



# Faculteit Bètawetenschappen

# Measurement of the centrality dependent nuclear modification factor of the $D^{\ast +}$ meson

# BACHELOR THESIS

Justus van der Velden Natuur- en Sterrenkunde



# Supervisors

Dr. Alessandro Grelli GRASP

 $\begin{array}{c} {\rm Syaefudin \ JAELANI, \ MSc} \\ {\rm GRASP} \end{array}$ 

June 10, 2020

#### Abstract

Quark-gluon plasma (QGP) is thought to have permeated the early universe briefly in the first moments after the Big Bang. Therefore, research into this new state of matter is pivotal to understanding the development of the cosmos. At the LHC, specialised detectors such as ALICE use heavy-ion collisions to simulate the conditions present at the Big Bang and thus produce and study this plasma. Since QGP is strongly interacting, it can be probed effectively by quantifying its influence on heavy-flavour quarks travelling through it. One of the most known observables is the nuclear modification factor ( $R_{AA}$ ), which quantifies the in-medium collisional and radiative energy loss of subatomic particles travelling through the hot and dense plasma. This thesis investigated the relation between  $R_{AA}$  and collision geometry for charm quarks, hypothesising that larger charm suppression would be found for more head-on (central) collisions of ions, indicating generation of larger and hotter QGP for these types of collisions. This was indeed found to be the case. The results were also compared to a theoretical model of  $R_{AA}$  including in-medium energy loss. It was found to be in good agreement with the data, within the relatively large statistical and systematic uncertainties of the measurement.

Cover image provided by CERN, for the benefit of the ALICE Collaboration [1].

# Contents

1	Intr	roduction 1
2	<b>Qua</b> 2.1	<b>rk-gluon plasma</b> 2 Standard Model
		2.1.1 Foundations of the Standard Model
		2.1.2 Particles in the Standard Model
	2.2	Quark matter and quark-gluon plasma
	2.3	QGP in the universe
	2.4	Measurable properties of QGP 5
3	Exp	periment.
-	3.1	Large Hadron Collider
	3.2	ALICE.
	0	3.2.1 ITS 8
		3.2.2 TPC 8
		323 TOF
		3 2 4 V0
	33	Worldwide LHC Computing Grid
	0.0	3.3.1 Baw data processing
		3.3.2 Data sets and Monte Carlo simulations
		3.3.3 AliBOOT & AliPhysics
4	Coll	lider physics 13
	4.1	Heavy ion collisions
		4.1.1 Initial and pre-equilibrium phases 13
		4.1.2 QGP phase
		4.1.3 Hadronisation and freeze-out 14
	4.2	Definitions
		4.2.1 Transverse momentum
		4.2.2 Rapidity and pseudo-rapidity
		4.2.3 Centrality and $\langle N_{part} \rangle$
		4.2.4 Nuclear modification factor $(R_{AA})$ 15
	4.3	$R_{AA}$ versus $\langle N_{part} \rangle$ hypothesis
5	Ana	lysis 17
0	5.1	Method 17
	5.2	$D^*$ decay channel
	5.3	Choosing the centrality and $p_{\rm T}$ ranges 18
	5.0	Data sets
	5.5	Original cut files
	5.6	Mass spectra analysis for unoptimised cuts
	5.0 5.7	Comparison MC
	0.1	$571  \text{Peak width } (\sigma) \qquad \qquad$
		5.7.1 Teak with $(0)$
	58	Significance and $1/S$
	5.0	Beconstruction officiency of propertimisation suits
	5.9	Reconstruction enciency of pre-optimisation cuts $\dots \dots \dots$
	5.10	$T_{AA}$ before optimisation $10.20\%$
	9.11	5 11 1  Durping the entimication
		5.11.1 running the optimisation
	F 10	$\begin{array}{c} 0.11.2  \text{Mass spectra analysis}  \dots  \dots  \dots  \dots  \dots  \dots  \dots  \dots  \dots  $
	5.12	$\begin{array}{c} \text{Optimisation } 20\text{-}30\% \\ \text{f 10.1} \\ \text{Burning the entire instantian} \end{array} $
		5.12.1 Running the optimisation
		$_{0.12.2}$ Mass spectra analysis $\ldots \ldots \ldots$

	5.13 Optimisation 60-80%	47
	5.13.1 Results of attempted optimisation	47
	5.14 Incorporating the optimised results	49
	5.14.1 Reconstruction efficiencies of optimised cuts	49
	5.14.2 Corrected yields and final nuclear modification factors	50
	5.15 Systematic uncertainties	55
	5.15.1 Yield extraction systematic	55
	5.15.2 PID systematic	61
	5.15.3 MC $p_T$ shape systematic	61
	5.15.4 B feed-down systematic	61
	5.15.5 Cut variation systematic	63
	5.15.6 Track reconstruction efficiency systematic	73
	5.15.7 Normalisation and branching ratio systematic	73
	5.15.8 Total systematic uncertainty	74
6	Results and conclusion	75
7	Discussion and outlook	77
	7.1 Discussion and outlook of this analysis	77
		79
	$(.2)$ Broader context of $R_{AA}$ measurements	10
Q	$A$ alwowled memory $R_{AA}$ measurements	70
8	<i>1.2</i> Broader context of $R_{AA}$ measurements	79
8 R	<i>Acknowledgements</i>	79 80
8 R A	Appendix	79 80 I
8 R A	<ul> <li>Acknowledgements</li> </ul>	79 80 I
8 R A	<ul> <li>Acknowledgements</li> <li>Acknowled</li></ul>	78 79 80 I I
8 R A	7.2 Broader context of R <sub>AA</sub> measurements         Acknowledgements         Acknowledgements         References         Appendix         A.1 Run numbers of data sets used         A.2 Cut files before optimisation         A.3 Mass spectra before optimisation	78 79 80 I II IV
8 R A	7.2 Broader context of R <sub>AA</sub> measurements         Acknowledgements         Acknowledgements         References         A.1 Run numbers of data sets used         A.2 Cut files before optimisation         A.3 Mass spectra before optimisation         A.3.1 As a function of momentum	78 79 80 I II IV IV
8 R A	7.2 Broader context of R <sub>AA</sub> measurements         Acknowledgements         Acknowledgements         References         A.1 Run numbers of data sets used         A.2 Cut files before optimisation         A.3 Mass spectra before optimisation         A.3.1 As a function of momentum         A.3.2 As a function of centrality	78 79 80 I II IV IV X
8 R A	7.2 Broader context of R <sub>AA</sub> measurements         Acknowledgements         Acknowledgements         References         A.1 Run numbers of data sets used         A.2 Cut files before optimisation         A.3 Mass spectra before optimisation         A.3.1 As a function of momentum         A.3.2 As a function of centrality         A.4 Reconstruction efficiencies before optimisation	78 79 80 I II IV IV X XII
8 R A	A.2 Broader context of R <sub>AA</sub> measurements         Acknowledgements         Acknowledgements         References         A.1 Run numbers of data sets used         A.2 Cut files before optimisation         A.3 Mass spectra before optimisation         A.3.1 As a function of momentum         A.3.2 As a function of centrality         A.4 Reconstruction efficiencies before optimisation         A.5 Corrected yield and cross section before optimisation	79 80 I II IV IV X XII XV
8 R A	A.2 Broader context of R <sub>AA</sub> measurements         Acknowledgements         Acknowledgements         References         A.1 Run numbers of data sets used         A.2 Cut files before optimisation         A.3 Mass spectra before optimisation         A.3.1 As a function of momentum         A.3.2 As a function of centrality         A.4 Reconstruction efficiencies before optimisation         A.5 Corrected yield and cross section before optimisation	79 80 I II IV IV X XII XV XVIII
8 R A	A.2 Broader context of R <sub>AA</sub> measurements         Acknowledgements         References         Appendix         A.1 Run numbers of data sets used         A.2 Cut files before optimisation         A.3 Mass spectra before optimisation         A.3.1 As a function of momentum         A.3.2 As a function of centrality         A.4 Reconstruction efficiencies before optimisation         A.5 Corrected yield and cross section before optimisation         A.6 Cut values after optimisation	79 <b>80</b> <b>I</b> II IV IV XXII XV XVIII XIX
8 R A	A.2 Broader context of R <sub>AA</sub> measurements         Acknowledgements         Acknowledgements         References         A.1 Run numbers of data sets used         A.2 Cut files before optimisation         A.3 Mass spectra before optimisation         A.3.1 As a function of momentum         A.3.2 As a function of centrality         A.4 Reconstruction efficiencies before optimisation         A.5 Corrected yield and cross section before optimisation         A.6 Cut values after optimisation         A.7 Mass spectra after optimisation         A.8 Reconstruction efficiencies after optimisation	78 79 80 I II IV IV X XII XV XVIII XV XVIII XIX XXI
8 R A	A.2 Broader context of R <sub>AA</sub> measurements         Acknowledgements         Acknowledgements         References         A.1 Run numbers of data sets used         A.2 Cut files before optimisation         A.3 Mass spectra before optimisation         A.3.1 As a function of momentum         A.3.2 As a function of centrality         A.4 Reconstruction efficiencies before optimisation         A.5 Corrected yield and cross section before optimisation         A.6 Cut values after optimisation         A.7 Mass spectra after optimisation         A.8 Reconstruction efficiencies after optimisation         A.7 Mass spectra after optimisation         A.8 Reconstruction efficiencies after optimisation         A.9 Corrected yield and cross section after optimisation	78 79 80 I II IV IV X XII XV XVIII XIX XXI XXII XXII
8 R A	A.2       Broader context of R <sub>AA</sub> measurements         Acknowledgements         References         A.1       Run numbers of data sets used         A.2       Cut files before optimisation         A.3       Mass spectra before optimisation         A.3.1       As a function of momentum         A.3.2       As a function of centrality         A.4       Reconstruction efficiencies before optimisation         A.5       Corrected yield and cross section before optimisation         A.6       Cut values after optimisation         A.7       Mass spectra after optimisation         A.8       Reconstruction efficiencies after optimisation         A.7       Mass spectra after optimisation         A.8       Reconstruction efficiencies after optimisation         A.9       Corrected yield and cross section after optimisation         A.10       Mass spectra from cut variation	79 80 I IV IV X XII XV XVIII XIX XXII XXII X
8 R A	Acknowledgements         Acknowledgements         Acknowledgements         Acknowledgements         Acknowledgements         Acknowledgements         Appendix         A.1 Run numbers of data sets used         A.2 Cut files before optimisation         A.3 Mass spectra before optimisation         A.3.1 As a function of momentum         A.3.2 As a function of centrality         A.3.2 As a function of centrality         A.4 Reconstruction efficiencies before optimisation         A.5 Corrected yield and cross section before optimisation         A.6 Cut values after optimisation         A.7 Mass spectra after optimisation         A.8 Reconstruction efficiencies after optimisation         A.9 Corrected yield and cross section after optimisation         A.10 Mass spectra from cut variation         A.10 Mass spectra from cut variation	79 80 I IV IV X XII XV XVIII XIX XXII XXII X

# **1** Introduction

On the border of Switzerland and France lies the largest and most complex machines ever assembled: the Large Hadron Collider (LHC). It has received international recognition for its discoveries of new elemental particles, for example famously through the discovery of the Higgs boson. Although the LHC itself is technically one single machine, the accelerator complex houses several independent research institutions which run detector installations along the accelerator. One of these institutions is *ALICE*, which is an acronym for A Large Ion Collider Experiment. Although the LHC is capable of colliding several different kinds of particles, ALICE specialises in analysing collisions involving the creation of quark-gluon plasma (QGP) [2]. This is because one of the main goals of ALICE is to further our understanding of this state of matter, which is thought to have permeated the universe just after the Big Bang.

In order to understand this state of matter, we need a basic understanding of the fundamental interactions of nature. These are of course the electromagnetic interaction, gravitational interaction, weak interaction and strong interaction. QGP is a strongly-interacting state of matter in which quarks and gluons are largely deconfined [3](6-10). The theory which describes the properties of the strong interaction is known as *quantum chromodynamics* (QCD). According to QCD, quarks and gluons do not strongly interact at high energies, while at low energies the strong force holds them together very tightly.

One might wonder how such a cosmological question could be resolved by a particle accelerator. The clue lies in the fact that we are able to simulate the circumstances of the universe just after the Big Bang by colliding heavy ions at relativistic velocities. Through heavy ion collisions, such as Pb-Pb, it is possible to study the properties of QGP [3](72-78). The plasma will affect any particles travelling through it generated by the collision itself through for instance energy loss. By measuring these effects carefully and coupling them to theoretical descriptions of the plasma, one can quantify these properties and apply boundary conditions which limits the number of plausible theories, until the most viable and physically relevant remain. Once a proper description of the plasma is found, we will be able to predict its other properties as well, for example its phase behaviour. Naturally, it is not reasonable to assume that one experiment will be able to provide all the data needed to fully quantify its behaviour. However, combined with all the studies conducted globally, the hope is that we will at some point be able to construct a solid understanding of this extreme phase of matter.

As mentioned, the goal of experiments in QGP is to provide boundary conditions for theoretical predictions in order to judge the viability of existing theorems. An existing theorem might predict that a particular observable should behave in a certain way and an experiment could provide the parameters for which that statement could be true. In this thesis, through analyses and considerations of existing ALICE collision data, a relation between the  $R_{AA}$  and the number of involved *participants* (directly involved nuclear partons) in the Pb-Pb collisions will be experimentally determined, giving another insight in the plasma's properties and a restriction on theoretical models that aim to describe it. The research question of this project therefore is: how does the nuclear modification factor for the  $D^{*+}$  meson in Pb-Pb collisions behave as a function of the number of participants?

This manuscript will first discuss the theory behind the experiment, which we touched upon in this introduction, in more detail. This includes a motivation of why QGP is interesting from a theoretical point of view and an introduction to the underlying physics of our measurements. We will also introduce the experimental setup itself, ALICE and the LHC, as well as a brief introduction to the underlying software and data processing infrastructure. Then we introduce a list of new definitions for the physical processes used specifically in this experiment, as well as our hypothesis based on earlier studies. This will be followed up by the analysis conducted for this measurement of the  $R_{AA}$  as a function of involved collision particles, which forms the bulk of the project. Finally, we will present our results and relate these to an existing theoretical prediction on the shape of the nuclear modification factor for different collision geometries.

# 2 Quark-gluon plasma

In the introduction we briefly touched upon the physical and cosmological concepts behind QGP. Here we will elaborate on them within the relevance of this analysis.

#### 2.1 Standard Model

The *Standard Model* of particle physics is a theory which provides a description of our universe at the smallest scales. It covers three of the four fundamental interactions in nature, accurately describing the strong, weak and electromagnetic interactions. We will introduce the Standard Model here, as it forms the central theoretical underpinning of this analysis.

#### 2.1.1 Foundations of the Standard Model

The Standard Model as we know it today has been around for several decades now. The first foundations for the theory were provided in 1954, when Chen-Ning Yang and Robert Mills first published a paper which generalised gauge theory for abelian groups, a fundamental mathematical concept used in quantum field theory, to non-abelian groups [4]. A theory of this kind is now classified as a *Yang-Mills theory*. Yang and Mills essentially attempted to add a description of strong interaction to the already existing description of electromagnetism (described by the gauge theory for abelian groups). In the two decades after this, the theory underwent several iterations which included a description for the electroweak<sup>1</sup> interaction and a theory describing this interaction was succesfully developed by Sheldon Glashow in 1961 [5], Steven Weinberg in 1967 [6] and Abdus Salam in 1968 [7]. The latter two incorporated the *Higgs mechanism*, an important part of the modern Standard Model, in the theory. The *Glashow-Weinberg-Salam model* as it was called was completed by Martinus Veldman and Gerard 't Hooft when they proved the renormalisability of the theory in 1971 [8].

The first experimental confirmation of the Glashow-Salam-Weinberg model took place in 1973, when scientists at CERN used the Gargamelle heavy-liquid bubble chamber to confirm the existence of weak neutral currents. This discovery was in line with predictions made by the Glashow-Weinberg-Salam model and had profound implications for the field of particle physics and studies at CERN in general, as it paved the way for investments in new experiments to study the weak interaction, including the Large Electron Positron collider (LEP) [9]. It even led to Nobel Prizes for the main contributors to the theory: for Glashow, Salam and Weinberg in 1979 and for 't Hooft and Veltman in 1999.

Around the time of the discoveries of weak neutral currents, quantum chromodynamics, which is a Yang-Mills theory based on SU(3), was being developed as an improved theory of the strong interaction. One of the fundamental insights of QCD is the concept of *asymptotic freedom*, which basically means that the coupling of the strong interaction decreases for increasing energies [10], further discussed in the next sections. It was far more difficult to study QCD experimentally than the weak or electromagnetic interactions at the time, since the calculational tools used for the latter at the time were generally not accurate enough for QCD experiments. However, QCD gained recognition through several experiments confirming its predictions. Notably, the asymptotic freedom property was successfully confirmed experimentally [11](56). Together with the description of the electroweak interaction from the Glashow-Weinberg-Salam model, this became the base of the Standard Model as it is used today.

The Standard Model has described many physical processes with great accuracy and predicted the existence of several particles which were experimentally confirmed later, granting the theory wide acceptance in the scientific community. However, it does not provide the elusive 'theory of everything', as it fails to account for some physical phenomena, such as Einstein's general relativity or the existence of dark matter.

 $<sup>^{1}</sup>$ The electroweak interaction is called as such because it incorporates both the electromagnetic and weak interaction; they were shown to be two aspects of the same force.



Figure 1: Elementary particles of the Standard Model [12].

#### 2.1.2 Particles in the Standard Model

The framework described by the Standard Model provides us with a number of elementary particles, currently believed to be the smallest constituents of all matter in the universe. All currently known elementary particles are shown in Figure 1. The Standard Model divides these particles into several categories. Firstly, a division can be made between particles with integer spin (*bosons*) and particles with half-integer spin (*fermions*) [13]. All particles belong in either one of these groups [14](440). Not just elementary particles carry this classification; for example, a proton made of three spin-1/2 particles (which are of course fermions) is also a fermion, but a D meson composed of a spin-1/2 particle and a spin-1/2 antiparticle is a boson. Right now however we are interested in elementary particles in the context of the Standard Model.

The elementary bosons of the Standard Model can be split into two further subdivisions: gauge bosons and scalar bosons. Currently, the Higgs boson is the only known scalar boson. Gauge bosons are the force mediators of the three interactions described by the Standard Model, where photons mediate the electromagnetic interaction, W and Z bosons the weak interaction and gluons the strong interaction.

Elementary fermions can also be divided into categories: quarks and leptons (and their antimatter counterparts). The basic difference between the two is their colour charge. Leptons carry no colour charge and quarks do. A direct result of this property in QCD is that quarks can interact via the strong force, mediated by gluons, while leptons cannot. We currently know of twelve quark flavours: (anti-)up, (anti-)down, (anti-)charm, (anti-)strange, (anti-)top and (anti-)beauty. In addition to interacting via the strong force, quarks also interact electromagnetically and via the weak interaction, as they all carry electric charge and weak isospin<sup>2</sup>. Similarly to quarks, twelve distinct flavours of leptons have been observed, namely electrons, electron neutrinos, muons, muon neutrinos, tau, tau neutrinos and their corresponding antiparticles. Leptons carry electrical charge and weak isospin as well, allowing them to interact via electromagnetic force and the weak force [15].

Force	Relative strength	Range (fm)		
Strong	1	$\sim 1$		
Weak	$\sim 10^{-16}$	$\sim 10^{-3}$		
Electromagnetic	$\sim 10^{-2}$	$\infty$		

Table 1: Properties of fundamental interactions described by the Standard Model [16](3). Note however that asymptotic freedom implies that the relative strength of the strong force decreases at high temperatures.

 $<sup>^{2}</sup>$ Clearly, these particles can also interact gravitationally, since they have mass. However, it was omitted in this discussion, as the hypothetical graviton which might mediate this field is yet to be discovered and not part of the current Standard Model.

As the name suggests, the strong interaction is the strongest interaction compared to the other fundamental forces [17]. The relative strengths of the interactions in the Standard Model are shown in Table 1. Although one might intuitively compare the fundamental forces to classical 'forces', it should be noted that within quantum mechanics the notion of such terms is a bit different. Physicists instead use *coupling constants* to gauge the strength of an interaction (without accounting for range that is). Thus, when we state that the strong interaction is 'strong', we mean that its coupling constant is high compared to the others. The strength of the strong force explains why quarks are bound together very strongly inside *hadrons* (particles composed of quarks). Quarks are never observed as isolated particles in nature. Instead, all properties of quarks we know of where measured indirectly using collision experiments in particle accelerators. Quarks (and gluons) cannot be observed directly because the strong force holds quarks together so tightly via gluons that 'splitting' a hadron under normal circumstances immediately generates new hadrons due to the high energies involved. This property of the strong interaction is known as *colour confinement* [3](5-10).

