

Identified particle yield  
associated with a high- $p_T$   
trigger particle at the LHC

ISBN: 978-90-393-6496-3

A catalogue record is available from the Utrecht University Library.

Art design: Sander Kloppenburg

This work is licensed under a *Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0)*

# Identified particle yield associated with a high- $p_T$ trigger particle at the LHC

Geïdentificeerde deeltjesproductie geassocieerd met een  
hoog- $p_T$  trigger deeltje bij de LHC

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van  
de rector magnificus, prof. dr. G.J. van der Zwaan, ingevolge het besluit van het  
college voor promoties in het openbaar te verdedigen op maandag 14 maart  
2016 des middags te 4.15 uur

DOOR

**Misha Veldhoen**

geboren op 31 mei 1985  
te Zevenaar

**Promotor:** Prof. dr. T. Peitzmann

**Copromotor:** Dr. ir. M. van Leeuwen

Dit werk maakt deel uit van het onderzoeksprogramma van de Stichting voor Fundamenteel Onderzoek der Materie (FOM), die financieel wordt gesteund door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

# Contents

<b>1. Introduction</b>	<b>1</b>
<b>2. Background Information</b>	<b>11</b>
2.1. Coordinates and Conventions . . . . .	12
2.2. Receding Pancakes . . . . .	13
2.3. Hydrodynamics and Collectivity . . . . .	16
2.3.1. Bjorken Model . . . . .	17
2.3.2. Anisotropic Flow . . . . .	17
2.4. Statistical Models of Particle Production . . . . .	20
2.5. The Standard Model . . . . .	21
2.6. Quantum Chromodynamics (QCD) . . . . .	23
2.6.1. Hadronization and Fragmentation Functions . . . . .	25
2.6.2. Lattice QCD and the QCD Phase Diagram . . . . .	26
2.7. Particle Recombination . . . . .	28
2.8. Hadrochemistry in Jets . . . . .	31
2.8.1. Modified Splitting Functions . . . . .	32
2.8.2. Shower-Thermal Recombination . . . . .	33
2.9. Did we discover the QGP? . . . . .	33
<b>3. Experimental Setup</b>	<b>37</b>
3.1. The Large Hadron Collider (LHC) . . . . .	37
3.2. The ALICE Detector . . . . .	38
3.2.1. Overview . . . . .	40
3.2.2. The Inner Tracking System (ITS) . . . . .	40
3.2.3. The Time Projection Chamber (TPC) . . . . .	40
3.2.4. The Time Of Flight (TOF) Detector . . . . .	44
3.2.5. The V0 detectors . . . . .	46

3.3.	Triggering . . . . .	46
3.4.	Offline Analysis . . . . .	47
3.4.1.	Centrality Determination . . . . .	47
3.4.2.	Track and Event Reconstruction . . . . .	47
<b>4.</b>	<b>Particle Identification</b>	<b>49</b>
4.1.	TPC Detector Response . . . . .	51
4.2.	TOF Detector Response . . . . .	52
4.3.	TOF-TPC Combined Fits . . . . .	56
4.4.	Systematic Uncertainties . . . . .	58
4.4.1.	Mismatched Background . . . . .	59
4.4.2.	Choice of Fit Region . . . . .	59
4.4.3.	Comparison with Other Fit Methods . . . . .	63
4.4.4.	Fixing the Tail Parameter . . . . .	64
4.4.5.	Contamination from Leptons . . . . .	65
4.5.	Fit Results . . . . .	66
4.6.	Correction for Secondary Particles . . . . .	68
4.7.	Summary . . . . .	70
<b>5.</b>	<b>Identified Associated Yield</b>	<b>71</b>
5.1.	Associated Particle Distributions . . . . .	71
5.2.	Identifying Associated Yield . . . . .	77
5.2.1.	Fit Templates . . . . .	78
5.2.2.	Fit Procedure . . . . .	80
5.3.	Systematic Uncertainty . . . . .	82
5.4.	Fit Results . . . . .	83
5.5.	Summary . . . . .	85
<b>6.</b>	<b>Associated Yield Analysis</b>	<b>87</b>
6.1.	A Model for the $\Delta\eta$ Projection . . . . .	88
6.2.	A Model for the $\Delta\phi$ Projection . . . . .	90
6.3.	Jet-Peak Yield Measurement . . . . .	91
6.4.	Choice of Peak Region in $(\Delta\eta, \Delta\phi)$ . . . . .	95
6.5.	A Toy-Model Study . . . . .	97
6.6.	Correction for Secondary Particles . . . . .	101
6.7.	Efficiency Correction . . . . .	103
6.8.	Systematic Uncertainties . . . . .	105
6.9.	Results . . . . .	108
6.10.	Summary . . . . .	114

---

7. Discussion and Outlook	115
A. Event and Track Cuts	119
B. Data Sample	121
C. The Discrete Cosine Transform	123
D. Associated Yield Projections	125
Bibliography	129
Summary	139
Samenvatting	143
Acknowledgements	147
Curriculum Vitae	149





## Chapter 1.

# Introduction

In the beginning of the 20th century, Geiger and Marsden carried out a number of experiments that were proposed by Rutherford to elucidate the structure of atoms. In these experiments, Geiger and Marsden bombarded a piece of gold foil with a beam of energetic  $\alpha$ -particles<sup>1</sup>, and measured the angle of deflection [1]. To their surprise, they observed very large angles of deflection, at a rate that could only be explained by a model of the atom where the bulk of the atomic mass is concentrated in a very small positively charged core, called the *nucleus*. In the years that followed, it was discovered by Rutherford that  $\alpha$ -particles could be used to knock off protons from heavier atoms, making it plausible that the nucleus was composed of protons. The mass of atoms, however, was not equal to the sum of the constituent protons. This discrepancy would only be resolved about a decade later, when Chadwick discovered the neutron [2, 3]. The neutron has approximately the same mass as the proton, but carries no net electrical charge. Assuming that the nucleus is composed of both protons and neutrons (collectively called *nucleons*), the masses in the periodic table could be understood.

In the middle of the 20th century, numerous new particles were discovered in collider experiments. Most of these new particles were unstable, quickly decaying to more stable particles such as protons and neutrons. In an attempt to create order in the growing particle population, as well as to understand the observed decay processes, Gell-Mann [4] and Zweig [5, 6] independently proposed a model where protons, neutrons, and all the newly discovered particles are composite particles, made up of *quarks*. To describe the proton and the neutron, two types of quarks are needed, the up quark ( $u$ ) with

---

<sup>1</sup>Helium ions.

electric charge  $\frac{2}{3}e$ , and the down quark ( $d$ ) with electric charge  $-\frac{1}{3}e$ . Within this scheme, the composition of the proton is  $uud$ , and the neutron is  $udd$ . To fit all the other particles in same the ordering scheme, more quarks are needed, as well as anti-quarks<sup>2</sup>. Currently six different types (or *flavors*) of quarks are known. All of them are spin- $\frac{1}{2}$  fermions, and carry the same electric charge as either the up or down quark. On the other hand, the masses of these quarks are very different. An overview of the approximate quark masses is shown in Table 1.1. Particles that are made up of quarks are collectively referred to as *hadrons*. To date, only two types of hadrons have been observed, ones consisting of a quark and an anti-quark, called *mesons*, and ones with three (anti-)quarks, called *baryons*.

In its original form, there was a problem with the quark model, which was related to baryons like  $\Delta^{++}(uuu)$ ,  $\Delta^{-}(ddd)$ . As these baryons are composed of three identical quarks and have a spin of  $\frac{3}{2}$ , the three quarks have to be in the same spin state, seemingly violating the Pauli exclusion principle. The issue was resolved by postulating a new quantum number called *color* to restore the anti-symmetry of the baryon states under quark exchange. There are six types of color, (anti-)red, (anti-)blue and (anti-)green, and baryons are in the three particle singlet state of this quantum number<sup>3</sup>.

At the time, Quantum Electrodynamics (QED) had already been developed, and it was soon realized that the force binding quarks together could be described by a very similar model, using the new color degree of freedom as the charge. In QED, the electromagnetic force between charged particles is mediated by a spin-1 boson, the photon. Analogously, it was proposed that there exists a force between particles with a net color charge, mediated by spin-1 bosons called *gluons*. This novel force is nowadays called the *strong force*, and the theory of quarks and gluons is called Quantum Chromodynamics (QCD). Color charge is similar to the electric charge in the sense that a composite particle with an equal number of positive and negative charges is color neutral. For example, all mesons are color neutral ( $\bar{r}r$ ,  $\bar{b}b$ ,  $\bar{g}g$ )<sup>4</sup>. There is, however, another way to form a color neutral composition, which is by combining three particles, each carrying a different positive (negative) type of charge<sup>5</sup>. Consequently, (anti-)baryons are color neutral (e.g.  $rgb$ ,  $\bar{r}\bar{g}\bar{b}$ ).

<sup>2</sup>An anti-particle has the same mass as the corresponding particle, but has opposite charge.

<sup>3</sup>This state is:  $rgb - rbg + brg - bgr + gbr - grb$ .

<sup>4</sup>The bar indicates an anti-particle

<sup>5</sup>The rule that a neutral object is created by combining three different charges was inspiration for the name “color”, in analogy with the three base colors.

Quark Flavor	Mass (MeV/ $c^2$ )	El. Charge
Up	$\sim 4$	$+\frac{2}{3}e$
Down	$\sim 7$	$-\frac{1}{3}e$
Charm	$\sim 1.5 \times 10^3$	$+\frac{2}{3}e$
Strange	$\sim 135$	$-\frac{1}{3}e$
Top	$\sim 175 \times 10^3$	$+\frac{2}{3}e$
Bottom	$\sim 5 \times 10^3$	$-\frac{1}{3}e$

**Table 1.1.:** The six quark flavors. Quark masses taken from Ref. [7].

The force that binds protons and neutrons together in nuclei, is understood as a direct consequence of the strong force. As nucleons are baryons, there is no color monopole interaction between them. The spatial distribution of color charge within the nucleons, however, causes multipole interactions, which, similarly to the van der Waals force, causes a net residual force between the nucleons. This residual force is identified with the nuclear force and its strength rapidly decreases with distance, which explains why the nuclear force does not seem to play an important role at large scales.

While the quark model was very successful in organizing all the hadron states, many physicists were not convinced that quarks were “real” particles. The main objection was that quarks had never been observed in the laboratory, i.e., there was no direct experimental proof of their existence. In an effort to probe the interior of nucleons, deep inelastic scattering experiments were carried out at the Stanford Linear Accelerator (SLAC) at the end of the 1960s. The principle of these experiments is the same as that of those carried out by Geiger and Marsden half a century earlier, except that now a much higher beam energy was used, thereby greatly increasing the resolving power. Also, instead of a beam of  $\alpha$ -particles, the SLAC experiments used electrons. As electrons only interact electromagnetically, the DIS experiments are sensitive to the distribution of electric charge within the target proton. These experiments led to the discovery that nucleons are indeed composite particles, and the constituents were called *partons*, a term coined by Feynman. Later, it was realized that these partons are in fact the same objects as quarks and gluons. Nevertheless, the term partons is still commonly used in particle physics.

Significant progress towards understanding why quarks are never observed as free particles, also referred to as *confinement*, was made in the early 1970s, when Gross, Wilczek [8] and Politzer [9], discovered that QCD was *asymptotically free*<sup>6</sup>. Asymptotic freedom means that the strength of the strong coupling decreases as the energy scale of a scattering experiment is increased (the probed length scale is decreased). Conversely, when the energy scale is decreased (length scale increased), the coupling becomes stronger. While the latter is a necessary condition for confinement, it should be noted that it is not sufficient. The equations of motion of QCD have so far not been solved exactly, which means that to date no analytical proof of quark confinement exists.

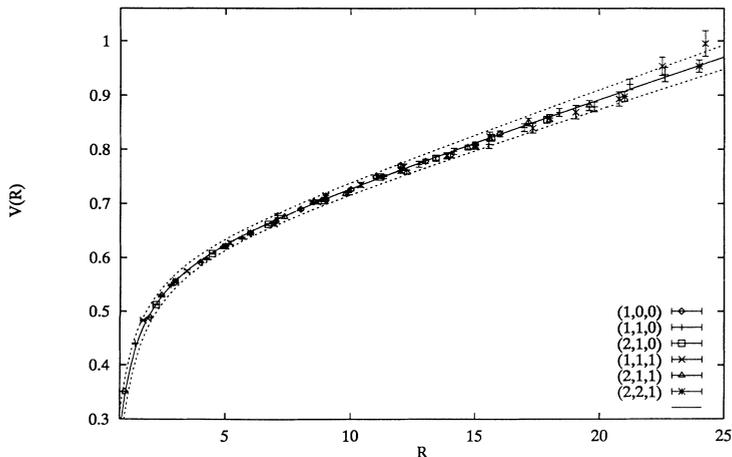
An alternative way to gain theoretical insight in the mechanism underlying confinement is to study a discretized version of QCD using Monte Carlo techniques. In this approach, called Lattice QCD, one generates an ensemble of tiny volumes of QCD matter<sup>7</sup> in thermal equilibrium, and measures various properties of the system. One important example is that one can measure the potential between a quark and an anti-quark  $V_{q\bar{q}}(r)$ , where  $r$  is the distance between the two quarks, shown in Fig. 1.1. While at short distances  $V_{q\bar{q}}(r)$  behaves approximately like a Coulomb potential, i.e.,  $V \sim 1/r$ , the large  $r$  behavior is very different. Instead of approaching a constant value at large  $r$ ,  $V_{q\bar{q}}(r)$  is shown to grow linearly, which is a hint (but again not a proof) of confinement.

Besides studying static quark configurations one can also use Lattice QCD to study how a volume of QCD matter behaves under changing thermodynamic circumstances, i.e., one can make a phase diagram. One interesting observation that was made, was that the effective number of degrees of freedom of QCD matter at extremely high temperatures and pressures far exceeds the one at low temperatures. Since the number of degrees of freedom seem to approach the number expected from a gas of free quarks and gluons, this is interpreted as a phase transition from a confined to a deconfined state. Intuitively, this phase transition can be compared to the phase transition from water droplets to vapor. It should be noted that the thermodynamic circumstances under which QCD matter is deconfined are so extreme that it is believed that deconfinement only occurs within neutron stars, or in the very early universe [10].

---

<sup>6</sup>t Hooft discovered asymptotic freedom one year before Gross, Wilczek and Politzer, but did not publish his results.

<sup>7</sup>This can also mean the QCD vacuum, which, due to the zero-point energy also contains QCD matter.



**Figure 1.1.:** Static quark potential from Lattice QCD as a function of the distance between the quarks  $R$ . The different symbols indicate the direction of  $R$  in terms of points of the lattice. Figure taken from Ref. [11].

The goal of relativistic heavy-ion collisions is to create a system of QCD matter in its deconfined phase in the laboratory, and to study its properties. By colliding highly energetic heavy ions, one can create a system of QCD matter with an energy density that is expected to be sufficiently high for the matter to undergo a phase transition to the deconfined state (see Sect. 2.3 and 2.6). Immediately after the collision, the hot QCD matter expands rapidly and cools down in the process. Long before reaching the detector, the QCD matter turns into hadrons again, a process known as *hadronization*. Since the system is only deconfined for a very short time, an important question that one can ask is whether the system exists long enough in the deconfined phase to be thermalized. As is discussed in Chapter 2, the currently available evidence favors very short thermalization times.

Another important open question is to understand the mechanism of hadronization. Hadronization of single quarks, generated for example in  $e^+e^- \rightarrow q\bar{q} \rightarrow \text{hadrons}$ <sup>8</sup>, is reasonably well understood. In such a collision, both quarks turn into a “spray” of hadrons called a *jet*, the properties of which can be determined experimentally. The theoretical understanding of

<sup>8</sup>The dominant hadronization mechanism in hadronic collisions such as  $\bar{p}p \rightarrow \text{hadrons}$  is assumed to be similar to the one in leptonic collisions.

this process in the perturbative regime is expressed in a system of integro-differential equations<sup>9</sup>, that relate the properties of jets coming from quarks with different virtualities<sup>10</sup>. On the other hand, experimental evidence suggests that the hot QCD matter produced in heavy-ion collisions hadronizes in a very different and much less understood way.

In a simplified picture, one can look at the hot QCD matter as being composed of thermalized low-momentum or *soft* partons, and high-momentum or *hard* partons. The soft partons are much more abundant than the hard partons, and their collective behavior can be understood as that of a relativistic fluid. The hard partons on the other hand, traverse the fluid of low momentum partons, thereby losing energy in the process. These hard partons can be used to probe the properties of the thermalized QCD matter, in a manner reminiscent of the way that leptonic probes were used to probe the interior of a nucleus. For example, one can study the jet spectrum in head-on heavy-ion collisions, and make a comparison with the jet spectrum measured in pp collisions [12], or peripheral heavy-ion collisions [13]. These studies show a significant suppression of the jet spectrum in head-on heavy-ion collisions, a phenomenon known as *jet quenching*.

There are two basic methods one can use to study jets. The first method is to reconstruct the jet, i.e., apply a jet-finding algorithm on the particles that are produced in a collision, that assigns every particle to a jet [14, 15]. This method works especially well in an environment where few jets are created, for example in pp collisions. It is in principle able to recover the kinematics on the partonic level, regardless of the details of the hadronization. On the other hand, in head-on heavy-ion collisions, where very many particles are created, it is much less obvious to relate the measured jets to physics on the partonic level. A second method to study jets is by specifically selecting those events which contain a hadron with large momentum (the *trigger hadron*), and to measure the angle of all other hadrons (the *associated hadrons*) in the event with respect to this trigger hadron. This probe is mainly sensitive to jets that have a large total momentum, distributed in relatively few hadrons<sup>11</sup>.

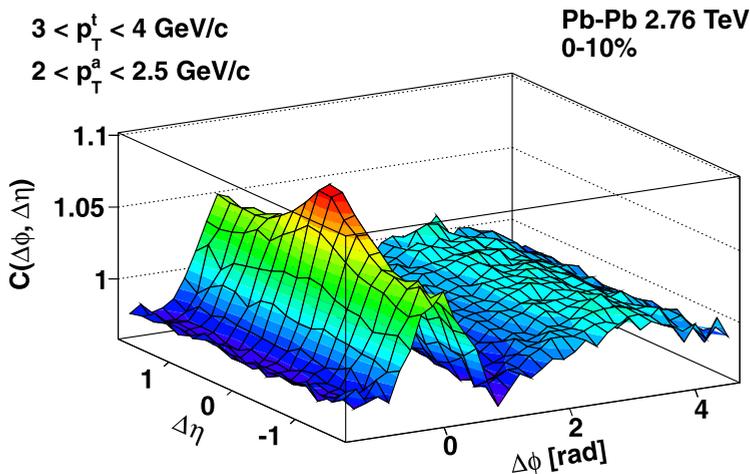
The second method has been studied extensively at the Relativistic Heavy-Ion Collider (RHIC) at Brookhaven [17–23] and at LHC [16, 24],

---

<sup>9</sup>DGLAP equations.

