INSTITUT POLYTECHNIQUE DE PARIS

NNT : 2021IPPAX107

Thèse de doctorat

Study of J/ψ and D^0 production in $\sqrt{s_{\rm NN}} = 69$ GeV PbNe collisions with the LHCb experiment

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 10 décembre 2021, par

FELIPE A. GARCÍA ROSALES

Composition du Jury :

Yves Sirois Directeur de Recherche au CNRS, LLR, École polytechnique	Président
Bruno Espagnon Professeur des Universités, IJCLab, Orsay	Rapporteur
Javier Castillo Castillanos Ingénieur-Chercheur, IRFU CEA, Saclay	Rapporteur
Elena Gonzalez Ferreiro Profesora de la Universidad De Santiago de Compostela	Examinatrice
Frédéric Fleuret Directeur de Recherche au CNRS, LLR, École polytechnique	Directeur de thèse

Declaration of Authorship

I, Felipe A. GARCÍA ROSALES, declare that this thesis titled, "Study of J/ψ and D^0 production in $\sqrt{s_{\mathrm{NN}}} = 69$ GeV PbNe collisions with the LHCb experiment" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: _____

Date: _____

"It's the questions we can't answer that teach us the most. They teach us how to think. If you give a man an answer, all he gains is a little fact. But give him a question and he'll look for his own answers."

Patrick Rothfuss

INSTITUT POLYTECHNIQUE DE PARIS

Abstract

Laboratoire Leprince–Ringuet
École Polytechnique

Doctor of Philosophy

Study of J/ψ and D^0 production in $\sqrt{s_{\mathrm{NN}}} = 69$ GeV PbNe collisions with the LHCb experiment

by Felipe A. García Rosales

At the end of 2015, the LHCb collaboration has recorded the first collisions induced by the LHC proton and lead beams on a fixed target (gaseous targets). This new research programme will allow a thorough test, for the first time, of the colour screening mechanism predicted by lattice QCD (when producing a quark-gluon plasma in heavy-ion collisions).

The LHCb detector is optimised for heavy flavour measurements. In particular, it allows extremely accurate measurements of bound states such as D mesons, J/ψ, ψ' and χ_c considered as very sensitive probes for quark-gluon plasma studies.

Thanks to the LHCb SMOG system (System for Measuring the Overlap with Gas), initially intended for luminosity measurements, noble gases such as He, Ne, Ar, can be injected inside the vertex detector VELO (Vertex Locator). Acting as "fixed targets" for the LHC beams, they give access to proton-nucleus and nucleus-nucleus collisions at an optimal energy to study the phase transition from normal nuclear matter to a quark-gluon plasma. In 2018, LHCb recorded the first PbNe collisions at a centre-of-mass energy of 69 GeV. The results of their analysis can be directly compared to the results obtained with previously recorded pNe collisions at the same centre-of-mass energy.

The work presented in this thesis encompasses the totality of the PbNe data treatment and analysis, from the data-taking, to the data quality determination, the signal extraction and the efficiency computation. In addition, in this thesis the development of a software tool to determine the centrality information of heavy-ion collisions is presented. This tool was developed with PbPb and PbNe data, and is available for usage within the LHCb collaboration. These results allow to make a much more in-depth analysis of the results by studying the behaviour of the J/ψ and D^0 production not only in a new system, but in different centrality regimes within the same PbNe system.

Acknowledgements

First of all I would like to begin by thanking the members of the jury: Prof. Bruno Espagnon and Dr. Javier Castillo for their time and dedication to carefully reading my manuscript; Prof. Elena Gonzalez for her availability to come to Paris and the interest in my work; Dr. Yves Sirois for presiding the jury and, as the laboratory director, for welcoming and hosting me during my thesis at LLR; Lastly I want to thank Dr. Frédéric Fleuret, my thesis advisor, for his continuous support, guidance and endless discussions on too many topics to be listed. More profoundly, I would like to thank him for believing in me since the very beginning and pushing with me for the Ph.D. grant.

Next off, I would like to thank the LHCb team at LLR. Émilie takes a special place here since she was my supervisor during my Master 1 internship, my first approach to experimental high energy physics. With always a good mood and good sense of humour she made sure I had a great experience and that I always felt a part of the team, which was reinforced daily with the 4pm tea and very interesting, albeit sometimes random, discussions with Qinhua, of whom I am also deeply grateful. Then as my Ph.D. began I was lucky enough to meet Benjamin at CERN, who more than a colleague has become a friend that I have been fortunate to have in my office to always bother with silly and not-so-silly questions. I would like to thank him for his help in navigating LHCb, his explanations, even if I had him repeating them 10 times, and of course for giving me a top tour of Sardinia. I would also like to thank Vladik, with whom I had less time to share since we only became office mates right before the pandemic struck. Nonetheless, he is always ready for a laugh and open to any question or discussion, be it about bread or physics.

I want to give special thanks to Yanxi Zhang, who is an infinite source of knowledge and records of LHCb, for his help throughout my thesis. No matter the hour of the day or how busy his schedule was, he was always willing to help, clear doubts, and answer any question related to any aspect of LHCb or particle physics. Without his support this work would have been significantly harder.

At the risk of having a name slip my mind, I would like to thank more generally the people of LLR. I would like to first mention Pascal, who was one of the people that welcomed me to the master's programme when I arrived from Chile in 2016. He is a great teacher with a contagious enthusiasm for particle physics. I also extend my gratitude to Hamid, Marie-Thérèse, Alimata, Sylvaine and Farid for always giving me a friendly hand and for having helped me in the times I have needed it. I owe a big thank you to the entire IT team of LLR, whose work and support has been of invaluable importance for my work.

Since the road to a Ph.D. is not all work, I would also like to thank my fellow Ph.D. candidates, the ones that were already there when I arrived that shared their experience with me, the ones with whom we began this journey together, and the ones who started after me. Even if the pandemic has taken more than half of my time as a Ph.D. candidate, there were several occasions to enjoy the incredible city that is Paris. Not limited to LLR I would like to thank Martin, who for a long time was my chilean family in France, and a fundamental support with whom we went through quite peculiar situations, particularly apartment-wise. I would like to thank my friends Christina A., Antonios, Yasser, Georgy, José(lito), Martina, Matteo, Filippo, Simone, Alberto and surely more to add to this list, for all the good times and for always making Paris feel like home.

I thank as well my chilean friends dispersed around the world Andrés, Vale, Chelo, Matías, Felipe Oyarce, Adolfo, Franco, Jorge, Tofi, Christian, Roberto, Catalina,

you have no idea how sporadic messages or short conversations lifted my spirits when the end of the road did not seem so clear. I want to thank the group of my high school friends who have always been close to me and supportive. Last but not least I would like to thank my family in Chile who, in my trips there, always does a great job of helping me charge my batteries and make me come back with a full heart. The list can go on and it is inevitable to forget some names but to all of you, a big thank you.

Finally I want to thank the fundamental pillars of my life: Alice, che da subito mi ha rubato il cuore e che ammiro profondamente, per il suo incessante sostegno, la sua capacità di schiarirmi i pensieri e farmi ragionare quando la vita sembra travolgente, per la sua tenerezza e gli infiniti momenti di risate, e per spingirmi ogni giorno ad essere una persona migliore; Agradezco también a mis padres, que han estado siempre conmigo desde que tengo memoria y aún antes, que me han empujado siempre a conseguir mis objetivos y que me han enseñado lo importante de ser una persona íntegra, que me han dado lecciones que no se aprenden en ninguna escuela, que han hecho todo por mí y por mis sueños y nunca han mostrado una gota de egoísmo, que han sabido guiarme y aconsejarme en medio de mis dudas, que han sido simplemente incondicionales. ¡Gracias por todo! Este logro es para ustedes.

Contents

List of Figures

List of Tables

List of Abbreviations

BCID	**B**unch-**C**rossing **I**dentifier
BGI	**B**eam-**G**as **I**maging
CB	**C**rystal **B**all (function)
CERN	**C**onseil **E**uropéen pour la **R**echerche **N**ucléaire
CGC	**C**olour **G**lass **C**ondensate
CNM	**C**old **N**uclear **M**atter
CRMC	**C**osmic-**R**ay **M**onte **C**arlo
ECAL	**E**lectromagnetic **CAL**orimeter
GEM	**G**as **E**lectron **M**ultiplier
GBR	**G**radient **B**oosted **R**eweighter
GPD	**G**eneral **P**urpose **D**etector
HCAL	**H**adronic **CAL**orimeter
HLT	**H**igh-**L**evel **T**rigger
IP	**I**nteraction **P**oint
IT	**I**nner **T**racker
LEP	**L**arge **E**lectron-**P**ositron (collider)
LEIR	**L**ow-**E**nergy **I**on **R**ing
LHC	**L**arge **H**adron **C**ollider
LHCb	**L**arge **H**adron **C**ollider **b**eauty (experiment)
LINAC	**LIN**ear **AC**celerator
lQCD	**l**attice **Q**uantum **C**hromo**D**ynamics
L0	**L**evel **0** (trigger)
MB	**M**inimum **B**ias
MC	**M**onte **C**arlo
$\overline{\textbf{MS}}$	modified **M**inimal **S**ubstraction (scheme)
NBD	**N**egative **B**inomial **D**istribution
NEG	**N**on-**E**vaporative **G**etter
nPDF	**n**uclear **P**arton **D**istribution **F**unction
OT	**O**uter **T**racker
PDF	**P**arton **D**istribution **F**unction
PE	**PE**nning (gauge)
PI	**PI**rani (gauge)
PID	**P**article **ID**entification
PRS	**PR**e-**S**hower (detector)
PS	**P**roton **S**ynchrotron
PSB	**P**roton **S**ynchrotron **B**ooster
PU	**P**ile-**U**p
PV	**P**rimary **V**ertex
PZ	**P**ie**Z**o (gauge)
QCD	**Q**uantum **C**hromo**D**ynamics
QFT	**Q**uantum **F**ield **T**heory
QGP	**Q**uark-**G**luon **P**lasma
QMA	**Q**uadrupole **M**ass **A**nalyser

RF	**R**adio-**F**requency (cavities)
RHIC	**R**elativistic **H**eavy-**I**on **C**ollider
RICH	**R**ing **I**maging **CH**erenkov (detector)
RMS	**R**oot-**M**ean-**S**quare
SM	**S**tandard **M**odel
SMOG	**S**ystem for **M**easuring the **O**verlap with **G**as
SPD	**S**cintillating-**P**ad **D**etector
SPS	**S**uper **P**roton **S**ynchrotron
TCK	**T**rigger **C**onfiguration **K**ey
TIS	**T**rigger **I**ndependent of **S**ignal
TOS	**T**rigger **O**n **S**ignal
TT	**T**racker **T**uricensis
UPC	**U**ltra-**P**eripheral **C**ollision
UT	**U**pstream **T**racker
VELO	**VE**rtex **LO**cator
WLS	**W**ave**L**ength-**S**hifting (fibres)
2pF	**2**-**p**arameter **F**ermi (distribution)

List of Symbols

q	Quark
$\overline{q}$	Antiquark
u, d, c, s, t, b	Quarks of specific flavour
p	Proton
n	Neutron
m_q	Mass of quark q
c	Speed of light
v	Speed
p_{T}	Transverse momentum
x	Momentum fraction carried by a parton
$y^{(*)}$	Rapidity (in the centre-of-mass)
η	Pseudorapidity
S	Sulfur
U	Uranium
Pb	Lead
Au	Gold
Ne	Neon
Ar	Argon
v_n	Hydrodynamic flow coefficients
E_{T}	Transverse energy
N_{part}	Number of participating nucleons
N_{coll}	Number of binary nucleon-nucleon collisions
N_{anc}	Number of ancestors
N_{out}	Number of outgoing particles
b	Impact parameter
$\sigma_{\mathrm{NN}}^{\mathrm{inel}}$	Nucleon-nucleon cross-section
PVZ	Primary vertex Z coordinate

To my parents

Chapter 1

From quarks to heavy-ion collisions

Particles are the building blocks of everything around us. Some of them are present in nature as they are, some are forced to be confined in groups, and some disappear in a matter of instants through the decay into stable particles. It is no ordinary task to make sense of what is observed, and in fact the organisation of these building blocks began over 2 millennia ago, with the greatest progress achieved in the last century, taking us to what is now known as the Standard Model of particle physics (SM).

In this chapter a brief review of the SM is given, followed by a description of quarkonia and the quark-gluon plasma (QGP). Finally, some relevant experimental results are summarised, which will help put in context the results presented in this thesis (Ch. 5).

1.1 The Standard Model of particle physics

The atom (size $\sim 10^{-10}$ m) is composed by electrons "orbiting" around a nucleus made up of protons and neutrons (size $\sim 10^{-15}$ m), which in turn are themselves made up of quarks, three valence quarks to be precise, held together by the strong interaction. Quarks and electrons however, are not the sole fundamental particles out there. The three fundamental interactions present in the SM (the electromagnetic, strong and weak interactions) have their own messengers, called the force carriers or gauge bosons. The photon mediates the electromagnetic interaction, the Z^0 and $W^{\pm}$ bosons mediate the weak interaction and 8 gluons mediate the strong interaction. There is another "lone" boson which is not a force carrier but gives mass to the particles, the famous Higgs boson, which was predicted to exist in 1964 and whose discovery at the Large Hadron Collider (LHC) was announced in 2012 [1, 2].

The particles that make up the SM are summarised in Fig. 1.1. The 12 fermions are divided into the quarks and the leptons. There are 6 different quarks, each with a given "flavour", and 6 different leptons, in pairs of flavour, with each pair consisting of a charged and a neutral lepton. Both the quark and lepton families have a corresponding set of antiparticles. The difference between the quark and lepton families resides in the fact that the quarks possess a colour charge, an intrinsic property fundamental for the strong interaction giving rise to the emission of the force carriers, the gluons. It is somewhat an analogue to the electric charge for the electromagnetic interaction. The gluons, as opposed to the photon, do carry colour charge, making it possible to also interact among themselves. Particles made up by quarks are called hadrons, which can be either mesons, made up of a quark (q) and an antiquark ($\bar{q}$), or baryons made up of 3 quarks or 3 antiquarks. Some more

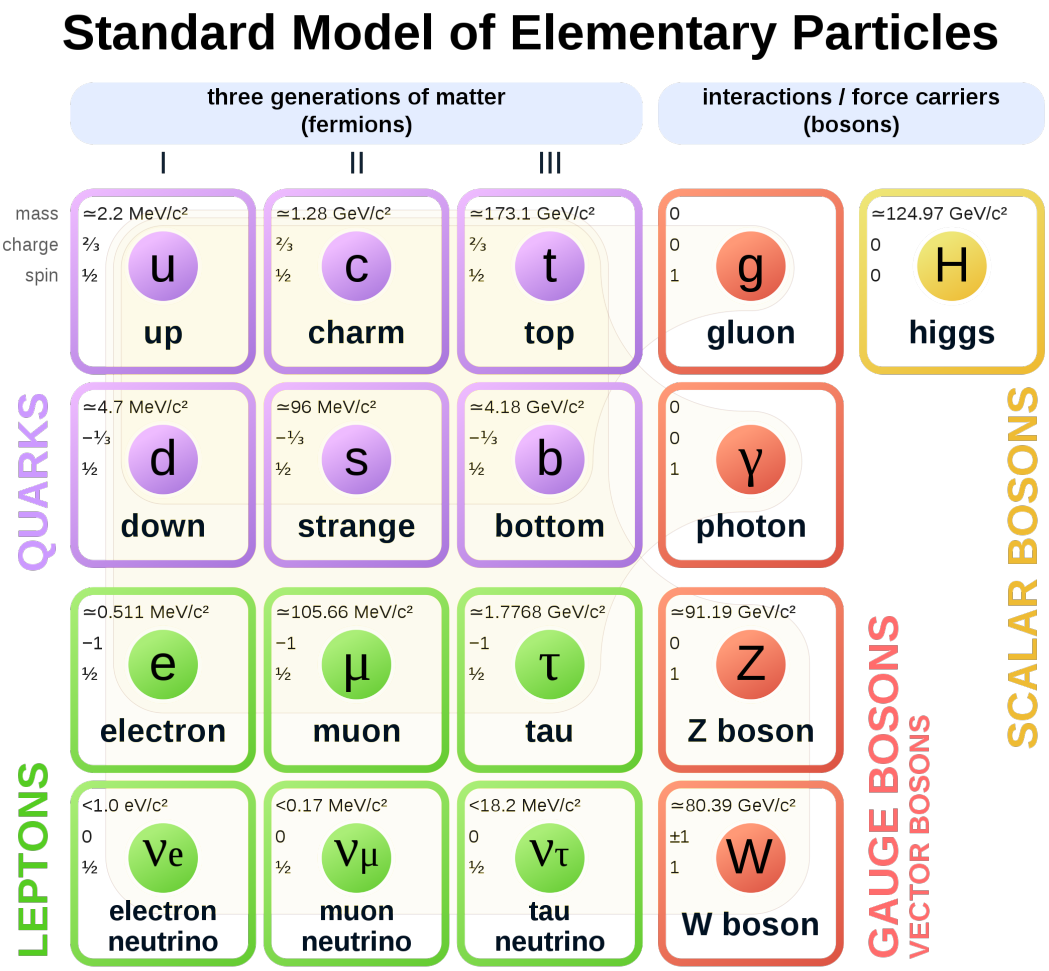

FIGURE 1.1: Fundamental particles of the Standard Model of particle physics. 12 fermions, 12 force carriers and the Higgs boson.

exotic arrangements of quarks/antiquarks have been seen as well, such as $q\bar{q}q\bar{q}$ or $qqqq\bar{q}$ [3–6].

The SM is a relativistic quantum field theory (QFT) [7] describing the particles as dynamic fields whose dynamics are given by the Lagrangian of the SM. The Lagrangian can more easily be navigated and used with the aid of Feynman path integral formalism, in which quantities can be computed as a perturbative expansion in powers of the coupling constants. Each term of the expansion can then be represented by a Feynman diagram, whose number of vertices corresponds to the power n of the given term.

The amplitude for a given process is then the sum of the amplitudes for each Feynman diagram. An illustration of how a given process can be decomposed is shown in Fig. 1.2. Naturally, this infinite sum can only be computed if the coupling constants are small compared to 1. In this scenario, already the lowest power of the expansion gives high precision results and more and more corrections are obtained as the higher order diagrams are computed.

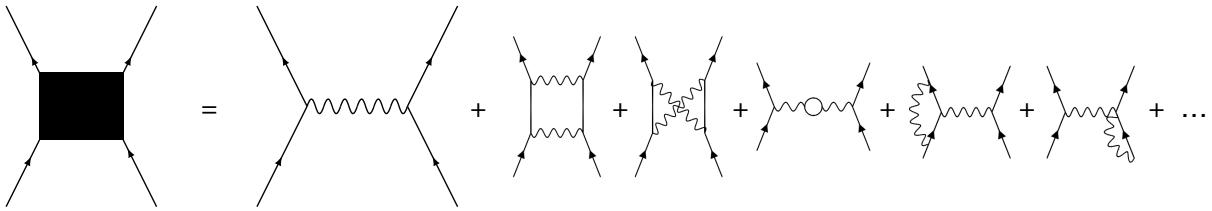

FIGURE 1.2: In the perturbative approach a given $2 \to 2$ process can be seen as the sum of all possible Feynman diagrams.

Quantum chromodynamics

Quantum chromodynamics (QCD) [8], is the branch of the SM which describes the theory of the strong interaction between quarks and gluons (colour charged particles, usually referred to as partons). It is a non-abelian[1] gauge theory with symmetry group SU(3). Quarks have 6 flavours and can have 3 different colour charges, which traditionally have been named red, blue and green, to make an analogy with the primary colours. The full QCD lagrangian can be written as

$$\mathcal{L}_{\text{QCD}} = \sum_{j=1}^{N_f} \overline{\psi}_j (i\gamma^\mu D_\mu - m_j)\psi_j - \frac{1}{4} F^a_{\mu\nu} F^{\mu\nu}_a . \tag{1.1}$$

The first term of eq. 1.1 describes the interactions between quarks, and between quarks and gluons through the covariant derivative $D_\mu = \partial_\mu - i g_s T_a G^a_\mu$. Here, T_a are the generators of the Lie algebra of SU(3), $G^{a=0,\dots 8}_\mu$ the 8 gluon fields, and $g_s = \sqrt{4\pi\alpha_s}$ is the strong interaction coupling parameter with α_s the renormalised QCD coupling constant. The sum goes to the number of possible quark flavours N_f. The second term describes the gluon-gluon interactions through the gluon field strength tensor F, defined as

$$F^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f^a{}_{bc} G^b_\mu G^c_\nu, \tag{1.2}$$

in which f^{abc} are the structure constants of SU(3). The first 2 terms of the gluon field strength tensor describe the gluon dynamics while the last term gives rise to the gluon self-interactions.

The possibility that gluons self-interact carries with it important consequences for the strong interactions. Most evident is the possibility that when 2 particles interact via the exchange of a gluon, this gluon can either produce a $q\overline{q}$ pair in a loop or split into 2 gluons (both at next-to-leading order, see Fig. 1.3). The former possibility gives rise to a colour charge screening effect (analogue to the effect present in quantum electrodynamics), while the latter leads to the dominant anti-screening effect. The screening effect leads to a decrease of the coupling constant as the distance between the interacting partons increase. Reversely, the anti-screening effect leads to an increase of the coupling constant with the distance between the partons. The overall effect of this is what is called "asymptotic freedom", where the QCD renormalised coupling constant α_s decreases with decreasing distance, or in terms of transferred energy Q, it decreases with larger transferred energy (the opposite is also true) [9], see Fig. 1.4.

The fact that α_s is not a fixed value, or even a tightly bound one, poses a problem since the perturbative approach will only be valid as long as α_s is sufficiently smaller

[1] A group is abelian if its generators commute.

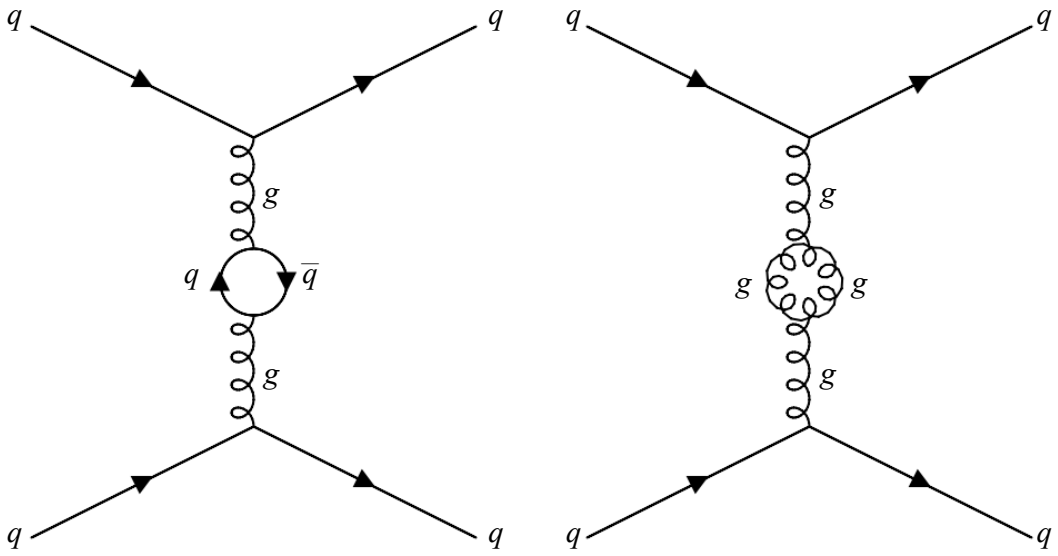

FIGURE 1.3: The diagram on the left shows a gluon exchange that would contribute to a screening effect. The diagram on the right shows an exchange in which a gluon self-interaction takes place. This results in an anti-screening effect.

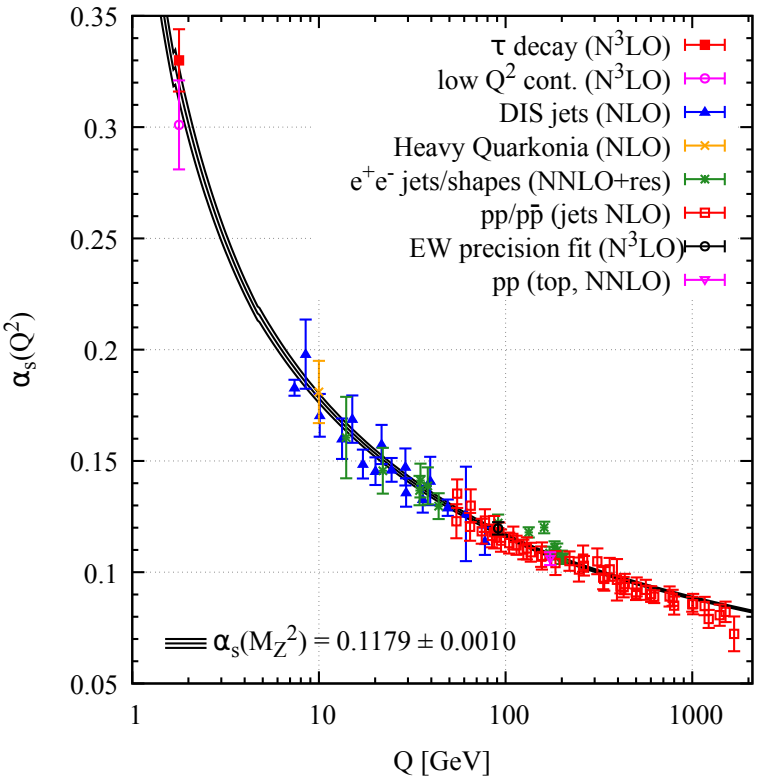

FIGURE 1.4: The "running" coupling constant α_s of QCD as a function of the transferred energy Q between the colour charges. Figure taken from the Particle Data Group [10].

TABLE 1.1: Quark masses as determined with the $\overline{\text{MS}}$ scheme and scales $\mu = 2\,\text{GeV}$ for light quarks, and $\mu = m_{Q=c,b,t}$ for heavy quarks. Values obtained from [10].

Light quark flavour	Up (u)	Down (d)	Strange (s)
Mass (MeV/c^2)	2.32 ± 0.10	4.71 ± 0.09	92.9 ± 0.7
Heavy quark flavour	**Charm (c)**	**Beauty (b)**	**Top (t)**
Mass (GeV/c^2)	1.280 ± 0.025	4.18 ± 0.03	172.76 ± 0.30

than 1. For this reason a distinction is made between different energy regimes, whose threshold is set at the QCD energy scale, given by $\Lambda_{\text{QCD}} \sim 200\,\text{MeV}$. This threshold separates the scale where $Q \gtrsim \Lambda_{\text{QCD}}$ and thus where $\alpha_s \lesssim 1$ and perturbation theory can be used, from the scale where $Q \lesssim \Lambda_{\text{QCD}}$ and thus where $\alpha_s \gtrsim 1$ and the strong interaction cannot be described with perturbative methods. The two extreme cases give rise to very different scenarios, these are the asymptotic freedom that was mentioned earlier and the colour confinement. The asymptotic freedom takes place at high energies where $Q \gg \Lambda_{\text{QCD}}$ (or small distances $\ll 1\,\text{fm}$), where the strong interaction becomes weaker. On the other hand, at low energies where $Q \lesssim \Lambda_{\text{QCD}}$ the colour confinement is present, which makes it impossible to observe a colour-charged particle individually. This is because as a pair of bound quarks begins to separate, the coupling constant grows to the point where it is more favourable to produce a new pair of quark-antiquark from the gluonic field, forming 2 new hadrons. As a matter of fact, quarks and gluons are always seen confined in composite systems that exhibit a neutral total colour charge. Colour confinement, being a non-perturbative effect is however still poorly understood, with the most accurate picture being the phenomenological Lund string model[2] [11] which describes the gluonic field between two quarks as strings and the confinement as a consequence of the break-up of the string to form a new quark pair, giving rise to 2 new independent composite systems.

Under normal conditions, the proton (p) and neutron (n) are the only stable hadrons in nature. When probed at intermediate momentum transfer they seem to be composed by 3 "valence" quarks, uud for the proton and udd for the neutron, but that hardly accounts for $\sim 1\%$ of their masses. In fact, when probed deeper, at low momentum transfer, the presence of gluons and "sea" quarks becomes important. They contribute to the total mass along with the energy involved in the QCD dynamics, thanks to the energy-mass equivalence. These sea quarks (mainly u, d and s) arise from pair production and individual gluons. Since coloured particles are confined, free quark masses (as they appear in the QCD Lagrangian) cannot be directly measured, and they are determined indirectly [10] from experimental values of hadron masses combined with a renormalisation scheme and a chosen scale, lattice QCD calculations and effective theories [12, 13]. Since there is always a dependence on the specific model used to determine the quark masses, it is important to always note which theoretical framework has been used to report any set of values. The most commonly used renormalisation scheme for QCD perturbation theory is the modified minimal subtraction scheme ($\overline{\text{MS}}$) [14], and with it the quark masses are determined to be the values reported in Tab. 1.1.

[2] Do not mistake with string theory.

TABLE 1.2: Some interesting charmonia states with their quantum numbers and mass [10].

Quantum state	Name	Mass (MeV/c^2)
1^3S_1	J/ψ or $\psi(1S)$	3096.900 ± 0.006
1^3P_0	χ_{c0}	3414.71 ± 0.30
1^3P_1	χ_{c1}	3510.67 ± 0.05
1^3P_2	χ_2	3556.17 ± 0.07
2^3S_1	ψ' or $\psi(2S)$	3686.10 ± 0.06

1.1.1 Quarkonia

Quarkonia are a type of meson that will take a central role in this thesis. These are quark-antiquark ($q\bar{q}$) bound states, but since the light quarks form quantum mechanical mixtures of their individual states[3], the term is generally reserved for $c\bar{c}$ and $b\bar{b}$ pairs, which are themselves called charmonia and bottomonia respectively, while the top quark does not form such states due to the fact that it decays so fast that $t\bar{t}$ pairs do not have time to form a bound state. Some charmonium states can be seen in Tab.1.2. The states are denoted by $n^{2S+1}L_J$ where n is the main quantum number, S the intrinsic angular momentum (spin quantum number), L is the orbital angular momentum in spectroscopic notation[4] and J the total angular momentum.

Quarkonia production

QCD allows for a factorisation of short and long distance interactions, also referred to as "hard" and "soft" scales respectively. This factorisation takes place both in the initial and final stages of the quarkonia production mechanism. It separates the initial colliding hadrons (soft scale), described by means of the parton distribution functions (PDFs), the parton scatterings that create the $q\bar{q}$ pairs (hard scale), described with perturbative QCD, and the hadronisation[5] of the $q\bar{q}$ pair into a specific state (soft scale). Hadronisation is described by effective theories such as the aforementioned Lund string model or non-perturbative QCD approaches, with the non-relativistic QCD framework [15], which disentangles the physics at the scale relevant for the heavy-quark production ($\sim m_q c^2$) from the physics at the scale of the quarkonium bound states energies ($\sim m_q v^2$ with $v \ll c$), being commonly used for quarkonium states.

The J/ψ can be produced either through quark annihilation or by gluon fusion. At high energies however, the dominant process is the gluon fusion [16]. The J/ψ production via a gluon fusion in a hadron-hadron collision is illustrated in Fig. 1.5. The production via gluon fusion can be described as a superposition of different interactions represented by the Feynman diagrams. The lowest order interactions, called t, u and s channel are shown in Fig. 1.6. These interactions have a consequence for the $q\bar{q}$ pair creation, by dictating if it is created as a colour singlet [17, 18] state

[3] For example consider the η meson which is a state defined as $\eta :\approx \frac{1}{\sqrt{6}}(u\bar{u} + d\bar{d} - 2s\bar{s})$.

[4] L takes alphabetic values which were originally assigned from characteristics of the spectroscopic lines corresponding to S, P, D and F orbitals (sharp, principal, diffuse, fundamental), followed alphabetically from G onwards. The 0 value corresponds to S, 1 to$P$, 2 to D and so on, with the exception of the letter J to avoid confusions.

[5] Hadronisation is the formation of a specific hadron from the resulting quarks of an interaction.

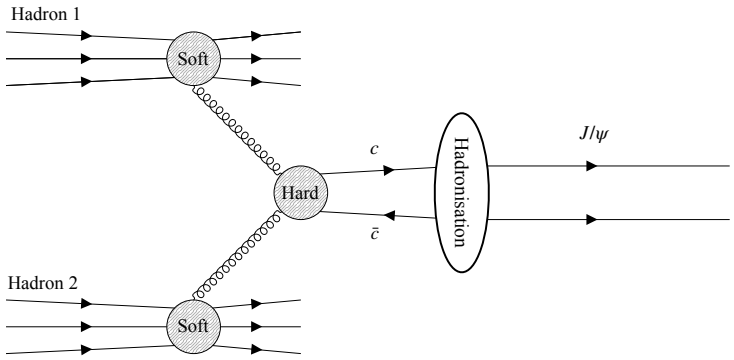

FIGURE 1.5: The QCD mechanism for the production of a $c\bar{c}$ pair hadronising into a J/ψ from gluon fusion in a hadron-hadron collision.

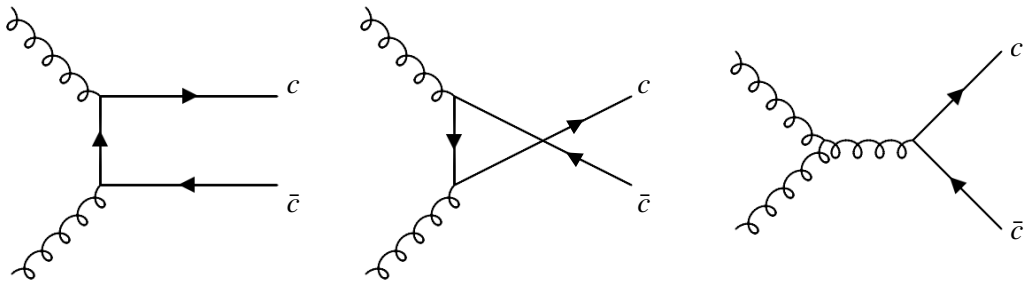

FIGURE 1.6: From left to right, the t, u and s channels for the $gg \to c\bar{c}$ process at leading order of the QCD perturbative expansion.

from the t and u channels, or as a colour octet state [19] from any channel.[6] The colour octet was introduced to fill the gap that existed between the data and the predictions made with only the colour singlet, which was originally thought to be the sole contributor given that only colour neutral particles are observed in nature. The inclusion of the colour octet contribution however, brought with it the necessity to include processes involving soft gluon exchanges with hadron remnants which are responsible of turning the colour octet state into a colour singlet. This resulted in a much better agreement with the data [20, 21], but the colour octet model introduces many new parameters which have to be determined from a fit to the data and also it still fails to describe satisfactorily the polarised charmonium data [22].

Quarkonia have been fundamental in the study of QCD. Being a $q\bar{q}$ pair, they constitute a symmetric system as opposed to baryons or asymmetrical mesons and, thanks to their high masses ($m_q \gg \Lambda_{QCD}$) which allow the factorisations of the different processes at different scales, they have provided crucial inputs for perturbative and non-perturbative QCD. Furthermore, they have been paramount in probing hadronic matter in high-energy heavy-ion collisions [23], particularly in the context of the quark-gluon plasma, as will be discussed in what follows.

[6] Colour singlet means that the outgoing quarks colour indices, i and k, are such that $i = \bar{k}$, while colour octet means that $i \neq k$.

1.2 The quark-gluon plasma

As predicted by QCD, in normal circumstances colour-charged particles are confined within hadrons, but at high enough temperature and energy density, accessible in the present day only through high energy heavy ion collisions, the formation of a QGP is expected [24, 25]. The QGP is a "deconfined" state of the nuclear matter, through which these colour charged particles are allowed to travel freely, thanks to the asymptotic freedom (see Fig. 1.4), not being bound to specific partners to form specific hadrons.

The quark-gluon plasma is a very special state of matter believed to have been in existence during the first moments after the Big-Bang, and on top of being a relatively newly discovered state of matter, its study and its phase transition into hadrons may shed light into the first moments of the Universe and its history.

Nowadays, far from the beginning of the Universe, QGP is a rare thing. It *might* be found in the core of neutron stars, and *may* be produced, albeit in small quantities, in heavy-ion collisions at high-energy particle colliders.[7]

Thermodynamics is used to study and describe the states of matter and their phase transitions, and based on an analogy between QFT and statistical mechanics a discrete formulation of QCD was developed, lattice QCD (lQCD) [26]. This is a useful tool to obtain theoretical predictions for non-perturbative phenomena, which are not easily approached analytically, by means of a partition function. The downside of lQCD however, is the rapidly increasing computation power needed to make more and more accurate predictions.[8] The QCD partition function reads [27]

$$Z(V,T,\mu_B) = \int \mathcal{D}A_\nu \mathcal{D}\overline{\psi}\mathcal{D}\psi e^{-S_E(V,T,\mu_B)}, \tag{1.3}$$

which is a Euclidean integral over the fields A_ν (gauge field), and $(\psi, \overline{\psi})$ (quark fields). To formally connect QFT and statistical mechanics, the classical time is replaced by the *imaginary time*, that is, by the Wick rotation $\tau = it$. [9] The thermodynamic variables that now describe the QCD system are the volume V, the temperature T and the baryon chemical potential μ_B, which as in statistical mechanics, represents the change in free energy with respect to a variation in the baryon number of the system.[10]

The Euclidean action can be decomposed as $S_E = S_G + S_F$, where S_G is a gluonic contribution which corresponds to

$$S_G(V,T) = \int_0^{T^{-1}} dx_0 \int_V d^3x \frac{1}{2}\mathrm{Tr}(F_{\mu\nu}F_{\mu\nu}) \tag{1.4}$$

and S_F is a fermionic contribution given by

$$S_F(V,T,\mu_B) = \int_0^{T^{-1}} dx_0 \int_V d^3x \sum_{f=1}^{N_f} \overline{\psi}_f(\gamma^\mu D_\mu + m_f - \mu_B\gamma_0)\psi_f, \tag{1.5}$$

[7] A QGP "fireball" under these conditions typically has a size of $\sim$ 10–20 fm, and lasts for about $\sim$ 5–10 fm/c.

[8] Nowadays, computing 1000 configurations on a lattice of $64^3 \times 16$ costs about 1 million core-hours. For a traditional supercomputer this means $\sim$ 30 000€ and 15 tons of CO_2, and for a GPU based super computer $\sim$ 10 000€ and 6 tons of CO_2.

[9] A Wick rotation (named after Gian Carlo Wick), is a method to find a solution to a mathematical problem in Minkowski space from a related problem in a Euclidean space by means of a substitution of an imaginary variable by a real number variable.

[10] The baryon number is a conserved quantity defined as $B = \frac{1}{3}(n_q - n_{\overline{q}})$, where n_q and $n_{\overline{q}}$ are the numbers of quarks and antiquarks respectively.

which couples the gauge and the fermion fields (N_f flavours). The system is then discretised into a 4D lattice, the spatial volume $V = (N_{\text{space}}a)^3$ and time $\tau = N_\tau a$,[11] where the degrees of freedom are now the lattice sites and the links connecting the sites. Like this the divergent path integrals are now regularised by the lattice spacing a and the results are extrapolated to the continuum limit $a \to 0$. With this at hand, the thermodynamic quantities such as the pressure p, the energy density ε, the free energy density f and the entropy density s can be obtained and related by means of the partition function.[12]

With these results from lQCD, and with the input from other models, a phase diagram can be put together to describe the phases of partonic matter, shown in Fig. 1.7. The phase diagram exhibits a cross-over phase transition between a hadronic gas and a deconfined partonic phase in a region of null baryon chemical potential, up to a critical value μ_B^{crit} [28]. Beyond μ_B^{crit} the transition is of first order [29],[13] that is, the temperature remains constant while the energy that is being released is used to transform one phase to the other. The critical temperature for the formation of QGP at zero chemical potential is of about $T_c \simeq 175 \pm 25\,\text{MeV}$ [30, 31]. While lQCD has helped achieve a good understanding of the equation of state of partonic matter, it has also brought with it important fluctuations in the predictions of other parameters and properties of the deconfined medium, such as the actual temperature of the QGP or its viscosity.

1.2.1 QGP in a heavy-ion collision

In high-energy heavy-ion collisions, due to the great number of colliding nucleons and the high multiplicity of particles produced, a deconfined state of matter may be reached if the conditions of the phase diagram of Fig. 1.7 are fulfilled.

From the moment of impact, a series of processes take place starting with the parton cascade. Then, the QGP formation and expansion is followed by the chemical freeze-out, and finally, the kinetic freeze-out. Fig. 1.8 shows a schematic view of the timeline of the collision.

From $t = 0$ to $t = \tau_0$, during the parton cascade, partons from both nuclei interact and produce a large number of particles. As the energy of the collision increases, gluons with a smaller momentum fraction x can interact, which effectively means that more and more gluons are "visible", even more than valence and sea quarks [34]. The initial parton distributions within a nucleon are described by PDFs $f^N_{g,u,d,\ldots}$. Within a nucleus however, a superposition of these PDFs is not enough to account for the distributions of the partons, and nuclear PDFs (nPDFs) $f^A_{g,u,d,\ldots}$ are defined taking into account possible interactions between the nucleons [35]. These interactions may result in an increase or a decrease of the gluon density relative to the one in a single nucleon. This effect is quantified by the nuclear modification ratio defined as

$$R_g^A(x, Q^2) = \frac{f_g^A(x, Q^2)}{A f_g^N(x, Q^2)}, \tag{1.6}$$

[11] There are several possible ways to discretise space-time, but a uniform regularly-spaced grid is the most commonly used (almost exclusively). This is because little is expected to be gained by introducing anisotropies or randomness to the lattice structure while the computing complexities grow significantly.

[12] The relation between different thermodynamic variables that describe a system is called an *equation of state*.

[13] A first order phase transition is characterised by discontinuity in the derivative of the free energy of the system.

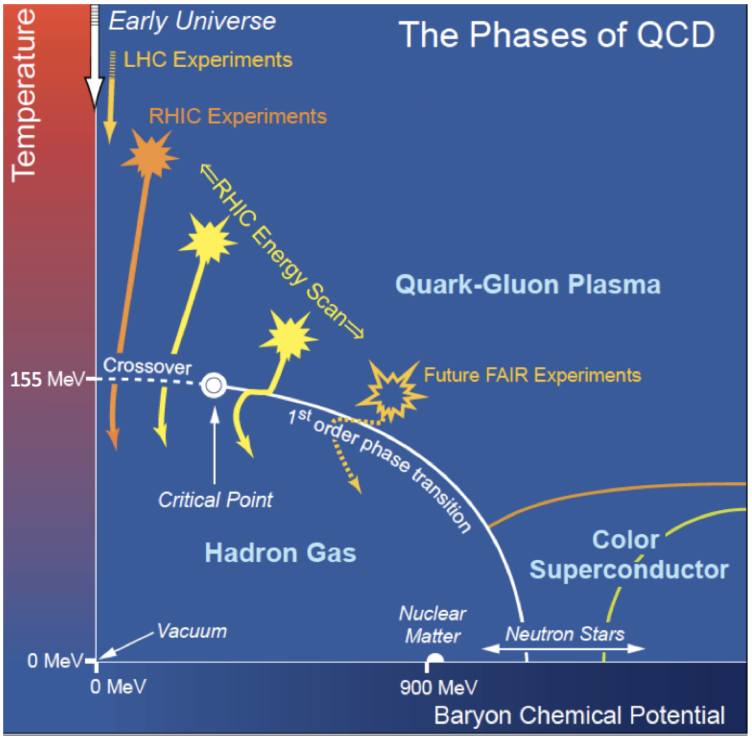

FIGURE 1.7: Phase diagram of QCD as function of temperature and baryon chemical potential μ_B. The regions covered by different experiments are shown. The region covered by the LHC corresponds to the collider configuration of LHC and does not consider the possible functioning of LHCb as a fixed-target experiment. Figure taken from Ref. [32].

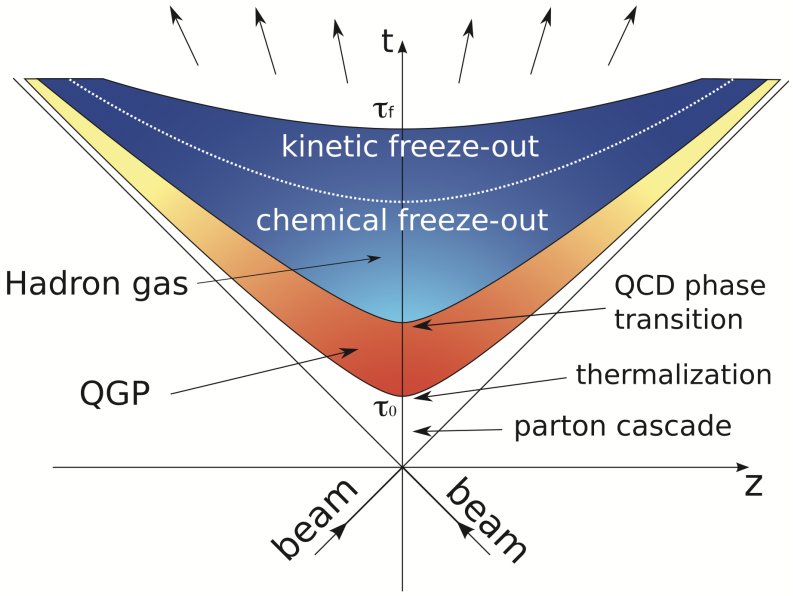

FIGURE 1.8: Space-time evolution of a heavy-ion collision. Figure taken from Ref. [33].

where A is the amount of nucleons present in the nucleus being considered.[14] When $R_g^A(x, Q^2)$ is less than 1, it means there is a *shadowing*, and conversely when it is greater than 1, there is an *anti-shadowing* [36]. Above a certain energy scale however, as the gluon density reaches a saturation point [37], the nPDFs are not enough to describe the multiple interactions that take place, and other non-perturbative methods, such as the Colour Glass Condensate (CGC), are used instead [38, 39]. The CGC model predicts the formation of the Glasma, which is composed of the original partons and the new particles (light and heavy quarks, photons, etc) originating mainly from gluon scatterings. The particle density and the temperature of the system increase dramatically in a very short time leading to the next stage of the collision, the QGP.[15]

After $t = \tau_0$, the chaotic Glasma thermalises (at least locally) giving rise to the QGP. The QGP properties depend on its temperature which, depending on the initial conditions of the collision, should go from a few T_c down to T_c. At very high temperatures ($T \gg T_c$), the coupling should be very weak $\alpha_s(T \gg \Lambda_{QCD}) \ll 1$, and the QGP may behave as a weakly interacting parton gas. At the temperatures of interest of a few T_c, the story is a bit different and lQCD predicts that partons may interact rather strongly, leading to a strongly-coupled QGP which would behave as an almost perfect fluid with low viscosity, and which can be described by quasi ideal hydrodynamics.

As the QGP expands and its temperature decreases reaching T_c, the phase transition into a hadron gas begins and the partons coalesce into hadrons (hadronisation stage). These hadrons will continue to interact inelastically with each other and with other nucleons that may have survived the collision intact, changing the nature of hadrons formed. Once the number and the species of hadrons becomes fixed, the so called chemical freeze-out is reached. Subsequently, when elastic interactions stop at $t = \tau_f$, and the particle kinematics is fixed, the last stage of the collision is reached, known as kinetic freeze-out.

After the kinetic freeze-out, the hadrons and leptons formed during the collision will either travel or decay into other particles and finally reach the detectors that are used to study the collisions.

1.2.2 Probes of QGP

Experimentally, there are several ways to probe the QGP of which only a brief description will be given with the exception of quarkonia suppression, which will be treated in more detail in Sec. 1.3.

Generally speaking, the probes of QGP can be classified into soft and hard probes. The soft probes consider the particles that are produced all along the QGP formation process and throughout its lifetime. These particles form the bulk of the medium and correspond mainly to light quarks and gluons but can also consider thermal photons and dileptons [40, 41] which can be used to probe the temperature of the medium. The soft probes can be studied in several ways, such as the hydrodynamic flow and strangeness enhancement which carry information of the collective behaviour that arises from the heavy-ion collision. On the other hand, hard probes such as the jet quenching and the quarkonia production can shed light on the early stages of the collision, since they are formed in the first moments of the collision, as well as on the characteristics of the medium as they must travel through it. Some of these probes will be discussed in what follows but this is by no means a complete list of probes.

[14] If the nPDFs were a superposition of PDFs, then $R_g^A(x, Q^2)$ would be equal to unity.

[15] The thermalisation itself however, is still poorly understood within the CGC model.

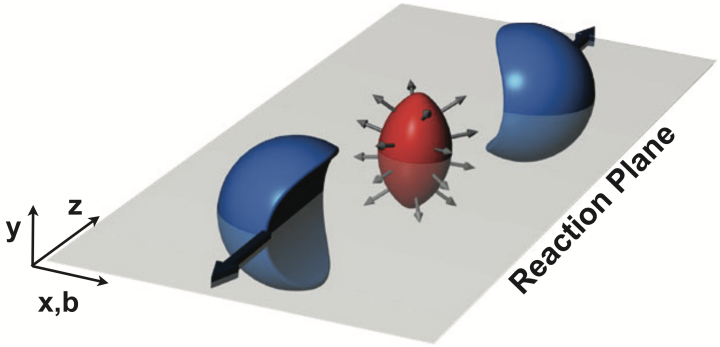

FIGURE 1.9: Reaction plane of a non-central heavy-ion collision. Figure taken from Ref. [42].

Hydrodynamic flow

Light hadrons made up of light quarks cannot be used individually to infer properties of the QGP. This is because their binding energy is much lower than the temperature of the medium, which leads to the "melting" of these hadrons entirely into the QGP and thus the light hadrons that are observed have been created either in the "shell" of the QGP (the phase boundary between the QGP and the vacuum) or during the final stages of the collision where the medium has already cooled down. Provided that the quarks and gluons undergo several interactions before hadronising, the medium will thermalise and exhibit a collective behaviour, which in turn can shed light on the dynamics of the QGP as a fluid, hence the name *Hydrodynamic flow*. This collective expansion, imprinted in the final kinematics of the light hadrons, can be correlated to the initial stage of the QGP, that is, to its initial energy density, its viscosity and the centrality of the collision. The presence of azimuthal anisotropy in the particle production is a clear sign of collective behaviour in heavy-ion collisions. This anisotropy is caused by the spatial asymmetry of the geometry of the collision, particularly in non-central events.[16]

To study the flow, an event (or reaction) plane is defined by the beam axis and the impact parameter of the nuclei,[17] see Fig. 1.9. The azimuthal anisotropy of the collective behaviour is studied by means of the multiplicity (N) distribution as a function of the azimuthal angle ϕ of the particles, which is decomposed into a Fourier series as

$$\frac{dN}{d\phi} = \frac{N}{2\pi}(1 + 2\sum_{n=1}^{\infty} v_n cos[n(\phi - \Psi_n)]), \tag{1.7}$$

where ϕ is the azimuthal angle and Ψ_n the azimuthal angle of the symmetry plane for the n^{th} harmonic. The coefficients v_1 and v_2 are known as the directed and elliptic flow respectively, which are the easiest to visualise, but experiments nowadays are capable of measuring much higher order terms. There are other methods to study flow, such as the cumulants [43], but these will not be treated in this thesis.

[16] The centrality of a collision will be the main topic of Chapter 4. For now it suffices to say that a non-central collision means that only part of the colliding nuclei actually undergo an inelastic collision and part of the nuclei remains intact. In other words, it means that the collision is not "head-on". It is usually quantified in percentages, 0% describing a full-central collision and 100% a peripheral collision where nuclei only graze each other.

[17] The impact parameter is the transverse distance of the two nuclei. This will be more clearly defined in Ch. 4.

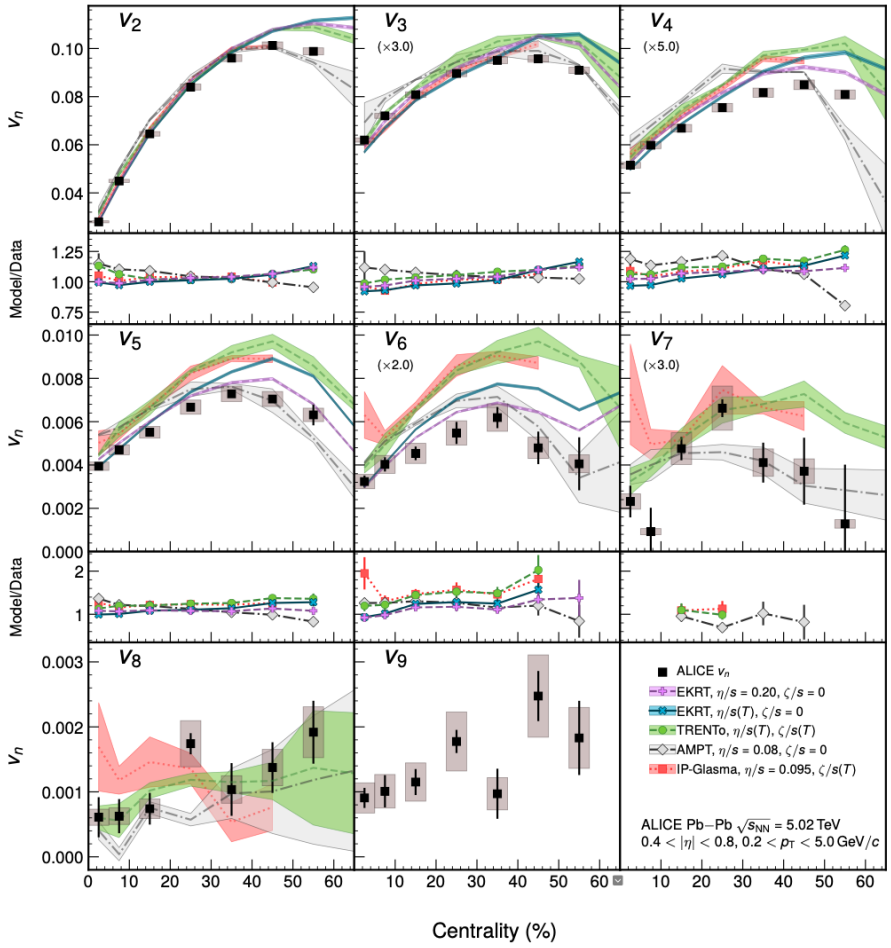

FIGURE 1.10: Flow coefficients up to the ninth order as a function of collision centrality, along with five different hydrodynamic calculations shown as color bands. On the bottom part of each panel, the ratios between model calculations and the data are shown with symbols. For some panels, the points are scaled by an indicated factor for better visibility across the panels. Figure taken from [44].

The flow present in heavy-ion collisions can be interpreted as an effect of multiple interactions between the partonic matter in the medium created by the collision. The higher the number of interactions the larger is the magnitude of the flow, and the closer the system is to thermalisation. Thus the flow is a very direct probe of the level of thermalisation [42].

In Fig. 1.10 the flow coefficients up to order 9 as measured by the ALICE experiment are shown. These are measured from lead-lead (PbPb) collision data from the 2015 heavy-ion run at the LHC, at a centre-of-mass energy of $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$ [44].

Strangeness enhancement

The strange quarks having a mass of roughly $93\,\mathrm{MeV}/c^2$ can very well be produced by the thermal medium in a heavy-ion collision. Indeed, their production mechanism is analogue to the one shown in Fig. 1.6 for the charm quark, whereas the quark-antiquark annihilation into a virtual gluon which then splits into an $s\bar{s}$ pair is a much slower contribution. For this reason the strange production in a QGP-forming

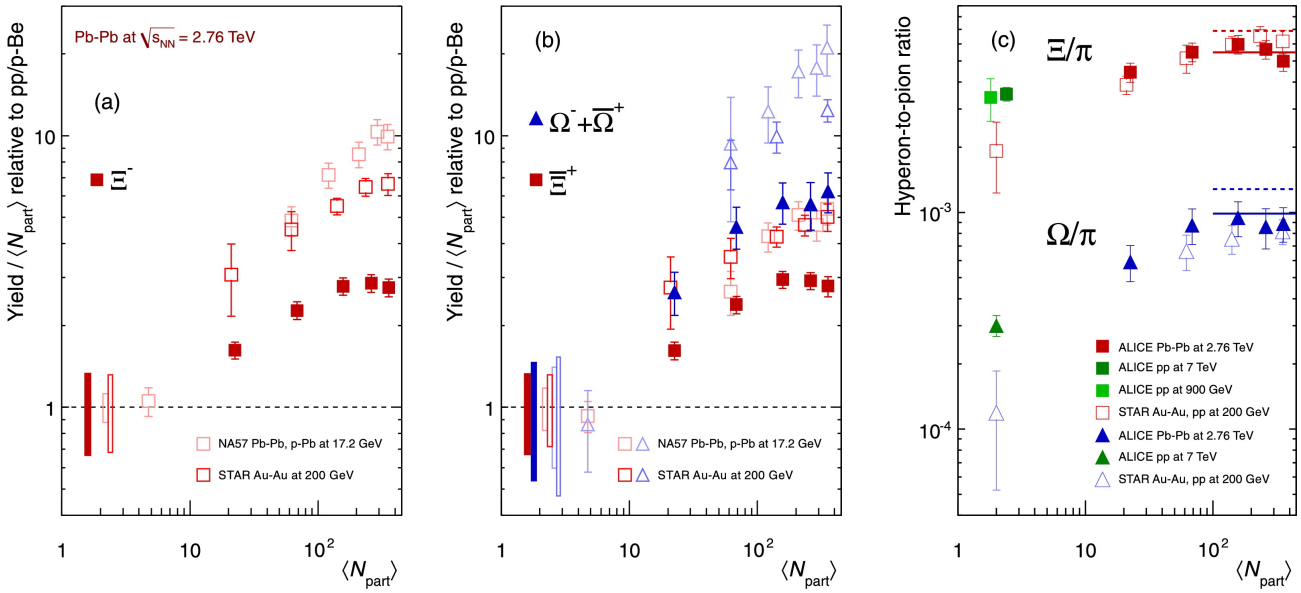

FIGURE 1.11: (*left and middle*) Enhancement of the strange baryon yields in the rapidity range $|y| < 0.5$ as a function of the number of participating nucleons in the collision ($\langle N_{\text{part}} \rangle$). The ALICE data are shown in solid symbols while the NA57 and STAR data are shown in open symbols. The boxes on the dashed line at unity indicate the statistical and systematic uncertainties on the reference value. (*right*) The hyperon-to-pion ratio as a function of $\langle N_{\text{part}} \rangle$. Figure taken from [46].

collision is expected to be heavily enhanced with respect to a pp collision,[18] since the strange production takes place all along the lifetime of the QGP where a high gluonic density is expected, and is proportional to the volume of the fireball and its duration.

Several experimental observations of this enhancement are available, notably from the Super Proton Synchrotron (SPS), the Relativistic Heavy-Ion Collider (RHIC) and from the LHC. In Fig. 1.11 results from ALICE, NA57 (at SPS) and STAR (at RHIC) are shown [46], depicting the enhancement of hyperon[19] production in heavy-ion collisions. The left and middle plots show the enhancement of the production with respect to the pp or pBe baseline as $\langle N_{\text{part}} \rangle$ increases, while the right-most plot shows the same effect with respect to the pion production.

Jet quenching

A jet is a set of energetic hadrons within a narrow region or "cone", arising from a single energetic parton and the product of its fragmentation into a parton shower, which subsequently results in a highly collimated ensemble of particles in the direction of the initial parton. This energetic parton usually comes from a hard process from which 2 partons are produced back-to-back. Naturally both partons should be roughly equally energetic, and should thus result in 2 high transverse momentum (p_{T}) back-to-back jets.

Jets can lose energy in a number of ways, among which radiative energy losses and elastic collisions which broaden the p_{T} spectrum of the jet. These energy loss mechanisms are of course much more effective if there is an extended medium

[18] There is evidence of QGP-like signatures such as strangeness enhancement in very high-energy and high multiplicity pp collisions [45], but for the scope of this thesis it is assumed that QGP is only formed in heavy-ion collisions since an *extended* medium is a fundamental requirement for QGP.

[19] Hyperons are baryons that contain one or more strange quarks but no charm, bottom or top quarks.

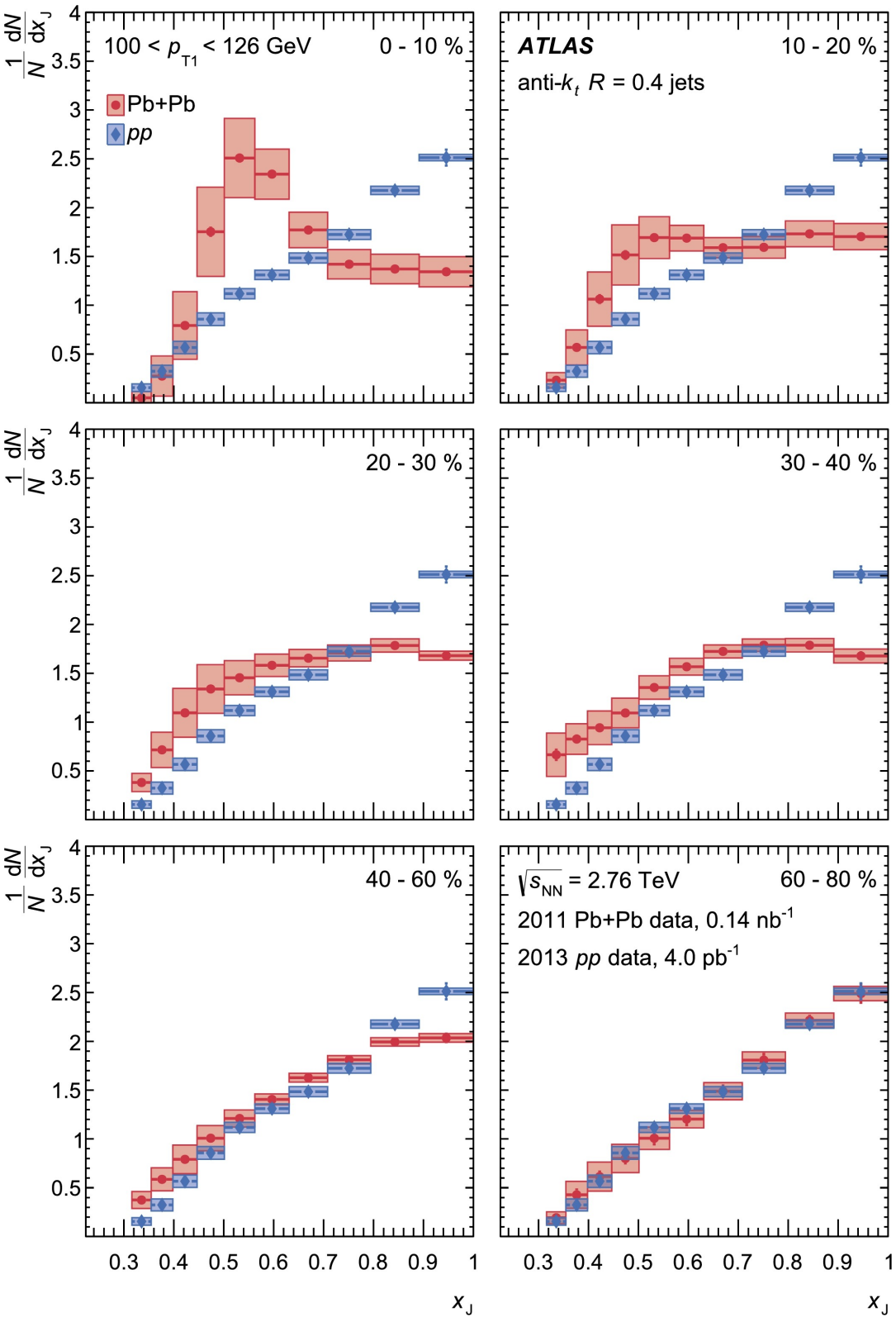

FIGURE 1.12: The normalised dijet yield distributions for jet pairs with $100 < p_{T_1} < 126$ GeV for different collision centralities. The red circles correspond to PbPb data and the blue diamonds to pp data. Figure taken from [47].

to traverse. As the parton moves it loses energy through medium-induced gluon emission (somewhat similar to Bremsstrahlung), and this depends directly on the parton momentum, the propagation length and a transport coefficient. This might explain why the energy loss at high-p_{T} increases with the centrality of the collision, and thus with the size of the partonic medium. This leads to the effect of jet quenching, when one of the two jets has to traverse a longer path inside the medium and is then detected with much lower energy than the partner jet, or in the extreme cases when one jet completely disappears while travelling through the medium, due to having lost all its energy.

As a result of these processes, the p_{T} spectra of final state hadrons and jets will be different with respect to the expected behaviour in pp collisions. More specifically, it will not coincide with what the result would be if a heavy-ion collision were a superposition of pp collisions. This change is quantified by the nuclear modification factor R_{AB} defined as[20]

$$R_{AB}(p_{\mathrm{T}}, \eta) = \frac{\frac{d^2N^{AB}}{dp_{\mathrm{T}}d\eta}}{\langle N_{\mathrm{coll}} \rangle \frac{d^2N^{pp}}{dp_{\mathrm{T}}d\eta}} \tag{1.8}$$

where A and B are the nuclear species colliding, N is the number of particles produced as function of p_{T} and η, and N_{coll} is the number of binary nucleon-nucleon collisions taking place, which will be further discussed in Sec. 4.2. R_{AB} is equal to 1 if the AB collision is a perfect superposition of pp collisions, and it is different from 1 if the medium effects modify the particle production or the particle kinematics. If $R_{AB} < 1$, one speaks of a production *suppression* and conversely, if $R_{AB} > 1$, of a production *enhancement*.

Another way to quantify the jet quenching effect is to define the p_{T} imbalance of the jets $x_{\mathrm{J}} \equiv p_{\mathrm{T}_2}/p_{\mathrm{T}_1}$, where p_{T_i} is the p_{T} of jet i.[21] Such measurement performed by the ATLAS experiment can be seen in Fig. 1.12, where each panel corresponds to a different centrality interval [47]. In the most peripheral case shown (60 - 80%), the x_{J} of the jets in PbPb coincides with the distribution found in pp collisions peaking at $x_{\mathrm{J}} = 1$, which means that in most cases the p_{T} of both jets are balanced. Conversely, as the centrality increases the PbPb jets become peaked at around $x_{\mathrm{J}} = 0.5$, deviating from the balance.

1.3 Quarkonia and QGP

Quarkonia are of particular interest when it comes to probing the QGP. This is because they have several qualities that can be exploited to infer information from the medium formed after the collision.

The $c\bar{c}$ pairs are formed at the very early stage of the collision before the formation of the QGP, and due to their high mass they are not produced thermally by the medium. This means that the $c\bar{c}$ bound states, of which the formation times are not fully known [48], will propagate through the medium and interact with it all along its evolution. The heavy flavours (c and b), being produced in small quantities cannot influence the dynamics of the medium, but there is compelling evidence that the medium does influence the fate of quarkonia. Observations of flow and quenching of

[20] η is the pseudorapidity defined as $\eta = -\ln(\tan(\theta/2))$, where θ is the angle with respect to the beam axis.

[21] Here 1 and 2 refer to the 2 highest-p_{T} jets in the event.

the open heavy flavours [49–51] and the suppression of quarkonia hint towards the fact that heavy quarks couple strongly with the medium [52].

Among the signatures given by quarkonia from possible QGP effects, the observation of how the quarkonia production is affected when the QGP is present is of particular interest. This is referred to as quarkonia suppression, which is characterised by a lower production of such states when compared to the production rate in pp collisions, where, due to the not high enough temperatures reached, no QGP is expected. Quarkonia states are thought to deconfine at larger temperatures than the QGP critical temperature [53], which offers then the possibility of evaluating the temperature of the medium above T_c.

There are several mechanisms that can contribute to the suppression of quarkonia even in the absence of QGP. These are the so called *cold nuclear matter effects* (CNM) which will be briefly described in Sec. 1.3.1. Subsequently, the sequential suppression of quarkonia, which is only present when the QGP is formed, is discussed in Sec. 1.3.2.

1.3.1 Cold nuclear matter effects

These effects play an important role in the quarkonia suppression both in the absence and presence of QGP and, as will be discussed, they are highly dependent on the kinematics of the probes. Their understanding is important to quantify and characterise the effects of QGP, since they are also present when QGP is formed, specially in the early stages of the collisions and in the final hadronisation phase. To study these effects, systems that exhibit collective behaviour but that do not result in such high multiplicity to form a QGP are needed, for example pA collisions, where A is a larger atomic nucleus.

Shadowing

The nPDFs that were already discussed, show that the gluon density within a nucleus is modified with respect to the pp baseline. This indeed affects the production of heavy quark pairs since they originate mainly from gluon fusion and are thus sensitive to the possible (anti)shadowing effects. As the collision energy increases and gluons of low x begin to saturate, the nPDFs predict a shadowing of the quarkonia production.

This effect is illustrated in Fig. 1.13 as a function of the momentum fraction carried by the parton (x) and for a given energy scale. At low momentum fraction ($x \lesssim 0.02$) the nuclear modification factor R_g is lower than one and thus represents a shadowing while at $0.02 \lesssim x \lesssim 0.3$ the R_g is greater than 1, representing an anti-shadowing effect or an enhancement. At higher momentum fraction between ~ 0.3 and ~ 0.8 there is another shadowing region due to the EMC effect, named after the European Muon Collaboration, which has different probable causes [36]. Finally, there is the region of $x \gtrsim 0.8$ where an anti-shadowing effect is predicted, called the Fermi motion region.

The gluon shadowing effect has a direct impact in the J/ψ production. This is illustrated in Fig. 1.14 where due to the gluon shadowing effect the J/ψ is suppressed ($R_{\mathrm{PbPb}} < 1$). According to the models this suppression is expected to depend on the kinematic region in which the J/ψ is produced.

Cronin effect

This effect, discovered in the 70's in pA collisions, leads to a broadening of the p_{T} spectrum of the colliding partons. This is explained by the increased number of elastic

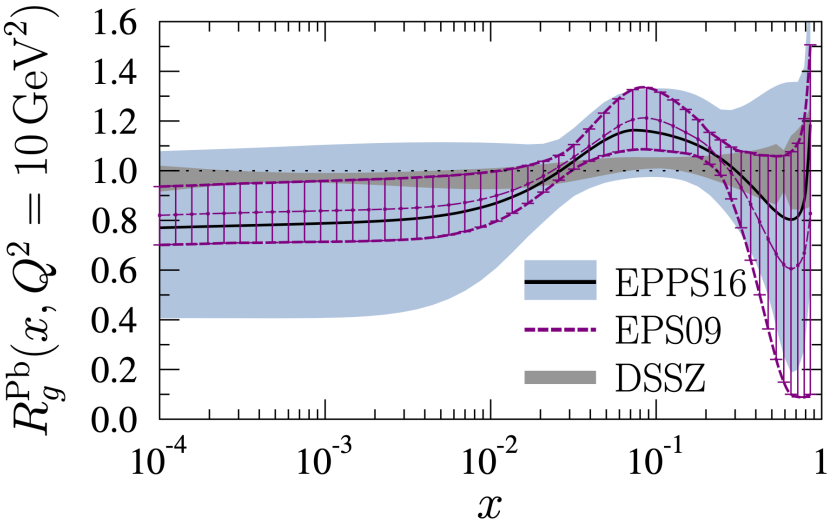

FIGURE 1.13: Nuclear modification of the gluon parton distribution function as predicted by different parametrisations for $A = 208$ and $Q^2 = 10\,\mathrm{GeV}^2$. Figure taken from [54].

scatterings that the initial partons undergo in pA (or AB) collisions as compared to pp collisions. As a result of such scatterings, the partons suffer a transverse momentum shift from lower to higher values, resulting in the observed broadening [56].

Parton energy loss

The propagation of a $q\bar{q}$ pair in the nuclear medium (nuclear matter or QGP) brings with it coherent medium-induced radiative energy losses, which result ultimately in quarkonia suppression. The modelling of these radiative losses has succeeded in describing the quarkonia suppression in pA collisions [57], as shown in Fig. 1.15, taking the leading role, hinting that the shadowing might be considered as a simple correction which might affect the magnitude but not the shape (in rapidity for example) of the suppression. In AB collisions however, even if the model cannot fully explain the observed suppression [58], the predictions still show a sizeable

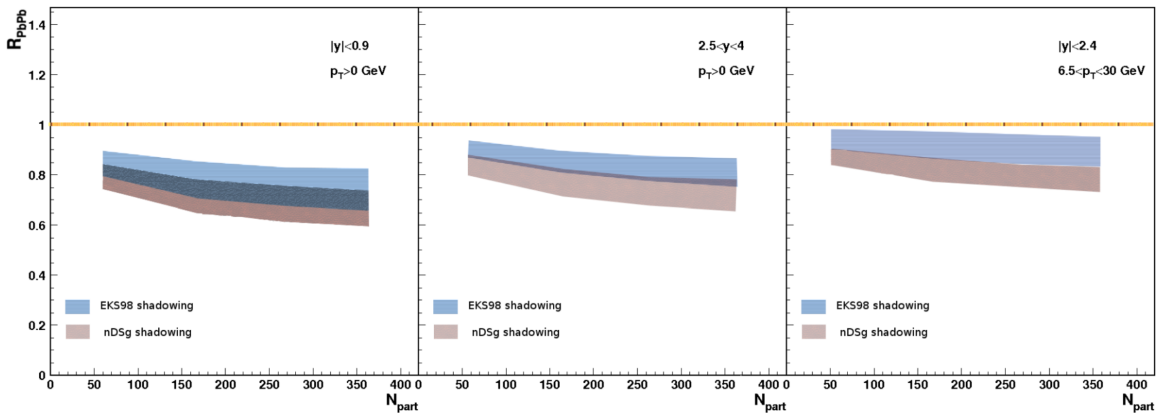

FIGURE 1.14: Shadowing effect on J/ψ suppression obtained with EKS98 (leading order) and nDSg (leading order) nPDFs in PbPb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$. Figure taken from [55].

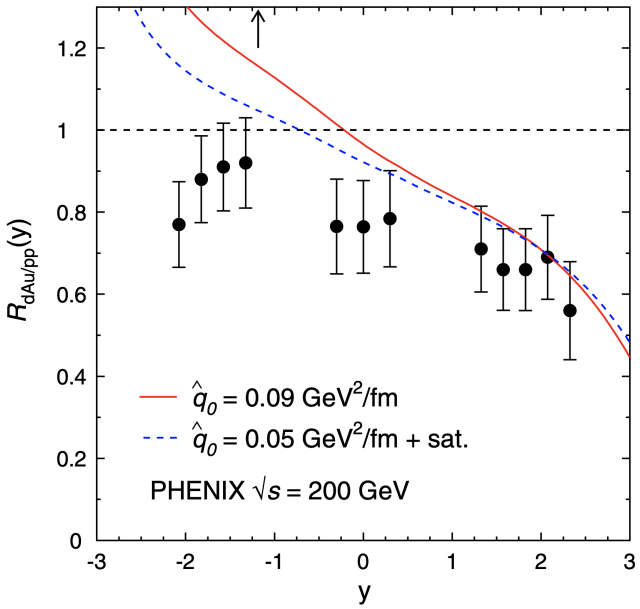

FIGURE 1.15: PHENIX J/ψ suppression data [59] in dAu collisions compared to the energy loss model. The upward-pointing arrow shows the point below which other suppression effects dominate. Figure taken from [57].

suppression from this effect alone, and thus should still be considered, specially at forward rapidities.

The upward-pointing arrow in Fig. 1.15 shows the point below which other suppression effects dominate. This point depends on the collision system and energy, the larger the system (lower the collision energy) the higher the rapidity threshold. For example in p-platinum collisions at $\sqrt{s_{\rm NN}} = 19.4\,$GeV this threshold is at $y = 1.14$ [57].

Nuclear absorption

This effect is based on the possible inelastic scattering that a quarkonium state might undergo with the colliding nuclei after its formation. In this framework, quarkonia have a survival probability which represents the probability that it will exit the nuclear matter unbroken. This survival probability is given by

$$S_{abs} = C \exp(-\rho \sigma_{abs} L), \tag{1.9}$$

where C is a normalisation constant, ρ is the nuclear density, σ_{abs} is the inelastic cross-section of a quarkonium with a nucleon and L is the mean propagation length, which depends directly on the size of the collision system. This effect has a clear dependence on the formation time of the quarkonium state under study, given that its path through nuclear matter would begin after it is formed and not when the $q\bar{q}$ is produced. The dependence of σ_{abs} with the collision energy has been studied and found to decrease rapidly as the energy increases [60], being negligible at LHC energies in collider mode.

Comovers

The Comovers model is similar to nuclear absorption in the sense that it considers the possibility that quarkonia be broken through inelastic scattering. This time however,

the scattering is not with the colliding nuclei but with the hadrons h_{co} produced that move along with the quarkonium, hence the name *comovers*. This scattering can lead to a dissociation of quarkonium through the process $q\bar{q} + h_{co} \rightarrow C + C + X$ where C is an open heavy flavour meson, and it can also lead to a recombination of quarkonium from two open mesons via $C + C \rightarrow q\bar{q} + X$. The number of comovers is directly related to the multiplicity of the collision and thus, to the centrality of the collision. This means that the quarkonia suppression due to the comovers increases as the energy density of the medium increases [61, 62].

It is worth mentioning that in the presence of QGP, an analogous process can take place with the gluons or free quarks present in the medium which can scatter inelastically with the quarkonium state, leading to its break-up. This effect, due to the large J/ψ break-up cross-section through gluon collisions and the very high gluon density in a deconfined medium, is considerably larger than the hadronic scattering and it increases with increasing energy density [63].

Before moving on to the QGP effects, it is useful to note that out of the aforementioned CNM effects, the gluon shadowing, nuclear absorption and the comovers are the ones that may have the greatest impact on the J/ψ production in PbNe collisions at a centre-of-mass energy $\sqrt{s_{\text{NN}}} = 69\,\text{GeV}$, which is the main system under study in the rest of this work. The parton energy loss effect would take place in a rapidity range that LHCb cannot probe in the case of PbNe fixed-target collisions, where only the backwards hemisphere is seen by the detector ($y^* < 0$).

1.3.2 Sequential suppression

Since the formation of the QGP requires a critical temperature $T_c \sim 175\,\text{MeV}$, and the heavy quark masses are significantly higher ($\sim 1.28\,\text{GeV}/c^2$ for c and $\sim 4.18\,\text{GeV}/c^2$ for b), the heavy quarks are produced only in the very early stages of the collision and can be used to probe the QGP. This is motivated by the fact that the QGP, assumed to have a larger formation time, should not play a role in the heavy quark yield. However, given that the QGP is a deconfined phase, the relative yield of quarkonia with respect to the open heavy mesons should be significantly modified when compared to a scenario where the QGP is not present. In fact, every quarkonium state is predicted to have a dissociation temperature $T_d > T_c$, sometimes called *screening temperature* due to the fact that the dissociation of the bound state comes from the colour screening, analogous to the Debye screening. More precisely, the binding between the $q\bar{q}$ pair depends on the strong interaction range, which decreases with the presence of colour charges in its surroundings. At the same time, the density of colour charge increases with temperature, resulting in the dissociation or "melting" of this state. Once this occurs, the now free heavy quarks hadronise most likely into an open meson state in the posterior hadronisation stage, such as a D^0, and consequently, the relative hidden-to-open yield is modified [64]. At this stage it is worth noting that the relative yield is interesting because when, for example, a $c\bar{c}$ pair is created, around 90% of the times they will form an open charm meson instead of a $c\bar{c}$ bound state [65]. So measuring the open charm yield is a way of estimating the total charm production.

