components that might pollute the signal region. However, the isospin zero of the Λ_b^0 baryon and the distinct topology due to the Λ hyperon decay leaves rather few possibilities for significant backgrounds that need to be studied. This section describes the investigation of the origin of these backgrounds, quantifying their amount and describing their invariant-mass shape, if they cannot be removed by applying additional selection cuts on top of the preselection. In addition, a BDT is trained to remove the combinatorial background, which is the dominant background source in this analysis. Besides the latter, there are four more physics background categories to be considered: tree-level decays proceeding through the dilepton resonances, partially reconstructed backgrounds, misidentification, and (double) semileptonic ones.

Combinatorial background

The dominant combinatorial background originates from the combination of random tracks in the detector. The applied preselection leads to the selection of mostly real Λ decays, therefore the composition of the combinatorial background is driven by random combinations of Λ candidates with muons. It is continuously distributed across the broad invariant mass range. In the lower sideband the combinatorial background can additionally interplay with the semileptonic backgrounds discussed below. As previously mentioned, a dedicated multivariate classifier is trained to suppress it, described later in Sec. 8.2.4. It should be noted that the composition of the combinatorial background for the resonant charmonia decay $\Lambda_b^0 \to \Lambda J/\psi$ ($\Lambda_b^0 \to \Lambda\psi(2S)$) changes due to additionally applied mass constraint for the J/ψ ($\psi(2S)$) decay. Therefore, the background level might even be higher since it consists out of random combinations from real Λ hyperons and J/ψ resonances. However, due to the much larger signal yield of the tree-level J/ψ decay the signal to background ratio is much larger compared to the rare $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay.

Resonances in the dilepton invariant mass

The $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay can proceed through several resonant modes $\Lambda_b^0 \to \Lambda X$ with the subleading decay $X \to \mu^+\mu^-$, producing the same final state particles. The two most prominent resonant components correspond to the $\Lambda_b^0 \to \Lambda J/\psi$ and $\Lambda_b^0 \to \Lambda\psi(2S)$ decays. The contribution of other resonances decaying to dilepton (in particular, $\rho(770)$, $\omega(782)$ and $\phi(1020)$) are expected to be negligible due to CKM suppression of the Λ_b^0 decays to a Λ and a light meson, as well as due to the small ($\sim 10^{-5}$-$\sim 10^{-4}$) branching fractions of light mesons to two muons. Additionally, all these backgrounds lay outside of the studied high q^2 region and therefore none of these backgrounds contribute to the signal region.
It is important to highlight that this assumption is not only true for the tree-level decays discussed above, but also for all other decays discussed in the following paragraphs that could have their dimuon pair arising from an $X \to \mu^+\mu^-$ decay.

Semileptonic backgrounds

The final state of interest can also be reached via the double semileptonic decay $\Lambda_b^0 \to \Lambda_c^+\mu^-\bar{\nu}_\mu$ with $\Lambda_c^+(\to \Lambda\mu^+\nu_\mu)$. It is expected to be shifted to lower mass values and not peaking in the signal region, since studies of semileptonic decays show that the neutrinos carry significant amount of energy. In addition, the decay would lead to an exponential shape that will be merged with the combinatorial background. In Fig. 8.7, the double semileptonic background should manifest itself as a bump at low mass values (below the mass of the Λ_c^+ baryon, $m_{PDG} = 2286.46$ MeV/c^2). Such a bump is not present in the fully selected sample. This means, double semileptonic background can be neglected in this analysis.

Lepton misidentification backgrounds

In general, backgrounds due to muon misidentifications such as $\Lambda_b^0 \to \Lambda_c^+\pi^-$ and $\Lambda_b^0 \to \Lambda_c^+K^-$ are expected to be negligible, after applying the stripping-level PID cuts. The cut `ProbNNmu>0.1` is applied on the signal muons to further suppress any possible background due to muon misidentification.
However, two possible background sources are studied in more detail, the single semileptonic decay $\Lambda_b^0 \to \Lambda_c^+(\to \Lambda\mu^+\nu_\mu)\pi^-$, where the pion is misidentified as lepton, and the charmless $\Lambda_b^0 \to \Lambda\pi^+\pi^-$ background with double misidentification. The single misidentification background has been checked by searching for the Λ_c^+ mass peak ($m_{PDG} = 2286.46$ MeV/c^2) in $\Lambda\mu^+$ combination with pion mass hypothesis on the muon, as one can see in Fig. 8.7. The absence of any Λ_c^+ mass peak also proves that the lepton misidentification backgrounds mentioned above can be neglected.
Similarly, the double misidentification background from the charmless decay mode can be checked the invariant mass distributions when mass hypotheses of both muons is replaced by the pion ones. However, due to the closeness of the muon and pion mass an additional particle identification requirement of `ProbNNpi` > 0.5 has been placed on both muon candidates. The resulting distributions is shown in Fig. 8.8. Since the peak is not close to the Λ_b^0 mass, it is assumed that it is some remaining $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ misidentification background and not a charmless $\Lambda_b^0 \to \Lambda\pi^+\pi^-$ decay. Therefore, the latter can be safely neglected.

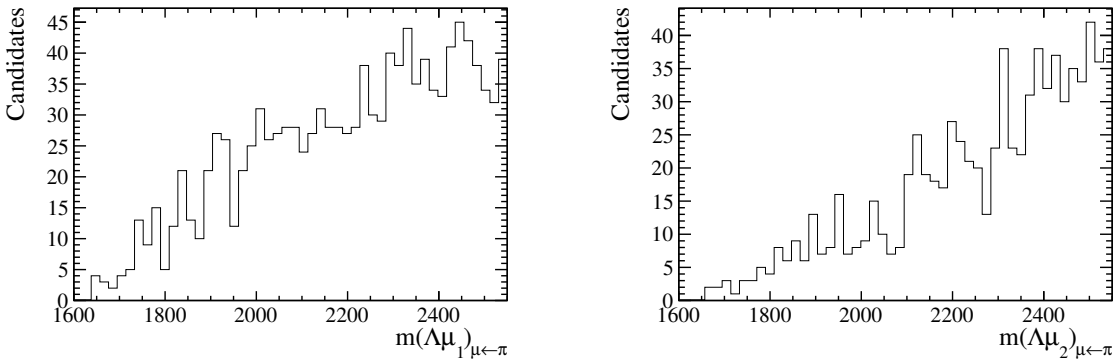


Figure 8.7: $\Lambda\mu^+$ combination with pion mass hypothesis on the muon using the full Run 1 and Run 2 dataset for both muon candidates with the full selection including the BDT cut applied. On the left for the first muon candidate and on the right for the second muon candidate.

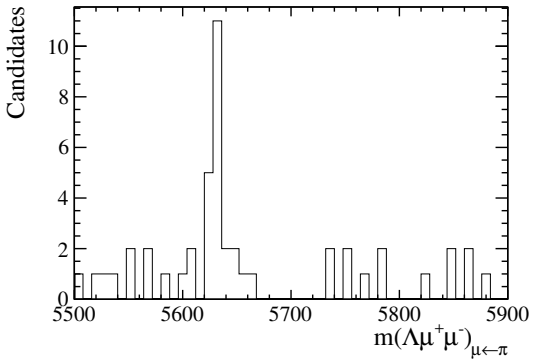


Figure 8.8: $\Lambda\mu^+\mu^-$ invariant mass with a pion mass hypothesis on both muons using the full Run 1 and Run 2 dataset with the full selection including the BDT cut and an additional particle ID requirement applied.

Hadron misidentification backgrounds

The long lifetime of the hyperon Λ helps significantly with the amount of backgrounds through hadron misidentification. Backgrounds from decays through Λ^* resonances, such as $\Lambda_b^0 \to \Lambda^*(1405, 1520, 1600)(\to pK^-)\mu^+\mu^-$, or B meson decays via $B^0 \to K^{*0}\mu^+\mu^-$ with $K^{*0} \to K^+\pi^-$, where a kaon from the resonance decays could be misidentified as pion or proton respectively, are suppressed to an negligible level by the cut on the Λ lifetime. The branching fraction of the decay mode $\Lambda_b^0 \to p\pi^-\mu^+\mu^-$ is two orders of magnitudes smaller and therefore together with the applied selection negligible.
However, the misidentification background from the $B^0 \to K_s^0(\to \pi^+\pi^-)\mu^+\mu^-$ decay needs to be studied more thoroughly. In previous analyses of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays it was found to be negligible, while it has been present in studies of the decay $\Lambda_b^0 \to \Lambda J/\psi$ [24]. Since the full Run 1 and Run 2 dataset is used in this analysis, the background source is carefully evaluated again. Therefore, the proton-pion (Λ) invariant mass is check for the presence of a significant K_S^0 mass peak ($m_{PDG} = 498$ MeV/c^2) after applying a pion mass hypothesis on the proton candidate. The resulting distribution after applying the full

selection, including the BDT cut discussed in Sec. 8.2.4, is displayed in the left plot of Fig. 8.9 and shows a small hint of a K^0_S mass peak.
As a second step a naive estimation of the yield in the full invariant mass range is performed. The $B^0 \to K^0_s \mu^+\mu^-$ yield can be naively estimated with

$$N(B^0 \to K^0_s \mu^+\mu^-) \sim \frac{\mathcal{B}(B^0 \to K^0_s \mu^+\mu^-)}{\mathcal{B}(\Lambda^0_b \to \Lambda \mu^+\mu^-)} \cdot \frac{f_{B^0}}{f_{\Lambda^0_b}} \cdot N(\Lambda^0_b \to \Lambda \mu^+\mu^-) \cdot \frac{\epsilon(B^0)}{\epsilon(\Lambda^0_b)}.$$

However, LHCb has only measured the production fraction of the Λ^0_b baryon with respect to the production fraction of the B^0 and B^+ meson, $f_{\Lambda^0_b}/(f_u + f_d) = 0.258 \pm 0.018$ [83]. Under the assumption of isospin asymmetry, one can use $f_u = f_d$ and therefore obtains about

$$N(B^0 \to K^0_s \mu^+\mu^-) \sim \frac{\mathcal{B}(B^0 \to K^0_s \mu^+\mu^-)}{\mathcal{B}(\Lambda^0_b \to \Lambda \mu^+\mu^-)} \cdot (2 \times \frac{f_{\Lambda^0_b}}{f_u + f_d})^{-1} \cdot N(\Lambda^0_b \to \Lambda \mu^+\mu^-) \cdot \frac{\epsilon(B^0)}{\epsilon(\Lambda^0_b)}.$$

For the final calculation the respective broad high q^2 bins for the branching fractions of $B^0 \to K^0_s \mu^+\mu^-$ [6] and $\Lambda^0_b \to \Lambda \mu^+\mu^-$ [7] have been used and combined with an estimated signal yield of $N(\Lambda^0_b \to \Lambda \mu^+\mu^-) \sim 1200$ for the full Run 1 and Run 2 dataset, leading to

$$N(B^0 \to K^0_s \mu^+\mu^-) \approx 183 \cdot \frac{\epsilon(B^0)}{\epsilon(\Lambda^0_b)}$$

background events for $B^0 \to K^0_s \mu^+\mu^-$. It should be noted, that this number accounts for the full mass range and is not restricted to the signal window. In order to estimate the number of events in the final signal window an efficiency estimation is used. A simple estimate of the efficiency for both the Λ^0_b signal and K^0_S background can be obtained by taking the number of events in a given simulation sample after the selection and dividing it through the number of initially produced events prior to any selection including the stripping line. The remaining number of backgrounds should be of the order of 1% and can therefore be neglected.
In light of the small hint of a K^0_S peak in the $m(p\pi^-)$ invariant mass under the pion mass hypothesis for the proton, a third approach is performed in order to evaluate the conclusion that the $B^0 \to K^0_s \mu^+\mu^-$ background can be neglected. The Armenteros-Podolanski plot [232] is an approach that enables to separate different two-body decays present in the dataset by visualising them with arc-like structures. The right plot of Fig. 8.9 visualises the Armenteros-Podolanski plot for the Λ hyperon decay. The transverse momentum of the proton with respect to the Λ is plotted as a function of the asymmetry of longitudinal momenta of proton and pion with respect to the Λ. The plot shows the remaining dataset after the full selection, including the BDT cut, is applied. Only one prominent arc-like structure remains corresponding to the Λ hyperon. Under careful inspection a second line can be identified crossing the Λ arc in the area of an asymmetry of longitudinal momenta between 0.7 and 0.75. This line corresponds to the K^0_S background and again shows the smallness of the contribution, which is a result of the tight Λ mass selection. The Armenteros-Podolanski plot also demonstrates that the only meaningful way to fully remove the remaining K^0_S pollution is through a dedicated mass veto. However, the latter are known to create artificial shapes in distributions and therefore this approach will only be studied as a source for a systematic uncertainty.

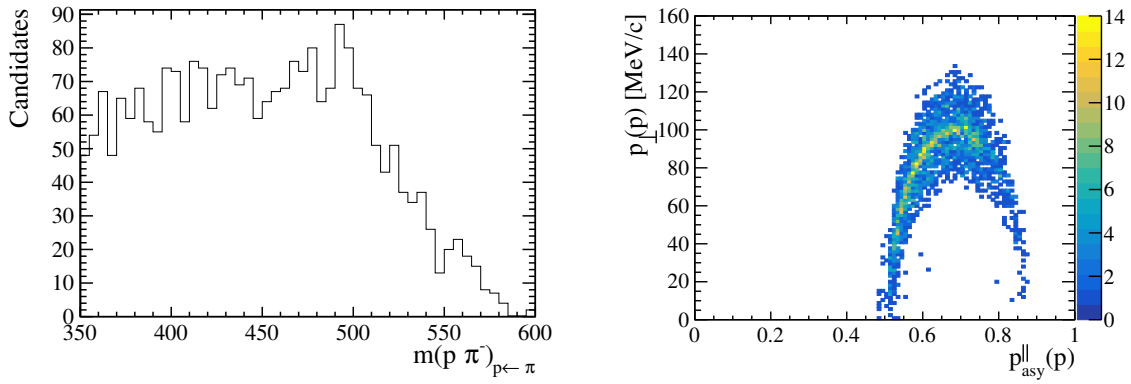


Figure 8.9: Left: $p\pi^-$ invariant mass with the proton assigned a pion mass hypothesis using the full Run 1 and Run 2 dataset and applying the full selection. A small sign of K^0_S mass peak is present. Right: Armenteros-Podolanski plot which is defined by the transverse momentum of the proton with respect to the Λ on the vertical axis, and the asymmetry of longitudinal (with respect to Λ) momenta of proton and pion on the horizontal one.

It should be noted, that Armenteros-Podolanski plots usually feature two arcs that are mirrored at zero for the Λ and the $\bar{\Lambda}$. However, the dataset has been created in a way, that the two decay products are defined by their kind as proton and pion, rather than as positive and negative charged particles. Combined with the fact that the proton has always a larger momentum compared to the pion, the Λ and the $\bar{\Lambda}$ candidates end up in the same arc.

Partially reconstructed backgrounds

A similarly important source of background arises from multibody decays, where some of the final state particles are lost because they were not reconstructed or they ended up outside of the detector acceptance. The topology of these decays can be almost identical to the one of the decay of interest. Therefore, a careful study of the so called partially reconstructed backgrounds is needed.
Decays like $\Lambda_b^0 \to \Lambda^*(\to \Lambda\pi^0)\mu^+\mu^-$ violate isospin and are consequently forbidden. However, for the higher Λ^* resonances, the decays $\Lambda_b^0 \to \Lambda^*(1520[1600])(\to \Lambda\pi^+\pi^-)\mu^+\mu^-$ are possible, but are shifted completely outside of the fit range due to two missing pions. In addition, the branching fractions are very small, making these decay modes negligible as background source. The last considered Λ_b^0 decay mode is $\Lambda_b^0 \to \Lambda^*(1405[1520,1600])\mu^+\mu^-$ with the subleading decay $\Lambda^* \to \Sigma^0(\to \Lambda\gamma)\pi^0$, where the π^0 is not reconstructed. This would shift the decay mode again outside of the fit range on top of the small branching fractions. Thus, all of the dominant decays of the Λ_b^0 involving Λ^* resonances are negligible.
A more tricky source of partially reconstructed backgrounds are those from heavier weakly decaying b-baryons. For the Ω_b^-, the decay mode $\Omega_b^- \to \Omega^-(\to \Lambda K^-)\mu^+\mu^-$, could in principle play a role as it provides the same final state in case of a soft kaon. However, due to the tiny production fraction of the Ω_b^-, this background is negligible. A bit more dangerous is the background from the decay $\Xi_b \to \Xi(\to \Lambda\pi)\mu^+\mu^-$ with a soft pion due to the

higher production fractions of both Ξ_b baryons compared to the Ω_b^- one. Additionally, the decay modes are expected to have most of their signal in the high q^2 region as well, motivating to perform a more elaborated study. This partially reconstructed background is expected to have its maximum on the left side of the signal peak outside of the fit range, but the tail could possibly reaching into the invariant mass window.
Similarly to the $B^0 \to K_s^0 \mu^+ \mu^-$ yield, the yield of the Ξ_b decays can be estimated under the use of the $SU(3)$ symmetry [212–214], which relates the partial decay width of both decay modes

$$\frac{\Gamma(\Xi_b^- \to \Xi^- \mu^+ \mu^-)}{\Gamma(\Lambda_b^0 \to \Lambda \mu^+ \mu^-)} = \frac{3}{2}.$$

While the $SU(3)$ symmetry assumption needs to be assigned an uncertainty of 30% due to the typical size of $SU(3)$-breaking effects, it is the most precise possibility to access the ratio as no absolute measurement of any Ξ_b^- branching fraction is available. Therefore, the ratio of the branching fractions of both decays can be expressed via the $SU(3)$ symmetry

$$\frac{\mathcal{B}(\Xi_b^- \to \Xi^- \mu^+ \mu^-)}{\mathcal{B}(\Lambda_b^0 \to \Lambda \mu^+ \mu^-)} = \frac{3}{2} \times \frac{\tau_{\Xi_b^-}}{\tau_{\Lambda_b^0}},$$

where τ_i is the respective lifetime of the b baryon. Therefore, one can express the yield of the total yield arising from both Ξ_b^- and Ξ_b^0 decays via

$$N(\Xi_b \to \Xi \mu^+ \mu^-) \sim \frac{3}{2} \cdot \frac{\tau_{\Xi_b^-}}{\tau_{\Lambda_b^0}} \cdot \frac{f_{\Xi_b^-}}{f_{\Lambda_b^0}} \times 2 \cdot N(\Lambda_b^0 \to \Lambda \mu^+ \mu^-) \cdot \frac{\epsilon(\Xi_b)}{\epsilon(\Lambda_b^0)}.$$

Under the assumption that the Ξ_b^0 background contributes in the same way, an additional factor 2 is added relying on the presumption that both Ξ_b production fractions are approximately equal. Therefore, an rough estimate of the yield is given by

$$\begin{aligned} N(\Xi_b \to \Xi \mu^+ \mu^-) &\sim 1.602 \times 0.082 \times 2 \cdot N(\Lambda_b^0 \to \Lambda \mu^+ \mu^-) \cdot \frac{\epsilon(\Xi_b)}{\epsilon(\Lambda_b^0)} \\ &\approx 315 \cdot \frac{\epsilon(\Xi_b)}{\epsilon(\Lambda_b^0)} \end{aligned}$$

using an approximate $N(\Lambda_b^0 \to \Lambda \mu^+ \mu^-)$ yield of 1200 events, the PDG [197] lifetimes of the Ξ_b^- and Λ_b^0, 1.57 ± 0.04 ps and 1.47 ± 0.01 ps, respectively, and the LHCb measurement of $f_{\Xi_b^-}/f_{\Lambda_b^0} \approx 0.082$ [66]. The efficiencies can then be estimated in the same manner as for the $B^0 \to K_S^0 \mu^+ \mu^-$ background in order to get an estimation on the possible contamination. Including the estimate of the efficiency ratio, the remaining number of backgrounds should be of the order of 2% and is therefore expected to be negligible.
However, it should be noted that an uncertainty of 30% needs to be assigned which is the typical size of $SU(3)$-breaking effects and additional 2.6% from the lifetime uncertainties. Since the Λ baryon arises as a product of the hyperon Ξ decay for Ξ_b decays, it is important to note that only about 10% of the Λ decay within the VELO, therefore this background will effect the downstream track category much more. Considering the large uncertainty on this estimate, the possible contribution of the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay will be covered by a dedicated systematic uncertainty. Considering the review state of the

first search of the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay, discussed earlier in Ch. 6, it might be possible to include the measured branching fraction for the estimation instead of relying on the $SU(3)$ symmetry assumption.
In addition, several excited states are known for the Λ_b^0 baryon: the $\Lambda_b^0(5912)$ and $\Lambda_b^0(5920)$, which both can decay to $\Lambda_b^0 \pi^+ \pi^-$ [233] and $\Lambda_b^0 \gamma$ final states. The latter decay mode assumes a possibility of Dalitz decays (or photon conversions) to $\Lambda_b^0 \mu^+ \mu^-$, creating a potential possibility of a swap between a bachelor lepton and a lepton from the Λ_b^0 decay. However, given that these decay modes are expected to have tiny branching fractions, together with the small production of excited Λ_b^0 states, this family of backgrounds is not considered in this analysis as well.
To conclude, it is expected that only the combinatorial background pollutes the signal mass window.

8.2.4 BDT selection against combinatorial background

The combinatorial background can be suppressed by using a BDT with gradient boosting. Its output provides a new variable that is then used as part of the selection requirements. In order to train the classifier a signal and background sample are provided. For the signal proxy truthmatched and calibrated simulation samples of the $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ decay, as described in the previous section, are used. The data sidebands outside of the signal region are utilised as signal proxy. In order to increase the statistics of the background proxy, the compatibility of the lower ($m(\Lambda \mu^+ \mu^-) < 5520\,\text{MeV}/c^2$) and upper ($m(\Lambda \mu^+ \mu^-) > 5720\,\text{MeV}/c^2$) sideband have been tested in various variables before using both of them in the training of the BDT classifier. Both the signal and background proxy use samples obtained and simulated under the Run 2 conditions. Additionally, the selection cuts other than the selected mass range are aligned between both samples.
One BDT is trained for both track type categories combined, but the output has been tested individually to ensure the selection efficiency is compatible for both categories. The variables (features) used in the training have been selected from a large set of variables removing those which have a low separation power or large correlations. The final set of variables includes the following 12 ones:

- $\chi^2_{\text{DTF}}(\Lambda_b^0)$
- $\texttt{DIRA}(\Lambda_b^0)$
- $\tau(\Lambda_b^0)$
- $\chi^2_{\tau}(\Lambda_b^0)$
- $\chi^2_{EVX}(\Lambda_b^0)$
- $p_{\text{T}}\ (\Lambda_b^0)$
- $\tau(\Lambda)$
- $p_{\text{T}}\ (\Lambda)$
- $\chi^2_{\text{IP}}(J/\psi)$
- $\min(\chi^2_{\text{IP}}(\mu^+, \mu^-))$
- $\min(\chi^2_{\text{IP}}(p, \pi^-))$
- $p(p)$

which mostly consist of variables from the Λ_b^0 candidate. This is a consequence of the fact that the combinatorial background consists of true Λ baryon candidates and, therefore, the separation power of variables from the Λ, π^- and proton is rather low. Additionally, some of these variables have significant correlations between themselves in the signal sample. But for the background sample these correlations are significantly smaller, consequently, the removal of one of these variables degrades significantly the BDT performance. After the feature selection, the optimal working point of the BDT training has been determined with a grid search for 250 trees with a depth of 5 and a learning rate of 0.17. More information about the general procedure for the determination of the optimal BDT are

given in App. A.2.
The classifier is checked for overtraining with Kolmogorov-Smirnov (KS) test values [217]. The implementation of the KS test enables to identify signs of overtraining if the test values get very close to zero (below 0.05). The left plot in Fig. 8.10 shows an example comparing the training and testing distributions, where no sign of overtraining is seen, neither visually, nor by means of the KS test.
In order to ensure the independence of the classifier performance from the choice of the training sample, a 10-fold cross-validation is adopted for the BDT training. Nine of the ten folds datasets are used to train the BDT, which afterwards tested on the remaining dataset. This procedure is repeated till every subset is used as test sample once. The resulting ten receiver operational characteristic (ROC) curves are shown in the right plot of Fig. 8.10, which show a very good agreement.

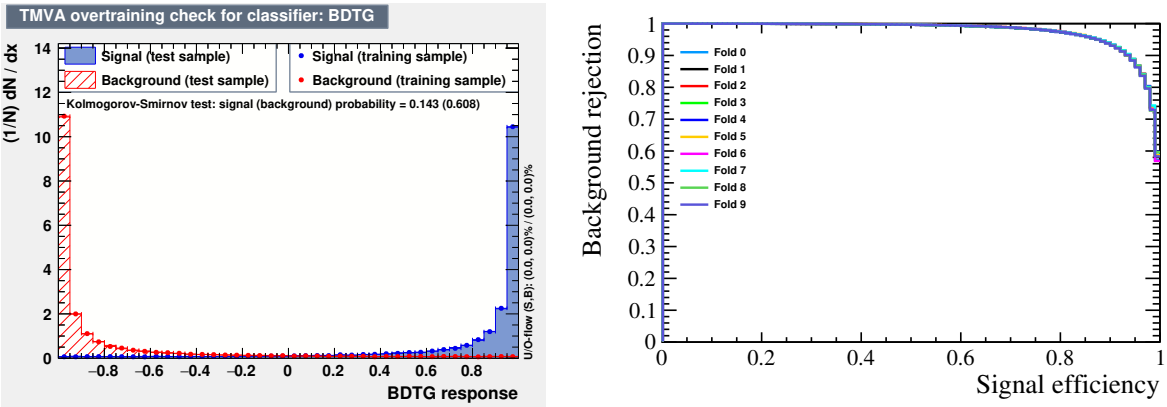


Figure 8.10: Comparison of a training and test sample distribution for the described BDT setup (left), showing no signs of overtraining. The ten ROC curves from the 10-fold cross-validation (right).

Additional information such as the BDT input variables distributions for the signal and background sample, their correlations and their importance for the BDT classifier can be found in App. G.1.
As a last step of the validation, the distribution of the BDT output has been compared between the sweighted data and the truthmatched and calibrated simulation (Sec. 8.2.2). This is done to ensure that no discrepancies between data and simulation have been picked up in the training of the BDT classifier resulting in a bias. This test is particularly important when using a large number of variables in the training, as in the presented analysis. The comparison of the BDT output distributions are shown for the Run 1 (left) and Run 2 (right) samples of the $\Lambda_b^0 \to \Lambda J/\psi$ decay in Fig. 8.11. It can be seen that add the edges of the classifier output there are some clear deviations between the distributions from data and from simulation. This is a sign of remaining mismodelling, that has not been corrected properly. Indeed, after revisiting the distribution of the BDT input variables, some residual residual disagreement between the data and simulation was found. This is in particular the case for the momentum of the Λ baryon. This finding highlights the importance of finalising the calibration before tuning the BDT. Additional, the observed mismodelling hints for a need of a separate calibration for the two track type categories depending if the Λ decayed within the VELO detector or not. This assumption is currently under investigation and will be followed by a test, if also two separate BDTs

will be needed for the two track type categories. However, for the demonstrative purpose of this thesis, the deviation will be set a side in order to demonstrate the developed analysis strategy.

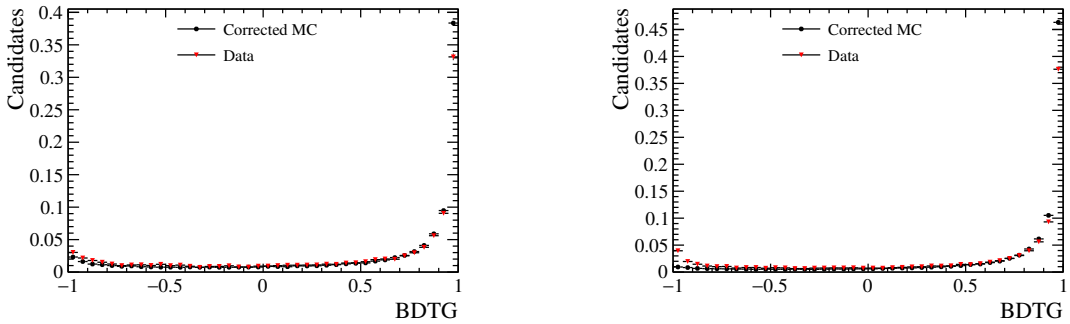


Figure 8.11: Comparison of the BDT output variable between the $\Lambda_b^0 \to \Lambda J/\psi$ data (red triangles) and simulation (black circles) for Run 1 (left) and Run 2 (right).

The cut value on the BDT classifier output is optimised in order to achieve the optimal sensitivity to the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay mode. The value is determined using a figure of merit (FOM) defined as $P = \frac{S}{\sqrt{S+B}}$, where S is the signal yield, and B is the background yield estimated from a fit to data at each cut value on the BDT output. As proven in Fig. 8.3, the signal peak of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay is already visible after the pre-selection. Therefore, the signal yield in the FOM is defined by $S = \epsilon_{\text{BDT}} \cdot S_0$. S_0 is the signal yield obtained from the fit to data prior to any BDT cut and the efficiency of each BDT cut ϵ_{BDT} is obtained from simulation at each BDT cut value. The cut optimisation is performed independently for the Run 1 and Run 2 dataset and the resulting FOMs are displayed in Fig. 8.12. The optimal BDT cut value according to the FOM is at > 0.45 for the Run 1 and > 0.6 for the Run 2 dataset. Due to the closeness of both optimal cut values it is decided to use the final BDT cut at > 0.55, when combining both datasets. This working point is chosen in order to keep as much signal as possible while still have a not too large background. The somewhat larger background level compared to a first measurement of decay channel is beneficial, as it helps with the fit stability of the angular fit discussed in the next section.
The peaks visible at certain cut values are nonphysical and arise from failed fits during the background estimation. The same BDT cut value is used for fits to other q^2 regions, *e.g.* when performing cross-checks with the J/ψ decay mode.

8.2.5 Summary of the signal selection

Once the entire selection is applied, as a last step, candidates sharing the same event and run numbers are identified. These candidates are called *multiple candidates* and are usually a result of reconstruction errors and not due to the presence of two candidates produced in the same event. This is because it is unlikely that the same rare decay happens for both b-hadrons produced in a given collision. Therefore, after finding such candidates, only one is retained randomly for this analysis, and all the others are removed. This procedure is only needed for the collision data, since the multiple candidates in simulation are removed by the truthmatching procedure. For the Run 1 (2) dataset about

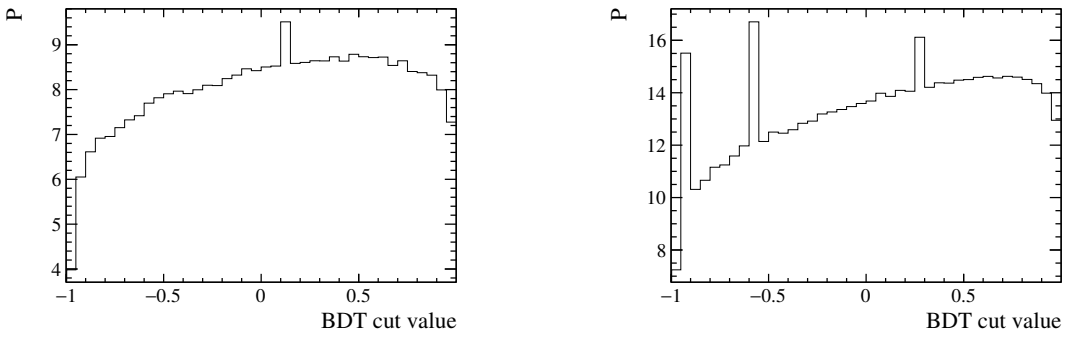


Figure 8.12: BDT output cut value optimisation: distributions of the FOM values as a function of the cut value. Left for the Run 1 and right Run 2 sample. Only central values are shown, without uncertainties. It should be noted that the peaks are non-physical and arise from failed fits during the calculation of the FOM.

0.3 (0.1) percent of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ candidates have been removed. The summary of the selection discussed in this section can be found in Tab. 8.2.

Table 8.2: Offline selection requirements for the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ angular analysis applied on top of the stripping and trigger selection.

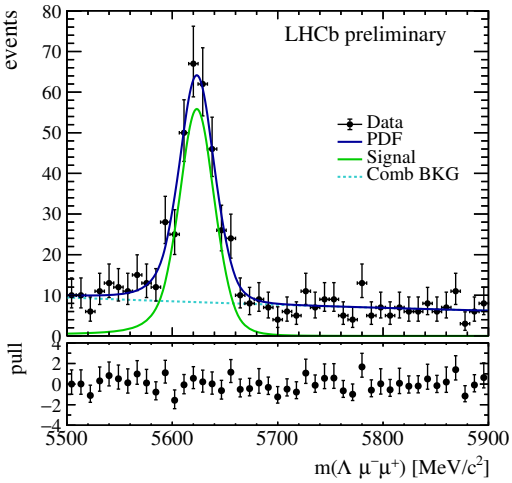
Type	Particle(s)	Requirement
Quality and acceptance	μ	$p > 3\,\text{GeV}/c$
		$p_T > 800$ MeV/c
		`InAccMuon, hasRich`
PID	μ	`ProbNNmu>0.1`
Mass windows	$\mu^+\mu^-$	$15 < q^2 < 20.28\,\text{GeV}^2/c^4$
	J/ψ	$3000 < \text{m}(\mu^+\mu^-) < 3150$ MeV/c^2
	Λ	$\lvert m(p\pi^-) - 1115.68\rvert < 8$ MeV/c^2
Background rejection	true Λ sel.	$0 < \tau(\Lambda) < 2\,\text{ps}$
		$\chi^2_{\text{FD}}(\Lambda) > 0$
		$0 < \text{EVX}_Z(\Lambda) < 2350\,\text{mm}$
		$\texttt{DIRA}_{PV}(\Lambda) > 0$
	Combinatorial	BDTG> 0.55

8.2.6 Λ_b^0 mass fits

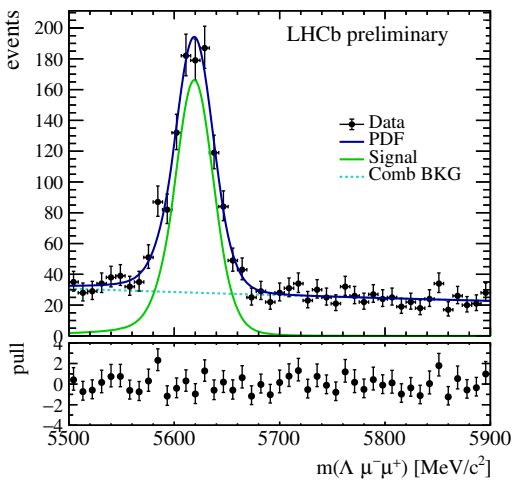
As discussed in the Sec. 4.4, the signal peak of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay is described by a linear combination of two double-sided Crystal Ball (dsCB) functions. The two PDFs share the same mean, while the tail parameters are allowed to be different. In addition, the signal width of the second component is expressed as a product of the first one and a fraction between the two: the first signal width is then kept floating in the fit to data. Before fitting the Λ_b^0 mass distribution with collision data, the tail parameters and the fractions of the widths of the two dsCB functions are fixed to the values obtained from simulation. An example fit to the Run 2 simulation can be found in Fig G.3 in App. G.2. As a result of the performed background studies, the nominal fit setup is finalised by including a fit component describing the combinatorial background. The background is modelled by an exponential distribution with a floating slope.
The fitted invariant mass shapes for the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ mode are shown in Fig. 8.13 for Run 1 in the top right, Run 2 in the top left and both combined in the bottom row. It can be seen that the pulls are rather stable around zero and show no indication of missed peaking background contributions in the chosen signal window due the absence of large fluctuations in the pulls. Finally, the obtained signal and background yields, as well as the signal resolution and the obtained mass values are shown in Tab. 8.3. Both the individual Run 1 and Run 2 values are in agreement with the current world average of the Λ_b^0 baryon mass, however the Run 1 result has a slightly larger shift. Additionally, it can be seen that the assumed signal yield used to evaluate the contributions from the partially reconstructed Ξ_b background and the misidentification background from the $B^0 \to K_S^0\mu^+\mu^-$ decay perfectly aligns with the obtained signal yield from the invariant mass fit.

Table 8.3: The obtained signal and background yields as well as the signal resolution and returned Λ_b^0 mass values from the mass fits to the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ datasets of Run 1 and Run 2, as well as both combined. The shown uncertainties are statistical only.

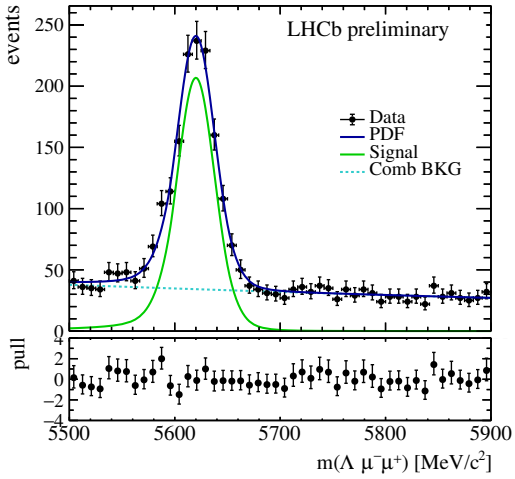
Dataset	Signal yield	BKG yield	Λ_b^0 mass, MeV/c^2	width, MeV/c^2
Run 1	285 ± 23	347 ± 25	5623.3 ± 1.5	12.5 ± 1.2
Run 2	924 ± 42	1194 ± 45	5619.0 ± 0.9	16.0 ± 0.8
Run 1 + 2	1208 ± 48	1543 ± 51	5620.1 ± 0.8	16.1 ± 0.7



(a) Run 1



(b) Run 2



(c) Run 1 + 2

Figure 8.13: Invariant mass fits of the fully selected $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ datasets for Run 1 (top left) and Run 2 (top right), as well as the combined dataset (bottom).

