THESE DE DOCTORAT

NNT : 2026UPASP076

Measurement of χ_c production in OO collisions at $\sqrt{s_{NN}}$ = 5.36 TeV with LHCb and tracking performance studies for the LHCb Upgrade II

Mesure de la production des χ_c dans les collisions OO à $\sqrt{s_{NN}}$ = 5,36 TeV avec LHCb et études des performances du système de trajectographie pour l'Upgrade II de LHCb

Thèse de doctorat de l'université Paris-Saclay et de l'université de Santiago de Compostela

École doctorale n◦576 : particules hadrons énergie et noyau : instrumentation, imagerie, cosmos et simulation (PHENIICS)
Spécialité de doctorat : Physique nucléaire
Graduate School : Physique. Référent : Faculté des sciences d'Orsay

Thèse préparée dans les unités de recherche **Département de Physique Nucléaire** (Université Paris-Saclay, CEA) et IGFAE (Universidade de Santiago de Compostela) sous la direction de **Javier CASTILLO CASTELLANOS**, Docteur, la co-direction de **Abraham Antonio GALLAS TORREIRA**, Docteur et le co-encadrement de **Benjamin AUDURIER**, Docteur

Thèse soutenue à Paris-Saclay, le 28 Septembre 2026, par

Carlos Javier BARBERO PRETEL

Composition du Jury

Membres du jury avec voix délibérative

Patrick ROBBE Directeur de Recherche, IJCLab, Université Paris-Saclay	Président
Jhon Matthew DURHAM Directeur de Recherche, Los Alamos National Laboratory	Rapporteur
Ginés MARTINEZ GARCIA Directeur de Recherche, Subatech, IMT Atlantique	Rapporteur & Examinateur
Elisabeth Maria NIEL Chargée de Recherche, LLR, Institut Polytechnique de Paris	Examinatrice
Kara MATTIOLI Maîtresse de Conferences, Subatech, Université de Nantes	Examinatrice
Elena GONZALEZ FERREIRO Professeur, IGFAE, Universidade de Santiago de Compostela	Examinatrice

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

Titre : Mesure de la production des χ_c dans les collisions OO à $\sqrt{s_{NN}}$ = 5,36 TeV avec LHCb et études des performances du système de trajectographie pour l'Upgrade II de LHCb

Mots clés : Plasma de Quarks et de Gluons, production de charmonium, trajectographie.

Résumé : Cette thèse présente la première mesure de la production des états χ_c(1P) dans des collisions ion-ion aux énergies des collisionneurs, à partir de données de collisions O-O enregistrées par l'expérience LHCb. La production relative des états χ_{c1} et χ_{c2} est mesurée à travers leurs désintégrations radiatives en J/ψ et photon, le J/ψ étant reconstruit dans l'état final dimuon. L'analyse a nécessité des stratégies de sélection dédiées fondées sur l'apprentissage automatique pour l'identification des photons de basse énergie, ainsi qu'une calibration dédiée de la réponse du calorimètre électromagnétique et de l'efficacité de détection des photons dans l'environnement à forte multiplicité des collisions ion-ion. Le rapport de production χ_{c2}/χ_{c1} et la fraction de J/ψ provenant des désintégrations de χ_c sont mesurés et comparés aux résultats obtenus dans d'autres systèmes de collision, apportant de nouvelles contraintes sur la production du charmonium dans les collisions d'ions lourds. La thèse présente également des études des performances de trajectographie pour l'Upgrade II de LHCb. Différentes configurations du détecteur Upstream Pixel sont évaluées dans des collisions pp et PbPb centrales à l'aide de simulations dédiées. L'efficacité de reconstruction et le taux de fausses traces sont étudiés pour différentes configurations du détecteur, granularités des capteurs et quantités de matière. De plus, un nouvel algorithme de trajectographie fondé sur un automate cellulaire est développé pour la reconstruction autonome dans le détecteur Upstream Pixel, atteignant des efficacités supérieures à 90 % dans les environnements d'ions lourds à très forte multiplicité considérés.

Title : Measurement of χ_c production in OO collisions at $\sqrt{s_{NN}}$ = 5.36 TeV with LHCb and tracking performance studies for the LHCb Upgrade II

Keywords : Quark Gluon Plasma, charmonium production, tracking

Abstract : This thesis presents the first measurement of χ_c(1P) production in ion-ion collisions at collider energies using Oxygen-Oxygen (OO) collision data collected by the LHCb experiment. The relative production of χ_{c1} and χ_{c2} is measured through their radiative decays into J/ψ and a photon, with the J/ψ reconstructed in the dimuon final state. The analysis required dedicated machine-learning-based strategies for low-energy photon identification, as well as a dedicated calibration of the electromagnetic calorimeter response and photon detection efficiency in the high-occupancy environment of ion-ion collisions. The χ_{c2}/χ_{c1} production ratio and the fraction of J/ψ mesons originating from χ_c feed-down decays are measured and compared with results from other collision systems, providing new constraints on charmonium production in heavy-ion collisions.

The thesis also presents tracking-performance studies for the LHCb Upgrade II. Different Upstream Pixel detector designs are evaluated in pp and central PbPb collisions using dedicated simulations. Tracking efficiency and ghost-track rates are studied for different detector layouts, sensor granularities, and material budgets. In addition, a novel Cellular Automaton Tracking algorithm is developed for standalone reconstruction in the Upstream Pixel detector, achieving efficiencies above 90% in the highest-occupancy heavy-ion environments considered.

ESCOLA DE DOUTORAMENTO
INTERNACIONAL DA USC

DOCTORAL THESIS

Measurement of χ_c production in OO collisions at $\sqrt{s_{NN}} = 5.36$ TeV with LHCb and tracking performance studies for the LHCb Upgrade II

Author

Carlos Javier Barbero Pretel

Supervisor/s: Abraham Antonio Gallas Torreira, Javier Castillo Castellanos

Tutor: Abraham Antonio Gallas Torreira

PHD PROGRAMME IN NUCLEAR AND PARTICLE PHYSICS

SANTIAGO DE COMPOSTELA - 2026

DECLARATION OF CONFLICTS OF INTERESTS

The author of this thesis declares no real, potential, or apparent conflict of interest in relation to the elaboration, content, data analysis, or conclusions presented in this work.

Carlos Javier Barbero Pretel

Santiago de Compostela, June 2026

Acknowledgements

First and foremost, I would like to thank my supervisor Benjamin. He has been the person who has always been there throughout the entire duration of my PhD, weathering the ups and downs that come with it and offering his support in every situation along the way. As an expert in heavy ions, Ben has guided me, and it is thanks to his help that the completion of this thesis has been possible. I would also like to thank my thesis directors Abraham and Javier for their support over these years. I am certain that without their help, getting to this point would not have been possible. A special mention goes to Sam, who, as my Master's thesis supervisor, first introduced me to experimental physics. With his help, I was able to aspire to this PhD position. Thank you for being a wonderful college, both inside and outside of physics.

My thesis was divided into two stages: the first took place in Paris, the second in Santiago de Compostela, but in both I felt at home thanks to the people who shared this time with me. On the Parisian side, I would like to thank all my colleagues at DPhN in CEA, Paul, Mostafa, Gabriel, Chi and Mika, as well as the senior members Michael, Alberto and Stefano. I would especially like to thank Imanol, Alisha and Carolina, who taught me to see Paris through better eyes. On the Galician side, I want to thank all the people in my LHCb group at IGFAE, Victoria, Javier, Cris, Clara and Adrián. Of course, I also want to thank the senior members at IGFAE, Cibran, Juan and Maria. Here, my most heartfelt thanks go to Lidia, Ivan and Diego for welcoming me into your wonderful land and making me feel like one more Galician, even if my accent never quite fit in. They have been my main support during my time in Santiago. A special mention also goes to Camille for being an incredible source of support both inside and outside of work.

I would also like to extend my thanks to everyone in the LHCb experiment, especially the people in the IFT group who have guided me and helped me with their expertise in the field of heavy ions, with a special mention to the best conveners Sam and Oscar.

Carlos Barbero

On a more personal note, I would like to thank and dedicate this thesis to all those people outside of academia who never stopped believing in me, even in my moments of doubt. To Papa and Mama, for their unconditional support in every circumstance, even when my research remained a mystery to them, and for giving me the education that made me the person I am today. To my Tata, for being the fundamental pillar of my life and always being there whenever I need her. To my niece Cloe and my nephew Eric, for being the light that brightens my days despite the distance, and for being my motivation for the future. To my cousin Lorena and my cousin Manuel, for walking this journey with me and making sure I never forget where I come from. I would also like to thank the family that one chooses. To my friends, or rather my younger brothers, Samuel, Manolo, David and Rubio, for being that breath of fresh air in the most difficult moments and for supporting me no matter what. Lastly, to my friends, or rather my older brothers, Jaza and Miguel, for giving me advice of immeasurable value in the most difficult moments, but also for joining me in the most important celebrations.

To each and every one of the people I mention here, there are no words enough to express my gratitude.

Thank you, merci beaucoup, graciñas. Una y mil veces gracias!

Abstract

This thesis presents the first measurement of $\chi_c(1P)$ production in ion-ion collisions at collider energies. The analysis is performed using Oxygen-Oxygen (OO) collision data collected by the LHCb experiment at CERN during the 2025 light-ion run, at a nucleon-nucleon center-of-mass energy of $\sqrt{s_{NN}} = 5.36$ TeV. The measurement provides new insight into charmonium production in heavy-ion collisions. Among the various probes used to study the Quark Gluon Plasma (QGP), charmonium states play a central role because their production is directly affected by the presence of the hot and dense medium. In particular, charmonium suppression arises from color screening in the QGP, with more weakly bound states expected to be more strongly suppressed than more tightly bound ones. Measurements of χ_c states constitute an essential ingredient for a complete understanding of charmonium production in ion-ion collisions and provide sensitivity to the thermodynamic properties of the QGP and to the sequential suppression pattern expected in a deconfined medium. Since a significant fraction of observed J/ψ mesons originates from the decays of excited charmonium states, measurements of χ_c production are crucial for disentangling direct and indirect J/ψ production. Furthermore, the different binding energies of the χ_{c1} and χ_{c2} states make them sensitive probes of the sequential suppression pattern expected in the QGP.

The relative production of the χ_{c1} and χ_{c2} states is measured through their radiative decays into a J/ψ meson and a photon, with the J/ψ reconstructed in the dimuon final state. The analysis required dedicated machine-learning-based selection strategies to identify low-energy photons from χ_c decays in the challenging environment of ion-ion collisions. A key aspect of the analysis was the dedicated calibration of the electromagnetic calorimeter response and the determination of the photon detection efficiency, both of which are particularly challenging for low-energy photons reconstructed in the high-occupancy environment of ion-ion collisions. From these measurements, the cross-section ratio $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ is extracted and compared with results obtained in other collision

systems. In addition, the fraction of J/ψ mesons originating from χ_c feed-down decays is measured. The measured feed-down fraction is found to be lower than previous determinations in pp and pPb collisions, although additional studies of systematic uncertainties are still required before quantitative conclusions on possible medium effects can be drawn. Together, these measurements provide the first experimental constraints on χ_c production in heavy-ion collisions at collider energies and establish the feasibility of precision studies of excited charmonium states in a high-occupancy heavy-ion environment.

The thesis also presents tracking-performance studies for LHCb Upgrade II, the next major upgrade of the experiment designed to operate under the conditions of the High-Luminosity LHC. Although the LHCb experiment will operate with luminosity leveling, at an instantaneous luminosity lower than the maximum that the HL-LHC can deliver, the resulting detector occupancies will still be extremely high. As part of this programme, several new detector concepts and reconstruction strategies are being investigated to cope with the extreme occupancies expected at the HL-LHC.

The studies focus on the impact of different design scenarios for the Upstream Pixel (UP) detector on the overall tracking performance. Detector performance is evaluated in both pp and central PbPb collisions using dedicated simulation samples, allowing the robustness of the reconstruction to be assessed in the highest occupancy conditions expected at Upgrade II. Tracking efficiency and ghost-track rates are used to compare different detector layouts, sensor granularities, and material budgets. These studies provided important input to the LHCb collaboration and contributed to the definition and optimization of the current Upgrade II detector design. Finally, a novel Cellular Automaton Tracking (CAT) algorithm was developed and implemented for standalone reconstruction in the UP detector. The algorithm achieves tracking efficiencies exceeding 90% even in the highest occupancy heavy-ion environments considered in this work, while maintaining an acceptable ghost-track rate. These results demonstrate the potential of cellular-automaton-based approaches for track reconstruction in the challenging conditions expected for future operation at the HL-LHC.

Taken together, the results presented in this thesis demonstrate both the feasibility of precision measurements of excited charmonium states in heavy-ion collisions and the readiness of novel tracking concepts for the demanding conditions foreseen at LHCb Upgrade II.

Keywords: Quantum Chromodynamics (QCD), LHCb, Oxygen-Oxygen collisions, Lead-Lead collisions, Quark–Gluon Plasma (QGP), charmonium production, charmonium feed-down, Upgrade II, tracking performance, Upstream Pixel

Résumé

Cette thèse présente la première mesure de la production de $\chi_c(1P)$ dans des collisions ion-ion à des énergies de collisionneur. L'analyse est réalisée à partir des données de collisions Oxygène-Oxygène (OO) collectées par l'expérience LHCb au CERN lors de la campagne d'ions légers de 2025, à une énergie dans le centre de masse nucléon-nucléon de $\sqrt{s_{NN}} = 5.36$ TeV. Cette mesure apporte un nouvel éclairage sur la production de charmonium dans les collisions d'ions lourds. Parmi les différentes sondes utilisées pour étudier le Plasma de Quarks et de Gluons (QGP), les états de charmonium jouent un rôle central car leur production est directement affectée par la présence du milieu chaud et dense. En particulier, la suppression du charmonium résulte de l'écrantage de couleur dans le QGP, les états les plus faiblement liés étant supposés plus fortement supprimés que les états plus fortement liés. Les mesures des états χ_c constituent un ingrédient essentiel pour une compréhension complète de la production de charmonium dans les collisions ion-ion et fournissent une sensibilité aux propriétés thermodynamiques du QGP ainsi qu'au schéma de suppression séquentielle attendu dans un milieu déconfiné. Étant donné qu'une fraction significative des mésons J/ψ observés provient des désintégrations d'états de charmonium excités, les mesures de la production de χ_c sont cruciales pour distinguer la production directe et indirecte de J/ψ. De plus, les différentes énergies de liaison des états χ_{c1} et χ_{c2} en font des sondes sensibles au schéma de suppression séquentielle attendu dans le QGP.

La production relative des états χ_{c1} et χ_{c2} est mesurée via leurs désintégrations radiatives en un méson J/ψ et un photon, le J/ψ étant reconstruit dans l'état final dimuon. L'analyse a nécessité des stratégies de sélection dédiées basées sur l'apprentissage automatique pour identifier les photons de basse énergie issus des désintégrations de χ_c dans l'environnement difficile des collisions ion-ion. Un aspect clé de l'analyse a été l'étalonnage dédié de la réponse du calorimètre électromagnétique et la détermination de l'efficacité de détection des photons, ces deux aspects étant particulièrement difficiles pour les photons de basse

énergie reconstruits dans l'environnement à forte occupation des collisions ion-ion. À partir de ces mesures, le rapport de section efficace $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ est extrait et comparé aux résultats obtenus dans d'autres systèmes de collision. De plus, la fraction de mésons J/ψ provenant des désintégrations en cascade ("feed-down") des χ_c est mesurée. La fraction de feed-down mesurée est trouvée inférieure aux déterminations précédentes dans les collisions pp et pPb, bien que des études supplémentaires des incertitudes systématiques soient encore nécessaires avant de pouvoir tirer des conclusions quantitatives sur les éventuels effets du milieu. Ensemble, ces mesures fournissent les premières contraintes expérimentales sur la production de χ_c dans les collisions d'ions lourds à des énergies de collisionneur et établissent la faisabilité d'études de précision des états de charmonium excités dans un environnement d'ions lourds à forte occupation.

Cette thèse présente également des études de performance de trajectographie pour LHCb Upgrade II, la prochaine mise à niveau majeure de l'expérience conçue pour fonctionner dans les conditions du LHC à haute luminosité (HL-LHC). Dans le cadre de ce programme, plusieurs nouveaux concepts de détecteurs et stratégies de reconstruction sont à l'étude afin de faire face aux occupations extrêmes attendues au HL-LHC.

Ces études se concentrent sur l'impact de différents scénarios de conception pour le détecteur Upstream Pixel (UP) sur la performance globale de trajectographie. La performance du détecteur est évaluée à la fois dans les collisions pp et PbPb centrales à l'aide d'échantillons de simulation dédiés, permettant d'évaluer la robustesse de la reconstruction dans les conditions d'occupation les plus élevées attendues pour Upgrade II. L'efficacité de trajectographie et les taux de fausses traces ("ghost tracks") sont utilisés pour comparer différentes configurations de détecteur, granularités de capteurs et budgets de matière. Ces études ont fourni une contribution importante à la collaboration LHCb et ont participé à la définition et à l'optimisation de la conception actuelle du détecteur Upgrade II. Enfin, un nouvel algorithme de trajectographie par automate cellulaire (Cellular Automaton Tracking, CAT) a été développé et mis en œuvre pour la reconstruction autonome dans le détecteur UP. L'algorithme atteint des efficacités de trajectographie supérieures à 90% même dans les environnements d'ions lourds à la plus forte occupation considérés dans ce travail, tout en maintenant un taux de fausses traces acceptable. Ces résultats démontrent le potentiel des approches basées sur les automates cellulaires pour la reconstruction de trajectoires dans les conditions exigeantes attendues pour les futures opérations sur ions lourds au HL-LHC.

Dans leur ensemble, les résultats présentés dans cette thèse démontrent à la fois la faisabilité de mesures de précision des états de charmonium excités dans les collisions d'ions lourds et la maturité des nouveaux concepts de trajectographie pour les conditions exigeantes prévues pour LHCb Upgrade II.

Mots-clés: Chromodynamique Quantique (QCD), LHCb, collisions Oxygène-Oxygène,

collisions Plomb-Plomb, Plasma de Quarks et de Gluons (QGP), production de charmonium, feed-down du charmonium, Upgrade II, performance de trajectographie, Upstream Pixel

Index

Acknowledgements **i**

Abstract **iii**

Résumé **v**

1 Resumo en galego da tese **5**
1.1 Introdución . . . 5
1.1.1 Motivación teórica . . . 6
1.1.2 O experimento LHCb no LHC . . . 7
1.1.3 Calibración do ECAL para colisións PbPb e OO . . . 7
1.1.4 Produción de χ_c en colisións OO . . . 8
1.1.5 Eficiencias e incertezas sistemáticas . . . 9
1.1.6 Resultados e discusión . . . 9
1.1.7 LHCb Upgrade II . . . 10
1.1.8 Estudos de rendemento de seguimento de trazas . . . 11
1.1.9 Conclusións . . . 12

2 Research Objectives, Hypotheses, and Methodology **15**
2.1 Objectives . . . 15
2.2 Hypotheses . . . 16
2.3 Methodology . . . 16

3 Introduction **19**

I χ_c analysis **23**

4 Theoretical motivation **25**
4.1 The Standard Model of Particle Physics . . . 25
4.2 Quantum Chromodynamics . . . 28
4.3 The Quark Gluon Plasma . . . 30
4.4 Heavy-Ions collisions . . . 31
4.4.1 Collision evolution . . . 33

4.4.2 Geometry of the collision . . . 35
4.5 Quarkonium production in heavy-ion collisions . . . 37
4.5.1 Quarkonium production in pp collisions . . . 38
4.5.2 Cold nuclear matter effects . . . 41
4.5.3 Quarkonium in AA collisions . . . 44
4.5.3.1 Overview of theoretical models . . . 46
4.5.3.2 Experimental overview of charmonium production in AA collisions . . . 48
4.5.3.3 The importance of measuring χ_c states . . . 50

5 LHCb experiment at the LHC 53
5.1 The Large Hadron Collider . . . 53
5.2 The LHCb Detector in Run 3 . . . 55
5.2.1 Dipole Magnet . . . 57
5.2.2 Tracking system . . . 57
5.2.2.1 Vertex Locator . . . 58
5.2.2.2 Upstream tracker . . . 59
5.2.2.3 Scintillating Fibre tracker . . . 60
5.2.3 Particle identification . . . 60
5.2.3.1 RICH detectors system . . . 61
5.2.3.2 Calorimeter system . . . 62
5.2.3.3 Muon system . . . 63
5.2.4 Track reconstruction . . . 64
5.2.4.1 HLT1 reconstruction . . . 65
5.2.4.2 HLT2 reconstruction . . . 66
5.3 LHCb software and simulation framework . . . 68
5.3.1 Trigger system and data flow . . . 68
5.3.2 Simulation framework . . . 69

6 ECAL calibration for PbPb and OO collisions 71
6.1 Calibration strategy for PbPb and OO . . . 73
6.2 Impact on photon PID . . . 74
6.3 Data sample and selection . . . 76
6.4 Signal extraction . . . 77
6.5 Calibration results . . . 79
6.5.1 Mass correction over time . . . 79
6.5.2 Mass correction as function of E^γ and ECAL region . . . 81

7 χ_c production in OO collisions 85
7.1 Data Sample and Selection . . . 86
7.1.1 Data and Monte Carlo samples . . . 86
7.1.2 Candidates selection in data and MC . . . 87

7.2 Machine learning enhanced signal extraction . . . 91
7.2.1 Fundamentals of Machine Learning . . . 91
7.2.2 Boosted Decision Tree . . . 92
7.2.2.1 XGBoost . . . 93
7.2.2.2 Training Samples . . . 94
7.2.2.3 Training Variables . . . 95
7.2.2.4 Hyperparameters Selection . . . 96
7.2.2.5 BDT model validation . . . 98
7.2.2.6 Predicted probabilities from the BDT . . . 101
7.2.2.7 Data filtering with the BDT . . . 103
7.3 Signal Extraction . . . 104
7.3.1 Combinatorial Background Description . . . 104
7.3.2 Modeling the Signal Shape . . . 107
7.3.3 Application of the photon energy calibration . . . 108
7.3.4 Fit to the invariant mass spectrum . . . 109
7.4 J/ψ production in OO collisions . . . 115

8 Efficiencies and Systematic Uncertainties 117
8.1 Geometrical Acceptance . . . 118
8.2 Reconstruction and selection efficiency . . . 119
8.3 Tracking and muon ID efficiency . . . 121
8.4 Photon detection efficiency . . . 127
8.4.1 Measurement strategy . . . 127
8.4.2 Data selection and MC simulations . . . 129
8.4.3 Measurement of fully reconstructed η^0 and ω . . . 130
8.4.4 Measurement of partially reconstructed η^0 and ω . . . 134
8.4.5 Kinematic projection matrices . . . 139
8.4.6 Results of the photon detection efficiency . . . 141
8.5 Systematic Uncertainties and prospects . . . 146
8.5.1 Signal extraction . . . 146
8.5.2 χ_c polarization . . . 147
8.5.3 Systematic uncertainty associated with the BDT selection . . . 148
8.5.4 Systematic uncertainty associated with the photon detection efficiency 148

9 Results and Discussion 149
9.1 Cross-section ratio $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ . . . 149
9.2 Fraction of χ_c decays to J/ψ . . . 152

II Tracking performance for LHCb Upgrade II 157

10 LHCb Upgrade II 159
10.1 Upgrade II Frame . . . 160

10.2 Physics program . . . 161
10.2.1 Heavy-ions physics program . . . 162
10.3 Subdetectors requirements . . . 164
10.3.1 Tracking system . . . 164
10.3.2 Particle Identification Systems . . . 166
10.3.3 Muon System . . . 166
10.3.4 Real-Time Analysis and Online Systems . . . 166
10.4 Expected performance . . . 167

11 Tracking performance studies **169**
11.1 Pattern Recognition workflow . . . 170
11.2 Tracking algorithms . . . 172
11.2.1 VELO tracks algorithm . . . 172
11.2.2 UP standalone algorithm . . . 173
11.2.3 TTracks algorithm . . . 174
11.2.4 Upstream matching algorithm . . . 174
11.2.5 Downstream and long tracks matching algorithm . . . 175
11.3 Performance results . . . 177
11.3.1 UP stations geometry comparison . . . 179
11.3.2 UP beam pipe hole comparison . . . 182
11.3.3 UP pixel size comparison . . . 185
11.3.4 Global performance in *pp* and PbPb collisions . . . 186
11.4 Cellular Automaton Tracking algorithm . . . 191
11.4.1 Cellular Automaton applied to tracking . . . 191
11.4.2 Implementation of UP CAT algorithm . . . 192
11.4.3 CAT algorithm tests . . . 196
11.4.4 CAT algorithm in different collision systems . . . 198
11.5 Summary . . . 200

12 Conclusions **203**

A Permissions and reuse licenses **207**

List of figures **213**

List of tables **223**

Glossary **225**

References **229**

1

Resumo en galego da tese

1.1 Introdución

A física de partículas subatómicas ocúpase das formas máis fundamentais da materia e das interaccións que as rexen a escalas de lonxitude inferiores á do átomo. Un dos fenómenos máis fascinantes que pode ocorrer cando os núcleos atómicos coliden entre si a enerxías extremas é a creación dun estado da materia coñecido como plasma de quarks e gluóns (QGP), no que os quarks e os gluóns xa non están confinados no interior dos hadróns, senón que se moven con liberdade nunha masa densa e quente. O estudo deste estado da materia constitúe a motivación central desta tese.

Os estados de charmonio excitados, designados como χ_c, son mesóns formados por un quark charm e o seu antiquark, e resultan de especial relevancia como sondas do QGP. Debido ás súas enerxías de ligadura comparativamente débiles con respecto ao J/ψ, os estados χ_c son particularmente sensibles aos efectos de disociación termal neste medio denso e quente. Ademais, estes estados contribúen de maneira significativa á produción de J/ψ a través da chamada alimentación descendente (*feed-down*): unha fracción considerable dos J/ψ observados experimentalmente non son producidos de forma directa, senón que proceden de decaementos radiativos dos χ_c. En consecuencia, comprender a supresión dos estados χ_c no QGP resulta imprescindible para interpretar correctamente as medicións de J/ψ en colisións de ións pesados.

A análise foi deseñada inicialmente para colisións PbPb, pero o elevado fondo combinatorio e a alta ocupación do calorímetro electromagnético nese sistema de colisión fixeron que a extracción do sinal do χ_c resultase particularmente difícil. A campaña de toma de datos Osíxeno-Osíxeno (OO), levada a cabo no LHC en 2025, brindou a oportunidade de realizar

a primeira medición de χ_c en colisións de ións pesados nunha contorna máis limpa, pero na que aínda é posible a produción dun medio denso e quente (QGP). Esta tese presenta, en consecuencia, a primeira observación e medición da produción de $\chi_c(1P)$ en colisións ión-ión a enerxías de colisor, e aporta os primeiros datos experimentais sobre a supresión de estados de charmonio excitado nun sistema nuclear a enerxías do LHC.

A tese estrutúrase en dúas partes diferenciadas. A primeira dedícase á medición da produción de χ_c en colisións OO utilizando o detector LHCb na súa configuración do Run 3. A segunda parte céntrase no desenvolvemento de detectores e algoritmos de reconstrución para o futuro LHCb Upgrade II, o que garantirá o mantemento das prestacións de seguimento de trazas nas condicións extremas do LHC de alta luminosidade (HL-LHC). Aínda que abordan preguntas distintas, as dúas partes ilustran o carácter complementario da física experimental de partículas: a extracción de novos resultados e o desenvolvemento dos instrumentos necesarios para obtelos discorren en paralelo.

1.1.1 Motivación teórica

O marco teórico que serve de base a esta tese é o Modelo Estándar (ME) da física de partículas, formulado como unha teoría cuántica de campos baseada no grupo de simetría gauge $\mathrm{SU}(3)_c \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. Dentro del, a Cromodinámica Cuántica (QCD) describe a interacción forte que goberna a dinámica de quarks e gluóns. Dúas propiedades fundamentais da QCD determinan a fenomenoloxía relevante para esta tese: a liberdade asintótica, pola que a constante de acoplamento forte diminúe ao aumentar a enerxía de transferencia, e o confinamento, pola que os quarks e os gluóns non poden observarse como partículas illadas a baixas enerxías.

Cando núcleos atómicos pesados coliden a enerxías ultrarelativistas, créase durante un breve instante un medio de densidade e temperatura extremas coñecido como QGP. A evolución temporal dunha colisión de ións pesados comprende diversas etapas diferenciadas: a produción de partículas no instante da colisión, a termalización do medio, a expansión hidrodinámica do fluído de quarks e gluóns e, finalmente, a hadronización cando a temperatura descende por baixo da temperatura crítica de transición de fase.

Os estados de charmonio, formados por un par $c\bar{c}$, constitúen sondas especialmente valiosas do QGP. Os mecanismos de produción en colisións pp inclúen os modelos de Cor-Octet e Non-Relativistic QCD (NRQCD), así como o modelo de Cor-Singlete. En colisións con núcleos, aparecen tanto efectos de materia nuclear fría (modificacións das funcións de distribución partónica nucleares, dispersión múltiple dos partóns antes da colisión e absorción nuclear) como efectos do medio quente (disociación termal por pantallaxe de cor no QGP). O patrón secuencial de supresión previsto para os distintos estados de charmonio en función das súas enerxías de ligadura constitúe unha das sinaturas máis directas do QGP. Os estados χ_c, con enerxías de ligadura intermedias entre o J/ψ e o $\psi(2S)$, son particularmente sensibles a este mecanismo. Medir a súa fracción de alimentación descendente ao J/ψ en distintos sistemas de colisión é fundamental para separar a contribución

directa da indirecta na produción de J/ψ e para interpretar as observacións de supresión como evidencia de deconfinamento.

1.1.2 O experimento LHCb no LHC

O Gran Colisionador de Hadróns (LHC) do CERN é o acelerador de partículas máis potente construído ata a data. Opera acelerando dous feixes de protóns ou núcleos pesados en sentidos opostos ao longo dun anel de 27 km de circunferencia e facéndoos colidir en catro puntos de interacción onde se sitúan os principais experimentos. LHCb é un destes experimentos e está deseñado como un espectrómetro de ángulo cara a diante, cubrindo a rexión de pseudorapidez $2 < \eta < 5$, o que lle outorga ventaxas cinemáticas únicas para o estudo de quarkonia en colisións de ións pesados.

Na súa configuración do Run 3, o detector LHCb incorpora un sistema de seguimento de trazas composto por tres subdetectores. O Localizador de Vértices (VELO) é un detector de píxeles de silicio que rodea o punto de interacción e permite distinguir con gran precisión os vértices primarios, onde ocorre a colisión inicial, dos vértices secundarios, onde decaen os hadróns con charm ou beleza. O Upstream Tracker (UT) é un detector de bandas de silicio situado entre o RICH1 e o imán dipolar, que mellorou as prestacións de reconstrución con respecto ao detector previo. O Scintillating Fibre tracker (SciFi) está situado despois do imán e proporciona o seguimento das trazas nas rexións de alta pseudorapidez e momento.

Para a identificación das partículas, o detector dispón de dous detectores RICH que permiten identificar kaóns, pións e protóns; o sistema de calorímetros, formado polo ECAL e o HCAL, que mide a enerxía de electróns, fotóns e hadróns; e o sistema de muóns, composto por catro estacións instrumentadas con detectores gasosos e fundamental para a identificación dos $J/\psi \to \mu^+\mu^-$ que se utilizan nesta análise. O sistema de selección en liña (HLT1 e HLT2) filtra e procesa os datos en tempo real a 40 MHz sen ningún disparador hardware, o que representa un avance substancial respecto aos runs anteriores e permite o acceso a todos os sucesos que se producen no detector.

1.1.3 Calibración do ECAL para colisións PbPb e OO

A detección de fotóns de baixa enerxía é un dos ingredientes cruciais para a reconstrución dos estados χ_c a través do decaemento radiativo $\chi_c \to J/\psi\,\gamma$. O calorímetro electromagnético (ECAL) é o subdetector responsable desta medición, e a súa calibración precisa resulta esencial para garantir a correcta reconstrución da masa invariante dos candidatos χ_c.

Durante as campañas de toma de datos de PbPb en 2024 e OO en 2025 observáronse desviacións significativas das masas reconstruídas en varias canles coñecidas respecto aos seus valores nominais. Este desprazamento de masa ten a súa orixe nunha sobreestimación da enerxía dos fotóns polo ECAL, debida a dúas causas principais: unha descalibración

electrónica celda a celda e a degradación progresiva do detector polo efecto acumulado da radiación (*ageing*), que reduce a ganancia dos fotomultiplicadores e empeora a transmisión da luz nos cintiladores e fibras. Os efectos de envellecemento son especialmente pronunciados na rexión interna do ECAL, onde os niveis de radiación son máis elevados.

A calibración estándar baseada en $\pi^0 \to \gamma\gamma$ que se realiza de forma rutineira en colisións protón-protón non pode aplicarse en liña en colisións de ións pesados, debido ao elevado fondo e á estatística insuficiente nalgunhas celdas do ECAL. Por esta razón, deseñouse por primeira vez un procedemento de calibración específico para as condicións de colisión PbPb e OO en LHCb. O método usa π^0 resoltos, é dicir, π^0 cuxa desintegración en dous fotóns separados pode reconstruírse individualmente. Axústanse histogramas da masa do difotón en bins de enerxía do fotón e rexión do ECAL (interna, media e externa), e calcúlanse factores de corrección para cada celda en función do desprazamento observado respecto ao valor nominal da masa do π^0.

Os mapas de corrección obtidos mostran un desprazamento de masa de ata o 15% en PbPb e do 18% en OO nas celdas da rexión interna e á maior enerxía de fotón, compatible coa degradación progresiva do detector. A estabilidade temporal das correccións ao longo de cada campaña de toma de datos foi verificada e as correccións resultantes son aplicadas en análises posteriores, sendo un paso imprescindible para situar os candidatos χ_c na súa masa nominal e garantir a calidade dos axustes de masa invariante.

1.1.4 Produción de χ_c en colisións OO

A análise do canal $\chi_c \to J/\psi\gamma$ en colisións OO a $\sqrt{s_{NN}} = 5.36$ TeV require superar varios retos experimentais. A selección de candidatos realizase en dúas etapas: a primeira, ao nivel do sistema de procesamento en liña DaVinci, aplica cortes sobre variables cinemáticas. A segunda etapa consiste nunha selección fóra de liña que completa a definición da mostra de candidatos χ_c.

Para a extracción do sinal, empregouse un clasificador de aprendizaxe automática baseado nun Boosted Decision Tree (BDT). Adestrouse con simulacións de Monte Carlo para distinguir candidatos de sinal fronte a fondo combinatorio, utilizando variábeis das trazas dos muóns, do fotón e da xeometría do vértice. O axuste de hiperparámetros e a validación do modelo realizáronse mediante técnicas de validación cruzada, verificando que o BDT non está sobreaxustado e que proporciona boa separación entre sinal e fondo. Tras aplicar a selección do BDT, constrúese o espectro de masa invariante da diferenza $\Delta M = M(\mu^+\mu^-\gamma) - M(\mu^+\mu^-)$, que presenta picos correspondentes aos estados χ_{c1} e χ_{c2}.

Para suprimir o fondo de forma eficaz sen recorrer exclusivamente á selección BDT, aplícase un algoritmo de mestura de sucesos (*event mixing*) que permite construír unha distribución de fondo combinatorio a partir de fotóns e candidatos J/ψ pertencentes a sucesos distintos. O rendemento do axuste de sinal foi validado con modelos alternativos para a función de distribución dos picos de sinal (Crystal Ball) e con distintas

parametrizacións do fondo. A medición de J/ψ inclusivo realízase en paralelo e proporciona o denominador necesario para calcular a fracción de alimentación descendente $F(\chi_c \to J/\psi)$.

1.1.5 Eficiencias e incertezas sistemáticas

A extracción dos observábeis físicos require corrixir os rendementos de sinal polas distintas eficiencias que afectan á detección dos candidatos χ_c. Estas eficiencias inclúen a aceptancia xeométrica para os estados χ_{c1} e χ_{c2} (ε_{acc}), a eficiencia de reconstrución e selección fóra de liña ($\varepsilon^{\text{offline}}_{\text{rec\&sel}}$) e a eficiencia de selección do BDT ($\varepsilon^{\text{BDT}}_{\text{sel}}$).

Particular atención merece a eficiencia de detección de fotóns ($\varepsilon^{\gamma}_{\text{det}}$), cuxa determinación precisa é especialmente complexa nas condicións das colisións de ións pesados. Para medila, utilízanse os decaementos $\eta \to \pi^+\pi^-\pi^0(\gamma\gamma)$ e $\omega \to \pi^+\pi^-\pi^0(\gamma\gamma)$, que presentan topologías similares ao fotón de baixa enerxía nos decaementos χ_c. A eficiencia determínase de forma diferencial en bins de momento transverso do fotón, pseudorapidez e multiplicidade do suceso, o que permite aplicar a corrección de forma máis precisa.

As incertezas sistemáticas avalíanse de forma coidadosa para cada fonte identificada. Entre as máis importantes figuran as asociadas ao procedemento de extracción do sinal (escolla do modelo de distribución dos picos e do fondo), a eficiencia de detección de fotóns e os factores de corrección de polarización dos estados χ_c. Os factores de corrección de polarización, que teñen en conta a distribución angular anisótropa dos produtos do decaemento, determínanse como médias ponderadas sobre o espazo de fase cinemático relevante para a análise. A combinación de todas as eficiencias e das súas incertezas asociadas proporciona os factores de corrección necesarios para o cálculo dos observábeis físicos finais.

1.1.6 Resultados e discusión

A partir dos rendementos de sinal, as eficiencias e as razóns de ramificación dos decaementos, extráense dous observábeis físicos fundamentais. O primeiro é a razón de seccions eficaces de produción en estado prompt:

$$\frac{\sigma(\chi_{c2})}{\sigma(\chi_{c1})} = 2.17 \pm 0.25\,(\text{stat.}) \pm 0.25\,(\text{sist.}), \tag{1.1}$$

medida no rango cinematico $p_T(\chi_c) > 1.5\,\text{GeV}/c$ e $2 < \eta(\chi_c) < 4.8$ en colisións OO a $\sqrt{s_{NN}} = 5.36\,\text{TeV}$. As incertezas estatística e sistemática son de magnitude comparable, o que reflicte tanto a relevancia dos datos recollidos como a calidade dos métodos de estimación sistemática desenvolvidos. O valor obtido é lixeiramente superior ás medidas equivalentes realizadas por LHCb en colisións pp e pPb, pero esta comparación debe terse en conta con cautela, xa que a presente análise estendese ata momentos transversos

menores ($p_T > 1.5\,\mathrm{GeV}/c$), onde se espera que a razón sexa maior en virtude da dependencia coa enerxía observada en pp. O resultado é consistente dentro de 1.5σ cunha simulación de Pythia Angantyr, o que aporta confianza no modelado da produción relativa de χ_{c1} e χ_{c2} neste rango cinematico.

O segundo observábel extraído é a fracción de mesóns J/ψ prompt que proceden de decaementos de estados χ_c (*feed-down*):

$$F(\chi_{c1} \to J/\psi) = 0.051 \pm 0.004\,(\mathrm{stat.}) \pm 0.001\,(\mathrm{sist.}), \tag{1.2}$$
$$F(\chi_{c2} \to J/\psi) = 0.073 \pm 0.004\,(\mathrm{stat.}) \pm 0.010\,(\mathrm{sist.}), \tag{1.3}$$
$$F(\chi_{c} \to J/\psi) = 0.124 \pm 0.004\,(\mathrm{stat.}) \pm 0.010\,(\mathrm{sist.}). \tag{1.4}$$

Trátase da primeira determinación desta fracción nun sistema núcleo-núcleo a enerxías de colisor. O valor obtido en OO sitúase por debaixo das medicións anteriores en pp e pPb realizadas por LHCb e está a uns 2.9σ do resultado en pPb coas incertezas actuais. Se esta redución se confirmase tras completar a avaliación das incertezas sistemáticas pendentes, sería consistente cunha supresión reforzada dos estados χ_c no medio formado en colisións OO con respecto á produción directa de J/ψ, en consonancia cos modelos de disociación termal do QGP. Non obstante, unha conclusión definitiva sobre a supresión requirirá o estudo completo das incertezas sistemáticas e a comparación coa dependencia no rango cinemático entre os diferentes sistemas de colisión.

1.1.7 LHCb Upgrade II

A segunda parte desta tese muda o foco cara ao futuro do experimento LHCb, abordando o deseño e a validación de tecnoloxías e algoritmos necesarios para o LHCb Upgrade II. Esta actualización representa a próxima grande evolución do experimento e está concibida para funcionar nas condicións do LHC de alta luminosidade (HL-LHC), que operará a partir de mediados dos anos 2030. O obxectivo principal é acumular polo menos $300\,\mathrm{fb}^{-1}$ de luminosidade integrada a unha luminosidade instantánea de pico de $(1.0\text{-}1.5) \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, o que supón un incremento de case unha orde de magnitude con respecto ao Upgrade I.

Para afrontar este ambiente extremo, o detector debe incorporar novas tecnoloxías de seguimento de trazas con maior granularidade, maior tolerancia á radiación e capacidades melloradas de recoñecemento de patróns. A flexibilidade no deseño do Upgrade II articúlase en tres escenarios de escalado: o escenario de liña base, de máximas prestacións e custo; o escenario medio, optimizado en custo; e o escenario baixo, que é o de referencia actual e implica as maiores restricións no sistema de subdetectores.

O programa de física do Upgrade II é igualmente ambicioso. No sector de ións pesados, ao que se lle presta especial atención nesta tese, prevense medicións de gran precisión dos factores de modificación nuclear de estados de charmonio e bottomonium, a determinación do espectro de fotóns térmicos, medicións de razóns de produción de barións charm e beleza, e o estudo de exóticos como o $\chi_{c1}(3872)$ en colisións nucleares. Para alcanzar

estes obxectivos no contorno de altísima ocupación das colisións PbPb no HL-LHC, é imprescindible un sistema de seguimento de trazas capaz de manter a eficiencia e a baixa taxa de trazas falsas (*ghost tracks*) que caracterizou os runs anteriores.

O detector Upstream Pixel (UP), situado entre o VELO e o imán dipolar, está baseado en tecnoloxía de sensores de píxeles activos monolíticos (MAPS), que integran as funcións de detección e lectura nun só dispositivo. Esta tecnoloxía ofrece menor material dentro da aceptancia e permite a operación en condicións de alta ocupación e radiación. Os estudos de rendemento realizados sobre distintas configuracións do UP constitúen o contido principal da segunda parte desta tese.

1.1.8 Estudos de rendemento de seguimento de trazas

Un marco de recoñecemento de patróns (*pattern recognition framework*) desenvolvido expresamente para esta tese permite avaliar as prestacións do detector Upgrade II en diferentes configuracións. Estudáronse varias alternativas de deseño do detector UP, variando a xeometría do detector, a apertura do tubo de feixe e a granularidade dos píxeles. Os resultados mostraron que a xeometría sen bordes (*Borderless*), combinada cun tamaño de píxel de $50 \times 50\ \mu\text{m}^2$, representa o mellor compromiso entre eficiencia de seguimento, aceptancia xeométrica, cantidade de material e factibilidade tecnolóxica, e foi adoptada como configuración nominal de referencia para os estudos posteriores.

As prestacións do detector nominal foron avaliadas en colisións *pp* e en colisións PbPb centrais. Os resultados amosan eficiencias de seguimento similares nos dous sistemas de colisión, con só un descenso dun 7% de eficiencia en PbPb respecto a *pp* a pesar da maior ocupación do detector no ambiente de ións pesados. Isto acredita a viabilidade do deseño proposto para o amplo abano de sistemas de colisión do programa de física do Upgrade II.

Un aspecto central desta contribución é o desenvolvemento dun novo algoritmo de seguimento de trazas baseado nun Autómata Celular (CAT, *Cellular Automaton Tracking*). Este algoritmo foi deseñado para a reconstrución autónoma no detector UP sen información adicional dos outros subdetectores. O seu principio de funcionamento baséase na formación de *tracklets* —pares de impactos en planos consecutivos do detector— e na posterior aplicación de dous pasos iterativos denominados *Run Forward* e *Run Backward*. No paso *Run Forward*, asígnase a cada *tracklet* un estado de evolución que aumenta cando dúas *tracklets* veciñas comparten o mesmo estado e non superan un ángulo de rotura máximo. No paso *Run Backward* constrúense secuencias ordenadas de *tracklets* que minimizan un estimador de calidade que pondera a lonxitude da traza e a suavidade da traxectoria. Finalmente, aplícase un axuste parabólico das trazas candidatas e un procedemento de eliminación de clons.

Os resultados demostran que o algoritmo CAT mellora substancialmente as prestacións de seguimento autónomo no detector UP. A eficiencia de seguimento aumenta dun $\sim 79\%$ co

algoritmo previo ata un $\sim 95\%$ en colisións *pp*, e dun $\sim 75\%$ ata máis do 91% en colisións PbPb centrais. O algoritmo é tamén robusto fronte a variacións realistas da eficiencia intrínseca dos píxeles (entre o 96% e o 100%), sen diferenzas significativas no rendemento entre eses escenarios. O incremento de eficiencia ven acompañado dun aumento na taxa de trazas falsas, que pasa do 0.65 % ao 4.48 % en *pp* e do 3.79 % ao 15.99 % en PbPb, o que indica a necesidade de desenvolvementos futuros, como a implementación dun filtro de Kalman e estratexias máis sofisticadas de eliminación de clons, para mellorar a pureza da mostra reconstruída sen sacrificar as ganancias xa obtidas en eficiencia.

1.1.9 Conclusións

Esta tese abordou dous aspectos complementarios da física experimental de partículas de alta enerxía no experimento LHCb: a medida de precisión da produción de estados de charmonio excitados nun sistema de colisión de ións lixeiros e os estudos de desenvolvemento de detectores e algoritmos orientados a preservar as prestacións de seguimento de trazas do experimento na era do HL-LHC.

No que respecta á medición de χ_c en colisións OO, preséntase a primeira medición de $\chi_c(1P)$ en colisións ión-ión a enerxías de colisor, que constitúe igualmente a primeira observación de estados χ_c en colisións OO. Os estados χ_{c1} e χ_{c2} foron reconstruídos a través do seu decaemento radiativo $\chi_c \to J/\psi\,\gamma$ con $J/\psi \to \mu^+\mu^-$, grazas a un programa de traballo que abarcou a calibración do calorímetro electromagnético, a determinación da eficiencia de detección de fotóns de baixa enerxía e o deseño de estratexias de selección baseadas en aprendizaxe automática. A razón de seccions eficaces de produción prompt $\sigma(\chi_{c2})/\sigma(\chi_{c1}) = 2.17 \pm 0.25\,(\text{stat.}) \pm 0.25\,(\text{sist.})$ resulta consistente dentro de 1.5σ cunha simulación de Pythia Angantyr. A fracción de J/ψ procedente de alimentación descendente de χ_c, $F(\chi_c \to J/\psi) = 0.124 \pm 0.004\,(\text{stat.}) \pm 0.010\,(\text{sist.})$, sitúase por debaixo das medidas anteriores en *pp* e *p*Pb, o que, de confirmarse coa avaliación completa das incertezas sistemáticas, sería consistente cunha supresión incrementada dos estados χ_c no medio formado en colisións OO. Independentemente da interpretación final, esta análise demostra por primeira vez a factibilidade de medicións de precisión de χ_c en colisións de ións pesados a enerxías de colisor e establece as ferramentas experimentais necesarias para avanzar cara a medicións máis precisas en PbPb.

No que respecta ao Upgrade II, os estudos de rendemento de seguimento de trazas para o detector Upstream Pixel identificaron a xeometría sen bordes con píxeles de $50 \times 50\,\mu\text{m}^2$ como a configuración óptima e demostraron que o detector proposto pode operar eficientemente tanto en colisións *pp* como en colisións de ións pesados centrais. O novo algoritmo CAT aumenta a eficiencia de seguimento autónomo dun $\sim 79\%$ ata un $\sim 95\%$ en *pp* e dun $\sim 75\%$ a máis dun 91% en PbPb central, establecendo o potencial dos enfoques baseados en autómatas celulares para a reconstrución de trazas nas esixentes condicións do Upgrade II, e sentando as bases para desenvolvementos futuros que melloren a pureza da mostra mantendo os incrementos de eficiencia xa conseguidos.

No seu conxunto, as dúas partes desta tese ilustran o carácter dual da investigación en física experimental de partículas: por un lado, aportan nova información experimental sobre a produción de charmonio e os posibles efectos do medio a través da primeira medición de estados χ_c nun sistema ión-ión a enerxías de colisor; por outro, contribúen ao desenvolvemento das tecnoloxías de detectores e dos algoritmos de reconstrución necesarios para estender o programa de ións pesados de LHCb cara á era do LHC de alta luminosidade. As dúas partes conectan así as medicións do presente coa preparación da seguinte xeración de investigacións en colisións nucleares no LHC.

2

Research Objectives, Hypotheses, and Methodology

2.1 Objectives

The general objective of this thesis is to advance two complementary aspects of the LHCb physics and detector-development programme: the experimental study of charmonium production in heavy-ion collisions as a probe of the Quark–Gluon Plasma (QGP), and the development of tracking concepts required to sustain LHCb performance during the High-Luminosity LHC (HL-LHC) era. From this general aim, the following specific objectives are derived.

Part I: χ_c production in OO collisions

- Perform the first measurement of χ_c production in ion-ion collisions at collider energies, using Oxygen–Oxygen (OO) collision data collected by LHCb during the 2025 light-ion run at $\sqrt{s_{NN}} = 5.36$ TeV.
- Develop the experimental tools required for this measurement: a dedicated calibration of the electromagnetic calorimeter for PbPb and OO data-taking conditions, a determination of the low-energy photon detection efficiency, and machine-learning-based selection strategies capable of suppressing combinatorial background while preserving signal efficiency.
- Extract the prompt cross-section ratio $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ and the fraction of prompt J/ψ mesons originating from χ_c feed-down decays, and to compare these observables with previous measurements in pp and pPb collisions and with theoretical predictions.

Part II: Tracking performance for LHCb Upgrade II

- Evaluate the tracking performance of candidate designs for the Upstream Pixel (UP) detector under the extreme occupancy conditions expected at Upgrade II, considering different detector geometries, beam-pipe apertures, and pixel granularities, in both *pp* and central PbPb collisions.
- To develop and validate a Cellular Automaton Tracking (CAT) algorithm for standalone reconstruction in the UP detector, aiming to improve tracking efficiency while keeping the ghost-track rate at an acceptable level.

2.2 Hypotheses

- Since the χ_c states are more weakly bound than the J/ψ, color screening in the QGP medium is expected to suppress χ_c production more strongly than direct J/ψ production. Consequently, the fraction of J/ψ mesons originating from χ_c feed-down should be lower in heavy-ion collisions than in *pp* and *p*Pb collisions.
- Oxygen-Oxygen collisions, despite producing a smaller and shorter-lived medium than PbPb collisions, are expected to be hot and dense enough to yield measurable medium effects, while remaining sufficiently low in occupancy to make a precision χ_c measurement experimentally feasible, unlike in central PbPb collisions.
- A finer-granularity, low-material Upstream Pixel (UP) detector, combined with a cellular-automaton-based pattern-recognition algorithm, can preserve high standalone tracking efficiency even at the highest occupancies expected at Upgrade II, without an unacceptable increase in the ghost-track rate.

2.3 Methodology

- **Data and simulation:** use of OO collision data collected by LHCb during the 2025 light-ion run, together with dedicated Monte Carlo samples (Part I). Simulated pp and central PbPb samples for the Upgrade II tracking studies (Part II).
- **Detector calibration:** dedicated calibration of the electromagnetic calorimeter response, adapted from PbPb studies to OO data-taking conditions, required for reliable reconstruction of low-energy photons.
- **Signal reconstruction and selection:** reconstruction of χ_c candidates by combining $J/\psi \to \mu^+\mu^-$ candidates with calorimeter photons, with machine-learning-based classifiers employed to suppress combinatorial background while preserving signal efficiency.

- **Efficiency determination and systematic uncertainties:** determination of reconstruction, selection, and photon-detection efficiencies from data and simulation, with systematic uncertainties evaluated at each stage of the analysis.
- **Extraction of physical observables:** computation of the cross-section ratio $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ and the χ_c feed-down fraction to prompt J/ψ production, compared with previous pp and pPb measurements and with theoretical predictions.
- **Tracking-performance studies:** use of the LHCb Pattern Recognition framework to compare several Upstream Pixel detector configurations (geometry, beam-pipe aperture, pixel granularity) in terms of tracking efficiency and ghost-track rate, followed by the development and benchmarking of a Cellular Automaton Tracking (CAT) algorithm against the existing standalone algorithm in pp and central PbPb simulation.

3
Introduction

Understanding the behavior of strongly interacting matter under extreme conditions of temperature and energy density is one of the central goals of contemporary particle physics. Quantum Chromodynamics (QCD), the theory describing the strong interaction, predicts that at sufficiently high temperatures and densities, ordinary hadronic matter undergoes a phase transition into a state in which quarks and gluons are no longer confined within individual hadrons. This state, known as the Quark Gluon Plasma (QGP), is believed to have existed in the early universe microseconds after the Big Bang, and can be recreated, for extremely short timescales, in ultra-relativistic heavy-ion collisions at facilities such as the Large Hadron Collider (LHC).

Among the experimental probes used to characterize the QGP, charmonium states (bound states of a charm quark and its antiquark) occupy a central position. Their production in heavy-ion collisions is sensitive to color screening [1], a mechanism by which the presence of the hot and dense medium reduces the binding potential between the quark and antiquark, leading to a sequential suppression pattern that depends on the binding energy of each state. The $\chi_c(1P)$ states, with masses just above the J/ψ but with comparatively smaller binding energies, are predicted to be more easily dissociated in the medium than the J/ψ itself [2], making them sensitive probes of the properties and temperature of the medium. Furthermore, because a sizeable fraction of the observed J/ψ yield originates from the radiative decay of χ_c states, a precise determination of χ_c production is a necessary ingredient for disentangling direct and feed-down contributions to J/ψ production, and therefore for a coherent interpretation of charmonium suppression measurements as a whole.

PbPb collisions at the LHC constitute the benchmark system for studies of the Quark Gluon Plasma. Their large system size, high energy density, and well-established collective

behavior have enabled many of the major advances in our understanding of the QGP at collider energies. For this reason, the χ_c analysis presented in this thesis was initially developed for PbPb collisions, with the objective of studying the suppression of excited charmonium states in the environment where deconfinement effects are expected to be most pronounced.

However, the reconstruction of χ_c states through their radiative decay into $J/\psi\gamma$ proved particularly challenging in this collision system. The very high detector occupancy in central PbPb collisions leads to a large combinatorial background and substantial degradation of the electromagnetic calorimeter performance, severely limiting the sensitivity to the low-energy photons emitted in χ_c decays. Moreover, the available PbPb data sample was not large enough to compensate for these effects by applying tighter selection criteria, as this would have resulted in an insufficient number of signal candidates for a meaningful measurement.

The studies performed in PbPb collisions were however very useful, as they highlighted the main limitations of the analysis and motivated dedicated calorimeter calibrations, photon-efficiency studies and machine-learning-based selection strategies. The Oxygen-Oxygen data-taking campaign performed at the LHC in 2025 then provided an opportunity to pursue the first measurement of χ_c production in a collision system still expected to produce a hot and dense medium, while benefiting from a substantially lower detector occupancy. In this cleaner environment, the reconstruction of low-energy photons and the extraction of the χ_c signal become experimentally tractable, making OO collisions an ideal intermediate step towards future measurements in PbPb collisions.

This thesis presents the first measurement of $\chi_c(1P)$ production in heavy-ion collisions at collider energies and the first observation of χ_c states in Oxygen-Oxygen collisions. These measurements provide new experimental information on charmonium production in a collision system expected to form a hot and dense medium, while remaining experimentally accessible thanks to its lower detector occupancy.

The OO system, intermediate in size between proton-proton and lead-lead collisions, offers a unique opportunity to investigate the onset of collective and medium effects in a controlled nuclear environment, complementing the broader heavy-ion programme. The analysis reconstructs the χ_{c1} and χ_{c2} states through their radiative decay to a J/ψ meson and a photon, with the J/ψ identified through its decay to two muons. Because the photons emitted in this decay are typically of very low energy, their identification in the high-occupancy environment characteristic of ion-ion collisions is particularly challenging. Addressing this required a dedicated calibration of the electromagnetic calorimeter, a measurement of the photon detection efficiency, and the development of machine-learning-based selection strategies to suppress combinatorial background while preserving signal efficiency. From these measurements, the cross-section ratio $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ is extracted and compared with results obtained in other collision systems. By measuring the χ_c and J/ψ yields produced in OO collisions, this analysis provides the first experimental

constraints on the contribution of χ_c feed-down to charmonium production in a high-energy ion-ion collision system.

Part I of this thesis is devoted to the first measurement of χ_c production in OO collisions. It begins with Chapter 4, which reviews the theoretical concepts relevant for this work, with particular emphasis on charmonium production mechanisms and the role of χ_c states as probes of the QGP. Chapter 5 introduces the LHCb detector in its Run 3 configuration, while Chapter 6 describes the dedicated electromagnetic calorimeter calibration developed for PbPb and OO data-taking conditions.

The analysis strategy is presented in Chapter 7, including the event selection, the machine-learning-based signal extraction procedure, and the measurement of inclusive J/ψ production in OO collisions. Chapter 8 discusses the determination of the efficiencies entering the measurement, together with the associated systematic uncertainties. Finally, Chapter 9 presents the measured cross-section ratio $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ and the fraction of J/ψ mesons originating from χ_c feed-down decays.

While the first part of this thesis focuses on present-day measurements, the second part addresses the longer-term evolution of the LHCb experiment through Upgrade II, the major detector upgrade foreseen for operation during the High-Luminosity LHC era. The substantially higher instantaneous luminosities expected at Upgrade II require tracking systems with finer granularity, improved radiation tolerance, and enhanced pattern-recognition capabilities to maintain excellent reconstruction performance in very high occupancy environments.

Within this context, tracking-performance studies are presented for the proposed Upstream Pixel (UP) detector, exploring several geometrical configurations and sensor granularities in terms of tracking efficiency and ghost-track rate. These studies contributed directly to the optimization of the UP detector layout within the LHCb collaboration. In addition, a new Cellular Automaton Tracking (CAT) algorithm is developed and implemented for standalone reconstruction in the UP detector, with the objective of improving reconstruction efficiency while maintaining an acceptable ghost-track rate.

Part II is composed of two chapters. Chapter 10 introduces the Upgrade II framework, its physics programme, and the detector requirements arising from operation at the HL-LHC, with particular emphasis on the heavy-ion programme. Chapter 11 presents the tracking-performance studies, including the Pattern Recognition framework, the comparison of several UP detector configurations, and the development and validation of the Cellular Automaton Tracking (CAT) algorithm in pp and PbPb collisions. The thesis closes with Chapter 12, which summarizes the main conclusions drawn from both the χ_c measurement and the Upgrade II tracking studies.

Taken together, the two parts of this thesis illustrate the complementary nature of modern experimental particle physics. On the one hand, they provide new experimental information on charmonium production and possible medium effects through the first

measurement of χ_c states in a high-energy ion-ion collision system. On the other hand, they contribute to the development of detector technologies and reconstruction algorithms required to extend the heavy-ion programme of LHCb into the High-Luminosity LHC era. In this sense, the work presented here connects present-day measurements with the preparation of the next generation of heavy-ion studies at the LHC.

Part I

χ_c analysis

4

Theoretical motivation

Subatomic physics is the branch of physics that investigates the fundamental constituents of matter and the interactions governing their behavior at length scales smaller than the atom. While early atomic models treated atoms as indivisible units, experiments in the late nineteenth and twentieth centuries revealed that atoms consist of a compact nucleus surrounded by electrons, and that the nucleus itself is composed of protons and neutrons, collectively known as nucleons. These nucleons are further understood to be bound states of more elementary particles called quarks, whose dynamics are described within the framework of Quantum Chromodynamics (QCD), a component of the broader theoretical structure known as the Standard Model (SM) of particle physics. In this chapter, the theoretical framework that describes the interactions between the particles that compose the subatomic matter will be introduced. Special attention will be paid to mechanisms that govern the state of matter created in extreme conditions of temperature and energy density, the Quark Gluon Plasma (QGP). The study of the behavior of QGP is the main motivation of this thesis.

4.1 The Standard Model of Particle Physics

The Standard Model (SM) of particle physics is a theoretical framework based on Quantum Field Theory that provides a unified description of the fundamental particles and the interactions between them. In this model, the fundamental forces are transmitted by mediator particles associated with gauge fields, allowing the theory to systematically account for the electromagnetic, weak, and strong interactions, while gravity remains outside its scope. The development of a unified theory describing all fundamental interactions remains one of the major open challenges in particle physics. Since its development in

the seventies, the SM of particle physics has successfully produced numerous predictions that have been confirmed experimentally. One of the most significant and recent confirmations is the discovery of the Higgs boson, whose existence was predicted by François Englert, Robert Brout, and Peter Higgs [3, 4]. The particle was observed in 2012 by the ATLAS Collaboration and the CMS Collaboration at the Large Hadron Collider located at CERN [5,6]. This discovery provided the final experimental confirmation of the mechanism responsible for generating the masses of elementary particles within the Standard Model.

The SM characterizes particles through intrinsic properties such as mass and electric charge, as well as quantum numbers including spin, color charge, hypercharge, baryon number, and lepton number. Within this framework, fundamental particles are classified into two main categories according to their spin: fermions, which have half-integer spin and constitute the matter content of the theory, and bosons, which possess integer spin and act as mediators of the fundamental interactions. Fermions all have their corresponding anti-particles of the same mass but with an opposite charge, and they can be further divided into two subcategories:

The quarks are fermions that participate in the electromagnetic, weak, and strong interactions and appear in three generations: (u, d), (s, c), and (t, b). These correspond respectively to the up, down, charm, strange, top, and bottom quarks. Each quark carries a fractional electric charge, a baryon number, and a quantum number known as color charge, which is associated with the theory of QCD. Quarks cannot be observed as free particles due to the phenomenon of confinement and instead combine to form composite particles known as hadrons. These include baryons, which consist of three quarks (or three antiquarks), such as the proton and neutron, and mesons, which are bound states of a quark and an antiquark. Through these combinations, quarks constitute the fundamental building blocks of strongly interacting matter.

The leptons are the other family of fermions that, unlike quarks, do not participate in the strong interaction. They are organized into three generations: (e, ν_e), (μ, ν_μ) and (τ, ν_τ). These correspond respectively to the electron and the electron neutrino, the muon and the muon neutrino, and the tau lepton and the tau neutrino. The charged leptons (electron, muon, and tau) carry an electric charge of -1 and interact through the electromagnetic and weak forces. In contrast, neutrinos are electrically neutral and interact only via the weak interaction. Each lepton carries a conserved lepton number associated with its flavor. Although neutrinos were long assumed to be massless within the Standard Model, experimental observations of neutrino oscillations have demonstrated that they possess small but nonzero masses [7], indicating physics beyond the minimal formulation of the Standard Model.

On the other side of the spin classification, **bosons** are the force carriers that mediate the fundamental interactions between fermions. The photon (γ) is the mediator of the electromagnetic force, while the eight gluons (g) are responsible for the strong interaction,

binding quarks together within hadrons. The weak interaction is mediated by the charged W bosons (W^+, W^-) and the neutral Z boson (Z^0). The Higgs boson (H) is associated with the Higgs field and provides mass to fermions through Yukawa couplings, as well as to the W and Z bosons via spontaneous electroweak symmetry breaking [3,4,8]. Importantly, the photon and gluons remain massless, reflecting the unbroken gauge symmetries of electromagnetism and the strong force, while the masses of the W and Z bosons encode the breaking of the electroweak symmetry. This framework of gauge bosons and the Higgs mechanism allows the SM to consistently describe particle interactions at both low and high energies. A summary of the different particles that compose the SM can be seen in 4.1.

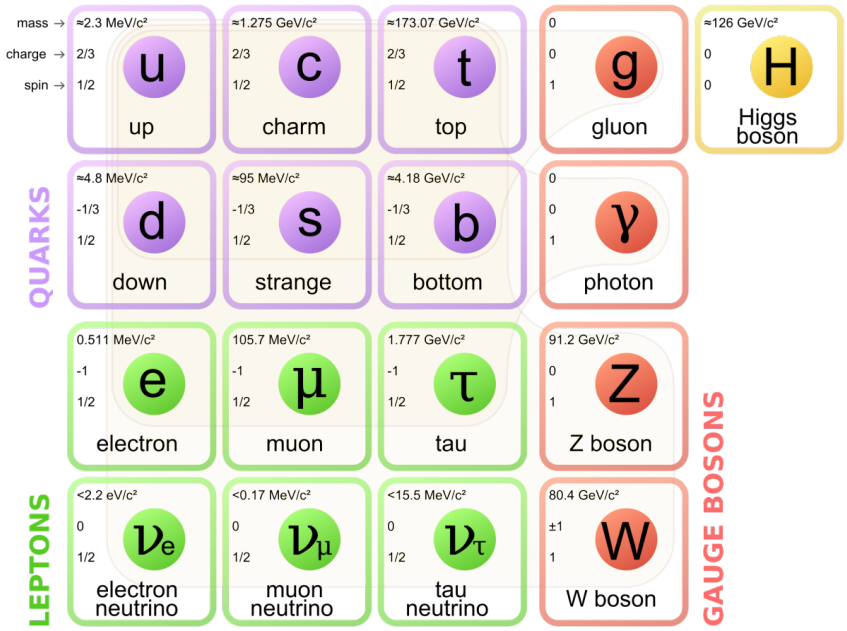


Figure 4.1: Summary of the different elementary particles of the Standard Model of particle physics. Figure taken from [9] under Creative Commons license.

From a mathematical perspective, the SM is formulated as a Quantum Field Theory based on the gauge symmetry group $SU(3)_c \times SU(2)_L \times U(1)_Y$. In this structure, $SU(3)_c$ corresponds to the strong interaction, governing the dynamics of quarks and gluons through the theory of QCD, while the electroweak interaction (unifying the weak and electromagnetic forces) is described by the $SU(2)_L \times U(1)_Y$ group. QCD accounts for the interactions among quarks and antiquarks via the exchange of eight massless gluons, incorporating the principles of color charge, confinement and asymptotic freedom. A more comprehensive explanation of this theory will be given below.

4.2 Quantum Chromodynamics

Quantum Chromodynamics (QCD) [10] constitutes the $SU(3)_c$ gauge sector of the SM and provides the theoretical framework describing the strong interaction. It governs the dynamics of quarks and gluons, which carry and conserve a charge called *color*. In QCD, quarks interact through the exchange of gluons, the gauge bosons associated with the non-Abelian $SU(3)_c$ symmetry group, where the subscript c denotes the color charge degree of freedom. Understanding the properties of the strong interaction is particularly important for the present work because it determines the internal dynamics and observable behavior of hadrons. These phenomena arise directly from the underlying QCD interactions among quarks and gluons. The dynamics of QCD are encoded in its Lagrangian density [11], which can be written as

$$\mathcal{L}_{\text{QCD}} = \sum_{i=u,d,s,c,b,t} \bar{q}_i(i\gamma^\mu D_\mu - m_i)q_i - \frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu}, \tag{4.1}$$

where $\bar{q}_i$ and q_i are the quark fields, D_μ is the covariant derivative:

$$D_\mu = \partial_\mu - ig_s T^a G^a_\mu, \tag{4.2}$$

with g_s being the strong coupling constant, T^a (for $a = 1, \ldots, 8$) the generators of $SU(3)_c$, and G^a_μ the eight gluon fields associated with the generators of the $SU(3)_c$ gauge group. The term $\frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu}$ describes the self-interactions of gluons, where $G^a_{\mu\nu}$ is the gluon field strength tensor:

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu, \tag{4.3}$$

and f^{abc} are the structure constants of the $SU(3)_c$ group [10,12]. Because quarks carry one of three color charges, combinations of quarks are classified according to the irreducible representations of the $SU(3)_c$ symmetry group. Observable hadrons must be color singlets, corresponding to color-neutral states, although intermediate colored configurations, such as color-octet states, can occur during parton interactions and hadronization.

A central feature of QCD is *asymptotic freedom*, whereby the interaction strength between quarks and gluons decreases as the energy scale increases. This behavior arises from the renormalization-group running of the strong coupling constant, $\alpha_s(Q^2)$, with the momentum-transfer scale Q^2. As a consequence, at very high energies, quarks and gluons behave as quasi-free particles. The strong coupling constant is a dimensionless parameter that characterizes the strength of the QCD interaction and is related to the gauge coupling g_s through:

$$\alpha_s(Q^2) = \frac{g_s^2(Q^2)}{4\pi}, \tag{4.4}$$

where Q represents the energy-momentum transfer scale, that is typically the momentum transfer in a scattering process. The energy dependence of α_s is driven by:

$$\alpha_s(Q^2) \sim \frac{1}{(11 - \frac{2}{3}n_f)\ln\left(\frac{Q^2}{\Lambda^2}\right)}, \tag{4.5}$$

where n_f is the number of quarks flavors and $\Lambda \sim$ 200-300 MeV is the QCD energy scale. Figure 4.2 shows a summary of the different determinations of the running coupling constant α_s as a function of the energy scale Q. The figure shows an excellent agreement between the experimental results and the theoretical calculations.

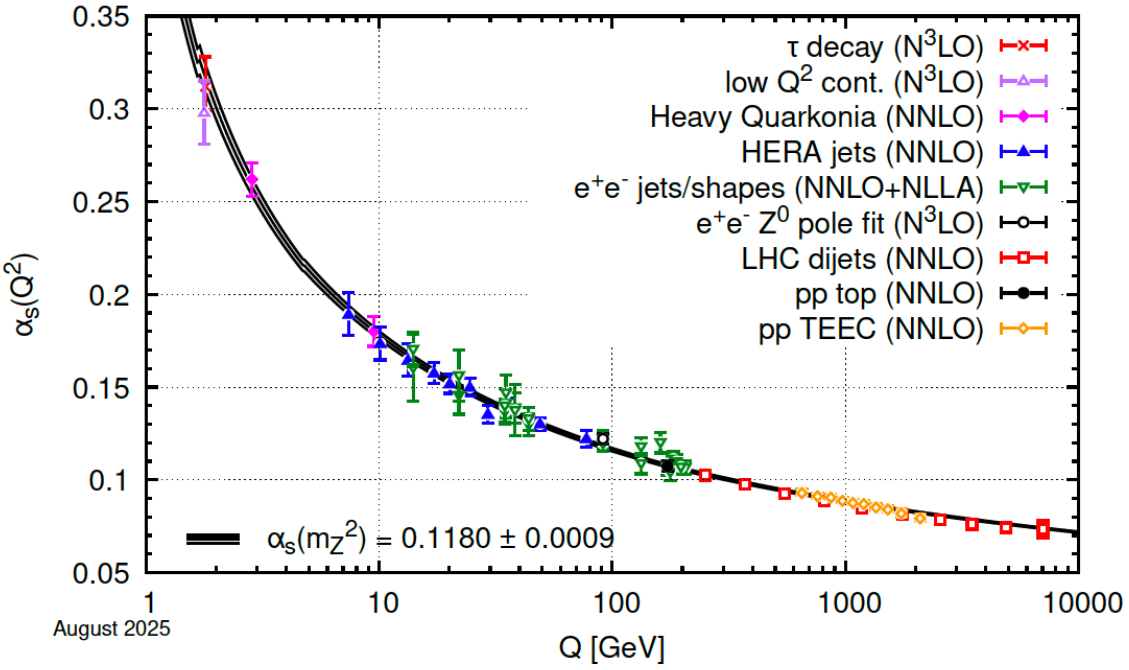


Figure 4.2: Summary of determinations α_s as a function of the energy scale Q compared to the running of the coupling computed at five loops taking as an input the current Particle Data Group (PDG) average, $\alpha_s(m_Z^2) = 0.1180 \pm 0.0009$. Figure taken from [13] under Creative Commons license.

The behavior of $\alpha_s(Q^2)$, under changes in the energy scale Q determines two of the defining features of Quantum Chromodynamics: asymptotic freedom and confinement. At large momentum transfer Q, the coupling α_s decreases logarithmically. As a consequence, the strong interaction becomes progressively weaker at short distances, and in the limit $Q \to \infty$, quarks and gluons behave as nearly free particles. This property is known as asymptotic freedom. Asymptotic freedom is a general feature of non-Abelian gauge theories, i.e., theories based on symmetry groups with non-commuting generators such as $SU(3)_c$. It originates from the self-interactions of the gauge bosons (gluons), which contribute negatively to the $\beta_0 = (11 - \frac{2}{3}n_f)$ function and drive the decrease of the coupling at high energies. Because the coupling becomes small in this regime, QCD observables at large Q^2 can be reliably computed using perturbation theory.

In contrast, at low energy scales (small Q), the coupling α_s grows large, and perturbation theory breaks down. This strong-coupling regime is associated with confinement, the phe-

nomenon by which quarks and gluons are permanently bound into color-neutral hadrons and cannot be observed as isolated particles.

4.3 The Quark Gluon Plasma

In order to study the behavior of the strongly interacting matter, one can examine the relations between the temperature T and the *baryon-chemical potential* μ_B. The baryon-chemical potential is defined as the energy required to increase the total baryon number by one unit and is proportional to the net baryon density. A schematic view of the QCD phase diagram can be seen in figure 4.3.

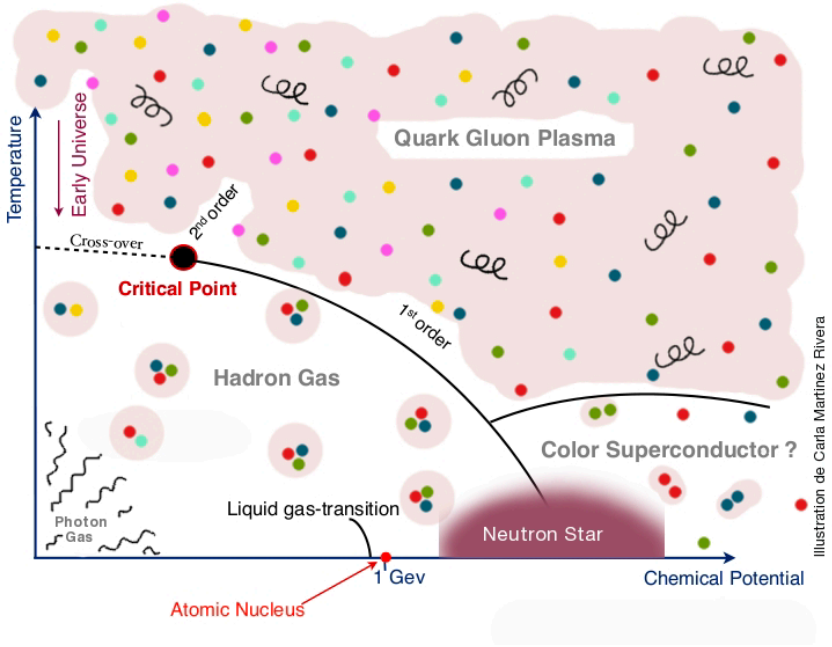


Figure 4.3: Layout of the hadronic matter phase diagram. Figure taken from [14] under Creative Commons license.

Ordinary nuclear matter is characterized by temperatures $T \approx 0$ MeV and a baryon chemical potential $\mu_B \approx 1$ GeV. As the baryon chemical potential increases, corresponding to extremely high densities and pressures on the order of 10^{32} Pa or higher, nuclear matter undergoes a transition into a regime that can be effectively described as a degenerate neutron gas. In this state, the Pauli exclusion principle dominates the behavior of the system, with neutrons occupying the lowest available energy states up to the Fermi surface. Such conditions are believed to be realized in the cores of neutron stars [15], where matter is compressed to densities exceeding that of atomic nuclei and exhibits strongly interacting, highly degenerate fermionic behavior. Further increasing the pressure to values on the order of 10^{35} Pa (corresponding to even larger baryon chemical potentials), the

internal Fermi pressure of the degenerate hadronic matter is eventually overcome by the extreme external compression. Under these conditions, hadrons are expected to dissolve into their constituent quarks and gluons, giving rise to a deconfined phase of matter. At sufficiently low temperatures, this high-density phase is predicted to exhibit color superconductivity, in which quarks form Cooper pairs through attractive strong interactions, analogous to the pairing of electrons in conventional superconductors. This results in a color-superconducting state characterized by the spontaneous breaking of color gauge symmetry [16].

At temperatures of the order $T \approx 100$ MeV, ordinary nuclear matter undergoes a transition into a hadron gas, where the relevant degrees of freedom are hadronic states such as mesons and baryons. As the temperature increases further, this hadronic phase experiences a transition to a deconfined state of matter known as the Quark–Gluon Plasma (QGP). In the QGP phase, the relevant degrees of freedom are the quarks and gluons, which are no longer confined within hadrons. According to the standard cosmological model, a hot and dense QGP is believed to have existed in the early Universe shortly after the Big Bang, during the first 10^{-6} seconds of its evolution [17], before cooling and hadronizing into the matter observed today.

As the baryon chemical potential increases, the QCD phase transition is expected to change in nature [18]. In particular, at sufficiently large μ_B, the transition from hadronic matter to the QGP is predicted to be of first order. The presence of a smooth cross-over transition at low μ_B, together with a first-order transition at high μ_B, implies the existence of a critical point in the QCD phase diagram [19], where the nature of the transition changes. According to lattice QCD calculations [20, 21], the properties of the phase transition depend sensitively on the number of quark flavors and their masses. In particular, the nature of the deconfinement and chiral transitions can be related to the quark mass values. In the limit of vanishing quark masses (the chiral limit), the transition is expected to be of second order for two quark flavors. By contrast, for finite quark masses, current results indicate that the transition at low μ_B is not a true phase transition but rather a smooth cross-over [22].