<sup>10</sup>Virtuality is defined as the amount of energy by which a particle is off its mass-shell. For a massless particle, the virtuality is therefore equal to  $E^2 - \mathbf{p}^2$ . As virtuality is the relevant energy scale in the evolution of a jet, the (perturbative) DGLAP equations can only be used for particles with a sufficiently high virtuality.

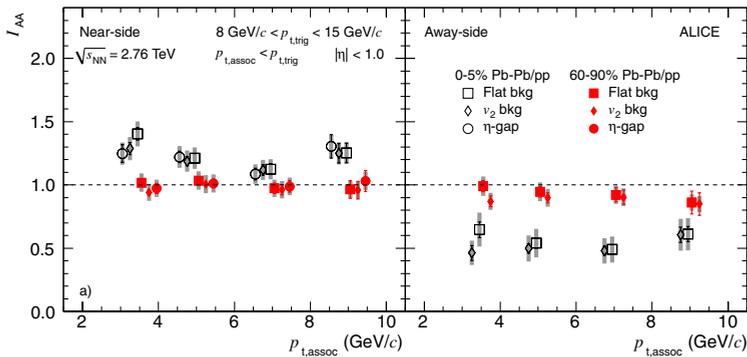
<sup>11</sup>In practice this does not induce a strong bias.



**Figure 1.2.:** Distribution of associated hadrons at  $2.0 < p_T < 2.5$  GeV/c with respect to trigger hadrons at  $3.0 < p_T < 4.0$  GeV/c in head-on Pb–Pb collisions. The correlation function  $C$  is normalized to one for uncorrelated pairs. The variables  $\Delta\phi$  and  $\Delta\eta$  are the relative azimuthal angle and the relative pseudorapidity of the associated hadron and the trigger hadron respectively. Pseudorapidity is related to the polar angle  $\theta$ , its precise definition is given in Sect. 2.1. Figure taken from Ref. [16].

and an example is shown in Fig. 1.2, where trigger particles are in the range  $3.0 < p_{T,\text{trig}} < 4.0$  GeV/c, and associated particles are in the range  $2.0 < p_{T,\text{assoc}} < 2.5$  GeV/c. The peak-shaped structure at  $(\Delta\eta, \Delta\phi) = (0, 0)$  is related to the jet, and is called the *jet peak*. The ridge-shaped structure underneath the jet peak is thought to originate from the collective motion or *flow* of the soft partons. On the other side ( $\Delta\phi \sim \frac{\pi}{2}$ ) another ridge-shaped structure is observed. This structure is believed to be mainly caused by flow, but also contains the jet coming from the parton that traversed the QCD medium in opposite direction compared to the one causing the jet at the origin<sup>12</sup>.

<sup>12</sup>While the opposite jet has to be produced at a relative azimuthal angle  $\Delta\phi = \pi$ , in the polar direction there is no such restriction. This is because the center-of-mass frame of the heavy ion collision is typically not the same as the center-of-mass frame of the initial partons that underwent a hard scattering.



**Figure 1.3.:** Left panel:  $I_{AA}$  of the near-side peak i.e.,  $|\Delta\phi| < \frac{\pi}{2}$ , for two centralities. Right panel:  $I_{AA}$  of the away-side peak, i.e.,  $|\Delta\phi| > \frac{\pi}{2}$ , for two centralities. Figure taken from Ref. [25].

One signature of jet quenching using a triggered correlation measurement such as in Fig. 1.2 is  $I_{AA}$ , which is defined as the per-trigger yield in the peak structure of an ion-ion collision, scaled by the same quantity in a pp collision. In Fig. 1.3 we show an example of an  $I_{AA}$  measurement performed by the ALICE collaboration, where trigger particles were selected in the range  $8 < p_{T,\text{trig}} < 15 \text{ GeV}/c$ , and  $p_{T,\text{assoc}} < p_{T,\text{trig}}$ . The results for the near ( $|\Delta\phi| < \frac{\pi}{2}$ ) and away-side ( $|\Delta\phi| > \frac{\pi}{2}$ ) peaks are shown, and a comparison is made between central and non-central Pb–Pb collisions<sup>13</sup>. While  $I_{AA}$  in the non-central collisions is close to unity, the  $I_{AA}$  in central collisions is significantly different. This means that the triggered correlation measurement can be used as a probe for jet quenching, and complement the observables based on fully reconstructed jets.

In this work we describe a measurement that is similar to the one shown in Fig. 1.2, but with the aim to measure the hadronic composition of the peak structure associated with a high- $p_T$  trigger particle. Using two of the main particle identification (PID) detectors in ALICE, we select the associated particles to be one of the three most abundantly produced hadrons, i.e., pions, kaons, and protons. We then subtract the background yield, and measure the hadron composition of the jet peak. The measurement is performed in head-on Pb–Pb and p–Pb collisions and is compared to a pp baseline.

<sup>13</sup>Central AA collisions are head-on, as opposed to non-central collisions. A lower percentage corresponds to a more central collision. The notion of centrality is made more precise in Sect. 2.2.

There are two main motivations for this measurement. First of all, it complements the current set of observables of jet quenching. Second, there are a number of theoretical reasons to believe that the QCD medium has a significant effect on the hadronic composition or *hadrochemistry* of jets [26].

In the next chapter we provide an overview of heavy-ion phenomenology, as well as an overview of several theoretical concepts relevant for the interpretation of this measurement. In Chapter 3, we give a description of the ALICE experiment at CERN, which was used to obtain the data for this work. The focus is on the sub-detectors of ALICE that were most important for this measurement. After that, in Chapters 4 and 5 we give a detailed description of the analysis that was carried out, followed by a discussion of the results of the measurement in Chapter 6. Some concluding remarks are given in Chapter 7.





## Chapter 2.

# Background Information

In this chapter, we provide a brief introduction to the field of heavy-ion physics. The field of heavy-ion physics is extensive, therefore we restrict ourselves to discussing concepts that are either important for a general understanding of the field, or that are relevant for the interpretation of the measurement that we performed. For a more in-depth overview on the topic, see for example Refs. [27, 28].

As mentioned in the previous chapter, the goal of heavy-ion collisions is to create and study a system of hot and deconfined QCD matter in the laboratory. The only things that we can measure experimentally, however, are the kinematics of the particles produced at these collisions and the correlations between them. Any interpretation of the data therefore requires an understanding of the time evolution of the QCD matter from its deconfined state back to hadrons. In principle the time evolution can be understood from the equations of motion of the Standard Model of elementary particles, and most importantly QCD. Unfortunately these equations cannot be solved exactly, and currently our best understanding of the time evolution is based on a patchwork of effective theories, valid only in a limited part of the time evolution, complemented with perturbative and numerical results from QCD.

Our first goal of this chapter is to give an overview of the most important effective models and approximations of QCD currently being used, in order to give a relatively complete picture of the time evolution of a heavy-ion collision. A secondary goal is to highlight some of the most important evidence that we have for the existence of a deconfined state in heavy-ion collisions. While all the evidence is exclusively circumstantial, it is sufficient for most physicists to believe that a deconfined state has in fact been created. Finally, the last goal of this chapter is to present some theoretical ideas that are related to

the measurement of identified hadrons correlated to a trigger particle at high momentum. While there has not yet been a quantitative prediction specifically for this measurement in heavy-ion collisions, there are several ideas that (with some modification) could lead to such a prediction.

This chapter is organized as follows. We start by discussing the coordinates and the commonly used variables in Sect. 2.1. In Sect. 2.2, we give a simple overview of a heavy-ion collision and we discuss the space-time evolution of the produced QCD matter, identifying several distinct stages that the matter is generally thought to pass through. We then continue by discussing two approaches for effective theories: relativistic hydrodynamics (Sect. 2.3), which has been shown to be a good description of the early stages, and statistical models (Sect. 2.4), which have been found to very accurately predict the ratios of identified particles produced in a heavy-ion collision. After this, we say a few words about the Standard Model (Sect. 2.5), with an emphasis on QCD (Sect. 2.6). We show how a perturbative treatment of QCD is applicable at high energies, and how this can be used to model certain aspects of particle production. Furthermore, we discuss some of the results from discretized QCD (Lattice QCD), which is currently the most important tool used in the theoretical study of the QCD phase diagram, and provides the strongest evidence for the existence of a deconfined phase. We then go on to show experimental evidence indicating a significantly altered mechanism of hadron production in heavy-ion collisions compared to elementary particle collisions, and discuss models for hadron production in heavy-ion collisions (Sects. 2.7, 2.8). We end the chapter with a discussion of the term Quark Gluon Plasma (QGP), and whether it has been discovered at RHIC and LHC (Sect. 2.9).

## 2.1. Coordinates and Conventions

In particle collision experiments, it is convenient to adopt a cylindrical coordinate system  $(z, r, \phi)$ , with the origin at the center of the detector,  $r$  being the radial direction,  $\phi$  the azimuthal angle, and the  $z$ -axis coinciding with the beam axis. Kinematics are typically described in terms of the momentum projected onto the plane transverse to the beam pipe,  $p_T = \sqrt{p_r^2 + p_\phi^2}$ , and the *rapidity*  $y$  with respect to the beam axis, i.e.:

$$y = \frac{1}{2} \ln \frac{E + p_z c}{E - p_z c} = \operatorname{arctanh}(v_z/c), \quad (2.1)$$

where  $E$  stands for energy,  $p$  for momentum,  $v$  for velocity and  $c$  for the speed of light. Using rapidity instead of velocity is especially useful as it is additive under Lorentz boosts along the  $z$ -axis.

While the theoretical advantage of using rapidity is clear, it is not the most straightforward quantity to measure experimentally. A related quantity that is easier to measure and therefore often used instead, is the *pseudorapidity*  $\eta$ , which is defined as:

$$\eta = -\ln \tan \theta/2, \quad (2.2)$$

where  $\theta$  is the angle with respect to the beam axis. After the identity of the particle has been determined with one or more particle identification (PID) detectors, the rapidity can be obtained as follows:

$$y = \ln \left( \frac{\sqrt{m^2 + p_T^2} \cosh^2 \eta + p_T \sinh \eta}{\sqrt{m^2 + p_T^2}} \right), \quad (2.3)$$

where  $m$  is the mass of the particle.

The unit for energy that is used throughout this work is the electron volt,  $1 \text{ eV} \approx 1.6 \times 10^{-19} \text{ J}$ . Similarly, we express momenta in terms of  $\text{eV}/c$ , and masses in terms of  $\text{eV}/c^2$ .

The initial system of two colliding particles is characterized by the collision energy in the center-of-mass frame:

$$\sqrt{s} = \sqrt{(p_1 + p_2)^\mu (p_1 + p_2)_\mu}, \quad (2.4)$$

where  $s$  is one of the Mandelstam variables, and we have used the Einstein summation convention. When one or two of the particles in the collision system is an ion, the center-of-mass energy is given per nucleon pair, and is denoted by  $\sqrt{s_{\text{NN}}}$ .

In this work we analyze data from proton-proton (pp), proton-lead (p-Pb) and lead-lead (Pb-Pb) collisions. In the following, we use the generic notation pA for proton-nucleus collisions, and AA for nucleus-nucleus collisions, whenever the mass number of the nucleus is not important for the discussion.

## 2.2. Receding Pancakes

The dimension of a highly relativistic object ( $\gamma \gg 1$ ) along the axis of propagation appears as contracted to the stationary observer. For this reason,

a highly relativistic ion looks like an almost flat disc, and is colloquially referred to as a “pancake” [29].

In Fig. 2.1 we present a simplified picture of a heavy-ion collision. In the top panel, the two Lorentz contracted ions approach each other at nearly the speed of light. Both ions are assumed to be lead ions, containing 208 nucleons which are represented by white spheres. The ions in the picture do not approach each other exactly head-on, and the perpendicular distance between their trajectories defines the *impact parameter*  $\mathbf{b}$ . The middle panel describes the situation at the time of the collision. Since the impact parameter is nonzero, but smaller than the nuclear diameter, some nucleons collide directly with other nucleons (the *participants*), while some continue relatively unharmed (the *spectators*). In the bottom panel the ions have passed through each other (“receding pancakes”), leaving a trail of hot QCD matter at large energy densities in their wake. Due to a large pressure in the QCD matter, it expands and cools down rapidly, and all QCD matter turns into hadrons again.

In collision experiments, the impact parameter cannot be controlled, however, since in pA and AA collisions a smaller impact parameter is correlated to a larger particle production, it can be estimated. Using the Glauber model [30], one can find the approximate mapping from the number of particles produced to the impact parameter. In the ALICE experiment, the particle multiplicity for this estimate is measured at large rapidity, using the V0 detectors, or at mid-rapidity, using the TPC (see Sect. 3.4.1) [31]. By comparing the particle multiplicity of a given event to the multiplicity distribution of all events, we can calculate its *centrality percentage*. An event is said to have a centrality of  $n\%$  when  $n\%$  of the events have a larger particle multiplicity, i.e., are more central.

The evolution of the hot QCD matter to hadrons can be divided into several stages. It is estimated that directly after the collision ( $t \sim 0$ ), the system has a sufficiently high energy density that it is in a deconfined state, see Sect. 2.3.1. After a short amount of time ( $t \sim \tau_{\text{th}}$ ), the system thermalizes. While the thermalization time  $\tau_{\text{th}}$  cannot be measured directly, it is believed to be very short  $\tau_{\text{th}} \sim 1 \text{ fm}/c$ . The evidence for this mainly comes from comparisons between experimental data<sup>1</sup> and fluid models that assume a short thermalization time (Sect. 2.3). The next step in the evolution is hadronization, where the QGP turns into a strongly interacting system of hadrons.

---

<sup>1</sup>Specifically: the observed final state anisotropies.

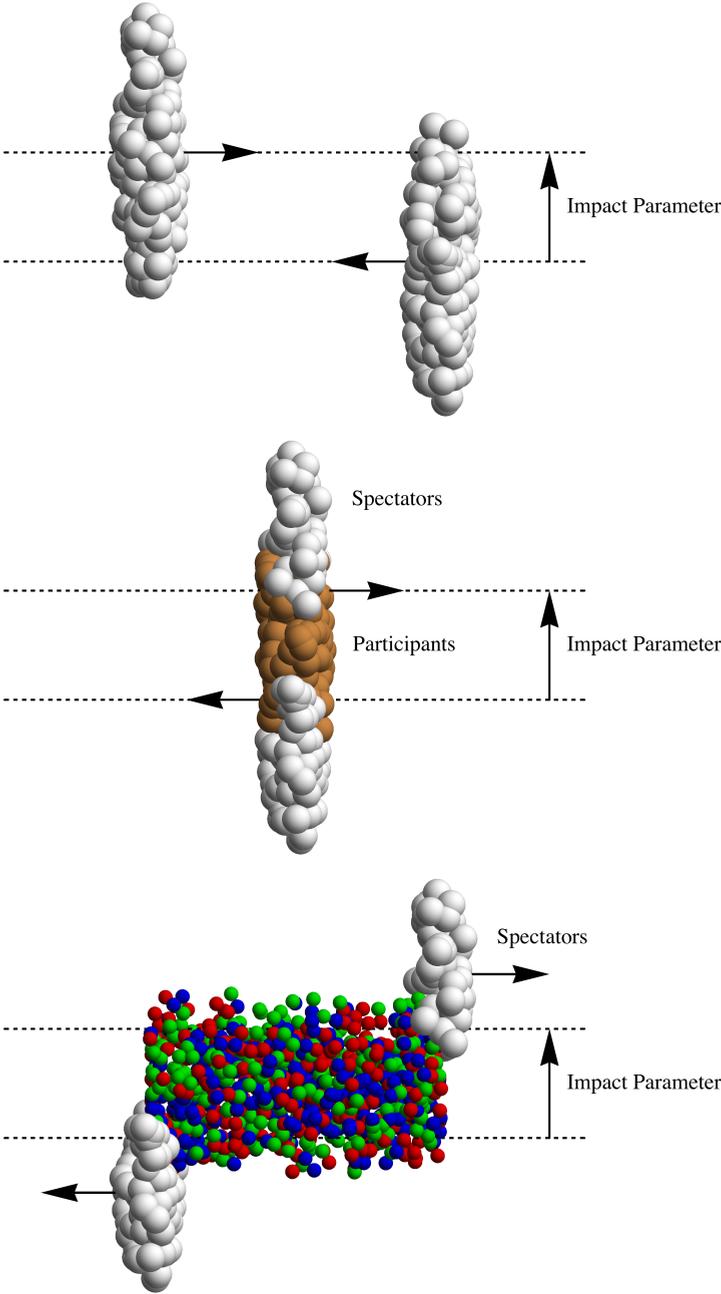


Figure 2.1.: Simple picture of a heavy-ion (Pb–Pb) collision, view facing the beam axis.

As we discuss in Sect. 2.6.2, the exact nature of this phase transition is still debated. The interacting hadrons can collide elastically, only exchanging momentum and energy, or they can collide inelastically, producing new particles. The rate of inelastic collisions is predicted to decrease much faster than the elastic ones [27], which means that at a certain time the hadrochemical composition is fixed. The time at which the hadrochemical composition is fixed, is called the *chemical freeze-out*, and it denoted by  $\tau_{\text{ch}}$ . After the chemical freeze-out the hadrons still undergo elastic collisions, and the momentum distributions of the individual hadron species can still change. Eventually the hadron gas becomes so dilute that also the elastic collisions cease. This is called the *kinetic freeze-out*.

All the interactions between the hadrons produced in a heavy-ion collision are referred to as *final state interactions* (FSI). In order to extract information about the deconfined phase, it is important to consider the impact that FSI have on the produced hadrons and on the correlations between them.

## 2.3. Hydrodynamics and Collectivity

As mentioned previously, it is believed that the hot QCD matter created in a heavy-ion collision thermalizes very quickly. If this is the case, then it makes sense to describe the system in terms of thermodynamic quantities, and its evolution in terms of (relativistic) hydrodynamics. One particularly simple implementation of this idea is the Bjorken model [29]. This model makes several approximations to the equation of state and the equations of motion (perfect fluid, one dimensional expansion only, boost invariance, etc.), which makes it possible to find analytic solutions. Modern hydrodynamic calculations are done numerically. They allow for a finite viscosity, use an equation of state taken from Lattice QCD (see Sect. 2.6) and do not assume boost-invariance [27, 32]. These models are very successful in predicting identified particle spectra up to  $p_{\text{T}} \sim 2.0 \text{ GeV}/c$ . In this section we first highlight the Bjorken model, as it is a simple model that provides an estimate of the energy density reached in heavy-ion collisions. After that we discuss the concept of *anisotropic flow*, which is the study of anisotropies in the particle production, caused by anisotropies in the density of the hot QCD matter.