Having different dissociation temperatures for the different quarkonia states gives rise to a suppression pattern which can be used as a QGP thermometer. Consider for example the J/ψ yield, which consists of prompt production, and feed down from the decay of the higher energy $c\bar{c}$ states (χ_c and ψ'). If there is no QGP, or at least the temperature is not high enough to melt any of these states and they all "survive" the QGP phase, around 10% of the J/ψ yield will come from the ψ' decay and about 30%

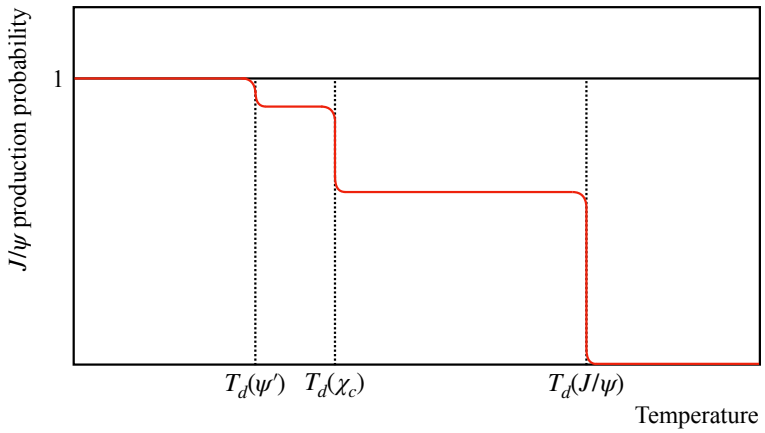

FIGURE 1.16: Illustration of the J/ψ sequential suppression as a function of the QGP temperature. $T_d(X)$ is the dissociation temperature of a state X.

will be due to the χ_c decay (the rest being direct J/ψ production). If the temperature of the QGP increases enough to be $T_d(\psi') < T_{QGP} < T_d(\chi_c)$, in the final J/ψ yield a 10% decrease would be expected. In the same way, if the temperature keeps rising to $T_d(\chi_c) < T_{QGP} < T_d(J/\psi)$, only 60% of the original expected yield would be measured. Finally if the temperature is $T_{QGP} > T_d(J/\psi)$ (see figure 1.16), all the charmonium states would have been melted into the QGP, and only open charm mesons would be detected. This step by step decrease in the yield of charmonium is the so called *sequential suppression* of charmonia, first predicted in 1986 by Matsui and Satz [53]. The experimental results related to this are discussed in Sec. 1.3.4.

1.3.3 Quarkonia recombination

Generally speaking, quarkonia are formed from a $q\bar{q}$ pair produced in the collision. When there is a collective medium present, typically where several nucleon-nucleon (NN) collision take place, it can happen that a q from one NN collision ends up binding to a $\bar{q}$ produced at another NN collision. This can lead for example to an enhanced J/ψ production at the hadronisation stage, provided that a sufficiently high charm density was reached, effectively covering the dissociative effect of the QGP.

The $c\bar{c}$ production scales as the number of binary NN collisions, while the light hadron production scales as the number of participating nucleons. This already means that in AB collisions, the relative amount of charm quarks with respect to the light flavours is higher than in pp collisions. This, together with the fact that the $c\bar{c}$ production cross-section increases with the collision energy, means that as the energy increases the probability that a c quark finds a $\bar{c}$ coming from another NN interaction becomes higher and higher with increasing energy [63], resulting in the hadronisation of an uncorrelated $c\bar{c}$ pair. Furthermore, going from the energies from the SPS, to RHIC and then to the LHC, the number of $c\bar{c}$ pairs produced per collision (in the most central heavy-ion collisions) grows from 0.1 to 10 to 100 [66], which makes the recombination a likely phenomenon as the energy increases.

A schematic representation of this effect in the J/ψ production probability is shown in Fig. 1.17. As a result this would imply that the hidden to open charm ratio increases with energy, and in the extreme cases it would lead to an overall enhancement of the J/ψ production [67].

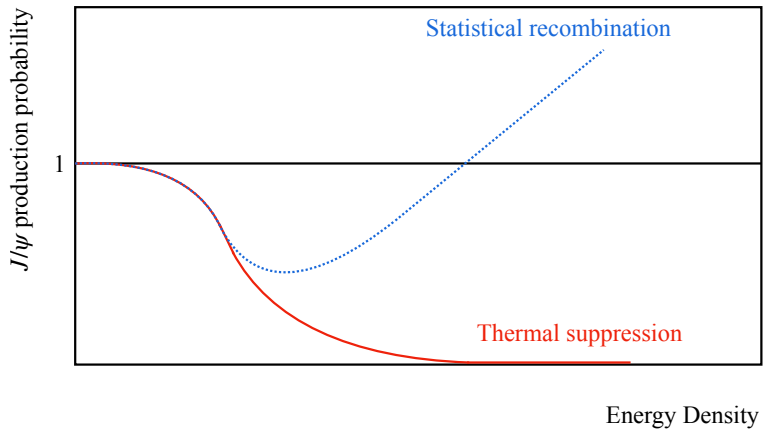

FIGURE 1.17: Illustration of the J/ψ statistical recombination in contrast to the overall suppression as a function of the energy density of the medium.

1.3.4 Experimental observations

For more than three decades there has been a massive collective effort to gather data to probe all these effects, particularly with the study of charmonia production at different energies and in different collision systems. These measurements have come mainly from experiments at the SPS, the RHIC and more recently at the LHC.

SPS results

At the SPS several experiments took place. The pioneers NA38 and NA50 were followed by the NA51 and NA60 experiments. Data were recorded from sulfur-uranium (SU) collisions (NA38) at a centre-of-mass energy of $\sqrt{s_{\mathrm{NN}}} = 19.4\,\mathrm{GeV}$ and of lead-lead (PbPb) collisions (NA50) at $\sqrt{s_{\mathrm{NN}}} = 17.3\,\mathrm{GeV}$.[22]

The observed relative charmonia production (as a ratio to Drell-Yan pair production) normalised to the expected value when normal nuclear absorption is taken into account for both J/ψ and ψ' in pA collisions, is shown in the left panel of Fig. 1.18. For the case of pA, the rates of J/ψ and ψ' are accounted for by the nuclear absorption. For the SU collisions, the J/ψ rate is still compatible with the pA results and their extrapolation to longer L. The ψ' however, exhibits some anomalous form of suppression, reaching almost a factor 5 in the most central cases. The data recorded with the larger system of PbPb, shows a significant suppression for the J/ψ, while still being compatible with the CNM suppression in the most peripheral points, indicating that this suppression is not an effect coming purely from the passage from pA to an AB system. The suppression seen in PbPb for the ψ' is compatible with the one found in SU and follows the same trend even for longer L.

The fact that for J/ψ the suppression was clear in the case of PbPb but not in the case of SU, motivated the study of different AB systems in between the two. Subsequent results from NA38, NA50 and the new NA60 confirmed the suppression, however, signs of suppression are present at centralities in which the SU system does not show this effect. These results are shown in the right panel of Fig. 1.18.

These anomalous suppression trends could be an effect of the charmonia break up due to interactions with the comovers. This would affect more easily the ψ' than the

[22] In addition, pA collisions were recorded to use as a reference.

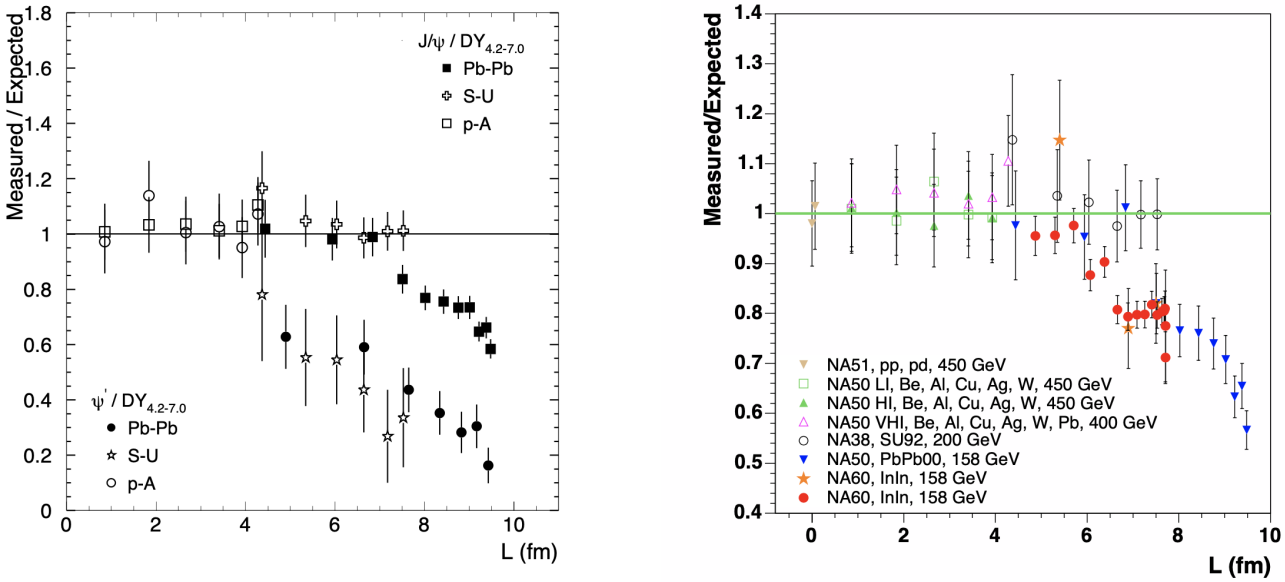

FIGURE 1.18: Relative charmonia production with respect to Drell-Yan pair production as a function of the mean propagation length L, normalised to the expected suppression from nuclear absorption. The left figure shows the results from NA38 and NA50 for J/ψ and ψ' while the figure on the right shows only the result for J/ψ and includes new results from NA38 and NA50 as well as data from NA60 and NA51. The quoted energies correspond to the incident beam energy per nucleon. Figures from [68, 69].

J/ψ due to its lower binding energy, thus suppressing the ψ' earlier. To test if the J/ψ suppression has any contribution from the sequential suppression due to the QGP, it would be necessary to measure the χ_c states, which is part of the LHCb physics agenda and of the objectives of its SMOG[23] system, since so far no experiment has been able to measure χ_c in AB collisions.

After more than a decade of data taking and careful analysis, in the year 2000 CERN announced the discovery of a new state of matter at the SPS, which features many of the characteristics of the theoretically predicted QGP [70].

RHIC results

Later on with the arrival of the results from RHIC the story became more puzzling. Collisions at RHIC occur at a centre-of-mass energy of about $\sqrt{s_{\mathrm{NN}}} = 200\,\mathrm{GeV}$, more than ten times larger than at the SPS, which should increase the energy density of the medium and thus increase as well the charmonia suppression due to the QGP effects. Surprisingly the suppression found for the J/ψ was equivalent to the one found at the SPS for mid-rapidities. The comparison of the results for the gold-gold (AuAu) collisions at RHIC and the PbPb SPS results are shown in Fig. 1.19 in the form of the nuclear modification factor R_{AA} as a function of the number of participating nucleons, which increases with the centrality of the collision.

This puzzling finding could be explained by a *suppression of the suppression*, that is, mechanisms that enhance the quarkonia production as the energy increases and thus counter the increase in suppression expected with increasing energy. The most popular candidate for this is the statistical recombination that was discussed in 1.3.3.

[23] The SMOG system of LHCb will be treated in Sec. 2.3.

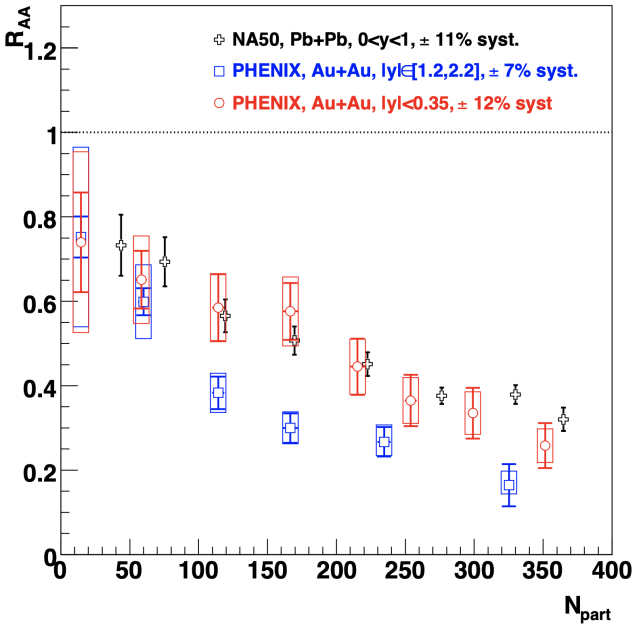

FIGURE 1.19: Nuclear modification factor of J/ψ production as a function of the mean number of participating nucleons from the PHENIX collaboration at RHIC and from NA50 at the SPS. Figure from [71].

At RHIC several aspects of the QGP were studied, ranging from the hydrodynamic flow, to the jet quenching, and the suppression of heavy-quark quarkonia.

LHC results

At the LHC, Pb beams have been made to collide at centre-of-mass energy of 2.76 and 5.02 TeV, once again pushing the energy by more than a factor ten when compared to the energy at RHIC. At these collision energies, the QGP is predicted to have an initial temperature of a few T_c, a longer lifetime and higher production rate of heavy-flavour $q\bar{q}$ pairs than at RHIC.

The ALICE detector, specifically designed for the study of heavy-ion collisions, has performed extensive measurements of quarkonia production and of collective properties of the hot medium. The J/ψ production is found to be less suppressed than what had been found by the PHENIX collaboration at RHIC, both at mid and forward rapidities. The measurement in the form of the nuclear modification factor as a function of the number of participating nucleons is shown in Fig. 1.20 for mid rapidities (left panel) and forward rapidities (right panel), where the different behaviour is apparent in the mid-central and central collisions. The suppression pattern at mid rapidities found by the ALICE collaboration might even hint at an increase with the centrality which is in agreement with what would be expected from statistical recombination.

When comparing the ALICE results from PbPb collisions at $\sqrt{s_{\rm NN}} = 2.76$ TeV and 5.02 TeV a very good agreement between the two is found for all $p_{\rm T}$ intervals where the comparison is possible, see Fig. 1.21. These results are in good agreement with transport models that include statistical recombination [75].

Another experimental observation that strengthens the recombination picture, either full statistical recombination or transport models, is the measurement of hydrodynamic flow of the J/ψ. Indeed, if the J/ψ is dissociated by the QGP, and the

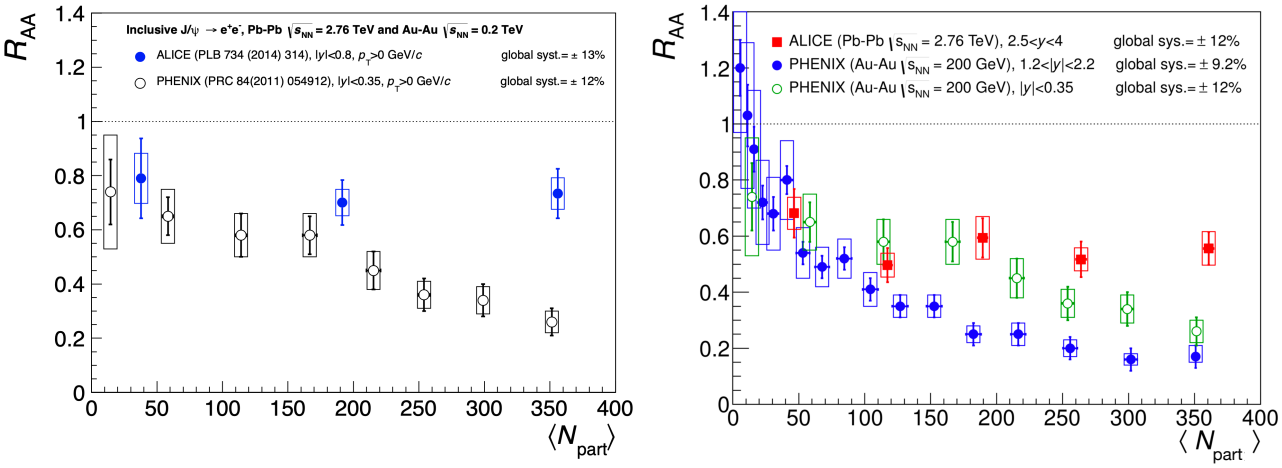

FIGURE 1.20: Nuclear modification factor of J/ψ production as a function of the mean number of participating nucleons from the ALICE collaboration at the LHC and from the PHENIX experiment at RHIC. Comparison at mid rapidity (*left*) and at forward rapidity (*right*), with PHENIX results at mid rapidity included. Figures from [72, 73].

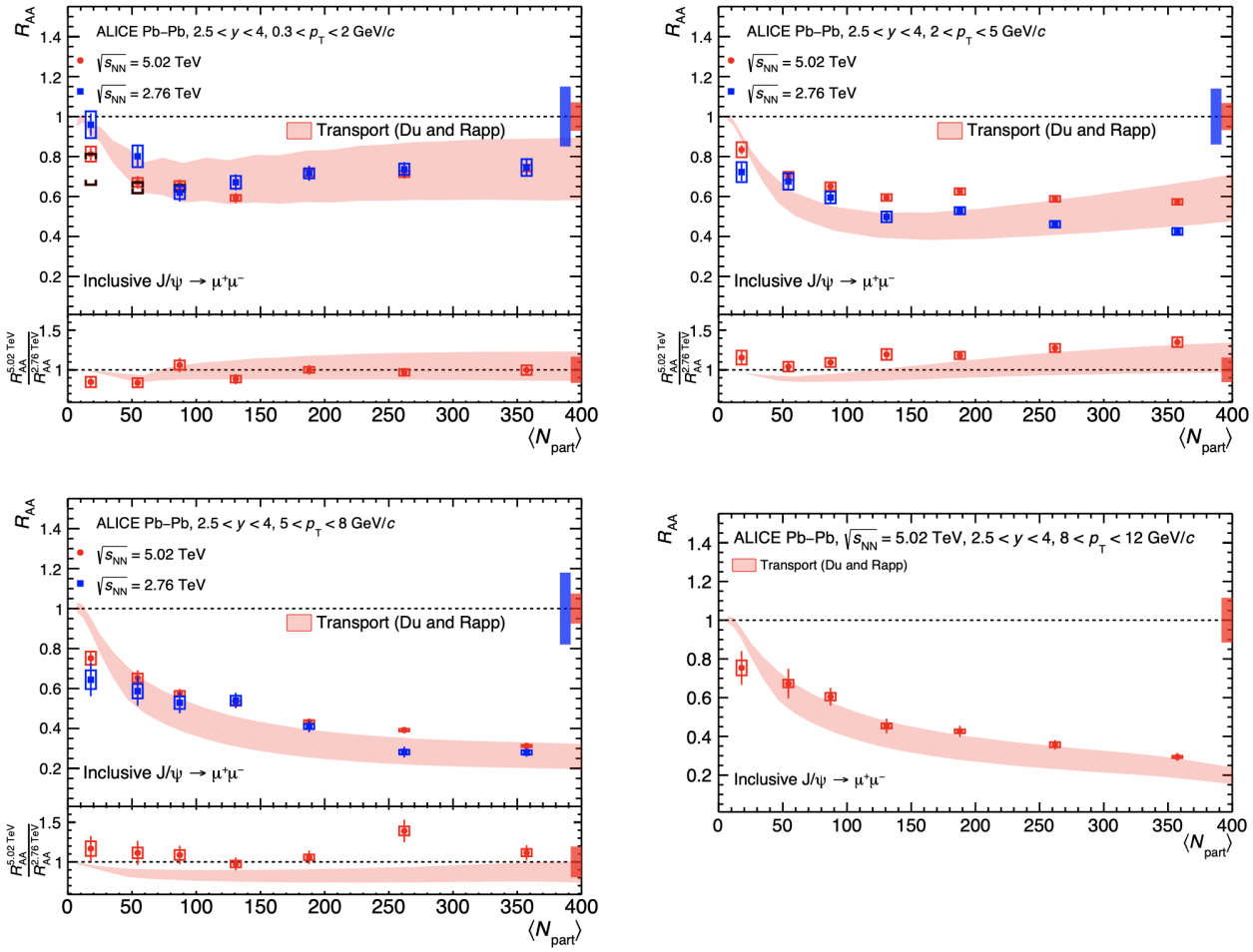

FIGURE 1.21: Nuclear modification factor of J/ψ production as a function of the mean number of participating nucleons from the ALICE collaboration at the LHC. Different panels show different p_{T} intervals. Figure from [74].

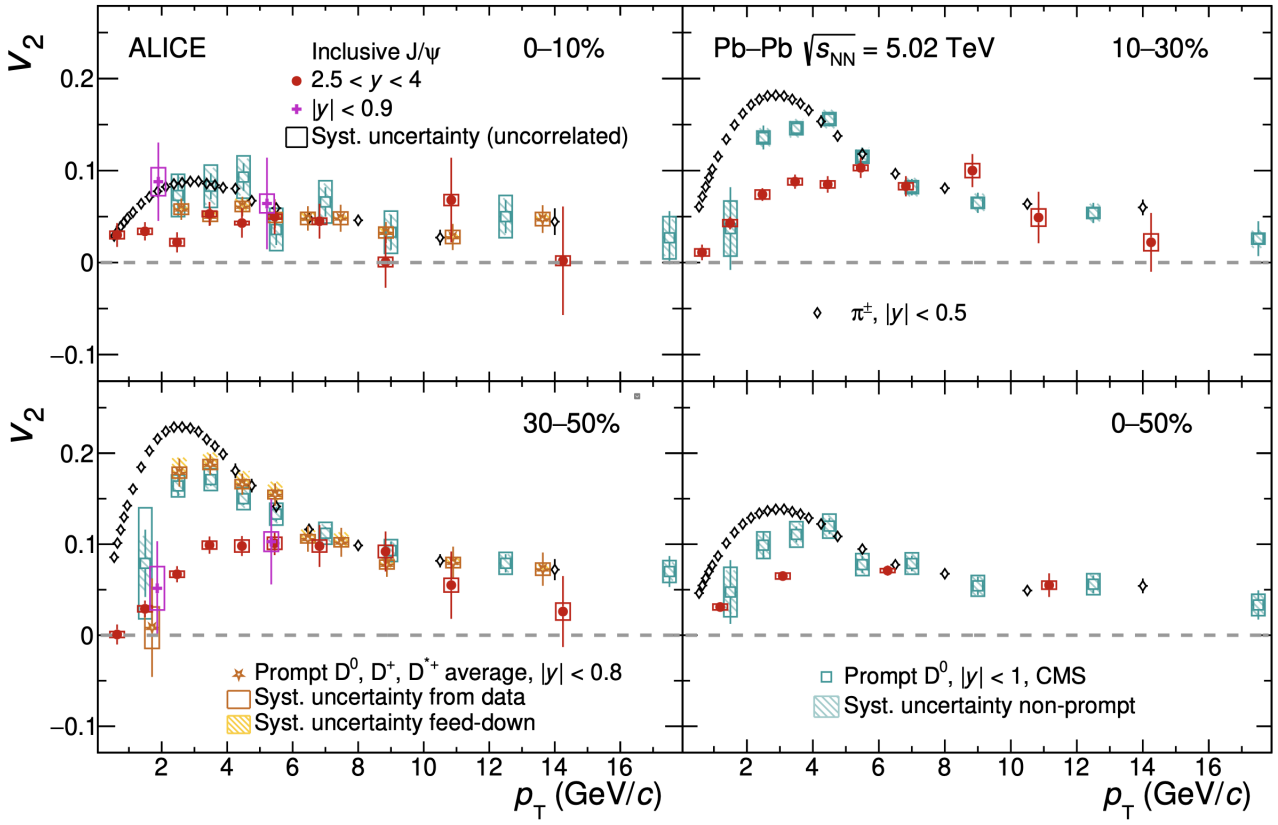

FIGURE 1.22: Inclusive J/ψ v_2 as function of p_T in different centrality intervals ($0-10\%$, $10-30\%$, $30-50\%$ and $0-50\%$) in PbPb collisions at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$. Both mid-rapidity and forward rapidity J/ψ v_2 measurements are shown. The results are compared with the v_2 coefficients at mid-rapidity for charged pions and prompt D^0 mesons. Figure from [76].

$c\bar{c}$ quarks travel freely through the deconfined medium, the charm quarks would inherit some radial and anisotropic flow from the medium. This would result in non-zero hydrodynamic flow coefficients (v_n) of the J/ψ, due to the recombined $c\bar{c}$ pairs. The ALICE collaboration has thoroughly studied flow in PbPb collisions at the LHC, and an analysis of the full Run 2 PbPb data set has found that J/ψ in non-central collisions exhibit non-zero elliptic flow (v_2) and triangular flow (v_3), the latter with a significance of 5.1σ [76]. The J/ψ v_2 measured in PbPb collisions at a centre-of-mass energy of $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ is shown in Fig. 1.22 for different centrality classes.

1.4 Quarkonia in fixed-target configuration with LHCb

The experimental scenario is still evolving and new efforts are being undertaken to further probe the QGP in colliders and to extend the current understanding of the abnormal quarkonia suppression and its true nature. The LHCb experiment has more recently entered into the world of heavy-ion collisions and has a promising future ahead.[24] In addition to its ability to exploit PbPb collisions at the LHC energy scale, it can also function as a fixed-target experiment thanks to the SMOG system, a unique

[24] This will be discussed in more detail in Sec. 2.

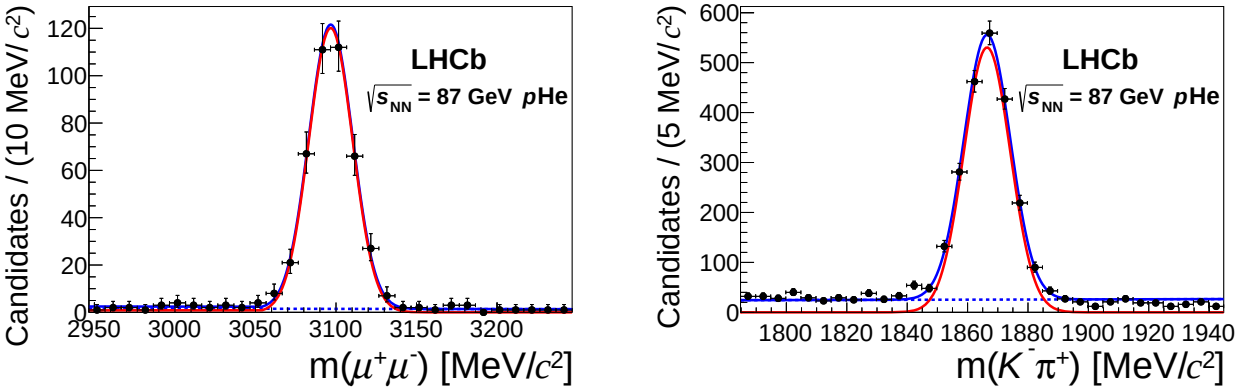

FIGURE 1.23: Invariant mass distributions for the $J/\psi \to \mu^+\mu^-$ (*left*) and $D^0 \to K^-\pi^+$ (*right*) in pHe collisions at a centre-of-mass energy $\sqrt{s_{\rm NN}} = 86.6\,\text{GeV}$. Figures from [65].

feature at the LHC, and whose main goal, in the context of the study of quarkonia, is to test the colour screening through the J/ψ sequential suppression. In this scenario, other than the possible QGP effects, the J/ψ production is expected to be affected mainly by the gluon anti-shadowing,[25] nuclear absorption and comovers.

The test of sequential suppression needs the measurements of J/ψ, ψ', χ_c and also of open charm mesons to use as reference. This must be done at sufficiently small energies so that there is no statistical recombination at play. Knowing that at the LHC energies $\sim 100\ c\bar{c}$ pairs are produced in central PbPb collisions, ~ 10 at RHIC in AuAu collisions and ~ 0.1 at SPS in PbPb collisions, the fixed-target collisions recorded by LHCb are the perfect scenario for this study, ranging from collision energies of 69 to 110 GeV, where only $\sim 1\ c\bar{c}$ pair is expected to be produced in central AB collisions, leading to a possible observation of the sequential suppression of J/ψ from the QGP without the statistical recombination effects. Furthermore, the fixed-target programme of LHCb could also contribute in covering the energy gap between 62.4 GeV and 130 GeV of the RHIC energy-scan programme and could serve to complement and to cross-check their results.

LHCb will take advantage of the target versatility of SMOG2 to thoroughly study the CNM effects using various targets, such as helium, neon, argon and other nobles gasses. At the collision energies accessible in fixed-target mode, the large fully-instrumented acceptance in the forward rapidity allows for an almost full coverage of the backward hemisphere in the centre-of-mass frame. LHCb has already shown its incredible capabilities in measuring open and hidden charm in pHe and pAr collisions [65]. The signal extraction from the invariant mass distribution can be seen in Fig. 1.23 for the $J/\psi \to \mu^+\mu^-$ (left) and $D^0 \to K^-\pi^+$ (right). The LHCb experiment, with these measurements, has already placed itself as an important contributor in this field, and has shown that it can reach unprecedented precision compared to other experiments, as it is evidenced by the measurement of the J/ψ cross-section and the $c\bar{c}$ cross-section shown in Fig. 1.24 from the same study.

The undeniably great performance of LHCb in fixed-target mode, together with the foreseeable increase in recorded data for the future LHC runs, may offer the first opportunity to measure all charmed hadrons, particularly the ψ' and χ_c, which should provide full control over the study of suppression mechanisms of charmonia.

[25] In its fixed-target configuration, LHCb is sensitive roughly to the target x between 0.02 and 0.3, where an anti-shadowing effect is expected. See Fig. 1.13.

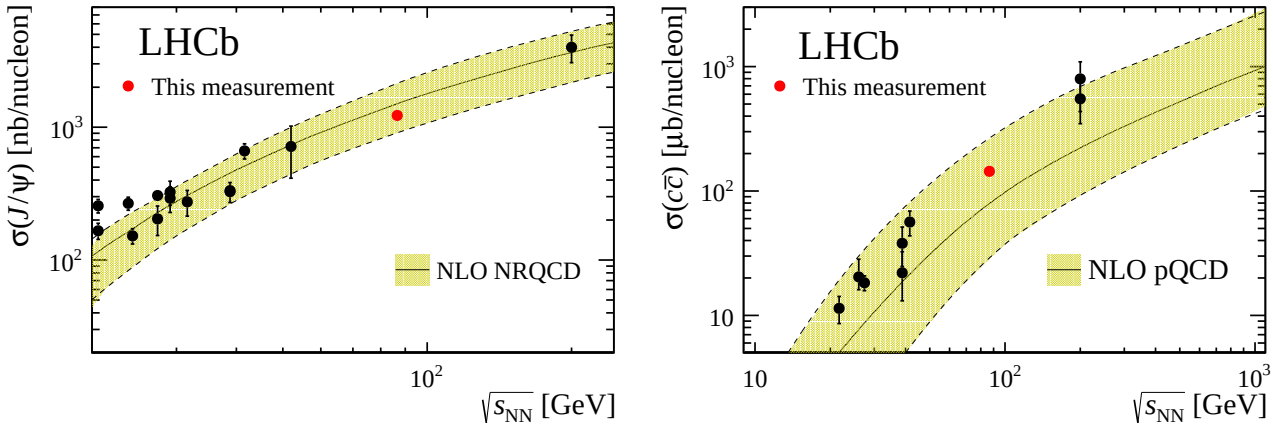

FIGURE 1.24: Cross-section measurements as a function of the center-of-mass energy for J/ψ (*left*) and for $c\bar{c}$ (*right*). Experimental data, represented by black points, are taken from other experiments (see details in Ref. [65]). The band corresponds to a fit based on NLO NRQCD calculations. Red points correspond to the pHe results from the LHCb measurement. Figures from [77].

This thesis finds its place in this experimental challenge, and presents the first measurement of hidden and open charm in AB collisions with the LHCb detector in its fixed-target configuration. This thesis opens this new era.

Chapter 2

The LHCb experiment

In the first part of this chapter, a brief description of the accelerator complex where the LHC is located, and its functioning, is illustrated. Then, the LHCb experiment is presented in detail with its particular features that make it a unique experiment in the world's largest particle collider. Finally, the SMOG, a very important part of LHCb for this thesis, is shown and described.

2.1 The LHC at CERN

Across the French and Swiss border, between the lake Léman and the Jura mountains in the Geneva area, lies the European Council for Nuclear Research (CERN, for its name in French). CERN is a scientific institution that has been at the forefront of fundamental research since it was founded in 1952.

At present, CERN harbours the largest and highest-energy particle-collider ever made, the LHC [78]. It makes use of the tunnel that once hosted the large electron-positron (LEP) collider, featuring a length of 26.7 km at a depth of 100 m underground.

The LHC began operations in 2008 with inaugural tests, which after a 9-day period came to an abrupt end when an electrical fault lead to a loss of 6 tonnes of liquid helium, which expanded explosively, damaging several magnets and portions of the machine [79]. It was not until late 2009 that the LHC was able to function again and only in 2010 the first physics run took place. The LHC first operational run saw proton beams ramp-up to 3.5 TeV giving rise to proton-proton (pp) collisions at 7 TeV, already the highest energy collision ever recorded in a man-made collider. In the following years during Run 1 (2009 - 2013) and Run 2 (2015 - 2018) the LHC has continued to improve and has provided pp collisions of up to 13 TeV. In addition to pp collisions, the LHC is also capable of accelerating Pb ions, and has been doing so for short periods since 2010, providing PbPb and pPb collisions.

The LHC is made up of 8 straight and 8 curved sections making one full loop. In the curved sections there are the bending magnets used to steer the trajectories of the beams. The straight sections on the other hand, denoted with P1 to P8, are used by experiments and for machine instrumentation and operation. The LHC brings the beams to collision in points P1, P2, P5 and P8 where, respectively, the experiments ATLAS [80], ALICE [81], CMS [82] and LHCb [83] record the collisions.

The LHC capabilities are a product of the state-of-the-art technologies that were used for its construction and subsequent upgrades. It features superconducting radio-frequency (RF) cavities which generate electric potentials of up to 2 MV each, focusing quadrupole and octupole magnets that keep the beams focused, 8 T dipole superconducting magnets to keep the beams in orbit and a superfluid helium cooling system that maintains the 1.9 K necessary for the functioning of the magnets. Thanks to this, the LHC is able to reach record-breaking energies for the collisions. For Run 3,

which is scheduled to begin in spring 2022, it is expected that the pp collisions will reach the design centre-of-mass energy of 14 TeV.

The particle beams need to go through a series of steps before entering the LHC and reaching the goal energy. The accelerator complex at CERN, shown in Fig. 2.1, provides all the necessary machines that are used to build and power the energy beams. The protons begin their journey at the linear accelerator (LINAC)[1] where they are accelerated to 50 MeV. Then they enter the proton synchrotron booster (PSB) reaching 1.4 GeV before being transferred to the proton synchrotron (PS) which further increases their energy to 26 GeV. The last step before arriving to the LHC is the SPS, which accelerates the protons to 450 GeV. Finally the beams are transferred to the LHC, forming 2 beams circulating in opposite directions in separate beam-pipes, and are further accelerated to the desired final energy. The ion accelerator chain starts at the LINAC3, a dedicated linear accelerator for ions, which then feeds the ions to the low-energy ion ring (LEIR) in which the beams are stored and accumulated before being passed on to the PS, from where the road of the ion beams to the LHC is the same as for the protons.

The RF cavities used by the LHC are operated at roughly 400 MHz. They accelerate the beams and confine the charged particles into RF buckets. The harmonic number[2] gives rise to 35640 RF buckets of 2.5 ns length of which nominally only 1 every 10 buckets is filled. The result is potentially filled bunches which span 10 RF buckets each and are numbered with a bunch crossing identifier (BCID) that goes from 1 to 3564.

Another important quantity in any collider is the instantaneous luminosity $\mathcal{L}(t)$ that it can deliver, which tells the number of collisions that take place as a function of time. There are different ways to increase the instantaneous luminosity, and at the LHC the increase in luminosity has been achieved thanks to increasing the bunch populations, increasing the number of circulating bunches, *i.e.*, the number of filled bunches and by reducing the β^*parameter, which is related to the transverse size of the beam in the interaction region. The nominal peak instantaneous luminosity of the LHC is $1 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

2.2 The LHCb detector

The Large Hadron Collider beauty (LHCb) experiment [83] is one of the main experiments at the LHC. LHCb was conceived to search for physics beyond the standard model by performing precise measurements of rare and CP-violating processes in heavy-flavour decays. The LHCb physics reach however, goes much beyond these topics and has contributed in many areas such as heavy baryon spectroscopy, discovery of tetraquark bound states, study of heavy-ion collisions and ultimately the study of fixed-target collisions, the latter being a unique feature among the LHC experiments.

The LHCb spectrometer has a particular geometry. The layout of the LHCb spectrometer can be seen in Fig. 2.2 as well as its sub-detectors. It covers the angular region where most $\overline{b}b$ pairs are produced, see Fig. 2.3. In the left plot the angular region covered by the LHCb spectrometer is highlighted in red, encompassing one of

[1] LINAC 4 in Fig. 2.1. It superseded LINAC2 which was used at CERN from 1978 until the end of the LHC Run 2 in 2018.

[2] The harmonic number is the proportionality constant between the RF frequency f_{RF} and the revolution frequency in the ring f_{rev} (the number of times per second a particle orbits the accelerator). f_{RF} must be an integer multiple of f_{rev} as $f_{RF} = h f_{rev}$, where h is the harmonic number.

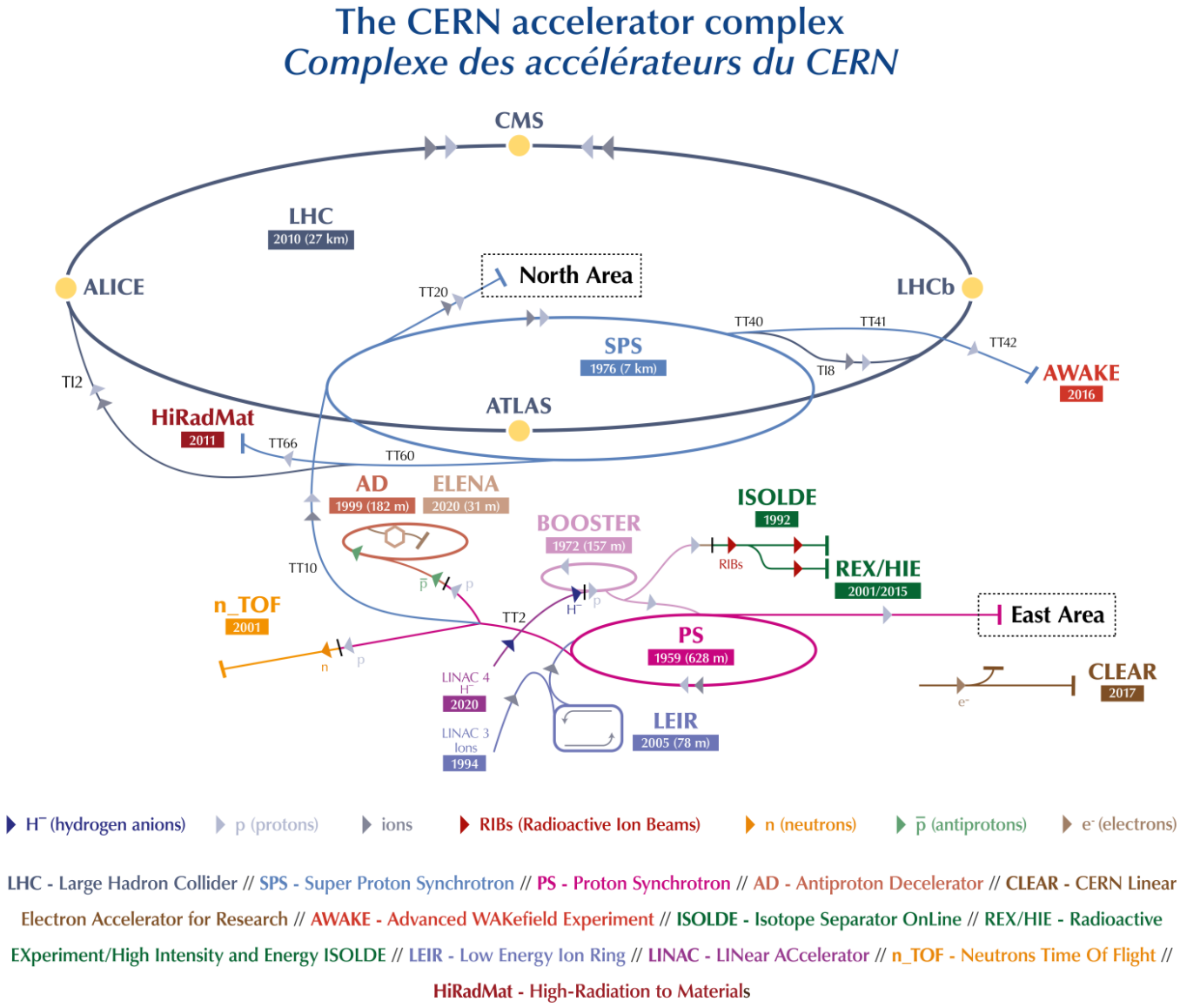

FIGURE 2.1: The CERN accelerator complex. The different accelerators that participate in the generation and ramp-up of the LHC beams are shown as well as other experiments that make use of particle beams. Figure taken from Ref. [84].

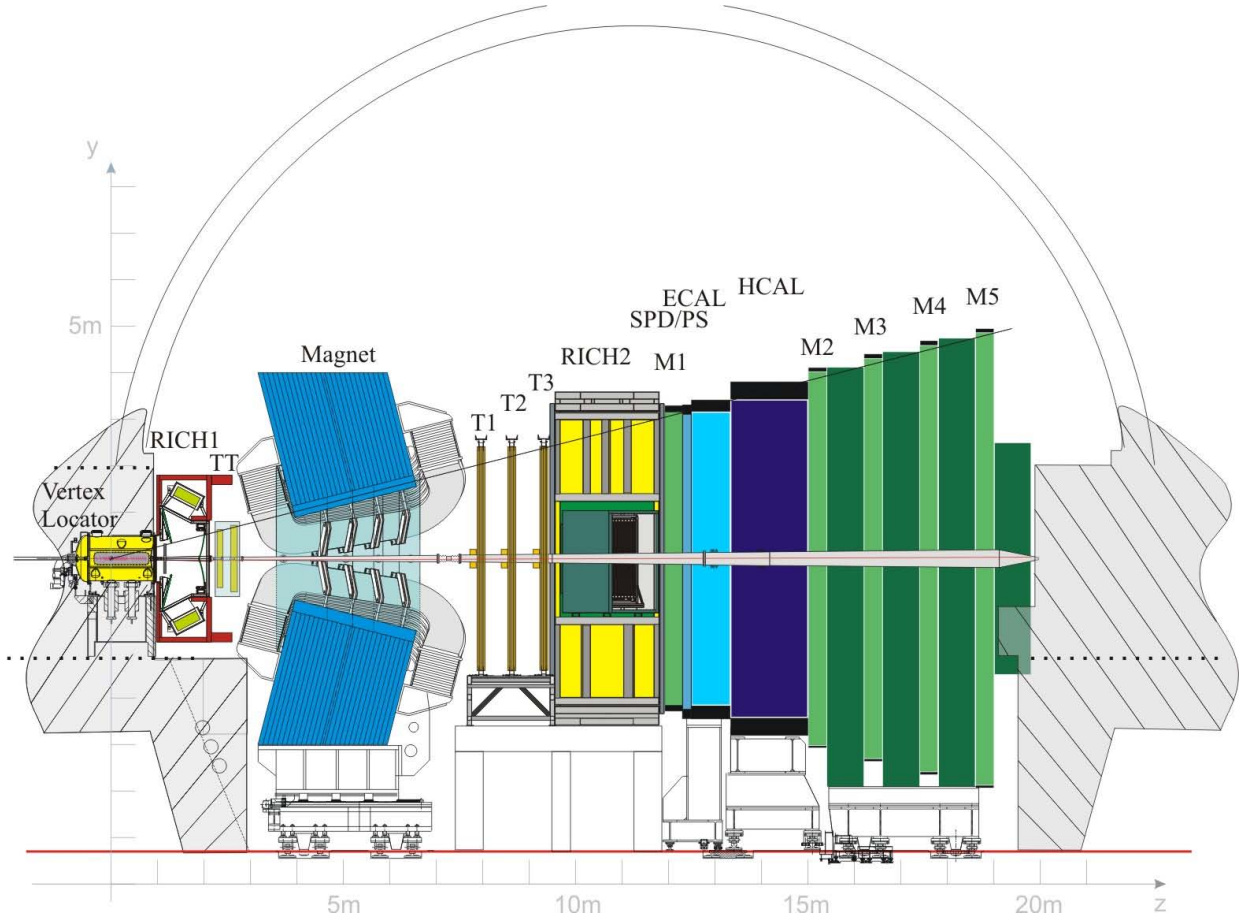

FIGURE 2.2: Lateral view of the LHCb detector layout. The collisions take place within the vertex locator (VELO) vessel. Figure taken from Ref. [86].

the two peaks in the production. In the plot on the right, the same effect is shown in units of pseudorapidity η. The red box shows the LHCb acceptance and the yellow box shows a typical general-purpose detector (GPD) acceptance, which is the case of most other experiments at the LHC. This shows that $\bar{b}b$ pairs are very likely to be produced with both quarks either in the very forward direction (or very backward), and due to the large production cross-section in pp collisions [85], LHCb is an optimal spectrometer for heavy-flavour studies.

The LHCb detector is a single-arm forward spectrometer covering the pseudorapidity range $\eta \in [2,5]$. LHCb has extremely good tracking capabilities. The tracking system is composed by the vertex locator (VELO), the only movable component of the LHC while in operation, the tracker turicensis (TT) and the tracking stations T1, T2 and T3. These, with the aid of a dipole bending magnet provide very precise momentum and charge measurements. Two ring imaging Cherenkov (RICH) detectors are responsible for the particle identification of charged hadrons. Photons, electrons and neutral hadrons are identified by a calorimeter system consisting of scintillating-pad (SPD) and pre-shower (PRS) detectors, an electromagnetic (ECAL) and a hadronic calorimeter (HCAL). Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers. The online event selection is performed by a trigger, which consists of a hardware stage, based on information from the calorimeter and muon systems, followed by a software stage, which applies a full event reconstruction.

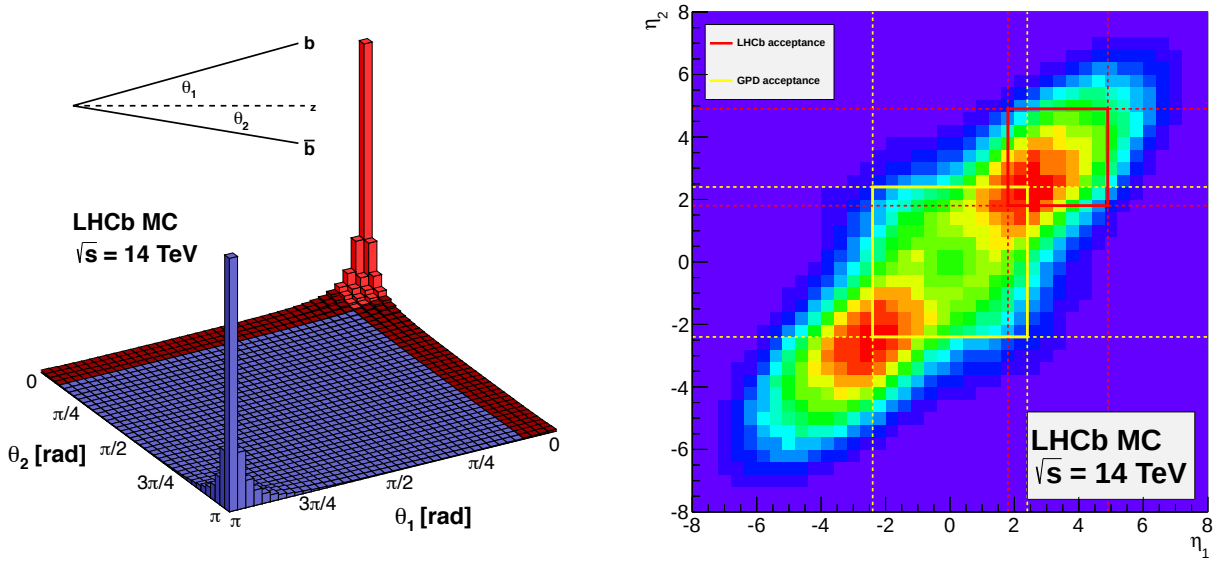

FIGURE 2.3: (*left*) Simulated $\bar{b}b$ production angles in pp collisions at $\sqrt{s} = 14\,\text{TeV}$. The red bars show the LHCb detector acceptance. (*right*) The $\bar{b}b$ pseudorapidity distribution. The yellow box represents the acceptance of a general-purpose detector and the red box the acceptance of the LHCb detector. Figures taken from Ref. [87].

2.2.1 The tracking system

The LHCb tracking system has as a starting point the VELO [88] which envelops the nominal interaction point (IP) of the beams. It is a detector which consists of two halves where a series of 21 half-disks are put in place in each half, making up the 21 modules. Each half module is a silicon strip detector with the strips arranged in such a way that they provide a radial measurement (R-sensor) and an azimuthal measurement (Φ-sensor) of the tracks hits. In order to guarantee a full azimuthal coverage, each half module spans 182° to have a small overlap and no acceptance gap. In addition to the 21 modules, there are 2 other upstream modules of only R-sensors, located opposite the rest of the LHCb spectrometer. These 2 extra modules are used to veto certain events and are called the pile-up (PU) stations. The two halves are normally about 6 cm apart to avoid unwanted interactions with the circulating beams, and only move in closer to the fully closed position when the beams are stable, approaching the beams to 8 mm distance. All of this makes the VELO a perfect detector to reconstruct displaced vertices of charm and beauty hadron decays, which occur very close to the primary vertex (PV). For a typical event, the spatial resolution in the beam direction is of 71 μm and in the transverse plane of 13 μm [89]. A schematic view of the VELO can be seen in Fig. 2.4.

After the VELO, the tracking system makes use of two other types of tracker; the silicon inner tracker (IT) [90] and the gas-tube outer tracker (OT) [91]. These trackers are arranged in four stations, the first one being the TT, which is located before the bending magnet and after the RICH 1 detector. It is made up of only IT modules, given its small dimension. The stations T1, T2 and T3 are located after the bending magnet and before the RICH 2 detector. These are composed of IT modules in the region closer to the beam-pipe, and OT modules in the regions farther from the beam-pipe. This is motivated by the fact that the detector occupancy is about 20 times

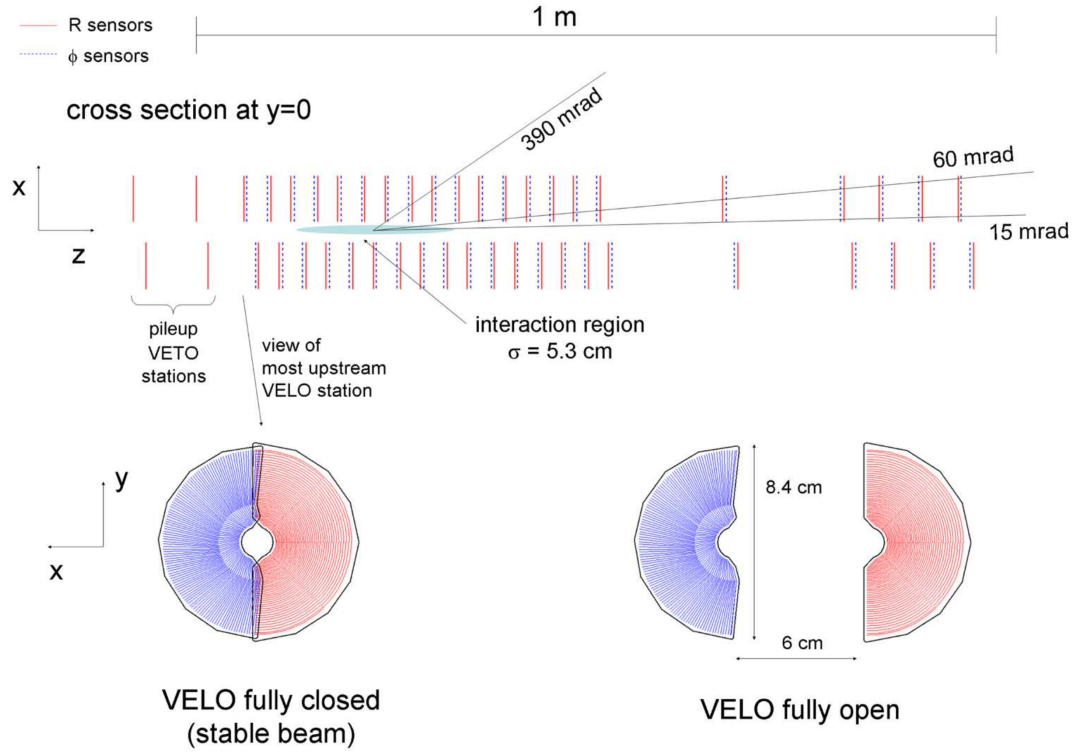

FIGURE 2.4: (*top*) The layout of the VELO detector. (*bottom*) The VELO sensors in the (*left*) closed and (*right*) open configurations. Figure taken from Ref. [83].

higher in the region occupied by the IT than in the region covered by the OT, due to the proximity to the beam-pipe. This justifies the choice of different technologies for different regions since the IT can achieve a much finer granularity and has a better radiation hardness. The tracking stations are made up of four layers with the strips vertically arranged in the first and last layers, while the two middle layers are tilted by $\pm 5^\circ$ to provide vertical information of the detected hits. In Fig. 2.5 one of the tilted layers of the TT can be seen.

The LHCb detector makes use of these tracking detectors to reconstruct tracks. The reconstructed tracks in a typical event at LHCb are shown in Fig. 2.6 as well as the hits used to reconstruct them in the different tracking stations. The track momentum is determined from the curvature generated by the magnetic field between the TT and the stations T1, T2 and T3. This field is the result of a dipole magnet made up of two trapezoidal coils placed symmetrically on the top and bottom with a bending power of 4 Tm. These parts working together provide a measurement of the momentum of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV/c. The minimum distance of a track to a PV, is measured with a resolution of $(15 + 29/p_{\mathrm{T}})$ µm, where p_{T} is the component of the momentum transverse to the beam, in GeV/c. In order to maintain the good track reconstruction at LHCb, the luminosity is levelled during the LHC fills by keeping a transverse displacement of the beams, ensuring that there is only about 1 collision per bunch-crossing. During Run 2, LHCb operated at a nominal average luminosity of $4 \times 10^{32}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$.

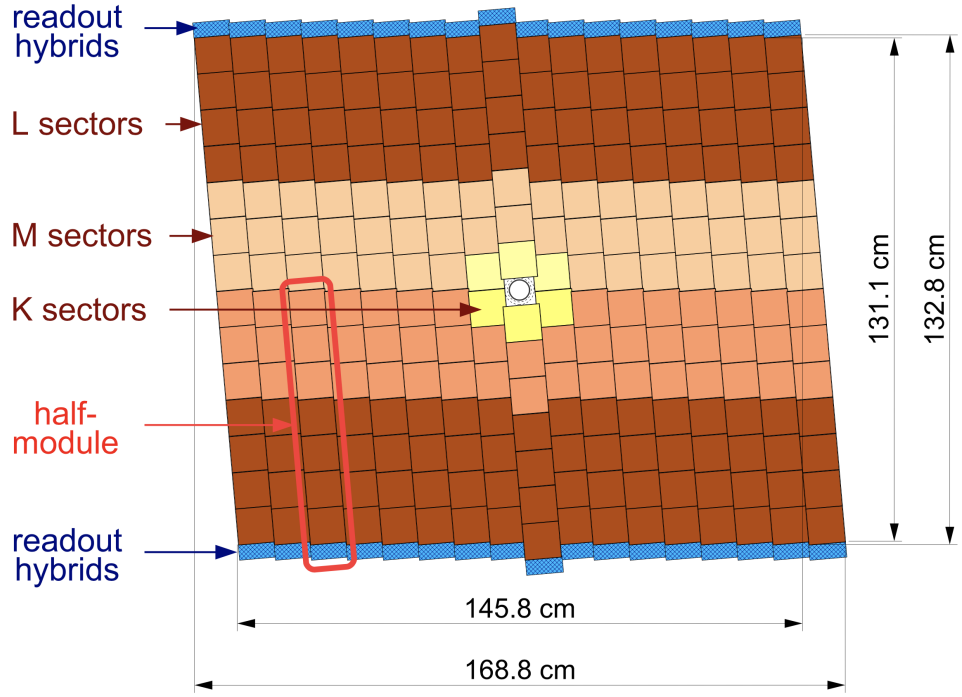

FIGURE 2.5: Layout of one of the TT detection layers. The different read-out regions are shown in different colours. Figure taken from Ref. [83].

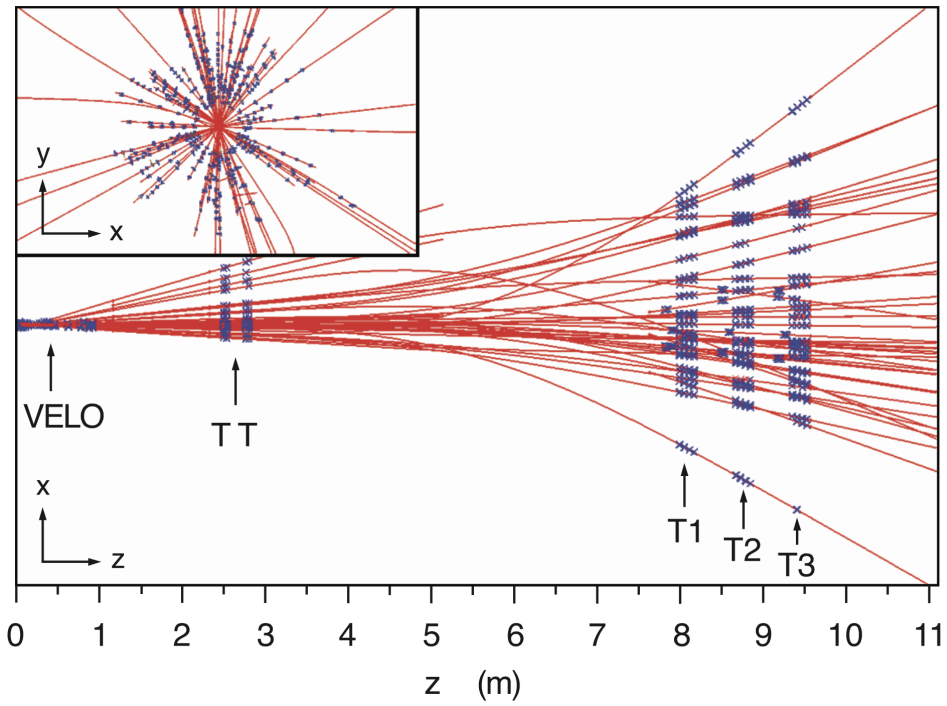

FIGURE 2.6: Display of reconstructed tracks in an event and the corresponding assigned hits in the different tracking stations. The inset shows a close-up inside the VELO region in the (x, y) plane. Figure taken from Ref. [83].

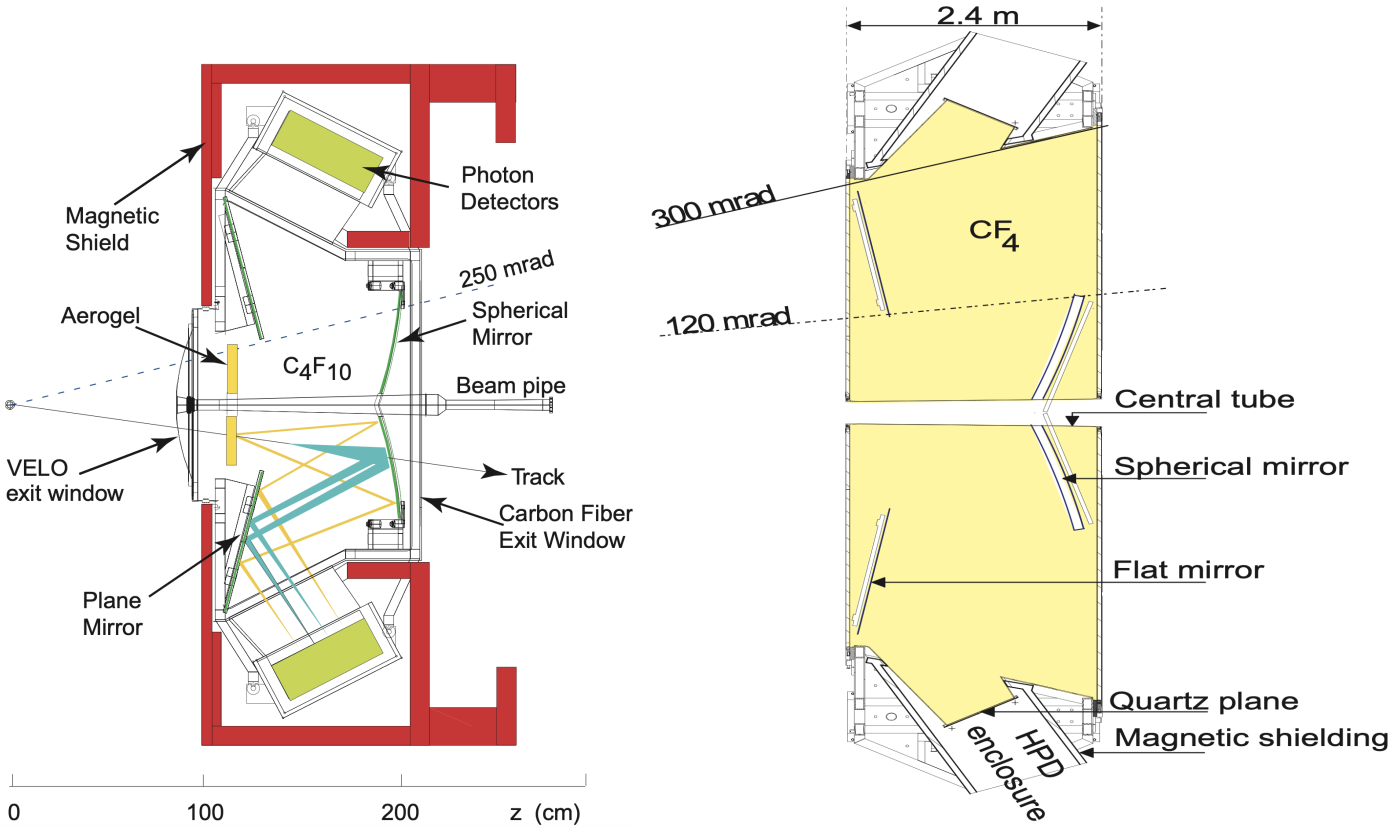

FIGURE 2.7: (*left*) RICH 1 and (*right*) RICH 2. Figures taken from Ref. [83].

2.2.2 The particle identification

For the particle identification of charged hadrons, LHCb has two different RICH detectors; RICH 1, located right after the VELO and RICH 2, located after the T3 tracking station [92]. These detectors are aimed at different momentum ranges, and they both make use of the Cherenkov light emitted by charged particles moving faster than the speed of light in the material medium. The angle of the emitted light with respect to the particle momentum is related to the particle speed as

$$\cos\left(\theta_C\right) = \frac{1}{n\beta}\,, \tag{2.1}$$

where n is the refractive index of the medium and β is the particle speed in natural units. This, in combination to the momentum measured by the tracking system allows to measure the mass of the particles, and thus their identification.

The RICH 1 detector uses C_4F_{10} as the radiator,[3] and performs well in identifying charged hadrons with a momentum of $p \in [2, 60]$ GeV/c covering an acceptance angle from 25 to 250 mrad in the vertical plane and to 300 mrad in the horizontal plane. The RICH 2 on the other hand has a smaller acceptance from 15 to 120 mrad horizontally and to 100 mrad vertically, and it is thus focused on the region where high momentum tracks, going up to 100 GeV, dominate. The layout of the two RICH detectors can be seen in Fig. 2.7.

The reconstructed Cherenkov angles for charged tracks are shown in the left plot of Fig. 2.8 as a function of the track momentum. There, clear bands can be seen which

[3] During Run 1 the RICH 1 had in addition silica aerogel as a radiator, but was removed for Run 2

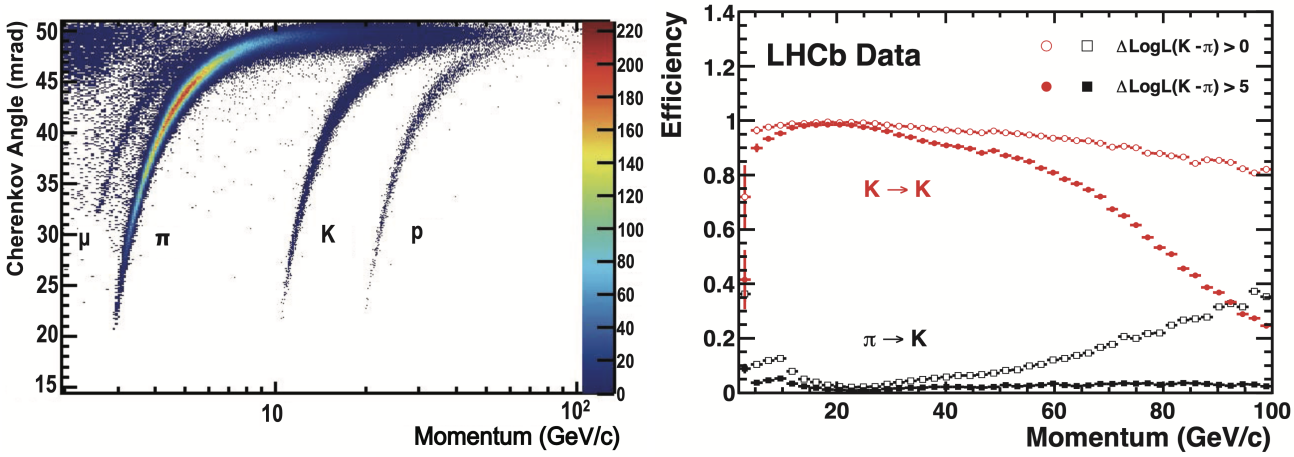

FIGURE 2.8: (*left*) Reconstructed Cherenkov angle as a function of the track momentum in the RICH 1. (*right*) The kaon identification efficiency and the pion miss-identification rate as a function of track momentum. Figures taken from Ref. [93].

correspond mainly to pions, kaons and protons. The kaon identification power as well as the pion miss-identification rate are shown in the right plot of Fig. 2.8. For the D^0 analysis that will be discussed in Sec. 5 it is particularly important to have a good identification capacity of kaons and pions, since the D^0 is studied in the $K^+\pi^-$ decay channel (and the charged conjugate decay).

2.2.3 The calorimeter system

The LHCb calorimeter system [94] consists of the SPD, PRS, ECAL and HCAL. These are used to measure precisely the position and the energy deposited by hadrons, electrons and photons and select high transverse energy candidates for the first trigger level (L0).

The ECAL is a sampling detector made up of 70 layers each consisting of a 2 mm-thick Pb sheet and a 4 mm-thick scintillator plate, whose light is collected by wavelength shifting (WLS) fibres through 5952 channels. The ECAL covers an area of about 50 m^2, and it provides the adequate granularity and resolution to reconstruct photons and neutral pions.

The HCAL is also a sampling device which exhibits 16 mm-thick iron plates and 4 mm-thick scintillating tiles, which are read-out by WLS fibres through 1468 channels. The peculiarity of the HCAL is the orientation of the iron and scintillating plates which are positioned parallel to the beam pipe.

The density of hits varies greatly on the calorimeters surface, almost 2 orders of magnitude between the central and outer regions. To compensate for this effect, different segmentations are chosen depending on the region they cover. For the ECAL three different segmentations are used with the most stringent demand being the resolution of the two showers of high-energy π^0 which decay to two photons. To cope with this, the cell size in the innermost region of the ECAL is close to the Molière radius,[4] like this, most of the shower energy is contained in about 4 cells. For the HCAL on the other hand, given the dimensions of the hadronic showers, and the performance needed, only two different segmentations are used. A quarter of the

[4]The Molière radius is the radius of a cylinder containing on average 90% of the shower energy deposit.

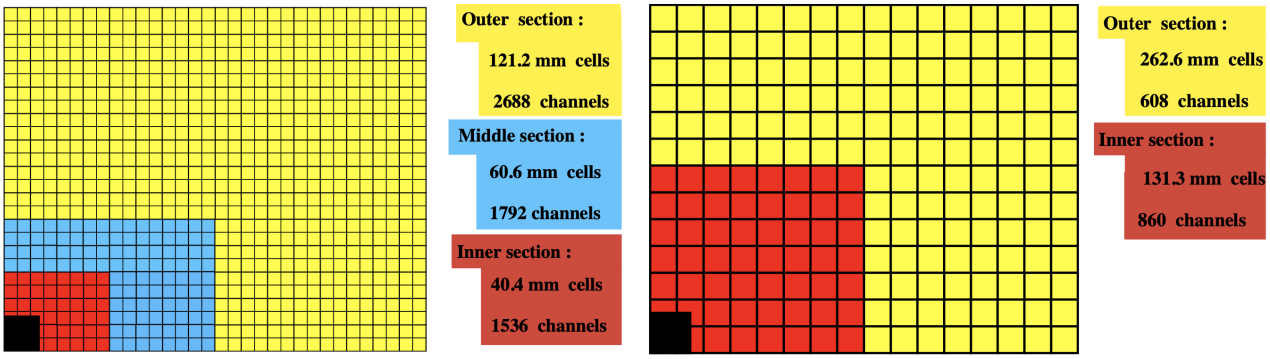

FIGURE 2.9: Segmentation of the (*left*) ECAL and (*right*) HCAL. Only the top-right quarter is shown for both detectors. The cell dimensions and read-out channels are shown for every segmentation region. Figures taken from Ref. [83].

front-face of the ECAL and HCAL can be seen in Fig. 2.9, with the different sections shown, intended for different levels of occupancy. The black regions represent the passage of the beam.

The SPD and the PRS are 15 mm-thick scintillator pads located before the ECAL and are separated by a 12 mm-thick Pb wall with the SPD being the first one. These detectors are used to provide an improved separation of electrons and photons at the SPD level, and the reduction in the charged-pion background at the PRS level. Their cell-layout is the same as the ECAL, with a slightly smaller cell size, resulting in 5952 channels for each detector.

2.2.4 The muon system

Lastly there is the muon system [95]. It consists of 5 stations denoted M1 to M5, covering in total an area of about 435 m^2, and it provides muon identification. The M1 station is located before the calorimeters, to improve the p_T measurements, and the stations M2 to M5 are located downstream the HCAL, being the last sub-detectors of LHCb. Each station is of rectangular shape and is mainly composed of multi-wire proportional chambers, except in the most central region of M1, where gas electron multiplier (GEM) detectors are used since it has to withstand a higher particle flux. The M2 to M5 stations are interleaved with 80 cm-thick iron absorbers to effectively stop all other charged particles. The total absorber thickness including the stations is of about 20 radiation lengths.

The geometry of all stations is projective, meaning that they increase in size proportionally to the distance to the nominal IP, to cover always the same angular acceptance, 20 to 306 mrad in the horizontal direction and 16 to 258 mrad in the vertical direction. Each station is divided in 4 regions R1 to R4 to account for the different occupancies as the distance from the beam-pipe grows. These linear dimensions and segmentation of the regions scale as 1:2:4:8 and with this geometry the channel occupancies are comparable in all regions. The layout of the muon stations M1 to M5 and the layout of the regions R1 to R4 in a single station can be seen in Fig. 2.10.

2.2.5 The trigger system

The LHCb trigger system [97] is designed to reduce the event rate from 40 MHz, which is the nominal bunch-crossing rate, down to 12.5 kHz, which is low enough to

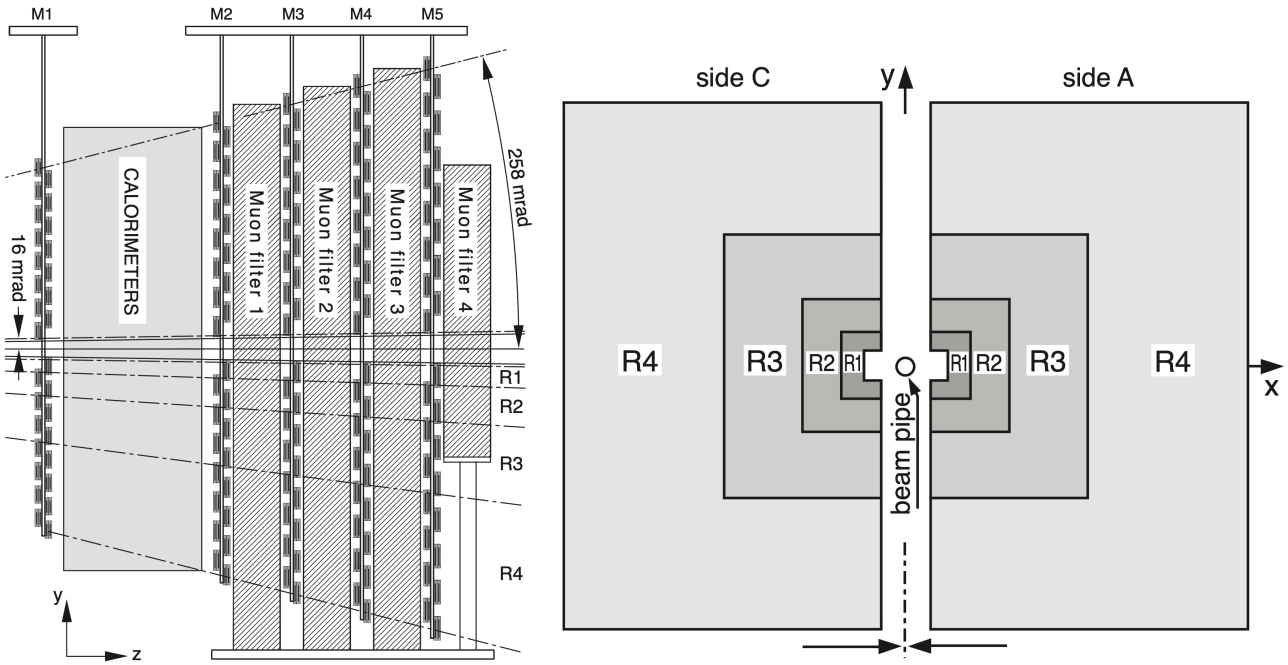

FIGURE 2.10: (*left*) Lateral view of the muon system and (*right*) the layout of the 4 regions R1–R4 in a single muon station. Figures taken from Ref. [83, 96].

allow the storage of the selected events for further offline analysis. A trigger system is always defined as a set of algorithms that classify events as being "interesting" or "uninteresting" for analysis, in such a way that most of the interesting processes are recorded efficiently. As a convention a single trigger algorithm is known as a "line", so a given trigger is effectively a collection of trigger lines.

The LHCb trigger system is divided in stages, the first one being the level 0, or L0, where some components of the detector need to be read out at 40 MHz, this means that the trigger decision is made purely of hardware requirements, using only information from the calorimeters, PU stations and muon stations. The following stage is the software high-level trigger (HLT) which in itself is further divided into the HLT1 and HLT2 stages, where the HLT1 consists of a partial event reconstruction and the HLT2 performs an offline-like full event reconstruction. A scheme of the trigger system flow can be seen in Fig. 2.11.

The L0 trigger is divided into three independent units, the L0-Calorimeter, L0-Muon and L0-PileUp. The latter is only used for luminosity determination purposes [99]. The calorimeter unit makes use of the SPD, PRS, ECAL and HCAL detectors. An event is triggered when the computed transverse energy (E_T),[5] deposited in clusters of 2×2 cells in the ECAL and HCAL, is above a certain threshold. From the found clusters different types of candidates are built. The L0Hadron candidates are the highest E_T clusters in the HCAL, which also contain the E_T from the corresponding ECAL cluster. The L0Photon are the highest E_T in the ECAL with 1 or 2 hits in the PRS and no hit in the corresponding SPD cells. The L0Electron are the same as for the L0Photon but with at least 1 hit in the corresponding SPD cells. Separately, the total number of hits in the SPD are computed and required to be

[5]The transverse energy is defined as $E_T = \sum_{i=1}^{4} E_i \sin(\theta_i)$, where E_i is the energy deposited in cell i and θ_i is the angle between the z-axis and a line going from the cell centre to the nominal IP.

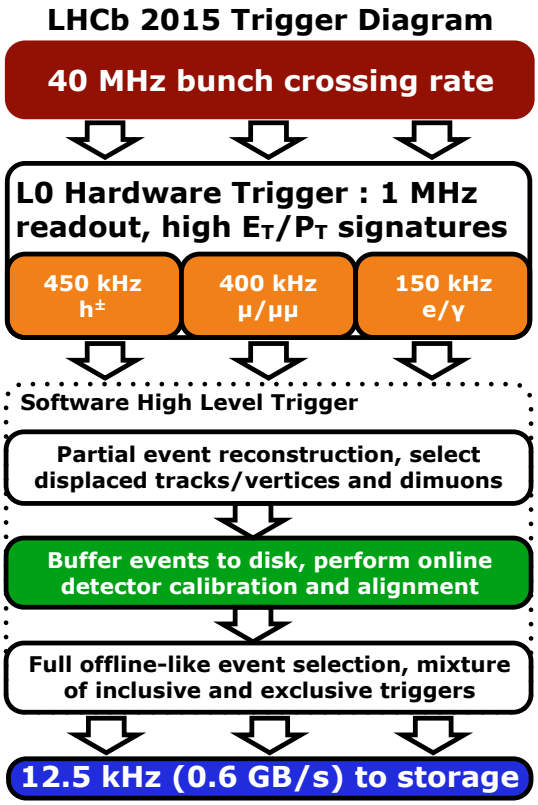

FIGURE 2.11: The LHCb trigger scheme for Run 2 of the LHC. Figure taken from Ref. [98].

below a certain threshold in order to avoid events that would require an exaggerated amount time to be processed in the HLT stage.

The L0Muon unit searches for straight-line tracks in the 5 muon stations. The direction of the reconstructed track is used to compute the particle p_T assuming it originated in the IP and that it only received a single kick by the LHCb magnet. The trigger decision is based on the two muons with the highest p_T. Either the highest momentum muon is above the L0Muon threshold or the product of the two largest momentum muons is above the L0DiMuon threshold.

In addition to these lines, there are other special trigger lines that select low multiplicity events to study central exclusive production and other inclusive jet measurements.

The HLT is a software trigger which takes place at the event filter farm which can accommodate $\sim 50\,000$ single-threaded processes using hyper-threading technology. The HLT1 is faster than the HLT2 in order to cope with the 1 MHz output from the L0 stage. The HLT1 uses one or two track signatures from the VELO or the TT to reconstruct particles that correspond to the L0 objects, or in the case of the L0Photon candidates confirm the absence of a charged particle that could be associated to these candidates. To illustrate this, consider the HLT1 line `Hlt1DiMuonHighMass`, which requires the muon track to have a p_T of at least 400 MeV/c, a total momentum of at least 3 GeV/c and that the track be reconstructed with a χ^2/ndf of less than 4. The HLT1 output is of about 30 kHz and at this rate the HLT2 can perform a more extensive offline-like event reconstruction to trigger the event. The events that were triggered by the HLT2 stage are stored at a rate of 12.5 kHz, for further offline analysis.

The different configurations for bandwidth allocations for different L0 lines or for

different HLT stages are managed by either pre- or post-scale factors, thus ensuring that the data-flow remains within the set volume. These different configurations are called the trigger configuration keys (TCKs) and they are identified by a hexadecimal number. During data-taking, different TCKs can be used even in the same fill. This is the case when for example, at the beginning of a fill the event rate is high and some trigger lines are heavily pre-scaled to fit within the allowed bandwidth, but then the interaction rates drop due to the degradation of the beams and a lower pre-scale can be used. It is important in any case to know the TCK used for the data-taking to be able to correct the results for possible scaling factors present.

Candidates for a specific decay or for a specific type of process are further identified in a process called "Stripping". Here, the reconstructed candidates are selected by certain stripping lines, which apply a specific selection, and stored in a data container for every line. This allows to work with much more reduced data volumes when performing an offline analysis.