Exploring the QGP experimentally requires recreating the extreme conditions of temperature and energy density under which it exists. This is achieved through ultrarelativistic heavy-ion collisions, where the energy deposited in the interaction region is large enough to drive nuclear matter across the deconfinement transition.

4.4 Heavy-Ions collisions

In the context of heavy-ion collisions, such as those involving lead (Pb) nuclei, the large number of nucleon–nucleon interactions leads to a substantial deposition of energy in a localized interaction region, commonly referred to as the *fireball*. The extreme energy densities and temperatures achieved during the early stages of the collision can exceed

the critical thresholds required for the formation of QGP. As a result, heavy-ion collision experiments provide a unique environment to create and study deconfined QCD matter under controlled laboratory conditions, making them essential for exploring the properties of the QGP.

First proposed in the latter half of the 20th century, heavy-ion experiments have since been carried out in experiments conducted at facilities such as CERN and Brookhaven National Laboratory. Early fixed-target experiments at the Super Proton Synchrotron (SPS) and Alternating Gradient Synchrotron (AGS) operated at energies $\sqrt{s_{NN}} \approx 7.7 - 17.3$ GeV, where nucleons stop in the collision region, creating a high baryon chemical potential $\mu_B \approx 200 - 500$ MeV, known as the stopping regime. This is key for studying dense matter and the QCD critical point. In contrast, higher-energy collider experiments at the Relativistic Heavy Ion Collider (RHIC) and the Large Hadron Collider (LHC) reach up to $\sqrt{s_{NN}} \approx 5.4$ TeV, where nuclei pass through each other, resulting in low baryon density $\mu_B \approx 0$ at midrapidity and only partial energy loss, making this regime suitable for studying QGP under conditions similar to the early universe. This particular regime is commonly referred to as the Bjorken regime or the transparency regime [23]. The forward rapidity region covered by LHCb probes a complementary regime, where the net baryon density is expected to be larger than at midrapidity due to baryon stopping, allowing the study of QCD matter under different thermodynamic conditions.

In order to study these high-energy collisions, it is important to define the observable variables of interest. Here, the energy E and momentum $\vec{p}$ of a particle can be described by the four-momentum vector p^μ that in natural units ($c = \hbar = 1$), is given by:

$$p = (E, \vec{p}) = (E, p_x, p_y, p_z), \tag{4.6}$$

satisfying the relativistic energy-momentum relation $p^\mu p_\mu = E^2 - |\vec{p}|^2 = m^2$. In high-energy collisions, momenta are described relative to the beam axis (z-axis):

$$\vec{p} = (\vec{p_{\mathrm{T}}}, p_z), \tag{4.7}$$

where the transverse momentum is:

$$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}. \tag{4.8}$$

The longitudinal dynamics are characterized by the rapidity:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \tag{4.9}$$

which precisely measures the velocity of a particle along the beam axis. However, the experimental determination of rapidity presents significant challenges. First, the energy requires knowledge of the mass of the particle. Second, while detectors can precisely measure momentum components (p_x, p_y, p_z), they typically cannot directly determine particle

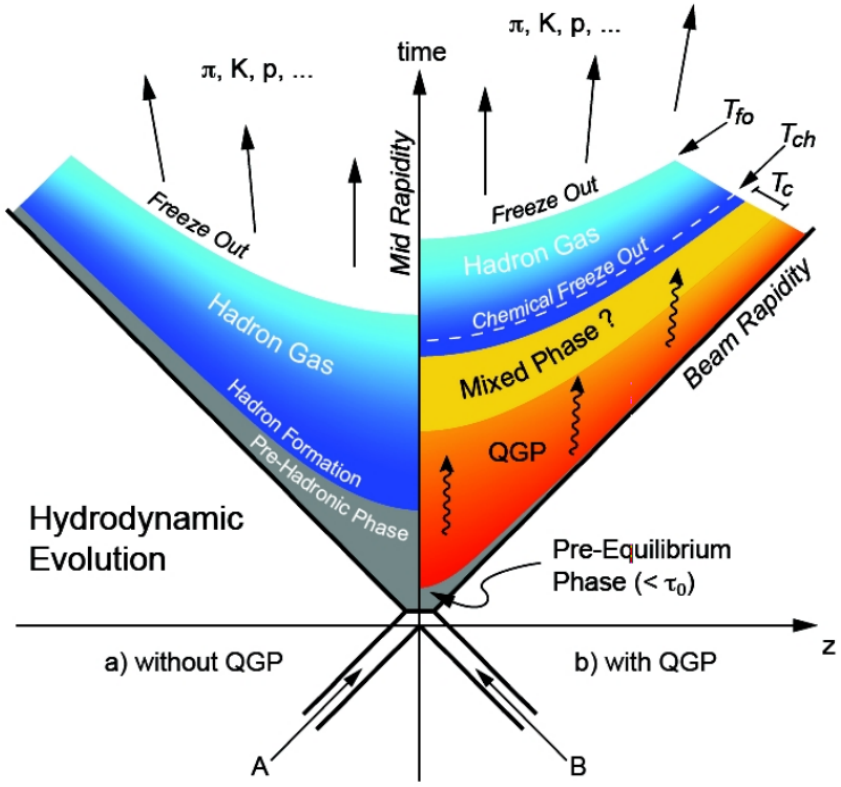


Figure 4.4: Diagram of the longitudinal heavy-ion collision evolution in the scenarios where no QGP is formed and the case where QGP is formed.

masses. For these reasons, experiments predominantly use the pseudorapidity:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right), \tag{4.10}$$

where θ is the polar angle relative to the beam axis. Pseudorapidity offers three principal advantages: it is directly measurable from track trajectories in the detector, it is independent of the particle mass, and for relativistic particles ($p \gg m$), it approximates the true rapidity ($\eta \sim y$). The relationship between rapidity and pseudorapidity becomes exact for massless particles.

4.4.1 Collision evolution

In collisions between two nuclei, a complex system of many particles is formed and undergoes a time-space evolution. James Bjorken proposed a simplified framework to describe this evolution within the context of the Landau hydrodynamic model [14, 23]. The central assumption of the model is boost invariance, where the fluid rapidity of each cell is identified with the space-time rapidity η. As a consequence, each fluid cell appears locally

at rest during the longitudinal expansion, as if the fireball were almost static in the rest frame, and the rapidity distribution of produced particles becomes flat at midrapidity [24]. This boost-invariant picture is further motivated by the transparency regime, in which the two colliding nuclei pass through each other, depositing energy in the central region. In this frame, the system's evolution can be outlined as follows:

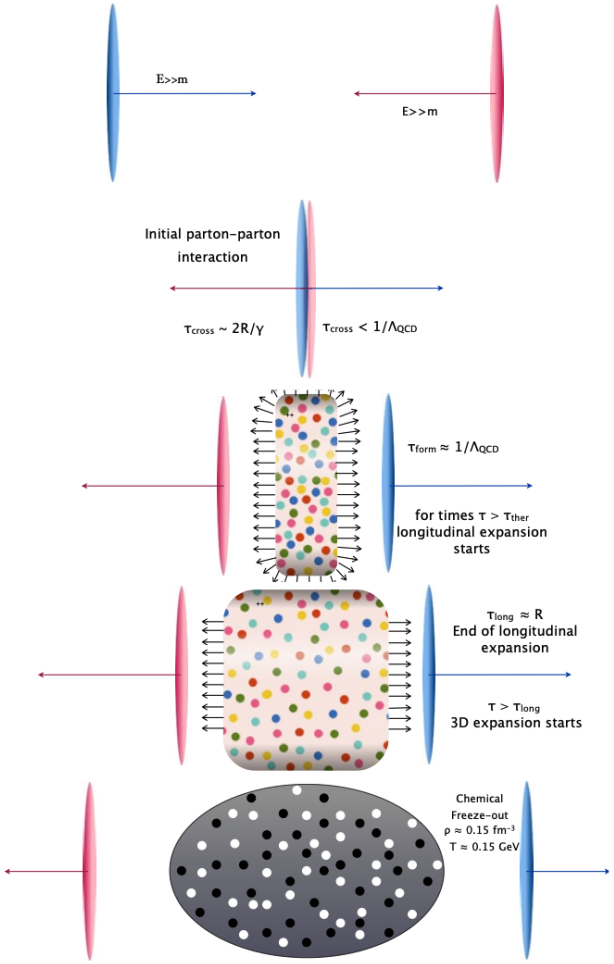


Figure 4.5: Scheme of the space-time evolution of a heavy-ion collision. Figure taken from [14] under Creative Commons license.

i. **Collision**: In this step, the colliding nuclei are contracted due to the Lorentz boost and typically they are referred to as *pancakes*. The collision takes place in a very short time computed as $\tau_{cross} = 2R_N/\gamma$, where R_N is the radius of the nucleus and γ the Lorentz factor. At LHC energies, this time lap is much smaller than the time scale of the strong interaction $\tau_{strong} = 1/\Lambda_{QCD} \approx 1$ fm/c. Therefore, particles generated by the strong interaction between partons in the colliding nucleons are produced only after the nuclei have already passed through each other.

ii. **Pre-equilibrium**: Here, the produced particles are not in thermal equilibrium, and the system formed is characterized by a high energy density and a low baryon density. In the mid-rapidity region, the net baryon density is close to zero, satisfying the condition of the Bjorken regime. On the other hand, the energy density reached at this point exceeds the pseudo-critical energy density necessary for the formation of QGP.

iii. **QGP formation**: The system of interacting particles reaches thermal equilibrium at a time τ_{eq}, as the energy density exceeds the pseudo-critical energy density. Under these conditions, the QGP is formed. The QGP then evolves according to the laws of relativistic hydrodynamics [25], expanding due to pressure gradients while simultaneously cooling. Hydrodynamic models indicate that thermal equilibrium is achieved at a proper time in the range $0.6 < \tau_{eq} < 1$ fm/c. At the LHC energies, the QGP phase is expected to last ~ 10 fm/c.

iv. **Hadron gas**: As the system expands, its temperature gradually decreases until it falls below the pseudo-critical value. At this stage, colour charges are no longer deconfined, and the QGP undergoes a transition into a hadronic system, in which partons become confined within hadrons. The newly formed hadrons then constitute an expanding hadron gas. In the case of a first-order phase transition, the system passes through a mixed phase in which confined hadronic matter coexists with the QGP.

v. **Freeze-out**: The hadron gas continues to expand and cool. At a certain stage, inelastic scatterings cease due to the decreasing particle density, and the relative abundances of the different hadronic species become fixed. This stage is referred to as chemical freeze-out. When the particle density decreases to the point where even elastic collisions cease, the momentum spectra of the hadrons become fixed. This stage is referred to as kinetic freeze-out (or thermal freeze-out). The hadrons then continue to expand freely until they reach the detector.

A schematic view of the evolution of the collisions can be seen in figures 4.4 and 4.5. The representation is done in the longitudinal direction and based on the Bjorken model.

4.4.2 Geometry of the collision

In heavy-ion collisions, the initial geometry of the interaction region plays a central role in determining the subsequent evolution described in section 4.4.1. The model used to describe the initial stages of a heavy-ion collision is the Glauber Model [26]. Initially developed at the end of the fifties, this model describes nucleus–nucleus collisions as a superposition of independent binary nucleon–nucleon interactions. The Glauber model is formulated in the so-called optical limit, where nuclei are treated as collections of point-like and non-interacting nucleons. In its original formulation, the model does not differentiate between nucleon species within the nucleus; protons and neutrons are therefore treated as

indistinguishable. Additionally, the optical limit assumes that the nuclei have sufficiently large momenta so that their trajectories are not significantly deflected during the collision, allowing the constituent nucleons to propagate along straight-line paths.

The Glauber model allows the calculation of several quantities that characterize the collision process. One of these is the number of participant nucleons, denoted N_{part}, which corresponds to the nucleons that take part in the interaction by experiencing at least one collision with a nucleon from the opposing nucleus. Another key quantity is the total number of binary nucleon–nucleon interactions, represented as $N_{coll}(b)$, where b is the impact parameter. The impact parameter is defined as the transverse distance between the centers of the two colliding nuclei. It plays a central role in determining the collision centrality and strongly affects how many binary nucleon–nucleon collisions occur. Based on this parameter, nucleus–nucleus collisions are generally categorized as either central or peripheral. Central collisions are characterized by having a very low impact parameter $b \sim 0$, which is interpreted as an almost complete overlap between the two nuclei in the transverse plane. In the other hand, peripheral collisions are characterized by having a high impact parameter $b \sim R_A + R_B$, resulting in a small superposition in the edge of the radius of the colliding A nucleus and B nucleus. The centrality of a collision is usually expressed as a percentage of the total inelastic cross section, where the most central collisions have the highest percentage. Figure 4.6 shows a schematic representation of two colliding nuclei in the Glauber model description.

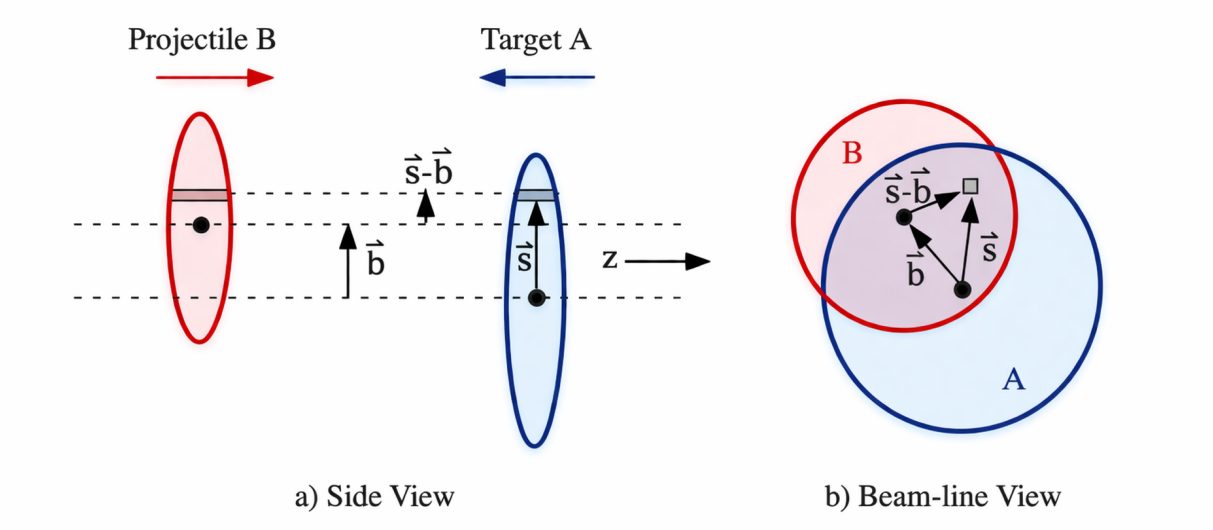


Figure 4.6: Schematic side view and beam-line view of a nucleus–nucleus collision according to the Glauber model description figure taken from [26] under Creative Commons license.

One direct example of the importance of the collision geometry, specially for this work, is *charged particle multiplicity* N_{ch}. It is defined as the number of charged particles produced

in a collision. The charged particle multiplicity acts as a global observable that reflects the energy density in the pre-equilibrium stage of the collision and provides information on the system's evolution toward the QGP phase. To enable meaningful comparisons across different collision systems while accounting for their varying sizes, the multiplicity is commonly normalized by the number of participant pairs $\langle N_{part} \rangle/2$, which is computed using the Glauber model. Experimentally, the multiplicity can be related to the occupancy variables in the different subdetectors of a high-energy physics experiment [27].

4.5 Quarkonium production in heavy-ion collisions

Quarkonia is the name given to the bound states of a heavy quark and its own antiquark ($q\bar{q}$), most notably charmonium ($c\bar{c}$) and bottomonium ($b\bar{b}$). Their large quark mass places them in a regime where perturbative QCD is applicable, making them theoretically controllable probes of the strong interaction. Furthermore, their production time scale, of order $\tau \sim 1/(2m_q)$, is short enough that they are formed in the very early stages of the collision, well before the QGP thermalizes. As a result, quarkonia are present throughout the entire space-time evolution of the hot and dense medium, and carry information about its properties at every stage, from the initial hard scattering to the final hadronization. An scheme of the broad charmonium spectrum is presented in figure 4.7.

In this work, the stability of quarkonium states in a thermal medium constitutes the primary subject of investigation, with a particular focus on charmonium states. The theoretically well-defined problem of determining the stability of quarkonium states at finite temperature has historically been closely associated with experimental studies of quarkonium yields in relativistic heavy-ion collisions. Following the influential work of Matsui and Satz [1], which proposed quarkonium suppression as a signature of QGP formation, and later developments such as the sequential suppression scenario introduced in [2], these ideas have become strongly intertwined. As a consequence, the question of quarkonium binding in a thermal medium under equilibrium conditions has often been conflated with the far more complex issue of quarkonium production within the highly non-equilibrium environment created in heavy-ion collisions. While acknowledging the significant role that quarkonium observables play in the study of heavy-ion collisions, it is important to make a clear distinction between quarkonium melting, which concerns in-medium binding properties, and quarkonium suppression, which refers to experimentally observed yield modifications. In order to understand quarkonia production in heavy-ion collisions in the presence of a hot and dense medium, first we need to understand quarkonia production in the absence of such medium. For this, it is necessary to establish a solid understanding of quarkonium production in *pp* collisions, which provides the baseline in the QCD vacuum free from any nuclear or medium effects. In addition, proton-nucleus (pA) collisions serve as an essential intermediate step, since they allow the so-called *cold nuclear matter effects* to be isolated from the genuine hot-medium effects. A detailed explanation of these effects is given below. Only by combining these three systems can

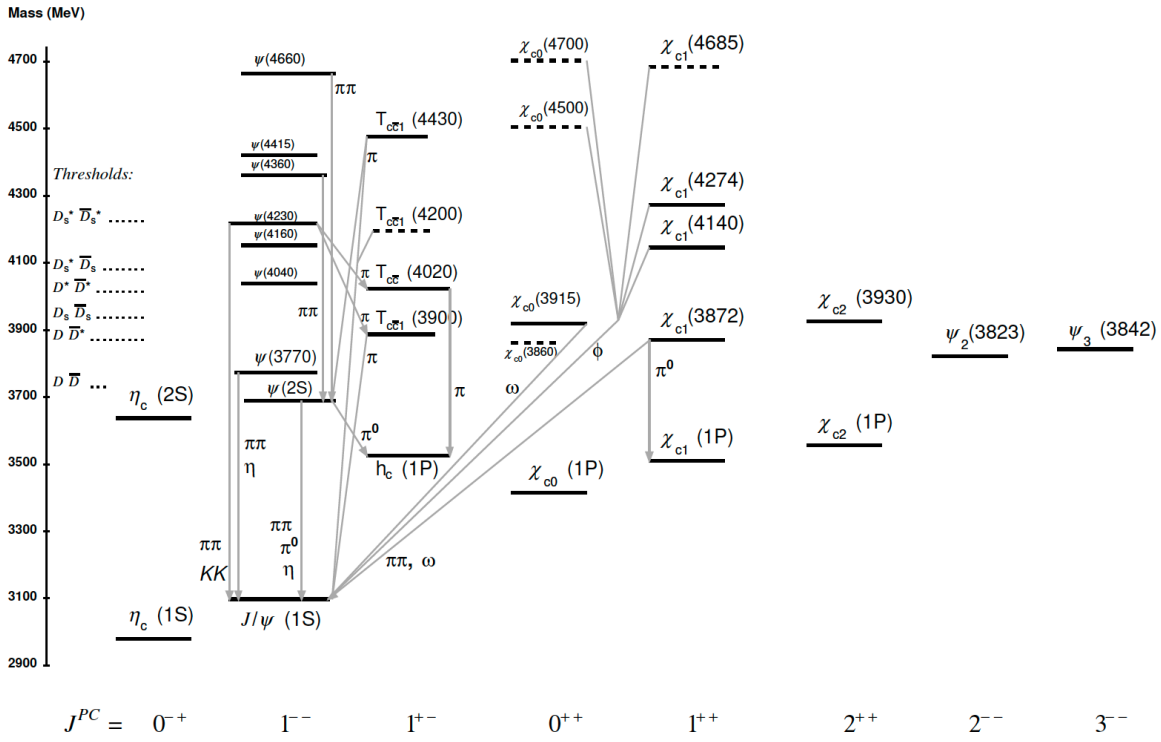


Figure 4.7: The level scheme of meson states containing a minimal quark content of $c\bar{c}$ as a function of their mass and their quantum numbers. Figure taken from [28] under Creative Commons license.

one reliably attribute modifications of quarkonium production in heavy-ion collisions to the formation of a QGP.

4.5.1 Quarkonium production in *pp* collisions

As an initial step, it is necessary to examine quarkonium production in the absence of a hot medium. In this context, *pp* collisions establish the baseline under vacuum QCD conditions free of any nuclear or medium effects. In addition, quarkonium production in *pp* collisions is an interesting subject in its own right, offering valuable insight into the internal structure of nucleons and the interplay between perturbative and nonperturbative QCD processes governing hadron formation. Moreover, a solid understanding of this baseline process is essential for interpreting quarkonium production in more involved systems such as proton–nucleus (pA) and nucleus–nucleus (AA) collisions.

Charmonium refers to any bound state composed of a charm quark and an anti-charm quark ($c\bar{c}$). Owing to the relatively large charm quark mass ($m_c \approx 1.3$ GeV/c^2), the

velocity of the quarks inside the bound state remains sufficiently small for the system to be approximately described using non-relativistic quantum mechanics. In this framework, the interaction between the quark and antiquark is commonly modeled using the Cornell potential,

$$V(r) = \sigma r - \frac{\alpha}{r}, \tag{4.11}$$

where σ represents the string tension associated with confinement and the second term corresponds to a Coulomb-like interaction at short distance. Solving the Schrödinger equation with such a potential leads to the spectrum of charmonium states shown in Fig. 4.7.

Among these states, the J/ψ meson corresponds to the ground state ($1S$) with a mass of approximately 3.1 GeV/c^2. Several excited states, such as $\psi(2S)$ and χ_c, can decay into a J/ψ meson and therefore contribute significantly to the inclusive J/ψ yield through feed-down processes. In addition, a non-negligible fraction of observed J/ψ mesons originates from the decay of beauty hadrons. Experimentally, it is therefore important to distinguish between prompt J/ψ mesons, produced either directly or through feed-down from excited charmonium states, and non-prompt J/ψ mesons originating from b-hadron decays. A proper interpretation of the measured J/ψ yields therefore requires accounting for these various sources.

A particularly important feature of quarkonium production is the very short formation time of the heavy quark pair, typically $\tau_f \sim 1/(2m_q) \approx 0.1$ fm/c in the case of charmonium, which is much shorter than the characteristic formation time of the QGP. Consequently, quarkonia are produced during the earliest stages of the collision and are therefore sensitive probes of the surrounding medium in heavy-ion collisions.

In general, the production of quarkonium states represents only a small fraction of the heavy quark-antiquark pairs created in high-energy collisions. Most heavy quarks hadronize into open heavy-flavor mesons or baryons. The possible hadronization channels are illustrated schematically in figure 4.8. The production of quarkonia is thus a relatively rare process involving both perturbative and nonperturbative QCD dynamics. The large heavy-quark mass provides a hard scale, $Q \gtrsim 2m_Q \gg \Lambda_{\mathrm{QCD}}$, allowing the initial production of the heavy quark pair to be described using perturbative QCD (pQCD). In contrast, the subsequent evolution of the pair into a color-neutral bound state remains intrinsically nonperturbative.

The first stage of the production mechanism consists in the creation of a heavy quark-antiquark pair through hard partonic scatterings involving quarks and gluons from the incoming protons. At leading order, the dominant processes are quark-antiquark annihilation and gluon fusion, as illustrated in Fig. 4.9. Additional contributions such as gluon splitting and flavor excitation also play an important role at high energies. Since the colliding partons are constituents of composite protons, their momentum distributions cannot be calculated using perturbative QCD. Instead, they are described by non-perturbative

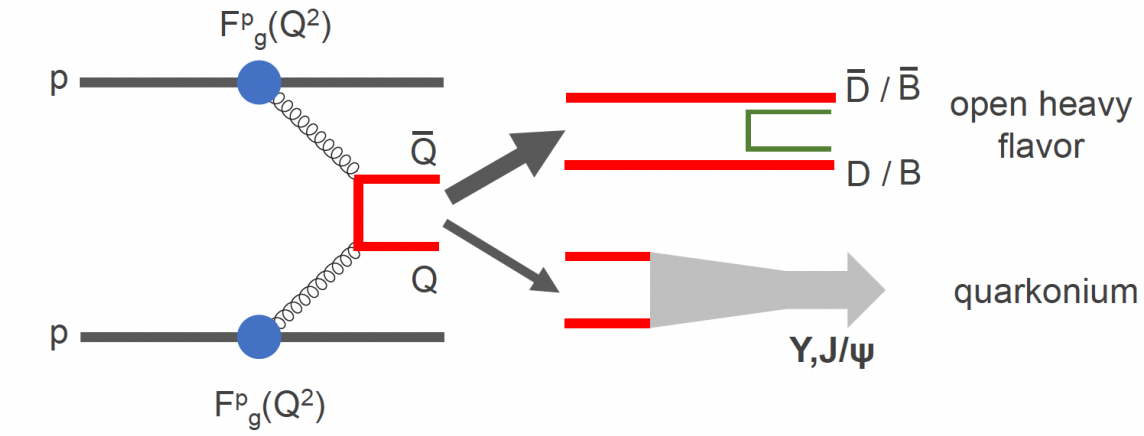


Figure 4.8: The two stages of the production process of prompt quarkonium: production of the color-bearing heavy quark-antiquark pair from hard partons and the subsequent formation of either open heavy flavor or quarkonium states. Figure taken from [29] under Creative Commons license.

parton distribution functions (PDFs), which encode the probability of finding a parton carrying a given fraction of the proton momentum.

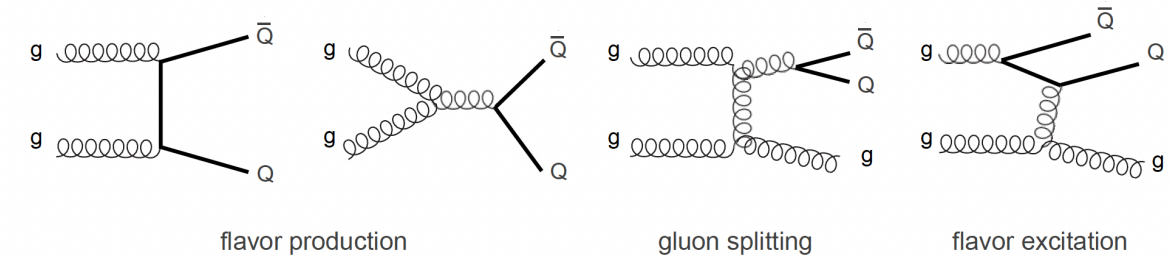


Figure 4.9: Scattering processes in the production of heavy quark anti-quark pairs at proton-proton colliders. Figure taken from [29] under Creative Commons license.

Following its production, the heavy quark pair can be created either in a color-singlet or color-octet state. Since physical quarkonia are color neutral, color-octet states must neutralize their color charge through soft gluon emissions before forming a bound state. The precise mechanism governing this hadronization process remains an open theoretical question and constitutes one of the major challenges in quarkonium physics.

Several theoretical approaches have been developed to describe quarkonium production. The Color Evaporation Model (CEM) [30, 31] assumes that the heavy quark pair neutralizes its color through interactions with the color field generated during the collision, and that the probability for the pair to hadronize into a given quarkonium state is universal

and phenomenological. In contrast, the Color-Singlet Model (CSM) [32,33] assumes that the heavy quark pair is produced directly with the same quantum numbers as the final quarkonium state. Finally, the Non-Relativistic QCD (NRQCD) [34–37] effective theory extends this description by including color-octet contributions and nonperturbative transitions between intermediate states through Long Distance Matrix Elements (LDMEs). While NRQCD successfully describes a wide range of quarkonium production cross-section measurements, particularly at LHC energies, significant tensions remain. In particular, the transverse polarization predicted at high-p_T is not observed in measurements by CMS and LHCb collaborations, which instead find little or no polarization over a broad kinematic range [38, 39]. In addition, the description of quarkonium production at low-p_T, where non-perturbative effects become more important, remains challenging [40].

4.5.2 Cold nuclear matter effects

The next step toward understanding quarkonium production in a hot and dense medium is to study proton-nucleus (pA) collisions, which represent an intermediate system between *pp* and nucleus-nucleus collisions. These systems allow for the isolation and characterization of the so-called *cold nuclear matter* effects, that refers to nuclear modifications of quarkonium production that are present even in the absence of a thermalized deconfined medium. Disentangling these cold nuclear matter effects from genuine hot-medium signatures is essential when study heavy-ion collisions, especially given that quarkonia are produced early in the collision, on a timescale of order $\tau \sim 1/(2m_q)$, and therefore also experience the pre-equilibrium phase of the system's evolution, before the QGP has fully thermalized. Any modification observed in heavy-ion collisions must thus be corrected for these cold nuclear matter baseline effects, which pA collisions are uniquely suited to constrain. Detailed discussions of these effects can be found in [41, 42].

The assumption of a nucleus as a mere incoherent sum of free proton and neutron Parton Distribution Functions (PDFs) fails to describe experimental data accurately. Deep Inelastic Scattering (DIS) experiments reveal that nuclear effects must be considered. These modifications are encoded in nuclear PDFs (nPDFs), which adjust the free-proton PDF with a nuclear modification factor

$$f_i^{p/A}(x, Q^2) = R_i^A(x, Q^2) \cdot f_i^p(x, Q^2), \tag{4.12}$$

where $f_i^p(x, Q^2)$ is the free-proton PDF for parton species i, and $R_i^A(x, Q^2)$ is the nuclear modification factor encapsulating the nuclear effects. The nuclear modification factor, $R_i^A(x, Q^2)$, varies across different regions of Bjorken-x and encapsulates distinct physical phenomena.

From an experimental view, differences between quarkonium yields measured in *pp* collisions and those observed in collisions involving nuclei, either pA or AA, retrieve the information about the nuclear effects. Therefore, in practice the nuclear modification factor is computed as

$$R_{xA}(p_T) = \frac{1}{\langle N_{coll} \rangle} \frac{dN_{xA}/dp_T}{dN_{pp}/dp_T}, \tag{4.13}$$

where the prefactor, which contains the number of binary collisions, serves merely as a normalization factor. Here, the subscript x can refer to both protons (p) or nucleus (A). If a collision involving a nucleus were nothing more than a superposition of many independent pp interactions, this ratio would remain equal to unity.

A useful classification distinguishes between initial-state and final-state effects. Initial-state effects arise from modifications to the partonic structure of nuclei relative to that of isolated protons. Final-state effects, in contrast, arise because quarkonium states produced in pA or AA collisions propagate through a nuclear environment and may interact with surrounding nuclear matter or with the fragments of the collision, processes that can destabilize the heavy quark–antiquark bound state.

Initial state effects

The determination of nPDFs remains an active area of research. Current knowledge relies on a close interplay between theoretical modeling and experimental measurements, through which constraints on the momentum fractions carried by quarks and gluons in nucleons bound within nuclei can be extracted from observed production yields.

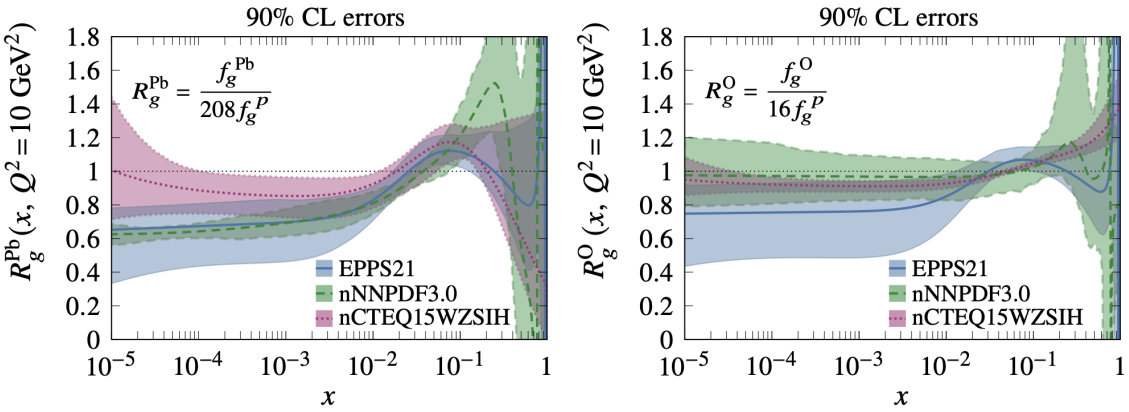


Figure 4.10: Nuclear modifications of the gluon PDFs in lead (left) and oxygen (right) nuclei. Results from the EPPS21, nNNPDF3.0 and nCTEQ15WZSIH analyses. Figure taken from [43] under Creative Commons license.

As illustrated in figure 4.10, several characteristic features appear in the ratio of nuclear to free gluon distribution function, R_g^A [44,45]. At small values of parton momentum fractions, $x < 10^{-2}$, this ratio falls below unity, a phenomenon known as small-x *shadowing*. At intermediate values, around $x \sim 0.1$, the ratio exhibits an enhancement often referred

to as *anti-shadowing*. Then, at larger momentum fractions, $x \gtrsim 0.6$, the ratio again decreases, forming a suppression first observed by the EMC collaboration and therefore known as the EMC effect [46]. Finally, at high values of $x \gtrsim 0.8$ the so-called *Fermi motion* effect manifests itself as an enhancement of the ratio. Unlike the other modifications discussed above, this effect is a purely kinematic consequence of the fact that nucleons inside the nucleus are not at rest but carry intrinsic momentum due to their confinement within the nuclear potential. This Fermi motion effectively allows nucleons to transfer momenta exceeding the kinematic limit of a free nucleon at rest, so that when the parton momentum distributions are convoluted with the nucleon momentum distribution inside the nucleus, the resulting nuclear structure function exhibits an enhancement over the free-nucleon baseline. A variety of nuclear PDF parametrizations are currently available, which differ both in their theoretical assumptions and in the strategies used to fit experimental data. Some examples include EPPS21, nNNPDF3.0 and nCTEQ15WZSIH as can be seen in figure 4.10.

The question at this point is how the modification of the nuclear PDF's affects the production of the individual quarkonium states, or equivalently, how the nuclear modification factor R_{pA} differs from unity, where values lower than unity implies a suppression and values greater than the unity implies an enhancement. Depending on the regime of x, shadowing or anti-shadowing, the gluon nuclear parton distribution functions (nPDFs) affects the $q\bar{q}$ production and a suppression or enhancement of the quarkonia production is expected respectively [47]. This provides one of the first instances in which the concept of quarkonium suppression appears. In this context, however, the suppression is unrelated to the formation of a hot medium and instead arises from modifications in the structure of the colliding nuclei. Such baseline effects must therefore be taken into account when analyzing quarkonium production in AA collisions, in order to distinguish genuine signatures of a hot QCD medium from those originating from cold nuclear matter.

Furthermore, more theoretical approaches to predict the initial conditions of the collision have also been proposed. A popular one worth mentioning is the Color Glass Condensate (CGC) framework [48]. The CGC framework is widely used to describe the initial color fields of colliding hadrons or nuclei at high energy and low Bjorken-x, where gluon densities dominate. In this regime, gluon radiation toward smaller x leads to a rapid growth of the gluon population. At sufficiently high density, nonlinear effects become important, and gluons begin to recombine, balancing further gluon splitting and leading to gluon saturation. As the effect of saturation depends on the number of nucleons A, it is expected to play an important role in the description of quarkonium production in heavy-ion collisions.

Final state effects

A variety of phenomenological models have been developed to describe final-state interactions affecting quarkonium production. At sufficiently low collision energies, where

quarkonium formation may occur within the spatial extent of the projectile nucleus, the bound state can interact with remnants of the nuclear environment, potentially leading to its dissociation. This mechanism is referred to as nuclear absorption or nuclear break-up [49]. However, this mechanism is expected to be negligible for quarkonia, as the crossing time is small compared to the formation time of the $c\bar{c}$ state.

For LHC energies, alternative mechanisms have been proposed. One of them is the Coherent Energy Loss (CEL) model [50, 51], which was introduced to explain charmonium suppression at high energy. In this framework, partons undergo medium-induced coherent gluon radiation while crossing cold nuclear matter, leading to an energy loss that can account for the observed suppression patterns without requiring modifications of the nPDFs. Both the initial-state modifications and coherent energy loss are expected to influence quarkonium ground states and excited states in a similar way, making them difficult to disentangle or combine. Experimentally, however, it has been observed that excited states of bottomonium and charmonium in pA collisions at the LHC energy exhibit a stronger suppression than their respective ground states.

To account for the remaining suppression effects at the LHC, an effective model called the comover interaction model [52] has been proposed and successfully reproduces this behavior. In this effective framework, the quarkonium state interacts with particles produced in the collision that share a similar rapidity and therefore propagate in its vicinity, called the co-moving medium. The interaction strength is parametrized through an effective quarkonium–comover cross section, while cold nuclear matter effects are incorporated via nPDFs. As we will see later, this model is also used in the context of heavy-ion collisions and QGP formation.

4.5.3 Quarkonium in AA collisions

After the review of the features of quarkonium production in pp and pA collisions, we will now discuss the quarkonium production in AA collisions, where QGP is formed. Specifically, we are going to focus on charmonium production as it is the object of study of this work.

The study of charmonium in heavy-ion collisions provides a clear example of how a complete understanding of all stages of the collision is required to describe quarkonium production. Early efforts to interpret charmonium behavior in such environments focused on the concept of quarkonium melting. As it has been mentioned before, Matsui and Satz examined the stability of charmonium states in a thermal medium and argued that the formation of a deconfined QGP would prevent the existence of bound states in the medium [1]. Their argument was originally formulated within a static picture of quarkonium dissociation. In general, the efficiency of quarkonium melting increases with the energy density of the medium. Moreover, states that are more weakly bound in vacuum are expected to dissolve more readily in the presence of the QGP. Building on these con-

siderations, Matsui further proposed in Ref. [1] that the disappearance of bound states in the medium should manifest experimentally as a reduction in charmonium yields, named quarkonium suppression and represented in figure 4.11, following a pattern similar to that of in-medium melting. This idea was later developed into the sequential suppression scenario, discussed in detail in Ref. [2]. According to this picture, the suppression should be strongest in central collisions and at small rapidities, where the density of medium constituents is highest.

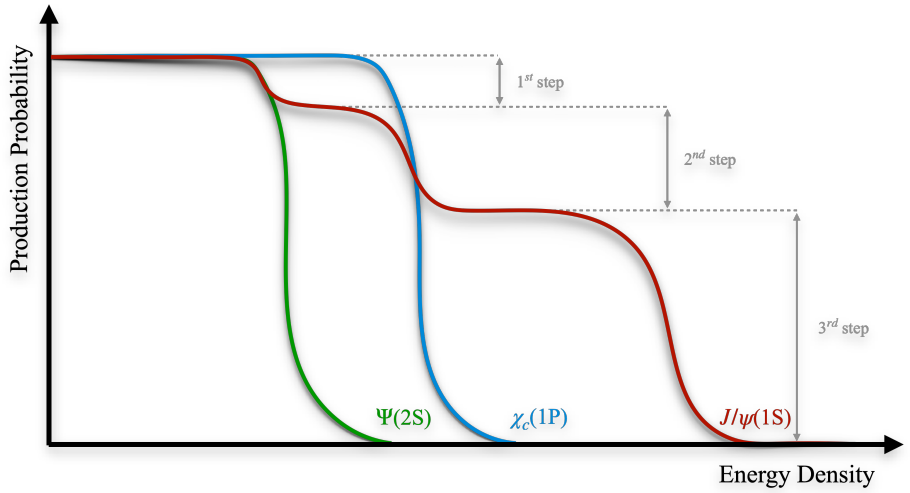


Figure 4.11: Scheme of the survival probability of J/ψ as a function of the energy density.

While the hierarchical structure of quarkonium melting remains a valid concept, it is now understood that suppression does not necessarily follow directly from melting, particularly at high collision energies. The reason lies in processes occurring both before and after the QGP phase originally considered by Matsui and Satz. As the collision energy increases, the production cross section for $c\bar{c}$ pairs grows rapidly. Consequently, a hotter plasma is typically accompanied by a significantly larger abundance of $c\bar{c}$ pairs. Although initially formed bound states may dissolve due to interactions with the medium, these abundant charm quarks can later recombine, either during the QGP phase itself or at hadronization, forming new charmonium states. This regeneration mechanism partially restores the charmonium yield.

The possibility of such recombination had already been mentioned by Matsui in early work, but at the time it was considered negligible because the experimentally accessible charm production cross section was relatively small. With the development of higher-energy facilities such as LHC, however, regeneration was revisited in greater detail and has

since been recognized as an essential component in understanding charmonium production in heavy-ion collisions. It is worth noting that in regimes where regeneration plays a minor role, such as at large transverse momentum, the nuclear modification factor R_{AA} of different quarkonium states still exhibits a suppression pattern that correlates with their respective in-medium binding energies [53].

4.5.3.1 Overview of theoretical models

Several theoretical frameworks have been developed to interpret quarkonium production in heavy-ion collisions [54]. Among these, the statistical hadronisation model, transport models, the collisional dissociation model, and the comover interaction model are commonly compared with experimental measurements of charmonium production. Their main features can be summarized as follows.

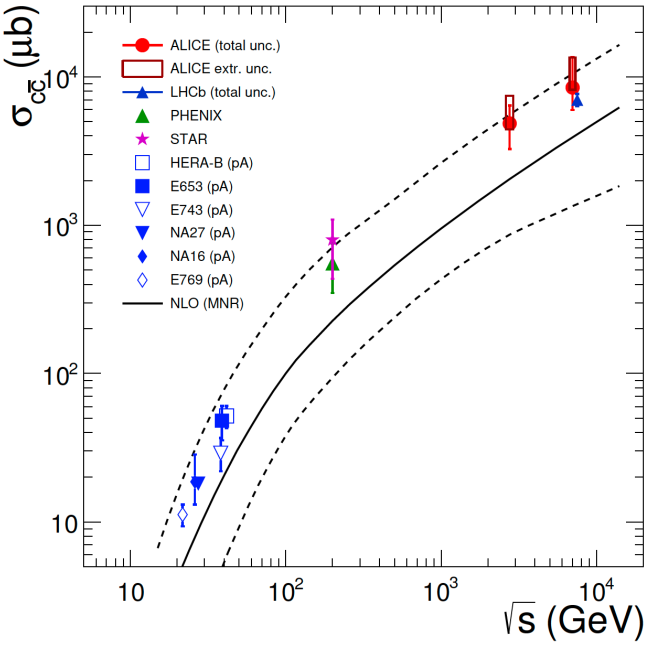


Figure 4.12: Total (extrapolated) $c\bar{c}$ cross section as a function of $\sqrt{s}$. Data in pA or deuteron–nucleus (dA) collisions were scaled down assuming no nuclear effect. Figure taken from [54] under Creative Commons license.

In the **Statistical Hadronisation Model** [55], charm quarks and antiquarks are produced in the initial hard scatterings and subsequently thermalize within the quark–gluon plasma. Hadrons are then formed at the stage of chemical freeze-out. A central assumption of this approach is that no quarkonium states survive within the deconfined phase, implying complete suppression before hadronization. Consequently, cold nuclear matter effects are not explicitly included in the model. An important element of this framework is

the canonical suppression of open charm or beauty hadrons, which governs the centrality dependence of the resulting yields. The overall normalization is determined by the input heavy-quark production cross section.

In **Transport Models** [56], an alternative production mechanism arises from kinetic *recombination* of heavy quarks and antiquarks within the QGP. Here, quarkonium states undergo continuous dissociation and regeneration throughout the entire lifetime of the deconfined medium. These models typically incorporate a hydrodynamic-like description of the expanding fireball, constrained by experimental data, together with the implementation of color screening effects informed by lattice QCD calculations. Additional key ingredients include parton-level interaction cross sections. Cold nuclear matter effects are accounted for through an effective absorption cross section that encapsulates contributions from shadowing (or antishadowing), nuclear absorption and Cronin effects.

The **collisional dissociation model** [57] considers several mechanisms affecting quarkonium production. Besides modifications of the quarkonium binding potential in the QGP and cold nuclear matter effects, it includes the radiative energy loss of the color-octet quarkonium precursor as well as collisional dissociation processes occurring within the deconfined medium.

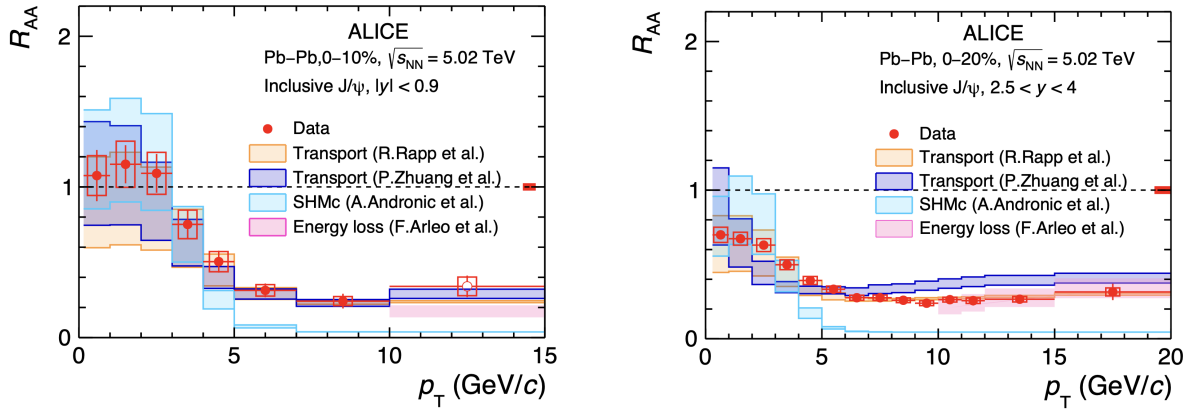
Finally, the **comover interaction model** [58] describes the dissociation of quarkonium states through interactions with the surrounding medium composed of co-moving hadronic and partonic matter. In this framework, regeneration processes are also incorporated, with their magnitude determined by the production cross section of $c\bar{c}$ pairs.

Within transport and comover interaction models, calculations indicate that at LHC energies a substantial portion of the observed J/ψ yield, exceeding about 50% in the most central collisions, originates from the recombination of charm quarks [54]. In contrast, the statistical hadronisation model assumes that all J/ψ mesons are produced at the moment of hadronization through purely statistical processes. To incorporate recombination effects, the transport and comover approaches typically employ a charm production cross section of $d\sigma^{pp}_{c\bar{c}}/dy \approx 0.6$ mb at midrapidity for $\sqrt{s_{NN}} = 2.76$ TeV. This corresponds to a total $c\bar{c}$ production cross section in pp collisions of roughly 5 mb, which is consistent with existing experimental measurements shown in figure 4.12. The cross section adopted in transport and comover models is approximately twice as large as the value used in the statistical hadronisation model. This difference does not imply an inconsistency between the approaches. In the statistical hadronisation framework, initial-state shadowing effects are not explicitly modeled. Instead, a smaller value of the charm production cross section, $d\sigma^{pp}_{c\bar{c}}/dy \approx 0.3 - 0.4$ mb, is chosen to effectively account for shadowing, which can reduce the charm production rate in PbPb collisions by up to a factor of two.

4.5.3.2 Experimental overview of charmonium production in AA collisions

In the charmonium sector, the state that can be measured with the highest precision in nucleus–nucleus collisions is the lowest-lying vector state J/ψ [59].

Results from LHC Run 2 provide strong support for the regeneration picture through several experimental observations. These include a larger nuclear modification factor at midrapidity compared with forward rapidity when integrated over transverse momentum, reflecting the higher charm quark density at midrapidity. An increase of the nuclear modification factor toward low transverse momenta, consistent with a thermal origin of the produced quarkonium states. The observation of a significant anisotropic flow of J/ψ, as well as characteristic centrality and collision-energy dependences. In figure 4.13 the Run 2 measurements of the nuclear modification factor [60].



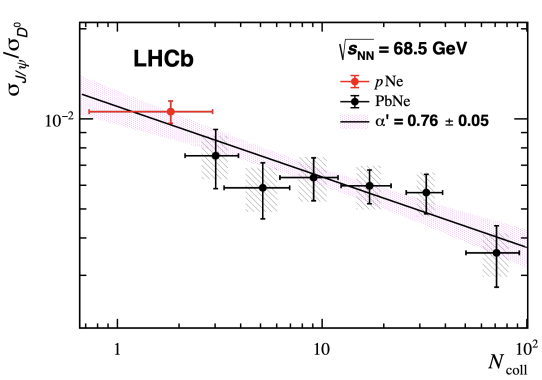
(a) R_{AA} of J/ψ at midrapidity.

(b) R_{AA} of J/ψ at forward rapidity.

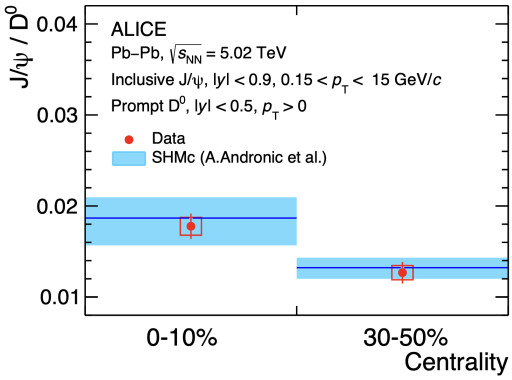
Figure 4.13: Results on the inclusive J/ψ nuclear modification factor from ALICE at midrapidity (a) and at forward rapidity (b) taken from [60] compared to model calculations. Figures taken from [60] under Creative Commons license.

Despite these observations, the experimental signatures discussed above do not yet provide a clear distinction between the two proposed mechanisms: quarkonium production occurring entirely at the phase boundary and the scenario of continuous dissociation and regeneration throughout the lifetime of the QGP. A crucial parameter in both approaches is the total charm production rate. At present, this quantity is not known with sufficient precision to fully exploit the predicted differences in the ratio of J/ψ production to the total charm yield. Fortunately, the total charm cross section can be measured experimentally and constitutes an important objective of the ALICE and LHCb programs during LHC Run 3, which focus on open heavy-flavor measurements both in collider and fixed-target configurations. As an initial step in this direction, analyses have begun to investigate the ratio of J/ψ to D^0 meson production, integrated over transverse momentum.

This observable has already been measured by ALICE in PbPb collisions at $\sqrt{s_{NN}}$ =



(a) Production ratio $J/\psi/D^0$ (LHCb).



(b) Production ratio $J/\psi/D^0$ (ALICE).

Figure 4.14: The production ratio of $J/\psi/D^0$ as function of N_{coll} in proton-Neon and lead-Neon collisions measured by LHCb (a) taken from [61] and the same ratio measured by ALICE at midrapidity compared with the statistical hadronization model (b) taken from [60]. Figures taken from indicated references under Creative Commons license.

5 TeV during Run 2 at midrapidity, and by LHCb in PbNe collisions at $\sqrt{s_{NN}} = 68.5$ GeV in fixed-target mode at forward rapidity, as it is shown in figure 4.14. The ratio obtained by ALICE is consistent with the expectations of the statistical hadronization model. Nevertheless, the comparison of the LHCb fixed-target measurements with both the statistical hadronization framework and transport-model calculations still needs to be done.

Now turning to measurements of excited charmonium states, become experimentally challenging, especially in heavy-ion collisions. These measurements enable the study of the feed-down chains of charmonium through studies of prompt χ_c and prompt $\psi(2S)$ production. Results from the LHCb collaboration indicate that, in pPb collisions at forward rapidity, where the charmonium is detected in the direction of the incoming proton, the contribution from χ_c decays is consistent with that measured in pp collisions [62]. In contrast, at backward rapidity, a larger χ_c feed-down contribution production is observed, as shown in figure 4.15.

This behavior can be understood in light of measurements of the $\psi(2S)$ state [63], which show a stronger suppression at backward rapidity than at forward rapidity. The observed suppression of $\psi(2S)$ has been interpreted in terms of rescattering processes involving the excited charmonium state and comoving particles or gluons in the medium.

Motivated by observations in the correlation and hadronization sector [64], which suggest that signatures commonly associated with heavy-ion collisions gradually appear in pp and pA systems as the multiplicity increases, quarkonium production ratios have been studied extensively as a function of the produced particle multiplicity in these smaller collision systems. Here, both CMS and LHCb observe a pronounced decrease of the $\psi(2S)/J/\psi$ production ratio with increasing multiplicity in pPb collisions, seen both at high transverse

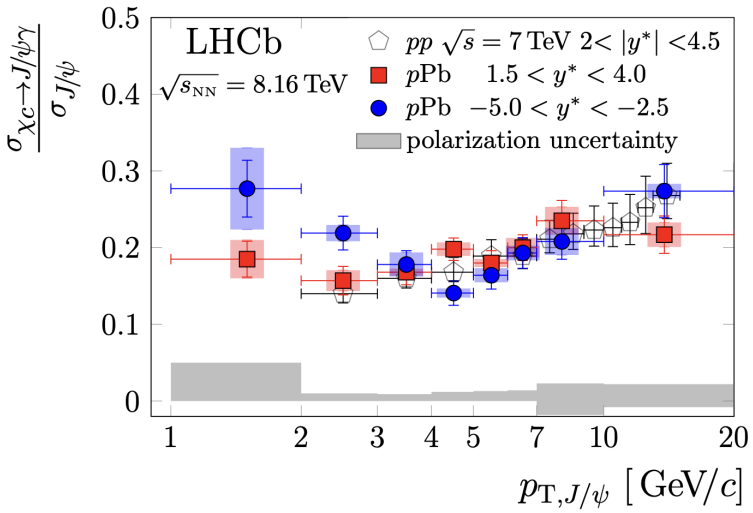


Figure 4.15: Fraction of χ_c decays in the prompt J/ψ yield in pPb and pp collisions as a function of $p_T^{J/\psi}$. Figure taken from [62] under Creative Commons license.

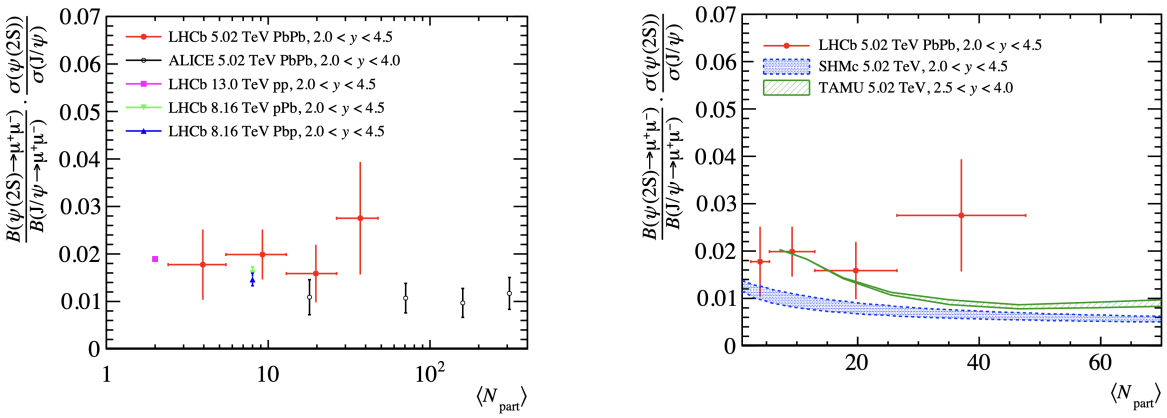
momentum in CMS data [65] and in transverse-momentum–integrated measurements by LHCb [63].

The study of higher charmonium states production in pPb collisions therefore, provides a crucial baseline for disentangling cold nuclear matter and small-system effects. With these results, the natural course of action is to study the production of these states in heavy-ion collisions, where the presence of QGP introduces additional mechanisms that can strongly influence quarkonium production. However, at the moment, very few and recent results are available due to the experimental difficulty involved in carrying out these analyses. Measurements of $\psi(2S)/J/\psi$ production ratio in PbPb collisions using Run 2 data were carried out by LHCb [66] and ALICE [67] collaborations. The results show no significant dependence with the centrality, as it can be seen in figure 4.16, but with the experimental precision in that data sample is difficult to reach a firm conclusion.

Regarding the P-wave charmonium states, χ_c states, they decay radiatively into J/ψ and a low-energy photon that is difficult to find in the background of thousands of photons resulting from neutral pion decays. For this reason, there are no experimental measurements of these states in heavy-ion collisions to date.

4.5.3.3 The importance of measuring χ_c states

The discussion above lead directly to the main motivation of this work. The J/ψ meson is one of the most extensively studied probes of the QGP produced in heavy-ion collisions. However, the interpretation of its measured yields is significantly complicated by the fact that an experimentally measured J/ψ signal does not consist exclusively of directly



(a) Comparison of $\psi(2S)/J/\psi$ production ratio with several collision systems.

(b) Comparison of $\psi(2S)/J/\psi$ production ratio with theoretical models.

Figure 4.16: Cross-section times branching fraction ratio between $\psi(2S)$ and J/ψ states as a function of the number of participating nucleon. Comparison with other collision systems (a) and with theoretical models (b). Figures taken from [66] under Creative Commons license.

produced J/ψ mesons. A substantial fraction of the observed yield originates from the decay of higher charmonium states, in particular the P-wave states χ_{c1} and χ_{c2}, through the radiative decay chain $\chi_c \to J/\psi\gamma$. This contribution is generically referred to as the *feed-down* from χ_c to J/ψ. As emphasized in the comprehensive review of Lansberg [68], an accurate and dedicated measurement of χ_c production is therefore an indispensable ingredient for a correct physical interpretation of any J/ψ measurement.

From detailed analyses of charmonium hadroproduction cross sections at mid-rapidity, the feed-down contribution from χ_c decays to the inclusive J/ψ yield in pp collisions has been estimated to be approximately 25% [69], as represented in the right-hand side of figure 4.17. This is a sizable fraction, comparable in magnitude to the contribution from $\psi(2S)$ decays and from b-hadron decays, and its precise determination directly affects all measurements of the nuclear modification factor R_{AA} of J/ψ in heavy-ion collisions. Indeed, if χ_c states are suppressed differently from the directly produced J/ψ in the QGP, then the measured inclusive J/ψ suppression pattern will necessarily reflect a mixture of these two effects, and cannot be attributed solely to the dissociation of the J/ψ itself.

This point is central to the sequential suppression picture [70] introduced previously, according to which the different charmonium and bottomonium states dissociate in the QGP at successive temperature thresholds determined by their respective binding energies. The χ_c states are P-wave charmonia with a smaller binding energy than the J/ψ ground state, their dissociation temperature is therefore expected to lie closer to the deconfinement threshold T_c. As a consequence, in a medium with temperature $T \gtrsim T_c$, the χ_c states may

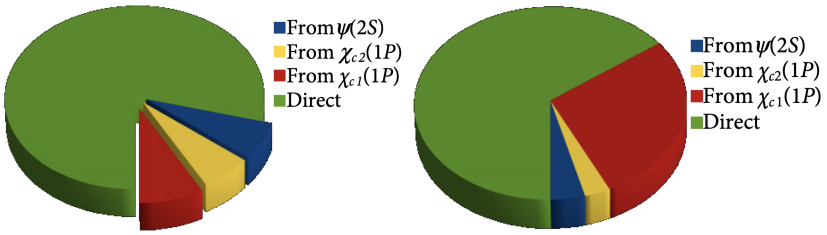


Figure 4.17: Scheme of the typical sources of hadroproduced J/ψ for low-p_T J/ψ in the left and high-p_T J/ψ in the right. Figure taken from [68] under Creative Commons license.

already be melted while the directly produced J/ψ remains intact. The suppression of the feed-down contribution then produces an effective suppression of the measured inclusive J/ψ yield, even in the absence of any direct suppression of the J/ψ itself, as shown in figure 4.11. Disentangling these two effects is only possible if the χ_c production cross section and its nuclear modification are measured independently.

Beyond the context of heavy-ion collisions, the measurement of χ_c production in pp collisions is valuable in its own right as a test of quarkonium production models. The measurement of χ_c states, as well as their feed-down contribution was achieved by LHCb and ATLAS [71, 72]. The knowledge of the χ_{c1} and χ_{c2} polarizations is an essential element for the determination of the acceptance correction of the corresponding cross section measurements, and the corrected production yields can vary by more than 50% depending on the polarization assumed. Constrains to the χ_{c1} and χ_{c2} polarizations in pp collision where measured by CMS [73]. This additional uncertainty propagates directly into the estimated feed-down fraction, making precise χ_c measurements critical not only for interpreting heavy-ion data, but also for constraining the theoretical frameworks, such as CSM, CEM and NRQCD, that are used to describe the production of all charmonium states. In summary, the χ_c states occupy a unique position in the charmonium family: they are directly produced in hard scatterings, they feed down into the most experimentally accessible quarkonium state, and their binding energies place them at the frontier of QGP-induced sequential suppression, making their measurement one of the most physically motivated and experimentally challenging tasks. Indeed, the purpose of this work is to gain deeper insight into the charmonium production mechanism in extreme conditions through the first measurement of the χ_c states in OO collisions, providing a new probe of both quarkonium formation and medium-induced effects.

5

LHCb experiment at the LHC

5.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the largest particle accelerator in the world, located at CERN (from French, Conseil Européen pour la Recherche Nucléaire) near Geneva, Switzerland. The LHC main accelerator ring is located in a 27-kilometer long tunnel. The accelerator tunnel is situated in-between Swiss and French border. This tunnel was originally built in the eighties for the Large Electron-Positron (LEP) collider and is situated between 50 and 175 meters underground. A complete and detailed description of the LHC can be found in [74].

Protons and heavy-ions are accelerated to speeds very close to that of light, with the purpose of performing hadronic collisions at extremely high energies. In the standard configuration, the machine accelerates two counter-rotating beams of protons organized into bunches in two separate rings. These beams are controlled by 1232 superconducting dipole magnets that bend the trajectory inside the machine and 392 quadrupole magnets that focus the beams.

LHC operation time is mostly dedicated to accelerate protons beams. These beams first pass through several and smaller accelerators before entering in the LHC main ring. First, protons are accelerated in the Linear Accelerator 2 (LINAC2) where they reach 50 MeV, then they go up to 1.4 GeV in the Proton Synchrotron Booster (PSB), 25 GeV in the Proton Synchrotron (PS), and 450 GeV in the Super Proton Synchrotron (SPS). Finally, they enter into the LHC, where each beam reaches energies of up to 7 TeV. Besides that, LHC also dedicate part of its operation time to accelerate ions, typically lead ($^{208}\mathrm{Pb}^{82+}$). Here the lead beam is produced using an Electron Cyclotron Resonance (ECR). Then it

is accelerated by the Linear Accelerator 3 (LINAC3) and passing through a carbon sheet of 1 μm width to strip some electrons from the lead atoms. Just after this step, the beam is composed by Pb^{54+} and stored in the Low Energy Ion Ring (LEIR). The ions are then accelerated by the PS and, after going through a second aluminium sheet that strip the rest of the electrons, are sent to the SPS before being injected in the LHC.

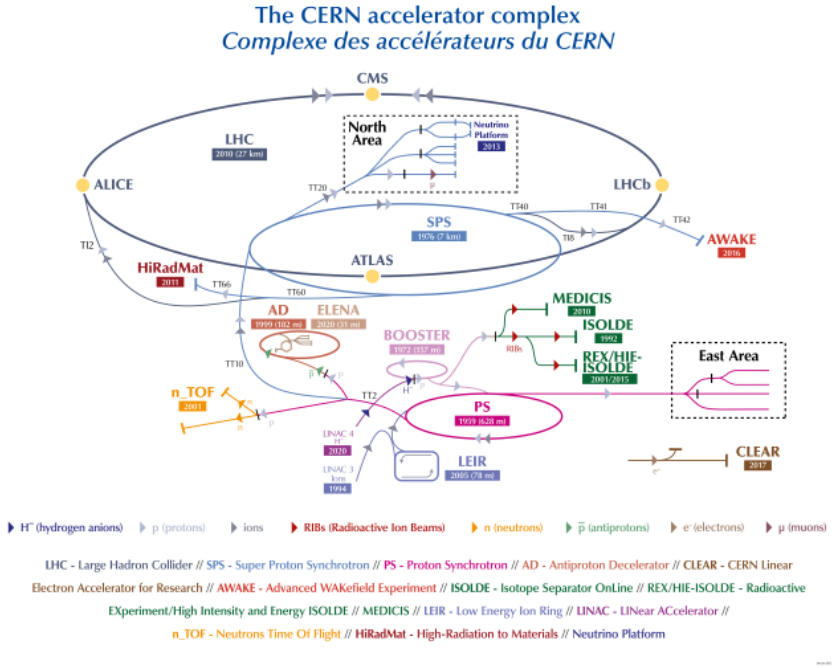


Figure 5.1: Diagram of the CERN accelerator complex in 2022. It shows the different accelerators that compose the accelerator complex and the flow of different particle beams. The figure also show the location of the four main experiments ALICE, ATLAS, CMS, and LHCb. Figure taken from [75] under Creative Commons license.

The collisions take place in four interaction points around the LHC ring, where the four main experiments of CERN are located. These are: ATLAS (A Toroidal LHC Apparatus), CMS (Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment) and LHCb (Large Hadron Collider beauty).

ATLAS [76] and CMS [77] are general-purpose detectors with similar research interests. Primarily, their research lines focus on Beyond the Standard Model (BSM) searches and the Higgs boson study. Their most remarkable achievement was the discovery of Higgs boson in 2012 [5,6]. ALICE [78] was designed for the study of the extremely hot and dense matter created in heavy-ion collisions, with the purpose of describing the Quark Gluon Plasma (introduced in section 4.3) dynamics. Lastly, LHCb [79] is principally designed for the study of CP-violation, heavy-flavor physics and the study of new physics BSM through indirect searches. Over the years LHCb has contributed with many results. One

of the most recent highlight results from the LHCb collaboration is the first observation of CP violation in baryon decays [80]. Furthermore, LHCb is involved in an ambitious heavy-ion program aiming to complement the results of the other experiments, taking advantage of its unique coverage on the forward region. A further description of LHCb experiment will be given in the following section 5.2.

5.2 The LHCb Detector in Run 3

The LHCb detector was originally designed for the study of particles composed by the heavy-quarks charm and beauty. It is located at point 8 of LHC, in the cavern where DELPHI [81] (Detector with Lepton, Photon and Hadron Identification) was operating while LEP was running. After its construction, the detector had an improvement of its subdetectors and software during the LHC Long Shutdown 2 (LS2). This is the *LHCb Upgrade I*. All the specifications that concern the LHCb Upgrade I can be found in [82]. The actual layout of the LHCb upgraded detector is shown in figure 5.2.

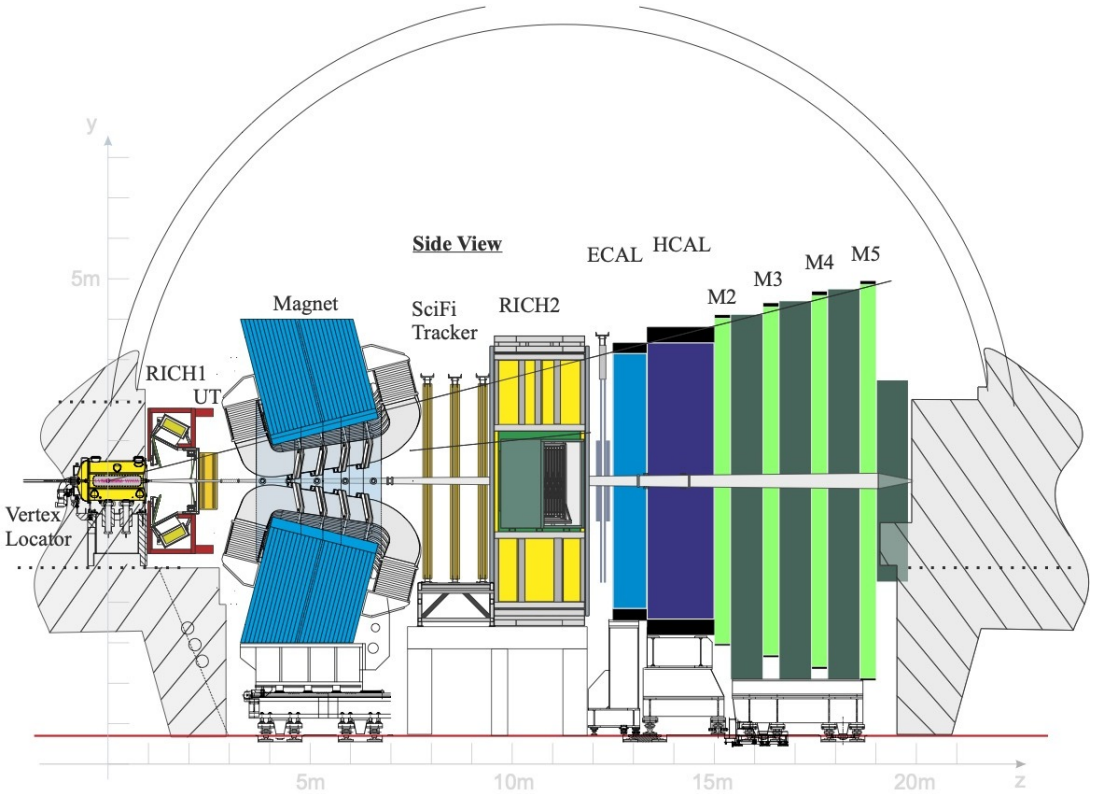
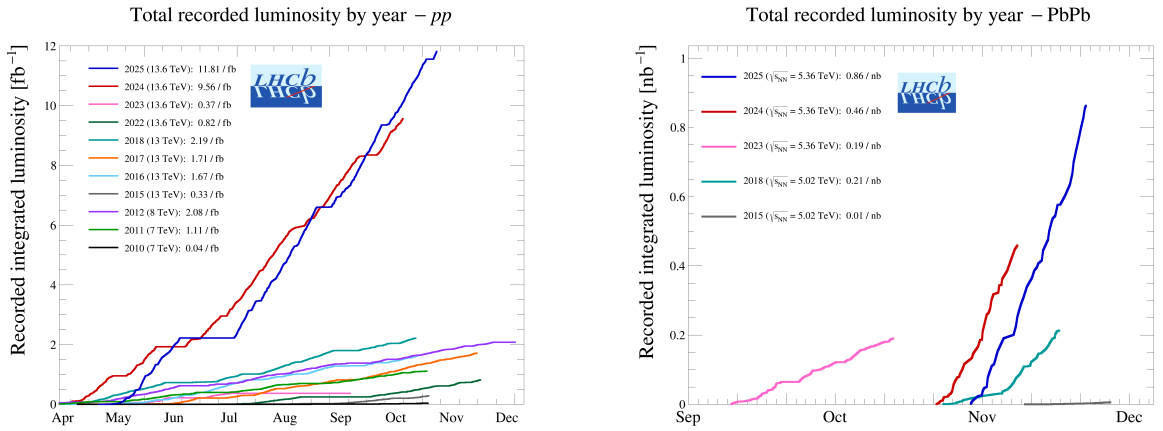


Figure 5.2: Side view of the layout of the LHCb upgraded detector. Figure taken from [82] under Creative Commons license.

LHCb is a single-arm forward spectrometer with a pseudorapidity coverage $2 < \eta < 5$. As mentioned before, this coverage in the forward region makes LHCb an outstanding

detector for the study of heavy-flavor particles. Furthermore, this feature is the distinction signature from the rest of the LHC experiments that are configured as 4π-detectors. The aim of covering this kinematic region is to optimize the acceptance of the detector for B-hadrons that are produced in the hadronization of b quarks at low angles.

Contrasting with the other detectors at LHC, LHCb reduces the incoming particle flux by reducing the number of collisions per interaction. This reduces the integrated luminosity in LHCb in comparison to that of the other experiments. As the occupancy in the detector is higher in the forward region than in central, this luminosity reduction ensures low-pile-up events to facilitate the reconstruction of the Primary Vertex (PV) and the Secondary Vertex (SV). This pile-up refers to multiple independent collisions within the same bunch crossing, so the detector records many overlapping events instead of a single pp collision. On the other hand, pile-up is not taken into account when lead ions are collided, since in this case there are much fewer than one collision per bunch crossing. As it can be seen in figure 5.3 the integrated luminosity has been increased for Run 3 compared to Run 1 and Run 2, leading to higher pile-up in the case of pp and higher multiplicity events in general. These high multiplicity events pushes some subdetectors to their limits, exceeding their granularity and radiation endurance capabilities. This effect is particularly noticeable in the case of PbPb, where the multiplicity of the event can be thousand times larger than in the case of pp.



(a) Recorded luminosity for pp collisions.

(b) Recorded luminosity for PbPb collisions.

Figure 5.3: Recorded luminosity for pp and PbPb collisions for each years. Figures taken from [83, 84] under Creative Commons license.

The coordinate system of LHCb is presented in figure 5.4. Here collisions occur at the origin of the coordinate system and detected by the Vertex Locator (VELO) which surrounds the interaction point. Then, the particles originated in the collision can pass through RICH1 (Ring Imaging Cherenkov detector 1) and the Upstream Tracker (UT). After that, the trajectories of the charged particles bend in the magnet and pass through the

Scintillating Fibre tracker (SciFi) and RICH2. Finally, particles go through the electromagnetic and hadronic calorimeters (ECAL and HCAL) and the muon stations (M2-M5). In the following a more detailed description of each subdetector will be given.

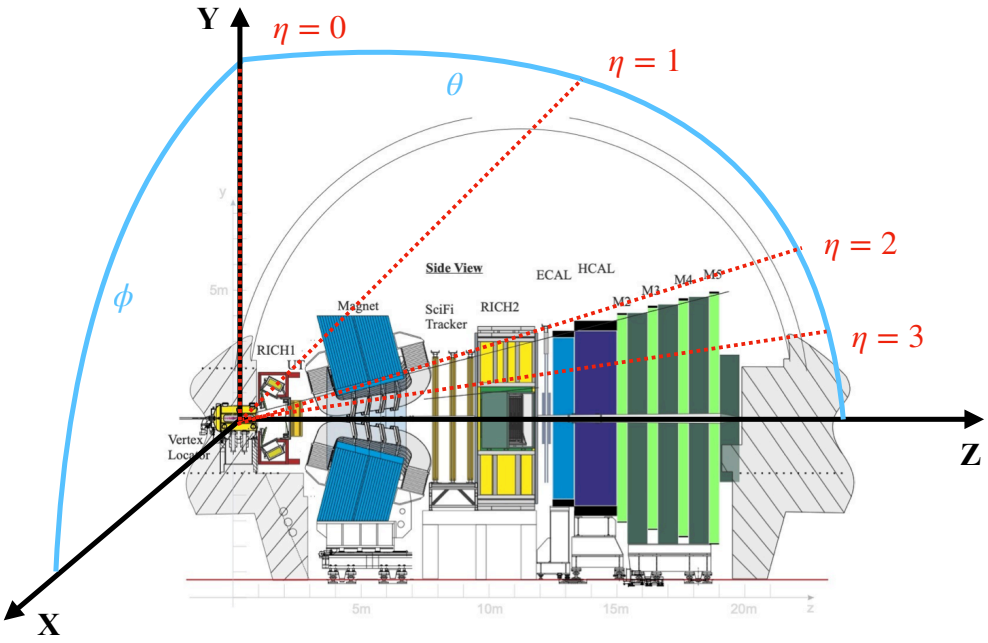


Figure 5.4: Overview of the coordinate system of LHCb, with the definition of the angles η, θ and ϕ.

5.2.1 Dipole Magnet

A dipole magnet is a fundamental piece in a spectrometer dedicated to the study of charged particles as it allow to bend their trajectory in order to measure their charge and momentum. The magnet has been maintained unchanged with respect to Run 1 and Run 2, so the specifications of its design can be found in [85]. It provides a vertical magnetic field with a bending power of $\simeq 4$ Tm. The magnetic field can be configured as *magnet up* (MU) and *magnet down* (MD), corresponding to the dipole magnetic field being along the positive or negative y-axis. The change of polarity is done in order to minimize the possible systematics uncertainties in measurements.

5.2.2 Tracking system

The tracking system of a particle physics experiment is one of the most important component in the detector. As it name indicates, this system allow to track the trajectories of the particles that pass through the detector. The tracking system is one of the components that is specially sensitive to the very high occupancy produced in PbPb collisions. The

LHCb tracking system is composed by three tracking subdetectors that are the VELO, the UP and the SciFi. Each element will be described in the following.

5.2.2.1 Vertex Locator

The Vertex Locator (VELO) [86], is the jewel in the crown of LHCb as it is its most precise subdetector. The VELO has been redesigned to be compatible with the luminosity increase and the trigger-less 40 MHz readout requirement of the upgraded experiment. The VELO is a silicon vertex detector surrounding the interaction point very closely. It consists of two VELO halves on either side of the beam axis, called side A and side C, each consisting out of 26 L-shaped pixelated ASICs arranged perpendicular to the beam line. An overview of the arrangement of the VELO stations in both A and C side and the geometry of VELO closed can be seen in figure 5.5.

