

# Thèse de doctorat

NNT : 2021IPPA130



INSTITUT  
POLYTECHNIQUE  
DE PARIS



## The $B_c^+$ meson in heavy-ion collisions with the CMS detector.

Thèse de doctorat de l'Institut Polytechnique de Paris  
préparée à l'École polytechnique

École doctorale n°626 École doctorale de l'Institut Polytechnique de Paris (EDIPP)  
Spécialité de doctorat : Physique des particules

Thèse présentée et soutenue à Palaiseau, le 15/12/2021, par

**GUILLAUME FALMAGNE**

Composition du Jury :

Giulia Manca Professeure associée, Universita di Cagliari and INFN (Cagliari)	Rapporteuse
Andrea Dainese Directeur de recherche, INFN (Padova)	Rapporteur
Pascal Paganini Directeur de recherche, CNRS, École polytechnique (LLR)	Président
Pol-Bernard Gossiaux Professeur, IMT Atlantique (Subatech)	Examinateur
Bruno Espagnon Professeur, Université Paris-Saclay (IJCLab)	Examinateur
Leticia Cunqueiro Professeure assistante, École polytechnique, (LLR)	Examinaterice
Raphaël Granier de Cassagnac Directeur de recherche, CNRS, École polytechnique (LLR)	Directeur de thèse
François Arleo Chargé de recherche, CNRS, École polytechnique (LLR)	Co-directeur de thèse



*“Imagination is more important than knowledge.  
Knowledge is limited.  
Imagination encircles the world.”*

---

Albert EINSTEIN<sup>1</sup>

---

<sup>1</sup>The Saturday Evening Post, October 26, 1929, interview by George S. Viereck, p. 117



# Declaration of authorship

I, Guillaume FALMAGNE, declare that this thesis titled “The  $B_c^+$  meson in heavy-ion collisions with the CMS detector.” and the work exposed in it are my own. In particular:

- The shown work was performed while in candidature for my PhD degree in École polytechnique, and for a small part during the Masters’ internship that preceded.
- When I have consulted or quoted the published work of others, I clearly reference their authors.
- The structure and authorship procedures of large collaborations in High Energy Physics are quite specific to this field. A significant part of the work presented here could not exist without the previous work of the thousands of authors in the CMS Collaboration. In the part of this thesis treating CMS data, I intend to show only my contributions to the endeavours of the collaboration. However, in a few places, the limits of the contributions from different people of the collaboration might be blurry or long to explain, or their work might concern the ground tools that I used, justifying in those cases to keep implicit the exact distribution of the works.
- Considering the previous caveat, I have acknowledged the primary sources of help. I have tried as much as possible to clearly state what my contributions were while keeping a reasonable impact on clarity and brevity.





Laboratoire Leprince-Ringuet

ÉCOLE POLYTECHNIQUE

Institut Polytechnique de Paris

## *Abstract*

Doctoral thesis in Physics

---

### **The $B_c^+$ meson in heavy-ion collisions with the CMS detector.**

---

by Guillaume FALMAGNE

This thesis addresses how heavy-quark hadronisation and high-energy partons are affected by the quark-gluon plasma (QGP), a hot and dense medium created in lead-lead (PbPb) collisions at the LHC. Data from the CMS detector are analysed to achieve the first observation of  $B_c^+$  mesons in heavy-ion collisions. Building on an existing scaling law from a model of radiative energy loss in the QGP, a second scaling law is also brought to light.

The analysis of CMS data from 2017 proton-proton and 2018 PbPb collisions at a centre-of-mass energy of 5.02 TeV per nucleon pair leads to the observation of  $B_c^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) \mu^+ \nu_\mu$  decays, and to the measurement of the  $B_c^+$  nuclear modification factor in two bins of the trimuon transverse momentum ( $p_T^{\mu\mu\mu}$ ) and of the PbPb collision centrality. It is presented in this thesis and summarised in Ref. [1]. Three main backgrounds are described either with simulation or with specifically-designed data-driven samples. A boosted decision tree (BDT) is trained on the selected background and simulated signal candidates. A likelihood fit is run on signal and background templates, binned in BDT, trimuon invariant mass, and  $p_T^{\mu\mu\mu}$  or centrality. The acceptance and efficiency of the selection chain are evaluated iteratively in each  $p_T^{\mu\mu\mu}$  or centrality bin with the simulated signal, whose  $p_T^{\mu\mu\mu}$  spectrum is first corrected with the one measured in a preliminary analysis. The  $B_c^+$  meson is found to be less suppressed than all measured open and hidden heavy flavour mesons, except the  $B_s^0$  meson. The results also hint at a softening of the  $p_T$  spectrum in PbPb collisions. These may indicate that heavy-quark recombination is a significant  $B_c^+$  production mechanism. As an amuse-gueule for this analysis, Fig. 1 shows a PbPb collision in CMS that has a very high probability of containing a  $B_c$  meson.

At high  $p_T$  ( $\gtrsim 10 - 15$  GeV), radiative energy loss should be the dominant source of suppression of hadrons in the QGP. An existing model for the radiative energy loss of partons, based on the BDMPS medium-induced gluon spectrum, predicts a universal  $p_T$ -dependence of the nuclear modification factor [2]. This fits collected measurements in systems of various geometric configurations and energies, from which the corresponding mean energy losses are determined [3]. A new scaling law is found consistent with the gathered measurements: it links the extracted mean energy loss with the average path length in the medium and the charged particle multiplicity. This can lead to an extraction of medium expansion and diffusion properties, and to a prediction of the azimuthal asymmetry coefficient  $v_2$  at high  $p_T$ .

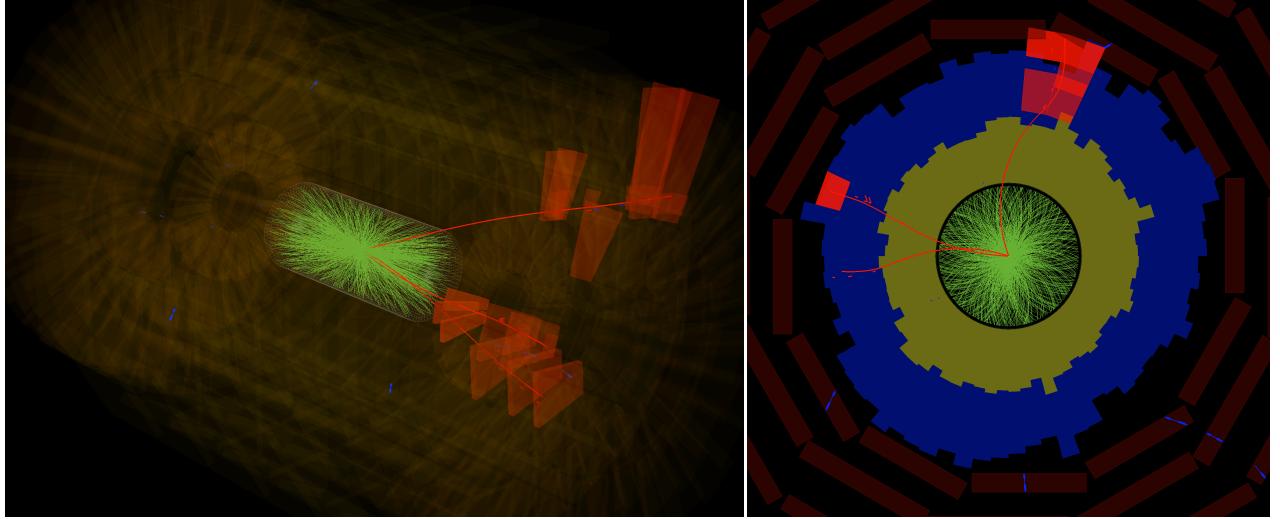


FIGURE 1: Display of the PbPb 2018 collision event #209030061 (run #327004, LS #497) in the CMS detector. This PbPb collision (at a 5% centrality) has a very high probability to contain a  $B_c$  meson decaying to three muons, shown in red. This is the selected PbPb event with the fourth-highest BDT value, whereas less than 0.1 background events are expected at or above this BDT value. The 3D (*left*) and transverse to the beam (*right*) views are shown. The trimuon from this  $B_c$  decay has  $(m^{\mu\mu\mu}, p_T^{\mu\mu\mu}, y^{\mu\mu\mu}) = (5.5 \text{ GeV}, 9.8 \text{ GeV}, 1.7)$ ; the high trimuon mass means the energy taken by the neutrino of the decay is small. One opposite-sign dimuon (that with largest opening angle) has a mass of 3.07 GeV, close to the  $J/\psi$  mass. The trimuon vertex is displaced by about 1 mm from the primary vertex, to which most charged tracks (shown in green when  $p_T > 1.2 \text{ GeV}$ ) point. The yellow and blue histograms show the transverse energy deposited in the calorimeters. The S-shape trajectory of the muons, inspiring the CMS logo, is due to the magnetic field changing sign in the return yoke.

---

This thesis is structured with an introduction (chapter 1) travelling from general physics considerations to concepts needed in the following, then a first part concerning the  $B_c$  analysis, and a second part concerning my work on the phenomenology of energy loss. Chapter 2 motivates the experimental search for  $B_c$  mesons in heavy ion collisions. Chapter 3 presents the LHC and the CMS detector, and the overall analysis strategy (section 3.3). Chapter 4 deals with the backgrounds blurring the  $B_c$  signal, and chapter 5 details the candidate selection established to discriminate against them. Chapter 6 shows the extraction of the signal yields from a template fit of the trimuon mass distributions. Chapter 7 computes the correction of the observed yields for the acceptance and efficiency of the signal reconstruction and selection. Chapter 8 summarises the sources of uncertainties and how they are estimated. Chapter 9 shows the results and their interpretation, including the  $B_c$  nuclear modification factor. In the phenomenology part, chapter 10 describes the energy loss model I start from, and chapter 11 shows two scaling laws it results in. Chapter 12 shows some improvements and additions to the model. Lists of abbreviations, acronyms, and some numerical values, can be found at the end of the document.



LM



Laboratoire Leprince-Ringuet

ÉCOLE POLYTECHNIQUE

Institut Polytechnique de Paris

## *Résumé*

Thèse de doctorat en Physique

---

**Le méson  $B_c^+$  en collisions d'ions lourds avec le détecteur CMS.**

---

by Guillaume FALMAGNE

Cette thèse traite de l'hadronisation des quarks lourds et des partons de haute énergie, et de leur modification par le plasma de quarks et gluons (QGP), un milieu chaud et dense créé dans les collisions plomb-plomb (PbPb) au LHC. La première observation de mésons  $B_c^+$  en collisions d'ions lourds est menée en analysant des données du détecteur CMS. Partant d'une loi d'échelle existante dans un modèle de pertes d'énergie radiatives dans le QGP, une seconde loi d'échelle est mise en évidence.

L'analyse de données CMS en collisions proton-proton (2017) et PbPb (2018) à une énergie au centre de masse de 5.02 TeV par paire de nucléons mène à l'observation de désintégrations  $B_c^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)\mu^+\nu_\mu$ , et à la mesure du facteur de modification nucléaire du méson  $B_c^+$  en deux intervalles d'impulsion transverse du trimuon ( $p_T^{\mu\mu\mu}$ ) ou de la centralité de la collision PbPb. L'analyse est présentée dans cette thèse et résumée en Ref. [1]. Trois bruits de fond principaux sont décrits, soit par simulation, soit par des méthodes fondées sur les données. Une de ces méthodes est spécifique à cette analyse, et décrit les  $J/\psi$  combinés avec des muons ne venant pas du même processus: elle pivote le dimuon  $J/\psi$  par différents angles avant de l'associer avec un troisième muon du même événement. Un arbre de décision amélioré (*boosted decision tree*, BDT) est entraîné sur les candidats sélectionnés du signal et des bruits de fond. Un ajustement d'histogrammes en probabilités (*likelihood template fit*) est opéré pour ajuster sur les données les modèles de signal et de bruits de fonds, séparés en intervalles de BDT, de masse invariante du trimuon, et de  $p_T^{\mu\mu\mu}$  ou de centralité. Des paramètres de nuisance dans le fit prennent en compte les incertitudes sur les formes des bruits de fond.

L'acceptance et l'efficacité du déclenchement, de la reconstruction et de la sélection, sont évaluées de manière itérative, dans chaque bin de  $p_T^{\mu\mu\mu}$  ou de centralité, avec le signal simulé dont le spectre en  $p_T^{\mu\mu\mu}$  est préalablement corrigé par le spectre mesuré dans une analyse préliminaire. L'efficacité des muons dans la simulation est corrigée par une méthode de tag-and-probe utilisant la résonance en masse du  $J/\psi$ , dont l'étude en collisions proton-proton est présentée en détail. L'incertitude sur la correction d'acceptance et efficacité appliquée au signal observé propage les autres sources d'incertitude à la correction du spectre en  $p_T^{\mu\mu\mu}$  de la simulation. Ce dernier est en effet varié dans les incertitudes de la mesure de l'analyse préliminaire, résultant en autant de variations de la correction d'acceptance et d'efficacité, qui représentent alors l'incertitude associée à cette correction.

Moins de suppression est observée pour le méson  $B_c^+$  que pour les autres saveurs lourdes ouvertes et fermées, excepté le méson  $B_s^0$ . Un probable adoucissement du spectre en  $p_T$  est mis en évidence. Ces résultats pourraient indiquer que la recombinaison de quarks lourds contribue significativement à la production de mésons  $B_c^+$  dans le QGP.

À haut  $p_T$  ( $\gtrsim 10 - 15$  GeV), la perte d'énergie radiative devrait dominer la suppression des hadrons dans le QGP. Un modèle de perte d'énergie radiative partonique existant, fondé sur le spectre gluonique BDMPS induit par le milieu, prédit une dépendance universelle en  $p_T$  du facteur de modification nucléaire [2]. Cette forme est confirmée par des mesures dans des systèmes de différentes géométries et énergies; elles sont collectées et ajustées à ce modèle, pour extraire la perte d'énergie moyenne correspondante [3]. Cette dernière est incluse dans une nouvelle loi d'échelle, qui la décrit comme dépendant uniquement de la taille et la géométrie du milieu et de la multiplicité en particules chargées. Différentes approches de la longueur de parcours du parton dans le milieu sont comparées. L'ajustement à des données de suppression hadronique en collisions plomb-plomb et xenon-xenon permet d'extraire des propriétés d'expansion et de diffusion du milieu. Le coefficient d'asymétrie azimuthale  $v_2$  à haut  $p_T$  pourrait aussi être prédit.

# Contents

<b>Declaration of authorship</b>	<b>v</b>
<b>Abstract/Résumé</b>	<b>vii</b>
<b>Contents</b>	<b>xi</b>
<b>Editorial choices</b>	<b>xv</b>
<b>Preamble: Why do we seek knowledge?</b>	<b>1</b>
<b>1 Where do we start? From apples to soup</b>	<b>5</b>
1.1 What interactions at what scale? . . . . .	5
1.2 The Standard Model . . . . .	8
1.3 The strong force . . . . .	10
1.3.1 Quantum Chromodynamics . . . . .	10
1.3.2 Parton distribution functions . . . . .	12
1.4 The quark-gluon plasma . . . . .	14
1.4.1 A state of matter . . . . .	14
1.4.2 Categories of experimental signatures . . . . .	15
1.5 Definitions and notations . . . . .	16
1.5.1 Detector and accelerator notions . . . . .	16
1.5.2 Coordinates . . . . .	17
1.5.3 Luminosity and cross sections . . . . .	18
1.5.4 Lifetime . . . . .	19
1.5.5 Types of pp or PbPb collisions . . . . .	20
1.5.6 The geometry and observables of heavy ion collisions . . . . .	21
<b>Part I The <math>B_c^+</math> meson in PbPb and pp collisions</b>	<b>25</b>
<b>2 Why? Heavy quarks in the QGP</b>	<b>27</b>
2.1 Heavy quarks and quarkonia . . . . .	27
2.2 $B_c$ meson properties and past measurements . . . . .	29
2.2.1 Production and spectroscopy . . . . .	29
2.2.2 Past measurements . . . . .	31
2.3 Energy loss . . . . .	32
2.4 Colour screening . . . . .	35
2.5 Recombination and predictions of the $B_c$ modification . . . . .	37
<b>3 How? Analysis strategy with CMS</b>	<b>41</b>
3.1 CERN and the LHC . . . . .	41
3.1.1 Brief history of CERN facilities and discoveries . . . . .	42
3.1.2 The Large Hadron Collider . . . . .	45

3.1.3	Main LHC experiments and results . . . . .	51
3.2	The Compact Muon Solenoid . . . . .	57
3.2.1	Global view . . . . .	57
3.2.2	Inner tracker and material budget . . . . .	61
3.2.3	Muon system . . . . .	63
3.2.4	Trigger . . . . .	65
3.3	The $B_c$ analysis strategy . . . . .	66
3.3.1	Global approach . . . . .	66
3.3.2	Hadronic channel feasibility . . . . .	68
3.3.3	Leptonic channel analysis workflow . . . . .	71
3.4	PbPb data blinding . . . . .	73
3.5	Samples and normalisations . . . . .	74
3.5.1	Data samples . . . . .	75
3.5.2	MC generation . . . . .	75
3.5.3	The ONIATREE analyser . . . . .	77
3.5.4	A priori MC normalisation and weighting . . . . .	77
3.5.5	Normalisation of PbPb yields . . . . .	80
<b>4</b>	<b>Whom do we fight? Backgrounds</b>	<b>83</b>
4.1	Categorisation . . . . .	83
4.2	Fake $J/\psi$ . . . . .	85
4.3	True $J/\psi$ and muon from the same $B$ decay . . . . .	88
4.3.1	$B$ decays . . . . .	88
4.3.2	Muon misidentification . . . . .	89
4.3.3	Other $B_c$ decays . . . . .	90
4.4	True $J/\psi$ and muon from different processes . . . . .	92
4.4.1	Combinatorial background from simulation . . . . .	93
4.4.2	Investigation of the $J/\psi$ MC in pp 2017 . . . . .	93
4.4.3	$J/\psi$ rotation method . . . . .	100
<b>5</b>	<b>What do we seek? Selection</b>	<b>105</b>
5.1	Selection overview . . . . .	105
5.2	Event-level and primary vertex selection . . . . .	106
5.2.1	Event-level selection . . . . .	106
5.2.2	Main primary vertex . . . . .	107
5.3	Single muon acceptance and selection . . . . .	108
5.3.1	Muon selection . . . . .	108
5.3.2	Muon acceptance . . . . .	108
5.3.3	Strategy for three muons . . . . .	112
5.4	Preselection . . . . .	113
5.5	Choice of the $J/\psi$ . . . . .	118
5.6	Analysis binning . . . . .	120
5.7	Boosted Decision Tree (BDT) . . . . .	122
5.7.1	Principle . . . . .	122
5.7.2	Discriminant variables after preselection . . . . .	123
5.7.3	BDT training and testing . . . . .	125
5.7.4	Check and correction of BDT distribution . . . . .	131
5.7.5	Decorrelate BDT from mass . . . . .	131
<b>6</b>	<b>How many do we see? Template fit</b>	<b>135</b>
6.1	Principle . . . . .	135
6.1.1	Likelihood fit with <code>combine</code> . . . . .	135

6.1.2	Structure, input and binning . . . . .	136
6.2	In-fit systematics (nuisance parameters) . . . . .	137
6.3	Fit results . . . . .	140
6.3.1	Post-fit mass distributions and signal yields . . . . .	140
6.3.2	Goodness of fit . . . . .	144
6.3.3	Observed significance . . . . .	144
6.3.4	Toy study for fit uncertainties . . . . .	146
6.3.5	Fit checks . . . . .	147
6.4	Fit method variations . . . . .	150
6.4.1	Mass and BDT binning . . . . .	150
6.4.2	BDT decorrelated from mass . . . . .	150
6.4.3	Statistical uncertainties on templates, shape regularisation . . . . .	151
6.4.4	Summary and calculation of systematic uncertainty . . . . .	154
<b>7</b>	<b>How much is produced? Acceptance and efficiency</b>	<b>159</b>
7.1	Tag-and-probe corrections to single-muon MC efficiencies . . . . .	160
7.1.1	Tag-and-probe principle . . . . .	160
7.1.2	Overview in pp and PbPb . . . . .	161
7.1.3	Single-muon scale factors for pp 2017 . . . . .	161
7.1.4	Scale factors for the $B_c$ analysis . . . . .	173
7.1.5	Tentative soft ID scale factors in PbPb . . . . .	174
7.2	Applying efficiency corrections to trimuons . . . . .	175
7.2.1	Applying scale factors on a MC trimuon . . . . .	175
7.2.2	Scale factors for trimuons with asymmetric muon cuts . . . . .	176
7.3	One-binned acceptance and efficiency corrections . . . . .	184
7.4	Acceptance and efficiency maps . . . . .	186
7.5	Feasibility of event-by-event corrections . . . . .	187
7.6	Two-steps procedure . . . . .	193
7.6.1	Lafferty-Wyatt: how to place the abscissa . . . . .	195
7.6.2	‘Fitting’ two points . . . . .	195
7.6.3	Choice of fit function . . . . .	196
7.6.4	Fit of toys (varied measurements) . . . . .	196
7.6.5	Resulting nominal value and uncertainty . . . . .	197
7.6.6	Additional remarks . . . . .	204
7.6.7	MC closure test of the iterative procedure . . . . .	205
<b>8</b>	<b>Are we sure? Summary of uncertainties</b>	<b>207</b>
8.1	Uncertainty sources . . . . .	207
8.2	Correlations . . . . .	208
8.2.1	Correlation factor of a ratio of random variables . . . . .	210
8.2.2	Correlation factor for the sum of independent uncertainty sources . . . . .	211
8.2.3	Separating correlated and uncorrelated uncertainties . . . . .	212
8.3	Potential correction of the visible $p_T$ . . . . .	213
<b>9</b>	<b>What do we learn? Results</b>	<b>215</b>
9.1	Cross sections and nuclear modification factors . . . . .	215
9.2	Conclusion . . . . .	222
<b>Part II</b>	<b>Partonic energy loss in the quark-gluon plasma</b>	<b>223</b>
<b>10</b>	<b>Are we lost in colour? Energy loss in the QGP</b>	<b>225</b>

10.1 BDMPS gluon emission . . . . .	225
10.2 Radiative energy loss model . . . . .	228
<b>11 Is it universal? Many systems</b>	<b>233</b>
11.1 Many systems . . . . .	233
11.1.1 Measurements to be fitted . . . . .	233
11.1.2 Extraction of the energy loss in various systems and centralities . . . . .	234
11.1.3 Universal $p_{\perp}$ shape of the modification factor . . . . .	239
11.2 A new scaling . . . . .	241
11.2.1 Principle: energy loss versus collision activity . . . . .	241
11.2.2 QGP density and multiplicity . . . . .	243
11.2.3 Glauber models . . . . .	244
11.2.4 Radius, density and nuclear thickness . . . . .	244
11.2.5 Number of participants and of binary collisions . . . . .	245
11.2.6 Centrality versus impact parameter . . . . .	249
11.2.7 Path length . . . . .	250
11.2.8 Transverse area . . . . .	256
11.2.9 Scaling results . . . . .	257
<b>12 Can we do better? Checks and improvements</b>	<b>263</b>
12.1 Checks of energy loss and path length formalism . . . . .	263
12.1.1 Alternative effective path length . . . . .	263
12.1.2 Independence of various integrals . . . . .	264
12.1.3 Separating time and space dependences . . . . .	267
12.2 $N_{\text{part}}$ and centrality in the Woods-Saxons model . . . . .	268
12.3 OO predictions . . . . .	268
12.4 Tests of scaling violation . . . . .	269
12.4.1 Logarithms of the energy . . . . .	269
12.4.2 Quarks vs gluons . . . . .	270
12.5 $v_2$ . . . . .	271
<b>List of Abbreviations</b>	<b>273</b>
<b>Values of Physical Constants</b>	<b>275</b>
<b>Bibliography</b>	<b>277</b>
<b>Acknowledgements</b>	<b>291</b>

# Editorial choices

Here are presented a few editorial shortcuts, choices, and conventions, that are adopted in this document, for clarity or to lighten the text. This list might not be exhaustive.

- The particle names might or might not be followed by a noun like ‘particle’, ‘candidate’, or ‘meson’. A general rule applied in this document is to always have some noun following the particle names, but not necessarily an equivalent of ‘particle’; for example, I could say ‘the  $B_c$  production is scarce...’ instead of ‘the  $B_c$  meson is produced scarcely...’.
- I will always consider a  $(+, -, -, -)$  Minkowski metric in relativistic calculations.
- I will use natural units in most of the text. This means that the values of  $c$  and  $\hbar$  are fixed to 1. The mass, momentum, and energy quantities can hence all be quantified in units of energy (typically GeV), while length and sometimes time will be in units of femtometers (fm), dropping the  $c$  factors that are transparent for the values given to these quantities. However, I will feel free to explicitly mention the factors of  $c$  and  $\hbar$  or to leave them implicit, depending on the context – I will typically include them only in introductory parts.
- The momentum transverse to the beam will be denoted as  $p_T$  or  $p_\perp$  in the first or second part of this thesis, respectively.
- I might, in some places where it lightens the text, personify accelerator names, *e.g.* ‘LHC is a great machine’ instead of ‘the LHC is a great machine’ which would be grammatically more correct. Similarly, to designate collaborations or experiments, I might use their names as subject, such as ‘observed by CMS’ instead of ‘observed by the CMS Collaboration/detector/experiment’.
- Similarly, I will often omit the word ‘collisions’ in the expression ‘in pp/PbPb collisions’, or use ‘pp/PbPb’ as an defining adjective, *e.g.* in ‘the pp distribution is shown ...’.
- I will sometimes write only ‘MC’ (or ‘MC sample’) instead of ‘MC simulation’ (or ‘MC simulated sample’).
- In general, the new (technical) terms that are introduced are italicised, or emboldened if it helps the readability of the current section.
- The inclusion of charge-conjugated processes (in particular the decay channels) is always implicit, unless explicitly mentioned. I will also freely mention or omit the charge of the considered particles: for example ‘ $B_c$  mesons’ or ‘ $B_c^+$  mesons’ should both be understood as including both charges ( $B_c^\pm$ ), and ‘D mesons’ as comprising  $D^0$ ,  $D^+$  and  $D^-$  mesons.



## Preamble

# Why do we seek knowledge?

I will present my work about the  $B_c^+$  meson in chapters 2 to 9 (it is also described in a public summary [1]) and about energy loss in chapters 10 to 12 (it is also partially shown in Ref. [3]). Before discussing it, let me share some thoughts concerning the root motivations for me – and, I believe, for many of my fellow humans – to understand the essence of our surrounding world, without stopping at bare appearances, nor contenting ourselves with appeasing our physical needs. ‘Motivations’ is not a light word: it means that the work I did for more than three years, the ensuing results, and this thesis would not exist without the state of mind I explain here.

Why don’t I fall through my chair, which would prevent me from writing these words? Because chairs are solid. But what does ‘solid’ mean? It means I do not pass through it. And if I try to avoid circular logic: the nature of its material is tight enough not to let through the material of my body. What is ‘tight enough’? What is the material really ‘made of’? Sentient beings like me, the reader, and their ancestors are and were faced with a daily choice. I could live without questions, acting on sole impulses and instincts, giving in my body to the maximum entropy it longs for or, at best, to my animal instincts pushing me to simply survive and reproduce. Or I could wonder – why am I here? – how do I know I am here? – and what does this stream of information from the outside to my consciousness mean? Without moral judgments (yet) about what is preferable, the second attitude, curiosity-driven, has led us to build civilisation, thanks to the progressive and cumulative understanding of our environment. But more importantly, as human beings are self-conscious, they have an inevitable thirst for understanding: wondering is not a choice. It is a self-imposed demand for purpose – or, in less finalist terms, a quest for the reason of our existence.

Who better than I. KANT can describe our insatiable thirst for knowledge? “Human reason has the particular fate in one aspect of its cognitions: that it is harassed by questions, that it cannot brush off; for they are posed for it by the nature of reason itself, but that it cannot answer; since they transcend all capacity of human reason.”<sup>1</sup> It is mainly because we are gifted with the capacity of reason and self-awareness that we harass ourselves with so many existential questions. It is not a useless tautology to notice that, without reason, we would not realise that we are gifted (or not) with reason. This resembles much the anthropic argument so famous in modern physics: B. CARTER [4], R. DICKE (for the age of the Universe [5]), S. WEINBERG (for the cosmological constant [6]), and others have hypothesised that the very fine tuning of the physical constants of our Universe, necessary for its stability given the current theories<sup>2</sup>, would

<sup>1</sup>“Die menschliche Vernunft hat das besondere Schicksal in einer Gattung ihrer Erkenntnisse: dass sie durch Fragen belästigt wird, die sie nicht abweisen kann; denn sie sind ihr durch die Natur der Vernunft selbst aufgegeben, die sie aber auch nicht beantworten kann, denn sie übersteigen alles Vermögen der menschlichen Vernunft.”, I. KANT in *Kritik der reinen Vernunft* (*Critique of Pure Reason*), 1781.

<sup>2</sup>A dire example of this in contemporary physics, relatively far from the main topic of this thesis, being the naturalness of the Higgs boson interactions: the hierarchy problem asks why the observed Higgs mass is so small, given that its bare mass is renormalised with corrections that could reach up to the Planck mass – if no other theory than the Standard Model is found at lower scales.

be possible because of an enormous multitude (maybe of order  $10^{500}$ ) of ‘parallel’ universes where other sets of physical constants are tried. Most of these universes die down quasi-instantaneously because these physical constants cannot yield stable matter, and among the tiny fraction of stable universes, a tiny fraction could present conditions favourable for the emergence of life – and maybe of an intelligent one. Only in such universes could some beings wonder why they exist, why they are ‘lucky’ enough to live at the right time, in the right place, in a stable Universe. I propose that it is mainly because we *can* reason (with such immense luck) that we should have a moral obligation to do so.

Coming back to my chair, wondering why I do not fall through it (and why I do fall when no chair supports me) rapidly led me to interrogate the nature of matter itself and its interactions. It brought to life topics and objects of study that are not of direct interest to my survival, for which I became curious and at which I started to marvel. Questioning the very basics of our existence leads to hidden unknowns, things we did not even suspect to be a matter of interest. This mindset is pure self-awareness; human beings would only be pretty weak animals if it was not for their ability to consider themselves as only a part of the world and not the centre of it. Being the sole subject of our sensitive relation to the surrounding world does not mean that we are the centre of it. Understanding this leads to a knowledge that is not directly linked to our mechanical body – even if it might be found later to be of great use to us. Civilisation is based on building things (such as scientific knowledge) that are not a direct answer to our physiological needs. The uncertainty about the potential application of some research topics to improve human comfort is what makes fundamental science one genuine specificity of human beings compared to other moving animals. This sheer curiosity ‘for the sake of it’ brings us higher than our physical condition, or at least maintains us in this illusion. This tentative escape from being trapped in a body helps – myself, at least – to fight the anxiety of the unanswerable question of the reason for our existence.

Now we have questions, but what answers are we able to reach? Are there answers that are not themselves other questions? Is there a consummation to the search for knowledge? I like to see the understanding of matter itself as what all other problems concerning objects made from this matter are based on; despite its naivety, this idea is what principally drove me to study this field. The resulting chain of ‘why’ questions naturally leads to long for universal phenomena acting on elementary objects. But are there really fundamental symmetries or pieces of matter that are ‘atomic’ in the etymological sense (*i.e.* indivisible) from which all known matter is made? It is not guaranteed that this ontological reductionism will satisfy our thirst for understanding – though travelling this way is how we can keep trying to answer our existential questions. As the circle of knowledge grows, its frontier with the things we ignore grows as well – and the space of all there is to know might be infinite. The most disappointing aspect of the gift of reason is to realise that it does not make us gods: Nature is probably infinitely complicated. Whatever stratagem we use to understand it with building blocks, we only extend the regime beyond which our theories cannot be tested – for example the LHC (see section 3.1) is a tool that pushes back this frontier, would it be only through its record energy density.

Nature is also extremely simple if taken as how it ‘looks’ to us: a sequence of interactions of the outside world with our living envelopes. In other words, a sequence of observations performed with our imperfect detectors – our eyes or the tools we invent. However, quantum physics historically made clear that Nature is more than what we see and that the outside world must be theorised to be understood better: purely mathematical tools were applied to physical systems never conceived of before and predicted their behaviour with great accuracy. Quantum physics, contrarily to, *e.g.*, the understanding that the Earth is round or that the passing of time depends on the observer, demonstrated an ‘invisible’ phenomenon only by using mathematical concepts. Phenomena call for theories, which can, in turn, predict phenomena that were not suspected before because they were too far from the regimes and scales we interact with daily

or through current experiments. Therefore, theories must be tightly linked to experiments and existing phenomena; and the observation of phenomena cannot be an actual quest for knowledge if it is not accompanied by models trying to understand them by progressive simplifications.

Around the end of the XIX<sup>th</sup> century, all the fundamental laws of Physics were thought to be discovered and soon to be archived knowledge – except for ‘minor’ remaining problems, as the threshold in the wavelength of the light absorbed in the photoelectric effect, or the blackbody radiation spectrum. The circle of knowledge was believed to be almost complete, except for some corridors leading to the unknown. This unknown happened to be modern physics based on quanta and probabilities. The situation could be compared to that of the Standard Model of particle physics (see section 1.2) that currently describes so well all current observations, but shows global inconsistencies and shortcomings, including the link of particle physics with the regimes where gravity is relevant. Particle physics and gravity are important in dramatically different regimes, as explained in section 1.1, so it is a central challenge of XXI<sup>st</sup> century Physics to gather them in a global theory. But is this a necessity or a vain attempt to seek simple explanations to literally everything? I do not think this can be answered, so let me stop here the chain of questions. Something we can rely on is that one goal of science consists in seeking generality or universality (could this be called ‘truth’?) in observed phenomena, to ultimately try to predict the future. This is based on the paradigm that similar initial conditions will yield similar consequences (excluding the chaotic systems for which initial conditions are never measured precisely enough to allow for predictions).

However, as K. POPPER made abundantly clear in his work [7], this quest is properly scientific only if it claims theories and models that are falsifiable. So any good theory, describing an object adequately in a specific regime, is made to be falsified in the future by a theory that will manage to encompass it and describe phenomena in a broader regime. In Ref. [8], C. ROVELLI notes that the ideas of ARISTOTLE can be understood as an approximation, in a certain regime, of the theory of gravitation of I. NEWTON, which is itself an approximation of the general relativity of A. EINSTEIN. Similarly, the chemistry of elements, keeping atomic nuclei intact, is a low-energy approximation of sub-atomic physics, though emerging phenomena in chemistry could hardly be guessed from first principles, as the collective phenomena in statistical physics. H. G. WELLS said that “The search for divine truth is like gold washing; nothing is of any value until most has been swept away”<sup>1</sup>. The Renaissance era progressively gave birth to the scientific method because it was based on the idea that we are “dwarves sitting on the shoulders of giants” in our efforts towards knowledge and seeing far away. Let me humbly try here to add a tiny ingredient to this knowledge recipe, along with I. NEWTON’s falling apple and the hot soup of the quark-gluon plasma (QGP, see section 1.4). Let me hope it will be removed in the future only for a more broadly-applicable ingredient to replace it, giving future dwarves a better taste of their Universe. I will observe a particle, the  $B_c$  meson, to try setting forth new phenomena that can produce it or melt it in the QGP medium it evolves in. I will also try to demonstrate a universal behaviour of the energy loss acting on particles in this medium.

---

<sup>1</sup>Herbert George WELLS in *God The Invisible King*, 1917.



“Pure logical thinking cannot yield us any knowledge of  
the empirical world; all knowledge of reality starts  
from experience and ends in it.”

---

Albert EINSTEIN<sup>1</sup>

## Chapter 1

# Where do we start? From apples to soup

### 1.1 What interactions at what scale?

I justified in the preamble, with the support of brilliant minds of the last centuries, that physical phenomena should be studied by progressive approximations, starting with understanding what is closest to our range of observations. A pragmatic goal of modern Physics is to find models that fit the reality such as it is observed with the experimental precision available in our epoch. To this end, it is important to clarify what phenomena are important in what regimes, which is illustrated in this section. The following sections of this chapter will gradually introduce the reader to the field that frames this thesis: section 1.2 briefly presents the ingredients of the Standard Model of particle physics, sections 1.3 and 1.4 respectively introduce the strong force and the quark-gluon plasma, and section 1.5 defines terms and concepts essential to this thesis.

The scales (of time, energy, or mass) at which objects are looked at are paramount to theorising their behaviour well enough to be predictive. This is related to the concept that any structure can be detected only with probes that have a wavelength smaller than the typical size of the structure, *i.e.* that have a high-enough energy. For example, visible light can only probe structures at length scales  $\gtrsim 0.5 \mu\text{m}$ . This is also a basic explanation of why searching for ever smaller elementary structures in matter calls for machines that break energy records (as the LHC presented in section 3.1.2). Another example of the importance of scales within particle physics, detailed in section 1.3.2, is how the perturbative and non-perturbative processes are separated in the cross section calculations at hadron colliders.

The interactions of particles and, by extension, of living beings, is what make them non-inert (in GALILEO’s sense), or alive. Let us illustrate here that the interactions relevant to the physics of small objects include the electromagnetic, weak and strong forces, but exclude gravity, by many orders of magnitude.

---

<sup>1</sup>The Herbert Spencer lecture, delivered at Oxford, June 10, 1933; reproduced in *Ideas and Opinions* by Albert Einstein (Alvin Redman Limited, London, 1954), p. 270. Published in *Mein Weltbild*. From this quote was probably derived this one, attributed to Einstein on the Internet: “*Information is not knowledge. The only source of knowledge is experience*”.

### Gravity vs electromagnetic force

It would be tempting to see the atom (let us consider the simple hydrogen atom, assuming a classical view of the electron rotating around the proton) as a system bound by gravitation like orbiting planets. However, the associated gravitational attraction is

$$G \frac{m_e m_p}{a_0^2} = 3.6 \times 10^{-47} \text{ N}$$

where  $a_0$  is the Bohr radius, *i.e.* the approximate radius of a ground state hydrogen atom. The Coulomb law gives the electromagnetic attraction between the proton and the electron as:

$$\frac{e^2}{4\pi\epsilon_0 a_0^2} = 8.2 \times 10^{-8} \text{ N}$$

which is similar to a weight of 10 micrograms (on Earth) pushing together the electron and the proton. It does not seem much, but is  $2 \times 10^{39}$  times stronger than the gravitational force!

### Gravity vs residual strong force

The gravitational force between a proton and a neutron (typically in the deuteron) is

$$G \frac{m_n m_p}{d^2} = 1.8 \times 10^{-35} \text{ N}$$

where the distance between the proton and neutron centres of mass is taken to be  $d = 2r_d - r_p = 3.2 \text{ fm}$ , where  $r_d$  is the measured charge radius of the deuteron and  $r_p = r_n \simeq 1 \text{ fm}$  is the approximate size of the proton or neutron.

The modern formulation of the strong force will be properly introduced in section 1.3.1. However, let us take here a historical perspective on the residual strong force binding nucleons in the atomic nuclei. It was modelled in the 1940's by YUKAWA [9]. He actually predicted the existence of the pion (discovered in 1947), which he thought was the strong force carrier. It indeed transmits a remainder of the strong force between nucleons. Considering a finite lifetime for the exchanged pion, he multiplied the Coulomb quadratic force by a decreasing exponential, such that the force between the proton and neutron would be:

$$F_{\text{strong}}(d) = -f_{\text{strong}} \frac{e^{-d/d_0}}{d^2} \quad (1.1)$$

where  $d$  is here the distance between the nucleons, and  $f_{\text{strong}}$  is the strong force magnitude to be determined. The Coulomb potential is linked to applying a Poisson equation of the di-nucleon system. A pion mass term can be added in the latter, making the solution deviate from the Coulomb potential, similarly to a drag term<sup>1</sup>; this yields  $d_0 \sim 1/m$ , where  $m$  is the mass of the actual exchanged meson<sup>2</sup>, leading to  $d_0 = 1.54 \text{ fm}$ . A detailed meson exchange modelling shows that the force becomes repulsive when  $d \lesssim 1 \text{ fm}$ , which is why the nucleons have a so-called hard core. So Eq. 1.1 empirical expression is only qualitatively valid at somewhat larger distances.

The binding energy of the deuteron was measured, via the energy of the photon needed to disintegrate it, to be 2.2 MeV. One can interpret this as the scale of the first energy level of the orbital angular momentum, hence equal to  $l\hbar$  with  $l = 1$ . Considering that the centrifugal force  $2mv^2/d$  (where  $m$  is the proton or neutron mass and  $v$  is the relative speed between the proton

<sup>1</sup>As I aim for orders of magnitude here, I neglect the subtlety that the decreasing exponential can be added to the potential or to the force depending on the phenomenological approaches, which lead to somewhat different expressions.

<sup>2</sup>Detailed calculations say it is mainly a pair of pions or one heavier  $\sigma$  meson.

and the neutron)<sup>1</sup> must compensate the strong force of Eq. 1.1, we find that at equilibrium  $v = \sqrt{f_{\text{strong}} \frac{e^{-d/d_0}}{2md}}$ . Using the angular momentum  $mvd = \hbar$ , we obtain:

$$f_{\text{strong}} e^{-d/d_0} md/2 = \hbar^2 \quad (1.2)$$

This results in  $f_{\text{strong}} = 3.3 \times 10^{-26} \text{ kg m}^3 \text{ s}^{-2}$ , if we take  $d = 3.2 \text{ fm}$  as the distance between the centres of the two nucleons. Now we finally get the magnitude of the residual strong force between the two nucleons of the deuteron:  $|F_{\text{strong}}| = 4.0 \times 10^2 \text{ N}$ . The weight of a child in Earth gravitation, at the scale of atomic nuclei, what an incredibly strong force indeed! The relative importance of the gravitational and residual strong force is about  $5 \times 10^{-38}$ , justifying by all means that it should be neglected when studying systems at nuclear length scales.

As a check of the order of magnitude of this force, one can compute the work needed to bring the two nucleons from a null distance (ignoring that the force is actually repulsive at small distances) to the equilibrium distance. It is of the order of the force times the distance, *i.e.* about  $1.3 \times 10^{-12} \text{ J}$ . Converting this into natural units indeed gives 8 MeV, of the same order of magnitude than the deuteron binding energy.

### Gravity vs strong force

It was possible to keep a perspective from classical physics in the above paragraph, with the phenomenological model of YUKAWA. However, within nucleons themselves where the strong force is the strongest, colour confinement rules, which is purely non-classical. There is however a simple model, considering that the strong force increases with distance, similar to an elastic string – that is, until this string breaks, when the potential energy is sufficient to create an particle-antiparticle pair (by the mass-energy equivalence  $E = mc^2$ ). This is fundamentally different than the residual strong force with a form close to a Coulomb force decreasing with distance, that is acting between uncoloured objects, the nucleons. This is intimately linked to the non-abelian nature of QCD (see section 1.3.1): the force carriers (the gluons) interact with themselves, and their number increases towards large distances.

Such a string carries a potential energy proportional to the distance between the two quarks:  $V = \frac{1}{2}kd^2$ . The string breaks when it carries enough energy to create a light hadron, so typically  $V_{\text{break}} \sim 140 \text{ MeV}$ . This happens at distances close to the size of mesons, around 0.5 fm. From these values of  $V$  and  $d$ , one estimates that  $k \simeq 2 \times 10^{20} \text{ N m}^{-1}$ . The equivalent classical force of a spring is  $F = kd \simeq 1 \times 10^5 \text{ N}$ . A spring constant this high is quite hard to conceive of: a force equivalent to a weight of 2 tons should be applied to extend this string by  $10^{-16} \text{ m}$ , *i.e.* about 20% of its assumed length. The weight of a large car in Earth gravitation is hidden within each nucleon of our bodies! All this so we are able to live from stable matter, to reflect about what we are composed of.

Now, comparing again this force to the gravitational force between two quarks, I need to estimate the quark masses. It is warned here that only toy mass values are used, because measuring the gravitational mass of quarks would imply that they can travel freely, which is not the case. Quark masses are usually quoted ‘naked’ or ‘bare’ in a given renormalisation scheme, but I choose here to quote rough masses that include the ‘dressing’ of the gluon interactions that surround them within hadrons. This way, the ‘dressed’ mass of a light quark ( $u$  or  $d$ ) is about half of the pion mass. One deduces a gravitational force between two light quarks at a distance  $d = 0.5 \text{ fm}$  to be  $G \frac{(m_\pi/2)^2}{d^2} \simeq 4 \times 10^{-36} \text{ N}$ , whose magnitude is hence  $4 \times 10^{-41}$  that of the strong force!

---

<sup>1</sup>Relativistic effects are neglected, as they are found a posteriori to be small, with  $v \simeq 0.1c$ .

### Strong vs electromagnetic forces

No wonder now that gravity can be fully ignored when looking at elementary constituents of matter – except if the energy experimentally available today was risen immensely so that the mass/energy of the system became relevant. But among the electromagnetic, weak, and strong forces, which dominates?

Their names are a good hint, but let us rather check the values of the coupling constants. Those drive the amplitude of the interaction from a given Feynman diagram. These ‘constants’ actually slightly depend on the energy scale of the process at hand. At the typical scales of processes studied at the LHC ( $1 - 100 \text{ GeV}$ ),  $\alpha_{\text{QED}} \sim 7 \times 10^{-3}$  for the electromagnetic force,  $\alpha_S \sim 0.1 - 1$  for the strong force, and  $\alpha_w \sim 10^{-6} - 10^{-7}$  for the weak force. The low coupling for the weak force is linked to the high mass of the W and Z bosons that carry it. Its order of magnitude is obtained here through the ratio of lifetimes of decays that do or do not change the quark flavours. As the coming section explains, the weak force is indeed the only one that can affect the quark flavour and that is why it is not ignored despite its low coupling.

## 1.2 The Standard Model

The Standard Model (SM) is describing incredibly well the behaviour of objects from scales of the inside of atoms ( $\lesssim 10^{-11} \text{ m}$ ) to the highest energy scales ever reached in a human laboratory, at LHC ( $13 \text{ TeV}$ , *i.e.*  $\gtrsim 10^{-20} - 10^{-19} \text{ m}$ , see section 3.1.2). A historical perspective on how CERN experiments were essential in building this theory is given in section 3.1.1. It is based on three gauge symmetry groups associated to the modelled interactions:  $SU(2)_{\text{left}} \times U(1)_{\text{hypercharge}} \times SU(3)_{\text{colour}}$ . Fig. 1.1 is a schematic overview of the elementary particles whose behaviour it describes. All currently known objects are made of these particles.<sup>1</sup> The behaviour and interactions of the SM particles are described by a Lagrangian. Let us give here a condensed form of it, the one printed on CERN’s merchandising<sup>2</sup>:

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

where *h.c.* designates the hermitian conjugate of the preceding term, and each term will be associated to particles and interactions in the following. If the action associated to this Lagrangian could be exactly minimised with perfectly known initial conditions, it would yield us the average behaviour<sup>3</sup> of all the particles of the Universe – up to the elements and regimes, of unknown nature yet, that the Standard Model fails to describe.

The elementary particles composing the SM can be categorised in multiple ways. First, the *gauge bosons* (light blue corona in Fig. 1.1) are the mediators of forces, while the *Higgs boson* holds a special role (centre of Fig. 1.1). The *fermions* (outer circle red particles) are the basic components of matter, and are subject to the Pauli exclusion principle (two fermions can never have the exact same state and quantum numbers). The elementary bosons of the SM are:

- The *gluon*, mediator of the strong interaction. It carries and interacts with the colour charge, associated to the  $SU(3)_{\text{colour}}$  sub-group of the SM.
- The *photon*, transmitting the electromagnetic force. It interacts with particles that carry electromagnetic charge.

<sup>1</sup>Except for the dark matter, that constitutes a large part of the mass of our Universe, but whose nature is still unknown.

<sup>2</sup>Developing explicitly each part of this symbolic form would be out of place in this thesis.

<sup>3</sup>By ‘average behaviour’, I mean here the average over many Universes with exactly similar initial conditions and physical constants. Particle physics indeed yields only probability distributions for the positions and momenta of particles. Knowing this, ‘Is the Lagrangian deterministic?’ is a philosophical question.

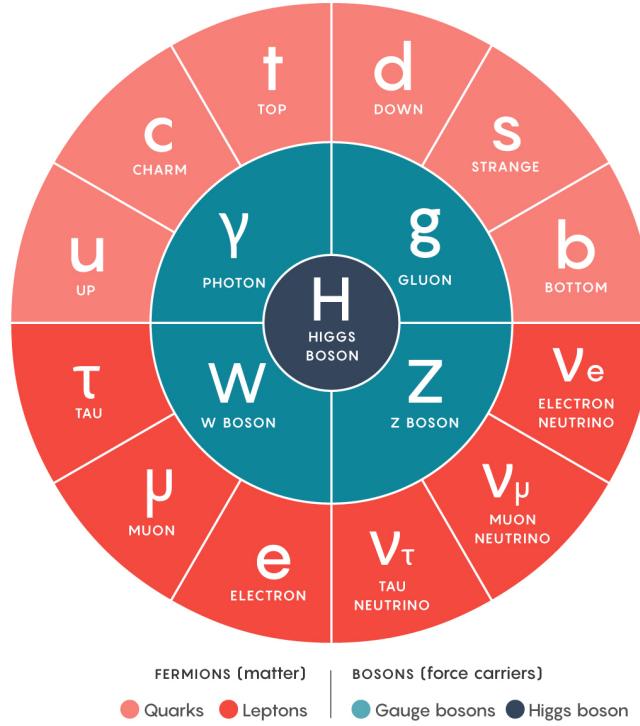


FIGURE 1.1: The particles composing the Standard Model of particle physics.

- The *weak bosons*  $W^\pm$  and  $Z^0$  carry the weak force. The weak force is unified with the electromagnetic force within the  $SU(2)_{\text{left}} \times U(1)_{\text{hypercharge}}$  symmetry group. They interact with all particles (except the gluon). However, fermions interact with the  $W^\pm$  bosons only through their left-handed fields – the weak interaction indeed violates the charge-parity symmetry. The  $W^\pm$  charged bosons are the only particles that can change the flavour of quarks; the  $Z^0$  neutral bosons cannot.
- The *Higgs* boson (of field  $\phi$  and kinetic term  $|D_\mu \phi|^2$  in Eq. 1.3) has a peculiar role. The spontaneous breaking of the symmetry of its potential ( $V(\phi)$  in Eq. 1.3) allows for the weak bosons to have mass. This allows, in turn, for the Yukawa interaction ( $\psi_i y_{ij} \psi_j \phi$  in Eq. 1.3) to give mass to all fermions (of fields  $\psi_i$ ). The Higgs boson interacts with all massive particles.

All elementary bosons have spin 1, except for the scalar (spin 0) Higgs boson.<sup>1</sup> The photon, gluon, and weak bosons are the gauge bosons of the  $U(1)$ ,  $SU(2)$  and  $SU(3)$  symmetries of the SM. The term  $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$  of Eq. 1.3 describes the kinetic energy and the self-interactions of these gauge bosons.

The fermions are all of spin 1/2. Their kinetic energies and interactions with other fermions and bosons are encoded in the term  $i\bar{\psi}\not{D}\psi$  in Eq. 1.3. Among the elementary SM fermions, there are first the *leptons*. They consist in three doublets composed of a *neutrino* and an *electron* or one of its heavier sisters, the *muon* and the *tau*. Their left-handed component interact with the weak force. Electrons, muons, and taus interact with the electromagnetic force, but not neutrinos. The mass of neutrinos has not been measured, but the observed oscillations between their flavours imply that at least two out of the three neutrinos have a finite mass.<sup>2</sup> They interact

<sup>1</sup>This excludes the graviton, that would be of spin 2. The graviton will be properly included in this microscopic framework only when a unified description of particle physics and gravitation is found.

<sup>2</sup>It is interesting to note that this phenomenon and the finite mass of neutrinos was not predicted in the original Standard Model – though it was not excluded either.

very feebly with matter (only through the weak interaction), such that they are detected only in specialised very large detectors, and pass unobserved through the CMS detector used in this thesis. Leptons do not interact with the strong force.

Next come the *quarks*, of most interest to this thesis. There are three generations of them, each composed of an up-type (of electromagnetic charge  $+\frac{2}{3}$ ) and a down-type quark (of charge  $-\frac{1}{3}$ ). The u, d (*up* and *down*) quarks are very light, the c, s (*charm* and *strange*) quarks are of intermediate mass, and the t, b (*top* and *beauty* – also called *bottom*) quarks are the heaviest. The top quark is the heaviest particle of the SM (173 GeV). Quarks undergo all SM interactions, but their nature (*a.k.a.* their *flavour*) can only be modified through the weak  $W^\pm$  bosons. This flavour mixing is encoded in the Cabibbo-Kobayashi-Maskawa matrix that acts between the down-type and up-type quark triplets [10]. It favours the mixing between quarks of the same generation (*e.g.* between u and  $\bar{d}$  or between t and  $\bar{b}$ ), and strongly disfavours mixing the first and third generations.

Quarks carry a *colour charge* (whose three states are conventionally called red, blue and green) that make them sensitive to the strong force. Their behaviour is hence described by *quantum chromo-dynamics* (QCD), which is the topic of the coming subsection. Each quark flavour exists in three different colour states, while there are eight gluons because they combine a colour and an anti-colour. Coloured states cannot travel long distances due to *colour confinement*. This means that the quarks must bind with partners that compensate their colour to form stable particles. Those quark composites are called *hadrons*. Existing hadrons are composed of quarks of any of the five lightest flavours, while the top quark decays before the typical time scale for hadronisation.

The original quark model described essentially hadrons that gather two quarks (a quark and an anti-quark of the associated anti-colour charge), called *mesons*, and hadrons binding three quarks (one of each colour), called *baryons*. Mesons that gather a quark and an antiquark of the same nature have no net flavour quantum number – they are *hidden flavour* mesons, contrarily to the *open flavour* mesons that contain quarks of different natures.

There are multiple hints that the Standard Model is only an excellent approximation of particle physics in the regime accessible in our experiments, of which some were suggested above. Let me non-exhaustively quote the mysteries of the nature of dark matter and dark energy (which accelerates the expansion of the Universe), the naturalness of the Higgs properties (linked to the hierarchy problem), and most importantly the inadequacy of this theory with the cosmological scales where gravity is relevant.

## 1.3 The strong force

### 1.3.1 Quantum Chromodynamics

The mediator of the strong force (the gluon) also carries the colour charge it is sensitive to. This is also the case for the weak force (though with a much smaller coupling), but not for the electromagnetic force as the photon is neutral. In other words, the QCD gauge group is non-abelian, and the associated Lagrangian contains vertices of three and four gluons. Each gluon carries a colour and an anti-colour, and hence combines the symmetry groups **3** and  **$\bar{3}$** . This results in a group of eight elements, so there are eight types of gluons – plus one arrangement that cancels the colour. This is why any arrangement of two colour-charged particles can be *colour-octets* interacting with the strong force, or *colour-singlets* that do not interact with it. The quarks carry only one colour charge, so they are part of *colour-triplets*. The elementary particles that carry a colour charge (quarks and gluons) are called *partons*.

The gluon self-interactions lead to another specificity of QCD: the strength of its coupling  $\alpha_S$ , despite being much larger than that of other interactions, decreases towards higher energy

scales (*i.e.* towards smaller distances). This is the **asymptotic freedom** of QCD<sup>1</sup>, illustrated in Fig. 1.2. This can be understood in analogy to the screening of the electric charge in a plasma at large distances, that makes the electromagnetic coupling decrease. Many  $e^+e^-$  pairs are indeed created close to a negative test charge. These pairs orient the positive part of their dipole towards the probe charge, which is hence attenuated at large distances; at small distances, some pairs are too far away to contribute to the screening. On the contrary, in QCD, the colour charge is anti-screened (higher coupling and effective charge) at large distances: the colour charge is spatially diluted by the gluons that, at small distances, are numerous and carry the colour away from the probe charge.

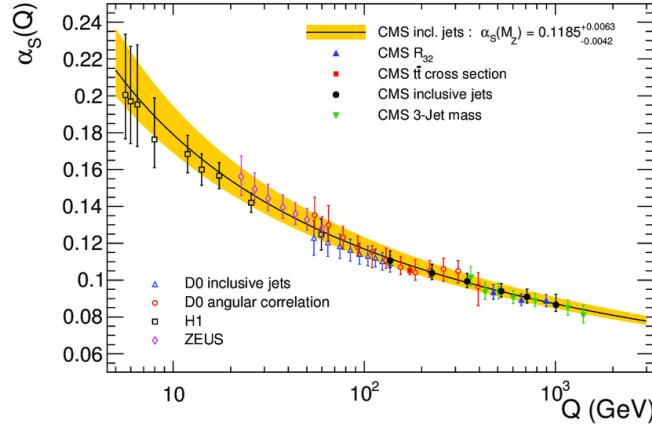


FIGURE 1.2: Asymptotic freedom in QCD, from various measurements of the strong coupling.

This running of the coupling can be calculated with the vacuum polarisation diagrams (responsible for the pair creation) and the renormalisation group equations, yielding, at scale  $Q$  (this was first done, independently, by GROSS and WILCZEK [11], and by POLITZER [12]):

$$\alpha_S(Q) = \frac{2\pi N_c}{(11N_c - 2N_f) \ln(Q/\Lambda_{\text{QCD}})} \quad (1.4)$$

where  $N_c = 3$  and  $N_f = 3$  to  $6$  (depending if  $Q$  is higher than the mass of the heavier quarks) are the number of colours and quark flavours, and  $(11N_c - 2N_f) < 0$  is responsible for the asymptotic freedom. Lowering the energy towards the hadronisation scale  $\Lambda_{\text{QCD}} \sim 0.2 \text{ GeV}$ , the energy scales are associated to non-perturbative processes. There, the coupling, relevant for interactions of individual partons and perturbative QCD, and the associated formula in Eq. 1.4 start to be ill-defined.

Finally, the phenomenon of **colour confinement**, meaning that coloured particles cannot travel long distances without binding with particles that neutralise their colour, is not well understood yet – despite it being paramount for the stability of all atomic nuclei! At energy scales smaller than  $\Lambda_{\text{QCD}}$  (*i.e.* at large distances), quarks must hadronise (*i.e.* form hadrons) through low-energy processes that cannot be described with perturbative techniques. The latter indeed assume a small coupling, which is not the case at small energies. The confinement is not directly linked to extrapolating the asymptotic freedom towards large  $\alpha_S$ : it can exist without a divergence of the coupling. Supposing confinement, there cannot be particle-antiparticle creation at distances larger than hadrons, so the  $\alpha_S$  increase towards large distances resulting from these

<sup>1</sup>As the electroweak coupling increases towards higher energy scales, this schematically supports the idea that all forces can unify at very high energy scales – similarly to the unification of the electromagnetic and weak forces at the scale of the weak boson masses. In practice, close to the Planck scale, where gravity becomes relevant and the forces should be unified, the extrapolated running couplings do not coincide, suggesting that other physical phenomena should intervene at scales lower than the Planck ones to modify these couplings.

pairs shall stop [13]. This illustrates why the separation of soft (non-perturbative, linked to hadronisation) and hard (linked to interactions of single partons at small distances) scales is necessary to understand QCD processes, in particular in the collisions of hadrons performed at the LHC. This separation of scales is detailed in section 1.3.2. Colour confinement is a very active research topic; current investigations address the ladder (*a.k.a.* ‘H’) graphs that are infrared-divergent – rather than the vacuum polarisation graphs responsible for the running coupling [13]. Another way to understand better the colour confinement is to study an extreme state of matter where it is not active: the *quark-gluon plasma* (QGP), introduced in section 1.4.

### 1.3.2 Parton distribution functions

The LHC (see section 3.1) accelerates and collides protons and lead nuclei. However, it can be argued that such a hadron collider actually collides **partons**, *i.e.* quarks or gluons contained in the nucleons, because those are the input particles of the initial hard processes taking place in the non-diffractive hadronic events (defined in section 1.5.5).

The proton is a composite of elementary particles. This was discovered experimentally by probing the proton structure with electrons of high enough energy to resolve its constituents. The energy and angular distributions of the final states of electron-proton scatterings were not consistent with elastic scatterings of point-like particles, and structure functions associated with point-like constituents of the proton were extracted. The time scale of the many interactions between the proton constituents is much larger than that associated with the high-energy electron, meaning that the electron probes a quasi-static proton content. These **deep inelastic scatterings** (DIS) measurements culminated with the HERA (*Hadron-Electron Ring Accelerator*) experiments at DESY. The kinematic framework of DIS scatterings is not detailed here; Ref. [14] gives a good description of its application to LHC. The proton structure probed by DIS is a great experimental discovery of a theoretical prediction – the quark model. It also confirmed the **BJORKEN** scaling predicted in 1969 [15]: the structure functions depend on the longitudinal momentum fraction  $x$  that the parton takes from the proton, but only marginally on the momentum transfer  $Q$  from the electron probe to the proton. This scale  $Q$  is the typical scale of the considered hard process (*e.g.* the dijet mass if this is the studied final state).

From the DIS form factors are extracted the **parton distribution functions** (PDF), which is a measure of the probability to find a parton of a given nature, at a given longitudinal momentum fraction  $x$ , when the proton is probed at a scale  $Q$ . The fraction  $x$  is usually identified with the full momentum fraction at high  $Q^2$ , where the partons from the protons have a negligible momentum transverse to the beam (the *generalised PDFs* include contributions from parton transverse momenta). The PDFs are non-perturbative functions that cannot be evaluated from first principles at a fixed order in  $\alpha_S$ , so they must be measured by DIS. Sets of PDFs are extracted from global fits to the available data, probing the proton content at multiple scales. Various theoretical collaborations perform these fits with subtle differences. Fig. 1.3 shows, for all parton natures, the most recent PDFs deduced from the fits of the NNPDF3.1 collaboration, that performs most calculations at NNLO [16].

A few characteristics of these curves are to be noticed. First, the ‘probability’ of finding **valence quarks** (the quark constituents in the initial quark model) peaks at  $x$  values of  $0.1 - 0.2$ , and is higher for the u than the d quark. This is reminiscent of the proton quark content  $uud$ , but already signifies that the momentum of the valence quarks is diluted and reduced (from the naive  $1/3$  value) by interactions with other proton constituents. Then there are so-called **sea quarks**, other quark flavours that are created in the many QCD interactions taking place within a nucleon. They are more numerous at low  $x$ , as they are easier to produce at small momenta. Finally, the **gluons** (scaled by 0.1 in Fig. 1.3) are the most numerous partons, especially at low  $x$  due to the low cost of radiating soft gluons from any parton. Gluons actually carry about

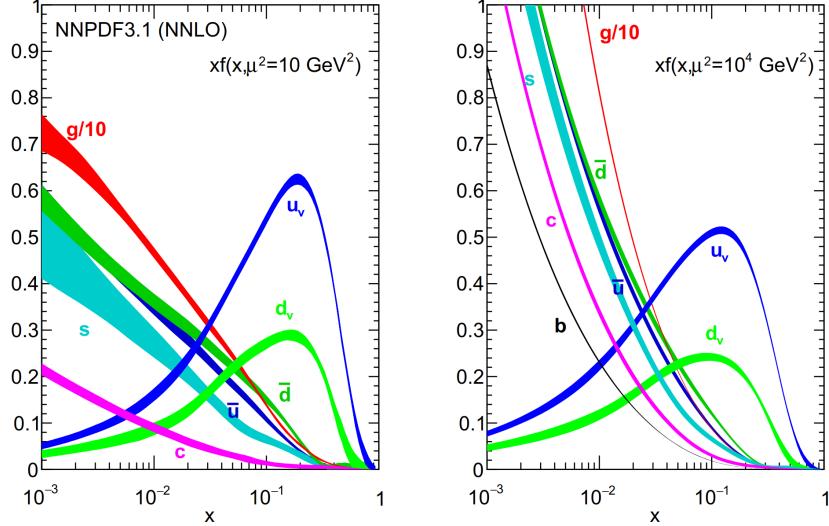


FIGURE 1.3: Parton distribution functions versus their momentum fraction  $x$  in the proton, at two different probing scales corresponding to typical energies of the final states of 3 or 100 GeV. Figure from Ref. [17].

half of the momentum of the proton.<sup>1</sup> In effect, mostly gluons are collided at LHC, except for processes of energy close to the centre-of-mass energy of the pp collision  $\sqrt{s}$ , so that the LHC might rather be called the *Large Gluon Collider*. Protons at the LHC are therefore far from the quark model vision, and resemble more a soup of gluons having low-energy interactions.

When multiple protons (and neutrons) are gathered in a nucleus, residuals of the strong force have an impact on the parton content of the nucleons, and modify their PDFs to so-called **nuclear PDFs** (nPDF). The latter are measured experimentally as well (and displayed as ratios to the PDFs of the proton), but with much larger uncertainties.

When hadrons collide at a centre-of-mass energy  $\sqrt{s}$ , the participants to the studied hard scattering (such as the production of a  $J/\psi$  meson or a high- $p_T$  jet) are in general one parton from each hadron. The key assumption of the results extracted from hadron colliders is the **factorisation** (*i.e.* the separation of scales) between the non-perturbative processes determining which partons from the protons will interact, and the hard scattering of partons whose cross section is calculated with perturbative methods. I denote  $x_1$  and  $x_2$  the momentum fractions of the interacting partons within the protons they come from, such that the (squared) energy available for the partonic hard process is  $\hat{s} = x_1 x_2 s$ . Then the cross section for producing a final state  $H$  of characteristic scale  $Q$  in pp collisions is given by:

$$\sigma_{pp \rightarrow HX}(s) = \int dx_1 dx_2 \sum_{i,j} f_{i/1}(x_1, Q) f_{j/2}(x_2, Q) \hat{\sigma}_{ij \rightarrow H}(\hat{s}) \quad (1.5)$$

where  $X$  can contain additional produced final states,  $\hat{\sigma}$  is the cross section of the partonic hard scattering at energy  $\hat{s}$ , and  $f_{i/k}(x_k, Q)$  is the PDF for a parton of type  $i$  to be found with momentum fraction  $x_k$  in the proton  $k$ . The part carrying the interesting hard-scattering information is the partonic cross section  $\hat{\sigma}_{ij \rightarrow X}(\hat{s})$ . No interference is assumed between the soft PDF processes and the hard scatterings. This hypothesis can break when considering energy loss of partons, typically that in cold nuclear matter [18], but discussing the consequences of the breakdown of factorisation goes beyond the scope of this thesis.

<sup>1</sup>Summing the integrals  $\int_0^1 dx f_i(x)$  for the parton distribution functions  $f_i(x, Q^2)$  (defined in this section) of all quark types  $i$  indeed yields only about 0.5 at the energy scales of processes studied at LHC, leaving 50% of the total momentum for gluons.

The Bjorken scaling is not exactly true: the PDFs depend logarithmically on the scale  $Q$  of the process at hand, due to the splitting of partons interfering with the hard process. This dependence is encoded in the evolution equations that were determined by DOKSHITZER [19], GRIBOV and LIPATOV [20], and ALTARELLI and PARISI [21], and that allow to determine  $f_i(x, Q^2)$  at all scales  $Q$  when it is known at some scale  $Q_0$ :

$$\frac{df_i(x, Q^2)}{d \log(Q^2)} = \frac{\alpha_S(Q^2)}{2\pi} \sum_j \int_x^1 \frac{dx'}{x'} P_{ij} \left( \frac{x}{x'} \right) f_j(x', Q^2). \quad (1.6)$$

The sum runs over all parton types  $j$  that can produce a parton of type  $i$ , and  $P_{ij}(x)$  are the *splitting functions* from ALTARELLI and PARISI. They determine the probability for a parton of type  $i$  to split into a final state containing a parton of type  $j$ . These splitting functions also lead the dynamics of *parton showers*, meaning how the partons created in hard processes split multiple times until the daughter partons reach the energy scale of hadronisation. At this point, the daughter partons hadronise into the many collimated hadrons observed in the experiments, reconstructed as a *jet*.

## 1.4 The quark-gluon plasma

### 1.4.1 A state of matter

At the densities and temperatures present in most of the current Universe, the coloured particles are confined into hadrons. However, at the very high energy density and temperature created in heavy ion collisions at the LHC, quarks and gluons move freely in a deconfined state of matter called quark-gluon plasma (QGP) [22]. It is a close-to-perfect fluid of partons<sup>1</sup> produced at the LHC at temperatures higher than any current astrophysical process ( $T \gg 10^{12}$  K, *i.e.*  $T \gg 100$  MeV). All matter in the Universe, during the first microseconds after the Big Bang, was in this state (at high temperature). It is also thought to constitute the core of neutron stars today (at high density).

A rapid transition takes places when QGP is formed or freezes out. The current understanding of the phase diagram of nuclear matter, including the QGP, is presented in Fig. 1.4. In the heavy ions collisions at LHC and RHIC, the baryon chemical potential (linked to the net density of baryons) is close to null and QGP is formed only due to the increase in temperature. In this case, there is no actual phase transition, but rather an analytical cross-over (without discontinuity in any derivative of the order parameter).

The cross over was observed with lattice QCD calculations, that implement QCD first principles on a space-time lattice, and then run computationally-expensive simulations. The order parameter of the QGP-hadron gas transition is the Wilson line parameter  $\propto e^{-Vt}$ , where  $V$  is the interaction energy of a sole quark in the medium, and  $t$  is the temporal length of the Wilson line [23]. Wilson lines are quantum field theory objects that schematically form the sides of the lattice elements. For the hadronic phase,  $V \sim \infty$ , so the order parameter is zero. It is non-zero (finite interaction energy  $V$ ) in the QGP.

During a heavy ion collision, nuclear matter visits a rich variety of states. First, the gluon density saturates in a colour-glass condensate; then, a pre-equilibrium plasma forms similarly to a disordered glass state; finally, the QGP is formed, on time scales of 0.2–1 fm. In the QGP, quarks and gluons thermalise and roam freely. When the medium temperature goes below the critical QGP temperature, a gas of hadrons forms. In the QGP, the number of degrees of freedom of the original QCD gauge theory is restored: partons, versus the hadrons in cold QCD. This change

---

<sup>1</sup>Contrarily to the weakly-interacting gas (coupling  $\alpha_S \ll 1$ ) that was initially expected, at the time of SPS discovery.

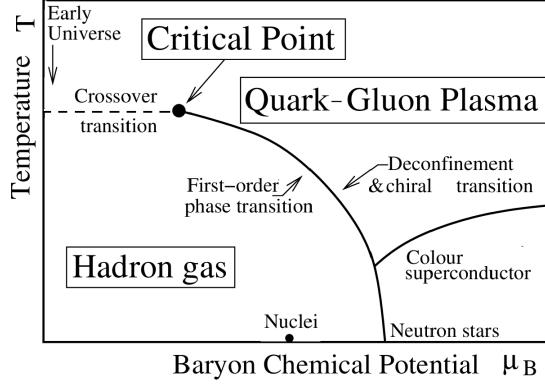


FIGURE 1.4: Phase diagram of nuclear matter, as a function of temperature and of the baryon chemical potential.

in degrees of freedom is observed with lattice QCD as well. It modifies the thermodynamical quantities (such as temperature, pressure, and internal energy) characterising the thermalised medium. This is illustrated with the temperature dependence of the energy density in Fig. 1.5: accounting for the number of colours, gluons, light quark flavours (including the strange quark when the temperature is high enough), and spins provides an order of magnitude more degrees of freedom above the critical temperature than below it.

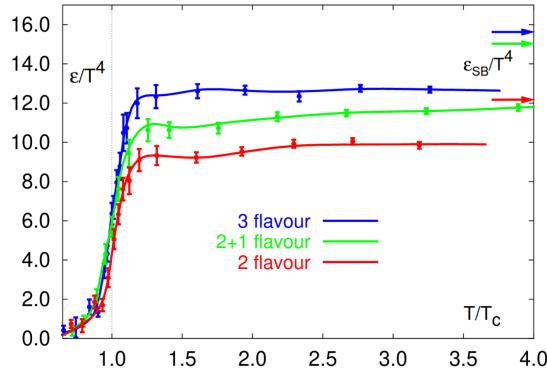


FIGURE 1.5: Energy density of QCD matter as a function of temperature. The energy density is proportional to the number of degrees of freedom of the medium, rising from that of hadrons to that of partons at high temperature. Fig. from Ref. [24].

#### 1.4.2 Categories of experimental signatures

The proof of the creation of a QGP in heavy ion collisions depends strongly on the definition of QGP. However, its formation is today widely admitted in the community, because a broad range of processes in principle associated to the QGP have been observed experimentally. Sections 3.1.1.3 and 3.1.3.3 will present a historical perspective of the discovery of the QGP at SPS, and of the experimental discoveries of some QGP signatures at the LHC, respectively. On the other hand, experiments at RHIC (*Relativistic Heavy-Ion Collider*) have already shown that the QGP is a near perfect liquid, with a viscosity over entropy ratio lower than any other known fluid [25]. Experimental discoveries due to RHIC are hence underlined here.

##### Elliptic flow

A typical signature of the QGP is the strong observed elliptic flow, which is linked to the azimuthal asymmetry of the distribution of particles, mostly encoded in its second Fourier coefficient  $v_2$ . The elliptic flow is due to the pressure gradient arising from the initial azimuthal

asymmetry of the collisions of large nuclei, and from the hydrodynamics of the QGP fluid. In 2000, RHIC measured a strong elliptic flow of charged hadrons [26].

### Jet quenching

Partons lose energy when passing through the coloured medium, which is the topic of section 2.3 (mainly for heavy quarks) and the second part of this thesis (mainly for high-energy partons). I discuss here only the effect of in-medium energy loss on jets (defined in section 1.5.1). It significantly modifies the parton showers constituting the jets observed in the detectors – the cone containing their energy becomes wider and part of their total energy is transferred to the medium. This jet quenching was first observed at RHIC in 2001 through the suppression of high- $p_T$  hadrons [27]. It was confirmed with the observation of the disappearance of (the highest- $p_T$  hadron of) away-side jets by RHIC in 2002 [28]: the di-jets produced close to the frontier of the QGP lose their momentum balance because only one jet is suppressed by the medium.

### Strangeness enhancement

The critical temperature to form a QGP is above the threshold for forming a strange quark-antiquark pair. These additional strange (anti)quarks can form final-state hadrons that would not be produced in the vacuum. Hints of the resulting strangeness enhancement (denoting an increase of the number of quarks carrying strangeness, and not only a different arrangement of quarks in hadrons) were seen already at SPS [29].

### Quarkonia

How the quarkonia and heavy quark interactions are affected by the QGP will be reviewed in chapter 2: the main effects are partonic energy loss, dissociation of the heavy quark bound states, and recombination of the heavy quarks freed in the QGP.

## 1.5 Definitions and notations

### 1.5.1 Detector and accelerator notions

A particle beam in a particle collider such as the LHC (see section 3.1) is actually composed of separated **bunches** of particles, of limited longitudinal size, so that collisions of beam particles happen only at regular time intervals. This delimits clearly the moments at which the detectors around the collision points can record a collision. Each bunch at the LHC contains about  $1 \times 10^{11}$  protons, or  $7 \times 10^7$  Pb nuclei. The colliding nuclei are composed of **nucleons**, namely protons and neutrons. The energy of a colliding nucleus is in general quoted normalised by the number of nucleons it contains, for comparison with proton collisions.

An **event** loosely designates all the activity recorded by a detector within a given bunch crossing. In PbPb collisions in CMS, this is often assimilated to the one (and usually only) PbPb collision that takes place in a given bunch crossing. In pp collisions however, multiple pp collisions take place in the same bunch crossing (due to pile-up, see section 5.2), and an event comprises all these collisions.

A **track** is the experimental estimate of the trajectory of a particle through a detector. It is in general a fit of discrete hits (*i.e.* points where the particle interacted visibly with a sub-detector) that are consistent with the passage of a charged particle (considering the possible magnetic field). It is more rare, but possible, to talk about tracks for neutral particles, for they can be detected only by absorbing their energy in calorimeters, which give a much less precise estimate of the direction than tracker detectors.

A **vertex** is a point in space from where multiple particles (or tracks, from the point of view of the detector) originate. During the vertex reconstruction, this common origin is generally inferred if the tracks spatially intersect within their fit uncertainties. The **primary vertex**

(PV) of a collision event is where the first hard interaction happened (between two partons from two protons or lead nuclei, in this thesis) and produced the majority of the particles in the event. It is hence experimentally reconstructed as a point through which many tracks of the event are passing. There can be multiple PVs in a pp event, due to multiple pp collisions in the same bunch crossing; this is called **pile-up**. A **secondary vertex** (SV) or **displaced vertex** is the decay vertex of a long-lived b- or c- hadron (typically 0.2 – 10 mm away from the PV). Separating it from the PV is an important experimental challenge. Detecting tracks that do not come from a given vertex can be done by measuring the **impact parameter** (IP), which is the distance of closest approach of this track to the vertex.

A **jet** is the experimental manifestation of the showering of a high- $p_T$  parton created in a hard scattering. Via various possible methods, jets reconstruct the history of each parton branching in the shower. They gather in a cone (that can have various opening angles, called the distance parameter  $R$ ) all final-states hadrons that are thought to originate from this initial hard parton, to estimate its energy.

In data analysis, one needs to simulate the processes of interest, as a help to understand better the data and estimate physical quantities. In particle physics, **Monte Carlo simulations** (MC) are used to generate events [30]. It is a method to draw events randomly from the true probability distributions implemented in the program. Those distributions consider both the physics one is interested in (*i.e.* the hard process generating the probe of interest), the softer underlying event activity of the collision, the particle decays, and the interaction of all resulting particles with the detector. The parameters of these programs are often tuned so that they describe better the available measurements of many different processes.

### 1.5.2 Coordinates

The cartesian coordinates  $x$ ,  $y$ , and  $z$  in the LHC detectors studied here are defined such that they form a direct basis, with the  $z$ -axis defined as the beam direction and the  $y$ -axis pointing to the sky. The  $z$ -axis points anti-clockwise in the LHC ring. I will denote as  $p$  the **3-momentum** of a particle (not to confused with the symbol  $p$  for the proton). Its component transverse to the beam is  $p_T = \sqrt{p_x^2 + p_y^2}$ .

The **rapidity**  $y$  (distinguished from the cartesian coordinate from context) of a particle is linked to the angle  $\theta$  between the positive  $z$ -axis (*i.e.* the beam direction) and the particle's momentum. It is defined as

$$y = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right) = \tanh^{-1} \left( \frac{p_z}{E} \right) = \tanh^{-1}(\beta_z) \quad (1.7)$$

where  $E$  is the total energy and  $p_z$  the longitudinal momentum. This is the equivalent of the standard rapidity in special relativity, except that it considers only longitudinal boosts. When a particle is close to the beam (at positive  $z$ ),  $E \sim \sqrt{p_z^2 + m^2}$ . This means that at relativistic energies  $y \rightarrow \frac{1}{2} \ln((2p_z/m)^2) \simeq \ln(\frac{2E}{m})$ ; this equals  $\ln(\frac{\sqrt{s}}{m_p})$  for the beam particles, about 8 to 9 at LHC energies. On the other hand, if a particle is nearly perpendicular to the beam, its  $p_z$  is small, so  $y \rightarrow 0$ .

The rapidity is interesting, because it is additive under longitudinal Lorentz transformations, *i.e.* a particle of rapidity  $y_1$  in a given frame has rapidity  $y_1 - y_\beta$  in a frame that is longitudinally boosted with velocity  $\beta$ . This implies that the difference between the rapidities of two particles is invariant under a longitudinal boost, and that histograms of this variable (and of the angle  $\phi$  introduced below) are not distorted by such boosts. In addition, away from the beam direction there is a roughly similar number of particles produced in each unit of rapidity in a collision, *i.e.* the rapidity distribution is only smoothly varying, contrarily to the angle distribution.

It is however more convenient in experiments to use only the beam-momentum angle  $\theta$ , through the **pseudorapidity**  $\eta$ . This angle is estimated with the momentum as  $\cos\theta = p_z/p$ . The pseudorapidity is then defined as

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right) \quad (1.8)$$

Pseudo-rapidity is an approximation of  $y$  that is exact in the ultra-relativistic limit ( $E \gg mc^2$ ). **Forward** and **backward** (for momenta, particles, detectors, etc.) mean at positive  $z$  and rapidity, or at negative  $z$  and rapidity, respectively.

The **azimuthal angle**  $\phi$  around the  $z$ -axis (from the  $x$ -axis direction) is generally the third dimension associated to a track, with the momentum  $p$  and the pseudorapidity  $\eta$ .

A useful quantity is the **pseudo-angle between two particles**, which is their distance evaluated in the  $(\eta, \phi)$  plane. For two particles of angle coordinates  $(\eta_1, \phi_1)$  and  $(\eta_2, \phi_2)$ , it is defined as

$$\Delta R = \sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2} \quad (1.9)$$

where  $(\phi_1 - \phi_2)$  is taken in the interval  $[-\pi, \pi]$ .

### 1.5.3 Luminosity and cross sections

The **cross section**  $\sigma$  of an interaction is the probability of it to happen in a given experiment, converted in units of area via the luminosity of this experiment, defined below. It is a universal property of the input and final particles of the studied process. An easy way to understand the cross section is with a (not so misleading) classical approach: considering the two initial particles as hard spheres of transverse sections equal to the cross section, the interaction will take place if the two hard spheres ‘touch’ in the classical sense.

The **instantaneous luminosity**  $\mathcal{L}$  of a collider (or fixed target) experiment is characterised through the cross section  $\sigma$  of the beam particles (proportional to their probability to interact with each other) and the rate of these interactions (*i.e.* their number per second)  $\mathcal{N}$ :

$$\mathcal{N} = \mathcal{L} \sigma. \quad (1.10)$$

Using the luminosity defined this way, one can then apply Eq. 1.10 with the cross section and rate of any physical process one wants to study. The cross section is a property of the studied particles that one generally quotes as a result, while the properties of the used beams only show in the luminosity. The luminosity can be understood as the flux of initial particles from one beam, taking into account the properties of the other beam or the target it collides with. If the effective transverse section shared by the colliding beams  $A_{\text{shared}}$  is known, then the instantaneous luminosity can be calculated from the number of particles in each bunch  $N$  and the frequency of the bunch crossings ( $f = 40 \text{ MHz}$  for the design LHC bunch separation of 25 ns):

$$\mathcal{L} = \frac{f N^2}{A_{\text{shared}}}. \quad (1.11)$$

The effective shared area depends only on the angle at which the beams cross, and on the beam quality parameters: the **emittance**  $\varepsilon$  and the amplitude **function**  $\beta$ , not detailed further in this thesis.  $\mathcal{L}$  is expressed in units of  $\text{cm}^{-2}\text{s}^{-1}$ , or its equivalent using the inverse barn unit ( $1 \text{ b}^{-1} = 10^{24} \text{ cm}^{-2}$ ). In the following, what will be simply called ‘luminosity’ is actually the **time-integrated luminosity**:

$$L = \int_{\text{run}} \mathcal{L}(t) dt \quad (1.12)$$

which is usually expressed in  $\text{fb}^{-1} = 10^3 \text{ pb}^{-1} = 10^6 \text{ nb}^{-1}$ . It is proportional to the number of collisions accumulated during a given **run** (*i.e.* a period during which the accelerator provides colliding beams).

The **differential cross section** element  $d\sigma_{2 \rightarrow n}$  for a process with 2 initial and  $n$  final particles yields the cross section after being integrated over the 4-momenta of the  $n$  output particles. It can be decomposed in three terms:

$$d\sigma_{2 \rightarrow n} = \frac{|\mathcal{M}|^2}{4F} d\Phi. \quad (1.13)$$

The **matrix element**  $\mathcal{M}$  contains the theoretical knowledge from the Standard Model or other theories of elementary particles. It is what one seeks to extract from the experiment. In the Standard Model, it is calculated from the Lagrangian of section 1.2 – at least in principle, considering that some of those calculations have kept physicists busy for decades! The flux factor  $F = \sqrt{(p.p')^2 - (m m')^2}$  contains the kinematic (relativistic) information about the conditions of the encounter of the two initial particles, of respective 4-momenta  $p$  and  $p'$ , and of masses  $m$  and  $m'$ . Finally,  $d\Phi$  is the Lorentz-invariant **phase space** element, containing the differential 4-momenta elements of the final particles and ensuring the conservation of energy and momentum. When partially integrated over a range of the 4-momenta of the final particles, for instance within the limits of the momenta accessible in a detector, it is simply called ‘(available) phase space’.

#### 1.5.4 Lifetime

In radioactivity, the half lifetime  $\tau_{1/2}$  is the time it takes for half a set of radioactive isotopes to decay into other states, based on the time evolution of the remaining initial isotopes  $N = N_0 (\frac{1}{2})^{-t/\tau_{1/2}}$ . In particle physics one rather uses the average **lifetime** of a particle  $\tau = \frac{\tau_{1/2}}{\ln(2)}$ , such that the number of particles that did not yet decay into other states evolves as

$$N = N_0 e^{-t/\tau}. \quad (1.14)$$

Relating this to the cross section formalism above, the lifetime is actually

$$\tau = \frac{\hbar}{\Gamma_{tot}} \quad (1.15)$$

where the transition rate  $\Gamma_{tot}$  can be seen as the cross section for a  $1 \rightarrow n$  process, or rather the sum of the transition rates  $\Gamma_{1 \rightarrow n}$  for all possible final states of the decays of this particle.

Another point of view on the transition rate [31] is that the energy of the unstable particle is  $E_0 - i\Gamma_{tot}/2$  and its wave function is  $\psi(t) = \psi(0)e^{-it(E_0 - i\Gamma_{tot}/2)}$ . This is justified by computing the resulting probability density  $|\psi(t)|^2 = |\psi(0)|^2 e^{-\Gamma_{tot} t} \propto e^{-t/\tau}$ , which fits to Eq. 1.14 defining the lifetime. Now if a Fourier transform is applied to the wave function to obtain its momentum-space equivalent:

$$\tilde{\psi}(E) = \frac{\psi(0)}{2\pi} \int dt e^{-it(E_0 - i\Gamma_{tot}/2)} = \frac{i\psi(0)}{2\pi} \frac{1}{(E - E_0) + i\Gamma_{tot}/2} \quad (1.16)$$

from which one obtains the BREIT-WIGNER function, giving the probability of finding this particle at energy  $E$ :

$$|\tilde{\psi}(E)|^2 \propto \frac{1}{(E - E_0)^2 + \frac{\Gamma_{tot}^2}{4}}. \quad (1.17)$$

This means that the energy of the decay products (*i.e.* their invariant mass) is spread around the nominal energy with a typical **decay width**  $\Gamma_{tot}$ , inversely proportional to its lifetime.

The spread of mass or energy of a supposedly resonant particle can be linked to the Heisenberg principle. If the particle lives only for a short time  $\sim \tau$ , then the energy is allowed to vary within a range  $\Delta E = \Gamma_{tot}$ , such that  $\Gamma_{tot}\tau \sim \hbar$ . The factor  $\hbar$  is the same than the one I kept explicit in Eq. 1.15. The smallest lifetimes directly detectable with the current detectors are of order  $\sim 10^{-13}$  s, corresponding to distances  $\sim 100 \mu\text{m}$  travelled in the detector, before decaying typically through the weak interaction. The intrinsic width of these particles is hence of order  $10^{-8}$  MeV, completely undetectable experimentally. Even the  $J/\psi$  resonance widely used in this thesis, despite its much smaller lifetime  $\tau_{J/\psi} = 7.1 \times 10^{-21}$  s, has an intrinsic width 0.93 MeV that is smeared by the experimental resolution on the reconstructed invariant mass of its decay products. Strong decays, on the contrary, happen at a time scale of  $\sim \frac{1}{\Lambda_{\text{QCD}}} \sim \frac{1}{200 \text{ MeV}} \sim 10^{-24} - 10^{-23}$  s, meaning the associated widths are often visible (*e.g.* 145 MeV for the  $\rho$  mesons). The lifetime of the weak bosons W and Z is even smaller, of order  $10^{-25}$  s, giving widths in the GeV range that are observed experimentally.

The **branching fraction** of an unstable particle for a decay to some final state is the probability that it will decay in this given final state. It is the ratio of the partial width  $\Gamma_{part}$  (for decaying into this final state) to the total width (for decaying into any final state).

It must be noted that the transition rate is not Lorentz-invariant, so it must be defined in a reference frame. The **proper lifetime**  $\tau_0$  is the lifetime defined in the rest frame of the particle, and is often simply called ‘lifetime’. When secondary vertices need to be distinguished from the PV in a detector, what is of interest is the lifetime in the frame of the detector, in which the particle is boosted by a Lorentz factor  $\gamma = p/m$ , where  $m$  and  $p$  are the particle mass and momentum in the lab frame. So the lifetime in the detector is  $\gamma\tau_0 > \tau_0$ , and the Lorentz boost makes it easier for experiments to distinguish small lifetimes.

### 1.5.5 Types of pp or PbPb collisions

The protons (but all the following concepts are valid for lead nuclei too) in the beams colliding in the LHC pipe (described in section 3.1) can interact in various ways. The less destructive one is the **elastic interaction**, where the particles come close enough to modify their momenta, but however keep their nature and integrity and do not generate other final state particles. At LHC at  $\sqrt{s} = 13$  TeV, about a quarter of the total pp cross section  $\sigma_{tot,pp} = (111 \pm 3)$  mb arises from elastic scatterings [32]. The total pp cross section at  $\sqrt{s} = 5.02$  TeV is about 90 mb. Elastic scatterings lead only to very low momentum transfers between the protons,  $p_T \simeq 0.2$  GeV in average [32]. Therefore, the final state protons have a very small angle with the beam ( $\theta \simeq 0.2/13000 = 1.5 \times 10^{-5}$  rad in average), which are detected only in very forward calorimeters a few hundred meters from the LHC interaction points.

Three quarters of all pp interactions at LHC are **inelastic scatterings** where one or both protons are broken, yielding particles of a different nature in the final state. 25% of these interactions are **diffractive**, which is defined experimentally by observing rapidity zones in the detector that do not contain any particles. This name is an analogy to optics; in the context of pp collisions, it can be understood as elastic collisions with point-like constituents of the protons. A so-called *pomeron* is exchanged between the protons, which then continue close to their initial direction. Zero, one, or the two protons might be dissociated and give showers of particles close to the beam; those cases respectively correspond to elastic, **single-diffractive**, or **double-diffractive** interactions. If two pomerons are exchanged, then they can interact and produce particles (only) at central rapidities, leaving the protons unbroken and close to the beam; these events are called **central-diffractive**.

The nature of the **pomeron** is still very debated (see Ref. [33] for a short summary of the current understanding of pomerons at high energy colliders); what is certain is that it is an exchange of a (maybe-composite) particle that is non-resonant in mass and carries the quantum

numbers of the vacuum, *i.e.* it does not carry a colour charge; this was historically understood as necessary to explain the logarithmically rising hadronic cross sections at high energy. Depending on the scale of the interaction, the pomeron could be understood perturbatively with a colourless set of gluons (close to the hypothesised *gluon ball*), or non-perturbatively with an interacting sea of gluons between the two protons. The colourlessness of the pomeron means that it does not break the coherence of the partonic components of the proton it is emitted from, so that even if the proton is dissociated, its decay products acquire only a small transverse momentum and rapidity gaps are observed in the detector.

The single-diffractive events, featuring high-rapidity particles from only one of the protons, are usually rejected by detecting activity in only one of the forward detectors. The double-diffractive, central-diffractive, and non-diffractive events are those selected by the event-level cuts of most analyses (detailed in section 5.2). The **non-diffractive** events, composing 75% of the inelastic interactions, are those of interest to this work; they involve direct interactions of the partons from the protons. What partons are interacting depends on their probability to exist in the proton or lead nucleus at given momentum fraction  $x$  and energy scale  $Q^2$  (see section 1.3.2 on the PDFs).

An experimental approach to this classification of events is the **minimum-bias** (MB) data. It consists in applying very minimal event trigger requirements (the CMS trigger is explained in section 3.2.4) to obtain a sample containing close to all inelastic hadronic collisions, except the single-diffractive ones. This excludes elastic hadronic interactions and electromagnetic interactions. The associated trigger mostly looks for deposited energy on both sides of the beam.

### 1.5.6 The geometry and observables of heavy ion collisions

Colliding heavy nuclei are not point-like, and the geometric configuration of a collision has important consequences, for example on the size and shape of the QGP volume they create, on the particle multiplicity, and on the transverse energy deposited in the detector.

This calls for the notion of **centrality**: when nuclei (both containing  $A$  nucleons) collide ‘head-on’, meaning when the overlap of the nuclei in the plane transverse to the beam is large, the collision is called **central**. On the contrary, when the overlap of nuclei is small, the collision is **peripheral**. When there is no spatial overlap but the nuclei still interact (typically electromagnetically), the collision is **ultra-peripheral**. The centrality of a collision is quoted as the fraction of the hadronic inelastic AA (nucleus-nucleus) interactions that have equal or larger overlap of the nuclei than this collision; for instance a 10% centrality means a close to full overlap, whereas an 80% centrality means a small geometric overlap. Fig. 1.6 illustrates the densities in the transverse plane for two centrality values, assuming that the nuclei are hard spheres with constant nucleonic density.

Centrality is theoretically associated to the **impact parameter**  $b$ , the distance between the two nucleus centres. However, it is experimentally estimated by assuming that this quantity fully correlates with the multiplicity of produced particles, or to the transverse energy deposited in the detector – mostly at high rapidities, which is why it mostly is evaluated using forward calorimeters (see section 3.2.1). Fig. 1.6 illustrates this relation between the theoretical and experimental definitions of centrality, as the number of participants (correlated with the number of final state particles of the collision) decreases with centrality. Fig. 1.7 shows how centrality is determined in PbPb collisions in the CMS detector, via the fraction of the total number of PbPb collisions, sorted according to the energy they deposited in the forward hadronic calorimeters (HF).

Let us review some important variables associated to the geometry of heavy ion collisions, of which a good review is performed in Ref. [34].  $N_{\text{part}}$  is the number of **nucleons participating** in the collisions (which underwent hadronic interactions with other nucleons), and  $N_{\text{coll}}$  the number

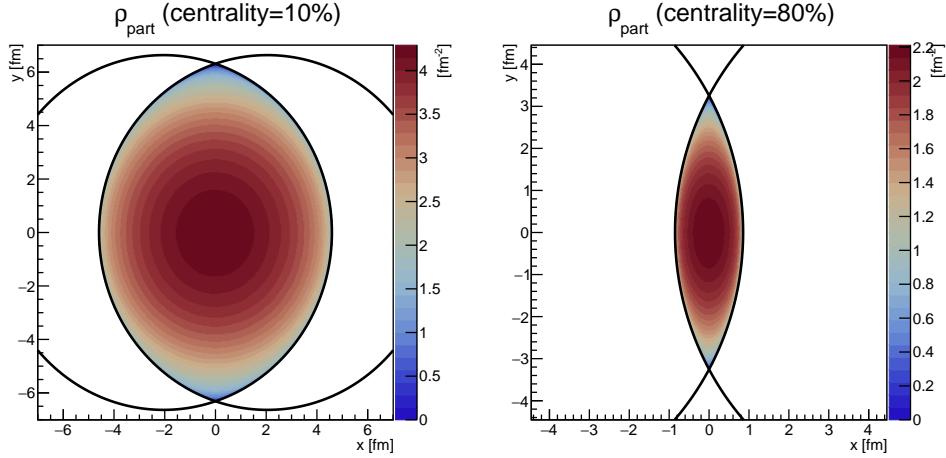


FIGURE 1.6: Number of participant nucleons per unit of transverse area  $\rho_{\text{part}}$ , at centrality 10% (left) or 80% (right), in my implementation of the hard-spheres heavy ion collisions optical model, detailed in section 11.2.3. The black circles show the transverse extent of the nuclei.  $\rho_{\text{part}}$  decreases towards more peripheral collisions, for which the impact parameter  $b$  is large.

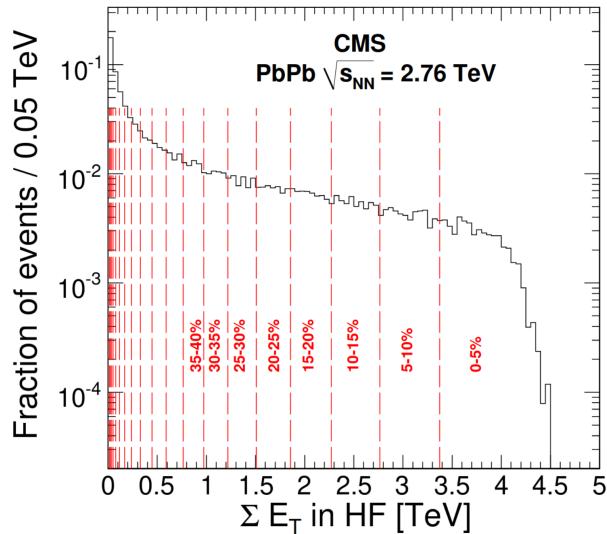


FIGURE 1.7: Determination of the centrality in 2011 PbPb data from CMS, using the sum of transverse energy depositions in the HF.

of **binary nucleon-nucleon (NN) collisions**. The maximal  $N_{\text{part}}$  and  $N_{\text{coll}}$  (at centrality 0%) equal respectively  $2A$  and  $A^2$ .

Another important quantity is the **nuclear thickness**

$$T_A(s) = \int dz \rho_A(s, z) \quad (1.18)$$

giving the number of nucleons per unit area in nucleus A, along the beam direction  $z$ , at a transverse distance  $s$  from the nucleus centre, for a nucleus of nucleonic density  $\rho_A$  (either a constant density or a Woods-Saxons profile, see section 11.2.3). The normalisation of the nuclear thickness is  $\int d^2s T_A(s) = A$ .

Then one can define the **nuclear overlap function** of a AB collision as the convolution of the thicknesses of the two nuclei:

$$T_{AB}(\vec{b}) = \int d^2\vec{s} T_A(\vec{s})T_B(\vec{b} - \vec{s}) \quad (1.19)$$

where  $\vec{s}$  is the transverse position in the first nucleus, and  $\vec{b}$  the transverse separation between the nuclei centres. It is normalised such that  $\int d^2\vec{b} T_{AB}(\vec{b}) = A \times B$ . This can be seen as a nucleon-nucleon (NN) luminosity per PbPb collision, as it is, roughly speaking, the number of NN pairs that *could* collide, per unit area. It results that the average number of NN collisions at impact parameter  $b$  is:

$$N_{\text{coll}}(b) = \sigma_{\text{NN}} T_{AB}(b) \quad (1.20)$$

Knowledge of  $\sigma_{AB}^{\text{geom}}$ , the total hadronic cross section (including elastic interactions) of two nuclei A and B, is important in order to calculate the centrality, and is estimated to  $7.66 \pm 0.03$  b in Ref. [35]. It is related to the NN inelastic cross section  $\sigma_{\text{NN}}$ , estimated to  $67.6 \pm 0.6$  mb (and assumed to be similar for protons and neutrons) in Ref. [35]. The probability of having at least one hadronic NN interaction in a AB collision at impact parameter  $b$  is

$$p_{\text{hadr}} = 1 - \left(1 - \frac{\sigma_{\text{NN}} T_{AB}(b)}{AB}\right)^{AB} \quad (1.21)$$

as the probability of a nucleon from nucleus A to interact with any nucleon from nucleus B is  $\sigma_{\text{NN}} \frac{T_{AB}(b)}{AB}$ . With the approximation  $AB \gg 1$ , it is equivalent to  $1 - e^{-\sigma_{\text{NN}} T_{AB}(b)}$ , so that the **total AB cross section** is

$$\int d^2b \left(1 - e^{-\sigma_{\text{NN}} T_{AB}(b)}\right) \quad (1.22)$$

which is close to the **geometric AB cross section** (but not equal to it, especially at large  $b$  when  $T_{AB}$  is small):

$$\sigma_{AB}^{\text{geom}} = \int d^2b \quad (1.23)$$

**Glauber models** calculate the above variables, making some assumptions on the nuclear density profile and the distribution of nucleons. They can be *optical*, considering analytical averaged densities of nucleons, or based on *Monte Carlo*, simulating many collisions in which the separate nucleons are randomly placed on a grid according to the chosen nuclear profile. The above variables will be more extensively discussed in the last part of this thesis, where I implemented Glauber models with varying assumptions.

Two broad types of processes can be studied in heavy ion collisions. First, soft processes associated to the bulk of low-energy particles whose collective behaviour due to the QGP is studied; the associated quantities typically **scale with  $N_{\text{part}}$** , which is proportional to the density of the QGP. An example of roughly  $N_{\text{part}}$ -scaling quantity is the total multiplicity of charged particles. The second type is the production of hard probes of relatively high energies, which are of most interest to this thesis. They **scale with  $N_{\text{coll}}$** , as each additional NN collision can be seen as an additional probability to produce a given hard probe, neglecting the fact that this nucleon might have interacted (in general in a soft interaction) already in a previous NN collision. This is consistent with the hard probes being generated in interactions of partons rather than of nucleons that kept their integrity.

Therefore, the effect of the QGP on the production of a hard probe  $X$  can be quantified by the **nuclear modification factor  $R_{AA}$**  (considering now a symmetric collision). It is the ratio between the production of  $X$  in AA collisions *per NN collision* and its production in pp collisions, where no QGP is expected to be formed. It can be written in two simple ways, when

integrated over all minimum-bias AA hadronic collisions. First:

$$R_{\text{AA}} = \frac{\mathcal{N}_{\text{AA}}^X / \langle N_{\text{coll}} \rangle_{\text{MB}}}{\mathcal{N}_{\text{pp}}^X} \quad (1.24)$$

where  $\mathcal{N}_{\text{pp}}^X = \sigma_{\text{pp}}^X / \sigma_{\text{NN}}$  is the pp per-event yield of  $X$ , *i.e.* the average number of  $X$  produced per pp collision, and  $\mathcal{N}_{\text{AA}}^X = \sigma_{\text{AA}}^X / \sigma_{\text{AA}}^{\text{geom}}$  is the AA per-event yield of  $X$ . Or  $R_{\text{AA}}$  is expressed as:

$$R_{\text{AA}} = \frac{\sigma_{\text{AA}}^X / A^2}{\sigma_{\text{pp}}^X} \quad (1.25)$$

Eq. 1.24 and 1.25 respectively state that the AA yield should be normalised (to be comparable to the pp ones) by the *effective* number of NN collisions, or that the AA cross section should be normalised by the number of *possible* NN collisions.

Eq. 1.24 and 1.25 relate by

$$\frac{1}{\langle N_{\text{coll}} \rangle_{\text{MB}}} = \frac{\sigma_{\text{AA}}^{\text{geom}} / A^2}{\sigma_{\text{NN}}} \quad (1.26)$$

which is consistent with the integration of Eq. 1.20 over AA collisions of all centralities. The minimum bias nuclear overlap function is indeed:

$$\langle T_{\text{AA}} \rangle_{\text{MB}} = \frac{\int d^2 b T_{\text{AA}}(b)}{\int d^2 b} = \frac{A^2}{\sigma_{\text{AA}}^{\text{geom}}} \quad (1.27)$$

Now, when  $R_{\text{AA}}$  is calculated for a given centrality class  $\mathcal{C}$  (for example 10–20%), associated to impact parameters in the range  $[b_1, b_2]$ , more care is needed. The integration over a class is  $\langle \cdot \rangle_{\mathcal{C}} = \int_{\mathcal{C}(b_1-b_2)} d^2 b$ . Making the per-event yield more explicit, averaged over class  $\mathcal{C}$ :

$$\langle \mathcal{N}_{\text{AA}}^X \rangle_{\mathcal{C}} = \frac{\langle N_{\text{AA}}^X \rangle_{\mathcal{C}}}{\langle N_{\text{AA}} \rangle_{\mathcal{C}}} = \frac{\langle N_{\text{AA}}^X \rangle_{\mathcal{C}}}{w(\mathcal{C}) \times N_{\text{MB,AA}}} \quad (1.28)$$

where  $\langle N_{\text{AA}}^X \rangle_{\mathcal{C}}$  is the raw number of  $X$  probes observed in an experiment,  $w(\mathcal{C})$  is the width of the centrality range (0.1 for  $\mathcal{C} = 10 - 20\%$ ), and  $N_{\text{MB,AA}}$  is the minimum-bias number of all (non-single-diffractive) hadronic AA collisions counted in the same experiment.

In centrality class  $\mathcal{C}$ , the equivalents of  $\sigma_{\text{AA}}^{\text{geom}}$  and of the nuclear overlap function in Eq. 1.27 are:

$$\langle \sigma_{\text{AA}}^{\text{geom}} \rangle_{\mathcal{C}} = \left( \int_{b_1}^{b_2} d^2 b \right) \frac{A^2}{\int_{b_1}^{b_2} d^2 b T_{\text{AA}}(b)} = \frac{A^2}{\langle T_{\text{AA}} \rangle_{\mathcal{C}}} \quad (1.29)$$

Then, in Eq. 1.25, one can replace

$$\langle \sigma_{\text{AA}}^X \rangle_{\mathcal{C}} / A^2 = \langle \mathcal{N}_{\text{AA}}^X \rangle_{\mathcal{C}} \times \langle \sigma_{\text{AA}}^{\text{geom}} \rangle_{\mathcal{C}} / A^2 = \langle \mathcal{N}_{\text{AA}}^X \rangle_{\mathcal{C}} / \langle T_{\text{AA}} \rangle_{\mathcal{C}} \quad (1.30)$$

which gives the  $R_{\text{AA}}$  in the main form used in this thesis:

$$\langle R_{\text{AA}} \rangle_{\mathcal{C}} = \frac{\frac{1}{w(\mathcal{C}) \times N_{\text{MB,AA}} \times \langle T_{\text{AA}} \rangle_{\mathcal{C}}} \langle \mathcal{N}_{\text{AA}}^X \rangle_{\mathcal{C}}}{\sigma_{\text{pp}}^X} \quad (1.31)$$

The yields and cross sections  $N^X$  and  $\sigma^X$  can be differential, often in  $p_T$  and rapidity.

## Part I

# The $B_c^+$ meson in PbPb and pp collisions



“Three quarks for Muster Mark!  
 Sure he hasn’t got much of a bark  
 And sure any he has it’s all beside the mark.  
 [...]”  
*Hohohoho, moulty Mark!*  
*You’re the rummest old rooster ever flopped out of a Noah’s ark*  
*And you think you’re cock of the wark.”*

---

James JOYCE, *Finnegans Wake*, Book II, Chapter 4.<sup>1</sup>

## Chapter 2

# Why? Heavy quarks in the QGP

Heavy ion collisions create a hot and dense matter, the quark-gluon plasma, in which colour confinement fades down so that partons behave as free particles – for a few femtometers. No external tool is at our disposal, considering the space-time scales over which it evolves, to seize the produced QGP before ‘looking’ at it. Therefore, we need to measure its properties with internal probes, produced and evolving on similar time scales. This chapter exposes how heavy quarks (*i.e.* charm and beauty), and the hadrons they form, are excellent probes of the formation, evolution, and other properties of the QGP.

Section 2.1 justifies why heavy quarks, and in particular heavy quarkonia, have been privileged probes of the QGP since its early studies. Section 2.2 presents properties and some past measurements of the  $B_c$  states. The subsequent sections motivate the experimental observation of  $B_c$  mesons in heavy ion collisions achieved in this thesis. They show the main effects of the QGP on heavy quarks and their bound states, and what light the  $B_c$  meson can shed on these topics: the energy loss of heavy quarks (section 2.3), the dissociation of quarkonia through the colour screening of their potential (section 2.4), and last but not least, the recombination of uncorrelated heavy quarks enhancing the production of heavy quark bound states (section 2.5).

### 2.1 Heavy quarks and quarkonia

Heavy quarks are interesting probes of the QGP, because in general, they exist through the whole history of the QGP. They are indeed mostly produced at time scales  $1/m_Q$  (where

---

<sup>1</sup>In this satiric novel mocking King Mark, the cuckolded husband in the Tristan legend, James Joyce uses many neologisms and puns that evoke the cries that birds throw at Mark. Murray GELL-MAN, when he thought of three fractionally charged nuclear constituents to explain the hadron zoo emerging at the time [36]. He initially named them ‘quarks’ with an ‘o’ sound, but he probably did not write this name down yet, as he testified that this was the way he called peculiar things/objects in his daily life. He realised only a few months later that this novel mentioning ‘three quarks’ was spot-on to describe this triplet of awkward fundamental constituents of matter.

$m_Q = m_b \sim 5 \text{ GeV}$  or  $m_c \sim 1.5 \text{ GeV}$ ), whereas the QGP is formed at time scales  $0.2 - 1 \text{ fm}$  (corresponding to energy scales from  $0.2 \text{ GeV}$  to  $1 \text{ GeV}$ ). As explained later, their production in jet fragmentation can however blur this simple time scale ordering. The bound states formed by heavy quarks, when they survive the interaction with the medium, decay much later than the cool-down of the QGP, on time scales of weak decays ( $100 - 500 \mu\text{m}$ ) for open heavy flavour mesons, and about  $7 \times 10^{-21} \text{ s} \sim 1/(10 \text{ MeV})$  for charmonium and bottomonium decays.

Heavy quarks can bind with light, strange, or charm quarks to form *open heavy flavour* hadrons (such as D and B mesons). Conversely, the quark colour can be neutralised by binding a quark and an antiquark of the same flavour, resulting in *quarkonia*: either *charmonia* composed of  $c\bar{c}$  quarks (e.g. the  $J/\psi$  meson, and its radially excited state  $\psi(2S)$ ), or *bottomonia* composed of  $b\bar{b}$  quarks (e.g. the  $\Upsilon(1S)$  meson, and the excited states  $\Upsilon(2S)$  and  $\Upsilon(3S)$ ).

Heavy quarkonia have been much studied because of the natural separation of scales they provide: the relative velocity  $v$  between the two heavy quarks is small due to the high bound state mass  $m$  (as the additional mass in the excited states is small compared to the ground state mass), hence the hierarchy  $m \gg mv \gg mv^2$  on which the *non-relativistic QCD* (NRQCD) theory is based [37]. The potential describing the interaction of the two quarks can be expressed through an effective field theory: the hard scales are integrated out ( $m$  for NRQCD, or even  $mv$  for potential-NRQCD [pNRQCD]), and an approximate potential can be obtained by matching the lowest-order terms of a  $1/m$  or  $1/mv$  power expansion to measurements. This is an advantage compared to the full QCD theory, in which hadronisation is a non-perturbative process, hence hard to calculate. NRQCD also includes all possible  $c\bar{c}$  colour states, contrarily to, e.g., the colour-singlet model [38]. However, a still-standing difficulty is for NRQCD (even at NLO) and other models to describe both the  $p_T$  spectrum and the close-to-null polarisation of  $J/\psi$  mesons [39–41], as shows Fig. 2.1.

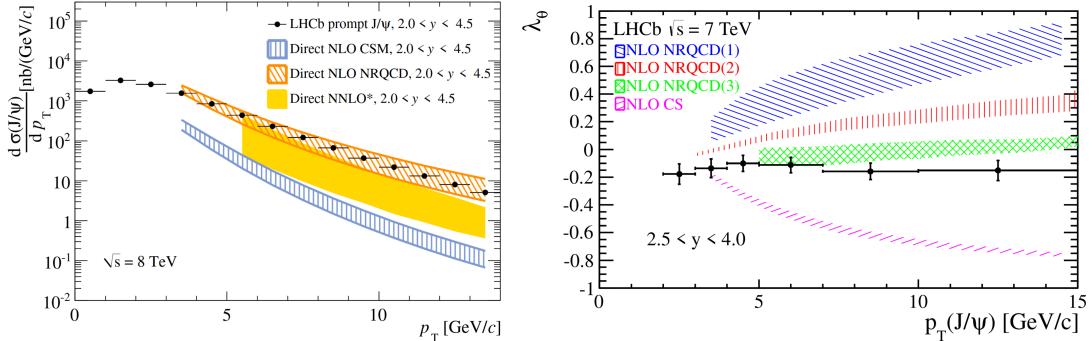


FIGURE 2.1: Comparisons of NRQCD and colour-singlet models to LHCb pp collision data. *Left:*  $p_T$  spectrum (from Ref. [40]). *Right:* polarisation (from Ref. [41]).

Despite these difficulties, quarkonia are of great interest as probes of the QGP. They interact with the medium via the strong force, possibly resulting in dissociation, one of the first proposed signatures of the QGP (see section 2.4). In the NRQCD framework, the modification of the heavy quark potential due to the presence of QGP can be quantified [42], including its imaginary part contributing to shorten the lifetime of the state in the medium.

The limitation of the NRQCD separation of scales in the context of QGP is that some quarkonia, or the heavy quarks they are composed of, might be produced at larger times than that of the formation of QGP. In that case, the parton parents of the quarkonium can be affected by the medium, and the effective time that the quarkonia spend in the QGP is reduced. The heavy quarks and the quarkonia they might form are produced on proper time scales  $\frac{1}{m}$  and  $\frac{1}{mv}$ , respectively. This scale can be dilated and exceed that of QGP formation when the partons are very boosted in the QGP frame: the typical quarkonium (or heavy quark) production scale

must then be  $\frac{p_T}{m} \times \frac{1}{mv} = \frac{p_T}{m^2 v}$  (or  $\frac{p_T}{m^2}$ ). This delay can also take place when the heavy quarks are produced later than the initial hard scattering, during the fragmentation of parton showers. CMS recently observed a high fraction of  $J/\psi$  mesons being generated in jets [43], suggesting an important role of fragmentation in the production of  $J/\psi$  mesons, that is not fully considered in the current simulations (notably PYTHIA). Therefore, high- $p_T$  quarkonia might not all be generated earlier than the timescale of the QGP evolution.

As a consequence, part of the very high- $p_T$   $J/\psi$  mesons undergo effects of the QGP while they still are the gluons (or simply colour-octet states) they originate from. They then endure medium-induced energy losses similar to the light hadrons, which mostly originate from gluons. Heavy quarks also endure energy loss, but with a different colour factor, and with a possible dependence on their mass. This energy loss reflects scattering properties of the QGP, and causes the *jet quenching* phenomenon observed experimentally; it is a historical signature of QGP, proposed by BJORKEN in 1982 [44]. The energy loss of single partons is discussed in section 2.3, and is the context of the phenomenological studies of the second part of this thesis.

Lastly, the b and c quarks are too heavy to be thermally produced in the medium, so their total multiplicity is roughly conserved over the lifetime of the QGP in a heavy ion collision (considering the above caveat about fragmentation and boosted production). The relative proportions of the hadrons they form are however affected by the medium. The open heavy flavour (D and B) mesons constitute the majority of the hadrons produced by heavy quarks, so their total production should be roughly conserved as is the number of heavy quarks. This is noticeable in the  $R_{AA}$  of D mesons, which rises above 1 towards null  $p_T$ , to compensate for the suppression at mid- and high- $p_T$  [45] – though the shadowing from cold nuclear matter, not discussed here, could modify the integrated  $R_{AA}$ . On the contrary, the low rate of heavy quarks that find another heavy quark to bind with can be modified by the QGP. Beyond suppression effects, the production can also be enhanced, thanks to heavy quarks roaming freely and recombining in the medium, as explained in section 2.5. This signature of the QGP can only be observed in bound states of multiple heavy quarks, benefiting from the abundance of heavy quarks in the QGP that contrasts with their rareness in pp collisions. A sole heavy quark can hadronise with a light quark to form open heavy flavour mesons; as light quarks are easily produced both in the QGP or in the vacuum, these states cannot probe recombination effects.

The effects of QGP on heavy quark bound states are sketched in Fig. 2.2. In this context, the  $B_c$  meson, a hybrid charm-beauty state that shares properties both with quarkonia and with open heavy flavour, is a connecting piece between the physics of charmonia, bottomonia, open charm and open beauty hadrons. Before demonstrating it can shed a new light on the above QGP phenomena, let us explain some of its properties and experimental history.

## 2.2 $B_c$ meson properties and past measurements

Let us review here some properties of the  $B_c$  meson states, along with the main observations and measurements that were performed on these states. Some measurements referenced here will be used in other sections.

### 2.2.1 Production and spectroscopy

The  $B_c^+$  meson contains a  $\bar{b}$  and a c quarks, making it (and its excited states) the only hadron observed so far that contains heavy quarks of two different flavours. Baryons containing net bottom and charm quantum numbers, *e.g.*  $\Xi_{bc}$  baryons, have indeed not been observed yet. The  $B_c$  meson has a net flavour number, which in principle excludes it from the quarkonium family. However, one could support that the proximity of its quark content to the charmonium and bottomonium states ('hidden' heavy flavour mesons) makes it an exotic heavy quarkonium.

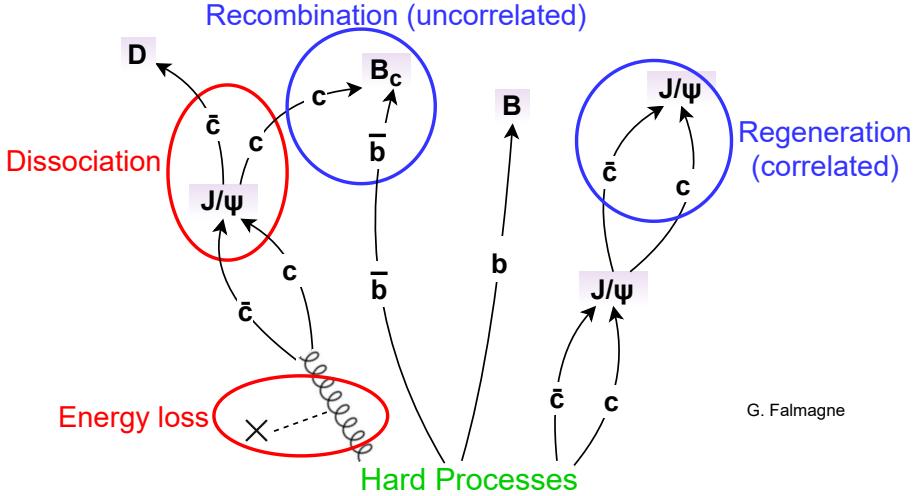


FIGURE 2.2: Mechanisms affecting charmonium and  $B_c$  production in the presence of QGP. The parton parents of the hadrons undergo medium-induced energy loss. The screening of the colour potential can dissociate or prevent the formation of the  $c\bar{c}$  or  $c\bar{b}$  bound states. Uncorrelated  $c$  and  $\bar{c}$  or  $\bar{b}$  can also bind after travelling in the deconfined medium, both if they come from different nucleon-nucleon scatterings or from a previously dissociated state.

The current estimation of the  $B_c$  ground state mass is  $m_{B_c}^{\text{PDG}} = 6274.9 \pm 0.8 \text{ MeV}$  [46], making it heavier than any other open heavy flavour meson.

The  $B_c$  meson has a particularly low **production** cross-section in pp collisions, because two heavy quark pairs ( $b\bar{b}$  and  $c\bar{c}$ ) need to be produced in the same hard-process. This means that additional production mechanisms in other conditions could easily dominate over this direct production, such as the possible quark recombination in heavy ion collisions (see section 2.5). In hadronic collisions, the dominant production mode of the  $B_c$  meson is a gluon fusion producing the two heavy quark pairs:  $gg \rightarrow B_c^+ + b + \bar{c}$ .

The  $B_c$  family has a rich **spectroscopy** of excited states. Fig. 2.3 shows the predictions of Ref. [47] of the masses of the various excited states, using the spectroscopic notation  $n^{2S+1}L_J$ .  $S$  is the spin, equal to 0 in the ground state, and 1 in some excited states.  $L$  is the azimuthal quantum number, indicating in principle the shape of the wave functions;  $L = 0, 1, 2$ , or 3 states are designated as  $S$ -wave (symmetric),  $P$ ,  $D$ , or  $F$ -wave functions.  $J$  is the total angular momentum, *i.e.* the vector sum of quantum numbers  $L$  and  $S$ . In this case where  $S = 0$  or 1,  $J$  can take any positive value among  $L - 1$ ,  $L$ , and  $L + 1$ . On Fig. 2.3, the principal quantum number  $n$  (indicating the radial excitation of the system) is kept implicit: it limits from above the values of  $L \leq n - 1$ , and higher  $n$  states have higher masses. In particular, the  $P$ -wave states of lowest mass have  $n = 2$ .

The first ( $n = 1$ ) excited state  ${}^3S_1$  decays with almost 100% branching fraction into the ground state and a photon of about 50 MeV energy that takes away its  $S = 1$  spin. Experimentally, this state directly contributes to the observed ground state cross section, as the low-energy photon is not reconstructed and barely affects the kinematic properties of its ground-state daughter. All other excited states feature decays to the ground-state  $B_c$  and other particles, with branching fractions that are in general unknown. These unknown **feed-down** fractions from the excited states to the observed ground state are problematic for all quarkonia studies, as an observed modification (*e.g.* from passing through QGP) of the ground state can reflect the modification of the ground state *and* of the excited states that decay into it. In particular, when interpreting the modification of the observed ground-state  $B_c$  using theoretical predictions of the effect of QGP on  $B_c$ , this effect should be applied to all ground and excited states.

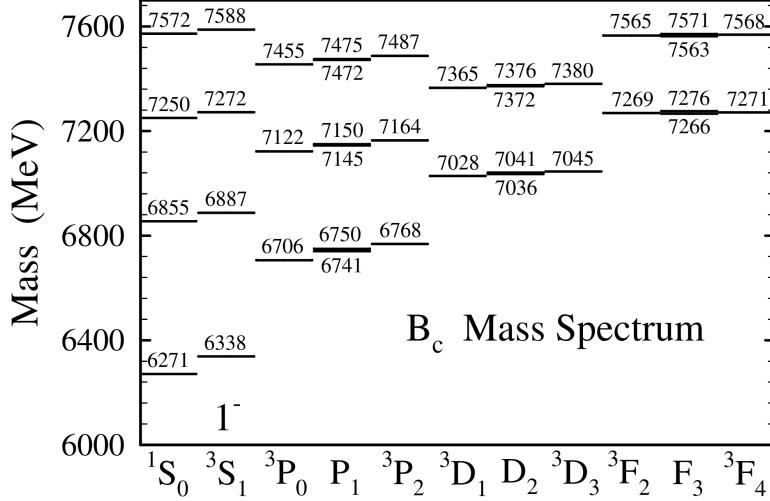


FIGURE 2.3: Spectroscopy of the  $\bar{b}c$  states, from Ref. [47]. The mass threshold for  $D + B$  meson production is at 7140 MeV.

As other quarkonia, the  $B_c$  is composed of a quark and an anti-quark, meaning its QCD colour state must be determined from summing a **3** (quark  $SU(3)_c$ ) group with a  **$\bar{3}$**  (anti-quark  $SU(3)_c$ ) group. This results in a group containing **1** element, the colour-singlet state (in which the colours of the two quarks cancel each other), or **8** elements, the colour-octet states. The colour-singlet states need partons whose colour charges cancel. Colour-octet states must emit a gluon to get rid of their colour charge before propagating at large distances (in the vacuum).

### 2.2.2 Past measurements

The  $B_c$  meson has been first detected in 1998 by the CDF experiment at TeVatron, in proton-antiproton collisions at  $\sqrt{s} = 1.8$  TeV [48], through the (semi-)leptonic channel  $B_c^+ \rightarrow J/\psi 1^+ \nu_1$  (see Fig. 2.4, left) – the one studied in this thesis. The channel  $B_c^+ \rightarrow J/\psi \pi^+$  (named ‘hadronic channel’ in this document) was then observed by CDF (4 $\sigma$  evidence in 2005 [49]) and D0 (in 2006 [50]) Collaborations. The latter channel is preferentially studied when high luminosities are available, as all decay products can be reconstructed by the detectors.

The  $B_c$  meson has then been studied in pp collisions at the LHC, especially by the LHCb Collaboration: it observed in 2012 the hadronic channel [51], and measured in 2014 its ratio to the leptonic channel [52], both at  $\sqrt{s} = 7$  TeV. The latter branching fraction ratio was extensively used in preliminary studies for this work (see sections 3.3.1.1 and 3.5.4.2). CMS also contributed, notably comparing at  $\sqrt{s} = 7$  TeV the hadronic channel to the  $B_c^\pm \rightarrow J/\psi \pi^\pm \pi^\pm \pi^\mp$  branching fraction, and to the  $B_c^\pm \rightarrow J/\psi K^\pm$  branching fraction times cross section [53]. The latter  $B_c^+/B_c^-$  ratio was also measured by the ATLAS [54] and LHCb [55] Collaborations at  $\sqrt{s} = 8$  TeV.

Many other decay channels have been observed in the LHCb detector, so only a few are cited here. The  $B_c^+ \rightarrow B_s^0 \pi^+$  decay has been observed already in 2013 with 7 and 8 TeV data [56]. More recently, the  $\chi_{c0} \pi^+$  [57] and  $D^0 K^+$  [58] final states have also been studied.  $B_c$  meson studies also contribute to look for processes beyond the Standard Model, notably with the ratio of the  $B_c^+ \rightarrow J/\psi \tau^+ \bar{\nu}_\tau$  to  $B_c^+ \rightarrow J/\psi \mu^+ \nu_\mu$  branching fractions measured by LHCb in 2018 [59]. As a side-comment, a discrepancy between the measurement of this ratio and its Standard Model expectation would indicate a violation of the lepton flavour universality, commanding that the tau and the muon in the two decays should have the same interactions with the Standard Model particles. The 2 $\sigma$  discrepancy found by LHCb adds to the other promising hints of lepton flavour universality violation recently observed by the BELLE and LHCb Collaborations.

LHCb also has measured  $B_c$  meson properties, such as its lifetime in 8 TeV data [60]. The world-average value of the  $B_c$  lifetime is  $c\tau_{B_c} = (153 \pm 3)\mu\text{m}$  [46], which is about three times smaller than that of other  $B$  mesons. This small lifetime will make the selection challenging in this analysis (chapter 5). The  $B_c$  charge production asymmetry has also been measured in 2019 from LHCb 7 and 13 TeV data, and it is found consistent with 0 within a few percent uncertainty [61].

All the above measurements concern the ‘ground’ states  $^1S_0$  and  $^3S_1$ , which are not distinguished experimentally. However,  $n = 2$  states were also measured. First, ATLAS observed in 2014, from 7 and 8 TeV data, a peak comprising the two  $2S$  states ( $2^1S_0$  and  $2^3S_1$ ), but did not distinguish their separate mass peaks [62]. Then, in 2019, CMS [63] (and LHCb [64], but with much lower yields) achieved, with 13 TeV data, the separation of the  $2^1S_0$  and  $2^3S_1$  states, respectively decaying to  $B_c(1^1S_0)\pi^+\pi^-$  and  $B_c(3^3S_1)\pi^+\pi^-$ . As the mass difference between  $^1S_0$  and  $^3S_1$  states is about 30 MeV larger for the  $n = 1$  than for the  $n = 2$  states, and as  $1S_0$  and  $3S_1$  are not experimentally distinguished, the  $B_c(1S)\pi^+\pi^-$  invariant mass shows two excited  $B_c$  peaks separated by 30 MeV (shown in Fig. 2.4, right), the lower peak corresponding to the heavier  $2^3S_1$  state. These measurements owe to an impressive mass resolution, partly due to the subtraction of the invariant mass of the  $B_c$  decay products and the associated experimental resolution, but also to the excellent performance of the pion and muon reconstruction of these detectors. CMS also measured more recently the ratio of cross sections of the two excited states [65].

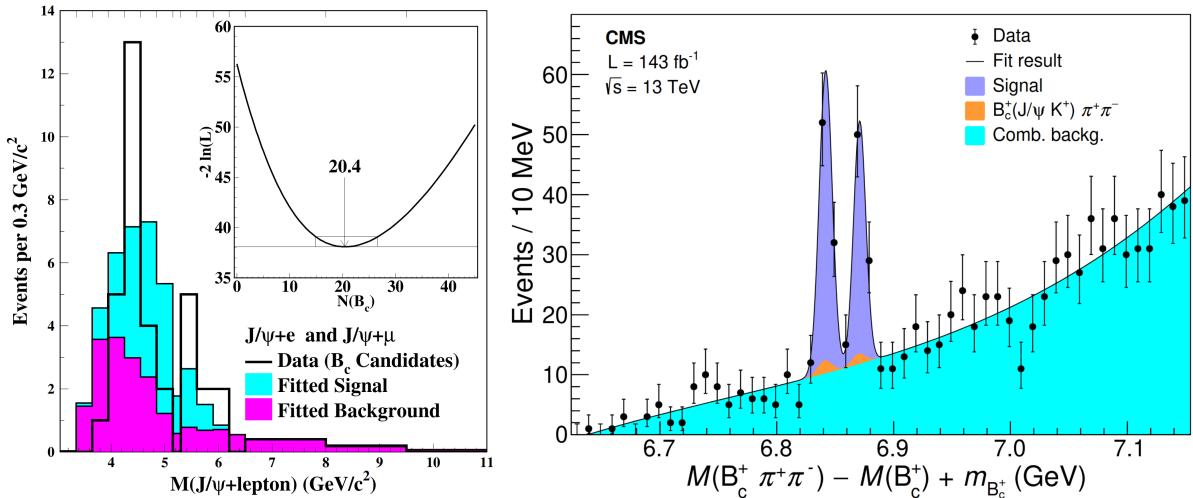


FIGURE 2.4: From the first to the most recent  $B_c$  experimental discovery. *Left:*  $B_c$  discovery with CDF in 1998 [48]. *Right:* Discovery of two  $B_c(2S)$  states with CMS in 2019 [63].

Fig. 2.4 is a snapshot of the journey between the first observation of the ground state with CDF, and the observation of two  $2S$  states by CMS. This thesis intends to make the next step of this experimental journey, namely the observation of (ground-state)  $B_c$  mesons in the much more challenging environment of heavy ion collisions. In this context, new  $B_c$  production (and dissociation) modes may arise, bringing new perspectives on the dynamics and hadronisation of heavy quarks in a QCD medium.

## 2.3 Energy loss

Quarks lose energy when they traverse the QGP, so large- $p_T$  hadrons appear suppressed by the QGP – they still leave the medium, but with a lower momentum. **Jet quenching** designates how this affects the production of jets [44]; it was first observed through the imbalance of the two jets in dijet events [66], when one jet is produced close to the medium boundary, and the other

crosses the whole QGP, hence losing significant energy. I rather concentrate here on the energy loss measured on single hadrons. At low- $p_T$ , there can be **collisional** energy loss, meaning the probe parton collides elastically with other partons in the medium. This is however negligible towards higher  $p_T$  (see its calculation for high- $p_T$  heavy quarks in Ref. [67]) compared to the **radiative** energy loss, in which partons emit gluons when traversing the QGP or of a parton shower in the vacuum, similarly as the Bremsstrahlung of electrons when accelerated. The second part of this thesis will develop on a model of radiative energy loss at high  $p_T$ .

The emission of gluons from a quark is thought to be lesser along the (downstream) flight direction for high quark masses. The original **dead-cone effect** [68] claimed that there is no gluon emission in a forward cone of angular size  $m/E$  around the direction of a quark of mass  $m$ . However this concerns the emission in the vacuum, whereas medium-induced radiation was found to partially fill this cone [69]. In practice, this cone is filled by the decay products of the heavy hadron, making the experimental observation of the angular dependence of radiation very challenging. The sheer existence of suppressed radiation for heavy quarks was also questioned, as quantum oscillations in a finite-size medium might cause an opposite effect [70]. Despite these caveats, a partial dead-cone effect could manifest as a lower suppression (integrated over angles) of heavy flavour hadrons compared to light hadrons [71].

The experimental evidence of the mass dependence of energy loss in heavy ion collisions is still ambiguous. As the dead-cone width decreases with  $p_T$ , its effect should be larger at low- and mid- $p_T$ , but collisional energy loss and non-perturbative hadronisation effects (such as recombination, see 2.5) also play a role in this regime. A relatively small effect is expected for D mesons in the range  $10 \lesssim p_T \lesssim 20$  GeV at LHC [71]. It should be more important for the heavier b quark, but B mesons are either measured with the limited statistics of exclusive decays, or inclusively through non-prompt D or J/ $\psi$  mesons. The latter implies that the kinematic distributions are smeared by the unreconstructed B decay products: the suppression of non-prompt J/ $\psi$  mesons at a given  $p_T$  does not represent the suppression of their B parents at this exact  $p_T$ .

The slope of the initial production momentum spectrum also affects the comparison of the suppression of heavy and light hadrons: for the same absolute energy loss, particles with a steeper (*i.e.* softer) spectrum will show more  $p_T$ -dependent suppression (at mid- to high- $p_T$ ), as the sketch of Fig. 2.5 illustrates. The light quark spectra are softer than the heavy quark ones, so this can enhance the ratio of heavy to light hadrons, as the dead-cone effect. The colour charge of the parent parton of the observed hadron also changes this heavy-to-light ratio: gluons (the dominant parent partons of light hadrons) interact more with the QCD medium than quarks (dominant parents of heavy hadrons), leading again to less suppression for heavy hadrons. One effect compensating the above effects and the dead-cone could come from the fragmentation fraction of heavy quarks to heavy hadrons being higher than that of gluons or light quarks to light hadrons; this fraction multiplies the partonic energy loss, causing more observed suppression for the final-state heavy hadrons [72].

Disentangling these contributions to extract the **dependence of energy loss on the colour charge and the mass** of the partons traversing the QGP is challenging. However, comparing the modification of open heavy flavour hadrons to that of light hadrons stays the best tool to evaluate experimentally those dependences – though new ideas emerge to probe angularly the dead-cone effect using jet substructure, even in heavy ion collisions, *e.g.* in Ref. [73].

Figure 2.6 shows a summary of CMS measurements, using 2015 PbPb data, of the nuclear modification factor of B mesons and hadrons they decay into compared to that of light hadrons. For the reasons evoked earlier, no strong conclusions on the mass dependence of energy loss can be drawn. The production of  $B_s$  mesons in PbPb collisions has also been measured [74, 75], but is excluded from this discussion, because the strangeness enhancement possibly contributing to its production blurs the energy loss picture (and because the uncertainties are larger).

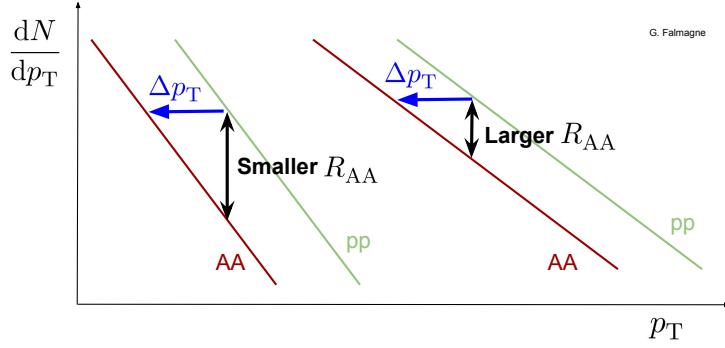


FIGURE 2.5: Harder  $p_T$  spectra lead to larger  $p_T$ -dependent nuclear modification factors, for the same energy loss  $\Delta p_T$ . Both axes are in logarithmic scales (assuming power law spectra).

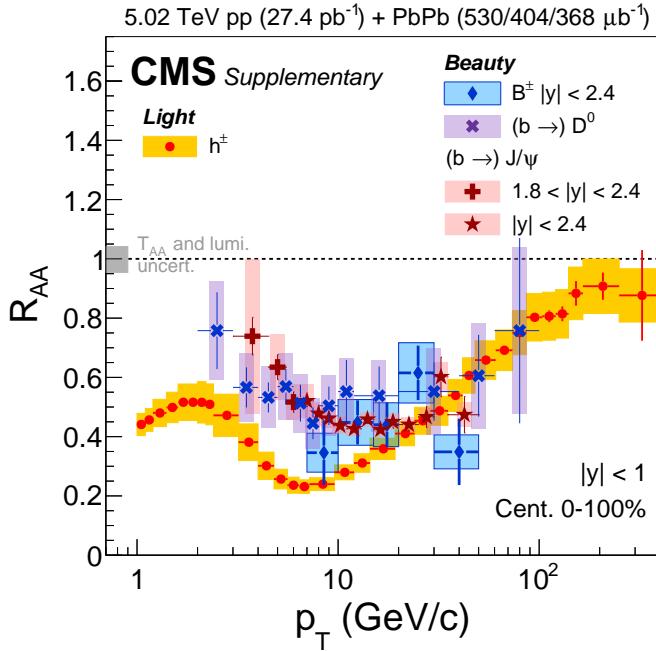


FIGURE 2.6: Nuclear modification factors of  $B^\pm$  [76], non-prompt  $D^0$  [77], and non-prompt  $J/\psi$  [78] mesons, compared to that of charged hadrons [79]. All shown  $R_{\text{PbPb}}$  are measured by CMS as a function of  $p_T$ , using 2015 data. The mass ordering of energy loss should appear at low-mid  $p_T$ , where the mass is relatively more important. The kinematics of non-prompt  $J/\psi$  and  $D^0$  mesons are not exactly those of their  $B$  meson parents.

The  $B_c$  meson studied at mid- or high- $p_T$  (where the energy loss starts to dominate the modification, but the dead-cone effect is still significant) can provide a new perspective on this mass dependence, as it might cumulate the dead-cone effects from both its charm and beauty component quarks. At high- $p_T$ , the  $B_c$  mesons should however mostly be produced through the fragmentation of a  $b$  quark (radiating a  $\bar{c}c$  pair at late times). This is to be compared to the  $J/\psi$  mesons that, in this range, are mostly produced by gluon fragmentation, and to other open heavy flavour hadrons, that are mostly produced by heavy quark fragmentation. The  $B_c$  meson might therefore also probe the colour charge dependence of energy loss, though this would still need a better understanding of the heavy flavour production mechanisms. Finally, at very high- $p_T$  (not accessible yet, considering the size of the available datasets), the  $B_c$  meson can be an additional probe of the universal  $p_T$  dependence of radiative energy loss (see chapter 10).

## 2.4 Colour screening

The QCD potential between two quarks at distance  $r$  (with a non-relativistic relative speed) is schematically, in the vacuum:

$$V = -\alpha_S C_F \frac{1}{r} + \sigma r \quad (2.1)$$

where only the first term matters at short distances, as the second term encodes the colour confinement at large distances through a tension  $\sigma$  of the colour strings. In the QGP, this colour potential is *screened* exponentially, such that:

$$V(r) = -\alpha_S C_F \frac{e^{-m_D r}}{r} \quad (2.2)$$

where  $m_D$  is the Debye mass. This means the attracting potential between two quarks at distance  $r$  decreases rapidly when  $r > \frac{1}{m_D} \equiv r_D$ , the Debye radius. In an electromagnetic plasma, the mobile charges (electrons) are repelled by a given probe electron, creating an effective region of opposite charge around this probe that partially cancels the probe charge. This probe has a significant interaction potential only with charges that are closer than the Debye radius:

$$r_D = 1/\sqrt{\beta n e^2} \quad (2.3)$$

where  $\beta = 1/(k_B T)$ ,  $e$  is the electron charge, and  $n$  is the number density of electrons (originally in Ref. [80] from HÜCKEL and DEBYE).

Similarly, in the QGP, the free partons screen the colour potential between heavy quarks with larger separation than  $r_D$ , in the non-relativistic approximation valid for heavy quarkonia (see section 2.1). The quarkonia of large-enough size hence dissociate into two open flavour mesons in the QGP, making their production directly suppressed; this is a historical signature of the QGP, proposed by MATSUI and SATZ [81], predicting lower quarkonia yields than in the vacuum. Let us find the equivalent of Eq. 2.3 in a plasma of colour charges. First, in analogy to an ideal gas, the number density in the QGP is  $n \propto P/T$ . Then, by conservation of the energy-momentum tensor, and considering the number of degrees of freedom, the pressure  $P$  is proportional to the energy density  $\varepsilon$  ( $P = \varepsilon/3$  in an ultra-relativistic gas, which is not exactly the case of QGP). By analogy to the STEFAN-BOLTZMANN law stating that the energy emitted by a thermal body per unit of its area and per time unit is proportional to  $T^4$ , one finds  $\varepsilon \propto T^4$  (where the three space dimensions in the denominator of  $\varepsilon$  are analogous to the area and time of the blackbody law). This results in

$$n \propto T^3 \propto \varepsilon^{\frac{3}{4}}. \quad (2.4)$$

Now, from Eq. 2.3 and using the strong constant  $g_S = \sqrt{4\pi\alpha_S}$ , analogous to the electric charge, the Debye radius for a medium containing free color charges is:

$$r_D = \frac{1}{m_D} \propto \frac{1}{g_S T}. \quad (2.5)$$

Therefore, when the temperature of the medium rises, the Debye radius decreases, and states of smaller size start to dissociate. This *sequential suppression* of the excited states of heavy quarkonia has long been considered to provide a ‘thermometer’ of the QGP, in so that the most excited states are larger and hence dissociate at a smaller temperature than the less excited states. Considering the progressive dilution of the medium, this also means that the most excited states need to be produced further from the centre of the medium not to be dissociated. Sequential suppression has been observed in bottomonium states, first by CMS [82, 83]:  $\Upsilon(nS)$  states are less suppressed for higher  $n$ . Fig. 2.7 shows the mass fits of the first two  $\Upsilon(nS)$  states in PbPb

2015 CMS data, where  $\Upsilon(2S)$  is more suppressed than  $\Upsilon(1S)$  compared to pp collisions. Let us here underline the excellent muon momentum resolution of CMS (giving the good separation of mass states in Fig. 2.7), important for this analysis and detailed in section 3.2.3. At the time of writing, a CMS analysis aims at observing the  $\Upsilon(3S)$  state. It is already known to be very suppressed ( $R_{AA} < 0.2$ ), but if it is measured to be non-zero whereas its loose binding should make it dissociate completely in the medium, it could elucidate the importance of the medium geometry (as states produced in the corona of the medium and directed away from it would not be dissociated) and of potential beauty quark recombination (see section 2.5).

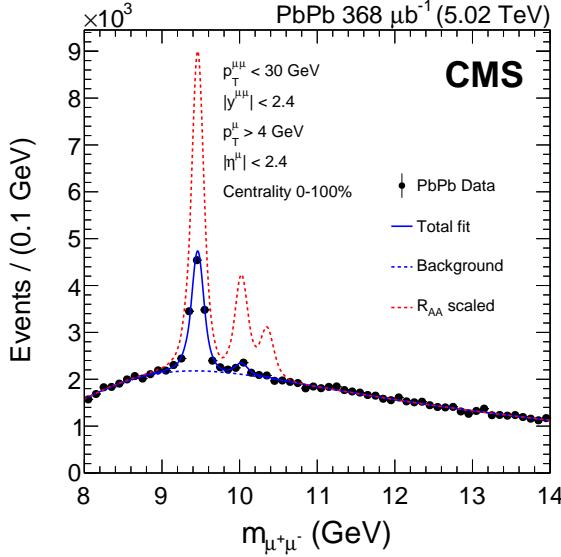


FIGURE 2.7: Dimuon mass fit of the  $\Upsilon(nS)$  states in 2015 CMS data in PbPb collisions, compared to the yields expected if  $R_{AA} = 1$ . Figure from Ref. [83].

The historical Debye screening approach however neglects the dynamics of the medium, as it considers the heavy quarks to be static. Many more approaches have since included dynamical processes [84]. When solving, on the lattice, a Schrödinger-like equation with the quarkonium potential, an imaginary part of the potential of Eq. 2.2 appears. The imaginary potential represents the Landau damping of gauge fields, according to the hard thermal loop formalism [85], and leads to a thermal increase of the width of the quarkonium peak. Gluo-dissociation (the dissociation of a quarkonium by absorption of a soft thermal gluon) also increases the quarkonium thermal width, and can be calculated *e.g.* in a pNRQCD framework [86]. Non-adiabatic and out-of-equilibrium effects might also blur the simple sequential suppression picture – see *e.g.* the open quantum systems approach [87]. Corrections from viscous QGP hydrodynamics have also been implemented [88]. Most recent approaches, such as the ‘heavy quarkonium quantum dynamics’ [89], aim at a realistic combination of these effects – also including the medium geometry and dynamics – to be compared to measurements, and also predict elliptic flow values. The quoted approaches mostly address bottomonium states, for which the non-relativistic approximation is solid and the three states are experimentally easy to study in parallel – but more importantly, because the recombination of charm quarks significantly counterbalances the charmonium suppression in some regimes, see section 2.5. Transport models [90] intend to dynamically include both opposing effects by calculating a detailed balance of states. They can be based on Boltzmann equations [91], or on Langevin processes, *i.e.* Brownian motion to describe the diffusion in a non-static medium [92].

The  $B_c$  meson has a peculiar nature of *asymmetric* bound state of heavy quarks, yielding an intermediate size and binding energy, so its dissociation temperature is expected to be above that of  $J/\psi$  and below that of  $\Upsilon$  mesons. It hence provides a new perspective on this rich topic.