### 2.3.1. Bjorken Model

The Bjorken model [29] aims to describe the particle production at mid-rapidity ( $y \sim 0$ ). It starts from the assumption that directly after the collision, when the distance between the receding pancakes does not exceed the nuclear size, the longitudinal velocity of the produced matter between the pancakes can be approximated by  $v_z = z/t$ . Using this, the initial energy density can be estimated as:

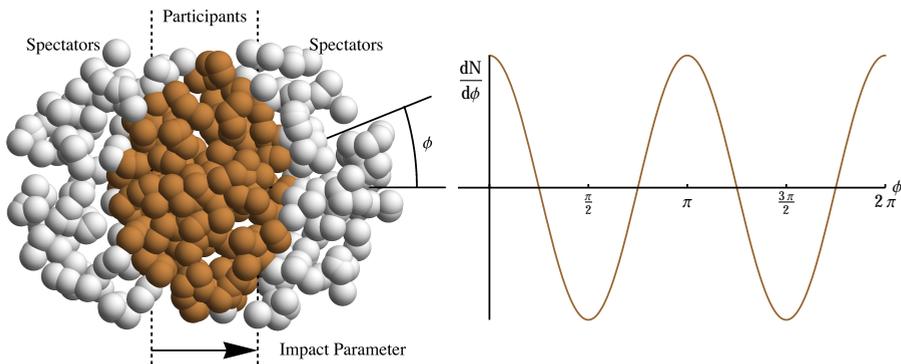
$$\epsilon = \frac{1}{c\tau_{\text{th}}A} \frac{dE_T}{dy}, \quad (2.5)$$

where  $y$  denotes rapidity and  $\tau_{\text{th}} \sim 1 \text{ fm}/c$  is the thermalization time. For the most central Pb–Pb collisions at  $\sqrt{s_{\text{NN}}} = 2.76 \text{ TeV}$ , it was measured by the CMS collaboration that  $dE_T/d\eta \approx 2 \text{ TeV}$  [33]. To transform the pseudo-rapidity to rapidity, knowledge about the identity of the particles is required. Using HYDJET 1.8 it was estimated that the Jacobian is  $d\eta/dy \approx 1.09$  at mid-rapidity. Together with the assumption of an thermalization time of  $\tau_{\text{th}} \sim 1 \text{ fm}$ , and an overlap area of  $A = \pi \times (7 \text{ fm})^2$ , the initial energy density is approximately  $\epsilon_i \approx 14 \text{ GeV}/\text{fm}^3$ . This exceeds the energy density at which the deconfining phase transition is expected to happen by an order of magnitude [34].

### 2.3.2. Anisotropic Flow

In the first panel of Fig. 2.2, we show the participants and the spectators as seen from along the beam pipe. In non-central collisions geometry of the system of participants is not circular, but instead has an almond shape. Due to the resulting anisotropic pressure gradient, particles produced in the direction of the impact parameter  $\mathbf{b}$  on average carries more momentum than the particles produced in perpendicular direction. This means that at similar momentum, the relative particle yield in the direction of the impact parameter is enhanced, as is illustrated in the second panel of Fig. 2.2. This effect is called *elliptic flow*. An observable sensitive to elliptic flow is the second order Fourier coefficient of the particle yield as a function of the angle with respect to the impact parameter  $\mathbf{b}$ .

If the matter distribution had a smooth almond shape, then the higher order Fourier coefficients of the particle yield would be very small. In reality however, the initial matter density distribution fluctuates significantly from event to event, causing also higher order harmonics to be non-zero.



**Figure 2.2.:** Simple picture of a heavy-ion (Pb–Pb) Collision, view along the beam axis.

The simplest model that can be used to study these initial density fluctuations is known as the *Glauber Model* [35, 36]. Within this model, the colliding nuclei are characterized by a distribution function for the nucleons.

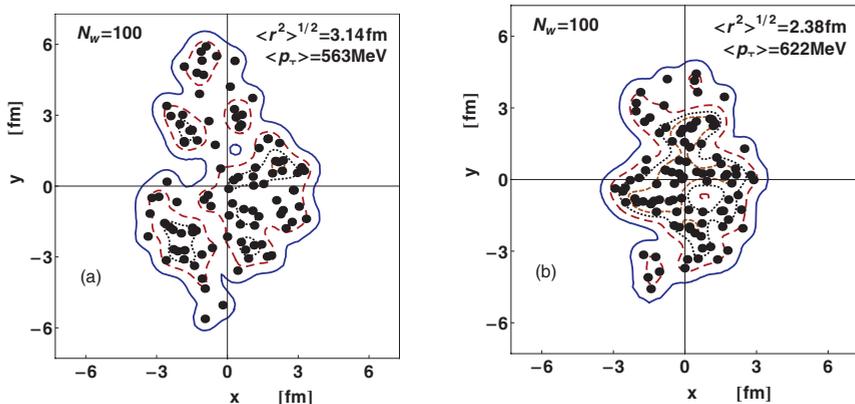
For large nuclei a parameterization that is found to agree well with experiments is the Woods-Saxon function [37], which is given by:

$$\rho_A(r) = \frac{\rho_0}{A(1 + \exp [(r - r_0)/a])}, \quad (2.6)$$

where  $A$ ,  $a$ ,  $\rho_0$  and  $r_0$  are parameters taken from experiment. Given the nucleon-nucleon inelastic cross section and an impact parameter, one can use Monte Carlo methods to calculate the probability that a nucleon interacts at least one time inelastically. These are the same nucleons that we previously referred to as participants, but the term *wounded nucleons* is often used interchangeably<sup>2</sup>.

In Fig. 2.3 two examples are given of wounded nucleon distributions. Note that even though the number of wounded nucleons is the same, which corresponds to a similar collision centrality, the density profiles are very different. The result of these density fluctuations, and the resulting non-trivial pressure gradients, is that the impact parameter is not the only direction in which particle production is enhanced. As a consequence, the

<sup>2</sup>In Figs. 2.2 and 2.3, the coloring of the spectators was determined by whether the nucleon is in the almond-shaped overlap region instead of whether it interacted. This is an approximation, however, due to the high packing density of the nucleons it is a reasonable one.



**Figure 2.3.:** Wounded nucleon distributions in the plane perpendicular to the beam axis with  $N_W = 100$ , generated by GLISSANDO [38]. Figures taken from Ref. [39].

particle yield can be described using a Fourier decomposition [32]:

$$\frac{dN}{d\phi}(p_T, \eta) \propto 1 + 2 \sum_{n=1}^{\infty} v_n \cos(n(\phi - \Psi_n)), \quad (2.7)$$

where  $\Psi_n$  is called the *flow angle*, and the Fourier coefficients  $v_n$  are often referred to as *flow coefficients*. While  $\Psi_2$  corresponds approximately to the direction of the impact parameter, the angles  $\Psi_n$  with  $n > 2$  are randomly distributed from event to event.

Explicitly referring to  $v_n$  and  $\Psi_n$  as “flow” parameters is suggestive that their values are completely determined by hydrodynamic flow. While the consensus is that this is certainly the largest contribution, it should be clear that there are also other effects that are not related to hydrodynamic flow (collectively referred to as *non-flow effects*), that can cause correlation structures. Examples of non-flow effects are jets (see Sect. 2.6.1), resonances, final state interactions, etc. There are several ways to perform  $v_n$  measurements, some of which are designed to be less sensitive to non-flow effects. These include higher order cumulants [40], or the analysis of particle pairs with a large separation in  $\eta$ . The latter is of importance for us, as we use this approach to single out the jet-peak contribution from the total particle yield.

## 2.4. Statistical Models of Particle Production

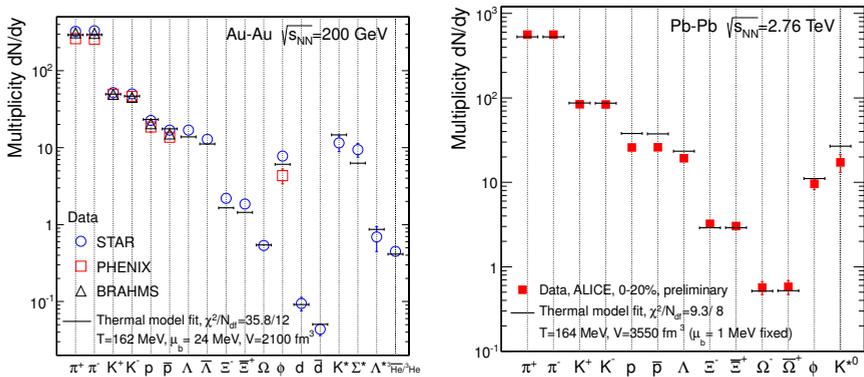
After the QCD matter cools down to the critical temperature it hadronizes. Without considering the microscopic mechanism of hadronization for the moment, one can study the thermal properties of the hadrons produced in a heavy-ion collision. If indeed these hadrons are created from a thermalized medium of quarks and gluons, then one would expect also the hadron gas to be in thermal equilibrium [41]. Under this assumption, the number of produced particles of species  $i$  can be written as:

$$N_i = g_i V \int \frac{d^3p}{(2\pi)^2} \left[ \exp \left( \frac{E_i(p) - \mu_i}{T} \right) - \epsilon \right]^{-1}, \quad (2.8)$$

where  $V$  is the volume of the fireball (QCD medium),  $\mu$  is the chemical potential,  $T$  is the temperature and  $\epsilon = -1, 0, 1$  indicates the statistics of the particles, i.e., Fermi-Dirac, (classical) Boltzmann and Bose-Einstein respectively [27, 41, 42]. In practice, one usually determines the most probable temperature and baryon chemical potential using a  $\chi^2$  fitting procedure to the measured particle ratios.

In Fig. 2.4 we show a comparison between the statistical model fits and data measured at RHIC (left panel) and LHC (right panel). In the comparison with the data the baryon and anti-baryon abundances are somewhat overestimated, most notably the (anti-)protons at LHC. In fact, in the comparison with the LHC data the (anti-)protons were excluded from the fit in order to improve the fit. One explanation for this overestimate is that baryon abundances still decrease somewhat after freeze-out due to hadronic final state interactions [42, 43].

Interestingly, the temperature found at both RHIC and LHC is  $T = 163$  MeV, which is only slightly below the temperature at which the phase transition is predicted to occur by Lattice QCD. On the other hand it has been shown that also the hadron production in  $e^+e^-$  [44] and pp [45] collisions is well reproduced using a statistical model. The temperature that was obtained is about  $T = 170$  MeV, and does not seem to depend much on the center-of-mass energy. The details of the statistical model used are slightly different from the model used in nuclear collisions. Most notably the canonical ensemble is used as opposed to the grand canonical ensemble, and a universal factor  $\gamma_s$  has to be introduced to account for the production of strange particles being lower than expected from a thermalized system.



**Figure 2.4.:** Particle multiplicities measured at RHIC (left panel) and at LHC (right panel), fitted with a statistical model described in Ref. [41]. Figures taken from Ref. [42].

An interesting recent idea of how to explain this universal hadronization temperature is that hadronization takes place through a mechanism that is very similar to Hawking-Unruh radiation [44]. Within this framework the strangeness suppression can be naturally accounted for by noting that the causal region in which particles can be produced is much larger in nuclear collisions than in collisions of elementary particles [46].

## 2.5. The Standard Model

In the preceding sections we saw that using very general assumptions and effective models, we could understand some basic properties of the hot QCD matter created in heavy-ion collisions. To improve on these models however, a better understanding of the microscopic theory of quarks and gluons is needed. In this section we briefly describe the Standard Model, which is the most fundamental description of subatomic physics that we have to date. As mentioned in Chapter 1, the most important part for understanding the behavior of quarks and gluons is QCD, which we discuss more extensively in Sect. 2.6.

The Standard Model is a quantum field theory, in which matter is described by spin- $\frac{1}{2}$  fermions, and all the known forces between matter (except for gravity) as mediated by spin-1 bosons. Within the Standard Model

there are two types of matter particles, those which couple to the previously discussed strong nuclear force (quarks), and those which do not (leptons). There are also three forces, the electromagnetic force which is mediated by photons, the weak nuclear force which is mediated by the  $W^\pm$  and  $Z$  bosons and the strong nuclear force, mediated by a family of eight gluons.

The structure of the Lagrangian of the Standard Model follows naturally if one starts from the free field equations of the matter fields, and then demands that these fields are locally invariant under so-called “gauge transformations”. To realize this local gauge invariance, additional bosonic fields have to be introduced, which correspond exactly to the force mediators for the three fundamental forces. For this reason these fields are collectively known as “gauge fields”, and the Standard Model is called a “gauge theory”.

Despite the unprecedented success of the Standard Model as a predictive theory, most physicists believe that the Standard Model provides only an approximate description of subatomic physics, valid only at limited energy and length scales. One of the reasons for this belief is the large number of seemingly arbitrary parameters in the Standard Model, for example the masses of the quarks and leptons [47], which have to be taken from experiment. Another reason is that the Standard Model does not include gravitational interactions between particles. The impact of gravity on earth-bound collision experiments is negligible, however, particle dynamics close to very heavy objects such as black holes are expected to be affected [48]. Finally, observations of the large-scale structure of the universe do not coincide with predictions from General Relativity, given the amount of observed mass and energy in the universe [49].

Much theoretical effort today is put into addressing these issues, and into the development theories beyond the Standard Model. One widely researched example is supersymmetry, where in its simplest implementation an additional bosonic/fermionic degree of freedom is added for every fermionic/bosonic degree of freedom in the standard model, see for example Ref. [50]. Another approach questions the assumption of the Standard Model that its fundamental particles are described as point-like particles, and instead models them as very small strings [51]. While these studies have had some theoretical successes, it has not been possible so far to find a solution that reduces to the Standard Model in the relevant limits.

## 2.6. Quantum Chromodynamics (QCD)

The part of the Standard Model that describes the strong force is called Quantum Chromodynamics (QCD). Its gauge group is  $SU(3)$ , which corresponds to rotations in “color space”. The number of different gluons (8) corresponds to the number of generators of the gauge group. The Lagrangian density of the quarks and the gluons is given by<sup>3</sup>:

$$\mathcal{L}_{\text{QCD}} = \sum_f \bar{\psi}_f \left[ i\gamma^\mu \partial_\mu - g_s \gamma^\mu \sum_a A_\mu^a \frac{\lambda_a}{2} - m_f \right] \psi_f - \frac{1}{4} \sum_a F_{\mu\nu}^a F_a^{\mu\nu}, \quad (2.9)$$

where  $\psi_f$  represents the quark fields of mass  $m_f$  and flavor  $f$ , with  $f \in \{u, d, s, c, b, t\}$ ,  $\gamma^\mu$  are the Dirac matrices,  $g_s$  is the QCD coupling,  $A_\mu^a$  represents the gluon fields of the eight different types of gluons ( $a = 1, \dots, 8$ ),  $F_{\mu\nu}^a$  is the field strength tensor of the gluon field<sup>4</sup> and  $\lambda_a/2$  are the generators of the group  $SU(3)$ .

The most straightforward way to calculate the expectation value of observables (correlation functions, scattering amplitudes, etc.) in a quantum field theory is by means of a path integral, i.e.:

$$\langle \mathcal{O} \rangle = \frac{\int \mathcal{D}\phi_i \mathcal{O}(\phi_i) \exp\{i \int d^4x \mathcal{L}\}}{\int \mathcal{D}\phi_i \exp\{i \int d^4x \mathcal{L}\}}, \quad (2.10)$$

where  $\phi_i$  are some fields,  $\mathcal{O}(\phi_i)$  is an observable depending on those fields, and the Lagrangian density  $\mathcal{L}$  describes the dynamics of those fields. For non-interacting theories, the terms in the Lagrangian density are at most of order 2, making this integral exactly solvable. Interacting theories, however, contain terms of higher order, rendering the path integral analytically unsolvable. An example is the term  $g_s \bar{\psi}_f \gamma^\mu A_\mu^a \frac{\lambda_a}{2} \psi_f$  in Eq. (2.9) describing the quark-gluon coupling. If the coupling  $g_s$  is small, a perturbative expansion in orders of the coupling gives a very good approximation of the solution. If  $g_s$  is large, then solutions to Eq. (2.10) can only be found numerically.

There are various ways to measure the coupling of QCD. One of the simplest ways is through the measurement of the three-jet fraction  $f_3$  in  $e^+e^-$  collisions. When an electron and a positron annihilate, a short-lived virtual

<sup>3</sup>Quarks also couple to the electromagnetic and weak nuclear force, but those terms are omitted here.

<sup>4</sup> $F_a^{\mu\nu} = \partial^\mu A_\nu^a - \partial^\nu A_\mu^a - g_s f_{abc} A_b^\mu A_c^\nu$ , where  $f_{abc}$  are the structure constants of the group  $SU(3)$ .

photon is created, which subsequently decays into either a lepton and an anti-lepton ( $l\bar{l}$ ), or a quark and an anti-quark ( $q\bar{q}$ )<sup>5</sup>. These two processes are easily distinguishable experimentally, because unlike the leptons, the quarks cannot exist as free particles. They instead undergo a process known as *fragmentation*, in which they turn into a collimated spray of hadrons, also known as *jets*.

Different experiments studied  $e^+e^-$  collisions at various center-of-mass energies, and they observed events with  $n \geq 2$  jets. At sufficiently high center-of-mass energy, it was observed that the relative frequency of  $n$ -jet events decreased with  $n$ . If we assume that each jet originates from one parton, then we may identify the di-jet events with the process  $e^+e^- \rightarrow q\bar{q}$ , the three-jet events with the process  $e^+e^- \rightarrow q\bar{q}g$ , etc. If we furthermore assume that the QCD coupling is sufficiently small to justify a perturbative approach, the differential cross section for the latter process is given by:

$$\frac{1}{\sigma_0} \frac{d^2\sigma}{dx_1 dx_2} = \frac{2\alpha_s}{3\pi} \frac{x_1^2 + x_2^2}{(1-x_1)(1-x_2)}, \quad (2.11)$$

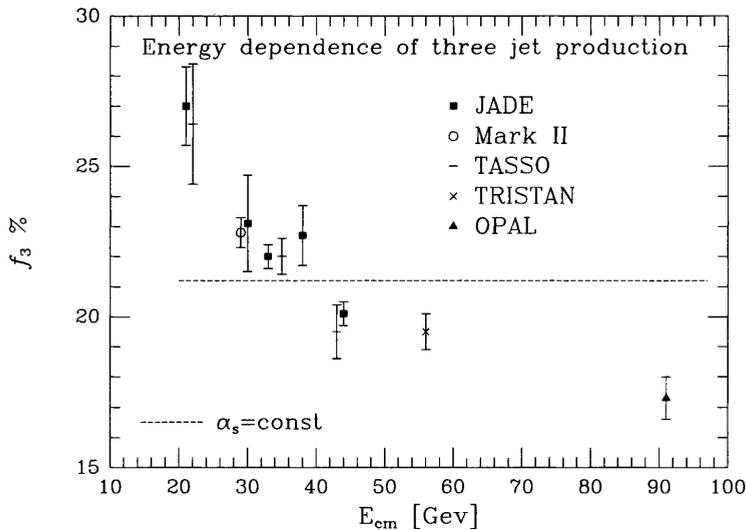
where  $\alpha_s = g_s^2/4\pi$ ,  $x_{1,2}$  denotes the energy fraction of the two quarks, i.e.,  $x_{1,2} = 2E_{1,2}/\sqrt{s}$ , and the energy fraction of the gluon  $x_3$  can be calculated from energy conservation, i.e.,  $2 = x_1 + x_2 + x_3$ . It is important to note that this cross section is divergent when  $x_{1,2} \rightarrow 1$ , which corresponds to one of the quarks radiating a soft collinear gluon. Note also that this singularity does not pose a problem for the definition of the three-jet cross section  $f_3$ , since the jets from two collinear partons are experimentally identified as one jet. In other words, we only call an event a three-jet event if the angle between the three jets is sufficient, and  $x_{1,2}$  are not close to one.