#### 2.2 Quark matter and quark-gluon plasma

When keeping the Standard Model in mind, it should come as no surprise that in quark-gluon plasma, the strong interaction plays an important role. As it turns out, QGP is in fact a phase of a more general state of matter known as *quark matter*. Quark matter is a term which covers all states of matter in which quarks are no longer bound in particles. We will for now focus on QGP, as this is within the relevance of this analysis. QGP can exist due to the asymptotic freedom described by QCD we discussed earlier. For increasing temperatures, the coupling constant of the strong force  $\alpha_s$  decreases asymptotically [18]. Mathematically this is described by

$$\alpha_s(q^2) = \frac{12\pi}{(11n_C - 2f)\ln(q^2/\Lambda_{QCD}^2)}$$
(2.1)

where q is the momentum transfer vector,  $n_C$  the number of colours,  $\Lambda_{QCD}$  a parameter currently estimated at roughly 200 MeV and f the number of quark flavours [16]. For high momentum transfer, i.e. at sufficiently high temperatures/densities, the strength of the strong force becomes small enough that it can no longer bind quarks together [3]. This results in the breakdown of colour confinement (i.e. *colour deconfinement*) and the addition of quarks and gluons as degrees of freedom in the matter in question.

As one might presume, quark matter can only exist in very extreme circumstances. In Figure 2, we can see a phase diagram of quark matter. Note that matter has to be either very densely packed together or have a very high temperature (or both) in order for QGP to exist. However, QGP has almost certainly been a part of the early universe, as we will discuss in the next section. The existence of ancient QGP warrants the research being done in this field, as a proper understanding of the strong force, quark matter and the Standard Model in general is required to build an accurate picture of the early universe.

#### 2.3 QGP in the universe

As mentioned, QGP occured naturally in the ancient, early universe. As a matter of fact, there was a brief period of time where it was the dominant state of matter. Shortly after the Big Bang, the universe underwent a series of large changes in structure within (very) brief timescales. In particular, the universe is thought to have been permeated by QGP in the so called *quark epoch* at around 10 µs in Hubble time [19](10). The quark epoch is interesting from a cosmological point of view, as it forms the bridge from the preceding electroweak epoch (when the universal temperature was low enough to allow electromagnetic and weak interactions to exist as separate forces) to primordial nucleosynthesis (the first formation of hadronic gas in the universe). During the quark epoch, the temperature of the universe had dropped enough for the strong interaction to exist as a separate force. However, the temperature was still too high for colour confinement to occur. Theoretical predictions from thermodynamics and quantum chromodynamics as well as previous collider experiments teach us that during the quark epoch, the universe was isotropic, consisted entirely of QGP and had a temperature of the order T = O(100 MeV), the minimum temperature for colour deconfinement during the conditions of the universe at that time [19](4-8) [20](146) [21]. Although the quark epoch only lasted



Figure 2: Phase diagram of quark matter [24]. The QCD transition is indicated by the orange area.

for a small fraction of a second, it is still important to understand how QGP behaves. Understanding the behaviour of QGP is a necessary ingredient to building an accurate model of both the quark epoch itself and its transitions to and from the other epochs.

The quark epoch is possibly not the only time in the universe in which quark matter existed naturally. For example, some speculate that the core of neutron stars consists of quark-gluon plasma [22](55-60). There are even some who hypothesise the existence of stars made entirely out of quark matter, called *quark stars*. Although this idea was formulated in the sixties [23], no conclusive evidence of such stars has been found yet.

In order to study QGP experimentally, we need to generate it in collision experiments. It is after all very difficult to study the core of a neutron star or the very early universe through direct observation. To generate QGP in collider experiments, it is necessary for the collision energies to exceed the minimum temperature required for colour deconfinement. This would very briefly generate QGP, allowing us to study its properties through how decay particles interact with it. The aim is to determine an equation of state of QGP through analyses of collisions in particle accelerators.

#### 2.4 Measurable properties of QGP

QGP has several different proposed properties which can be tested via experiments. Some of these properties center around the phase transition from QGP to hadronic gas, the *QCD transition*, illustrated in Figure 2. For example, it is currently not known whether the QCD transition is of first or higher order. This is still an active research topic and there is no conclusive evidence for either at this moment. The focus of this particular project will be on constraining theoretical models by analysing the in-medium energy loss of partons as a function of the number of nucleons involved in the heavy-ion collision. The number of nucleons involved is a proxy for the achieved energy density and is therefore directly related to the heat and density of the produced QGP. We will specify this further in Chapter 4.



Figure 3: The full CERN accelerator complex in 2019 [25].

# 3 Experiment

#### 3.1 Large Hadron Collider

In this chapter we discuss the experimental setup for this project and the data processing infrastructure used by the CERN collaboration. As we discussed earlier, the Large Hadron Collider is one of the largest machines ever created. In essence it is a circular tunnel in which hadrons are accelerated in opposite directions, colliding at of the experiments along the ring. A schematic of the CERN accelerator complex including the LHC is depicted in Figure 3. The LHC is 26 659 m long and can run over 1 billion particle collisions per second [26]. The LHC can accelerate several types of hadrons, such as protons, gold ions and lead ions. The choice of particle depends on the type of analysis being done by the detector in question. For this analysis for instance, collisions of two lead ions were considered (*Pb-Pb collisions*). These particles are accelerated along the ring using an array of magnets. 1232 dipole magnets are used to keep the particle beam from hitting the side of the tunnel, while 392 quadrupole magnets focus the particle beam, increasing its density and maximising collision likelihood. Since these particles achieve ultra-relativistic velocities before colliding, the magnets have to be strong enough to keep them in a circular path. The LHC therefore consumes an enormous amount of energy: roughly 1.3 TWh worth of energy per year, while for example the municipality of Utrecht, including transport, electricity and natural gas, consumed 6.315 TWh in 2017 [27]. Naturally, most metals are not capable of sustaining these currents without overheating, which is why the LHC uses superconductors at 1.9 K.

When particles are accelerated, they do not enter the main ring directly. A system of smaller accelerators is used to increase the energy of the particles before entry [28]. The system used depends on the particle being accelerated; since we will use Pb-Pb collisions, we will focus on that. Lead ions are first accelerated in a small linear accelerator called LINAC3 in order to ionise them further and give them an energy of 4.5 MeV per nucleon (note that for heavy ions, one usually refers to the energy per nucleon, rather than the energy of the ion as a whole). Then they are accelerated further in the Low Energy Ion Ring (LEIR). Once they reach an energy of 72 MeV per nucleon, they proceed through the same accelerating process as protons would. They are accelerated up to 5.9 GeV per nucleon in the Proton Synchrotron (PS) and then up to 177 GeV per nucleon in the Super Proton Synchroton (SPS). Finally, the particles are injected in the LHC, where they are accelerated up to a centre of mass energy of 5.02 TeV in our case. Two beams of ions are accelerated in opposite directions and will collide in one of the four main detectors along the LHC: ATLAS, CMS, LHCb or, in our case, ALICE.

## 3.2 ALICE

As stated above, ALICE is one of the four main detectors along the LHC. As can be seen in Figure 4, the detector is a large device consisting of many components. Despite being called one of the smaller detectors of the four, the detector has dimensions of 16 m by 16 m by 26 m and weighs 10 000 tons [29]. ALICE is operated by the ALICE Collaboration, which is an independent research organisation (just like the other detector institutions) in order to allow for unbiased reproduction of results by other detectors along the LHC. Currently, over 1000 scientists are part of the collaboration, distributed across over 100 institutions in roughly 30 countries. ALICE was designed specifically to analyse heavy-ion collisions in order to further our understanding of QGP and other phenomena of QCD at extreme energy densities [2]. Studying QGP effectively requires ALICE to have at least the following detection capabilities [29](4):

- Reconstruct all particle tracks;
- Measure particle momenta, both very low and very high;
- Identify particles;
- Observe decay vertices of signature quarks very close to the collision point.

In order to reconstruct all these unknowns, ALICE uses several different detectors. We will now discuss the most important ones used in this analysis. Note that at the moment of writing, the ALICE detector is undergoing several upgrades during Long Shutdown 2 (LS2) of the LHC. These include an upgrade to the TPC, ITS and MFT and a new component: FIT (Fast interaction trigger) [30]. These upgrades were not yet employed in this analysis, as they are set to complete in 2021.



Figure 4: Schematic of the ALICE detector [31].



Figure 5: Schematic of the ITS [32].

#### 3.2.1 ITS

At the heart of ALICE lies the *beam pipe*, through which the particle beams pass. Directly around this pipelies the ITS (*Inner Tracking System*). The ITS is one of the detectors which reconstructs particle trajectories in the central barrel, a schematic of which is depicted in Figure 5. The ITS consists of several components. The innermost detectors are two *silicon pixel detectors* (SPDs). An SPD records the x and y position of a particle passing through it<sup>3</sup>. It is a checker-board-like device with individual detection elements of 0.05 mm by 0.5 mm, resulting in roughly 10 to 100 million different detection channels. Through its very high resolution and close proximity to the collision it is able to give very accurate coordinates of particle tracks and the collision vertex. The next two layers around the SPDs are *silicon drift detectors* (SDDs). An SDD is also capable of pixel-like recording the x and y position of particles passing through it with high accuracy. Although it is not as accurate as an SDD, it was chosen as a cost-effective option for the intermediate layer of the ITS [29](6). Finally, the outer layers of the ITS consist of two *silicon strip detectors* (SSDs), which are not pixel detectors like the SPDs and SDDs but still reliably trace track positions of particles passing through them. Additionally, the SPDs and SDDs contribute to *particle identification* (PID) as well. This is because they can measure the energy loss of particles travelling through them, which is a unique property of every particle.

#### 3.2.2 TPC

The next layer around the beam pipe is the *Time Projection Chamber* (TPC). The TPC is the main device used for tracking particles within ALICE. It consists of a large cylinder filled with gas (85% neon gas, 10% carbon dioxide and 5% nitrogen gas [33]), separated into two halves by an electrode, as depicted in Figure 6. When a particle travels through the gas, it ionizes some of the gas molecules. The free electrons then drift towards the detectors of the TPC, which triggers a signal. By combining the drift time (determined by the geometrical location of the ionized gas molecule) and trigger location of the signal, particle tracks can be calculated. Due to the large number of particles travelling through the TPC per collision, it generates a huge amount of data. This is compensated for through its reliability and ability to track almost all particles travelling through it after a collision. Additionally, the TPC has a role in particle identification similar to the SPDs and SDDs due to its ability to measure energy loss of particles travelling through it.

<sup>&</sup>lt;sup>3</sup>The ALICE detector uses Cartesian coordinates to describe particle tracks. Beams are set to travel in the z direction, while x is taken along the width of the detector and y along the height.



Figure 6: Schematic of the TPC [29].



Figure 7: Schematic of a TOF module within the ALICE frame [34].

#### 3.2.3 TOF

Another layer in the detector consists of the *Time Of Flight system* (TOF). Whereas the previously mentioned detectors serve primarily to track particles with a complementary PID function, the TOF's primary function is particle identification. The TOF system is split into 18 sectors in the azimuthal direction featuring in total 1638 *Multi-gap Resistive Plate Chambers* (MRPC strips). These strips each contain 10 gaps of 0.2 mm and are used to measure the time between collision and the particle entering the TOF detector (*time of flight*) with a resolution which can be as small as 56 ps, depending on several factors [35]. By using the momentum and time of flight of a particle, its mass can be determined, allowing for particle identification.

#### 3.2.4 V0

Although the ALICE detector contains many more modules, the final one we will discuss here is the V0 detector. The V0 is of particular importance to this thesis as it is used to determine the centrality<sup>4</sup> of a given collision and provides two specialised triggers for this: a central trigger and a semi-central trigger. Besides that it also measures the charges of particles and their arrival times and provides a minimum-bias and multiplicity trigger for the detectors around the beam pipe in the centre [36]. The V0 detector consists of two arrays of 32 scintillating counters. They are placed along the z axis of the detector, on opposing sides of the collision point in the beam pipe. The two components of the V0 detector are called VO-A and V0-C and have a slightly different design, as can be seen in Figure 8.



Figure 8: Schematic of V0-A and V0-C [37].

<sup>&</sup>lt;sup>4</sup>Centrality relates to the geometry of the collision. A more thorough definition of centrality will be given in the next chapter.

## 3.3 Worldwide LHC Computing Grid

The ALICE detector generates very large volumes of data, through a combination of very sensitive detectors and the high number of collision events. It would not be feasible for normal computers to process all this data. This is why CERN uses a computer network called the Worldwide LHC Computing Grid, or for short the Grid, to process the raw data from the detectors. The Grid is a worldwide collaboration of roughly 170 computing centres spread across 40 nations [38], displayed in Figure 9. Here we will briefly discuss the basics of how data analysis works specifically for the ALICE Collaboration.

#### 3.3.1 Raw data processing

During LHC operation, the Grid processes between 50 and 70 petabytes of data per year. The Grid deals with this by preselecting data before saving it permanently. During operation, data goes through a two-step process [39]. In the first step, algorithms are employed to broadly preselect events which could be interesting. This generates a data set which is a lot smaller than the original detector output and thus easier to handle; only one in every ten thousand events is stored in this step. This is then again processed by algorithms which are more specialised to filter for certain events, further reducing the number of events by a factor of one hundred. Effectively, only the events which could yield interesting physics are saved, roughly one in a million collision events. After this preselection, the data is reconstructed digitally and also copied to permanent tape storage at the *CERN Computing Centre*. This phase of data processing is called *Tier 0*. The data then proceeds to *Tier 1*; distribution across 11 international computing centres for storage, reprocessing and analysis. Finally, the data enters *Tier 2* processing. This is where simulations are generated and where end-users such as researchers can access and use the data. For this analysis, we are involved in the Tier 2 data sets generated from data from the ALICE detector.



Figure 9: Map displaying data centres used by the Grid as shown by the ALIMonitor on March 24th, 2020.

#### 3.3.2 Data sets and Monte Carlo simulations

The data sets available to ALICE researchers are still too large to download on individual computers. Because of this, ALICE employs a system where users can submit analysis jobs to the Grid. For example, a researcher writes a particular analysis task which they would like to run over all 2018 Pb-Pb data. This analysis can then be submitted to the Grid via either a private job or via a *LEGO train*. In a private job, a user can use (a fraction of) the computing power of the Grid to run a customised analysis algorithm across specified data sets, in this case particular parts (runs) of the 2018 Pb-Pb data. It is however generally advisable to use LEGO trains where possible. A LEGO train is a way of more efficiently using the computing capacity of the Grid. In principle, rather than running the task themselves, the researcher can submit a waqon for a particular train with the settings of their analysis task. They can then specify over which data sets this should be run. A couple times per week, a train operator then runs all submitted wagons across the specified data sets at the same time. That way, the Grid can work more efficiently, saving computing power for other tasks. The limitation of the LEGO trains is however that one can only run algorithms which are universally implemented in AliPhysics; a topic which will be discussed in the next section. Note that the use of this system implies that the Grid is not just important for raw detector data processing; it is also vital for the other studies on data being conducted at ALICE, as can be seen in Figure 10. A great deal of data is being processed, despite the LHC currently being shut down.

ALICE researchers also have access to simulated data sets called *Monte Carlo simulations* (MC). MC simulations are generated using several simulation programs such as PYTHIA, FONLL and the actual detector output [40]. The MC sets are available to researchers just like 'normal' data sets. There are several reasons to employ MC, despite it being simulations:

- MC simulations are used in several algorithms to calculate the efficiency of an analysis, as well as the acceptance levels of certain results.
- MC sets are typically much smaller and easier to handle than real data sets, allowing researchers to run tests on whether their algorithms are behaving as expected<sup>5</sup>.
- Since MC sets are generated, it is possible to separate signal and background noise completely in the results. That way, MC can be used as a 'sanity check' for data results. If there is a large discrepancy between data and MC when run through the same analysis, that could indicate a mistake in the analysis.



#### Last 24 hours throughput

Figure 10: Grid usage in the previous 24 hours at 14:30, March 25th, 2020 [38]

 $<sup>^{5}</sup>$ Note that Grid analyses can also be run in 'test mode' in private jobs to test them, which is a more common way to test algorithms before their submission to the Grid. In test mode, a private job simulates a Grid node on the users computer to check compatibility and catch errors.

The analyses referenced above are all written in ROOT. ROOT is a modular scientific software toolkit built in C++. It is used by researchers within CERN to perform analyses on collider data and was developed for this specific purpose. Therefore ROOT is able to process very large data sets and at the same time provide some options for plotting several types of graphs, the most used ones being histograms. ROOT allows users to access and analyse data sets, save new data sets, create graphical representations of results, calculate fitting models and has some integration with other popular scientific programming languages [41]. ALICE also has a software toolkit which is implemented in ROOT called *AliROOT*. AliROOT has more built-in functionalities that are used to analyse ALICE data. The collection of available classes provided by ALICE is called *AliPhysics*. Now we can clarify what was stated in the previous section about LEGO trains. Although LEGO trains are a powerful analysis tool, the analyses which can be run are limited to those available inside AliPhysics, which means that customised macros that need to be run over entire datasets need to be run in private jobs.

This analysis was conducted using AliPhysics to analyse data and MC through Grid calculations, as well as modifying existing tasks for specific analysis purposes. The bulk of calculations could be performed using wagons on LEGO trains but processing the data from these wagons was done locally by writing or editing analysis macros written in ROOT.

# 4 Collider physics

In this chapter we will discuss physical definitions and processes relevant to this analysis.

#### 4.1 Heavy ion collisions

As we discussed in the previous chapters, we can use the LHC te generate QGP by colliding heavy ions. In such a collision, the high energies of the colliding relativistic particles transform these nucleons into a 'fireball' of QGP. We have up until now mostly discussed how and why we want to generate QGP, now we will discuss the evolution of QGP once generated in a collision. In Figure 11 we can see the time evolution of QGP in a heavy ion collision. If the nuclear matter does not reach the energy density required for deconfinement, no QGP is generated and the collision follows the evolution on the left of the time axis. If however deconfinement is achieved, we follow the evolution on the right. We now discuss this evolution phase by phase.



Figure 11: Schematic of the evolution of a collision experiment with and without QGP [42].

#### 4.1.1 Initial and pre-equilibrium phases

Ions involved in collisions are initially travelling very close to light speed before crashing into each other. Due to this they experience strong Lorentz contraction in the z direction, effectively turning them into discs in the laboratory frame. Because the quarks inside these ions have an internal velocity much lower than the velocity of the ions as a whole, we can describe the ions as being solid on short timescales, since they experience time much slower than the laboratory frame [33](6-8). On longer timescales, we can classify this matter as a fluid, with the matter as a whole being classified as *colour glass condensate* (CGC). At some point, the particle collision will start and the nucleons will start to overlap. Then the matter enters a state called *pre-equilibrium*. According to the theory of CGC, the initial moment of collision allows for hard/elastic scattering (collisions where most of the momentum is transferred rather than absorbed as energy) between individual quarks of the two nucleons.

#### 4.1.2 QGP phase

After the hard scattering of the pre-equilibrium phase, the matter expands and an increasing number of soft-scattering particles are created. This increases the temperature until thermal equilibrium is achieved (*thermalisation*). If the requirements in terms of density and energy are met, this is where QGP is created. This QGP only lasts for a brief amount of time, as it quickly expands and thus cools down. According to measurements by the ALICE Collaboration, the QGP phase only lasts for  $\mathcal{O}(10^{-23} \text{ s})$  [43]. During the QGP

phase, the hot and dense matter behaves as a perfect fluid and is well described by relativistic hydrodynamics [16](25).

#### 4.1.3 Hadronisation and freeze-out

After forming, the plasma enters its phase transition towards a hadronic gas. This stage is called *hadronisation*. At some point in time, the requirements for the existence of QGP are no longer met. However, colour confinement does not abruptly reappear. First, the matter enters a *mixed phase*, in which the plasma hadronizes but some soft scattering still occurs, allowing QGP to still exist locally. Finally, once the plasma has expanded enough, the QGP disappears entirely as the temperature drops below the minimum required for inelastic collisions. The matter completely transforms into hadronic gas. Since the hadronic formation has now fully happened and hadrons are from now on confined again, the composition of hadrons at this time is 'frozen'. This is called *chemical freeze-out*. After the chemical freeze-out, the matter consists of a hadronic gas which is still dense enough for individual particles to interact inelastically, which is important to account for as it changes their momentum. The medium continues to expand and at some point it is dilute enough that particles no longer interact. The paths of individual particles are then 'frozen', leading to the name of this state being *kinetic freeze-out*. After kinetic freeze-out, the only thing left to do for the particles is to be picked up by the ALICE detector systems.

The hadrons which were created in the hadronisation and freeze-out phases carry information about the initial conditions of the collision and formation of QGP. That is why it is crucial to measure both the particle identities and properties of decay products of Pb-Pb collisions. Certain particles strongly interact with the plasma and can therefore be used as probes of the existence of QGP, as well as the heat and density of the plasma produced.

#### 4.2 Definitions

In order to quantify the amount of QGP generated as a function of centrality, we need proper definitions of these observables. Here we will state the basic definitions used in heavy-ion physics.