## 2.5 Recombination and predictions of the $B_c$ modification

Central PbPb collisions at 5.02 TeV can produce up to 200 charm quark pairs<sup>1</sup> – this can be compared to about 8 beauty quark pairs produced per PbPb collision of centrality 0–5%<sup>2</sup>. As the deconfined medium allows quarks to move further than the scale of a nucleon (1 fm), this could lead to low- $p_T$  charm quarks that, despite coming from different uncorrelated hard processes, would recombine into charmonium  $c\bar{c}$  states. This is sketched in Fig. 2.2. It is expected to act mostly on low- $p_T$  heavy quarks, as they are more numerous (increasing the combinatorics) and have in general a smaller relative speed (the proximity in phase space facilitates the binding of the quarks into a hadron). This *recombination*<sup>3</sup> is the common understanding of the puzzles that emerged from comparing measurements of the  $J/\psi$  modification factor at the different energies of LHC and RHIC and from the non-null  $J/\psi$  elliptic flow measured at LHC.

### Charm recombination: clues and models

The existence of recombination of charm quarks in the medium is indeed almost a consensus today, mainly thanks to the combination of three observations, all made possible by the LHC heavy ion program. The most striking one was the  $R_{AA}$  of inclusive  $J/\psi$  mesons (both prompt and non-prompt, and mostly at low- $p_T$ ) measured with ALICE [93, 94] to be undoubtedly higher than in the PHENIX experiment at RHIC [95, 96] towards central collisions. And this despite the 10-fold increase in energy resulting in a significant increase of the QGP temperature from RHIC to LHC! This was difficult to interpret without a high-temperature process that enhanced  $J/\psi$  production rather than diminishing it. The fact that ALICE also showed a much higher  $R_{AA}$  than PHENIX at low  $p_T$  (and not particularly at high  $p_T$ ) gave a solid sign that this enhancement was consistent with a charm recombination mechanism, that is more important at low- $p_T$ . The last clue was the large positive  $J/\psi$  elliptic flow observed with ALICE [97], that can be interpreted as  $J/\psi$  mesons inheriting the flow from the charm quarks, which were themselves ‘pushed’ by the pressure gradient of the medium (this being possible because the quarks are deconfined). As explained in section 2.1, neglecting the Lorentz boost and fragmentation, charm quarks are produced on a time scale of order  $1/m_{J/\psi}$ , meaning typically before the medium formation. Though some primordial  $J/\psi$  mesons could be produced somewhat later (on scales  $1/mv$ ) and traverse the medium as color-octet states, a large fraction of them will traverse it as color-singlet states. These states do not thermalise, so they must obtain their flow from recombining charm quarks, whose elliptic flow came from being partially thermalised. Fig. 2.8 gathers this impressive beam of signs, where the whole is more than the sum of the parts.

Many models nevertheless compete to quantitatively describe the  $J/\psi$  data with various implementations of recombination processes. I will only shortly describe them here. Let us start with the statistical hadronisation model, which describes impressively well the abundance of many particle species through Fermi-Dirac (for fermions) and Bose-Einstein (for bosons) distributions depending on the ratio of the particle mass to the medium temperature – see the comparison to ALICE measurements of the cross sections of hadrons composed of u, d, and s quarks over a wide mass range in Ref. [98]. Contrarily to the light quarks that have a mass

<sup>1</sup>A first order estimation can be obtained from the  $c\bar{c}$  cross section in pp at 5 TeV,  $\sigma_{cc}^{pp} \simeq 7$  mb, and the average nuclear overlap function in a centrality class (for example  $\langle T_{PbPb} \rangle_{0-5\%} \simeq 26 \text{ mb}^{-1}$ ) that can be seen as a luminosity of nucleon binary collisions per PbPb collision. There are hence in average  $\sigma_{cc}^{pp} \times \langle T_{PbPb} \rangle_{0-5\%} \simeq 180$   $c\bar{c}$  pairs produced in a single PbPb collision of 0–5% centrality class at 5 TeV.

<sup>2</sup>With the same value of  $T_{PbPb}$ , and a total  $b\bar{b}$  cross section in pp at 5 TeV of  $\sigma_{bb}^{pp} \simeq 0.3$  mb, there are in average 8 beauty quark pairs produced per 0–5% centrality PbPb collision.

<sup>3</sup>Concerning semantics, the production of heavy quark bound states that do not come from primary production in hard scatterings is sometimes rather called *regeneration*. I do not distinguish here the agnostic term of regeneration with the term of recombination, designating the major process thought to lead to regeneration. The term of *coalescence* is more often used to designate quarks that form bound states simply for being close in phase space.

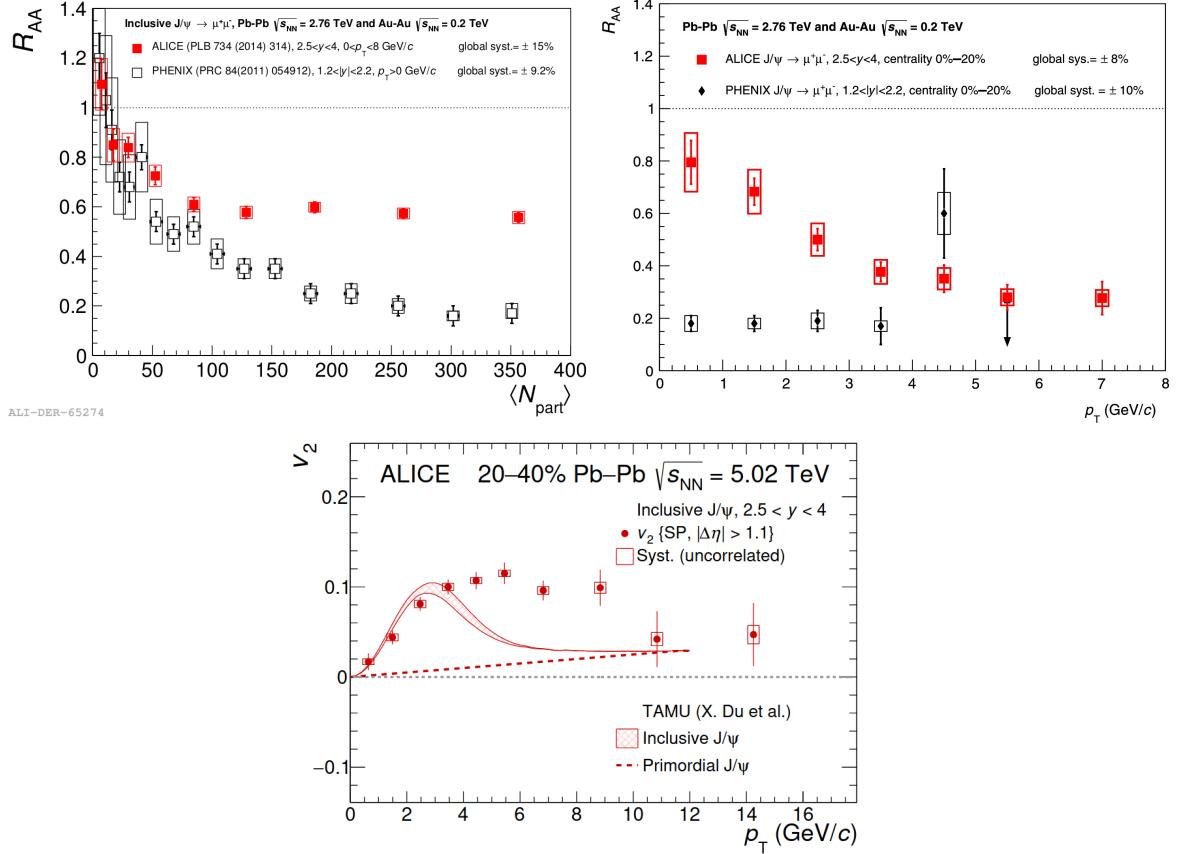


FIGURE 2.8: Signs of charm recombination from ALICE measurements of inclusive  $J/\psi$  mesons. *Top left:* comparison of the  $N_{\text{part}}$  dependence of  $R_{AA}(J/\psi)$  between RHIC (PHENIX [96]) and LHC (ALICE [94]) energies.  $R_{AA}$  is higher in ALICE than in PHENIX towards central events. *Top right:* same comparison, but for the  $p_T$  dependence.  $R_{AA}$  is higher at low- $p_T$  in ALICE than in PHENIX. *Bottom:*  $J/\psi$  elliptic flow measured by ALICE [97], compared to a transport model.

smaller than the QGP temperature, charm quarks are too heavy to be significantly produced thermally to satisfy this statistical model. However, assuming a full thermalisation of the charm quarks produced in the primary hard scatterings, and that the bound states form close to the freezing of the medium, the statistical hadronisation predicts an enhancement of charmonia via recombination of the charm quarks at hadronisation time [99]. The comovers model [100] solves a rate equation using a cross section of final state interactions with surrounding particles (the comovers) fitted on measurements. It stays agnostic with regard to the nature of the particles that trigger  $J/\psi$  dissociations or recombinations. Finally, as mentioned in section 2.4, transport models [90–92, 101] also solve a rate equation including dissociation and recombination processes, implementing the diffusion properties of the QGP through Boltzmann or Langevin equations. The typical form of the time-dependent rate (from ZHAO and RAPP in Ref. [91]) is:

$$\frac{dN_\psi}{d\tau} = -\Gamma_\psi(T)[N_\psi - N_\psi^{\text{eq}}(T)] \quad (2.6)$$

where the equilibrium yield  $N_\psi^{\text{eq}}(T)$  is damped (from the purely statistical expectation) with a relaxation time implementing a partial charm thermalisation, and the dissociation/formation rate  $\Gamma_\psi(T)$  considers all the processes  $g + \psi \leftrightarrow g + c + \bar{c}$ , where  $g$  can be a gluon or any light (anti)quark.

### Beauty recombination

The  $b\bar{b}$  cross section is much smaller, so the recombination of beauty quarks, even at the high LHC temperature and energy, would be small if any. However, the primary production of the most excited bottomonium states is very suppressed, and recombination is more efficient for states of larger size (in which quarks that are less close in phase space can bind). Therefore, recombination might produce a significant part of the observed excited bottomonium states – especially when considering the recombination of beauty quarks that were previously dissociated by QGP interactions, as does Ref. [102]. Most of the models quoted above for charmonia also computed the expected modification (and elliptic flow) of bottomonia, including its possible regeneration – *e.g.* the DU, RAPP, and HE transport model [103], or the heavy quark quantum dynamics from ISLAM and STRICKLAND [89].

### Charm-beauty recombination

The cross section for direct production of  $B_c$  mesons in pp collisions is small compared to other  $B$  mesons and to charmonia (see section 2.2). But in PbPb collisions, heavy quark recombination could be a significant  $B_c$  production mechanism. A  $\bar{b}$  quark produced in a given hard scattering could recombine with one of the numerous  $c$  quark from another nucleon-nucleon collision (see the sketch of Fig. 2.2). In terms of purely statistical hadronisation, this process scales linearly with the number of charm quarks, contrarily to the recombination of  $J/\psi$  mesons that scales quadratically with it. However, if this is a significant mechanism, the  $B_c$  yields might be dramatically augmented compared to the expectations from  $N_{\text{coll}}$ -scaling, considering the small primary production expected from the pp cross section. SCHROEDTER, THEWS and RAFELSKI [104] predicted already in 2000 (not long after the discovery of the  $B_c$  meson) an enhancement of up to  $10^3 - 10^4$ . They used a simple kinetic model for the recombination and a generic QGP expansion model; an important caveat is that they neglected final state interaction with the medium that could suppress these produced states. LIU, GREINER and KOSTYUK [105] predicted in 2013, both with a statistical coalescence model and with a transport model, a modification factor reaching 2 to 18 in the most central collisions, depending on how much heat exchange between the heavy quarks and the medium is considered in the potential. The authors do not provide a  $p_T$  dependence of the predicted  $B_c$  yields. Fig. 2.9 shows these two predictions.

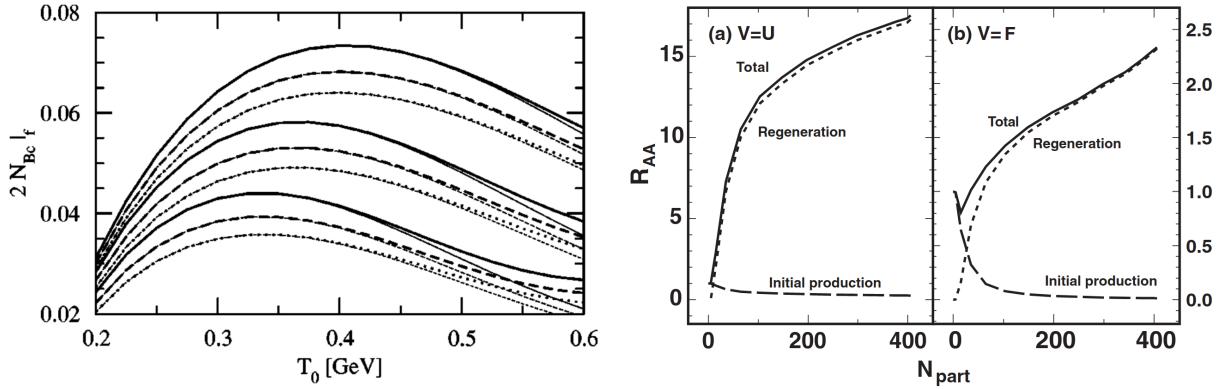


FIGURE 2.9: *Left:* Number of  $B_c$  produced per  $b\bar{b}$  pair predicted for RHIC by Ref. [104], versus the initial QGP temperature, and for various QGP volume and relative quark speed assumptions. The values of order 0.05 should be compared to the values  $10^{-5} - 10^{-4}$  expected at RHIC in the vacuum. *Right:*  $R_{\text{PbPb}}(B_c^+)$  versus  $N_{\text{part}}$  predicted at LHC at  $\sigma_{\text{NN}} = 2.76 \text{ TeV}$  by Ref. [105], for two hypotheses ( $V = U$  or  $V = F$ ) on the heavy quark potential inputted in the Schrödinger equation solved to provide the dissociation temperature.

RAPP et al. provided, upon request, a prediction of the inclusive  $R_{\text{PbPb}}(B_c^+)$ , based on a transport model already applied to the modification of charmonia [91] and bottomonia [103] in

the QGP. Their paper is in preparation, along with the  $p_T$  dependence that would be needed for a comparison to the results of this thesis. As they predict a large fraction of the produced yield to come from regenerated  $B_c$  mesons rather than from primary production in hard scatterings, the shown results depend a lot on the assumed  $B_c$  cross section in pp in the denominator of the  $R_{AA}$ . Two different cross sections were hence considered, based on existing measurements. I provided the pp measurement of this thesis to the authors for more consistent future estimates. Fig. 2.10 shows that both assumed pp cross sections lead to  $R_{PbPb}$  of the order of unity, without strong dependence on  $N_{\text{part}}$ . YAO also provided a set of predictions, based on Ref. [102]; it includes a  $p_T$  dependence which makes it comparable to the results of chapter 9, so it is discussed there.

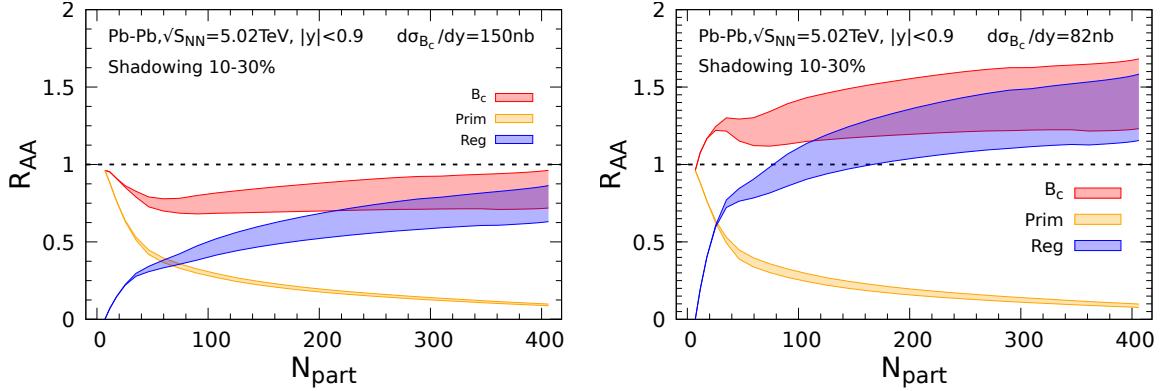


FIGURE 2.10:  $R_{PbPb}(B_c^+)$  versus  $N_{\text{part}}$ , predicted by RAPP et al. (unpublished work from the authors of Refs. [91, 103]) at  $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$ , for two assumptions on the  $B_c$  cross section in pp. Most  $B_c$  mesons escaping the QGP originate from recombination processes in this model.

A central caveat for all these predictions is that recombination processes are relevant mostly at  $p_T \lesssim m_{B_c}$ : first because most of the charm quarks present in the QGP have low  $p_T$ , and second because the recombining quarks need to have a small relative momentum (typically not larger than the binding energy) to be able to recombine. However, as explained in section 5.3, the acceptance of the CMS detector only allows to detect  $B_c$  mesons of relatively high transverse momentum, typically  $p_T > 6 \text{ GeV}$ . Therefore, a majority of the  $B_c$  mesons produced by recombination should have a  $p_T$  lower than the acceptance threshold of CMS. Some remnant recombined  $B_c$  mesons can though be hoped for in the lowest reachable  $p_T$  region. The possible resulting enhancement shall be partly compensated by all suppression mechanisms mentioned earlier. Most models predict that the majority of  $B_c$  mesons produced in the primary hard scatterings are dissociated in the medium; however this study [106], considering no regeneration and only the gluon fusion production mechanism, still shows substantial expected yields in PbPb collisions.

Let us summarise what motivates the experimental study of  $B_c$  mesons in heavy ion collisions. At high- $p_T$ , it can bridge the gap between charm and beauty energy loss. At lower  $p_T$ , it can provide a unique insight into the recombination mechanism. It also connects the physics of charmonia and bottomonia, notably in terms of the colour screening of the heavy quark potential. The measurement of the modification of the  $B_c$  meson in PbPb collisions performed in this thesis will thus help understanding both **energy loss and hadronisation dynamics of heavy quarks** in the QGP. The PbPb nuclear modification factor of the  $B_c$  is measured here both as a function of the (partially reconstructed)  $p_T$  and of the collision centrality. The QGP mechanisms discussed in this chapter are indeed expected to depend on these two variables.

“Neuf l’outil n’est pas fait, il faut que s’établisse entre lui et les doigts qui le tiennent cet accord né d’une possession progressive de gestes légers et combinés, d’habitudes mutuelles et même d’une certaine usure.”

---

Henri FOCILLON, *Vie des formes, éloge de la main*,  
1934, p. 34.<sup>1</sup>

## Chapter 3

# How? Analysis strategy with CMS

As shown in chapter 2, studying the  $B_c$  meson in heavy ion collisions sheds light on heavy quark hadronisation and energy loss. However, measuring its production in heavy ion collisions has never been achieved, notably because of the lower available luminosities and the important underlying event activity in these collisions. This thesis demonstrates the first observation of  $B_c$  mesons in PbPb and pp collisions at a centre-of-mass energy per nucleon pair of  $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$ , and the measurement of the corresponding cross sections and nuclear modification factor. This chapter shows the means used for this accomplishment, starting with the context of CERN and its main accelerator, the LHC, in section 3.1, and the apparatus of the CMS detector in section 3.2. The global analysis workflow is reviewed in section 3.3.3, referencing each part of the coming analysis. In section 3.4, I explain the partial blinding used for the PbPb dataset. Finally, section 3.5 details the extraction and generation of the data and simulated samples used in the analysis, along with their normalisations.

### 3.1 CERN and the LHC

CERN, the *European Council for Nuclear Research (Conseil Européen pour la Recherche Nucléaire)*, was founded at the French-Swiss border close to Geneva in June 1953, when the CERN Convention was signed by 12 member states. Its missions and principles were then declared, including a policy of open access, a budget shared among member states, and more importantly peaceful and world-class fundamental research in nuclear physics. Most of its current research now deals with sub-nuclear physics, justifying the less used but more appropriate CERN name of *European Laboratory for Particle Physics*. Nowadays, 23 countries are member states, and 10 are associate member states. CERN has 2500 employees, but about 12 000 researchers of 110 nationalities participate in its activities, operating the experiments or analysing their results. In this section, I first present a brief history of CERN and its achievements (section 3.1.1), then

---

<sup>1</sup>Translation by Luc DEITZ: “When new, a tool is not ready; between itself and the fingers that hold it, that kind of agreement must come about which is the result of a progressive mastery of gestures both gentle and combined, of mutual habits, and even of a certain wear.”

I describe CERN headliner: the LHC (section 3.1.2), that manages to host both the largest cryogenic system of the world – below the temperature of interstellar space – and the hottest place of today’s Universe, in heavy ion collisions! Section 3.1.3 describes the main experiments at the LHC and the discoveries they have carried out.

### 3.1.1 Brief history of CERN facilities and discoveries

#### 3.1.1.1 First generation accelerators

After Felix BLOCH laid the first stone of CERN in July 1955, the international collaboration kept up to its stated mission of pushing the boundaries of human knowledge by building and maintaining unique facilities for unparalleled research. For most of its history, CERN provided the research field with accelerators at the highest energy in the world, as shown in Fig. 3.1.

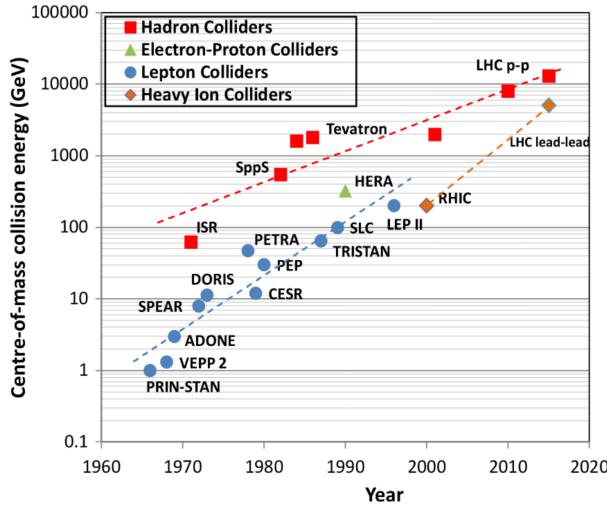


FIGURE 3.1: Center-of-mass energy of lepton and hadron accelerators versus the time they began to operate. The tribute goes to CERN for the following accelerators: ISR, SppS, LEP and LHC. Figure from Ref. [107].

The *Synchro-cyclotron* (SC) was the first accelerator built, and started to accelerate protons to 600 MeV in 1957 thanks to magnets of 5 m diameter. It soon started discoveries, observing the rare **electron decay of the pion** [108], for which a few previous experiments had wrongly claimed an upper limit of its branching fraction lower than the theoretical prediction. Its focus shifted to nuclear physics from 1964, including the production of rare isotopes, and it was decommissioned only in 1990.

The *Proton Synchrotron* (PS) was the second CERN project, using more modern technologies compared to the SC, in a ring of 230 m circumference. Contrarily to a cyclotron where the accelerated particles spiral out as they are accelerated, a synchrotron increases progressively its electromagnetic field to match the energy of the particles, allowing them to circulate in a circular beam of constant radius. The PS started accelerating a beam of protons to an energy of 24 GeV in 1959. In its first years, it provided beams for a neutrino experiment, a muon storage ring, and multiple bubble chamber experiments, including *Gargamelle* that was receiving neutrinos from a target hit by the PS beam. It then served as pre-accelerator successively for the ISR, SPS, LEP, and LHC accelerators, and still functions nowadays after more than 60 years of service.

The *Gargamelle* bubble chamber, containing  $12 \text{ m}^3$  of liquid Freon, detected neutrinos from the PS beam hitting a fixed beryllium target, from 1971 to 1979. It discovered in 1973 both leptonic ( $\nu_\mu + e \rightarrow \nu_\mu + e$ , with a yet-undiscovered  $Z^0$  boson mediator) [109] and hadronic (such as  $\nu + p/n \rightarrow \nu + p/n$ ) [110] **neutral currents**, by comparing the candidate events with the

equivalent charged current candidates mediated by the then-hypothesised  $W^\pm$  boson. The electroweak unification in the Standard Model, evoked in section 1.2, was conceived in the preceding years by SALAM and WARD [111], GLASHOW [112], and WEINBERG [113], notably to explain the observed charged currents (now interpreted as the exchange of a  $W^\pm$  boson). Its credibility was increased by the proposition of the Higgs mechanism (in three papers in 1964 by HIGGS [114], ENGLERT and BROUT [115], and GURALNIK et al. [116]) that allows for the electroweak gauge bosons to be massive through the spontaneous breaking of the electroweak symmetry, and by the proof of renormalisability of the resulting gauge interactions by T'HOOFT [117]. However, a neutral weak boson was needed to complete the  $SU(2) \times U(1)$  group assumed to unify the electroweak interactions.<sup>1</sup> The discovery of neutral currents by *Gargamelle* was hence the central experimental evidence for the Standard Model.

### 3.1.1.2 The collider era

The *Intersecting Storage Rings* (ISR), consisting of two rings of 942 m circumference, was the world's first hadron collider. It ran from 1971 to 1984 at a center-of-mass (c.o.m.) energy of 62 GeV, this high energy being allowed by colliding beams.<sup>2</sup> It demonstrated many **collider technologies**, particularly directed to increase or monitor the beam luminosity. For example, stochastic cooling damps the betatron oscillations of the beam to increase the beam quality [118], and is used at the LHC as well in CERN decelerators such as the AD; and the VAN DER MEER scan for luminosity measurement, still used at the LHC, was conceived for the ISR [119]. The ISR also established large magnets and  $4\pi$  geometric acceptances as the central features of detectors at a particle collider.

The 6.9 km-circumference ring of the *Super Proton Synchrotron* (SPS) accelerated proton beams to a 300 GeV energy from 1976, and now serves as a pre-accelerator for the LHC, accelerating protons to 450 GeV. It nowadays also provides proton beams for fixed target experiments as COMPASS, NA61, and NA62. It also accelerated electrons and positrons to be injected in the LEP (1989–2000, cf section 3.1.1.3). During the 1981–1991 period, it was called the *Super Proton–Antiproton Synchrotron* (Sp $\bar{p}$ S) and also accelerated antiprotons, to collide them with protons up to a c.o.m. energy of 630 GeV. Carlo RUBBIA, Peter MCINTYRE and David CLINE had the idea of this transformation, suggesting this new machine would be able to produce the  $W$  and  $Z$  bosons, so long awaited since the Standard Model was conceived and arguably confirmed by the observation of neutral currents in *Gargamelle*. The experiments UA1 and UA2 placed at the interaction points of the Sp $\bar{p}$ S beams achieved the **observation of the  $W$**  [120, 121] and  **$Z$**  [122, 123] **bosons**, unequivocally proving that the Standard Model was the best available understanding of the electromagnetic and weak forces. In parallel to the theoretical advances on the electroweak unification, the quark model was developed by GELL-MANN [36] to explain the zoo of ‘strange’ hadrons discovered in the 1950’s. The original three quarks were based on an approximate flavour- $SU(3)$  group, broken by the mass differences of the u, d, and s quarks [124]. But to understand the existence of the newly-discovered  $\Omega^-$ (sss) baryon with

<sup>1</sup>I warmly recommend reading this excellent review from GLASHOW on the emergence of the Standard Model: <https://inference-review.com/article/the-standard-model>.

<sup>2</sup>Let us show the potential of colliding two accelerated beams of particles of respective energies  $E$ , compared to sending a beam of energy  $E_{\text{beam}} = 2E - m_2$  of particles of mass  $m_1$  onto a fixed target of particles of mass  $m_2$ . In the first case, the laboratory frame coincides with the centre-of-mass frame, so the c.o.m. energy is simply  $\sqrt{s} = 2E$ . In the second case,  $\sqrt{s}$  is obtained via a Lorentz transformation from the lab to the c.o.m. frame. The c.o.m. energy is the invariant mass of the 2-particle system of energy  $2E$  and momentum  $|p| = \sqrt{(2E - m_2)^2 - m_1^2}$  in the lab frame. This gives  $\sqrt{s} = \sqrt{(2E)^2 - (2E - m_2)^2 + m_1^2} = \sqrt{4Em_2 + m_2^2 + m_1^2} = \sqrt{2E_{\text{beam}}m_2 + 3m_2^2 + m_1^2}$ . The ratio, at fixed lab frame energies, between c.o.m. energies in the collider or fixed target mode is hence  $1/\sqrt{\frac{m_2}{E} + \frac{m_1^2 + m_2^2}{4E^2}}$ . Assuming  $E \gg m_i$ , this ratio is  $\sqrt{\frac{E}{m_2}} = 32$  or  $83$ , for proton beams at  $E = 980$  GeV (TeVatron) or  $E = 6.5$  GeV (LHC Run II).

regard to the Pauli exclusion principle, the three strange quarks needed to have different quantum numbers. The colour charge was therefore introduced, resulting in the QCD theory (see section 1.3.1) based on the colour- $SU(3)$  group [125]. With this last piece, the Standard Model group  $SU(2)_{\text{left}} \times U(1)_{\text{hypercharge}} \times SU(3)_{\text{colour}}$  as we still know it today was finally formed; it is further explained in section 1.2.

### 3.1.1.3 Towards modern machines

Of more direct interest to this thesis is that the SPS accelerated heavy ion beams from 1986 to 2003. In a first phase (until 1993), existing facilities were used to accelerate oxygen ( $^{16}\text{O}$ ) and sulphur ( $^{32}\text{S}$ ) beams to 60 and 200 GeV per nucleon, collided on fixed targets (of sulphur, oxygen, and others) in multiple experiments.<sup>1</sup> Theoretical predictions of the consequences of the existence of QGP were flourishing in the 1980's, including a central rapidity plateau of hadron production (due to the longitudinal expansion of the medium foreseen by BJORKEN [126]), and the enhanced production of hadrons containing strange quarks [127]. Evidence for the latter was one of the first experimental signals of the existence of QGP: the ratio of  $\bar{\Xi}$  baryons (containing two strange quarks) and  $\bar{\Lambda}$  baryons (containing only one) was found (with a relatively high probability) to be larger in sulphur-tungsten collisions than in pp collisions by the WA85 experiment [29]. In a second period, from 1994 to 2002, SPS accelerated lead ( $^{207}\text{Pb}$ ) beams to 158 GeV per nucleon, which collided on lead fixed targets in seven experiments specifically conceived for heavy ion physics. Other signatures of the QGP were then brought to light, such as a strong elliptic flow, and the suppression of  $J/\psi$  mesons in central collisions (see Fig. 9 of Ref. [128] that shows a ratio of  $J/\psi$  to Drell-Yan events, indeed decreasing from peripheral to central collisions). This finally lead to the CERN announcement of the **discovery of a “new state of matter”** in February 2000.<sup>2</sup> The SPS still provides today proton and ion beams for many fixed-target experiments as COMPASS (notably studying the nucleon spin and the parton transverse momentum distributions in the proton) or AWAKE (studying the feasibility of plasma wakefield acceleration for future accelerators).

CERN also worked at decelerating particles for antimatter studies. The *Low Energy Antiproton Ring* (LEAR) ran from 1982 to 1996, decelerating antiprotons coming from the *Antiproton Accumulator Complex* (AAC), that stored antiprotons produced by the PS beam. The LEAR allowed the PS210 experiment to create the **first antihydrogen atoms** in 1995 [129]. The LEAR and the AAC decelerating functions were taken by the *Antiproton Decelerator* (AD) in 2000, which, still today, brings antiprotons down to a kinetic energy of 5.3 MeV. Its experiments are now dedicated to studying the spectra of small antimatter atoms, and to find out if gravity acts the same on antimatter than on matter. The LEAR machine was converted in 2005 to the *Low Energy Ion Ring* (LEIR), used to accelerate the ions that are then injected into the PS, and ultimately into the LHC.

The large and ever-increasing number of staff, users, projects, and experiments, could be a reason why CERN was the stage of the creation of one of the most revolutionary technology for information management and sharing: the **WorldWideWeb**. It was invented by Tim BERNERS-LEE in 1989 as a “wide-area hypermedia information retrieval initiative aiming to give universal access to a large universe of documents”, according to the first ever website created in 1991, that CERN has recreated<sup>3</sup>. The first project proposal can also be found there.<sup>4</sup>

Getting now closer to the LHC era, the next large accelerator at CERN was the *Large Electron–Positron Collider* (LEP), colliding electrons with positrons at c.o.m. energies of 91 to

---

<sup>1</sup>A beam of energy 200 GeV per nucleon on a fixed target corresponds to a c.o.m. energy per nucleon pair of  $\sqrt{s_{\text{NN}}} = 19.5$  GeV, hence one (two) orders of magnitude smaller than at RHIC (LHC).

<sup>2</sup><https://home.cern/news/press-release/cern/new-state-matter-created-cern>

<sup>3</sup><http://info.cern.ch/hypertext/WWW/TheProject.html>

<sup>4</sup><http://info.cern.ch/hypertext/WWW/Proposal.html>

209 GeV, from 1989 to 2000. Its tunnel accommodates today the LHC. LEP still holds the high-energy record for lepton colliders, the potential future competitors being the linear *International Linear Collider* (ILC) project in Japan or the circular FCC-ee (*Future Circular Collider*) at CERN. Beyond the precise measurement of many electroweak boson properties, one of its main accomplishment was to measure the **number of light neutrinos** (with  $m_\nu < m_Z/2$ ) thanks to the measurement of the mass width of the Z boson production, via the Z cross section as a function of the c.o.m. energy. The legacy measurement [130] of  $N_{\text{light}\nu} = 2.9840 \pm 0.0082$  leaves no doubt for other values than 3, hence strongly constraining models that add a fourth fermion generation to the SM. The high precision of various Z observables also allowed for the prediction of the mass of the top quark with a Standard Model fit ( $172^{+22}_{-24}$  GeV in 1994, p.137 in Ref. [131]) just one month before the first indication of its production at TeVatron [132]. LEP was also on the verge of finding the Higgs boson, the accessible phase space giving 95% confidence lower limits to the Higgs mass of 114 GeV; false hopes even emerged from a  $1.7\sigma$  statistical fluctuation at 115 GeV. It almost called for an upgrade of the LEP magnets to slightly rise the centre-of-mass energy, which was abandoned not to delay the construction of the LHC.

### 3.1.2 The Large Hadron Collider

#### 3.1.2.1 The LHC injection system

Fig. 3.2 shows the current complex of accelerators at CERN. Many pre-accelerators aim at progressively rising the energy of protons or ions before injecting them in the *Large Hadron Collider* (LHC). This is the largest machine ever built by humans, in a 26.7 km tunnel lying from 50 to 175 m underground in the Geneva area, both in France and Switzerland. It accelerates both protons (hydrogen nuclei), and lead ( $^{208}\text{Pb}$ ) and xenon ( $^{129}\text{Xe}$ ) nuclei, to unparalleled energies (up to 6.5 TeV per nucleon).

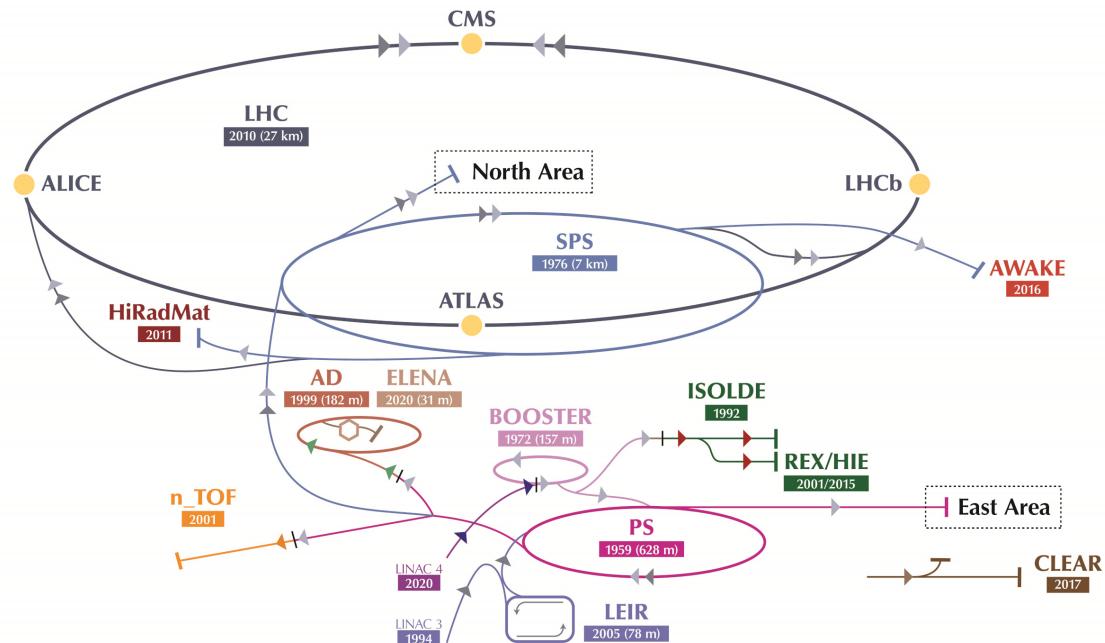


FIGURE 3.2: The CERN accelerator complex.

Let us follow the acceleration path of protons or ions until the LHC beam. It all starts with a bottle of gaseous  $\text{H}_2$  or a from a solid block of lead that is heated into a vapour, transformed into a plasma by microwaves before being accelerated. Only half a gram of lead is used in two weeks of running heavy ion collisions!

### Chain of accelerators

In proton runs from 1978 to 2018, the hydrogen was stripped off its electron with an electric field, leaving only  $H^+$  ions to be accelerated to 50 MeV in the 36 m long linear accelerator **Linac2**. The latter has been replaced for the future runs by Linac4 (a 86 m long line, 12 m under the surface), which rather accelerates  $H^-$  ions to 160 MeV before stripping away its two electrons. The protons are then injected in the *Proton Synchrotron Booster* (**PSB**) to be accelerated to 1.4 GeV (or 2 GeV when coming from Linac4), before injection in the PS. The Linac4 upgrade will allow to inject more protons in the subsequent accelerators, in view of the coming high-luminosity phase of LHC. The PSB can also direct its beam to fixed-target experiments as ISOLDE, which produces many different radioactive isotopes for later study by 90 associated experiments.

In lead (or xenon, but it is ignored in the numbers given in this paragraph) runs from 1994 until now, the ions are partially stripped, mostly to  $^{208}Pb^{29+}$ , thanks to collisions with heated electrons in the plasma of the *Electron Cyclotron Resonance* ion source. Then, they enter **Linac3** with an energy of 2.5 keV per nucleon, and are accelerated to 4.2 MeV per nucleon. Before being injected into the LEIR, they need to be stripped down to  $Pb^{53+}$  by passing through a carbon foil about 0.5  $\mu m$  thick. Before the construction of LEIR in 2005, Linac3 injected ions in the PSB.

**LEIR** takes in from Linac3 *bunches* (defined in section 1.5.1) of ions long of hundreds of microseconds, and shapes them into four much shorter (about 200 ns) and denser bunches for future injection in the PS [133]. This increase of beam quality (*i.e.* the decrease of emittance, that describes the beam transverse size and momentum spread) uses electron cooling, in which a beam of electrons is coupled to the ion beam, and the lower-momentum electrons take momentum from the ions to reach thermal equilibrium. The increase of beam intensity comes from accumulating pulses (*i.e.* bunches from Linac3) during many turns of the beam. Each 2.5 seconds, two bunches of  $2.2 \times 10^8$  ions are formed, then accelerated to 72 MeV per nucleon, and then stored in the PS.

The **PS** takes in bunches of protons or ions, combines them into *trains* of bunches and accelerates them to 26 GeV for protons, or to 5.9 GeV per nucleon for Pb nuclei, thanks to about 300 magnets at room temperature for beam deflection and focusing, and straight sections with radiofrequency cavities for acceleration (section 3.1.2.2 will give some explanations on beam deflection and acceleration). The **SPS** then brings the protons to 450 GeV or the Pb nuclei to 177 GeV per nucleon, before finally injecting them in the LHC where they are accelerated to their final energy: 5.02 TeV per nucleon pair (2.51 TeV per nucleon for each beam) for the data studied in this analysis, but protons were accelerated up to 13 TeV in other runs. The design c.o.m. proton collision energy is 14 TeV, which is aimed for in future runs. Two beams are actually extracted from SPS and sent in opposite directions in two parallel beam pipes of the LHC. The PS and SPS beams are also distributed to many fixed-target experiments (such as AWAKE or COMPASS) or to the AD (for antimatter studies).

### Bunch schemes

The LHC beam is actually composed of bunches of particles (see section 1.5.1), clearly delimiting the periods in which collisions take place in the detectors. Each bunch contains about  $1 \times 10^{11}$  protons, or  $7 \times 10^7$  Pb nuclei [133]. The bunches are about 2 ns (60 cm) long, and about 200  $\mu m$  large (this transverse width depends on the machine performance, and is strongly reduced at the collision points). Depending on the performance of the machine, running conditions, and target luminosity, the scheme of trains of bunches during the injections in each (pre-)accelerator vary [134].

In optimal pp running conditions, the bunches are separated by 25 ns in the LHC – with some spots along the whole circumference being kept empty for various security reasons (including a secure time to dump the beam in case of urgent failure). In this optimal case, the PS forms trains

of 72 bunches<sup>1</sup>, and groups of two to four trains (or *batches*) are then sent into the SPS. They are then injected in the LHC, which constitutes one SPS cycle. After 12 cycles (each finishing with the insertion of the batches), of 22 seconds each, the LHC length (equivalent to 89  $\mu\text{s}$  at speed  $c$ ) is fully filled (except for security gaps) with about 2800 bunches. In the 2017 pp run studied in this thesis, the number of bunches was somewhat smaller than in these optimal conditions.

In PbPb conditions, the filling scheme is different, following the constraints of the output of LEIR. The two bunches inserted from LEIR are shaped into four bunches in the PS. Those bunches of  $\text{Pb}^{53+}$  are then extracted and passed through a 0.8 mm thick aluminium foil to strip the remaining electrons and obtain lead nuclei  $^{208}\text{Pb}^{53+}$ . Successive injections into the SPS progressively fills it with a train of 8 to 13 sets of 4 bunches separated by 100 ns. Sending this train in the LHC ends the 54 s SPS cycle – this time being limited by the injection time of the two LEIR bunches in the PS. It hence takes 10 minutes to fill the LHC with the 592 bunches needed in the nominal ion runs. At the end of the 2018 PbPb run, a 75 ns bunch spacing could be used to increase the luminosity, thanks to the smooth operations of the machine.

### 3.1.2.2 The LHC apparatus

The two main functions of the LHC, as well as of the PS and SPS, are to accelerate and focus the beams, using magnets and radiofrequency cavities, respectively. The beam also needs to circulate in a very high vacuum. Many superlatives can be used for each of these systems, as suggests the large size of the whole apparatus.

#### Electromagnets

The beam first needs to be bent – the whole idea of a circular accelerator! To this goal, the LHC uses 1232 dipole magnets of 15 m length and weighing about 25 tons each. They apply a 8.3 T magnetic field to the nuclei, not so far from the current world records of the order of 40 T despite the fact that they were conceived twenty years ago. The two beams circulate close to each other, and are both contained in parallel tunnels within the one-piece stainless steel container of the ‘2-in-1’ dipoles.

The beam also needs to keep a small transverse section to obtain high luminosities at the collision points. Mainly 392 quadrupole magnets handle this focusing task, but there are also smaller hexapole, octupole and decapole magnets close to the dipoles to correct the imperfections of their field. To reduce even more the beam section at the point where the two beams collide head-on, sets of three quadrupole magnets are used. Two such *inner triplets* are placed around each collision point, and reduce the beam width from 200 to a few micrometers. Additional dipoles separate again the beams after they collided.

All these electromagnets provide a magnetic field produced by the passage of electric current in coils made of niobium-titanium (NbTi) wires. To transport the 100 MW of electricity needed to run the accelerator (about a fifth of the electricity consumption of Geneva), it is wise to avoid losses by Joule heating, using superconducting coils. This is possible thanks to 120 tons of superfluid helium bringing them at a temperature of 1.9 K (colder than outer space!). This is not so heavy compared to the 36,000 tons of magnets that need to be cooled; the helium superfluidity indeed gives it a very high thermal conductivity.

#### Radiofrequency cavities

Almost all the acceleration steps of the presented accelerator chain use the technology of radiofrequency (RF) cavities. Their main characteristic is that they amplify electromagnetic standing waves, which accelerate the (positively charged) particles that are positioned in the

---

<sup>1</sup>The 72 bunches are separated by 25 ns at close to the speed of light, meaning they total a length of 540 m, slightly less than the PS circumference.

half cycle of the wave that shows a positive electric field, and decelerate particles that are in the other half cycle, where the field is negative. This conveniently creates the bunches of particles mentioned above, and smooths out their longitudinal momentum and position dispersion. The ‘late’ particles are indeed more accelerated than the ‘in-time’ particles, and the particles ‘in advance’ are less accelerated.

The frequency and position of the wave is tuned so that the bunches are positioned in the accelerating half of the wave cycle. The potential energy of the RF waves is hence converted into kinetic energy of the beam particles: the beam accelerates. The LHC cavities oscillate at 400 MHz, so their wavelength is ten times smaller than a 25 ns separation between two bunches. As a consequence, the bunches are made about ten times smaller than the bunch separation.

A *klystron* attached to a hole in the cavity generates these RF waves. It holds a high-power electron beam that passes through (relatively low-power) input RF waves, in which electrons are bunched as in the main cavities. Then an output cavity of the same resonant frequency is placed in the half cycle where the wave decelerates the electrons; in other words the electrons excite the RF waves of the output cavity by yielding them their kinetic energy. These excited high-power RF waves are transmitted to the main RF cavities through waveguides, that are rectangular metal pipes.

The peculiar cavity shape makes the electric field resonate to a maximum intensity, with a high quality factor. The latter is notably possible thanks to a superconducting regime at 4.5 K for the niobium coating of the copper structure; it makes the losses in the currents of the cavity walls negligible compared to the power transferred to the beam. On the LHC ring, there is one straight section with 16 such cavities (8 per beam). With a voltage of 2 MV each (that is however not fully transferred to the particles), they accelerate the beams from the SPS energy to the final energy in about 20 minutes, corresponding to more than ten million turns in the tunnel.

### Vacuum system

There are three vacuum systems along the LHC, totalling 104 km of pipes and a volume of 15.000 m<sup>3</sup>. 50 km of those pipes maintain a 10<sup>-9</sup> bar vacuum to thermally insulate the cooling system of the superconducting magnets, and the helium line that distributes the coolant to them.

The other 54 km of pipes (27 km per beam line) welcome a 10<sup>-13</sup> to 10<sup>-14</sup> bar vacuum around the beams so that they do not collide with gas particles; this would lower the luminosity available for the experiments and possibly damage the beam pipes. This pressure, as low as the one on surface of the Moon, is maintained through cryogenic pumping in the 24 km of the ring that welcome superconducting bending magnets: the gases condense on the refrigerated walls and adhere to their surface. The vacuum system in the remaining 3 km of the ring that are at room temperature is ‘baked’ at 300 °C. This pumps out residual gas, thanks to a special coating, developed at CERN, covering the beam pipe.

#### 3.1.2.3 LHC pre-history: path to the first collisions

The idea of LHC existed 30 years before its first physics runs: in 1977, John Adams, as technical director of CERN, initially pushing for building a hadron collider rather than LEP, pointed out [135] that the LEP tunnel should be large enough to welcome magnets for a hadron collider at a few TeV. In 1984, as the construction of LEP was beginning, the seminal Lausanne ECFA (*European Committee for Future Accelerators*) workshop [136] was held. It aimed at kick-starting the necessary R&D efforts on magnets, and solidifying the LHC physics case, which was already centred on the Higgs boson and mentioning supersymmetry. The project was then competing with the *Superconducting Super Collider* (SSC), a USA project approved in 1987 but abandoned in 1993 after 23 km of tunnel were already dug (out of the 87 km of planned tunnel). The SSC was aiming at  $\sqrt{s} = 40$  TeV pp collisions, but with a lower luminosity than LHC; it

was also lacking international support, as it was politically presented as a US-centered project rather than an international one as the LHC.<sup>1</sup>

In 1987, the possibility of  $p\bar{p}$  collisions was set aside, for pp collisions at 13 to 15 TeV with an instantaneous luminosity of  $10^{33} \text{ cm}^{-2}\text{s}^{-1}$  (multiple interactions per bunch crossing were initially not considered). In the LHC Conceptual Design Report in 1995, the original idea of installing LHC on top of LEP to run both simultaneously was also abandoned, because the settings of the necessary magnets would have been too complicated. The machine design, construction, and budget were approved in 1994. Letters of intent for possible detectors were written in 1992, and the ATLAS and CMS detectors were approved in 1997. Approval for the LHCb and ALICE experiments followed a few years later.

The on-site construction started after LEP stopped taking data in 2000. The first beam circulated in the LHC ring on September 10<sup>th</sup>, 2008. The joy was short, as on September 19<sup>th</sup>, a fault in an electrical bus joining two magnets caused local resistive heat losses, which caused in turn a quenching of the neighbour magnet – when the magnet regime (here the temperature) is changed too fast, making it lose superconductivity. The electric discharge also produced an electric arc that pierced the helium cryogenic system, due to which a total of 6 tons of helium were lost, and many magnets heated up and quenched. The chain of events lead to quite dramatic damage, and operations have been stopped for more than a year to repair or replace dozens of dipole magnets, repair the cooling system, and install safety procedures to avoid such events in the future. Finally, the first collisions unrolled in November 2009, and the first pp collisions at 7 TeV happened in March 2010, without significant incident. It was chosen to run at half the nominal energy, because the full energy would have needed more preparation time for the machine. The first PbPb collisions took place soon after, in November 2010.

### 3.1.2.4 Past and future runs

The running schedule of the LHC depends on the needed technical stops, and on the end-of-year stops, traditionally happening in the heart of winter when electricity is more expensive. At the restart each year, the machine needs to be warmed up and tuned, so that the first physics collisions generally take place in April. Between each run of two or three years, a long shutdown of similar duration serves to repair or upgrade the LHC and its detectors. Each year, 6 to 7 months are dedicated to pp collisions at the highest possible luminosities, and typically one month at the end of the year concerns heavy ion collisions or pp reference runs at low pile-up (see section 5.2 for explanations on pile-up). Fig. 3.3 details the integrated luminosities accumulated in most of the runs that LHC performed. The definition of luminosity is given in section 1.5.3. The given luminosity numbers correspond to the one delivered by the beam, but due to minor detector failures, the one recorded by the experiment is often a few percent smaller.

The first run (**Run I**) saw data collection in pp collisions from April to November 2010 and 2011 at a c.o.m. energy  $\sqrt{s} = 7 \text{ TeV}$ , then from April to December 2012 at  $\sqrt{s} = 8 \text{ TeV}$ . About 75% of the  $10^{34} \text{ cm}^{-2}\text{s}^{-1}$  nominal instantaneous luminosity was reached steadily in 2012. There were also two 1-month PbPb runs in 2010 and 2011, of respective luminosities  $9$  and  $184 \mu\text{b}^{-1}$ , at a c.o.m. energy per nucleon pair of  $\sqrt{s_{\text{NN}}} = 2.76 \text{ TeV}$ . pPb and pp collisions at  $\sqrt{s} = 5.02 \text{ TeV}$  were also recorded during respectively three weeks and three days in February 2013.

A two-year shutdown (**LS1**) followed, during which magnets were trained to reach the field needed to increase the beam energy, and the wear and tear damage of Run I was repaired. Some sub-detectors were also upgraded. The pixel tracker and endcap RPC of CMS got a fourth layer – see section 3.2 for a description of the CMS detector. The ATLAS detector benefited from an additional innermost tracker layer (IBL) and a new luminosity detector (LUCID)– see

---

<sup>1</sup>This is, to me, a bright example of the importance of international peaceful collaborations (such as the CERN ones) in carrying out breakthrough scientific experiments.

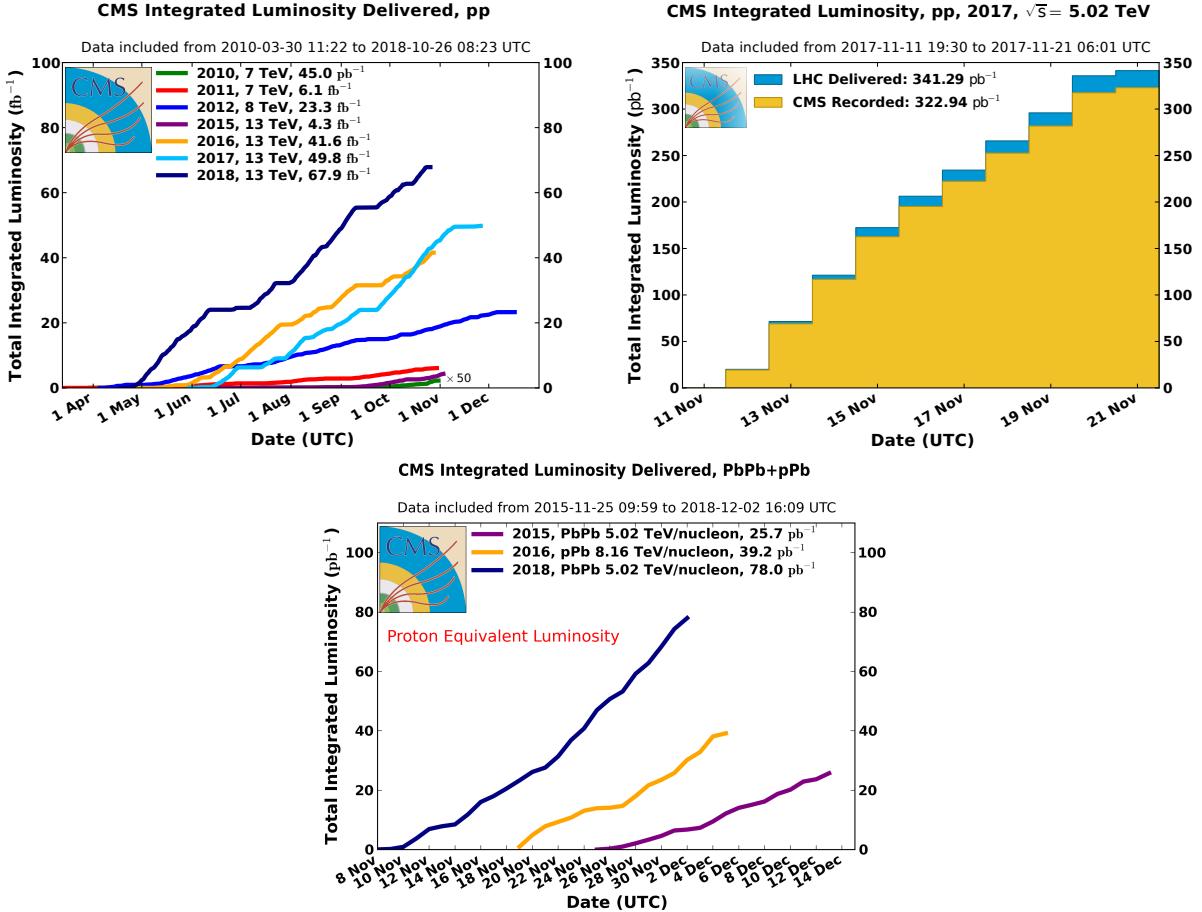


FIGURE 3.3: Luminosity delivered by LHC to the CMS detector in the high-luminosity pp runs from 2010 to 2018 (top left), in the 2017 pp reference run at 5.02 TeV (top right), and in the heavy ion runs of Run II (bottom; here the proton-equivalent luminosities are given, to be divided by  $A = 208$  for pPb and  $A^2$  for PbPb). The luminosity increased significantly from year to year. Figures from Ref. [137].

section 3.1.3.1 for a short description of ATLAS, ALICE, and LHCb detectors. The ALICE detector installed an extended electromagnetic calorimeter for larger acceptance and got most of its electronics replaced. LHCb will undergo a major upgrade in the second long shutdown (LS2).

In the second run (**Run II**) from 2015 to 2018, five times more pp luminosity was collected than in Run I.<sup>1</sup> The bunch separation improved to 25 ns, compared to the 50 ns used in Run I. The energy was risen to  $\sqrt{s} = 13 \text{ TeV}$  – and not to 14 TeV due to the magnet training being slower than expected. In 2017 and 2018, the typical instantaneous luminosity was double the design one (*i.e.*  $2 \times 10^{34} \text{ cm}^{-2} \text{s}^{-1}$ ). This run was also very rich for heavy ion physics, as there were two 3-week PbPb runs in December 2015 ( $0.55 \text{ nb}^{-1}$ ) and 2018 ( $1.8 \text{ nb}^{-1}$ ) at  $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$ ,<sup>2</sup> but also two pp reference runs at the same energy in 2015 ( $28 \text{ pb}^{-1}$ ) and 2017 ( $340 \text{ pb}^{-1}$ ). A pPb

<sup>1</sup>A weasel tried to stop this successful operation on April 29th, 2016: it had sneaked in a 66 kV to 18 kV transformer of the LHC electric network, and sacrificed to cause a short circuit to ground in the transformer. The LHC started back normal operations ten days later, after the repairs of the local damage caused by the electrical arc of the short circuit. See <https://home.cern/news/news/accelerators/lhc-report-stoat-ally-back-track>.

<sup>2</sup>The reachable c.o.m. energy in heavy ion collisions per nucleon pair depends on the ratio of the number  $A$  of nucleons in the nucleus to its accelerating power proportional to its charge  $Z$ . So with the same magnet bending power than for pp 13 TeV collisions, PbPb collisions can reach  $\sqrt{s_{\text{NN}}} = \frac{Z=82}{A=208} \times 13 \text{ TeV} \sim 5 \text{ TeV}$ . In pPb collisions, only the energy of one of the two beams must be multiplied by  $Z/A$ , but the c.o.m. energy is boosted (in the lab) towards the Pb nucleus, which has less energy than the more efficiently accelerated proton. This slightly diminishes the energy per nucleon that would be available with a symmetric pPb collision.

run also collected data at  $\sqrt{s} = 8.16$  TeV per nucleon for three weeks in 2016. An original xenon-xenon (XeXe) run of 8 hours also took place in 2017, recording  $3 \mu\text{b}^{-1}$  of luminosity (equivalent to  $50 \text{ pb}^{-1}$  of pp luminosity). The data used in the  $B_c$  analysis presented in this thesis was collected in the 2017 pp reference run and in the 2018 PbPb run. I had the chance to participate in the 2018 PbPb data-taking, from the CMS control room.

The current **LS2** started in 2019. Magnets are being trained in order to slightly rise the LHC energy (probably to 13.6 TeV rather than 14 TeV). Detectors are also undergoing major upgrades, notably to start preparing them for the approaching High-Luminosity LHC phase (HL-LHC). LS2 was extended by one year (mainly due to the delays from the Covid-19 pandemic), so that Run III physics collisions will start in March 2022. ALICE mainly aims at running without trigger, necessitating an upgrade of the DAQ (*Data AcQuisition system*), in order to enlarge by a factor 100 the recorded minimum bias samples; it also renews its silicon tracker and inner tracker. LHCb increases its readout capacity by a factor 40, to install a software-only trigger system. It also renews with upgraded technologies all its silicon-based tracking systems and the RICH particle identification detector. ATLAS installs new muon small wheels at its endcaps and upgrades its liquid-argon-based ECAL to sustain the higher pile-up (definition in section 5.2) foreseen in future runs. Finally, CMS changes the photo-detectors in the hadronic calorimeter, as well as the material of the beam pipe; its major upgrade will take place in LS3 (planned in 2025–2027), when a highly-granular calorimeter (HGCal) will replace both the electromagnetic and hadronic calorimeters in the endcaps, the silicon tracker will be upgraded with trigger capabilities, and a MIP (*minimum-ionisation particle*) timing detector will be installed for better pile-up discrimination and particle identification.

Now what is next? The Higgs boson being discovered, and all searches for phenomena beyond the Standard Model being inconclusive, future projects at higher luminosities and energies seem to be the most promising to make a discovery. The luminosity will sharply increase from Run IV, intended to start in 2027. This upgraded machine, baptised *High-Luminosity LHC* (or Phase II), aims at progressively increasing the instantaneous luminosity to  $5 - 7 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . Heavy ion runs will continue with a similar pattern, except that a short oxygen-oxygen run is also planned in Run III. The schedules of Run V and VI are unclear yet, and depend on the approval and timelines of the future projects. A supposed end-date is 2038, at which point the total recorded pp luminosity will be about  $3000 - 4000 \text{ fb}^{-1}$ , which is 10 to 20 times what has been collected until 2021!

The successor of HL-LHC, not yet approved nor funded, is thought to be the *Future Circular Collider* (FCC), first as a lepton collider version (FCC-ee), then as a hadron collider (FCC-hh). A 100 km tunnel would be excavated and ultimately welcome pp collisions at 100 TeV. The main competitors of FCC-ee are the ILC or the *Circular Electron Positron Collider* (CEPC) in China, that would have similar capabilities for precise measurements of Higgs boson properties. The *Compact Linear Collider* (CLIC) is also another option at CERN. The main competitor of the FCC-hh would be the *Super proton-proton Collider* (SppC), the continuation of the CEPC project, using the same 100 km tunnel.

### 3.1.3 Main LHC experiments and results

#### 3.1.3.1 The Big Four (detectors)

The LHC beams cross paths at four interaction points along the ring (see Fig. 3.2), where four large complementary detectors are placed. The **CMS** detector [138] is the heaviest one, and its 14000 tonnes are disposed in concentric cylindrical layers around the beam, with a close to  $4\pi$  solid angle geometric coverage. A more detailed description of this general-purpose experiment is given in section 3.2, along with some explanations of the technologies of the sub-detectors (some of them being also used in the other experiments). As a spoiler and to compare to the

numbers given in this section, the low- $p_T$  muon threshold goes from 1.2 to 3.5 GeV depending on  $\eta$ . In Run I data, the resolution on charged track transverse impact parameter to the PV was 80  $\mu\text{m}$  (160  $\mu\text{m}$ ) at  $p_T = 1 \text{ GeV}$  and 22  $\mu\text{m}$  (30  $\mu\text{m}$ ) at  $p_T = 10 \text{ GeV}$  in the barrel (endcaps) region. The transverse momentum resolution was 1% in the barrel and 2.5% in the endcap region, for  $p_T = 1 - 10 \text{ GeV}$  (worsening to 4 and 20% at  $p_T = 100 \text{ GeV}$ ).

The largest LHC detector is **ATLAS** (*A Toroidal LHC Apparatus*). Its layout is shown in Fig. 3.4. It is also a general-purpose cylindrical detector, of 46 m length and 25 m diameter, weighting 7000 tonnes. It has a similar physics reach than CMS, but with many differences in design. This redundancy, although with different experimental methods, was fundamental to have confidence in the Higgs boson discovery, and will be as important in case of a future discovery.

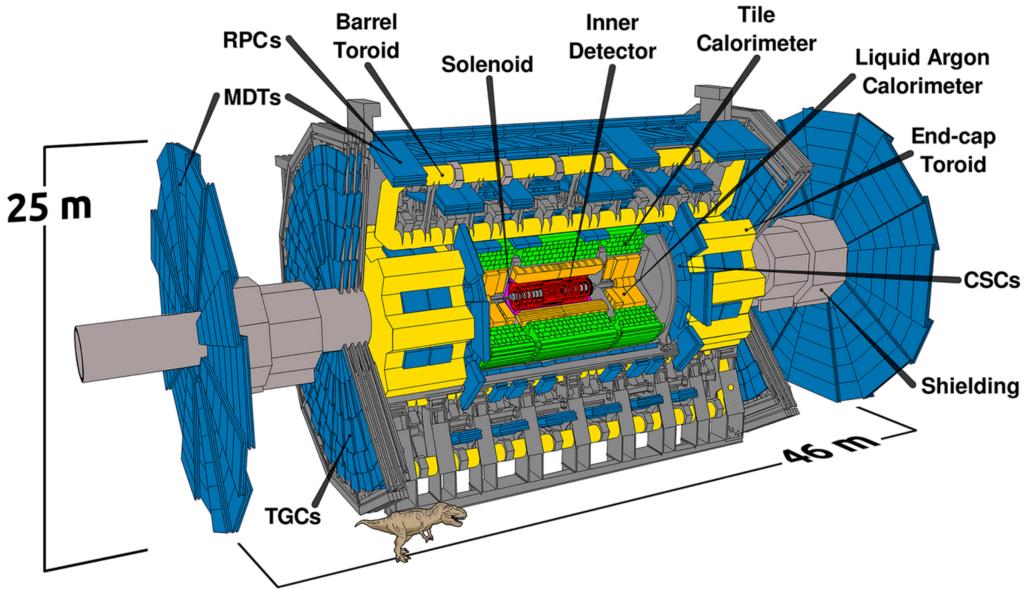


FIGURE 3.4: Layout of the ATLAS detector at the LHC.

ATLAS tracking system contains three sub-systems based on silicon technologies as in the CMS tracker, and an outermost Transition Radiation Tracker based on transition radiation (to differentiate electrons from pions) and on straws of ionising gas. In the EM calorimeter, the accordion-shaped cells sample with liquid argon the energy that was deposited in the interspersed lead and steel absorbers. The steel of the huge hadronic calorimeter absorbs energy, that is then measured with scintillating tiles. There are two magnet systems to measure the momentum of particles: first a 2 T superconducting solenoid only 5 cm thick is placed between the tracking system and the calorimeters. Then the huge toroidal field, created by the eight 25 m long superconducting coils so emblematic of the experiment, surrounds the whole detector with a non-constant field reaching up to 3.5 T. The two endcap toroidal magnets (also each made of 8 coils) extend the toroidal field lines at forward rapidities. The muon detection system is composed of sub-detectors placed both inside and outside of the toroidal magnet; they use four different technologies, including the three used by CMS (RPC, CSC, and drift tubes, see section 3.2.3). *Small muon wheels* are currently added (during LS2) to the endcap muon system. Four sub-detectors close to the beamline help monitoring the luminosity and measuring forward processes. The resolution of the transverse impact parameter for tracks is typically from 60  $\mu\text{m}$  at  $p_T = 1 \text{ GeV}$  to 25  $\mu\text{m}$  at  $p_T = 10 \text{ GeV}$ , in 2015 pp collisions at 13 TeV [139]. The muon transverse momentum resolution was about 4% at  $p_T = 10 \text{ GeV}$  and 10% at  $p_T = 4 \text{ GeV}$  in Run I [140]. ATLAS covers roughly the pseudorapidity range  $|\eta| < 2.5$  for all detected particles, but also

measures the energy of forward particles, to be close to hermetic as CMS. The low- $p_T$  threshold for detecting muons is about 3 GeV.

The **LHCb** detector is a forward single-arm spectrometer covering the pseudo-rapidity range  $2 < \eta < 5$  (where the Lorentz boost helps the measurement of the displacement of B mesons, hence the detector name), dedicated to heavy flavour physics, including asymmetry measurements for CP (charge-parity symmetry) violation studies. To this purpose, it has the interesting feature of reversing its magnet polarity to cancel most systematic uncertainties in asymmetry measurements. It also has excellent momentum and vertex resolutions, better than ATLAS and CMS, and can differentiate protons, kaons, and pions, which is impossible with ATLAS and CMS (until Run III). However the clear disadvantages of LHCb compared to these general-purpose detectors are the limited  $\eta$  acceptance (notably preventing any missing energy measurement) and the high sensitivity of the sub-detectors to radiation and occupancy. This is the main reason why the luminosity delivered to the interaction point is at least ten times smaller than to ATLAS and CMS, and why only the most peripheral PbPb collisions have been studied (only in Run II), with centrality  $> 50\%$ . This limit might be improved to centrality  $> 30\%$  in Run III thanks to detector upgrades. A transverse view of this detector is shown in Fig. 3.6.

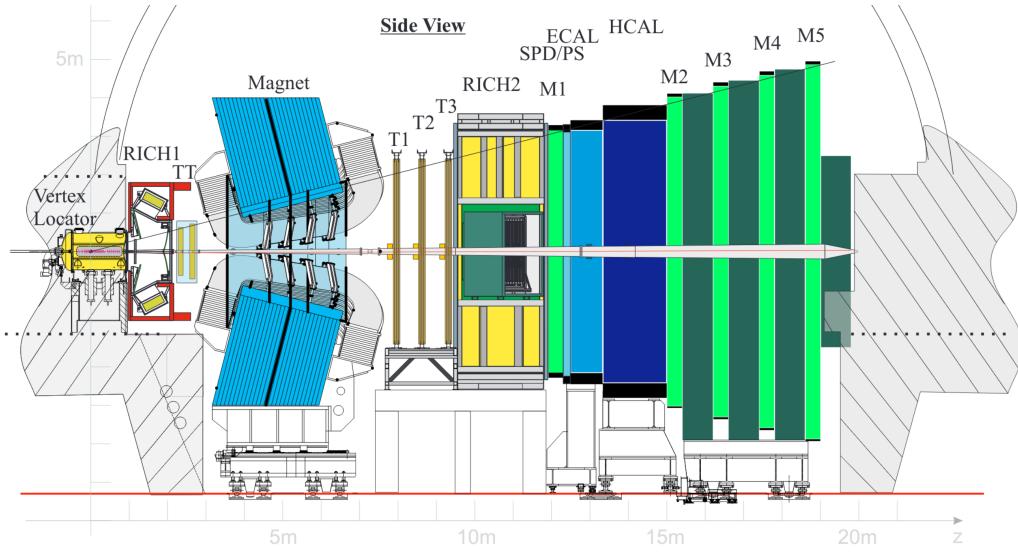


FIGURE 3.5: Layout of the LHCb detector at the LHC in Run I.

The most upstream (closest to the interaction point) LHCb sub-detector is the mobile *Vertex Locator* composed of silicon strips set in perpendicular directions. Then the first *Ring Imaging Cherenkov* (RICH) detector identifies particles at low momentum through CHERENKOV cones. Other silicon strips are then placed right before the superconducting magnet of integrated magnetic field 4 Tm. Then come three 4-layer *T-stations*, composed of silicon strips, and straw tubes for those further from the interaction point. Then the second RICH identifies particles at higher momentum than RICH1. Muon stations (one before the calorimeters, and four downstream of them) use iron absorbers and multi-wire proportional chambers. The calorimeters alternate lead and iron absorbers with scintillating material. In Run I, the momentum resolution of *long tracks* (reconstructed using all tracking-related sub-detectors) was about 0.5% for  $p < 20$  GeV and < 1% at all momenta; and the resolution of the track transverse impact parameter went from 50  $\mu\text{m}$  at  $p_T = 1$  GeV to 20  $\mu\text{m}$  at  $p_T = 10$  GeV [141]. It also has an average misidentification rate of pions into kaons of 3% for a kaon efficiency of 90%, over a large  $p_T$  range.

Lastly, the **ALICE** (*A Large Ion Collider Experiment*) detector is dedicated to heavy ion studies, though it also studies pp collisions, notably for topics concerning low- $p_T$  behaviour, total cross sections, or particle multiplicity. A global view of this detector is shown in Fig. 3.6. The

strongest advantage it has over heavy ion analyses at CMS and ATLAS is its ability to reconstruct low- $p_T$  particles and to identify them (including pion, kaon, and proton discrimination). This allows one to measure for instance D and  $J/\psi$  mesons down to null  $p_T$ , *e.g.* with muons reconstructed with reasonably high efficiency down to  $p_T = 1\text{ GeV}$ . ALICE mainly reconstructs the bulk of charged particles and electromagnetic probes in the central rapidity region (typically  $|\eta| < 1$ ), and muons in the forward region  $-4 < \eta < -2.5$  to measure heavy-quark resonances with a good momentum resolution.

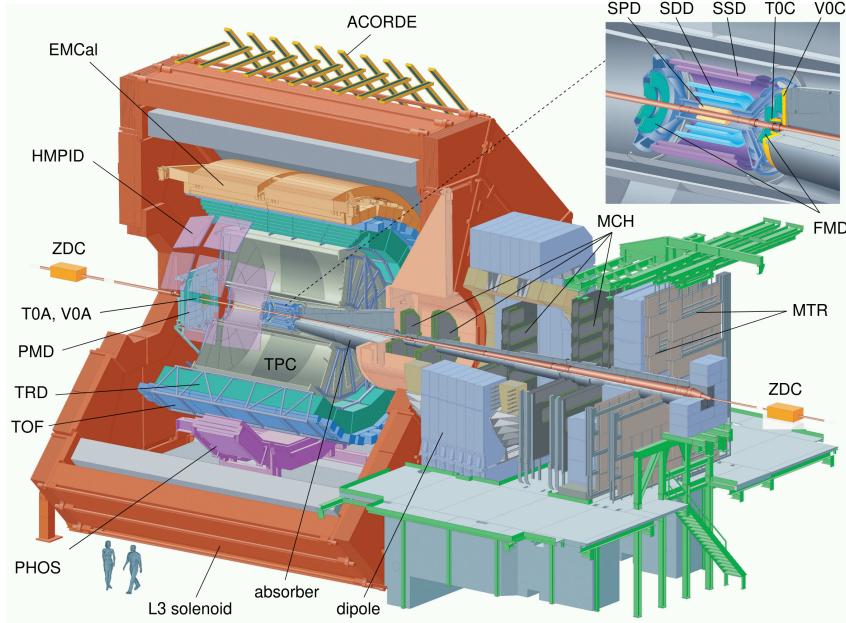


FIGURE 3.6: Layout of the ALICE detector at the LHC in Run I.

In ALICE, the forward muons first pass through composite absorbers, then through a dipole magnet of 3 Tm integrated field and ten cathode strip tracking stations placed before and after the magnet. After a second absorber in iron, RPCs are used for muon identification and triggering. In the central barrel, the *Inner Tracking System* (ITS) contributes to a good momentum resolution and has good vertex reconstruction capabilities. It also identifies non-relativistic particles, thanks to six silicon pixel and drift layers. The *Time Projection Chamber* (TPC) covering the barrel region is emblematic of ALICE; it is designed to track and identify (with  $dE/dx$  measurements, effective up to  $p_T = 50\text{ GeV}$ ) particles in environments of a few thousand tracks, thanks to electrons from ionised gas drifting towards the readout endcap plates. The outermost layer is a large solenoid magnet of 0.5 T field, bending the particles for the tracking system to measure their momentum. The TPC and the ITS are very lightweight, which limits the multiple scattering with detector material, hence improving the tracking (see section 3.2.2). However, it is a slow detector as the electrons take about  $100\mu\text{s}$  to drift to the readout system. Even if there are electrons from tracks from a few (about 5) collisions drifting at the same time in the TPC, this is to be compared with the bunch separation of 75 ns in 2018 PbPb collisions. It is hence the limiting factor for the frequency of collisions (proportional to the luminosity) ALICE can record. Three other particle identification systems are placed in the barrel, based on time-of-flight, transition radiation, and Cherenkov radiation techniques. Scintillating crystals (PHOS) and a Pb-scintillator sampling calorimeter help detect electromagnetic probes and jets, and contribute to triggering the events. A set of sub-systems at forward rapidities helps trigger the events and gives multiplicity information on uncovered parts of the solid angle. In Run I data [142], the charged track resolution of the transverse impact parameter was  $70\mu\text{m}$  at  $p_T = 1\text{ GeV}$  and  $20\mu\text{m}$  at  $p_T = 10\text{ GeV}$  in the ITS+TPC system in PbPb collisions, slightly worsening in pp collisions

due to pile-up. Displacement from the PV cannot be measured for forward muons. The charged track momentum resolution is 0.8% at  $p_T = 1 \text{ GeV}$  and 1.5% at  $p_T = 10 \text{ GeV}$  in the ITS+TPC system. The mass resolution of dimuons from  $J/\psi$  mesons in the forward muon system is about 70 MeV.

### 3.1.3.2 Outstanding discoveries in pp collisions

The most famous accomplishment of the LHC was to discover the long-awaited **Higgs boson**, whose mechanism was predicted in 1964 in three different papers from six different authors [114–116]. On July 4<sup>th</sup>, 2012, two seminars from the ATLAS and CMS collaborations followed by a press conference announced the discovery of a boson at 125–126 GeV with properties similar to the predicted Higgs boson, soon resulting in publications [143, 144]. This discovery came from only one and a half years of data-taking for physics. Along the history of modern physics, the 48 years lapse between the theoretical prediction and its experimental proof has probably been only beaten by the discovery of gravitational waves in 2015 by the LIGO and Virgo collaborations [145], a century after their prediction by Albert EINSTEIN! Those two discoveries made the headlines of newspapers worldwide, more than any other in fundamental physics. With the still-accumulating data, the Higgs boson properties, in particular its couplings to other Standard Model particles, were measured with increasing precision. All of its properties were found consistent with the Standard Model predictions. François ENGLERT and Peter HIGGS did not wait much longer for their Nobel Prize, in 2013 – it was already too late for their collaborator Robert BROUT, who passed away in 2011.

**Tetraquarks and pentaquarks** are arrangements of quarks into bound states that were in principle allowed by the original quark model, but never actually observed until the efforts from the Belle collaboration at KEKB accelerator and the LHCb collaboration at LHC, which discovered and studied many of these exotic hadrons. The first pentaquark state was observed in the LHCb detector [146], as well as the first tetraquark containing two charm quarks, with quark content  $cc\bar{u}\bar{d}$  [147]. The observation of these states is a great step for the understanding of the strong force; it remains to be understood if these bound states are compact states of tightly bound quarks or simply two mesons (or a meson and a baryon) weakly bound by a remainder of the strong force, as VAN DER WAALS forces bind neutral atoms into molecules. As mentioned in the coming section, the nature of tetraquarks might be understood via heavy ion collisions.

### 3.1.3.3 Main feats in heavy ion collisions

Somewhat less publicised to the large public than the Higgs boson were the discoveries in heavy ion collisions at the LHC. The main ones are presented in this section – whose perspective is the first observations of some phenomena, so the reader should be warned that the given references (intentionally) mostly correspond to the results that were published the earliest.

#### Flow

The strong positive elliptic flow (see section 1.4.2) is a typical signature of the QGP. Another azimuthal anisotropy, called **triangular flow** (the Fourier coefficient  $v_3$  of the  $\phi$  distribution of particles), was initially thought not to be significant because it does not respect a  $\phi \rightarrow \phi + \pi$  symmetry as the even coefficients do ( $v_2, v_4, \dots$ ). However, it has later been understood as a sign of fluctuations in the initial distribution of nucleons. It has been observed first in 2011 in the ALICE detector [148]. The correlations from  $v_n$  coefficients are most easily measured on two particles, but measuring *e.g.* the  $v_2$  coefficient from **correlations of 4, 6, or even 8 particles** can help reject contributions that do not come from the flow of the medium, but rather from 2-particle effects such as remainders of jet correlations. This was measured by many analyses at LHC, notably with the ALICE detector; they confirmed that the observed azimuthal asymmetries are indeed partly due to pure flow effects.

Elliptic flow can be understood as the correlation of particles such that pairs of particles of an event are preferentially separated by  $\Delta\phi = \pi$  (with respect to the segment joining the two nuclei centres) in the azimuthal angle. In addition, there can be long-range correlations along pseudo-rapidity, at  $\Delta\phi = \pi$ , and even at  $\Delta\phi = 0$ , compared to a jet of the event. This so-called ***ridge*** in the 2D ( $\Delta\eta, \Delta\phi$ ) correlations plane was measured in heavy ion collisions before the LHC, and was thought to be associated with final-state effects in the QGP, linked to radial collective motion – though initial-state correlations could also be responsible [149]. What a surprise it then was for the CMS Collaboration to observe a similar phenomenon in pp collisions with high track multiplicity [150]! Even if it concerns only a fraction  $10^{-4}$ – $10^{-3}$  of the highest-multiplicity pp events, it is still a puzzle today in the community. It actually opened a new approach of the field, that aims at finding continuity of the heavy ion observables between **systems of all sizes**; typically in pp, pPb, and PbPb collisions, but also in symmetric systems with smaller nuclei as XeXe or the OO collisions that are considered for future LHC runs. Since then, the strangeness enhancement, also initially thought to signify the presence of QGP, have also been observed in high-multiplicity pp collisions by ALICE [151].

### New probes

The LHC has also set forth new probes in PbPb collisions, allowing for new heavy ion observables to be tested. Let us start with those that were clearly made accessible by the jump in energy from RHIC to LHC. The electroweak probes are particularly interesting to study the initial state of the collision, and provide a reference modification (compared to pp collisions) that cannot be caused by strong interactions with the medium. They are in that sense the standard candles of QGP studies. The modification in PbPb collisions of **isolated photons** [152] and **W bosons** [153] was measured very early from the first CMS heavy ion data, as was the one of Z bosons with ATLAS and CMS [154, 155]. This excludes the low- $p_T$  photons, which were studied at RHIC [156], and also more precisely with ALICE [157]. In this regime, it is not a standard candle any more, but gives a hold on the average temperature of the medium (only slightly higher at LHC,  $T \sim 300$  MeV, than at RHIC,  $T \sim 240$  MeV): the low- $p_T$  excess of photons in PbPb collisions can indeed be interpreted as its thermal radiation. Concerning the initial state of the collision, the distribution of partons in the nucleus before the first partonic interactions is encompassed in the **nuclear parton distribution functions** (nPDF, see section 1.3.2). Data at the LHC constrained them in previously unexplored parts of the  $(Q^2, x_{\text{Bjorken}})$  phase space, mainly with measurements of the charge and forward-backward asymmetries of the W boson in pPb collisions [158]. W bosons at backward or forward rapidities come from either the proton or the lead nucleus, which gives a way to compare PDFs and nPDFs.

In LHC experiments, the first **jets** were reconstructed in heavy ion collisions. The jet studies first addressed jet quenching through the momentum imbalance of dijets [159, 160]. Then the jet shapes were studied, showing that jets get wider in the QGP due to the redistribution of the momenta of the jet constituents interacting with the medium [161]. The nuclear modification factor of jets was then measured [162, 163]. The jet substructure studies are now flourishing (see *e.g.* Ref. [164]), to understand details of the first parton splittings in the QGP. These jet-level measurements provide more information on the partonic energy loss processes than the previous measurements of single high- $p_T$  hadrons (in experiments that had no jet reconstruction capabilities), which are notably affected by fragmentation functions (the probabilities for partons to hadronise in a given hadron).

The LHC has also brought to light new heavy flavour probes, that are of most interest to this thesis.  $D^0$  mesons have been measured at RHIC [45], though at much lower  $p_T$  than what LHC experiments can reach; and the latter were already showing D results at the time of this measurement. The low- $p_T$  shape of the  $D^0$  nuclear modification factor still tells a lot about the expansion of the medium, as it rises above 1 at very low  $p_T$  from collective effects, and then

decreases at medium  $p_T$  because of energy loss processes.<sup>1</sup> However, b-hadrons at RHIC were only indirectly seen through a mix of charm and beauty hadrons (typically with displaced leptons), dominated by charm at low- $p_T$ . The **b- and c-hadrons** were hence differentiated only at LHC, in particular thanks to the CMS detector; first inclusively with non-prompt  $J/\psi$  mesons (coming from b-hadrons) [78] and b-tagged jets [165], then in exclusive decays with the  $D^{0,\pm}$  [77, 166] and the  $B^\pm$  [76] mesons. These probes shed light on many hypothesised phenomena, as the dependence of the radiated energy loss on the mass (*a.k.a.* the *dead-cone* effect) and colour charge of the parton parent of the observed hadron (see section 2.3). The mass dependence of the flow, linked to how heavy quarks inherit collective flow from light quarks in the expanding medium, has also been studied via the azimuthal anisotropy of open charm mesons [167]. Going further, LHC experiments were able to investigate how the medium modifies the hadronisation of heavy quarks by studying more exotic mesons, that mix non-valence quarks.  $D_s$  mesons were first observed by ALICE [168], the  $B_s$  modification factor was first measured by CMS [74], and last but not least, the  $B_c$  mesons were studied in this thesis with the CMS detector [1]. CMS even observed the  $X(3872)$  [169], whose nature could be elucidated in heavy ion collisions: a more loosely-bound state (close to a  $D^0 - \bar{D}^{*0}$  molecule, compared to a compact tetraquark state) would mean that it dissociates more easily in the medium due to its larger size (though it might also be associated to a higher coalescence rate or lower the equilibrium limit in transport models, which have opposite consequences).

### Standard quarkonia

The suppression of the excited states of heavy quarkonia can constitute a ‘thermometer’ of the QGP, as section 2.4 explains. This **sequential suppression** has been first observed in bottomonium states by CMS [82]: the suppression of  $\Upsilon(nS)$  states is smaller for higher  $n$ .

The **recombination** of uncorrelated charm quarks in the medium is a central phenomenon motivating the  $B_c$  analysis carried out in this thesis, as explained in section 2.5 including a historical perspective. The set of evidences for recombination can be summarised in three observations: the comparison of the  $R_{PbPb}$  of inclusive  $J/\psi$  mesons in ALICE [93, 94] and in PHENIX at RHIC [96] both at high  $N_{\text{part}}$  and at low- $p_T$ , and the large positive  $J/\psi$  elliptic flow observed with ALICE [97].

### QED

Advances in particle physics thanks to the LHC concern mainly the strong and weak forces; however heavy ion collisions allowed to observe for the first time a phenomenon purely concerning QED: the elastic **light-by-light scattering**, consisting in two photons scattering into two photons. This quantum effect breaks the linearity of Maxwell equations yielding the superposition principle, and is rare because it contains four QED vertices at leading order. It is therefore proportional to  $\alpha_{\text{QED}}^4 \sim 3 \times 10^{-9}$ , but is also enhanced by  $Z^4$  ( $Z$  being the lead nucleus charge), making heavy ion collisions the right place to look for it. It has been observed by ATLAS and CMS (first evidence in Ref. [170]).

## 3.2 The Compact Muon Solenoid

### 3.2.1 Global view

The *Compact Muon Solenoid* (CMS) is a multi-purpose cylindrical detector placed 100 m underground, opposite to ATLAS (which is on the Meyrin site of CERN) on the LHC ring. CMS has a broad scope of physics goals. It is quasi-hermetic (covering almost a  $4\pi$  solid angle), is structured in layers of subsystems aiming at detecting different particles or particle properties,

---

<sup>1</sup>The total number of charm quarks is actually conserved, and a large majority of them hadronise into open charm, therefore the inclusive modification factor of  $D$  mesons (integrated on phase space) should be close to 1.

and has similar capabilities as ATLAS but with a much more compact design: 14,000 tonnes<sup>1</sup> are contained in a cylinder of 21 m length and 15 m diameter. Most subsystems are configured in a cylindrical layer in the *barrel* region, covering roughly  $|\eta| < 1.2$ , or an *endcap* disk, covering typically  $0.9 < |\eta| < 2.4 - 3.0$ . The  $0.9 < |\eta| < 1.2$  is hence called the *transition region*. CMS was built in fifteen longitudinal ‘slices’, to facilitate assembly; they can also be moved during shutdown to access the inside of the detector. Fig. 3.7 shows a longitudinal and a transverse section of CMS, as well as a sector of a transverse slice of the barrel where the different subsystems are broken down. Listing the main subsystems starting from those closest to the beam, CMS features a silicon inner tracker, an electromagnetic calorimeter, a hadronic calorimeter, a surrounding superconducting solenoid, and muon chambers interspersed with the return yoke of the solenoid. This section reviews the characteristics of each subsystem. The muons used in this analysis are reconstructed from the inner tracker and muon chamber subsystems, which are hence more extensively discussed in sections 3.2.2 and 3.2.3, respectively. Section 3.2.4 describes the event triggering system. Ref. [138] provides a full description of CMS.

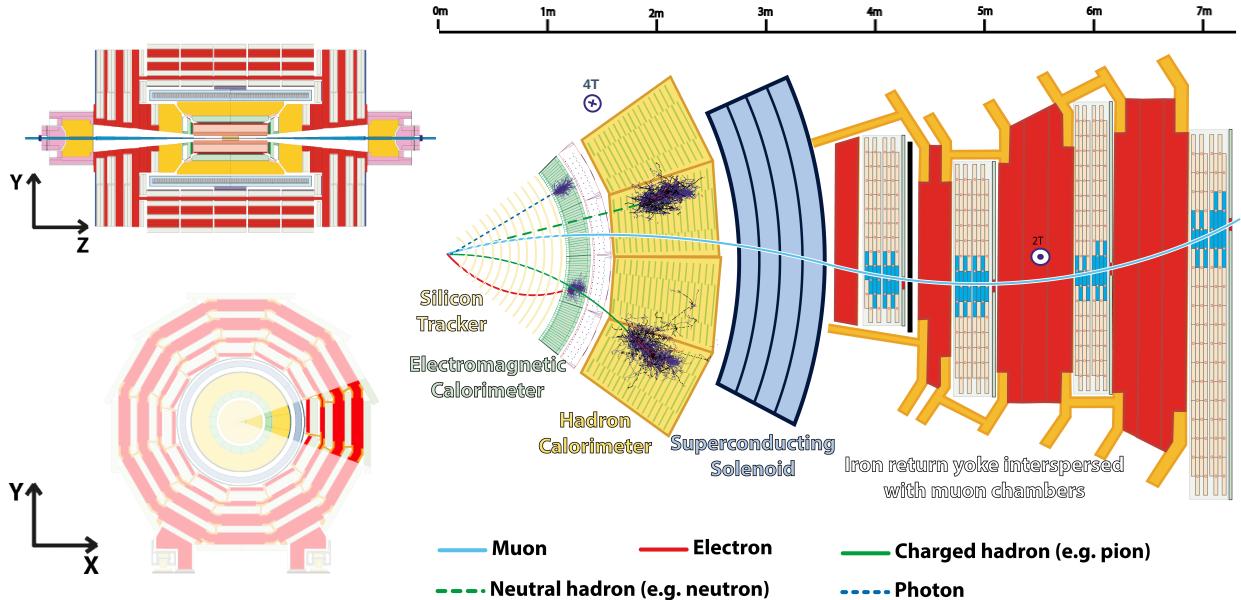


FIGURE 3.7: Various views of the CMS detector. *Top left:* Longitudinal section. *Bottom left:* Transverse slice in the barrel region. *Right:* Focus on a disk sector of the slice. The paths through each sub-detector of various detectable particles are also shown.

The cartesian coordinates are taken as an orthonormal direct basis  $(x, y, z)$ , where the *longitudinal*  $z$  direction follows the LHC beam, and the  $x$ -axis is directed towards the centre of the LHC ring. In this thesis, variables are called *transverse* when projected on the  $(x, y)$  plane. The directions of detected particles will be described with the pseudorapidity  $\eta$  and azimuthal angle  $\phi$  defined in section 1.5.2;  $\phi \in [-\pi, \pi]$  and  $\phi = 0$  on the positive  $x$ -axis, rising towards the positive  $y$ -axis.

CMS owes part of its name to the 6 m-diameter superconducting solenoid, which surrounds all subsystems of the barrel except the muon chambers. It provides a quasi-uniform **magnetic field** of 3.8 T along the beam direction, that bends charged particle so that their momentum can be measured. The winding is composed of four layers of niobium-titanium NbTi conductor, amounting to only 31 cm in radial thickness. The winding is maintained at about 7 K by liquid helium, to keep it superconducting for the 18 kA current it endures. CMS is compact notably because the layers of the iron yoke (red in Fig. 3.7) that return and close the magnetic field are

<sup>1</sup>Twice the weight of the metal structure of the 324 m-high Eiffel Tower!

spaced so that the muon detection chambers can be inserted between them. The magnet cold mass weighs 220 tons, but the 10,000-ton yoke makes most of the total CMS weight.

All long-lived particles of the Standard Model are detected by CMS, except the neutrinos whose presence is only noticed in pp collisions, through **missing energy** in the total energy balance, *i.e.* the energy of the incoming protons minus the one of all particles reconstructed in the quasi-hermetic detector and assumed to come from this pp collision. This quantity is only significant when neutrinos (or other unknown and undetected particles) carry a  $p_T$  of a few tens of GeV, as in decays of electroweak bosons; so it will be irrelevant for this thesis dealing with objects of  $p_T$  of a few GeV.

The detector part closest to the beam is the **inner tracker**. Section 3.2.2 describes it and how the amount of material of the detector affects the reconstruction of tracks (defined in section 1.5.1). All particles carrying electromagnetic charge (as the muons of interest to this thesis) can ionise the silicon atoms of the tracker components. The inner tracker hence reconstructs the **tracks of all charged particles**, and their curvature (due to the magnetic field) provides a measurement of their **momentum**.

The tracks that are not associated to electrons or muons (which can be identified through other subsystems) are either **pions**, **kaons**, or **protons**, the only remaining long-lived charged particles of the Standard Model. About 80 – 90% of the tracks are pions, as they have the smallest mass and are produced abundantly by strong interactions (contrarily to the electrons), and there are more kaons than protons, as the latter are heavier and are subject to baryon number conservation, contrarily to mesons as the kaon.<sup>1</sup> Therefore, by default, unidentified tracks are attributed the charged pion mass. At variance with other types of tracker, the CMS tracker does not discriminate kaons, protons, and pions; except through energy loss in the tracker material  $dE/dx$  for  $p_T \lesssim 2\text{ GeV}$  (corresponding to velocities significantly lower than  $c$ ), but it is rarely used and of low discrimination power in the busy PbPb environment.

To continue, I need to define the **radiation length**  $X_0$ , characteristic of a material. It is the distance for which a high-energy electron loses a fraction  $1 - 1/e$  of its energy by Bremsstrahlung. Bremsstrahlung is the photon radiation due to the acceleration (*i.e.* the deviation) induced by Coulomb interactions with the atomic nuclei of charge  $\pm Z$  of the material, and it is the dominant energy loss process for electron energies relevant in CMS.  $X_0$  is also a fraction  $9/7$  of the mean path length of an electron before emitting a photon (or equivalently the mean path length of a photon before converting in a  $e^+e^-$  pair). Finally,  $X_0$  is also used as a reference distance for multiple Coulomb (elastic) scatterings of any charged particle with the detector material. Most of the electromagnetic interactions of charged particles with the detector material is due to nuclei, with which the cross section is higher than with the electrons of the material.

Electrons interacting with matter of the detector emit photons, which themselves can ‘decay’ (thanks to a kick given to a nucleus of the detector) into a  $e^+e^-$  pair. Such conversions happen on a length scale  $\frac{9}{7}X_0$ , leading to final states that each carry about half of the energy of the parent particle, with slightly modified directions compared to the parent particle. **Photons** and **electrons** of high energy  $E$  hence undergo multiple such scatterings until daughter photons are not energetic enough to produce a  $e^+e^-$  pair, at  $E_{min} \simeq 1\text{ MeV}$ . The length of the resulting electromagnetic shower is hence roughly proportional to  $\ln(E/E_{min})$ , and the transverse spread of the shower is reasonably small, meaning the direction of the incoming electron or photon is reconstructed. **Electromagnetic calorimeters** (ECAL) intend to trigger this shower with

---

<sup>1</sup>It can be underlined here that at LHC energies, the baryon number carried by the initial protons is firmly dragged by the beams towards high rapidities, such that the net baryon density at CMS typical rapidities  $|\eta| \lesssim 2.5$  is close to null. This is also why the baryonic density  $\mu_B$ , a parameter of the phase diagram of the QGP, is close to null at LHC.

absorbing material until only low-energy photons and electrons remain, and to measure the total **energy** of these numerous secondary particles with detecting material. Part of the **neutral pions** can also be detected in electromagnetic calorimeters, as they decay with 99% branching fraction into two photons.

The CMS **ECAL** is homogeneous, in the sense that the absorbing and detecting functions are carried by the same material, namely lead-tungstate  $\text{PbWO}_4$  crystal. It is mostly transparent to light (so that low-energy photons can reach the sensitive material), has high-density (important for  $X_0$  to be small and for the shower to end within the depth of the ECAL), and is resistant to accumulated radiation. The ECAL is a 92-ton hermetic layer covering  $|\eta| < 3.0$ . It is mostly composed of 61,200 quasi-rectangular crystals of length  $220 - 230$  mm pointing to the interaction point (corresponding to  $25 - 26X_0$ ) and section  $22 \times 22$  mm to  $30 \times 30$  mm. The smallest sections are closer to the beam (about 1.3 m from it in the barrel) so that a given crystal covers a fixed solid angle. The crystals absorb the shower energy and re-emit it in the form of visible light. This light is detected by photomultipliers (avalanche photodiodes in the barrel and phototriodes in the endcaps) attached downstream of each crystal. The response of the crystals depends on temperature, so the crystals are cooled to  $18^\circ\text{C}$ . The response also depends on the radiation damage endured during a run; this is continuously monitored and corrected for by a laser. A 20 cm deep preshower, alternating lead radiators and silicon strips, is placed before the endcap ECAL to help determining the position of electron and photons and identifying neutral pion decays. The relative resolution on the energy of a showering photon or electron decreases with the energy; in pp collisions at 7 TeV [171], it was in the range 1 – 5% for a transverse energy of about 50 GeV, with worse resolution in the endcaps.

Last before the magnet comes the hadronic calorimeter (HCAL), between radii 1.8 and 3.0 m from the beam. It is essential to measure the energy from hadrons, especially the **neutral hadrons** that cannot be detected in the other subsystems. This energy completes the estimation from other subsystems, notably to estimate the missing energy from neutrinos (or unknown particles). However, hadronic calorimetry is notoriously difficult and imprecise, mostly because the fraction of the electromagnetic and hadronic components of the showers fluctuates a lot, as does the part of the absorbed energy that is undetectable, coming from nuclear recoil or late nuclear de-excitation. The shower is initiated by an inelastic interaction with a nucleus of the material, and its development is driven by the *nuclear interaction length*  $\lambda_I$  (the typical length for one nuclear interaction to occur). The latter equals 16 cm in brass. About a third of the secondary particles of the shower are neutral pions, which decay into two photons that give rise to the electromagnetic part of the shower. The hadronic part, mostly charged pions and nuclear fragments, has a much larger transverse spread compared to the electromagnetic shower.

The barrel and endcap parts of CMS **HCAL** are composed of a brass absorber, of thickness  $6\lambda_I$  at  $\eta = 0$  and about  $10\lambda_I$  for  $\eta > 1$ , and plastic scintillating plates of 4 to 9 mm thickness, whose emitted light is brought to photodiodes through external optic fibres. The upstream ECAL adds about  $1\lambda_I$  to the total thickness. To fully contain the showers at mid-rapidity, an outer calorimeter (HO) covers the range  $|\eta| < 1.3$ , just outside of the magnet which takes the role of absorber. It is composed of one or two layers of scintillator tiles. The total thickness is consequently larger than  $12\lambda_I$  on the whole covered range  $|\eta| < 3.0$ . As the fraction of undetectable energy intrinsically limits the energy resolution, the constraints on the HCAL design are mostly the mechanical stress of the heavy structure and the long-term resistance to radiation.

The forward calorimeters (**HF**) complete the covered pseudorapidity range, covering  $3 < |\eta| < 5$ . They are placed at larger  $z$  than the endcap disks. The active material is made of quartz fibres, chosen for their resistance to the very high radiation rates near the beam, and the steel absorber is 165 cm deep and has a 1.3 m radius. The fibres, placed parallel to the beam, collect part of the light from the Cherenkov radiation emitted by the charged particles – meaning that mostly the electromagnetic part of the showers is detected. The fibres are read-out in groups

called *towers*, directed towards the interaction point. The HF is surrounded by a steel, concrete, and polyethylene shielding. The transverse energy deposited in the HFs serves to monitor the **luminosity**, and to evaluate the global activity of the event which determines the **centrality** in PbPb collisions (see section 1.5.6).

The **muon** chambers constitute the outermost subdetector. They are described in section 3.2.3 along with the different algorithms for identifying muons. The advantage of CMS is its excellent muon reconstruction, giving it part of its name. The four muon stations are embedded in the flux-return yoke of the magnet. Muons are the detected particles that interact least with the detector material, and the great majority of particles are stopped by upstream subsystems, so that hits in the muon chambers strongly indicate the passage of a muon.

In pp collisions, the rate of bunch crossings reaches 40 MHz, meaning that events take place every 25 ns (and up to every 75 ns in PbPb). However, the data writing speed (5 to 8 Gb/s to the CERN storage facility called Tier0), the computing resources needed for its reconstruction, and the storage capacity, do not allow to record each bunch crossing. During the run, fast decisions (*online*) must hence be taken as to which events are stored and deserve their full data to be extracted for later (*offline*) analysis. This is the role of the **trigger** system, described in section 3.2.4.

### 3.2.2 Inner tracker and material budget

#### Tracking principles

Tracking systems are composed of many sensitive channels. Each channel can have its atoms ionised by charged particles and is connected to a separate electronic output, and its position in the detector is well-known. In a given event, a channel ionised by a crossing particle is called a *hit*. A global fit is performed on all hits of the tracker, to gather into *tracks* the hits that are consistent with the passage of one particle through the known magnetic field. Such a reconstructed track is the estimate of the true trajectory of the particle in the detector, built from consecutive discrete hits.

The fit of tracks is imperfect. Due to the finite size of the channels, some hits might be shared by multiple particles and a hit from a given track can be attributed to another close-by track, deteriorating the quality of both tracks. Fake tracks can be reconstructed from combining uncorrelated hits that are aligned by chance. This combinatorial background increases with the multiplicity of particles in the event, getting especially high in central PbPb collisions – as well as increases the duration of track reconstruction. Those fake tracks are mostly rejected thanks to the *high-purity* variable (defined in Ref. [172]), that combines low-level information from the tracking subdetectors and is optimised for the majority of CMS analyses. The positions of each channel and subdetector are also not perfectly known, so they must be corrected at software-level. This is done using particles reconstructed with high confidence in a clean environment, typically with tag-and-probe methods with Z boson decays into high- $p_T$  muons (similarly to section 7.1.1). The positions of the channels hit by these clean particles are corrected in the software to the fitted particle trajectory. How these difficulties are dealt with in the CMS tracking algorithm will not be detailed here.

#### CMS tracking system

The inner tracker of CMS [172] is composed of two technologies: a **silicon pixel detector** placed very close to the beam (covering radii from 3 to 16 cm around the beam axis), and **silicon strip modules** (from 16 to 110 cm from the beam axis). In Run I, it achieved a track reconstruction efficiency above 90% (80%) for pions of  $p_T > 0.4$  GeV in the barrel (endcap) region, and above 98% for muons of  $p_T > 0.8$  GeV, in the range  $|\eta| < 2.4$ . The resolution on charged track transverse impact parameter to the PV was 80  $\mu\text{m}$  (160  $\mu\text{m}$ ) at  $p_T = 1$  GeV and

$22\text{ }\mu\text{m}$  ( $30\text{ }\mu\text{m}$ ) at  $p_{\text{T}} = 10\text{ GeV}$  in the barrel region (endcaps region). The transverse momentum resolution was 1% in the barrel and 2.5% in the endcap region, for  $p_{\text{T}} = 1 - 10\text{ GeV}$  (worsening to 4 and 20% at  $p_{\text{T}} = 100\text{ GeV}$ ). These two resolutions are shown as a function of  $p_{\text{T}}$  for isolated muons in Fig. 3.8.

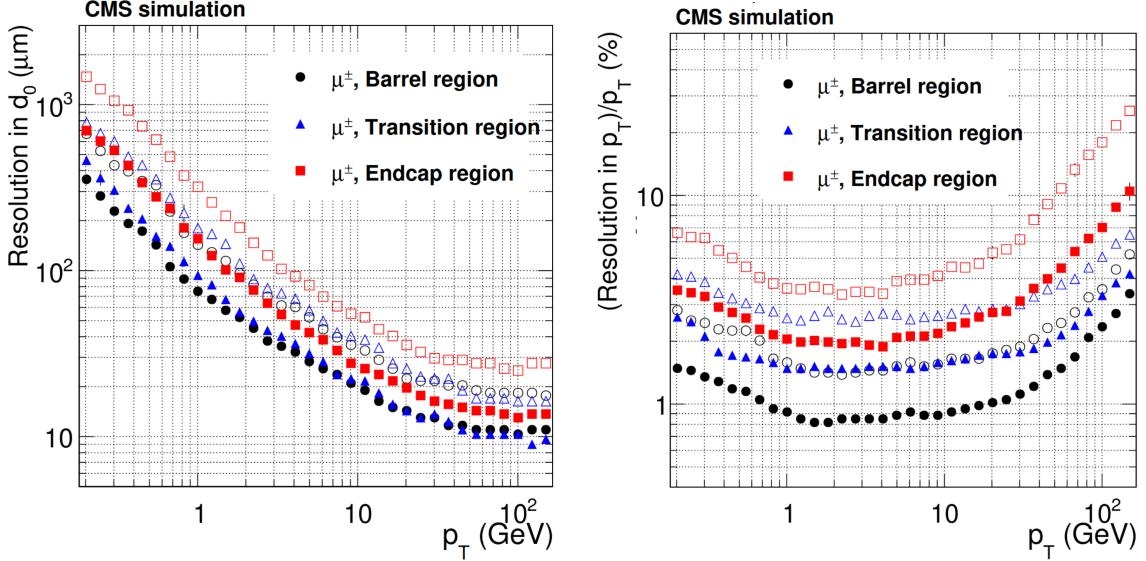


FIGURE 3.8: Resolution in the transverse impact parameter (*left*) and relative  $p_{\text{T}}$  resolution (*right*) for isolated muons in the barrel, transition, and endcap regions, as a function of  $p_{\text{T}}$ . Solid (open) symbols correspond to 68% (90%) confidence intervals. Both figures are from Ref. [172].

The pixel detector is composed of 4 barrel cylindrical layers (covering  $|z| < 27\text{ cm}$ ) and 3 endcap disks (covering  $30 < |z| < 51\text{ cm}$ ), containing a total of 120 million pixels of size  $100 \times 150\text{ }\mu\text{m}$  in the  $\phi \times z$  directions. This is a 2017 upgrade [173] from the initial tracker, that had only 3 barrel layers and 2 endcap disks; this upgrade slightly improved the performance numbers given above. The pixels have a transverse and longitudinal hit position resolution of 10 and  $20 - 40\text{ }\mu\text{m}$ , respectively. A pixel is hit when a charged particle gives it enough energy for an electron-hole pair to be created. The freed electron is collected by a small electric field towards the readout chips, laid out in a layer on top of the silicon layer.

The silicon strips are composed of the tracker inner barrel and disks (TIB and TID) covering  $16 < r < 55\text{ cm}$  and  $|z| < 118\text{ cm}$ , the tracker outer barrel (TOB) covering  $55 < r < 110\text{ cm}$  and  $|z| < 118\text{ cm}$ , and the tracker endcaps (TEC) covering  $124 < |z| < 282\text{ cm}$ . Each of these subsystems contains from 3 to 9 cylindrical or disk layers. Contrarily to pixels, the silicon strips only collect the ionised charge in one dimension. The known position of the strip gives a second direction, but to get the third spatial dimension, a second layer of strips is placed close to a given layer, with a small angle between the directions of the two sets of strips. This is the case for part of the strip layers. Within a layer, the pitch (*i.e.* the separation between two strips) goes from 80 to  $180\text{ }\mu\text{m}$ , depending on the subsystem.

The track reconstruction algorithm starts from seeds composed of two or three hits in the pixel detector, which has less channel occupancy than the strip tracker and misses less low-momentum tracks. Then hits from successive detector layers are added to the track when consistent with the propagated trajectory of the seeds, and track parameters are updated at each layer. The algorithm is based on the Kalman filter method [174], but is quite sophisticated and detailed in Ref. [172].

## Material budget

It is important for a tracking detector to contain as few material as possible, to limit:

- The electromagnetic scatterings of charged tracks with material from the detector, which deviate the track and complicate its reconstruction. The multiple **Coulomb interactions** of a charged particle of momentum  $p$  with atomic nuclei of the detector material add up to deviate the track by a typical angle

$$\theta_0 \sim \frac{13.6 \text{ MeV}/c}{p} \sqrt{x/X_0}, \quad (3.1)$$

considering velocities close to  $c$  and neglecting a term logarithmic in  $x/X_0$  (in [46], section 33 derives this from the BETHE-HEITLER formula for the energy loss of charged particles in a material). The distance travelled in the detector  $x$  must hence be reduced to limit this angle, and the radiation length  $X_0$  of the detector material should be maximised. The radiation length is higher for atoms with lower charge  $Z|e|$ ; Beryllium is for example used for the LHC beam pipe. The spread in angle reduces with the momentum of particles, because the energy loss in the material is only logarithmically dependent on the incident momentum. For hadrons, the elastic nuclear interactions also contribute to the spread in angle.

- The **inelastic nuclear interactions** with the detector, that can stop hadrons before they cross the full tracker, reducing the probability of reconstructing them. The total nuclear interaction length  $\lambda_I$  is the typical distance on which these interactions (and the elastic ones mentioned above) take place. The hadrons that interacted before the end of the tracker are less efficiently reconstructed because they left less hits in the detector.
- The **loss of energy** of particles (from the two effects above) before the calorimeters, which will therefore measure only part of the energy those particles were produced with. This is less important to the analysis of this thesis, but paramount for all analyses related to electrons, photons, and jets – even if these losses are somewhat corrected for. Concerning muons, the loss of momentum before reaching the muon chambers can slightly affect the momentum estimate of standalone muons, which has limited consequences because the momentum estimation relies on the inner tracker except at very high  $p_T$  (see section 3.2.3).

What matters most for the reconstruction of muons in this analysis is the multiple scattering in the inner tracker that limits the resolution of the muon momentum. It depends directly on the thickness of the detector, expressed in units of  $X_0$ , the characteristic length of Coulomb interactions with the nuclei of the traversed material. The inner tracker material budget is shown in Fig. 3.9. Contributions from each sub-system are separated. The thickness in units of the total nuclear interaction length  $\lambda_I$  is also shown, and adds to the electromagnetic interactions that hadrons undergo.

### 3.2.3 Muon system

#### 3.2.3.1 The apparatus

As most other subsystems, the muon system is shaped as a cylindrical barrel, and two endcap disks. The return yoke in which it is embedded closes the magnetic field lines. In the barrel yoke, where the field lines are simply parallel to the beam but opposite to the field inside the solenoid, there is a field of about 2 T that gives muon trajectories the ‘S’ shape shown in Fig. 3.7 and in the CMS logo. In the endcaps, the fields lines are more complicated and not fully contained by the yoke, so a more robust technology is needed for the muon detectors used therein. The whole system contains  $25.000 \text{ m}^2$  of detection planes, which hence have to be robust and inexpensive.

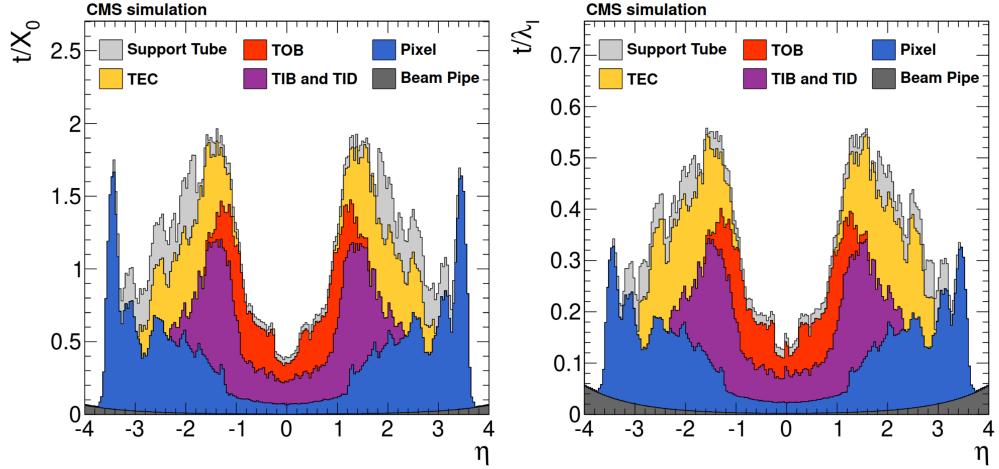


FIGURE 3.9: Total thickness  $t$  of the tracker material for a particle produced in the centre of the detector, as a function its pseudorapidity, and expressed in units of  $X_0$  (left) and  $\lambda_1$  (right). Contributions from each of the tracker subsystems, from the beam pipe, and from the support tube that surrounds the tracker are shown. Figure from Ref. [172]. The contribution from the pixel detector has been slightly reduced in the endcaps since the 2017 upgrade [173].

In between the plates of the yoke (red in Fig. 3.7) are four slots for four *muon stations*. Each station in the barrel ( $|\eta| < 1.2$ ) is composed of two or three groups of four drift tube (DT) *chambers*. They contain 172,000 sensitive wires in total, each of length 2.4 m (corresponding to the longitudinal size of a CMS slice) and width 21 mm. The latter means that an electron ejected from the Ar+CO<sub>2</sub> gas mixture of the tube by a passing muon takes up to 380 ns to drift to the collecting cathode. The drift time is measured to estimate where the muon traversed the wire. As for the strip tracker, drift tubes give one-dimensional position information, complemented by their known position in the detector. One or two groups of DT measure the muon coordinates in the transverse plane ( $r$  with the position of the chamber, and  $\phi$  with the position of the hit drift tube) whereas the other group provides the longitudinal position of the hit. The chambers are offset so that they slightly overlap, avoiding dead spots.

Each endcap contains four Cathode Strip Chambers (CSC) each of which contains six staggered layers. The CSC layers are multi-wire proportional chambers. The longitudinal position of the hit is given by the position of the hit chamber. The collecting strip cathodes run radially, so their segmentation measures the  $\phi$  coordinates, whereas the multiple anode wires per chamber run along the  $\phi$  direction, hence providing a radial position. The CSC chambers use a gas mixture of CO<sub>2</sub>, argon, and CF<sub>4</sub>. The DT and CSC stations provide muon tracking information, with a position resolution of the order of 100  $\mu\text{m}$ .

The muon system is central to the first-level trigger of CMS (see section 3.2.4). The information given to the L1 system must be attached to the right bunch crossing. The main goal of the Resistive Plate Chambers (RPC) is to complement the determination of the bunch crossing associated to detected activity, and to add redundancy in the muon tracking. The RPC double-gap chambers indeed have a timing resolution better than the bunch crossing time, 25 ns, contrarily to the DT and CSC. They cover the range  $|\eta| < 1.9$ . Six cylindrical plates are placed in between the DT barrel chambers, and five concentric disks are placed in each endcap. Each plate is composed of two 2 mm gas gaps (mostly freon, with isobutane and SF<sub>6</sub>), each framed by two 2 mm bakelite plates coated with conductive graphite. When a muon ionises the gas, an electron rapidly triggers an avalanche due to the electric field applied in the plates.

### 3.2.3.2 Identification and reconstruction

The performance of muon reconstruction in Run II is detailed in Ref. [175]. In the offline reconstruction, muon tracks are first reconstructed separately in the inner tracker and in the muon chambers. The inner tracker reconstruction is mentioned in the previous section. In the muon chambers, a Kalman filter technique [174] is used with input from the complete DT, CSC, and RPC information to reconstruct *standalone* muons.

The reconstruction of *tracker* muons starts from a track in the inner tracker and propagates it to the muon stations. It is identified as *tracker* muon if at least one segment (a tracklet in one muon station) matches with the extrapolated track, with a difference of position of less than 3 cm or 4 standard deviations in the better-measured transverse coordinate. Tracker muons can be reconstructed at lower  $p_T$  than with other identification methods, as they include muons that only reach the first muon station. However, they often represent a misidentified hadron rather than a true muon (see section 4.3.2), especially in events containing many inner tracks, each being a potential seed for a reconstructed tracker muon.

A *global* muon is reconstructed starting from a standalone muon, which is propagated to the inner tracker where a matching track is looked for. If matched, the two tracks are refitted into a *global* track.

Reconstructed muons are passed to the Particle-Flow (PF) algorithm, that uses combined information from all subsystems to identify all types of particles [176]. This *particle-flow* muon identification includes criteria such as isolation (from calorimeter information). It is used mostly for high- $p_T$  muons. Only *tracker* and *global* muons will be relevant to the  $B_c$  analysis.

The *best track* of a muon is the one yielding the best momentum resolution. It is usually the inner track (or the global track when the muon is reconstructed as global), that dominates the determination of the momentum because of the much higher spatial resolution of the inner tracker components. However, the momentum estimation of muons with  $p_T \gtrsim 100 - 200$  GeV starts to benefit from the standalone track in the muon chambers, thanks to the size of these detectors that make the curvature of the high- $p_T$  muon observable – contrarily to the shorter inner track, that appears too straight.

The muons identified in the muon chambers might not come from the beam collisions, but from cosmic sources. Muons interact feebly with matter, so they can cross the 100 m of rock between CMS and the atmosphere. Cosmic muons are relatively easy to reject, with cuts on the proximity of the muon to the primary vertex (see section 5.3.1). The reconstructed muons might also not represent true muons; the three types of misidentification are explained in section 4.3.2.

### 3.2.4 Trigger

The goal of the trigger system is to keep as high an efficiency as possible for the interesting physics probes while keeping the rate of triggered events below the maximum allowed by the computing installations. The trigger system is composed of a first level, L1, based on hardware processors using basic information from the calorimeters and muon stations, and a High Level Trigger (HLT) running a fast version of the offline event reconstruction. Only if an event passes L1 and HLT, the complete data from the detector subsystems is extracted; even HLT uses mostly local detector information, that is close to areas identified as active in L1.

**L1** reduces the 10 – 40 MHz bunch crossing rate to 100 kHz and takes decisions in less than 4  $\mu$ s, the time that events are pipelined before being discarded or sent to the HLT. The L1 trigger system was upgraded before Run II [177]. To be passed to HLT, the event must fulfil at least one of many possible conditions – those conditions being aimed at various physics programs. The decisions are taken by FPGA boards, that analyse coarse information collected by specifically-designed front-end electronics in the calorimeters (typically the deposited transverse energy in

a group of ECAL crystals and HCAL towers) and in the muon chambers (typically tracks built only in the muon system).

The **HLT** further reduces the rate to 1 kHz, with a latency of a few hundred milliseconds for the dedicated computing farm (32,000 CPU cores) to run a simplified version of the reconstruction algorithm and deduce the final trigger decision. The HLT tests, for each event, hundreds of paths corresponding to various physics goals, and an event is stored if it respects the requirements of at least one path. As an example, the trigger paths used in the  $B_c$  analysis (see section 3.5.1) primarily require two muons identified at L1 or during HLT reconstruction. The tests of the paths and the production of high-level objects (such as muons) are however sequential, with the fastest tasks ran first; a path is dropped as soon as one of the requirement fails.

In average, each decision should take no more than about 300 ms, but there is a large tail of high-activity events that take up to a few seconds. In the PbPb 2018 run, about  $100 \mu\text{b}^{-1}$  of luminosity were lost for muon-dedicated physics because some extreme events made the reconstruction slower than expected. The HLT muon reconstruction algorithm started from track seeds in the inner tracker, and extrapolated them to the muon chambers to find a compatible muon track. This was very slow for PbPb events with large inner track combinatorics; the solution was to start rather from muon tracks to find matching tracks in the inner tracker.

Once the event is selected, it is stored in the Tier0 infrastructure of CERN, in a dataset specific to the trigger path it fired, and undergoes a *prompt* offline reconstruction. The datasets are often reconstructed again later, *e.g.* to include better calibrations of physics objects.

### 3.3 The $B_c$ analysis strategy

Let us dive now into how the first observation of  $B_c$  mesons in heavy ion collisions is reached. Section 3.3.1 presents a global view of the adopted strategy, notably which decay channel is the most promising to reach an observation. Section 3.3.2 is a short feasibility study of the hadronic fully reconstructed decay  $B_c^+ \rightarrow J/\psi \pi^+$ . This thesis actually studies the partially reconstructed leptonic channel  $B_c^+ \rightarrow J/\psi \mu^+ \nu_\mu$ , using 2017 and 2018 data from the CMS detector. Section 3.3.3 presents the full global workflow of this analysis, and points to the explanations of each analysis block.

The complete code used for the  $B_c$  analysis presented in this thesis is available in this github repository: <https://github.com/gfalmagn/Bc-HeavyIons>.

#### 3.3.1 Global approach

##### 3.3.1.1 Chosen decay channel

As was made clear in section 2.2.2, mostly two  $B_c$  decay channels were studied experimentally: the leptonic channel  $B_c^+ \rightarrow J/\psi l^+ \nu_l$  and the hadronic channel  $B_c^+ \rightarrow J/\psi \pi^+$ . At LHC, as muons are easier to measure than electrons, the leptonic channel was only studied with a muon as the charged lepton – and what we call ‘the leptonic channel’ in the following actually refers to the muonic channel  $B_c^+ \rightarrow J/\psi \mu^+ \nu_\mu$ ; similarly ‘the hadronic channel’ refers to  $B_c^+ \rightarrow J/\psi \pi^+$ , the fully reconstructable channel with the highest known branching fraction. The leptonic branching fraction has been measured by LHCb to be about 20 times higher than that of the hadronic channel [52]. This is why it was the discovery channel at Tevatron [48]; it also justifies the use of the leptonic channel in this analysis, considering the relatively low luminosity in PbPb compared to that in pp collisions. As it would be significantly easier to analyse, the hadronic channel was however investigated, as reported in section 3.3.2. This rapid study in pp collisions reveals a small signal with large statistical uncertainties, however consistent with the pp cross section obtained in chapter 9 via the leptonic channel. This provides a satisfying check of the overall

analysis method. The PbPb luminosity is however too low to see the signal peak emerge above the large background.

In the leptonic channel, the neutrino is not reconstructed. Other  $B_c$  decays could contain the same visible final state, for example featuring excited charmonia decaying into a  $J/\psi$  and unreconstructed products. A systematic uncertainty arising from such decays is calculated in section 4.3.3, such that the cross section time branching fraction results that will be quoted correspond to the exclusive  $B_c^+ \rightarrow J/\psi \mu^+ \nu_\mu$  decay, and *not* to  $B_c^+ \rightarrow J/\psi \mu^+ X$  decays where  $X$  designates any set of unreconstructed particles, even if the latter is mostly what the CMS experiment is sensitive to. In addition, the  $B_c$  lifetime (about 150  $\mu\text{m}$ ) is three times smaller than that of other  $B$  mesons, but is detectable in CMS.

Therefore, the experimental signature to look for is a displaced vertex made of three muons of total charge  $\pm 1$  and whose invariant mass is a broad distribution between  $m_{J/\psi} + m_\mu = 3.203 \text{ GeV}$  and  $m_{B_c}^{\text{PDG}} = 6.275 \text{ GeV}$  [46], the mass lacking from the true  $B_c$  mass being carried by the neutrino. Fig. 3.10 shows the diagram of this channel, as well as the shape of the sought-after signal in the invariant mass of the trimuon final state. The partial reconstruction also implies that the  $p_T$ -dependent measurements are performed as a function of  $p_T^{\mu\mu\mu}$  (the  $p_T$  of the reconstructed trimuon), which is on average lower than the true  $p_T$  of the  $B_c$ . A possible average correction of  $p_T^{\mu\mu\mu}$  to the full  $p_T(B_c)$  is presented in section 8.3; it is not applied in this analysis. The rapidity cuts on the  $B_c$  are also performed using  $y^{\mu\mu\mu}$ , the rapidity of the visible trimuon.

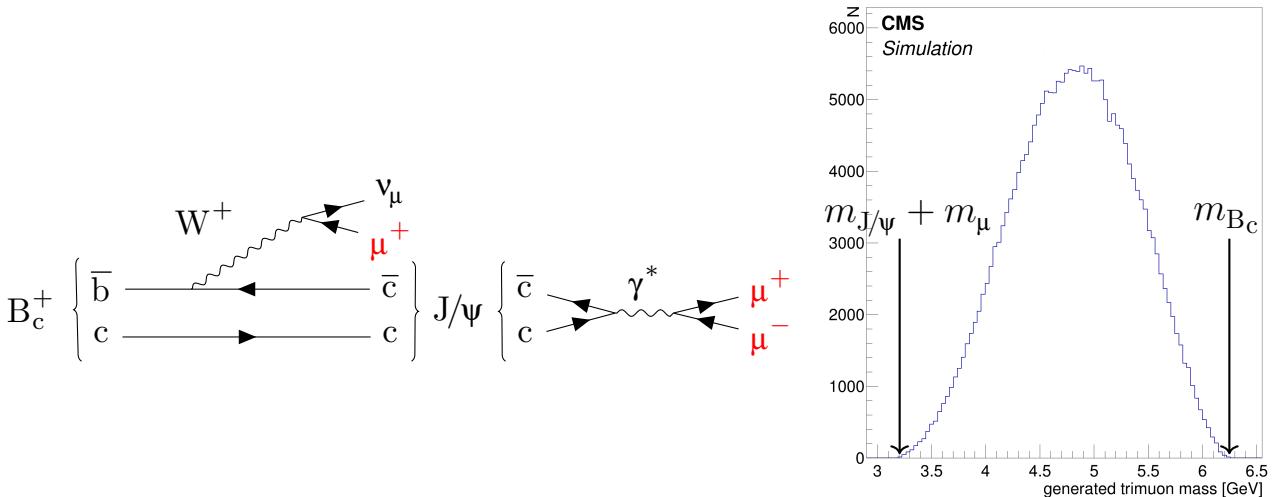


FIGURE 3.10: The  $B_c$  leptonic decay channel. *Left:* Feynman diagram of the  $B_c^+$  leptonic decay. The three detected final states muons are in red. *Right:* Simulated mass of the trimuon from the  $B_c^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) \mu^+ \nu_\mu$  decay. Only the generated  $B_c$  mesons in the trimuon kinematic region studied in this analysis, including  $p_T^{\mu\mu\mu} > 6 \text{ GeV}$  and detailed in section 5.6, are plotted. See section 3.5.2 for details on the  $B_c$  generation.

### 3.3.1.2 Finding $B_c$ mesons with CMS

CMS has many advantages for this measurement, see section 3.2. It is specialized in detecting muons, with muon chambers providing excellent muon identification and triggering. Its silicon tracker also demonstrates excellent  $p_T$  resolution, as well as secondary vertex reconstruction – the latter being essential given the finite but small  $B_c$  lifetime. It also collected a high luminosity with efficient dimuon triggers for  $J/\psi$  studies (accepting 100% of the triggering events), providing the data sample for this analysis. However, CMS has a quite high  $p_T$  threshold for the muon acceptance (around 1.5 GeV in the forward region, and 3.5 GeV for muons of pseudo-rapidity  $|\eta| < 1.2$ , see section 5.3.2). This translates into a low acceptance for signal trimuons of  $p_T \lesssim$

6 GeV (see section 7.4 for maps of the acceptance and efficiency), whereas a large majority of the produced  $B_c$  mesons lie in this kinematic region.

In this analysis, the cross sections are determined from CMS data in 2017 pp collisions at a center-of-mass energy per nucleon pair of  $\sqrt{s_{\text{NN}}} = 5.02$  TeV, and in 2018 PbPb collisions at the same energy in the 0-90% centrality range. The mean pile-up (defined in section 5.2) during data taking was 3.5 for pp collisions, and negligible for PbPb collisions. The runs associated to these datasets collected respective luminosities of  $302 \text{ pb}^{-1}$  and  $1605 \mu\text{b}^{-1}$ , for the trigger requirements used in this analysis. The latter is equivalent to about  $70 \text{ pb}^{-1}$  of  $N_{\text{coll}}$ -scaled pp collisions, so about 4.3 times more  $B_c$  candidates are expected in the pp data sample than in the PbPb one, if they are not modified by the medium and the pp and PbPb efficiencies are similar.

Let us first present a short feasibility study on the hadronic channel, before detailing the global workflow of the main leptonic channel analysis (section 3.3.3).

### 3.3.2 Hadronic channel feasibility

The hadronic channel  $B_c^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) \pi^+$  is rapidly studied, mainly in pp, to:

- independently cross-check the cross section measured in the leptonic channel;
- check that there is not a dramatic enhancement in the  $R_{\text{PbPb}}(B_c^+)$  that would give a visible peak in the hadronic channel in PbPb.
- and hence confirm that the extrapolated yields in PbPb are insufficient for this channel to be worth a full analysis;

This feasibility study performs similar first analysis steps than the leptonic channel  $B_c$  analysis: essentially a preselection followed by a BDT selection. I will hence refer to the sections of the main analysis description where these steps are detailed. I thank Natalie BLOT for her contribution to this study.

The **0niaTree** procedure (that extracts to a tractable format preliminary dimuon or trimuon candidates, see section 3.5.3) is adapted to reconstruct dimuon+track  $B_c$  candidates. Similar cuts than for the main pp preselection (shown in section 5.4 and Table 5.1) are used, except of course all cuts related to the muon identification of the third muon (not coming from the  $J/\psi$  decay). Instead, the track combined with the  $J/\psi$  is required to pass the *high-purity* requirement (see section 3.2.2), and the part of the muon identification (detailed in section 5.3.1) that is related to the inner track. This track is attributed the pion mass. In addition, we require this track to have  $p_T > 0.5 \text{ GeV}$  in pp and  $p_T > 0.7 \text{ GeV}$  in PbPb (the cut being tighter in PbPb due to the higher track background). Another difference in the selection is that all variables related to the secondary vertex (vertex probability, pointing angle  $\alpha$ , lifetime significance; see section 5.4) use a vertex fitted with a kinematic constraint that forces the dimuon to have the nominal  $J/\psi$  PDG mass. This improves the resolution on the candidate mass, and on the vertex-related variables. The last differences are that the 3D lifetime significance is required  $> 1$ , and the 2D and 3D pointing angles are required to be  $< 0.6 \text{ rad}$  (tighter than the leptonic channel cut, because there is no unreconstructed neutrino taking energy from the decay).

A signal  $B_c$  MC sample is produced with the hadronic decay, in pp (similarly to the MC simulation of the main analysis, see section 3.5.2). A MC simulation was not produced in PbPb, to limit the resources taken for this rapid study. The pp signal MC distributions will hence be used as approximate PbPb distributions. The pp simulation and the PbPb data are though expected to differ, due to the smaller reconstruction efficiencies in PbPb and the potential medium modification; this makes the PbPb study only a coarse first look. The MC sample is scaled as in section 3.5.4, using the branching fraction of the hadronic channel. In PbPb, it is scaled to the PbPb luminosity, assuming there is no modification of  $B_c$  in the QGP.

Signal and sideband regions are defined with respect to the dimuon+track invariant mass, considering an expected peak around the  $B_c$  mass  $m_{B_c,PDG} = 6.275$  GeV. The signal region is  $|m_{J/\psi+\pi} - m_{B_c,PDG}| < 0.13$  GeV, and the background sideband region is  $0.17 < |m_{J/\psi+\pi} - m_{B_c,PDG}| < 0.30$  GeV. A  $J/\psi$  mass constraint is included in the vertex fit, so the dimuon mass is close to the  $J/\psi$  mass due to the used vertex fit probability cut. In pp, there are 219k preselected candidates in the data signal region, and 125 expected signal candidates from MC. In PbPb, there are 417k preselected candidates in data, and 27 expected signal candidates from MC (assuming  $R_{\text{PbPb}}(B_c^+) = 1$  and that the PbPb efficiencies are similar to the pp ones).

A BDT is trained both in pp and PbPb, similarly to the main analysis BDT of section 5.7. The signal is taken from the pp MC sample, and the background is the dimuon+track mass sidebands in data, defined above. The eight variables used in the BDT training are the pion  $p_T$ , the dimuon+track vertex probability, the distance of closest approach between the two muons, the 3D pointing angle  $\alpha$  between the candidate momentum and its displacement vector, the significance of the 3D displacement, the sum (over the three daughters) of the significances of the transverse displacements from the PV, the  $\Delta R$  between the  $J/\psi$  and the pion, and finally the ratio  $\frac{\Delta R(J/\psi)}{\Delta R(\mu^+\pi) + \Delta R(\mu^-\pi)}$ . These variables are further explained in sections 5.4 and 5.7. The BDT output value in data does not significantly depend on the candidate mass. In addition, if there is a dependence, it is monotonous, such that it cannot affect the finding of a peak; in other words, there cannot be artificial peaks created by the BDT selection. The BDT distributions are shown for pp and PbPb in Fig. 3.11 for the signal MC and for the data in the mass signal region.

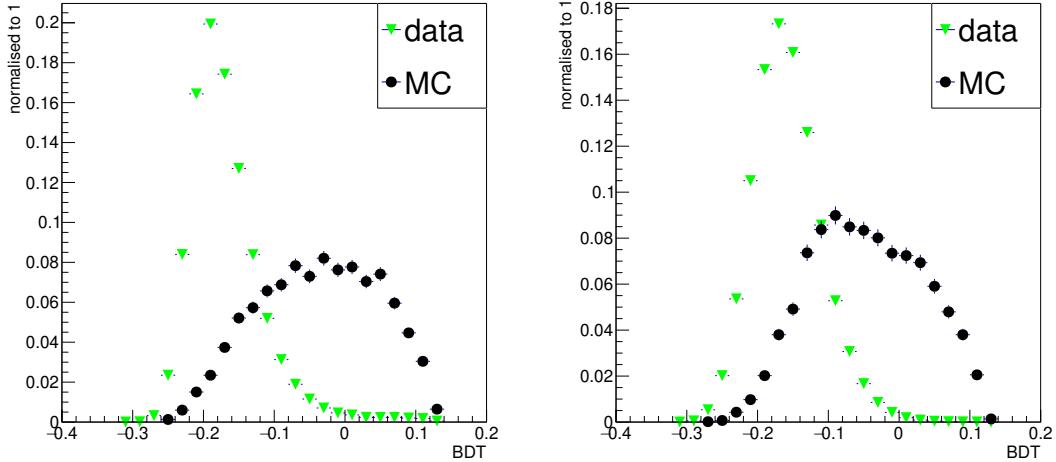


FIGURE 3.11: BDT distribution of the mass signal region in data and in the signal MC in pp (left) and in PbPb (right). The distributions are normalised to 1.

Finally, a cut on the BDT output value is applied to obtain a signal-enriched sample. This cut is chosen by optimising  $\frac{S}{\sqrt{S+B}}$ , where  $S$  and  $B$  are respectively the number of MC events and data sidebands events in the mass signal region passing the BDT cut. In pp (resp. PbPb) we obtain  $\frac{S}{\sqrt{S+B}} = 3.7$  (resp. 0.56), which is to be interpreted as an approximate significance expected for the observation of  $B_c$  mesons in this channel. It is reminded here that the expected PbPb significance depends directly on the assumption of no modification in PbPb: a very large  $R_{\text{PbPb}}(B_c^+)$  might give a significance high enough for an observation. Fig. 3.12 compares the mass distributions for data and MC in pp and PbPb collisions.

In pp, we obtain a peak that has a significance close to  $4\sigma$ . In the mass signal region, 25 signal events (from MC) and  $21 \pm 4.6$  background events (from data sidebands) are expected, and 50 are observed. No acceptance and efficiency study is performed here, but this still means

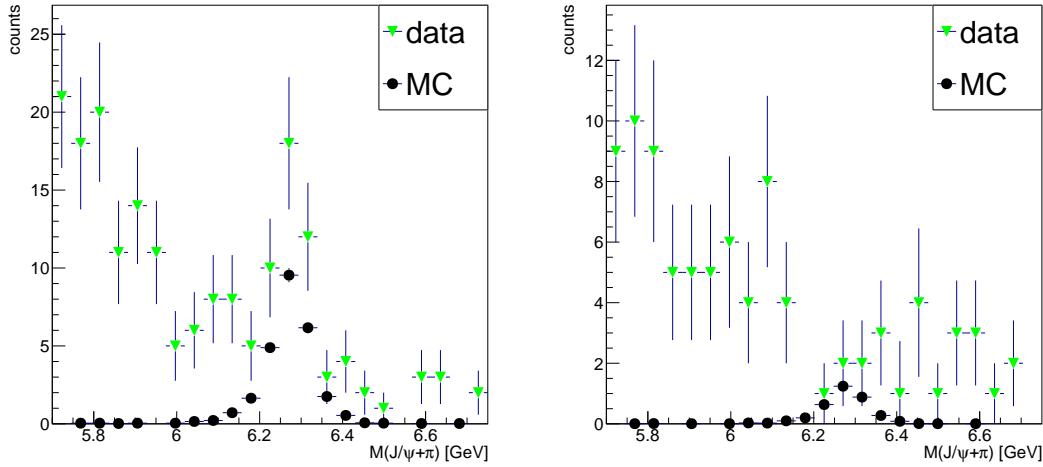


FIGURE 3.12: Dimuon+track mass of the data and the signal MC after a cut  $BDT > 0.045$  in pp (left) and  $BDT > 0.055$  in PbPb (right)

the observed signal is consistent, within uncertainties of the order of 30%, with the *a priori* normalisation of the  $B_c$  MC (calculated from previous measurements, see section 3.5.4). In chapter 9, I will show that the pp cross section measured in the leptonic channel is less than 30% different from the *a priori* signal normalisation. By transitivity, this means that the leptonic and hadronic channels give consistent pp cross sections, within large uncertainties.

In PbPb, no signal shows, with even a probable background underfluctuation: in the mass signal region, 3.3 signal events are expected (from MC, assuming  $R_{\text{PbPb}}(B_c^+) = 1$ ) and 25 background events (from sidebands), but only 9 events are observed in data. This underfluctuation could be due to the simple sidebands background estimate that assumes a linear background; this might be solved with a fit. To be conservative, let us consider a lower limit on the background of  $14 \pm 5.3$ , from doubling the content of the upper sideband that only shows 7 events. Now, one can set a rough upper limit on  $R_{\text{PbPb}}(B_c^+)$  in the kinematic range favoured by the acceptance and efficiency of CMS (typically  $p_T > 5$  GeV and  $1 < |y| < 2.3$ ).

This limit is computed using the FELDMAN-COUSINS procedure [178], with 14 expected background events and 9 observed events. It gives a number of observed  $B_c < 3.8$  at 95% confidence level, resulting in  $R_{\text{PbPb}}(B_c^+) < 1.2$ . However, this limit depends on the background estimation that has a large statistical uncertainty (if we shift down the expected background by  $1\sigma$ , to 9 expected events, the limit becomes  $R_{\text{PbPb}}(B_c^+) < 2.3$ ). It also depends on the estimation of the signal yield from MC, which was calculated here using a pp efficiency for a rapid study. The PbPb efficiency is expected to be lower, meaning the expected yield for a given PbPb modification is lower, and this upper limit shall have a larger value. However, the drop in efficiency from pp to PbPb is expected to be less than an order of magnitude (see their comparisons for the leptonic channel in section 7.4), so a strong conclusion is that, in this kinematic range, the  $B_c$  modification factor in PbPb does not exceed 10 to 20 (depending if the number of background events is shifted from 14 to the more conservative estimate of 9)).

To conclude, this coarse result is consistent with the leptonic channel result in pp, which is a satisfying check of the overall analysis method. The hadronic channel gives a too small signal significance in PbPb to be worth a full analysis, though it is useful to check in an independent manner that the nuclear modification factor of the  $B_c$  in PbPb is less than an order of magnitude higher than 1 for  $p_T \gtrsim 5$  GeV.

### 3.3.3 Leptonic channel analysis workflow

A flowchart showing the multiple steps of the analysis is shown in Fig. 3.13. Subsequent to the preselection of the sample, the core of the analysis flow features the training of a BDT discriminant, a template fit, variations of the fit method, and corrections for the acceptance and efficiency of the selection. This core analysis is performed twice, where in the second run, the  $p_T^{\mu\mu\mu}$  distribution of the signal MC is corrected thanks to a first version of the cross section measurement obtained in the first analysis run (see section 7.6). This correction has an impact mostly on the calculation of acceptance and efficiency estimations, but also on the training of the BDT and the check of its distribution. The analysis modules run after this correction are tagged as ‘second-step’ in the whole text (and are red in Fig. 3.13). Some second-step modules are ran separately in centrality bins and in  $p_T^{\mu\mu\mu}$  bins.

The data used in this analysis is triggered through a double muon requirement (detailed in section 3.5.1). Then, the necessary **MC simulated samples** are generated, for the signal  $B_c$  and for the prompt and non-prompt  $J/\psi$  processes used in background estimations. This is detailed in section 3.5.2. Then, the relevant data and MC samples are treated with the so-called **OniaTree analyser** that produces the trimuon candidates and computes and records the relevant variables (cf section 3.5.3). The **OniaTree** is also run on data in which  $J/\psi$  mesons are rotated; this is used for background estimations (section 4.4.3). A priori normalisations for the simulated samples are determined from previous measurements (section 3.5.4).

The understanding and modelling of the **backgrounds** are key to this analysis, and described in chapter 4. These backgrounds can be split in 3 categories:

- Fake  $J/\psi$ , featuring a dimuon pair that falls accidentally in the  $J/\psi$  mass region. It is estimated with the dimuon mass sidebands in data, and described in section 4.2.
- B-meson decays, in which a true displaced  $J/\psi$  meson associates with a (fake or true) muon from the same displaced decay. It is estimated from simulation in section 4.3.
- True  $J/\psi$  that associate with a (fake or true) muon from another decay or from the primary vertex. This is estimated from a data-driven method, and described in section 4.4.

A **preselection** (section 5.4) is first applied to these samples. To increase chances of observing the  $B_c$ , whose cross section rises strongly at low  $p_T$ , emphasis is put on lowering the single muon acceptance threshold in  $p_T$ . This acceptance depends on the muon selection; in particular, one of the three muons is not required to trigger the event and can be measured in a looser acceptance. The chosen strategy is justified in section 5.3. The ambiguity in the choice of the  $J/\psi$  dimuon within a trimuon candidate is treated with candidate weights explained in section 5.5. The trimuon kinematic region quoted in the results is chosen in section 5.6.

A **Boosted Decision Tree** (BDT) is then trained (see section 5.7) on the pre-selected samples, separately in each  $p_T^{\mu\mu\mu}$  bin, to compute a unique variable gathering the signal/background discrimination power from discriminant variables that were or were not used in the preselection. The samples will be binned in this BDT variable. The training is performed again in a second analysis run, where the normalisation and weighting of the signal and background samples are improved with the information of the first-step fit. Section 5.7.4 jumps a bit ahead of the coming fit by describing how the BDT distribution of the sum of second-step post-fit signal and backgrounds is then checked against data, and reweighted in the pp case, separately in each  $p_T^{\mu\mu\mu}$  or centrality bin. In case the BDT distribution is reweighted, a last iteration of the fit is then performed.