From the quark-gluon interaction term in Eq. (2.9) we know that the probability of gluon radiation is proportional to the QCD coupling. As shown in Fig. 2.5, measurements of the three-jet fraction  $f_3$  show a steady decrease with the center-of-mass energy of the collision system. This means that QCD is a strongly coupled when probed at low energies, but weakly coupled at high energies, providing direct evidence for asymptotic freedom. It also explains why the probability of radiating a soft (low energy) gluon cannot be calculated with perturbative methods.

In the following two subsections, we highlight two applications of QCD. First, in Sect. 2.6.1 we give a brief overview of hadronization in elementary

---

<sup>5</sup>To avoid complications, we only consider collision systems with  $\sqrt{s} \ll M_{Z^0}$ . Above this energy different decay channels become possible and in fact become dominant.



**Figure 2.5.:** Observed three-jet rate ( $f_3$ ) in  $e^+e^-$  collisions [7, 52].

particle collisions such as  $e^+e^- \rightarrow$  hadrons. The concepts developed in this section are used later when we discuss ways in which hadronization is modified by the presence of the hot QCD matter. Then, in Sect. 2.6.2 we briefly discuss Lattice QCD, which aims at calculating the thermal properties of QCD matter by numerically solving a discretized version of QCD. As this is a purely numerical approach, there is no problem with exploring the low energy (strong coupling) limit of QCD.

### 2.6.1. Hadronization and Fragmentation Functions

Let us now look a bit closer at the mechanism of hadron production in  $e^+e^-$  collisions. As we discussed before, the hadrons produced in this process are produced in  $n \geq 2$  collimated jets, where  $n$  corresponds directly to the number of partons created at large angles. If the center-of-mass energies of the collision are sufficiently high, then perturbative methods can be used to calculate the first stage of the hadronization process, where the initial partons branch off other partons. The final step of the process, where hadrons are created from the partons, involves low-energy processes, and can therefore only be treated in a phenomenological way.

Factorization theorems show that the cross section for the process  $e^+e^- \rightarrow h + X$  can be written as the convolution of the partonic cross section  $d\sigma_i/dy$ , calculable using perturbative methods, and a non-perturbative *fragmentation function*  $D_i^h(z, Q^2)$ , which gives the probability for a parton  $i$  to create a hadron  $h$  with a momentum fraction  $z$  at energy scale  $Q^2$  [7], i.e.:

$$\frac{1}{\sigma_{\text{tot}}} \frac{d\sigma_{e^+e^- \rightarrow h+X}}{dx}(x, s) = \sum_i \int_x^1 \frac{dz}{z} D_i^h(z, s) \frac{d\sigma_i}{dz}(x/z, s). \quad (2.12)$$

The energy scale in this expression is identified with the center-of-mass energy, and  $x$  denotes the energy fraction of the produced hadron, that is,  $x = 2E_h/\sqrt{s}$ . The fragmentation function is not calculable, however, it can be determined by comparing measured hadronic cross sections with Eq. (2.12), see for example Refs. [53, 54]. In order to use  $e^+e^-$  data with a wide range of center-of-mass energies, one needs to be able to relate fragmentation functions at different energies. This is done by using the so-called DGLAP<sup>6</sup> evolution equations [55]:

$$Q^2 \frac{\partial}{\partial Q^2} D_i^h(x, Q^2) = \sum_j \int_x^1 \frac{dz}{z} \frac{\alpha_s}{2\pi} P_{ji}(z, \alpha_s) D_j^h(x/z, Q^2), \quad (2.13)$$

where  $P_{ji}(z, \alpha_s)$  is called the *splitting function*, and it gives the probability that a parton of type  $i$  branches to a parton of type  $j$  with momentum fraction  $z$ . Splitting functions can be calculated perturbatively, in orders of  $\alpha_s$ . Intuitively, this equation can be understood by realizing that by increasing the energy, the parton is probed at a higher resolution. At this increased resolution, additional collinear partons can be resolved, that “dress” the original parton.

In Sect. 2.7 and 2.8 we discuss hadronization in the presence of a hot QCD medium, by generalizing the concepts that were introduced in this section.

### 2.6.2. Lattice QCD and the QCD Phase Diagram

To gain insight in the phase diagram of QCD matter, one needs to calculate its partition function. This can be done by evaluating a path integral similar to the one in Eq. (2.10), where the time coordinate is rotated to the imaginary

<sup>6</sup>Dokshitzer, Gribov, Lipatov, Altarelli, Parisi.

axis<sup>7</sup>, and interpreted as inverse temperature  $\beta$ . In the most interesting regions of the phase diagram, non-perturbative effects dominate, hence we cannot determine the partition function analytically. Fortunately, due to the Boltzmann factor  $\exp\{-\beta E\}$ , most configurations do not contribute much to the partition function, which justifies a Monte Carlo approach [56].

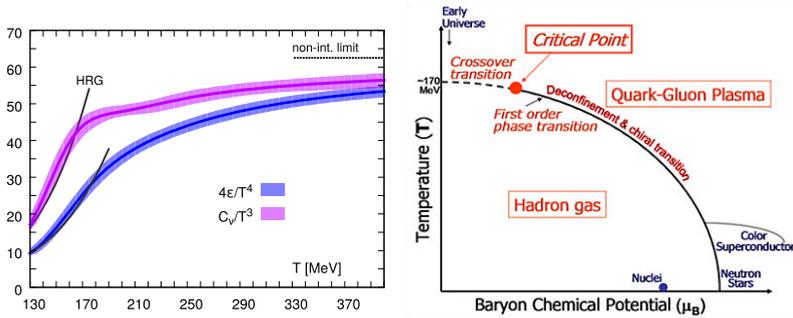
In this approach, pioneered by Kenneth Wilson in the 1970s [57, 58], the QCD Lagrangian in Eq. (2.9) is replaced by a discretized version on a lattice with lattice spacing  $a$ . By “measuring” observables on a lattice with decreasing  $a$ , one can get a good idea of the behavior of those observables in the continuum limit. The advantage of Lattice QCD is that there is in principle no restriction on the energy scale at which QCD can be studied. The discretization procedure introduces a natural largest energy scale, which is inversely proportional to  $a$ , but this can be chosen as small as computer resources and patience allow. The problems that can be studied using Lattice QCD are limited to the equilibrium properties of QCD under given thermodynamical circumstances. Examples are: the mass spectrum of hadronic states, the static potential between quarks, and the equation of state (EoS) [59].

In the first panel of Fig. 2.6 we show the energy density  $\epsilon$ , scaled with  $T^4$ , and the specific heat  $C_V$ , scaled with  $T^3$ , calculated at baryon chemical potential  $\mu_B = 0$ . For an ideal, non-interacting gas of quarks and gluons at zero chemical potential, the energy density is proportional to  $g_{\text{QGP}} \frac{\pi^2}{30} T^4$  [27], where  $g_{\text{QGP}} = 37(47.5)$  for two (three) flavors<sup>8</sup>. The non-interacting limit indicated in the figure as a dashed line, and for increasing temperatures the EoS seems to approach this value. The steeply increasing value of  $\epsilon/T^4$  around  $T \sim 160$  MeV indicates that the effective degrees of freedom change, which is identified with the confining-deconfining phase transition. The EoS determined by Lattice QCD can be parameterized, and used in the previously discussed hydrodynamic models.

The order of the phase transition is still being debated, however, recent Lattice QCD calculations strongly favor a crossover at  $\mu_B = 0$  [61, 62]. On the other hand, at large  $\mu_B$  the transition is expected to be of first order. At non-zero  $\mu_B$  Lattice QCD calculations cannot be performed exactly, and the location of the critical point that joins the crossover with the first order phase transition can only be obtained approximately [56]. A situation with very high  $\mu_B$  and relatively low  $T$ , is expected to occur in neutron stars.

<sup>7</sup>This procedure is called a Wick rotation.

<sup>8</sup> $g_{\text{QGP}} = \frac{7}{8} \cdot 2_{q,\bar{q}} \cdot 2_{\text{spin}} \cdot 3_{\text{color}} \cdot n_f + 2_{\text{spin}} \cdot 8_{\text{color}}$ , where  $n_f$  stands for the number of flavors.

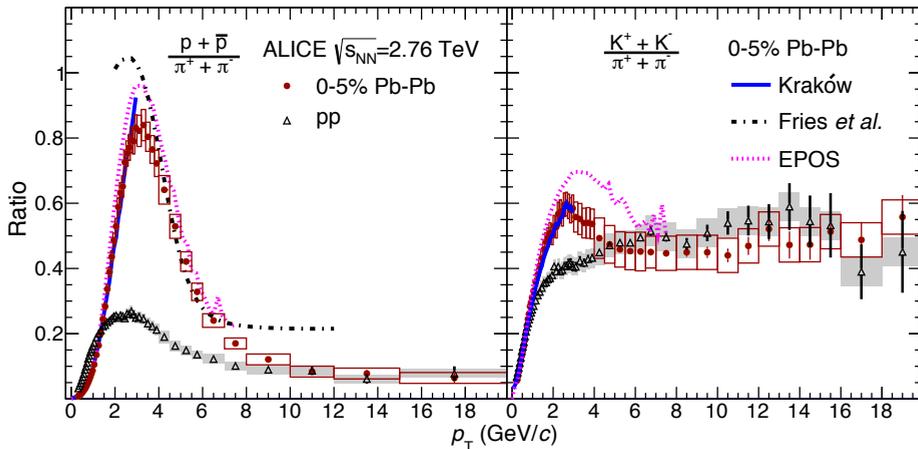


**Figure 2.6.:** The thermodynamics of QCD. First panel: QCD EoS at  $\mu_B = 0$ , figure taken from Ref. [59]. Second panel: QCD phase diagram, based on Lattice QCD ( $\mu_B \ll 1$ ), and on phenomenological input for larger chemical potential. Figure taken from Ref. [60].

Using the information presented, one can compose a phase diagram of QCD, which we show in the second panel of Fig. 2.6. Experimentally, one can make an attempt to probe the location of the critical point, by varying the beam energy. Such a measurement is also known as a *beam-energy scan* (BES). The system created in a nuclear collision in principle evolves along a “path” in the phase diagram. If the center-of-mass energy is sufficient, this path starts in the deconfined phase, after which it crosses the phase transition and converges towards the temperature and chemical potential associated with nuclear matter. By changing the center-of-mass energy of the nuclear collision, one can alter this path. Certain observables have been predicted to alter substantially, depending on the nature of the phase transition, and studying these observables as a function of the center-of-mass energy can shed light on the position of the critical point. Currently, there is a BES ongoing at RHIC and recent results are reviewed in Ref. [63]. Unfortunately, these results do not yet provide conclusive evidence for the location (or in fact for the existence) of the critical point.

## 2.7. Particle Recombination

There are clear experimental signatures that imply that the primary mechanism of hadron production in AA collisions is different from that in pp collisions. One of the clearest examples of such a signature is the enhanced baryon-to-meson ratio compared to a pp reference, that is observed in central



**Figure 2.7.:** Particle ratios, measured by the ALICE experiment [64]. The measurements are compared to calculations from a hydrodynamical model (Kraków) [67], a recombination model (Fries *et al.*) [68] and a hydrodynamical model which also describes jet-medium interaction (EPOS) [69].

AA collisions at intermediate  $p_T$ . This effect is known as the *baryon anomaly*. In Fig. 2.7 we show recent results from the ALICE experiment [64], however, the effect has been observed before at RHIC, both for strange and non-strange particles [65, 66]. It is generally thought that in the range  $p_T < 2.0$  GeV/ $c$  collective flow is the natural explanation of the baryon anomaly, but that in the range  $2.0 < p_T < 8.0$  GeV/ $c$  the effect is more likely caused by a different mechanism of hadron production becoming dominant.

In this section we discuss a relatively simple yet quite successful alternative model for hadronization, known as *recombination*. The essential mechanism of hadron production proposed in this model is that the quarks in the deconfined system cluster together (recombine) and form hadrons. We will start by describing the basics of this model, and then focus our attention on the comparison of this model with experimental observations. There are several different implementations of the recombination model, however here we will restrict ourselves to discussing the work of Rudolph Hwa and collaborators, as described in Refs. [70–74].

The starting point of the recombination model is to write the meson spectra as:

$$E \frac{dN^M}{dp_T} = \int \frac{dp_1}{p_1} \frac{dp_2}{p_2} F_{q_1 \bar{q}_2}(p_1, p_2) R_{q_1 \bar{q}_2}^M(p_1, p_2, p_T), \quad (2.14)$$

and the baryon spectra as:

$$E \frac{dN^B}{dp_T} = \int \frac{dp_1}{p_1} \frac{dp_2}{p_2} \frac{dp_3}{p_3} F_{q_1 q_2 q_3}(p_1, p_2, p_3) R_{q_1 q_2 q_3}^B(p_1, p_2, p_3, p_T), \quad (2.15)$$

where  $M$  and  $B$  refer to meson and baryon respectively, with energy  $E$  and transverse momentum  $p_T$ , the function  $F$  denotes the probability of finding 2 (or 3) quarks at momenta  $p_{1,2(,3)}$ , and the function  $R$ , which is called the *recombination function* (RF) denotes the probability of those quarks to recombine into a meson (baryon).

In the most recent versions of the recombination model, it is assumed that the medium consists of *thermal partons* which we denote by  $\mathcal{T}$  and *shower partons* which we denote by  $\mathcal{S}$ . All partons are allowed to recombine, regardless of their origin, i.e., in the case of a meson the distribution function  $F$  can be written as:

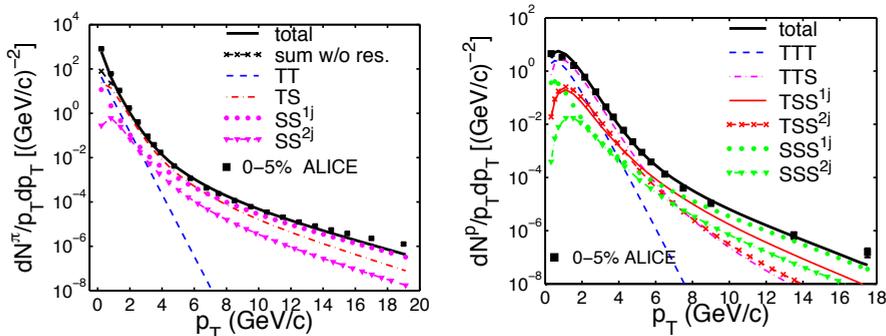
$$F = \mathcal{T}\mathcal{T} + \mathcal{T}\mathcal{S} + \mathcal{S}\mathcal{S}, \quad (2.16)$$

Thermal partons are assumed to be distributed according to a thermal distribution  $\mathcal{T} \propto \exp\{-p_T/T\}$ . The distribution of the shower partons is assumed to be of the form:

$$\mathcal{S}^j(p) = \int \frac{dq}{q} \sum_i \hat{F}_i(q) S_i^j(p, q), \quad (2.17)$$

where  $S_i^j(p, q)$  denotes the probability that a jet coming from a parton of type  $i$  and momentum  $q$  produces a parton of type  $j$  and momentum  $p$ , and  $\hat{F}_i(q)$  is the distribution of (semi-)hard partons at the surface of the hot QCD matter. The distribution  $S_i^j(p, q)$  can be determined from the previously discussed fragmentation functions. The function  $\hat{F}_i(q)$  can be further decomposed into an initial distribution of hard partons, convoluted with a suitable energy loss model, which takes into account that partons need to traverse a certain distance  $L$  through the hot QCD matter before they

reach the surface. For a complete description of the distribution of shower partons in heavy-ion collisions, we refer to Refs. [72, 74].



**Figure 2.8.:** Pion and proton yield measured by ALICE, and compared to the recombination prediction. Figures taken from Ref. [74]

While this formalism adequately reproduces RHIC data at  $\sqrt{s_{NN}} = 200$  GeV, in Ref. [74] it is shown that at the LHC energy  $\sqrt{s_{NN}} = 2.76$  TeV, shower partons become relatively more important, and one has to also consider recombination from shower partons originating from two separate showers to account for the observed spectra. In Fig. 2.8, the identified pion and proton spectra from Ref. [64] are compared to the recombination model. It is shown that hadrons containing at least one shower parton have a significant contribution to the entire  $p_T$  range. This suggests that the picture where low- $p_T$  hadrons are solely thermal is oversimplified.

## 2.8. Hadrochemistry in Jets

So far we have discussed hadron production in two different settings. We started with hadron production in jets that were essentially produced in a vacuum background<sup>9</sup>. Then, we observed that the hadrochemistry in AA collisions is significantly different at intermediate  $p_T$ , compared to pp collisions. We now pose the following question: how do jets created in hot QCD matter hadronize?

<sup>9</sup>The cleanest environment is an  $e^+e^-$  collision, however we assume that the main mechanism of particle production in a pp collision is also vacuum fragmentation.

There are several observables that show that the characteristics of the jets produced in heavy-ion collisions are significantly different from those created in vacuum, a phenomenon referred to as *jet quenching*. In Chapter 1 we already discussed several such observables, including jet- $R_{AA}$ , i.e., the suppression of the jet spectrum, and a significant di-jet imbalance in the most central collisions.

Jet energy loss in the QGP is modeled as being radiative (enhanced gluon radiation), collisional (with softer or thermal partons) or as a combination of both. These models successfully account for the aforementioned observables. On the other hand, it is likely that current models of jet quenching also cause a change in the hadronic composition of the jet [26]. In this section we touch upon two ideas that have been proposed to quantify the change in the jet hadrochemistry induced by the hot QCD medium. The first idea is to introduce a modified version of the vacuum splitting functions that were introduced in Sect. 2.6.1. The second idea is to relate the hadrons in the jet to hadrons that contain at least one shower parton. Neither of these ideas have been sufficiently developed at the moment to make quantitative predictions of the outcome of the measurement that we describe in this work, however it is likely that this is possible in the near future.

### 2.8.1. Modified Splitting Functions

In Sect. 2.6.1 we saw that the jet evolution was governed by the splitting functions  $P_{ij}(z, \alpha_s)$ , which could be calculated using perturbation theory. These calculations however are done in the QCD vacuum, and the splitting functions are expected to take a different form in the presence of a hot QCD medium. Several prescriptions for modifying the splitting function exist, however, instead of elaborating on the details, we argue that a simple but reasonable modification is sufficient to significantly alter the hadrochemistry of a jet [26]. The suggested modification is of the type:

$$P_{qq} = \frac{4}{3} \left\{ \frac{2(1 + f_{\text{med}})}{(1 - z)_+} - (1 - z) \right\}, \quad (2.18)$$

where the constant  $f_{\text{med}} \geq 0$  is the only modification, i.e.,  $f_{\text{med}} = 0$  corresponds to the vacuum case. Note that the presence of the additional factor increases the likelihood of a splitting, which is in accord with the observed soft-

ening of the jet spectrum<sup>10</sup>. The other splitting functions (i.e.,  $P_{qg}$ ,  $P_{gq}$ ,  $P_{gg}$ ) are modified in a similar way. For the details of the calculation we refer to the original paper and references therein.