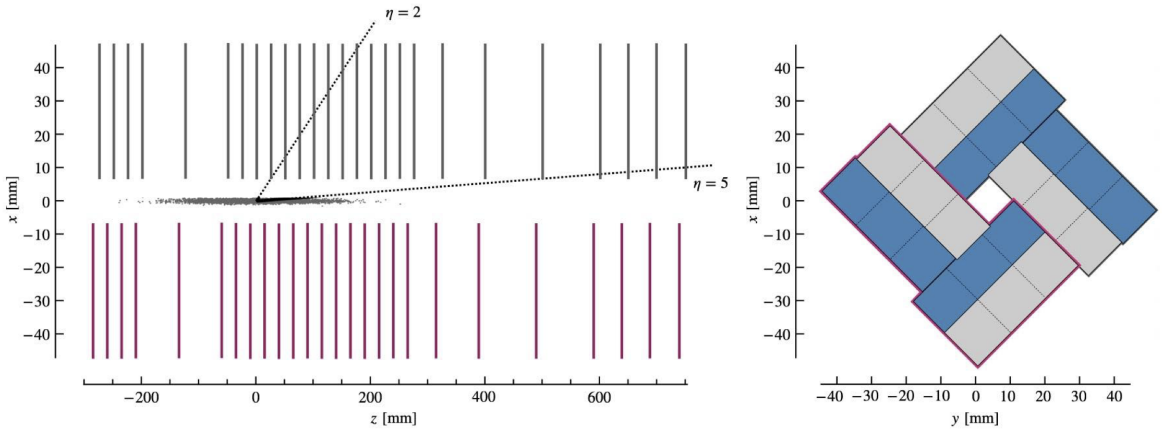


Figure 5.5: In the left, a schematic layout of the VELO foils around the interaction point. In the right, the layout of the ASICs around the z axis in the closed VELO configuration. The modules on the Side C are highlighted in purple on both sketches. Figure taken from [82] under Creative Commons license.

During operations, the inner part of the VELO is positioned 5 mm around the LHC beams, being the subdetector positioned more closely to the beam pipe among all the detectors in LHC. To avoid damage to the sensitive detector during unstable beam conditions such as during ramp-up or injection, both VELO halves can be moved by about 3 cm on a horizontal axis such that both halves are closed during measurement and open as soon as conditions destabilize. While the VELO is closed, there is a small overlap between the two VELO halves for alignment purposes. The VELO sensors are located in a secondary vacuum, separated from the primary vacuum of LHC where the primary collisions occur.

The VELO is a crucial subdetector of LHCb as it allows to distinguish with great precision primary vertices (PV), where the initial pp or PbPb collision takes place, and secondary

vertices (SV), where beauty or charmed hadrons decay. In addition, tracks reconstructed in the VELO are used as a seed for the rest of the tracking system. It is also a key piece of the High-Level Trigger (HLT) system, that will be described later. This subdetector provides very high hit resolution and excellent track reconstruction efficiency (higher than 97%) and very low ghost rate (lower than 3%).

5.2.2.2 Upstream tracker

The Upstream Tracker (UT) [87] is located between the RICH1 detector and the dipole magnet. After the detector upgrade, it replaces the Tracking Turicensis (TT). The UT detector comprises four planes of silicon detectors organized in two stations, as shown in figure 5.6. They are arranged in vertical units, called *staves*. The first station (labeled 'a') is composed of an x-measuring layer (UTaX) with vertical strips and a stereo layer (UTaU) with strips inclined by 5°. Both layers are made of 16 staves each. The second station ('b') is similar, with first a stereo layer (UTbV) with opposite inclination, and a layer with vertical strips (UTbX). Both layers contain 18 staves each. The two pairs of stations are symmetrically arranged around z =2485 mm.

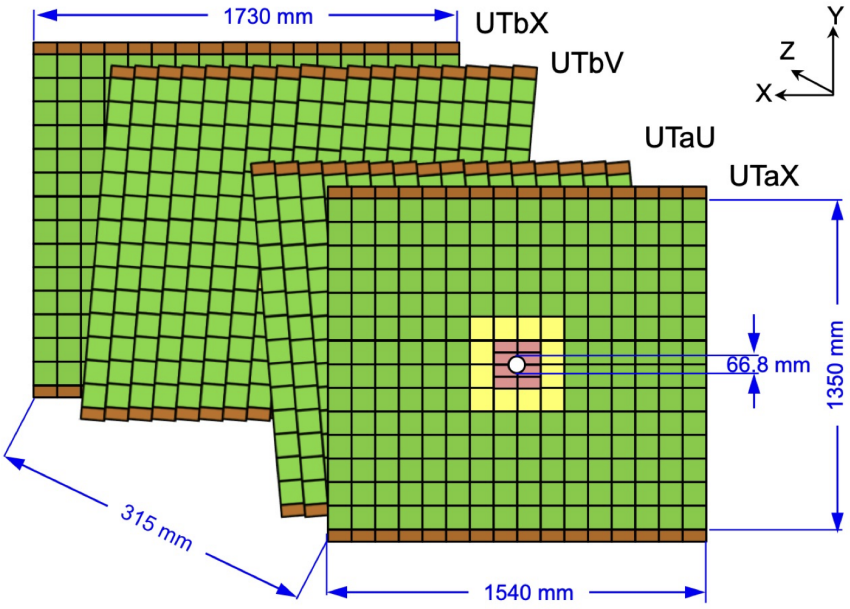


Figure 5.6: Drawing of the four UT silicon planes with indicative dimensions. Different colors designate different types of sensors. Figure taken from [82] under Creative Commons license.

The purpose of the UT is to help in the reconstruction of *long tracks* (tracks with information of all the tracking subdetectors) but also build other type of tracks as *upstream*

tracks (with VELO and UT information) and *downstream tracks* (with UT and SciFi information). These last mentioned tracks are extremely useful for long-lived particles analyses. Furthermore, the UT exploit the magnetic field between the interaction region and the dipole magnet to obtain a first determination of the track momentum with moderate precision ($\sim$15%).

UT subdetector play an essential role during Run 3 data taking. Adding the information of the hits registered in the UT allow to improved the over all tracking performance, increasing the tracking efficiency and reducing the ghost rate[1]. This is specially noticeable in PbPb collisions where the occupancy of the detector is extremely high. Nevertheless, due to some delays in the installation of the LHCb upgrade, UT couldn't be used during the first period of Run 3 data taking. The usage of this subdetector started in April 2024.

5.2.2.3 Scintillating Fibre tracker

The Scintillating Fibre tracker, or SciFi [87] is the tracker located right after the dipole magnet and it is the upgraded version of the tracking stations that operated during Run 1 and Run 2. This subdetector uses a single detector technology based on 250 μm diameter plastic scintillating fibres arranged in multilayered fibre mats. It has in total 12 detection planes arranged in 3 stations called T1, T2 and T3, with 4 layers in each one, as seen in figure 5.7. These layers are distributed in a X-U-V-X configuration. The X-layers have their fibres oriented vertically and are used for determining the deflection of the charged particle tracks caused by the magnetic field. The inner two stereo layers,U and V,have their fibres rotated by $\pm 5°$ in the plane of the layer for reconstructing the vertical position of the track hit.

The SciFi has a very high single hit detection efficiency above 95% and single hit spatial resolution better than 100 μm in the bending plane of the magnet. As mentioned before, this subdetector play a role in the reconstruction of long tracks and downstream tracks and gives a more accurate determination of the momentum of the charged particles that go through it. In contrast to *pp* collisions case, where SciFi performance remains excellent, in PbPb collisions this subdetector is one of the most affected due the very high occupancy in this collision system. In this regime, the SciFi saturates when the collision centrality reaches 30%, which severely degrades its efficiency and prevents it from providing reliable measurements.

5.2.3 Particle identification

Particle identification (PID) consists on the classification of the detected particles according to their mass. In LHCb the identifiable particles are K, π, p, μ, e and γ. The LHCb PID system is composed by two Ring Imaging Cherenkov detectors called RICH1 and RICH2, that are dedicated to the identification of the mentioned hadrons. Then,

[1]Ghost rate refers to the fraction of reconstructed tracks that do not belong to a real particle.

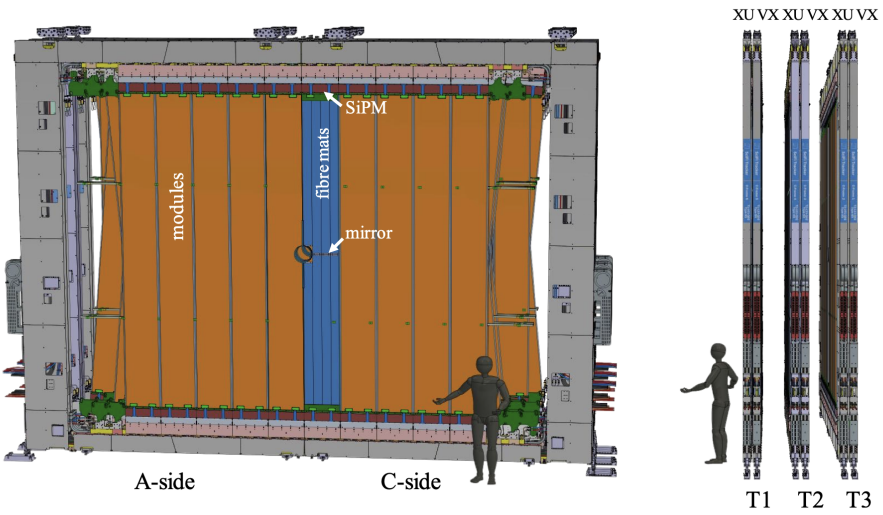


Figure 5.7: In the left, a schematic of four SciFi layers making up a station behind each other. In the right, the three stations, T1, T2 and T3, behind each other. Figure taken from [82] under Creative Commons license.

the electromagnetic and hadronic calorimeters (ECAL and HCAL) are in charge of the identification of e and γ and finally the muon stations identify the μ. The PID system is crucial for the present work, since all the analysis focus on muons and photons overall, but also pions.

5.2.3.1 RICH detectors system

The RICH system of LHCb provides particle identification of charged hadrons over the momentum range 1.5–100 GeV/c [88]. RICH1 is located between the VELO and the UT, while RICH2 is placed between the SciFi and the Calorimeters. The overall structure of both RICH detectors remained unchanged in the upgrade, however with significant modifications to RICH 1 optical system. An schematic view of RICH detectors can be seen in figure 5.8. RICH1 employs C_4F_{10} gas radiator at standard temperature and pressure allowing to provide PID in the momentum range between 1.5-60 GeV/c. RICH2 uses CF_4 gas radiator at standard temperature and pressure and is designed to provide PID for higher momentum particles, between 15-100 GeV/c. RICH system measure the momentum and the mass of the particles going through it via the emission of Cherenkov radiation when those particles travel faster than the speed of light in a medium different from the vacuum.

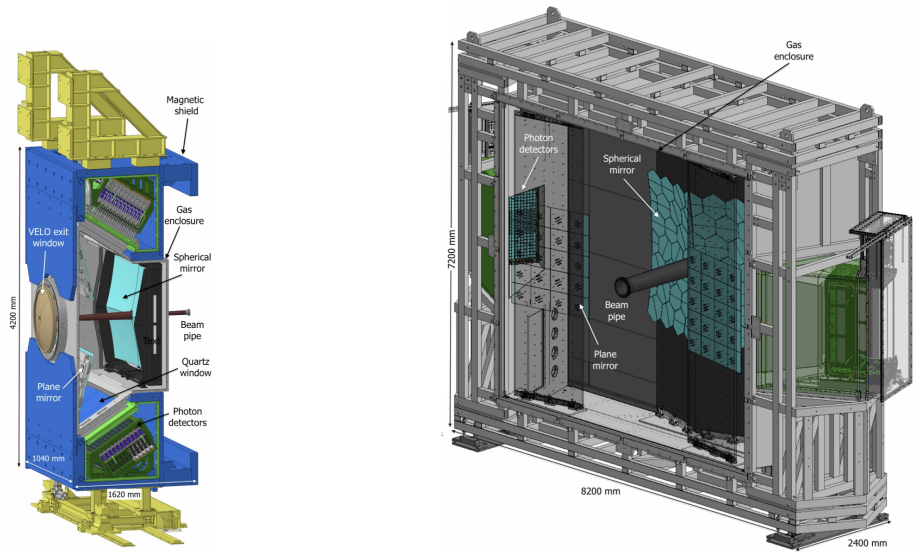


Figure 5.8: An schematic view of RICH1 in the left and RICH2 in the right. Figure taken from [82] under Creative Commons license.

5.2.3.2 Calorimeter system

The calorimeter system play a key role in a PID process and is going to be a crucial part of this analysis, especially in the identification/separation of γ from π^0. The electromagnetic and the hadronic calorimeters, named as ECAL and HCAL, are located between the RICH2 and the muon stations [88]. The general layout of both ECAL and HCAL [89], shown in figure 5.9, remains unchanged for the upgrade. Nevertheless, some changes have been done in the calorimeter system. The readout electronics of the ECAL and HCAL have been entirely redesigned and replaced. Moreover, two subdetectors of the previous calorimeter system, namely the Scintillating PadDetector (SPD) and the PreShower (PS), have been removed, given their reduced role in the new LHCb all-software trigger.

The ECAL consists of scintillator layers alternating with lead layers as absorption material. There are three differentiated regions, that are outer, middle and inner, with different cell granularity being 120x120, 60x60 and 40x40 mm^2 cells respectively. These dimensions are optimized to provide an uniform measurement of the transverse energy. Its role is to absorb photons and electrons within the detector, inducing the development of electromagnetic showers and providing information for the identification of electrons and photons. The total energy deposited in the ECAL in an event also can be linked to N_{part} using the Glauber Model in order to determine the centrality of a PbPb collision.

The HCAL alternates its scintillator layers with thin iron layers. It has two different regions that are outer and the inner with 260x260 and 130x130 mm^2 cells respectively. These dimensions are given by the typical spread of hadronic showers and also optimized to provide an uniform measurement of the transverse energy. It absorbs hadrons traversing

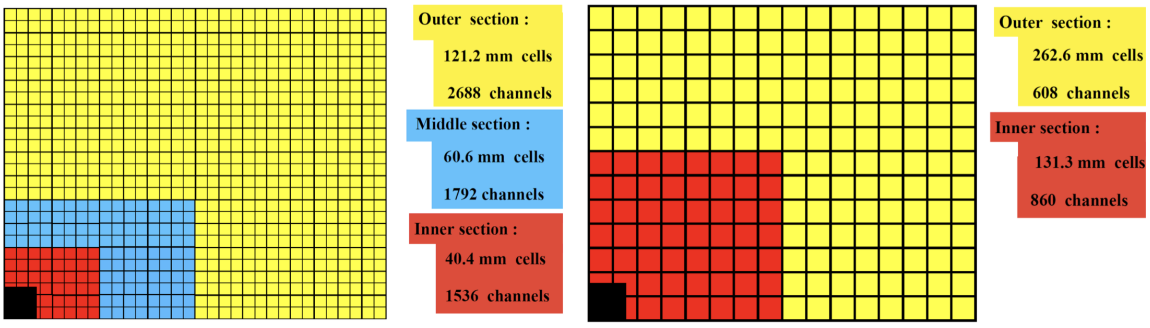


Figure 5.9: A lateral segmentation of the ECAL in the left and the HCAL in the right. One quarter of the detector front face is shown. Figure taken from [82] under Creative Commons license.

the detector by initiating hadronic showers that are subsequently contained within the iron layers. It also supplies complementary information to enhance hadron identification, which is primarily performed by the RICH detectors.

5.2.3.3 Muon system

After passing through all the detector, the only remaining particles with a little energy loss are muons. The muon system [88] is the one in charge to measure those muons and it is going to be the other crucial subdetector for this analysis. The muon stations (M2-M5) are located at the end of the detector. The general layout of the muon stations can be seen in figure 5.10. The LHCb Run 1 and Run 2 muon system additionally included a station M1 located before the calorimeters. Station M1 was utilized in the hardware L0 trigger and thus is no longer needed in the upgraded system. Its supporting structure has been maintained to host the SciFi neutron shielding. Each station is composed of two mechanically independent halves, Side A and Side C.

Muon stations consist of multiwire proportional chambers (MWPCs) filled with a mixture of carbon dioxide, argon and tetrafluoromethane. The MWPCs are interposed with iron absorbers to filter low energy particles. Each station is divided into four regions, R1 to R4, of increasing area moving from the central beam axis outwards. As mentioned, as muons are the only charged particles that can pass through all the detector without being absorb, their detection can be achieve by the muon system with a very high efficiency. They can also be employed for muon tracking, as they provide geometrical hit information. However, this capability is used almost exclusively to evaluate the track reconstruction efficiency.

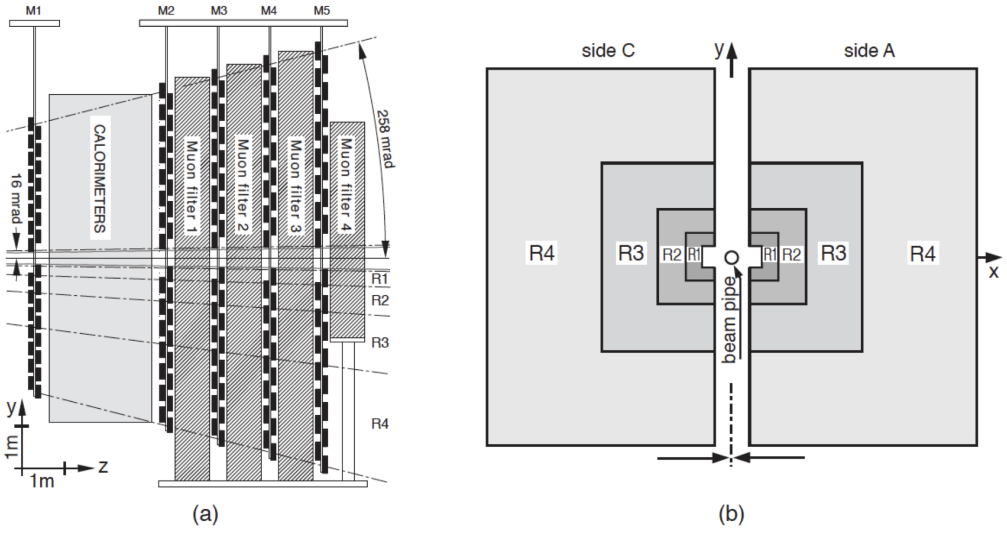


Figure 5.10: In the left, the side view of the muons stations. In the right, the front view of on muon station. Figure taken from [88] under Creative Commons license.

5.2.4 Track reconstruction

As mentioned before, during Run 3 there has been an increase of the luminosity delivered by LHC. This motivated the upgraded subdetectors to reduce their material budget, since minimizing the material in the detector reduce the interaction with it. This leads to an increase of the efficiency and a more precise energy estimation. Taking advantage of this, LHCb stands out from others detectors at CERN, allowing to measure low energy particles more precisely.

Track reconstruction is the process of determining the trajectories of charged particles as they traverse the detector and estimating their momenta. This is achieved by combining the electrical signals generated by the ionization of the detector material as a charged particle passes through it. Each of these electric signals is called a *Hit* in the detector. By combining hits from different subdetectors it is possible to build the different types of tracks in LHCb. High track reconstruction efficiency is essential, particularly for decay channels that produce a large number of daughter particles. This track reconstruction efficiency is another feature specially affected by high multiplicity events, as very central PbPb collisions, since the large number of tracks generated in those events make the reconstruction more difficult. An scheme of the different types of tracks that can be reconstructed in LHCb is shown in figure 5.11. The details of each of these tracks will be given later in section 5.2.4.2 where the reconstruction step in which they are built is introduced.

In Run 3 the data taken by LHCb performs Real-Time-Analysis (RTA) and is processed

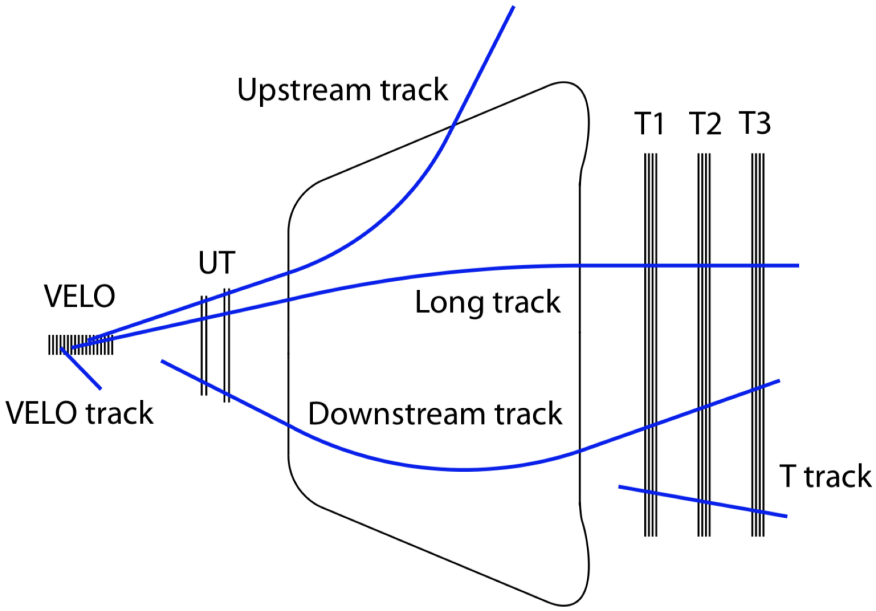


Figure 5.11: Scheme of the different track types defined in LHCb tracking system. Figure taken from [90] under Creative Commons license.

purely using software-based triggers. The processing is done in two steps of High Level Trigger (HLT). The details of each HLT step, namely HLT1 and HLT2, are given in the following.

5.2.4.1 HLT1 reconstruction

The first step of the reconstruction in LHCb is HLT1, which is carried out by the *Allen* project [91]. Allen is an implementation of partial reconstruction sequences on GPUs. Allen reconstructs tracks using hits in the VELO, UT, and SciFi subdetectors, and determines both primary and secondary vertices. It identifies electrons through ECAL clusters, recovers bremsstrahlung photons, and matches tracks to hits in the muon stations for muon identification. Predefined event-selection criteria in the HLT1 trigger lines use this information to classify and filter events.

Before any reconstruction is performed, a Global Event Cut (GEC) is applied, rejecting incoming events with high detector occupancy, as they demand a huge amount of computing resources for reconstruction. As an example, in PbPb collisions, this GEC is done in order to prevent the subdetectors to saturate at high multiplicities.

The reconstruction starts decoding and clustering the VELO information in the detector readout boards, called TELL40, the raw event information is then transferred to the Event Builder (EB). Here, a track-seeding method is used to find straight trajectories from VELO

hits, trajectories that are used to find primary vertices (PV) in the VELO [92].

In the subsequent step, hits from the UT are decoded, if the UT is included in the reconstruction sequence. This applies to data collected after August 2024. The decoded UT hits are then associated with the VELO tracks. Once the UT hits have been incorporated, an initial estimate of the track charge and momentum is obtained by extrapolating the tracks to the SciFi detector using a parametrization of the magnetic field. For data recorded before August 2024, the lack of UT information implies that no charge or momentum estimate is available in the HLT1 reconstruction sequence at this stage. In this case, a matching algorithm is employed: standalone SciFi tracks are first reconstructed from SciFi hits, then extrapolated back to the VELO and matched to the VELO tracks. This procedure is optimized for tracks with transverse momentum greater than 1 GeV/c and total momentum greater than 5 GeV/c [93].

From August 2024 onwards, UT reconstruction is integrated into Allen. In addition to the matching algorithm, the forward algorithm is employed. This algorithm is optimized for tracks with transverse momentum greater than 0.5 GeV/c and total momentum greater than 3 GeV/c. The forward algorithm uses the charge and momentum estimates obtained from VELO tracks matched with UT hits to extrapolate the tracks to the SciFi detector. It then searches for compatible SciFi hits around the extrapolated trajectory and associates them with the combined VELO + UT tracks. Thereafter, the matching algorithm is executed to recover tracks that can be reconstructed without UT information. The matching algorithm is executed after the forward algorithm to recover tracks that can be reconstructed without UT information. After that, a Kalman filter [94] is applied to the partial trajectory in the VELO to improve the momentum estimate.

The track with the improved momentum estimation is then extrapolated to the muon stations for possible matching with muon hits [95]. Electrons are identified by associating the tracks with cell clusters in the ECAL, and bremsstrahlung photons are recovered to correct the momentum of electron tracks.

In the last step of HLT1 reconstruction, secondary vertices (SV) are reconstructed from pairs of VELO track segments. The resulting tracks are then forwarded to HLT2 if they satisfy the criteria of any predefined HLT1 trigger line. In HLT2, the reconstruction is performed from scratch, without using any HLT1 reconstruction information except for which trigger line was triggered. The primary purpose of HLT1 reconstruction is to filter out non-interesting events and to perform an initial classification into trigger line categories based on the event content.

5.2.4.2 HLT2 reconstruction

After HLT1 reconstruction, HLT2 reconstruction is performed. The HLT2 takes events that have been partially reconstructed using the Allen algorithm and runs the full online reconstruction on them. The HLT2 runs in the Gaudi-based Moore project of LHCb,

which has three main components: charged track reconstruction, calorimeter reconstruction and particle identification. The full detector information is utilized, although the HCAL contributes little to the reconstruction and is primarily employed for monitoring and improving particle identification. In this step, several track types are reconstructed, as it can be seen in figure 5.11. The details of each of these tracks will be given in the following.

VELO tracks are generated by reconstructing straight-line trajectories from three-dimensional hits recorded in the VELO. A simplified Kalman filter is applied, taking into account multiple scattering and using noise parameters based on the average momentum spectrum.

Seed tracks, also called T tracks, are the standalone tracks reconstructed in the SciFi. These tracks use the hybrid algorithm, which hits of two SciFi layers to build a track and then adding a hit from the layer in between. If this is successful, additional hits are added using a third-order polynomial as a fitting model.

Upstream tracks are reconstructed using the VeloUT tracking algorithm. Here, VELO tracks are used as a baseline to do a linear extrapolation to a UT searching window. All the hits within this searching window around the extrapolated track are collected and used in a least-square fit with the matched VELO track. If the fit quality is good enough, a maximum of one upstream track per VELO track is combined.

Downstream tracks are reconstructed by extrapolating a seed track to the UT using a parametric model of the expected trajectory, based on the momentum estimated from the hybrid seeding fit. All UT hits near the extrapolated track position are added, and a Kalman filter is then applied to remove outliers hits and improve the momentum estimate.

Long tracks are tracks that go through the entire tracking system and can be reconstructed using either the *matching* or the *forward* algorithm. In the *matching* algorithm all the possible combinations of VELO and T tracks are done, extrapolating them to a position in the magnet and estimating their distance. Tracks pairs with a large distance are rejected and then a machine learning algorithm qualifies the quality of the remaining pairs. At this point, UT hits are searched to form part of the long track. After matching, all utilized VELO tracks and SciFi hits are removed. The *forward* algorithm is applied to reconstruct long tracks from the remaining VELO tracks and SciFi hits. If upstream tracks are available for the remaining VELO tracks, their momentum estimate is used. Otherwise, the algorithm searches for compatible SciFi hits without UT information. The VELO track is extrapolated to the SciFi, and compatible SciFi hits are added based on a least-squares fit. The forward algorithm requires at least nine SciFi hits. After combining the VELO or upstream track with the SciFi hits, the momentum is re-estimated. Using the updated momentum, the track is extrapolated once more from the VELO to the UT to search for additional UT hits. Finally, a Kalman filter is applied to reject outliers and further improvement of the momentum estimation.

Fake tracks, also called ghost tracks, are reconstructed tracks with hits that cannot be associated with a simulated particle. These type of tracks represent a big background source, specially in high multiplicity events.

Clone tracks are duplicate track candidates formed from overlapping subsets of detector hits, where a significant fraction of hits is shared between them, and neither candidate is a ghost track.

Regarding to the calorimeter reconstruction, it is done after the reconstruction of the standard tracks. This reconstruction form clusters around energy peaks in the calorimeters using graph-based clustering algorithms [96]. Then, a multivariate network is used to distinguish photons from merged π^0, hadrons or pile-up. Electrons are identified extrapolating reconstructed tracks to the ECAL and combining them with ECAL clusters.

Finally, muons are reconstructed by extrapolating the reconstructed charged track trajectories to the muon stations and combining with hits there. Here, the minimum number of muon hits required for muon identification depends on the momentum of the tracks. In the top of that, a multivariate classifier is used to reduce misidentifications.

5.3 LHCb software and simulation framework

The LHCb software framework collects from all the subdetectors to reconstructs tracks and adding them into the event information. Here, only the physics-interesting decays are saved in order economize storage space. Therefore, is extremely important to constantly verifying the quality of event reconstruction and selection.

5.3.1 Trigger system and data flow

In order to achieve precision measurements it is needed a proper trigger that select only interesting events for study. For this, LHCb uses a pure software trigger that works at a rate up to 40 MHz. As mentioned in the previous section, the software trigger works on two stages: HLT1 and HLT2, programmed on GPUs and CPUs respectively.

Since events that do not pass the trigger stage are permanently discarded, real-time alignment and calibration of the detector are essential. This ensures that issues arising during the trigger stage are identified and corrected promptly, before significant data loss occurs. Achieving this requires fast, precise, and efficient track reconstruction, done partially in HLT1 and completed in HLT2.

HLT1 produces an output rate of approximately 1 MHz, which is written to a buffer with a capacity of about 30 PB. This buffer stores the data for up to 80 hours and then transfer it to HLT2 at an output rate of 1 MHz. In addition to buffering, the data are used immediately for calibration and alignment, enabling real-time monitoring of HLT1 performance. As illustrated in the full trigger scheme 5.12, the data are divided

into three categories: *Full Events*, in which the complete event information is retained. *Turbo Events*, where only the decay of interest is stored to reduce storage requirements. *Calib Events*, which are dedicated to alignment and calibration tasks. HLT2 reduces the combined data stream of these three categories to approximately 10 GB/s. Measurements of the track reconstruction efficiency are used for detector calibration and therefore rely on the Calib Events stream, which preserves the full event information.

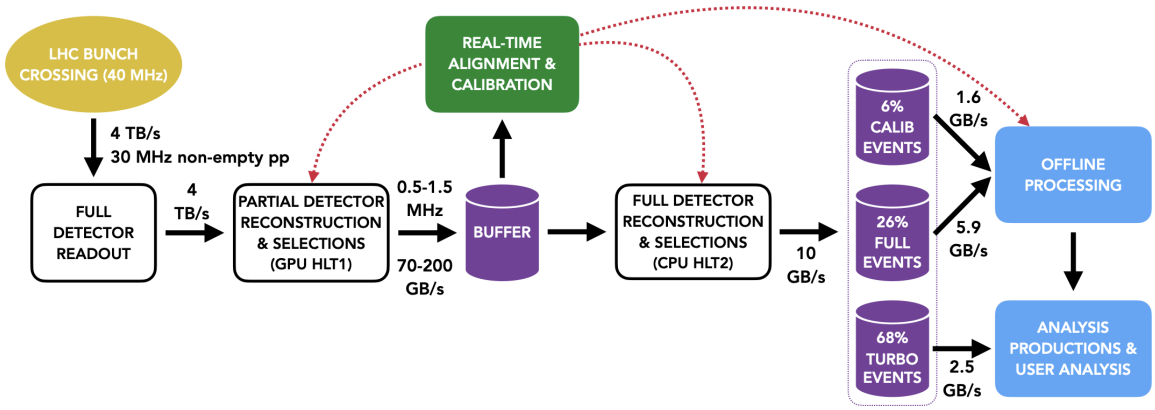


Figure 5.12: Scheme of the Online dataflow. Figure taken from [97] under Creative Commons license.

5.3.2 Simulation framework

Simulations done via Monte-Carlo (MC) generated samples is fundamental to conduct physics analysis in a high energy physics experiment. In LHCb there is a standard way to produce reliable simulations that reproduce the data to perform fundamental studies as estimation of detection efficiencies, determination of background contributions or evaluation of the detector resolution, etc.

The simulation framework that LHCb uses is Gauss [98]. This framework encapsulates different applications used in the different steps of the simulation chain. Gauss uses Pythia8 event generator [99] to simulate pp collisions. In order to simulate PbPb collisions, a pp event is first simulated with Pythia8 and then embedded to an underlying event generated with EPOS [100]. This allows to reproduce all the processes that do not come from the primary vertex hard scattering in a PbPb collision. The decay and time evolution of the produced particles is implemented with EvtGen [101]. Then, the Photos package [102] is used to model the electromagnetic radiation of the final-state particles. Once all the final-state particles are generated, the toolkit GEANT4 [103] is used to reproduce the interactions of these particles passing through the detector. The detector geometry and data-taking conditions are configured to match those of the

targeted data sample through the detector description database (DDDB) and the condition database (CondDB). The output of Gauss is passed to another framework called Boole [104], that is the one in charge of digitalizing the simulated data into the same format provided by the experiment electronics.

6

ECAL calibration for PbPb and OO collisions

The detection and identification of neutral particles play a fundamental role in this work, in which we study a radiative decay involving photons. For this reason, it is essential to ensure proper calibration of the ECAL, as it is the subdetector responsible for detecting the photons of interest. As explained in the Introduction chapter, the analysis was initially developed for PbPb collisions. Consequently, the ECAL calibration procedure presented in this chapter is performed for both PbPb and OO collisions. In particular, the calibration in PbPb provides the necessary foundation for the continuation of this analysis in future studies and can also serve as a valuable reference for other analyses involving photons in PbPb collisions.

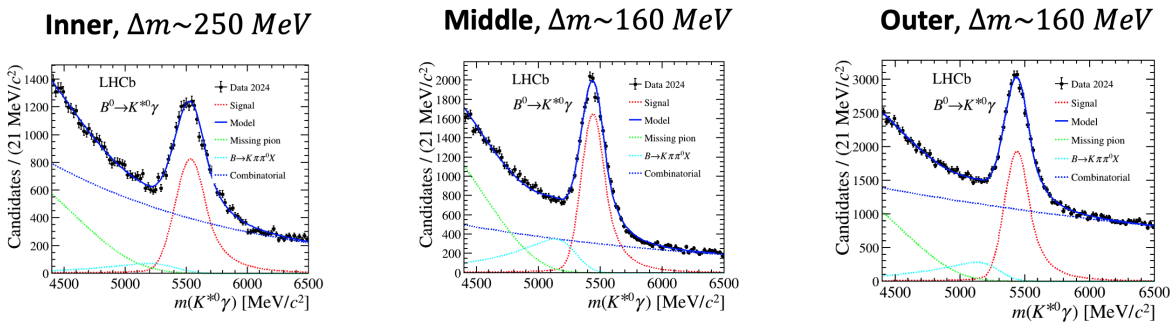


Figure 6.1: B^0 mass shift with respect to the PDG value in the different regions of the ECAL measured in pp collisions data in 2024. Figure taken from [105] under Creative Commons license.

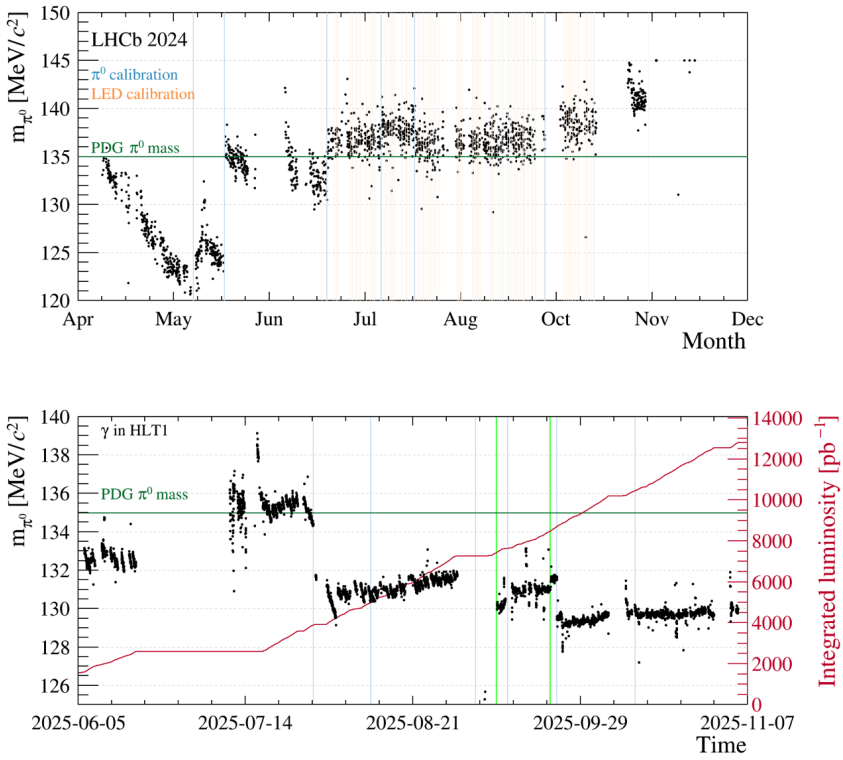


Figure 6.2: Measured π^0 mass as a function of time in regular proton-proton collisions at $\sqrt{s} = 14.5$ TeV. The upper plot corresponds to the 2024 data-taking period, and the vertical blue lines represent when the π^0 calibration was applied. The lower plot corresponds to the 2025 data-taking period, and vertical green lines represent the π^0 calibration. Unpublished figure provided by LHCb collaboration under LHCb policy on the use of LHCb figures in members' theses.

During the 2024 and 2025 data-taking periods, deviations of reconstructed masses from their nominal values were observed in several well-known decay channels, such as $\pi^0 \to \gamma\gamma$ or $B^0 \to K^{*0}\gamma$, as it can be seen in figure 6.1, where the measured mass of the B^0 meson is overestimated with respect to the PDG value in the different regions of the ECAL. This mass shift comes from an overestimation of the photon energy when it is measured by the ECAL. This overestimation of the photon energy has two main sources. On the one hand, there is purely a miscalibration of the ECAL, which, in the case of pp collisions, is addressed periodically, as detailed below. The performance of the ECAL in Run 1 and Run 2 [106] shows that this miscalibration has several causes. Variations in the electronic response from cell to cell can introduce differences in the measured energy. Furthermore, other effects as pile-up contamination, can cause energy deposits from unrelated interactions that may overlap with the signal of interest, artificially increasing the measured cluster energy. These effects are constant over time. On the other hand, the ECAL performance deteriorates over time due to ageing effects. Sustained exposure to radiation

damages the detector components, reducing the gain of the photomultiplier tubes and degrading the light transmission properties of the scintillators and fibres. As a result, the signal amplitude decreases, leading to a gradual change in the detector response. These ageing effects are most pronounced in the inner region of the ECAL, where the radiation levels are higher.

As mentioned, the calibration of the ECAL is performed periodically in the case of *pp* collisions. Here, *resolved* π^0*'s* are used to compute a mass correction in an iterative process. The term *resolved* π^0 refers to π^0 mesons that are detected as the decay of two photons and not merged as a single cluster in the calorimeter. The strategy is the following: first, resolved π^0's are reconstructed. Then, a mass histogram is filled for each cell in the ECAL. Those mass histograms are fitted, and a mass correction is computed for each cell depending on the mass shift observed with respect to the nominal value of the π^0. Finally, the corrections are implemented into the reconstruction framework, and the process is repeated with the updated conditions.

However, these calibrations, which are routinely performed during proton-proton data taking in LHCb, rely on a low and stable background environment. In larger collision systems, the background conditions evolve significantly with collision energy, while the limited integrated luminosity results in insufficient statistics in some ECAL cells. As a consequence, performing the π^0-based calibration online becomes particularly challenging. For this reason, no online calibration was carried out during the PbPb run in November 2024 or during the OO run in the first week of July 2025. The periods during which the π^0 calibration was applied online are shown in figure 6.2 for the 2024 and 2025 data-taking campaigns.

The offline calibration procedure described in the following sections constitutes the first calorimeter calibration performed specifically for heavy-ion collisions, i.e., the PbPb 2024 and OO 2025 data sets in LHCb. Its purpose is to correct the observed mass shift by accounting for both residual miscalibrations and detector ageing effects.

6.1 Calibration strategy for PbPb and OO

Following the strategy used in *pp* collisions, we use resolved π^0's to measure the π^0 mass shift with respect to its nominal value. This mass shift depends on the energy of the photons but also on the region of the ECAL where the cluster of the detected photon is located. As explained in section 5.2.3.2, there are three different regions in the ECAL, according to the distance to the beam-pipe and the size of the cells. Therefore, each region of the ECAL is treated independently.

In the laboratory frame, the squared mass of a π^0 that decays into two photons can be written as

$$m_{\pi^0}^2 = 2E_1E_2(1 - \cos\theta_{12}), \tag{6.1}$$

where E_1 and E_2 refer to the energy of the photons from the π^0 decay and θ_{12} refers to the angle formed by the photons' momenta after the decay. We assume that the energy of the photons is not estimated properly, overestimated as observed in this case although it could also be underestimated, so we correct the energy of each photon with a scale factor λ_1 and λ_2, respectively

$$m_{\pi^0}^{corr} = \sqrt{2\lambda_1 E_1 \lambda_2 E_2 (1 - \cos\theta_{12})}, \tag{6.2}$$

or equivalently

$$m_{\pi^0}^{corr} = \sqrt{\lambda_1 \lambda_2} m_{\pi^0}^{data}. \tag{6.3}$$

This equation gives the corrected mass of the π^0 as a function of the measured mass $m_{\pi^0}^{data}$ and the scale factors. This corrected mass must match the value of the π^0 mass. Finally, assuming both photons are in the same region of the ECAL and bin of energy E^γ, the scale factors become $\lambda_1 = \lambda_2 = \lambda$. Therefore, the mass correction can be written as

$$\lambda(E^\gamma, \text{region}) = \frac{m_{\pi^0}^{PDG}}{m_{\pi^0}^{data}(E^\gamma, \text{region})}. \tag{6.4}$$

6.2 Impact on photon PID

To perform the calibration, it is essential to know how this miscalibration affects the particle identification performance. In particular, as we use photons from a π^0 decay to perform the calibration, it is crucial to understand how the photon PID is affected by the miscalibration. This understanding is also useful in the χ_c measurement discussed in the following chapter.

Photons and neutral pions are detected exclusively using the calorimeter system. However, not every reconstructed ECAL cluster labeled as neutral (meaning it has no associated track) actually originates from a photon or a neutral pion. In practice, hadrons, both charged and neutral, constitute a significant source of background in photon identification.

A particularly problematic background for high-energy photons (with transverse energy $E_T > 2$ GeV) arises from neutral pions. When a neutral pion carries high energy, the two photons produced in its decay can be close enough so that they are reconstructed as a single ECAL cluster. This situation is known as a *merged* π^0 and can be mistakenly identified as a single photon. In contrast, if the two decay photons are reconstructed as distinct clusters, the event is referred to as a resolved π^0, as mentioned before.

For Run 3, the removal of the SPD and PS subdetectors required the development of new photon PID methods that rely solely on information from the ECAL and HCAL. Unlike earlier approaches, these methods use low-level detector features as inputs to classifiers rather than derived high-level variables. In addition, the higher luminosity leads to increased detector occupancy and pile-up, making it necessary to treat pile-up clusters,

where multiple particles deposit significant energy in the same cluster, as an additional source of background. The different types of backgrounds that can be misidentified with photon clusters are presented in figure 6.3. The two variables used within the LHCb software framework in Run 3 are IsPhoton and IsNotHadron. Both of these tools consist of Neural Networks that take multiple input variables to classify a neutral candidate as a photon or another particle species labeled as background as detailed below.

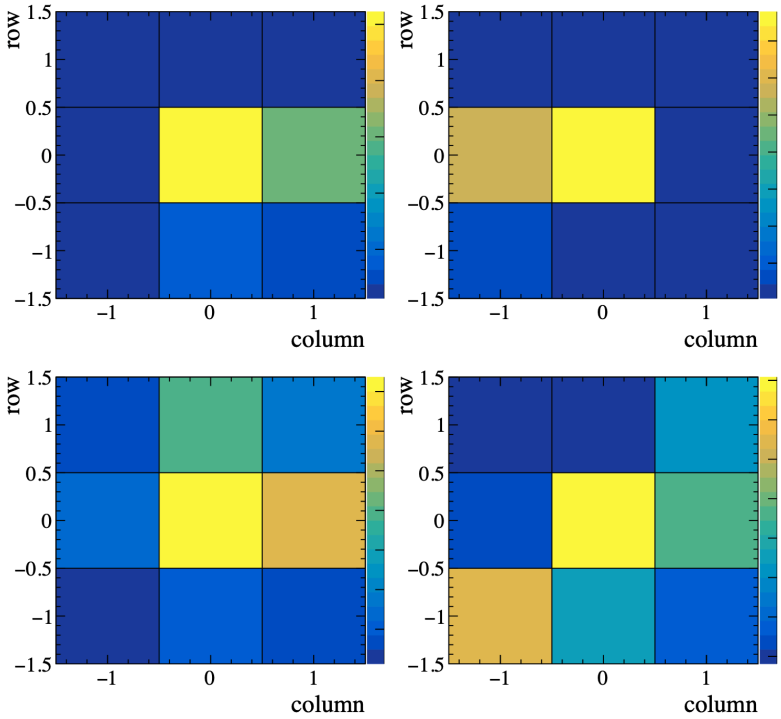


Figure 6.3: Examples of energy deposition of a photon (top left), a merged π^0 (top right), a hadron (bottom left), and a pile-up cluster (bottom right) from a MC Upgrade simulation sample of $B^0 \to K^{*0}\gamma$. Figure taken from [107] under Creative Commons license.

The **IsNotHadron** variable addresses both photon–hadron and photon–pile-up discrimination across all E_T ranges. This is based on the observation that, in both cases, hadrons and pile-up clusters tend to produce wider clusters than genuine photons. This variable is computed using the normalized energy of each cell in a 3x3 cell cluster.

The **IsPhoton** variable is in charge of separating photons from merged π^0's. At very high energies, the angle between the photons coming from a π^0 decay is very small, so it is very difficult to distinguish from a single photon. This separation becomes more difficult in the inner part of the ECAL where the π^0's are very energetic. Similarly to IsNotHadron, this variable is computed using the normalized energy of each cell in a 3x3 cells cluster and, in addition, the transverse energy E_T that can identify most of the merged π^0.

To select π^0 candidates for the calibration, it is necessary to determine whether these PID

variables are affected by the miscalibration. Studies in *pp* collisions have shown a significant discrepancy between data and MC simulation in the distribution of IsPhoton in the inner region of the ECAL, arising from ageing effects. In contrast, the IsNotHadron distributions are found to be in good agreement between data and simulation. The selection criteria described below are therefore done with these observations in mind. This effect is also taken into account in the χ_c measurement presented in the next chapter. Since IsPhoton is the variable affected by the miscalibration, no stringent requirement is imposed on it.

6.3 Data sample and selection

As mentioned, the data samples used for this calibration are the 2024 PbPb data sample at $\sqrt{s_{NN}} = 5.36$ TeV corresponding to an integrated luminosity of $\mathcal{L} \sim 440$ μb^{-1} and the OO data sample at $\sqrt{s_{NN}} = 5.36$ TeV corresponding to an integrated luminosity of $\mathcal{L} \sim 5.1$ nb^{-1}. In both cases, not the full data sample is used to carry out the calibration, but only a subset of a ∼10% of the whole data sample is used. These subsets contain all the RunNumber associated with the 2024 PbPb and 2025 OO data-taking periods, which ensures a representative sample of these periods.

The selection of the π^0 candidates is presented in table 6.1. The selections are the same for both the PbPb and OO datasets, with the only difference being the requirement in the transverse momentum of the π^0. Here, the larger luminosity delivered in the OO sample allows to do a tighter cut in this variable, reducing the background from low-energy photons while being under the merged-π^0 threshold ($p_T > 2$ GeV). The ECAL acceptance used for the photon and π^0 selection corresponds to the region considered suitable for physics analyses by the ECAL experts, excluding the innermost part of the calorimeter, which has been affected by radiation damage.

PbPb selection	
Variable	Requirement
Photon p_{T}	$p_T^{\gamma} > 300$ MeV/c
Photon η	$2 < \eta^{\gamma} < 4$
Pion p_{T}	$p_T^{\pi^0} > 800$ MeV/c
Pion η	$2 < \eta^{\pi^0} < 4$
Photon ID	IsPhoton > 0.4 IsNotH > 0.7 $E_{HCAL}/E_{ECAL} < 0.1$

OO selection	
Variable	Requirement
Photon p_{T}	$p_T^{\gamma} > 300$ MeV/c
Photon η	$2 < \eta^{\gamma} < 4$
Pion p_{T}	$p_T^{\pi^0} > 1200$ MeV/c
Pion η	$2 < \eta^{\pi^0} < 4$
Photon ID	IsPhoton > 0.4 IsNotH > 0.7 $E_{HCAL}/E_{ECAL} < 0.1$

Table 6.1: π^0 candidate selection criteria for PbPb and OO datasets.

6.4 Signal extraction

The π^0 signal is extracted by fitting the diphoton mass spectrum. This is done by combining two probability density functions, one to describe the signal and the other to describe the combinatorial background.

A simple Gaussian distribution is insufficient to describe the signal shape. The π^0 mass distribution is affected by residual miscalibrations, which introduce a non-uniform smearing and lead to an asymmetric peak. This asymmetry is well modeled by an Exponentially Modified Gaussian (EMG) function [108], defined as

$$\mathrm{EMG}(m;\mu,\sigma,\lambda) = \frac{\lambda}{2} e^{\frac{\lambda}{2}\left(2\mu+\lambda\sigma^2-2m\right)} \cdot \mathrm{erfc}\left(\frac{\mu+\lambda\sigma^2-m}{\sqrt{2}\,\sigma}\right), \tag{6.5}$$

where the parameters μ and σ represent the mean and the standard deviation of the distribution, respectively, and λ represents the exponential parameter that controls the asymmetric tail. These three parameters are obtained in the fit for each bin of E^γ and the region of the ECAL. Here, it should be noted that the parameter μ does not coincide with the maximum of the distribution as in a Gaussian distribution. We therefore need to calculate the maximum value of the distribution, as well as its error propagation, based on the parameters obtained from the fit.

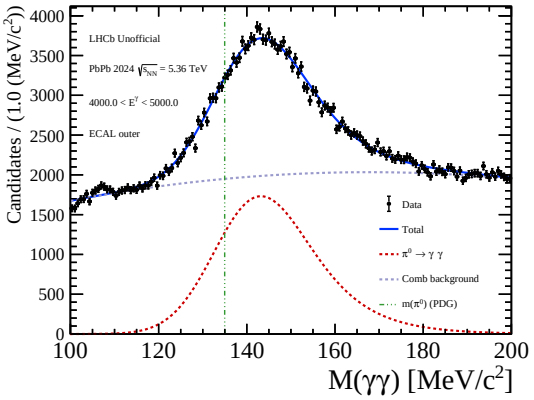
The combinatorial background shape is described with a second-order polynomial function. The three parameters that compose the second-order polynomial function remain free and are obtained during the fit procedure, as these vary significantly depending on the bin in which the fit is performed.

Both data samples are fitted in a 2D-grid of photon energy $E^\gamma \in [1.1, 100]$ GeV and ECAL region $\in$ [outer, middle, inner]. Different binning schemes are used at low energy for PbPb and OO samples, as they had to be adjusted depending on the statistics in these regions.

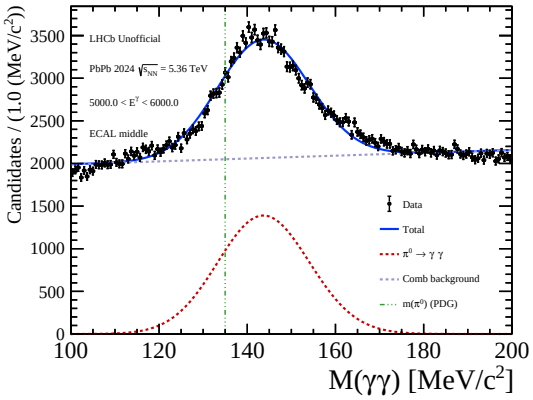
Each fit is classified as usable or not depending on some requirements. First, it is required that the number of candidates in each bin be larger than 1500, ensuring that the fit has enough statistics to be reliable. Second, to avoid non-convergent fits, it is required that $\chi^2_{NDOF} > 0$ and a positive-definite covariance matrix. If the fit does not fulfill these conditions, no correction is applied in this bin.

Due to the large combinatorial background in ion–ion collisions, performing fits to the diphoton invariant mass distribution for low-energy photons in the inner region of the ECAL is particularly challenging. A possible approach to calibrate this region is to use the *pp* reference run, where the lower ECAL occupancy reduces the combinatorial background, making it easier to isolate and reconstruct low-energy photons. However, the *pp* reference run does not produce the same occupation conditions in the detector as PbPb or OO collisions, which would need to be taken into account if this approach is used.

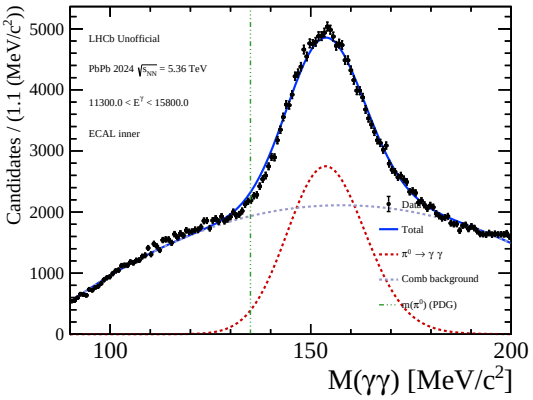
Usable fits are shown in figures 6.4 and 6.5 for the invariant mass of the π^0 for the three different regions of the ECAL in PbPb and OO collisions, respectively. In those figures, the PDG value of the π^0 mass is marked with a vertical dotted green line. Here, it can already be seen that the π^0 mass shift is higher when we go closer to the beam-pipe, i.e., it has a higher mass shift in the inner region than in the outer region. This dependency on the position and the energy of the photon is studied in detail in the following section.



(a) Fit of the π^0 mass in the outer region.

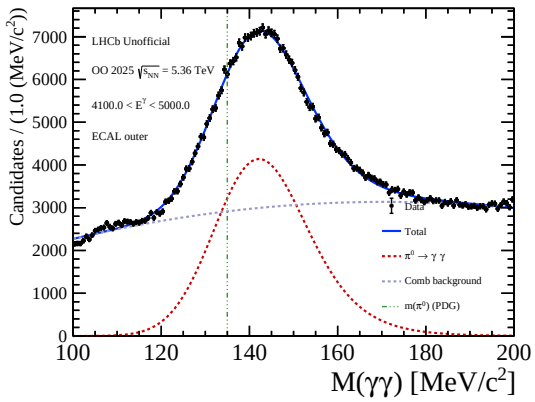


(b) Fit of the π^0 mass in the middle region.

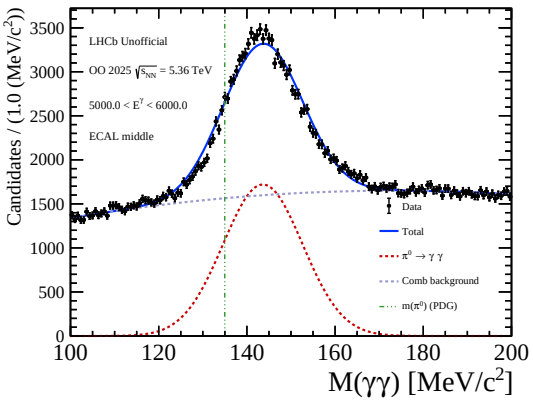


(c) Fit of the π^0 mass in the inner region.

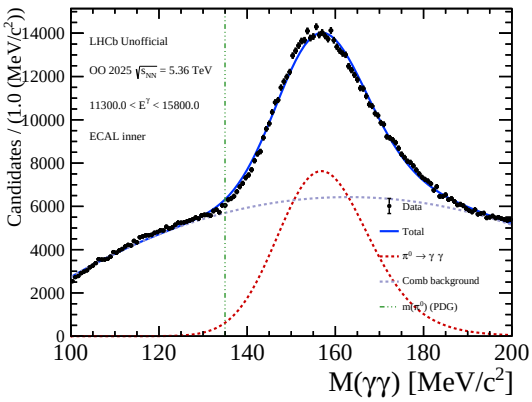
Figure 6.4: Invariant mass fit of π^0 candidates in PbPb collisions in the outer region with $E^\gamma \in [4.0, 5.0]$ GeV (a), middle region with $E^\gamma \in [5.0, 6.0]$ GeV (b) and inner region with $E^\gamma \in [11.3, 15.8]$ GeV (c). The vertical dotted green line indicates the nominal π^0 mass



(a) Fit of the π^0 mass in the outer region.



(b) Fit of the π^0 mass in the middle region.



(c) Fit of the π^0 mass in the inner region.

Figure 6.5: Invariant mass fit of π^0 candidates in OO collisions in the outer region with $E^{\gamma} \in [4.1, 5.0]$ GeV (a), middle region with $E^{\gamma} \in [5.0, 6.0]$ GeV (b) and inner region with $E^{\gamma} \in [11.3, 15.8]$ GeV (c). The vertical dotted green line indicates the nominal π^0 mass

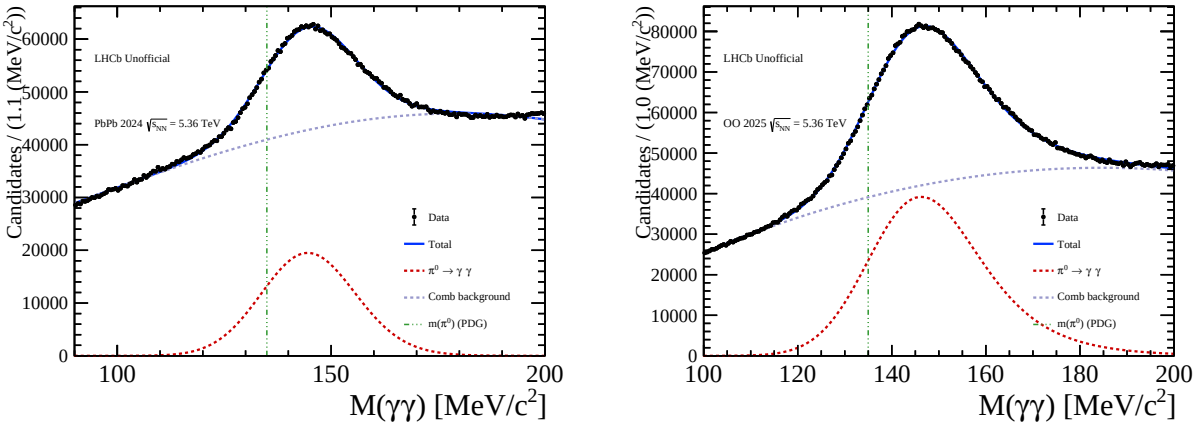
6.5 Calibration results

In this section, the results of the calibration are presented. The corrections to the photon energy are computed according to equation 6.4 using the results of the fits in the different bins of E^{γ} and ECAL region described in the previous section. Furthermore, the calibration over time is also studied.

6.5.1 Mass correction over time

Before computing a correction as a function of E^{γ} and the ECAL region, it is important to check if the average π^0 mass shift is the same for all the data-taking period. In the

case where the mass shift changes over time, a different calibration would be needed for different time periods. To check the average mass correction as a function of time, we fit the diphoton mass spectrum as a function of the RunNumber being integrated over photon energy and all the ECAL regions. Examples of such fits in a range of RunNumber within the PbPb 2024 or OO 2025 data taking periods are shown in figure 6.6.

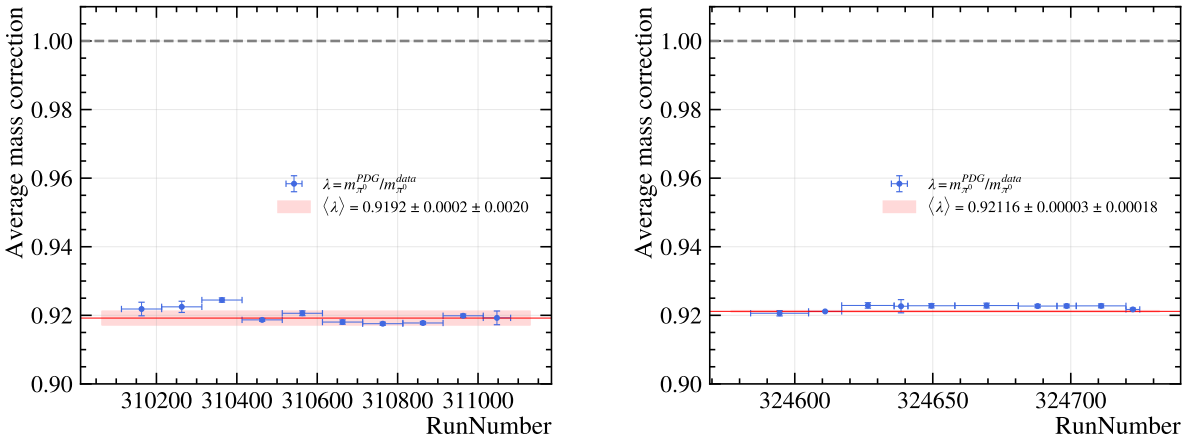


(a) Fit of the π^0 mass within the 2024 PbPb collisions data-taking period.

(b) Fit of the π^0 mass within the 2025 OO collisions data-taking period.

Figure 6.6: Fit of the π^0 mass integrated over E^γ and all the ECAL regions with a RunNumber within 2024 PbPb (a) and 2025 OO (b) data-taking periods.

After fitting the diphoton mass spectrum for different ranges of RunNumber covering the entire period of PbPb and OO data taking, the average mass correction is computed for each RunNumber range as indicated in equation 6.4. The results of the average mass correction as a function of the RunNumber are presented in figure 6.7. Here, the blue points are the computed mass correction for each bin of RunNumber, while the red line is the weighted mean computed from the set of blue points for each case. The red shaded zone around the line corresponds to the Root Mean Squared (RMS) of all of the corrections. In both cases, we see that the mass corrections remain stable during both data-taking periods, meaning that there is no need to calibrate the ECAL for different time periods during PbPb or OO data taking. The average mass correction is $\langle\lambda\rangle = 0.9192 \pm 0.0002 \pm 0.0020$ in the case of PbPb and $\langle\lambda\rangle = 0.92116 \pm 0.00003 \pm 0.00018$ in the case of OO, where the first uncertainty is associated with the statistical error and the second is the RMS. Here, we note that the measured average mass corrections in both PbPb and OO collisions are compatible within uncertainties, indicating that the ECAL is performing consistently during heavy or light ion runs.



(a) Average mass corrections in PbPb.

(b) Average mass corrections in OO.

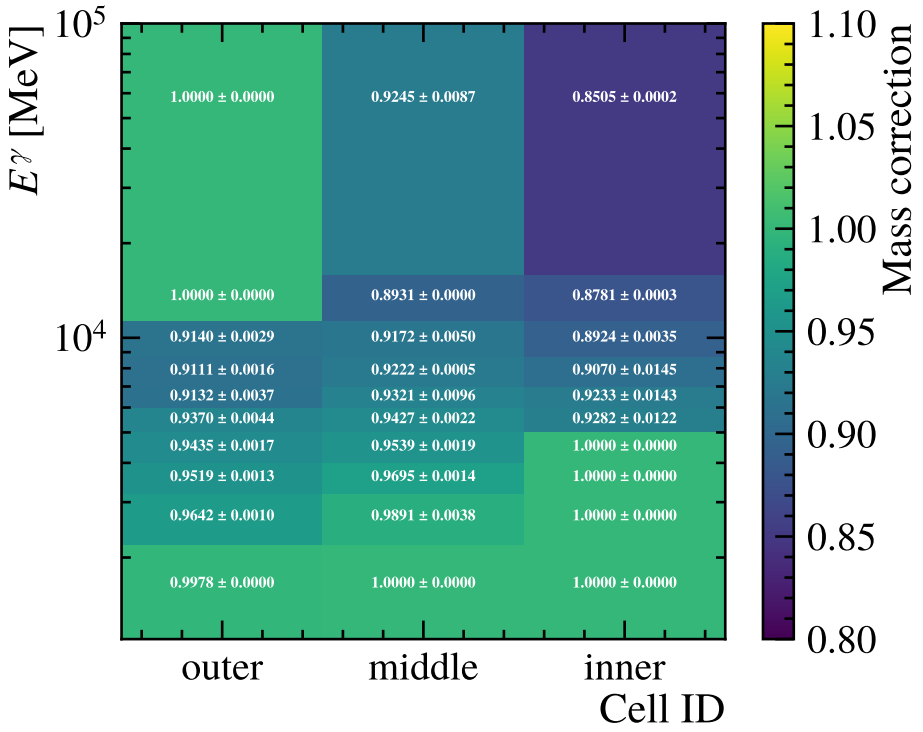
Figure 6.7: Average mass corrections as a function of RunNumber for PbPb 2024 (a) and OO 2025 (b) data taking periods.

6.5.2 Mass correction as function of E^{γ} and ECAL region

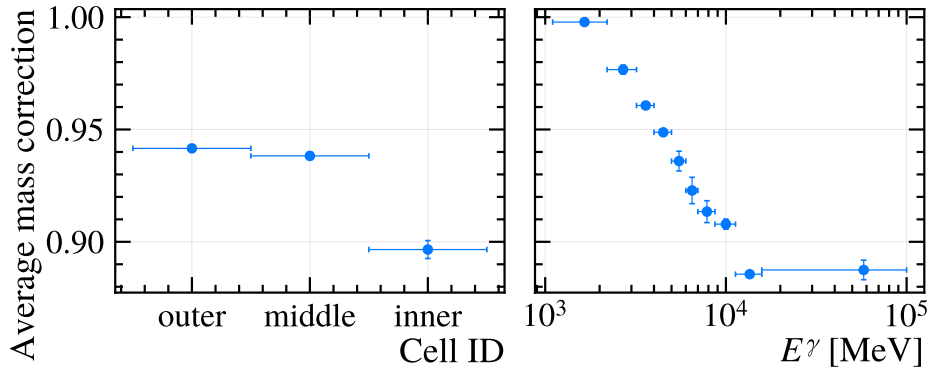
After verifying that the mass correction remains stable throughout the data-taking period, the correction factors are determined as a function of photon energy, E^{γ}, and ECAL region. To this end, the diphoton invariant-mass spectrum is fitted in bins of photon energy and ECAL region. The resulting fit parameters are then used to compute the mass corrections according to Eq. 6.4. The corresponding correction maps are shown in the top panels of Figures 6.8 and 6.9 for PbPb and OO collisions, respectively. Bins with a correction factor equal to unity correspond to regions where no correction is applied, either because the fit did not satisfy the quality criteria described in Section 6.4 or because insufficient statistics were available.

In these tables, we can see that the mass shift increases, i.e., the mass corrections further deviate from unity when the photon has higher energy and gets closer to the beam-pipe. The mass shift goes up to 15% in the case of PbPb and 18% in the case of OO from the nominal mass, in the highest photon energy bin in the inner region. This trend can be seen more clearly if we look at the projections of the computed mass corrections in the photon energy and ECAL region (also named as Cell ID) axes, presented in the bottom row of figures 6.8 and 6.9. The correction maps reveal a noticeable degradation of the ECAL performance in the inner and middle regions between the PbPb 2024 and OO 2025 data-taking periods. This observation is consistent with the progressive ageing of the calorimeter in the regions most exposed to radiation. Nevertheless, this behaviour remains compatible with the time evolution of the average mass correction. Indeed, the average correction is statistically dominated by candidates reconstructed in the outer ECAL region, where the detector response remains stable over the period considered.

The mass corrections obtained for the OO case are applied in the measurement of χ_c later in this work. Applying this correction is a crucial step in the analysis, as this will ensure the particles under study are at their nominal masses. The details of how these corrections are applied in the case of the χ_c analysis are given in section 7.3.3.

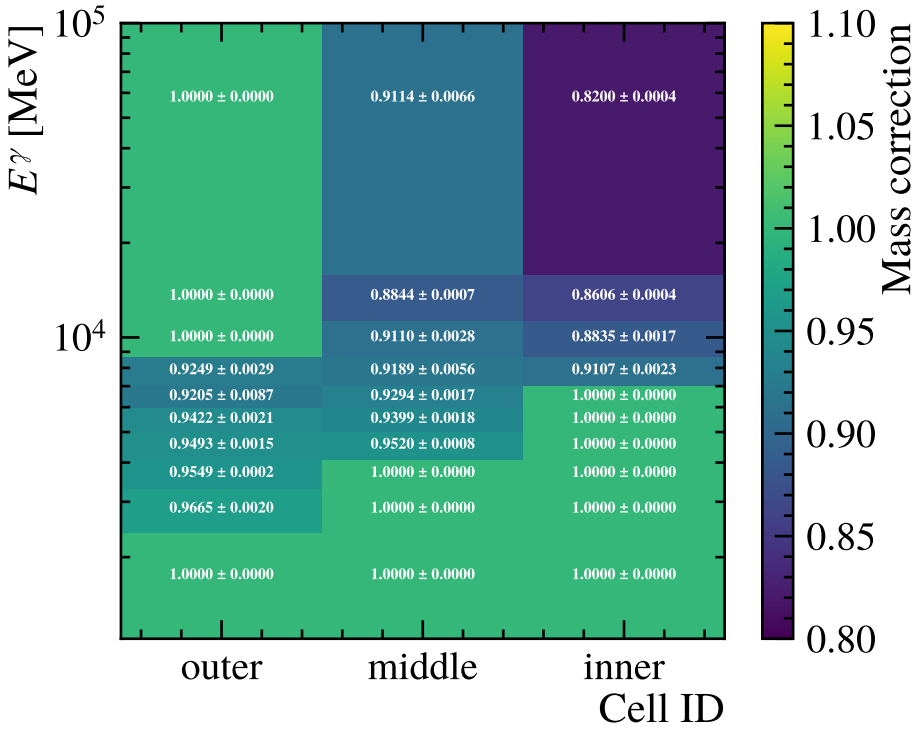


(a) Calibration table with mass correction factors in (E^γ, region) bins.

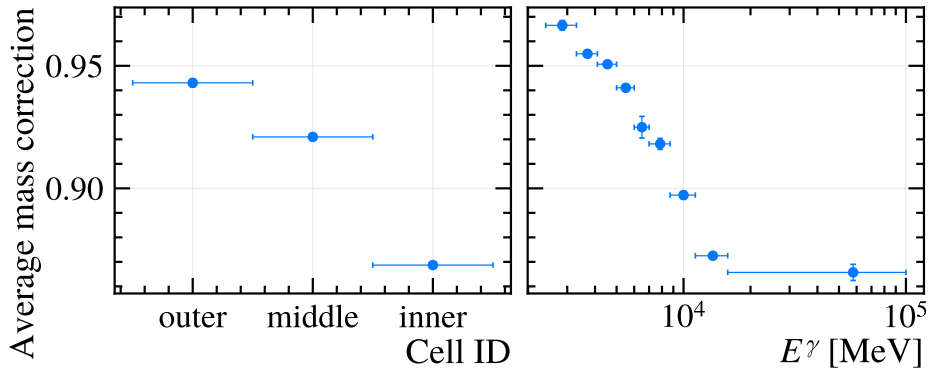


(b) Projection of the mass correction in the E^γ and Cell ID axes.

Figure 6.8: Calibration table with mass correction factors in (E^γ, region) bins (a) and the projection of the values presented in this table in the E^γ and Cell ID axes (b) for PbPb collisions. Only bins with mass correction have been taken into account for the projection.



(a) Calibration table with mass correction factors in (E^{γ}, region) bins.



(b) Projection of the mass correction in the E^{γ} and Cell ID axes.

Figure 6.9: Calibration table with mass correction factors in (E^{γ}, region) bins (a) and the projection of the values presented in this table in the E^{γ} and Cell ID axes (b) for OO collisions. Only bins with mass correction have been taken into account for the projection.

7

χ_c production in OO collisions

Previously in chapter 4, the importance of measuring charmonium states to have a better understanding of the behavior of QGP was discussed. In particular, measurements of the charmonium ground state J/ψ have been developed for many collision systems, allowing the comparison between them. When comparing the J/ψ production, especially in heavy-ion collisions, with the production in pp collisions, we can infer how the created QGP affects the J/ψ production. Unfortunately, it is not possible to have a complete picture of this effect without studying also the excited charmonium states, since the decay of these states contributes to the feed-down production of J/ψ.

Excited charmonium states, such as χ_c, have been measured in very few collision systems at LHC energies, such as pp [109, 110] or pPb [111] collisions. The goal of this analysis is to make a first measurement of χ_c states production in OO collisions and compare it with other collision systems. In particular, this analysis will study the radiative decay channel

$$\chi_{cJ} \to J/\psi\gamma, \tag{7.1}$$

where the subscript $J = \{1, 2\}$ denotes two members of the P-wave charmonium triplet χ_c. The J/ψ meson produced in this decay is reconstructed through its dimuon decay channel, $J/\psi \to \mu^+\mu^-$. The accompanying photon can be reconstructed using two complementary methods. The first method exploits so-called *converted photons*, in which the photon undergoes an electromagnetic interaction with the detector material and converts into an electron-positron pair, $\gamma \to e^-e^+$, before reaching the ECAL. The second method relies on *calorimetric photons*, also referred to as non-converted photons, which are reconstructed from the energy deposited in the ECAL. Both approaches exhibit distinct advantages and limitations. Calorimetric photons are more abundant and can be reconstructed solely from their energy deposit in the ECAL, without requiring associated charged-particle

tracks. Their performance, however, is limited by the intrinsic energy resolution and granularity of the calorimeter, as described previously in chapter 6. In contrast, converted photons are reconstructed from charged tracks, allowing for a more precise determination of the photon kinematics and significantly improving the angular resolution and vertex-pointing capabilities. Their reconstruction efficiency is nevertheless strongly dependent on the detector material budget. In the case of LHCb, the relatively low conversion probability, combined with the challenges associated with reconstructing electrons affected by Bremsstrahlung, results in a lower yield of converted photons compared to calorimetric photons. Therefore, to maximize the available statistics of χ_c candidates, only calorimetric photons are considered in this analysis. The properties of the two χ_{cJ} states under study, together with their mass differences with respect to the J/ψ, are listed in Table 7.1.

	mass (MeV)	$\Delta M(\chi_{cJ} - J/\psi)$ (MeV)	width (MeV)	$\mathcal{B}(\chi_{cJ} \to J/\psi\gamma)$
χ_{c1}	3510.67 ± 0.05	413.77 ± 0.01	0.88 ± 0.05	(33.9 ± 1.2) %
χ_{c2}	3556.17 ± 0.07	459.27 ± 0.01	2.00 ± 0.11	(19.0 ± 0.8) %

	mass (MeV)	width (keV)	$\mathcal{B}(J/\psi \to \mu^+\mu^-)$
J/ψ	3096.900 ± 0.006	92.6 ± 1.7	(5.961 ± 0.033) %

Table 7.1: Basic properties of χ_{cJ} and J/ψ mesons. Values taken from Ref. [28]

7.1 Data Sample and Selection

7.1.1 Data and Monte Carlo samples

The data used to perform this analysis were recorded by LHCb during LHC Run 3 in July 2025 with OO collisions at a center-of-mass energy $\sqrt{s_{NN}} = 5.36$ TeV corresponding to an estimated integrated luminosity of $\mathcal{L} \sim 5.1$ nb^{-1}. The events were reconstructed at HLT1, requiring at least one PV. Furthermore, at this step, a Global Event Cut (GEC) was imposed, where only events containing less than 30000 SciFi clusters were reconstructed. This requirement was imposed to avoid events with very high particle multiplicities, where the occupancy of the SciFi detector becomes sufficiently large to lead to saturation effects and degraded reconstruction performance. Due to the limited luminosity delivered by the LHC, all the events reconstructed at HLT1 passed through the steps of HLT2 and sprucing, i.e., all the events reconstructed were reconstructed by HLT2 and went through the sprucing process with no other selection criteria.

In this analysis, two simulation samples were used, one for the χ_{c1} signal and the other for the χ_{c2} signal. In both cases, Monte Carlo simulations (MC) were generated using a special LHCb tuned version of PYTHIA8 [99] to generate the χ_{cJ} and J/ψ signals, which were superimposed to an underlying OO event from EPOS [100] to reproduce a realistic environment similar to what would happen in a true OO collision. Then the simulated

event is coupled to the GEANT4 transport code [103] to simulate the particles passing through the simulated geometry of the detector. The decay of the χ_{cJ} states was performed by EvtGen [101]. The use of the MC samples is detailed in the appropriate section of this document. Furthermore, a generator-level private MC simulation was used for efficiency computation. The information of each data sample is detailed in Table 7.2. After the GEANT4 simulation, the detector response is emulated using the Boole [104] software and subsequently reconstructed with the standard LHCb reconstruction algorithms.

Process	Generator	Sim version	Beam	# Events
$\chi_{c1} \to J/\psi(\mu\mu)\gamma$	EPOS	Sim10g	OO	4M
$\chi_{c1} \to J/\psi(\mu\mu)\gamma$	Pythia	private	pp	20000
$\chi_{c2} \to J/\psi(\mu\mu)\gamma$	EPOS	Sim10g	OO	4M
$\chi_{c2} \to J/\psi(\mu\mu)\gamma$	Pythia	private	pp	20000
inclusive J/ψ	EPOS	Sim10f	OO	4M
$J/\psi \to \mu\mu$	Pythia	private	pp	10000

Table 7.2: List of MC simulation datasets for this analysis.

7.1.2 Candidates selection in data and MC

Once the events have been reconstructed, both in data and MC, the objective of the candidate selection is to enhance the signal-to-background ratio by applying suitable selection requirements. The same selection criteria are used for all states under study, allowing for a consistent comparison of the measured observables. A first pre-selection is performed at the DaVinci [112] level, that is the software where reconstruction-level information is used to build and select the decay chains. The requirements applied at this stage are optimized using the true information available in the simulated samples. For this, all the variables of interest were compared with the truth-matched signal distributions, in order to search for differences between signal and background and apply a discriminative requirement on them. The complete list of cuts used in this pre-selection is given in Table 7.3.