8.3 Towards the angular fits

While the unfinished calibration requires to update the selection and mass fits again, this section is dedicated to the angular analysis part of the presented analysis of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay. Consequently, the presented studies are of preliminary nature and performed under the assumption that the results from the previous section are final. However, the studies are not expected to change significantly and are assumed to largely align with the final approach.
In the following subsection the analysis strategy, the acceptance, background modelling and the angular fit model validation will be discussed.

8.3.1 Analysis strategy

In order to extract the nine angular coefficients[4] from the CP averaged angular distributions in Eqn. (8.12), introduced in Sec. 8.1.2, a four-dimensional simultaneous unbinned maximum likelihood fit will be performed to the invariant mass $m(\Lambda\mu^+\mu^-)$ and the three angles $\cos(\theta_l), \cos(\theta_\Lambda)$ and ϕ. The full fit to data is performed in the mass window $m(\Lambda\mu^+\mu^-) \in [5500, 5900]\,\text{MeV}/c^2$ which enables to determine which of the events are signal and which are background from the mass fit, which is crucial to correctly fit the individual angles. In addition, the larger mass window enables to increase the fit stability. While the invariant mass distribution has been looked at, in order to ensure that the signal and background are correctly modelled, the angular distributions of $\cos(\theta_l), \cos(\theta_\Lambda)$ and ϕ are blinded and the fit procedure is validated with pseudoexperiments.
The full fit function is obtained as the sum of the signal and the combinatorial background component, as this is the only remaining background. The fraction of the latter is obtained directly from the fit by fitting the invariant mass $m(\Lambda\mu^+\mu^-)$. The total PDF is, therefore, defined by:

$$\begin{aligned}\mathcal{PDF}_{\text{tot}} = f_{\text{sig}} \times \mathcal{F}_{\text{sig}}(m) \times \mathcal{A}_{\text{sig}}(\Omega) \\ + (1 - f_{\text{sig}}) \times \mathcal{F}_{\text{comb}}(m) \times \mathcal{A}_{\text{comb}}(\Omega),\end{aligned} \tag{8.19}$$

where $\mathcal{F}(m)$ and $\mathcal{A}(\Omega)$ refer to the mass and angular fit components, respectively. In this case Ω refers to the three angles $\Omega = \cos(\theta_l), \cos(\theta_\Lambda)$ and ϕ.
As previously discussed, the analysis will be performed using the full Run 1 and Run 2 LHCb dataset and focuses solely on the high q^2 region above the $\psi(2S)$ peak.
Lastly, the angular distributions of the background are studied with data sidebands of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ dataset, blinding a 3σ mass window around the known Λ_b^0 baryon mass. This is done in order to obtain an estimate of the angular shapes in the background and increase the fit stability of the four-dimensional likelihood fit.
It should be noted that the datasets for decay mode $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ have not been unblinded for the years 2017+2018 prior to this analysis. However, it was decided to not blind the mass fits to full data in order to ensure a correct estimation of the background contributions and control of the latter. In addition, the Run 1, 2015 and 2016 datasets have been used in the previous observation of the decay and the preformed angular analyses.

[4]The previously mentioned ten coefficients can be reduced to nine by utilising the normalisation discussed in Sec. 8.1.2.

As of the moment of writing, the angular distributions and the coefficients have not been unblinded due to the pending LHCb internal review of the analysis strategy.
One crucial pre-unblinding cross-check, discussed in this thesis, is the validation of the fitting procedure with the use of pseudoexperiments and will be covered in Sec. 8.3.4. Furthermore, several other cross-checks are currently in preparation, but will only be briefly mentioned in the summary 8.5 of this chapter.

8.3.2 Angular acceptance

Using the Eqn. (8.12) in Sec. 8.1.2 and the normalisation constraint, the full angular signal PDF can be expressed as:

$$\begin{aligned}\mathcal{A}_{sig}(\Omega) = \epsilon(\Omega) \times [&\Big(K_{1ss} + (1 - 3K_{1ss})\cos^2(\theta_\ell) + K_{1c}\cos(\theta_\ell)\Big) \\ &+ (K_{2ss} + (K_{2cc} - K_{2ss})\cos^2(\theta_\ell) + K_{2c}\cos(\theta_\ell))\cos(\theta_\Lambda) \\ &+ (K_{3sc}\sqrt{1-\cos^2(\theta_\ell)}\cos(\theta_\ell) + K_{3s}\sqrt{1-\cos^2(\theta_\ell)})\sqrt{1-\cos^2(\theta_\Lambda)}\sin\phi \\ &+ (K_{4sc}\sqrt{1-\cos^2(\theta_\ell)}\cos(\theta_\ell) + K_{4s}\sqrt{1-\cos^2(\theta_\ell)})\sqrt{1-\cos^2(\theta_\Lambda)}\cos\phi],\end{aligned} \tag{8.20}$$

with $\epsilon(\Omega)$ being the acceptance function. Since phase-space simulation is usually produced flat within ϕ and the cosine of the other two helicity angles, it enables to study the acceptance effects of the geometry of the LHCb detector, along with the processes of reconstruction and selection, on the three angles considered: $\cos(\theta_\ell)$, $\cos(\theta_\Lambda)$, and ϕ. Therefore, phase-space simulation samples of the signal mode $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ are processed with the entire signal selection described in Sec. 8.2.5 and the correction weights of Sec. 8.2.2 applied. The resulting distributions give access to the angular acceptance of the three angles $\epsilon(\Omega)$.
It is assumed that there is no significant variation of the efficiency in a given q^2 bin and that the acceptance factorises:

$$\epsilon(\Omega) = \epsilon_\ell(\cos(\theta_\ell)) \times \epsilon_\Lambda(\cos(\theta_\Lambda)) \times \epsilon_\phi(\phi). \tag{8.21}$$

Legendre polynomials are utilised to model the individual acceptances, as the fitted coefficients have only small correlations. Concerning the $\cos(\theta_\ell)$ and $\cos(\theta_\Lambda)$ angles, the acceptance is modelled using Legendre polynomials of the order two. Due to the asymmetric cuts on the proton and pion in the stripping selection an asymmetric shape of the $\cos(\theta_\Lambda)$ distribution is expected. Therefore, the distribution is modelled using even and odd coefficients:

$$\epsilon(x = \cos(\theta_\Lambda)) = 1 + c_{1\Lambda}x + \frac{c_{2\Lambda}}{2}(3x^2 - 1). \tag{8.22}$$

For the $\cos(\theta_\ell)$ distribution no such asymmetry is expected and a Legendre polynomial of the order two with only even coefficients,

$$\epsilon(x = \cos(\theta_\ell)) + 1 + \frac{c_{2\ell}}{2}(3x^2 - 1), \tag{8.23}$$

sufficiently describes the distribution. The two coefficients for $\cos(\theta_\Lambda)$ and the one for $\cos(\theta_\ell)$ are found to be not compatible with zero, higher orders have been tested but

no non-zero coefficients have been found. Finally, the ϕ angle is expected to be rather unaffected by the reconstruction and selection effects. However, it is known that the ϕ acceptance has a modulation at large q^2 values above the J/ψ resonance. As discussed in Sec. 8.1.1, the angle ϕ is defined between the dilepton and the Λ plane in the Λ_b^0 rest-frame rather than with respect to single particles. If the angle ϕ between the two planes in the Λ_b^0 rest-frame is small ($\phi \to 0, \pm\pi$) then the corresponding angle between the two planes will also tend to be small in the lab-frame. Consequently, the azimuthal angle of the proton and the muon will be very similar in the lab-frame. Therefore, if $\phi \sim \pm\frac{\pi}{2}$ all four particles will have large differences in the lab-frame. If this is combined with an azimuthal angle dependence of the reconstruction efficiency in the lab-frame, one of the four particles will most likely end up in the $\phi \sim \pm\frac{\pi}{2}$ region leading to an acceptance effect in ϕ. Therefore, the flatness of the ϕ angular acceptance is tested against $\cos(2\phi), \sin(2\phi), \cos(\phi)$ and $\sin(\phi)$ effects. The total ϕ angular acceptance PDF is given by

$$\epsilon(\phi) = 1 + c_{c2}\cos(2\phi) + c_{s2}\sin(2\phi) + c_c\cos(\phi) + c_s\sin(\phi). \tag{8.24}$$

The fits for all three angles $\cos(\theta_\Lambda)$, $\cos(\theta_\ell)$ and ϕ (left to right) are shown in Fig. 8.14 for Run 1 (top), Run 2 (middle) and both data taking periods merged (bottom). For the ϕ acceptance in all three samples, three of the four parameters are found to be compatible with zero. Only c_{c2} is found to be non-zero, and therefore, the acceptance function in ϕ is modelled by $\epsilon(\phi) = 1 + \cos(2\phi)$.

Factorisation of the acceptance

For simplicity of the procedure, the angular acceptances of each angle have been determined independently under the assumption of factorisation. It is therefore important to check the presumption in two steps. Since the most prominent effects on the individual acceptances arise from $\cos(\theta_\ell)$ and $\cos(\theta_\Lambda)$, it is checked if both angular acceptances factorise, i.e. $\epsilon(\cos(\theta_\ell), \cos(\theta_\Lambda)) = \epsilon(\cos(\theta_\ell)) \cdot \epsilon(\cos(\theta_\Lambda))$. In a second step, it is then checked if the ϕ angular acceptance factorises with respect to $\cos(\theta_\ell)$ or $\cos(\theta_\Lambda)$.
To check the factorisation of angular acceptances $\cos(\theta_\ell)$ and $\cos(\theta_\Lambda)$, the flat generated distributions of the phase-space $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ simulation are attempted to be recovered by applying the inverse of the fitted acceptance as a weight, *e.g.* $(\epsilon(\cos(\theta_\ell)) \cdot \epsilon(\cos(\theta_\Lambda)))^{-1}$. Any deviation from a flat distribution in $\cos(\theta_\ell)$ or $\cos(\theta_\Lambda)$ would indicate a correlation in the angular acceptances and thus would prevent their factorisation. The recovered $\cos(\theta_\ell)$ and $\cos(\theta_\Lambda)$ distributions, using a simple $(\epsilon(\cos(\theta_\Lambda)))^{-1}$ or $(\epsilon(\cos(\theta_\ell)))^{-1}$ weight, are then tested for flatness via a fit to an order four Legendre polynomial. All parameters are found to be compatible with zero and the reweighted distributions using $(\epsilon(\cos(\theta_\ell)) \cdot \epsilon(\cos(\theta_\Lambda)))^{-1}$ weights have similar shapes as the reweighted distributions using $(\epsilon(\cos(\theta_\Lambda)))^{-1}$ or $(\epsilon(\cos(\theta_\ell)))^{-1}$ individually. Thus, no indication of strong correlation is found and the factorisation hypothesis is validated.
The factorization of the ϕ angular acceptance with respect to $\cos(\theta_\ell)$ and $\cos(\theta_\Lambda)$ is checked in a similar way. A weight $(\epsilon(\phi) \cdot \epsilon(\cos(\theta_\Lambda)))^{-1}$ or $(\epsilon(\phi) \cdot \epsilon(\cos(\theta_\ell)))^{-1}$ is applied to the phase-space $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ simulation. If the angular acceptances are uncorrelated, the reweighted ϕ is expected to be similar to the weighted with $(\epsilon(\phi))^{-1}$ distribution. A fit with the function $1 + c_{c2}\cos(2\phi) + c_{s2}\sin(2\phi) + c_c\cos(\phi) + c_s\sin(\phi)$ is performed. Again, all the fit parameters are compatible with the ones using a simple $(\epsilon(\phi))^{-1}$ weight, thus

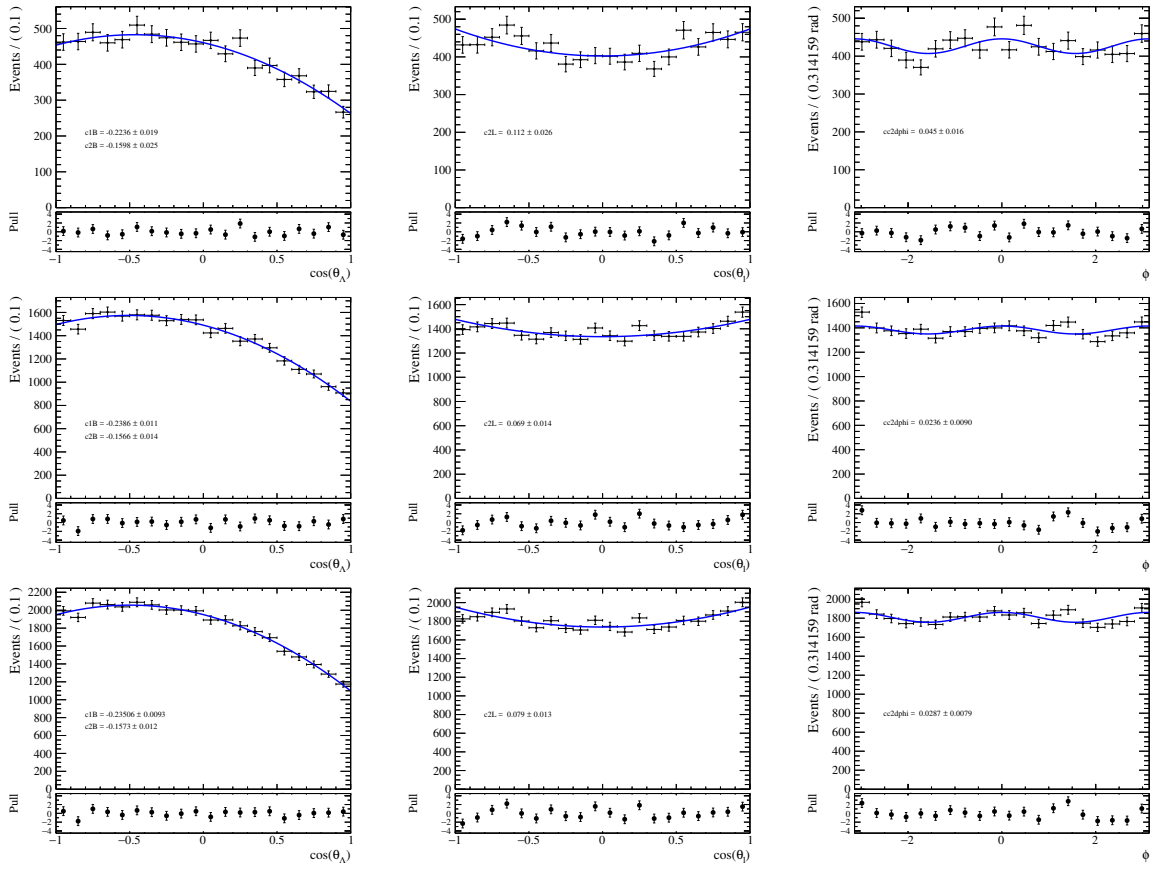


Figure 8.14: Legendre polynomial fits to the angle distribution of the phase-space simulation to extract the angular acceptance for the Run 1 (top row), Run 2 (middle row) and combined dataset (bottom) row. The $\cos(\theta_\Lambda)$ distribution is on the left, the $\cos(\theta_\ell)$ distribution is in the middle and he $\cos(\theta_\Lambda)$ distribution is on the right. The full selection and the discussed calibration maps are applied. It should be noted that the coefficients for the $\cos(\theta_\Lambda)$ fits are labelled with a B to distinguish them more easily from the $\cos(\theta_\ell)$ ones.

validating the factorisation hypothesis.
Fig 8.15 finally shows, the distributions of the angular acceptance as determined in the beginning of this section on the left and the obtained flat distributions for each of the three angles $\cos(\theta_\ell)$ (top), $\cos(\theta_\Lambda)$ (middle) and ϕ (bottom) after a reweighting with the inverse of the efficiencies $(\epsilon(\cos(\theta_\ell)) \cdot \epsilon(\cos(\theta_\Lambda)) \cdot \epsilon(\phi))^{-1}$. This test was done as a last step to successfully confirm the factorisation of all three angular acceptances using the full Run 1 and Run 2 simulation samples.

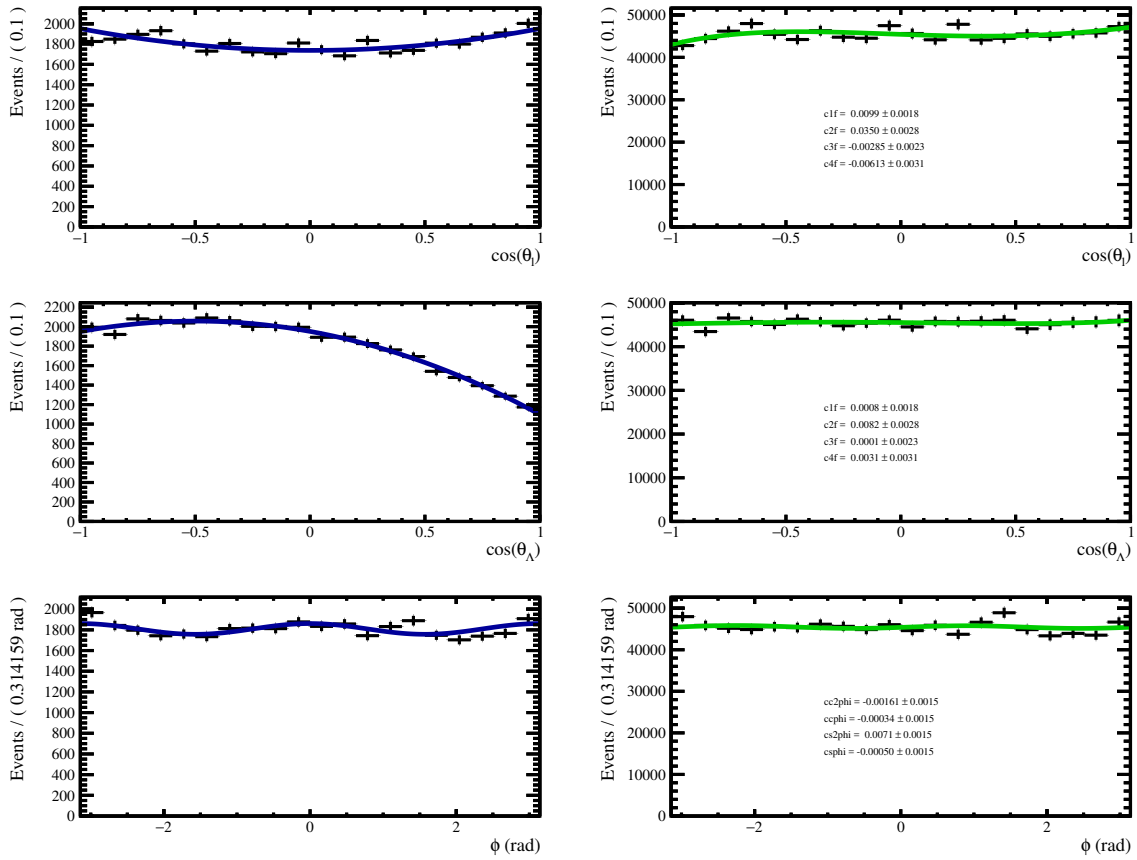


Figure 8.15: Left: Angular acceptance (blue) obtained from the calibrated phase-space simulation for the Run 1 and Run 2 sample combined. Right: The same angular distributions reweighted (green) with the inverse of the determined acceptance shapes in order to re-obtain the initially produced flat angular distributions.

8.3.3 Combinatorial background

As previously mentioned, in the mass fit the combinatorial background is modelled with an exponential distribution and its slope is kept free in the fit. A more delicate situation arises for the correct description of the background contribution in the angular distributions. The goal is to benefit from a wider fit range and, consequently, larger background level, which enables to leave the fit components from the background floating in the final four-dimensional simultaneous likelihood fit. However, it turns out that the composition of the combinatorial background with real Λ hyperons poses certain challenges, which will be discussed further in Sec. 8.3.4. In dedicated cross-checks all angular parameters of the background are constrained to the values obtained from the fit to the data sidebands, in order to study how observed biases are related to the background component and the fit stability. Afterwards, the background components will all left floating in the final fit to the data. Thus, to obtain the angular distributions of the background, the data sidebands ($5500 < m(\Lambda\mu^+\mu^-) < 5520\,\text{MeV}/c^2$ and $5720 < m(\Lambda\mu^+\mu^-) < 5900\,\text{MeV}/c^2$), are combined and fitted with Legendre polynomials, similarly to the procedure used in angular acceptance determination. The values of the coefficients of the functions, determined in these fits, are used for the background component to generate realistic pseudoexperiments. Additionally, the obtained fit function will be used for the background component in the final fit model.
The $\cos(\theta_\ell)$ distribution is fitted with an order two Legendre polynomial with only the even coefficient $1 + \frac{c_{2\ell}}{2}(3x^2 - 1)$ and the $\cos(\theta_\Lambda)$ distribution with an order three Legendre polynomial $1 + c_{1\Lambda}x + \frac{c_{2\Lambda}}{2}(3x^2 - 1) + \frac{c_{3\Lambda}}{2}(5x^3 - 3x)$. The ϕ distribution is fitted with the function $1 + c_c\cos(\phi) + c_{cc}\cos 2\phi + c_s\sin(\phi) + c_{ss}\sin(2\phi)$. The full nominal selection is applied to the data sidebands and the resulting fits for Run 1 (top), Run 2 (middle) and both combined (bottom) are shown in Fig. 8.16. For the $\cos(\theta_\ell)$ and ϕ distributions, the coefficients are compatible with zero and, therefore, both distributions are assumed to be flat for the final fit, both in pseudoexperiment studies and the final fit to data. The asymmetric shape in the $\cos(\theta_\Lambda)$ distribution is caused by the asymmetric selection requirements on the proton and pion in the stripping selection and the fact that the cut-based preselection selects real Λ candidates. The obtained values for the three coefficients for the fit to $\cos(\theta_\Lambda)$ distribution will be used to generate the pseudoexperiments and fixed to the obtained values in the some tests for the validation of the fit model. This is done to investigate the effects on the obtained values of the angular coefficients due to the similarity between the shapes of the signal and background components in their $\cos(\theta_\Lambda)$ distribution.
Additionally, the compatibility of the background shapes between the two track type categories was checked and found to be within 2 standard deviations for the splitting of the Run 1, Run 2, as well as the combined dataset. Furthermore, a broader mass range, especially for the lower sideband, has been tested and found to provide compatible results with the upper sideband mass shapes. This test was done to ensure that no significant physics backgrounds, such as partially reconstructed ones, remain after the selection, since they only pollute the lower mass sideband. The obtained values of the coefficients from the fitted functions of the different subsets are in agreement within 1σ.

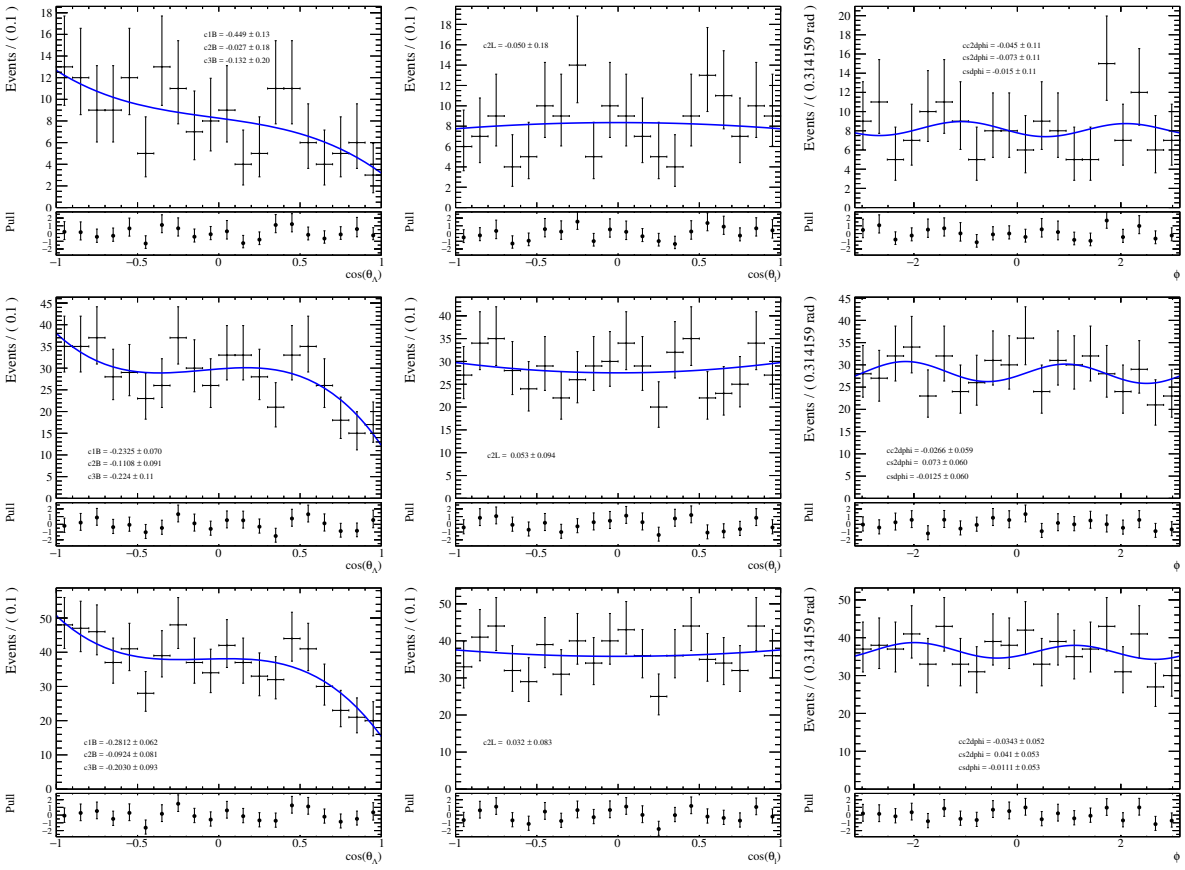


Figure 8.16: Legendre polynomial fits to the angular distributions of the data sidebands in the invariant mass fit range for the Run 1 (top row), Run 2 (middle row) and combined dataset (bottom row). The $\cos(\theta_\Lambda)$ distribution is on the left, the $\cos(\theta_\ell)$ distribution is in the middle and he $\cos(\theta_\Lambda)$ distribution is on the right. It should be noted that the coefficients for the $\cos(\theta_\Lambda)$ fits are labelled with a B to distinguish them more easily from the $\cos(\theta_\ell)$ ones.

8.3.4 Fit model validation

The fit model based on the discussed PDF in Eqn. (8.19) can be tested by performing a set of pseudoexperiments, simulating data with the same statistical properties as the actual dataset. In order to do that the determined shapes from the mass fits, the angular acceptance and background shapes for the angles are combined. In these pseudoexperiments, the values of the nine angular parameters are generated based on the SM values reported in Eqn. 8.14. The obtained dataset is then fitted with the full fit model and should again return the SM values. Any deviation needs to be investigated in order to understand the origin of the bias.
In order to validate the fitting procedure, several steps are performed with pseudoexperiments.

As a first step, signal-only pseudoexperiments are produced according to the angular distribution and fitted with the determined angular acceptance described in Sec. 8.3.2. This additional step is done to verify that no bias is introduced to the fit, when extracting nine coefficients from the fit. Therefore, 1000 pseudoexperiments with each a signal yield

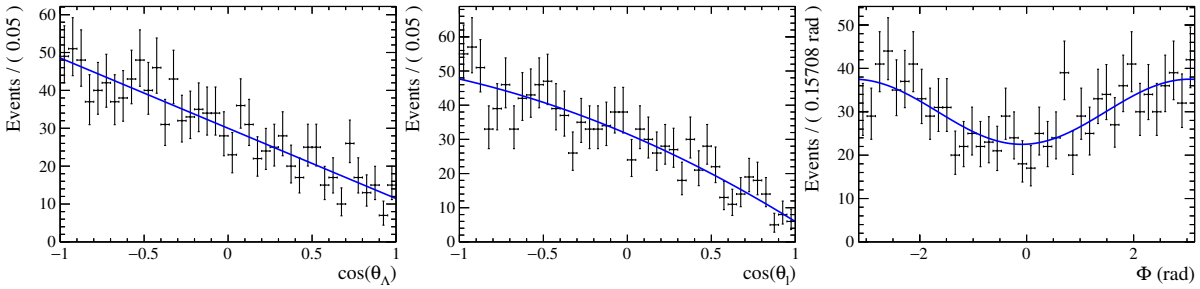


Figure 8.17: One random signal pseudoexperiment using the Run 1 and Run 2 combined values of the angular acceptance. The top to bottom are the fitted $\cos(\theta_\Lambda)$, $\cos(\theta_\ell)$ and ϕ distributions.

of 1200 events are produced and the fit for one random example is displayed in Fig. 8.17. The values of the fitted angular coefficients

$$\begin{aligned} K_{1ss} &= 0.351 \pm 0.009 \\ K_{1c} &= -0.232 \pm 0.025 \\ K_{2ss} &= -0.214 \pm 0.021 \\ K_{2cc} &= -0.191 \pm 0.032 \\ K_{2c} &= 0.212 \pm 0.025 \\ K_{3sc} &= -0.018 \pm 0.039 \\ K_{3s} &= 0.012 \pm 0.019 \\ K_{4sc} &= -0.019 \pm 0.039 \\ K_{4s} &= -0.135 \pm 0.019 \end{aligned} \tag{8.25}$$

are all in agreement within one standard deviation with the initial values displayed in Eqn. (8.14).
Finally, the pulls of all nine individual angular coefficients are checked for any biases by fitting them with a Gaussian distribution. As it can be seen in Fig. 8.18, all nine coefficients have a mean consistent with zero and a standard deviation consistent with unity, being fully in agreement with a Gaussian distribution. Consequently, it is possible to extract all nine angular coefficients from 1200 signal events without introducing a bias from the signal side of the angular fit. After evaluating the signal distribution for the angular distributions, 1000 realistic pseudoexperiments including a proper background contribution are produced. This strategy enables to test the full PDF from Eqn. (8.19), as it will be later used in the fit to the real LHCb dataset collected during Run 1 and Run 2. As previously mentioned, the $\cos(\theta_\ell)$ and ϕ distribution of the background are assumed and therefore generated flat. In a first step, only the background component for the $\cos(\theta_\Lambda)$ background distribution are left floating, due to the flatness of the other two angular background components. In this setup, a bias is observed in the pull distributions of the angular coefficients of K_{1c} and K_{2c}. Therefore, in order to investigate the origin of the bias, the background angular parameters have all been fixed in the fit to the generated samples. Thus, besides the angular coefficients, only the background slope and the signal fraction are kept floating. One random example for a realistic pseudoexperiment with the signal and background components included can be found in Fig. 8.19.

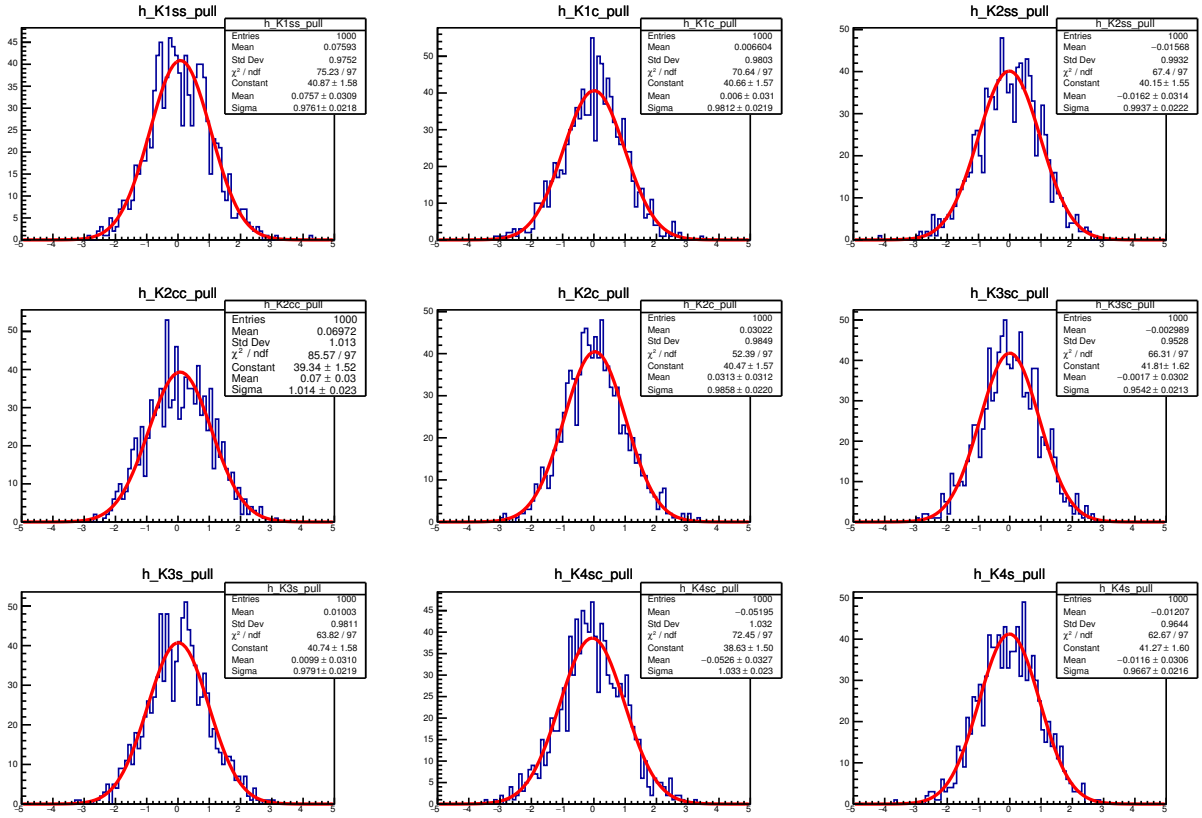


Figure 8.18: Pulls of all nine angular coefficients obtained from the 100 signal-only pseudoexperiments. No significant biases are found, according to the fitted Gaussian distributions (shown in red).