In Fig. 2.9 we show the main results of the study. It is clear that the simple ansatz on the medium modified splitting functions has a significant effect on the particle composition of a jet, for a wide range of jet energies. It should be mentioned that this study is presented as an exploratory study, and the results should be seen as a qualitative rather than quantitative.

### 2.8.2. Shower-Thermal Recombination

One can also use the recombination framework discussed in Sect. 2.7 to study the hadrochemistry in jets. In Ref. [75] a formalism is proposed to treat two-particle correlations with a high- $p_T$  trigger particle. While the study is limited to the pion yield associated with a high- $p_T$  trigger pion, one can expect that similar calculations can be performed for other particle species. In this formalism, the pion yield associated with a high- $p_T$  trigger pion is assumed to contain at least one shower parton. Furthermore, Eq. (2.14) is generalized to a two-particle distribution, where one of the particles is a trigger particle, and was assumed to have a large transverse momentum, i.e.:  $p_T > 4$  GeV/ $c$ . The main results of this study are shown in Fig. 2.10. In this figure a good agreement is found with the data from the STAR collaboration, however it is also clear that a much more extensive study is warranted.

## 2.9. Did we discover the QGP?

In this chapter we set ourselves two goals: to give a brief introduction to the field of heavy-ion physics in general and to provide some theoretical motivation to the study of hadron production in jet-like structures in heavy-ion collisions. We hope that we succeeded in the latter, but the former is incomplete without mentioning the status of the discovery of the quark-gluon plasma (QGP).

Throughout this chapter we avoided using the term quark-gluon plasma, instead we used “hot QCD matter”. This term only implies that QCD is the correct theory to describe the matter created in relativistic AA collisions, and that the matter created has a temperature and energy density exceeding the

---

<sup>10</sup>There seems to be no intuitive explanation why the increased splitting directly leads to a change in jet hadrochemistry [26].

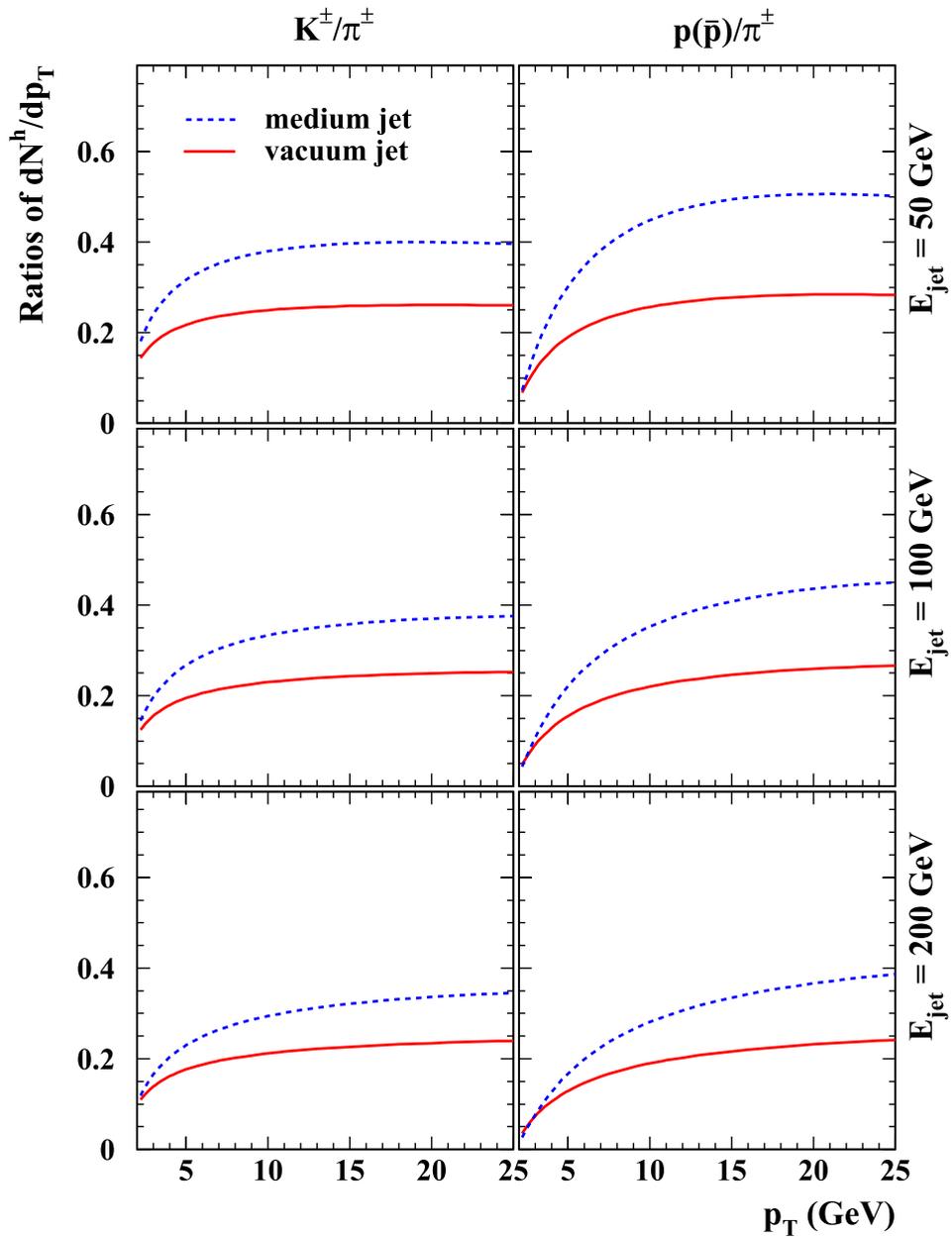
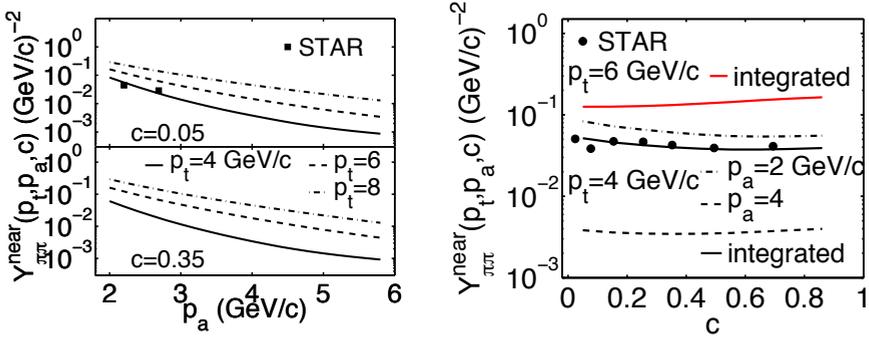


Figure 2.9.: Particle ratios in jets of several energies. Figure taken from Ref. [26].



**Figure 2.10.:** The pion yield associated with a high- $p_T$  trigger pion. In this figure  $c$  refers to centrality. Figures taken from Ref. [75].

critical values of the deconfining phase transition predicted by Lattice QCD. The latter is strongly supported by measurements of  $dN/dy$  at RHIC and LHC, combined with simple kinematic arguments presented in Sect. 2.3.1.

There is also strong evidence that the effective degrees of freedom of the matter created in relativistic AA collisions are not hadronic. The success of relativistic fluid dynamics in describing soft observables, using an equation of state (EoS) from Lattice QCD, suggests that the number of degrees of freedom is much larger than what one would expect from a hadron gas. Also, the success of the recombination picture points to the existence of deconfined partons in the produced matter. There are many more arguments, and for a comprehensive overview see for example Refs. [28, 76].

In order to discuss whether a QGP was created at RHIC or LHC we first need to establish what we exactly mean by QGP. Soon after the discovery of asymptotic freedom, physicists started exploring the possibility that under certain thermodynamic conditions quarks and gluons could exist as a gas of weakly interacting particles. It was in this context that the name QGP was first conceived in 1975 by Shuryak [77]. On the other hand, subsequent experimental observations at RHIC showed evidence for strong collective behavior of the hot QCD matter. This can only be explained when there is significant interaction between the relevant degrees of freedom.

For this reason, the hot QCD matter is much better described by a fluid model than by a weakly interacting gas. Comparisons of fluid models with the data, in fact suggest that the fluid has a very low viscosity, possibly close to the value that was recently found to be the lowest possible viscosity in

AdS/CFT [78]. Some authors<sup>11</sup> therefore prefer to use the term *sQGP*, where the *s* refers to “strong”, to make a clear distinction with the originally proposed weakly interacting QGP, while others caution against the introduction of imprecise new terminology [79].

The consensus is that the data at RHIC and LHC cannot be explained under the assumption that the relevant degrees of freedom are hadrons. On the other hand, the working definition of the QGP usually also includes the requirement of thermalization, which has not been experimentally established. While hydrodynamic calculations seem to favor an early thermalization time of  $\tau_{\text{th}} \approx 1 \text{ fm}/c$ , this cannot be directly measured. Furthermore, Lattice QCD also predicts that the deconfinement phase transition should occur almost simultaneously with the restoration of chiral symmetry, of which there is also no experimental evidence to date.

In conclusion, the current experimental evidence is sufficient to claim that a previously unknown type of matter has been produced in relativistic heavy-ion collisions. Some important questions regarding the details of this matter are however still unanswered, and warrant further investigation.

---

<sup>11</sup>Most notably Shuryak himself.



## Chapter 3.

# Experimental Setup

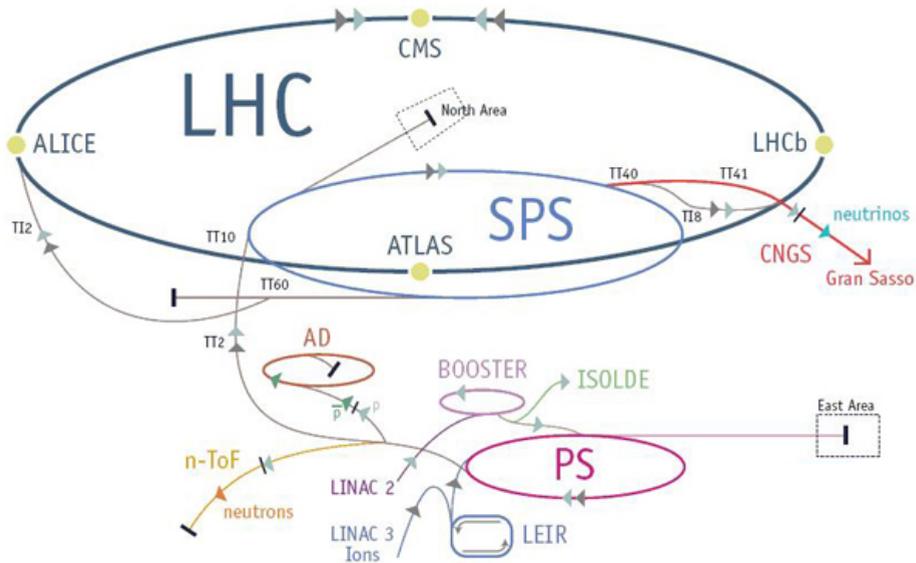
The ALICE detector is one of the four large detectors at the Large Hadron Collider (LHC) at CERN. It is a general purpose detector, with a strong focus on measurements aimed at elucidating the nature of the hot QCD matter that is expected to be created in central heavy-ion collisions. For this reason, two of the most important design criteria were: the ability to cope with the large track density  $dN/dy \sim 8000$  expected to be produced at the most central Pb–Pb collisions, and good PID capabilities at low to intermediate transverse momentum.

The rest of this chapter is built up as follows: after a short overview of the LHC, we proceed by describing the ALICE detector, emphasizing the parts that are most relevant for this work. We then finalize with a description of the techniques used to reconstruct events and tracks from the raw data.

## 3.1. The Large Hadron Collider (LHC)

The Large Hadron Collider is a particle accelerator at CERN, close to Geneva. It is built in a nearly circular tunnel with a circumference of approximately 27 km, and it was designed to create proton beams with an energy of 7 TeV, and lead (Pb) ion beams with an energy of 2.76 TeV per nucleon. As of 2014, the following collision systems were produced at the LHC: pp collisions at  $\sqrt{s} = 2.76, 7,$  and 8 TeV, p-Pb collisions at  $\sqrt{s_{NN}} = 5.02$  TeV, and Pb–Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV.

In Fig. 3.1 a schematic overview is shown of the LHC complex. Before particles enter the LHC, they are pre-accelerated by several other machines at the LHC complex. The first acceleration stage for protons happens at a



**Figure 3.1.:** Schematic overview of the LHC complex [80].

linear accelerator, called LINAC2, which injects them into the booster ring at an energy of 50 MeV. The booster increases the beam energy to 1.4 GeV, after which the protons are transferred to the Proton Synchrotron (PS), and subsequently the Super Proton Synchrotron (SPS). From the SPS, the protons are sent to the LHC at a beam energy of 450 GeV, and they are accelerated to the desired beam energy, with a maximum of 7 TeV. For a lead-ion beam the initial stages of acceleration are different from the proton beam. Lead ions are first injected into the Low Energy Ion Ring (LEIR) by LINAC3. From LEIR they are injected into the PS, and the rest of the sequence is the same as for protons.

## 3.2. The ALICE Detector

The ALICE detector consists of a large number of smaller sub-detectors, each being optimized for one or more specific task(s). The main tasks that need to be performed by these sub-detectors are:

- **Triggering** - Determining whether there actually was a collision<sup>1</sup>, and whether this collision has (one of) the properties that we are currently interested in. When a trigger fires, the rest of the ALICE detector is read out, and an *event* is recorded. There are triggers that only fire when a rare event occurs, for example the measurement of a jet with very large energy. In this analysis, however, we only make use of the so-called *minimum-bias trigger* (MB), which selects events with a minimal bias on the final state<sup>2</sup>. Triggering is discussed in more detail in Sect. 3.3.
- **Tracking** - Directly after recording, an event is nothing more than a collection of pixels and analog signals, measured by the various sub-detectors. Fitting algorithms are applied in order to reconstruct the precise location and time of the collision and to determine the momentum and pseudorapidity of the particles that were produced. These algorithms become more reliable as the number of measurements made along each particles track increases. Detectors that provide input to these algorithms are called *tracking detectors*. We come back to the methods used in primary vertex and track reconstruction in Sect. 3.4.
- **Event Characterization** - Certain characteristics of the event can be immediately deduced from the reconstructed tracks. Other characteristics, such as event centrality and the direction of the event plane (see Sect. 3.4) can be determined with greater accuracy using dedicated detectors.
- **Particle Identification (PID)** - Finally, there are several sub-detectors which perform measurements that are sensitive to the identity of a particle, the so-called *PID detectors*. The PID detectors of ALICE are designed to complement one another, so that the combined information of these detectors allows to distinguish between a number of different particle species in a wide range of momenta. To achieve this goal, the PID detectors mostly rely on different techniques, or are strategically located relative to the primary interaction point. The two detectors used

---

<sup>1</sup>There are also triggers which are designed to fire when a cosmic ray has been detected, and have therefore nothing to do with the occurrence of a collision. These triggers are, however, not directly important to us.

<sup>2</sup>The efficiency of the MB trigger is not 100% for collisions where only a few particles are produced.

in this analysis for PID are the Time Projection Chamber (Sect. 3.2.3) and the Time of Flight Detector (Sect. 3.2.4).

We focus in this section on how these main tasks are accomplished by the various sub-detectors of ALICE. The scope of this discussion is limited to the sub-detectors that were most important for our analysis, and for a more comprehensive description of the ALICE detector, its design specifications and its performance we refer to Refs. [81–84].

### 3.2.1. Overview

In Fig. 3.2, a schematic overview of the ALICE detector is shown, labeling all the sub-detectors. The largest structure in the ALICE detector is the L3 solenoid magnet. It provides a nearly homogeneous magnetic field of  $B = 0.5$  T parallel to the beam pipe, causing charged particles to travel on curved paths with radius  $R = p/(qB)$ . The sub-detectors enclosed within the solenoid magnet are often referred to as the central barrel detectors. Most detectors that are relevant for our analysis are central barrel detectors.

### 3.2.2. The Inner Tracking System (ITS)

The detector that is closest to the interaction point is the Inner Tracking System (ITS). It consists of six layers, and makes use of three different detection technologies. The inner- and outermost layers are located at  $r = 3.9$  cm and  $r = 43.0$  cm respectively, and it has a pseudorapidity coverage of  $|\eta| < 0.9$ . The two innermost layers are the Silicon Pixel Detector (SPD). The SPD is essential for determining the exact location of the collision, i.e., the primary vertex (see Sect. 3.4.2). As is discussed in Sect. 3.3, the SPD is also used as trigger detector for certain MB triggers. The next two layers of the ITS are the Silicon Drift Detector (SDD), and the outer two layers are the Silicon Strip Detector (SSD). The SDD and SSD are mainly used for tracking, but also for particle identification of very low momentum ( $p < 200$  MeV/ $c$ ) particles.