#### 4.2.1 Transverse momentum

As any informed reader should know, all particles carry momentum and momentum must always be conserved. Within the coordinate frame of the LHC (z along the beam axis), two beams which will collide inside the ALICE detector have a great deal of momentum along the z axis. Most of this momentum will be cancelled in the collision since the particles travel in opposing directions. Before collision, there is no momentum in any other direction apart from this axis. However, after collision, some decay particles will of course carry some momentum perpendicular to the z axis. This component of their total momentum is called *transverse momentum* ( $p_{\rm T}$ ) and is generally measured in GeV/c. For this particular study, we are interested in particles which carry between 3 GeV/c and 36 GeV/c of transverse momentum. The amount of transverse momentum carried by a particle depends on several factors, including the geometrical position and elasticity of the collision of its mother particle(s).

#### 4.2.2 Rapidity and pseudo-rapidity

Since particles inside the LHC travel at relativistic velocities, it is not practical to refer to particle velocities using standard units. As basic Special Relativity teaches us, it is not relativistically viable to simply add up two velocities to obtain a total velocity. However, making the full calculation using the Lorentz factor every time would be very cumbersome and heavy on computers performing calculations. It is much easier computationally to add velocities linearly instead by defining a new quantity which incorporates a general Lorentz boost in its definition, taking care of the relativistic component of velocity addition. Noting that a Lorentz boost can be represented by

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh \chi & \sinh \chi \\ \sinh \chi & \cosh \chi \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix} , \qquad (4.1)$$

where accents indicate the different reference frames and  $\chi$  is some boost parameter, a suitable definition for the sought after quantity turns out to be the following. We define *rapidity* (y) as

$$y = \frac{1}{2} \ln \frac{E + p_{\rm T}c}{E - p_{\rm T}c} , \qquad (4.2)$$

where E is the particle energy and  $p_{\rm T}$  the transverse momentum. The natural logarithm is a natural consequence of incorporating a Lorentz boost with rotational generalisation as seen in Equation 4.1. Using rapidity allows for linear addition and also satisfies Equation 4.1 when taking  $\chi = y$ . Although rapidity allows for mathematically correct calculations, an often used substitute is called *pseudo-rapidity* ( $\eta$ ). The pseudo-rapidity is found by making the assumption that for sufficiently high speeds, i.e.  $v \to c$ , we have that  $E \approx p$ , which directly leads to

$$\eta = -\ln(\tan\frac{\theta}{2}) , \qquad (4.3)$$

where  $\theta$  is the angle between a particles momentum and the beam direction. In practice, pseudo-rapidity is a very good approximation and is therefore extensively used in LHC particle physics. Also, we can now tie in to the components of the detector we discussed earlier. Different components of ALICE are optimised for different ranges of  $\eta$  and altogether they provide near-total coverage, but notably not for  $|\eta| > 0.9$ .

#### 4.2.3 Centrality and $\langle N_{part} \rangle$

A very important parameter in particle collisions is the geometry of the collision. In practice, we define this by observing how many nucleons are actually involved in the collision. From now on, we define *participants* as nucleons which are involved in the collision, and *spectators* as ones which do not directly interact with nucleons from the colliding ions. We quantify the number of participants by using the *centrality*, which is a directly related quantity. First however we must define the *impact parameter* (b). This is the distance between the two centers of the colliding nucleons perpendicular to the beam axis, as displayed in Figure 12. Note the disc-like shape of nucleons, which is caused by Lorentz contraction at high velocities. *Centrality* can be defined as the effective fraction of the total nuclear interaction cross section, such that

$$C = \frac{1}{\sigma} \int_{b_{thr}}^{b} \frac{d\sigma}{db'} db' , \qquad (4.4)$$

where C is the centrality, b the impact parameter,  $b_{thr}$  the lower threshold (sometimes taken as 0) and  $\sigma$  the total nuclear interaction cross section. Strictly speaking, this is a value between 0 and 1. However, in this study we will usually implicitly express it as a percentage between 0 and 100, since this is conventional and allows easier use in the algorithms we employ for our study. A collision with low impact parameter and thus low centrality is called a *central* collision, while for larger b we refer to them as *semi-central* or *peripheral* collisions. Now we introduce two other important quantities:  $N_{coll}$ , which is the number of binary nucleon-nucleon collisions, and  $N_{part}$ , which is the number of participants in a collision. The latter is directly related to the centrality through its relation to the cross section and is used to relate the centrality to the  $R_{AA}$ , which we will discuss in the next section.

#### 4.2.4 Nuclear modification factor $(R_{AA})$

The nuclear modification factor  $(R_{AA})$  is a way of measuring the energy loss of particles travelling through QGP generated in a heavy ion collision. Since QGP is strongly interacting, heavy flavour quarks like charm in particular are affected by the medium. This means that these quarks will interact with the plasma and lose some energy, resulting in their transverse momentum  $p_{\rm T}$  being modified. The  $R_{AA}$  as a function of momentum will change shape as a result, which is intrinsically related to the heat and density of QGP generated in heavy ion collisions.  $R_{AA}$  is defined as

$$R_{AA} = \frac{1}{\langle T_{AA} \rangle} \frac{dN_{AA}/dp_{\rm T}}{d\sigma_{AA}/dp_{\rm T}}$$
(4.5)



Figure 12: Layout of a heavy ion collision. On the left, b is the impact parameter. On the right, we see the difference between spectators and participants [44].

where  $\langle T_{AA} \rangle$  is a function accounting for the thickness of the nucleus,  $dN_{AA}/dp_{\rm T}$  the differential yield of the heavy ion collisions and  $d\sigma_{AA}/dp_{\rm T}$  the differential cross section of proton-proton collisions. Note that in this equation, the proton-proton collisions serve as a baseline to compare the heavy ion collisions to, and the thickness function as a normalisation of the quantity. Proton-proton collisions are a valid choice of reference since no QGP is expected to be generated in these events. Thus, for relatively cold or low-density plasma, one expects to find that the nuclear modification factor is close to 1. For hotter and denser QGP, we expect the nuclear modification factor to drop. If no QGP is produced, the  $R_{AA}$  is expected to be 1, in the case that cold-nuclear-matter effects are negligible [45].

# **4.3** $R_{AA}$ versus $\langle N_{part} \rangle$ hypothesis

Previous studies into the  $R_{AA}$  have indicated a relation between the  $R_{AA}$  and  $\langle N_{part} \rangle$ . If a collision is more central, we expect  $N_{part}$  to be larger which causes a hotter and denser QGP to be created, which causes a larger suppression in particle yield and therefore a lower  $R_{AA}$ . Indeed, using the data from the LHC running at  $\sqrt{s_{NN}} = 2.76$  TeV, some centrality dependence of the nuclear modification factor of  $D^{*+}$  was confirmed as seen in Figure 13. In this study we will go through a similar process, but this time we use the 2018 set of collision data and MC simulations.



Figure 13: Centrality dependence of the nuclear modification factor as found in a 2012 ALICE study for several D mesons. Both figures show a different momentum range [46].

# 5 Analysis

Using the theoretical background described in previous chapters, an analysis was conducted to investigate the  $R_{AA}$  dependence on  $\langle N_{part} \rangle$ . First we will give an overview of the method and its steps.

#### 5.1 Method

For this project, we wish to fix the transverse momentum and instead focus on the centrality as a variable rather than a parameter. This results in this analysis principally consisting of many iterations of often-used processes, repeated for all centrality classes we are interested in. Therefore, we will usually show only one of the several classes of results in this chapter as examples, to prevent a cluttered and illegible thesis. A complete set of results is available in the appendices.

Firstly, we choose our particular decay channel, which determines which decay reaction we will use as a probe in order to build the  $R_{AA}$ . The next step of this analysis is to determine which centrality and  $p_{\rm T}$  ranges we wish to analyse. Then, the principal processing of data is done by analysing the raw ALICE data and MC simulations and cutting out all events which are not (presumably) related to our decay channel. This leaves us with many  $p_{\rm T}$  dependent *invariant mass spectra* which display the measured mass difference between  $D^{*+}$  and  $D^0$ . An important part of that process is determining what parameters to use to find our particle. These mass spectra contain relevant information on the significance of the measurement.

In order to build the nuclear modification factor, we also need the efficiency of each measurement. This is calculated using MC simulations of the data. Additionally, some work on rebinning the pp reference cross section is required to make our analysis compatible with the  $R_{AA}$  calculation. Together with the mass spectra and efficiency data, we can then reconstruct the true particle spectra and finally the nuclear modification factor. Once we have repeated this process for all centrality ranges, we can plot these as a function of centrality and then  $\langle N_{part} \rangle$  for fixed  $p_{\rm T}$ .

We are however not quite done yet. The cuts we use are not optimised for all centrality classes we use them in. Therefore, we will next proceed to optimise them using specialised algorithms and analysis over raw data. This will in turn allow us to construct new, improved cuts which yield higher significance than the unoptimised cuts. We validate this by comparing the new cuts to earlier MC calculations and then repeat the entirety of the  $R_{AA}$  calculation process (from mass spectra in data and MC to  $R_{AA}$  as a function of  $\langle N_{part} \rangle$ ).

Once we have our centrality dependent nuclear modification factor, we calculate the systematic error of our measurement. This requires several different analyses varying from the stability of our cut file to the effect of PID on our results. Once we have our results including errors, we can finally compare our results to theoretical predictions and discuss them.

#### **5.2** $D^*$ decay channel

For this analysis, we will use this specific decay channel to analyse the nuclear modification factor:

$$D^{*+} \to D^0 \pi^+ \to K^- \pi^+ \pi^+$$
 (5.1)

This particular decay channel can be identified using the properties of its decay products. A schematic of this decay channel can be seen in Figure 14.

We look at the cases in which the  $D^*$  immediately decays into a  $D^0$  and a pion. We focus on the case where this  $D^0$  meson decays into a kaon and another pion. These decays are not the only ones possible. In fact, they are quite rare. In a typical collision of 0-10% centrality 60 charm-anticharm pairs are produced, of which roughly 20% fragment on  $D^{*+}$ . The decay into  $D^0$  and a pion has a chance of 67.7% of occuring and the decay of  $D^0$  into a kaon and a pion has only a 3.9% probability [47]. Since additionally not all charm-anticharm pairs come through the acceptance levels of the ALICE detector, finding this decay chain is not very common.



Figure 14: Schematic of the  $D^{*+}$  decay chain in a heavy ion collision. Note that in this schematic we did not account for momentum conservation or magnetic fields in order to keep the figure legible.

The ALICE detector can identify specific decay chains by identifying its decay products, as this is what is measured by the detector. However, not every measurement of two pions and a kaon is an indication of this decay chain, especially considering its slim probability of occurring. Therefore we use *topological cuts* to separate background from signal. These cuts use characteristic properties of this decay to split the actual events from the background. It is therefore very important to use accurate cut files as they basically determine not only the number of events successfully detected but also the significance of this measurement.

#### 5.3 Choosing the centrality and $p_{\rm T}$ ranges

Since we want to analyse centrality as a variable, we have to maximize the range of centralities we perform our measurement at. In practice, we choose to measure from 0% up till 80% centrality. Within this determined centrality range, we have to chose a binning which is suitable for our measurements. We will use the binning 0-10%, 10-20%, 20-30%, 30-50%, 50-60% and 60-80%. This is a balance between the need to have a fine a binning as possible and the need to have enough data in every bin to perform a solid invariant mass analysis.

We must also define our  $p_{\rm T}$  range. Our total range will cover  $3 \,{\rm GeV/c}$  up till  $36 \,{\rm GeV/c}$ . For momenta lower than  $3 \,{\rm GeV/c}$ , there would be too much background to make a good fit of the mass spectra and therefore to reliably reconstruct the  $D^*$  meson, while for momenta higher than  $36 \,{\rm GeV/c}$  there will be too few events to get a significant result. Within this momentum range, we choose the binning of 3-5, 5-8, 8-12, 12-16 and 16-36 GeV/c. For our actual centrality-dependent studies later, we will fix momentum to 3-5 GeV/c and 8-12 GeV/c. We choose to make the binning large enough to have plenty of events in a given bin, while still allowing us to distinguish lower and higher momentum events and determine the shape of the  $R_{AA}$ .

#### 5.4 Data sets

The data sets used in this analysis were LHC18q\_pass1 and LHC18r\_pass1. These data sets consists of several (sub)runs, which are distinguished through their different run numbers. The runs used in this analysis are listed in a table in the appendix. LHC18q\_pass1 consists of 126 such subruns and LHC18r\_pass1 of 90 subruns. The data sets were of course of Pb-Pb collisions with a centre of mass energy of  $\sqrt{s_{NN}} = 5.02$  TeV. They are indeed very large: the total data input from these data sets to be processed by the Grid is usually around 200 to 400 terabytes, depending on the specific set. Naturally, only a very small part of this is actually interesting to us. We will run our analyses over these data sets and separate out the interesting  $D^*$  meson events using

cut files. Once this is done, we can see how many events we managed to find in these data sets.

#### 5.5 Original cut files

For this analysis we have to create suitable cut files specifically for our centrality ranges. These cuts are based on three sources. From earlier preliminary studies by the ALICE collaboration [47], we take the cuts optimised for 0-10% and 30-50%. We also source a 60-80% cut file from an earlier analysis note and rebin it for our purposes. We then apply the 0-10% cut values for 0-10% and 10-20%, the 30-50% cut values to 20-30%, 30-50% and 50-60%, and the 60-80% cut values to just 60-80%. This method is justified by presuming that these groups of centrality intervals have enough in common to use a similar cut file. Our intention is to analyse the feasibility of this project first using these cuts and optimising them later if we deem it to be so. Cut files take 16 values per momentum bin. These 16 variables are physical observables we use to separate  $D^{*+}$  from background. They are (in order):

- 0. Invariant mass  $(\text{GeV}/\text{c}^2)$ ,
- 1. Distance between the two decay lines (cm),
- 2.  $\cos(\theta^*)$  where  $\theta^*$  is the angle where  $\theta_P$  is the angle between the actual  $D^0$  flight path and the reconstructed momentum of  $D^0$  from the kaon and pion momenta in the  $D^0$  rest frame,
- 3.  $p_{\rm T}$  of kaon from  $D^0$  decay (GeV/c),
- 4.  $p_{\rm T}$  of pion from  $D^0$  decay (GeV/c),
- 5. d0 of kaon from  $D^0$  decay<sup>6</sup> (cm),
- 6. d0 of pion from  $D^0$  decay (cm),
- 7. The dot product of the d0 for the kaon and the d0 for the pion from  $D^0$  decay (cm<sup>2</sup>),
- 8.  $\cos(\theta_P)$ , where  $\theta_P$  is the angle between the actual  $D^0$  flight path and the reconstructed momentum of  $D^0$  from the kaon and pion momenta in the laboratory rest frame,
- 9. Invariant mass half width of  $D^*$  (GeV),
- 10. Half width of  $(M_{K\pi\pi} M_{D^0})$  (GeV),
- 11. Minimum momentum of soft  $pion^7$  (GeV/c),
- 12. Maximum momentum of soft pion (GeV/c),
- 13. Angle between soft pion and decay plane of  $D^0$  (rad),
- 14.  $|\cos(\theta_{PXY})|$  where  $\theta_{PXY}$  is the projection of  $\theta_P$  in the (x, y) plane<sup>8</sup>,
- 15. Normalised decay length of  $D^0$

Note that some of the cuts use a different  $p_{\rm T}$  binning from ours. However, this is no problem, as the binning used is finer than or the same as the one we intend to use. This means that we can easily rebin the results later when fitting the results. The next page contains the parameters of the cut file used for 0-10% and 10-20% as an example. The other cuts can be found in the appendix.

<sup>&</sup>lt;sup>6</sup>This is found by drawing a straight line through the kaon/pion vector and the  $D^0$  decay point and then taking the minimum distance of that line to the primary vertex of the collision.

<sup>&</sup>lt;sup>7</sup>The 'soft' pion is the one generated during  $D^*$  decay rather than  $D^0$  decay.

<sup>&</sup>lt;sup>8</sup>Although this might seem superfluous, this variable exists separately from  $\cos(\theta_P)$  because the ALICE detector is much more accurate in the (x, y) plane.

Cut for 0-10%, 10-20% centrality	Invariant mass	dca	$\cos \theta^*$	$pT_K$	$pT_{\pi}$	$d0_K$	$d0_{\pi}$	d0d0
1-2 GeV/c	0.025	0.024	0.8	0.9	1.0	0.1	0.1	-0.00045
$2-2.5~{ m GeV/c}$	0.025	0.024	0.8	0.9	1.0	0.1	0.1	-0.00045
2.5-3 GeV/c	0.025	0.024	0.8	0.9	1.0	0.1	0.1	-0.00045
3-3.5 GeV/c	0.024	0.022	0.8	0.9	1.0	0.1	0.1	-0.00045
$3.5-4  \mathrm{GeV/c}$	0.024	0.022	0.8	0.9	1.0	0.1	0.1	-0.00045
$4-4.5~{ m GeV/c}$	0.03	0.021	0.9	0.9	0.9	0.1	0.1	-0.00035
4.5-5 GeV/c	0.03	0.021	0.9	0.9	0.9	0.1	0.1	-0.00035
5-5.5 GeV/c	0.032	0.021	1.0	0.9	0.9	0.1	0.1	-0.0003
$5.5-6 \mathrm{GeV/c}$	0.032	0.021	1.0	0.9	0.9	0.1	0.1	-0.0002
$6-6.5~{ m GeV/c}$	0.034	0.021	1.0	0.9	0.9	0.1	0.1	-0.0002
6.5-7 GeV/c	0.034	0.021	1.0	0.9	0.9	0.1	0.1	-0.0002
$7-7.5~{ m GeV/c}$	0.036	0.021	1.0	0.9	0.9	0.1	0.1	-0.000127
$7.5-8  \mathrm{GeV/c}$	0.036	0.021	1.0	0.9	0.9	0.1	0.1	-0.000127
8-9 GeV/c	0.055	0.021	1.0	0.9	0.9	0.15	0.15	-7.5e-05
9-10 GeV/c	0.055	0.021	1.0	0.9	0.9	0.15	0.15	-7.5e-05
10-12  GeV/c	0.055	0.021	1.0	0.9	0.9	0.15	0.15	-7.5e-05
12-16 GeV/c	0.074	0.021	1.0	0.7	0.7	0.15	0.15	-7.5e-05
16-24 GeV/c	0.074	0.021	1.0	0.5	0.5	0.15	0.15	-5e-05
24-36 GeV/c	0.084	0.02	1.0	0.5	0.5	0.2	0.2	0.0004
36-50 GeV/c	0.094	0.02	1.0	0.5	0.5	0.2	0.2	0.0004
50-70 GeV/c	0.094	0.02	1.0	0.5	0.5	0.2	0.2	0.0004

	$\cos \theta_P$	$M_{inv} hw D^*$	$hw \ \Delta M$	PtMin $\pi_s$	$PtMax \ \pi_s$	$\theta \pi_s D^0$	$ \cos \theta_{PXY} $	NDL
1-2  GeV/c	0.985	0.3	0.15	0.05	1.0	1.0	0.998	7.7
$2-2.5 \ \mathrm{GeV/c}$	0.985	0.3	0.15	0.05	1.0	1.0	0.998	7.7
$2.5-3 ~\mathrm{GeV/c}$	0.985	0.3	0.15	0.05	1.0	1.0	0.998	7.7
3-3.5  GeV/c	0.98	0.3	0.15	0.05	1.0	1.0	0.998	7.5
3.5-4  GeV/c	0.98	0.3	0.15	0.05	1.0	1.0	0.998	7.5
4-4.5  GeV/c	0.98	0.3	0.15	0.1	10	1.0	0.998	7
4.5-5 GeV/c	0.98	0.3	0.15	0.1	10	1.0	0.998	7
5-5.5  GeV/c	0.95	0.3	0.3	0.25	10	1.0	0.998	6.5
5.5-6 GeV/c	0.93	0.3	0.15	0.3	100	1.0	0.998	6.5
6-6.5 GeV/c	0.93	0.3	0.15	0.3	100	1.0	0.998	6.5
6.5-7  GeV/c	0.93	0.3	0.15	0.3	100	1.0	0.998	6.5
7-7.5 GeV/c	0.93	0.3	0.15	0.3	100	1	0.998	6
7.5-8 GeV/c	0.93	0.3	0.15	0.3	100	1	0.998	6
8-9 GeV/c	0.93	0.3	0.15	0.3	100	1	0.998	5
9-10 GeV/c	0.93	0.3	0.15	0.3	100	1	0.998	5
10-12 GeV/c	0.93	0.3	0.15	0.3	100	1	0.998	5
12-16 GeV/c	0.93	0.3	0.15	0.3	100	1	0.99	3.7
16-24 GeV/c	0.92	0.15	0.15	0.3	100	1	0.99	1
24-36 GeV/c	0.87	0.15	0.15	0.3	100	1	0.9	0.5
36-50  GeV/c	0.8	0.15	0.15	0.3	100	1	0.9	0
50-70  GeV/c	0.8	0.15	0.15	0.3	100	1	0.9	0

Table 2: Cut values used for  $0\mathchar`-10\%$  and  $10\mathchar`-20\%$  analysis.