The fitted angular coefficients for this realistic random pseudoexperiment,

$$\begin{aligned} K_{1ss} &= 0.355 \pm 0.012 \\ K_{1c} &= -0.207 \pm 0.019 \\ K_{2ss} &= -0.224 \pm 0.028 \\ K_{2cc} &= -0.167 \pm 0.041 \\ K_{2c} &= 0.229 \pm 0.031 \\ K_{3sc} &= 0.023 \pm 0.052 \\ K_{3s} &= 0.034 \pm 0.025 \\ K_{4sc} &= -0.072 \pm 0.050 \\ K_{4s} &= -0.105 \pm 0.024 \end{aligned} \tag{8.26}$$

are all in agreement between one and two standard deviations with the initial values displayed in Eqn. (8.14).
However, it was found that the pseudoexperiments have a non-negligible rate of non-converged fits, which is currently under investigation. Furthermore, looking again at the pulls from all angular coefficients determined with the 1000 realistic pseudoexperiments, a bias remains in the distributions of K_{1c} and K_{2c}. When investigating Eqn. (8.12) in Sec. 8.1.2, it should be noted that these two angular coefficients are only distinguishable

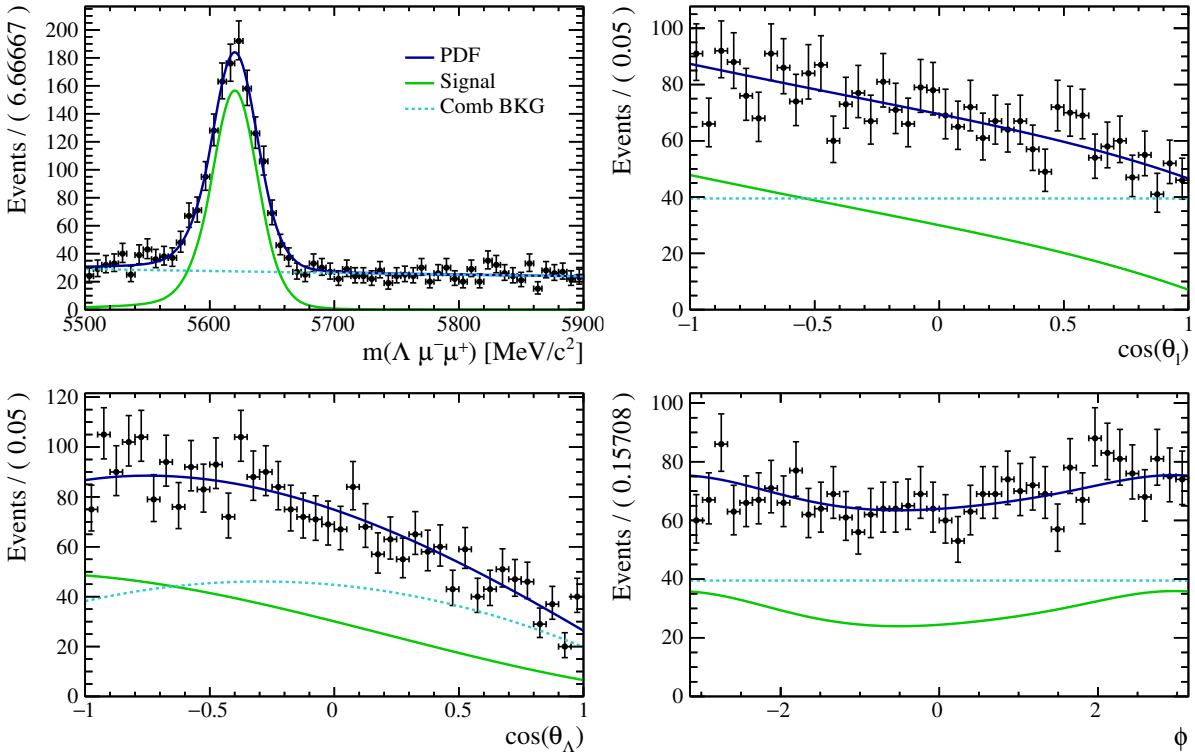


Figure 8.19: One random realistic pseudoexperiment using the Run 1 and Run 2 combined values of the angular acceptance and background estimates for the angular shapes. The $m(\Lambda\mu^+\mu^-)$ invariant mass distribution is shown in the top left, the $\cos(\theta_\ell)$ one in the top right, the $\cos(\theta_\Lambda)$ one in the bottom left and the ϕ distribution in the bottom right. The signal contribution is displayed with a solid green line, the background contribution by a dashed light blue line and the combined fit combining both components by a dark blue solid line.

by their $\cos(\theta_\Lambda)$ distributions. As already discussed in this section, the shape of $\cos(\theta_\Lambda)$ is expected to be very similar between signal and background components due to selection that can only be passed by real Λ candidates. Consequently, with the available amount of statistics collected during Run 1 and Run 2, the fit struggles to distinguish K_{1c} and K_{2c} sufficiently. However, this bias can be corrected and is therefore not a problem for the measurement, when all angular distributions of the background component are left floating in the final fit to the collision data. The exact size of the bias will be investigated after finalising the calibrations of the simulation and updating the fit strategy.
Several tests have been performed in order to investigate this claim. As a first step, the sum of the angular coefficients K_{1c} and K_{2c} has been calculated, leading to absence of a significant remaining bias. This proves that the biases between both coefficients are linked. In addition, it was tested how using a broader invariant mass window affects the likelihood fit. It was concluded that it helps to increase the fit stability but the bias in K_{1c} and K_{2c} remained. As a last test, the yields used as an assumption in the generation of the pseudoexperiments have been increased by factor 3 and, similarly, the signal fraction has been set to 70%. In this case, the bias of K_{1c} and K_{2c} fully disappears and the

fit is very stable with a high convergence rate. Thus, it is concluded that the observed bias is a result from the difficulty to distinguish the K_{1c} and K_{2c} angular coefficients in low-statistics regime, as it is the case in this analysis. The bias will be corrected for the final result and a dedicated systematic uncertainty assigned.
Additionally, further tests are planned as a part of the internal review process before unblinding the angular coefficients, which well help to evaluate the fit procedure. As a first step, the physics simulation samples used for the BDT training and determination of the signal shapes of the mass distribution can be used to validate the angular fit procedure. However, in order to do that, the generator-level counterparts of these $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ simulation samples need to be accessed and fitted as well. The obtained values for all angular coefficients for both the generator and reconstruction-level samples should agree with each other, as well as with the predicted values reported in Eqn. (8.14) in Sec. 8.1.
In addition, the data can be used in two separate ways to evaluate the discussed procedures. The first step is to test the compatibility with the Run 1 result reported in Ref. [7]. Therefore, the three-dimensional likelihood Eqn. (8.12) needs to be integrated over ϕ. It can then be expressed via the fraction of longitudinal polarisation F_0 and the three forward-backward asymmetries with respect to the leptonic scattering A_{FB}^{ℓ}, to the baryonic scattering angle $A_{\mathrm{FB}}^{\Lambda}$ and a combination of both asymmetries $A_{\mathrm{FB}}^{\ell\Lambda}$, which have been defined in Eqn. 8.9. The final two-dimensional likelihood is then defined by:

$$\begin{aligned}
&\frac{1}{\frac{\mathrm{d}\Gamma}{\mathrm{d}q^2} + \frac{\mathrm{d}\bar{\Gamma}}{\mathrm{d}q^2}} \left\langle \frac{\mathrm{d}^3(\Gamma + \bar{\Gamma})}{\mathrm{d}q^2 \mathrm{d}\cos\theta_\ell \mathrm{d}\cos\theta_\Lambda} \right\rangle_{CP\text{ averaged}} \\
&\qquad = \frac{3}{4}\Big[\frac{1}{4}(1+\cos^2\theta_\ell) + \frac{F_0}{4}(1-3\cos^2\theta_\ell) + \frac{2}{3}A_{\mathrm{FB}}^{\ell}\cos\theta_\ell \\
&\qquad + \frac{4}{3}A_{\mathrm{FB}}^{\ell\Lambda}\cos\theta_\ell\cos\theta_\Lambda + 2A_{\mathrm{FB}}^{\Lambda}\cos^2\theta_\ell\cos\theta_\Lambda \\
&\qquad + \frac{S_{2ss}}{2S_{1ss}+S_{1cc}}(1-3\cos^2\theta_\ell)\cos\theta_\Lambda\Big],
\end{aligned} \tag{8.27}$$

In a second approach, the full Run 1 and Run 2 dataset is fitted with the nominal fitting procedure, blinding only the angular distributions and the nominal values of the angular coefficients but not their uncertainties. This enables to test the convergence of the final fitting procedure by studying the pulls of the angular distributions. Furthermore, it is considered to perform an error test, where the uncertainties derived from the fit to data are compared to the distributions of uncertainties obtained from the pseudoexperiments with comparable yields.
These tests have so far not been performed, as the goal is to complete the proper calibration of the simulation samples and then re-evaluate the BDT classifier as well as the discussed studies in this section, in order to remove any unnecessary bias.

8.4 Systematic uncertainties

The final step of this analysis will be the determination of the systematic uncertainties. Rather than reporting the actual numbers, this section only reports the dedicated approaches considered. This is done, since meaningful uncertainties can only reported after the calibration and all following steps are finalised. In general, two different types of

systematic uncertainties need to be considered. The first class is the uncertainties due to the statistical power of the simulation and data samples that have been utilised to study the angular shapes and acceptance. The second class considers the ones related to fit model choices.
The systematic uncertainties related to the size of the simulation samples can be accessed with a bootstrapping technique [204]. As a part of this method *e.g.* 1000 resampled datasets can be created by replacing each event by a randomly chosen event from the same sample and resampling it. The 1000 resampled datasets are then used each individually to model the angular shape and acceptance each and redo the full fit. This procedure can be performed with pseudoexperiments, where the statistics have been increased in order to avoid fluctuation due to the available statistics in the Run 1 and Run 2 LHCb dataset. Similarly, the uncertainty related to the limited statistics of the datasets used to calibrated the simulation samples can be accessed.
The uncertainties related to the fit model can be split in two categories, those depending on the simulation samples and those obtained from data. For the uncertainty related to the simulation samples, they arise due to the calibration of the simulation and the choice of the function used to model the acceptance. The latter can be evaluated by choosing a higher order Legendre polynomial to model the acceptance. The uncertainty due to the correction can be determined by simply using the uncorrected simulation samples to access the acceptance and mass shapes. Additionally, it might be useful to choose a decay with larger statistical power compared to the $\Lambda_b^0 \to \Lambda J/\psi$ decay (such as $\Lambda_b^0 \to pK^- J/\psi$) to calibrate the kinematics of the Λ_b^0 candidate. This would enable to obtain a correction map for each data-taking year individually, allowing to catch possible difference between data-taking periods. Finally, the acceptance could be split into smaller q^2 bins in the high-q^2 region in order to catch deviations from the expected uniformity across the entire bin.
The last systematic uncertainties arise due to the choices made for the background modelling. As previously discussed, the misidentificaton background $B^0 \to K_S^0 \mu^+ \mu^-$ is expected to be negligible, but some hints for residual background remain. The related systematic uncertainty can be accessed by explicitly vetoing the background with a K_S^0 mass veto. The choice of the background modelling for the combinatorial background can be accessed with three different approaches: first, the invariant-mass window can be enlarged; the BDT cut can be tightened or relaxed; and, lastly, an alternative background samples from the same-sign $\Lambda \mu^\pm \mu^\pm$ candidates in data can be used to model the combinatorial background. Since no physics process is known to create the same-sign muon final state, the full dataset is of combinatorial nature. However, this dataset would lack certain types of combinatorial events, such as that created by combining products of two correlated b-hadron decays produced in the same event.

8.5 Summary

To summarise, this chapter reported the current status of the angular analysis of the $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ decay in the high-q^2 region above $15\,\mathrm{GeV}^2/c^4$ using the full Run 1 and Run 2 dataset collected by the LHCb experiment. The analysis is currently blind and in preparation for the internal review.
The default selection and the discussion of relevant physics backgrounds have been dis-

cussed. A dedicated mass fit model has been developed and tested, which is crucial in order to classify the events as signal and background in the angular distributions of the final four-dimensional maximum-likelihood fit. In addition, the current status of the calibration of the used simulation samples has been discussed. It should be noted that the calibration and, therefore, all related results are only preliminary, as the calibration strategy is not finalised yet. Nevertheless, the determination of the angular acceptance and its factorisation have been presented. Similarly, the angular fit procedure has been discussed, as well as its validation with the help of pseudoexperiments. This was crucial, as it uncovered the bias for the K_{1c} and K_{2c} angular coefficients. They can only be distinguished by the $\cos(\theta_\Lambda)$ distribution, which has very similar shapes for signal and background. Finally, the systematic uncertainties that need to be studied have been discussed.

As a summary, this chapter successfully demonstrated the full analysis strategy in order to extract all nine angular coefficients of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay for the first time using a four-dimensional simultaneous unbinned maximum-likelihood fit. The analysis will be updated shortly after the calibration is finalised in order to start the internal review process.

Chapter 9

Future aspects of beauty baryon decays

The measurements presented in this thesis highlight the physics reach of rare decays of weakly decaying beauty baryons. They offer unique possibilities to improve the knowledge of the baryons properties and help to clarify the nature of the B anomalies. But as the presented work has shown, this particular field of flavour physics is only starting due to the small production fractions of the different types of weakly decaying beauty baryons. Therefore, the final results of the discussed analyses will not only create further constraints themselves, but they will also mark the starting ground for new analyses. Some of these future possibilities will be outlined in this chapter.

9.1 $\Lambda_b^0 \to \Lambda\ell^+\ell^-$ decays

Since the Λ_b^0 baryon has the highest production fraction out of the four weakly decaying beauty baryons, it dominates the current knowledge of weakly-decaying b-baryons. However, there is still room for improvements. As discussed the dielectron decay mode of the $\Lambda_b^0 \to \Lambda\ell^+\ell^-$ decays has never been observed before. Since the lepton flavour universality of the SM was proven to hold also in $b \to s\ell^+\ell^-$ transitions [107, 108], studying differential branching fractions of electron decay modes is crucial to clarify the B anomalies observed in differential branching fractions of muonic $b \to s\ell^+\ell^-$ transitions. Additionally, the very low q^2 bin close to the photon pole, can be used to directly access the photon polarisation with electron decay modes as it has been proven in the measurement of the $B^0 \to K^{*0}e^+e^-$ [111]. Baryonic $b \to s\ell^+\ell^-$ transitions are particularly useful to access the photon polarisation, since the subleading weak decays of the hyperons preserve the underlying helicity structure of the $b \to s\ell^+\ell^-$ transition. The $\Lambda_b^0 \to \Lambda e^+e^-$ decay mode is the most promising one, due to the higher production fraction of the Λ_b^0 baryon.
However, the more interesting challenge will be to perform the angular analysis of the $\Lambda_b^0 \to \Lambda e^+e^-$ decay. This is not only the case because the angular observables are a more powerful tool to disentangle Wilson coefficients in order to set limits on NP scenarios, but to provide a novel way to test lepton flavour universality. If the difference between the angular observables of the muon and electron mode $K_{LFU} = K_i^\mu - K_i^e$ is not compatible with zero, there is a sign for lepton flavour non-universality and NP. Beside the

possibility to search for NP, the difference has a second benefit. This particular way to test the LFU of the SM will help to shine a light on the size of charm loop contributions to the size of the B anomalies as they completely cancel in the difference between the angular coefficients. This is particularly important since the unbinned amplitude analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay [110, 234] did not yet help to fully clarify the picture as intended (with the available amount of data). However, it proved that the tensions of the SM predictions with the measured Wilson coefficients remain at a level of $1.4 - 1.9$ standard deviations even if significant QCD effects are allowed, motivating the continued investigation of the nature of the deviations. This kind of measurements based on the difference between the angular coefficients of the electron and muon decay modes have previously been only performed by the Belle collaboration [10].

9.2 Prospects with the isospin doublet Ξ_b

It seems that the differential branching fraction of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay as a function of q^2 [7], displayed in Fig. 1.7, features a lower branching fraction with respect to the SM prediction in the central q^2 bins below the J/ψ resonance. While the statistical precision is not sufficient to derive any conclusion if this is indeed a deviation, it is consistent with the ones seen in other B anomalies in B meson decays. While the size of the possible deviation can only be accessed after more data have been analysed, it is equally helpful to start investigating the differential branching fractions of other weakly decaying b-baryons. The most promising candidate is the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay. However, in order to measure the differential branching fraction, the production fraction of the Ξ_b^- baryon needs to be measured first. As already discussed in Sec. 6.2.2 the most promising approach is via semileptonic decay modes as it has been done in the case of the Λ_b^0 baryon [82, 83]. Similarly, the angular analysis of the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay is particularly interesting due to its rich angular structure arising from the two subsequent weak decays, as displayed in Fig. 6.9.

While the observation and measurement of $b \to s\ell^+\ell^-$ transitions in general is very interesting in the context of indirect searches, the isospin doublet Ξ_b^- and Ξ_b^0 offers additional ways to test the SM. In general, it is believed that the spectator quarks of the b-hadrons do not influence the $b \to s\ell^+\ell^-$ transition itself as it has been tested in mesonic decays [6, 13]. Therefore, there should be no obvious difference in the behaviour of isospin doublets. In the case of the Ξ_b baryons, their isospin doublet nature can be used to test this assumption by measuring the isospin asymmetry:

$$\mathcal{A}_{iso} = \frac{\Gamma(\Xi_b^0 \to \Xi^0\mu^+\mu^-) - \Gamma(\Xi_b^- \to \Xi^-\mu^+\mu^-)}{\Gamma(\Xi_b^0 \to \Xi^0\mu^+\mu^-) + \Gamma(\Xi_b^- \to \Xi^-\mu^+\mu^-)}. \tag{9.1}$$

This measurement would be complementary to isospin asymmetry measurements currently performed in the mesonic sector using $B \to K^{(*)}\mu^+\mu^-$ decays. However, there are several challenges that need to be tackled first. First of all, both rare decays need to be observed. Secondly, the asymmetry in the Ξ_b^- and Ξ_b^0 production rates at the LHC is not yet known and needs to be measured. The production fractions of both Ξ_b baryons are expected to be not equal due to contributions of their excited states, as discussed in Sec. 1.2.2. In order to determine the asymmetry in the production fractions between the

Ξ_b^- and Ξ_b^0 baryon precisely, the $\Xi_b^0\pi^0$ final state needs to be probed. This study should enable to discover the missing $\Xi_b'^0$ state and measure its feed-down rate contributing to the asymmetry in the production fractions.
Finally, the studies of the resonant charmonia decay modes of the Ξ_b^0 baryon, discussed in Ch. 7, indicate that the available sub-dataset from Run 2 will not have sufficient statistical power to directly observe the rare decay mode $\Xi_b^0 \to \Xi^0\mu^+\mu^-$. Therefore, this channel is of high experimental interest for Run 3. However, there are drawbacks that need to be considered. In particular, the high background levels due to the abundant production of neutral pions in various unrelated processes are expected to increase significantly with the increase of the pile-up for Run 3. Therefore, more sophisticated background suppressing techniques will be needed with respect to the conditions during the Run 2 data-taking. A promising development will be the timing information of the calorimeter planned to be implemented in future upgrades [235]. The timing information would enable to disentangle neutral pions from different primary vertices and consequently enable a handle on the π^0 background level.

9.3 Towards rare Ω_b^- decays

The study of rare Ω_b^- decays such as $\Omega_b^- \to \Omega^-\mu^+\mu^-$ are particularly interesting since the Ω^- is the only known baryon of a spin 3/2 which decays weakly. All the other baryons of spin 3/2 would decay strongly (or electromagnetically) to their respective ground states of the spin 1/2, or are yet to be discovered (e.g. the bbb state). Therefore, angular analysis of the $\Omega_b^- \to \Omega^-\mu^+\mu^-$ and $\Omega_b^- \to \Omega^- J/\psi$ decay modes offer a possibility to reveal the rich angular structures. Additionally, the branching fraction measurement of the $\Omega_b^- \to \Omega^-\mu^+\mu^-$, especially when performed differentially in q^2, enables to cross-check the deviations from the SM predictions observed in the $B_s^0 \to \phi\mu^+\mu^-$ decay in the central q^2 region [16] with a baryonic decay mode[1].
Additionally, there is a broad theoretical interest in decays of $(1/2) \to (3/2)\gamma$ [236, 237], as they are convenient to probe the photon polarisation in b-baryon decays and search for devaitions from the left-handed nature of the weak interaction. The $\Omega_b^- \to \Omega^-\gamma$ is particularly interesting due to the weakly decaying Ω hyperon.

9.4 Resonant decay modes

Additionally, the J/ψ decay modes of the Ξ_b^- and Ω_b^- baryon have been proven to be very clean from contamination due to different backgrounds and the overall background level is quite low. Combined with the low Q-values due to the large mass of the J/ψ resonance, these decay modes are therefore ideal candidates to study the properties of the beauty baryons in the future. The situation is similar for the $\Lambda_b^0 \to \Lambda J/\psi$ decay, but more backgrounds need to be considered. The same assumption does not hold for the Ξ_b^0 decay mode due to the overall background level being higher due to the backgrounds caused by the π^0 and the challenging reconstruction of neutral particles.

[1] While is was not subject to the presented thesis, Dr. Vitalii Lisovskyi and the author are currently also working on the first search for the $\Omega_b^- \to \Omega^-\mu^+\mu^-$ decay integrated over q^2.

Conclusion

Weak decays of beauty hadrons offer an attractive laboratory to search for effects beyond the SM. FCNCs such as $b \to s\ell^+\ell^-$ transitions are of particular interest due to their high suppression in the SM and their resulting sensitivity for smallest deviations from the SM predictions. In the last decade, a consistent set of deviations from the SM predictions has been seen in differential branching fractions and angular observables of dominantly mesonic $b \to s\ell^+\ell^-$ transitions, called B anomalies. While most of those deviations individually are not statistically significant, their consistent nature may point to either effects beyond the SM, or yet (properly) unaccounted SM effects like *e.g.* charm-loop contributions. Further exploration of these anomalies can be achieved by either collecting larger datasets, or studying additional processes. In particular, baryonic $b \to s\ell^+\ell^-$ transitions offer an attractive setting for complementary tests of the anomalies, while simultaneously enabling to test a possible spin dependence of NP. However, so far only baryonic $b \to s\ell^+\ell^-$ transitions of the Λ_b^0 baryon have been studied due to its higher production cross-section compared to the other four weakly-decaying beauty baryons. The low production cross-sections also limit the precision with which the properties of weakly decaying b-baryons are determined yet. Uncertainties on b-baryon properties propagate not only to all experimental measurements of their decays, but also to their theoretical predictions. Current studies of b-baryons are limited to the LHC experiments due to constraints in the energy reach of current lepton colliders.
This thesis presents recent updates on precision measurements of the Ω_b^- baryon properties, as well as studies of several baryonic loop-level $b \to s\ell^+\ell^-$ transitions and their tree-level $b \to s(c\bar{c} \to \ell^+\ell^-)$ counterparts. These studies include the first search for the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay, the observation of the $\Xi_b^- \to \Xi^-\psi(2S)$ and $\Xi_b^0 \to \Xi^0 J/\psi$ decays, as well as first promising hints for the $\Xi_b^0 \to \Xi^0\psi(2S)$ decay and the angular analysis of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay. Beside the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay, the other three decay modes have not been observed at the start of this thesis and only the $\Xi_b^- \to \Xi^-\psi(2S)$ has just recently been observed for the first time by the CMS collaboration [18]. All the presented analyses are performed with data collected by the LHCb experiment.

The first analysis presented in this thesis contains two parts and has been published in Ref. [19]. The first part is the mass measurement of the Ω_b^- baryon using the full Run 1 and Run 2 dataset of the LHCb experiment, corresponding to an integrated luminosity of $9\,\text{fb}^{-1}$. The fact that the mass difference between the initial b-baryon and its decay products (energy release, or the Q-value) is rather small, for both the $\Omega_b^- \to \Omega^- J/\psi$ and $\Xi_b^0 \to \Xi^0 J/\psi$ decays, makes these modes excellent for mass measurements. To further

improve the sensitivity, the mass difference between the Ω_b^- and Ξ_b^- baryons is extracted, rather than the absolute Ω_b^- mass value. This approach comes with the benefit that a lot of the systematic uncertainties cancel due to the nearly identical topologies of both decay modes. They only differ by one kaon versus a charged pion in the final state. The selection developed for both decay modes largely benefits from the unique topologies that cannot be mimicked by any mesonic decay mode. The mass difference is extracted from a simultaneous fit with four projections, the Run 1 and Run 2 datasets for both the Ω_b^- and Ξ_b^- decay. The most precise mass measurement of the Ξ_b^- mass [20] to that date, has then been used to determine the Ω_b^- mass:

$$m(\Omega_b^-) - m(\Xi_b^-) = 248.54 \pm 0.51\,(\text{stat}) \pm 0.38\,(\text{syst})\,\text{MeV}/c^2\ ,$$
$$m(\Omega_b^-) = 6045.9 \pm 0.5\,(\text{stat}) \pm 0.6\,(\text{syst})\,\text{MeV}/c^2.$$

This result is in agreement with the world-average value, but over a factor two more precise: this proves that LHCb is the main player in the Ω_b^- baryon physics. Additionally, the LHCb average of the Ω_b^- mass $[m(\Omega_b^-)]_{\text{LHCb average}} = 6045.7 \pm 0.5\,(\text{stat}) \pm 0.6\,(\text{syst})\,\text{MeV}/c^2$ has been determined and correctly account for correlations between different LHCb measurements.

Using the same decay modes, the first measurement of the relative production fraction times the branching fraction of the Ω_b^- baryon at the LHC was performed. Due to the energy dependence of production fractions and the limited statistics of the Ω_b^- sample, the measurement is only performed with the Run 2 dataset corresponding to $6\,\text{fb}^{-1}$. The result

$$\frac{f_{\Omega_b^-}}{f_{\Xi_b^-}} \times \frac{\mathcal{B}(\Omega_b^- \to J/\psi\Omega^-)}{\mathcal{B}(\Xi_b^- \to J/\psi\Xi^-)} = 0.120 \pm 0.008(\text{stat}) \pm 0.008(\text{syst})$$

is obtained from a simultaneous fit with two projections in the defined kinematic region of $2 < \eta < 6$ and $p_T < 20$ GeV/c for the Ξ_b^- and Ω_b^- baryons. The biggest challenge of the analysis were the limited knowledge about strange baryon properties, and the large fraction of downstream tracks. Due to their lower tracking and vertexing efficiency, their fraction needs to be precisely known in each dataset to ensure no bias is introduced when calibrating and subsequently merging different datasets to extract the final results. A good stability of the result was reached across different sub-datasets, proving the successful calibration.

The next chapter was dedicated to the study of rare Ξ_b^- decay modes. In the first part the observation of the $\Xi_b^- \to \Xi^-\psi(2S)$ decay mode was reported for the first time at the LHCb experiment, using the full Run 1 and Run 2 dataset corresponding to $9\,\text{fb}^{-1}$. The branching fraction of the decay mode with respect to that of the $\Xi_b^- \to \Xi^- J/\psi$ has been evaluated and found to be

$$r_{BR}^{\psi}(\Xi_b^-) = 0.492 \pm 0.041(\text{stat}) \pm 0.021(\text{syst}).$$

When the result is compared to the first observation performed by the CMS collaboration [18], they differ about 1.5 standard deviations, but the LHCb result presented in this thesis is a factor 5 more precise. Additionally, the ratio r_{BR}^{ψ} also differs from the quark model predictions [22], similarly to the measurements in Refs. [23, 24] of the ratio of the charmonia branching fractions of Λ_b^0 decays (6.3). The result highlights the importance of studying the charmonia decay modes as well and their potential to test the quark-model

predictions.
The second part was dedicated to the current status of the first search for the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay mode pre-unblinding. Due to the lack of any absolute branching fraction measurement of the Ξ_b^- decays, the $\Xi_b^- \to \Xi^- J/\psi$ decay mode was chosen as normalisation to cancel as many systematic uncertainties as possible. An estimation of the $\Xi_b^- \to \Xi^- J/\psi$ branching fraction was then used to determine the ratio r_{BR} and the expected signal yields in the three different q^2 regions. These predictions have then been used to test the simultaneous fit with six projections, two decay modes and three subsets of data, with pseudoexperiments generated according to the predicted SM yields. Lastly, the expected statistical significance of the SM-like signal has been evaluated and the limit setting procedure has been put into place, preparing for the unblinding. The biggest challenges in this process arise due to the q^2 dependence of the efficiencies and the missing form factor predictions for the hadronic part of the decay modes. The latter leads to phase-space simulation samples, that have been calibrated according to the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ physics model, to better account for helicity angles and the q^2 distribution of the decay. The q^2 dependence of the efficiencies has been successfully studied with the simulation and is included in the fitting procedure to extract the branching fraction.
When this thesis was written, the analysis was in the review by the LHCb Collaboration, and final unblinding was expected to happen on a short timescale. It is expected that this analysis will lead to not only the second measurement challenging the quark model predictions with charmonia decays of weakly decaying particles but to the first observation of baryonic $b \to s\ell^+\ell^-$ transitions with Ξ_b^- baryons. This would only be the second of the four weakly decaying beauty baryons, that can be used to study $b \to s\ell^+\ell^-$ transitions in the future. Consequently, it highlights the importance of studying rare baryon decays.
The next analysis presented describes the first steps taken for the search of the $\Xi_b^0 \to \Xi^0 J/\psi$ and $\Xi_b^0 \to \Xi^0\psi(2S)$ decay modes. The analysis is performed with a partial Run 2 dataset, corresponding to $5.4\,\text{fb}^{-1}$. The main challenge of this analysis arises due to the combination of the low production cross-section of the Ξ_b^0 baryon and presence of neutral particles in the decay chain, especially a π^0 in the final state. Since the Ξ^0 decays into neutral particles, its decay vertex cannot be reconstructed via tracks, therefore the π^0 is assumed to originate from the primary vertex. This has several impacts: first, the background level due to the abundant production of π^0 mesons in all kind of processes is very high. The second impact is that it causes a mass-dependent failure rate of the decay-tree fitter, which is crucial to improve the precision of reconstructed quantities in decays with neutral particles. Finally, presence of a π^0 significantly worsens the mass resolution. In the work presented in this thesis, a first cut-based selection has been developed and the $\Xi_b^0 \to \Xi^0 J/\psi$ decay mode has been observed for the first time. Interestingly, this might be the first time the Ξ^0 baryon is reconstructed in an LHC experiment. Additionally, two major challenges were tackled: the development of a first truthmatching strategy for the signal simulation, and significantly reducing the DTF failure rate. The latter enabled the development of the BDT for the selection which lead to some strong hints for the first observation of the $\Xi_b^0 \to \Xi^0\psi(2S)$ decay as well. However, these results are only preliminary and dedicated studies of the challenges introduced by the π^0 reconstruction are needed. Among them are background studies to examine the contribution from partially reconstructed Ξ_b^- decays combined with random π^0 mesons, as well as the calibration of the simulation to ensure no unaccounted biases are introduced. Furthermore, these changes

lead to the possibility to improve the selection further. Nevertheless, this study is a huge step forward and proves the capability of reconstructing the $\Xi_b^0 \to \Xi^0 \mu^+ \mu^-$ decay in the future with larger available datasets.
In the last analysis chapter of this thesis, the current status of the angular analysis of the $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ decay using full Run 1 and Run 2 dataset is presented. The analysis is currently still blinded, but the full analysis strategy has been evaluated in order to prepare for the internal review. The analysis is performed in the dilepton invariant mass squared above $q^2 > 15\,\mathrm{GeV}^2/c^4$, where most of the signal is expected according to the differential branching fraction measurement of $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ [7]. The angular analysis attempts to extract all ten angular coefficients using a maximum likelihood fit, for the first time. A dedicated set of selection requirements is developed in order to obtain clean signals of the rare decay, and suppress all the dominant background modes. This includes a dedicated multivariate classifier to suppress the background consisting of random track combinations. A mass fit model has been developed, crucial to later distinguish between the signal and background events in the angular fit components of the four-dimensional maximum-likelihood fit. Dedicated phase-space simulation samples have been used to extract the angular acceptance of the LHCb detector, and a fit model has been developed to extract the angular coefficients. The latter has been studied extensively, especially with pseudoexperiments, to ensure no bias is introduced due to the relatively small signal yield compared to the complexity of the fit model. In fact a small bias was uncovered, caused by the similarity of the $\cos(\theta_\Lambda)$ distribution in signal and background due to the selection of real Λ hyperons in both cases, which will be corrected in the final fit to data. The remaining open point is to finalise the calibration of the simulation as residual differences have been uncovered. Afterwards, the full analysis will be updated in order to start performing the missing validation steps of the fit procedure relying on proper data and evaluate systematic uncertainties. These steps are currently performed insight of the pending internal review.
Beside the measurements presented in this thesis I also contributed to the assembly of the SciFi detector and a failure investigation during the end of the year shutdown in 2023. The optical fibres needed for an efficient communication and data readout between the SciFi frontend-boards and the data-center had to undergo a thorough optical inspection to ensure no dirt or scratches on the optical fibres would reduce the data quality. After taking part in the optical inspection during the assembly process in 2021, I was able to apply the obtained skills to the failure investigation of the SciFi detector in 2023. During the data-taking several half-ROBs and data links had to be excluded from the data-taking due to instabilities reducing the data quality. During the optical inspection, I was able to trace back the instability to outgasing of the VTRx blocking the center of the optical fibres. This was the case for all excluded half-ROBs and data links, where no reason for the instability was found prior. Consequently, all affected and accessible optical fibres were cleaned leading to only two excluded data links after the intervention.

To conclude, the presented work proved that the LHCb experiment is ideally suited for the analysis of rare decays of all four weakly-decaying beauty baryons due to its high trigger efficiency, as well as excellent tracking and particle identification performance. The study of weakly-decaying beauty baryons shows great potential to help to deepen the understanding of the B anomalies and reveal their true nature. Additionally, the described

contributions to the SciFi detector were important to ensure a smooth data-taking with the SciFi detector during ongoing Run 3 and upcoming Run 4 of the LHC. The SciFi detector is particularly important for the reconstruction of long and downstream tracks and therefore enables the future studies of weakly-decaying beauty baryons that decay into long-lived hyperons, as presented in this thesis.

Appendix A

Additional information about used tools

A.1 Momentum scale calibration

The precise knowledge of particle momenta is crucial when performing precision measurements. This is particularly important, when trying to access characteristics like a particle mass or natural width. Since the momentum of a particle track is directly calculated from the curvature ω of the track, any bias in the conversion factor between the two quantities can affect the precision. This factor is called *momentum scale* and determined with the knowledge of the magnetic field map, the detector alignment, and the amount of detector material crossed by the particle (which introduces an energy loss). The nominal value of the momentum scale was determined before the start of the data taking. Thanks to recently available measuring systems, the magnetic field map is known more precisely nowadays. Additionally, detailed studies showed, that the alignment of the tracking system changes during runs due to the influence of the magnetic field. Lastly, the detector material is only know to a precision of $5-10\%$ [129] leading to possible underestimation of the energy loss correction applied in the tracking algorithms. As a consequence, a momentum scale calibration procedure is applied in LHCb to reduce the bias. The procedure is explained in detail in Ref. [203].
For a two-body decay $A \to BC$, the mass of the particle A is calculated via relativistic kinematics from the mass, energy and momentum of the decay products B and C. Therefore, the reconstructed invariant mass changes with respect to the magnitude of the momentum of one daughter particle p_B:

$$\frac{\mathrm{d}\, m_{BC}}{\mathrm{d}\, p_B} = \frac{m_{BC}^2 - m_B^2 - m_C^2 - 2\, m_B^2\, E_C/E_B}{2\, m_{BC}\, p_B}$$

Consequently, if the momentum scale is biased by a scale factor $p_{B/C} \to (1+\alpha_{B/C})p_{B/C}$, the two-body invariant mass changes in first order of $\alpha_{B/C}$ with respect of Eqn. (A.1). Under the assumption that the bias is equal for both daughter particles, the total mass bias in the first order is obtained by summing the individual ones. Therefore, the shift in the mass is given by:

$$\Delta m = \alpha \frac{m_A^2 - m_B^2 - m_C^2 - m_B^2\, E_C/E_B - m_C^2\, E_B/E_C}{m_A}. \tag{A.1}$$

To obtain the scale factor from the observed shifts, certain decays can be targeted to further simplify the calculation. For example, one can use particle decays with the equal final-state masses, like $K^0_S \to \pi^-\pi^+$, leading to

$$\Delta m = \alpha \frac{m_A^2 - m^2(2 + E_C/E_B + E_B/E_C}{m_A}. \tag{A.2}$$

For decays where the initial particle is much heavier than the daughter particles, like $J/\psi \to \mu^+\mu^-$, this simplifies even further to

$$\Delta m \approx \alpha \cdot m_A. \tag{A.3}$$

This particular mass bias is caused by limited knowledge of the integrated magnetic field or the energy loss in the detector. With similar calculation, one can determine a curvature bias $\Delta\omega$

$$\Delta m_{BC} = \frac{1}{2}\,\Delta\omega\; m_{BC}\;(p_C - p_B) \tag{A.4}$$

due to misalignment of the detector.
The correction of the momentum scale was determined from detached $J/\psi \to \mu^+\mu^-$ and $B^+ \to K^+ J/\psi$ decays. The uncertainty on the procedure is judged to be about 0.03% by studying the decays of well-known resonances. Fig. A.1 shows the average momentum scale bias for several decay modes for 2011 (left) as Run 1 and 2016 (right) as Run 2 proxy.

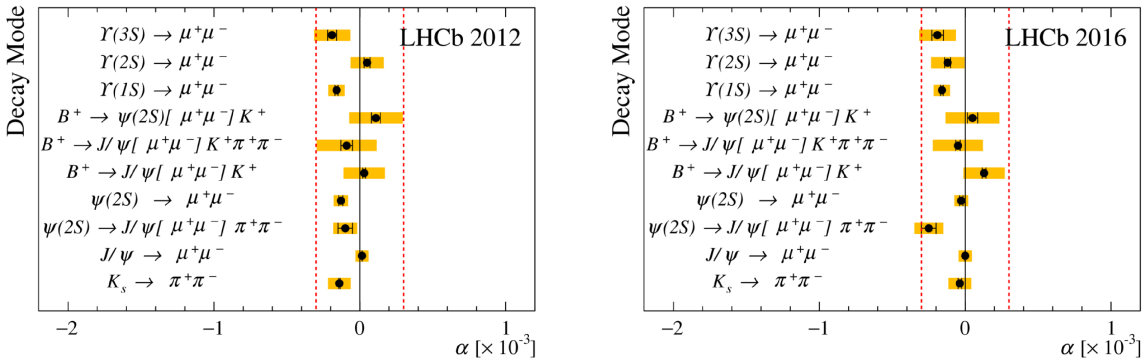


Figure A.1: Average momentum scale bias α extracted from the reconstructed mass of various decay modes after the momentum calibration procedure for dataset from 2012 (left) and 2016 (right). The statistical uncertainties are indicated by the black error bars and the systematic uncertainties are by the yellow bars. The assigned uncertainty of $\pm 0.03\%$ on the momentum scale is indicated by the red dashed lines. Taken from Ref. [203].

A.2 Multivariate analysis tools

One of the most important tasks in particle physics is to efficiently separate signal and background candidates, while losing as little signal as possible. This is especially true when searching and analysing rare decay modes, such as those discussed in this thesis. Due to their different nature, the distributions of physics parameters differs between signal and background. For example, a common background source consists of random track combinations, which often involve tracks not originating from b-hadron decays and therefore have lower momenta or a poor vertex quality. This difference can be utilised in selections based on rectangular cuts. However, this approach does not take into account any dependencies or correlation between the individual variables and is therefore inefficient.
This problem can be overcome by performing the separation with a multivariate analysis (MVA). It is a statistical procedure analysing several variables simultaneously, while taking correlations into account, and provides a score for each candidate. In high-energy physics, machine-learning approaches such as Neutral Networks (NN), Multi-Layer Perceptrons (MLP) and Boosted Decision Trees (BDTs) are the most commonly used approaches to perform MVAs.

A.2.1 Boosted Decision Tree

As the name indicates, BDTs are based on decision trees [238], providing a simple structure, that is fast to train and has a high separation efficiency. Therefore, they have been used as the MVA approach for the analyses presented in this thesis. While they contribute to rejecting different background sources, their main purpose is the rejection of combinatorial background[1].
A Decision Tree starts at the root note with one random variable and applies an optimal rectangular cut splitting the entire provided dataset into two subsets. It is then followed by further internal nodes iteratively splitting each subset in two new partitions, till in the end each terminal node provides a signal-like and background-like subdatasets. Since the separation quality can depend on the order in which the variables are accessed, BDTs are not based on one single tree but an ensemble of K of them, called forest. In this process the prediction obtained from the tree n is passed on to the next tree $n + i$ in order to reduce possible bias by assigning misclassified events with a larger weight. In the end, all predictions of the single trees are combined into an average score. This procedure is called boosting and displayed in Fig. A.2.
Consequently, for each BDT there is a number of so-called hyperparameters that can be chosen. This includes *e.g.* the number of trees, the depth of each individual tree (equal to the number of nodes), the learning rate of the algorithm (called shrinkage) and the number of input variables. In addition, there are different boosting algorithms available. For the work presented in this thesis, the gradient boosting has been chosen, as it provides a higher robustness against outliers and mislabeled data points [238]. All BDTs in this thesis have been implemented in the TMVA framework [238] developed at CERN.

[1] Combinatorial background is a background consisting out of combinations of random particle tracks picked up by the reconstruction software.