2.3 SMOG

The System for Measuring the Overlap with Gas (SMOG) is a feature of LHCb [100–102] and it is unique at the LHC. It was originally designed to measure the overlap integral for the luminosity determination using the beam-gas imaging (BGI) technique. Its development was necessary in order to increase the beam-gas interactions, and it achieves this by injecting noble gasses into the VELO vacuum chamber. Additionally, SMOG gives LHCb the unique capability to function as a fixed-target experiment, by making proton-nucleus and nucleus-nucleus collisions possible between the LHC beams and various possible target types at different centre-of-mass energies.

2.3.1 The setup

The SMOG system, shown in Fig. 2.12, consists of a series of valves, pressure gauges, storage volumes and pumps, which allow for a very controlled gas-feeding into the pre-VELO volume and subsequently to the VELO vacuum chamber. A drawing representing the full SMOG system and its couplings until the VELO is shown in Fig. 2.13 using neon (Ne) as a source, but in principle other noble gasses can be and have been used. The valves PV501 ("Fill"), PV502 ("Bypass") and GV302 are controlled with the VELO vacuum control software, whereas the manual MV502 ("Isolate") valve is kept open. The circuit is equipped with different Penning (PE) and Pirani (PI) pressure gauges to monitor the low pressure in the different stages after the Fill valve, and a Piezo (PZ) pressure gauge for the high pressure volume before the Fill valve. The passage to the pre-VELO volume is through a non-evaporative getter (NEG) filter, which filters out residual gasses from the SMOG local circuit. The pre-VELO volume is pumped with a turbomolecular pump (TP301).The two PE gauges, PE411 and PE412 that are located in the VELO vessel, monitor the beam vacuum pressure at the IP. During laboratory testing of the gas-feed system, quadrupole mass analysers (QMA) were installed in the laboratory pre-VELO and VELO volumes in order to precisely measure not only the Ne presence and pressure, but also the relative abundance and pressures of other gasses produced by the outgassing of the circuit components such as hydrogen, water and carbon dioxide.

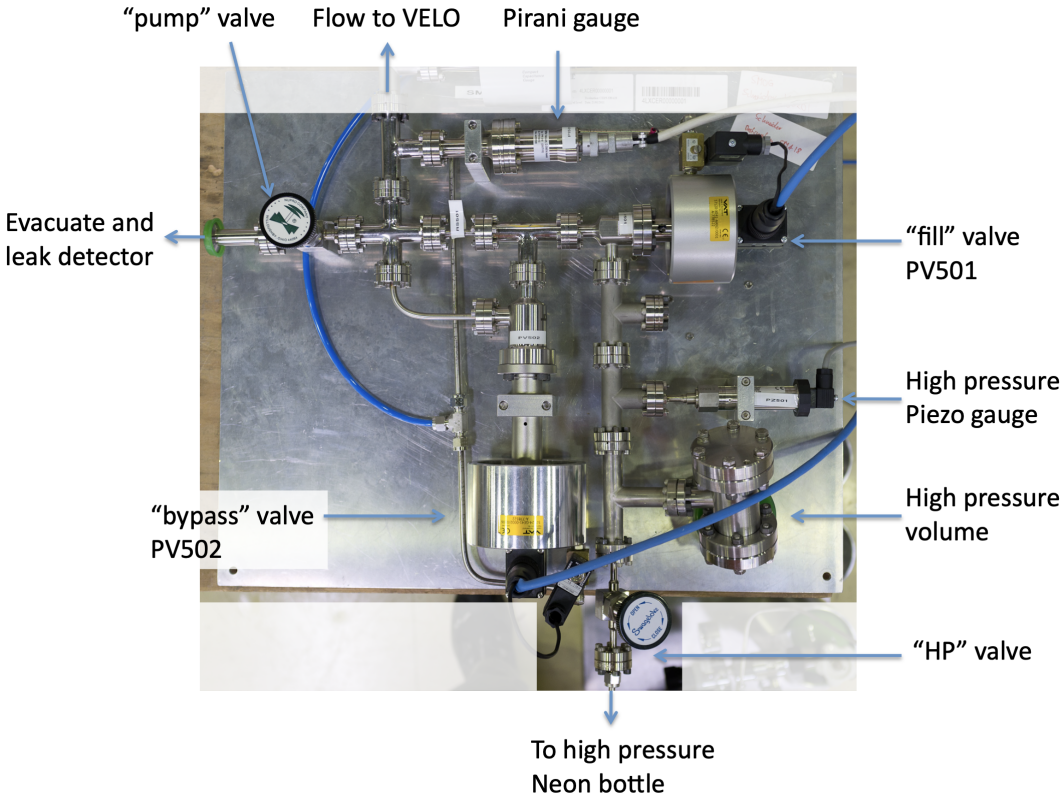

FIGURE 2.12: The SMOG device and its components. Figure taken from Ref. [102].

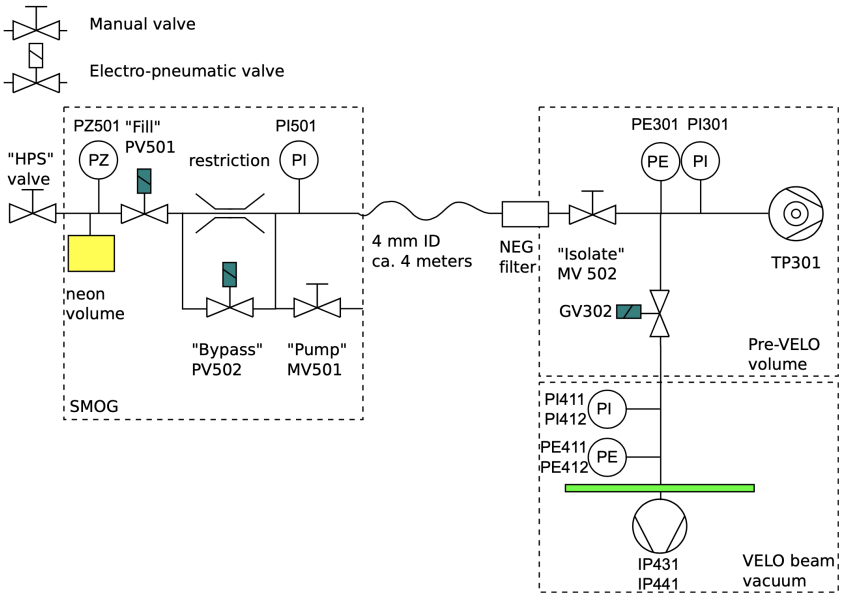

FIGURE 2.13: Drawing of the SMOG device organisation and the connection to the pre-VELO and VELO volumes. The gas to be injected is in yellow. Figure taken from Ref. [102].

2.3.2 The gas injection

Before gas injection, with the valve GV302 closed, the two ionic pumps of the VELO are switched off, increasing the pressure by a factor 5. Before the opening of the valve GV302, the SMOG is already injecting gas into the pre-VELO volume. Then when the GV302 valve is opened, the Ne gas starts to flow into the VELO and it reaches a stable pressure within a few minutes. Then the injected gas can flow freely into the beam-pipes, where it is pumped out at ± 20 m. When the injection is stopped, the valve GV302 is closed, and the Bypass valve open, to evacuate the remaining gas in the circuit. Subsequently the valve GV302 is opened again to allow the TP301 pump to evacuate out the Ne gas from the VELO.

A gas injection cycle can be seen in Fig. 2.14 (top left), where the pressure measured by the 2 PE gauges in the VELO is shown, as well as the beam-gas interaction rates as seen by the hardware triggers ODIN B1 and B2, dedicated to trigger on beam-gas interactions for beam 1 (LHCb-going) and beam 2 respectively, and the calorimeter rates. The pressure and interaction rate increase roughly by two orders of magnitude. The pressure shown by the PE gauges needs to be corrected by a factor of 4.1, to account for the different sensitivity of the gauges to Ne with respect to nitrogen gas.[6]

The time needed to return to the nominal background pressure after an injection was measured in the laboratory with the aid of a QMA in the pre-VELO volume. This procedure, starting from a steady state of injected gas, can be seen in Fig. 2.14 (bottom left). It takes $\sim$ 20 minutes for the Ne pressure to drop below the residual gas level.

In the long term, the Ne contained in the storage volume, and the injection pressure, will decrease due to the Ne being lost at every injection, and ultimately a refill of Ne will be needed. This Ne loss however, happens at a rate of 1%/day during continuous gas injection and is thus not critical for the stability of interaction rates or the pressure during a given fill of the LHC. The pressure evolution during a period of roughly 170 hours of gas injection can be seen in Fig. 2.14 (right). The top panel shows the pressure measured by the PZ gauge before the Fill valve, which reflects the changes in pressure due to the temperature changes. This happens because the PZ gauge measures the force which is proportional to the pressure and is thus sensitive to the temperature. On the other hand the PE gauges and the QMA measurements shown in the middle panel, are proportional to the gas density, and consequently they are not sensitive to the temperature. The bottom panel shows the residual gas pressure measured with the QMA.

2.3.3 Data-taking with SMOG

The SMOG gas injection is activated and initiated from the control room by the shift leader. There is a software protocol in place to make this task as simple as possible. Once the LHC has declared stable beams and the VELO is aligned and closed around the beams, the SMOG activation can begin. Once the gas injection has started, the shift leader will manually stop the current run,[7] and start a new one. This is to ensure that the data collected with SMOG will have been recorded with a stable gas pressure. This results in the first few minutes of gas injection being dumped, which is a negligible amount of time.

The SMOG system made it possible to drastically expand the LHCb physics programme by allowing it to function as a fixed-target experiment. Since 2015 several

[6] The calibration of the penning gauges has been done with nitrogen gas.

[7] During data-taking, runs are the live periods where data is recorded. These are usually one hour long by default and can be manually ended. Not to be confused with the LHC Run 1 or Run 2.

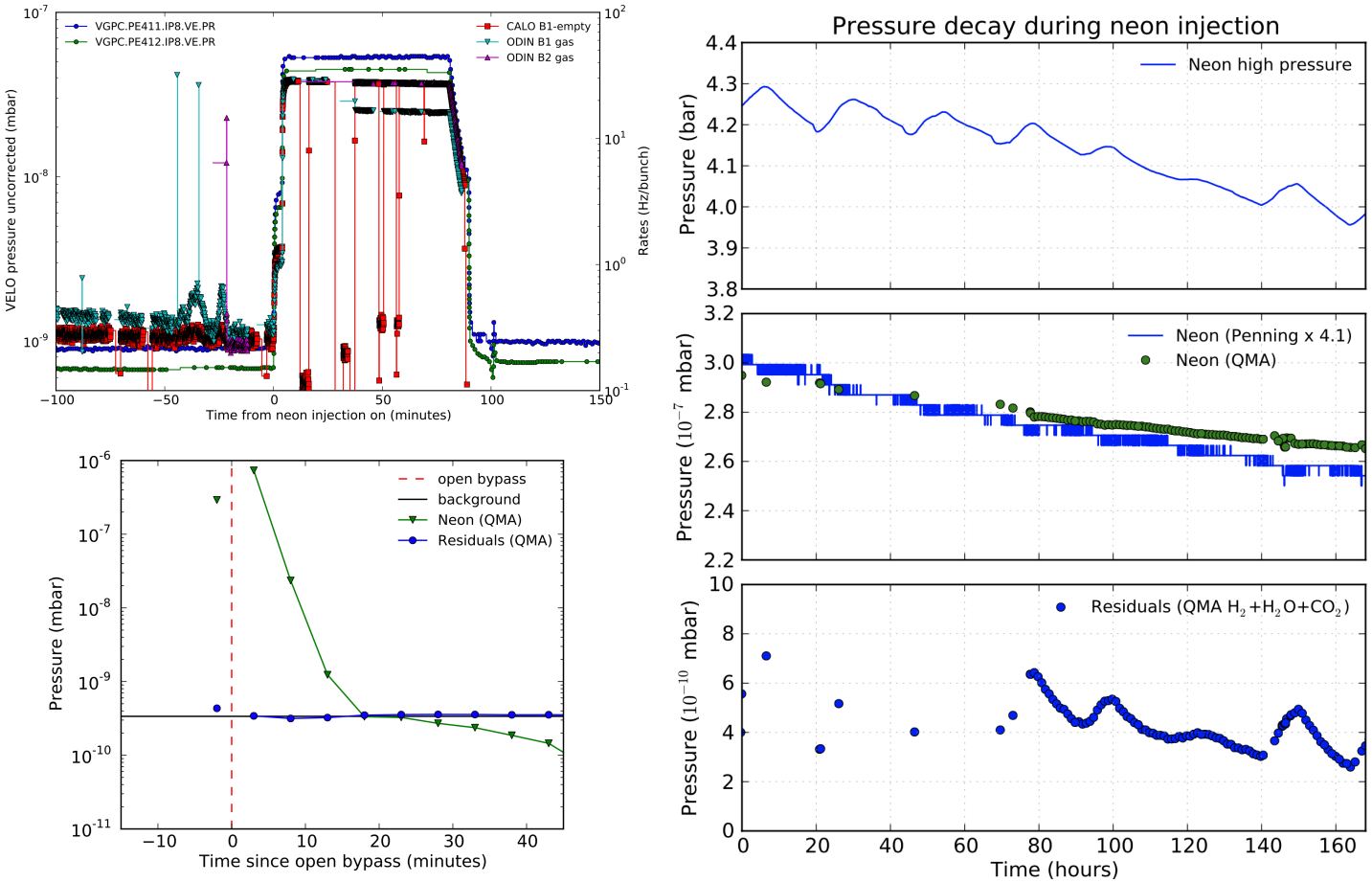

FIGURE 2.14: (*top left*) Pressure and beam-gas interaction rates during gas injection cycle from start to finish. The two penning gauges are shown in blue and green circles, the calorimeter trigger rate in red squares and the ODIN triggers in turquoise and purple triangles. (*bottom left*) Residual gasses and Ne pressure when the injection is stopped. The vertical dashed line marks the opening of the Bypass valve. The first measurement at t less than 0 was performed during an injection steady state. (*right*) Pressure decrease due to the periodic loss of Ne during long term injection. The top panel shows the pressure measured by the PZ gauge, which is sensitive to the temperature and exhibits the oscillating behaviour. The middle panel shows the Ne pressure measured by the PE gauge and the QMA. The bottom panel shows the residual gas pressure measured by the QMA. Figures taken from Ref. [102].

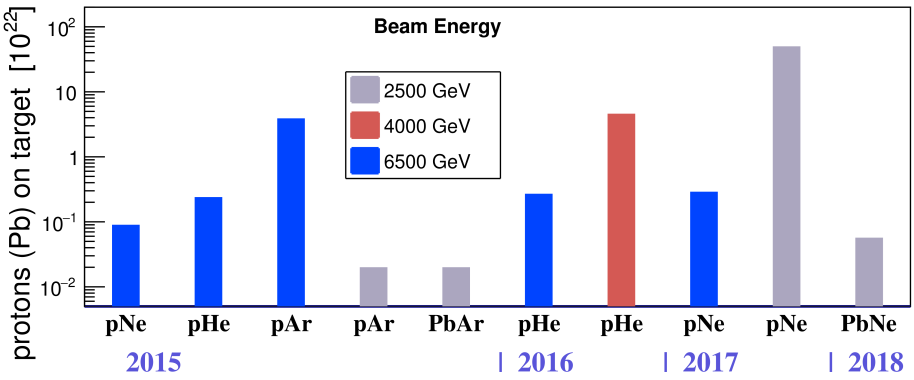

FIGURE 2.15: Different datasets collected since 2015 to date with the different SMOG runs. Different gas types and beam energies have been used.

dedicated runs have taken place with helium, argon and neon targets with p and Pb beams of energies ranging from 2.5 to 6.5 TeV. The recorded datasets to date can be seen in Fig. 2.15. This extended physics programme has already produced high-quality results, such as heavy-flavour production in proton-nucleus collisions [65] and measurements of prompt antiproton production in proton-helium collisions [103]. It has also opened new doors in the search for intrinsic charm in the proton at high Bjorken-x and more importantly for this thesis, it has opened a new window into the study of QGP formation in Pb-nucleus collisions, specifically in PbNe collisions.

2.4 LHCb upgrade and SMOG2

During the LHC long shutdown 2, in 2019-2021, the LHCb detector is undergoing major upgrades of all its subdetectors. This will allow LHCb to handle an instantaneous luminosity a factor 5 higher in collider mode and also it will be able to be read out at the full bunch-crossing rate, no longer needing a hardware trigger.

In order to have a low pile-up of collisions in the IP, LHCb has worked until now with a levelled instantaneous luminosity of $4 \times 10^{-32}\,\mathrm{cm}^{-2}s^{-1}$. This is because the low pile-up condition helps to avoid the risk of mismatching a secondary vertex by another pp interaction. Now however, with the current reconstruction algorithms, this mismatching can be kept at the percent level while working with an instantaneous luminosity of $2 \times 10^{-33}\,\mathrm{cm}^{-2}s^{-1}$. This on the other hand, means that all the subdetectors must be able to withstand the higher radiation hardness and have a greater granularity for the increased particle density to be expected.

The trigger system will be completely remodelled [104]. During Runs 1 and 2 only limited information from the calorimeter and muon stations was read out at 40 MHz and then in another processing stage the HLT1 would reduce the candidate rate to 1 MHz, which is the maximum rate at which the full detector could be read out. With this upgrade all subdetectors will be read out at 40 MHz feeding the data directly into a software trigger which will allow to have much more complex trigger decisions as the first triggering stage.

The VELO will be upgraded [105] and will consist of 26 stations made up of 41 million $55 \times 55\,\mu\mathrm{m}^2$ hybrid pixel detectors. It will approach even closer to the beams, reaching a minimum distance of 5.1 mm. The TT will be replaced by the Upstream Tracker (UT) [106] which, as its predecessor, has 4 layers but now features 3 different

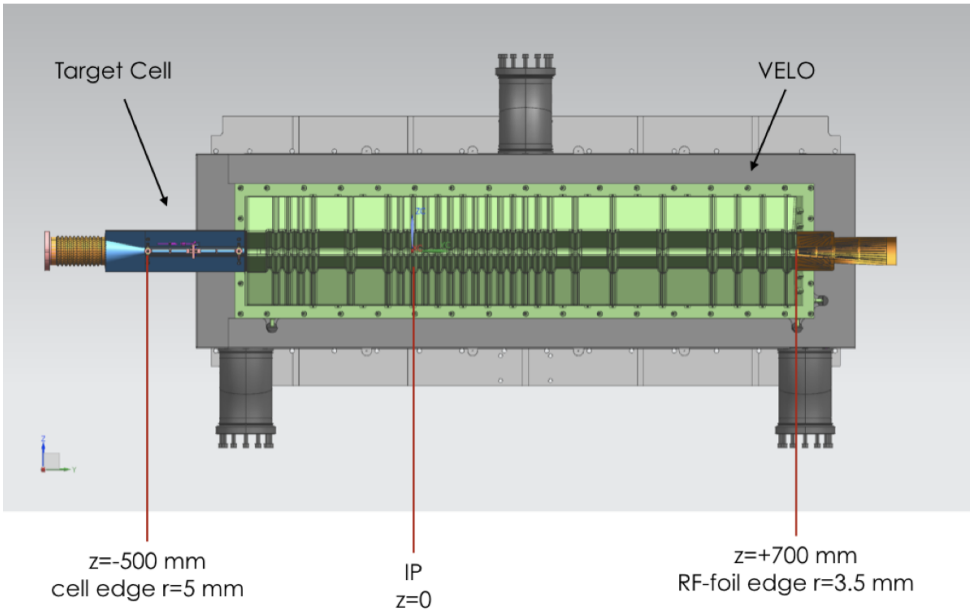

FIGURE 2.16: Schematic view of the SMOG2 setup attached to the VELO. The target cell in the figure is the SMOG2 storage cell. Figure taken from [108].

granularity regions depending on the distance to the beam-pipe. The new tracking stations downstream of the magnet, called the SciFi, will employ a single detector technology, the scintillating fibres. For particle identification [107], the RICH detectors will be at the same locations and use the same gas radiators as the current ones. The system of light-collection of RICH 1 will be upgraded to deal with the greater particle density and in both RICH detectors the photo multipliers will be changed in order to cope with the 40 MHz read-out. The muon stations M2-M5 remain the same with new electronics for the read-out, while the M1 station will be removed since its purpose was to provide information for the L0 hardware trigger. The only change will be in the central region of M2 where GEM detectors will be employed to cope with the high particle density. The calorimeter upgrade mainly consists on changing the electronics and getting rid of the SPD and PRS, whose function was mainly to provide information to the L0 trigger.

The SMOG system will become SMOG2 [108, 109], which will consist in a storage cell located upstream of the VELO. This will provide much more precise knowledge of the gas pressure and increase the beam-gas interaction rate by a factor of 10^2 due to the higher achievable pressures. Also the new configuration with the SMOG displaced from the nominal pp IP, provides a very clean separation of the PV originating from beam-beam interactions and from beam-gas interactions. This results in an important reduction of the background and also makes parallel running of collider and fixed target mode possible for the entire data-taking period. In addition, with this more sophisticated and controlled gas feed system, many more gas species can be used, such as hydrogen and deuterium, nitrogen or oxygen, going beyond the noble gasses.

The SMOG2 storage cell is made up of 2 halves attached to the VELO halves, this allows the cell to open and close along with the VELO. A schematic view of the storage cell and its attachment to the VELO can be seen in Fig. 2.16.

With LHCb increasing so drastically its capabilities and with the new improved SMOG2, the expectations remain high for Run 3 and even higher for the future upgrades already scheduled.

Chapter 3

SMOG data-taking and quality

In this chapter the heavy-ion data-taking period of 2018 is described, including the machine set-up and the filling schemes used. Before using the recorded data for the analysis, it is important to evaluate the quality of the stored data. For this, all runs are processed and it is verified that the rates of events are coherent between the different runs and within a given fill. This filtering step described here allows to leave out runs whose rate of candidate selection is abnormal or not understood.

3.1 The data

The data used were recorded during the heavy-ion run of the LHC which took place late 2018, with the first stable beams declared on November 8^{th} until December 1^{st}. Lead nuclei are accelerated to 2.51 TeV and made to collide at a centre-of-mass energy of $\sqrt{s_{\text{NN}}} = 5.02\,\text{TeV}$. Neon gas was injected into the VELO chamber where the nominal IP is located, thanks to the SMOG system. When the detector is set up to function with the injected gas the ionic vacuum pumps are switched off and the gas can flow freely up to $\pm 20\,\text{m}$ from the IP until it reaches the closest active pumps and is evacuated. This results in an average pressure of the order of 1×10^{-7} mbar, roughly two orders of magnitude higher than the nominal vacuum pressure of the LHC beam-pipes.

The Ne atoms act as a fixed-target for the forward going Pb beam resulting in a collision with a centre-of-mass energy of $\sqrt{s_{\text{NN}}} = 69\,\text{GeV}$ which are the events of interest for this analysis. The Pb ions are arranged in bunches within the beam, and the beam is characterised by the spacing (in ns) between bunches, the number of filled bunches, and the number of colliding bunches in each experiment. All of these factors affect the event rate that is recorded.

3.1.1 Filling schemes

The filling schemes are defined in order to provide significant instant luminosity to the four main LHC experiments; ATLAS, ALICE, CMS and LHCb. Before the beam is fully loaded in the LHC, ramped up to the operational energy and ready for physics, a series of steps need to be performed in the different accelerator systems [110–112]. At each stage, several bunches of Pb nuclei need to be accumulated from the previous stage, filling the available space before being passed to the following accelerator. Like this the train of bunches becomes longer and eventually builds the final LHC beam.

The train of bunches has different possible values for bunch spacing, and also since it is possible to have empty bunches, there are different possible numbers of filled Pb bunches and different numbers of colliding bunches at each IP. The filling schemes used during the 2018 heavy ion run can be seen in Tab. 3.1, where FS1 to FS10 denote the different schemes used. The number of non-colliding bunches, shown

TABLE 3.1: Filling schemes used during the 2018 data-taking period. The number of colliding and non-colliding bunches refers specifically to LHCb.

Filling scheme	FS1	FS2	FS3	FS4	FS5	FS6	FS7	FS8	FS9	FS10
Bunch spacing (ns)	100	100	100	100	100	100	100	75	75	75
Total bunches	64	260	484	592	648	632	648	460	670	733
Colliding bunches	45	35	53	132	75	126	52	304	384	468
Non-colliding bunches	19	225	431	460	573	506	596	156	286	265

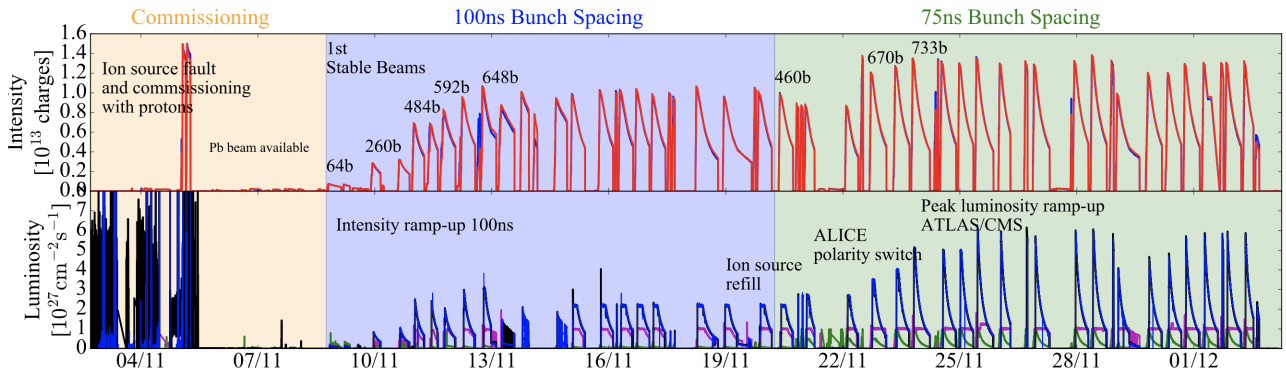

FIGURE 3.1: Overview of the 2018 heavy-ion run at the LHC. The beam intensity is shown on the top panel and the individual instantaneous luminosity of the experiments is shown (ATLAS and CMS in blue; ALICE in magenta; LHCb in green). Major changes are indicated. Figure taken from Ref. [113].

in Tab. 3.1, corresponds to the number of bunches that go through LHCb without crossing a bunch in the opposite direction. This number is important for the PbNe collisions, since they are only recorded when bunches from beam 1 (the LHCb-going beam) do not cross a bunch travelling in the opposite sense (a bunch from beam 2) and thus the higher the number of non-colliding bunches, the higher the rate of PbNe collisions available.

The first part of the heavy-ion run was marked by an ion source fault and the commissioning had to begin with proton beams to perform some beam optics measurements and corrections for the future Pb beams [113]. When the Pb beams became available, the final adjustments on the optics, collimation set-up and the validation of the collision configuration for the machine protection were performed. From the first stable beams, with a bunch spacing of 100 ns, the number of filled Pb bunches in the beam was increased steadily going from 64 in the first fills, to 648. Once this was reached, it was the norm until the beams were upgraded to a bunch spacing of 75 ns. In this new regime the filled bunches were increased from 460 to 733, surpassing greatly the initial design goal of 592 bunches. This 75 ns scheme resulted in many more bunch collisions at LHCb, but it was detrimental for the PbNe events since it meant less non-colliding bunches. An outline of the entire 2018 heavy-ion run can be seen in Fig. 3.1.

TABLE 3.2: Stripping and HLT1 lines definition for the J/ψ candidate selection.

HeavyIonDiMuonJpsi2MuMu	
$\mu^{\pm}$	$p_{\mathrm{T}} > 500\,\mathrm{MeV}/c$, track $\chi^2/\mathrm{ndof} < 10$
J/ψ	vertex $\chi^2/\mathrm{ndof} < 25$
Invariant mass	$\in [2900, 4000]\,\mathrm{MeV}/c^2$
Hlt1SMOGDiMuonHighMass	
μ track	$p_{\mathrm{T}} > 400.0\,\mathrm{MeV}/c$, $p > 3\,\mathrm{GeV}/c$
	track $\chi^2/\mathrm{ndof} < 4$
	track ghost probability < 999

3.2 Event rates

To have a measure of the data quality from the PbNe collisions certain event rates are studied. This is done with the aim of finding runs that present an unexpected behaviour with respect to the trend, or runs in which the D^0 candidate rate changed differently than the J/ψ candidate rate. The D^0 and J/ψ candidates are used as a proxy for what the main analysis will be, described in Sec. 5, but it should be noted that at this stage no signal selection is performed.

Only runs in which Ne was injected, via the SMOG system, were considered. This is important since the goal is to study PbNe interactions, and in order to detect any anomaly in the runs analysed by looking at the candidate rates it is necessary that the gas injection be already finished. For this reason, runs during which the gas is being injected are completely left out. This does not pose a big loss because usually the runs in which the gas is injected are of very short duration since a few minutes after the gas injection the run is manually cut and a new run, potentially valid for physics analysis, is initiated.

In order to obtain the candidate rates to be studied, the first step is to consider the already stripped data. D^0 candidates are obtained from reconstructed D^0 in the $K^-\pi^+$ decay channel (the charge conjugate is implied) from events that triggered the `HeavyIonOpenCharmD02kpi` stripping line and the HLT1 `Hlt1BEHasTrack` line. In addition, the events must have taken place during the crossing of a filled Pb bunch from beam 1 and an empty Pb bunch from beam 2 (`BCType` 1), there should be at least one reconstructed primary vertex (`nPV` > 0) and only the best candidate found in the event is kept (`nCandidate` = 0). The best candidate corresponds to the candidate with the best reconstructed primary vertex. Since the vertex reconstruction involves the fitting of a set of tracks that lead to roughly the same point in space (the primary vertex), the best vertex is the one for which the fit χ^2/ndf value is minimum.

The J/ψ candidates on the other hand, are obtained from reconstructed J/ψ in the $\mu^+\mu^-$ decay channel from events that triggered the `HeavyIonDiMuonJpsi2MuMu` stripping line and the HLT1 `Hlt1SMOGDiMuonHighMass` line. As for the D^0, the events must have taken place during a `BCType` 1 beam-crossing, have at least one reconstructed primary vertex and only the best candidate found is kept.

The definition of the stripping and the HLT1 lines used are summarised in Tab. 3.2 and 3.3 for J/ψ and D^0 respectively. In the tables IP refers to the impact parameter, that is, the transverse distance of the track to the PV. The χ^2/ndof of tracks and vertices refers to their reconstruction quality. The track ghost probability, as the name suggests, refers to the probability that the track is a fake track reconstructed by the reconstruction algorithm.

TABLE 3.3: Stripping and HLT1 lines definition for the D^0 candidate selection.

HeavyIonOpenCharmD02kpi	
K^-	$p_T > 250\,\text{MeV}/c$, $p > 3\,\text{GeV}/c$, $\eta \in [0, 10]$, PIDK-PIDpi > 0.0, track $\chi^2/\text{ndof} < 5$, IP $\chi^2/\text{ndof} > 4$
π^+	$p_T > 250\,\text{MeV}/c$, $p > 3\,\text{GeV}/c$, $\eta \in [0, 10]$, PIDK-PIDpi < 5.0, track $\chi^2/\text{ndof} < 5$, IP $\chi^2/\text{ndof} > 4$
DOCA	$< 2\,\text{mm}$
At least 1 daughter	$p_T > 500\,\text{MeV}/c$
D^0	vertex $\chi^2/\text{ndof} < 16$
Invariant mass	$\in [1764.84, 1964.84]\,\text{MeV}/c^2$
Hlt1BEHasTrack	
π^+ or K^- track	$p_T > 500.0\,\text{MeV}/c$, track $\chi^2/\text{ndof} < 4$, track ghost probability < 0.8

Once the absolute number of candidates have been obtained for every run, by means of the aforementioned selection, it is divided by the run duration effectively resulting in the D^0 and J/ψ candidate rate per minute. Each candidate has a time-stamp given by the variable `GpsTime`, and thus the run duration is taken to be the difference of the maximum `GpsTime` and the minimum found for every run.

As another criterion for the run pre-selection it was required that the run duration be longer than 5 minutes to ensure a stable condition of the data-taking. The only exception to this being runs that were taken at the end of the fill and that were ended by the beam-dump. After this pre-selection stage the integrated recorded time left was 224.45 hours and the time left out due to the runs being less than 5 minutes long was 0.71 hours ($\sim 0.3\%$).

In Fig. 3.2 the D^0 and J/ψ candidate rates are shown for every run of the 2018 heavy-ion data-taking period during which there was gas injected. There is a clear decreasing trend within each fill, which reflects the beam intensity degradation with time due to the continued collisions. Another striking feature is the high-rate and low-rate regions, specifically the period before fill 7460 (high-rate) and after run 7460 (low-rate). This is a reflection of the different bunch spacing used in the filling schemes. Before fill 7460 only filling schemes FS1 to FS7 were used, which have a bunch spacing of 100 ns and from fill 7460 onwards only filling schemes FS8 to FS10 were used, with a bunch spacing of 75 ns. The only exception being fill 7483 where filling scheme FS7 was used for a *van der Meer* scan[1] performed for the ALICE experiment.

In Fig. 3.2, as opposed to the behaviour shown in Fig. 3.1, a higher event rate is evident when the beams are filled with the 100 ns filling schemes. This is because the recorded PbNe collision rates are higher when the number of non-colliding bunches is high, which is contrary to the rates in normal beam-beam collisions where the more filled bunches, the higher the collision rate.

In order to identify problematic runs different approaches were implemented. The first one is to identify runs which exhibit a rate that does not fit in the decreasing trend within each fill. As a concrete example, the expected behaviour after the beginning of

[1] A *van der Meer* scan is a method to determine the luminosity in collider experiments and it consists in moving the beams across each other in the vertical and horizontal directions to measure the collision rates at different beam positions. This allows one to estimate the collider luminosity [114].

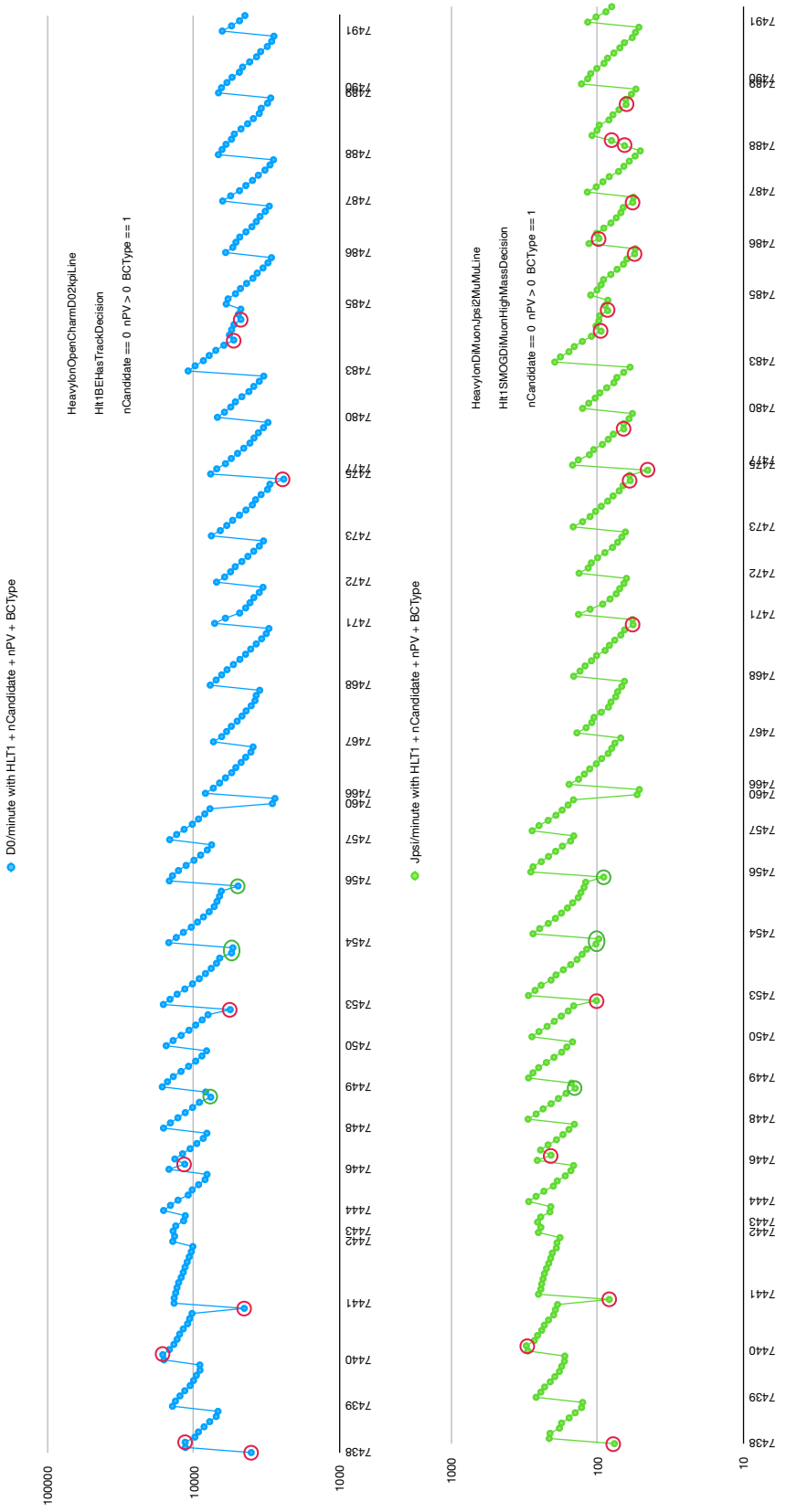

FIGURE 3.2: Rates for (*top*) D^0 and (*bottom*) J/ψ candidates per minute which fulfil the respective HLT1 requirements. In both cases, it is required that only one candidate per event is kept and only from events originating from a beam-empty crossing with at least one primary vertex reconstructed. Identified problematic runs are circled with red if they were finally rejected and with green if they were kept (see text).

TABLE 3.4: Runs deemed fit for analysis, that present an out-of-trend behaviour. The reason that explains the apparent problem is detailed.

Fill	Run	Problem	Explanation
7448	218210	Too low	Duration 9:31 minutes and 59s lost ⇒ 9% lost.
7453	218325/6	Too low	SMOG was closed and re-opened 40 minutes later. Beam degraded.
7454	218371	Too low	LHC in adjust mode for 5 hours. SMOG was closed and re-opened at the end of the fill.

a fill is a steady decrease in the recorded rate, so if a given run presents an increase in the rate with respect to the previous run from the same fill, or an excessive drop which falls below the decrease trend, it is flagged as a potential bad run. In Fig. 3.2 the second run of fill 7440 and the last run of the same fill are examples of the increase and excessive drop respectively.

Several runs were flagged as showing a non-standard behaviour, and after a careful search in the LHCb data-taking logbook, they were classified as being fit for analysis (circled in green in Fig. 3.2) if there was a reasonable explanation for the value of the rate, or being not-fit (circled in red in Fig. 3.2) due to the fact that the behaviour is not clearly explained with the available information and thus, not understood.

The runs that were deemed as fit for analysis are shown in Tab. 3.4 with the corresponding explanation for their behaviour.

Another sign of runs that could pose a problem are rates that are affected differently for D^0 and for J/ψ, that is runs in which the rate drops more for D^0 than for J/ψ or vice versa. To study the possibility of runs being affected in this way the ratio of the rates of D^0 over J/ψ candidates was computed. There are two runs for which this ratio is visibly out of the ordinary, these are the first two runs from fill 7488 (see Fig. 3.2) which are runs 219015 and 219018. Here, a very low rate for the J/ψ is seen and the absence of this effect in the D^0 rates, results in a very large ratio. Since all runs have a different duration and thus a different statistical uncertainty, the weighted mean of the ratios of all the runs was computed, excluding the 2 initial runs of fill 7488. The weighted mean and the weighted standard deviation are defined as

$$\mu_{\mathrm{w}} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}, \quad \sigma_{\mathrm{w}} = \sqrt{\frac{\sum_{i=1}^{n} w_i (x_i - \mu_{\mathrm{w}})^2}{\frac{N-1}{N} \sum_{i=1}^{n} w_i}}, \tag{3.1}$$

resulting in $\mu_{\mathrm{w}} = 53.97$ and $\sigma_{\mathrm{w}} = 1.16$. To put it in perspective, the 2 initial runs of fill 7488 had a ratio of 103.88 and 79.69 respectively, being completely off the norm.

The ratio found for every run was finally normalised by the weighted mean. The resulting value for every run considered can be seen in Fig. 3.3. Here the uncertainty assigned to every run corresponds to the propagation for a quotient of the statistical uncertainty of each run and the weighted standard deviation σ_{w}. With this, all runs that are farther than 3σ away from 1 are to be left out.

3.3 Results

With the above-mentioned criteria, several runs were left out of the analysis. All the runs that were left out, either because of individually exhibiting an out-of-the-ordinary behaviour or because of the ratio being too far from the mean, are shown in Tab. 3.5. Due to the ratio being too far from 1, only two runs were rejected, these are

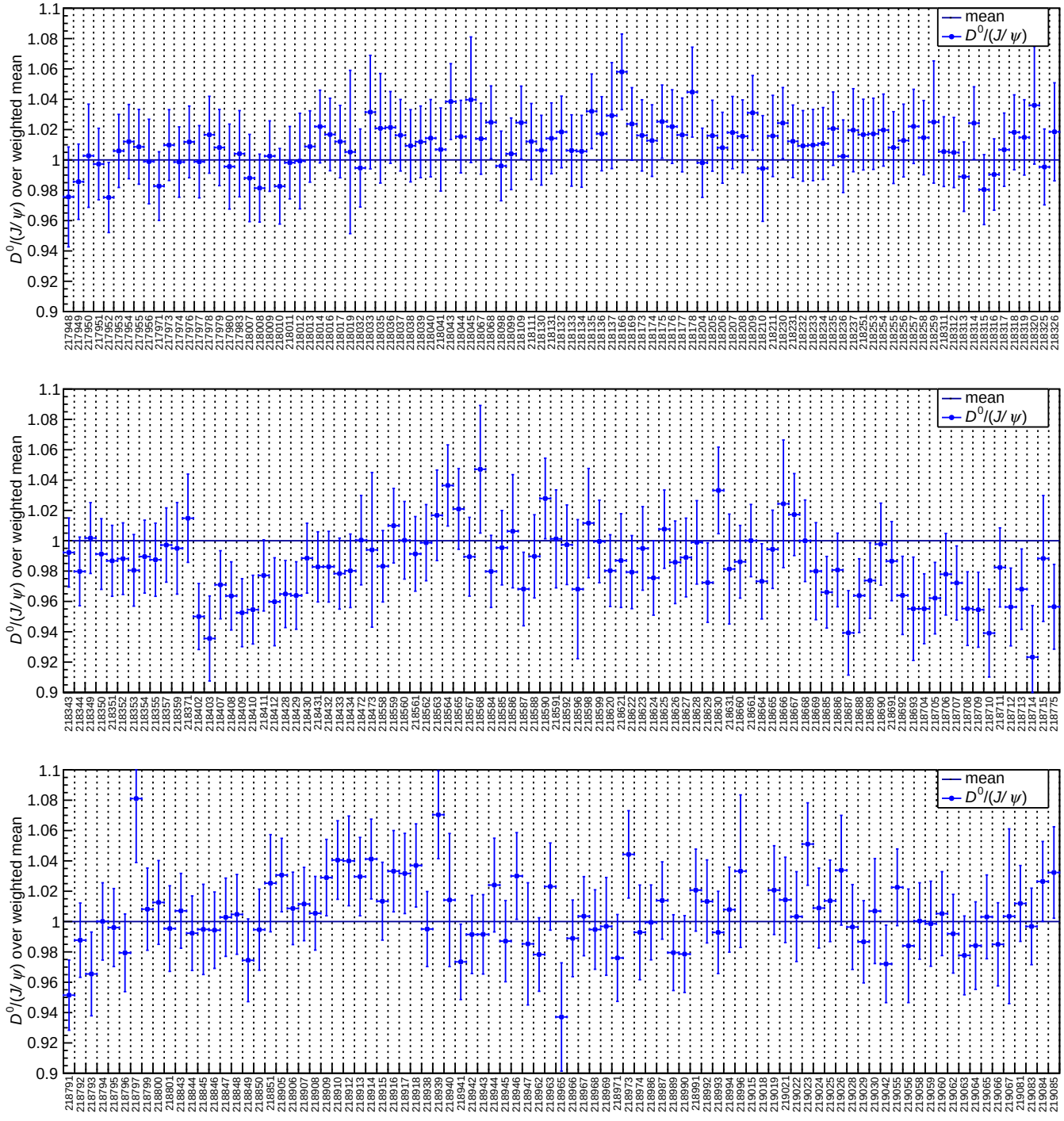

FIGURE 3.3: D^0 over J/ψ ratio for every run, normalised by the weighted mean of all runs (excluding runs 219015 and 219018). Only the ratios from runs 219015 and 219018 are farther than 3σ away from 1. The x-axis represents the run number.

TABLE 3.5: List of runs classified as not-fit for analysis. The duration of each individual run is shown. These runs correspond to 10.08 hours of data-taking out of a total of 224.45 hours.

Fill	Run	Duration [minutes]
7438	217948	19.73
7438	217950	7.03
7440	218008	60.05
7440	218019	5.11
7446	218169	59.99
7450	218259	9.28
7468	218630	60.02
7473	218713	60.05
7473	218715	17.34
7477	218797	14.59
7483	218912	31.36
7483	218916	57.61
7485	218946	60.41
7486	218963	32.25
7486	218973	60.01
7488	219015	8.84
7488	219018	20.34
7488	219026	21.02
Total time		10.08h

runs 219015 and 219018 from fill 7488. Amounting to 10.08 hours, the rejected runs correspond to 4.49% of the recorded data.

The full list of runs fit for analysis within each fill can be seen in detail in Tab. 3.6, where also the TCKs (discussed in Sec. 2.2.5) used in each fill are shown.

3.4 Conclusions

A careful study was performed on all the runs of the 2018 heavy-ion data-taking period where Ne gas was injected inside the VELO tank thanks to the SMOG system. Only runs which were longer than 5 minutes were considered to ensure the stable conditions of the runs. After a delicate selection, 18 runs were deemed not-fit for analysis due to their out-of-trend behaviour being unexplained by the available information. These runs amount to 4.49% of the data-taking time. The remaining runs correspond to 214.37 hours of recorded data and are now ready to be exploited for the main analysis.

TABLE 3.6: Fill numbers, TCKs and list of good runs of PbNe data. These runs amount to 214.37 hours of data-taking.

Fill	TCK	Good run list
7438	0x11851822	217949, 217951, 217952, 217953, 217954, 217955, 217956
7439	0x11851822	217971, 217973, 217974, 217976, 217977, 217978, 217979, 217980, 217983
7440	0x11851822-25	218007, 218009, 218010, 218011, 218012, 218013, 218014, 218016, 218017
7441	0x11851822-25	218032, 218033, 218035, 218036, 218037, 218038, 218039, 218040, 218041, 218043, 218044, 218045
7442	0x11851822-25	218067, 218068
7443	0x11851822-25	218098, 218099, 218109, 218111
7444	0x11851825	218130, 218131, 218132, 218133, 218134, 218135, 218136, 218137
7446	0x11871825	218166, 218173, 218174, 218175, 218176, 218177, 218178
7448	0x11871825	218204, 218205, 218206, 218207, 218208, 218209, 218210, 218211
7449	0x11871825	218230, 218231, 218232, 218233, 218234, 218235, 218236, 218237
7450	0x11871825	218251, 218253, 218254, 218255, 218256, 218257, 218258
7453	0x11871825	218311, 218312, 218313, 218314, 218315, 218316, 218317, 218318, 218319, 218320, 218325, 218326
7454	0x11871825	218343, 218344, 218349, 218350, 218351, 218352, 218353, 218354, 218355, 218357, 218359, 218371
7456	0x11871825	218402, 218403, 218407, 218408, 218409, 218410, 218411, 218412
7457	0x11871825	218428, 218429, 218430, 218431, 218432, 218433, 218434
7460	0x11871822-25	218472, 218473
7466	0x11871822-25	218558, 218559, 218560, 218561, 218562, 218563, 218564, 218565, 218567, 218568
7467	0x11871822-25	218584, 218585, 218586, 218587, 218588, 218590, 218591, 218592, 218596, 218598, 218599
7468	0x11871822-25	218620, 218621, 218622, 218623, 218624, 218625, 218626, 218627, 218628, 218629, 218631
7471	0x11871822-25	218660, 218661, 218664, 218665, 218666, 218667, 218668, 218669
7472	0x11871822-25	218685, 218686, 218687, 218688, 218689, 218690, 218691, 218692, 218693
7473	0x11871822-25	218704, 218705, 218706, 218707, 218708, 218709, 218710, 218711, 218712, 218714
7475	0x11871822	218775
7477	0x11871822-25	218791, 218792, 218793, 218794, 218795, 218796, 218799, 218800, 218801
7480	0x11871822-25	218843, 218844, 218845, 218846, 218847, 218848, 218849, 218850, 218851
7483	0x11871825	218905, 218906, 218907, 218908, 218909, 218910, 218913, 218914, 218915, 218917, 218918
7485	0x11871822-25	218938, 218939, 218940, 218941, 218942, 218943, 218944, 218945, 218947
7486	0x11871822-25	218962, 218965, 218966, 218967, 218968, 218969, 218971, 218974
7487	0x11871822-25	218986, 218987, 218989, 218990, 218991, 218992, 218993, 218994, 218996
7488	0x11871822-25	219019, 219021, 219022, 219023, 219024, 219025, 219028, 219029, 219030
7489	0x11871822	219042
7490	0x11871822-25	219055, 219056, 219058, 219059, 219060, 219062, 219063, 219064, 219065, 219066, 219067
7491	0x11871822-23	219081, 219083, 219084, 219085

Chapter 4

Centrality

The collision centrality in the context of heavy-ion collisions is of particular importance since it is directly related to the medium formed after the collision. This is a fundamental quantity in QGP studies because for its formation, an extended medium is needed, and this extension in turn is precisely dependent on the centrality of the collision.

After defining the centrality, the Glauber model and the assumptions on which it relies are described. Then, the datasets will be presented in detail, followed by the centrality determination procedure with its results and the study of systematic uncertainties, for PbPb first and subsequently for PbNe. Finally, the results are used to build and release a tool for the LHCb analysis framework, which is briefly described at the end of the chapter.

4.1 Definition

In the context of heavy-ion collisions, centrality is a quantity of relevance since it is directly related to the medium formed by the colliding nuclei, and measures the overlap region between the two nuclei in a collision. The centrality of a collision is characterised by the impact parameter (b) between the two nuclei, *i.e.* the distance between their centres in the plane transverse to the beam axis. The impact parameter defines the overlap region of the nuclei and thus determines also the size and shape of the resulting medium. A schematic view of a heavy-ion collision is shown in Fig. 4.1. The smaller the distance b, the more central is the collision. Centrality classes can be defined based on these geometric quantities, but in order to do this the knowledge of their distributions is needed, which is obtained through simulation.

The geometry of the collision is related to the number of nucleons that participate in it and the number of nucleon-nucleon collisions. These quantities are not directly accessible and hence need to be derived from the data recorded during the collisions by making use of other quantities that scale approximately with the number of participating nucleons, such as the outgoing particle multiplicity. For this purpose, a Glauber model is often used [115].

4.2 Glauber model

The centrality of a nucleus-nucleus collision is related to the overlap region between the nuclei where the nucleons are colliding. In practice, the particles produced by the collisions are not originating purely from hadronic interactions between the nuclei, but also from electromagnetic processes. Therefore, a model is needed to isolate the hadronic part and subsequently define the centrality classes. The most common approach in heavy-ion physics to model the collision of the two nuclei is to consider

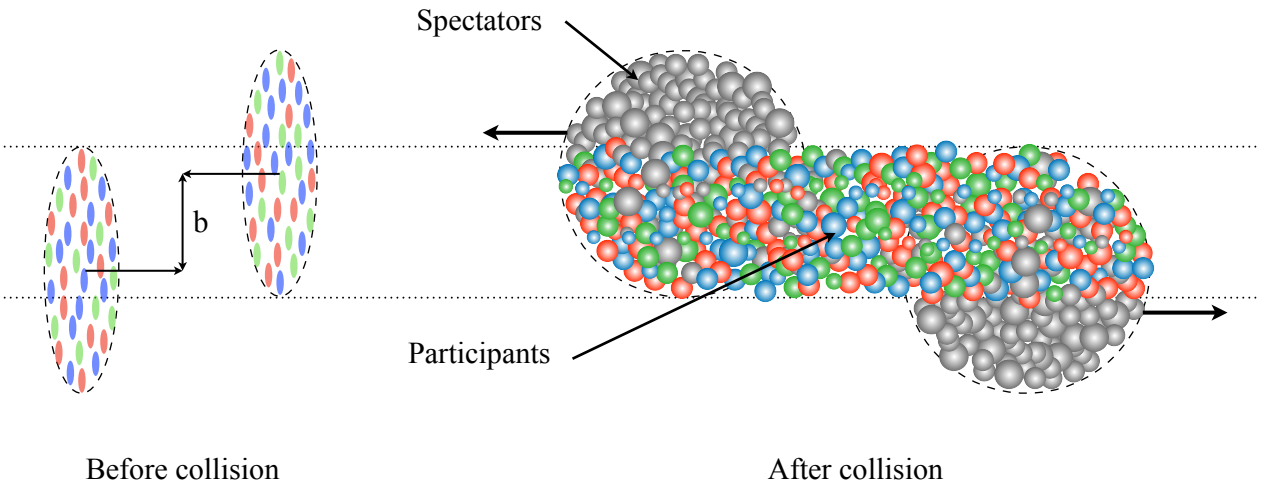

FIGURE 4.1: A schematic view of a heavy-ion collision. The impact parameter b is shown as well as the spectator nucleons and the participant nucleons.

the transverse shape of the nuclei. This shape is determined by a nuclear density function, described by a two-parameter Fermi distribution (2pF), also known as Woods–Saxon distributions, for each nuclear species considered, defined as

$$\rho(r) = \rho_0 \frac{1 + w(r/R)^2}{1 + \exp(\frac{r-R}{a})}, \tag{4.1}$$

where r stands for the radial distance from the centre of the nucleus. The parameter R corresponds to the nuclear radius, which is approximately the radial extension of the bulk of the nucleus. The diffusivity a describes the radial tail of the density distribution. The third parameter, w, is used to describe nuclei whose maximum density is reached at a radius $r > 0$. The values of these parameters are taken from other experiments, typically involving lepton-nucleus collisions and other types of nuclear spectroscopy [116, 117].

The Glauber model is generally approached in two ways, the optical Glauber model and the Monte Carlo (MC) Glauber model. The two colliding nuclei are labelled A and B. In the optical model, it is considered that nucleons from projectile A see the target B as a continuous distribution, which is described by an analytical function, and vice versa. This is also called the optical limit approximation. Subsequently, the overlap area, the number of participating nucleons (N_{part}) and the number of binary nucleon-nucleon collisions (N_{coll}) can be obtained analytically.

On the other hand, in the MC Glauber model, nucleons from each nuclei A and B are generated as hard spheres (there is a variation of the model that can also take into account the sub-nucleonic dynamics called Glauber–Gribov [118]) and are placed around the respective centres of the nuclei following the 2pF density distributions. Then a random impact parameter b is sampled from the distribution $\mathrm{d}\sigma/\mathrm{d}b = 2\pi b$,

where σ represents the area of the circle of radius b. Finally, the nuclei are made to collide with the following assumptions:

- nucleus-nucleus collisions are considered to be a superposition of several independent nucleon-nucleon collisions;
- nucleons are treated as hard spheres moving in straight lines all along the process, even if they have undergone a collision;
- nucleons have a geometrical transverse cross-section ($\sigma_{\text{NN}}^{\text{inel}}$) and two nucleons collide if the transverse distance between their centres is $d < \sqrt{\sigma_{\text{NN}}^{\text{inel}}/\pi}$.

The average values of the participating nucleons $\langle N_{\text{part}} \rangle$, the number of binary collisions $\langle N_{\text{coll}} \rangle$ and impact parameter $\langle b \rangle$, are obtained by simulating many nucleus-nucleus collisions. The distributions of N_{part}, N_{coll} and b can then be obtained for any centrality class.

The Glauber model has two relevant external inputs, the nucleon-nucleon inelastic cross-section $\sigma_{\text{NN}}^{\text{inel}}$ and the spatial distribution given by the 2pF distribution with its two parameters R and a. The cross-section is obtained from a data-driven parametrisation, possible thanks to the broad range of energies already available ranging from $\sim 20\,\text{GeV}$ to $\sim 60\,\text{TeV}$, given by $\sigma_{\text{NN}}^{\text{inel}}(s) = A + B\,\ln^2(s)$, with $A = 25.0 \pm 0.9\,\text{mb}$ and $B = 0.146 \pm 0.004\,\text{mb}$ [119].

4.3 PbPb data selection for centrality determination

In order to tune the MC Glauber model to reproduce the heavy ion collisions, a careful treatment of the data is needed. For this, a minimum bias (MB) data sample needs to be constructed with the available recorded data.

In theory, any observable from the data could be used for classifying events according to centrality, as long as it scales monotonously with the impact parameter, which is the main quantity of interest in this context. In reality, the reach in centrality classes possible within LHCb is limited by its tracking performances in a high multiplicity environment. In the case of PbPb collisions, this means that the VELO information cannot be used for this purpose, since it saturates under these conditions. The ECAL has the advantage of not saturating even in the most central collisions. Because of this, special attention is set on the energy deposited in the ECAL since it will be the observable that will be used for the centrality determination, described in Sec. 4.4.

The data sample used for this analysis corresponds to the data recorded in a special run of the 2018 PbPb data taking period, where the main L0 trigger lines were not pre- nor post-scaled. The PbPb collisions take place at a centre-of-mass energy of $\sqrt{s_{\text{NN}}} = 5.02\,\text{TeV}$. Simultaneously, data were being recorded of PbNe collisions that occur between the Pb beams and the Ne atoms injected into the VELO chamber. The run used was 217730, where the hardware L0 trigger configuration was TCK 0x1827, in which the only pre-scaled trigger line was `"L0SPDLowMultDecision"`. This is a low-multiplicity line which only triggers events with less than 50 hits in the SPD, which, as will be shown later, does not affect the centrality determination since it corresponds to the region in ECAL energy lower than 0.2 TeV, with an average energy deposit of 0.02 TeV, peaking at very low energy compared to the full range which goes up to $\sim 52\,\text{TeV}$.

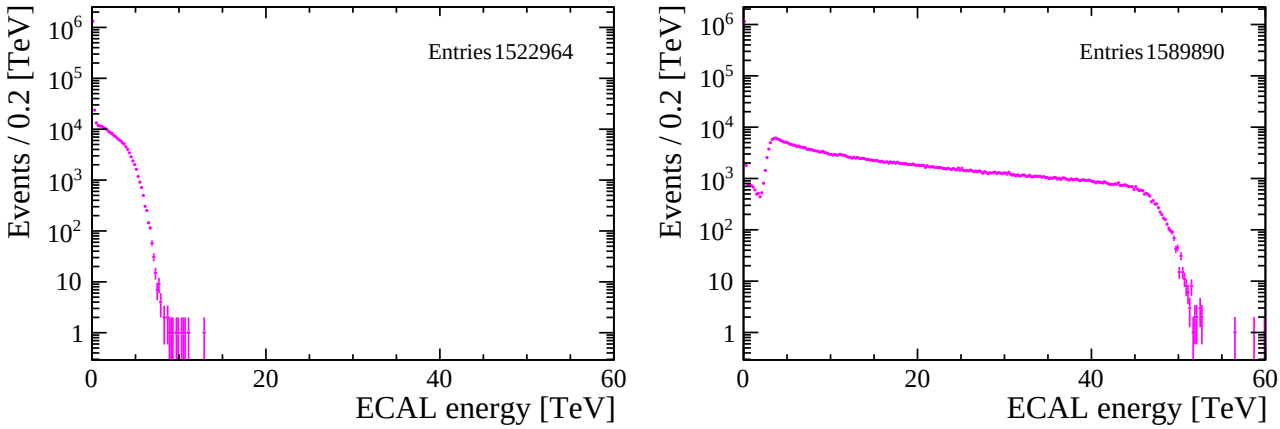

FIGURE 4.2: Energy distribution for events on IFT stream (*left*) and MINIBIAS stream (*right*). Data from PbPb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$.

At the stripping stage, the data are divided into two different streams, "IFT" and "MINIBIAS". In the "IFT" stream the data have been stored in containers according to the regular stripping lines. The "MINIBIAS" stream has two data containers for the lines `HeavyIonTopologyGammaLowActivityLine` and `MBHighMult`. The latter stripping line selects events with more than 10 000 `nVeloClusters` which are not usually used for the analyses because of the high detector occupancy, and are thus not included in the IFT stream. For this analysis, the two data streams are merged together using an appropriate combination of events from both streams, multiplicity cuts and scaling factors, in particular, the data from the IFT stream has to be scaled by a factor 5 to account for the post-scaling of the stripping line used to select this data.

The raw energy distribution of the events from each stream can be seen in Fig. 4.2. The resulting MB distribution of energy deposited in the ECAL on the full range of event multiplicity can be seen in the right plot of Fig. 4.3.

The merging of the two data streams, as shown in Fig. 4.3, results in a smooth distribution both for the energy deposited in the ECAL and for the distribution of the VELO clusters. The VELO clusters distribution exhibits a saturation structure with a peak in the tail and a sharp drop. The ECAL energy distribution on the other hand, shows no sign of saturation and is in a proper shape to move on to the centrality determination.

4.4 PbPb centrality determination

Centrality classes are quantiles of the inelastic PbPb cross section, and since the data contain contributions to the deposited energy in the ECAL from both hadronic and electromagnetic origin, the latter originating from peripheral collisions[1] where the electromagnetic interactions dominate and ultra-peripheral collisions (UPC),[2] the energy spectrum cannot be used straight away to define the desired quantiles for centrality. To overcome this issue, the MC Glauber model [119] is used to simulate

[1] Peripheral collisions refers to collisions where the impact parameter b is almost as large as the sum of the radii of the nuclei, and only a few nucleon-nucleon inelastic collisions take place.

[2] Ultra-peripheral collisions refers to collisions where the impact parameter b is larger than the sum of the radii of the nuclei, and the interaction is purely electromagnetic. There is no nucleon-nucleon inelastic collision.

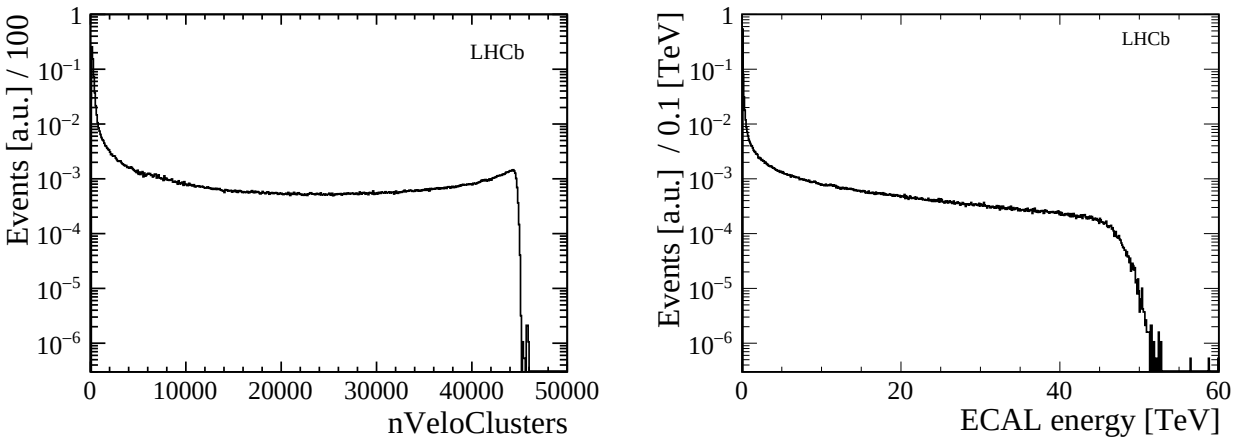

FIGURE 4.3: The number of VELO clusters (*left*) and the energy deposited in the ECAL (*right*) from PbPb collisions. The distribution of the VELO clusters exhibits a peak structure with a sharp fall at 45 000 clusters. This is due to the saturation of the VELO for high occupancy events.

the colliding nuclei and, from the resulting quantities, the corresponding energy distribution from hadronic origin can be constructed and fit to the data.

Bearing all of this in mind, the strategy is the following. The PbPb collision is simulated using the MC Glauber model. Geometric quantities coming out of this MC, such as b, N_{coll} or N_{part} can be used to construct the expected observable, which is in this case the energy deposited in the ECAL. The parameters of the model are then tuned to fit the ECAL energy distribution coming from the data. Finally, the centrality quantiles can be defined from the simulated distribution that corresponds only to the hadronic part of the interaction. Once this is done, the geometric quantities from the Glauber MC can be mapped to the data for each centrality class considered.

4.4.1 Methodology

In this section the simulation of the events is described first, then the generation of the simulated ECAL energy distribution and the steps to fit it to the data are explained. Once the fit has been performed, the simulated distribution is split into centrality classes based on the fraction of the total hadronic distribution integral, and the geometric variables of each class are mapped to the measured events falling in the same class.

Simulating the events

The first step of the way is to simulate the PbPb collisions using the TGLAUBERMC software from Ref. [119].[3] For ^{208}Pb the 2pF density distribution takes as input parameters $R = 6.68\,\text{fm}$ and $a = 0.447\,\text{fm}$ for the protons, $R = 6.69\,\text{fm}$ and $a = 0.560\,\text{fm}$ for the neutrons, and $w = 0$ for both. One million collisions were simulated using the corresponding nucleon-nucleon cross-section $\sigma_{\text{NN}}^{\text{inel}} = 67.6 \pm 0.6\,\text{mb}$ for a centre-of-mass energy of $\sqrt{s_{\text{NN}}} = 5.02\,\text{TeV}$. After the events have been simulated, the

[3]For this work, in the Glauber code the Pb nuclei was specified as "Pbpnrw", which considers slightly different distributions for protons and neutrons in the nucleus, and a reweighting of the nucleons positions to make the centre of mass coincide with the nominal position of the nucleus.

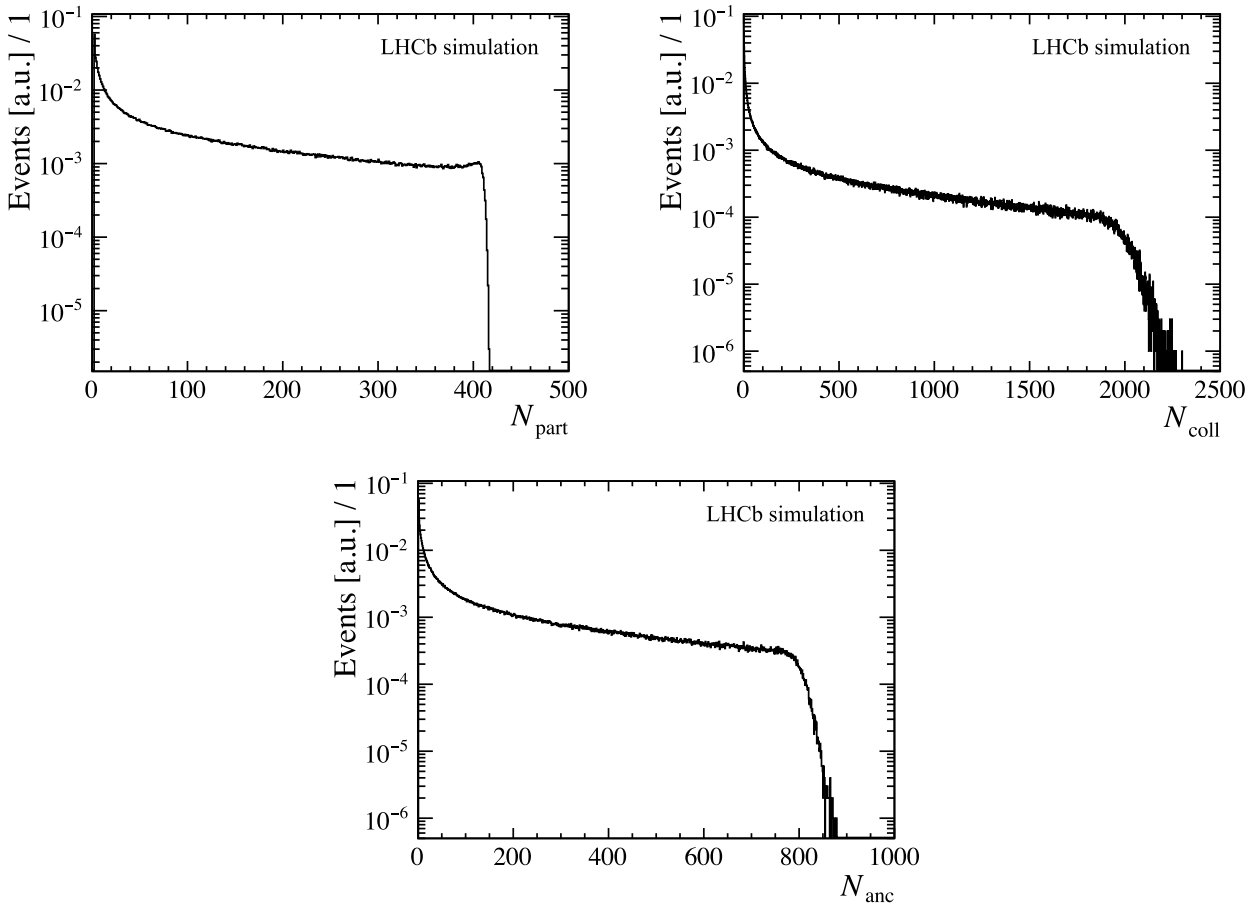

FIGURE 4.4: N_{part}, N_{coll} and N_{anc} from the MC Glauber model. For this N_{anc} distribution a value of $f = 0.75$ was used.

N_{part} and N_{coll} of every simulated collision can be accessed and with these numbers the number of ancestors N_{anc} is defined as

$$N_{\text{anc}} = f \times N_{\text{part}} + (1 - f) \times N_{\text{coll}} \tag{4.2}$$

which is taken to be the number of particle production sources. This is motivated by the fact that the particle multiplicity is expected to scale with N_{part} when soft processes dominate, and to scale with N_{coll} when hard processes dominate [120–124]. Here f is a free parameter to be fit later on. In Fig. 4.4 the obtained distributions of N_{part}, N_{coll} and N_{anc} are displayed with, as an example, $f = 0.75$.

To get the distribution of particles coming out of the collision, N_{anc} is convoluted with a negative binomial distribution (NBD) which has been extensively used to model particle production and has been shown to be a reasonable approach at diverse energy and rapidity regimes [125–129]. The NBD is defined as

$$P_{\mu,k}(n) = \frac{\Gamma(n+k)}{\Gamma(n+1)\Gamma(k)} \frac{(\mu/k)^n}{(\mu/k+1)^{n+k}} , \tag{4.3}$$

which can be rewritten in a discretised manner, considering $p = \left(\frac{\mu}{k} + 1\right)^{-1}$, as

$$P_{p,k}(n) = \frac{(n+k-1)!}{n!(k-1)!} p^k (1-p)^n , \tag{4.4}$$

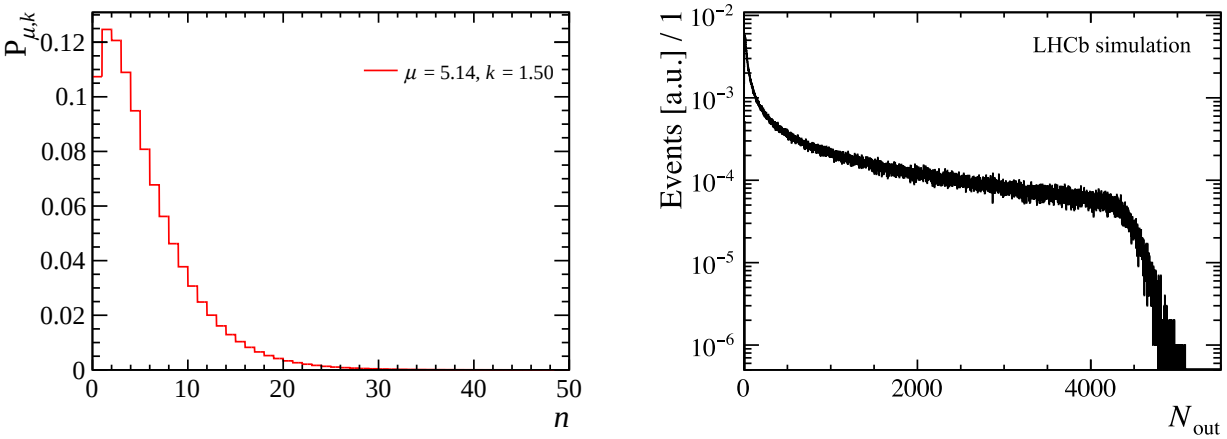

FIGURE 4.5: Negative Binomial Distribution (*left*) and distribution of the number of outgoing particles from the MC Glauber model (*right*).

where μ and k are free parameters and n is the number of particles that are produced which deposit energy in the ECAL. Since there are N_{anc} particle sources for each nucleus-nucleus collision, and each producing particles following a NBD, the NBD is sampled N_{anc} times to get the particle multiplicity distribution. Fig. 4.5 illustrates the NBD function and the result after sampling it N_{anc} times for each event to obtain the distribution of the number of outgoing particles N_{out} which deposit energy in the ECAL.

The last ingredient needed to make a simulated energy distribution to be able to compare with the data, is the mean energy deposited by each of these particles in the ECAL, which is computed in the following section.

Mean energy deposit

In the environment of the collision, most of the outgoing particles are pions which deposit energy in the HCAL ($\pi^{\pm}$) and in the ECAL ($\pi^0 \rightarrow \gamma\gamma$). The π^0 energy spectrum is approximated by the charged pion spectrum seen in MB pp collisions at 5.02 TeV recorded by LHCb in 2015. Charged pion tracks are studied to obtain their momentum, and given that their mass is negligible in comparison, it holds that $E^2 = m^2c^4 + p^2c^2 \Rightarrow E \simeq pc$, so the average energy deposit per particle in the ECAL can be determined.

Electrons ($e^{\pm}$) from the same dataset were studied in order to compare abundances and confirm that pions are the dominant species. The momentum distribution of pions ($\pi^{\pm}$) and electrons ($e^{\pm}$) can be seen in Fig. 4.6.

Neutral pions account for about one third of the total pion yield. In the considered dataset there were $\sim$ 38 million charged pions, this in turn means that there are about $\sim$ 19 million neutral pions, whose energy distribution follows the same trend as for the charged pions. On the other hand, $\sim$ 0.42 million electrons were found in the same dataset, which correspond to about 2.2% of the amount of neutral pions. Consequently the major part of the energy deposited on the ECAL comes from the neutral pions.

Bearing this in mind, the mean energy deposited per particle on the ECAL is taken to be $\langle E^{\text{PbPb}} \rangle = 10.44\,\text{GeV}$, and this is the value used in what follows. A simulated ECAL energy distribution can be seen in Fig. 4.7, which already approaches the shape found in the data.

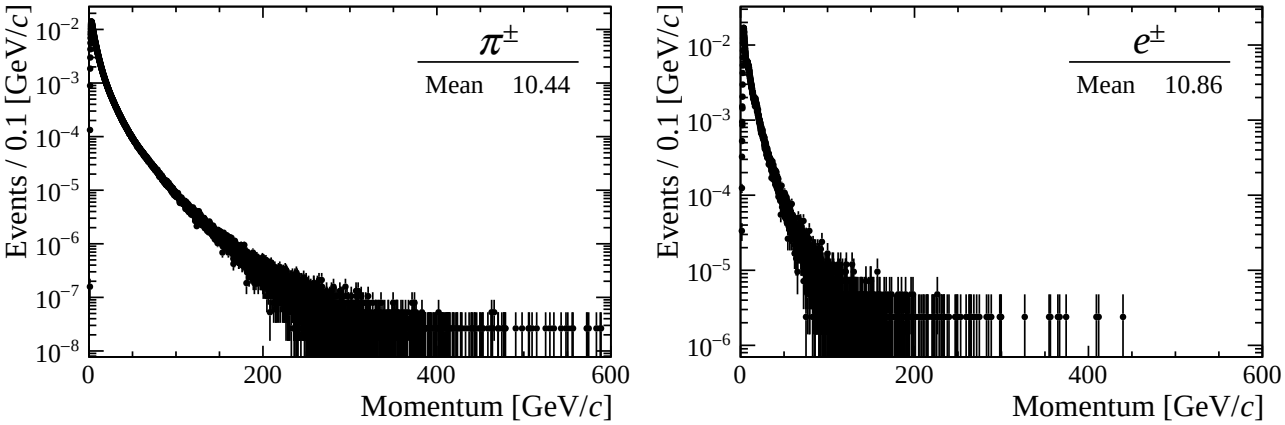

FIGURE 4.6: Momentum distribution for charged-pion tracks (*left*) and electron tracks ($e^{\pm}$) (*right*). Data from 2015 pp collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$.

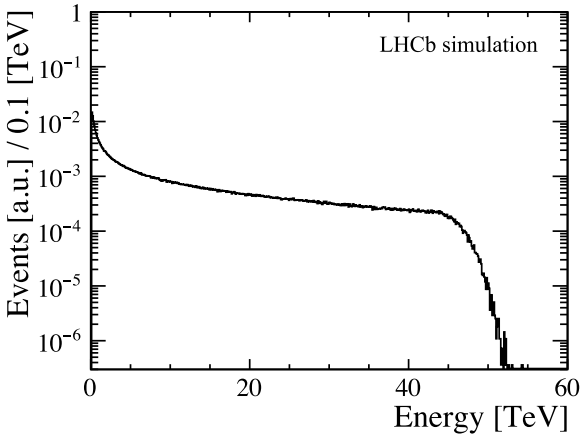

FIGURE 4.7: Simulated distribution of the energy deposited in the ECAL. A mean energy deposit per particle of $\langle E \rangle = 10.44\,\mathrm{GeV}$ is considered.

Finalising the model

In the model, to simulate the particle production, a NBD was used, see eq. 4.4, whose parameters were μ and k. The parameter k is linked to the width of the NBD distribution. The ALICE collaboration has used in their analysis the value $k = 1.6$ [130], and a previous LHCb study performed on 2015 PbPb data used $k = 1.5$ [131]. A simple comparison was made between simulated distributions varying k between 1.0 and 2.0 while everything else was kept constant, and it was found that there was no significant dependence on this parameter. The resulting distributions can be seen in Fig. 4.8. For this reason, the parameter k was fixed to $k = 1.5$.

Finally the model has only two free parameters, f from eq. 4.2 and μ from the NBD.

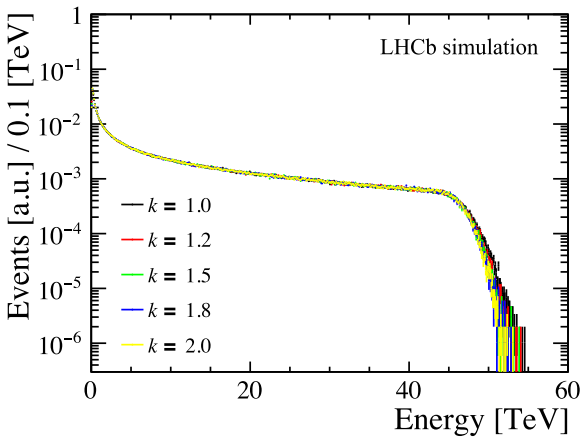

FIGURE 4.8: Resulting simulated energy distribution in the ECAL for $k \in [1.0, 2.0]$.