#### 5.6 Mass spectra analysis for unoptimised cuts

Using the presented topological cuts, several tasks were submitted to the Grid to calculate the (invariant) mass spectra of  $D^*$  particles in 2018 data. As stated earlier, the data sets used were LHC18q\_pass1 and LHC18r\_pass1. Four tasks were run, over 10-20%, 20-30%, 50-60% and 60-80%, since the 0-10% and 30-50% data were already available from a previous preliminary study [47]. Upon completion, the results were taken and analysed using a macro which fits the signal to a Gaussian and background to a power and exponential function, allowing determination of for instance the mass of the particle, the width of the peak ( $\sigma$ ) and the significance. In Table 3, the number of events for each centrality using these cuts and data samples are displayed. The mass spectra created by this fitting macro of 10-20% are presented in Figure 15 as an example.

From Figure 15 we learn several things. Firstly, it is important to note that we show the measured difference in invariant mass of  $D^{*+}$  and  $D^0$  in the mass spectra displayed in this study, rather than the measured masses of both particles. We perform the invariant mass analysis on this mass difference  $M(K\pi\pi) - M(K\pi)$  in order to partially negate resolution effects on the  $D^0$  measurement, which in turn yields a sharper signal peak. Furthermore, it seems that for increasing momentum, fewer events are measured, decreasing the accuracy of the fit. Additionally, for low momentum the signal/background ratio is higher than for higher momentum. We also have to note that for 60-80% centrality (see appendix), the fitter was unable to construct a solid fit for the momentum range of 16-36 GeV/c. Figures 16 and 17 show that our choice of momentum range has given us mostly well-fitted peaks. However, we can now strongly see the decrease in signal quality for increasing centrality. This will factor into the statistical error of our final measurement later on. Especially the 60-80% centrality bin at 8-12 GeV/c has a notably worse fit than the other data points.

Centrality	0-10%	10-20%	20-30%	30-50%	50-60%	60-80%
Number of events	$8.58907 \cdot 10^{7}$	$1.874767 \cdot 10^{7}$	$1.463399 \cdot 10^{7}$	$7.38447 \cdot 10^{7}$	$1.53441 \cdot 10^{7}$	$2.33451 \cdot 10^{7}$

Table 3: Number of events for different centrality ranges analysed. Note that 0-10% and 30-50% are notably large due to the special triggers used in those regions.



Invariant mass spectra  $D^*$  at 10-20% centrality

Figure 15: Mass spectra at 10-20% centrality. The blue line indicates the signal fit and the red line the background fit. Note that these spectra display the detected mass difference between  $D^0$  and  $D^{*+}$ .



Invariant mass spectra  $D^*$  at 3-5 GeV/c

Figure 16: Mass spectra as a function of centrality at 3-5 GeV/c, where red lines indicate background and blue lines signal fits. Note that these spectra display the detected mass difference between  $D^0$  and  $D^{*+}$ .



Invariant mass spectra  $D^*$  at 8-12 GeV/c

Figure 17: Mass spectra as a function of centrality at 8-12 GeV/c, where red lines indicate background and blue lines signal fits.

#### 5.7 Comparison MC

In order to analyse the quality of the fit more thoroughly, we will analyse some of the properties of these fits more quantitatively. First, we generate the same mass spectra, but this time using MC simulations. Since MC simulations allow us to fully separate signal and background, we can use these as a control for the data.

#### 5.7.1 Peak width $(\sigma)$

First we look whether the widths of the mass peaks are in agreement. From Figure 18 we can see that the peak widths of 3-5 GeV/c are in agreement with the MC simulations. The biggest outlier is the entry at 30-50% centrality, but this is still within its statistical uncertainty. There also seems to be a slight downward trend in  $\sigma$  for data. This might be caused by the fact that central collisions generate relatively more background. Especially in the 3-5 GeV/c range, this is a large part of the measurement. Some of the background will be picked up as signal, causing the peak to be slightly wider.



Figure 18: Width of peaks in the 3-5 GeV/c range. Figure 19: Width of peaks in the 8-12 GeV/c range.

In Figure 19 it is evident that data and MC are in agreement as well for 8-12 GeV/c. The largest outlier is the data point at 60-80% centrality for data. However, we noted earlier while analysing the mass spectra that this fit in particular was not very accurate compared to the others. Note that the downward trend caused by the large background in 3-5 GeV/c is not present in this momentum range. In order to address the poor fit, we will try fixing  $\sigma$  in the data fit to the value in the MC fit for 60-80% centrality, to see the difference.

From Figure 20 we can see the effect of fixing  $\sigma$ . Although there are only slight changes in the significance value, the fit in the 8-12 GeV/c bin has not dramatically improved by fixing  $\sigma$ . Furthermore, it seems that the fit for 3-5 GeV/c, 5-8 GeV/c and 12-16 GeV/c are very slightly worse than they are without fixing  $\sigma$ . Therefore we believe there not to be enough justification to fix the peak width in these mass fits based on their MC counterparts.



Mass spectra  $D^*$  at 60-80% centrality for fixed  $\sigma$  from MC.

Figure 20: Mass fits in 60-80% centrality with fixed  $\sigma$ . Note that for 16-36 GeV/c, there were too few events for the fitter to construct a meaningful fit for. This plot can be compared to the unfixed  $\sigma$  fits found in the appendix.

#### 5.7.2 Mass

Now we will compare the masses as found by the data and MC fits. Although most of the masses are around the same expected value for 3-5 GeV/c in Figure 21, the MC point in 20-30% centrality seems to be an outlier. The point is not entirely outside of its uncertainty range, but if this point is still present after optimising the 20-30% cut, an explanation should be found. The relatively high volatility of the mass in this momentum bin could be another effect of the large amount of background present. If this is indeed the case, this will become evident upon studying the systematic error of this bin in particular.

In Figure 22 it seems that the masses are well in line with the expected values. It seems to be less volatile than Figure 21 and almost all masses lie comfortably within their MC counterpart uncertainty ranges. We do however again see that the 60-80% bin of this momentum range is noticeably different from the other values, presumably due to its relatively poor fitting.

Based on the MC data, we can proceed with these fits, while noting that for peripheral events the fits are notably less of quality than central fits, but that low momentum fits seem to be less stable due to high background. We should also take into account that 20-30% MC has shown a small outlier in mass for 3-5 GeV/c, which should be addressed by optimisation of that cut or otherwise further studied.



Figure 21: Position of peaks (i.e. measured mass) in the 3-5 GeV/c range.

Figure 22: Position of peaks (i.e. measured mass) in the 8-12 GeV/c range.

#### 5.8 Significance and 1/S

From the fitted spectra we can calculate the significance of our measurement, which will give us insight in the dependability of the results and the statistical precision we can achieve for our final result.

When looking at the significances in Figure 23, the values of 0-10% and 30-50% for both  $p_{\rm T}$  bins stand out as being significantly higher than others. This is because of two reasons. It is partly because we were using cuts which were optimised for these centrality ranges specifically, but mostly because the ALICE detector uses special V0 triggers for events with centrality 0-10% and 30-50%. This means that any measurement done within these two ranges will yield many more events and thus a higher significance than other centrality bins will.

One of the most important take-aways from Figure 23 is that the significance of our fits is never extremely low. If the significance would have been around 3 for example, it might not have been feasible to construct any meaningful  $R_{AA}$  from these cut files and data samples. However, the fits are strong enough to continue. Furthermore, the significance of the 8-12 GeV/c measurement is higher across all centralities, except for very peripheral events. This can be attributed to the fact that there is far less background in 8-12 GeV/c, allowing for more distinct signal peaks to form. In peripheral events however, the lower number of entries caused the fitted lines to align less precisely with the data, lowering the significance.

The inverse significance (1/S) shown in Figure 24 is interesting for us as well. Although it in essence shows the same information as Figure 23, the inverse significance is a more 'natural' value, since a normal distribution (on which of course the concept of significance is based) contains a factor 1/S and  $(1/S)^2$ . The values from this graph give us a quantitatively more intuitive insight in the quality of the fit.



Figure 23: Significance of mass fits in 3-5 GeV/c and 8-12 GeV/c.



Figure 24: Inverse significance of mass fits in 3-5 GeV/cand 8-12 GeV/c.

# 5.9 Reconstruction efficiency of pre-optimisation cuts

In this study we wish to study the effect QGP has on charm quarks travelling through it. We study these charm quarks by analysing the  $D^*$  decay channel. However,  $D^*$  can also come to exist through fragmentations involving beauty quarks, which means that not every  $D^*$  we detect actually indicates a charm quark affected by QGP after being created in the collision. The ones that *are* of interest are called *prompt charm*. We need to take into account the so-called charm/beauty fraction in our measurements when building the  $R_{AA}$ . This is done using existing theoretical predictions of these fragmentations and one of the ingredients we use in this process is the *reconstruction efficiency*. The efficiency from prompt charm is called *charm efficiency* or *prompt efficiency*, while the efficiency from beauty decay is called the *beauty efficiency* or *feed-down efficiency*.

Efficiencies of cuts are calculated by using a specialised macro and running our cut files through the available MC simulations of the datasets. These macros will yield a particular output, which we process in a similar fashion to the output from the mass spectra macro. From this we will retrieve two histograms per centrality range, displaying the cut efficiency in all our bins. We will naturally be interested in the efficiency as a function of centrality as well. The hope is that the efficiency does not depend too much on centrality, since this might add a bias in the final  $R_{AA}$  result.

The MC simulations used were LHC19c3a\_q, LHC19c3a\_r (for 0-10% and 10-20%), LHC19c3b\_q, LHC19c3b\_r (for 20-30% and 30-50%) and LHC16i2c\_AOD198 (for 50-60% and 60-80%). Note that the latter is in fact an MC simulation for the 2015 data rather than 2018 data. Unfortunately, the MC simulations for peripheral events in 50-60% and 60-80% centrality were not available at the time. Therefore, we have to make due with these for now. A later study could redo the peripheral events using the 2018 MC simulations. After running the calculations on the Grid, the efficiencies were calculated.

There are several important remarks we have to make about the efficiencies. Firstly, we have applied a rebinning algorithm. This is because we need our efficiencies to have the correct binning to be compatible with our  $R_{AA}$  construction macro. Another aspect of the efficiency calculation is considering which 'steps' are taken into account. The efficiency macro uses several calculation steps or levels to factor out different effects and calculate the final efficiency. They all represent a different effect which influences the final result. These steps are:

- 0. MC Limit Acceptance
- $1. \ \mathrm{MC}$
- 2. MC Acceptance
- 3. Reconstruction Vertex
- 4. Reconstruction Refit
- 5. Reconstruction
- 6. Reconstruction Acceptance
- 7. Reconstruction ITS Cluster
- 8. Reconstruction Cuts
- 9. Reconstruction PID

The steps are calculated for different sensitive variables. We are naturally interested in  $p_{\rm T}$  as the sensitive variable. The final efficiency was calculated by dividing the histogram generated in step 9 by the histogram in step 0. That way, all effects were accounted for. This yields several histograms for all centrality ranges; the result for 10-20% was included in Figure 25 as an example. From Figure 25 we can see that the charm efficiency is generally lower than the beauty efficiency. This was found to be the case for all centralities. In order to answer the question whether the efficiency varies strongly over centrality, the following ratio plots were created.

Figures 26 and 27 were created by taking all efficiencies per centrality, dividing them by the efficiency of 30-50% centrality and plotting them together. 30-50% was chosen as a baseline since it is a previously used cut which is known to give dependable results from earlier studies [47]. As we can see, there does seem to be some centrality dependence of the efficiency. These differences are especially pronounced in the 3-5 GeV/c region in beauty and 3-5 GeV/c and 16-36 GeV/c regions in charm.

Upon calculating the  $R_{AA}$ , the efficiencies will be taken into account, which should prevent bias caused by these differences in efficiency. It is however good to know that there are effects at work here, which could be investigated further if desired.



#### Efficiencies 10-20 cut

Figure 25: Efficiencies of the 10-20% centrality cut for the full  $p_{\rm T}$  range.



Figure 26: Efficiencies per centrality (beauty) as a function of  $p_{\rm T}$ , divided by the efficiencies of 30-50%. Note that in the 16-36 GeV/c range, 10-20% and 60-80% overlap almost perfectly, which is why 10-20% is hard to see.



Figure 27: Efficiencies per centrality (charm) as a function of  $p_{\rm T}$ , divided by the efficiencies of 30-50%. Note that in the 8-12 GeV/c and 12-16 GeV/c regions, 50-60% and 60-80% overlap almost perfectly, which is why 50-60% is hard to see in these ranges.

#### Efficiencies as a ratio of 30-50 - beauty

### **5.10** $R_{AA}$ before optimisation

Calculating the nuclear modification factor requires not only the efficiencies and mass spectra. We use two separate macros to come to our end results for every centrality range. The first macro reconstructs the  $p_{\rm T}$  spectra and  $D^*$  production cross sections of the measurements using our results for the efficiencies and mass spectra. We then take these and combine them with the proton-proton collision reference file using another macro in order to come to our nuclear modification factor.

The reconstruction of the true spectra and cross sections requires specifically the following input: a set of MC predictions on the properties of our decay reaction (this is available from earlier ALICE measurements), the efficiencies for charm and beauty, the reconstructed mass spectra, the centrality range, the number of events (this is included in the initial analysis results), total inelastic cross section  $\sigma$  (a constant for our purposes here) and some other settings. Using this input, the corrected yields and cross sections were calculated. As an example we include the results from 10-20% in Figures 28 and 29, the rest are in the appendix.

As expected, the yield and cross section still decrease for increasing momentum after correcting for beauty. We also see the effects of the central and semi-central triggers in these results, as 0-10% and 30-50% centrality still have noticeably larger yields compared to the other centralities. These yields and cross sections will play a role in the systematic error analysis later as well. Note that the 2015 systematics were included as an indication of what to expect, but that they will not be used for our analysis.

The next macro which calculates the  $R_{AA}$  as a function of momentum has the following input: the results from above (including settings like centrality again) and importantly the p-p reference file. Although this file is available from the ALICE Collaboration, it uses a different momentum binning as a default. Therefore, this file had to be rebinned. This was done in a similar fashion to the rebinning of the efficiencies; rebinning and then normalising for the number of merged bins. Additionally, since the systematic errors were not available yet at this time, we will for now only show the  $R_{AA}$  including statistical errors.



#### Reconstructed yield and reconstructed cross section at 10-20% centrality

Figure 28: Corrected yield at 10-20% centrality. The red bar represents the 2015 systematics.

Figure 29: Corrected cross section at 10-20% centrality. The green bar represents the 2015 systematics.


Figure 30:  $R_{AA}$  as a function of  $p_{\rm T}$  for the full centrality range, before optimisation. The dotted line indicates the maximum physical value  $R_{AA}$  can take. The missing data point for 60-80% at 16-36 GeV/c is due to the mass fitter not being able to make a fit in that bin. "Eloss hypothesis" refers to a particular energy loss effect which we did not take this into account for this study.

Figure 30 shows that measured charm quarks indeed lose some energy to quark-gluon plasma, which is indicated by the strong suppression measured in the 0-10% centrality range. Furthermore, we see that for increasingly peripheral collisions, a less hot and dense plasma is produced than for the more central ranges, resulting in less charm suppression in these regions. We see that mostly the nuclear modification factor increases for more peripheral events but remains under 1, exactly as we expected. The shape of the  $R_{AA}$  we measure, which indicates the effect of the in-medium energy loss of partons for different momentum regions, is affected by several phenomena, two of which are *flow* and *shadowing*.

Flow and shadowing are two separate effects which have a large influence on the shape of the  $R_{AA}$  [45]. These effects have various underlying causes and are mostly relevant for low momentum ( $p_{\rm T} < 4 \, {\rm GeV/c}$ ). Shadowing is an *initial-state effect*, meaning it affects the initial state of the collision. It is a so-called *cold*nuclear-matter effect. Oversimplifying, we can say that shadowing causes a decrease in parton density if we have small enough momentum per parton in a heavy ion collision. For increasing squared momentum transfer and nuclear mass number, this effect becomes stronger. Although quantification of this effect is studied in a phenomenological fashion, the most accurate theoretical description of this is given using the description of CGC effective theory stated previously, because the PDFs of partons will at some point saturate the phase space during a collision and thus no longer accurately describe the matter. In our measurements, shadowing causes additional charm suppression and thus a decreased  $R_{AA}$  in the lower momentum range. Whereas shadowing is an initial state effect, flow is a *final-state effect*. It is an experimentally-confirmed effect which is also described by CGC effective theory through hydrodynamics but other descriptions also exist. According to this, collective expansion of the medium could affect heavy-flavour hadrons as well, giving them some radial flow outwards. Experiments and calculations have shown that this effect is again the most prominent for  $p_T < 4 \,\mathrm{GeV/c}$ , but it might be present for higher momenta to some extent as well. The effects of flow are indicated in  $R_{AA}$  versus  $p_{\rm T}$  results by a peak at low momentum.



Figure 31:  $R_{AA}$  as a function of centrality before optimisation and without systematic errors.

Now that we have the momentum-dependent  $R_{AA}$ , we now plot the centrality dependent  $R_{AA}$ . This will give us insight in the behaviour for our momentum bins of interest. Figure 31 depicts the information from Figure 30 from bins 3-5 GeV/c and 8-12 GeV/c more clearly. Overall,  $R_{AA}$  does indeed seem to increase for increasingly peripheral events. Although this figure is useful, we will now plot the final result (pre-optimisation): the nuclear modification factor as a function of  $\langle N_{part} \rangle$ . In order to do this, we take the relation between centrality and  $\langle N_{part} \rangle$  calculated by the ALICE Collaboration. Using these numbers, we can create Figure 32.

Figure 32 is essentially what our end result will look like, but with added systematic errors and slightly changed values from optimisation. This can be seen as a preliminary result. Although we will not use these results to compare theory to, we have shown that it is feasible to obtain a significant and physical result using this approach, which also seems to be in line with previous studies such as the one shown in Figure 13. Now that we know this, it is time to start with the cut optimisation and then repeat the process for the new cuts.

Centrality	0-10%	10-20%	20-30%	30 - 50%	50-60%	60-80%
$\langle N_{part} \rangle$	357.3	262	187.9	109	54.34	23.35

Table 4: The relation between  $\langle N_{part} \rangle$  and centrality [48].



Figure 32:  $R_{AA}$  as a function of  $\langle N_{part} \rangle$  before optimisation and without systematic errors.

## 5.11 Optimisation 10-20%

In order to improve the significance of our end result, we will perform a *cut optimisation* on 10-20%, 20-30% and 60-80% centrality. These particular ranges were chosen because 0-10% and 30-50% were already optimised. The reason that we did not optimise 50-60% is mostly because the low quality of the earlier fit would not likely provide much improvement through optimisation and time constraints had to be taken into account as well. However, we did choose to optimise 60-80% centrality, because we would like to have a probe for peripheral events which is as significant as we can using these cut files.

The cut optimisation is performed by using the same macro which generates the regular cut files. This macro is configured to generate a special optimisation file, which takes in two sets of cuts: slightly tighter and slightly looser. The tighter values are entered for the number of variables which one wants to optimise; in our case, we chose to optimise for three variables. Naturally more variables is better, but also more time and computationally intensive. Considering we had to run this optimisation three times, we chose to optimise for the variables 'd0d0', 'cos  $\theta_P$ ' and 'Norm Decay Length' for all momentum bins and given centralities. These variables are known to be very 'strong' cuts, i.e. they cut out a great deal of background and therefore have a large impact on the significance.

The optimisation cut file is then run on the Grid over all relevant data, in our case LHC18q-pass1 and LHC18r-pass1. The data from this is then downloaded and another, local macro analyses which variations between these two boundary values yield the maximum significance fit. Note that optimising for significance does not automatically mean that the fit will be better; this is why we have to do some manual checks before assuming these values.

### 5.11.1 Running the optimisation

For 10-20%, the cut file previously optimised for 0-10% was rebinned into our desired binning (this saves a great deal of time later) and then run on the Grid. This yielded data which provided an improved cut upon analysis.

Our first check of the results is by checking the data from Figure 33. Here we see a plot of two of the three optimisation variables on both axes, and the resulting significance from using a fit with that particular combination of the two. Naturally, three sets of these plots were generated, but since they are in essence similar we will only discuss one of them here. There are several factors to take into account here. Firstly it is important to note that the plots all use different scales. Thus, all colours are relative to their own plots, meaning that for 3-5 GeV/c the scale goes up to roughly 11.2 significance, while for 8-12 GeV/c it almost reaches 14 significance. In particular, we see that there are no values within the given range where the significance is unacceptably low. This is an early indication that this cut is somewhat stable between the given intervals.