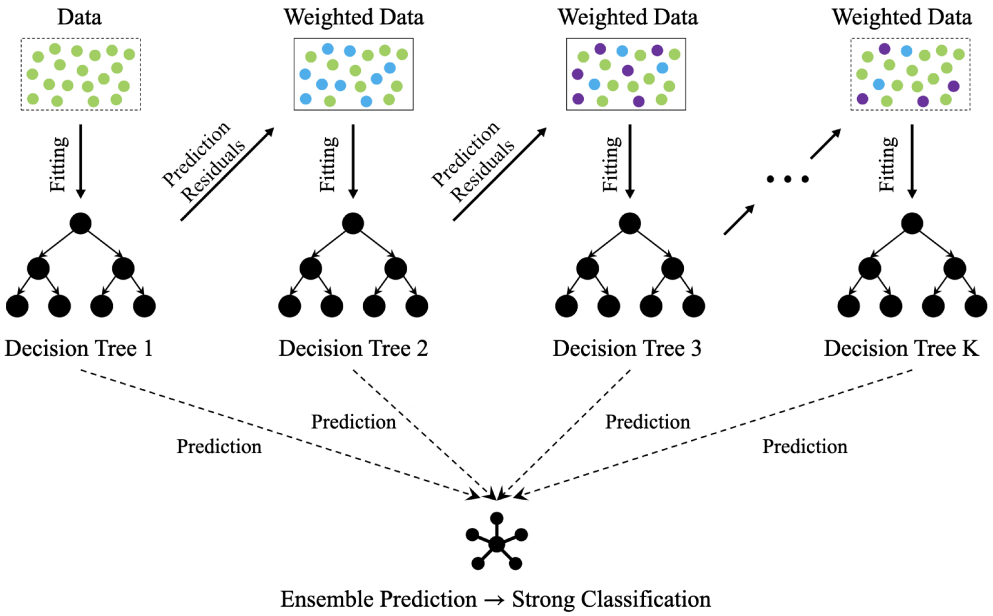


Figure A.2: Schematic presentation of the architecture of boosted decision trees. Adaptation from Ref. [239].

A.2.2 BDT evaluation

As explained in the previous subsection, there are several parameters that need to be optimised in order to choose the best working point of a BDT. As a first step a larger set of input variables should be considered. These variables should have a specific physics motivation, making them more unlikely to look similar in the background distributions. For all these variables the distributions for the signal and background should be checked to avoid biases from any unusual and nonphysical shapes. In particular, the used variables should be well modelled in the (calibrated) simulation. The number of input variables should be optimised with respect to the size of the dataset. In order to reduce the number of input variables, their correlations with each other can be checked. Highly correlated pairs of variables offer little additional information and one of them can therefore be safely removed. At this stage, the training stage of the BDT can start, where it is handed a dedicated dataset to learn classifying the likeliness of each event between background (assigned a score of -1 in TMVA) and signal (1). The provided dataset includes a signal and background proxy. In this thesis, the signal proxy is usually described by the simulation of the targeted decay mode, and the background proxy is obtained by providing data sidebands while excluding the signal region. After the training, the BDT provides the feature importance of each variable, which can be used to further reduce the used set of them by rejecting the unimportant observables. In the next step, the trained BDT classifier is tested on an independent testing dataset not used in the training. This enables to test the classifier for overtraining, meaning that it simply remembers the exact data as provided in the training. This possible dependence can be reduced by for example a k-fold cross-validation. The full dataset is split in k subsets, and while the $k-1$ subsets are used in the training, the last one is used for the testing. This procedure is repeated till each subset has been used as a validation sample once. In order to compare these individual classifiers, the actual performance of each BDT classifier needs to be evaluated. While

there are many different ways, for the work presented in this thesis the *receiver operating characteristic* (ROC) curve and the *area under the ROC curve* (AUC) score where chosen. The ROC curve plots the background rejection against the signal efficiency by counting the correctly assigned signal events against the background events classified as signal. The higher the signal efficiency and background rejection, the closer to unity the AUC score gets. Therefore, this procedure can be used to validate that a BDT is performing well and in the case of k-folding, that each individual classifier has a good performance. The final BDT is then obtained by averaging over the k individual classifiers.

Appendix B

Trigger selections

This chapter includes all tables with the HLT trigger lines used to select analysed decays, but not given in the main text of the thesis. It should be noted that while the L0 trigger lines are applied the dimuons, the HLT trigger lines are applied to the beauty hadrons. For the Ω_b^- mass and production fraction measurement, the HLT trigger lines are listed in Tab B.1.

Run 1	Run 2
HLT1	
`TrackMVA`	`TrackMVA`
`DiMuonHighMass`	`DiMuonHighMass`
HLT2	
`Topo2Body`	
`TopoMu2Body`	`DiMuonDetachedJPsi`
`DiMuonDetached`	
`DiMuonDetachedHeavy`	

Table B.1: Summary of HLT lines used in the Ω_b^- mass and relative production fraction measurement.

For the first search for the rare decay $\Xi_b^- \to \Xi^-\mu^+\mu^-$, the trigger selection is nearly identical to that for the Ω_b^- mass and production fraction, but for Run 2 the trigger selection also relies on the muon detached and topological triggers, as for Run 1. The used trigger lines are summarised in Tab. B.2. The same trigger selection is used for the angular analysis of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay.
For the first study of the $\Xi_b^0 \to \Xi^0 J/\psi$ and $\Xi_b^0 \to \Xi^0\psi(2S)$ decays, the dedicated trigger lines are given in Tab. B.3.

Run 1	Run 2
HLT1	
TrackAllL0 TrackMuon DiMuonHighMass DiMuonLowMass	TrackMVA TrackMuon TwoTrackMVA
HLT2	
Topo2Body TopoMu2Body DiMuonDetached DiMuonDetachedHeavy	Topo2Body TopoMu2Body TopoMuMu2Body DiMuonDetached DiMuonDetachedHeavy

Table B.2: Summary of HLT lines used in the angular analysis of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay, as well as the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ and $\Xi_b^0 \to \Xi^0\psi(2S)$ measurements.

2016, 2017 & 2018
HLT1
TrackMVA DiMuonHighMass
HLT2
DiMuonDetachedJpsi DiMuonDetachedPsi2S

Table B.3: Summary of HLT lines used in the first search for the $\Xi_b^0 \to \Xi^0 J/\psi$ and $\Xi_b^0 \to \Xi^0\psi(2S)$ decays.

Appendix C

Stripping selection requirements

This chapter contains the tables with versions of the Stripping software and the Stripping selection requirements applied in each dedicated analysis.
The specific cuts of the `Bu2LKK_mmLine` stripping line used for the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ angular analysis are displayed in Tab. C.1.
The specific cuts of the `B2XMuMu` stripping line used for the Ω_b^- mass and relative production fraction measurement as well as the search for the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay are displayed in Tab. C.2 and Tab. C.3.
The specific cuts of the `RareBaryonicMuMu` stripping line used for the first study of the Ξ_b^0 charmonia decay modes $\Xi_b^0 \to \Xi^0 J/\psi$ and $\Xi_b^0 \to \Xi^0\psi(2S)$ are displayed in Tab. C.4.

Table C.1: Summary of the requirements of the `Bu2LLK_mmLine` stripping, where the cuts are identical between all years.

	Requirement
Event	nSPDHits < 600
Λ_b^0	$\lvert m - m^{PDG}\rvert < 1500\,\text{MeV}/c^2$
	end vertex χ^2/ndf < 27
	$\chi^2_{\text{IP}}(primary) < 25$
	`DIRA` > 0.9995
	$\chi^2_{(DV\leftrightarrow PV)} > 100$
Λ	$\lvert m - m_{PDG}\rvert < 64\,\text{MeV}/c^2$
	vertex χ^2/ndf < 30 LL (25 DD)
	$p_T > 400$ MeV/c
p, π^-	$p > 2$ GeV/c **(only for π^-)**
	$\chi^2_{\text{IP}}(primary) > 9$ LL (4 DD)
$\mu^+\mu^-$	$m < 5500\,\text{MeV}/c^2$
	end vertex χ^2/ndf < 9
	$\chi^2_{(DV\leftrightarrow PV)} > 16$
	χ^2_{IP} (primary) >0
	$p_T > 0\,\text{MeV}/c$
μ	`hasMuon && isMuon`
	$p_T > 300\,\text{MeV}/c$
	χ^2_{IP} (primary) > 9

Table C.2: Summary of the requirements of the `B2XMuMu_Line` stripping line in 2011-16 for both $\Xi_b^- \to \Xi^- J/\psi(\to \mu^+\mu^-)$ and $\Omega_b^- \to \Omega^- J/\psi(\to \mu^+\mu^-)$ decays; and in 2017-18 for the $\Omega_b^- \to \Omega^- J/\psi(\to \mu^+\mu^-)$ only.

	Requirement
Ξ_b^- / Ω_b^-	$\lvert\sum(\text{charge})\rvert < 3$
	$4700 < m < 7000\,\text{MeV}/c^2$
	end vertex $\chi^2/\text{ndf} < 8$
	$\chi^2_{\text{IP}}(primary) < 16$
	`DIRA` > 0.9999
	χ^2 separation from related PV > 64
	$\max(\chi^2_{\text{IP}}(daughters, PV)) > 9$
Ξ^- / Ω^-	$\lvert m - m_{PDG}\rvert < 50\,\text{MeV}/c^2$
	vertex $\chi^2/\text{ndf} < 20$
	decay time > 2 ps
π_Ξ / K	`GhostProb` < 0.5
	$\chi^2_{\text{IP}}(primary) > 6$ (In Run1, only for L tracks)
Λ	$\lvert m - m_{PDG}\rvert < 30\,\text{MeV}/c^2$
	vertex $\chi^2/\text{ndf} < 30$ LL (25 DD)
	decay time > 2 ps
	χ^2 separation from related PV > 4 (LL)
p, π_Λ	$p > 2$ GeV/c
	$\chi^2_{\text{IP}}(primary) > 9$ LL (4 DD)
	`GhostProb` < 0.4 (only Run 2)
$\mu^+\mu^-$	$m < 7100\,\text{MeV}/c^2$
	vertex $\chi^2/\text{ndf} < 12$
	`DIRA` > -0.9
	χ^2 separation from related PV > 9
	$\max(\chi^2_{\text{IP}}(daughters, PV)) > 6$
μ	`hasMuon && isMuon && PIDmu> -3`
	`GhostProb` < 0.5
	$\chi^2_{\text{IP}}(primary) > 9$

Table C.3: Summary of the requirements of the `B2XMuMu_Line` stripping line in 2017-18, only for the $\Xi_b^- \to \Xi^- J/\psi(\to \mu^+\mu^-)$ channel. The cuts which differ from the 2011-16 version, are marked in bold.

	Requirement
Ξ_b^-	$\lvert\sum(\text{charge})\rvert < 3$
	$\mathbf{4900} < m < 7000\,\text{MeV}/c^2$
	end vertex $\chi^2/\text{ndf} < 8$
	$\chi^2_{\text{IP}}(primary) < 16$
	`DIRA` > 0.9999
	χ^2 separation from related PV $> \mathbf{121}$
	$\max(\chi^2_{\text{IP}}(daughters, PV)) > 9$
Ξ^-	$\lvert m - m_{PDG}\rvert < 50\,\text{MeV}/c^2$
	vertex $\chi^2/\text{ndf} < 20$
	decay time > 2 ps
π_Ξ	`GhostProb` < 0.5
	$\chi^2_{\text{IP}}(primary) > 6$
Λ	$\lvert m - m_{PDG}\rvert < 30\,\text{MeV}/c^2$
	vertex $\chi^2/\text{ndf} < 30$ LL (25 DD)
	decay time > 2 ps
	χ^2 separation from related PV > 4 (LL)
p, π_Λ	$p > 2$ GeV/c
	$\chi^2_{\text{IP}}(primary) > 9$ LL (4 DD)
	`GhostProb` < 0.4
$\mu^+\mu^-$	$m < 7100\,\text{MeV}/c^2$
	vertex $\chi^2/\text{ndf} < 12$
	`DIRA` > -0.9
	χ^2 separation from related PV > 9
	$\max(\chi^2_{\text{IP}}(daughters, PV)) > \mathbf{9}$
μ	`hasMuon && isMuon && PIDmu> -3`
	`GhostProb` < 0.5
	$\chi^2_{\text{IP}}(primary) > 9$

Table C.4: Summary of the requirements of the `RareBaryonicMuMu` stripping line in 2016-18.

	Requirement
Ξ_b^0	$\lvert m - m_{PDG}\rvert < 1600\,\text{MeV}/c^2$
	end vertex $\chi^2/\text{ndf} < 20$
	$\chi^2_{\text{IP}}(primary) < 25$
	`DIRA` > 0.999
	χ^2 separation from related PV > 64
Ξ^0	$\lvert m - m_{PDG}\rvert < 260\,\text{MeV}/c^2$
	$p_{\text{T}} < 200\,\text{MeV}/c$
	$\chi^2_{\text{MTDOCA}} < 16$
π^0	$p_{\text{T}} > 100\,\text{MeV}/c$
Λ	$1105\,\text{MeV}/c^2 < m < 1130\,\text{MeV}/c^2$
	$\lvert m - m_{PDG}\rvert < 35\,\text{MeV}/c^2$ (LL) $64\,\text{MeV}/c^2$ (DD)
	end vertex $\chi^2 < 30$ (LL) 25 (DD)
	χ^2 separation from related PV > 4 (LL)
p, π^-	$p > 2$ GeV/c
	$\chi^2_{\text{IP}}(primary) > 9$ LL (4 DD)
	`GhostProb` < 0.4
$\mu^+\mu^-$	$p_{\text{T}} > 0\,\text{MeV}/c$
	vertex $\chi^2/\text{ndf} < 10$
	χ^2 separation from related PV > 10
	$\min(\chi^2_{\text{IP}}(daughters, PV)) > 9$
μ	`hasMuon && isMuon`
	$\min(p_{\text{T}}) > 250\,\text{MeV}/c$
	`GhostProb` < 0.5
	$\chi^2_{\text{IP}}(primary) > 9$

Appendix D

Additional information on the Ω_b^- mass and relative production fraction

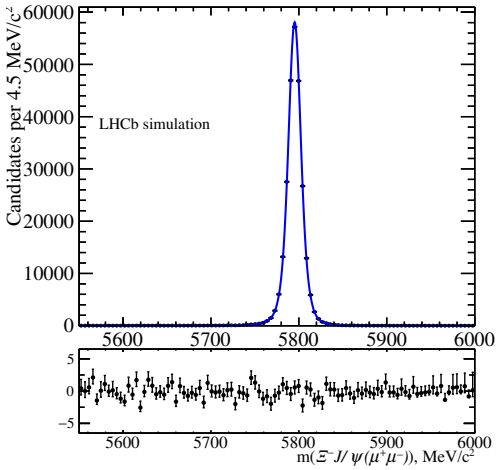
This appendix chapter summarises additional information for the measurement of the Ω_b^- mass and its relative production fraction with respect to the Ξ_b^- baryon.

D.1 Fit validation

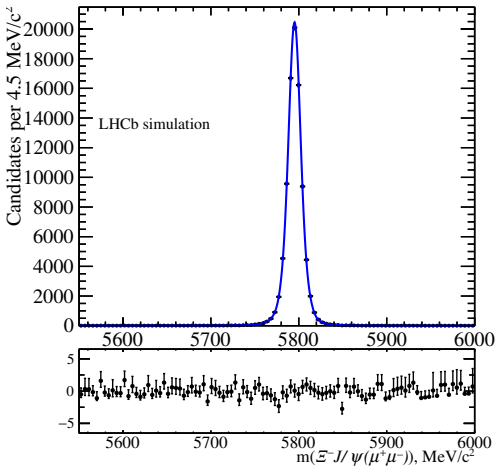
Fig. D.1 shows the four projections of the simultaneous fit to the $\Omega_b^- \to \Omega^- J/\psi$ and $\Xi_b^- \to \Xi^- J/\psi$ simulation samples used to test the fit setup and constrain the discussed fit parameters. This helps to improve the fit stability despite the low statistics, in particular the $\Omega_b^- \to \Omega^- J/\psi$ Run 1 sample.

D.2 MC calibration

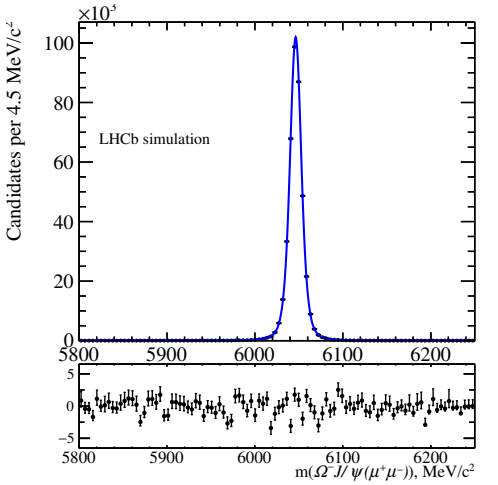
In order to calibrate known mismodelling of the simulation in a data-driven manner, the ${}_s\mathcal{P}lot$ technique was utilised. Looser selections are used in order to obtain background-subtracted data samples for both decay modes. The details of both selections can be found in Tab. D.1 for the $\Xi_b^- \to \Xi^- J/\psi$ decay and in Tab. D.2 for the respective Ω_b^- decay. The signal shapes are modelled by the sum of two double-sided Crystal Ball functions and after a fit to data, the tail parameters are fixed, before the total fit model is constructed with an exponential distribution describing the background. The result of the fit to Run 2 data can be seen in Fig. D.2, after which all parameters but the yields are fixed for the ${}_s\mathcal{P}lot$ fit.
Fig. D.3 displays the `nTracks` distribution of the $\Xi_b^- \to \Xi^- J/\psi$ dataset before (left) and after (right) applying a correction factor of 1.3.
The Run 2 $\Xi_b^- \to \Xi^- J/\psi$ ($\Omega_b^- \to \Omega^- J/\psi$) simulation is compared to the ${}_s\mathcal{P}lot$ of the $\Xi_b^- \to \Xi^- J/\psi$ ($\Omega_b^- \to \Omega^- J/\psi$) signal in data in order to evaluate the transverse momentum correction of the simulation. The comparison of the data and simulation distributions and the resulting weights are shown in Fig. D.4 for the Ξ_b^- case, and in Fig. D.5 for the Ω_b^- case.



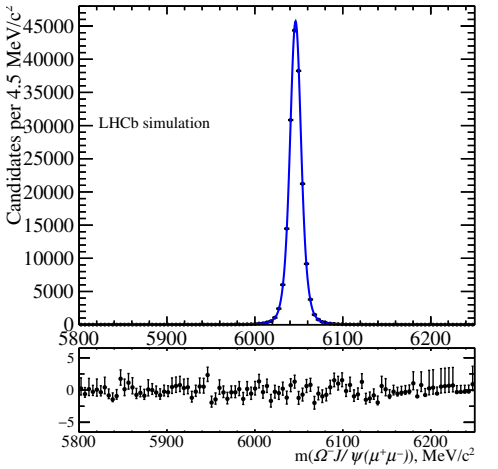
(a) $\Xi_b^- \to \Xi^- J/\psi$ Run 2



(b) $\Xi_b^- \to \Xi^- J/\psi$ Run 1



(c) $\Omega_b^- \to \Omega^- J/\psi$ Run 2



(d) $\Omega_b^- \to \Omega^- J/\psi$ Run 1

Figure D.1: Invariant mass distributions of the simulated $\Xi_b^- \to \Xi^- J/\psi$ and $\Omega_b^- \to \Omega^- J/\psi$ signals, with the results of the simultaneous fit overlaid.

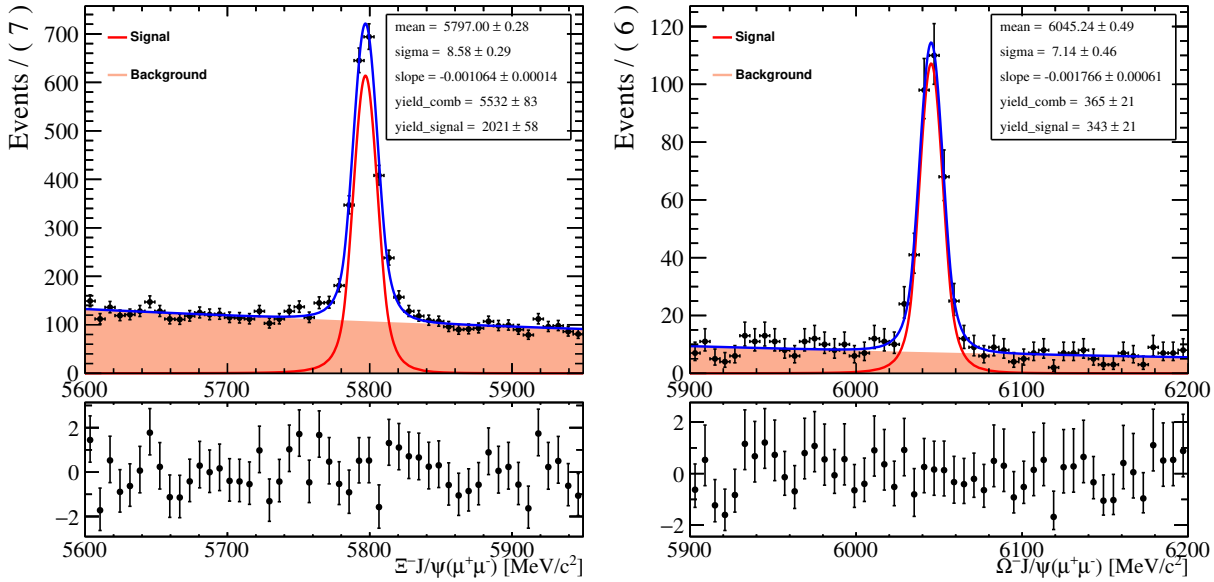
The correction weight maps for the Ξ_b^-, Ω^- (left) and respective Λ (right) momentum asymmetries can be found in Fig. D.6 and Fig. D.7. They are used to correct residual mismodelling of the helicity angles in the hyperon decays.
The final calibration corrects for remain mismodelling in the quality of the DTF χ^2_{DTF}. The comparison plots and weights can be found in Fig. D.8 (Ξ_b^-) and Fig. D.9 (Ω_b^-).

Type	Cut values
L0TOS	L0Muon,DiMuon
Hlt1TOS	TrackMuon,TrackMVA,TwoTrackMVA
Hlt2TOS	Topo(Mu)(Mu)2Body,DiMuonDetached(Heavy)
Mass windows	$3000 < m(\mu^+\mu^-) < 3150\,\text{MeV}/c^2$ $\lvert m(p\pi^-) - 1116\rvert < 8\,\text{MeV}/c^2$ $\lvert m(\Lambda\pi^-)_\Lambda - 1322\rvert < 10\,\text{MeV}/c^2$ $5600 < m(\Xi^-\mu^+\mu^-)_{\Xi,\Lambda,J/\psi} < 6000\,\text{MeV}/c^2$
PID	`ProbNNmu`$(\mu^\pm) > 0.1$
Muon kinematics	$p_T(\mu) > 800\,\text{MeV}/c$, $3 < p(\mu) < 150\,\text{GeV}/c$
Muon acceptance	`HasRich, InAccMuon` for $\mu^\pm$
Track quality	all tracks `GhostProb` < 0.3

Table D.1: Summary of selection applied to the $\Xi_b^- \to \Xi^- J/\psi$ data, for data-simulation comparison studies.

Type	Cut values
L0TOS	L0Muon,DiMuon
Hlt1TOS	TrackMuon,TrackMVA,TwoTrackMVA
Hlt2TOS	Topo(Mu)(Mu)2Body,DiMuonDetached(Heavy)
Mass windows	$3000 < m(\mu^+\mu^-) < 3150\,\text{MeV}/c^2$ $\lvert m(p\pi^-) - 1116\rvert < 8\,\text{MeV}/c^2$ $\lvert m(\Lambda K^-) - 1672.5 - m(\Lambda) + 1115.7\rvert < 10\,\text{MeV}/c^2$ $5900 < m(\Omega^-\mu^+\mu^-)_{\Omega,\Lambda,J/\psi} < 6200\,\text{MeV}/c^2$
Muon PID	`ProbNNmu`$(\mu^\pm) > 0.1$
Kaon PID	`PIDK>5`
Muon kinematics	$3 < p(\mu) < 150\,\text{GeV}/c$
Muon acceptance	`HasRich, InAccMuon` for $\mu^\pm$
Track quality	all tracks `GhostProb` < 0.3
Hyperon quality	$\tau(\Omega^-) > 0.5$ ps, $\tau(\Lambda) > 0$ ps

Table D.2: Summary of selection applied to the $\Omega_b^- \to \Omega^- J/\psi$ data, for data-simulation comparison studies.



(a) $\Xi_b^- \to \Xi^- J/\psi$ (b) $\Omega_b^- \to \Omega^- J/\psi$

Figure D.2: Preliminary fits to the $\Xi^- J/\psi$ and $\Omega^- J/\psi$ invariant mass with Ξ^- (Ω^-), Λ and J/ψ mass constraints applied; used for further background subtraction via ${}_s\mathcal{P}lot$ method in order to evaluate corrections to the simulation, and to quantify data-simulation agreement.

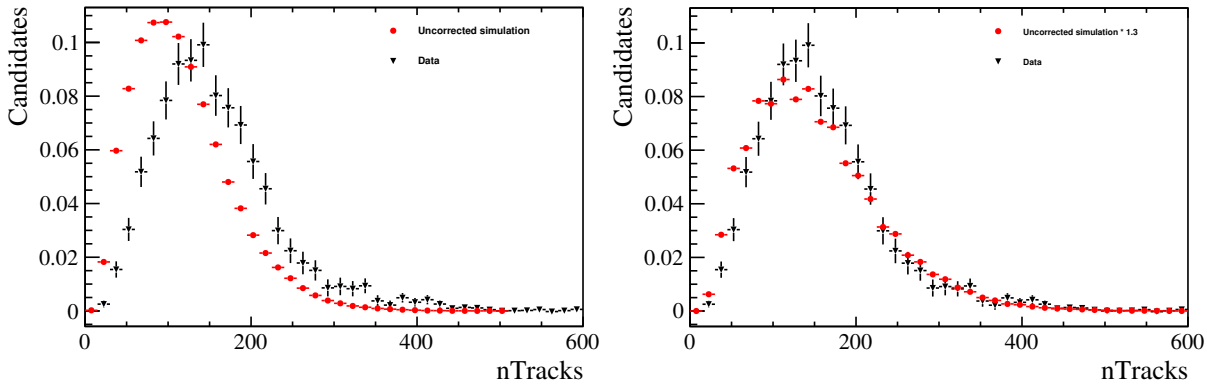
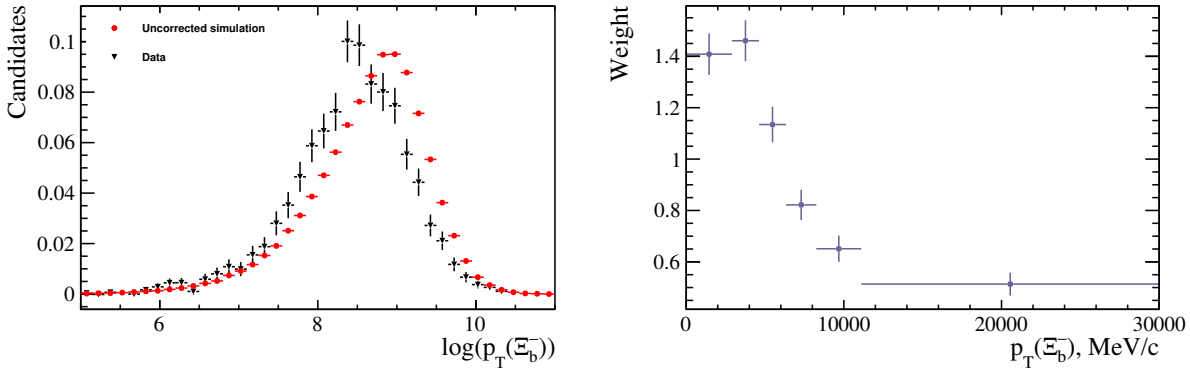
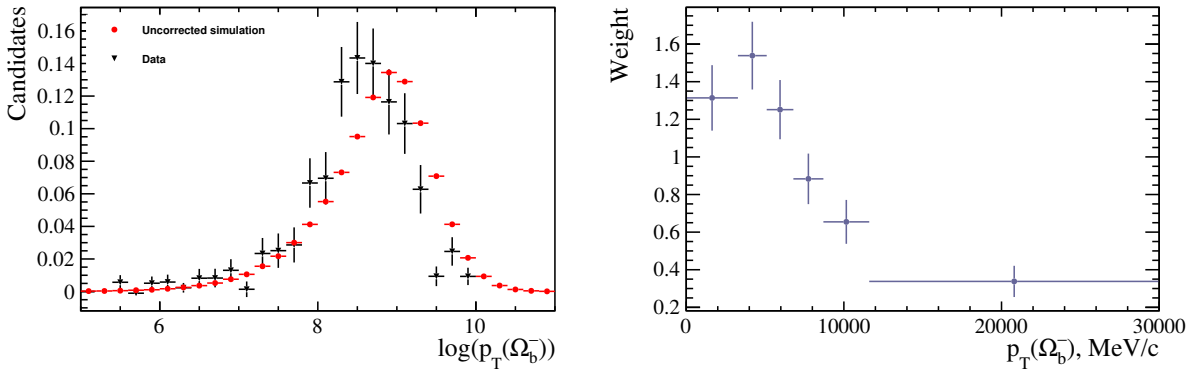


Figure D.3: nTracks variable distribution in the $\Xi_b^- \to \Xi^- J/\psi$ channel in Run 2, compared between data and (left) raw simulation, (right) simulation rescaled by a factor 1.3.



(a) (b)

Figure D.4: a): Normalised distributions of $p_T(\Xi_b^-)$ in Run 2 data and simulation; b): resulting Ξ_b^- kinematics correction for Run 2. Note that the plot a) has the logarithmic X scale (natural logarithm of the p_{T} measured in MeV/c).



(a) (b)

Figure D.5: a): Normalised distributions of $p_T(\Omega_b^-)$ in Run 2 data and simulation; b): resulting Ω_b^- kinematics correction for Run 2. Note that the plot a) has the logarithmic X scale (natural logarithm of the p_{T} measured in MeV/c).

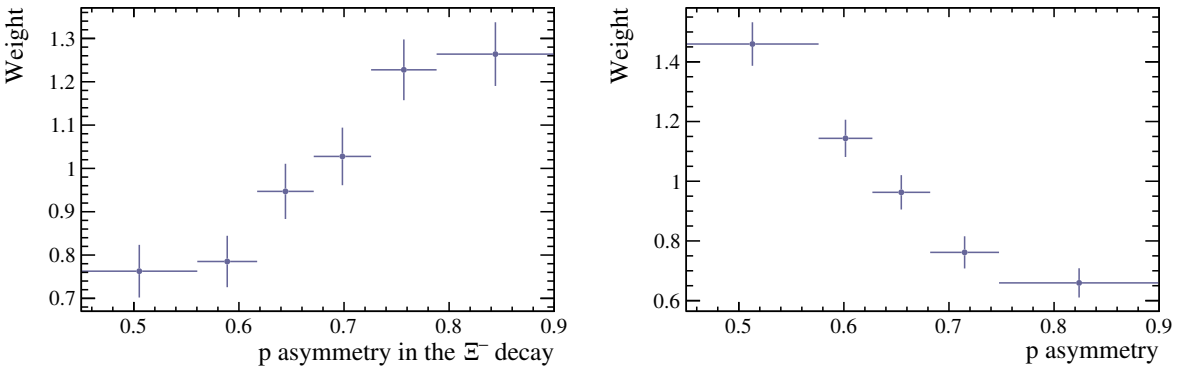


Figure D.6: Momentum asymmetry weights in hyperon decays from the Ξ_b^- decay chain (left: Ξ^-, right: Λ) for Run 2.

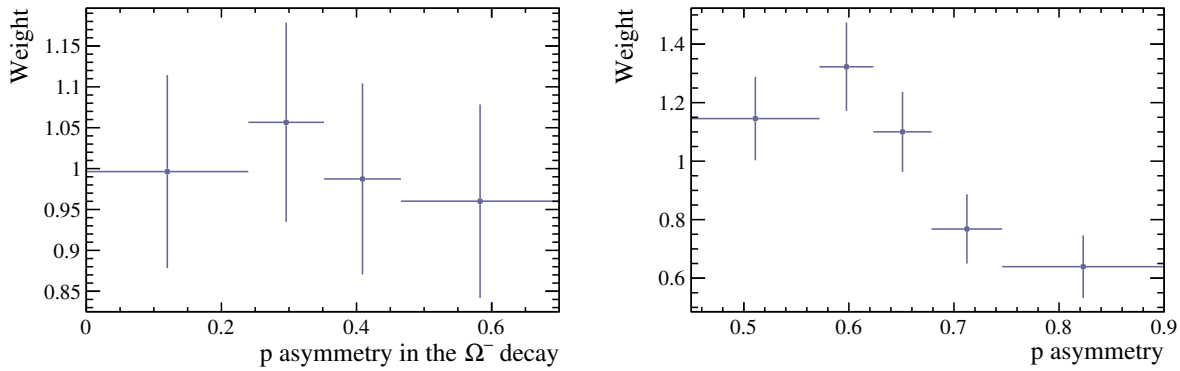
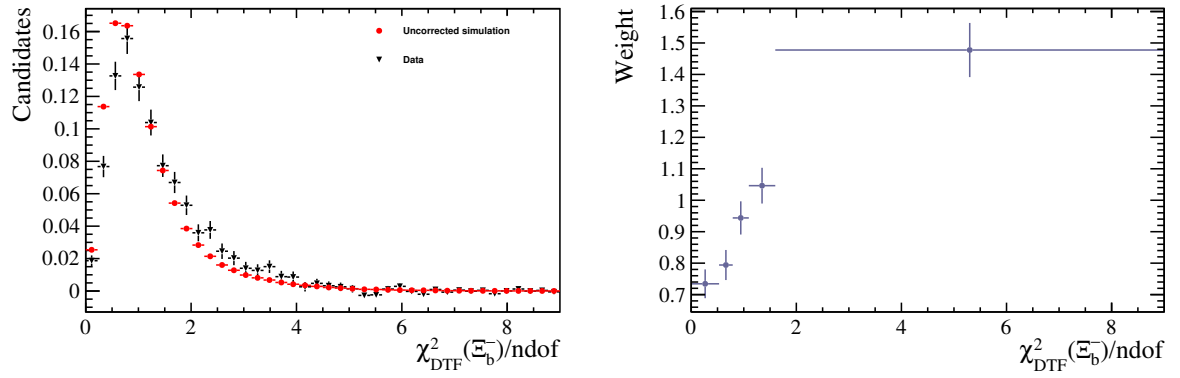
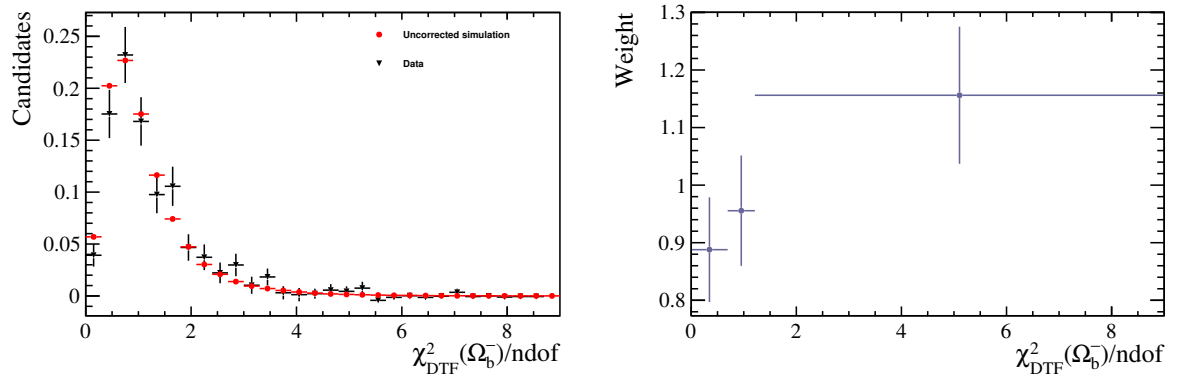


Figure D.7: Momentum asymmetry weights in hyperon decays from the Ω_b^- decay chain (left: Ω^-, right: Λ) for Run 2. The former is compatible with unity in all bins and therefore not applied.



(a) (b)

Figure D.8: a) χ^2_{DTF} distributions in Run 2 $\Xi_b^- \to \Xi^- J/\psi$ data and simulation, normalised to unit area; b) resulting χ^2_{DTF} correction for Run 2.



(a) (b)

Figure D.9: a) χ^2_{DTF} distributions in Run 2 $\Omega_b^- \to \Omega^- J/\psi$ data and simulation, normalised to unit area; b) resulting χ^2_{DTF} correction for Run 2.

Appendix E

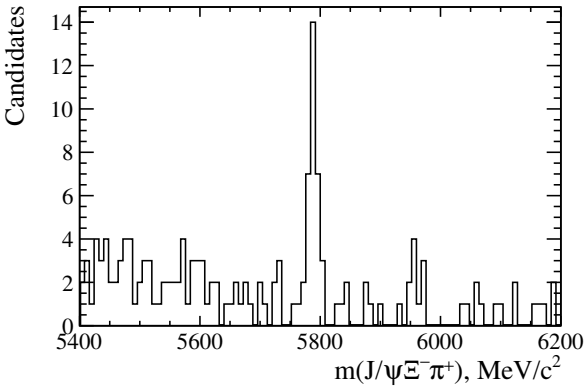
Additional information on rare Ξ_b^- decay modes

This appendix chapter for the study of the $\Xi_b^- \to \Xi^-\psi(2S)$ decay and the first search for the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay includes additional background studies, further details on the BDT setup, the calibration, the fit evaluation, as well as the efficiency determination. Lastly, some cross-checks on the ratio r^{ψ}_{BR} are presented and the plots for the determination of the expected signal significance for the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ are shown.