Fitting to data

To fit the simulated distribution to the data a χ^2 function between both distributions is minimised. The χ^2 is defined as

$$\chi^2 = \sum_{i \in \text{bins}} \frac{(E_i - O_i)^2}{E_i}, \tag{4.5}$$

where E_i and O_i are the "expected" and "observed" values for the i^{th} bin, *i.e.*, the simulated and measured values for a given energy bin. As has been mentioned before, in order to avoid possible contamination from electromagnetic origin present at low energy, the fitting range was chosen to be from 2 to 52 TeV. The MC Glauber energy distribution was normalised to the data in the energy range of 5 to 15 TeV to the extremes of the distributions.

The fit procedure is a very delicate one since the variables f and μ are highly correlated in the sense that they both modulate the horizontal reach of the distribution. This is easily seen due to the fact that μ is related to the mean value of the NBD, and thus the higher μ is, the higher the number of particles produced per collision and consequently the more energy is deposited in the ECAL. On the other hand, f controls how much the final distribution is "like" N_{part} or N_{coll}, which have a different reach on the x-axis. However, because of the same reason f controls the shape of the right shoulder, and that is one feature that can be exploited to overcome the correlation issue between f and μ. The fit is made in two stages. In the first one, only the right shoulder of the distributions is considered for the χ^2 computation, namely the region between 35 and 52 TeV, and with it the allowed range for f is defined. Then a range for μ is defined to match the range in f and a grid search is performed where the χ^2 is computed from 2 to 52 TeV.

In the first step μ is fixed to a test value, $\mu = 3.85$, and the similitude between the right shoulder of the data and the simulated distribution is evaluated for 1000 different values of f ranging from 0 to 1. For this procedure the simulated distribution was horizontally scaled to match the data by a factor H_s defined as:

$$H_s = \frac{E_{300}^{Data}}{E_{300}^{Glauber}}, \tag{4.6}$$

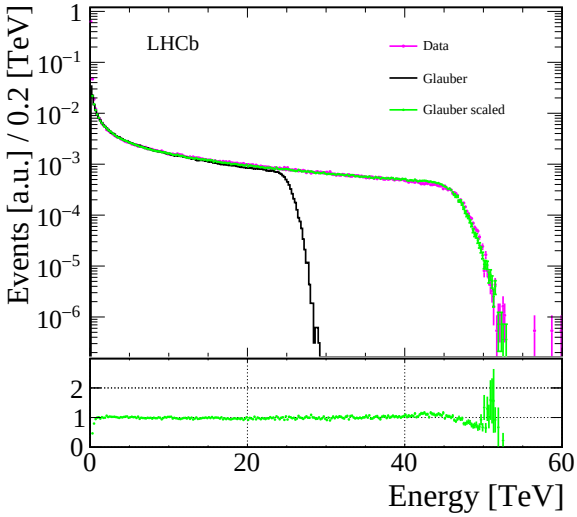

FIGURE 4.9: ECAL energy distribution comparison between data and MC Glauber. The black histogram corresponds to the simulated distribution with $f = 0.9$ which is then rescaled by H_s (green dots) to match the data and compare the right shoulders. The bottom panel shows the ratio of the MC Glauber over the data. For the whole process μ was fixed to 3.85.

where E_{300}^{Data} is the energy of the last bin containing more than 300 events in the data (bin centre), and $E_{300}^{Glauber}$ is the same but in the simulated distribution. This rescaling procedure can be seen in Fig. 4.9, where the data have been normalised to 1, meaning that 300 events now correspond to roughly 1.6×10^{-4}. The value of 300 events was chosen to pick a place in the distribution where the decrease is sharp, and thus provides an unambiguous point to the right of the "knee" of the distribution, and where there are still enough statistics to not be bothered by fluctuations. For this step the χ^2 was computed from 35 to 52 TeV. The resulting values for the χ^2 as a function of f can be seen in Fig. 4.10.

The minimum of the χ^2 distribution was found to be at $f = 0.83$. This result was

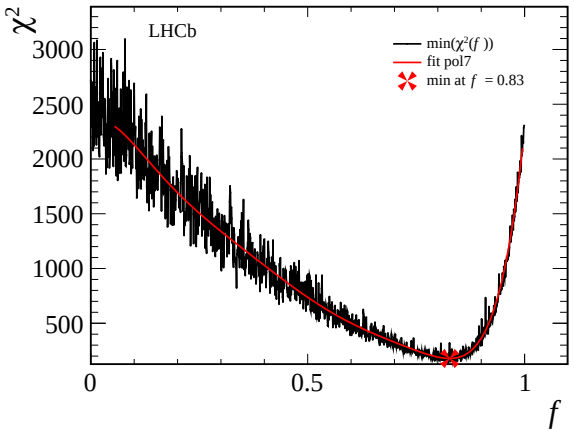

FIGURE 4.10: The χ^2 values for 1000 steps in $f \in [0, 1]$. The χ^2 distribution has been fit by a 7^{th} degree polynomial whose minimum is at $f = 0.83$.

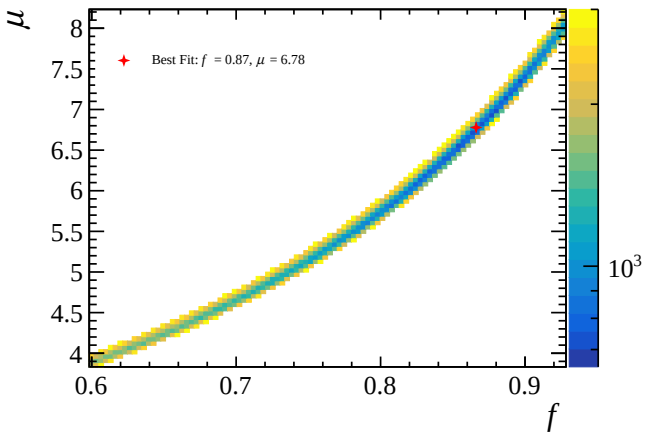

FIGURE 4.11: The χ^2 map for the grid search in $f \in [0.60, 0.93]$ and $\mu \in [3.7, 9.0]$. The "best fit" corresponds to the values $f = 0.866$ and $\mu = 6.778$.

taken as a reference to reduce the range in f for the subsequent grid search. The allowed range for f was set to be $[0.60, 0.93]$ (roughly the region where $\chi^2 < 0.5 \times 10^3$), from which the range for μ was chosen by estimating the needed value of μ for the simulation to match the data for the extremes cases of f. The resulting range for μ was set to be $[3.7, 9.0]$.

A grid of 100×100 was defined in the previously mentioned ranges, and the χ^2 was computed from 2 to 52 TeV at every point of the grid. The result of this grid search can be seen in Fig. 4.11.

The "best fit" found is not the point of the grid with the lowest value, since this is prone to be affected heavily by the unavoidable fluctuations from the random NBD sampling. Instead the χ^2 map from Fig. 4.11 is considered and the minimum parametrised by f and by μ separately. The strategy is as follows. In order to get the parametrisation as a function of f, the minimum of the χ^2 distribution as a function of μ in bins of f is found. Like this, the value of the minimum in each slice along μ (at fixed f) is assigned to the corresponding value of f. The same is done for the μ dependence. Doing this for all values of f and μ the f-parametrised minimum and μ-parametrised minimum are constructed. In Fig. 4.12 the parametrisations as a function of f and μ are shown and an example slice is displayed to illustrate the process. From this procedure, the best fit is found at $(f, \mu) = (0.866, 6.778)$ with a $\chi^2/\text{ndf} = 3.025$, which is the one shown in Fig. 4.11. By adopting this approach, the number of points on the grid is not limiting the result.

Finally another grid search was performed, the same amount of points was used but on a narrower range, namely $f \in [0.79, 0.92]$ and $\mu \in [5.7, 7.9]$. The resulting χ^2 map can be seen in Fig. 4.13. Two results are shown as a best fit. These results are computed using different methods that will be now discussed.

Since a finer grid was obtained, there are more non-empty bins in the histograms made up from the slices of f and μ. This can be exploited to go one step further and fit the slices with a 3^{rd} degree polynomial. The slices and their respective fits can be seen on the top row of Fig. 4.14. A 3^{rd} degree polynomial describes the slices well. Now in addition to saving the minima of the slices histograms, the minima of the fit to the slices can be saved. In the second row of Fig. 4.14 a comparison between the actual minima of the slices and the minima from the fit to the slices can be seen. For f the distribution is less noisy when considering the fit to the slices, whereas for

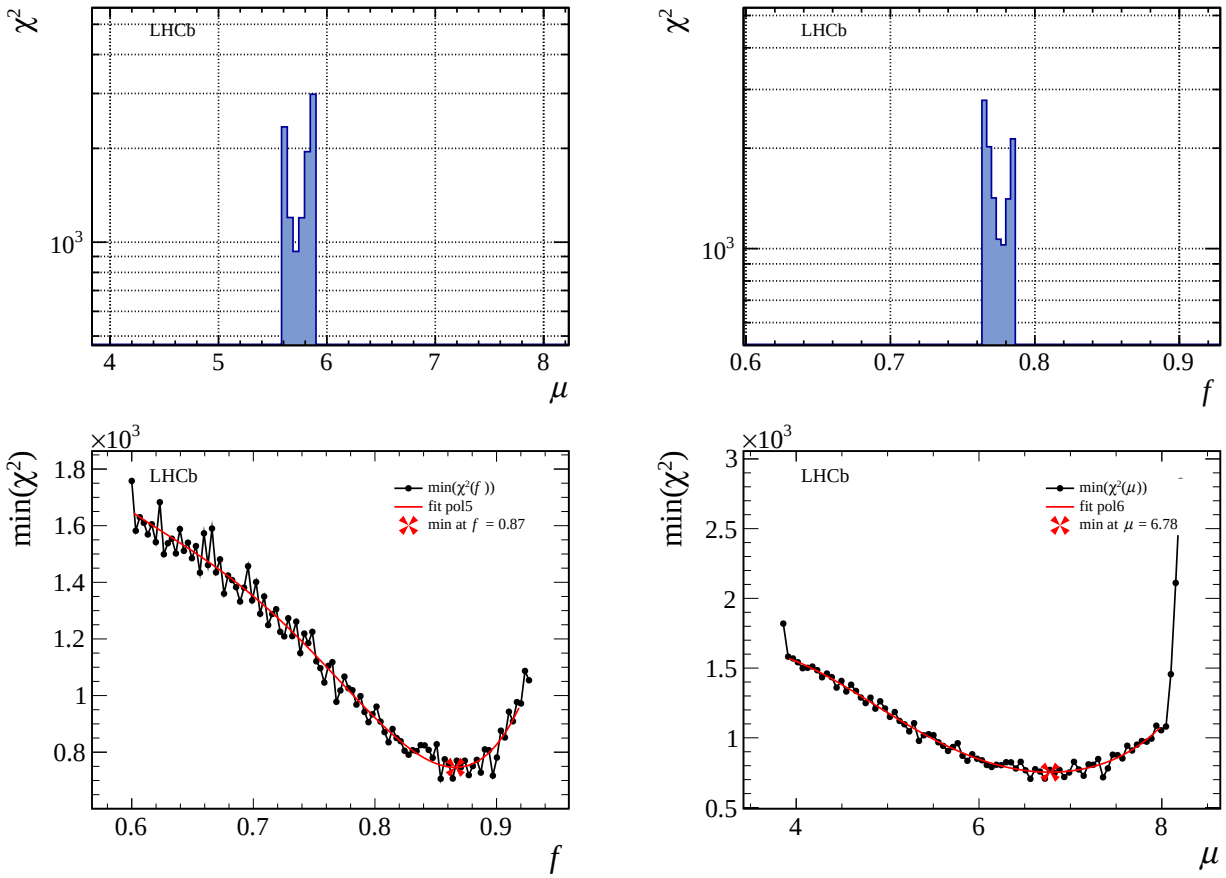

FIGURE 4.12: The top plots show a slice for $f = 0.798$ (*left*) and for $\mu = 5.45$ (*right*). From each slice the minimum of the histogram is kept. The bottom plots show the result of doing this for all values of f and μ, that is, the f-parametrised minimum (*left*) and the μ-parametrised minimum (*right*). These are fit by a 5^{th} and 6^{th} degree polynomial respectively whose minima are at $f = 0.866$ and $\mu = 6.778$.

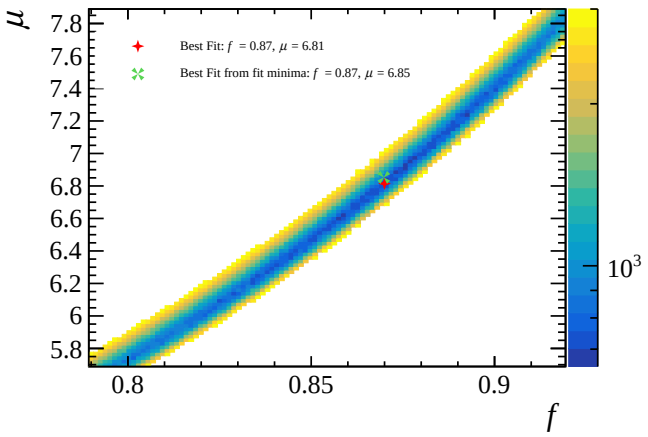

FIGURE 4.13: The χ^2 map for the grid search in $f \in [0.79, 0.92]$ and $\mu \in [5.7, 7.9]$. The two shown "best fits" correspond to the results from two different methods (see text).

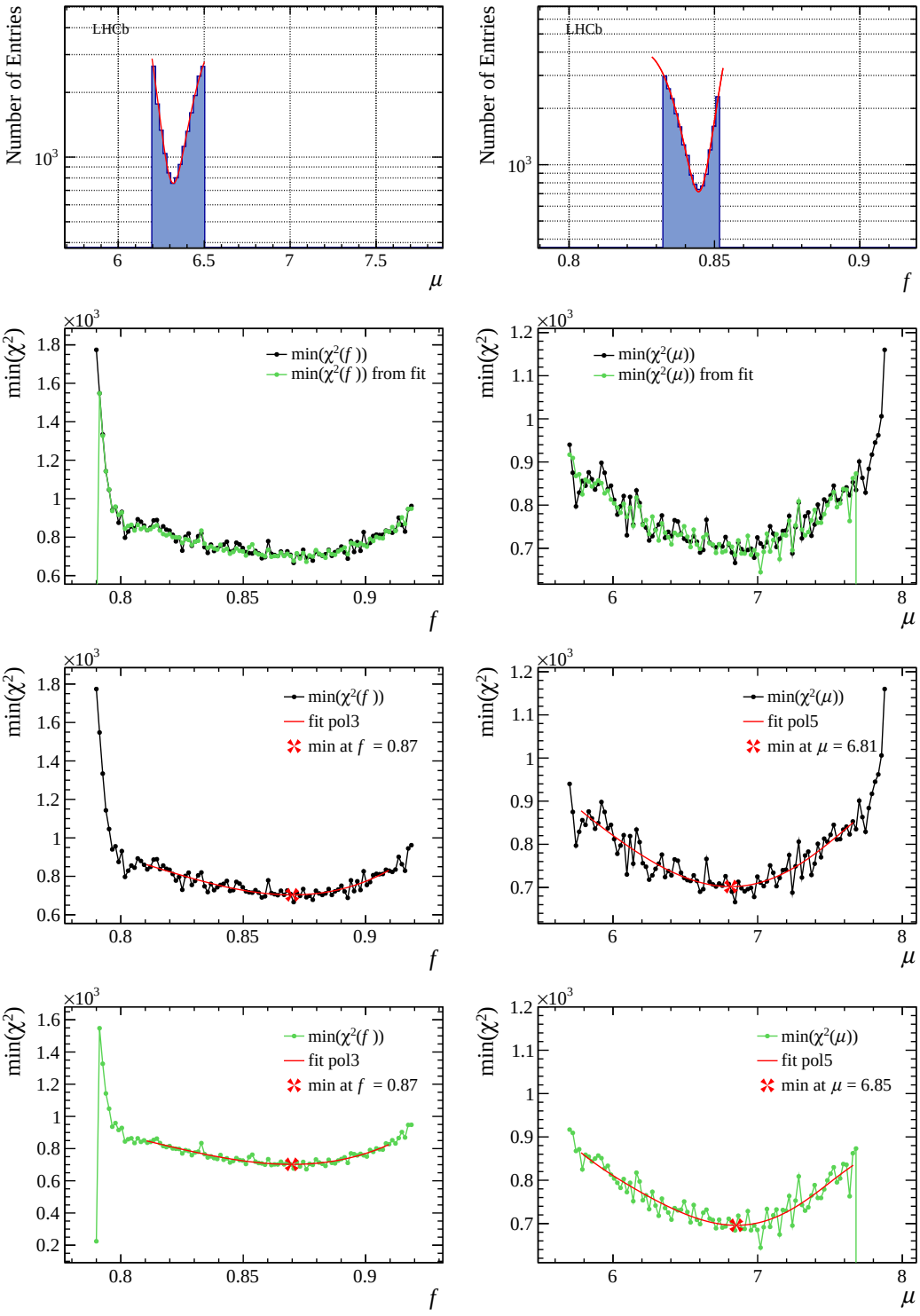

FIGURE 4.14: The top plots show a slice for $f = 0.840$ (*left*) and for $\mu = 6.382$ (*right*). The second row shows the comparison between the actual minima from the histograms and the minimum from the fit function. The third row shows the polynomial fits to the minimum parametrised by f and μ. The bottom row shows the polynomial fits to the minima of the fit slices.

μ the improvement is less evident. On the third row of Fig. 4.14 a fit to the f and μ-parametrised minimum from the slices has been performed (as was done for the previous grid), finding the best fit point at $(f, \mu) = (0.869, 6.814)$ with $\chi^2/\text{ndf} = 2.82$. Finally, in the bottom row the f and μ-parametrised minimum from the fit to the slices have also been fitted, finding the best fit at $(f, \mu) = (0.869, 6.853)$ with $\chi^2/\text{ndf} = 2.83$. These best fits correspond to the red star and the green cross from Fig. 4.13 respectively. Since the goodness of fit is virtually the same for the two best fits, the one which does not fit the slices histograms is kept, that is, $(f, \mu) = (0.869, 6.814)$.

The fitting procedure previously described can be summarised as follows:

1. Fit only the right shoulder shape (χ^2 computed from 35 to 52 TeV).
2. Define the (f, μ) grid and perform a grid search (χ^2 computed from 2 to 52 TeV).
3. From the coarse grid results, consider every slice in the f and μ directions:
   (a) Get the minimum of every slice and construct the f and μ-parametrised minimum.
   (b) Get the minimum of each parametrisation and keep that as the best fit.
4. Narrow parameter ranges around the best fit to make a finer grid.
5. From the fine grid results, consider every slice in the f and μ directions:
   (a) Get the minimum of every slice and construct the f and μ-parametrised minimum.
   (b) Get the minimum of each parametrisation and keep that as one best fit.
   (c) Get the minimum of the fit to every slice and construct the f and μ-parametrised minimum from the fitted slices.
   (d) Get the minimum of each parametrisation and keep that as another best fit.

The final result of the fit can be seen in Fig. 4.15. On the right plot of Fig. 4.15 a close-up of the low-energy part of the distribution can be seen, where the discrepancy between the MC Glauber and the data, due to the presence of events of electromagnetic origin, becomes clear. This will be addressed in more detail in section 4.4.2. However, it is worth mentioning that this discrepancy point is well outside the fitting range, which starts at 2 TeV. Now that the simulated distribution is already fit to the data, it can be used to define the centrality classes.

Defining centrality classes

After the simulated distribution of energy deposited in the ECAL is fitted, which now corresponds only to the hadronic contribution, it can be divided into the centrality classes. To determine the ECAL energy cut for each class the simulated distribution is integrated from a value of deposited energy (starting from the first bin of non-zero energy) to infinity, until a starting value for which a given percentage of the total integral is found. The total integral of the energy distribution is denoted I_{T}. The ECAL energy cut for the 90% centrality class is the value of E_{90} such that

$$0.9\, I_{\text{T}} = \int_{E_{90}}^{\infty} \frac{\text{dN}}{\text{d}E} \text{d}E, \tag{4.7}$$

and equivalently, to define the energy cut E_p for any percentage p of centrality

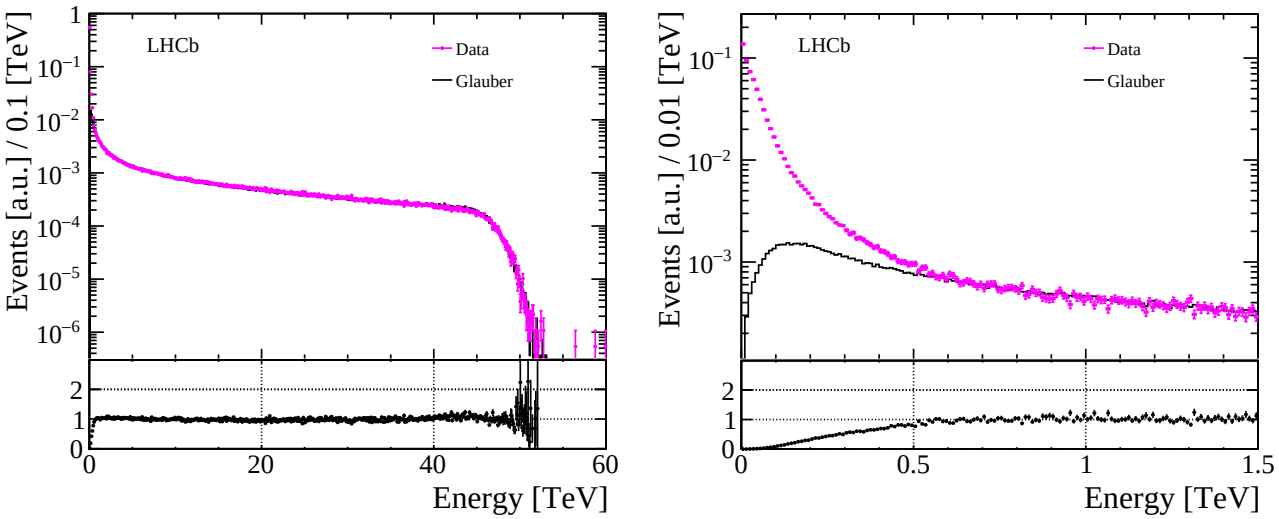

FIGURE 4.15: Final fit of the simulated energy distribution to the data. The best fit found was $(f, \mu) = (0.869, 6.814)$ with a corresponding $\chi^2/\text{ndf} = 2.82$. The right plot shows a close-up to the lower-energy part of the distribution.

$$(p \times 10^{-2})\, I_{\text{T}} = \int_{E_p}^{\infty} \frac{\text{dN}}{\text{d}E} \text{d}E, \tag{4.8}$$

then, for example, the centrality class $10 - 20\%$ corresponds to the events depositing an energy E such that $E_{20} < E < E_{10}$.

4.4.2 Results

The centrality classification of the MB dataset of run 217730 in percentile intervals of 10%, with the cuts in energy obtained as in the previously described procedure, is shown in Fig. 4.16 as well as the b, N_{part} and N_{coll} distributions for each class obtained from the MC Glauber model. For each class one can assign a mean number for each of the quantities of interest and their corresponding standard deviations.

In this way, one can define as many classes as desired and of arbitrary width in percentiles. The values of the geometric quantities for each class, as well as the corresponding energy cuts, can be seen in Tab. 4.1 for ten classes.

One important caveat to bear in mind is the fact that at low energy the events that dominate are of electromagnetic nature or from UPC events. Because of this, it is important to exclude in the analyses the energy region where there is a sizeable contamination from these events. If no further selection is applied to get rid of UPC events, their contamination will be below 5% at energies higher than 585 GeV, that is, at centralities lower than 84% (more central than 84%).

To determine this threshold, the data was compared to the fit MC Glauber as in Fig. 4.15 (see the plot on the right). The point from which the two distributions match was found by computing a centred mean of the Data/MC ratio around each bin, then when this ratio was below a chosen tolerance of 1.05 (meaning 5% contamination of UPC events) for three consecutive bins, the centre of the bin of lower energy was chosen as the energy threshold.[4]

[4] For a given bin n, the ratio of Data/MC was computed for bins $n-1$, n and $n+1$, and averaged.

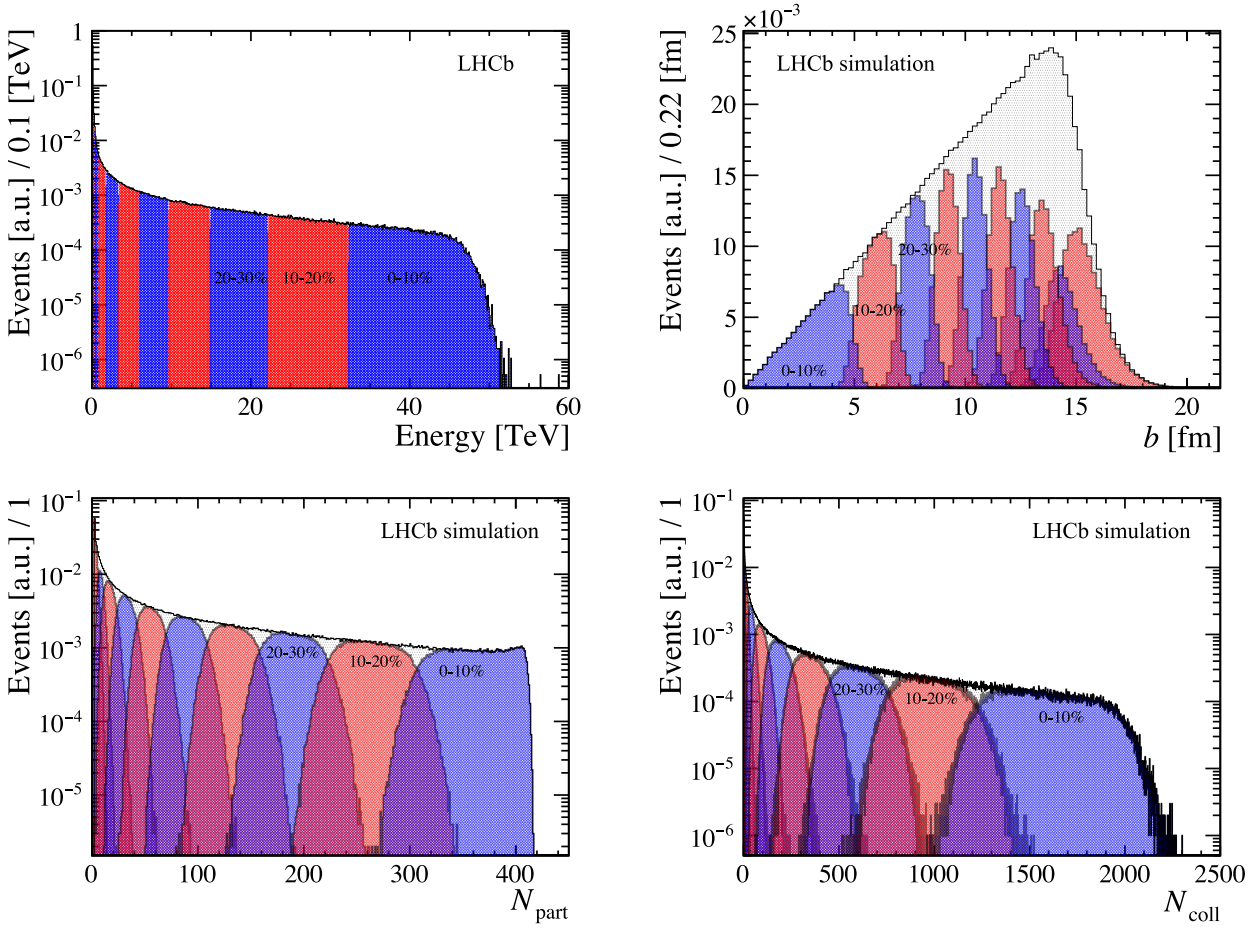

FIGURE 4.16: Classification of PbPb events from data according to the defined centrality classes (*top left*). Distribution of the impact parameter (*top right*), N_{part} (*bottom left*) and N_{coll} (*bottom right*) values for the corresponding centrality classes.

TABLE 4.1: Geometric quantities (N_{part}, N_{coll} and b) of PbPb collisions for centrality classes defined from a MC Glauber model fit to the data. The classes correspond to sharp cuts in the energy deposited in the ECAL. Here σ stands for the standard deviation of the corresponding distributions.

Centrality %	E [GeV]	N_{part}	$\sigma_{N_{\text{part}}}$	N_{coll}	$\sigma_{N_{\text{coll}}}$	b	σ_b
$100-90$	$0-310$	2.91	1.24	1.83	1.16	15.41	1.03
$90-80$	$310-800$	7.03	2.85	5.77	3.10	14.56	0.89
$80-70$	$800-1750$	15.92	4.75	16.44	7.04	13.59	0.70
$70-60$	$1750-3360$	31.26	7.11	41.28	14.72	12.61	0.59
$60-50$	$3360-5900$	54.65	9.98	92.59	27.70	11.59	0.54
$50-40$	$5900-9630$	87.54	13.25	187.54	46.66	10.47	0.51
$40-30$	$9630-14860$	131.24	16.91	345.53	71.62	9.23	0.51
$30-20$	$14860-22150$	188.02	21.51	593.92	105.22	7.80	0.55
$20-10$	$22150-32280$	261.84	27.10	972.50	151.91	6.02	0.66
$10-0$	$32280-\infty$	357.16	32.21	1570.26	236.80	3.31	1.20

If one is able to identify and take out UPC events from the data sample then this limit of 84% can be pushed to more peripheral events.

The results obtained for the geometrical quantities in the PbPb case are in very good agreement with the results obtained by the ALICE [132, 133], ATLAS [134] and CMS [135] collaborations at the same centre-of-mass energy.

4.5 PbPb centrality determination uncertainties

To compute uncertainties, some aspects that may affect the cuts on the deposited energy to define the centrality classes will be analysed. However, the focus is set not on how these aspects change the selections on the energy but on how they affect the geometric quantities of interest for each centrality class, that is, how the mean values of N_{part}, N_{coll} and b are affected for each class. Like this, the previously found energy selections are kept, but the systematic uncertainties can be quantified on the geometric properties. In what follows, the systematic uncertainties are reported in tables of ten classes of ten percentiles each. This is done for the sake of illustration of the concept and of clarity. However, in the tuple tool that is described in section 4.9, the systematic uncertainties are included for twenty classes of five percentiles each.

4.5.1 Bin-width dependence

To find the cuts in energy an integration procedure is performed on the histogram filled with the simulated energy deposit per event. The idea for a given percentage p is to find the bin from which integrating until the end, p% of the total integral is obtained. It is intuitive to realise that the broader the bins of the histogram are, the less precise the cut on the energy for that percentage will be.

The "true" energy cut for percentage p is denoted by E_{T}, which is an existing but unknown value, and the desired bin is denoted by B_p, which contains the energy E_{T}. When the desired bin B_p is found, the events that are being integrated are the ones which have a deposited energy greater than the low edge of B_p, which is denoted by E_{LE}. Clearly, if the binning is broad, the energy range in which E_{T} could be is big, conversely, if a finer binning is used, the range is narrower and thus E_{LE} would be closer to E_{T}. In theory, if the number n of bins in the histogram goes to infinity, then $E_{\text{LE}} \longrightarrow E_{\text{T}}$, but there is a limit in computation power and in the time used to perform the calculation. To deal with this situation, this effect is quantified by a "miss" percentage, and a reasonable criterion is defined to pick the desired binning to use. With this done, the percentage of events that leaks from one class to another and how this affects the geometric properties can be determined and assigned as a systematic uncertainty.

For the binning schemes to test, four different schemes were chosen, spanning three orders of magnitude. Since the energy histogram goes from 0 to 60 TeV, schemes of 300, 600, 6000 and 60 000 bins were considered. The results from these binning schemes can be seen in Tab. 4.2. Only values for ten different percentiles are shown to illustrate the effect, but the quantities were computed for every single percentile from 100 to 1.

From Tab. 4.2 it can be seen that the binning schemes of 6000 and 60 000 bins give reasonable mean "miss" percentages, less than 0.1%. Due to the computational cost of the procedure with 60 000 bins, the binning scheme of 6000 bins is chosen, which has an average discrepancy of 0.04%. Since each percentile corresponds to a 1% interval, an average miss of 0.04% means that on average 4% of the events of one percentile

TABLE 4.2: Real percentage selected for a given desired percentage p and the discrepancy referred to as "Miss %", for ten values of p. The cited mean of the miss % in the last row does not correspond to the mean of the table values but to the mean of every percentile from 100 to 1.

	300 bins		600 bins		6000 bins		60000 bins	
p%	Real %	Miss %	Real %	Miss %	Real %	Miss %	Real %	Miss %
90	87.45	2.55	87.45	2.55	89.92	0.08	89.98	0.02
80	79.92	0.08	79.92	0.08	79.92	0.08	79.98	0.02
70	69.53	0.47	69.53	0.47	69.92	0.08	69.99	0.01
60	59.79	0.21	59.79	0.21	59.99	0.01	60.00	0.00
50	49.64	0.36	49.97	0.03	49.97	0.03	50.00	0.00
40	39.60	0.40	39.83	0.17	39.99	0.01	40.00	0.00
30	29.77	0.23	29.93	0.07	29.99	0.01	30.00	0.00
20	19.93	0.07	19.93	0.07	19.99	0.01	20.00	0.00
10	9.90	0.10	9.98	0.02	10.00	0.00	10.00	0.00
Mean		0.77%		0.40%		0.04%		0.00%

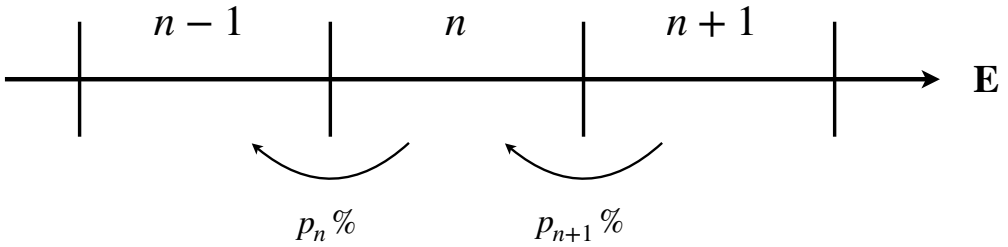

FIGURE 4.17: Schema of the meaning of "miss" percentages in the bins. In the figure p_n% of events that should be in bin n are considered to be in bin $n-1$.

"migrate" to the class immediately below in energy. The next step is to study the effect on the geometric quantities assigned to the centrality classes from the model.

A given general quantity is denoted by O. The true mean value of this quantity in a percentile interval n (the word "true" here refers to the value that would be found if there were no discrepancy between the desired percentiles and the real ones, *i.e.*, "miss" % equal to 0) is O_n^* and O_n the measured mean value in the same percentile, but this one is affected by the migration of events to and from other percentile intervals. On a given bin n, the measured value O_n is composed of the true value O_n^*, a gain due to the migration from bin $n+1$ to n of events which have a mean value of O_{n+1}^*, and a loss due to the migration of events from bin n to $n-1$ which have a mean value of O_{n-1}. The percentage of events of bin n which migrate to the bin $n-1$ is denoted by p_n%. A schema of this situation can be seen in Fig. 4.17.

The measured value O_n can be approximated as

$$\begin{aligned} O_n &= O_n^* - O_{n-1}\, p_n + O_{n+1}^*\, p_{n+1} \\ \Rightarrow O_n^* &= O_n + O_{n-1}\, p_n - O_{n+1}^*\, p_{n+1}, \end{aligned} \tag{4.9}$$

like this a recursive relation can be established, since the value O_n^* depends on the

TABLE 4.3: Systematic uncertainties due to the binning effect for the geometric quantities (N_{part}, N_{coll} and b) of PbPb collisions for centrality classes defined from a MC Glauber model fit to the data.

Centrality %	$N_{\text{part}} \pm \sigma^{\text{syst.}}_{bin}$	$N_{\text{coll}} \pm \sigma^{\text{syst.}}_{bin}$	$b \pm \sigma^{\text{syst.}}_{bin}$
$100-90$	2.91 ± 0.54	1.83 ± 0.34	15.41 ± 2.96
$90-80$	7.03 ± 0.78	5.77 ± 0.64	14.56 ± 1.80
$80-70$	15.92 ± 0.63	16.44 ± 0.67	13.59 ± 0.52
$70-60$	31.26 ± 0.66	41.28 ± 0.88	12.61 ± 0.28
$60-50$	54.65 ± 1.12	92.59 ± 1.91	11.59 ± 0.24
$50-40$	87.54 ± 0.99	187.54 ± 2.03	10.47 ± 0.14
$40-30$	131.24 ± 1.12	345.53 ± 2.88	9.23 ± 0.08
$30-20$	188.02 ± 1.45	593.92 ± 4.61	7.80 ± 0.06
$20-10$	261.84 ± 1.79	972.50 ± 6.55	6.02 ± 0.04
$10-0$	357.16 ± 1.66	1570.26 ± 7.34	3.31 ± 0.01

true value of the bin $n+1$.[5] If the maximum value that n can take is M, then the expressions that are useful to describe this recursive relation are

$$\begin{aligned} O^*_0 &= O_0 - O^*_1\, p_1 \\ O^*_1 &= O_1 + O_0\, p_1 - O^*_2\, p_2 \\ &\vdots \\ O^*_{M-1} &= O_{M-1} + O_{M-2}\, p_{M-1} - O^*_M\, p_M, \end{aligned} \tag{4.10}$$

and finally the value that allows to compute every term recursively

$$O^*_M = O_M + O_{M-1}\, p_M \tag{4.11}$$

If O^*_M is plugged into equation 4.10, an expression for O^*_{M-1} is obtained, which depends only on the observed quantities, which are known, resulting in

$$O^*_{M-1} = O_{M-2}\, p_{M-1} + O_{M-1}\left(1 - p^2_M\right) - O_M\, p_M, \tag{4.12}$$

then the process can be repeated for O^*_{M-2} and so on, until O^*_0 is reached. By computing the first few terms a general expression for the true values can be found recursively, valid for $n \in [1, M-2]$, which takes the form

$$\begin{aligned} O^*_n = {} & O_{n-1}\, p_n + O_n\left(1 - p^2_{n+1}\right) \\ & + \left\{ \sum_{j=1}^{M-n-1} (-1)^{M-n+j} \left(1 - p^2_{M-j+1}\right) O_{M-j} \prod_{k=n+1}^{M-j} p_k + (-1)^{M-n}\, O_M \prod_{l=n+1}^{M} p_l \right\}. \end{aligned} \tag{4.13}$$

For simplicity reasons, instead of using equation 4.13 for every term, O^*_M is computed numerically, plugged into O^*_{M-1}, and this is repeated until O^*_0.

[5] The reasoning behind this, is that if there were no migration then $O_n = O^*_n$, then it is necessary to account for the migrations. The events going to bin $n-1$ are used in the computation of the measured value O_{n-1} and thus it goes in the second term. The extra events coming from bin $n+1$ are part of the (unknown) true value, and thus O^*_{n+1} is used for the third term.

TABLE 4.4: Systematic uncertainties due to the hadronic cross-section uncertainty for the geometric quantities (N_{part}, N_{coll} and b) of PbPb collisions for centrality classes defined from a MC Glauber model fit to the data.

Centrality %	$N_{\text{part}} \pm \sigma^{\text{syst.}}_{had}$	$N_{\text{coll}} \pm \sigma^{\text{syst.}}_{had}$	$b \pm \sigma^{\text{syst.}}_{had}$
$100-90$	2.91 ± 0.02	1.83 ± 0.02	15.41 ± 0.02
$90-80$	7.03 ± 0.04	5.77 ± 0.06	14.56 ± 0.01
$80-70$	15.92 ± 0.07	16.44 ± 0.11	13.59 ± 0.01
$70-60$	31.26 ± 0.08	41.28 ± 0.26	12.61 ± 0.01
$60-50$	54.65 ± 0.14	92.59 ± 0.59	11.59 ± 0.01
$50-40$	87.54 ± 0.16	187.54 ± 1.24	10.47 ± 0.01
$40-30$	131.24 ± 0.23	345.53 ± 2.50	9.23 ± 0.01
$30-20$	188.02 ± 0.30	593.92 ± 4.65	7.80 ± 0.01
$20-10$	261.84 ± 0.33	972.50 ± 7.92	6.02 ± 0.01
$10-0$	357.16 ± 0.30	1570.26 ± 13.61	3.31 ± 0.00

To take this relation to the case at hand, O is taken to be the quantities of interest, that is, N_{part}, N_{coll} and b. For every percentile a mean value for these quantities is measured, and in addition the corresponding migration percentage for each percentile is known. The estimated true values for these quantities can be computed in every percentile, using the methodology described above, and the discrepancy with the measured values for each percentile can be obtained. Finally, the root-mean-square (RMS) of these discrepancies over consecutive sets of ten percentiles can be computed to assign this final value as a systematic uncertainty for each interval. The result can be seen in Tab. 4.3 for all three quantities.

4.5.2 Hadronic cross-section uncertainty

One of the main ingredients for the MC Glauber model was the nucleon-nucleon cross-section at a centre-of-mass energy of $\sqrt{s_{\text{NN}}} = 5.02\,\text{TeV}$, which corresponds to $\sigma^{\text{inel}}_{\text{NN}} = 67.6 \pm 0.6\,\text{mb}$, where the uncertainty comes from the data-driven parametrisation that was described in Ref. [119]. To quantify the effect this uncertainty has in the geometric properties, the simulated energy distribution is generated from the MC Glauber simulations made with $\sigma^{\text{inel}}_{\text{NN}} \pm 1\sigma$, *i.e.*, with $\sigma^{\text{inel}}_{\text{NN}} = 67.0\,\text{mb}$ and with $\sigma^{\text{inel}}_{\text{NN}} = 68.2\,\text{mb}$. Then the centrality classes are defined with these new distributions and the effect on the mean values for N_{part}, N_{coll} and b is taken as the associated systematic uncertainty.

The mean values obtained with $\sigma^{\text{inel}}_{\text{NN}} = 67.0\,\text{mb}$ are denoted by O^-_n, the mean values obtained with $\sigma^{\text{inel}}_{\text{NN}} = 68.2\,\text{mb}$ by O^+_n and the mean values obtained with $\sigma^{\text{inel}}_{\text{NN}} = 67.6\,\text{mb}$ by O_n, where n is the percentile number. For each percentile the values $D^-_n = O_n - O^-_n$ and $D^+_n = O_n - O^+_n$ are computed to then obtain $D_n = \sqrt{\left[(D^-_n)^2 + (D^+_n)^2\right]/2}$, which is the value assigned as uncertainty to each percentile. Finally, as was done for the binning uncertainty, the RMS of D_n for consecutive sets of ten percentiles is obtained. The resulting systematic uncertainties for N_{part}, N_{coll} and b coming from the hadronic cross-section uncertainty can be seen in Tab. 4.4.

TABLE 4.5: Systematic uncertainties due to the fitting procedure for the geometric quantities (N_{part}, N_{coll} and b) of PbPb collisions for centrality classes defined from a MC Glauber model fit to the data.

Centrality %	$N_{\text{part}} \pm \sigma_{fit}^{\text{syst.}}$	$N_{\text{coll}} \pm \sigma_{fit}^{\text{syst.}}$	$b \pm \sigma_{fit}^{\text{syst.}}$
$100-90$	2.91 ± 0.02	1.83 ± 0.02	15.41 ± 0.01
$90-80$	7.03 ± 0.04	5.77 ± 0.04	14.56 ± 0.01
$80-70$	15.92 ± 0.06	16.44 ± 0.08	13.59 ± 0.01
$70-60$	31.26 ± 0.07	41.28 ± 0.14	12.61 ± 0.01
$60-50$	54.65 ± 0.04	92.59 ± 0.17	11.59 ± 0.01
$50-40$	87.54 ± 0.11	187.54 ± 0.43	10.47 ± 0.01
$40-30$	131.24 ± 0.13	345.53 ± 0.65	9.23 ± 0.01
$30-20$	188.02 ± 0.15	593.92 ± 0.76	7.80 ± 0.00
$20-10$	261.84 ± 0.11	972.50 ± 1.11	6.02 ± 0.00
$10-0$	357.16 ± 0.20	1570.26 ± 1.38	3.31 ± 0.01

4.5.3 Fit uncertainty

One of the most important steps in the process of determining centrality, is the choice of the parameters f and μ. As a reminder, f is the parameter that mixes N_{part} with N_{coll} to give the number of particle producing sources N_{anc} (see eq. 4.2), and μ is a parameter of the NBD that is related to the mean value of this distribution, directly affecting how many particles are produced at each source (see eq. 4.4). Recalling from section 4.4.1, there were three best fits that were found. The first one, with a grid of 10 000 points encompassing the values of $f \in [0.60, 0.93]$ and $\mu \in [3.7, 9.0]$, which is now denoted by $\text{BF}^{\text{coarse}}$, the second one with a finer grid of 10 000 points encompassing the values of $f \in [0.79, 0.92]$ and $\mu \in [5.7, 7.9]$, which is the one that was kept and is now called BF^{*}, and finally, from the same fine grid, the one where the slices of f and μ were fitted with a 3^{rd} degree polynomial and the minimum of the fit function was used, which is now called BF^{fine}. These three best fits are used to compute the systematic uncertainties due to the choice of a given set of (f, μ). The values for these best fits were:

- $\text{BF}^{\text{coarse}} = (f = 0.866, \mu = 6.778)$
- $\text{BF}^{*} = (f = 0.869, \mu = 6.814)$
- $\text{BF}^{\text{fine}} = (f = 0.869, \mu = 6.853)$

In order to compute the uncertainty, the MC Glauber energy distribution is generated with the three sets of values, the centrality classes are defined for each case and finally the mean values of the geometric quantities (N_{part}, N_{coll} and b) are compared from $\text{BF}^{\text{coarse}}$ and BF^{fine} with respect to BF^{*}.

Similar to what was done before, the mean values obtained with $\text{BF}^{\text{coarse}}$ are denoted by O_n^{coarse}, the mean values obtained with BF^{fine} by O_n^{fine} and the mean values obtained with BF^{*} by O_n, where as before n is the percentile number. For each percentile the values $D_n^{\text{coarse}} = O_n - O_n^{\text{coarse}}$ and $D_n^{\text{fine}} = O_n - O_n^{\text{fine}}$ are computed to obtain $D_n = \sqrt{\left[(D_n^{\text{coarse}})^2 + (D_n^{\text{fine}})^2\right]/2}$, which is the value assigned as uncertainty to every percentile. Finally as was done for the previous uncertainties, the RMS of D_n is computed for consecutive sets of ten percentiles. The resulting systematic

TABLE 4.6: Systematic uncertainties due to the random sampling of the NBD for the geometric quantities (N_{part}, N_{coll} and b) of PbPb collisions for centrality classes defined from a MC Glauber model fit to the data.

Centrality %	$N_{\text{part}} \pm \sigma^{\text{syst.}}_{NBD}$	$N_{\text{coll}} \pm \sigma^{\text{syst.}}_{NBD}$	$b \pm \sigma^{\text{syst.}}_{NBD}$
$100-90$	2.91 ± 0.01	1.83 ± 0.01	15.41 ± 0.01
$90-80$	7.03 ± 0.02	5.77 ± 0.03	14.56 ± 0.01
$80-70$	15.92 ± 0.04	16.44 ± 0.06	13.59 ± 0.01
$70-60$	31.26 ± 0.05	41.28 ± 0.10	12.61 ± 0.01
$60-50$	54.65 ± 0.05	92.59 ± 0.14	11.59 ± 0.00
$50-40$	87.54 ± 0.07	187.54 ± 0.23	10.47 ± 0.00
$40-30$	131.24 ± 0.08	345.53 ± 0.35	9.23 ± 0.00
$30-20$	188.02 ± 0.09	593.92 ± 0.52	7.80 ± 0.00
$20-10$	261.84 ± 0.10	972.50 ± 0.64	6.02 ± 0.00
$10-0$	357.16 ± 0.10	1570.26 ± 0.73	3.31 ± 0.00

uncertainties for N_{part}, N_{coll} and b coming from this fit uncertainty can be seen in Tab. 4.5.

4.5.4 NBD uncertainty

Another unavoidable process in the generation of the simulated energy distribution is the random sampling of a NBD. This sampling introduces fluctuations that affect directly the observed χ^2 value when comparing the MC Glauber simulation and the data. This effect is noticeable even when the χ^2 is computed at the same point of the (f, μ) parameter space. To have an idea of how these fluctuations affect the final computed geometric quantities, ten extra simulated energy distributions are generated and the geometric quantities are computed for all of them.

For each percentile the standard deviation for N_{part}, N_{coll} and b was computed. Finally as was done before, the RMS of the standard deviations for consecutive sets of ten percentiles are obtained. The resulting systematic uncertainties for N_{part}, N_{coll} and b coming from this NBD uncertainty can be seen in Tab. 4.6.

4.5.5 Total systematic uncertainties

To put these uncertainties together they were added in quadrature for each centrality class, like this just one final systematic uncertainty is reported. The result can be seen in Tab. 4.7.

The uncertainties on the geometric quantities are dominated by the systematic uncertainties, as it would be expected, the dominant one being the uncertainty due to the binning effect.

These systematic uncertainties are included in the centrality tool developed for LHCb, presented in Sec. 4.9.

4.6 PbNe data selection for centrality determination

The same process that was used for the PbPb centrality determination is implemented for PbNe with a few particular differences. In order to do this, as before, the data must be selected to construct a PbNe clean MB data sample to tune the MC Glauber

TABLE 4.7: Total uncertainties for the geometric quantities (N_{part}, N_{coll} and b) of PbPb collisions for centrality classes defined from a MC Glauber model fit to the data. The statistical and systematic uncertainties are added in quadrature, and shown individually as well.

Centrality %	$N_{\text{part}} \pm \sigma^{\text{stat.}}_{\text{syst.}}$	$N_{\text{coll}} \pm \sigma^{\text{stat.}}_{\text{syst.}}$	$b \pm \sigma^{\text{stat.}}_{\text{syst.}}$
$100-90$	$2.91 \pm 0.54^{0.00}_{0.54}$	$1.83 \pm 0.34^{0.00}_{0.34}$	$15.41 \pm 2.96^{0.00}_{2.96}$
$90-80$	$7.03 \pm 0.78^{0.01}_{0.78}$	$5.77 \pm 0.64^{0.01}_{0.64}$	$14.56 \pm 1.80^{0.00}_{1.80}$
$80-70$	$15.92 \pm 0.64^{0.02}_{0.64}$	$16.44 \pm 0.69^{0.02}_{0.69}$	$13.59 \pm 0.52^{0.00}_{0.52}$
$70-60$	$31.26 \pm 0.67^{0.02}_{0.67}$	$41.28 \pm 0.93^{0.05}_{0.93}$	$12.61 \pm 0.28^{0.00}_{0.28}$
$60-50$	$54.65 \pm 1.13^{0.03}_{1.13}$	$92.59 \pm 2.01^{0.09}_{2.01}$	$11.59 \pm 0.24^{0.00}_{0.24}$
$50-40$	$87.54 \pm 1.01^{0.04}_{1.01}$	$187.54 \pm 2.43^{0.15}_{2.43}$	$10.47 \pm 0.14^{0.00}_{0.14}$
$40-30$	$131.24 \pm 1.15^{0.05}_{1.15}$	$345.53 \pm 3.89^{0.23}_{3.88}$	$9.23 \pm 0.08^{0.00}_{0.08}$
$30-20$	$188.02 \pm 1.49^{0.07}_{1.49}$	$593.92 \pm 6.62^{0.33}_{6.61}$	$7.80 \pm 0.06^{0.00}_{0.06}$
$20-10$	$261.84 \pm 1.83^{0.09}_{1.83}$	$972.50 \pm 10.37^{0.48}_{10.36}$	$6.02 \pm 0.04^{0.00}_{0.04}$
$10-0$	$357.16 \pm 1.70^{0.10}_{1.70}$	$1570.26 \pm 15.56^{0.75}_{15.54}$	$3.31 \pm 0.01^{0.00}_{0.01}$

simulation. This section is dedicated to the definition of an unbiased and pure PbNe sample.

The PbNe data correspond to collisions recorded during the whole 2018 data taking period. The PbNe collisions take place at a centre-of-mass energy of $\sqrt{s_{\text{NN}}} = 69\,\text{GeV}$ where the Ne atoms act as a fixed target. As it has been previously stated, throughout the whole period of data-taking there were PbPb collisions simultaneously being recorded. This poses some contamination to the data sample, and so it has to be cleaned to contain only PbNe collisions. The runs used for the analysis are the ones classified as fit for analysis in Chapter 3 and summarised in Tab. 3.6.

The first constraint that needs to be considered is that PbNe events of interest must take place within the VELO, specifically in the region $Z \in [-200, 200]$ mm since in this region the tracking efficiency is well understood and stable. The different topology between the PbPb and PbNe events, allows to disentangle these two types of events by setting an upper limit to the hits on the pile-up stations (`nPUHits`) which are located upstream from the nominal interaction point. Naturally since the PbNe collisions are all boosted downstream towards the detector, a very low `nPUHits` value is expected. On the other hand, since PbPb collisions are upstream-downstream symmetric, for these collisions a greater value for `nPUHits` is expected.

Another source of contamination are the same SMOG events that take place far upstream. Since gas can travel up to 20 m in either direction, the incoming Pb beam can undergo interactions with the gas long before arriving into the VELO tank. These events can shoot particles forward, hitting the pile-up stations, and depositing energy in the detector. The way to deal with these events is also to set an upper limit on the `nPUHits`.

To ensure the purity of PbNe collisions in the sample, the events are limited to have `nPUHits` < 1, which means to keep only events which did not produce any hits in the PU stations. The effect of this cut can be seen in Fig. 4.18. On the left 3 different populations can clearly be seen, the high-slope population which corresponds to the very upstream events (indicated by the red line), the middle-slope population

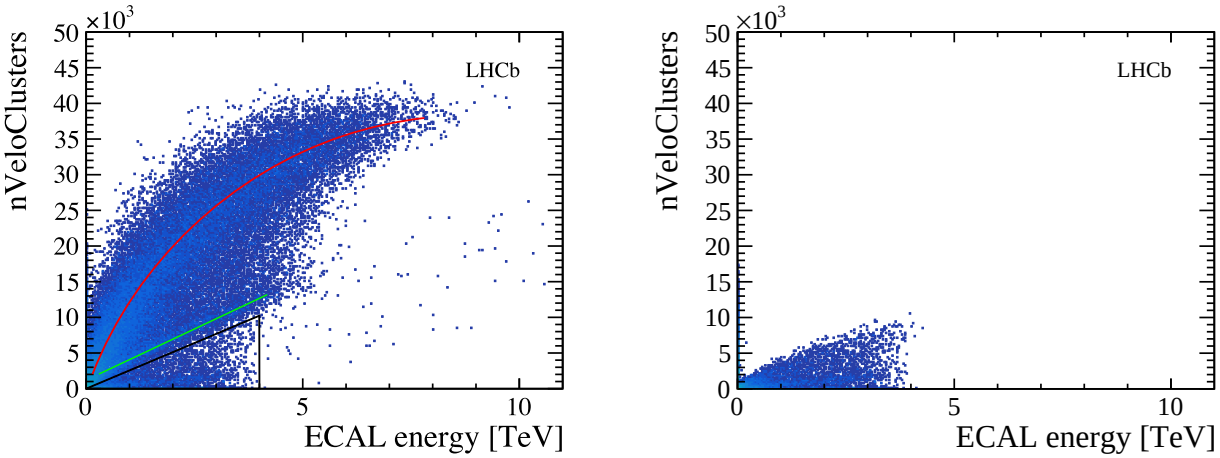

FIGURE 4.18: Number of VELO clusters as a function of ECAL energy for PbNe events without any requirement (*left*) and without any cluster in the PU stations (*right*). The red line indicates the population which corresponds to the very upstream events, the green line indicates the population which corresponds to ghost PbPb collisions and the black lines enclose the PbNe collisions of interest which present no clusters in the PU stations.

which corresponds to ghost PbPb collisions (indicated by the green line),[6] and finally, the continuum below PbPb which corresponds to the SMOG collisions of interest (enclosed by the black lines), which have been isolated in the plot on the right by requiring `nPUHits` < 1.

In order to justify the requirement of `nPUHits` < 1, and not some less stringent value, a closer look at the SMOG events from the right plot of Fig. 4.18 is needed. To understand the following, one prior caveat needs to be settled. Because of an issue during the reconstruction stage of the PbNe data, the pp configuration was used and events with more than 6000 `nVeloClusters` were not fully reconstructed (these are called "non-reconstructed" in the following). However, these events are needed to construct the MB sample for the centrality determination, and the only way to reliably use them, is to limit the possible contamination to these events. The events remaining after the `nPUHits` selection, are further divided into 4 groups; non-reconstructed events for which an additional requirement of having more than 300 hits in the SPD is applied, upstream events which have a primary vertex Z_{PV} position at $Z_{PV} < -200\,\text{mm}$, central events with a $Z_{PV} \in [-200, 200]\,\text{mm}$ and downstream events with a $Z_{PV} > 200\,\text{mm}$. This distinction is made because upstream and downstream events are outside the Z window of interest. The events displayed according to this classification can be seen in the Fig. 4.19.

From the left plot of Fig. 4.19 it is clear that the non-reconstructed events (red squares) are not downstream events (blue squares) given that they reside in very different regions of the phase space. The non-reconstructed events are mainly central-region events (black squares) with a small contamination from upstream events (green squares). Among the reconstructed events (green, black and blue) a few events that reside in the very low energy part and that have a high amount of `nVeloClusters` can be seen. These are fake events originating far from the passage of the Pb beam. Also being only about 20 events in the region $Z_{PV} \in [-200, 200]\,\text{mm}$, affecting only

[6] Ghost PbPb collisions occur when Pb atoms escape their nominal bunches within the beam, leading to collisions when there should be none.

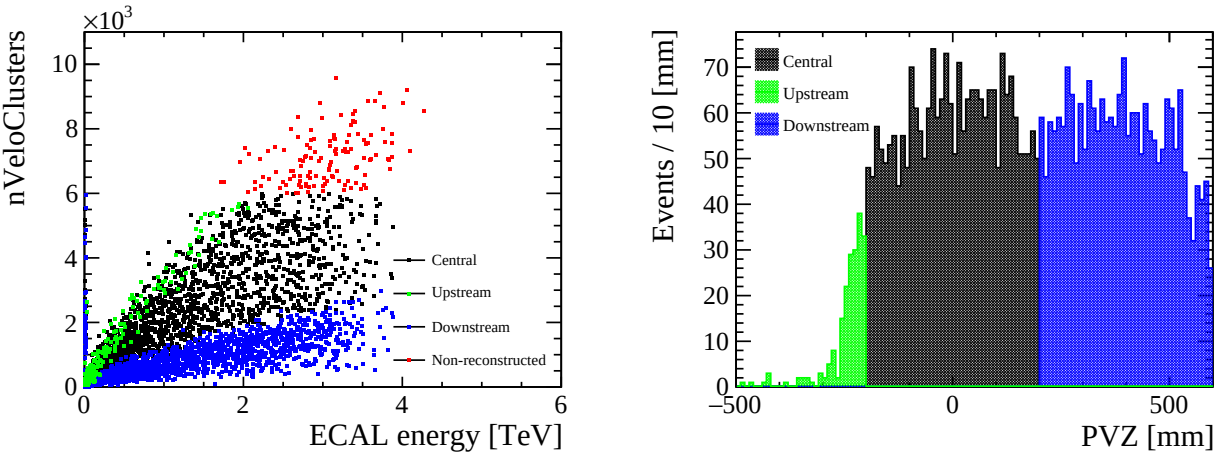

FIGURE 4.19: The event distribution in the `nVeloClusters` and ECAL energy plane with `nPUHits` < 1 (*left*) and their Z_{PV} distribution (*right*). Upstream events have a primary vertex at $Z_{PV} < -200\,\text{mm}$, central events at $Z_{PV} \in [-200, 200]\,\text{mm}$ and downstream events at $Z_{PV} > 200\,\text{mm}$.

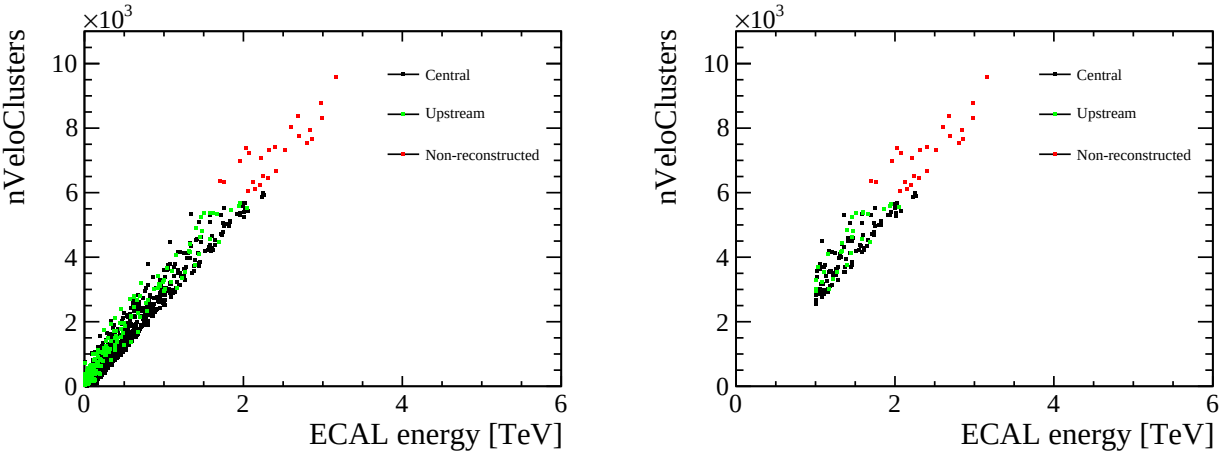

FIGURE 4.20: The event distribution in the `nVeloClusters` and ECAL energy plane within a selected region of space without an energy lower limit (*left*), and with ECAL energy deposit higher than 1 TeV (*right*).

the lowest energy bin in the energy distribution, they are not considered since as was mentioned before, there is an already expected discrepancy between the data and the MC Glauber at low energy. These types of events are also present in the non-reconstructed sample, but due to the kinematics of the non-reconstructed events, where all events from the bulk have more than 1000 hits on the SPD, events with `nSPDHits` < 300 are left out, thus obtaining the red squares from Fig. 4.19, free of these fake events.

To estimate the contamination of upstream events in the non-reconstructed events, a window is defined in the `nVeloClusters` versus ECAL energy plane that roughly encompasses only the space where upstream events lie. This window can be seen in Fig. 4.20 on the left plot. In addition, the window is restricted to events that have deposited more than 1 TeV in the ECAL (see right plot of Fig. 4.20) in order to have a better proxy of the events above 6000 `nVeloClusters`. From the reconstructed events (black and green squares) in this region, an estimate of the proportion of those

TABLE 4.8: Population in numbers of the 4 identified groups of events, for different limits applied on `nPUHits`. Also the contamination of non-reconstructed events by upstream events is shown.

`nPUHits` <	1	5	10	15
Upstream	205	556	1037	1465
Central	2330	2615	2623	2633
Downstream	2260	2469	2477	2481
Non-reconstructed	131	180	198	214
Upstream in Non-reconstructed	5.45	13.69	20.51	27.00
Contamination %	4.16	7.61	10.36	12.62

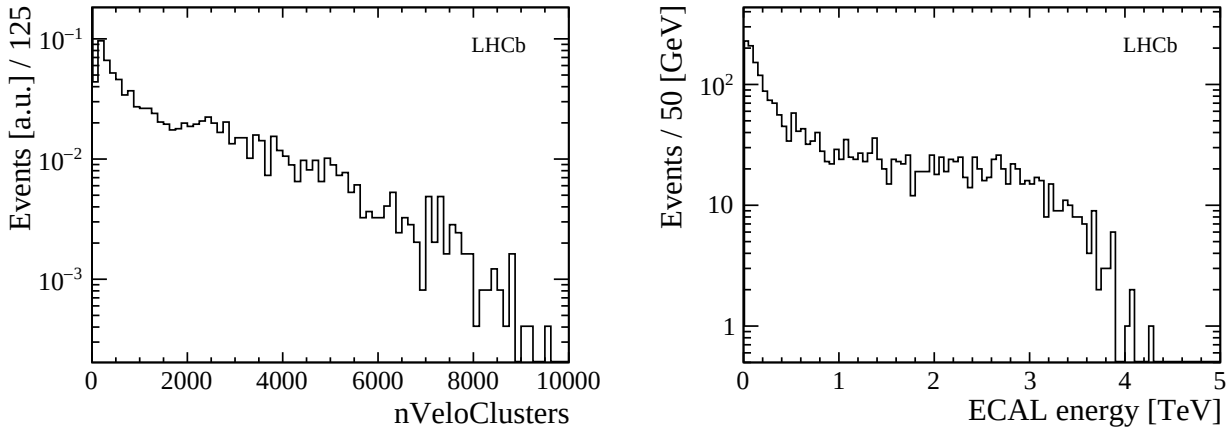

FIGURE 4.21: The number of VELO clusters (*left*) and the energy deposited in the ECAL (*right*) from PbNe collisions. The distributions correspond to the stacked distributions of central-region events and non-reconstructed events.

events in the non-reconstructed group within the window can be obtained. Then by knowing how many non-reconstructed events (red squares) are within this window, an approximate number of events that actually come from upstream collisions can be determined. With the requirement of `nPUHits` < 1 there are 29 upstream events, 104 central-region events and 25 non-reconstructed events in the restricted window. This leads to the estimate that 22% of the non reconstructed events in this window come from upstream collisions, meaning 5.45 events. This translates into a contamination of 4.16% in the total non-reconstructed events (red squares in the left plot of Fig. 4.19) by events that are upstream. A summary of these numbers and the corresponding quantities for other `nPUHits` limits are shown in Tab. 4.8.

From Tab. 4.8 it can be seen that by relaxing the requirement on `nPUHits` the gain that can be expected in statistics is not so significant while the contamination of the non-reconstructed events increases considerably. Hence a very stringent `nPUHits` requirement is needed and `nPUHits` < 1 is chosen.

Now that these groups of events have been properly characterised, only the central-region events and the non-reconstructed events will be used in what follows. Their `nVeloClusters` and ECAL energy distributions can be seen in Fig. 4.21.

To obtain the MB distribution shown in Fig. 4.21 it is required that:

- For central-region events:
  - ▷ `nPUHits` < 1

- ▷ $Z_{PV} \in [-200, 200]$ mm

- For non-reconstructed events:
  - ▷ `nPUHits` < 1
  - ▷ `nSPDHits` > 300 to exclude the very low energy events.
  - ▷ `nVeloClusters` > 6000

4.7 PbNe centrality determination

In the same way as it was done for PbPb the energy deposited in the ECAL is used as the variable to determine the centrality of the PbNe collisions. In this scenario the VELO does not saturate, but since the collisions do not occur only at the nominal collision point ($Z = 0$) and occur uniformly in an extended region several centimetres long, the VELO cannot be used for this purpose. The reason behind this is that a given event would look completely different if it takes place at $Z = -200$ mm, $Z = 0$ mm or $Z = 200$ mm from the point of view of the VELO.

Centrality classes are defined in the same manner as for PbPb but now as quantiles of the inelastic PbNe cross-section. Because of the same issue that the data contains contributions to the deposited energy in the ECAL from both hadronic and electromagnetic origin, the energy spectrum cannot be used straight away to define the desired quantiles for centrality and a MC Glauber must be used. From the resulting quantities the corresponding energy distribution from hadronic origin can be constructed and fit to the data.

The strategy and mapping procedure of the geometric quantities from the MC Glauber to the data remain unchanged from the PbPb case.

4.7.1 Methodology

Here, the collisions are simulated, the geometric quantities computed from the MC Glauber model and used to construct a simulated distribution of the energy deposited in the ECAL. This distribution is fitted to the data, centrality classes defined and the corresponding values of the geometric quantities assigned to each class.

Simulating the events

The first step of the way is to simulate the PbNe collisions. These are again simulated using the TGLAUBERMC software.

To model the collision of the two nuclei, Pb and Ne, their transverse shape needs to be considered. As for PbPb, the shape is determined by a nuclear density function, described by the 2pF, see eq. 4.1. For ^{20}Ne, the parameters R and a, take the values $R = 3.005$ fm [116, 117] and $a = 0.54$ fm [136–138]. The third parameter w is assumed to be 0. For ^{208}Pb the same parameters from section 4.4.1 were used.

One million collisions were simulated using the corresponding nucleon-nucleon cross-section $\sigma_{\mathrm{NN}}^{\mathrm{inel}} = 35.4 \pm 0.9$ mb for a centre-of-mass energy of $\sqrt{s_{\mathrm{NN}}} = 69$ GeV. After the events have been simulated, the N_{part} and N_{coll} of every simulated collision can be accessed and N_{anc} can be obtained (see eq. 4.2).

To get the distribution of particles coming out of the collision, N_{anc} is convoluted with a NBD. Since there are N_{anc} particle sources for each nucleus-nucleus collision, and each producing particles following a NBD, the NBD is sampled N_{anc} times to get the particle multiplicity distribution.

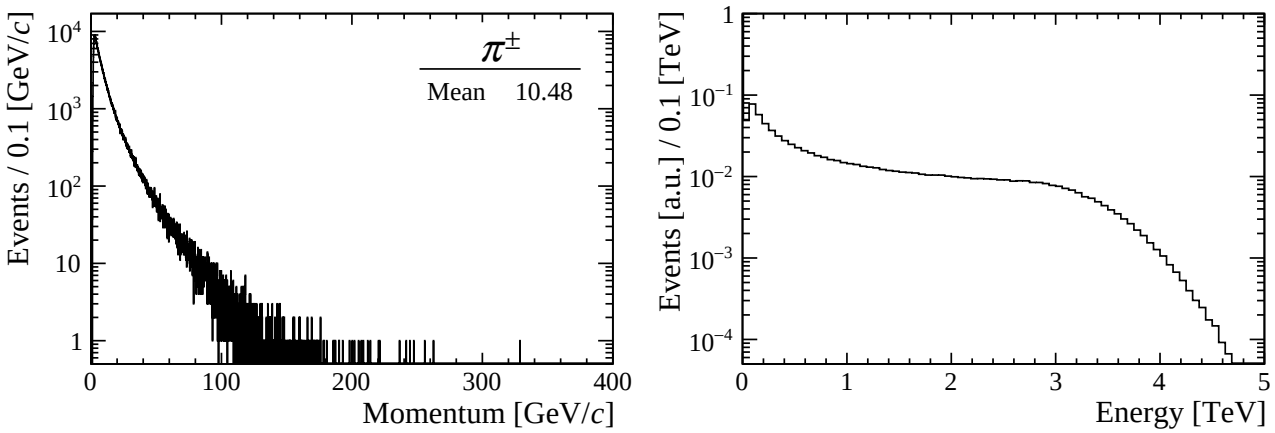

FIGURE 4.22: (*left*) Momentum distribution for charged-pion tracks. Data from 2017 pNe collisions at $\sqrt{s_{\mathrm{NN}}} = 69\,\mathrm{GeV}$. (*right*) Simulated energy distribution using the mean energy per particle obtained from the pion tracks.

As for the case of PbPb, the last ingredient needed to make a simulated energy distribution is the mean energy deposited by each of the outgoing particles in the ECAL, which is computed in the following section in the same way that was done for PbPb.

Mean energy deposit

To obtain the mean energy deposit per particle in the ECAL, the same principle as for PbPb was used. However, since there are no pp collisions at $\sqrt{s_{\mathrm{NN}}} = 69\,\mathrm{GeV}$ the most similar collision system available was used, that is, pNe collisions at $\sqrt{s_{\mathrm{NN}}} = 69\,\mathrm{GeV}$.

Since in this scenario the pions under study must come from SMOG collisions a few additional requirements are needed to have a clean sample. Only events of `BCType` $= 1$ were considered, meaning events for which nominally only beam 1 crosses the interaction point in the downstream direction and the corresponding bunch from beam 2 is empty. To exclude possible contamination from parasitic pp collisions due to debunched protons the primary vertices of the events are restricted to be in the range $Z_{PV} \in [-200, -150] \cup [150, 200]$ mm given that in the very central region an important contamination is expected [139], and it is additionally required that `nPUHits` < 1 and `nVeloClusters` > 0.

The momentum distribution of pions ($\pi^{\pm}$) can be seen in Fig. 4.22. Neutral pions account for about one third of the total pion yield. In the considered dataset $\sim 683 \times 10^3$ charged pions were found, which in turn means that there are about $\sim 341 \times 10^3$ neutral pions, whose energy distribution follows the same trend as for the charged pions.

Bearing this in mind, the mean energy deposited per particle in the ECAL is taken to be $\langle E \rangle = 10.48\,\mathrm{GeV}$, and this is the value that will be used in what follows. It is worth noting that this value is roughly the same as it was for the PbPb at $\sqrt{s_{\mathrm{NN}}} = 5.02\,\mathrm{TeV}$. However, the number of particle emitting sources (N_{anc}) and the average number of particles produced in each source are much lower in the case of PbNe, which leads to a much lower energy deposit in the ECAL. A simulated ECAL energy distribution can be seen in the right plot of Fig. 4.22, which resembles the shape found in data.

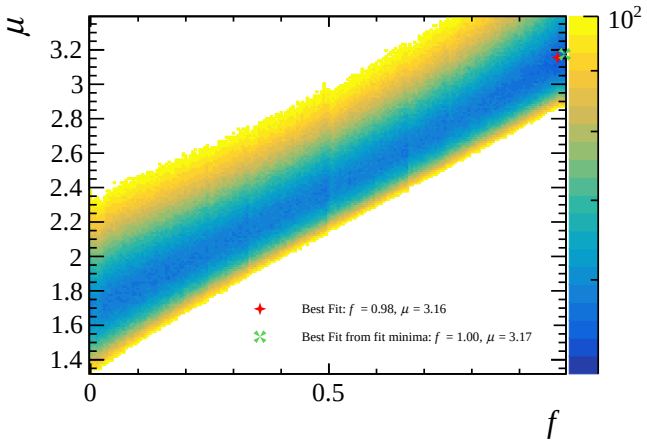

FIGURE 4.23: The χ^2 map for the grid search in $f \in [0.0, 1.0]$ and $\mu \in [1.0, 3.4]$. The two "best fits" shown correspond to the results of two different methods (see text).

Finalising the model

For PbNe the parameter k continues to be fixed to 1.5. This is because again there is no significant dependence of the final energy distribution on k. Finally the model has only two free parameters, f from eq. 4.2 and μ from the NBD.

Fitting to data

The same χ^2 function was used to evaluate the goodness of fit of the MC Glauber to the data. The fitting range was chosen to be from 0.5 to 3.9 TeV in order to avoid possible contamination from electromagnetic origin present at low energy. The MC Glauber energy distribution was normalised to the data in the energy range of 0.5 to 2 TeV to not consider the tails of the distributions, even in the extreme cases.

Since in this scenario the shape of the energy distribution at high energy is not as characteristic as it was for PbPb, the approach of trying to fit f first cannot be used, and the allowed range remains $f \in [0.0, 1.0]$. The range in μ is chosen accordingly to be $\mu \in [1.0, 3.4]$.

A grid of 200×200 was defined in the previously mentioned ranges, and the χ^2 was computed at every point of the grid. The result of this grid search can be seen in the Fig. 4.23.

Here to find the "best fits" shown in Fig. 4.23, both approaches described for the PbPb case were used. The f and μ-parametrised minima are obtained by taking the minimum value for every slice (first method, red star in Fig. 4.23), and by fitting each slice and getting the minimum of the fit (second method, green cross in Fig. 4.23). However, in the following step only the μ-parametrised minima was fit to get the optimal μ whereas the f-parametrised minima was not fit and only the f value where the χ^2 is minimum was picked. This is because the f-parametrised distributions show a steep decrease as f approaches 1, and fitting a function around that region is difficult. In Fig. 4.24 the parametrisations by f and μ of the minima can be seen. The black dots are the values obtained by looking at the actual minima of the slices and the green dots are the values obtained by fitting each slice and finding the minimum of this fit. From the black dots the best fit at $(f, \mu) = (0.980, 3.156)$ is found with a $\chi^2/\text{ndf} = 1.039$, and from the green dots the best fit at $(f, \mu) = (0.995, 3.174)$ is found

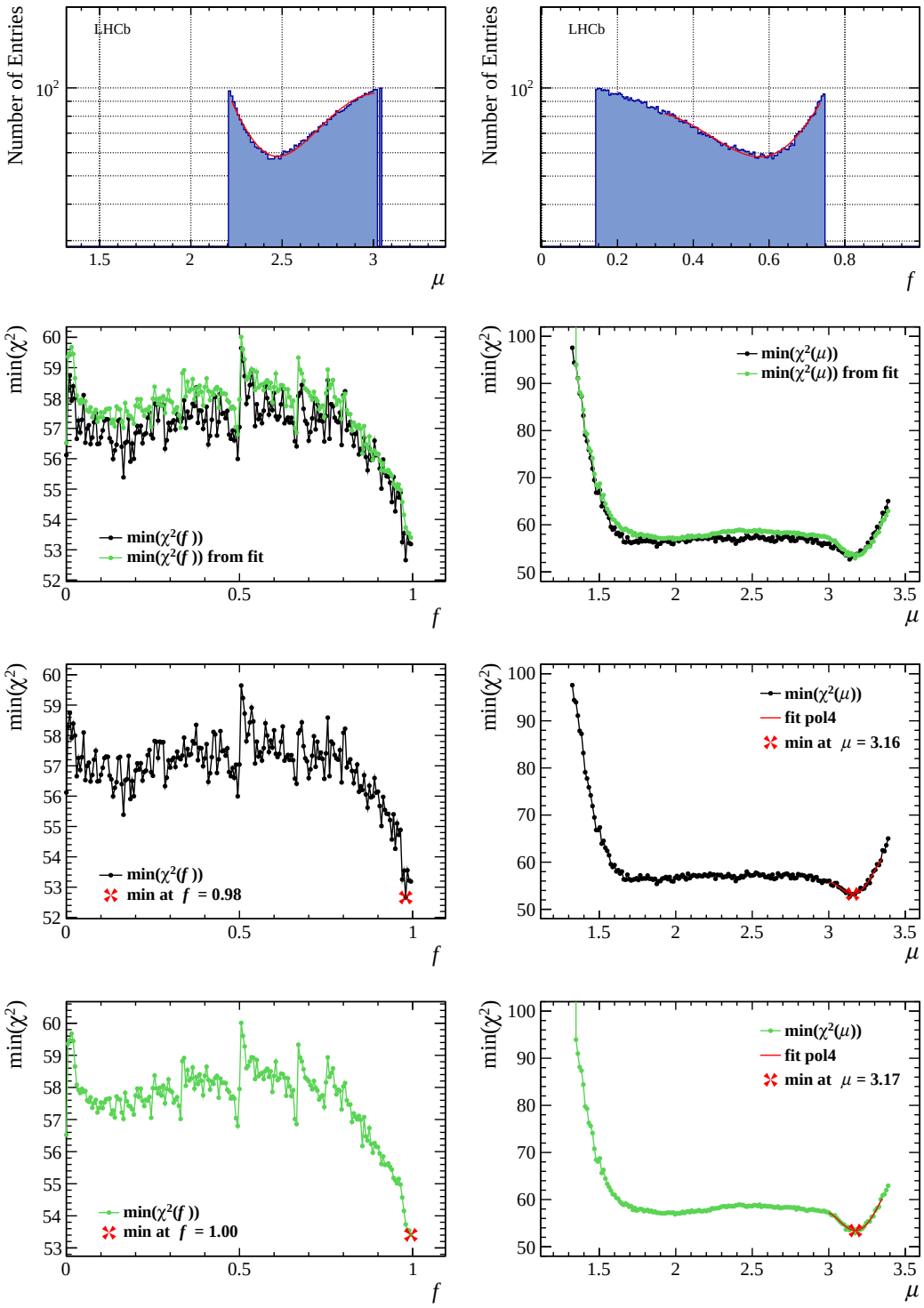

FIGURE 4.24: The top plots show a slice for $f = 0.546$ (*left*) and for $\mu = 2.512$ (*right*). The second row shows the comparison between the actual minima from the histograms (black) and the minimum from the fit function (green). The third row shows the minimum χ^2 reached by the f and μ-parametrised minimum. The bottom row shows the same but using the fits of the slices.