Furthermore, little 'islands' of highest significance can be seen in several plots. These indicate the 'sweet spots' where the significance is highest and also hint at some correlation between the two variables. The optimisation fitter will choose the highest significance values, thus they will be located inside these islands in plots where they exist. For some plots, especially, in higher momentum, we see that the correlation between  $\cos \theta_P$  and Norm Decay Length decreases, until there is little left. We can also see from the 16-36 GeV/c plot that the upper bound of Norm Decay Length might have been slightly too low in this momentum bin, since the significance seems to constantly rise for increases in that variable. However, we should also remember that the number of events and significance generally decreases for higher momentum and that there is probably little to be gained here in terms of significance. A further study could choose wider boundary values and investigate that further if desired.



### Some heat maps from the optimisation of 10-20%

Significance wrt cosThetaPoint vs NormDecayLenghtXY (Ptbin1 5.0<pt<8.0)

Figure 33: Heat maps depicting the optimal values for 10-20% as found by the optimisation fitter for  $\cos \theta_P$ and Norm Decay Length.

The heat maps serve mostly as a check whether the chosen optimisation boundaries were somewhat logical. Besides these heat maps, the macro used for analysis generates a great number of mass spectra using the different fit values displayed in the heat maps. It also gives the cut values and spectra for which the significance is highest.

As displayed in Figure 34, the new mass peaks seem to fit the data very well. Indeed, when we compare them to the mass spectra from Figure 15, we can see that especially the high momentum bins have seen a great deal of improvement in fit quality. Interestingly, the significance values at 5-8 GeV/c and 16-36 GeV/c actually decreased very slightly. This can be attributed to the fact that the original, pre-optimisation fit used a finer binning than this optimised fit. However, since the other bins (and especially the ones we care about for this study, 3-5 GeV/c and 8-12 GeV/c) did see some improvement and the fits also seem to follow the data a bit more closely, we will proceed by taking the resulting cut values as calculated by the optimisation. This cut is the one we will from now on use for 10-20% centrality. We will now run the same analyses over it as we did for the unoptimised cut and then we will compare the results to the old cut.

Cut optimised 10-20%	Invariant mass	dca	$\cos \theta^*$	$pT_K$	$pT_{\pi}$	$d0_K$	$d0_{\pi}$	d0d0
$3-5  \mathrm{GeV/c}$	0.024	0.022	0.8	0.9	1.0	0.1	0.1	-0.000513
5-8  GeV/c	0.032	0.021	1.0	0.9	0.9	0.1	0.1	-0.0003
8-12 GeV/c	0.055	0.021	1.0	0.9	0.9	0.15	0.15	-7.5e-05
12-16 GeV/c	0.074	0.021	1.0	0.7	0.7	0.15	0.15	-7.5e-05
16-36 GeV/c	0.074	0.021	1.0	0.5	0.5	0.15	0.15	-5e-05

	$\cos \theta_P$	$M_{inv} hw D^*$	$hw \Delta M$	$PtMin \ \pi_s$	$PtMax \ \pi_s$	$\theta \pi_s D^0$	$ \cos \theta_{PXY} $	NDL
$3-5  \mathrm{GeV/c}$	0.987125	0.3	0.15	0.05	1.0	1.0	0.998	7.875
$5-8  \mathrm{GeV/c}$	0.98	0.3	0.3	0.25	10	1.0	0.998	6.8125
8-12  GeV/c	0.973125	0.3	0.15	0.3	100	1	0.998	5.25
12-16  GeV/c	0.9375	0.3	0.15	0.3	100	1	0.99	4.85
16-36 GeV/c	0.98125	0.15	0.15	0.3	100	1	0.99	2.75

Table 5: Parameters of optimised 10-20% cut.



Optimised mass spectra 10-20%

Figure 34: New mass spectra for optimised cuts for 10-20%.

### 5.11.2 Mass spectra analysis

Using the new cut, a new analysis on the Grid was run. The new cut file was used to calculate new mass spectra in data and MC as well as new efficiencies. We will first discuss the new mass spectra as displayed in Figure 34. The data sets used are again LHC18q\_pass1 and LHC18r\_pass1. At a first glance these spectra seem to fit the data quite well, slightly better than the original fits did. We will now compare several aspects of the new pre- and post-optimisation cut to see if there was any further improvement.

In Figure 35 we see the comparison of data and MC of the mass spectra seen in Figure 34. The same MC simulations were used as before, namely LHC19c3a\_q, LHC19c3a\_r (for 0-10% and 10-20%), LHC19c3b\_q, LHC19c3b\_r (for 20-30% and 30-50%) and LHCi2c\_AOD198 (for 50-60% and 60-80%). Mostly, data and MC seem in agreement. However, the mass at 3-5 GeV/c does still seem to be slightly different. This implies that the background in data might have shifted the peak of the Gaussian slightly to the left. This is important to keep in mind if we run into discrepancies between different results later. Additionally, high momentum bins in the peak width also show some deviation. As we can see in the mass spectra of Figure 34, these fits are still of lower quality due to lack of statistics (too few events). This can be seen as another justification why we do not use these momentum bins for the nuclear modification factor calculations as a function of  $\langle N_{part} \rangle$ .



Figure 35: Comparison data and MC for the optimised 10-20% cut.

Figure 36 shows the improvement in significance for all momentum bins. As we can see, not all bins seem to have actually improved using this cut. We must again remember that this new cut has a larger binning, which will cost us some significance because the binning is less 'finely tuned'. We can however also see a slight increase in significance for 3-5 GeV/c and especially 8-12 GeV/c, which is naturally the most important for us.

Finally, Figure 37 displays the inverse significance of the optimised cut for completeness. Naturally its improvement scales inversely to the improvement of the significance, which is why we do not show a ratio plot of this as well. Using the information above and also keeping in mind that the fits are accurate according to Figure 34, we choose to use this improved cut for our final analysis.



Figure 36: Significance of the original and optimised cut for 10-20%.



1/S optimised 10-20

Figure 37: Inverse significance of optimised cut for 10-20%.

# 5.12 **Optimisation 20-30%**

Now we will perform the exact same steps for the optimisation of the 20-30% centrality range as for the 10-20% region, i.e. run the improver task over the specialised cut file, analyse the results, construct the new cut, rerun this optimised cut over data and MC and examine the results.

## 5.12.1 Running the optimisation

Once again, the improver task was run over the 2018 data in sets LHC18q\_pass1 and LHC18r\_pass1 and for the same variables, 'd0d0', 'cos  $\theta_P$ ' and 'Norm Decay Length'. This yielded several sets of heat maps, one of which is depicted in Figure 38. We once again see some islands of higher significance, however, we can also see that for increasing momentum correlation decreases, which again could be addressed by taking into account more variables or by increasing the interval between optimisation boundaries. Using the current settings, we came to the cutting parameters displayed in Table 6. In Figure 39 we see the mass spectra of the optimised cut. From these peaks we can see that the fit seems to be accurate for most bins. 16-36 GeV/c remains troublesome for the fitter as expected and 12-16 GeV/c is not a perfect fit either. However, the bins of interest for our study seem to carry strong fits. Up till now the optimisation looks promising but we will need to analyse it further before deciding whether to use it or not.

Cut optimised 20-30%	Invariant mass	dca	$\cos \theta^*$	$pT_K$	$pT_{\pi}$	$d0_K$	$d0_{\pi}$	d0d0
$3-5 \mathrm{GeV/c}$	0.032	0.022	0.8	1.0	1.0	0.1	0.1	-0.0003
$5-8  \mathrm{GeV/c}$	0.04	0.021	1.0	1.0	1.0	0.1	0.1	-0.00023
$8-12  \mathrm{GeV/c}$	0.055	0.021	1.0	0.9	0.9	0.12	0.12	-0.000116
12-16 GeV/c	0.074	0.021	1.0	0.7	0.7	0.15	0.15	-7.5e-05
16-36 GeV/c	0.074	0.021	1.0	0.5	0.5	0.15	0.15	-5e-05

	$\cos \theta_P$	$M_{inv} hw D^*$	$hw \ \Delta M$	$PtMin \ \pi_s$	$PtMax \ \pi_s$	$\theta \pi_s D^0$	$ \cos \theta_{PXY} $	NDL
$3-5  \mathrm{GeV/c}$	0.975	0.3	0.15	0.05	0.5	0.5	0.998	7.025
5-8  GeV/c	0.9825	0.3	0.15	0.25	10	0.5	0.998	7.25
8-12 GeV/c	0.96	0.3	0.15	0.3	100	0.5	0.998	4.7
12-16 GeV/c	0.955	0.3	0.15	0.3	100	1	0.99	4.8375
16-36 GeV/c	0.96375	0.3	0.15	0.3	100	1	0.99	3.75

Table 6: Parameters of optimised 20-30% cut.



### Some heat maps from the optimisation of 20-30%

Significance wrt cosThetaPoint vs NormDecayLenghtXY (Ptbin1 5.0<pt<8.0)

Figure 38: Heat maps depicting the optimal values for 20-30% as found by the optimisation fitter for  $\cos \theta_P$ and Norm Decay Length.



Figure 39: New mass spectra for optimised cuts for 20-30%.

Once again we take a closer look at the mass spectra to investigate the effects of the optimisation as well as the resulting fits. Taking the values from Figure 39 we perform the same calculations as earlier for 10-20%. Naturally, we used the same data sets as for 10-20% and also the same MC simulations. Although the masses and peak widths line up well for most bins, there is a large difference in the 3-5 GeV/c bin for mass. Just like with 10-20%, it seems that this cut file has a bias to underestimate the middle of the peak of the Gaussian. Luckily, the other momentum bin we are interested in, 8-12 GeV/c, does seem to show expected results. The fact that the underestimation of the 3-5 GeV/c bin occurs in both centrality ranges might indicate a background effect which is causing this difference. It is therefore important to note that this momentum bin might provide biased results and will probably need a relatively strong systematic error to take this into account. We will still use both of the momentum bins but should keep in mind that although the lower one seems to provide *more* statistics, it does not implicitly provide *better* statistics.



Mass versus  $p_T^2$  20-30 optimised

Figure 40: Comparison data and MC for the optimised 20-30% cut.



Figure 41: Significance of the original and optimised cut for 20-30%.



## 1/S optimised 20-30

Figure 42: Inverse significance of optimised cut for 20-30%.

Once again we calculate the improvement in significance and inverse significance, which are displayed in Figures 41 and 42. We see that the improvement is a bit more volatile, depending more strongly on the momentum bin. It has become notably better at low momentum but then worse for 5-8 GeV/c and 8-12 GeV/c. Like before we attribute this decrease to the rebinning used. Since the decrease in significance is small for 8-12 GeV/c and the increase in significance large for 3-5 GeV/c, we proceed with this optimised cut.

# 5.13 **Optimisation 60-80%**

The next step would be to list the method and results of the optimisation of the 60-80% cut. However, this optimisation did not yield any improvement on the original. We will list the steps taken anyway in order to be able to propose how one could improve upon this process in a further study.

## 5.13.1 Results of attempted optimisation

Much like the other centrality ranges, a specialised cut file was created with optimisation intervals. This file was run through the improver task on the Grid using data sets LHC18q\_pass1 and LHC18r\_pass1 and variables 'd0d0', ' $\cos \theta_P$ ' and 'Norm Decay Length'. The results of this were then analysed. However, immediately it was clear that this optimisation had not yielded anything useful. This could be seen from the following figure.

Momentum bin	0	1	2	3	4
Max sign found in bin:	0 0 0	000	0 0 0	0 0 0	0 0 0
d0d0	-0,000350	-0,000100	-0,000075	-0,000075	-0,000075
cosThetaPoint	0,920000	0,930000	0,930000	0,850000	0,700000
NormDecayLength	7,000000	6,500000	5,000000	5,000000	2,000000
Significance	9,679097	11,245783	8,264448	5,180366	4,146502
Sign. Error	0,121172	0,127517	0,160007	0,215305	0,246483
Purity	0,692923	0,781506	0,861897	0,964376	0,983928
Purity Error	0,013892	0,015387	0,030591	0,078073	0,115638
Global Address	0 (0)	0 (512)	0 (1024)	0 (1536)	0 (2048)

Figure 43: Output from the optimisation analysis macro depicting several parameters from the optimisation performed. The momentum bins are numbered such that 0 refers to 3-5 GeV/c up till 4, which refers to 16-36 GeV/c.

Figure 43 shows several parameters from the 60-80% optimisation. The relevant row for us is the second one, 'max significance found in bin:'. Here, we see that for all momentum bins and all three variables, the first iteration (i.e. the original cut) yielded the highest significance. This means that the improver task was unable to find any setting of the three variables for which the significance was higher than initially. To get a more intuitive feeling for this and to explore the reason why, we can take a look at some of the resulting heat maps.

The heat maps shown in Figure 44 show a distinct lack of the 'islands' of high significance seen in previous optimisations. For 12-16 GeV/c we even see that the whole grid is basically the same significance level with only a very small difference between the two regions. It seems that in the given interval, no obvious improvements could be made on this cut. It is unlikely that, also taking into account the rebinning, this cut is the 'perfect' cut for this range. Therefore it is plausible that the optimisation parameters should be set up differently if one wishes to optimise this peripheral region. For instance, more than three variables could be optimised. Furthermore, intervals with different boundary parameters should be explored according to the shape of the significances in the heat maps (for Norm Decay Length and  $\cos \theta_P$  for example, the heat maps in Figure 44 show that one should apply lower boundaries). Finally, once could run several iterations, the first one to pinpoint the region of the 'island' of highest significance and next ones to find the optima within these boundaries. For now, we will continue to use the original 60-80% cut. Since we did not change anything, it has the same mass spectra, properties and significances as we found and discussed earlier.



## Some heat maps from the optimisation of 60-80%

Figure 44: Heat maps depicting the optimal values for 60-80% as found by the optimisation fitter for  $\cos \theta_P$  and Norm Decay Length. Note that the 16-36 GeV/c bin still had too few events for a proper fit to be constructed, hence its omission.

## 5.14 Incorporating the optimised results

Now that the optimised cuts are available, we will use them for our final analysis. Essentially this means recalculating all of the steps required for the  $R_{AA}$ . Since we already analysed the mass spectra of the original and optimised mass spectra, we will not do this again. There are several other factors we need to consider. Firstly, in Table 7 we can see the new number of events for the optimised centrality regions. Furthermore, using the new results, we will calculate the improved efficiencies and corrected yields and see how the final nuclear modification factor looks.

Centrality	10-20%	20-30%
# events	$1.93601 \cdot 10^{7}$	$1.51197 \cdot 10^{7}$

Table 7: Number of events for optimised centrality regions. Note that the numbers have slightly increased compared to the unoptimised results in Table 3.

### 5.14.1 Reconstruction efficiencies of optimised cuts

To calculate the efficiencies we run the new cuts over the specialised efficiency task on the Grid. For 10-20% we use the simulation data from LHC19c3a\_q and LHC19c3a\_r, while for 20-30% we use LHC19c3b\_q and LHC19c3b\_r. Since we have made our cuts in our desired binning this time, we do not have to rebin everything again here (which is the time saving indicated a few sections back). Proceeding through the same steps as before we come to the following efficiency plots for both cuts for RecoPID/MCLimAcc. These efficiencies are of the same order as the unoptimised ones, which is unsurprising as the optimisation only slightly tweaked the variables of the cut. Figure 46 shows the ratios of 30-50% for all efficiencies, including the new optimised 10-20% and 20-30% efficiencies. When we compare Figure 46 to the ratio plots of the unoptimised efficiencies, one can see that both 10-20% and 20-30% have smaller efficiencies than before. Notably, 20-30% now sits below 30-50% across the full momentum range.



Figure 45: New efficiency plots for 10-20% and 20-30% after optimisation.



Efficiencies as a ratio of 3050 - beauty

Figure 46: New efficiency ratios after optimisation. Like before, some points in 50-60% and 10-20% are difficult to see due to them overlapping with other efficiencies very closely; their precise values can also be seen in the appendix.

### 5.14.2 Corrected yields and final nuclear modification factors

Using the newly acquired raw yields and efficiencies from the optimised cuts, we can again compute  $R_{AA}$  versus  $\langle N_{part} \rangle$ . To do this we first calculate the new corrected yields and cross sections of the optimised centralities (we already computed the others after all). This was done in close analogy to earlier, resulting in the following cross sections and corrected yields. The results depicted in Figures 47 through 50 are rather in line with the expectations based on the unoptimised cuts. Again, it would have been strange if they had been very different; in principle mostly the significance of our measurement should be affected rather than the actual outcome. Keeping this in mind we can proceed to calculate the new  $R_{AA}$  as a function of momentum using the same methods as earlier.



### Reconstructed yield and reconstructed cross section at 10-20% centrality (optimised)

Figure 47: Corrected yield at 10-20% centrality after optimisation. The red bar represents the 2015 systematics.

Figure 48: Corrected cross section at 10-20% centrality after optimisation. The green bar represents the 2015 systematics.





Figure 49: Corrected yield at 20-30% centrality after optimisation. The red bar represents the 2015 systematics.

Figure 50: Corrected cross section at 20-30% centrality after optimisation. The green bar represents the 2015 systematics.



 $R_{AA}$  vs  $p_{_{T}}$  (no Eloss hypothesis) after optimisation

Figure 51:  $R_{AA}$  as a function of  $p_{\rm T}$  for the full centrality range, after optimisation. The dotted line indicates the maximum physical value  $R_{AA}$  can take.



Figure 52:  $R_{AA}$  as a function of centrality after optimisation and without systematic errors.

Figure 51 shows the new nuclear modification factor as a function of momentum. As can be seen, our momentum bins of interest seem to still behave as predicted. Naturally there is still a data point missing at high momentum for 60-80% as we did not change that cut. The next step is to compute this new data as a function of centrality.

In Figure 52 we can more clearly see the effect of our optimisation. Although the results have shifted very slightly in 10-20% and 20-30%, this is not a significant change within the statistical uncertainties. However, we do see that the statistical errors of this measurement are slightly smaller in the centrality regions we optimised for. Although it is a shame that 60-80% was not successfully optimised, we can see the effects of a successful optimisation here and would therefore strongly suggest any future studies to carefully consider the optimisation when performing this analysis, especially for more peripheral centrality regions. There might still be some significance to gain in this data that way.

According to an earlier study [47], the results given by these calculations give an underestimation of the actual  $R_{AA}$ . This correction was 5% for 0-10% and 7% for 30-50%, which arose due to a distortion of the electric field in the TPC that resulted in a lower track reconstruction efficiency. Currently, researchers are reworking the 2018 data and the 'pass 3' sets are expected to no longer require this correction. These sets are however not yet available. Although the appropriate thing to do would be to recalculate this bias for the all centralities separately, due to time constraints we could not do this. Therefore we had to assume that this bias evolves linearly across an increasing centrality, leading to the corrections in Table 8. Finally, we can calculate our definitive results (before systematic errors that is). In Figure 53 we see our results. Again, we used data from Table 4 to convert centrality to number of participants.

As expected, we still see that for fewer participants, less hot and dense QGP is generated. We must not forget that these results are not yet complete. Without systematic error bars, these results are essentially educated guesses. Therefore, the next step is to calculate the systematic uncertainties (*systematics*) of these measurements.

Centrality	0-10%	10-20%	20-30%	30-50%	50-60%	60-80%
Correction	+5%	+5.7%	+6.3%	+7%	+7.7%	+8.3%

Table 8: Correction applied to final results.



Figure 53:  $R_{AA}$  as a function of  $\langle N_{part} \rangle$  after optimisation and without systematic errors.

## 5.15 Systematic uncertainties

Up till now, our calculations have incorporated an uncertainty based on statistical errors. This error is represented through for instance the thin lines in Figure 53. Although they definitely play a role in the accuracy of the results, they do not tell us everything. After all, any bias in the detectors is not included in this error, as there is generally no way of deducing this from the regular detector output. This is why we have to perform some calculations to get an estimation of the systematic uncertainties of the study. We can then add these to our final plot and that will give some weight to the results. If it turns out that these uncertainties are for example very large peripherally, then it will be more difficult to conclude anything concerning the behaviour of the  $R_{AA}$  in this region.

The systematics we will consider require considerable calculations. Essentially we have to perform the entire calculation again for several variations of our cut files to analyse some of the biases incorporated in our results. There is a problem with this however. The systematic analyses have to be thorough, but since this thesis has a strict time limit, it was not possible to fully calculate all systematics. For instance, as we will discuss in more detail later, the cut variation systematic alone would technically require sending 288 separate wagons to the Grid across data and MC, if one were to do it fully for all centralities and the right number of variation variables. Even if the COVID-19 crisis, which restricted some of the work we could do, had not happened, this would have been an impossible task for this bachelor thesis (and the train operators) within the given time frame. Therefore, we have had to cut corners in several places by taking for instance results from earlier studies or by approximating some results by comparing similar centrality bins. Note however that all these choices were made sensibly and therefore the given systematic uncertainties are representative of what the 'true' results would be. We will always describe the method in which the systematics should be calculated and then what was done for this particular study.

For this study we will consider the following systematics: yield extraction, PID, MC  $p_T$  shape, B feed-down, cut variation, track reconstruction efficiency, normalisation and branching ratio.