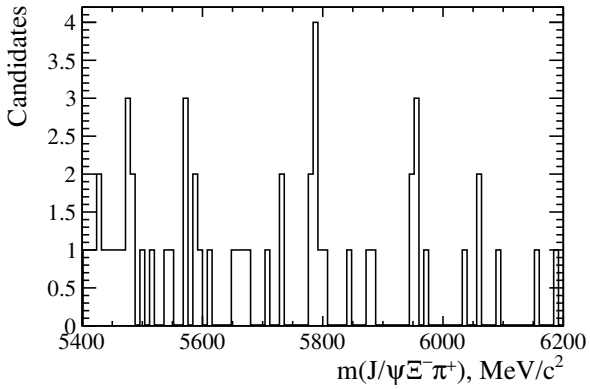
E.1 Study of $\Xi_b^0 \to \Xi^-\pi^+ J/\psi$ channel

A search for the $\Xi_b^0 \to \Xi^-\pi^+ J/\psi$ decay[1] is performed in order to prove the suppression of Ξ_b decays to the lowest-lying excited Ξ resonance ($\Xi(1530)$) due to symmetry constraints on the wave functions. The search follows the approach of the analysis in Ref. [199], but uses 2018 data (the old search was performed on Run 1 data) and the full selection as per Tab. 5.1 is applied. Due to the similarity in the final state particles and the topology of the decay with the $\Xi_b^- \to \Xi^- J/\psi$ one (one extra pion, that is always a long track) the efficiency of the selection is expected to be similar. A tight BDT cut at 0.95 is applied, and the decay time of Ξ^- is required to exceed 5 ps. The resulting distribution with mass constraints applied to Ξ^-, Λ and J/ψ, is shown in Fig. E.1(a). It can be seen that the signal peak is rather small, with only few tens events in the full 2018 data. This is to be compared to more than 1000 events seen in the $\Xi_b^- \to \Xi^- J/\psi$ channel in the same dataset. Even taking into account lower acceptance and vertexing efficiency for the six-track final state compared to the five-track one, the small yield suggests a small branching fraction of the $\Xi_b^0 \to \Xi^-\pi^+ J/\psi$ decay. The dataset is then split into regions of the $\Xi^-\pi^+$ invariant mass to prove the suppression of the lowest resonance. It can be seen from Fig. E.1(b-c) that most of signal is indeed found above $m(\Xi^-\pi^+) > 1800\,\text{MeV}/c^2$ and is not coming from the lowest excited states.

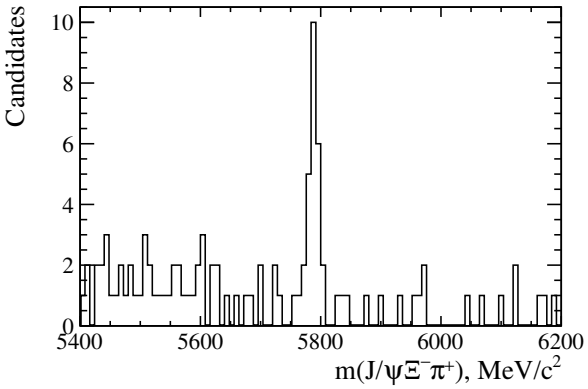
[1] This choice is made since isospin partner $\Xi_b^- \to \Xi^-\pi^0 J/\psi$ is much harder to reconstruct due to the neutral pion in the decay.



(a) 2018 data



(b) $m(\Xi^-\pi^+) < 1800\,\mathrm{MeV}/c^2$



(c) $m(\Xi^-\pi^+) > 1800\,\mathrm{MeV}/c^2$

Figure E.1: Invariant mass distribution of $J/\psi\Xi^-\pi^+$ data, with the Ξ_b^0 mass peak visible. The bottom row shows data split in two regions of the $\Xi^-\pi^+$ invariant mass.

E.2 Supplementary information on the BDT classifier

This section provides additional information on the BDT classifier used to search for the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ decay and study the $\Xi_b^- \to \Xi^-\psi(2S)$ branching fraction.

E.2.1 Variables used for the BDT training

After a dedicated study, 15 variables were used for the BDT training. The list of variables includes:

- $\chi^2_{DTF}/n_{dof}(\Xi_b^-)$
- $\chi^2_{vertex}(\Xi^-)$
- $\chi^2_{IP}(\Xi^-)$
- $p_T(\Xi_b^-)$
- $\eta(\Xi^-)$
- $\chi^2_{IP}(\pi_\Xi)$
- $\chi^2_{vertex}(\Xi_b^-)$
- $\chi^2_{IP}(\Xi_b^-)$
- $\chi^2_{IP}(\mu^+) + \chi^2_{IP}(\mu^-)$

- $(\chi^2_{IP})^{min}(\pi_\Lambda, p)$
- $\cos\theta(\Xi^-)$
- $p_T^{min}(\mu^+, \mu^-)$
- $DIRA(\Xi_b^-)$
- $\cos\theta(\mu^+)$
- $p(p)$

where the two angles, $\theta(\mu^+)$ and $\theta(\Xi^-)$, are defined as the opening angle between the muon (Λ) momentum in the dimuon (Ξ^-) rest frame and the direction of the dilepton (Ξ^-) in the Ξ_b^- rest frame.
The distributions of these variables are compared between the signal and background proxy samples.

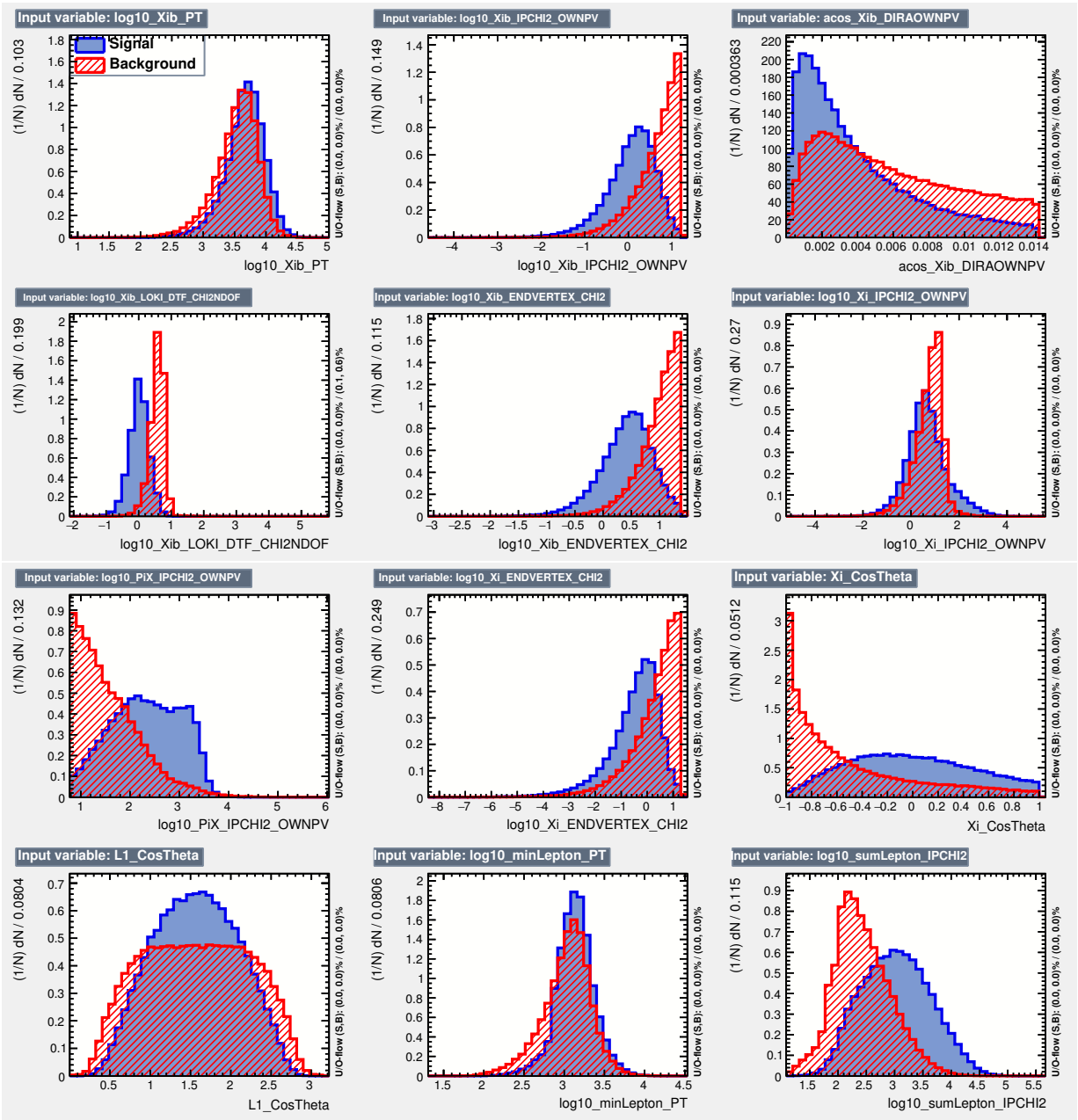


Figure E.2: Part 1

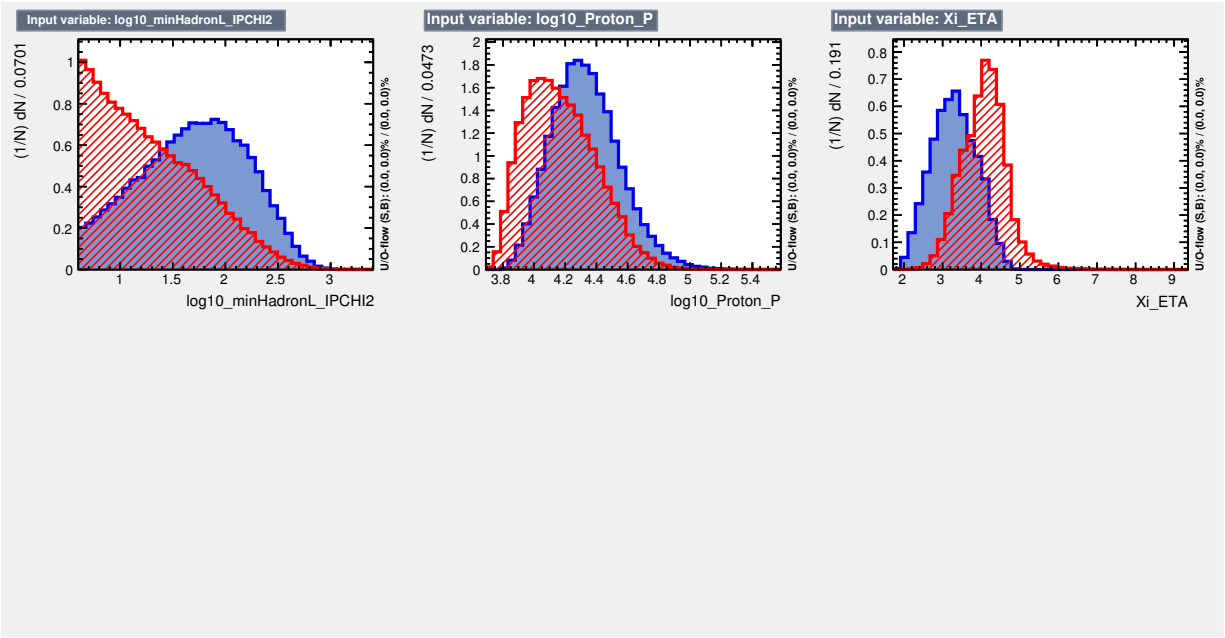


Figure E.3: Part 2

Additionally, the ranking of the input variables according to their importance for the BDT training is provided.

```
--------------------------------------------------------------------
: Rank : Variable                          : Variable Importance
: ----------------------------------------------------------------
:    1 : log10(abs(Xib_LOKI_DTF_CHI2NDOF)) : 9.600e-02
:    2 : Xi_ETA                            : 8.793e-02
:    3 : log10(abs(PiX_IPCHI2_OWNPV))      : 8.340e-02
:    4 : log10(abs(Proton_P))              : 7.497e-02
:    5 : sumLepton_IPCHI2                  : 7.405e-02
:    6 : log10(abs(Xib_ENDVERTEX_CHI2))    : 7.021e-02
:    7 : minLepton_PT                      : 6.966e-02
:    8 : minHadronL_IPCHI2                 : 6.839e-02
:    9 : log10(abs(Xib_IPCHI2_OWNPV))      : 6.818e-02
:   10 : Xi_CosTheta                       : 6.739e-02
:   11 : log10(abs(Xi_IPCHI2_OWNPV))       : 5.988e-02
:   12 : acos(Xib_DIRA_OWNPV)              : 5.564e-02
:   13 : log10(abs(Xib_PT))                : 4.871e-02
:   14 : log10(abs(Xi_ENDVERTEX_CHI2))     : 4.284e-02
:   15 : acos(L1_CosTheta)                 : 3.274e-02
: ----------------------------------------------------------------
```

E.2.2 Correlation matrices of the BDT input variables

Similarly, Fig. E.4 provides the correlation matrices of the BDT input variables for the signal and background proxies.

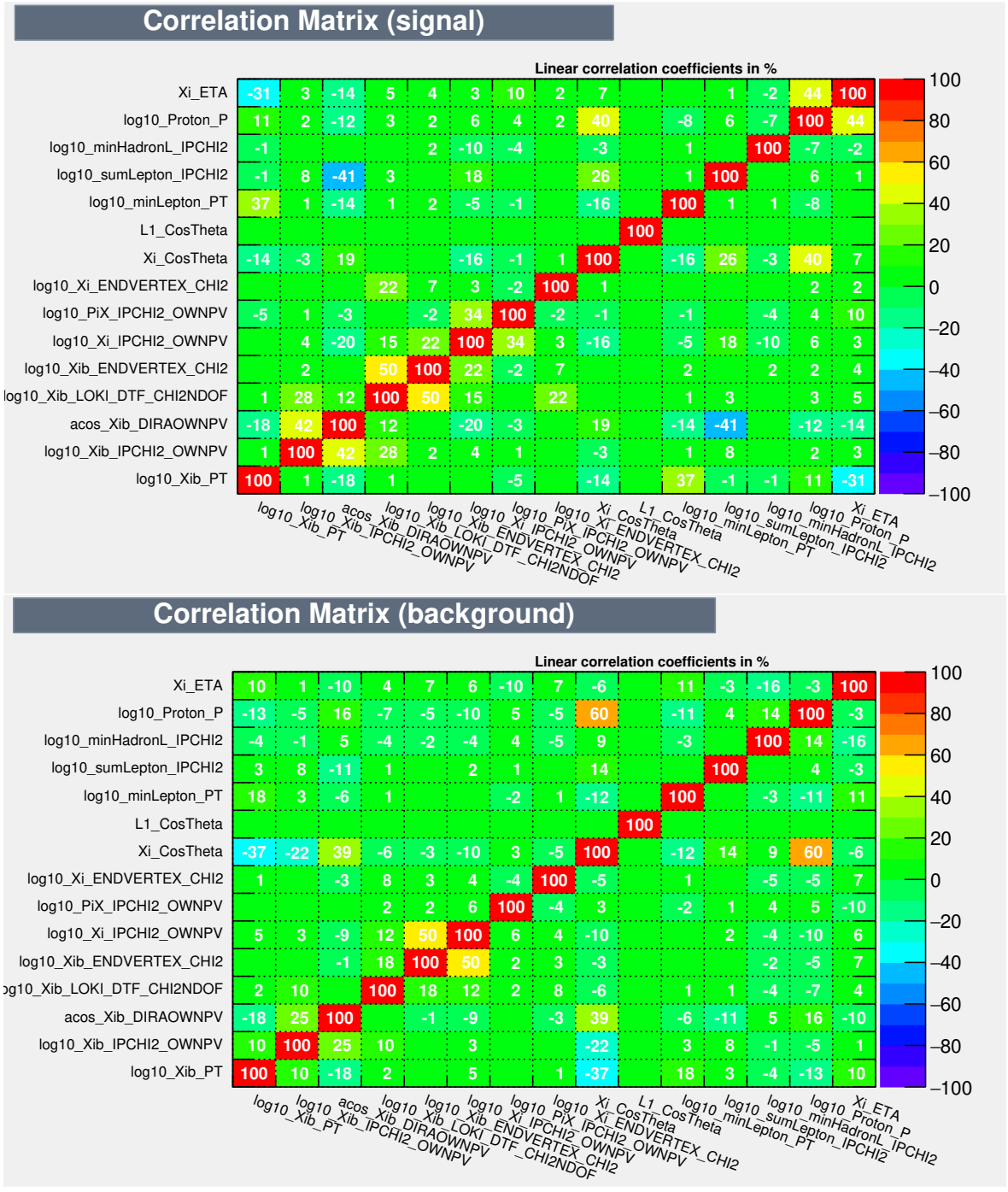


Figure E.4: Correlation matrices (signal and background)

E.3 Calibration of the simulation

In this section, some further details on the calibration of the simulation for the Ξ_b^- analyses discussed in Ch. 6 are provided. Fig. E.5 shows the comparison of the `nTracks` variable for the Run 1 $\Xi_b^- \to \Xi^- J/\psi$ simulation before (left) and after (right) rescaling with a factor 1.3 to match the distribution obtained from background subtracted data.

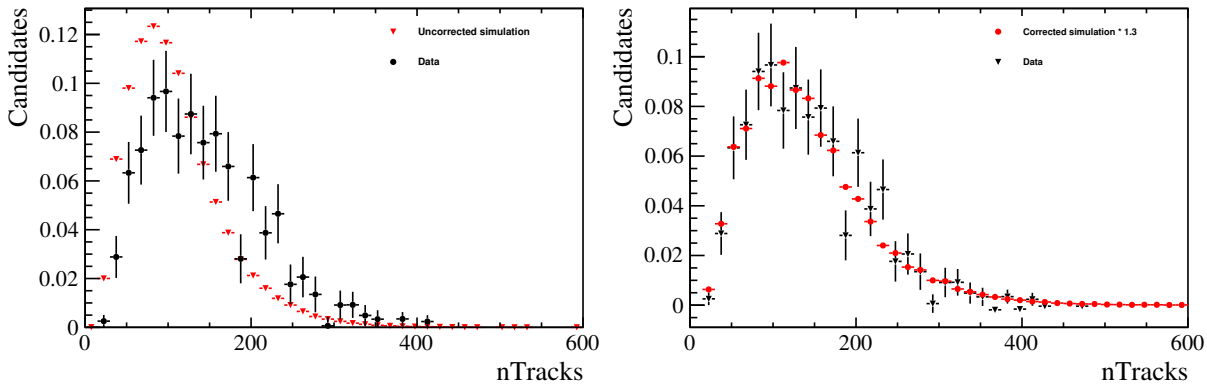


Figure E.5: nTracks variable distribution in the $\Xi_b^- \to \Xi^- J/\psi$ channel in Run 1, compared between data and raw simulation (left), and after rescaling by a factor 1.3 (right).

E.4 Fit evaluation

Each of the following figure contains to plots, the left being the fit to the truthmacthed, fully calibrated and selected simulation samples to obtain the parameters, which are fixed in the fit to data shown in the right plots. Fig. E.6 shows the fits for the Run 1 dataset of the $\Xi_b^- \to \Xi^- J/\psi$ decay mode, and similarly Fig. E.7 the ones for Run 2. In both cases, the J/ψ mass constraints are applied. Similarly, the fits for the $\Xi_b^- \to \Xi^- \psi(2S)$ samples are shown in Fig. E.8 for Run 1 and Fig. E.9 for Run 2. However, in the data fit to Run 1, the width has been fixed to 1.1 times the width obtained in simulation. This value is chosen after performing different fits to different subsets of data. The obtained values of the ratio data/simulation for the signal width are presented in Tab. E.1. It can be seen that simulation underestimates the value of the resolution by about ten percent, regardless of whether the J/ψ mass constraint is applied. The evaluation of the effect of removing the J/ψ and $\psi(2S)$ mass constraint on the respective sample are studied with the Run 2 samples only. The corresponding fits can be found in Fig. E.10 for the $\Xi_b^- \to \Xi^- J/\psi$ samples and in Fig. E.11 for the $\Xi_b^- \to \Xi^- \psi(2S)$ ones. In both cases, the obtained yields agree with the ones from the fits with the charmonia mass constraints and only the signal width get wider.

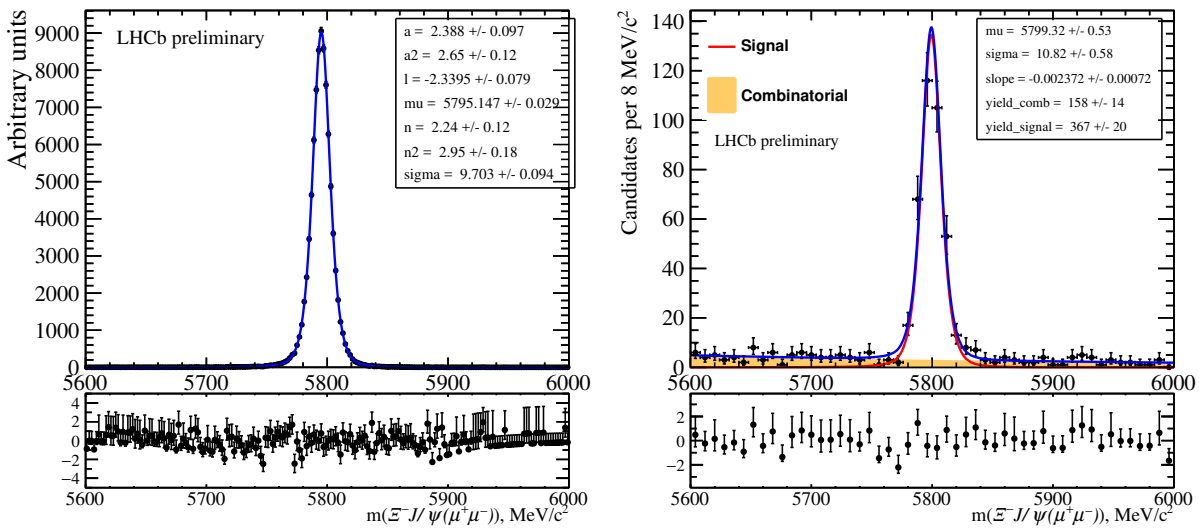
For completeness Tab. E.2 obtains the signal yields of the $\Xi_b^- \to \Xi^- J/\psi$ fits determined with different Run 2 subsamples with and without width kept floating, proving that the do not differ.

Finally, the fits to $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ signal simulation in the three q^2 regions are shown in

Fig. E.12. For simplicity, here only the 2011-12 (Run 1) simulated dataset is shown, since the models for the 15-16 and 17-18 datasets are very similar.

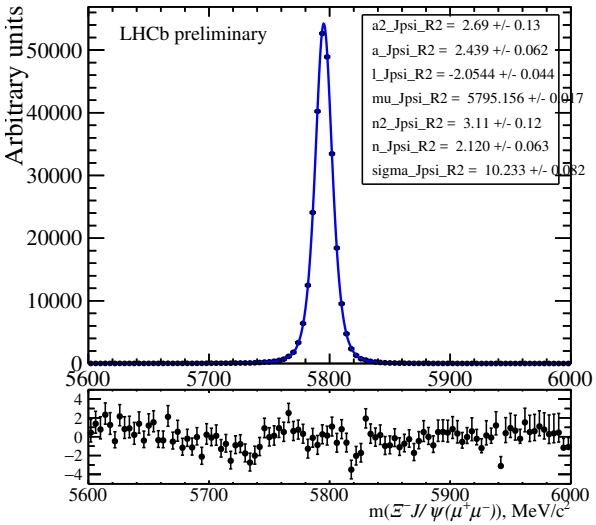
Table E.1: Ratios of signal widths in data and simulation.

Constraint	Dataset	Width ratio
J/ψ	DDD (Run 2)	1.07 ± 0.04
	LDD (Run 2)	1.13 ± 0.04
	DDD + LDD (Run 2)	1.10 ± 0.03
	LLL (Run 2)	1.14 ± 0.10
	DDD + LDD + LLL (Run 2)	1.10 ± 0.03
	DDD + LDD (Run 1)	1.11 ± 0.06
	DDD + LDD + LLL (Run 1)	1.12 ± 0.05
not J/ψ	DDD + LDD (Run 2)	1.10 ± 0.03
	DDD + LDD + LLL (Run 2)	1.10 ± 0.03

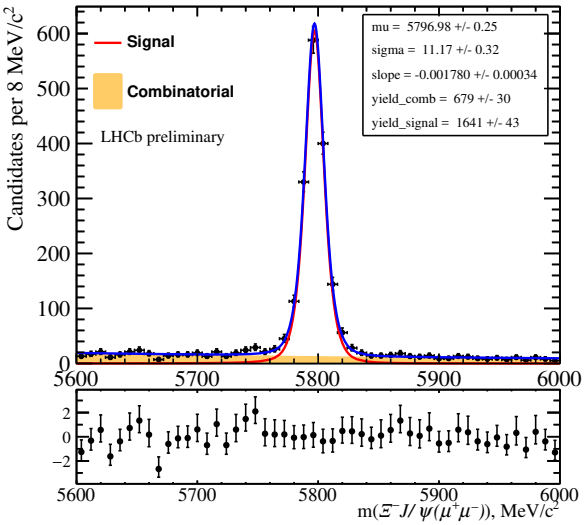


(a) All three categories

Figure E.6: Invariant mass shapes of the simulated $\Xi_b^- \to \Xi^- J/\psi(\to \mu^+\mu^-)$ signal, with the J/ψ mass constraint applied, for the Run 1 conditions.

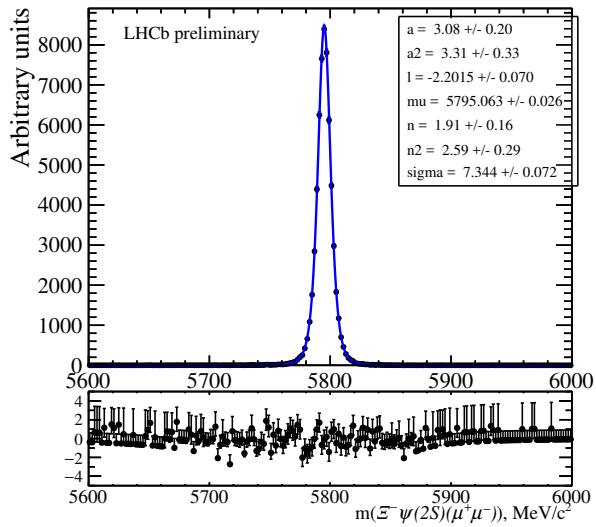


(a) All three categories



(b) All three categories

Figure E.7: Invariant mass shapes of the simulated $\Xi_b^- \to \Xi^- J/\psi(\to \mu^+\mu^-)$ signal, with the J/ψ mass constraint applied, for the Run 2 conditions.



(a) All three categories

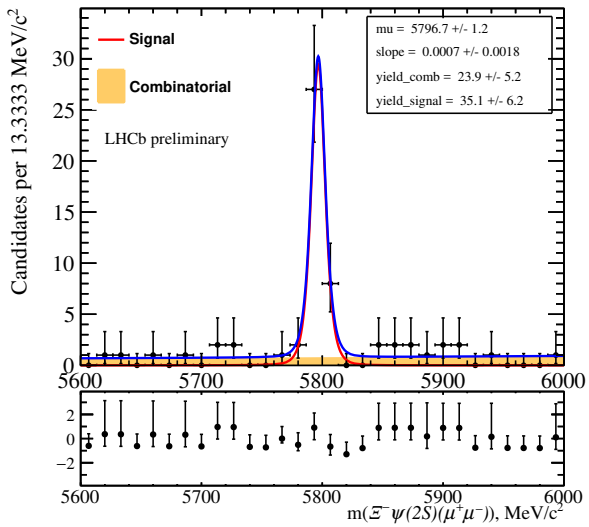
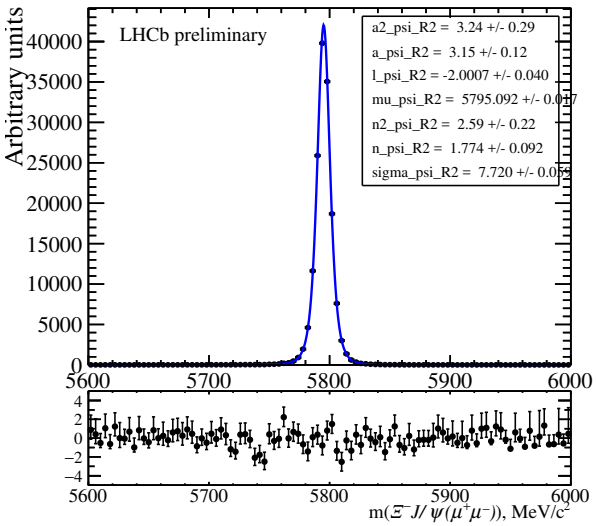
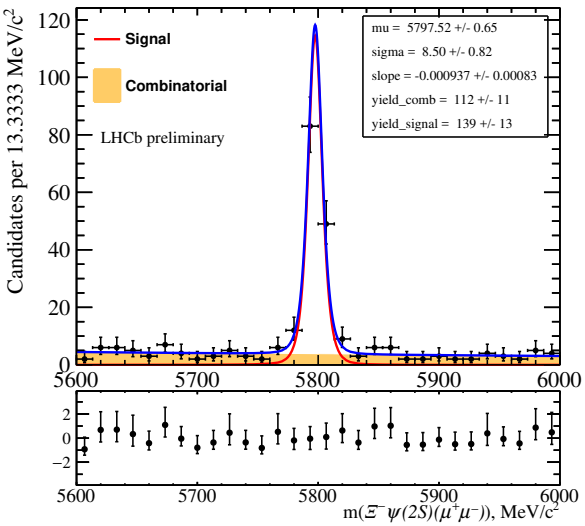


Figure E.8: Invariant mass shapes of the $\Xi_b^- \to \Xi^-\psi(2S)(\to \mu^+\mu^-)$ sample, with the $\psi(2S)$, Ξ^- and Λ mass constraints applied, for the Run 1 conditions, for simulation (left) and data (right). The signal width is fixed to a 10% larger value then the one obtained from simulation.

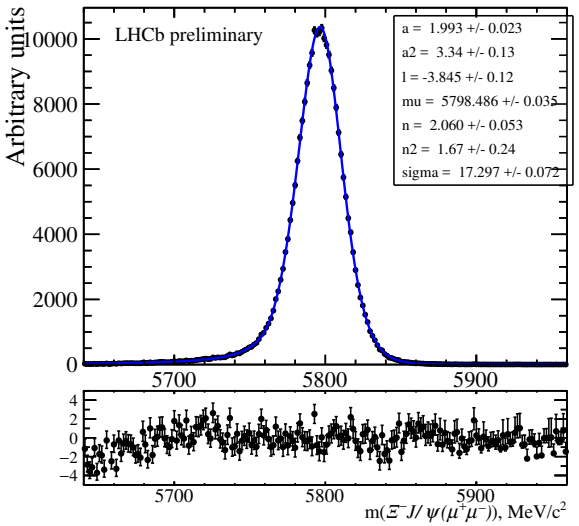


(a) All three categories



(b) All three categories

Figure E.9: Invariant mass shapes of the $\Xi_b^- \to \Xi^-\psi(2S)(\to \mu^+\mu^-)$ sample, with the $\psi(2S)$, Ξ^- and Λ mass constraints applied, for the Run 2 conditions, for simulation (left) and data (right). The signal width is allowed to float in fits to data.



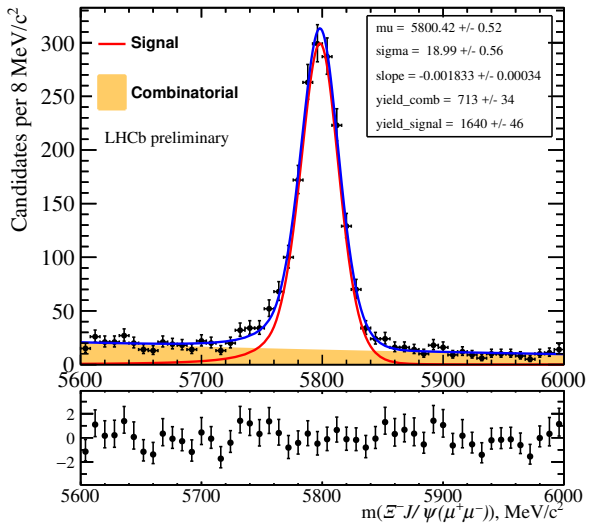


Figure E.10: Invariant mass fits of $\Xi_b^- \to \Xi^- J/\psi$ samples, with no J/ψ mass constraint applied (but the Ξ^- and Λ constraints still applied), for the Run 2 conditions in simulation (left) and data (right).

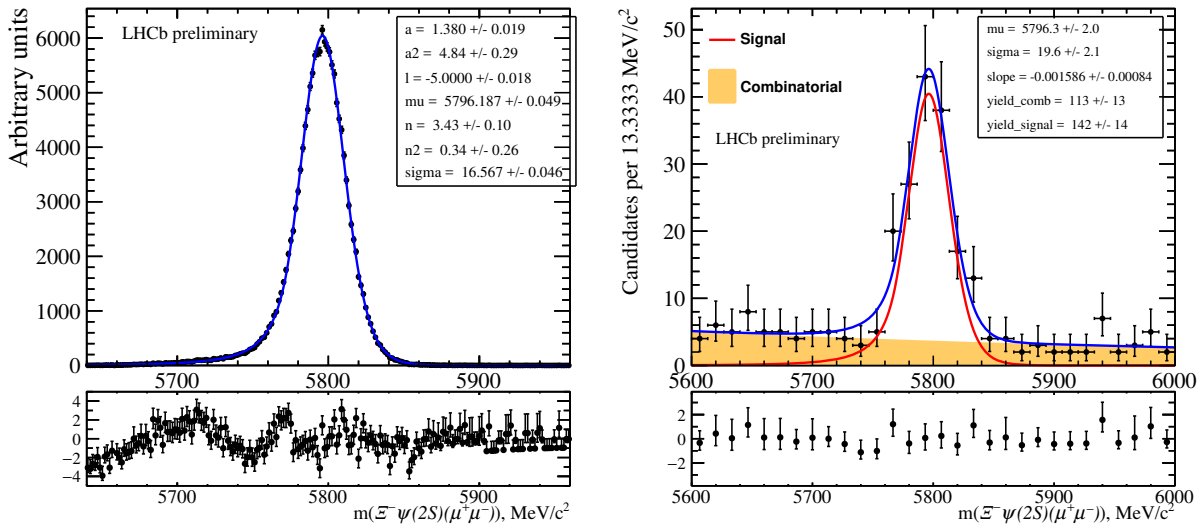
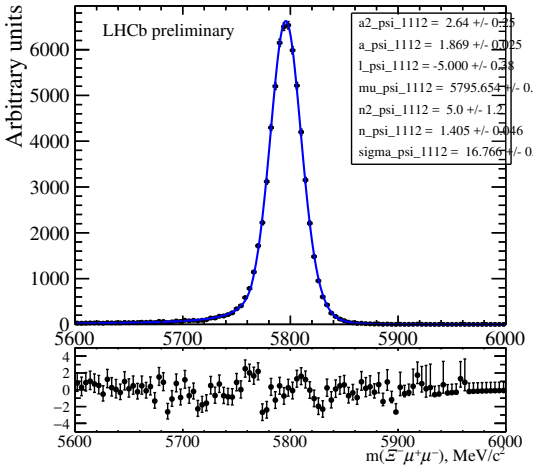


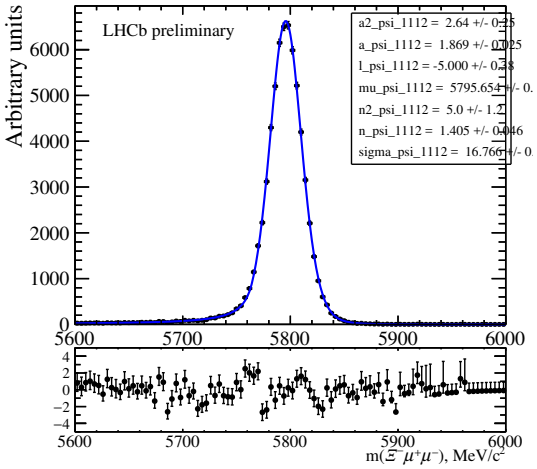
Figure E.11: Invariant mass fits $\Xi_b^- \to \Xi^- \psi(2S)$ samples, with no $\psi(2S)$ mass constraint applied (but Ξ^- and Λ constraints still applied), for the Run 2 conditions for simulation (left) and data (right).

Table E.2: $\Xi_b^- \to \Xi^- J/\psi$ signal yields obtained in fits under the nominal fit model. Two scenarios are considered: signal width is kept floating; or is fixed to be 1.1 times larger than the value from the relevant simulation sample.

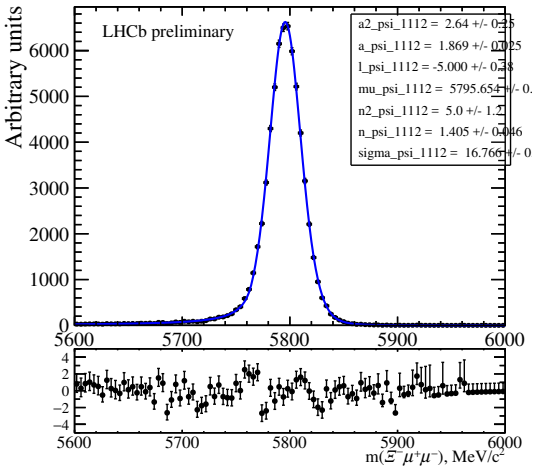
Dataset	Yield (free width)	Yield (fixed width)
DDD (Run 2)	739 ± 29	742 ± 28
LDD (Run 2)	715 ± 28	713 ± 28
DDD + LDD (Run 2)	1455 ± 40	1455 ± 40
LLL (Run 2)	186 ± 15	185 ± 15
DDD + LDD + LLL (Run 2)	1641 ± 43	1641 ± 43
DDD + LDD (Run 1)	320 ± 19	319 ± 19
DDD + LDD + LLL (Run 1)	367 ± 20	367 ± 20



(a) region 1



(b) region 2



(c) region 3

Figure E.12: Fits to the $\Xi_b^- \to \Xi^-\mu^+\mu^-$ signal simulation, for the three q^2 regions for the Run 1 sample.