### 3.2.3. The Time Projection Chamber (TPC)

Surrounding the ITS is the Time Projection Chamber (TPC), which is the main tracking detector of the ALICE experiment. The TPC is a large cylindrical barrel with an inner radius of 85 cm, and an outer radius of 250 cm. It has a volume of  $90$  m<sup>3</sup>, and is filled with Ne-CO<sub>2</sub> gas. In the middle of the

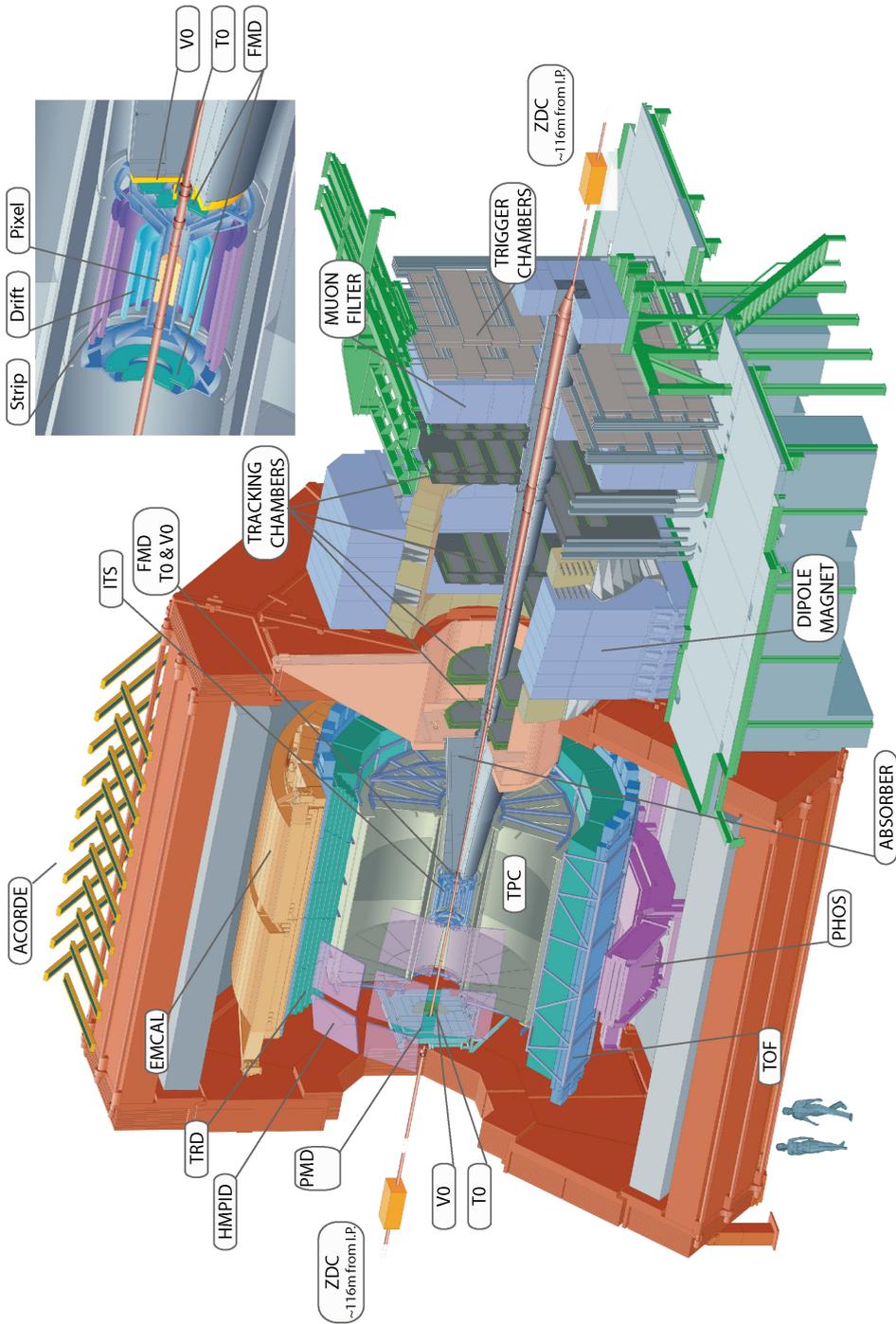
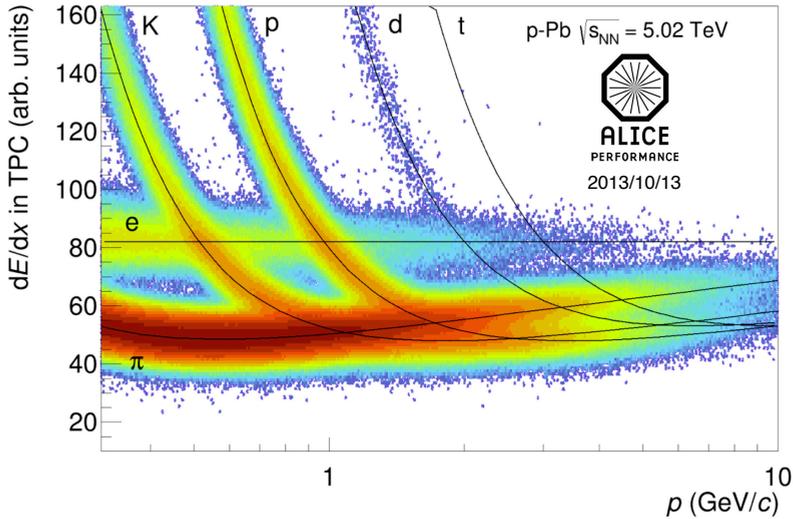


Figure 3.2.: Schematic overview of the ALICE detector.

TPC there is a central membrane, which is held at a potential difference of  $\Delta V \sim 100$  kV compared to the end plates. The working principle of the TPC is that charged particles which traverse the gas ionize the molecules they encounter. The freed electrons then drift to the end plates due to the large electric field in the TPC, thereby creating a projection of the trajectory on the end plates. An end plate consists of 18 different readout chambers, each covering an angular range of 20 degrees, and having 159 rows perpendicular to the radial direction. For particles produced within the pseudorapidity window of  $|\eta| < 0.9$ , the TPC can therefore record up to 159 points along its track. Due to this large number, the TPC is ideally suited to measure the momentum and pseudorapidity of the particles, as well as disentangle tracks in high-density events. In principle, the TPC can also track particles in a pseudorapidity interval  $0.9 < |\eta| < 1.5$ , however, with a decreased resolution. Since the TPC has a significant readout time, it is less suited for coping with very high luminosities.



**Figure 3.3.:** Specific energy loss measurement with the ALICE TPC detector. The expected energy loss for the different particle species is shown.

The TPC can also be used for particle identification, by measuring a particle's *specific energy loss*<sup>3</sup>. The specific energy loss of a charged particle

<sup>3</sup>Or *ionization energy loss*.

is defined as the energy that it loses per unit path-length ( $dE/dx$ ) due to electromagnetic interaction with the material it traverses. This can be determined by the TPC, since it is directly related to the number of electrons that are freed, and eventually drift to the end-plates. The TPC can perform up to 159 specific energy loss measurements per track. These measurements are not symmetrically distributed around the true value of the specific energy loss, instead the distribution exhibits a long *Landau tail* towards large energy loss. This is due to a small but significant probability that a particle transfers a lot of energy in a collision with one of the gas molecules. In practice, the specific energy loss is therefore estimated by a truncated mean, ignoring the highest values.

The measured specific energy loss can then be compared to the expected specific energy loss, which, for a heavy particle (much heavier than an electron), can be calculated using the Bethe-Bloch formula. The Bethe-Bloch formula only depends on material properties of the gas in the TPC, and on the velocity  $\beta = v/c$  of the particle:

$$\left\langle \frac{dE}{dx} \right\rangle \propto \frac{1}{\beta^2} \left[ \ln \left( \frac{2m_e c^2 \beta^2 \gamma^2}{I} \right) - \frac{\beta^2}{2} - \frac{\delta}{2} \right], \quad (3.1)$$

where  $I$  is the mean excitation energy of the gas,  $m_e$  is the mass of the electron, and  $\delta$  is a factor taking into account density effects [85, 86]. As the specific energy loss depends only on velocity, it can be combined with a momentum measurement, to estimate a particle's mass ( $mc^2 = pc/(\beta\gamma)$ ) and hence its identity.

In Fig. 3.3 we show an ALICE measurement of the specific energy loss as a function of momentum, with parameterizations of the expected energy loss of the most abundant particle species superimposed<sup>4</sup>. At small velocities ( $\beta \ll 1$ ) Eq. (3.1) falls off as  $\propto \beta^{-2}$ , whereas at large velocities it rises logarithmically. This causes the expected energy loss curves for different particle species to intersect, meaning that at specific momenta the separation power for two species drops to zero. For these momenta, the TPC measurement needs to be complemented with information from a different PID detector to create sufficient separation power. For large momenta ( $p > 5$  GeV/c), the separation between the most abundant hadrons ( $\pi$ ,  $K$ ,  $p$ ) is small, but nearly constant. This makes the TPC the most important detector for PID at large momenta.

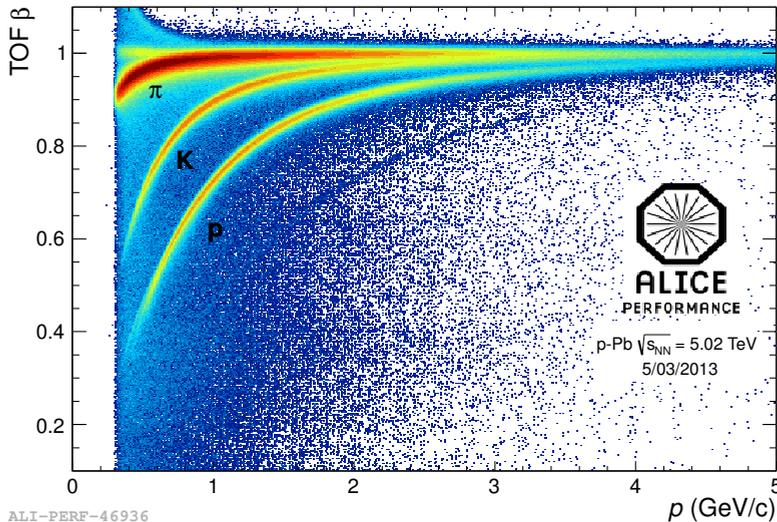
---

<sup>4</sup>Due to several experimental complications, the actual signal of the TPC is not exactly proportional to the expected energy loss, see for example Ref. [87].

The resolution of the  $dE/dx$  measurement depends mainly on the number of freed electrons that reach each TPC cluster. For this reason an increased gas pressure will improve the TPC's resolution, due to the increased number of gas molecules encountered by the particle. For the same reason, tracks which traverse more gas (large  $\eta$ ) also have a somewhat better resolution. We discuss the PID resolution of the TPC in some more detail in Sect. 4.1.

### 3.2.4. The Time Of Flight (TOF) Detector

Surrounding the Transition Radiation Detector (TRD), which does not play a major role in our analysis, is the Time Of Flight detector (TOF). The TOF detector is a cylindrical Multi-gap Resistive-Plate Chamber (MRPC), located at a radius of  $370 < r < 399$  cm, covering a pseudorapidity range of  $|\eta| < 0.9$ . It has a large number of readout pads ( $\sim 150,000$ ), each measuring  $2.5 \times 3.5$  cm, so that even high-density events have a reasonable occupancy rate of about 14% [83, 86]. Its main purpose is to endow TPC tracks with a time of flight measurement. The integrated path length of a track and the time of flight can be used in combination with a momentum measurement to determine the mass of the particle.



**Figure 3.4.:** Particle velocity measurement with the ALICE TOF detector. The pion, kaon, and proton bands are indicated.

Fig. 3.4 shows a velocity measurement, where the integrated path length was used as well as the time of flight measurement. Notice that a TOF measurement can only be performed for particles with momenta  $p \gtrsim 300$  MeV/ $c$ , as particles with a smaller momentum do not reach the TOF detector due to their small radius of curvature. For momenta  $p \leq 2.5$  GeV/ $c$ , the TOF separation power is sufficient to separate kaons and pions using a fitting procedure. Protons and kaons can be separated in the same way, up to about  $p \leq 4.0$  GeV/ $c$ . Note that these momentum ranges cover the crossing points of the Bethe-Bloch curves of the TPC in Fig. 3.3, so that an effective strategy would be to combine both measurements.

There are a significant number of measurements that do not fall on the bands. These measurements consist mostly of mismatched tracks, i.e., tracks where the incorrect hit in the TOF detector was associated with a TPC track, causing a faulty velocity measurement.

A TOF measurement consists of two independent time measurements, i.e., the time of the interaction  $t_{\text{int}}$  and the arrival time  $t_{\text{arr}}$ . To measure  $t_{\text{int}}$ , ALICE is equipped with a dedicated T0 detector, consisting of two Cherenkov rings at forward rapidities, one on each side of the interaction point [88]. An alternative way to measure  $t_{\text{int}}$  is to consider one TOF measurement  $j$ , and perform a minimization procedure, using the following  $\chi^2$ :

$$\chi^2(t_{\text{int}}; m_1, m_2, \dots) = \sum_{i \in \text{tracks} \setminus j} \frac{((t_{\text{arr},i} - t_{\text{int}}) - \Delta t_{\text{exp}}(m_i))^2}{\sigma_{\text{arr}}^2 + \sigma_{\text{texp}}^2}, \quad (3.2)$$

where  $i$  sums over all tracks except track  $j$ , and every particle is assumed to be a pion, kaon, or a proton. For physics analyses, both  $t_{\text{int}}$  measurements are taken into account [89].

The resolution of the total TOF measurement, i.e.  $\sigma_{\Delta t}$ , depends both on the uncertainty of the interaction time and arrival time measurements, i.e.:

$$\sigma_{\Delta t}^2 = \sigma_{t_{\text{int}}}^2 + \sigma_{t_{\text{arr}}}^2. \quad (3.3)$$

The uncertainty of the interaction time measurement decreases rapidly as more tracks are used in the measurement ( $\sigma_{t_{\text{int}}} \sim 1/\sqrt{n}$ ) [89]. High density events, such as (semi-)central Pb–Pb collisions, therefore have a significantly smaller uncertainty of the TOF measurement compared to pp collisions, where an event often has only a few reconstructed tracks.

The uncertainty of the arrival time is to a large extent hardware related. The contribution from the hardware to the uncertainty is  $\sim 80$  ps, and is

largely independent of the details of the collision system. On top of that, there is a contribution from the uncertainty of the energy that a particle loses before reaching the TOF detector. For reasons that are discussed in Sect. 4.2 there is some dependence of the arrival time resolution on the particle's velocity and pseudorapidity.

### 3.2.5. The V0 detectors

Surrounding the beam pipe, at a distance of  $z = 3.4$  m, and  $z = -0.9$  m from the nominal interaction point there are two scintillator counters installed, called the V0A and V0C. As is discussed in Sect. 3.4.1, one of the main functions of the V0 detectors is to determine the centrality of an event. In case of Pb–Pb collisions, both detectors are used, while for p-Pb only the V0A is used for this purpose. The V0 detectors can also be used to determine the direction of the impact parameter, also referred to as the *event plane angle*. Finally, they are used for triggering, and for distinguishing beam-beam interactions from beam-gas interactions.

## 3.3. Triggering

During data taking, the ALICE detector is programmed to respond to one or more trigger conditions. Every 25 ns the Central Trigger Processor (CTP) of ALICE decides whether (one of) the trigger conditions are met, based on the signals it receives from the various trigger detectors. After data taking, the recorded events are analyzed further, and additional filtering is applied in order to remove trigger noise and beam-gas interactions. In this work, only the minimum bias (MB) trigger is used. The definition of the MB trigger depends on the collision system and the luminosity. In general, when an interaction is expected to produce a small number of particles, or the luminosity is low, a more sensitive trigger definition is used.

The MB trigger for pp collisions is the most sensitive MB trigger, it fires when there is a signal in either V0 detector or in one of the layers in the SPD. For p-Pb collisions, it is defined as a signal in both V0 detectors. During the Pb–Pb data taking of 2010 three different MB hardware triggers were used, namely:

- A signal in both V0 detectors.
- Two or more signals in the SPD, and a signal in either V0 detector.

- Two or more signals in the SPD, and a signal in both V0 detectors.

depending on the run period.

For high-multiplicity events, the efficiency of the MB trigger is practically 100%. This holds for all but the most peripheral p-Pb and Pb–Pb events. The track multiplicity in pp events on the other hand can be very low, thus significantly reducing the trigger efficiency. Furthermore, different MB trigger definitions typically have a different sensitivity for non-diffractive (ND), single diffractive (SD), and double diffractive events, which together make up the set of inelastic events. For a comprehensive overview of the MB trigger efficiencies in pp collisions, see Ref. [90].

## 3.4. Offline Analysis

Most of the processing of the raw data requires significant computing resources, and is therefore postponed until after the data-taking has been completed. Here we briefly discuss the data processing steps that are essential for our work, namely, the determination of the collision centrality and the reconstruction of the collision vertex and the tracks.

### 3.4.1. Centrality Determination

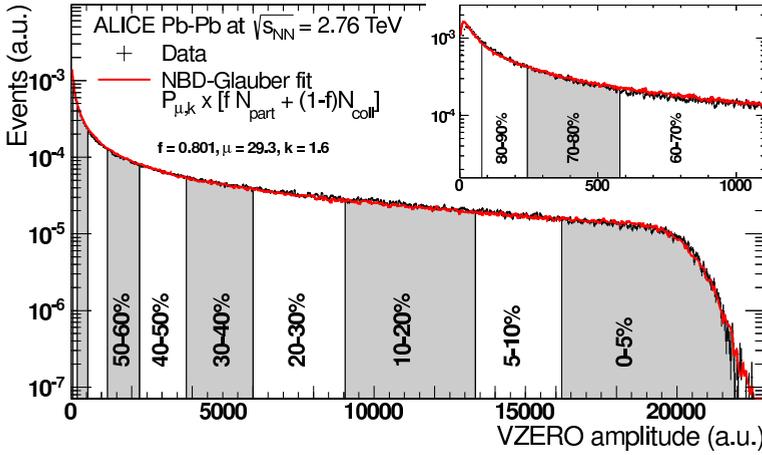
There are several detectors that can be used to determine the centrality of a collision. In this work we use the centrality as it is determined by the V0 detectors, while the other centrality measurements are only used to filter out faulty V0 measurements. As is shown in Fig. 3.5, a more central collision corresponds to a larger amplitude in the V0 scintillators. By fitting the distribution of the V0 amplitude with a Glauber model (see Sect. 2.3), one can relate the centrality percentile to the value of the impact parameter. The exact procedure is described in great detail in Ref. [31].

### 3.4.2. Track and Event Reconstruction

The first step in the reconstruction of an event is to estimate the location of the primary vertex. This is done by extrapolating *tracklets* from the SPD<sup>5</sup>, i.e., lines connecting two points from the SPD, and locating the position where most lines coincide.

---

<sup>5</sup>The term “tracklets” is not reserved exclusively for the SPD, and more generally refers to tracks consisting of a small number of points.



**Figure 3.5.:** Centrality determination in ALICE, using the V0 scintillators. Figure taken from Ref. [31]

After that, tracks are reconstructed from the points in the ITS and the TPC, using an inward-outward-inward procedure [91] based on a Kalman filter [92]. After the final inward fit has been performed, the momentum and pseudorapidity (curvature, position) of the track are determined. Using the fully reconstructed tracks, an improved measurement of the primary vertex is made. Using the precise location of the interaction vertex, the distance of closest approach (DCA) of all (extrapolated) tracks can be determined. This information can be used to filter out tracks produced in weak decays or tracks originating from material interactions, as these tracks typically have a large DCA.

The inward-outward-inward reconstruction procedure was designed to be suitable for the large track density environment of a central Pb–Pb collision. Efficiency studies indeed show no significant difference in track finding efficiency between pp and central Pb–Pb events [84].



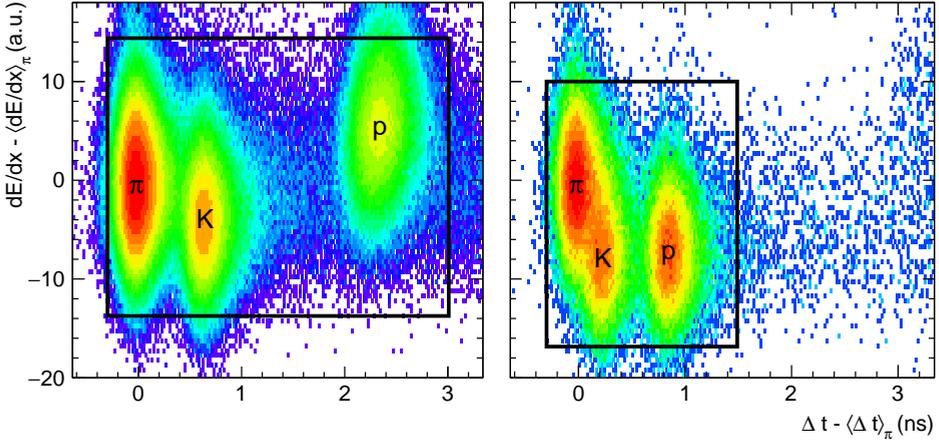
## Chapter 4.