### 5.15.1 Yield extraction systematic

The yield extraction systematic takes into account some of the 'randomness' of our fit to the data. This is done by changing several aspects of the fit and then randomly displacing the fits very slightly through a Poisson process in a so-called multitrial yield extraction macro. In total just over ten thousand such variations are run for every centrality and every momentum bin for 5 different versions of the fit. Through these random variations we can analyse the stability of our particular choice of parameters. If the chosen fit is very unstable, these results will vary greatly. If it is very stable, there will be hardly any change.

We will look at five fit variations and one additional effect for this systematic. They are:

- Standard fit. Here, we use the normal fit parameters and apply the randomisation without extra changes.
- Alternative background. Rather than using a power and exponent background function, we use just a power function.
- 160. For this variation we set the upper limit of the mass for the fit to  $0.160 \,\text{GeV}/\text{c}^2$  rather than the default  $0.155 \,\text{GeV}/\text{c}^2$ .
- 165. Similarly to the previous one, we set the upper limit of the mass for the to  $0.165 \,\text{GeV}/\text{c}^2$  rather than the default  $0.155 \,\text{GeV}/\text{c}^2$ .
- Rebinned. This time we change the binning parameters of the fit.
- Residual bin counting. This is not a variation, but rather an effect we calculate for the standard fit. In principle it tells us how much signal is lost on the boundaries of bins due to the choice of binning.

For the variations, we will run each of them a couple thousand times through the Poisson process and calculate the yield. This will generate 5 distinct Poisson distributions. Additionally, we take the residual bin counting of the standard fit. This is then combined with the 5 means of the Poisson distributions in the yield histograms, which we take as our final systematic by calculating the RMS. To summarize: we run the variations for yield, we run the residual for the standard fit, combine that with the means variations of the yield and calculate the RMS to come to our systematic. Naturally this generates many different graphs, but not all were included as they are all roughly similar. Rather, we will show the 30-50% systematic at 3-5 GeV/c as an example and discuss any odd results where they arise for other centralities.

In Figure 54 we see an example of the systematic yield extraction. In this example, the cuts behave as expected. The randomisation is clearly visible through the 'smearing' of the results. Furthermore we can already somewhat see the effects of the fit parameter changes. However, we need to construct the yields to analyse this properly. Figure 55 shows all different yields and how often they occur for the same yield extraction. These are again clearly random distributions but their means are what we are interested in. We see that there is no huge difference between the peaks but that the alternative background has the largest impact. The means all lie within the statistical acceptances of the other Poisson peaks. Therefore, we take their means for our systematic in 30-50% at 3-5 GeV/c.



Figure 54: Example of systematic yield extraction for 30-50% centrality at 3-5 GeV/c. On the left we see the sigmas of all fits, sorted by trial number. On the right we see the masses sorted by trial number.



Figure 55: Another example from the 30-50% yield extraction at 3-5 GeV/c. Here we display the number of trials with particular yield values. The RMS of the means of this plot combined with the residual forms the final systematic. Note the distinctive Poisson distributions.



Figure 56: Residual plot from the 30-50% yield extraction at 3-5 GeV/c.

The residual bin counting results of the 30-50% centrality systematic at at 3-5 GeV/c are displayed in Figure 56. If there were no effect, this Poisson distribution would have a mean of 0. This is clearly not the case; hence we take the mean of this peak as another factor in our final calculation for this particular systematic. The process we have just described was repeated twice per centrality region (for both momentum bins). This resulted in 12 different RMS values. These values and their components can be seen in Figure 57.



Figure 57: Results of systematic yield extraction. The RMS is the value we take as the systematic for every centrality bin for both momenta.

As we can see, the systematic seems to be largest at low momentum, which is reasonable as there is more background there, causing loss of fit stability, especially for more central collisions. We also see that the 0-10% bin of the 8-12 GeV/c range is exceptionally stable. This can be attributed to the combination of high multiplicity from the centrality setting and lower background due to the momentum range. One might notice that for 3-5 GeV/c, we did not include the alternative background variation at 0-10% centrality. There is in fact a reason for this which we shall now clarify.



Figure 58: Yield peaks of 3-5 GeV/c range at 0-10% centrality. Note the discrepancy of the alternative background variation in red.

Initially, this particular multitrial yield extraction had almost no entries for the alternative background variation. As can be seen in Figure 58, the peak of the alternative background variation is much lower and shifted to the right. Clearly, something is going wrong. At a first glance, it seemed that the fitter was fitting far wider peaks than allowed by the multitrial macro. This meant that not many fits were taken in the above diagram due to the upper limit of peak width, but the ones that were allowed registered a very high yield. This indicates wide peaks since then a lot of background is interpreted as signal, increasing the measured yield. Naturally, before we can conclude anything, several fixes have to be tried. First we tried allowing larger values of  $\chi^2$  and peak width since these are limited by the multitrial macro, but this did not solve the issue as it just increased the multiplicity of the alternative background peak but left it in the same place. Fixing the peak width to values from the original variation did not improve the situation significantly either. Finally, we investigated the mass spectra from this fit to see what was happening. Upon investigation it became clear that the alternative background function used was not appropriate for this fit.

From the two spectra in Figure 59, we can see the problem. The fit was taking a large amount of background noise as signal, widening the peak and distorting the result. The difference might seem subtle and it is in a way, but appearances deceive; a quick look at the difference in yield provides the evidence we need. The 'thick tail' of the signal peak for the power function background was causing an overestimation of the amount of signal in the data. This was something which had been an issue in the past as well for studies of this decay channel; power functions just don't fit background well enough at low momenta and central events. One might wonder why the other centralities and momenta did not give this problem. However, one should consider that the amount of background in this bin is particularly high, making the fitter highly sensitive to the background function choice.

Based on the nonphysical fit assigned by the alternative background function, we decided to discard that variation for this bin. Note that the bin still gets useful systematic; this bin still incorporates the other 8000 variations and the residual bin counting. The numerical results of the systematic yield variation are displayed in Tables 9 and 10. Note that we round off our RMSs to whole percentages as this procedure does not provide more accountable accuracy than this.



Mass spectrum 0-10% at 3-5 GeV/c, power background function

Mass spectrum 0-10% at 3-5 GeV/c, original background function



Figure 59: Mass spectra with both background types. At first the spectra look the same but upon closer inspection we can see that the peak with the power background is a lot wider and includes entries which are clearly not part of the Gaussian signal peak. Note the considerable differences in signal, width and background. Interestingly, the alternative background function provides a higher significance, once again proving that high significance is not everything.

Yield extraction	0-10%	10-20%	20-30%	30-50%	50-60%	60-80%
$3-5  \mathrm{GeV/c}$	7%	3%	5%	2%	2%	2%
$8-12  \mathrm{GeV/c}$	1%	4%	2%	2%	1%	1%

Table 9: Systematics from yield extraction assigned to our final results.

3-5 GeV/c	0-10	10-20	20-30	30-50	50-60	60-80
Original signal	11423.200	1401.300	1372.280	3867.48	319.320	151.90100
0 std fit	11465.200	1432.080	1398.150	3864.87	322.081	155.11100
1 alt bkg	x	1492.820	1457.840	4011.58	326.905	157.48300
2 160	11048.900	1415.700	1348.920	3855.78	325.273	155.39900
3 165	10331.500	1351.020	1270.350	3826.83	321.151	156.10000
4 rebin	11423.300	1436.430	1397.400	3910.55	321.562	155.24500
0 std fit ratio	0.003671	0.021970	0.018850	-0.000676	0.008645	0.02113
1 alt bkg ratio	x	0.065310	0.062340	0.037258	0.023752	0.03675
2 160 ratio	-0.032772	0.010270	-0.017023	-0.003025	0.018641	0.023029
3 165 ratio	-0.095569	-0.035884	-0.074282	-0.010513	0.005734	0.02764
4 rebin ratio	0.000005	0.025070	0.018305	0.011137	0.007020	0.02201
Residual binCounting abs	0.11300	0.01849	0.06637	0.04514	0.02287	0.00739
RMS (total)	0.067809	0.034428	0.049648	0.024732	0.016287	0.02460
%	7	3	5	2	2	2
8-12 GeV/c	0-10	10-20	20-30	30-50	50-60	60-80
Original signal	4721.140	781.010	510.039	1859.96	193.144	89.03520
0 std fit	4712.430	779.202	512.436	1856.63	195.353	88.56060
1 alt bkg	4792.970	788.498	506.106	1881.57	194.216	89.68630
2 160	4723.730	779.793	516.246	1872.94	194.456	90.10380
3 165	4679.100	768.695	504.890	1863.36	195.407	87.94110
4 rebin	4718.220	782.555	510.288	1842.82	196.353	88.95350
0 std fit ratio	-0.001869	-0.002315	0.004700	-0.001793	0.011438	-0.00533
1 alt bkg ratio	0.015215	0.009590	-0.007711	0.011614	0.005549	0.00731
2 160 ratio	0.000548	-0.001558	0.012171	0.006974	0.006791	0.01200
3 165 ratio	-0.008903	-0.015768	-0.010095	0.001824	0.011719	-0.01229
4 rebin ratio	-0.000626	0.001980	0.000490	-0.009217	0.016616	-0.00092
Residual binCounting abs	0.02383	0.09178	0.04662	0.03517	0.00177	0.01216
RMS (total)	0.012130	0.038245	0.020434	0.015874	0.010200	0.00936
%	1	4	2	2	1	1

Table 10: Full numerical results of the systematic yield extraction.

### 5.15.2 PID systematic

As we discussed during earlier in this thesis, the TPC is the main detector used for particle identification within the ALICE detector. The PID technique used might be prone to a systematic error. In order to account for this systematic in the TPC, an attempt was made to calculate the effect of changing the sensitivity of the TPC. However, this analysis did not produce correct results and unfortunately there was not enough time left for us to investigate this further. We of course still need a PID systematic, which we will therefore source from an earlier analysis [47]. This study analysed pure samples of pions and kaons traveling through the TPC and TOF. From this, they discovered a bias caused by an imperfect calibration for the TPC in recognising the signature energy loss of hadron species and calculated a correction. The final results from this study were that for  $D^*$  the PID systematics are 1% for 0-10% centrality and zero elsewhere, Thus, we assign a 1% systematic to 0-10% and 10-20%, since they use similar cut files, and no PID systematic elsewhere.

PID	0-10%	10-20%	20-30%	30-50%	50-60%	60-80%
3-5 GeV/c	1%	1%	0%	0%	0%	0%
8-12 GeV/c	1%	1%	0%	0%	0%	0%

Table 11: Systematics from PID assigned to our final results.

## 5.15.3 MC $p_T$ shape systematic

While generating the MC simulations needed for the corrections of the raw yield, we are forced to make an assumption on shape of the transverse momentum of the  $D^*$  meson, which creates a systematic. We source this systematic from an earlier study [47]. In 0-10% centrality, the systematic is 1% for 3-3.5 GeV/c, 0.5% for 3.5-4 GeV/c and zero for other momenta. For 30-50%, the results are that the systematic is zero above 3 GeV/c. We again apply 0-10% to 10-20% and 30-50% elsewhere. This gives the MC  $p_{\rm T}$  shape systematics displayed in Table 12.

MC $p_{\rm T}$ shape	0-10%	10-20%	20-30%	30-50%	50-60%	60-80%
3-5 GeV/c	1%	1%	0%	0%	0%	0%
8-12  GeV/c	0%	0%	0%	0%	0%	0%

Table 12: Systematics from MC  $p_{\rm T}$  shape assigned to our final results.

### 5.15.4 B feed-down systematic

As stated before in this analysis, we use efficiencies to separate the yield generated from the decay of beauty quarks from those of  $D^*$  particles generated via fragmentation of charm quarks produced directly during the hard scattering. However, the final corrected yield takes some other calculations into account as well. In particular, we also implicitly use beauty-production cross sections from the FONLL model and some properties of the beauty decay to  $D^*$  from other simulations. The fraction of the yield from prompt charm (also called the fraction of prompt  $D^*$  meson) is given by the following equation [47]:

$$f_{charm} = 1 - \left( N_{raw}^{Dfeed-down} / N_{raw}^{D} \right), \tag{5.2}$$

where  $N_{raw}^D$  indicates the raw yield from prompt  $D^*$  and  $N_{raw}^{Dfeed-down}$  the  $D^*$  yield from B decays. The expected feed-down fraction of  $D^* \frac{dN}{dp_T}$  is obtained by taking  $f_{feed-down} = 1 - f_{charm}$ . These  $D^*$  from B have a different suppression; beauty quarks are expected to be less suppressed than charm quarks. When calculating the  $R_{AA}$ , this suppression of feed-down  $D^*$  has to be assumed. This assumed suppression is combined with the expected feed-down fraction to yield the systematic. This systematic accounts for the expected change of the prompt  $D^* R_{AA}$  due to this feed-down fraction and is depicted in Figure 60.



Figure 60: Charm fractions with systematic from feed-down. Note that the line through the points starts in the origin, which is not shown in these plots. Furthermore, the line drops to the x axis in 60-80% because the last momentum bin is empty.

B feed-down	0-10%	10-20%	20-30%	30-50%	50-60%	60-80%
$3-5 \mathrm{GeV/c}$	+4.7% $-5.9%$	+4.9% $-6.1%$	+4.2% $-5.3%$	+3.8% $-4.8%$	+3.7% $-4.7%$	+3.7% $-4.6%$
8-12  GeV/c	+11.9% $-15.7%$	+10.6% $-13.3%$	+7.9% $-10.4%$	+5.1% $-6.7%$	+3.8% $-5%$	+4% -5.2%

From Figure 60, we can infer the systematics from feed-down. Note that these are asymmetric, unlike our earlier systematics. This is why we make a distinction between positive and negative for this result.

Table 13: Systematics from B feed-down assigned to our final results.

### 5.15.5 Cut variation systematic

The next systematic we discuss is the cut variation systematic. This systematic takes into account the stability of our cut file and cutting parameters. If our cut is very unstable, the yield will vary greatly for small changes to the cut. If these changes are too large, it might indicate that the results from our cut are an outlier, rather than an accurate measurement. To take this into account, we will analyse the stability of our cuts and see whether this is the case.

The cut variation was started for this thesis, but due to time constraints we were unable to complete it. A problem was found in the efficiencies of this study, but we could not identify or address it in time. Therefore, we will use the cut variation systematic found by the earlier study we have used before [47]. However, since a great deal of the cut variation was completed, we will show the progress made with it in this section. In the outlook we will discuss what the next steps would be if we had more time.

For the cut variation, the strategy was as follows. For the centrality ranges 0-10%, 30-50% and 60-80%, we take the original cut file and make four variations<sup>9</sup>. Each variation takes four cutting variables and varies them in some way. In our case, we chose to make the variations of 10% looser, 20% looser, 10% tighter and 20% tighter in 3-5 GeV/c and 8-12 GeV/c. The variables in question are 'd0d0', 'cos  $\theta_P$ ', 'NormDecayLength' and ' $p_{Tkaon}$ '. They were chosen as such for several reasons. Firstly, we chose four variables because this prevents us from being strongly affected by the unique characteristics of one variable, which is not something we desire since we are looking for the volatility of our overall cut. Secondly, the first three variables are the variables we optimised for, so presumably they have some measurable effect on the yield. The last variable was not optimised for, but we chose this one as well because it is not a linear combination of the other variables. Finally, the variables we chose are known as relatively sensitive variables, which means that they will have a significant impact on the yield and are thus appropriate to choose as cut variation variables.

Using this cutting strategy, 24 new cut files were generated (4 per centrality range). For these cuts, several analyses were started. In total, 12 wagons were launched in data to generate mass spectra for all these cuts and 24 wagons were launched in MC to calculate their efficiencies. The data sets used were the same as for the initial study; for data we used LHC18q\_pass1 and LHC18r\_pass1 and for MC LHC19c3a\_q, LHC19c3a\_r (for 0-10% and 10-20%), LHC19c3b\_q, LHC19c3b\_r (for 20-30% and 30-50%) and LHC16i2c\_AOD198 (for 50-60% and 60-80%). Using these results, we reconstructed the prompt yield and cross sections. For the new mass fits, we chose to fix the peak width of the signal peaks to the MC values from the initial study to prevent too much statistical fluctuations in the fitter from affecting our final systematic. The respective RMSs of the ratios of the yields for 3-5 GeV/c and 8-12 GeV/c were calculated and intended for use as cut variation systematics. Like earlier we would assign the 0-10% results to 10-20% as well and 30-50% to 20-30% and 50-60% too, to cover the full range. However, we ran into problems when comparing the results. In order to illustrate this, we will take a look at some of the graphs generated by this analysis.

 $<sup>^{9}</sup>$ Naturally, for full accuracy, this should be done for *every* centrality range separately and we did in fact prepare cuts for those as well. However, performing these analyses too would have doubled the number of wagons and due to the high memory profile of Pb-Pb jobs combined with generally busy trains and the COVID-19 crisis, we were unable to run these as well within our given time frame.



Figure 61: Results of cut variation analysis for all analysed centralities. Left: corrected yields calculated with efficiencies and raw yields for all variations and original cut. Right: ratio plots of the corrected yields with the original yield. The RMS values displayed are the systematics we would have assigned for cut variation.

In Figure 61, we can see the corrected yields of the cut variations and the ratio plots with the original yield. Although they might seem reasonable at first, they are not quite what we were expecting. Based on earlier studies, we expected these values to be around 10%, but they are all larger than this, in particular in the 3-5 GeV/c bin. This systematic was found to be broken in some way, and several analyses were performed to diagnose the problem. First, we tried fixing the peak width to the ones from the original data sample rather than MC, which actually worsened the results. Not fixing  $\sigma$  at all also granted no improvement. Then, we decided to investigate the mass spectra to see whether they were somewhat logical. We will show the mass spectra for 0-10% here; the others are in the appendix.

Figure 62 shows very neat-looking mass spectra. We can clearly see the effect of making cuts tighter and looser. For looser cuts, less background is filtered. This can be seen in the size of the peaks, but more quantitatively through looking at the S/B ratio. It rises for increasing tightness, which makes sense, since we cut out more background for tighter cuts. It seems that for these fits though the significance is still very high, despite the cut variation. Combined with the physical interpretation of the peaks as being accurate, this means that we can conclude that the cut itself seems suitable for our purpose. One might wonder why we don't use the tighter cuts as baselines, but we must also note that we lose some signal by using tighter cuts which costs us statistics. Either way, the problem with the cut variation analysis is not a cut parameter issue; otherwise we would have seen some anomaly here. As can be seen in the appendix, 30-50% and 60-80% behave as expected as well, with the fitter working well on the given data.

In order to quantitatively check whether there were problems with the fits which could not be seen from the spectra, some additional investigations were conducted. In particular, we checked the masses and peak widths of the fits to see whether all was well. In Figures 63, 64 and 65 the results are depicted. Unfortunately no indications of the issue could be found here either. The peak widths were all correctly fixed to the MC values, which are not wildly different from the original data values. Also, the masses are all within statistical uncertainty practically the same as the original mass. Altogether we can conclude that the problem does not lie in the mass spectra or in the cut parameters. This leaves two places where the issue can be found; the corrected spectra macro and the efficiencies.

First we analyse the effects of the corrected yield macro on our results, for which we will construct some new observables. These are the raw yield comparisons (as opposed to corrected yield) and 'manual' cross sections (as opposed to corrected cross sections). The manual cross sections are calculated by dividing the raw yield by the prompt efficiency. Creating these diagrams allows us some insight in the functioning of the corrected yield macro.

Figures 66 and 67 show several effects. Firstly, we see that the raw yields display the same tighter and looser variations we expect. While the spreads are different from what we found through the corrected calculations, this is in principle in line with expectations as this manual cross section does not compensate for feed-down effects. Furthermore, we see that the cross sections seem to be shifted by the corrected yield macro, by comparing Figure 67 with the corrected cross sections shown in the appendix. This shift is unexpected and indicates some problem with the corrected yield calculation macro which is influencing our systematics analysis which has not yet been diagnosed precisely. What we also clearly see from the cross sections is that they do no longer vary nicely along the original but seem to have strong shifts away from the original, which is not at all what should be happening physically. This strongly indicates an issue with the efficiencies on top of an issue with the corrected yield macro.



Mass spectra from cut variation 0-10%

Figure 62: Mass spectra from 0-10% cut variation. Please refer to the appendix for the 30-50% and 60-80% variations.



Figure 63: Mass spectra analysis for the cut variation of 0-10%. The flux in the 20% tight cuts for unchanged bins is most likely due to statistical fluctuations, but we do not use these bins anyway. Note that all variation peak widths are fixed to the exactly the same value.



Figure 64: Mass spectra analysis for the cut variation of 30-50%. The flux in the 20% tight cuts for unchanged bins is most likely due to statistical fluctuations, but we do not use these bins anyway. Note that all variation peak widths are fixed to the exactly the same value.