We then extract **trimuon mass histograms**, for the signal MC and for each background, for each analysis bin ( $p_T^{\mu\mu\mu}$  or centrality), and for each BDT bin. The used mass range is [3.5, 7.8], keeping a 1.6 GeV wide control region at high mass, to help study the backgrounds and stabilise the fit. These histograms are the shape inputs of the **template fit**, that is performed with the

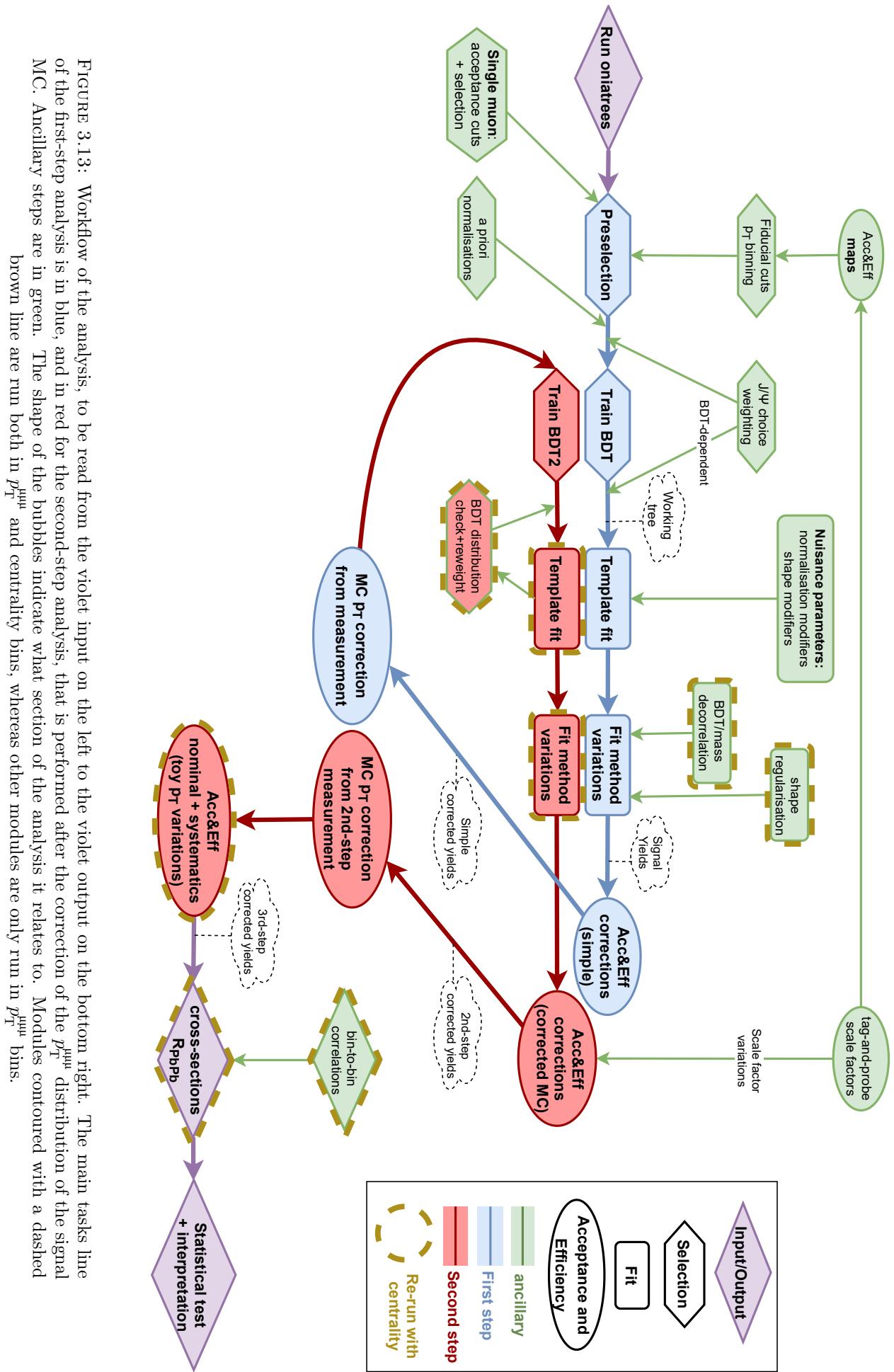


FIGURE 3.13: Workflow of the analysis, to be read from the violet input on the left to the violet output on the bottom right. The main tasks line of the first-step analysis is in blue, and in red for the second-step analysis, that is performed after the correction of the  $p_T^{\mu\mu}$  distribution of the signal MC. Ancillary steps are in green. The shape of the bubbles indicate what section of the analysis it relates to. Modules contoured with a dashed brown line are run both in  $p_T^{\mu\mu}$  and centrality bins, whereas other modules are only run in  $p_T^{\mu\mu}$  bins.

combine package in chapter 6, in six categories (two analysis bins times three BDT bins). The BDT binning provides regions with more background or more signal, to constrain separately the first and the latter. Multiple sources of systematic uncertainty are plugged in the fit procedure, in the form of nuisance parameters modifying the shape or normalisation of the templates; they are listed in section 6.2. The second-step fit is improved due to a more performant BDT discriminant than in the first step, to a more accurate  $p_T^{\mu\mu\mu}$  distribution in the signal MC, and to a BDT distribution closer to data in case it is reweighted. Section 6.4 assesses the uncertainty on the choice of fit methods by running alternative fits.

Once the signal yields are measured, they are corrected for the **acceptance and efficiency** of the reconstruction, triggering, and selection (chapter 7). The corrections are straightforward one-binned estimations in MC (section 7.3), except that the  $p_T^{\mu\mu\mu}$  distribution of MC is corrected thanks to a first-step cross section estimation, as explained in section 7.6. This correction (and the subsequent estimation of the acceptance and efficiency) is repeated after running the second-step analysis, to benefit from the improved fit results and corrected yields. The associated uncertainty is obtained through estimates of the acceptance and efficiency correction for many toy  $p_T^{\mu\mu\mu}$ -differential cross section measurements, varied within the uncertainties of the nominal measurement and resulting in varied  $p_T^{\mu\mu\mu}$ -spectrum corrections. An event-by-event correction procedure using acceptance times efficiency maps was also implemented, but was dropped because it gives too large uncertainties (section 7.5).

The **systematic uncertainties** that are not included via nuisance parameters in the fit are detailed in section 6.4 for the fit method, and in section 7.6.5 for the acceptance and efficiency correction method. A summary of all uncertainties is provided in section 8.1. Finally, the yields are normalised to obtain the cross-sections and **nuclear modification factors**  $R_{\text{PbPb}}(B_c^+)$  as a function of  $p_T^{\mu\mu\mu}$  or of centrality, and these results are interpreted (chapter 9).

## 3.4 PbPb data blinding

Data analysts might sometimes want (more or less consciously) to see something in the datasets they study, *e.g.* to obtain a positive (statistically significant) result. This might occur when an analysis method (in a broad sense) is preferred to another because it shows more statistical significance; this example could lead to an overestimation of the supposedly-observed signal. To limit the human bias linked to the choice of analysis methods, in this study, three quarters of PbPb data were hidden to the analysts (blinded) until a late stage of the analysis, when the great majority of methods were settled and approved by colleagues. Only the signal region was blinded, meaning events with relatively high BDT values and having a trimuon mass in the region where  $B_c$  mesons can be expected.

In the searches for a hypothetical signal beyond the Standard Model, with CMS or other LHC experiments, the blinding usually affects the whole signal region in data, because this potential bias is especially important to avoid when looking for new physical probes in many places at a time (this is linked to the *Look Elsewhere Effect* [179]). In this analysis though,  $B_c$  mesons are known states that are simply looked for in a busier environment. Only their cross sections are unknown, though it is probable that they are not order of magnitudes different from the values in the absence of QGP. In addition, the backgrounds were quite difficult to understand in the early studies, so being able to check their behaviour in the signal region (with only a quarter of the data) was relatively important – this applies mostly to the background from combinations of a  $J/\psi$  and a muon from different processes (described by a data-driven method in section 4.4.3). For these reasons, only three quarters of the signal region data was blinded.

In practice, a quarter of PbPb data in the signal region were given a weight 4, and the rest was given a null weight, such that the weighted data showed yields that were representative of

the unblinded ones, only with statistical uncertainties that are  $\sqrt{4} = 2$  times larger. The pp data was not blinded in such manner, because it was used as a benchmark in early stages of the analysis to understand the background sources in a data-driven way, and because less surprises were expected than in PbPb. The unblinding of the whole PbPb dataset was performed on April 13th, 2021. The only significant changes to the analysis decided after this date were:

- The use of the Kaplan spectrum to fit the measured  $p_T^{\mu\mu\mu}$  spectrum used to correct the signal MC (see section 7.6.3). The impact on the new fitted spectrum is negligible.
- In section 7.6.5, the RMS of the ratios of PbPb/pp variations is directly used as the acceptance and efficiency uncertainty on  $R_{\text{PbPb}}$ , instead of the quadratic sum of the separate PbPb and pp uncertainties. The RMS is also evaluated from all toys from the three different  $p_T^{\mu\mu\mu}$  spectrum fit functions, rather than from the set of toys based on the fit function that gives the largest RMS. These two changes have a small but non-negligible impact on the acceptance and efficiency uncertainty, increasing or decreasing them depending on the correlations between pp and PbPb spectra. These changes were however motivated and reasonable enough to be performed after unblinding.

### 3.5 Samples and normalisations

Though it is quite technical, for completeness, I give in Table 3.1, intended for internal use by CMS members, the names of all the datasets used in this analysis. For data, the names are those of the primary datasets, that receive the output of a set of triggers during data taking. The PbPb dataset strategy for the trigger of interest was changed halfway through the run, hence the two different datasets. The dataset names also indicates the reconstruction time: in April 2019 for PbPb, and in the ‘Ultra-Legacy’ campaign in August 2019 for pp, some time after the run.

For MC samples, the names indicate the main MC parameters, concerning the tuning and what tools are used for what steps. Three MC processes were generated in pp and PbPb: signal  $B_c$  events, and prompt and non-prompt  $J/\psi$  events used for background estimations (see sections 4.3 and 4.4).

TABLE 3.1: All used data and MC samples in this analysis

category	description	dataset path
PbPb 2018 5.02 TeV	data	/HIDoubleMuon/HIRun2018A-04Apr2019-v1/AOD for runs 326295-327122 /HIDoubleMuonPsiPeri/HIRun2018A-04Apr2019-v1/AOD for runs 327123-327564
	MC signal $B_c^+$	/BcToJpsiMuNu_TuneCP5_5p02TeV_BCVEGPY_pythia8-evtgen/ HINPbPbAutumn18DR-FixL1CaloGT_103X_upgrade2018_realistic_HI_v13-v2/AODSIM
	MC non-prompt $J/\psi$	/BToJpsi_pThat-2_TuneCP5-EvtGen_HydjetDrumMB_5p02TeV_pythia8/ HINPbPbAutumn18DR-mva98_103X_upgrade2018_realistic_HI_v11-v1/AODSIM
	MC prompt $J/\psi$	/JPsi_pThat-2_TuneCP5_HydjetDrumMB_5p02TeV_Pythia8/ HINPbPbAutumn18DR-mva98_103X_upgrade2018_realistic_HI_v11-v1/AODSIM
pp 2017 5.02 TeV	data	/DoubleMuon/Run2017G-09Aug2019_UL2017-v1/AOD
	MC signal $B_c^+$	/BcToJpsiMuNu_BCVEGPY_pythia8/ RunIIpp5Spring18DR-94X_mc2017_realistic_forppRef5TeV_v1-v2/AODSIM
	MC non-prompt $J/\psi$	/BjPsiMM_TuneCUETP8M1_5p02TeV_pythia8/ RunIIpp5Spring18DR-94X_mc2017_realistic_forppRef5TeV_v1_ext1-v1/AODSIM /BjPsiMM/gfalmagn-BjPsiMM_TuneCUETP8M1_5p02TeV_pythia8_ 16052018_withLambda_ptHatMin10_reco-679d0c7a470a830e53fe749c4e83e359/USER
	MC prompt $J/\psi$	/JPsiMMXXXTuneCUETP8M1_5p02TeV_pythia8/ RunIIpp5Spring18DR-94X_mc2017_realistic_forppRef5TeV-v2/AODSIM where XXX = _, _pThat-15_, _pThat-25_, _pThat-35_, and _pThat-45_
acceptance (pp 2017 conditions)	MC $B_c^+$ acceptance no kinematic cuts	/BcToJpsiMuNu_BCVEGPY_TuneCP5_Pythia8-EvtGen /RunIIpp5Spring18DR-94X_mc2017_realistic_forppRef5TeV_v1-v2/AODSIM

### 3.5.1 Data samples

The used datasets contain events triggered with a double muon criterion aimed at  $\text{J}/\psi$  mesons: `HLT_HIL1DoubleMu0` in pp, and `HLT_HIL3Mu0NHitQ10_L2Mu0_MAXdR3p5_M1to5` in PbPb. All triggered events were stored. The pp trigger only requires **two muons** reconstructed at L1 (see section 3.2.4 about L1 and HLT), of basic quality and without  $p_T$  cuts. In PbPb, the rate of triggered events needed to be lowered, so the trigger requires a L2 (HLT) muon, plus another L3 muon (of higher quality than in L2) with hit quality requirements, both without  $p_T$  cuts. Very loose criteria on the dimuon mass ( $1 < m < 5 \text{ GeV}$ ) and opening angle ( $\Delta R < 3.5$ ) are also applied. It is reminded that only two of the three muons from the  $B_c$  selected candidates are required to fire this trigger, these two muons not necessarily coming from the  $\text{J}/\psi$  decay.

When extracting events from these datasets, a JSON file masks the parts of the run in which the subdetectors used for muon physics were not fully functional. There were also short periods in which the trigger of interest was not active, so these periods are removed as well. For example, section 3.2.4 explains the loss of  $100 \mu\text{b}^{-1}$  of PbPb due to HLT timing problems; about  $15 \mu\text{b}^{-1}$  could be recovered using an additional trigger, but this low luminosity is not worth the complications. Taking into account these removed events, the luminosities are  $302.3 \text{ pb}^{-1}$  in pp and  $1.605 \text{ nb}^{-1}$  in PbPb. The latter can be compared to the  $1.8 \text{ nb}^{-1}$  of PbPb luminosity delivered by the LHC.

This luminosity will be used to normalise the corrected yields (obtained in chapter 7) in pp. In PbPb however, the corrected yield will rather be normalised with the nuclear overlap function  $T_{\text{PbPb}}$  and the total number of hadronic PbPb collisions events (see section 3.5.5.2).

### 3.5.2 MC generation

The simulated sample are generated with a Monte Carlo (MC) method (defined in section 1.5.1), primarily with the PYTHIA8 program using the `CP5` tune to CMS data [180]. The MC samples are produced ‘centrally’, meaning within the CMS-wide MC generation framework, except for one private complementary sample for the non-prompt  $\text{J}/\psi$  MC in pp (see section 3.5.2.2). For the prompt and non-prompt  $\text{J}/\psi$  MC in pp or PbPb, 40 to 60 million events are generated. For the signal  $B_c$  MC, about 3.5 to 4.0 million events are generated, which are reduced to about 500k after applying generator-level loose kinematic cuts. However, in the acceptance sample used to evaluate the fraction of generated  $B_c$  events whose muons fall in the CMS kinematic acceptance (see chapter 7), all generated events, without any cuts, are kept – but their detection and reconstruction does not need to be simulated.

To provide high-enough statistics at high  $p_T$ , some samples are biased towards generating harder processes (with higher  $\hat{p}_T$ , that represent the typical energy scale of a process). The PbPb samples use a continuous  $\hat{p}_T$  biasing, while in pp, multiple samples are produced with a different  $\hat{p}_T$  threshold (see Table 3.1). Event weights are later used to unbias the sample (in PbPb) and produce a smooth  $p_T$  spectrum (in pp)<sup>1</sup>.

#### 3.5.2.1 The $B_c$ hard process

As explained in section 2.2, gluon fusion  $gg \rightarrow B_c^+ + b + \bar{c}$  is the dominant production mode in pp collisions. This process is computed via an optimised specific Monte Carlo (MC) generator BCVEGPY2.2 [181], simulating the  $B_c$  production hard process in pp collisions at 5.02 TeV. As the production of  $B_c$  mesons is rare in PYTHIA events, it is more efficient to use a generator specialised for this process.

---

<sup>1</sup>In the pp case, discrete weights are applied: indexing with  $j$  the samples of crescent minimal  $\hat{p}_{T,j}$ , the weight applied to sample  $j$  is  $(\sigma_j - \sigma_{j+1})/N_j$ , where  $N_j$  is the total number of events between  $\hat{p}_{T,j}$  and  $\hat{p}_{T,j+1}$ , and  $\sigma_j$  is the cross section (from PYTHIA) for generating events above  $\hat{p}_{T,j}$ .

Multiple states are generated, mostly colour-singlet:

- the ground state  $^1S_0$ , the first excited state  $^3S_1$ , with spin 1 (decaying to the ground state and an unreconstructed low-energy photon);
- the four  $n = 2$  (lowest-mass) P-wave states  $^1P_1$ ,  $^3P_0$ ,  $^3P_1$ ,  $^3P_2$ ;
- the colour-octet  $^1S_0$  and  $^3S_1$  (with  $n = 1$ );
- the (colour-singlet)  $n = 2$  and  $L = 0$  states  $2^1S_0$  and  $2^3S_1$  (of significantly higher mass).

These are the eight states with lowest masses in Fig. 2.3 of section 2.2.1, plus the colour-octet for the two lowest-energy states. The excited states are considered to decay with 100% branching fraction to the ground state, and the energy-momentum carried by the by-products  $X$  of the decays  $B_c^* \rightarrow B_c X$  is ignored: in practice, the mass of the quarks forming the excited  $B_c$  are set to  $m_b^- = 5\text{ GeV}$  and  $m_c = 1.2749\text{ GeV}$ , and they are considered to be of the same nature as the ground state. This amounts to considering that the excited states have similar kinematic distributions as the ground state. This was decided after a discussion with the authors.

Only the hard subprocess is generated, and the kinematics of the outgoing  $B_c$ ,  $b$ , and  $\bar{c}$  are stored in a LHE (Les Houches Event) file. The latter is handed to the rest of the MC generation chain, including the softer part of the collision dealt with PYTHIA (see section 3.5.2.3). The initial and final state radiation effects are disabled in BCVEGPY, and rather generated with PYTHIA.

### 3.5.2.2 The $J/\psi$ hard process

In the prompt  $J/\psi$  MC, PYTHIA8 simulates the whole hard process of a  $J/\psi$  produced in the primary nucleon-nucleon collision. In the non-prompt  $J/\psi$  MC, a  $b\bar{b}$  pair is produced, and one of the quarks hadronizes into a  $B^0$ ,  $B^+$ , or  $B_s^0$ ; then only decays that contain a  $J/\psi$  are allowed for this  $B$  meson. It was found that including the  $\Lambda_b \rightarrow J/\psi X$  decays in the simulation did not change significantly the properties of the generated  $J/\psi$  (in terms of lifetime and multiplicity), so they are not included. In pp we however produce a small (2M events) private MC sample, with the same parameters as the central non-prompt  $J/\psi$  MC, except that the  $\Lambda_b \rightarrow J/\psi X$  decays are considered too, and the minimal  $p_T^\gamma$  is set to 10 GeV. This covers the case of the  $\Lambda_b$  unexpectedly giving a trimuon background with different properties than from  $B$  mesons, and adds statistics at high  $p_T$ .

### 3.5.2.3 Following generation steps

For the MC used in this analysis, the generation chain starts with PYTHIA8, except for the signal  $B_c$  MC where a LHE file is used as input. It first simulates the hard process (typically outputting a  $J/\psi$  or a  $B$  meson along with a few hard partons). Then these partons are showering and hadronise with the Lund string fragmentation model. The underlying event is also simulated (multiple parton interaction, and various soft processes). The particles with typical decay length smaller than the closest detector parts are in general decayed by PYTHIA. However, in the non-prompt  $J/\psi$  sample, the  $B$  decays to  $J/\psi X$  and other final states are performed with EVTGEN1.3 [182]; and in the  $B_c$  signal MC, EVTGEN1.3 computes the  $B_c$  decays with the latest lifetime estimation  $(5.10 \pm 0.09) \times 10^{-13}\text{ s}$  [46].

Concerning the used  $B_c$  decay model in EVTGEN, three models are tested in Ref. [60], and checked to cause variations in the trimuon mass shape that are negligible for this analysis; the model from EBERT et al. [183] is chosen here, because it is the most recent and specific to semileptonic decays.

In pp, the simulated nucleon-nucleon collision then needs to be embedded in other pp collisions taking place in the same event (to simulate the pile-up explained in section 5.2); minimum-bias collisions (*i.e.* any collision containing a hadronic interaction) simulated by PYTHIA collisions are used for these accompanying collisions. In PbPb, the nucleon-nucleon interaction is

then embedded in the other nucleon-nucleon interactions of a PbPb collision, simulated with HYDJET1.8 [184].

The generated particles then propagate and interact with the detector, and this is simulated with GEANT4 [185]. The hits left in the detector are then digitised, simulating the response of the electronics to the deposited energy. Then the same CMS software reconstruction algorithms as in data are applied to build tracks and other objects usable in analysis, stored in the same format as actual data.

### 3.5.3 The `OniaTree` analyser

The `OniaTree` analyser is conceived and mainly used for dimuon (quarkonium) analyses in the CMS heavy ion group, but I adapted it to build trimuon candidates containing a  $J/\psi$ . The core of the analyser builds a list of muons with loose selection and kinematic acceptance cuts, and then runs over all possible pairs of muons to build  $J/\psi$  candidates. I implemented an additional loop, combining each such  $J/\psi$  candidate with a third muon (called  $\mu_W$  in the rest of the text, because it comes from the  $W$  decay), to form trimuon candidates.

In addition, some variables concerning the dimuon or trimuon candidates are computed at this stage, and only the information relevant for this analysis is kept. Part of the preselection (detailed in section 5.4) is already applied at this level to reduce the computing time. For the primary datasets, the data reduction thanks to the `OniaTree` procedure is dramatic ( $> 10^4$ ), from tens of TB to a few GB. For the MC samples, the cuts, if any, are much looser, typically to allow for acceptance and efficiency estimations from uncut events.

I was responsible of developing and maintaining the `OniaTree` code during my thesis. This code constitutes modules that are plugged in the standard CMS software. These modules can be found on github (<https://github.com/CMS-HIN-dilepton/cmssw>, packages `HiSkim` and `HiAnalysis`), in the branches corresponding to the CMS software releases used in this analysis (`Onia_pp_9_4_X` for pp and `Onia_AA_10_3_X` for PbPb). All above samples are treated with this analyser.

### 3.5.4 A priori MC normalisation and weighting

An *a priori* normalisation of the MC samples is needed, notably to form input samples for the BDT training (section 5.7), but also to guide the choice of initial values of the normalisation parameters in the template fit (chapter 6). Early in the analysis, this also provided a coarse sanity check of the measured pp cross section. Previous measurements are exploited to provide these *a priori* MC normalisations.

#### 3.5.4.1 Prompt and non-prompt $J/\psi$

CMS measured the prompt and non-prompt  $J/\psi$  cross sections in pp and PbPb in a similar phase space as the  $J/\psi$  mesons of this analysis ( $|y(J/\psi)| < 2.4$  and  $p_T(J/\psi) > 6.5$  GeV) [78]. With this measurement, the MC can be normalised accurately for both pp and PbPb, and prompt and non-prompt  $J/\psi$  samples. For each of these four samples, the  $p_T$ -dependent  $J/\psi$  cross section measurement in data and the distribution of MC events are plotted (see Figs. 3.14 and 3.15). The *a priori* normalisation is derived from the data/MC ratio (scaled by the appropriate luminosity and  $J/\psi$  branching fraction). One sees that this ratio is not constant (except maybe for the non-prompt PbPb sample), so it is fitted with a function defined as a constant above a certain  $p_T$  (about 20-30 GeV), and a power law below this threshold, requiring continuity at the threshold. The fitted function is extrapolated to  $p_T(J/\psi) < 6.5$  GeV, and taken as a  $p_T(J/\psi)$ -dependent weight applied on the MC – which is hence weighted rather than simply normalised.

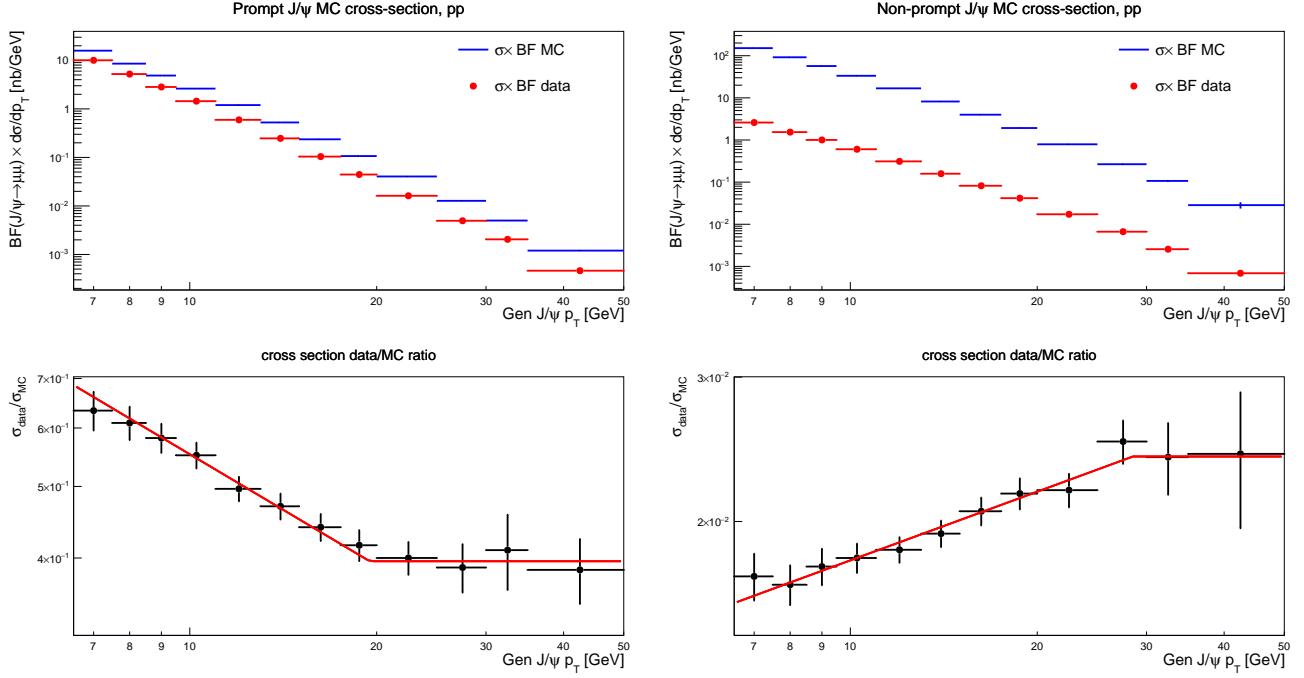


FIGURE 3.14: Weighting of the prompt (*left*) and non-prompt (*right*)  $J/\psi$  MC in pp. *Top*: Branching fraction times cross section for data and MC, versus  $p_T(J/\psi)$ . *Bottom*: ratio of the data and MC cross sections. The fit function (red line), is extrapolated to  $p_T < 6.5$  GeV and used to weight the MC. Power laws show as straight lines in the log-log canvas.

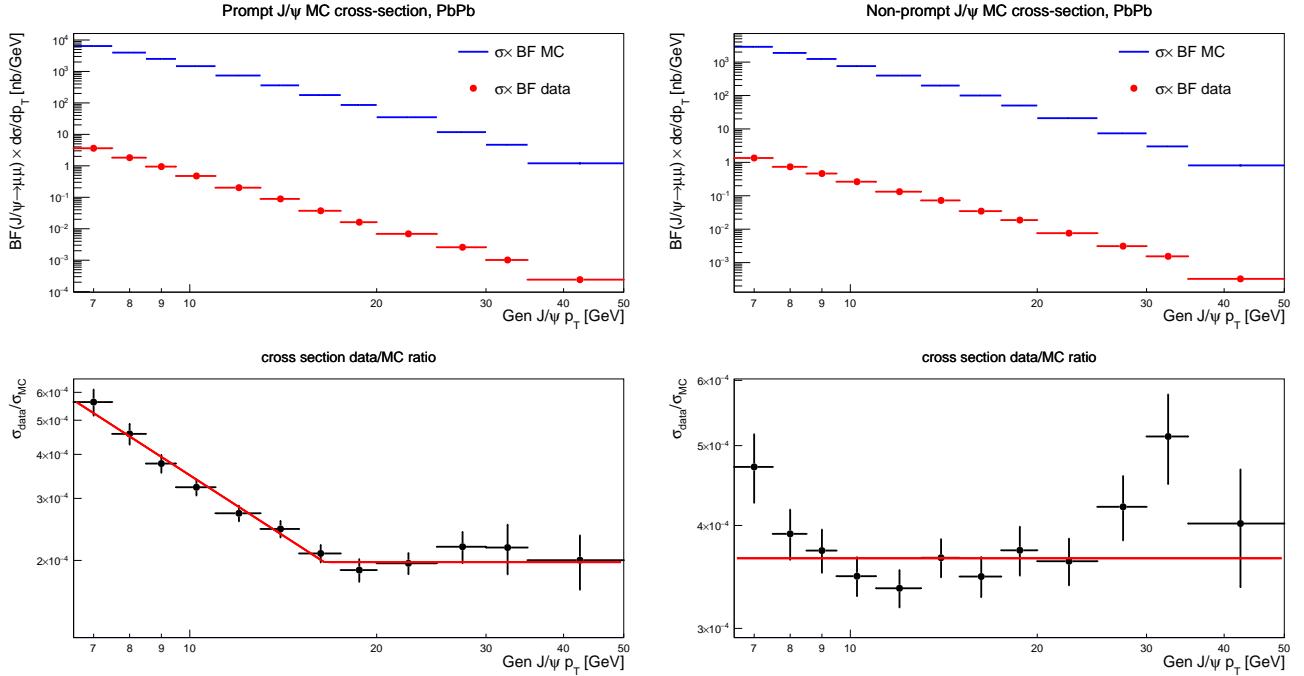


FIGURE 3.15: Weighting of the prompt (*left*) and non-prompt (*right*)  $J/\psi$  MC in PbPb. *Top*: Branching fraction times cross section for data and MC, versus  $p_T(J/\psi)$ . *Bottom*: ratio of the data and MC cross sections. The fit function (red line), is extrapolated to  $p_T < 6.5$  GeV and used to weight the MC. Power laws show as straight lines in the log-log canvas.

The ratio of the PbPb non-prompt  $J/\psi$  MC to data might not be flat as considered in its

nominal fit (Fig. 3.15, right). An alternative fit with a decreasing power law below 10.5 GeV (and a constant above) is tried, and the resulting weighting of MC is used in an alternative template fit (only the first-step fit, see chapter 6). The post-fit signal normalisation from the alternative MC weighting is only 3.5% (resp. 0.1%) different from the nominal template fit result for the first (resp. second)  $p_T^{\mu\mu\mu}$  bin. These deviations are at least an order of magnitude smaller than the total uncertainties, so they are not considered in systematic uncertainties.

### 3.5.4.2 Simulated $B_c$

We also extract an *a priori* normalisation for the signal MC in pp and PbPb that comes from previous pp measurements of the  $B_c$  cross section (meaning that this normalisation assumes no modification of the  $B_c$  production in PbPb). The signal cross section to which the MC is scaled is the weighted average of a CMS measurement at  $\sqrt{s} = 7$  TeV [186] and a LHCb measurement at  $\sqrt{s} = 8$  TeV [55, 187]. This normalisation does not enter directly in the final result: it is only used to normalise the input signal sample of the BDT training and to give an initial value to the signal normalisation parameter in the fit. In early stages, it was also used to assess the feasibility of an observation of  $B_c$  mesons in PbPb.

The  $B_c$  cross section to which the signal  $B_c$  MC is scaled is the weighted average of a CMS measurement at  $\sqrt{s} = 7$  TeV [186] and a LHCb measurement at  $\sqrt{s} = 8$  TeV [55, 187].

As these measurements are quoted in different kinematic regions, the cross section is extrapolated to the whole phase space using BCVEGPY samples without any cuts. These measurements are done in the hadronic channel  $B_c^+ \rightarrow J/\psi \pi^+$ , so I first compare the two cross section times branching fraction measurements for this channel. Then I multiply their weighted average by the ratio of the leptonic and hadronic channels, measured by LHCb [52], to get a cross section times branching fraction for the leptonic channel.

CMS measured  $\sigma \times BF_{B_c^+ \rightarrow J/\psi \pi^+} = (8.16 \pm 1.09) \times 10^{-2}$  nb at 7 TeV [186]. The fiducial cuts of this cross section are  $p_T > 10$  GeV and  $|y| < 1.5$ , so I get the inclusive (in the full phase space) cross section by dividing by the fraction of all BCVEGPY-generated  $B_c$  mesons that lie in this kinematic region, which gives

$$\sigma_{B_c^\pm} \times BF_{B_c^+ \rightarrow J/\psi \pi^+} = (2.47 \pm 0.32) \text{ nb} \quad (3.2)$$

where the error only comes from the original measurement: the uncertainty from rescaling to the inclusive cross section is neglected.

The LHCb measurement at  $\sqrt{s} = 8$  TeV [55] is:

$$\frac{\sigma_{B_c^\pm} BF_{B_c^+ \rightarrow J/\psi \pi^+}}{\sigma_B \times BF_{B \rightarrow J/\psi K}} = (6.83 \pm 0.20) \times 10^{-3} \quad (3.3)$$

which is measured in the region  $p_T < 20$  GeV and  $2 < y < 4.5$ . This is multiplied by the measurement of Ref. [187]:

$$\sigma_B \times BF_{B \rightarrow J/\psi K} = (4.30 \pm 0.30) \times 10^4 \text{ nb} \quad (3.4)$$

which is measured at 7 TeV, in the same kinematic region (except that  $p_T < 40$  GeV instead of  $p_T < 20$  GeV, but the associated difference in cross section is negligible). The variation of the  $B_c$  to  $B$  ratio from  $\sqrt{s} = 8$  TeV to 7 TeV is neglected. The product of the quantities in Eqs. 3.3 and 3.4 is the  $B_c$  cross section times branching fraction in the mentioned phase space, to be divided by the fraction of  $B_c$  generated in this kinematic region as for CMS measurement.

This assumes that the  $B$  and  $B_c$  have similar kinematics. The inclusive result is then:

$$\sigma \times BF_{B_c^+ \rightarrow J/\psi \pi^+} = (1.90 \pm 0.16) \text{ nb} \quad (3.5)$$

The results from Eqs. 3.2 and 3.5 are only 1.6 standard deviations away, which is acceptable considering that I neglected some uncertainties. The integrated cross section is then a weighted average of LHCb and CMS results:

$$\sigma_{B_c^{\pm}, tot} \times BF_{B_c^+ \rightarrow J/\psi \pi^+} = (2.01 \pm 0.14) \text{ nb} \quad (3.6)$$

at 7 TeV in pp collisions.

However, the centre-of-mass energy per nucleon-nucleon collision is 5.02 TeV in the data studied here, which shall yield a lower cross section than at 7 TeV. This is corrected with BCVEGPY, by generating two  $B_c$  samples at 5.02 and 7 TeV, resulting in  $\frac{\sigma(\sqrt{s}=5.02 \text{ TeV})}{\sigma(\sqrt{s}=7 \text{ TeV})} = 0.668$ .

Finally, I divide by the ratio of the hadronic to leptonic channels  $(4.69 \pm 0.56) \times 10^{-2}$  from LHCb [52] (whose uncertainty is included), and multiply by  $BF_{J/\psi \rightarrow \mu\mu} = (5.93 \pm 0.06) \times 10^{-2}$  [46], to get the integrated pp production cross section times branching fraction for the leptonic channel for both charges:

$$\sigma_{pp, B_c^{\pm}} \times BF_{B_c^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) \mu^+ \mu^-} = (1.70 \pm 0.23) \text{ nb} \quad (3.7)$$

The signal MC is hence multiplied by

$$\mathcal{L}_{\text{eff}} \times (\sigma_{pp} \times BF) / N_{\text{gen}} \quad (3.8)$$

where  $N_{\text{gen}}$  is the number of generated  $B_c$  in the MC sample before any cuts.  $\mathcal{L}_{\text{eff}}$  is simply  $\mathcal{L}_{pp}$  in pp, and is given in PbPb in Eq. 3.10.

### 3.5.5 Normalisation of PbPb yields

#### 3.5.5.1 MC

In PbPb collisions, the centrality of the collisions and a normalisation of the yields to pp-equivalent cross sections are needed for comparison to the numerator of the  $R_{\text{PbPb}}$  of Eq. 1.31. This is done using the concepts explained in section 1.5.6 (straightforwardly adapted to PbPb collisions). As hard probe cross sections are studied,  $N_{\text{coll}}$ -scaling must be reflected in the pp-equivalent cross section, which is not the case in MC where the centrality distribution is flat. For the signal MC, this is taken into account by using the effective luminosity of Eq. 3.10. The  $J/\psi$  MC was normalised directly to PbPb yields, but the  $N_{\text{coll}}$ -scaling must still be implemented, through a weight  $\frac{N_{\text{coll}}}{\langle N_{\text{coll}} \rangle_{\text{MB}}}$ . The quantities  $N_{\text{coll}}$ ,  $N_{\text{part}}$ ,  $T_{\text{PbPb}}$ ,  $\sigma_{\text{NN}}$  and their relationships, are explained in section 1.5.6, and their values are taken from the MC Glauber model of Ref. [35] (which is the model used in the standard centrality estimations in CMS).

As a reminder of Eq. 1.31, and considering a trimuon  $p_T$  and rapidity dependence, the  $R_{\text{PbPb}}$  quoted in the results will be:

$$R_{\text{PbPb}}(B_c^+(p_T^{\mu\mu\mu}, y^{\mu\mu\mu})) = \frac{\frac{1}{\langle T_{\text{PbPb}} \rangle_C \times w(C) \times N_{\text{PbPb, MB}}} \frac{d^2 N_{\text{PbPb}}^{B_c}}{dp_T^{\mu\mu\mu} dy^{\mu\mu\mu}}}{\frac{d^2 \sigma_{pp}^{B_c}}{d^2 p_T^{\mu\mu\mu} dy^{\mu\mu\mu}}} \quad (3.9)$$

The pp-equivalent PbPb cross section in a centrality class  $\mathcal{C}$  is hence the total yield divided by  $\langle T_{\text{PbPb}} \rangle_{\mathcal{C}} \times w(\mathcal{C}) \times N_{\text{PbPb,MB}}$ : the nucleonic luminosity per PbPb collision for the studied centrality range  $\mathcal{C}$ , times the number of (minimum-bias) hadronic PbPb collisions in this centrality class.

The considered centrality class for  $T_{\text{PbPb}}$  and  $w(\mathcal{C})$  is 0 – 90%, except when the centrality dependence is studied (in the ranges 0 – 20% and 20 – 90%, cf section 5.6). The MC in PbPb is normalised assuming that there is no QGP modification ( $R_{\text{PbPb}} = 1$ ). The total expected signal yield in MC is the ratio of the pp  $B_c$  cross section from previous measurements (from section 3.5.4.2), and the pp-equivalent of the minimum-bias luminosity in PbPb collisions:

$$\mathcal{L}_{\text{eff}} = \langle N_{\text{PbPb,MB}} T_{\text{PbPb}} \rangle_{\mathcal{C}} = \langle (\mathcal{L}_{\text{PbPb}} \times \sigma_{\text{PbPb}}^{\text{geom}}) \times T_{\text{PbPb}} \rangle_{\mathcal{C}} = \mathcal{L}_{\text{PbPb}} \times \frac{\sigma_{\text{PbPb}}^{\text{geom}}}{\sigma_{NN}} \times \langle N_{\text{coll}} \rangle_{\mathcal{C}} \quad (3.10)$$

where Eq. 1.20 is used to replace  $T_{\text{PbPb}}$ , and  $\mathcal{L}_{\text{PbPb}}$  is given in section 3.5.1. This effective luminosity, to be used in Eq. 3.8, indeed incorporates a  $N_{\text{coll}}$  weight (function of centrality) in the simulated signal samples.

### 3.5.5.2 Data

Section 3.5.5.1 concerns the normalisation of the (signal) MC samples, but let us now consider how the signal yields extracted from data will be normalised, to obtain the equivalent cross sections entering the  $R_{\text{PbPb}}$  in Eq. 3.9. In pp, the corrected yield is divided by the luminosity given in section 3.5.1, for the used dimuon trigger and run periods. The uncertainty on the pp luminosity is 1.9% [188].

The luminosity is determined with a LHC run dedicated to a VAN DER MEER scan [119]. The beams are progressively deviated to modify their transverse intersection area, and the associated change in the rate of the L1 trigger (defined in section 3.2.4) allows for a luminosity measurement, as explained in Ref. [189].

Only events in the 0 – 90% centrality range are considered in this analysis (see the event selection in section 5.2). So in PbPb, the corrected signal yield is normalised by  $\langle N_{\text{PbPb,MB}} \times T_{\text{PbPb}} \rangle_{0-90\%}$ . The number of minimum bias events (defined in section 1.5.5) in the 0 – 90% centrality range is calculated as:

$$\langle N_{\text{PbPb,MB}} \rangle_{0-90\%} = 0.9 \times N_{\text{PbPb,MB}} = N_{\text{MB}}^{\text{ref}}(0-90\%) \times \frac{\mathcal{L}_{\mu\mu \text{ trigger}}}{\mathcal{L}^{\text{ref}}} = 1.1194 \times 10^{10} \quad (3.11)$$

The quantities  $\mathcal{L}^{\text{ref}}$  and  $N_{\text{MB}}^{\text{ref}}$  correspond to the luminosities and number of minimum bias for the reference set of run periods that consists in all the runs where the subdetectors used for muons were operational.  $N_{\text{MB}}^{\text{ref}}(0-90\%) = 1.1968 \times 10^{10}$  was calculated for all muon analyses, and its uncertainty is 1.26%, due to the removal of electromagnetic or beam scraping events and the centrality calibration.<sup>1</sup>

The nuclear overlap function  $T_{\text{PbPb}}(0-90\%) = 6.27 \pm 0.14 \text{ mb}^{-1}$  is taken from [35], where the 2.2% error is from the Glauber modelling. Including the  $N_{\text{MB}}^{\text{ref}}$  uncertainty, this amounts to a 2.6% uncertainty on the global normalisation of the PbPb signal yields. In the two centrality ranges 0 – 20% and 20 – 90%, the  $T_{\text{PbPb}}$  factors are respectively  $18.79 \pm 0.36$  and  $2.700 \pm 0.098 \text{ mb}^{-1}$ , which give total PbPb normalisation uncertainties of 2.3% and 3.8%. These factors are similarly multiplied by  $N_{\text{PbPb,MB}}$  times the width of the centrality range.

---

<sup>1</sup>For  $N_{\text{MB}}$  and its error, see [https://indico.cern.ch/event/935265/contributions/3930641/attachments/2068596/3472117/PAG\\_200703\\_NMB.pdf](https://indico.cern.ch/event/935265/contributions/3930641/attachments/2068596/3472117/PAG_200703_NMB.pdf) or [https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideHeavyIonCentrality#PbPb\\_Data\\_2018](https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideHeavyIonCentrality#PbPb_Data_2018) (limited access).



*“Je vous le dis, mes frères, l’ère de la peur est terminée. Ce que l’homme craint, c’est ce qu’il ne connaît pas. Je sais qu’il y a tant de choses que nous ignorons, tant de choses que la machine savait pour nous. Je sais que depuis toujours, vous vivez dans la crainte, la crainte du Virus, la crainte de l’extérieur, la crainte de ne pas survivre. Pourquoi? Parce que vous ne savez pas ce que veulent vraiment dire ces mots.”*

---

Raphaël GRANIER DE CASSAGNAC,  
*Eternity Incorporated*<sup>1</sup>

## Chapter 4

# Whom do we fight? Backgrounds

The leptonic channel  $B_c^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)\mu^+\nu_\mu$  we aim to study is partially reconstructed, meaning the signal is spread over a wide mass range. Therefore, its signal extraction will be done with a template fit of the trimuon mass ( $m^{\mu\mu\mu}$ ) distribution, in chapter 6. The trimuon mass distribution is studied in the [3.5, 7.8] GeV region: the sum of the shapes of the simulated signal and the different background sources is fitted to the data. This means the **background trimuon mass shapes** must be well understood, which is the key challenge of this analysis – along with the consideration of the true signal kinematics in the acceptance and efficiency estimation (see chapter 7.6). The signal normalisation is fitted alongside the normalisations of some of the backgrounds, and alongside nuisance parameters to account for the uncertainties on the background shapes (see section 6.2). This chapter aims at determining the background shapes, their normalisations for some of them, and the uncertainties on these shapes that will be considered in the fit. Section 4.1 explains how the backgrounds are categorised into three sources, and the three following sections detail how reliable trimuon mass distributions are extracted for each of these backgrounds.

### 4.1 Categorisation

The various backgrounds are categorised with respect to these sequential criteria:

1. Is there a true  $J/\psi$  among the two opposite-sign muon pairs of the trimuon? Or a pair only accidentally has a mass close enough to the  $J/\psi$  mass?
2. If yes, do the true  $J/\psi$  and (fake or true)  $\mu$  come from the same B-meson decay vertex?

---

<sup>1</sup>*“I tell you, brothers, the era of fear is over. What humans dread is what they don’t know. I know there are so many things we ignore, so many things the machine knew for us. I know that you’ve always lived in fear, fear of the Virus, fear of the outside, fear not to survive. Why? Because you don’t know what those words really mean.”*. Personal note: I do not assert that those words relate to any current worldwide situation.

As a side note, it is warned here that ‘from the same B-decay vertex’ includes the case where the  $\mu$  is displaced compared to the B-decay vertex, *e.g.* if the B decay products contain a  $J/\psi$  and a long-lived particle that itself decays in a muon. It is however a very small contribution if any (no obvious B decay, excluding the  $B_c$ , having been thought of with such characteristics), justifying the use of the expression ‘from the same B-decay vertex’.

This leads to three categories of backgrounds, that are described in the following sections. A flowchart with the categorisation strategy is shown in Fig. 4.1. The **fake  $J/\psi$  background** is the most straightforward, and described with the dimuon mass sidebands, in section 4.2. The background from **B mesons decaying to a  $J/\psi$  and a (fake or true) muon** is described with the non-prompt  $J/\psi$  MC, which is reliable for these well-known decays. This association of a displaced  $J/\psi$  with a charged particle from the same parent decay is described in 4.3. The case of  $B_c$  mesons decaying to a  $J/\psi$ , a muon, and other particles, is taken into account with a global systematic uncertainty, detailed in section 4.3.3. The rest of the backgrounds, described in section 4.4, comprises prompt or non-prompt  **$J/\psi$  mesons combining with a muon from the PV or from a different secondary vertex** (typically from the decay of the ‘companion’ B meson). It is estimated by a data-driven method in section 4.4.3, as the  $J/\psi$  MC samples were found not to be able to describe it for pp collisions.

All plots shown in this section contain the events passing the preselection of section 5.4, but are not required to pass the fiducial cuts of section 5.6.

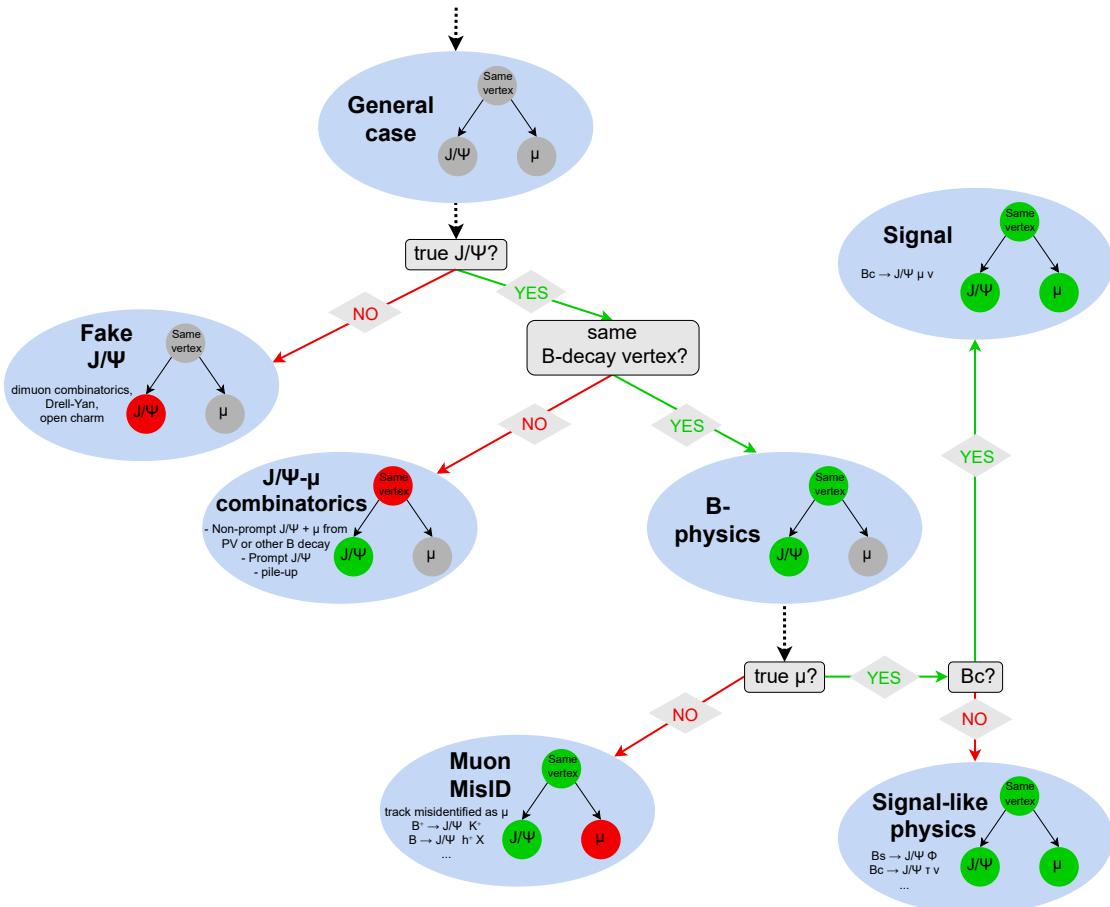


FIGURE 4.1: Flowchart showing the categorisation of backgrounds.

## 4.2 Fake J/ $\psi$

The fake J/ $\psi$  background is described with trimuon candidates from the **sidebands of the dimuon mass spectrum**. These trimuons are estimated to have similar properties as the fake J/ $\psi$  events under the J/ $\psi$  peak. The width of the sidebands region is such that this background has the same normalisation as the background under the J/ $\psi$  peak, therefore it does not require a fitted normalisation parameter in the template fit.

Considering the relatively low statistics, one can reasonably assume that the dimuon mass spectrum close to the J/ $\psi$  peak is linear. Therefore, one should get an equivalent number of fake J/ $\psi$  in the sidebands than under the peak, by simply choosing the lower and upper sidebands each of a size equal to half the width of the signal mass region. The peak and sideband regions are also defined tighter when the muons are closer to the central barrel, because the muon  $p_T$  (hence dimuon mass) resolution is better in this region [175]. Due to final state radiation, the J/ $\psi$  dimuon mass can be dragged to lower values, resulting in an asymmetric peak, so I also slightly shift down the mass window of the lower-mass sideband. One needs to limit the contamination of signal into the sidebands background sample, while keeping the sidebands close enough to the nominal J/ $\psi$  mass  $m_{\text{PDG}}^{\text{J}/\psi} = 3.096 \text{ GeV}$  [46] so that their properties (especially the kinematic distributions) stay similar; I hence choose a 50 MeV (resp. 40 MeV) gap between the signal region and the lower (resp. higher) sideband.

Finally, the signal and sideband regions are respectively:

- $m_{\mu\mu} - m_{\text{PDG}}^{\text{J}/\psi} \in [-100, 80] \text{ MeV}$  and  $m_{\mu\mu} - m_{\text{PDG}}^{\text{J}/\psi} \in [-240, -150] \cup [120, 210] \text{ MeV}$ , when all three muons have  $|\eta| < 1.5$ ;
- $m_{\mu\mu} - m_{\text{PDG}}^{\text{J}/\psi} \in [-150, 110] \text{ MeV}$  and  $m_{\mu\mu} - m_{\text{PDG}}^{\text{J}/\psi} \in [-330, -200] \cup [150, 280] \text{ MeV}$  otherwise.

Selecting events in such a mass peak region has an overall 95.7% (resp. 96.3%)<sup>1</sup> efficiency on the signal MC in pp (resp. in PbPb) with other preselection cuts applied (listed in section 5.4). On the other hand, only 1.2% (resp. 1.8%) of the preselected signal MC falls in the sidebands in pp (resp. in PbPb), which is a negligible signal contamination considering the amount of fake J/ $\psi$  it will be mixed with.

Care has been taken not to bias the J/ $\psi$  candidate mass shape, by using no mass criteria to select the J/ $\psi$  dimuon among the two possible opposite-sign muon pairs of the candidate trimuon. This is necessary to keep the correct fraction of fake J/ $\psi$  events in the signal and sideband regions and ensure their cancellation. Section 5.5 explains how a weight is attributed to the candidates whose J/ $\psi$  choice is ambiguous. This strategy also removes artificial structures in the trimuon mass shapes due to excessive kinematic constraints. The dimuon mass spectra of preselected data are shown in Fig. 4.2, taking into account this weighting.

### Differences in the two sidebands

When using the trimuon mass distribution of this background, one would like to reduce the effect of the displaced dimuon mass that makes the kinematic properties of sideband events slightly different than under the J/ $\psi$  peak. This is partially accomplished by fixing the dimuon mass to the PDG J/ $\psi$  mass in the trimuon invariant mass calculation (this is done for all samples, because it also improves the trimuon mass resolution by removing the experimental resolution on the J/ $\psi$  dimuon mass). Ideally, the trimuon mass shapes in the two sidebands would be similar, so that one could be confident in interpolating this shape to the background between the two dimuon sidebands. However, the shapes in the two sidebands are somewhat different, as shown in Fig. 4.3, with a smaller difference in PbPb than in pp.

<sup>1</sup>Weights concerning the choice of the J/ $\psi$  dimuon (from section 5.5) are included in those numbers.

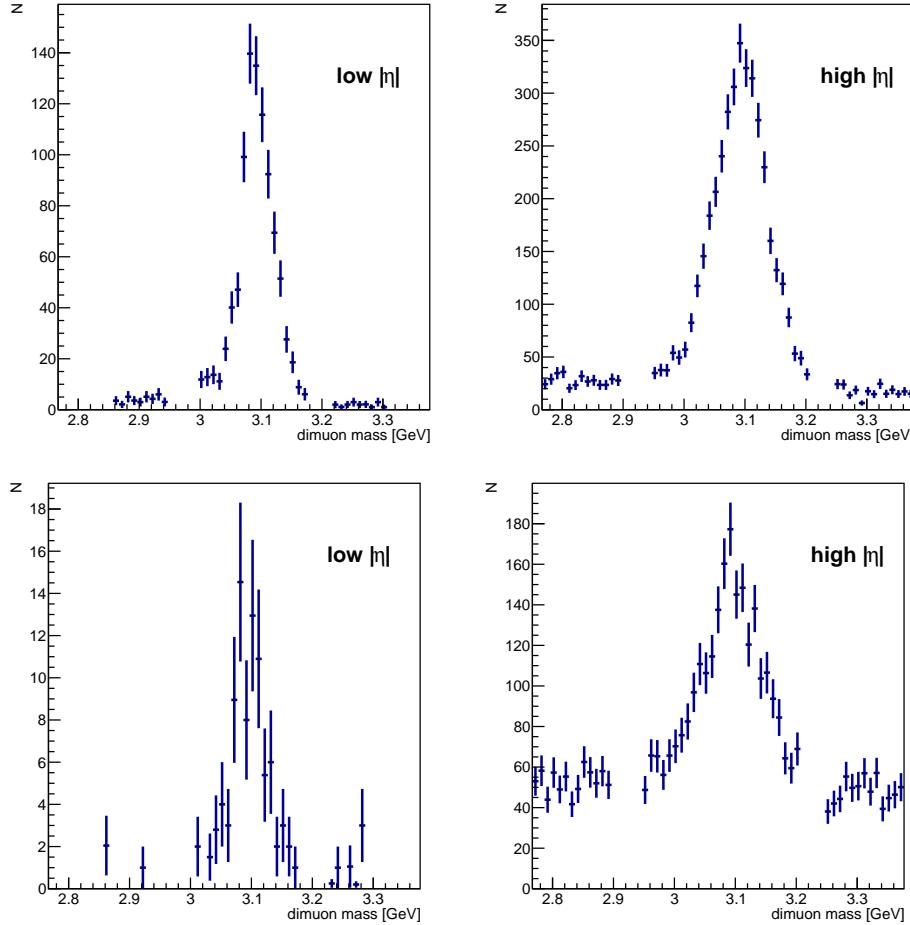


FIGURE 4.2: Dimuon mass in pp (*top*) and PbPb (*bottom*) preselected data, for tight (all three muons have  $|\eta| < 1.5$ , *left*) and loose (other cases, *right*) ranges, when the trimuon mass is in  $[3.5, 6.2]$  GeV. Only the events in the signal or sideband region are shown.

Ultimately, the sum of these histograms serves as the fake  $J/\psi$  sample, but it has been verified that the introduced bias is small. I thank here Natalie BLOT, who implemented this check during her internship at LLR. To estimate the true shape under the  $J/\psi$  peak, one needs to interpolate between the two observed sidebands. Let us call  $f(m_{\mu\mu})$  the trimuon mass probability density of fake  $J/\psi$ , depending on the dimuon mass. The shape  $f(m_{\mu\mu}(\text{lower sideband})) + f(m_{\mu\mu}(\text{higher sideband}))$  is *not* the same as the  $f(m_{\mu\mu}(\text{peak region}))$  one. First the two sidebands are fitted with an error function times decreasing exponential, with a fixed error function threshold (3.28 GeV). The fitted functions are  $f(\{\alpha_{i,\text{low}}\})$  and  $f(\{\alpha_{i,\text{high}}\})$ , where  $\{\alpha_i\}$  are the fitted parameters. Then we interpolate the parameters to have an estimation of the central shape  $f(\{\alpha_{i,\text{central}}\})$ . The parameter interpolation is a weighted average (taking into account the different number of events  $N_{\text{low}}$  and  $N_{\text{high}}$ ) of the two shapes:  $\alpha_{i,\text{central}} = \frac{N_{\text{low}}\alpha_{i,\text{low}} + N_{\text{high}}\alpha_{i,\text{high}}}{N_{\text{low}} + N_{\text{high}}}$ . The two sideband fits and the function with averaged parameters are shown in Fig. 4.4 (*left*), for the more problematic pp case. The fits are of good quality:  $\chi^2/\text{ndf} = 1.7$  or 1.1 for the left or right sideband, respectively.

However, comparing the function with interpolated parameters to the sum of the fits of the two sidebands as shown in Fig. 4.4 (*right*), our sophisticated strategy appears to make only a relatively small change in the trimuon mass shape. Moreover, the fitting cannot be used with lower statistics, which will be the case later when the BDT binning is applied, producing low-background BDT bins with very low fake  $J/\psi$  yields. This interpolation strategy is therefore

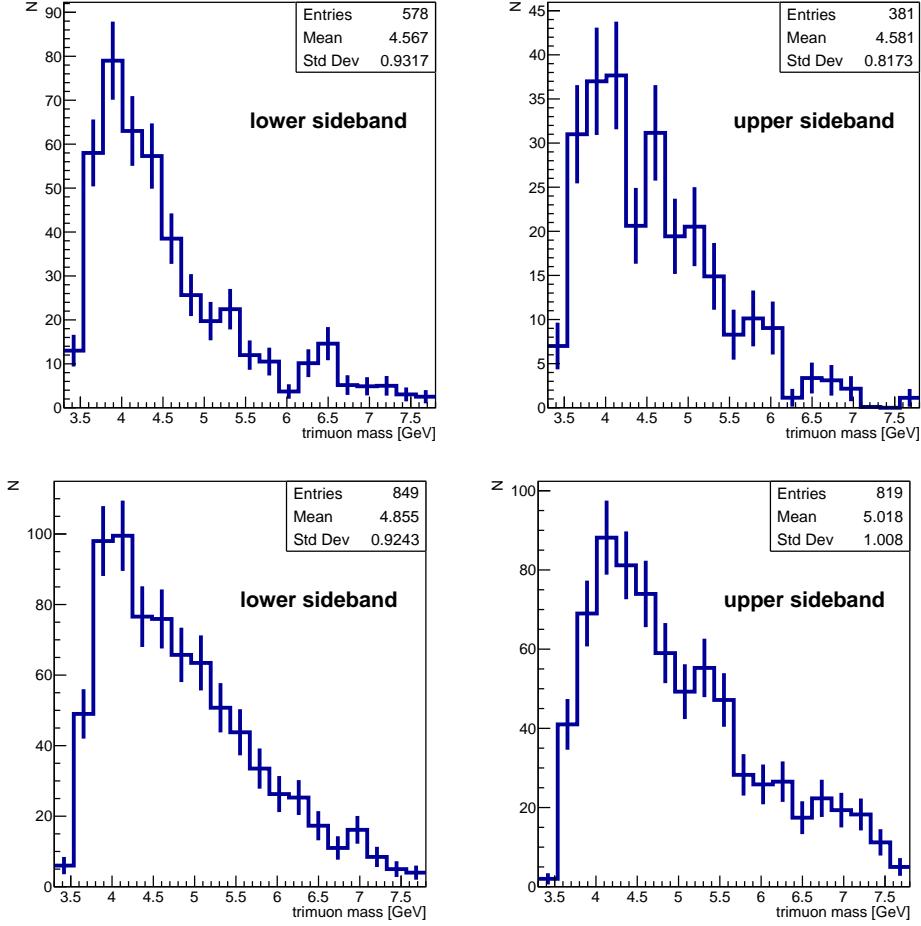


FIGURE 4.3: Trimuon mass for pp (*top*) and PbPb (*bottom*) events from the lower (*left*) and higher (*right*) sidebands.

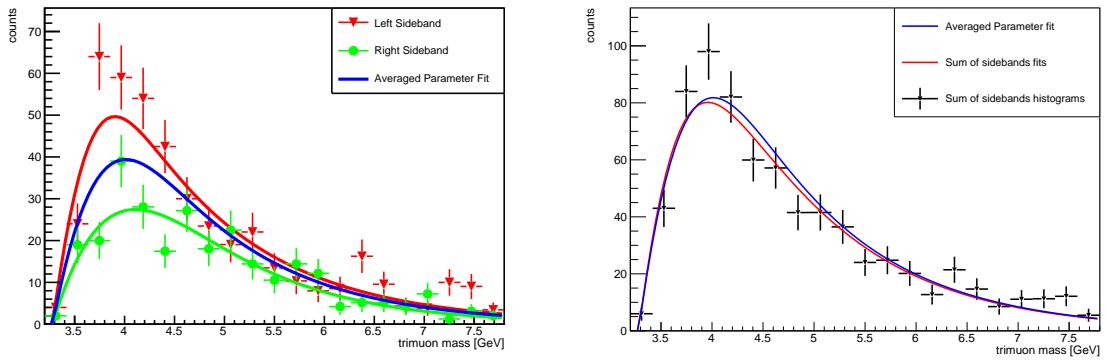


FIGURE 4.4: Fitting sidebands in order to get a better fake J/ $\psi$  trimuon mass shape in pp collisions. *Left*: Separate fits of the left and right sidebands, compared to the function with parameters interpolated between the two sideband fits (blue). *Right*: Fit with interpolated parameters (blue), compared to the sum of the fits of separate sidebands (red), and to the histogram of summed sidebands (black). The two functions are normalised to the integral of the summed histogram (total number of fake J/ $\psi$ ).

dropped as the central method<sup>1</sup>, and the shapes from the **separate sidebands are kept as systematic variations in the template fit** (see section 6.2), which is sufficient to cover the uncertainty on this shape.

#### Note on the wrong-sign sample

Fake J/ $\psi$  candidates can come from physics source (e.g. a Drell-Yan  $\mu^+\mu^-$  continuum, or  $D^+D^-$  pairs that decay leptonically) or from pure combinatorics (uncorrelated muons which are mostly misidentified hadrons, and possibly from pile-up). The latter can be estimated by the wrong-sign sample, consisting of trimuons of charge  $\pm 3$  (with other selections identical to the signal region). This sample will be displayed for illustration in the template fit plots (in section 6.3) but does not participate to the fit, since the associated background is already included in the fake J/ $\psi$  contribution (the dimuon mass sidebands).

Looking at an ordered list of three  $\pm 1$  charges ( $+++,---,++-,-+ +,+-+,--+$ ,  $+--,-++$ ), one observes that only two out of eight combinations of three random muons lead to a charge  $\pm 3$  trimuon. This means that there are three times more trimuons of charge  $\pm 1$  coming from random muons than those of charge  $\pm 3$  (the latter constituting the wrong-sign sample). Therefore, a weight 3 is given to the wrong-sign candidates, for a correct normalisation of the estimated purely combinatorial background in the right-sign sample.

### 4.3 True J/ $\psi$ and muon from the same B decay

Let us now consider backgrounds that feature a true J/ $\psi$  meson, starting with a J/ $\psi$  meson associated to an identified muon arising from the hard process. If produced at the primary vertex, such J/ $\psi$ -muon pairs could leak into our displaced vertex selection, because of the finite vertex resolution. However, no prompt hard mechanisms can strictly produce a J/ $\psi$  with a third muon with yields relevant for this analysis, so I limit this second background category to non-prompt J/ $\psi$  mesons, namely to **true  $B \rightarrow J/\psi X$  decays** where the  $X$  decays to or contains a muon or a hadron misidentified as a muon. This excludes  $B_c$  decays for which a systematic uncertainty is estimated in section 4.3.3. In any case, prompt J/ $\psi$  mesons combined with a (true or fake) muon coming from the same process or the underlying event are included in the data-driven background of section 4.4.3.

#### 4.3.1 B decays

So we now deal with a non-prompt J/ $\psi$  combined with a muon from the same displaced vertex, meaning B meson decays (a displaced vertex could mean an open charm decay, but those are not heavy enough to produce a J/ $\psi$ ). The cases where a B decays to a J/ $\psi$  and a long-lived particle (e.g. a D meson), which then decays to a muon, are also included but negligible (no such significant decay was thought of except for  $B_c$  decays). The **B meson decays to three muons** (and possibly other particles) are rare (1% of the true B decays in the non-prompt J/ $\psi$  MC in pp), the predominant one being  $B_s \rightarrow J/\psi \phi \rightarrow \mu\mu\mu\mu$ . More abundant are the B meson decays with a **J/ $\psi$  and a hadron misidentified as a muon**, such as  $B^+ \rightarrow J/\psi K^+$  (see section below). These processes are included in the non-prompt J/ $\psi$  MC, in which all  $B^0$ ,  $B^\pm$ , and  $B_s$  decays are simulated. This sample will thus provide the second template for our fit, with a quasi-free normalization parameter and an important characteristic: it does not extend above  $\sim 5.4$  GeV, that is the  $B_s$  mass.

---

<sup>1</sup>Another solution would be to use the *sPlot* technique [190]. This technique does not need the assumption of linear background, and predicts through *sWeights* (discriminating signal and background) the shape of the background under the J/ $\psi$  peak. Its caveat is however that the discriminant (dimuon mass) and extracted (trimuon mass) variables need to be uncorrelated for the technique to be correct, which is clearly not the case in this analysis. In addition, the complications linked to the choice of the J/ $\psi$  dimuon among the trimuon make unclear the dimuon mass distribution that is to be used for the extraction of the *sWeights*.

### 4.3.2 Muon misidentification

About 90% of the simulated trimuon candidates from these same-decay processes indeed have a fake as their third identified muon (the one not coming from  $J/\psi$ ), that is a misidentified hadron. As detailed in Ref. [175], the misidentification of hadrons into muons stems from three sources: *mismatching* of a standalone track in the muon chambers with a hadron track in the inner tracker, *punch-through* (high- $p_T$  hadrons that do not loose all their momentum before reaching the first muon station), and *decay-in-flight* (hadrons decaying to muons before the muon stations). The latter is the most important source at low  $p_T$ . For a loose muon selection, the probability (that depends on  $p_T$ ) to identify a proton as a muon is  $\lesssim 0.1\%$ , while it is respectively 0.5-1% and 0.2-0.5% for the kaon and pion.<sup>1</sup>

The non-prompt  $J/\psi$  MC is considered reliable for the **true  $B \rightarrow J/\psi h^\pm X$  and  $B \rightarrow J/\psi \mu^\pm X$  decays** (where  $B$  is any  $B$  meson except the  $B_c$ ), because these are well-known processes, for which the MC has been checked by other analyses. In this simulation, one needs to discard these candidates from the non-prompt  $J/\psi$  mesons combining with a muon from a different decay or from the PV (this background source is studied in section 4.4). It is done by matching, in the ONIATREE, the simulated particles (i.e. simulated hits that are reconstructed into tracks as in data) with the true particles generated in the PYTHIA processes, for which the decay genealogy is known.

The trimuon invariant mass shapes are presented in Fig. 4.5. The bulk below 5 GeV corresponds to partially reconstructed  $B \rightarrow J/\psi (h^\pm/\mu^\pm) X$  decays where the missing particles make the trimuon mass lower than the  $B$  mass. The peak observed around 5.2 GeV corresponds to the decay  $B^+ \rightarrow J/\psi K^+$  where the kaon is misidentified as a muon. As will be clear in the template fit of chapter 6, this background is lower in PbPb– notably because  $B$  mesons are suppressed in the QGP.

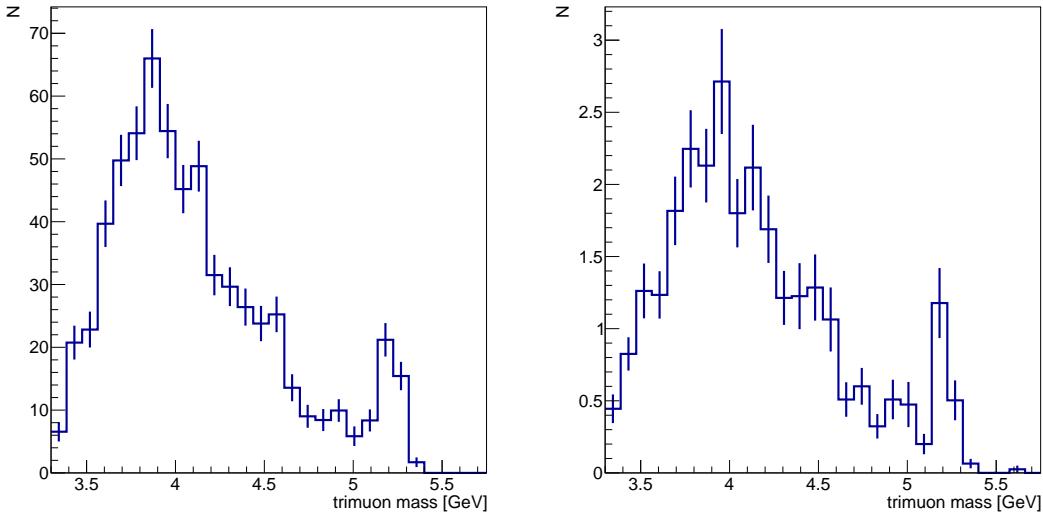


FIGURE 4.5: Trimuon mass shapes of preselected of simulated  $B \rightarrow J/\psi (h^\pm/\mu^\pm) X$  candidates in pp (*left*) and PbPb (*right*).

In Fig. 4.5, the simulations are scaled by the *a priori* normalisation calculated in section 3.5.4 (with the help of previous non-prompt  $J/\psi$  cross section measurements) to give a first estimate of the expected background. However, since this normalisation and the misidentification rate are imperfectly known, the normalisation of this background source will be left free in the template

<sup>1</sup>The kaon has a higher misidentification probability than the pion because of its smaller lifetime, giving it a larger probability to decay into a muon within the tracker volume.

fits (though with a loose constraint to guide them). It is to be noted that the  $B^+ \rightarrow J/\psi K^+$  peak helps the fit to determine the normalization (at least in pp in the background-enriched low-BDT samples), but that it is not always visible, given its low statistics and the other backgrounds.

It is beyond the reach of this analysis to put in place a full data-driven check of the misidentification rate, for example with a tag-and-probe on the  $\phi \rightarrow K^+ K^-$  (resp.  $K_S^0 \rightarrow \pi^+ \pi^-$ ) decays for the misidentified kaons (resp. pions).<sup>1</sup> A global factor of data/simulation discrepancy in this rate is covered by the normalisation of this background being free in this fit. However, to cope with possible momentum- or rapidity-dependent errors in the simulated misidentification, modifications of the trimuon mass shape of this background will also be considered as systematic uncertainties (see section 6.2), except in PbPb where it is subdominant compared to other background sources.

### 4.3.3 Other $B_c$ decays

One should not forget a decay that resembles very much the signal: the  $B_c \rightarrow J/\psi \tau \bar{\nu}_\tau$  decay, with the  $\tau$  decaying in a  $\mu$  (and neutrinos). One could produce an extra signal sample with this decay, but one would still need to input the coarsely-known ratio of the tauonic and muonic  $B_c$  decays to get an inclusive signal MC; furthermore, I argue here that such a sample is not necessary. The ratio of the tauonic to muonic  $B_c$  decays is bound to be strictly less than unity both by lepton universality combined with a tighter tauonic phase space, and by a LHCb measurement of  $0.71 \pm 0.25$  [59]. As a side comment, this measurement is part of the set of hints for lepton flavour universality that is currently agitating the High Energy Physics community: the Standard Model predictions for this ratio, some of which are listed in Ref. [59], are indeed in the range  $0.25 - 0.28$ , which is  $1.8\sigma$  lower than the measurement. This ratio then needs to be multiplied by the branching fraction of  $\tau$  to  $\mu$  of  $17.3\%$  [46]. This means that 12% (or less than 5% if one rather believes Standard Model predictions) of the true  $B_c \rightarrow \mu\mu\mu X$  events come from this decay. Moreover, for multiple reasons, the tauonic decay will have a much smaller acceptance times efficiency than the muonic decay:

- the muon from the tau decay is in general of lower  $p_T$  than the one in the muonic decay because some energy is lost to the neutrinos from the tau decay. This yields a lower muon acceptance times efficiency in general (see maps of section 5.3.2);
- the  $\tau$  lifetime ( $ct_\tau = 0.087$  mm at rest) is slightly above the CMS vertex resolution, meaning that the cuts on the probability of the trimuon vertex `VtxProb` (probability that the three muon have a common vertex, see section 5.4) will tend to reject tauonic decays;
- the decay kinematics will be slightly different than the signal channel, which further helps the BDT to discriminate against it.

Estimating precisely the effect of all these additional inefficiencies in the  $\tau$  channel would require its full simulation. However, one can get a lower limit through a rough estimate of two contributing inefficiencies, helped by the MC sample for the  $B_c$  muonic decay, considering only trimuons passing pre-selection (section 5.4) and fiducial cuts (section 5.6).

First, I estimate that a muon decaying from a tau has in average  $p_T^\mu \simeq 0.6p_T^\tau$ ; this is starting from an ultrarelativistic estimation that half of the energy is taken away by the neutrinos of the decay, and leaving some margin for a contribution of the  $\tau$  mass to the  $p_T$ . Then I check with the muonic  $B_c$  simulation the proportion of signal events that pass the single muon acceptance cuts (of section 5.3) when the  $p_T$  of the muon not coming from  $J/\psi$  is multiplied by 0.6; it is about 67%. This assumes that the kinematic distributions of the decay products are similar in

---

<sup>1</sup>Previous studies like Ref. [191] at 7 TeV and Ref. [175] at 13 TeV have compared the rate of muons that are actually hadrons between data and MC, and found consistency, though within relatively large errors, in different detector conditions than ours, and in general for muons with  $p_T > 3.5$  GeV.

the muonic and tauonic decays, despite the non-negligible tau mass. To this factor could be added the additional inefficiency of muon identification, that decreases towards low  $p_T$  (see maps of section 5.3.2).

Then, I estimate how the additional displacement of the muon coming from  $\tau$  affects its efficiency of passing the selection, in particular the trimuon vertex probability cut (`VtxProb`). For this, I look at the inefficiency of a tighter `VtxProb` cut compared to the preselection, as a function of the distance of closest approach (DCA) between the two muons of the J/ $\psi$ ; this DCA serves as an approximation for the DCA between the J/ $\psi$  and the muon coming from the  $\tau$  decay. Assuming  $p_\tau \gg m_\tau$  (and small  $\tau$  decay opening angles), the angle between the tau and its muon daughter is of order  $m_\tau/p_\tau$  (obtained from the decay kinematics), but also of order  $\frac{DCA}{ct_\tau \times p_\tau/m_\tau}$  (considering the tau boost), which means  $DCA \sim ct_\tau$ , the proper time. Now, Fig. 4.6 (left plot) shows in pp (conclusions are similar in PbPb) the inefficiency of the cut  $\text{VtxProb} > 0.04$  (*left plot*) or  $\text{VtxProb} > 0.06$  (*right plot*), from the preselection cut  $\text{VtxProb} > 0.01$ . It is fitted to show this inefficiency rises linearly with the  $DCA(\text{J}/\psi)$ . Most of the trimuons have a  $DCA(\text{J}/\psi)$  very close to 0, therefore adding a  $DCA \sim ct_\tau$  between two muons of the trimuon affects the inefficiency of a `VtxProb` cut similarly as an increase of  $DCA(\text{J}/\psi)$  from 0 to  $ct_\tau$ . The ratio of the inefficiency fit function between  $DCA = ct_\tau$  and  $DCA = 0$  gives about 6 for both tested `VtxProb` cuts. Multiplying this by 5%, the inefficiency of the preselection `VtxProb` cut (see tables of section 5.4), I get an efficiency of this cut on the  $\tau$  channel of about 65 – 70%.

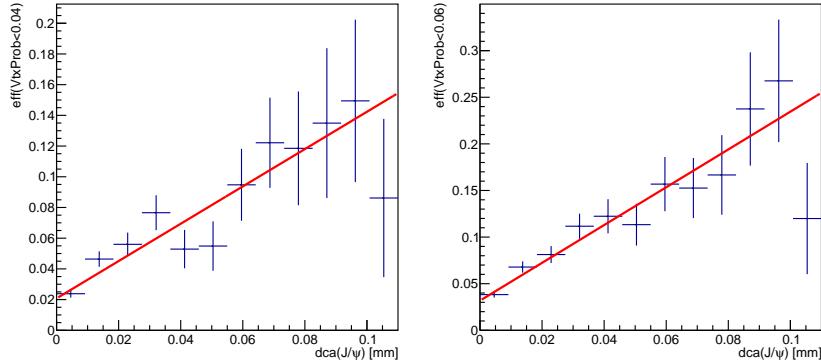


FIGURE 4.6: Inefficiency, on the pp signal MC sample (preselected and in fiducial cuts), of a cut of the trimuon vertex probability  $> 0.04$  (*left*) or  $> 0.06$  (*right*), as a function of the distance of closest approach between the two muons from the J/ $\psi$ .

Multiplying these two efficiencies, present only in the  $\tau$  channel, gives 45%. This estimation is similar in pp and in PbPb. It is reasonable to assert that the contributions not taken into account (in particular the sophisticated BDT selection) drive this efficiency down to 25%. Multiplied by 12% (the ratio of tauonic to muonic decays), I hence estimate that the overall contamination of our signal by these channels should not exceed 3% and will assign this as an asymmetric (downward only) uncertainty on the cross sections, fully correlated along analysis bins. Considering that the efficiency of reconstructing these channels is similar in pp and PbPb, these contributions approximately cancel in the ratio of pp and PbPb yields. This is partially confirmed by finding similar inefficiencies in pp and PbPb for the two contributions roughly estimated above. A symmetric 1% error is hence assigned in the  $R_{\text{PbPb}}$  to cover for an imperfect cancellation of these contributions.

Decays of the type  $B_c \rightarrow J/\psi(D \rightarrow \mu X)X$  have been observed, but in addition to low (unmeasured) branching fractions, the D meson travels more than the  $\tau$  before decaying to a muon, which reduces enough its efficiency for these channels to be ignored.

There can also be feed-down from  $B_c \rightarrow h_{cc} \mu X$ , where  $h_{cc}$  is an excited charmonium that decays to a  $J/\psi$  and other particles. It is to be noted that no such decays have been observed yet, but Ref. [192] summarises some theoretical predictions for the branching fractions of these decays. The  $B_c \rightarrow \psi(2S)\mu\nu$  decay has a negligible branching fraction compared to our main decay, despite the high feed-down  $BF(\psi(2S) \rightarrow J/\psi X) = 0.61$ . However, the  $B_c \rightarrow \chi_{c0,1,2} \mu X$  has a significant branching fraction: roughly 30% of our signal channel, considering the various (and quite dispersed) predictions in Ref. [192]. The radiative feed-down fractions  $\chi_{c0,1,2} \rightarrow J/\psi \gamma$  are resp. 1.4%, 34.3%, and 19.0% for the  $\chi_{c0}$ ,  $\chi_{c1}$ , and  $\chi_{c2}$  states [46], so I roughly consider that 20% of the  $\chi_c$  mesons from  $B_c$  decays result in a  $J/\psi$ . The ratio of the total  $B_c \rightarrow (\chi_{c0,1,2} \rightarrow J/\psi X) \mu X$  branching fraction to our signal channel is therefore about 6%.

However, the acceptance and efficiency for reconstructing these decays will be lower than for the main signal channel, with similar arguments than for the  $\tau$  channel above. Namely, the  $\chi_c$  states (of masses 3.41 GeV to 3.56 GeV) take more energy from the  $B_c$ , notably through the unreconstructed radiated photon, leaving a smaller  $p_T$  for the final-state muons; and the muon acceptance and efficiency reduces rapidly with  $p_T$ . The selection variables might also vary from the main signal events; for example, the resulting trimuon mass shape will be shifted to lower values due to the energy taken by the photon. It is therefore reasonable to estimate that the acceptance and efficiency for reconstructing and selecting these channels as signal is at least twice lower than for the studied signal channel, resulting in a maximum  $0.5 \times 6\% = 3\%$  contribution to the measured signal yields, taken as a one-sided (downward) systematic uncertainty on the cross sections. It is considered to partially cancel in the  $R_{PbPb}$  such that a 1% symmetric systematic uncertainty is applied there.

As for the  $\tau$  channel, the efficiency loss compared to the signal channel is considered to be very similar in both  $p_T^{\mu\mu\mu}$  bins (and even more so in the centrality bins, because these decays happen outside of the QGP). Relevant for the excited charmonium decays, the feed-down fractions to  $J/\psi$  are also assumed to have no significant  $p_T$  dependence.<sup>1</sup> This translates in a full correlation of this uncertainty along analysis bins. The final uncertainty associated to other  $B_c$  decays being included in the signal yield is the quadratic sum of the  $B_c \rightarrow \chi_{c0,1,2} \mu X$  and  $B_c \rightarrow J/\psi \tau \bar{\nu}_\tau$  contributions, amounting to a downward 4.5% uncertainty on the cross sections, and a 1.5% symmetric uncertainty on the  $R_{PbPb}$ , fully correlated along analysis bins.

## 4.4 True $J/\psi$ and muon from different processes

Let us now look into the remaining true  $J/\psi$  mesons, that combine with a muon that *does not* originate from the same decay vertex, including the case where they both come from the primary vertex (but are reconstructed as displaced due to the finite vertex resolution). An important characteristic of this background is that, contrary to the  $B$  decays, it should provide the high trimuon mass (above the  $B_c$  mass) background events that the fake  $J/\psi$  background (data-driven and self-normalised) do not provide. Two approaches were followed to estimate this background:

- In principle, one could treat this case with a perfect simulation. Due to problems in the  $J/\psi$  MC with the pp 2017 running conditions (manifesting itself as a bad agreement with data in control regions), this method was complicated to deal with. It was abandoned for the pp case, but found efficient in the PbPb case. This possible method is described in the following section 4.4.1, while the shortcomings of the  $J/\psi$  simulation in pp are demonstrated in section 4.4.2.
- Alternatively, this (mostly) combinatorial background can be estimated directly from data. I developed a method consisting in rotating, in each  $J/\psi$  dimuon event, the  $J/\psi$  candidate

---

<sup>1</sup>This is an approximation, but a  $p_T$  dependence of the feed-down fractions would probably concern low- $p_T$  regions that are not accessed in this analysis.

by some angle before trying to combine it with an additional muon from the same event. This *rotated  $J/\psi$*  method is described in section 4.4.3, and will give its name to this background source in the rest of the analysis.

#### 4.4.1 Combinatorial background from simulation

The prompt  $J/\psi$  MC should cover the case where a prompt  $J/\psi$  meson (surviving the non-zero displacement cuts of the analysis selection) is associated with a muon coming from the underlying event (typically multiple partonic interactions), or from other vertices (pile-up, or non-prompt muons from another  $B$  decay in the event). Similarly, the non-prompt  $J/\psi$  MC should describe the candidates with a true non-prompt  $J/\psi$  and a muon from the PV (passing close enough to the displaced vertex) or from another  $B$  decay, in general from the ‘companion’  $\bar{b}$  that recoiled against the  $b$ -parent of the  $J/\psi$ .

For pp collisions, the study in section 4.4.2 however shows the incapacity of the  $J/\psi$  MC in pp 2017 to predict the yields in the control region at high trimuon mass ( $> 6.3 \text{ GeV}$ ). In the pp 5.02 TeV 2017 run, I found large discrepancies in the variables linked to vertexing and displacement, which were also seen in earlier pp 2017 datasets, and whose source could not be precisely established. The strategy of the Collaboration for high-statistics pp 2017 analyses<sup>1</sup> was to correct the MC in 6 variables simultaneously. However, this is technically heavy, and too demanding for the limited statistics available in this dataset. A weighting in ‘3.5’ variables ( $DCA$ ,  $p_T$ , and lifetime significance, in three rapidity regions) was envisioned, but considering less variables means that non-weighted variables (possibly behaving fine initially) could get distorted due to correlations with the weighted variables; and the discrepant variables are significantly correlated, as is shown in section 4.4.2 too.

In addition, getting perfectly correct  $J/\psi$  distributions on an inclusive  $J/\psi$  sample does not give any guarantee that the distributions of  $J/\psi$  mesons within a preselected trimuon candidate will be correct, as the latter could have different properties than inclusive  $J/\psi$  mesons. I hence concluded that the existing  $J/\psi$  MC in pp could not correctly describe data, and turned to an alternate data-driven method to describe the background of true  $J/\psi$  mesons combining with a muon from a different vertex (or both from the PV).

The relief is that these discrepancies were not observed in pp 2018 runs, and in PbPb 2018 data the high trimuon mass control region can indeed be fitted correctly with the  $J/\psi$  simulations.<sup>2</sup> The latter could possibly be used as the nominal description of this background in PbPb. However, for the sake of consistency between the pp and PbPb methods, and because the data-driven rotated  $J/\psi$  sample is proven in section 4.4.3 to be even more reliable in PbPb than in pp, the  $J/\psi$  rotation is taken as the nominal method. The  $J/\psi$  simulations are considered only as a shape variation of the nominal sample, to cover the systematic uncertainties of this shape (see section 6.2).

#### 4.4.2 Investigation of the $J/\psi$ MC in pp 2017

The problem with the  $J/\psi$  MC in the conditions of the pp 2017 run first manifested itself in the trimuon mass shapes of the prompt and non-prompt  $J/\psi$  MC (supposed to describe all true  $J/\psi$  backgrounds) summed with the  $J/\psi$  sidebands shape (describing the fake  $J/\psi$  background) not being able to make the template fit converge, due to an insufficient total background yield in the high trimuon mass control region. No signal is expected in this  $m^{\mu\mu\mu} > m_{B_c}$  region, so that the sum of the predicted background shapes (with a free normalisation for the MC)

<sup>1</sup>See this talk: <https://indico.cern.ch/event/870866/contributions/3673333/attachments/1963354/3263973/17-12-2019.pdf> (restricted access).

<sup>2</sup>See Fig. 4.16 at the end of this chapter, that shows similar mass shapes for the simulation and the data-driven method in PbPb, while the latter performs well in the template fit of chapter 6.

should fit data, but it does not. Several possible flaws were investigated and reported below: muon identification, track multiplicity, and  $J/\psi$  variables. I investigate this mainly with the non-prompt  $J/\psi$  MC, but most of the presented arguments are valid for the prompt MC too. It is underlined that this section presents an ancillary study, that helped the choice of analysis methods but is not directly used in the main analysis.

#### 4.4.2.1 Dimuon+track mass to test the muon identification

One could think that this problem is related to the muon description in the simulation. However, I checked that this underestimation at high mass still exists in a  $J/\psi$ +track selection in the non-prompt MC, where the simulated track is not affected by possible faults in the muon selection in MC. `OniaTree`'s similar to the trimuon ones were generated to build dimuon+track candidates, as in section 3.3.2. The track that the  $J/\psi$  combines with is still attributed the muon mass. Weights from a simple sideband subtraction (to cancel out fake  $J/\psi$  candidates, as in section 4.2) were also applied, along with a cut on the significance of the 3D lifetime of the dimuon ( $> 3\sigma$ , to eliminate prompt  $J/\psi$ ). The MC was normalised to data through the integral of the  $B^+ \rightarrow J/\psi K^+$  peak.

The data/MC comparison of the dimuon+track mass shapes are shown, both for the trimuon (here, the dimuon+track) vertex probability nominal selection (Fig. 4.7 *left*), and for a low vertex probability range (*right*). The latter sample is enriched in candidates where the  $J/\psi$  and the track do not come from the same vertex, which is the background targeted in this section. At low vertex probability, the underestimation in the non-prompt  $J/\psi$  MC of the high dimuon+track mass yields is clear, not yielding enough combinatorial background along the entire mass range. With the nominal `VtxProb` cut (left plot), it certainly fails to produce enough high-mass trimuons. The shape is also observed to be wrong at lower masses (more events in data at low mass), as the potential  $B_c$  signal in data (not simulated in the non-prompt  $J/\psi$  MC) would be completely dominated by random dimuon+track combinations, which are very numerous when no muon identification is required on the track.

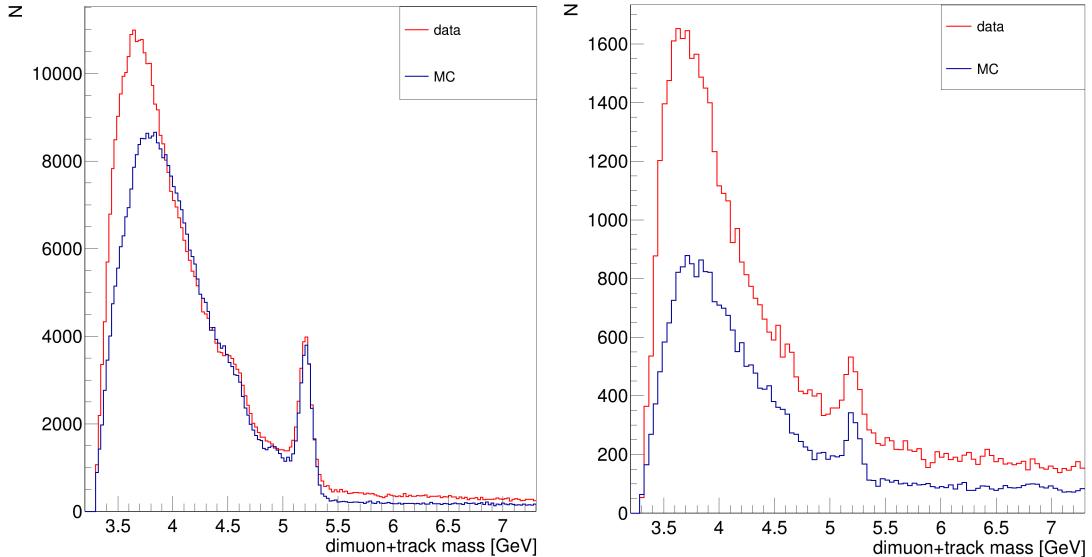


FIGURE 4.7: Dimuon+track mass shapes of non-prompt  $J/\psi$  MC candidates, with preselection cuts except the muon selection on the non- $J/\psi$  track, and normalised with the height of the  $B^+ \rightarrow J/\psi K^+$  peak, in pp collisions. *Left*: vertex probability cut of the preselection ( $VtxProb > 0.01$ ). *Right*: Low vertex probability range ( $0.01 < VtxProb < 0.04$ ).

#### 4.4.2.2 MC multiplicity

As the problem is not (or at least not fully) related to muon identification, I turned to charged track multiplicity: if it is underestimated in the MC, either globally in the full event or locally in a cone around the b quark, then the  $J/\psi$  has less possibilities to form a dimuon+track candidate that passes cuts. This would impact the yields of trimuons where the  $J/\psi$  and muon come from uncorrelated processes, but not the trimuons where the third muon comes from the ‘companion’  $\bar{b}$  shower. I look into the PYTHIA generation method for a possible bias of the multiplicity in events containing a  $J/\psi$ . The non-prompt  $J/\psi$  MC uses the option `HardQCD` which generates a  $b\bar{b}$  pair in each event; this option could trim higher-multiplicity events compared to a simulation where the hard process is just a rare event occurring in a minimum-bias (MB) generation. When a hard process is enforced with `HardQCD`, it is generated first, and then other processes (*e.g.* multiple partonic interactions) are generated with decreasing scales; and the scale of charm production is relatively low compared to underlying event processes that could but will not be produced in this mode due to higher scales than the first generated process.

I hence also generated MB events and selected only the ones with a  $b\bar{b}$  pair (about 0.1% of events), and forced a B meson to decay into a  $J/\psi$ . Fig. 4.8 (*left*) shows the charged track multiplicity associated to the chosen PV, for PYTHIA MB events, for `HardQCD` events with a non-prompt  $J/\psi$ , and for MB events that contain a B meson (forced to decay with a  $J/\psi$ ). The mean multiplicities in the last two samples are close (resp. 39 and 42), the small difference (about  $3\sigma$  significant) not being sufficient to explain the factor 2 discrepancy in the high dimuon+track mass region (reported above in Fig. 4.7). For completeness, I also checked (Fig. 4.8 *right*) the self-normalised  $J/\psi$  yields as a function of the self-normalised track multiplicity, which is a quantity often used to address how well the underlying event activity close to a quarkonia is simulated. This quantity does not show significant differences between the two generation modes.

The other possibility for a wrong track environment around the non-prompt  $J/\psi$  is an inaccurate b quark showering. PYTHIA `HardQCD` only simulates a leading-order hard process. However, in other B analyses with much more statistics, a NLO  $b\bar{b}$  generation was not found to be more useful than the LO samples to describe data. This lead is therefore not followed.

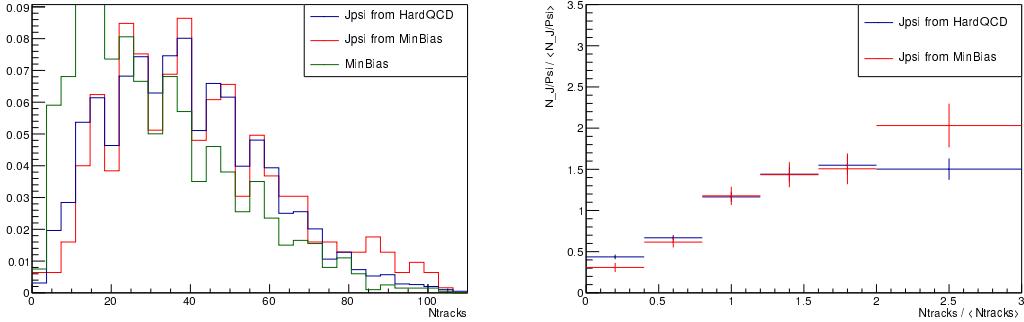


FIGURE 4.8: Multiplicities in PYTHIA events containing a non-prompt  $J/\psi$  meson. *Left*: Multiplicity of charged tracks in a MB event (green), in a MB event containing a  $B \rightarrow J/\psi X$  (red), and in a `HardQCD` event where the  $B \rightarrow J/\psi X$  generation is forced (blue). *Right*: Number of  $J/\psi$  mesons divided by the average number of  $J/\psi$  mesons in the whole sample, versus the track multiplicity in the corresponding event divided by the average multiplicity.

No strong discrepancy between `HardQCD` and MB generation modes is found.

#### 4.4.2.3 Dimuon $J/\psi$ distributions

I have hence shown that the high trimuon mass underestimation in the pp  $J/\psi$  simulation does not come exclusively from the  $J/\psi$  having to combine with a muon, nor from a lack of

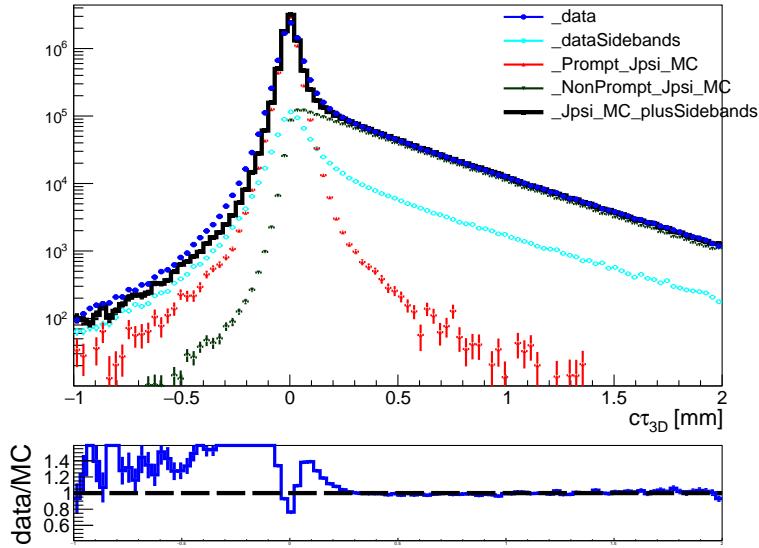


FIGURE 4.9: 3D dimuon lifetime. The data (dimuon mass signal region), the data sidebands, the prompt and non-prompt simulations, and the sum of the last three are shown. This distribution is used for the normalisation of the MC samples. The bottom pad shows the ratio between data (dark blue in the top panel) and the sum of sidebands and MC samples (thick black line).

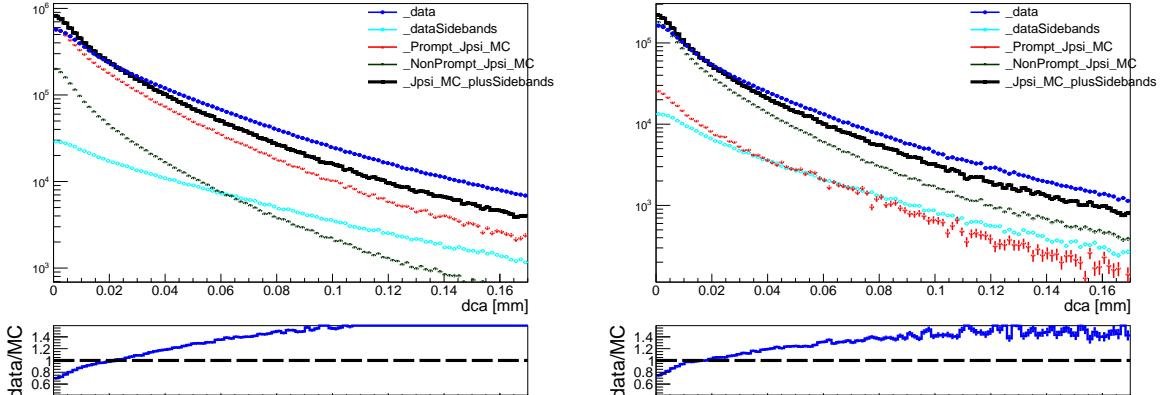


FIGURE 4.10: Distance of closest approach (dca) between the two muons, compared between data (dark blue) and the sum of sidebands and  $J/\psi$  simulations (black). *Left:* no lifetime cut. *Right:* with lifetime cuts.

charged tracks to combine with. I also saw that the PYTHIA generation mode does not affect enough the multiplicity to be the main cause of this underestimation. Therefore, the problem probably lies with the  $J/\psi$  distributions themselves, so that  $J/\psi$  mesons are less easily associated with a track in the MC, or are excessively rejected by the preselection. Problems were indeed detected with the  $J/\psi$  distributions in the 2017 pp run at 13 TeV (that preceded the 5.02 TeV run), whose source is not explicitly understood, but might have been linked to data-taking incidents in the pixel tracker commissioning (2017 RunB) or with DCDC converters (2017 RunF) that were not properly simulated in the MC. I check now that the problems in the 2017 RunG used in this analysis are similar.

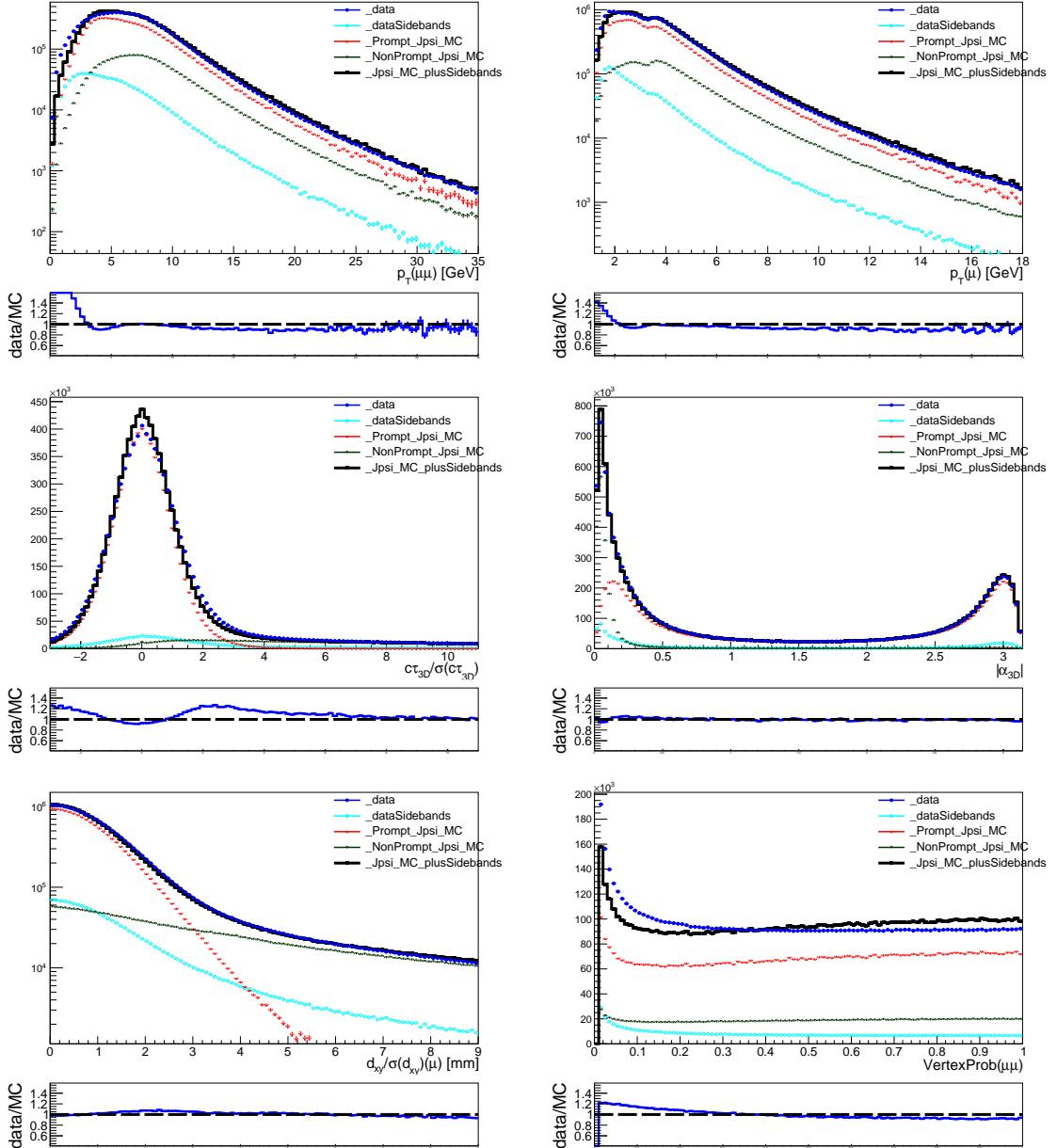


FIGURE 4.11: Comparison of  $J/\psi$  distributions between data (dark blue) and the sum of sidebands and simulations (black), without lifetime cuts. Variables are: the dimuon  $p_T$ , the  $p_T$  of muons from the dimuon, the 3D dimuon lifetime significance, the 3D dimuon angle between the momentum and the vertex displacement, the significance of the transverse displacement of muons from the PV, and the vertex probability.

`Oniatree`'s are run to gather inclusive  $J/\psi$  candidates from the 2017 RunG dataset, as well as from the prompt and non-prompt  $J/\psi$  MC described in section 3.5. I apply cuts similar to the ones on the  $J/\psi$  within the trimuon selection, meaning:

- Dimuon vertex probability  $> 0.01$ .
- Dimuon distance of closest approach  $dca < 0.3$  mm.
- Two hybrid-soft muons (defined in section 5.3.1).
- One muon passing the tight acceptance and firing the trigger, and the other might pass only the loose acceptance (single muon acceptance definitions in section 5.3.2). The event

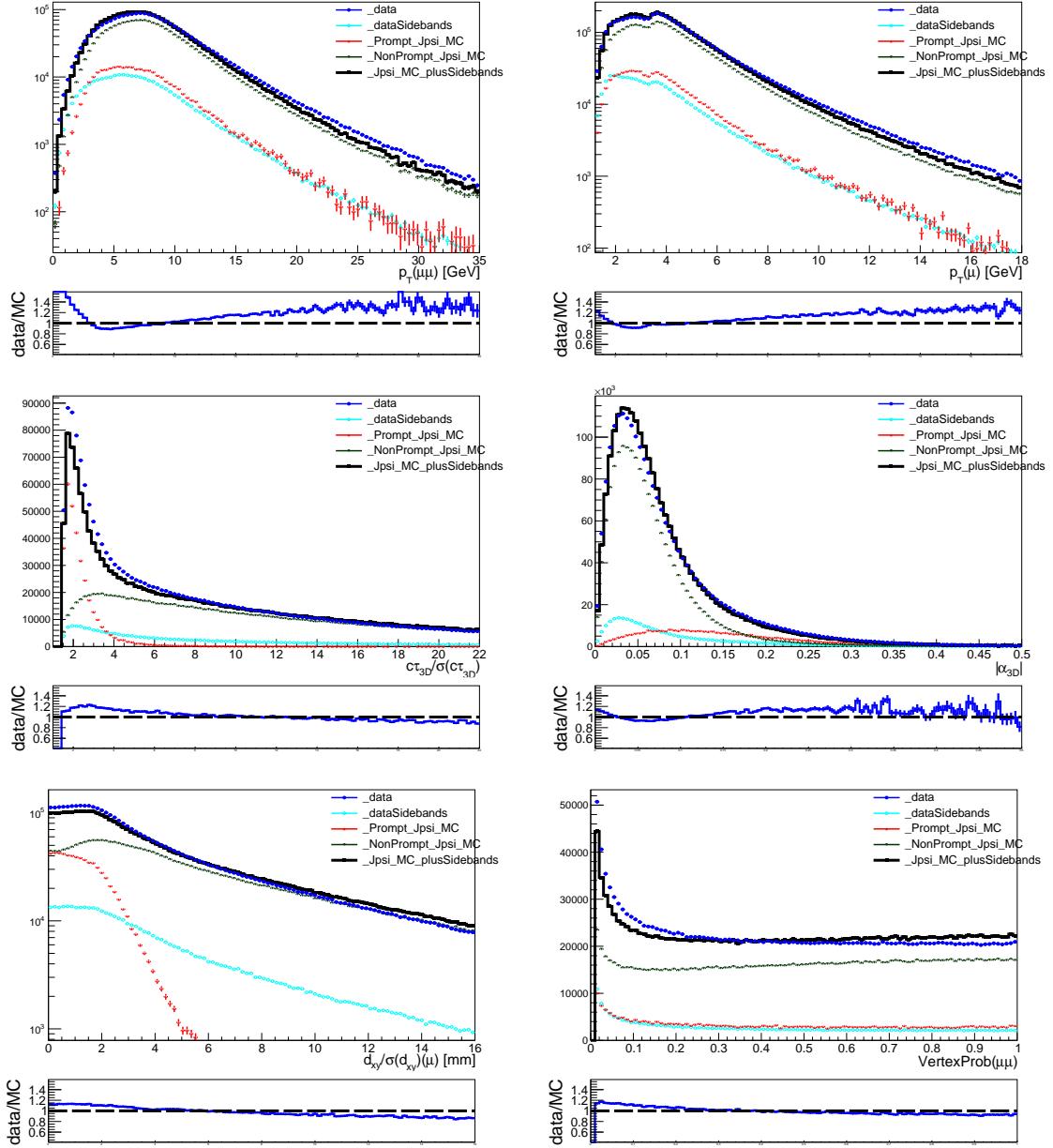


FIGURE 4.12: Comparison of  $J/\psi$  distributions between data (dark blue) and the sum of sidebands and simulations (black), with lifetime cuts applied. Variables are: the dimuon  $p_T$ , the  $p_T$  of muons from the dimuon, the 3D dimuon lifetime significance, the 3D dimuon angle between the momentum and the vertex displacement, the significance of the transverse displacement of muons from the PV, and the vertex probability.

needs to fire the double-muon trigger used in the  $B_c$  analysis.

- Dimuon mass signal and sideband regions defined as in section 4.2.

I draw variables separately for the mass signal region data, for prompt and non-prompt  $J/\psi$  simulations, and for the sidebands. Then I compare the signal region data to the sum of the sidebands and the simulations. To do this, a normalisation of the MC is needed, which I obtain in two steps, through the 3D lifetime distribution in Fig. 4.9:

- Scale the non-prompt MC (plus the sidebands with fixed normalisation) to data events at high lifetime ( $c\tau_{3D} > 0.9$  mm)

- Scale the prompt MC (plus the sidebands with fixed normalisation) to data events in the complementary region, after having subtracted the scaled non-prompt contribution.

I compare the data and MC distributions with the cuts stated above, but also with cuts on the dimuon lifetime such that the dimuon lifetime distribution is close to the one of the preselected trimuon sample (with cuts of section 5.4). The first sample is dominated by prompt  $J/\psi$ , and the second by non-prompt  $J/\psi$ . Discrepancies in the second sample are directly meaningful for this analysis. For instance, Fig. 4.10 presents the distance of closest approach ( $dca$ ) between the two muons, for the two sets of cuts, and shows large data/MC discrepancies (up to 50%).

Other figures (4.11 for inclusive and 4.12 for non-prompt-enriched samples) reported hereafter show other variables, with the following also exhibiting discrepancies: dimuon vertex probability and lifetime significance for both selections, and the dimuon  $p_T$  and 3D pointing angle for the sample with lifetime cuts (see section 5.4 for the definition of these variables).

The correlation matrices between the  $J/\psi$  variables for data and for prompt and non-prompt MC samples are also presented in Fig. 4.13 for the two sets of cuts. They show quite high correlations which make a reweighting strategy particularly difficult, especially considering the limited statistics of the 2017 pp RunG (in particular if the  $J/\psi$  is required to be in the vicinity of a third muon), which mean that the number of reweighted variables must be limited.

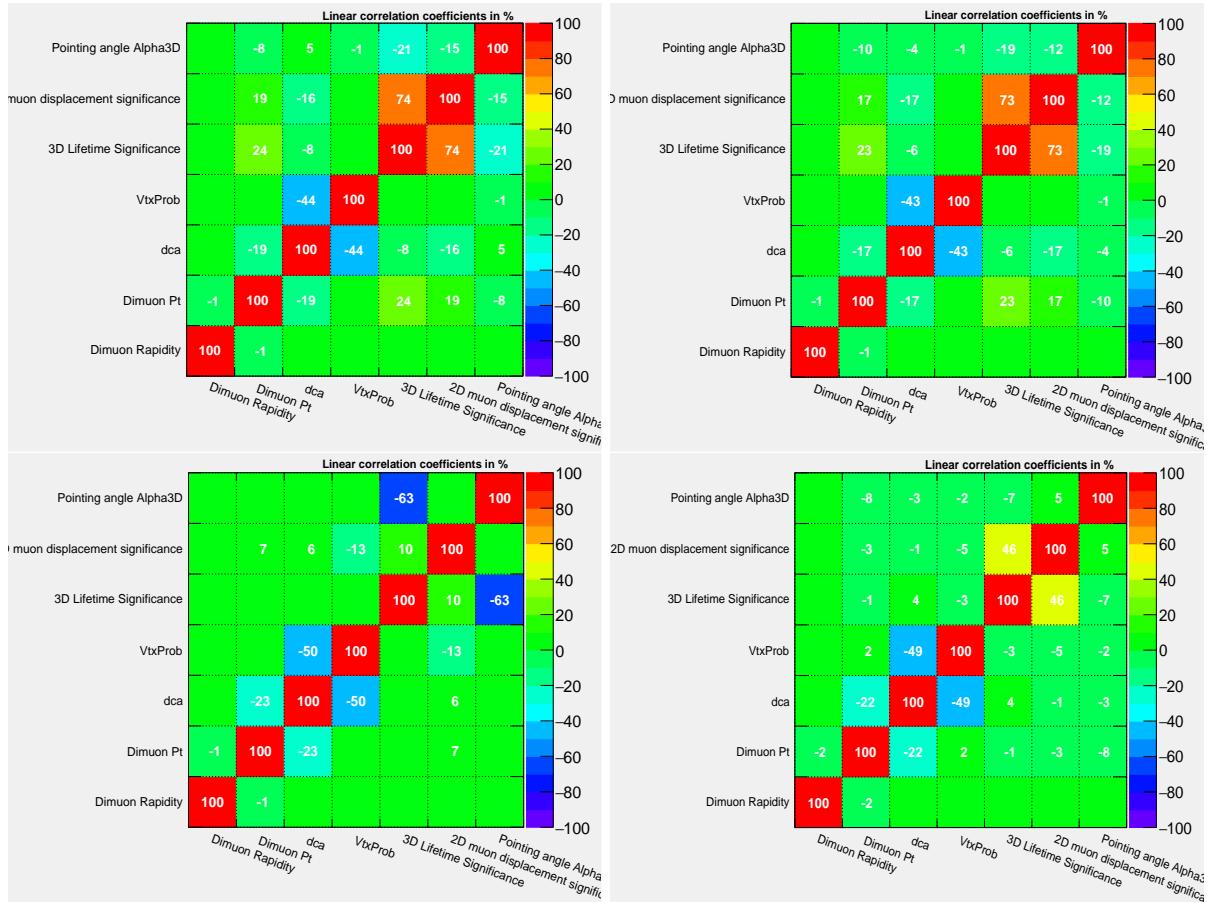


FIGURE 4.13: Correlation matrices between  $J/\psi$  variables, for non-prompt MC (*top*), and for prompt MC (*bottom*). For brevity, the correlation matrices for sidebands-subtracted data are not displayed, but show correlations of the same order than those of MC. *Left:* no lifetime cuts. *Right:* non-prompt lifetime cuts are applied.

### 4.4.3 $\text{J}/\psi$ rotation method

The  $\text{J}/\psi$  MC samples were shown above not to be able to describe the background from  $\text{J}/\psi$  mesons combining with muons from a different hard process. Therefore, let us now rather look for a data-driven method. The background at hand can be decomposed in two categories:

- The random combinations of a true  $\text{J}/\psi$  with a (fake or true) muon whose production is **uncorrelated** with the  $\text{J}/\psi$ . There are multiple cases that will not need to be distinguished:
  - a prompt  $\text{J}/\psi$  passing the displacement selection + a (non)prompt muon;
  - a non-prompt  $\text{J}/\psi$  + a prompt muon whose track is close to the  $\text{J}/\psi$  vertex.
- When the muon is prompt, it is mostly a misidentified hadron, especially in PbPb collisions where the track multiplicity at the primary vertex is very large.
- The combinations of a true  $\text{J}/\psi$  with a (fake or true) muon that shows some level of correlation with the  $\text{J}/\psi$ . One can think of two **loose correlations**:
  - particles linked by their parent hard process, in particular  $b\bar{b}$  processes where one  $b$ -hadron decays in a  $\text{J}/\psi$  meson, and the companion  $\bar{b}$ -hadron decays in a muon that is wrongly associated to the  $\text{J}/\psi$  vertex;
  - particles linked by the longitudinal boost of the center-of-mass of the parent partonic interaction.

The first category justifies well a data-driven approach. I envisaged the widely-used *event-mixing* method. It consists in combining a  $\text{J}/\psi$  from a given event with a muon from another event, whose properties are similar to the original event. One hence needs to categorise the events in terms of the position of the PV (and possibly of the secondary vertex as well), multiplicity/centrality, and event plane angle in PbPb. This method is however technically heavy (in particular for a displaced decay) and it can be biased by the choice of the categories which need a lot of tuning to encompass all relevant event characteristics while keeping a reasonable number of binning dimensions. Therefore, it was not deployed in this analysis.

Event-mixing also has the caveat to ignore the second background type above, which should be quite hidden by the numerous nucleon-nucleon interactions in PbPb collisions, but important in pp collisions. To estimate its relative contribution, several rotation angles will be used in the method presented in the following. To describe at least part of this **loosely correlated background**, one needs to leave the  $\text{J}/\psi$  in the same event, but break the strong correlation between the  $\text{J}/\psi$  and its parent and sister particles. This is done by rotating the  $\text{J}/\psi$  mesons, with angles large enough to make the original decay ( $\text{J}/\psi + X$ ) too broken to be reconstructed; but when kept also relatively small, some angles can capture the loose correlations of the  $\text{J}/\psi$  with the surrounding event activity (low- $p_T$   $b\bar{b}$  pairs with a large opening angle, or multiplicity increased by the boosted center-of-mass of the hard process).

The rotation method takes all  $\text{J}/\psi$  candidates in the triggered dataset, rotates the direction of their momentum and of their displacement vector (from the primary to the dimuon vertex) by a common angle, and then lets the rotated  $\text{J}/\psi$  candidates combine with muons from the same event; in other words I run the ONIATREE to reconstruct trimuons in inclusive  $\text{J}/\psi$  data events where the  $\text{J}/\psi$  was rotated. Thirteen rotation angles are tried: for eight rotations the sign of  $\eta$  is reversed and the azimuthal angle is changed by  $\Delta\phi = -\frac{3\pi}{4}, -\frac{\pi}{2}, -\frac{\pi}{4}, 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}$ , or  $\pi$ , and for five other rotations the pseudorapidity is unchanged and the azimuthal angle is modified by  $\Delta\phi = -\frac{3\pi}{4}, -\frac{\pi}{2}, -\frac{\pi}{4}, \frac{3\pi}{4}$ , or  $\pi$ . The rotations that do not change  $\eta$  have a large enough angle to limit contamination from  $\text{J}/\psi$  mesons combining with particles from the same decay or shower. Two rotations of angles  $\Delta\phi_1 = \pm\Delta\phi_2$  are of course expected to give samples of the exact same nature, but considering both provides more statistics. The absolute pseudo-rapidity is kept unchanged, to probe similar event and detector regions (as CMS is a cylindrical detector). Mass

sideband subtraction weights ( $\pm 1$  depending if the dimuon mass is in the signal or sidebands region defined in section 4.2) are applied so that the sample has, in average, the properties of only true  $J/\psi$  dimuons. The entire trimuon mass range of study is kept ( $3.3 - 7.8$  GeV). The nominal background sample will contain the  $J/\psi + \mu$  combinations from all the rotation angles, with weights of  $1/n_{\text{rotations}} = 1/13$ .

A possibly worrisome case is if the  $J/\psi$  is not rotated enough, and the  $B \rightarrow J/\psi \mu X$  decay it originates from still passes the preselection despite the  $J/\psi$  rotation. It is however not problematic, as the  $J/\psi + \mu$  vertex and invariant mass will still have background-like properties (similar to the background one seeks to describe) because the  $J/\psi$  was moved away from its muon sister. Therefore it will not be a double-counting of the background with  $J/\psi$  and muon coming from the same  $B$  vertex (section 4.3), but just a change of normalisation of the background described in this section; this can be ignored as the normalisation is a free parameter in the template fit in pp, where those cases could contribute significantly compared to the purely combinatorial background (contrarily to PbPb where it is a small contribution, as shown below).

The rotation method could also be run on the signal  $B_c$  MC to check the level of expected signal contamination in the rotated  $J/\psi$  sample. However, as for the  $B$  decays, the properties of potential  $B_c$  events that would still pass the preselection despite the rotation of the  $J/\psi$  daughter would be far from the ones of signal and close to the ones of the background described in this section. In addition, these cases are expected to be very rare, so this check has not been done.

The resulting **rotated  $J/\psi$  yields** are shown for each rotation angle in Fig. 4.14, where a dimuon-sidebands subtraction and preselection cuts are applied. In PbPb collisions (*right*), the obtained yields do not depend significantly on the rotation angle. All angles give similar yields because the correlated effects are swamped in the underlying event, meaning that a majority of candidates are random combinations of a  $J/\psi$  and a muon from a different hard process. This makes the method particularly reliable in PbPb, because the loosely correlated background (mostly from other  $B$  decays close to a  $J/\psi$ ) is negligible – whereas it is important in pp collisions.

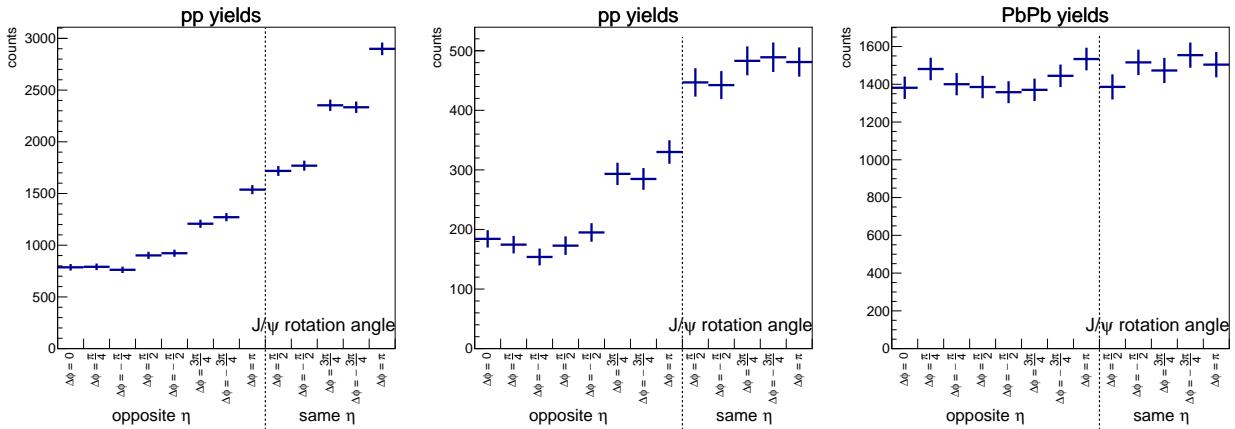


FIGURE 4.14: Yields for the thirteen  $J/\psi$  rotation angles, the first eight featuring a  $\eta$  reversal, and the last five with only a  $\phi$  rotation. *Left:* pp. *Center:* pp, restricted to the high mass region. *Right:* PbPb. There are no significant differences of trends in the PbPb high-mass or complete mass regions.

And indeed for pp (Fig. 4.14, *left*), a large yield variation is observed versus the angle. Two effects are revealed:

- About twice more events are observed in pp when the  $J/\psi$  pseudorapidity is not changed (last five points) than when it is reversed (first eight points). This higher activity on the side of the original  $J/\psi$  can be understood as an effect of the longitudinal boost of

the partonic center-of-mass, which drags more particles near the rapidity of the hardest process.

- Yields are increasing when the rotation angle is getting closer to  $\Delta\phi = \pi$ . Momentum conservation in the transverse plane raises the probability for the  $J/\psi$  (from a b quark) to find a muon partner (from a companion  $\bar{b}$  quark) when rotated to opposite  $\phi$ .

In the middle plot, the yields are restricted to the high-mass control region (6.3–7.8 GeV) for pp collisions. The second effect (close to  $\Delta\phi = \pi$ ) is dampened because these trimuon candidates have larger opening angles, meaning that the angle correlations are smeared out. It is however remarkable that the yield obtained from the non-rotated data in the same mass region (320 after sideband subtraction to cancel the fake  $J/\psi$  contribution) is of the same magnitude as the ones obtained for the close-to-opposite- $\phi$  and opposite- $\eta$  rotations (around 300), and roughly in the center of the range of values taken by the different rotation angles. This gives confidence that this method finally captures the high-mass background that was not well described in MC, and that an average over rotation angles is a good approach for the nominal background description.

The trimuon mass shapes for different rotation angles are compared in Fig. 4.15 in pp and PbPb collisions. In the PbPb case, the mass shapes show no significant differences. This confirms that the rotation angle does not matter and that the background is mostly random combinations of a  $J/\psi$  and a muon, uncorrelated but happening to have a consistent common vertex. I will therefore not vary the rotation angle for systematic study of this shape, and just take the average over candidates from all angles as the nominal rotated- $J/\psi$  shape. Moreover, the yield studies above show that the normalisation of this sample is well-controlled and close to the yields of the high-mass control region; this justifies to fix its normalisation in the fit to the yield of the data-driven method.

In pp, despite the different total yields, no clear difference appears in the shape between the rotations on the same- or opposite- $\eta$  side (Fig. 4.15, *top left*). However, the shape for large  $|\Delta\phi|$  rotations is different than other  $\phi$  rotations (*right*): it is skewed to lower masses. This probably reflects the association of a non-prompt  $J/\psi$  with decay products of the companion B, going back-to-back (in the transverse plane) compared to the B parent of the  $J/\psi$  (the small masses coming from small angles between the companion B and the rotated  $J/\psi$ ). In the pp template fit, I will allow the rotated- $J/\psi$  shape to vary between these two shapes, which is a way to leave free the ratio of the normalisations of the combinatorial and the loosely correlated background. The global normalisation will be left quasi-free too, because the choice of angles considered in the nominal background sample affects the rotated- $J/\psi$  yields in pp (see Fig. 4.14).

Other discriminant variables (used in the preselection and in the BDT training) were compared in the high-mass control region between the full rotated- $J/\psi$  sample and data (no signal being expected in this mass region). In the limits of the low statistics of data in this mass region, few to no significant differences were observed. The BDT output variable, as it is the variable used to separate signal and background in preselected data, is the most important to check; this is done in section 5.7.4, comparing data to the sum of all backgrounds and expected signal, and a weighting of this variable is applied in pp.

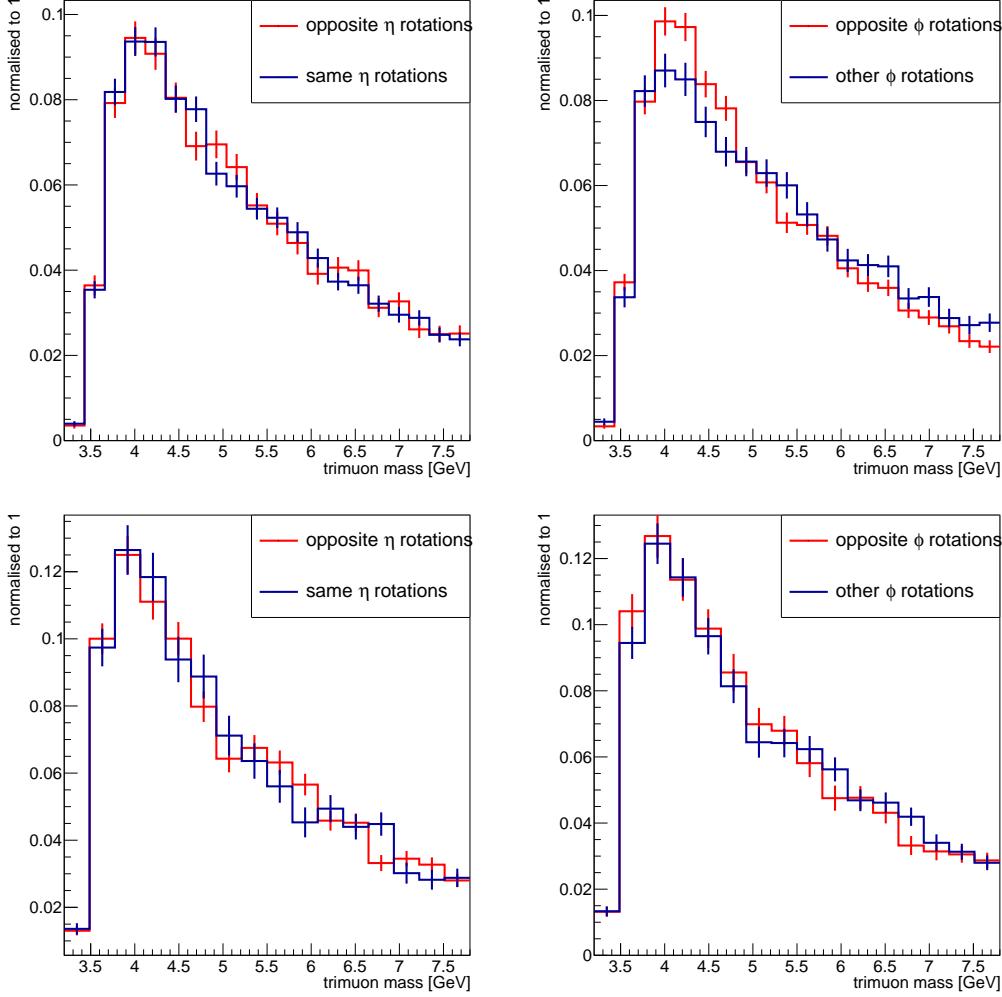


FIGURE 4.15: Trimuon mass for rotated- $J/\psi$  candidates in pp (*top*) and PbPb (*bottom*), compared for different rotation angles. *Left:* Opposite- $\eta$  side rotations versus same- $\eta$  side. *Right:* Large  $\phi$  rotations ( $|\Delta\phi| \geq \frac{3}{4}\pi$ ) versus smaller  $\phi$  rotations ( $|\Delta\phi| \leq \pi/2$ ).

To conclude, the fake  $J/\psi$  background is described with the dimuon mass sidebands (section 4.2), and the backgrounds containing a true  $J/\psi$  are nominally described by two samples: the fully correlated part of the non-prompt  $J/\psi$  MC sample ( $J/\psi$  and muon from the same  $B$  decay, section 4.3) and the rotated- $J/\psi$  sample, averaged over candidates from all thirteen rotation angles (this section). In pp, systematic variations of the shape of the last background are taken as the rotated- $J/\psi$  with different rotation angles. In PbPb the full prompt plus non-prompt  $J/\psi$  (excluding the true  $B \rightarrow J/\psi \mu X$  decays) MC is taken as a shape variation, as it is believed to be accurate in PbPb. This simulated background is compared to the rotated- $J/\psi$  sample in Fig. 4.16, where only a small difference (probably due to an imperfect relative normalisation of the prompt and non-prompt MC's) is showing for PbPb, which adds to our confidence in the method in PbPb. The difference is much larger in pp, confirming that the  $J/\psi$  MC is not to be used as such there. These alternative choices of trimuon mass shape are taken into account with nuisance parameters in the fit, which are summarised (along with the normalisation parameters) in Table 6.1 of section 6.2.

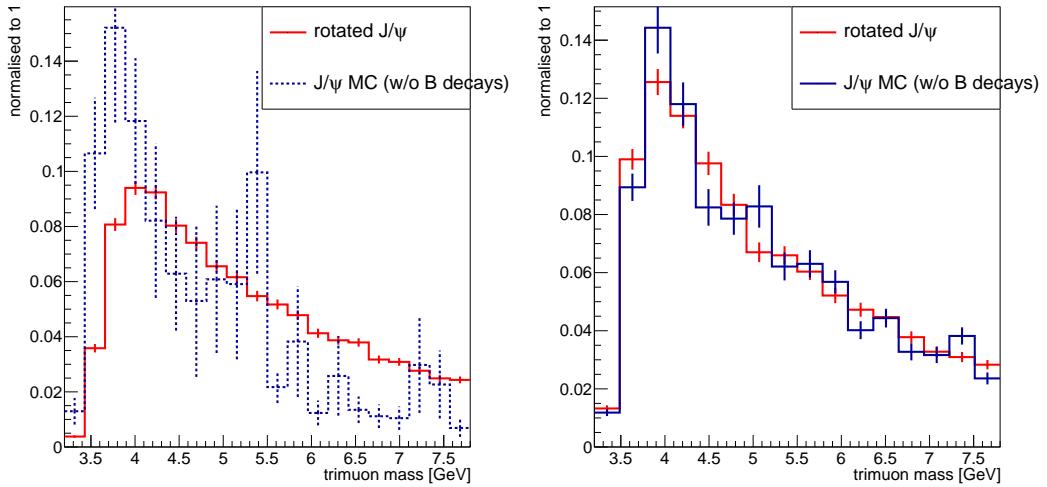


FIGURE 4.16: Trimuon mass shape for the rotated- $J/\psi$  sample compared to the sum of the prompt  $J/\psi$  MC and the combinatorial part of the non-prompt  $J/\psi$  MC (*i.e.* where the B decays are excluded), in pp (*left*) and PbPb (*right*). The prompt and non-prompt MC have their *a priori* (respective) normalisations.

“*Choisir, c’était renoncer pour toujours, pour jamais,  
à tout le reste.*”

---

André GIDE, *Nourritures terrestres*,  
Livre 4, Chapitre 1.<sup>1</sup>

## Chapter 5

# What do we seek? Selection

### 5.1 Selection overview

The dimuon-triggered datasets first endure an event-level selection to reject the non-inelastic and background events. Then a primary vertex (PV) needs to be found and chosen. These steps are addressed in section 5.2.

A signal trimuon is defined as three muons coming from a common vertex, with a total charge  $\pm 1$ , and an invariant mass between  $m_{J/\psi} + m_\mu \simeq 3.2 \text{ GeV}$  and  $m_{B_c} \simeq 6.2 \text{ GeV}$ . The trimuons of charge  $\pm 3$  are kept for cross-checking the purely random part of the fake  $J/\psi$  background (see section 4.2; this sample will be shown for illustration in the post-fit mass plots in chapter 6). The mass range used in the fit is however  $[3.5, 7.8] \text{ GeV}$ , to keep a high mass control region  $[6.2, 7.8] \text{ GeV}$ , and because the candidates below  $3.5 \text{ GeV}$  end up being overwhelmed by background and rejected very fast by the BDT (cf section 5.7). The trimuon must also feature at least one opposite-sign (OS) muon pair (there are two such pairs) whose mass is in the  $J/\psi$  mass signal region defined in section 4.2. The candidates with an OS dimuon in the dimuon mass sidebands are kept to describe fake  $J/\psi$  background. Choosing which OS dimuon is the  $J/\psi$  is ambiguous when the two OS pairs are in the signal and/or sidebands regions; this is dealt with a candidate weighting, as described in section 5.5.

A *global muon* (defined in section 3.2.3.2) is most often a *tracker muon* too. Tracker muons contain more fakes, especially in PbPb data where the underlying event results in a huge background in the inner tracker, but can be reconstructed to lower  $p_T$  which is a key lever to get enough statistics. In this analysis, two muon selections are considered: the *soft-muon* selection (described in Ref. [175]) requiring a tracker muon, and the tighter *hybrid-soft-muon* selection requiring a global muon that passes (almost) all soft-muon cuts. A strategy requiring the trimuons to feature three hybrid-soft muons was chosen over a strategy requiring two hybrid-soft and one soft muon. Section 5.3 discusses these muon selection choices and shows the single muon acceptance associated to each selection.

Many variables can help discriminating trimuons that come from a signal  $B_c$  or from background processes. They are presented in section 5.4, along with the chosen cuts; the trimuon

---

<sup>1</sup>“To choose was to renounce everything else, for ever and for never.”

candidates passing this preselection are the ones used and studied in the core of the analysis. These preselected candidates can be more finely discriminated with a *Boosted Decision Tree* (BDT), that compiles all the discriminating information from the input variables into one eponym variable. It is explained in section 5.7, in particular the used variables, the training strategy, and how the output variable is checked and/or weighted.

Section 5.6 shows our choice of the analysis bins in the trimuon transverse momentum  $p_T^{\mu\mu\mu}$ , along with the associated trimuon rapidity cuts. These  $p_T^{\mu\mu\mu}$  and rapidity cuts define the phase space for which the cross sections and modification factors results will be quoted (in chapter 9). The BDT training and testing is done separately in two  $p_T$  bins (even for the results binned in centrality rather than in  $p_T^{\mu\mu\mu}$ ).

## 5.2 Event-level and primary vertex selection

### 5.2.1 Event-level selection

A very loose selection is applied at the event level, to reject events that are only diffractive (containing only very high rapidity particles) or elastic pp or PbPb collisions, or that come from cosmic muons or noise from the detector. It is standard to most of the heavy ion analyses in CMS (see *e.g.* Ref. [43]).

In pp collisions, it consists in requiring:

- a fitted PV with at least two tracks, situated in the bunch crossing region;
- at least a quarter of the tracks to be tagged with *high-purity* (defined in section 3.2.2), except for vertices with less than 10 tracks, which are kept whatsoever. This rejects background from beam scraping, which is the interaction of the beam with residual gas in the pipe or with beam collimators (that act as diaphragms that diffract the outlier beam particles far from the beam axis), producing bad quality tracks, with momentum mostly in the  $z$  direction. Such particles giving significant tracks in the detector are mainly muons because of their long lifetime. These muons are also rejected by the muon selection cuts aimed at cosmic muons, which require the muon track to be close to the PV (see section 5.3).

In PbPb collisions, the requirements are:

- a fitted PV with at least two tracks, situated in the bunch crossing region;
- at least one tower (*i.e.* a cell covering a transverse area) above the 4 GeV energy threshold in each of the forward hadron calorimeters, which are sensitive to a large majority of the inelastic interactions. This notably rejects single-diffractive events (defined in section 1.5.5) that cannot deposit energy in both the forward and backward calorimeters, but also electromagnetic or elastic PbPb collisions (also called ultra-peripheral because the nuclei do not touch in the collision, see section 1.5.6);
- a good compatibility between the PV and the clusters (*i.e.* a group of fired cells) in the pixel detector. One considers the ratio of the vertex-clusters compatibility (*i.e.* how much the pixel detector clusters point towards this vertex) for the default PV position, and for a vertex shifted by  $\pm 10$  cm along the  $z$  axis; it must be above a threshold that depends on the number of pixel hits. This rejects beam-scraping events, and other events from detector noise.

In PbPb, the selection of the tracks used to fit the PV's during data reconstruction was tight and in some peripheral events, no PV is reconstructed, so it is cut by the PV requirement. A recovery (standard to the heavy ion analyses) of those peripheral events that do have a primary vertex is performed.

The selection done with the forward hadron calorimeters has a significant inefficiency in very peripheral collisions. This is the main reason for this analysis to measure the cross sections only in the 0–90% centrality range, where the event selection efficiency is very close to 1, especially for events that contain three muons and a well-reconstructed PV. This restriction makes easier the counting of minimum bias events, needed for the normalisation of PbPb yields (see section 3.5.5).

### 5.2.2 Main primary vertex

Multiple identification and discrimination variables are defined with respect to the main PV. A primary vertex (PV) is defined as a point consistent with the beam line ( $x = y = 0$ ), and to which multiple charged tracks are consistently pointing to. The average number of pp collisions (*i.e.* of possible PVs) per bunch crossing is called *pile-up*. The pp 2017 RunG at 5.02 TeV had a low pileup compared to other pp runs, namely  $\langle n \rangle \simeq 3.5$  in average in selected events (2.5 in minimum-bias events [137, 193]); its distribution is shown in Fig. 5.1. In other pp runs at higher energies, the objective is to accumulate the most possible luminosity rather than keeping backgrounds low, so the average pile-up reached about 38 in the 2017 pp run at 13 TeV. However even at  $\langle n \rangle \simeq 3.5$ , multiple PVs are found in most of the collisions. In PbPb the PV is (almost always) uniquely defined, because the pile-up was  $\langle n \rangle < 0.006$  in all running conditions [193], meaning that a PbPb bunch crossing already containing one actual PbPb collision (which concerns less than 0.6% of the bunch crossings!) has less than 0.3% probability (calculated from a Poisson distribution of mean  $\langle n \rangle$ ) to feature two or more PVs.

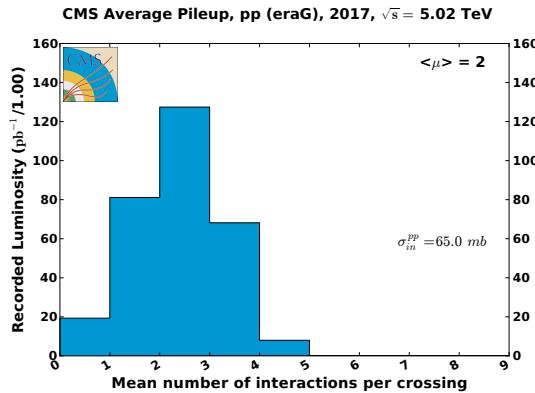


FIGURE 5.1: Distribution of pile-up (number of pp interactions per bunch crossing) for the pp reference run at 5.02 TeV. Figure from Ref. [137].

The multiple PVs in pp are spread along a longitudinal region of about 25 cm, which means the 2.5 additional collisions (in average) are most of the time a few centimeters apart. This is significantly more than the typical flight distance of the B meson we look for: even a  $B^+$  with a boost factor of 10 (momentum  $\sim 50$  GeV) travels only 5 mm, and the  $B_c$  has a three times smaller lifetime (and many selected candidates have a momentum smaller than 50 GeV). Hence, if there are multiple PVs for a given  $J/\psi$  candidate, we choose the main PV (hereafter called ‘the PV’) to be the one closest to the dimuon vertex in terms of longitudinal distance ( $z$ ); the numbers above ensure that this is the vertex that the displaced  $J/\psi$  actually comes from, except in less than 1% of the cases. If the wrong PV is chosen, then the variables associated to the displacement of the candidate (lifetime and pointing angle) will be somewhat wrong – but not dramatically different, as the wrongly chosen PV is close to the right one along  $z$ . In addition, even if this candidate still manages to pass the preselection, the variables will be similarly wrong for signal and background, meaning it should not significantly affect the fit procedure.

In pp, for each dimuon candidate, the PV is refitted without its two muon tracks (if there are at least two additional tracks to be fitted), to remove the potential influence of a displaced dimuon on the PV position. In PbPb, there are in general many more tracks associated to the PV, so this bias is considered negligible (and the PV refitting would be computationally heavy).

## 5.3 Single muon acceptance and selection

### 5.3.1 Muon selection

The main muon selection of this analysis is the *hybrid-soft identification*. It requires that:

- The muon passes both the *global* and *tracker* identifications;
- $d_{xy} < 0.3\text{ cm}$  and  $d_z < 20\text{ cm}$ , where  $d_{xy}$  and  $d_z$  are the minimal distance between the best muon track and the main PV, projected on the transverse plane or on the beam direction, respectively;
- At least six layers of the inner tracker have hits associated to this track. This is necessary for a good  $p_T$  measurement and to reject decay-in-flight muons (whose tracks show a kink where the pion or kaon decays into a muon, see section 3.2.3);
- Among the inner tracker hits associated to the muon, at least one must come from one of the three layers of the pixel detector (the part of the inner tracker closest to the beam). This also helps rejecting decay-in-flight muons.

The required *global* identification rises the  $p_T$  threshold of the identified muons: it requires a standalone track coming from hits in muon stations, so global muons must have relatively high momentum not to be deviated out of the acceptance of the muon stations or loose all its momentum before reaching the second station. The alternative is the *soft identification*, similar to *hybrid-soft*, except that the *global* requirement is replaced by:

- a tracker muon whose track tightly matches with hits in one muon station;
- a *high-purity* requirement on the inner track.

This second algorithm gives many more fakes, especially in PbPb where the occupation of the inner tracker is very high (so that random hits that align by chance might be fitted as a track), but has a much higher efficiency at low- $p_T$ .

### 5.3.2 Muon acceptance

Different aspects of the CMS detector, such as its material budget between the tracker and the muon chambers (that makes low-energy muons loose a significant fraction of their energy), its intense magnetic field strongly bending low-momentum tracks, and its limited geometrical coverage, introduce an effective muon momentum threshold depending on pseudorapidity (typically 1 to 3.5 GeV). The muon reconstruction efficiency is therefore very low (or null) and non-uniform in some kinematic regions. We want to exclude those by defining acceptance regions in the  $(p_T, \eta)$  plane where the muon efficiency is high enough, so that the efficiency of trimuons whose three muons are in the muon acceptance is not too low. If the trimuon efficiency reaches very low values, the associated corrections to be applied to the observed yields (the inverse of the efficiency, see chapter 7) are too high and vary too fast with  $p_T$  and  $\eta$ , which is to be avoided in finite-size bins (see section 7.3). In addition, the corrections to the single muon efficiencies, from the data/MC discrepancies calculated with a tag-and-probe method in section 7.1, are large and not well-known in these low-efficiency regions.

This section presents muon kinematic cuts (and how they were determined) that avoid such low-efficiency regions, for different muon selections. The choice of identification criteria (see

section 5.3.3) and the determination of the associated kinematic acceptance was however lead carefully to obtain single-muon  $p_T$  thresholds as low as possible, as the muons from  $B_c$  decays have relatively low  $p_T$ . The generated  $B_c$  mesons indeed have a trimuon  $p_T$  peaking at  $\simeq 3$  GeV. The three muons typically do not share equally the available  $p_T$ , one of them often having  $p_T < 2$  GeV: Fig. 5.2 (*left*) shows that when the two muons with highest  $p_T$  are well reconstructed (as *hybrid-soft*), the third (softest) muon has a very low  $p_T$  compared to the typical muon acceptance of CMS. Simple muon  $p_T$  cuts have a dramatic inefficiency on signal. To illustrate this, if we apply the tight acceptance cut detailed at the end of this section (in which the lowest reachable  $p_T$  is 1.5 GeV) on the softest muon only, when the other two pass *hybrid-soft*, only 16% of the signal yields are preserved. In addition, interesting enhancement phenomena could take place at low  $p_T(B_c)$  in PbPb (see section 2.5). Lowering the muon  $p_T$  acceptance thresholds, and allowing a low- $p_T$  acceptance for one of the muons, is therefore crucial to increase the observable  $B_c$  yields. In particular, the standard  $p_T$  thresholds in the  $1.2 < |\eta| < 2.4$  rapidity region for *hybrid-soft* muons passing the  $J/\psi$  trigger were lowered by 0.3 GeV compared to the ones that were used for similar analyses on 2015 PbPb data.