In Table 7.3, the cuts done in the Flight Distance (FD) and Impact Parameter have the objective to select prompt χ_c candidates. The cut applied to the DIRA variable refers to the cosine of the angle between the momentum of the particle and the flight direction. For the J/ψ candidate and the muon tracks, requirements are imposed on the maximum vertex and track χ^2/NDOF, respectively. In addition, two criteria related to muon identification are applied to the muon candidates, together with a stringent requirement on the ghost probability to suppress fake tracks. In the case of the photons, the cuts applied are the ones considered safe in line with what was discussed in chapter 6.

After this pre-selection, an offline selection is done in order to increase the signal-over-background of our already filtered sample. The information of the offline selection is

listed in Table 7.4. In this offline selection, all the particles are required to be within LHCb acceptance, except the photons, to which a tighter cut has been applied to avoid the damaged inner region of the ECAL. Another important cut applied is the one in the z position of χ_c vertex (VZ) that removes contamination coming from SMOG2 collisions. It has also been applied a tight cut around the J/ψ nominal mass in order to remove background coming from random combinations of muons. In the case of muons, a tighter cut in muon ID has been applied. Finally, the cut on the p_T of the photons helps to reduce the combinatorial background coming from low-energy photons.

After all the cuts done at the pre-selection and the offline selection steps, the invariant mass spectrum $M(\mu\mu\gamma) - M(\mu\mu)$ is computed (see Fig. 7.1). The invariant mass difference, $M(\mu\mu\gamma) - M(\mu\mu)$, is studied instead of the absolute invariant mass $M(\mu\mu\gamma)$ to improve the mass resolution. In this quantity, the contribution of the muon momentum resolution largely cancels, such that the remaining resolution is dominated by the photon measurement. As we can see in Fig. 7.1, the invariant mass spectrum under study is completely dominated by the combinatorial background; no resonance can be distinguished there. Applying just tighter linear cuts will result in a critical reduction of χ_c candidates statistics. Therefore, we need a smarter way to remove the background, keeping as much signal as possible. With this purpose, a machine learning-based selection is applied. All the details about this selection technique are detailed in the following section.

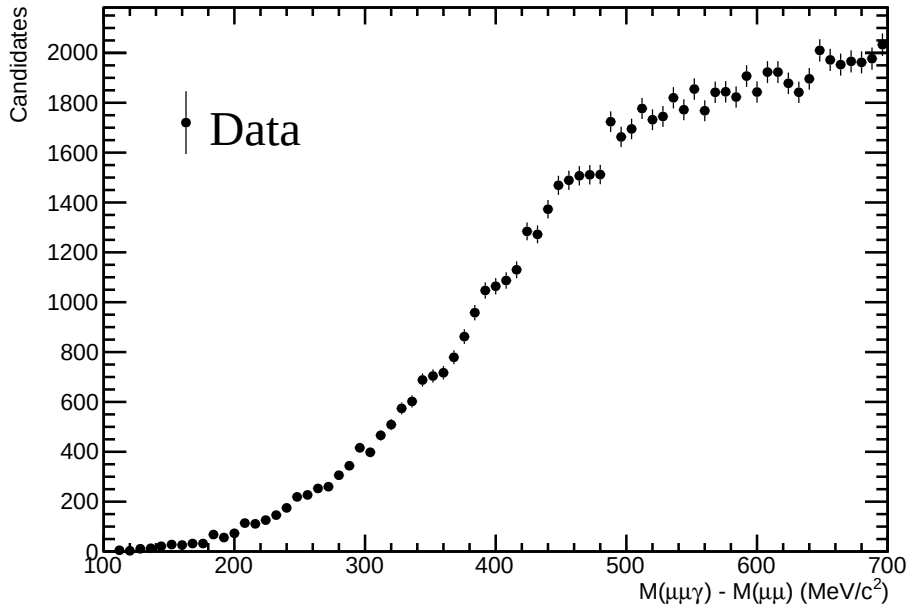


Figure 7.1: Invariant mass spectrum for $M(\mu\mu\gamma) - M(\mu\mu)$ after pre-selection and offline selection in 2025 OO collisions.

Pre-Selection	
χ_c cuts	$\lvert\text{DIRA}(\chi_c)\rvert > 0.95$ $\text{FD}(\chi_c) < 1$ mm $\text{IP}(\chi_c) < 0.05$ mm
J/ψ cuts	$p_T^{J/\psi} > 1000$ MeV/c $\chi^2_{J/\psi} < 6$ $\text{IP}(J/\psi) < 0.05$ mm $\text{FD}(J/\psi) < 1$ mm
Muons cuts	$p_T^{\mu} > 700$ MeV/c $\chi^2_{\mu} < 2$ $\text{IP}(\mu) < 0.1$ mm $\text{ISMUON} > 0.9$ $\text{PIDMU} > -5$ $\text{GHOSTPROB}(\mu) < 0.1$
Photons cuts	$p_T^{\gamma} > 300$ MeV/c $E_{Hcal}/E_{Ecal} < 0.3$ $\text{IsPhoton} > 0.4$ $\text{IsNotHadron} > 0.6$

Table 7.3: Selection done at DaVinci level according to the MC truth information.

Offline Selection	
χ_c cuts	$2 < \eta^{\chi_c} < 4.8$ $p_T^{\chi_c} > 1500$ MeV/c $3200 < M(\chi_c) < 4200$ MeV/c^2 VZ(χ_c) > -200 mm
J/ψ cuts	$2 < \eta^{J/\psi} < 4.8$ $3050 < M(J/\psi) < 3150$ MeV/c^2
Muons cuts	$2 < \eta^{\mu} < 4.8$ PIDMU > 0
Photons cuts	$2 < \eta^{\gamma} < 4$ $p_T^{\gamma} > 600$ MeV/c

Table 7.4: Offline χ_c candidates selection.

7.2 Machine learning enhanced signal extraction

Machine Learning (ML) is a specialized area within computer science focused on developing algorithms that can learn independently from data. Its main goal is to enable computers to automatically improve and adjust their behavior in a way that resembles human cognitive processes. The following sections will explain the main principles of ML and how ML techniques have been applied to this analysis.

7.2.1 Fundamentals of Machine Learning

During the past decade, ML methods have become highly valuable in particle physics, leading to significant progress in the field. Their success is largely due not only to the improved performance achieved in data analysis compared with conventional linear selection methods, but also to their ability to enhance detector performance [113–115]. ML approaches are especially well-suited for classification and regression tasks, which are among the most common challenges in particle physics research. In this Thesis, classification techniques are applied to identify χ_c candidates by separating genuine physical signals from background events.

The classification task belongs to the class of *supervised learning algorithms* [116]. Supervised learning follows a well-defined workflow composed of four main stages: training, testing, validation, and application. In the training stage, the algorithm is provided with a collection of n examples, referred to as the *training set* (A), which is defined as:

$$A = \{(\vec{x}_1, y_1), (\vec{x}_2, y_2), ..., (\vec{x}_n, y_n)\}. \tag{7.2}$$

Each element of A consists of a set of features $\vec{x}_i$, representing the input variables provided to the algorithm (for instance, the physical properties associated with the sample), together with the corresponding target value y_i. In classification tasks, the target corresponds to a class label. Assuming the existence of a deterministic or stochastic function f that associates the input features with the output value, the relationship between the two can be written as $y = f(\vec{x})$. The primary goal of the training phase is to approximate the mapping function f with an objective function to be used as a predictor in the following phases. The objective function can be defined as:

$$\text{obj}(\vec{\theta}) = L(\hat{y} - y, \vec{\theta}) + \Omega(\vec{\theta}), \tag{7.3}$$

where $\vec{\theta}$ are the internal parameters of the ML algorithm. The objective function consists of two components: the loss function $L(\hat{y} - y, \vec{\theta})$ and the regularization term $\Omega(\vec{\theta})$. The loss function quantifies the difference between the predicted value, $\hat{y}$, and the true target value, y. Various types of loss functions can be employed, such as the Mean Squared Error, Binary Cross-Entropy, or Huber loss [117, 118], each characterized by specific strengths and limitations. To achieve the best possible approximation of the function f, the algorithm iteratively updates its parameters to minimize the loss function. This optimization

process can be carried out using different algorithms. The regularization term, instead, is introduced to control the complexity of the model and to mitigate the risks of overfitting and underfitting. Overfitting arises when the model becomes excessively complex and effectively memorizes the training dataset, leading to poor generalization when applied to unseen data. In such cases, the algorithm performs well only on the examples contained in A, while failing to make reliable predictions for new instances. Conversely, underfitting occurs when the model is too simple to capture the underlying relationships within the training data, resulting in poor performance even on the training set itself. Under these conditions, the algorithm is unable to learn the mapping between the input features and the corresponding outputs. Ideally, the algorithm should identify general patterns from the training data that can be successfully transferred to previously unseen examples.

After the training phase has been completed, the algorithm can generate predictions for previously unseen examples by exploiting the relationship between features and output values learned from the training dataset. Before deploying the trained model, however, its performance must be carefully evaluated. This assessment is carried out using a separate and independent dataset, referred to as the *test set*. Evaluating the model on the same data used during training may produce artificially optimistic results, since the algorithm could suffer from overfitting or underfitting, thus providing a misleading estimate of its actual performance. Similarly to the training dataset, the true output values y are also known for the test set. This makes it possible to objectively measure the capability of the algorithm to generalize to unseen data.

At this stage, the test dataset is employed to tune the hyperparameters of the algorithm. Hyperparameters are quantities that cannot be automatically learned during training and therefore must be defined by the analyzer before the start of the training procedure.

Finally, the trained algorithm is applied to the dataset of interest, referred to as the *application set*, for which the true output values are not known. In binary classification problems, the algorithm returns a single numerical output, commonly called a score, which represents the probability that a given instance belongs to the reference class. To assign an instance to a given class, a threshold value must be defined. Depending on the selection strategy, instances with scores either above or below this threshold are accepted or rejected.

7.2.2 Boosted Decision Tree

The Boosted Decision Tree (BDT) [119, 120] is a widely used supervised learning method applicable to both classification and regression tasks. Its basic building block is the Decision Tree (DT) [121], which extends traditional linear selection approaches by allowing events that do not strictly satisfy individual selection criteria to still be considered. This flexibility is particularly important since many candidates do not exhibit all the features that clearly distinguish signal from background, making a single cut-based selection insufficient for effective separation.

A DT is constructed as a sequence of binary decision rules, referred to as nodes. These nodes are arranged hierarchically: the initial nodes, or root nodes, are connected to the input features, while the terminal nodes, or leaves, produce the final output score. At each node, a single feature of the candidate is compared to a predefined threshold. Rather than simply accepting or rejecting an event, the DT assigns a score based on the outcome of this comparison, which is then passed to the subsequent node where another feature-based decision is applied. The more discriminating the sequence of splits, the higher the resulting score for signal-like candidates and the better the separation between classes. This procedure continues iteratively until a stopping condition is reached or all instances are assigned to a leaf.

As with other supervised learning techniques, DTs follow the standard training, testing, validation, and application workflow. During training, the structure of the tree, including the ordering of nodes and the optimization of threshold values, is adjusted to maximize class separation in classification problems or to minimize the difference between predicted and true values in regression tasks. The interpretability of DTs, derived from their sequential decision-making process, makes them more transparent than many more complex machine learning models, whose internal logic is often difficult to interpret due to their "black-box" nature. However, DTs have limited capacity to capture highly complex relationships between inputs and outputs, which is why they are typically considered weak learners. Increasing the depth of the tree (i.e., the number of nodes) improves model complexity but also increases the risk of overfitting.

To address this limitation, boosting techniques are employed. Boosting constructs an ensemble of weak learners and combines them into a stronger predictive model, the BDT. In this framework, trees are trained sequentially, with each new DT using both the original training data and the output of the previous tree. The final BDT prediction is obtained by aggregating the outputs of all individual trees in the ensemble.

BDTs provide a powerful balance between predictive performance and robustness against overfitting compared to deep single decision trees. For this reason, they are the chosen algorithm for the analysis presented in this Thesis. A schematic representation of a BDT composed of two decision trees is shown in Fig. 7.2.

7.2.2.1 XGBoost

The implementation of the BDT for this work has been made from *Extreme Gradient Boosting* (XGBoost) [123] open-source package. XGBoost enables parallel computation on both CPUs and GPUs, significantly reducing training time. This BDT implementation is built upon the gradient boosting framework introduced previously and is characterized by strong flexibility and scalability. It relies on the CART (Classification and Regression Tree) algorithm [124], a decision tree approach capable of addressing both classification and regression tasks.

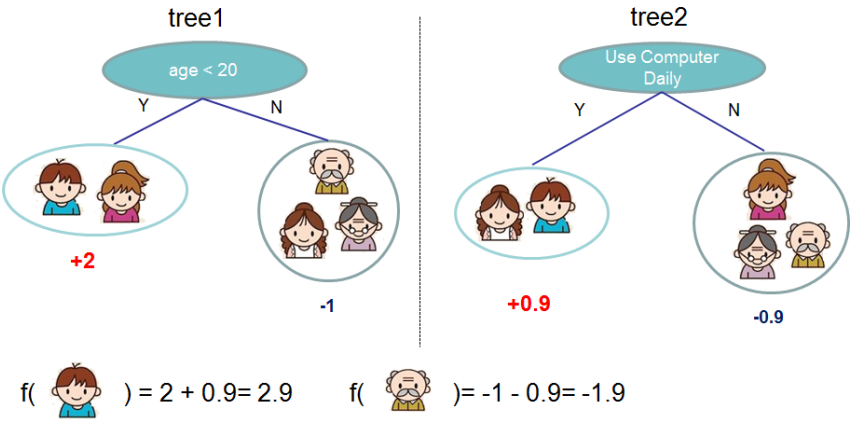


Figure 7.2: Schematic representation of a BDT. Figure taken from [122] under Creative Commons license.

The XGBoost framework provides a wide range of hyperparameters that can be adjusted to control model complexity and optimize performance for a given problem. A complete list of these hyperparameters, along with their default values and detailed descriptions, is available in the official XGBoost documentation [122]. In this Thesis, the following hyperparameters are considered during the optimization of the BDT model:

- **Number of estimators** (`n_estimators`): refers to the number of individual trees that are sequentially combined to form the final model.
- **Max depth** (`max_depth`): defines the maximum number of levels each individual decision tree can have. It controls how complex each tree is allowed to become.
- **Learning rate** (`learning_rate`): controls how much each new tree contributes to the overall model.

7.2.2.2 Training Samples

To train the BDT classifier, dedicated samples are used to model the signal and background components. For the signal, the simulated χ_{c1} and χ_{c2} samples in OO collisions listed in Table 7.2 are employed. These samples are processed through the full detector simulation and reconstruction chain, and the same pre-selection and offline selection requirements applied to data are imposed. Only candidates matched to a true χ_{c1} or χ_{c2} decay, are retained for the signal training.

For the background training, a reduced sample of pre-filtered data (first selection step) is used. This choice is motivated by the fact that the pre-filtered sample is overwhelmingly dominated by combinatorial background, with a negligible contribution from signal candidates. The signal and background samples are then combined to form the final training sample. Since the two samples contain different numbers of candidates, the BDT is

trained using event weights to compensate for this imbalance. In particular, signal candidates are assigned a weight corresponding to the ratio between the number of background and signal candidates, $N_{\text{bkg}}/N_{\text{sig}}$.

As mentioned in section 7.2.1, the training sample is split into two subsets: 70% of the original sample will be dedicated to the training itself, while the remaining 30% will be dedicated to test and validate the trained BDT model. The separation between training and test samples is extremely important to avoid bias during the training procedure.

7.2.2.3 Training Variables

Once a proper training sample is defined, the next step for the BDT training procedure is to choose an appropriate set of variables that will be the features used to train the BDT. The feature selection is a very sensitive step during the training procedure. Here, some considerations have to be taken into account to perform an accurate training. Firstly, the training variables should not be correlated with the variable of interest or under study (the χ_c mass in our case) to avoid any bias. If the BDT training includes variables carrying significant information about the invariant mass, the classifier may sculpt the mass distribution and introduce artificial structures in the spectrum after the BDT selection. The second important aspect to consider when selecting the training features is their discriminating power. The chosen variables should exhibit significantly different distributions for signal and background candidates, allowing the BDT to effectively separate the two classes. Having taken into account these two aspects, the chosen training variables are shown in Table 7.5.

Training Variables	
Kinematics	$p_T(J/\psi)$, $p_T(\mu^{\pm})$, $p_T(\gamma)$
Decay Geometry	DIRA(J/ψ) IP(J/ψ), IP($\mu^{\pm}$) Life Time (J/ψ) DOCA(J/ψ) χ^2_{DOCA} (J/ψ) χ^2_{IP} (J/ψ), χ^2_{IP} ($\mu^{\pm}$)
Photon ID	CaloNeutral1To9EnergyRatio CaloNeutral4To9EnergyRatio CaloNeutralShowerShape CaloNeutralEcalEnergy CaloTrackMatchChi2 CaloNeutralHcal2EcalEnergyRatio

Table 7.5: Variables for BDT training. The variables are divided into three categories: Kinematics, decay geometry and photon ID.

Among all the available variables, those selected for the training are the most discriminating ones. Their discriminating power was assessed using a feature-importance metric provided by the XGBoost library. In this approach, the importance of a given variable is defined as the average, over all candidates, of the total reduction in the loss function associated with the splits involving that variable across all decision trees [122].

To check the correlation among variables and the χ_c mass, as well as other variables of interest (e.g., $p_T(\chi_c)$, $\eta(\chi_c)$, the total energy deposited in the ECAL per event, E_{ECAL}), a Pearson Correlation Coefficient is used. In general, this coefficient computes the correlation between a variable X and another variable Y, and it is given by:

$$r_{XY} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2}}. \tag{7.4}$$

The computation of this coefficient for all the training variables leads to the correlation matrix shown in Figure 7.3. This correlation matrix is symmetric; its diagonal will always present the highest correlation value (correlation of the variable with itself).

The most relevant information provided by the correlation matrix is contained in its first column (or equivalently, its first row), which gives the correlation coefficients between the χ_c invariant mass and the remaining training variables. As shown, none of the variables used in the training exhibits a strong correlation with the χ_c mass. Notice that the transverse momentum and the pseudorapidity of the χ_c, as well as the deposited energy of the ECAL are not used as training variables. Furthermore, the training variables do not display significant correlations with the other observables for which potential biases should be avoided. The only notable exception is the strong correlation observed between $p_T(J/\psi)$ and $p_T(\chi_c)$, as expected, since the J/ψ carries most of the momentum of the χ_c, while the accompanying photon is typically soft.

7.2.2.4 Hyperparameters Selection

As mentioned in section 7.2.1, the hyperparameters are the settings that control how the model is built and how it learns from the training sample. They are not learned automatically during the training procedure but must be set before the learning process begins. The hyperparameters to be used, mentioned in section 7.2.2.1, are crucial because they determine the model's complexity, learning speed, and generalization ability.

When tuning the hyperparameters, the correlations between them must be taken into account. In particular, there is a trade-off between the learning rate and the number of estimators. A smaller learning rate generally improves the generalization capability of the BDT by slowing down the learning process. However, if the number of estimators is not sufficiently large, too small a learning rate may prevent the model from fully converging during training. The maximum depth controls the complexity of the model, with larger

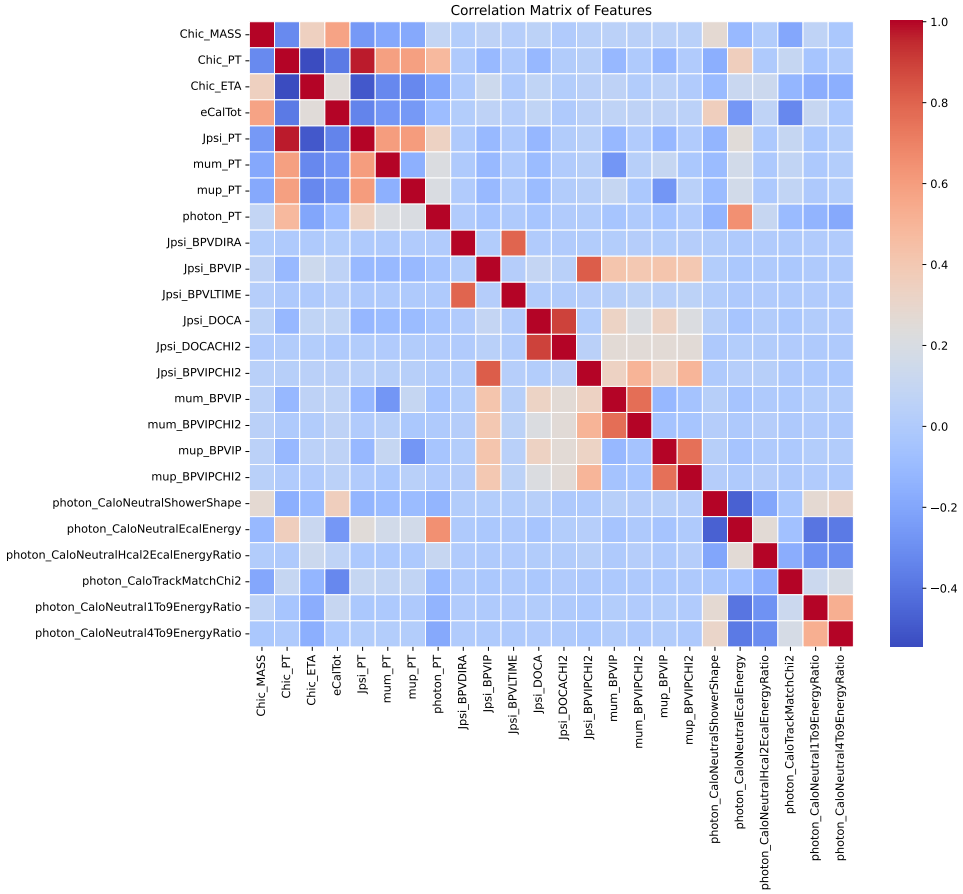


Figure 7.3: Correlation matrix for all the training variables, χ_c mass, $p_T(\chi_c)$, $\eta(\chi_c)$ and E_{ECAL}. Matrix computed using the Pearson Correlation Coefficient in the training sample (signal + background).

values allowing for more intricate decision boundaries. Nevertheless, excessively large values of this hyperparameter may lead to over-training of the BDT.

As the hyperparameters are an external selection from the training procedure, it is crucial to choose the optimal set of values for them before the training. For that purpose, several models with different hyperparameters are compared. The features used to determine which is the best of the models are the Receiver Operator Characteristic (ROC) curve and the Area Under the Curve (AUC). The ROC curve shows how well the model can distinguish between signal and background, while the AUC summarizes this performance information in a single number, with 1 being a perfect classification and 0.5 being almost a random classification. A summary of the models and their hyperparameters together with their AUC associated value is shown in Table 7.6.

The hyperparameters of the models considered in the comparison were chosen according

Model	Number of Estimators	Max Depth	Learning Rate	AUC
Model 1	200	5	0.2	0.953479
Model 2	400	6	0.15	0.96513
Model 3	600	7	0.1	0.972705
Model 4	900	8	0.05	0.975914

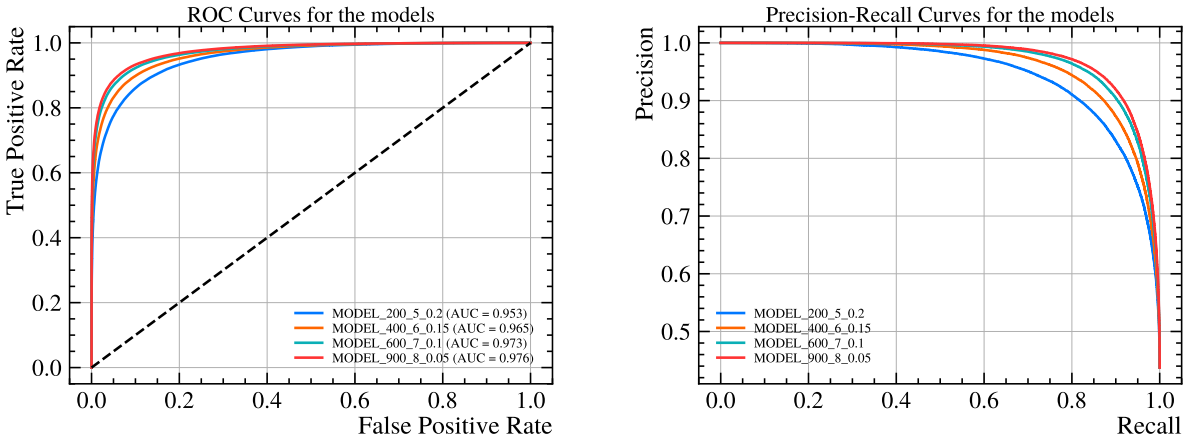
Table 7.6: Different models for the BDT depending on the values of their hyperparameters and their AUC value.

to the considerations discussed above. The first model is trained with a relatively small number of estimators and a correspondingly large learning rate. Successive models are then obtained by increasing the number of estimators while decreasing the learning rate. Regarding the maximum depth, the first model employs a moderately low value, which is progressively increased for the subsequent models in order to study the effect of a higher model complexity. The largest value considered for the maximum depth corresponds to the threshold beyond which the risk of overfitting becomes significant, which is why no larger values were investigated. Furthermore, the AUC values reported in Table 7.6 are observed to converge as the number of estimators increases. For this reason, models with a larger number of estimators were not considered.

Figure 7.4 (left) displays a comparison between all the models considered in Table 7.6. Based on the AUC value, Model 4 seems to be the best, as it has the highest value among all the considered models. However, the ROC curve could give an overly optimistic view when the training samples are imbalanced, as it is the case here. Therefore, it is complemented by the precision-recall curve shown on the right-hand side of Figure 7.4, which focuses on the positive class (signal) and provides a clearer assessment of the model's ability to identify signal candidates. The precision quantifies the fraction of candidates identified as signal that are truly signal, while the recall (or sensitivity) measures the fraction of true signal candidates correctly identified by the model. From these results, it can be concluded that Model 4 still provides the best overall performance. Consequently, this model is selected for the remainder of the analysis.

7.2.2.5 BDT model validation

The validation of the algorithm performance is a crucial step in the ML workflow. To perform this validation, the validation sample reserved for this purpose is used. Once the optimal hyperparameters are chosen, one can study the performance of the BDT model for this concrete set of hyperparameter values. The first thing that can be checked is the Confusion Matrix. This matrix is a table used to evaluate the performance of the BDT classification model by comparing the predicted class labels with the true class labels. Here, there are four possible cases of classification: Correctly predicted signal (true positive), wrongly predicted signal (false positive), correctly predicted background



(a) ROC curves.

(b) Precision-recall curves.

Figure 7.4: Comparison between the different BDT models depending on their hyperparameters. Comparison between the ROC curves for each model (a) and comparison of the precision-recall curves for each model (b).

(true negative), and wrongly predicted background (false negative). The goal of the BDT training is to maximize the number of correctly classified candidates while minimizing the number of misclassified ones. The confusion matrix corresponding to the BDT model selected in this analysis is shown in Fig. 7.5. As can be observed, the matrix is predominantly populated along its diagonal, which contains the true positive and true negative classifications. This indicates that the BDT efficiently discriminates between signal and background candidates, resulting in a very low probability of misidentifying signal as background, or vice versa.

Another useful validation consists in studying the distribution of the candidates classified by the BDT as a function of the χ_c invariant mass, which is the variable of interest in this classification. The fractions of the four classification categories as a function of the χ_c mass are shown in Fig. 7.6. In each mass bin, the number of candidates in a given category is normalized to the total number of candidates in that bin.

As expected, the fraction of true positive candidates peaks around the nominal masses of the χ_{c1} and χ_{c2} states, while the fraction of true negative candidates decreases in the same region. In contrast, the fractions of false positive and false negative candidates remain small and approximately constant over the entire mass spectrum. However, the small variation of the fractions of false positive and false negative candidates could be considered as a systematic uncertainty in the future. The observed behavior indicates that the trained BDT model performs as intended, efficiently identifying signal candidates in the mass regions where the χ_c resonances are expected.

The final aspect to be investigated is the evolution of the training throughout the suc-

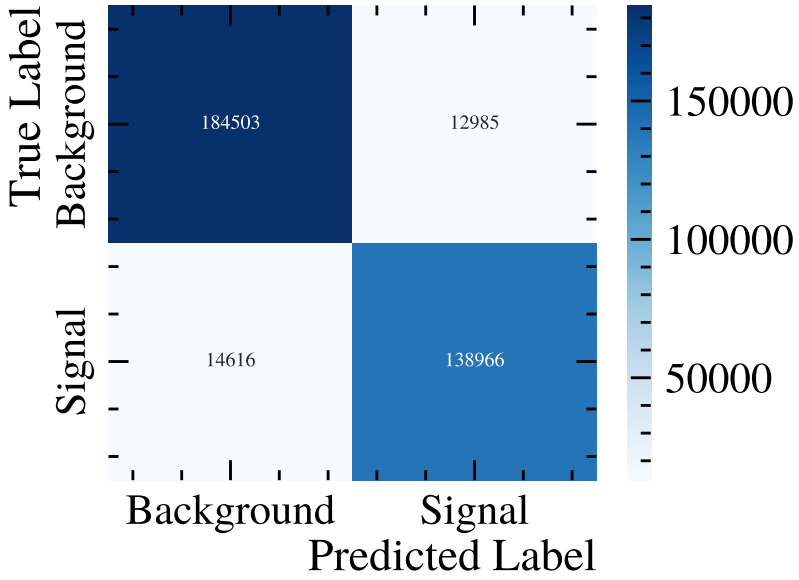


Figure 7.5: Confusion Matrix for the considered BDT model in this analysis.

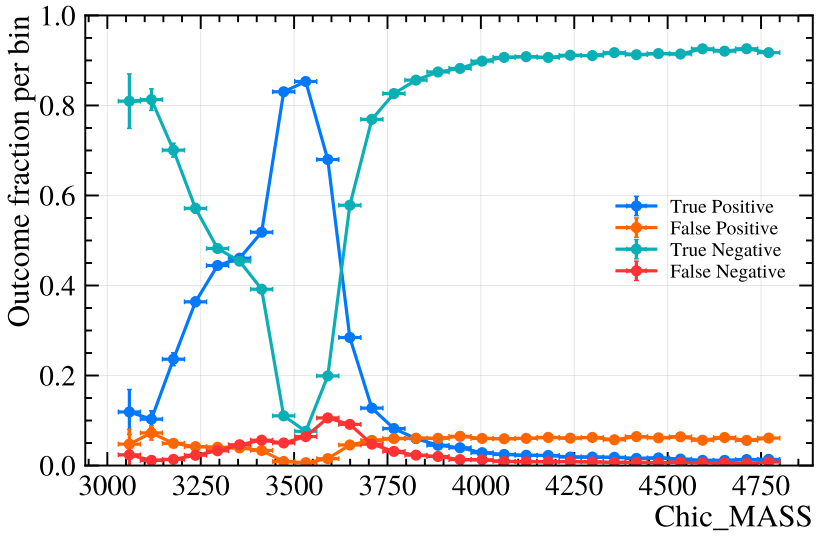


Figure 7.6: Fraction of classified candidates as a function of χ_c mass for the considered BDT model in this analysis.

cessive boosting iterations. To this end, a suitable loss function must be chosen, as introduced in Eq. 7.3. In this analysis, the performance of the model is evaluated using the logarithmic loss, commonly referred to as *Log Loss*. Following the notation introduced

in Section 7.2.1, the Log Loss function is defined as

$$L(\theta) = \sum_i [y_i \ln(1 + e^{-\hat{y}_i}) + (1 - y_i) \ln(1 + e^{\hat{y}_i})]. \tag{7.5}$$

Figure 7.7 displays the Log Loss applied to the training and validation samples. In the case of Log Loss, lower values of this metric are the result of a low error, i.e., a small difference between the predicted and the actual true values of the target variables. As we can see in Fig. 7.7, both train and validation curves start at a relatively high value of Log Loss. However, as the number of iterations increases, we see that both curves decrease rapidly. This sharp drop at the beginning means that the BDT model learns the most important structure in the early stages of the training. Afterward, the improvements become smaller, and the optimization enters a refinement phase, where the curves flatten gradually. Another important thing to notice here is that both curves follow the same decreasing trend, which is a sign of healthy training with no overfitting. This also reflects that the BDT model retains the generalization capability. The fact that the validation curve is above the training one is something expected, as the validation sample has been unseen by the model, giving more optimistic results when the BDT model is applied to the training sample. The difference between the train and validation curve remains very small during the training, which is another positive signature of a well-trained BDT model.

With all the proves exposed above, we can conclude that the BDT model has been trained successfully, unbiased and with a high generalization capability, what allow the final trained BDT model to separate between signal and background efficiently.

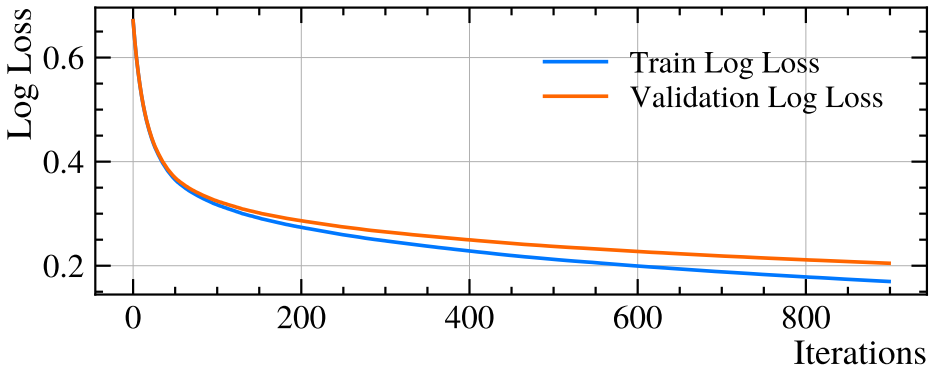


Figure 7.7: Log Loss functions as a function of the iterations of the training. In blue, the loss computed for the training sample and in orange, the loss computed for the validation sample.

7.2.2.6 Predicted probabilities from the BDT

At this stage, all the necessary steps for sample and feature selection, hyperparameter tuning, and validation of the BDT model have been completed. The trained BDT can

then be applied candidate by candidate, providing a prediction based on the values of the input features. Since the BDT was trained as a binary classifier to distinguish signal from background, its output can be interpreted as the probability that a given candidate belongs to the signal class.

The distributions of the predicted probabilities for signal and background candidates in the training, testing, and combined samples are shown in Fig. 7.8. A clear separation between the two classes can be observed: background candidates are predominantly assigned low predicted probabilities, whereas signal candidates are concentrated at high values. This demonstrates the good discriminating power of the trained BDT model.

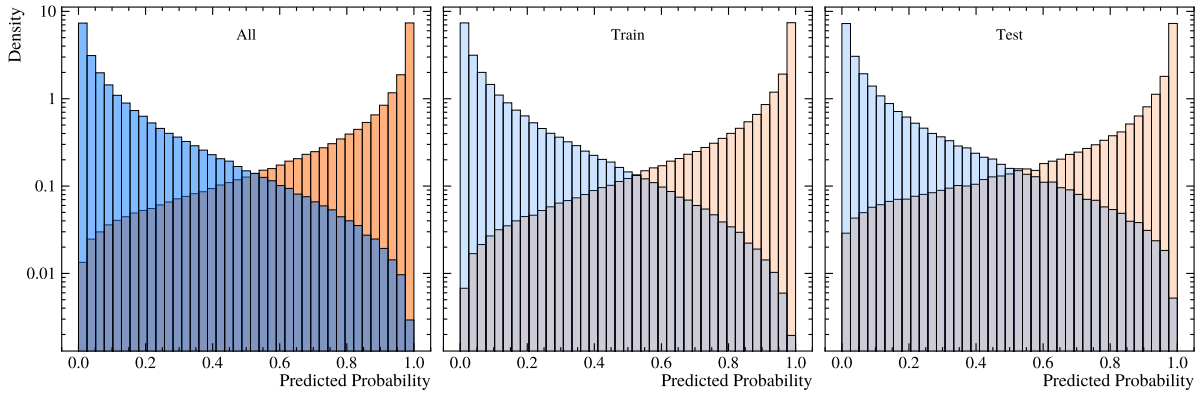


Figure 7.8: Distributions of the BDT-predicted probabilities for signal (orange) and background (blue) candidates. From left to right, the distributions are shown for the full sample, the training sample, and the testing sample.

The BDT model is going to retrieve a predicted probability value, candidate by candidate, once the model is applied to our data sample. In view of the later filtering of the data, we need to choose a threshold from which we accept each candidate as a signal, or we reject it as background. In this sense, we have to take into account the existing tradeoff between the chosen threshold and the available statistics in our data sample. Tighter cuts on the predicted probability lead to a higher signal-to-background ratio but considerably reduce the available statistics, whereas looser cuts retain more statistics at the expense of a lower signal-to-background ratio. In order to address this issue, a Figure of Merit (FoM) is used, taking the significance as a metric. The significance of the FoM is defined as

$$\mathrm{FoM} = \frac{S}{\sqrt{S+B}}, \tag{7.6}$$

where S denotes the number of signal candidates and B the number of background candidates, both obtained from the dedicated test and validation sample. The resulting figure

of merit (FoM), computed using the validation sample, is shown in Fig. 7.9. The optimal threshold on the predicted probability is defined as the value that maximizes the significance. In this case, the maximum is reached for a threshold of 0.5. This value provides the best compromise between signal purity and statistical significance, and is therefore chosen as the working point for the selection of the data sample.

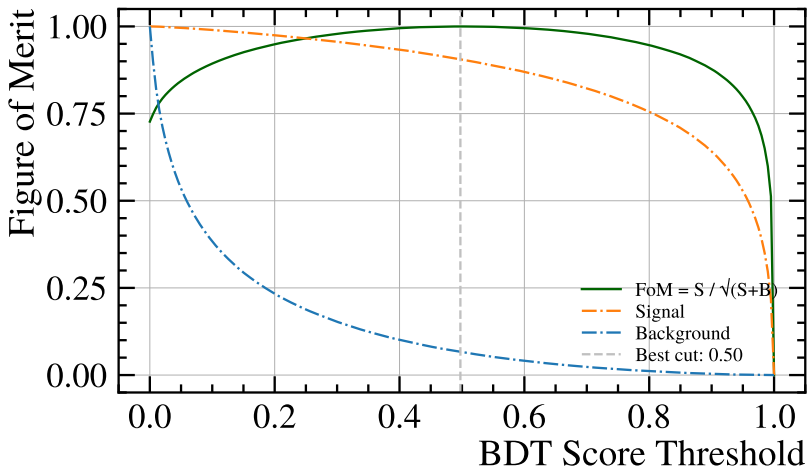


Figure 7.9: Figure of Merits using the significance as metric (green solid line) as a function of the predicted probability. The number of signal candidates (dashed orange line) and the number of background candidates (dashed blue line), both normalized to their integral, are also shown.

7.2.2.7 Data filtering with the BDT

The last step of this procedure is to apply the trained BDT model to our dataset. Candidate by candidate, the BDT assign a score to each candidate based on their feature variables. If the candidate's score surpasses the threshold discussed in the previous section, then the candidate is going to be selected as a signal.

Candidates who pass the pre-selection and offline selection requirements are filtered using the trained BDT model. A comparison between the candidates classified as signal and those classified as background is shown in Fig. 7.10. As can be observed, the candidates identified as background are predominantly located in the region of high invariant mass and low transverse momentum, and tend to exhibit slightly larger pseudorapidities than the candidates identified as signal. In the invariant mass distribution, an enhancement can already be observed in the mass region corresponding to the χ_{c1} and χ_{c2} states, indicating the emergence of a signal. This motivates the signal extraction procedure described in the following section.

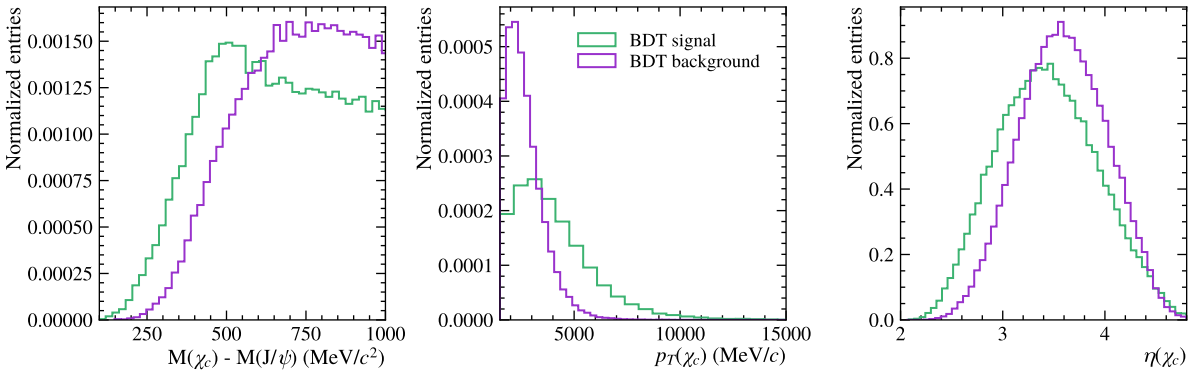


Figure 7.10: Normalized number of candidates filtered by the BDT selected as signal (green) and selected as background (purple) using 2025 OO data sample. From left to right, the invariant mass, transverse momentum, and pseudorapidity distributions of χ_c candidates.

7.3 Signal Extraction

7.3.1 Combinatorial Background Description

After applying the BDT selection, a residual background contribution remains and must be accounted for during the signal extraction procedure. This background does not arise from any physical process of interest, but rather from random combinations of the decay products considered in the reconstruction of the parent particle. Such a contribution is referred to as *combinatorial background*. Its importance increases with the event multiplicity and with the number of particles in the final state under study.

In the present analysis, the final state is composed of $\mu^+\mu^-\gamma$ candidates, implying that the number of possible combinatorial combinations scales approximately as $N_{\mu^+}N_{\mu^-}N_{\gamma}$. This contribution can nevertheless be significantly reduced by applying additional constraints, such as a tight selection around the J/ψ mass, which suppresses combinations of uncorrelated muons, or a higher transverse momentum requirement on the photon candidate, which rejects low-energy prompt photons and photons originating from π^0 decays.

A widely used approach to model the combinatorial background, and the one adopted throughout this analysis, is the **Mixed-Event Technique** [125, 126], commonly referred to as *event mixing*. The underlying principle of this method is straightforward: if the combinatorial background originates from random combinations of uncorrelated particles, an estimate of this background can be obtained by combining particles from different physical events. Since particles produced in different collisions cannot originate from the same decay, any genuine resonance contribution is naturally removed. The resulting sample therefore, provides an empirical description of the uncorrelated combinatorial background.

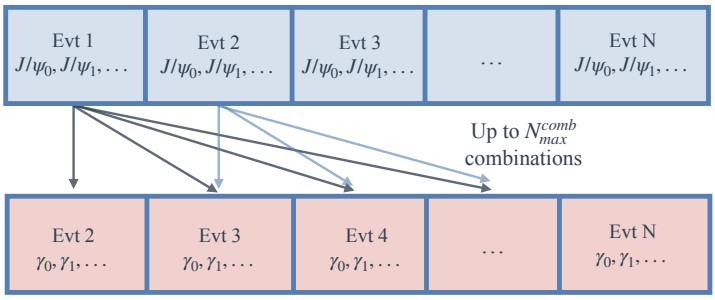


Figure 7.11: Schematic view of the event mixing algorithm.

How to put this event mixing technique into practice depend on the analysis itself. For simplicity, we are going to consider two-particle combinations. Due to the tight cut around the J/ψ mass, the random combinations of muons are negligible compared to the photons contribution to the combinatorial background. Therefore, in this case, the event mixing will be carried out by uncorrelated combinations of dimuons with photons. A diagram showing how the algorithm works can be seen in Fig. 7.11. To ensure the absence of correlation between the mixed particles and the proper operation of the event mixing algorithm some conditions are required:

- Particles to be combined must not belong to the same event. This condition is implemented requiring that the `RUNNUMBER` or the `EVENTNUMBER` of the events of the combination candidate have to be different.

- Particles to be combined have to have similar z position of their associated primary vertex. This condition requires $|PVZ_{J/\psi} - PVZ_{\gamma}|$ to be lower than a certain value $|\Delta PVZ|_{\text{max}}$. This ensure that the events that are being combined have similar initial conditions, what leads in a more accurate reproduction of the combinatorial background.

- The events of the particles to be combined have to have similar multiplicity. This condition is imposed with the same purpose as the previous one. Here, the number of tracks in the VELO, `nVeloTracks`, is used as multiplicity estimator. The `nVeloTracks` distribution is divided in four regions of multiplicity that contain the 25% of the total candidates each one, as shown in Fig. 7.12. Only particles that belongs to the same region of multiplicity can be combined.

If the data sample is large, doing all the possible combinations still take a long computational time. To speed up the process, the event mixing algorithm takes a number of

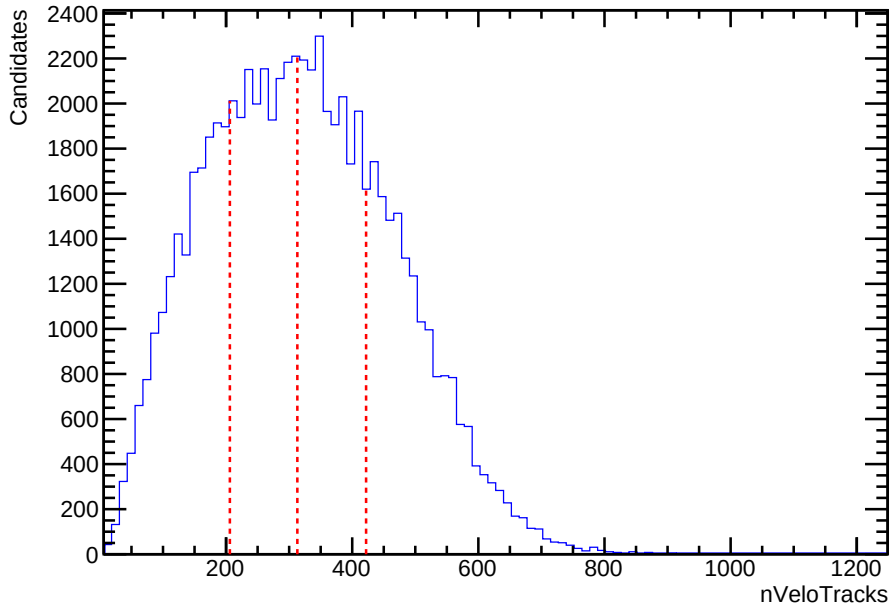


Figure 7.12: Multiplicity distribution of the number of tracks in the VELO for 2025 OO data sample. The vertical dotted red lines separate the four multiplicity regions, containing the 25% of the total candidates each one.

candidates $N_{\text{comb}}^{\text{max}}$ (bunch of candidates) and make all the allowed combinations according with the requirements above. Once all the possible combinations are done in this subset of candidates, the processed candidates are discarded and the algorithm takes a new subset of $N_{\text{comb}}^{\text{max}}$ candidates, iterating this process until reach the total number of candidates of the data sample.

In order to take into account the systematics uncertainties introduced in the signal extraction process several conditions of $|\Delta PVZ|_{\text{max}}$ and $N_{\text{comb}}^{\text{max}}$ where tested. A list of the used conditions is listed in Table 7.7. All the combinations of these values are considered subsequently.

$\|\Delta PVZ\|_{\text{max}}$ (mm)	$N_{\text{comb}}^{\text{max}}$
80	5000
50	2500

Table 7.7: List of values considered in the event mixing algorithm.

The shape of the resulting invariant mass distribution from the event mixing procedure will be used as a Probability Density Function (PDF) in the fitting process to model the combinatorial background. Only the normalization of this PDF will remain free and will

be determined in the fit. This normalization coincide with the number of background candidates N_{bkg}.

7.3.2 Modeling the Signal Shape

With the aim of modeling the shape of the signal, first, we are going to fit the signal from truth reconstructed MC candidates in order to constrain the fit parameters. The signal is described by a generalized Asymmetrical Double-Sided Crystal Ball (CB) function [127] which contains a gaussian core with asymmetric tails to take into account the radiative energy loss. The generalized asymmetrical double-sided CB is defined as:

$$f(m; m_0, \sigma_L, \alpha_L, n_L, \sigma_R, \alpha_R, n_R) = \begin{cases} A_L \cdot \left(B_L - \dfrac{m - m_0}{\sigma_L} \right)^{-n_L}, & \text{for } \dfrac{m - m_0}{\sigma_L} < -\alpha_L \\ \exp\left(-\dfrac{1}{2} \left[\dfrac{m - m_0}{\sigma_L} \right]^2 \right), & \text{for } \dfrac{m - m_0}{\sigma_L} \leq 0 \\ \exp\left(-\dfrac{1}{2} \left[\dfrac{m - m_0}{\sigma_R} \right]^2 \right), & \text{for } \dfrac{m - m_0}{\sigma_R} \leq \alpha_R \\ A_R \cdot \left(B_R + \dfrac{m - m_0}{\sigma_R} \right)^{-n_R}, & \text{otherwise} \end{cases} \tag{7.7}$$

where the A_i and B_i factors, with $i \in \{L, R\}$ referring to the left and right tail respectively, are defined as

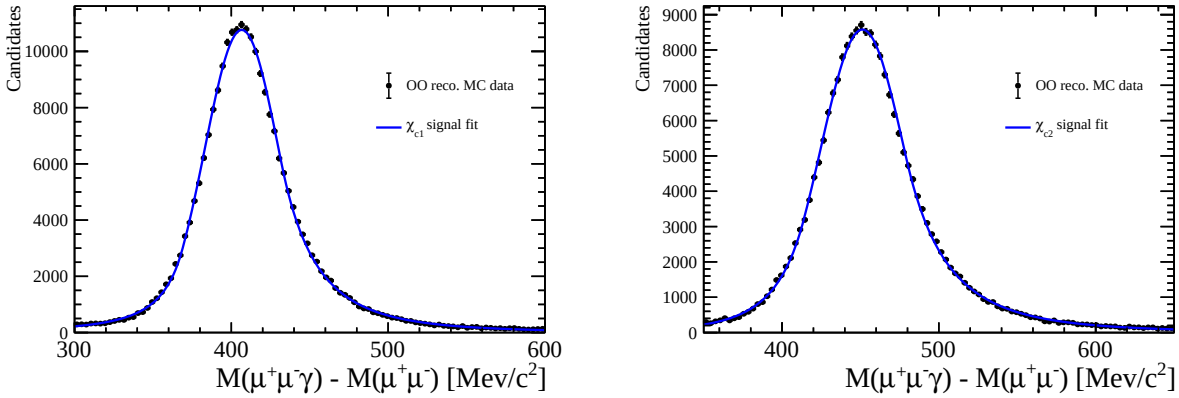
$$A_i = \left(\frac{n_i}{|\alpha_i|} \right)^{n_i} \cdot \exp\left(-\frac{|\alpha_i|^2}{2} \right) \tag{7.8}$$

$$B_i = \frac{n_i}{|\alpha_i|} - |\alpha_i| \tag{7.9}$$

The parameters $\alpha_L, n_L, \alpha_R, n_R$ will modulate the shape of the left and right tail of the distribution, while m_0 will be the mean of the distribution that coincide with the mass nominal value of the particle under study. In this case $\sigma_L = \sigma_R = \sigma$ will be the width, or the resolution, of the distribution.

In Fig. 7.13 is shown the fits to the invariant mass in reconstructed MC for both χ_{c1} and χ_{c2}. The results from the fits of the parameters for both CB are listed in Table 7.8. In order to constrain as much as possible the fit in data, the parameters σ, α_L, n_L, α_R and n_R will be fixed to the values obtained from the fit to the MC reconstructed data.

As in the case of the combinatorial background, a Gaussian shape will be tested also in order to take into account the systematic uncertainties coming from the signal extraction



(a) Fit for χ_{c1} signal.

(b) Fit for χ_{c2} signal.

Figure 7.13: Fits to the invariant mass distribution in reconstructed OO collisions MC simulation for χ_{c1} (a) and χ_{c2} (b) signal.

procedure. The σ of the gaussian distribution will be fixed to the same value obtained from CB fit to the MC reconstructed data listed in Table 7.8. For both CB and gaussian models, the normalization will remain free and to be determined in the fit to the data, what will give us the yields $N_{\chi_{c1}}$ and $N_{\chi_{c2}}$.

Fit parameter	χ_{c1}	χ_{c2}
χ^2/NDOF	3.7	2.8
m_0 (MeV/c^2)	406.6 $\pm$ 0.1	451.2 $\pm$ 0.1
σ (MeV/c^2)	23.5 $\pm$ 0.1	26.6 $\pm$ 0.2
α_L	1.61 $\pm$ 0.02	1.41 $\pm$ 0.03
n_L	2.3 $\pm$ 0.2	5.8$\pm$0.7
α_R	1.08 $\pm$ 0.01	1.05 $\pm$ 0.01
n_R	3.8 $\pm$ 0.1	4.1 $\pm$ 0.1

Table 7.8: List of the fit results for the parameter of the CB model for χ_{c1} and χ_{c2}.

7.3.3 Application of the photon energy calibration

As described in chapter 6 the OO data need to be recalibrated, so in this section the results of the mass correction obtained before are applied. This calibration will not recompute all the variables related with the photon energy, but it will correct the invariant mass spectrum to ensure us that what we are measuring actually coincide with the nominal masses of χ_{c1} and χ_{c2}. This calibration will be applied to the last step of the data filtering, that is to say, after the BDT filter. In the same way, the calibration has to be applied to the results of the event mixing too.

The strategy to correct the invariant mass will be as follows: the results from the calibration table presented in Fig. 6.9 will be applied candidate by candidate to the BDT filtered data and the event mixing output as

$$\Delta m^{\text{corr}} = [m(\mu\mu\gamma) - m(\mu\mu)] \cdot \lambda(E^{\gamma}, \text{reg}), \tag{7.10}$$

or equivalently

$$\begin{aligned} m^{\text{corr}}(\mu\mu\gamma) &= \Delta m^{\text{corr}} + m(\mu\mu) \\ &= m(\mu\mu\gamma) \cdot \lambda(E^{\gamma}, \text{reg}) + m(\mu\mu)[1 - \lambda(E^{\gamma}, \text{reg})]. \end{aligned} \tag{7.11}$$

The corrected invariant mass spectrum for the BDT filtered data and for the event mixing output is shown in Fig. 7.14. In this figure, we can see that the invariant mass distributions are shifted to lower mass values after the calibration. This is something that we expected, as the mass correction factors applied were equal or lower to the unity. Now, in the corrected invariant mass spectrum the peak of the distribution is located around the nominal mass of χ_{c1} and χ_{c2}. The fit to the invariant mass spectrum will be done in these corrected samples.

7.3.4 Fit to the invariant mass spectrum

The yields of χ_{c1} and χ_{c2} are extracted from fits to the corrected invariant mass spectrum integrated over transverse momentum, pseudorapidity, and event multiplicity. The signal and background models introduced in the previous sections are employed, with additional constraints imposed in the fitting procedure.

The shape of the combinatorial background is fixed using the event-mixing technique for different choices of $|\Delta PVZ|_{\text{max}}$ and $N_{\text{comb}}^{\text{max}}$, while its normalization is left free and determined by the fit. For the signal component, two models are considered. The first model employs a Crystal Ball function, whose tail parameters and width are fixed to the values obtained from fits to the simulated samples listed in Table 7.8. The second model uses a Gaussian function with its width fixed to the same value as in the Crystal Ball model. In both cases, the mean position and normalization of the χ_{cJ} signals are allowed to vary freely in the fit.

In addition, the mean position of the χ_{c2} signal is constrained by fixing the mass difference with respect to χ_{c1} to the value reported by the PDG,

$$m(\chi_{c2}) = m(\chi_{c1}) + (45.5 \pm 0.1)\ \text{MeV}/c^2. \tag{7.12}$$

The different fits performed are shown in Figures 7.15 and 7.16. The fit range is chosen according to the phase-space limitations of the selected χ_c candidates arising from the kinematic requirements applied in the analysis.

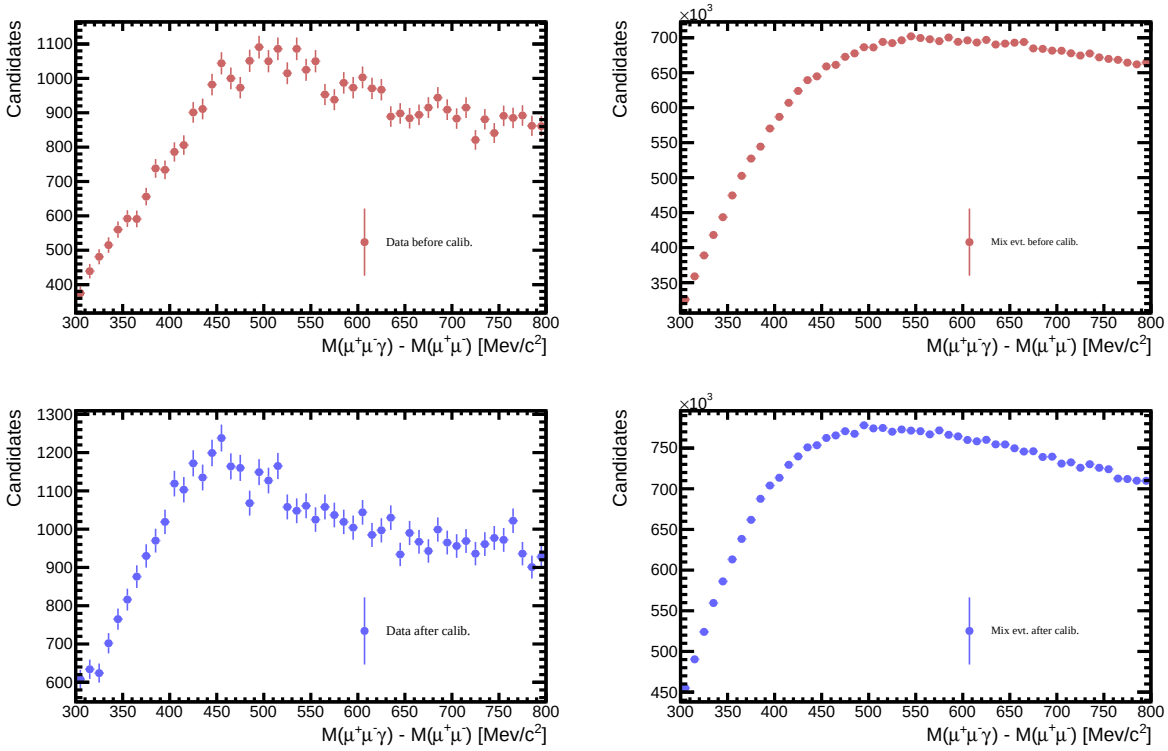


Figure 7.14: Corrected $\mu\mu\gamma - \mu\mu$ invariant mass spectrum in the 2025 OO data sample. Red points represent data (or mixed events) before calibration and blue points after the calibration.

The results of the various fits are summarized in Table 7.9. All fits provide a satisfactory description of the data, with $\chi^2/\text{NDOF} < 2$. Furthermore, the fitted χ_{c1} mass is found to be compatible within at least 2σ with its nominal value in all cases. The extracted χ_{c1} yields remain stable among the different models considered, whereas the χ_{c2} yields exhibit a stronger dependence on the assumed signal parameterization.

As discussed previously, the spread of the results obtained with the different signal and background models is assigned as a systematic uncertainty associated with the signal extraction procedure. To estimate this uncertainty, a weighted average is computed using the fit results obtained with the different models. The weights w_i are defined from the uncertainties σ_i associated with each measurement as

$$w_i = \frac{1}{\sigma_i^2}. \tag{7.13}$$

The weighted mean $\bar{x}$ is then calculated as

Models	χ^2/NDOF	$m_0(\chi_{c1})$	$N_{\chi_{c1}}$	$N_{\chi_{c2}}$	N_{bkg}
CB-5000-80	1.13	416.5 ± 8.7	931 ± 295	1225 ± 298	47309 ± 266
CB-5000-50	1.15	416.0 ± 9.3	922 ± 314	1241 ± 315	47301 ± 266
CB-2500-80	1.13	417.7 ± 9.1	945 ± 299	1176 ± 307	47343 ± 266
CB-2500-50	1.14	416.7 ± 9.0	931 ± 301	1207 ± 305	47326 ± 266
GS-5000-80	1.24	422.6 ± 6.2	907 ± 168	912 ± 182	47647 ± 256
GS-5000-50	1.26	422.2 ± 6.3	903 ± 170	923 ± 184	47640 ± 255
GS-2500-80	1.23	423.5 ± 6.2	906 ± 165	883 ± 180	47677 ± 255
GS-2500-50	1.25	422.7 ± 6.3	903 ± 169	900 ± 184	47664 ± 255

Table 7.9: Fit results for the different models of signal and background. The nomenclature of the models follows: Crystal Ball (CB) or Gaussian (GS) signal shape, $N_{\text{comb}}^{\text{max}} \in \{2500, 5000\}$ and $|\Delta PVZ|_{\text{max}} \in \{50, 80\}$ mm.

$$\bar{x} = \frac{\sum_i w_i x_i}{\sum_i w_i}, \tag{7.14}$$

where the weighted uncertainty associated with this mean is

$$\sigma_{\bar{x}} = \sqrt{\frac{1}{\sum_i w_i}}. \tag{7.15}$$

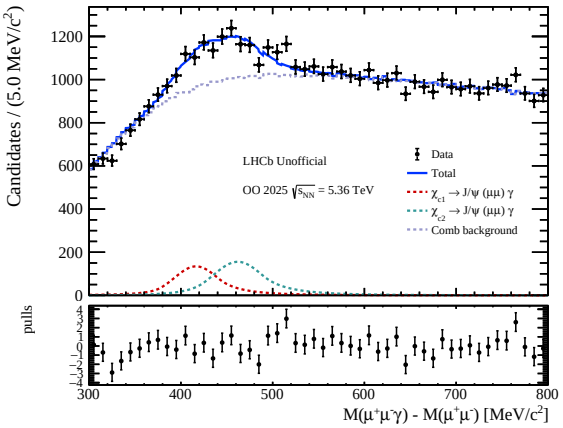
The systematic uncertainty is estimated from the weighted root mean square (RMS),

$$\text{RMS} = \sqrt{\frac{\sum_i w_i (x_i - \bar{x})^2}{\sum_i w_i}} \tag{7.16}$$

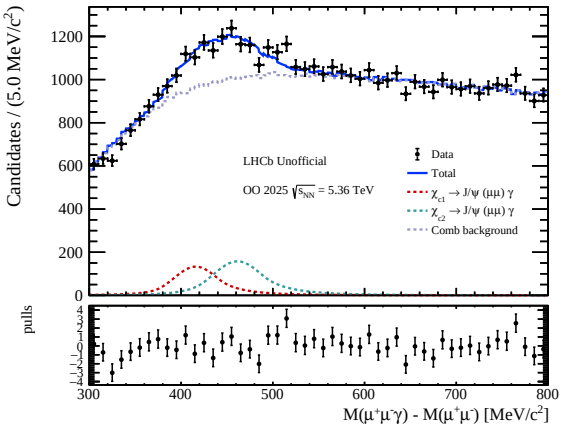
The results of this procedure are presented in Fig. 7.17. The extracted values of $N(\chi_{c1})$, $N(\chi_{c2})$, and their ratio, together with their statistical and systematic uncertainties, are summarized in Eq. 7.17.

$$\begin{aligned} N(\chi_{c1}) &= 911 \quad \pm 73\ (8.01\%\ \text{stat.}) \quad \pm 12\ (1.32\%\ \text{syst.}) \\ N(\chi_{c2}) &= 985 \quad \pm 78\ (7.92\%\ \text{stat.}) \quad \pm 136\ (13.81\%\ \text{syst.}) \\ N(\chi_{c2})/N(\chi_{c1}) &= 1.06 \quad \pm 0.12\ (11.32\%\ \text{stat.}) \quad \pm 0.12\ (11.32\%\ \text{syst.}) \end{aligned} \tag{7.17}$$

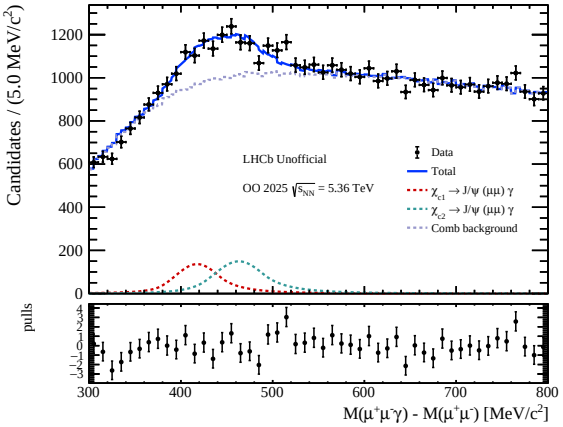
The significances of both signals exceed 5σ, establishing the first observation of χ_{c1} and χ_{c2} production in OO collisions and, more generally, in nucleus-nucleus collisions at LHC energies. These measurements open the possibility of studying quantities such as the production cross-section ratio $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ and the fraction of J/ψ mesons originating from feed-down decays of χ_c states.



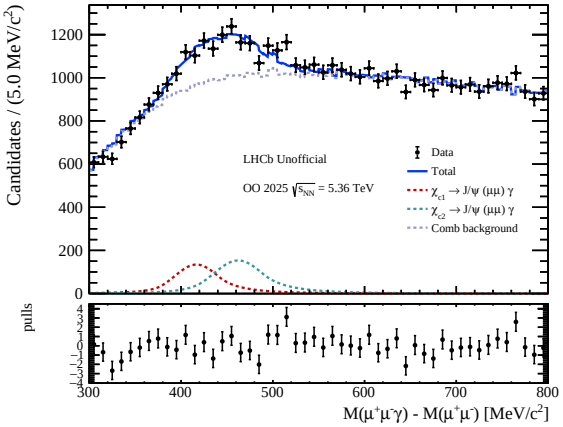
(a) $N^{\max}_{\text{comb}} = 5000$, $|\Delta PVZ|_{\max} = 80$ mm



(b) $N^{\max}_{\text{comb}} = 5000$, $|\Delta PVZ|_{\max} = 50$ mm

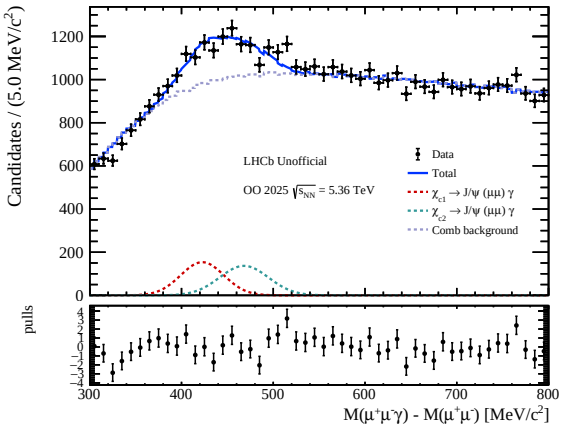


(c) $N^{\max}_{\text{comb}} = 2500$, $|\Delta PVZ|_{\max} = 80$ mm

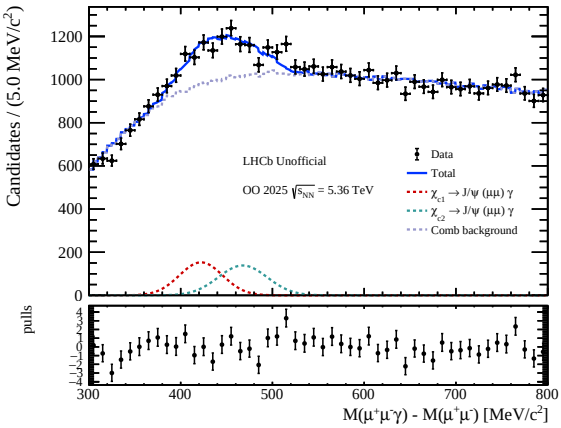


(d) $N^{\max}_{\text{comb}} = 2500$, $|\Delta PVZ|_{\max} = 50$ mm

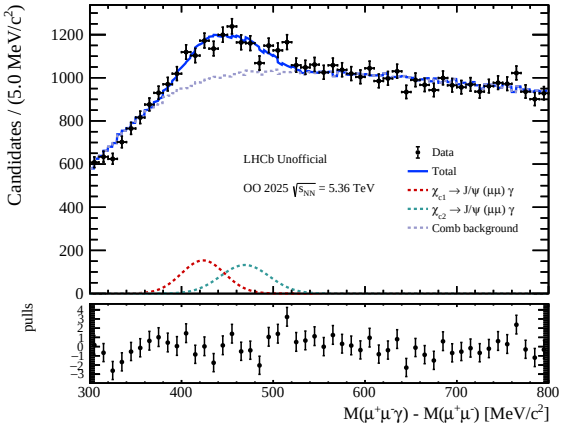
Figure 7.15: Fits to the $\mu\mu\gamma - \mu\mu$ invariant-mass distribution in the 2025 OO data sample, using a CB model for the χ_c signal and different parameters for the combinatorial background.



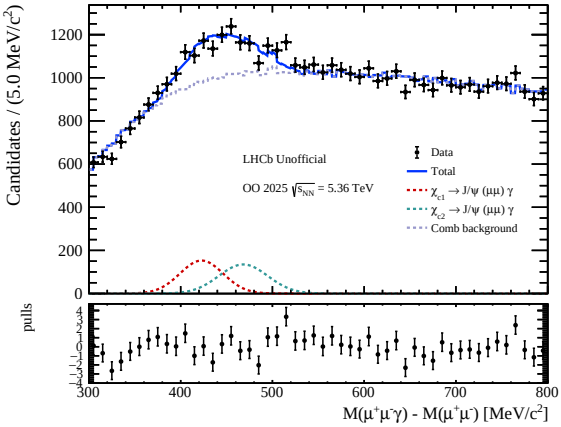
(a) $N_{\text{comb}}^{\max} = 5000$, $|\Delta PVZ|_{\max} = 80$ mm



(b) $N_{\text{comb}}^{\max} = 5000$, $|\Delta PVZ|_{\max} = 50$ mm



(c) $N_{\text{comb}}^{\max} = 2500$, $|\Delta PVZ|_{\max} = 80$ mm



(d) $N_{\text{comb}}^{\max} = 2500$, $|\Delta PVZ|_{\max} = 50$ mm

Figure 7.16: Fits to the $\mu\mu\gamma - \mu\mu$ invariant-mass distribution in the 2025 OO data sample, using a Gaussian model for the χ_c signal and different parameters for the combinatorial background.

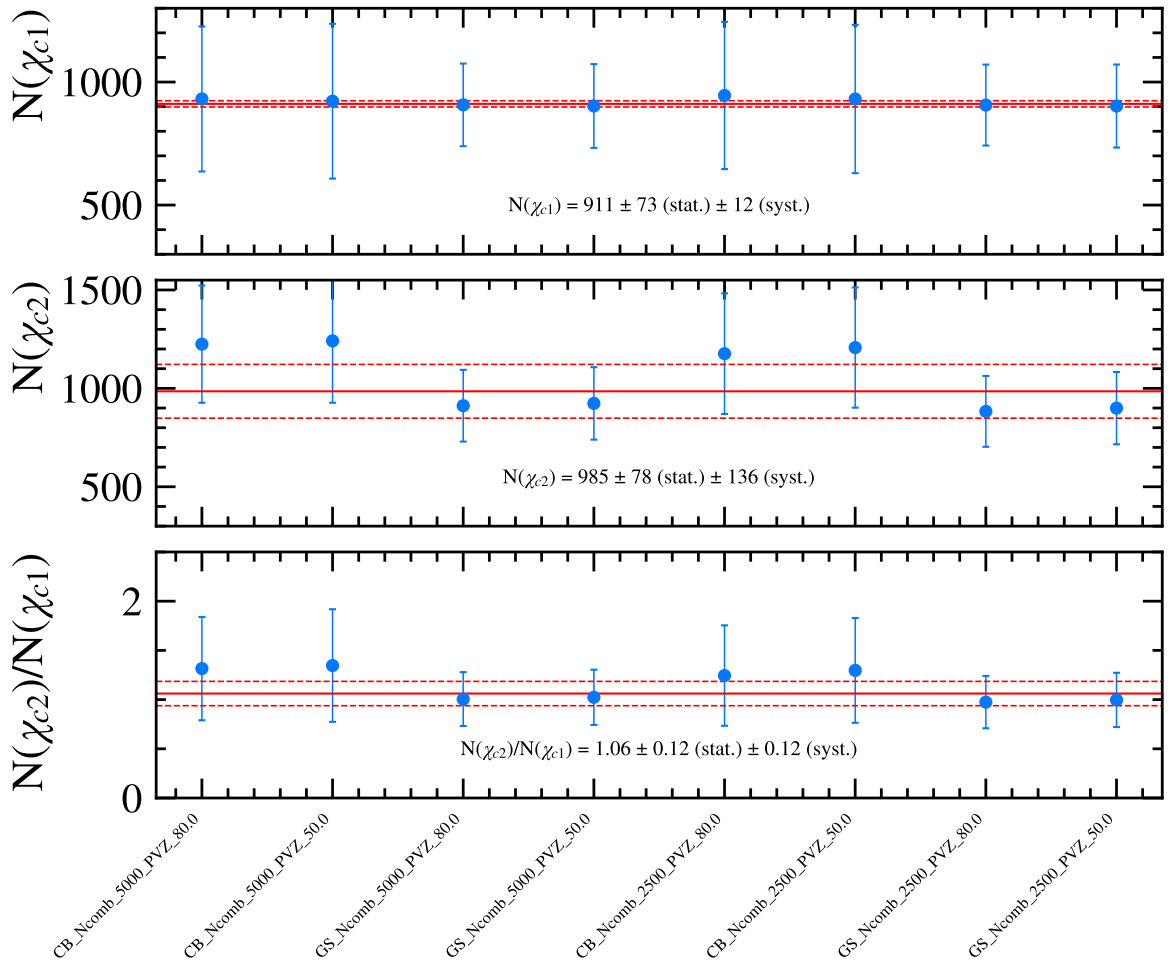


Figure 7.17: Spread of results of $N(\chi_{c1})$, $N(\chi_{c2})$ and $N(\chi_{c2})/N(\chi_{c1})$ in OO collisions. The solid red line represents the weighted mean, the red dashed lines show the RMS, which is taken as the systematic uncertainty.

7.4 J/ψ production in OO collisions

One of the goals of this analysis is to disentangle the prompt production of J/ψ mesons from the contribution arising from feed-down decays of higher charmonium states, in particular the χ_c states. To this end, prompt J/ψ mesons are measured under the same kinematic conditions as the reconstructed χ_c candidates. The decay channel considered is

$$J/\psi \to \mu^+\mu^-. \tag{7.18}$$

The data sample is the same as in the study of χ_c described in section 7.1.1 and the selection of J/ψ candidates and muons is the same applied in the pre-selection and the offline selection listed in tables 7.3 and 7.4. Thanks to the two-body final state, compared to the three-body decay of χ_c, and since there is no need to take into account the photon ID and the resolution of the ECAL, the measurement of J/ψ is much more straightforward. Here, the signal appear clearly uniquely applying the cuts mentioned before. Therefore, the main difference with the χ_c analysis is that no BDT selection is done in the case of the J/ψ measurement. This difference will be taken into account in later steps.

In the dimuon mass spectrum the dominant source of background is again the combinatorial background, in this case coming from random combinations of muons with opposite charge. In order to model or subtract this background, several approaches are possible, as already mentioned event mixing method or the same-sign method, where muons with the same charge in the same event are combined to describe the uncorrelated background. For simplicity, in this analysis, we are going to model the dimuon combinatorial background with an analytical function, since the observed trend of this background is very smooth. As done in other J/ψ analysis, the background will be modeled with an exponential function $\propto e^{\tau \cdot m}$. In the case of the signal, we are going to use a generalized Asymmetrical Double-Sided Crystal Ball function defined in equation 7.7 to take into account the radiative energy loss, as in the case of χ_c signal. Since no photons are involved in the decay, no further calibration is needed.

As the statistics of J/ψ candidates are high enough, all of the parameters of the CB, exponential background, as well as the normalizations of these PDFs (yields of J/ψ and background) will remain free and to be determined in the fit. The fit to the dimuon mass spectrum can be seen in Fig. 7.18, while a list of the results of the fit is shown in Table 7.10. The quality of the fit is good, with a χ^2/NDOF very close to unity. All the fit parameters were computed successfully during the fit procedure. From now on, the value of the J/ψ yield $N_{J/\psi} = 77594 \pm 454$ obtained from the fit will be considered as nominal in this analysis, to do further calculations. The only error associated to this value is the statistical error computed in the fit. In other J/ψ analyses, the systematic error associated with the signal extraction is between 1-2 %. However, in this work there are no studies of systematics uncertainties, which remains a task for the future.