Figure 65: Mass spectra analysis for the cut variation of 60-80%. The flux in the 20% tight cuts for unchanged bins is most likely due to statistical fluctuations, but we do not use these bins anyway. Note that all variation peak widths are fixed to the exactly the same value.


Figure 66: Comparisons of raw yield. These can be cross-referenced with the RMSs as found through the actual corrected yield calculations in Figure 61.



Figure 67: Comparisons of manual cross sections (raw yield divided by prompt efficiency). These can be cross-referenced with the RMSs as found through the actual corrected cross section calculations in the appendix.

We will now investigate these efficiencies. In order to do this, we calculated the efficiency ratio plots of the variations and the original cut. That way we can spot any anomalies in our cut. We will depict the 0-10% result here as there seems to be an issue with it; the others can be found in the appendix.

In Figure 68 we see the efficiencies as a ratio of the original. Immediately, there is an indication that something unexpected is happening. For the looser cuts, the 8-12 GeV/c bin seems to actually yield a tighter efficiency rather than a looser one. This is highly unexpected, since we know from the mass spectra that the actual cut is in fact looser. First, we checked whether the rebinning was done correctly; it was. Then we reran the efficiency by merging some of the output files from the Grid first, but to no avail. Finally, since the efficiency is calculated by dividing the RecoPID step by the LimAcc step, we investigate how those steps actually look.



Figure 68: Efficiency ratios for 0-10% cut variation.



Figure 69: MCLimAcc and RecoPID step for 0-10% centrality. Here we display the prompt efficiency since that is used to calculate the crude corrected yield from before.

In Figure 69 we can see that the RecoPID step seems to behave as expected. However, the LimAcc step shows strange behaviour, since the shape is much more erratic than we would usually find from efficiency calculations. This indicates strongly that something is wrong with the efficiencies, either in the Grid calculations or somewhere during the data processing. It is not a rebinning normalisation problem; this is something we checked and which would besides cancel due to taking the ratio of the two plots. Although Figure 69 only depicts the 0-10% centrality, we use the same algorithms for 30-50% and 60-80% as well, which means that the problem in the efficiency calculations from 0-10% propagates to the other centralities, with no way of knowing in what way it is affecting our result exactly. Unfortunately, we ran out of time to further investigate the root of the problem. The two issues we found with the cut variation, the corrected yield calculation problem and the efficiency problem, combined with the lack of time to properly investigate them, led us to use the cut variation results from a previous study instead.

These previous results were once again sourced from the same earlier study as we sourced the other systematics from [47]. Their approach was broadly similar. However, this study varied different variables (dca, d0d0,  $p_{Tkaon}$  and  $p_{Tpion}$ ) and varied them one by one, four times per centrality and variation. The analysed centralities were 0-10% and 30-50%. We again run into the problem that the systematic varies within a bin and like before we will take the highest value to prevent inaccurate claims on our results. Using the same centrality distribution as earlier, where we use the 0-10% result for 0-10% and 10-20% and 30-50% for the rest, we come to the systematics in Table 14.

Cut variation	0-10%	10-20%	20-30%	30-50%	50-60%	60-80%
$3-5 \mathrm{GeV/c}$	12%	12%	6%	6%	6%	6%
8-12  GeV/c	8%	8%	6%	6%	6%	6%

Table 14: Systematics from cut variation assigned to our final results.

#### 5.15.6 Track reconstruction efficiency systematic

The next systematic we address is the track reconstruction efficiency. This systematic was also calculated in the earlier study we cite for most of our systematics. It in principle consists of two aspects: the tracking matching between the TPC and ITS detectors and the track-quality selections [47]. The former is caused by a difference between data and MC regarding tracking performance of the detectors. This performance is quantified through the *ITC-TPC matching efficiency*. The number of tracks reconstructed correctly in the TPC is different for *primary particles* produced in collisions and *secondary particles* produced through strong hadron decay and material interactions. Primary and secondary particles carry different 'weights' in the total matching efficiency and the first part of this systematic calculates these weights, yielding the 'true' matching efficiency. This is then compared to the matching efficiency as calculated by dividing the number of tracks fitted by the ITS and TPC through the so-called Kalman filter which also gave a hit in an SPD by the number of tracks fitted in the TPC.

The second part of this systematic, the track-quality selections, is caused by a difference between data and MC simulations of the efficiency of these selections. A small systematic was indeed observed and added to the ITS-TPC matching efficiency systematic to form the track reconstruction efficiency systematic. From this study, the values are displayed in Table 15. Since this study was also performed for 0-10% and 30-50% centrality we use the same distribution of these values as for the other systematics. Because the study was done for data sets LHC18\_q and LHC18\_r separately, we will take the RMS of the values found for these data sets for each momentum bin since we use both of them for our study. From this finer momentum binning, we take the highest value within our wider binning as our systematic like before.

Track reconstruction	0-10%	10-20%	20-30%	30-50%	50-60%	60-80%
$3-5  \mathrm{GeV/c}$	11%	11%	9.3%	9.3%	9.3%	9.3%
$8-12  \mathrm{GeV/c}$	8%	8%	8.5%	8.5%	8.5%	8.5%

Table 15: Systematics from track reconstruction efficiency assigned to our final results.

#### 5.15.7 Normalisation and branching ratio systematic

The final systematics we will discuss are the normalisation and branching ratio systematics. First we will explain the normalisation systematic. During operation, the ALICE detector measures several parameters of a collision, including for example  $\frac{dN}{dp_{\rm T}}$ . This value has to be normalised using the total inelastic cross section of the collision. However, there is an uncertainty on this as the value has to be measured by the detector as well. The uncertainty in this measurement is the normalisation systematic and its value was taken from ALICE documents. Next we discuss the branching ratio systematic. The branching ratio of a decay channel is essentially the ratio of the number of decays following that specific channel divided by those following the total number of possible decay channels. The best known value for the branching ratio of our decay channel still carries some uncertainty and this uncertainty will be taken in our systematics as well. We source the value and uncertainty level of the branching ratio from the Particle Data Group [49]. Both systematics are shown in Table 16. They will be treated separately from the other systematics as our specific study has no way of accounting for them.

Normalisation	4%
Branching ratio	1.3%

Table 16: Systematics from normalisation and branching ratio.

#### 5.15.8 Total systematic uncertainty

Using the values obtained above, we can calculate the total systematic uncertainty to be assigned to our results. They were calculated by quadratically summing over all uncertainties twice for every bin, to account for the asymmetric B feed-down systematic. The normalisation and branching ratio systematics were not taken into account for this sum as they will be added as separate uncertainties. Table 17 displays the total systematic uncertainties, excluding normalisation and branching ratio. The total systematics are also displayed graphically in Figure 70.

<b>Systematics</b> $R_{AA}$	0-10%	10-20%	20-30%	30-50%	50-60%	60-80%
3-5 GeV/c	+18% - 19%	+17% - 18%	+13% $-13%$	+12% - 12%	+12% - 12%	+12% - 12%
8-12  GeV/c	+16% - 19%	+16% - 18%	+14% - 15%	+12% -13%	+11% -12%	+11% -12%

Table 17: Systematic errors of this measurement of  $R_{AA}$  versus  $\langle N_{part} \rangle$ .



Figure 70: Systematic errors of this measurement of  $R_{AA}$  versus  $\langle N_{part} \rangle$  represented graphically. Note that the systematic error is larger for more central events.

### 6 Results and conclusion

Using the data from the previous section and the corresponding systematic errors, we are finally in a position to calculate our final  $R_{AA}$  versus  $\langle N_{part} \rangle$ . The measurement is compared to a theoretical model prediction [50].



Figure 71:  $R_{AA}$  as a function of  $\langle N_{part} \rangle$  for 3-5 GeV/c and 8-12 GeV/c, including systematic errors. The thin lines through the points indicate the statistical errors. Note that the branching ratio and normalisation systematic were taken as a percentage of unity and applied globally for all data points.

In Figure 71 we see the final results from our analysis. As we calculated, the relative systematic error is larger for 3-5 GeV/c than 8-12 GeV/c, especially in the central region. Our 8-12 GeV/c results are somewhat similar to the 6-12 GeV/c  $D^*$  values from the 2012 study seen in Figure 13, our original hypothesis, although we measured more charm suppression than they did. This could be explained by their different  $p_{\rm T}$  range; this wider range includes lower momentum  $D^*$  mesons than ours. However, it is difficult to conclude this because these deviations could be caused by statistical fluctuations as well. Our peripheral events, where these effects are smaller, still measure below the 2012 values, but the error bars do comfortably overlap. All together, we have determined the shape of  $R_{AA}$  versus  $\langle N_{part} \rangle$  and we measure more charm suppression and thus hotter and denser QGP for increasingly central events.

In Figure 72 we see our 8-12 GeV/c measurements compared to a model of the nuclear modification factor's behaviour for different values of  $\langle N_{part} \rangle$  [50]. We use our 8-12 GeV/c results for comparison to a model of the shape of  $R_{AA}$ . The model in question predicts a relation between  $R_{AA}$  and  $\langle N_{part} \rangle$  for both central and peripheral events. The model was built to strongly incorporate in-medium (kinetic) energy loss of hadrons while simplifying the evolution of QGP. The researchers argued that there was "robust agreement" between their model and the  $\sqrt{s_{NN}} = 2.76$  TeV Pb-Pb  $D^*$  results available at the time.



Figure 72:  $R_{AA}$  as a function of  $\langle N_{part} \rangle$  for 8-12 GeV/c, including systematic errors and theoretical predictions [50]. The thin lines through the points indicate the statistical errors. The branching ratio and normalisation systematic were taken as a percentage of unity and applied globally for all data points.

When we compare our results to this charm suppression model, we see that the shape of the predictions seems to be in accordance with our results. For peripheral events we see a slight outlier in 50-60%, however, the position of the two surrounding points within the predictions and the large error bars indicate that it is most likely a statistical fluctuation. The 20-30% point overlaps with the error bar of the model, meaning that the predictions are in agreement with data here as well. For 0-10% and 10-20%, we calculate the deviation from the model to properly determine its accuracy, since it seems at first glance that we measure more charm suppression than the theory would predict. In order to quantify this deviation, we take the uncertainties in our own measurements as a basis. Then, we approximate  $\sigma_{total}$  of our data points by quadratically summing the statistical and systematic uncertainties.<sup>10</sup> For 0-10% we obtain a value of  $\sigma_{total}^{0-10\%} \approx 0.025$ , while for 10-20% we have  $\sigma_{total}^{10-20\%} \approx 0.039$ . Then, we calculate the distance expressed in  $\sigma_{total}$  between the model and both data points. As it turns out, the 0-10% data point lies within 2  $\sigma_{total}^{0-10\%}$  of the theory's minimum value and within 3  $\sigma_{total}^{0-10\%}$  of its central value. For 10-20% we also find that the data point lies within 2  $\sigma_{total}^{10-20\%}$  of the predicted minimum and 3  $\sigma_{total}^{10-20\%}$  of the predicted center. In general, we would take more than 5  $\sigma_{total}$ deviation as 'true' deviation and less than 3  $\sigma_{total}$  as statistical fluctuation. This means that in principle the theory seems to be in agreement with data in 0-10% and 10-20%, since they both lie within 3  $\sigma_{total}$  of the model's centre. Thus, this model was found to be an accurate portrayal of the nuclear modification factor for the full centrality range within the statistical and systematic uncertainties of this measurement.

 $<sup>^{10}</sup>$ Note that the statistic uncertainties are actually asymmetric. However, since the points in question lie below the theoretical values, we take the upper uncertainties for our calculation.

## 7 Discussion and outlook

This study consisted of many small parts, the outcome of each of which was vital for the final result. This is why during the description of the analysis method, intermediate results were discussed as they presented themselves. During these discussions we occasionally mentioned that several phenomena could or should be investigated further. In this chapter, we will succinctly recall these and discuss how a further study could improve on the results, without repeating ourselves too much with the specific anomalies we mentioned during the analysis already. Furthermore, we will discuss new possibilities for research in the broader context of  $R_{AA}$  versus centrality measurements.

#### 7.1 Discussion and outlook of this analysis

The first item we will mention are the topological selections (cuts). The cuts are very important to reject combinatorial background and form a crucial part of any study at ALICE. In this study, we based large parts of our measurements on previously optimised cuts and we optimised two of them for use in other centralities, while leaving the other cuts in their unoptimised state. A clear way in which we could improve this study is by spending more time optimising the cuts. One could optimise for more than three cutting variables and investigate which variables are the most useful for optimisation. Furthermore, this study did not actually (successfully) optimise the 60-80% cut; this is something a future study could spend more time on, since for the peripheral data points the statistical uncertainty could probably be reduced significantly. Additionally, we could not obtain a reliable mass fit on the 16-36 GeV/c  $p_{\rm T}$  bin in the 60-80% centrality range. Careful optimisation of the selections might solve this problem in a new study. Another choice we made concerning cuts is the momentum and centrality binning. If one had another six months time, a finer centrality binning could be chosen. It would be relatively easy to rebin the momentum if desired. For instance, one might simply make two or three momentum bins of interest and only create cuts for them. This could make the process for centrality-dependent studies a lot easier and thus less prone to errors. However, rebinning also affects the effectiveness of the cut, since a much broader binning makes a cut crude. Thus, upon serious momentum rebinning, one should optimise every cut file again, even the 0-10% and 30-50% cuts. Furthermore, the most effective way to decrease the work load of these projects is by implementing centrality-dependent macro's in AliPhysics, which let us set momentum rather than centrality. This would be the best way to streamline any future investigation of centrality-dependent phenomena. Implementing these macros and data infrastructures in AliPhysics might allow single wagons to be run over all centralities, which for example would have reduced the computational workload of this project by roughly 60%.

Next, we can discuss the invariant mass spectra. When analysing mass spectra, choices have to be made on what makes a 'good fit' and what doesn't. A new study making these calls would add to the legitimacy of these results. Furthermore, in this study, the differences between data and MC were sometimes larger than we expected. A future study could spend more time investigating the  $D^{*+}$  signal properties in comparison to MC and find out what might be causing these discrepancies. We also saw in the yield extraction and cut variation that some configurations of the mass fitter might be more appropriate than others; this could also be varied and studied further to see if other fitting parameters are more appropriate for use as defaults.

Some additional work could be carried out on the reconstruction efficiencies in order to improve the solidity of the measurement. In this study there was a fundamental problem that the 2018 MC simulations for peripheral events were not yet available. Therefore, the measurements in 50-60% and 60-80% centrality were not corrected with the proper MC simulation. While using the 2015 MC simulations was sufficient for a first study, writing a final paper on this will require a new set of simulations anchored to the 2018 peripheral data.

Lastly, we will discuss the corrected spectra and  $R_{AA}$  calculation. One of the main open points left by the measurement presented in this thesis is the correct calculation of the systematic uncertainties. Firstly, the principle assumption of applying the 0-10% to 10-20% as well and 30-50% to 20-30%, 50-60% and 60-80% naturally should not be done in a more in-depth study; rather, every centrality region should be systematically analysed. Furthermore, while the systematic from yield extraction as well as the one from B feed-down were properly calculated as presented in this manuscript, the other sources of systematic uncertainty require additional work. In particular, as was made clear during the systematic uncertainty analysis of this thesis, the two sources of systematic uncertainty that require additional work are the cut variation and PID. In both cases, evidence was found of some tension on the calculation of the reconstruction efficiencies. Ideally, this is the point were we should restart the work to make this measurement ready for publication.

#### 7.2 Broader context of $R_{AA}$ measurements

Naturally, particle physics is a very active research field and therefore developing constantly. Within the context of this particular study, the peripheral MC simulations anchored to 2018 data will allow for more appropriate calculations in the 50-80% centrality region. Beyond the scope of this project, there are many processes at work which will enhance our ability to investigate these and other phenomena at the ALICE detector. An obvious change is the ongoing upgrade of the ALICE detector. New modules are being added and the ITS is undergoing an upgrade, which will dramatically increase the statistics generated from collisions and the tracking resolution, allowing a far more precise insight into the behaviour of QGP. The upgrades will naturally not only benefit the nuclear modification factor studies, but all analyses done with the ALICE detector. Altogether, physicists are steadily improving our understanding of QGP and with the ongoing projects in mind it is a very exciting time to be working with ALICE at CERN.

## 8 Acknowledgements

It has been a great challenge and an absolute pleasure to work on this project. Working with the very building blocks of nature through actual data from the most famous and advanced machine used in physics has been a privilege not many young students are granted. Therefore I would like to wholeheartedly express my gratitude to everyone who has made this thesis possible. In particular, I would like to start by thanking Dr. Alessandro Grelli for accepting me as a bachelor student and helping me on countless occasions with the project, starting almost a year ago with explaining to me what I would be doing to reading the entirety of this thesis, as well as being extraordinarily involved and supportive during the entire process. Furthermore, I would like to thank Syaefudin Jaelani, my daily supervisor, for answering my many questions on the basic functions and more subtle intricacies of ALICE analyses and never losing patience with me during all of it. When the COVID-19 crisis hit, Dr. Henrique Zanoli built a virtual machine from scratch with all relevant ALICE software installed on it, saving all our bachelor projects, for which we are of course all very grateful indeed. I would also like to thank Dr. Fabrizio Grosa explicitly, the train operator who has launched all of my analysis wagons on the Grid, which must have been a lot of work considering there were 127 of them in total. Finally, I would like to thank everyone at GRASP and the HF group in particular for generally helping out during group meetings and elsewhere, as well as Olaf Massen, my friend, housemate and coincidentally fellow thesis student of Dr. Grelli, for providing help and humour during my daily activities at GRASP.

# References

- [1] Cover image. https://cds.cern.ch/record/2646381. Apr. 2020.
- [2] ALICE Collaboration website. https://home.cern/science/experiments/alice. Feb. 2020.
- Johann Rafelski Jean Letessier. Hadrons and Quark-Gluon Plasma. Cambridge University Press. 9780521385367. 2002.
- [4] Robert L. Mills Chen N. Yang. Conservation of Isotopic Spin and Isotopic Gauge Invariance. Phys. Rev. 96, 191. 1954.
- [5] Sheldon L. Glashow. Partial-symmetries of weak interactions. Nuclear Physics 22 (4), 579–588. 1961.
- [6] Steven Weinberg. A Model of Leptons. Phys. Rev. Lett. 19, 1264. 1967.
- [7] Abdus Salam Nils Svartholm. Elementary Particle Theory, Relativistic Groups and Analyticity (Eighth Nobel Symposium). John Wiley Sons Inc. 9780470838426. 1969.
- [8] Nobel Prize in Physics 1999. https://www.nobelprize.org/prizes/physics/1999/summary/. Feb. 2020.
- [9] Dieter Haidt. The discovery of the weak neutral currents. CERN Courier. 2004.
- [10] Frank Wilczek David J. Gross. Ultraviolet Behavior of Non-Abelian Gauge Theories. Phys. Rev. Lett. 30, 1343. 9780470838426. 1973.
- [11] National Research Council (U.S.) Elementary-Particle Physics : Revealing the Secrets of Energy and Matter. The National Academies Press. 9780309174169. 1998.
- [12] University of Zurich. https://www.physik.uzh.ch/en/researcharea/lhcb/outreach/StandardModel.html. Mar. 2020.
- [13] Donald H. Perkins. Introduction to High Energy Physics, second edition. Cambridge University Press. 9780521621960. 2001.
- [14] Michel Le Bellac. Quantum Physics. Cambridge University Press. 9781107602762. 2007.
- [15] M. Tanabashi et al. (Particle Data Group). Review of particle physics. Phys. Rev. D 98, 030001. 2019.
- [16] Alberto Caliva. Low-mass dielectron measurement in Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV with ALICE at the LHC. Utrecht University. 2017.
- [17] D.A. Greenwood W.N. Cottingham. An Introduction to Nuclear Physics. Cambridge University Press. 0511040466. 2000.
- [18] Siegried Bethke. World Summary of  $\alpha_S$  (2012. Nucl. Phys. Proc. Supll.234, 229. arXiv:1210.0325 [hep-ex]. 2013.
- [19] Dominic J. Schwarz. The first second of the Universe. Annalen Phys. 12, 220-270. arXiv:astro-ph/0303574. 2003.
- [20] Viatcehslav Mukhanov. Physical Foundations of Cosmology, second edition. Cambridge University Press. 9780511200991. 2002.
- [21] George M. Fuller Kevork N. Abazajian. Bulk QCD thermodynamics and sterile neutrino dark matter. Phys. Rev. D 66, 023526. arXiv:astro-ph/0204293. 2002.
- [22] Marcelo J. Reboucas Cé Vasconcellos Valdir B. Bezerra. Astronomy and Relativistic Astrophysics : New Phenomena and New States of Matter in the Universe. World Scientific Publishing Co Pte Ltd. 9789814304887. 2010.
- [23] D. F. Kurdgelaidze D. D. Ivanenko. Hypothesis concerning quark stars. Astrophysics. 1 (4), 251-252. 1965.
- [24] QGP phase diagram 2011. https://cds.cern.ch/record/2025215. Mar. 2020.
- [25] CERN accelerator complex in 2019. https://cds.cern.ch/record/2684277. Mar. 2020.
- [26] LHC facts and figures. https://home.cern/sites/home.web.cern.ch/files/2018-07/factsandfigures-en\_0.pdf. Mar. 2020.