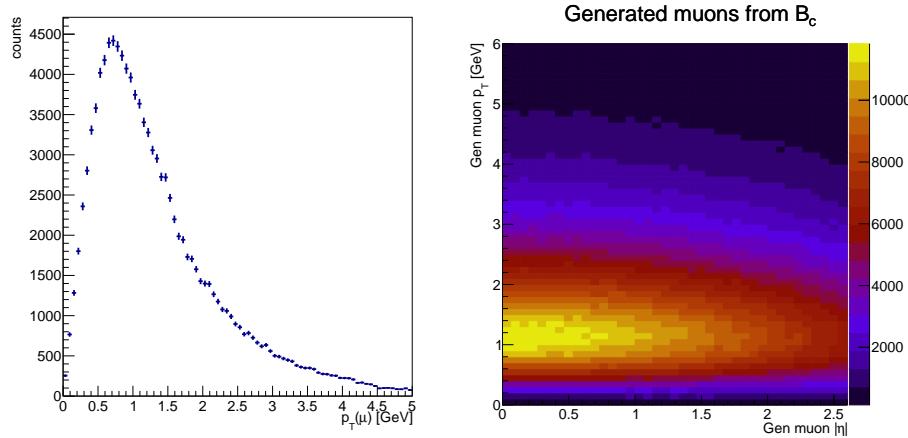


FIGURE 5.2: *Left:*  $p_T$  of the softest muon from generated  $B_c$  trimuons, when the two other muons are already reconstructed as *hybrid-soft*. *Right:*  $(p_T, |\eta|)$  distribution of any muon from the generated  $B_c$  mesons; this can be compared to the acceptance maps shown in the next figures. A  $B_c$  simulation (without PbPb medium modification) without any kinematic cuts on the generated muons from  $B_c$  is used in both plots.

The reader is to be warned here that the analysis was actually performed with three *hybrid-soft* muons, among which at least two need to fire a dimuon trigger, but the possibility to release the selection of the non-triggering muon to the *soft* identification was seriously considered; the choice is discussed in section 5.3.3. We thus describe here three sets of single-muon acceptance cuts corresponding to three muon selections: *soft* muons, *hybrid-soft* muons, and *hybrid-soft* muon that participate in firing the double muon trigger. Even if the  $B_c$  analysis eventually used only the last two, the three selections and the kinematic acceptance cuts associated to them (and determined in this section) are standard to reconstruct low- $p_T$  muons in 2017 and 2018 CMS heavy ion data.

The single-muon reconstruction, selection, and trigger efficiencies are studied in pp and PbPb, in order to define acceptance cuts in the  $(p_T, |\eta|)$  plane that keep only regions with a total efficiency above 10%. This 10% threshold is somewhat arbitrary, but means the worst possible efficiency due to muon selection for trimuons with accepted muons is 0.1%, though in practice it only reaches 1% (resp. 7%) in the less efficient corner of the fiducial phase space in PbPb (resp. pp), as can be seen in the efficiency maps of section 7.4. It is stressed here that the single-muon efficiency maps shown below serve only to determine the acceptance cuts, and are not directly used in the final acceptance and efficiency estimation of the trimuons.

The single-muon efficiencies are here measured on prompt J/ $\psi$  simulations in pp and PbPb conditions, described in section 3.5. The J/ $\psi$  process is embedded in minimum-bias collisions in PbPb and in pile-up collisions in pp; it is also corrected to data by the scale factors calculated with a tag-and-probe method in section 7.1. No kinematics cuts are applied to the dimuons `OniaTrees` extracted from this dataset. The PbPb MC events (generated with a flat centrality) are weighted with  $N_{\text{coll}}$  so that it has the centrality distribution of minimum-bias data. The reconstructed muons are matched to the generated ones if they respect the conditions  $\Delta R < 0.05$  and  $\frac{\Delta p_T}{p_T} < 0.5$  (and the reconstructed muon with smallest  $\Delta R$  is preferred if multiple matches are found).

For the reconstruction and identification as *hybrid-soft*, all generated muons are used as probes, and the efficiency is the fraction of the probes that are matched to a *hybrid-soft* reconstructed muon; and similarly for the *soft* identification. The efficiency of the reconstruction and identification of *soft* muons is shown in Fig. 5.3, and the one of *hybrid-soft* muons in Fig. 5.4, in pp and PbPb conditions. The efficiency is calculated in bins of  $p_T$  and  $|\eta|$  of the muon, and the regions where it is below 10% (that should be rejected) are printed in pale pink. The corresponding acceptance cuts, detailed below, are superimposed.

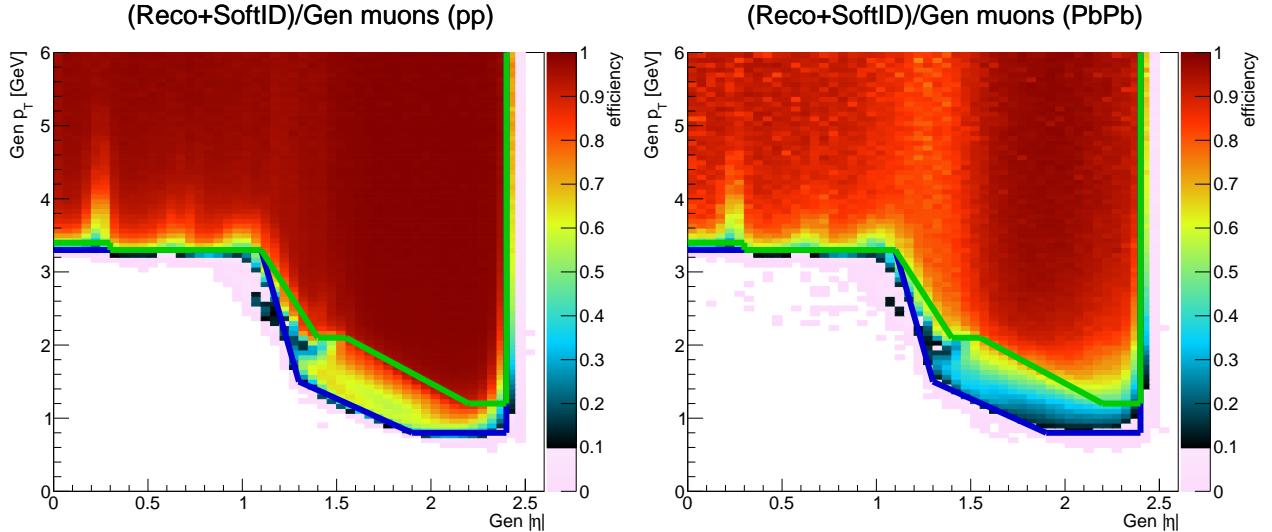


FIGURE 5.3:  $(p_T, |\eta|)$  maps of the efficiency of single-muon reconstruction + *soft-muon* identification. *Left:* pp, *right:* PbPb. The blue line shows the chosen acceptance cuts for the *soft* muons that are not required to trigger; it is to be compared to the acceptance for *hybrid-soft* identification in green (which is decided based on the maps of Fig. 5.4).

The next measured efficiency is for a single muon to be *hybrid-soft* and to fire the (double) muon trigger, when there is another muon in the event that can help firing it. The evaluated triggers are the ones used for the  $B_c$  analysis, presented in section 3.5.1. The pp trigger requires two muons at the L1 (hardware) level without  $p_T$  cuts. The PbPb trigger requires a L3 muon with quality cuts (`L3Mu0_NHitQ10`), plus a L2 muon (`L2Mu0`), without  $p_T$  cuts. For the PbPb case, to be conservative, the  $\varepsilon > 10\%$  criterium must be passed considering the tighter L3 branch of the trigger (as only a single-muon efficiency can be checked with this method). The efficiency of the L2 branch, slightly looser than the L3 one, are found to give similar maps. The efficiency map for the L1 (hardware-only) trigger has been calculated and shows significantly higher efficiencies than the L3 trigger, due to some inefficiency of the software trigger. The cuts of the PbPb trigger on the dimuon mass and  $\Delta R$  are ignored, but known to be very efficient.

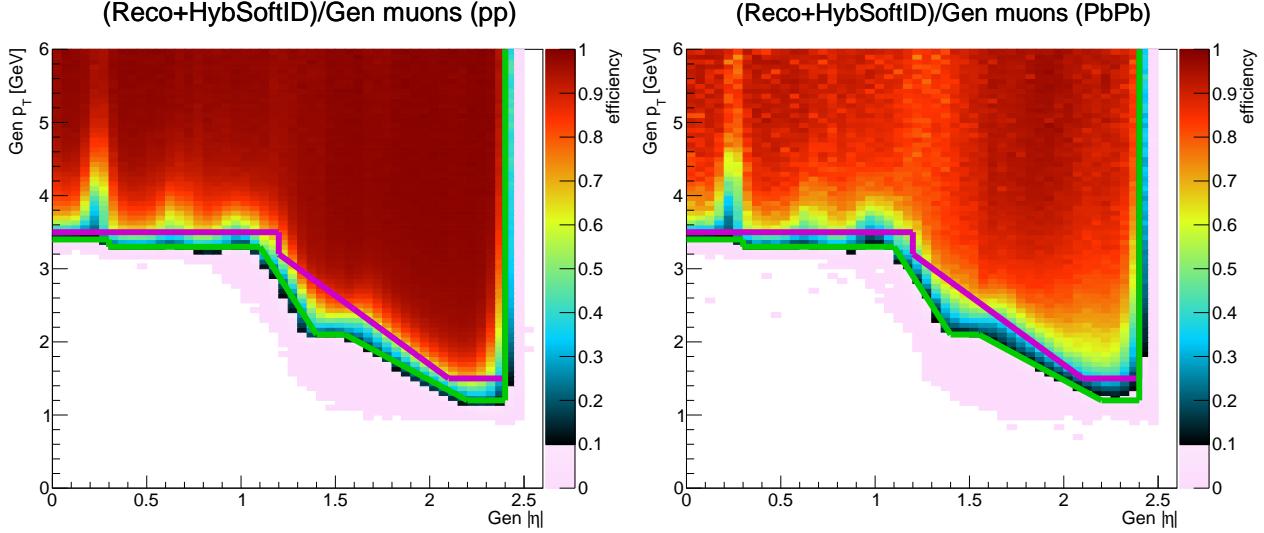


FIGURE 5.4:  $(p_T, |\eta|)$  maps of the efficiency of single-muon reconstruction + *hybrid-soft* identification. *Left:* pp, *right:* PbPb. The green line shows the chosen acceptance cuts for the *hybrid-soft* muons that are not required to trigger; it is to be compared to the acceptance for triggering muons in purple (which is decided based on the maps of Fig. 5.5).

As a double muon trigger is to be probed, looking directly at a single muon efficiency is wrong. Let us consider an extreme case, where the first muon is fully in the acceptance, and the second muon fully out of it; then the dimuon trigger would not fire (assuming there are only two muons) and probing the trigger efficiency on the first muon would artificially bring it down, despite the fact it lies in a high-efficiency region. To avoid this, we apply a simplified tag-and-probe method to emulate the efficiency of a single muon that fired a double muon trigger. Only generated  $J/\psi$  are considered, and one of the muon must pass tag cuts, the second muon being the probe on which the efficiency is measured. The tag must pass *hybrid-soft* and fire the L3Mu3 single-muon trigger in pp, or the L3Mu3\_NHitQ10 one in PbPb; these triggers, tighter than any branch of the double-muon triggers, guarantee that the tag provides a second HLT muon (companion of the possibly-triggering probe) able to fire double-muon trigger. The resulting pp and PbPb maps are shown in Fig. 5.5.

Three sets of acceptance cuts are hence established and shown in Figs. 5.3, 5.3, and 5.5, following the  $\varepsilon > 10\%$  criterium. As the pp efficiencies are higher than in PbPb, it is sufficient to check that this criterium is passed in PbPb, and the same cuts are kept in pp collisions for simplicity, and to avoid adding differences in the kinematic distributions of pp and PbPb events.

The single-muon kinematic acceptance cuts determined from the efficiency map of Fig. 5.3 (and represented there by a blue line) correspond to the **soft muon** selection, and are:

- $(0 < |\eta| \leq 1.1 \text{ } \& \text{ } p_T > 3.3 \text{ GeV}) \text{ OR}$
- $(1.1 < |\eta| \leq 1.3 \text{ } \& \text{ } p_T > (13.2 - 9.0) \times |\eta| \text{ GeV}) \text{ OR}$
- $(1.3 < |\eta| \leq 2.4 \text{ } \& \text{ } p_T > 0.8 \text{ GeV} \text{ } \& \text{ } p_T > (3.02 - 1.17 \times |\eta|) \text{ GeV})$

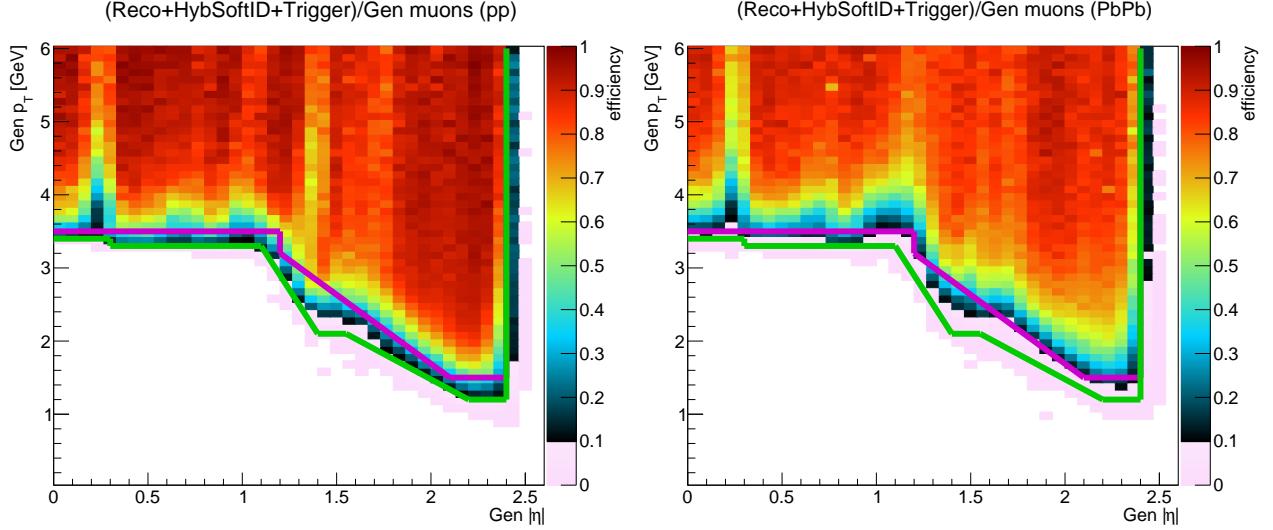


FIGURE 5.5:  $(p_T, |\eta|)$  maps of the efficiency of *hybrid-soft* identification + triggering of the L1Mu0 (pp, *left*) or L3Mu0\_NHitQ10 (PbPb, *right*) branch of the double muon trigger, when another muon able to fire the trigger is present. The purple line draws the acceptance cuts for triggering muons; it is to be compared to the acceptance for *hybrid-soft* muons that are not required to trigger, in green.

The acceptance cuts for a ***hybrid-soft* muon** (green line in Figs. 5.4 to 5.5, hereafter designated as ‘loose acceptance’) are:

- $(0 < |\eta| \leq 0.3 \text{ } \& \text{ } p_T > 3.4 \text{ GeV}) \text{ OR}$
- $(0.3 < |\eta| \leq 1.1 \text{ } \& \text{ } p_T > 3.3 \text{ GeV}) \text{ OR}$
- $(1.1 < |\eta| \leq 1.4 \text{ } \& \text{ } p_T > (7.7 - 4.0 \times |\eta|) \text{ GeV}) \text{ OR}$
- $(1.4 < |\eta| \leq 1.55 \text{ } \& \text{ } p_T > 2.1 \text{ GeV}) \text{ OR}$
- $(1.55 < |\eta| \leq 2.2 \text{ } \& \text{ } p_T > (4.25 - 1.39 \times |\eta|) \text{ GeV}) \text{ OR}$
- $(2.2 < |\eta| \leq 2.4 \text{ } \& \text{ } p_T > 1.2 \text{ GeV})$

The acceptance cuts for a ***hybrid-soft* muon that is required to trigger** (purple line in Figs. 5.4 and 5.5, hereafter designated as ‘tight acceptance’) are:

- $(0 < |\eta| \leq 1.2 \text{ } \& \text{ } p_T > 3.5 \text{ GeV}) \text{ OR}$
- $(1.2 < |\eta| \leq 2.1 \text{ } \& \text{ } p_T > (5.47 - 1.89 \times |\eta|) \text{ GeV}) \text{ OR}$
- $(2.1 < |\eta| \leq 2.4 \text{ } \& \text{ } p_T > 1.5 \text{ GeV})$

It is underlined again that lower  $p_T$  thresholds in the muon acceptance raises significantly the observable  $B_c$  yields. The *soft* muon acceptance  $p_T$  cuts are much lower than the *hybrid-soft* one, which are in turn lower than those of triggering muons.

### 5.3.3 Strategy for three muons

The *soft* muons contain too much fakes (especially in PbPb where there is a high number of charged particles possibly misidentified as muons) for an analysis using only this muon selection to be doable. However, as shows Fig. 5.2, asymmetric muon cuts can be very advantageous to reach  $B_c$  mesons with lower- $p_T$ , which is desired for physics interests (to get closer to the

recombination regime) and for the observable yields (because the mean  $p_T$  of the produced  $B_c$  mesons is lower than that of the softer accepted candidates).

The trigger strategy naturally makes the muon cuts asymmetric: only two muons are required to fire the trigger, and are not required to be the muons from the  $J/\psi$  dimuon. This brings some (manageable) complications in the application of scale factors on the MC efficiency corrections (see section 7.2), but keeps about a quarter more signal. The trigger is more efficient on higher-quality muons, so it is reasonable to require the two triggering muons to be tightly identified, as *hybrid-soft*. Requiring the muons to trigger decreases the efficiency compared to only-identified muons, and the triggering muons hence need the tight acceptance cuts defined in section 5.3.2.

The third muon does not need to trigger, so it has a higher efficiency and a possibly looser acceptance. Considering the large background induced by keeping *soft* muons, the strategy that was finally chosen is to require it to be *hybrid-soft*, giving a reasonably good muon purity; the associated loose acceptance cuts are detailed in section 5.3.2 and are significantly looser than the ones for triggering muons.

However, the low- $p_T$  efficiency is much higher for *soft* muons, so the associated acceptance is much looser (see section 5.3.2). To further check the option of allowing one of the muons to be simply *soft*, the signal efficiency of requiring *hybrid-soft* on the three muons (and the respective acceptance cuts) when the trimuon passes the ‘2 *hybrid-soft* + 1 *soft*’ strategy and the preselection (from section 5.4) was measured. It is only 54% (resp. 63%) in pp (resp. PbPb), but also cuts a large part of the background (see Tables 5.1 and 5.2 of next section). We chose the ‘3 *hybrid-soft*’ selection because of the more manageable background, but also because the scale factors to correct the MC efficiency of *soft* muons are challenging to measure, especially in the PbPb environment. A trial was done (shown in section 7.1.5), and gives about 10% errors on the *soft* muon scale factors, which is considered a show-stopper.

## 5.4 Preselection

A trimuon signal candidate thus consists in three muons identified as *hybrid-soft*, of total charge  $\pm 1$  (though the candidates of charge  $\pm 3$  are kept for illustration in the template fit, as explained in section 4.2), and for which a common vertex can be fitted. As a reminder, only two muons are required to fire the trigger and to pass the tight acceptance (defined in section 5.3.2 along with the loose one), and these are not necessarily from the  $J/\psi$  decay. The third muon must be in the loose acceptance. The trigger must be fired at the event-level as well. Only the 0-90% centrality range is kept in PbPb.

For a trimuon of charge  $\pm 1$ , there are two choices of opposite-sign (OS) dimuons, meaning two  $J/\psi$  candidates. For reasons related to the fake  $J/\psi$  background study (cf section 4.2), the  $J/\psi$  is **not** chosen as the closest to  $m_{J/\psi}^{\text{PDG}}$ . Instead, a trimuon with ambiguous  $J/\psi$  choice (this concerns about 7% of the weighted preselected candidates of the signal region, and 5% of the weighted  $J/\psi$  sidebands candidates) is split in two candidates having weights of sum 1, as explained in section 5.5.

The trimuon candidates undergo a preselection before being inputted in the core analysis steps that follow (BDT, template fit, acceptance and efficiency corrections...). The term ‘pre’-selection is used even if all preselected events are considered in the following analysis steps, because the binning in BDT (see sections 5.7 and 6) separates the samples in background- or signal-enriched regions as a selection does.

The key challenge of the selection is to find variables that have significantly different distributions for the signal and for the backgrounds. Fig. 5.6 sketches the topology of a  $B_c$  trimuon decay and points out some standard variables that make it stand out versus the backgrounds. The variables used in the preselection cuts are:

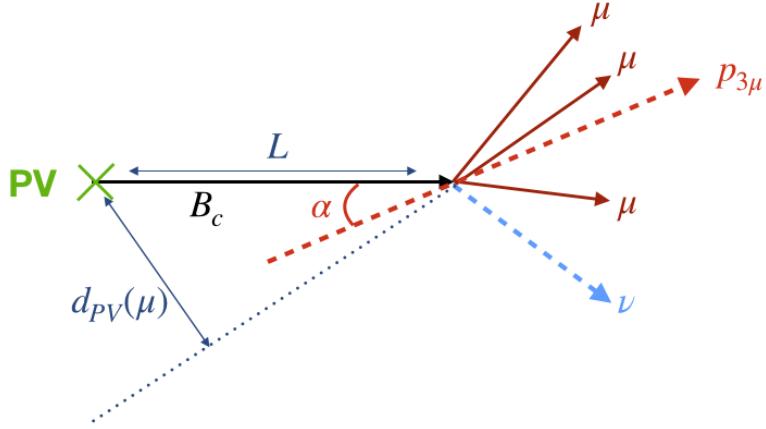


FIGURE 5.6: Sketch of the spatial topology of a  $B_c$  decay to three muons and an unreconstructed neutrino. The flight distance  $L$  is also named  $c\tau$  in the text. The distance of closest approach of a muon to the PV  $d_{PV}(\mu)$  is called  $d_{xy}(\mu)$  and  $d_z(\mu)$  when projected in the transverse plane or on the  $z$  direction. The pointing angle  $\alpha$  is the angle between the  $B_c$  flight direction and the reconstructed trimuon momentum.

- The **probability** (`VtxProb`) of the trimuon **vertex fit**, and of the  $J/\psi$  dimuon vertex fit. A failed fit or a fit with low probability indicates muons not coming from the same vertex.
- The absolute **angle  $\alpha$  between the  $B_c$  flight direction** (*i.e.* the direction of the segment from the PV to the trimuon vertex) **and the trimuon momentum**. This should be very close to 0 for signal when the neutrino takes no momentum and the length of the flight segment is well above the spatial resolution of the detector. When the neutrino takes more energy, it can though sometimes get larger than in the fully reconstructed decays of the non-prompt  $J/\psi$  MC background. The same variable projected on the transverse plane is also used.
- The **trimuon lifetime significance**  $\tau/\sigma_\tau$ , *i.e.* the distance from the PV to the trimuon vertex divided by its uncertainty, both in 3D and projected on the transverse plane (2D). The ‘experimental’ proper lifetime is defined as:

$$\tau = L \cdot \cos \alpha \cdot \frac{m_{B_c}^{\text{PDG}}}{p_{\mu\mu\mu}}$$

where  $\frac{m_{B_c}^{\text{PDG}}}{p_{\mu\mu\mu}}$  corrects the travelled distance  $L$  to the proper lifetime (up to the missing neutrino momentum), and  $\cos \alpha$  drives down this quantity for events where the flight distance of the  $B_c$  is not well resolved and/or when the neutrino takes a part of the energy. The  $m/p$  term cancels in the ratio with the lifetime error. This effective lifetime can be negative when the flight distance is not resolved ( $\cos \alpha < 0$ ).

- The **distance of closest approach** of each **muon to the PV** in the longitudinal direction  $|d_z(\mu)|$  must not be too large, to avoid pile-up (in pp) and cosmic muons. The choice of PV was detailed in section 5.2.
- The **corrected mass**. It can get closer to the PDG  $B_c$  mass than the trimuon mass, by correcting for the momentum of the neutrino transverse to the flight direction of the  $B_c$ . It is defined, from the trimuon momentum transverse to the flight direction  $p_{\perp}^{\mu\mu\mu}$ , as:

$$m_{\text{corr}} = \sqrt{(m^{\mu\mu\mu})^2 + (p_{\perp}^{\mu\mu\mu})^2 + |p_{\perp}^{\mu\mu\mu}|} \quad (5.1)$$

If we look at the decay in the center of mass of the  $B_c$ , then the momentum conservation reads  $\vec{p}_v^* = -\vec{p}_{\mu\mu\mu}^*$ , and the energy conservation gives:

$$m_{B_c} = E_{\mu\mu\mu}^* + E_v^* = \sqrt{(m^{\mu\mu\mu})^2 + \vec{p}_{\mu\mu\mu}^{*2}} + |\vec{p}_{\mu\mu\mu}^*|.$$

As the momentum of the trimuon transverse to the flight direction of the  $B_c$  is not modified by a boost along this flight direction, we can remove from it the mention of center-of-mass, and

$$m_{B_c} = \sqrt{(m^{\mu\mu\mu})^2 + (p_{\perp}^{\mu\mu\mu})^2 + p_{z,v}^{*\perp 2}} + \sqrt{(p_{\perp}^{\mu\mu\mu})^2 + p_{z,v}^{*2}}$$

from which we find the expression of the corrected mass of Eq. 5.1, after neglecting the projection of the momentum of the neutrino along the flight direction of the  $B_c$ . The resolution on the  $B_c$  flight distance additionally smears this corrected mass.

- The **sum of dimuon opening angles**  $\sum_{i \neq j} \Delta R(\mu_i, \mu_j)$ , where  $\Delta R$  is the pseudo-angle in  $(\eta, \phi)$  between two particles (defined in Eq. 1.9), and the sum runs on the three possible muon pairs. For the medium/high  $p_T$  range accessible to this analysis, a signal trimuon is expected to be somewhat collimated.

The  $\Delta R$  is however modified to correct for the difference between the  $J/\psi$  dimuon mass and the  $J/\psi$  PDG mass, as the dimuon mass is roughly proportional to the  $\Delta R(J/\psi)$  variable (the dimuon opening angle is of order  $m_{\mu\mu}/p_{\mu\mu}$  when  $p_{\mu\mu} \gg m_{\mu\mu}$ ). This becomes important in the case of the dimuon mass sidebands, where the  $J/\psi$  dimuon mass and  $\Delta R$  are not representative of those of the actual fake  $J/\psi$  events to be described (that have masses in the  $J/\psi$  peak region). Therefore, the  $\Delta R$  of the  $J/\psi$  dimuon must be modified<sup>1</sup> into  $\Delta R'$  such that:

$$1 - \cos \Delta R' = \left( \frac{m_{J/\psi}^{\text{PDG}}}{m_{\mu\mu}} \right)^2 (1 - \cos \Delta R) \quad (5.2)$$

Then the modifications of the  $\Delta R$  of the other two muon pairs are calculated geometrically: they are assimilated to the two sides  $a$  and  $b$  of a triangle formed by the three muons, the non- $J/\psi$  muon vertex being fixed between the  $a$  and  $b$  sides. The direction of the third side  $c$  (representing  $\Delta R(J/\psi)$ ) is also fixed, but its length changes according to Eq. 5.2. If  $c$  is modified to  $c' = c\varepsilon$  with  $\varepsilon \ll 1$ , then the other two sides are modified as:

$$\begin{aligned} a' &= a \left( 1 + \frac{p_T(\mu_b)}{p_T(\mu_a) + p_T(\mu_b)} \frac{\varepsilon}{2} \left( 1 + \frac{c^2 - b^2}{a^2} \right) \right) \\ b' &= b \left( 1 + \frac{p_T(\mu_a)}{p_T(\mu_a) + p_T(\mu_b)} \frac{\varepsilon}{2} \left( 1 + \frac{c^2 - a^2}{b^2} \right) \right) \end{aligned}$$

where the  $\frac{p_T(\mu_{a/b})}{p_T(\mu_a) + p_T(\mu_b)}$  term indicates that the modification of opening angle is assumed more important when it concerns a  $J/\psi$  muon with smaller  $p_T$ . It is reminded that these corrections are negligible, except in the case of a dimuon from the  $J/\psi$  sidebands.

These variables are plotted along with other discriminant variables for the preselected signal and background samples in section 5.7.2. Tables 5.1 and 5.2 present for each cut in pp and PbPb the cut value, as well as the ‘n-1’ inefficiency, *i.e.* the fraction of trimuon candidates having already passed all the other preselection cuts and that do *not* pass this last cut.

The inefficiencies are presented for signal and for the three background samples separately. In pp these are the nominal samples: the data dimuon sidebands (fake  $J/\psi$ ), the non-prompt  $J/\psi$  MC

<sup>1</sup>To find this expression, let us consider the mass of a dimuon, assuming massless muons of momenta  $\vec{p}_1$  and  $\vec{p}_2$  with an opening angle  $\theta$ : it is  $m^2 = 2|\vec{p}_1||\vec{p}_2|(1 - \cos \theta) \simeq |\vec{p}_1||\vec{p}_2|\theta^2$  (where the approximation  $m \propto |\theta|$  is valid at small angle). Keeping constant muon momenta, this means that a mass modification  $m \rightarrow m'$  implies  $1 - \cos \theta \rightarrow (\frac{m'}{m})^2(1 - \cos \theta)$ . It is also assumed that  $\Delta R$  is a proxy for the euclidean opening angle.

TABLE 5.1: Preselection cuts in pp, and the associated inefficiencies on candidates that have passed all the other cuts. The inefficiencies of signal and of the three nominal background samples (see text for details) are shown. Some ‘n-1’ inefficiencies are also mentioned for pairs of combined cuts. Some inefficiencies show only lower limits, along with the initial cut from which this minimal inefficiency was measured. The last row corresponds to all cuts except the *hybrid-soft* selections. All inefficiencies are in percents.

Variable	cut value	pp inefficiencies [%]			
		signal	fake J/ $\psi$	B $\rightarrow$ J/ $\psi$ + X non-prompt MC	rotated J/ $\psi$
(1) trimuon VtxProb	> 0.008	4.5	(from $prob > 0.005$ ) >5.1	13.1	(from $prob > 0.005$ ) >4.8
(2) J/ $\psi$ VtxProb	> 0.005	0.73	1.2	0.83	(from $prob > 0.002$ ) >0.52
(1) and (2)		7.9	>6.8	16.8	>5.4
(3) $\alpha_{3D}$ [rad]	< 0.3	1.5	(from $\alpha_{3D} < 1.57$ ) >2.1	0.49	(from $\alpha_{3D} < 1.57$ ) >4.2
(4) $\alpha_{2D}$ [rad]	< 0.6	0.50	(from $\alpha_{2D} < 1.37$ ) >5.5	0.16	(from $\alpha_{2D} < 1.37$ ) >4.8
(3) and (4)		2.1	>8.3	0.65	>9.6
(5) $\tau_{3D}/\sigma_{\tau_{3D}}$	> 1.2	0.62	(from $\tau/\sigma > 0$ ) >2.9	0.22	(from $\tau/\sigma > 0$ ) >4.1
(6) $\tau_{2D}/\sigma_{\tau_{2D}}$	> 1.2	3.1	(from $\tau/\sigma > 0$ ) >7.8	0.92	(from $\tau/\sigma > 0$ ) >12.2
(5) and (6)		7.0	>22.0	2.5	>31.8
(7) $max_i(d_z(\mu_i))$	< 6 mm	0.02	0.45	2.4	0.04
(8) $m_{corr}$	< 20 GeV	0.39	1.0	0.16	0.97
(9) $\sum \Delta R(\mu_i, \mu_j)$	< 4.5	1.7	10.0	0.33	8.9
(10) 3 <i>hybrid-soft</i> in loose acc.		46.1	87.3	85.6	79.3
cuts (1) to (9)		<b>32.2</b>	<b>&gt;60.2</b>	<b>29.9</b>	<b>&gt;68.4</b>

events in which the J/ $\psi$  and muon are daughters of the same B meson, and the rotated J/ $\psi$  sample. In PbPb, we look at the dimuon sidebands, and the prompt and (full) non-prompt MC. The last two are not the central methods of the template fit, but are still good estimators of the true-J/ $\psi$  background and separate the effects of variables with respect to their link to lifetime. The three backgrounds are not merged for this study, because the relative background normalisations are precisely known only after the template fit. Only candidates of masses between 3.5 GeV and 6.2 GeV are considered for the efficiency estimations, corresponding to the signal mass range used in the template fit. Some inefficiencies cannot be exactly measured, as some of the preselection cuts were applied during the processing of the data **Oniatree**’s to lower dramatically the output file size and running time; for these, only a lower limit on the inefficiency is given, calculated starting from the mentioned initial cut. The last shown cut (10) corresponds to the efficiency of 3 *hybrid-soft* muons in the loose acceptance, from the 2 *hybrid-soft* + 1 *soft* sample, as discussed in section 5.3.3.

The tables also present the ‘n-1’ inefficiencies of multiple simultaneous cuts. It is worth explaining why this inefficiency is always higher than the sum of inefficiencies of the cuts considered separately. Let us consider the inefficiencies  $1 - \varepsilon_1$ ,  $1 - \varepsilon_2$ ,  $1 - \varepsilon_{1+2}$ , respectively for cuts 1, 2, and the combination of 1 and 2. Considering small inefficiencies, we have for each cut  $i$  that

TABLE 5.2: Preselection cuts in PbPb, and the associated inefficiencies on candidates that have passed all the other cuts. See caption of Table 5.1 for details.

Variable	cut value	PbPb inefficiencies [%]			
		signal	fake J/ $\psi$	non-prompt MC	prompt MC
(1) trimuon VtxProb	> 0.008	4.6	(from $prob > 0.005$ ) >4.4	73.4	34.6
(2) J/ $\psi$ VtxProb	> 0.005	0.60	1.4	0.19	2.1
(1) and (2)		7.6	>6.1	74.7	39.5
(3) $\alpha_{3D}$ [rad]	< 0.3	0.56	(from $\alpha_{3D} < 1.16$ ) >2.3	0.81	2.4
(4) $\alpha_{2D}$ [rad]	< 0.6	0.28	(from $\alpha_{2D} < 1.57$ ) >15.5	2.2	7.2
(3) and (4)		0.85	>18.7	3.2	10.8
(5) $\tau_{3D}/\sigma_{\tau_{3D}}$	> 1.2	0.51	(from $\tau/\sigma > 0$ ) >5.6	0.84	13.8
(6) $\tau_{2D}/\sigma_{\tau_{2D}}$	> 1.2	2.6	(from $\tau/\sigma > 0$ ) >20.0	5.4	22.6
(5) and (6)		8.3	>42.4	13.0	60.4
(7) $max_i(d_z(\mu_i))$	< 6 mm	0.02	0.13	0.14	0.00
(8) $m_{corr}$	< 20 GeV	0.09	0.07	0.06	0.84
(9) $\sum \Delta R(\mu_i, \mu_j)$	< 4.5	0.45	21.4	2.6	8.4
(10) 3 hybrid-soft in loose acc.		37.3	96.3	96.2	95.6
cuts (1) to (9)		<b>27.6</b>	<b>&gt;79.9</b>	<b>81.7</b>	<b>94.4</b>

$1 - \varepsilon_i = \frac{N_{cut,i}}{N_{cut,i} + N_{pass}} \simeq \frac{N_{cut,i}}{N_{pass}}$ , where we defined  $N_{pass}$  as the number of trimuons passing all final cuts, and  $N_{cut,i}$  the number of trimuons recovered by releasing the set of cuts  $i$ . The point is that  $N_{cut,1+2} > N_{cut,1} + N_{cut,2}$  (implying that  $(1 - \varepsilon_{1+2}) > (1 - \varepsilon_1) + (1 - \varepsilon_2)$ ), because there are events that do not pass either of the two cuts, and are hence not recovered in  $N_{cut,1}$  nor in  $N_{cut,2}$ .

The criterium used to decide the cut values is mostly qualitative, looking for tails of the distributions that are sparsely populated by the signal MC, but contains a significant amount of background. Some cuts (*e.g.*  $m_{corr}$  or  $d_z(\mu)$ ) do not have much effect, and are mostly set to eliminate outliers with pathological topology. Some other cuts (*e.g.* on  $\alpha_{3D}$ ) are set quite stringent in the preselection, because they are not included as input variables of the BDT (to eliminate variables that are too correlated, cf section 5.7.2). A majority of the background rejection comes from cuts on the probabilities of the vertex fits, on the angle  $\alpha$ , and on the significance of the flight distance.

Overall, the preselection is satisfying because it cuts less than a third of signal, but cuts at least 60% of pp background, and more than 80% of PbPb background (the numbers for backgrounds are only lower limits due to the `Oniatree`'s having initial cuts, and the imprecise relative background normalisations before the fit). This number excludes in pp the MC describing true  $B \rightarrow J/\psi + X$  candidates, which is not well discriminated against with the preselection cuts. The BDT procedure will perform better, using higher-level topological variables.

After the preselection and the  $J/\psi$  choice weighting of next section, we obtain in pp (in the

signal region  $3.5 < M_{\text{trimuon}} < 6.2 \text{ GeV}$ ) 4883 candidates in data, 1788 in the expected signal (with a MC normalisation from previous measurements), 637 in sidebands, 690 in (true  $B \rightarrow J/\psi + X$ ) non-prompt MC, and 1149 in rotated- $J/\psi$  sample (averaged over rotation angles). Similarly in PbPb, we get about 2655 candidates in data, 194 expected signal candidates (normalising the MC assuming  $R_{\text{PbPb}}(B_c^+) = 1$ ), 1320 in sidebands, 288 in non-prompt MC, 340 in prompt MC, and 1134 in the rotated- $J/\psi$ . These numbers are pre-fit and not exactly representative of the final background normalisations that will be determined through the template fit. In particular, the fit will show that the background normalisations from the true  $B \rightarrow J/\psi + X$  simulations are underestimated.

## 5.5 Choice of the $J/\psi$

Care has been taken in keeping unbiased the mass shape of  $J/\psi$  candidates, by using no mass criterion to select the  $J/\psi$  dimuon. This is important in order to get the right proportions of candidates in the dimuon mass signal or sideband region. In fact, if a criterion on the dimuon mass is used, artificial peaks appear in the *trimuon* mass distributions of candidates having a dimuon in the  $J/\psi$  signal mass region, due to excessive kinematic constraints; and the dimuon mass distribution is dragged towards the  $J/\psi$  peak region.

To avoid this bias, the  $J/\psi$  is *not* chosen as the closest to  $m_{\text{PDG}}(J/\psi)$ . Instead, a trimuon is considered to lie in the  $J/\psi$  signal or sideband region if at least one of its two opposite-sign (OS) dimuons is in this region. In case the two pairs are in the dimuon signal and/or sideband region (which concerns 5 to 7% of the weighted preselected data candidates, depending on the sample), then both trimuon candidates are kept (corresponding to the two  $J/\psi$  choices), and two weights of sum 1 are applied on them. These weights estimate the probability that this pair is a true  $J/\psi$  meson, from the dimuon mass distribution of unambiguous events in data.

The dimuon mass probability distribution  $f(m_{\mu\mu})$  is drawn for preselected trimuons where the  $J/\psi$  choice is unambiguous, meaning that there cannot be two OS pairs (passing the dimuon-specific preselection cuts) in the signal and/or sideband region. Let us now consider an ambiguous candidate with OS dimuons of masses  $m_1$  and  $m_2$ ; this candidate is split in two candidates where the  $J/\psi$  is chosen as the OS pair 1 or 2. These sub-candidates are respectively attributed weights  $\frac{f(m_1)}{f(m_1) + f(m_2)}$  and  $\frac{f(m_2)}{f(m_1) + f(m_2)}$ , which are indeed probabilities that the first or second candidate has chosen the right  $J/\psi$ , according to unambiguous preselected data. If the two candidates were attributed weights of  $\frac{1}{2}$ , the fake  $J/\psi$  candidates in the sidebands would be overestimated, and the true  $J/\psi$  candidates would be underestimated.

The dimuon mass histograms for candidates with an unambiguous  $J/\psi$  are drawn separately for the tight and loose choices of dimuon mass regions (see the ranges defined in section 4.2 depending on the maximal muon pseudo-rapidity), and after fiducial cuts (of section 5.6). First, these two distributions are drawn from the whole preselected data, and used to apply weights on all preselected samples. However, once the BDT is trained (see section 5.7) with this simple  $J/\psi$  choice weighting, one can split data in three bins of the BDT output value as will be done for the template fit (see section 6.1), and these bins will contain different proportions of fake  $J/\psi$ . We hence draw again, for each BDT bin, the (unambiguous) dimuon mass distributions for the loose and tight mass ranges, and recalculate the weights of ambiguous candidates therefrom. Some of these categories have too small statistics to obtain a reasonable mass shape, so:

- If the histogram for a category has less than 100 entries, we take the dimuon mass histogram for the sample integrated on BDT.
- If the latter has less than 70 entries, we attribute weights  $0.8 + i \times 0.05$  and  $0.2 - i \times 0.05$  respectively to the signal region and sideband candidates, where  $i$  is the BDT bin number

(higher  $i$  meaning a bin more pure in signal). This choice is a bit arbitrary but concerns very few events.

This procedure is ran only after the first-step BDT training (and not at the beginning of the second step analysis). The dimuon mass distributions  $f(m_{\mu\mu})$  used for the  $J/\psi$  choice weighting are shown for pp and PbPb in Fig. 5.7, for the three BDT bins and for the tight and loose mass ranges. A unique bin per sideband is used, as they contain low statistics and should show an approximately flat shape.

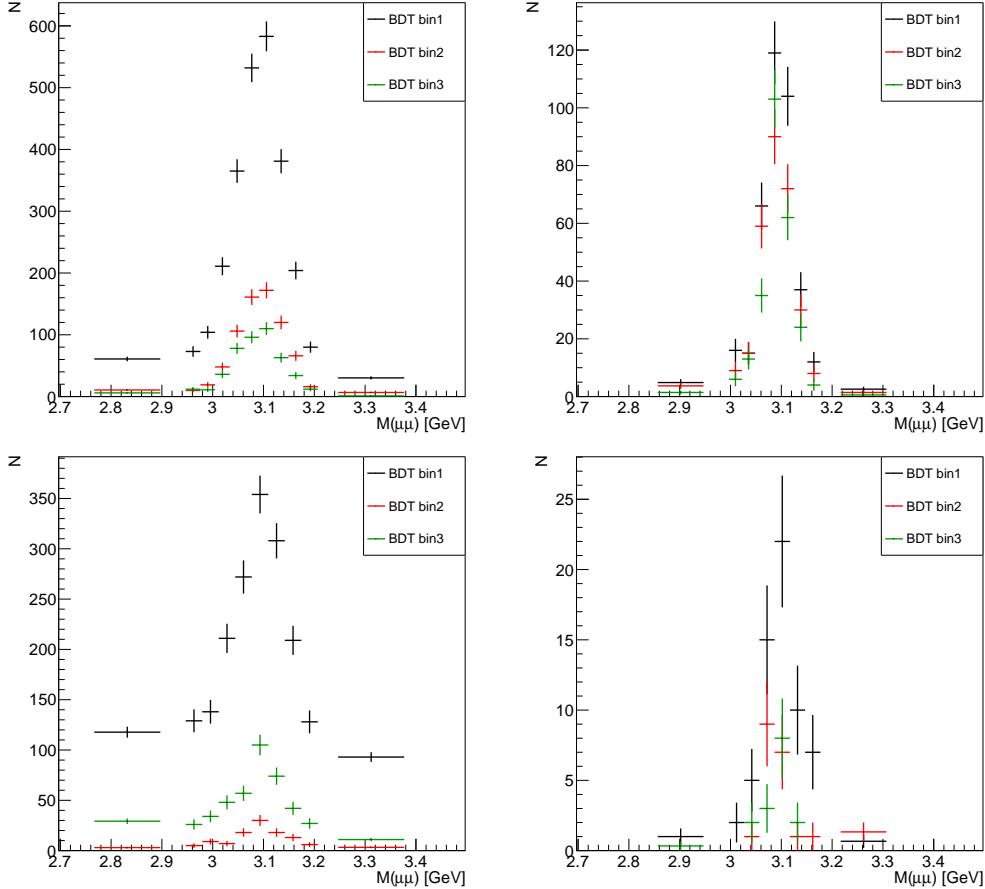


FIGURE 5.7:  $J/\psi$  dimuon mass distributions in the signal and sideband regions in pp (*top*) and PbPb (*bottom*), for the trimuons with unambiguous  $J/\psi$  choice, separated in three BDT bins, for the loose (*left*) and tight (*right*) dimuon mass regions. These distributions are used as probabilities for an ambiguous candidate to have made the right  $J/\psi$  choice.

These weights, always extracted from the unambiguous events in data, are applied to all signal and background samples. It nevertheless has the most importance for the data signal region and sidebands, for which it is pivotal to the proper cancellation of the fake  $J/\psi$  candidates having a dimuon mass in the  $J/\psi$  peak region. It is also applied to the signal MC and other background samples, even if we know in these cases which dimuon pair is the true  $J/\psi$  (either by matching it to the generated  $J/\psi$ , or because we know which dimuon we chose to rotate). This is to account for the effect of the weighting on these samples as well, as they similarly contain about 5-10% of candidates with an ambiguous  $J/\psi$  choice, in which the second OS pair is accidentally in the dimuon mass signal or sidebands regions. It is underlined here that the contamination from the signal MC to the  $J/\psi$  sidebands is taken into account, as for the rotated- $J/\psi$  sample, with negative weights for true  $J/\psi$  mesons having their mass in the sidebands.

For data, the signal MC, and the rotated-J/ $\psi$  sample, the dimuon mass shape of corrected (ambiguous) candidates, was checked against the one of unambiguous candidates. The shapes (not shown here for brevity) are found to be rather similar within the large uncertainties of the scarce ambiguous events. The potential discrepancy for the data shapes probably comes from the fact that ambiguous candidates undergo an additional constraint on their kinematics, which could bias their mass shape – but this would be due to singling out those events and not to the above weighting.

## 5.6 Analysis binning

Now that the preselection is set, we can decide (using the signal MC) in which phase space the cross section and modification factor results will be quoted (often called *fiducial cuts*). Within these fiducial cuts is also chosen the analysis binning. The observable trimuon kinematic variables are used for these cuts and binning, because the precision of this analysis is not sufficient for an evaluation of the true  $B_c$  kinematic variables (see section 8.3) to be relevant. The choice of the studied kinematic region and binning is based on the following criteria:

- There must be some expected signal events in all corners of the chosen phase space (one cannot quote a measurement on a region where no observed event is expected).
- In the case of event-by-event acceptance and efficiency corrections (which are *not* taken as the nominal strategy, see section 7.5), the differences between the corrections for each of the candidates must not be too large between parts of the same analysis bin, otherwise very different acceptance and efficiency weights are applied. This decreases the statistical significance of the yields corrected with this method, by relying on a small part of events with very dominant weights. However in practice, the achievable binning results in large kinematic regions, where the corrections vary a lot; this lead to dropping this method.
- It is important both for physics interests and for the expected significance of the observation to try lowering the trimuon  $p_T^{\mu\mu\mu}$  threshold of the measured phase space. The  $B_c$  cross section is sharply rising around the low- $p_T$  trimuon acceptance threshold, and the low- $p_T$  region is particularly interesting because different effects than at high- $p_T$  could appear, such as recombination.
- Since the modifications due to the presence of QGP are expected to depend more on  $p_T$  (due to different expected phenomena) than on rapidity, it is preferred to optimise the slicing in  $p_T^{\mu\mu\mu}$ , to assess this dependence. The rapidity cuts that will be associated to the  $p_T^{\mu\mu\mu}$  bins should not prevent this binning from assessing a  $p_T^{\mu\mu\mu}$  dependence, because the rapidity spectrum is expected to be much flatter than the  $p_T$  spectrum.

The map of the number of simulated signal trimuons as a function of the visible  $p_T^{\mu\mu\mu}$  and absolute rapidity  $|y^{\mu\mu\mu}|$  is shown in pp and PbPb in Fig. 5.8, and the map of the trimuon acceptance times efficiency in Fig. 5.9. The  $p_T^{\mu\mu\mu}$  distribution of the used signal MC is here weighted with the final  $p_T^{\mu\mu\mu}$ -spectrum correction, meaning at the end of the second-step analysis (just before the final acceptance and efficiency computation), as is explained in section 7.6; maps without any  $p_T^{\mu\mu\mu}$  weighting (that do not contain this preliminary form of the measured yields) however lead to the same conclusions. The binning of these maps is tuned to get a similar number of expected candidates in each bin. Details on acceptance and efficiency maps are given in section 7.4.

According to the criteria above, we define rectangular regions in trimuon  $p_T$  and rapidity over which the full acceptance times efficiency is always higher than  $2 \times 10^{-4}$ , and in which a  $p_T^{\mu\mu\mu}$  binning is relatively simple. Keeping in mind the low statistics, it is not realistic to evaluate the cross sections in more than two bins. The two chosen analysis bins correspond to different rapidity cuts to take advantage of the shape of the acceptance. They are defined as:

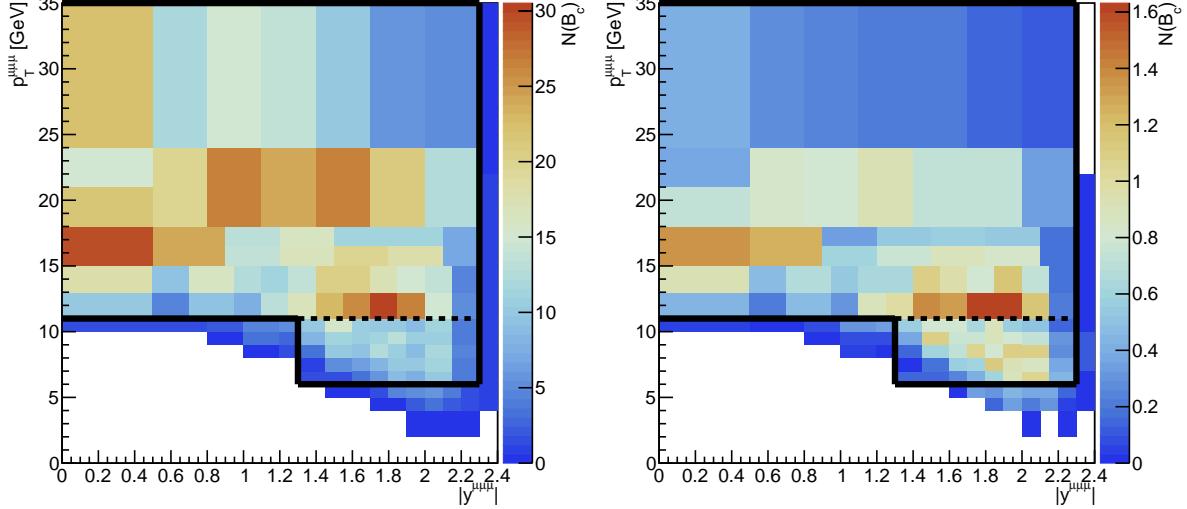


FIGURE 5.8: Trimuon ( $p_T^{\mu\mu\mu}, y^{\mu\mu\mu}$ ) map of the number of simulated signal candidates (using an MC corrected with the final measured  $p_T^{\mu\mu\mu}$ -spectrum, see section 7.6), in pp (left) and PbPb (right). The black lines show the regions chosen as the two  $p_T$  analysis bins.

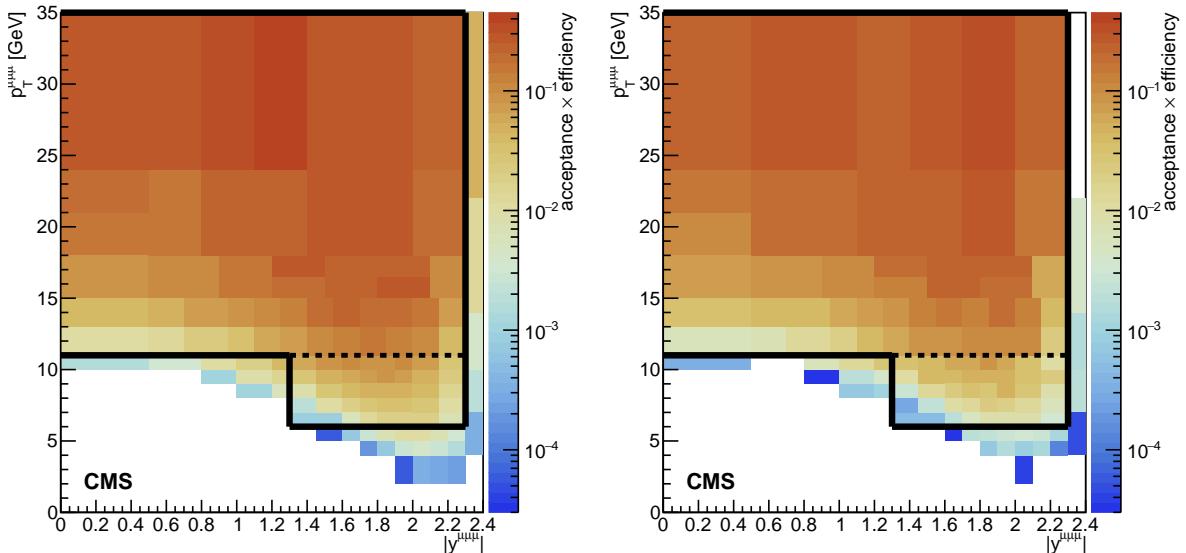


FIGURE 5.9: ( $p_T^{\mu\mu\mu}, y^{\mu\mu\mu}$ ) map of the trimuon acceptance times efficiency (from the signal MC corrected to the final measured  $p_T^{\mu\mu\mu}$ -spectrum, see section 7.6), in pp (left) and PbPb (right). The black lines show the regions chosen as the two  $p_T$  analysis bins.

- $6 < p_T^{\mu\mu\mu} < 11 \text{ GeV}$  for  $1.3 < |y^{\mu\mu\mu}| < 2.3$
- $11 < p_T^{\mu\mu\mu} < 35 \text{ GeV}$  for  $0 < |y^{\mu\mu\mu}| < 2.3$

The signal yields expected within  $m^{\mu\mu\mu} \in [3.5, 6.2] \text{ GeV}$ , with the pre-fit normalisation (implying  $R_{\text{PbPb}} = 1$ ) and no  $p_T^{\mu\mu\mu}$  weighting, and the J/ $\psi$ -choice weighting, are respectively 50 and 137 in the first and second  $p_T^{\mu\mu\mu}$  bin in PbPb, and respectively 594 and 1058 in the first and second  $p_T$  bin in pp. Increasing the upper limit of the second  $p_T$  bin to 50 GeV only increases the signal yields in this bin by about 1%, which will have a negligible effect on the result, so the upper limit of 35 GeV to lower the size of this bin.

It is to be reminded here that the centrality is restricted to 0-90% in this analysis (justified

in section 5.2). An important physics message can also be carried by a binning in centrality. The ranges 0-20% and 20-90% are natural regarding their average  $N_{\text{coll}}$  (1271 in 0-20% and 183 in 20-90% [35]) times range width (0.2 and 0.7), according to which the yields of the studied hard process should scale. This amounts to 254 and 128 equivalent  $N_{\text{coll}}$  in the 0-20% and 20-90% ranges, but this difference will be somewhat compensated by a lower efficiency and higher background in the most central events. It would be ideal to measure the  $p_{\text{T}}$  dependence in each centrality bin, but the statistics do not allow a measurement with reasonable precision, so the two centrality bins are integrated over the  $p_{\text{T}}^{\mu\mu\mu}$  (and rapidity) bins spelt out above. The 0-20% centrality range will be called ‘first centrality bin’, by convention.

Many analysis procedures are performed separately both for the two centrality and the two  $p_{\text{T}}$  bins, and some are performed separated only in  $p_{\text{T}}^{\mu\mu\mu}$  bins, such as the BDT training and the correction of the  $p_{\text{T}}$  spectrum from the first-step measurement. An overview of which analysis modules ones are performed for which binning is included in section 3.3.3 and Fig. 3.13.

## 5.7 Boosted Decision Tree (BDT)

### 5.7.1 Principle

A *Boosted Decision Tree* is trained with the ROOT TMVA module [194] from the preselected signal and background samples, to build a unique discriminant variable (called BDT or BDT output variable in the following) encompassing all the discriminant power of the provided input variables. The template fit (chapter 6) will be performed in three bins of the BDT variable, to benefit from background-enriched bins that stabilise the background parameters, and signal-enriched bins that will drive the determination of the signal yield.

The *training* algorithm starts from decision trees. The first step is to look for the most discriminating input variable, on which the best cut is found according to some signal/background separation criterion. This criterion is to minimise the *Gini* index, equal to  $p(1 - p)$  where  $p$  is the purity (*i.e.* the fraction of signal in a sample); this index reaches 0 for a sample containing only signal or only background. The sample is separated via this cut into a signal-enriched and a background-enriched subsample (*a.k.a* a *leaf*). Then this procedure is repeated in each subsample, and iterated until some stopping criterion is met. I just described one decision tree, but *boosting* means that many such trees are computed, reweighting the events before each new tree computation to give more importance to the events that were classified in a wrong end-leaf in the previous trees. The trained BDT is a function that takes values for the input variables of a given event, and outputs a value that represents how probable it is for this event to be classified as signal by a majority of these trees. To summarise with a simplified view of a BDT trained in  $n$  variables, it draws rectangular  $n$ -dimensional regions (the leaves) that are tagged with a high fraction of signal or of background, and those fractions schematically give the probability of being a signal event for any point in this space.

A BDT performs better if the input variables already discriminate well, so I build some higher-level variables to that end in section 5.7.2, that will be inputted along some variables used in the preselection; and these variables are plotted for preselected samples.

The training maximises the rejection of the background while minimizing the rejection of the inputted signal. With given input samples, the training can in principle get sophisticated enough to discriminate perfectly the signal and background; however this would exactly fit the statistical fluctuations and systematic uncertainties of the input samples, and would not perform as well (and often worse than less optimised BDT’s) on data it was not trained on. This is called *overtraining*. Section 5.7.3 details the choice of the input signal and background samples, and of the training parameters and other ways to deal with the trade-off between performance and overtraining.

The BDT is actually trained twice, where in the second training, the input signal sample benefits from the weighting of its  $p_T^{\mu\mu\mu}$  distribution using the first-step cross-section measurement (see section 7.6). The final template fit (chapter 6) uses the BDT variable from the second-step training. The training is performed separately on the two  $p_T^{\mu\mu\mu}$  bins (even for the centrality dependence, so that the BDT value of an event does not depend on which binning is considered). Section 5.7.4 shows a comparison of the distributions of the second-step BDT for data versus the sum of signal and the various backgrounds with fitted parameters. A weighting is applied to correct the BDT distribution, only in pp, before running the final template fit.

Section 5.7.5 calculates a modified BDT variable that is uncorrelated with the trimuon mass, which leaves to the mass its full discriminant power in the template fit procedure. This variable is used for the BDT binning in an alternate fit, which participates in the estimation of the systematic uncertainty on the fit procedure (section 6.4).

### 5.7.2 Discriminant variables after preselection

We will use these variables from the preselection (defined in section 5.4) in the training input:

- The corrected mass  $m_{\text{corr}}$ ;
- The logarithm of the trimuon lifetime significance;
- The logarithm of the trimuon vertex probability;
- The projection of the angle  $\alpha$  (between the flight segment and the trimuon momentum) on the transverse plane. It is less correlated with the lifetime significance than its 3D version, and hence provides a slightly better performance increase;
- The sum of the three pseudo-angles between the muon pairs  $\sum_{i \neq j} \Delta R(\mu_i \mu_j)$ .

The cuts tried by the training algorithm on the input variables can only be discrete, and the number of cuts equally distributed on the full variable range cannot be set too high, because this favours overtraining. In case of variables sharply peaking, the details of the signal/background separation could be hidden in the range containing most of the events, but where only a few cuts are tried. To avoid this phenomenon, we smoothen with a logarithm the variables showing most of their signal/background differences in a small part of their range (the lifetime significance and the vertex probability). This is seen to improve the performance.

In addition to these variables, we input these higher-level discriminant variables:

- The imbalance between the transverse momenta of  $J/\psi$  and of the third muon, named  $\mu_W$ :

$$\text{Imbal}_{J/\psi - \mu_W} = \left| \frac{p_T(J/\psi) - p_T(\mu_W)}{p_T(J/\psi) + p_T(\mu_W)} \right|.$$

In a background event with a true  $J/\psi$  meson, the  $J/\psi$  will have much higher  $p_T$  than *e.g.* a random track it combines with, whereas the  $p_T$  of two true daughters of a  $B_c$  meson will be more balanced;

- The ratio of  $\Delta R(J/\psi)$  to the  $\Delta R$  of the other two muon pairs:

$$\frac{\Delta R(\mu_1 \mu_2)}{\Delta R(\mu_2 \mu_W) + \Delta R(\mu_1 \mu_W)}.$$

At relatively low  $p_T$ , the mass of the  $J/\psi$  should result in a larger  $\Delta R$  for the  $\mu_1 \mu_2$  dimuon than for the other dimuons. This quantity hence tends to be close to 1 for signal (this value corresponding to the  $\mu_W$  having its direction exactly in-between the two  $J/\psi$  muons), and has low values when the  $J/\psi$  muons are collimated and the third muon is at a large angle (typically for background events containing a true  $J/\psi$  meson);

- The 3D significance of the displacement to the primary vertex for the non-J/ $\psi$  muon  $d_{xyz}(\mu_W)/\sigma_{d_{xyz}}$ . This limits the combinations of a true J/ $\psi$  with a random track from the PV. Considering uncorrelated transverse and longitudinal measurements and simple error propagation, the 3D displacement significance of a muon can be obtained from the significances of the (more standard) transverse and longitudinal displacements via:

$$d_{xyz}/\sigma_{d_{xyz}} = \frac{d_{xy}^2 + d_z^2}{\sqrt{d_{xy}^2 \sigma_{d_{xy}}^2 + d_z^2 \sigma_{d_z}^2}}$$

To limit overtraining, we checked the correlations between input variables (very correlated variables can bring ambiguity in the determination of the cuts chosen by the various trees). We removed too correlated variables, reducing the training input to the above 8 variables, whose correlation matrices (including the correlations with the mass, that are a motivation for the decorrelation performed in section 5.7.5) are plotted for signal and total background in Figs. 5.10 for pp and PbPb. The additional variables were included in a test version of the BDT, and were then progressively removed, by looking for the most correlated variables (typically with a correlation factor  $\gtrsim 0.5$ ) whose removal did not degrade significantly the performance. The following variables were removed from the training input but are still relatively good discriminants (the last two are less so, but mentioned for completeness):

- The 3D pointing angle  $\alpha_{3D}$  between the flight segment and the trimuon momentum;
- $\sum_i d_{xy}(\mu_i)/\sigma_{d_{xy,i}}$ : the sum of the significances of the transverse displacements of each muon to the primary vertex;
- $\sum_i d_z(\mu_i)/\sigma_{d_{z,i}}$ : idem, but for the longitudinal displacements;
- The minimum instead of the sum for the two above variables, and the 3D displacement instead of the transverse or longitudinal one.
- The distance of closest approach of the J/ $\psi$  dimuon.

Flat trees (one row per trimuon candidate) were built from the `Oniatree`'s, containing all these variables as well as other variables useful for various studies, for events passing the preselection of section 5.4. The eight BDT discriminant variables mentioned above are plotted (normalised to 1) for signal MC and for the relevant background samples in Figs. 5.11 and 5.12, together with the BDT itself and the pointing angle  $\alpha_{3D}$ . The other mentioned variables are plotted in Figs. 5.13 and 5.14, along with the  $\Delta R$  of any muon pair in trimuon candidates. The latter variable confirms that the muons of a trimuon candidate are separated enough to be reconstructed mostly independently in the detector (meaning that in general, they do not share hits). The ROC curves, illustrating the discriminant power of the variables, are shown along with the distributions. They represent the background rejection ( $1 - \varepsilon_B$ ) versus the signal efficiency, for all possible cuts; they stick to the borders of the canvas (*i.e.* their integral is 1) for perfectly discriminant variables, and are separated from the diagonal for cuts that are somewhat discriminant. Some backgrounds have a ROC curve below the reference diagonal indicating equal signal and background distributions; this means that the inverse cut (*e.g.* cutting from below rather than from above) can be picked as discriminant by the BDT training. The ROC curve of the BDT variable is satisfying. In all these plots, the second-step  $p_T^{\text{muon}}$  weighting of the signal MC is applied, and the BDT is from the second-step training. The fiducial cuts in trimuon  $p_T$  and rapidity from section 5.6 are also applied. The used non-prompt J/ $\psi$  MC sample is inclusive (not only the candidates with J/ $\psi$  and muon coming from the same B decay). The integral of the ROC curve is also printed for each background and each variable.

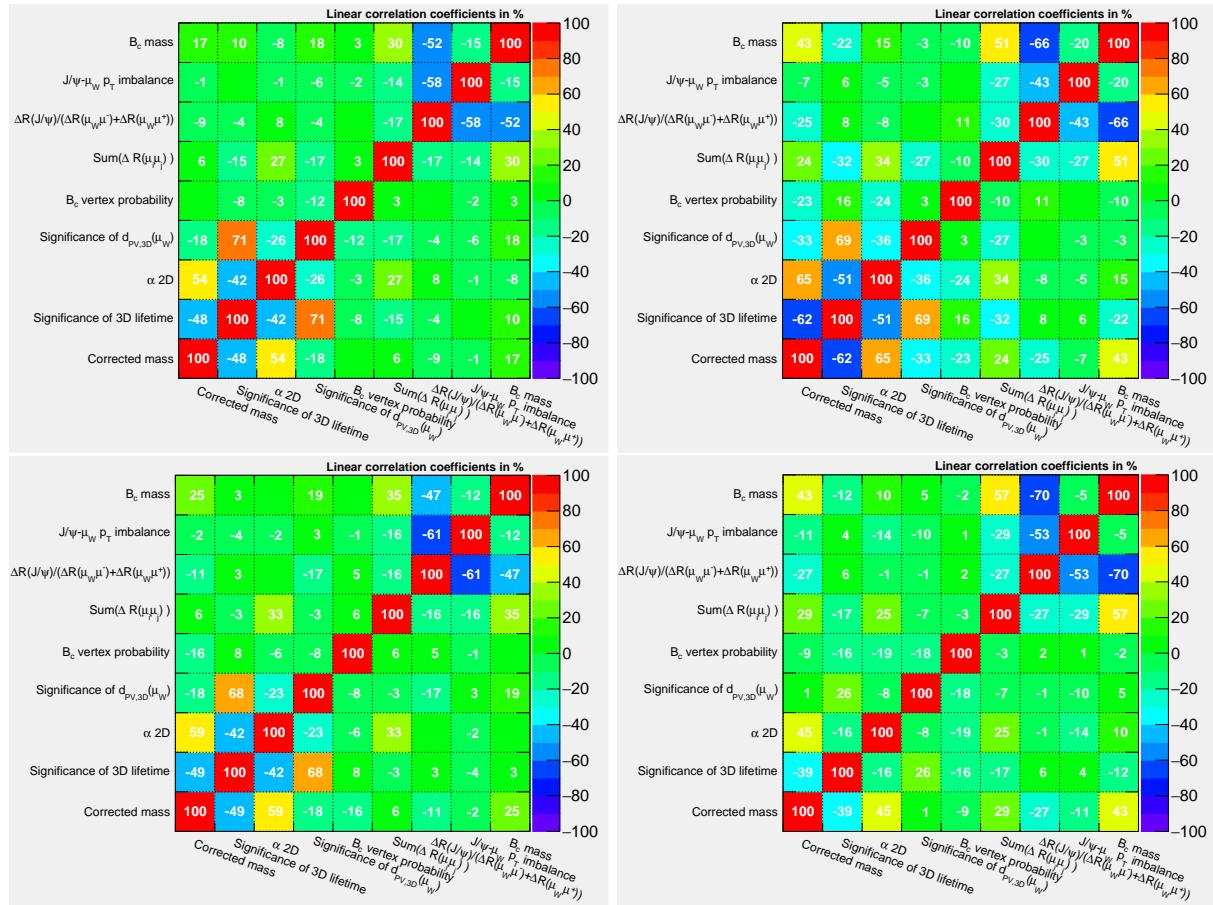


FIGURE 5.10: Correlation matrices of the input variables of the second-step BDT training and the trimuon mass, for the signal (*left*) and background (*right*) training samples in pp (*top*) and PbPb (*bottom*). See section 5.7.3 for the (pre-fit) composition of the background sample. The trimuon mass is not a training variable and only included for illustration.

### 5.7.3 BDT training and testing

A priori normalisations of the signal and background samples (from section 3.5.4) are needed for the BDT training. These normalisations are not exact because they are not fitted yet. Only in the second-step training, the signal MC is corrected to have the  $p_T^{\mu\mu\mu}$  spectrum of the signal measured in the first-step (see section 7.6). The imperfect respective signal and background normalisations mean that the training optimisation might not be perfect. But it does not mean that the output variable is *wrong* – it is only a discriminating tool. We however applied rough global factors to the input background samples, based on preliminary fits.

The samples used as input for the BDT are the simulated  $B_c$  signal, as well as the following background samples, after preselection:

- The  $J/\psi$  sidebands.
  - The wrong-sign sample, with candidates of charge  $\pm 3$ . This sample describes purely combinatorial background, which should be present in the sidebands sample, but it is still used as a complement.
  - The rotated- $J/\psi$  sample, averaged over the thirteen rotation angles, with a weight of 0.7. We give it a lower normalisation than the post-fit one, to leave room to the candidates of the  $J/\psi$  MC featuring a  $J/\psi$  and a muon from different processes.

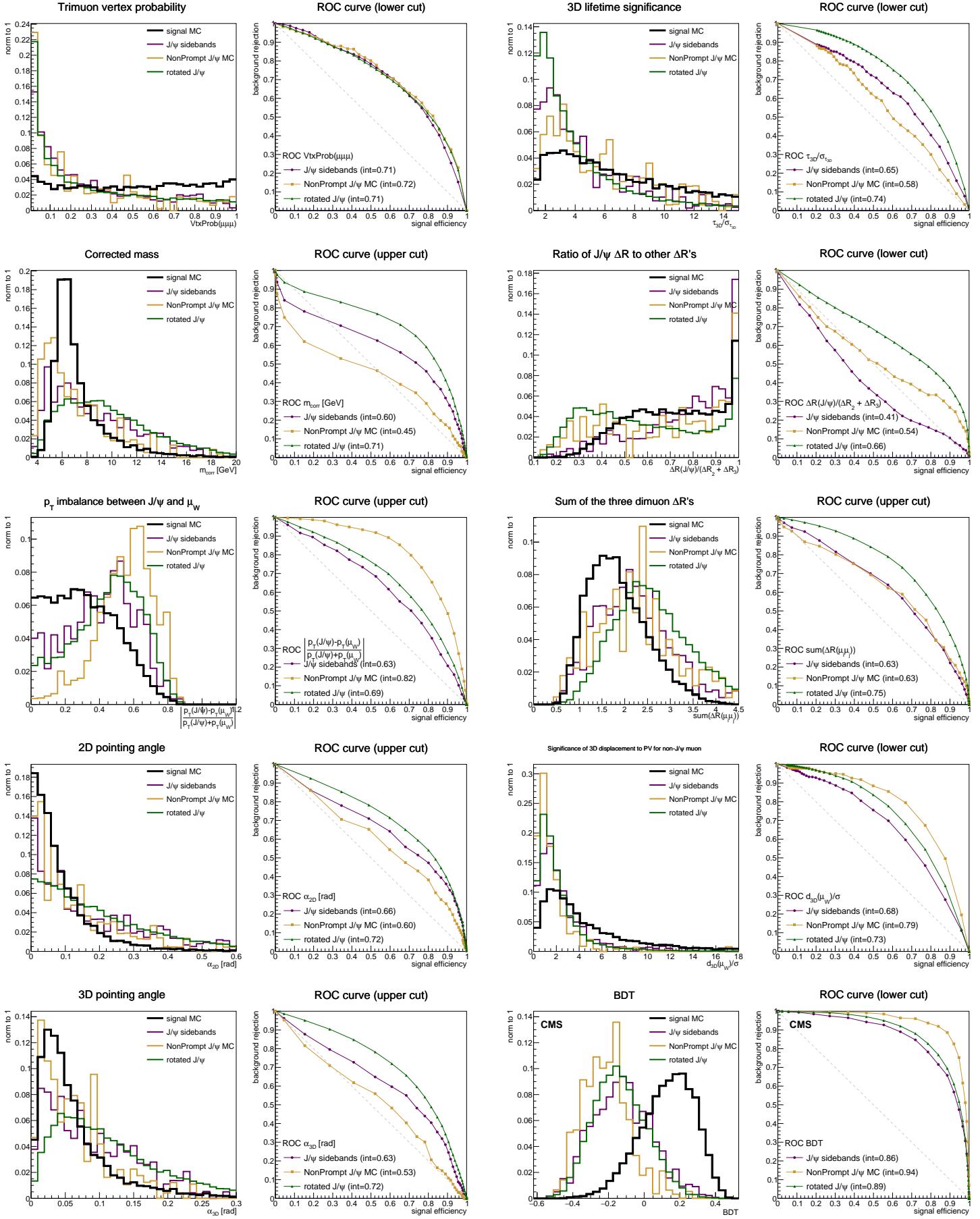


FIGURE 5.11: Distributions of discriminant variables for preselected signal and background samples in pp, along with their ROC curve, for lower or upper cuts. The top 8 variables are used for training the bottom-right BDT.

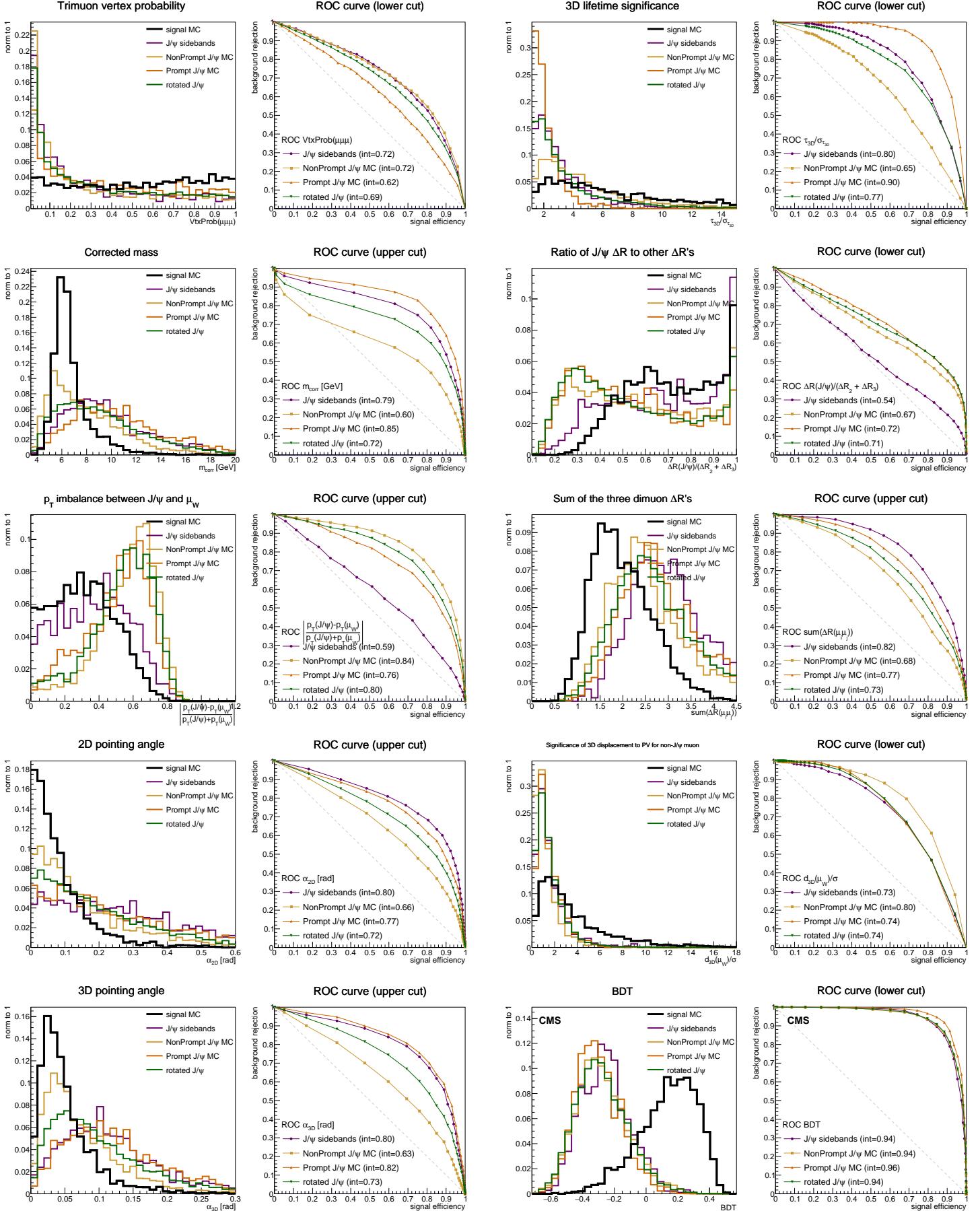


FIGURE 5.12: Distributions of discriminant variables for preselected signal and background samples in PbPb, along with their ROC curve, for lower or upper cuts. The top 8 variables are used for training the bottom-right BDT.

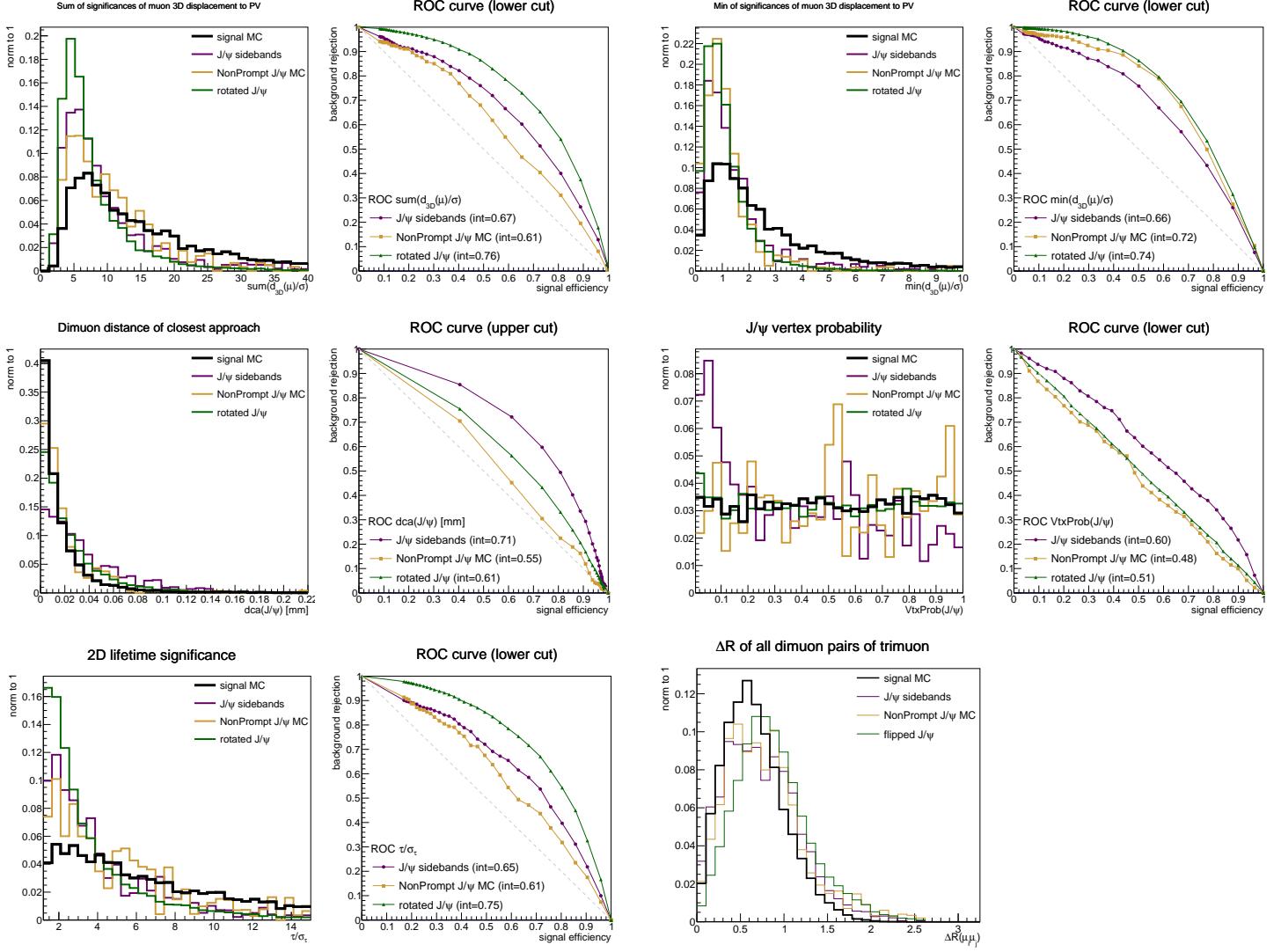


FIGURE 5.13: Distributions of additional discriminant variables for preselected signal and background samples in pp, along with their ROC curve, for lower or upper cuts. The  $\Delta R$  of all dimuons of trimuon candidates is also shown.

- The non-prompt J/ $\psi$  MC, with a weight 2.6 in PbPb (1.75 in pp) for the true  $B \rightarrow J/\psi + X$  decays, and 0.6 (which is the 30% we left room for in the rotated-J/ $\psi$  normalisation, doubled because of its underestimated normalisation) for the other (less correlated) events.
- The prompt J/ $\psi$  MC, with a weight 0.6 for similar reasons as the uncorrelated non-prompt MC.

The global normalisation parameters of the J/ $\psi$  MC will get values between 1.7 and 2.7 in the template fits, which justifies *a posteriori* the weights given to those input samples. There can however be some imprecise counting in the sum of above backgrounds. However, as explained before, this only means that the optimisation will not be as good as it could be, but cannot bias the results. We keep these partially redundant backgrounds to be sure to cover all possible background properties: if some background events are not well described in a given sample, the BDT can train as well against the events of another sample, that might better describe those events.

After the training (which builds the BDT variable, concentrating the discriminant power of all input variables), the BDT needs to be tested for overtraining, and then evaluated (*i.e.* the

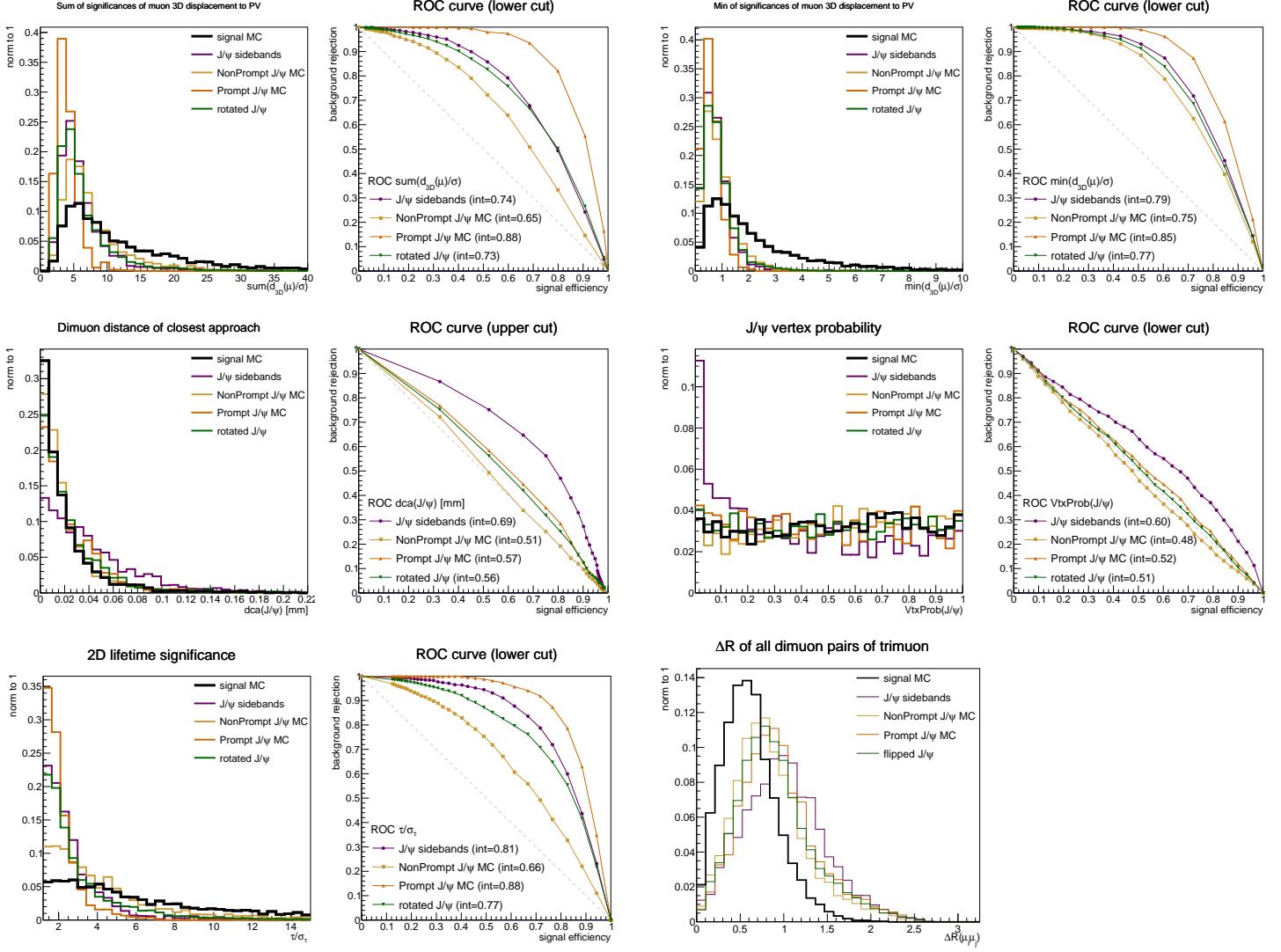


FIGURE 5.14: Distributions of additional discriminant variables for preselected signal and background samples in PbPb, along with their ROC curve, for lower or upper cuts. The  $\Delta R$  of all dimuons of trimuon candidates is also shown.

output BDT variable is calculated in all samples, before drawing the mass distributions used in the template fit). The input samples listed above are used for these three tasks. However, evaluating the BDT on the same sample that it has been trained on leads to bad consequences even for a light overtraining. The first stage of overtraining means that some property inferred by the training is due to a statistical fluctuation of the training sample, and not to an authentic feature of the true distributions. When this slightly overtrained BDT is applied to the training sample, it will reject it more easily than it will reject the actual background from data, so that the training sample cannot give a correct background template to compare to data. To circumvent this, all samples are randomly separated in two, and two separate BDTs are trained in each  $p_T^{\mu\mu\mu}$  bin: one uses the first half as training sample and the second half as the testing and evaluation sample, and vice-versa for the second half.

The parameters of the BDT training in TMVA are:

- Number of trees: 600.
- Maximum depth of the trees: 2. This is a stopping criterion of the training.

- Minimal size of an end leaf (tagged as signal- or background-enriched): 7% of the total integral of the input samples. This is a stopping criterion too.
- Signal separation criterion: Gini index  $p(1 - p)$ .
- Number of cut values tried for each variable: 30 in pp, 15 in PbPb.
- The Boosting uses AdaBoost (*i.e.* adaptative boosting, known to be performant with shallow trees [194]) with a slow-learning rate  $\beta = 0.2$  (limiting the overtraining).

The integral of the BDT output ROC curve and the signal efficiency at 99% background rejection are two quantities for which we optimised the above parameters. While optimizing, we also need to keep the second quantity similar in the training and testing samples, so that there is no important difference in the BDT performance when evaluating it on the statistically independent training and testing samples. Those quantities are given for the two  $p_T^{\mu\mu\mu}$  bins in pp and PbPb in Table 5.3, for the second-step BDT training. There are signs of light overtraining here, with the signal efficiencies at fixed background rejection for training and testing samples being somewhat different, but these differences are of the same order than the variability (due to low statistics) of this efficiency between the trainings of the first and second half of the samples. In addition, as different samples are used for training and testing/evaluating, this small overtraining has negligible consequences. Too tight training parameters could however decrease the performance (which is a second – worse – stage of overtraining); but we checked that releasing them indeed makes the performance decrease. The ROC integral is between 0.84 and 0.87 in pp and between 0.90 and 0.94 in PbPb, depending on the  $p_T^{\mu\mu\mu}$  bin and which random half is considered. The BDT distributions are shown in Figs. 5.11 and 5.12 (*bottom right*).

TABLE 5.3: Second-step BDT performance and overtraining tests in pp and PbPb. The integral of the ROC curve in the test sample, and the signal efficiency at 90% and 99% background rejections in the test and training samples, are shown for the two sample splittings (train on the first half and test on the second half, or vice-versa), along with the average of values over the two splittings.

		ROC integral test sample	$\varepsilon_{sig}$ at $\varepsilon_{bkg} = 0.01$ test (train) sample	$\varepsilon_{sig}$ at $\varepsilon_{bkg} = 0.1$ test (train) sample
pp $6 < p_T < 11 \text{ GeV}$	train half A, test half B	0.845	0.15 (0.23)	0.55 (0.55)
	train B, test A	0.835	0.17 (0.21)	0.54 (0.59)
	average	<b>0.840</b>	<b>0.16 (0.22)</b>	<b>0.55 (0.57)</b>
pp $11 < p_T < 35 \text{ GeV}$	train half A, test half B	0.863	0.12 (0.25)	0.54 (0.66)
	train B, test A	0.872	0.20 (0.17)	0.59 (0.56)
	average	<b>0.867</b>	<b>0.16 (0.21)</b>	<b>0.57 (0.61)</b>
PbPb $6 < p_T < 11 \text{ GeV}$	train half A, test half B	0.907	0.36 (0.47)	0.76 (0.77)
	train B, test A	0.901	0.41 (0.43)	0.77 (0.77)
	average	<b>0.904</b>	<b>0.38 (0.45)</b>	<b>0.76 (0.77)</b>
PbPb $11 < p_T < 35 \text{ GeV}$	train half A, test half B	0.937	0.40 (0.37)	0.82 (0.82)
	train B, test A	0.931	0.24 (0.49)	0.78 (0.84)
	average	<b>0.934</b>	<b>0.32 (0.43)</b>	<b>0.80 (0.83)</b>

The BDT training, testing, and evaluation are done separately for each  $p_T^{\mu\mu\mu}$  bin. As mentioned in section 5.5, after the first-step BDT is evaluated, we run the weighting for the J/ $\psi$  choice a second time, taking into account the BDT binning used in the template fit. The J/ $\psi$  choice weighting is *not* repeated in the second-step analysis.

### 5.7.4 Check and correction of BDT distribution

The BDT distribution might be wrong in the signal MC and in the various background samples. This could affect significantly the performance and reliability of the fit, by migrating events among the three BDT bins. One would like to compare, and correct if necessary, the BDT distribution of the sum of the signal and all backgrounds, to the distribution of data. This is done only with the second-step BDT variable, used in the final template fit.

Right after the second BDT training, I run a preliminary second-step template fit, to extract the post-fit normalisations of signal and of the various backgrounds, and apply them to the samples plotted here. The shape morphing parameters (see section 6.2) are also used, to account for the contributions of the background shape variations in the BDT distributions of backgrounds. Avoiding to use the signal and background normalisations from a previous fit was tried, by getting these weights from the high-mass control region. However, in addition to the low statistics of this region, these weights are not necessarily the same than in the signal region.

Figs. 5.15 and 5.16 present the comparison of the second-step BDT distribution of data versus the sum of post-fit templates, in pp and PbPb, separately in the two  $p_T^{\mu\mu\mu}$  bins, and in two centrality bins for PbPb.

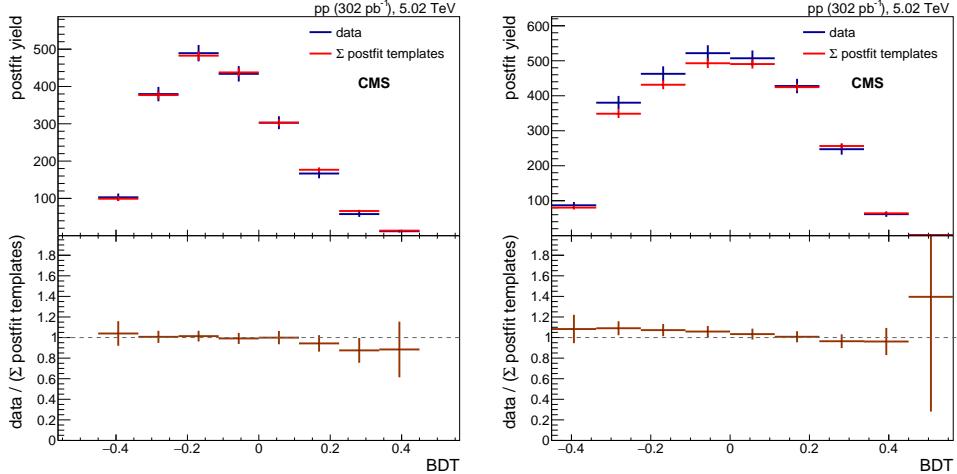


FIGURE 5.15: Comparison (*top*) and ratio (*bottom*) of the BDT distributions of data and of the sum of all post-fit templates (signal MC and the three backgrounds), for the first (*left*) and second (*right*)  $p_T^{\mu\mu\mu}$  bins, in pp. The ratio of distributions is used to weight the samples entering the final fit.

Within the low statistics of PbPb data, no correction seems necessary for the BDT distribution in PbPb – it is checked in bins of  $p_T^{\mu\mu\mu}$  or centrality, and on the integrated sample. Only the pp samples deserve a correction of the BDT distribution. This weighting is calculated, and applied, separately in each  $p_T^{\mu\mu\mu}$  bin. The same weights are applied to all signal and background samples (meaning only the total BDT distribution is corrected). After this weighting, the final second-step fit is performed.

### 5.7.5 Decorrelate BDT from mass

Despite our efforts to reject these, the BDT uses variables that can be significantly correlated to the trimuon mass, as shown in the correlation matrices of Fig. 5.10. During the optimisation, the training can hence partially ‘guess’ the typical mass of the signal events, and make the output discriminant variable dependent on the mass. This biases the trimuon mass shapes towards central masses (where most signal lies) for the highest BDT bins. The fit strongly exploits the fact that the signal and background mass shapes have different features to extract the signal

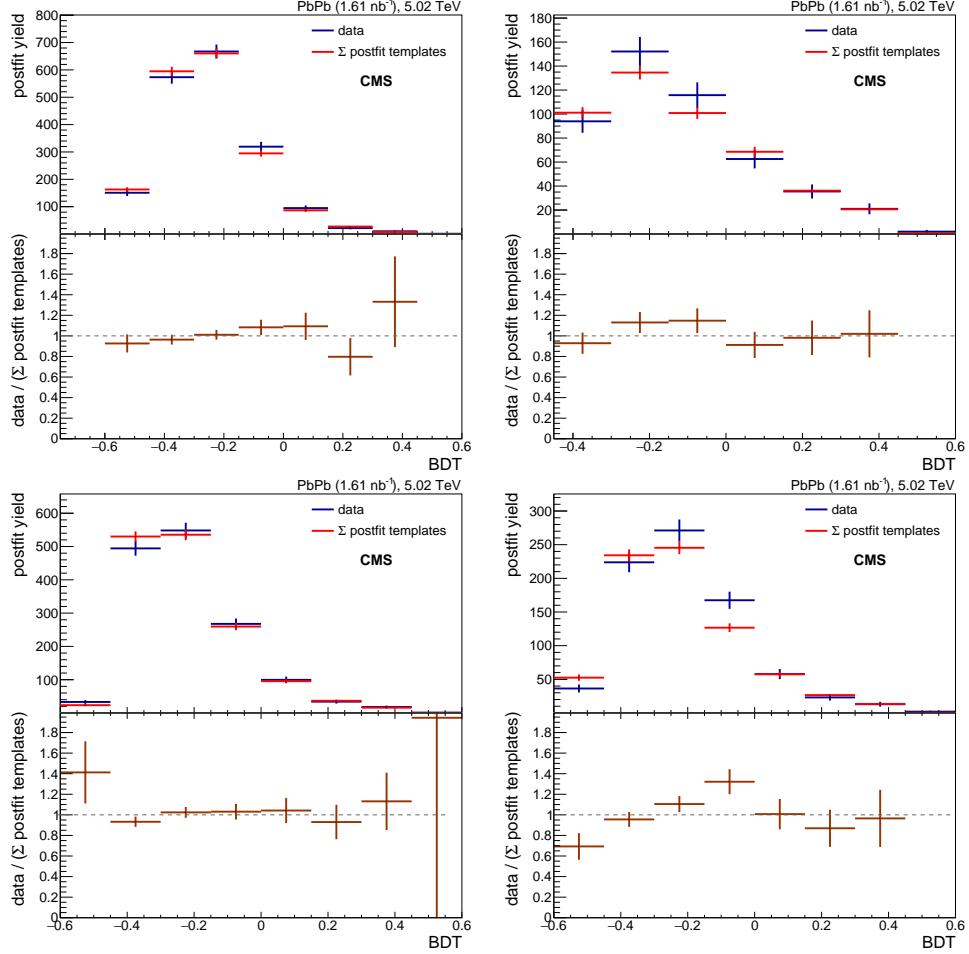


FIGURE 5.16: Comparison and ratio of the BDT distributions of data and of the sum of all post-fit templates (signal MC and the three backgrounds), for the first (top left) and second (top right)  $p_T^{\mu\mu}$  bins, and for the first (bottom left) and second (bottom right) centrality bins, in PbPb.

normalisation, and such bias hence transfers some discriminative power from the fit to the BDT binning.

One would like to check that leaving this discriminative power to the fit yields similar results than the nominal method. Therefore, I design here a modified BDT variable that is uncorrelated with the mass, when plotted for the summed backgrounds. This variable will be used to bin the samples in an alternative fit, whose result is used in the estimation of the systematic uncertainty on the template fit method (see section 6.4). This variable is computed separately for the two analysis bins. The decorrelation is applied to the BDT variables from both first- and second-step trainings, because the alternative fit is used in both steps to evaluate the systematics on the fit method. All plots shown in this subsection correspond to the second-step BDT.

To decorrelate the BDT from the mass, I start by plotting the BDT versus the mass for the summed backgrounds (Fig. 5.17). The mass binning is the same as the most fine-grained one used in the template fit, and only events in the fiducial cuts (of section 5.6) are used. The events are weighted to reproduce the post-fit mass distributions and normalisations; this decorrelation is used only in variations of the fit method, so the results of the first- or second-step nominal fit can be used.

One can then plot the average BDT versus mass for the summed backgrounds, with the finest binning used in the template fit (*i.e.* the 2D histograms of Fig. 5.17 profiled along the

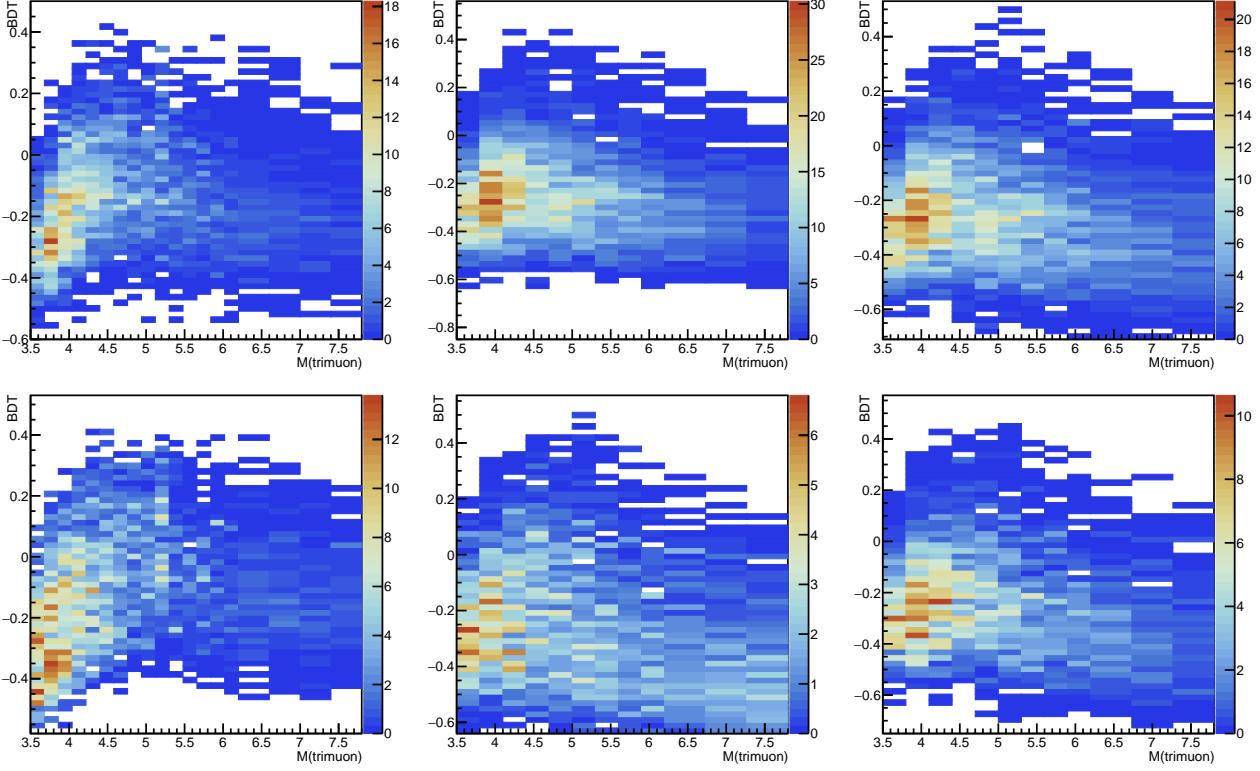


FIGURE 5.17: 2D distributions of BDT versus the trimuon mass, for the summed backgrounds in pp (left), in PbPb  $p_T^{\mu\mu\mu}$  bins (middle), and in PbPb centrality bins (right), for the first (top) and second (bottom) bins.

mass). This function  $\langle bdt \rangle_{background}(m^{\mu\mu\mu})$  is shown in Fig. 5.18, separated in  $p_T^{\mu\mu\mu}$  bins, for pp. The RMS of the BDT distribution in each mass bin can also be computed from Fig. 5.17; the resulting plots are not shown for brevity; the RMS changes smoothly between 0.11 and 0.17 over the mass range.

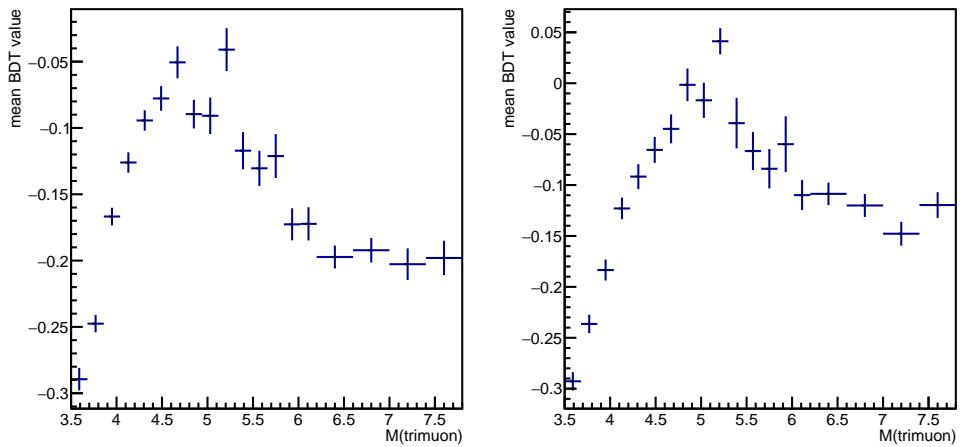


FIGURE 5.18: Mean BDT value in each trimuon mass, for the summed backgrounds in pp, in the first (left) and second (right)  $p_T^{\mu\mu\mu}$  bins.

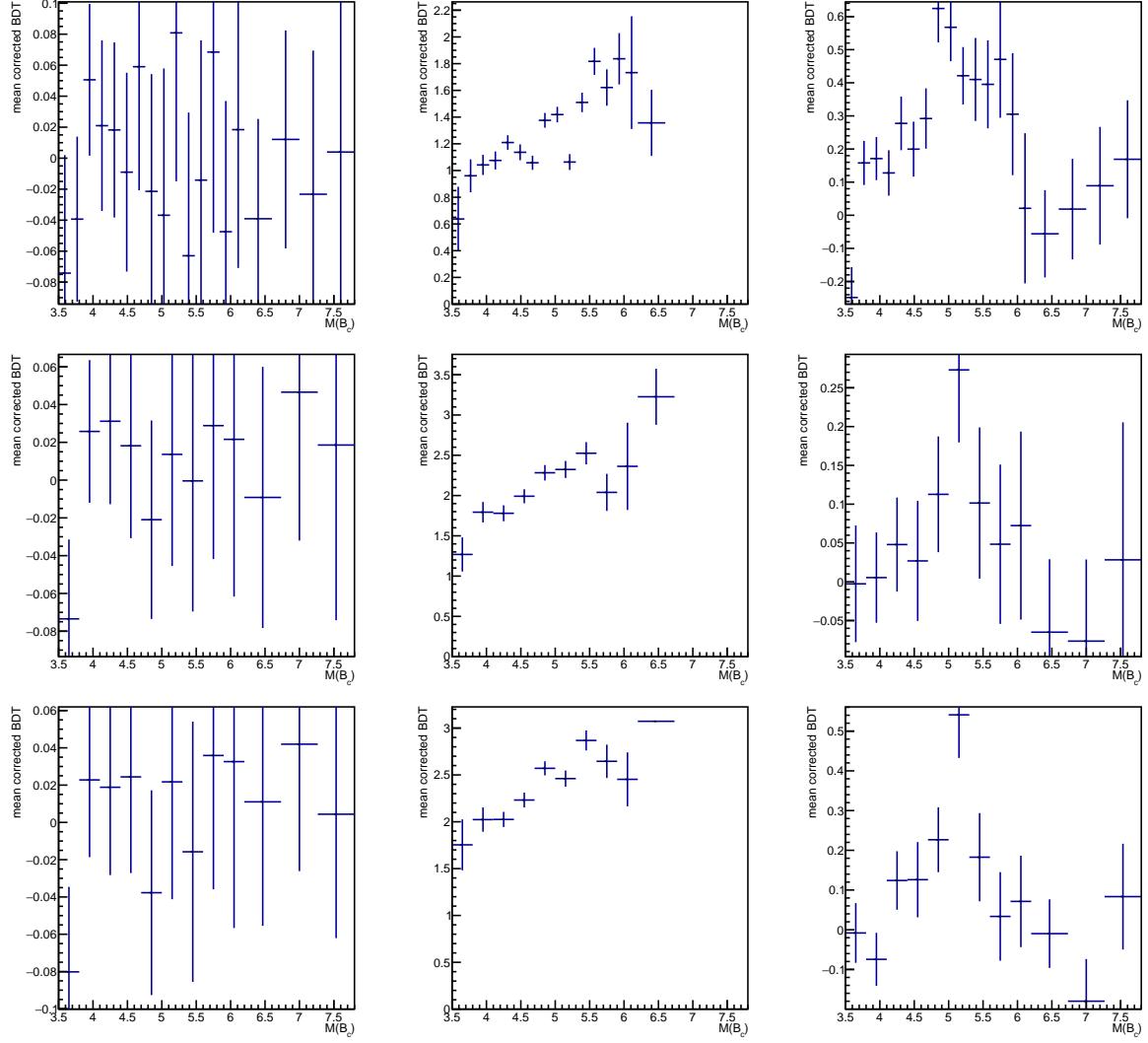


FIGURE 5.19: Mean corrected BDT versus the trimuon mass in the first  $p_T^{\mu\mu\mu}$  bin, in pp (top) and PbPb (middle), and in the first PbPb centrality bin (bottom), for the summed backgrounds (left), signal MC (center), and data (right).

I can now compute a corrected BDT value that is (in average) uncorrelated with the mass for background candidates, and that has a similar spread for all masses:

$$bdt' = \frac{bdt - \langle bdt \rangle_{background}(m^{\mu\mu\mu})}{RMS_{bdt}(m^{\mu\mu\mu})} \quad (5.3)$$

Binning in this variable instead of the original BDT in the template fit gives similar background shapes in all bins of corrected BDT. There is however a remaining mass dependence of the BDT variable in the signal MC, because the aim was only to keep the total background shape stable. Fig. 5.19 shows the mean corrected BDT value versus mass in the first  $p_T^{\mu\mu\mu}$  bin, in pp and PbPb, and in the first PbPb centrality bin, for the total background, the signal MC, and data.

This procedure is repeated separately for each  $p_T^{\mu\mu\mu}$  or centrality bin, and for the integrated samples, in pp and PbPb. Each of these categories hence has a different corrected BDT value. The fits using this BDT decorrelated from the mass are shown in section 6.4.2.

“Do you really believe that the moon exists  
only when you look at it?”

---

Albert EINSTEIN to Abraham PAIS<sup>1</sup>

## Chapter 6

# How many do we see? Template fit

Trimuon mass histograms for data, for the simulated signal, and for the various backgrounds (detailed in chapter 4) are now extracted. The core of the analysis, presented in this chapter, is then to extract the signal normalisation through a likelihood template fit, using the extracted signal and background trimuon mass histograms as input templates. Section 6.1 explains the principle and structure of the fit. Section 6.2 presents the nuisance parameters included in the fit to take into account the background uncertainties. Section 6.3 shows the fit results and associated crosschecks. Section 6.4 details the alternative fit methods that are tried, and the resulting systematic uncertainty.

### 6.1 Principle

#### 6.1.1 Likelihood fit with `combine`

The likelihood fit is run in the framework of the `combine` tool of the CMS software, based on ROOFIT [195]. This framework was conceived to combine the statistical power of multiple Higgs measurements [196], but is now very complete and used for general purposes. Its principle is to fit data with a model featuring probability distributions for each signal or background source, of parameters  $\vec{\beta}$ , simultaneously in many *channels*, which can be individual bins of histograms.

*Bayesian* logic is used via the *likelihood*  $\mathcal{L}(\vec{\beta})$ , which is proportional to the probability of obtaining the observed data if  $\vec{\beta}$  are the *true parameters* of the model chosen to describe the data. Each data event (or bin content in the case of a template fit) contributes to the likelihood with a factor equal to the probability of observing it given the considered model parameters. The core of the likelihood fit consists in simple counting experiments, where the observed number of events is compared to the number of signal and background events expected from the model. For this simple case the model is a Poisson probability with mean the number of expected events. More complicated models include various parametrised yield-modifying constraints in the likelihood. The most probable value (in a Bayesian sense)  $\vec{\beta}_0$  of the parameters is then found by maximising  $\mathcal{L}$ , or equivalently minimising the conventionally-used (double of the) negative log-likelihood (NLL)  $-2 \ln(\mathcal{L})$ . Wilkes’ theorem [197] tells that, at infinite statistics, the NLL approaches a  $\chi^2$  distribution with the number of degrees of freedom corresponding to the number of parameters

---

<sup>1</sup>Told by A. PAIS in Rev. Mod. Phys. **51**, 863–914 (1979), p. 907

left free. Therefore, the region of the parameter space defined by  $-2(\ln \mathcal{L}(\vec{\beta}) - \ln \mathcal{L}(\vec{\beta}_0)) < 1$  is a 68% credibility region.<sup>1</sup> The frontiers of this region can be interpreted as the  $1\sigma$  (*i.e.* 1 standard deviation) asymmetric uncertainty on the best-fit values  $\vec{\beta}_0$  of the parameters, as detailed in [198]; this is called the MINOS technique in the MINUIT minimiser [199] used in this fit.

The  $\vec{\beta}$  parameters comprise the ones we want to measure (*parameters of interest*, or POIs) and the ones that only implement constraints and uncertainties of the model (*nuisance parameters*). We typically do not need to know the best-fit values of the nuisances, so the likelihood is *profiled*: for each value of the parameters of interest, the NLL is minimised along nuisance parameters. This profiled NLL, depending now only on POIs, is used for the minimisation and credible interval estimation. When nuisance parameters correlated with the POIs are included in the fit procedure, the global shape of the likelihood is changed, so the profiled likelihood ‘trajectory’ in the parameter space is modified along with the credible intervals on POIs.

The template fit is a binned shape analysis, *i.e.* it is based on principles stated in [200] except histograms are given as the input rather than multiple expected yields. Mathematically it is equivalent to a simple counting experiment for each bin of each inputted histogram of each channel. However, the nuisance parameters implemented in the likelihood model affect the yields of various bins in a correlated way.

### 6.1.2 Structure, input and binning

The two parameters of interest (named  $r_1$  and  $r_2$ ) in the pp or PbPb fit are multipliers of the normalisations of the simulated signal templates in the two  $p_T^{\mu\mu\mu}$  or centrality bins. They should be close to 1 in pp if our result agrees with the previous measurements (used for the a priori signal normalisation, see section 3.5.4) and if the signal MC has correct kinematics, and in PbPb too if  $B_c$  mesons are not modified in the QGP. In addition, multiple nuisance parameters take into account the uncertainties on the normalisations and shapes of the templates; they are detailed in section 6.2.

In the trimuon invariant mass calculation, the mass of the dimuon chosen as the  $J/\psi$  (cf section 5.5) is fixed to the PDG  $J/\psi$  mass, both to reduce the effect of the different kinematic distributions in the dimuon sidebands (see section 4.2), and to improve the  $m^{\mu\mu\mu}$  resolution by removing the dimuon mass smearing by the imperfect detector reconstruction. The input shape histograms are defined on the [3.5, 7.8] GeV mass range. The [3.5, 6.2] GeV region is the mass signal region (less than 0.2% of the PbPb preselected signal events have a mass outside this range). The [6.2, 7.8] GeV region above the  $B_c$  mass is used as a control region to get a better hold on the background. The mass binning is made dependent on the BDT bin, because low-BDT bins contain many more events: [15, 13, 12] (or [9, 7, 6]) bins in the mass signal region in pp (PbPb), and [4, 3, 2] (or [3, 2, 1]) bins in the control region in pp (PbPb), respectively for the [first, second, third] BDT bins, from background-enriched to signal-enriched.

Mass histograms are drawn separately for each of the two analysis bins (*i.e.*  $p_T^{\mu\mu\mu}$  or centrality), and for each of the three BDT bins, which sums to six ‘channels’ that are simultaneously fitted. Fits on the integrated samples are performed as well, with only three channels corresponding to the BDT bins. Many parameters affect all the channels concomitantly. This can lead, for example, to the backgrounds being constrained mainly by the low-BDT bins and propagated to the high-BDT bins, in which the signal can be evaluated with a constrained background. In that sense, it is a two-dimensional fit in mass and BDT, or even a three-dimensional fit considering the two analysis bins. The limits of the BDT bins change according to the channel or fit method (*e.g.* when the mass-decorrelated BDT is used, as in section 6.4.2), and are always defined such that the first, second, and third bin respectively contain approximately 25%, 40%, and 35% of

---

<sup>1</sup>Keeping in mind, as explained later with the likelihood profiling, that only one parameter (usually the parameter of interest, or a linear combination of multiple parameters) is varied at a time.

the expected signal. The lowest considered BDT value is set such that 0.1% of signal MC events (of the considered analysis bin) are excluded, which rejects from the fit the events that are almost without doubt background (and that 0.1% is included in the preselection inefficiency). For fit stability reasons, the bins of all shape histograms are forced to have a positive content: if there is a bin with negative content (for example due to the  $J/\psi$  sidebands subtraction in the rotated- $J/\psi$  sample), it is set to zero as its uncertainty, and its old content is subtracted from a neighbouring positive-content bin.

In practice, the details of the model, parameters, and input shapes, are gathered in a *datacard*. The input shapes are:

- The **signal region data**. Before the data unblinding (that took place in a late stage of the analysis work), the PbPb data was partially blinded except in the high-mass control region and for very low BDT values for which almost no signal was expected. Blinded data contained only a quarter of the available data, and were attributed a weight <sup>1</sup>.
- The **expected signal** from the MC (section 3.5.2.1).
- The  **$J/\psi$  dimuon mass sidebands**, describing the fake  $J/\psi$  background (section 4.2).
- The **non-prompt  $J/\psi$  MC**, where only the candidates where  $J/\psi$  and muon come from the same B meson (section 4.3) are kept.
- The **data-driven rotated- $J/\psi$**  sample is used to describe the background from a  $J/\psi$  meson associating to a muon from a different decay vertex (section 4.4). In pp, the ‘combinatorial’ part (*i.e.* excluding true B decays) of the non-prompt MC, known to be underestimated (cf section 4.4.2), is added to it; it is redundant with it, but a nuisance parameter deals with this overlap through alternative shapes (section 6.2). Similarly, in PbPb, the shape of this background will be allowed to vary to the prompt plus non-prompt  $J/\psi$  MC samples (without the true B decays), which is considered a good background description too.

In addition to some preliminary fits in nominal conditions, the full fit (with fit method variations of section 6.4) is performed twice. In the second-step of the analysis, the fit benefits from the improved BDT and the correction of the  $p_T^{\mu\mu\mu}$  spectrum of signal MC. In the second-step, I also run a preliminary nominal fit, whose signal and background parameters are used in a check of the BDT distribution of the sum of templates versus the one of data; in pp, it results in a weighting of the BDT value, before the full second-step fit (see section 5.7.4). Unless mentioned otherwise, all plots of this section show the results of this final second-step fit. Five types of fit are run: integrated over the fiducial kinematic region in pp and PbPb, in  $p_T$  bins in pp and PbPb, and in centrality bins in PbPb. The centrality dependence and integrated fits are run only in the second step.

## 6.2 In-fit systematics (nuisance parameters)

Nuisance parameters take into account the uncertainties on the background templates, so that the final POI uncertainties given by the fit reflect these background uncertainties. The datacard also contains the initial guess and constraining ranges for the nuisance parameters, as well as how they shall affect specific background or signal shapes. If a background normalisation  $N$  is not precisely known, it is associated a parameter that modifies this rate, according to a log-normal distribution with a relative width of  $\sigma_N^{rel}$ . This means a  $+1\sigma$  variation will raise the rate by  $\sigma_N^{rel}\%$ . In practice, the normalisation is changed to  $N(1 + \sigma_N^{rel})^\theta$ , where the nuisance parameter  $\theta$  is constrained via a gaussian term  $e^{-\frac{1}{2}\theta^2}$  multiplying the likelihood.

<sup>1</sup>This meant the central values of the PbPb data points in blinded regions were a proxy of what was expected after unblinding, but the error bars displayed on the plots were  $\sqrt{4} = 2$  times larger.

Similarly, if a background shape is poorly known, one can provide  $\pm 1\sigma$  alternative shapes to the datacard, and a gaussian-constrained nuisance parameter will allow the shape to vary consistently between the nominal and these two extreme shapes. More explicitly, a parameter  $\theta$ , leading the so-called *shape morphing*, is considered in the likelihood with a term  $e^{-\frac{1}{2}\theta^2}$ . When  $\theta = 0$  or  $\pm 1$ , the fit takes respectively the nominal or alternative shapes for this background. For other values, a vertical interpolation of the background fractions in each mass bin is done, with a spline for  $|\theta| < 1$  and linearly outside this range.

The MC samples have *a priori* normalisations that correspond to a value of 1 for their normalisation modifier parameter. The data-driven rotated-J/ $\psi$  sample should be self-normalised (*i.e.* directly given by the method performed on data), but we still consider a normalisation modifier for it in pp (section 4.4 motivates this choice). The following nuisance parameters are included in the likelihood model to take into account the uncertainties on normalisations:

- The normalisation of the (true B decays part of the) **non-prompt J/ $\psi$  MC** is multiplied by a log-normal-distributed parameter of initial value 2 in pp (2.2 in PbPb) and relative width 40%<sup>1</sup>. The initial value and loose constraint aim at guiding the fit. The initial value is based on preliminary fits. This variable normalisation could correct for a mismodelled muon misidentification rate in the MC (see section 4.3), but also in pp for the underestimation of MC yields due to wrong J/ $\psi$  distributions (those discrepancies are described in section 4.4.2).
- The data-driven **rotated-J/ $\psi$  sample** is expected to have close to correct normalisation, especially in PbPb where all rotation angles give similar yields (see section 4.4.3). However, in pp, one would like to leave freedom to the fit for this background, considering the somewhat arbitrary choice of the mix of rotation angles (which can each give a different normalisation). Therefore this background has a log-normal normalisation modifier of initial value 1 and relative width 40%. The normalisation is fixed to 1 in PbPb; when leaving it completely free, this parameter gets a value very close to 1 and an uncertainty of 7%.
- The global normalisation uncertainties (on the luminosity and Glauber model for centrality, and on the contribution of other B<sub>c</sub> decays), are *not* implemented here, but rather at the end of the analysis, as detailed in chapter 8. They are given special status because of their correlation between pp and PbPb (the uncertainty from other B<sub>c</sub> channels partially cancels in the  $R_{\text{PbPb}}$ ) or because they affect all  $p_T^{\mu\mu\mu\mu}$  bins with the same relative uncertainty.

In pp, the 40% pre-fit uncertainty on the rate parameters (for the two true J/ $\psi$  backgrounds), that we would like to leave free while guiding it with initial values, is a bit arbitrary. However it was checked that doubling or halving it does not change the post-fit value, nor the POI values nor their post-fit uncertainties (or by less than 4% of the POI uncertainty). It should however stay significantly above the post-fit uncertainty, equal to 0.05 for the B decays and 0.05 for the rotated J/ $\psi$  background in pp, so that the pp rate modifiers are not significantly constrained.

The pre-fit value and uncertainty of the normalisation of the true B decays MC sample in PbPb are more challenging, because it has subdominant yields compared to the other backgrounds, so it is hardly constrained by the fit. Leaving it completely free gives a value  $3.2 \pm 1.0$ . One can however consider that it is somewhat constrained by the pp fit giving a value around 1.9, even if multiple running conditions differ. I therefore decided to set the initial value to 2.2 with a pre-fit (multiplicative) uncertainty of 40%, which leaves both values accessible. Halving or doubling this uncertainty has a very small impact on the POIs post-fit central values and uncertainties (less than 2% of the fit uncertainty).

---

<sup>1</sup>In `combine`, it is a `rateParam` normalisation modifier for the initial value, fixed in the fit, and a unique `lnN` nuisance parameter with 40% width applied on all channels of the related background.

The following nuisance parameters<sup>1</sup> allow for shape morphings that fully accommodate the uncertainty on the chosen shapes:

- The nominal shape for the B-decays background is the full non-prompt MC in pp, and only the fully correlated part of it ( $J/\psi$  and muon coming from the same B decay) in PbPb. In pp, the  $\pm 1\sigma$  shape variations are the non-prompt MC where the combinatorial part ( $J/\psi$  and muon from different decay vertices, about 18% of the non-prompt MC) is removed, and the summed prompt and non-prompt  $J/\psi$  MC samples. In PbPb, a shape variation including the combinatorial part of the non-prompt MC (that constitutes 92% of the full non-prompt MC) in the shape of the B decays background was tried. However, this is a modification of the shape at fixed normalisation, and the yields of this background are subdominant in PbPb. Consequently, this shape variation was seen to have no observable effect on the fit central value and uncertainty, so it was removed.
- The initial shape describing the background from combinations of a  $J/\psi$  and a muon from different decay vertices is the rotated  $J/\psi$  sample, integrated over all thirteen rotation angles. In pp the  $\pm 1\sigma$  shape variations only integrate over the six rotations that reverse or are close to reversing the azimuthal angle ( $|\Delta\phi| \geq \frac{3}{4}\pi$ ), or integrate over the seven other rotations; those shapes are compared in Fig. 4.15 (top). In PbPb, varying the rotation angles has very little effect (see Fig. 4.15, bottom), so we simply set a unique  $\pm 1\sigma$  shape variation as the prompt plus non-prompt  $J/\psi$  MC (without its fully correlated B decays part), which is considered as a good description of this background in PbPb, as shows the shape comparison of Fig. 4.16 (right).
- The fake  $J/\psi$  background is nominally described with the sum of the two dimuon mass sidebands. Separately, the shapes from the two sidebands (shown in Fig. 4.3) are extreme limits for the shape of this background, therefore they are taken as  $\pm 2\sigma$  variations.

In each analysis bin and BDT bin, these varied shapes are normalised to have the same integral as the nominal shape in this bin. The fit can hence favour the nominal or varied shape without affecting simultaneously the normalisation parameter of this background.

The use of data-driven or even MC histogram templates that have low statistics can significantly affect the fit results, for the backgrounds are estimated only by sampling the true background distributions. This is accommodated through the `autoMCstats` parameters in `combine`, following the Barlow-Beeston procedure detailed in Ref. [201]. One nuisance parameter is added per mass bin, and possibly per signal or background template. For each mass bin showing a non-empty prediction, an effective number of total background events (evaluating the statistical power of this bin content)  $n_{\text{tot}}^{\text{eff}} = n_{\text{tot}}^2/e_{\text{tot}}^2$  is calculated from the error  $e_{\text{tot}}$  on the total background in this bin. If  $n_{\text{tot}}^{\text{eff}} > 10$ , a single parameter constrained by a Gaussian of width  $e_{\text{tot}}$  will add to the sum of signal and background templates in this bin. If not, a Poisson parameter is to multiply each non-empty background separately in this bin. These parameters follow a Poisson distribution of mean the number of events expected from this background  $n_{bkg,i}$ , and the variable is divided by  $n_{bkg,i}$  so that the mean value of this multiplier is indeed 1. A parameter is also added for signal, Gaussian- or Poisson-distributed depending on whether the signal bin content is above threshold.

These parameters are numerous, but are strongly constrained, especially when the backgrounds are high (and hence have small statistical uncertainties). The Poisson threshold is though chosen not too high to limit the number of parameters. These parameters lead to worrying instabilities with templates having very low statistics (typically only a few MC events over the whole mass range in a BDT bin, corresponding to less than 1 or 2 expected events), in versions of the fit where the shape variations are normalised only to the nominal shape integrated on the

---

<sup>1</sup>Keyword `shape` in the nuisance entered in the COMBINE datacard

BDT bins, and not normalised in each BDT bin as is nominally done. It is understood as many largely correlated parameters building up large contents for supposedly-low background bins. This appears for background templates that have a very irregular shape due to the low statistics. When removing these template-statistics parameters in the fit method systematics or changing how the shape variations are normalised to the nominal shapes (as is done in section 6.4.3), we will need to regularise the low-statistics shapes before the fit, which solves the issue.

All nuisance parameters included in the fit are summed up in Table 6.1.

TABLE 6.1: Summary of systematic uncertainties implemented as nuisance parameters in the likelihood fit. Each row corresponds to one fit parameter, except for the last row which sets one parameter per mass bin.

Affected samples	Name	Initial value	$-1\sigma$	$+1\sigma$	Comments
<b>Fake J/<math>\psi</math></b>	shape	<b>full dimuon sidebands</b>	low-mass sideband	high-mass sideband	$\pm 2\sigma$ variations
<b>B decays</b>	normalisation	2. (pp) 2.2 (PbPb)	40%		multiplicative, log-normal
		<b>non-prompt J/<math>\psi</math> MC</b> full (pp) or keep only true B $\rightarrow$ J/ $\psi$ X (PbPb)	keep only true B $\rightarrow$ J/ $\psi$ X (pp) no variation (PbPb)	add prompt J/ $\psi$ MC (pp) no variation (PbPb)	
<b>J/<math>\psi</math>-<math>\mu</math> combinatorics</b>	normalisation	1 (pp) fixed to 1 (PbPb)	40% (pp)		multiplicative, only in pp
		<b>rotated J/<math>\psi</math></b>	rotated-J/ $\psi$ , only $ \Delta\phi  \geq \frac{3}{4}\pi$ rotations (pp) no variation (PbPb)	rotated-J/ $\psi$ , other rotations (pp) uncorrelated (non-)prompt MC (PbPb)	
All	<b>Statistical uncertainties</b> on binned shapes	$\geq 1$ Poisson or Gaussian parameter per mass bin Poisson threshold 10			<b>autoMCstats</b> tool

## 6.3 Fit results

Before unblinding (explained in section 3.4), the results shown here for PbPb used only one quarter of the signal region data: the signal region data points had central values similar to the expected unblinded ones (except with higher statistical fluctuations), but two times larger uncertainties. The likelihood fit however considers a Poisson constraint between data points and the signal plus background yields from the model, so only the central value (*i.e.* a simple number of events) matters for the pull exerted by a data point on the model in the likelihood. The eventuality of the fit uncertainties decreasing due to unblinding in PbPb (bringing smoother data mass histograms) was discussed with toy datasets in section 6.3.4.

### 6.3.1 Post-fit mass distributions and signal yields

In each mass bin, the post-fit yields of the signal and of each background are taken from the second-step fit result, and the post-fit templates are thence built. The fitted templates in pp for both  $p_T^{\mu\mu\mu}$  bins, and integrated over both analysis bins, are shown respectively in Figs. 6.2 and 6.1. They are also shown in PbPb in both  $p_T^{\mu\mu\mu}$  bins, both centrality bins, and integrated over the

whole phase space, respectively in Figs. 6.4, 6.5, and 6.3. Each of these figures correspond to a different simultaneous fit. The wrong-sign sample is only shown for illustration and not added in the summed backgrounds. For each sub-plot, the fraction of selected signal or of the total selected background that falls in this BDT bin is printed, as well as the purity (the fraction of fitted signal yields in the total yield) and the post-fit observed signal yield.

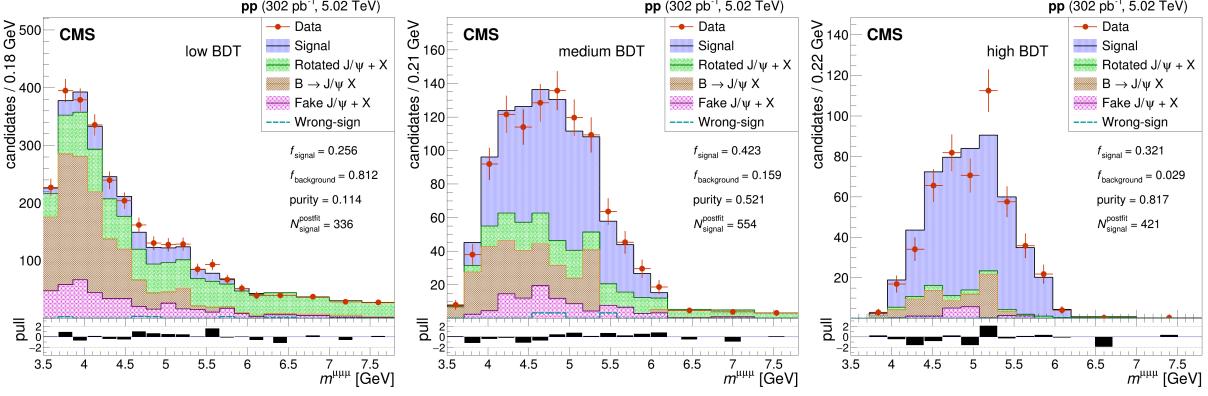


FIGURE 6.1: Template fit in pp, for all candidates within the studied kinematic region (defined in section 5.6).

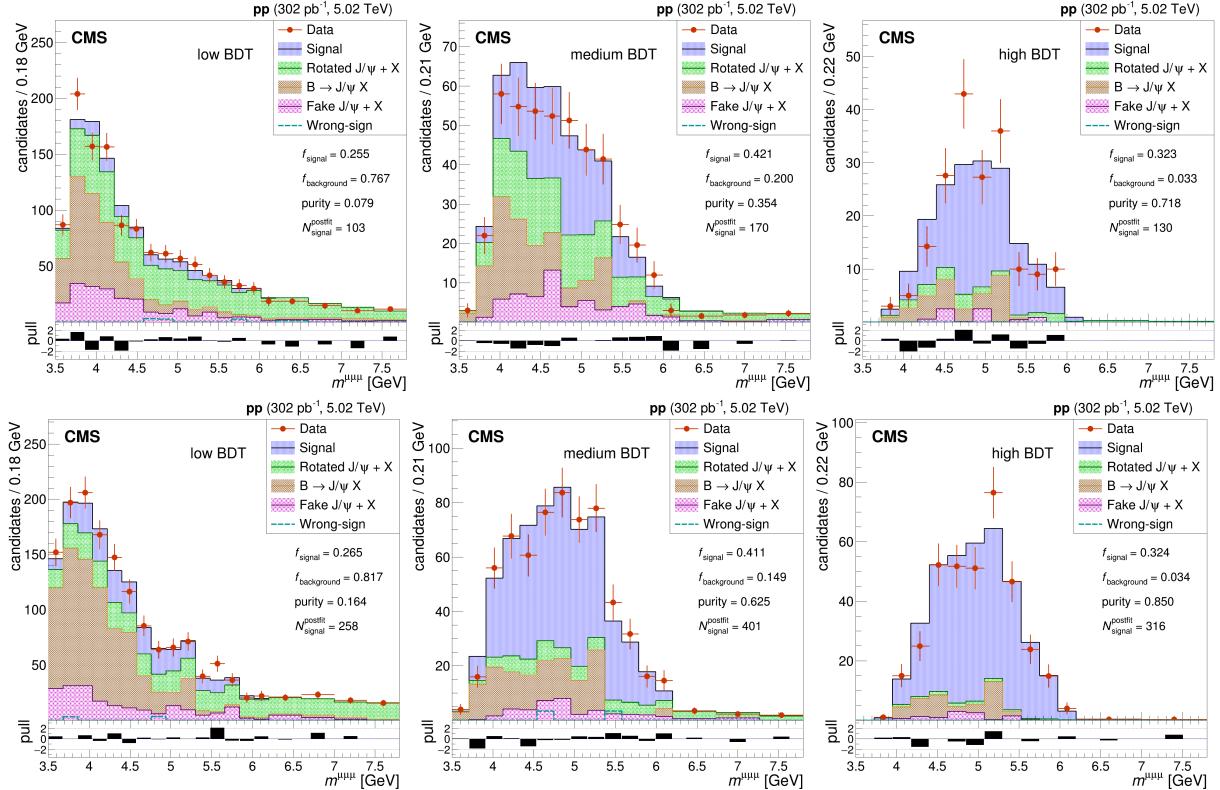


FIGURE 6.2: Template fit in pp for the  $p_T^{\mu\mu\mu\mu}$  dependence.  
Top: candidates with  $6 < p_T^{\mu\mu\mu\mu} < 11$  GeV. Bottom: candidates with  $11 < p_T^{\mu\mu\mu\mu} < 35$  GeV.

The sum of fitted background templates are to be compared to the data points, separately in each of the six ‘channels’ (*i.e.* in each BDT bin and analysis bin). The pull graphs, showing for each bin the difference between the data and the post-fit model divided by the uncertainty on the data point, are included in the bottom panels of the fit plots. The pulls already suggest that the fits are good, but this is not the appropriate statistical test for a likelihood fit. A proper

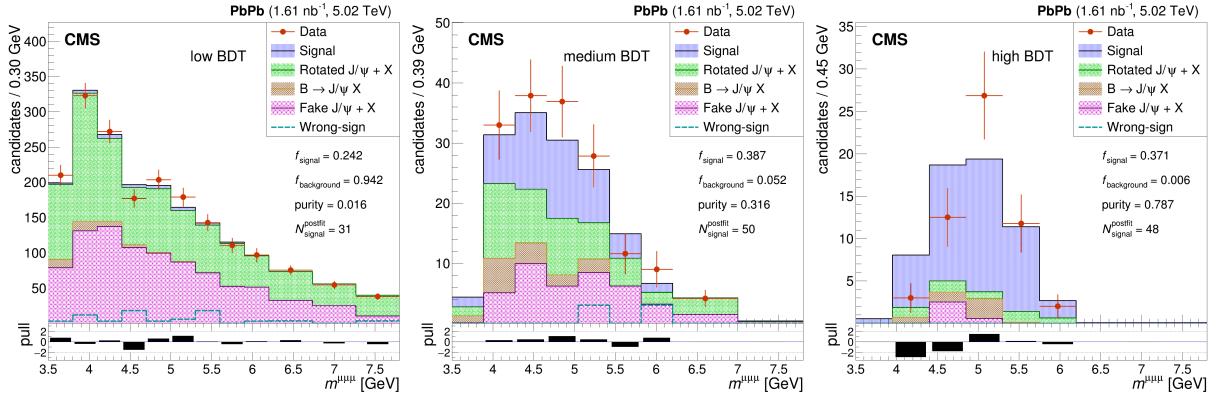


FIGURE 6.3: Template fit in PbPb, for all candidates within the studied kinematic region (defined in section 5.6).

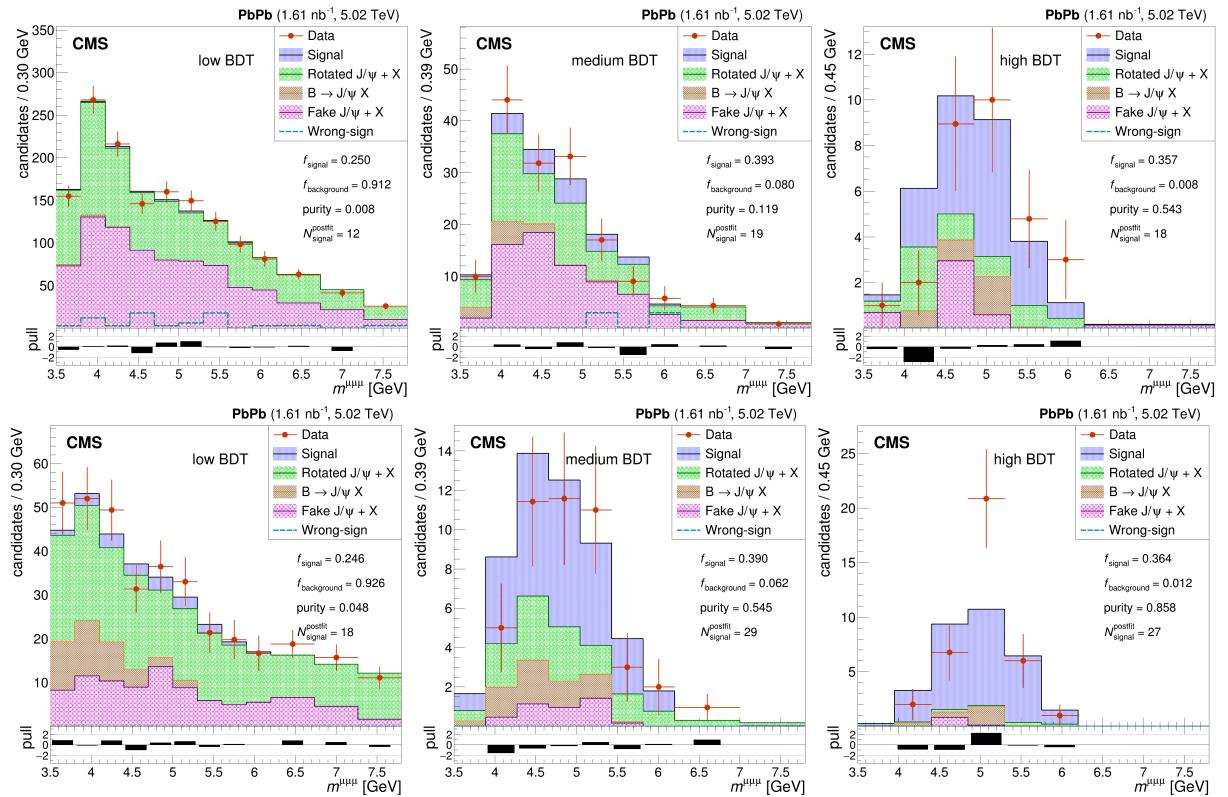


FIGURE 6.4: Template fit in PbPb for the  $p_T^{\mu\mu\mu}$  dependence.  
Top: candidates with  $6 < p_T^{\mu\mu\mu} < 11$  GeV. Bottom: candidates with  $11 < p_T^{\mu\mu\mu} < 35$  GeV.

goodness-of-fit test is performed in the coming section 6.3.2 and is satisfying. One feature making the fit reliable is that the low-signal BDT bin stabilises the background parameters, and the signal-enriched BDT bin mostly determines the signal yields using the background yields constrained by the other BDT bins.

The post-fit signal normalisation multipliers (the two parameters of interest  $r_1$  and  $r_2$ ) are shown for both fit steps in Table 6.2, along with the number of observed signal events from the final second-step fit, its uncertainties, and the correlation factors between analysis bins.

Given that the input signal for the second-step fit is weighted by the first-step result, the product of the first- and second-step  $r_i$  is expected to be 1 if the measurement gives exactly

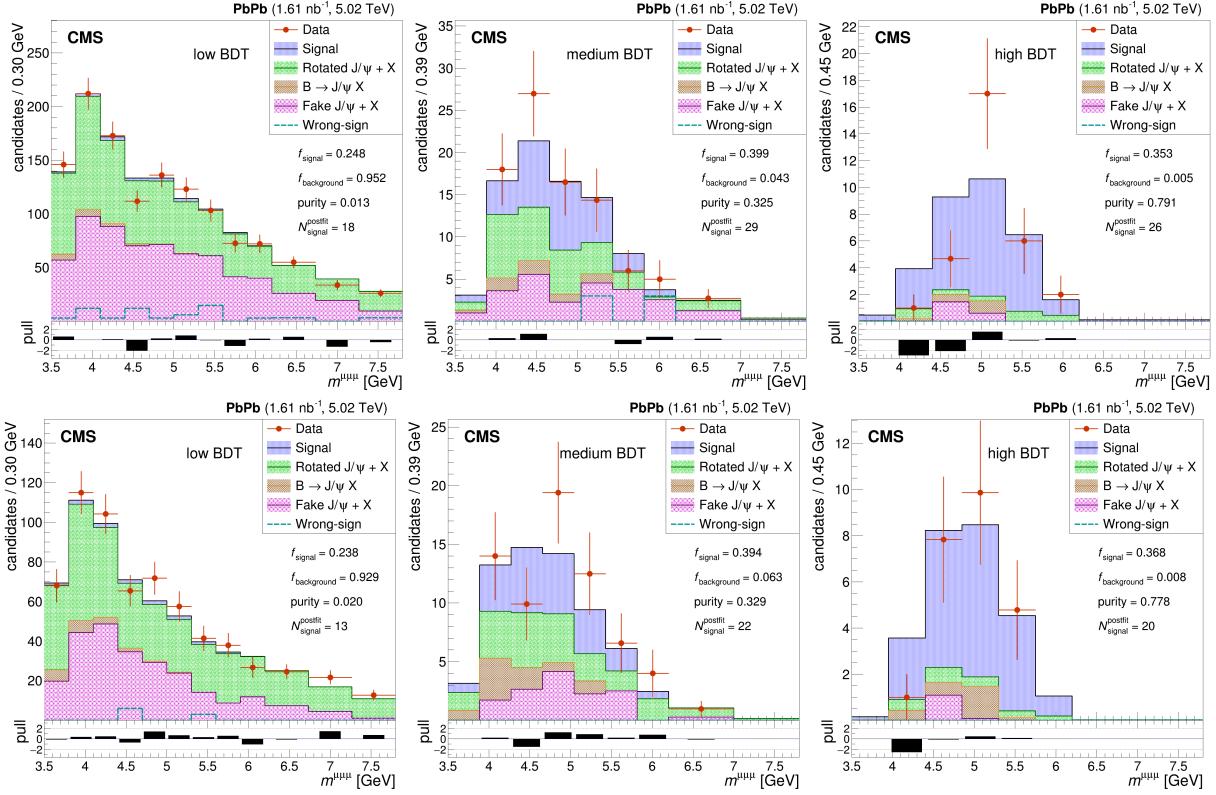


FIGURE 6.5: Template fit in PbPb for the centrality dependence, integrated over  $p_T^{\mu\mu\mu}$  bins.  
*Top:* candidates with event centrality in 0 – 20%. *Bottom:* centrality in 20 – 90%.