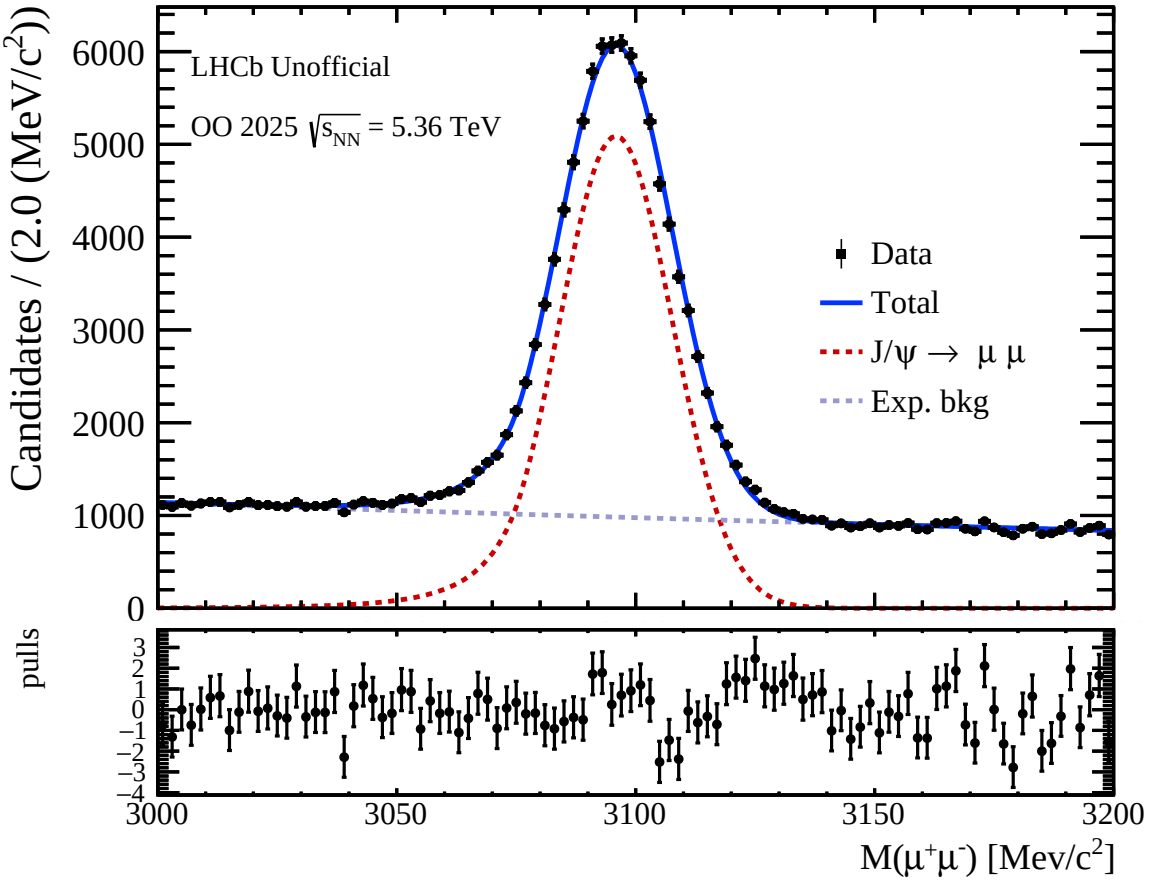


Figure 7.18: Fit to the dimuon invariant mass in the 2025 OO data sample.

Fit parameter	Value
χ^2/NDOF	1.18
m_0 (MeV/c^2)	3095.87 ± 0.06
σ (MeV/c^2)	11.83 ± 0.06
α_L	1.61 ± 0.05
n_L	5.5 ± 1.2
α_R	4.8 ± 1.7
n_R	2.6 ± 8.4
τ	(-1.53 ± 0.05) ·10^{-3}
$N_{J/\psi}$	77594 ± 454
N_{bkg}	98034 ± 475

Table 7.10: List of the fit results for the J/ψ fit.

8

Efficiencies and Systematic Uncertainties

The signal yields extracted in the previous sections do not directly correspond to the number of particles produced in the collisions, as the detector only records a fraction of the generated particles and introduces distortions through its finite acceptance and imperfect reconstruction performance. To compute the true production yields, the measured quantities must therefore be corrected for the detector response, which is commonly referred to as efficiencies.

The computation of these efficiencies quantifies the impact of the detector on the measurement, accounting for effects such as the geometrical and kinematical acceptance of the detector, as well as its response and reconstruction performance. The total efficiency can typically be factorized into acceptance (ε_{acc}), reconstruction and selection ($\varepsilon_{\text{rec\&sel}}$), tracking ($\varepsilon_{\text{trk}}$), and particle identification (ε_{PID}) efficiencies. It is then obtained as the product of the individual contributions:

$$\varepsilon_{\text{Tot}} = \varepsilon_{\text{acc}} \times \varepsilon_{\text{rec\&sel}} \times \varepsilon_{\text{trk}} \times \varepsilon_{\text{PID}}. \tag{8.1}$$

In general, the estimation of the efficiencies is done using the truth information of MC simulation, but they also can be estimated from data-driven methods, specially if the detector response or the underlying physics processes are not well reproduced in the MC simulations. The details of the computation of each efficiency will be given in the following sections.

On the other hand, systematic uncertainties encode the incomplete knowledge of the experimental inputs to the analysis. Typically, the main sources of systematic uncertainties

come from the signal extraction, but also from the computation of the mentioned efficiencies. However, the exact sources of systematic errors depend on the analysis and, in the case of this work, will be detailed in the corresponding section.

8.1 Geometrical Acceptance

All the χ_c produced in the fiducial acceptance in an OO collision do not necessarily have their decay products in the fiducial acceptance as well. To account for this, the observed yields are corrected by the acceptance of the detector. The acceptance efficiency, or geometrical acceptance, is defined as

$$\varepsilon_{acc} = \frac{\chi_c \text{ and all (grand)daughters in fiducial acceptance } 2 < \eta < 4.8}{\chi_c \text{ in fiducial acceptance } 2 < \eta < 4.8}. \tag{8.2}$$

This efficiency is computed from a generator-level MC simulation, meaning that the generated particles are not reconstructed or propagated through the detector. This approach is intended to capture purely kinematic effects and isolate the acceptance from any inefficiency induced by interactions with the detector material or limitations of the reconstruction algorithms. Since the centrally produced MC simulations listed in Table 7.2 were produced with an acceptance cut on the muons, required to be within LHCb acceptance, private only generated MC simulations with no acceptance cuts, also listed in Table 7.2, were used to compute this efficiency. Here, we compare the number of all χ_c states within fiducial acceptance to those that also have decay and have all their (grand)daughters in fiducial acceptance. To be considered within fiducial acceptance, a particle generated randomly in 4π needs to be within a pseudorapidity window $2 < \eta < 4.8$. Furthermore, generated χ_c are required to have a transverse momentum $p_T > 1.5$ GeV, as this is the fiducial cut done in the analysis.

The values of the acceptance efficiency for χ_{c1} and χ_{c2} states, as well as their ratio, are listed in Table 8.1. The uncertainties associated to each efficiency reflect the limited number of events in the MC simulation used to compute these acceptance values. In this table it can be seen that the ratio between the acceptance efficiency of χ_{c1} and χ_{c2} is very close to the unity but still differ from it. This tell us that the geometry of the decay of both χ_c states are very similar, but with little differences we have to take into account with this acceptance efficiency.

	χ_{c1}	χ_{c2}	χ_{c1}/χ_{c2}
ε_{acc}	0.53 ± 0.01	0.50 ± 0.01	1.07 ± 0.02

Table 8.1: Geometrical acceptance for χ_{c1} and χ_{c2} in OO collisions at $p_T > 1.5$ GeV.

In order to compute the fraction of J/ψ that come from the feed-down of the χ_c, it is necessary to take into account the amount of photons from a χ_c decay that are in

the detector acceptance, knowing that the J/ψ from the decay is within the detector acceptance. This photon from χ_c acceptance correction is defined as

$$\text{Acc.}(\gamma \text{ from } \chi_c) = \frac{\chi_c \to J/\psi + \gamma \text{ in LHCb acceptance}}{\chi_c \to J/\psi \text{ in LHCb acceptance}} \tag{8.3}$$

This acceptance correction is also computed from the same private only generated MC simulations with no cuts in the acceptance as the acceptance efficiency, in view of only take into account pure kinematic effects at this stage. In this case, we compare the number of χ_c states within acceptance whose J/ψ and muons are also within detector acceptance to the number of χ_c that also have the photon from their decay within the detector acceptance. As in the case of the acceptance efficiency, it is required to the generated χ_c states to fulfill the fiducial cut of $p_T > 1.5$ GeV. This correction takes into account the prompt J/ψ coming from the χ_c feed-down in the cases when the photon of the χ_c decay ends up out of the acceptance of LHCb.

The values of this acceptance correction are listed in Table 8.2. The uncertainties associated with each acceptance correction reflect the limited number of events in the MC simulation used to compute these values. In this case, we also have computed the case of the sum of both χ_{c1} and χ_{c2}, as the computation of the fraction of J/ψ that comes from the feed-down of the χ_c will be done later for χ_c states in general and not individually for each χ_c state.

	χ_{c1}	χ_{c2}	χ_c
Acc.(γ from χ_c)	0.870 ± 0.005	0.866 ± 0.005	0.868 ± 0.003

Table 8.2: Acceptance correction due to the photons for χ_{c1}, χ_{c2} and their sum ($\chi_c \equiv \chi_{c1} + \chi_{c2}$) in OO collisions at $p_T > 1.5$ GeV.

8.2 Reconstruction and selection efficiency

As detailed in section 7.1.2, in this analysis the reconstruction and selection of the χ_c candidates is done in several steps. For this reason, the reconstruction and selection efficiency can be factorized as the product of the efficiencies in each step. Therefore the total reconstruction and selection efficiency is defined as

$$\varepsilon^{\text{tot}}_{\text{rec\&sel}} = \varepsilon^{\text{offline}}_{\text{rec\&sel}} \times \varepsilon^{\text{BDT}}_{\text{sel}}, \tag{8.4}$$

where $\varepsilon^{\text{offline}}_{\text{rec\&sel}}$ encapsulate the inefficiencies at the reconstruction, pre-selection and offline selection stages, while $\varepsilon^{\text{BDT}}_{\text{sel}}$ compute the inefficiencies introduced by the BDT selection process. Notice that the reconstruction effects are only included in the first step, not in the BDT. The computation of both efficiencies will be done using the centrally produced

MC simulations for χ_{c1}, χ_{c2} listed in Table 7.2. In these MC simulations the generated χ_c are propagated through the detector, to mimic the interaction with the detector material, and reconstructed with the same algorithms as the data, in order to emulate a realistic reconstruction. Although the simulated samples do not perfectly reproduce all features of the data, no data-driven reweighting corrections are applied in the present analysis. Since the measurement is based on ratios of yields, many potential mismodelling effects are expected to largely cancel. In addition, the χ_{c1} and χ_{c2} states have very similar kinematic distributions, leading to comparable reconstruction and selection efficiencies. Therefore, the efficiencies obtained from simulation are used directly.

The reconstruction and selection efficiency at the pre-selection and offline selection steps is defined as

$$\varepsilon^{\text{offline}}_{\text{rec\&sel}} = \frac{\chi_c \text{ that are reconstructed and passed selections in MC}}{\chi_c \text{ and all (grand)daughters in LHCb acceptance } 2 < \eta < 4.8}. \tag{8.5}$$

Here, we compare the number of generated χ_c with all their (grand)daughters within acceptance to the number of reconstructed χ_c that were successfully reconstructed and passed all the cuts, listed in tables 7.3 and 7.4, from the pre-selection and offline selection. In this case the denominator coincide with the numerator in equation 8.2. Then, the selection efficiency of the BDT will be given by

$$\varepsilon^{\text{BDT}}_{\text{sel}} = \frac{\chi_c \text{ selected by the BDT}}{\chi_c \text{ that are reconstructed and passed selections in MC}}. \tag{8.6}$$

In this last step of selection, we compare the number of reconstructed χ_c that were successfully reconstructed and passed all the cuts of pre-selection and offline selection to the number of χ_c that also were selected by the BDT. Here, the denominator coincide with the numerator in equation 8.5. A summary of all the reconstruction and selection efficiencies can be seen in Table 8.3. The uncertainties associated to these efficiencies again reflect the limited number of generated or reconstructed χ_c in the MC simulation used to compute these values. From the results in the table of the ratio between χ_{c1} and χ_{c2} efficiencies, we can see that is very close to the unity. This point a similar performance of the reconstruction and selection processes for both χ_{c1} and χ_{c2}, as expected.

	χ_{c1}	χ_{c2}	χ_c	χ_{c1}/χ_{c2}
$\varepsilon^{\text{offline}}_{\text{rec\&sel}}$	0.0911 ± 0.0003	0.0860 ± 0.0003	0.0887 ± 0.0002	1.059 ± 0.004
$\varepsilon^{\text{BDT}}_{\text{sel}}$	0.9716 ± 0.0005	0.9463 ± 0.0007	0.9601 ± 0.0004	1.027 ± 0.001

Table 8.3: Reconstruction and selection efficiencies for χ_{c1}, χ_{c2}, their ratio and their sum ($\chi_c \equiv \chi_{c1} + \chi_{c2}$) in OO collisions.

As mentioned in 7.4 the only difference between the selection process in the χ_c production measurement and the J/ψ production measurement is the BDT step, that is only applied in the case of the χ_c. Consequently, in the later computation of the fraction of J/ψ that come from the feed-down of the χ_c, it will be necessary to take also into account corrections using BDT selection efficiency.

8.3 Tracking and muon ID efficiency

The tracking efficiency is defined as the probability of a charged particle which actually passed through the detector to be successfully reconstructed as a track by the tracking algorithm. The baseline tracking efficiency determination assumes that track reconstruction performance is identical in data and simulation. However, the most accurate way to estimate tracking efficiencies is using data-driven techniques. In the case of tracking efficiency for muons, the most widely used method is the *Tag-and-Probe* technique [128]. The tag-and-probe method uses decays as $K_s^0 \to \pi^+\pi^-$ or $J/\psi \to \mu^+\mu^-$. For example, in the case of K_s^0 one of the charged pions (tag) is reconstructed as a long track and the other (probe) is reconstructed as a VELO track. The tracking efficiency is then the fraction of K_s^0 decays with a probe track matched to the long track produced by the corresponding pion.

On the other hand, we understand the muon PID efficiency as the probability of a muon to be correctly identified as a muon by the particle identification system. The determination of muon PID efficiency is also performed using the tag-and-probe method. In this case, $J/\psi \to \mu^+\mu^-$ is the decay used for the computation of this efficiency. Here, one of the muons (tag) is selected with all the proper PID requirements while the in other muon (probe) no PID requirements are applied. Then the muon PID efficiency is computed as the fraction of J/ψ reconstructed with muon PID compared with the J/ψ where no muon PID were applied.

Within LHCb, there are are tools for the direct application of these efficiencies to an analysis. These are `TrackCALIB` and `PIDCALIB` packages, where the already computed tracking and PID efficiencies respectively are listed in tables as a function of several variables, as momentum, pseudorapidity or multiplicity. By the time being, these tools are not available yet for the OO data sample. Therefore, the proper computation of these efficiencies remain as a task to do in the future. However, in the following we will show that, for the statistical precision we can reach with the OO data sample, we can consider that the kinematics of the χ_c states are identical. Then the corrections needed due to tracking and PID would cancel out.

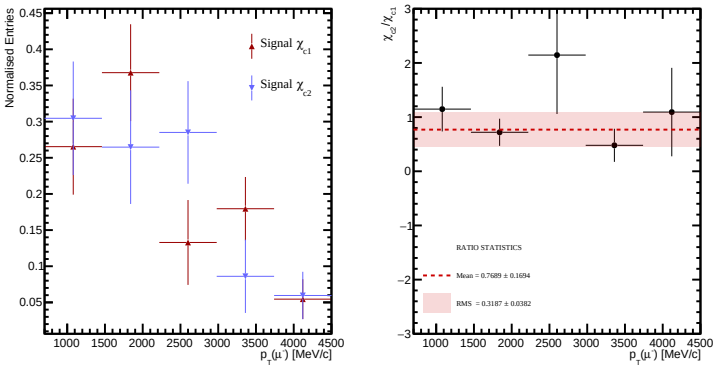
In order to show this we are going to do a comparison of the kinematic variables of the measured χ_c states. The limited statistics of the OO sample do not allow for differential fits. Instead, we obtained the kinematic distributions for individual daughters from sPlot technique [129]. The sPlot assigns a statistical weight (sWeight) to each event based on a

likelihood fit to the invariant mass, such that when these weights are applied to any other variable, the contribution of background events cancels out and the true distribution of the signal can be recovered. For this comparison the model used for the fit is a CB for the signal of both χ_c states and $N_{\text{comb}}^{\text{max}} = 5000$, $|\Delta PVZ|_{\text{max}} = 80$ mm as the parameters for the combinatorial background. The sWeights are applied to the transverse momentum and the pseudorapidity of χ_c states daughters.

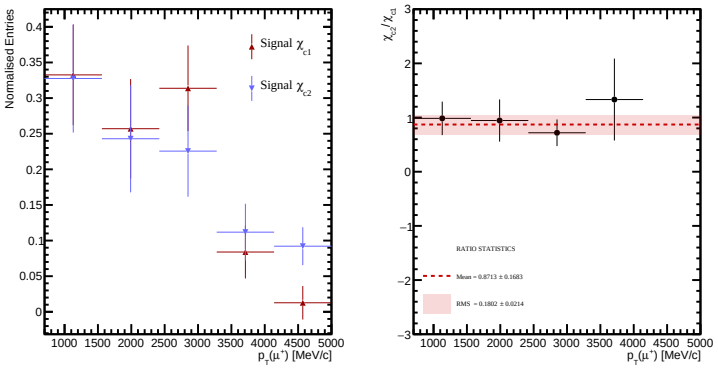
The kinematic distributions and their ratios are shown in Figures 8.1 and 8.2. The ratios obtained from the sPlot distributions are displayed in the right-hand panels of the figures, together with their weighted mean (solid red line) and RMS (shaded red band). In all cases, the mean ratio is found to be compatible with unity within uncertainties. This observation indicates that, within the current precision, the kinematic distributions of the daughter particles from the two χ_c states are indistinguishable. Consequently, tracking and PID corrections are expected to largely cancel in the efficiency ratio between χ_{c1} and χ_{c2} and are therefore neglected in the comparison of the two states.

The same study is repeated by comparing the kinematic distributions of the χ_c daughters with those of the J/ψ daughters, both extracted using the sPlot technique. The corresponding comparisons are presented in Figures 8.3 and 8.4. The larger statistics available in the J/ψ sample lead to a more precise determination of the daughter-particle kinematic distributions. In this case, differences are observed between the muon kinematics originating from χ_c decays and those from inclusive J/ψ decays, particularly at low transverse momentum and pseudorapidity. Such a behavior is expected, since the J/ψ produced in χ_c decays carries most of the momentum of the parent particle, resulting in more energetic decay muons than those produced directly in J/ψ decays.

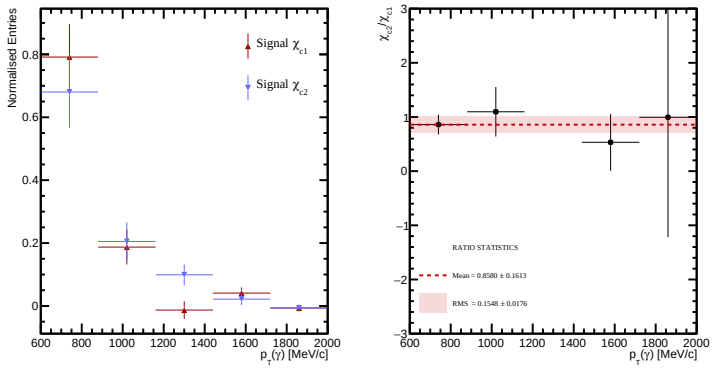
Nevertheless, the ratios of the distributions remain compatible with unity within uncertainties. Therefore, for the integrated measurement presented here, the tracking and PID efficiency corrections are assumed to cancel to first order in the ratio between χ_c and J/ψ decays. Any residual effect is expected to be small compared to the current statistical uncertainty. In future developments of the analysis, the tracking and muon PID efficiencies will be computed explicitly using dedicated calibration samples, enabling a more accurate evaluation of the corresponding corrections and systematic uncertainties.



(a) Transverse momentum of μ^-.

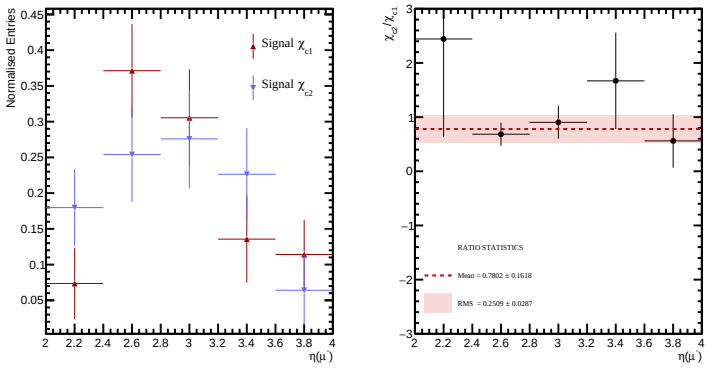


(b) Transverse momentum of μ^+.

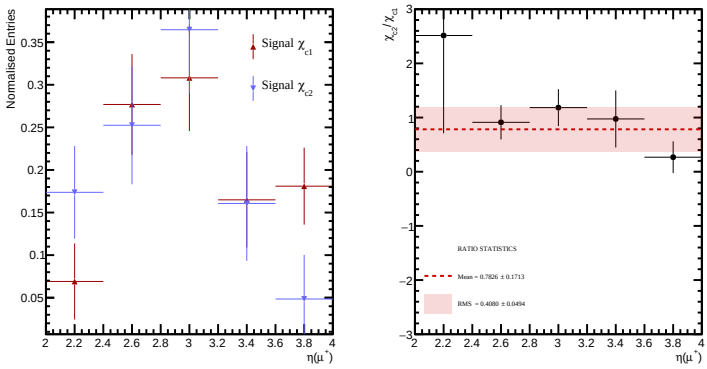


(c) Transverse momentum of γ.

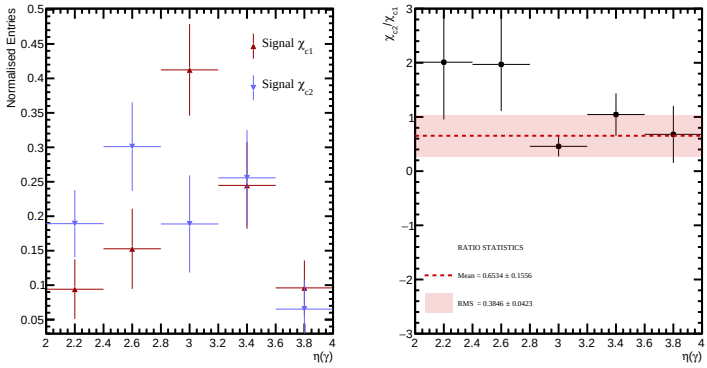
Figure 8.1: Comparison of the transverse momentum distribution from sPlot of the daughters of χ_{c1} and χ_{c2} in the 2025 OO data sample. The right-side plots show the ratio between both distributions.



(a) Pseudorapidity of μ^-.

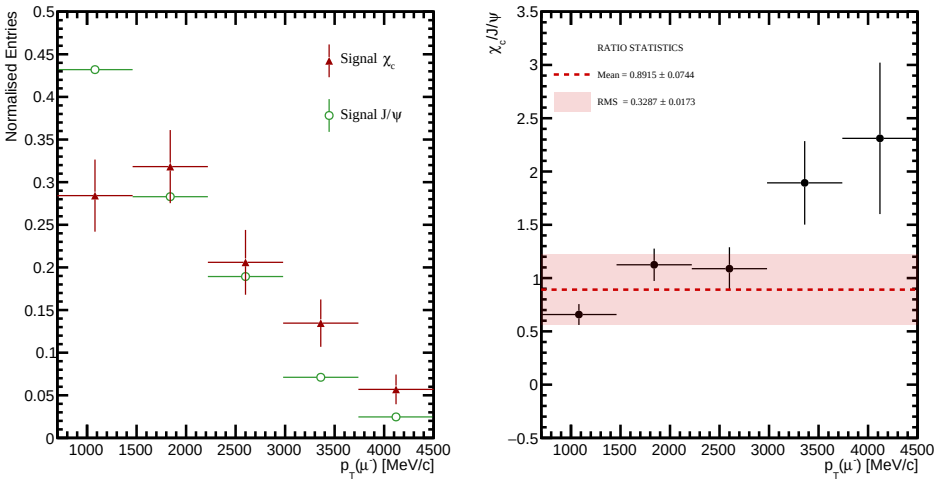


(b) Pseudorapidity of μ^+.

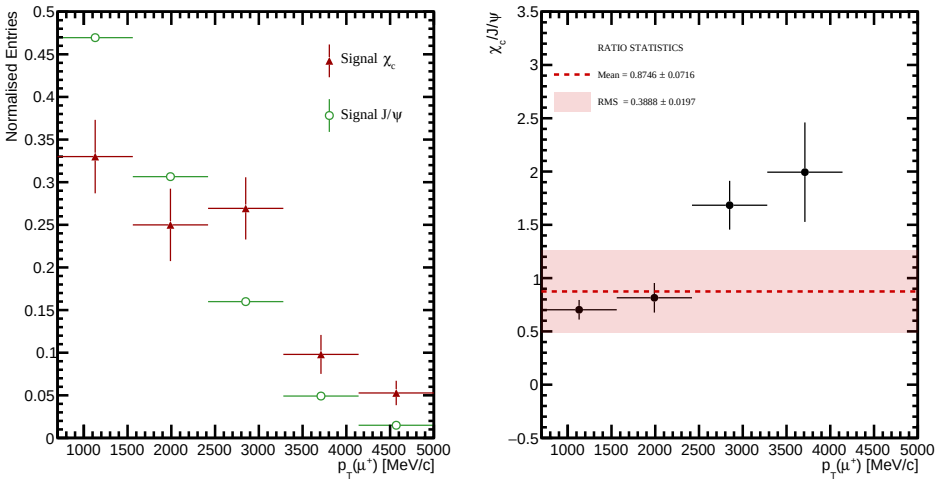


(c) Pseudorapidity of γ.

Figure 8.2: Comparison of the pseudorapidity distribution from sPlot of the daughters of χ_{c1} and χ_{c2} in the 2025 OO data sample. The right-side plots show the ratio between both distributions.

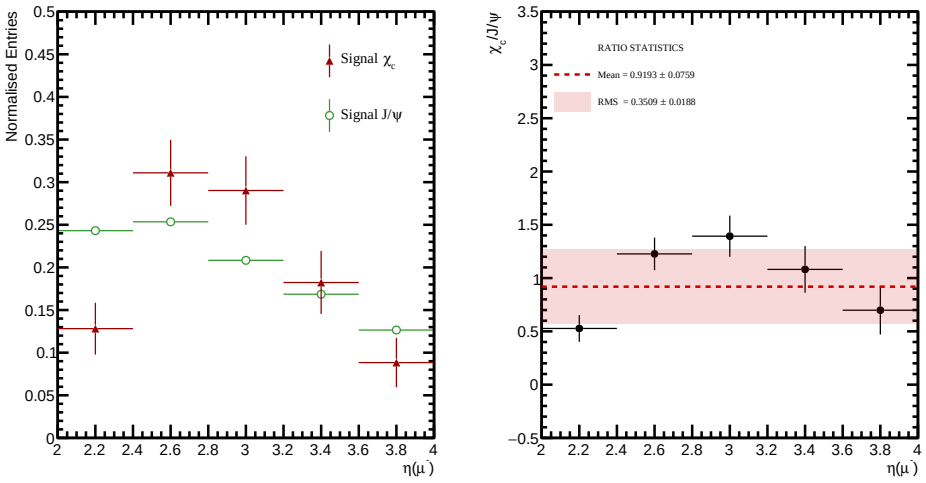


(a) Transverse momentum of μ^-.

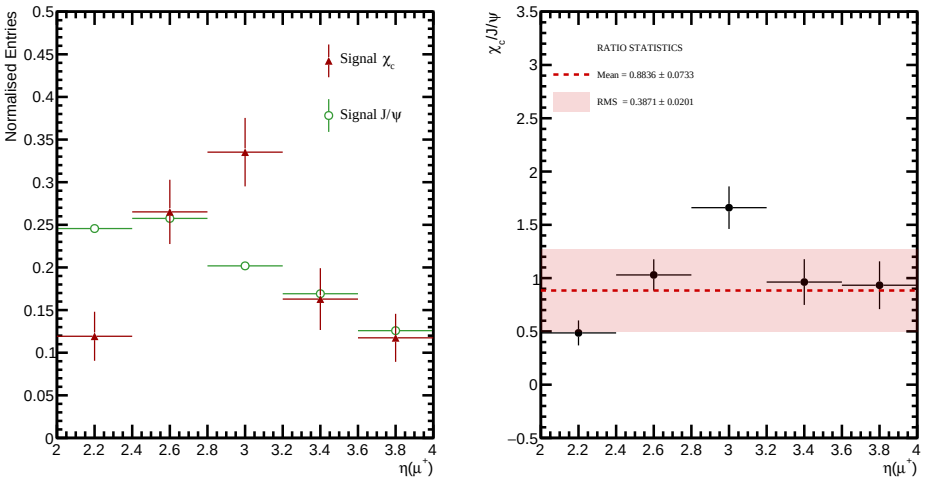


(b) Transverse momentum of μ^+.

Figure 8.3: Comparison of the transverse momentum distribution from sPlot of the daughters of χ_c and J/ψ in the 2025 OO data sample. The right-side plots show the ratio between both distributions.



(a) Pseudorapidity of μ^-.



(b) Pseudorapidity of μ^+.

Figure 8.4: Comparison of the pseudorapidity distribution from sPlot of the daughters of χ_c and J/ψ in the 2025 OO data sample. The right-side plots show the ratio between both distributions.

8.4 Photon detection efficiency

Photon reconstruction plays a central role in radiative decay measurements. Although the photon efficiency cancels to first order in the comparison between the χ_{c1} and χ_{c2} states, it remains the dominant efficiency contribution when comparing χ_c production with prompt J/ψ production. A reliable determination of the photon detection efficiency is therefore essential.

Estimating photon efficiencies only from detector simulation is challenging due to uncertainties associated with the detector material budget, the calorimeter response, electromagnetic shower development, and cluster merging effects arising from the large abundance of photons originating from π^0 decays. In LHCb, these effects have traditionally been constrained using control samples involving radiative decays, such as $B^+ \to J/\psi K^{*+}$. However, these channels provide only limited statistics in the OO data sample.

In this work, the photon detection efficiency is determined using the abundant decays

$$\eta^0 \to \pi^+\pi^-\pi^0, \qquad \omega \to \pi^+\pi^-\pi^0, \tag{8.7}$$

following a strategy previously developed by LHCb for pPb collisions [62]. The measured efficiency includes both the probability for a photon to produce a reconstructed ECAL cluster and the probability for this cluster to satisfy the photon identification requirements used in the analysis.

8.4.1 Measurement strategy

The determination of the photon detection efficiency relies on the tag-and-probe method. Since the decay topology is identical for the η^0 and ω mesons, the strategy will be illustrated using the ω decay, the procedure being fully analogous for the η^0 case.

For these decays, two categories of candidates can be distinguished. The first one corresponds to the fully reconstructed decay

$$\omega \to \pi^+\pi^-(\pi^0 \to \gamma_{\text{tag}}\gamma_{\text{prob}}), \tag{8.8}$$

where both photons originating from the π^0 decay are successfully reconstructed. The second category corresponds to the partially reconstructed decay

$$\omega \to \pi^+\pi^-(\pi^0 \to \gamma_{\text{tag}}\cancel{\gamma_{\text{prob}}}), \tag{8.9}$$

where one of the photons is not reconstructed. The reconstructed photon is referred to as the *tag photon*, while the missing photon is referred to as the *probe photon*.

The yields of these two populations are denoted as $N_\omega^{\text{full(rec)}}$ and $N_\omega^{\text{part(rec)}}$, respectively. In terms of efficiencies, they can be expressed as

$$N_\omega^{\text{full(rec)}} = N_\omega^{\text{tot}} \times \varepsilon_{\text{tot}}(\pi^\pm) \times A^{\gamma_{\text{tag}}} \times \varepsilon_{\text{det}}^{\gamma_{\text{tag}}} \times A^{\gamma_{\text{prob}}} \times \varepsilon_{\text{det}}^{\gamma_{\text{prob}}}, \tag{8.10}$$

$$N_\omega^{\text{part(rec)}} = N_\omega^{\text{tot}} \times \varepsilon_{\text{tot}}(\pi^\pm) \times A^{\gamma_{\text{tag}}} \times \varepsilon_{\text{det}}^{\gamma_{\text{tag}}}, \tag{8.11}$$

where N_ω^{tot} denotes the total number of produced ω mesons, $\varepsilon_{\text{tot}}(\pi^\pm)$ represents all efficiencies associated with the charged pions, $A^{\gamma_{\text{tag/prob}}}$ is the geometrical and kinematical acceptance for a given photon, and $\varepsilon_{\text{det}}^{\gamma_{\text{tag/prob}}}$ is the corresponding photon detection efficiency.

Taking the ratio of the two populations leads to

$$\frac{N_\omega^{\text{full(rec)}}}{N_\omega^{\text{part(rec)}}} = A^{\gamma_{\text{prob}}} \times \varepsilon_{\text{det}}^{\gamma_{\text{prob}}}. \tag{8.12}$$

Therefore, the ratio of fully and partially reconstructed yields provides access to the product of the probe-photon acceptance and detection efficiency. Since both quantities depend on the probe-photon kinematics and on the event multiplicity, Eq. 8.12 can be rewritten as

$$\frac{N_\omega^{\text{full(rec)}}(p_T^{\gamma_{\text{prob}}}, \eta^{\gamma_{\text{prob}}}, N_{\text{VeloTracks}})}{N_\omega^{\text{part(rec)}}(p_T^{\gamma_{\text{prob}}}, \eta^{\gamma_{\text{prob}}}, N_{\text{VeloTracks}})} = A^{\gamma_{\text{prob}}}(p_T^{\gamma_{\text{prob}}}, \eta^{\gamma_{\text{prob}}}) \times \varepsilon_{\text{det}}^{\gamma_{\text{prob}}}(p_T^{\gamma_{\text{prob}}}, \eta^{\gamma_{\text{prob}}}, N_{\text{VeloTracks}}). \tag{8.13}$$

The event multiplicity is characterized using the number of reconstructed VELO tracks, $N_{\text{VeloTracks}}$.

Although the efficiency is ultimately determined as a function of the probe-photon kinematics, these quantities are not directly accessible in the partially reconstructed sample, since the probe photon is not reconstructed. Consequently, the partially reconstructed yields are first extracted as functions of the measurable kinematic variables of the reconstructed $\pi^+\pi^-\gamma_{\text{tag}}$ candidate and are subsequently projected onto the probe-photon phase space using a kinematic projection matrix derived from simulation. The details of this procedure are presented in Section 8.4.4.

Finally, combining the ratio of fully and partially reconstructed yields measured in data with the acceptance factor obtained from simulation provides an estimate of the photon detection efficiency as a function of the probe-photon transverse momentum, pseudorapidity, and event multiplicity.

8.4.2 Data selection and MC simulations

The efficiency determination was performed using a reduced subset of the OO data sample. The full data sample was collected at $\sqrt{s_{NN}} = 5.36$ TeV and corresponds to an integrated luminosity of approximately 5.1 nb^{-1}.

Regarding the simulated samples, there are no dedicated MC simulations for the decays $\eta^0 \to \pi^+\pi^-\pi^0$ and $\omega \to \pi^+\pi^-\pi^0$. However, these decays are abundantly produced in minimum-bias collisions, meaning that they are naturally present in any inclusive OO minimum-bias simulation. A list of the simulation samples used throughout the photon efficiency determination is shown in Table 8.4. In addition, a private minimum-bias generated sample containing 10^6 events and produced without acceptance requirements is used in several steps of the efficiency measurement.

Process	Sim version
inclusive $\Upsilon(1S)$	Sim10f
inclusive $\Upsilon(2S)$	Sim10f
inclusive $\Upsilon(3S)$	Sim10f
inclusive J/ψ	Sim10f
inclusive $\psi(2S)$	Sim10f
$D^0 \to K^\pm\pi^\mp$	Sim10f
$\chi_{c1} \to J/\psi(\mu^+\mu^-)\gamma$	Sim10g
$\chi_{c2} \to J/\psi(\mu^+\mu^-)\gamma$	Sim10g
$\psi(2S) \to J/\psi(\mu^+\mu^-)\pi^0\pi^0$	Sim10g
$D^0 \to K^\pm\pi^\mp$	Sim10g
minbias	Sim10g

Table 8.4: List of MC simulation datasets for photon efficiency.

The selection criteria were defined using truth-matched MC candidates and then applied identically to both the fully and partially reconstructed samples. A list of all the cuts applied is shown in Table 8.5. Since the photon detection efficiency obtained in this chapter will be used to correct the measured χ_c yields, the photon identification requirements are chosen to be the same as those used in the χ_c analysis.

Tight requirements are imposed on the charged pions in order to suppress contamination from misidentified kaons, protons and ghost tracks. In addition, a minimum transverse momentum requirement is applied to improve the signal-to-background ratio. The cuts applied to the pion pair are intended to select prompt candidates while reducing contributions from long-lived particles, such as K^0_S. Despite these tight selection criteria, the η^0 and ω signals are still barely visible above the combinatorial background. The extraction of the fully and partially reconstructed η^0 and ω signals will be described in the following sections.

ω candidates selection	
Mother cuts	$M(\omega) < 1000$ MeV/c^2 $p_T(\omega) > 3200$ MeV/c $2 < \eta(\omega) < 4.8$
Pion pair cuts	$300 < M(\pi\pi) < 600$ MeV/c^2 $p_T(\pi\pi) > 2000$ MeV/c DOCA$(\pi\pi) < 0.1$ mm $\lvert$DIRA$(\pi\pi)\rvert > 0.995$ FD$(\pi\pi) < 8$ mm IP$(\pi\pi) < 0.1$ mm
Pion cuts	$p_T(\pi) > 1000$ MeV/c $2 < \eta(\pi) < 4.8$ $\chi^2(\pi) < 4$ GHOSTPROB$(\pi) < 0.1$ PROBNNPI$(\pi) > 0.9$ PIDK < -10 PIDP < -20
Photons cuts	$p_T(\gamma) > 600$ MeV/c $2 < \eta(\gamma) < 4$ $E_{Hcal}/E_{Ecal} < 0.3$ IsPhoton > 0.4 IsNotHadron > 0.6

Table 8.5: ω candidates selection done according to the MC truth information.

8.4.3 Measurement of fully reconstructed η^0 and ω

The invariant mass distribution of the fully reconstructed sample is composed of several contributions. The first ones are the η^0 and ω signals at their nominal masses in the $\pi^+\pi^-\gamma\gamma$ invariant mass spectrum. The background contribution can be divided into two categories: a background associated with the charged pions and a background associated with the photons. The pion background originates from uncorrelated combinations of π^0 mesons with charged pion pairs, while the photon background arises from uncorrelated

combinations of photon pairs.

In order to subtract the combinatorial background, the like-sign $\pi^{\pm}\pi^{\pm}\gamma\gamma$ invariant mass distribution is used. The underlying assumption is that the like-sign distribution reproduces the combinatorial background as long as the two pions forming the pair are uncorrelated, which is expected since no resonant structure is observed in the like-sign invariant mass spectrum. A comparison between the opposite-sign and like-sign distributions is shown in Figure 8.5.

Usually, the like-sign distribution is computed as $2\sqrt{N^{++}_{\text{raw}}N^{--}_{\text{raw}}}$, where N^{++}_{raw} and N^{--}_{raw} correspond to the invariant mass distributions built from positive and negative pion pairs, respectively. In this analysis, the like-sign distribution is built directly at the DaVinci level from the same data sample, applying the same selection criteria as for the opposite-sign sample and considering all possible combinations of same-sign pion pairs with two photons. We denote this distribution as $N^{\pm\pm}_{\text{raw}}$.

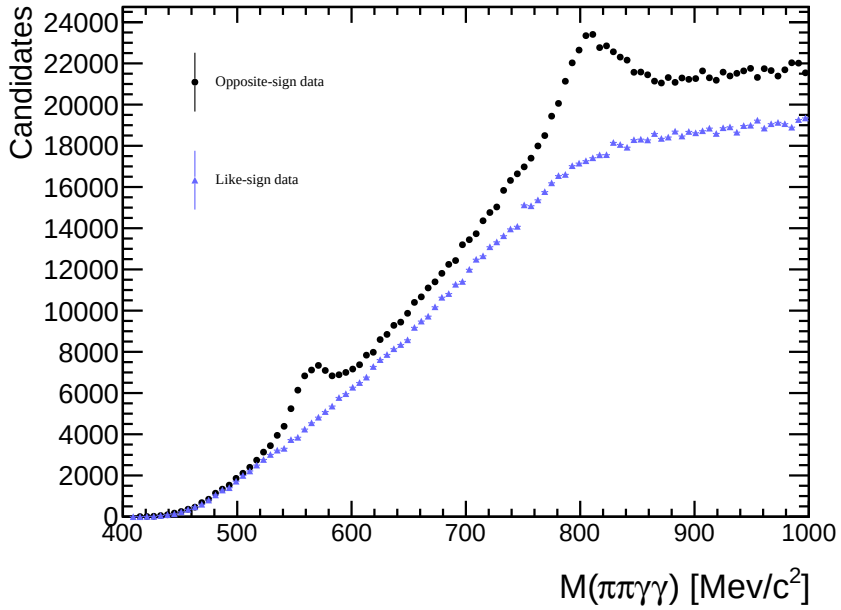


Figure 8.5: Comparison between the opposite-sign and like-sign in the $\pi\pi\gamma\gamma$ invariant mass spectrum in the 2025 OO data sample.

In general, the like-sign distribution $N^{\pm\pm}_{\text{raw}}$ does not contain the same number of combinations as the opposite-sign distribution, denoted as N^{+-}_{raw}. Therefore, a proper normalization of the like-sign distribution is required in order to extract the signal contribution

$$N_{\text{sig}} = N^{+-}_{\text{raw}} - R_{\text{mix}} \cdot N^{\pm\pm}_{\text{raw}}. \tag{8.14}$$

The normalization factor R_{mix} is introduced as a bin-by-bin correction accounting for charge-dependent differences between opposite-sign and like-sign pairs in mixed events,

including detector acceptance and reconstruction asymmetries, and is defined as

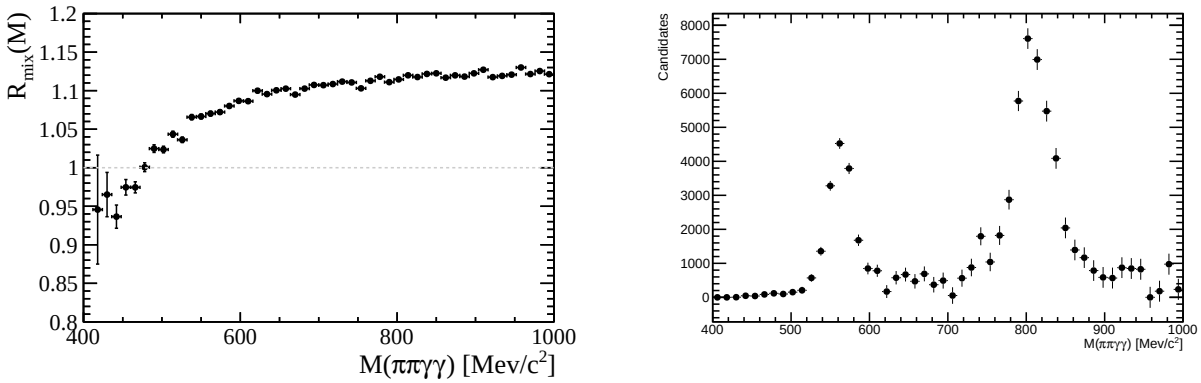
$$R_{\text{mix}} = \frac{N^{+-}_{\text{mix}}}{N^{\pm\pm}_{\text{mix}}}, \tag{8.15}$$

Here, N^{+-}_{mix} and $N^{\pm\pm}_{\text{mix}}$ are the invariant mass distributions obtained from event mixing of opposite-sign and like-sign pion pairs, respectively. The same event mixing algorithm introduced in Section 7.3.1 is used to compute this factor, with a few modifications adapted to the present case.

In this case, the event mixing algorithm combines charged pion pairs from one event with photon pairs from a different event. In addition, only events with similar PVZ and multiplicity are allowed to be combined. The parameters of the event mixing algorithm are fixed to $N^{\text{max}}_{\text{comb}} = 2000$ and $|\Delta PVZ|_{\text{max}} = 80$ mm.

To account for the multiplicity dependence, the $N_{\text{VeloTracks}}$ distribution is divided into five intervals containing approximately 20% of the total candidates each, and only events belonging to the same multiplicity interval can be combined.

The resulting R_{mix} factor and the invariant mass distribution after the background subtraction are shown in Figure 8.6. After the subtraction, clear η^0 and ω signals are observed, although a small residual background remains due to uncorrelated photon combinations. This contribution is modeled during the signal extraction procedure.



(a) R_{mix} normalization factor. (b) Invariant mass distribution.

Figure 8.6: R_{mix} normalization factor as a function of the invariant mass (a) and invariant mass distribution after the background subtraction (b) using the like-sign distribution normalized by the R_{mix} factor for the fully reconstructed case in the 2025 OO data sample.

Once most of the background contribution has been removed, the fully reconstructed η^0 and ω yields are extracted by fitting the invariant mass spectrum.

The signal shape is modeled using truth-matched simulated candidates, from which the fit parameters used to constrain the fit to data are obtained. The signal model is chosen

to be a Voigtian function, defined as the convolution of a Breit-Wigner distribution with a Gaussian function. This parameterization simultaneously accounts for the natural width of the particle (Γ) and the detector resolution (σ). Only the Γ parameter is fixed to the value obtained from simulation, while the mean and resolution are left free in the fit to data.

The fit to the reconstructed simulated sample is shown in Figure 8.7, and the resulting values of Γ are listed in Table 8.6. The residual background contribution is modeled using a generic analytical function. In this analysis, a second-order exponential function is used,

$$f_{\text{bkg}}(m; a_1, a_2) \propto e^{m(a_1+a_2 m)}, \tag{8.16}$$

where the parameters a_1 and a_2 are left free in the fit to data.

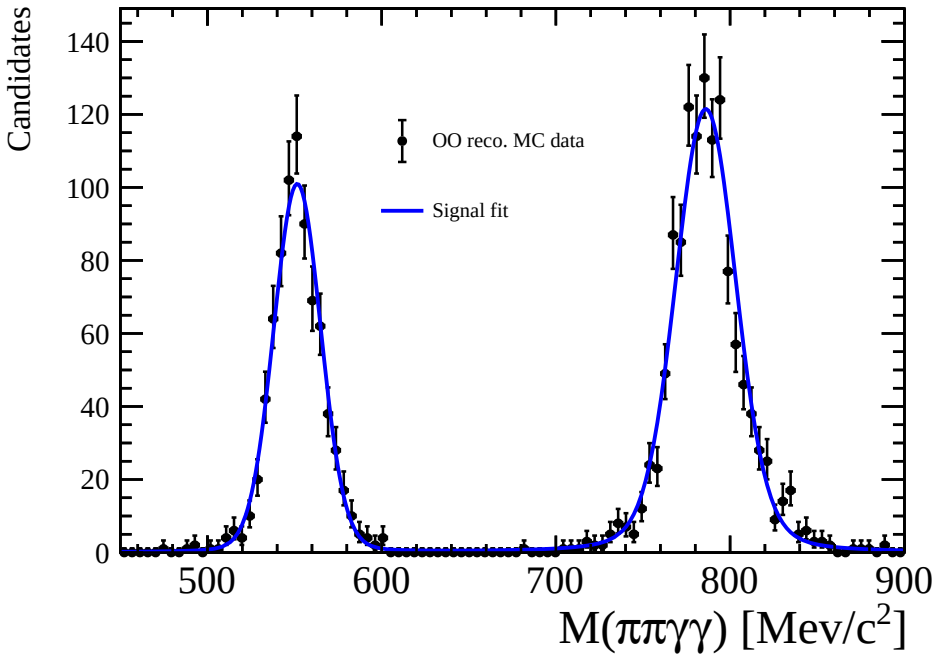


Figure 8.7: Fit to the invariant mass distribution in reconstructed OO MC simulation data for η^0 and ω signal for fully reconstructed case.

	ω	η^0
Γ (MeV/c^2)	7.55 $\pm$ 1.52	2.35 $\pm$ 0.97

Table 8.6: Γ parameter result from the fit to the reconstructed truth MC data.

Having defined the signal and background models, the η^0 and ω yields are extracted by

fitting the invariant mass spectrum obtained after the background subtraction. The fit is performed differentially as a function of $p_T^{\gamma_{\rm prob}}$, $\eta^{\gamma_{\rm prob}}$ and $N_{\rm VeloTracks}$.

The binning in $\eta^{\gamma_{\rm prob}}$ and $N_{\rm VeloTracks}$ is chosen to study the behavior of the photon detection efficiency in regions of low and high pseudorapidity and multiplicity, respectively. The transverse momentum binning is selected according to the available statistics and covers the same kinematic range as the photons from χ_c decays.

Examples of fits corresponding to different regions of transverse momentum, pseudorapidity and multiplicity are shown in Figure 8.8. The uncertainties in the invariant mass distributions are propagated from the background subtraction procedure and become larger at high pseudorapidity and multiplicity. The fitted η^0 and ω yields obtained in the different kinematic intervals are listed in Table 8.7.

$p_T^{\gamma_{\rm prob}}$ (MeV/c) bin	$\eta^{\gamma_{\rm prob}}$ bin	$N_{\rm VeloTracks}$ bin	$N_\omega^{\rm full(rec)}$	$N_\eta^{\rm full(rec)}$
[600, 1000]	[2, 3.2]	[0, 180]	6161 ± 143	3297 ± 69
[600, 1000]	[2, 3.2]	[180, 420]	4504 ± 125	2779 ± 58
[600, 1000]	[3.2, 4]	[0, 180]	2904 ± 95	1144 ± 42
[600, 1000]	[3.2, 4]	[180, 420]	642 ± 74	755 ± 38
[1000, 1500]	[2, 3.2]	[0, 180]	3823 ± 93	1840 ± 48
[1000, 1500]	[2, 3.2]	[180, 420]	2735 ± 71	1180 ± 41
[1000, 1500]	[3.2, 4]	[0, 180]	1528 ± 62	617 ± 28
[1000, 1500]	[3.2, 4]	[180, 420]	653 ± 38	214 ± 20
[1500, 5000]	[2, 3.2]	[0, 180]	2614 ± 74	954 ± 34
[1500, 5000]	[2, 3.2]	[180, 420]	2533 ± 82	826 ± 32
[1500, 5000]	[3.2, 4]	[0, 180]	1019 ± 44	226 ± 19
[1500, 5000]	[3.2, 4]	[180, 420]	503 ± 36	212 ± 16

Table 8.7: Results from the invariant mass fit of the η^0 and ω yields for the different bins of transverse momentum, pseudorapidity and multiplicity in the fully reconstructed case.

8.4.4 Measurement of partially reconstructed η^0 and ω

The measurement of the partially reconstructed η^0 and ω signals follows the same strategy as in the fully reconstructed case, although the procedure becomes considerably more involved. The raw mass distribution of the partially reconstructed sample is composed by several components. First, the partially reconstructed η^0 and ω signals appear at masses lower than their nominal values, since the energy of the probe photon is missing. In addition, two important background contributions are present: the combinatorial background due to uncorrelated combinations of pion pairs and photons and the correlated background coming from other decays.

In this case, we also want to do a background subtraction using the like-sign distribution

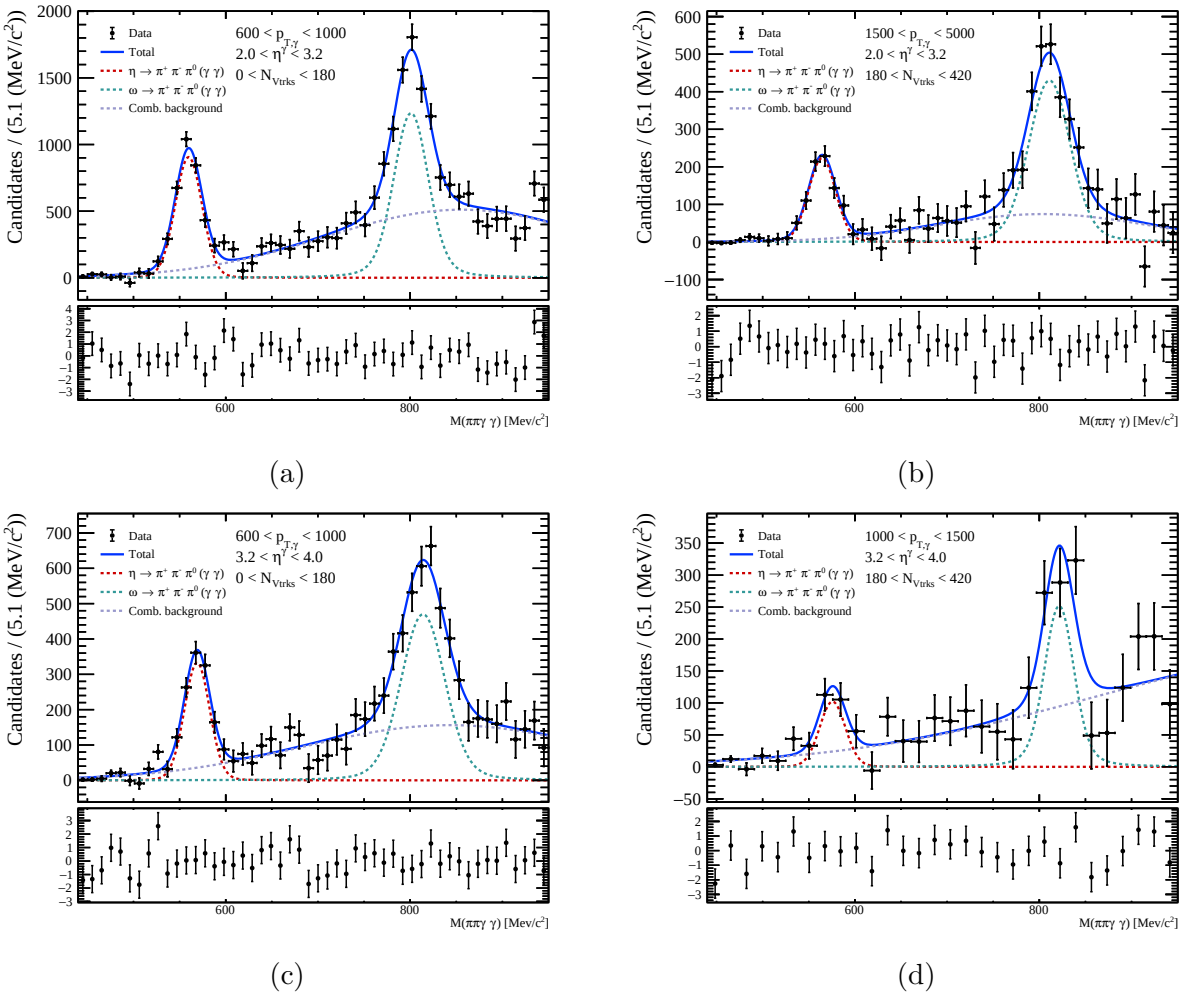


Figure 8.8: Examples of fits to the $\pi\pi\gamma\gamma$ invariant mass for different bins of transverse momentum, pseudorapidity and multiplicity in the 2025 OO data sample. The (a) and (b) fits are cases of low pseudorapidity, while (c) and (d) are at high pseudorapidity. In addition, (a) and (c) fits are cases of low multiplicity, while (b) and (d) are high multiplicity.

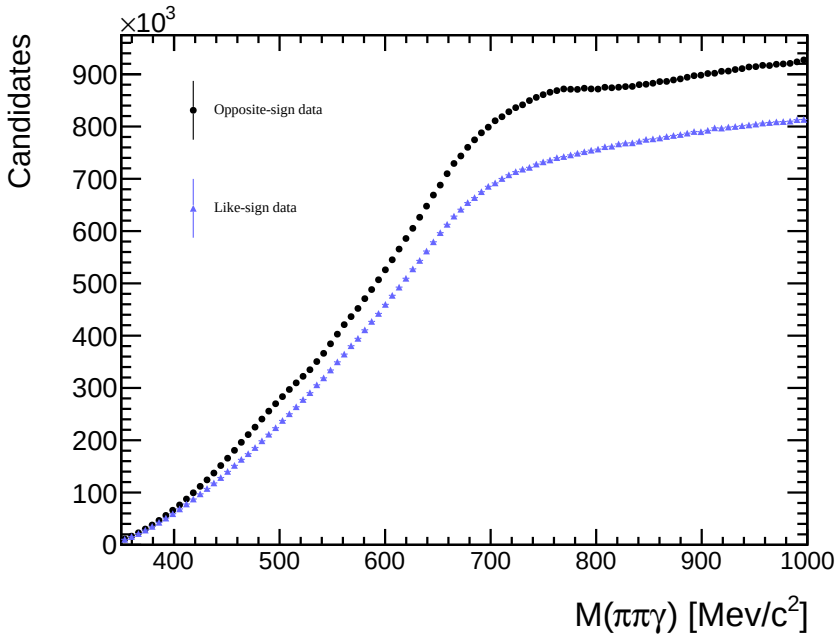
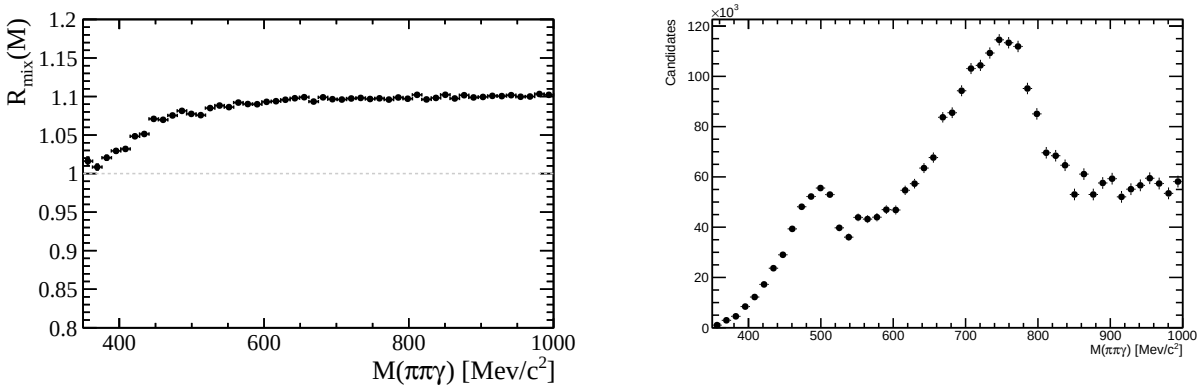


Figure 8.9: Comparison between the opposite-sign and like-sign in the $\pi\pi\gamma$ invariant mass spectrum in the 2025 OO data sample.

normalized with the R_{mix} factor computed from the event mixing. The raw $\pi\pi\gamma$ invariant mass distribution and the like-sign distribution before the normalization are presented in figure 8.9. No clear η^0 or ω resonance can be identified in the opposite-sign distribution. In this case, since the photon is not constrained to come from a π^0, the number of possible combinations with the pion pair increases, leading to an opposite-sign distribution almost entirely dominated by the combinatorial background.

For the computation of the R_{mix} factor in the partially reconstructed case, the event mixing algorithm combines charged pion pairs from one event with photons in a different event. The condition of similar multiplicity and similar PVZ remains the same as in the fully reconstructed case in the event mixing algorithm. Since the partially reconstructed sample contains significantly larger statistics than the fully reconstructed one, a smaller value of the parameter $N_{\text{comb}}^{\text{max}}$ will be used in order to save computational time. Specifically, the chosen value is $N_{\text{comb}}^{\text{max}} = 100$. The resulting R_{mix} factor and the invariant mass distribution after background subtraction are shown in Fig. 8.10.

After the background subtraction, the two partially reconstructed resonances of η^0 and ω are clearly visible in the $\pi\pi\gamma$ invariant mass spectrum. However, we can see there is still a significant background contribution. Studies performed with simulated samples indicate that this residual background is mainly associated with correlated pion pairs originating from $\rho^0 \to \pi^+\pi^-$ decays. Here, the correlated pion pairs from the ρ^0 decay are combined with random photons, which creates a wide distribution around the ρ^0 nominal mass.



(a) R_{mix} normalization factor.

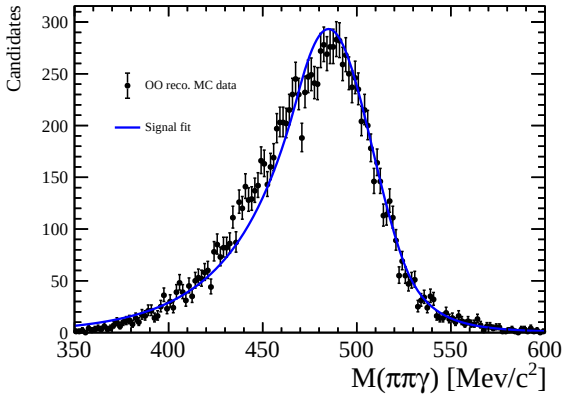
(b) Invariant mass distribution.

Figure 8.10: R_{mix} normalization factor as a function of the invariant mass (a) and invariant mass distribution after the background subtraction (b) using the like-sign distribution normalized by the R_{mix} factor for the partially reconstructed case in the 2025 OO data sample.

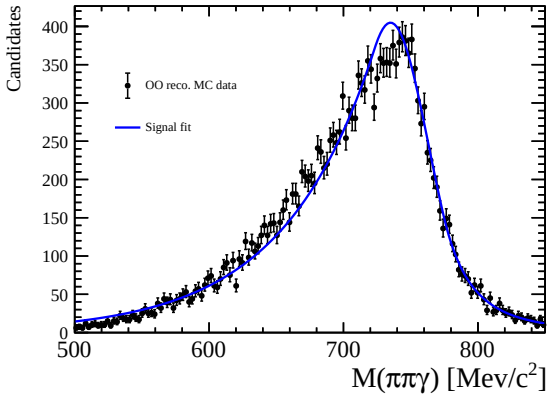
Knowing the origin of this correlated background, we are going to take advantage of the already computed opposite-sign event mixing distribution N^{+-}_{mix}, since the event-mixing procedure reproduces the same topology, namely correlated pion pairs combined with uncorrelated photons. Therefore, the shape of the event mixing opposite-sign distribution will be used as a PDF to model the correlated background in the signal extraction. Apart from the partially reconstructed resonances of η^0 and ω and the correlated background, in the signal extraction process we have to take into account the fully reconstructed decay of $\eta^0 \to \pi^+\pi^-\gamma$, that also contributes in the invariant mass spectrum.

As discussed above, the $\pi\pi\gamma$ invariant mass spectrum shows several contributions that have to be taken into account during the signal extraction. In order to constrain as much as possible the fit procedure to the data, we are going to model the signals shape using truth information of the reconstructed MC simulation. The partially reconstructed resonances are particularly challenging to model because of the missing energy carried away by the probe photon. Several models were tested in order to reproduce the partially reconstructed resonances observed in simulation. Finally, the chosen model was a CB defined in equation 7.7. The tail parameters of the CB model obtained from the fit to the MC will be fixed in the fit to the data. The fits to the partially reconstructed MC data are shown in figure 8.11.

Even with the chosen parameterization, the CB model does not perfectly reproduce the shape observed in simulation, especially in the peak region and the left tail. Consequently, a sizeable systematic uncertainty associated with the signal extraction procedure is expected for the partially reconstructed sample. The results of the parameters from the MC fit to be fixed in the fit to the data are listed in Table 8.8.



(a) Partially reconstructed η^0 truth signal.



(b) Partially reconstructed ω truth signal.

Figure 8.11: Fit to the invariant mass distribution in reconstructed OO MC simulation data for η^0 (a) and ω (b) signal for partially reconstructed case.

	ω	η^0
α_L	0.42 ± 0.01	0.71 ± 0.02
n_L	123.29 ± 0.02	129.31 ±0.06
α_R	1.40 ± 0.07	1.69 ± 0.09
n_R	3.57 ± 0.95	6.6 ± 2.5

Table 8.8: Parameter results from the fit to the partially reconstructed truth MC data.

Ideally, the fit to the MC sample would be performed differentially using the same variables and the same binning as in data. Since the CB tail parameters depend on the transverse momentum, pseudorapidity and multiplicity, these parameters would ideally be determined independently in each kinematic bin. However, the available MC statistics are not sufficient to perform such a differential fit. Instead, the results obtained from the integrated fit are used in all bins. Lastly, the fully reconstructed $\eta^0 \to \pi^+\pi^-\gamma$ contribution will be modeled with a Gaussian shape where both mean and width parameters will remain free in the fit to the data.

In the partially reconstructed sample, the kinematic information of the probe photon is not accessible, since this particle is not reconstructed. Consequently, the differential fit cannot be performed directly as a function of $p_T^{\gamma_{\rm prob}}$ and $\eta^{\gamma_{\rm prob}}$. Instead, the differential fit is performed as a function of the kinematic variables of the partially reconstructed candidate, $p_T^{\pi\pi\gamma}$ and $\eta^{\pi\pi\gamma}$, and the yields are obtained as a function of the probe photon variables from these results. A kinematic projection matrix derived from simulation is subsequently used to express these yields in the probe-photon phase space. The details of this procedure are discussed below. Some representative examples of the fits from the different bins of transverse momentum, pseudorapidity, and multiplicity are shown in

figure 8.12. The uncertainties in the invariant mass distribution were propagated from the background subtraction procedure. The results from the invariant mass fit of the η^0 and ω yields for the different bins of transverse momentum, pseudorapidity and multiplicity are shown in Table 8.9.

$p_T^{\pi\pi\gamma}$ (MeV/c) bin	$\eta^{\pi\pi\gamma}$ bin	$N_{\text{VeloTracks}}$ bin	$N_\omega^{\text{part(rec)}}$	$N_\eta^{\text{part(rec)}}$
[3200, 4500]	[2, 3.2]	[0, 180]	140015 ± 1516	33435 ± 306
[3200, 4500]	[2, 3.2]	[180, 420]	109608 ± 386	32811 ± 206
[3200, 4500]	[3.2, 4]	[0, 180]	68779 ± 739	16688 ± 222
[3200, 4500]	[3.2, 4]	[180, 420]	28917 ± 13	14678 ± 6
[4500, 6000]	[2, 3.2]	[0, 180]	72600 ± 614	27637 ± 250
[4500, 6000]	[2, 3.2]	[180, 420]	61299 ± 1643	26926 ± 342
[4500, 6000]	[3.2, 4]	[0, 180]	39290 ± 481	11793 ± 173
[4500, 6000]	[3.2, 4]	[180, 420]	32212 ± 518	9050 ± 169
[6000, 12000]	[2, 3.2]	[0, 180]	39027 ± 392	15816 ± 185
[6000, 12000]	[2, 3.2]	[180, 420]	51942 ± 504	15867 ± 195
[6000, 12000]	[3.2, 4]	[0, 180]	18617 ± 334	5132 ± 109
[6000, 12000]	[3.2, 4]	[180, 420]	18395 ± 344	4425 ± 115

Table 8.9: Results from the invariant mass fit of the η^0 and ω yields for the different bins of transverse momentum, pseudorapidity and multiplicity in the partially reconstructed case.

8.4.5 Kinematic projection matrices

Once the partially reconstructed yields of the η and ω mesons have been determined as functions of $p_T^{\pi\pi\gamma}$ and $\eta^{\pi\pi\gamma}$, the next step is to express these measurements in terms of the probe-photon kinematics, namely $p_T^{\gamma_{\text{prob}}}$ and $\eta^{\gamma_{\text{prob}}}$. To achieve this, a minimum-bias simulated sample containing 10^6 generated events is used to construct a kinematic projection matrix relating the kinematics of the probe photon to those of the partially reconstructed $\pi^+\pi^-\gamma_{\text{tag}}$ candidate. For a given pseudorapidity interval, this matrix is defined as

$$\begin{bmatrix} N(p_{T_0}^{\gamma_{\text{prob}}}) \\ N(p_{T_1}^{\gamma_{\text{prob}}}) \\ \vdots \\ N(p_{T_i}^{\gamma_{\text{prob}}}) \end{bmatrix} = \begin{bmatrix} c_{\gamma_0,(\pi\pi\gamma)0} & c_{\gamma_0,(\pi\pi\gamma)1} & \cdots & c_{\gamma_0,(\pi\pi\gamma)j} \\ c_{\gamma_1,(\pi\pi\gamma)0} & c_{\gamma_1,(\pi\pi\gamma)1} & \cdots & c_{\gamma_1,(\pi\pi\gamma)j} \\ \vdots & \vdots & \ddots & \vdots \\ c_{\gamma_i,(\pi\pi\gamma)0} & c_{\gamma_i,(\pi\pi\gamma)1} & \cdots & c_{\gamma_i,(\pi\pi\gamma)j} \end{bmatrix} \cdot \begin{bmatrix} N(p_{T_0}^{\pi\pi\gamma}) \\ N(p_{T_1}^{\pi\pi\gamma}) \\ \vdots \\ N(p_{T_j}^{\pi\pi\gamma}) \end{bmatrix}, \tag{8.17}$$

$$c_{\gamma_k,(\pi\pi\gamma)l} = \frac{N(p_{T_k}^{\gamma_{\text{prob}}}, p_{T_l}^{\pi\pi\gamma})}{\sum_{l=0}^{j} N(p_{T_k}^{\gamma_{\text{prob}}}, p_{T_l}^{\pi\pi\gamma})}, \tag{8.18}$$

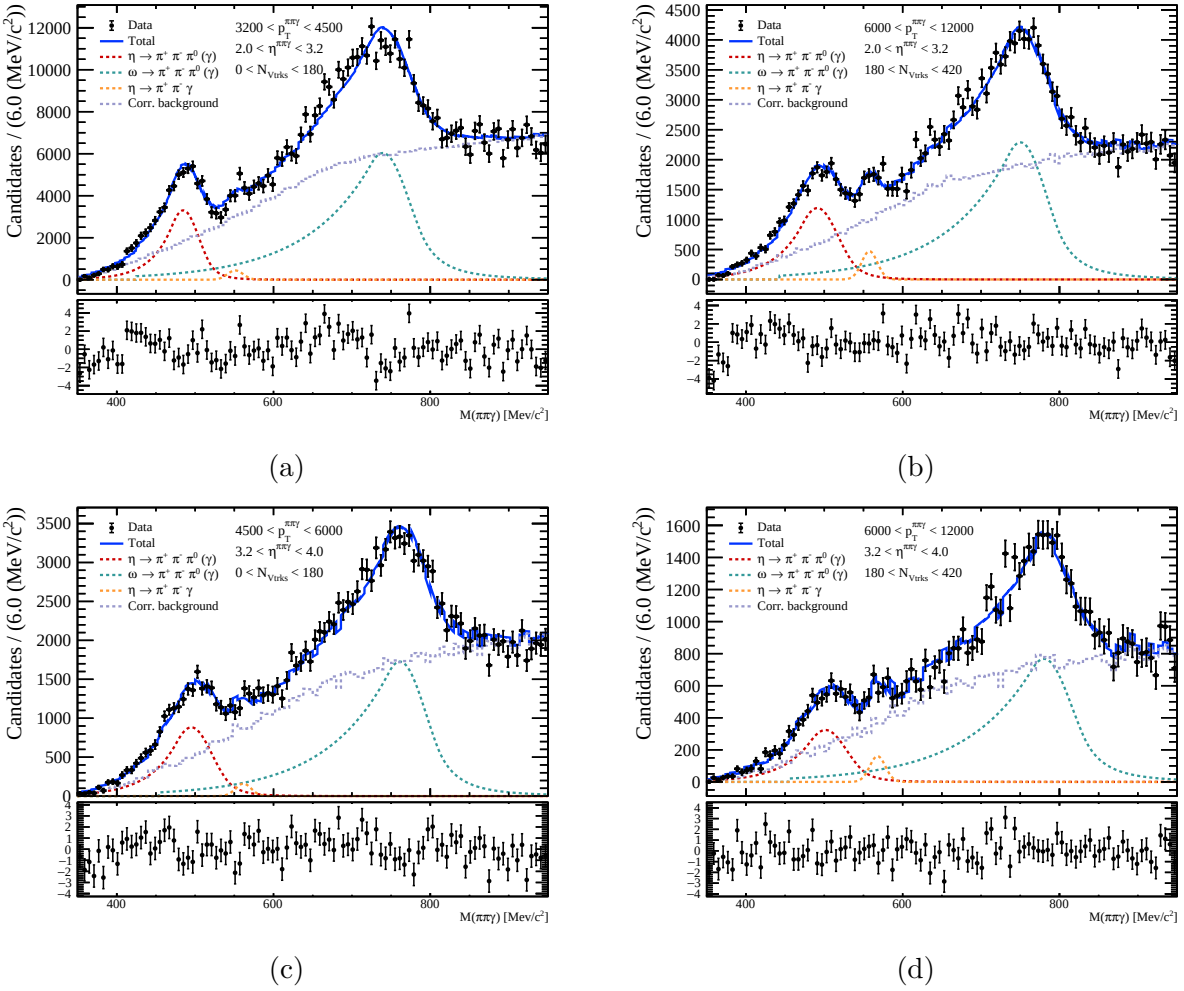


Figure 8.12: Examples of fits to the $\pi\pi\gamma$ invariant mass for different bins of transverse momentum, pseudorapidity and multiplicity in the 2025 OO data sample. The (a) and (b) fits are cases of low pseudorapidity, while (c) and (d) are at high pseudorapidity. In addition, (a) and (c) fits are cases of low multiplicity, while (b) and (d) are high multiplicity.

The matrix is normalized such that

$$\sum_l c_{\gamma_k,(\pi\pi\gamma)l} = 1, \tag{8.19}$$

for every probe-photon transverse momentum interval $p_{T_k}^{\gamma_{\text{prob}}}$. For a given pseudorapidity interval, the kinematic projection matrix relates the distributions in $p_T^{\pi\pi\gamma}$ and $p_T^{\gamma_{\text{prob}}}$. This means that we have to build a kinematic projection matrix for each pseudorapidity bin and for each particle. In total, four kinematic projection matrices are constructed, corresponding to the two pseudorapidity intervals and the two particle species. The four kinematic projection matrices are shown in Figures 8.13 and 8.14. Applying these matrices to the partially reconstructed yields listed in Table 8.9 provides the corresponding yields as functions of the probe-photon kinematics, which can then be used in the determination of the photon detection efficiency.

8.4.6 Results of the photon detection efficiency

During the signal extraction of the partially reconstructed η^0 and ω signals, we have seen that the procedure is subject to sizeable systematic uncertainties. To reduce their impact as much as possible, the η^0 and ω signals are combined into a single calibration sample, instead of determining the photon detection efficiency independently for each particle. Taking also into account the kinematic projection matrices (denoted as C_p, with $p \in \{\eta^0, \omega\}$), equation 8.13 can be rewritten as

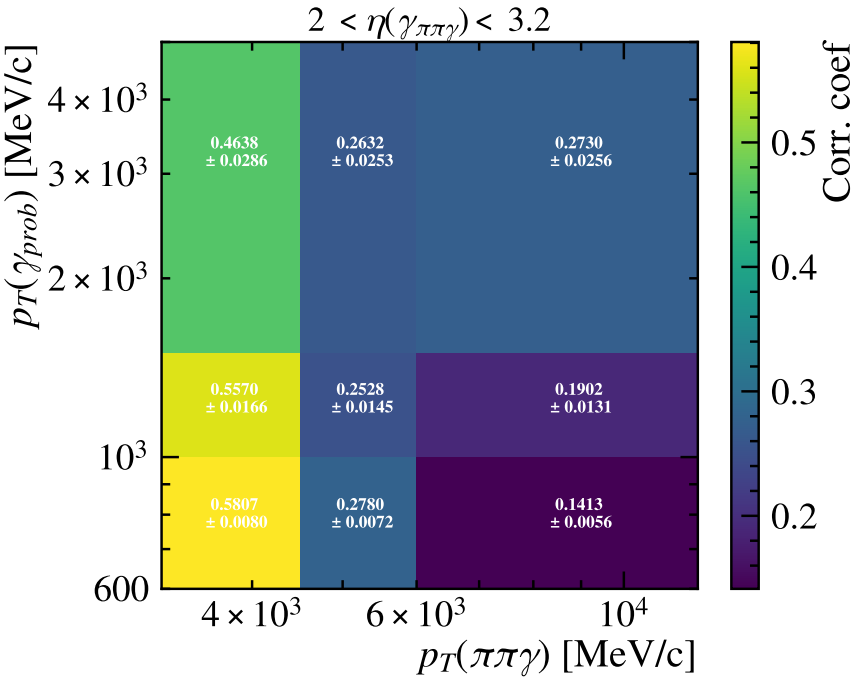
$$\frac{N_\omega^{\text{full(rec)}} + N_\eta^{\text{full(rec)}}}{C_\omega \cdot N_\omega^{\text{part(rec)}} + C_\eta \cdot N_\eta^{\text{part}}} = A^\gamma \times \varepsilon_{\text{det}}^\gamma. \tag{8.20}$$

The ratio on the left-hand side of equation 8.20 provides a direct measurement of the product of the photon acceptance and detection efficiency. The result is shown in figure 8.15. However, in order to obtain an estimate of the photon detection efficiency itself, acceptance effects have to be taken into account.