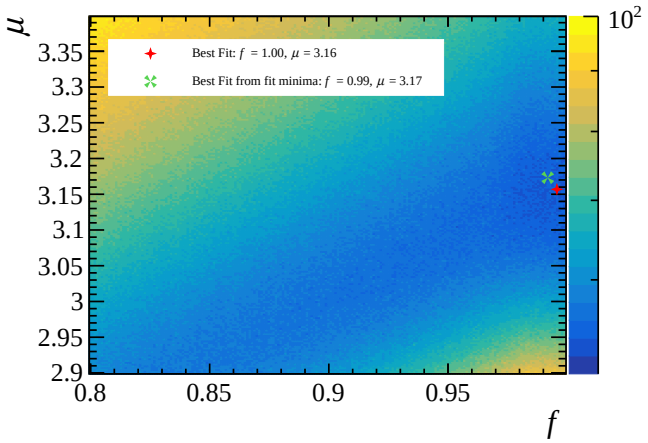

FIGURE 4.25: The χ^2 map for the grid search in $f \in [0.8, 1.0]$ and $\mu \in [2.9, 3.4]$. The two shown "best fits" correspond to the results from two different methods (see text).

with a $\chi^2/\text{ndf} = 1.031$. These best fits are shown in Fig. 4.23 by the red star and green cross respectively.

Finally another grid search was performed, the same amount of points was used but on a narrower range, namely $f \in [0.8, 1.0]$ and $\mu \in [2.9, 3.4]$. The resulting χ^2 map can be seen in Fig. 4.25. The two results that are shown as best fit are computed using the same previously described methods.

In Fig. 4.26 the parametrisations by f and μ of the minima in the fine grid can be seen. As before, the black dots are the values obtained by looking at the actual minima of the slices and the green dots are the values obtained by fitting each slice and finding the minimum of this fit. From the black dots the best fit at $(f, \mu) = (0.996, 3.157)$ is found with a $\chi^2/\text{ndf} = 1.026$, and from the green dots the best fit at $(f, \mu) = (0.992, 3.173)$ is found with a $\chi^2/\text{ndf} = 1.031$. These best fits are shown in Fig. 4.25 by the red star and green cross respectively. Since the goodness of fit is virtually the same for all best fits in the coarse and fine grid, the one which results in the smallest χ^2 value is kept, that is, $(f, \mu) = (0.996, 3.157)$.

The fitting procedure previously described can be summarised as follows:

1. Define the (f, μ) grid and perform a grid search (χ^2 computed from 0.5 to 3.9 TeV).
2. From the coarse grid results, consider every slice in the f and μ directions:
   (a) Get the minimum of every slice and construct the f and μ-parametrised minimum.
   (b) Get the minimum of each parametrisation and keep that as the best fit.
   (c) Get the minimum of the fit to every slice and construct the f and μ-parametrised minimum from the fitted slices.
   (d) Get the minimum of each parametrisation and keep that as another best fit.
3. Narrow parameter ranges around the best fits to make a finer grid.
4. From the fine grid results, consider every slice in the f and μ directions:

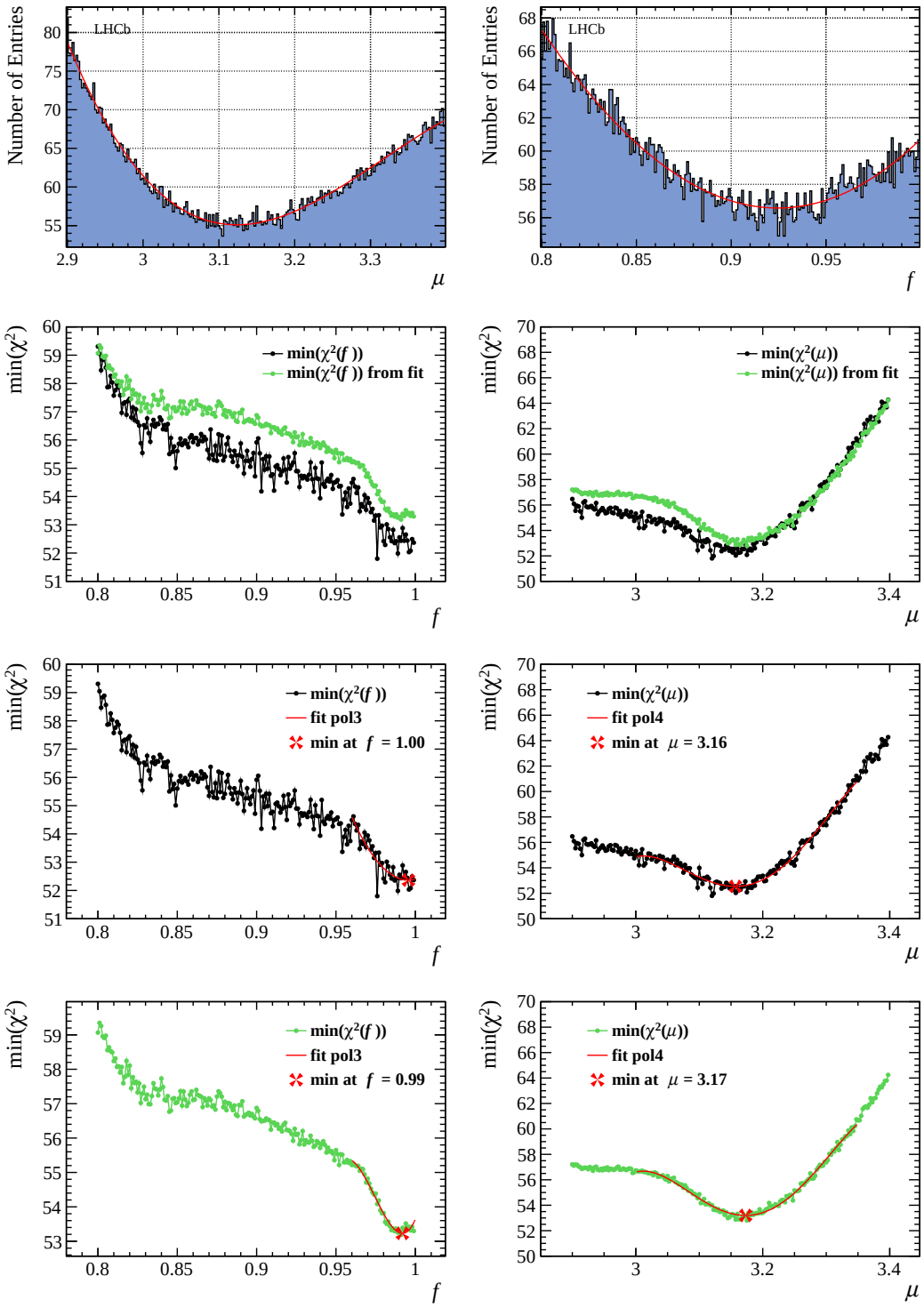

FIGURE 4.26: The top plots show a slice for $f = 0.969$ (left) and for $\mu = 3.045$ (right). The second row shows the comparison between the actual minima from the histograms (black) and the minimum from the fit function (green). The third row shows the polynomial fits to the minimum parametrised by f and μ. The bottom row shows the polynomial fits to the minima of the fit to the slices.

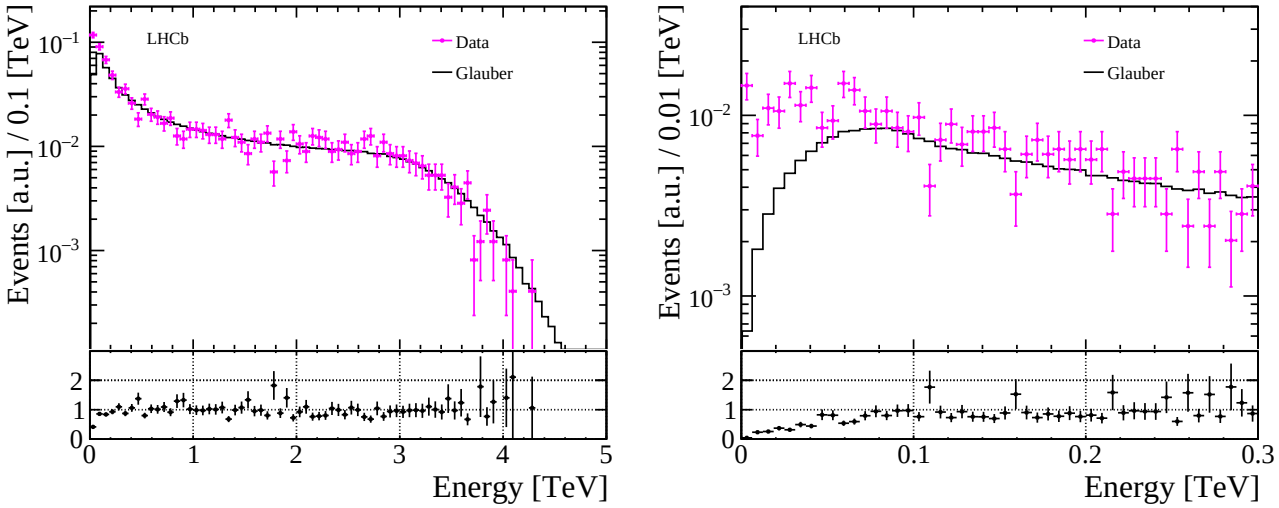

FIGURE 4.27: Final fit of the simulated energy distribution to the data. The best fit found was $(f, \mu) = (0.996, 3.157)$ with a corresponding $\chi^2/\text{ndf} = 1.026$. The right plot shows a close-up to the lower-energy part of the distribution.

(a) Get the minimum of every slice and construct the f and μ-parametrised minimum.

(b) Get the minimum of each parametrisation and keep that as the 3^{rd} best fit.

(c) Get the minimum of the fit to every slice and construct the f and μ-parametrised minimum from the fitted slices.

(d) Get the minimum of each parametrisation and keep that as the 4^{th} best fit.

The final result of the fit can be seen in Fig. 4.27. On the right plot of Fig. 4.27 a close-up of the low-energy part of the distribution can be seen, where the discrepancy between the MC Glauber and the data, due to the presence of events of electromagnetic origin, becomes clear. This discrepancy point is well outside the fitting range, which starts at 0.5 TeV.

Now that the simulated distribution is already fitted to the data, it can be used to define the centrality classes.

The fact that f is very close to 1 (hence $N_{\text{anc}} \sim N_{\text{part}}$) is reasonable since it is expected that below $\sqrt{s_{\text{NN}}} = 100\,\text{GeV}$ the particle production is dominated by soft processes, as was mentioned in section 4.4.1.

4.7.2 Results

The centrality classification of the MB dataset of PbNe collisions in percentile intervals of 10%, with the cuts in energy obtained in the same way as they were found for PbPb (see section 4.4.1), is shown in Fig. 4.28 as well as the b, N_{part} and N_{coll} distributions for each class obtained from the MC Glauber model. For each class one can assign a mean number for each of the quantities of interest and their corresponding standard deviations.

The values of the geometric quantities for each class, as well as the corresponding energy cuts, can be seen in Tab. 4.9 for ten classes.

As it was shown in the right plot of Fig. 4.27, for PbNe there are still events of electromagnetic nature or from UPC events at low energy. Because of this, it is

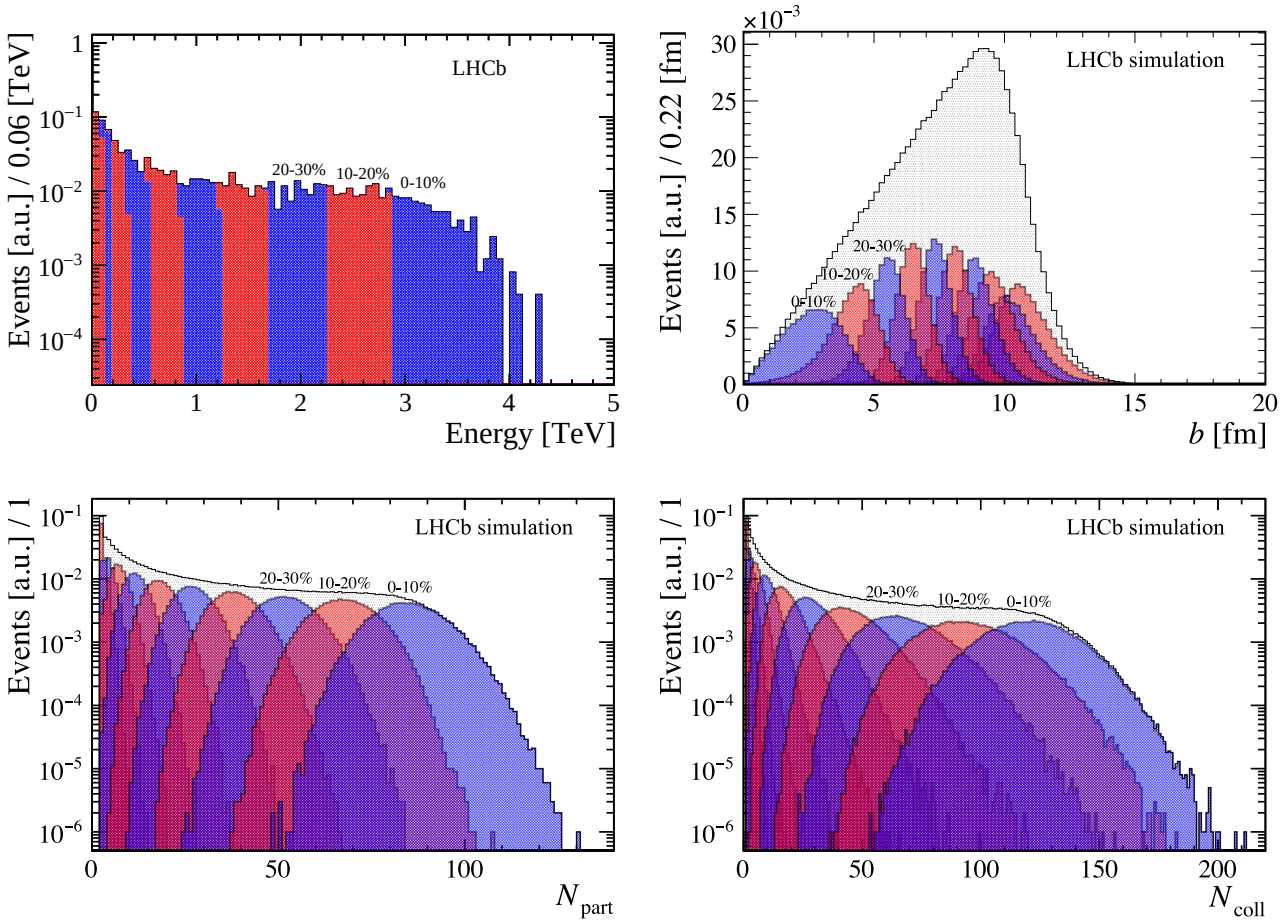

FIGURE 4.28: Classification of PbNe events from data according to the defined centrality classes (*top left*). Distribution of the impact parameter (*top right*), N_{part} (*bottom left*) and N_{coll} (*bottom right*) values for the corresponding centrality classes.

TABLE 4.9: Geometric quantities (N_{part}, N_{coll} and b) of PbNe collisions for centrality classes defined from a MC Glauber model fit to the data. The classes correspond to sharp cuts in the energy deposited in the ECAL. Here σ stands for the standard deviation of the corresponding distributions.

Centrality %	E [GeV]	N_{part}	$\sigma_{N_{\text{part}}}$	N_{coll}	$\sigma_{N_{\text{coll}}}$	b	σ_b
$100-90$	$0-94$	2.45	0.79	1.41	0.71	10.85	1.09
$90-80$	$94-184$	3.93	1.61	2.67	1.45	10.37	1.03
$80-70$	$184-324$	6.80	2.40	5.21	2.40	9.69	0.89
$70-60$	$324-533$	11.34	3.21	9.73	3.77	8.95	0.75
$60-50$	$532-828$	17.88	4.15	17.25	5.87	8.19	0.67
$50-40$	$828-1213$	26.72	5.20	28.95	8.68	7.38	0.63
$40-30$	$1213-1690$	37.99	6.32	45.62	12.25	6.48	0.65
$30-20$	$1690-2250$	51.72	7.48	67.78	16.10	5.44	0.76
$20-10$	$2250-2879$	67.30	8.26	94.10	18.88	4.14	1.01
$10-0$	$2879-\infty$	84.84	9.53	120.43	18.61	2.67	1.08

important to exclude in the analyses the energy region where there is a sizeable contamination from these events. If no further selection has been applied to get rid of UPC events, their contamination will be below 5% at energies higher than 98.9 GeV, that is at centralities lower than 89% (more central than 89%).

To determine this threshold, the data was compared to the fit MC Glauber as in Fig. 4.27 (see the plot on the right). The point from which the two distributions match was found by computing a centred mean of the Data/MC ratio around each bin, then when this ratio was below a chosen tolerance of 1.05 (meaning 5% contamination of UPC events) for three consecutive bins, the centre of the bin of lower energy was chosen as the energy threshold.[7]

If one is able to identify and take out UPC events from the data sample then this limit of 89% can be pushed to more peripheral events.

4.8 PbNe centrality determination uncertainties

For the uncertainty computations the same approach as for PbPb will be used, and the same sources of uncertainties will be considered. The principle is that the focus is set not on how the different sources change the cuts in energy but on how they affect the geometric quantities of interest for each centrality class, that is, how the mean N_{part}, N_{coll} and b are affected for each class. Like this the previously found energy cuts can be kept, but the systematic uncertainties on the geometric properties can be quantified. As for PbPb, in what follows the systematic uncertainties are reported in tables of ten classes of ten percentiles each. This is done for the sake of illustration of the concept and of clarity. However, in the software tool that will be described in section 4.9, the systematic uncertainties are included for twenty classes of five percentiles each in order to provide finer intervals and still have several percentiles from which to compute an average.

4.8.1 Bin-width dependence

For the centrality classification the binning scheme from PbPb is kept, namely 6000 bins. For PbPb this binning scheme resulted in an average miss percentage of 0.04%. For PbNe however, this same binning scheme results in an average miss percentage of 0.02%. The decrease in this miss percentage average is likely to come from the fact that the PbNe energy distribution does not present such a steep slope in the low energy part, as was the case for PbPb.

The same procedure considering the migration to and from each individual percentile was carried out for the PbNe simulated energy distribution to quantify how this migration affects the resulting geometric quantities.

Making use of the equations 4.10, 4.11 and the procedure described there, O is taken to be a quantity of interest, that is, either N_{part}, N_{coll} or b. For every percentile a mean value has been measured for these quantities, and in addition, the corresponding migration percentage for each percentile is known. The estimated true values for these quantities are computed in every percentile and the discrepancy with the measured values for each percentile is obtained. Finally, the RMS of these discrepancies is computed over consecutive sets of ten percentiles to assign this final value as a systematic uncertainty for each interval. The result can be seen in Tab. 4.10 for all three quantities.

[7] As opposed to the PbPb case, since there are less statistics, for a given bin n, the ratio of Data/MC was computed for the 5 precedent bins, for bin n, and for the 5 subsequent bins, and averaged.

TABLE 4.10: Systematic uncertainties due to the binning effect for the geometric quantities (N_{part}, N_{coll} and b) of PbNe collisions for centrality classes defined from a MC Glauber model fit to the data.

Centrality %	$N_{\text{part}} \pm \sigma^{\text{syst.}}_{bin}$	$N_{\text{coll}} \pm \sigma^{\text{syst.}}_{bin}$	$b \pm \sigma^{\text{syst.}}_{bin}$
$100-90$	2.45 ± 0.07	1.41 ± 0.04	10.85 ± 0.33
$90-80$	3.93 ± 0.14	2.67 ± 0.10	10.37 ± 0.41
$80-70$	6.80 ± 0.24	5.21 ± 0.19	9.69 ± 0.34
$70-60$	11.34 ± 0.26	9.73 ± 0.22	8.95 ± 0.22
$60-50$	17.88 ± 0.20	17.25 ± 0.19	8.19 ± 0.09
$50-40$	26.72 ± 0.21	28.95 ± 0.23	7.38 ± 0.06
$40-30$	37.99 ± 0.45	45.62 ± 0.54	6.48 ± 0.08
$30-20$	51.72 ± 0.33	67.78 ± 0.44	5.44 ± 0.03
$20-10$	67.30 ± 0.50	94.10 ± 0.70	4.14 ± 0.03
$10-0$	84.84 ± 0.46	120.43 ± 0.66	2.67 ± 0.01

4.8.2 Hadronic cross-section uncertainty

For PbNe collisions at a centre-of-mass energy of $\sqrt{s_{\text{NN}}} = 69\,\text{GeV}$, the nucleon-nucleon cross-section corresponds to $\sigma^{\text{inel}}_{\text{NN}} = 35.4 \pm 0.9\,\text{mb}$, where the uncertainty comes from the data-driven parametrisation that was described in Ref. [119]. To quantify the effect this uncertainty has in the geometric properties, the simulated energy distribution was generated from the MC Glauber simulations made with $\sigma^{\text{inel}}_{\text{NN}} = 34.5\,\text{mb}$ and with $\sigma^{\text{inel}}_{\text{NN}} = 36.3\,\text{mb}$. Then, the centrality classes were defined with these new distributions and the effect on the mean values for N_{part}, N_{coll} and b was quantified, taking this as the associated systematic uncertainty.

The mean values obtained with $\sigma^{\text{inel}}_{\text{NN}} = 34.5\,\text{mb}$ are again denoted by O^{-}_{n}, the mean values obtained with $\sigma^{\text{inel}}_{\text{NN}} = 36.3\,\text{mb}$ by O^{+}_{n} and the mean values obtained with $\sigma^{\text{inel}}_{\text{NN}} = 35.4\,\text{mb}$ by O_n, where n is the percentile number. For each percentile the values $D^{-}_{n} = O_n - O^{-}_{n}$ and $D^{+}_{n} = O_n - O^{+}_{n}$ are computed to then obtain $D_n = \sqrt{\left[(D^{-}_{n})^2 + (D^{+}_{n})^2\right]}/2$, which is the value assigned as uncertainty to each percentile. Finally, as was done for the binning uncertainty, the root mean square of D_n for consecutive sets of ten percentiles is obtained. The resulting systematic uncertainties for N_{part}, N_{coll} and b coming from the hadronic cross-section uncertainty can be seen in Tab. 4.11.

4.8.3 Fit uncertainty

As before, one important step in the centrality determination, is the choice of the parameters f and μ. Recalling from section 4.7.1, there were four best fits found. The first two, with a grid of 40 000 points encompassing the values of $f \in [0.0, 1.0]$ and $\mu \in [1.0, 3.4]$, denoted by $\text{BF}^{\text{coarse}}_{1}$ and $\text{BF}^{\text{coarse}}_{2}$. The last two with a finer grid of 40 000 points encompassing the values of $f \in [0.8, 1.0]$ and $\mu \in [2.9, 3.4]$, of which one is the one kept and now denoted byBF^{*}, and the other from the same fine grid now denoted by BF^{fine}. These four best fits are used to compute the systematic uncertainties due to the choice of a given set of (f, μ). The values for these best fits were:

- $\text{BF}^{\text{coarse}}_{1} = (f = 0.980, \mu = 3.156)$
- $\text{BF}^{\text{coarse}}_{2} = (f = 0.995, \mu = 3.174)$

TABLE 4.11: Systematic uncertainties due to the hadronic cross-section uncertainty for the geometric quantities (N_{part}, N_{coll} and b) of PbNe collisions for centrality classes defined from a MC Glauber model fit to the data.

Centrality %	$N_{\text{part}} \pm \sigma^{\text{syst.}}_{had}$	$N_{\text{coll}} \pm \sigma^{\text{syst.}}_{had}$	$b \pm \sigma^{\text{syst.}}_{had}$
$100-90$	2.45 ± 0.01	1.41 ± 0.01	10.85 ± 0.02
$90-80$	3.93 ± 0.03	2.67 ± 0.03	10.37 ± 0.02
$80-70$	6.80 ± 0.05	5.21 ± 0.06	9.69 ± 0.02
$70-60$	11.34 ± 0.09	9.73 ± 0.14	8.95 ± 0.02
$60-50$	17.88 ± 0.14	17.25 ± 0.30	8.19 ± 0.02
$50-40$	26.72 ± 0.21	28.95 ± 0.56	7.38 ± 0.01
$40-30$	37.99 ± 0.31	45.62 ± 0.95	6.48 ± 0.01
$30-20$	51.72 ± 0.44	67.78 ± 1.52	5.44 ± 0.01
$20-10$	67.30 ± 0.58	94.10 ± 2.19	4.14 ± 0.01
$10-0$	84.84 ± 0.86	120.43 ± 2.95	2.67 ± 0.01

TABLE 4.12: Systematic uncertainties due to the fitting procedure for the geometric quantities (N_{part}, N_{coll} and b) of PbNe collisions for centrality classes defined from a MC Glauber model fit to the data.

Centrality %	$N_{\text{part}} \pm \sigma^{\text{syst.}}_{fit}$	$N_{\text{coll}} \pm \sigma^{\text{syst.}}_{fit}$	$b \pm \sigma^{\text{syst.}}_{fit}$
$100-90$	2.45 ± 0.01	1.41 ± 0.01	10.85 ± 0.01
$90-80$	3.93 ± 0.02	2.67 ± 0.02	10.37 ± 0.02
$80-70$	6.80 ± 0.02	5.21 ± 0.02	9.69 ± 0.01
$70-60$	11.34 ± 0.04	9.73 ± 0.06	8.95 ± 0.01
$60-50$	17.88 ± 0.04	17.25 ± 0.05	8.19 ± 0.01
$50-40$	26.72 ± 0.04	28.95 ± 0.08	7.38 ± 0.01
$40-30$	37.99 ± 0.06	45.62 ± 0.13	6.48 ± 0.01
$30-20$	51.72 ± 0.11	67.78 ± 0.20	5.44 ± 0.01
$20-10$	67.30 ± 0.13	94.10 ± 0.21	4.14 ± 0.02
$10-0$	84.84 ± 0.10	120.43 ± 0.25	2.67 ± 0.02

- $\text{BF}^{*} = (f = 0.996, \mu = 3.157)$
- $\text{BF}^{\text{fine}} = (f = 0.992, \mu = 3.173)$

In order to compute this uncertainty, the MC Glauber energy distribution is generated with the four sets of values. The centrality classes are defined for each case and finally, the mean values of the geometric quantities (N_{part}, N_{coll} and b) from $\text{BF}^{\text{coarse}}_1$, $\text{BF}^{\text{coarse}}_2$ and BF^{fine} are compared to the ones from BF^{*}.

Similar to what was done before, the mean values obtained with $\text{BF}^{\text{coarse}}_1$ is denoted by $O^{\text{coarse},1}_n$, the mean values obtained with $\text{BF}^{\text{coarse}}_2$ by $O^{\text{coarse},2}_n$, the mean values obtained with BF^{fine} by O^{fine}_n and the mean values obtained with BF^{*} by O_n, where as before n is the percentile number. For each percentile the values $D^{\text{coarse},1}_n = O_n - O^{\text{coarse},1}_n$, $D^{\text{coarse},2}_n = O_n - O^{\text{coarse},2}_n$ and $D^{\text{fine}}_n = O_n - O^{\text{fine}}_n$ are computed to then obtain $D_n = \sqrt{\left[(D^{\text{coarse},1}_n)^2 + D^{\text{coarse},2}_n)^2 + (D^{\text{fine}}_n)^2\right]/3}$, which is the

TABLE 4.13: Systematic uncertainties due to the random sampling of the NBD for the geometric quantities (N_{part}, N_{coll} and b) of PbNe collisions for centrality classes defined from a MC Glauber model fit to the data.

Centrality %	$N_{\text{part}} \pm \sigma^{\text{syst.}}_{NBD}$	$N_{\text{coll}} \pm \sigma^{\text{syst.}}_{NBD}$	$b \pm \sigma^{\text{syst.}}_{NBD}$
$100-90$	2.45 ± 0.01	1.41 ± 0.01	10.85 ± 0.01
$90-80$	3.93 ± 0.01	2.67 ± 0.01	10.37 ± 0.01
$80-70$	6.80 ± 0.02	5.21 ± 0.02	9.69 ± 0.01
$70-60$	11.34 ± 0.02	9.73 ± 0.03	8.95 ± 0.01
$60-50$	17.88 ± 0.03	17.25 ± 0.04	8.19 ± 0.01
$50-40$	26.72 ± 0.04	28.95 ± 0.07	7.38 ± 0.01
$40-30$	37.99 ± 0.05	45.62 ± 0.10	6.48 ± 0.01
$30-20$	51.72 ± 0.06	67.78 ± 0.14	5.44 ± 0.01
$20-10$	67.30 ± 0.07	94.10 ± 0.16	4.14 ± 0.01
$10-0$	84.84 ± 0.05	120.43 ± 0.15	2.67 ± 0.01

value assigned as uncertainty to every percentile. Finally, as was done for the previous uncertainties, the RMS of D_n are obtained for consecutive sets of ten percentiles. The resulting systematic uncertainties for N_{part}, N_{coll} and b coming from this fit uncertainty can be seen in Tab. 4.12.

4.8.4 NBD uncertainty

It has been discussed that an unavoidable process in the generation of the simulated energy distribution is the random sampling of a NBD. This sampling introduces fluctuations that affect directly the observed χ^2 value when comparing the MC Glauber simulation and the data. To have an idea of how these fluctuations affect the final computed geometric quantities, ten simulated energy distributions are generated and the geometric quantities are computed for all of them.

For each percentile the standard deviation for N_{part}, N_{coll} and b was computed. Finally, as was done before, the RMS of the standard deviations for consecutive sets of ten percentiles is obtained. The resulting systematic uncertainties for N_{part}, N_{coll} and b coming from this NBD uncertainty can be seen in Tab. 4.13.

4.8.5 Total systematic uncertainties

To put these uncertainties together they are added in quadrature for each centrality class, like this just one final systematic uncertainty is reported. The result can be seen in Tab. 4.14.

The uncertainties on the geometric quantities are dominated by the systematic uncertainties, the dominant being the uncertainty due to the binning effect in more peripheral collisions (centrality higher than 50%) and the uncertainty due to the hadronic cross-section uncertainty for more central events (centrality lower than 50%).

These systematic uncertainties are included in the centrality tool developed for LHCb, presented in Sec. 4.9.

TABLE 4.14: Total uncertainties for the geometric quantities (N_{part}, N_{coll} and b) of PbNe collisions for centrality classes defined from a MC Glauber model fit to the data. The statistical and systematic uncertainties are added in quadrature, and shown individually as well.

Centrality %	$N_{\text{part}} \pm \sigma^{\text{stat.}}_{\text{syst.}}$	$N_{\text{coll}} \pm \sigma^{\text{stat.}}_{\text{syst.}}$	$b \pm \sigma^{\text{stat.}}_{\text{syst.}}$
$100-90$	$2.45 \pm 0.07^{0.00}_{0.07}$	$1.41 \pm 0.04^{0.00}_{0.04}$	$10.85 \pm 0.33^{0.00}_{0.33}$
$90-80$	$3.93 \pm 0.15^{0.01}_{0.14}$	$2.67 \pm 0.11^{0.00}_{0.11}$	$10.37 \pm 0.41^{0.00}_{0.41}$
$80-70$	$6.80 \pm 0.25^{0.01}_{0.25}$	$5.21 \pm 0.20^{0.01}_{0.20}$	$9.69 \pm 0.34^{0.00}_{0.34}$
$70-60$	$11.34 \pm 0.28^{0.01}_{0.28}$	$9.73 \pm 0.27^{0.01}_{0.27}$	$8.95 \pm 0.22^{0.00}_{0.22}$
$60-50$	$17.88 \pm 0.25^{0.01}_{0.25}$	$17.25 \pm 0.36^{0.02}_{0.36}$	$8.19 \pm 0.09^{0.00}_{0.09}$
$50-40$	$26.72 \pm 0.30^{0.02}_{0.30}$	$28.95 \pm 0.62^{0.03}_{0.61}$	$7.38 \pm 0.06^{0.00}_{0.06}$
$40-30$	$37.99 \pm 0.55^{0.02}_{0.55}$	$45.62 \pm 1.11^{0.04}_{1.10}$	$6.48 \pm 0.08^{0.00}_{0.08}$
$30-20$	$51.72 \pm 0.56^{0.02}_{0.56}$	$67.78 \pm 1.60^{0.05}_{1.60}$	$5.44 \pm 0.03^{0.00}_{0.03}$
$20-10$	$67.30 \pm 0.78^{0.03}_{0.78}$	$94.10 \pm 2.32^{0.06}_{2.31}$	$4.14 \pm 0.04^{0.00}_{0.04}$
$10-0$	$84.84 \pm 0.98^{0.03}_{0.98}$	$120.43 \pm 3.04^{0.06}_{3.04}$	$2.67 \pm 0.03^{0.00}_{0.03}$

4.9 A tuple tool for LHCb

The centrality information obtained for the PbPb and the PbNe case has been included in the form of a tuple tool named `TupleToolCentrality` into the main LHCb analysis software. The tool works as a look-up table, where the values for the energy cuts of the 100 percentiles are stored, along with the values of $\langle N_{\text{part}} \rangle$, $\sigma_{N_{\text{part}}}$, $\langle N_{\text{coll}} \rangle$, $\sigma_{N_{\text{coll}}}$, $\langle b \rangle$ and σ_b for all possible centrality classes and for PbPb and PbNe collisions at their respective energies. The energy cuts are used to classify the events into their corresponding percentiles. Then, the information about the resulting percentile and the associated geometric quantities are written into the output tuple. The respective leaves names are `CentralityPercentile`, `N_part`, `N_part_error`, `N_coll`, `N_coll_error`, `B` and `B_error`.

The tuple tool was written such that the user can set the desired number of centrality classes (N), with the restriction that $100/N$ is integer. This is to ensure that every class will span an integer number of percentiles. If $N = 10$ for example, then the mean values of the geometric quantities of the spanned percentiles get averaged, and this is the value returned for each class. The effect of doing this is negligible as can be seen in Tab. 4.15 for the three geometric quantities in the PbPb case. In the table the overlined quantities refer to the averaged values of the ten percentiles comprised in each centrality class.

The `error` leaves are computed taking into account the statistical uncertainties as well as the systematic uncertainties for the associated percentile. Concretely, in the tuple tool the systematic uncertainties are stored for twenty classes of five percentiles each, covering all the one hundred percentiles, that is, of the intervals $95-100$ %, $95-90$ %, and so on until $0-5$ %. If ten classes have been requested from the tuple tool and an event is found to belong to the percentile 74, it will be classified into the centrality class $70-80$ %. The uncertainty that will be reported for this event will be the quadrature addition of the statistical uncertainty from the $70-80$ % class and the systematic uncertainty of the $70-75$ % percentiles, for each of the geometric quantities. In the same way an event belonging to the percentile 77, will take into account the statistical uncertainty from the same $70-80$ % class and the systematic

TABLE 4.15: Comparison between the geometric quantities from the calculation performed specifically for ten classes and the values obtained by averaging ten percentiles (overlined values). The quantities used here are from PbPb.

Centrality %	N_{part}	$\overline{N_{\text{part}}}$	N_{coll}	$\overline{N_{\text{coll}}}$	b	$\overline{b}$
$100-90$	2.91	2.93	1.83	1.85	15.41	15.41
$90-80$	7.03	7.03	5.77	5.77	14.56	14.56
$80-70$	15.92	15.90	16.44	16.42	13.59	13.59
$70-60$	31.26	31.25	41.28	41.26	12.61	12.62
$60-50$	54.65	54.63	92.59	92.53	11.59	11.59
$50-40$	87.54	87.50	187.54	187.43	10.47	10.47
$40-30$	131.24	131.24	345.53	345.55	9.23	9.23
$30-20$	188.02	188.01	593.92	593.85	7.80	7.80
$20-10$	261.84	261.82	972.50	972.40	6.02	6.02
$10-0$	357.16	357.17	1570.26	1570.37	3.31	3.31

uncertainty of the $75-80$ % percentiles.

The returned value for the `CentralityPercentile` corresponds to the upper edge of the centrality interval, so for `NClasses` $= 20$, an event having a value of 85 for `CentralityPercentile`, would be an event corresponding to the centrality class $80-85$ %.

4.9.1 Special case: working with VELO clusters

As it has been argued, the VELO clusters cannot be used for centrality measurements either in PbPb, due to saturation, nor in PbNe, due to the extension in space of the region where the collisions take place. However, if one wants to perform an analysis not in centrality bins but in bins of another quantity, such as `nVeloClusters`, and one still wants to have the centrality information of the results, there is a way of proceeding. The tuple tool can be used on a MB sample to get the mean values of the centrality percentiles and geometric quantities for each of the `nVeloClusters` bins. This will be illustrated with an example.

As an example, PbPb data is to be analysed and the `nVeloClusters` intervals $[1000, 4000]$, $[4000, 6000]$ and $[6000, 10\,000]$ are used, which are denoted 1^{st}, 2^{nd} and 3^{rd} interval. The first step is to store in the data tuple the centrality information for single percentiles from a MB data sample, that is, the centrality information having selected `NClasses` $= 100$. Now when one event is analysed, it is classified according to the `nVeloClusters` bins, and for each of these bins the event centrality percentile, N_{part}, N_{coll} and b is saved. Finally, for each `nVeloClusters` interval there is a distribution for each of the relevant quantities. An example of this is shown in Fig. 4.29. Tab. 4.16 shows the centrality information for the 1^{st}, 2^{nd} and 3^{rd} intervals. For the centrality percentiles the statistical uncertainty is computed. For N_{part}, N_{coll} and b, the statistical uncertainty is computed, and added in quadrature with the root mean square of the uncertainties of all events that go into each category (the `error` leaves values), like this the systematic uncertainties are also taken into account.

This approach has already been used in different studies at LHCb [140, 141] where the data are binned in the `nVeloClusters` variable.

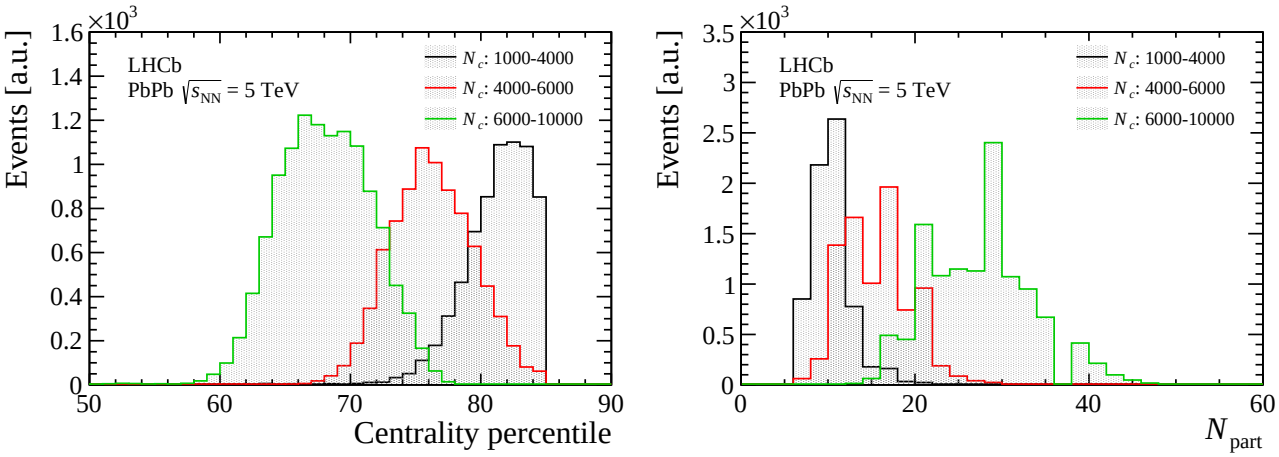

FIGURE 4.29: Centrality (*left*) and N_{part} (*right*) distributions corresponding to the VELO intervals from the text.

TABLE 4.16: Centrality classes and the corresponding geometrical quantities for the selected VELO intervals. The uncertainties for the centrality percentiles are purely statistical. The uncertainties for N_{part}, N_{coll} and b are the quadrature addition of the statistical uncertainty and the root mean square of the uncertainties (the values of the `error` leaves) of all events that went into each class.

nVeloClusters	Cent. percentile	N_{part}	N_{coll}	b
$1000-4000$	80.69 ± 0.03	10.59 ± 0.72	9.72 ± 0.64	14.10 ± 1.08
$4000-6000$	75.74 ± 0.04	15.70 ± 0.65	16.23 ± 0.70	13.62 ± 0.61
$6000-10000$	67.51 ± 0.03	27.80 ± 0.67	35.41 ± 0.91	12.81 ± 0.32

4.10 Conclusions

A procedure to determine centrality in PbPb collisions at $\sqrt{s_{NN}} = 5.02\,\text{TeV}$ and PbNe collisions at $\sqrt{s_{NN}} = 69\,\text{GeV}$ with the LHCb detector was implemented. It consisted in fitting a simulated energy distribution, based on a MC Glauber model, to the energy deposited in the ECAL. The ECAL is used since it does not saturate in the case of PbPb and it is not sensitive to the wide Z range in which the PbNe collisions take place as opposed to the VELO. After the fit is performed, the simulated distribution is divided in percentiles, which are delimited by sharp energy cuts obtained by integrating the distribution. These energy cuts allow to classify the data into the same percentiles and finally the geometric quantities from the MC Glauber model can be mapped to the real data, bearing in mind the limitation at 84% for PbPb and 89% for PbNe due to the UPC contamination. Sources of uncertainty for this procedure are identified and quantified. Finally, the energy values found to classify the data in percentiles, as well as the corresponding geometric quantities, are used to create a tuple tool named `TupleToolCentrality`, which has already been incorporated into the main LHCb analysis software. The results obtained in the PbPb case are in very good agreement with the results obtained by the ALICE [132, 133], ATLAS [134] and CMS [135] collaborations at the same centre-of-mass energy. The PbNe results correspond to the first centrality measurements in fixed-target collisions at the LHC.

Chapter 5

J/ψ and D^0 in PbNe collisions

In this chapter the study of J/ψ and D^0 production in PbNe collisions is studied in the decay channels $J/\psi \to \mu^+\mu^-$ and $D^0 \to K^-\pi^+$. The dataset is described, followed by the signal extraction, the efficiencies and the systematic uncertainties. Finally, the results are presented.

5.1 Data

The data used for this analysis are the same described in 3.1 and the runs used are the ones summarised in Tab. 3.6. They correspond to fixed-target collisions of Pb ions with Ne atoms acting as a fixed-target thanks to the SMOG system of LHCb. The nucleon-nucleon centre-of-mass energy is $\sqrt{s_{\text{NN}}} = 69\,\text{GeV}$.

5.2 Monte Carlo simulations

The different efficiencies, detailed in Sec. 5.5, are estimated with the aid of a MC simulation, which corresponds to a full-event simulation of the desired event using the LHCb simulation software, GAUSS, through a generation step based on several external tools, a subsequent step of the propagation of the particles through the detector based on the GEANT4 package [142, 143] and their reconstruction with the official LHCb software tools. The propagation step is the same as the one used for pp collisions, described in Ref. [144], but the generation step is specific to heavy-ion analyses.

The MB samples are generated with the EPOS event generator, using the LHC model [145]. This generator is interfaced with the GAUSS simulation software via the CRMC (Cosmic Ray Monte Carlo) interface library. All short lived particles are decayed with the EVTGEN decay package [146], similarly to what is done for pp simulation in LHCb. Radiative QED corrections to the decays containing charged particles in the final state are applied with the PHOTOS package [147], which is particularly important for $J/\psi \to \mu^+\mu^-$ decays. Since the instantaneous luminosity of the collisions recorded by the experiment in the different heavy-ion configurations is low, no pile-up is generated, and events contain only one interaction.

Signal samples are generated using an embedding technique: MB events are generated using the PYTHIA (version 8) generator [148], with colliding proton beams having momenta equal to the momenta per nucleon of the heavy-ion beams or targets. The sought signal particles (J/ψ or D^0 in this case) are then extracted from these MB events, discarding all other particles in the events. Their decays are forced to the signal decay modes using the EVTGEN package, and the resulting decay chain is added to a single MB EPOS event generated with beam parameters identical to those seen in data.

TABLE 5.1: MC data samples generated with PYTHIA plus EPOS.

EventType	Decay chain	Number of events	Details
22162000	$D^0 \to K^-\pi^+$	4 009 848	Official MC PbNe, within LHCb acceptance
22162000	$D^0 \to K^-\pi^+$	4 000 000	Private MC PbNe, in 4π
24142001	$J/\psi \to \mu^-\mu^+$	1 005 420	Official MC PbNe, within LHCb acceptance
24142001	$J/\psi \to \mu^-\mu^+$	4 000 000	Private MC PbNe, in 4π

For the analysis it is required that the PV be in the region $[-200, 200]$ mm, so to avoid any edge effects MC events are generated uniformly in a larger PVZ range of $[-250, 250]$ mm. The MC samples used are listed in Table 5.1.

5.3 Signal extraction

When the collisions take place in the detector, the events are either saved or discarded depending on whether they have triggered an L0 trigger line or not. Even then, the event rate would be too high and a system of HLT is used to further discriminate and keep events of interest. Subsequently, the data are stripped, which corresponds to a broad selection to organise the data into different categories depending on more complicated measures of the event topology. Finally, the data that have passed the previous criteria are carefully selected in order to isolate as much as possible the sought signal and perform the desired measurement.

The full selection chain corresponds thus to a combination of trigger lines (L0 and HLT), stripping lines, and a candidate selection. Additionally, to ensure that the data correspond to PbNe collisions and not to parasitic PbPb collisions, a series of requirements are needed, called the global event selection.[1] In the following the GEC will be described first, followed by trigger, stripping and candidate selection for J/ψ and D^0.

It is worth mentioning here that in the final candidate selection the required rapidity range is chosen to be $y \in [2, 4.29]$. The reason behind this is twofold. First, it will allow to compare the results with the parallel analysis of pNe collisions at the same centre-of-mass energy, and secondly, given that the collision system is Lorentz boosted towards the detector, this rapidity range translates to $y^* \in [-2.29, 0]$, encompassing the negative hemisphere of the centre-of-mass rapidity. Additionally, limiting the forward y acceptance has the advantage that it removes possible edge effects due to the beam-pipe, which could become even more important for events that take place at large Z close to the limit of the allowed region (200 mm).

5.3.1 Global event selection

As already established, the topology from SMOG events is completely different from regular beam-beam collisions. Hence, a series of requirements can be set in order to ensure the purity of the PbNe data sample.

The collisions are classified according to a beam crossing type (BCType), which depends on the status of the bunches crossing the IP, that is, if they are filled or empty. Nominally, to keep only PbNe collisions that are boosted towards the detector, a BCType = 1 is needed. This means that a filled bunch from beam 1 crosses an empty

[1] This selection will be referred to as GEC in the following, from its more colloquial name "global event cuts".

bunch from beam 2, and thus no PbPb collision is expected. This however, is not always the case since some Pb ions can migrate from a filled bunch into an empty bunch and cause parasitic PbPb collisions, commonly referred to as "ghost" collisions. In order to get rid of these the following requirements are implemented:

- nPVs > 0: at least one primary vertex is reconstructed.
- PVNTRACKS$[0] > 13$: the best quality vertex has been reconstructed with at least 13 tracks.
- $-200 < Z_{PV} < 200\,\text{mm}$: the primary vertex position is in the region $[-200, 200]$ mm with respect to the nominal IP.
- nPUHits ≤ 10: a veto on the backward activity. In the PbNe collisions due to the boost of the system very few hits in the PU stations are expected as compared to PbPb collisions.

An overall systematic uncertainty arising from these requirements is determined to be 0.23%. For more details on the definition of the GEC and the uncertainty determination see Ref. [149].

5.3.2 J/ψ selection

To extract the J/ψ signal among all the data the first step is to select the relevant L0 trigger line. Since the decay channel under study presents a dimuon final state, the trigger line considered is the L0MuonBE line, which requires that:

- $\mu_{p_T} > 10\,\text{MeV}/c$.
- nSPDHits < 3000.[2]

The definition of this line is constant throughout the data-taking period.

The following step is the HLT1 selection. In this case as in Sec. 3.2 the considered line is the `Hlt1SMOGDiMuonHighMass`, defined in Tab. 3.2, which imposes constrains in the p and p_T of the J/ψ, as well as on the track reconstruction quality. For the J/ψ candidates it is required that this line is triggered-on-signal (TOS) meaning that the track that fulfilled the HLT1 requirements is a track that makes up the J/ψ candidate. For the HLT2 selection, only the `Hlt2SMOGPassThrough` line is considered, which only requires to have passed an HLT1 line.

For the stripping stage, only J/ψ candidates fulfilling the requirements of the stripping line `HeavyIonDiMuonJpsi2MuMu` are considered. The detailed definition of this stripping line is given in Tab. 3.2.

Finally, the candidate selection is applied. It corresponds to a tighter selection to reduce the background level. The final candidate selection requirements are listed in Tab. 5.2. Here, ProbNNmu is a neural-network-based particle identification variable which outputs the probability of the particle being a muon, considering detector information and track quality. The vertex χ^2/ndf variable is a measure of the vertex reconstruction quality and the IP χ^2/ndf is related to how close the J/ψ back-propagated path is to a PV. This is important to leave out possible feed-down coming from b decays, which would be displaced from the PV.

[2] This is to avoid the reconstruction of spurious events that would require important computing time.

TABLE 5.2: J/ψ candidate selection requirements.

J/ψ selection requirements	
$\mu^{\pm}$	$2 < \eta < 4.6$, ProbNNmu > 0.5
J/ψ	vertex $\chi^2/\text{ndf} < 16$, IP $\chi^2/\text{ndf} < 6$, $y \in [2, 4.29]$

TABLE 5.3: D^0 candidate selection requirements.

D^0 selection requirements	
K^-	$2 < \eta < 4.6$, PIDK > 5, IP $\chi^2/\text{ndf} > 4$, $p_T > 650\,\text{MeV}/c$
π^+	$2 < \eta < 4.6$, PIDK < 0, IP $\chi^2/\text{ndf} > 4$, $p_T > 650\,\text{MeV}/c$
D^0	vertex $\chi^2/\text{ndf} < 16$, IP $\chi^2/\text{ndf} < 6$, $y \in [2, 4.29]$, $\tau > 0.4\,\text{ps}$, DIRA > 0.9998

5.3.3 D^0 selection

To extract the D^0 signal, the work-flow is identical to that of J/ψ. Since the decay channel under study presents a final state consisting in two hadrons, the trigger line considered is the L0HadronBE line, which requires that:

- transverse energy in HCAL $> 15\,\text{MeV}$.
- $4 <$ nSPDHits < 3000.

The definition of this line is constant throughout the data-taking period.

The following step is the HLT1 selection. In this case as in Sec. 3.2 the considered line is the `Hlt1BEHasTrack`, defined in Tab. 3.3, which imposes constrains in the p_T of the D^0, as well as on the track reconstruction quality. For the D^0 candidates it is required that this line is TOS. For the HLT2 selection, only the `Hlt2SMOGPassThrough` line is considered, which only requires to have passed an HLT1 line.

For the stripping stage, only D^0 candidates fulfilling the requirements of the stripping line `HeavyIonOpenCharmD02kpi` are considered. The detailed definition of this stripping line is given in Tab. 3.3.

Finally, the candidate selection is applied. The final candidate selection requirements are listed in Tab. 5.3, where τ is the proper lifetime of the D^0, the variable DIRA (which stands for direction angle) is the cosine of the angle between the direction of the D^0 reconstructed momentum and the direction defined by the PV position and the decay vertex of the D^0. For the D^0, the particle identification of its decay products is made with the variable PIDK, which only uses information from the RICH, muon stations and calorimeters to asses the ID of the particle. In general PIDx means $LL(P, x) - LL(P, \pi)$, with $LL(P, x)$ the log-likelihood of particle P being x. Like this, PIDK > 5 means that it is more likely for the particle in question to be a kaon over a pion.

5.3.4 Determination of signal yields

The $J/\psi \to \mu^+\mu^-$ candidates are evaluated with an unbinned maximum likelihood fit to the invariant mass using the ROOFIT software, available in ROOT. To model the signal a crystal-ball (CB) function is used. It is defined as:

$$F_{CB}(x; \mu, \sigma, \alpha, n) = C \begin{cases} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) & \text{if } \frac{x-\mu}{\sigma} > -|\alpha| \\ \frac{\left(\frac{n}{|\alpha|}\right)^n \exp\left(-\frac{1}{2}\alpha^2\right)}{\left(\frac{n}{|\alpha|} - |\alpha| - \frac{x-\mu}{\sigma}\right)^n} & \text{otherwise} \end{cases} \tag{5.1}$$

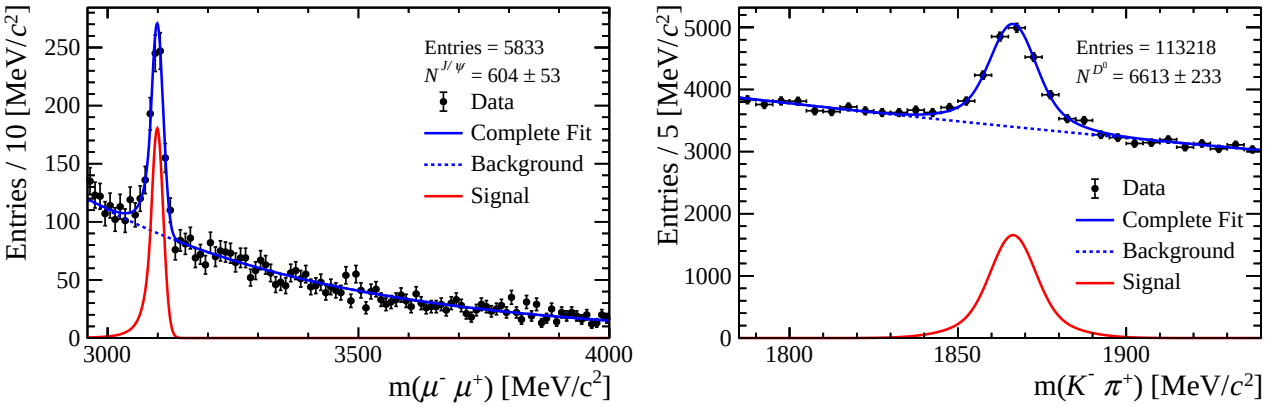

FIGURE 5.1: Invariant mass fit for $J/\psi \to \mu^+\mu^-$ (*left*) and for $D^0 \to K^-\pi^+$ (*right*).

This function, developed by the Crystal Ball Collaboration [150, 151], binds a Gaussian function with a power-law tail at low mass to account for the radiative losses and energy loss fluctuations in the detection process of the muons.

On the other hand, $D^0 \to K^-\pi^+$ candidates are evaluated with two Gaussian functions centred at the same mean value. For both J/ψ and D^0 the background is modelled with an exponential function.

The fit to the invariant mass distributions are shown in Fig. 5.1. In total there are 604 J/ψ candidates and 6613 D^0 candidates in the data sample.

Binned signal yields

Given the low statistics of J/ψ in the recorded PbNe collisions, it is not viable to perform this measurement in double differential bins. Hence, the signal will be measured in one single variable at a time. The binning scheme chosen is the following:

- nSPDHits: [0, 200, 300, 446, 715, 960, 1700].
- y: [2.0, 2.79, 3.29, 3.79, 4.29].
- p_{T}: [0, 1000, 2000, 3000, 8000] MeV/c.
- PVZ: [-200, -120, -40, 40, 120, 200] mm.

The nSPDHits binning scheme was initially chosen to be [0, 446, 715, 960, 1700] in order to have a similar number of J/ψ candidates in each bin. However, the first bin was further split into 3 bins to be sensitive to the low multiplicity behaviour, which allows a better comparison to the pNe system.

Sometimes it will be useful to present results in the centre-of-mass rapidity. Due to the Lorentz boost of the system, to go from the laboratory frame rapidity to the centre-of-mass it is necessary to consider the relation[3]

$$y = y^* + \ln\left(\frac{\sqrt{s_{\mathrm{NN}}}}{m_p}\right), \tag{5.2}$$

where $\sqrt{s_{\mathrm{NN}}}$ is the centre-of-mass energy and m_p is the proton mass. With this at $\sqrt{s_{\mathrm{NN}}} = 69\,\mathrm{GeV}$ the centre-of-mass rapidity is given by $y^* = y - 4.29$ and thus with

[3] Provided that $\sqrt{s_{\mathrm{NN}}} \gg m_p$.

the binning scheme used only the backward rapidity hemisphere is considered, that is, $y^* \in [-2.29, 0]$.

In Figs. 5.2 and 5.3 the fits to the invariant mass in the different nSPDHits bins are shown, for J/ψ and D^0 respectively. In the case of the J/ψ, the sum of the candidates from all fits results in $N^{J/\psi} = 589 \pm 35$ and in the case of the D^0 the sum results in $N^{D^0} = 6573 \pm 282$. In both cases the sum is compatible with the number of candidates of the integrated fit (from Fig. 5.1) and the slightly lower number can be explained by the lower quality of the fit when binning the data sample due to the low statistics. The same plots in bins of y can be seen in Fig. 5.4. For the corresponding plots for p_{T} and PVZ bins see the Appendix A.

5.4 Detector occupancy

Before the computation of efficiencies it is important to check the occupancy of the detector. This is done in order to check that in PbNe collisions the occupancy regime is below the occupancy of Pbp collisions at a centre-of-mass energy of $\sqrt{s_{\mathrm{NN}}} = 8\,\mathrm{TeV}$, restricting the Pb$p$ data sample to events that have less than 8000 nVeloClusters, which is the usual limit taken for Pbp analyses. If this is the case, then it is certain that the detector performance in the context of PbNe is understood and characterised and in particular, it indicates that certain results from Pbp, such as the weights for particle identification (PID) efficiency, can be applied to the PbNe case. This is important since the PID weights and some correction factors for the tracking efficiency, which come from standard LHCb tools for efficiency computation,[4] are not yet available for the SMOG data samples, and thus they are taken from an available dataset that is similar in occupancy which, in this case, is the Pbp data.

In order to evaluate the occupancy of the detector, the nSPDHits and the number of VELO tracks (nVeloTracks) are studied. This gives an idea of the number of particles traversing the detector in the VELO region and also farther downstream at the level of the SPD (also indicative of the calorimeter region). These are shown in Fig. 5.5, where in both cases the distributions have been normalised to unity. It thus confirms that the detector occupancy even in most central PbNe collisions is lower, albeit similar, to the occupation of high multiplicity Pbp events. In particular:

- $MAX(\mathrm{nSPDHits(PbNe)}) \approx 2150 < MAX(\mathrm{nSPDHits(Pb}p\mathrm{)}) \approx 2350$
- $MAX(\mathrm{nVeloTracks(PbNe)}) \approx 520 < MAX(\mathrm{nVeloTracks(Pb}p\mathrm{)}) \approx 800$

In the following section the efficiencies are computed, and thanks to this result when evaluating the PID and track reconstruction efficiency the tables obtained from Pbp data with TrackCalib and PIDCalib can be safely used in the PbNe context.

5.5 Efficiencies

In this section the efficiencies are defined and computed. For the calculation of the efficiencies the MC must be reweighted to reproduce the multiplicity distribution found in the data. The final efficiency will be used to correct the measured particle yields to obtain the real yields. The computation is separate for J/ψ and D^0.

[4] These are TrackCalib and PIDCalib. More details about them are given in Secs. 5.5.5 and 5.5.8.

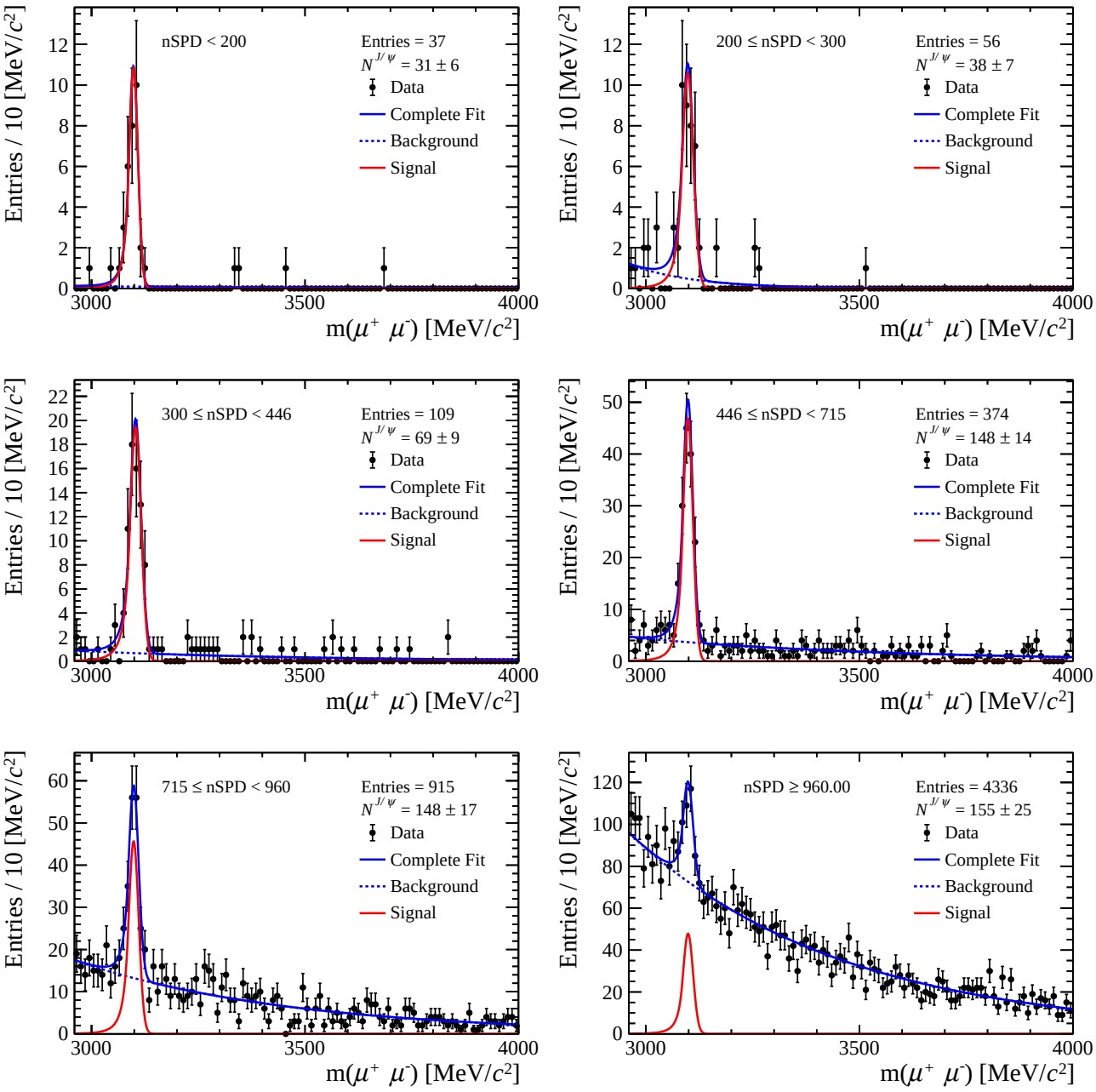

FIGURE 5.2: Invariant mass fit for $J/\psi \to \mu^+\mu^-$ in different nSPDHits bins.

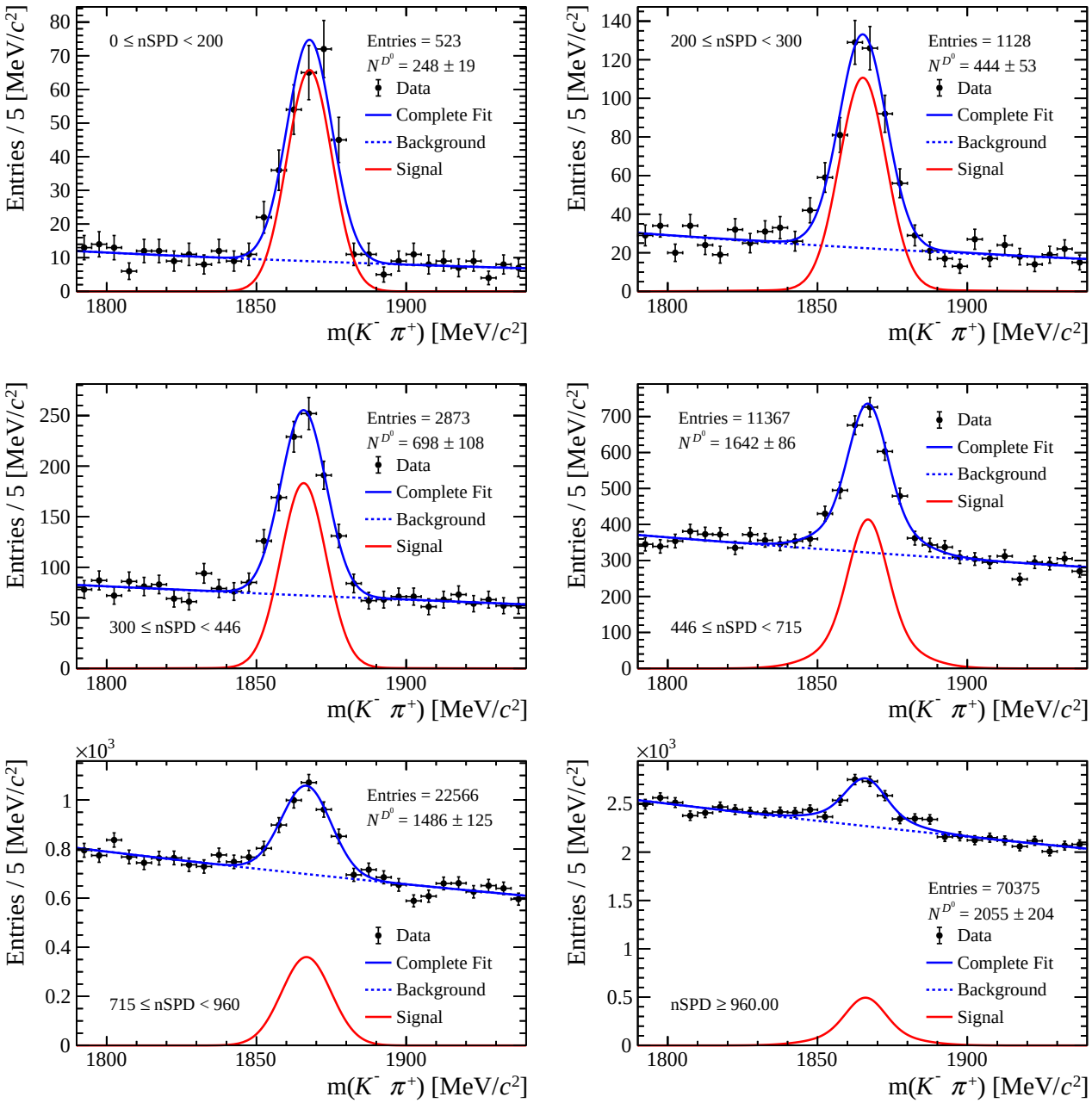

FIGURE 5.3: Invariant mass fit for $D^0 \to K^-\pi^+$ in different nSPDHits bins.

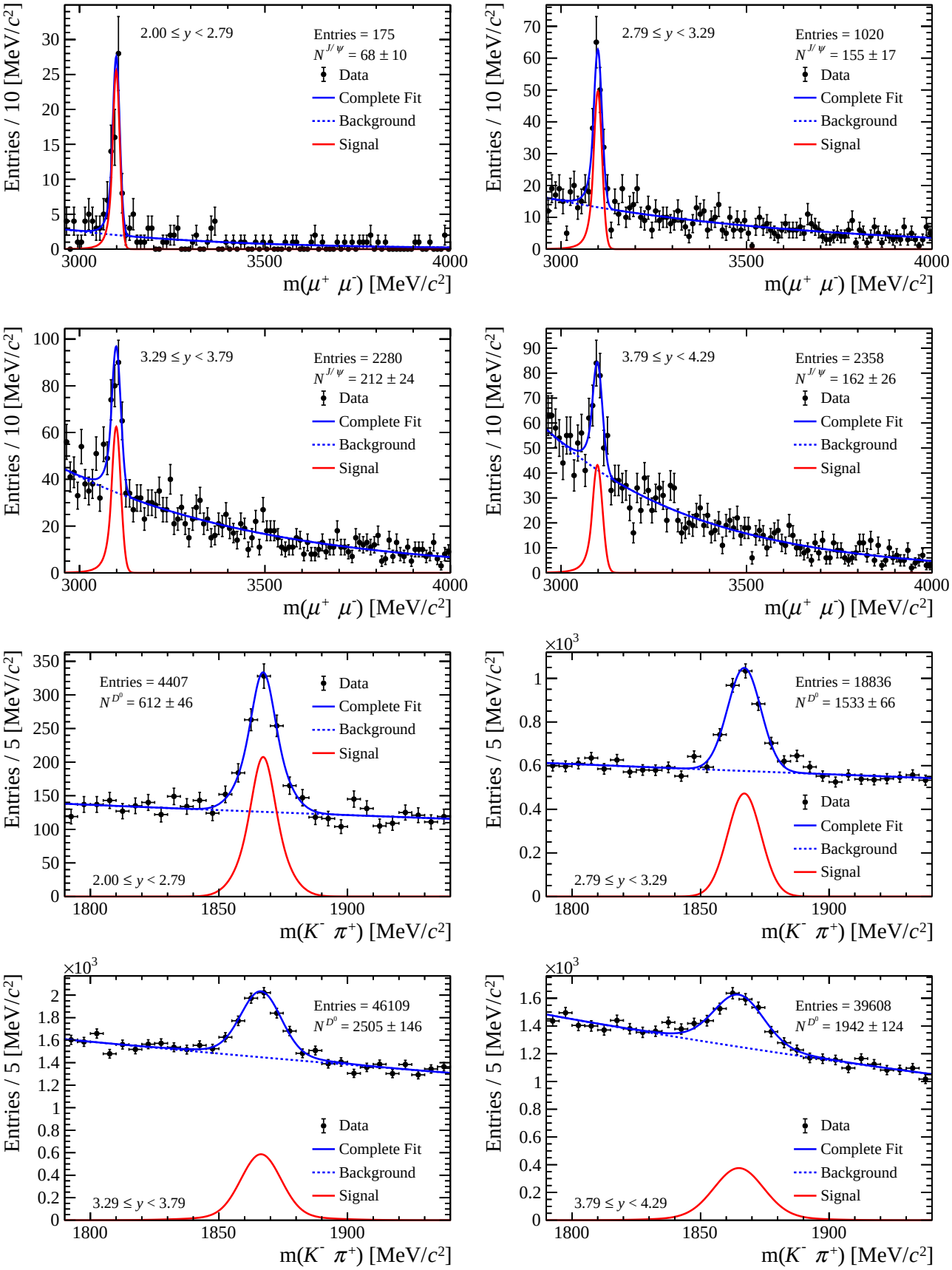

FIGURE 5.4: Invariant mass fit for $J/\psi \to \mu^+\mu^-$ (*top* 4) and $D^0 \to K^-\pi^+$ (*bottom* 4) in different rapidity bins.

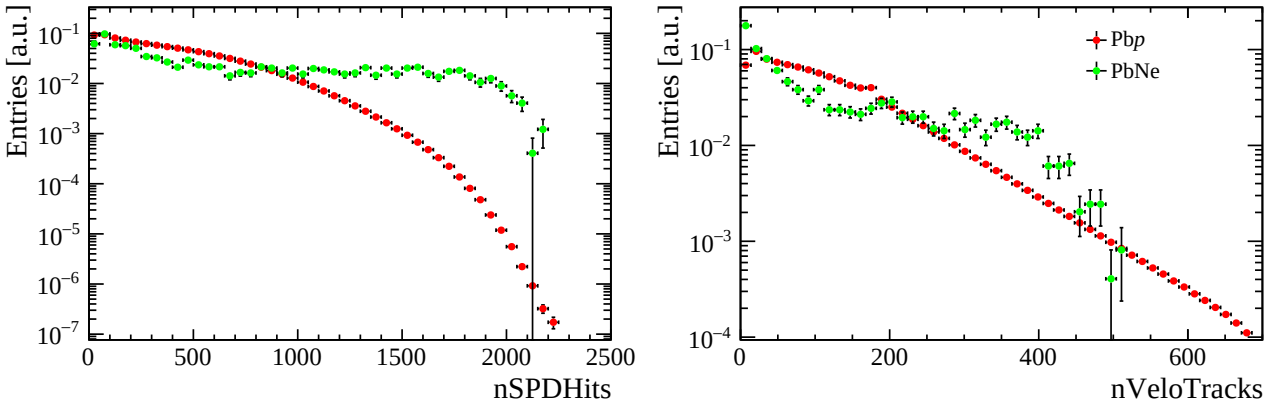

FIGURE 5.5: Detector occupancy comparison between Pbp (red) and PbNe (green) with nSPDHits (left) and nVeloTracks (right). All distributions have been normalised to unity.

5.5.1 Definitions

The correction to the measured yields is done as

$$Y^i = \frac{N^i}{\varepsilon^i_{\text{acc}}\varepsilon^i_{\text{tot}}\mathcal{B}^i}, \tag{5.3}$$

where Y^i is the real particle yield, N^i the measured yield, ε^i_{acc} is the geometrical acceptance, computed from a private MC simulation at generator level, ε^i_{tot} is the total detection efficiency and $\mathcal{B}^i$ the branching ratio of the specific decay channel under consideration. The index i stands for either J/ψ or D^0.

The total detection efficiency ε^i_{tot} can be decomposed as follows:[5]

$$\varepsilon_{\text{tot}} = \varepsilon_{\text{PV}} \times \varepsilon_{\text{Track}} \times \varepsilon_{\text{Stripping}} \times \varepsilon_{\text{GEC,Sel}} \times \varepsilon_{\text{PID}} \times \varepsilon_{\text{Trigger}}. \tag{5.4}$$

The different terms are defined as:

- ε_{PV}: The primary vertex efficiency, computed from the full LHCb simulation.
- $\varepsilon_{\text{Track}}$: Track reconstruction efficiency, computed from the full LHCb simulation.
- $\varepsilon_{\text{Stripping}}$: The stripping efficiency without the PID requirements, computed from the full LHCb simulation. The truth-matching (in)efficiency is considered in this step. The PID requirements are left out because they are considered later together with other PID requirements from the selection stage.
- $\varepsilon_{\text{GEC,Sel}}$: The GEC and selection efficiency, computed from the full LHCb simulation.
- ε_{PID}: The particle identification efficiency, computed from the full LHCb simulation and with the tables obtained from PIDCalib.
- $\varepsilon_{\text{Trigger}}$: The trigger efficiency, which is a "remaining" trigger efficiency after all the other selection criteria has been applied. It is computed from the full LHCb simulation.

[5] This decomposition is done in order to have a better understanding of how each selection step affects the signal, and to be able to cross-check step-by-step the results with other similar analyses.

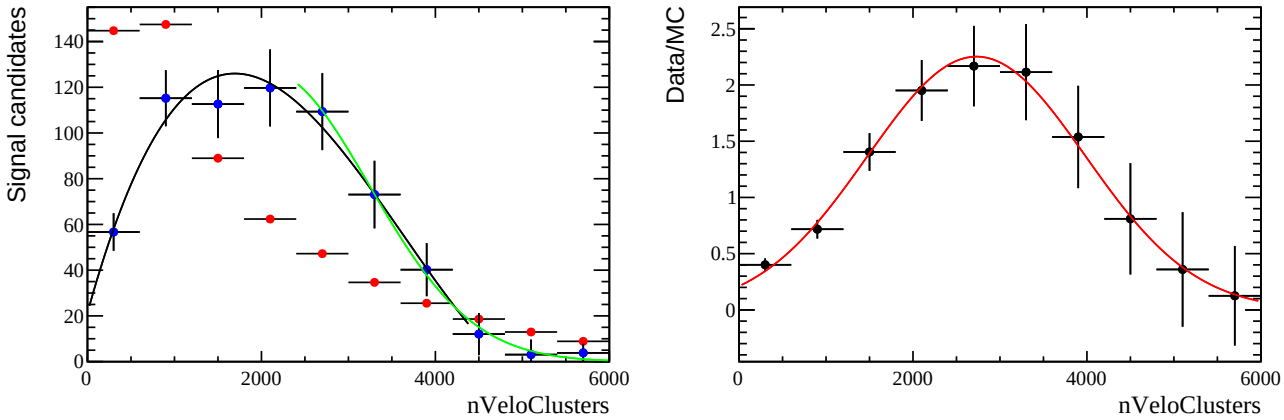

FIGURE 5.6: (*left*) nVeloClusters distributions in MC (red) and data (blue) for the J/ψ. The data have been fit empirically with a 3rd polynomial (black line) and a gaussian tail (green line). (*right*) The ratio between the fit to the data and the MC distribution. This ratio has been fitted with a gaussian function, and weights are sampled from this gaussian.

In order to compute the efficiencies, the first step is to reweight the MC, which will be described in the following section. This is due to the well known problem that the multiplicity of the MC events does not reproduce correctly the one found in the data. After the reweighting of the MC, the efficiencies are computed in bins of p_{T}, y, PVZ and nSPDHits independently, always in the PVZ region of $[-200, 200]$ mm.

5.5.2 MC reweighting

To reweight the MC, two approaches were considered. The first one consisted in comparing the distribution of the variable nVeloClusters between the data and the MC, since it is an important multiplicity variable. The data distribution was fitted empirically (see left panels of Figs. 5.6 and 5.7) and then the ratio fit(Data)/MC was

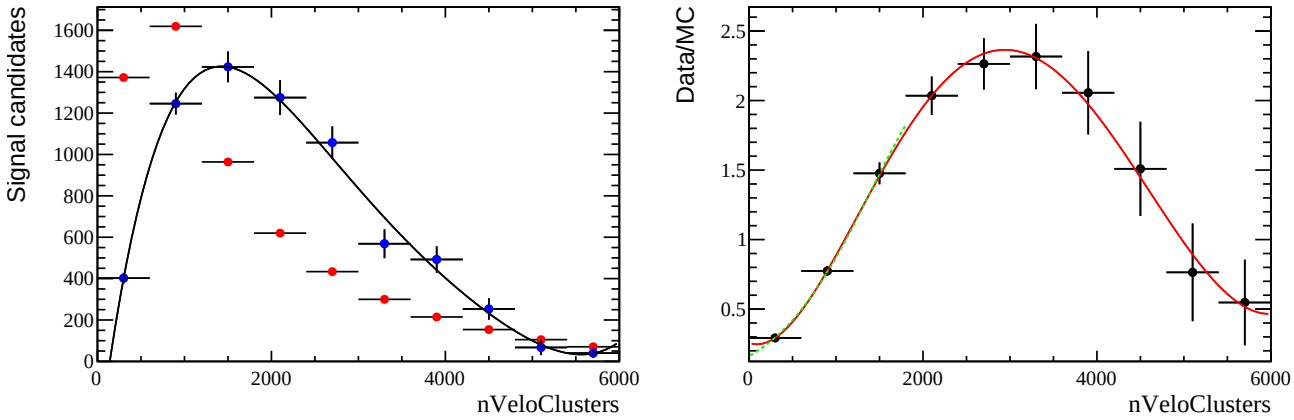

FIGURE 5.7: (*left*) nVeloClusters distributions in MC (red) and data (blue) for the D^0. The data have been fit empirically with a 5th degree polynomial. (*right*) The ratio between the fit to the data and the MC distribution. This ratio has been fitted with a 5th degree polynomial (red line) plus a gaussian tail (dashed green line) on the left side. Weights are sampled from these functions.

computed. Finally, as shown in the right panels of Figs. 5.6 and 5.7, the ratio was fitted again either by a Gaussian, in the case of J/ψ, or by a 5^{th} degree polynomial plus a Gaussian tail, in the case of D^0. This approach however, does not satisfactorily reproduce other multiplicity or kinematic variables, see Figs. 5.8 and 5.9.