# Particle Identification

In this chapter we develop a PID method based on the combined signal of the TOF and TPC detectors. Both the TOF and TPC detectors have been used individually to identify charged particles, however, using only one of these detector limits the  $p_T$ -range over which identification can be done. While a fit of the TOF signal can separate protons from pions and kaons up to 4.5 GeV/ $c$ , separating pions from kaons becomes difficult already at 3.0 GeV/ $c$  (see Sect. 3.2.4). The TPC detector on the other hand, can separate pions from kaons and pions from protons much better than kaons from protons (see Sect. 3.2.3).

As shown in Fig. 4.1, the signals of the TOF and TPC detector complement one another and the combined signal can be used to identify particles with a greater separation power and over a larger  $p_T$  range compared to when a single detector is used. To get the most accurate results and to be mostly independent of calibration issues, we chose to fit two dimensional distributions in (TOF, TPC) instead of using a cut on the detector signal. From these figures it is clear that at low  $p_T$  the TPC alone cannot distinguish pions from kaons, while the TOF provides good separation. At higher  $p_T$ , the TOF of pions and kaons becomes comparable, while the TPC has good separation.

There are a number of difficulties that need to be overcome when identifying particles through a fitting procedure. First of all, the expected distance between the peaks in the PID signal typically depends on the kinematic variables ( $p_T$  and  $\eta$ ). This means that the (TOF, TPC) distributions should only contain tracks selected from a narrow window in those variables, to avoid “smearing” of the peaks. Furthermore, the resolution of both the TOF and TPC detectors depends on  $p_T$  and  $\eta$  (see Sects. 4.1 and 4.2). Finally, since we use two PID detectors simultaneously, the dimension of the parameter



**Figure 4.1.:** Examples of combined distributions of the time of flight  $\Delta t$ , and energy-loss  $dE/dx$  in the TPC, relative to the expected time of flight and energy loss for pions. The data is collected from Pb–Pb collisions, and the kinematical cuts on the tracks are:  $0.0 < \eta < 0.1$  (both panels),  $1.4 < p_T < 1.5$  GeV/ $c$  (left panel) and  $2.5 < p_T < 2.6$  GeV/ $c$  (right panel). The fit region is indicated with a black rectangle.

space becomes quite large. In our fit function, each peak in the (TOF, TPC) distribution is described by a function of six parameters, leading to parameter space of dimension six times the number of included peaks.

The aim in this chapter is to outline the procedure of obtaining identified particle yields in the interval  $1.0 < p_T < 4.0$  GeV/ $c$ , using both TOF and TPC information. The methods developed in this chapter provide the basis for the method described in Chapter 5, where we describe a procedure for identifying the particle yield associated with a high- $p_T$  trigger particle. We also use the fit procedure described in this chapter to obtain identified spectra which are corrected for detector inefficiencies. These spectra are compatible with published spectra from ALICE.

This chapter is set up as follows: in Sect. 4.1 and Sect. 4.2 we introduce the model for the TPC and TOF response. After that, in Sect. 4.3 we discuss the model used to fit the combined (TOF, TPC) response, as well as the procedure used to fit this model to the data. We then continue by discussing the systematic uncertainties of this procedure in Sect. 4.4, and in Sect. 4.5 we show examples of fit results. In Sect. 4.6 we discuss a method to determine

the fraction of secondary particles and in Sect. 4.7 we conclude with some final remarks.

## 4.1. TPC Detector Response

As mentioned in Sect. 3.2.3, the final  $dE/dx$  measurement for a track is a truncated mean of the individual measurements from the TPC clusters associated with the track. Provided that the tracks with a small number of associated TPC clusters are not used, these final  $dE/dx$  measurements are nearly Gaussian-distributed. In this work we only use tracks with at least 60 TPC clusters. This is sufficient to suppress the Landau-tail of the  $dE/dx$  distribution (see Sect. 3.2.3), but decreases the efficiency by 1 – 2%.

There are several factors that determine the resolution of the final  $dE/dx$  measurement. First of all, the resolution scales with the inverse square root of the number of TPC clusters that were used to calculate the truncated mean. This is another important reason for the requirement of a minimum number of TPC clusters. In principle the minimum requirement of 60 clusters still leaves a track sample with a range of TPC resolutions, however, in practice the vast majority of tracks have a very large number of TPC clusters<sup>1</sup>.

Furthermore, particles that traverse the TPC gas ionize the gas molecules, and the signal in each TPC cluster is proportional to the number of freed electrons that is detected. Therefore, the absolute uncertainty on the single cluster measurement grows as the square root of the number of observed electrons, but the relative uncertainty decreases. The number of electrons that are liberated by a particle is determined by the number of scattering centers it encounters on its path. The resolving power of the TPC can therefore be increased by increasing the pressure of the gas.

Note also that due to an increase in path length, a particle at large pseudorapidity liberates more electrons than a particle at mid-rapidity. Moreover, the electrons freed by a particle at large pseudorapidity have on average a shorter drift path, which suppresses potential loss of electrons due to diffusion or absorption by ions. These effects lead to a somewhat better resolution (and separation power) at large pseudorapidity.

In the rest of this work, we only use the actual signal in the TPC minus the expected signal for a certain particle species, which can be calculated using the Bethe-Bloch formula. It is therefore convenient to define the following

---

<sup>1</sup>Performing the analysis on several subsets of the data, each containing tracks with a similar number of TPC clusters, did not significantly improve the results.

variable:

$$X_i \equiv \frac{dE}{dx} - \left\langle \frac{dE}{dx} \right\rangle_i, \quad (4.1)$$

where the  $\langle \cdot \rangle_i$  denotes the expected value, given the reconstructed momentum and assuming particle species  $i \in \{\pi, K, p\}$ . Using this variable  $X_i$ , we can write the response of the TPC as follows:

$$\frac{dN_j}{dX_i} = \frac{N_{\text{tot},j}}{\sqrt{2\pi}\sigma_{j|i}} \exp \left\{ -\frac{(X_i - \mu_{j|i})^2}{2\sigma_{j|i}^2} \right\}, \quad (4.2)$$

where  $N_{\text{tot},j}$  is the total number of particles of species  $j$ , and  $\mu_{j|i}$  and  $\sigma_{j|i}$  stand for the position and width of the peak of species  $j$  in a plot with mass assumption  $i$ . Note that if the detector calibration is perfect, then  $\mu_{i|i} = 0$ , however, as we show Sect. 4.5, some systematic deviations are observed. For brevity, we omitted explicitly writing the  $(p_T, \eta)$  dependence of  $\mu$  and  $\sigma$ , however for the various reasons mentioned in this section, this dependence cannot be neglected.

To obtain the number of particles in an interval  $a < X_i < b$ , we simply integrate Eq. (4.2), i.e.:

$$N_j|_a^b = \frac{N_{\text{tot},j}}{2} \left[ \text{erf} \left\{ \frac{b - \mu_{j|i}}{\sqrt{2}\sigma_{j|i}} \right\} - \text{erf} \left\{ \frac{a - \mu_{j|i}}{\sqrt{2}\sigma_{j|i}} \right\} \right]. \quad (4.3)$$

## 4.2. TOF Detector Response

In Sect. 3.2.4 we already mentioned that part of the TOF resolution is due to hardware effects and due to uncertainty on the amount of energy lost by a particle before reaching the TOF detector. In this section we discuss these points further. We also introduce the model that is commonly used for the TOF detector response, namely, a Gaussian distribution, smoothly joined with an exponential tail towards higher TOF<sup>2</sup>.

Just as for the TPC, the signal strength measured by a MRPC depends on the amount of electrons that are freed when a particle traverses the detector. A particle at large pseudorapidity has a longer path through the hardware of the TOF detector, and therefore frees on average more electrons than a

<sup>2</sup>The cause of this deviation from a perfect Gaussian is unfortunately not well understood.

particle at mid-rapidity. The TOF detector records a “hit” whenever the measured signal strength reaches a certain threshold, and for stronger signals the exact time that the threshold is reached can be determined with higher accuracy<sup>3</sup>. This effect causes a small improvement in the TOF resolution at large pseudorapidity.

The uncertainty on the amount of energy lost by a particle before arriving at the TOF detector is related to the amount of energy loss itself. According to the Bethe-Bloch formula (Eq. (3.1)), the amount of energy loss depends on the particle’s velocity (or mass and momentum) and on the properties of the material that it traverses. A substantial part of the energy that the particle loses is due to traversing the TPC drift gas. Fig. 3.3 therefore implies that in the range  $1.0 < p_T < 4.0$  GeV/ $c$  the TOF resolution will be roughly similar for pions and kaons, but somewhat worse for protons towards the lower end of that range.

As in the case of the TPC, it is convenient to define the following variable for the TOF measurement:

$$T_i \equiv \Delta t - \langle \Delta t \rangle_i. \quad (4.4)$$

The TOF signal for particle species  $j$  under mass assumption  $i$  is modeled with a piecewise function consisting of a Gaussian, smoothly joined with an exponential tail in the direction of positive  $T_i$ , i.e.:

$$\frac{dN_j}{dT_i} = \frac{N_{\text{tot},j} A_{j|i}}{\sqrt{2\pi}\sigma_{j|i}} \begin{cases} \exp\left\{-\frac{(T_i - \mu_{j|i})^2}{2\sigma_{j|i}^2}\right\} & T_i < \kappa_{j|i} + \mu_{j|i} \\ \exp\{-\lambda_{j|i}(T_i - \mu_{j|i}) + B_{j|i}\} & T_i > \kappa_{j|i} + \mu_{j|i}, \end{cases} \quad (4.5)$$

where  $\mu_{j|i}$  and  $\sigma_{j|i}$  have similar meanings as before,  $\lambda_{j|i}$  is the slope of the exponential tail,  $\kappa_{j|i}$  is the place relative to  $\mu_{j|i}$  where the Gaussian is joined to the exponential. The parameter  $A_{j|i}$  corrects for the exponential having a larger integral than the Gaussian, consequently we have that  $A_{j|i} \leq 1$ .

Note that the exponent introduces three extra parameters on top of the parameters from the Gaussian, however we can solve for two of the parameters in terms of the others by using the continuity and smoothness condition. The continuity conditions can be written as:

$$\kappa = \lambda\sigma^2, \quad B = \lambda^2\sigma^2/2. \quad (4.6)$$

---

<sup>3</sup>The accuracy is related to the first time derivative of the signal as it reaches the threshold.

The two possibilities we have investigated were fixing  $\lambda$  and  $B$  (indices suppressed), or  $\kappa$  and  $B$ . We noted that the latter resulted in somewhat faster convergence in the minimization procedure, so from now on we only use this choice.

The number of particles in an interval  $a < T_i < b$  integral of Eq. (4.5) is given by:

$$N_j|_a^b = N_{\text{tot},j} A \begin{cases} \frac{1}{2} \left[ \text{erf} \left\{ \frac{b-\mu}{\sqrt{2}\sigma} \right\} - \text{erf} \left\{ \frac{a-\mu}{\sqrt{2}\sigma} \right\} \right] & a, b < \mu + \kappa \\ \frac{1}{2} \left[ \text{erf} \left\{ \frac{\kappa}{\sqrt{2}\sigma} \right\} - \text{erf} \left\{ \frac{a-\mu}{\sqrt{\sigma}} \right\} \right] \\ \quad + \frac{e^B}{\sqrt{2\pi\lambda\sigma}} (e^{-\lambda\kappa} - e^{-\lambda(b-\mu)}) & a < \mu + \kappa, b > \mu + \kappa \\ \frac{e^B}{\sqrt{2\pi\lambda\sigma}} (e^{-\lambda(a-\mu)} - e^{-\lambda(b-\mu)}) & a, b > \mu + \kappa. \end{cases} \quad (4.7)$$

From this equation we find that the total normalization of the probability distribution function (PDF) is given by:

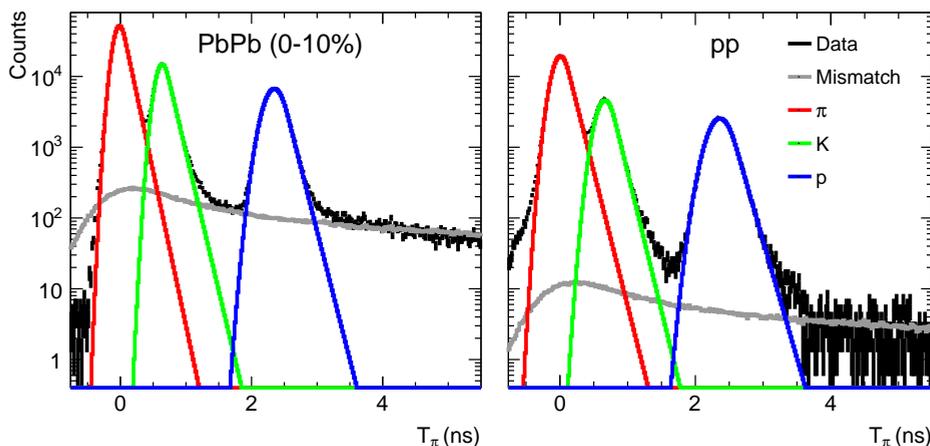
$$A = \frac{2}{\text{erf} \left\{ \frac{\kappa}{\sqrt{2}\sigma} \right\} + \frac{2}{\sqrt{2\pi\lambda\sigma}} e^{-B} + 1}. \quad (4.8)$$

Note that as expected

$$\lim_{\kappa \rightarrow \infty} A = 1.$$

The PDF in Eq. (4.5) only describes the tracks that were matched with the correct TOF hit. The two other possibilities are that a track is not matched with a TOF hit at all, or that a track was matched with an incorrect TOF hit, either with a hit produced by another particle or with a fake hit due to detector noise. The latter is commonly referred to as a *mismatch*. The probability for a mismatch typically gets larger in a high-track density environment, such as a central Pb–Pb collision. In our analysis, tracks without a TOF hit are not selected, however, tracks with a mismatched TOF signal cause a background that needs to be disentangled from the signal.

In Fig. 4.2, we show a typical example of a TOF signal distribution. The logarithmic scale on the vertical axis is used to emphasize the presence of the mismatched background. Note that the fraction of mismatched tracks is much larger in central Pb–Pb collisions than in pp collisions. To correctly subtract mismatches in the fits, we need a suitable model for it. Unfortunately, it is not trivial to approximate it with a simple analytic function, and template



**Figure 4.2.:** Examples of the TOF response in Pb–Pb and pp. The kinematical cuts on the tracks are:  $0.0 < \eta < 0.1$  and  $1.4 < p_T < 1.5$  GeV/ $c$  (both panels). The gray line is a template which roughly matches the expected shape of the TOF mismatch band.

histograms, generated using Monte Carlo events and a model of the ALICE detector in GEANT3 [93], do not describe its shape correctly.

An alternative is to use a data-driven method to generate a template histogram. We now briefly describe this method, which, at the moment, is the standard way of describing the TOF mismatches in physics analyses from the ALICE collaboration [94]. In principle, only those TOF hits that are matched to a reconstructed track are stored for further analysis. During data taking however, a time distribution is created, using information from all TOF hits, including the ones that are not matched<sup>4</sup>. To make this a sensible distribution, all TOF hits are corrected for the distance between the interaction point and the precise location of the TOF cluster involved<sup>5</sup>. One can generate a mismatch template by considering all reconstructed tracks, and assigning an arbitrary TOF measurement to them, which is randomly taken from the time distribution of all TOF hits. This procedure turns out to describe the mismatched background reasonably well, especially at larger momenta.

<sup>4</sup>This distribution was originally intended to be used for an interaction time estimate based on the TOF hits.

<sup>5</sup>Unfortunately all particles are assumed to travel on straight paths, which is not a very good approximation at low momenta.

### 4.3. TOF-TPC Combined Fits

To describe the peaks in the  $(T_i, X_i)$ -distributions, we essentially take the product of the TOF and TPC response, which were described in Sects. 4.1 and 4.2. A template histogram describing the mismatches can be generated in the same manner as described in Sect. 4.2, however, now the distribution is filled with a random TOF time, but the true specific energy loss measurement. The total fit function for the  $(T_i, X_i)$ -distribution can be written as:

$$\frac{d^2 N_{\text{tot}}}{dT_i dX_i} = \sum_j \frac{d^2 N_j}{dT_i dX_i} + M_i, \quad (4.9)$$

where,  $j$  sums over all the particle species that are included in the fit, and where  $M_i$  stands for the mismatch template generated assuming the peak of species  $i$  is in the origin, and:

$$\frac{d^2 N_j}{dT_i dX_i} \propto \frac{dN_j}{dT_i} \times \frac{dN_j}{dX_i}. \quad (4.10)$$

To determine the total yield for particle species  $i$ , we only use the histogram with mass assumption  $i$ , i.e., with  $T_i$  and  $X_i$  on the axes. The reason for this is that the distance between the peaks depends on momentum, which causes all peaks except for the one corresponding to species  $i$  to be “smeared out” somewhat, i.e., their resemblance to our fit model diminishes. The impact of this smearing on our fit results can be minimized by taking narrow bins in  $p_T$  and  $\eta$ , and by excluding the proton peak in the pion and kaon fits, and vice versa in the range  $p_T < 1.2 \text{ GeV}/c$ . In the remainder of this section we outline the procedure that we used to find the optimal parameters to describe the data, with the emphasis on obtaining the most reliable values for the pion, kaon, and proton yields.

The first step in the procedure is to find the number of mismatches in the data. This is achieved by normalizing the template mismatch histogram to its number of entries, and then to determine the scale factor that is needed to make the mismatch template fit the data. Note that in this way also data that does not fall within the domain of the  $(T_i, X_i)$  histogram is taken into account. In principle one should use this scale factor as one of the parameters in a fitting procedure, however this does not produce good results. Most notably, since neither the description of the peaks nor the description of the mismatched background are perfect, the scale factor is typically estimated

to be much higher than that it should be. For this reason we estimated the scale factor by comparing the mismatch template with the data in a region far away from the peaks. A simple check that is described in Sect. 4.4.1 shows that this way of determining the mismatched yield only gives rise to a very small bias on the fitted particle yields  $N_{\text{tot},i}$ .

After the mismatch template is fixed, what remains is the estimation of the parameters of the peaks. This is done by maximizing the logarithm of the likelihood  $L$  of the model, assuming the data in the bins of the  $(T_i, X_i)$  histogram are Poisson distributed. In practice, we minimize the following function:

$$\log L = -\nu_{\text{tot}} + \sum_{i \in R} n_i \log \nu_i, \quad (4.11)$$

where  $R$  is the set of bins that fall within the fit region,  $\nu_{\text{tot}}$  is the total fitted yield in region  $R$ ,  $n_i$  is the measured yield in bin  $i$  and  $\nu_i$  is the yield the model predicts in bin  $i$ . For all PID fits in this work we use the MINUIT minimizer.

The domain of the  $(T_i, X_i)$  histograms used in our analysis is generally large compared to the typical distance between the peaks that we are interested in  $(\pi, K, p)$ . The main reason for this is to get a substantial number of bins that are far away from the peaks, which improves the estimate of the scale factor of the mismatch template. For the sake of computational efficiency however, one should aim at keeping the region  $R$  to be as small as possible, i.e., much smaller than the full domain of the histogram, but sufficiently large to fully include the peaks that are included in the fit. The default choice of the region  $R$  is a rectangle shape (See Sect. 4.4.2). Performing the fits with a differently shaped region, leads to small deviations in the fitted particle yields compared to the default shape. This is most likely caused by imperfections in our model.