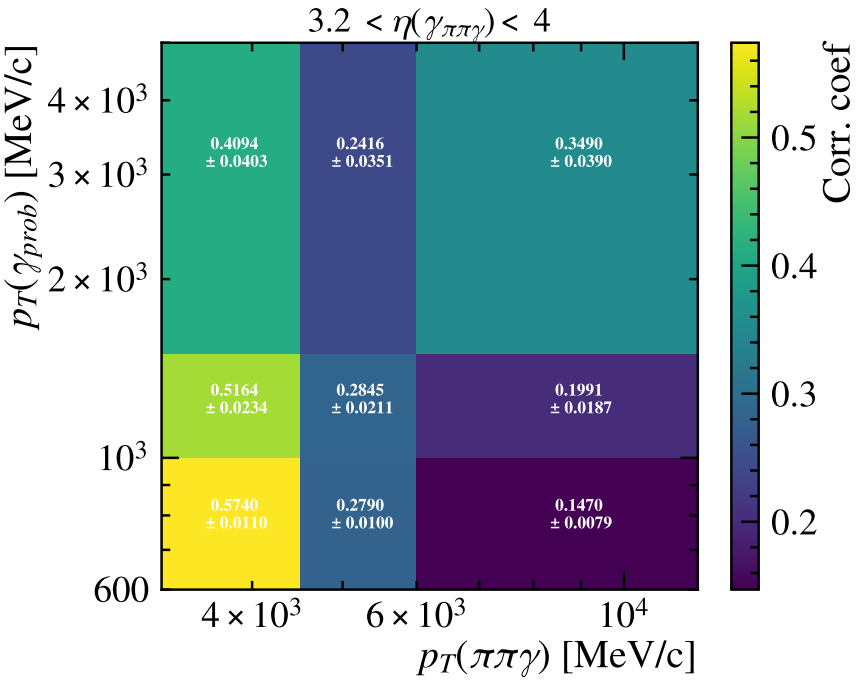
The photon acceptance factor A^γ represents the probability that, given one photon from an η^0 or ω decay lying in a given p_T and η interval, the second photon is reconstructed within the fiducial acceptance of LHCb. It is defined as

$$A^\gamma(p_T^\gamma, \eta^\gamma) = \frac{(\gamma_1 \text{ in } p_T \text{ and } \eta \text{ bin}) \,\&\, (\gamma_2 \text{ in LHCb acceptance})}{(\gamma_1 \text{ in } p_T \text{ and } \eta \text{ bin})}. \tag{8.21}$$

In equation 8.21, a photon is considered to be within fiducial acceptance if $2 < \eta < 4$, corresponding to the pseudorapidity requirement applied throughout this analysis. The

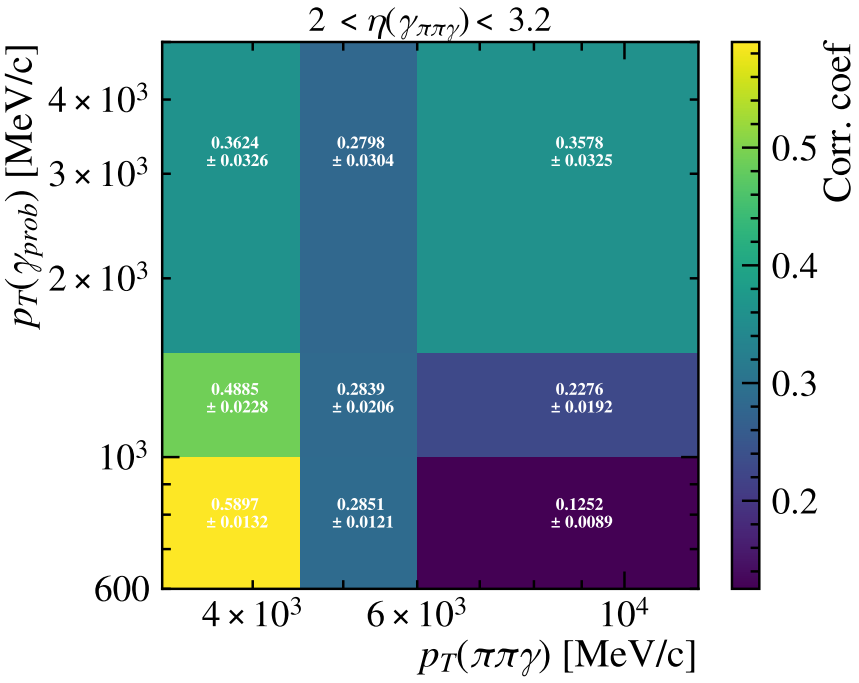


(a) Low pseudorapidity bin.

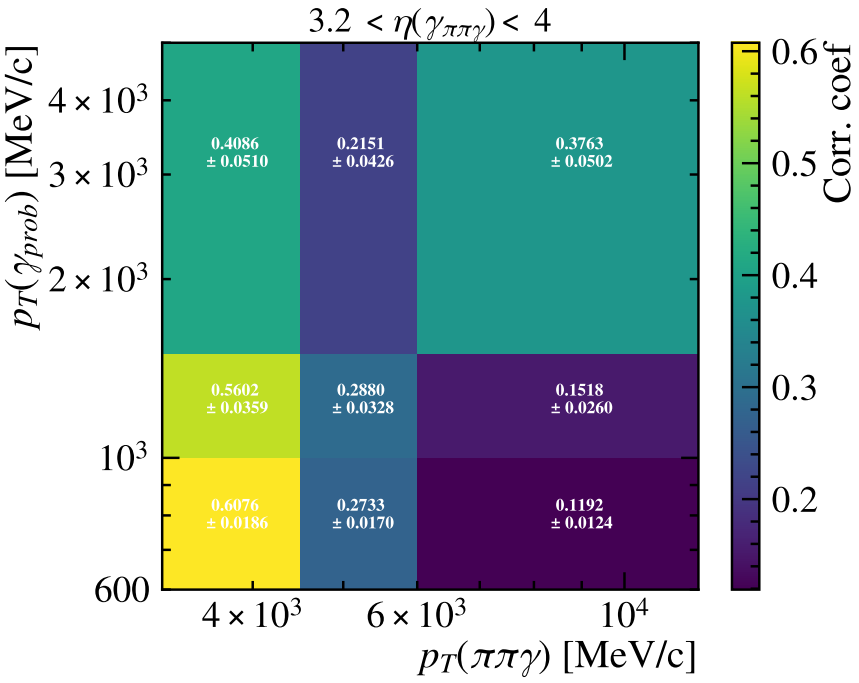


(b) High pseudorapidity bin.

Figure 8.13: Kinematic projection matrices relating the $p_T^{\pi\pi\gamma}$ and $p_T^{\gamma_{\text{prob}}}$ distributions for the two pseudorapidity intervals of the ω meson using a generator-level OO MC simulation.



(a) Low pseudorapidity bin.



(b) High pseudorapidity bin.

Figure 8.14: Kinematic projection matrices computed for each bin of pseudorapidity of η^0 using a generator-level OO MC simulation.

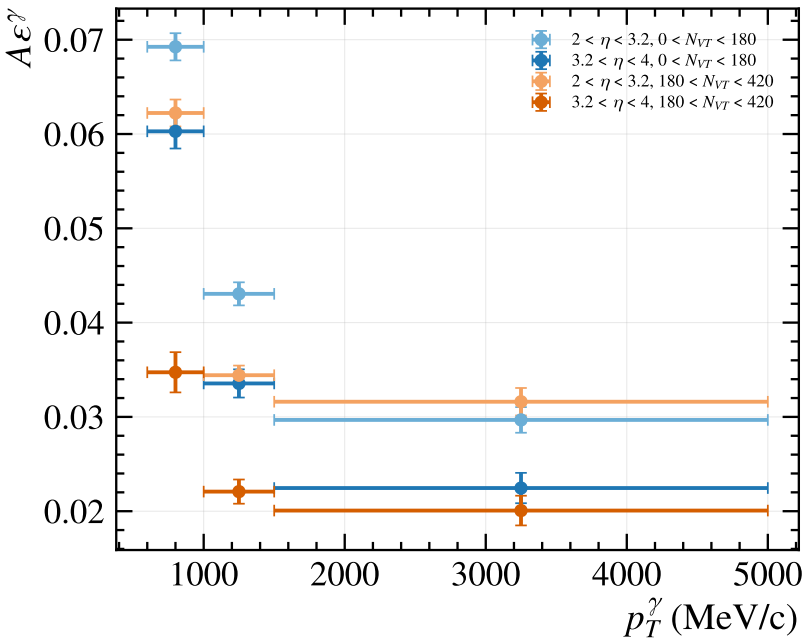


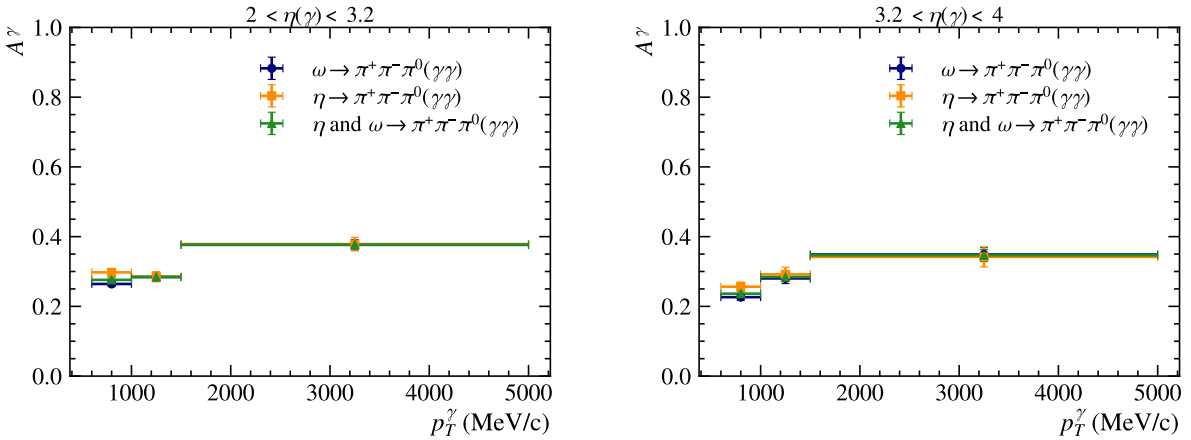
Figure 8.15: Acceptance times photon detection efficiency as a function of the photon transverse momentum for different bins of pseudorapidity and multiplicity in the 2025 OO data sample.

condition "in p_T and η bin" refers to the transverse momentum and pseudorapidity intervals used in the efficiency determination.

The acceptance factor is computed using the generated minimum-bias sample produced without acceptance requirements. All particles except the photons from the η^0 and ω decays are required to satisfy the same transverse momentum and pseudorapidity requirements as in data. Since the photon detection efficiency is determined using the combined η^0 and ω samples, the acceptance factor is also computed using photons originating from both decays. The results are shown in figure 8.16. The acceptance factors obtained using η^0, ω, and the combination of both decays are compatible within statistical uncertainties.

Finally, the η^0 and ω yields extracted from the fits to the fully and partially reconstructed samples, together with the acceptance factor computed from the generated MC sample, are used to estimate the photon detection efficiency. The final results, presented in figure 8.17, show a decreasing trend of the efficiency with increasing transverse momentum. This behavior is expected, since energetic photons are more likely to merge with nearby photons from π^0 decays.

Comparing the two pseudorapidity intervals, we observe lower efficiencies at larger pseudorapidities. This behavior is consistent with the degraded performance of the ECAL inner region, which is more affected by ageing effects. Finally, lower efficiencies are also



(a) Low pseudorapidity bin.

(b) High pseudorapidity bin.

Figure 8.16: Photon acceptance factor as a function of the transverse momentum in the low (a) and high(b) pseudorapidity bins from a generator-level OO MC simulation. The results using η^0, ω and both decays at the same time are presented separately.

observed at high multiplicities, indicating that the occupancy of the ECAL impacts the photon reconstruction performance.

These efficiency corrections will be applied to the measured χ_c yields to determine the fraction of J/ψ mesons originating from χ_c decays.

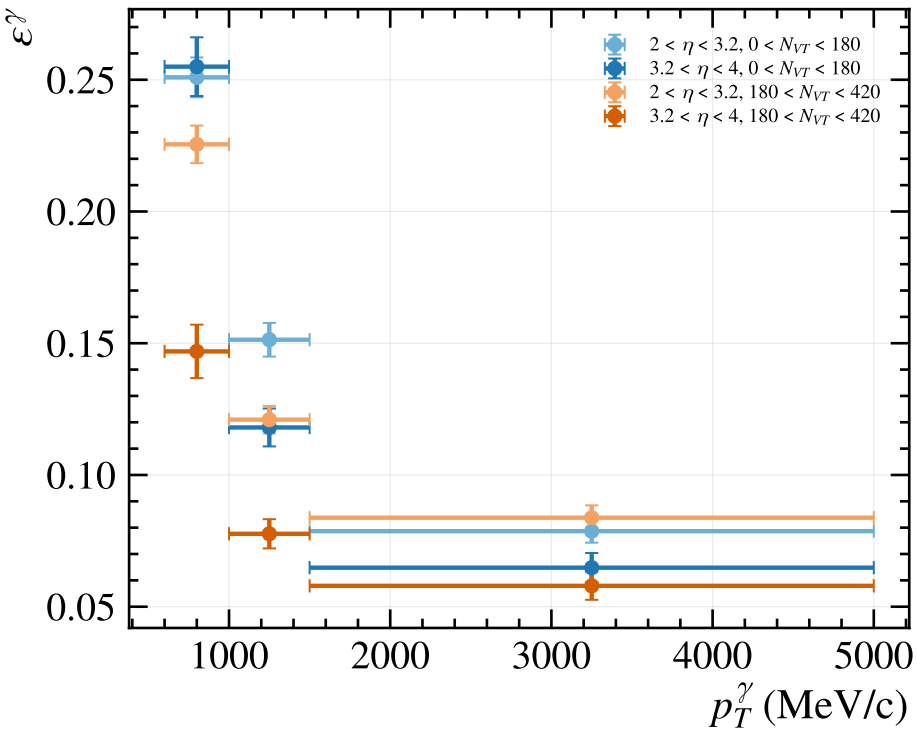


Figure 8.17: Photon detection efficiency as a function of the photon transverse momentum for different bins of pseudorapidity and multiplicity in the 2025 OO data sample.

8.5 Systematic Uncertainties and prospects

Due to time limitations, not all sources of systematic uncertainty have been evaluated in this work. The systematic uncertainties considered in this thesis are those associated with the χ_c signal extraction and the possible polarization of the χ_c states. Other sources of uncertainty, such as the BDT selection, the J/ψ signal extraction and the photon detection efficiency, were not evaluated within the scope of this work and are left for future studies.

8.5.1 Signal extraction

The systematic uncertainty associated with the χ_c signal extraction was introduced in Section 7.3.4. Several signal parameterizations and event-mixing configurations were considered in order to study the stability of the extracted χ_c yields. The systematic uncertainty is estimated as the RMS of the yields obtained with all the tested models.

For the determination of the cross-section ratio $\sigma(\chi_{c2})/\sigma(\chi_{c1})$, this is the only source of systematic uncertainty considered, since most of the remaining uncertainties cancel in the ratio.

In the determination of the fraction of J/ψ mesons originating from χ_c feed-down, the systematic uncertainty associated with the J/ψ signal extraction also contributes. As discussed in Section 7.4, no dedicated study of the J/ψ signal extraction systematic uncertainty was performed. Instead, the nominal J/ψ yield obtained from the fit was kept fixed, while the fraction of J/ψ mesons coming from χ_c decays was recomputed for each of the tested χ_c fit models. Again, the RMS of the obtained results was taken as an estimate of the associated systematic uncertainty. The results of this study are presented together with the final feed-down measurements in the following section.

8.5.2 χ_c polarization

The effect of χ_c polarization on the detection efficiency was first discussed in the LHCb measurement performed in pp collisions [71]. At that time, neither the J/ψ polarization nor the polarization states of χ_{c1} and χ_{c2} were experimentally constrained. The J/ψ polarization parameter $\lambda_{J/\psi}$ was later measured by LHCb [39], while CMS reported evidence for a sizeable polarization of χ_{c1} and χ_{c2} for $p_T(J/\psi) > 8$ GeV/c [73]. Both measurements were performed in the helicity frame, where the polarization axis is chosen along the direction of the quarkonium momentum in the centre-of-mass frame of the colliding beams, and the decay angles are defined in the quarkonium rest frame with respect to this axis.

These results were used in [62] to construct a toy Monte Carlo model for estimating the systematic uncertainty associated with χ_c polarization. Two extreme polarization scenarios were considered, corresponding to the azimuthal angular-momentum quantum numbers $m_{\chi_{cJ}} = 0, 1$ for χ_{c1} and $m_{\chi_{cJ}} = 0, 2$ for χ_{c2}. In addition, the value of $\lambda_{J/\psi}$ was varied within one standard deviation of the measurement reported in [39]. The purpose of this study was not to predict the physical χ_c polarization, but rather to estimate the possible change in acceptance if the true decay angular distribution differs from that assumed in the unpolarized simulation.

In the present analysis, the results of the χ_c polarization studies reported in [62] are used to compute an integrated correction factor by performing a weighted average over all transverse-momentum bins considered in the pPb analysis. The weighted correction factor is defined as

$$\langle C_{\mathrm{pol}} \rangle = \frac{\sum_i N_i(J/\psi \leftarrow \chi_c), C_i^{\mathrm{pPb}}}{\sum_i N_i(J/\psi \leftarrow \chi_c)}, , \tag{8.22}$$

where $N_i(J/\psi \leftarrow \chi_c)$ denotes the number of J/ψ mesons originating from χ_c decays and C_i^{pPb} is the polarization correction factor obtained in the pPb analysis, both evaluated in the i-th bin of J/ψ transverse momentum. The resulting integrated correction factors for the two extreme polarization scenarios are listed in Table 8.10.

The largest deviation from unity is taken as the systematic uncertainty associated with

Scenario	Correction factor
χ_{c1} max. polarization	1.038 ± 0.013
χ_{c2} max. polarization	1.062 ± 0.013

Table 8.10: Integrated polarization correction factors.

χ_c polarization. Consequently, a systematic uncertainty of 6.2% is assigned and added in quadrature to the other systematic uncertainties considered in the analysis.

8.5.3 Systematic uncertainty associated with the BDT selection

An ideal BDT model would perfectly separate signal from background, without any misidentification between the two classes. In practice, as discussed in Section 7.2.2.5, the BDT selection is not perfect and therefore introduces an additional source of systematic uncertainty.

This uncertainty was not evaluated within the scope of this work. A possible strategy to estimate it would consist of varying different aspects of the BDT model, including the set of input variables, the hyperparameters, and the selection threshold. For each variation, the complete analysis chain would have to be repeated and the corresponding χ_c yields compared with the nominal result.

8.5.4 Systematic uncertainty associated with the photon detection efficiency

As discussed in Section 8.4, the determination of the photon detection efficiency is affected by sizeable uncertainties, mainly associated with the signal extraction in the partially reconstructed sample.

This source of uncertainty was not evaluated within the scope of this work. A possible approach to estimate it would consist of varying the photon efficiencies within their uncertainties and studying the corresponding variation of the fraction of J/ψ mesons originating from χ_c feed-down.

9 Results and Discussion

Having presented all the ingredients of the analysis, this chapter combines them to derive the final observables of interest. In particular, two quantities are determined: the prompt-production cross-section ratio $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ and the fraction of prompt J/ψ mesons originating from χ_c feed-down, $F(\chi_c \to J/\psi)$. The details of the computation of these observables are presented in the following sections.

9.1 Cross-section ratio $\sigma(\chi_{c2})/\sigma(\chi_{c1})$

As discussed in section 4.5, the binding energies of χ_c states are significantly smaller than that of the J/ψ state and greater than the binding energy of the $\psi(2S)$ state. The small difference in the binding energies of χ_{c1} and χ_{c2} charmonia makes the ratio of their production cross-sections, $\sigma(\chi_{c2})/\sigma(\chi_{c1})$, a useful tool to study their sensitivity to final-state nuclear effects, which are expected to be similar for both states. Furthermore, the χ_c states contribute significantly to the feed-down component of prompt J/ψ production, so measurements of nuclear effects on χ_c states can clarify the interpretation of the J/ψ data. The computation of this ratio is especially convenient, as various efficiency factors and sources of uncertainty cancel out in the ratio, allowing for a more precise measurement.

The prompt-production cross-section ratio is defined as the ratio of the cross sections for prompt χ_{c1} and χ_{c2} production. The cross-section ratio is defined as

$$\frac{\sigma(\chi_{c2})}{\sigma(\chi_{c1})} = \frac{N(\chi_{c2})}{N(\chi_{c1})} \cdot \frac{\varepsilon(\chi_{c1})}{\varepsilon(\chi_{c2})} \cdot \frac{\mathcal{B}(\chi_{c1} \to J/\psi\gamma)}{\mathcal{B}(\chi_{c2} \to J/\psi\gamma)}, \tag{9.1}$$

where $N(\chi_{c1})$ and $N(\chi_{c2})$ represent the signal yields of the χ_{c1} and χ_{c2} states, respectively,

and $\varepsilon(\chi_{c1})$ and $\varepsilon(\chi_{c2})$ denote the efficiencies to reconstruct and select the corresponding state. Lastly, $\mathcal{B}(\chi_c \to J/\psi\gamma)$ represents the branching ratios for the χ_c states decays. The values of these branching ratios are listed in table 7.1 and are taken from [28].

After considering the efficiency cancellations in the ratio, the efficiency term can be written as

$$\frac{\varepsilon(\chi_{c1})}{\varepsilon(\chi_{c2})} = \frac{\varepsilon_{acc}(\chi_{c1})}{\varepsilon_{acc}(\chi_{c2})} \cdot \frac{\varepsilon^{\text{offline}}_{\text{rec\&sel}}(\chi_{c1})}{\varepsilon^{\text{offline}}_{\text{rec\&sel}}(\chi_{c2})} \cdot \frac{\varepsilon^{\text{BDT}}_{\text{sel}}(\chi_{c1})}{\varepsilon^{\text{BDT}}_{\text{sel}}(\chi_{c2})}. \tag{9.2}$$

The details of the computation of these efficiencies were described in the previous chapter and the obtained values are listed in tables 8.1 and 8.3. Finally, the yield ratio is computed in section 7.3.4 with its associated systematic uncertainty due to the signal extraction procedure. The value of the yield ratio $N(\chi_{c2})/N(\chi_{c1})$ is taken from 7.17.

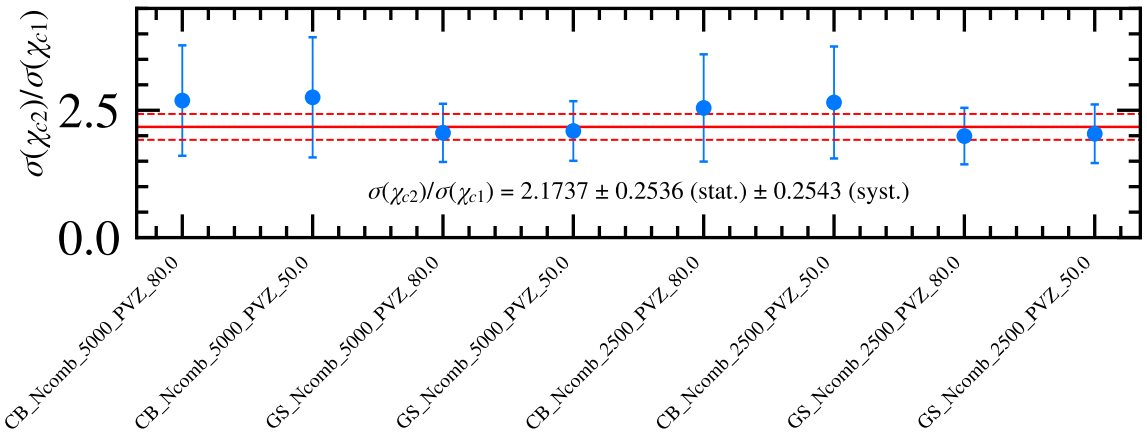


Figure 9.1: Spread of results of $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ in OO collisions. The solid red line represents the weighted mean, and the red dashed lines show the RMS, which is taken as the systematic uncertainty.

Following the same procedure we applied before to account for the systematic uncertainty, the result of the cross-section ratio is presented in figure 9.1. The final result for the integrated prompt-production cross-section ratio $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ in OO collisions at $\sqrt{s_{NN}} = 5.36$ TeV is

$$\frac{\sigma(\chi_{c2})}{\sigma(\chi_{c1})} = 2.17 \pm 0.25(11.67\%\ \text{stat.}) \pm 0.25(11.70\%\ \text{syst.}) \tag{9.3}$$

This result, computed using calorimetric photons in the kinematic range $p_T(\chi_c) > 1.5$ GeV/c and $2 < \eta(\chi_c) < 4.8$, is measured with comparable statistical and systematic uncertain-

ties. The statistical and systematic uncertainties contribute almost equally to the total uncertainty. The cross-section ratio obtained as a result of this work can be compared with the same cross-section ratio in other collision systems. Here, it will be compared with the corresponding ratio measured in pp collisions at $\sqrt{s} = 7$ TeV using converted photons by the LHCb collaboration [109] and with the measurement in pPb collisions at $\sqrt{s_{NN}} = 8.16$ TeV using both converted and calorimetric photons also by LHCb [111]. The comparison as a function of the rapidity is shown in figure 9.2.

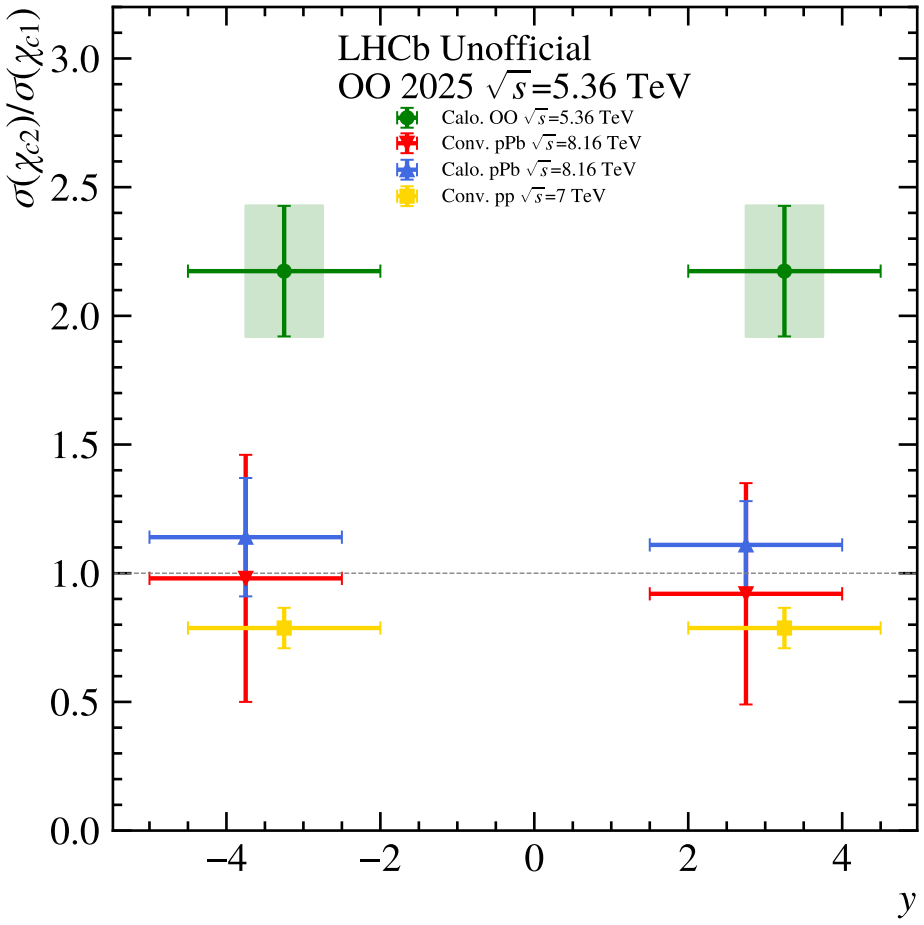


Figure 9.2: Cross-section ratio, $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ as a function of the rapidity, for the χ_{c1} and χ_{c2} promptly produced in OO collisions. The OO data is compared with the results of the converted sample in pp [109] and with the converted and calorimetric samples in pPb [111]. Green error bars correspond to statistical uncertainty, and the green shaded region corresponds to systematic uncertainty. The rest of the error bars correspond to total uncertainties.

Both OO and pp measurements were made in an integrated forward rapidity range $2 < y < 4.5$, but the points for these collision systems at backward rapidity were mirrored with

respect to $y = 0$ to be compared with pPb data at this rapidity. This comparison must be made carefully, taking into account that the different measurements are integrated in different transverse momentum ranges. On the one hand, the result in pp is integrated in a transverse momentum of the J/ψ $3 < p_T(J/\psi) < 20$ GeV/c. On the other hand, the results of pPb using converted photons are within a transverse momentum range $3 < p_T(\chi_c) < 15$ GeV/c, while the results of pPb using calorimetric photons are within a transverse momentum range $5 < p_T(\chi_c) < 15$ GeV/c.

Whereas the pPb measurements are compatible with the result obtained in pp collisions, the value measured in OO collisions tends to be larger. As mentioned, this comparison should be interpreted with caution, since the measurements are integrated over different transverse momentum ranges. In particular, the present analysis extends down to $p_T(\chi_c) > 1.5$ GeV/c, while previous measurements were restricted to higher transverse momenta. The dependence of the cross-section ratio on transverse momentum observed in pp collisions [109] suggests that lower values of p_T favor larger values of $\sigma(\chi_{c2})/\sigma(\chi_{c1})$, which could naturally explain the enhancement observed in OO collisions.

The χ_{c2}/χ_{c1} production ratio was also investigated using simulated events. Samples of χ_{c1} and χ_{c2} production in OO collisions at $\sqrt{s_{NN}} = 5.36$ TeV were generated with Pythia Angantyr [130] within the fiducial region defined by $1.5 < p_T < 10$ GeV/c and $2 < y < 4.5$. A total of 1600 χ_{c2} and 585 χ_{c1} mesons were generated in this region, corresponding to a production ratio of $N(\chi_{c2})/N(\chi_{c1}) = 2.74 \pm 0.13$. This prediction is consistent with the value measured in this analysis within 1.5σ, indicating no statistically significant discrepancy between the simulation and the experimental result. The observed agreement provides confidence that the simulation reproduces the relative production of χ_{c1} and χ_{c2} mesons in the kinematic region considered.

9.2 Fraction of χ_c decays to J/ψ

The primary motivation of this work is to constrain the feed-down contribution from χ_c states to prompt J/ψ production in a heavy-ion environment (see Sec. 4.5.3.3). The measurement of J/ψ production as a probe of the QGP produced in heavy-ion collisions is especially difficult to interpret, since the experimentally measured prompt J/ψ signal does not consist exclusively of directly produced J/ψ mesons. One of the most significant yet least known contributions to the J/ψ production in ion collisions is the feed-down from χ_c states. The observable carrying information about this contribution is the fraction of prompt J/ψ mesons originating from χ_c decays. Following the previous analysis performed in pPb collisions at LHCb [62], this fraction is obtained from

$$F(\chi_c \to J/\psi) \equiv \frac{\sigma(\chi_c \to J/\psi\gamma)}{\sigma(J/\psi)} = \frac{N(\chi_c \to J/\psi\gamma)}{N(J/\psi) \cdot \varepsilon_{\chi_c/J/\psi}}, \tag{9.4}$$

where $N(\chi_c \to J/\psi\gamma)$ and $N(J/\psi)$ are the prompt $\chi_c \to J/\psi\gamma$ and J/ψ yields, and $\varepsilon_{\chi_c/J/\psi}$ is the fraction of $\chi_c \to J/\psi\gamma$ decays which are detected in the LHCb acceptance relative to the decays where only the J/ψ is detected. Here, the measured χ_c yield is the sum of the χ_{c1} and χ_{c2} yields, as the contribution from χ_{c0} is negligible due to its small branching ratio. Combining both χ_c states in the determination of the feed-down fraction reduces the uncertainty associated with their separation. The prompt yields are extracted in sections 7.3 and 7.4. The $\varepsilon_{\chi_c/J/\psi}$ term takes into account all the acceptance and selection effects that affect the prompt χ_c and J/ψ measurements. The only differences between the prompt χ_c and prompt J/ψ measurements arise from the photon acceptance, photon detection efficiency, and BDT selection, so this term can be expressed as

$$\varepsilon_{\chi_c/J/\psi} = \text{Acc.}(\gamma \text{ from } \chi_c) \times \varepsilon^{\gamma}_{\text{det}} \times \varepsilon^{\text{BDT}}_{\text{sel}}. \tag{9.5}$$

The photons from χ_c acceptance Acc.(γ from χ_c) value is taken from Table 8.2, while the BDT selection efficiency is taken from Table 8.3. As the photon detection efficiency $\varepsilon^{\gamma}_{\text{det}}$ was computed differentially as a function of the transverse momentum, pseudorapidity, and multiplicity, it cannot be applied directly to the results of the integrated yields. Instead, the correction is applied using the sPlot technique. Using the χ_c signal sWeights, the yields are determined in bins of transverse momentum, pseudorapidity and multiplicity. Then, these yields are corrected by the photon detection efficiency in the corresponding bin. The final corrected integrated yield is the sum of the yields in all the bins. Following the same procedure we applied before to account for the systematic uncertainty, the result of the fraction of J/ψ coming from χ_c is presented in figure 9.3.

The final results for the integrated fraction of J/ψ coming from χ_c feed-down in OO collisions at $\sqrt{s_{NN}} = 5.36$ TeV are

$$\begin{aligned} F(\chi_{c1} \to J/\psi) &= 0.0510 \pm 0.0036\ (7.06\%\ \text{stat.}) \pm 0.0004\ (0.78\%\ \text{syst.}) \\ F(\chi_{c2} \to J/\psi) &= 0.073 \pm 0.004\ (5.48\%\ \text{stat.}) \pm 0.010\ (13.70\%\ \text{syst.}) \\ F(\chi_{c} \to J/\psi) &= 0.124 \pm 0.004\ (3.23\%\ \text{stat.}) \pm 0.010\ (8.06\%\ \text{syst.}) \end{aligned} \tag{9.6}$$

From these results, we see that the χ_{c2} state has a larger contribution to J/ψ feed-down than the χ_{c1} state. However, the fraction result for χ_{c2} is the most affected by the systematic uncertainty. The same measurement was performed by LHCb in pp collisions at $\sqrt{s} = 7$ TeV differentially as a function of the J/ψ transverse momentum [131]. That analysis does not have an integrated measurement, so it is difficult to compare with the current result in OO. Even so, the result in every bin of $p_T(J/\psi)$ in the analysis in pp is above the integrated result in OO. Other measurements of the integrated fraction of J/ψ coming from χ_c in nucleus collisions were done by HERA-B and PHENIX experiments. HERA-B measured this fraction in proton on C, Ti, and W targets at $\sqrt{s_{NN}} = 41.6$ GeV [132], showing no dependence on the target material. PHENIX measured the same

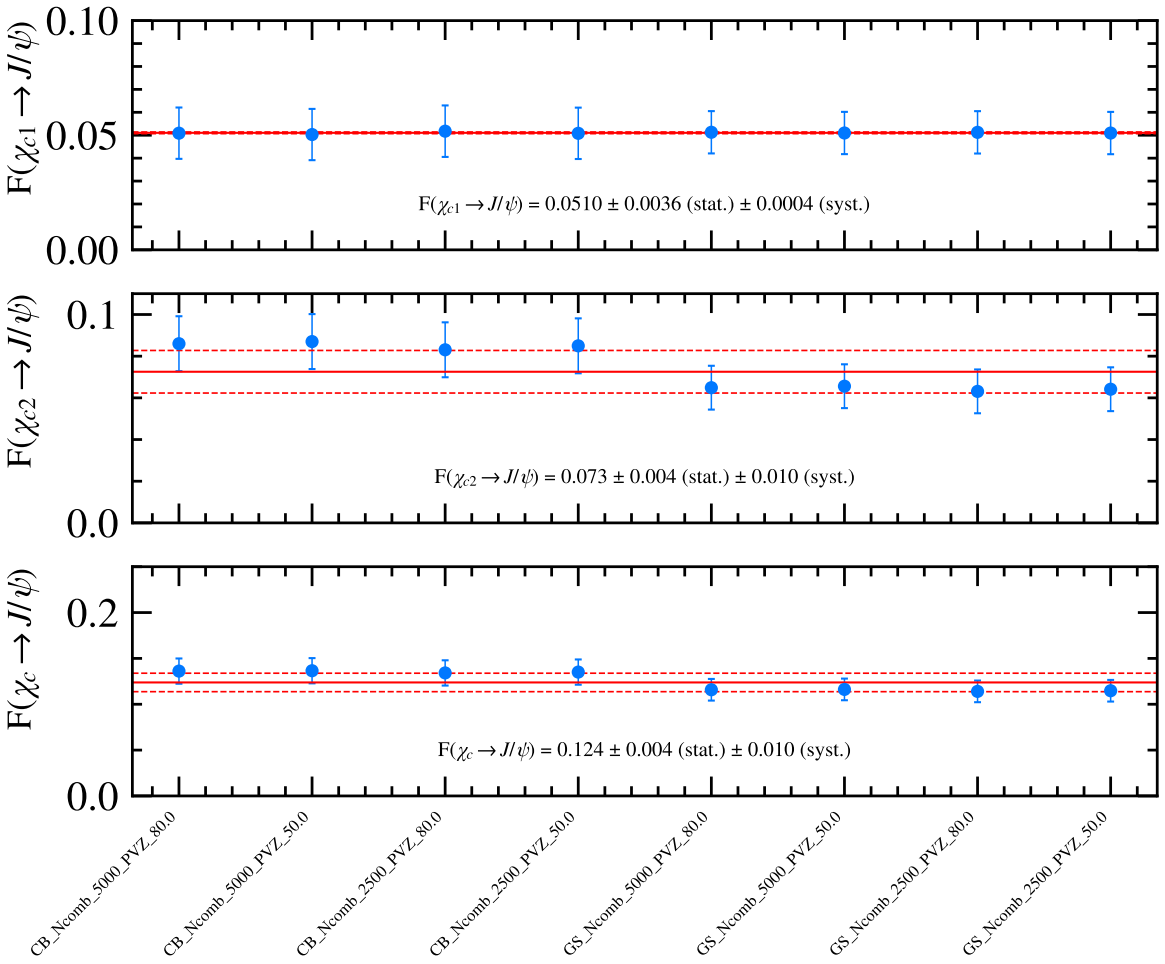


Figure 9.3: Spread of results of $F(\chi_{c1} \to J/\psi)$, $F(\chi_{c2} \to J/\psi)$ and $F(\chi_c \to J/\psi)$ in OO collisions. The solid red line represents the weighted mean, and the red dashed lines show the RMS, which is taken as the systematic uncertainty.

fraction in dAu collisions at $\sqrt{s_{NN}} = 200$ GeV [133], consistent with measurements in pp collisions at the same energy within the large statistical uncertainty.

Figure 9.4 shows a comparison of the fraction of prompt J/ψ mesons originating from χ_c decays for different collision systems. Here, we focus on the measurements performed at LHC energies (pp, pPb and OO at LHCb), where the conditions required for the formation of a QGP medium can be fulfilled. The $p_T(J/\psi)$-dependent results obtained in pPb collisions were found to be compatible with those measured in pp collisions, suggesting that neither directly produced J/ψ mesons nor χ_c states experience a sizeable suppression in the medium created in pPb collisions.

The present analysis constitutes the first reconstruction and measurement of χ_c mesons

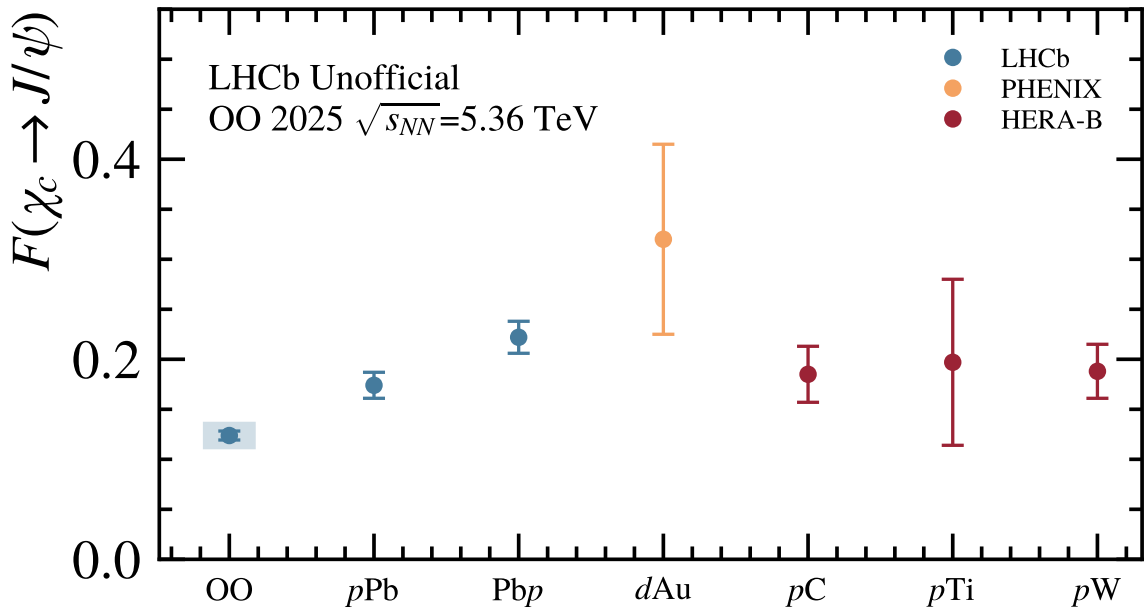


Figure 9.4: Fraction of J/ψ from χ_c for the different collision systems at the different experiments. The shaded box of the OO point represent the current systematic uncertainty.

in nucleus-nucleus collisions at collider energies. The current OO result lies below all previous measurements performed in pp and pPb collisions and is lower than the pPb result by about 2.9σ when considering the presently available uncertainties. However, a quantitative comparison is complicated by the different transverse momentum ranges covered by the various measurements. In addition, several sources of systematic uncertainty have not yet been evaluated in the present analysis.

Nevertheless, the analysis chain is complete and already allows the extraction of the χ_{c1} and χ_{c2} signals, the determination of their production ratio, and an estimate of the fraction of prompt J/ψ mesons originating from χ_c feed-down. If confirmed after including the remaining systematic uncertainties, the observed reduction of the feed-down fraction in OO collisions could be interpreted as evidence for an enhanced suppression of χ_c states in the medium produced in OO collisions. Regardless of the final value obtained after the inclusion of the remaining systematic uncertainties, this work demonstrates for the first time the feasibility of performing precision measurements of χ_c production in a heavy-ion environment at collider energies. It establishes the experimental tools required to study χ_c suppression in nucleus-nucleus collisions and opens the way to future measurements with larger heavy-ion datasets collected by LHCb.

Part II

Tracking performance for LHCb Upgrade II

10

LHCb Upgrade II

The second part of this thesis shifts focus away from the physics measurement presented above to address a complementary, yet equally essential, aspect of experimental high-energy physics: the development of the detectors that will enable future measurements. Rather than analyzing data from an existing detector, this part looks forward to LHCb Upgrade II.

The LHCb Upgrade II represents the next major evolution of the LHCb experiment and is designed to operate under the challenging conditions of the High-Luminosity LHC (HL-LHC). Although the LHCb experiment will operate with luminosity leveling, at an instantaneous luminosity lower than the maximum that the HL-LHC can deliver, the resulting detector occupancies will still be extremely high. Its primary objective is to fully exploit the flavor-physics potential of the HL-LHC by collecting an integrated luminosity of at least 300 fb^{-1} at a peak instantaneous luminosity of $(1 - 1.5) \times 10^{34}$ $\text{cm}^{-2}\text{s}^{-1}$. Maintaining the detector performance under these conditions requires the introduction of new detector technologies, including precision timing capabilities and radiation-hard sensor architectures.

The Upgrade II physics program is equally ambitious, aiming at unprecedented precision in CP-violation measurements, the study of extremely rare decays, and investigations of hadron spectroscopy and heavy-ion physics. A complete description of the detector design and the associated physics program can be found in the LHCb Upgrade II Scoping Document [134].

Particular emphasis is placed in this thesis on the tracking system and, more specifically, on the Upstream Pixel detector, whose performance in the high-occupancy environment expected in heavy-ion collisions is the main subject of the studies presented in the following

chapters.

10.1 Upgrade II Frame

The LHCb Upgrade II is the second major phase of advancement for the LHCb experiment, positioned as a critical component of the European Strategy for Particle Physics to maximize the scientific program of the HL-LHC [135–137]. The primary vision for Upgrade II is to provide the experiment with the sensitivity required to probe physics beyond the SM. This is dictated by the evolving landscape of high-energy physics, where precision measurements in the flavor sector complement direct searches at the energy frontier.

The HL-LHC project aims to increase the delivered luminosity of the LHC by a factor of ten beyond its original design value. For the LHCb experiment, this transition represents a fundamental shift in the operational environment. While the initial LHCb upgrades (Upgrade I) prepared the detector for an instantaneous luminosity of 2×10^{33} cm^{-2}s^{-1}, Upgrade II must handle nearly an order of magnitude more. This extreme environment is characterized by a significantly higher pile-up in regular proton-proton collisions, which complicates track reconstruction and vertex identification. As represented in Fig. 10.1, the peak luminosity targeted is 1.5×10^{34} cm^{-2}s^{-1}, which translates to a pile-up of approximately 40 interactions per bunch crossing.

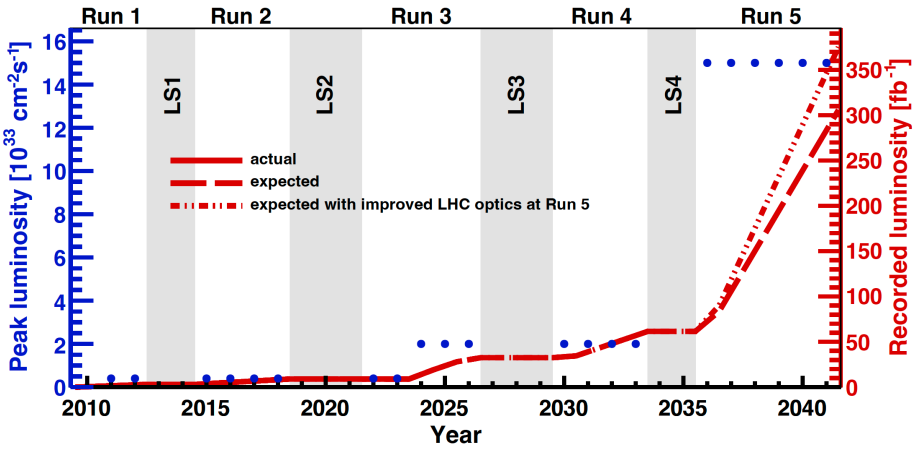


Figure 10.1: Instantaneous (peak) and integrated (recorded) luminosity profiles for the original LHCb (Runs 1 and 2), Upgrade I (Runs 3 and 4) and Upgrade II (Run 5) detectors, in the context of the LHC and HL-LHC schedule. Figure taken from [134] under Creative Commons license.

The upgraded detector is scheduled for installation during Long Shutdown 4 (LS4), currently projected for the mid-2030s. It is expected to operate at the aforementioned peak luminosity, enabling the accumulation of a total integrated luminosity of at least 300

fb^{-1} throughout the HL-LHC's operation. Such a massive data sample is essential for achieving the statistical precision required for the experiment's core physics goals, particularly in the study of extremely rare decays and high-precision CP violation observables. The vision is to maintain the current detector's high efficiency and resolution despite the significantly harsher conditions.

The design of Upgrade II is flexible and is currently being evaluated through three main scoping scenarios, which differ in detector complexity, cost, and peak luminosity capability.

The Baseline scenario corresponds to the full Upgrade II configuration and is designed to operate at a peak luminosity of 1.5×10^{34} $\text{cm}^{-2}\text{s}^{-1}$. It includes the complete set of subdetectors, notably the TORCH detector for low-momentum particle identification and a fully instrumented Mighty Tracker with silicon pixels in the highest occupancy regions. This configuration provides the largest physics reach and the highest robustness against occupancy effects.

The Middle scenario is a cost-optimized alternative targeting a peak luminosity of 1.0×10^{34} $\text{cm}^{-2}\text{s}^{-1}$. While it retains most of the core detector capabilities, it involves moderate compromises in detector granularity and acceptance. The objective is to preserve sensitivity to the key physics observables while reducing the overall cost.

The Low scenario is the most constrained configuration and also assumes a peak luminosity of 1.0×10^{34} $\text{cm}^{-2}\text{s}^{-1}$. In this scenario, more significant compromises are made, such as the removal of the TORCH detector and a simplified tracking system. Consequently, it offers reduced robustness and physics sensitivity, particularly for low-momentum particle identification and complex decay topologies. As of today, the Low scenario serves as the reference configuration under study. Ongoing efforts within the collaboration aim to maximize the detector performance while remaining within the constraints imposed by this option.

10.2 Physics program

The LHCb Upgrade II physics program is designed to deliver world-leading precision across a diverse set of observables, providing unique probes for new physics. The sheer volume of data will allow LHCb to move from a discovery machine to a precision laboratory for flavor physics.

A major goal is to perform extremely precise measurements of CP violation, particularly through the CKM angle γ and the weak phase ϕ_s. Measuring γ with a precision of about 0.35° using theoretically clean tree-level decays will provide one of the strongest tests of the SM, while an $\mathcal{O}(1 \text{ mrad})$ measurement of ϕ_s in B_s-meson oscillations will be highly sensitive to contributions from new heavy particles. Any inconsistency between these measurements and SM predictions could provide clear evidence for new physics.

The upgrade will also greatly enhance studies of rare decays, charm physics, and hadron spectroscopy. Precise measurements of rare processes such as $B_s^0 \to \mu^+\mu^-$ and searches for $B^0 \to \mu^+\mu^-$ will probe new particles and interactions beyond the SM. In the charm sector, LHCb aims to measure CP violation at the 10^{-5} level, testing whether observed charm CP violation can be fully explained by known physics. In addition, the large dataset will enable the discovery and detailed study of exotic hadrons, including doubly heavy hadrons and tetraquarks, improving our understanding of the strong interaction and the structure of matter.

The unique forward geometry of LHCb and the high granularity of this upgrade will also significantly benefit the heavy-ion program, providing access to the full centrality range of AA collisions. The prospects for heavy-ion physics at LHCb after Upgrade II are presented in Ref. [138]. As the primary focus of this thesis is heavy-ion physics, the relevant aspects of the LHCb heavy-ion program are discussed in more detail in the following.

10.2.1 Heavy-ions physics program

One of the most important scientific goals is the study of the thermodynamic properties of strongly interacting matter. Because particle multiplicities vary with rapidity, the forward acceptance of LHCb naturally probes different effective temperatures within the QGP. This creates a unique opportunity to investigate the QCD equation of state over a temperature range not fully accessible elsewhere. Measurements of average transverse momentum as a function of rapidity and the extraction of the speed of sound in the medium also can be done. These observables connect directly to lattice-QCD calculations and hydrodynamic models. The experiment is also expected to provide strong constraints on transport coefficients such as shear viscosity and bulk viscosity through precise anisotropic-flow measurements as functions of rapidity and particle species.

Direct photons and dileptons are among the cleanest probes of heavy-ion collisions because they interact only weakly with the surrounding medium after production. They therefore preserve information about the earliest stages of the collision. To study the electromagnetic probes, four major photon sources are considered: prompt photons, thermal photons, fragmentation photons, and bremsstrahlung radiation. Thermal photons are particularly valuable because their spectrum contains information about the temperature of the medium. LHCb benefits from forward kinematics, which improves photon-energy resolution, and from photon conversion measurements using electron–positron pairs reconstructed in the tracking system. Novel analysis techniques, including machine-learning classifiers and photon–hadron correlation studies, are being developed to separate different photon sources, as seen in Fig. 10.2. Dilepton measurements in the intermediate mass range are also emphasized because they can probe pre-equilibrium dynamics and the approach to thermalization. Upgrade II is expected to provide the tracking, muon identification, and heavy-flavour rejection capabilities required to make these measurements feasible in central PbPb collisions.

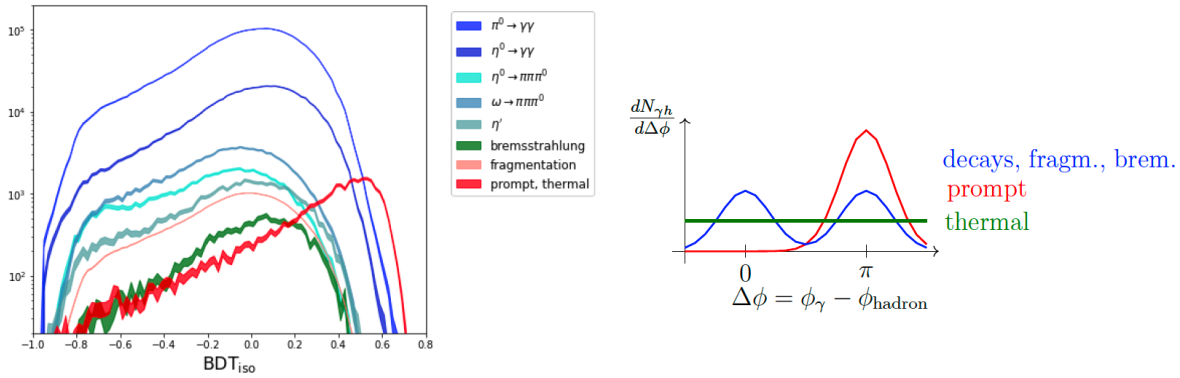


Figure 10.2: Identification of photon sources using a BDT trained using Pythia and EPOS generated events (left) and the azimuthal correlation between the photons and hadrons produced in the event (right). Figure taken from [138] under Creative Commons license.

Heavy-flavour particles containing charm or beauty quarks remain central probes of QCD matter. Their large masses allow perturbative calculations, while their interactions with the medium provide information about diffusion, transport, and hadronization mechanisms. Unresolved questions surrounding coalescence models, which describe how quarks combine into hadrons in dense environments, are subjects of interest. Measurements of charm baryon-to-meson ratios, such as Λ_c^+/D^0 and Ξ_c/D^0, are projected with high precision in central PbPb collisions. Upgrade II will also extend heavy-flavour studies into the beauty sector, enabling measurements of B_s^0/B^0 and Λ_b^0/B^0 production ratios over the full centrality range. As discussed in detail in the first part of this thesis, quarkonium states provide complementary information about deconfinement and color screening. Therefore, the measurement of J/ψ, χ_c and Υ states are a central pillar of the research program. A major objective is to measure feed-down contributions from excited states, which are essential for interpreting quarkonium suppression patterns. Upgrade II should greatly improve statistical precision for nuclear-modification-factor measurements, according to studies based on simulations as shown in Fig. 10.3, and may enable new studies of η_c production in nuclear collisions.

The upgraded detector will also support an extensive jet program. Measurements of jet quenching at forward rapidity are especially important because they probe the temperature dependence of parton energy loss in cooler regions of the medium than those typically studied by ATLAS and CMS. Heavy-flavour jets, jet substructure observables, and photon-tagged heavy-flavour jets will provide detailed insight into quark energy loss mechanisms. Heavy-ion collisions may provide a unique environment for studying the production and internal structure of exotic hadrons as $\chi_{c1}(3872)$. Measurements of these exotic states could discriminate between molecular, tetraquark, transport, and statistical-hadronization scenarios. Ultra-peripheral and central-exclusive collisions form another

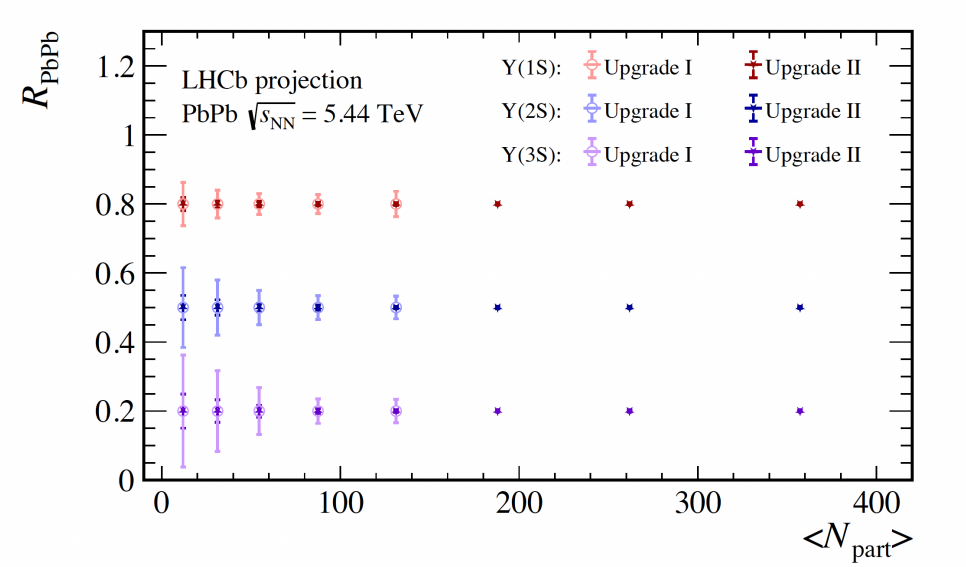


Figure 10.3: Expected statistical uncertainty on R_{PbPb} of $\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$ states in the LHCb Upgrade I with $\mathcal{L}_{\mathrm{int}} = 1$ nb^{-1} and LHCb Upgrade II with $\mathcal{L}_{\mathrm{int}} = 10$ nb^{-1} with PbPb collisions. Figure taken from [138] under Creative Commons license.

major research direction. These processes provide access to generalized parton distributions, gluon-density fluctuations, and possible saturation effects. High-statistics samples of exclusive quarkonium production are expected in both PbPb and pPb collisions.

10.3 Subdetectors requirements

Achieving the goals of Upgrade II requires substantial improvements to the detector subsystems, integrating state-of-the-art sensors, fast electronics, and high-speed data transmission. A layout of the Upgrade II detector is shown in Fig. 10.4.

10.3.1 Tracking system

The tracking system, responsible for identifying the trajectories of charged particles and reconstructing decay vertices, will replace and improve the existing subdetectors, but will also introduce some new ones in some of the scoping scenarios.

Closest to the interaction region is VELO, which is designed to operate in an environment characterized by high particle densities and significant radiation exposure. To meet these challenges, it employs radiation-tolerant silicon sensor technology combined with precise timing capabilities that enhance track and vertex reconstruction. Its compact mechanical design, with a reduced inner radius down to 7.2 mm, minimizes material within the acceptance region while protecting sensitive components from beam-related effects.

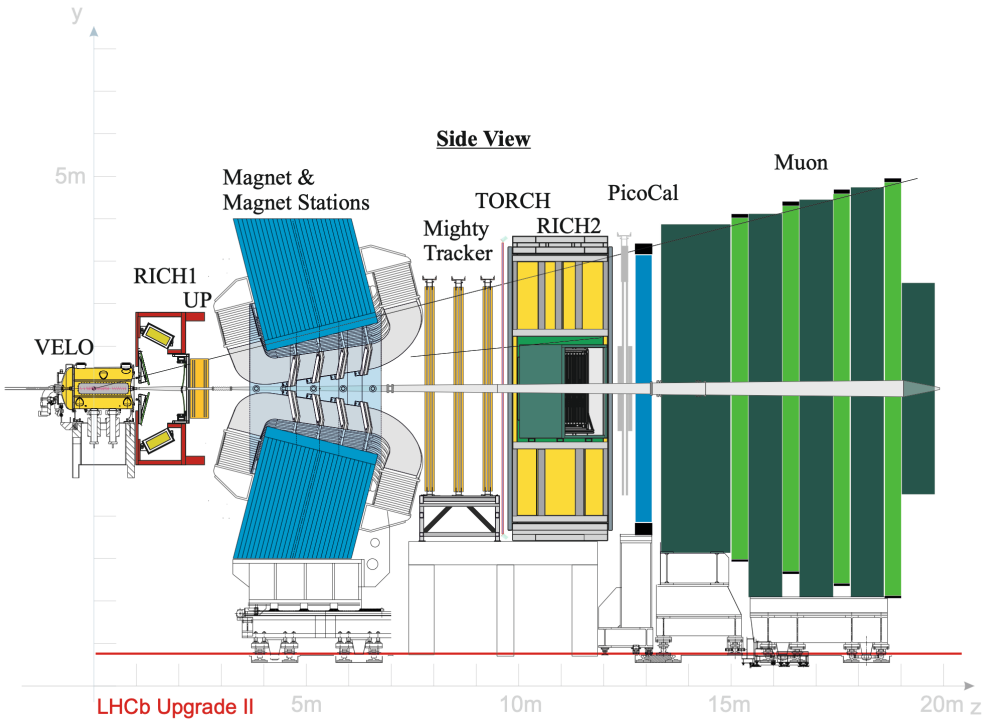


Figure 10.4: Schematic side-view of the Upgrade II baseline detector. Figure taken from [134] under Creative Commons license.

Upstream of the magnet, the Upstream Pixel (UP) detector provides additional tracking information before particles enter the magnetic field. The system is based on monolithic active pixel-sensor (MAPS) technology, in which sensing and readout functions are integrated within a single device. This approach reduces detector material and supports operation under demanding occupancy and radiation conditions while maintaining efficient track matching.

To improve the reconstruction of particles with relatively low momentum, the Upgrade II tracking baseline scenario also includes Magnet Stations positioned within the magnetic field region. These detectors provide additional measurement points that improve momentum determination and increase the reconstruction efficiency for tracks that may otherwise leave the detector acceptance before reaching downstream tracking elements. Their contribution is particularly valuable in complex events containing large numbers of particles.

Downstream tracking is provided by the Mighty-Tracker (MT), a hybrid system designed to cope with varying occupancy and radiation conditions across its large area. The regions exposed to the highest particle fluxes use silicon-based detector technology, while lower-occupancy areas employ scintillating-fibre detectors. This combination balances

performance, radiation tolerance, and cost, while providing the redundancy required for robust and efficient track reconstruction.

10.3.2 Particle Identification Systems

To preserve performance under the more demanding operating conditions of the upgraded experiment, several detector subsystems are being enhanced with improved sensing, timing, and readout capabilities.

The RICH system remains the primary tool for charged-hadron identification. Its upgrade focuses on maintaining particle-identification performance in an environment with significantly higher event occupancy. This involves the adoption of more advanced photodetector technologies, upgraded readout electronics, and modifications to the optical system. Enhanced timing capabilities are expected to improve background rejection and sustain the precision required for accurate Cherenkov-angle measurements.

Complementing the RICH detectors is TORCH, a dedicated time-of-flight detector designed to extend particle-identification performance in momentum regions where traditional Cherenkov techniques become less effective. By combining precise timing measurements with Cherenkov-photon detection, TORCH provides an additional means of distinguishing between particle species and improves overall identification coverage across a broad range of momenta.

The electromagnetic calorimeter, the PicoCal, is also being redesigned to address the increased radiation levels and particle densities anticipated during future operation. The upgraded system employs different calorimeter technologies in regions exposed to varying radiation and occupancy conditions, ensuring optimal performance throughout the detector acceptance. In addition to measuring electromagnetic energy deposits, the calorimeter incorporates precise timing information to improve photon reconstruction and reduce background contributions in complex collision events.

10.3.3 Muon System

To cope with the Upgrade II rates, the inner regions (R1 and R2) of the muon stations will be upgraded from Multi-Wire Proportional Chambers (MWPCs) to μ-RWELL detectors, which offer better rate capability and longevity. Additional iron-concrete shielding will be installed in front of the M2 station to reduce the background from hadrons.

10.3.4 Real-Time Analysis and Online Systems

The massive data rate from the detector (up to 40 MHz) requires a revolutionary approach to data processing. To address this, a triggerless readout architecture has been adopted, in which all data are streamed directly to HLT at the full collision rate rather than relying on a hardware pre-selection stage.

The HLT level itself is structured as a two-stage system, as the actual system. The first stage, HLT1, leverages GPU hardware to carry out rapid track reconstruction, vertexing, and preliminary event selection. The second stage, HLT2, operates on a dedicated CPU farm and performs complete event reconstruction, incorporating particle identification and calorimeter information alongside real-time calibration.

Crucially, the RTA design maintains consistency across operational scenarios, ensuring that the most physically significant events are reliably processed and stored even under peak luminosity conditions. This architecture underpins the broader RTA paradigm, in which raw detector data is reduced to high-level physics objects on the fly, rather than stored in bulk for offline processing.

10.4 Expected performance

Understanding the expected performance of the Upgrade II project is essential for assessing its ability to achieve the targeted physics goals under future operating conditions. Performance projections provide quantitative estimates of key detector capabilities, such as tracking efficiency; momentum and vertex resolution; particle-identification performance; timing precision; and trigger efficiency. These metrics determine the experiment's sensitivity to rare processes, precision measurements, and searches for new phenomena. Furthermore, evaluating the expected performance allows the collaboration to validate design choices, identify potential limitations, optimize detector configurations, and estimate the impact of increased luminosity and occupancy on data quality. In order to address this, several studies based on simulations have been carried out.

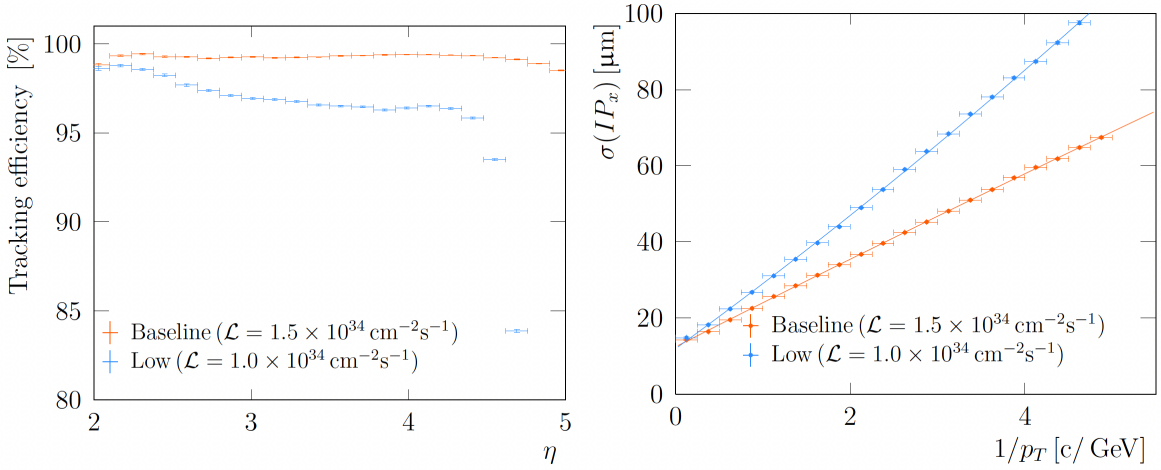


Figure 10.5: VELO tracking efficiency as a function of η comparing the Baseline and Low scenarios(left) and IP resolution in the Baseline and Low scenarios. Figure taken from [134] under Creative Commons license.

In particular, tracking performance is one of the most critical ingredients of the LHCb physics program. Simulation studies have shown that the integration of the 50 ps timing in the VELO is the key to maintaining performance in high pile-up conditions. As shown in Fig. 10.5, the Baseline VELO is expected to achieve a tracking efficiency of $\sim$99% for charged particles. The 4D tracking capability restores the PV reconstruction efficiency to near-perfect levels, allowing for the precise association of secondary heavy-flavour decays to their origin vertices. Furthermore, in the Baseline scenario, the IP resolution will remain comparable to the current detector, approximately 20-30 μm.

Assuming near-perfect VELO performance, this thesis investigates the global tracking performance of the Upgrade II tracking system, including studies of the tracking efficiency for upstream, downstream, and long tracks. The results of these studies are reported in [134, 138]. Building on this, the thesis then focuses on the performance of the UP detector. The following chapter provides a detailed account of the tracking performance studies.

11

Tracking performance studies

In the previous chapter, the importance of understanding the expected performance of the Upgrade II detector was introduced. Tracking performance acts as a bridge between the detector response and the physics analyses, providing an estimate of the detector capabilities and limitations under the operating conditions foreseen at the HL-LHC. In particular, track reconstruction plays a crucial role in momentum determination, primary vertex reconstruction, and the identification of displaced decay vertices, all of which are essential ingredients for the Upgrade II physics programme.

At the luminosities expected during HL-LHC operation, the number of particles traversing the detector will increase considerably in both pp and heavy-ion collisions. The resulting detector occupancy makes the reconstruction task more challenging, as disentangling hits belonging to different particles becomes increasingly difficult. Consequently, the probability of reconstructing fake tracks, commonly referred to as ghost tracks, increases while the tracking efficiency tends to decrease. In this context, the Upstream Pixel (UP) detector is expected to play a central role by providing additional high-granularity measurements before the magnet, thereby improving track reconstruction, momentum estimation, and ghost-track rejection.

Within the context of the preparation of the LHCb Upgrade II Scoping Document [134], several tracking-performance studies were carried out. In this work, the tracking efficiency and ghost rate are evaluated for the different track categories reconstructed at LHCb under a variety of Upgrade II detector configurations, with particular emphasis on the geometry and design of the Upstream Pixel (UP) detector. In addition, the tracking performance is studied in both pp and PbPb collision environments in order to assess the robustness of the detector under the wide range of occupancies expected during Upgrade II operation.

The results presented in this chapter were included in the LHCb Upgrade II Scoping Document and provided important input to the collaboration, contributing to the validation of detector layouts, the optimization of detector configurations within cost constraints, and the evaluation of the impact of increased luminosity and occupancy on the reconstruction performance.

11.1 Pattern Recognition workflow

At the time these studies were performed, the Upgrade II software framework was still under development. Consequently, they could not be carried out within the official LHCb reconstruction software. Instead, a lightweight Python-based framework, referred to as the *Pattern Recognition* framework, was used to emulate the reconstruction chain locally. The Pattern Recognition framework provides a simplified implementation of the LHCb reconstruction workflow while offering the flexibility to study detector configurations that are difficult to implement in the official software, such as alternative definitions of reconstructible tracks or modifications of the intrinsic characteristics of the tracking subdetectors. An important advantage of this framework is that modifications of the detector geometry can be emulated directly at the tuple level, avoiding the need to regenerate dedicated Geant4 simulation samples.

The Pattern Recognition workflow takes as input simulated events produced with the Gauss framework [98]. These samples contain the generator-level information, the Geant4 simulation of the baseline detector geometry, and the hit collections associated with each subdetector. The simulated events are processed and converted into tuples with a more convenient format for further studies. These tuples store the position, time, and deposited energy of each simulated particle in the different tracking detectors.

At this stage, the hit information corresponds to the true Geant4 response of the detector. In order to emulate the finite detector resolution, each hit is smeared according to the pixel size of the corresponding subdetector. The smearing of the x and y coordinates is performed by assigning a random value generated from a Gaussian distribution centered on the original position and with a standard deviation

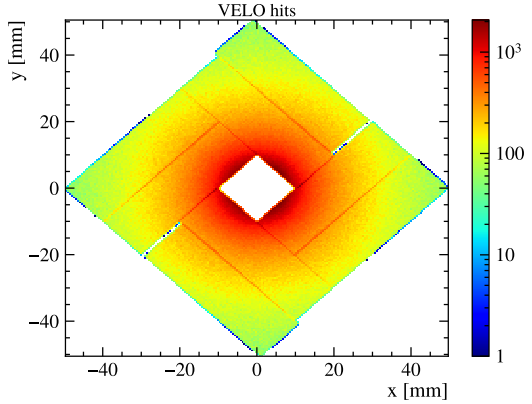
$$\sigma_{x,y} = \frac{\text{pixel size (x,y)}}{\sqrt{12}}, \tag{11.1}$$

where the factor $1/\sqrt{12}$ corresponds to the standard deviation of a uniform distribution, assuming that the true impact position is uniformly distributed within a pixel cell.

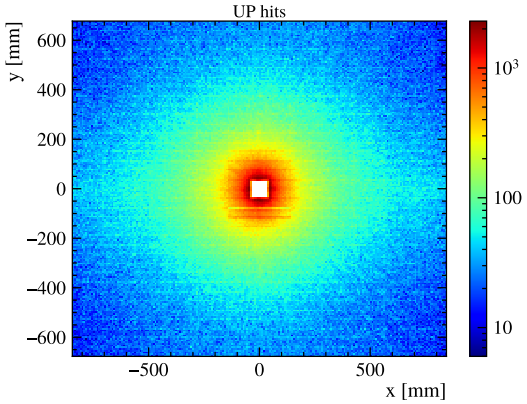
As mentioned previously, the detector geometry is fixed to the Baseline scenario during the Gauss simulation step. A heatmap of the simulated hits in PbPb collisions in the (x, y) plane for each tracking subdetector is shown in Fig. 11.1. Nevertheless, the Pattern Recognition framework allows modifications of the detector acceptance, provided that

they correspond to a reduction of the active area. This feature makes it possible to emulate alternative detector geometries without the need to generate dedicated simulation samples.

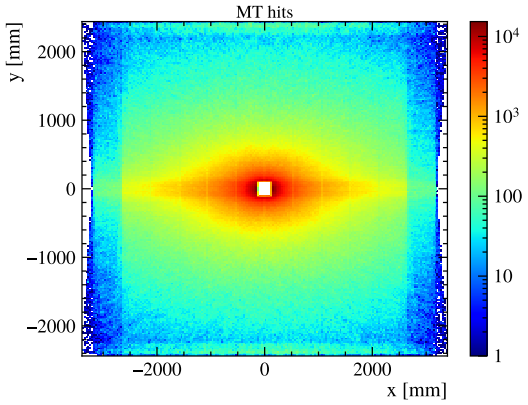
The smeared hits are then used as input to the reconstruction algorithms corresponding to the different track categories. Although the hit positions have been smeared, the association between hits and MC particles is preserved, allowing the computation of the two quantities used throughout this chapter to assess the tracking performance: the *tracking efficiency* and the *ghost rate*.



(a)



(b)



(c)

Figure 11.1: Spatial distribution of simulated hits in the (x, y) plane for a central (0-5%) PbPb collision generated with Angantyr [130]. The distributions are shown for the VELO (a), UP (b), and Mighty Tracker (c) detectors using the Baseline Upgrade II geometry.

The tracking efficiency is defined as the probability for a charged particle traversing the detector to be successfully reconstructed as a track. In this work, the tracking efficiency is defined as

$$\varepsilon = \frac{N_{\text{reco}}}{N_{\text{reconstructible}}}, \tag{11.2}$$

where N_{reco} is the number of reconstructed tracks and $N_{\text{reconstructible}}$ is the number of reconstructible particles. A reconstructible particle is an MC-simulated charged particle satisfying a set of geometrical requirements. In this work, these requirements correspond to the presence of at least three hits in the VELO and UP detectors, respectively, and at least one hit in each station of the Mighty Tracker. A particle is considered reconstructed if the corresponding fitted track shares at least 70% of its hits with the original MC particle.

The ghost-track rate is defined as the fraction of reconstructed tracks that cannot be associated with any MC particle. It is computed as

$$GR = \frac{N_{\text{fake}}}{N_{\text{reco}}}, \tag{11.3}$$

where N_{fake} is the number of fake tracks and N_{reco} is the total number of reconstructed tracks.

11.2 Tracking algorithms

The general strategy for track reconstruction in Upgrade II consists of reconstructing standalone tracks in each tracking subdetector and subsequently matching them to build upstream, downstream, and long tracks. In this section, the standalone reconstruction algorithms and the corresponding matching procedures are described.

11.2.1 VELO tracks algorithm

The performance of the VELO track reconstruction was studied independently of this work. As discussed in Sec. 10.4, the expected tracking efficiency of the VELO detector is close to 99%. To emulate this nearly ideal performance, a *Cheated Algorithm* is employed. In this algorithm, hits are grouped according to the MC particle from which they originate, and only collections of hits associated with the same MC particle are considered during the reconstruction. The Cheated Algorithm is therefore not intended to represent a realistic reconstruction procedure, but rather provides an estimate of the maximum performance achievable by the VELO detector. Besides VELO tracks, this algorithm is also used later to evaluate the performance of the matching procedures.

Once the hits have been grouped according to their originating MC particle, they are fitted using a model that depends on the track category under consideration. In the case of VELO tracks, the effect of the magnetic field is negligible and the trajectory can be described by two independent linear functions in the xz and yz planes.

By construction, the Cheated Algorithm ghost-track rate is zero. However, its efficiency is not strictly equal to 100%, since a small fraction of tracks may fail the fitting procedure or not satisfy the quality requirements imposed on the fitted parameters.

11.2.2 UP standalone algorithm

The current design of the UT detector is based on silicon-strip technology, making the reconstruction of standalone UT tracks particularly challenging. Nevertheless, the UT provides hit information that is exploited during long-track reconstruction and plays an important role in reducing the ghost-track rate. In contrast, the Upgrade II UP detector is based on high-granularity pixel technology, which enables a more precise reconstruction of standalone tracks.

The standalone UP reconstruction algorithm starts by selecting a reference hit in the first station. Search windows with an annular-sector shape, parameterized by $\Delta\rho$ and $\Delta\phi$, are then defined in stations 2 to 4. As illustrated in Fig. 11.2, the value of $\Delta\rho$ depends on the pair of stations considered, whereas $\Delta\phi$ is kept fixed for all combinations. In addition, the search windows are required to satisfy a maximum slope condition with respect to the reference hit. All possible combinations of hits satisfying these requirements are built, yielding a collection of track candidates.

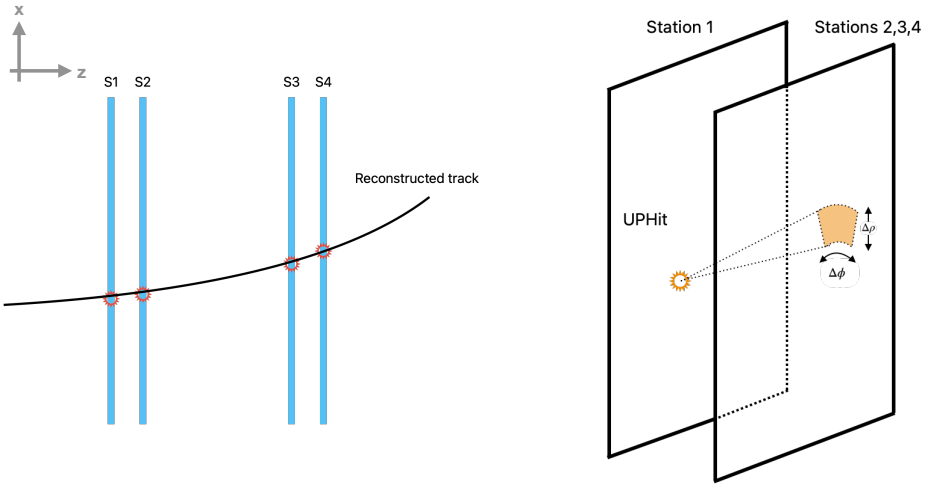


Figure 11.2: Scheme of the UP standalone track reconstruction algorithm.