The second reweighting approach makes use of a machine learning algorithm from the `hep_ml` python library.[6] Specifically the Gradient Boosted Reweighter (GBR) tool is used. It is an algorithm capable to work in high dimensions while keeping reweighting reliable and precise. To cross-check the two methods, a first reweighting with the GBR was performed using only nVeloClusters as a reweighting variable and all efficiencies were computed and compared to the ones obtained with the first approach. They were all found to be compatible within the percent level.

As it was stated before, the reweighting using nVeloClusters alone is not enough to close the gap between MC and data, and thus finally the variables chosen to reweight the MC were the nVeloClusters, the nSPDHits, and the p_{T}. The nSPDHits is important since it is a multiplicity indicator near the ECAL, which is the detector that was used to estimate the centrality of the collisions, and thus this variable can be used to classify events according to their centrality. The p_{T} is the only kinematic variable that does not match between MC and data, so it is used to reweight the MC.

The advantage of using the GBR is that it can simultaneously consider several variables, ensuring identical multidimensional distributions. In order to achieve this, the reconstructed MC and the data (with the full selection applied) are split in half, to form a training set and a test set. Usually one would choose 80% for training and 20% for testing, but since the statistics of the data are so low, it was preferred to halve the datasets. The strategy of this procedure is to:

1. Train the GBR with the MC^{train} and the $\text{data}^{\text{train}}$ sets.
2. Produce weights from the trained GBR for the MC^{test} set.
3. Compare the weighted MC^{test} set to the $\text{data}^{\text{test}}$.
4. If the procedure works with these 2 separate samples, the model is validated and the weights are predicted for the entire reconstructed MC (without selections) and for the entire truth MC, which are then used to compute the efficiencies.

The reweighting results can be seen in Figs. 5.10 and 5.11 for the variables that were used to train the GBR, and in Figs. 5.12 and 5.13 for other multiplicity and kinematic variables, which are in good agreement considering the uncertainties of the data.

[6] The documentation can be found here `https://arogozhnikov.github.io/hep_ml/reweight.html`. It is an open source library under the Apache 2.0 license.

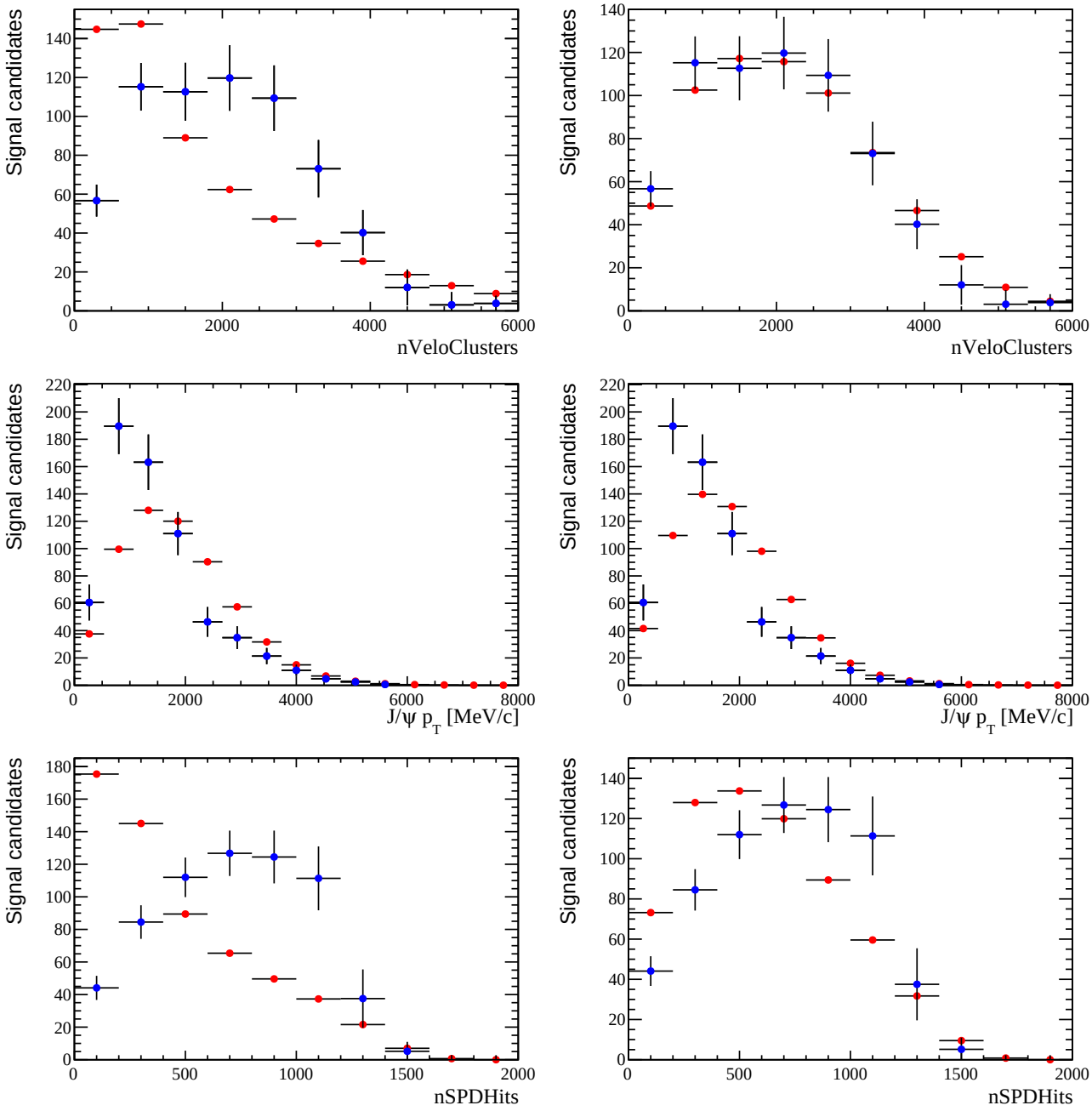

FIGURE 5.8: Comparison of the MC (red) and data (blue) distributions for the J/ψ. On the left side, the distributions before the reweighting are shown, while on the right side the distributions after the reweighting using approach 1 are shown. The top panels show the nVeloClusters distribution, which is the variable used for the reweighting, where a good agreement is found after reweighting. For the other variables, after reweighting there is not a good agreement between data and MC.

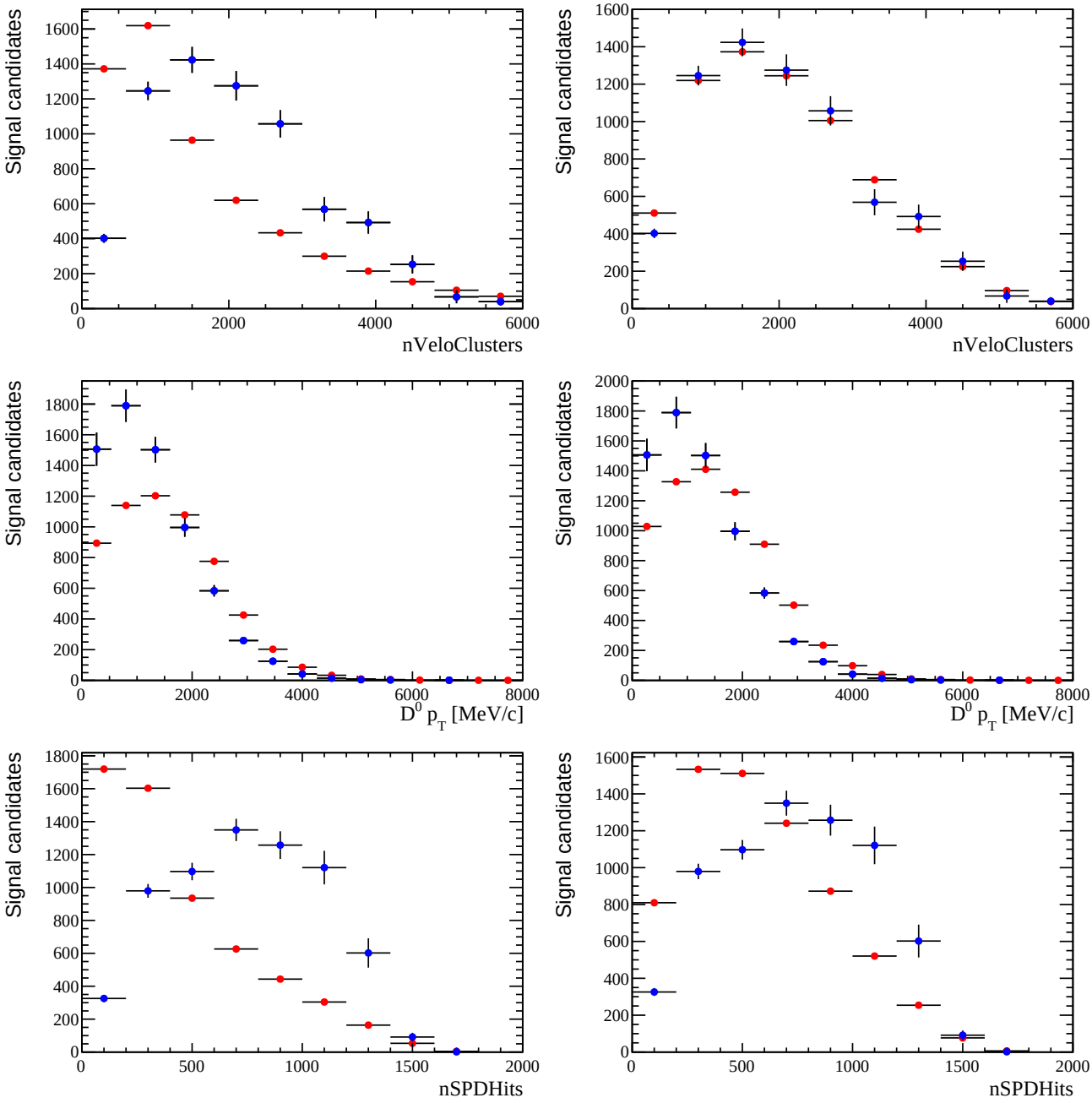

FIGURE 5.9: Comparison of the MC (red) and data (blue) distributions for the D^0. On the left side, the distributions before the reweighting are shown, while on the right side the distributions after the reweighting using approach 1 are shown. The top panels show the nVeloClusters distribution, which is the variable used for the reweighting, where a good agreement is found after reweighting. For the other variables, after reweighting there is not a good agreement between data and MC.

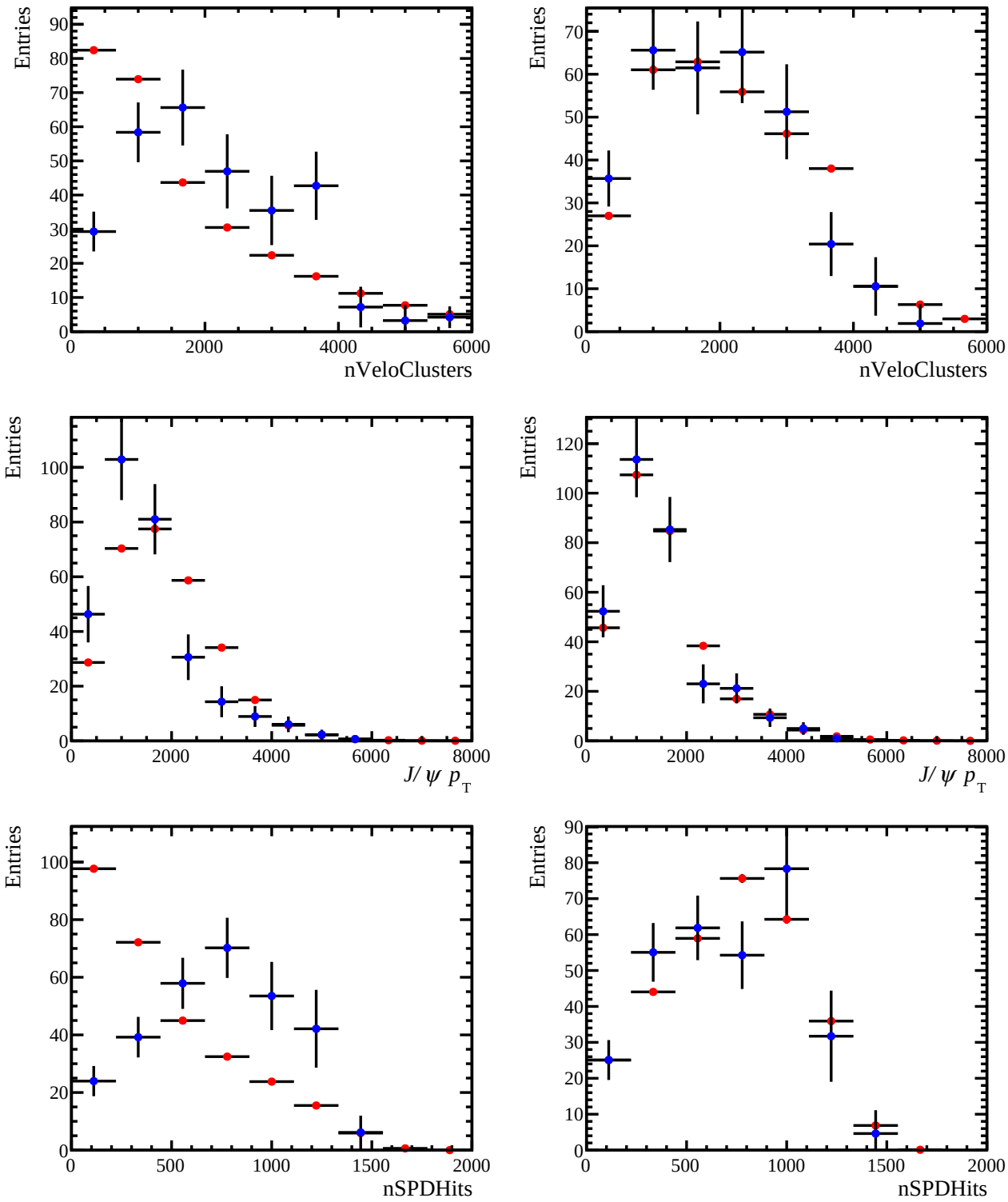

FIGURE 5.10: Comparison of the MC (red) and data (blue) distributions of the reweighting variables for the J/ψ with the GBR. On the left side, the distributions before the reweighting are shown, while on the right side the distributions after the reweighting are shown.

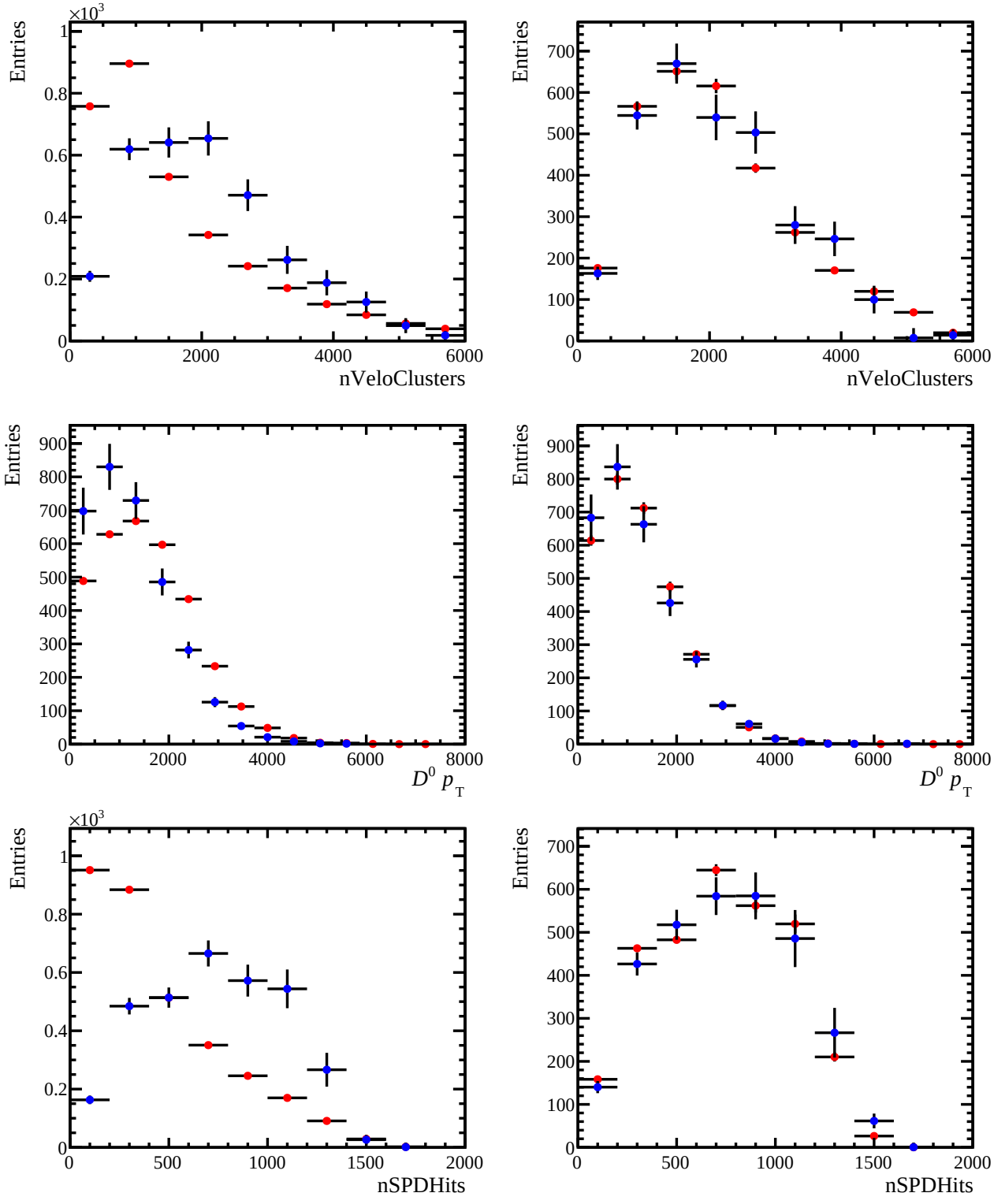

FIGURE 5.11: Comparison of the MC (red) and data (blue) distributions of the reweighting variables for the D^0 with the GBR. On the left side, the distributions before the reweighting are shown, while on the right side the distributions after the reweighting are shown.

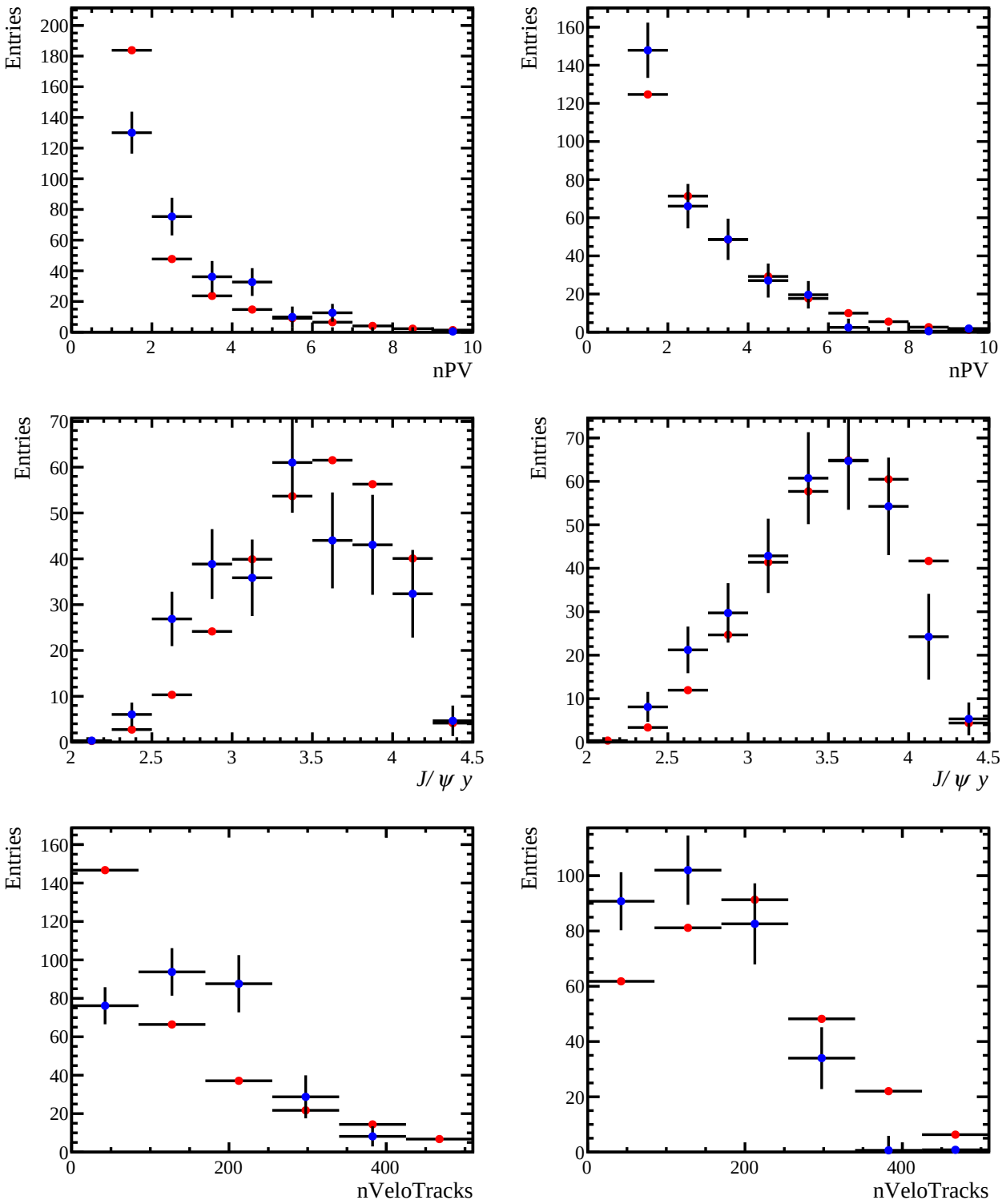

FIGURE 5.12: Comparison of the MC (red) and data (blue) distributions for the J/ψ with the GBR. On the left side, the distributions before the reweighting are shown, while on the right side the distributions after the reweighting are shown.

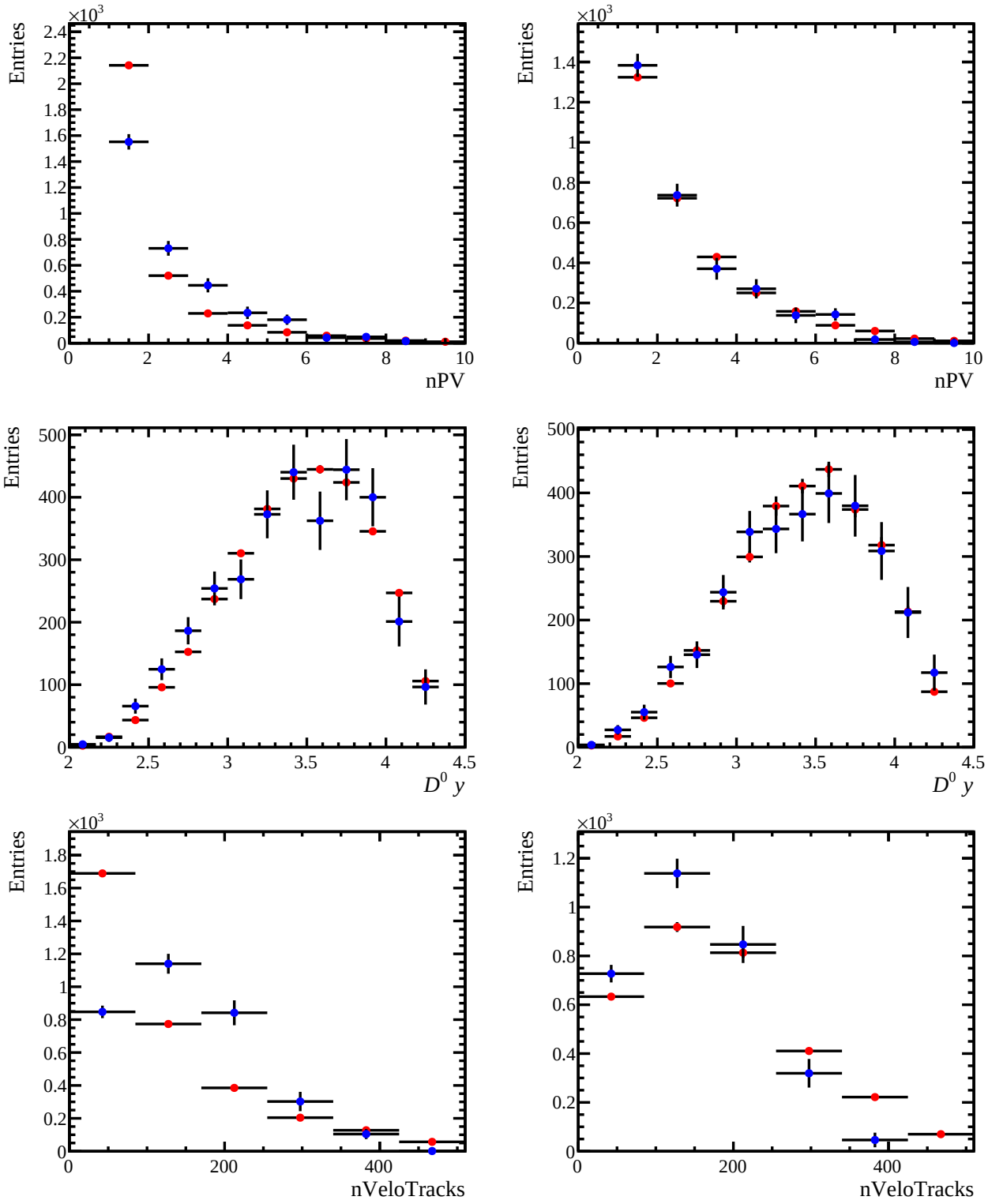

FIGURE 5.13: Comparison of the MC (red) and data (blue) distributions for the D^0 with the GBR. On the left side, the distributions before the reweighting are shown, while on the right side the distributions after the reweighting are shown.

5.5.3 Acceptance

The first quantity to be computed is the geometrical acceptance. With a vertex at PVZ $= 0$, an outgoing track is within the LHCb acceptance if its pseudorapidity is within $2 < \eta < 5$. It is common practice, to avoid possible border effects, to limit the tracks to $2 < \eta < 4.6$.

On the other hand, due to the fact that PbNe collisions are fixed-target collisions, the PVZ range is extended and there is a boost in the rapidity of 4.29 due to the 2.5 TeV Pb beam. Therefore, SMOG kinematics range mostly covers negative y^* hemisphere. Because of this, the study is limited to candidates with a rapidity in the range $y \in [2, 4.29]$, which in the centre-of-mass becomes $y^* \in [-2.29, 0]$, effectively considering only the backward hemisphere of the collision.

The acceptance is determined with the simulated data where J/ψ and D^0 signal candidates generated with PYTHIA are embedded in EPOS generated PbNe events. Since the acceptance depends only on the particles (pseudo)rapidities, it is calculated using generator level samples only, so no reweighting is needed. The 4π acceptance is calculated as follows:

$$\varepsilon_{4\pi\,\mathrm{acc}} = \frac{\text{Events with both daughters within } 2 < \eta < 4.6 \text{ and signal } y^* \in [-2.29, 0]}{\text{Events with both daughters within } 4\pi}. \tag{5.5}$$

In addition, both in the numerator and the denominator there are the additional requirements that the p_T be lower than 8 GeV, to limit it to the region of interest for the analysis (only 0.001% of events is left out) and that nVeloClusters be less than 6000, since in the data events with more than 6000 were not reconstructed.

J/ψ acceptance

The J/ψ acceptance is:

- $\varepsilon^{J/\psi}_{4\pi\,\mathrm{acc}} = (33.23 \pm 0.03)\%$.

As a cross-check, the acceptance within the rapidity range (y^* requirement both in numerator and denominator) is computed and found to be:

- $\varepsilon^{J/\psi}_{\mathrm{acc}} = (66.71 \pm 0.08)\%$.

As expected, there is a factor 0.498 between the two, which is explained by the fact that in the centre-of-mass only the negative rapidity hemisphere is considered.

The resulting distributions of the $\varepsilon^{J/\psi}_{4\pi\,\mathrm{acc}}$ as a function of y, p_T, PVZ and (p_T,y) can be seen in Fig. 5.14. It is important to mention that in Fig. 5.14 when visualising the efficiency as a function of y, there is no change if one considers the 4π acceptance or the LHCb acceptance. This is because the resulting efficiency is already limited to the considered rapidity bins set on the rapidity axis.

D^0 acceptance

The D^0 acceptance is:

- $\varepsilon^{D^0}_{4\pi\,\mathrm{acc}} = (35.08 \pm 0.03)\%$.

As a cross-check, the acceptance within the rapidity range (y^* requirement both in numerator and denominator) is computed and found to be:

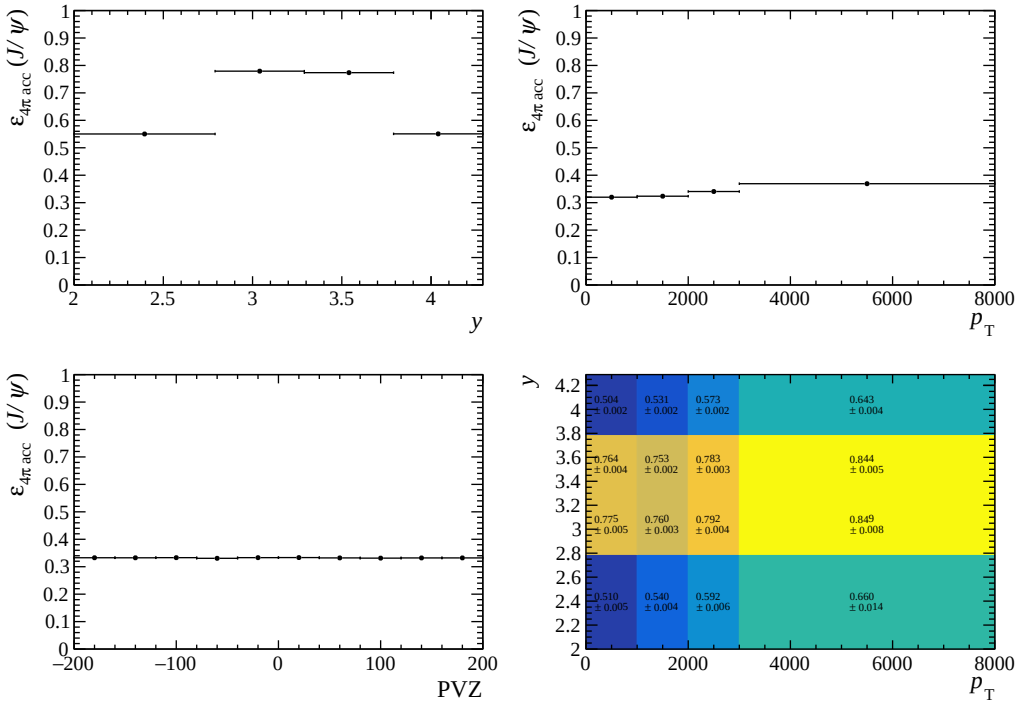

FIGURE 5.14: J/ψ 4π acceptance efficiency as a function of different variables. The top-left plot shows a greater efficiency since it is not affected by the change from LHCb acceptance to 4π (see text).

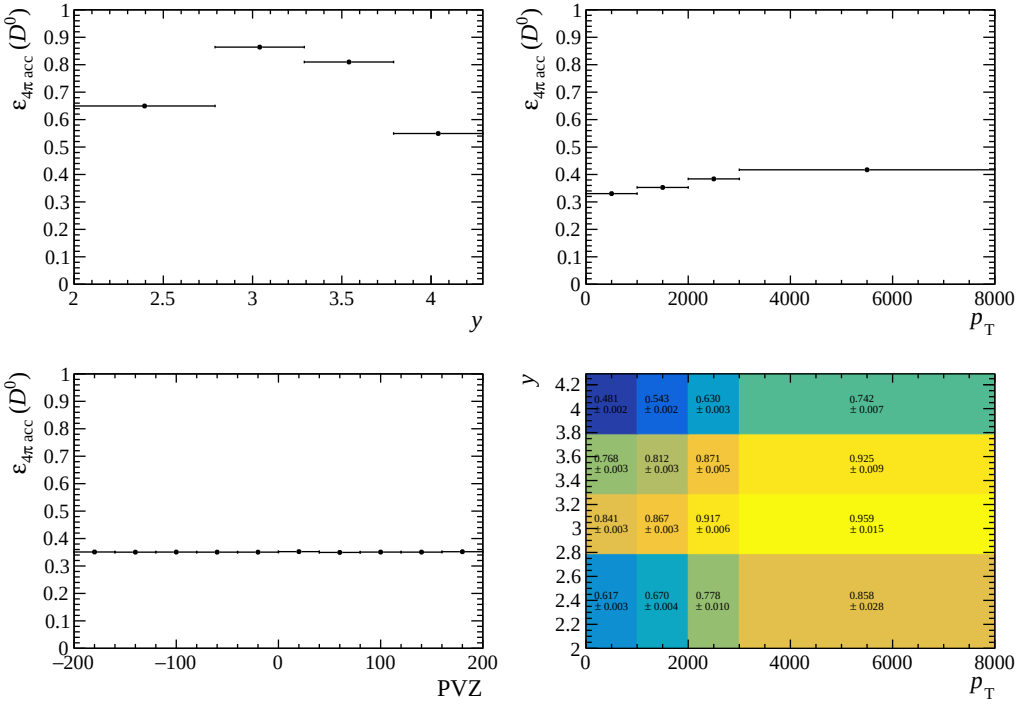

FIGURE 5.15: D^0 4π acceptance efficiency as a function of different variables. The top-left plot shows a greater efficiency since it is not affected by the change from LHCb acceptance to 4π (see text).

- $\varepsilon^{D^0}_{\text{acc}} = (71.16 \pm 0.09)\%$.

As expected, there is a factor 0.493 between the two, which is again due to the fact that only the negative rapidity hemisphere is considered.

The resulting distributions of the $\varepsilon^{D^0}_{4\pi\,\text{acc}}$ as a function of y, p_{T}, PVZ and (p_{T},y) can be seen in Fig. 5.15. As in the J/ψ case, in Fig. 5.15 when visualising the efficiency as a function of y, there is no change if one considers the 4π acceptance or the LHCb acceptance.

5.5.4 PV efficiency

The weighted MC samples are used to evaluate PV efficiency, which is computed as:

$$\varepsilon_{\text{PV}} = \frac{\text{Events with nPV} > 0}{\text{Events with both daughters within } 2 < \eta < 4.6 \text{ and signal } y^* \in [-2.29, 0]}. \tag{5.6}$$

The resulting PV efficiency is then:

- $\varepsilon^{J/\psi}_{\text{PV}} = (99.98 \pm 0.00)\%$.
- $\varepsilon^{D^0}_{\text{PV}} = (99.99 \pm 0.00)\%$.

Since the PV efficiency should not depend on the candidate MC, the average PV efficiency is computed and the difference between J/ψ and D^0 is used to compute a systematic uncertainty. The resulting efficiency as a function of several variables can be seen in Fig. 5.16 for the J/ψ and in Fig. 5.17 for the D^0. The final overall value for the PV efficiency, considering both results and their difference is:

- $\varepsilon_{\text{PV}} = (99.99 \pm 0.00\ (\text{stat}) \pm 0.01\ (\text{syst}))\%$.

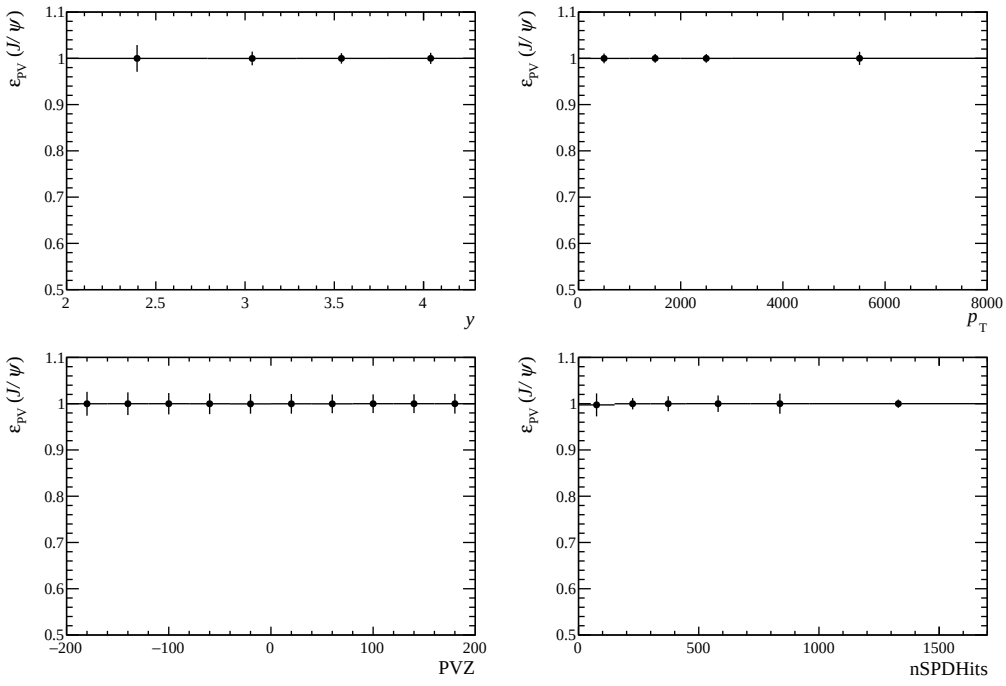

FIGURE 5.16: J/ψ PV efficiency as a function of y, p_{T}, PVZ and nSPDHits.

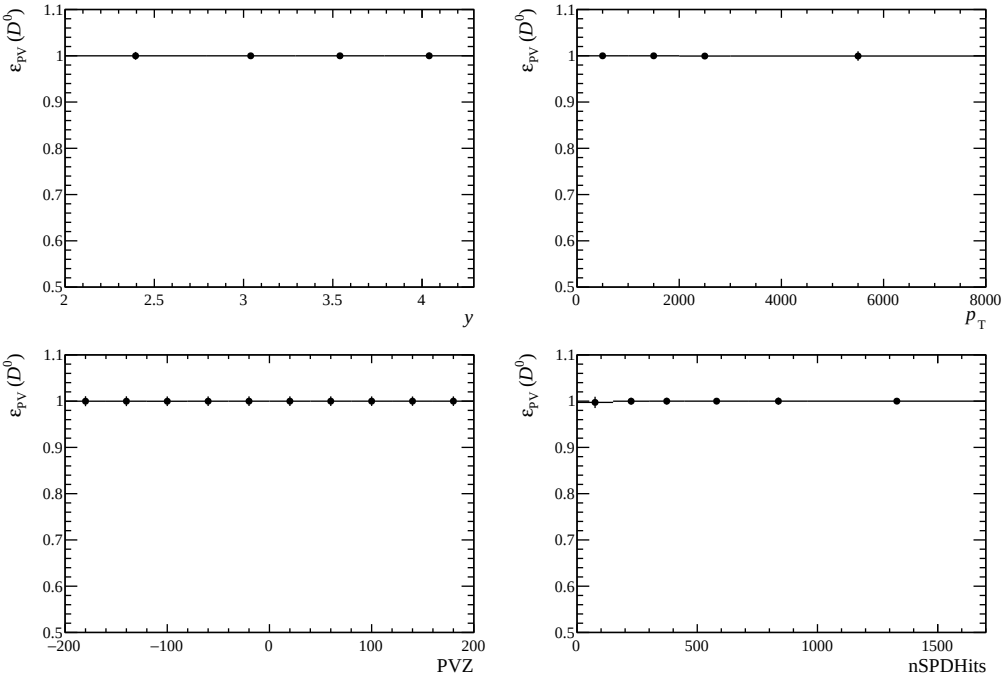

FIGURE 5.17: D^0 PV efficiency as a function of y, p_T, PVZ and nSPDHits.

5.5.5 Tracking efficiency

Tracking efficiencies encompass the reconstruction efficiency and the requirement that a track be reconstructed as a long-track. They are computed using the MC samples, including a track-by-track MC/data correction. More precisely, correction factors must be applied to simulation-based tracking efficiencies due to well-known discrepancies between data and simulation. These factors, $f^{\text{Track}} = \varepsilon^{\text{data}}_{\text{Track}}/\varepsilon^{\text{MC}}_{\text{Track}}$, have been computed using a tag-and-probe method, specifically the *Long* method,[7] with data from Pbp collisions at $\sqrt{s_{\text{NN}}} = 8\,\text{TeV}$. These corrections are computed

[7] For details of this method see Ref. [152].

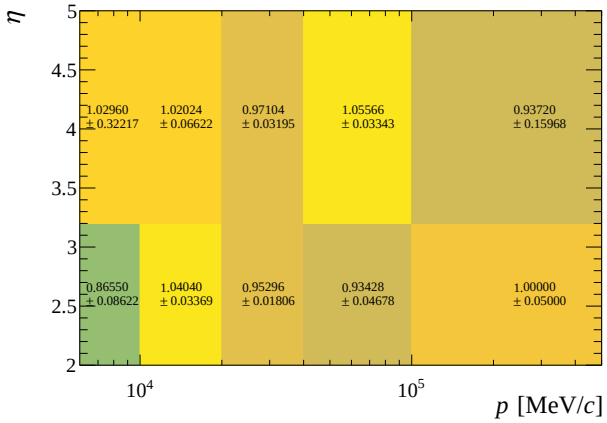

FIGURE 5.18: Correction factors to account for the discrepancy between track reconstruction in MC and data. These have been obtained using a tag-and-probe method on data from Pbp collisions at $\sqrt{s_{\text{NN}}} = 8\,\text{TeV}$.

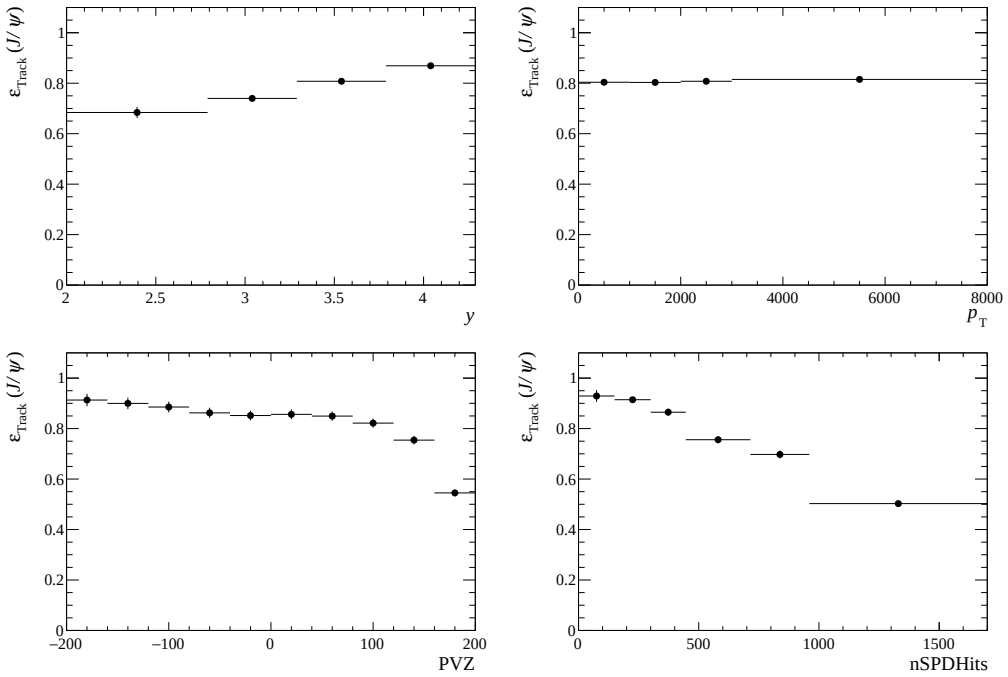

FIGURE 5.19: J/ψ Track efficiency as a function of y, p_{T}, PVZ and nSPDHits.

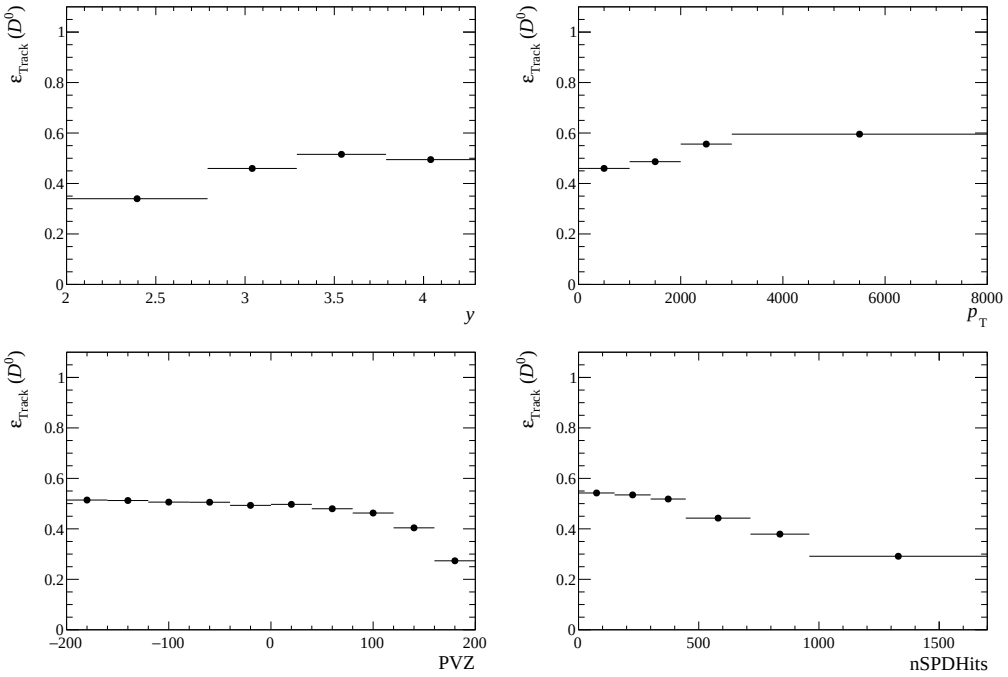

FIGURE 5.20: D^0 Track efficiency as a function of y, p_{T}, PVZ and nSPDHits.

as a function of the track momentum (p) and pseudorapidity (η), see Fig. 5.18. An additional systematic uncertainty of 0.5% has to be assigned on top of the uncertainties for the μ tracks, and of 2% for the π or K tracks.

The truth MC samples are used to compute the tracking efficiency as:

$$\varepsilon_{\text{Track}} = \frac{(\text{Daughter 1 reconstructed \& Daughter 2 reconstructed}) \times (f_1 \times f_2)}{\text{Both daughters with } 2 < \eta < 4.6 \text{ \& mother } y^* \in [-2.29, 0] \text{ \& nPV} > 0}, \tag{5.7}$$

with f_1, f_2 the correction factors previously discussed for daughter 1 and daughter 2 respectively. Both tracks have to be reconstructed as long-tracks.

The resulting integrated efficiencies are:

- $\varepsilon^{J/\psi}_{\text{Track}} = (80.42 \pm 0.05)\%$.
- $\varepsilon^{D^0}_{\text{Track}} = (46.63 \pm 0.01)\%$.

The efficiencies as function of y, p_{T}, PVZ and nSPDHits are shown in Figs. 5.19 and 5.20 for the J/ψ and the D^0 respectively.

5.5.6 Stripping efficiency

Usually the stripping efficiency would be computed in a straightforward manner using the reconstructed truth-matched MC. However, in this case since the PbNe events are of high multiplicity, the reconstructed MC tuple can only be filled with events that pass the stripping selection, in order to reduce the amount of candidates arising from combinatorial background. This means that the reconstructed tuple is already stripped and it cannot be directly disentangled from the truth-matching. Unfortunately, on the data only the stripping efficiency should be applied in the correction and so it needs to be isolated. Ideally, the stripping efficiency would be computed as:

$$\varepsilon_{\text{Strip}} = \frac{\text{Reconstructed events passing the stripping requirements, with a PV and both tracks reconstructed}}{\text{Reconstructed events with a PV and both tracks reconstructed}}, \tag{5.8}$$

with truth-matched events both in the numerator and denominator. However, since the reconstructed tuple is stripped from the very beginning the denominator is not accessible and instead a combined efficiency $\varepsilon_{\text{Stripping,truth}-\text{match}}$ is obtained. To isolate the stripping efficiency the following is computed:

$$\varepsilon_{\text{Strip}} = \frac{\varepsilon_{\text{Stripping,truth}-\text{match}}}{\varepsilon_{\text{truth}-\text{match}}}, \tag{5.9}$$

where $\varepsilon_{\text{truth}-\text{match}}$ is computed from the true reconstructed events that failed to be matched to the generated events (unmatched events). The generated signal events are fitted and the signal shape is used as a template to fit the signal in the unmatched events. The percentage of the candidates found in the unmatched events with respect to the total matched and unmatched candidates corresponds to the truth-matching inefficiency, $\bar{\varepsilon}_{\text{truth}-\text{match}}$, which is then used to obtain the truth-matching efficiency as $\varepsilon_{\text{truth}-\text{match}} = 1 - \bar{\varepsilon}_{\text{truth}-\text{match}}$.

For the computation of $\varepsilon_{\text{Stripping,truth}-\text{match}}$ both the truth MC and the reconstructed MC are used, and it is thus useful to define two requirements:

- For the reconstructed MC:
  - $MC^{Reco}_{Stripping,truth-match} = Strip + truth\text{-}match + PV + acceptance + track$,

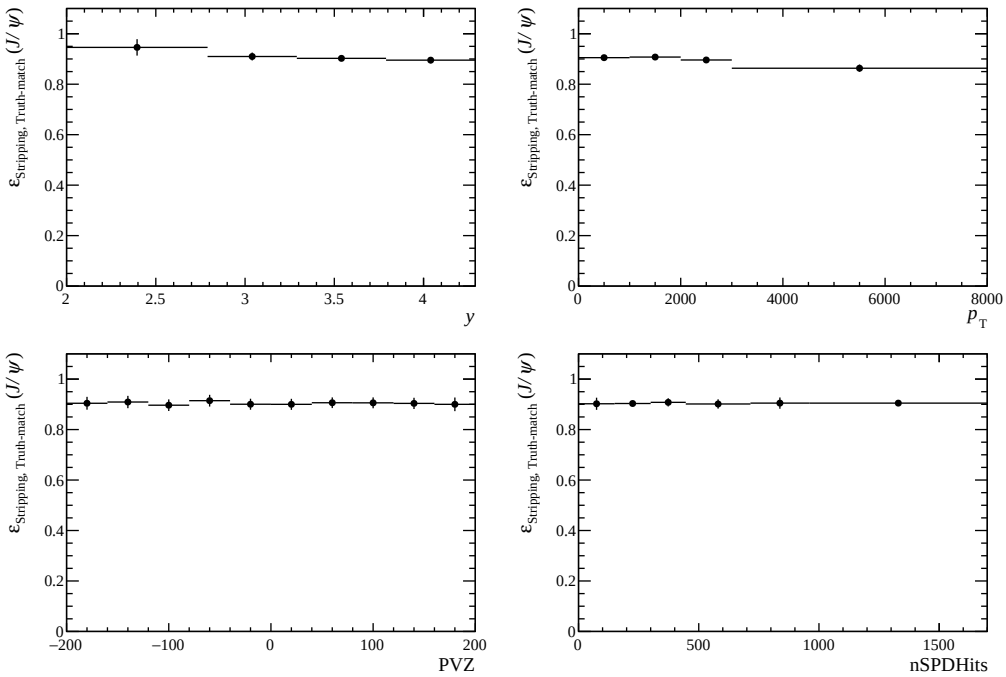

FIGURE 5.21: J/ψ stripping and truth-matching efficiency as a function of y, p_{T}, PVZ and nSPDHits.

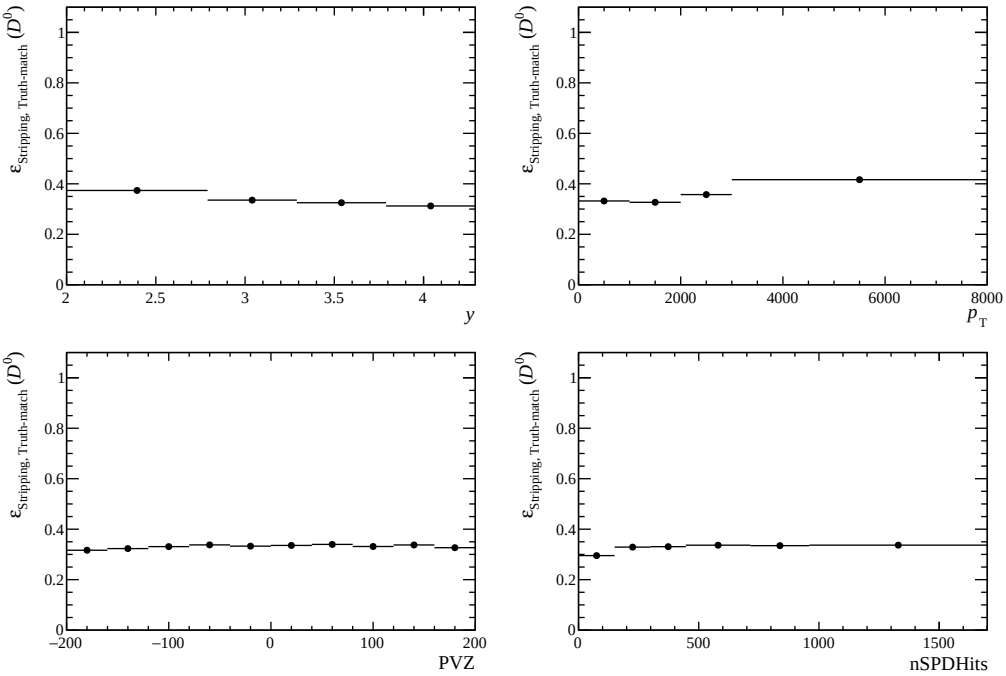

FIGURE 5.22: D^0 stripping and truth-matching efficiency as a function of y, p_{T}, PVZ and nSPDHits.

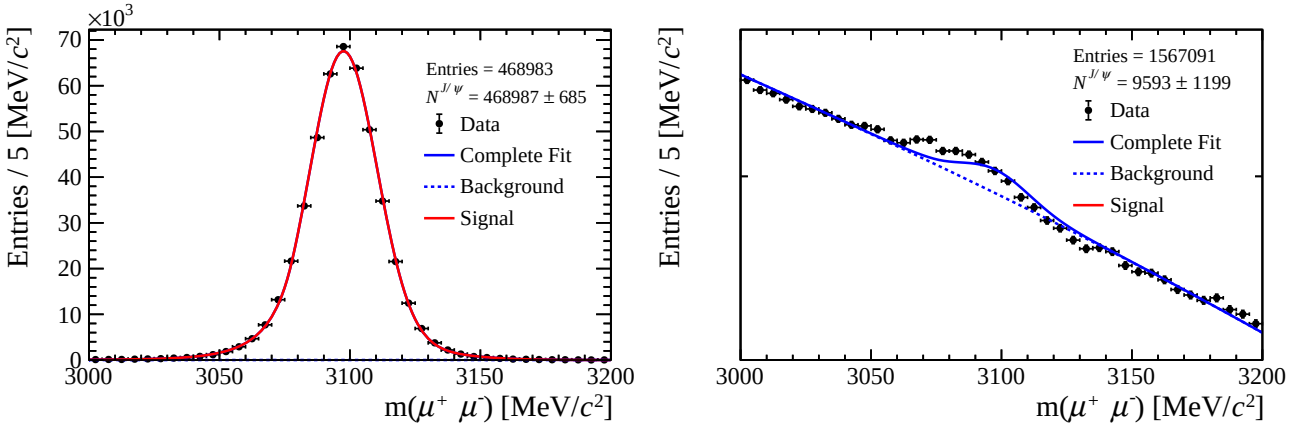

FIGURE 5.23: Fit to the J/ψ truth-matched candidates (*left*) and unmatched candidates (*right*). The function that fits the truth-matched candidates has been used as a template on the right with an exponential background.

- For the truth MC:
  - $MC^{truth}_{Track} = PV + acceptance + track$,

where *PV*, *acceptance* and *track* stand for the requirements previously defined for each respective efficiency. With these requirements, the $\varepsilon_{\text{Stripping,truth}-\text{match}}$ is obtained as:

$$\varepsilon_{\text{Stripping,truth}-\text{match}} = \frac{MC^{Reco}_{Stripping,truth-match}}{MC^{truth}_{Track}}. \tag{5.10}$$

The integrated $\varepsilon_{\text{Stripping,truth}-\text{match}}$ is found to be:

- $\varepsilon^{J/\psi}_{\text{Stripping,truth}-\text{match}} = (90.40 \pm 0.04)\%$.
- $\varepsilon^{D^0}_{\text{Stripping,truth}-\text{match}} = (33.13 \pm 0.01)\%$.

The resulting efficiencies are shown in Figs. 5.21 and 5.22 for J/ψ and D^0 respectively.

To obtain the correction factor due to the truth-matching inefficiency the first step is to fit the reconstructed MC signal using only truth-matched candidates. To do this the signal is modelled by a CB shape plus a gaussian function in the case of the J/ψ, and by two gaussian functions in the case of the D^0. Then, with the template from the fit to the truth-matched candidates, another fit is performed on the unmatched candidates with the addition of an exponential background. The results of this procedure are shown in Fig. 5.23 for the J/ψ and in Fig. 5.24 for the D^0.

For the D^0, 17805 ± 1644 candidates are found unmatched. This represents 4.69% of the total, resulting an efficiency of $\varepsilon^{D^0}_{\text{truth}-\text{match}} = (95.31 \pm 0.43)\%$. For the J/ψ on the other hand, the fit shown in Fig. 5.23 results in an efficiency of $\varepsilon^{J/\psi}_{\text{truth}-\text{match}} = (97.99 \pm 0.25)\%$. However, there is an apparent tail on the left side of the mass peak which is not covered by the fit function. To account for this and to include it as a systematic uncertainty, another fit was performed leaving free the means of the CB and of the gaussian, resulting in an efficiency of $\varepsilon^{J/\psi}_{\text{truth}-\text{match}} = (97.10 \pm 0.16)\%$. Combining these two results, the final value for the J/ψ truth-matching efficiency

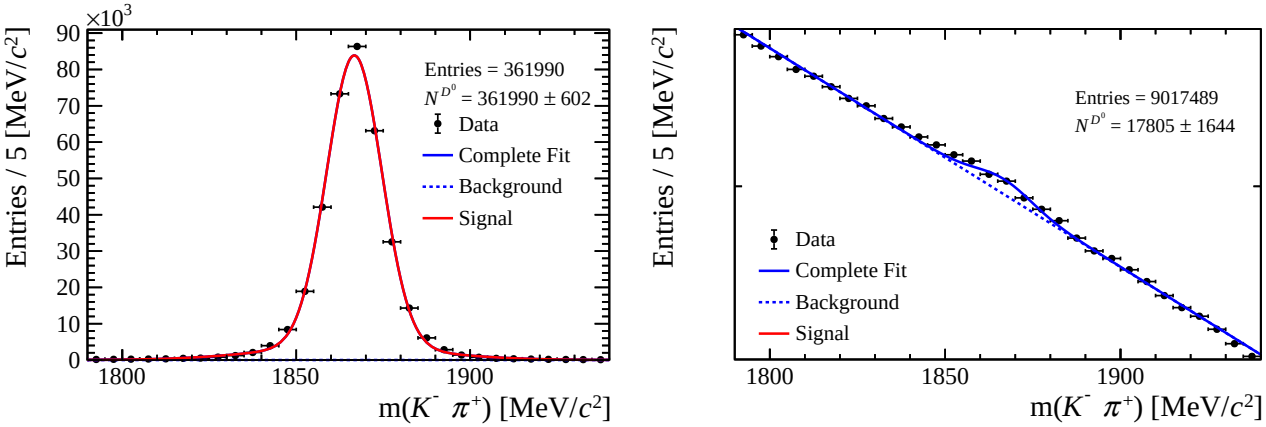

FIGURE 5.24: Fit to the D^0 truth-matched candidates (*left*) and unmatched candidates (*right*). The function that fits the truth-matched candidates has been used as a template on the right with an exponential background.

is $\varepsilon^{J/\psi}_{\text{truth}-\text{match}} = (97.55 \pm 0.29(\text{stat}) \pm 0.45(\text{syst}))\%$. These values are finally used to correct the $\varepsilon_{\text{Stripping,truth}-\text{match}}$ efficiencies, resulting in an integrated value of:

- $\varepsilon^{J/\psi}_{\text{Stripping}} = (92.67 \pm 0.28(\text{stat}) \pm 0.43(\text{syst}))\%$.
- $\varepsilon^{D^0}_{\text{Stripping}} = (34.76 \pm 0.16)\%$.

The same factors found are used to correct the efficiencies shown in Figs. 5.21 and 5.22.

5.5.7 GEC and selection efficiency

The GEC and selection efficiency is computed entirely with the reconstructed MC. The PID requirements from the stripping and the selection are not applied here. It is convenient to define:

$$MC^{Reco}_{GEC,Sel} = MC^{Reco}_{Stripping,truth-match} + GEC + Selection.$$

Then the GEC and selection efficiency is computed as:

$$\varepsilon_{GEC,Sel} = \frac{MC^{Reco}_{GEC,Sel}}{MC^{Reco}_{Stripping,truth-match}}. \tag{5.11}$$

The overall efficiencies found are:

- $\varepsilon^{J/\psi}_{\text{GEC,Sel}} = (91.45 \pm 0.04)\%$.
- $\varepsilon^{D^0}_{\text{GEC,Sel}} = (38.33 \pm 0.02)\%$.

The GEC and selection efficiency as a function of y, p_{T}, PVZ and nSPDHits is shown in Figs. 5.25 and 5.26 for J/ψ and D^0 respectively.

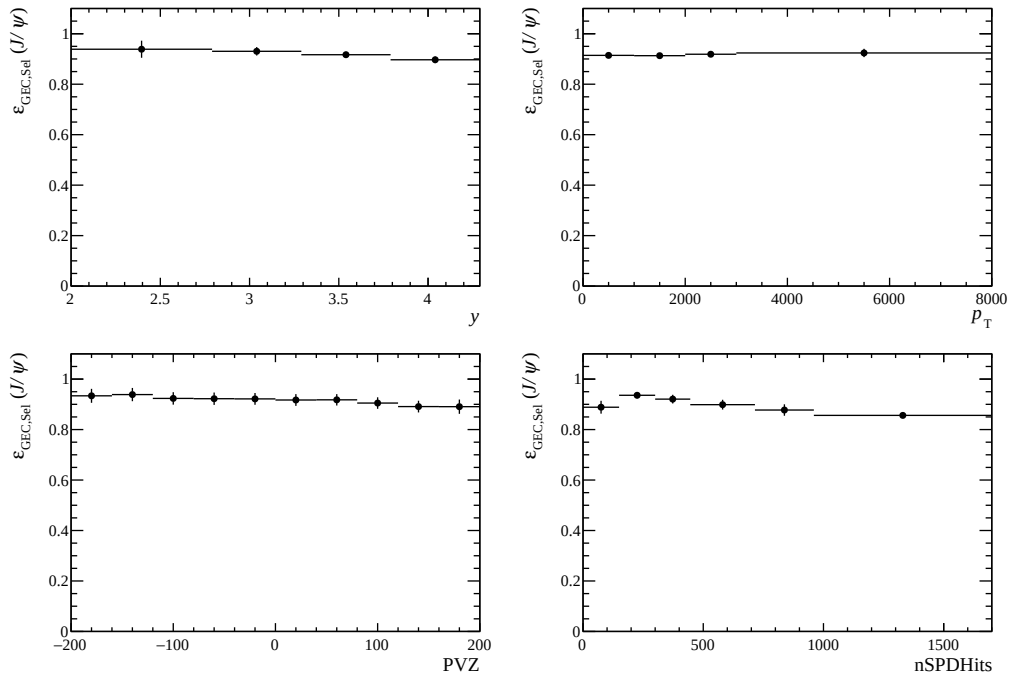

FIGURE 5.25: J/ψ GEC and selection efficiency as a function of y, p_T, PVZ and nSPDHits.

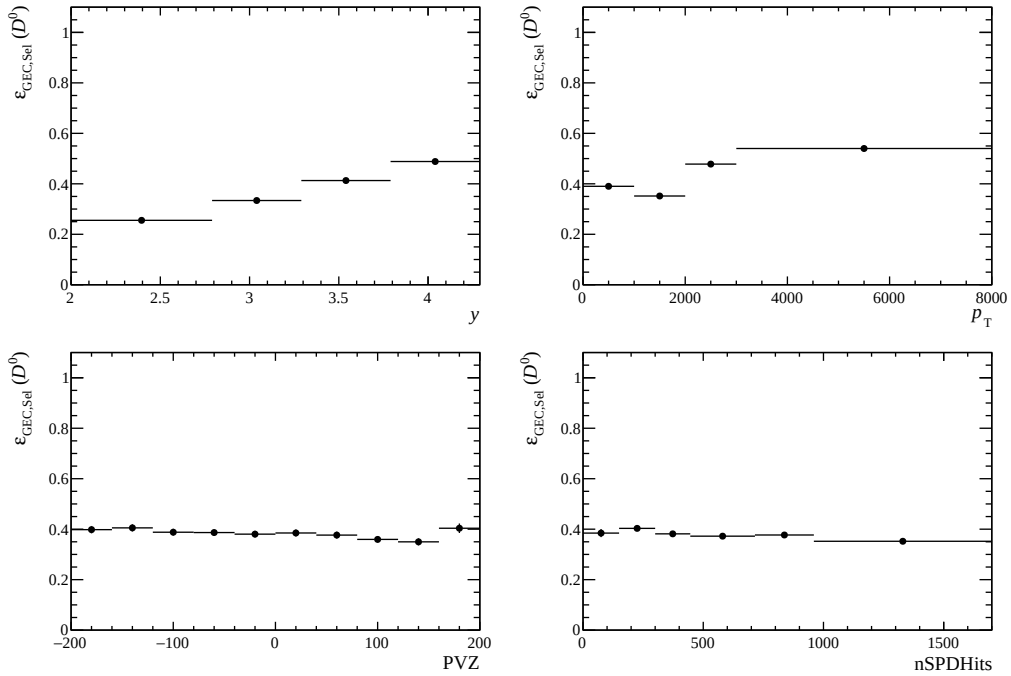

FIGURE 5.26: D^0 GEC and selection efficiency as a function of y, p_T, PVZ and nSPDHits.

5.5.8 PID efficiency

As previously discussed, current simulations do not reproduce well the detector occupancy, and thus they cannot be used to compute reliably the PID efficiencies. To overcome this issue, the PID efficiencies for the different particle species ($\mu^{\pm}$, $\pi^{\pm}$ and $K^{\pm}$) are obtained with the PIDCalib package. Since PbNe data are not present among the PIDCalib available datasets due the low statistics of the sample, the efficiencies are computed with data from Pbp collisions at $\sqrt{s_{\rm NN}} = 8$ TeV (see Sec. 5.4). The single track PID efficiencies are computed in bins of p, η and nSPDHits. The binning scheme, selection and PID selection applied are:

- The binning scheme:
  - p: [0 : 25000 : 50000 : 75000 : 110000 : 170000] MeV/c.
  - η: [2 : 2.79 : 3.29 : 3.79 : 4.29 : 4.6].
  - nSPDHits: [0 : 446 : 715 : 960 : 1700 : 2250].
- The selection:
  - For $\mu^{\pm}$: $p_{\rm T} > 400\,$MeV/c & $p > 3000\,$MeV/c & nPV > 0 & InMuonAcc $== 1$.
  - For $\pi^{\pm}$: $p_{\rm T} > 500\,$MeV/c.
  - For $K^{\pm}$: $p_{\rm T} > 500\,$MeV/c.
- The PID selection:[8]
  - For $\mu^{\pm}$: MC15TuneV1_ProbNNmu > 0.5 & IsMuon $== 1$.
  - For $\pi^{\pm}$: DLLK < 0.
  - For $K^{\pm}$: DLLK > 5.

The PIDCalib outputs are three dimensional (3D) histograms denoted as *passed* and *total*, which are filled with events passing the PID selection (*passed*) or all events (*total*). The ratio between *passed* and *total* constitutes the PID efficiency in a given bin. To obtain the PID efficiency maps, the 3D histograms are projected onto a two dimensional (2D) histogram and then divided. This can be done integrated in nSPDHits or for each nSPDHits bin. In Fig. 5.27 this procedure is shown, integrated in nSPDHits. The top row shows the 3D histograms, *passed* on the left and *total* on the right. The middle row shows their 2D projection integrated in nSPDHits. The bottom plot shows the PID map which consists on the division of the two 2D histograms.[9] Fig. 5.28, shows the same maps but for each nSPDHits bin (first four bins) and also the one dimensional projection into nSPDHits to illustrate the dependence of the PID efficiency on the multiplicity. Figs. 5.29 and 5.30 show the same things for the $\pi^{\pm}$, and Figs. 5.31 and 5.32 for the $K^{\pm}$.

[8] DLLK is the same variable as the aforementioned PIDK.

[9] It can occur that in some bins with low statistics the efficiency is greater than 1. This is because PIDCalib uses sWeighted data to evaluate the efficiencies. Consider as an example a given bin with 5 signal events with weights 1 and 10 background events with weights -0.4 giving a total bin content of 1. Once the selection is applied, 5 background events are removed while the signal is kept. The content of the bin would now be 3. This leads to the "total" histogram having bins with less content than in the "passed" histogram and thus an efficiency greater than 1. This however, does not occur in phase-space regions with high statistics.

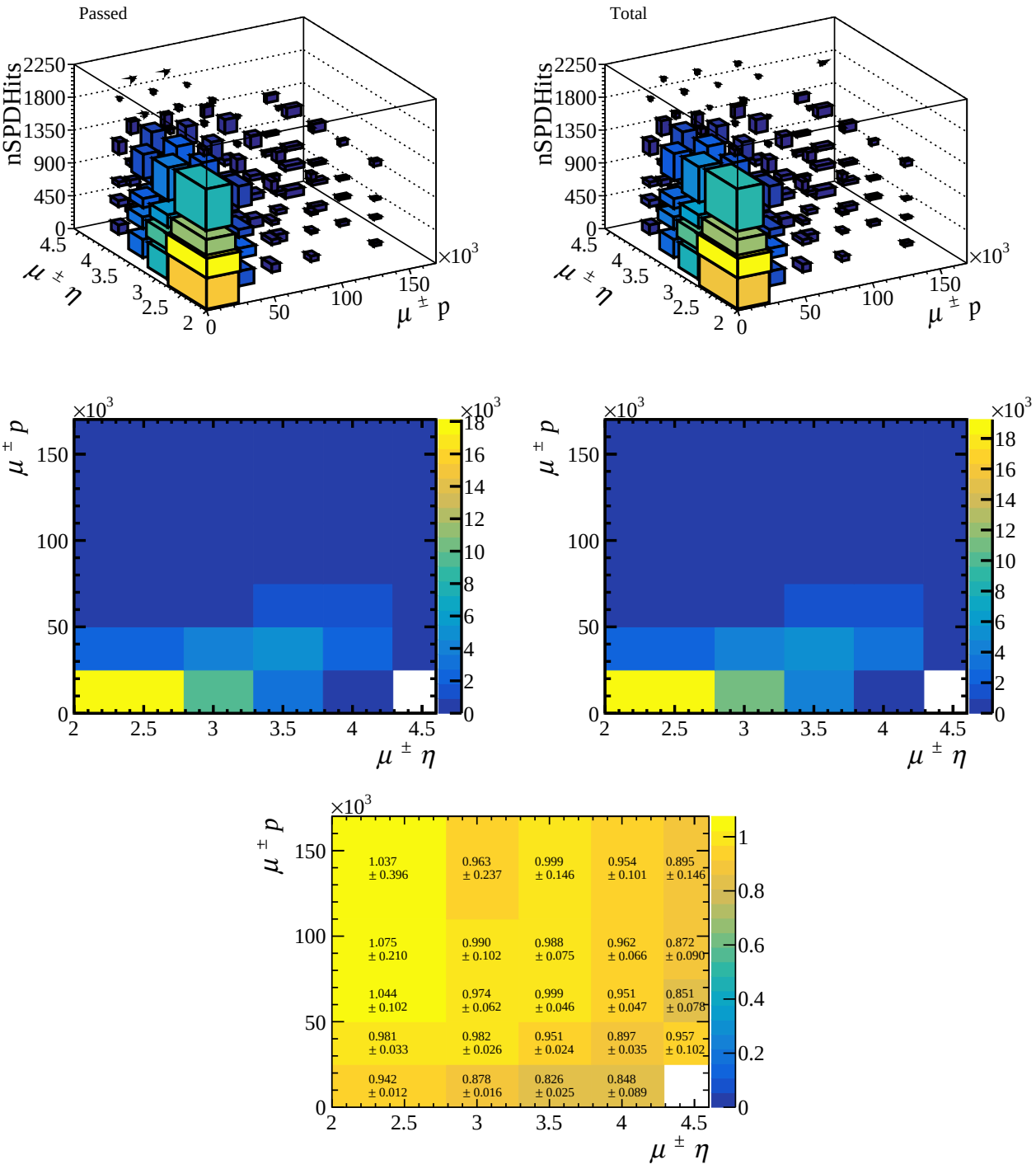

FIGURE 5.27: The output of PIDCalib for the $\mu^\pm$ tracks. The top row shows the 3D histograms filled with events passing the PID selection (*left*) and all the events (*right*). The middle row shows the 2D projection integrated in nSPDHits respectively for the *passed* and *total* histograms. The bottom plot shows the ratio between *passed* and *total*, which corresponds to the PID efficiency map integrated in nSPDHits.

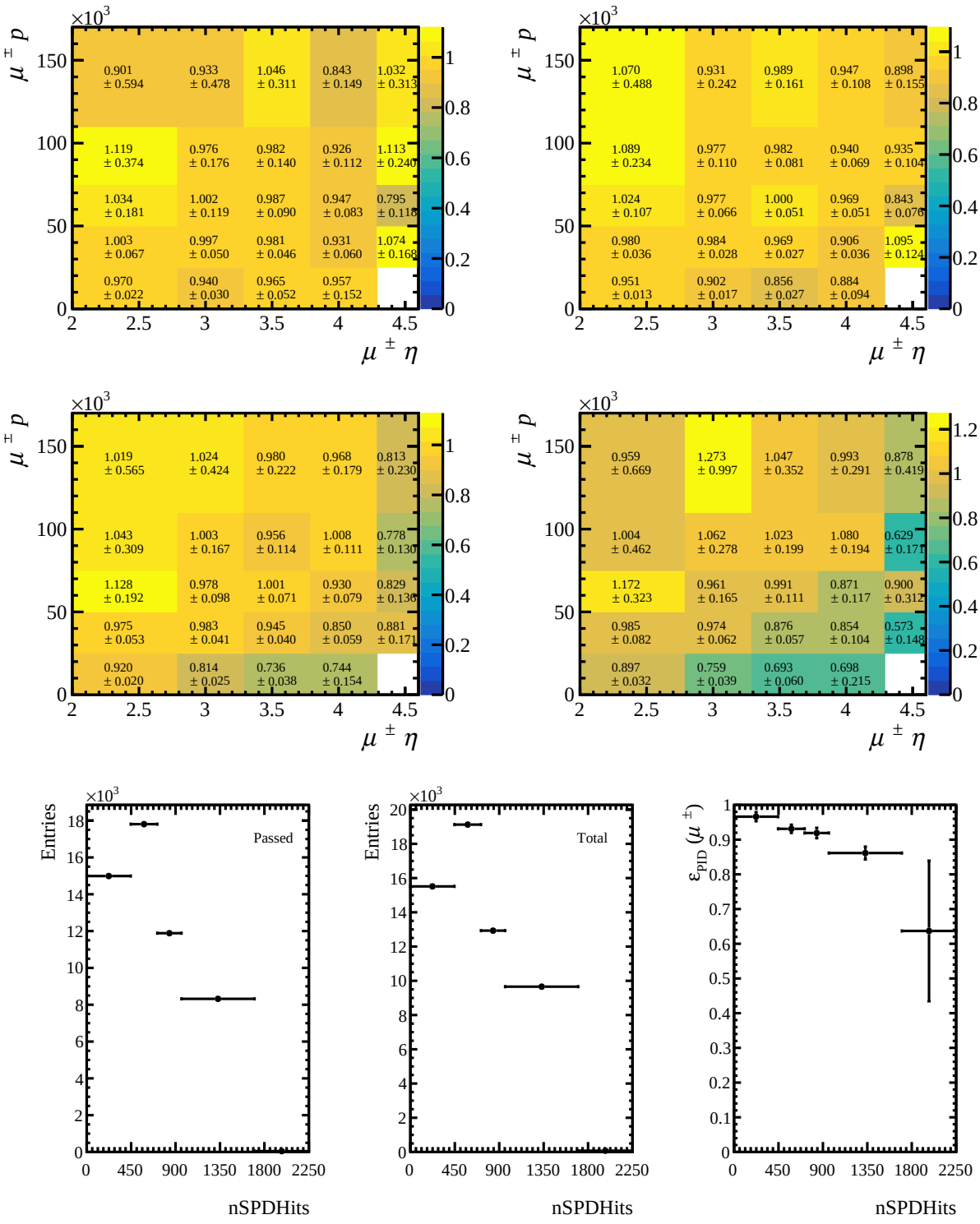

FIGURE 5.28: The 2D $\mu^{\pm}$ efficiency maps for the nSPDHits bins [0,446] (*top-left*), [446, 715] (*top-right*), [715, 960] (*middle-left*) and [960, 1700] (*middle-right*). On the bottom row, the 1D projections of the 3D histograms integrated in p and η are shown as well as their ratio to illustrate the PID efficiency dependence on multiplicity.