TABLE 6.2: Normalisation modifiers  $r_1$  and  $r_2$  (*i.e.* the ratio to the *a priori* signal normalisations) and the associated second-step number of observed signal  $B_c$  events, in all differential and integrated fits. The modifiers from the fits in the two different steps (before and after corrections of the MC  $p_T^{\mu\mu\mu}$  spectrum) are shown, along with their product which should be 1 if the *a-priori* MC  $p_T$ -differential yields are exact. The quoted errors are only the ones on  $r_i$  parameters from the second-step fit. The (second-step) correlation factor between the normalisations of the two bins is also quoted.

		$r_i^{\text{first-step}}$	$r_i^{\text{second-step}}$	$r_i^{\text{first}} \times r_i^{\text{second}}$	$N_{B_c}^{\text{observed}}$	correlation $\rho(r_1, r_2)$
pp	$6 < p_T^{\mu\mu\mu} < 11 \text{ GeV}$	0.597	1.238	$0.739^{+0.068}_{-0.066}$	$403^{+37}_{-36}$	0.10
	$11 < p_T^{\mu\mu\mu} < 35 \text{ GeV}$	0.895	1.091	$0.977^{+0.050}_{-0.049}$	$976^{+50}_{-49}$	
	integrated	0.798	1.067	$0.851^{+0.041}_{-0.040}$	$1310^{+63}_{-62}$	—
PbPb	$6 < p_T^{\mu\mu\mu} < 11 \text{ GeV}$	0.839	1.306	$1.096^{+0.352}_{-0.323}$	$49^{+16}_{-15}$	0.05
	$11 < p_T^{\mu\mu\mu} < 35 \text{ GeV}$	0.524	1.140	$0.597^{+0.105}_{-0.098}$	$74^{+13}_{-12}$	
	centrality 0-20%	0.695	1.363	$0.947^{+0.192}_{-0.179}$	$74^{+15}_{-14}$	0.05
	centrality 20-90%	0.623	1.125	$0.701^{+0.168}_{-0.156}$	$55^{+13}_{-12}$	
	integrated	0.634	1.258	$0.797^{+0.124}_{-0.118}$	$129^{+20}_{-19}$	—

the *a priori* cross section of section 3.5.4.2. We observe that the normalisation modifier in pp is consistent with 1 for the second  $p_T^{\mu\mu\mu}$  bin, but not for the first  $p_T^{\mu\mu\mu}$  bin, which implies a discrepancy with the *a priori* spectrum from BCVEGPY. That discrepancy in the slope of the spectrum will be discussed in chapter 9. However, the normalisation of the integrated sample is only 15% lower than the *a-priori* one, and this is consistent with 1 if we take into account the (probably under-estimated) 14% uncertainty on the normalisation from previous measurements given in Eq. 3.7. The normalisation modifiers differ from 1 (the normalisation based on pp measurements) in PbPb, which is expected in the presence of modifications compared to pp.

A satisfying cross-check from Table 6.2 is that, both in pp and PbPb, the number of  $B_c$  mesons observed in the separate fit on the integrated sample is consistent within uncertainties with the sum of observed yields from the simultaneous fits of the two  $p_T^{\mu\mu\mu}$  or centrality bins.

The relative uncertainties on the yields from the fit are 5-9% in pp, and 17-31% in PbPb, depending on the bin. The uncertainties can be measured with the HESSE method or the MINOS method. The first calculates symmetric uncertainties via the second derivative of the likelihood at the found minimum (assuming the NLL has a parabolic shape near the minimum). The second looks numerically for the values of the parameters for which the NLL equals 1, which are the frontier values of the  $\pm 1\sigma$  region,<sup>1</sup> as explained in section 6.1.1. This method can give asymmetric uncertainties. We observe that the low-side uncertainties are about 10% smaller than the high-side ones in PbPb, so we keep the asymmetric MINOS uncertainties.

### 6.3.2 Goodness of fit

The  $\chi^2$  value of the fit is not given, because it is not a valid goodness-of-fit test in the context of a likelihood fit with many strongly constrained parameters. Instead, we use a test statistic  $t$  based on the saturated algorithm [202], that is a likelihood ratio test similar to a  $\chi^2$  but more adapted to non-Gaussian bin contents and to a likelihood with parameter constraints. 500 toy datasets are generated following the post-fit signal plus background distributions (from the second-step fit), meaning in a given toy, the content of a given bin is drawn from a Poisson distribution of mean equal to the bin content in the post-fit model.  $t$  is measured for each toy to estimate the test statistic distribution  $f(t)$ . The value for data  $t_0$  is then calculated, and compared to the distribution to get the p-value  $\int_{t_0}^{\infty} f(t) dt$ . The *p-value* corresponds to the probability to obtain a test statistic as bad as in data if the fit model is right. The data value  $t_0$  is compared to the distribution of  $t$  in Fig. 6.6 for pp and PbPb. The p-value is very good in pp (34%) and in PbPb (76%), indicating that the fit model is well-chosen, up to the statistics available in data.

In PbPb when the data was partially blinded (see strategy in section 3.4), this test was run via fits of the blinded dataset with the true associated (blinded) yields, meaning that the events in the partially blinded regions were *not* multiplied by 4 (contrarily to samples used in the rest of the analysis), and the events in non-blinded regions had their weight *divided* by 4. If we had used the blinded sample weighted to seemingly have unblinded yields, the p-value would have been bad: the likelihood test indeed knows only about the number of events in the bins, whose fluctuations would have been artificially large because of the use of only a quarter of the data.

### 6.3.3 Observed significance

Let us now evaluate the significance of the observation in PbPb, from the  $p_T^{\mu\mu\mu}$ -dependent fit. We plot in Fig. 6.7 the scan of  $-2\Delta \ln \mathcal{L}$  (twice the difference of NLL between a point and the best-fit point) versus the parameters of interest divided by their best-fit value  $r_1/r_1^{best}$

---

<sup>1</sup>It is reminded here that this assumes that the Wilkes' theorem applies, so that the log likelihood-ratio is close to a  $\chi^2$  distribution.

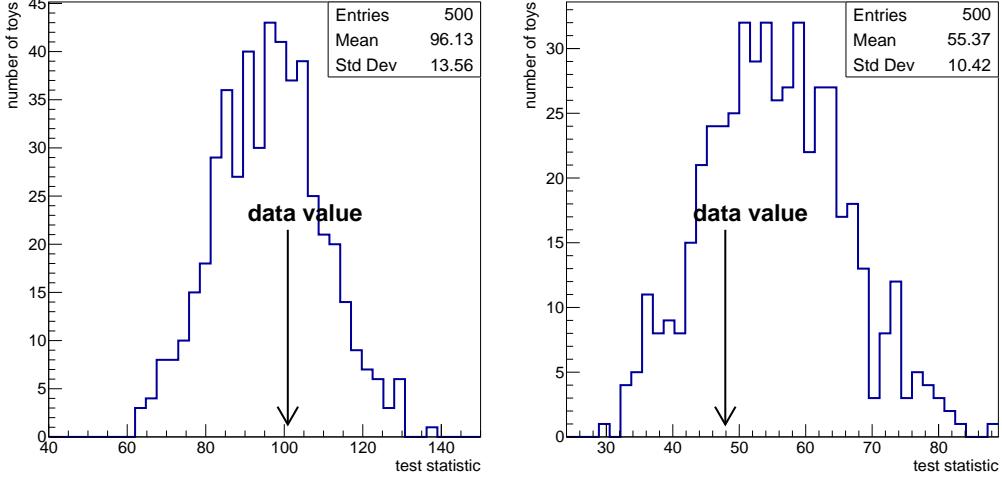


FIGURE 6.6: Goodness of fit test in pp (*left*) and PbPb (*right*). The p-value (probability to observe a test statistic worse than the data value if our post-fit model is right) is the fraction of toys having a larger test statistic than the data value, which is 101 in pp and 48 in PbPb.

and  $r_2/r_2^{best}$ . For each  $(r_1, r_2)$ , the nuisance parameters are profiled again. As explained in section 6.1.1,  $-2\Delta \ln \mathcal{L}$  approaches a  $\chi^2$  distribution (via Wilks' theorem), meaning that the value 1 provides a  $1\sigma$  contour, and the value  $25 = 5^2$  sets a  $5\sigma$  contour. The latter is far away from the  $(r_1 = 0, r_2 = 0)$  point, which shows that the observation is  $> 5\sigma$  significant. Assuming the likelihood has a Gaussian shape, the observed significance in PbPb can be calculated as  $\sqrt{-2 \ln \mathcal{L}}$  at the  $(r_1 = 0, r_2 = 0)$  point; this results in a significance of the observation of  $B_c$  mesons in PbPb collisions of  $7.4\sigma$ .

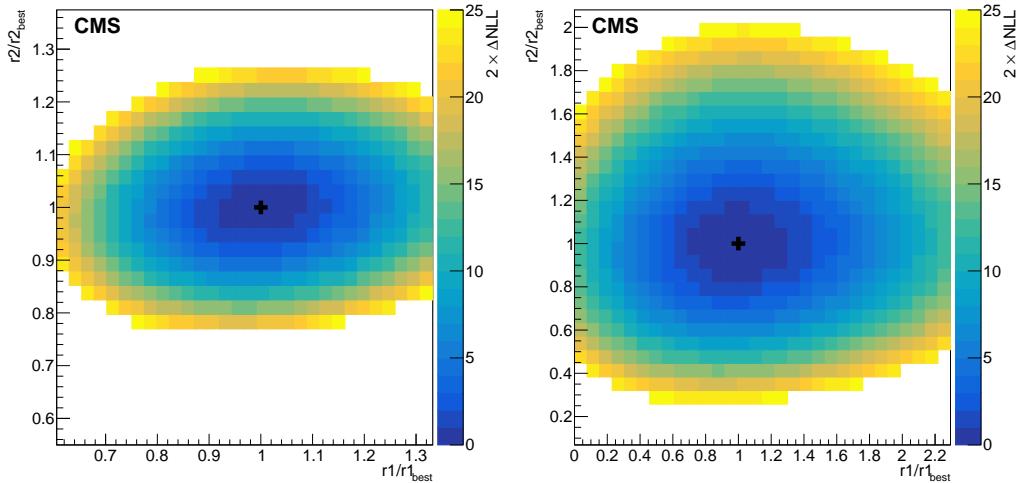


FIGURE 6.7:  $-2\Delta \ln \mathcal{L}$  versus the parameters of interest  $r_1$  and  $r_2$ , in pp (*left*) and PbPb (*right*). Only the regions with  $-2\Delta \ln \mathcal{L} < 25$  are coloured, which approximately correspond to a  $5\sigma$  contour. The first (dark blue) colour level corresponds to a  $1\sigma$  contour.

This value of the significance takes into account only the fit uncertainty. However, I intend to estimate here the significance of the observation of the  $B_c$  *yields*, not the  $B_c$  *cross-sections*, so we can neglect other uncertainties if they are mostly of multiplicative nature, like those affecting luminosity, acceptance, and efficiency. The only other uncertainty that is mainly of additive nature is the *metafit* one dealing with the choice of fit method, presented in the next section 6.4. We use the values of the relative uncertainties  $\sigma_{1,metafit,rel}$  and  $\sigma_{2,metafit,rel}$  in the two  $p_T^{\mu\mu\mu}$  bins

in PbPb, and their correlation factor  $\rho_{1-2}$  from Table 6.4. The significance quoted above is the z-score of a one-dimensional equivalent gaussian approximating the likelihood, evaluated at the ( $r_1 = 0, r_2 = 0$ ) point. So  $z = \sqrt{-2 \ln \mathcal{L}} = \frac{r_{best}}{\sigma_{fit}} = \frac{1}{\sigma_{fit,rel}}$ , where  $\sigma_{fit,rel}$  is the 1D relative fit uncertainty equivalent to the 2D likelihood. One needs to sum in quadrature the 1D-equivalent relative uncertainties from the fit and from the fit method, which hence gives a final significance of

$$z_{final} = \frac{1}{\sqrt{\sigma_{fit,rel}^2 + \sigma_{metafit,rel}^2}} = \frac{1}{\sqrt{\frac{1}{-2 \ln \mathcal{L}} + \frac{1}{-2 \ln \mathcal{G}}}} \quad (6.1)$$

where  $\mathcal{G}$  is the value at  $(0, 0)$  of a 2D gaussian centered at the best fit point and that has the relative uncertainties and correlation factor from the section 6.4 fit method uncertainty.<sup>1</sup> The fit method uncertainty being sub-dominant, we obtain a **significance of the observation of  $B_c$  mesons in PbPb of  $z = 6.8\sigma$** , close to the one using only the fit. The associated one-sided gaussian p-value (probability of the data coming only from a statistical fluctuation of the background) is  $5 \times 10^{-12}$ . The significance evaluated from the centrality dependence should yield similar values; the integrated fit might have given somewhat smaller values, due to the removal of the  $p_T$  binning information.

### 6.3.4 Toy study for fit uncertainties

When the PbPb dataset was partially blinded, it was important to assess if the unblinding would affect the fit uncertainties. This initially motivated a study performing the  $p_T^{\mu\mu\mu}$ -dependent fit on toy datasets, but it also serves after unblinding as a crosscheck of the stability of the uncertainty and of the potential bias of the fit.

We hence generate 300 toy datasets drawn from second-step post-fit background distributions, plus the signal distribution where  $(r_1^{true}, r_2^{true}) = (1., 1.)$ . These values assume the second-step fit gives results close to the first-step one, and give more agnostic results than taking the post-fit POI values; it also allows to interpret the spreads and uncertainties directly as relative deviations. These datasets use the statistics of the unblinded data. A full fit is performed on each toy, and the values of the POIs  $(r_1, r_2)$  and their uncertainties are recorded. Figs. 6.8 and 6.9 show their distributions in pp and PbPb, that represent the POI values one could get by running another time the same experiment in the same conditions.

The means and RMSs printed on these plots provide multiple checks. The distribution of POIs and uncertainties in the pp toys are useful to check that the fit uncertainties are reliable. First, the RMSs of the  $r_1$  and  $r_2$  distributions (printed on the left panels of Figs. 6.8 and 6.9) are close to the uncertainties on these parameters in the nominal (*i.e.* using actual data) fit shown in Table 6.2; the small difference could come from the initial  $r_i^{true}$  value being set to 1 instead of the post-fit value. Second, the means of the  $r_1$  and  $r_2$  distributions (printed on the left panels) are very close to the input  $r_i^{true}$  values from which the toys were drawn: the relative difference is 0.2-0.9% in pp and 0.9-1.5% in PbPb, *i.e.* much smaller than the fit uncertainties and possibly caused by the low statistics of 300 toys. This proves the absence of significant bias in the fit convergence. The uncertainties on POIs in the nominal fit (Table 6.2) are also consistent with the distributions of the  $r_1$  and  $r_2$  uncertainties of the toys (middle and right panels). Lastly, the

---

<sup>1</sup>A 2D gaussian with a correlation  $\rho$  between the two variables is

$$\mathcal{G}(x, y) = \exp \left( -\frac{1}{2} \frac{1}{1 - \rho^2} \left[ \left( \frac{x - x_0}{\sigma_x} \right)^2 + \left( \frac{y - y_0}{\sigma_y} \right)^2 - \frac{2\rho(x - x_0)(y - y_0)}{\sigma_x \sigma_y} \right] \right)$$

When evaluating such a gaussian centered on the best fit point, and evaluated at the ( $x = r_1 = 0, y = r_2 = 0$ ) point, we obtain

$$-2 \ln \mathcal{G} = \frac{1}{1 - \rho^2} \left[ \frac{1}{\sigma_{1,metafit,rel}^2} + \frac{1}{\sigma_{2,metafit,rel}^2} - \frac{2\rho}{\sigma_{1,metafit,rel} \sigma_{2,metafit,rel}} \right]$$

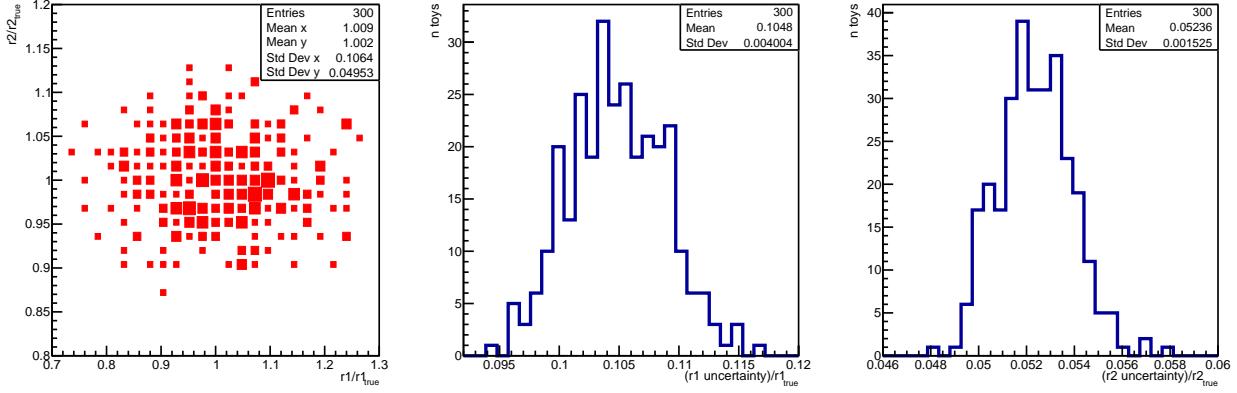


FIGURE 6.8: Results of the fits of 300 toys in pp. *Left*:  $r_1$  versus  $r_2$ . *Middle and right*: distributions of the fit uncertainties on  $r_1$  and  $r_2$ . The  $r_1$  and  $r_2$  axes have the same graduation width.  $r_i^{\text{true}}$  was set to 1.

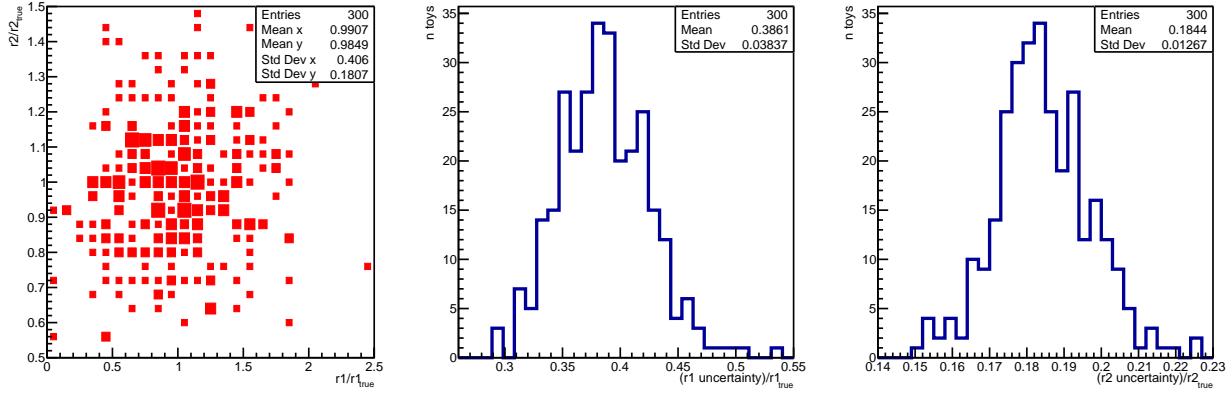


FIGURE 6.9: Results of the fits of 300 toys in PbPb. *Left*:  $r_1$  versus  $r_2$ . *Middle and right*: distributions of the fit uncertainties on  $r_1$  and  $r_2$ .  $r_i^{\text{true}}$  was set to 1.

correlation factor of the 2D POIs distribution is 0.10 (-0.03), which is close to the 0.10 (0.05) shown in Table 6.2 for pp (PbPb) collisions.

In PbPb, the uncertainty we had obtained from the partially blinded dataset was consistent with the spread of uncertainties from the toys (middle and right panels), which indicated that it was a good estimate of the unblinded fit uncertainty. Even after unblinding, it is interesting to interpret the RMS of the signal normalisation uncertainties of the toys as a variability (from 3 to 10% relative to the mean uncertainty) expected on the fit uncertainties that one would obtain by reproducing the experiment, meaning by re-running the LHC and CMS in similar conditions giving similar-sized datasets.

### 6.3.5 Fit checks

Various checks have been performed to check that the fit (and notably the multi-dimensional likelihood) is well-behaved.

First, the tool `ValidateDatacards.py` provided by `combine` has been run on all datacards and issued no warnings. Then, a fit on an Asimov toy (*i.e.* a toy that has exactly the content predicted by the model in each bin, removing any statistical fluctuations) with post-fit backgrounds, and signal normalisations fixed to 0, gave fitted normalisations  $r_1$  and  $r_2$  that are all  $< 10^{-3}$ , *i.e.* very close to 0 compared to our uncertainties. Similarly, fitting an Asimov toy generated with  $r_1 = 1$  and  $r_2 = 1$  gives fitted  $r_1$  and  $r_2$  values exactly equal to 1.

The tool `combineTool.py -M FastScan` has also been run to scan the shape of the NLL in each of its many parameters. The goal is mainly to check that the shape is approximately parabolic around the minimum (or at least convex, meaning the derivative of the NLL versus each parameter is monotonous) so that the minimisation does not run into a local minimum, or give pathologic uncertainties (due to incorrect behaviour of the NLL in the minimum  $\pm 1\sigma$  region). Most parameters have a NLL shape very close to parabolic. Some parameters, mostly the ones linked to shape morphings, show less conventional shapes, but are always convex around the minimum. There is one exception in PbPb for the parameter determining if the shape of the combinatorial J/ $\psi$ - $\mu$  background is better described by MC or the rotated-J/ $\psi$  sample; it shows a local minimum close to the true minimum, but with a very small barrier to reach the true minimum. We checked that the nominal fit reaches the correct minimum for this parameter.

The impacts of the nuisance parameters on the parameters of interest have been computed as well, with the post-fit parameter values. They consist in showing how much the variation of a given nuisance parameter within its post-fit uncertainties affects the values of the POIs – in other words, how much a nuisance parameter contributes to the fit uncertainties on POIs. The nuisance parameters having the largest impacts on  $r_1$  (which has larger uncertainties than  $r_2$ , so potentially a more abnormal behaviour) in pp and PbPb are shown in Figs. 6.10 and 6.11. Many parameters (among the 92 parameters in PbPb, and 136 in pp) have very small impacts and are not shown; this is especially the case of the parameters implementing the statistical uncertainties of the templates (whose names start with `prop_bin` in the figures) in the high-statistics background-enriched BDT bins.

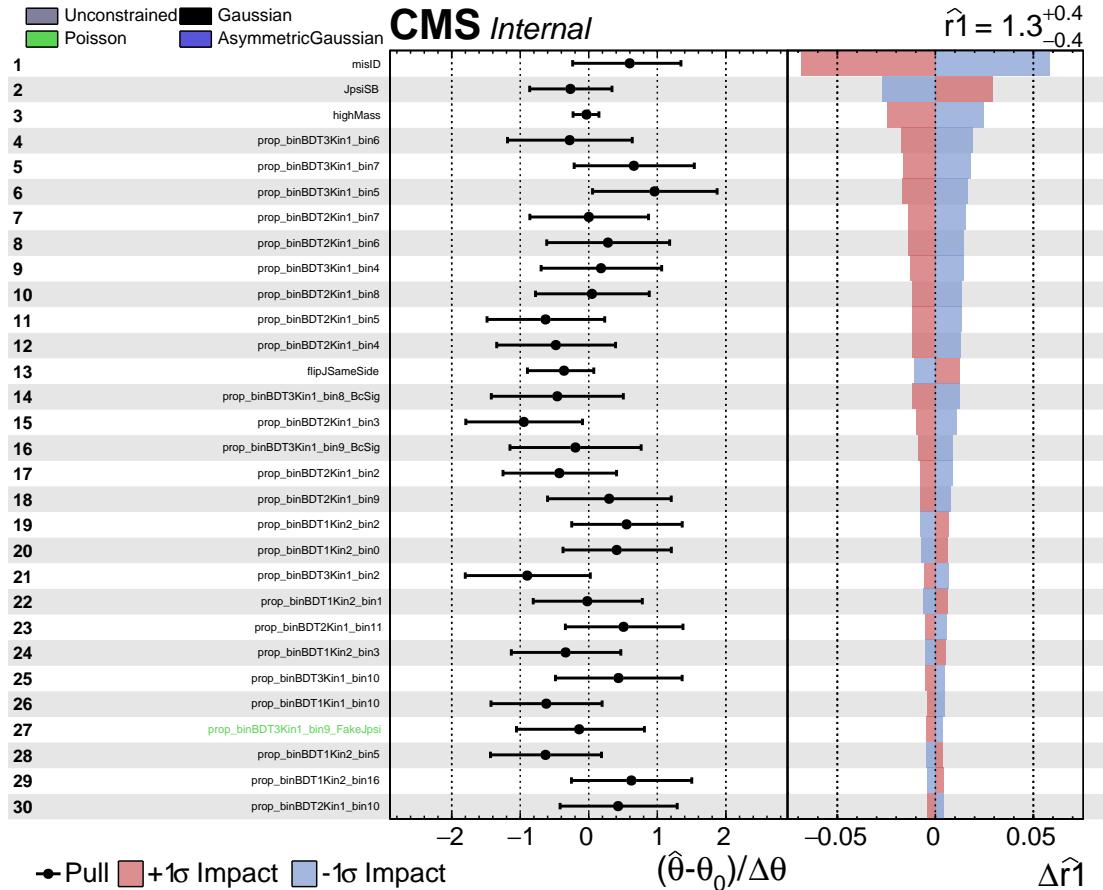


FIGURE 6.10: Impacts of the nuisance parameters on  $r_1$  (signal normalisation in the first  $p_T^{\mu\mu\mu}$  bin), in the pp second-step fit.

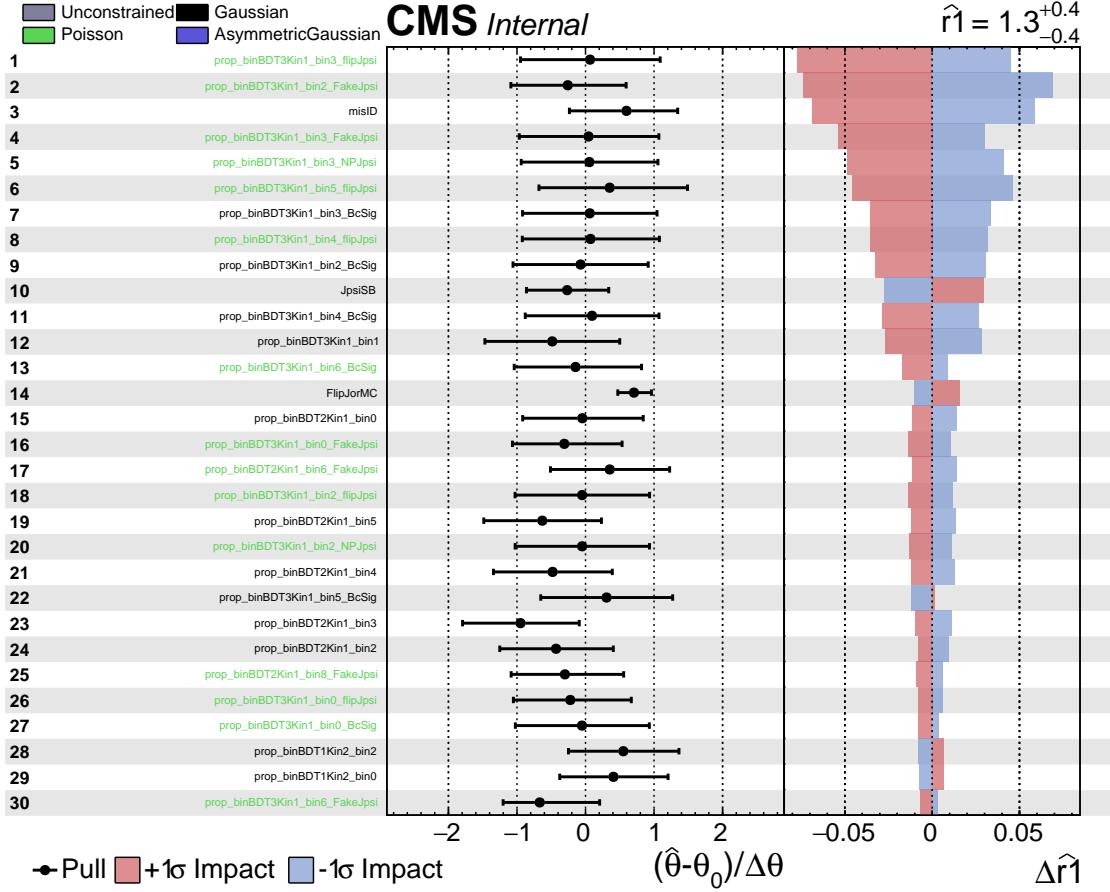


FIGURE 6.11: Impacts of the nuisance parameters on  $r_1$  (signal normalisation in the first  $p_T^{\mu\mu\mu}$  bin), in the PbPb second-step fit.

In PbPb, the parameters dealing with the statistical uncertainties of background templates have the largest impacts on POI uncertainties, especially those varying the very low statistics of background in the signal-enriched BDT bin. In particular, in the centre of the signal mass region, a modification of the background bin content is very correlated to a significant modification of the signal fraction. The normalisation of the  $B \rightarrow J/\psi X$  simulation (`misID`) also contributes significantly, because it is too small to be well-constrained by the fit in the background-enriched bins. The shape of the  $J/\psi$  sidebands (`JpsiSB`) and the shape of the combinatorial  $J/\psi - \mu$  background (`FlipJorMC`) have only the 10<sup>th</sup> and 14<sup>th</sup> largest impacts, respectively.

In pp, the  $B$  decays MC normalisation `misID` is largely the first contributor, followed by `JpsiSB` (2<sup>nd</sup> largest) and the normalisation of the rotated- $J/\psi$  sample (`highMass`, 3<sup>rd</sup> largest). Parameters implementing the statistical uncertainties of the templates in the signal-enriched BDT bins have the next largest impacts.

Some of the nuisance parameters for shape morphings and global background normalisations are significantly constrained by the fits, as can be seen in the middle column of the figures, which presents the ratio of the post-fit uncertainties (and values) compared to the pre-fit ones. This is nevertheless expected, considering the prior lack of knowledge on the backgrounds of this analysis, that we implemented in the fit procedure by leaving very loose constraints on the background shapes and normalisations. For example, the muon misidentification in MC (affecting directly the  $B$  decays simulated background) has not been measured in CMS in the conditions of this analysis. The combinatorial  $J/\psi - \mu$  background was also never studied before in these conditions.

## 6.4 Fit method variations

The fit provides uncertainties on the observed signal yields, but these depend on the assumptions of the fit model and method. The goal of this section is to vary these assumptions and run again the full fit on each variation, in order to extract a systematic uncertainty associated to the choice of the fit method. Multiple variations are presented here, some of which address difficulties with the fit method that were raised earlier. The post-fit signal yields for each method are then compared. Some variations are used only to check the fit method, and others are used to calculate the final systematic uncertainty. The central value of the observed yield is obtained with the nominal fit method presented until here (this is more stable than an average over some fit methods, which initially was a considered strategy). The multiple fit methods are run in both steps of the analysis. In the first step, it is only used for displaying the full uncertainties, as the central value used in the  $p_T$  spectrum fit of section 7.6 comes from the nominal method only. In the second step, it is used to determine the final fit method systematic uncertainty. The variations are run in  $p_T$  bins, in centrality bins, and for the integrated samples.

### 6.4.1 Mass and BDT binning

A first variation is to remove the first BDT bin (background-enriched) in all channels, to check that the background parameters still manage to stabilise only using the less-populated second BDT bin (which is harder due to lower stats). It also checks that the background shapes in the first BDT bin do not pull the parameters to a significantly different state than without the data of this bin.

Two other fit variations keep the first BDT bin but vary the limits of the BDT bins, such as to have approximately [20, 35, 45]% or [30, 45, 25]% of the signal MC sample in the [first, second, third] BDT bin (from less to more signal-pure, the nominal split being [25, 40, 35]%). This is particularly important for the first  $p_T^{\mu\mu\mu}$  bin in PbPb, where almost only the third BDT bin shows significant signal, to control that migrating an event between signal-pure and background-enriched bins does not modify its counting by the fit.

Two other fit variations use a different trimuon mass binning. In each BDT bin, the nominal binning is modified for the finer (coarser) binning by adding (removing) two bins in the mass signal region, and one bin in the high-mass control region; a minimum of one bin is kept in the control region.

### 6.4.2 BDT decorrelated from mass

As explained in section 5.7.5, one can decorrelate the BDT value from the trimuon mass, in order to leave more discriminative power to the mass during the template fit. A corrected BDT variable is used to distribute events in the BDT bins, so the input templates of the fit are significantly different. These fits are shown for the second-step in Figs. 6.12, 6.13 and 6.14, for the  $p_T^{\mu\mu\mu}$  and centrality dependences in pp and PbPb.

The background shapes indeed look different, but the fits are still good. It is reminded here that only the mean and RMS of the BDT distribution at a given mass are corrected, so, for example, some high-BDT tails can still be mass-dependent. In addition, only the total background is made mass-independent in average, not the separate background sources nor the signal.

Another variation consists in using this corrected BDT variable, but removing the first BDT bin from the fit input as in section 6.4.1.

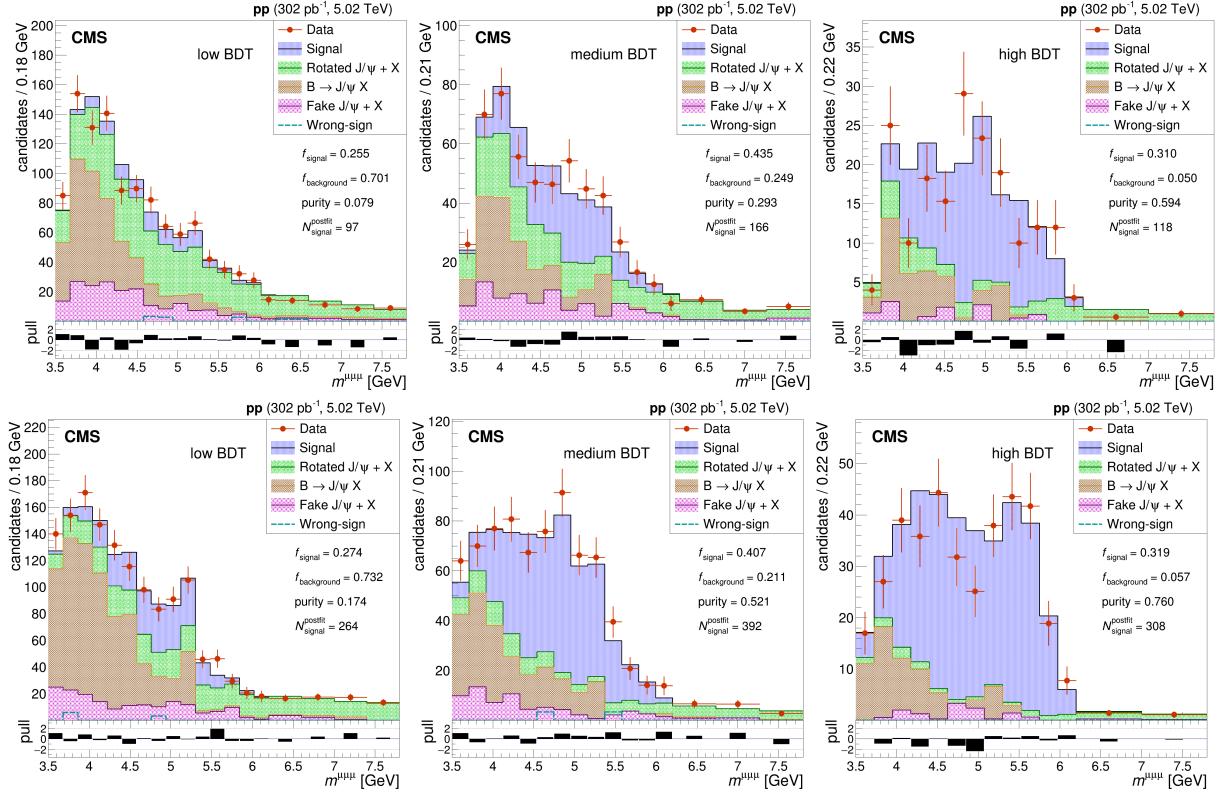


FIGURE 6.12: Template fit in pp, for the  $p_T^{\mu\mu\mu}$  dependence. The candidates are split in three bins of the mass-decorrelated BDT value. *Top*: candidates with  $6 < p_T^{\mu\mu\mu} < 11 \text{ GeV}$ . *Bottom*: candidates with  $11 < p_T^{\mu\mu\mu} < 50 \text{ GeV}$ .

### 6.4.3 Statistical uncertainties on templates, shape regularisation

Some background templates in high-BDT bins have very low statistics, which gives a lot of leverage in the minimisation to the parameters encoding the statistical uncertainties on the background content of each mass bin (as is made clear by the impacts of nuisance parameters shown in section 6.3.5). The role of these parameters is to take into account these low-statistics bins, but in practice, the large number of parameters and their large correlations can sometimes give rise to questionable behaviours in the nominal fit. Typically, some background bins with only a fraction of an expected event can surge to a few expected events to improve the agreement with a data fluctuation in this bin. This is seen to happen especially when the statistics are low enough that some background bins are empty but neighbouring bins are not (which is not expected of the true background for the smooth underlying distributions we are dealing with).

These oddities were quite worrying with partially blinded data, but they do fade with the smaller data fluctuations of unblinded data. I however check the effect of these parameters on the fitted signal yield by removing them in a fit variation, in the lowest-statistics BDT bin (*i.e.* in the third bin), or in the second and third BDT bins. When they are removed, the irregular shapes of the low-statistics backgrounds must be dealt with in another way; I choose to regularise them via a 3-bin floating average.

An input background shape in a given channel (*i.e.* BDT and analysis bin) is regularised if it does not contain any bin with relative error  $< 30\%$  and if it follows one of these conditions:

- All non-empty bins have a relative error  $> 50\%$ ;
- Or there is a bin that is empty or has  $> 95\%$  relative error, and there is at least one non-empty bin at lower mass *and* one at higher mass.

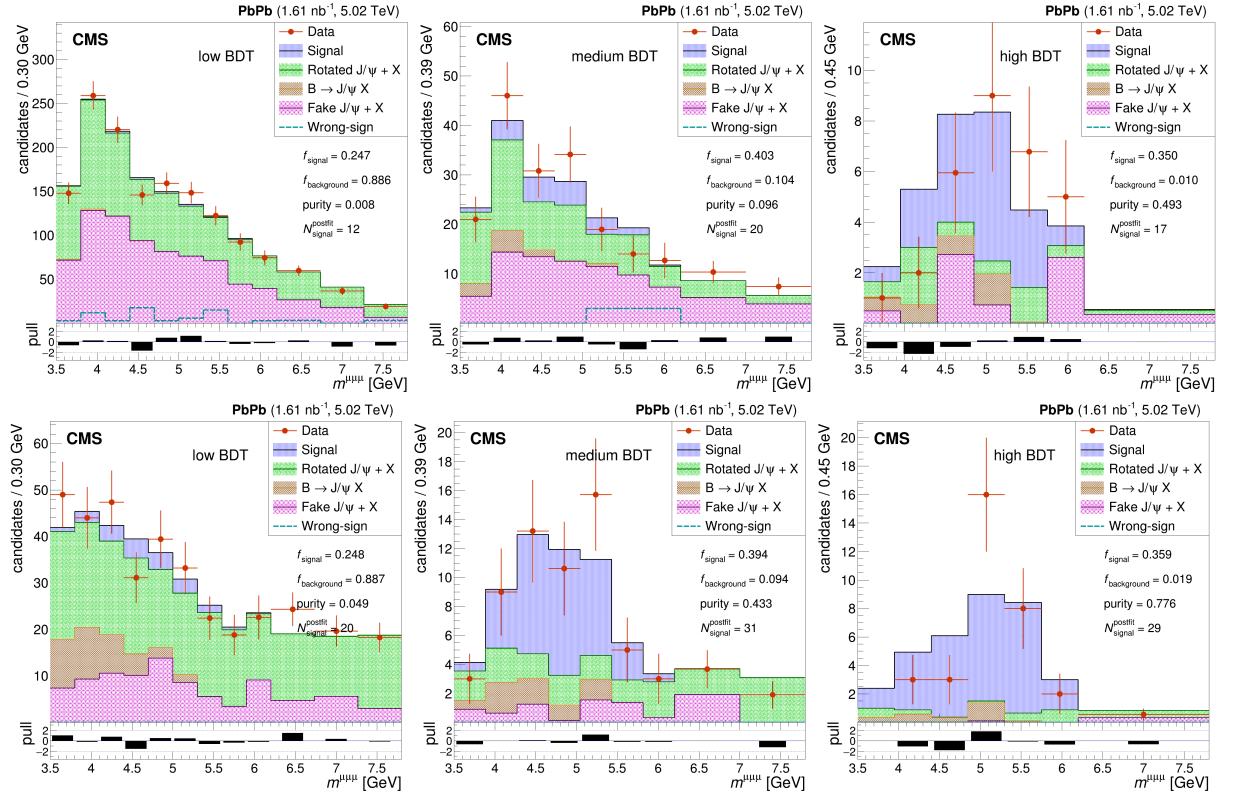


FIGURE 6.13: Template fit in PbPb, for the  $p_T^{\mu\mu\mu}$  dependence. The candidates are split in three bins of the mass-decorrelated BDT value. Top: candidates with  $6 < p_T^{\mu\mu\mu} < 11$  GeV. Bottom: candidates with  $11 < p_T^{\mu\mu\mu} < 50$  GeV.

If a template for a given background in a given channel follows these conditions, it is regularised with a 3-bin floating average. Calling  $b_i$  the content of bin number  $i$  and  $\sigma_i$  its uncertainty,  $b'_i$  and  $\sigma'_i$  the new bin content and uncertainty, and  $n$  the number of mass bins, the regularised histogram has:

- $b'_1 = \frac{2b_1+b_2}{3}$  and  $\sigma'_1 = \max(\sigma_1, \sigma_2)$
- $b'_n = \frac{2b_n+b_{n-1}}{3}$  and  $\sigma'_n = \max(\sigma_{n-1}, \sigma_n)$
- $b'_i = \frac{b_{i-1}+b_i+b_{i+1}}{3}$  and  $\sigma'_i = \max(\sigma_{i-1}, \sigma_i, \sigma_{i+1})$  for  $1 < i < n$ .

The new uncertainties are not set to the quadratic sum of the old ones because this would assume the bin contents are independent, which is not the case after the floating average is applied. The maximum of the old uncertainties on the averaged bins is an upper limit on the new uncertainty but not a precise estimate. The correlation between bins induced by this procedure implies that the parameters that were used to take into account independently the statistical uncertainty on each mass bin cannot be used. Fortunately, the aim of this fit variation is to remove those uncertainties.

Two variations are set: a fit with regularised shapes where the parameters for statistical uncertainties in each mass bin are removed (fixed to 0 in practice) for all shapes of the third BDT bin, and a fit where they are removed for shapes of the second and third BDT bins. The regularisation indeed mostly happens for shapes in the third BDT bin, so the statistical uncertainty parameters in the first and/or second BDT bins can be kept.

When varying the shape of a background with a shape morphing parameter, as is explained in section 6.2, the varied background histogram is normalised to the yield of the nominal background

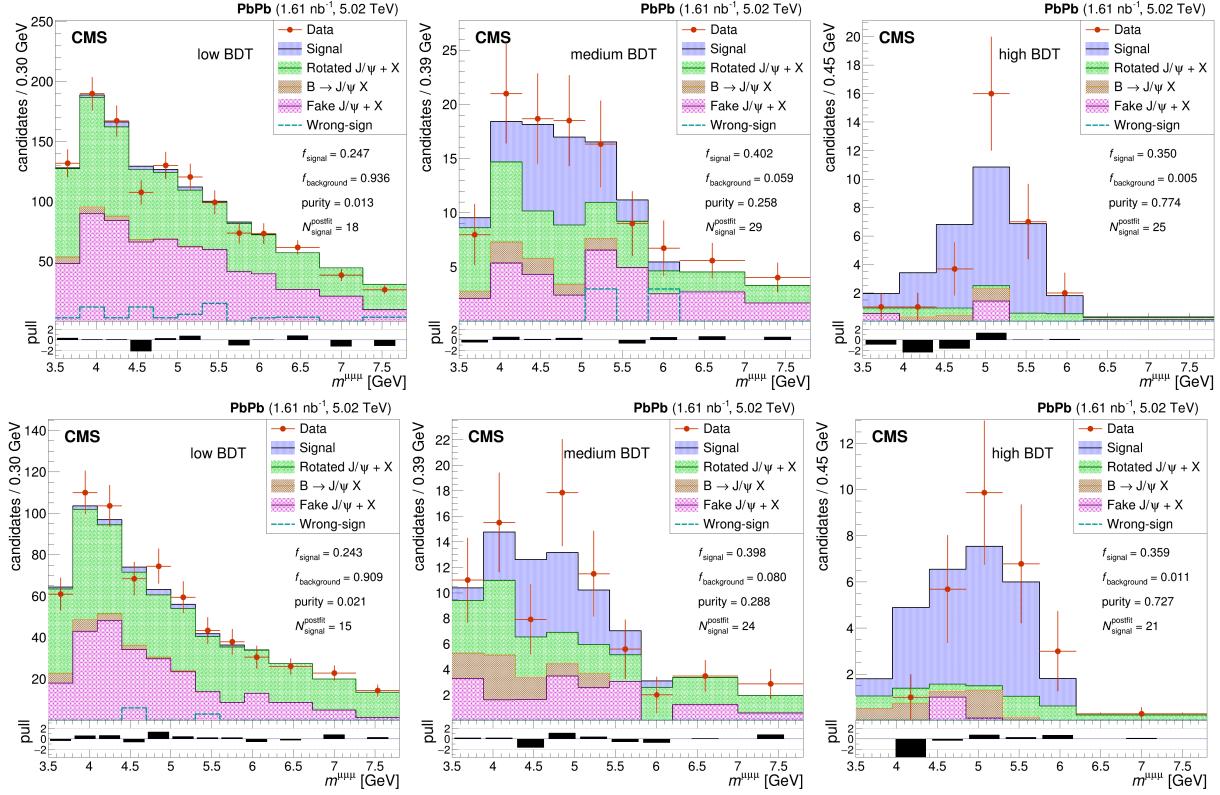


FIGURE 6.14: Template fit in PbPb, for the centrality dependence. The candidates are split in three bins of the mass-decorrelated BDT value. *Top:* candidates with centrality in [0, 20]%. *Bottom:* candidates with centrality in [20, 90]%. *Bottom:*

histogram in each channel, meaning *in each BDT bin* (and each analysis bin). This is meant to stabilise the nominal fit, by limiting the correlations between the shape morphing parameters and the template-statistics parameters (which can modify the integral of a background histogram). However, if I remove the most sensitive of those template-statistics parameters, I can leave more freedom to the shape variations, by normalising only their yield *integrated over the BDT bins* to the yield of the nominal shape, in each analysis bin.

I thus make two more variations of the fit method, in which the shape variations are normalised to the nominal shape through yields integrated over BDT bins, and in which the template-statistics parameters are removed either in the third BDT bin, or in the second and third BDT bins. When these parameters are removed, the low-statistics shapes need some regularisation, so I apply as well the 3-bin floating average for these fit variations. This way of normalising the shape variations gives more freedom to the fit, by giving the shape morphing parameters a larger impact on the background templates: they can indeed, in this case, change how the yields of a background are distributed in the BDT bins.

#### 6.4.4 Summary and calculation of systematic uncertainty

The previous sections presented 11 variations of the fit method, which are summed up in Table 6.3.

TABLE 6.3: Summary of the 11 fit method variations. See the text for details. The last column says if this variation is counted in the RMS calculation of the systematic uncertainty or is only used as a check.

Type	Summary	Details	Check or Systematic?
<b>Mass binning</b>	coarser	-2 bins in mass signal region -1 bin in high-mass control region	check
	finer	+2 bins in SR, +1 bin in CR	check
<b>BDT binning</b>	wider signal bin	[20, 35, 45)% signal efficiencies in the [1st,2nd,3rd] BDT bins	check
	thinner signal bin	[30, 45, 25)% signal efficiencies	check
<b>BDT-mass decorrelation</b>	no BDT bin 1	remove the (background-enriched) BDT bin 1 from the fit	syst.
	decorrelation	use a BDT variable decorrelated from mass	syst.
<b>Template statistical uncertainties</b>	decorrelation + no BDT 1	idem + remove BDT bin 1	syst.
	<b>regularisation</b> + no stat. parameters for BDT3	regularisation of low-stats shapes + no stat. parameters in BDT bin 3	syst.
<b>Template statistical uncertainties</b>	<b>regularisation</b> + no stat. parameters for BDT2-3	regularisation + no stat. parameters in BDT bins 2 and 3	syst.
	<b>BDT-integrated norm.</b> + no stat. parameters for BDT3	regularisation + no stat. parameters in BDT bin 3 + systematic shape variations normalised with BDT-integrated yields	syst.
	<b>BDT-integrated norm.</b> + no stat. parameters for BDT2-3	regularisation + no stat. parameters in BDT bins 2 and 3 + systematic shape variations normalised with BDT-integrated yields	syst.

Let us now look at how much the post-fit yields vary with the fit method compared to the nominal yields, for pp, PbPb, and for the ratio of PbPb yields to pp yields (equivalent to  $R_{\text{PbPb}}$ , because the additional normalisation factors are not related to the fit method). The latter allows the uncertainty to partially cancel in the  $R_{\text{PbPb}}$  thanks to partial correlations between pp and PbPb. Some of the fit methods are taken only as checks, and the others will enter the root-mean-square (RMS) calculation of the systematic uncertainty associated to the fit method.

The nominal method presented in sections 6.1 to 6.3 provides the central value and the fit uncertainty. A strategy with an average over the methods that were judged equivalently good was tested but not kept in the final strategy.

Fig. 6.15 shows the post-fit yield and associated fit uncertainty, divided by the yield from the nominal method, for each method variation including the nominal one. It is shown in parallel for each  $p_T^{\mu\mu\mu}$  or centrality bin and for the integrated sample, and in pp, PbPb, and for the

ratio of PbPb to pp yields. Similarly, Fig. 6.16 shows the yields for all fit method variations in  $p_T^{\mu\mu\mu}$  bins, in pp and PbPb, after correction by the one-binned acceptance $\times$ efficiency (explained in section 7.3) obtained from the signal MC after the first-step (and before the second-step)  $p_T^{\mu\mu\mu}$ -spectrum correction (see section 7.6).



FIGURE 6.15: Post-fit yield and its uncertainty, divided by the nominal post-fit yield, for all fit methods. The pp and PbPb yields as well as the PbPb/pp ratio are shown, for both  $p_T^{\mu\mu\mu}$  bins (*top*), for both centrality bins (*middle*, where the integrated pp fit used as denominator of  $R_{\text{PbPb}}$  is also shown), and for the fit integrated on analysis bins (*bottom*). For each bin, a vertical dotted line at 1 indicates the nominal value. The horizontal lines group similar variations. The names of all variations are mentioned in the left column. The orange variations are included in the uncertainty calculation; the others, in violet, are only used as checks.

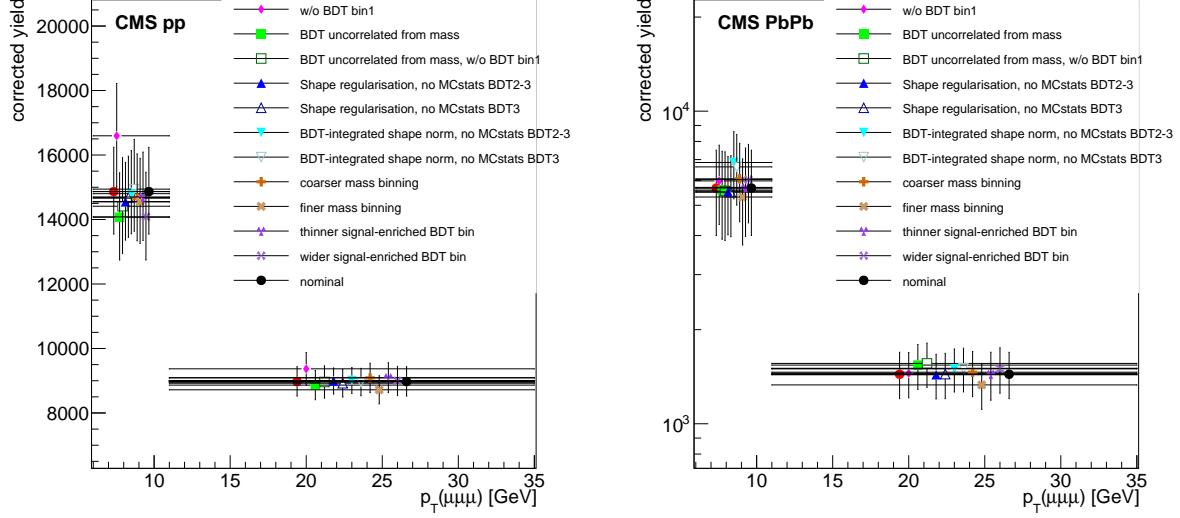


FIGURE 6.16: Post-fit yields corrected by the one-binned acceptance×efficiency corrections (after the first-step  $p_T^{\mu\mu\mu}$  correction of the signal MC) as a function of the trimuon  $p_T^{\mu\mu\mu}$ , in pp (left) and PbPb (right). All fit methods are compared. The markers for each method are spread on the  $p_T^{\mu\mu\mu}$  axis to avoid superposition.

The variations changing the granularity of the mass binning or the limits of the signal-enriched BDT bin are only considered as checks of the good behaviour of the fit, because there are no strong reasons for these to affect the fit convergence. It is checked that in all bins the POIs  $r'_1$  and  $r'_2$  from these methods all deviate from the nominal values  $r_1$  and  $r_2$  by less than  $0.4\sigma$  in terms of the fit uncertainties of the nominal and varied POIs. More explicitly,  $\frac{|r'_i - r_i|}{\sqrt{\sigma'^2_i + \sigma^2_i}} < 0.4$  for each POI of each of these four variations, with asymmetric uncertainties (if  $r'_i < r_i$ , the high uncertainty is taken for  $\sigma'_i$  and the low one for  $\sigma_i$ , and vice versa). However, the fit uncertainties are close to fully correlated between the methods

To obtain the relative systematic uncertainty from the fit method, I start by getting the maximal deviation to the nominal value, separately in each of these three blocks of variations: the removal of the first BDT bin, the two variations linked to the BDT-mass decorrelation, and the four variations dealing with the statistical uncertainties of the templates. The variations in each block are considered very correlated, as they come from a similar source. Then the final uncertainty is the RMS over the maxima of the  $n = 3$  blocks (hence covering all the 7 variations labelled with ‘syst.’ in Table 6.3). More explicitly, it is:

$$\sigma_{syst,rel} = \sqrt{\frac{\sum_n (y_i/y_{nom} - 1)^2}{n}}$$

where  $y_i$  is the post-fit yield of block  $i$  that deviates the most from the nominal, and  $y_{nom}$  is the yield from the nominal fit. The denominator is  $n$  and not  $n - 1$  because the RMS is calculated compared to a given method and not to the average of the sampled points. This uncertainty is symmetrised (*i.e.* both the up and down uncertainty are taken equal to the RMS). A similar calculation gives the correlation factor between the two analysis bins ( $b = 1$  or  $2$ ) for this systematic uncertainty:

$$\rho_{1-2} = \frac{\sum_n (y_i^{b=1}/y_{nom} - 1)(y_i^{b=2}/y_{nom} - 1)}{n \sigma_{syst,rel}^{b=1} \sigma_{syst,rel}^{b=2}}$$

This correlation can also be seen graphically in Fig. 6.15 when, in multiple methods, the yields deviate in the same direction for both analysis bins.

The resulting systematic uncertainties and their correlations between the two analysis bins are shown in Table 6.4.

TABLE 6.4: Relative systematic uncertainties from the choice of fit method in all analysis bins, on pp and PbPb yields, and on the PbPb/pp yield ratio. The correlation factors between the two  $p_T^{\mu\mu\mu}$  or centrality bins are also quoted.

	pp	PbPb	PbPb/pp
$6 < p_T^{\mu\mu\mu} < 11 \text{ GeV}$	7.5%	12.4%	<b>12.9%</b>
$11 < p_T^{\mu\mu\mu} < 35 \text{ GeV}$	2.7%	5.5%	<b>5.9%</b>
correlation $\rho_{1,2}$	0.97	0.30	0.57
centrality 0-20%	—	6.8%	<b>5.5%</b>
centrality 20-90%	—	10.3%	<b>9.4%</b>
correlation $\rho_{1,2}$	—	0.37	0.42
integrated	2.3%	9.7%	8.2%



“Le premier [principe] était de ne recevoir jamais aucune chose pour vraie que je ne la connusse évidemment être telle. [...]”

Le dernier [principe était] de faire partout des dénombrements si entiers et des revues si générales, que je fusse assuré de ne rien omettre.”

---

René DESCARTES, *Discours de la méthode*,  
Deuxième partie, 1637.<sup>1</sup>

## Chapter 7

# How much is produced? Acceptance and efficiency

The observed signal yields in each analysis bin were obtained in the previous chapter. The next step is to correct these yields by the acceptance and efficiency of the detector to reconstruct and select trimuons that come from a signal  $B_c$ . This will provide the true number of  $B_c$  mesons produced in the pp or PbPb collision, in the trimuon ( $p_T^{\mu\mu\mu}$ ,  $y^{\mu\mu\mu}$ ) phase space defined in section 5.6 (the so-called ‘fiducial cuts’ where CMS is able to reconstruct  $B_c$  mesons). These corrected yields will then only need to be scaled by the (equivalent) luminosities (see section 3.5.5) to obtain the cross sections and nuclear modification factor.

To evaluate the acceptance and efficiency, we mainly rely on the signal  $B_c$  MC, whose production is detailed in section 3.5.2. The core of the calculation of the acceptance  $\alpha$  and the efficiency  $\varepsilon$  is always:

$$\alpha = \frac{\#\text{accepted}}{\#\text{generated}} \quad , \quad \varepsilon = \frac{\#\text{passing cuts}}{\#\text{accepted}} \quad (7.1)$$

where a generated trimuon comes from a true simulated  $B_c$ , an accepted trimuon is a generated trimuon that has its three muons in the kinematic single-muon acceptance (defined in section 5.3.2), and a passing trimuon passes all the single-muon reconstruction, identification, and trigger requirements, as well as the trimuon-specific selection cuts (chapter 5).

Let us consider a true  $B_c$  meson of given kinematics (which will be the input of the acceptance and efficiency maps of section 7.4). The acceptance of this  $B_c$  depends only on the kinematics of its true generated trimuon, which are due to EVTGEN. It simulates the  $B_c$  decay into three muons and a neutrino; but section 3.5.2 mentions that the trimuon kinematics do not change significantly (for the precision of this analysis) with different decay models. This gives confidence

---

<sup>1</sup>Translation by Ian JOHNSTON: “The first rule was that I would not accept anything as true which I did not clearly know to be true. [...] And the last was to make my calculations throughout so complete and my examinations so general that I would be confident of not omitting anything.”. The word ‘dénombrements’ translated here as ‘calculations’ could also be translated as ‘enumerations’ or ‘counts’.

in the acceptance value for a  $B_c$  of given kinematics. On the contrary, the simulated efficiencies must be checked against data, and corrected with the associated scale factors; this is done in section 7.1 for single muons with a tag-and-probe technique based on the  $J/\psi$  resonance. The pp scale factors, that I extracted, are more detailed than the PbPb ones, that were extracted for multiple analyses by a collaborator. The application of these corrections is not straightforward due to asymmetric cuts on the three muons, and section 7.2 explains the related strategy. These scale factors correct only for the single-muon efficiencies; the efficiency of the trimuon-specific cuts is indirectly checked by the background studies in chapter 4 and by the check of the total BDT distribution in section 5.7.4.

Two correction strategies were initially envisioned. The first (simplest) one is ‘one-binned’: it directly measures the acceptance and efficiency on the signal MC sample, providing a simple  $\alpha \times \varepsilon$  number for each analysis bin in pp and PbPb. Section 7.3 treats this method. However, it relies on the kinematic distributions of the  $B_c$  mesons simulated in MC. It is challenging to check them precisely in the pp signal MC, and they are not simulated in PbPb. And, as the acceptance and efficiency depend strongly on the kinematics, wrong simulated distributions can affect a lot the acceptance and efficiency values integrated over an analysis bin. This could in principle be circumvented by building acceptance and efficiency maps (computed in section 7.4), and applying event-by-event corrections to each observed trimuon based on these maps. The method is explained in section 7.5 along with its limitations: the needed background subtraction is not straightforward, which participates in increasing the statistical uncertainties to a stage where the PbPb results are unreliable.

I hence turned to a two-steps procedure, that makes the ‘one-binned’ method more reliable by roughly correcting the  $p_T^{\mu\mu\mu}$  distribution in the signal MC with a first iteration of our  $p_T^{\mu\mu\mu}$ -differential cross-section measurement; this avoids the complications and very large uncertainties of the event-by-event corrections. Improving the acceptance and efficiency ( $\alpha \times \varepsilon$ ) estimation is the main motivation for correcting the  $p_T^{\mu\mu\mu}$  distribution of signal MC before running a second-step of the analysis (see the analysis flowchart in Fig. 3.13), but this correction also improves the BDT and template fit. After the second-step fit and fit method variations, the nominal yield is corrected by a one-binned  $\alpha \times \varepsilon$  from the first-step-corrected MC, and this corrected yield is used to perform a second-step correction of the  $p_T^{\mu\mu\mu}$  distribution of MC. The resulting corrected MC is used to estimate the final  $\alpha \times \varepsilon$  correction (that could be called third-step). Variations of the  $p_T^{\mu\mu\mu}$  distribution correction of MC (within the uncertainties of the second-step measurement) provide the systematic uncertainty on  $\alpha \times \varepsilon$ . The fitting of the  $p_T^{\mu\mu\mu}$  distribution, the associated correction of signal MC, and the estimation of the final  $\alpha \times \varepsilon$  and its uncertainty are explained in section 7.6.

## 7.1 Tag-and-probe corrections to single-muon MC efficiencies

### 7.1.1 Tag-and-probe principle

The reliance of the simulated efficiencies of reconstruction, selection, and triggering of physics objects is key to many data analysis. A perfect description of the detector and how particle interact with it is impossible in MC. *Tag-and-probe* (TnP) methods can compensate for these imperfections, by deriving a data-driven correction of the MC efficiencies.

The TnP method is based on a known decay of mass resonance ( $J/\psi$  in this case, but  $Z$  bosons are used for high- $p_T$  muons). Among the two particle candidates from this decay (muons in this case), a *tag* and a *probe* are designated, and the efficiency of a particular selection criterion is evaluated on the probe. In general, the tag needs to pass a tight selection criteria, so that they very rarely misidentified ( $\ll 1\%$ ). The probe usually endures a much looser selection criteria. Probes are selected by pairing them with tags such that the invariant mass of the combination is consistent with the mass of the studied resonance.

The invariant mass distribution of the tag-probe pairs is fitted for three cases: ‘all pairs’, ‘failed pairs’ and ‘passing pairs’, according to whether the probe passes or fails the selection criteria whose efficiency is tested. The  $\text{J}/\psi$  yields are fitted in each of the three cases and the efficiency is the ratio of the fitted passing pairs to all pairs. The efficiency is calculated in this way for various kinematic ranges ( $p_{\text{T}}$  and  $\eta$ ) of the probe as well as in bins of the event activity (pile-up in pp or centrality in PbPb). The TnP efficiencies in data and MC are compared and their ratios give the scale factors correcting the MC muon efficiencies.

### 7.1.2 Overview in pp and PbPb

The scale factors used in this analysis (and extracted in section 7.1.3 for pp) correct the efficiency of muons identified with the *hybrid-soft* selection defined in section 5.3.1, which requires a *global* muon. These scale factors are used by multiple analyses studying  $\text{J}/\psi$  mesons in CMS. The scale factors for the *soft* muon identification (based on tracker muons) are not calculated in pp since no analysis plans on using this identification in the pp 5.02 TeV 2017 dataset, but they would be too challenging. However, scale factors for tracker muons (with *soft* ID) have never been achieved, which is a main argument against using this identification for one of the three muons from the  $B_c$  decay (see section 5.3.3). This is attempted in a fast study in section 7.1.5, yielding too large uncertainties for *soft* ID to be a viable option.

The *tight* kinematic acceptance cuts of single muons (defined in section 5.3.2) is applied for the efficiencies of reconstruction, hybrid-soft, and triggering of the muon that are required to fire the double muon trigger defined in section 3.5.1. The *loose* kinematic acceptance of muons (cf section 5.3.2) is used for the reconstruction and *hybrid-soft* identification of muons that are not required to fire the trigger.

Section 7.1.3 details the extraction of the TnP scale factors for the MC efficiencies of low- $p_{\text{T}}$  muons with the hybrid-soft identification in pp 2017 data at 5.02 TeV. The existing PbPb scale factors will directly be used, and shortly explained in section 7.1.4.

Each source of muon inefficiency is assumed to be independent, so that the associated efficiencies can be factorised. In general, the inefficiency is separated in two identification steps (the reconstruction, and the offline selection) and the trigger step, but the classification is in practice more subtle. Separate scale factors are calculated for each efficiency.

### 7.1.3 Single-muon scale factors for pp 2017

#### 7.1.3.1 General settings

All results and plots can be found at [https://gfalmagn.web.cern.ch/gfalmagn/cms\\_private/pp2017\\_5TeV\\_Muon\\_scaleFactors/](https://gfalmagn.web.cern.ch/gfalmagn/cms_private/pp2017_5TeV_Muon_scaleFactors/); only a subset is shown here. A header file containing the corrections to be applied on the single muon efficiencies in MC, depending on  $|\eta|$  and  $p_{\text{T}}$ , is given in [https://github.com/gfalmagn/MuonAnalysis-TagAndProbe/blob/92X/test/results/tnp\\_weight.h](https://github.com/gfalmagn/MuonAnalysis-TagAndProbe/blob/92X/test/results/tnp_weight.h), and usage instructions are in [https://twiki.cern.ch/twiki/pub/CMS/HIMuonTagProbe/TnP\\_pp\\_summary\\_12.07.2019.pdf](https://twiki.cern.ch/twiki/pub/CMS/HIMuonTagProbe/TnP_pp_summary_12.07.2019.pdf). These results are planned to be published along with other muon performance studies in heavy ions in CMS (including the PbPb 2018 low- $p_{\text{T}}$  muon scale factors).<sup>1</sup>

#### Factorisation of efficiencies

The total muon efficiency is factorised in three pieces, for which the ‘all probes and ‘passing probes’ requirements are described in Table 7.1, along with the main parameters of the procedure. The tag cuts are the same in all cases: a high-quality muon, passing all analysis cuts (acceptance

---

<sup>1</sup>Internal reference in CMS: MUO-21-001 for the future paper, and AN-18-316 for the analysis note (<https://gitlab.cern.ch/tdr/notes/AN-18-316>).

and muon ID selection), and which is matched to a single muon trigger. Events firing this trigger are stored in another dataset than with the double muon trigger, ensuring a small overlap – a few % – between the two samples (the first being used for TnP, the second for the analysis). The tight single muon trigger requirement also ensures that the tag muon will always be able to help firing the double muon trigger, along with the probe whose trigger efficiency is tested.

Name	Tracking	Glb	MuID	MuIdTrg
<b>Efficiency type</b>	inner-tracking in. track-standalone matching	global muon reconstruction	hybrid-soft ID	hybrid-soft ID+trigger
<b>Probe selection</b>	standalone muon # valid muon stations > 1 tight acceptance	inner track <i>not</i> from a muon seed # of meas. tracker layers > 5 $d_{xy} < 0.3$ cm tight acceptance	global loose acceptance	global hybrid-soft ID tight acceptance
<b>Passing Probe</b>	global	global	hybrid-soft ID	L1DoubleMu0 trigger
<b>SF binning</b>	No correction	$p_T$ in $\eta$ bins	$p_T$ in $ \eta $ bins	$p_T$ in $ \eta $ bins
<b>Nominal fit</b>				
signal PDF	Triple Gaussian	Double Gaussian	CrystalBall+Gaussian	CrystalBall+Gaussian
bkg PDF	2nd order polynomial	2nd order polynomial	1st order polynomial	1st order polynomial
mass range [GeV]	[2.0,5.0]	[2.6,3.5]	[2.6,3.5]	[2.6,3.5]

TABLE 7.1: Overview of the settings for the separate efficiencies studied in this section.

The factorised efficiency chain for triggering muons features **Glb** and **MuIdTrg**, which give respectively the efficiencies for global muon reconstruction (starting from good-quality tracks), and for the hybrid-soft identification and double muon trigger, all using the tight muon acceptance. The chain for muons not required to trigger features **Glb** and **MuID**, giving respectively the global reconstruction and hybrid-soft identification efficiencies, both in the loose muon acceptance.

The probe of the first factorised efficiency has to be a reconstructed object in the detector (in this case, a general track), and this reconstruction has an inefficiency that needs to be checked as well. That is why a fourth **Tracking** efficiency, not following the factorisation paradigm, is also studied. The derived scale factors should be close to 1, as they will cannot be used to correct data: some inefficiencies would be double-counted.

### Data and MC samples

For data, the primary datasets used are `/SingleMuon/Run2017G-17Nov2017-v1/AOD` and `/SingleMuonTnP/Run2017G-17Nov2017-v1/AOD`, using the TnP-specific single-muon triggers. The first contains all events passing high- $p_T$  triggers, whereas the second contains a small fraction of the events from low- $p_T$  single-muon triggers. After processing, the two datasets are combined (removing duplicates). As for the main dataset (see section 3.5.1), only the parts of the run where all muon detectors were functional are kept. The centrally produced prompt J/ $\psi$  MC samples (see section 3.5.2) are used.

The following sections (7.1.3.2 to 7.1.3.5) detail the extraction of the four categories of scale factors of Table 7.1, and section 7.1.3.6 shows a closure test of the procedure in MC.

#### 7.1.3.2 Efficiency of global muon reconstruction

This section calculates the efficiency and scale factor for reconstructing a global muon track from a *general track* (*i.e.* any track in the inner track). The probe inner track is requested not to be reconstructed starting from a track seed in the muon chambers. To reduce the large track background, two cuts from the hybrid-soft selection (that proved best to remove this background) are also applied to the probes (see Table 7.1). These cuts were shown to modify

the MC efficiencies by less than 0.1% in all bins, which is negligible compared to the final uncertainties. Finally, the probes need to pass either the tight or the loose muon kinematic acceptance (defined in section 5.3.2), to be applied respectively to triggering or non-triggering muons. Only the plots for the tight acceptance are shown for brevity.

The probe is considered to pass if it is also reconstructed as a global muon, and to fail otherwise. A simultaneous fit of the dimuon (tag-probe) invariant mass around the  $J/\psi$  peak is done in the all/pass/fail categories for data and MC, in different bins of  $\eta_{\text{probe}}$  and  $p_T, \text{probe}$ . The signal shape is modelled by a sum of two gaussians of common mean. The background is described by a second order polynomial. An example of this fit, for the integrated sample, is shown in Fig. 7.1. The integrated efficiencies for data and MC are 90.6% and 90.0%, respectively.

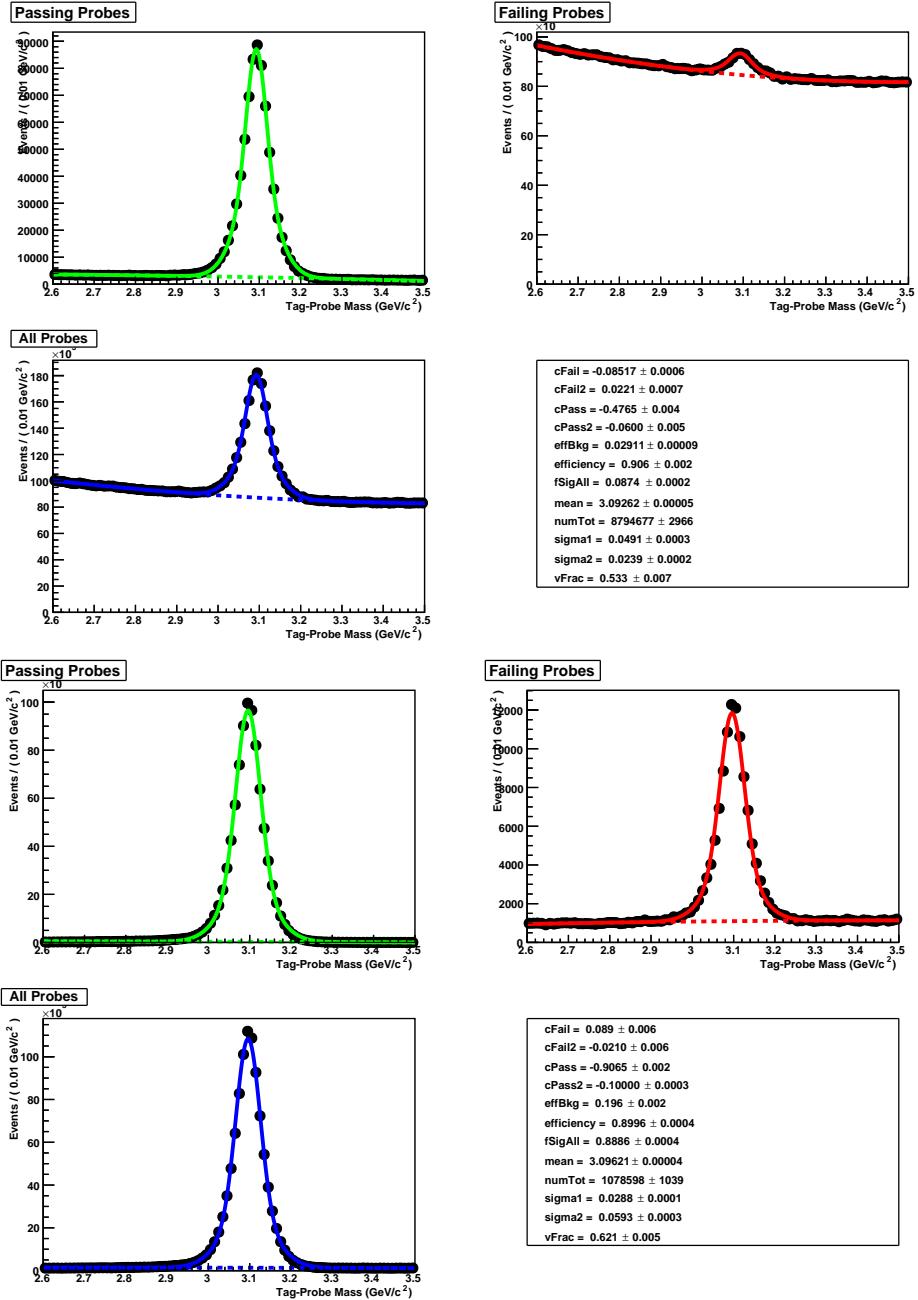


FIGURE 7.1: Dimuon mass fit of the global muon reconstruction efficiencies, integrated over kinematics, in data (top) and MC (bottom).

The efficiencies are calculated in  $\eta$  bins, shown in Fig. 7.2. As the dependence is found to be weak and as the statistical uncertainty start to dominate if the measurement is too differential, only three bins in  $|\eta|$  are considered ( $[0, 1.2]$ ,  $[1.2, 1.8]$ , and  $[1.8, 2.4]$ ), compared to four in the following sections. The  $p_T$  dependence is then shown in Fig. 7.3, separately for the three  $|\eta|$  bins.

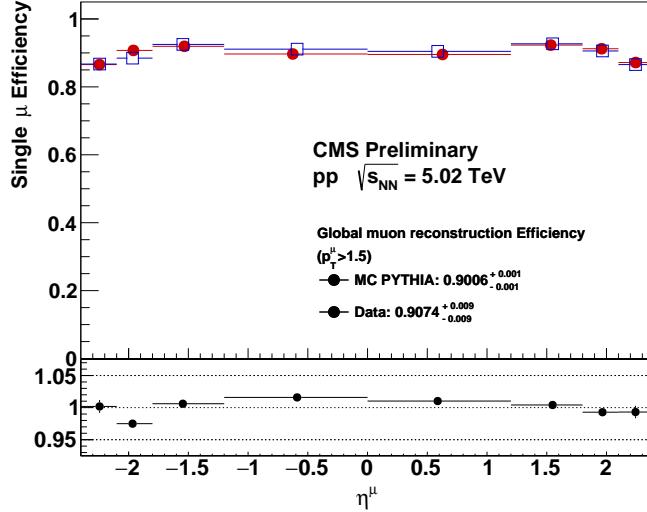


FIGURE 7.2: Global muon reconstruction efficiencies for data and MC, versus pseudorapidity. The bottom panel shows the ratios of data over MC efficiencies, *i.e.* the scale factors.

The scale factors (ratio between data and MC efficiencies) are also shown in the lower panel of these plots. The scale factors from Fig. 7.3, binned in  $(|\eta|, p_T)$ , will be applied as weights on the muon efficiency in MC; their deviation from 1 is mostly below 2% for  $|\eta| < 1.8$  and below 4% for  $1.8 < |\eta| < 2.4$ .

The systematic uncertainties on these scale factors are based on three variations of the fitting procedure. The signal shape is changed to a sum of a Gaussian and a Crystal Ball<sup>1</sup> of common mean, or the background shape is changed to a third order polynomial, or the mass range used in the fit is changed from  $[2.6, 3.5]$  GeV to  $[2.8, 3.4]$  GeV. Binned scale factors are calculated separately for these three alternative methods, and compared to the nominal results. Fig. 7.4 shows this comparison.

In a given bin, the statistical uncertainty on a scale factor is a simple error propagation from the uncertainties on the fit efficiency parameter for data and MC; for a ratio of uncorrelated quantities, this is the quadratic sum of the relative errors. As the variations from the change of fit methods are reasonable, the systematic uncertainties on the scale factors are calculated in each bin as the largest deviation from the nominal result. To be conservative, the largest deviation is symmetrised (*i.e.* it gives both the up and down uncertainties).

A possible dependence on pile-up has also been studied. The efficiency as a function of the number of reconstructed primary vertices is shown in Fig. 7.5, for data and MC. No dependence is observed.

### 7.1.3.3 Efficiency of muon identification

The efficiencies and scale factors calculated in this section concern the hybrid-soft selection, starting from a global muon. The probes need to pass the loose muon kinematic acceptance (see

<sup>1</sup>A Crystal Ball is a combination of a Gaussian (of parameters the mean and width), and a power law (of exponent  $n$ ) below some threshold ( $\alpha$ ) below the mean. See [https://en.wikipedia.org/wiki/Crystal\\_Ball\\_function](https://en.wikipedia.org/wiki/Crystal_Ball_function).

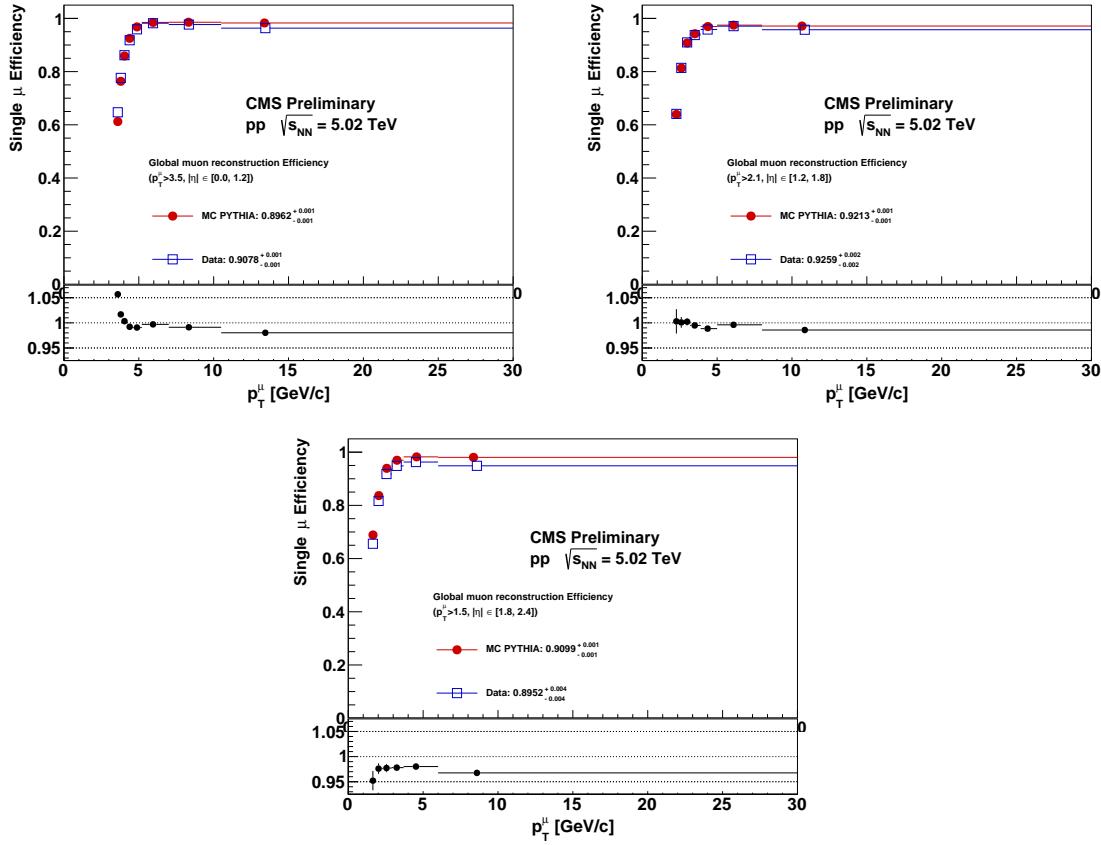


FIGURE 7.3: Global muon reconstruction efficiencies for data and MC versus  $p_T$ , in the three  $|\eta|$  bins. The bottom panels show the ratios of data over MC efficiencies, *i.e.* the scale factors.

section 5.3.2), meaning that the scale factors derived here should be used only for muons that are *not* required to fire the trigger, and are selected with this acceptance.

The global muon probe is considered to pass if it satisfies the hybrid-soft ID, and to fail otherwise. A simultaneous fit of the dimuon invariant mass is performed in the all/pass/.fail categories for data and MC, in different bins of  $\eta_{\text{probe}}$  and  $p_T, \text{probe}$ . The signal shape is modelled by a sum of a Crystal Ball and a Gaussian of common mean. The background is described by a first order polynomial. An example of this fit, for the integrated sample, is shown in Fig. 7.6. The integrated efficiencies for data and MC are 98.35% and 98.74%, respectively.

The  $p_T$  dependence is shown in Fig. 7.7, separately in four  $|\eta|$  bins (of limits 0, 1.2, 1.8, 2.1, 2.4). The  $\eta$  dependence is not shown, for brevity. The scale factors (ratio between data and MC efficiencies) are also shown in the lower panel of these plots. These scale factors binned in  $(|\eta|, p_T)$  are the weights to be applied to the muon efficiencies in MC. Their deviation from 1 is below 1% in all bins.

To evaluate the systematic uncertainties on these scale factors, three variations of the fitting procedure are performed. The signal shape is changed to a simple Crystal Ball, or the background shape is changed to a second order polynomial, or the mass range used in the fit is changed from [2.6, 3.5] GeV to [2.8, 3.4] GeV. Binned scale factors are calculated for these three methods, and compared to the nominal results. This comparison is shown in Fig. 7.8.

The statistical uncertainties on scale factors are calculated as for the global muon scale factors of section 7.1.3.2. As the variations from the change of fit methods are small (always less than 0.15%), the systematic uncertainties on the scale factors are the symmetrised largest deviation from the nominal result, as in section 7.1.3.2.

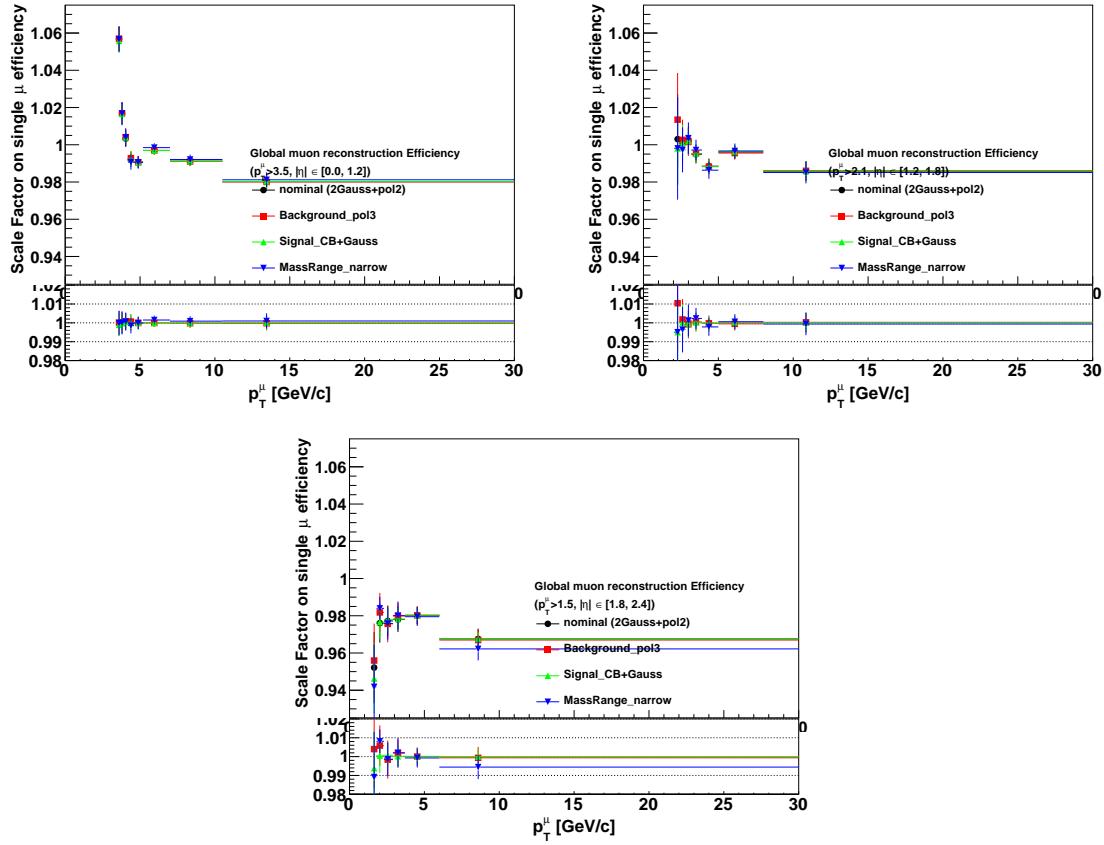


FIGURE 7.4: Scale factors and their variations, for the global muon reconstruction efficiency, versus  $p_T$ , in the three  $|\eta|$  bins. The nominal result is shown in black, and compared to the scale factors derived from the alternative methods. The ratios between the nominal and the varied scale factors are shown in the lower panels.

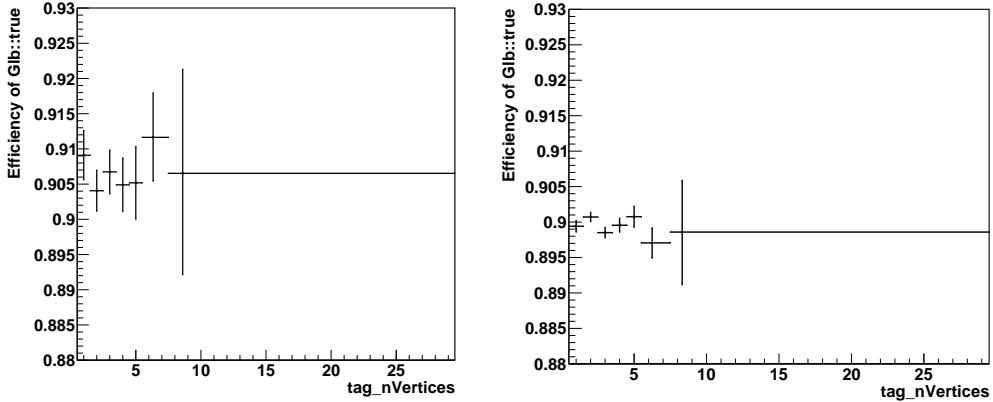


FIGURE 7.5: Global muon reconstruction efficiency as a function of the number of primary vertices, for data (left) and MC (right).

#### 7.1.3.4 Efficiency of muon identification and event trigger

This section computes the efficiencies and scale factors for the hybrid-soft ID and the firing of the HLT\_HIL1DoubleMu0 double muon trigger (when another firing muon already exists), starting from a Global muon. The probes must pass the tight muon kinematic acceptance (section 5.3.2), so that these scale factors be only for muons that are required to fire the event trigger, and are selected with this acceptance.

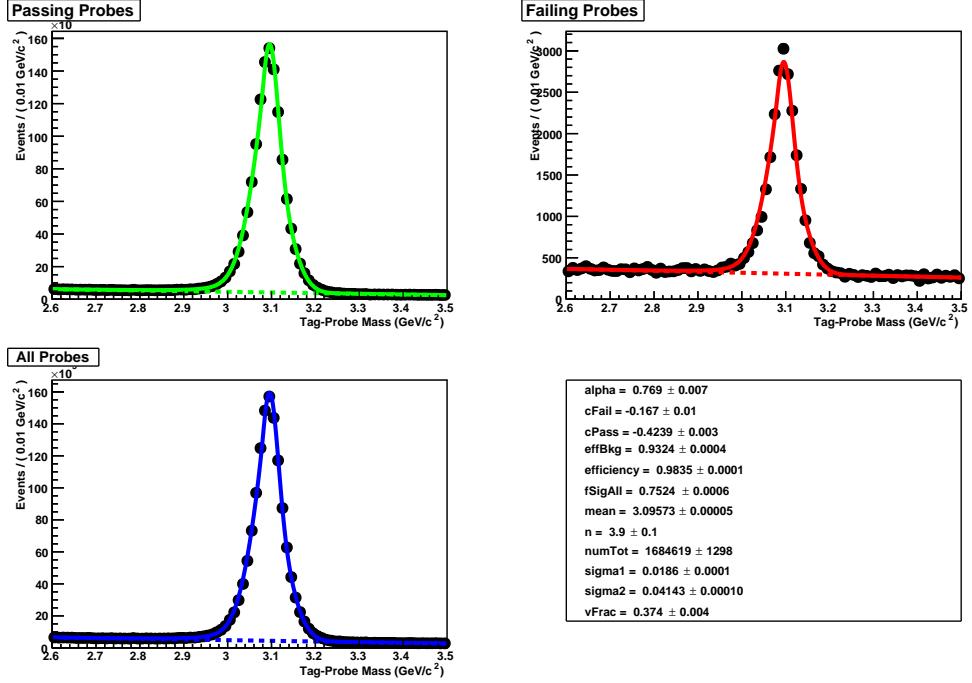


FIGURE 7.6: Dimuon mass fit of the hybrid-soft identification efficiencies, integrated over kinematics, in data. The MC fits look similar, except with a very small background.

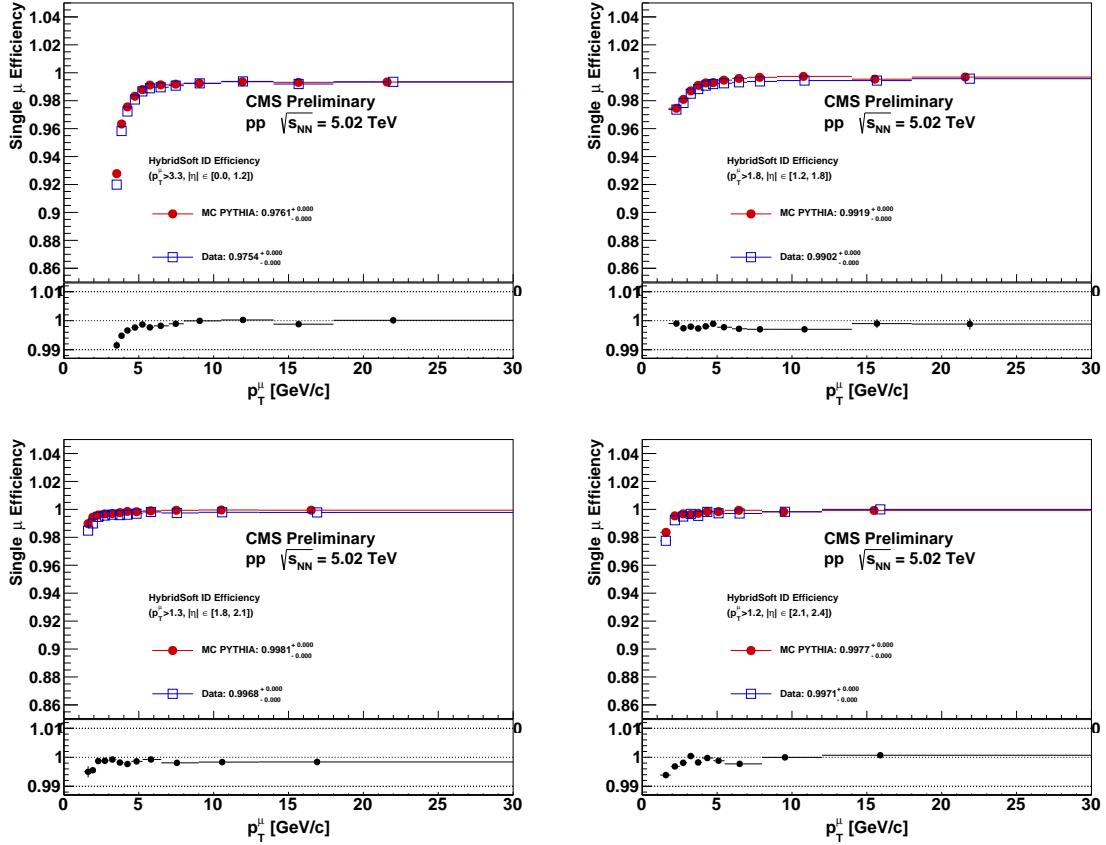


FIGURE 7.7: Hybrid-soft identification efficiency for data and MC, versus  $p_T$ , in four  $|\eta|$  bins

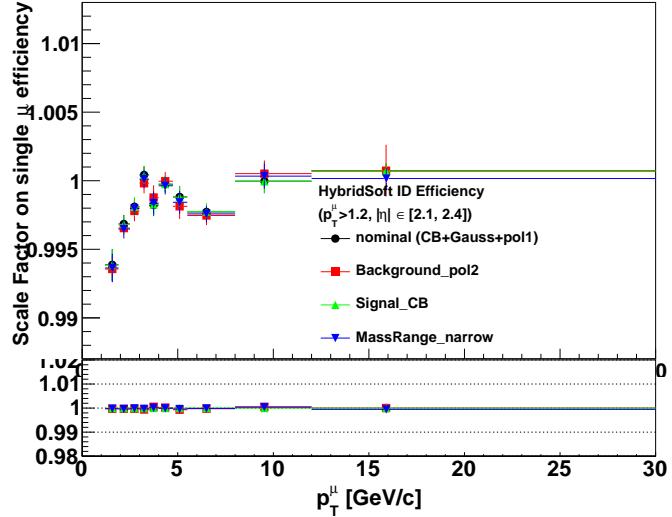


FIGURE 7.8: Scale factors and their variations, for the hybrid-soft identification, versus  $p_T$  in the range  $2.1 < |\eta| < 2.4$  (the plots for the other  $|\eta|$  bins being very similar). See caption of Fig. 7.4 for details.

The probes pass if they satisfy the hybrid-soft ID requirements and fire the studied trigger, and they fail otherwise. The dimuon mass fit for data and MC are similar to those of section 7.1.3.3, including the same signal and background nominal shapes. An example of this fit, for the integrated sample, is shown in Fig. 7.9. The integrated efficiencies for data and MC are 87.89% and 85.58%, respectively.

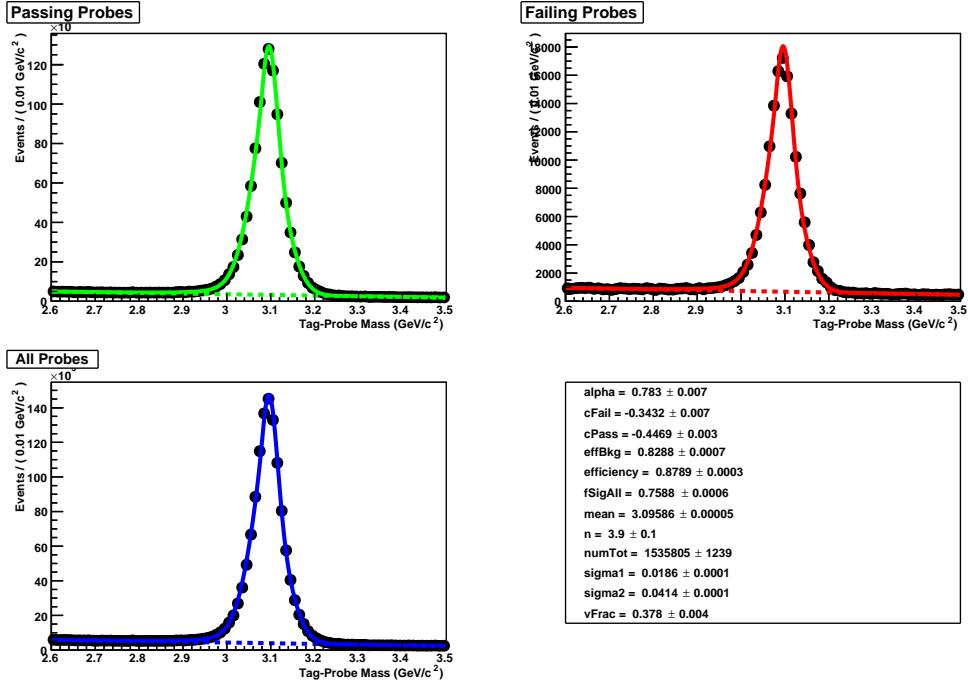


FIGURE 7.9: Dimuon mass fit of the hybrid-soft identification + double muon trigger efficiencies, integrated over kinematics, in data. The MC fits look similar, except with a very small background.

The  $\eta$  and  $p_T$  dependences of the efficiencies are shown in Figs. 7.10 and 7.11, respectively. The  $p_T$  dependence is shown in the same four  $|\eta|$  bins than in section 7.1.3.3. The binned scale factors (ratio between data and MC efficiencies) are also shown in the lower panel of these plots; they are the weights to be applied to the muon efficiencies in MC. Their deviation from 1 goes up to 15% at very low  $p_T$ , but are always below 2% for  $p_T > 4$  GeV.

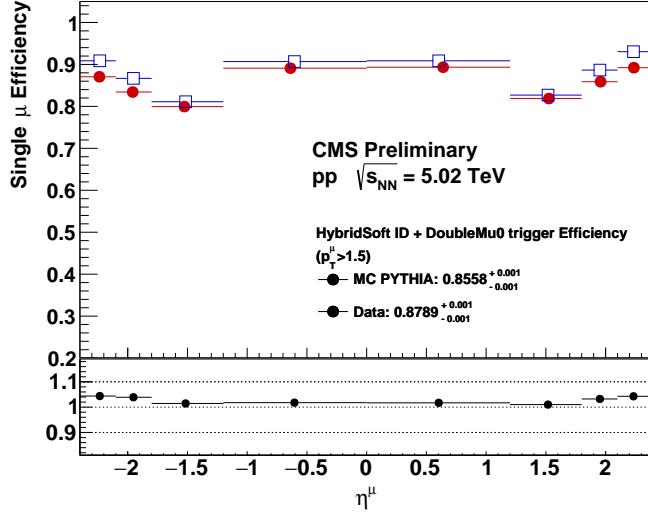


FIGURE 7.10: Hybrid-soft identification + trigger efficiencies for data and MC, versus  $\eta$

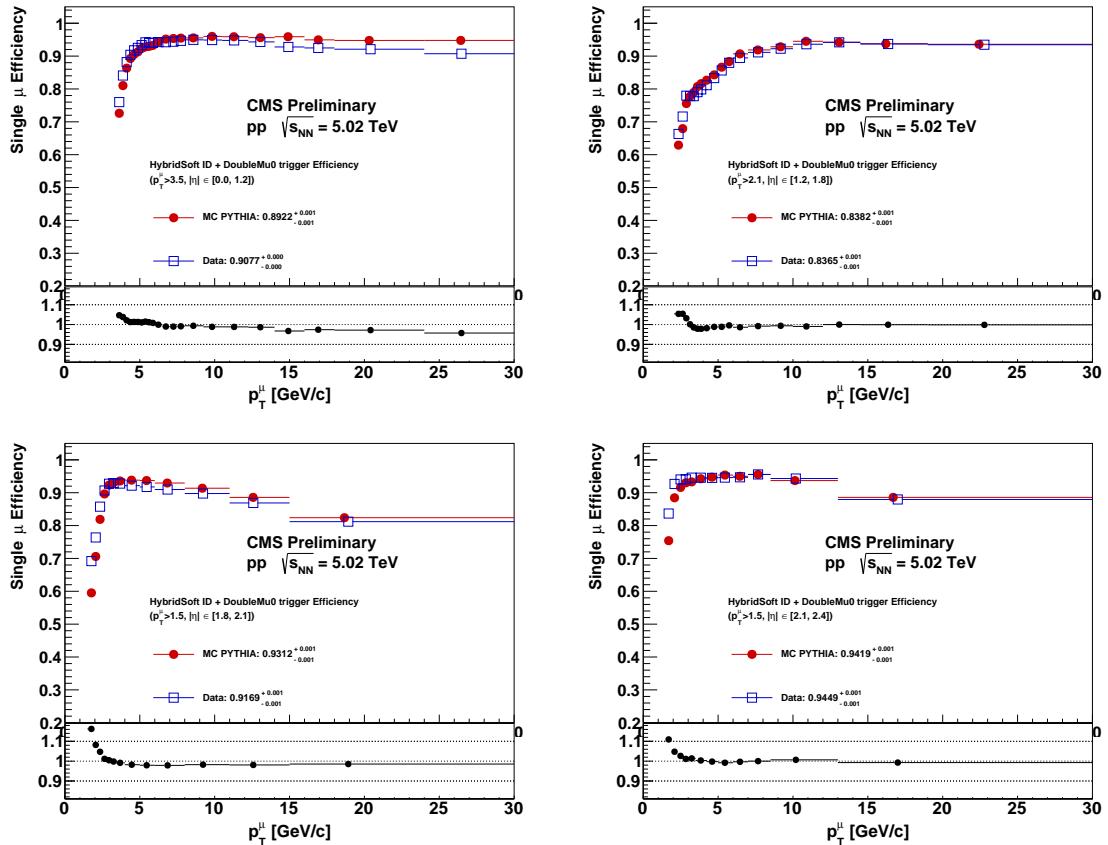


FIGURE 7.11: Hybrid-soft identification + trigger efficiencies for data and MC versus  $p_T$ , in the four  $|\eta|$  bins

The statistical uncertainties on scale factors are calculated as for the global muon scale factors of section 7.1.3.2. The procedure to calculate the systematic uncertainties on the scale factors is also the same as in section 7.1.3.3, using the same three fit method variations. The comparisons of the latter to the nominal fit are shown in Fig. 7.12. As the variations from the change of fit methods are small (always less than 0.5% except for the lowest and highest  $p_T$  bins), the systematic uncertainties on the scale factors are the symmetrised largest deviation from the nominal result, as in section 7.1.3.2.

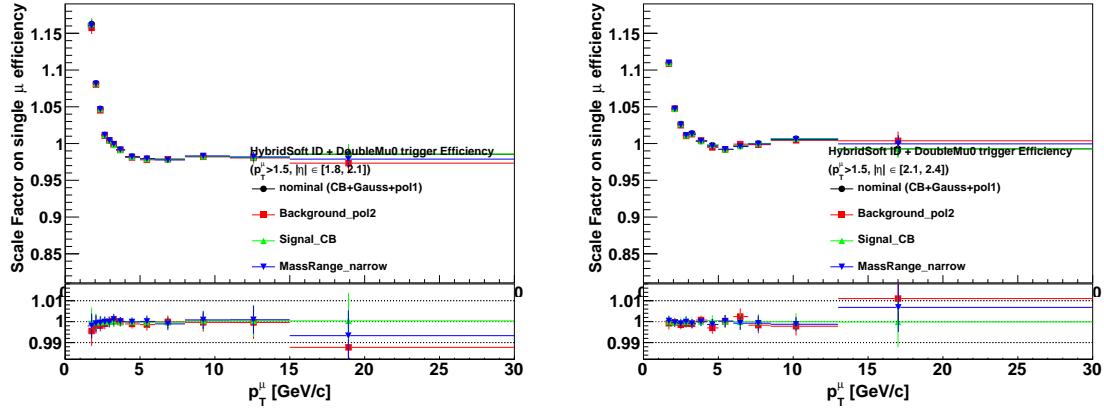


FIGURE 7.12: Scale factors and their variations, for the hybrid-soft ID + trigger, versus  $p_T$  in the last two  $|\eta|$  bins ( $1.8 < |\eta| < 2.1$  and  $2.1 < |\eta| < 2.4$ ); the scale factors in the first two  $|\eta|$  bins are smaller and vary less with the fit method variations. See caption of Fig. 7.4 for details.

A possible pileup dependence has also been studied. The efficiency as a function of the number of PVs is shown in Fig. 7.13, for data and MC. No significant dependence is observed.

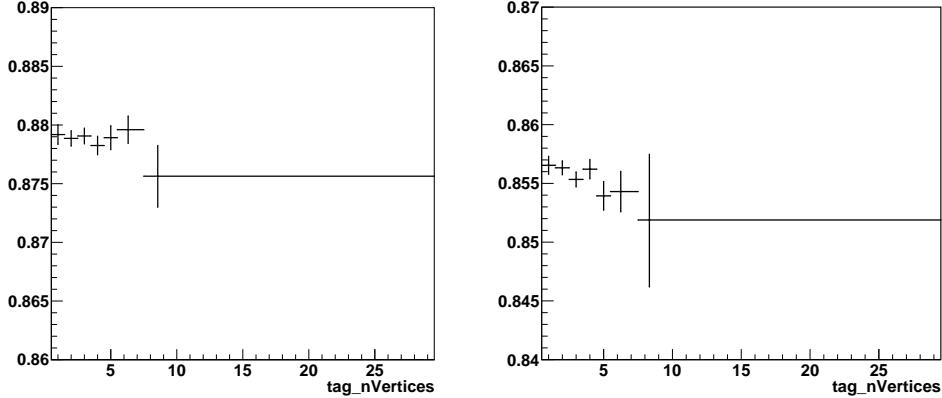


FIGURE 7.13: Hybrid-soft ID + trigger efficiency as a function of the number of primary vertices, for data (left) and MC (right).

### 7.1.3.5 Efficiency of inner track reconstruction and global muon matching

This section checks the efficiency of both the inner track reconstruction, and the matching of a standalone track to an inner track and refitting of the resulting global track (see definitions in section 3.2.3). The considered probes are standalone tracks in the muon chambers. The probes must pass the tight kinematic acceptance (section 5.3.2). To reduce the background from misidentified standalone muons and to stabilise the fit, the standalone track must show in at least two valid muon stations (see Table 7.1).

A probe passes if it is also reconstructed as a global muon, and to fail otherwise. In each bin  $\eta_{\text{probe}}$  or  $p_T, \text{probe}$ , dimuon (tag-probe) mass fits are performed simultaneously in the all/pass/fail categories for data and MC. The large  $p_T$  resolution of standalone muons (hence the large  $J/\psi$  mass width) makes the fit more complicated than in the previous sections, notably because the  $J/\psi$  signal reaches the  $\psi(2S)$  mass. The signal shape is modelled by two Gaussians of common mean around  $m_{J/\psi}^{\text{PDG}}$ , plus a third Gaussian whose mean is fixed to  $m_{\psi(2S)}^{\text{PDG}}$ ; the normalization of the latest is constrained to be 1 to 8% of the double gaussian normalization, around typical values for the  $\psi(2S)/J/\psi$  ratio. The background is described by a second order polynomial. An example of this fit, for the integrated sample, is shown in Fig. 7.14, in data and MC. The integrated efficiencies for data and MC are 98.8% and 99.7%, respectively.

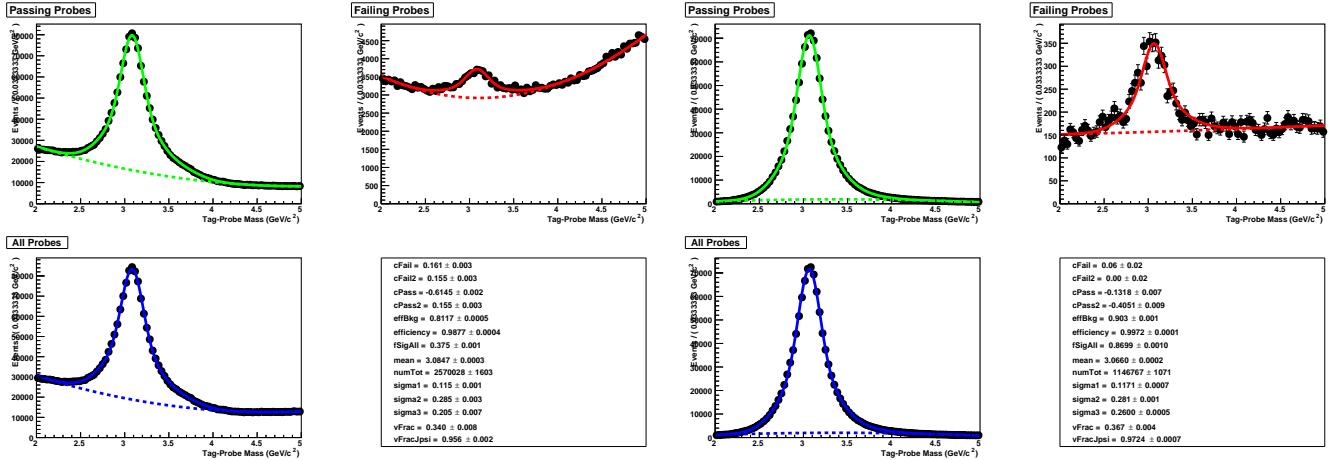


FIGURE 7.14: Fit of the inner track reconstruction and global muon matching efficiencies, integrated over kinematics, for data (left) and MC (right).

The  $\eta$  dependence of the efficiencies is shown in Fig. 7.15. The  $p_T$  dependence is shown in Fig. 7.16 (left), integrated on  $\eta$ . Considering the  $p_T$  resolution of standalone tracks, the  $p_T$  dependence should however not be fully trusted.

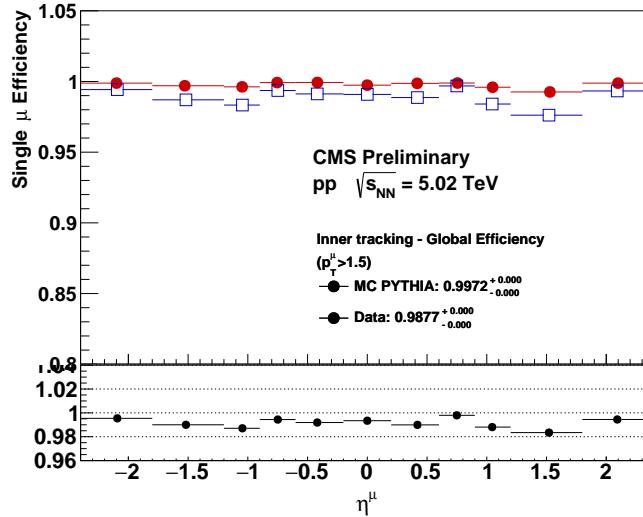


FIGURE 7.15: Inner track reconstruction and global muon matching efficiencies for data and MC, versus  $\eta$ .

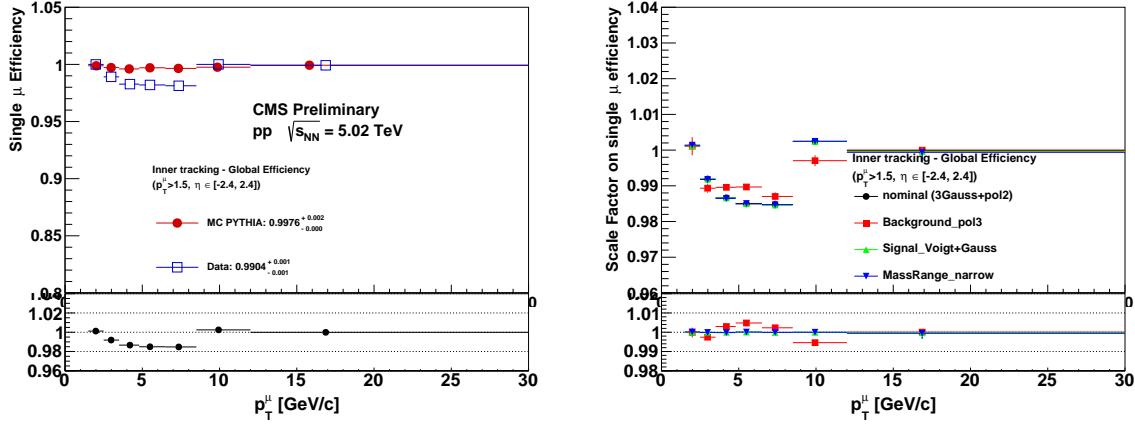


FIGURE 7.16: *Left:* Inner track reconstruction + global muon matching efficiencies for data and MC, versus  $p_T$ . *Right:* Scale factors and their variations, for inner track reconstruction + global matching, versus  $p_T$ . See caption of Fig. 7.4 for details.

The scale factors (ratio between data and MC efficiencies) are also shown in the lower panel of Figs. 7.15 and 7.16. These scale factors can deviate up to 1 or 2%, especially in the transition region. However, one can assume that the inefficiency observed to be wrongly simulated in MC originates mostly from the matching of the inner track to the standalone track, and from the global muon refitting; the inefficiency of inner tracking for muons was indeed shown to be smaller than 0.5% in all  $\eta$  regions, and consistent with simulation, in pp data (see *e.g.* Ref. [172] for Run I). This matching and refitting efficiencies are part of the global muon efficiency, whose scale factors are calculated in section 7.1.3.2 and considered as MC weights in this analyses. Therefore, the scale factors shown here are only checks, and are not considered in the analyses.

Systematic uncertainties on these scale factors are nevertheless evaluated, via three variations of the fitting procedure. The signal shape is changed to a sum of a Voigtian function and a Gaussian (ignoring the  $\psi(2S)$  peak), or the background shape is changed to a third order polynomial, or the mass range used in the fit is changed from [2, 5] GeV to [2.3, 4.7] GeV. Fig. 7.16 (*right*) shows the comparison of the scale factors from the nominal or the alternative fit methods. Only the variation of the background shape gives substantial variations of the scale factors.

#### 7.1.3.6 Closure test in MC

A closure test is performed, comparing the MC traditional efficiencies with the ones fitted in the TnP procedure. It checks the validity of the full TnP method, and if the extracted scale factors can be used as corrections in the standard efficiency measurements.

First, a closure test is performed for the muon ID + trigger efficiencies, comparing:

- The ‘traditional’ efficiency, which starts from the number of global muons in acceptance and that are in the same event than another muon passing all tag cuts detailed in section 7.1.3.1. It then measures the fraction of these probe muons that pass the full ID and trigger cuts.
- The ‘TnP’ efficiency, which is extracted from the MuID+Trigger efficiency measurements from the TnP procedure (of section 7.1.3.4). In practice, each muon from the denominator of the ‘traditional’ efficiency is filled in a histogram with a weight equal to the TnP efficiency corresponding to the  $(p_T, \eta)$  of this muon. A linear interpolation between the binned TnP efficiencies is assumed.

The pseudo-TnP method in the traditional efficiency is needed to calculate a double-muon trigger efficiency for a single muon. The results are shown for the four studied  $|\eta|$  regions in Fig. 7.17, and shows good closure.

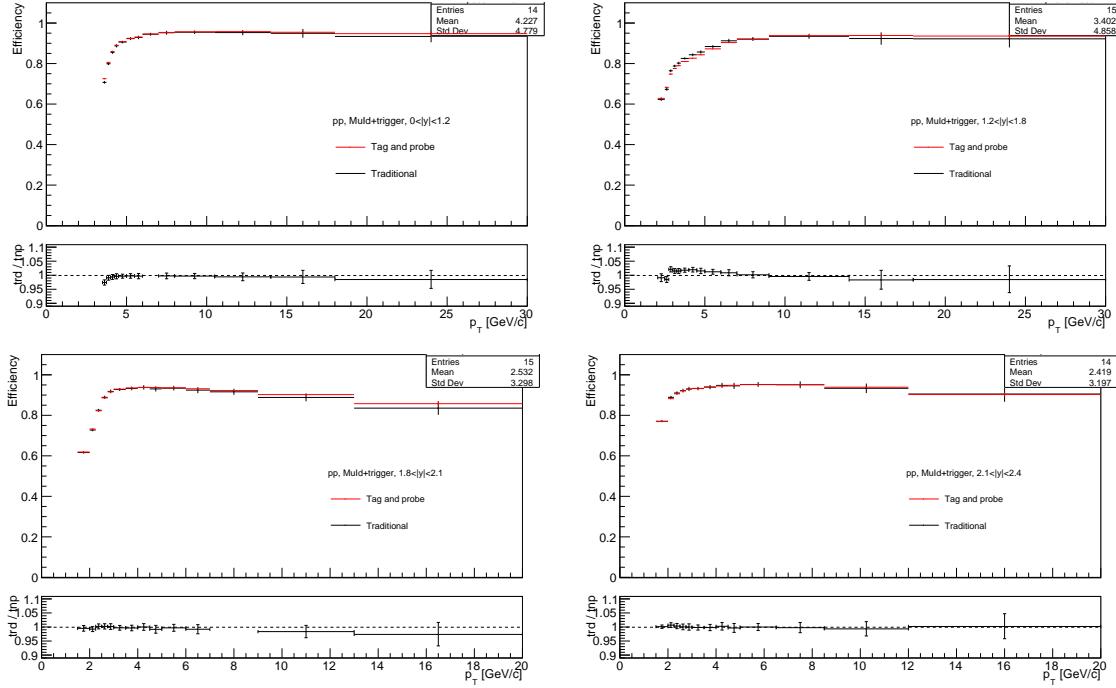


FIGURE 7.17: MC closure test for the hybrid-soft ID + trigger efficiency, for the four  $|\eta|$  regions.

A similar closure test is performed for the Global muon efficiency, comparing:

- The ‘traditional’ efficiency, which starts from the number of tracks in the muon acceptance, passing the loose ID cuts mentioned in section 7.1.3.2, matched to a generated muon, and that are in the same event than a second muon that passes all the tag cuts used in the TnP method. It then measures what fraction of these tracks are reconstructed as global muons.
- The ‘TnP’ efficiency, which is extracted from the the global muon efficiency measurements in the TnP framework (from section 7.1.3.2). Each generated muon in the acceptance is filled in a histogram with a weight equal to the TnP efficiency at the  $(p_T, \eta)$  values of this muon. A linear interpolation between the binned TnP efficiencies is assumed.

The pseudo-TnP method in the traditional efficiency maintains consistency with the muon ID + trigger closure test. The results are shown for the four studied  $|\eta|$  regions in Fig. 7.18, and show a good closure.

A closure test for the full efficiency was also performed (multiplying the global muon and muon ID + trigger histograms separately for the traditional and the TnP method). The resulting plots are not included for brevity, but show good closure; this also indicates that the efficiencies are properly factorizing.

#### 7.1.4 Scale factors for the $B_c$ analysis

The  $p_T$ - and  $\eta$ -dependent scale factors, computed above for pp data, must be applied to each muon entering any efficiency calculation in the MC.

As a reminder, three types of scale factors are used to correct muon efficiencies of the pp MC in the  $B_c$  analysis. They concern the reconstruction of the global muon (from a track probe), the

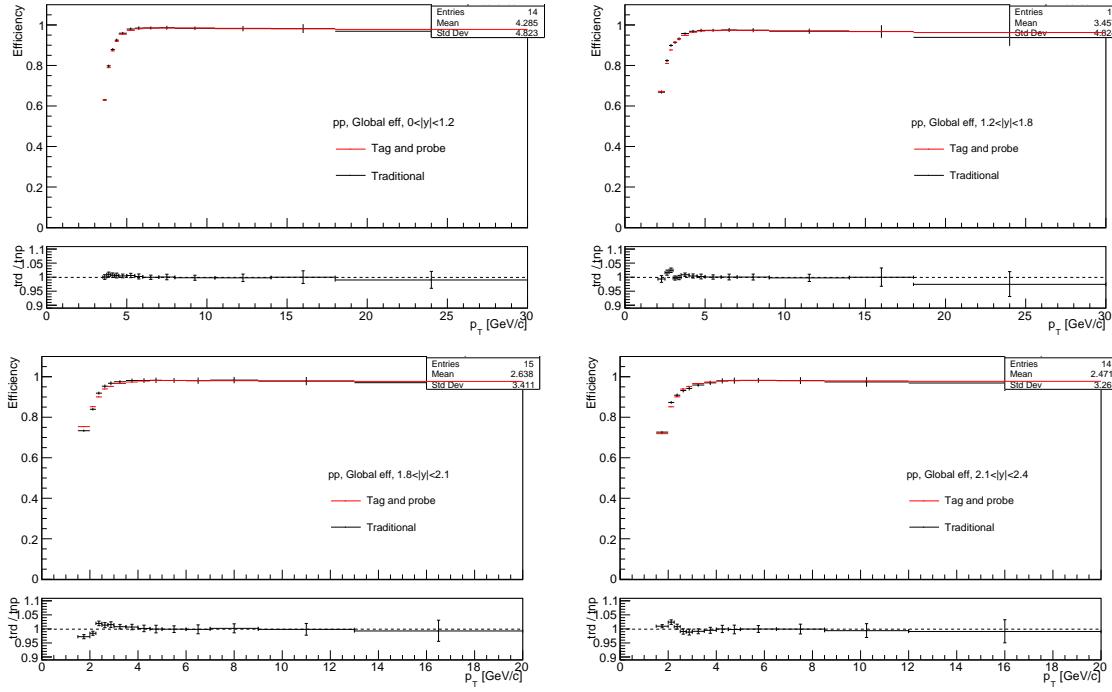


FIGURE 7.18: MC closure test for the global muon efficiency, for the four  $|\eta|$  regions.

identification as hybrid-soft (from a global muon), and the firing of the trigger (from a hybrid-soft probe). (At least) two muons from the  $B_c$  trimuon must fire the double-muon trigger and pass the tight acceptance, and a third muon is not required to trigger, and is hence allowed to be in a looser acceptance than the two triggering muons (see section 5.3.3). The *global* (section 7.1.3.2) and *ID* (section 7.1.3.3) scale factors, both in the loose acceptance, are hence applied to the third muon; whereas the *global* and *ID+trigger* (section 7.1.3.4) scale factors, both in the tight acceptance, are applied to the two triggering muons. In section 7.2, I will call ‘loose’ ( $l$ ) the total scale factor and efficiency associated to the first case (non-triggering muon), and ‘tight’ ( $t$ ) the ones associated to the second case (triggering muons) including the probability of passing the tight acceptance cuts from the loose acceptance.

In PbPb, no scale factor is used for the reconstruction of a global muon from a track (*i.e.* the reconstruction of a standalone muon track and its matching to a track of the inner tracker), but it was checked to be close to 1 within uncertainties. The scale factors for track reconstruction (from a standalone muon) and hybrid-soft identification (from a global muon) are used for the identification in the loose muon acceptance; the associated efficiencies, for the identification of non-triggering muons, is named  $l$  in section 7.2. The official TnP results in PbPb, only calculated in the tight acceptance, were rerun to extend them to the loose acceptance. Then the scale factors for the L3Mu0NHitQ10\_L2Mu0\_MAXdR3p5\_M1to5 trigger must be separated for the efficiencies of its L2 and L3 branches,<sup>1</sup> reflecting the asymmetric requirements on the two HLT muons. In section 7.2, the L2 (+identification) efficiency (and scale factor) is named  $t$  as ‘trigger’, and the L3 (+L2+identification) efficiency is named  $t'$ . These two efficiencies also include the probability to pass the tight acceptance from the loose one.

### 7.1.5 Tentative soft ID scale factors in PbPb

Section 5.3.3 mentions an alternative strategy for the identification of the three muons, where one muon is allowed to pass only the *soft* identification (in an even looser muon acceptance),

<sup>1</sup>The efficiency of the  $\Delta R < 3.5$  and  $1 < M_{\mu\mu} < 5$  GeV requirements of this trigger is very close to one, so potential associated scale factors are neglected.

while the other two muons must fire the trigger and pass hybrid-soft ID. This section presents the difficulties linked to measuring scale factors for this selection, through tentative estimations of scale factors for the *soft* identification in PbPb. I thank Noémie PILLEUX, who performed this study during her internship at LLR, and re-ran the official TnP PbPb scale factors in the loose acceptance (see section 7.1.4).

Scale factors for the soft identification efficiency in PbPb were never calculated. It is indeed difficult to deal with the very large background from charged tracks, when counting what fraction of charged tracks probes are passing the tracker-muon requirement. As a comparison, the charged track background in a similar setup in pp is manageable, cf section 7.1.3.2.

The strategy for the soft ID scale factors is to factorise the efficiency in two parts:

- The efficiency of tracker muons. The probes are charged tracks with a basic purity selection (coinciding with the part of the soft identification that concerns the inner track, see section 5.3.1).
- The efficiency of the *soft* identification. The probes are here tracker muons.

The efficiencies and scale factors are estimated with the TnP method used previously in this section. Fig. 7.19 (*left*) shows the pseudo-rapidity dependence for the tracker muon efficiency, and Fig. 7.19 (*right*) shows the  $p_T$  dependence for the soft ID efficiency, in one of the three existing bins of absolute pseudo-rapidity.

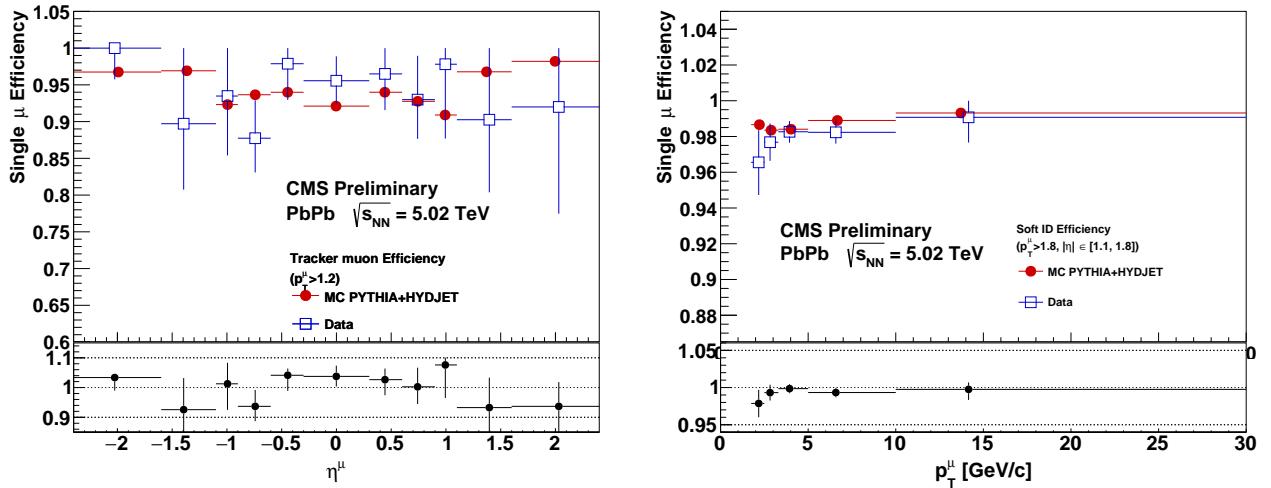


FIGURE 7.19: *Left*: pseudo-rapidity dependence of tracker muon efficiencies, and the resulting data/MC scale factors. *Right*: transverse momentum dependence of *soft*-ID muon efficiencies and the resulting scale factors, in the region  $1.1 < |\eta| < 1.8$ .

The uncertainties on the tracker muon efficiency were judged too large ( $> 10\%$ ) for this option to be viable.

## 7.2 Applying efficiency corrections to trimuons

### 7.2.1 Applying scale factors on a MC trimuon

The signal efficiency in MC is the ratio of the number of passing trimuons, each corrected with a scale factor, to the number of accepted trimuons. A scale factor must be attributed to each trimuon entering the numerator. It should be the ratio of the ‘true’ trimuon efficiency (from data) to its MC efficiency. The full efficiency of a trimuon can be assumed to factorise in

the efficiencies for identification and triggering of the three separate muons and in a trimuon-specific efficiency (corresponding to the preselection cuts of Tables 5.1 and 5.2, excluding the muon identification). No scale factor is calculated for the latter, because that would necessitate  $B_c$  probes in data.

On the other hand, scale factors are available for the single-muon efficiencies. When the single-muon requirements are symmetric, *i.e.* the same cuts of efficiency  $\varepsilon_\mu$  (and associated scale factor  $S(\mu)$ ) are applied on each of the three muons, the single-muon part of the trimuon efficiency is  $\varepsilon_\mu(\mu_1)\varepsilon_\mu(\mu_2)\varepsilon_\mu(\mu_3)$ . The factorisation of the efficiencies of the three muons is justified by the relatively low- $p_T$  of the trimuons of this analysis, so that the muons are not too collimated (see the plots of the  $\Delta R$  between all muon pairs of trimuon candidates in Figs. 5.13 and 5.14) and rarely overlap in the detector, allowing for independent reconstruction of each muon. In this simple case, the scale factor to be applied is plainly  $S(\mu_1)S(\mu_2)S(\mu_3)$ . I clarify that, here and in the following,  $\varepsilon^c(\mu_i) \equiv c_i$  and  $S^c(\mu_i) \equiv S_i^c$  are the efficiency and associated scale factor of cuts  $c$ , for a typical muon having the kinematics ( $p_T, \eta$ ) of  $\mu_i$ , extracted from the TnP procedure.

## 7.2.2 Scale factors for trimuons with asymmetric muon cuts

### 7.2.2.1 General procedure for $B_c$ trimuons

However, our requirements on the three muons are asymmetric: only two muons out of three need to trigger, and in PbPb one muon needs to fire a tighter trigger branch than the second muon. These telescopic efficiencies were named earlier  $l$  (identification),  $t$  (identification + (L2) trigger) and  $t'$  (identification + (L2+L3) trigger)<sup>1</sup>; in pp the double muon trigger is symmetric so  $t = t'$ . The efficiency of such a triplet requirement is **not** the simple product of the three muon efficiencies  $l(\mu_1)t(\mu_2)t'(\mu_3) \equiv l_1t_2t'_3$ , and the scale factor is not  $S_1^lS_2^tS_3^{t'}$ . How would one choose on which muon the scale factor for loose efficiency is applied? An intuitive solution would be to take the trimuon scale factor as the average of the cases where the different scale factors are applied in turn to each muon:

$$(S_1^lS_2^tS_3^{t'} + S_1^lS_3^tS_2^{t'} + S_2^lS_1^tS_3^{t'} + S_2^lS_3^tS_1^{t'} + S_3^lS_2^tS_1^{t'} + S_3^lS_1^tS_2^{t'})/6 \quad (7.2)$$

One could think as well of selectively applying the scale factors for the tight efficiency  $t$  only to muons that pass the trigger, and for the efficiency  $t'$  only to muons passing the L3 trigger. For example, if  $\mu_1$  does not pass the trigger, one would correct it with the scale factor  $(S_1^lS_2^tS_3^{t'} + S_1^lS_3^tS_2^{t'})/2$ . The two intuitive ideas above are nevertheless wrong because the scale factor must be the ratio of efficiencies of a typical trimuon whose muons have the same kinematics than the trimuon to be corrected (without considering the cuts passed by the particular trimuon to be corrected). Eq. 7.2 is already wrong, for two reasons. First, the trimuon scale factor should be the ratio of data and MC trimuon efficiencies, not a linear combination of single-muon scale factors. Then the full trimuon efficiency is not the simple average of  $l_i t_j t'_k$  for all  $(i, j, k)$  muon permutations: binomial probabilities need to be taken into account. Sections 7.2.2.2 to 7.2.2.4 will show that the actual efficiency whose data/MC ratio must be considered is:

$$t'_1 t'_2 t'_3 + \{t'_1 [(l_2 - l_1)t_3 + (t_2 - t'_1)t_3] + perm(1, 2, 3 \rightarrow 2, 3, 1) + perm(1, 2, 3 \rightarrow 3, 1, 2)\} \quad (7.3)$$

where the first term of the curly brackets is repeated three times with each possible rotation of the  $(1, 2, 3)$  indices.

The signal efficiency from MC is the number of (accepted) trimuon candidates passing the cuts and weighted by a scale factor, divided by the number of accepted trimuons. Three methods

---

<sup>1</sup>When a muon does not pass the tight acceptance, then we set  $t = t' = 0$  (this is equivalent to saying that  $t$  and  $t'$  are the efficiencies of triggering and being in the tight acceptance).

are tested for this trimuon scale factor. The first one (blue in Fig. 7.20), that is correct, is

$$\frac{\varepsilon_{trimu}^{data}}{\varepsilon_{trimu}^{MC}} = \frac{\varepsilon_{trimu}(L \times S^l, T \times S^t, T' \times S^{t'})}{\varepsilon_{trimu}(L, T, T')} \quad (7.4)$$

where the capital letters designate the MC efficiencies (whereas the lower case efficiencies are the [true] data ones), and the trimuon efficiency is that of Eq. 7.3. One practical consequence is that the data and MC efficiencies themselves, and not only their ratio, must be stored in the output of the TnP procedure. The second method (red in Fig. 7.20) is based on the previous one, but selectively applies the single-muon efficiencies in the full trimuon efficiency, *e.g.* I set  $t'_1 = T'_1 = 0$  if  $\mu_1$  does not pass L3 or  $t_2 = T_2 = 0$  if  $\mu_2$  does not trigger. The third method (green in Fig. 7.20) expands on the latter, but using the averaged scale factor of Eq. 7.2 – this was at some point the wrongly recommended method for  $J/\psi$  dimuons with asymmetric double-muon trigger.

The distribution of trimuon scale factors (from preselected candidates of the second-step-corrected MC) is drawn for these three methods in Fig. 7.20. The mean applied scale factor for the nominal method in pp or PbPb is 0.964 or 1.028, respectively; the difference of this mean with the simplistic method is 1% in pp (6% in PbPb), *i.e.* of the order of the scale factor correction itself.

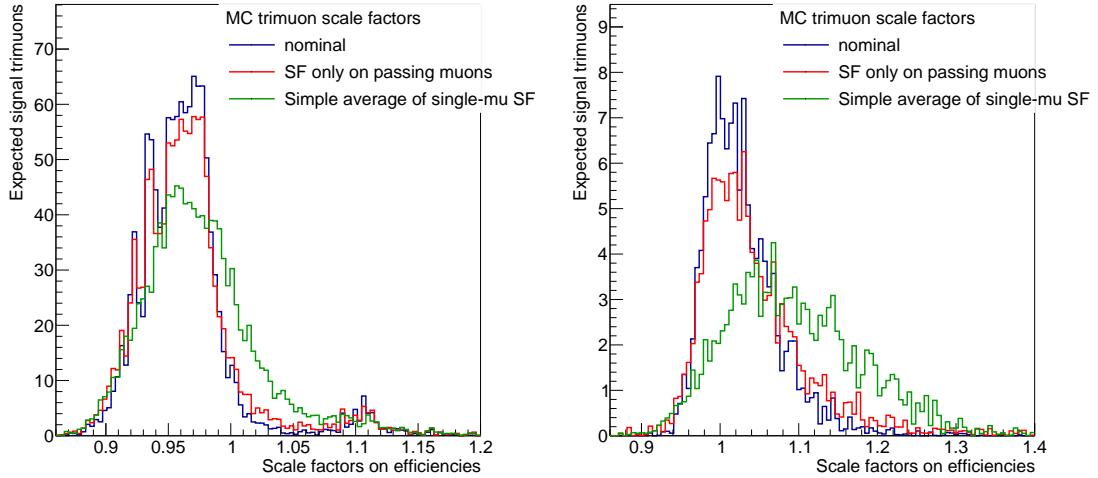


FIGURE 7.20: Scale factors applied to signal MC trimuons in the signal efficiency estimation, in pp (*left*) and PbPb (*right*). Three methods for calculating the trimuon scale factors are compared (see text for details). The blue distribution is from the correct method (Eqs. 7.3 and 7.4) and is taken as nominal.

The maps of the acceptance times efficiency of the signal are shown in section 7.4. The efficiency map is affected by the scale factors shown previously. The ratio of the nominal efficiency map to the one without scale factors, or to the one with the two alternative scale factor calculations, are shown for pp and PbPb in Fig. 7.21.

The single-muon TnP scale factors are attributed uncertainties, depending on the muon  $p_T$  and  $\eta$ . The efficiency measurements must be run four additional times, using the up and down values of the statistical or systematic uncertainties on the scale factors. These uncertainties are symmetrised (the mean of the up and down deviations from the nominal efficiency is taken). The uncertainty is the quadratic sum of the statistical and systematic deviations of the efficiencies, from each of the three separate scale factors. This is quoted as the tag-and-probe systematic

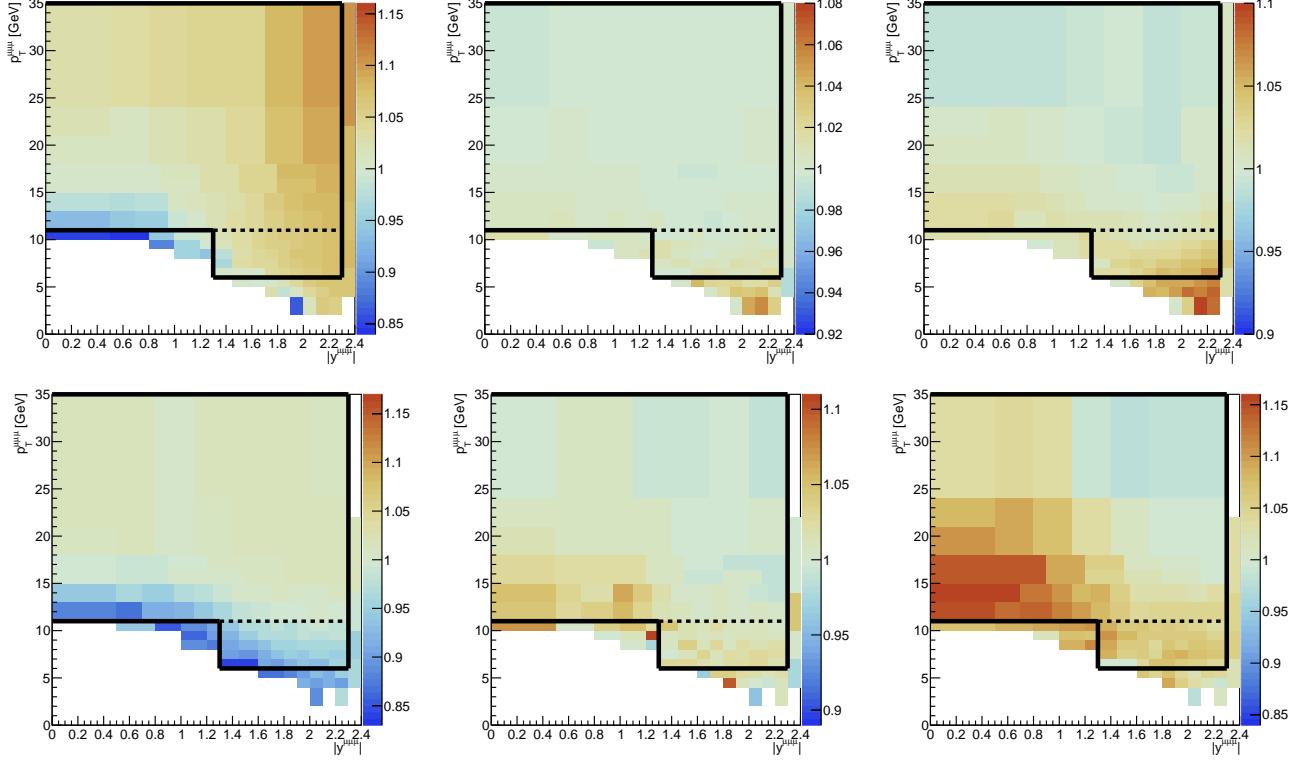


FIGURE 7.21: Ratio of the nominal pp (*top*) or PbPb (*bottom*) trimuon efficiency map to the maps using alternative trimuon scale factor methods. The black lines are the fiducial cuts chosen in section 5.6. *Left:* no scale factor is applied. *Middle:* in the scale factor of Eq. 7.4, non-zero efficiencies are used only with muons that pass the associated cuts. *Right:* idem, but using the scale factor from Eq. 7.2.

uncertainty in section 8.1. A 80% correlation between the two  $p_T$  bins is assumed because there is a large overlap between the kinematics of the single muons for the two  $p_T^{\mu\mu\mu}$  bins.<sup>1</sup>

This way to apply the TnP scale factor uncertainties on the efficiency is the one recommended by the group, but it is clearly overestimated: it considers that the scale factor errors are 100% correlated along all kinematic bins. It could be true for the systematic errors, but is wrong for the statistical errors. I am convinced that toys should be in principle used to apply these uncertainties, calculating many efficiency values, each using random variations of the scale factors within their uncertainties. However, the TnP uncertainty is among the smallest contributions to the full uncertainty (see chapter 8), so this can be overlooked.

Sections 7.2.2.2 to 7.2.2.4 calculate exactly the trimuon efficiency and scale factor as a function of the various single-muon efficiencies (of Eq. 7.3), from the simpler case of the asymmetric dimuon trigger in PbPb (a case useful for all  $J/\psi$  analyses in PbPb, using mostly this trigger) to the full case of the trimuon efficiency with three different efficiencies on each muon. The efficiencies can be computed using explicit decision trees or using the inclusion–exclusion principle, yielding identical results.

### 7.2.2.2 Case of asymmetric PbPb dimuon trigger

#### Decision trees

What scale factor should be applied to a dimuon in which two muons must fire an asymmetric

<sup>1</sup>It might be possible to calculate it, but it would be quite sophisticated, hence not judged necessary considering the uncertainty is subdominant.

trigger, one muon firing an L2 branch (efficiency  $t_i$  on muon  $i$ ) and the other firing an L3 branch (efficiency  $t'_i < t_i$ ), given that firing the L3 branch implies firing the L2 one? Fig. 7.22 draws a decision tree with probabilities  $t'_1, t_1 - t'_1$  (the probability of triggering L2 but not L3) and  $1 - t_1$  (probability of not triggering at all) in the first layer, and  $t'_2, t_2 - t'_2$  and  $1 - t_2$  in the second layer. At each node (*i.e.* a decision for one of the muons), the probabilities sum to 1.

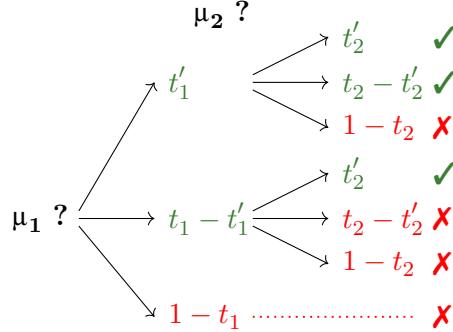


FIGURE 7.22: Decision tree for the calculation of the efficiency of a dimuon with asymmetric requirements on its two muons, representative of the PbPb trigger case. The paths leading to the dimuon passing the cuts are green, others are red.

The full probability for a dimuon to pass the asymmetric cuts is the sum of the green paths of Fig. 7.22, where the probabilities of the first and second muon decisions are multiplied (because the reconstruction of the two muons are assumed independent, see section 7.1 for a justification). The dimuon efficiency is hence

$$t'_1 t_2 + (t_1 - t'_1) t'_2 = t'_1 t'_2 \left( \frac{t_1}{t'_1} + \frac{t_2}{t'_2} - 1 \right) \quad (7.5)$$

The scale factor is the above efficiency for data, divided by the same expression for MC (see Eq. 7.4). With the notations of section 7.2.2, the full scale factor is hence:

$$SF = S'_1 S'_2 \frac{\frac{t_1}{t'_1} + \frac{t_2}{t'_2} - 1}{\frac{T_1}{T'_1} + \frac{T_2}{T'_2} - 1} \quad (7.6)$$

Next, I calculate an order of magnitude of the ratio of this scale factor to the intuitive – but false – one that was advised until now for J/ $\psi$  analyses with the PbPb trigger. The latter is the average over applying the attributing the L3 scale factor to one muon and the L2 to the other one, and vice versa:  $\frac{1}{2}(S'_1 S^t_2 + S^t_1 S'_2)$ . With the rough approximation of equal efficiency ratios  $t/t'$  and  $T/T'$  for both muons, I get:

$$\frac{SF_{full}}{SF_{simple}} = \frac{\frac{t_1}{t'_1} + \frac{t_2}{t'_2} - 1}{\frac{T_1}{T'_1} + \frac{T_2}{T'_2} - 1} \frac{2}{\frac{S^t_1}{S'_1} + \frac{S^t_2}{S'_2}} \sim \frac{1 - \frac{1}{2} \frac{t'}{t}}{1 - \frac{1}{2} \frac{T'}{T}} \quad (7.7)$$

From there, concrete values of scale factors are needed to quantitatively estimate the difference between methods – this is done in Fig. 7.21.