E.5 Details on the efficiencies

The (average over q^2) generator-level efficiency is provided as a single number for each data taking period and magnet polarity are provided in Tab. E.3. The efficiencies for two magnet polarities are averaged. The large difference between efficiencies between the two charmonium channels indicates a significant trend with q^2. Therefore, a dedicated study is performed to obtain the q^2 dependence of the generator-level efficiency.
There are two obstacles with the simulation samples used in this analysis when try to access the q^2 trend. First, the samples have been produced in a dedicated *TightCut* setting which is used to reduce the computing resources by only processing long and downstream tracks within the LHCb acceptance. Furthermore, the samples have been filtered with the stripping line as described in Sec. 5.1.3. Therefore, two dedicated samples are produced to access the trend of the efficiency with q^2: one without the *TightCut* requirement and one with it applied. Both samples have not been filtered and there is a dedicated sample produced for Run 1 and Run 2 conditions. Additionally, the decays have only been matched to the Ξ^- decay and not for the full decay chain. This is done to correctly account for material interactions, more details can be found in App. E.5.3. Both distributions are normalised to unity, divided in 15 equally populated bins in q^2 and the ratio of them is taken. Finally, to restore the proper normalisation, this resulting efficiency map is multiplied by the 'average' generator-level efficiency provided in Tab. E.3.
The summary of the (average over q^2) filtering efficiencies for all year and channels, merged for the magnet polarities can be found in Tab. E.5 and the normalised to the $\Xi_b^- \to \Xi^- J/\psi$ filtering efficiency ones in Tab. E.6. The q^2 dependence of the filtering efficiency is studied with the unfiltered samples, produced for the generator-level efficiency study, by passing it through the filtering script. The ratio of the distributions before and after passing the filtering script enables to obtain the correctly normalised efficiency maps.

Table E.3: Generator-level efficiencies, %.

Dataset	$\Xi_b^- \to \Xi^- J/\psi$	$\Xi_b^- \to \Xi^- \psi(2S)$	$\Xi_b^- \to \Xi^- \mu^+ \mu^-$
2011	7.34 ± 0.02	7.95 ± 0.02	7.15 ± 0.02
2012	7.44 ± 0.01	8.09 ± 0.02	7.22 ± 0.01
2015	7.63 ± 0.02	8.33 ± 0.02	7.39 ± 0.02
2016	7.66 ± 0.02	8.36 ± 0.02	7.42 ± 0.02
2017	7.66 ± 0.02	8.34 ± 0.02	7.44 ± 0.02
2018	7.64 ± 0.02	8.32 ± 0.02	7.43 ± 0.02

The q^2 distribution of the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay is not correctly modelled in simulation due to the missing form factor prediction for a dedicated physics model. However, it is expected to be very similar to the one of the $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ decay. Therefore, the ratio of the generator-level simulation of the $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ and $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay displayed in the left plot of Fig. E.13 are taken to calculate the q^2 weight shown in the right plot to correct for the q^2 shape. This q^2 weight is then added to the generator-level sample of $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ without any acceptance cuts applied to obtain the total efficiency-ratio averages over each of three given q^2 bins.

Table E.4: Ratio of generator-level efficiencies with respect to the $\Xi_b^- \to \Xi^- J/\psi$ mode. The values are consistent between data taking periods.

Dataset	$\Xi_b^- \to \Xi^- \psi(2S)$	$\Xi_b^- \to \Xi^- \mu^+ \mu^-$
2011	1.083 ± 0.004	0.974 ± 0.003
2012	1.087 ± 0.003	0.970 ± 0.003
2015	1.092 ± 0.004	0.968 ± 0.003
2016	1.093 ± 0.004	0.969 ± 0.003
2017	1.089 ± 0.004	0.970 ± 0.004
2018	1.090 ± 0.004	0.973 ± 0.004

Table E.5: Filtering efficiencies, %

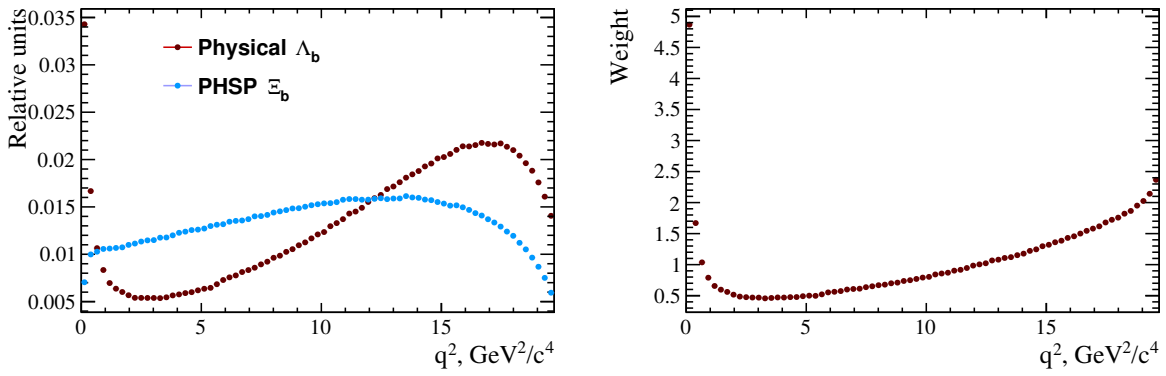
Channel	$\Xi_b^- \to \Xi^- J/\psi$	$\Xi_b^- \to \Xi^- \psi(2S)$	$\Xi_b^- \to \Xi^- \mu^+ \mu^-$
2011	14.55 ± 0.02	19.00 ± 0.03	14.31 ± 0.02
2012	14.48 ± 0.02	19.00 ± 0.02	14.23 ± 0.02
2015	15.41 ± 0.02	20.06 ± 0.03	15.03 ± 0.02
2016	15.56 ± 0.01	20.26 ± 0.03	15.23 ± 0.02
2017	12.95 ± 0.01	17.28 ± 0.02	12.70 ± 0.01
2018	12.93 ± 0.01	17.26 ± 0.02	12.67 ± 0.01

Table E.6: Ratio of filtering efficiencies with respect to the $\Xi_b^- \to \Xi^- J/\psi$ mode. The values are consistent between 2011-16 data taking periods, and slightly different in 2017-18 due to the difference in the stripping selection.

Dataset	$\Xi_b^- \to \Xi^- \psi(2S)$	$\Xi_b^- \to \Xi^- \mu^+ \mu^-$
2011	1.306 ± 0.003	0.984 ± 0.002
2012	1.312 ± 0.002	0.982 ± 0.002
2015	1.302 ± 0.003	0.975 ± 0.002
2016	1.303 ± 0.002	0.979 ± 0.001
2017	1.335 ± 0.002	0.981 ± 0.001
2018	1.335 ± 0.002	0.980 ± 0.001

E.5.1 Total efficiencies split per track category

Tab. E.7 shows the total efficiencies split per track category, as they are used for certain cross-checks.



(a) (b)

Figure E.13: Comparison of q^2 distributions in the physical $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ [102] and phase-space $\Xi_b^- \to \Xi^-\mu^+\mu^-$ simulated samples at generator-level (left). Ratio of the two, taken as a q^2 weight (right). Note that the upper edge of phase-space (19.7 to 20.1 GeV$^2/c^4$) is excluded to avoid bias from a slight difference in kinematic ranges of the two decay modes. The weight of the last bin is assigned in that region.

Table E.7: Total efficiencies (including geometric, filtering, reconstruction, selection), for different track type configurations. These efficiencies are only used for cross-checks. All efficiencies below have to me multiplied by $\mathbf{10^{-3}}$.

Period	Channel	DDD	LDD	DDD+LDD	LLL	all three
Run 1	$\Xi_b^- \to \Xi^- J/\psi$	0.181	0.315	0.496	0.095	0.591
	$\Xi_b^- \to \Xi^- \psi(2S)$	0.228	0.414	0.643	0.140	0.782
	ratio $\psi(2S)/J/\psi$	1.262	1.320	1.296	1.467	1.323
Run 2	$\Xi_b^- \to \Xi^- J/\psi$	0.369	0.399	0.768	0.108	0.876
	$\Xi_b^- \to \Xi^- \psi(2S)$	0.478	0.524	1.001	0.156	1.157
	ratio $\psi(2S)/J/\psi$	1.295	1.325	1.304	1.445	1.321

E.5.2 Sources of efficiency trends in Study of q^2

The trend of the generator level efficiency is not flat as a function of q^2 due to the requirements on the daughters of long-lived particles to fall into the acceptance of LHCb. In order to evaluate the source of the efficiency trends with q^2 in more depth, the effect of certain cuts on the q^2 distribution in $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ simulation is studied. Fig. E.14 shows that if no cut is applied a significant non-flatness with q^2 is visible (black).

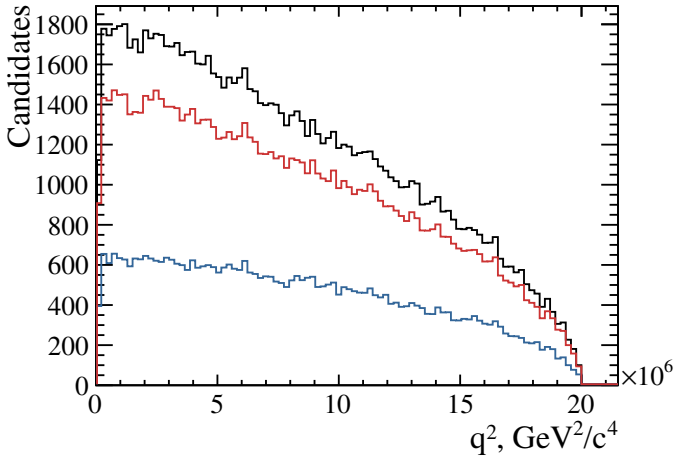


Figure E.14: Distribution of q^2 in unfiltered generator-level $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ sample. Black: no acceptance (nor any other) cuts applied. Red: decay vertex of Ξ^- and Λ required to have z coordinate below 3 m. Blue: same, with additional requirement of soft pions from Ξ^- and Λ to have $1.5 < \eta < 7.5$. One can see clearly that the efficiency is lower at low q^2.

E.5.3 Matching of the decay descriptor

When reconstructing the $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ decay, there are two ways: one where the decay is only matched to the Ξ^- hyperon and one where the subleading decays of the Ξ^- are also matched ("fully matched"). The top left plot in Fig. E.15 shows a visible deficit of events in the "fully-matched" sample being especially pronounced at Ξ^- pseudorapidity of 4-5, where a lot of material is located. This in it turn reflects in the measured trend of generator-level efficiency as a function of pseudorapidity as it can be seen in the top right plot of Fig. E.15. Since in the LHCb simulation software the decay of the Ξ^- is propagated from EvtGen to Geant4, if the material interaction happen before the Ξ^- decay, the fully-matched descriptor does not select these events anymore while the Xi-matched descriptor still does.

As the lack of fully-matched events correlates with the kinematics of the Ξ^-, it is natural that it also correlates with the q^2. It can clearly be seen in the bottom plot of Fig. E.15 that the generator-level efficiency of the $\Xi_b^- \to \Xi^- J/\psi$ and $\Xi_b^- \to \Xi^- \psi(2S)$ decays, extracted from the official generation tables, fit better the "Xi-matched" trend rather than the fully-matched one. Since the $\Xi_b^- \to \Xi^- J/\psi$ point is used for normalisation, Xi-matched samples need to be used for calculation of all generator-level and filtering efficiencies to avoid any (even q^2 dependent) bias. Since samples are not fully matched some weights

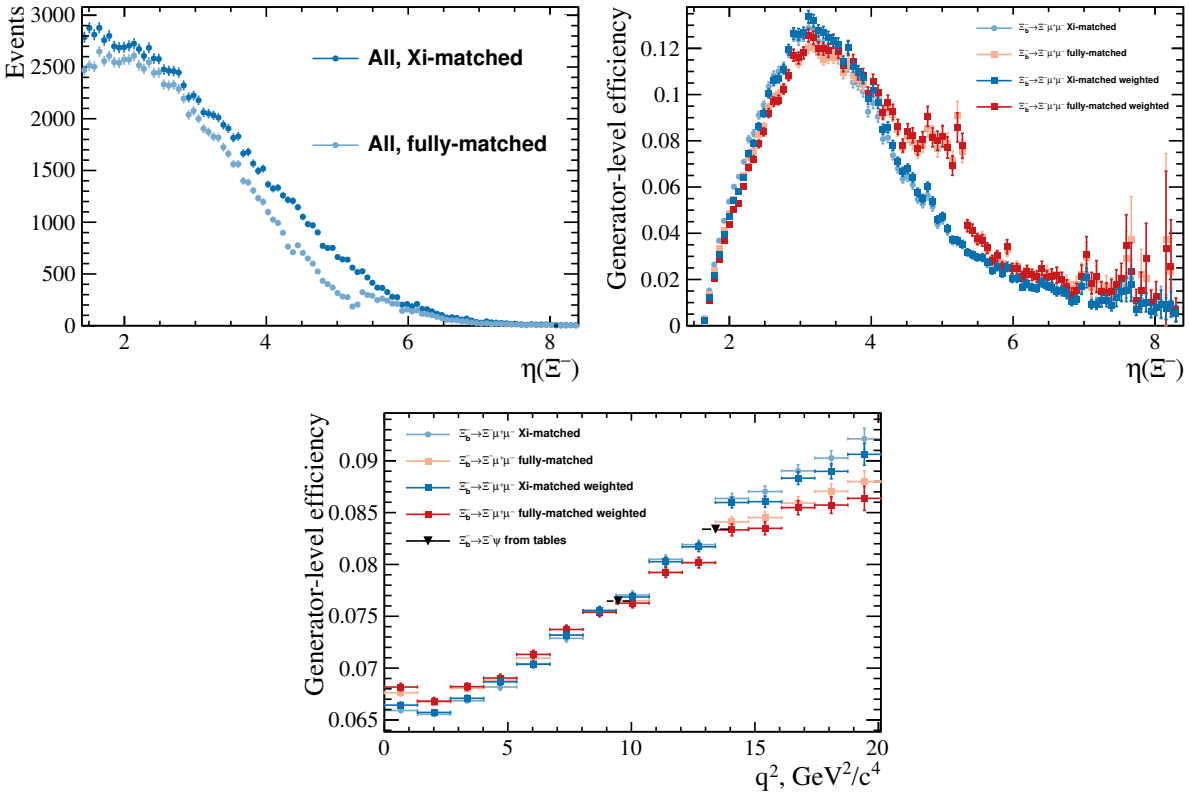


Figure E.15: Top left: distribution of the $\eta(\Xi^-)$ in the "fully-matched" and "Xi-matched" samples. Top right: generator-level (detector acceptance) efficiency measured in the "fully-matched" and "Xi-matched" samples. Bottom: generator-level efficiency as a function of q^2 in the "fully-matched" and "Xi-matched" samples. "Weighted" refers to the Ξ_b kinematic weighting.

like the momentum asymmetries cannot be applied, while they have only minor effects on the generator-level and filtering efficiency they are important for the selection efficiencies. Therefore, the selection efficiencies are calculated with the fully-matched and reweighted samples (especially to account for the BDT), and a systematic uncertainty is assigned for these choices.

E.6 Cross-checks for the stability of r^{ψ}_{BR}

The stability of the r^{ψ}_{BR} result is checked for the different categories of data used. Fig. E.16 compares the values of $\mathcal{B}(\Xi_b^- \to \Xi^-\psi(2S))/\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi)$ between the different data taking periods, and an excellent agreement is seen. Results are shown in Fig. E.17 for comparing LDD, DDD and LLL track type categories, as well as their combinations. It can be seen that the results are in agreement within $< 2\sigma$. Lastly, Fig. E.18 compares the results for Ξ_b^- and $\overline{\Xi}_b^+$ decays, which are also agreeing very well.

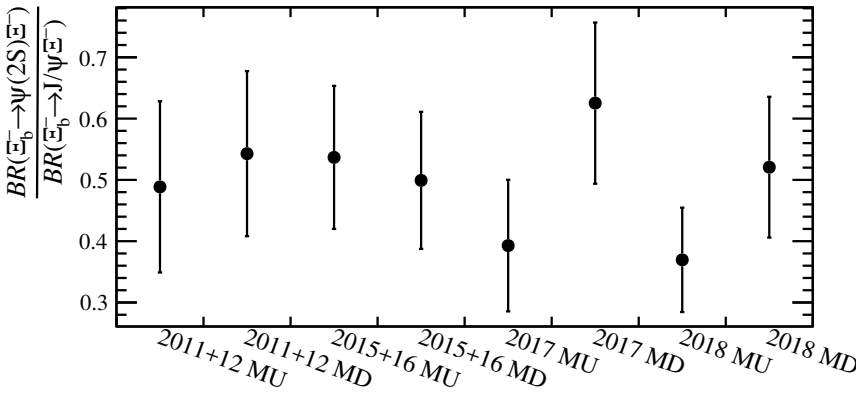


Figure E.16: Comparison of the $\mathcal{B}(\Xi_b^- \to \Xi^-\psi(2S))/\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi)$ values in different data taking periods (all track categories, uncertainties statistical only).

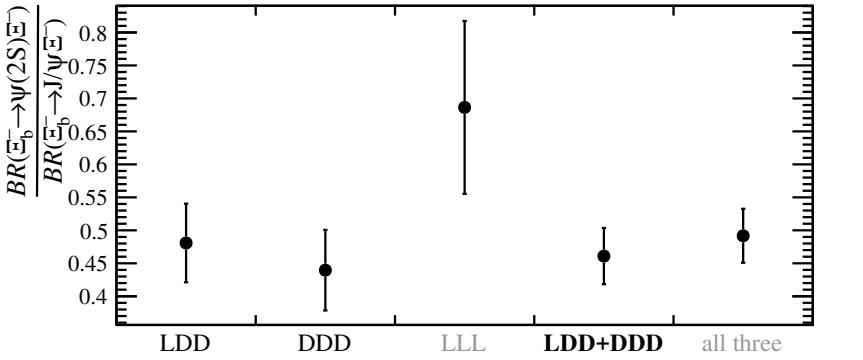


Figure E.17: Comparison of the $\mathcal{B}(\Xi_b^- \to \Xi^-\psi(2S))/\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi)$ values in different track type categories (uncertainties statistical only). Only the first three points are statistically independent.

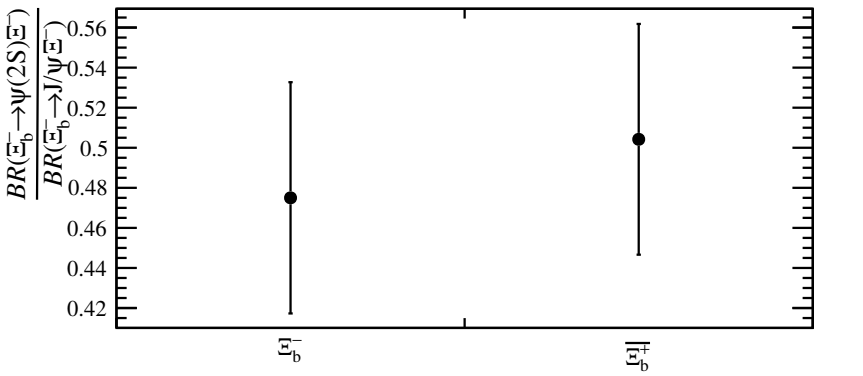
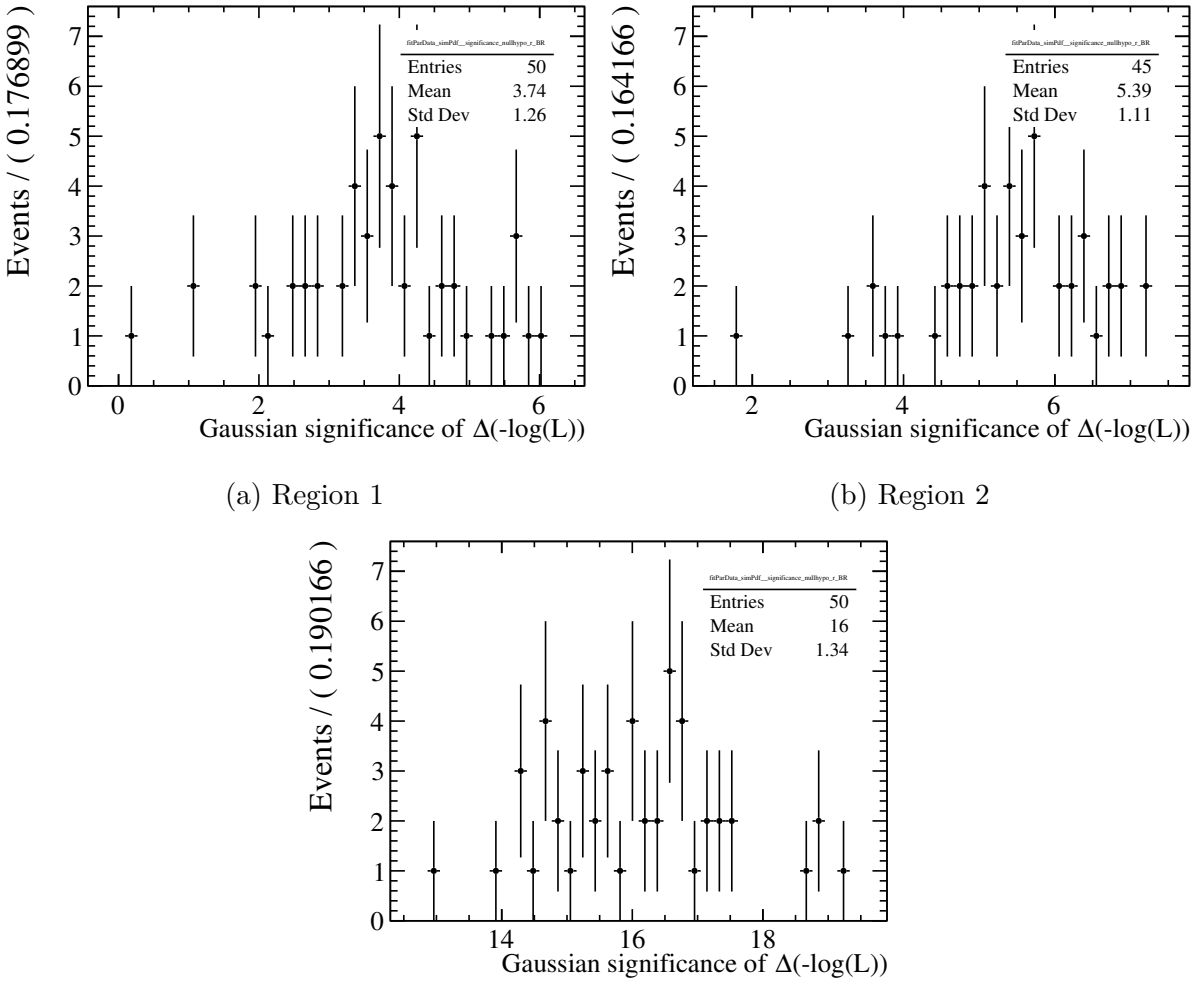


Figure E.18: Comparison of the $\mathcal{B}(\Xi_b^- \to \Xi^-\psi(2S))/\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi)$ values for Ξ_b decays and anti-Ξ_b decays (uncertainties statistical only).

E.7 Determination of the SM expected significance

The obtained means of the $\sqrt{-2\Delta LL}$ distributions from the 50 pseudoexperiments for each of the three q^2 regions are shown in Fig. E.19. It should be noted, that the SM predictions used to generate the pseudeexperiments have large uncertainties due to the assumptions on the $\Xi_b^- \to \Xi^- J/\psi$ branching fractions (see Sec. 6.2.2). Therefore, the true signal significance obtained after the unblinding might significantly differ.



(a) Region 1

(b) Region 2

(c) Region 3

Figure E.19: Expected (from 50 toys) $\Xi_b^- \to \Xi^- \mu^+ \mu^-$ signal significance in the three q^2 regions, assuming the SM rate from Ref. [27], converted into r_{BR} with our estimate of $\mathcal{B}(\Xi_b^- \to \Xi^- J/\psi)$; and using the efficiency values from Tab. 6.6. The means of these distributions for each region are about 3.7, 5.4 and 16 standard deviations, respectively.

Appendix F

Additional information on the study of Ξ_b^0 decays to charmonia

This section includes more information on the first study of the $\Xi_b^0 \to \Xi^0 J/\psi$ and $\Xi_b^0 \to \Xi^0 \psi(2S)$ decays.
The fits to simulation in order to constrain the tail parameters for the fully cut-based selection can be found in Fig. F.1.
Similarly, the fits to simulation in order to constrain the tail parameters for the selection including a BDT cut can be found in Fig. F.2. In both cases it should be noted, that the resolution is similar to that obtained from fits of the charged Ξ_b^- decays. However, due to the presence of the π^0 and its approximation to originate from the PV instead of the Ξ^0 decay vertex that cannot be reconstructed, the resolution is expected to be wider. Therefore, a more detailed study is needed to exclude possible differences in simulation with respect to the data due to issues with the truthmatching.

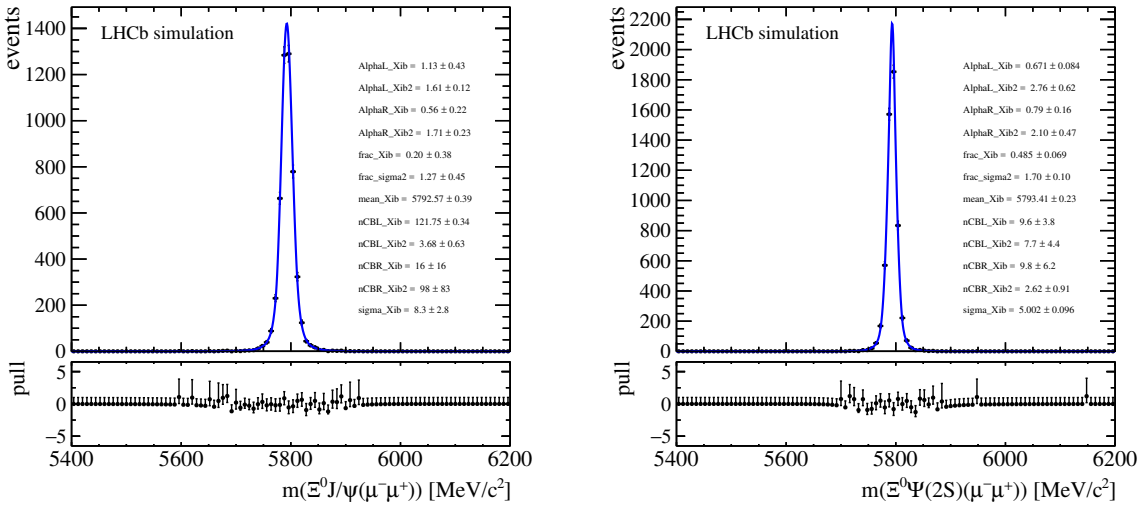


Figure F.1: Invariant mass distributions of the simulated $\Xi_b^0 \to \Xi^0 J/\psi$ (left) and $\Xi_b^0 \to \Xi^0 \psi(2S)$ (right) signals obtained with a fully cut-based selection.

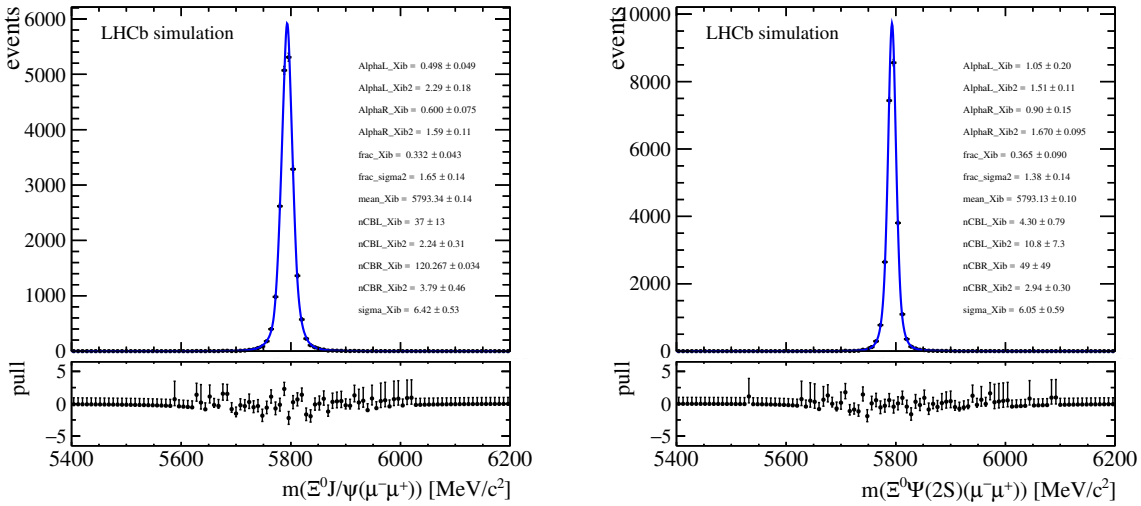


Figure F.2: Invariant mass distributions of the simulated $\Xi_b^0 \to \Xi^0 J/\psi$ (left) and $\Xi_b^0 \to \Xi^0 \psi(2S)$ (right) signals obtained with a selection including a BDT cut.

F.1 Details on the Ξ_b^0 BDT setup

The BDT discussed in this section was developed as part of a master thesis of Leandra Moeser [220]. Therefore, the plots have been taken from the thesis document.
The best working point of the BDT is determined with a grid search and found for 200 trees, a depth of 4 for each tree and a learning rate of 0.12.
It should be noted that the variable `Xib_DTF_Lpi0_Xi0_ctau_Best_PV` is obtained with the new BDT setup and refers to the Ξ^0 decay time calculated with applying a constraint on the Λ and π^0 masses as well as requiring the Ξ_b^0 to originate from the PV. The *Best_PV* indicated that the best primary vertex candidate has been selected for the latter constraint. Similarly, the variable the variable `Xib_DTF_XiLpi0_chi2_Best_PV` translated to the quality of the DTF fit χ^2_{DTF} with a Ξ^0, Λ and π^0 masses constrained to their known values, as well as a PV constraint applied to the Ξ_b^0 candidate.
Fig. F.3 displays the correlations for the final set of variables used in the BDT training for the signal (top) and background (bottom) sample. Additionally, the distributions of all variables for signal and background can be found in Fig. F.4 and F.5.
Finally, Fig. F.6 shows the determined cut values for the BDT selection with a Punzi FOM. Since the channel is unobserved the value $a = 5$ has been chosen. The used datasets are the $\Xi_b^0 \to \Xi^0 J/\psi$ data and since no peak was visible in the broad range (left), a zoomed plot is produced (right). The FOM maximum is found at the value 0.9925. With the same strategy the optimal BDT cut value for the $\Xi_b^0 \to \Xi^0 \psi(2S)$ sample is evaluated and determined to be at 0.990. The corresponding plots are shown in Fig. F.7.

F.2 Data-simulation agreement

While it is expected that the kinematics of the Ξ_b^0 baryon is mismodelled in the simulation and the PID variables need additional calibration, also the BDT input variables have been

checked for mismodelling. The obtained plots can be seen in Fig. F.8. While some of the variables show reasonable agreement, others like the transverse momentum of the Λ hyperon show disagreements. Therefore, a dedicated calibration of the simulation is

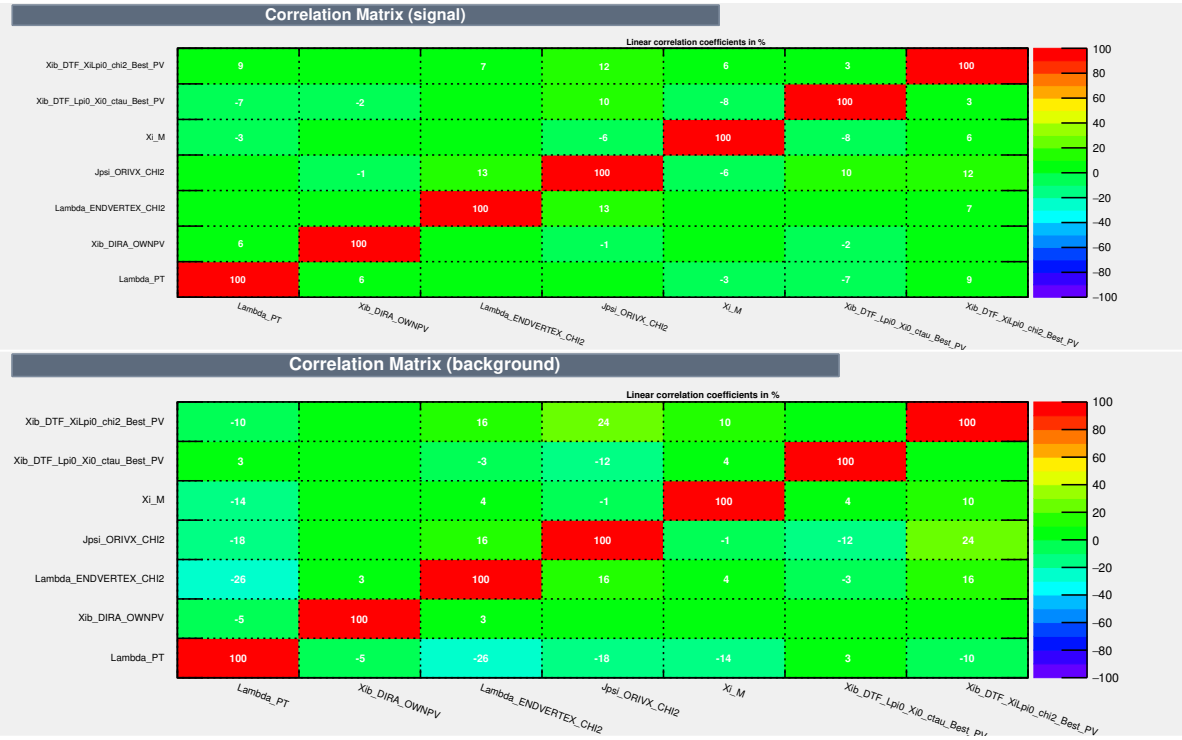


Figure F.3: Correlation matrix for the input variables of the BDT for signal (top) and background (bottom). Taken from [220].

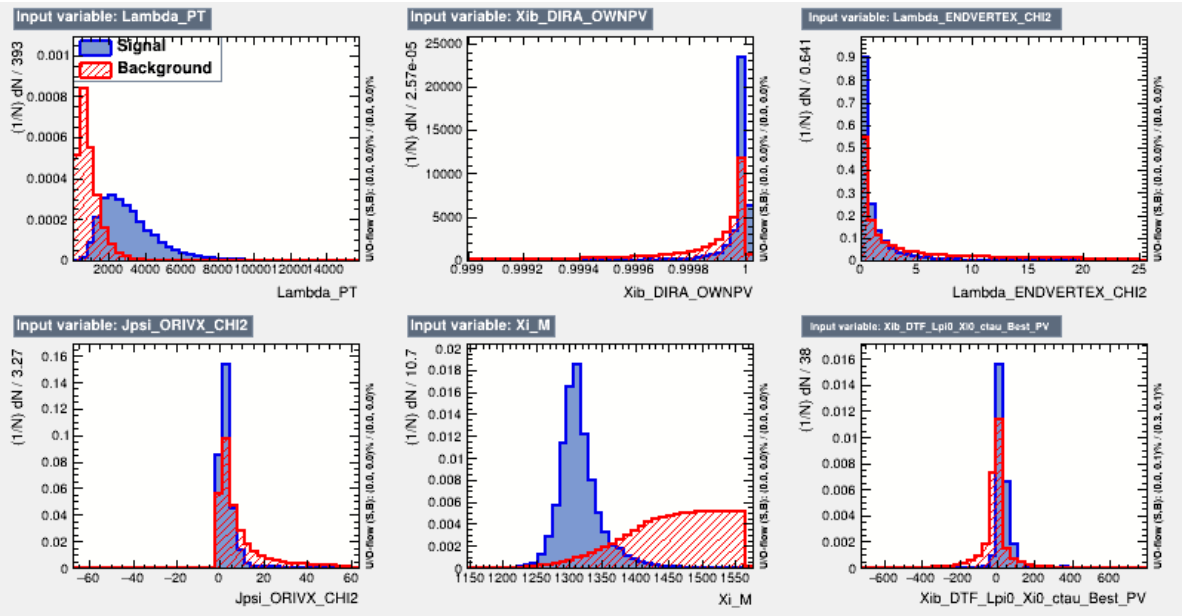


Figure F.4: Distribution of the input variables used in the BDT training for signal (blue) and background (red). Taken from [220]. (Part I)

needed in the future. It should be noted that the Ξ^0 mass is not fully uncorrelated to the Ξ_b^0 mass used to perform the sPlot procedure and obtain the background-subtracted data, due to the poor resolution in the studied decay.

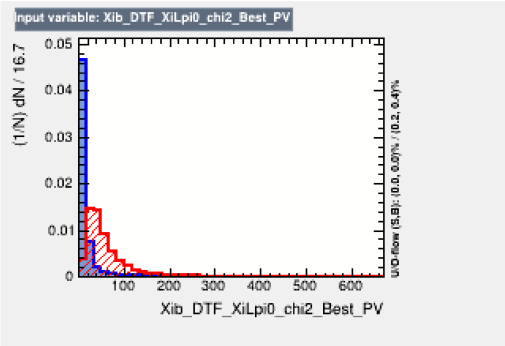


Figure F.5: Distribution of the input variables used in the BDT training for signal (blue) and background (red). Taken from [220]. (Part II)

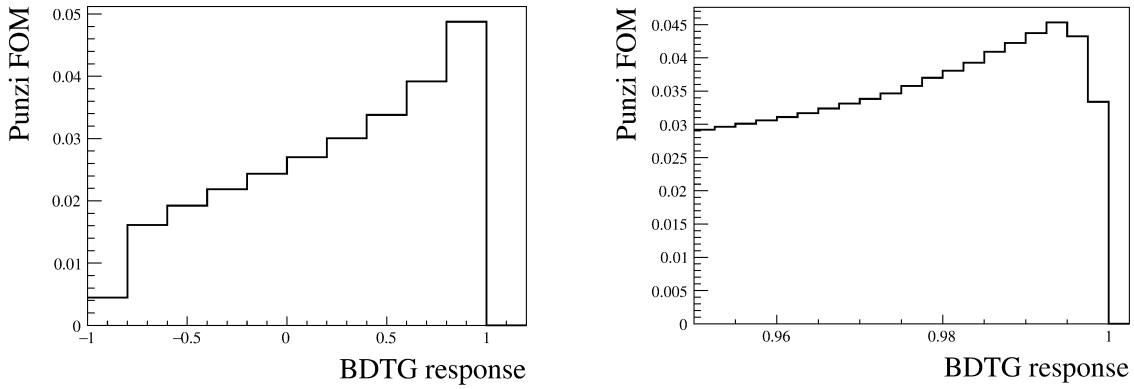


Figure F.6: Punzi FOM to determine the optimal BDT cut value for the $\Xi_b^0 \to \Xi^0 J/\psi$ decay (left). The full 2018 dataset is used and the final cut is determined by zooming into the Punzi FOM maximum region (right). Taken from [220].