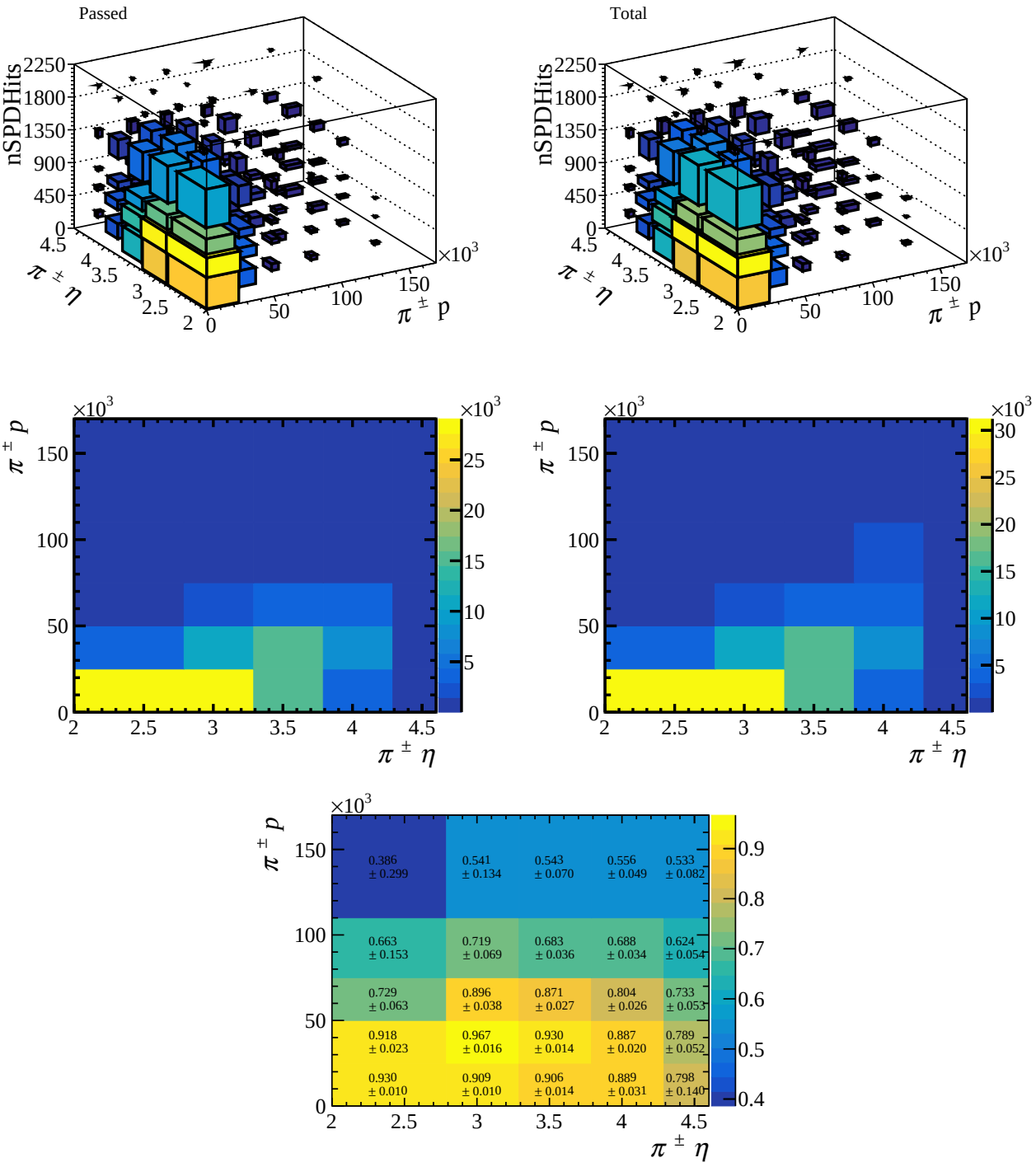

FIGURE 5.29: The output of PIDCalib for the $\pi^\pm$ tracks. The top row shows the 3D histograms filled with events passing the PID selection (*left*) and all the events (*right*). The middle row shows the 2D projection integrated in nSPDHits respectively for the *passed* and *total* histograms. The bottom plot shows the ratio between *passed* and *total*, which corresponds to the PID efficiency map integrated in nSPDHits.

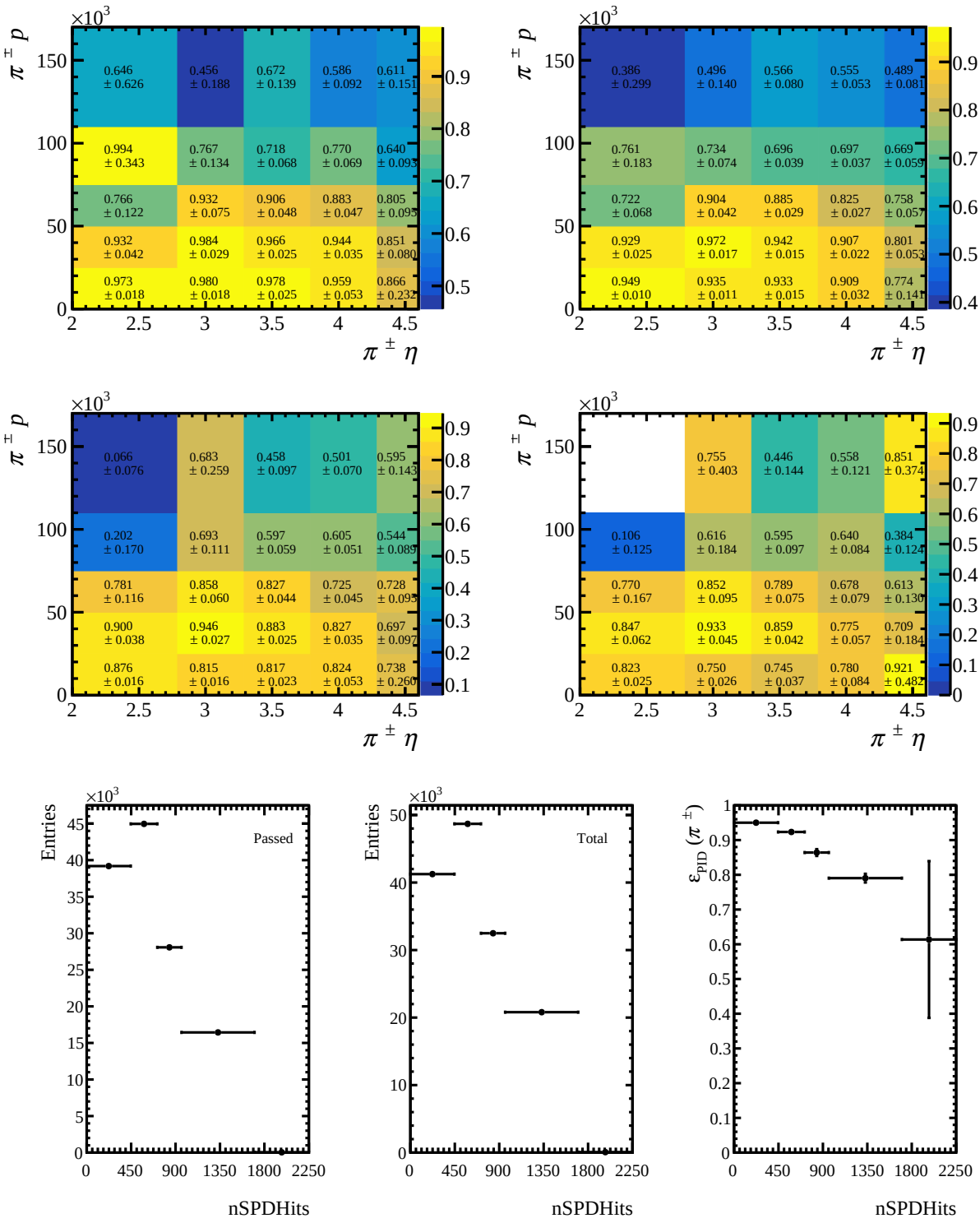

FIGURE 5.30: The 2D $\pi^{\pm}$ efficiency maps for the nSPDHits bins [0,446] (*top-left*), [446, 715] (*top-right*), [715, 960] (*middle-left*) and [960, 1700] (*middle-right*). On the bottom row, the 1D projections of the 3D histograms integrated in p and η are shown as well as their ratio to illustrate the PID efficiency dependence on multiplicity.

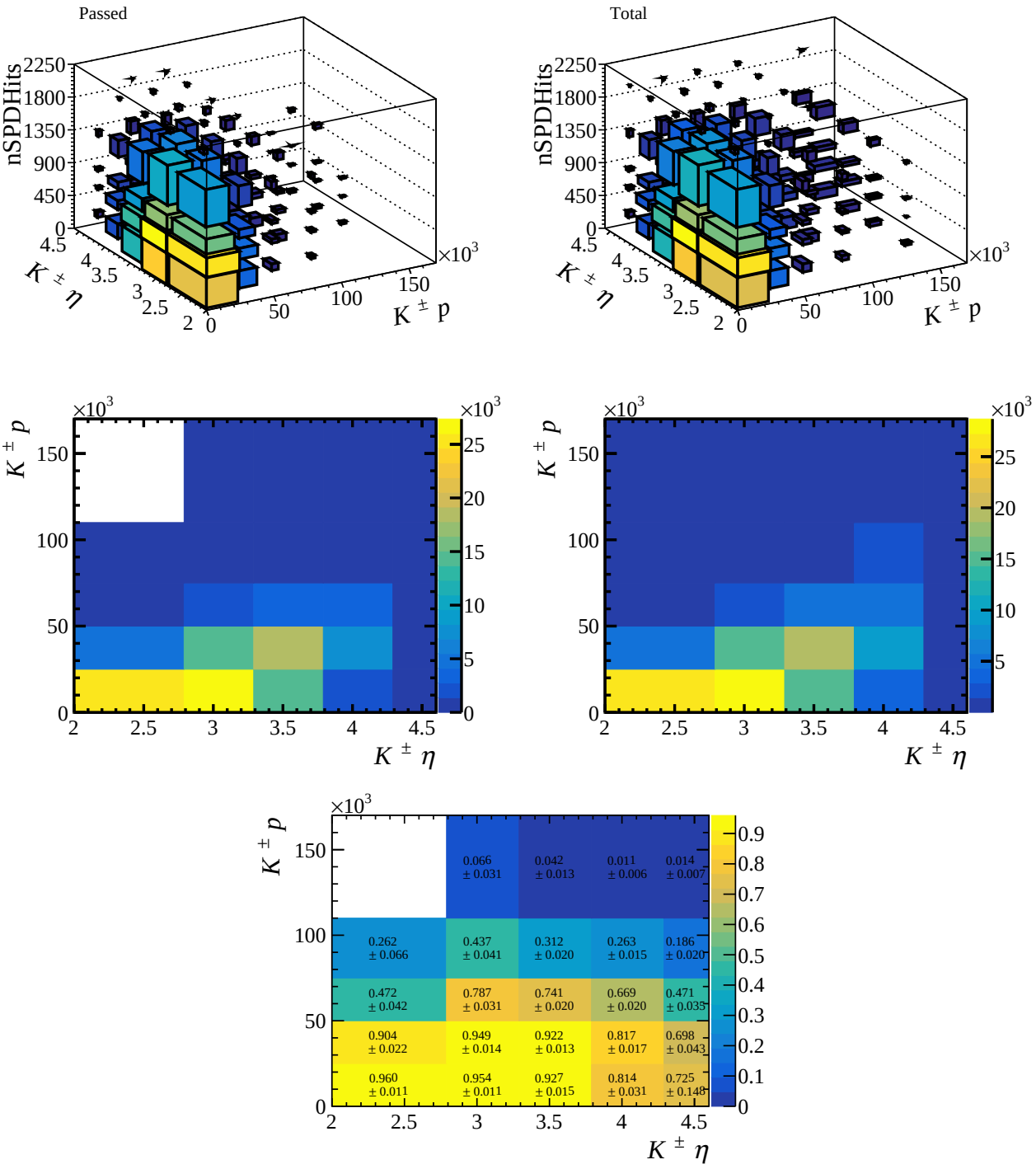

FIGURE 5.31: The output of PIDCalib for the $K^{\pm}$ tracks. The top row shows the 3D histograms filled with events passing the PID selection (*left*) and all the events (*right*). The middle row shows the 2D projection integrated in nSPDHits respectively for the *passed* and *total* histograms. The bottom plot shows the ratio between *passed* and *total*, which corresponds to the PID efficiency map integrated in nSPDHits.

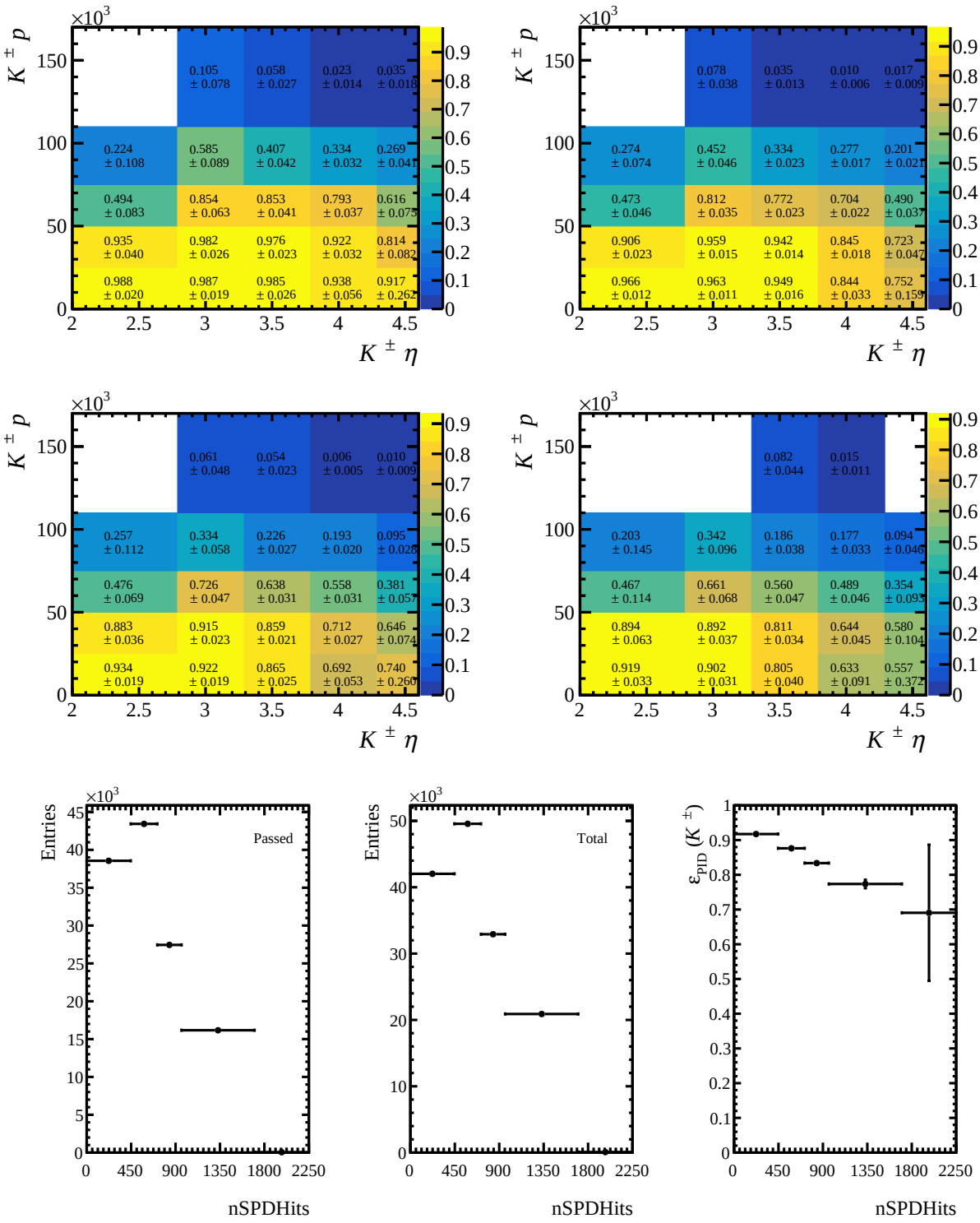

FIGURE 5.32: The 2D $K^{\pm}$ efficiency maps for the nSPDHits bins [0,446] (*top-left*), [446, 715] (*top-right*), [715, 960] (*middle-left*) and [960, 1700] (*middle-right*). On the bottom row, the 1D projections of the 3D histograms integrated in p and η are shown as well as their ratio to illustrate the PID efficiency dependence on multiplicity.

Finally, the candidate PID efficiency is computed with the PID maps in bins of nSPDHits. The procedure is as follows:

1. Each final track in the MC sample, with the $MC^{Reco}_{GEC,Sel}$ + PID selections applied, is assigned an efficiency from the PID maps according to their (η, p) and the event multiplicity (nSPDHits).

2. For each candidate, the PID efficiency is extracted for both of its daughter tracks, so $\varepsilon^{daughter1}_{\text{PID}}$ and $\varepsilon^{daughter2}_{\text{PID}}$.

3. Each candidate is then weighted by the product of their daughter efficiency, that is, by $\varepsilon^{signal}_{\text{PID}} = \varepsilon^{daughter1}_{\text{PID}} \times \varepsilon^{daughter2}_{\text{PID}}$.

In other words, defining $MC^{Reco}_{PID} = MC^{Reco}_{GEC,Sel} + PID$, the PID efficiency is computed as:

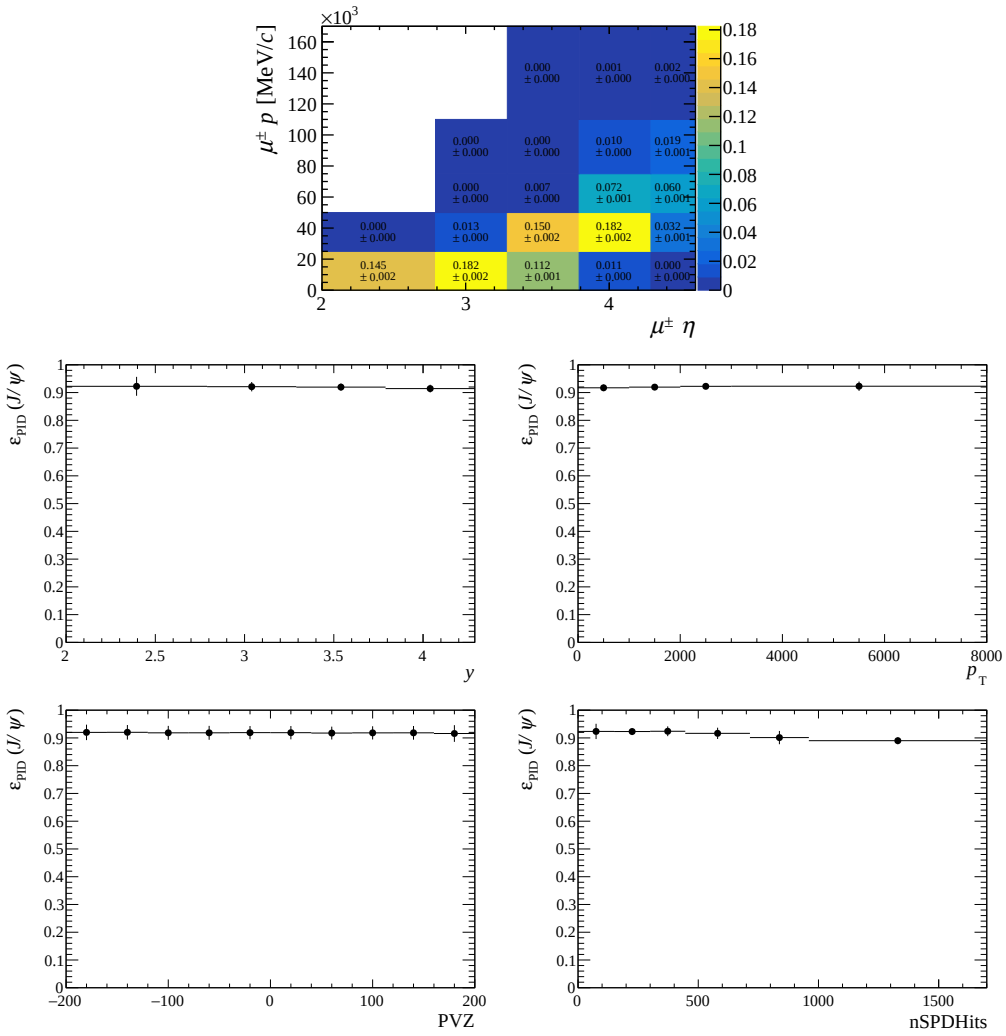

FIGURE 5.33: The top row shows the daughter occupancy of the (p, η) space for the J/ψ, where the content of each bin represents the fraction of the total number of daughters within the given bin. The middle and bottom rows show the PID efficiency as a function of y, p_T, PVZ and nSPDHits.

$$\varepsilon_{\text{PID}} = \frac{MC^{Reco}_{PID} \times \varepsilon^{daughter1}_{\text{PID}}(p,\eta,\text{nSPDHits}) \times \varepsilon^{daughter2}_{\text{PID}}(p,\eta,\text{nSPDHits})}{MC^{Reco}_{PID}}. \tag{5.12}$$

It is useful to consider the occupation of the (p,η) space by the single tracks, since it can help understand the resulting efficiencies for the final candidates. For the J/ψ, the fraction of $\mu^{\pm}$ tracks occupying each (p,η) bin is shown on the top row of Fig. 5.33. On the bottom and middle row the PID efficiency is shown as a function of y, p_{T}, PVZ and nSPDHits. Indeed the decrease as a function of nSPDHits is less pronounced than what the bottom right plot of Fig. 5.28 suggests, and this can be explained by the fact that most muons occupy a (p,η) region, where the efficiency does not vary dramatically as the nSPDHits increases. For the D^0, the daughter occupancy plots and PID efficiency as a function of y, p_{T}, PVZ and nSPDHits are shown in Fig. 5.34.

The resulting integrated efficiencies are:

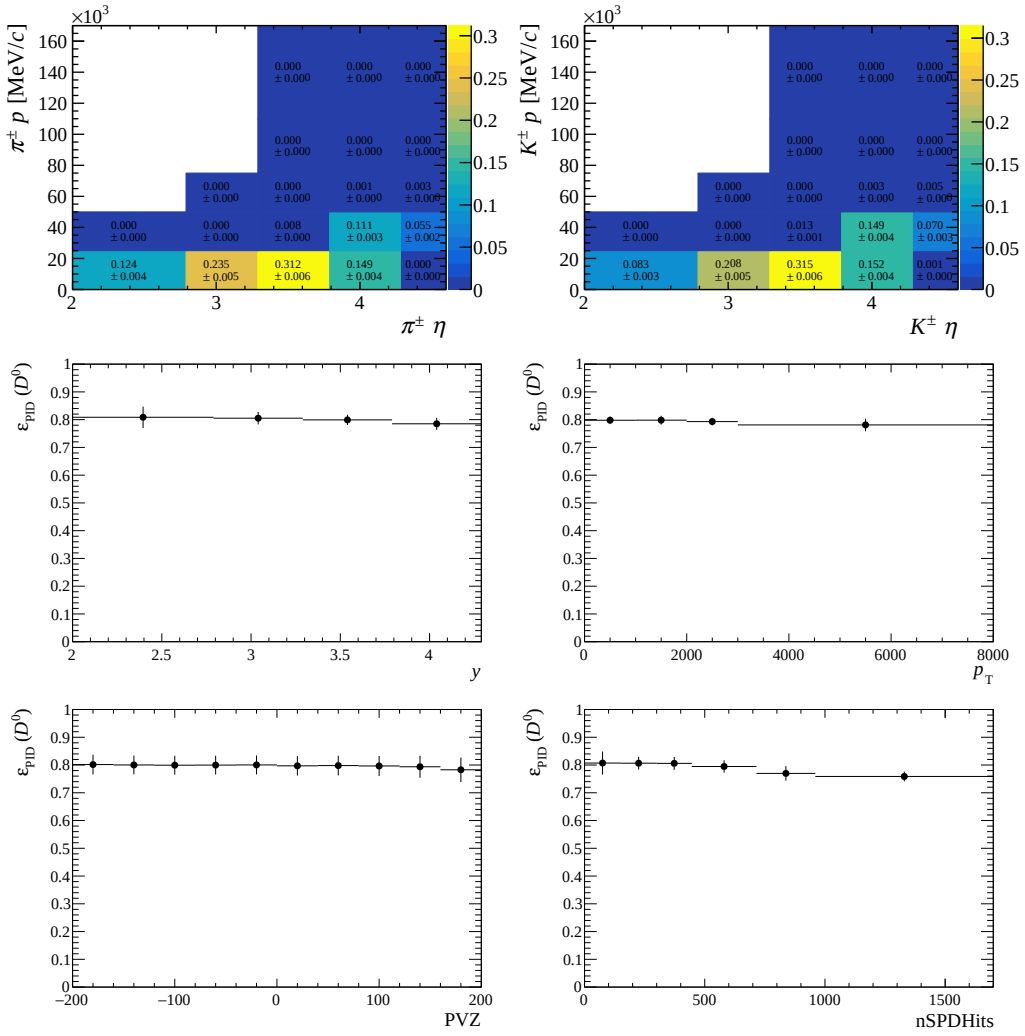

FIGURE 5.34: The top row shows the daughter occupancy of the (p,η) space for the D^0, where the content of each bin represents the fraction of the total number of daughters within the given bin. The middle and bottom rows show the PID efficiency as a function of y, p_{T}, PVZ and nSPDHits.

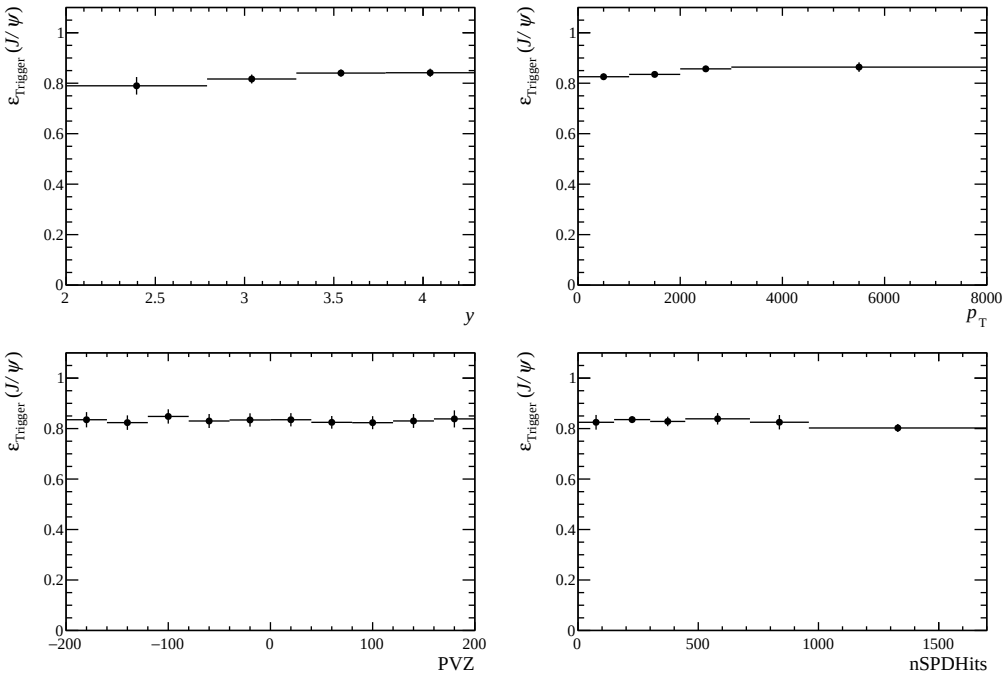

FIGURE 5.35: J/ψ trigger efficiency as a function of y, p_T, PVZ and nSPDHits.

- $\varepsilon^{J/\psi}_{\text{PID}} = (91.84 \pm 0.01)\%$.
- $\varepsilon^{D^0}_{\text{PID}} = (84.40 \pm 0.04)\%$.

5.5.9 Trigger efficiency

The trigger efficiencies are determined with the reconstructed MC samples, right after the PID efficiency evaluation to avoid introducing any bias due to the PID cuts at the trigger level. The trigger efficiencies include L0 and HLT1 requirements. No HLT2 requirements has been applied, all events are HLT2 "pass-through".

Defining $MC^{Reco}_{Trigger} = MC^{Reco}_{PID} + \text{L0} + \text{HLT1}$, the trigger efficiency is computed as:

$$\varepsilon_{\text{Trigger}} = \frac{MC^{Reco}_{\text{Trigger}}}{MC^{Reco}_{\text{PID}}}. \tag{5.13}$$

The overall integrated trigger efficiency is found to be:

- $\varepsilon^{J/\psi}_{\text{Trigger}} = (83.17 \pm 0.06)\%$.
- $\varepsilon^{D^0}_{\text{Trigger}} = (99.66 \pm 0.01)\%$.

The trigger efficiencies as a function of the y, p_T, PVZ and nSPDHits are shown in Figs. 5.35 and 5.36 for J/ψ and D^0 respectively.

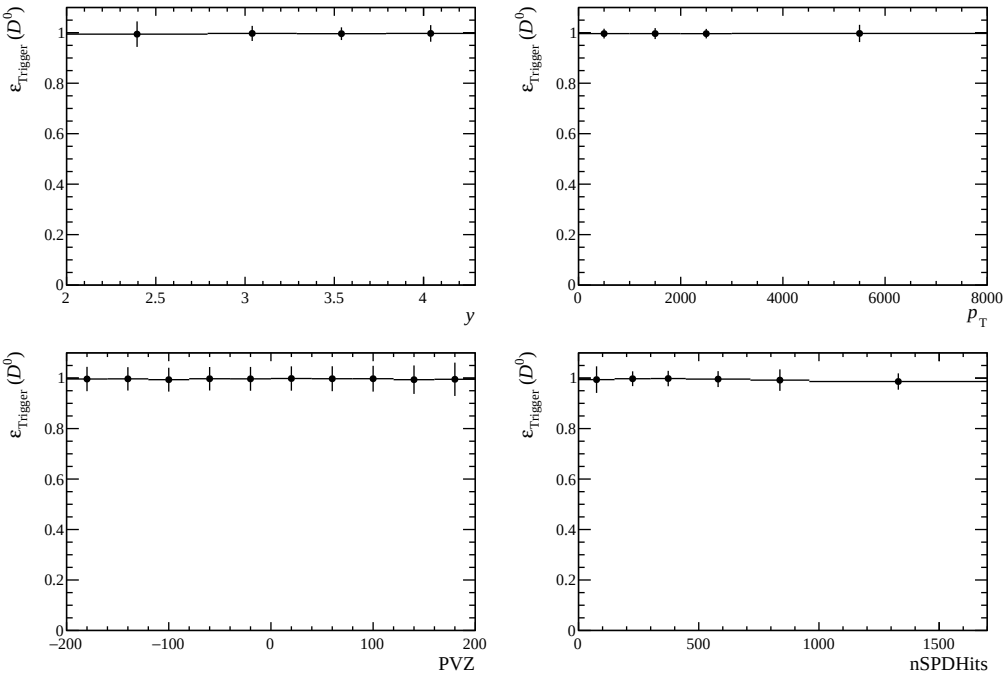

FIGURE 5.36: D^0 trigger efficiency as a function of y, p_T, PVZ and nSPDHits.

TABLE 5.4: Summary of efficiencies for J/ψ and D^0. The two rightmost columns show the efficiencies for the nSPDHits interval [0,446].

Efficiency	J/ψ	D^0	$J/\psi_{[0,446]}$	$D^0_{[0,446]}$
Acceptance	33.23 ± 0.03 %	35.08 ± 0.03 %	33.23 ± 0.03 %	35.08 ± 0.03 %
PV	99.98 ± 0.00 %	99.99 ± 0.00 %	99.97 ± 0.94 %	99.98 ± 0.50 %
Track	80.42 ± 0.05 %	46.63 ± 0.01 %	89.84 ± 0.87 %	52.66 ± 0.31 %
Stripping	92.67 ± 0.29 %	34.76 ± 0.16 %	90.47 ± 0.92 %	32.84 ± 0.32 %
GEC, Sel	91.45 ± 0.04 %	38.33 ± 0.02 %	92.89 ± 0.99 %	39.16 ± 0.63 %
PID	91.84 ± 0.01 %	84.40 ± 0.04 %	92.06 ± 0.02 %	80.34 ± 0.07 %
Trigger	83.17 ± 0.06 %	99.66 ± 0.01 %	83.28 ± 1.07 %	99.78 ± 2.04 %
Total	17.30 ± 0.06 %	1.83 ± 0.01 %	23.04 ± 0.66 %	1.90 ± 0.05 %

5.5.10 Total efficiencies

The total efficiencies are obtained as a product of all the previously computed efficiencies. In the case of the efficiency as a function of nSPDHits, there is a factor applied to account for the overall acceptance efficiency. The resulting total efficiency as a function of y, p_T, PVZ and nSPDHits can be seen in Figs. 5.37 and 5.38 for J/ψ and D^0 respectively. The signal is corrected using the efficiencies of the corresponding binning variable.

The overall total efficiencies for J/ψ and D^0 are summarised in Tab. 5.4. The efficiencies for the nSPDHits interval [0,446] are also shown.

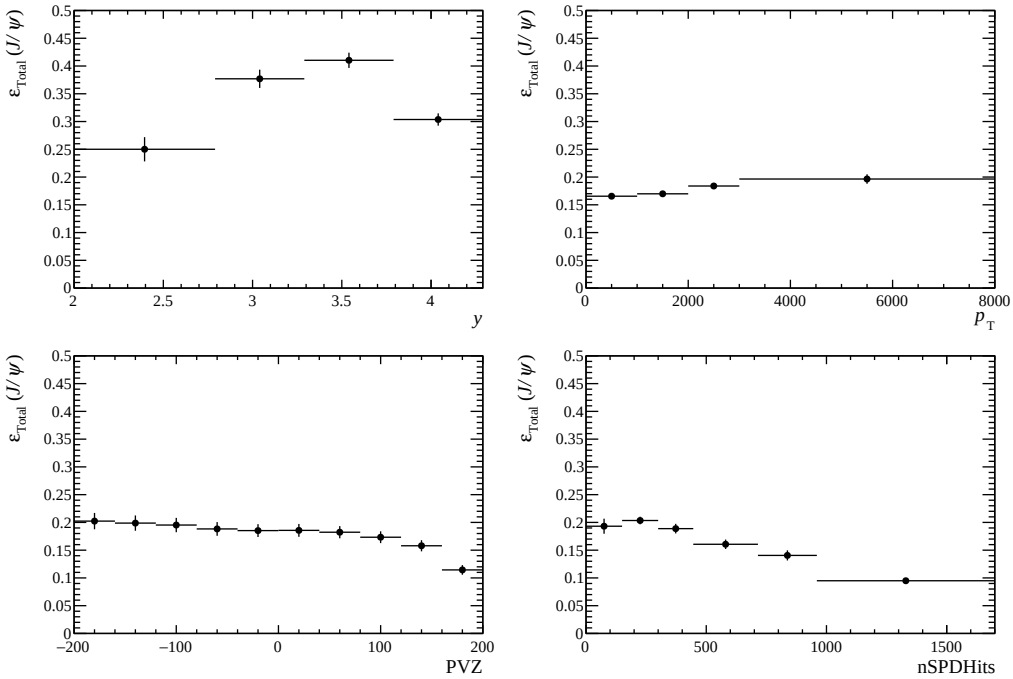

FIGURE 5.37: J/ψ total efficiency as a function of y, p_{T}, PVZ and nSPDHits.

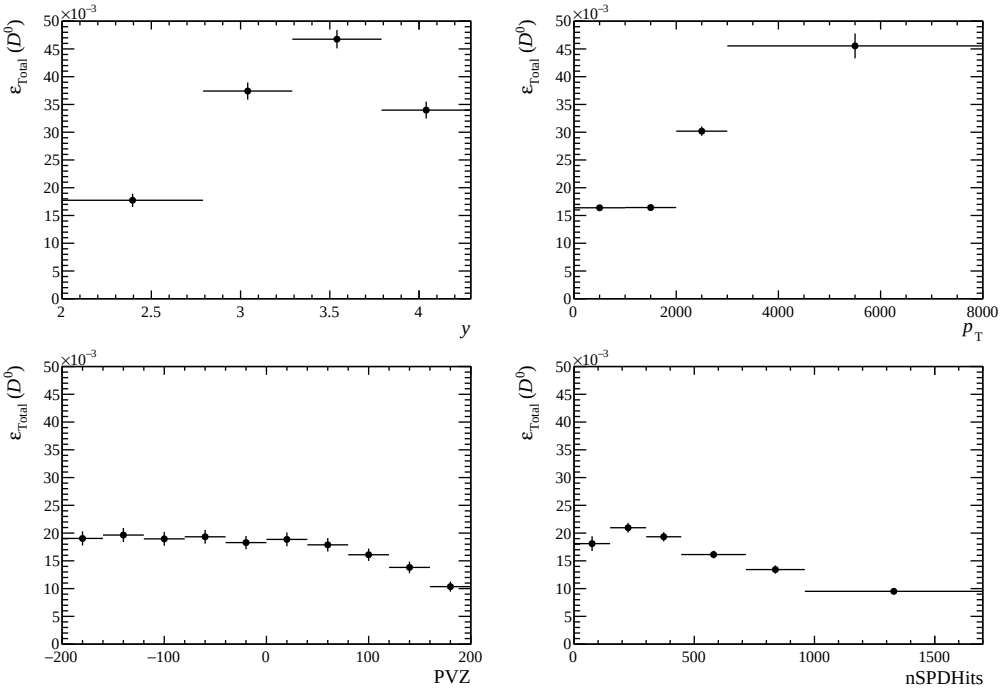

FIGURE 5.38: D^0 total efficiency as a function of y, p_{T}, PVZ and nSPDHits.

TABLE 5.5: Systematic uncertainties due to the MC statistics in nSPDHits bins. Values are in percentages.

Particle	[0, 200]	[200, 300]	[300, 446]	[446, 715]	[715, 960]	[960, 1700]
J/ψ	0.20	0.60	0.57	0.47	0.53	0.39
D^0	0.19	0.43	0.40	0.32	0.34	0.25

TABLE 5.6: Systematic uncertainties due to the MC statistics in y bins. Values are in percentages.

Particle	[2.0, 2.79]	[2.79, 3.29]	[3.29, 3.79]	[3.79, 4.29]
J/ψ	0.71	0.35	0.25	0.25
D^0	0.32	0.23	0.21	0.22

5.6 Systematic uncertainties

In this section the systematic uncertainties affecting the efficiency calculations and the signal extraction are identified and computed. In order to do this, the MC samples used, the models to extract the signal, the MC reweighting and the different efficiency computation procedures are considered.

5.6.1 Monte Carlo

The MC samples are a key part of the measurement and, as such, any uncertainties arising from the MC may affect the measurements. Three kinds of uncertainties are considered: the statistics, the truth-matching efficiency and the agreement between MC and data.

Monte Carlo statistics

To evaluate this systematic uncertainty, the ratio of the MC sample with all the requirements applied with respect to the MC sample with only the acceptance requirements is considered. The uncertainty in each bin corresponds to the statistical uncertainty of the sample used. In addition, for each bin the uncertainty is added in quadrature with the statistical uncertainty from the private MC sample which was used for the computation of the acceptance efficiency.

The total uncertainties for each of the variables are shown in Tabs. 5.5, 5.6, 5.7 and 5.8 for the nSPDHits, y, p_{T} and PVZ respectively. All values are in percentages.

MC truth-match

The stripping efficiency had to be corrected by a truth-matching efficiency correction. This factor carried with it a statistical and systematic part. From the point of view of the results, both correspond to a systematic uncertainty, of which one part comes from the statistics of the unmatched candidates (the one that was called "statistical" in

TABLE 5.7: Systematic uncertainties due to the MC statistics in p_{T} bins. The bins are in MeV/c and values are in percentages.

Particle	[0, 1000]	[1000, 2000]	[2000, 3000]	[3000, 8000]
J/ψ	0.34	0.24	0.30	0.45
D^0	0.19	0.18	0.31	0.49

TABLE 5.8: Systematic uncertainties due to the MC statistics in PVZ bins. Values are in percentages.

Particle	$[-200,-120]$	$[-120,-40]$	$[-40,40]$	$[40,120]$	$[120,200]$
J/ψ	0.50	0.48	0.47	0.48	0.52
D^0	0.40	0.37	0.35	0.35	0.42

Sec. 5.5.6), and another part from the modelling of the signal to extract the remaining unmatched signal (the one referred to as "systematic" in Sec. 5.5.6) which was null for the D^0. To assign an overall systematic value from the truth-matching, the statistic and systematic uncertainties are added in quadrature resulting in an overall factor of

- 0.53% for the J/ψ
- 0.43% for the D^0

These systematic uncertainties will be applied homogeneously to all bins.

The MC reweighting

The reweighting of the MC sample brings with it fluctuations due to the random splitting of the dataset into training and test sets. To account for this effect, the reweighter was trained 10 times with random dataset splittings and then the weighted rapidity distributions were compared for all the different weight-sets. For each bin the standard deviation was computed and then averaged over all bins to obtain a systematic uncertainty. This was done separately for D^0 and J/ψ, and for each case, the reported uncertainty is the mean of the values obtained in the truth and the reconstructed MC samples. The resulting values are

- 0.29% for J/ψ
- 0.12% for D^0

These systematic uncertainties will be applied homogeneously to all bins. The rapidity distribution plots can be seen in Appendix B.1.

5.6.2 Primary vertex

Given that SMOG events present a PV displaced from $Z = 0$, it is important to consider how this affects the final result. Two different effects are considered. First, the difference in the result from the different MC samples (J/ψ or D^0) and second, the variation of the efficiencies with respect to the PV position. The latter effect is of particular importance because when results are binned in any other variable, they are integrated in PVZ.

PV efficiency

The PV efficiency was computed in Sec. 5.5.4. This efficiency, as was mentioned, should not depend on the particle under study, that is, it should be the same for J/ψ and D^0. Because of this the difference of the values found in each case is taken as a systematic uncertainty, resulting in an uncertainty of 0.01%.

TABLE 5.9: Systematic uncertainty for each efficiency due to the PVZ dependence for J/ψ and D^0. Values are in percentages.

Efficiency	J/ψ	D^0
acceptance	0.09	0.09
PV	0.02	0.01
Track	5.95	4.22
Strip	0.47	0.57
GEC, Sel	1.18	1.35
PID	0.10	0.33
Trigger	0.59	0.14
Total	6.11	4.48

PVZ dependence

To compute this systematic uncertainty the variation in PVZ of every efficiency is considered. This is done separately for J/ψ and D^0. For any given efficiency, the value obtained at PVZ $= 0$, which will be referred as $\varepsilon_{Z=0}$, is estimated by averaging the values obtained for PVZ $\in [-40, 0]$ and PVZ $\in [0, 40]$. Then, the absolute value of the difference of the value of each bin with respect to $\varepsilon_{Z=0}$ is computed and averaged over all bins. Finally, the average for each efficiency is added in quadrature to obtain an overall factor. The results are reported in Tab. 5.9, with a total systematic uncertainty of 6.11% for the J/ψ and 4.48% for the D^0. These values will be applied homogeneously to all bins.

5.6.3 Signal extraction

The number of signal candidates used for the analysis may be affected by the models used to describe the data, and also by possible contamination coming from ghost PbPb collisions. These effects will be discussed in the following.

Signal modelling

To obtain the signal the following component description is used:

For J/ψ:

- Signal: Crystal ball function.
- Background: Exponential function.

For D^0:

- Signal: 2 Gaussian functions.
- Background: Exponential function.

These models are fit in every bin separately.[10] To evaluate the systematic uncertainty related to the modelling, different approaches are used to describe the different components. In the case of the J/ψ, the different approaches were:

- Template fit: use the parameters for the crystal ball function found in the integrated data for the fits performed in each bin.

[10] For the J/ψ, when performing the fit in the binned data, the parameters for the radiative tail were fixed to the values found for the fit of the integrated data.

TABLE 5.10: Systematic uncertainties due to the signal extraction model in nSPDHits bins. Values are in percentages.

Particle	[0, 200]	[200, 300]	[300, 446]	[446, 715]	[715, 960]	[960, 1700]
J/ψ	3.28	3.84	2.36	3.20	3.41	3.90
D^0	2.63	2.10	3.06	2.97	1.36	5.54

TABLE 5.11: Systematic uncertainties due to the signal extraction model in y bins. Values are in percentages.

Particle	[2.0, 2.79]	[2.79, 3.29]	[3.29, 3.79]	[3.79, 4.29]
J/ψ	3.70	2.54	3.99	4.77
D^0	6.86	7.21	2.37	5.43

- Describe the background with a 1^{st} degree polynomial instead of an exponential function.
- Describe the signal with a single Gaussian function. To do this, the width of the Gaussian, in the case of the fits to the binned data, was fixed to the value found for the integrated case.

In the case of the D^0, these were:

- Template fit: use the parameters for the double Gaussian function found in the integrated data for the fits performed in each bin.
- Describe the background with a 1^{st} degree polynomial instead of an exponential function.
- Describe the signal with a single Gaussian function.

Finally, the extracted signal with each of the fit variations is compared to the signal obtained with the nominal model, and the RMS of the differences is taken to be the systematic uncertainty. The resulting systematic uncertainty found for each bin in nSPDHits, y, p_{T} and PVZ can be seen in Tabs. 5.10, 5.11, 5.12 and 5.13 respectively. All values are in percentages. The extracted yields can be seen in Appendix B.2.

PbPb ghost contamination

As discussed in Sec. 5.3.1, during the data taking there is some residual contamination from parasitic PbPb collisions due to the Pb nuclei that migrate to a nominally empty bunch in beam 2. The procedure to determine the GEC requirements to limit this contamination and the uncertainty that arises from it is detailed in Ref. [149]. The uncertainty due to this effect is determined to be 0.23%. However, this uncertainty cancels out in the J/ψ and D^0 ratio measurement; therefore, no systematic uncertainty is assigned for the ratio measurement.

TABLE 5.12: Systematic uncertainties due to the signal extraction model in p_{T} bins. The bins are in MeV/c and values are in percentages.

Particle	[0, 1000]	[1000, 2000]	[2000, 3000]	[3000, 8000]
J/ψ	6.68	3.74	4.06	9.48
D^0	1.65	4.75	1.57	1.83

TABLE 5.13: Systematic uncertainties due to the signal extraction model in PVZ bins. Values are in percentages.

Particle	$[-200,-120]$	$[-120,-40]$	$[-40,40]$	$[40,120]$	$[120,200]$
J/ψ	3.21	3.90	3.00	2.91	3.71
D^0	0.33	4.20	3.48	3.70	3.25

TABLE 5.14: Systematic uncertainties due to the tracking correction factors in nSPDHits bins. Values are in percentages.

Particle	$[0,200]$	$[200,300]$	$[300,446]$	$[446,715]$	$[715,960]$	$[960,1700]$
J/ψ	7.26	7.09	6.86	5.88	5.38	3.94
D^0	6.57	6.52	6.37	5.49	4.76	3.66

5.6.4 Tracking efficiency

The tracking efficiency is computed from the MC samples, with correction factors (see Fig. 5.18) computed with TrackCalib, a software tool developed within LHCb that has been extensively used. For this analysis, the correction factor table used was initially developed for Pbp collisions, which is compatible in multiplicity with the present PbNe data. A systematic uncertainty of 0.5% for each muon track and of 2% for each kaon or pion track is reported, which translates into 0.70% for the J/ψ and 2.83% for the D^0.[11] The greater uncertainty in the case of the kaon and pion tracks comes from the fact that hadrons are prone to be lost due to hadronic interactions, and thus from the uncertainty on the material budget, there is an additional uncertainty ultimately affecting the D^0.

In addition, the statistical uncertainty in the computation of these correction factors should be considered. To do this, the tracking efficiency is computed with the correction factors modified within their statistical uncertainties, that is, taking their central value minus 1σ and subsequently plus 1σ. Finally the RMS of the differences with respect to the nominal value is considered as the associated systematic uncertainty. The values obtained are summarised in Tabs. 5.14, 5.15, 5.16 and 5.17 for nSPDHits, y, p_{T} and PVZ respectively.

5.6.5 PID efficiency

In a similar way to the tracking efficiency, the PID efficiency was also computed with weights obtained with the PIDCalib software, a tool developed by the LHCb collaboration. The weights were obtained from Pbp data, and their statistical uncertainty is used as a systematic uncertainty. To do so, the PID efficiency is computed by varying the weights within their uncertainty, and the RMS with respect to the nominal value is quoted as the corresponding systematic uncertainties. These are summarised in Tabs. 5.18, 5.19, 5.20 and 5.21 for nSPDHits, y, p_{T} and PVZ respectively. Contrary to

[11] These values are recommended by Ref. [153].

TABLE 5.15: Systematic uncertainties due to the tracking correction factors in y bins. Values are in percentages.

Particle	$[2.0,2.79]$	$[2.79,3.29]$	$[3.29,3.79]$	$[3.79,4.29]$
J/ψ	7.24	6.08	6.04	6.57
D^0	4.13	6.09	6.59	5.22

TABLE 5.16: Systematic uncertainties due to the tracking correction factors in p_T bins. The bins are in MeV/c and values are in percentages.

Particle	$[0, 1000]$	$[1000, 2000]$	$[2000, 3000]$	$[3000, 8000]$
J/ψ	6.17	6.19	6.97	8.14
D^0	5.71	5.92	5.88	5.99

TABLE 5.17: Systematic uncertainties due to the tracking correction factors in PVZ bins. Values are in percentages.

Particle	$[-200, -120]$	$[-120, -40]$	$[-40, 40]$	$[40, 120]$	$[120, 200]$
J/ψ	7.17	6.82	6.68	6.62	4.98
D^0	6.32	6.17	6.08	5.78	4.32

what is expected, the uncertainties for the muon identification are larger than the ones found for the kaons and pions, resulting in a systematically higher uncertainty for the J/ψ as compared to the D^0. This arises from the fact that in the data samples used to create the PID maps the J/ψ statistics are very low and thus results in some bins with large uncertainties, directly affecting this result. This effect is still under study.

5.6.6 Gas purity

Outgassing from the beam-pipes or components of the SMOG device, which is mainly hydrogen, could in principle dilute the Ne gas affecting its purity and with it, the purity of the data sample. In order to study this effect some runs were recorded with the full SMOG injection configuration, but without the actual gas injection. These runs[12] amount to 8 hours and 23 minutes of data taking with the filling scheme FS10 (See Tab. 3.1). Considering the filling scheme used, where 8 hours and 23 minues of data taking with FS10 correspond to 2.29% of all the recorded data,[13] the contamination due to outgassing can be estimated. After the full selection described in Sec. 5.3, including the GEC to avoid a double counting of PbPb ghost contamination, is applied to the recorded events of these runs, there are no remaining J/ψ nor D^0 candidates. To set a limit on the possible contamination it can be argued that:

- The expected number of selected D^0 candidates in this time window with Ne SMOG injection would be 3289 events.
- No candidate is selected, so assuming a scenario where there are two candidates detected during the no-gas runs, these would correspond to 0.06% (2/3289) of the expected SMOG events.

[12]The runs were all recorded during fill 7481, and correspond to runs 218872, 218871, 218870, 218869, 218868, 218867, 218866, 218865, 218864, 218863, 218862, 218861 and 218859.

[13]One hour of data taking with this filling scheme corresponds to 0.27% of the total data recorded.

TABLE 5.18: Systematic uncertainties due to the PID weights in nSPDHits bins. Values are in percentages.

Particle	$[0, 200]$	$[200, 300]$	$[300, 446]$	$[446, 715]$	$[715, 960]$	$[960, 1700]$
J/ψ	9.02	9.05	9.02	8.52	8.95	9.53
D^0	3.05	3.16	3.14	2.87	3.03	3.39

TABLE 5.19: Systematic uncertainties due to the PID weights in y bins. Values are in percentages.

Particle	$[2.0, 2.79]$	$[2.79, 3.29]$	$[3.29, 3.79]$	$[3.79, 4.29]$
J/ψ	8.36	8.61	8.87	9.44
D^0	2.81	2.87	3.07	3.43

TABLE 5.20: Systematic uncertainties due to the PID weights in p_{T} bins. The bins are in MeV/c and values are in percentages.

Particle	$[0, 1000]$	$[1000, 2000]$	$[2000, 3000]$	$[3000, 8000]$
J/ψ	8.94	8.96	9.10	9.24
D^0	3.09	3.07	3.07	3.08

- In addition, given that the possible contamination comes mainly from H, and due to the factor 20 between the heavy-flavour cross-section of Ne and H, the contamination limit would be further reduced to $(0.06/20)\% \sim 0.003\%$.

With this approximation the contamination due to the outgassing is considered to be negligible.

5.6.7 Total systematic uncertainties

The different systematic uncertainties previously described are summarised in Tab. 5.22. In total, the correlated systematic uncertainties stand for 6.16% and 4.92% for J/ψ and D^0 respectively, while the uncorrelated are within $[\pm 10.23, \pm 13.90]$ and $[\pm 5.81, \pm 9.87]$ for J/ψ and D^0 respectively, as detailed in Tabs. 5.23, 5.24, 5.25 and 5.26 for nSPDHits, y, p_{T} and PVZ respectively. The ratio of J/ψ over D^0 could in principle benefit from some systematic uncertainty cancellation, such as the PV efficiency and ghost contamination, but the effect in the final systematic uncertainties is marginal.

From Tab. 5.22 it is apparent that the systematic uncertainties are of the same order as the statistical uncertainties previously found for the data. The systematic uncertainties are dominated by the uncorrelated uncertainties and, in particular, by the statistical uncertainty related to the tracking correction factors and to the weights for the PID efficiency which are computed from Pbp data samples. These could be greatly reduced in the future either by including larger Pbp data samples in their computation, or by generating these factors and weights specifically from PbNe (or other system under study) assuming that there will be much more data available with the future upgrades of LHCb.

TABLE 5.21: Systematic uncertainties due to the PID weights in PVZ bins. Values are in percentages.

Particle	$[-200, -120]$	$[-120, -40]$	$[-40, 40]$	$[40, 120]$	$[120, 200]$
J/ψ	9.02	9.07	9.01	8.91	8.87
D^0	3.12	3.07	3.16	3.03	3.04

TABLE 5.22: Summary of the systematic uncertainties on J/ψ and D^0 candidates. Systematic quoted within brackets are the uncorrelated systematic uncertainties, which have a dependence on the variables under study. Values are in percentages.

Systematics summary	J/ψ	D^0
MC		
MC stat	$[\pm 0.20, \pm 0.60]$	$[\pm 0.18, \pm 0.49]$
MC-truth matched	± 0.53	± 0.43
MC-reweighting	± 0.29	± 0.12
PV		
PV efficiency	± 0.01	± 0.01
PVZ-dependence	± 6.11	± 4.48
Signal determination		
Signal and Bkg modelling	$[\pm 2.36, \pm 9.48]$	$[\pm 0.33, \pm 7.21]$
Ghost contamination	± 0.23	± 0.23
Tracking		
Tracking overall syst.	± 0.70	± 2.83
Tracking stat	$[\pm 7.26, \pm 3.94]$	$[\pm 3.66, \pm 6.59]$
PID		
PID stat	$[\pm 8.36, \pm 9.53]$	$[\pm 2.81, \pm 3.43]$
Total correlated uncertainties	± 6.18	± 5.32
Total correlated uncertainties for ratio	± 6.18	± 5.32
Total uncorrelated uncertainties	$[\pm 10.84, \pm 15.55]$	$[\pm 5.81, \pm 9.87]$

TABLE 5.23: Total uncorrelated systematic uncertainties in nSPDHits bins. Values are in percentages.

Particle	$[0, 200]$	$[200, 300]$	$[300, 446]$	$[446, 715]$	$[715, 960]$	$[960, 1700]$
J/ψ	12.04	12.14	11.59	10.85	11.00	11.03
D^0	7.71	7.56	7.74	6.88	5.81	7.46

TABLE 5.24: Total uncorrelated systematic uncertainties in y bins. Values are in percentages.

Particle	$[2.0, 2.79]$	$[2.79, 3.29]$	$[3.29, 3.79]$	$[3.79, 4.29]$
J/ψ	11.68	10.85	11.45	12.45
D^0	8.49	9.87	7.65	8.28

TABLE 5.25: Total uncorrelated systematic uncertainties in p_{T} bins. The bins are in MeV/c and values are in percentages.

Particle	$[0, 1000]$	$[1000, 2000]$	$[2000, 3000]$	$[3000, 8000]$
J/ψ	12.76	11.52	12.16	15.55
D^0	6.70	8.19	6.82	7.00

TABLE 5.26: Total uncorrelated systematic uncertainties in PVZ bins. Values are in percentages.

Particle	$[-200, -120]$	$[-120, -40]$	$[-40, 40]$	$[40, 120]$	$[120, 200]$
J/ψ	11.97	12.01	11.62	11.49	10.84
D^0	7.07	8.08	7.69	7.51	6.22

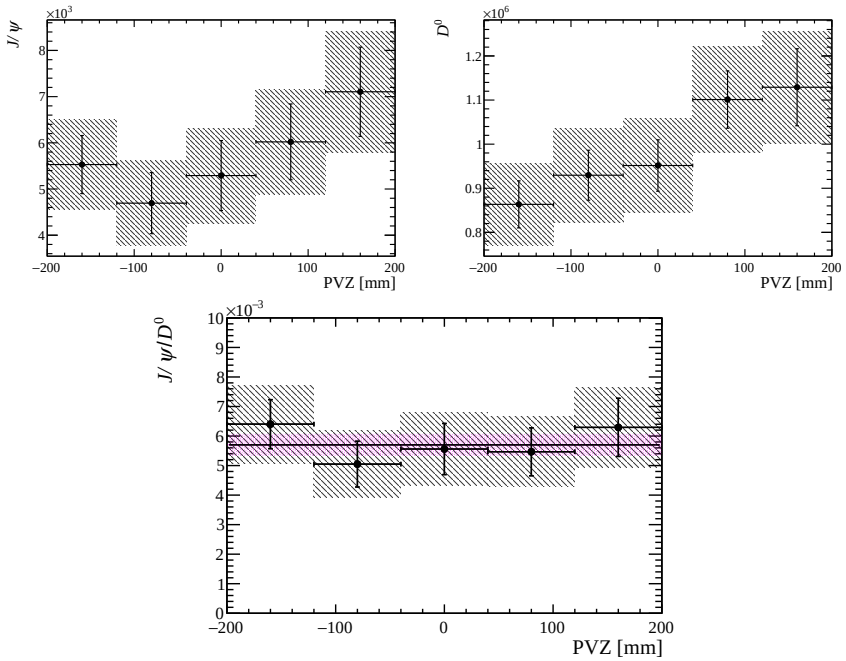

FIGURE 5.39: J/ψ and D^0 corrected yields (*top*) and ratio (*bottom*) as a function of PVZ. The error bars represent the statistical uncertainty while the height of the shaded area represents the quadrature addition of the statistical and systematic uncertainties. The bottom plot is fitted by a constant function and the pink dotted region shows the 1σ confidence region.

5.7 Results

With all the efficiencies at hand, the measured yields can be corrected to obtain the J/ψ and D^0 corrected yields. Since the luminosity measurement for the PbNe dataset is not available yet the J/ψ or D^0 cross-sections cannot be determined. The results presented here will be focused in the J/ψ to D^0 ratio. This can be done in a binned manner according to the bins defined in Sec. 5.3.4, or integrated.

The corrected yields are determined according to Eq. 5.3, where the branching ratios are $\mathcal{B}^{J/\psi} = (5.961 \pm 0.033)\%$ and $\mathcal{B}^{D^0} = (3.93 \pm 0.04)\%$ [10]. The total corrected yields are thus:

- $Y^{J/\psi} = 58\,587 \pm 5132\,(\text{stat})$
- $Y^{D^0} = 9\,194\,469 \pm 340\,943\,(\text{stat})$

The ratio as a function of PVZ, shown in Fig. 5.39, is compatible with a flat distribution, as it would be expected since the vertex position should not influence the J/ψ to D^0 ratio. The corrected yields binned in the variables nSPDHits, y and p_{T}, considering the efficiencies and the branching ratios, are shown in Figs. 5.40, 5.41 and 5.42 respectively as well as the respective J/ψ to D^0 ratios. The error bars represent the statistical uncertainty while the height of the shaded area represents the quadrature addition of the statistical and systematic uncertainties, including both correlated and uncorrelated. Within the current uncertainties there is no conclusive dependence of the ratio in y, but a clear decreasing trend is seen as a function of multiplicity (in the nSPDHits variable) and a strong increase as a function of p_{T}.

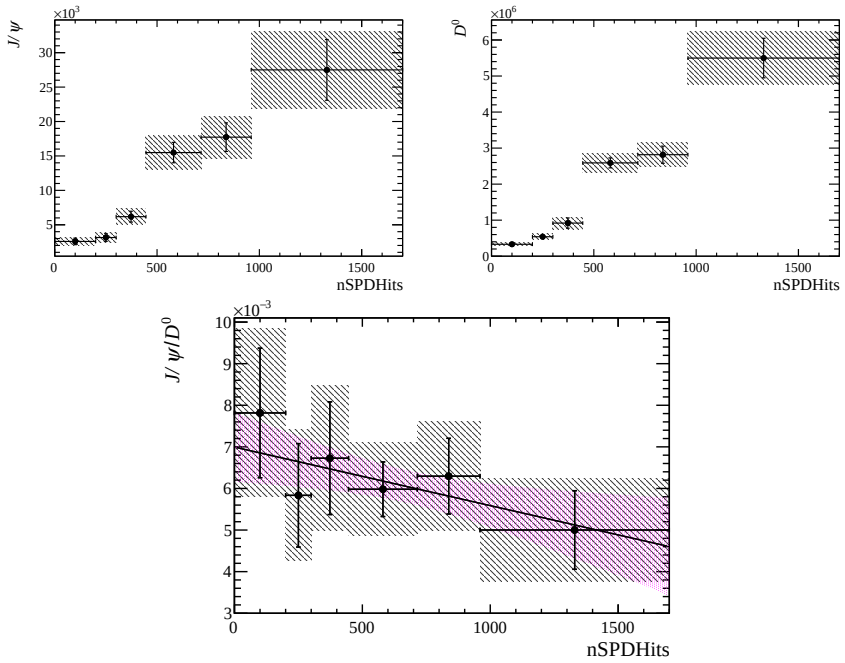

FIGURE 5.40: J/ψ and D^0 corrected yields (*top*) and ratio (*bottom*) as a function of nSPDHits. The error bars represent the statistical uncertainty while the height of the shaded area represents the quadrature addition of the statistical and systematic uncertainties. The bottom plot is fitted by a constant function and the pink dotted region shows the 1σ confidence region.

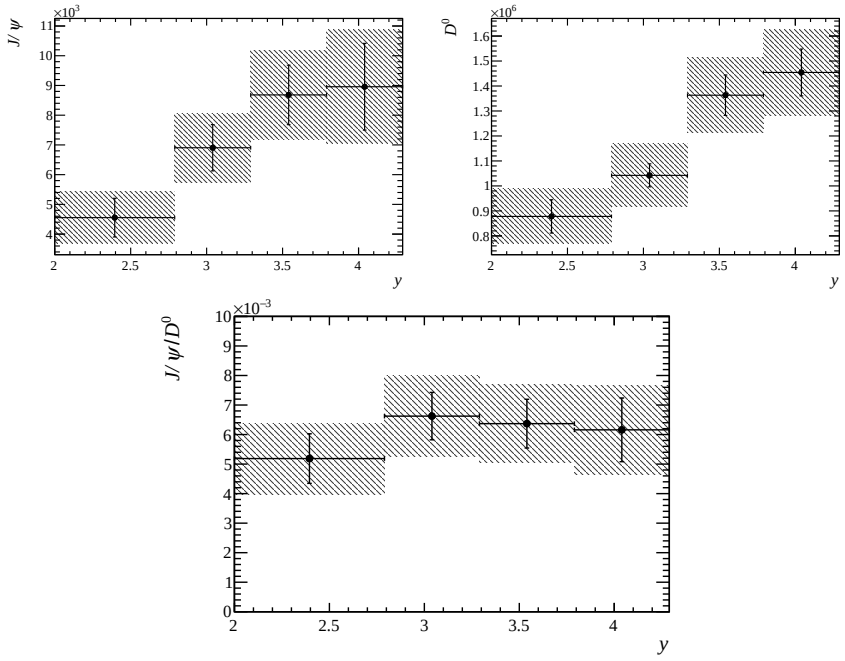

FIGURE 5.41: J/ψ and D^0 corrected yields (*top*) and ratio (*bottom*) as a function of y. The error bars represent the statistical uncertainty while the height of the shaded area represents the quadrature addition of the statistical and systematic uncertainties.

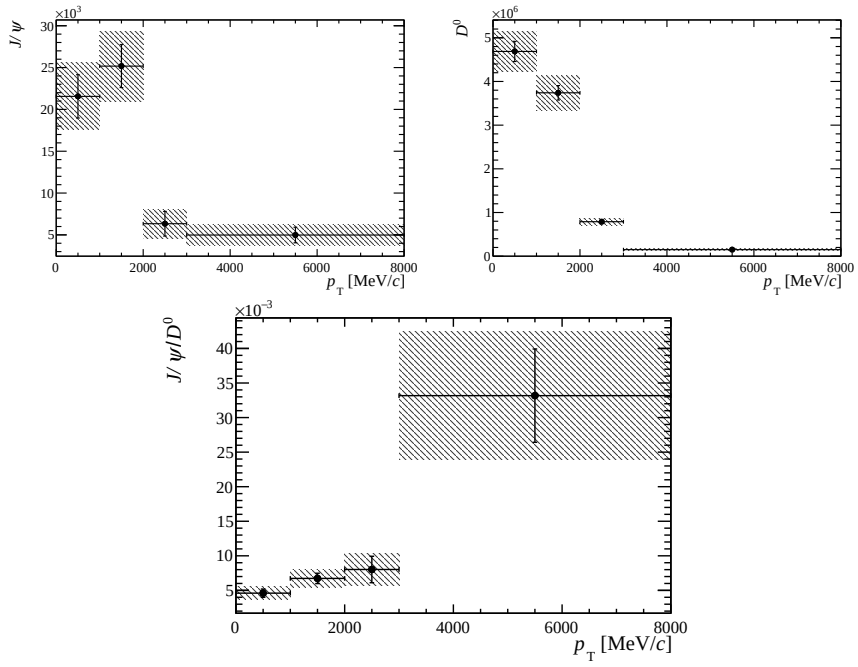

FIGURE 5.42: J/ψ and D^0 corrected yields (*top*) and ratio (*bottom*) as a function of p_{T}. The error bars represent the statistical uncertainty while the height of the shaded area represents the quadrature addition of the statistical and systematic uncertainties.

TABLE 5.27: Mapped centrality quantities to the nSPDHits bins used in the analysis.

nSPDHits	$\langle N_{\mathrm{part}} \rangle$	rms(N_{part})	$\langle N_{\mathrm{coll}} \rangle$	rms(N_{coll})	mean cent. %	rms(cent.)
0-200	4.35	1.02	3.02	0.88	83.68	3.97
200-300	6.70	1.93	5.13	1.81	76.05	5.36
300-446	10.68	2.72	9.09	2.87	66.88	5.21
446-715	17.65	3.80	17.04	4.67	55.89	5.04
715-960	29.05	4.51	32.26	6.51	43.20	4.25
960-1700	53.57	12.52	71.12	20.70	24.53	8.50

5.7.1 Comparison to pNe

Proton-nucleus data provide an important baseline to evaluate the modification of the J/ψ over D^0 ratio in PbNe data. Therefore, the ratio from PbNe data is compared to the one obtained for pNe at the same centre-of-mass energy.[14] In order to be able to evaluate the suppression experienced by the J/ψ as the medium becomes larger, both results need to be expressed as a function of N_{part} and N_{coll}. For pNe collisions, the obtained values for $\langle N_{\mathrm{part}} \rangle$ and $\langle N_{\mathrm{coll}} \rangle$, with the aid of the MC Glauber model, are $\langle N_{\mathrm{part}} \rangle = 2.81$ with an rms$(N_{\mathrm{part}}) = 1.10$, and $\langle N_{\mathrm{coll}} \rangle = 1.81$ with an rms$(N_{\mathrm{coll}}) = 1.10$.

The yields from PbNe have been strategically computed as a function of the nSPDHits. This allows to make a mapping between multiplicity of the events and the centrality information, which is important to study the centrality dependence of the J/ψ to D^0 ratio and to compare the results with pNe. This mapping is done with the aid of the centrality software tool following the same principle from the

[14] The pNe results are preliminary and not yet officially approved by LHCb.

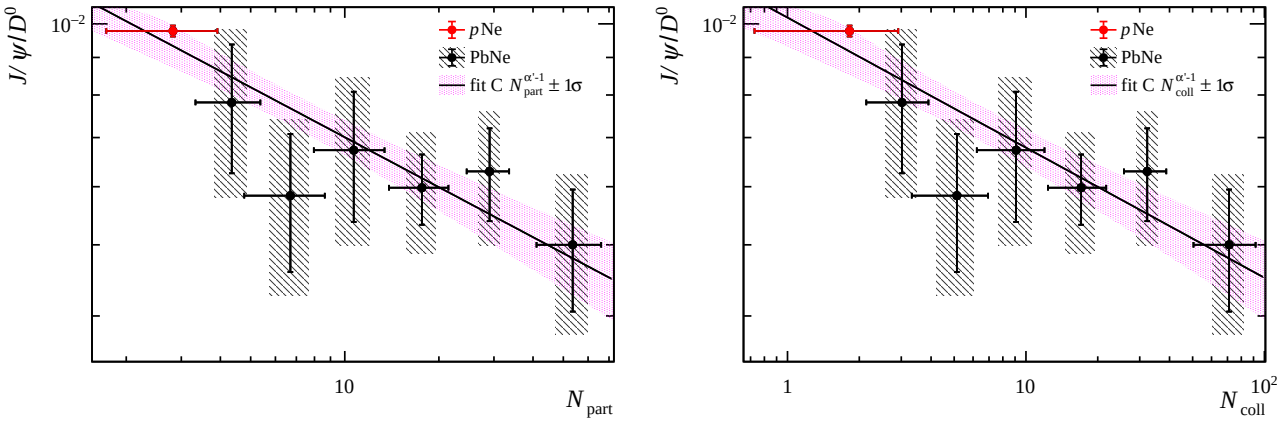

FIGURE 5.43: J/ψ to D^0 ratio as a function of N_{part} (*left*) and N_{coll} (*right*) for pNe and PbNe. The error bars represent the statistical uncertainty while the height of the shaded area represents the quadrature addition of the statistical and systematic uncertainties. The data points on both plots are fitted by a power law function and the pink dotted region shows the 1σ confidence region.

example discussed in Sec. 4.9.1, but with nSPDHits instead of nVeloClusters. The mapping results are summarised in Tab. 5.27 where the mean values of N_{part}, N_{coll} and centrality percentile are shown for each nSPDHits bin, as well as the corresponding RMS values.

With the information from Tab. 5.27, the resulting J/ψ to D^0 ratios as a function of N_{part} and N_{coll} are shown in Fig. 5.43 along with the result obtained in pNe. The ratios from pNe and PbNe as a function of N_{part} have been fitted with a power law function, and the 1σ confidence interval is shown as the pink dotted area. The same has been done in the case of N_{coll}. As before, the error bars represent the statistical uncertainty while the height of the shaded area represents the quadrature addition of the statistical and systematic uncertainties.

The fit model is motivated by the fact that for D^0 the production cross-section is expected to scale as N_{coll}, that is, $\sigma_{\text{AB}}^{D^0} \propto N_{\text{coll}}$ and being an open-charm meson it is not expected to suffer much from the break up caused by the CNM effects or the QGP.[15] On the other hand, the J/ψ production in the presence of any suppression mechanism is expected to be suppressed, that is $\sigma_{\text{AB}}^{J/\psi} \propto N_{\text{coll}}^{\alpha'}$, with $0 < \alpha' < 1$.[16] Hence the ratio between the two would yield $\sigma^{J/\psi}/\sigma^{D^0} \propto N_{\text{coll}}^{\alpha'-1}$. The result of the fit to the N_{coll} variable gives $\alpha' = 0.820 \pm 0.041$, which indeed corresponds to a suppression of J/ψ.

The same result as a function of N_{coll} but for the centrality-integrated system can be seen in Fig. 5.44. In the case of the PbNe data point the error bar represents the statistical uncertainty while the vertical square brackets show the quadrature addition of statistical and systematic uncertainty. The systematic uncertainty has been taken to be the largest found among all the binned measurements. Here, the resulting α' takes the value $\alpha' = 0.854 \pm 0.062$, which is compatible with the non-integrated result.

This model is analogous to a more vastly studied phenomenological model as a function of AB, the product of the mass number of the nuclei A and B, which has been widely used in the comparison of other collision systems. With the same argument as before, it holds that:

[15] This is because naïvely in the event that it breaks, it is highly likely that it will hadronise again with one of the very abundant light quarks available.

[16] Here, AB stands for a general nucleus-nucleus system. In the case at hand it would be PbNe.

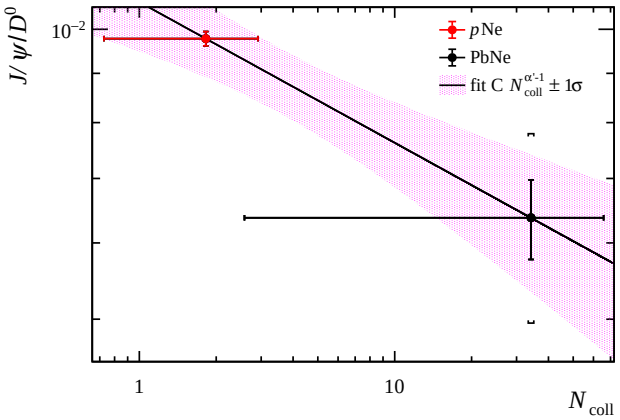

FIGURE 5.44: J/ψ to D^0 ratio as a function of the $\langle N_{\text{coll}} \rangle$ of the system, for pNe and PbNe. The data points on both plots are fitted by a power law function and the pink region shows the 1σ confidence region. The error bar of the PbNe data point represents the statistical uncertainty while the vertical square brackets show the quadrature addition of statistical and systematic uncertainty.

$$\frac{\sigma^{J/\psi}_{\text{AB}}}{\sigma^{D^0}_{\text{AB}}} = \frac{(AB)^{\alpha}\sigma^{J/\psi}_{pp}}{(AB)\sigma^{D^0}_{pp}} = (AB)^{\alpha-1}\frac{\sigma^{J/\psi}_{pp}}{\sigma^{D^0}_{pp}}. \tag{5.14}$$

In the literature, the value found for α under normal suppression conditions of the J/ψ is close to 0.92 [154–157] for $y^* \in [-0.5, 0.5]$. Considering the systems pNe and PbNe, where $AB(p\text{Ne}) = 20$ and $AB(\text{PbNe}) = 4160$, the fit using this model is shown in Fig. 5.45, and α is found to be $\alpha = 0.920 \pm 0.018\,(\text{stat}) \pm 0.030\,(\text{syst})$, which agrees with the value present in the literature. The error bar of the PbNe data point shows the statistical uncertainty while the vertical square brackets show the quadrature addition of the statistical and systematic uncertainty (as was done for Fig. 5.44).

To better put in context the N_{coll} dependence and to evaluate its agreement to the literature value for α, it is useful to relate N_{coll} with AB. It is assumed that $N_{\text{coll}} = c_1(AB)^{\beta}$ with c_1 and β parameters to be fitted. Replacing AB in eq. 5.14 one obtains,

$$\frac{\sigma^{J/\psi}_{\text{AB}}}{\sigma^{D^0}_{\text{AB}}} = (AB)^{\alpha-1}\frac{\sigma^{J/\psi}_{pp}}{\sigma^{D^0}_{pp}} = \left(\frac{N_{\text{coll}}}{c_1}\right)^{\frac{\alpha-1}{\beta}}\frac{\sigma^{J/\psi}_{pp}}{\sigma^{D^0}_{pp}} \propto N_{\text{coll}}^{\frac{\alpha-1}{\beta}}. \tag{5.15}$$

From here it can be deduced that

$$\alpha' - 1 = \frac{\alpha - 1}{\beta} \Rightarrow \alpha' = \frac{\alpha + \beta - 1}{\beta}. \tag{5.16}$$

With the aid of the MC Glauber model the $N_{\text{coll}}(AB)$ dependence can be studied. This is shown in Fig. 5.46. From the fit, the value found is $\beta = 0.687 \pm 0.026$ and results in an expected value for $\alpha' = 0.883 \pm 0.057$. This is in agreement to the values found from the fits in Figs. 5.43 (*right*) and 5.44 where for the integrated case $\alpha' = 0.854 \pm 0.062$ and for the non-integrated case $\alpha' = 0.820 \pm 0.041$. In both cases the difference between the expected and the obtained α' is compatible with 0.

In Fig. 5.47, a binned analysis of α in y^* bins is shown. In the same figure the results from the fixed-target experiments E866/NuSea [154], NA3 [155], NA38 [156]

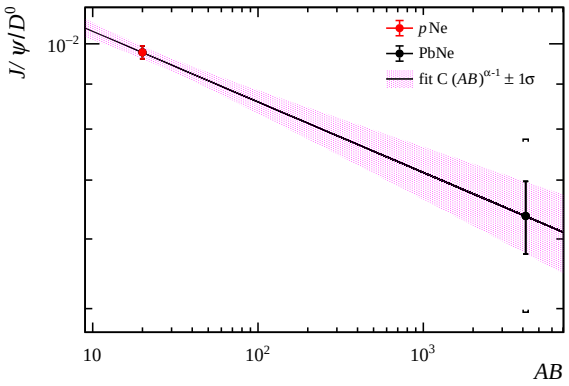

FIGURE 5.45: J/ψ to D^0 ratio as a function of AB, with data from pNe and PbNe collisions. The data points are fitted by a power law function and the pink region shows the 1σ confidence region. The error bar of the PbNe data point represents the statistical uncertainty while the vertical square brackets show the quadrature addition of statistical and systematic uncertainty.

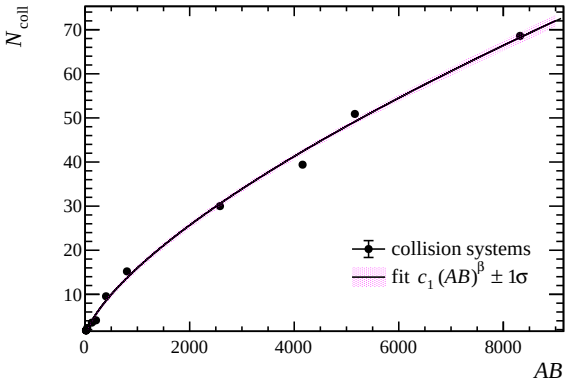

FIGURE 5.46: N_{coll} as a function of AB from MC Glauber model calculations. The points are fitted by a power law function and the pink region shows the 1σ confidence region.

and NA50 [157], covering a complementary y^* range, are shown. The present LHCb results are in very good agreement with these previous results. The E866/NuSea data points are systematically higher than the results of all other experiments, but this is explained by the fact that they have included very small systems in the determination of the value of α (hydrogen and deuterium as targets) [157, 158]. In the figure, the energies quoted in parenthesis for each experiment refer to the beam energy per nucleon.

The results presented here show that there is no evidence of anomalous suppression in PbNe collisions at 69 GeV, since the value found for α is in agreement with the normal suppression scenario. In other words, in the presence of QGP, the J/ψ is expected to be anomalously suppressed, resulting in a lower value for α as was found in PbPb collisions by the NA50 experiment [159], which is not the case here. However, by the present results PbNe becomes the SMOG LHCb baseline for future QGP studies. In addition, among the fixed-target experiments suited for the study of the charmonia suppression, the present result from LHCb complements the y^* space explored up to now.

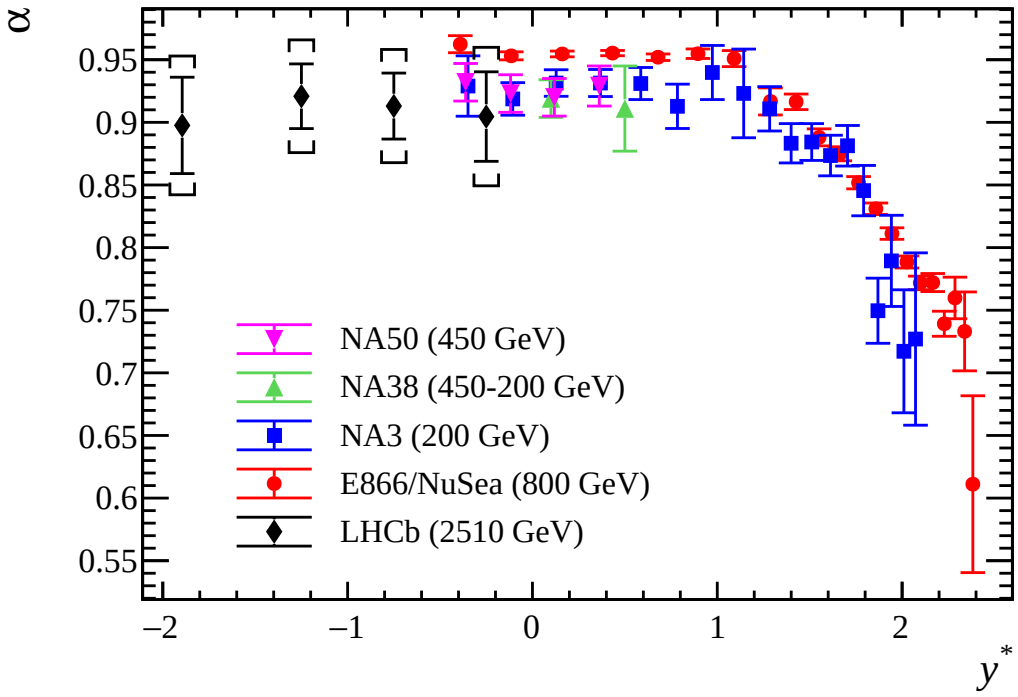

FIGURE 5.47: α as a function of the rapidity in the centre-of-mass frame y^*. The data are from LHCb (diamonds), E866/NuSea (circles), NA3 (squares), NA38 (upward-pointing triangles) and NA50 (downward-pointing triangles). The energy in parenthesis refers to the beam energy per nucleon used in the corresponding experiment. In the case of NA38, the data with the 450 GeV beam results in the lower rapidity point ($y^* = 0.1$) while the data from the 200 GeV beam results in the higher rapidity point ($y^* = 0.5$). The error bars of the LHCb data points represent the statistical uncertainties while the vertical square brackets show the quadrature addition of statistical and systematic uncertainties.