Doing however the same exercise directly with the efficiencies (with the same assumption that both muons behave the same way), the ratio between the full method and a simple average of the efficiencies applied in turn to each muon ( $(t_1 t'_2 + t'_1 t_2)/2$ ) is more dramatic. It is of order  $2 - \frac{t'}{t}$ , which tends to 2 when  $t' \ll t$ !

### Inclusion–exclusion principle

The same results can be obtained using the simple inclusion–exclusion principle. I thank François ARLEO for this crosscheck of the decision tree formulae (including in the two following sections). Define the two events:

- $A = \mu_1 L_2 \cap \mu_2 L_3$
- $B = \mu_1 L_3 \cap \mu_2 L_2$

where the notation  $\mu_i L_k$  means that the muon  $\mu_i$  fires the  $L_k$ . The wanted probability is that of A or B occurring,  $P(A \cup B)$ , given by

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \quad (7.8)$$

Since  $\mu_1$  and  $\mu_2$  are independent,  $P(A) = P(\mu_1 L_2)P(\mu_2 L_3) = t_1 t'_2$  and similarly for  $B$ , so:

$$P(A \cup B) = t_1 t'_2 + t'_1 t_2 - P(A \cap B). \quad (7.9)$$

Let us compute  $P(A \cap B)$ , knowing that events on  $\mu_1$  or  $\mu_2$  are independent:

$$\begin{aligned} P(A \cap B) &= P(\mu_1 L_2 \cap \mu_2 L_3 \cap \mu_1 L_3 \cap \mu_2 L_2) \\ &= P(\mu_1 L_2 \cap \mu_1 L_3) \times P(\mu_2 L_2 \cap \mu_2 L_3). \end{aligned} \quad (7.10)$$

Using  $P(\mu_1 L_2 \cap \mu_1 L_3) = P(\mu_1 L_3) = t'_1$  (since  $L_3 \subset L_2$ ), and idem for  $\mu_2$ , and Eq. 7.10, one gets

$$P(A \cup B) = t_1 t'_2 + t'_1 t_2 - t'_1 t'_2, \quad (7.11)$$

which is indeed Eq. 7.5.

#### 7.2.2.3 Symmetric dimuon trigger plus a third muon (pp case)

##### Decision trees

A similar efficiency calculation can be performed for the pp case, where the double muon trigger is symmetric (efficiency  $t$ , including identification), but one of the three muons is not required to fire it (efficiency of the identification  $l > t$ ). Passing the  $t$  efficiency requirements implies passing the  $l$  efficiency.

Fig. 7.23 draws a decision tree with probabilities  $t_1$ ,  $l_1 - t_1$  (the efficiency of passing the loose efficiency, but not the trigger) and  $1 - l_1$  (the probability of not passing the loose cuts) in the first layer, and  $t_2$ ,  $l_2 - t_2$  and  $1 - l_2$  in the second layer. At each node, the probabilities sum to 1.

The full probability for a trimuon to pass these asymmetric cuts is the sum of the green paths of Fig. 7.23, where the probabilities of the first, second, and third muon decisions are multiplied. The trimuon efficiency in this case is hence

$$t_1 t_2 t_3 + t_1 t_2 (l_3 - t_3) + t_1 t_3 (l_2 - t_2) + t_2 t_3 (l_1 - t_1) = t_1 t_2 t_3 (l_1/t_1 + l_2/t_2 + l_3/t_3 - 2) \quad (7.12)$$

Now, let us calculate the scale factor to be applied to this trimuon, with the notations of section 7.2.2 (capital and small letters for MC and data efficiencies, respectively). The scale factor is the trimuon efficiency of Eq. 7.12 for data, divided by that in MC:

$$SF = S_1^t S_2^t S_3^t \frac{\frac{l_1}{t_1} + \frac{l_2}{t_2} + \frac{l_3}{t_3} - 2}{\frac{L_1}{T_1} + \frac{L_2}{T_2} + \frac{L_3}{T_3} - 2} \quad (7.13)$$

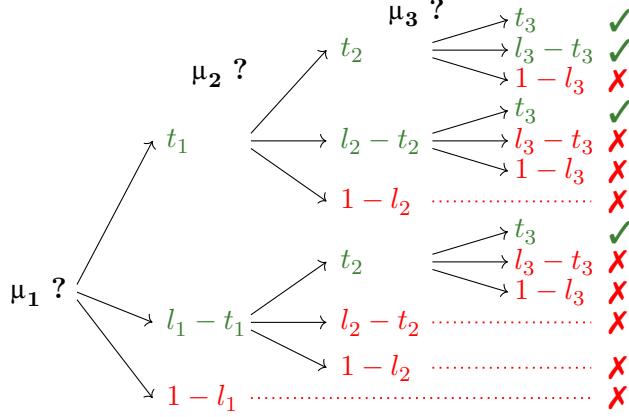


FIGURE 7.23: Decision tree for the calculation of the efficiency of a trimuon with one muon having a different efficiency ( $l$ ) than the other two ( $t$ ), representative of the pp case. The paths leading to a trimuon passing the cuts are green, others are red.

or, introducing the scale factor for loose cuts:

$$SF = \frac{\frac{L_1}{T_1} S_1^l S_2^t S_3^t + \frac{L_2}{T_2} S_1^t S_2^l S_3^t + \frac{L_3}{T_3} S_1^t S_2^t S_3^l - 2S_1^t S_2^t S_3^t}{\frac{L_1}{T_1} + \frac{L_2}{T_2} + \frac{L_3}{T_3} - 2} \quad (7.14)$$

In practice the explicit SF formula above is not used: the scale factor is the ratio of the trimuon efficiencies from Eq. 7.12 calculated separately in data and MC (see Eq. 7.4).

As for the PbPb dimuon trigger case, let us compare this exact scale factor to the simple average over applying the different efficiencies to each of the three muons in turn (as in Eq. 7.2). Considering (for a rough estimate) similar efficiencies for each muon, this ratio is

$$\frac{SF_{full}}{SF_{simple}} \sim \frac{\frac{2}{3}(l-t)}{l - \frac{2}{3}t} = \frac{1 - \frac{t}{l}}{\frac{3}{2} - \frac{t}{l}} \quad t \ll l \quad (7.15)$$

For a quantitative estimate, actual values of the efficiencies and scale factors are needed. Fig. 7.20 (*left*) compares the exact method with the simplistic method in pp, and the difference in the mean applied scale factor is about 4% in the pp case, and the exact method shows lower scale factors than the simplistic one as predicted by the analytical approximation of Eq. 7.15.

### Inclusion–exclusion principle

Define the three events:

- $A = \mu_1 L_1 \cap \mu_2 L_1 \cap \mu_3 L_1$
- $B = \mu_1 L_1 \cap \mu_2 L_0 \cap \mu_3 L_1$
- $C = \mu_1 L_0 \cap \mu_2 L_1 \cap \mu_3 L_1$

Here,  $\mu_k L_1$  means  $\mu_k$  fired the trigger and passed ID cuts (probability  $t_k$ ), and  $\mu_k L_0$  means  $\mu_k$  passed the loose efficiency (ID cuts, probability  $l_k$ ). As before, the probability that one of these events occurs,  $P(A \cup B \cup C)$ , is needed:

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C) \\ &= t_1 t_2 t_3 + t_1 l_2 t_3 + l_1 t_2 t_3 - P(A \cap B) - P(B \cap C) - P(A \cap C) + P(A \cap B \cap C). \end{aligned}$$

In addition,

$$P(A \cap B) = P(\mu_1 L_1) \times P(\mu_2 L_0 \cap \mu_2 L_1) \times P(\mu_3 L_0 \cap \mu_3 L_1) = t_1 t_2 t_3,$$

where  $P(\mu_2 L_0 \cap \mu_2 L_1) = P(\mu_2 L_1) = t_2$  (and idem for  $\mu_3$ ) since  $L_1 \subset L_0$ , and events concerning different muons are independent. Similarly,  $P(B \cap C) = P(A \cap C) = t_1 t_2 t_3$ . Since one has in addition  $P(A \cap B \cap C) = P(\mu_1 L_1 \cap \mu_2 L_1 \cap \mu_3 L_1) = t_1 t_2 t_3$ , this leads to

$$P(A \cup B \cup C) = t_1 t_2 l_3 + t_1 l_2 t_3 + l_1 t_2 t_3 - 2t_1 t_2 t_3, \quad (7.16)$$

which is the same as Eq. (7.12).

#### 7.2.2.4 Full PbPb case

##### Decision trees

Let us now make the full calculation where the two above cases combine in PbPb. The decision tree displays for each of the three layers (*i.e.* for each muon) the probabilities  $t'_i$  (L3 trigger),  $t_i - t'_i$  (L2 trigger but not L3),  $l_i - t_i$  (pass ID but not trigger),  $1 - l_i$  (does not pass ID), of sum 1. It is more cumbersome to draw, but the method is the same as for the two simpler cases above. The resulting probability for having at least three identified muons, among which two are triggering, including one muon triggering L3, is expressed via rotations of the 1, 2, 3 indices as:

$$\begin{aligned} & t'_1 t'_2 t'_3 + [t'_1 [(l_2 - l_2)t_3 + (t_2 - t'_2)l_3] + perm(1, 2, 3 \rightarrow 2, 3, 1) + perm(1, 2, 3 \rightarrow 3, 1, 2)] = \\ & t'_1 t'_2 t'_3 + [l_1 [t'_2(t_3 - t'_3) + t_2 t'_3] - t'_1 t_2 t_3 + perm(1, 2, 3 \rightarrow 2, 3, 1) + perm(1, 2, 3 \rightarrow 3, 1, 2)] \end{aligned} \quad (7.17)$$

One can understand this probability as expressed in the first line through the combination of different paths of the decision tree. The first term with three  $t'$  is the case where all muons trigger the L3. The next term, which considers the case where  $\mu_1$  triggers L3, is repeated three times with rotations of the concerned muons. It is the addition of the following cases:

- if  $\mu_2$  passes ID but not trigger (efficiency  $(l_2 - t_2)$ ), then  $\mu_3$  must trigger (L2 or L3, efficiency  $t_3$ );
- and if  $\mu_2$  triggers L2 but not L3, the  $\mu_3$  only needs to be identified but might trigger (efficiency  $l_3$ ).

Let us check the consistency with the other rotations of indices:

- $\mu_2$  cannot fire L3 in this term, because this case is covered by the (2, 3, 1) rotation, where  $t'_2$  is in factor;
- The case where  $\mu_3$  triggers L3 is not redundant with the permutation (3, 1, 2) where  $t'_3$  is in factor: in this term  $\mu_1$  never triggers L3 (only efficiencies  $(l_1 - t_1)$  and  $(t_1 - t'_1)$  will show).

The only sound way of checking the difference between this exact method and the simplistic method is to test them with the actual scale factors values. This is shown in Fig. 7.20, and the difference in the mean applied scale factor is about 6% in the PbPb case.

##### Inclusion–exclusion principle

Define the 6 events:

- $A_\ell = \mu_i L_0 \cap \mu_j L_2 \cap \mu_k L_3$

where  $\ell = 1 \dots 6$  stands for the permutation of indices. As before, one defines the probabilities  $l_k = P(\mu_k L_0)$ ,  $t_k = P(\mu_k L_2)$ ,  $t'_k = P(\mu_k L_3)$  and one has  $L_3 \subset L_2 \subset L_0$ .

We want to compute the probability that (at least) one of these events occurs, namely  $P(\cup_\ell A_\ell)$ , given by the inclusion–exclusion principle [203]:

$$P(\cup_\ell A_\ell) = \sum_{m=1}^6 (-1)^{m+1} S_m \quad (7.18)$$

where  $S_m$  is the sum, over all collections of precisely  $m$  events, of the probabilities that all these  $m$  events occur (for each  $m$ , there are  $\binom{6}{m}$  permutations), *i.e.*

$$\begin{aligned} S_1 &= \sum_{m=1}^6 P(A_m) \\ S_2 &= P(A_1 \cap A_2) + P(A_1 \cap A_3) + \cdots + P(A_5 \cap A_6) \\ S_3 &= P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_4) + \cdots + P(A_4 \cap A_5 \cap A_6) \\ \dots &= \dots \\ S_6 &= P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5 \cap A_6) \end{aligned}$$

It can be shown that

$$\begin{aligned} S_1 &= \{l t t'\} \\ S_2 &= \{l t' t'\} + \{t t t'\} + 3 \{t t' t'\} \\ S_3 &= 4 \{t t' t'\} + 8 \{t' t' t'\} \\ S_4 &= \{t t' t'\} + 12 \{t' t' t'\} \\ S_5 &= 6 \{t' t' t'\} \\ S_6 &= \{t' t' t'\} \end{aligned}$$

where  $\{\dots\}$  refers to all permutations  $(1, 2, 3)$ . Summing all the contributions according to (7.18), one obtains in the end

$$P(\cup_\ell A_\ell) = \{l t t'\} - \{l t' t'\} - \{t t t'\} + \{t' t' t'\}, \quad (7.19)$$

or in a less compact way,

$$\begin{aligned} P(\cup_\ell A_\ell) &= l_1 t_2 t'_3 + l_1 t_3 t'_2 + l_2 t_1 t'_3 + l_2 t_3 t'_1 + l_3 t_1 t'_2 + l_3 t_2 t'_1 \\ &\quad - l_1 t'_2 t'_3 - l_2 t'_1 t'_3 - l_3 t'_1 t'_2 \\ &\quad - t_1 t_2 t'_3 - t_1 t_3 t'_2 - t_2 t_3 t'_1 \\ &\quad + t'_1 t'_2 t'_3. \end{aligned}$$

This is equivalent to Eq. (7.17).

### 7.2.2.5 Check of the application of asymmetric SF with toys

One would like to confirm with a closure check on the signal MC that the procedure to apply the scale factors when the muon cuts are asymmetric is correct. Toy efficiency functions for data and MC are considered, for three telescopic cuts as in the full PbPb case. The goal is to start from these muon efficiencies, known both in data and MC, to then correct the MC with the procedure of section 7.2.2.1, and finally, to check that the trimuon efficiency resulting from the corrected MC is exactly equal to the trimuon efficiency from data. I thank Pedro GONZALEZ and Lizardo VALENCIA PALOMO from Sonora University for their help on performing this test.

The considered muon efficiency functions are inspired from the fit of the pp efficiencies for global muon identification (named  $l$  here), hybrid-soft ID (whose product with  $l$  is named  $t$ ),

and trigger (whose product with  $t$  is named  $t'$ ). They depend both on the muon  $p_T$  and  $\eta$ . The notations imply that  $l < t < t'$ . To have a larger data/MC discrepancy to test,  $0.2 \times (1 - t')$  is added to the MC efficiency  $t'$ , and the data efficiency  $l$  is multiplied by  $\frac{2}{\pi} \arctan(p_T[\text{GeV}]/0.5)$ .

We start, as for the standard efficiency measurement, by looping over the accepted trimuons in the signal MC sample (the pp sample is used, but it does not matter as the muon efficiencies are artificially inputted). We then throw three random numbers  $u_1$ ,  $u_2$ , and  $u_3$  between 0 and 1, associated to the three muons. The muon  $i$  passes the efficiency  $\varepsilon$  if  $u_i < \varepsilon(p_T(\mu_i), \eta(\mu_i))$ . This test is applied to the three muons, and for the three efficiencies, which emulates an actual event, where the interaction with the detector settles which muon passes what cuts. We now apply the trimuon cut as in the nominal procedure for the full PbPb case: the trimuon passes the cuts if all muons pass  $l$ , if at least two muons pass  $t$ , and if at least one muon passes  $t'$  (if this test is satisfactory, then it validates as well the simpler procedures in pp or with PbPb dimuons). If the trimuon passes the cuts, it is filled in the selected events histogram.

The efficiency is then the ratio of the selected events over the accepted ones. To increase the statistics, we run 100 times the random number generation, each leading to a trimuon cut test and the filling of the selected histogram; then this histogram is divided by 100.

This procedure is ran twice: once with the data muon efficiency functions (giving the true trimuon efficiency that we would like to reach), and once with the MC muon efficiency functions. In the latter case, the selected histogram is filled with a weight equal to the scale factor defined in Eq. 7.4, except that we replace  $SF \times \varepsilon_{MC}$  in the numerator by the corresponding data efficiency (which assumes that the extracted *single muon* scale factors are exact). We then compare the trimuon efficiency from data muon efficiency functions, to the trimuon efficiency from MC functions corrected with the asymmetric scale factors.

Fig. 7.24 shows the outcome of this toy procedure. The method closes very well: from a trimuon efficiency showing data/MC differences of order 100% (*bottom right* panel), we obtain a difference between data and the corrected MC significantly smaller than 1% (*bottom left* plot). This confirms in an independent way that applying the scale factors as explained in section 7.2.2.1 is correct.

### 7.3 One-binned acceptance and efficiency corrections

From the number of observed signal trimuons obtained with the template fit, to the corrected yield, a correction by the acceptance and efficiency defined in Eq. 7.1 is needed. The corrected yield is the actual number of  $B_c$  mesons produced and leaving the QGP; dividing it by the (equivalent) luminosity (see sections 3.5.1 and 3.5.5.2) results in the final cross section.

We evaluate the acceptance with a signal MC where no cut is applied on the generated  $B_c$  mesons, but the detector reconstruction can be skipped (see Table 3.1). The efficiency can be measured with an MC where events with muons out of the loose acceptance are cut off.

The one-binned method to measure the acceptance and efficiency correction is the most straightforward. For the efficiency, in a given analysis bin (*i.e.* one of the two  $p_T^{\mu\mu\mu}$  or centrality bins or integrated on the studied phase space), the passing trimuons (corrected by the scale factor of section 7.2.2.1) are counted and divided by the number of accepted muons. Similarly, for the acceptance, all accepted trimuons in this bin are counted and divided by the number of generated trimuons. The values of the acceptance, efficiency, and their product, are given for pp and PbPb in Table 7.2 of section 7.6.5, where they are compared to the values obtained after the first-step or second-step  $p_T^{\mu\mu\mu}$ -spectrum corrections of the MC. The acceptance only depends on the kinematics of the muon daughters of a generated  $B_c$ , meaning that in the first step it does not depend on the detector conditions (pp or PbPb); it will however differ for pp and PbPb in the second and third steps, due to different corrections of the  $p_T^{\mu\mu\mu}$  spectrum of MC.

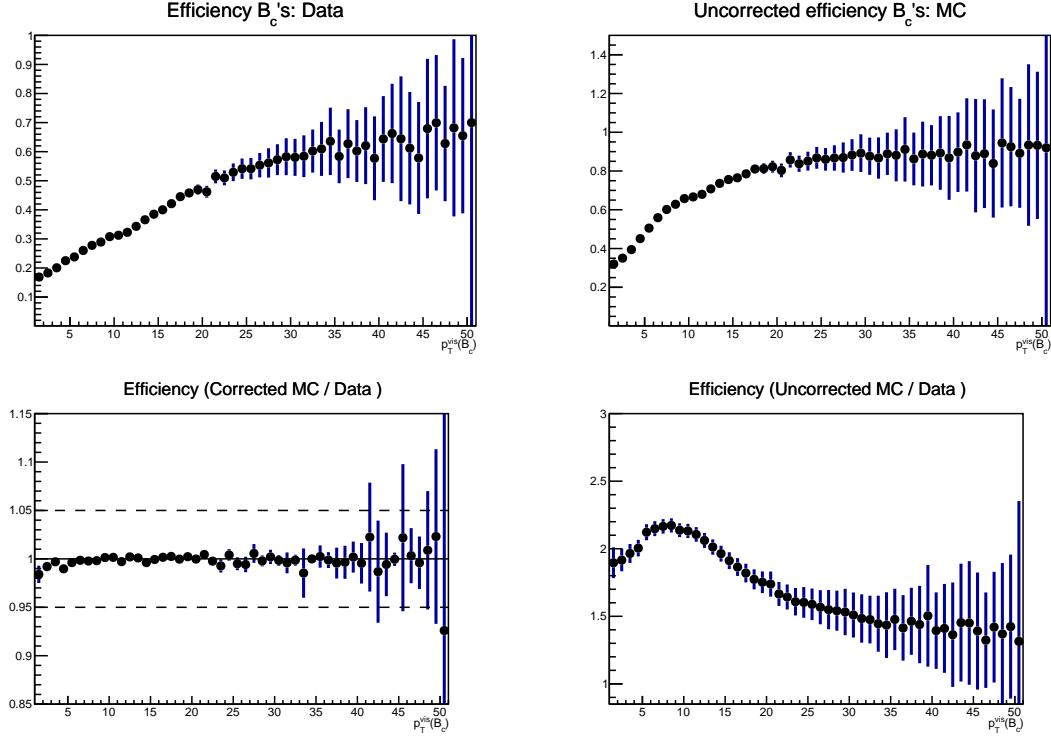


FIGURE 7.24: Closure test with toy muon efficiency functions, to check the application of scale factors when the muon cuts are asymmetric. *Top*: trimuon efficiency from data (*left*) and uncorrected MC (*right*), verus  $p_T^{\mu\mu\mu}$ . *Bottom right*: ratio of the uncorrected MC to the data efficiencies, *i.e.* the trimuon scale factor. *Bottom left*: ratio of the corrected MC to the data efficiencies, satisfyingly close to 1.

Calling  $k$  the kinematic variables on which the acceptance and efficiency depend, and  $f(k)$  the distribution of the generated  $B_c$  mesons in these variables, and bin  $i$  a given analysis bin, then the correction is the inverse of:

$$(\alpha\varepsilon)_{bin\,i} = \frac{\int_{bin\,i} dk f(k) \alpha\varepsilon(k)}{\int_{bin\,i} dk f(k)} \quad (7.20)$$

The issue with the one-binned method is that the bin correction depends strongly on the simulated signal kinematics  $f(k)$ .

The BCVEGPY2.2 simulation of the  $B_c$  production is the most complete one existing nowadays (see section 3.5.2.1), but even if it was exactly correct in pp (which can actually be checked by this measurement, within our uncertainties) it does not intend to describe the PbPb mechanisms. Therefore, one might have reasonable confidence in the  $B_c$  kinematic distributions in pp, but not in PbPb.

To illustrate the effect of possibly wrong distributions in the signal MC on the correction in PbPb, the  $p_T^{\mu\mu\mu}$  distribution of the signal MC is biased, and the acceptance and efficiency calculation is rerun. A  $\frac{1}{N}$  weight is attributed to all signal trimuons in MC. Fig. 7.25 then shows the ratio of the nominal  $\alpha \times \varepsilon$  to the correction using the MC samples biased with varied values of  $N$ . An assumed ratio of 2 between the two  $p_T^{\mu\mu\mu}$  bins of the  $R_{\text{PbPb}}$  corresponds to an exponent  $N = 1.1$  in the power law biasing, and results in a 15% change in the acceptance times efficiency values in both  $p_T$  bins. If this ratio is 3 (exponent  $N = 1.7$ ) the acceptance times efficiency is modified by about 20%.

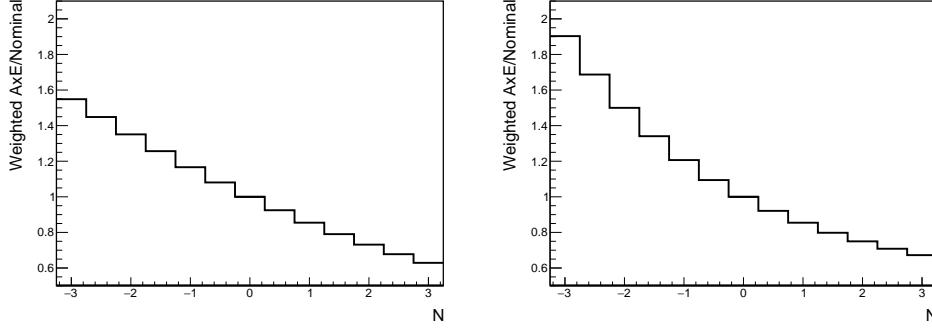


FIGURE 7.25: PbPb acceptance times efficiency with a  $\frac{1}{(p_T^{\mu\mu\mu})^N}$  biasing of the signal MC, divided by the nominal one, as a function of  $N$ . *Left:* first  $p_T$  bin. *Right:* second  $p_T$  bin.

The MC kinematics can therefore strongly impact the acceptance times efficiency corrections. The first idea to circumvent this was to apply event-by-event corrections, that take for each data trimuon the acceptance and efficiency value from a MC map read at the kinematics of this data trimuon. The maps are shown in section 7.4 and the event-by-event method in section 7.5. This method however increases too much the statistical uncertainties, and the solution is detailed in section 7.6: the two-steps method based on this simple one-binned correction method. The principle is to use a first estimate of the  $p_T^{\mu\mu\mu}$ -dependent cross sections to roughly correct the  $p_T^{\mu\mu\mu}$  distribution of MC, and to subsequently re-run the analysis (including the acceptance and efficiency correction) with the corrected MC. The final acceptance and efficiency is calculated from a MC that is corrected one more time, by the second-step measurements.

## 7.4 Acceptance and efficiency maps

To be independent of the signal MC kinematics, one would like to use the exact kinematics of data, by correcting each data candidate by the acceptance and efficiency corresponding to the kinematics of this given candidate. To that end, and to understand better how the correction depends on the kinematics, maps of the acceptance times efficiency versus the trimuon kinematics are built. They are shown in Figs. 7.26 and 7.27, and were calculated exactly as the one-binned corrections in the previous section, except in much finer bins. The black lines show the phase space and bins in which the cross sections are measured, chosen from these maps in section 5.6. The shown maps use the third-step signal MC, meaning after the final second-step correction of the  $p_T^{\mu\mu\mu}$  distribution (see section 7.6).

The binning of these maps was chosen so that, in the bulk of the fiducial phase space, a similar number of events is expected in each bin, which hence show similar and reasonable statistical uncertainties on the acceptance times efficiency estimation.

The smallest acceptance times efficiency reached within the fiducial cuts is  $1.7 \times 10^{-4}$  in PbPb and  $9 \times 10^{-4}$  in pp. This is a low value, but considering our low statistics (and the *bottom left* panels of Figs. 7.26 and 7.27, showing the expected number of signal events), it is very probable that no event with such kinematics (close to  $p_T = 6$  GeV and  $y = 1.3$ ) will appear in data. The acceptance times efficiency however depends a lot on the kinematics, which underlines both that correct kinematic distributions in MC are important if the one-binned method is used, and that the event-by-event corrections within an analysis bin will be widely spread. The latter implies that the statistical power of data events will be strongly diminished; this drawback made the two-steps strategy preferable.

For the event-by-event correction in PbPb, only the third BDT bin is considered, because the first two bins contain too much background to provide a reliable corrected signal yield.

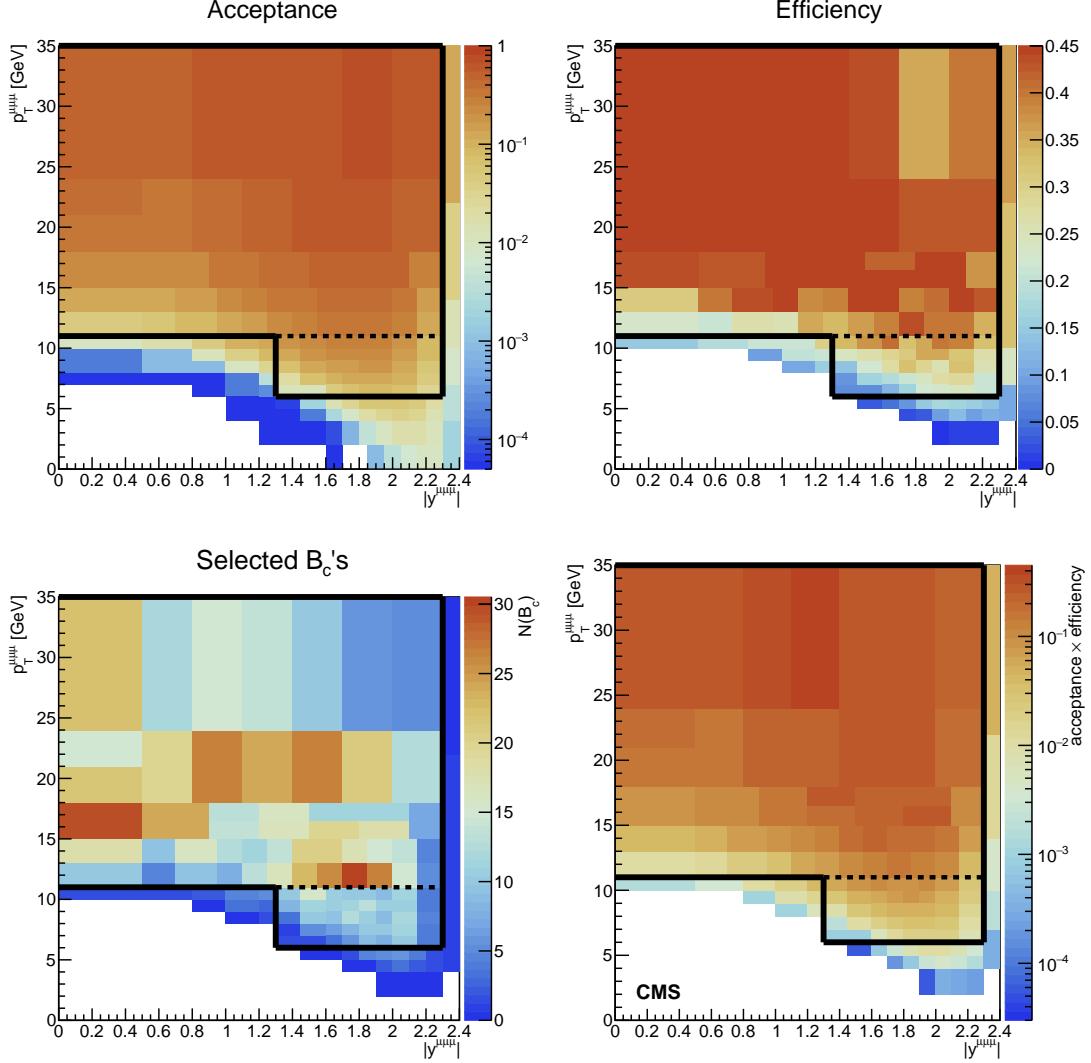


FIGURE 7.26: pp acceptance and efficiency maps of third-step signal MC trimuons, corrected by the scale factors of section 7.2.2.1, versus their transverse momentum and rapidity. *Top left:* acceptance. *Top right:* efficiency. *Bottom right:* acceptance times efficiency. *Bottom left:* expected number of signal trimuons in data (normalised with the  $p_T^{\mu\mu\mu}$ -corrected MC).

Therefore, the efficiency of a preselected event to be in the third BDT bin must be corrected for. Fig. 7.28 shows the corresponding map. This efficiency is around 35% by construction of the BDT binning, but shows a slight  $p_T^{\mu\mu\mu}$  dependence. The map of the efficiency of being in the second or third BDT bin was also produced.

## 7.5 Feasibility of event-by-event corrections

Correcting for acceptance times efficiency candidate-by-candidate allows to directly consider the kinematics of data, instead of the ones of MC. Usually, this method is rather straightforward to apply. First, each data candidate is weighted by the inverse of the acceptance times efficiency from the bin of the  $\alpha \times \varepsilon$  map corresponding to the kinematics of this candidate. Then, the signal extraction procedure is applied, and the obtained signal yields are hence already corrected.

Two things differ in our case: the necessity of running the fit before applying the correction, and of properly subtracting the background. Concerning the first point, I did not manage to

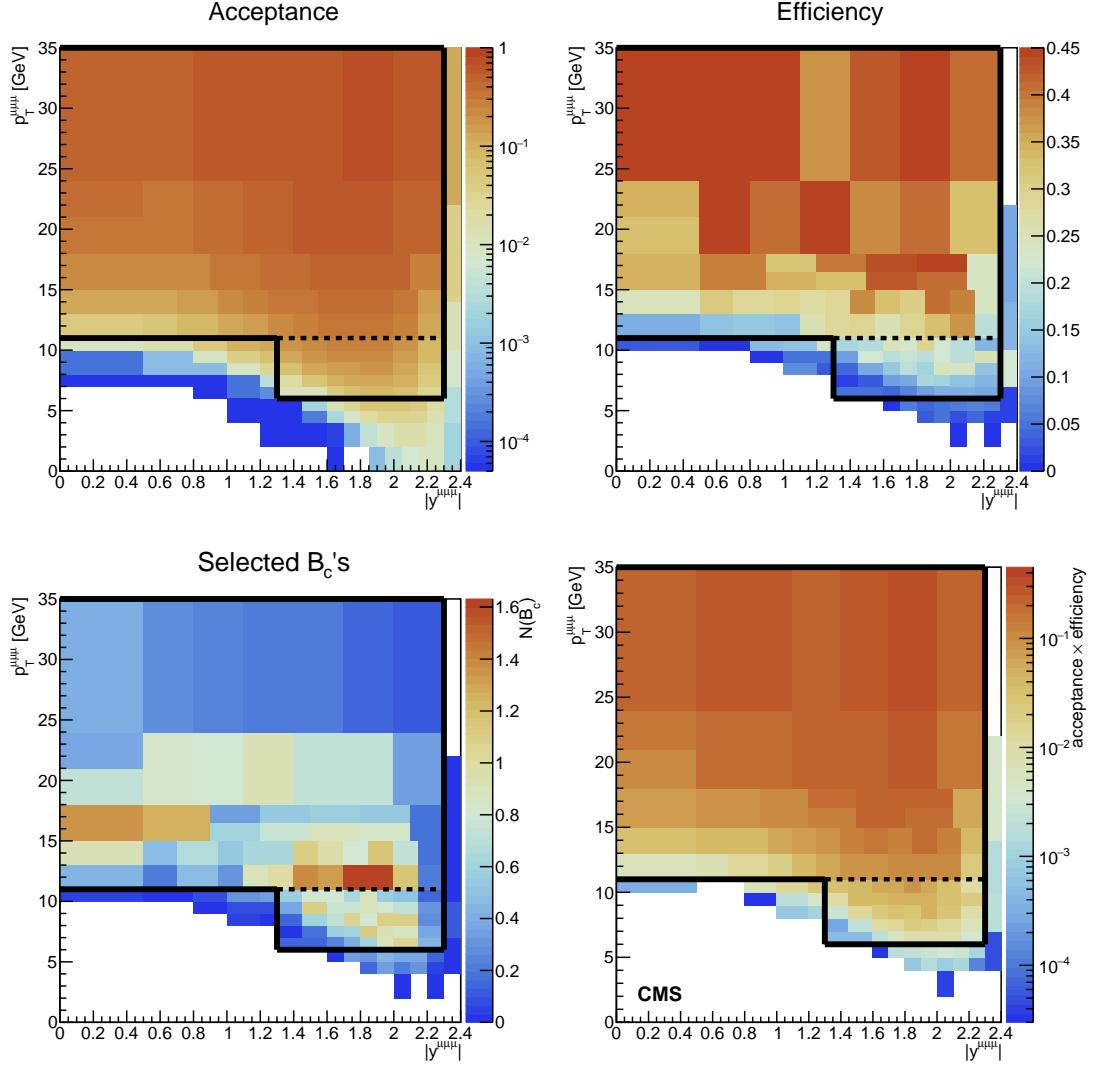


FIGURE 7.27: PbPb acceptance and efficiency maps of third-step signal MC trimuons, corrected by the scale factors of section 7.2.2.1, versus their transverse momentum and rapidity. *Top left:* acceptance. *Top right:* efficiency. *Bottom right:* acceptance times efficiency. *Bottom left:* expected number of signal trimuons in data (normalised with the  $p_T^{\mu\mu\mu}$ -corrected MC).

stabilise the template fit (chapter 6) when using trimuon mass templates in which the candidates were already weighted by  $1/\alpha\varepsilon$ . The statistical uncertainties on each mass bin in this case are too large (often close to 100%, due to some events having a much larger weight than the rest) for a reliable fit convergence. The corrections hence need to be applied after the fit.

However, after having extracted the signal yield in an analysis bin it is not known exactly which data candidates are from signal or background, so which individual candidates must be corrected. One cannot simply apply corrections for all data candidates (and then deduce an average correction per candidate) because the kinematics of background candidates, and hence their corrections, can be different than for signal candidates.

The trick is to correct all events, including background, and to correct as well the events from the background templates. Then the corrected background template events can be subtracted from the data corrected events, to obtain the corrected signal yield. Of course, the corrections are determined from the signal MC, so they do not adequately correct the background events for  $\alpha \times \varepsilon$ ; nevertheless the corrections are considered similarly wrong in the background template events

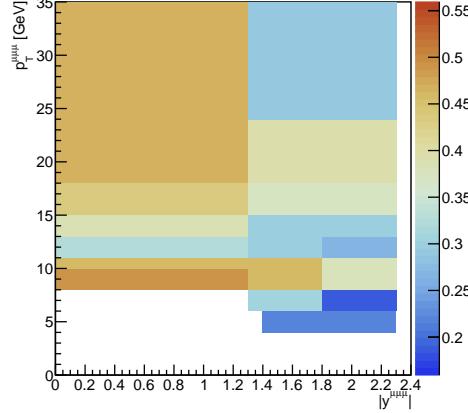


FIGURE 7.28: PbPb map for the efficiency of a preselected MC signal trimuon to be in the third BDT bin, versus its kinematics.

and in the background data events, ensuring a proper cancellation of the corrected background.

#### Bin-by-bin corrected background subtraction

The details of the background subtraction are however subtle. The template fit affects the proportions of the different background sources in each analysis, BDT, and mass bin; therefore the corrected background yield for a given source cannot be directly obtained by multiplying its fitted background normalisation by the sum of corrected candidates from the corresponding background sample. The driver of the values of the correction of a given candidate is its kinematics, which can be different for each background and in each analysis, BDT, and mass bin. Therefore, the background subtraction must be performed separately in each of these bins, for which the fit indeed gives the fraction of signal and of each background.

I call  $N_{b,\text{postfit}}(\text{bin}_i)$  the yield for a given background  $b$  ( $b \in 1, 2, 3$ ) in a mass bin  $\text{bin}_i$ , obtained from the fit. The corrected yield in this bin is  $N_{b,\text{postfit}}(\text{bin}_i) \times \mathcal{C}(b, \text{bin}_i)$ , where  $\mathcal{C}(b, \text{bin}_i)$  is the acceptance and efficiency correction for background  $b$  in  $\text{bin}_i$ . In a bin, the total prediction of the post-fit model might be different than the data yield; to get a more realistic estimate of the background corrected yield, we rescale the sum of  $N_{b,\text{postfit}}(\text{bin}_i)$  and of the signal post-fit yield to the content of this bin in data.

The correction can also depend on the systematic shape variation that is considered for a given background. Therefore, I redefine the corrections as  $\mathcal{C}(b, s, \text{bin}_i)$  for background  $b$ , undergoing a shape variation  $s$  ( $s = 0$  for the nominal shape,  $s = \pm 1$  for the varied shapes) in mass bin  $\text{bin}_i$ .  $\mathcal{C}(b, s, \text{bin}_i)$  is the yield for this background and this systematic variation in the mass bin  $\text{bin}_i$ , corrected candidate-by-candidate by  $1/\alpha\epsilon$  from the maps of Figs. 7.26 or 7.27, and divided by the uncorrected yield. If the uncertainty on  $\mathcal{C}(b, s \neq 0, \text{bin}_i)$  is larger than the one on  $\mathcal{C}(b, s = 0, \text{bin}_i)$  and larger than  $|\mathcal{C}(b, s \neq 0, \text{bin}_i) - \mathcal{C}(b, s = 0, \text{bin}_i)|$ , then I set  $\mathcal{C}(b, s \neq 0, \text{bin}_i) = \mathcal{C}(b, s = 0, \text{bin}_i)$ . Optionally, we limit the size of the correction to  $500/k$ , where  $k \in 1, 2, 3$  is the BDT bin number. If a candidate has a higher correction, the correction is set to this limit. This affects only a few events (mostly in the background-enriched BDT bin,  $k = 1$ ), that make the correction fluctuate too much in some mass bins. This is applied only in the PbPb first  $p_T^{\mu\mu\mu}$  bin where these fluctuations are the largest (and lead to no significant signal).

Now I introduce  $S_b$ , the continuous nuisance parameters leading the shape morphing of background  $b$  (particular cases are: only the nominal shape is used when  $S_b = 0$ , and only the shape  $s = \pm 1$  is used when  $S_b = \pm 1$ ). I force this to be a probability by using  $p_{b,s \neq 0} =$

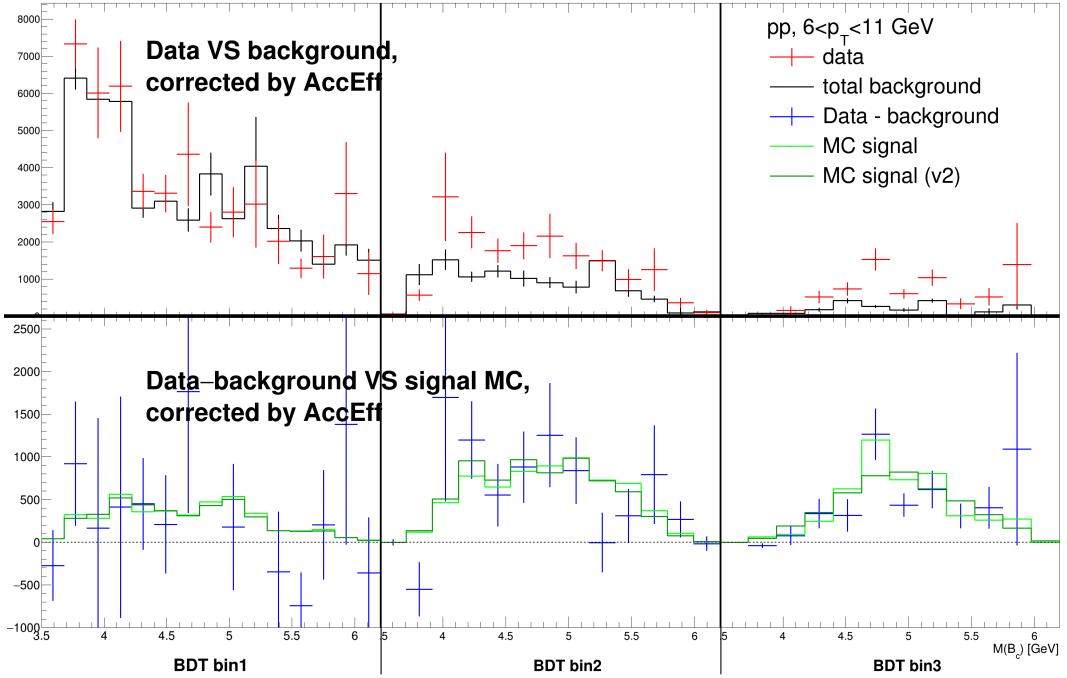


FIGURE 7.29: Determination of the signal yield corrected candidate-by-candidate for the acceptance and efficiency, including the subtraction of the post-fit background, in the first  $p_T^{\mu\mu\mu}$  bin in pp. The *left*, *middle*, and *right* panels show the three BDT bins. *Top*: corrected data (red) and summed backgrounds (black). *Bottom*: Corrected signal yield, from the background-subtracted data (blue) or from signal MC (two versions in light and dark green).

$\min(1, |S_b|)$ . The corrected yield of background  $b$  is finally:

$$B_{b,corr}(bin_i) = N_{b,postfit}(bin_i) \times [(1 - p_{b,s \neq 0}) \mathcal{C}(b, s = 0, bin_i) + p_{b,s \neq 0} \mathcal{C}(b, s = sign(S_b), bin_i)] \quad (7.21)$$

For backgrounds that have two such shape morphing parameters  $S_b$  and  $S'_b$ , with similar definitions the corrected yield of background  $b$  is:

$$\begin{aligned} B_{b,corr}(bin_i) = N_{b,postfit}(bin_i) \times & [(1 - p_{b,s \neq 0})(1 - p_{b,s' \neq 0}) \mathcal{C}(b, s = 0, s' = 0, bin_i) + \\ & p_{b,s \neq 0}(1 - \frac{p_{b,s' \neq 0}}{2}) \mathcal{C}(b, s = sign(S_b), s' = 0, bin_i) + \\ & (1 - \frac{p_{b,s \neq 0}}{2})p_{b,s' \neq 0} \mathcal{C}(b, s = 0, s' = sign(S'_b), bin_i)] \end{aligned} \quad (7.22)$$

The two above formulas give back a simple  $B_{b,corr}(bin_i) = N_{b,postfit}(bin_i) \times \mathcal{C}(b, s^{(')} = 0, bin_i)$  when the background exactly has the nominal shape ( $p_{b,s^{(')} \neq 0} = 0$ ). There is actually little difference in the results when we assume  $p_{b,s \neq 0} = 0$  for all backgrounds, because the kinematics (and hence the corrections) of the systematic background variations have only small differences with the nominal background.

I similarly define  $\mathcal{C}(data, bin_i)$  and  $\mathcal{C}(sig, bin_i)$ , the corrections in each mass bin for the data and signal MC samples. The corrected signal yield in  $bin_i$  is then:

$$S_{corr}(bin_i) = N_{data}(bin_i) \mathcal{C}(data, bin_i) - B_{b,corr}(bin_i) \quad (7.23)$$

Finally, the corrected signal yield for an analysis bin is the sum of  $S_{corr}(bin_i)$  where  $bin_i$  runs over the BDT and mass bins.

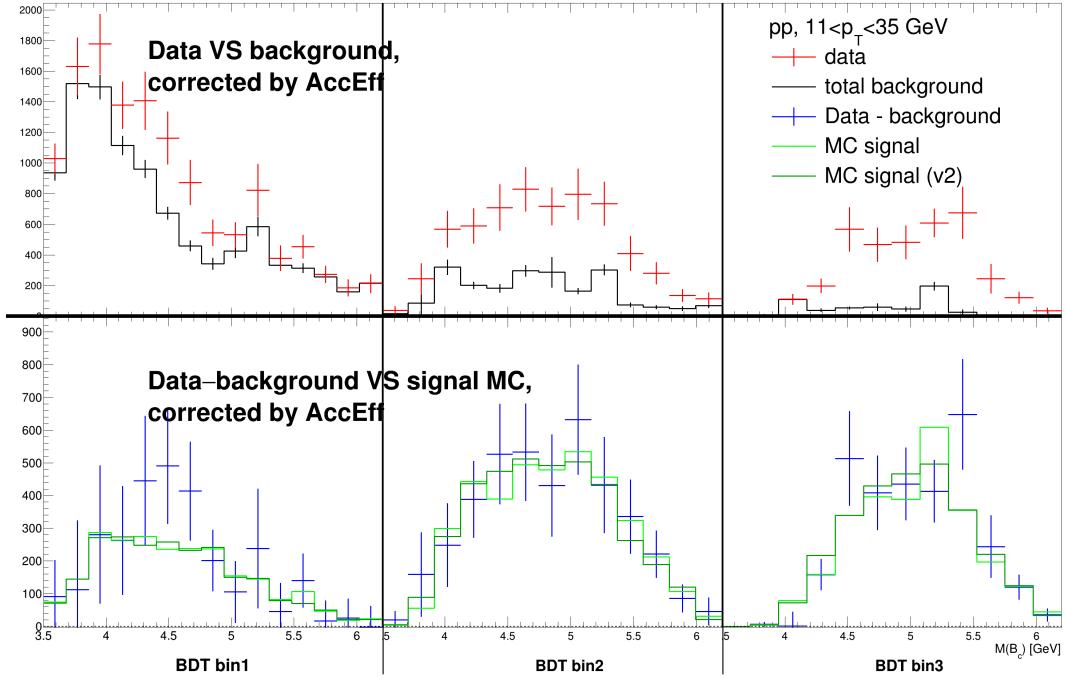


FIGURE 7.30: Determination of the signal yield corrected candidate-by-candidate for the acceptance and efficiency, including the subtraction of the post-fit background, in the second  $p_T^{\mu\mu\mu}$  bin in pp. The *left*, *middle*, and *right* panels show the three BDT bins. *Top*: corrected data (red) and summed backgrounds (black). *Bottom*: Corrected signal yield, from the background-subtracted data (blue) or from signal MC (two versions in light and dark green).

We also run this procedure on the signal MC sample as a crosscheck. In  $bin_i$ , the signal yield from MC is  $N_{sig,postfit}(bin_i) \times \mathcal{C}(sig, bin_i)$ , where  $N_{sig,postfit}$  is the post-fit yield (after rescaling the total signal+background prediction to data yields), and  $\mathcal{C}(sig)$  is the ratio of the sum of acceptance-and-efficiency-weighted signal MC candidates to the unweighted sum. In a second version of this MC crosscheck,  $N_{sig,postfit}$  is not rescaled with the full model to the data content. This crosscheck allows to see the result of the method with large statistics. The signal MC kinematics are corrected with the first-step measured  $p_T^{\mu\mu\mu}$  distribution (contrarily to the final  $\alpha \times \varepsilon$  that is calculated with the second-step-corrected MC).

Figs. 7.29 to 7.32 illustrate the method for the two  $p_T^{\mu\mu\mu}$  bins in pp and PbPb, by showing for each  $p_T^{\mu\mu\mu}$ , BDT, and mass bin, both the corrected data and summed background, and their difference, that is the corrected signal yield. The full signal yield for an analysis bin is the sum of this corrected yield for all BDT and mass bins. The two MC crosschecks are also shown, the difference between them showing the effect of rescaling the total prediction to the actual data content in each bin. The difference between the corrected signal yield from the background-subtracted data and from the signal MC mostly comes from the statistical fluctuations of data, but can partly come from the kinematics of the MC which lacks the second-step  $p_T$  spectrum correction.

In PbPb the first BDT bins have too many fluctuations to be included in the corrected signal yield estimation. Therefore, I consider either only the third BDT bin, or the second and third BDT bin, and I then correct the obtained yield for the efficiency of having a BDT value in these bins. This efficiency is taken either from an efficiency integrated on MC, or from a map like Fig. 7.28, but this does not change significantly the resulting yield. The latter is significant only in the second  $p_T$  bin. In the first  $p_T$  bin, isolated background events with large corrections

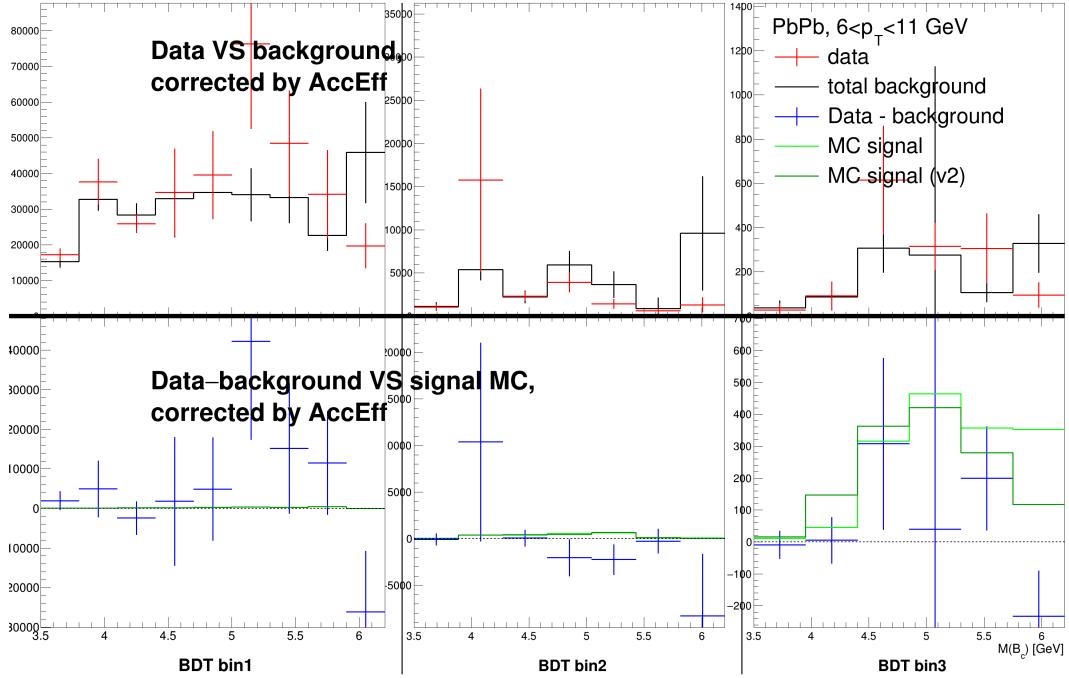


FIGURE 7.31: Determination of the signal yield corrected candidate-by-candidate for the acceptance and efficiency, including the subtraction of the post-fit background, in the first  $p_T^{\mu\mu\mu}$  bin in PbPb. The *left*, *middle*, and *right* panels show the three BDT bins. *Top*: corrected data (red) and summed backgrounds (black). *Bottom*: Corrected signal yield, from the background-subtracted data (blue) or from signal MC (two versions in light and dark green). In this  $p_T^{\mu\mu\mu}$  bin, no significant signal is observed after corrections, so a cutoff of the largest corrections is applied to reduce the fluctuations.

cause fluctuations larger than the expected signal – despite indulging in limiting the maximum single-candidate correction.

### Comparison of methods

The corrected signal yields obtained from the various methods were compared: the nominal event-by-event corrections (only in pp), the version where only the third (or second and third) BDT bins are considered (the only way to get significant PbPb yields), and the one-binned method of section 7.3. The event-by-event method results in large uncertainties (mainly from the inhomogeneous  $1/\alpha \times \epsilon$  corrections to the candidates within a bin), and no significant yield in the first  $p_T^{\mu\mu\mu}$  bin in PbPb, so this method was abandoned. This outcome was already clear with blinded data, so this decision was taken before PbPb data unblinding. The comparison plot, only relevant in pp where the event-by-event yields are significant, is not shown here (it is not judged informative, considering the many caveats). The yields from the various event-by-event methods are consistent with those from the one-binned method, within large uncertainties.

The comparison is slightly flawed by the uncertainty calculation for the event-by-event method: it was calculated as the quadratic sum of the uncertainties in each bin, hence neglecting the correlations between the corrected yields in the different BDT and mass bins. To have a proper estimate of these uncertainties, one could run the MC crosscheck with only a fraction of the statistics (representative of the data statistics), generated many times. On each small MC dataset, the event-by-event method could be run, and the distribution of the results would be an estimate of the expected uncertainty. However, as this method is not kept as the nominal one, this proper treatment was not implemented.

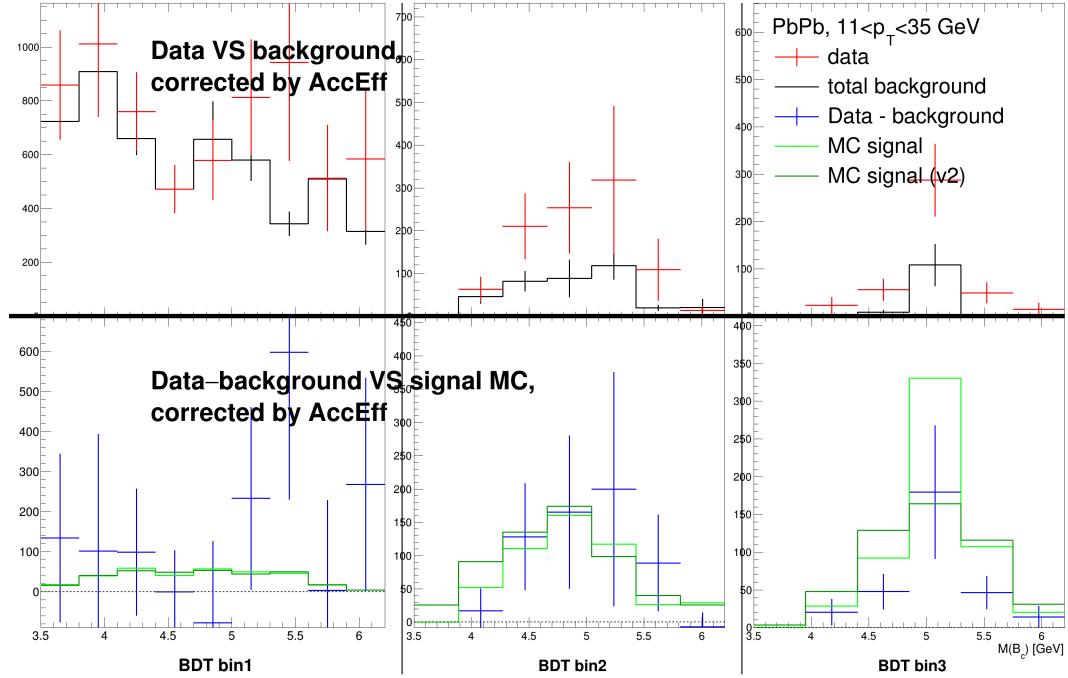


FIGURE 7.32: Determination of the signal yield corrected candidate-by-candidate for the acceptance and efficiency, including the subtraction of the post-fit background, in the second  $p_T^{\mu\mu\mu}$  bin in PbPb. The *left*, *middle*, and *right* panels show the three BDT bins. *Top*: corrected data (red) and summed backgrounds (black). *Bottom*: Corrected signal yield, from the background-subtracted data (blue) or from signal MC (two versions in light and dark green).

As the event-by-event method proved not feasible, let us now rather improve the one-binned method by correcting the distributions of MC with a first estimate of the  $p_T^{\mu\mu\mu}$  distribution from the data measurement, as explained in section 7.6.

## 7.6 Two-steps procedure

I implemented two methods to correct the observed signal yields (from the fit of chapter 6) for the effect of the acceptance and efficiency of reconstructing, selecting, and triggering signal trimuons. The most straightforward, named *one-binned* (section 7.3), is simpler and keeps all the statistical power of the extracted yield; but its drawback is that it relies on the kinematic distributions of the signal MC. Wrong MC kinematics bias a lot the corrections if the acceptance and efficiency depend on the kinematics, which is strongly the case (see the maps of Figs 7.26 and 7.27). Beyond the fact that the  $B_c$  kinematics in pp might not be perfectly simulated, the QGP effects on  $B_c$  production are not considered in the PbPb MC and influence the kinematics. Using the kinematics of data directly, by correcting separately each data event, has been implemented (section 7.5), but in a non-standard way because the corrections are applied after the fit, making the background subtraction relatively sophisticated; moreover, it strongly reduces the statistical power of the yields, resulting in too large uncertainties.

Therefore, I rather correct the kinematics of the signal MC, up to the knowledge of the spectra from preliminary analyses, before applying the one-binned method. The acceptance and efficiency depends more on  $p_T^{\mu\mu\mu}$  than on rapidity, and the modification of the  $B_c$  production (which increases the difference between the PbPb data and MC spectra) is expected to depend more on  $p_T$  than rapidity, so the  $p_T^{\mu\mu\mu}$  distribution of MC is corrected rather than the rapidity. This correction is run both at the end of the first-step and at the end of the second-step analysis

(see Fig. 3.13). In the first step, only the correction of the  $p_T^{\mu\mu\mu}$  spectrum is needed for running the second-step analysis (*i.e.* the part of the procedure linked to the acceptance and efficiency uncertainty is run only at the end of the second step). A simplified sketch of the procedure to extract the uncertainties from the  $p_T^{\mu\mu\mu}$  spectrum correction is presented in Fig. 7.33.

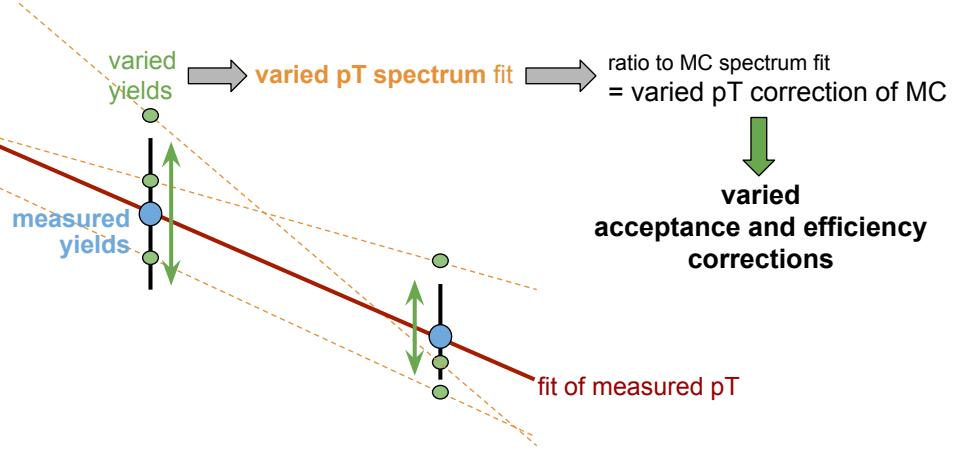


FIGURE 7.33: Extraction of the uncertainties on the acceptance and efficiency corrections from variations of the measured  $p_T^{\mu\mu\mu}$  spectrum.

The two-steps procedure is the following:

- Some assumption on the spectrum shape (used as fit function) is made, with the help of a fit of the signal MC spectrum.
- I fit with this assumed shape the corrected yields measured in two coarse  $p_T^{\mu\mu\mu}$  bins in a first iteration of the analysis. The yields are divided by the  $p_T^{\mu\mu\mu}$  and rapidity bin widths. I use for this measurement one-binned  $\alpha \times \varepsilon$  corrections resulting from the uncorrected signal MC, or from the first-step-corrected one (for the second-step  $p_T^{\mu\mu\mu}$  spectrum fit).
- (Only in the second-step run:) the corrected yields in the two bins are varied within what the uncertainties from the second-step measurement allow. The uncertainty comes from the fit (sampling the fit likelihood), from the fit method variations, and from the tag-and-probe (via Gaussian-distributed random numbers multiplied by the uncertainties). The uncertainties from the luminosity and other  $B_c$  decays are excluded because they are small and fully correlated along  $p_T^{\mu\mu\mu}$ , so they do not affect significantly the shape of the  $p_T^{\mu\mu\mu}$  spectrum. The uncertainty from  $\alpha \times \varepsilon$  corrections is to be determined from this method, so it is not included. The correlations between the uncertainties on the two  $p_T^{\mu\mu\mu}$  bins are taken into account when varying the two data points.
- (Only in the second-step run:) 500 variations are run for each shape assumption. Each two-points toy measurement is fitted, resulting in 500  $p_T^{\mu\mu\mu}$  spectra per shape assumption.
- For each variation (only in the second-step run) and for the nominal, the fit to the (toy or true) measurement is divided by the fit of the MC with the same shape assumption (to eliminate most of the potential error in the choice of the fit function). This ratio is the correction applied to the  $p_T^{\mu\mu\mu}$  distribution of MC.
- For each of the  $p_T^{\mu\mu\mu}$  spectrum corrections determined from the varied (only in the second-step run) and the nominal two-points measurements, the one-binned acceptance and efficiency are computed with the resulting corrected MC.
- In each analysis bin ( $p_T^{\mu\mu\mu}$ , centrality, or integrated), the nominal acceptance and efficiency correction is determined from the MC whose spectrum was corrected with the fit

of the nominal (non-varied) measurement. This means the  $\alpha \times \varepsilon$  correction is calculated a third time, after the second-step  $p_T^{\mu\mu\mu}$ -correction of MC, to provide the final nominal value. (Only in the second-step run:) the spread of the acceptance and efficiency (or directly of the corrected yields) values from the variations of the  $p_T^{\mu\mu\mu}$ -differential measurement (cf Fig. 7.36 and 7.37) results in the final uncertainty on acceptance and efficiency.

### 7.6.1 Lafferty-Wyatt: how to place the abscissa

To fit the measurement in two coarse  $p_T^{\mu\mu\mu}$  bins, one needs an appropriate prescription for placing the points along the  $p_T^{\mu\mu\mu}$  axis. Intuitively, a hypothetical continuous version of the spectrum should coincide with the two points. The center of the bin or the average  $p_T^{\mu\mu\mu}$  within a bin do not satisfy this requirement. One instead needs to follow the prescription from LAFFERTY and WYATT [204], which consists in placing the point  $x_{i,LW}$  such that:

$$f(x_{i,LW}) = \int_{bin\ i} f(x) dx \quad (7.24)$$

where  $f$  is the true spectrum, and the integral is performed within the limits of the studied bin.

This is the points are placed along  $p_T^{\mu\mu\mu}$  for this study, and for the final plots of the cross-section and modification factor. For this study, these abscissa positions are though only set for display purposes, as the fit of the two points is performed using only the integrals of the spectrum over each bin range. For the results plots, the abscissa are extracted with  $f$  being the fit of the nominal two-points measurement.

### 7.6.2 ‘Fitting’ two points

The true spectrum must be inferred from a measurement that features only two points. Literally any function with two parameters could exactly fit this. I thus need a justified assumption on the fit function, taken from fitting the continuous spectrum of MC. However, we need to limit this function to only two parameters, in order not to have more parameters to determine than the number of independent data points. I then ‘fit’ two parameters on two points, which amounts to solving a system of two equations with two unknowns, and gives an exact result (with reasonable functions).

As suggested by the Lafferty-Wyatt prescription, this system of equations involves the integrals of the spectrum in the two bins, which should be equal to our measured points  $y_i$  – those are indeed estimations of the integrals of the produced signal events over the respective bin ranges. So the parameters  $\{a_i\}$  of function  $f$  are determined by solving these equations (one per bin  $i$ ):

$$\int_{bin\ i} f_{\{a_i\}}(p_T^{\mu\mu\mu}) dp_T^{\mu\mu\mu} = y_i \quad (7.25)$$

Parts of this equation system can be done analytically: *e.g.* the integral of the power law (function mentioned later) is simple. However, the complete resolution is not feasible analytically, so I look numerically for the solutions of this system with the `HybridSJ` algorithm of the `GSLMultiRootFinder` class of ROOT. It is based on the GNU Scientific Library.<sup>1</sup> The fits of some variations are failing, but tweaking the initial values of the parameters for these cases makes all of them converge (with only 2 exceptions for which the toy measurement is randomly drawn again).

---

<sup>1</sup>see <https://www.gnu.org/software/gsl/doc/html/multiroots.html> for details on the algorithm.

### 7.6.3 Choice of fit function

An intuitive way to fit a spectrum with a two-parameter spectrum is a power law  $f_1(p_T) = a p_T^{-n}$ . However this does not fit well the continuous spectrum of MC (see the dashed green thick line in Figs. 7.34 and 7.35); it is still kept as a possible fit function, because it has only two parameters, and dividing two such functions (the fit on data and the fit on MC) should cancel out most of the modelling error.

The power law is a straight line in a log-log display, but the curved shape of the spectrum invites to try a function that is quadratic when displayed in a log-log canvas:

$$f_2(p_T^{\mu\mu\mu}) = N (p_T^{\mu\mu\mu})^{n+m \times \ln(p_T^{\mu\mu\mu})} \quad (7.26)$$

This function fits very well the MC spectrum (see the dashed thick red line of Figs. 7.34 and 7.35). However, it has three parameters, so I need to fix one parameter, not to have more unknowns than independent equations. I choose to fix  $m$  to the value obtained in the fit of the MC spectrum, but fixing  $n$  was tried too and the change in the data/MC ratio is of the same order than changing the model to other fit functions.

Another possibility is to use the so-called KAPLAN spectrum, that has more physical grounds than the log-log-quadratic function:

$$f(p_T^{\mu\mu\mu}) = \frac{N}{(1 + (p_T^{\mu\mu\mu}/a)^2)^n} \quad (7.27)$$

It also fits very well the MC spectrum, except a possible small deviation at high  $p_T^{\mu\mu\mu}$ , that mostly cancels in the data/MC ratio. It is shown in violet in Figs. 7.34 and 7.35. Here again, to fit the two data points, I fix one parameter to the one found in the fit of the continuous MC spectrum. I fix  $a$ , but when fixing  $n$ , the changes are of the same order than when the fit function is changed.

A satisfying check is that, using the last two functions (whichever parameter is fixed), the fit of the continuous MC spectrum gives parameters very close to the fit on the two-points MC spectrum (integrated on the bin ranges). The small difference between the fit of a continuous spectrum and of two points comes from the definition of the abscissa of the two points: they should be the  $x_{LW}$  corresponding exactly to the true spectrum, whereas they are actually defined on the pre-fit spectrum. This test is not satisfying in the power law case because of the modelling error.

### 7.6.4 Fit of toys (varied measurements)

This concerns only the procedure run at the end of the second-step analysis. I hence have three different functions, and for each of them, 500 toy measurements (obtained by varying the second-step measurement within uncertainties) are fitted, along with the nominal second-step measurement and the MC. The ratio between the fits of the second-step measurement and the MC is close to 1, because a correction was already applied to MC in the first step. Figs. 7.34 and 7.35 show these fits, for the MC in dashed lines, for the nominal measurement in solid lines, and for the variations in thin dash-dotted lines (only 40 out of 500 per function type are displayed, for readability). Three colour palettes are used respectively for the three fit function types. The bottom pad shows the ratio of the fit of the nominal or varied measurements to the fit of the MC spectrum. Similar plots exist for the first-step correction, with higher data/MC difference and without the variations.

The spectrum of the first-step-corrected MC (whose fit gives the reference MC function) is also shown. In practice, it is the signal MC without any acceptance nor selection cuts. An interesting feature to notice here is the small discontinuity at the frontier of the two  $p_T^{\mu\mu\mu}$  bins,

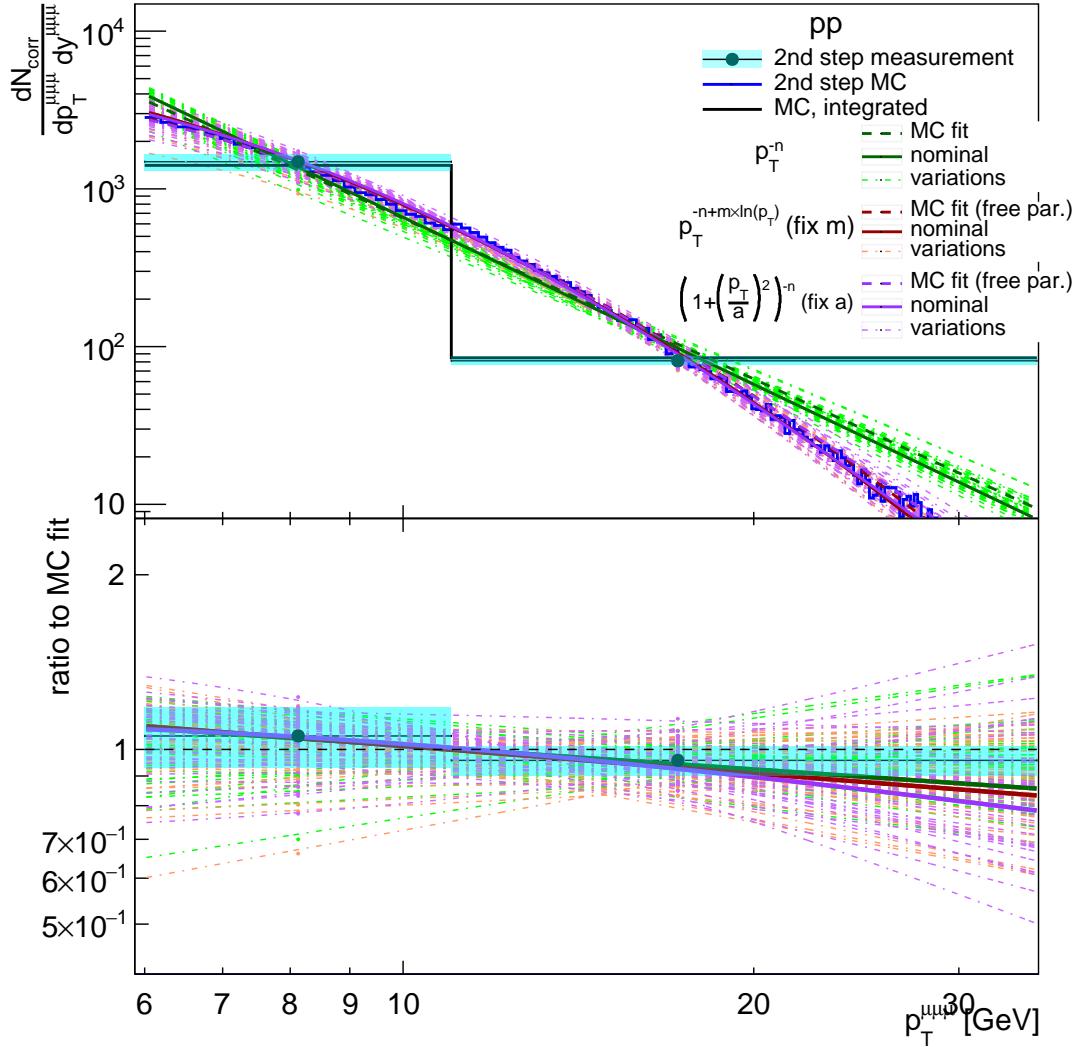


FIGURE 7.34: Nominal and variations of the second-step correction of the  $p_T^{\mu\mu\mu}$  spectrum of signal MC in pp. *Top:* fits of the yields corrected by first-step-corrected one-binned  $\alpha \times \varepsilon$ , differential in  $p_T^{\mu\mu\mu}$  and  $y^{\mu\mu\mu}$ . *Bottom:* ratio of the fit of measured corrected yields to the fit of the spectrum of the first-step-corrected signal MC. The shape is first fitted (dashed lines) on the continuous MC (solid blue histogram), and then the two shape parameters are estimated from the two measured points. Three different function shape assumptions are used (in red, green, and violet). The thin dash-dotted lines show for each function type 40 (out of the 500 actually run) fits to two-bin toy measurements, from varying the data points within their uncertainties. The thick line for each shape is the fit of the nominal (non-varied) measurement.

which points to the difference in rapidity range used for the two  $p_T^{\mu\mu\mu}$  bins: at lower  $p_T^{\mu\mu\mu}$ , the range  $1.3 < y^{\mu\mu\mu} < 2.3$  is used, where a lower differential yield is expected than in the more central rapidity range  $0 < y^{\mu\mu\mu} < 2.3$  that is used for the higher  $p_T^{\mu\mu\mu}$  bin. This discontinuity is small enough to be neglected compared to the dominant uncertainties of the fitted measurement.

### 7.6.5 Resulting nominal value and uncertainty

For each of the three fit methods, I have now 500 varied functions (ratio of the toy fits to the MC fit) and one nominal function to correct the  $p_T^{\mu\mu\mu}$  spectrum of the original (or first-step-corrected) signal MC. I hence run (in the second step), both in pp and PbPb, 1503 computations

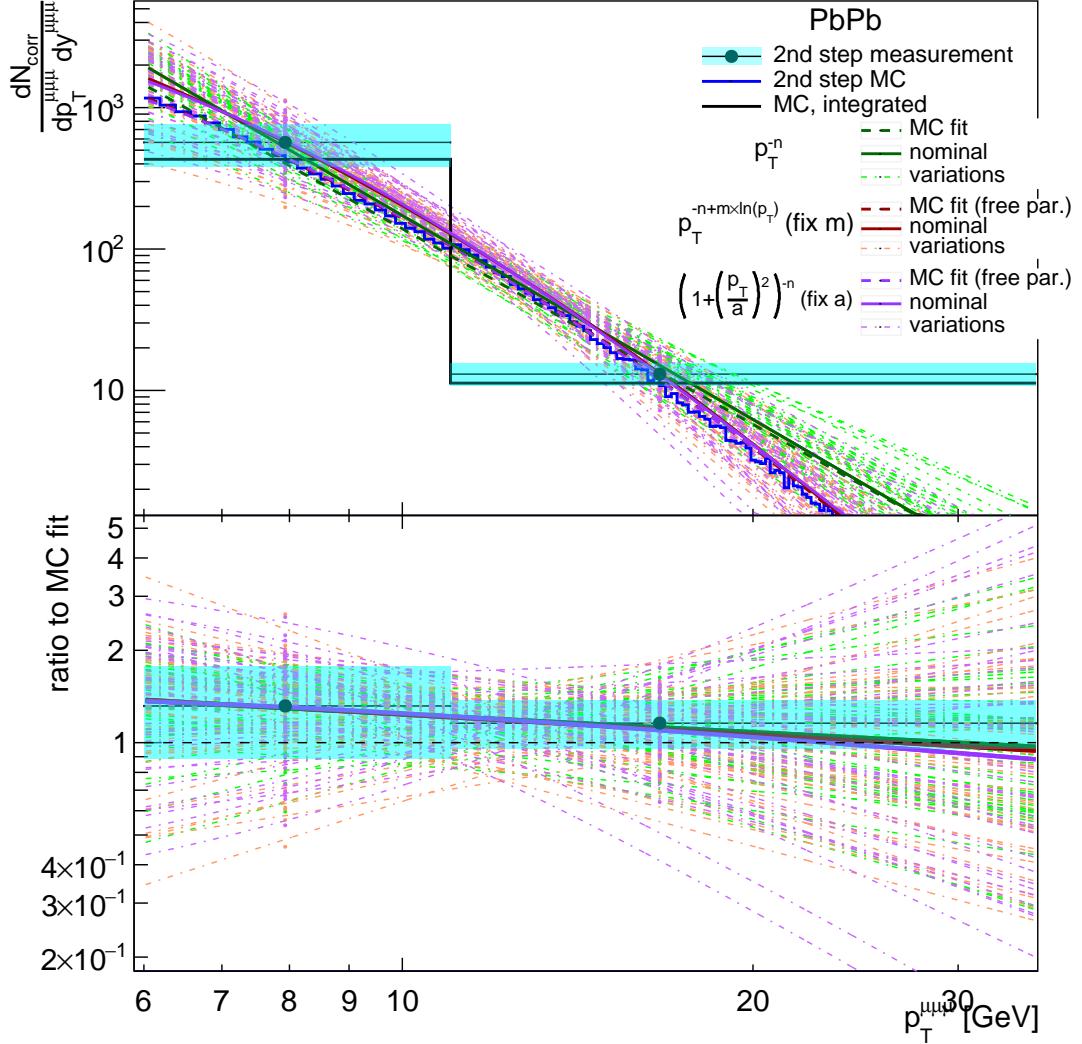


FIGURE 7.35: Nominal and variations of the second-step correction of the  $p_T^{\mu\mu\mu}$  spectrum of signal MC in PbPb. See caption of Fig. 7.34 for details.

of the acceptance and of the efficiency, performed as in section 7.3 but with the MC corrected by one of the  $p_T^{\mu\mu\mu}$  spectrum function. The acceptance previously depended only on the kinematics of the generated MC events and was the same in pp and PbPb, but now the MC is different in the two cases, so both are run separately. In the first step, only one acceptance and efficiency value is calculated, with the MC corrected with one of the  $p_T^{\mu\mu\mu}$  fit method (the one quadratic in log-log and where  $m$  is fixed to the value in the continuous MC spectrum).

This results in a distribution of the  $\frac{1}{\alpha \times \varepsilon}$  corrections representative of the uncertainty of the  $p_T^{\mu\mu\mu}$  spectrum correction, which follows from the uncertainty on the second-step measurement. Fig. 7.36 shows these distributions for the integrated samples and the two centrality bins in PbPb, separated for the three function types. The coloured dashed vertical lines show the value resulting from the  $p_T^{\mu\mu\mu}$  spectrum corrected by the fit of the nominal second-step measurement. They are much closer to each other than the spread due to varying the fitted measurement, which is a sign of stability of the procedure. Those values, as well as the spread, can be compared to the value from the simple one-binned procedure without correcting the  $p_T^{\mu\mu\mu}$  spectrum of MC, or with the first-step correction. The separation of the value associated to the original MC from the new nominal is larger than the RMS of the distribution of the toys, which means that this is a useful correction. The difference between the second-step value and the first-step value is

much smaller than the effect of the first-step correction on the original MC, indicating that the procedure converges (see section 7.6.7 for more on this convergence).

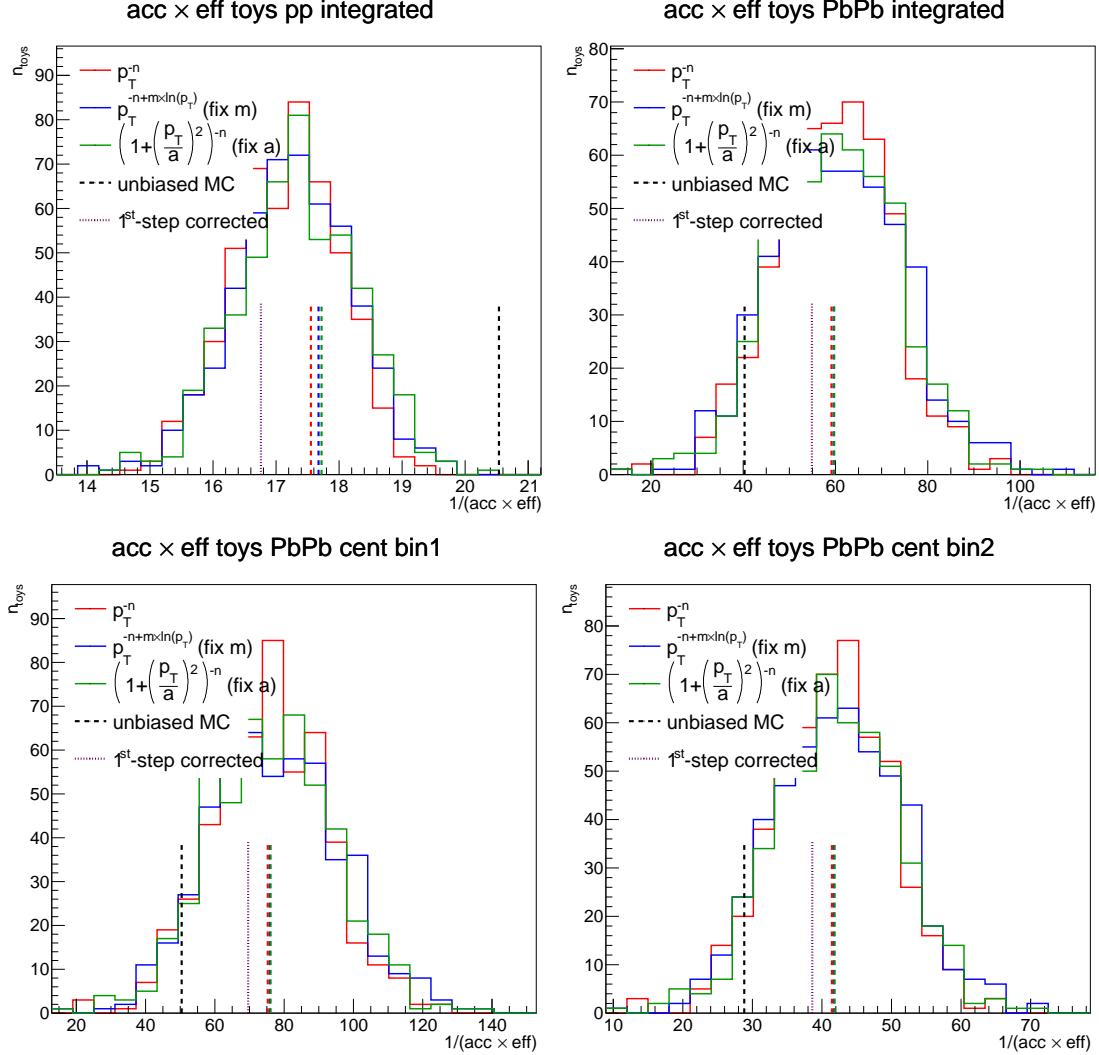


FIGURE 7.36: Variations of the second-step data measurement are associated to variations of the  $p_T^{\mu\mu\mu}$  spectrum correction in MC. Here is shown in the integrated sample in pp (*top left*) or PbPb (*top right*) or in the two PbPb centrality bins (*bottom left and right*) the effect of these variations on the inverse of the acceptance times efficiency which will multiply the post-fit yields. The coloured dashed lines are the nominal results from the three fit functions (and no variation of the data points), compared to the result without any correction of the  $p_T^{\mu\mu\mu}$  spectrum (dashed black line) or with the first-step correction (dotted line).

The spread (RMS) of this distribution of acceptance and efficiency values will be quadratically summed to the other uncertainties for the centrality-binned and the integrated results. However, for the  $p_T^{\mu\mu\mu}$ -differential measurement, one can and should take into account the correlation between the variation of the  $p_T^{\mu\mu\mu}$ -dependence measurement (the 1500 toy two-binned measurements) and the associated variation of acceptance and efficiency (the 1500  $\alpha \times \varepsilon$  values extracted from the MC corrected with each of these varied fitted  $p_T^{\mu\mu\mu}$  spectra). This is done by calculating the final uncertainty directly as the spread of the varied corrected yields, which are the products of the varied yields times the associated varied corrections. Only the uncertainties from luminosity and other  $B_c$  channels (that were not included for the variation of the  $p_T^{\mu\mu\mu}$ -differential measurement) need then to be quadratically summed to obtain the final uncertainty.

Fig. 7.37 shows the distribution of corrected yields from the toys, in the two  $p_T^{\mu\mu\mu}$  bins in pp and PbPb. The final uncertainty (excluding the luminosity and  $\tau$  channel ones) is calculated for the  $p_T^{\mu\mu\mu}$ -differential measurements as the RMS of this distribution (including toys from the three spectrum-fit methods). The correlation between the yield variation and the associated  $\alpha \times \varepsilon$  value changes substantially the final uncertainty. For PbPb, a correlation increases strongly the total uncertainty in the first  $p_T^{\mu\mu\mu}$  bin (compared to the simple quadratic sum of the  $\alpha \times \varepsilon$  uncertainty from the equivalent of Fig. 7.36 and the other uncertainties), but an anti-correlation decreases it in the second bin down to a final uncertainty smaller than when ignoring the acceptance and efficiency. For the  $p_T^{\mu\mu\mu}$  dependence, the  $\alpha \times \varepsilon$  uncertainty is quoted as the quadratic subtraction of the other uncertainties from the final uncertainty, which is hence negative in the case of the second  $p_T^{\mu\mu\mu}$  bin in PbPb. In pp, the uncertainty in the two  $p_T^{\mu\mu\mu}$  bins is decreased with this procedure, but the  $\alpha \times \varepsilon$  contribution to the total uncertainty stays positive.

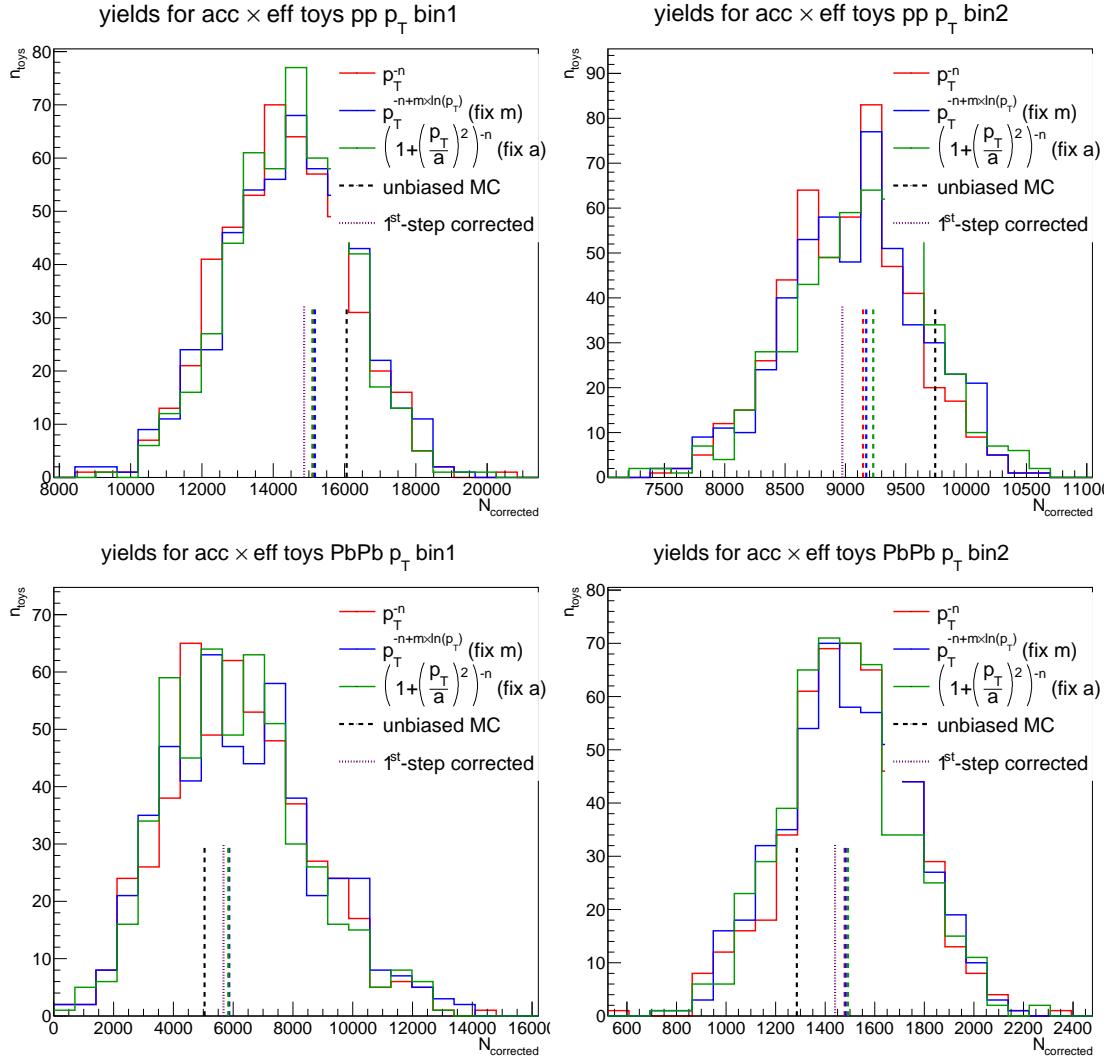


FIGURE 7.37: Variations of the second-step data measurement are associated to variations of the  $p_T^{\mu\mu\mu}$  spectrum correction in MC. Here is shown in the two  $p_T^{\mu\mu\mu}$  bins in pp (*top*) and PbPb (*bottom*) the effect of these variations on the corrected yields, including the uncertainty on the measurement of the observed yields. The coloured dashed lines are the nominal results from the three  $p_T^{\mu\mu\mu}$ -spectrum fit functions (and no variation of the data points), compared to the original result without any correction of the  $p_T^{\mu\mu\mu}$  spectrum (dashed black line) or with the first-step correction (dotted line).

Applying this procedure to the integrated results as well was considered (summing the variations of the two  $p_T^{\mu\mu\mu}$ -bin varied yields), but it would be somewhat hybrid to use the integrated fit for the central value, and the variations of the  $p_T^{\mu\mu\mu}$ -dependent measurements for the uncertainties. I therefore keep this acceptance and efficiency uncertainty decorrelated from the other uncertainties, as for the centrality bins (that are integrated on  $p_T^{\mu\mu\mu}$ ). The main impediment for applying the strategy considering these correlations, is that in the latter, the  $p_T^{\mu\mu\mu}$  spectrum is smeared with all the non- $\alpha\varepsilon$  uncertainties in a given  $p_T^{\mu\mu\mu}$  bin, and the combined effects of all uncertainties (including  $\alpha\varepsilon$ ) on the *same*  $p_T^{\mu\mu\mu}$  bin is considered. This is impossible to apply in the  $p_T$ -integrated sample (and in the centrality bins) for which the non- $\alpha\varepsilon$  uncertainties are *not* the ones that are used to vary the  $p_T^{\mu\mu\mu}$  spectrum and the acceptance and efficiency values.

One can visualise the correlation between the  $\alpha\varepsilon$  uncertainty and all the other ones (except that from luminosity and other  $B_c$  decays) by displaying jointly the acceptance and efficiency correction (first uncertainty) and the observed yield (second uncertainty) associated to each toy. Toys from the three methods are simply summed here. For the  $p_T^{\mu\mu\mu}$  bins, the final uncertainty is the RMS of the distribution of the corrected yields, *i.e.* the product of these two quantities. This means that if the two quantities are positively (negatively) correlated, the final uncertainty on the corrected yields will be larger (smaller) than the simple quadratic sum of the RMS values of the two quantities. This mechanism is illustrated in Fig. 7.38, where the two-dimensional plots that show a positive (negative) correlation factor are associated to a final uncertainty that is larger (smaller) than if the acceptance and efficiency was simply considered as another uncorrelated source of uncertainty.

The said correlation is positive (negative) for the low- $p_T^{\mu\mu\mu}$  (high- $p_T^{\mu\mu\mu}$ ) bins, both in pp and PbPb. This can be qualitatively understood: if the *observed yield in the low- $p_T^{\mu\mu\mu}$  bin rises* (meaning if the true value is higher than the nominal measured value in this bin), and considering to simplify that the high- $p_T^{\mu\mu\mu}$  bin yield is fixed to the measured value, then the fitted  $p_T^{\mu\mu\mu}$ -differential corrected yields shall have a *stronger slope*, favouring low- $p_T^{\mu\mu\mu}$  signal. This larger slope will be reflected in the signal MC, that will in turn be biased towards lower- $p_T^{\mu\mu\mu}$  candidates and give *lower acceptance and efficiency values* (because the acceptance and efficiency sharply drop towards low  $p_T^{\mu\mu\mu}$  values), and hence even *higher corrected yields*. Inversely, if the high- $p_T^{\mu\mu\mu}$  observed yield rises and the low- $p_T^{\mu\mu\mu}$  yield is fixed, then the slope of the fitted  $p_T^{\mu\mu\mu}$  spectrum will be smaller, resulting in higher acceptance and efficiency values from the associated biased MC.

The same logic can be applied to the corrected yields in pp and PbPb, whose ratio is proportional to the  $R_{\text{PbPb}}$  quoted in the results. As there are toys distributed with the final uncertainties (on the  $p_T$ -dependent corrected yields), I can check if the pp and PbPb ones are correlated, and therefore check if it is correct to simply add the associated uncertainties in quadrature. Correlations between pp and PbPb could arise *e.g.* from similar shapes of the acceptance and efficiency versus  $p_T^{\mu\mu\mu}$ . Fig. 7.39 shows the two-dimensional distributions of the corrected yields in PbPb and pp in both  $p_T^{\mu\mu\mu}$  bins, and the distributions of the acceptance and efficiency corrections for the PbPb centrality bins versus the pp integrated sample. The correlation factors are quite small in both cases, revealing that it would be acceptable to take the final uncertainties on the  $R_{\text{PbPb}}$  as the quadratic sum of the uncertainties on the pp and PbPb measurements. I however take the final uncertainty on  $R_{\text{PbPb}}$  from the RMS of the distribution of the ratio of varied (toy) PbPb to pp corrected yields for the  $p_T^{\mu\mu\mu}$ -dependent results, and the acceptance and efficiency uncertainty on  $R_{\text{PbPb}}$  from the RMS of the ratio of the varied pp to PbPb acceptance and efficiency values for the  $p_T^{\mu\mu\mu}$ -integrated results.

The nominal  $\frac{1}{\alpha \times \varepsilon}$  correction (or corrected yield, in the case of the  $p_T^{\mu\mu\mu}$  dependence) is defined as the average of the nominal values from the three  $p_T^{\mu\mu\mu}$  spectrum fit methods. It is shown in Table 7.2 along with the values from the simple one-binned method, and the values after the first-step correction. The associated systematic uncertainty is the quadratic sum of:

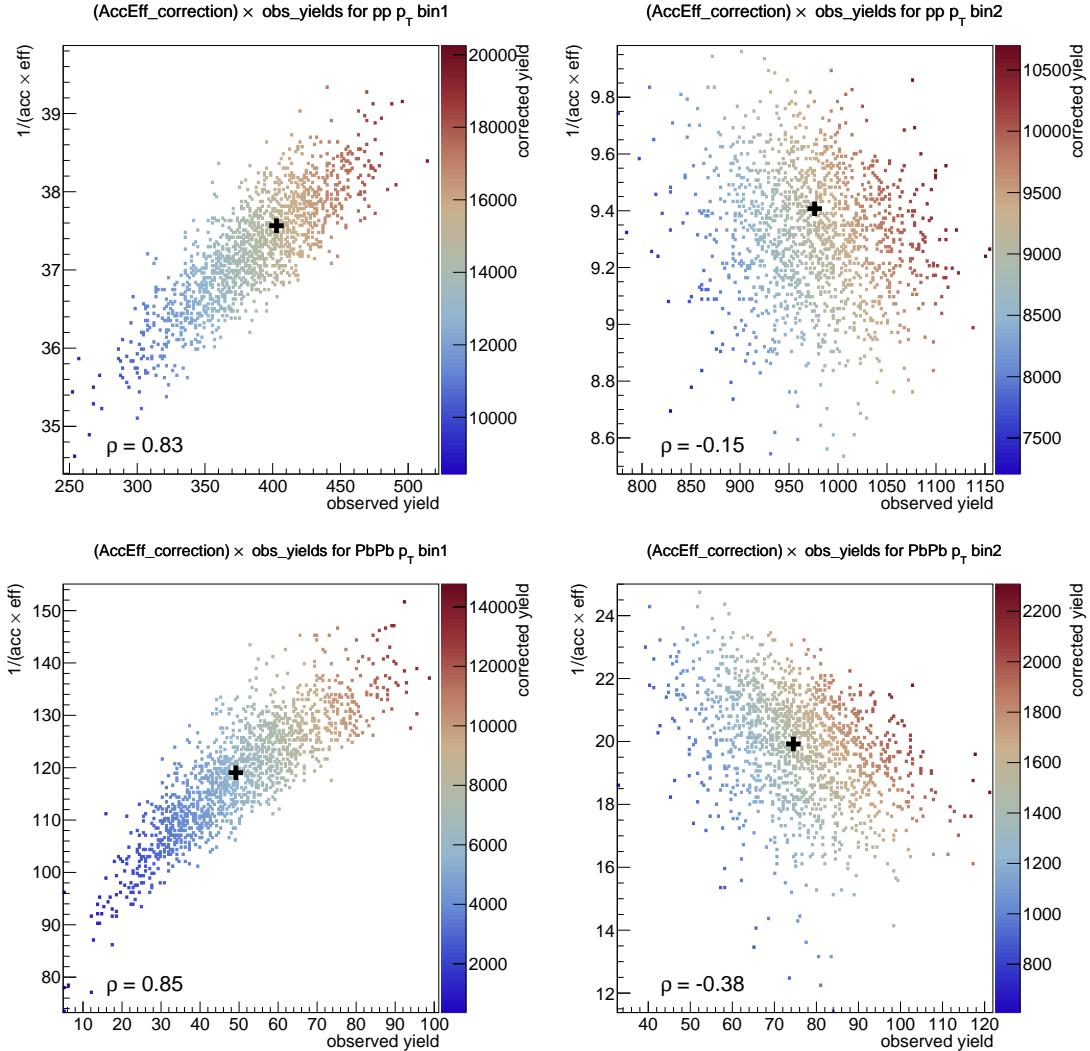


FIGURE 7.38: Acceptance and efficiency correction associated to each toy variation of the observed yield, the latter arising from all relevant uncertainty sources other than the acceptance and efficiency, in pp (*top*) and PbPb (*bottom*), for the first (*left*) and second (*right*)  $p_{\text{T}}^{\mu\mu\mu}$  bins. The z-axis, represented in colour nuances, shows the product of the two plotted quantities, namely the variation of the corrected yield, whose RMS will constitute the final uncertainty on the  $p_{\text{T}}^{\mu\mu\mu}$ -differential cross sections. The black cross shows the nominal value.

The correlation factors between the two axes are also displayed.

- The RMS of all toys from the three fit methods. The three RMS values for the three methods are very close, so it is reasonable to consider the common RMS of all methods. As the spread of values around the nominal is not symmetric, I use an asymmetric RMS determination (meaning for *e.g.* the upper error, I use the root of the mean of squares of only upper-going deviations to the nominal). For the modification factor  $R_{\text{PbPb}}$ , the RMS is that of the distribution of the ratios of the varied acceptance and efficiency corrections (or the varied corrected yields for the  $p_{\text{T}}^{\mu\mu\mu}$  dependence) in PbPb over the pp ones, taking into account possible correlations between pp and PbPb uncertainties.
- The maximal deviation of the nominal values from the three methods to the average of the three values. This uncertainty is much smaller than the RMS, which indicates that the procedure is reliable.

Along with the RMS, the correlation between the two  $p_{\text{T}}^{\mu\mu\mu}$  bins for all  $p_{\text{T}}^{\mu\mu\mu}$ -spectrum variations can be calculated with the standard formula (based on a symmetric RMS), with  $n$  variations:

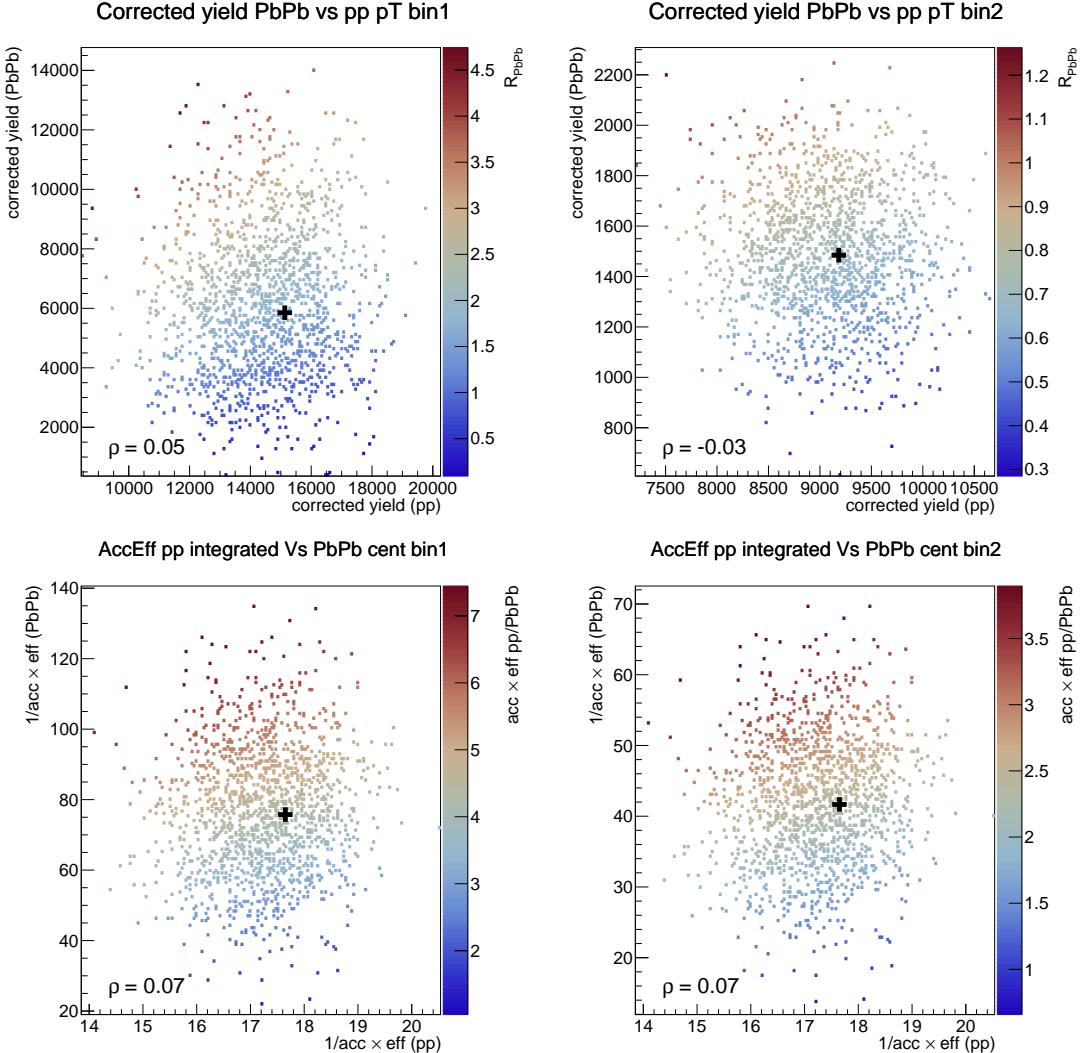


FIGURE 7.39: *Top:* Corrected yields in PbPb versus the ones in pp associated to toy variations of the observed yields, for the low- $p_T^{\mu\mu\mu}$  (*left*) and high- $p_T^{\mu\mu\mu}$  (*right*) regions. The z-axis colours show the ratio of the two quantities, whose RMS will be taken as the final uncertainty on the  $p_T^{\mu\mu\mu}$ -differential  $R_{\text{PbPb}}$ . *Bottom:* Acceptance and efficiency corrections in PbPb versus the ones in pp associated to toy variations of the observed yields. The first (*left*) and second (*right*) PbPb centrality bins are shown versus the  $p_T^{\mu\mu\mu}$ -integrated pp sample; the two bins give very similar plots because only the efficiency changes with centrality (and not the acceptance, which brings the largest  $p_T^{\mu\mu\mu}$  dependence). The z-axis colours shows the distribution of the ratio of the two corrections, whose RMS will be used as the uncertainty from the acceptance and efficiency corrections for the centrality-dependent  $R_{\text{PbPb}}$ . The nominal values (black cross) and the correlation factors between the two axes are also displayed.

$$\rho_{1,2} = \frac{1}{n-1} \frac{\sum_i (y_{1,i} - \langle y_1 \rangle)(y_{1,2} - \langle y_2 \rangle)}{\sigma_1 \sigma_2} \quad (7.28)$$

As a given  $p_T^{\mu\mu\mu}$  spectrum correction has a very similar effect on the spectrum within each  $p_T^{\mu\mu\mu}$  bins, this correlation is very close to 1 for all three fit methods ( $\rho_{1,2} > 0.99$ ). We add to this correlation factor (via the sum of the covariances) an assumed correlation factor of 0.3 for the maximal deviation of the three nominal values to the final nominal correction (which is sub-dominant compared to the RMS). The resulting correlation factor is still  $> 0.95$ . The bin-to-bin correlation factors for the corrected yield uncertainty on the  $p_T^{\mu\mu\mu}$ -dependent  $R_{\text{PbPb}}$  and for the

TABLE 7.2: Acceptance and efficiency values calculated on the signal MC, integrated over the kinematics of the first or second  $p_T^{\mu\mu\mu}$  bin. The values are from the (final) third step, except in the two columns showing the results without  $p_T^{\mu\mu\mu}$ -spectrum correction or with a first-step correction.

system	bin	acceptance $\alpha$ (final)	efficiency $\varepsilon$ (final)	$\alpha \times \varepsilon$ (no correction)	$\alpha \times \varepsilon$ (2 <sup>nd</sup> step)	$\alpha \times \varepsilon$ (final)
PP	$6 < p_T^{\mu\mu\mu} < 11 \text{ GeV}$	0.098	0.27	0.025	0.0271	<b>0.0266</b>
	$11 < p_T^{\mu\mu\mu} < 35 \text{ GeV}$	0.249	0.43	0.100	0.109	<b>0.106</b>
	integrated	0.155	0.36	0.049	0.0597	<b>0.0567</b>
PbPb	$6 < p_T^{\mu\mu\mu} < 11 \text{ GeV}$	0.086	0.097	0.0098	0.0087	<b>0.0084</b>
	$11 < p_T^{\mu\mu\mu} < 35 \text{ GeV}$	0.22	0.23	0.058	0.052	<b>0.050</b>
	centrality 0-20%	0.113	0.117	0.0198	0.0144	<b>0.0132</b>
	centrality 20-90%	0.11	0.21	0.035	0.026	<b>0.024</b>
	integrated	0.113	0.148	0.0248	0.0182	<b>0.0168</b>

acceptance and efficiency uncertainty on the centrality-dependent  $R_{\text{PbPb}}$  are also calculated this way. This is possible thanks to the complete distribution of the ratios of corrected yields (or  $1/\alpha\varepsilon$  corrections) of PbPb over pp. Only the calculation of bin-to-bin correlation of the fit uncertainty of the centrality-dependent  $R_{\text{PbPb}}$  needs a small MC procedure that is explained in section 8.2, where the three correlations are shown (Table 8.1).

This almost-full bin-to-bin correlation of the  $\alpha \times \varepsilon$  uncertainty is the reason why it is much larger in the centrality bins and in the  $p_T^{\mu\mu\mu}$ -integrated results: the  $p_T^{\mu\mu\mu}$ -spectrum variations have cumulating effects on the  $\alpha \times \varepsilon$  values along  $p_T^{\mu\mu\mu}$  bins. The correlation between the two centrality bins is high because the uncertainties of the two bins come from the same  $p_T^{\mu\mu\mu}$ -spectrum variations and similar  $p_T^{\mu\mu\mu}$  distributions. It is assumed that the large uncertainties (of  $p_T$ -integrated bins) cover for the possibility of the  $p_T^{\mu\mu\mu}$  spectra in the two centrality bins being different.

The fact that the extracted uncertainty on the acceptance and efficiency is very correlated between the  $p_T^{\mu\mu\mu}$  bins renders the iterative procedure very stable, as will also demonstrate section 7.6.7: the change seen between the first-step and second-step  $\frac{1}{\alpha \times \varepsilon}$  can indeed only come from the changes in the BDT and fit of the second-step analysis, because the change of the acceptance and efficiency from the previous step (used to calculate the corrected yields that are fitted) affects the yields mostly as a global factor (that does not impact the  $\frac{1}{\alpha \times \varepsilon}$  determination from MC). In other words, running again this procedure with unchanged BDT and fit will result in very similar acceptance and efficiency values.

### 7.6.6 Additional remarks

A few other items of the analysis, linked to the BDT, are improved by the input of a first-step measurement:

- The second-step BDT training and fit benefit from a better description of the  $p_T^{\mu\mu\mu}$  spectrum of signal MC.
- The background normalisations during the training of the BDT (section 5.7.3) are coarsely inspired from the (first-step) post-fit values of the corresponding parameters.
- The average BDT value and RMS per mass bin, extracted in Fig. 5.18 of section 5.7.5, uses background samples with post-fit parameters to build the full background sample. The

second-step post-fit background is more accurate, giving a corrected BDT that is more exactly decorrelated from the mass.

- In the second-step, one can check in section 5.7.5 that the BDT distribution of data is consistent with the BDT distribution from the sum of the post-fit signal MC and backgrounds, and apply an associated correction in pp before a last re-fit.

### 7.6.7 MC closure test of the iterative procedure

In total, three estimations of the nominal acceptance and efficiency correction are performed: with the original MC, and after the first-step or second-step  $p_T^{\mu\mu\mu}$  spectrum correction. The difference between the last two estimates mostly comes from the BDT and the template fit being modified in the second-step analysis.

Figs. 7.36 and 7.37 already show that this iterative procedure converges, as the difference between the second-step and third-step nominal values (of the correction or of the corrected yield) is smaller than the difference between the second-step value and the value from the original MC. However, it is worth checking this convergence, as well as the fact that it converges to the right value, through a closure check on the signal MC. To do this, I create two toy MC datasets in PbPb with biased  $p_T^{\mu\mu\mu}$  spectrum, run the whole iterative procedure from the original MC as in our main analysis, and intend to find back the  $p_T^{\mu\mu\mu}$  spectrum that was injected. More precisely:

1. I strongly bias the  $p_T^{\mu\mu\mu}$  spectrum of the MC to create a new ‘true’ value of the corrected yield, taking the role of the true data yield that this analysis aims at measuring. The two toys are biasing the MC with event weights  $\left(\frac{p_T^{\mu\mu\mu} [\text{GeV}]}{11}\right)^{\pm 1.7}$ , which is a stronger variation than from the original MC to the final measured PbPb spectrum.
2. From this biased MC, I estimate the true acceptance and efficiency correction with the one-binned method. Multiplying the known true yield by this, I get an equivalent observed yield, that takes the role of the signal yield extracted from the template fit in the main analysis. I then forget temporarily the true value of the acceptance and efficiency, and rather try to evaluate it as is done on actual data.
3. I hence first evaluate the acceptance and efficiency on the original MC, as in the first-step analysis. This divides the toy observed yields mentioned above to give the initial corrected yields. I then fit the  $p_T^{\mu\mu\mu}$ -differential corrected yields to obtain the first-step correction to the  $p_T^{\mu\mu\mu}$  spectrum, to be applied to the original MC.
4. At this stage, in the nominal analysis, I would re-run most of the analysis steps including the template fit, giving slightly modified observed yields. In this closure test, I ignore this, *i.e.* the effect of the  $p_T^{\mu\mu\mu}$  correction of MC on the extracted observed signal yields. This crosscheck indeed aims at confirming the convergence of the acceptance and efficiency determination procedure.
5. From the first-step-corrected MC, I estimate again the acceptance and efficiency and the resulting corrected yields. Those yields are fitted to obtain the second-step  $p_T^{\mu\mu\mu}$  correction of the MC.
6. As in the main analysis, the final acceptance and efficiency is evaluated on the second-step-corrected MC.

I hence obtained, for each of the two toy MC datasets with arbitrarily biased corrected yields, three values for the corrected yields, from the original MC and the first-step or second-step corrected MC. These values can be compared with the true corrected yield, corresponding to the true acceptance and efficiency calculated with the known bias used to create the toy datasets. The convergence of the iterative procedure is also tested. Fig. 7.40 shows the iterated values of

the corrected yields, and that they indeed converge to the true value, already after the first-step correction. This validates the global procedure. In the actual measurement, the last two values of the corrected yields are more different than in this closure test, because the observed yields were modified in the second run of the analysis (including the template fit), which this closure test does not intend to reproduce.

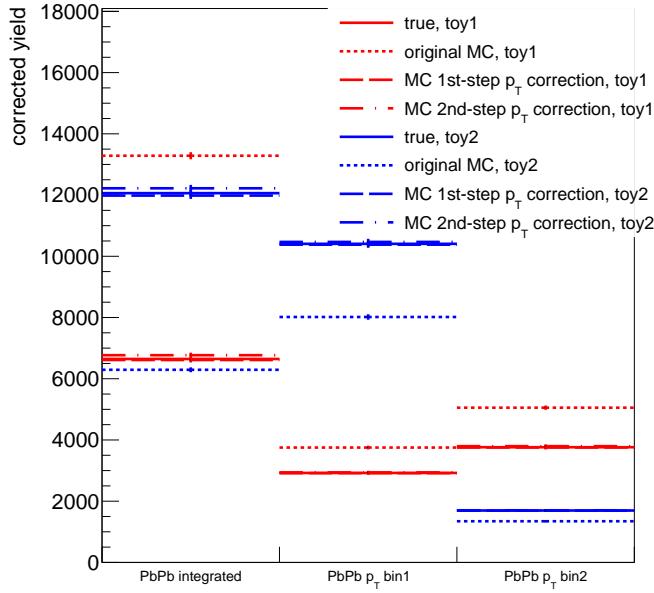


FIGURE 7.40: MC closure test of the iterative procedure for determining the nominal acceptance and efficiency corrections, in PbPb. Two toy datasets (red and blue) undergo the same procedure as in the main analysis, each giving rise to three sets of corrected yields for the three steps of the analysis. These corrected yields are compared to the true toy corrected yields, for both  $p_T^{\mu\mu\mu}$  bins and for the integrated sample.

“Entre la foi et l’incrédulité, un souffle,  
 entre la certitude et le doute, un souffle.  
 Sois joyeux dans ce souffle présent où tu vis,  
 car la vie elle-même est dans le souffle qui passe.”

---

Omar KHAYYAM (died 1123),  
*Les 144 Quatrains d’Omar Khayyam*,  
 traduits par Claude ANET et Mirza MUHAMMAD,  
<sup>1</sup>  
 Quatrain 130, 1920.

## Chapter 8

# Are we sure? Summary of uncertainties

### 8.1 Uncertainty sources

I summarise here the various uncertainties that are considered in the final uncertainty on the cross sections and  $R_{\text{PbPb}}$ :

- The uncertainty from the fit procedure is given by MINOS when running `combine`. MINOS gives asymmetric uncertainties, which are kept as such; they are shown in Table 6.2. MINOS also outputs a correlation factor between the two  $p_T^{\mu\mu\mu}$  or centrality bins. These fit uncertainties take into account uncertainties such as the statistical fluctuations of data, or the morphing variations of the background shapes within reasonable hypotheses.
- The uncertainties from varying the fit method are treated in section 6.4 and values are summarised in Table 6.4. These uncertainties are calculated from an RMS of some variations, so it is symmetric by construction, and also provides the correlation between the two analysis bins through an RMS-like calculation.
- The uncertainty on the acceptance and efficiency corrections is determined in sections 7.6.4 and 7.6.5. It results from the RMS of the re-computations of the  $\frac{1}{\alpha \times \varepsilon}$  correction after modifying the  $p_T^{\mu\mu\mu}$  spectrum of the signal MC within what it allowed by the second-step measurement uncertainties. A small contribution from the potential mis-modelling of the  $p_T^{\mu\mu\mu}$  spectrum is also included. For the  $p_T^{\mu\mu\mu}$ -dependent result, this uncertainty is actually

---

<sup>1</sup>I prefer this French traduction (from Farsi or Arabic, as Khayyam wrote this piece in both languages) over the English one found in quatrain 20 of *The quatrains of Omar Khayyam translated by E.H. WHINFIELD*, 1903: “The distance which separates incredulity from faith is but a breath, – that which separates doubt from certainty is equally but a breath. – Let us, then, pass this precious space of a breath gaily, – for our life also is only separated [from death] by the space of a breath.”. I thank Batoul DIAB for checking that the meaning is similar in the original quatrain in Arabic.

convolved with (most) other uncertainties, such that the final uncertainty (excluding luminosity and other  $B_c$  decay channels) is the RMS of the varied corrected yields, that are based on MC toy samples whose  $p_T^{\mu\mu\mu}$  spectra are each corrected with a different variation of the second-step  $p_T^{\mu\mu\mu}$ -differential measurement. The bin-to-bin correlation is calculated from the sample of varied acceptance and efficiency values or of varied yields.

- The uncertainty on the correction of single-muon efficiencies through tag-and-probe scale factors is explained in section 7.2.2.1.
- Concerning the possible contribution of other  $B_c$  decays as  $B_c \rightarrow J/\psi \tau \bar{\nu}_\tau$  and  $B_c \rightarrow \chi_{c,0,1,2} \mu X$  to the measured signal yield, section 4.3.3 justified to assign an asymmetric (yield-decreasing) 4.5% uncertainty for the cross sections, and a symmetric 1.5% for  $R_{\text{PbPb}}$ . This uncertainty is assumed 100% correlated between the two  $p_T^{\mu\mu\mu}$  or centrality bins.
- The uncertainties on the normalisation of the extracted yield are 1.9% in pp (on luminosity), 2.6% in PbPb (on  $N_{\text{MB}} \times T_{\text{PbPb}}$  for the 0 – 90% centrality range), and respectively 2.3% and 3.8% for the 0 – 20% and 20 – 90% PbPb centrality ranges, as explained in section 3.5.5.2. This (multiplicative) luminosity uncertainty is summed in quadrature to the other uncertainties after all corrections. It is fully correlated along analysis bins.
- There is an uncertainty in the calibration of the centrality ranges, that amounts to varying the limits of the sum of transverse energy in the forward calorimeter corresponding to a given centrality range. The varied centrality limits were calculated, but have no effect on the result in the 0 – 90% centrality, as there is no preselected event in data that is close to 90% centrality. Concerning the centrality binning, there are zero events from the third BDT bin that migrates centrality bin (*i.e.* crosses the 20% boundary) with the varied centrality calibration, and only one event in the second BDT bin (containing a minority of signal events). The centrality has hence a negligible systematic effect on the results.
- It is chosen in this analysis to present all results as a function of the partially reconstructed kinematic variables. A possible (imperfect) correction to the full  $B_c$  kinematics, that is not considered in the nominal results, is presented in section 8.3.

In the  $p_T^{\mu\mu\mu}$ -dependence case, the main part of the uncertainty is obtained via the spread of the corrected yields resulting from varying the  $p_T^{\mu\mu\mu}$  spectrum within the uncertainties from the fit, the fit method, and the tag-and-probe. I then simply add in quadrature the uncertainties from luminosity/Glauber model and from other  $B_c$  decays, and combine the corresponding bin-to-bin correlation factors as explained in section 8.2.2. In other cases (integrated results and centrality dependence), the first six uncertainty sources considered above are quadratically summed into the final uncertainties, and their correlation factors are combined into the final correlation.

Figs. 8.1 and 8.2 present the contributions of these sources to the relative asymmetric uncertainties for the  $p_T^{\mu\mu\mu}$  and centrality dependences, in pp, PbPb, and in the  $R_{\text{PbPb}}$ . The dominant contribution is from the fit in the  $p_T^{\mu\mu\mu}$  dependence, and from the acceptance and efficiency when integrating on  $p_T^{\mu\mu\mu}$  (including the centrality bins). For the  $p_T^{\mu\mu\mu}$ -integrated bins, the fit is the second-largest contribution. For the  $p_T^{\mu\mu\mu}$  dependence, depending on the bins, the uncertainty from the acceptance and efficiency corrections or from the fit method variations is the second-largest one.

## 8.2 Correlations

For the cross sections, the bin-to-bin correlations for the fit, fit method, and acceptance and efficiency uncertainties are directly provided by the procedures that compute these uncertainties. For the  $R_{\text{PbPb}}$ , they are also given by the procedures, for the fit method, for the acceptance and efficiency, and, in the  $p_T^{\mu\mu\mu}$  dependence, for the uncertainty on the corrected yields (gathering

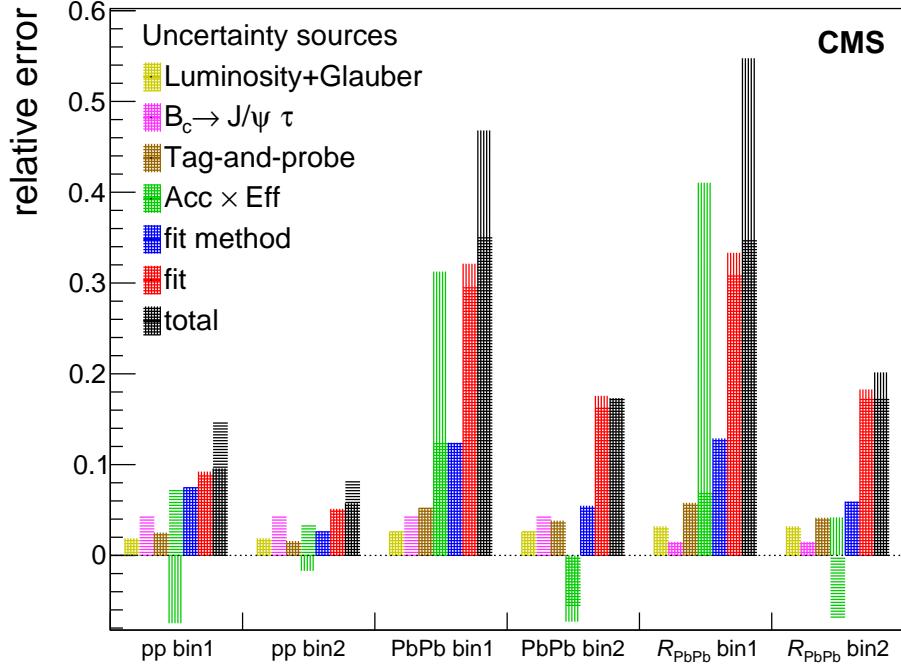


FIGURE 8.1: Contribution of the various sources to the relative uncertainties on the cross sections and on the  $R_{\text{PbPb}}$ , compared to the total uncertainty (in black), for the  $p_{\text{T}}^{\mu\mu\mu}$  dependence. Bin1 is low- $p_{\text{T}}^{\mu\mu\mu}$  and bin2 is high- $p_{\text{T}}^{\mu\mu\mu}$ . The vertical-hash bars show the high-side uncertainties, and the horizontal-hash bars the low-side ones. The negative values for the acceptance and efficiency in some bins represent the effect of anti-correlations compensating the effect of the other uncertainties (cf section 7.6.5).

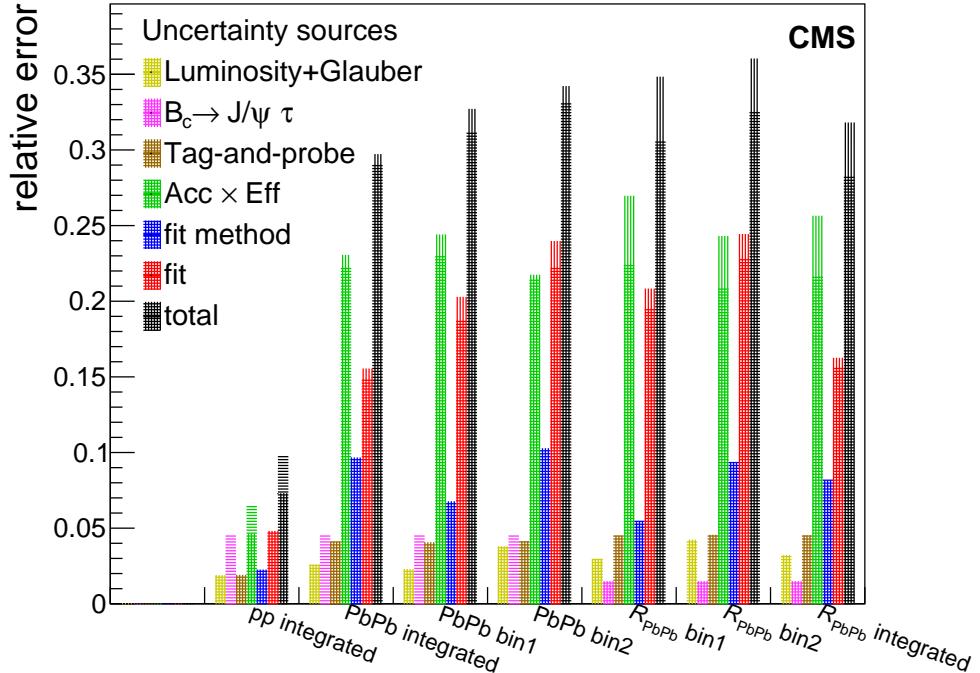


FIGURE 8.2: Contribution of the various sources to the relative uncertainties on the cross sections and on the  $R_{\text{PbPb}}$ , compared to the total uncertainty (in black), for the centrality dependence and the integrated samples. Bin1 contains more central events and bin2 more peripheral events. The vertical-hash bars show the high-side uncertainties, and the horizontal-hash bars the low-side ones.

the three mentioned sources). In these cases, the correlation is indeed computed with a RMS-like formula on the distribution of ratios of the varied PbPb and pp values. In the centrality dependence however, the bin-to-bin correlation misses for the fit uncertainty. This is calculated through a simple MC in section 8.2.1, that is also applied to display the correlation for the  $R_{\text{PbPb}}$  fit uncertainty in the  $p_T^{\mu\mu\mu}$  dependence (unused in the final result).

The uncertainty from other  $B_c$  decays has been assumed fully correlated along bins, including for the  $R_{\text{PbPb}}$  where the uncertainty partially cancels in the PbPb/pp ratio. Similarly, for the tag-and-probe uncertainty, the bin-to-bin correlation factor is assumed to be 0.8, including for the  $R_{\text{PbPb}}$ .

The bin-to-bin correlation factors from the different sources need to be combined into the correlation quoted for the final uncertainty. The procedure to do so is explained in section 8.2.2. The resulting correlation factors for all uncertainty sources and the final uncertainties are displayed in Table 8.1.

TABLE 8.1: Correlation factor between the two  $p_T^{\mu\mu\mu}$  or centrality bins for pp, PbPb, and  $R_{\text{PbPb}}$ , for all uncertainty sources and for the total uncertainties. For the  $p_T^{\mu\mu\mu}$  dependence the final correlation combines the one from the ‘full’ uncertainty with the ones from luminosity/Glauber model and other  $B_c$  decays. For the centrality dependence, the correlations of the sources in the first five lines are combined.

uncertainty source	pp ( $p_T^{\mu\mu\mu}$ )	PbPb ( $p_T^{\mu\mu\mu}$ )	$R_{\text{PbPb}}$ ( $p_T^{\mu\mu\mu}$ )	PbPb (centrality)	$R_{\text{PbPb}}$ (centrality)
fit	0.10	0.05	0.06	0.05	0.10
fit method	0.97	0.30	0.57	0.37	0.42
acc. and eff. correction	0.98	0.99	0.99	1.00	1.00
$B_c \rightarrow J/\psi \tau \bar{\nu}_\tau$ decay + lumi.	1.0	1.0	1.0	1.0	1.0
tag-and-probe scale factors	0.8	0.8	0.8	0.8	0.8
full (w/o lumi.+ $B_c \rightarrow J/\psi \tau$ )	0.67	0.44	0.42	—	—
total	<b>0.69</b>	<b>0.45</b>	<b>0.43</b>	<b>0.54</b>	<b>0.57</b>

### 8.2.1 Correlation factor of a ratio of random variables

If the correlation factors between the two analysis bins for the pp and the PbPb yields  $\rho_{1,2}(y_{\text{pp}})$  and  $\rho_{1,2}(y_{\text{PbPb}})$  are known, what is the correlation factor for the ratio of the two yields  $\rho_{1,2}(y_{\text{PbPb}}/y_{\text{pp}})$ ? There is no exact analytical answer to this question, so I implement a simple MonteCarlo to measure it. This is needed for the correlation of the fit uncertainty between the two centrality bins of  $R_{\text{PbPb}}$  – and only for display purposes for the  $p_T^{\mu\mu\mu}$  dependence.

For  $n = 2 \times 10^5$  iterations, I draw from a bivariate Gaussian distribution, separately for the pp yields in the two studied bins, and for the PbPb yields. Each 2D Gaussian is biased to contain the known correlations of the pp and PbPb yields. The  $R_{\text{PbPb}}$  (ratio of the two yields, normalised as in Eq. 3.9) resulting from each set of toy PbPb and pp yields is then plotted on a 2D histogram. Finally, the correlation factor of this 2D histogram is measured; it is that of the  $R_{\text{PbPb}}$ , corresponding to the input pp and PbPb correlations.

In practice, four random numbers  $u_1$ ,  $u_2$ ,  $v_1$ , and  $v_2$ , are drawn for each iteration, from a unit gaussian distribution of mean 0. Then the pp yield distribution is reproduced from the

nominal yields  $y_{b=1,2}^{\text{pp}}$ , their uncertainties  $\sigma_{b=1,2}^{\text{pp}}$ , and their correlation factor  $\rho_{1,2}^{\text{pp}}$ , and similarly for PbPb:

$$\begin{aligned} Y_1^{\text{pp}} &= y_1^{\text{pp}} + u_1 \times \sigma_1^{\text{pp}} \\ Y_2^{\text{pp}} &= y_2^{\text{pp}} + (u_1 \rho_{1,2}^{\text{pp}} + u_2) \sqrt{1 - \rho_{1,2}^{\text{pp}}} \times \sigma_2^{\text{pp}} \end{aligned}$$

and similarly in PbPb, but using  $(v_1, v_2)$  instead of  $(u_1, u_2)$ . Three two 2D histograms are filled, respectively with the pp or PbPb yields, or with the resulting  $R_{\text{PbPb}}$  values. The wanted correlation factor is extracted from the last histogram.

This method is run for the correlation of the fit uncertainty in the  $R_{\text{PbPb}}$ . The resulting correlation is only used for the centrality dependence, which is actually a special case, as the pp yields entering the denominator are exactly the same in the two bins (*i.e.* the integrated yield). This means the correlation factor is exactly 1 for the pp yields, resulting in an increase of the correlation from the PbPb yields to the  $R_{\text{PbPb}}$ . The mentioned histograms of the procedure in this case are shown in Fig. 8.3, and the same plots for the  $p_T$  dependence are omitted.

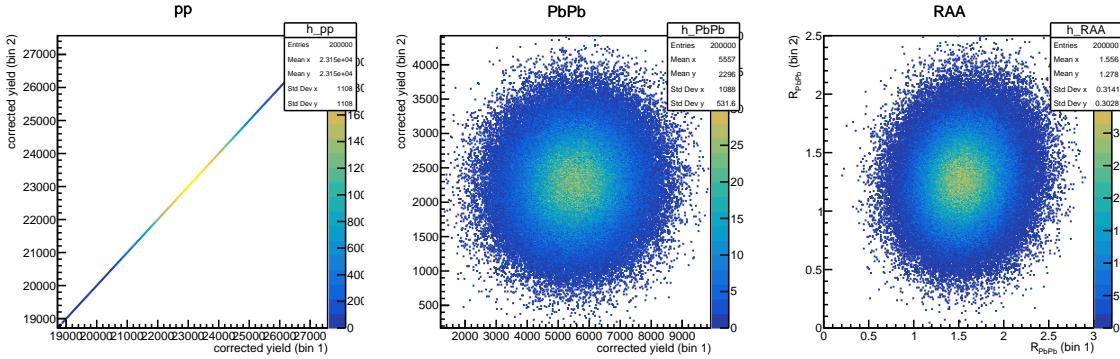


FIGURE 8.3: Determination of the correlation of the fit uncertainty between the two centrality bins for  $R_{\text{PbPb}}$ . *Left:* identical corrected yields for the two bins in pp, drawn from a 1D Gaussian distribution with the uncertainty of the pp integrated yield. *Center:* yields in the PbPb centrality bins, drawn from a bivariate Gaussian distribution using the relevant uncertainties. *Right:* Values of the  $R_{\text{PbPb}}$  from the ratio of pp and PbPb yields drawn from their respective distributions.

### 8.2.2 Correlation factor for the sum of independent uncertainty sources

Let us start from the bilinearity of the covariance of two random variables. Let  $X, X', Y$  be three random variables, *i.e.* uncertainty sources; then:

$$\text{Cov}(X + X', Y) = \text{Cov}(Y, X + X') = \text{Cov}(X, Y) + \text{Cov}(X', Y) \quad (8.1)$$

In passing, this gives the classic result that the variances (defined by  $\text{Var}(A) = \text{Cov}(A, A)$ ) of independent sources  $A$  and  $B$  (so that  $\text{Cov}(A, B) = 0$ ) can be summed:

$$\text{Var}(A + B) = \text{Cov}(A, A) + \text{Cov}(B, B) + 2\text{Cov}(A, B) = \text{Var}(A) + \text{Var}(B) \quad (8.2)$$

Defining now the variables  $A_1, B_1, A_2$ , and  $B_2$  corresponding to independent sources  $A$  and  $B$  that influence respectively the first and second analysis bins, I get :

$$\text{Cov}(A_1 + B_1, A_2 + B_2) = \text{Cov}(A_1, A_2) + \text{Cov}(B_1, B_2) + \text{Cov}(A_1, B_2) + \text{Cov}(B_1, A_2) \quad (8.3)$$

The last two terms are null because of the independence of  $A$  and  $B$ , so I get the correlation factor between the two analysis bins for a total uncertainty resulting from independent sources  $A^i$ :

$$\rho_{1,2} = \frac{Cov(\sum_i A_1^i, \sum_i A_2^i)}{\sqrt{Var(\sum_i A_1^i)Var(\sum_i A_2^i)}} = \frac{\sum_i Cov(A_1^i, A_2^i)}{\sigma_{tot,1}\sigma_{tot,2}} \quad (8.4)$$

where  $\sigma_{tot,k} = \sigma(\sum_i A_k^i)$  sums in quadrature the uncertainties of all (assumed independent) sources, for each bin  $k$ . These standard deviations are symmetrised if they were extracted as asymmetric (through the average of upper and lower uncertainties).

This allows to simply sum the covariances (between the two analysis bins) of each source and then divide by the product of the total standard deviations in each bin, to obtain the final bin-to-bin correlation for all sources. As shown in Table 8.1, the correlations from all six sources are combined for the centrality dependence, and from only three sources in the  $p_T^{\mu\mu\mu}$  dependence, in which some uncertainties were already combined in that on the corrected yields.

### 8.2.3 Separating correlated and uncorrelated uncertainties

As chapter 9 will explain, the final uncertainties of the results are presented differently than the usual separation of statistical and systematic uncertainties. This is motivated by two facts. First, the `combine` fit outputs uncertainties that already account for some systematic sources, and separating them would be somewhat artificial. Second, the majority of the systematic uncertainties actually depend significantly on the available statistics – meaning they would decrease with larger datasets. The characteristic that I want to keep from the ‘statistical versus systematic’ paradigm is that systematic uncertainties are often considered fully correlated along bins. The information most useful for the results to be used (typically for comparisons with models) is the correlation factor between the two bins, needed for any rigorous fit. For display purposes, let us find a way to separate the uncertainty into a part that is fully correlated between the two analysis bins, and another part that is uncorrelated between the two bins. The latter is mostly associated to statistical sources, but also partly to uncorrelated systematic sources.

Therefore, I intend to extract the correlated ( $\sigma_{1,c}$  and  $\sigma_{2,c}$ ) and uncorrelated ( $\sigma_{1,unc}$  and  $\sigma_{2,unc}$ ) uncertainties, from the total uncertainties on the two bins  $\sigma_{1,tot}$  and  $\sigma_{2,tot}$ , and the total correlation factor between the two bins  $\rho_{1,2}$ . The quadratic sum of these two parts should equal the total uncertainty:

$$\sigma_{1,unc}^2 + \sigma_{2,unc}^2 = \sigma_{i,tot}^2$$

I characterise the correlated part of an uncertainty by requiring that it is the only part that cancels in the uncertainty on the subtraction (for a positive  $\rho_{1,2}$ ) or the addition (for a negative  $\rho_{1,2}$ ) of the two bin contents. Defining the uncertainty  $\sigma_{1\pm 2,tot}$  on the addition or subtraction of the bin contents, I get:

$$\begin{aligned} \sigma_{1\pm 2,tot}^2 &= \sigma_{1,tot}^2 + \sigma_{2,tot}^2 \pm 2\rho_{1,2}\sigma_{1,tot}\sigma_{2,tot} \\ &= \sigma_{1,unc}^2 + \sigma_{2,unc}^2 + \sigma_{1,c}^2 + \sigma_{2,c}^2 \pm 2\rho_{1,2}\sigma_{1,tot}\sigma_{2,tot} \\ &\stackrel{\text{require}}{=} \sigma_{1,unc}^2 + \sigma_{2,unc}^2 \end{aligned} \quad (8.5)$$

To satisfy the last equation, and considering that one would like either the bin addition or subtraction to cancel the correlated uncertainties, one needs:

$$\sigma_{1,c}^2 + \sigma_{2,c}^2 = 2|\rho_{1,2}|\sigma_{1,tot}\sigma_{2,tot} \quad (8.6)$$

I also assume that the correlated fraction of the uncertainty is the same for the first and second bin, so that  $\frac{\sigma_{1,c}}{\sigma_{1,tot}} = \frac{\sigma_{2,c}}{\sigma_{2,tot}}$ . Inserting this in Eq. 8.6, I get the correlated part of the uncertainty as:

$$\sigma_{2,c}^2 = \sigma_{1,c}^2 \frac{\sigma_{2,tot}^2}{\sigma_{1,tot}^2} = \frac{2|\rho_{1,2}|}{1 + \frac{\sigma_{1,tot}^2}{\sigma_{2,tot}^2}} \sigma_{1,tot} \sigma_{2,tot} \quad (8.7)$$

The correlation factor is calculated after symmetrising the uncertainties, but this formula is applied separately to each side of the asymmetric uncertainties: the correlated part of the low-side uncertainty uses the symmetrised correlation factor and the low-side total uncertainties in both bins.

### 8.3 Potential correction of the visible $p_T$

All the results presented in chapter 9 and before show dependences along the visible kinematic variables of the  $B_c$ , meaning those of the trimuon from the  $B_c$  decay. The energy taken by the neutrino makes these variables differ from those of the true  $B_c$ . This effect adds to the reconstruction by the detector, which smears the reconstructed kinematics; this is however negligible compared to the effect of partial reconstruction.

One could imagine correcting the  $p_T^{\mu\mu\mu}$  distribution so that it corresponds in average to the  $p_T$  of the generated  $B_c$ , relying on the signal MC. In LHCb, it is usual to correct by the average of the ratio of visible and generated kinematics, as a function of the visible mass: the neutrino indeed takes away more energy at lower visible masses, yielding a smaller correction.

Fig. 8.4 shows the reconstructed trimuon  $p_T$  and the full  $p_T$  of the generated  $B_c$ . The signal MC is weighted using the nominal second-step  $p_T^{\mu\mu\mu}$  correction obtained in section 7.6.

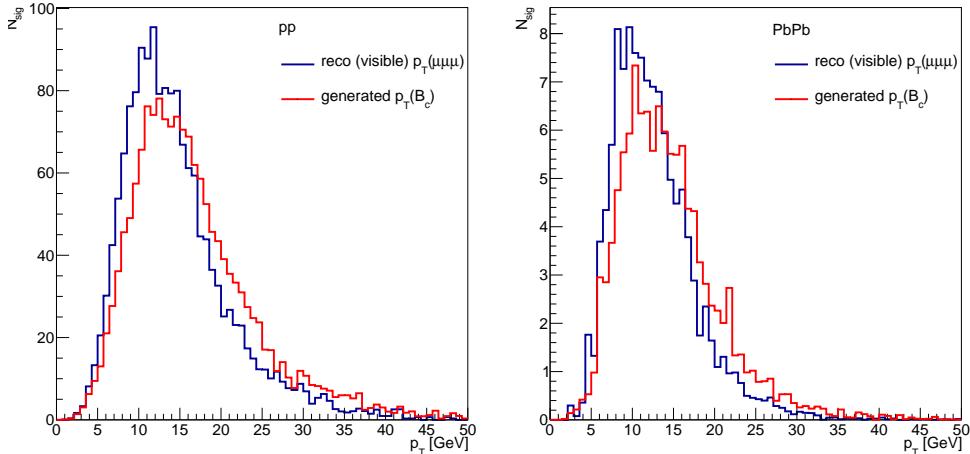


FIGURE 8.4: Reconstructed trimuon  $p_T$  and the full  $p_T$  of the generated  $B_c$ , for all preselected signal MC events, in pp (*left*) and PbPb (*right*).

Fig. 8.5 shows the ratio of the visible and generated  $p_T$ , as a function of the visible (trimuon) mass. The average  $p_T$  ratio is also shown. Despite the different  $p_T^{\mu\mu\mu}$  distribution corrections, the results are similar in pp and PbPb. One could divide the  $p_T^{\mu\mu\mu}$  by this average ratio (red line in Fig. 8.5), so that the corrected signal events would have in average the generated  $p_T$ . This average correction is below 15%, and even below 10% for  $m > 5.4$  GeV. One might rather correct with the median ratio instead of the mean; it would give a smaller correction, because of the long asymmetric tails in the distribution of the ratio.

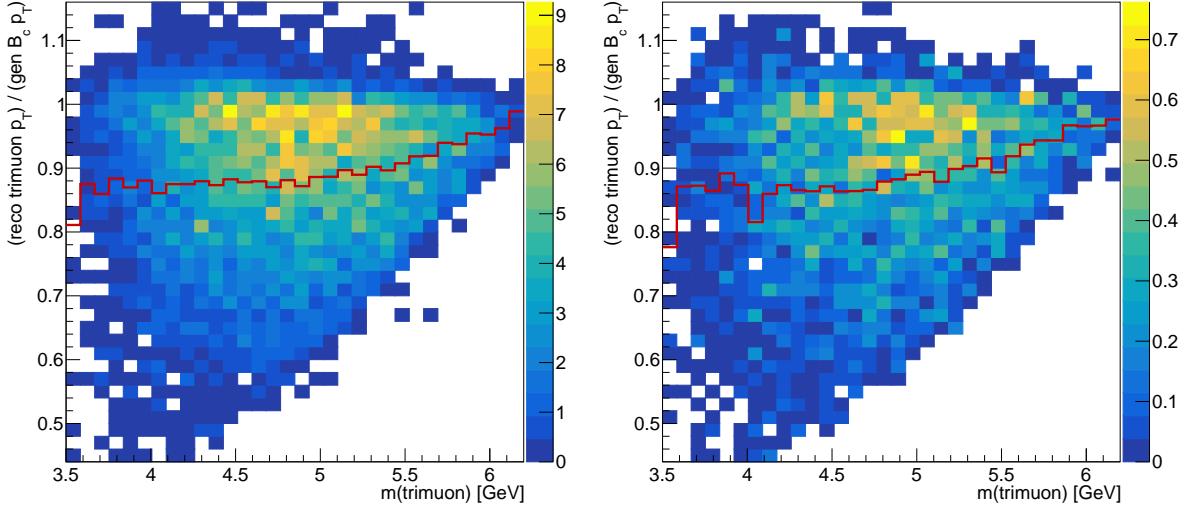


FIGURE 8.5: Ratio of the reconstructed trimuon  $p_T$  to the full  $p_T$  of the generated  $B_c$ , as a function of the visible trimuon mass, for preselected signal MC events, in pp (*left*) and PbPb (*right*). The red line shows the average ratio.