The track candidates associated with a given reference hit are subsequently fitted. Since the UP detector is located immediately upstream of the magnet, the effect of the magnetic field, although still modest, cannot be neglected. Therefore, the trajectories are fitted with a parabolic function in the xz plane, usually referred to as the bending plane, and with a linear function in the yz plane. For all successfully fitted candidates, the best track

is selected according to the sum of the χ^2 values obtained in the two projections. The hits belonging to the selected track are removed from the global hit collection, and the procedure is repeated with another reference hit in station 1.

The algorithm gives priority to tracks composed of four hits, one per station. However, according to the reconstructibility criteria adopted in this work, tracks containing only three hits are also considered reconstructible. Consequently, if no four-hit combination is found within the search windows, the reconstruction procedure is repeated while omitting one of the stations. Although this algorithm achieves a low ghost-track rate, its overall efficiency remains limited, especially in high-occupancy environments. This limitation motivated the development of an alternative standalone reconstruction strategy based on a Cellular Automaton, which is presented later in this chapter.

11.2.3 TTracks algorithm

The Mighty Tracker employs a hybrid technology combining silicon pixels in the high-occupancy region close to the beam pipe and scintillating fibres in the outer acceptance. This detector layout allows the reconstruction of three different categories of TTracks: pixel-pixel, fibre-fibre and pixel-fibre tracks. The detector is composed of three stations (T1-T3), each containing four measurement planes, providing up to twelve hits per reconstructed track.

The standalone reconstruction performance of the Mighty Tracker was studied independently of this work and was found to exceed 90% in all Upgrade II scenarios considered, as shown in Fig. 11.3. Consequently, TTracks are reconstructed using the Cheated Algorithm throughout this chapter.

The fit model associated with the Cheated Algorithm consists of a parabolic trajectory in the xz plane, with a larger bending between T1 and T2 than between T2 and T3 to account for the decrease of the magnetic field with increasing z, and a linear trajectory in the yz plane.

11.2.4 Upstream matching algorithm

One of the objectives of this work is to evaluate the performance of the matching strategy used to reconstruct upstream tracks from already reconstructed VELO and UP tracks. To assess the intrinsic performance of the matching procedure, both VELO and UP tracks are reconstructed using the Cheated Algorithm.

The upstream matching algorithm considers all possible combinations of VELO and UP tracks. For each pair, the hits belonging to the two standalone tracks are refitted simultaneously using a third-order polynomial in the xz plane and a second-order polynomial in the yz plane. These parameterizations account for small deviations from the ideal trajectory induced by interactions with the detector material. The best upstream track

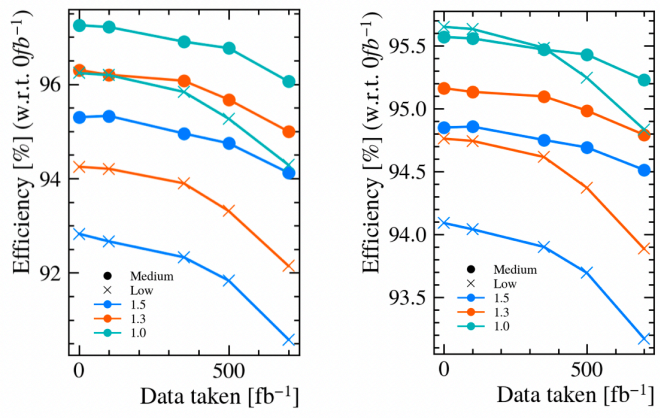


Figure 11.3: Expected track reconstruction efficiency in the MT as a function of the collected integrated luminosity, in (left) the fibre region only and (right) the full MT acceptance. For each plot, different values of peak luminosity are indicated by different colors. Figure taken from [134] under Creative Commons license.

candidate is then selected according to the sum of the χ^2 values obtained in the two projections. A schematic representation of the upstream matching procedure is shown in Fig. 11.4.

This algorithm also includes a BDT-based filter to suppress ghost tracks. The BDT model is trained following the procedure described in Sec. 7.2.1. Only quantities available from the standalone tracks are used as input variables, namely the hit information in each subdetector and the parameters obtained from the standalone-track fits. The trained model then classifies each VELO-UP pair as either matched or unmatched.

The training is performed using an independent simulation sample in order to avoid potential biases and overly optimistic performance estimates. The selection applied to the BDT response is intentionally kept loose, yet it significantly reduces the ghost-track rate while preserving the tracking efficiency.

11.2.5 Downstream and long tracks matching algorithm

Similarly to the upstream tracks, this work also aims to evaluate the performance of the matching strategy used to reconstruct downstream and long tracks, which rely on the same reconstruction algorithm. Both track categories are built from standalone tracks reconstructed with the Cheated Algorithm. In the case of downstream tracks, all possible combinations of UP tracks and TTracks are considered, whereas long tracks are formed from all possible combinations of upstream tracks and TTracks.

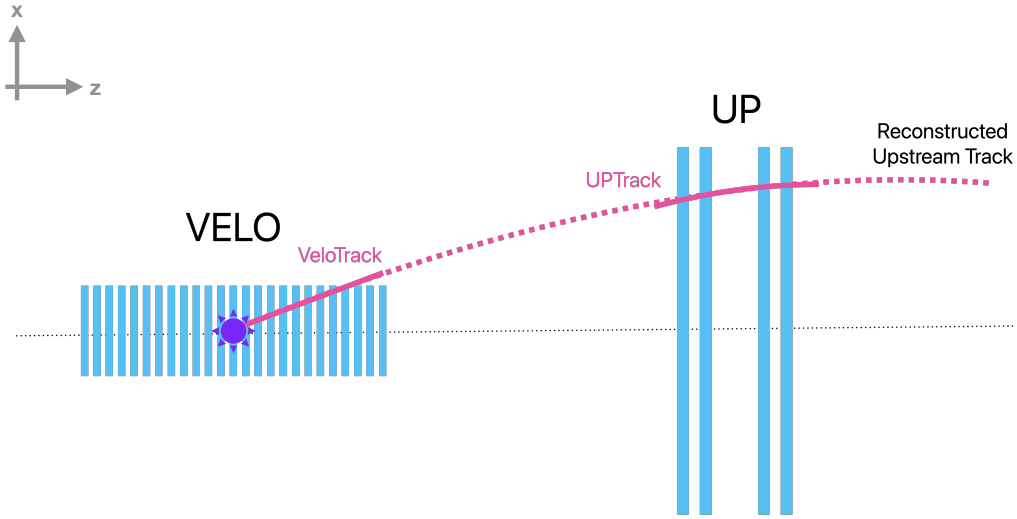


Figure 11.4: Scheme of the upstream matching using already reconstructed VELO and UP tracks.

For each pair of tracks, the incoming track (UP or upstream) and the TTrack are extrapolated to a reference plane located inside the magnet, as illustrated in Fig. 11.5. The position of this plane, denoted as z_{mag}, has been optimized independently for downstream and long tracks and is chosen such that the distance between the extrapolated incoming track and the TTrack is minimized. The tuning of z_{mag} is essential for the momentum estimation during the reconstruction.

The momentum estimation is performed using the so-called *p-kick method*, which approximates the effect of the magnetic field by an instantaneous kick of the momentum vector at a given position inside the magnet. The p-kick approximation is motivated by the fact that most of the magnetic field integral is concentrated around the center of the dipole magnet.

The distance between the two extrapolated tracks is defined as

$$\chi^2 = \frac{dx^2}{d\sigma_x^2}\bigg|_{z_{\text{mag}}} + \frac{dy^2}{d\sigma_y^2}\bigg|_{z_{\text{mag}}} + \frac{dt_y^2}{d\sigma_{t_y}^2}, \tag{11.4}$$

where dx and dy are the differences between the x and y positions of the two extrapolated tracks evaluated at z_{mag}, while $d\sigma_x$ and $d\sigma_y$ correspond to the typical standard deviations of these quantities obtained during the tuning of z_{mag}. The quantity dt_y represents the

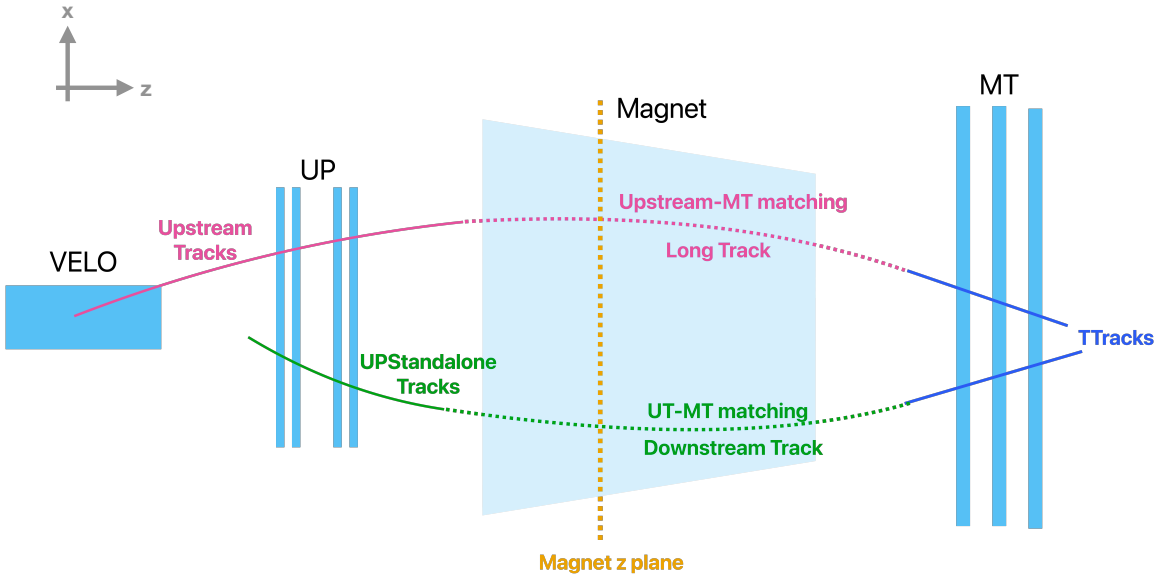


Figure 11.5: Scheme of the matching algorithm for downstream and long tracks.

difference between the slopes of the incoming track and the TTrack in the yz plane, and $d\sigma_{t_y}$ is its corresponding standard deviation. The best downstream or long-track candidate is therefore the pair of tracks that minimizes the quantity defined in Eq. 11.4.

Similarly to the upstream matching strategy, the downstream and long-track matching algorithm also includes a BDT-based filter to suppress ghost tracks. The model is trained using the same dedicated simulation sample employed for the upstream matching and only exploits information available from the standalone tracks, namely the hit information in each subdetector and the fit parameters of the incoming track and the TTrack.

The BDT filter is applied before computing the quantity defined in Eq. 11.4 in order to reduce the computational cost. As for the upstream matching, the selection on the BDT response is intentionally kept loose, while still providing a significant reduction of the ghost-track rate with a negligible impact on the tracking efficiency.

11.3 Performance results

Several studies were carried out in this work to evaluate the tracking performance under different Upgrade II scenarios. Particular emphasis is placed on the reconstruction of upstream, downstream, and long tracks for different geometries and layouts of the Upstream Pixel detector. In addition, the tracking performance is compared in pp and high-centrality PbPb collisions in order to assess the robustness of the detector in the

highest occupancy environments expected during Upgrade II operation.

These studies make use of three different MC simulation samples generated under the HL-LHC operating conditions and using the Baseline geometry of the LHCb Upgrade II detector. As described in Sec. 11.1, these simulations contain the truth information associated with the hits in each tracking subdetector. A sample of 950 pp collisions generated with Pythia is used to study the tracking performance in proton-proton interactions. In addition, two high-centrality PbPb collision samples generated with the Angantyr heavy-ion model implemented in Pythia [130] are employed. The first sample, containing 100 events, is used to evaluate the tracking performance in PbPb collisions, while the second sample, containing 10 events, is dedicated to the training of the BDT models introduced in the previous sections.

The PbPb simulations correspond to the highest centrality interval considered in this work, namely the 0-5% centrality class, which is equivalent to an impact parameter of $b < 3.5$ fm. This choice was made to assess the tracking performance under the most challenging occupancy conditions. The results obtained for central PbPb collisions can therefore be regarded as conservative estimates of the expected Upgrade II performance. A comparison of the hit multiplicity in pp and PbPb collisions is shown in Fig. 11.6. As expected, the detector occupancy is substantially larger in central PbPb collisions, making the reconstruction task considerably more difficult.

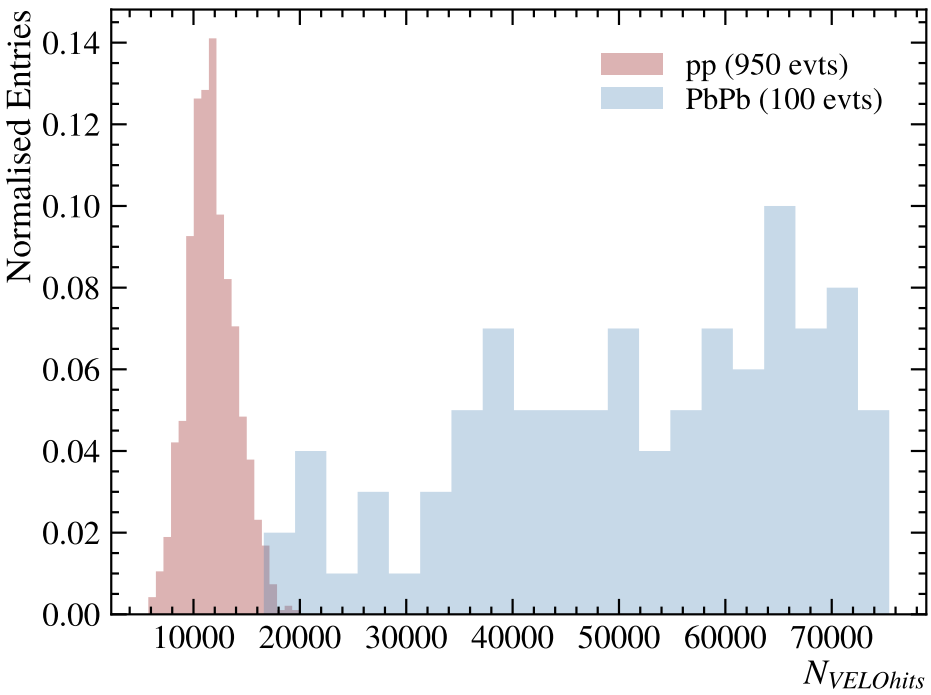


Figure 11.6: Distribution of the number of VELO hits in pp and central PbPb collisions.

At the time these studies were performed, the official Upgrade II reconstruction software was still under development and the reconstruction was therefore carried out locally using the Pattern Recognition framework. Due to the limited computational resources available, only subsets of the simulated samples, containing 50 pp events and 10 PbPb events, were processed. Although these reduced samples lead to larger statistical uncertainties, they are sufficient to compare detector configurations and evaluate the relative impact of different reconstruction strategies.

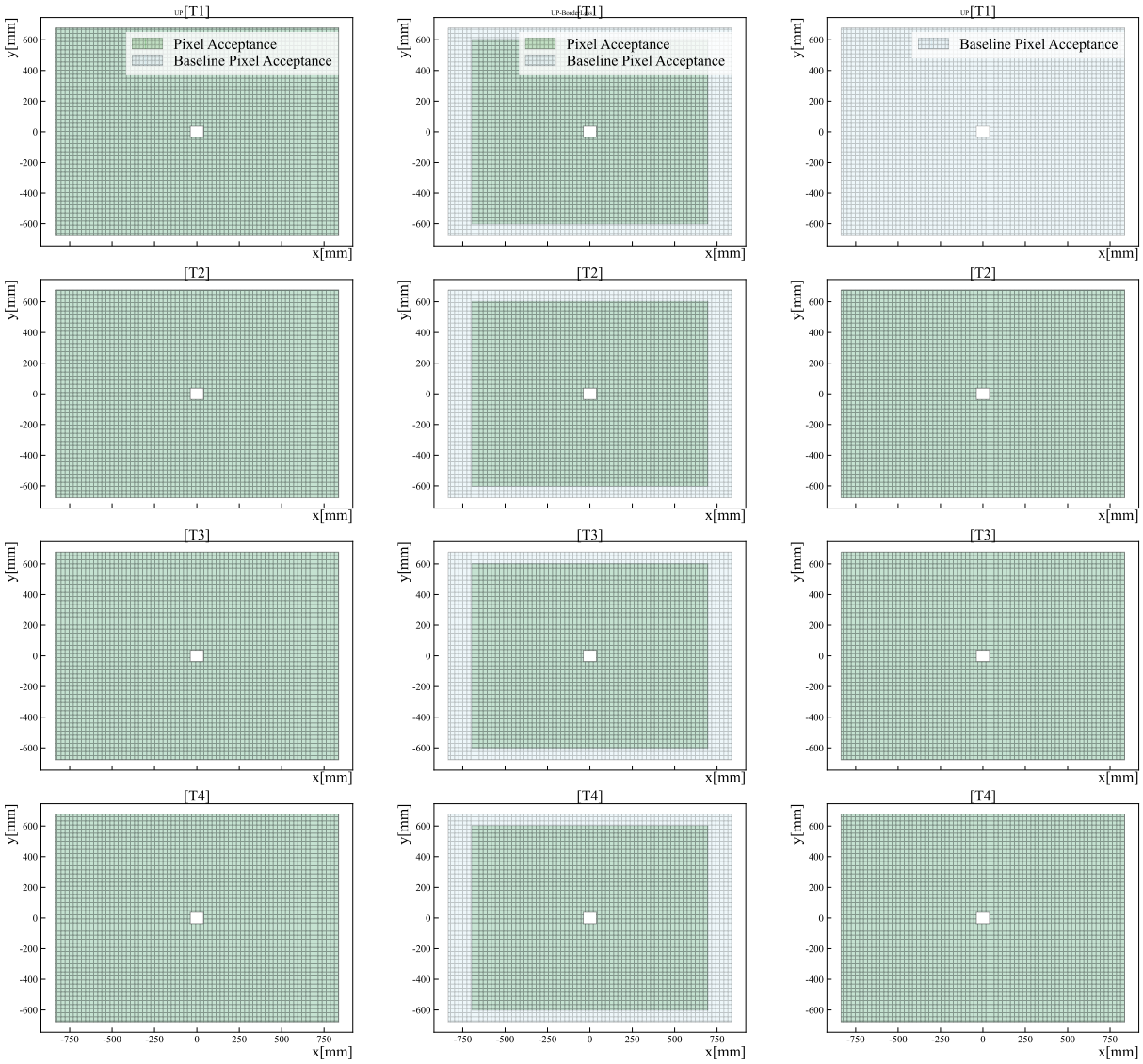
11.3.1 UP stations geometry comparison

The performance of the UP standalone reconstruction is evaluated for three different UP geometries. The first geometry, referred to as the Baseline scenario, consists of four square stations covering the full LHCb acceptance and corresponds to the layout implemented in the simulated samples. The second geometry, called the Borderless scenario, is obtained by removing one row of pixels along the border of each UP station, resulting in a slightly reduced acceptance. Finally, in the NoT1 scenario, the first UP station is removed, while the remaining stations retain the Baseline configuration. A schematic representation of the three geometries considered is shown in Fig. 11.7.

The impact of the detector geometry on the tracking performance is evaluated by computing the tracking efficiency and ghost-track rate of standalone UP tracks reconstructed with the algorithm described in Section 11.2.2. The efficiency is evaluated for tracks with a total momentum larger than 5 GeV/c and within the pseudorapidity range $2 < \eta < 5$. These requirements ensure that the selected tracks have a sufficiently large momentum to traverse the magnetic field and potentially be reconstructed as downstream or long tracks.

The tracking efficiencies obtained in pp and PbPb collisions are shown in Fig. 11.8. The Baseline and Borderless geometries exhibit very similar performances, yielding an overall efficiency of approximately 78% in pp collisions. In contrast, the tracking efficiency is significantly degraded when the first station is removed in the NoT1 configuration. This effect becomes even more pronounced in PbPb collisions, where the efficiency decreases from about 75% in the Baseline and Borderless layouts to approximately 24% in the NoT1 scenario. As expected, the efficiency is slightly higher in pp collisions than in PbPb collisions, although the difference remains modest except in the NoT1 case.

The overall ghost-track rate is also studied for the three geometries considered. The results obtained in pp and PbPb collisions are presented in Fig. 11.9, where the ghost-track rate is shown as a function of the number of hits associated with the reconstructed tracks. As discussed in Section 11.2.2, the standalone UP reconstruction algorithm gives priority to tracks composed of four hits. Consequently, most tracks reconstructed with the Baseline and Borderless geometries contain four hits, as can be observed from the overlaid distributions. In contrast, the NoT1 geometry is forced to reconstruct tracks using only three hits.



(a) Baseline. (b) Borderless. (c) NoT1.

Figure 11.7: Layout of the three different UP geometries considered. The Baseline scenario (a), the Borderless scenario (b) with reduced acceptance and the NoT1 scenario (c) where the first station was removed. Each row corresponds to one of the four UP stations.

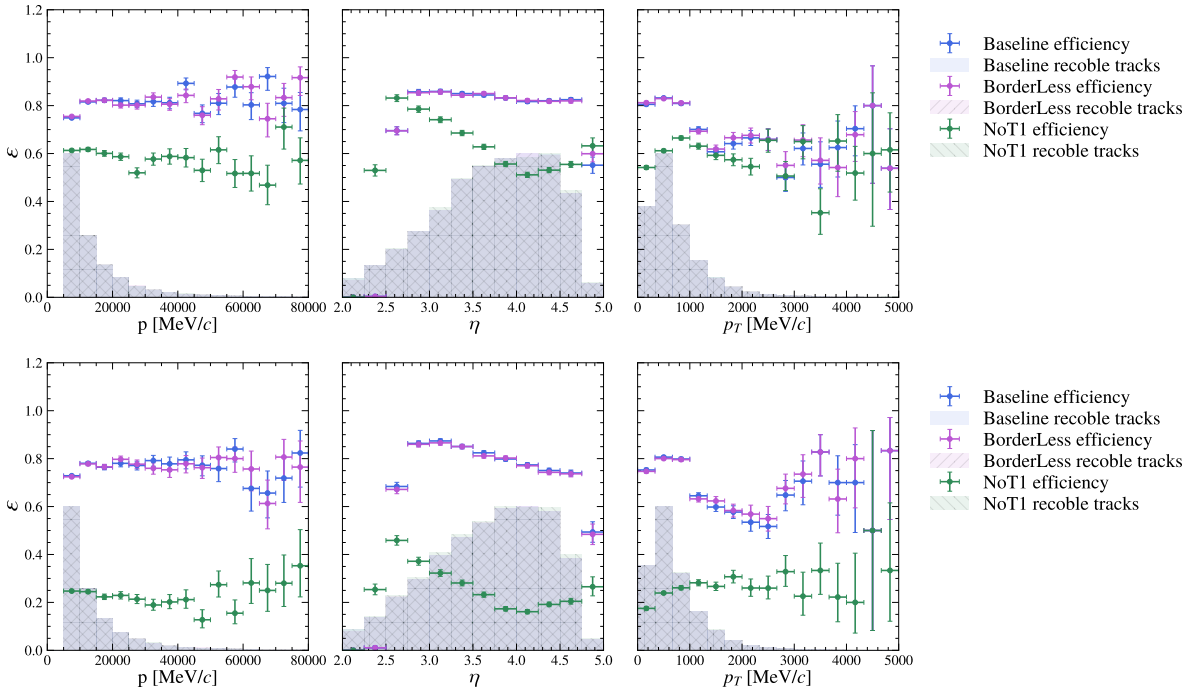


Figure 11.8: UP standalone tracking efficiency as a function of the momentum, pseudorapidity and transverse momentum together with the corresponding distributions of reconstructible tracks. for the Baseline, Borderless and NoT1 UP geometries. The top row show the results for *pp* collisions and the bottom row the results for PbPb collisions.

Consistent with the efficiency results, the Baseline and Borderless geometries exhibit very similar ghost-track rates. In both configurations, the algorithm achieves an excellent performance, yielding ghost-track rates of approximately 0.65% in *pp* collisions and 4% in PbPb collisions. In contrast, the ghost-track rate increases dramatically in the NoT1 scenario, reaching values as high as 86% in central PbPb collisions. This demonstrates that the presence of a fourth measurement plane is essential to suppress combinatorial ambiguities and maintain an acceptable ghost-track rate, particularly in high-occupancy environments.

Overall, the results indicate that the Baseline and Borderless geometries exhibit nearly identical tracking performances. In contrast, removing the first UP station leads to a severe degradation of both the tracking efficiency and the ghost-track rate. The NoT1 configuration is therefore excluded from further consideration, as the potential reduction in material and cost is largely outweighed by the substantial loss in tracking performance.

Since the Borderless geometry preserves the tracking performance of the Baseline layout while reducing the active detector area, it is preferred as the nominal UP geometry owing to its smaller material budget and lower construction cost. Consequently, all studies presented in the following sections are performed using the Borderless configuration.

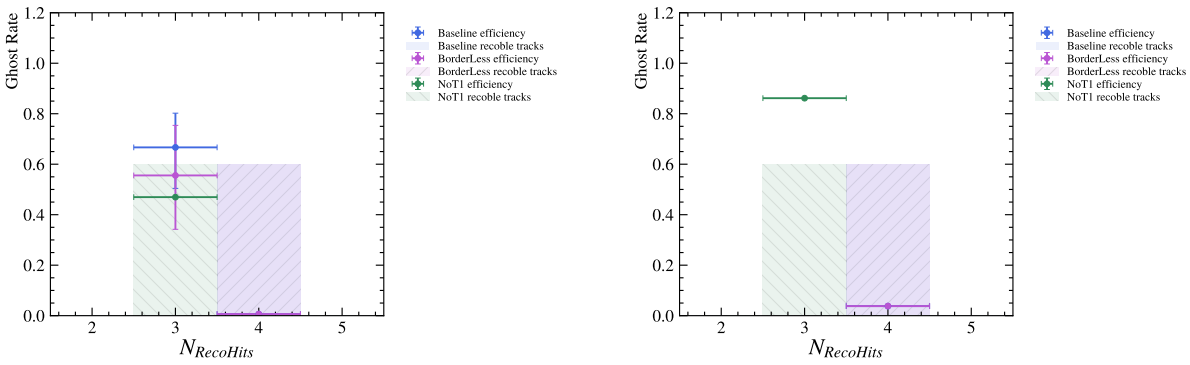


Figure 11.9: UP standalone ghost rate as a function of the number of reconstructed hits together with the distributions of the reconstructible (recoble) tracks for the Baseline, Borderless and NoT1 UP geometries. The left column shows the results for *pp* collisions and the right column the results for PbPb collisions.

11.3.2 UP beam pipe hole comparison

In order to further optimize the material budget of the Upgrade II detector, a dedicated study of the beam-pipe hole size in the UP detector is performed. In the previous section, the nominal inner aperture was fixed to a rectangular opening of 39×37 mm^2. Here, the possibility of enlarging this aperture to 70×70 mm^2 is investigated. The comparison is performed between the nominal Borderless geometry with the original hole size or a modified geometry featuring the enlarged aperture. The impact of the hole size on the reconstruction performance of upstream, downstream, and long tracks is evaluated in both *pp* and central PbPb collisions.

The tracking efficiencies obtained in *pp* and PbPb collisions are presented in Figs. 11.10 and 11.11, respectively. The enlarged-hole geometry exhibits a slightly better performance for upstream and downstream tracks, leading to higher reconstruction efficiencies. This behavior can be explained by the fact that the innermost region of the UP detector experiences the highest occupancy, particularly in central PbPb collisions. Removing this region alleviates the combinatorial complexity of the reconstruction, thereby improving the tracking performance.

As can be seen from the overlaid distributions, enlarging the beam-pipe hole also reduces the detector acceptance for upstream and downstream tracks. The most significant degradation is observed for long tracks. In this case, the efficiency decreases substantially at large momentum and pseudorapidity values when using the enlarged aperture, since the most energetic particles predominantly traverse the innermost region of the detector.

The ghost-track rates obtained in *pp* and PbPb collisions are shown in Fig. 11.12. The results indicate that the reconstruction performance remains very similar for both beam-pipe hole sizes. As expected, the ghost-track rates are slightly lower in *pp* collisions than

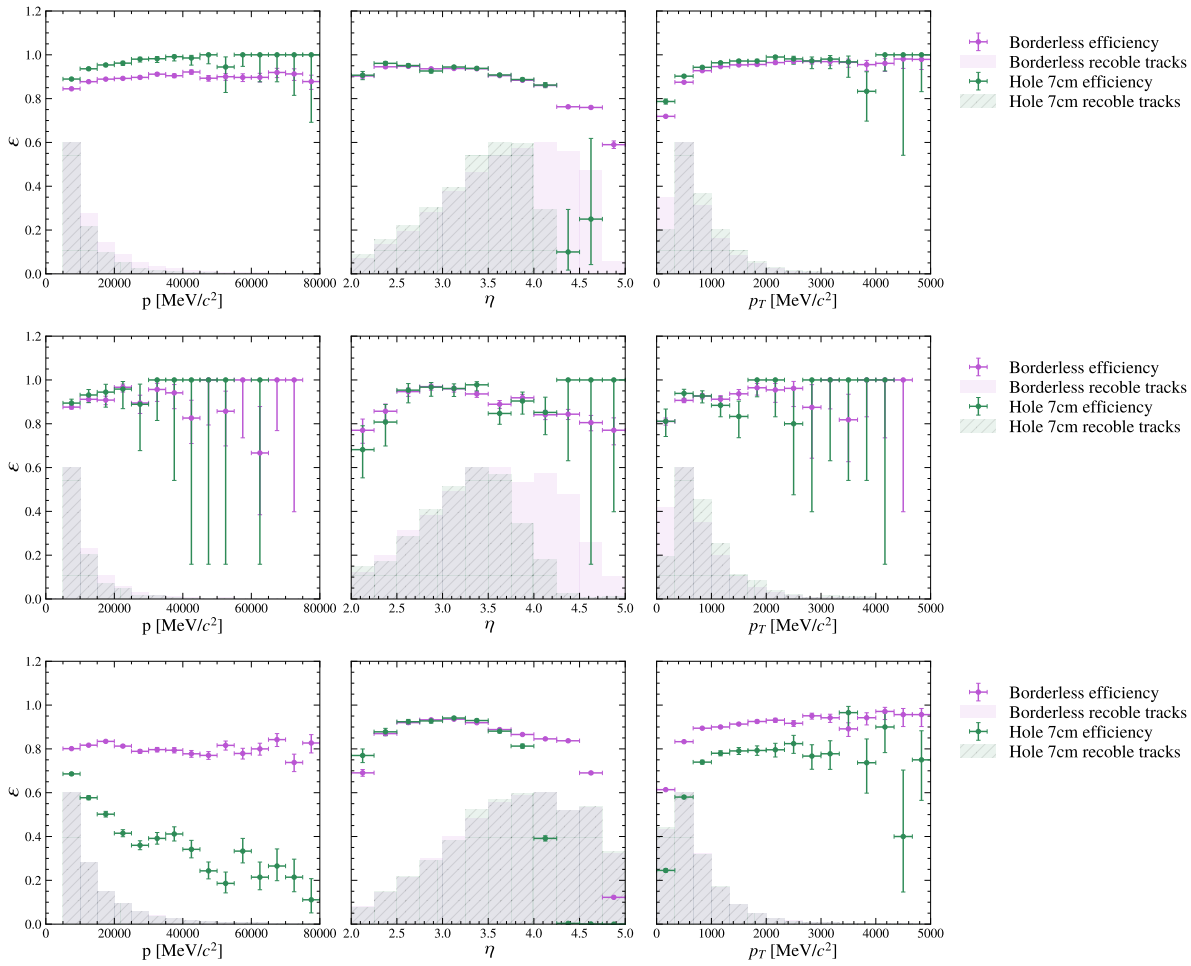


Figure 11.10: Comparison of the tracking efficiency obtained with the nominal and enlarged UP beam-pipe apertures. for pp collisions. From top to bottom, the results of the tracking efficiencies of upstream, downstream and long tracks.

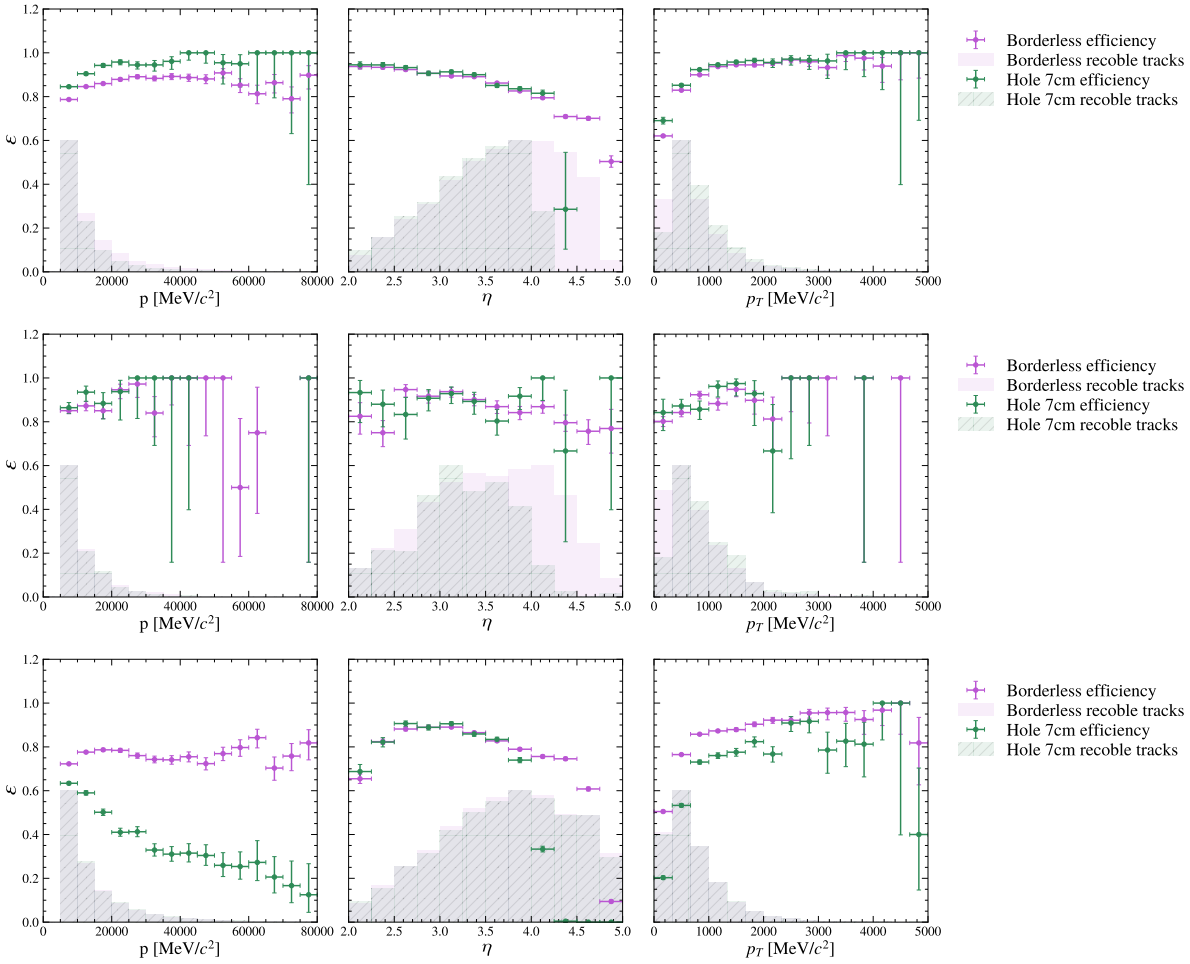


Figure 11.11: Comparison of the tracking efficiency obtained with the nominal and enlarged UP beam-pipe apertures for PbPb collisions. From top to bottom, the results of the tracking efficiencies of upstream, downstream and long tracks.

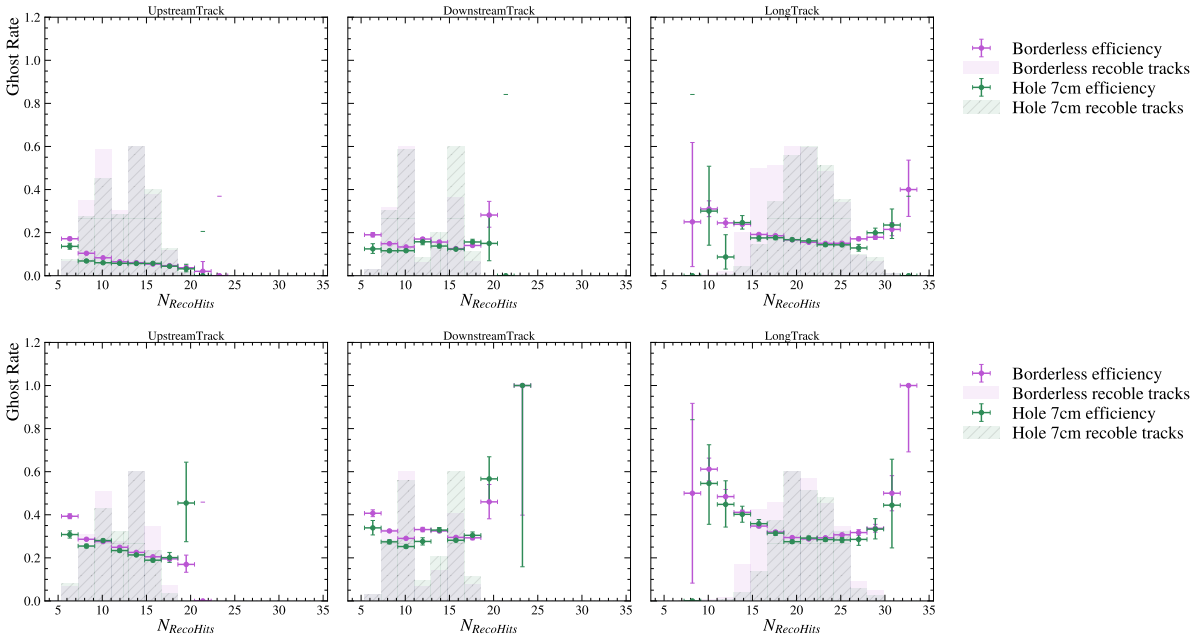


Figure 11.12: Comparison of the ghost rate as a function of the number of reconstructed hits between the original size and the enlarged UP beam pipe hole for upstream, downstream and long tracks. In the top row, the results for *pp* collisions and in the bottom row the results for PbPb collisions.

in central PbPb collisions for all track categories. Overall, enlarging the beam-pipe hole does not significantly affect the ghost-track rate.

Although the enlarged-hole geometry slightly improves the reconstruction efficiency of upstream and downstream tracks, it also reduces the detector acceptance and significantly degrades the performance of long-track reconstruction. Since the ghost-track rate remains largely unaffected, the loss in long-track efficiency becomes the dominant effect. Consequently, the enlarged-hole geometry was not retained for further studies.

11.3.3 UP pixel size comparison

Apart from the detector geometry, the granularity of the UP detector also plays an important role in the tracking performance. A finer granularity may improve the separation of nearby tracks, particularly in high-occupancy environments. Previous studies were performed assuming square pixels with dimensions of 30×30 μm^2. However, such a configuration represents a rather optimistic scenario, as the production of pixels with these dimensions is technologically challenging.

This study investigates the impact of the pixel size on the tracking performance. In the Pattern Recognition framework, the pixel dimensions directly affect the hit resolution

through the smearing procedure described in Eq. 11.1. The reconstruction performance of upstream, downstream and long tracks is compared in *pp* and PbPb collisions for four different pixel sizes: 30×30, 50×50, 50×150 and 100×100 μm^2.

The tracking efficiencies obtained in *pp* and PbPb collisions are presented in Figs. 11.13 and 11.14, respectively. The different pixel configurations exhibit very similar performances for upstream and long tracks, while more significant differences are observed for downstream tracks. In general, smaller pixel sizes lead to higher reconstruction efficiencies. The largest difference is found between the 30×30 μm^2 and 100×100 μm^2 configurations, particularly in the reconstruction of downstream tracks. As in the previous studies, the tracking efficiency is found to be slightly lower in PbPb collisions than in *pp* collisions.

The ghost-track rates obtained for the different pixel configurations are shown in Fig. 11.15. A behavior similar to that observed for the tracking efficiency is found. Most pixel sizes provide comparable performances, with the exception of the 100×100 μm^2 configuration, which exhibits a noticeably larger ghost-track rate in the reconstruction of downstream tracks. This makes the 100×100 μm^2 pixel size the least favorable option.

In contrast, the 30×30, 50×50 and 50×150 μm^2 configurations yield very similar performances, indicating that increasing the pixel dimensions within this range does not significantly degrade the tracking capabilities. Based on these results, a pixel size of 50×50 μm^2 is adopted as the nominal UP configuration, as it provides a realistic compromise between tracking performance and technological feasibility.

11.3.4 Global performance in *pp* and PbPb collisions

Based on the conclusions of the previous optimization studies, the performance of upstream, downstream and long tracks is finally evaluated in *pp* and PbPb collisions using the nominal UP configuration. This configuration corresponds to the Borderless geometry, the original beam-pipe hole dimensions and a pixel size of 50×50 μm^2.

As discussed in the previous sections, the matching algorithms rely on standalone tracks reconstructed with the Cheated Algorithm. Consequently, realistic inefficiencies associated with standalone track reconstruction are not taken into account. The results presented in this section should therefore be regarded as an upper limit on the tracking performance achievable with Upgrade II.

The comparison of the tracking efficiencies obtained in *pp* and PbPb collisions is shown in Fig. 11.16. As observed in the previous studies, the performance in *pp* collisions is slightly better than in PbPb collisions for upstream, downstream and long tracks. This behavior is expected, as the large number of particles produced in central PbPb collisions increases the detector occupancy and makes the reconstruction task more challenging. The difference is more pronounced at large pseudorapidities, where the occupancy is highest. Nevertheless, the overall efficiency difference between *pp* and PbPb collisions

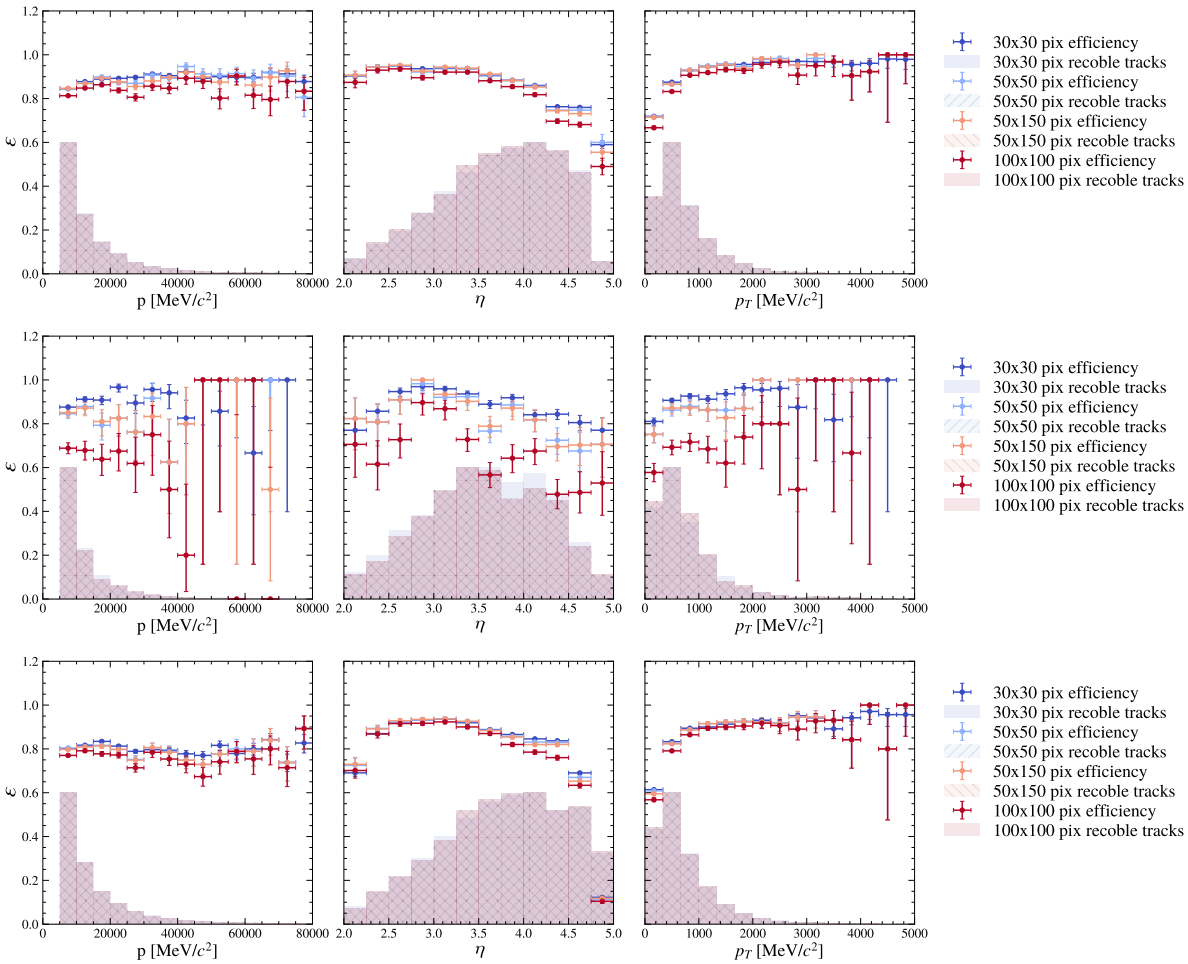


Figure 11.13: Comparison of the tracking efficiency between the four different UP pixel sizes for *pp* collisions. From top to bottom, the results of the tracking efficiencies of upstream, downstream and long tracks.

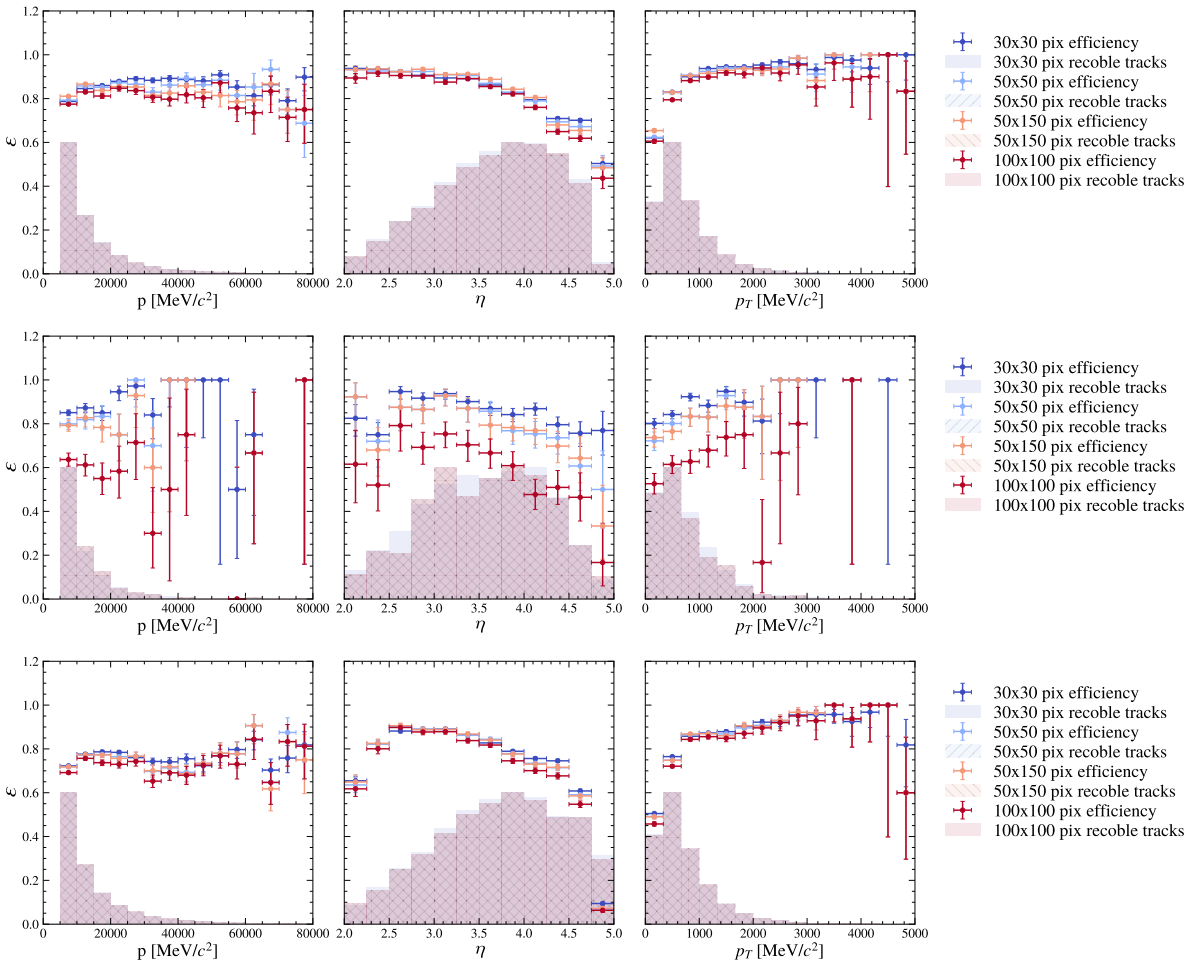


Figure 11.14: Comparison of the tracking efficiency between the four different UP pixel sizes for PbPb collisions. From top to bottom, the results of the tracking efficiencies of upstream, downstream and long tracks.

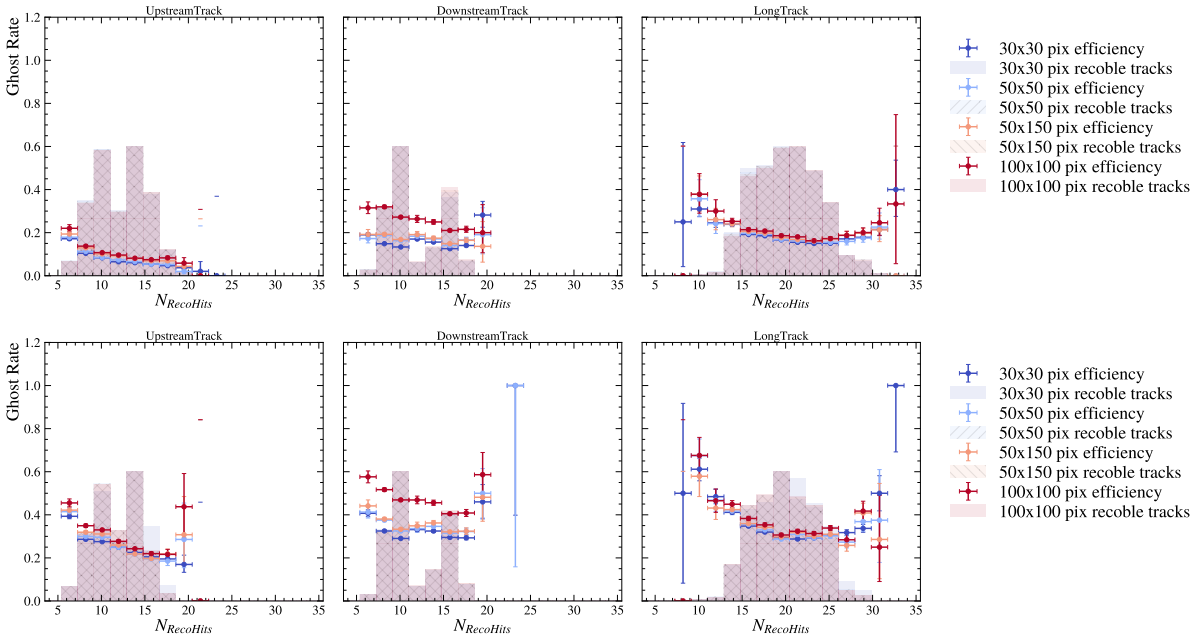


Figure 11.15: Comparison of the ghost rate as a function of the number of reconstructed hits between the four different UP pixel sizes for upstream, downstream and long tracks. In the top row, the results for *pp* collisions and in the bottom row the results for PbPb collisions.

remains modest, at approximately 7%, demonstrating a robust tracking performance for all three track categories.

The comparison of the ghost-track rates between *pp* and PbPb collisions is presented in Fig. 11.17. As expected, the ghost-track rates are lower in *pp* collisions than in PbPb collisions. The difference between the two collision systems is more visible than for the tracking efficiency, reflecting the larger probability of forming fake tracks in the high-occupancy environment of central PbPb collisions. Despite this increase, the overall reconstruction performance remains relatively close to that obtained in *pp* collisions, illustrating the robustness of the Upgrade II tracking system in extreme multiplicity conditions.

These results were incorporated into the LHCb Upgrade II Scoping Document [134] and the internal note describing the heavy-ion physics prospects of LHCb following Upgrade II [138]. They provided important input to the collaboration for evaluating detector performance under different operating conditions and physics scenarios.

Overall, these studies indicate that the increased granularity of the Upgrade II tracking detectors and the adoption of silicon-pixel technology will enable the reconstruction of central PbPb collisions with an efficiency approaching that achieved in *pp* collisions. This represents a major improvement with respect to the current detector capabilities and considerably enhances the heavy-ion physics potential of LHCb at the HL-LHC.

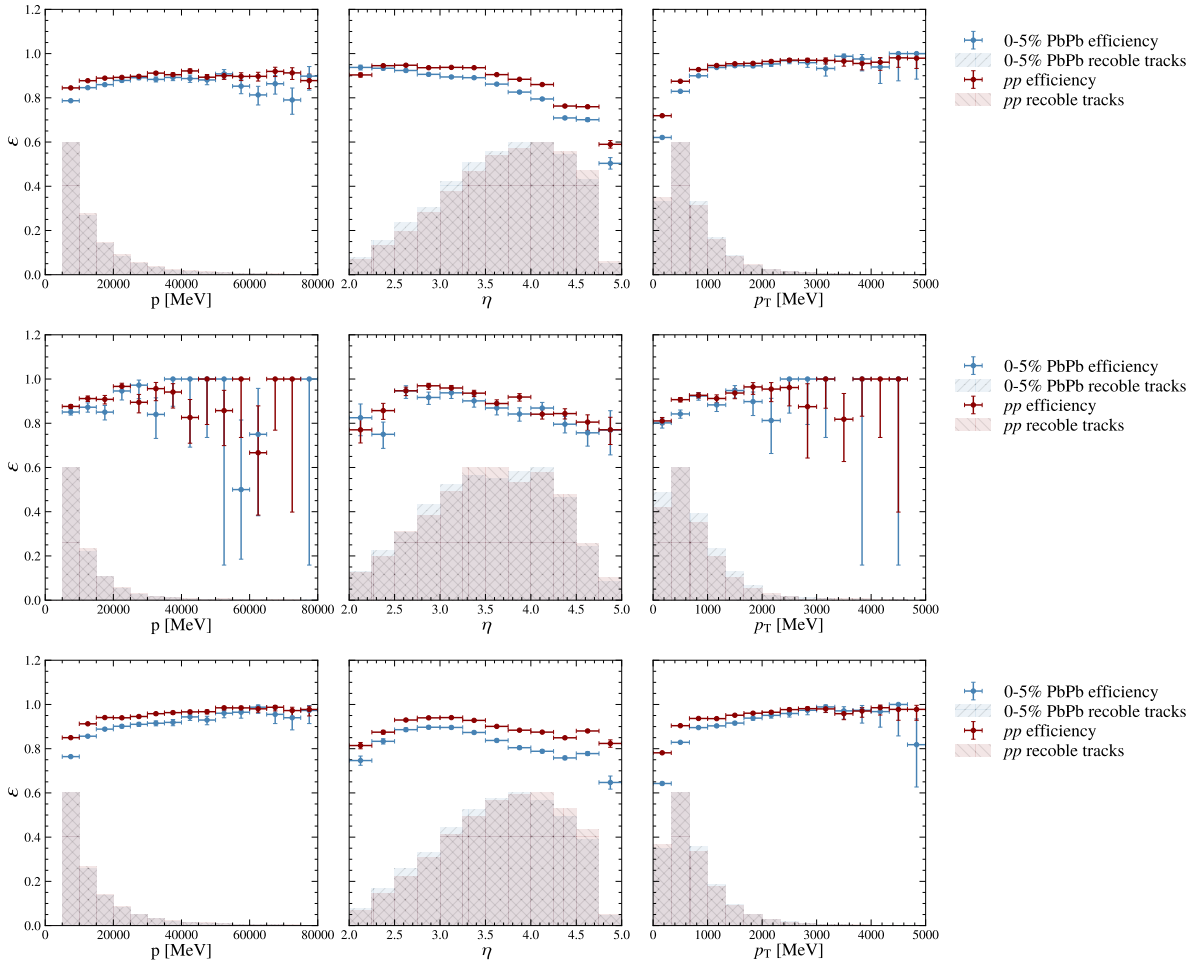


Figure 11.16: Comparison of the tracking efficiency between *pp* and PbPb collisions. From top to bottom, the results of the tracking efficiencies of upstream, downstream and long tracks.

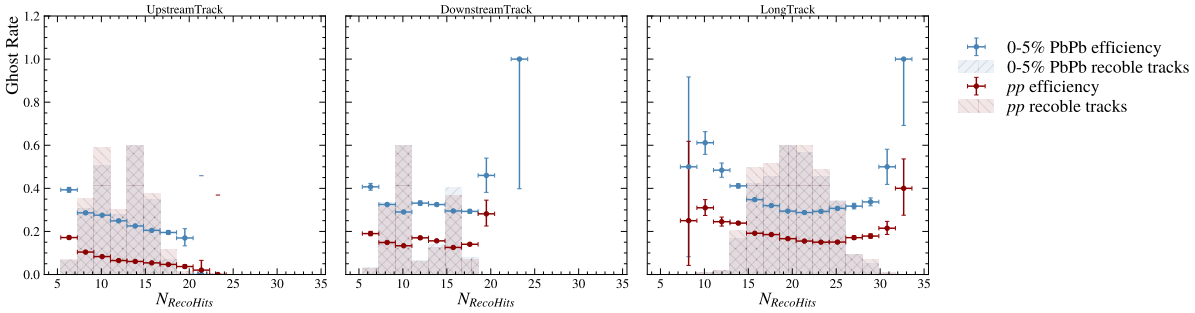


Figure 11.17: Comparison of the ghost rate as a function of the number of reconstructed hits for upstream, downstream and long tracks between *pp* and PbPb collisions.

11.4 Cellular Automaton Tracking algorithm

In the previous section several studies of how the geometry and the design of the detector affect to the tracking performance. However, the tracking performance not only rely on the hardware, but also depends on the algorithm used to perform the track reconstruction. The individual tracking efficiency of the standalone tracks in each tracking subdetector has an impact in the subsequent reconstruction of upstream, downstream and long tracks. Therefore, the improvement and optimization of the standalone algorithms is an essential part of the development of the software for the LHCb Upgrade II. In the case of UP standalone track reconstruction, Sec. 11.3.1 demonstrated that the current reconstruction algorithm is limited to a maximum overall tracking efficiency of 78% in pp collisions. This result motivated the development of a new UP standalone track reconstruction algorithm based on a Cellular Automaton [139, 140].

11.4.1 Cellular Automaton applied to tracking

A cellular automaton (CA) is a discrete computational model describing the evolution of a complex system through the repeated application of simple local rules. Despite its simple construction, a cellular automaton can exhibit highly non-trivial emergent behavior. One of the best-known examples is the *Conway's Game of Life* [141]. Formally, a cellular automaton consists of:

- **Lattice of cells**: a regular grid of cells (minimal elements that compose the system) arranged in one, two, or more dimensions. Each cell occupies a fixed position in the lattice.
- **Finite set of states**: each cell can be in one of a finite number of possible states.
- **Neighborhood definition**: for each cell, a set of neighboring cells is specified.
- **Local update rule**: a deterministic or probabilistic rule determines the state of each cell at the next time step as a function of its current state and the states of its neighbors.
- **Discrete time evolution**: all cells are updated simultaneously at successive time steps according to the update rule.

Cellular automata possess several properties that make them attractive for computational applications. Their evolution is governed by local interactions, meaning that the state of a cell depends only on its immediate neighbors rather than on the global state of the system. This locality naturally enables parallel execution, as all cells can be updated simultaneously at each time step. Furthermore, cellular automata are highly scalable, with a computational complexity typically growing approximately linearly with the number of cells. Despite the simplicity of their local update rules, they are capable of exhibiting emergent behavior, whereby complex global structures and patterns arise from the

repeated application of simple interactions. Finally, cellular automata offer considerable flexibility, as their evolution can be described by either deterministic or stochastic transition rules.

These properties make cellular automata particularly suitable for track-finding algorithms. Detector hits can be represented as cells or graph nodes, and local compatibility criteria is used to connect neighboring hits into track segments. Through iterative rule application, long particle trajectories emerge from locally constructed connections, providing a computationally efficient pattern-recognition technique in environments with high detector occupancy. With an appropriate choice of evolution rules, a cellular automaton can naturally accommodate small deviations from ideal trajectories induced by multiple scattering in the detector material or by the magnetic field.

Cellular automata have been successfully employed for parallelized track reconstruction, particularly in vertex detectors, in experiments such as HERA-B [142, 143], CMS [144] and ALICE [145]. However, this reconstruction strategy has never been explored within the LHCb experiment. In the following, a Cellular Automaton Tracking (CAT) algorithm is developed for standalone reconstruction in the UP detector.

11.4.2 Implementation of UP CAT algorithm

The CAT algorithm starts by constructing the elementary cells that will compose the cellular automaton. In this implementation, the cells correspond to *tracklets*, defined as hit pairs connecting two UP stations. Each tracklet stores the information of the hits it comprises together with the parameters of the straight line joining them, namely the intercept and slope in the xz and yz projections. Tracklets connecting non-consecutive stations are also allowed, enabling the reconstruction of tracks with one missing hit. Starting from a hit in station j, with $j \in \{1, 2, 3, 4\}$, all possible combinations with hits belonging to stations $j + 1$ and $j + 2$ and lying inside predefined search windows are considered. This leads to five possible track topologies, depending on the station combinations used to build the tracklets. A schematic representation of these configurations is shown in Fig. 11.18. Different search windows, characterized by opening angles $\Delta\phi_x$ and $\Delta\phi_y$ in the xz and yz projections, are defined for the various station combinations. Since this stage requires testing all possible hit combinations within these windows, it constitutes the most computationally demanding part of the algorithm.

To mitigate this issue, a BDT-based pre-selection is introduced, following a strategy similar to that employed in the matching algorithms described in Sections 11.2.4 and 11.2.5. The BDT is trained exclusively using the hit coordinates in the different UP stations and predicts whether two hits are compatible with originating from the same particle trajectory. The influence of this BDT filter on the CAT performance is studied in Section 11.3. At this stage, all tracklets are initialized with an evolution state, which serves as an internal score and is updated throughout the subsequent steps of the algorithm. Their state will subsequently evolve according to local compatibility rules during the Run Forward

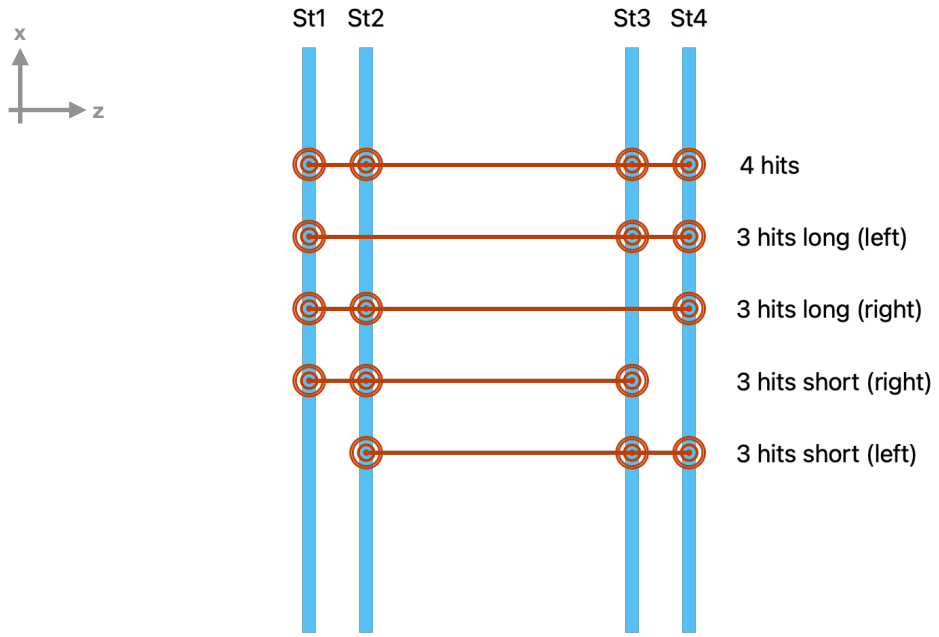


Figure 11.18: Scheme of the different possibilities of UP standalone track based on the tracklets that make it up.

procedure.

Run Forward

In this step, the evolution rules of the cellular automaton are applied and all tracklets evolve simultaneously according to these rules. Since the UP detector is composed of four stations, a complete standalone track can contain either two or three tracklets. For each tracklet state s_j, with $j \in \{1, 2, 3\}$, neighboring tracklets are searched from left to right, i.e. along the positive direction of the z axis. Two tracklets are considered neighbors if they share a common hit.

At the beginning of the procedure, all tracklets are assigned an initial evolution state equal to one. A tracklet located on the right evolves to a new evolution state, increasing its score by one unit, if the following conditions are satisfied:

- The two neighboring tracklets have the same evolution state.
- The breaking angle between them does not exceed a maximum value $\varphi_{\max}$.

The breaking angle between two tracklet states is defined as

$$\varphi(s_j, s_{j+1}) = \sqrt{(t_{x,j+1} - t_{x,j})^2 + (t_{y,j+1} - t_{y,j})^2}, \tag{11.5}$$

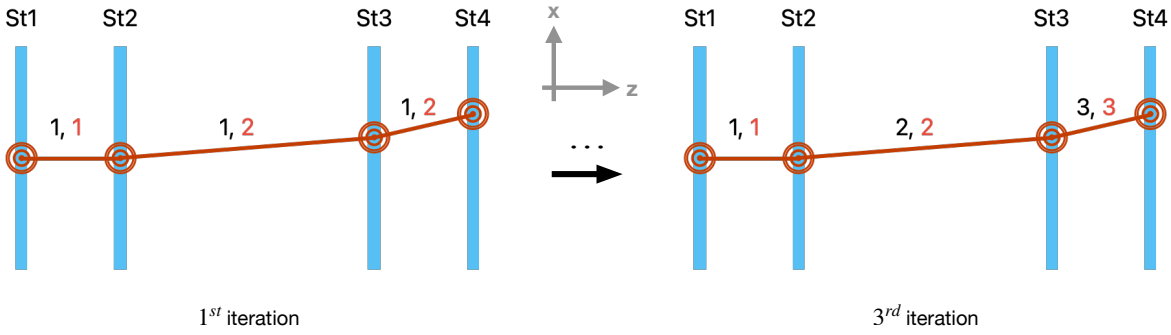


Figure 11.19: Evolution of the tracklets during the Run forward step. The current evolution states are presented in black, while the new evolution states to be updated at the end of the iteration are presented in red.

where $t_{x,j}$ and $t_{y,j}$ are the slopes of tracklet s_j in the xz and yz planes, respectively.

The evolution procedure is iterated until no further changes in the evolution states occur. At the end of each iteration, all evolution states are updated synchronously. A schematic representation of the Run Forward procedure is shown in Fig. 11.19. Once the Run Forward step has converged, each tracklet is associated with an evolution state and the track-finding procedure proceeds to the *Run Backward* step.

Run Backward

In this step, track candidates are constructed by searching ordered sequences of neighboring tracklets from right to left, i.e. along the negative direction of the z axis. A valid sequence must satisfy the requirement that the evolution state of the tracklet on the right is larger than that of the tracklet on the left, as illustrated in Fig. 11.20. Several compatible sequences may satisfy the ordering requirement. A quality estimator is therefore introduced to identify the most likely particle trajectory, J, defined as

$$J = N - \gamma \sum_{j=1}^{N-1} \frac{\varphi(s_{j+1}, s_j)}{z_{R,j+1} - z_{L,j}} \tag{11.6}$$

where N denotes the number of tracklets composing the sequence, $\varphi(s_{j+1}, s_j)$ is the breaking angle defined in Eq. 11.5, and $z_{R,j+1} - z_{L,j}$ corresponds to the separation in the z direction between two neighboring tracklets. The parameter γ controls the trade-off between favoring long sequences and penalizing large trajectory deviations. Since the magnetic field in the UP region is relatively weak, only small curvatures are expected, while additional deviations may arise from multiple scattering in the detector material. The value

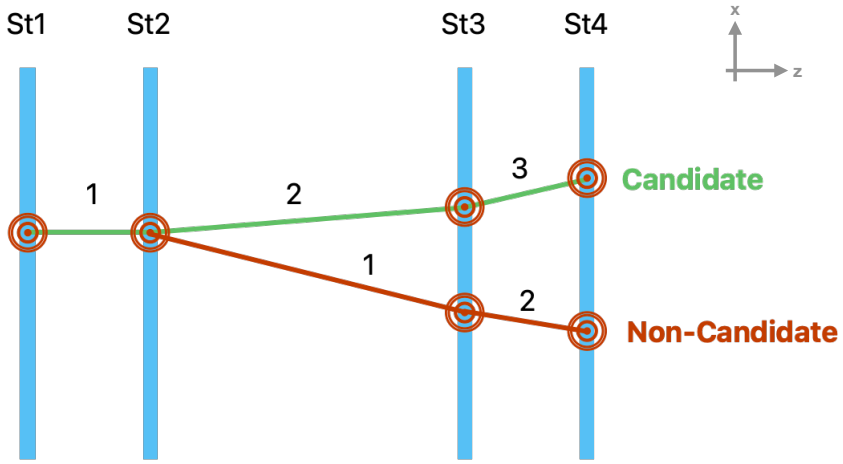


Figure 11.20: Schematic illustration of the tracklet-sequence search during the Run Backward step. The green sequence corresponds to a valid track candidate, whereas the red sequence is rejected because it does not satisfy the requirement of an ordered sequence.

of γ was therefore tuned empirically to account for these effects.

At the end of the Run Backward procedure, a complete collection of track candidates is obtained. The final stage of the CAT algorithm is then dedicated to assessing their quality and selecting the best candidates.

Track quality check

The final stage of the UP CAT algorithm evaluates the quality of the track candidates selected during the Run Backward step. Although the Run Backward procedure identifies geometrically compatible sequences of tracklets, additional quality criteria are required to suppress residual fake combinations.

Each tracklet sequence is fitted using a parabolic model in the xz plane and a linear model in the yz plane. Track candidates are retained only if the fit χ^2 is below a predefined threshold.

Ideally, this step should rely on a Kalman Filter [146, 147] to determine the optimal track parameters. The Kalman filter sequentially propagates the track state through the detector layers and updates the track parameters using the measured hit positions, providing an optimal estimate of the trajectory together with its associated uncertainties. The implementation of such a procedure is, however, left for future developments.

Finally, a clone-removal procedure is applied. Tracks sharing two or more hits are considered clones. Within each clone family, only the candidate with the best fit quality is

retained, while candidates with larger χ^2 values are discarded.

11.4.3 CAT algorithm tests

The performance studies presented in this section are performed using the nominal UP configuration determined in Section 11.3. This configuration corresponds to the Borderless geometry, the original beam-pipe hole dimensions and a pixel size of 50×50 μm^2.

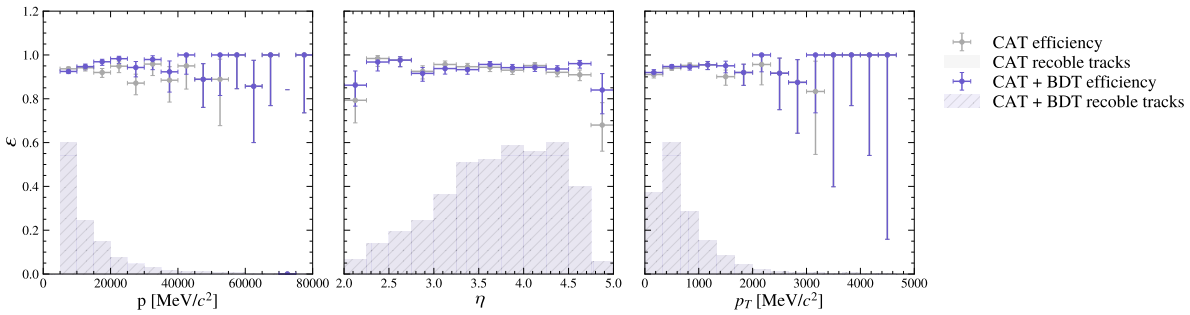


Figure 11.21: Comparison of the UP standalone tracking efficiency when the CAT algorithm uses or not a BDT filter in *pp* collisions.

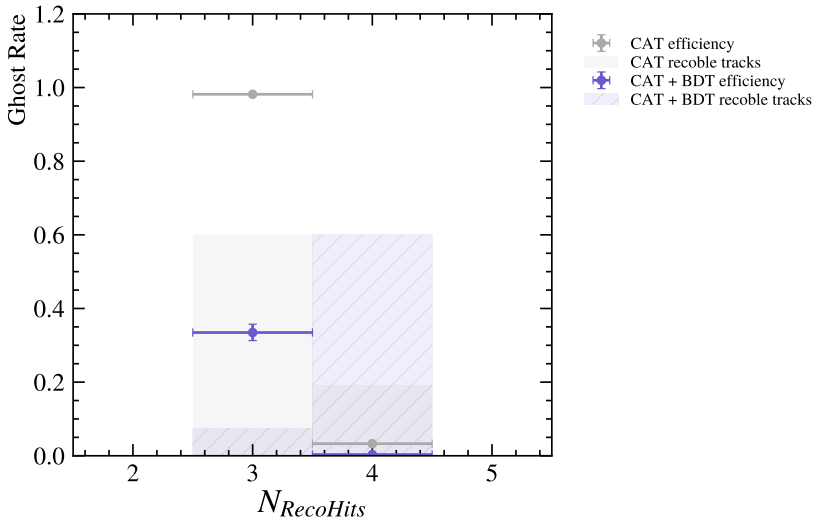


Figure 11.22: UP CAT standalone ghost rate as a function of the number of reconstructed hits together with the distributions of the reconstructible (recoble) tracks for the cases that include or not a BDT filter in *pp* collisions.

As discussed previously, the construction of tracklets requires evaluating all possible hit combinations within the search windows and therefore dominates the execution time of the CAT algorithm. A BDT-based pre-selection is introduced to alleviate this limitation.

In the following, the impact of the BDT filter on the CAT performance is evaluated by comparing the tracking efficiency and ghost-track rate obtained with and without the BDT selection.

The tracking efficiencies obtained in both configurations are shown in Fig. 11.21. No significant difference is observed, indicating that the BDT filter has a negligible impact on the tracking efficiency.

The corresponding ghost-track rates are presented in Fig. 11.22. In this case, the impact of the BDT filter is substantial. Without the BDT pre-selection, the algorithm reconstructs a larger fraction of three-hit tracks, leading to an overall ghost-track rate of approximately 75%. In contrast, when the BDT filter is applied, the reconstruction is dominated by four-hit tracks and the ghost-track rate is reduced to about 4%. In addition, the BDT considerably decreases the execution time of the algorithm by reducing the number of hit combinations that need to be explored.

These results demonstrate that the BDT filter preserves the tracking efficiency while being essential for suppressing ghost tracks and reducing the execution time of the CAT algorithm.

To assess the robustness of the CAT algorithm, its performance is studied for different values of the intrinsic pixel efficiency, defined as the probability for a charged particle traversing a pixel sensor to produce a recorded hit. Since the CAT algorithm operates only on the reconstructed hits available in the UP detector, its performance is expected to be only weakly dependent on the intrinsic pixel efficiency. Variations of this quantity mainly affect the number of available hits rather than the reconstruction procedure itself.

The performance of the CAT algorithm is therefore evaluated for hit efficiencies of 100%, 98% and 96%. The corresponding tracking efficiencies obtained in pp collisions are presented in Fig. 11.23. No significant differences are observed among the three configurations, with an overall tracking efficiency above 92% in all cases.