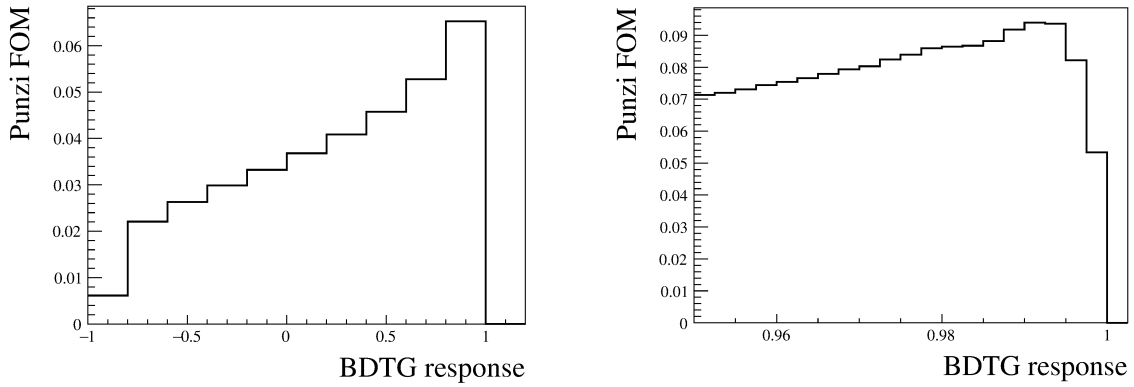


Figure F.7: Punzi FOM to determine the optimal BDT cut value for the $\Xi_b^0 \to \Xi^0 \psi(2S)$ decay (left). The full 2018 dataset is used and the final cut is determined by zooming into the Punzi FOM maximum region (right). Taken from [220].

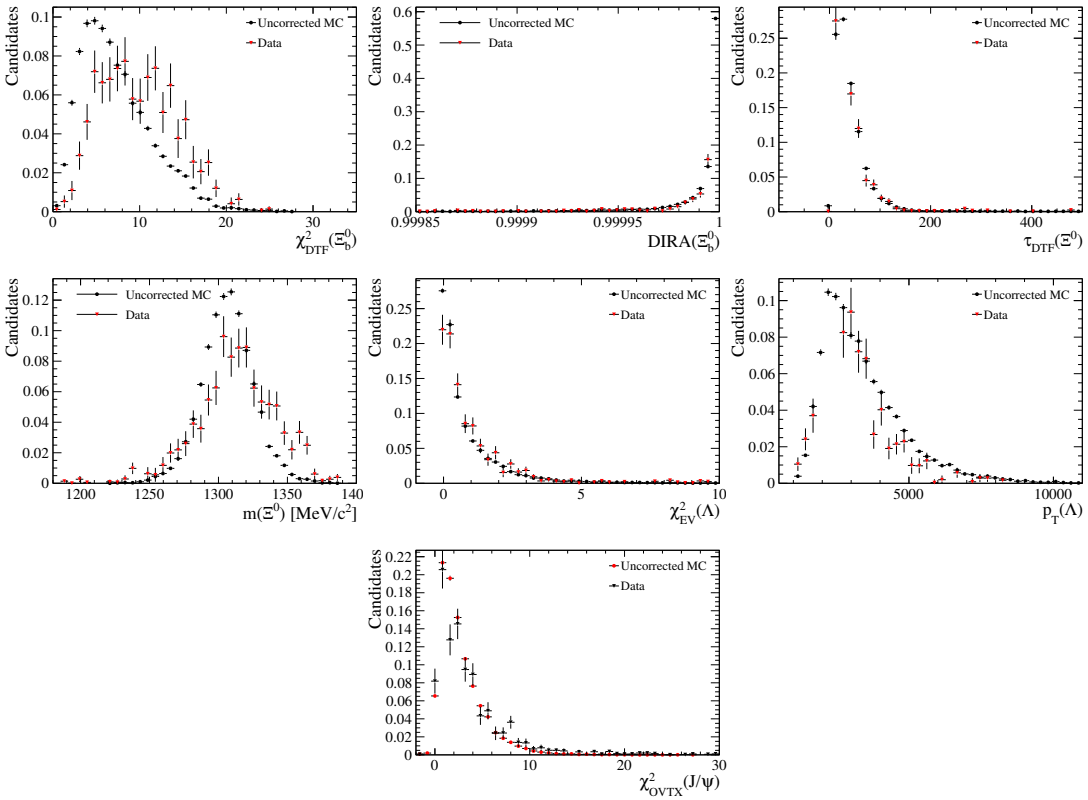


Figure F.8: Comparison of the background-subtracted data (red) and uncalibrated simulation (black) distributions of the variables used in the BDT training after the DTF fix.

Appendix G

Additional information for the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ angular analysis

This chapter provides additional information for the angular analysis of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay.

G.1 BDT classifier

The distributions of the variables used in the training of the BDT are compared between the signal and background proxy samples and displayed in Fig. G.1.
Additionally, the ranking of the input variables according to their importance for the BDT training is provided.

```
--------------------------------------------------------------------
: Rank : Variable                          : Variable Importance
: ----------------------------------------------------------------
:    1 : Lambdab_TAUCHI2                   : 1.014e-01
:    2 : Lambdab_DIRA_OWNPV                : 9.362e-02
:    3 : L12_IPCHI2_OWNPV                  : 9.354e-02
:    4 : Lambda_TAU                        : 8.572e-02
:    5 : log10(abs(Lambda_PT))             : 8.358e-02
:    6 : log10(abs(Lambdab_PT))            : 8.310e-02
:    7 : Lambdab_TAU                       : 7.915e-02
:    8 : Lambdab_LOKI_DTF_CHI2NDOF         : 7.706e-02
:    9 : log10(abs(Jpsi_IPCHI2_OWNPV))     : 7.640e-02
:   10 : Lambdab_ENDVERTEX_CHI2            : 7.605e-02
:   11 : minIPCHI2_PPi                     : 7.599e-02
:   12 : log10(abs(Proton_P))              : 7.442e-02
: ----------------------------------------------------------------
```

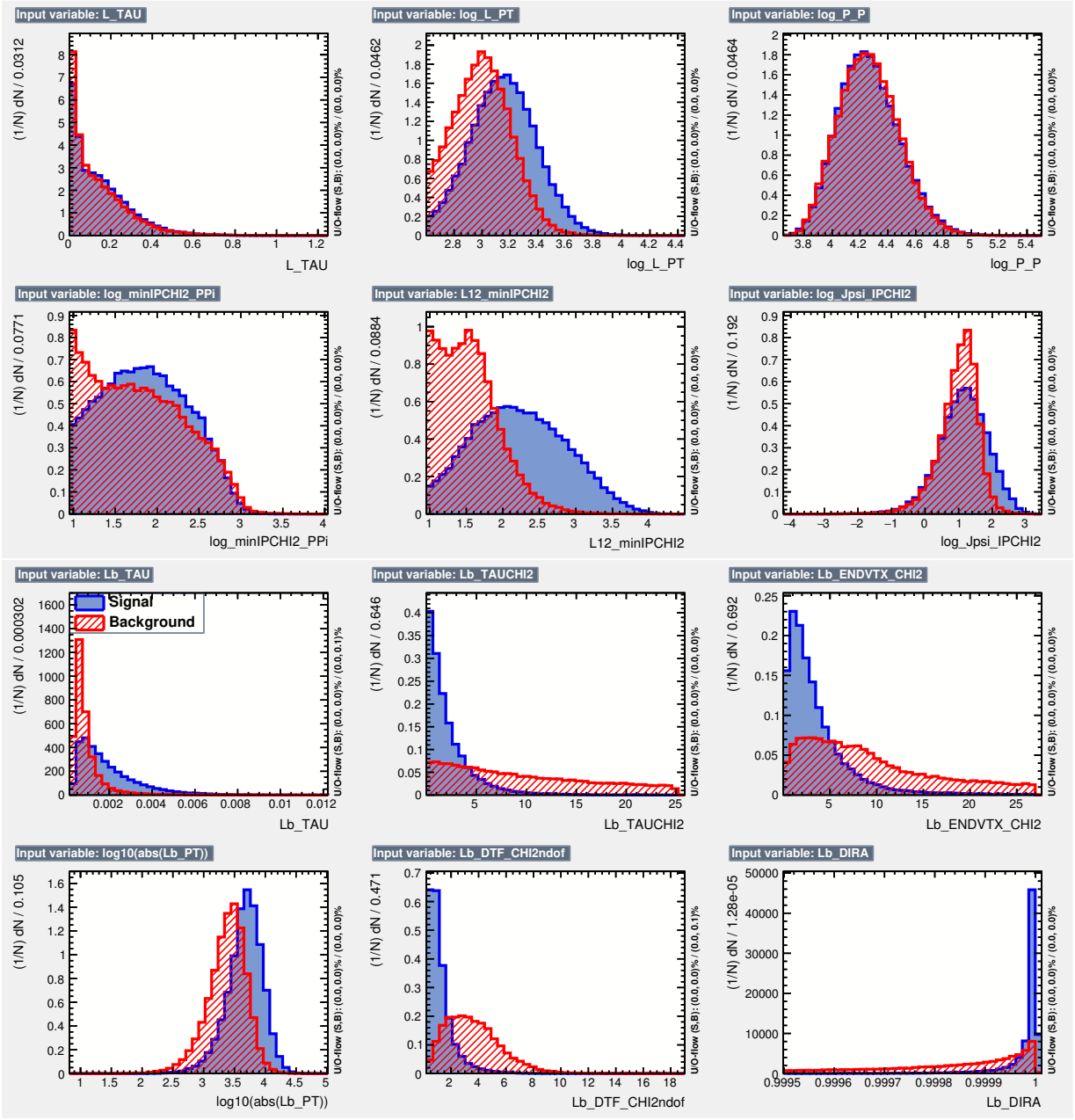


Figure G.1: Signal and background shapes of the input variables in the BDT of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ angular analysis

Finally, Fig. G.2 provides the correlation matrices of the BDT input variables for the signal and background proxies.

G.2 Constraint on the shape of the signal distribution

The shape of the signal contributions in the mass fits are constrained using the corresponding physics simulation samples of the $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decay. An example fit of the signal PDF to the Run 2 simulation can be found in Fig G.3.

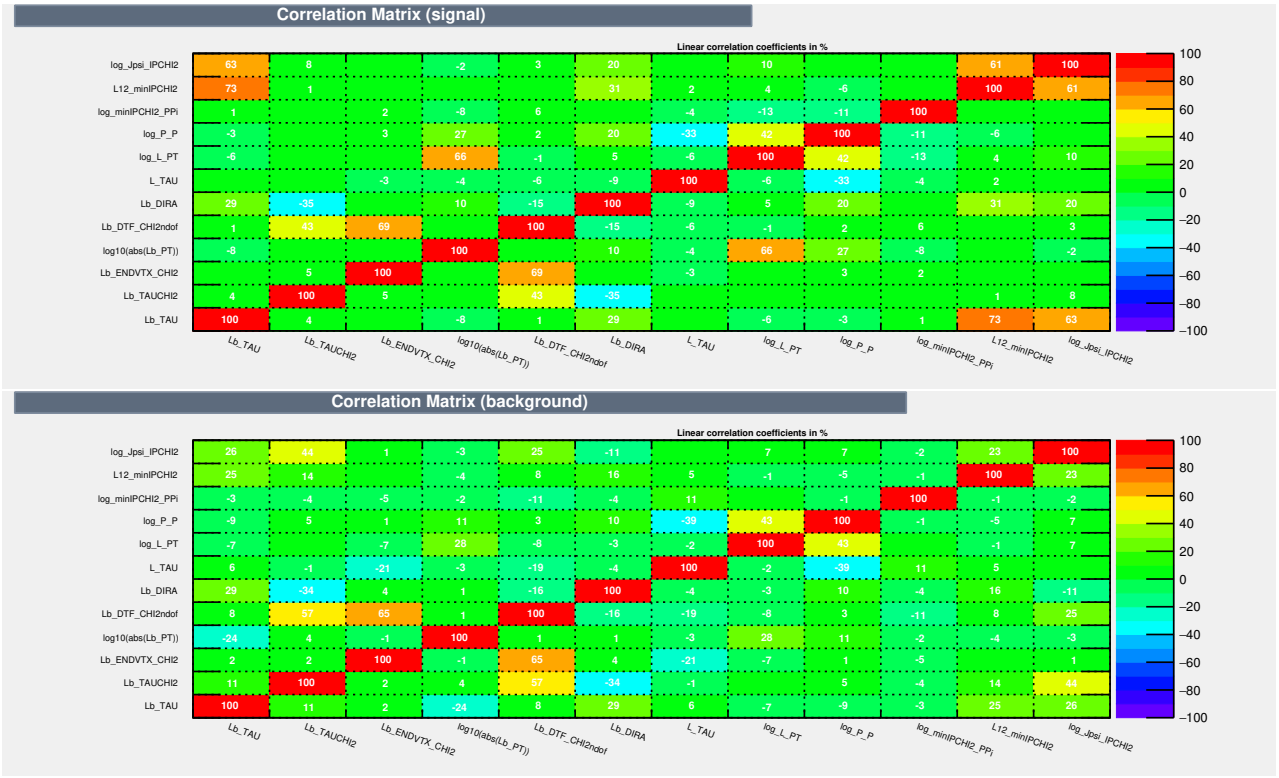


Figure G.2: Correlation matrices (signal and background)

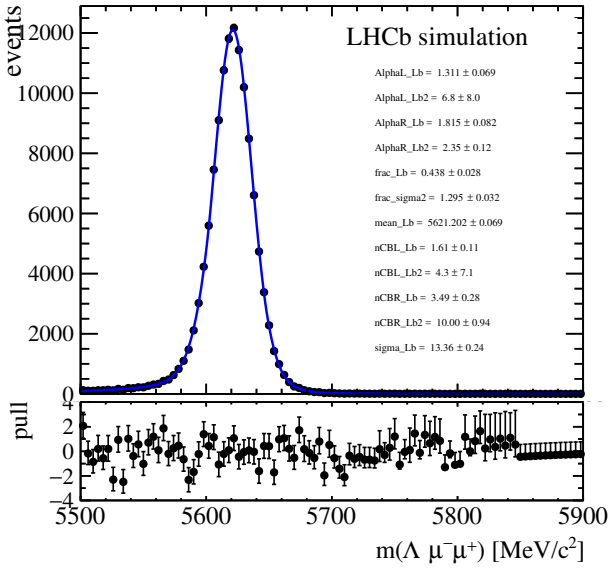


Figure G.3: Fit to the Run 2 $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ simulation samples to determine the tail parameters of the signal distribution.

Appendix H

Electrons vs muons: important differences

Two of the most important differences between muons and electrons arise from the about 200 times smaller mass of the latter and the fact that they are one of the key components in ordinary matter. The stopping power of a particle describes the lost energy per travelled path through matter and is therefore a good measure to study the difference in the material interaction between electrons and muons. Fig. H.1 shows the different stopping powers of muons and electrons in copper. The blue square in the middle marks the typical energy range of leptons in the LHCb experiment. It can be seen, that for muons the energy loss (red) is very small while electrons loose a bit of energy due to collision loss (teal) and significantly due to bremsstrahlung (light blue). Consequently, when electrons travel through a detector they loose a big fraction of their energy due to the emission of photons.
Due to the electric charge of the electron it interacts with the electrical charged particle of the detector material, causing the electron to slow done and loosing it energy by emitting bremsstrahlung colinearly to its flight direction. This effect not only complicates the reconstruction of electrons, it also leads to a lower momentum resolution. In order to understand, why this is the case it is important to understand, where bremsstrahlung is emitted and how it gets reconstructed to improve the resolution of electrons.
Bremsstrahlung is emitted every time an electron passes material in the LHCb detector. This effect can even be used to visualise the LHCb detector, when plotting the number of bremsstrahlung photons as function of their origin z-coordinate. In Fig. H.2 it can be seen that below $z < 1.5\,\mathrm{m}$ most of the photons are emitted, this is the position of the VELO which has one of the biggest material budgets in the whole experiment. Between $2 < z < 3\,\mathrm{m}$, the two TT stations are visible and lastly between $7 < z < 10\,\mathrm{m}$ the three tracking stations of the downstream tracker. The region between the VELO and TT has some photons emitted due to interactions with the RICH. The very small amount of photons in the region of the magnet arise from very few interactions with the beampipe, it transport structure and the air.
The place of original of a bremsstrahlung photon largely impacts how it is reconstructed. When the photon is only emitted in material after the magnet, the flight direction of the electron is not altered anymore. As a consequence the electron and bremsstrahlung photon end up in the same calorimeter cells and no dedicated reconstruction is needed.
Contrary, if the photon is emitted before the magnet it flies in the original flight direction

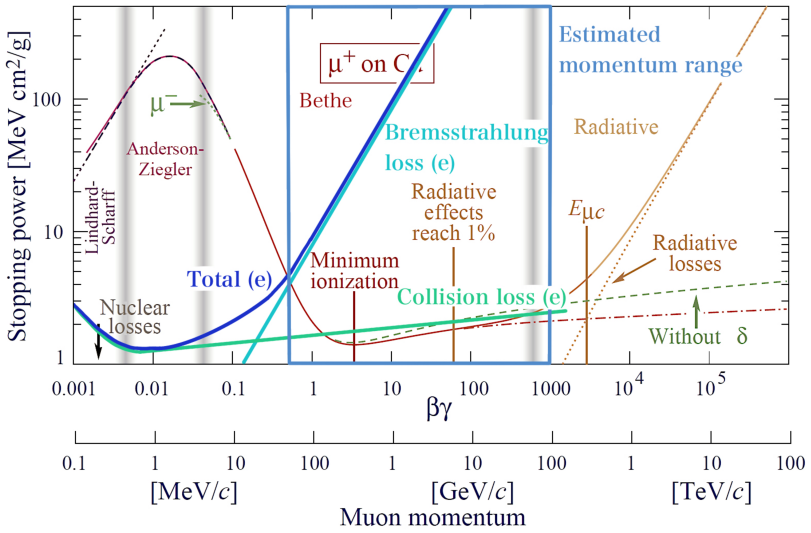


Figure H.1: Stopping power ($\langle -dE/dx \rangle$) of an electron compared to the one of a positive charge muon in copper as a function of the momentum. The solid red to orange line is the total stopping power of the muon, while the solid dark blue line is the one for the electron. For electrons the two main components are collision losses (solid line in turquoise) and the bremsstrahlung loss (solid line in light blue). The squared window shows the estimated momentum range for analysis of $b \to s\ell^+\ell^-$ transitions. The graphic is a personal adaption from stopping power for muons [197].

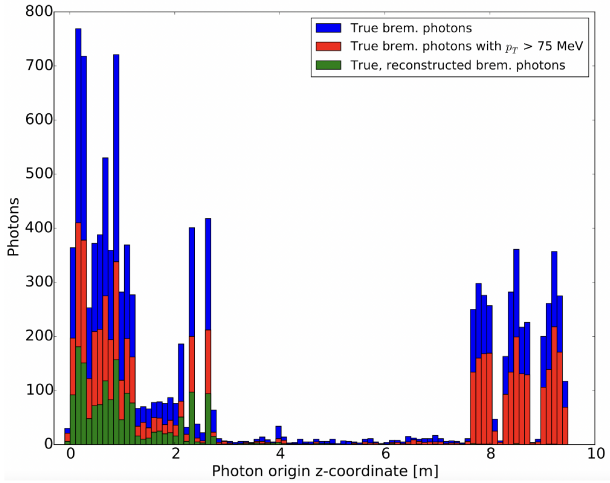


Figure H.2: Bremsstrahlung photons origin in the LHCb detector as function of the z-coordinate. All bremsstrahlung photons are shown in blue; potentially reconstructable photons (those having $p_{\mathrm{T}} > 75\,\mathrm{MeV}/c$) are shown in red; and photons which are actually reconstructed properly are shown in green. Taken from Ref. [240].

of the electron, while the electrons flight direction is altered by the magnet. As a result the photon and electron do not end up in the same calorimeter cell and the momentum of the electron is underestimated due to the fraction given to the photon.
The impact of bremsstrahlung can be visualised when comparing the reconstruction efficiency of muons and electrons as a function of their transverse momentum. Fig. H.3 shows the VELO and long track reconstruction efficiency. While in the VELO both muons and electrons are nearly untouched by the effects from the magnetic field, they travel through the detector in the same way leading to approximately the same reconstruction efficiency and close to 100% (left plot). However, when looking at long tracks electrons loose a fraction of their momentum to the photons leading to a lower reconstruction efficiency (right plot).

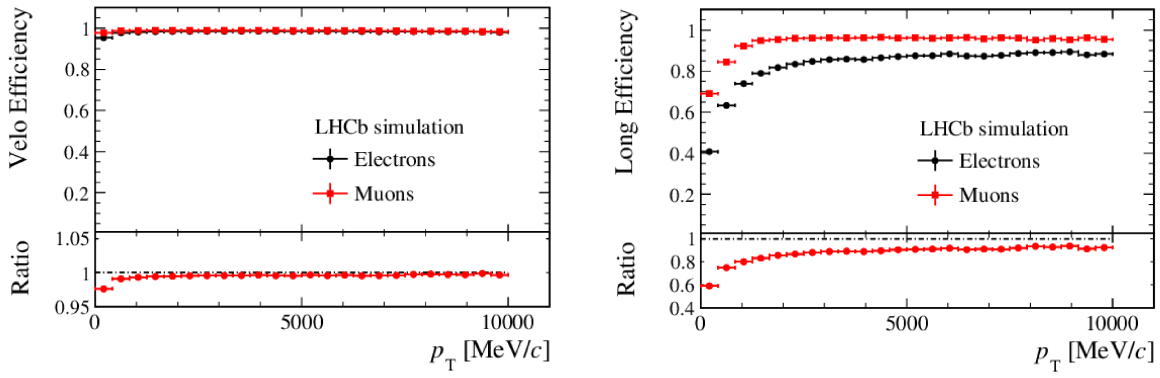


Figure H.3: Reconstruction efficiency for VELO (left) and long tracks (right) for muons and electrons (as a function of the lepton's transverse momentum. The used sample are obtained from simulation of $B^+ \to J/\psi K^+$ decays. Taken from Ref. [241].

A dedicated reconstruction algorithm, called `BremAdder`, helps to associate a bremsstrahlung photon with its source electron, enabling to improve the invariant mass resolution in decays involving electrons. The left part of Fig. H.4 shows how the signature of an electron event with a bremsstrahlung photon emitted before the magnet looks like. The `BremAdder` algorithm defines a search window for a photon cluster in the ECAL by extrapolating the electron track before the magnet to the ECAL. This procedure is illustrated in right side of Fig. H.4. The slope of the electron track can then be used to reconstruct the four-momentum of the photon, which is afterwards used to corrected the four-momentum of the electron track.
A second important algorithm is the `DiElectronMaker`. It is run as part of the stripping if a channel has has an electron pair to ensure a bremsstrahlung photon is only added to one electron per event. However, the invariant mass of a decay involving electron is always at least diluted by the limited ECAL resolution and the cut on the minimal transverse momentum of the photon ($p_T > 75\,\text{MeV}/c$) as photons below this threshold are not reconstructed [240]. In addition, the performance of the algorithms is not perfect and largely impacted by the high occupancy in the electromagnetic calorimeter. Two extreme configurations can occur and can be seen visualised, when comparing the resolution of the dilepton invariant mass or the invariant mass of the studied decay. When a random photon is picked up ad bremsstrahlung photon, the electron momentum is overestimated

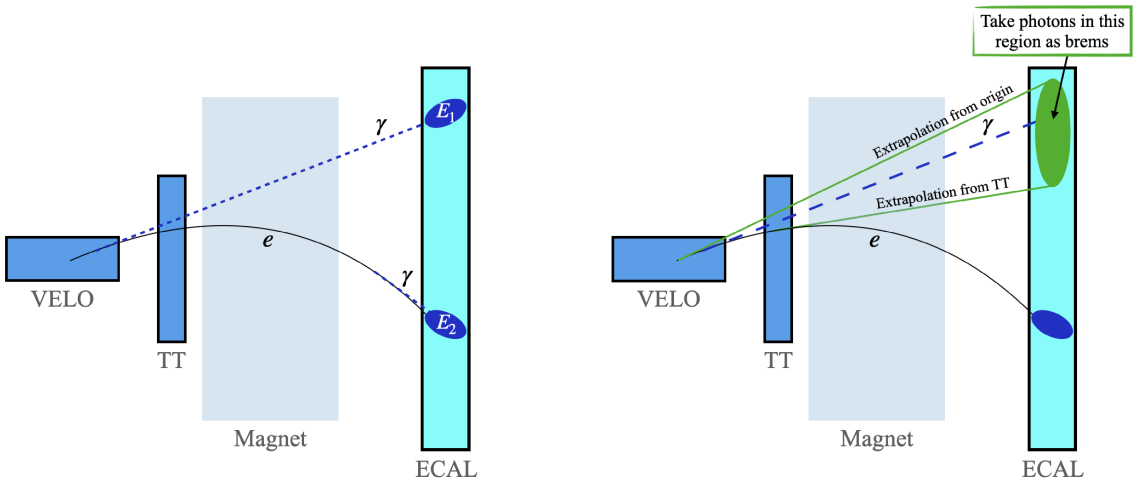


Figure H.4: Schematic representation of the bremsstrahlung photons emitted by an electron before and after the magnet (left). Schematic illustration of the `BremAdder` algorithm search window (right). Personal adaptation from Ref. [240].

leading to a higher invariant mass and the tail on the right side from the signal peaks. But sometime the algorithms does not find the bremsstrahlung photon, leading to an underestimation of the electron momentum and consequently in the dielectron mass $m(e^+e^-)$ as well as the depending invariant mass of the studied decay mode.
What is important for the studies with electrons in the decay is that the bremsstrahlung recovery might not be perfect, but it is a powerful discrimination variable for the electron identification. This is a result of the fact that only electron tracks have bremsstrahlung photons associated to them, improving the reliability of the particle identification performance.

Bibliography

[1] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak interactions with lepton-hadron symmetry*, Physical Review D **2** (1970) 1285.

[2] Belle collaboration, J.-T. Wei *et al.*, *Measurement of the Differential Branching Fraction and Forward-Backward Asymmetry for $B \to K^{(*)}\ell^+\ell^-$*, Phys. Rev. Lett. **103** (2009), no. BELLE-PREPRINT-2009-7, KEK-PREPRINT-2008-56 171801, arXiv:0904.0770.

[3] CDF collaboration, T. Aaltonen *et al.*, *Observation of the Baryonic Flavor-Changing Neutral Current Decay $\Lambda_b^0 \to \Lambda\mu^+\mu^-$*, Phys. Rev. Lett. **107** (2011), no. FERMILAB-PUB-11-328-E 201802, arXiv:1107.3753.

[4] BaBar collaboration, J. P. Lees *et al.*, *Measurement of Branching Fractions and Rate Asymmetries in the Rare Decays $B \to K^{(*)}l^+l^-$*, Phys. Rev. D **86** (2012), no. SLAC-PUB-14957, BABAR-PUB-12-002 032012, arXiv:1204.3933.

[5] CMS collaboration, S. Chatrchyan *et al.*, *Angular Analysis and Branching Fraction Measurement of the Decay $B^0 \to K^{*0}\mu^+\mu^-$*, Phys. Lett. B **727** (2013), no. CMS-BPH-11-009, CERN-PH-EP-2013-152 77, arXiv:1308.3409.

[6] LHCb collaboration, R. Aaij *et al.*, *Differential branching fractions and isospin asymmetries of $B \to K^{(*)}\mu^+\mu^-$ decays*, JHEP **06** (2014) 133, arXiv:1403.8044.

[7] LHCb collaboration, R. Aaij *et al.*, *Differential branching fraction and angular analysis of $\Lambda_b^0 \to \Lambda\mu^+\mu^-$ decays*, JHEP **06** (2015) 115, Erratum ibid. **09** (2018) 145, arXiv:1503.07138.

[8] CMS collaboration, V. Khachatryan *et al.*, *Angular analysis of the decay $B^0 \to K^{*0}\mu^+\mu^-$ from pp collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. B **753** (2016), no. CMS-BPH-13-010, CERN-PH-EP-2015-178 424, arXiv:1507.08126.

[9] LHCb collaboration, R. Aaij *et al.*, *Measurements of the S-wave fraction in $B^0 \to K^+\pi^-\mu^+\mu^-$ decays and the $B^0 \to K^*(892)^0\mu^+\mu^-$ differential branching fraction*, JHEP **11** (2016) 047, Erratum ibid. **04** (2017) 142, arXiv:1606.04731.

[10] Belle collaboration, S. Wehle *et al.*, *Lepton-Flavor-Dependent Angular Analysis of $B \to K^*\ell^+\ell^-$*, Phys. Rev. Lett. **118** (2017), no. BELLE-PREPRINT-2016-15, KEK-PREPRINT-2016-54 111801, arXiv:1612.05014.

[11] CMS collaboration, A. M. Sirunyan *et al.*, *Measurement of angular parameters from the decay* $\mathrm{B}^0 \to \mathrm{K}^{*0}\mu^+\mu^-$ *in proton-proton collisions at* $\sqrt{s} = 8$ *TeV*, Phys. Lett. B **781** (2018), no. CMS-BPH-15-008, CERN-EP-2017-240 517, arXiv:1710.02846.

[12] ATLAS collaboration, M. Aaboud *et al.*, *Angular analysis of* $B_d^0 \to K^*\mu^+\mu^-$ *decays in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, JHEP **10** (2018), no. CERN-EP-2017-161 047, arXiv:1805.04000.

[13] Belle collaboration, S. Choudhury *et al.*, *Test of lepton flavor universality and search for lepton flavor violation in* $B \to K\ell\ell$ *decays*, JHEP **03** (2021), no. BELLE-CONF-1904, Belle Preprint 2020-11, KEK Preprint 2020-12 105, arXiv:1908.01848.

[14] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP-averaged observables in the* $B^0 \to K^{*0}\mu^+\mu^-$ *decay*, Phys. Rev. Lett. **125** (2020) 011802, arXiv:2003.04831.

[15] 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.

[16] LHCb collaboration, R. Aaij *et al.*, *Branching fraction measurements of the rare* $B_s^0 \to \phi\mu^+\mu^-$ *and* $B_s^0 \to f_2'(1525)\mu^+\mu^-$ *decays*, Phys. Rev. Lett. **127** (2021) 151801, arXiv:2105.14007.

[17] LHCb collaboration, R. Aaij *et al.*, *Angular analysis of the rare decay* $B_s^0 \to \phi\mu^+\mu^-$, JHEP **11** (2021) 043, arXiv:2107.13428.

[18] CMS collaboration, *Observation of the* $\Xi_\mathrm{b}^- \to \psi(2S)\Xi^-$ *decay and studies of the* Ξ_b^{*0} *baryon in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, 2024.

[19] LHCb collaboration, R. Aaij *et al.*, *Measurement of the mass difference and relative production rate of the* Ω_b^- *and* Ξ_b^- *baryons*, Phys. Rev. **D108** (2023) 052008, arXiv:2305.15329.

[20] LHCb collaboration, R. Aaij *et al.*, *Observation of a new* Ξ_b^0 *state*, Phys. Rev. **D103** (2021) 012004, arXiv:2010.14485.

[21] D0 collaboration, V. M. Abazov *et al.*, *Observation of the doubly strange b baryon* Ω_b^-, Phys. Rev. Lett. **101** (2008), no. FERMILAB-PUB-08-335-E 232002, arXiv:0808.4142.

[22] H.-Y. Cheng, *Nonleptonic weak decays of bottom baryons*, Phys. Rev. D **56** (1997), no. IP-ASTP-06-96 2799, arXiv:hep-ph/9612223, [Erratum: Phys.Rev.D 99, 079901 (2019)].

[23] ATLAS collaboration, G. Aad *et al.*, *Measurement of the branching ratio* $\Sigma(\Lambda_b^0 \to \psi(2S)\Lambda)/\Sigma(\Lambda_b^0 \to J/\psi\Lambda)$ *with the ATLAS detector*, Physics Letters B **751** (2015) 63–80.

[24] LHCb collaboration, R. Aaij *et al.*, *Measurement of the ratio of branching fractions of the decays* $\Lambda_b^0 \to \psi(2S)\Lambda$ *and* $\Lambda_b^0 \to J/\psi\Lambda$, JHEP **03** (2019) 126, arXiv:1902.02092.

[25] M. Kenzie *et al.*, *GammaCombo framework for combinations of measurements and computation of confidence intervals: Public Release v1.3*, Feb., 2020. doi: 10.5281/zenodo.3660991.

[26] LHCb collaboration, R. Aaij *et al.*, *Measurement of the CKM angle γ from a combination of LHCb results*, JHEP **12** (2016) 087, arXiv:1611.03076.

[27] P. Nayek, S. Biswas, P. Maji, and S. Sahoo, *Imprints of New Physics in Baryonic Decay $\Xi_b \to \Xi l^+ l^-$ in Non-universal Z' Model*, Int. J. Theor. Phys. **59** (2020), no. 9 2712.

[28] C. Rovelli, *The First Scientist: Anaximander and His Legacy*, Westholme Publishing, 2011.

[29] J. J. Thomson, *Cathode rays*, Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character **180** (1893) 649.

[30] D. Walsh, R. F. Carswell, and R. J. Weymann, *0957+561 A, B: Twin Quasistellar Objects or Gravitational Lens?*, Nature **279** (1979) 381.

[31] V. C. Rubin and W. K. Ford Jr. *Rotation of the Andromeda Nebula from a spectroscopic survey of emission regions*, The Astrophysical Journal **238** (1980) 471.

[32] UA1, G. Arnison *et al.*, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*, Phys. Lett. **B122** (1983), no. CERN-EP-83-13 103.

[33] Particle Data Group, R. L. Workman *et al.*, *Review of particle physics*, Prog. Theor. Exp. Phys. **2022** (2022), no. 8 083C01.

[34] K. Nishijima, *Charge Independence Theory of V Particles**, Progress of Theoretical Physics **13** (1955) 285.

[35] M. Gell-Mann, *The interpretation of the new particles as displaced charge multiplets*, Nuovo Cim. **4** (1956), no. S2 848.

[36] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[37] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) .

[38] S. L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961), no. 4 579.

[39] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[40] A. Salam, *Weak and electromagnetic interactions*, Conf. Proc. C **680519** (1968) 367.

[41] N. Cabibbo, *Unitary symmetry and leptonic decays*, Phys. Rev. Lett. **10** (1963) 531.

[42] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progress of Theoretical Physics **49** (1973) 652.

[43] SNO collaboration, S. N. Ahmed *et al.*, *Measurement of the total active B-8 solar neutrino flux at the Sudbury Neutrino Observatory with enhanced neutral current sensitivity*, Phys. Rev. Lett. **92** (2004) 181301, arXiv:nucl-ex/0309004.

[44] B. Pontecorvo, *Inverse beta processes and nonconservation of lepton charge*, Zh. Eksp. Teor. Fiz. **34** (1957) 247.

[45] Z. Maki, M. Nakagawa, and S. Sakata, *Remarks on the unified model of elementary particles*, Prog. Theor. Phys. **28** (1962) 870.

[46] L. Calefice, *Standalone track reconstruction on GPUs in the first stage of the upgraded LHCb trigger system; Preparations for measurements with strange hadrons in Run 3*, PhD thesis, TU Dortmund University, 2022, CERN-THESIS-2022-343, Presented 13 Dec 2022.

[47] D. Griffiths, *Introduction to Elementary Particles*, 2nd revised Edition, Wiley-VCH, 2017.

[48] ATLAS collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012), no. CERN-PH-EP-2012-218 1, arXiv:1207.7214.

[49] CMS collaboration, S. Chatrchyan *et al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012), no. CMS-HIG-12-028, CERN-PH-EP-2012-220 30, arXiv:1207.7235.

[50] G. Gamow, *The evolution of the universe*, Nature **162** (1948) 680.

[51] A. Penzias and R. W. Wilson, *A measurement of excess antenna temperature at 4080 mc/s*, The Astrophysical Journal **142** (1965) 419.

[52] A. Penzias and R. W. Wilson, *Microwave noise from rainstorms*, The Astrophysical Journal **142** (1965) 419.

[53] L. Canetti, M. Drewes, and M. Shaposhnikov, *Matter and antimatter in the universe*, New Journal of Physics **14** (2012) 095012.

[54] E. Corbelli and P. Salucci, *The extended rotation curve and the dark matter halo of M33*, Monthly Notices of the Royal Astronomical Society **311** (2000) 441.

[55] L. D. Duffy and K. v. Bibber, *Axions as dark matter particles*, New Journal of Physics **11** (2009) 105008.

[56] J. R. Ellis, *Supersymmetric dark matter*, AIP Conf. Proc. **562** (2000) 9.

[57] S.-K. Choi *et al.*, *Observation of a Narrow Charmoniumlike State in Exclusive $B^{\pm} \to K^{\pm}\pi^{+}\pi^{-}J/\psi$ Decays*, Physical Review Letters **91** (2003) .

[58] LHCb collaboration, R. Aaij *et al.*, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi p K^-$ decays*, Phys. Rev. Lett. **115** (2015) 072001, arXiv:1507.03414.