Another way to understand the possible correction is via the ratio of the generated to the visible  $p_T$  as a function of the visible  $p_T$ , shown in Fig. 8.6. Here the average ratio ranges again from 10 to 15%, and is close to flat versus  $p_T^{\mu\mu\mu}$ . As this correction shows no significant dependence on  $p_T^{\mu\mu\mu}$ , it could be applied without worrying about bin migration compared to the results versus  $p_T^{\mu\mu\mu}$ . This actually means that a zeroth order correction can be to simply scale the  $p_T^{\mu\mu\mu}$  axis in the shown results with this constant ratio (neglecting its mass dependence).

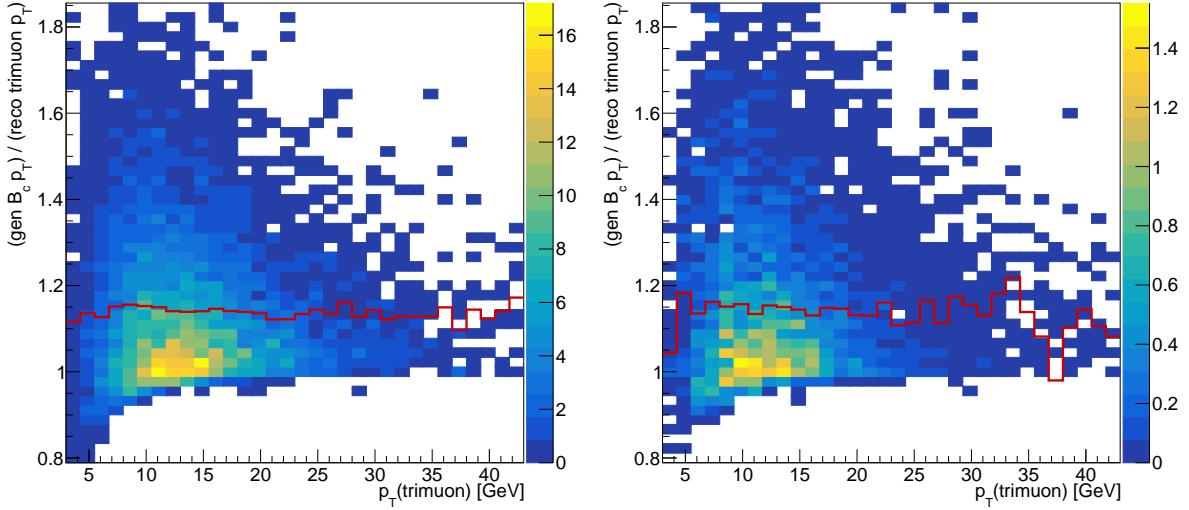


FIGURE 8.6: Ratio of the full  $p_T$  of the generated  $B_c$  to the reconstructed trimuon  $p_T$ , as a function of the visible trimuon  $p_T$ , for preselected signal MC events, in pp (*left*) and PbPb (*right*). The red line shows the average ratio.

At first order, this average correction (as a function of the mass) could be applied to each signal MC event to obtain in average the true  $p_T$  of  $B_c$ . However, there is a large dispersion around the average  $p_T$  ratio, which means events would be attributed to the wrong true  $p_T$  bin. The proper correction would involve unfolding the  $p_T$  distribution, but the available data statistics are very insufficient for this. A systematic uncertainty linked to the  $p_T$  bin migration could be relatively high due to the steepness of the  $p_T$  distribution.

“*La beauté, le charme  
Fugace association  
Ne durent qu’un temps.*<sup>1</sup>”

---

François ARLEO

## Chapter 9

# What do we learn? Results

### 9.1 Cross sections and nuclear modification factors

The corrected yields obtained in section 7.6 are then normalised as explained in section 3.5.5.2. In pp, dividing by the luminosity gives the cross section times the branching fraction for the  $B_c^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) \mu^+ \nu_\mu$  decay. In PbPb, the yields are divided by  $T_{\text{PbPb}} \times N_{\text{PbPb}}^{\text{MB}}$  for the appropriate centrality range, giving the per-event pp-equivalent yield times the branching fraction. The cross section or pp-equivalent yield is then divided by the width of the considered (2D) bin  $\Delta p_T^{\mu\mu\mu} \times \Delta y^{\mu\mu\mu}$  to obtain the differential cross section or pp-equivalent yield. The cross sections and yields include both charges of the  $B_c$  meson. Finally, the  $R_{\text{PbPb}}(B_c^+)$  is calculated from Eq. 3.9, dividing the pp-equivalent PbPb yield by the pp cross section.

The central values for the results integrated on  $p_T^{\mu\mu\mu}$  (and centrality) are obtained with the integrated fit, and the uncertainty from acceptance and efficiency is taken as uncorrelated from the other uncertainties (as for the centrality dependence, see section 7.6.5). I have considered accounting for the correlations between the acceptance and efficiency and the fit, fit method, and tag-and-probe uncertainties, as is done for the  $p_T^{\mu\mu\mu}$  dependence – by taking the uncertainty on the four mentioned sources as the RMS of the sum of the varied corrected yields in each  $p_T^{\mu\mu\mu}$  bin. However, section 7.6.5 (p. 200) explains how this method is somewhat ill-defined, so it is discarded.

This measurement is performed as a function of the visible (trimuon)  $B_c$  transverse momentum, and the results are *not claimed to represent a dependence in the true  $p_T(B_c)$* . The trimuon  $p_T$  is in general lower than the true transverse momentum of the  $B_c$  due to the unreconstructed neutrino of the decay. Section 8.3 has compared the two variables and proposed a correction, that is not applied in the shown results. Considering the relatively large uncertainties on the results, the absence of this correction is a rather sub-dominant systematic change compared to a true  $p_T$  dependence. A zeroth-order approximation of the result versus the true  $p_T$  would be to scale the  $p_T^{\mu\mu\mu}$  axis by  $\sim 1.1 - 1.15$ . This neglects the dispersion of the  $p_T^{\mu\mu\mu}/p_{T,\text{true}}$  ratio around

---

<sup>1</sup>“Beauty, charm  
Fleeting combination  
Last but a while.”

the average as well as the trimuon mass dependence of the average correction. However, for an adequate comparison of the results with theoretical predictions, the latter should perform the leptonic channel decay and be quoted using the trimuon kinematic variables – it is easier to fold than to unfold kinematic distributions.

Table 9.1 presents for the  $p_T^{\mu\mu\mu}$  and centrality bins and for the integrated sample the results of the cross sections and  $R_{\text{PbPb}}(B_c^\pm)$  along with the full uncertainty on these quantities, and the correlation factor between the two bins. This measurement is performed in a phase space divided in two  $p_T^{\mu\mu\mu}$  bins that each have cuts on the rapidity  $|y^{\mu\mu\mu}|$  (see the cuts in section 5.6); the integrated and centrality-binned results integrate over these two  $p_T^{\mu\mu\mu}$  and rapidity regions.

TABLE 9.1: Branching fraction  $BF$  of the  $B_c^+ \rightarrow (\text{J}/\psi \rightarrow \mu^+\mu^-)\mu^+\nu_\mu$  decay times the  $B_c^\pm$  meson pp cross section and PbPb per-event yield, and the  $B_c^\pm$  nuclear modification factor, in the two  $p_T^{\mu\mu\mu}$  and centrality bins, and integrated over the studied regions. The total uncertainties and the correlation between the two bins are also displayed.

system	bin	quantity	value	$\rho_{1,2}$
pp	$6 < p_T^{\mu\mu\mu} < 11 \text{ GeV}$	$BF \times \frac{d\sigma_{\text{pp}}}{dp_T^{\mu\mu\mu} dy^{\mu\mu\mu}} [\text{pb}/\text{GeV}]$	$5.01^{+0.48}_{-0.73}$	0.69
	$11 < p_T^{\mu\mu\mu} < 35 \text{ GeV}$		$0.275^{+0.016}_{-0.023}$	
	integrated	$BF \times \sigma_{\text{pp}} [\text{pb}]$	$76.6^{+5.6}_{-7.5}$	–
PbPb	$6 < p_T^{\mu\mu\mu} < 11 \text{ GeV}$	$BF \times \frac{1}{N_{\text{MB}}^C T_{\text{PbPb}}^C} \frac{dN_{\text{PbPb}}^{B_c}}{dp_T^{\mu\mu\mu} dy^{\mu\mu\mu}} [\text{pb}/\text{GeV}]$	$8.3^{+3.9}_{-2.9}$	0.45
	$11 < p_T^{\mu\mu\mu} < 35 \text{ GeV}$		$0.192^{+0.033}_{-0.033}$	
	centrality 0-20%	$BF \times \frac{1}{N_{\text{MB}}^C T_{\text{PbPb}}^C} N_{\text{PbPb}}^{B_c}(\mathcal{C}) [\text{pb}/\mathcal{C}]$	$119^{+39}_{-37}$	0.54
	centrality 20-90%		$98^{+33}_{-32}$	
	integrated		$110^{+33}_{-32}$	–
$\frac{\text{PbPb}}{\text{pp}}$	$6 < p_T^{\mu\mu\mu} < 11 \text{ GeV}$	$R_{\text{PbPb}}(p_T^{\mu\mu\mu})$	$1.67^{+0.91}_{-0.58}$	0.43
	$11 < p_T^{\mu\mu\mu} < 35 \text{ GeV}$		$0.70^{+0.14}_{-0.12}$	
	centrality 0-20%	$R_{\text{PbPb}}(\mathcal{C})$	$1.55^{+0.54}_{-0.47}$	0.57
	centrality 20-90%		$1.28^{+0.46}_{-0.42}$	
	integrated	$R_{\text{PbPb}}$	$1.43^{+0.46}_{-0.41}$	–

The differential  $B_c^\pm$  branching fraction times cross section and per-event yield are plotted in Fig. 9.1 for the  $p_T^{\mu\mu\mu}$  dependence and in Fig. 9.2 for the integrated samples. In these plots, the fit uncertainty (containing mostly the purely statistical uncertainty) is isolated from the total uncertainty. The pp cross section is compared to the prediction from BCVEGPY2.2, further detailed later in this section. The  $B_c$  nuclear modification factor is shown in Fig. 9.3 as a function of  $p_T^{\mu\mu\mu}$ , and in Fig. 9.4 versus the centrality of the PbPb collision. In these plots, the part of the uncertainty that is uncorrelated between the two bins (such that the uncertainty on the subtraction of the two bins is the quadratic sum of the uncorrelated uncertainties) is isolated from the total uncertainty, as explained in section 8.2.3.

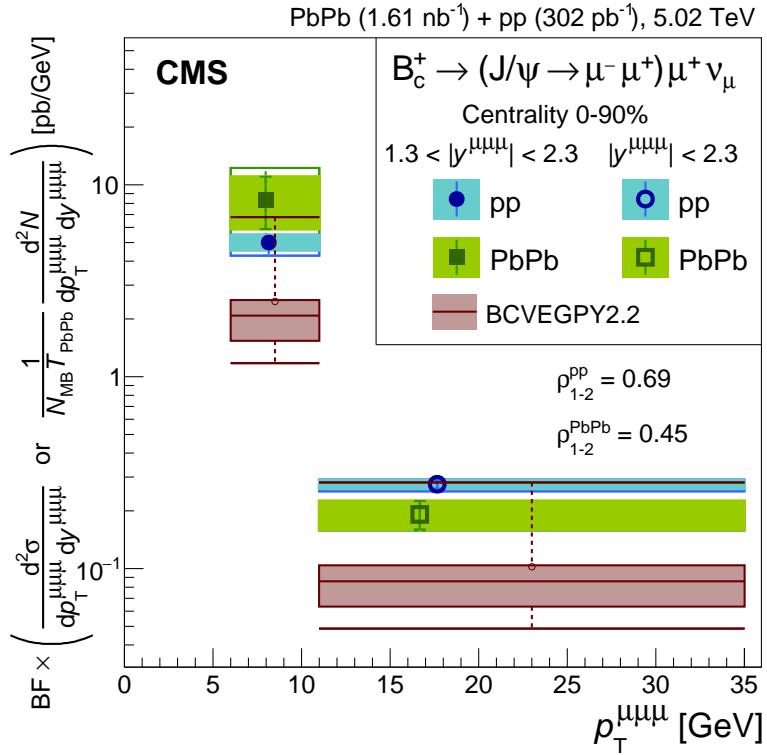


FIGURE 9.1: The cross section or pp-equivalent yield for  $B_c^+$  meson production times the branching fraction of the  $B_c^+ \rightarrow (J/\psi \rightarrow \mu^-\mu^+)\mu^+\nu_\mu$  decay, in pp and PbPb collisions, as a function of  $p_T^{\mu\mu\mu}$ , compared to BCVEGPY2.2 predictions [181]. The bin-to-bin correlation factors are printed. For data, the solid and lighter rectangles show the fit and total uncertainties, respectively. See the text concerning the placement of markers. For BCVEGPY, the red horizontal bars and small open circle respectively mark out the four quartiles and the mean of the distribution of about thirty different  $B_c^+ \rightarrow J/\psi \mu^+\nu_\mu$  branching fraction predictions that multiply the BCVEGPY predicted cross sections.

For the  $p_T^{\mu\mu\mu}$  dependent cross sections, the abscissa markers are drawn at the  $p_T^{\mu\mu\mu}$  obtained in section 7.6.1 applying the Lafferty-Wyatt prescription to the second-step-corrected  $p_T^{\mu\mu\mu}$  spectrum (cf section 7.6.4). The resulting  $p_T^{\mu\mu\mu}$  bin positions are relatively close for the pp or the PbPb spectrum, so their arithmetic average is taken as the position of the  $p_T^{\mu\mu\mu}$  markers for  $R_{PbPb}$ . The centrality bin markers are placed at the minimum-bias average number of participants  $N_{part}$  in the corresponding centrality range, calculated from the model used in section 3.5.5.2. The low statistics do not allow for a doubly-differential distribution (in centrality and  $p_T^{\mu\mu\mu}$ ). The uncertainty is dominated by the acceptance and efficiency correction in the centrality dependence, because the  $p_T^{\mu\mu\mu}$  spectrum variations are integrated on  $p_T^{\mu\mu\mu}$ , but it is very correlated for the two bins. This is the case as well for the integrated cross-sections and modification factor.

The measured values of  $R_{PbPb}$  in the two  $p_T^{\mu\mu\mu}$  bins are probably different, with a significance of  $1.8\sigma^1$ . This suggests that the  **$p_T$  spectrum of  $B_c$  mesons is softened in the QGP**. The  $R_{PbPb}$  in the  $p_T^{\mu\mu\mu} \in [6, 11] \text{ GeV}$  bin stands  $1.2\sigma$  above 1, and is **consistent with high values**, possibly exceeding 2.5. If  $R_{PbPb}$  continues rising at low- $p_T$ , as most  $B_c$  mesons are produced below the  $p_T$  threshold of this measurement, the resulting  $R_{PbPb}$  values integrated over phase space could reach values consistent with the generous predictions of section 2.5. Future measurements could consolidate this hint of  $R_{PbPb}(B_c^+) > 1$ , which would be a totally

<sup>1</sup>Significance values given in this section assume an (asymmetric) Gaussian behaviour of the total uncertainties, taking into account the asymmetric uncertainties, and the bin-to-bin correlations.

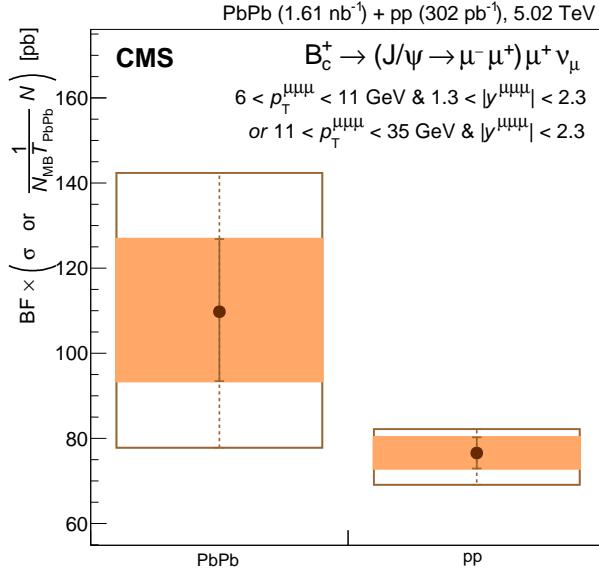


FIGURE 9.2: The cross section or pp-equivalent yield for  $B_c^+$  meson production times the branching fraction of the  $B_c^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)\mu^+\nu_\mu$  decay, in pp and PbPb collisions, integrated over the  $p_T^{\mu\mu\mu}$  and  $y^{\mu\mu\mu}$  regions defined in section 5.6. The solid and lighter rectangles show the fit and total uncertainties, respectively.

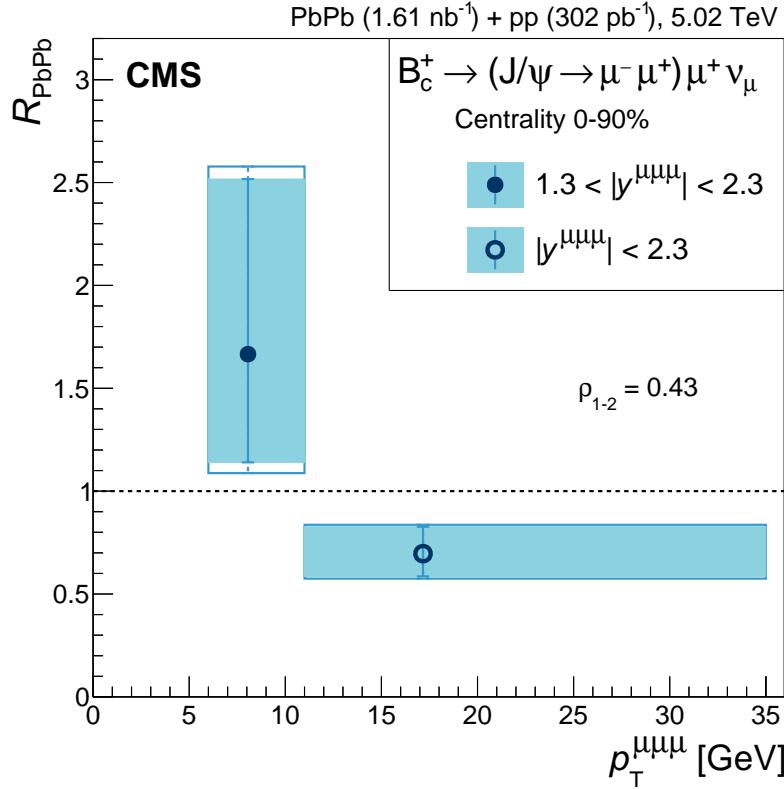


FIGURE 9.3: Nuclear modification factor of  $B_c^+$  mesons in two  $p_T^{\mu\mu\mu}$  bins, each corresponding to a different trimuon rapidity range. The solid and lighter rectangles respectively show the bin-to-bin-uncorrelated and total uncertainties. The bin-to-bin correlation factor is printed. See the text for the placement of markers.

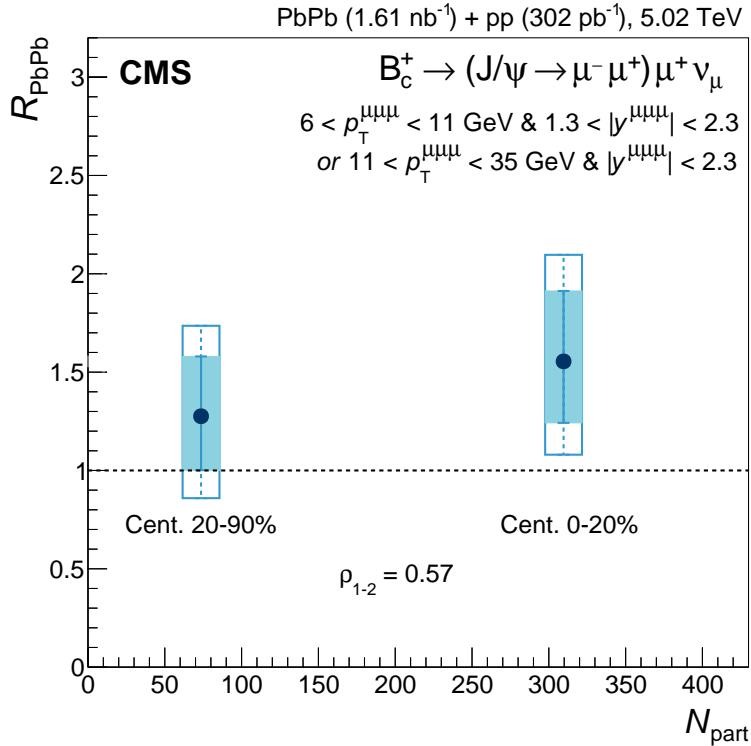


FIGURE 9.4: Nuclear modification factor of  $B_c^+$  mesons in two PbPb centrality bins, integrated on the  $p_T^{\mu\mu\mu}$  and  $y^{\mu\mu\mu}$  regions of section 5.6. See the text for the placement of markers. The solid and lighter rectangles respectively show the bin-to-bin-uncorrelated and total uncertainties.

The bin-to-bin correlation factor is printed.

**unambiguous proof of recombination** – or at least of a new  $B_c$  production mechanism in heavy ion collisions.

The pp cross section can be compared to theoretical expectations from BCVEGPY2.2 [181] (obtained from the samples detailed in section 3.5.2.1), where the  $B_c$  mesons are decayed with EVTGEN1.3. The easiest to compare is the ratio between the differential cross sections in the low- $p_T^{\mu\mu\mu}$  and high- $p_T^{\mu\mu\mu}$  bins: it equals 24.1 in the original BCVEGPY simulation, whereas our measurement gives  $18.2^{+1.3}_{-2.1}$ . This suggests that BCVEGPY overestimates the high- $p_T^{\mu\mu\mu}$  spectrum slope, hence generating  $B_c$  mesons with an underestimated  $p_T^{\mu\mu\mu}$ .

Comparing the absolute cross sections is however less straightforward, because the total cross section given by BCVEGPY must be multiplied by the branching fraction of the studied decay before being compared to this measurement. The  $J/\psi \rightarrow \mu\mu$  branching fraction is precisely known ( $5.96 \pm 0.03\%$  [46]), but not the one for the  $B_c^+ \rightarrow J/\psi \mu^+ \nu_\mu$  decay: it was never measured, and its theoretical predictions range from 1.3% to 7.5%, based on cross-studies in Refs. [192, 205] quoting about thirty different values. Values from Ref [205] actually need to be combined with the ratio of the hadronic to leptonic branching fractions from Ref. [52], whose 15% uncertainty is negligible compared to the spread of the branching fraction predictions.

The total cross section given by BCVEGPY (after integrating it with VEGAS) for both  $B_c^\pm$  charges is 160 nb for the ground state  ${}^1S_0$  and the  ${}^3S_1$  state (almost all decaying to the ground state and a soft photon). When including the other eight excited states simulated in BCVEGPY, the cross section rises to 235 nb. This measurement observes only the ground state, but let us assume here that the feed-down from the excited states to the ground state is close to 100%, hence that the observable total  $B_c$  cross section is 235 nb.

I compare this cross section with two measurements from LHCb, up to the values of the branching fraction. Ref [192] measures the fraction of b quarks that fragment into a  $B_c$  times the leptonic branching fraction  $f_c \times BF(B_c^+ \rightarrow J/\psi \mu^+ \nu_\mu) = 5.0 \times 10^{-5}$  with a 5% uncertainty. The total production cross section of hadrons containing a  $b$  quark was measured in Ref. [206] to be  $72 \mu\text{b}$  for the pseudorapidity range  $2 < \eta < 5$  at  $\sqrt{s} = 7 \text{ TeV}$ . I get two extrapolation factors from BCVEGPY: 0.69 to get to  $\sqrt{s} = 5.02 \text{ TeV}$ , and 4.1 to translate this into a cross section integrated on the whole phase space, giving  $204 \mu\text{b}$ . Multiplying by the fragmentation to  $B_c$  states, I obtain a total cross section times leptonic branching fraction of 10.2 nb. Dividing by branching fractions from 1.3 to 7.5% yields total cross section values from 136 to 785 nb, compatible with BCVEGPY for the highest branching fraction values. However, it should be noted that these highest values are relative outliers, and excluded from considerations in Ref. [192] on grounds that it is not compatible with previous measurements of the width of  $B$  mesons to semi-leptonic final states.

As BCVEGPY seems to overestimate the high- $p_T^{\mu\mu\mu}$  slope, it is fair to correct the fraction of the total cross section that reaches the experimentally studied phase space before attempting a comparison to an absolute normalisation of the cross section from BCVEGPY. As relatively high  $p_T^{\mu\mu\mu}$  ranges are measured here, this fraction is very sensitive to the  $p_T^{\mu\mu\mu}$  spectrum slope, if the reference is the total cross section (which considers mostly low- $p_T^{\mu\mu\mu}$   $B_c$  mesons). So I correct the  $p_T^{\mu\mu\mu}$  spectrum of BCVEGPY with the second-step measurement before recalculating this fraction, which is then found to be 1.5 times higher than the original value. This constant factor multiplies both  $p_T^{\mu\mu\mu}$  bins of the BCVEGPY expectation displayed in Fig. 9.1.

The BCVEGPY absolute cross section seems to be lower than our result, despite two hypothesised corrections that increase its value (the 100% feed-down of excited states into the ground state, and the preliminary correction of the fraction of simulated events that reach the studied phase space). The cross section propagated to the phase space measured here are consistent – within the above caveats – with our result when taking the (relatively disfavoured) highest branching fraction predictions for the  $B_c^+ \rightarrow J/\psi \mu^+ \nu_\mu$  decay. The fact that only the dominant production mechanism in pp collisions (gluon-gluon fusion) is simulated in BCVEGPY could participate in this probable underestimation.

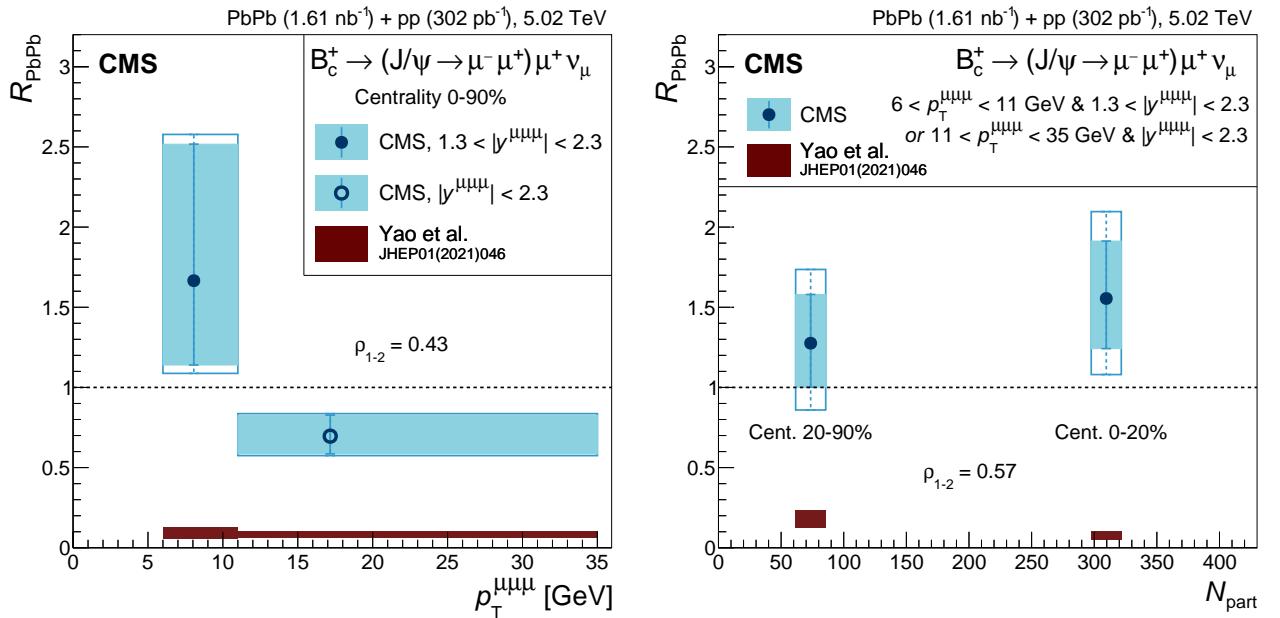


FIGURE 9.5: Comparison of theoretical predictions of YAO et al. (based on Ref. [102]) of the  $B_c$  nuclear modification factor in the two  $p_T^{\mu\mu\mu}$  bins (left) and in the two centrality bins (right). The kinematic distributions of the prediction are those of the full  $B_c$  meson.

Only two theory predictions were received. The first, from RAPP et al., was presented in section 2.5 (Fig. 2.10), because it is integrated over the phase space, so it cannot be directly compared to our results. The second, from YAO et al., is based on the transport model of Ref. [102], a paper containing the  $B_c$  predictions being in preparation. This model includes both correlated ( $b$  and  $c$  from a previously dissociated  $B_c$ ) and uncorrelated ( $b$  and  $c$  from independent hard processes) recombination. It does not however include cold nuclear matter effects like nPDFs, nor takes into account the modification of excited  $B_c$  states that would decay into the ground state measured here. The kinematics are the ones of the full  $B_c$ , and not of the daughter trimuon; as shown in section 8.3, this is non-negligible but sub-dominant effect (10 to 15% difference in  $p_T$  in average) considering the uncertainties of this measurement.

Fig. 9.5 shows the comparison of these predictions to our measurements, in  $p_T^{\mu\mu\mu}$  and centrality bins. The predicted suppression is much higher than what we measure. This might be due to the feed-down from excited states, whose potential recombination is ignored in this prediction. This could point to the importance of recombination of excited states, and in particular to the cross-talk between recombined excited states: even with reasonably small feed-down fractions, if the excited states are more recombined than the primary ground states, the enhancement of ground states will be larger. Ref. [102] declares that this cross talk recombination is key to describe the bottomonium suppression.

Putting aside the global magnitude of the suppression, Yao's prediction shows no  $p_T$  dependence, contrarily to the  $p_T$  dependence having a  $1.8\sigma$  significance in this measurement. Yao also predicts a  $R_{PbPb}$  in 0 – 20% centrality to be twice lower than that in 20 – 90% centrality, whereas this measurement finds no significant centrality dependence. Overall, as the recombination is probably underestimated in this model (that hence mainly shows the effects suppressing the  $B_c$  production), this comparison could suggest that most observed  $B_c$  mesons are produced by recombination rather than in the primary hard scatterings.

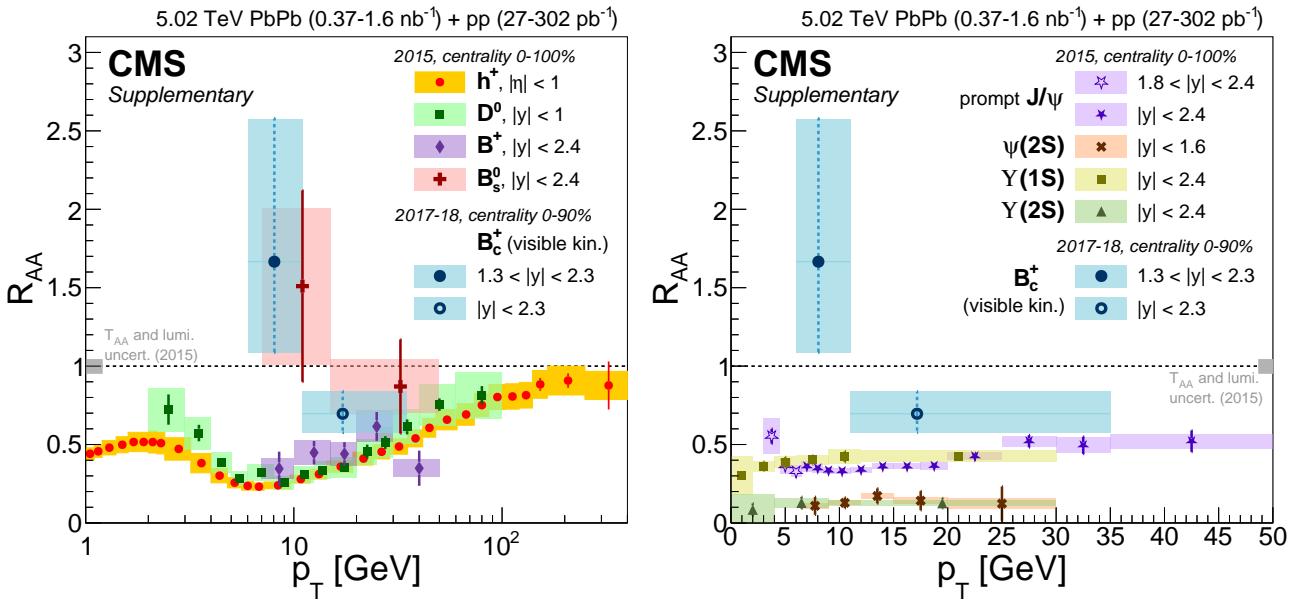


FIGURE 9.6: Comparison of the  $B_c$  modification to the nuclear modification factors measured with CMS for various open heavy flavour mesons (left), or for heavy quarkonia ground and excited states (right). Only the total uncertainty is shown for the  $B_c$ , whereas all other results show the statistical uncertainties as bars and the systematic uncertainties as shaded boxes.

It is also interesting to compare these results to previous CMS results on the modification of heavy flavour mesons. Fig 9.6 shows both a comparison of the  $B_c^\pm$  modification to that of

multiple open flavour mesons (light hadrons [79], and  $D^0$  [77],  $B^\pm$  [76], and  $B_s^0$  [74] mesons, *left*), and to heavy quarkonia ground and excited states (1S and 2S states for charmonia [78] and bottomonia [83], *right*). The suppression of quarkonia is strikingly larger than that of  $B_c$  mesons, pointing towards a different behaviour in the QGP despite a similar quark content. The  $B_c$  modification is found to be close to the  $B_s$  meson one (within the large uncertainties), even though the expected enhancement for these mesons comes from a different source. That of  $B_s$  mesons points to strangeness enhancement, which is mostly thermal and changes the net strangeness number in the QGP; that of  $B_c$  mesons points to charm recombination, which is mostly combinatorial and is not linked to a change in the total charm number.

$B_c$  mesons also appear to be less suppressed than open heavy flavour mesons (excluding  $B_s$ ) and light hadrons. Let us calculate the p-value for the hypothesis that the measured  $R_{\text{PbPb}}$  is below 0.5 (the typical value of the modification of light hadrons, and the maximal value reached by current quarkonia measurements) in both analysis bins. I use a 2D asymmetric Gaussian implementing the final  $R_{\text{PbPb}}$  values and uncertainties, including the bin-to-bin correlation, and calculate its integral in the region where the values of both bins are below 0.5. The resulting p-value (0.5%) yields a  $2.8\sigma$  significance (both when using the  $p_T^{\mu\mu\mu}$  or the centrality bins), tending to confirm that  **$B_c$  mesons are less suppressed than most hadrons**. Considering the current uncertainties and the possible recombination, it is hard to use the  $B_c$  measurements to inform the mass dependence of energy loss at low- to mid- $p_T$ . The overall view of Fig 9.6 also suggests that the  $R_{\text{AA}}$  of all hadrons converge at very high  $p_T$ , supporting the idea of a universal  $p_T$  dependence of the radiative energy loss, which is discussed in the second part of this thesis.

## 9.2 Conclusion

This analysis fulfilled in this thesis studies the  $B_c^\pm$  meson in its semi-leptonic decay  $B_c^+ \rightarrow (J/\psi \rightarrow \mu^+ \mu^-) \mu^+ \nu_\mu$ , where the neutrino is not reconstructed. I measured its pp and PbPb cross sections, as well as its nuclear modification factor in PbPb collisions, as a function of the transverse momentum of the trimuon decay daughter and of the PbPb collision centrality. The presence of  $B_c$  signal in PbPb collisions is significant at the  $6.8\sigma$  level, which makes this analysis **the first to observe  $B_c$  mesons in heavy-ion collisions**. The  $B_c$  meson is significantly less suppressed than most other light and heavy hadrons, which may indicate that the heavy-quark recombination is a significant  $B_c$  production mechanism. The hypothesis of a modification factor independent of  $p_T$  is rejected with a  $1.8\sigma$  significance, favouring a softening of the  $p_T$  spectrum, and the modification in the low- $p_T$  region stands  $1.2\sigma$  above 1. Future measurements with larger datasets will be able to confirm or infirm these hints, potentially leading to an unambiguous proof of heavy quark recombination, namely  $R_{\text{AA}} > 1$  for heavy hadrons.

## Part II

# Partonic energy loss in the quark-gluon plasma



“Off the flying quarks  
In the cosmogonic bath  
Colour is blooming”

---

François ARLEO

## Chapter 10

# Are we lost in colour? Energy loss in the QGP

The last part of this thesis addresses the energy loss that partons undergo when they radiate gluons, due to their interaction with the QGP. Some aspects of energy loss in the QGP were already discussed in section 2.3. Here, I develop on a model focusing on the radiative energy loss of high- $p_T$  partons. At high  $p_T$ , this process should indeed dominate over other collisional energy loss, as well as over other processes affecting partons in heavy ion collisions. The model is that of F. ARLEO in Ref. [2], and predicts a universal dependence of high- $p_T$  hadron suppression in the QGP, from which an average energy loss can be extracted. This chapter explains this result after some generalities about radiative energy loss. In the next two chapters, I explain first a new scaling that I set forth from the energy loss values extracted in the first scaling, then other contributions I brought to the model.

### 10.1 BDMPS gluon emission

The starting point of the model is the **medium-induced gluon emission** spectrum determined by BAIER, DOKSHITZER, MUELLER, PEIGNÉ and SCHIFF in Ref. [207]. Let us present here some aspects of it. First, one broad goal of the study of energy loss in the QGP is to quantify its diffusion properties, which can be encoded in the **transport coefficient**  $\hat{q}$ . It is the typical squared transverse momentum kick received from the medium by a traversing parton in one rescattering ( $\mu^2$ ), in units of the mean path length of this parton in the medium ( $\lambda$ ). Explicitly:

$$\hat{q} = \frac{\mu^2}{\lambda} \tag{10.1}$$

where  $\lambda \equiv \lambda_g$  is by convention here the mean path length of a gluon. The mean free path for any traversing parton in representation  $R$  is  $\frac{N_c}{C_R} \lambda$  (as  $N_c$  is the gluon colour charge), so its associated transport coefficient is rather  $\hat{q}_R = \frac{C_R}{N_c} \frac{\mu^2}{\lambda}$ . The momentum kick  $\mu^2$  in one rescattering is of order of the Debye mass:

$$\mu^2 \sim m_D^2 \sim \alpha_S T^2 \tag{10.2}$$

where Eq. 2.3 is used for the Debye mass.

### Gluon formation time

Let us compute the typical formation time of an emitted gluon, of use in the following. The formation time  $t_f$  is linked to the phase of the  $e^{ik \cdot x}$  term of the wave function of the gluon, where  $k$  and  $x = (t, z, x_2, x_3)$  are the energy and position 4-vectors. Let us consider light-cone coordinates, where  $x_{\pm} = \frac{t \pm z}{\sqrt{2}}$  and the 4-vector dot product is

$$a \cdot b = a_- b_+ + a_+ b_- - a_2 b_2 - a_3 b_3. \quad (10.3)$$

Let us assume for simplification that the gluon is on-shell, so that it is on the light-cone and  $x_- = 0$ . Considering it travels only along the  $z$  direction, the dot product of interest becomes  $k \cdot x = x_+ k_-$ . Looking now at the 4-vector  $k$ , we have  $\omega = k_t = |\vec{k}| = \sqrt{k_z^2 + k_{\perp}^2}$  on the light-cone, so that

$$k_- k_+ = \frac{1}{2}(k_t^2 - k_z^2) = \frac{k_{\perp}^2}{2}. \quad (10.4)$$

In addition,  $k_+ = \sqrt{2}\omega$  and  $x_+ = \sqrt{2}t$ , so the phase of the gluon wave-function is

$$x_+ k_- = \frac{tk_{\perp}^2}{2\omega} \quad (10.5)$$

The phase is also of order  $t/t_f$ , therefore

$$t_f \sim \frac{\omega}{k_{\perp}^2} \quad (10.6)$$

Another approach to obtain this result is to consider  $t_f$  to be the time necessary for the distance  $d$  between the gluon and the parton it is emitted from to be larger than the (inverse of the) gluon transverse wave length, so that the gluon and the parent parton are not coherent any more. This amounts to

$$d \sim v_{\perp} t_f = \frac{k_{\perp}}{\omega} t_f \gtrsim \frac{1}{k_{\perp}} \quad (10.7)$$

which is consistent with Eq. 10.6.

### Different gluon emission regimes

The gluon energy spectrum is expressed as  $\omega \frac{dI}{d\omega}$ . Integrating this spectrum directly provides the total energy lost by a parton traversing the QGP. Depending on the formation time of the emitted gluons, this spectrum behaves differently. Three regimes can be separated, illustrated in Fig. 10.1. In this thesis, the total energy loss is always assumed small compared to the initial parton energy (this is the *soft gluon approximation*), *i.e.*  $\omega \ll E$ , where  $E$  is the energy (or the transverse momentum) of the traversing parton.

The shape of the gluon emission spectrum depends on how the gluon formation time compares to the probability for the traversing parton to interact with the medium (through the mean path length  $\lambda_R$ ), and to the finite length  $L$  of the medium:

- At small formation times ( $t_f \ll \lambda_R$ ), the parton scatters independently with each source of radiation. This is the **Bethe-Heitler** (BH) regime, where the parton scatters  $L/\lambda_R$  times.
- The **Landau-Pomeranchuk-Migdal** (LPM) regime corresponds to intermediate formation times  $\lambda_R \ll t_f \lesssim L$ . In this regime, groups of  $t_f/\lambda_R$  scattering centres behave as single coherent radiators. There are hence  $L/t_f$  gluon emissions.
- In the **fully coherent regime** at large formation times ( $t_f \gg L$ ), the whole medium acts as a unique coherent scattering center.

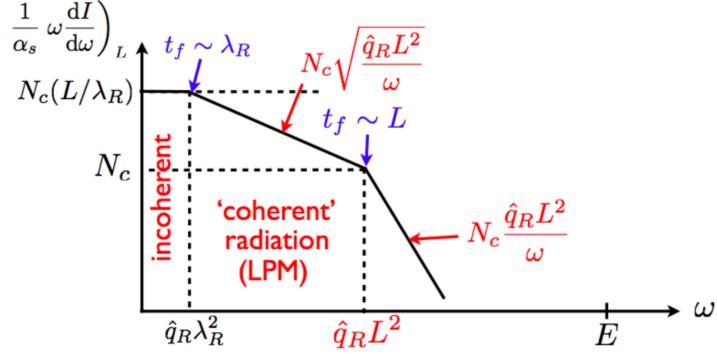


FIGURE 10.1: Regimes in the energy spectrum of the gluons emitted in the medium. Figure from Ref. [18].

Now, the squared transverse momentum kick due to one emitted gluon  $k_\perp^2$  in Eq. 10.6 is the product of the broadening due to one rescattering in the medium  $\mu^2$  with the number of coherent scattering centres that the gluon is emitted from: 1,  $t_f/\lambda_R$  and  $L/\lambda_R$  respectively in the BH, LPM and fully coherent regime. Therefore, in the fully coherent regime:

$$k_\perp^2 = \frac{\mu^2 L}{\lambda_R} = \hat{q}_R L \quad (10.8)$$

and

$$t_f = \frac{\omega}{\hat{q}_R L} . \quad (10.9)$$

In the LPM regime,  $t_f = \frac{\omega}{\hat{q}_R t_f}$  implies

$$t_f = \sqrt{\omega/\hat{q}_R} . \quad (10.10)$$

In the BH regime, we get  $t_f = \frac{\omega}{\hat{q}_R \lambda} = \frac{\omega}{\mu^2}$ , so the probability of emitting a gluon should be proportional to  $1/\omega$  in a fixed-size medium. The gluon energy spectrum  $\omega \frac{dI}{d\omega} \propto \alpha_S$  hence does not depend on the emitted gluon energy, as shows Fig. 10.1.

Conversely, in the LPM regime, the gluon energy spectrum scales with the total number of gluons emitted by the traversing parton, so:

$$\frac{1}{\alpha_S} \omega \frac{dI}{d\omega} \propto \frac{L}{t_f} = \sqrt{\frac{L^2 \hat{q}_R}{\omega}} = \sqrt{\frac{2(C_R/N_c)\omega_c}{\omega}} \quad (10.11)$$

where

$$\omega_c = \frac{1}{2} \hat{q} L^2 \quad (10.12)$$

is the **characteristic total energy loss** for a traversing gluon, as explained below.

### Total energy loss

From integrating over the energies of all emitted gluons, the total energy loss of a parton of colour charge  $C_R$  in a medium of size  $L$ , from Eq. (6.8) of Ref. [207] (corrected by a factor  $\frac{1}{2}$  as explained in Ref. [208]), is, averaging over fluctuations, but for a fixed parton path:

$$\varepsilon \equiv \Delta E = \frac{\alpha_S C_R}{4} \frac{\mu^2}{\lambda_g} L^2 \ln \left( \frac{L}{\lambda_g} \right) \simeq \frac{\alpha_S C_R}{2} \omega_c \quad (10.13)$$

where the term  $\ln(L/\lambda_g)$ , neglected here, is indeed of order 1 in practice. This is interestingly independent of the energy of the considered parton, up to logarithmic corrections (discussed at the end of this section). It is reminded here that this is valid only in the soft gluon approximation, *i.e.* when the fractional energy loss is small.

The dependence  $\varepsilon \sim \hat{q}L^2$  can be understood from the discussion of (in)coherent regimes above and from the spectrum schematised in Fig. 10.1. The LPM regime is the main contributor to the total energy loss, in particular at the largest energies allowed in this regime, corresponding to  $t_f \sim L$ . The latter yields  $\omega = \hat{q}L^2$  using Eq. 10.10. This hence supports the result of Eq. 10.13, based on the integral

$$\int_{\hat{q}\lambda^2}^{\hat{q}L^2} \frac{d\omega}{\sqrt{\omega}} \sqrt{2\omega_c} = 2\sqrt{2\omega_c} \left( \sqrt{\hat{q}L^2} - \sqrt{\hat{q}\lambda^2} \right) \simeq 4\omega_c. \quad (10.14)$$

where here  $\lambda$ ,  $\hat{q}$ , and  $\omega_c$  are those of the traversing parton, and the condition  $L \gg \lambda$  was used.

For  $t_f \gg L$ , the energy spectrum rapidly decreases as  $\omega \frac{dI}{d\omega} \propto 1/\omega$ , which integrates (in the range  $[\hat{q}L^2, E]$ ) into a term  $2\omega_c \ln(\frac{E}{2\omega_c})$  (this logarithmic dependence is also found in the high-energy limit in Ref. [209]). This weak contribution is due to most of the emitted gluon radiation originating from interactions with the vacuum (as a jet in pp collisions) rather than with the medium. Therefore, this regime does not contribute much to the **induced gluon spectrum**, *i.e.* the difference between the AA and pp spectra, and the energy loss from gluons in this regime can be neglected in the integral of the energy spectrum. This logarithmic dependence on the energy of the traversing parton is neglected in this study. However, one could test how this term affects the scaling explained below, by scaling the standard energy loss of Eq. 10.13 by  $1 + \frac{1}{2} \ln(\frac{E}{2\omega_c})$  (comparing the obtained integral with that of Eq. 10.14) in the quenching weight defined later; see section 12.4. Similarly, the Bethe-Heitler regime does not contribute significantly, because the energy of the emitted gluons is too small ( $< \hat{q}\lambda_R^2$ ).

## 10.2 Radiative energy loss model

Let me remind the main message of Ref. [2]. The goal is to determine the average energy loss from measurements of single hadron modification, by modelling only the dominant process at high  $p_T$ : the medium-induced gluon radiation. Cold nuclear matter effects indeed fade away above the saturation scale (of order  $1 - 2$  GeV) of the initial colliding nuclei [210]. So do other effects of hot nuclear matter, such as quark recombination – including that of light quarks, but the principle is the same as that explained in section 2.5. Another advantage of studying high- $p_T$  partons is that the gluon emission spectrum becomes independent of the energy, up to logarithms (see section 10.1). Finally, the pp cross section for producing single hadrons has a power law behaviour at high  $p_T$ , which is key for the scaling explained below.

### Spectrum and quenching weight

The effect of energy loss on the production cross section of a particle species  $i$  in minimum-bias AA collisions is modelled as:

$$\frac{d\sigma_{AA}^i}{dp_\perp}(p_\perp) = A^2 \int d\varepsilon' P_i(\varepsilon') \frac{d\sigma_{pp}^i}{dp_\perp}(p_\perp + \varepsilon') \quad (10.15)$$

where the **quenching weight**  $P_i(\varepsilon')$  is the probability for particle  $i$  to lose the energy  $\varepsilon'$  when traversing the medium. As explained section 10.1, it is independent of the parton energy  $E$  in

the large  $E$  approximation [211]. Therefore, one can define a scaleless  $\bar{P}_i$  such that:

$$P(\varepsilon) \equiv \frac{1}{\omega_c} \bar{P}\left(x \equiv \frac{\varepsilon}{\omega_c}\right) \quad (10.16)$$

Knowing the characteristic energy loss  $\omega_c$  and the mean of the  $\bar{P}$  distribution ( $\langle x \rangle = \frac{1}{2}\alpha_S C_R$  from Ref. [208] or from Eq. 10.13), one can obtain the average energy loss:

$$\langle \varepsilon \rangle = \langle x \rangle \omega_c \quad (10.17)$$

It is denoted as  $\langle \varepsilon \rangle$  when it is related to this phenomenological model and the associated fits, where the behaviour

### Fragmentation into hadrons

What is observed in the experiment is not the parton that escapes the medium, but the hadron that it hadronised into, which only carries of fraction  $z$  of the momentum of its parent parton. To obtain the spectrum of observed hadrons, one must integrate the partonic cross section over the fragmentation functions of every parton species to the considered hadron species. Here, it is assumed that only one parton species dominates the production of a given hadron (*e.g.* gluons dominate the production of light hadrons); this assumption will be challenged in Eq. 12.4. It is also assumed that  $\frac{1}{z}\bar{P}(x/z)$  is a smooth function of  $z$ , so that it can be evaluated at a typical value  $\langle z \rangle$  and taken out of the integral over the fragmentation function. With these assumptions and a change of variables  $x = \varepsilon/\omega_c$ , Eq. 10.15 is only modified due the  $p_\perp$  loss of the hadron being multiplied by  $\langle z \rangle$  compared to that of its parton parent:

$$\frac{d\sigma_{AA}}{dp_\perp} (p_\perp) = A^2 \int dx \bar{P}(x) \frac{d\sigma_{pp}}{dp_\perp} (p_\perp + \langle z \rangle \omega_c x) \quad (10.18)$$

### pp spectrum

Now let us determine the  $p_\perp$  spectrum at high energy in pp collisions. It is assumed to follow a power law at large  $p_\perp$ , which is consistent with the measurements that are fitted for various hadron species and energies:

$$\frac{\sigma_{pp}}{p_\perp dp_\perp} \propto \frac{1}{p_\perp^{n(h,\sqrt{s})}} \quad (10.19)$$

The values of the exponent  $n$  are obtained from fits to measurements of the production of the considered hadron in pp collisions. The  $E \frac{d\sigma}{d^3 p} \propto \frac{d\sigma}{p_\perp dp_\perp}$  measured spectra are actually fitted with the Kaplan spectrum of Eq. 7.27, which tends to a power law at large  $p_\perp$ , and gives the same value of the exponent  $n$  as that of a power-law fit using only very high- $p_\perp$  measurements. A lower limit on  $p_\perp$  of 5 GeV is used in these fits. CMS measurements of Refs. [79, 212, 213] are used for light hadrons production, respectively for  $\sqrt{s} = 2.76, 5.02$ , and  $5.44$  TeV. The  $n$  values are  $5.55 \pm 0.02$ ,  $5.54 \pm 0.02$ , and  $5.29 \pm 0.04$ , respectively. The apparent discrepancy (5%) between the exponents at 5.02 and 5.44 TeV might come from the 5.02 TeV measurements reaching higher  $p_\perp$  values, where some phase-space limitations might start to increase the spectrum slope; however, this difference has a negligible impact on the coming results. A measurement from PHENIX at RHIC [214] is used for the pp spectrum of  $\pi^0$  at  $\sqrt{s} = 200$  GeV, yielding an exponent  $n = 7.70 \pm 0.28$ ; the accessible  $p_\perp$  range being closer to the total available energy at RHIC could be the cause of this faster spectrum decrease. CMS measurements are used for the pp spectrum of  $J/\psi$  [78] and  $D$  [77] at  $\sqrt{s} = 5.02$  TeV. The resulting  $n$  exponents are respectively  $5.90 \pm 0.11$  and  $5.28 \pm 0.11$ . The associated uncertainty will be ignored in the uncertainty on the extracted mean energy loss. Fig. 10.2 shows the  $p_\perp$  spectrum fit in pp for charged hadrons (at two energies), and  $J/\psi$  and  $D$  mesons (at  $\sqrt{s} = 5.02$  TeV). For the last two, I fitted more recent measurements than the 7 TeV data originally considered (see Table 11.1).

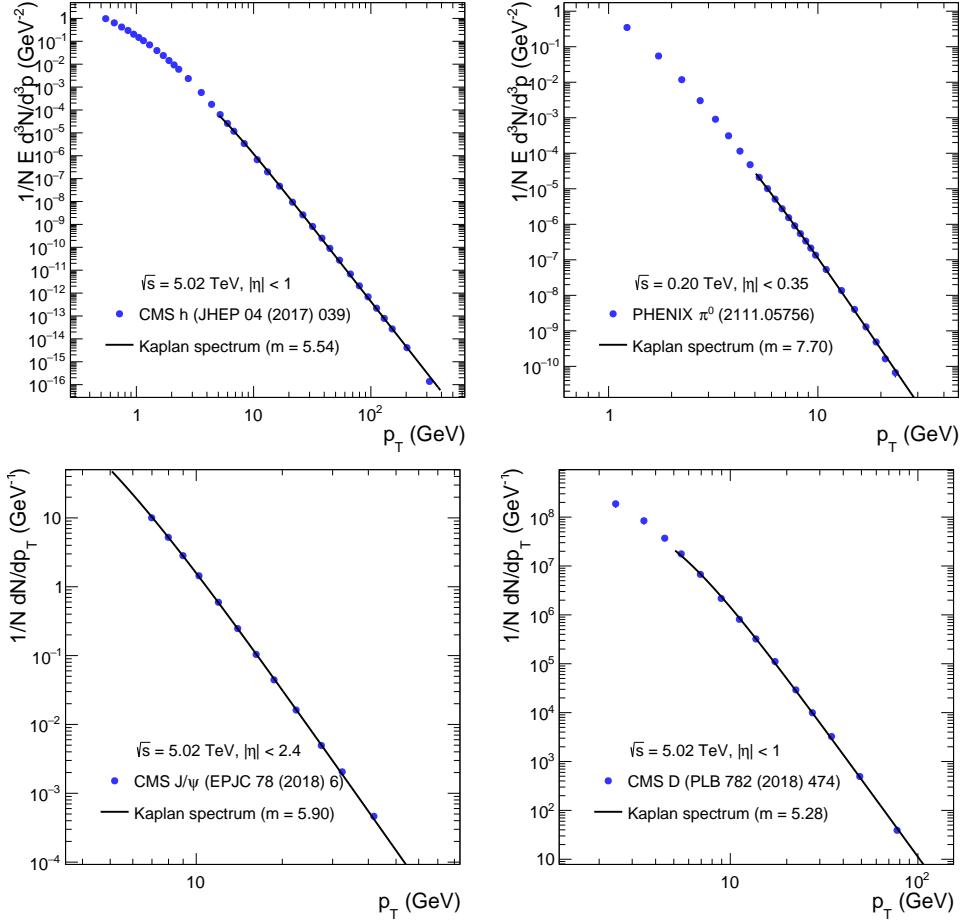


FIGURE 10.2: Fit of the  $p_{\perp}$  spectrum in pp collisions with a Kaplan spectrum (equivalent to a power law at high  $p_{\perp}$ ), at various energies and for various hadrons. *Top right:*  $\pi^0$  mesons at  $\sqrt{s} = 0.20$  TeV. *Top left:* light hadrons at  $\sqrt{s} = 5.02$  TeV. *Bottom:*  $J/\psi$  (*left*) and  $D$  (*right*) mesons at  $\sqrt{s} = 5.02$  TeV. The fit of pp collisions at  $\sqrt{s} = 2.76$  TeV and  $\sqrt{s} = 5.44$  TeV are similar and not shown for brevity.

### Nuclear modification factor scaling

Now one can write the  $p_{\perp}$ -dependent nuclear modification factor  $R_{AA}$  using Eq. 1.25 and the pp cross section dependence of Eq. 10.19:

$$R_{AA}(p_{\perp}) = \int dx \bar{P}(x) \left( \frac{1}{1 + x \frac{\langle z \rangle \omega_c}{p_{\perp}}} \right)^n \quad (10.20)$$

At fixed  $n$ , this only depends on  $\frac{p_{\perp}}{\langle z \rangle \omega_c}$ . Considering that small energy losses dominate the quenching weight, one can exponentiate the power law as:

$$R_{AA}(p_{\perp}) = \int dx \bar{P}(x) \exp \left( -x \frac{n \langle z \rangle \omega_c}{p_{\perp}} \right). \quad (10.21)$$

This assumption was checked in Ref. [2] to be very good for the  $n$  and  $p_{\perp}$  ranges of the considered measurements. It will be checked again on each set of measurements fitted in the following. Eq. 10.20 therefore gives a  $R_{AA}$  that is only a function of the scaling variable

$$\frac{p_{\perp}}{n \langle z \rangle \omega_c}. \quad (10.22)$$

An explicit expression of  $R_{AA}$  is obtained by using the quenching weight computed in Ref. [215] (for ‘outgoing quarks’) as a log-normal empirical parametrisation.

The shape of Eq. 10.20 can be fitted on the  $p_T$ -differential modification of high- $p_T$  hadrons in various heavy ion collision systems. For each  $\sqrt{s}$  and hadron species,  $n$  is known by fitting pp spectrum measurements (shown in Fig. 10.2). Therefore, the values of  $\bar{\omega}_c \equiv \langle z \rangle \omega_c$  can be extracted from fits to various systems. For now, no assumption is made on the average hadronisation fraction  $\langle z \rangle$  (expected to be of order 0.5 for light hadrons [216]), and only values of  $\bar{\omega}_c$  are quoted.

Examples of such fits are displayed in Fig. 10.3. The exact model of Eq. 10.20 is used as fit function. The difference with using the model of Eq. 10.21 including the exponentiation is small. I have superimposed on these plots a fit (dashed red line) using this approximate model along with the associated fitted energy loss, and the resulting  $R_{AA}$  functions are indeed very close, especially towards peripheral collisions and high  $p_\perp$  (where the argument of the exponential  $\frac{p_\perp}{n\bar{\omega}_c}$  is small thanks to a small energy loss). In addition, a somewhat arbitrary lower limit of about 10 GeV (varying slightly for different systems, see section 11.1.2.2) is set on the data points considered in the fit, corresponding to the conditions  $p_\perp \gg m_h, \varepsilon$  in which the model is assumed to hold, and to the fact that radiative energy loss might not be the dominant effect any more at lower momenta. This lower limit is discussed in section 11.1.2.2, yielding a systematic uncertainty on the extracted energy loss. I have superimposed on the fit plots an alternative fit (dotted green line) and the resulting energy loss using a higher  $p_\perp$  threshold. No uncertainties on the  $x$ -axis are considered in the fit. For each bin, the geometric average of the  $p_\perp$  bin limits is taken as the  $x$ -axis bin centre.

Fits of the modification of light hadrons in multiple collision systems, centrality ranges, and at different energies have been performed. I have fitted more sets of measurements than in the original work [2], in particular from collisions of other nuclei than lead (xenon and gold) and at other energies (including at RHIC). Examples of fits in various centrality classes and systems are presented in section 11.1.2.4, along with details of the extraction of the fitted energy loss. Section 11.1.3 illustrates how various systems all scale with the variable of Eq. 10.22 at high  $p_\perp$ . This indicates that radiative energy loss is the dominant process modifying the production in this regime and that it is well described by the model adopted here.

The  $p_\perp$ -dependent modification of J/ $\psi$  [78] and D [77] mesons from CMS have also been fitted, and scale with the  $\frac{p_\perp}{n\bar{\omega}_c}$  variable as the light hadrons. I have added additional J/ $\psi$  and D measurements to be fitted compared to the original publication (see chapter 11).

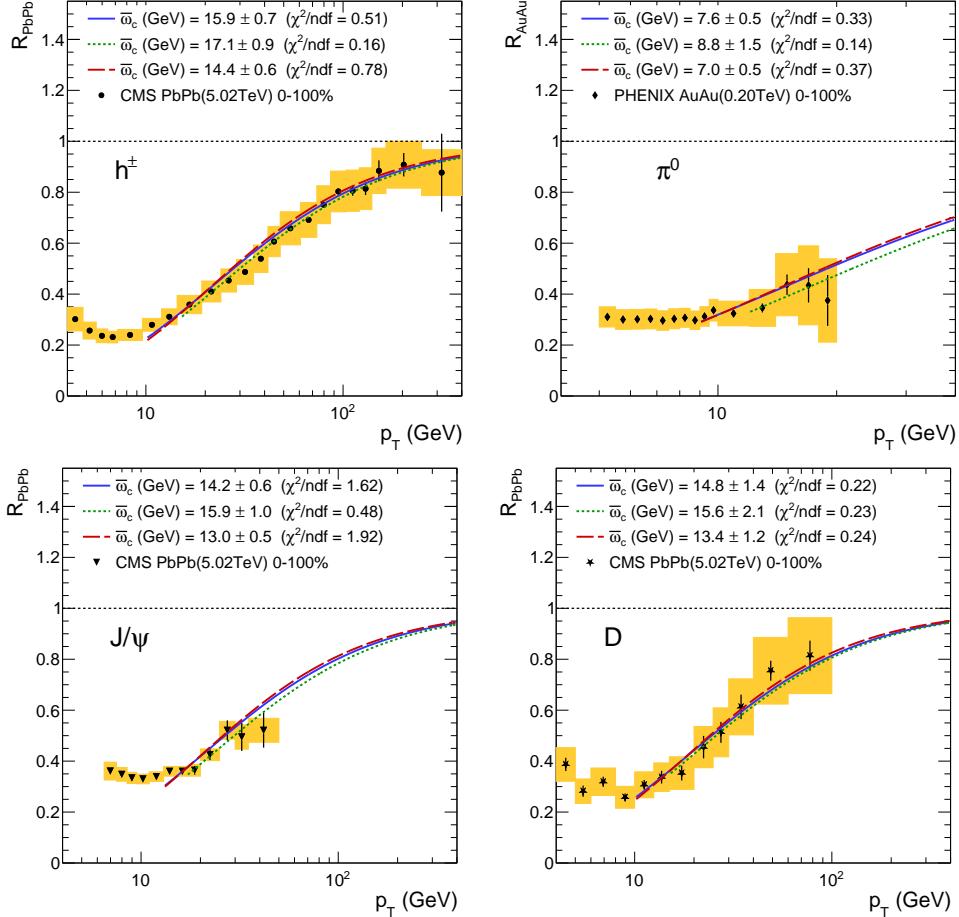


FIGURE 10.3: Inclusive (minimum-bias)  $R_{AA}$  of hadrons, fitted by the energy loss model. The nominal fit function is in blue (Eq. 10.20), while a fit with the exponentiated model (Eq. 10.21) is shown in dashed red, and a fit with a higher  $p_\perp$  threshold is shown in dotted green. *Top left:* light charged hadrons in  $\text{PbPb}$  collisions at  $\sqrt{s_{\text{NN}}} = 5.02 \text{ TeV}$  [79]. *Bottom:*  $J/\psi$  [78] (left) and  $D$  [77] (right) mesons in the same conditions. *Top right:* Neutral pions in  $\text{AuAu}$  collisions at  $200 \text{ GeV}$  [217].

“You may think it strange that I publish anything about colours, without a particular theory of them. But the professed design of this treatise is to deliver things rather historical than dogmatical [...]. This is a treatise, wherein I do not pretend to present my reader with a compleat fabrick, or so much as a model; but only to bring in materials proper for the building.”

---

Robert BOYLE, *Experimenta et considerationes de coloribus... ceu initium historiae experimentalis de coloribus*, Foreword, 1671.<sup>1</sup>

## Chapter 11

# Is it universal? Many systems

### 11.1 Many systems

#### 11.1.1 Measurements to be fitted

Data from more systems than in the original model have been fitted with the predicted  $R_{AA}(p_T)$  shape of Eq. 10.20. Measurements in a different collision system have been fitted, namely in xenon-xenon (with  $A = 154$ ) collisions at  $\sqrt{s} = 5.44$  TeV, and in gold-gold (AuAu, with  $A = 197$ ) collisions at  $\sqrt{s} = 200$  GeV at RHIC. More measurements of the modification of  $J/\psi$  and D mesons have also been included, as well as ALICE measurements of the modification of charged hadrons. All measurements have been extracted from the HepData database. Table 11.1 provides a full view of the measurements fitted with the model of Eq. 10.20. One energy loss value per dataset per centrality class is extracted from these fits. The datasets used for the fits of the  $p_\perp$  spectra in pp are also listed. The observation of a scaling of the high- $p_\perp$  modification in PbPb as well as in XeXe collisions, discussed in section 11.1.3, has been published in Ref. [3], complementing Ref [2].

I need to comment some of these datasets. The results extracted from the  $\pi^0$  measurements with PHENIX at RHIC will be considered on the same footing than those from inclusive charged hadrons, as the latter at LHC energies contains 80-90% of pions. This is equivalent to considering that the average parton parents of pions behave similarly in the QGP as the partons that fragment into inclusive hadrons. In addition, in these PHENIX measurements, the available modification factors at centrality  $> 60\%$  are discarded. This is because the global uncertainties (from Glauber

<sup>1</sup>English translation (from latin) from *The works of the honourable Robert Boyle*, 1744, Vol. 2, p. 3. The original title translates as “Experiments and considerations on colours... meaning the beginning of the experimental history of colours”.

TABLE 11.1: Datasets to be fitted with the radiative energy loss model, each providing a fitted energy loss value per centrality class. The centrality class column also indicates if the fitted pp spectrum originates from this dataset.

Particle	System	$\sqrt{s_{\text{NN}}}$	experiment	already in Ref. [2]? Ref.	pp fit and centrality classes	Kinematic range	Ref.
Light charged hadrons $h^\pm$	<b>PbPb</b> 2.76 TeV	CMS	yes	<b>pp</b> , 0-5%, 5-10%, 10-30%, 30-50%, 50-70%, 70-90%	$ \eta  < 1$ , $p_\perp < 103 \text{ GeV}$	[212]	
		ATLAS	no	0-5%, 5-10%, 10-20%, 20-30%, 30-40%, 40-50%, 50-60%, 60-80%	$ \eta  < 2$ , from $p_\perp < 95 \text{ GeV}$ to $p_\perp < 150 \text{ GeV}$	[218]	
		ALICE	no	0-5%, 5-10%, 10-20%, 20-30%, 30-40%, 40-50%, 50-60%, 60-70%, 70-80%	$ \eta  < 0.8$ , $p_\perp < 50 \text{ GeV}$	[219]	
	5.02 TeV	ALICE	no	<b>pp</b> , 0-5%, 5-10%, 10-30%, 30-50%, 50-70%, 70-90%, 0-100%	$ \eta  < 1$ , $p_\perp < 400 \text{ GeV}$ (250 GeV for centr.> 70%)	[79]	
		CMS	yes	<b>pp</b> , 0-5%, 5-10%, 10-30%, 30-50%, 50-70%, 70-80%, 0-80%	$ \eta  < 1$ , $p_\perp < 103 \text{ GeV}$ (48 GeV for centr.> 50%)	[213]	
$\pi^0$	<b>AuAu</b> 0.20 TeV	PHENIX	no	<b>pp</b>	$ y  < 0.35$ , $p_\perp < 25 \text{ GeV}$	[214]	
				0-10%, 10-20%, 20-30%, 30-40%, 40-50%, 50-60%	$ y  < 0.35$ , $p_\perp < 20 \text{ GeV}$	[217]	
				<b>pp</b> , 0-10%, 0-100%	$ y  < 1$ , $p_\perp < 100 \text{ GeV}$	[77]	
$D^0$	<b>PbPb</b> 5.02 TeV	CMS	yes (except pp)	0-10%, 30-50%, 60-80%	$ y  < 0.5$ , $p_\perp < 50 \text{ GeV}$ (35 GeV for centr.> 50%)	[220]	
		ALICE	no	<b>pp</b> , 0-10%, 10-30%, 30-100%, 0-100%	$ y  < 2.4$ , $p_\perp < 30 \text{ GeV}$ (50 GeV for centr. 0-100%)	[78]	
	<b>PbPb</b> 5.02 TeV	CMS	no (except 0-100%)	0-10%, 20-40%, 40-80%, 0-80%	$ y  < 2$ , $p_\perp < 40 \text{ GeV}$	[221]	
		ATLAS	no				
$J/\psi$	5.02 TeV	ATLAS	no				

modelling) are large in peripheral collisions and the energy loss is very small, which results in uncertainties of order 100% on the extracted energy loss; the used multiplicity measurements (see section 11.2.2) are also unavailable above centralities of 60%.

The  $D$  meson modification factors measured with ALICE actually average over  $D^0$ ,  $D^+$ , and  $D^{*+}$  meson results. The quoted  $D$  measurements from CMS only include  $D^0$  mesons. All  $J/\psi$  and  $D$  meson measurements include cuts to select prompt mesons, that are not daughters of  $B$  mesons.

### 11.1.2 Extraction of the energy loss in various systems and centralities

Figs. 11.1, 11.2, and 11.3 show the ( $\chi^2$  minimisation) fits of multiple systems respectively in the most central, in mid-central, and in the most peripheral considered centrality classes. The experiment yielding each measurement is printed on each canvas. In case two experiments measured the same system, in general the dataset reaching the highest  $p_\perp$  is shown, except for PbPb collisions at 2.76 TeV for which ALICE measurements are shown. The uncertainties shown on the data points are the statistical (black line) and systematic (yellow boxes, excluding luminosity and Glauber) quoted by the experiments. In all Figures, the nominal fit function is in blue (from Eq. 10.20), while a fit with the exponentiated model (from Eq. 10.21) is shown in

dashed red, and a fit with a higher  $p_\perp$  threshold (section 11.1.2.2) is shown with a green dotted line. The energy loss values resulting from these three fits are printed. In Fig. 11.3, the two fits where all points are moved either up or down with the global uncertainties (section 11.1.2.3) are also shown, as these uncertainties have a large impact in peripheral collisions. These peripheral fits are mainly shown to illustrate the current limits of the procedure, as they are somewhat less informative due to the approximations of the model come to a limit (see section 11.1.2.1) and to the large data uncertainties.

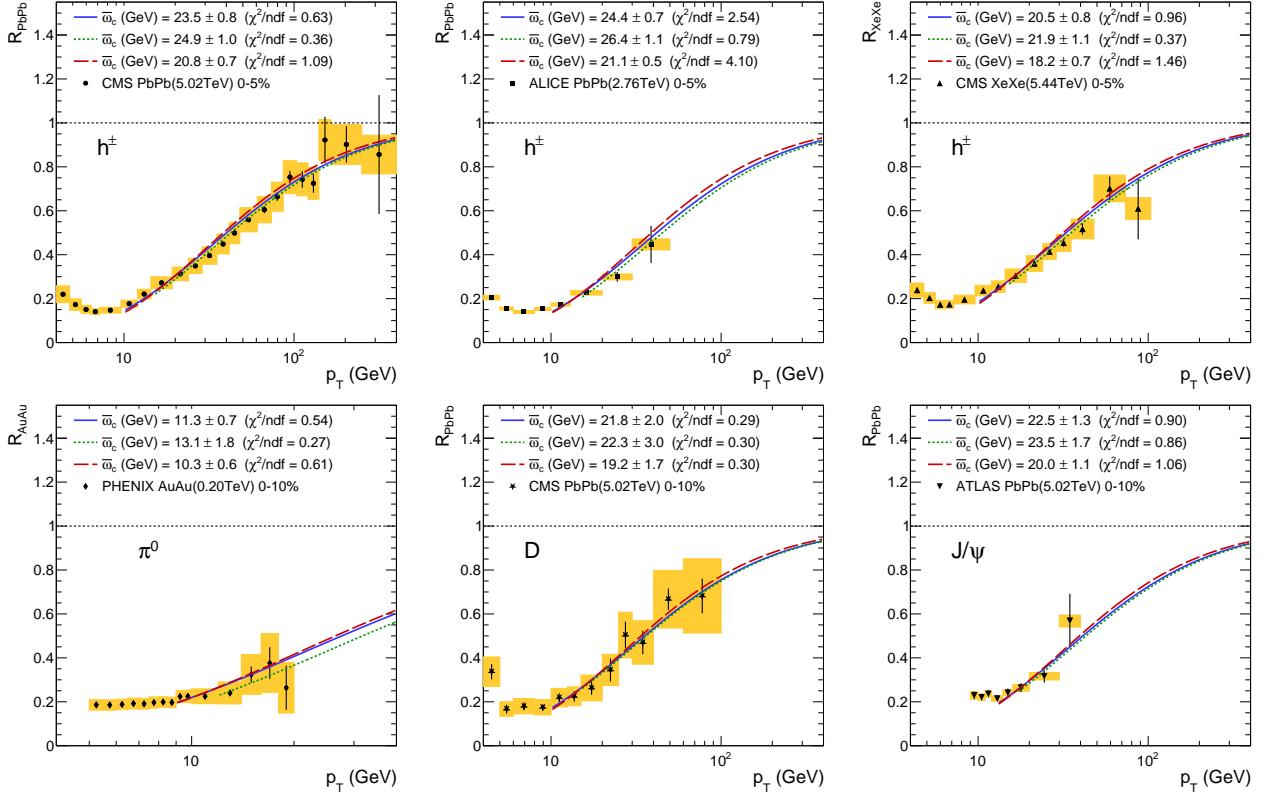


FIGURE 11.1:  $R_{AA}$  of various hadrons in various systems, fitted by the energy loss model, for the most central (0-5% or 0-10%) collisions considered. *Top:* light charged hadrons in PbPb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV (*left*) and at 2.76 TeV (*middle*), and in XeXe collisions at 5.44 TeV (*right*). *Bottom left:*  $\pi^0$  in AuAu collisions at 200 GeV. *Bottom middle and right:* D and  $J/\psi$  mesons in PbPb collisions at 5.02 TeV.

All fits show a good  $\chi^2$  per number of degrees of freedom, below 2.5 (see in section 11.1.2.3 what part of the uncertainties is considered in these nominal fits), except those on ALICE  $h^\pm$  measurements in PbPb 5.02 TeV below 30% of centrality, on ATLAS  $h^\pm$  measurements in PbPb 2.76 TeV below 20% of centrality and on the CMS measurements of  $J/\psi$  below 30% of centrality, which show a less steep rising slope than the model. Fortunately, there are measurements of the same or similar systems by other experiments, that do not confirm these isolated discrepancies. In addition, for all of these tensions (except for the CMS  $J/\psi$  0-10% and the ATLAS  $h^\pm$  2.76 TeV central collision measurements), one of the three fit variations mentioned in sections 11.1.2.2 and 11.1.2.3 leads to  $\chi^2/ndf < 1$ . The relatively low  $p_\perp$  reach of the  $J/\psi$  measurements compared to the  $J/\psi$  mass hinders a strong conclusion on the discrepancy of the model with the CMS measurements in central collisions (not seen in ATLAS datasets).

The following sections (11.1.2.1, 11.1.2.2, and 11.1.2.3) detail how the fit is performed and what uncertainties are considered on the fitted measurements and on the extracted energy loss.

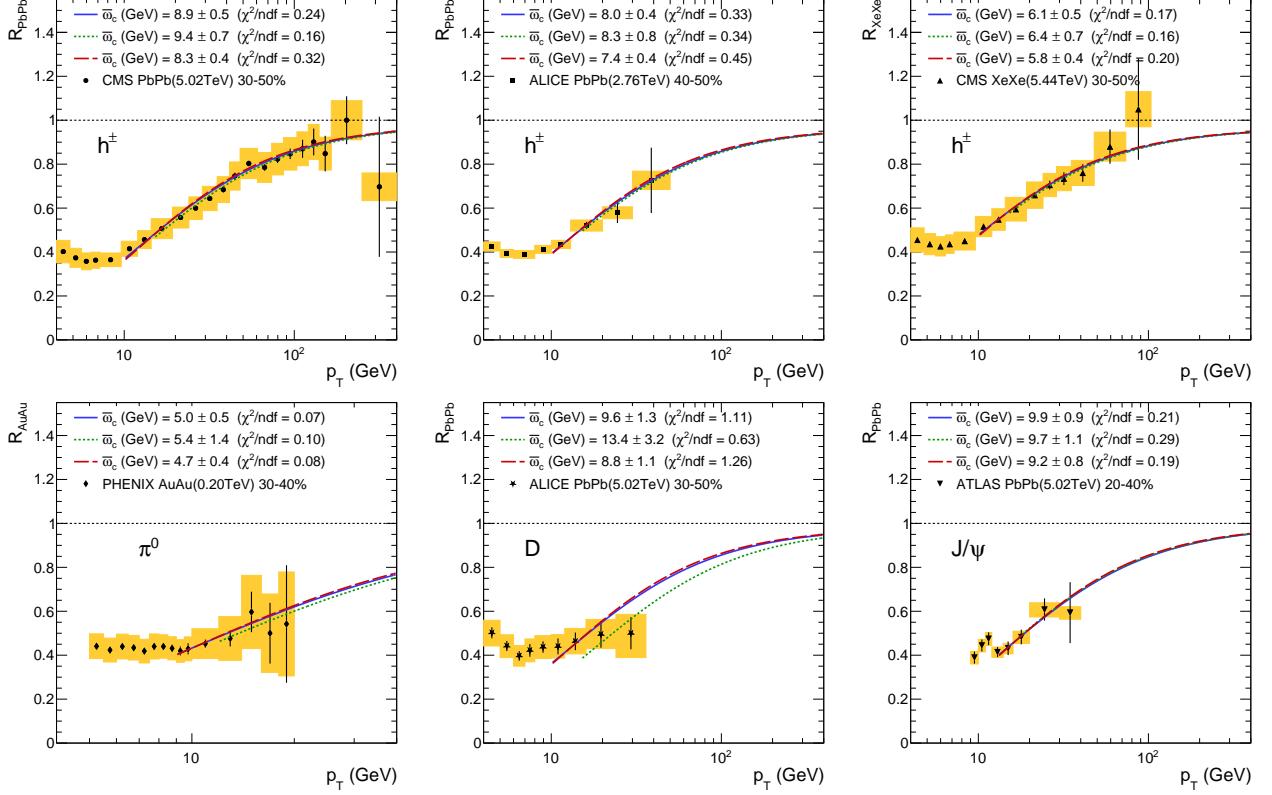


FIGURE 11.2:  $R_{AA}$  of various hadrons in various systems, fitted by the energy loss model, for mid-central collisions (between 30% and 50% centrality). *Top:* light charged hadrons in PbPb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV (*left*) and at 2.76 TeV (*middle*), and in XeXe collisions at 5.44 TeV (*right*). *Bottom left:*  $\pi^0$  in AuAu collisions at 200 GeV. *Bottom middle and right:* D and  $J/\psi$  mesons in PbPb collisions at 5.02 TeV.

### 11.1.2.1 Correction for peripheral collisions

A geometric bias due to the trigger selection of hadronic heavy ion collisions has been demonstrated by LOIZIDES and MORSCH in Ref. [222]. It tends to show an apparent suppression in the  $R_{AA}$  in peripheral centrality classes. I used the correction factors given by the authors, that are less than 2% in centrality classes below 50%, but rise to respectively 7 and 18% for the centrality classes 50 – 70% and 70 – 90% in PbPb.

However, it should be underlined that at these centralities, the framework of the BDMPS energy loss explained in section 10.1 starts to break down: the medium is indeed assumed to contain a few scattering centres – in other words, the path length of the parton through the medium should be large enough. This assumption might be at its limit for centralities above 70–80%, corresponding to path lengths in the QGP of order 1–1.5 fm (as shows Fig. 11.13).

### 11.1.2.2 $p_\perp$ minimal threshold

The used model should hold only at high  $p_\perp$ , because of the two approximations  $p_\perp \gg m, \Lambda_{QCD}$  and  $p_\perp \gg \varepsilon$ . The latter requires a small energy loss compared to the initial parton energy. The  $p_\perp$  lower limit used in the fits is 10 GeV in general, except for the heavier  $J/\psi$  mesons (13 GeV), for the pion measurements from PHENIX (9 GeV, to compensate the fact that the measurement only reaches 20 GeV), and for the centrality ranges above 50% where the energy loss is smaller (2 GeV smaller than the quoted thresholds).

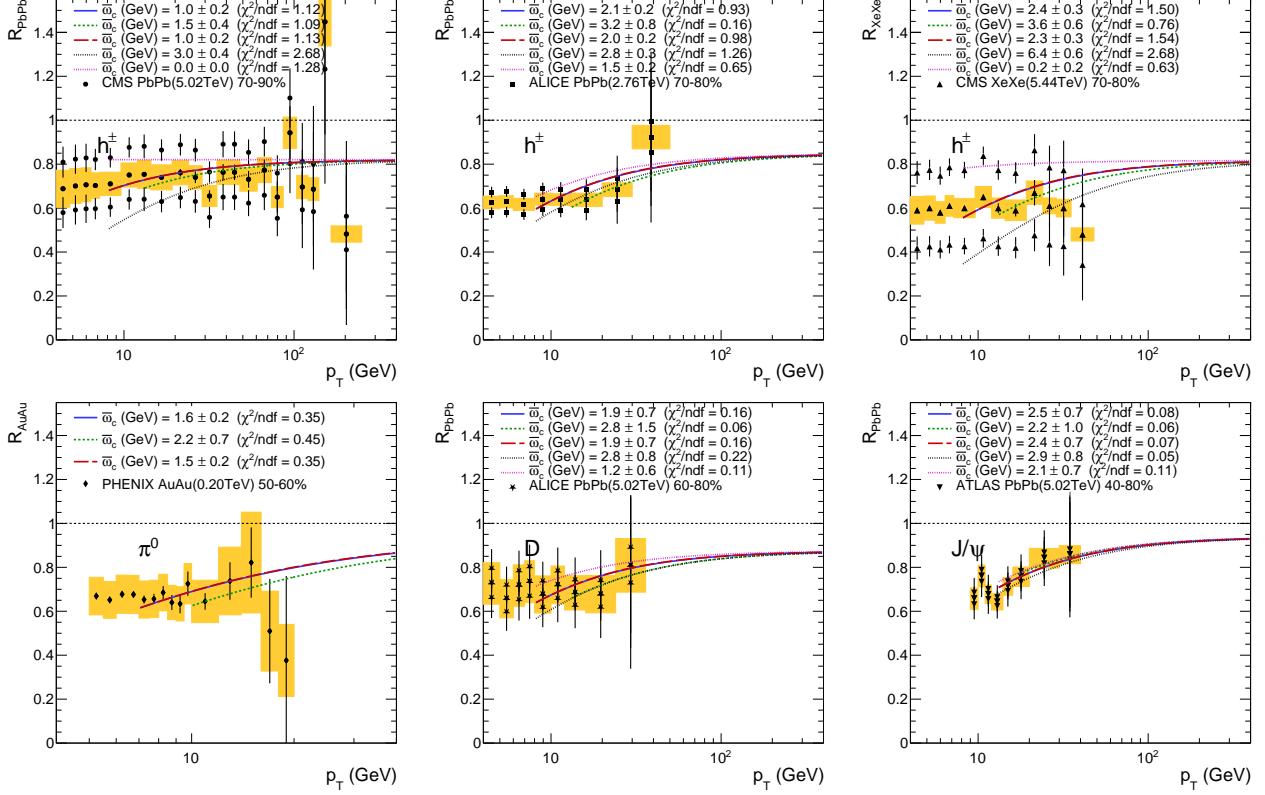


FIGURE 11.3:  $R_{AA}$  of various hadrons in various systems, fitted by the energy loss model, in the most peripheral collisions considered. The data points all moved either up (fitted with the pink dotted line) or down (fitted with the black dotted line) are also drawn. *Top:* light charged hadrons in PbPb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV (*left*) and at 2.76 TeV (*middle*), and in XeXe collisions at 5.44 TeV (*right*). *Bottom left:*  $\pi^0$  in AuAu collisions at 200 GeV. *Bottom middle and right:* D and  $J/\psi$  mesons in PbPb collisions at 5.02 TeV.

As a systematic variation, alternative fits are run with a higher  $p_\perp$  threshold: 5 GeV higher in general, except for the  $J/\psi$  mesons (4 GeV higher) and the pions from PHENIX (3 GeV higher). The absolute difference of the extracted energy loss with that extracted from the nominal fit is taken as a symmetric systematic uncertainty on the energy loss value.

### 11.1.2.3 Global uncertainties on the measurements

Global uncertainties are often quoted on the  $R_{AA}$  measurements, linked to the luminosity and the Glauber modelling (that affects the  $T_{AA}$  values used for the normalisation of the AA yields). The energy loss acts on the model as a rescaling of the  $x$ -axis variable (*i.e.* a shift in the  $\ln(x)$  variable). Therefore, these uncertainties, fully correlated along all bins, have a stronger impact on the extracted energy loss than uncorrelated uncertainties (such as the statistical ones), and should be given a special status.

In addition, the other systematic uncertainties in the used datasets often have high bin-to-bin correlations, as some uncertainty sources act similarly on each  $p_\perp$  bin. This is qualitatively suggested by the data points showing a relatively smooth evolution compared to what the uncertainties would allow if they were uncorrelated along bins. To fully take these correlations into account, the full correlation matrix (containing  $n(n - 1)/2$  correlation factors for  $n$  bins) should be used in the fit; it is however not given, in any of the used measurements. I hence assume that the bin-to-bin correlations are all equal to a single correlation factor of  $f_{\text{corr}} = 0.3$  (or  $f_{\text{corr}} = 0.5$  for the measurements that do not quote a global uncertainty, meaning that it is included in the

quoted systematic uncertainties; this is the case for the two sets of ATLAS measurements). This is equivalent to assuming that a fraction  $\sqrt{f_{\text{corr}}}$  of the systematic uncertainty of each bin is fully correlated to all other bins.

The fully-correlated uncertainty is taken in each bin as  $\sqrt{f_{\text{corr}} \sigma_{\text{syst}}^2 + \sigma_{\text{glob}}^2}$ , while the nominal fit providing the fit uncertainty on the energy loss uses the fully-uncorrelated uncertainties  $\sqrt{(1 - f_{\text{corr}}) \sigma_{\text{syst}}^2 + \sigma_{\text{stat}}^2}$ .  $\sigma_{\text{stat}}$ ,  $\sigma_{\text{syst}}$ , and  $\sigma_{\text{glob}}$  are respectively the statistical, systematic, and global uncertainties quoted by the measurements, where the global one is a simple fraction of the value of the considered bin. Some measurements quote asymmetric uncertainties, which are taken into account.

I could modify the  $\chi^2$  function minimised in the fit to accommodate for these global uncertainties. However, as the correlations are global (*i.e.* it acts only on a global normalisation), it is easier to perform on each  $R_{\text{AA}}(p_\perp)$  dataset two alternative fits, in which the data values in all bins are shifted either up or down with the fully-correlated uncertainties. The resulting uncertainty on the energy loss is half the absolute difference between the energy loss extracted from the two alternative fits.

The fits of measurements in peripheral collisions suffer most from these global uncertainties, notably because the Glauber model uncertainties on  $T_{\text{AA}}$  rise significantly towards peripheral collisions. Fig. 11.3 shows fits in the most peripheral collisions considered here, along with the data points all shifted either up or down, and the fits of these two alternative datasets. The resulting energy loss values show that the associated uncertainty can be of order 50-100% in the most peripheral collisions. As mentioned in section 11.1.2.1, the approximations of the energy loss framework are at their limit in the most peripheral collisions, so the large uncertainties mentioned above might cover for the potential resulting deviations from the energy loss model.

The total uncertainty on the fitted energy loss for each dataset is the quadratic sum of the uncertainty on the energy loss parameter from the nominal fit, that from the  $p_\perp$  threshold variation (section 11.1.2.2), and that from the global uncertainties. The latter (and maybe the one from the  $p_\perp$  threshold) are partially correlated between datasets from the same experiment and/or from the same run, in particular that from luminosity. For now, this correlation will be ignored in the conclusions drawn from these energy loss values, but it could be considered in the future.

#### 11.1.2.4 Resulting energy loss

I show in Fig. 11.4 all the extracted energy loss values, along with their uncertainties. It is reminded here that the fitted characteristic energy loss is  $\bar{\omega}_c$ , *i.e.* the characteristic energy loss times the average fragmentation fraction  $\langle z \rangle$  of the traversing parton into the observed hadron. To obtain the values of the mean energy loss times fragmentation fraction  $\langle \bar{\varepsilon} \rangle$  plotted *e.g.* in Fig. 11.4, the characteristic loss  $\bar{\omega}_c$  is multiplied by the first moment of the quenching weight  $\langle x \rangle \sim \frac{1}{2} C_A \alpha_S \simeq 0.289$ , evaluated from the parametrisation of Ref. [215]. This amounts to  $\alpha_S \sim 0.2$ , evaluated at the scale of the energy loss, *i.e.* 1 – 10 GeV, and considering that the observed hadrons mostly originate from gluons.

Fig. 11.4 makes clear that the energy loss decreases with centrality, as expected due to a smaller medium density and volume; the latter is associated to a smaller path length  $L$  of partons through the QGP entering the energy loss of Eq. 10.13. The aim of section 11.2 is to find variable representative of the activity of each system, in which the energy loss does scale for a large range of systems. The centrality is only a proxy of this variable; it does not take into account the density of the medium, nor the proper dependence on its size.

Another noteworthy feature (already mentioned in Refs. [2, 3]) is that the minimum-bias energy loss values of light hadrons in PbPb and XeXe collisions is of the same order of those

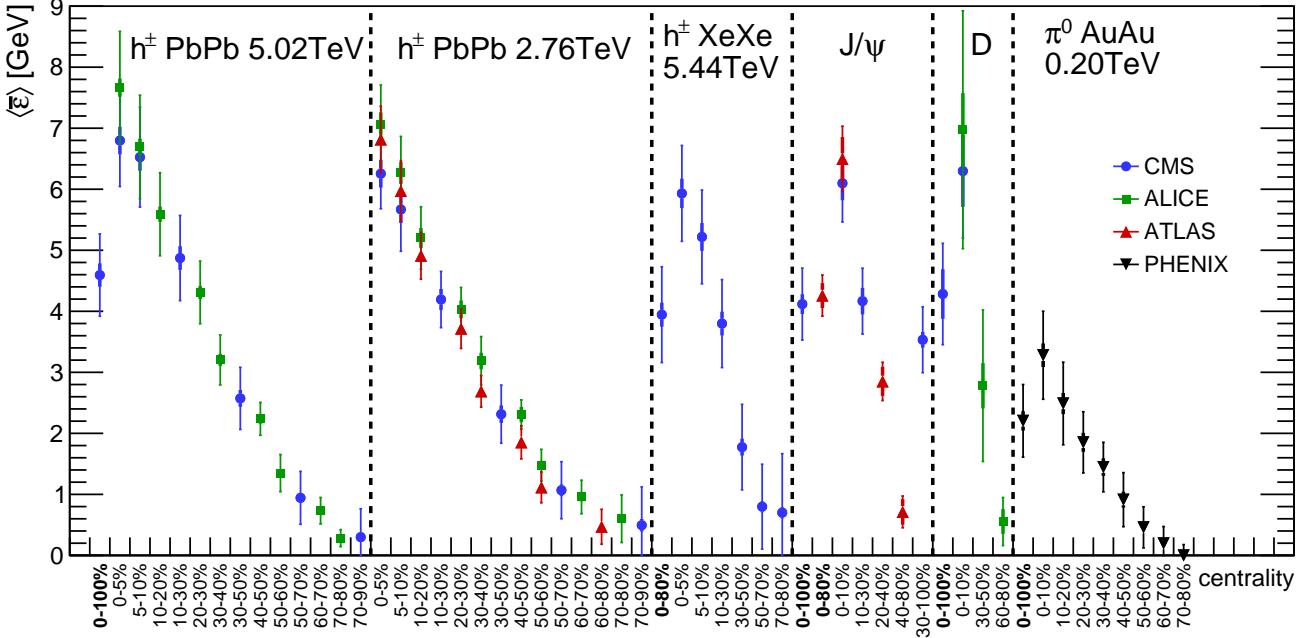


FIGURE 11.4: Summary of all extracted average energy loss values (multiplied by the average fragmentation fraction  $\langle z \rangle$ ), for all considered systems and centrality ranges, from various experiments. Table 11.1 details the fitted datasets yielding these values. The thin vertical bars show the total uncertainty, whereas the thick vertical bars show the uncertainty of the nominal fit (that accounts only for the uncorrelated part of the uncertainties from the measurements).

of  $J/\psi$  and  $D$  mesons in  $PbPb$  collisions at similar  $\sqrt{s}$  values. The  $D$  mesons mostly come from charm fragmentation (associated to a colour charge  $C_F/C_A = 4/9$  smaller than gluons), whereas  $J/\psi$  mostly come from gluon fragmentation, so the latter should naively yield a  $9/4$  higher energy loss. The non-negligible gluon fragmentation for  $D$ , as well as the colour-singlet contribution (that do not lose energy) to  $J/\psi$  mesons, could dampen this factor such that the observed ratio is consistent with 1. A potential dead-cone effect (see section 2.3), decreasing radiation from heavy partons, would have the opposite effect, diminishing the energy loss of  $D$  mesons – but it would affect  $p_\perp$  values much lower than those considered here. More precise  $D$  and  $J/\psi$  measurements towards higher  $p_\perp$  might help enlighten this discussion.

### 11.1.3 Universal $p_\perp$ shape of the modification factor

Fig. 11.5 and 11.6 gather fitted measurements of hadron modification in various systems, where for each set, the  $p_\perp$  is scaled with the corresponding energy loss extracted in section 11.1.2.4 and the value of the pp slope  $n$ . As can be seen on all fit plots above, the exact model of Eq. 10.20 scaling with  $p_\perp/\bar{\omega}_c$  is very close to the universal shape scaling with  $p_\perp/(n\bar{\omega}_c)$ . However, there are slight deviations at lower  $p_\perp$  and higher energy loss (so towards central collisions). Therefore, as the values of  $n$  are close for the studied systems, I still show here the exact model, with the  $n$  value of one of the printed systems. The  $x$ -axis variables of the measurements are though still scaled with the (energy loss and the)  $n$  slope corresponding to their collision system and energy.

The Loizides-Morsch factor (mostly relevant in peripheral collisions) has been applied to the data measurements in this scaling plot (contrarily to the  $R_{AA}(p_\perp)$  fit plots where the function itself was scaled, and the data points were shown as quoted by the experiments). Only the statistical uncertainties are printed; this is motivated by the systematic uncertainties usually being significantly correlated, resulting in a shift in the fitted energy loss that is compensated

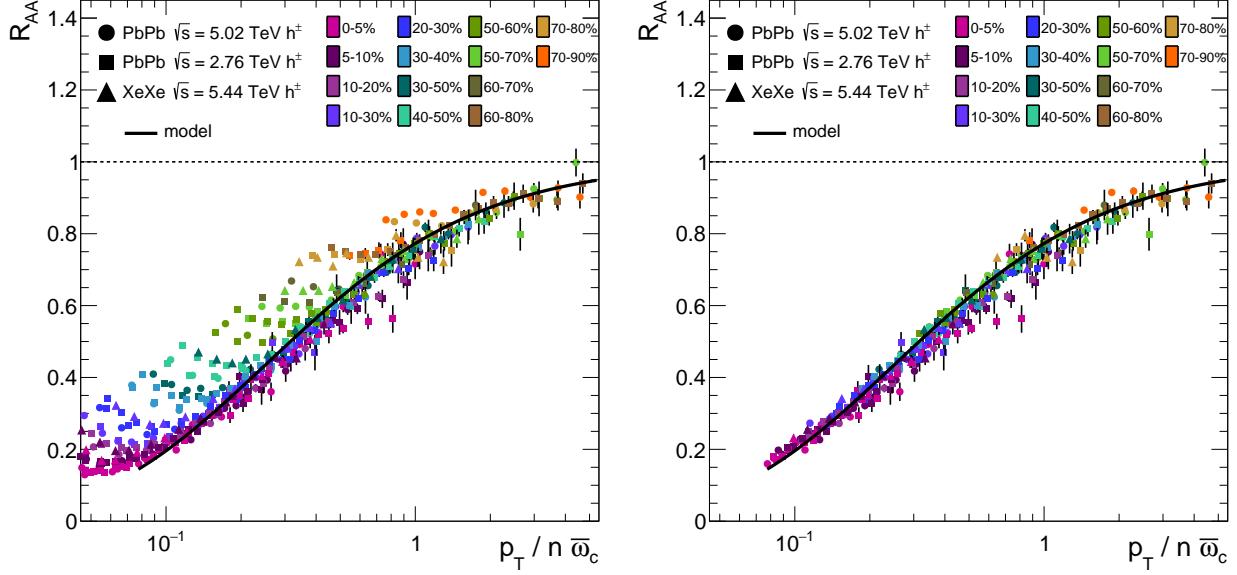


FIGURE 11.5:  $R_{\text{PbPb}}$  of charged hadrons measured in PbPb and XeXe collisions at various energies, where  $p_\perp$  is scaled by the fitted mean energy loss and  $n$  value. The shape predicted by the model is superimposed. On the left (or right) plot, only data points with  $p_\perp > 4 \text{ GeV}$  (or  $p_\perp$  above the threshold of the fit, see section 11.1.2.2) are shown.

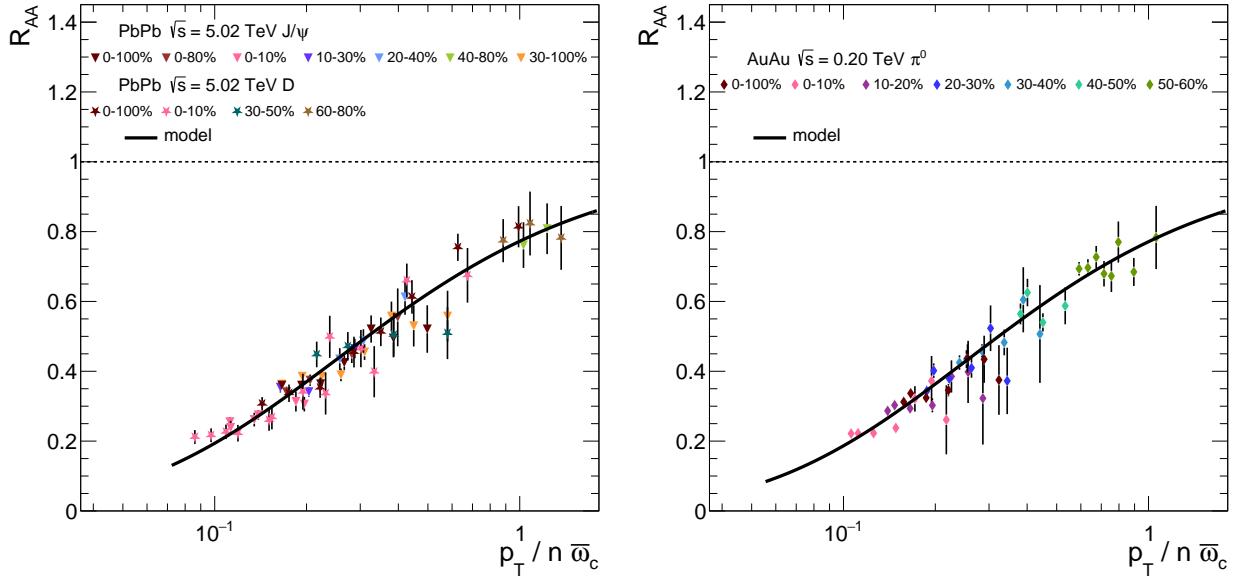


FIGURE 11.6:  $R_{\text{PbPb}}$  of D and  $\text{J}/\psi$  mesons in PbPb collisions at 5.02 TeV (left) or of  $\pi^0$  mesons in AuAu collisions at 200 GeV (right), where  $p_\perp$  is scaled by the fitted mean energy loss and  $n$  value. The shape predicted by the model is superimposed. Only data points with  $p_\perp$  above the threshold of the fit (see section 11.1.2.2) are shown.

by using the  $p_\perp/(n\bar{\omega}_c)$  variable. Data points with a statistical uncertainty larger than 0.05 (or 0.15 for PHENIX measurements, or 0.1 for D and  $\text{J}/\psi$ ) are not printed, for readability reasons.

Ref. [2] whose model this study is based on pointed out that all measurements coincide with the universal  $R_{\text{AA}}$  shape of Eq. 10.21 at high  $p_\perp$ . The multiple additional measurements added here are consistent with this conclusion, even at the much lower energy (200 GeV) reached at RHIC (though these measurements do not reach very high  $p_\perp$ ). This supports the fact that

radiative energy loss is the dominant process leading the suppression of hadrons in this regime, and that the main characteristics of this process are well encompassed by the BDMPS model. In Fig. 11.5 (*left*), the scaling violations are clear when I consider  $p_{\perp}$  values below those chosen as the threshold for the individual fits (between 8 and 13 GeV depending on the considered system).

## 11.2 A new scaling

### 11.2.1 Principle: energy loss versus collision activity

In this section, I aim at finding a variable describing the collision activity, in which the energy loss scales in many different collision systems and energies. The starting point is to consider that the transport coefficient (or the density of scattering centres) is not constant in the medium, and dilutes along with the QGP expansion – in other words, that the QGP created in heavy ion collisions is not a static ‘brick’. This evolution with the time  $\tau$  after the collision was discussed by SALGADO and WIEDEMANN [223]:

$$\hat{q}(\tau) = \hat{q}_0 \left( \frac{\tau_0}{\tau} \right)^{\alpha} \quad (11.1)$$

where the **expansion parameter**  $\alpha$  is 0 for a static medium, while  $\alpha = 1$  is consistent for a typical BJORKEN expansion [126], *i.e.* when the medium dilutes longitudinally (along the beam axis) at speed  $c$ . When the medium volume (proportional to  $1/\tau^{\alpha}$ ) decreases, a traversing parton meets less scatterers, so it loses less energy. The transport coefficient  $\hat{q}_0$  is that of the medium at formation time,  $\tau_0$ . It is actually quoted in Ref. [223] as:

$$\hat{q}_0 \propto n_0 C \quad (11.2)$$

where  $n_0$  is the **initial density of scatterers** in the medium, and  $C$  enters the dipole cross section  $\sigma(\vec{r}) = Cr^2$ . Eq. 11.2 can be qualitatively understood through the use of  $\lambda_g \sim 1/\sigma n_0$  in  $\hat{q} = \mu^2/\lambda_g = \mu^2 n_0 C r^2$ . More precisely, the initial transport coefficient is quoted in Ref. [224] as:

$$\hat{q}_0 = \frac{9\pi}{2} \alpha_S^2 n_0 \quad (11.3)$$

where the term  $\alpha_S^2$  is understandable as coming from the cross section of a hard parton with a parton of the medium.

Now the main point of Ref. [223] is the **dynamical scaling** of the transport coefficient in a non-static medium. The transport coefficient equivalent to the static case is therein:

$$\langle \hat{q} \rangle = \frac{2}{L^2} \int_{\tau_0}^{\tau_0+L} d\tau (\tau - \tau_0) \hat{q}(\tau) = \frac{2\tau_0^{\alpha}}{L^2} \hat{q}_0 \int_{\tau_0}^{\tau_0+L} d\tau \frac{(\tau - \tau_0)}{\tau^{\alpha}} \quad (11.4)$$

where  $L$  is the **path length**, *i.e.* the length of the trajectory of the considered parton in the QGP, in the transverse plane. When the medium is static, *i.e.*  $\alpha = 0$ , the integral simplifies to  $L^2/2$  and one indeed recovers the standard constant transport coefficient. In the integral, the time after the QGP formation is considered equivalent to the distance travelled by the parton (moving at speed  $c$ ). When the medium is large compared to  $c\tau_0$ , which will be assumed from now on (except in specific checks), the equivalent dynamic transport coefficient reads:

$$\langle \hat{q} \rangle \simeq \frac{2\tau_0^{\alpha}}{L^2} \hat{q}_0 \int_0^L d\tau \tau^{1-\alpha} = \frac{2}{2-\alpha} \hat{q}_0 \left( \frac{\tau_0}{L} \right)^{\alpha}. \quad (11.5)$$

Eq. 11.5 and the expression of  $\hat{q}_0$  versus the medium density of Eq. 11.3 can then be incorporated in the standard total energy loss  $\varepsilon$  from Eq. 10.13 to get:

$$\varepsilon = \frac{\alpha_S C_R}{4} \langle \hat{q} \rangle L^2 = \frac{9\pi}{4(2-\alpha)} \alpha_S^3 C_R n_0 \tau_0^\alpha L^{2-\alpha} \quad (11.6)$$

This is the mean energy lost by a parton on a given trajectory of length  $L$  through the QGP, but I will from now on mostly use the path length  $\langle L \rangle$  averaged over all possible parton paths. This assumption, concerning independence of various integrals, will be discussed in section 12.1.2.1.

This general intent is here to prove, from fitting measurements, that the energy loss indeed scales as  $\varepsilon \propto L^\beta$  with

$$\beta \equiv 2 - \alpha \quad (11.7)$$

and possibly  $\beta \neq 2$ , contrarily to the static BDMPS picture commanding  $\varepsilon \propto L^2$ . To this goal, I need more explicit expressions of the density and the average path length. The latter will be discussed at length in section 11.2.7. Concerning the density of partons than the probe parton can scatter with, it should be that of the medium at creation time. At that time, the medium has a volume  $\tau_0 A_\perp$ , where  $A_\perp$  is the area (transverse to the beam) in which nucleons from the nuclei did interact. The number of partons in that medium is a non-perturbative quantity that is hard to evaluate from first principles, so it is approximated as the multiplicity of charged particles observed in the detector, considering that, at first order and in average, each parton will hadronise into one hadron. This quantity is multiplied by  $3/2$  to account for the neutral hadrons that are also produced. This multiplicity should be that at central rapidities ( $\eta = 0$ ), more representative of the early-time medium dynamics (but Bjorken expansion yields a rapidity plateau, so measurements at any rapidity  $|\eta| \lesssim 3$  would be acceptable). This multiplicity is extracted from measurements in each of the studied collision configurations, see section 11.2.2. So, using

$$n_0 = \frac{\frac{3}{2} \frac{dN_{ch}}{d\eta} \Big|_{\eta=0}}{A_\perp \tau_0}, \quad (11.8)$$

I get:

$$\langle \varepsilon \rangle = \frac{27\pi}{8(2-\alpha)} \alpha_S^3 C_R \tau_0^{\alpha-1} \frac{dN_{ch}}{d\eta} \frac{\langle L \rangle^{2-\alpha}}{A_\perp}. \quad (11.9)$$

The transverse area  $A_\perp$  will be discussed and computed in section 11.2.8.

The fits results of section 11.1.2.4 provide values of the energy loss  $\langle \bar{\varepsilon} \rangle_{fit} = \langle x \rangle \bar{\omega}_c = \langle x \rangle \langle z \rangle \omega_c$ , where  $\langle x \rangle = \langle x \rangle_{ref} = \frac{1}{2} \alpha_S C_A \simeq 0.289$  was taken as the scaleless average quenching weight for a gluon. However, let us switch to a general parton, and input

$$\langle \varepsilon \rangle = \frac{\alpha_S C_R}{2} \omega_c = \frac{\alpha_S C_R}{2} \frac{\langle \bar{\varepsilon} \rangle_{fit}}{\langle x \rangle_{ref} \langle z \rangle} \quad (11.10)$$

in Eq. 11.9, so that

$$\frac{\langle \bar{\varepsilon} \rangle_{fit}}{\langle x \rangle_{ref} \langle z \rangle} = \frac{\langle \bar{\omega}_c \rangle_{fit}}{\langle z \rangle} = \frac{27\pi \alpha_S^2}{4(2-\alpha)} \tau_0^{\alpha-1} \frac{dN_{ch}}{d\eta} \frac{\langle L \rangle^{2-\alpha}}{A_\perp}. \quad (11.11)$$

This first equality assumes that only gluons fragment into the hadrons whose energy loss is measured. Were only quarks fragmenting, the energy loss would be reduced by a factor  $C_A/C_F = 9/4$ . However, for the second equality, this hypothesis is not needed: the colour factor can be unknown and the characteristic energy loss  $\langle \bar{\omega}_c \rangle$  be extracted. The impact of the nature of the parton on the conclusions of the used energy loss model will be discussed in section 12.4.

I will show in section 11.2.9 that the **energy loss behaviour** of Eq. 11.11 is **consistent with**

**all the measurements of light hadrons** listed in Table 11.1, with an expansion parameter  $\alpha$  not far from 1, **disproving the static medium** hypothesis. The prefactor could in principle allow for extracting the formation time  $\tau_0$ , but in practice  $\alpha$  is too close to 1 to have a lever arm on this quantity (mind that  $\langle \bar{\varepsilon} \rangle \propto \tau_0^{\alpha-1}$  in Eq. 11.9). I will only provide an order of magnitude of the static-equivalent transport coefficient for typical systems.

To this goal, I need to compute the multiplicities (shown in section 11.2.2), the transverse area and path length for each collision centrality and system. For the last two, an explicit Glauber model is needed to describe the geometry of the collisions in various centrality ranges. As explained in section 11.2.3, I will consider the MC Glauber model of Ref. [35], and three versions of an optical Glauber model I implemented: first considering that nuclei are hard spheres, with a constant QGP density in the interaction region, or with a variable density, and finally using nuclei densities parametrised as Woods-Saxons (*a.k.a* two-parameter Fermi) distributions. Sections 11.2.4, 11.2.5 and 11.2.6 explain the calculations within the various models of the nuclear thickness, of the number of participants and of binary collisions, and of the centrality, respectively. Sections 11.2.7 and 11.2.8 are dedicated to the computation of the path length and the transverse area, respectively, and compare these quantities in the different systems and using the different models.

### 11.2.2 QGP density and multiplicity

The number density of the QGP entering Eq. 11.6 (and ultimately Eq. 11.9) is estimated from the charged particle multiplicity at central rapidities, that is well measured in experiments, in each of the studied collision systems and energies. For PbPb at  $\sqrt{s} = 5.02$  TeV, ALICE [225] measurements are used. For PbPb at 2.76 TeV, ALICE [226] and CMS [227] measurements are averaged, with a weighting considering their respective uncertainties (BLUE method). Similarly, ALICE [228] and CMS [229] measurements are weighted-averaged for XeXe collisions at 5.44 TeV. Finally, PHENIX [230] measurements are used for AuAu at 200 GeV. Fig. 11.7 shows all these measurements of  $\frac{dN_{ch}}{d\eta}$  at mid-rapidity, as a function of the collision centrality.

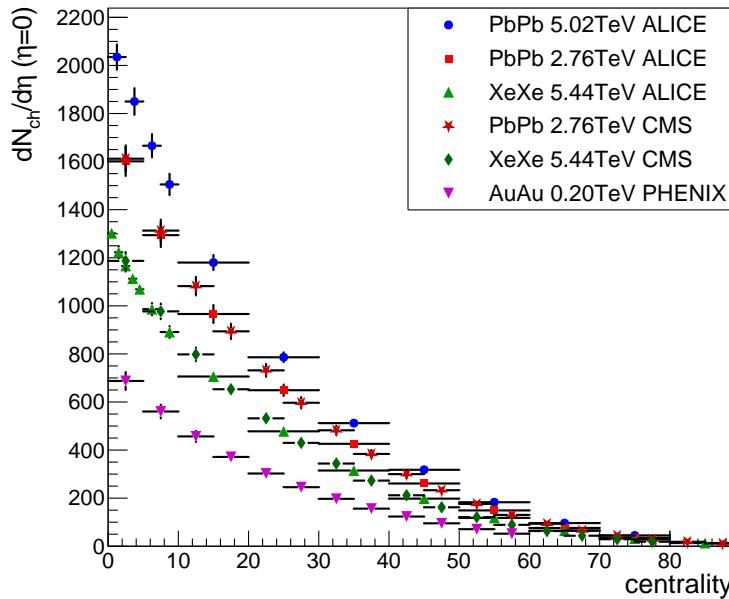


FIGURE 11.7: Multiplicity at central rapidity for various systems and at different energies, measured by ALICE, CMS, and PHENIX.

### 11.2.3 Glauber models

From now on, I will extensively use the Glauber model concepts and definitions introduced in section 1.5.6, such as centrality, the nuclear thickness  $T_A$ , the nuclear overlap function  $T_{AA}$ , the impact parameter  $\vec{b}$  between the two nuclei centres (defining the horizontal  $x$ -axis) and the transverse position compared to the first nucleus  $\vec{s}$ . To contract slightly the expressions, and because this is the case in the studied measurements, I will consider only symmetric collisions of nuclei of same number of nucleons  $A$ .

When fitting all the measurements together with the intent of finding that they scale with Eq. 11.11, I will test four Glauber models, listed below. However, for a given fit, the computations of the area and of the path length will always use the same Glauber model, such that the model variations partially cancel in the ratio  $L^\beta/A_\perp$ .

The first model considered is the **MC Glauber** of Ref. [35], that tabulates  $\langle A_\perp \rangle$  and  $\langle L \rangle$  versus centrality. The MC Glauber does not consider a continuous density of nucleons, but rather simulates events with discrete positions of each nucleon from the colliding nuclei.

Then, I simply consider that nuclei are perfect **hard spheres with a constant density** – both the nucleus density and that of the QGP created in the overlap region. In this case,  $\langle A_\perp \rangle$  and  $\langle L \rangle$  are straightforward and unambiguous to calculate: they are directly based on the geometric limits of the region where the hard spheres cross. No Glauber model is needed in this simple case. No assumption is made in this case on the standard Glauber parameters such as the nucleon-nucleon cross section  $\sigma_{NN}$ .

Still considering **hard sphere nuclei**, I implemented a full **optical Glauber model**, meaning with continuous nuclear densities. It allows for calculating the area and path length considering a non-constant QGP density.

Finally, I run my **optical Glauber** model considering a **Woods-Saxons** distribution for the nucleonic density of the nuclei. In this case, I will need relatively arbitrary choices of regularisation for the path length calculation of peripheral collisions, as the Woods-Saxons profile does not feature a clear cut of the active area for the path length to end.

### 11.2.4 Radius, density and nuclear thickness

#### 11.2.4.1 Numerical values of the radius, $a$ , and $\sigma_{NN}$

Let me first discuss the numerical values of the core parameters of the Glauber models: the radius of the nucleus  $R$ , the skin depth  $a$  (which is the length scale on which the density fades out in the Woods-Saxons) and the hadronic nucleon-nucleon cross section  $\sigma_{NN}$ .

For the Woods-Saxons nuclear profile explained in section 11.2.4.3, I take the nucleus radius and skin depth of the  $^{208}\text{Pb}$  lead nucleus to be  $R = 6.66\text{ fm}$  and  $a = 0.46\text{ fm}$ ,  $R = 6.42\text{ fm}$  and  $a = 0.44\text{ fm}$  for the  $^{197}\text{Au}$  gold nucleus, and  $R = 2.61\text{ fm}$  and  $a = 0.51\text{ fm}$  for the  $^{16}\text{O}$  oxygen nucleus, from Ref. [231], with a  $A^{1/3}$  scaling to get  $R(^{208}\text{Pb})$  from  $R(^{207}\text{Pb})$ , and taking into account the finite nucleon distribution. Ref. [35] also provides the parameters of  $^{129}\text{Xe}$  xenon nucleus:  $R = 5.36 \pm 0.10\text{ fm}$  and  $a = 0.45\text{ fm}$ , where  $a$  was scaled through the skin depths of gold in both references to account for the finite number of nucleons.

In the models using hard sphere nuclei, the number density of nucleons  $\rho_0$  is constant and equal to

$$\rho_0 = \frac{A}{V} = \frac{A}{\frac{4\pi}{3}R^3} = \frac{1}{\frac{4\pi}{3}r_0^3} \quad (11.12)$$

where the value  $r_0 = 1.12\text{ fm}$  (resulting in  $\rho_0 = 0.170\text{ fm}^{-3}$ ) is assumed here, as in frequent conventions. From these values and the number of nucleons  $A$ , the radius  $R = r_0 A^{1/3}$  is calculated, resulting in  $R = 6.64\text{ fm}$ ,  $6.52\text{ fm}$ ,  $5.66\text{ fm}$  and  $2.82\text{ fm}$  for the lead, gold, xenon and oxygen nuclei.

The nucleon-nucleon cross sections in the Glauber model are taken from Ref. [35]: 67.6 mb for PbPb at  $\sqrt{s} = 5.02$  TeV (taken to be the same at  $\sqrt{s} = 2.76$  TeV for simplicity, as this has a sub-dominant effect on the results), 68.4 mb for XeXe at  $\sqrt{s} = 5.44$  TeV, and 70.8 mb for OO at  $\sqrt{s} = 7$  TeV. The uncertainties on  $R$  and  $\sigma_{\text{NN}}$  are negligible for this thesis.

#### 11.2.4.2 Density and nuclear thickness: hard spheres

Using Eq. 1.18 with the constant density  $\rho_0$  of Eq. 11.12, the nuclear thickness at position  $\vec{s}$  from the nucleus centre is:

$$T_A(\vec{s}) = \rho_0 \int_{-\sqrt{R^2 - s^2}}^{\sqrt{R^2 - s^2}} dz = 2\rho_0 \sqrt{R^2 - s^2} \quad (11.13)$$

and it is normalized to  $A$ , so that:

$$\int T_A(\vec{s}) d\vec{s} = \rho_0 V = A \quad (11.14)$$

#### 11.2.4.3 Density and nuclear thickness: Woods-Saxons

The main difference between the last two Glauber models that I will consider is the use of a more realistic number density of nucleons in the nucleus: the standard Woods-Saxons distribution (also called 2-parameter-Fermi-Model) that reads

$$\rho(r) = \frac{\rho_0}{1 + \exp \frac{r-R}{a}}. \quad (11.15)$$

So the nuclear overlap is

$$T_A(\vec{s}) = \rho_0 \int \frac{dz}{1 + \exp \frac{\sqrt{s^2 + z^2} - R}{a}} \quad (11.16)$$

This integral has no analytic form at fixed  $\vec{s}$ , but it should be normalized to  $A$ . Therefore:

$$\int T_A(\vec{s}) d\vec{s} = A = \int \rho(r) dr \quad (11.17)$$

$$= 4\pi\rho_0 \int_0^\infty \frac{r^2 dr}{1 + \exp \frac{r-R}{a}} \stackrel{r'=r/R}{=} 4\pi R^3 \rho_0 \int_0^\infty \frac{r'^2 dr'}{1 + \exp \frac{(r'-1)R}{a}} \quad (11.18)$$

The last integral has a analytical expression with polylogarithm functions  $Li$ , giving:

$$I \equiv \int_0^\infty \frac{r'^2 dr'}{1 + \exp \frac{(r'-1)R}{a}} = -\frac{2}{(R/a)^3} Li_3(-e^{-R/a}) = \frac{2e^{-R/a}}{(R/a)^3}$$

where the last expression uses the limit  $R/a \gg 1$ . However, in practice I integrated  $T_A$  numerically and normalised  $T_A$  by the result.

#### 11.2.5 Number of participants and of binary collisions

##### 11.2.5.1 In Glauber models

For the area and path length calculation, it will be important to know the density of the QGP (in the transverse plane) that a parton travels through. This density will be assumed to be that of the participant nucleons, despite other hypotheses being possibly legitimate. Considering

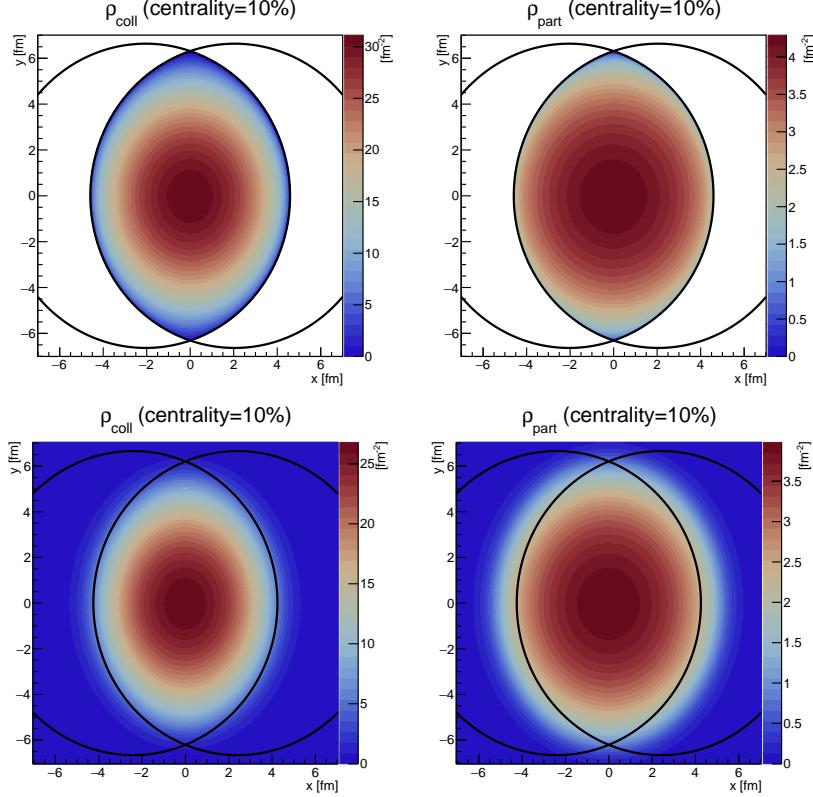


FIGURE 11.8: Comparison of hard spheres (*top*) and the Woods-Saxons (*bottom*) optical Glauber models with lead nuclei. The transverse surface number density of participants  $\rho_{\text{coll}}$  (*left*) and of binary collisions  $\rho_{\text{part}}$  (*right*) are shown for impact parameters corresponding to a centrality of exactly 10%.

the finite nucleon-nucleon cross section, and with similar arguments as in Eq. 1.21, it is:

$$\begin{aligned} \rho_{\text{QGP}}(\vec{b}, \vec{s}) &\propto \rho_{\text{part}}(\vec{b}, \vec{s}) = \rho_{\text{part}}^{\text{A}}(\vec{b}, \vec{s}) + \rho_{\text{part}}^{\text{B}}(\vec{b}, \vec{s}) \\ &= T_{\text{A}}(\vec{s}) \left[ 1 - \left( 1 - \frac{\sigma_{\text{NN}} T_{\text{A}}(\vec{s} - \vec{b})}{AB} \right)^{AB} \right] + T_{\text{B}}(\vec{s} - \vec{b}) \left[ 1 - \left( 1 - \frac{\sigma_{\text{NN}} T_{\text{B}}(\vec{s})}{AB} \right)^{AB} \right] \end{aligned} \quad (11.19)$$

The density of binary collisions, using Eq. 1.19 and 1.20, is

$$\rho_{\text{coll}}(\vec{b}, \vec{s}) = \sigma_{\text{NN}} T_{\text{A}}(\vec{s}) T_{\text{B}}(\vec{b} - \vec{s}) \quad (11.20)$$

These two quantities, as well as their integrals over ranges in  $b$ , are calculated with numerical integrals.

Figs. 11.8, 11.9 and 11.10 show the surface density of nucleon participants and binary collisions for impact parameters  $b$  corresponding to three different centrality classes, for the Glauber models I implemented with hard sphere or Woods-Saxons nucleonic distributions. The black lines show the disks of radius  $R$  of the two nuclei: a hard limit of the presence of nuclei in hard spheres, but only the typical radius at which the density drops in Woods-Saxons. The conversion from  $b$  to a given centrality is explained in section 11.2.6.

In the Woods-Saxons case, the number of participants in peripheral collisions is much lower than what is expected from the hard spheres model, that gives similar values of  $N_{\text{part}}$  for a given centrality range than the MC Glauber model from Ref. [35]. These lower  $N_{\text{part}}$  values in the Woods-Saxons case is due to the very high impact parameters allowed in peripheral classes

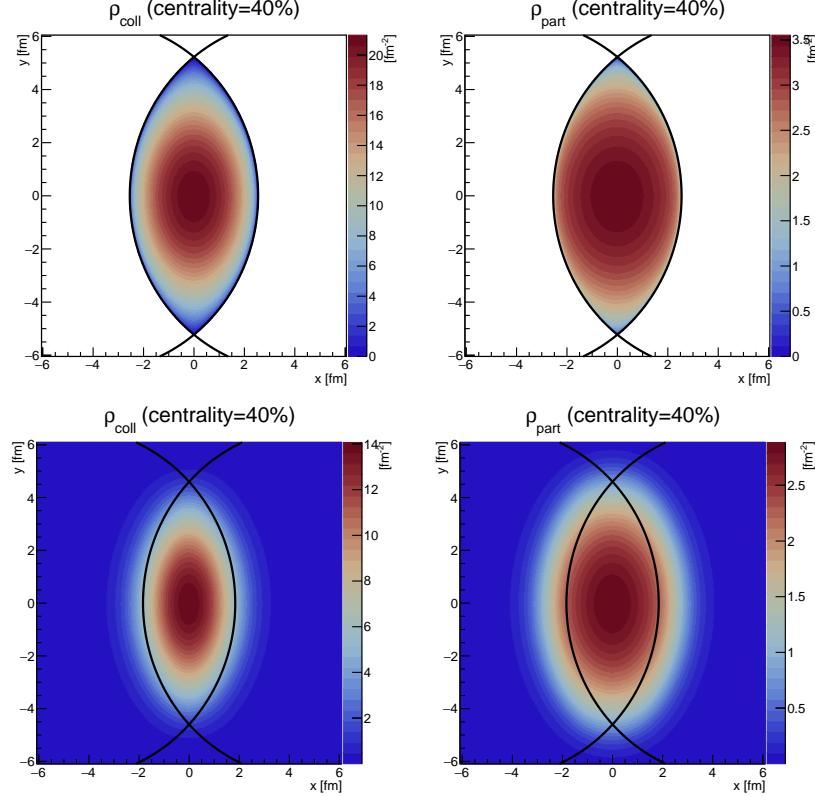


FIGURE 11.9: Comparison of hard spheres (*top*) and Woods-Saxons (*bottom*) optical Glauber models with lead nuclei. The transverse surface number density of participants  $\rho_{\text{coll}}$  (*left*) and of binary collisions  $\rho_{\text{part}}$  (*right*) are shown for impact parameters corresponding to a centrality of exactly 40%.

(see section 11.2.6). This has not been fully investigated yet, but a potential fix is discussed in section 12.2. The  $N_{\text{part}}$  values are kept as such for now, despite the impact on the area and path length estimations; keeping this consistent in the estimations of these two quantities (*i.e.* both in the Woods-Saxons model) will indeed reduce the impact, as the final result depends on the ratio  $L^{\beta}/A_{\perp}$ . In addition, peripheral collisions are associated to small energy loss values, and hence do not have a large impact on the scaling results.

#### 11.2.5.2 $N_{\text{part}}$ in hard spheres

The number of participants in hard spheres is computed here analytically, but will not be used in the later text: it was calculated initially to obtain the impact parameters corresponding to given centralities, but in hard spheres a simple trick is sufficient and  $N_{\text{part}}$  is not needed (see section 11.2.6.1).

In the simple case where all touching nucleons are assumed to interact,  $\rho_{\text{part}}$  is proportional to  $T_A(\vec{s}) + T_B(\vec{s}) \propto \sqrt{R^2 - s^2} + \sqrt{R^2 - |\vec{s} - \vec{b}|^2}$ , simplifying Eq. 11.19. From Eq. 11.20, one also has  $\rho_{\text{coll}}(\vec{b}, \vec{s}) \propto \sqrt{R^2 - s^2} \sqrt{R^2 - |\vec{s} - \vec{b}|^2}$ . This must be integrated over  $\vec{s}$  to obtain  $N_{\text{part}}$  (or  $N_{\text{coll}}$ ). I nevertheless prefer here to calculate from scratch the volume of nucleus A that ‘touches’ or ‘wounds’ the nucleus B, starting exceptionally with asymmetric nuclei. I will then normalise it to  $2A$  at  $b = 0$  (both the volume of both nuclei are fully counted).

Two different volumes must be calculated. The  $x$  axis is taken along  $\vec{b}$ , and  $x = 0$  at the centre of nucleus A. I first call  $d = \frac{R_A^2 - R_B^2 + b^2}{2b}$  the intersection of the spheres ( $d = b/2$  if

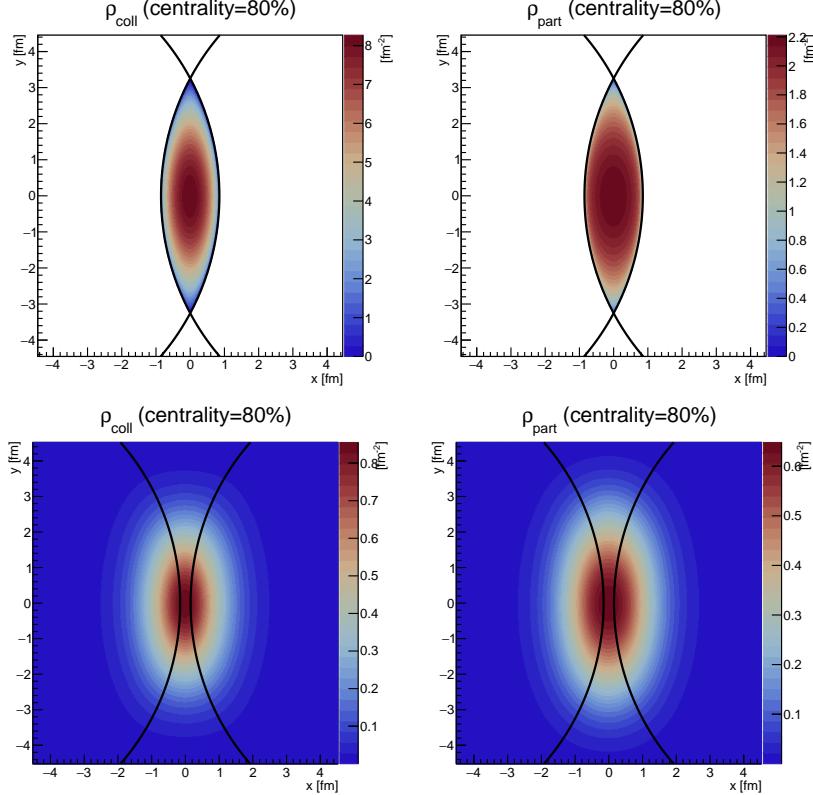


FIGURE 11.10: Comparison of hard spheres (*top*) and Woods-Saxon (*bottom*) optical Glauber models with lead nuclei. The transverse surface number density of participants  $\rho_{\text{part}}$  (*right*) and of binary collisions  $\rho_{\text{coll}}$  (*left*) are shown for impact parameters corresponding to a centrality of exactly 80%.

$R_A = R_B$ ). I call  $V_1$  the touched volume of A that is cut by the almond intersection, but is at  $x < d$ , and  $V_2$  the volume of the endcap of nucleus A, from  $x = d$  to  $x = R_A$ . For symmetric nuclei,  $V_1 = V_2$ . Then:

$$N_{\text{part}} = 2\rho_{\text{constant}}(V_1 + V_2) \quad (11.21)$$

$V_2$  is the sum, for all relevant  $x$  values, of the disk surfaces  $\pi(R_A^2 - x^2)$ :

$$\begin{aligned} V_2 &= \int_d^{R_A} dx \pi(R_A^2 - x^2) \\ &= \pi \left( \frac{2}{3} R_A^3 - \frac{b}{8} (3R_A^2 + R_B) + \frac{b^3}{24} + \frac{1}{8b} (R_B^4 + 2R_A^2 R_B^2 - 3R_A^4) + \frac{1}{24} (R_A^6 + 3R_A^2 R_B^4 - 3R_A^4 R_B^2 - R_B^6) \right) \end{aligned}$$

which, when  $R_A = R_B$ , amounts to

$$V_2 = \pi \left( \frac{2}{3} R^3 - \frac{b}{2} R^2 + \frac{b^3}{24} \right)$$

The other volume is harder to calculate. The  $z$ -directed line touching the edges of A has a length of  $2\sqrt{R_A^2 - x^2 - y^2}$ , for  $x$  and  $y$  being in the low- $x$  part of the almond (on the side of nucleus A). I must integrate the  $z$ -line over the  $(x, y)$  that are in this half-almond surface.  $x$

runs from  $b - R_B$  to  $d$ , and for a given  $x$ ,  $|y|$  runs from 0 to  $\sqrt{R_B^2 - (b-x)^2}$ . Hence:

$$\begin{aligned} V_1 &= \int_{b-R_B}^d dx \int_0^{\sqrt{R_B^2-(b-x)^2}} dy 2\sqrt{R_A^2 - x^2 - y^2} \\ &= 2 \int_{b-R_B}^d dx \left[ (R_A^2 - x^2) \arcsin \left( \sqrt{\frac{R_B^2 - (b-x)^2}{R_A^2 - x^2}} \right) + \sqrt{R_B^2 - (b-x)^2} \sqrt{R_A^2 - R_B^2 + (b-x)^2 - x^2} \right] \end{aligned}$$

and when  $R_A = R_B = R$ :

$$V_1 = 2 \int_{b-R}^{b/2} dx \left[ (R^2 - x^2) \arcsin \left( \sqrt{\frac{1 - (\frac{b-x}{R})^2}{1 - (\frac{x}{R})^2}} \right) + \sqrt{R^2 - (b-x)^2} \sqrt{b^2 - 2bx} \right]$$

This integral has no analytical solution, so it is integrated numerically.

### 11.2.6 Centrality versus impact parameter

The goal of this section is to calculate the  $b$  limits of a given centrality range  $[c_1, c_2]\%$ , *i.e.* to compute  $b$  as a function of centrality. This amounts to calculating the number of PbPb collisions  $N$  versus  $b$  (or  $N_{\text{part}}$ ). The number of participants at given impact parameter  $N_{\text{part}}(b)$  is (proportional to) the wounded volume of Eq. 11.21 for the hard spheres case without Glauber model, or is the numerical integration of Eq. 11.19 over the transverse plane for the Glauber models.

#### 11.2.6.1 Hard spheres without Glauber model

In hard spheres without Glauber model, the explicit expression of  $N_{\text{part}}(b)$  (which is the wounded volume) is not needed. The distribution  $N(b)$  is readily obtained by noticing that the distribution  $\frac{dN}{db^2} \propto \frac{dN}{db}$  is constant, because the nuclei are uniformly distributed in the transverse plane, so the relative position  $\vec{b}$  of two nuclei is uniformly distributed in the space of  $\vec{b}$  vectors. This is true for nuclei that do touch (equivalent to interacting, in this simple model), *i.e.* for  $b < 2R$ . After integration,  $N(b) \propto b^2$ , *i.e.*  $b \propto \sqrt{N(b)}$ . This was checked with a very simple MC. So if I define  $N(b_{\mathcal{C}=100\%}) = 1$  (corresponding to 100% centrality, *i.e.*  $b_{\mathcal{C}=100\%} = 2R$ ), I directly get the  $b$  corresponding to a given centrality  $\mathcal{C}$  as:

$$b_{0,\mathcal{C}\%} = \sqrt{\mathcal{C}} \times b_{100\%} \quad (11.22)$$

These simple  $b$  limits of the centrality ranges are used in the case of hard spheres without Glauber model, because it does not depend on any assumption on  $N_{\text{part}}$  and  $\sigma_{\text{NN}}$ . As explained below, for the Glauber models both with hard spheres and with Woods-Saxons, the full  $N_{\text{part}}$  from integrating Eq. 11.19 and the distribution of AA collisions from Eq. 11.23 are rather used.

#### 11.2.6.2 Glauber model case

In case of a full Glauber model, I start from the function  $N_{\text{part}}(b)$ , that is strictly monotonous, so the function  $b(N_{\text{part}})$  (and  $\frac{db}{dN_{\text{part}}}$ ) is obtained by (numerical) inversion. The goal is then to obtain the distribution  $\frac{dN}{db}$ , from which the  $b$  limits of the centrality classes can be deduced. We start with:

$$\frac{dN}{db} = \frac{dN}{db^2} \frac{db^2}{db} = 2b \frac{dN}{db^2}. \quad (11.23)$$

The term  $\frac{dN}{db^2}$  would be constant if the nucleon-nucleon cross section  $\sigma_{NN}$  was infinite, as was assumed in section 11.2.6.1, yielding  $\frac{dN}{dN_{\text{part}}} \propto b$ . However,  $\sigma_{NN}$  is finite, so  $\frac{dN}{db^2}$  should be proportional to the probability of two nuclei to have at least one hadronic interaction when they are at distance  $b$ . This probability  $p_{\text{hadr}}(b)$  is exactly that of Eq. 1.21, depending on  $\sigma_{NN}$  and  $T_{AA}(b)$ . The latter is calculated as a numerical integral for each  $b$ . Fig. 11.11 shows  $\frac{dN}{db} \propto b p_{\text{hadr}}(b)$  for both Glauber models in PbPb collisions.

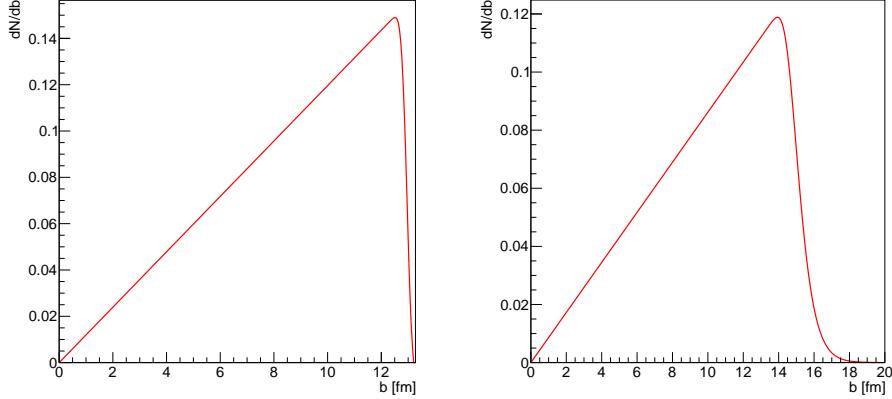


FIGURE 11.11: The distribution of hadronic PbPb collisions simulated in the custom Glauber model with hard spheres (*left*) or Woods-Saxons (*right*) nuclei, versus  $b$ . Integrating this normalised distribution from 0 to  $b'$  yields the centrality corresponding to  $b'$ , in percent. Higher  $b$  values than  $2R$  are of course reached in the Woods-Saxons case.

Fig. 11.12 shows the more standard centrality plots, meaning the number of events versus the event activity; here it is based on this  $N_{\text{part}}$  dependence:

$$\frac{dN}{dN_{\text{part}}} = \frac{dN}{db} \frac{db}{dN_{\text{part}}} . \quad (11.24)$$

In this expression, I calculate numerically the derivative  $\frac{db}{dN_{\text{part}}}(N_{\text{part}}) = \frac{b(N_{\text{part}}(1+w)) - b(N_{\text{part}}(1-w))}{2wN_{\text{part}}}$ , with  $w = 0.002$  giving a stable result.

The  $\frac{dN}{dN_{\text{part}}}$  distribution in the Woods-Saxons case shows an unusual feature at the highest  $N_{\text{part}}$  values: it does not decrease sharply as do published Glauber models and the hard spheres model implemented here, which show significantly larger  $N_{\text{part}}$  values for given centrality ranges. It might be a numerical instability or an actual feature of a Woods-Saxons optical model, in which all very high  $b$  values still result in substantial NN interactions. This is not investigated yet; temporary solutions are discussed in section 12.2, but they are not applied here and the resulting  $b$  and  $N_{\text{part}}$  limits of centrality classes are kept as such.

The  $b$  (or  $N_{\text{part}}$ ) value corresponding to a centrality of  $C\%$  is that for which the cumulative of the distribution of Eq. 11.23 (or Eq. 11.24), normalised to 1 and starting from the lowest  $b$  values (or highest  $N_{\text{part}}$  values), equals  $C\%$ . These  $b$  limits will be those of the  $b$  integrals in the calculation of the path length and transverse area in a given centrality class, for the two custom Glauber models.

### 11.2.7 Path length

When a parton is created at a given point in the transverse overlap area in a given direction, the distance it covers before exiting the QGP created in the collision is the *path length*. The *average path length* of particles created in collisions at a given  $b$  enters Eq. 11.9 that I intend

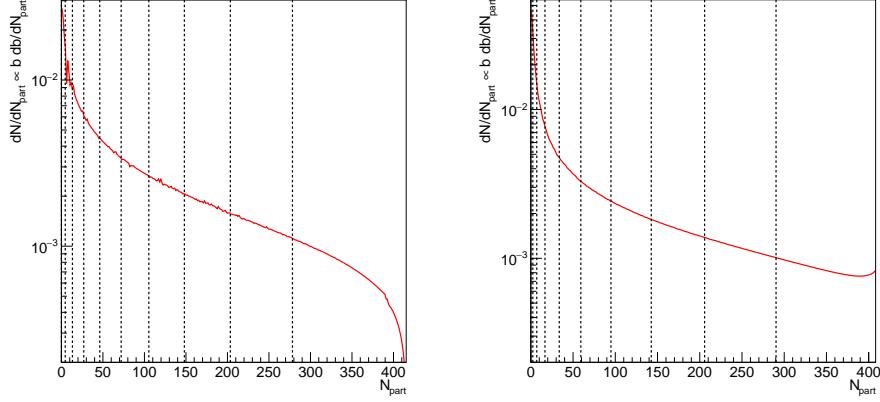


FIGURE 11.12: The distribution of hadronic PbPb collisions simulated in the custom Glauber model with hard spheres (*left*) or Woods-Saxons (*right*) nuclei, versus  $N_{\text{part}}$ . The vertical lines show the  $N_{\text{part}}$  limits of all 10%-wide centrality classes.

to prove consistent with the measurements, and it is at the core of the scaling demonstrated in section 11.2.9.

Determining the appropriate method for computing the path length has not been a straight path. It is unambiguous in the case of hard spheres and of constant QGP density in the overlap region. However, it requires a careful definition when the QGP density is realistic (meaning in the hard spheres Glauber model), and even more so in the Woods-Saxons case for which a somewhat arbitrary regularisation parameter is needed. However, the final result will prove relatively stable with the various (non-)Glauber models, at the condition that the path length and transverse area (section 11.2.8) are computed consistently in the same framework. I start with spelling out the average path length formula in section 11.2.7.1. Then section 11.2.7.2 computes analytically the path length in the hard-spheres case with constant QGP density. Last, sections 11.2.7.3 and 11.2.7.4 compute it numerically from a path length formula that is justified, in the hard-spheres and Woods-Saxons case.

### 11.2.7.1 Average path length

It is reminded here that  $\vec{b}$  joins the centres of the two colliding nuclei, and defines the  $x$ -axis. Points in the transverse plane are parametrised with  $\vec{s}(s, \theta) = s e^{i\theta}$ , where  $\theta$  is the angle between the  $\vec{b}$  and  $\vec{s}$  directions. I also define  $\phi$  as the angle between  $\vec{b}$  and the trajectory of a particle in the transverse plane. This trajectory is assumed straight, according to the approximation of small energy loss  $\varepsilon \ll E$ .