- [27] Municipal Energy Consumption. https://klimaatmonitor.databank.nl/Jive?report=energiegebruikinp\_geo=gemeente\_344. Mar. 2020.
- [28] LHC brochure. https://home.cern/sites/home.web.cern.ch/files/2018-07/CERN-Brochure-2017-002-Eng\_0.pdf. Mar. 2020.
- [29] J. Schukraft C. Fabjan. The story of ALICE: Building the dedicated heavy ion detector at LHC. 'The Large Hadron Collider: A marvel technology', EPFL-Press Lausanne, chapter 5.4. arXiv:1101.1257 [physics.ins-det]. 2009.
- [30] ALICE upgrades during LS2. https://home.cern/news/news/experiments/upgrading-alice-whats-storenext-two-years. Mar. 2020.
- [31] ALICE schematic. https://cds.cern.ch/record/2263642. Mar. 2020.
- [32] Paul Kuijer. The inner tracking system of the Alice experiment. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 530, Issues 1–2, Pages 28-32. 2004.
- [33] Naghmeh Mohammadi. Higher Harmonic Anisotropic Flow of Identified Particles. Utrecht University. 2017.
- [34] ALICE Collaboration. Determination of the event collision time with the ALICE detector at the LHC. Eur. Phys. J. Plus 132, 99. arXiv:1610.03055 [physics.ins-det]. 2007.
- [35] Francesca Carnesecchi. Performance of the ALICE Time-Of-Flight detector at the LHC. Journal of Instrumentation, volume 14. arXiv:1806.03825 [physics.ins-det]. 2018.
- [36] Details on the ALICE V0 detector. http://alice.web.cern.ch/detectors/more-details-alice-v0-detector. Mar. 2020.
- [37] ALICE Collaboration. Performance of the ALICE VZERO system. JINST 8 P10016. arXiv:1306.3130v2 [nucl-ex]. 2014.
- [38] WLCG website. https://wlcg.web.cern.ch/. Mar. 2020.
- [39] CERN general information on data processing. https://home.cern/science/computing/processing-whatrecord. Mar. 2020.
- [40] ALICE Analysis tutorial documentation. https://alice-doc.github.io/alice-analysis-tutorial/. Mar. 2020.
- [41] ROOT website. *https://root.cern.ch/about-root.* Mar. 2020.
- [42] Johannes H. Stiller. Full kinematic reconstruction of charged B mesons with the upgraded Inner Tracking System of the ALICE Experiment. Heidelberg University. 2016.
- [43] J. Schukraft. Heavy ion physics at the Large Hadron Collider: what is new ? What is next? Phys. Scr. 014003. arXiv:1311.1429 [hep-ex]. 2013.
- [44] Alberica Toia. Participants and spectators at the heavy-ion fireball. CERN Courier. 2013.
- [45] ALICE Collaboration. *D*-meson production in p-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV and in pp collisions at  $\sqrt{s} = 7$  TeV. Phys. Rev. C 94, 054908. arXiv:1605.07569 [nucl-ex]. 2016.
- [46] ALICE Collaboration. Suppression of high transverse momentum D mesons in central Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV. JHEP 09, 112. arXiv:1203.2160 [nucl-ex]. 2012.
- [47] Syaefudin Jaelani (for the ALICE Collaboration). Measurement of non-strange D-meson production and azimuthal anisotropy in Pb-Pb collisions with ALICE at the LHC. 18th International Conference on Strangeness in Quark Matter. arXiv:1910.04504 [nucl-ex]. 2019.
- [48] ALICE Collaboration. Centrality determination in heavy ion collisions (notes). ALICE-PUBLIC-2018-011. 2018.
- [49] C. Patrignani et al. *Review of particle physics*. Chinese Physics C, Volume 40, Number 10. 2016.
- [50] Bojana Blagojevic Magdalena Djordjevic Marko Djordjevic. RHIC and LHC jet suppression in noncentral collisions. Phys. Lett. B 737 298. arXiv:1405.4250v2 [nucl-th]. 2014.

# A Appendix

# A.1 Run numbers of data sets used

Data set	LHC18q_pass1	LHC18r_pass1
	296623, 296622, 296621,	1
	296619, 296618, 296616,	
	296615, 296594, 296553,	
	296552, 296551, 296550,	
	296549, 296548, 296547,	
	296516, 296512, 296511,	
	296510, 296509, 296472,	297595, 297590, 297588,
	296433, 296424, 296423,	297558, 297544, 297542,
	296420, 296419, 296415,	297541, 297540, 297537,
	296414, 296383, 296381,	297512, 297483, 297481,
	296380, 296379, 296378,	297479, 297452, 297451,
	296377, 296376, 296375,	297450, 297446, 297442,
	296312, 296309, 296304,	297441, 297415, 297414,
	296303, 296280, 296279,	297413, 297406, 297405,
	296273, 296270, 296269,	297380, 297379, 297372,
	296247, 296246, 296244,	297367, 297366, 297363,
	296243, 296242, 296241,	297336, 297335, 297333,
	296240, 296198, 296197,	297332, 297317, 297311,
	296196, 296195, 296194,	297310, 297278, 297222,
	296192, 296191, 296143,	297221, 297218, 297196,
Bun numbers	296142, 296135, 296134,	297195, 297193, 297133,
	296133, 296132, 296123,	297132, 297129, 297128,
	296074, 296066, 296065,	297124, 297123, 297119,
	296063, 296062, 296060,	297118, 297117, 297085,
	296016, 295942, 295941,	297035, 297031, 296966,
	295937, 295936, 295913,	296941, 296938, 296935,
	295910, 295909, 295861,	296934, 296932, 296931,
	295860, 295859, 295856,	296930, 296903, 296900,
	295855, 295854, 295853,	296899, 296894, 296852,
	295831, 295829, 295826,	296851, 296850, 296848,
	295825, 295822, 295819,	296839, 296838, 296836,
	295818, 295816, 295791,	296835, 296799, 296794,
	295788, 295786, 295763,	296793, 296790, 296787,
	295762, 295759, 295758,	296786, 296785, 296784,
	295755, 295754, 295725,	296781, 296752, 296694,
	295723, 295721, 295719,	296693, 296691, 296690
	295718, 295717, 295714,	
	295712, 295676, 295675,	
	295673, 295668, 295667,	
	295666, 295615, 295612,	
	295611, 295610, 295589,	
	295588, 295586, 295585	

7-7.5 GeV/c

7.5-8 GeV/c

8-9 GeV/c

9-10 GeV/c

10-12 GeV/c

12-16 GeV/c

16-24 GeV/c

24-36 GeV/c

36-50 GeV/c

50-70 GeV/c

0.93

0.93

0.93

0.93

0.93

0.93

0.92

0.87

0.8

0.8

0.3

0.3

0.3

0.3

0.3

0.3

0.15

0.15

0.15

0.15

# A.2 Cut files before optimisation

Cut 0-10%, 10	-20%	Invariant mass	dca	$\cos \theta^*$	$pT_K$	$pT_{\pi}$	$d0_{II}$	$d0_{\pi}$	d0d0	
1-2  GeV/c		0.025	0.024	0.8	0.9	1.0	0.1	0.1	-0.00045	
$2-2.5 \mathrm{GeV/c}$		0.025	0.024	0.8	0.9	1.0	0.1	0.1	-0.00045	
2.5-3  GeV/c		0.025	0.024	0.8	0.9	1.0	0.1	0.1	-0.00045	
3-3.5  GeV/c		0.024	0.022	0.8	0.9	1.0	0.1	0.1	-0.00045	
3.5-4  GeV/c		0.024	0.022	0.8	0.9	1.0	0.1	0.1	-0.00045	
4-4.5  GeV/c		0.03	0.021	0.9	0.9	0.9	0.1	0.1	-0.00035	
4.5-5  GeV/c		0.03	0.021	0.9	0.9	0.9	0.1	0.1	-0.00035	
5-5.5  GeV/c		0.032	0.021	1.0	0.9	0.9	0.1	0.1	-0.0003	
5.5-6  GeV/c		0.032	0.021	1.0	0.9	0.9	0.1	0.1	-0.0002	
6-6.5 GeV/c		0.034	0.021	1.0	0.9	0.9	0.1	0.1	-0.0002	
6.5-7  GeV/c		0.034	0.021	1.0	0.9	0.9	0.1	0.1	-0.0002	
7-7.5 GeV/c		0.036	0.021	1.0	0.9	0.9	0.1	0.1	-0.000127	
7.5-8  GeV/c		0.036	0.021	1.0	0.9	0.9	0.1	0.1	-0.000127	
8-9 GeV/c		0.055	0.021	1.0	0.9	0.9	0.1	5 0.15	-7.5e-05	
9-10 GeV/c		0.055	0.021	1.0	0.9	0.9	0.1	5 0.15	-7.5e-05	
10-12 GeV/c		0.055	0.021	1.0	0.9	0.9	0.1	5 0.15	-7.5e-05	
12-16 GeV/c		0.074	0.021	1.0	0.7	0.7	0.1	5 0.15	-7.5e-05	
16-24  GeV/c		0.074	0.021	1.0	0.5	0.5	0.1	5 0.15	-5e-05	
24-36 GeV/c		0.084	0.02	1.0	0.5	0.5	0.2	0.2	0.0004	
36-50  GeV/c		0.094	0.02	1.0	0.5	0.5	0.2	0.2	0.0004	
50-70  GeV/c		0.094	0.02	1.0	0.5	0.5	0.2	0.2	0.0004	
								·		
	$\cos \theta_P$	$M_{inv} hw D^*$	$hw \ \Delta M$	PtMi	$n \pi_s$	PtMax	$\pi_s$	$\theta \pi_s D^0$	$ \cos \theta_{PXY} $	NDL
1-2  GeV/c	0.985	0.3	0.15	0.05		1.0		1.0	0.998	7.7
$2-2.5~{ m GeV/c}$	0.985	0.3	0.15	0.05		1.0		1.0	0.998	7.7
$2.5-3 { m ~GeV/c}$	0.985	0.3	0.15	0.05		1.0		1.0	0.998	7.7
$3-3.5 \ \mathrm{GeV/c}$	0.98	0.3	0.15	0.05		1.0		1.0	0.998	7.5
3.5-4 GeV/c	0.98	0.3	0.15	0.05		1.0		1.0	0.998	7.5
4-4.5 GeV/c	0.98	0.3	0.15	0.1		10		1.0	0.998	7
4.5-5  GeV/c	0.98	0.3	0.15	0.1		10		1.0	0.998	7
5-5.5  GeV/c	0.95	0.3	0.3	0.25		10		1.0	0.998	6.5
$5.5-6 \ \mathrm{GeV/c}$	0.93	0.3	0.15	0.3		100		1.0	0.998	6.5
$6-6.5~{ m GeV/c}$	0.93	0.3	0.15	0.3		100		1.0	0.998	6.5
$6.5-7  \mathrm{GeV/c}$	0.93	0.3	0.15	0.3		100		1.0	0.998	6.5

0.3

0.3

0.3

0.3

0.3

0.3

0.3

0.3

0.3

0.3

100

100

100

100

100

100

100

100

100

100

1

1

1

1

1

1

1

1

1

1

0.998

0.998

0.998

0.998

0.998

0.99

0.99

0.9

0.9

0.9

6

6

5

5

5

1

0

0

3.7

0.5

0.15

0.15

0.15

0.15

0.15

0.15

0.15

0.15

0.15

0.15

Cut 20-30%, 30-50%, 50-60%	Invariant mass	dca	$\cos \theta^*$	$pT_K$	$pT_{\pi}$	$d0_K$	$d0_{\pi}$	d0d0
1-2 GeV/c	0.025	0.023	0.8	1.0	1.0	0.1	0.1	-0.00040
2-3  GeV/c	0.025	0.023	0.8	1.0	1.0	0.1	0.1	-0.00035
3-4 GeV/c	0.032	0.022	0.8	1.0	1.0	0.1	0.1	-0.0003
4-5  GeV/c	0.032	0.022	0.8	1.0	1.0	0.1	0.1	-0.0003
5-6 GeV/c	0.04	0.021	1.0	1.0	1.0	0.1	0.1	-0.00023
6-7 GeV/c	0.043	0.021	1.0	1.0	1.0	0.12	0.12	-0.0001
$7-8  \mathrm{GeV/c}$	0.045	0.021	1.0	1.0	1.0	0.12	0.12	-0.0001
$8-10  \mathrm{GeV/c}$	0.055	0.021	1.0	0.9	0.9	0.12	0.12	-7.5e-05
$10-12 \mathrm{GeV/c}$	0.06	0.021	1.0	0.9	0.9	0.15	0.15	-7.5e-05
12-16 GeV/c	0.074	0.021	1.0	0.7	0.7	0.15	0.15	-7.5e-05
$16-24  \mathrm{GeV/c}$	0.074	0.021	1.0	0.5	0.5	0.15	0.15	-5e-05
24-36 GeV/c	0.094	0.02	1.0	0.5	0.5	0.2	0.2	0.0004
36-50 GeV/c	0.098	0.03	1.0	0.5	0.5	0.15	0.15	0.0001

	$\cos \theta_P$	$M_{inv}$ hw $D^*$	hw $\Delta M$	PtMin $\pi_s$	PtMax $\pi_s$	$\theta \pi_s D^0$	$ \cos \theta_{PXY} $	NDL
1-2  GeV/c	0.97	0.3	0.15	0.05	0.5	0.5	0.998	7.8
$2-3  \mathrm{GeV/c}$	0.94	0.3	0.15	0.05	0.5	0.5	0.99	7
3-4  GeV/c	0.93	0.3	0.15	0.05	0.5	0.5	0.998	6.7
4-5  GeV/c	0.93	0.3	0.15	0.05	10	0.5	0.998	6.5
5-6 GeV/c	0.93	0.3	0.15	0.25	10	0.5	0.998	6.5
6-7 GeV/c	0.93	0.3	0.15	0.3	100	0.5	0.998	6.4
7-8 GeV/c	0.93	0.3	0.15	0.3	100	0.5	0.998	6.4
8-10 GeV/c	0.93	0.3	0.15	0.3	100	0.5	0.998	4.7
10-12 GeV/c	0.93	0.3	0.15	0.3	100	1	0.998	4.7
12-16 GeV/c	0.93	0.3	0.15	0.3	100	1	0.99	3.7
16-24 GeV/c	0.92	0.3	0.15	0.3	100	1	0.99	2
$24-36  \mathrm{GeV/c}$	0.85	0.15	0.15	0.3	100	1	0.9	0
36-50 GeV/c	0.77	0.15	0.15	0.3	100	999	0.9	0

Cut 60-80%	Invariant mass	dca	$\cos  heta^*$	$pT_K$	$pT_{\pi}$	$d0_K$	$d0_{\pi}$	d0d0
$3-5  \mathrm{GeV/c}$	0.032	0.02	0.8	0.9	1.0	0.1	0.1	-0.00035
$5-8  \mathrm{GeV/c}$	0.043	0.021	1.0	1.0	1.0	0.12	0.12	-0.0001
8-12 GeV/c	0.055	0.02	1.0	0.9	0.9	0.12	0.12	-7.5e-05
12-16 GeV/c	0.074	0.02	1.0	0.7	0.7	0.15	0.15	-7.5e-05
16-36 GeV/c	0.094	0.02	1.0	0.5	0.5	0.2	0.2	-7.5e-05

	$\cos \theta_P$	$M_{inv}$ hw $D^*$	hw $\Delta M$	PtMin $\pi_s$	PtMax $\pi_s$	$\theta \pi_s D^0$	$ \cos \theta_{PXY} $	NDL
$3-5 \mathrm{GeV/c}$	0.92	0.3	0.15	0.05	1.0	1.0	0.998	7
5-8  GeV/c	0.93	0.3	0.3	0.25	10	1.0	0.998	6.5
8-12 GeV/c	0.93	0.3	0.15	0.3	100	1	0.998	5
12-16  GeV/c	0.85	0.3	0.15	0.3	100	1	0.99	5
16-36 GeV/c	0.7	0.15	0.15	0.3	100	1	0.9	2

#### A.3 Mass spectra before optimisation

#### A.3.1 As a function of momentum



Mass spectra  $D^*$  at 0-10% centrality



Mass spectra  $D^*$  at 10-20% centrality





Mass spectra  $D^*$  at 20-30% centrality













Mass spectra  $D^*$  at 60-80% centrality



#### A.3.2 As a function of centrality



Mass spectra  $D^*$  at 3-5 GeV/c





Mass spectra  $D^*$  at 8-12 GeV/c

# A.4 Reconstruction efficiencies before optimisation



Efficiencies 10-20 cut





Efficiencies 20-30 cut



Efficiencies 30-50 cut



Efficiencies 50-60 cut



## A.5 Corrected yield and cross section before optimisation



Reconstructed yield and reconstructed cross section at 0-10% centrality

Reconstructed yield and reconstructed cross section at 10-20% centrality





#### Reconstructed yield and reconstructed cross section at 20-30% centrality

Reconstructed yield and reconstructed cross section at 30-50% centrality





#### Reconstructed yield and reconstructed cross section at 50-60% centrality

Reconstructed yield and reconstructed cross section at 60-80% centrality



## A APPENDIX

# A.6 Cut values after optimisation

Cut optimised 10-20%	Invariant mass	dca	$\cos \theta^*$	$pT_K$	$pT_{\pi}$	$d0_K$	$d0_{\pi}$	d0d0
$3-5 \mathrm{GeV/c}$	0.024	0.022	0.8	0.9	1.0	0.1	0.1	-0.000513
$5-8  \mathrm{GeV/c}$	0.032	0.021	1.0	0.9	0.9	0.1	0.1	-0.0003
$8-12  \mathrm{GeV/c}$	0.055	0.021	1.0	0.9	0.9	0.15	0.15	-7.5e-05
12-16 GeV/c	0.074	0.021	1.0	0.7	0.7	0.15	0.15	-7.5e-05
16-36 GeV/c	0.074	0.021	1.0	0.5	0.5	0.15	0.15	-5e-05

	$\cos \theta_P$	$M_{inv} hw D^*$	$hw \Delta M$	$PtMin \ \pi_s$	$PtMax \ \pi_s$	$\theta \pi_s D^0$	$ \cos \theta_{PXY} $	NDL
$3-5  \mathrm{GeV/c}$	0.987125	0.3	0.15	0.05	1.0	1.0	0.998	7.875
5-8  GeV/c	0.98	0.3	0.3	0.25	10	1.0	0.998	6.8125
8-12 GeV/c	0.973125	0.3	0.15	0.3	100	1	0.998	5.25
12-16 GeV/c	0.9375	0.3	0.15	0.3	100	1	0.99	4.85
16-36 GeV/c	0.98125	0.15	0.15	0.3	100	1	0.99	2.75

Cut optimised 20-30%	Invariant mass	dca	$\cos \theta^*$	$pT_K$	$pT_{\pi}$	$d0_K$	$d0_{\pi}$	d0d0
$3-5  \mathrm{GeV/c}$	0.032	0.022	0.8	1.0	1.0	0.1	0.1	-0.0003
$5-8  \mathrm{GeV/c}$	0.04	0.021	1.0	1.0	1.0	0.1	0.1	-0.00023
$8-12  \mathrm{GeV/c}$	0.055	0.021	1.0	0.9	0.9	0.12	0.12	-0.000116
12-16 GeV/c	0.074	0.021	1.0	0.7	0.7	0.15	0.15	-7.5e-05
16-36 GeV/c	0.074	0.021	1.0	0.5	0.5	0.15	0.15	-5e-05

	$\cos \theta_P$	$M_{inv} hw D^*$	$hw \ \Delta M$	PtMin $\pi_s$	$PtMax \ \pi_s$	$\theta \pi_s D^0$	$ \cos \theta_{PXY} $	NDL
$3-5  \mathrm{GeV/c}$	0.975	0.3	0.15	0.05	0.5	0.5	0.998	7.025
5-8  GeV/c	0.9825	0.3	0.15	0.25	10	0.5	0.998	7.25
8-12 GeV/c	0.96	0.3	0.15	0.3	100	0.5	0.998	4.7
12-16 GeV/c	0.955	0.3	0.15	0.3	100	1	0.99	4.8375
16-36 GeV/c	0.96375	0.3	0.15	0.3	100	1	0.99	3.75

#### A.7 Mass spectra after optimisation



Optimised mass spectra 10-20%







# A.8 Reconstruction efficiencies after optimisation



## A.9 Corrected yield and cross section after optimisation



Reconstructed yield and reconstructed cross section at 10-20% centrality (optimised)

Reconstructed yield and reconstructed cross section at 20-30% centrality (optimised)



### A.10 Mass spectra from cut variation






30-50%









## A.11 Corrected cross sections of cut variation

## A.12 Reconstruction efficiency ratios of cut variation







Efficiency ratio original and L20 60-80

## Efficiency ratio original and L10 60-80