[59] LHCb collaboration, R. Aaij *et al.*, *Observation of exotic $J/\psi\phi$ structures from amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays*, Phys. Rev. Lett. **118** (2017) 022003, arXiv:1606.07895.

[60] Y.-R. Liu *et al.*, *Pentaquark and tetraquark states*, Progress in Particle and Nuclear Physics **107** (2019) 237.

[61] R. L. Anderson and et al. *Experimental study of the inelasticity of electron scattering by nucleons*, Physical Review Letters **23** (1969), no. 12 662.

[62] M. Gell-Mann, *A schematic model of baryons and mesons*, Physics Letters **8** (1964), no. 3 214.

[63] V. Lisovskyi, *Study of rare b-baryon decays and test of lepton universality at LHCb*, PhD thesis, Université Paris-Saclay, 2019, CERN-THESIS-2019-202, Presented 09 Sep 2019.

[64] C. Albajar *et al.*, *First observation of the beauty baryon Λ_b^0 in the decay channel $\Lambda_b^0 \to J/\psi\Lambda$ at the CERN proton-antiproton collider*, Physics Letters B **273** (1991), no. 4 540.

[65] CDF collaboration, T. Aaltonen *et al.*, *Observation and mass measurement of the baryon Ξ_b^-*, Phys. Rev. Lett. **99** (2007), no. FERMILAB-PUB-07-336-E 052002, arXiv:0707.0589.

[66] LHCb collaboration, R. Aaij *et al.*, *Measurement of the mass and production rate of Ξ_b^- baryons*, Phys. Rev. **D99** (2019) 052006, arXiv:1901.07075.

[67] LHCb collaboration, R. Aaij *et al.*, *Measurements of the Λ_b^0, Ξ_b^-, and Ω_b^- baryon masses*, Phys. Rev. Lett. **110** (2013) 182001, arXiv:1302.1072.

[68] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the mass and lifetime of the Ξ_b^- baryon*, Phys. Rev. Lett. **113** (2014) 242002, arXiv:1409.8568.

[69] LHCb collaboration, R. Aaij *et al.*, *Observation of the decay $\Xi_b^- \to pK^-K^-$*, Phys. Rev. Lett. **118** (2017) 071801, arXiv:1612.02244.

[70] LHCb collaboration, R. Aaij *et al.*, *Search for CP violation in $\Xi_b^- \to pK^-K^-$decays*, Phys. Rev. **D104** (2021) 052010, arXiv:2104.15074.

[71] LHCb collaboration, R. Aaij *et al.*, *Search for the radiative $\Xi_b^- \to \Xi^-\gamma$ decay*, JHEP **01** (2022) 069, arXiv:2108.07678.

[72] CDF collaboration, T. Aaltonen *et al.*, *Observation of the Ξ_b^0 baryon*, Phys. Rev. Lett. **107** (2011) 102001.

[73] LHCb collaboration, R. Aaij *et al.*, *Precision measurement of the mass and lifetime of the Ξ_b^0 baryon*, Phys. Rev. Lett. **113** (2014) 032001, arXiv:1405.7223.

[74] LHCb collaboration, R. Aaij *et al.*, *Study of beauty baryon decays to D^0ph^- and $\Lambda_c^+h^-$ final states*, Phys. Rev. **D89** (2014) 032001, arXiv:1311.4823.

[75] LHCb collaboration, R. Aaij *et al.*, *Measurement of branching fractions of charmless four-body Λ_b^0 and Ξ_b^0 decays*, JHEP **02** (2018) 098, arXiv:1711.05490.

[76] LHCb collaboration, R. Aaij *et al.*, *Search for CP violation using triple product asymmetries in $\Lambda_b^0 \to pK^-\pi^+\pi^-$, $\Lambda_b^0 \to pK^-K^+K^-$, and $\Xi_b^0 \to pK^-K^-\pi^+$ decays*, JHEP **08** (2018) 039, arXiv:1805.03941.

[77] LHCb collaboration, R. Aaij *et al.*, *Isospin amplitudes in $\Lambda_b^0 \to J/\psi\Lambda(\Sigma^0)$ and $\Xi_b^0 \to J/\psi\Xi^0(\Lambda)$ decays*, Phys. Rev. Lett. **124** (2020) 111802, arXiv:1912.02110.

[78] CDF collaboration, T. Aaltonen *et al.*, *Observation of the Ω_b^- and Measurement of the Properties of the Ξ_b^- and Ω_b^-*, Phys. Rev. D **80** (2009), no. FERMILAB-PUB-09-256-E 072003, arXiv:0905.3123.

[79] CDF collaboration, T. A. Aaltonen *et al.*, *Mass and lifetime measurements of bottom and charm baryons in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. D **89** (2014), no. FERMILAB-PUB-14-061-E 072014, arXiv:1403.8126.

[80] LHCb collaboration, R. Aaij *et al.*, *Measurements of the mass and lifetime of the Ω_b^- baryon*, Phys. Rev. **D93** (2016) 092007, arXiv:1604.01412.

[81] LHCb collaboration, R. Aaij *et al.*, *Observation of excited Ω_c^0 baryons in $\Omega_b^- \to {\Xi_c}^+K^-\pi^+$ decays*, Phys. Rev. **D104** (2021) L091102, arXiv:2107.03419.

[82] LHCb collaboration, R. Aaij *et al.*, *Study of the kinematic dependences of Λ_b^0 production in pp collisions and a measurement of the $\Lambda_b^0 \to \Lambda_c^+\pi^-$ branching fraction*, JHEP **08** (2014) 143, arXiv:1405.6842.

[83] LHCb collaboration, R. Aaij *et al.*, *Measurement of b-hadron fractions in 13 TeV pp collisions*, Phys. Rev. **D100** (2019) 031102(R), arXiv:1902.06794.

[84] LHCb collaboration, R. Aaij *et al.*, *Measurement of the B_c^- production fraction and asymmetry in 7 and 13 TeV pp collisions*, Phys. Rev. **D100** (2019) 112006, arXiv:1910.13404.

[85] M. Karliner, B. Keren-Zur, H. J. Lipkin, and J. L. Rosner, *The quark model and b baryons*, Annals of Physics **324** (2009), no. 1 2.

[86] LHCb collaboration, R. Aaij *et al.*, *Measurement of the properties of the Ξ_b^{*0} baryon*, JHEP **05** (2016) 161, arXiv:1604.03896.

[87] LHCb collaboration, R. Aaij *et al.*, *First branching fraction measurement of the suppressed decay $\Xi_c^0 \to \pi^-\Lambda_c^+$*, Phys. Rev. **D102** (2020) 071101(R), arXiv:2007.12096.

[88] S. Stone and M. Wilkinson, *Observation of the suppressed decay $\Xi_c^0 \to \pi^-\Lambda_c^+$*, LHCb-ANA-2020-016.

[89] LHCb collaboration, R. Aaij *et al.*, *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*, Phys. Rev. Lett. **118** (2017) 052002, Erratum ibid. **119** (2017) 169901, arXiv:1612.05140.

[90] LHCb collaboration, R. Aaij *et al.*, *Measurement of Υ production cross-section in pp collisions at $\sqrt{s}$ =13 TeV*, JHEP **07** (2018) 134, arXiv:1804.09214.

[91] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $B^{\pm}$ production cross-section in pp collisions at $\sqrt{s}$ =7 and 13 TeV*, JHEP **12** (2017) 026, arXiv:1710.04921.

[92] P. Colangelo and A. Khodjamirian, in *QCD sum rules, a modern perspective*, p. 1495–1576, World Scientific, 04, 2001. doi: 10.1142/9789812810458_0033.

[93] N. Mathur, R. Lewis, and R. M. Woloshyn, *Charmed and bottom baryons from lattice nonrelativistic qcd*, Physical Review D **66** (2002) .

[94] D. Bečirević *et al.*, *Scalar leptoquarks from grand unified theories to accommodate the b-physics anomalies*, Physical Review D **98** (2018) .

[95] J. Fuentes-Martín, G. Isidori, J. Pagès, and K. Yamamoto, *With or without U(2)? Probing non-standard flavor and helicity structures in semileptonic B decays*, Physics Letters B **800** (2020) 135080.

[96] A. J. Buras, *Weak Hamiltonian, CP violation and rare decays*, in *Les Houches Summer School in Theoretical Physics, Session 68: Probing the Standard Model of Particle Interactions*, no. TUM-HEP-316-98, pp. 281–539, 6, 1998. arXiv:hep-ph/9806471.

[97] A. Pich, *Effective field theory: Course*, in *Les Houches Summer School in Theoretical Physics, Session 68: Probing the Standard Model of Particle Interactions*, no. FTUV-98-46, IFIC-98-47, pp. 949–1049, 6, 1998. arXiv:hep-ph/9806303.

[98] N. Isgur and M. B. Wise, *Weak transition form factors between heavy mesons*, Physics Letters B **237** (1990), no. 3 527.

[99] N. Isgur and M. B. Wise, *Weak decays of heavy mesons in the static quark approximation*, Physics Letters B **232** (1989), no. 1 113.

[100] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $\Lambda_b^0 \to \Lambda(1520)\mu^+\mu^-$ differential branching fraction*, Phys. Rev. Lett. **131** (2023) 151801, arXiv:2302.08262.

[101] S. Descotes-Genon and M. N. Brunet, *Angular analysis of the rare decay $\Lambda_b \to \Lambda(1520)(\to N\overline{K})\ell^+\ell^-$*, Journal of High Energy Physics **2019** (2019) .

[102] W. Detmold and S. Meinel, *$\Lambda_b^0 \to \Lambda\ell^+\ell^-$ form factors, differential branching fraction, and angular observables from lattice QCD with relativistic b quarks*, Phys. Rev. D **93** (2016), no. 7 074501, arXiv:1602.01399.

[103] T. Blake, S. Meinel, and D. van Dyk, *Bayesian Analysis of $b \to s\mu^+\mu^-$ Wilson Coefficients using the Full Angular Distribution of $\Lambda_b^0 \to \Lambda(\to p\,\pi^-)\mu^+\mu^-$ Decays*, Phys. Rev. D **101** (2020), no. 3 035023, arXiv:1912.05811.

[104] Belle collaboration, A. Abdesselam *et al.*, *Test of Lepton-Flavor Universality in $B \to K^* \ell^+ \ell^-$ Decays at Belle*, Phys. Rev. Lett. **126** (2021), no. BELLE-CONF-1901, Belle Preprint 2020-14, KEK Preprint 2020-16 161801, arXiv:1904.02440.

[105] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality in beauty-quark decays*, Nature Physics **18** (2022) 277, arXiv:2103.11769.

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

[107] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality in $b \to s \ell^+ \ell^-$ decays*, Phys. Rev. Lett. **131** (2023) 051803, arXiv:2212.09152.

[108] 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.

[109] LHCb collaboration, R. Aaij *et al.*, *Test of lepton universality using $\Lambda_b^0 \to p K^- \ell^+ \ell^-$ decays*, JHEP **05** (2020) 040, arXiv:1912.08139.

[110] 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.

[111] 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.

[112] LHCb collaboration, R. Aaij *et al.*, *Angular moments of the decay $\Lambda_b^0 \to \Lambda \mu^+ \mu^-$ at low hadronic recoil*, JHEP **09** (2018) 146, arXiv:1808.00264.

[113] BaBar collaboration, D. Boutigny *et al.*, *BaBar technical design report*, Tech. Rep. SLAC-R-0457, SLAC-R-457, SLAC-0457, SLAC-457, SLAC-R-95-457, 3, 1995.

[114] A. Abashian *et al.*, *The belle detector*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **479** (2002), no. 1 117, Detectors for Asymmetric B-factories.

[115] T. Abe *et al.*, *Belle II Technical Design Report*, Tech. Rep. KEK REPORT 2010-1, 2010. Comments: Edited by: Z. Doležal and S. Uno.

[116] DELPHI collaboration, P. Abreu *et al.*, *Measurement of Λ_b Production and Lifetime in Z^0 Hadronic Decays*, Phys. Lett. B **311** (1993), no. CERN-PPE-93-032 379.

[117] ALEPH collaboration, D. Buskulic *et al.*, *Measurement of the mass of the Λ_b baryon*, Phys. Lett. B **380** (1996), no. CERN-PPE-96-028, CERN-PPE-96-28 442.

[118] DELPHI collaboration, P. Abreu *et al.*, *Λ_b polarization in Z^0 decays at LEP*, Phys. Lett. B **474** (2000), no. CERN-EP-99-155 205.

[119] LHCb collaboration, *Framework TDR for the LHCb Upgrade: Technical Design Report*, CERN-LHCC-2012-007.

[120] J. P. G. L. Tock *et al.*, *The Second LHC Long Shutdown (LS2) for the Superconducting Magnets*, in *Proc. 9th International Particle Accelerator Conference (IPAC'18), Vancouver, BC, Canada, April 29-May 4, 2018*, no. 9 in International Particle Accelerator Conference, (Geneva, Switzerland), pp. 240–243, JACoW Publishing, June, 2018. https://doi.org/10.18429/JACoW-IPAC2018-MOPMF056, doi: doi:10.18429/JACoW-IPAC2018-MOPMF056.

[121] J. Haffner, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, , General Photo.

[122] E. Boltezar *et al.*, *THE NEW CERN 50-MEV LINAC.*, in *10th International Linear Accelerator Conference*, pp. S2–2, 1980.

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

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

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

[126] W. J. Stirling, private communication, *Plot: proton-(anti)proton cross sections*, https://www.hep.ph.ic.ac.uk/~wstirlin/plots/crosssections2012_v5.pdf.

[127] LHCb collaboration, S. Amato *et al.*, *LHCb technical proposal: A Large Hadron Collider Beauty Experiment for Precision Measurements of CP Violation and Rare Decays*, .

[128] LHCb collaboration, C. Elsässer, *$b\bar{b}$ production angle plots*, https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html.

[129] LHCb collaboration, R. Aaij *et al.*, *LHCb detector performance*, Int. J. Mod. Phys. **A30** (2015) 1530022, arXiv:1412.6352.

[130] LHCb collaboration, *LHCb Performance plots*, https://lbgroups.cern.ch/online/OperationsPlots/index.htm.

[131] C. Abellan Beteta *et al.*, *Monitoring radiation damage in the LHCb Tracker Turicensis*, arXiv:1809.05063.

[132] LHCb collaboration, A. A. Alves Jr. *et al.*, *The LHCb detector at the LHC*, JINST **3** (2008), no. LHCb-DP-2008-001 S08005.

[133] LHCb collaboration, *LHCb VELO (VErtex LOcator): Technical Design Report*, CERN-LHCC-2001-011.

[134] R. Aaij *et al.*, *Performance of the LHCb Vertex Locator*, JINST **9** (2014) P09007, arXiv:1405.7808.

[135] LHCb collaboration, P. Collins, *Approved VELO conference plots*, https://lbtwiki.cern.ch/bin/view/VELO/VELOConferencePlots.

[136] LHCb collaboration, *LHCb reoptimized detector design and performance: Technical Design Report*, CERN-LHCC-2003-030.

[137] LHCb collaboration, *LHCb inner tracker: Technical Design Report*, CERN-LHCC-2002-029.

[138] LHCb collaboration, *LHCb outer tracker: Technical Design Report*, CERN-LHCC-2001-024.

[139] R. Arink *et al.*, *Performance of the LHCb Outer Tracker*, JINST **9** (2014) P01002, arXiv:1311.3893.

[140] P. d'Argent *et al.*, *Improved performance of the LHCb Outer Tracker in LHC Run 2*, JINST **12** (2017) P11016, arXiv:1708.00819.

[141] R. Calabrese *et al.*, *Performance of the LHCb RICH detectors during LHC Run 2*, JINST **17** (2022), no. LHCb-DP-2021-004 P07013, arXiv:2205.13400.

[142] LHCb collaboration, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036.

[143] 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.

[144] P. Perret and X. Vilasis-Cardona, *Performance of the LHCb calorimeters during the period 2010-2012*, J. Phys. : Conf. Ser. **587** (2015), no. LHCb-PROC-2014-019, CERN-LHCb-PROC-2014-019 012012, see LHCb-TALK-2014-236.

[145] V. V. Gligorov, *Conceptualization, implementation, and commissioning of real-time analysis in the High Level Trigger of the LHCb experiment*, arXiv:1806.10912.

[146] M. Borsato, *Study of the $B^0 \to K^{*0}e^+e^-$ decay with the LHCb detector and development of a novel concept of PID detector: the Focusing DIRC*, PhD thesis, Paris U., IV, 2015, CERN-THESIS-2015-219, Presented 08 Sep 2015.

[147] LHCb collaboration, *LHCb muon system: Technical Design Report*, CERN-LHCC-2001-010.

[148] LHCb collaboration, *LHCb: Second addendum to the muon system technical design report*, Tech. Rep. CERN-LHCC-2005-012, 4, 2005.

[149] A. A. Alves Jr. *et al.*, *Performance of the LHCb muon system*, JINST **8** (2013) P02022, arXiv:1211.1346.

[150] F. Archilli *et al.*, *Performance of the muon identification at LHCb*, JINST **8** (2013) P10020, arXiv:1306.0249.

[151] W. D. Hulsbergen, *The global covariance matrix of tracks fitted with a Kalman filter and an application in detector alignment*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **600** (2009) 471–477.

[152] O. Callot, *FastVelo, a fast and efficient pattern recognition package for the Velo*, Tech. Rep. LHCb-PUB-2011-001, CERN-LHCb-PUB-2011-001, CERN, Geneva, 2011. LHCb collaboration.

[153] O. Callot and M. Schiller, *PatSeeding: A Standalone Track Reconstruction Algorithm*, Tech. Rep. LHCb-2008-042, CERN-LHCb-2008-042, CERN, Geneva, 2008.

[154] O. Callot and S. Hansmann-Menzemer, *The Forward Tracking: Algorithm and Performance Studies*, Tech. Rep. LHCb-2007-015, CERN-LHCb-2007-015, CERN, Geneva, 2007.

[155] M. Needham and J. Van Tilburg, *Performance of the track matching*, Tech. Rep. LHCb-2007-020, CERN-LHCb-2007-020, CERN, Geneva, 2007.

[156] M. Needham, *Performance of the Track Matching*, Tech. Rep. LHCb-2007-129, CERN-LHCb-2007-129, CERN, Geneva, 2007.

[157] E. E. Bowen, B. Storaci, and M. Tresch, *VeloTT tracking for LHCb Run II*, Tech. Rep. LHCb-PUB-2015-024, CERN-LHCb-PUB-2015-024, LHCb-INT-2014-040, CERN, Geneva, 2016.

[158] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, arXiv:1408.1251.

[159] R. Aaij *et al.*, *Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019), no. LHCb-DP-2019-001 P04013, arXiv:1812.10790.

[160] O. Callot, *Downstream Pattern Recognition*, Tech. Rep. LHCb-2007-026, CERN-LHCb-2007-026, CERN, Geneva, 2007.

[161] M. De Cian, S. Farry, P. Seyfert, and S. Stahl, *Fast neural-net based fake track rejection in the LHCb reconstruction*, LHCb-PUB-2017-011.

[162] M. Needham, *Clone Track Identification using the Kullback-Liebler Distance*, Tech. Rep. LHCb-2008-002, CERN-LHCb-2008-002, LPHE-2008-002, CERN, Geneva, 2008.

[163] A. Perieanu, *A Fast Algorithm to Identify and Remove Clone Tracks*, Tech. Rep. LHCb-2008-020, CERN-LHCb-2008-020, CERN, Geneva, 2008.

[164] L. M. Garcia *et al.*, *Tracking performance for long-lived particles at lhcb*, Journal of Physics: Conference Series **1525** (2020) 012095.

[165] LHCb collaboration, R. Aaij *et al.*, *Measurement of the inelastic pp cross-section at a centre-of-mass energy of $\sqrt{s}$ =7 TeV*, JHEP **02** (2015) 029, arXiv:1412.2500.

[166] LHCb collaboration, R. Aaij *et al.*, *Measurement of the inelastic pp cross-section at a centre-of-mass energy of $\sqrt{s}$ =13 TeV*, JHEP **06** (2018) 100, arXiv:1803.10974.

[167] LHCb collaboration, D. Derkach, M. Hushchyn, and N. Kazeev, *Machine Learning based Global Particle Identification Algorithms at the LHCb Experiment*, EPJ Web Conf. **214** (2019) 06011.

[168] LHCb collaboration, F. Dordei, *LHCb detector and trigger performance in Run II*, EPJ Web Conf. **164** (2017) 01016.

[169] R. Aaij *et al.*, *The LHCb trigger and its performance in 2011*, JINST **8** (2013) P04022, arXiv:1211.3055.

[170] J. Albrecht, V. V. Gligorov, G. Raven, and S. Tolk, *Performance of the lhcb high level trigger in 2012*, Journal of Physics: Conference Series **513** (2014) 012001.

[171] LHCb collaboration, *LHCb trigger plots and diagrams for conferences*, https://twiki.cern.ch/twiki/bin/viewauth/LHCb/LHCbTriggerConferenceDiagramsPlots.

[172] P. Robbe, *Generators, Calorimeter Trigger and J/ψ production at LHCb*, PhD thesis, Orsay U., 2012, Presented 12 Mar 2012.

[173] O. Callot, M. Ferro-Luzzi, and P. Perret, *Using the SPD multiplicity in the Level-0 trigger*, Tech. Rep. LHCb-2003-022, CERN, Geneva, 2003.

[174] LHCb collaboration, R. Aaij *et al.*, *The LHCb upgrade I*, arXiv:2305.10515.

[175] LHCb collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021.

[176] LHCb collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001.

[177] R. Aaij *et al.*, *Allen: A high level trigger on GPUs for LHCb*, Comput. Softw. Big Sci. **4** (2020), no. 1 7, arXiv:1912.09161.

[178] D0 collaboration, B. Pifer *et al.*, *Design Report: An Experiment at D0 to Study Antiproton-Proton Collisions at* 2 *TeV*, .

[179] J. P. Cachemiche *et al.*, *The PCIe-based readout system for the LHCb experiment*, Journal of Instrumentation **11** (2016) P02013.

[180] LHCb collaboration, L. Henry, *The SciFi detector - Status report during the LHCb week in September 2023*, .

[181] LHCb collaboration, *OnlineMon/Shift/SCIFI: 1 - Hitmap - Page Information*, https://https://lbwebmonet.cern.ch.

[182] D. A. Berninghoff, *Commissioning of the Front-End Electronics of the LHCb Scintillating Fibre Tracker*, PhD thesis, Heidelberg University, 2022, CERN-THESIS-2022-053.

[183] CMS collaboration, G. Cummings, *CMS HCAL VTRx-induced communication loss and mitigation*, JINST **17** (2022), no. 05 C05020.

[184] K. Wyllie, *Status of VTRX issue – LHCb Technical Board Meeting*, https://indico.cern.ch/event/1040629/.

[185] FASER collaboration, A. Ariga *et al.*, *Technical Proposal for FASER: ForwArd Search ExpeRiment at the LHC*, Tech. Rep. CERN-LHCC-2018-036, LHCC-P-013, UCI-TR-2018-22, KYUSHU-RCAPP-2018-07, LHCC-P-013, CERN, Geneva, 2018.

[186] LHCb collaboration, *OnlineMon/Shift/SCIFI: 1 - Hitmap* , https://https://lbwebmonet.cern.ch.

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

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

[189] L. Anderlini *et al.*, *The PIDCalib package*, LHCb-PUB-2016-021.

[190] A. Poluektov, *Correction of simulated paritcle identification response in LHCb using kernel density estimation*, LHCb-INT-2017-007.

[191] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. **A555** (2005), no. LAL-04-07 356, arXiv:physics/0402083.

[192] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, *Data driven trigger efficiency determination at LHCb*, LHCb-PUB-2014-039.

[193] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. **A552** (2005) 566, arXiv:physics/0503191.

[194] W. D. Hulsbergen, *Decay chain fitting with a kalman filter*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **552** (2005), no. 3 566.

[195] D. Martínez Santos and F. Dupertuis, *Mass distributions marginalized over per-event errors*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **764** (2014) 150–155.

[196] Y.-K. Hsiao and C.-C. Lih, *Fragmentation fraction f_{Ω_b} and the $\Omega_b \to \Omega J/\psi$ decay in the light-front formalism*, Phys. Rev. D **105** (2022), no. 5 056015, arXiv:2110.00945.

[197] Particle Data Group, P. A. Zyla *et al.*, *Review of particle physics*, Prog. Theor. Exp. Phys. **2020** (2020), no. 8 083C01.

[198] L. Oliver, J.-C. Raynal, and R. Sinha, *Note on new interesting baryon channels to measure the photon polarization in $b \to s\gamma$*, Phys. Rev. **D82** (2010) 117502, arXiv:1007.3632.

[199] R. Marki, *Measurements of b-baryon properties at LHCb*, PhD thesis, Jan, 2014, CERN-THESIS-2014-029, Presented 06 Mar 2014.

[200] LHCb collaboration, R. Aaij *et al.*, *Evidence of a $J/\psi\Lambda$ structure and observation of excited Ξ^- states in the $\Xi_b^- \to J/\psi\Lambda K^-$ decay*, Science Bulletin **66** (2021), no. 13 1278, arXiv:2012.10380.

[201] M.-S. Liu, K.-L. Wang, Q.-F. Lü, and X.-H. Zhong, *Ω baryon spectrum and their decays in a constituent quark model*, Phys. Rev. D **101** (2020), no. 1 016002, arXiv:1910.10322.

[202] V. A. Epanechnikov, *Non-parametric estimation of a multivariate probability density*, Theory of Probability & Its Applications **14** (1969), no. 1 153, arXiv:https://doi.org/10.1137/1114019.

[203] LHCb collaboration, R. Aaij *et al.*, *Momentum scale calibration of the LHCb spectrometer*, JINST **19** (2024) P02008, arXiv:2312.01772.

[204] B. Efron, *Bootstrap methods: Another look at the jackknife*, Ann. Statist. **7** (1979) 1.

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

[206] L. M. García Martín *et al.*, *Radiative b-baryon decays to measure the photon and b-baryon polarization*, Eur. Phys. J. C **79** (2019), no. 7 634, arXiv:1902.04870.

[207] Y.-l. Liu, L.-f. Gan, and M.-q. Huang, *The exclusive rare decay b->s gamma of heavy b-baryons*, Phys. Rev. D **83** (2011) 054007, arXiv:1103.0081.

[208] K. Azizi, Y. Sarac, and H. Sundu, *Light cone QCD sum rules study of the semileptonic heavy Ξ_Q and Ξ'_Q transitions to Ξ and Σ baryons*, Eur. Phys. J. **A48** (2012) 2, arXiv:1107.5925.

[209] R.-M. Wang, Y.-G. Xu, C. Hua, and X.-D. Cheng, *Studying $\mathcal{B}_1(\frac{1}{2}^+) \to \mathcal{B}_2(\frac{1}{2}^+)\ell^+\ell^-$ Semi-leptonic Weak Baryon Decays with the SU(3) Flavor Symmetry*, Phys. Rev. D **103** (2021), no. 1 013007, arXiv:2101.02421.

[210] K. Azizi, A. T. Olgun, and Z. Tavukoglu, *Impact of scalar leptoquarks on heavy baryonic decays*, Adv. High Energy Phys. **2017** (2017) 7435876, arXiv:1609.09678.

[211] I. I. Bigi, T. Mannel, and N. Uraltsev, *Semileptonic width ratios among beauty hadrons*, Journal of High Energy Physics **2011** (2011) .

[212] M. J. Savage and M. B. Wise, *SU(3) Predictions for Nonleptonic B Meson Decays to Charmed Baryons*, Nucl. Phys. B **326** (1989), no. CALT-68-1544 15.

[213] M. B. Voloshin, *Remarks on measurement of the decay $\Xi_b^- \to \Lambda_b\pi^-$*, arXiv:1510.05568.

[214] Y. K. Hsiao, P. Y. Lin, L. W. Luo, and C. Q. Geng, *Fragmentation fractions of two-body b-baryon decays*, Phys. Lett. B **751** (2015) 127, arXiv:1510.01808.

[215] D0 collaboration, V. M. Abazov *et al.*, *Measurement of the production fraction times branching fraction $f(b \to \Lambda_b) \cdot \mathcal{B}(\Lambda_b \to J/\psi\Lambda)$*, Phys. Rev. D **84** (2011), no. FERMILAB-PUB-11-206-E 031102, arXiv:1105.0690.

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

[217] F. J. Massey, *The Kolmogorov-Smirnov Test for Goodness of Fit*, Journal of the American Statistical Association **46** (1951), no. 253 68.

[218] A. Rogozhnikov, *Reweighting with Boosted Decision Trees*, J. Phys. Conf. Ser. **762** (2016), no. 1 012036, arXiv:1608.05806, https://github.com/arogozhnikov/hep_ml.

[219] A. L. Read, *Presentation of search results: The CL_s technique*, J. Phys. **G28** (2002) 2693.

[220] L. Moeser, *Search for tree-level Ξ_b^0 decays at the LHCb experiment*, Master thesis, 2024, Master thesis handed in 23 Feb 2024.

[221] T. Gutsche *et al.*, *Polarization effects in the cascade decay $\Lambda_b^0 \to \Lambda(\to p\pi) + J/\psi(\to \ell^+\ell^-)$in the covariant confined quark model*, Physical Review D **88** (2013) .

[222] T. Gutsche *et al.*, *Towards an assessment of the ATLAS data on the branching ratio $\Gamma(\Lambda_b^0 \to \psi(2S)\Lambda)/\Gamma(\Lambda_b^0 \to J/\psi\Lambda)$*, Physical Review D **92** (2015) .

[223] Z.-T. Wei, H.-W. Ke, and X.-Q. Li, *Evaluating decay rates and asymmetries of Λ_b into light baryons in the light-front quark model*, Phys. Rev. D **80** (2009) 094016.

[224] L. Mott and W. Roberts, *Rare dilepton decays of Λ_b^0 in a quark model*, International Journal of Modern Physics A **27** (2012) 1250016.

[225] LHCb collaboration, R. Aaij *et al.*, *Measurement of the $\Lambda_b^0 \to J/\psi\Lambda$ angular distribution and the Λ polarisation in pp collisions*, JHEP **06** (2020) 110, arXiv:2004.10563.

[226] ATLAS collaboration, G. Aad *et al.*, *Measurement of the parity-violating asymmetry parameter α_b and the helicity amplitudes for the decay $\Lambda_b^0 \to J/\psi\Lambda^0$ with the ATLAS detector.* , Phys. Rev. D **89** (2014), no. CERN-PH-EP-2014-034, CERN-PH-EP-2014-034 092009, arXiv:1404.1071.

[227] CMS collaboration, A. M. Sirunyan *et al.*, *Measurement of the Λ_b polarization and angular parameters in $\Lambda_b \to J/\psi\,\Lambda$ decays from pp collisions at $\sqrt{s} = 7$ and 8 TeV*, Phys. Rev. D **97** (2018), no. CMS-BPH-15-002, CERN-EP-2017-331 072010, arXiv:1802.04867.

[228] P. Böer, T. Feldmann, and D. van Dyk, *Angular analysis of the decay $\Lambda_b^0 \to \Lambda(\to N\pi)\ell^+\ell^-$*, Journal of High Energy Physics **2015** (2015) .

[229] D. van Dyk *et al.*, *EOS: a software for flavor physics phenomenology*, The European Physical Journal C **82** (2022) .

[230] D. van Dyk. Private communication.

[231] Heavy Flavor Averaging Group Collaboration, Y. Amhis *et al.*, *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2021*, Phys. Rev. D **107** (2023) 052008.

[232] J. Podolanski and R. Armenteros, *Iii. analysis of v-events*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **45** (1954), no. 360 13.

[233] Particle Data Group, M. Tanabashi *et al.*, *Review of particle physics*, Phys. Rev. **D98** (2018) 030001.

[234] LHCb collaboration, R. Aaij *et al.*, *Amplitude analysis of the $B^0 \to K^{*0}\mu^+\mu^-$ decay*, Phys. Rev. Lett. **132** (2024) 131801, arXiv:2312.09115.

[235] 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.

[236] G. Hiller, M. Knecht, F. Legger, and T. Schietinger, *Photon polarization from helicity suppression in radiative decays of polarized Λ_b to spin-3/2 baryons*, Phys. Lett. **B649** (2007), no. DO-TH-07-01, LPHE-2007-01 152, arXiv:hep-ph/0702191.

[237] T. Mannel and Y.-M. Wang, *Heavy-to-light baryonic form factors at large recoil*, JHEP **12** (2011), no. SI-HEP-2011-12 067, arXiv:1111.1849.

[238] A. Hocker *et al.*, *TMVA - Toolkit for Multivariate Data Analysis*, arXiv:physics/0703039.

[239] H. Deng, Y. Zhou, L. Wang, and C. Zhang, *Ensemble learning for the early prediction of neonatal jaundice with genetic features*, BMC Medical Informatics and Decision Making **21** (2021) .

[240] J. Albrecht, D. Berninghoff, and V. V. Gligorov, *Bremsstrahlung recovery of electrons using multivariate methods*, LHCb-INT-2016-018, 2016.

[241] LHCb collaboration, R. Aaij *et al.*, *Measurement of the electron reconstruction efficiency at LHCb*, JINST **14** (2019) P11023, arXiv:1909.02957.

Acknowledgements

The journey of my PhD, did not start with the PhD project itself, but rather with the start of my Master thesis project and the COVID-19 pandemic. Sitting on packed boxes, the closing of the European boarders was announced and my initial plans fell through. While this time was characterised by uncertainty, Johannes offered me a Master thesis project on beauty baryons. All obstacles of teleworking and meeting the Dortmund LHCb group not a single time in the full year, I really enjoyed my work and it was clear that I wanted to continue my research on beauty baryons.
At this point, I want to thank you, Johannes, for offering me a PhD position and being open for my wish to proceed the project as a cotutelle as well as based at CERN. Thank you for contacting people you trust to offer me possibilities to choose from. I enjoyed and still enjoy working with you. I appreciate the freedom and trust you offered me to try out my own ideas and uncommon approaches. This helped me a lot to become the scientist I am today. Simultaneously, I am very grateful, that you still took the time, whenever I needed help and discussed many things with me. Your advise guided me through some difficult decisions and taught me to always think about the bigger picture. And last but not least, thank you for suffering through an uncommonly long PhD thesis especially for Dortmund standards, I hope you will forgive me for it.
Similarly, I want to give a big thank you to you, Marie-Hélène. Thank you for agreeing to the experiment of a cotutelle PhD project with a student interested in too many projects at the same time. I really appreciate our physics discussions and learned a lot about angular analyses from you. You trained me to not lose focus and trust in my own capabilities, as well as how to actually write a PhD thesis and not an internal documentation with too many details. Especially, I want to thank you for being supportive through uncertain times of my PhD that were not related to work.
At this point, I would like to express my gratitude to Chris Delitzsch and Jonas Rademacker for reviewing my thesis as referees. I am sure you will help me to improve the quality of this document further. In addition, I would like to thank Justine Serrano, Tim Gershon, Danny van Dyk and Fabien Cavalier to join my defence as examiners. It is a pleasure for me that all of you attend the thesis defence in person.
I would also like to thank Dirk for introducing me to the work at the SciFi detector and trusting me to do the job right. It was a pleasure to work on the detector and actively contribute to the first Upgrade.
But my biggest thank you goes to you, Vitalii. I cannot express enough the gratitude I feel about all the support I got from you over the last four years. Looking back, I am wondering whether I would have made it so far, if you hadn't had the patience to answer

all my little and annoying questions especially in the beginning of my Master thesis. I cannot imagine a better supervisor through a pandemic and beyond. I learned a lot from you and still am. Your fascination for rare baryon decays sparked a lot of research interest and our discussions helped me to deepen my understanding for a lot of things. Thank you for your patience with me, it was an absolute pleasure to work with you and under your guidance on so many projects. Without you and our work together, this thesis would probably look very different. Thank you for taking the time reviewing this whole document and ensuring that everything is discussed clearly. But the best part for me is, that we started as advisor and student, became colleagues and ended this journey as friends. I am looking forward to what comes next.
I would like to thank all members of the LHCb groups in Dortmund and at the IJCLab for the welcoming atmosphere and always providing support. I really appreciate it, especially since I was on the move so much. I would like to thank Yasmine for being an amazing person and always supportive even in the toughest times. You make the LHCb collaboration a better place. I would also like to thank Felicia for our chats, Gaëlle for helping me when I hit the wall while coding again and Piera and Marie for always giving you a reason to smile. Thank you to Chiara, Valeriia and Vsevolod for the great evenings and journeys back in the RER B from Paris. I would like to thank Lukas for all rants about dealing with German and French administration simultaneously, as well as all the great after-work events. I would like to thank Micol for making CERN more fun and Elena for being an amazing person and introducing me to the operational side of LHCb. A special thanks also goes to Anja for always missing you in Dortmund and platonically falling in love with you after meeting for the first time in person. I am always happy to start the next riot with you. I also want to thank all the other LHCb colleagues that ensure that LHCb is always a friendly working environment.
A special thanks also goes to my CERN friends, that became a second family. You made all of this so much easier. Thank you Grace, Jonas, Andrés, Elisabeth, Sam, Sophie, Dylan, Tara, Bing and Anders. You make academia a more enjoyable place. I am equally thankful for my friends back home, Melina, Yvonne, Julia, Julian and Robert. Thank you for putting up with the long distance and knowing that I will always travel back to Germany to people that make me feel home. And last but not least, there have been a couple of friends, that have not been in my life for a very long time, but they are a core reason I did not go insane in the last year of my thesis. Thank you Martino, Wise, Sebastian, Aleksandra, Alice, Federica, Flo, Hendrik, Aleksandra and Ginny. And of course a big thank you to my family, your support and encouragement got me here in the first place. I am grateful to had you on my side through the whole process.
Kevin a simple thank you is not enough, but I hope you know how much you and your unconditional support mean to me. You are the reason I did not give up and stood back up every single time.
And congratulations to every reader that made it so far. If I can give you one piece of advise, do not write your PhD thesis in 2 months. And if you have to, remember that there is more in life than a PhD! You got this.