From the expression of the path length of a single parton through the QGP, the average over all produced partons at given impact parameter is needed: this is the quantity used for comparisons with measurements. To obtain it, partons are shot in the transverse plane, with a probability  $\rho_{\text{coll}}(b, \vec{s})$ , the nuclear overlap in the longitudinal direction from Eq. 11.20. I consider uniformly distributed  $\phi$  angles for the trajectories of particles in the transverse plane (otherwise, that would amount to  $v_2 \neq 0$ , see section 12.5). The average path length of the particle in the medium is then [35]:

$$\langle L \rangle = \frac{\int L(\vec{s}, \phi) \rho_{\text{coll}}(\vec{b}, \vec{s}) d\vec{s} d\phi}{2\pi \int \rho_{\text{coll}}(\vec{b}, \vec{s}) d\vec{s}} \quad (11.25)$$

where  $\vec{s}$  covers the whole transverse plane. This assumes a  $N_{\text{coll}}$ -scaling for the produced yields of hard partons. This is integrated numerically once  $L(\vec{s}, \phi)$  for a given parton path is known. It is performed by drawing partons in a conservative region encompassing the interaction region,

uniformly in the  $\phi$  and  $\vec{s}$  variables; then the single path length values are summed with a weight  $\rho_{\text{coll}}$  to obtain the numerator.

### 11.2.7.2 Constant QGP density in hard spheres (analytic calculation)

Now I calculate, for a particle with given  $(s, \theta, \phi)$ , the length of the path in the medium, assumed to have a constant density in the whole hard-spheres interaction region.  $\theta$  and  $\phi$  are defined w.r.t. the  $\vec{b}$  direction, in the interval  $[-\pi, \pi]$ . So the problem reduces to finding, from a given point and direction in the transverse intersection of the two nuclei, the distance to the edge of this almond region. I present here an analytical solution, based on standard trigonometry, but for which care is needed in the definition of limit cases. A personal sketch of the geometry will help the reader understanding the following expressions.

Two cases must be considered: if the touched edge is the limit of nucleus A (case A) or of nucleus B (case B). At fixed  $\vec{s}$ , the discrimination between these cases can be done via min and max values of  $\phi$ . If the particle is produced closer to the centre of A than of B (*i.e.*  $s \cos \theta < \frac{b}{2}$ ), then the condition for case A is:

$$-\frac{\pi}{2} < -\arctan \left( \frac{\sqrt{R^2 - (\frac{b}{2})^2} + s \sin \theta}{\frac{b}{2} - s \cos \theta} \right) < \phi < \arctan \left( \frac{\sqrt{R^2 - (\frac{b}{2})^2} - s \sin \theta}{\frac{b}{2} - s \cos \theta} \right) < \frac{\pi}{2}.$$

When the particle is produced in the other half of the almond ( $s \cos \theta > \frac{b}{2}$ ), the condition for case A is rather:

$$-\pi < -\frac{\pi}{2} - \arctan \left( \frac{s \cos \theta - \frac{b}{2}}{\sqrt{R^2 - (\frac{b}{2})^2} + s \sin \theta} \right) < \phi < \frac{\pi}{2} + \arctan \left( \frac{s \cos \theta - \frac{b}{2}}{\sqrt{R^2 - (\frac{b}{2})^2} - s \sin \theta} \right) < \pi.$$

And if  $s \cos \theta = \frac{b}{2}$ , the min and max values of  $\phi$  are  $\pm \frac{\pi}{2}$ .

Let me now calculate the path length  $l$  in case A. I start from the vector equality  $\vec{R} = \vec{s} + \vec{l}$  (with  $\vec{l} = (l, \phi)$  expressed in polar coordinates of origin the particle starting point), which implies  $R^2 = s^2 + l^2 + 2sl \cos(\theta - \phi)$  where  $\theta - \phi$  is the angle between  $\vec{l}$  and  $\vec{s}$ . This gives a second order equation in  $l$ :

$$l^2 + l(2s \cos(\theta - \phi)) + s^2 - R^2 = 0 \quad (11.26)$$

of discriminant

$$\Delta = 4s^2 \cos(\theta - \phi)^2 + 4(R^2 - s^2) = 4(R^2 - s^2 \sin(\theta - \phi)^2) > 0. \quad (11.27)$$

The only positive solution is then

$$l_A = -s \cos(\theta - \phi) + \sqrt{R^2 - s^2 \sin(\theta - \phi)^2} \quad (11.28)$$

The solution for the case B (*i.e.*  $\vec{l}$  rather touches the side of nucleus B), is found from the solution of case A. By symmetry, I just need to replace  $s \rightarrow |\vec{s} - \vec{b}|$  and  $\cos(\theta - \phi) \rightarrow \cos(\theta_B - \phi)$ , where  $\theta_B - \phi$  is the angle between  $\vec{l}$  and  $\vec{s} - \vec{b}$ . Care is needed for the calculation of  $\theta_B$ , which must be counted from the  $\vec{b}$  reference direction – and this slightly breaks the symmetry of the problem. I hence have:

$$l_B = -|\vec{b} - \vec{s}| \cos(\theta_B - \phi) + \sqrt{R^2 - |\vec{b} - \vec{s}|^2 \sin(\theta_B - \phi)^2} \quad (11.29)$$

where

$$\begin{aligned} |\vec{b} - \vec{s}|^2 &= b^2 + s^2 - 2bs \cos \theta \\ \text{if } b \geq s \cos \theta: \quad \theta_B &= \pi - \arcsin\left(\frac{s \sin \theta}{|\vec{b} - \vec{s}|}\right) \\ \text{if } b < s \cos \theta: \quad \theta_B &= \arcsin\left(\frac{s \sin \theta}{|\vec{b} - \vec{s}|}\right). \end{aligned}$$

The limiting case for the definition of  $\theta_B$  is when the point  $(s, \theta)$  is right (towards positive  $x$ ) of the centre of nucleus B. The definition by parts is needed because  $\arcsin(\cdot) \in [-\pi/2, \pi/2]$ .

With the Monte Carlo that I will use in the non-constant QGP density calculation, I checked that the path length expression calculated above is correct.

### 11.2.7.3 Effective path length for non-constant QGP density

The path length calculated in the previous section assumed that the QGP affects the partons in the same way in the whole transverse intersection region (*i.e.* that the QGP transverse density is constant). I will now discuss a definition of the path length that takes into account the inhomogeneity of the medium that the parton passes through, using  $\rho_{\text{QGP}} = \rho_{\text{part}}$  of Eq. 11.19 as a proxy for the QGP density.

As the path length is now not only geometric, and that the borders of the interaction region are not even definite in the Woods-Saxons case, one needs to define an effective path length that takes into account the density of the traversed medium. Let me start with an *intuitive but wrong* definition, that integrates over the trajectory while weighting each point with the density:

$$L_{\text{eff,naive}}(\vec{b}, \vec{s}, \phi) = \frac{\int_0^\infty \rho_{\text{QGP}}(\vec{s} + \vec{l}) dl}{\rho_{\text{QGP,ref}}}. \quad (11.30)$$

Unfortunately, this calls for an arbitrary decision for the reference density in the denominator. Five possibilities were thought of, that depend or not on  $\vec{s}$  and  $\vec{b}$ :

- The starting point of the produced particle  $\rho_{\text{QGP}}(\vec{s}, l = 0)$ . However, this is problematic when the parton crosses regions with higher density than at its start: the reference density in the denominator can get very small compared to the integral over the whole path in the numerator.
- The maximal density found on the path. It avoids the above problem, but depends on  $s$ , so that there would no absolute reference for the effect of QGP on partons.
- The maximum density of the interaction region (at  $x = b/2$  and  $y = 0$ ). This would give much smaller lengths, but is a true reference density, working also in the Woods-Saxons case.
- $N_{\text{part}}/A_T$ , the average density in the lentil. The area is however not well-defined in the Woods-Saxons case.
- A true reference density for all impact parameters.

All above solutions seem intuitive, but are intrinsically arbitrary. In addition, it is not consistent with the energy loss formalism leading to  $\varepsilon \propto L^2$ , as shown below. This is why the following **effective path length** is adopted in the literature [35, 232] and will be used here:

$$L_{\text{eff}}(\vec{b}, \vec{s}, \phi) = \frac{2 \int_0^\infty \rho_{\text{QGP}}(\vec{s} + \vec{l}) l dl}{\int_0^\infty \rho_{\text{QGP}}(\vec{s} + \vec{l}) dl} \quad (11.31)$$

from which the geometric length is recovered when inputting hard spheres nuclei with constant density.

The (best-behaved) path length definitions based on Eq. 11.30 were implemented and tested against measurements, but the consistency of the results with the nominal path length calculation was not fully investigated. It is not used in the literature because it is inconsistent with the BDMPS path dependence of the energy loss  $\varepsilon \propto L^2$ , with which Eq. 11.31 is consistent as shown below. However, future work could be done to adapt this effective length to a  $\varepsilon \propto L$  dependence due to the close-to-longitudinal medium expansion set forth at the end of this chapter.

Let us start from

$$\varepsilon = \frac{1}{2} \langle \hat{q} \rangle L_{\text{eff}}^2 = \int_0^\infty \frac{dE}{dz} dz \quad (11.32)$$

where the left term is the BDMPS estimation at constant QGP density (proportional to the average transport coefficient  $\langle \hat{q} \rangle$ ), and the right term can take into account a variable density. Let us find the effective length needed for Eq. 11.32 to hold, that is fully justified only in the case of hard spheres with constant density, but later generalised to the other cases. No expansion of the medium is assumed here, so only the space dependence of the QGP density is averaged.

The BDMPS energy loss on a small trajectory is taken as

$$dE = z \hat{q}(z) dz \quad (11.33)$$

consistent with Eq. 11.32 for constant densities. Then, the average density on a path  $[0, z]$  is

$$\langle \hat{q} \rangle(z) = \frac{\int_0^z \hat{q}(l) dl}{z} \quad (11.34)$$

which translates into

$$\langle \hat{q} \rangle(L_{\text{eff}}) = \frac{\int_0^\infty \hat{q}(l) dl}{L_{\text{eff}}} \quad (11.35)$$

where the integral boundary has been generalised from  $L_{\text{eff}}$  to infinity, on the basis of a null density beyond the limits of hard spheres nuclei. As a consequence:

$$\varepsilon = \frac{1}{2} \int_0^\infty \hat{q}(l) dl \times L_{\text{eff}} = \int_0^\infty \frac{dE}{dz} dl \quad (11.36)$$

Inputting  $dE/dz$  from Eq. 11.33, one indeed finds back the effective length of Eq. 11.31. In section 12.1.1, I attempt a justification of this without constant density arguments, and how this might affect the definition of the effective path length.

The path length of Eq. 11.31 with  $\rho_{\text{QGP}} = \rho_{\text{part}}$  is hence used for the two Glauber models. It is computed numerically, making the particle travel from its production point to the outside of the medium in small steps. At each step  $\rho_{\text{part}}$  times the distance from the production point is summed to give the numerator, and similarly for the denominator.

#### 11.2.7.4 Minimal QGP density for Woods-Saxons

In the Woods-Saxons optical Glauber model, the length from Eq. 11.31 grows un-physically in peripheral collisions, because towards small densities, the denominator decreases almost as much as the numerator does. This is due to the optical Glauber model having no hard phase space limits: nucleon interactions can possibly take place at very large distances from the centre of the interaction region. To be able to extract a length and a transverse area for this Glauber model, I need to regularise the formula with a minimal-density cutoff, to prevent the integral in the numerator to give non-null weight to very large distance. A somewhat arbitrary choice is

needed for the regularisation parameter, but I have checked that it does not affect significantly the results.

What could be the minimal transverse surface density of nucleons for a QGP to be created, or simply for any hadronic activity to occur? The area in Ref. [232] is taken as  $A = \pi r_0^2 (\frac{1}{2} N_{\text{part}})^{2/3}$  (considering the overlap area as a disk). Assuming that the QGP cannot exist for nucleonic densities lower than the  $\langle \rho_{\text{part}} \rangle$  for  $N_{\text{part}} = 1$  (corresponding to half the activity of typical pp collisions), I get

$$\rho_{\text{part,min}} = \frac{N_{\text{part}} = 1}{\pi r_0^2 ((N_{\text{part}} = 1)/2)^{2/3}} \simeq 0.35 \text{ fm}^{-2} \quad (11.37)$$

with  $r_0 \simeq 1.2 \text{ fm}$ . In practice, a lower value  $\rho_{\text{part,min}} = 0.1 \text{ fm}^{-2}$  is taken, to limit the influence of this parameter on the path length values. I tested that taking rather  $\rho_{\text{part,min}} = 0.01 \text{ fm}^{-2}$  has few effect on the results, and only in the most peripheral collisions.

#### 11.2.7.5 Comparison of methods

I compare in Fig. 11.13 the average path length in the four considered models (listed in ), including that tabulated for the MC Glauber model in Ref. [35]. The uncertainties quoted in the latter are large, about 50% at all centralities, but will not be considered in the fit results of section 11.2.9. I present it as a function of centrality, because the discrepancies of path length functions of  $b$  partially compensate thanks to the different  $b$  limits of the centrality classes associated to the various models.

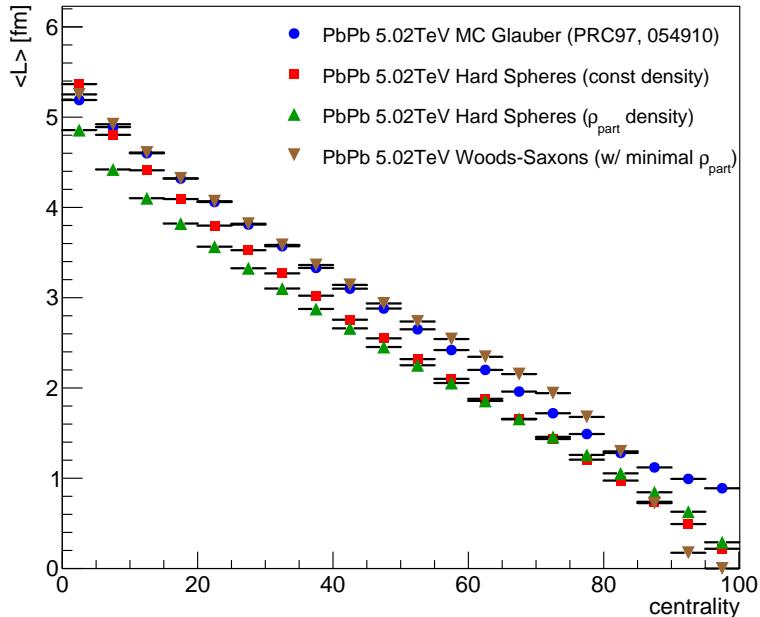


FIGURE 11.13: Comparison of the average path length in the medium as a function of centrality for the four Glauber models.

Despite all the differences of the models, they result in similar path length estimations. In particular, the ratio of the lengths for different models is close to constant, up to centralities of 80%. For most centralities the MC Glauber and optical Woods-Saxons models are very close, whereas the two hard spheres model are particularly close at higher centralities, where the QGP density comes closer to constant.

For the most peripheral collisions, the models start to diverge, as the path length definition is slightly ambiguous where few hadronic interactions occur. The MC Glauber model [35] is the only model showing there non-null lengths, of order 1 fm. The sharp drop in the optical

Woods-Saxons model is due to the minimal-density regularisation (section 11.2.7.4); smaller values of this threshold (such as  $0.01 \text{ fm}^{-2}$ ) push this drop closer to 100% centrality, but with a significantly higher length than the other models in the 70-90% range. However, these most peripheral collisions have only a small impact on the results of the scaling.

I show in Fig. 11.14 the comparison of the average path length for the four collision systems present in the fitted measurements, in the case of the full Glauber model in hard spheres. As expected, smaller nuclei show smaller path lengths. However, especially in the MC Glauber and Woods-Saxons cases, the path length AuAu collisions at 200 GeV starts to decrease below that of smaller nuclei in peripheral collisions, because the smaller  $\sigma_{\text{NN}}$  at this smaller energy starts to affect the probability for hadronic interactions of nucleons.

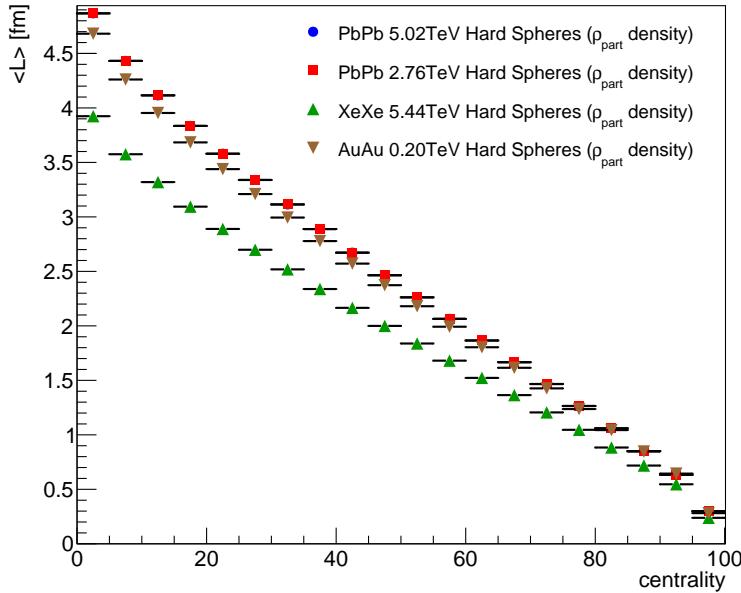


FIGURE 11.14: Comparison of the average path length in the medium as a function of centrality for the four studied collision systems, calculated with the full Glauber model in hard spheres.

## 11.2.8 Transverse area

### 11.2.8.1 Hard spheres and constant density

Let us start by calculating, for hard sphere nuclei of constant density, the transverse area of the nuclei wherein the hadronic activity occurs. This corresponds to the textbook almond/lentil shape of centre ( $x = b/2, y = 0$ ), width  $2R - b$ , and height  $2\sqrt{R^2 - (\frac{b}{2})^2}$ , e.g. shown in Fig. 1.6. I calculate it here analytically, in the hard-spheres view where it is non-ambiguously defined as the area of a nucleus that is ‘touched’ by the other nucleus. I will then compare these to the values of the transverse area versus centrality given in the MC Glauber model of Ref. [35].

The transverse area of the interaction region in the collision of two hard-sphere nuclei of radius  $R$ , with a separation  $b$  between the nuclei centres, is:

$$A_{\perp}(b, R) = 2R^2 \arccos\left(\frac{b}{2R}\right) - b\sqrt{R^2 - \left(\frac{b}{2}\right)^2}. \quad (11.38)$$

The first term is twice the area of the sector of the first nucleus joining the two tips of the almond, of angle  $2 \arccos(\frac{b}{2R})$ . The second term subtracts the unwanted area from this sector,

which, by playing with triangles of hypotenuse  $\vec{R}$ , is half of a rectangle of sides  $b$  and the height of the almond  $2\sqrt{R^2 - (\frac{b}{2})^2}$ .

The area in a class of centrality between impact parameters  $b_1, b_2$  is:

$$A_{\perp}(\mathcal{C}_{b_1, b_2}, R) = \frac{\int_{b_1}^{b_2} b A_{\perp}(b, R) db}{\int_{b_1}^{b_2} b db}. \quad (11.39)$$

The integral in the denominator is trivial, and I obtain the one in the numerator thanks to integral calculator (which was checked manually):

$$\int_{b_1}^{b_2} db b A_{\perp}(b) = \left[ R^4 \arcsin\left(\frac{b}{2R}\right) + b^2 R^2 \arccos\left(\frac{b}{2R}\right) - b R^3 \sqrt{1 - \left(\frac{b}{2R}\right)^2} \left(\frac{1}{2} + \left(\frac{b}{2R}\right)^2\right) \right]_{b_1}^{b_2} \quad (11.40)$$

### 11.2.8.2 Glauber models

The area in the MC Glauber [35] is that where any nucleon has a hadronic interaction. It is tabulated in 5%-wide centrality ranges. The quoted uncertainties are large, but smaller than for the path length: they run from 5-10% in central collisions to 100% in the most peripheral collisions.

For the hard spheres optical Glauber, it could be calculated as the geometrical intersect of section 11.2.8.1. However, for consistency with the path length determination, the area is rather defined based on the path length of Eq. 11.31 for non-constant QGP densities. It is the integral

$$\int_0^{2\pi} d\phi \int_{r=0}^{r=l(\phi)} r dr = \int_0^{2\pi} d\phi \frac{l^2(\phi)}{2} \quad (11.41)$$

where  $l(\phi)$  is calculated from the centre of the almond interaction region along the direction  $\phi$ . This amounts to taking the limits of the area of the medium to be one path length away from the centre of the intersection region.

In the Woods-Saxons case, the area is calculated in the same way, with the path length based on the Woods-Saxons densities. The same minimal  $\rho_{\text{part}} > 0.1 \text{ fm}^{-2}$  threshold is used as in the average length calculation.

### 11.2.8.3 Comparison of models

In Fig. 11.15, the transverse area as a function of centrality is compared for all (non-)Glauber models. It is systematically higher with the MC Glauber than in the custom Glauber models, probably due to the MC allowing for isolated interacting nucleons far from the nuclei centres. The Woods-Saxons model shows a flatter transverse area versus centrality in peripheral collisions than the other models, which would be worse with a looser minimal  $\rho_{\text{part}}$  threshold.

As I discussed the consistency of the calculations of the path length and of the area, let me discuss a simple variable that should mitigate the modelling uncertainties: the ratio of the area to the squared path length. It is shown in Fig. 11.16. It takes values that are comfortably between  $\pi$  and  $2\pi$ , and are relatively constant until a centrality of 80%.

### 11.2.9 Scaling results

I finally have all the ingredients to test the proposed scaling of Eq. 11.11 on all the light hadron measurements listed in Table 11.1. All the extracted energy loss values of light hadrons

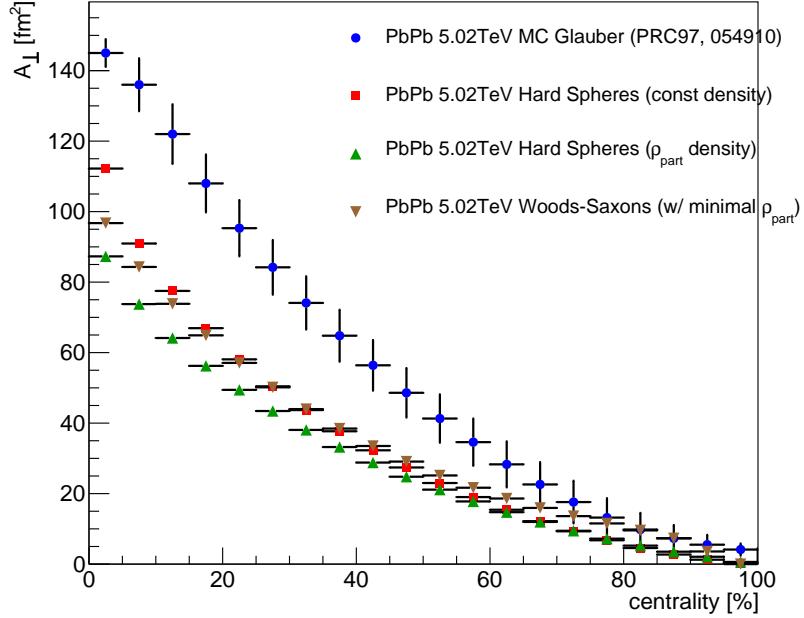


FIGURE 11.15: Transverse area  $A_{\perp}$  of the active medium versus centrality, compared for the four discussed Glauber models. The erratum of Ref. [35] is used, where the area is that where any participant nucleon interacts, without necessarily breaking a nucleon of the other nucleus.

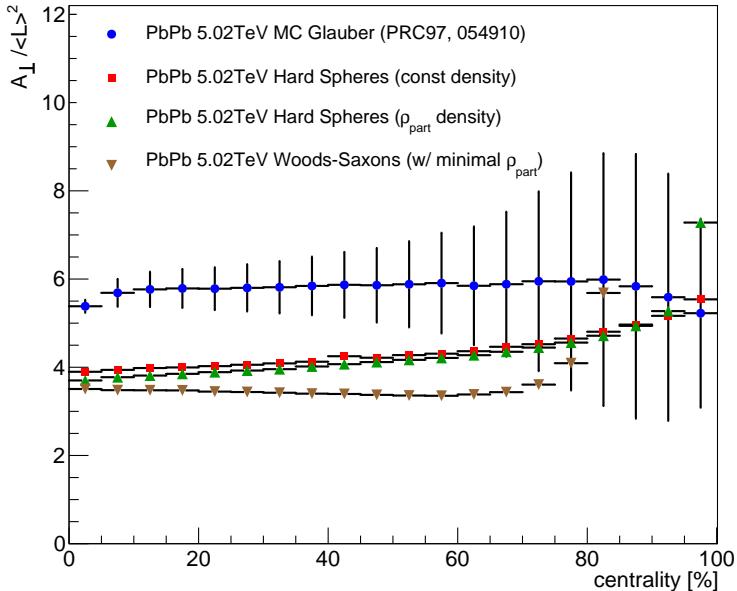


FIGURE 11.16: Transverse area  $A_{\perp}$  divided by the squared average path length, versus centrality, compared for the four Glauber models.

summarised in Fig. 11.4 are used, except the inclusive measurements when they are redundant with the centrality-dependent ones, and except the centralities above 60% measured with PHENIX (because the measured energy loss is almost null).

Each energy loss value is fitted from light hadron measurements in a given collision system, at a given energy, measured by a given experiment, in a given centrality range. In each of these specific conditions, the multiplicity at central rapidities, the transverse area, and the path length are calculated as mentioned in the above sections, consistently within each Glauber model. The hard-spheres Glauber model (with non-constant QGP density) will be used for the nominal results shown here, but the extracted parameters for all four models will be mentioned.

I then perform a fit to determine the value of  $\beta = 2 - \alpha$  and the prefactor  $K$  entering the scaling:

$$\langle \bar{\varepsilon} \rangle_{\text{fit}} = K \frac{dN_{\text{ch}}}{d\eta} \frac{\langle L \rangle^\beta}{A_\perp}. \quad (11.42)$$

So I draw the values of  $\langle \bar{\varepsilon} \rangle_{\text{fit}} A_\perp / (\frac{dN_{\text{ch}}}{d\eta})$  as a function of  $\langle L \rangle$ , and fit them with a power law function with positive exponent, constrained to have null energy loss for null medium length. This fit is shown in Fig. 11.17.

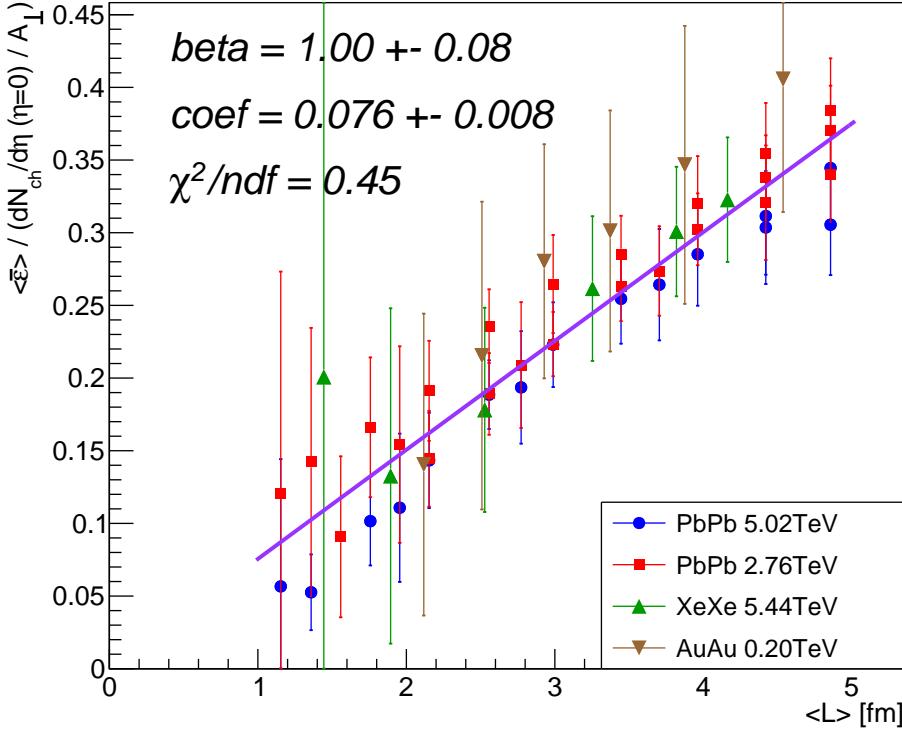


FIGURE 11.17: Mean energy loss divided by the multiplicity and the transverse area as a function of the average path length. The energy loss values are extracted from light hadron measurements (Table 11.1), and the other quantities are determined with a custom hard-spheres optical Glauber model. The power law fit is shown with a violet solid line, yielding the exponent  $\beta$ . See the text for details on the uncertainties.

Only the quadratic sum of the total uncertainties on the extracted energy loss (see section 11.1.2.3) and on the multiplicities are included in this fit, and are displayed in the error bars. In the MC Glauber fit, the uncertainties on the area and the path length are not included. In the custom Glauber models, no uncertainties on the path length and area are considered: I prefer to extract the final result using each Glauber model, and to quote a systematic uncertainty as the spread of these results.

The measurements already scale well in Fig. 11.17, but with large uncertainties, especially for points originating from peripheral collision systems. This scaling is better illustrated when the dependence of interest is plotted: the energy loss as a function of the activity of the QGP  $\frac{dN_{\text{ch}}}{d\eta} \frac{\langle L \rangle^\beta}{A_\perp}$ . The value of  $\beta$  is fixed from the fit of Fig. 11.17, and only the prefactor of Eq. 11.42 is fitted once more, and yields values very close to the previous fit. This new scaling law is shown in Fig. 11.18, and is valid over more than an order of magnitude of difference in the collision energy, and for nuclei of sizes up to two times smaller or larger. I hence show here that all existing measurements, in very different collision systems, energies, and geometric configurations, do scale with the ‘activity’ of the QGP as explained in this chapter.

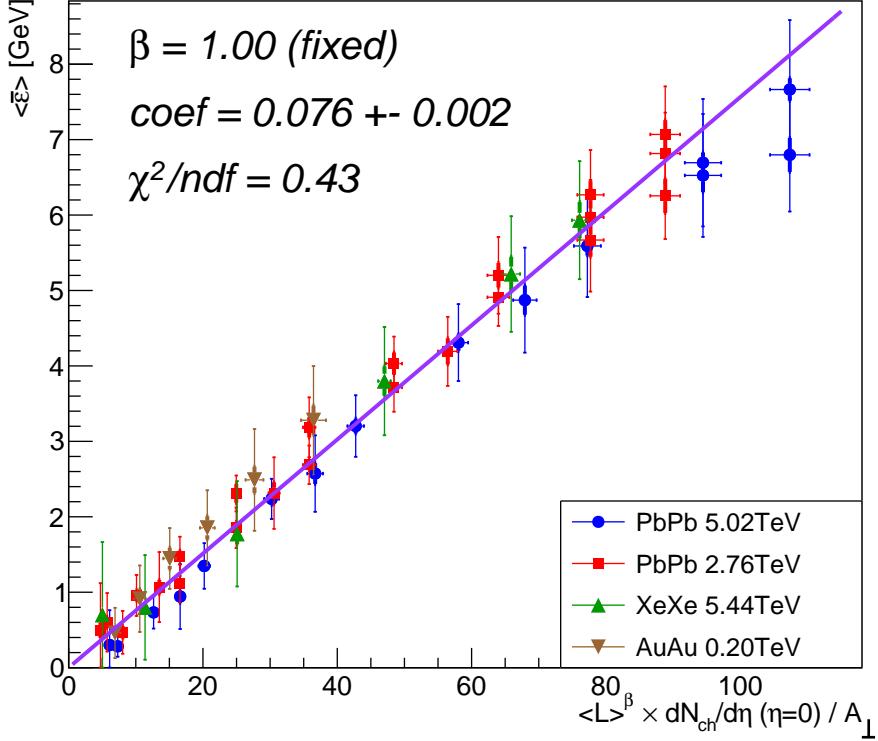


FIGURE 11.18: Mean energy loss as a function of  $\frac{dN_{ch}}{d\eta} \frac{\langle L \rangle^\beta}{A_\perp}$ . All light hadron measurements of Table 11.1 follow this scaling law, fitted in violet. The thin vertical lines show the total uncertainties on the fitted energy loss values, while the thick ones show the uncertainty originating only from the uncorrelated uncertainties of the measured datasets. The horizontal lines show only the uncertainties on the measured multiplicities.

The  $\chi^2$  of the fit is very good, and is similar for all four Glauber models. It is actually slightly underestimated due to the correlations of the energy loss uncertainties in groups of points that originate from the same Run and experiment. Typically, the luminosity or Glauber model uncertainties are correlated for all measurements of a data-taking period, and the systematic uncertainties linked to a given analysis method is very correlated along the bins of the same experimental analysis. I shall take into account these correlations in the future, but it should not affect significantly the results discussed here.

They are not shown here, but the energy loss values extracted for  $J/\psi$  and D mesons are also consistent with the found scaling, within large uncertainties, despite the differences that could emerge from different quark/gluon ratio or fragmentation function (mentioned in section 11.1.2.4).

Let me discuss now the values of the parameters extracted in the fit of Fig. 11.17, with the hard spheres Glauber model. The most striking conclusion is the value  $\beta = 1.00 \pm 0.08$ , (fortuitously) equalling the value  $\alpha = 2 - \beta = 1$  expected with a perfect **Bjorken expansion in the longitudinal direction** [126]. If the medium was expanding in the transverse direction as well, the density would decrease faster so  $\alpha$  would be larger than 1, and if the longitudinal expansion was slower than  $c$ , then  $\alpha < 1$  (*i.e.*  $\beta > 1$ ). With the observed  $\alpha = 1.00 \pm 0.08$ , the expansion could be a perfect longitudinal one, or a compensation of an imperfect longitudinal expansion with some transverse expansion.

The values of  $\beta$  obtained in hard spheres with constant QGP density (no Glauber model), the MC Glauber model, and the custom Woods-Saxons optical Glauber model, are  $1.10 \pm 0.07$ ,  $0.97 \pm 0.10$ , and  $1.03 \pm 0.09$ , respectively. The value for the simple hard spheres case deviates

by about  $1\sigma$  from the other models. However, it is not worrying considering the significantly different hypotheses made for the path length and area estimations in this case; it rather shows the robustness of the procedure. So I deduce a systematic uncertainty on  $\beta$  of 0.03, which yields a total uncertainty of 0.09. This **definitively excludes the view of a static block of QGP** created in heavy ion collisions.

The prefactor coefficient  $K$  equals  $0.073 \pm 0.001$ ,  $0.076 \pm 0.002$ ,  $0.127 \pm 0.003$ , and  $(0.073 \pm 0.001) \text{ GeV.fm}^{2-\beta}$ , in the simple hard-spheres, the optical hard spheres, the MC, and the Woods-Saxons optical Glauber models, respectively. The coefficient is significantly higher in the MC Glauber model mostly due to the transverse area (dividing the prefactor) being systematically larger than for the other models. From Eq. 11.11, the prefactor is:

$$K = \langle x \rangle_{\text{ref}} \langle z \rangle \frac{\langle \bar{\omega}_c \rangle_{\text{fit}}}{\langle z \rangle} = \langle x \rangle_{\text{ref}} \langle z \rangle \frac{27\pi\alpha_S^2}{4(2-\alpha)} \tau_0^{\alpha-1}. \quad (11.43)$$

In principle, this scaling could have constrained the formation time  $\tau_0$  of the QGP. However, as the fitted  $\alpha$  value is very close to 1, the factors  $\tau_0^{\alpha-1}$  and  $(2-\alpha)$  are considered to be 1 (and dimensionless) in this discussion. The term  $\langle x \rangle_{\text{ref}} = 0.289$  only cancels the factor used to convert the energy loss scale  $\omega_c$  to the energy loss  $\varepsilon$ . The average fragmentation fraction  $\langle z \rangle$  is estimated to be of order 0.5 (with 10-15% uncertainties) for light hadrons at the LHC [216].

Let us compare the numerical value of  $K$  versus expectations for the factors of Eq. 11.43. Taking the above-mentioned values of  $\langle x \rangle_{\text{ref}}$  and  $\langle z \rangle$ ,  $K = 0.076$  and  $\tau_0^{\alpha-1} = 1$  results in  $\alpha_S \simeq 0.16$  close to what could be expected. If the factor  $\langle x \rangle_{\text{ref}}$  is not fixed, one rather gets a constraint  $C_R \alpha_S^3 = 0.014$ , which is consistent with the values  $C_R = 3$  and  $\alpha_S = 0.17$  – the colour factor is nevertheless not well constrained, as the factor  $\alpha_S^3$  is very sensitive to variations of the assumed value of  $\alpha_S$ .

Values of the transport coefficient  $\hat{q}$  cannot be extracted from this result, as the length dependence is not  $L^2$ , as needed for a comparison to the BDMPS formalism. This is because some properties of the transport coefficient are absorbed in the multiplicity, length and area quantities.

To conclude, I have shown here that the **average energy loss of partons in any collision condition creating a QGP**, assuming that the mean fragmentation fraction into hadrons  $\langle z \rangle$  is similar to that of the light hadrons of the used measurements, can be determined as:

$$\langle \varepsilon \rangle = (0.26 \pm 0.03) \frac{\alpha_S C_R}{2} \frac{dN_{\text{ch}}}{d\eta} \frac{\langle L \rangle^{1.00 \pm 0.09}}{A_\perp} [\text{GeV}] \quad (11.44)$$

where  $C_R$  is the colour factor of the dominant parton parent of the considered hadron (and  $\langle L \rangle$  and  $A_\perp$  are measured in fm and fm<sup>2</sup>). The uncertainty on the prefactor is strongly correlated with that on the exponent of  $\langle L \rangle$ , and it is recommended to use the values of  $\langle L \rangle$  and  $A_\perp$  of a Glauber model close to the two custom ones presented here.



“Theory is when you know everything but nothing works. Practice is when everything works but no one knows why.”

---

Attributed to Albert Einstein or Hermann Hesse<sup>1</sup>

## Chapter 12

# Can we do better? Checks and improvements

This chapter presents checks of the formalism of the previous chapter, as well as work directions and ideas that are promising but have not been fully exploited yet.

## 12.1 Checks of energy loss and path length formalism

### 12.1.1 Alternative effective path length

Let us try here to justify the formula of the effective length of section 11.2.7.3 without using the argument of constant QGP density. I generalize the formula of the energy loss on a path  $[0, z]$  from Eq. 11.33 to a variable density as  $\varepsilon(z) = \langle \hat{q} \rangle(z) z^2 = z \int_0^z \hat{q}(l) dl$ , meaning I have:

$$\frac{d\varepsilon}{dz}(z) = z\hat{q}(z) + \int_0^z \hat{q}(l) dl \quad (12.1)$$

which gives an energy loss on the range  $[0, \infty]$  of

$$\varepsilon = \int_0^\infty z\hat{q}(z) dz + \int_0^\infty dz \int_0^z \hat{q}(l) dl \quad (12.2)$$

Equating this with the total energy loss of Eq. 11.35, I get

$$L_{\text{eff}} = \frac{\int_0^\infty z\hat{q}(z) dz + \int_0^\infty dz \int_0^z \hat{q}(l) dl}{\int_0^\infty \hat{q}(z) dz} \quad (12.3)$$

I do not use this path length expression, that has not been tested yet.

---

<sup>1</sup>The source is very unclear. The main internet trend is to add at the end of this quote “*Here, theory and practice are combined: nothing works and no one knows why!*”, and to still attribute it to Einstein. The version attributed to Hermann Hesse is “Theorie ist, wenn man alles weiss, aber nichts funktioniert. Praxis ist, wenn alles funktioniert, aber niemand weiss warum.”

### 12.1.2 Independence of various integrals

I revisit here with more explicit expressions the principle of the scaling (section 11.2.1), and assess the validity of the assumption of constant QGP density and the successive averaging of multiple involved variables.

#### 12.1.2.1 Effect of spatial inhomogeneities

In Ref. [223] the transport coefficient is:

$$\hat{q}(z, t) \equiv C n_0 f(z) g(t) = C \left( \frac{\frac{dN_{ch}}{dy} \Big|_{y=0}}{A_\perp \tau_0} \right) \left( \frac{\rho_{\text{part}}(z)}{N_{\text{part}}/A_\perp} \right) \left( \frac{\tau_0}{\tau_0 + t} \right)^\alpha \quad (12.4)$$

In the second term, the surface density of nucleons is normalised by the total  $N_{\text{part}}$  from the same model to cancel some modelling uncertainties.

I now impose  $z = t$  because of the medium expands at light-speed, and the energy loss on a given path is:

$$\varepsilon(b, X_0, \phi) = \int_0^{L(b, X_0, \phi)} \hat{q}(z, z) z dz \quad (12.5)$$

and integrating on all possible (isotropic)  $\phi$  directions of emission, and on the probability of production of a hard probe at  $X_0$ , the average energy loss for a given impact parameter  $b$  is:

$$\langle \varepsilon \rangle(b) = C \left( \frac{\frac{dN_{ch}}{dy} \Big|_{y=0}}{A_\perp \tau_0} \right) \frac{1}{N_{\text{coll}}(b)} \int dX_0 \rho_{\text{coll}}(X_0) \frac{1}{2\pi} \int_0^{2\pi} d\phi \int_0^{L(X_0, \phi)} \frac{\rho_{\text{part}}(X_0, \phi, l)}{N_{\text{part}}(b)/A_\perp(b)} \left( \frac{\tau_0}{\tau_0 + l} \right)^\alpha l dl \quad (12.6)$$

The goal of our scaling is to show that this can be approximated by

$$\langle \varepsilon \rangle(b) \propto C \frac{\tau_0^{\alpha-1}}{2-\alpha} \left( \frac{\frac{dN_{ch}}{dy} \Big|_{y=0}}{A_\perp \tau_0} \right) \langle L \rangle_{(b)}^{2-\alpha} \quad (12.7)$$

where we approximated  $\frac{l}{(\tau_0+l)^\alpha} \simeq l^{1-\alpha}$ , and where  $\langle L \rangle$  is an average length in the interaction region:

$$\langle L \rangle_{(b)} = \frac{1}{N_{\text{coll}}(b)} \int dX_0 \rho_{\text{coll}}(X_0) \frac{1}{2\pi} \int_0^{2\pi} d\phi L_{\text{geom}}(X_0, \phi) \quad (12.8)$$

or, taking into account an effective length for a non-constant QGP density:

$$\langle L \rangle_{(b)} = \frac{\int dX_0 \rho_{\text{coll}}(X_0) \int d\phi \int_0^\infty \rho_{\text{part}}(X_0, \phi, l) l dl}{2\pi \int dX_0 \rho_{\text{coll}}(X_0) \frac{1}{2} \int_0^\infty \rho_{\text{part}}(X_0, \phi, l) dl} \quad (12.9)$$

Note that, even neglecting the spatial dependence of  $\hat{q}$ , having a scaling of the integrated  $\langle \varepsilon \rangle$  on  $\langle L \rangle^{2-\alpha}$  is not exactly equivalent to having a dependence  $\varepsilon(z) \propto L^{2-\alpha}$  on a given path! This is because for the  $\rho_{\text{coll}}$  and  $\phi$  integrals,  $\int dx x^\alpha \neq (\int dx x)^\alpha$ . Luckily,  $\alpha$  being found close to 1 limits *a posteriori* this problem.

Now, let me assess how the spatial inhomogeneity of the density can affect the final energy loss values, and if this depends on parameter  $\alpha$ , on centrality ( $b$ ), and on the average length  $\langle L \rangle_{(b)}$ . This amounts to checking the difference between the hard spheres model with or without the assumption of constant density. In hard spheres, the effect of inhomogeneity  $F$ , for a given

path, on a given lentil ( $b$ ), is:

$$\int_0^{L(X_0,\phi)} \frac{\rho_{\text{part}}(X_0, \phi, l)}{N_{\text{part}}(b)/A_{\perp}(b)} \left(\frac{\tau_0}{\tau_0 + l}\right)^{\alpha} l \, dl \equiv F(X_0, \phi, b) \int_0^{L(X_0,\phi)} \left(\frac{\tau_0}{\tau_0 + l}\right)^{\alpha} l \, dl \quad (12.10)$$

Except this depends on the production point and direction of the particle. So I rather consider take the ratio of fully integrated energy losses, with or without the spatial inhomogeneity  $\rho_{\text{part}}$ :

$$F_{\text{full}}(\alpha, b) = \frac{1}{N_{\text{part}}(b)/A_{\perp}(b)} \frac{\int dX_0 \rho_{\text{coll}}(X_0) \int d\phi \int_0^{L(X_0,\phi)} \rho_{\text{part}}(X_0, \phi, l) \left(\frac{\tau_0}{\tau_0 + l}\right)^{\alpha} l \, dl}{\int dX_0 \rho_{\text{coll}}(X_0) \int d\phi \int_0^{L(X_0,\phi)} \left(\frac{\tau_0}{\tau_0 + l}\right)^{\alpha} l \, dl} \quad (12.11)$$

Let me test analytically some limits of  $F$  (before integration over the lentil), with a particle drawn at a given  $\phi$  from the centre of the QGP region, so that it travels distance  $L$  (in hard spheres). I assume the spatial dependence of  $\hat{q}$  to be  $f(z) = 1 - (\frac{z}{L})^2$ , qualitatively close to the numerical hard spheres density, and the time dependence as  $g(t) = (\frac{l_0}{l+l_0})^{\alpha}$ , where  $\alpha \neq 1$  and  $\neq 2$ . I assume that  $L \gg l_0$ ,  $L(\alpha - 1) \gg l_0$ , and  $L(\alpha - 2) \gg l_0$ . I then obtain

$$F \simeq \frac{1 - \frac{\alpha^2 - 7\alpha + 10}{\alpha - 4} \left(\frac{L}{l_0}\right)^{2-\alpha}}{1 - (\alpha - 1) \left(\frac{L}{l_0}\right)^{2-\alpha}} \simeq \frac{2}{4 - \alpha} \quad (12.12)$$

which gives values between  $2/3$  and  $1$ , for  $1 < \alpha < 2$ , and values between  $1/2$  and  $2/3$  for  $0 < \alpha < 1$ . This is satisfyingly close to a constant. Doing the same calculation with  $\alpha = 1$  gives

$$F = 1 - \frac{l_0^2}{L^2} - \frac{1}{3} \frac{L - \frac{3}{2}l_0}{L - l_0 \ln |1 + \frac{L}{l_0}|} \simeq 2/3 \quad (12.13)$$

which is comforting as well:  $0.5 < F < 1$  for  $0 < \alpha < 2$ .

Fig. 12.1 shows the integrated  $F_{\text{full}}$  in the simulated hard spheres case. For the  $\alpha = 1$  that is observed in the previous chapter, it varies only of  $2 - 3\%$  in the most central collisions, but is constant at larger  $b$  values. This is consistent with the relatively small difference found in the results of the scaling when using a constant or a  $\rho_{\text{part}}$  QGP density.

### 12.1.2.2 Order of averages, formation time

Still in hard spheres, I quantify here the difference between the expected scaling (section 11.2.1) and a full calculation of energy loss, that considers the energy loss of each individual parton separately and includes the effect of the formation time. This amount to the ratio

$$1 + \delta\langle\varepsilon\rangle = \frac{\int \rho_{\text{coll}} \int d\phi \int_0^L \frac{\rho_{\text{part}}(X_0, \phi, l)}{N_{\text{part}}(b)/A_{\perp}(b)} \left(\frac{\tau_0}{\tau_0 + l}\right)^{\alpha} l \, dl}{\frac{\tau_0^{\alpha}}{2-\alpha} \left(\int \rho_{\text{coll}} \int d\phi L\right)^{2-\alpha}} \quad (12.14)$$

except I checked above that the spatial dependence of the density does not have a strong effect, so I need only to look at

$$1 + \delta\langle\varepsilon\rangle = \frac{\int \rho_{\text{coll}} \int d\phi \int_0^L \frac{l}{(\tau_0 + l)^{\alpha}} \, dl}{\frac{1}{2-\alpha} \left(\int \rho_{\text{coll}} \int d\phi L\right)^{2-\alpha}} \quad (12.15)$$

This ratio, numerically calculated, is shown in Fig. 12.2 in the  $(\alpha, b)$  plane, for  $\tau_0 = 0.1\text{fm}$ . It ranges from 0.4 to 1.7 in the  $(b, \alpha)$  space, for  $0.3 < \alpha < 1.6$  and  $\tau_0 = 0.3\text{ fm}$  (and a density that starts being diluted only from  $\tau_0$ , meaning  $\frac{\tau_0}{l+\tau_0}$  becomes 1 for  $l < \tau_0$ ). Let us investigate what dominates in this ratio.

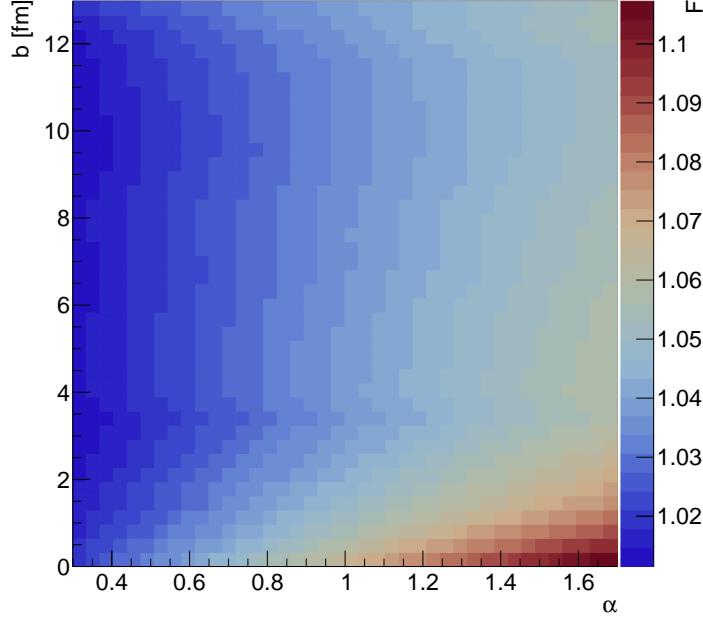


FIGURE 12.1: Influence (factor modifying the energy loss) of the inhomogeneity on the energy loss, as a function of the expansion parameter  $\alpha$  and the impact parameter.

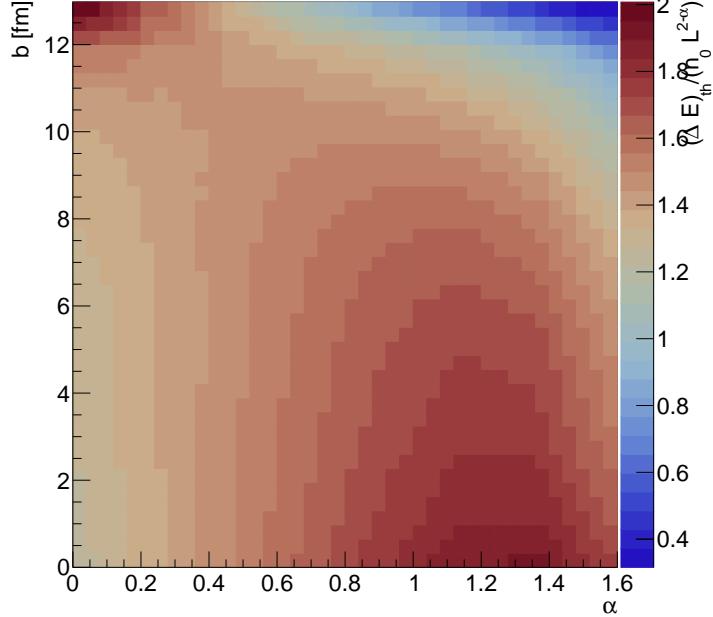


FIGURE 12.2: Ratio between the full energy loss calculation for individual parton trajectories and including the formation time (0.3 fm), and the energy loss as calculated from measurements, as a function of the expansion parameter  $\alpha$  and the impact parameter.

First, I study the position of the  $\rho_{\text{coll}}, \phi$  integrals compared to the power  $2 - \alpha$  is negligible when  $\alpha$  is close to 1. For values up to  $\alpha < 1.6$ , the ratio

$$1 + \delta\langle\varepsilon\rangle = \frac{\int \rho_{\text{coll}} \int d\phi L^{2-\alpha}}{\left(\int \rho_{\text{coll}} \int d\phi L\right)^{2-\alpha}} \quad (12.16)$$

deviates down by max 5-6%. However, for  $\alpha < 1$ , it deviates up in peripheral collisions, and

ranges from 1.15 to 1.5 for  $\alpha = 0.3$ . However, for  $\alpha = 0.85$ , the deviation for large  $b$  goes only up to 5-7%. This is hence not the main contribution to the large ratio of Fig. 12.2.

Now, I study the dependence on  $\tau_0$ . Even with a very small  $\tau_0 = 0.01$  fm, the ratio varies up to 60% in the  $(b, \alpha)$  plane, giving lower values for small  $\alpha$  or large  $b$ . This can be understood with the equivalence large  $b \leftrightarrow$  small  $L$  and, ignoring the position of the  $\rho_{\text{coll}}$  integral and the power:

$$\frac{\int_0^L \frac{l \, dl}{(l+\tau_0)^\alpha}}{\int_0^L l^{1-\alpha} \, dl} = \left(1 + \frac{\tau_0}{L}\right)^{1-\alpha} \left(1 + \frac{\tau_0}{L(\alpha-1)}\right) - \frac{1}{\alpha-1} \left(\frac{\tau_0}{L}\right)^{2-\alpha} \quad (12.17)$$

which we develop at 2nd order in  $\tau_0/L$  for  $\alpha < 2$ :

$$\frac{\int}{\int} = 1 + \frac{\tau_0}{L(1-\alpha)} (\alpha^2 - 2\alpha + (\tau_0/L)^{1-\alpha}) - \frac{1+\alpha}{2} \left(\frac{\tau_0}{L}\right)^2 \quad (12.18)$$

which is mostly lower than 1, due to  $\alpha^2 - 2\alpha$  for  $\alpha < 1$ , and to  $(\tau_0/L)^{1-\alpha}$  for  $\alpha > 1$ . In the later case, the main contribution is  $\propto (\frac{\tau_0}{L})^{2-\alpha}$ , which gives significant deviations for  $\alpha$  close to 2 and small  $L$  (ie large  $b$ ).

To conclude, it is not straightforward to use the Salgado-Wiedemann effective density (with medium expansion) as a scaling on the average length, especially in peripheral collisions. Even close to  $\alpha = 1$ , the ratio has variations of order 50% along  $b$ . This still needs to be investigated further.

### 12.1.3 Separating time and space dependences

In hard spheres, for a given parton produced at point and direction  $(x_0, y_0, \phi)$  and with a density  $\hat{q} = \hat{q}_0 g(z) f(t)$ , the energy loss is:

$$\varepsilon = \hat{q}_0 \int_0^L g(l) f(l) l \, dl \quad (12.19)$$

and to separate the temporal ( $f$ ) and spatial ( $g$ ) dependences in the integral I define  $G$  such that

$$\varepsilon = \hat{q}_0 G \int_0^L g(l) l \, dl \quad (12.20)$$

so

$$G = \frac{\int_0^L f(l) g(l) l \, dl}{\int_0^L g(l) l \, dl} \quad (12.21)$$

If  $G$  does not depend strongly on the shape of  $f(l) = (\frac{l_0}{l_0+l})^\alpha$ , then it justifies the separation between  $f$  and  $g$  integrals. The hope is also that it does not depend too much on  $L$ . If it has a weak but power dependence  $G \propto L^\gamma$ , then this can be interpreted as  $\varepsilon \propto L^{\gamma+(2-\alpha)}$ , where  $\alpha$  is only linked to the time expansion, and  $\gamma$  is mostly due to the spatial shape of the medium density. The shape of  $G$  differ for various  $\phi$  and centralities.

With the custom Glauber models of section 11.2, I check explicitly that the space and time dependences act on different scales, so that their integrals can effectively be performed independently. Fig. 12.3 shows the  $\rho_{\text{part}}$  and  $\rho_{\text{coll}}$  densities in hard spheres at centrality 40% along the  $x$  and  $y$  directions (passing through the centre of the interaction region), compared to the term  $\frac{\tau_0}{\tau_0+\tau} \sim \frac{l_0}{l_0+l}$ . The latter evolves indeed on much smaller length scales than the change of density. This conclusion is less strong in the most peripheral collisions, but still holds.

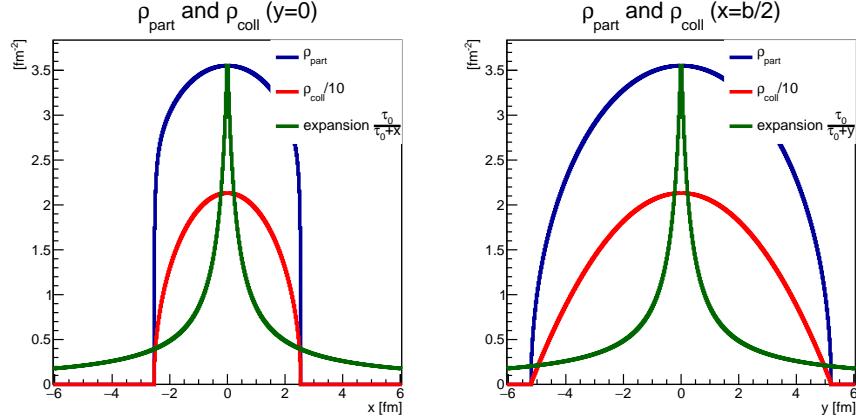


FIGURE 12.3: Hard spheres  $\rho_{\text{part}}$  and  $\rho_{\text{coll}}$  along the  $x$  and  $y$  directions, at centrality 40%. The spatial variation of  $\rho_{\text{part}}$  is slow compared to that of the medium expansion close to the centre of the medium, and inversely close to the edge of the medium.

## 12.2 $N_{\text{part}}$ and centrality in the Woods-Saxons model

I come back in this section on the difference of  $N_{\text{part}}$  (densities) between hard spheres and Woods-Saxons mentioned in section 11.2.6.2. I considered changing the definition of centrality for Woods-Saxons, so that the integrated  $N_{\text{part}}$  at a given centrality is similar in hard-spheres and in Woods-Saxons. This new centrality variable will still correspond to different values of  $b$  for the two models, but lower  $b$  (so higher  $N_{\text{part}}$ ) in peripheral collisions in Woods-Saxons than previously.

This fix could be justified by the fact that with the optical (continuous) approximation, contrarily to a Glauber MC where the probabilities of single discrete nucleon-nucleon collisions are calculated, the criterium for ‘a hadronic AA interaction took place’ (which defines the ‘centrality=100%’ reference) could be biased. The hadronic interaction probability might decrease unrealistically slowly with  $b$  because in reality, the nucleons at these low densities are scarce and might not meet. In a classical sense, there could be two nucleons present in a given area with reasonably high nucleon density, but they would be far away to meet head-on.

I could cut away the very high- $b$  collisions, defining a ‘100% centrality’ reference point as the  $b$  where the maximal  $\rho_{\text{part}}$  is lower than the reference density mentioned in section 11.2.7.3. Alternatively, I could simply redefine the  $b$  limits of the centrality classes as those corresponding to the  $N_{\text{part}}$  limits of the centrality classes of the hard spheres model. However, this was not tried yet, because the Woods-Saxons Glauber model I implemented gives very similar results than the other models despite the observed differences in the  $N_{\text{part}}$  limits, meaning that the modelling uncertainties partially cancel in the ratio of the length and the area.

## 12.3 OO predictions

From the scaling of section 11.2, more precisely from the result of Eq. 11.44, the radiative energy loss of high- $p_{\perp}$  hadrons in the OO collisions at  $\sqrt{s_{\text{NN}}} = 7 \text{ TeV}$  can be predicted, along with the resulting  $R_{\text{PbPb}}$ , in various centrality classes. These OO collisions will take place in the Run3 of LHC, probably in 2024-2025.

I only need the path length, the transverse area, and the charged multiplicities. The first two are already calculated exactly as in the other systems, possibly with any of the three custom models, but I keep the hard spheres Glauber model (with non-constant density) as the nominal.

The NN cross section is only slightly higher, and is calculated from the  $\sqrt{s_{\text{NN}}}$  dependence of Ref. [35].

The charged particle multiplicity is harder to obtain as it was not measured in this system at this energy. However, I can estimate it at first order with Fig. 12.4 from Ref. [229]. The ratio  $dN_{ch}/d\eta/(2A\langle N_{\text{part}} \rangle^2)$  is of order  $2.5 - 5$ , so I can take  $dN_{ch}/d\eta$  in OO collisions to be

$$dN_{ch}/d\eta(\text{OO}) = dN_{ch}/d\eta(\text{PbPb}) \times \frac{(N_{\text{part}}^2/A)(\text{OO})}{(N_{\text{part}}^2/A)(\text{PbPb})} \quad (12.22)$$

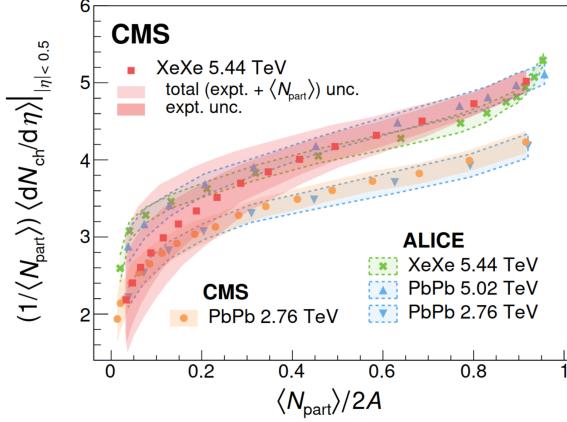


FIGURE 12.4: The charged particle multiplicity divided by  $N_{\text{part}}$ , as a function of  $N_{\text{part}}/2A$ , for PbPb and XeXe collisions. Figure from Ref. [229].

The ratio  $dN_{ch}/d\eta/\langle N_{\text{part}} \rangle$  slowly rises with the collision energy: Ref. [233] gives a dependence in  $\sqrt{s}^{n=0.11 \text{ or } 0.158}$  respectively in pp or (various) AA collisions. This gives a factor  $(7/5.44)^{0.11} = 1.028$  or  $(7/5.02)^{0.158} = 1.054$  to multiply the right-hand side of Eq. 12.22 for OO collisions at 7 TeV. In the future, the resulting values of  $N_{\text{part}}$  could be compared to predictions from the EPOS generator.

The resulting energy loss and  $R_{\text{AA}}$  will imminently be calculated.

## 12.4 Tests of scaling violation

I show here two tests, that will soon be performed, of the robustness of the scaling of section 10.2 versus the considered assumptions.

### 12.4.1 Logarithms of the energy

Here I challenge the assumption that the logarithmic dependence of the BDMPS energy loss on the energy of the parton does not affect significantly the scaling and the extracted energy loss values.

In the following modification factor:

$$R_{\text{AA}} = \int d\varepsilon P(\varepsilon) \left( \frac{p_\perp + \varepsilon}{p_\perp} \right)^{-n} \quad (12.23)$$

one can modify the quenching weight  $P$  into  $P'$  such that

$$P'(\varepsilon) = \frac{1}{r} P \left( \frac{\varepsilon}{r} \right) \quad (12.24)$$

which indeed has an average energy loss  $r$  times larger:  $\langle \varepsilon' \rangle = \int d\varepsilon \varepsilon P'(\varepsilon) = r \langle \varepsilon \rangle$ . Therefore, the modified  $R_{AA}$  is:

$$R_{AA} = \int \frac{d\varepsilon}{r} P\left(\frac{\varepsilon}{r}\right) \left(\frac{p_\perp + \varepsilon}{p_\perp}\right)^{-n} = \int d\varepsilon P(\varepsilon) \left(1 + \frac{r\varepsilon}{p_\perp}\right)^{-n} \quad (12.25)$$

Let me modify the mean energy loss to take into account the logarithmic dependence on the parton energy mentioned at the end of section 10.1. This implies, taking  $p_\perp$  for the parton energy:

$$r = 1 + \frac{1}{2} \ln \left( \frac{p_\perp}{\omega_c} \right). \quad (12.26)$$

where  $p_T \gg \omega_c$  is always assumed (the logarithm is forced to 0 if  $p_T < \omega_c$ ). The shape of  $P$  is assumed, at first order, to stay the same, and only the average of  $P$  to be shifted. Including the average fragmentation fraction  $\langle z \rangle$  ignored above, this reads

$$R_{AA} = \int d\varepsilon P(\varepsilon) \left( 1 + \varepsilon \langle z \rangle \frac{1 + \frac{1}{2} \ln \left( \frac{p_\perp}{\omega_c} \right)}{p_\perp} \right)^{-n} \quad (12.27)$$

In this context,  $\omega_c$  should be the energy scale at which the spectrum goes from the LPM to the fully coherent regime (see section 10.1). I take it here as  $\omega_c = \langle \varepsilon \rangle = \int d\varepsilon' \varepsilon' P(\varepsilon')$ .

One can then check if the  $R_{AA}$  breaks significantly the scaling observed in the  $p_\perp/n\omega_c$  variable, by drawing this  $R_{AA}$  for various realistic values of  $\omega_c$ ,  $n$ , and  $p_\perp$ . These scaling violations are expected to be much smaller than the sensitivity of the current measurements, but it needs to be checked quantitatively, which will be done soon.

### 12.4.2 Quarks vs gluons

The energy loss model starts from only one species of parton to produce the hadron of which we measure the  $R_{AA}$ . Let us consider that this parton can either be a gluon or a quark, and how this affects the  $R_{AA}$  model and the scaling in  $p_\perp/n\omega_c$ .

I introduce the fraction  $x_q(p_\perp) = \sigma_q^h/\sigma_{tot}^h$  of final state partons that are quarks (meaning that  $\sigma_g^h = (1 - x_q)\sigma_{tot}^h$ ). The fraction  $x_q$  ranges from 0.6 at 10 GeV to 0.4 at 100 GeV, *e.g.* from the NLO calculations of Ref. [216]. Then:

$$\sigma_{AA} = A^2 \int d\varepsilon (P_q(\varepsilon)x_q(p_\perp) + P_g(\varepsilon)(1 - x_q(p_\perp))) \sigma_{pp}^h(p_\perp + \varepsilon) \quad (12.28)$$

The mean energy loss is also proportional to the colour charge, so that  $\langle \varepsilon_q \rangle / \langle \varepsilon_g \rangle = C_F/C_A = \frac{4}{9} = r$ . Therefore:

$$\langle \varepsilon_q \rangle = \int P_q(\varepsilon) \varepsilon d\varepsilon = C_F = rC_A = r\langle \varepsilon_g \rangle = r \int P_g(\varepsilon) \varepsilon d\varepsilon = \int \frac{1}{r} P_g\left(\frac{\varepsilon'}{r}\right) \varepsilon' d\varepsilon' \quad (12.29)$$

where the last equality comes from a change of variable  $\varepsilon' = r\varepsilon$ . Assuming that the functional form of  $P_q$  and  $P_g$  are the same, I get  $P_q(\varepsilon) = \frac{1}{r} P_g\left(\frac{\varepsilon}{r}\right)$ . Then:

$$\sigma_{AA} = A^2 \int d\varepsilon \left( \frac{1}{r} P_g\left(\frac{\varepsilon}{r}\right) x_q(p_\perp) + P_g(\varepsilon)(1 - x_q(p_\perp)) \right) \sigma_{pp}^h(p_\perp + \varepsilon) \quad (12.30)$$

Now, defining  $P_g(\varepsilon) = \frac{1}{\langle \varepsilon \rangle} \bar{P}_g(\frac{\varepsilon}{\langle \varepsilon \rangle})$  and changing the variable  $x = \frac{\varepsilon}{\langle \varepsilon \rangle}$  or  $x = \frac{\varepsilon}{r\langle \varepsilon \rangle}$  in the first or second terms:

$$\sigma_{AA} = A^2 \int dx \bar{P}_g(x) x_q(p_\perp) \sigma_{pp}^h(p_\perp + xr\langle \varepsilon \rangle) + A^2 \int dx \bar{P}_g(x) (1 - x_q(p_\perp)) \sigma_{pp}^h(p_\perp + x\langle \varepsilon \rangle) \quad (12.31)$$

I then divide by the pp cross section:

$$R_{AA} = \int dx \bar{P}_g(x) \left[ x_q(p_\perp) \frac{1}{\left(1 + \frac{xr\langle \varepsilon \rangle}{p_\perp}\right)^n} + (1 - x_q(p_\perp)) \frac{1}{\left(1 + \frac{x\langle \varepsilon \rangle}{p_\perp}\right)^n} \right] \quad (12.32)$$

and after exponentiation:

$$R_{AA} = \int dx \bar{P}_g(x) \left[ x_q(p_\perp) \exp\left(-x \frac{nr\langle \varepsilon \rangle}{p_\perp}\right) + (1 - x_q(p_\perp)) \exp\left(-x \frac{n\langle \varepsilon \rangle}{p_\perp}\right) \right] \quad (12.33)$$

which varies now not only with  $\frac{n\langle \varepsilon \rangle}{p_\perp}$ , due to  $x_q(p_\perp)$  not being constant. I project to compare soon the resulting shapes in the typical  $p_\perp$  range of the measurements, for conservative assumptions on the  $p_\perp$  dependence of  $x_q$  and different values of  $n$ , and quantify the scaling violation. I will also refit the measurements with this improved model.

## 12.5 $v_2$

The Glauber models of the previous chapter allow compute a detailed  $\phi$  dependence of the high- $p_\perp$  energy loss, hence of the  $R_{AA}$  of light hadrons. The  $v_2$  coefficient is then straightforward to extract from  $R_{AA}(\phi)$ , at a given  $p_\perp$ , or integrated over  $p_\perp$ . I could *e.g.* use the simple formula for the  $v_2$  Fourier coefficient in Ref. [234]. The  $v_2$  is the second coefficient of the Fourier series of the distribution of particles in AA collisions as a function of  $\phi$ . However, this is equivalent to the  $R_{AA}(\phi)$ , as the distribution of particles can be considered isotropic in  $\phi$  in pp collisions, giving a constant denominator for the  $R_{AA}$ . This  $v_2$  calculation could motivate future measurements of  $v_2$  at very high- $p_\perp$ .



# List of Abbreviations

<b>AA</b>	nucleus-nucleus
<b>AD</b>	Antiproton Decelerator
<b>ALICE</b>	A Large Ion Collider Experiment
<b>ATLAS</b>	A Toroidal LHC ApparatuS
<b>BDT</b>	Boosted Decision Tree
<b>CERN</b>	Conseil Européen pour la Recherche Nucléaire (Eur. Council for Nuclear Research)
<b>CSC</b>	Cathode Strip Chambers
<b>CMS</b>	Compact Muon Solenoid
<b>c.o.m.</b>	centre-of-mass
<b>DAQ</b>	Data AcQuisition system
<b>DCA</b>	Distance of Closest Approach
<b>DT</b>	Drift Tubes
<b>ECAL</b>	Electromagnetic Calorimeter
<b>Eq.</b>	Equation
<b>Fig.</b>	Figure
<b>HCAL</b>	Hadronic Calorimeter
<b>HLT</b>	High Level Trigger
<b>ID</b>	particle IDentification
<b>IP</b>	impact parameter
<b>L1</b>	first Level of trigger
<b>LEP</b>	Large Electron–Positron Collider
<b>LHC</b>	Large Hadron Collider
<b>LS</b>	Long Shutdown
<b>MB</b>	Minimum Bias
<b>MC</b>	Monte Carlo (simulation)
<b>MIP</b>	Minimum-Ionising Particle
<b>NLL</b>	Negative Log-Likelihood
<b>NN</b>	Nucleon-Nucleon
<b>nPDF</b>	nuclear Parton Distribution Functions
<b>NRQCD</b>	Non-Relativistic QCD
<b>OO</b>	Oxygen–Oxygen
<b>OS</b>	Opposite-Sign
<b>pp</b>	proton–proton
<b>PbPb</b>	lead–lead
<b>PHENIX</b>	Pioneering High Energy Nuclear Interaction eXperiment
<b>POI</b>	Parameter Of Interest
<b>pNRQCD</b>	perturbative-NRQCD
<b>PV</b>	Primary Vertex
<b>QED</b>	Quantum ElectroDynamics
<b>QCD</b>	Quantum ChromoDynamics
<b>QGP</b>	Quark-Gluon Plasma
<b>Ref.</b>	Reference
<b>RF</b>	RadioFrequency

<b>RHIC</b>	Relativistic Heavy-Ion Collider
<b>RMS</b>	Root-Mean-Square
<b>RPC</b>	Resistive Plate Chambers
<b>SF</b>	Scale Factor
<b>SM</b>	Standard Model
<b>SPS</b>	Super Proton Synchrotron
<b>TnP</b>	Tag-and-Probe
<b>XeXe</b>	xenon–xenon

# Values of Physical Constants

This table shows numerical values of some physical constants, particle properties, and unit conversions that are used in this document. The uncertainties are not displayed when lower than  $\simeq 10^{-4} - 10^{-5}$  relative to the central value.

<b>Physical constants</b>	<b>Symbol = Value (SI units)</b>	<b>Value (natural units)</b>
Speed of light	$c = 2.997\,924\,58 \times 10^8 \text{ m s}^{-1}$	$c = 1$ , exact
Reduced Planck constant	$\hbar = h/2\pi = 1.054\,571\,8 \times 10^{-34} \text{ Js}$	$\hbar = 1$
Gravitational constant	$G = 6.674\,08 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$	
Elementary charge	$e = 1.602\,18 \times 10^{-19} \text{ C} \text{ (or J V}^{-1}\text{)}$	$(1 \text{ eV})/(1 \text{ V})$
Vacuum permittivity	$\varepsilon_0 = 8.854\,187\,81 \times 10^{-12} \text{ F m}^{-1}$	
Fine structure constant	$\alpha_{\text{QED}} = \frac{1}{4\pi\varepsilon_0\hbar c} = 7.297\,352\,569 \times 10^{-3}$	
Bohr radius	$a_0 = \frac{\hbar}{m_e c a} = 5.291\,772\,109 \times 10^{-11} \text{ m}$	
<b>Particle properties</b>		
Neutron mass	$m_n = 1.6749 \times 10^{-27} \text{ kg}$	$939.56 \text{ MeV}/c^2$
Proton mass	$m_p = 1.6726 \times 10^{-27} \text{ kg}$	$938.27 \text{ MeV}/c^2$
Pion mass	$m_\pi = 2.4881 \times 10^{-28} \text{ kg}$	$139.57 \text{ MeV}/c^2$
Electron mass	$m_e = 9.1093 \times 10^{-31} \text{ kg}$	$0.51100 \text{ MeV}/c^2$
Muon mass	$m_\mu = 1.8835 \times 10^{-28} \text{ kg}$	$105.66 \text{ MeV}/c^2$
$B_c$ mass	$m_{B_c} =$	$(6274.9 \pm 0.8) \text{ MeV}/c^2$
$B_c$ lifetime	$c\tau_{B_c} = (153 \pm 3) \mu\text{m}$	
$J/\psi$ mass	$m_{J/\psi} =$	$3096.9 \text{ MeV}/c^2$
<b>Conversions</b>		
fermi vs MeV	$1 \text{ fm} = (197 \text{ MeV})^{-1}$	
second vs MeV	$1 \text{ s} = (6.57 \times 10^{-22} \text{ MeV})^{-1}$	
barn vs cm <sup>-2</sup>	$1 \text{ b} = 10^{-24} \text{ cm}^{-2}$	



# Bibliography

- [1] CMS COLLABORATION, *Observation of the  $B_c^+$  meson in  $PbPb$  and  $pp$  collisions at  $\sqrt{s_{NN}} = 5.02$  TeV*. arXiv:[2201.02659](#), 2022. CMS-HIN-20-004, CERN-EP-2021-259.
- [2] F. ARLEO, *Quenching of Hadron Spectra in Heavy Ion Collisions at the LHC*. Phys. Rev. Lett., **119**(6), 062302, 2017. doi: [10.1103/PhysRevLett.119.062302](#).
- [3] F. ARLEO and G. FALMAGNE, *Quenching of hadron spectra in  $XeXe$  and  $PbPb$  collisions at the LHC*. PoS, **HardProbes2018**, 075, 2019. doi: [10.22323/1.345.0075](#).
- [4] B. CARTER. *Large number coincidences and the anthropic principle in cosmology*. In *Confrontation of Cosmological Theories with Observational Data*, volume **63**, pages 291–298, 1974.
- [5] R. H. DICKE, *Dirac's Cosmology and Mach's Principle*. Nature, **192**, 440–441, 1961.
- [6] S. WEINBERG, *Anthropic Bound on the Cosmological Constant*. Phys. Rev. Lett., **59**, 2607–2610, 1987. doi: [10.1103/PhysRevLett.59.2607](#).
- [7] K. POPPER. *Conjectures and Refutations: The Growth of Scientific Knowledge*. 1963. ISBN 0-415-04318-2.
- [8] C. ROVELLI, *Aristotle's Physics: A Physicist's Look*. Journal of the American Philosophical Association, **1**(1), 23–40, 2015. doi: [10.1017/apa.2014.11](#).
- [9] H. YUKAWA, *On the interaction of elementary particles*. Proceedings of the Physico-Mathematical Society of Japan, Third Series, **17**(2), 48–57, 1935.
- [10] S. DESCOTES GENON and P. KOPPENBURG, *The CKM Parameters*. Ann. Rev. Nucl. Part. Sci., **67**, 97–127, 2017. doi: [10.1146/annurev-nucl-101916-123109](#).
- [11] D. J. GROSS and F. WILCZEK, *Ultraviolet Behavior of Non-Abelian Gauge Theories*. Phys. Rev. Lett., **30**, 1343–1346, 1973. doi: [10.1103/PhysRevLett.30.1343](#).
- [12] H. D. POLITZER, *Reliable Perturbative Results for Strong Interactions?* Phys. Rev. Lett., **30**, 1346–1349, 1973. doi: [10.1103/PhysRevLett.30.1346](#).
- [13] A. DEUR, S. BRODSKY, and G. de TERAMOND, *The QCD running coupling*. Progress in Particle and Nuclear Physics, **90**, 2016. doi: [10.1016/j.ppnp.2016.04.003](#).
- [14] J. M. CAMPBELL, J. W. HUSTON, and W. J. STIRLING, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*. Rept. Prog. Phys., **70**, 89, 2007. doi: [10.1088/0034-4885/70/1/R02](#).
- [15] J. D. BJORKEN, *Asymptotic Sum Rules at Infinite Momentum*. Phys. Rev., **179**, 1547–1553, 1969. doi: [10.1103/PhysRev.179.1547](#).
- [16] NNPDF COLLABORATION, *Parton distributions for the LHC Run II*. JHEP, **04**, 040, 2015. doi: [10.1007/JHEP04\(2015\)040](#).
- [17] L. DEL DEBBIO, *Parton distributions in the LHC era*. EPJ Web Conf., **175**, 01006, 2018. doi: [10.1051/epjconf/201817501006](#).

- [18] F. ARLEO, R. KOLEVATOV, S. PEIGNÉ, and T. SAMI, *Nuclear suppression in p-A collisions from induced radiation*. EPJ Web Conf., **112**, 04005, 2016. doi: 10.1051/epjconf/201611204005.
- [19] Y. L. DOKSHITZER, *Calculation of the Structure Functions for Deep Inelastic Scattering and  $e+e-$  Annihilation by Perturbation Theory in Quantum Chromodynamics*. Sov. Phys. JETP, **46**, 641–653, 1977.
- [20] V. N. GRIBOV and L. N. LIPATOV, *Deep inelastic e p scattering in perturbation theory*. Sov. J. Nucl. Phys., **15**, 438–450, 1972.
- [21] G. ALTARELLI and G. PARISI, *Asymptotic Freedom in Parton Language*. Nucl. Phys. B, **126**, 298–318, 1977. doi: 10.1016/0550-3213(77)90384-4.
- [22] R. PASECHNIK and M. ŠUMBERA, *Phenomenological Review on Quark–Gluon Plasma: Concepts vs. Observations*. Universe, **3**(1), 7, 2017. doi: 10.3390/universe3010007.
- [23] C.-Y. WONG, *Signatures of quark gluon plasma phase transition in high-energy nuclear collisions*. Nucl. Phys. A, **681**, 22–33, 2001. doi: 10.1016/S0375-9474(00)00477-2.
- [24] F. KARSCH, *Lattice QCD at high temperature and density*. Lect. Notes Phys., **583**, 209–249, 2002. doi: 10.1007/3-540-45792-5\_6.
- [25] H. SONG and U. W. HEINZ, *Extracting the QGP viscosity from RHIC data - A Status report from viscous hydrodynamics*. J. Phys. G, **36**, 064033, 2009. doi: 10.1088/0954-3899/36/6/064033.
- [26] STAR COLLABORATION, *Elliptic flow in Au + Au collisions at  $(\sqrt{s_{NN}} = 130 \text{ GeV})$* . Phys. Rev. Lett., **86**, 402–407, 2001. doi: 10.1103/PhysRevLett.86.402.
- [27] PHENIX COLLABORATION, *Suppression of hadrons with large transverse momentum in central Au+Au collisions at  $\sqrt{s_{NN}} = 130\text{-GeV}$* . Phys. Rev. Lett., **88**, 022301, 2002. doi: 10.1103/PhysRevLett.88.022301.
- [28] STAR COLLABORATION, *Disappearance of back-to-back high  $p_T$  hadron correlations in central Au+Au collisions at  $\sqrt{s_{NN}} = 200\text{-GeV}$* . Phys. Rev. Lett., **90**, 082302, 2003. doi: 10.1103/PhysRevLett.90.082302.
- [29] WA85 COLLABORATION,  *$Xi$ -, anti- $xi$ -, Lambda and anti-Lambda production in sulphur-tungsten interactions at 200 GeV/c per nucleon*. Phys. Lett. B, **270**, 123–127, 1991. doi: 10.1016/0370-2693(91)91548-A.
- [30] S. ALIOLI et al., *Monte Carlo event generators for high energy particle physics event simulation*. arXiv:**1902.01674**, 2019. MCnet-19-02.
- [31] P. PAGANINI. *Fundamentals of Particle Physics*. 2022. Textbook to be published in Cambridge University Press.
- [32] TOTEM COLLABORATION, *First measurement of elastic, inelastic and total cross-section at  $\sqrt{s} = 13 \text{ TeV}$  by TOTEM and overview of cross-section data at LHC energies*. Eur. Phys. J. C, **79**(2), 103, 2019. doi: 10.1140/epjc/s10052-019-6567-0.
- [33] M. G. ALBROW, T. D. COUGHLIN, and J. R. FORSHAW, *Central Exclusive Particle Production at High Energy Hadron Colliders*. Prog. Part. Nucl. Phys., **65**, 149–184, 2010. doi: 10.1016/j.ppnp.2010.06.001.
- [34] D. G. d'ENTERRIA, *Hard scattering cross-sections at LHC in the Glauber approach: From pp to pA and AA collisions*. arXiv:**nucl-ex/0302016**, 2003.
- [35] C. LOIZIDES, J. KAMIN, and D. d'ENTERRIA, *Improved Monte Carlo Glauber predictions at present and future nuclear colliders*. Phys. Rev. C, **97**(5), 054910, 2018. doi: 10.1103/PhysRevC.97.054910. [Erratum: Phys.Rev.C 99, 019901 (2019)].

- [36] M. GELL MANN, *A Schematic Model of Baryons and Mesons*. Phys. Lett., **8**, 214–215, 1964. doi: 10.1016/S0031-9163(64)92001-3.
- [37] N. BRAMBILLA, A. PINEDA, J. SOTO, and A. VAIRO, *Potential NRQCD: An Effective theory for heavy quarkonium*. Nucl. Phys. B, **566**, 275, 2000. doi: 10.1016/S0550-3213(99)00693-8.
- [38] J. P. LANSBERG, *J/ψ production at  $\text{sqrt}(s)=1.96$  and 7 TeV: Color-Singlet Model, NNLO\* and polarisation*. J. Phys. G, **38**, 124110, 2011. doi: 10.1088/0954-3899/38/12/124110.
- [39] M. BUTENSCHOEN and B. A. KNIEHL, *J/ψ production in NRQCD: A global analysis of yield and polarization*. Nucl. Phys. B Proc. Suppl., **222-224**, 151–161, 2012. doi: 10.1016/j.nuclphysbps.2012.03.016.
- [40] LHCb COLLABORATION, *Production of J/ψ and Upsilon mesons in pp collisions at  $\text{sqrt}(s) = 8$  TeV*. JHEP, **06**, 064, 2013. doi: 10.1007/JHEP06(2013)064.
- [41] LHCb COLLABORATION, *Measurement of J/ψ polarization in pp collisions at  $\sqrt{s} = 7$  TeV*. Eur. Phys. J. C, **73(11)**, 2631, 2013. doi: 10.1140/epjc/s10052-013-2631-3.
- [42] N. BRAMBILLA, J. GHIGLIERI, A. VAIRO, and P. PETRECZKY, *Static quark-antiquark pairs at finite temperature*. Phys. Rev. D, **78**, 014017, 2008. doi: 10.1103/PhysRevD.78.014017.
- [43] CMS COLLABORATION, *Fragmentation of jets containing a prompt J/ψ meson in PbPb and pp collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  TeV*. arXiv:[2106.13235](https://arxiv.org/abs/2106.13235), 2021. CMS-HIN-19-007, CERN-EP-2021-104.
- [44] J. D. BJORKEN, *Energy Loss of Energetic Partons in Quark - Gluon Plasma: Possible Extinction of High  $p(t)$  Jets in Hadron - Hadron Collisions*. 1982. FERMILAB-PUB-82-059-THY, FERMILAB-PUB-82-059-T.
- [45] STAR COLLABORATION, *Observation of  $D^0$  Meson Nuclear Modifications in Au+Au Collisions at  $\sqrt{s_{\text{NN}}} = 200$  GeV*. Phys. Rev. Lett., **113(14)**, 142301, 2014. doi: 10.1103/PhysRevLett.113.142301. [Erratum: Phys.Rev.Lett. 121, 229901 (2018)].
- [46] PARTICLE DATA GROUP COLLABORATION, *Review of Particle Physics*. Progress of Theoretical and Experimental Physics, **2020(8)**, 2020. doi: 10.1093/ptep/ptaa104. 083C01.
- [47] S. GODFREY, *Spectroscopy of  $B_c$  mesons in the relativized quark model*. Phys. Rev. D, **70**, 054017, 2004. doi: 10.1103/PhysRevD.70.054017.
- [48] CDF COLLABORATION, *Observation of  $B_c$  mesons in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.8$  TeV*. Phys. Rev. D, **58**, 112004, 1998. doi: 10.1103/PhysRevD.58.112004.
- [49] CDF COLLABORATION, *Evidence for the exclusive decay  $B_c^\pm \rightarrow J/\psi \pi^\pm$  and measurement of the mass of the  $B_c$  meson*. Phys. Rev. Lett., **96**, 082002, 2006. doi: 10.1103/PhysRevLett.96.082002.
- [50] D0 COLLABORATION, *Observation of the  $B_c$  Meson in the Exclusive Decay  $B_c \rightarrow J/\psi \pi$* . Phys. Rev. Lett., **101**, 012001, 2008. doi: 10.1103/PhysRevLett.101.012001.
- [51] LHCb COLLABORATION, *Measurements of  $B_c^+$  production and mass with the  $B_c^+ \rightarrow J/\psi \pi^+$  decay*. Phys. Rev. Lett., **109**, 232001, 2012. doi: 10.1103/PhysRevLett.109.232001.
- [52] LHCb COLLABORATION, *Measurement of the ratio of  $B_c^+$  branching fractions to  $J/\psi \pi^+$  and  $J/\psi \mu^+ \nu_\mu$* . Phys. Rev. D, **90**, 032009, 2014. doi: 10.1103/PhysRevD.90.032009.
- [53] CMS COLLABORATION, *Measurement of the ratio of the production cross sections times branching fractions of  $B_c^\pm \rightarrow J/\psi \pi^\pm$  and  $B^\pm \rightarrow J/\psi K^\pm$  and  $\mathcal{B}(B_c^\pm \rightarrow J/\psi \pi^\pm \pi^\pm \pi^\mp)/\mathcal{B}(B_c^\pm \rightarrow J/\psi \pi^\pm)$  in pp collisions at  $\sqrt{s} = 7$  TeV*. JHEP, **01**, 063, 2015. doi: 10.1007/JHEP01(2015)063.

- [54] ATLAS COLLABORATION COLLABORATION, *Measurement of the relative  $B_c^\pm/B^\pm$  production cross section with the ATLAS detector at  $\sqrt{s} = 8$  TeV*. Phys. Rev. D, **104**, 012010, 2021. doi: 10.1103/PhysRevD.104.012010.
- [55] LHCb COLLABORATION, *Measurement of  $B_c^+$  Production in Proton-Proton Collisions at  $\sqrt{s} = 8$  TeV*. Phys. Rev. Lett., **114**, 132001, 2015. doi: 10.1103/PhysRevLett.114.132001.
- [56] LHCb COLLABORATION, *Observation of the Decay  $B_c^+ \rightarrow B_s^0\pi^+$* . Phys. Rev. Lett., **111** (18), 181801, 2013. doi: 10.1103/PhysRevLett.111.181801.
- [57] LHCb COLLABORATION, *Study of  $B_c^+$  decays to the  $K^+K^-\pi^+$  final state and evidence for the decay  $B_c^+ \rightarrow \chi_{c0}\pi^+$* . Phys. Rev. D, **94**(9), 091102, 2016. doi: 10.1103/PhysRevD.94.091102.
- [58] LHCb COLLABORATION, *Observation of  $B_c^+ \rightarrow D^0K^+$  decays*. Phys. Rev. Lett., **118**(11), 111803, 2017. doi: 10.1103/PhysRevLett.118.111803.
- [59] LHCb COLLABORATION, *Measurement of the ratio of branching fractions  $\mathcal{B}(B_c^+ \rightarrow J/\psi\tau^+\nu_\tau)/\mathcal{B}(B_c^+ \rightarrow J/\psi\mu^+\nu_\mu)$* . Phys. Rev. Lett., **120**(12), 121801, 2018. doi: 10.1103/PhysRevLett.120.121801.
- [60] LHCb COLLABORATION, *Measurement of the  $B_c^+$  meson lifetime using  $B_c^+ \rightarrow J/\psi\mu^+\nu_\mu X$  decays*. Eur. Phys. J. C, **74**(5), 2839, 2014. doi: 10.1140/epjc/s10052-014-2839-x.
- [61] LHCb COLLABORATION, *Measurement of the  $B_c^-$  meson production fraction and asymmetry in 7 and 13 TeV pp collisions*. Phys. Rev. D, **100**(11), 112006, 2019. doi: 10.1103/PhysRevD.100.112006.
- [62] ATLAS COLLABORATION, *Observation of an excited  $B_c^\pm$  meson state with the ATLAS detector*. Phys. Rev. Lett., **113**, 212004, 2014. doi: 10.1103/PhysRevLett.113.212004.
- [63] CMS COLLABORATION, *Observation of two excited  $B_c^+$  states and measurement of the  $B_c^+(2S)$  mass in pp collisions at  $\sqrt{s} = 13$  TeV*. Phys. Rev. Lett., **122**, 132001, 2019. doi: 10.1103/PhysRevLett.122.132001.
- [64] LHCb COLLABORATION, *Observation of an excited  $B_c^+$  state*. Phys. Rev. Lett., **122**, 232001, 2019. doi: 10.1103/PhysRevLett.122.232001.
- [65] CMS COLLABORATION, *Measurement of  $B_c(2S)^+$  and  $B_c^*(2S)^+$  cross section ratios in proton-proton collisions at  $\sqrt{s} = 13$  TeV*. Phys. Rev. D, **102**(9), 092007, 2020. doi: 10.1103/PhysRevD.102.092007.
- [66] STAR COLLABORATION, *Evidence from  $d + Au$  measurements for final state suppression of high  $p(T)$  hadrons in  $Au+Au$  collisions at RHIC*. Phys. Rev. Lett., **91**, 072304, 2003. doi: 10.1103/PhysRevLett.91.072304.
- [67] S. PEIGNE, *Collisional Energy Loss of a Fast Parton in a QGP*. AIP Conf. Proc., **1038** (1), 139–148, 2008. doi: 10.1063/1.2987166.
- [68] Y. L. DOKSHITZER, V. A. KHOZE, and S. I. TROIAN, *On specific QCD properties of heavy quark fragmentation ('dead cone')*. J. Phys. G, **17**, 1602–1604, 1991. doi: 10.1088/0954-3899/17/10/023.
- [69] N. ARMESTO, C. A. SALGADO, and U. A. WIEDEMANN, *Medium induced gluon radiation off massive quarks fills the dead cone*. Phys. Rev. D, **69**, 114003, 2004. doi: 10.1103/PhysRevD.69.114003.
- [70] P. AURENCHE and B. G. ZAKHAROV, *Anomalous mass dependence of radiative quark energy loss in a finite-size quark-gluon plasma*. JETP Lett., **90**, 237–243, 2009. doi: 10.1134/S0021364009160048.

- [71] N. ARMESTO, A. DAINESI, C. A. SALGADO, and U. A. WIEDEMANN, *Testing the color charge and mass dependence of parton energy loss with heavy-to-light ratios at RHIC and CERN LHC*. Phys. Rev. D, **71**, 054027, 2005. doi: 10.1103/PhysRevD.71.054027.
- [72] ALICE COLLABORATION, *Perspectives for the study of charm in-medium quenching at the LHC with ALICE*. Eur. Phys. J. C, **33**, 495–503, 2004. doi: 10.1140/epjc/s2004-01645-4.
- [73] L. CUNQUEIRO and M. PŁOSKOŃ, *Searching for the dead cone effects with iterative declustering of heavy-flavor jets*. Phys. Rev. D, **99**(7), 074027, 2019. doi: 10.1103/PhysRevD.99.074027.
- [74] CMS COLLABORATION, *Measurement of  $B_s^0$  meson production in pp and PbPb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV*. Phys. Lett. B, **796**, 168–190, 2019. doi: 10.1016/j.physletb.2019.07.014.
- [75] CMS COLLABORATION, *Observation of  $B_s^0$  mesons and measurement of the  $B_s^0/B^+$  yield ratio in PbPb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV*. arXiv:[2109.01908](https://arxiv.org/abs/2109.01908), 2021. CMS-HIN-19-011, CERN-EP-2021-141.
- [76] CMS COLLABORATION, *Measurement of the  $B^\pm$  Meson Nuclear Modification Factor in Pb-Pb Collisions at  $\sqrt{s_{NN}} = 5.02$  TeV*. Phys. Rev. Lett., **119**(15), 152301, 2017. doi: 10.1103/PhysRevLett.119.152301.
- [77] CMS COLLABORATION, *Nuclear modification factor of  $D^0$  mesons in PbPb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV*. Phys. Lett. B, **782**, 474–496, 2018. doi: 10.1016/j.physletb.2018.05.074.
- [78] CMS COLLABORATION, *Measurement of prompt and nonprompt charmonium suppression in PbPb collisions at 5.02 TeV*. Eur. Phys. J. C, **78**(6), 509, 2018. doi: 10.1140/epjc/s10052-018-5950-6.
- [79] CMS COLLABORATION, *Charged-particle nuclear modification factors in PbPb and pPb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV*. JHEP, **04**, 039, 2017. doi: 10.1007/JHEP04(2017)039.
- [80] E. HÜCKEL and P. DEBYE, *The theory of electrolytes: I. lowering of freezing point and related phenomena*. Phys. Z., **24**(185-206), 1, 1923.
- [81] T. MATSUI and H. SATZ,  *$J/\psi$  Suppression by Quark-Gluon Plasma Formation*. Phys. Lett. B, **178**, 416–422, 1986. doi: 10.1016/0370-2693(86)91404-8.
- [82] CMS COLLABORATION, *Observation of Sequential Upsilon Suppression in PbPb Collisions*. Phys. Rev. Lett., **109**, 222301, 2012. doi: 10.1103/PhysRevLett.109.222301. [Erratum: Phys.Rev.Lett. 120, 199903 (2018)].
- [83] CMS COLLABORATION, *Measurement of nuclear modification factors of  $\Upsilon(1S)$ ,  $\Upsilon(2S)$ , and  $\Upsilon(3S)$  mesons in PbPb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV*. Phys. Lett. B, **790**, 270–293, 2019. doi: 10.1016/j.physletb.2019.01.006.
- [84] A. MOCSY, P. PETRECZKY, and M. STRICKLAND, *Quarkonia in the Quark Gluon Plasma*. International Journal of Modern Physics A, **28**, 2013. doi: 10.1142/S0217751X13400125.
- [85] M. LAINE, O. PHILIPSEN, and M. TASSLER, *Thermal imaginary part of a real-time static potential from classical lattice gauge theory simulations*. JHEP, **09**, 066, 2007. doi: 10.1088/1126-6708/2007/09/066.
- [86] N. BRAMBILLA, M. A. ESCOBEDO, J. GHIGLIERI, and A. VAIRO, *Thermal width and gluodissociation of quarkonium in pNRQCD*. JHEP, **12**, 116, 2011. doi: 10.1007/JHEP12(2011)116.
- [87] C. YOUNG and K. DUSLING, *Quarkonium above deconfinement as an open quantum system*. Physical Review C, **87**, 2010. doi: 10.1103/PhysRevC.87.065206.

- [88] Q. DU, A. DUMITRU, Y. GUO, and M. STRICKLAND, *Bulk viscous corrections to screening and damping in QCD at high temperatures.* JHEP, **01**, 123, 2017. doi: 10.1007/JHEP01(2017)123.
- [89] A. ISLAM and M. STRICKLAND, *Bottomonium suppression and elliptic flow using Heavy Quarkonium Quantum Dynamics.* JHEP, **21**, 235, 2020. doi: 10.1007/JHEP03(2021)235.
- [90] P. ZHUANG, L. YAN, and N. XU,  *$J/\psi$  continuous regeneration and suppression in quark gluon plasma.* J. Phys. G, **34**, S487–S494, 2007. doi: 10.1088/0954-3899/34/8/S38.
- [91] X. ZHAO and R. RAPP, *Medium Modifications and Production of Charmonia at LHC.* Nucl. Phys. A, **859**, 114–125, 2011. doi: 10.1016/j.nuclphysa.2011.05.001.
- [92] M. HE, H. van HEES, P. B. GOSSIAUX, R. J. FRIES, and R. RAPP, *Relativistic Langevin Dynamics in Expanding Media.* Phys. Rev. E, **88**, 032138, 2013. doi: 10.1103/PhysRevE.88.032138.
- [93] ALICE COLLABORATION,  *$J/\psi$  suppression at forward rapidity in Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV.* Phys. Rev. Lett., **109**, 072301, 2012. doi: 10.1103/PhysRevLett.109.072301.
- [94] ALICE COLLABORATION, *Centrality, rapidity and transverse momentum dependence of  $J/\psi$  suppression in Pb-Pb collisions at  $\sqrt{s_{NN}}=2.76$  TeV.* Phys. Lett. B, **734**, 314–327, 2014. doi: 10.1016/j.physletb.2014.05.064.
- [95] PHENIX COLLABORATION,  *$J/\psi$  Production vs Centrality, Transverse Momentum, and Rapidity in Au+Au Collisions at  $\sqrt{s_{NN}} = 200$  GeV.* Phys. Rev. Lett., **98**, 232301, 2007. doi: 10.1103/PhysRevLett.98.232301.
- [96] PHENIX COLLABORATION,  *$J/\psi$  suppression at forward rapidity in Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV.* Phys. Rev. C, **84**, 054912, 2011. doi: 10.1103/PhysRevC.84.054912.
- [97] ALICE COLLABORATION,  *$J/\psi$  elliptic flow in Pb-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV.* Phys. Rev. Lett., **119(24)**, 242301, 2017. doi: 10.1103/PhysRevLett.119.242301.
- [98] A. ANDRONIC, P. BRAUN MUNZINGER, K. REDLICH, and J. STACHEL, *Decoding the phase structure of QCD via particle production at high energy.* Nature, **561(7723)**, 321–330, 2018. doi: 10.1038/s41586-018-0491-6.
- [99] P. BRAUN MUNZINGER and J. STACHEL, *(Non)thermal aspects of charmonium production and a new look at  $J/\psi$  suppression.* Physics Letters B, **490(3-4)**, 196–202, 2000. doi: 10.1016/s0370-2693(00)00991-6.
- [100] E. G. FERREIRO, *Charmonium dissociation and recombination at LHC: Revisiting co-movers.* Phys. Lett. B, **731**, 57–63, 2014. doi: 10.1016/j.physletb.2014.02.011.
- [101] P. B. GOSSIAUX, V. GUIHO, and J. AICHELIN, *Charmonia enhancement in quark-gluon plasma with improved description of c-quarks phase-distribution.* J. Phys. G, **31**, S1079–S1082, 2005. doi: 10.1088/0954-3899/31/6/062.
- [102] X. YAO, W. KE, Y. XU, S. A. BASS, and B. MÜLLER, *Coupled Boltzmann Transport Equations of Heavy Quarks and Quarkonia in Quark-Gluon Plasma.* JHEP, **21**, 046, 2020. doi: 10.1007/JHEP01(2021)046.
- [103] X. DU, R. RAPP, and M. HE, *Color Screening and Regeneration of Bottomonia in High-Energy Heavy-Ion Collisions.* Phys. Rev. C, **96(5)**, 054901, 2017. doi: 10.1103/PhysRevC.96.054901.
- [104] M. SCHROEDTER, R. L. THEWS, and J. RAFELSKI,  *$B_c$  meson production in nuclear collisions at RHIC.* Phys. Rev., **C62**, 024905, 2000. doi: 10.1103/PhysRevC.62.024905.

- [105] Y. LIU, C. GREINER, and A. KOSTYUK,  *$B_c$  meson enhancement and the momentum dependence in  $Pb+Pb$  collisions at energies available at the CERN Large Hadron Collider*. Physical Review C, **87**(1), 014910, 2013. doi: 10.1103/physrevc.87.014910.
- [106] G. CHEN, C.-H. CHANG, and X.-G. WU,  *$B_c(B_c^*)$  meson production via the proton-nucleus and the nucleus-nucleus collision modes at the colliders RHIC and LHC*. Phys. Rev. D, **97**(11), 114022, 2018. doi: 10.1103/PhysRevD.97.114022.
- [107] F. ZIMMERMANN, *Future colliders for particle physics —“Big and small”*. Nucl. Instrum. Meth. A, **909**, 33–37, 2018. doi: 10.1016/j.nima.2018.01.034.
- [108] T. FAZZINI, G. FIDECARO, A. W. MERRISON, H. PAUL, and A. V. TOLLESTRUP, *Electron Decay of the Pion*. Phys. Rev. Lett., **1**, 247–249, 1958. doi: 10.1103/PhysRevLett.1.247.
- [109] F. J. HASERT et al., *Search for Elastic  $\nu_\mu$  Electron Scattering*. Phys. Lett. B, **46**, 121–124, 1973. doi: 10.1016/0370-2693(73)90494-2.
- [110] GARGAMELLE NEUTRINO COLLABORATION, *Observation of Neutrino Like Interactions Without Muon Or Electron in the Gargamelle Neutrino Experiment*. Phys. Lett. B, **46**, 138–140, 1973. doi: 10.1016/0370-2693(73)90499-1.
- [111] A. SALAM and J. C. WARD, *Weak and electromagnetic interactions*. Nuovo Cim., **11**, 568–577, 1959. doi: 10.1007/BF02726525.
- [112] S. L. GLASHOW, *Partial Symmetries of Weak Interactions*. Nucl. Phys., **22**, 579–588, 1961. doi: 10.1016/0029-5582(61)90469-2.
- [113] S. WEINBERG, *A Model of Leptons*. Phys. Rev. Lett., **19**, 1264–1266, 1967. doi: 10.1103/PhysRevLett.19.1264.
- [114] P. W. HIGGS, *Broken Symmetries and the Masses of Gauge Bosons*. Phys. Rev. Lett., **13**, 508–509, 1964. doi: 10.1103/PhysRevLett.13.508.
- [115] F. ENGLERT and R. BROUT, *Broken Symmetry and the Mass of Gauge Vector Mesons*. Phys. Rev. Lett., **13**, 321–323, 1964. doi: 10.1103/PhysRevLett.13.321.
- [116] G. S. GURALNIK, C. R. HAGEN, and T. W. B. KIBBLE, *Global Conservation Laws and Massless Particles*. Phys. Rev. Lett., **13**, 585–587, 1964. doi: 10.1103/PhysRevLett.13.585.
- [117] G. ’t HOOFT, *Renormalization of Massless Yang-Mills Fields*. Nucl. Phys. B, **33**, 173–199, 1971. doi: 10.1016/0550-3213(71)90395-6.
- [118] P. BRAMHAM et al., *Stochastic Cooling of a Stored Proton Beam*. Nucl. Instrum. Meth., **125**, 201–202, 1975. doi: 10.1016/0029-554X(75)90575-3.
- [119] S. Van der MEER, *Calibration of the Effective Beam Height in the ISR*. CERN-ISR-PO-68-31, 1968.
- [120] UA1 COLLABORATION, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at  $\sqrt{s} = 540$  GeV*. Phys. Lett. B, **122**, 103–116, 1983. doi: 10.1016/0370-2693(83)91177-2.
- [121] UA2 COLLABORATION, *Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider*. Phys. Lett. B, **122**, 476–485, 1983. doi: 10.1016/0370-2693(83)91605-2.
- [122] UA1 COLLABORATION, *Experimental Observation of Lepton Pairs of Invariant Mass Around 95  $GeV/c^2$  at the CERN SPS Collider*. Phys. Lett. B, **126**, 398–410, 1983. doi: 10.1016/0370-2693(83)90188-0.
- [123] UA2 COLLABORATION, *Evidence for  $Z^0 \rightarrow e^+e^-$  at the CERN  $\bar{p}p$  Collider*. Phys. Lett. B, **129**, 130–140, 1983. doi: 10.1016/0370-2693(83)90744-X.

- [124] G. ZWEIG, *An SU(3) model for strong interaction symmetry and its breaking, Version 2. In Developments in the quark theory of hadrons, Vol. 1. 1964 - 1978*, pages 22–101. 1964.
- [125] H. FRITZSCH, M. GELL MANN, and H. LEUTWYLER, *Advantages of the color octet gluon picture*. Physics Letters B, **47**(4), 365–368, 1973. doi: [https://doi.org/10.1016/0370-2693\(73\)90625-4](https://doi.org/10.1016/0370-2693(73)90625-4).
- [126] J. D. BJORKEN, *Highly Relativistic Nucleus-Nucleus Collisions: The Central Rapidity Region*. Phys. Rev. D, **27**, 140–151, 1983. doi: [10.1103/PhysRevD.27.140](https://doi.org/10.1103/PhysRevD.27.140).
- [127] J. RAFELSKI and B. MULLER, *Strangeness Production in the Quark-Gluon Plasma*. Phys. Rev. Lett., **48**, 1066, 1982. doi: [10.1103/PhysRevLett.48.1066](https://doi.org/10.1103/PhysRevLett.48.1066). [Erratum: Phys.Rev.Lett. 56, 2334 (1986)].
- [128] NA50 COLLABORATION, *Charmonium production in Pb Pb interactions at 158-GeV/c per nucleon*. Nucl. Phys. A, **638**, 261–278, 1998. doi: [10.1016/S0375-9474\(98\)00360-1](https://doi.org/10.1016/S0375-9474(98)00360-1).
- [129] G. BAUR et al., *Production of antihydrogen*. Physics Letters B, **368**(3), 251–258, 1996. doi: [https://doi.org/10.1016/0370-2693\(96\)00005-6](https://doi.org/10.1016/0370-2693(96)00005-6).
- [130] ALEPH, DELPHI, L3, OPAL, SLD, LEP ELECTROWEAK WORKING GROUP, SLD ELECTROWEAK GROUP AND SLD HEAVY FLAVOUR GROUP COLLABORATIONS, *Precision electroweak measurements on the Z resonance*. Phys. Rept., **427**, 257–454, 2006. doi: [10.1016/j.physrep.2005.12.006](https://doi.org/10.1016/j.physrep.2005.12.006).
- [131] J. TRAN THANH VAN, editor. *'94 electroweak interactions and unified theories. Proceedings, Leptonic Session of the 29th Rencontres de Moriond, Moriond Particle Physics Meeting, Meribel les Allues, France, 1994*. Ed. Frontieres.
- [132] CDF COLLABORATION, *Evidence for top quark production in  $\bar{p}p$  collisions at  $\sqrt{s} = 1.8$  TeV*. Phys. Rev. D, **50**, 2966–3026, 1994. doi: [10.1103/PhysRevD.50.2966](https://doi.org/10.1103/PhysRevD.50.2966).
- [133] A. BEURET et al. *The LHC Lead Injector Chain*. CERN-LHC-Project-Report-776, 2004.
- [134] R. BAILEY and P. COLLIER. *Standard Filling Schemes for Various LHC Operation Modes*. LHC-PROJECT-NOTE-323, CERN, Geneva, 2003.
- [135] C. LLEWELLYN SMITH, *Genesis of the Large Hadron Collider*. Phil. Trans. Roy. Soc. Lond. A, **373**(2032), 20140037, 2014. doi: [10.1098/rsta.2014.0037](https://doi.org/10.1098/rsta.2014.0037).
- [136] *Proceedings, ECFA-CERN Workshop on large hadron collider in the LEP tunnel: Lausanne and Geneva, Switzerland*. 1984. CERN-84-10-V-2. doi: [10.5170/CERN-1984-010-V-2](https://doi.org/10.5170/CERN-1984-010-V-2).
- [137] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>. CMS public results on luminosity and pile-up.
- [138] CMS COLLABORATION, *The CMS experiment at the CERN LHC*. JINST, **3**, S08004, 2008. doi: [10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004).
- [139] *Track Reconstruction Performance of the ATLAS Inner Detector at  $\sqrt{s} = 13$  TeV*. ATL-PHYS-PUB-2015-018, CERN, Geneva, 2015.
- [140] A. SALVUCCI, *Measurement of muon momentum resolution of the ATLAS detector*. EPJ Web Conf., **28**, 12039, 2012. doi: [10.1051/epjconf/20122812039](https://doi.org/10.1051/epjconf/20122812039).
- [141] LHCb COLLABORATION, *LHCb Detector Performance*. Int. J. Mod. Phys. A, **30**(07), 1530022, 2015. doi: [10.1142/S0217751X15300227](https://doi.org/10.1142/S0217751X15300227).
- [142] ALICE COLLABORATION, *Performance of the ALICE Experiment at the CERN LHC*. Int. J. Mod. Phys. A, **29**, 1430044, 2014. doi: [10.1142/S0217751X14300440](https://doi.org/10.1142/S0217751X14300440).

- [143] CMS COLLABORATION, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*. Phys. Lett. B, **716**, 30–61, 2012. doi: 10.1016/j.physletb.2012.08.021.
- [144] ATLAS COLLABORATION, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*. Phys. Lett. B, **716**, 1–29, 2012. doi: 10.1016/j.physletb.2012.08.020.
- [145] VIRGO AND LIGO SCIENTIFIC COLLABORATIONS, *Observation of Gravitational Waves from a Binary Black Hole Merger*. Phys. Rev. Lett., **116**(6), 061102, 2016. doi: 10.1103/PhysRevLett.116.061102.
- [146] LHCb COLLABORATION, *Observation of  $J/\psi p$  Resonances Consistent with Pentaquark States in  $\Lambda_b^0 \rightarrow J/\psi K^- p$  Decays*. Phys. Rev. Lett., **115**, 072001, 2015. doi: 10.1103/PhysRevLett.115.072001.
- [147] LHCb COLLABORATION, *Observation of an exotic narrow doubly charmed tetraquark*. arXiv:**2109.01038**, 2021. CERN-EP-2021-165, LHCb-PAPER-2021-031.
- [148] ALICE COLLABORATION, *Higher harmonic anisotropic flow measurements of charged particles in Pb-Pb collisions at  $\sqrt{s_{NN}}=2.76$  TeV*. Phys. Rev. Lett., **107**, 032301, 2011. doi: 10.1103/PhysRevLett.107.032301.
- [149] F. GELIS, *Some Aspects of the Theory of Heavy Ion Collisions*. Rept. Prog. Phys., **84**(5), 056301, 2021. doi: 10.1088/1361-6633/abec2e.
- [150] CMS COLLABORATION, *Observation of Long-Range Near-Side Angular Correlations in Proton-Proton Collisions at the LHC*. JHEP, **09**, 091, 2010. doi: 10.1007/JHEP09(2010)091.
- [151] ALICE COLLABORATION, *Enhanced production of multi-strange hadrons in high-multiplicity proton-proton collisions*. Nature Phys., **13**, 535–539, 2017. doi: 10.1038/nphys4111.
- [152] CMS COLLABORATION, *Measurement of isolated photon production in pp and PbPb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV*. Phys. Lett. B, **710**, 256–277, 2012. doi: 10.1016/j.physletb.2012.02.077.
- [153] CMS COLLABORATION, *Study of W boson production in PbPb and pp collisions at  $\sqrt{s_{NN}} = 2.76$  TeV*. Phys. Lett. B, **715**, 66–87, 2012. doi: 10.1016/j.physletb.2012.07.025.
- [154] ATLAS COLLABORATION, *Measurement of the centrality dependence of  $J/\psi$  yields and observation of Z production in lead-lead collisions with the ATLAS detector at the LHC*. Phys. Lett. B, **697**, 294–312, 2011. doi: 10.1016/j.physletb.2011.02.006.
- [155] CMS COLLABORATION, *Study of Z boson production in PbPb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV*. Phys. Rev. Lett., **106**, 212301, 2011. doi: 10.1103/PhysRevLett.106.212301.
- [156] PHENIX COLLABORATION, *Centrality dependence of low-momentum direct-photon production in Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV*. Phys. Rev. C, **91**(6), 064904, 2015. doi: 10.1103/PhysRevC.91.064904.
- [157] ALICE COLLABORATION, *Direct photon measurements in pp and Pb-Pb collisions with the ALICE experiment*. Nucl. Phys. A, **967**, 696–699, 2017. doi: 10.1016/j.nuclphysa.2017.05.094.
- [158] CMS COLLABORATION, *Study of W boson production in pPb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV*. Phys. Lett. B, **750**, 565–586, 2015. doi: 10.1016/j.physletb.2015.09.057.
- [159] ATLAS COLLABORATION, *Observation of a Centrality-Dependent Dijet Asymmetry in Lead-Lead Collisions at  $\sqrt{s_{NN}} = 2.77$  TeV with the ATLAS Detector at the LHC*. Phys. Rev. Lett., **105**, 252303, 2010. doi: 10.1103/PhysRevLett.105.252303.

- [160] CMS COLLABORATION, *Observation and studies of jet quenching in PbPb collisions at nucleon-nucleon center-of-mass energy = 2.76 TeV*. Phys. Rev. C, **84**, 024906, 2011. doi: 10.1103/PhysRevC.84.024906.
- [161] CMS COLLABORATION, *Modification of Jet Shapes in PbPb Collisions at  $\sqrt{s_{NN}} = 2.76$  TeV*. Phys. Lett. B, **730**, 243–263, 2014. doi: 10.1016/j.physletb.2014.01.042.
- [162] ATLAS COLLABORATION, *Measurements of the Nuclear Modification Factor for Jets in Pb+Pb Collisions at  $\sqrt{s_{NN}} = 2.76$  TeV with the ATLAS Detector*. Phys. Rev. Lett., **114** (7), 072302, 2015. doi: 10.1103/PhysRevLett.114.072302.
- [163] CMS COLLABORATION, *Measurement of inclusive jet cross sections in pp and PbPb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV*. Phys. Rev. C, **96**(1), 015202, 2017. doi: 10.1103/PhysRevC.96.015202.
- [164] ALICE COLLABORATION, *Exploration of jet substructure using iterative declustering in pp and Pb-Pb collisions at LHC energies*. Phys. Lett. B, **802**, 135227, 2020. doi: 10.1016/j.physletb.2020.135227.
- [165] CMS COLLABORATION, *Evidence of b-Jet Quenching in PbPb Collisions at  $\sqrt{s_{NN}} = 2.76$  TeV*. Phys. Rev. Lett., **113**(13), 132301, 2014. doi: 10.1103/PhysRevLett.113.132301. [Erratum: Phys. Rev. Lett. 115, 029903 (2015)].
- [166] ALICE COLLABORATION, *D meson nuclear modification factors in Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV with the ALICE detector*. Nucl. Phys. A, **904-905**, 635c–638c, 2013. doi: 10.1016/j.nuclphysa.2013.02.096.
- [167] ALICE COLLABORATION, *Azimuthal anisotropy of D meson production in Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV*. Phys. Rev. C, **90**(3), 034904, 2014. doi: 10.1103/PhysRevC.90.034904.
- [168] ALICE COLLABORATION, *D-meson nuclear modification factor and elliptic flow measurements in Pb-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV with ALICE at the LHC*. Nucl. Phys. A, **967**, 612–615, 2017. doi: 10.1016/j.nuclphysa.2017.05.062.
- [169] CMS COLLABORATION, *Evidence for X(3872) in PbPb collisions and studies of its prompt production at  $\sqrt{s_{NN}} = 5.02$  TeV*. arXiv:**2102.13048**, 2021. CERN-EP-2021-023.
- [170] ATLAS COLLABORATION, *Evidence for light-by-light scattering in heavy-ion collisions with the ATLAS detector at the LHC*. Nature Phys., **13**(9), 852–858, 2017. doi: 10.1038/nphys4208.
- [171] CMS COLLABORATION, *Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at  $\sqrt{s} = 7$  TeV*. JINST, **8**, P09009, 2013. doi: 10.1088/1748-0221/8/09/P09009.
- [172] CMS COLLABORATION, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*. JINST, **9**(10), P10009, 2014. doi: 10.1088/1748-0221/9/10/P10009.
- [173] CMS TRACKER GROUP, *The CMS Phase-1 Pixel Detector Upgrade*. JINST, **16**(02), P02027, 2021. doi: 10.1088/1748-0221/16/02/P02027.
- [174] R. FRUHWIRTH, *Application of Kalman filtering to track and vertex fitting*. Nucl. Instrum. Meth. A, **262**, 444–450, 1987. doi: 10.1016/0168-9002(87)90887-4.
- [175] CMS COLLABORATION, *Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at  $\sqrt{s} = 13$  TeV*. JINST, **13**(06), P06015, 2018. doi: 10.1088/1748-0221/13/06/P06015.

- [176] CMS COLLABORATION, *Particle-flow reconstruction and global event description with the CMS detector*. JINST, **12(10)**, P10003, 2017. doi: 10.1088/1748-0221/12/10/P10003.
- [177] CMS COLLABORATION, *Performance of the CMS Level-1 trigger in proton-proton collisions at  $\sqrt{s} = 13$  TeV*. JINST, **15(10)**, P10017, 2020. doi: 10.1088/1748-0221/15/10/P10017.
- [178] G. J. FELDMAN and R. D. COUSINS, *A Unified approach to the classical statistical analysis of small signals*. Phys. Rev., **D57**, 3873–3889, 1998. doi: 10.1103/PhysRevD.57.3873.
- [179] E. GROSS and O. VITELLS, *Trial factors for the look elsewhere effect in high energy physics*. Eur. Phys. J. C, **70**, 525–530, 2010. doi: 10.1140/epjc/s10052-010-1470-8.
- [180] CMS COLLABORATION, *Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements*. Eur. Phys. J. C, **80(1)**, 4, 2020. doi: 10.1140/epjc/s10052-019-7499-4.
- [181] C.-H. CHANG, X.-Y. WANG, and X.-G. WU, *BCVEGPY2.2: A newly upgraded version for hadronic production of the meson  $B_c$  and its excited states*. Comput. Phys. Commun., **197**, 335–338, 2015. doi: 10.1016/j.cpc.2015.07.015.
- [182] D. J. LANGE, *The EvtGen particle decay simulation package*. Nucl. Instrum. Meth. A, **462**, 152, 2001. doi: 10.1016/S0168-9002(01)00089-4.
- [183] D. EBERT, R. N. FAUSTOV, and V. O. GALKIN, *Rare semileptonic decays of  $B$  and  $B_c$  mesons in the relativistic quark model*. Phys. Rev. D, **82**, 034032, 2010. doi: 10.1103/PhysRevD.82.034032.
- [184] I. P. LOKHTIN and A. M. SNIGIREV, *A model of jet quenching in ultrarelativistic heavy ion collisions and high- $p_T$  hadron spectra at RHIC*. Eur. Phys. J. C, **45**, 211, 2006. doi: 10.1140/epjc/s2005-02426-3.
- [185] GEANT4 COLLABORATION, *GEANT4—a simulation toolkit*. Nucl. Instrum. Meth. A, **506**, 250, 2003. doi: 10.1016/S0168-9002(03)01368-8.
- [186] CMS COLLABORATION, *Measurement of production cross sections times branching fraction of  $B_c^+ \rightarrow J/\psi\pi^+$  and  $B^+ \rightarrow J/\psi K^+$  in pp collisions at  $\sqrt{s} = 7$  TeV at CMS*. CMS-PAS-BPH-13-002, CERN, Geneva, 2018.
- [187] LHCb COLLABORATION, *Measurement of the  $B^\pm$  production cross-section in pp collisions at  $\sqrt{s} = 7$  TeV*. Journal of High Energy Physics, **2012(4)**, 93, 2012. doi: 10.1007/JHEP04(2012)093.
- [188] CMS COLLABORATION, *Luminosity measurement in proton-proton collisions at 5.02 TeV in 2017 at CMS*. <http://cds.cern.ch/record/2765655>, CERN, 2021.
- [189] CMS COLLABORATION, *Van der Meer calibration of the CMS luminosity detectors in 2017*. CMS-CR-2018-339, 2019.
- [190] M. PIVK and F. R. LE DIBERDER, *SPlot: A Statistical tool to unfold data distributions*. Nucl. Instrum. Meth. A, **555**, 356–369, 2005. doi: 10.1016/j.nima.2005.08.106.
- [191] CMS COLLABORATION, *Performance of CMS muon reconstruction in pp collision events at  $\sqrt{s} = 7$  TeV*. JINST, **7**, P10002, 2012. doi: 10.1088/1748-0221/7/10/P10002.
- [192] LHCb COLLABORATION, *Measurement of the  $B_c^-$  meson production fraction and asymmetry in 7 and 13 TeV pp collisions*. Phys. Rev. D, **100(11)**, 112006, 2019. doi: 10.1103/PhysRevD.100.112006.
- [193] <https://cmsoms.cern.ch/cms/fills/summary>. (Internal) fill summary from the CMS web-based monitoring, fills 6372-6399 and 7427-7497.

- [194] H. Voss, A. Höcker, J. Stelzer, and F. Tegenfeldt, *TMVA, the Toolkit for Multivariate Data Analysis with ROOT*. XIth International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT), 2020.
- [195] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*. eConf, **C0303241**, MOLT007, 2003.
- [196] ATLAS, CMS, AND LHC HIGGS COMBINATION GROUP COLLABORATIONS. *Procedure for the LHC Higgs boson search combination in summer 2011*. ATL-PHYS-PUB-2011-011, CMS-NOTE-2011-005, 2011.
- [197] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*. Annals Math. Statist., **9**(1), 60–62, 1938. doi: 10.1214/aoms/1177732360.
- [198] CMS COLLABORATION, *Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV*. Eur. Phys. J. C, **75**, 212, 2015. doi: 10.1140/epjc/s10052-015-3351-7.
- [199] F. James. *MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1*. CERN-D506, 1994.
- [200] J. Conway. *Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra*. In *PHYSTAT 2011*, pages 115–120, 2011. doi: 10.5170/CERN-2011-006.115.
- [201] R. J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*. Comput. Phys. Commun., **77**, 219–228, 1993. doi: 10.1016/0010-4655(93)90005-W.
- [202] R. D. Cousins. *Generalization of Chisquare Goodness-of-Fit Test for Binned Data Using Saturated Models, with application to Histograms*. [http://www.physics.ucla.edu/~cousins/stats/cousins\\_saturated.pdf](http://www.physics.ucla.edu/~cousins/stats/cousins_saturated.pdf), University of California, Los Angeles, 2010.
- [203] J. Haigh. *Probability Models (Second Edition)*. Springer, 2013.
- [204] G. D. Lafferty and T. R. Wyatt, *Where to stick your data points: The treatment of measurements within wide bins*. Nucl. Instrum. Meth. A, **355**, 541–547, 1995. doi: 10.1016/0168-9002(94)01112-5.
- [205] C.-F. Qiao, P. Sun, D. Yang, and R.-L. Zhu,  *$B_c$  exclusive decays to charmonium and a light meson at next-to-leading order accuracy*. Phys. Rev. D, **89**(3), 034008, 2014. doi: 10.1103/PhysRevD.89.034008.
- [206] LHCb COLLABORATION, *Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions*. Phys. Rev. Lett., **118**(5), 052002, 2017. doi: 10.1103/PhysRevLett.118.052002. [Erratum: Phys.Rev.Lett. 119, 169901 (2017)].
- [207] R. Baier, Y. L. Dokshitzer, A. H. Mueller, S. Peigne, and D. Schiff, *Radiative energy loss of high-energy quarks and gluons in a finite volume quark - gluon plasma*. Nucl. Phys. B, **483**, 291–320, 1997. doi: 10.1016/S0550-3213(96)00553-6.
- [208] R. Baier, Y. L. Dokshitzer, A. H. Mueller, and D. Schiff, *Medium induced radiative energy loss: Equivalence between the BDMPS and Zakharov formalisms*. Nucl. Phys. B, **531**, 403–425, 1998. doi: 10.1016/S0550-3213(98)00546-X.
- [209] B. G. Zakharov, *On the energy loss of high-energy quarks in a finite size quark - gluon plasma*. JETP Lett., **73**, 49–52, 2001. doi: 10.1134/1.1358417.
- [210] F. Arleo, *Aspects of hard QCD processes in proton–nucleus collisions*. Nucl. Part. Phys. Proc., **289-290**, 71–76, 2017. doi: 10.1016/j.nuclphysbps.2017.05.014.
- [211] R. Baier, Y. L. Dokshitzer, A. H. Mueller, and D. Schiff, *Quenching of hadron spectra in media*. JHEP, **09**, 033, 2001. doi: 10.1088/1126-6708/2001/09/033.

- [212] CMS COLLABORATION, *Study of high- $pT$  charged particle suppression in  $PbPb$  compared to  $pp$  collisions at  $\sqrt{s_{NN}} = 2.76$  TeV*. Eur. Phys. J. C, **72**, 1945, 2012. doi: 10.1140/epjc/s10052-012-1945-x.
- [213] CMS COLLABORATION, *Charged-particle nuclear modification factors in  $XeXe$  collisions at  $\sqrt{s_{NN}} = 5.44$  TeV*. JHEP, **10**, 138, 2018. doi: 10.1007/JHEP10(2018)138.
- [214] PHENIX COLLABORATION, *Systematic study of nuclear effects in  $p+Al$ ,  $p+Au$ ,  $d+Au$ , and  $^3He+Au$  collisions at  $\sqrt{s_{NN}} = 200$  GeV using  $\pi^0$  production*. arXiv:**2111.05756**, 2021.
- [215] F. ARLEO, *Tomography of cold and hot QCD matter: Tools and diagnosis*. JHEP, **11**, 044, 2002. doi: 10.1088/1126-6708/2002/11/044.
- [216] R. SASSOT, P. ZURITA, and M. STRATMANN, *Inclusive Hadron Production in the CERN-LHC Era*. Phys. Rev. D, **82**, 074011, 2010. doi: 10.1103/PhysRevD.82.074011.
- [217] PHENIX COLLABORATION, *Neutral pion production with respect to centrality and reaction plane in  $Au+Au$  collisions at  $\sqrt{s_{NN}}=200$  GeV*. Phys. Rev. C, **87**(3), 034911, 2013. doi: 10.1103/PhysRevC.87.034911.
- [218] ATLAS COLLABORATION, *Measurement of charged-particle spectra in  $Pb+Pb$  collisions at  $\sqrt{s_{NN}} = 2.76$  TeV with the ATLAS detector at the LHC*. JHEP, **09**, 050, 2015. doi: 10.1007/JHEP09(2015)050.
- [219] ALICE COLLABORATION, *Transverse momentum spectra and nuclear modification factors of charged particles in  $pp$ ,  $p-Pb$  and  $Pb-Pb$  collisions at the LHC*. JHEP, **11**, 013, 2018. doi: 10.1007/JHEP11(2018)013.
- [220] ALICE COLLABORATION, *Measurement of  $D^0$ ,  $D^+$ ,  $D^{*+}$  and  $D_s^+$  production in  $Pb-Pb$  collisions at  $\sqrt{s_{NN}} = 5.02$  TeV*. JHEP, **10**, 174, 2018. doi: 10.1007/JHEP10(2018)174.
- [221] ATLAS COLLABORATION, *Prompt and non-prompt  $J/\psi$  and  $\psi(2S)$  suppression at high transverse momentum in 5.02 TeV  $Pb+Pb$  collisions with the ATLAS experiment*. Eur. Phys. J. C, **78**(9), 762, 2018. doi: 10.1140/epjc/s10052-018-6219-9.
- [222] C. LOIZIDES and A. MORSCH, *Absence of jet quenching in peripheral nucleus–nucleus collisions*. Phys. Lett. B, **773**, 408–411, 2017. doi: 10.1016/j.physletb.2017.09.002.
- [223] C. A. SALGADO and U. A. WIEDEMANN, *A Dynamical scaling law for jet tomography*. Phys. Rev. Lett., **89**, 092303, 2002. doi: 10.1103/PhysRevLett.89.092303.
- [224] M. GYULASSY, I. VITEV, and X. N. WANG, *High  $p(T)$  azimuthal asymmetry in noncentral  $A+A$  at RHIC*. Phys. Rev. Lett., **86**, 2537–2540, 2001. doi: 10.1103/PhysRevLett.86.2537.
- [225] ALICE COLLABORATION, *Centrality dependence of the charged-particle multiplicity density at midrapidity in  $Pb-Pb$  collisions at  $\sqrt{s_{NN}} = 5.02$  TeV*. Phys. Rev. Lett., **116**(22), 222302, 2016. doi: 10.1103/PhysRevLett.116.222302.
- [226] ALICE COLLABORATION, *Centrality dependence of the charged-particle multiplicity density at mid-rapidity in  $Pb-Pb$  collisions at  $\sqrt{s_{NN}} = 2.76$  TeV*. Phys. Rev. Lett., **106**, 032301, 2011. doi: 10.1103/PhysRevLett.106.032301.
- [227] CMS COLLABORATION, *Dependence on pseudorapidity and centrality of charged hadron production in  $PbPb$  collisions at a nucleon-nucleon centre-of-mass energy of 2.76 TeV*. JHEP, **08**, 141, 2011. doi: 10.1007/JHEP08(2011)141.
- [228] ALICE COLLABORATION, *Centrality and pseudorapidity dependence of the charged-particle multiplicity density in  $Xe-Xe$  collisions at  $\sqrt{s_{NN}} = 5.44$  TeV*. Phys. Lett. B, **790**, 35–48, 2019. doi: 10.1016/j.physletb.2018.12.048.

- [229] CMS COLLABORATION, *Pseudorapidity distributions of charged hadrons in xenon-xenon collisions at  $\sqrt{s_{NN}} = 5.44$  TeV*. Phys. Lett. B, **799**, 135049, 2019. doi: 10.1016/j.physletb.2019.135049.
- [230] PHENIX COLLABORATION, *Transverse energy production and charged-particle multiplicity at midrapidity in various systems from  $\sqrt{s_{NN}} = 7.7$  to 200 GeV*. Phys. Rev. C, **93**(2), 024901, 2016. doi: 10.1103/PhysRevC.93.024901.
- [231] C. LOIZIDES, J. NAGLE, and P. STEINBERG, *Improved version of the PHOBOS Glauber Monte Carlo*. SoftwareX, **1-2**, 13–18, 2015. doi: 10.1016/j.softx.2015.05.001.
- [232] A. DREES, H. FENG, and J. JIA, *Medium induced jet absorption at RHIC*. Phys. Rev. C, **71**, 034909, 2005. doi: 10.1103/PhysRevC.71.034909.
- [233] CMS COLLABORATION, *Pseudorapidity distributions of charged hadrons in proton-lead collisions at  $\sqrt{s_{NN}} = 5.02$  and 8.16 TeV*. JHEP, **01**, 045, 2018. doi: 10.1007/JHEP01(2018)045.
- [234] J. JIA, W. A. HOROWITZ, and J. LIAO, *A study of the correlations between jet quenching observables at RHIC*. Phys. Rev. C, **84**, 034904, 2011. doi: 10.1103/PhysRevC.84.034904.

# Acknowledgements

My thirst for research has led me to start a thesis, but only the presence and support of the people quoted here have brought me to its successful end. I would like to thank them warmly for backing me in this long journey; especially since it has been a roller-coaster going down with the debugging dead-ends and the crossroad questions, and going up with the ‘it is going to work!’ realizations and the communication of results to the heavy ion physics community.

There are so many things I would like to say to thank my supervisors Raphaël GRANIER DE CASSAGNAC and François ARLEO. This thesis would never have begun nor been finished without their infallible support, their sincere interest for the progress of my research and for my well-being, and their availability whenever needed. From my short experience, guided fortune seems to be a major driver of paths of research careers. I probably would not have studied the quark gluon plasma at LLR without the warm welcome of Raphaël and François in the peculiar environment of academic research, during my Bachelor internship. With François, it all began with a remake of the legend of Guillaume Tell using a tennis ball. It continued during the PhD with vivid discussions on energy loss where his capacity to bear my stream of often annoying questions (that his stoicism in front of a tennis ball had augured) was key. His theorist perspective on my experimental analysis, including in the introduction of the resulting paper, was also very helpful. With Raphaël, it started with a competition to bring more train station pictures to François. It continued with debates on the categorisation of backgrounds and very useful schedules when I was overwhelmed by the number of tasks ahead. I am also very grateful for his help on the writing and the review of the  $B_c^+$  paper during the critical last months of my PhD. Before starting my PhD, I heard many stories about bad relationships with thesis supervisors; I thank immensely François and Raphaël for proving that an efficient professional relationship does not exclude an enriching friendship. They have also shown me how one does physics alongside theatre, science fiction, games, and even ecology; I am beholden to them for giving legitimacy to my interests in other fields than particle physics.

The atmosphere in the CMS heavy-ion team of LLR has always been very welcoming and cheerful, as well as a safe space for asking questions. I warmly thank André STÅHL for his generous help in the beginning of my PhD and the preceding internship; it largely contributed to bringing me up to pace on technical CMS aspects. I am also grateful to Matthew NGUYEN for his help on CMS technicalities, his advice on analysis pathways, his many re-reads of the  $B_c^+$  paper, and his friendly and fuss-less leadership of the team. I also thank Batoul DIAB for her immutable good mood, contagious smile and lively discussions including on food and PhD setbacks, and Florian DAMAS for his sympathetic support during the last rough thesis months.

I supervised Natalie BLOT and Noémie PILLEUX as interns in the lab, and I thank them for their help on the  $B_c$  analysis, and for teaching me to teach and make clear. I also thank my colleagues Pedro GONZALEZ and Lizardo VALENCIA PALOMO for their contribution to parts of the acceptance and tag-and-probe studies. I owe Kevin STENSON a thorough, useful and always on point internal review of the  $B_c^+$  paper. Last but not least, I thank the CMS Collaboration at large for the exceptional opportunity to do great science within a great community.

I am grateful to the whole LLR, and in particular to its director Yves SIROIS, for having provided comfortable working conditions, including in covid times, and for having supported financially the conferences I attended and my six-month stay at CERN. I thank the French *Ministère de l’Enseignement Supérieur et de la Recherche* for funding my whole PhD.

My defence had the exceptional luck to take place with all members of my jury being physically present, and I sincerely thank them for this, for their careful reading of my manuscript, and for accepting to stand in my jury: Giulia MANCA, Andrea DAINESSE, Pascal PAGANINI, Pol-Bernard GOSSIAUX, Leticia CUNQUEIRO, and Bruno ESPAGNON. Their questions during the defence were spot-on and triggered very interesting discussions; it was a pleasant occasion for me to share details of the analysis that I had few occasions to mention. Additional thanks go to the *rappoeteurs* Giulia and Andrea for their deep questions and their report on my manuscript, but also to Pascal for offering me to help him teach his particle physics course in masters, and for his lecture notes that were almost my introduction to particle physics during my SLAC internship.

I am deeply obliged to my dear friends for having brought me so many invaluable out-of-thesis moments, that are more than vital for a full life – and a successful thesis. Immense thanks to Pablo for being such a great thesis-writing buddy, it was a great luck to live together in the last months of the thesis and to be able to share worries with a friend understanding so well the situation. Genuine thanks to Martin for epic bar outings, music sharing, festivals, and games, that started in our *classes prépas* with the great *Down the road* from C2C; he always reminds me of what good life is. Thanks to Luc, Théo and Martin for the nightly games of *Age of Empires II*, which almost managed to replace evening parties. Thanks to Anne-Kim for her profound friendship, always listening, joyful, and available, and for her constant support with thesis concerns. Thanks to Augustin for the outdoor and board game outings, and for the conversations on ecology that might have pushed me a step further. Thanks to my CERN friends, Emilien, Olivier, Alice, Xavier, Nadir, and Kinga, with whom I had intensely funny lunches, and in particular to Emilien who many times gave me guidance from his own ENS-to-CMS experience. And thanks to so many others for all the things we shared and will share: Antoine, Vincent, Rachel, Simon, Émilie, Lucas, Kevin, Auriane, Felipe, Marc, Joyce...

I would not *be* without my parents Thomas FALMAGNE and Isabelle DRAELANTS, but I would certainly not be the same person without their intense and exceptional upbringing. I got the perseverance and stubbornness necessary for being able to finish my PhD work from my father, and the thoroughness and perfectionism needed to do it well from my mother. I got my thirst for research from both of them and I thank them affectionately for having watered that seed of curiosity, including during a joyful childhood. My father told me “Le boeuf est lent, mais la terre est patiente” (“The ox is slow, but the land is patient”) many times to hearten my perseverance in the thesis. I have and will always follow this image of patiently ploughing my soil to always go forward, inspired by them being live models of excellent and passionate researchers. I also thank my father for his contribution to the quotes introducing each chapter of this thesis, and our friend Luc DEITZ for some of their translations. Special thanks for the support from my sisters Élizabeth and Bérénice, who tolerated my stubbornness and my daydreaming all those years; I also owe Élizabeth the good choices in many of my life decisions.

The strongest feelings are the hardest to put in words. But I owe my young-adult life path (in personal matters, of course, but also in professional ones) to my loved one, Maylis AVARO. I am grateful for where she took me, but even more for the present and future that she offers me. Short after we met, she already pushed me to reach out to SLAC researchers in Stanford, where I did my first masters internship. She always gave me unconditional help in re-reading or writing important emails, CVs or research projects. During the PhD, and particularly towards the end, she helped me refocus when needed. But more importantly, she guided me to take salutary pauses, notably by being my reason to pause. These pauses often offered me to take a step back on a (debugging) problem I was making no progress on. Covid-19 times were quite merry for us as we home-worked together – at least in the beginning. When these times dragged on towards the tough last thesis year, her warm presence and our caring conversations became essential for me to avoid breakdown and to focus on any good work. Her guidance in choices on any matter has always been relevant, while impressively accounting for my messy feelings. I feel blessed to be given to enjoy this guidance and all the things we share, for decades to come.



**Titre :** Le méson  $B_c^+$  en collisions d'ions lourds avec le détecteur CMS.

**Mots clés :** plasma de quarks et gluons,  $B_c$ , QGP, CMS, ions lourds, pertes d'énergie

**Résumé :** Cette thèse traite de comment l'hadronisation des quarks lourds et les partons de haute énergie sont affectées par le plasma de quarks et gluons (QGP) créé dans les collisions plomb-plomb (PbPb) au LHC. La première observation de mésons  $B_c^+$  en collisions d'ions lourds est menée en analysant des données du détecteur CMS. Partant d'une loi d'échelle existante dans un modèle de pertes d'énergie radiatives dans le QGP, une seconde loi d'échelle est mise en évidence.

L'analyse de données CMS en collisions proton-proton (2017) et PbPb (2018) à  $\sqrt{s_{NN}} = 5.02$  TeV mène à l'observation de désintégrations  $B_c^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)\mu^+\nu_\mu$ , et à la mesure du facteur de modification nucléaire du  $B_c^+$  en deux intervalles d'impulsion transverse du trimuon ( $p_T^{\mu\mu\mu}$ ) ou de la centralité de la collision PbPb. Trois bruits de fond principaux sont décrits par simulation ou par des méthodes spécifiques et fondées sur les données. Un arbre de décision amélioré (BDT) est entraîné sur le signal et les bruits de fond sélectionnés. Un ajustement d'histogrammes en probabilités (*likelihood template fit*) est opéré en intervalles de BDT, de masse invariante du trimuon, et de  $p_T^{\mu\mu\mu}$  ou de centralité. L'acceptance et l'efficacité de la chaîne de sélection sont évaluées

itérativement, avec le signal simulé dont le spectre en  $p_T^{\mu\mu\mu}$  est préalablement corrigé par une analyse préliminaire. Moins de suppression est observée pour le méson  $B_c^+$  que pour les autres saveurs lourdes ouvertes et fermées, excepté le méson  $B_s^0$ . Un probable adoucissement du spectre en  $p_T$  est trouvé. Ces résultats pourraient indiquer que la recombinaison de quarks lourds contribue significativement à la production de mésons  $B_c^+$ .

À haut  $p_T$  ( $\gtrsim 10$  GeV), la perte d'énergie radiative devrait dominer la suppression des hadrons dans le QGP. Un modèle de perte d'énergie radiative partonique existant, fondé sur le spectre BDMPS des gluons induits par le milieu, prédit une dépendance en  $p_T$  universelle du facteur de modification nucléaire. Des mesures dans des systèmes de différentes géométries et énergies sont collectées et ajustées à ce modèle, pour en extraire la perte d'énergie moyenne. Cette dernière est incluse dans une nouvelle loi d'échelle, qui la décrit comme dépendant uniquement de la taille du milieu et de la multiplicité en particules chargées. Cela permet d'extraire des propriétés d'expansion et de diffusion du milieu et potentiellement de prédire le coefficient d'asymétrie azimuthale  $v_2$  à haut  $p_T$ .

**Title :** The  $B_c^+$  meson in heavy-ion collisions with the CMS detector.

**Keywords :** quark gluon plasma,  $B_c$ , QGP, CMS, heavy ions, energy loss

**Abstract :** This thesis addresses how heavy-quark hadronisation and high-energy partons are affected by the quark-gluon plasma (QGP), a hot and dense medium created in lead-lead (PbPb) collisions at the LHC. Data from the CMS detector are analysed to achieve the first observation of  $B_c^+$  mesons in heavy-ion collisions. Building on an existing scaling law from a model of radiative energy loss in the QGP, a second scaling law is also brought to light.

The analysis of CMS data from 2017 proton-proton and 2018 PbPb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV leads to the observation of  $B_c^+ \rightarrow (J/\psi \rightarrow \mu^+\mu^-)\mu^+\nu_\mu$  decays, and to the measurement of the  $B_c^+$  nuclear modification factor in two bins of the trimuon transverse momentum ( $p_T^{\mu\mu\mu}$ ) and PbPb collision centrality. Three main backgrounds are described either with simulation or with specifically-designed data-driven samples. A boosted decision tree (BDT) is trained on the selected background and simulated signal candidates. A likelihood fit is run on templates binned in BDT, trimuon invariant mass, and  $p_T^{\mu\mu\mu}$  or centrality. The acceptance and efficiency of the selection chain are evaluated iteratively in each  $p_T^{\mu\mu\mu}$  or centrality bin with the simula-

ted signal, whose  $p_T^{\mu\mu\mu}$  spectrum is first corrected with the one measured in a preliminary analysis. The  $B_c^+$  meson is found to be less suppressed than all measured open and hidden heavy flavour mesons, except the  $B_s^0$  meson. The results also hint at a softening of the  $p_T$  spectrum in PbPb collisions. These may indicate that heavy-quark recombination is a significant  $B_c^+$  production mechanism.

At high  $p_T$  ( $\gtrsim 10$  GeV), radiative energy loss should be the dominant source of suppression of hadrons in the QGP. An existing model for the radiative energy loss of partons, based on the BDMPS medium-induced gluon spectrum, predicts a universal  $p_T$ -dependence of the nuclear modification factor. This fits measurements collected in systems of various geometries and energies, whose mean energy losses are extracted. A new scaling law is found consistent with the gathered measurements: it links the mean energy loss with the medium size and the inclusive charged particle multiplicity. This leads to an extraction of medium expansion and diffusion properties, and could predict the azimuthal asymmetry coefficient  $v_2$  at high  $p_T$ .