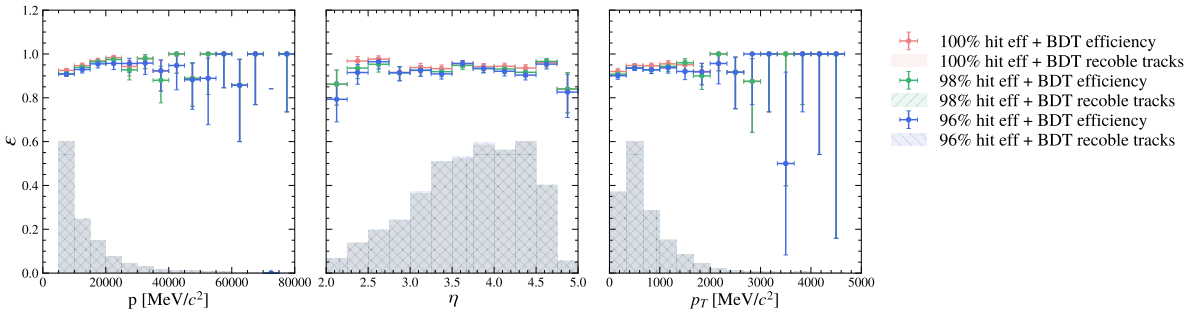


Figure 11.23: Comparison of the UP CAT tracking efficiency for three different scenarios of hit efficiency of the UP pixels in pp collisions.

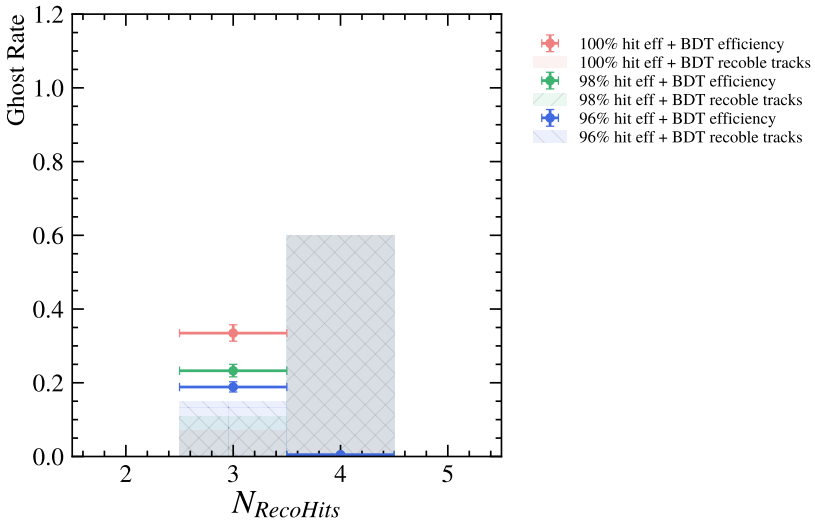


Figure 11.24: UP CAT standalone ghost rate as a function of the number of reconstructed hits together with the distributions of the reconstructible (recoble) tracks for the three different cases of hit efficiency of the UP pixels in *pp* collisions.

A similar behavior is observed for the ghost-track rate, shown in Fig. 11.24. The ghost-track rates remain very similar for the three hit-efficiency scenarios, staying below 5% in all cases. These results indicate that the CAT algorithm is largely insensitive to realistic variations of the intrinsic pixel efficiency and therefore appears robust against moderate detector inefficiencies.

11.4.4 CAT algorithm in different collision systems

As a final performance study, the CAT algorithm is evaluated in *pp* and very central PbPb collisions. This comparison allows its behavior to be assessed under significantly different occupancy conditions. In addition, the overall performance is compared with that of the standalone UP reconstruction algorithm introduced in Section 11.2.2.

The comparison of the tracking efficiencies obtained in *pp* and PbPb collisions is presented in Fig. 11.25. As expected, the efficiency achieved in *pp* collisions is slightly higher than in PbPb collisions owing to the increased detector occupancy in central heavy-ion events. Nevertheless, the efficiency difference between the two collision systems remains below 4%, indicating that the CAT algorithm preserves a high reconstruction efficiency even in extreme multiplicity environments.

The comparison of the ghost-track rates between *pp* and PbPb collisions is shown in Fig. 11.26. In this case, the difference between the two collision systems is more pronounced. The larger ghost-track rate observed in PbPb collisions is expected, as the increase in detector occupancy results in a larger number of possible hit combinations

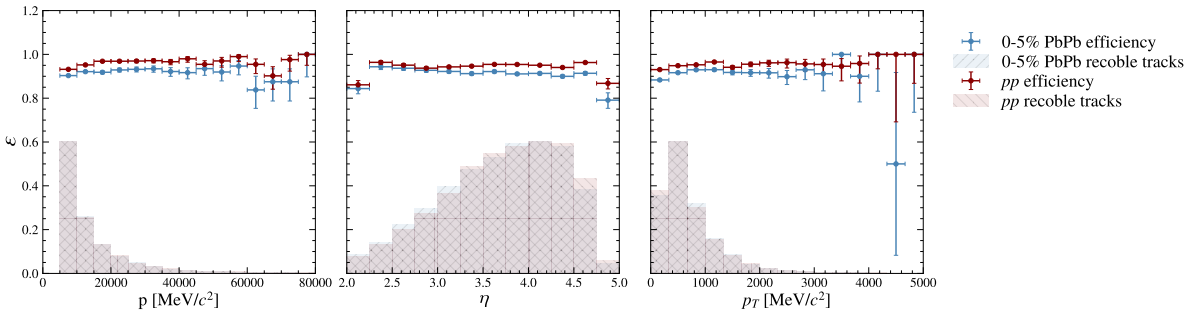


Figure 11.25: Comparison of the UP CAT tracking efficiency in *pp* and high centrality PbPb collisions.

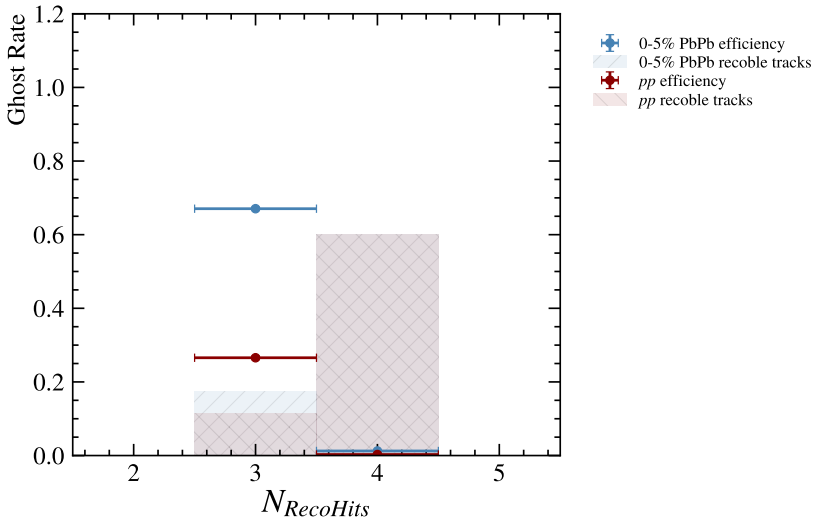


Figure 11.26: UP CAT standalone ghost rate as a function of the number of reconstructed hits together with the distributions of the reconstructible (recoble) tracks for *pp* and high centrality PbPb collisions.

and therefore enhances the probability of reconstructing fake tracks. Despite this degradation, the CAT algorithm maintains an acceptable ghost-track rate in the most challenging occupancy conditions considered in this work.

The performance of the CAT algorithm is finally compared with that of the previous standalone UP reconstruction algorithm. The overall tracking efficiencies and ghost-track rates achieved by the two approaches are summarized in Tables 11.1 and 11.2, respectively.

A substantial improvement in tracking efficiency is obtained with the CAT algorithm. In *pp* collisions, the efficiency increases from approximately 79% to 95%, while in central PbPb collisions it rises from 75% to more than 91%. Inspection of Figs. 11.8 and 11.25 further indicates that the largest gains are achieved at low and high pseudorapidities,

together with a more stable performance at large transverse momentum.

This improvement comes at the cost of an increased ghost-track rate. While the previous standalone algorithm strongly favors the reconstruction of four-hit tracks, the CAT algorithm reconstructs a larger fraction of three-hit candidates, which constitute the dominant source of ghost tracks. Consequently, the two algorithms exhibit different compromises between tracking efficiency and purity.

Tracking efficiency		
Algorithm	ε (pp) (%)	ε (PbPb) (%)
UP standalone	78.69 ± 0.30	75.05 ± 0.36
CAT	94.70 ± 0.17	91.32 ± 0.23

Table 11.1: Overall tracking efficiency obtained with the standalone UP and CAT algorithms in *pp* and central PbPb collisions. Uncertainties are statistical only.

Ghost rate		
Algorithm	GR (pp) (%)	GR (PbPb) (%)
UP standalone	0.65 ± 0.05	3.79 ± 0.12
CAT	4.48 ± 0.10	15.99 ± 0.18

Table 11.2: Overall ghost-track rate obtained with the standalone UP and CAT algorithms in *pp* and central PbPb collisions. Uncertainties are statistical only.

The results obtained with the CAT algorithm are highly encouraging, not only in *pp* collisions but also in the highest-centrality PbPb events considered in this work. Although the ghost-track rate remains larger than that achieved with the previous standalone algorithm, the substantial gain in tracking efficiency clearly demonstrates the potential of cellular-automaton-based approaches for standalone reconstruction in Upgrade II. Further developments, such as the implementation of a Kalman filter and more sophisticated clone-rejection strategies, are expected to improve the purity of the reconstructed track sample.

11.5 Summary

In this chapter, the tracking performance expected for the LHCb Upgrade II detector was studied using a dedicated Pattern Recognition framework.

Several detector configurations of the Upstream Pixel detector were investigated, including alternative geometries, beam-pipe apertures and pixel granularities. These studies showed that the Borderless geometry with a 50×50 μm^2 pixel size provides an optimal compromise between tracking performance, detector acceptance and material budget.

The performance of the nominal Upgrade II detector was evaluated in both *pp* and central PbPb collisions. Similar tracking efficiencies were obtained in the two collision systems, demonstrating the capability of Upgrade II detector to operate efficiently in the extreme occupancy conditions expected during heavy-ion running.

Finally, a novel Cellular Automaton Tracking algorithm was developed for standalone reconstruction in the UP detector. The algorithm increases the standalone tracking efficiency from approximately 79% to 95% in *pp* collisions and from 75% to more than 91% in central PbPb collisions, while maintaining an acceptable ghost-track rate. These results demonstrate the potential of cellular-automaton-based approaches for standalone tracking in Upgrade II and provide a basis for future developments including Kalman-filter fitting and improved clone-rejection procedures.

12
Conclusions

This thesis has addressed two complementary aspects of experimental high-energy physics at LHCb: a precision measurement of excited charmonium production in a light-ion collision system and detector-development studies aimed at preserving the tracking performance of the experiment throughout the High-Luminosity LHC era. The two parts of this thesis, although addressing distinct questions, reflect the dual character of experimental particle physics research, in which the extraction of new physics results and the development of the instrumentation needed to obtain them proceed hand in hand.

χ_c production in OO collisions

The first part of this thesis has presented the first measurement of $\chi_c(1P)$ production in ion-ion collisions at collider energies, using Oxygen-Oxygen data collected by LHCb during the 2025 light-ion run at $\sqrt{s_{NN}} = 5.36$ TeV. The χ_{c1} and χ_{c2} states were reconstructed through their radiative decay to $J/\psi\,\gamma$, with the J/ψ identified in the dimuon final state. This measurement was made possible by a dedicated programme of work spanning the calibration of the electromagnetic calorimeter, the determination of the low-energy photon detection efficiency, and the design of machine-learning-based selection strategies capable of isolating the χ_c signal in the high-occupancy environment characteristic of nucleus-nucleus collisions.

The analysis was originally conceived for PbPb collisions, where the χ_c signal proved difficult to extract due to the high combinatorial background and calorimeter occupancy inherent to that system, compounded by the limited available statistics. This thesis has shown that redirecting the measurement towards the smaller and cleaner OO system resolves these difficulties while still providing access to a hot and dense medium relevant to

QGP studies, establishing OO collisions as a tractable and physically interesting environment.

Two principal observables were extracted from the analysis. The prompt-production cross-section ratio was measured to be

$$\frac{\sigma(\chi_{c2})}{\sigma(\chi_{c1})} = 2.17 \pm 0.25\,(\text{stat.}) \pm 0.25\,(\text{syst.}), \tag{12.1}$$

with statistical and systematic uncertainties of comparable size. This result is found to lie above the corresponding ratios measured by LHCb in pp and pPb collisions, although the comparison must be qualified by the different transverse-momentum ranges covered by the various measurements: the present analysis extends down to $p_T(\chi_c) > 1.5$ GeV/c, below the range accessible in earlier measurements, and the observed p_T-dependence of the ratio in pp collisions suggests that this lower reach can largely account for the apparent enhancement. The measured ratio is consistent within 1.5σ with a dedicated Pythia Angantyr simulation, providing confidence in the modeling of relative χ_{c2}/χ_{c1} production in this kinematic regime.

The fraction of prompt J/ψ mesons originating from χ_c feed-down was determined to be

$$F(\chi_c \to J/\psi) = 0.124 \pm 0.004\,(\text{stat.}) \pm 0.010\,(\text{syst.}), \tag{12.2}$$

with the individual χ_{c1} and χ_{c2} contributions measured separately. This is the first such determination in a nucleus-nucleus collision system at collider energies, and the result is found to lie below previous measurements in pp and pPb collisions, by approximately 2.9σ with respect to the pPb result when considering the uncertainties currently available. If confirmed, this reduction would be consistent with an enhanced suppression of χ_c states relative to direct J/ψ production in the medium formed in OO collisions, in line with the expectation that the more weakly bound χ_c states should be more strongly affected by the QGP medium effects than the J/ψ. However, this interpretation must be regarded as preliminary: several sources of systematic uncertainty affecting the feed-down measurement have not yet been fully evaluated, and a quantitative comparison across collision systems is complicated by their differing kinematic acceptances. A firm conclusion regarding medium-induced suppression of χ_c states in OO collisions will therefore require the completion of the systematic-uncertainty programme begun in this work.

Independently of the final interpretation of this comparison, the analysis chain developed in this thesis is complete and has been shown to function reliably, from candidate selection through to the extraction of both observables. Beyond the measured values themselves, an important achievement of this work is the demonstration that precision χ_c measurements are feasible in a heavy-ion environment at collider energies. This thesis establishes the experimental tools required to pursue this physics case with larger heavy-ion datasets collected by LHCb, including dedicated calorimeter calibrations, low-energy photon efficiency determinations, and background-rejection methodologies. The sequential suppression pattern expected for excited charmonium states in the QGP remains one

of the most direct experimental probes of the thermodynamic properties of the medium, and the present work lays the necessary groundwork for exploiting this probe in light and heavy nuclear systems at the LHC.

Tracking performance for LHCb Upgrade II

The second part of this thesis has addressed a different, yet equally necessary, ingredient of the LHCb physics programme: ensuring that tracking performance can be preserved under the extreme occupancy conditions expected at the High-Luminosity LHC. Using a dedicated Pattern Recognition framework, this work has evaluated the performance of several designs for the Upstream Pixel (UP) detector and has developed a new reconstruction algorithm for its standalone operation.

Several detector layouts, beam-pipe apertures and pixel granularities were investigated. These studies led to the selection of a Borderless geometry equipped with 50×50 μm^2 pixels as the nominal UP configuration, providing an optimal compromise between tracking performance, detector acceptance, material budget and technological feasibility. Using this configuration, the performance of upstream, downstream and long-track reconstruction was evaluated in both *pp* and central PbPb collisions. A modest efficiency reduction of approximately 7% was observed in the heavy-ion environment relative to *pp* collisions, despite the substantially larger detector occupancy. For comparison, the tracking efficiency of the current Run 3 detector can degrade by more than 30% when comparing *pp* and PbPb collisions. This result demonstrates the robustness of the proposed Upgrade II detector and reconstruction strategy, supports the viability of the proposed UP design across the broad range of collision systems relevant for the Upgrade II physics programme, and contributed directly to the optimization of the UP detector layout within the LHCb collaboration.

Building on these results, a Cellular Automaton Tracking (CAT) algorithm was developed and validated for standalone reconstruction in the UP detector. The CAT algorithm increases the standalone tracking efficiency from approximately 79% to 95% in *pp* collisions, and from 75% to more than 91% in the highest-occupancy central PbPb collisions considered in this work, while keeping the associated ghost-track rate at an acceptable level. This represents a substantial improvement over the existing standalone algorithm and demonstrates the potential of cellular-automaton-based approaches for track reconstruction under the demanding conditions expected at Upgrade II. The increase in ghost-track rate accompanying this gain in efficiency also points towards a clear direction for future developments, notably the implementation of a Kalman-filter-based track fit and more sophisticated clone-rejection strategies, which should further improve the purity of the reconstructed sample without sacrificing the efficiency gains already achieved.

Although the two parts of this thesis address different aspects of the LHCb physics programme, they ultimately pursue a common objective: improving our understanding of

charmonium production in nuclear collisions. The first part demonstrates that precision measurements of excited charmonium states in ion-ion collisions at collider energies have now become experimentally accessible, while the second part contributes to the detector and reconstruction developments that will enable such measurements to be extended to the High-Luminosity LHC era. Together, these studies connect present-day measurements with the preparation of the next generation of heavy-ion investigations at the LHC.

A

Permissions and reuse licenses

This annex lists the figures that are not original from this work with the source, rights holder and license. Material under Creative Commons licences is reused in accordance with the corresponding licence terms.

Figure	**Page**	**Description**	**Source**	**Rights holder**	**Licence**
Fig. 4.1	21	Elementary particles of the Standard Model	Wikimedia Commons contributors, Standard model of elementary particles (file by MissMJ)	MissMJ	CC BY 3.0
Fig. 4.2	23	Determinations of α_s as a function of energy scale Q	Particle Data Group, Review of Particle Physics	Particle Data Group	CC BY 4.0
Fig. 4.3	24	Layout of the hadronic matter phase diagram	G. Martinez Garcia, Advances in Quark Gluon Plasma, arXiv:1304.1452	The author / arXiv	CC BY 4.0
Fig. 4.5	28	Space-time evolution of a heavy-ion collision	G. Martinez Garcia, Advances in Quark Gluon Plasma, arXiv:1304.1452	The author / arXiv	CC BY 4.0

Figure	Page	Description	Source	Rights holder	Licence
Fig. 4.6	30	Glauber model: side view and beam-line view	M. L. Miller et al., Glauber Modeling in High-Energy Nuclear Collisions, Annu. Rev. Nucl. Part. Sci. 57 (2007) 205	The author(s) / arXiv	CC BY 4.0
Fig. 4.7	32	Level scheme of $c\bar{c}$ meson states	Particle Data Group, S. Navas et al., Review of Particle Physics, Phys. Rev. D 110 (2024) 030001	The Author(s), Particle Data Group	CC BY 4.0
Fig. 4.8	34	Two stages of prompt quarkonium production	A. Rothkopf, Heavy Quarkonium in Extreme Conditions, Phys. Rep. 858 (2020) 1	The author (A. Rothkopf)	CC BY 4.0
Fig. 4.9	34	Heavy quark pair production processes in pp collisions	A. Rothkopf, Heavy Quarkonium in Extreme Conditions, Phys. Rep. 858 (2020) 1	The author (A. Rothkopf)	CC BY 4.0
Fig. 4.10	36	Nuclear modifications of gluon PDFs (Pb and O)	P. Paakkinen, Nuclear PDFs at the beginning of LHC Run 3, PoS LHCP2022 (2023) 137	The author (P. Paakkinen)	CC BY-NC-ND 4.0
Fig. 4.12	40	Total (extrapolated) $c\bar{c}$ cross section vs $\sqrt{s}$	A. Andronic et al., Heavy-flavour and quarkonium production in the LHC era, Eur. Phys. J. C 76 (2016) 107	The Author(s)	CC BY 4.0
Fig. 4.13	42	Inclusive J/ψ nuclear modification factor, ALICE (a, b)	ALICE Collaboration, S. Acharya et al., Phys. Lett. B 849 (2024) 138451	The Author(s), ALICE Collaboration	CC BY 4.0
Fig. 4.14	43	$J/\psi/D^0$ ratio vs Ncoll: LHCb (a) and ALICE (b)	LHCb Collaboration, R. Aaij et al., Eur. Phys. J. C 83 (2023) 658; ALICE Collaboration, S. Acharya et al., Phys. Lett. B 849 (2024) 138451	CERN for the benefit of the LHCb Collaboration; The Author(s), ALICE Collaboration	CC BY 4.0

Figure	**Page**	**Description**	**Source**	**Rights holder**	**Licence**
Fig. 4.15	44	Fraction of χ_c decays in prompt J/ψ, pPb and pp	LHCb Collaboration, R. Aaij et al., Phys. Rev. Lett. 132 (2024) 102302	The Author(s), LHCb Collaboration	CC BY 4.0
Fig. 4.16	45	$\psi(2S)/J/\psi$ cross-section ratio vs number of participants (a, b)	LHCb Collaboration, R. Aaij et al., JHEP 07 (2025) 235	CERN for the benefit of the LHCb Collaboration	CC BY 4.0
Fig. 4.17	46	Typical sources of hadroproduced J/ψ	J.-P. Lansberg, New observables in inclusive production of quarkonia, Phys. Rep. 889 (2020) 1	The author / arXiv	CC BY 4.0
Fig. 5.1	48	CERN accelerator complex in 2022	E. Lopienska, The CERN accelerator complex, layout in 2022, CERN-GRAPHICS-2022-001	CERN	© CERN. Image freely available for non-commercial use
Fig. 5.2	49	Side view of the LHCb upgraded detector	LHCb Collaboration, R. Aaij et al., The LHCb Upgrade I, JINST 19 (2024) P05065	CERN for the benefit of the LHCb Collaboration	CC BY 4.0
Fig. 5.3	50	Recorded luminosity for pp and PbPb collisions (a, b)	CERN, LHCb online luminosity and fill summary (2025 operation test) [83, 84]	CERN / LHCb	Reproduced with permission of the LHCb Collaboration (collaboration policy on the use of LHCb figures in members' theses)
Fig. 5.5	52	VELO foils layout and ASICs layout	LHCb Collaboration, R. Aaij et al., The LHCb Upgrade I, JINST 19 (2024) P05065	CERN for the benefit of the LHCb Collaboration	CC BY 4.0
Fig. 5.6	53	Four UT silicon planes	LHCb Collaboration, R. Aaij et al., The LHCb Upgrade I, JINST 19 (2024) P05065	CERN for the benefit of the LHCb Collaboration	CC BY 4.0
Fig. 5.7	55	SciFi layers and stations T1-T3	LHCb Collaboration, R. Aaij et al., The LHCb Upgrade I, JINST 19 (2024) P05065	CERN for the benefit of the LHCb Collaboration	CC BY 4.0

Figure	Page	Description	Source	Rights holder	Licence
Fig. 5.8	56	RICH1 and RICH2 schematic view	LHCb Collaboration, R. Aaij et al., The LHCb Upgrade I, JINST 19 (2024) P05065	CERN for the benefit of the LHCb Collaboration	CC BY 4.0
Fig. 5.9	57	Lateral segmentation of ECAL and HCAL	LHCb Collaboration, R. Aaij et al., The LHCb Upgrade I, JINST 19 (2024) P05065	CERN for the benefit of the LHCb Collaboration	CC BY 4.0
Fig. 5.10	58	Muon stations, side and front views	LHCb Collaboration, LHCb PID Upgrade Technical Design Report, CERN-LHCC-2013-022	CERN	CC BY 4.0
Fig. 5.11	59	Track types in the LHCb tracking system	S. Aiola et al., Hybrid seeding, Comput. Phys. Commun. 260 (2021) 107713	Elsevier	CC BY-NC-ND 4.0 (accepted manuscript)
Fig. 5.12	63	Online dataflow scheme	LHCb Collaboration, RTA and DPA dataflow diagrams for Run 1, Run 2, and the upgraded LHCb detector, LHCB-FIGURE-2020-016	CERN for the benefit of the LHCb Collaboration	CC BY 4.0
Fig. 6.1	65	B^0 mass shift by ECAL region, pp data 2024	M. Calvo, Photon PID in Run 3, presentation at the LHCb Collaboration Meeting, 13 February 2025	M. Calvo / LHCb Collaboration	Reproduced with permission of the LHCb Collaboration (collaboration policy on the use of LHCb figures in members' theses)
Fig. 6.2	66	Measured π^0 mass as a function of time 2024/2025	LHCb internal calibration plots	Jean-Francois Marchand / LHCb Collaboration	Reproduced with permission of the LHCb Collaboration (collaboration policy on the use of LHCb figures in members' theses)

Figure	**Page**	**Description**	**Source**	**Rights holder**	**Licence**
Fig. 6.3	69	Energy deposition: photon, merged π^0, hadron, pile-up	M. Calvo, Photon identification in Run 3, LHCb-INT-2025-004	M. Calvo / LHCb Collaboration	Reproduced with permission of the LHCb Collaboration (collaboration policy on the use of LHCb figures in members' theses)
Fig. 7.2	88	Schematic representation of a BDT	XGBoost developers, XGBoost documentation (release 3.2.0)	XGBoost contributors	Apache License 2.0
Fig. 10.1	153	LHCb luminosity profiles, Runs 1-5	LHCb Collaboration, LHCb Upgrade II Scoping Document, CERN-LHCC-2024-010	CERN for the benefit of the LHCb Collaboration	CC BY 4.0
Fig. 10.2	156	Photon source identification with BDT and azimuthal correlation	LHCb Collaboration, Opportunities for Heavy-Ion and Fixed-Target Physics with LHCb Upgrade II, LHCb-PUB-2024-007	CERN for the benefit of the LHCb Collaboration	CC BY 4.0
Fig. 10.3	157	Expected statistical uncertainty on R_{PbPb} of Υ states	LHCb Collaboration, Opportunities for Heavy-Ion and Fixed-Target Physics with LHCb Upgrade II, LHCb-PUB-2024-007	CERN for the benefit of the LHCb Collaboration	CC BY 4.0
Fig. 10.4	158	Side view of the Upgrade II baseline detector	LHCb Collaboration, LHCb Upgrade II Scoping Document, CERN-LHCC-2024-010	CERN for the benefit of the LHCb Collaboration	CC BY 4.0
Fig. 10.5	160	VELO tracking efficiency and IP resolution, Baseline vs Low	LHCb Collaboration, LHCb Upgrade II Scoping Document, CERN-LHCC-2024-010	CERN for the benefit of the LHCb Collaboration	CC BY 4.0
Fig. 11.3	168	Expected track reconstruction efficiency in the Mighty Tracker	LHCb Collaboration, LHCb Upgrade II Scoping Document, CERN-LHCC-2024-010	CERN for the benefit of the LHCb Collaboration	CC BY 4.0

List of Figures

4.1 Summary of the different elementary particles of the Standard Model of particle physics. Figure taken from [9] under Creative Commons license. . 27

4.2 Summary of determinations α_s as a function of the energy scale Q compared to the running of the coupling computed at five loops taking as an input the current Particle Data Group (PDG) average, $\alpha_s(m_Z^2) = 0.1180 \pm 0.0009$. Figure taken from [13] under Creative Commons license. 29

4.3 Layout of the hadronic matter phase diagram. Figure taken from [14] under Creative Commons license. 30

4.4 Diagram of the longitudinal heavy-ion collision evolution in the scenarios where no QGP is formed and the case where QGP is formed. 33

4.5 Scheme of the space-time evolution of a heavy-ion collision. Figure taken from [14] under Creative Commons license. 34

4.6 Schematic side view and beam-line view of a nucleus–nucleus collision according to the Glauber model description figure taken from [26] under Creative Commons license. 36

4.7 The level scheme of meson states containing a minimal quark content of $c\bar{c}$ as a function of their mass and their quantum numbers. Figure taken from [28] under Creative Commons license. 38

4.8 The two stages of the production process of prompt quarkonium: production of the color-bearing heavy quark-antiquark pair from hard partons and the subsequent formation of either open heavy flavor or quarkonium states. Figure taken from [29] under Creative Commons license. 40

4.9 Scattering processes in the production of heavy quark anti-quark pairs at proton-proton colliders. Figure taken from [29] under Creative Commons license. 40

4.10 Nuclear modifications of the gluon PDFs in lead (left) and oxygen (right) nuclei. Results from the EPPS21, nNNPDF3.0 and nCTEQ15WZSIH analyses. Figure taken from [43] under Creative Commons license. 42

4.11 Scheme of the survival probability of J/ψ as a function of the energy density. 45

4.12 Total (extrapolated) $c\bar{c}$ cross section as a function of $\sqrt{s}$. Data in pA or deuteron–nucleus (dA) collisions were scaled down assuming no nuclear effect. Figure taken from [54] under Creative Commons license. 46

4.13 Results on the inclusive J/ψ nuclear modification factor from ALICE at midrapidity (a) and at forward rapidity (b) taken from [60] compared to model calculations. Figures taken from [60] under Creative Commons license. 48
4.14 The production ratio of $J/\psi/D^0$ as function of N_{coll} in proton-Neon and lead-Neon collisions measured by LHCb (a) taken from [61] and the same ratio measured by ALICE at midrapidity compared with the statistical hadronization model (b) taken from [60]. Figures taken from indicated references under Creative Commons license. 49
4.15 Fraction of χ_c decays in the prompt J/ψ yield in pPb and pp collisions as a function of $p_T^{J/\psi}$. Figure taken from [62] under Creative Commons license. 50
4.16 Cross-section times branching fraction ratio between $\psi(2S)$ and J/ψ states as a function of the number of participating nucleon. Comparison with other collision systems (a) and with theoretical models (b). Figures taken from [66] under Creative Commons license. 51
4.17 Scheme of the typical sources of hadroproduced J/ψ for low-p_T J/ψ in the left and high-p_T J/ψ in the right. Figure taken from [68] under Creative Commons license. 52

5.1 Diagram of the CERN accelerator complex in 2022. It shows the different accelerators that compose the accelerator complex and the flow of different particle beams. The figure also show the location of the four main experiments ALICE, ATLAS, CMS, and LHCb. Figure taken from [75] under Creative Commons license. 54
5.2 Side view of the layout of the LHCb upgraded detector. Figure taken from [82] under Creative Commons license. 55
5.3 Recorded luminosity for pp and PbPb collisions for each years. Figures taken from [83, 84] under Creative Commons license. 56
5.4 Overview of the coordinate system of LHCb, with the definition of the angles η, θ and ϕ. 57
5.5 In the left, a schematic layout of the VELO foils around the interaction point. In the right, the layout of the ASICs around the z axis in the closed VELO configuration. The modules on the Side C are highlighted in purple on both sketches. Figure taken from [82] under Creative Commons license. 58
5.6 Drawing of the four UT silicon planes with indicative dimensions. Different colors designate different types of sensors. Figure taken from [82] under Creative Commons license. 59
5.7 In the left, a schematic of four SciFi layers making up a station behind each other. In the right, the three stations, T1, T2 and T3, behind each other. Figure taken from [82] under Creative Commons license. 61
5.8 An schematic view of RICH1 in the left and RICH2 in the right. Figure taken from [82] under Creative Commons license. 62

5.9 A lateral segmentation of the ECAL in the left and the HCAL in the right. One quarter of the detector front face is shown. Figure taken from [82] under Creative Commons license. 63
5.10 In the left, the side view of the muons stations. In the right, the front view of on muon station. Figure taken from [88] under Creative Commons license. 64
5.11 Scheme of the different track types defined in LHCb tracking system. Figure taken from [90] under Creative Commons license. 65
5.12 Scheme of the Online dataflow. Figure taken from [97] under Creative Commons license. 69

6.1 B^0 mass shift with respect to the PDG value in the different regions of the ECAL measured in pp collisions data in 2024. Figure taken from [105] under Creative Commons license. 71
6.2 Measured π^0 mass as a function of time in regular proton-proton collisions at $\sqrt{s} = 14.5$ TeV. The upper plot corresponds to the 2024 data-taking period, and the vertical blue lines represent when the π^0 calibration was applied. The lower plot corresponds to the 2025 data-taking period, and vertical green lines represent the π^0 calibration. Unpublished figure provided by LHCb collaboration under LHCb policy on the use of LHCb figures in members' theses. 72
6.3 Examples of energy deposition of a photon (top left), a merged π^0 (top right), a hadron (bottom left), and a pile-up cluster (bottom right) from a MC Upgrade simulation sample of $B^0 \to K^{*0}\gamma$. Figure taken from [107] under Creative Commons license. 75
6.4 Invariant mass fit of π^0 candidates in PbPb collisions in the outer region with $E^\gamma \in [4.0, 5.0]$ GeV (a), middle region with $E^\gamma \in [5.0, 6.0]$ GeV (b) and inner region with $E^\gamma \in [11.3, 15.8]$ GeV (c). The vertical dotted green line indicates the nominal π^0 mass 78
6.5 Invariant mass fit of π^0 candidates in OO collisions in the outer region with $E^\gamma \in [4.1, 5.0]$ GeV (a), middle region with $E^\gamma \in [5.0, 6.0]$ GeV (b) and inner region with $E^\gamma \in [11.3, 15.8]$ GeV (c). The vertical dotted green line indicates the nominal π^0 mass 79
6.6 Fit of the π^0 mass integrated over E^γ and all the ECAL regions with a RunNumber within 2024 PbPb (a) and 2025 OO (b) data-taking periods. 80
6.7 Average mass corrections as a function of RunNumber for PbPb 2024 (a) and OO 2025 (b) data taking periods. 81
6.8 Calibration table with mass correction factors in (E^γ, region) bins (a) and the projection of the values presented in this table in the E^γ and Cell ID axes (b) for PbPb collisions. Only bins with mass correction have been taken into account for the projection. 83

6.9 Calibration table with mass correction factors in (E^γ, region) bins (a) and the projection of the values presented in this table in the E^γ and Cell ID axes (b) for OO collisions. Only bins with mass correction have been taken into account for the projection. 84

7.1 Invariant mass spectrum for $M(\mu\mu\gamma)-M(\mu\mu)$ after pre-selection and offline selection in 2025 OO collisions. 88
7.2 Schematic representation of a BDT. Figure taken from [122] under Creative Commons license. 94
7.3 Correlation matrix for all the training variables, χ_c mass, $p_T(\chi_c)$, $\eta(\chi_c)$ and E_{ECAL}. Matrix computed using the Pearson Correlation Coefficient in the training sample (signal + background). 97
7.4 Comparison between the different BDT models depending on their hyper-parameters. Comparison between the ROC curves for each model (a) and comparison of the precision-recall curves for each model (b). 99
7.5 Confusion Matrix for the considered BDT model in this analysis. 100
7.6 Fraction of classified candidates as a function of χ_c mass for the considered BDT model in this analysis. 100
7.7 Log Loss functions as a function of the iterations of the training. In blue, the loss computed for the training sample and in orange, the loss computed for the validation sample. 101
7.8 Distributions of the BDT-predicted probabilities for signal (orange) and background (blue) candidates. From left to right, the distributions are shown for the full sample, the training sample, and the testing sample. . . 102
7.9 Figure of Merits using the significance as metric (green solid line) as a function of the predicted probability. The number of signal candidates (dashed orange line) and the number of background candidates (dashed blue line), both normalized to their integral, are also shown. 103
7.10 Normalized number of candidates filtered by the BDT selected as signal (green) and selected as background (purple) using 2025 OO data sample. From left to right, the invariant mass, transverse momentum, and pseudo-rapidity distributions of χ_c candidates. 104
7.11 Schematic view of the event mixing algorithm. 105
7.12 Multiplicity distribution of the number of tracks in the VELO for 2025 OO data sample. The vertical dotted red lines separate the four multiplicity regions, containing the 25% of the total candidates each one. 106
7.13 Fits to the invariant mass distribution in reconstructed OO collisions MC simulation for χ_{c1} (a) and χ_{c2} (b) signal. 108
7.14 Corrected $\mu\mu\gamma-\mu\mu$ invariant mass spectrum in the 2025 OO data sample. Red points represent data (or mixed events) before calibration and blue points after the calibration. 110

7.15 Fits to the $\mu\mu\gamma - \mu\mu$ invariant-mass distribution in the 2025 OO data sample, using a CB model for the χ_c signal and different parameters for the combinatorial background. 112
7.16 Fits to the $\mu\mu\gamma - \mu\mu$ invariant-mass distribution in the 2025 OO data sample, using a Gaussian model for the χ_c signal and different parameters for the combinatorial background. 113
7.17 Spread of results of $N(\chi_{c1})$, $N(\chi_{c2})$ and $N(\chi_{c2})/N(\chi_{c1})$ in OO collisions. The solid red line represents the weighted mean, the red dashed lines show the RMS, which is taken as the systematic uncertainty. 114
7.18 Fit to the dimuon invariant mass in the 2025 OO data sample. 116

8.1 Comparison of the transverse momentum distribution from sPlot of the daughters of χ_{c1} and χ_{c2} in the 2025 OO data sample. The right-side plots show the ratio between both distributions. 123
8.2 Comparison of the pseudorapidity distribution from sPlot of the daughters of χ_{c1} and χ_{c2} in the 2025 OO data sample. The right-side plots show the ratio between both distributions. 124
8.3 Comparison of the transverse momentum distribution from sPlot of the daughters of χ_c and J/ψ in the 2025 OO data sample. The right-side plots show the ratio between both distributions. 125
8.4 Comparison of the pseudorapidity distribution from sPlot of the daughters of χ_c and J/ψ in the 2025 OO data sample. The right-side plots show the ratio between both distributions. 126
8.5 Comparison between the opposite-sign and like-sign in the $\pi\pi\gamma\gamma$ invariant mass spectrum in the 2025 OO data sample. 131
8.6 R_{mix} normalization factor as a function of the invariant mass (a) and invariant mass distribution after the background subtraction (b) using the like-sign distribution normalized by the R_{mix} factor for the fully reconstructed case in the 2025 OO data sample. 132
8.7 Fit to the invariant mass distribution in reconstructed OO MC simulation data for η^0 and ω signal for fully reconstructed case. 133
8.8 Examples of fits to the $\pi\pi\gamma\gamma$ invariant mass for different bins of transverse momentum, pseudorapidity and multiplicity in the 2025 OO data sample. The (a) and (b) fits are cases of low pseudorapidity, while (c) and (d) are at high pseudorapidity. In addition, (a) and (c) fits are cases of low multiplicity, while (b) and (d) are high multiplicity. 135
8.9 Comparison between the opposite-sign and like-sign in the $\pi\pi\gamma$ invariant mass spectrum in the 2025 OO data sample. 136
8.10 R_{mix} normalization factor as a function of the invariant mass (a) and invariant mass distribution after the background subtraction (b) using the like-sign distribution normalized by the R_{mix} factor for the partially reconstructed case in the 2025 OO data sample. 137

8.11 Fit to the invariant mass distribution in reconstructed OO MC simulation data for η^0 (a) and ω (b) signal for partially reconstructed case. 138
8.12 Examples of fits to the $\pi\pi\gamma$ invariant mass for different bins of transverse momentum, pseudorapidity and multiplicity in the 2025 OO data sample. The (a) and (b) fits are cases of low pseudorapidity, while (c) and (d) are at high pseudorapidity. In addition, (a) and (c) fits are cases of low multiplicity, while (b) and (d) are high multiplicity. 140
8.13 Kinematic projection matrices relating the $p_T^{\pi\pi\gamma}$ and $p_T^{\gamma_{\rm prob}}$ distributions for the two pseudorapidity intervals of the ω meson using a generator-level OO MC simulation. 142
8.14 Kinematic projection matrices computed for each bin of pseudorapidity of η^0 using a generator-level OO MC simulation. 143
8.15 Acceptance times photon detection efficiency as a function of the photon transverse momentum for different bins of pseudorapidity and multiplicity in the 2025 OO data sample. 144
8.16 Photon acceptance factor as a function of the transverse momentum in the low (a) and high(b) pseudorapidity bins from a generator-level OO MC simulation. The results using η^0, ω and both decays at the same time are presented separately. 145
8.17 Photon detection efficiency as a function of the photon transverse momentum for different bins of pseudorapidity and multiplicity in the 2025 OO data sample. 146

9.1 Spread of results of $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ in OO collisions. The solid red line represents the weighted mean, and the red dashed lines show the RMS, which is taken as the systematic uncertainty. 150
9.2 Cross-section ratio, $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ as a function of the rapidity, for the χ_{c1} and χ_{c2} promptly produced in OO collisions. The OO data is compared with the results of the converted sample in pp [109] and with the converted and calorimetric samples in pPb [111]. Green error bars correspond to statistical uncertainty, and the green shaded region corresponds to systematic uncertainty. The rest of the error bars correspond to total uncertainties. . . 151
9.3 Spread of results of $F(\chi_{c1} \to J/\psi)$, $F(\chi_{c2} \to J/\psi)$ and $F(\chi_c \to J/\psi)$ in OO collisions. The solid red line represents the weighted mean, and the red dashed lines show the RMS, which is taken as the systematic uncertainty. 154
9.4 Fraction of J/ψ from χ_c for the different collision systems at the different experiments. The shaded box of the OO point represent the current systematic uncertainty. 155

10.1 Instantaneous (peak) and integrated (recorded) luminosity profiles for the original LHCb (Runs 1 and 2), Upgrade I (Runs 3 and 4) and Upgrade II (Run 5) detectors, in the context of the LHC and HL-LHC schedule. Figure taken from [134] under Creative Commons license. 160

10.2 Identification of photon sources using a BDT trained using Pythia and EPOS generated events (left) and the azimuthal correlation between the photons and hadrons produced in the event (right). Figure taken from [138] under Creative Commons license. 163
10.3 Expected statistical uncertainty on R_{PbPb} of $\Upsilon(1S)$, $\Upsilon(2S)$ and $\Upsilon(3S)$ states in the LHCb Upgrade I with $\mathcal{L}_{\mathrm{int}} = 1$ nb^{-1} and LHCb Upgrade II with $\mathcal{L}_{\mathrm{int}} = 10$ nb^{-1} with PbPb collisions. Figure taken from [138] under Creative Commons license. 164
10.4 Schematic side-view of the Upgrade II baseline detector. Figure taken from [134] under Creative Commons license. 165
10.5 VELO tracking efficiency as a function of η comparing the Baseline and Low scenarios(left) and IP resolution in the Baseline and Low scenarios. Figure taken from [134] under Creative Commons license. 167

11.1 Spatial distribution of simulated hits in the (x, y) plane for a central (0-5%) PbPb collision generated with Angantyr [130]. The distributions are shown for the VELO (a), UP (b), and Mighty Tracker (c) detectors using the Baseline Upgrade II geometry. 171
11.2 Scheme of the UP standalone track reconstruction algorithm. 173
11.3 Expected track reconstruction efficiency in the MT as a function of the collected integrated luminosity, in (left) the fibre region only and (right) the full MT acceptance. For each plot, different values of peak luminosity are indicated by different colors. Figure taken from [134] under Creative Commons license. 175
11.4 Scheme of the upstream matching using already reconstructed VELO and UP tracks. 176
11.5 Scheme of the matching algorithm for downstream and long tracks. 177
11.6 Distribution of the number of VELO hits in *pp* and central PbPb collisions. 178
11.7 Layout of the three different UP geometries considered. The Baseline scenario (a), the Borderless scenario (b) with reduced acceptance and the NoT1 scenario (c) where the first station was removed. Each row corresponds to one of the four UP stations. 180
11.8 UP standalone tracking efficiency as a function of the momentum, pseudorapidity and transverse momentum together with the corresponding distributions of reconstructible tracks. for the Baseline, Borderless and NoT1 UP geometries. The top row show the results for *pp* collisions and the bottom row the results for PbPb collisions. 181
11.9 UP standalone ghost rate as a function of the number of reconstructed hits together with the distributions of the reconstructible (recoble) tracks for the Baseline, Borderless and NoT1 UP geometries. The left column shows the results for *pp* collisions and the right column the results for PbPb collisions. 182

11.10Comparison of the tracking efficiency obtained with the nominal and enlarged UP beam-pipe apertures. for *pp* collisions. From top to bottom, the results of the tracking efficiencies of upstream, downstream and long tracks. 183
11.11Comparison of the tracking efficiency obtained with the nominal and enlarged UP beam-pipe apertures for PbPb collisions. From top to bottom, the results of the tracking efficiencies of upstream, downstream and long tracks. 184
11.12Comparison of the ghost rate as a function of the number of reconstructed hits between the original size and the enlarged UP beam pipe hole for upstream, downstream and long tracks. In the top row, the results for *pp* collisions and in the bottom row the results for PbPb collisions. 185
11.13Comparison of the tracking efficiency between the four different UP pixel sizes for *pp* collisions. From top to bottom, the results of the tracking efficiencies of upstream, downstream and long tracks. 187
11.14Comparison of the tracking efficiency between the four different UP pixel sizes for PbPb collisions. From top to bottom, the results of the tracking efficiencies of upstream, downstream and long tracks. 188
11.15Comparison of the ghost rate as a function of the number of reconstructed hits between the four different UP pixel sizes for upstream, downstream and long tracks. In the top row, the results for *pp* collisions and in the bottom row the results for PbPb collisions. 189
11.16Comparison of the tracking efficiency between *pp* and PbPb collisions. From top to bottom, the results of the tracking efficiencies of upstream, downstream and long tracks. 190
11.17Comparison of the ghost rate as a function of the number of reconstructed hits for upstream, downstream and long tracks between *pp* and PbPb collisions. 190
11.18Scheme of the different possibilities of UP standalone track based on the tracklets that make it up. 193
11.19Evolution of the tracklets during the Run forward step. The current evolution states are presented in black, while the new evolution states to be updated at the end of the iteration are presented in red. 194
11.20Schematic illustration of the tracklet-sequence search during the Run Backward step. The green sequence corresponds to a valid track candidate, whereas the red sequence is rejected because it does not satisfy the requirement of an ordered sequence. 195
11.21Comparison of the UP standalone tracking efficiency when the CAT algorithm uses or not a BDT filter in *pp* collisions. 196
11.22UP CAT standalone ghost rate as a function of the number of reconstructed hits together with the distributions of the reconstructible (recoble) tracks for the cases that include or not a BDT filter in *pp* collisions. 196

11.23Comparison of the UP CAT tracking efficiency for three different scenarios of hit efficiency of the UP pixels in *pp* collisions. 197
11.24UP CAT standalone ghost rate as a function of the number of reconstructed hits together with the distributions of the reconstructible (recoble) tracks for the three different cases of hit efficiency of the UP pixels in *pp* collisions. 198
11.25Comparison of the UP CAT tracking efficiency in *pp* and high centrality PbPb collisions. 199
11.26UP CAT standalone ghost rate as a function of the number of reconstructed hits together with the distributions of the reconstructible (recoble) tracks for *pp* and high centrality PbPb collisions. 199

List of Tables

6.1 π^0 candidate selection criteria for PbPb and OO datasets. 76

7.1 Basic properties of χ_{cJ} and J/ψ mesons. Values taken from Ref. [28] 86
7.2 List of MC simulation datasets for this analysis. 87
7.3 Selection done at DaVinci level according to the MC truth information. . . 89
7.4 Offline χ_c candidates selection. 90
7.5 Variables for BDT training. The variables are divided into three categories: Kinematics, decay geometry and photon ID. 95
7.6 Different models for the BDT depending on the values of their hyperparameters and their AUC value. 98
7.7 List of values considered in the event mixing algorithm. 106
7.8 List of the fit results for the parameter of the CB model for χ_{c1} and χ_{c2}. . 108
7.9 Fit results for the different models of signal and background. The nomenclature of the models follows: Crystal Ball (CB) or Gaussian (GS) signal shape, $N_{\text{comb}}^{\text{max}} \in \{2500, 5000\}$ and $|\Delta PVZ|_{\text{max}} \in \{50, 80\}$ mm. 111
7.10 List of the fit results for the J/ψ fit. 116

8.1 Geometrical acceptance for χ_{c1} and χ_{c2} in OO collisions at $p_T > 1.5$ GeV. . 118
8.2 Acceptance correction due to the photons for χ_{c1}, χ_{c2} and their sum ($\chi_c \equiv \chi_{c1} + \chi_{c2}$) in OO collisions at $p_T > 1.5$ GeV. 119
8.3 Reconstruction and selection efficiencies for χ_{c1}, χ_{c2}, their ratio and their sum ($\chi_c \equiv \chi_{c1} + \chi_{c2}$) in OO collisions. 120
8.4 List of MC simulation datasets for photon efficiency. 129
8.5 ω candidates selection done according to the MC truth information. 130
8.6 Γ parameter result from the fit to the reconstructed truth MC data. 133
8.7 Results from the invariant mass fit of the η^0 and ω yields for the different bins of transverse momentum, pseudorapidity and multiplicity in the fully reconstructed case. 134
8.8 Parameter results from the fit to the partially reconstructed truth MC data. 138
8.9 Results from the invariant mass fit of the η^0 and ω yields for the different bins of transverse momentum, pseudorapidity and multiplicity in the partially reconstructed case. 139
8.10 Integrated polarization correction factors. 148

11.1 Overall tracking efficiency obtained with the standalone UP and CAT algorithms in *pp* and central PbPb collisions. Uncertainties are statistical only. 200
11.2 Overall ghost-track rate obtained with the standalone UP and CAT algorithms in *pp* and central PbPb collisions. Uncertainties are statistical only. 200

Glossary

All acronyms and abbreviations used in this thesis are defined here. Some of these are re-defined also in the text for completeness.

LHCb Large Hadron Collider beauty experiment

QGP Quark-Gluon Plasma

HL-LHC High-Luminosity Large Hadron Collider

LHC Large Hadron Collider

OO Oxygen-Oxygen collisions

PbPb Lead-Lead collisions

pp Proton-proton collisions

UP Upstream Pixel detector

CAT Cellular Automaton Tracking

QCD Quantum Chromodynamics

SM Standard Model

ATLAS A Toroidal LHC Apparatus

CMS Compact Muon Solenoid

CERN Conseil Européen pour la Recherche Nucléaire

PDG Particle Data Group

SPS Super Proton Synchrotron

AGS Alternating Gradient Synchrotron

RHIC Relativistic Heavy Ion Collider

AA Ion-Ion collisions (generic)

pQCD Perturbative Quantum Chromodynamics

PDFs Parton Distribution Functions

CEM Colour Evaporation Model

CSM Colour-Singlet Model

NRQCD Non-Relativistic Quantum Chromodynamics

PDF Parton Distribution Function

CGC Colour Glass Condensate

CEL Coherent Energy Loss

ALICE A Large Ion Collider Experiment

LEP Large Electron-Positron Collider

LINAC2 Linear Accelerator 2

PSB Proton Synchrotron Booster

PS Proton Synchrotron

ECR Electron Cyclotron Resonance

LINAC3 Linear Accelerator 3

LEIR Low Energy Ion Ring

BSM Beyond the Standard Model

CP Charge-Parity (symmetry)

LS2 Long Shutdown 2

PV Primary Vertex

SV Secondary Vertex

VELO Vertex Locator

UT Upstream Tracker

SciFi Scintillating Fibre Tracker

ECAL Electromagnetic Calorimeter

HCAL Hadronic Calorimeter

HLT High-Level Trigger

TT Tracking Turicensis (predecessor of the UT)

PID Particle Identification

RICH Ring-Imaging Cherenkov detector

SPD Scintillating Pad Detector

MWPC Multi-Wire Proportional Chamber

RTA Real-Time Analysis

HLT1 High-Level Trigger stage 1

HLT2 High-Level Trigger stage 2

GEC Global Event Cut

EB Event Builder

MC Monte Carlo

GEANT4 GEometry ANd Tracking toolkit (version 4)

EMG Exponentially Modified Gaussian

RMS Root Mean Square

FD Flight Distance

DIRA Direction Angle (cosine of the angle between the momentum vector and the flight direction)

NDOF Number of Degrees of Freedom

SMOG2 System for Measuring Overlap with Gas (second generation fixed-target system)

IP Impact Parameter

PIDMU Muon particle-identification variable

ML Machine Learning

BDT Boosted Decision Tree

DT Decision Tree

DOCA Distance Of Closest Approach

ROC Receiver Operating Characteristic curve

AUC Area Under the (ROC) Curve

CB Crystal Ball (probability density function)

GS Gaussian function

PIDK Kaon particle-identification variable

PHENIX Pioneering High Energy Nuclear Interaction eXperiment

LS4 Long Shutdown 4

TORCH Time Of internally Reflected CHerenkov light detector

MAPS Monolithic Active Pixel Sensor

MT Mighty Tracker

GPU Graphics Processing Unit

AuAu Gold-Gold collisions

ISR Initial-State Radiation

FSR Final-State Radiation

MPI Multiple Parton Interactions

DPS Double Parton Scattering

UE Underlying Event

CNM Cold Nuclear Matter

TRUEID Monte Carlo Particle Data Group ID of the particle

References

[1] T. Matsui and H. Satz, *J/ψ suppression by quark-gluon plasma formation*, Physics Letters B **178** (1986) 416.

[2] F. Karsch, D. Kharzeev, and H. Satz, *Sequential charmonium dissociation*, Physics Letters B **637** (2006) 75, `arXiv:hep-ph/0512239`.

[3] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Physical Review Letters **13** (1964) 508.

[4] P. W. Higgs, *Broken Symmetries, Massless Particles and Gauge Fields*, Physics Letters **12** (1964) 132.

[5] A. Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Physics Letters B **716** (2012) 1.

[6] C. Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Physics Letters B **716** (2012) 30.

[7] Y. Fukuda and others (Super-Kamiokande Collaboration), *Evidence for Oscillation of Atmospheric Neutrinos*, Phys. Rev. Lett. **81** (1998) 1562.

[8] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321.

[9] Wikimedia Commons contributors, *Standard model of elementary particles*, `https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg`, 2025. Accessed: 2026-07-01.

[10] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Advantages of the color octet gluon picture*, Physics Letters B **47** (1973) 365.

[11] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. **30** (1973) 1343.

[12] M. Gell-Mann, *Symmetries of Baryons and Mesons*, Phys. Rev. **125** (1962) 1067.

[13] S. N. et al. (Particle Data Group), *Review of Particle Physics*, Progress of Theoretical and Experimental Physics **2024** (2024) 083C01, 2025 update available at `https://pdg.lbl.gov/2025/reviews/standard_model_and_related.html`.

[14] G. Martinez Garcia, *Experimental overview on quark-gluon plasma*, `arXiv:1304.1452arXiv:1304.1452`.

[15] A. Li *et al.*, *Neutron star equation of state: Quark mean-field (QMF) modeling and applications*, Journal of High Energy Astrophysics **28** (2020) 19.

[16] K. RAJAGOPAL and F. WILCZEK, in *The Condensed Matter Physics of QCD*, pp. 2061–2151, World Scientific, 2001.

arXiv:https://www.worldscientific.com/doi/pdf/10.1142/9789812810458$_0$043.

[17] L. Wang, R. R. Caldwell, J. P. Ostriker, and P. J. Steinhardt, *Cosmic Concordance and Quintessence*, The Astrophysical Journal **530** (2000) 17.

[18] P. Braun-Munzinger and J. Wambach, *Colloquium: Phase diagram of strongly interacting matter*, Rev. Mod. Phys. **81** (2009) 1031.

[19] M. A. Stephanov, *QCD phase diagram: an overview*, PoS **LAT2006** (2006) 024, arXiv:hep-lat/0701002.

[20] A. Bazavov *et al.*, *Equation of state and QCD transition at finite temperature*, Phys. Rev. D **80** (2009) 014504, arXiv:0903.4379.

[21] F. Karsch and E. Laermann, *Thermodynamics and in-medium hadron properties from lattice QCD*, in *Quark-Gluon Plasma 3* (R. C. Hwa and X.-N. Wang, eds.). World Scientific, 2004. arXiv:hep-lat/0305025.

[22] Y. Aoki *et al.*, *The order of the quantum chromodynamics transition predicted by the standard model of particle physics*, Nature **443** (2006) 675, arXiv:hep-lat/0611014.

[23] J. D. Bjorken, *Highly relativistic nucleus-nucleus collisions: The central rapidity region*, Phys. Rev. D **27** (1983) 140.

[24] J.-Y. Ollitrault, *Relativistic hydrodynamics for heavy-ion collisions*, Eur. J. Phys. **29** (2008) 275, arXiv:0708.2433.

[25] P. Romatschke and U. Romatschke, *Relativistic Fluid Dynamics In and Out of Equilibrium: And Applications to Relativistic Nuclear Collisions*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, 2019.

[26] M. L. Miller, K. Reygers, S. J. Sanders, and P. Steinberg, *Glauber Modeling in High-Energy Nuclear Collisions*, Annual Review of Nuclear and Particle Science **57** (2007) 205, arXiv:nucl-ex/0701025, https://arxiv.org/abs/nucl-ex/0701025.

[27] L. Collaboration *et al.*, *Centrality determination in heavy-ion collisions with the LHCb detector*, arXiv:2111. 01607 [nucl-ex] (2021) arXiv:2111.01607, Also published as JINST **17** (2022) P05009, DOI:10.1088/1748-0221/17/05/P05009.

[28] Particle Data Group, S. Navas *et al.*, *Review of Particle Physics*, Phys. Rev. D **110** (2024) 030001, 2025 update available at https://pdg.lbl.gov.

[29] A. Rothkopf, *Heavy Quarkonium in Extreme Conditions*, Physics Reports **858** (2020) 1, arXiv:1912.02253.

[30] H. Fritzsch, *Producing heavy quark flavors in hadronic collisions: A test of quantum chromodynamics*, Physics Letters B **67** (1977) 217.

[31] F. Halzen, *CVC for gluons and hadroproduction of quark flavors*, Physics Letters B **69** (1977) 105.

[32] R. Baier and R. Rückl, *Hadronic production of J/ψ and Υ: Transverse momentum distributions*, Zeitschrift für Physik C Particles and Fields **19** (1983) 251.

[33] J.-P. Lansberg, *J/ψ, ψ' and Υ production at hadron colliders: A review*, Journal of Physics G: Nuclear and Particle Physics **38** (2011) 124110, arXiv:1107.0292.

[34] W. E. Caswell and G. P. Lepage, *Effective Lagrangians for bound state problems in QED, QCD, and other field theories*, Physics Letters B **167** (1986) 437.

[35] G. T. Bodwin, E. Braaten, and J. Lee, *Comparison of the color-evaporation model and the NRQCD factorization approach in charmonium production*, Physical Review D **72** (2005) 014004, arXiv:hep-ph/0504014.

[36] B. Gong, L.-P. Wan, J.-X. Wang, and H.-F. Zhang, *Polarization for prompt J/ψ and $\psi(2S)$ production at the Tevatron and LHC*, Physical Review Letters **110** (2013) 042002, arXiv:1205.6682.

[37] B. A. Kniehl and G. Kramer, *Color-octet contributions to J/ψ photoproduction via fragmentation at HERA*, Physics Letters B **413** (1997) 416, arXiv:hep-ph/9703280.

[38] CMS Collaboration, *Measurement of the prompt J/ψ and $\psi(2S)$ polarizations in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. B **727** (2013) 381.

[39] LHCb collaboration, R. Aaij *et al.*, *Measurement of J/ψ polarization in pp collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. C **73** (2013) 2631, arXiv:1307.6379.

[40] N. Brambilla *et al.*, *Heavy Quarkonium: Progress, Puzzles, and Opportunities*, Eur. Phys. J. C **71** (2011) 1534, arXiv:1010.5827.

[41] J. L. Albacete *et al.*, *Predictions for cold nuclear matter effects in p+Pb collisions at sNN=8.16 TeV*, Nuclear Physics A **972** (2018) 18.

[42] E. G. Ferreiro, *Quarkonium: a theory overview*, Nuclear Physics A **982** (2019) 127, The 27th International Conference on Ultrarelativistic Nucleus-Nucleus Collisions: Quark Matter 2018.

[43] P. Paakkinen, *Nuclear PDFs at the beginning of LHC Run 3*, in *Proceedings of the 10th Annual Conference on Large Hadron Collider Physics (LHCP2022)*, **LHCP2022** 137, 2023.

[44] K. J. Eskola, *Global analysis of nuclear parton distribution functions*, Nucl. Phys. A **967** (2017) 269, arXiv:1704.05063.

[45] N. Armesto, *Nuclear shadowing*, J. Phys. G **32** (2006) R367, arXiv:hep-ph/0604108.

[46] European Muon Collaboration, *Measurement of the nucleon structure function in nuclei*, Phys. Lett. B **123** (1983) 275.

[47] M. Klasen and H. Paukkunen, *Nuclear Parton Distribution Functions After the First Decade of LHC Data*, Annu. Rev. Nucl. Part. Sci. **74** (2024) 49, arXiv:2311.00450.

[48] F. Gelis, E. Iancu, J. Jalilian-Marian, and R. Venugopalan, *The Color Glass Condensate*, Annual Review of Nuclear and Particle Science **60** (2010) 463.

[49] D. Kharzeev and H. Satz, *Charmonium composition and nuclear suppression*, Phys. Lett. B **366** (1996) 316, arXiv:hep-ph/9508276.

[50] F. m. c. Arleo, S. Peigné, and T. Sami, *Revisiting scaling properties of medium-induced gluon radiation*, Phys. Rev. D **83** (2011) 114036.

[51] F. Arleo, G. Jackson, and S. Peigné, *Impact of fully coherent energy loss on heavy meson production in pA collisions*, JHEP **01** (2022) 164, arXiv:2107.05871.

[52] E. G. Ferreiro, *Charmonium dissociation and recombination at LHC: Revisiting comovers*, Physics Letters B **731** (2014) 57.

[53] CMS Collaboration, *Observation of sequential Υ suppression in PbPb collisions*, Phys. Rev. Lett. **120** (2018) 142301, arXiv:1706.05984.

[54] A. Andronic *et al.*, *Heavy-flavour and quarkonium production in the LHC era: from proton–proton to heavy-ion collisions*, European Physical Journal C **76** (2016) 107, arXiv:1506.03981.

[55] J. Stachel, A. Andronic, P. Braun-Munzinger, and K. Redlich, *Confronting LHC data with the statistical hadronization model*, Journal of Physics: Conference Series **509** (2014) 012019.

[56] L. Grandchamp, R. Rapp, and G. E. Brown, *In-Medium Effects on Charmonium Production in Heavy-Ion Collisions*, Phys. Rev. Lett. **92** (2004) 212301.

[57] R. Sharma and I. Vitev, *High transverse momentum quarkonium production and dissociation in heavy ion collisions*, Phys. Rev. C **87** (2013) 044905.

[58] A. Capella, E. G. Ferreiro, and A. B. Kaidalov, *Nonsaturation of the J/ψ Suppression at Large Transverse Energy in the Comovers Approach*, Phys. Rev. Lett. **85** (2000) 2080.

[59] M. Winn, *Experimental overview: quarkonia as probes for deconfinement*, doi: 10.48550/arXiv.2510.09075 arXiv:2510.09075.

[60] ALICE Collaboration, S. Acharya *et al.*, *Measurements of inclusive J/ψ production at midrapidity and forward rapidity in Pb–Pb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV*, Physics Letters B **849** (2024) 138451, arXiv:2303.13361.

[61] LHCb Collaboration, R. Aaij *et al.*, *J/ψ and D^0 production in $\sqrt{s_{\rm NN}} = 68.5$ GeV PbNe collisions*, European Physical Journal C (2023) arXiv:2211.11652.

[62] LHCb Collaboration, R. Aaij *et al.*, *Fraction of χ_c decays in prompt J/ψ production measured in pPb collisions at $\sqrt{s_{\rm NN}} = 8.16$ TeV*, Physical Review Letters **132** (2024) 102302, arXiv:2311.01562.

[63] LHCb Collaboration, R. Aaij *et al.*, *Prompt and nonprompt $\psi(2S)$ production in pPb collisions at $\sqrt{s_{\rm NN}} = 8.16$ TeV*, Journal of High Energy Physics **04** (2024) 111, arXiv:2401.11342.

[64] J. F. Grosse-Oetringhaus and U. A. Wiedemann, *A Decade of Collectivity in Small Systems*, arXiv:2407.07484arXiv:2407.07484.

[65] C. Collaboration, *Observation of the charged-particle multiplicity dependence of $\sigma_{\psi(2S)}/\sigma_{J/\psi}$ in pPb collisions at 8.16 TeV*, Physical Review Letters **135** (2025) 092301, arXiv:2503.02139.

[66] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the $\psi(2S)$ to J/ψ cross-section ratio as a function of centrality in PbPb collisions at $\sqrt{s_{\rm NN}} = 5.02$ TeV*, Journal of High Energy Physics **07** (2025) 235, arXiv:2411.05669.

[67] ALICE Collaboration, *$\psi(2S)$ Suppression in Pb-Pb Collisions at the LHC*, Physical Review Letters **132** (2024) 042301.

[68] J.-P. Lansberg, *New observables in inclusive production of quarkonia*, Phys. Rep. **889** (2020) 1, arXiv:1903.09185.

[69] P. Faccioli, C. Lourenço, J. Seixas, and H. K. Wöhri, *Study of ψ' and χ_c decays as feed-down sources of J/ψ hadro-production*, Phys. Rev. D **79** (2009) 074001, arXiv:0809.2153.

[70] S. Digal, P. Petreczky, and H. Satz, *Quarkonium feed-down and sequential suppression*, Phys. Rev. D **64** (2001) 094015, arXiv:hep-ph/0106017.

[71] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the ratio of prompt χ_c to J/ψ production in pp collisions at $\sqrt{s} = 7$ TeV*, Physics Letters B **718** (2012) 431, arXiv:1204.1462.

[72] ATLAS Collaboration, G. Aad *et al.*, *Measurement of χ_{c1} and χ_{c2} production with $\sqrt{s} = 7$ TeV pp collisions at ATLAS*, Journal of High Energy Physics **07** (2014) 154, arXiv:1404.7035.

[73] CMS Collaboration, A. M. Sirunyan *et al.*, *Constraints on the χ_{c1} versus χ_{c2} polarizations in proton-proton collisions at $\sqrt{s} = 8$ TeV*, Physical Review Letters **124** (2020) 162002, arXiv:1912.07706.

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

[75] E. Lopienska, *The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022*, , `https://cds.cern.ch/record/2800984`.

[76] A. Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[77] C. Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.

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

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

[80] LHCb Collaboration, *Observation of charge–parity symmetry breaking in baryon decays*, Nature **643** (2025) 1223.

[81] DELPHI Collaboration, P. Abreu, W. Adam *et al.*, *The DELPHI detector at LEP*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **303** (1991) 233.

[82] R. e. a. Aaij, *The LHCb Upgrade I*, Journal of Instrumentation **19** (2024) P05065.

[83] CERN, *LHCb online luminosity and fill summary (2025 operation test, square layout)*, 2025.

[84] CERN, *LHCb online luminosity and fill summary (2025 operation test)*, 2025.

[85] LHCb Collaboration, *LHCb magnet: Technical Design Report*, CERN-LHCC-2000-007, LHCb-TDR-1, CERN, Geneva, 2000. doi: 10.17181/CERN.MCVC.Q12G.

[86] LHCb Collaboration, *LHCb VELO Upgrade Technical Design Report*, CERN-LHCC-2013-021, LHCb-TDR-013, CERN, Geneva, 2013. doi: 10.17181/CERN.4DGI.MZN4.

[87] LHCb Collaboration, *LHCb Tracker Upgrade Technical Design Report*, CERN-LHCC-2014-001, LHCB-TDR-015, CERN, Geneva, 2014. doi: 10.17181/CERN.OOOE.9O9T.

[88] LHCb Collaboration, *LHCb PID Upgrade Technical Design Report*, CERN-LHCC-2013-022, LHCB-TDR-014, CERN, Geneva, Switzerland, 2013. doi: 10.17181/CERN.BD4G.I3VP.

[89] S. Amato *et al.*, *LHCb calorimeters: Technical Design Report*, CERN-LHCC-2000-036, LHCb-TDR-002, CERN, Geneva, Switzerland, 2000. Technical Design Report of the LHCb experiment.

[90] S. Aiola *et al.*, *Hybrid seeding: A standalone track reconstruction algorithm for scintillating fibre tracker at LHCb*, Computer Physics Communications **260** (2021) 107713.

[91] R. Aaij *et al.*, *Allen: A High-Level Trigger on GPUs for LHCb*, Computing and Software for Big Science **4** (2020) 7.

[92] LHCb Collaboration, *Run3 Primary Vertex reconstruction description and performance*, LHCB-FIGURE-2020-005, CERN, Geneva, Switzerland, 2020.

[93] A. Gunther, *Track Reconstruction Development and Commissioning for LHCb's Run 3 Real-time Analysis Trigger*, PhD thesis, Heidelberg University, Geneva, Switzerland, 2023.

[94] P. Billoir, M. D. Cian, P. A. Günther, and S. Stemmle, *A parametrized Kalman filter for fast track fitting at LHCb*, Computer Physics Communications **265** (2021) 108026.

[95] L. Anderlini *et al.*, *Muon identification for LHCb Run 3*, Journal of Instrumentation **15** (2020) T12005.

[96] N. V. Canudas, M. C. Gómez, X. Vilasís-Cardona, and E. G. Ribé, *Graph Clustering: a graph-based clustering algorithm for the electromagnetic calorimeter in LHCb*, European Physical Journal C **83** (2023) 179, `arXiv:2212.11061`.

[97] LHCb Collaboration, *RTA and DPA dataflow diagrams for Run 1, Run 2, and the upgraded LHCb detector*, LHCB-FIGURE-2020-016, CERN, Geneva, Switzerland, 2020.

[98] M. Clemencic *et al.*, *The lhcb simulation application, gauss: Design, evolution and experience*, Journal of Physics: Conference Series **331** (2011) 032023.

[99] T. Sjöstrand, S. Mrenna, and P. Skands, *A brief introduction to PYTHIA 8.1*, Computer Physics Communications **178** (2008) 852.

[100] T. Pierog *et al.*, *EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider*, Phys. Rev. C **92** (2015) 034906.

[101] D. J. Lange, *The EvtGen particle decay simulation package*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **462** (2001) 152, BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.

[102] A. J. Buras *et al.*, *FCNC Processes: Theory and Implications for New Physics*, arXiv e-prints (2005) `arXiv:hep-ph/0506026`.

[103] S. Agostinelli *et al.*, *Geant4—a simulation toolkit*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506** (2003) 250.

[104] LHCb Collaboration, *Boole: the LHCb digitization program*, `https://lhcbdoc.web.cern.ch/lhcbdoc/boole/`, 2006. Software documentation; accessed 2026-02-25.

[105] M. Calvo, *Photon PID in Run 3*, Presentation at the LHCb Collaboration Meeting, 2025. Presented on 13 February 2025.

[106] C. Abellán Beteta *et al.*, *Calibration and performance of the LHCb calorimeters in Run 1 and 2 at the LHC*, arXiv preprint arXiv:2008. 11556 (2020) `arXiv:2008.11556`.

[107] M. Calvo, LHCb Internal Note *Photon identification in Run 3*, LHCb-INT-2025-004, LHCb, 2025.

[108] E. Grushka, *Characterization of Exponentially Modified Gaussian Peaks in Chromatography*, Anal. Chem. **44** (1972) 1733.

[109] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the relative rate of prompt χ_{c0}, χ_{c1} and χ_{c2} production at $\sqrt{s} = 7$ TeV*, doi: 10.1007/JHEP10(2013)115.

[110] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the ratio of prompt χ_c to J/ψ production in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. B **718** (2012) 431, `arXiv:1204.1462`.

[111] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the prompt-production cross-section ratio $\sigma(\chi_{c2})/\sigma(\chi_{c1})$ in pPb collisions at $\sqrt{s_{NN}} = 8.16$ TeV*, Physical Review C **103** (2021) 064905.

[112] *Davinci analysis software documentation*, .

[113] A. Radovic *et al.*, *Machine learning at the energy and intensity frontiers of particle physics*, Nature **560** (2018) 41.

[114] D. Guest, K. Cranmer, and D. Whiteson, *Deep learning and its application to LHC physics*, Annual Review of Nuclear and Particle Science **68** (2018) 161.

[115] K. Albertsson *et al.*, *Machine learning in high energy physics community white paper*, Journal of Physics: Conference Series **1085** (2018) 022008.

[116] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016. `http://www.deeplearningbook.org`.

[117] F. Pedregosa *et al.*, *Scikit-learn: Machine Learning in Python*, Journal of Machine Learning Research **12** (2011) 2825.

[118] scikit-learn developers, *User Guide — scikit-learn 1.8.0 documentation*, `https://scikit-learn.org/stable/user_guide.html`, 2025. Accessed: 2026-05-11.

[119] L. Breiman, *Random Forests*, Machine Learning **45** (2001) 5.

[120] S. B. Kotsiantis, *Decision trees: a recent overview*, Artificial Intelligence Review **39** (2013) 261.

[121] J. Fürnkranz, *Decision Tree*, in *Encyclopedia of Machine Learning and Data Mining* (C. Sammut and G. I. Webb, eds.), pp. 330–335. Springer, Boston, MA, 2017. doi: 10.1007/978-1-4899-7687-$1_6$6.

[122] XGBoost developers, *XGBoost documentation*, `https://xgboost.readthedocs.io/en/release_3.2.0/index.html`, 2026. Accessed: 2026-05-11.

[123] T. Chen and C. Guestrin, *XGBoost: A Scalable Tree Boosting System*, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (New York, NY, USA), 785–794, ACM, 2016.

[124] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*, Chapman and Hall/CRC, Boca Raton, FL, 1 ed., 1984.

[125] D. Drijard, H. G. Fischer, and T. Nakada, *Study of event mixing and its application to the extraction of resonance signals*, Nuclear Instruments and Methods in Physics Research **225** (1984) 367.

[126] D. L'Hôte, *About resonance signal extraction from multiparticle data: combinatorics and event mixing methods*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **337** (1994) 544.

[127] T. Skwarnicki, *A Study of the Radiative Cascade Transitions Between the Upsilon-Prime and Upsilon Resonances*, PhD thesis, Institute of Nuclear Physics, Krakow, 1986.

[128] LHCb Collaboration, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, `arXiv:1408.1251`.

[129] M. Pivk and F. R. Le Diberder, *sPlot: A statistical tool to unfold data distributions*, Nucl. Instrum. Meth. A **555** (2005) 356, `arXiv:physics/0402083`.

[130] C. Bierlich, G. Gustafson, L. Lonnblad, and H. Shah, *The Angantyr model for heavy-ion collisions in Pythia8*, JHEP **10** (2018) 134, `arXiv:1806.10820`.

[131] LHCb Collaboration, R. Aaij *et al.*, *Measurement of the ratio of prompt χ_c to J/ψ production in pp collisions at $\sqrt{s} = 7$ TeV*, Physics Letters B **718** (2012) 431.

[132] HERA-B Collaboration, I. Abt *et al.*, *Production of the Charmonium States χ_{c1} and χ_{c2} in Proton-Nucleus Interactions at $\sqrt{s} = 41.6$ GeV*, Physical Review D **79** (2009) 012001, arXiv:0807.2167.

[133] PHENIX Collaboration, A. Adare *et al.*, *Cold-Nuclear-Matter Effects on Heavy-Quark Production in d + Au Collisions at $\sqrt{s_{NN}} = 200$ GeV*, Physical Review Letters **111** (2013) 202301, arXiv:1305.5516.

[134] LHCb Collaboration, *LHCb Upgrade II Scoping Document*, CERN-LHCC-2024-010, LHCB-TDR-026, CERN, 2024. doi: 10.17181/CERN.2RXP.HDK0.

[135] European Strategy Group, *Deliberation Document on the 2020 Update of the European Strategy for Particle Physics*, Geneva, 2020.

[136] European Strategy Group, *2020 Update of the European Strategy for Particle Physics*, Geneva, 2020.

[137] *European Strategy for Particle Physics: 2026 Update – Recommendations*, CDS 2950671, CERN Council, 2026. CERN Document Server.

[138] LHCb Collaboration, LHCb Collaboration, *Opportunities for Heavy-Ion and Fixed-Target Physics with LHCb Upgrade II*, CERN-LHCb-PUB-2024-007, LHCb-PUB-2024-007, CERN, 2025.

[139] J. von Neumann, *Theory of Self-Reproducing Automata*, University of Illinois Press, Urbana, IL, 1966.

[140] S. Wolfram, *Cellular Automata as Models of Complexity*, Nature **311** (1984) 419.

[141] M. Gardner, *Mathematical Games: The Fantastic Combinations of John Conway's New Solitaire Game "Life"*, Scientific American **223** (1970) 120.

[142] I. Abt, D. Emeliyanov, I. Kisel, and S. Masciocchi, *CATS: A cellular automaton for tracking in silicon for the HERA-B vertex detector*, Nucl. Instrum. Meth. A **489** (2002) 389.

[143] D. Emeliyanov *et al.*, *Cellular automaton and Kalman filter based track search in the HERA-B pattern tracker*, Nucl. Instrum. Meth. A **490** (2002) 546.

[144] D. Funke *et al.*, *Parallel track reconstruction in CMS using the cellular automaton approach*, J. Phys. Conf. Ser. **513** (2014) 052010.

[145] D. Rohr, S. Gorbunov, M. O. Schmidt, and R. Shahoyan, *GPU-based Online Track Reconstruction for the ALICE TPC in Run 3 with Continuous Read-Out*, arXiv:1905.05515arXiv:1905.05515.

[146] R. E. Kalman, *A New Approach to Linear Filtering and Prediction Problems*, Transactions of the ASME–Journal of Basic Engineering **82** (1960) 35.

[147] R. Frühwirth, *Application of Kalman Filtering to Track and Vertex Fitting*, Nucl. Instrum. Meth. A **262** (1987) 444.





This thesis presents the first measurement of $\chi_c(1P)$ production in ion–ion collisions at collider energies, using OO data from LHCb. Color screening suppresses weakly bound χ_c states more than the J/ψ, so their relative production probes the sequential suppression pattern of the QGP. The χ_{c2}/χ_{c1} ratio and the J/ψ feed-down fraction are extracted, although further systematic studies are needed. An important achievement is demonstrating that precision χ_c measurements are feasible in a heavy-ion environment at collider energies. The thesis also presents Upstream Pixel tracking studies for LHCb Upgrade II in simulated pp and Pb–Pb collisions, informing the current detector design. A novel Cellular Automaton Tracking algorithm achieves standalone efficiencies above 90% even at the highest occupancies.