5.8 Discussion and prospects

The results presented above consist in the first Pb-nucleus collisions in fixed-target configuration at the LHC. These go on to show the feasibility of the measurement of the hidden and open charm production, and the capabilities of the LHCb detector to function as a heavy-ion fixed-target experiment.

The J/ψ suppression found in PbNe collisions suggests that there is no QGP formation in this system at this centre-of-mass energy, which opens the door to pursue this measurement in increasingly larger systems such as PbAr and so on. A possibility that can be easily explored with the LHCb detector. The observed suppression is nonetheless in a very good agreement with the CNM effects found by other experiments, quantified by the α parameter, despite the very limited statistics. Furthermore, it validates the J/ψ to D^0 ratio as a relevant measurement of the charmonia suppression.

The presented result in this chapter shows the great potential for the upcoming Run 3 of the LHC, where LHCb will feature the brand new SMOG2 collecting 2 orders of magnitude more data[17] allowing for a significantly more precise measurement.

[17] Assuming the same beam conditions and run time.

SMOG2 will reach higher gas pressures and thus higher instantaneous luminosities will be recorded. In addition to this, SMOG2 could in principle allow for a simultaneous data taking of fixed-target and collider mode collisions given its spatial displacement with respect to the nominal collision point at $Z = 0$, increasing also the total data taking time and with it the data available for analysis. The increased statistics that are expected with SMOG2, place within the reach of LHCb the measurement of ψ' and χ_c in fixed-target collisions. This, together with the possibility to easily change targets, will allow LHCb to thoroughly study the QGP and test the colour-screening as a suppression mechanism of quarkonia.

Chapter 6

Conclusions

The study of QGP has never been an easy task, and with each step taken forward, new effects and mechanisms seem to come into play. Big efforts are being put in place to further probe the QGP and to thoroughly understand the quarkonia suppression. The main challenges present in this endeavour have to do with the interplay of effects, many times competing in the opposite sense: charmonia are an important probe of the QGP and colliders need to produce them in large quantities in order to acquire enough statistics, but as the production increases, more and more likely becomes the statistical recombination scenario, hiding the signatures of the colour screening.

In order to have a complete picture of the suppression mechanisms affecting charmonia, it is paramount to (1) have a solid baseline through the understanding of the CNM effects, (2) be free of the possible recombination effects to have an untampered scenario in which only the dissociative effects are present, and (3) have control over open charm production and all charmonia states — the J/ψ, ψ' and χ_c — in order to test the sequential suppression. The LHCb detector is opening a new door into the future of these studies, providing the right environment and tools needed. Thanks to the SMOG system, a unique feature at the LHC, fixed-target collisions at a centre-of-mass energy of the order of $69 - 100\,\text{GeV}$ between p or Pb beams and a target nucleus are possible. At these energies only $\sim 1\ c\bar{c}$ pair per collision is expected to be produced, ruling out recombination effects and, due to the different available choices of beams and targets, a full study of the phase transition is possible.

LHCb has already shown its capabilities in measuring open and hidden charm in p-nucleus fixed-target collisions [65]. This thesis takes another step, presenting the first measurements in Pb-nucleus fixed-target collisions which, by moving onto a larger system, is a fundamental piece in the understanding of CNM effects. The road to these results began at the heavy-ion data-taking period at the end of the Run 2 of the LHC. PbNe collisions were recorded from 2.51 GeV Pb beams against Ne atoms acting as the target.

In the study of heavy-ion collisions, where the colliding system forms a relatively large extended medium whose dimensions depend on the impact parameter, it is important to be able to access the centrality information of the collision. With the PbNe data and data from the PbPb collisions recorded during the same period, a procedure to determine the centrality of the collisions was developed and implemented, based on a Glauber MC simulation, where the energy deposit in the ECAL was simulated and fitted to the data. After the fit was performed, the simulated distribution was divided in percentiles, which are delimited by sharp energy cuts obtained by integrating the distribution. These energy cuts allow to classify the data into centrality percentiles and a mapping procedure can be performed to link the geometric quantities from the MC Glauber model to the real data. Sources of systematic uncertainty for this procedure were identified and quantified. The results of this analysis culminated in the creation of a software tool which was incorporated

into the main LHCb framework and is now available for use within the collaboration. The centrality measurement is in agreement with the results obtained by ALICE [132, 133], ATLAS [134] and CMS [135] in the case of PbPb, and the results presented for PbNe correspond to the first measurement of centrality performed on a fixed-target system at the LHC.

In the last chapter, the same PbNe data is analysed to study the J/ψ production with respect to the D^0 production, which acts as an indicator of the total $c\bar{c}$ production. This is studied as a function of the event multiplicity, through the nSPDHits variable, which allows to relate the results to geometric quantities such as the N_{coll} or N_{part} that facilitate a possible comparison to other experiments. This translation into geometric variables is possible thanks to the centrality software tool that was developed in the context of this thesis. The J/ψ to D^0 ratio measured in PbNe data was complemented by the preliminary (not officially approved yet) results from pNe collisions, which made possible the extraction of the phenomenological parameter α, which has been extensively used by other experiments to describe the global suppression by CNM effects in the charmonia production. In this context, the present result of LHCb is in remarkable agreement with the results present in the literature [154–157] which suggests that in PbNe at $\sqrt{s_{\text{NN}}} = 69\,\text{GeV}$ there is no sign of abnormal suppression and thus, no presence of QGP. The limitations of this result are of statistical nature.

The findings presented here open a new era in the study of QGP, and motivate the pursuit of this measurement in increasingly larger collision systems. This is a possibility that the LHCb detector can bring to fruition from Run 3 and beyond, where LHCb will feature the brand new SMOG2 collecting at least 2 orders of magnitude more data in its fixed-target configuration. The versatility of the possible nuclear targets and the increased statistics that are expected to be recorded with SMOG2, place within the reach of LHCb the measurement of ψ' and χ_c in fixed-target collisions, which will allow LHCb to thoroughly study the QGP and test the colour-screening as a suppression mechanism of quarkonia.

Appendix A

Invariant mass fits

A.1 Invariant mass fits in p_{T} bins

The invariant mass fits in the bins of p_{T} can be seen in Fig. A.1 for J/ψ and D^0. The sum of candidates from all fits results in 593 ± 41 and 6640 ± 186 for J/ψ and D^0 respectively, which is compatible with the total number of candidates from the integrated fit.

A.2 Invariant mass fits in PVZ bins

The invariant mass fits in the bins of PVZ can be seen in Fig. A.2 for J/ψ and in Fig. A.3 for D^0. The sum of candidates from all fits results in 598 ± 36 and 6645 ± 186 for J/ψ and D^0 respectively, which is compatible with the total number of candidates from the integrated fit.

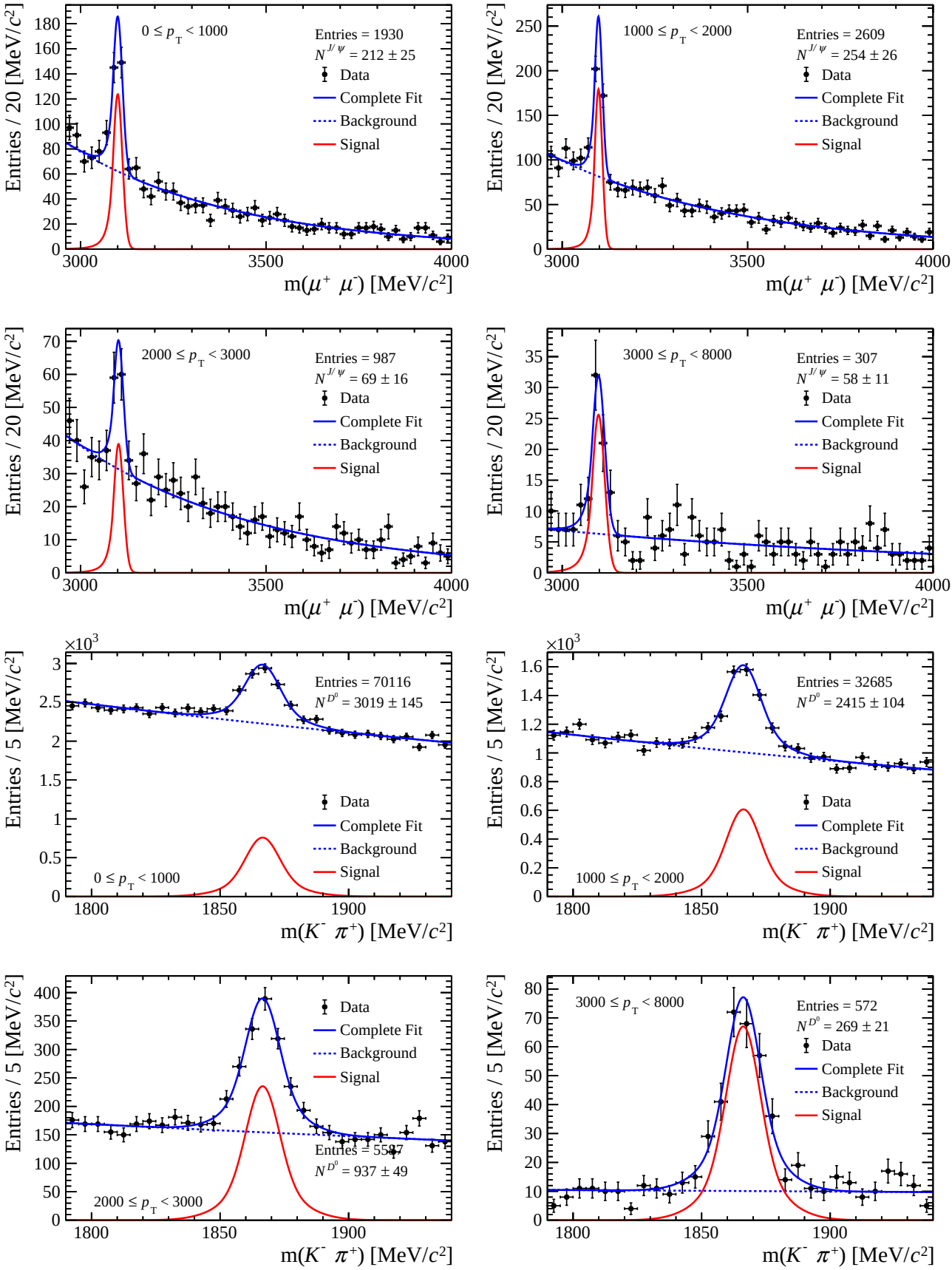

FIGURE A.1: Invariant mass fit for $J/\psi \to \mu^+\mu^-$ (*top* 4) and $D^0 \to K^-\pi^+$ (*bottom* 4) in different p_T bins.

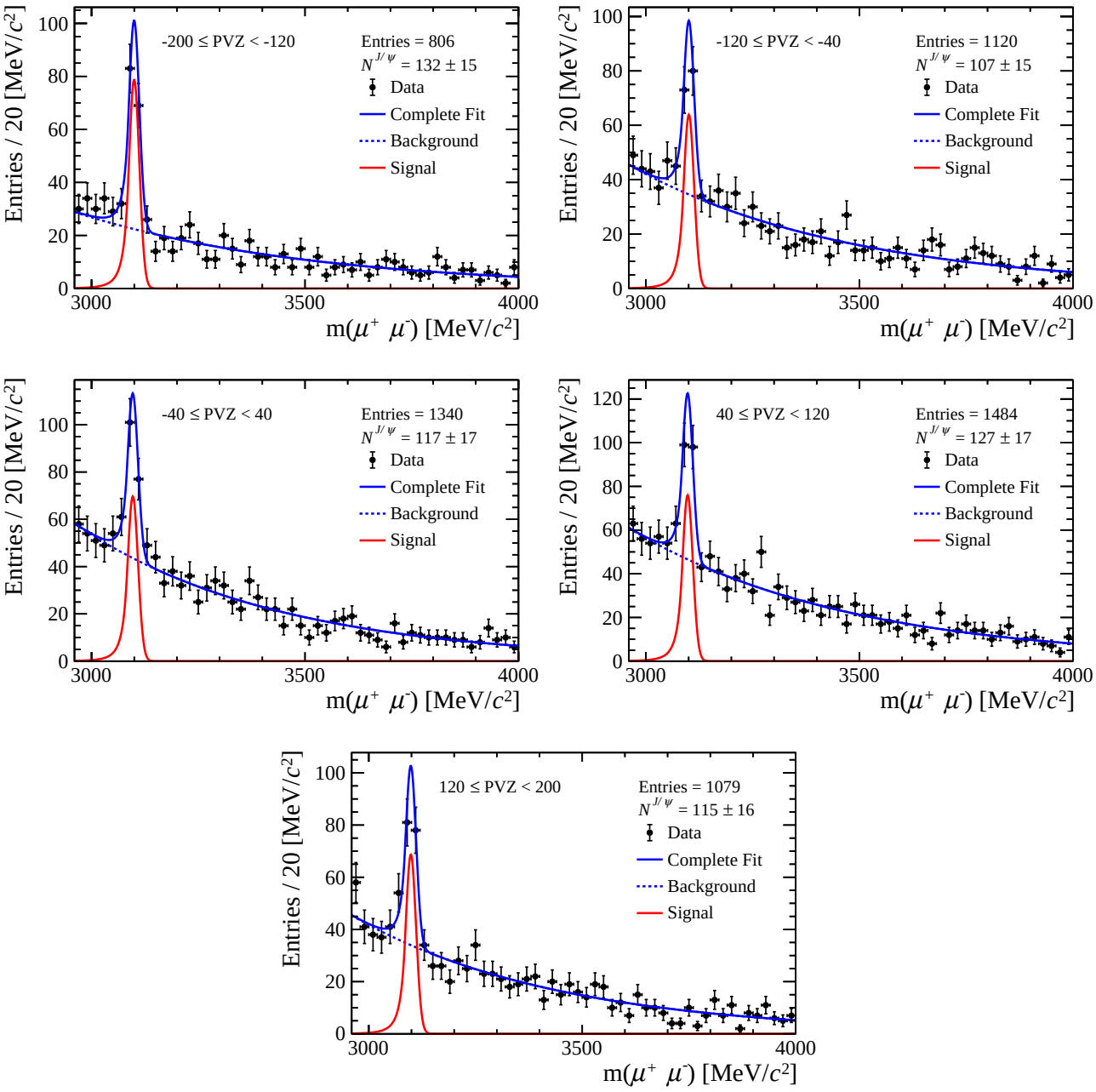

FIGURE A.2: Invariant mass fit for $J/\psi \to \mu^+\mu^-$ in different PVZ bins.

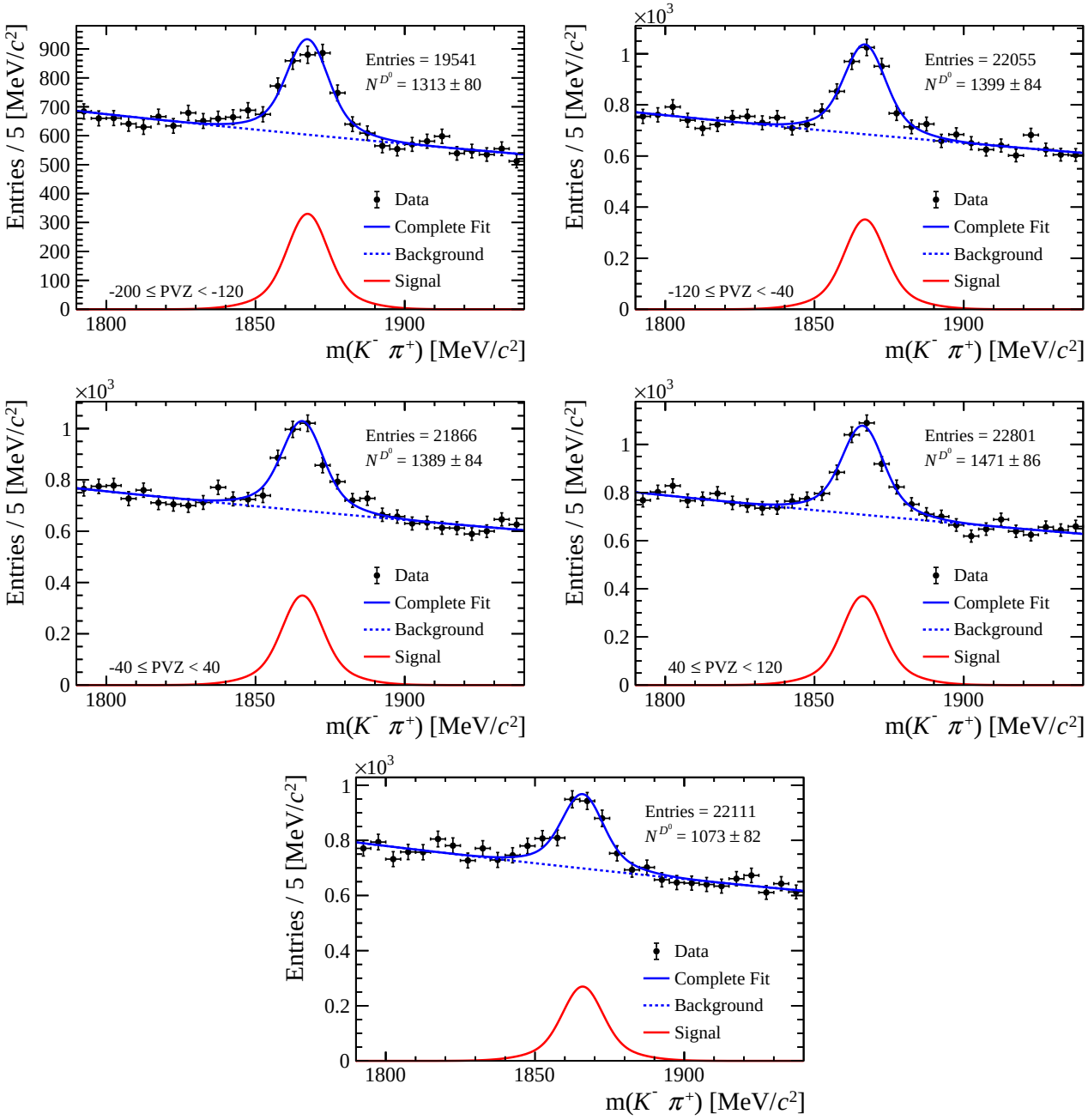

FIGURE A.3: Invariant mass fit for $D^0 \to K^-\pi^+$ in different PVZ bins.

Appendix B

Plots for systematic uncertainties

B.1 Rapidity distributions with different weight sets

The systematic uncertainty assigned to account for the reweighting of the MC was computed from the standard deviation obtained for each rapidity bin after weighting the MC samples with 10 different sets of weights. The resulting reported uncertainty corresponds to the mean value of the standard deviation of all bins, including the MC reconstructed and truth samples. This is done separately for the J/ψ and the D^0.

In Fig. B.1, the rapidity distributions with all 10 sets of weights are shown for all MC samples (see the caption for details).

B.2 Signal extraction using different models

In order to evaluate the systematic uncertainty associated to the choice of a specific model, the signal extraction was performed using different models for the background and for the mass peak. Fig. B.2 shows the yields extracted with the different models for the J/ψ and D^0, in all variables. The RMS of the differences of each model with respect to the nominal one is assigned as a systematic uncertainty.

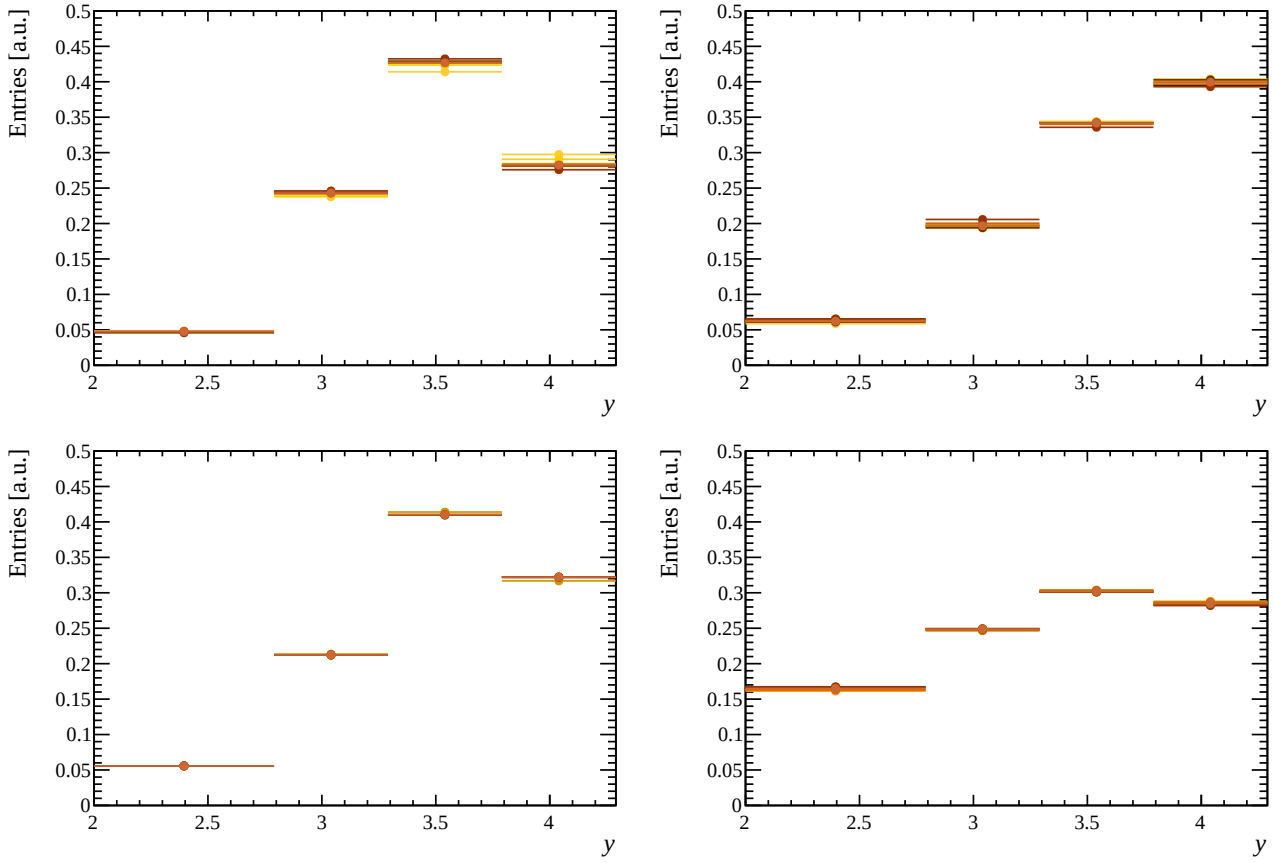

FIGURE B.1: y distribution with 10 different sets of weights for the J/ψ reconstructed MC (*top left*), J/ψ truth MC (*top right*), D^0 (*bottom left*) and D^0 truth MC (*bottom right*).

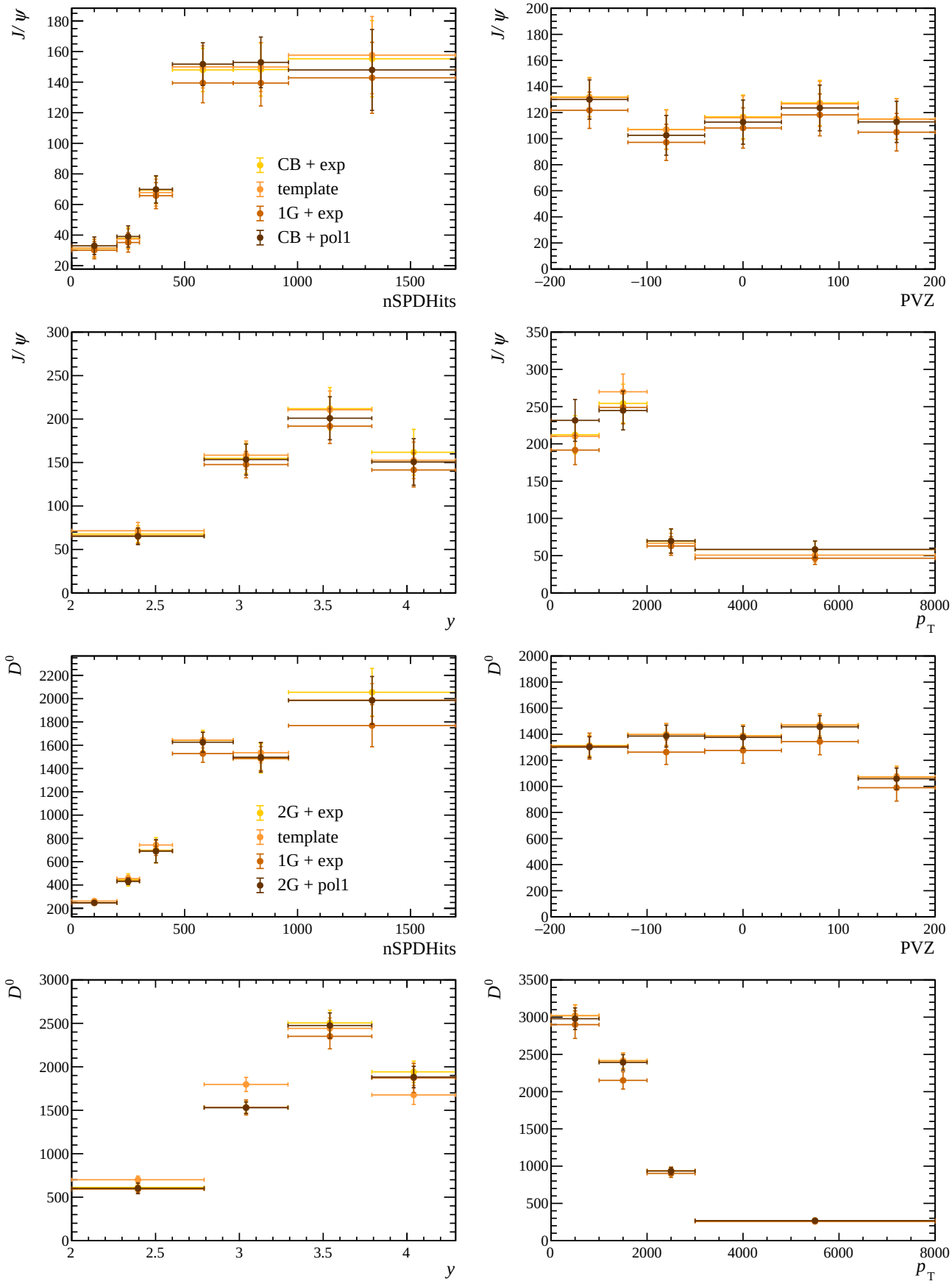

FIGURE B.2: Extracted yields obtained using different models for the mass peak and background description. The top four plots correspond to the J/ψ signal extraction and the bottom 4 to the D^0.

Bibliography

[1] S. Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 30–61. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2012.08.021. URL: https://www.sciencedirect.com/science/article/pii/S0370269312008581.

[2] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 1–29. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2012.08.020. URL: https://www.sciencedirect.com/science/article/pii/S037026931200857X.

[3] LHCb collaboration. "Observation of structure in the J/ψ-pair mass spectrum". In: *Science Bulletin* 65.23 (2020), pp. 1983–1993. ISSN: 2095-9273. DOI: https://doi.org/10.1016/j.scib.2020.08.032. URL: https://www.sciencedirect.com/science/article/pii/S2095927320305685.

[4] R. Aaij et al. "Observation of New Resonances Decaying to $J/\psi K^+$ and $J/\psi\phi$". In: *Physical Review Letters* 127.8 (Aug. 2021). ISSN: 1079-7114. DOI: 10.1103/physrevlett.127.082001. URL: http://dx.doi.org/10.1103/PhysRevLett.127.082001.

[5] LHCb Collaboration. "Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays". In: *Phys. Rev. Lett.* 115 (7 Aug. 2015), p. 072001. DOI: 10.1103/PhysRevLett.115.072001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.115.072001.

[6] R. Aaij et al. "Observation of a Narrow Pentaquark State, $P_c(4312)^+$, and of the Two-Peak Structure of the $P_c(4450)^+$". In: *Physical Review Letters* 122.22 (June 2019). ISSN: 1079-7114. DOI: 10.1103/physrevlett.122.222001. URL: http://dx.doi.org/10.1103/PhysRevLett.122.222001.

[7] Michael E. Peskin et al. *An Introduction To Quantum Field Theory*. Addison-Wesley Publishing Company, 1995.

[8] Walter Greiner et al. *Quantum Chromodynamics*. Springer, 2007.

[9] S. Bethke. "Experimental tests of asymptotic freedom". In: *Progress in Particle and Nuclear Physics* 58.2 (Apr. 2007), pp. 351–386. ISSN: 0146-6410. DOI: 10.1016/j.ppnp.2006.06.001. URL: http://dx.doi.org/10.1016/j.ppnp.2006.06.001.

[10] Particle Data Group. "Review of Particle Physics". In: *PTEP* 2020.8 (2020), p. 083C01. DOI: 10.1093/ptep/ptaa104.

[11] B. Andersson et al. "Parton fragmentation and string dynamics". In: *Physics Reports* 97.2 (1983), pp. 31–145. ISSN: 0370-1573. DOI: https://doi.org/10.1016/0370-1573(83)90080-7. URL: https://www.sciencedirect.com/science/article/pii/0370157383900807.

[12] Rainer Sommer. "Non-perturbative Heavy Quark Effective Theory: Introduction and Status". In: *Nuclear and Particle Physics Proceedings* 261-262 (2015). Advances in Computational Particle Physics, pp. 338–367. ISSN: 2405-6014. DOI: https://doi.org/10.1016/j.nuclphysbps.2015.03.022. URL: https://www.sciencedirect.com/science/article/pii/S2405601415002722.

[13] Szabolcs Borsányi et al. "SU(2) chiral perturbation theory low-energy constants from 2+1 flavor staggered lattice simulations". In: *Phys. Rev. D* 88 (1 July 2013), p. 014513. DOI: 10.1103/PhysRevD.88.014513. URL: https://link.aps.org/doi/10.1103/PhysRevD.88.014513.

[14] Stephan Narison. "MS-bar renormalization scheme". In: *QCD as a Theory of Hadrons: From Partons to Confinement*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2004, pp. 84–90. DOI: 10.1017/CBO9780511535000.013.

[15] Geoffrey T. Bodwin et al. "Rigorous QCD analysis of inclusive annihilation and production of heavy quarkonium". In: *Phys. Rev. D* 51 (3 Feb. 1995), pp. 1125–1171. DOI: 10.1103/PhysRevD.51.1125. URL: https://link.aps.org/doi/10.1103/PhysRevD.51.1125.

[16] Partha Pratim Bhaduri et al. "J/ψ production in proton-induced collisions at energies available at the GSI Facility for Antiproton and Ion Research". In: *Physical Review C* 84.5 (Nov. 2011). ISSN: 1089-490X. DOI: 10.1103/physrevc.84.054914. URL: http://dx.doi.org/10.1103/PhysRevC.84.054914.

[17] R. Baier et al. "Hadronic collisions: A quarkonium factory". In: *Zeitschrift für Physik C Particles and Fields* 19.3 (1983), pp. 251–266. DOI: 10.1007/BF01572254. URL: https://doi.org/10.1007/BF01572254.

[18] B. Guberina et al. "Rare decays of the Z0". In: *Nuclear Physics B* 174.2 (1980), pp. 317–334. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(80)90287-4. URL: https://www.sciencedirect.com/science/article/pii/0550321380902874.

[19] Eric Braaten et al. "Color-Octet Fragmentation and the ψ' Surplus at the Fermilab Tevatron". In: *Phys. Rev. Lett.* 74 (17 Apr. 1995), pp. 3327–3330. DOI: 10.1103/PhysRevLett.74.3327. URL: https://link.aps.org/doi/10.1103/PhysRevLett.74.3327.

[20] Peter Cho et al. "Color-octet quarkonia production". In: *Physical Review D* 53.1 (Jan. 1996), pp. 150–162. ISSN: 1089-4918. DOI: 10.1103/physrevd.53.150. URL: http://dx.doi.org/10.1103/PhysRevD.53.150.

[21] Peter Cho et al. "Color-octet quarkonia production. II". In: *Physical Review D* 53.11 (June 1996), pp. 6203–6217. ISSN: 1089-4918. DOI: 10.1103/physrevd.53.6203. URL: http://dx.doi.org/10.1103/PhysRevD.53.6203.

[22] Mathias Butenschoen et al. "J/ψ Polarization at the Tevatron and the LHC: Nonrelativistic-QCD Factorization at the Crossroads". In: *Physical Review Letters* 108.17 (Apr. 2012). ISSN: 1079-7114. DOI: 10.1103/physrevlett.108.172002. URL: http://dx.doi.org/10.1103/PhysRevLett.108.172002.

[23] Z. Conesa del Valle et al. "Quarkonium production in high energy proton-proton and proton-nucleus collisions". In: *Nuclear Physics B - Proceedings Supplements* 214.1 (May 2011), pp. 3–36. ISSN: 0920-5632. DOI: 10.1016/j.nuclphysbps.2011.03.053. URL: http://dx.doi.org/10.1016/j.nuclphysbps.2011.03.053.

[24] J. C. Collins et al. "Superdense Matter: Neutrons or Asymptotically Free Quarks?" In: *Phys. Rev. Lett.* 34 (21 May 1975), pp. 1353–1356. DOI: 10.1103/PhysRevLett.34.1353. URL: https://link.aps.org/doi/10.1103/PhysRevLett.34.1353.

[25] E.V. Shuryak. "Quark-gluon plasma and hadronic production of leptons, photons and psions". In: *Physics Letters B* 78.1 (1978), pp. 150–153. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(78)90370-2. URL: https://www.sciencedirect.com/science/article/pii/0370269378903702.

[26] Claudia Ratti. "Lattice QCD and heavy ion collisions: a review of recent progress". In: *Reports on Progress in Physics* 81.8 (July 2018), p. 084301. ISSN: 1361-6633. DOI: 10.1088/1361-6633/aabb97. URL: http://dx.doi.org/10.1088/1361-6633/aabb97.

[27] F. Karsch. "Lattice QCD at high temperature and density". In: *Lect. Notes Phys.* 583 (2002). Ed. by Willibald Plessas et al., pp. 209–249. DOI: 10.1007/3-540-45792-5_6. arXiv: hep-lat/0106019.

[28] M. A. Stephanov. "QCD phase diagram and the critical point". In: *International Journal of Modern Physics A* 20.19 (2005), pp. 4387–4392. DOI: 10.1142/S0217751X05027965. eprint: https://doi.org/10.1142/S0217751X05027965. URL: https://doi.org/10.1142/S0217751X05027965.

[29] Zoltán Fodor. "Lattice QVD results at finite temperature and density". In: *Nuclear Physics A* 715 (Mar. 2003), pp. 319c–328c. ISSN: 0375-9474. DOI: 10.1016/s0375-9474(02)01442-2. URL: http://dx.doi.org/10.1016/S0375-9474(02)01442-2.

[30] Helmut Satz. "The Quark-Gluon Plasma – A Short Introduction". In: *Nuclear Physics A* 862-863 (July 2011), pp. 4–12. ISSN: 0375-9474. DOI: 10.1016/j.nuclphysa.2011.05.014. URL: http://dx.doi.org/10.1016/j.nuclphysa.2011.05.014.

[31] A. Bazavov et al. "Equation of state and QCD transition at finite temperature". In: *Phys. Rev. D* 80 (1 July 2009), p. 014504. DOI: 10.1103/PhysRevD.80.014504. URL: https://link.aps.org/doi/10.1103/PhysRevD.80.014504.

[32] Yasuyuki Akiba et al. "The Hot QCD White Paper: Exploring the Phases of QCD at RHIC and the LHC". In: (2015). arXiv: 1502.02730 [nucl-ex].

[33] Shusu Shi. "Event anisotropy v_2 at STAR". Presented Mar 2010. 2010. URL: https://www.bnl.gov/userscenter/thesis/past-competitions/2010/Shi/PhDthesis_ShusuShi.pdf.

[34] F. D. Aaron et al. "Combined measurement and QCD analysis of the inclusive e±p scattering cross sections at HERA". In: *Journal of High Energy Physics* 2010.1 (2010), p. 109. DOI: 10.1007/JHEP01(2010)109. URL: https://doi.org/10.1007/JHEP01(2010)109.

[35] Kari J. Eskola. "Global analysis of nuclear PDFs – latest developments". In: *Nuclear Physics A* 910-911 (2013). Hard Probes 2012, pp. 163–170. ISSN: 0375-9474. DOI: https://doi.org/10.1016/j.nuclphysa.2012.12.029. URL: https://www.sciencedirect.com/science/article/pii/S0375947412003521.

[36] Néstor Armesto. "Nuclear shadowing". In: *Journal of Physics G: Nuclear and Particle Physics* 32.11 (Sept. 2006), R367–R393. DOI: 10.1088/0954-3899/32/11/r01. URL: https://doi.org/10.1088/0954-3899/32/11/r01.

[37] A.H. Mueller et al. "Gluon recombination and shadowing at small values of x". In: *Nuclear Physics B* 268.2 (1986), pp. 427–452. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(86)90164-1. URL: https://www.sciencedirect.com/science/article/pii/0550321386901641.

[38] FRANÇOIS GELIS. "COLOR GLASS CONDENSATE AND GLASMA". In: *International Journal of Modern Physics A* 28.01 (2013), p. 1330001. DOI: 10.1142/S0217751X13300019. eprint: https://doi.org/10.1142/S0217751X13300019. URL: https://doi.org/10.1142/S0217751X13300019.

[39] Larry McLerran. "A Brief Introduction to the Color Glass Condensate and the Glasma". In: (2008). arXiv: 0812.4989 [hep-ph].

[40] Thomas Peitzmann et al. "Direct photons from relativistic heavy-ion collisions". In: *Physics Reports* 364.3 (2002), pp. 175–246. ISSN: 0370-1573. DOI: https://doi.org/10.1016/S0370-1573(02)00012-1. URL: https://www.sciencedirect.com/science/article/pii/S0370157302000121.

[41] R Shahoyan et al. "NA60 results on charm and intermediate mass dimuons production in In–In 158 GeV/A collisions". In: *Journal of Physics G: Nuclear and Particle Physics* 34.8 (July 2007), S1029–S1032. DOI: 10.1088/0954-3899/34/8/s149. URL: https://doi.org/10.1088/0954-3899/34/8/s149.

[42] Raimond Snellings. "Elliptic flow: a brief review". In: *New Journal of Physics* 13.5 (May 2011), p. 055008. DOI: 10.1088/1367-2630/13/5/055008. URL: https://doi.org/10.1088/1367-2630/13/5/055008.

[43] Zuzana Moravcova et al. "Generic algorithm for multiparticle cumulants of azimuthal correlations in high energy nucleus collisions". In: *Physical Review C* 103.2 (Feb. 2021). ISSN: 2469-9993. DOI: 10.1103/physrevc.103.024913. URL: http://dx.doi.org/10.1103/PhysRevC.103.024913.

[44] S. Acharya et al. "Higher harmonic non-linear flow modes of charged hadrons in Pb-Pb collisions at $$ \sqrt{s_{\mathrm{NN}}} $$= 5.02 TeV". In: *Journal of High Energy Physics* 2020.5 (2020), p. 85. DOI: 10.1007/JHEP05(2020)085. URL: https://doi.org/10.1007/JHEP05(2020)085.

[45] J. Adam et al. "Enhanced production of multi-strange hadrons in high-multiplicity proton–proton collisions". In: *Nature Physics* 13.6 (2017), pp. 535–539. DOI: 10.1038/nphys4111. URL: https://doi.org/10.1038/nphys4111.

[46] "Multi-strange baryon production at mid-rapidity in Pb–Pb collisions at sNN=2.76 TeV". In: *Physics Letters B* 728 (2014), pp. 216–227. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2013.11.048. URL: https://www.sciencedirect.com/science/article/pii/S0370269313009544.

[47] M. Aaboud et al. "Measurement of jet pT correlations in Pb+Pb and pp collisions at sNN=2.76 TeV with the ATLAS detector". In: *Physics Letters B* 774 (2017), pp. 379–402. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2017.09.078. URL: https://www.sciencedirect.com/science/article/pii/S0370269317307906.

[48] A. Andronic et al. "Heavy-flavour and quarkonium production in the LHC era: from proton–proton to heavy-ion collisions". In: *The European Physical Journal C* 76.3 (2016), p. 107. DOI: 10.1140/epjc/s10052-015-3819-5. URL: https://doi.org/10.1140/epjc/s10052-015-3819-5.

[49] Raphaelle Bailhache. "Heavy-flavour elliptic flow measured in Pb–Pb collisions at sNN=2.76 TeV with ALICE". In: *Nuclear Physics A* 931 (2014). QUARK MATTER 2014, pp. 530–534. ISSN: 0375-9474. DOI: https://doi.org/10.1016/j.nuclphysa.2014.09.001. URL: https://www.sciencedirect.com/science/article/pii/S0375947414003704.

[50] PHENIX Collaboration. "Heavy-quark production in $p + p$ and energy loss and flow of heavy quarks in Au + Au collisions at $\sqrt{s_{NN}} = 200$ GeV". In: *Phys. Rev. C* 84 (4 Oct. 2011), p. 044905. DOI: 10.1103/PhysRevC.84.044905. URL: https://link.aps.org/doi/10.1103/PhysRevC.84.044905.

[51] B. Abelev et al. "Suppression of high transverse momentum D mesons in central Pb-Pb collisions at $ \sqrt{{{s_{\mathrm{NN}}}}}=2.76\;\mathrm{TeV} $". In: *Journal of High Energy Physics* 2012.9 (2012), p. 112. DOI: 10.1007/JHEP09(2012)112. URL: https://doi.org/10.1007/JHEP09(2012)112.

[52] PHENIX Collaboration. "Energy Loss and Flow of Heavy Quarks in Au + Au Collisions at $\sqrt{s_{\mathrm{NN}}} = 200$ GeV". In: *Phys. Rev. Lett.* 98 (17 Apr. 2007), p. 172301. DOI: 10.1103/PhysRevLett.98.172301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.98.172301.

[53] T. Matsui et al. "J/ψ suppression by quark-gluon plasma formation". In: *Physics Letters B* 178.4 (1986), pp. 416–422. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(86)91404-8. URL: https://www.sciencedirect.com/science/article/pii/0370269386914048.

[54] Kari J. Eskola et al. "EPPS16: nuclear parton distributions with LHC data". In: *The European Physical Journal C* 77.3 (2017), p. 163. DOI: 10.1140/epjc/s10052-017-4725-9. URL: https://doi.org/10.1140/epjc/s10052-017-4725-9.

[55] Roland Katz. "A quantum approach to dynamical quarkonia suppression in high energy heavy ion collisions". Theses. Ecole des Mines de Nantes, Dec. 2015. URL: https://tel.archives-ouvertes.fr/tel-01278863.

[56] Jean-Paul Blaizot et al. "High energy pA collisions in the color glass condensate approach I: gluon production and the Cronin effect". In: *Nuclear Physics A* 743.1-3 (Oct. 2004), pp. 13–56. ISSN: 0375-9474. DOI: 10.1016/j.nuclphysa.2004.07.005. URL: http://dx.doi.org/10.1016/j.nuclphysa.2004.07.005.

[57] François Arleo et al. "J/ψ suppression in pA collisions from parton energy loss in cold QCD matter". In: *Physical Review Letters* 109.12 (Sept. 2012). ISSN: 1079-7114. DOI: 10.1103/physrevlett.109.122301. URL: http://dx.doi.org/10.1103/PhysRevLett.109.122301.

[58] François Arleo et al. "Quarkonium suppression in heavy-ion collisions from coherent energy loss in cold nuclear matter". In: *Journal of High Energy Physics* 2014.10 (Oct. 2014). ISSN: 1029-8479. DOI: 10.1007/jhep10(2014)073. URL: http://dx.doi.org/10.1007/JHEP10(2014)073.

[59] PHENIX Collaboration. "Cold Nuclear Matter Effects on J/ψ Yields as a Function of Rapidity and Nuclear Geometry in $d + A$ Collisions at $\sqrt{s_{NN}} = 200$ GeV". In: *Phys. Rev. Lett.* 107 (14 Sept. 2011), p. 142301. DOI: 10.1103/PhysRevLett.107.142301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.107.142301.

[60] Carlos Lourenço et al. "Energy dependence of J/ψ absorption in proton-nucleus collisions". In: *Journal of High Energy Physics* 2009.02 (Feb. 2009), pp. 014–014. ISSN: 1029-8479. DOI: 10.1088/1126-6708/2009/02/014. URL: http://dx.doi.org/10.1088/1126-6708/2009/02/014.

[61] E.G. Ferreiro. "Charmonium dissociation and recombination at LHC: Revisiting comovers". In: *Physics Letters B* 731 (Apr. 2014), pp. 57–63. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2014.02.011. URL: http://dx.doi.org/10.1016/j.physletb.2014.02.011.

[62] E.G. Ferreiro. "Excited charmonium suppression in proton–nucleus collisions as a consequence of comovers". In: *Physics Letters B* 749 (2015), pp. 98–103. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2015.07.066. URL: https://www.sciencedirect.com/science/article/pii/S0370269315005766.

[63] Helmut Satz. "Colour deconfinement and quarkonium binding". In: *Journal of Physics G: Nuclear and Particle Physics* 32.3 (Feb. 2006), R25–R69. ISSN: 1361-6471. DOI: 10.1088/0954-3899/32/3/r01. URL: http://dx.doi.org/10.1088/0954-3899/32/3/R01.

[64] S. Digal et al. "Heavy quark interactions in finite temperature QCD". In: *The European Physical Journal C* 43.1-4 (July 2005), pp. 71–75. ISSN: 1434-6052. DOI: 10.1140/epjc/s2005-02309-7. URL: http://dx.doi.org/10.1140/epjc/s2005-02309-7.

[65] LHCb collaboration. "First Measurement of Charm Production in its Fixed-Target Configuration at the LHC". In: *Phys. Rev. Lett.* 122 (13 Apr. 2019), p. 132002. DOI: 10.1103/PhysRevLett.122.132002. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.132002.

[66] Partha Pratim Bhaduri. "J/ψ Suppression in Nuclear Collisions: a Unique Signal of Quark-Gluon-Plasma (QGP)". In: *AAPPS Bull.* 30.5 (2020), pp. 14–18. DOI: 10.22661/AAPPSBL.2020.30.5.14.

[67] R. L. Thews. "Quarkonium formation in statistical and kinetic models". In: *The European Physical Journal C - Particles and Fields* 43.1 (2005), pp. 97–102. DOI: 10.1140/epjc/s2005-02219-8. URL: https://doi.org/10.1140/epjc/s2005-02219-8.

[68] NA50 Collaboration. "Charmonia Production at the CERN/SPS". In: *J. Phys. G* 32 (May 2005). 4 pages, 5 figures, XXXXth Rencontres de Moriond conference, S381–S390. 4 p. DOI: 10.1088/0954-3899/32/12/S47. URL: https://cds.cern.ch/record/837684.

[69] Roberta Arnaldi et al. "Anomalous J/ψ suppression in In-In collisions at 158 GeV/nucleon". In: *Nuclear Physics A* 774 (2006). QUARK MATTER 2005, pp. 711–714. ISSN: 0375-9474. DOI: https://doi.org/10.1016/j.nuclphysa.2006.06.120. URL: https://www.sciencedirect.com/science/article/pii/S0375947406003204.

[70] Ulrich W. Heinz et al. "Evidence for a new state of matter: An Assessment of the results from the CERN lead beam program". In: (Jan. 2000). arXiv: nucl-th/0002042.

[71] Raphael Granier de Cassagnac. "Heavy flavours and quarkonia". In: *Journal of Physics: Conference Series* 509 (May 2014), p. 012006. DOI: 10.1088/1742-6596/509/1/012006. URL: https://doi.org/10.1088/1742-6596/509/1/012006.

[72] Steffen Georg Weber. "Charmonium production at mid-rapidity in Pb-Pb and p-Pb collisions with ALICE". In: *Journal of Physics: Conference Series* 668 (Jan. 2016), p. 012096. ISSN: 1742-6596. DOI: 10.1088/1742-6596/668/1/012096. URL: http://dx.doi.org/10.1088/1742-6596/668/1/012096.

[73] ALICE Collaboration. "J/ψ Suppression at Forward Rapidity in Pb-Pb Collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV". In: *Phys. Rev. Lett.* 109 (7 Aug. 2012), p. 072301. DOI: 10.1103/PhysRevLett.109.072301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.109.072301.

[74] S. Acharya et al. "Studies of J/ψ production at forward rapidity in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV". In: *Journal of High Energy Physics* 2020.2 (Feb. 2020). ISSN: 1029-8479. DOI: 10.1007/jhep02(2020)041. URL: http://dx.doi.org/10.1007/JHEP02(2020)041.

[75] Xiaojian Du et al. "Sequential regeneration of charmonia in heavy-ion collisions". In: *Nuclear Physics A* 943 (2015), pp. 147–158. ISSN: 0375-9474. DOI: https://doi.org/10.1016/j.nuclphysa.2015.09.006. URL: https://www.sciencedirect.com/science/article/pii/S0375947415002055.

[76] S. Acharya et al. "J/ψ elliptic and triangular flow in Pb-Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV". In: *Journal of High Energy Physics* 2020.10 (2020), p. 141. DOI: 10.1007/JHEP10(2020)141. URL: https://doi.org/10.1007/JHEP10(2020)141.

[77] LHCb Collaboration. "Erratum: First Measurement of Charm Production in Its Fixed-Target Configuration at the LHC [Phys. Rev. Lett. 122, 132002 (2019)]". In: *Phys. Rev. Lett.* 123 (23 Dec. 2019), p. 239901. DOI: 10.1103/PhysRevLett.123.239901. URL: https://link.aps.org/doi/10.1103/PhysRevLett.123.239901.

[78] Oliver Sim Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: 10.5170/CERN-2004-003-V-1. URL: https://cds.cern.ch/record/782076.

[79] L Rossi. "Superconductivity: Its Role, Its Success and Its Setbacks in the Large Hadron Collider of CERN". In: *Supercond. Sci. Technol.* 23.CERN-ATS-2010-006 (Jan. 2010), 034001. 27 p. DOI: 10.1088/0953-2048/23/3/034001. URL: http://cds.cern.ch/record/1235168.

[80] ATLAS collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08003–S08003. DOI: 10.1088/1748-0221/3/08/s08003. URL: https://doi.org/10.1088/1748-0221/3/08/s08003.

[81] ALICE collaboration. "The ALICE experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08002–S08002. DOI: 10.1088/1748-0221/3/08/s08002. URL: https://doi.org/10.1088/1748-0221/3/08/s08002.

[82] CMS collaboration. "The CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08004–S08004. DOI: 10.1088/1748-0221/3/08/s08004. URL: https://doi.org/10.1088/1748-0221/3/08/s08004.

[83] LHCb collaboration. "The LHCb detector at the LHC". In: *JINST* 3.LHCb-DP-2008-001 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[84] Esma Mobs. "The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019". In: (July 2019). General Photo. URL: https://cds.cern.ch/record/2684277.

[85] LHCb collaboration. "Measurement of $\sigma(pp \to b\bar{b}X)$ at $\sqrt{s} = 7\,\text{TeV}$ in the forward region". In: *Physics Letters B* 694.3 (2010), pp. 209–216. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2010.10.010. URL: https://www.sciencedirect.com/science/article/pii/S0370269310012074.

[86] Rolf Lindner. "LHCb layout 2". LHCb Collection. Feb. 2008. URL: http://cds.cern.ch/record/1087860.

[87] LHCb collaboration. *$\bar{b}b$ production angle plots*. URL: https://lhcb.web.cern.ch/speakersbureau/html/bb_ProductionAngles.html.

[88] *LHCb VELO (VErtex LOcator): Technical Design Report*. Geneva, 2001.

[89] LHCb collaboration. "Performance of the LHCb Vertex Locator". In: *Journal of Instrumentation* 9.09 (Sept. 2014), P09007–P09007. DOI: 10.1088/1748-0221/9/09/p09007. URL: https://doi.org/10.1088/1748-0221/9/09/p09007.

[90] *LHCb inner tracker: Technical Design Report*. Geneva, 2002.

[91] *LHCb outer tracker: Technical Design Report*. Geneva, 2001.

[92] *LHCb RICH: Technical Design Report*. Geneva, 2000.

[93] LHCb collaboration. "LHCb detector performance". In: *Int. J. Mod. Phys.* A30 (2015), p. 1530022. DOI: 10.1142/S0217751X15300227. arXiv: 1412.6352 [hep-ex].

[94] *LHCb calorimeters: Technical Design Report*. Geneva, 2000.

[95] *LHCb muon system: Technical Design Report*. Geneva, 2001.

[96] A A Alves Jr. et al. "Performance of the LHCb muon system". In: *JINST* 8 (2013), P02022. DOI: 10.1088/1748-0221/8/02/P02022. arXiv: 1211.1346 [physics.ins-det].

[97] LHCb collaboration. "Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC". In: *Journal of Instrumentation* 14.04 (Apr. 2019), P04013–P04013. DOI: 10.1088/1748-0221/14/04/p04013. URL: https://doi.org/10.1088/1748-0221/14/04/p04013.

[98] *Trigger Schemes*. URL: http://lhcb.web.cern.ch/lhcb/speakersbureau/html/TriggerScheme.html.

[99] LHCb collaboration. "Absolute luminosity measurements with the LHCb detector at the LHC". In: *Journal of Instrumentation* 7.01 (Jan. 2012), P01010–P01010. DOI: 10.1088/1748-0221/7/01/p01010. URL: https://doi.org/10.1088/1748-0221/7/01/p01010.

[100] Massimiliano Ferro-Luzzi. "Proposal for an absolute luminosity determination in colliding beam experiments using vertex detection of beam–gas interactions". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 553.3 (2005), pp. 388–399. ISSN: 0168-9002. DOI: https://doi.org/10.1016/j.nima.2005.07.010. URL: https://www.sciencedirect.com/science/article/pii/S0168900205014130.

[101] "Precision luminosity measurements at LHCb". In: *Journal of Instrumentation* 9.12 (Dec. 2014), P12005–P12005. DOI: 10.1088/1748-0221/9/12/p12005. URL: https://doi.org/10.1088/1748-0221/9/12/p12005.

[102] Colin Barschel. "Precision luminosity measurement at LHCb with beam-gas imaging". Presented 05 Mar 2014. 2014. URL: https://cds.cern.ch/record/1693671.

[103] LHCb Collaboration. "Measurement of Antiproton Production in $p-$He Collisions at $\sqrt{s_{NN}} = 110$ GeV". In: *Phys. Rev. Lett.* 121 (22 Nov. 2018), p. 222001. DOI: 10.1103/PhysRevLett.121.222001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.222001.

[104] *LHCb Trigger and Online Technical Design Report*. Geneva, 2014.

[105] *LHCb VELO Upgrade Technical Design Report*. Geneva, 2013.

[106] *LHCb Tracker Upgrade Technical Design Report*. Geneva, 2014.

[107] *LHCb PID Upgrade Technical Design Report*. Geneva, 2013.

[108] LHCb Collaboration. *LHCb SMOG Upgrade*. Tech. rep. CERN-LHCC-2019-005. LHCB-TDR-020. Geneva: CERN, May 2019. URL: https://cds.cern.ch/record/2673690.

[109] Albert Bursche et al. *Physics opportunities with the fixed-target program of the LHCb experiment using an unpolarized gas target*. Tech. rep. LHCb-PUB-2018-015. CERN-LHCb-PUB-2018-015. Geneva: CERN, Dec. 2018. URL: https://cds.cern.ch/record/2649878.

[110] A Beuret et al. "The LHC Lead Injector Chain". In: LHC-Project-Report-776. CERN-LHC-Project-Report-776 (Aug. 2004). revised version submitted on 2004-09-23 14:33:06, 4 p. URL: https://cds.cern.ch/record/792709.

[111] Django Manglunki et al. "Ions for LHC: status of the injector chain". In: (Jan. 2007).

[112] Hannes Bartosik et al. "The LHC Injectors Upgrade (LIU) Project at CERN: Ion Injector Chain". In: CERN-ACC-2017-206 (2017), TUPVA020. 4 p. DOI: 10.18429/JACoW-IPAC2017-TUPVA020. URL: https://cds.cern.ch/record/2289479.

[113] J.M. Jowett et al. "The 2018 Heavy-Ion Run of the LHC". In: *Proc. 10th International Particle Accelerator Conference (IPAC'19), Melbourne, Australia, 19-24 May 2019* (Melbourne, Australia). International Particle Accelerator Conference 10. https://doi.org/10.18429/JACoW-IPAC2019-WEYYPLM2. Geneva, Switzerland: JACoW Publishing, June 2019, pp. 2258–2261. ISBN: 978-3-95450-208-0. DOI: doi:10.18429/JACoW-IPAC2019-WEYYPLM2. URL: http://jacow.org/ipac2019/papers/weyyplm2.pdf.

[114] Vladislav Balagura. "Van der Meer scan luminosity measurement and beam–beam correction". In: *The European Physical Journal C* 81.1 (2021), p. 26. DOI: 10.1140/epjc/s10052-021-08837-y. URL: https://doi.org/10.1140/epjc/s10052-021-08837-y.

[115] Michael L. Miller et al. "Glauber modeling in high energy nuclear collisions". In: *Ann. Rev. Nucl. Part. Sci.* 57 (2007), pp. 205–243. DOI: 10.1146/annurev.nucl.57.090506.123020. arXiv: nucl-ex/0701025 [nucl-ex].

[116] *Nuclear Data Services*. https://www-nds.iaea.org/radii/. Last accessed: 2020-11-04.

[117] I. Angeli et al. "Table of experimental nuclear ground state charge radii: An update". In: *Atomic Data and Nuclear Data Tables* 99.1 (2013), pp. 69–95. ISSN: 0092-640X. DOI: https://doi.org/10.1016/j.adt.2011.12.006. URL: http://www.sciencedirect.com/science/article/pii/S0092640X12000265.

[118] C. Loizides et al. "Improved version of the PHOBOS Glauber Monte Carlo". In: *SoftwareX* 1-2 (2015), p. 13. ISSN: 2352-7110. DOI: https://doi.org/10.1016/j.softx.2015.05.001. arXiv: 1408.2549 [nucl-ex]. URL: http://www.sciencedirect.com/science/article/pii/S2352711015000047.

[119] Constantin Loizides et al. "Improved Monte Carlo Glauber predictions at present and future nuclear colliders". In: *Phys. Rev. C* 97 (5 May 2018), p. 054910. DOI: 10.1103/PhysRevC.97.054910. URL: https://link.aps.org/doi/10.1103/PhysRevC.97.054910.

[120] Loyal Durand et al. "QCD and rising cross sections". In: *Phys. Rev. Lett.* 58 (4 Jan. 1987), pp. 303–306. DOI: 10.1103/PhysRevLett.58.303. URL: https://link.aps.org/doi/10.1103/PhysRevLett.58.303.

[121] Wei R. Chen et al. "Particle productivity in pp and pA collisions". In: *Phys. Rev. D* 38 (11 Dec. 1988), pp. 3394–3396. DOI: 10.1103/PhysRevD.38.3394. URL: https://link.aps.org/doi/10.1103/PhysRevD.38.3394.

[122] Wei R. Chen et al. "Geometrical branching model: Phenomenology with jets". In: *Phys. Rev. D* 39 (1 Jan. 1989), pp. 179–186. DOI: 10.1103/PhysRevD.39.179. URL: https://link.aps.org/doi/10.1103/PhysRevD.39.179.

[123] Xin-Nian Wang. "Role of multiple minijets in high-energy hadronic reactions". In: *Phys. Rev. D* 43 (1 Jan. 1991), pp. 104–112. DOI: 10.1103/PhysRevD.43.104. URL: https://link.aps.org/doi/10.1103/PhysRevD.43.104.

[124] Wei-Tian Deng et al. "Hadron production in $p+p$, $p+$Pb, and Pb$+$Pb collisions with the hijing 2.0 model at energies available at the CERN Large Hadron Collider". In: *Phys. Rev. C* 83 (1 Jan. 2011), p. 014915. DOI: 10.1103/PhysRevC.83.014915. URL: https://link.aps.org/doi/10.1103/PhysRevC.83.014915.

[125] J Cugnon et al. "Validity of the Negative Binomial Distribution in Particle Production". In: *Europhysics Letters (EPL)* 4.10 (Nov. 1987), pp. 1127–1131. DOI: 10.1209/0295-5075/4/10/008. URL: https://doi.org/10.1209%2F0295-5075%2F4%2F10%2F008.

[126] Premomoy Ghosh. "Negative binomial multiplicity distribution in proton-proton collisions in limited pseudorapidity intervals at LHC up to $\sqrt{s} = 7$ TeV and the clan model". In: *Phys. Rev.* D85 (2012), p. 054017. DOI: 10.1103/PhysRevD.85.054017. arXiv: 1202.4221 [hep-ph].

[127] ALICE collaboration. "Charged-particle multiplicity measurement in proton-proton collisions at $\sqrt{s} = 0.9$ and 2.36 TeV with ALICE at LHC". In: *Eur. Phys. J.* C68 (2010), pp. 89–108. DOI: 10.1140/epjc/s10052-010-1339-x. arXiv: 1004.3034 [hep-ex].

[128] ALICE collaboration. "Charged-particle multiplicity measurement in proton-proton collisions at $\sqrt{s} = 7$ TeV with ALICE at LHC". In: *Eur. Phys. J.* C68 (2010), pp. 345–354. DOI: 10.1140/epjc/s10052-010-1350-2. arXiv: 1004.3514 [hep-ex].

[129] LHCb collaboration. "Measurement of charged particle multiplicities and densities in pp collisions at $\sqrt{s} = 7$ TeV in the forward region". In: *Eur. Phys. J.* C74.5 (2014), p. 2888. DOI: 10.1140/epjc/s10052-014-2888-1. arXiv: 1402.4430 [hep-ex].

[130] ALICE collaboration. "Centrality determination of Pb-Pb collisions at $\sqrt{s_{\rm NN}} =$ 2.76 TeV with ALICE". In: *Phys. Rev.* C88.4 (2013), p. 044909. DOI: 10.1103/PhysRevC.88.044909. arXiv: 1301.4361 [nucl-ex].

[131] Francesco Bossù et al. *J/ψ studies in PbPb collisions at $\sqrt{s_{NN}} = 5$ TeV with the LHCb experiment*. Tech. rep. LHCb-ANA-2016-067. Geneva: CERN, 2016.

[132] "Centrality Dependence of the Charged-Particle Multiplicity Density at Midrapidity in Pb-Pb Collisions at $\sqrt{s_{NN}} = 5.02$ TeV". In: *Phys. Rev. Lett.* 116 (22 June 2016), p. 222302. DOI: 10.1103/PhysRevLett.116.222302. URL: https://link.aps.org/doi/10.1103/PhysRevLett.116.222302.

[133] ALICE collaboration. *Centrality dependence of the charged-particle multiplicity density at midrapidity in Pb-Pb collisions at $\sqrt{s_{\mathrm{NN}}}$ = 5.02 TeV*. Tech. rep. ALICE-PUBLIC-2015-008. Dec. 2015. URL: https://cds.cern.ch/record/2118084.

[134] "Measurement of the nuclear modification factor for inclusive jets in Pb+Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV with the ATLAS detector". In: *Physics Letters* B790 (2019), pp. 108–128. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2018.10.076. URL: https://www.sciencedirect.com/science/article/pii/S037026931830995X.

[135] "Constraints on the Initial State of Pb-Pb Collisions via Measurements of Z-Boson Yields and Azimuthal Anisotropy at $\sqrt{s_{NN}} = 5.02$ TeV". In: *Phys. Rev. Lett.* 127 (10 Aug. 2021), p. 102002. DOI: 10.1103/PhysRevLett.127.102002. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.102002.

[136] S. Hatakeyama et al. "Nuclear surface diffuseness revealed in nucleon-nucleus diffraction". In: *Phys. Rev. C* 97 (5 May 2018), p. 054607. DOI: 10.1103/PhysRevC.97.054607. URL: https://link.aps.org/doi/10.1103/PhysRevC.97.054607.

[137] Aage Bohr et al. *Nuclear Structure*. World Scientific Publishing Company, 1998. DOI: 10.1142/3530. eprint: https://www.worldscientific.com/doi/pdf/10.1142/3530. URL: https://www.worldscientific.com/doi/abs/10.1142/3530.

[138] W. M. Seif et al. "Systematics of nucleon density distributions and neutron skin of nuclei". In: *International Journal of Modern Physics E* 24.11 (Nov. 2015), p. 1550083. ISSN: 1793-6608. DOI: 10.1142/s0218301315500834. URL: http://dx.doi.org/10.1142/S0218301315500834.

[139] Frederic Fleuret et al. *Global event cuts for 2017 pNe SMOG data*. Tech. rep. LHCb-INT-2020-012. CERN-LHCb-INT-2020-012. Geneva: CERN, June 2020. URL: https://cds.cern.ch/record/2720461.

[140] Samuel Belin et al. "Study of low-p_T J/ψ events in PbPb collisions at $\sqrt{(s_{NN})} =$ 5 TeV with the LHCb experiment". In: (Jan. 2020). URL: https://cds.cern.ch/record/2707982.

[141] Benjamin Audurier et al. "Λ_c^+ over D^0 production in peripheral PbPb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV". In: (July 2020). URL: https://cds.cern.ch/record/2724110.

[142] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(03)01368-8. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[143] J. Allison et al. "Geant4 developments and applications". In: *IEEE Transactions on Nuclear Science* 53.1 (2006), pp. 270–278. DOI: 10.1109/TNS.2006.869826.

[144] I Belyaev et al. *Handling of the generation of primary events in Gauss, the LHCb simulation framework*. Tech. rep. Geneva: CERN, Nov. 2010. URL: https://cds.cern.ch/record/1307917.

[145] T. Pierog et al. "EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider". In: *Phys. Rev. C* 92 (3 Sept. 2015), p. 034906. DOI: 10.1103/PhysRevC.92.034906. URL: https://link.aps.org/doi/10.1103/PhysRevC.92.034906.

[146] David J. Lange. "The EvtGen particle decay simulation package". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 462.1 (2001). BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines, pp. 152–155. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(01)00089-4. URL: https://www.sciencedirect.com/science/article/pii/S0168900201000894.

[147] P. Golonka et al. "PHOTOS Monte Carlo: a precision tool for QED correctionsin Z and W decays". In: *The European Physical Journal C* 45.1 (Jan. 2006), pp. 97–107. ISSN: 1434-6052. DOI: 10.1140/epjc/s2005-02396-4. URL: http://dx.doi.org/10.1140/epjc/s2005-02396-4.

[148] Torbjörn Sjöstrand et al. "A brief introduction to PYTHIA 8.1". In: *Computer Physics Communications* 178.11 (2008), pp. 852–867. ISSN: 0010-4655. DOI: https://doi.org/10.1016/j.cpc.2008.01.036. URL: https://www.sciencedirect.com/science/article/pii/S0010465508000441.

[149] Frederic Fleuret et al. *Global event cuts for 2018 PbNe SMOG data*. Tech. rep. Geneva: CERN, Apr. 2021. URL: https://cds.cern.ch/record/2765569.

[150] R. Partridge et al. "Decay $J\psi \to 3\gamma$ and a Search for the η_c". In: *Phys. Rev. Lett.* 44 (11 Mar. 1980), pp. 712–716. DOI: 10.1103/PhysRevLett.44.712. URL: https://link.aps.org/doi/10.1103/PhysRevLett.44.712.

[151] J. E. Gaiser. "Charmonium Spectroscopy from Radiative Decays of the J/ψ and ψ'". 1982. URL: https://www.slac.stanford.edu/cgi-bin/getdoc/slac-r-255.pdf.

[152] LHCb collaboration. "Measurement of the track reconstruction efficiency at LHCb". In: *Journal of Instrumentation* 10 (2015), P02007–P02007. DOI: 10.1088/1748-0221/10/02/p02007. URL: https://doi.org/10.1088/1748-0221/10/02/p02007.

[153] *TrackCalib for Pbp*. Last accessed: 2021-10-07. URL: https://twiki.cern.ch/twiki/bin/view/LHCbPhysics/IonsFixedTarget%7B%5C#%7DTracking_efficiency_calibration.

[154] FNAL E866/NuSea Collaboration. "Measurement of Differences between J/ψ and ψ' Suppression in $p-A$ Collisions". In: *Phys. Rev. Lett.* 84 (15 Apr. 2000), pp. 3256–3260. DOI: 10.1103/PhysRevLett.84.3256. URL: https://link.aps.org/doi/10.1103/PhysRevLett.84.3256.

[155] J. Badier et al. "Experimental J/ψ hadronic production from 150 to 280 GeV/c". In: *Zeitschrift für Physik C Particles and Fields* 20.2 (1983), pp. 101–116. DOI: 10.1007/BF01573213. URL: https://doi.org/10.1007/BF01573213.

[156] M.C. Abreu et al. "J/ψ and ψ' production in p, O and S induced reactions at SPS energies". In: *Physics Letters B* 466.2 (1999), pp. 408–414. ISSN: 0370-2693. DOI: https://doi.org/10.1016/S0370-2693(99)01108-9. URL: https://www.sciencedirect.com/science/article/pii/S0370269399011089.

[157] The NA50 Collaboration. "Charmonium production and nuclear absorption in p-A interactions at 450 GeV". In: *The European Physical Journal C - Particles and Fields* 33.1 (2004), pp. 31–40. DOI: 10.1140/epjc/s2003-01539-y. URL: https://doi.org/10.1140/epjc/s2003-01539-y.

[158] Ruben Shahoyan. "J/ψ and ψ' production in 450 GeV pA interactions and its dependence on the rapidity and x_F". PhD thesis. Lisboa U., 2001.

[159] M.C. Abreu et al. "Anomalous J/ψ suppression in Pb-Pb interactions at 158 GeV/c per nucleon". In: *Physics Letters B* 410.2 (1997), pp. 337–343. ISSN: 0370-2693. DOI: https://doi.org/10.1016/S0370-2693(97)00915-5. URL: https://www.sciencedirect.com/science/article/pii/S0370269397009155.

Résumé en français

À la fin de l'année 2015, la collaboration LHCb a démarré un programme de physique basé sur des collisions induites par les faisceaux de protons et de plomb du LHC sur une cible fixe (cibles gazeuses). Ce programme permettra un test approfondi, pour la première fois, du mécanisme d'écrantage de couleur prédit par la QCD sur réseau (lors de la production d'un plasma de quarks et gluons dans des collisions d'ions lourds).

Le détecteur LHCb est optimisé pour les mesures de saveurs lourdes. En particulier, il permet des mesures extrêmement précises d'états liés tels que les mésons D, J/ψ, ψ' et χ_c considérés comme des sondes très sensibles pour les études du plasma de quarks et gluons.

Grâce au système LHCb SMOG (System for Measuring the Overlap with Gas), initialement destiné à la mesure de la luminosité, des gaz rares tels que He, Ne, Ar, peuvent être injectés à l'intérieur du détecteur de vertex VELO (Vertex Locator). Agissant comme des « cibles fixes » pour les faisceaux du LHC, ils donnent accès aux collisions proton-noyau et noyau-noyau à une énergie optimale pour étudier la transition de phase de la matière nucléaire normale vers un plasma de quarks et gluons. En 2018, LHCb a enregistré, pour la première fois, des premières collisions PbNe à une énergie dans le centre de masse de 69 GeV ; les résultats de l'analyse de ces données pouvant être directement comparés aux résultats obtenus avec des collisions pNe précédemment enregistrées dans la même configuration et à la même énergie.

Le travail présenté dans cette thèse englobe la totalité du traitement et de l'analyse des données PbNe, depuis la prise de données à leur sélection suivant différents critères de qualité, l'extraction du signal et le calcul des efficacités. De plus, dans cette thèse, le développement d'un outil logiciel pour déterminer l'information sur la centralité des collisions d'ions lourds est présenté. Cet outil a été développé avec des données PbPb et PbNe, et est disponible pour utilisation au sein de la collaboration LHCb. Il a permis de faire une analyse beaucoup plus approfondie des données PbNe, en permettant l'étude du comportement de la production du J/ψ et D^0 non seulement dans un nouveau système, mais aussi dans différents régimes de centralité.

La thèse est organisée de la manière suivante :

Le chapitre 1 situe le contexte physique dans lequel cette étude est menée. Les bases du modèle standard de la physique des particules et de la chromodynamique quantique (QCD) sont présentées conduisant à l'une des prédictions de la QCD : la formation d'un plasma de quarks et de gluons (QGP) lors de collisions d'ions lourds ultra-relativistes. Les principales observables utilisées dans l'étude du QGP sont décrites et un intérêt particulier est accordé à la production de quarkonia en tant que sonde de ce nouvel état de la matière, ainsi qu'aux différents effets possibles pouvant contribuer à sa suppression. Enfin, plusieurs résultats expérimentaux sont présentés en provenance du SPS, du RHIC et du LHC.

Le chapitre 2 présente le CERN et son complexe d'accélérateurs suivi d'une description détaillée de l'expérience LHCb et de ses parties constitutives. Une attention

particulière est accordée au système SMOG (System for Measuring the Overlap with Gas) qui permet à LHCb de fonctionner comme une expérience à cible fixe.

Le chapitre 3 décrit les différents types de modèles de remplissage de la machine LHC pendant les collisions d'ions lourds de 2018, ainsi que leur impact sur les données enregistrées. Les critères de sélection des runs (période de prise de données) et de contrôle de la qualité des données sont présentés.

Le chapitre 4 traite de la détermination de la centralité dans les collisions PbPb et PbNe. La centralité est un paramètre fondamental dans les études de collisions d'ions lourds ultra-relativistes. A partir de la détermination expérimentale de ce paramètre, il est possible d'estimer le nombre de nucléons participants, de collisions binaires et le paramètre d'impact, variables largement utilisées dans l'étude du QGP. Le modèle de Glauber est présenté et utilisé pour déterminer les classes de centralité. La méthodologie pour le faire est expliquée en détail et les incertitudes systématiques associées sont calculées.

Le chapitre 5 présente la mesure de J/ψ et D^0 dans les collisions PbNe à 69 GeV. L'extraction du signal est présentée ainsi que le calcul des efficacités nécessaires pour corriger le signal de particules mesurés. Les incertitudes systématiques associées à ces mesures sont également décrites et calculées en détail. Enfin, le rapport de J/ψ sur D^0 est calculé et une diminution de ce rapport est observée à mesure que le nombre de nucléons participants ou de collisions binaires nucléon-nucléon augmente.

Le chapitre 6 donne une conclusion générale sur ce travail et une perspective sur l'avenir des collisions d'ions lourds à LHCb.

ECOLE DOCTORALE

Titre : Étude de la production des J/ψ et D^0 dans les collisions PbNe à $\sqrt{s_{\rm NN}} = 69$ GeV avec l'expérience LHCb

Mots clés : Quark charme, LHCb, Cible fixe, PbNe

Résumé : À la fin de l'année 2015, la collaboration LHCb a démarré un programme de physique basé sur des collisions induites par les faisceaux de protons et de plomb du LHC sur une cible fixe (cibles gazeuses). Ce programme permettra un test approfondi, pour la première fois, du mécanisme d'écrantage de couleur prédit par la QCD sur réseau (lors de la production d'un plasma de quarks et gluons dans des collisions d'ions lourds).
Le détecteur LHCb est optimisé pour les mesures de saveurs lourdes. En particulier, il permet des mesures extrêmement précises d'états liés tels que les mésons D, J/ψ, ψ' et χ_c considérés comme des sondes très sensibles pour les études du plasma de quarks et gluons.
Grâce au système LHCb SMOG (System for Measuring the Overlap with Gas), initialement destiné à la mesure de la luminosité, des gaz rares tels que He, Ne, Ar, peuvent être injectés à l'intérieur du détecteur de vertex VELO (Vertex Locator). Agissant comme des ≪ cibles fixes ≫ pour les faisceaux du LHC, ils donnent accès aux collisions proton-noyau et noyau-noyau à une énergie optimale pour étudier la transition de phase de la matière nucléaire normale vers un plasma de quarks et gluons. En 2018, LHCb a enregistré, pour la première fois, des premières collisions PbNe à une énergie dans le centre de masse de 69 GeV ; les résultats de l'analyse de ces données pouvant être directement comparés aux résultats obtenus avec des collisions pNe précédemment enregistrées dans la même configuration et à la même énergie.
Le travail présenté dans cette thèse englobe la totalité du traitement et de l'analyse des données PbNe, depuis la prise de données à leur sélection suivant différents critères de qualité, l'extraction du signal et le calcul des efficacités. De plus, dans cette thèse, le développement d'un outil logiciel pour déterminer l'information sur la centralité des collisions d'ions lourds est présenté. Cet outil a été développé avec des données PbPb et PbNe, et est disponible pour utilisation au sein de la collaboration LHCb. Il a permis de faire une analyse beaucoup plus approfondie des données PbNe, en permettant l'étude du comportement de la production du J/ψ et D^0 non seulement dans un nouveau système, mais aussi dans différents régimes de centralité.

Title : Study of J/ψ and D^0 production in $\sqrt{s_{\rm NN}} = 69$ GeV PbNe collisions with the LHCb experiment

Keywords : Charm quark, LHCb, Fixed-target, PbNe

Abstract : At the end of 2015, the LHCb collaboration has recorded the first collisions induced by the LHC proton and lead beams on a fixed target (gaseous targets). This new research programme will allow a thorough test, for the first time, of the colour screening mechanism predicted by lattice QCD (when producing a quark-gluon plasma in heavy-ion collisions).
The LHCb detector is optimised for heavy flavour measurements. In particular, it allows extremely accurate measurements of bound states such as D mesons, J/ψ, ψ' and χ_c considered as very sensitive probes for quark-gluon plasma studies.
Thanks to the LHCb SMOG system (System for Measuring the Overlap with Gas), initially intended for luminosity measurements, noble gases such as He, Ne, Ar, can be injected inside the vertex detector VELO (Vertex Locator). Acting as "fixed targets" for the LHC beams, they give access to proton-nucleus and nucleus-nucleus collisions at an optimal energy to study the phase transition from normal nuclear matter to a quark-gluon plasma. In 2018, LHCb recorded the first PbNe collisions at a centre-of-mass energy of 69 GeV. The results of their analysis can be directly compared to the results obtained with previously recorded pNe collisions at the same centre-of-mass energy.
The work presented in this thesis encompasses the totality of the PbNe data treatment and analysis, from the data-taking, to the data quality determination, the signal extraction and the efficiency computation. In addition, in this thesis the development of a software tool to determine the centrality information of heavy-ion collisions is presented. This tool was developed with PbPb and PbNe data, and is available for usage within the LHCb collaboration. These results allow to make a much more in-depth analysis of the results by studying the behaviour of the J/ψ and D^0 production not only in a new system, but in different centrality regimes within the same PbNe system.

Institut Polytechnique de Paris
91120 Palaiseau, France