Finally, we found that the parameter  $\lambda_{K,i}$ , described in Eq. (4.5) cannot be determined in a reliable way at high  $p_T$  by means of the fitting procedure. The distribution in Fig. 4.1 reveals the reason for this, namely that at high  $p_T$  the proton peak largely overlaps the tail of the kaon peak. This means that the optimal value of  $\lambda_{K,i}$  is very sensitive to imperfections in the fit of the proton peak. Note that the same is not true for the parameters  $\lambda_{\pi,i}$  and  $\lambda_{p,i}$ . In our procedure we therefore fix the value of  $\lambda_{K,i}$ . The systematic uncertainty caused by this approximation is discussed in Sect. 4.4.4.

Parameter	Value
$\mu_{\text{TOF}}, \mu_{\text{TPC}}$	Taken from the detector calibration
$\sigma_{\text{TOF}}$	85 ps (p-Pb, Pb-Pb) and 115 ps (pp)
$\sigma_{\text{TPC}}$	3.5 a.u.
$\lambda_{\text{TOF}}$	0.008 ps <sup>-1</sup>
<i>Yield</i>	Estimated from the bins around the expected peak position

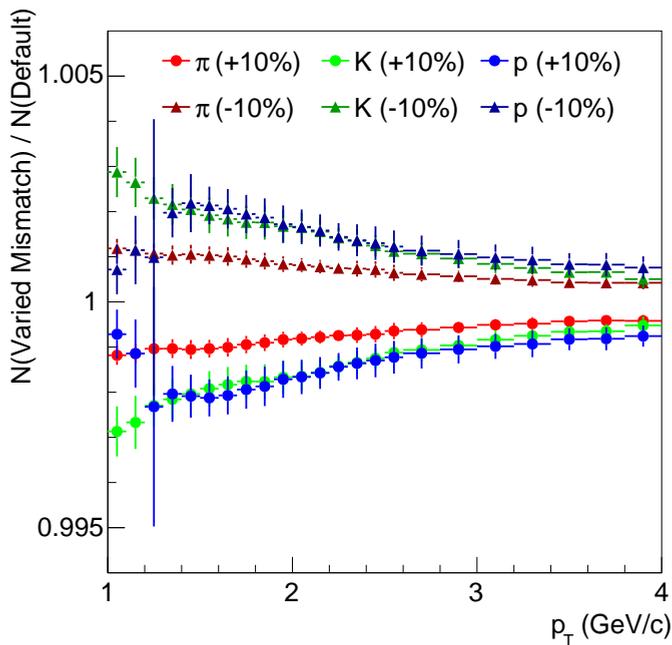
**Table 4.1.:** The initial values that were used in every fit, or the method used to obtain them.

Table 4.1 summarizes the initial values used in every fit. For the peak position ( $\mu_{\text{TOF}}$  and  $\mu_{\text{TPC}}$ ) the value from the detector calibration is sufficiently close to the optimal value to get a stable fit. We estimate the initial values for the yield by comparing a 2D Gaussian distribution with the initial  $\sigma_{\text{TOF}}$  and  $\sigma_{\text{TPC}}$ , to the values in the bins around the initial peak position. This estimate could significantly deviate from the final fitted value, especially if the detector calibration was not optimal. We find, however, that this again did not lead to instabilities in the fit. To guide the minimizer to the correct minimum, we perform all fits in two steps. First, the fit is performed with  $\lambda_{\text{TOF}}$  fixed to its initial value, for all peaks. After this, all parameters are released, except for  $\lambda_{K,i}$ .

The particle identification method described in this section has not been used before by the ALICE collaboration. For this reason, we performed a number of crosschecks, including a comparison of our procedure with bin counting and with stand-alone TOF fits. The most notable crosschecks are discussed in Sect. 4.4.

## 4.4. Systematic Uncertainties

Apart from a few assumptions, the fit function described in the previous sections is fairly generic. In this section we elaborate on these assumptions, and give an estimate of the uncertainty that they cause on the measured pion, kaon, and proton yields.



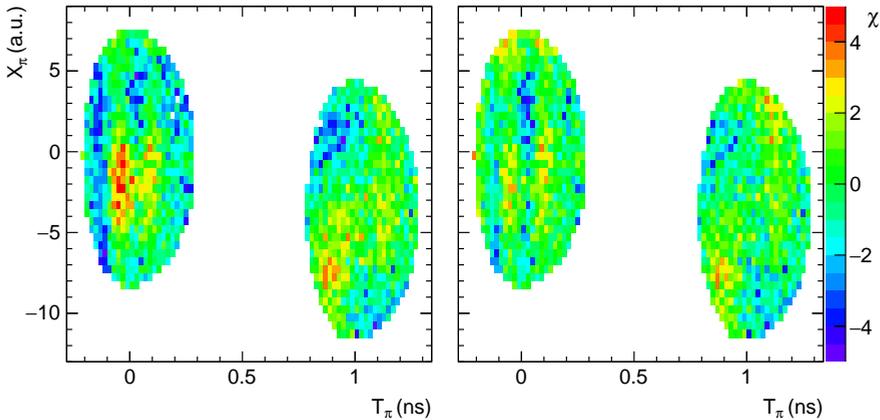
**Figure 4.3.:** Ratio of identified yield after varying the fixed mismatch yield by 10% in Pb–Pb collisions.

#### 4.4.1. Mismatched Background

As mentioned in Sect. 4.3, the mismatch template is fixed in a region far away from the peaks. To estimate the resulting uncertainty, we refitted the spectra, but increased/decreased the estimated mismatch yield by 10%. From Fig. 4.3 it is clear that the effect of varying mismatches is negligible for Pb–Pb collisions. For p–Pb and pp this also holds, as TOF mismatches are significantly less abundant in these systems.

#### 4.4.2. Choice of Fit Region

The fit region that is used in this analysis is a rectangle, containing all peaks up to several times the detector resolution. In Fig. 4.1, these rectangular fit regions are shown as black boxes around the peaks. If our fit model (Eq. (4.9)) would be a perfect description of the data, then choosing the fit region differently would yield a statistically compatible optimized value



**Figure 4.4.:** Examples of  $\chi$ -distributions in  $(T_\pi, X_\pi)$ . Mass assumption: pion,  $1.0 < p_T < 1.1$  GeV/c,  $-0.8 < \eta < -0.7$ . Left panel: standard rectangular fit region, right panel: refit using the smaller fit region #1, defined in Eqs. (4.12) and (4.13), and Table 4.2.

for all fit parameters, provided that the different fit region includes all the important features of the distribution.

To estimate the goodness of the fit, we evaluated the  $\chi$  value<sup>6</sup> in all bins in the region around the center of the fitted peak. An example of this is shown in the first panel of Fig. 4.4. Taking a closer look at the pion peak, located approximately in the origin, we see a clear structure in the  $\chi$ -distribution. The data in the center of the peak is generally underestimated by our model, while towards the edge it is sometimes overestimated. This implies that the tails of the distribution, which are included in the fit region, are not well described by our model. The other peak in the figure, which is the kaon peak, is also not described perfectly, however this is less important for our analysis, and can at least partially be attributed to residual smearing due to the finite size of the  $(p_T, \eta)$  bin.

To estimate the influence of the choice of fit region on our results, we added a third fitting step to the procedure described in Sect. 4.3, where the model is refitted in a smaller fit region. This smaller fit region is composed of two half ellipses, including all the bins which have centers obeying the

<sup>6</sup>For an observed yield  $N_{\text{obs}}$  and fitted model estimate  $N_{\text{mod}}$  we have that:  

$$\chi = (N_{\text{obs}} - N_{\text{mod}}) / \sqrt{N_{\text{obs}}}.$$

	Region No.			
	1	2	3	4
$r_{\text{TOF,pos}}$ (ps)	300	400	300	400
$r_{\text{TOF,neg}}$ (ps)	200	300	200	300
$r_{\text{TPC}}$ (a.u.)	8	8	12	12

**Table 4.2.:** Definition of the smaller fit region parameters.

following inequality:

$$\sqrt{\left(\frac{T - \mu_{\text{TOF}}}{r_{\text{TOF,neg}}}\right)^2 + \left(\frac{X - \mu_{\text{TPC}}}{r_{\text{TPC}}}\right)^2} < 1, \quad (4.12)$$

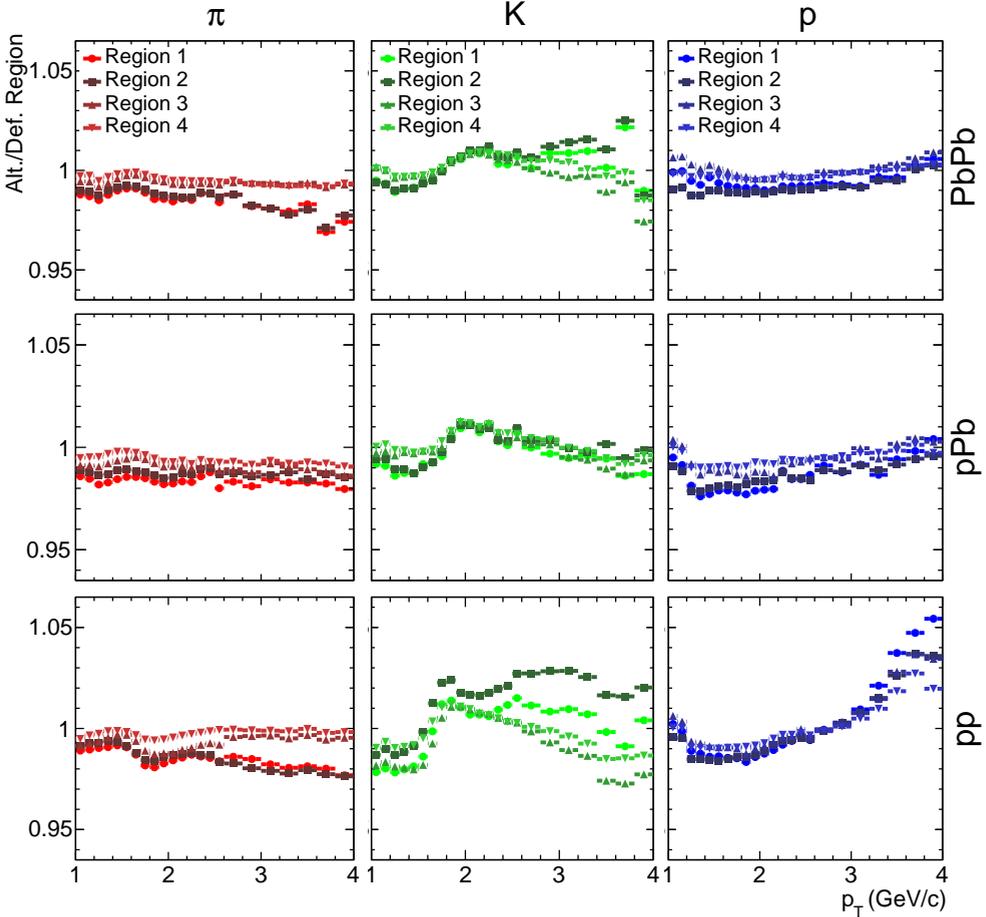
if  $T < \mu_{\text{TOF}}$ , and:

$$\sqrt{\left(\frac{T - \mu_{\text{TOF}}}{r_{\text{TOF,pos}}}\right)^2 + \left(\frac{X - \mu_{\text{TPC}}}{r_{\text{TPC}}}\right)^2} < 1, \quad (4.13)$$

if  $T > \mu_{\text{TOF}}$ , where the radii  $r_{\text{TOF,neg}}$ ,  $r_{\text{TOF,pos}}$  and  $r_{\text{TPC}}$  are given in Table 4.2. As indicated in this table, the four different regions are significantly smaller than the default rectangular region, but sufficiently large to contain at least 90% of the estimated yield. The radii of the smallest region, which is the first configuration, correspond to roughly twice the detector resolution. Notice that in the TOF direction, the radius in positive direction is taken to be larger than the one in negative direction. This ensures that the fit region contains sufficient data for the fit of the parameter  $\lambda$ .

The right panel of Fig. 4.4 shows the  $\chi$ -distribution after refitting using fit region #1. The agreement between the model and the data around the origin has improved compared to the left panel. In Fig. 4.5 we show the result of this study for all four alternative regions, compared to the original region for pp, p-Pb, and Pb-Pb. The deviation in the yield for the largest region (region #4) compared to the regular fit region is typically small ( $\sim 1 - 2\%$ ), with a small dependence on particle species and collision system. From there, decreasing the fit region in the TOF direction does not seem to alter the results more than  $\sim 1\%$  (region #3). Decreasing the size of the fit region in the TPC direction has a significant impact on the fitted yield in certain

cases. This is especially true at high  $p_T$  and most notably for pions in pp and Pb–Pb, as well as for kaons and protons in pp, with discrepancies in the fitted yield up to  $\sim 5\%$ .



**Figure 4.5.:** Ratio of the identified yield using an alternative fit region and the default region in pp, p-Pb, and Pb–Pb. The alternative fit regions are defined in Eqs. (4.12) and (4.13), and Table 4.2.

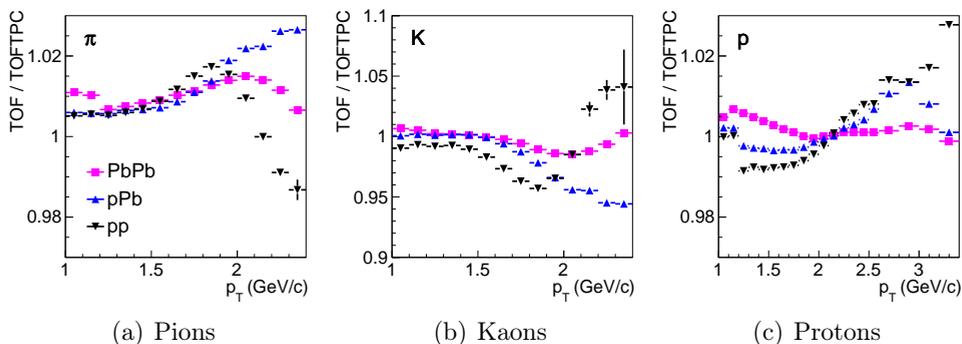
Closer inspection of the fit results at high  $p_T$  for pions and kaons does not show any obvious features in the  $(T_i, X_i)$  distributions that are not described by the model, hence, for Pb–Pb, we choose to assign a systematic uncertainty to the pion and kaon yield of 1 – 2%, and to the proton yield of 1%. For p-Pb the same uncertainties are used as for Pb–Pb, except for protons for

which the uncertainty is  $\sim 2\%$ . For pp the systematic uncertainty on the pions is also 1 – 2%, while for kaons and protons it is 2 – 3% and 2 – 5% respectively.

### 4.4.3. Comparison with Other Fit Methods

In the lowest  $p_T$  bins ( $1.0 < p_T < 1.4$  GeV/ $c$ ) the proton peak is far removed from the pion and kaon peaks in the (TOF, TPC) plane, and a very good estimate of the proton yield can be obtained simply by bin counting. These estimates generally compare very well ( $\lesssim 1\%$ ) with the 2D fit method, and no significant dependence on  $\eta$  is observed.

While the bin-counting comparison provides evidence for the validity of the fit function, it is limited in its scope. We therefore also compared our 2D fit method with a TOF standalone (SA) fit procedure. For each  $p_T$  and  $\eta$  bin, after the regular fit procedure described in Sect. 4.3 is completed, a TOF SA fit was performed using the fit parameters from the 2D method as initial values. In Fig. 4.6 we show how with TOF SA fit and the combined fit procedure compare for pp, p-Pb, and Pb-Pb. The  $p_T$  range of this study is limited by the decreasing resolving power of the TOF detector towards higher  $p_T$ . In general, we found that without making additional assumptions on the fit function, the pion and kaon TOF SA fits converge up to  $p_T \sim 2.4$  GeV/ $c$ , while proton fits converge up to  $p_T \sim 3.4$  GeV/ $c$ .



**Figure 4.6.:** Comparison between the identified yield using a TOF stand-alone fit method and the (TOF, TPC) fit method for pp, p-Pb, and Pb-Pb.

The first thing that stands out from Fig. 4.6 is that the observed difference for pions and kaons are anti-correlated. A possible explanation of this effect

is that when the kaon peak moves closer to the pion peak at higher  $p_T$ , the value of  $\lambda_\pi$  becomes more difficult to estimate in the TOF SA fit, and may be somewhat overestimated, effectively absorbing some kaons into the pion peak. As the pion production is larger than the kaon production, the relative effect on the kaons is larger.

Since, upon closer inspection, both the TOF SA and combined fit procedures give reasonable fit results and since it is difficult to precisely pinpoint the origin of the discrepancy, it is not possible to decide which of the two methods is superior. While the combined fit procedure clearly has more resolving power at high  $p_T$ , imperfections in the description of the TPC response may in fact make it somewhat less accurate in the  $p_T$  region where both methods are competitive. For this reason we take the difference between the two methods to as systematic uncertainty<sup>7</sup>.

Since both the systematic uncertainty that is estimated from comparing different fit regions (Sect. 4.4.2) and the systematic uncertainty that is estimated from comparing to TOF SA fits are both a measure of how well the fit function captures the data, it is reasonable to not simply add these two uncertainties in quadrature, but instead to only use the largest of the two. The only case where the TOF SA comparison yields larger uncertainties is for kaons in pp and p-Pb where the uncertainty becomes 2 – 5%, and the maximum value of 5% is attained around  $p_T \sim 2$  GeV/ $c$ . For protons the uncertainties remain unchanged.

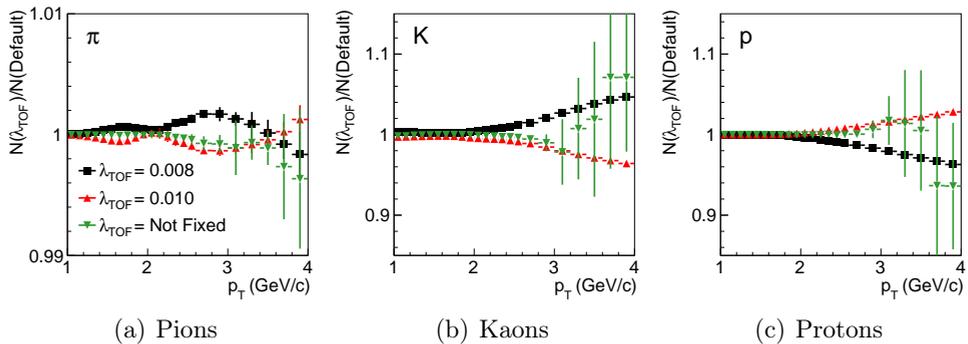
#### 4.4.4. Fixing the Tail Parameter

In Sect. 4.3 we mentioned that the parameters  $\lambda_{K,i}$  are very sensitive in the kinematic region where the proton peak overlaps with the exponential kaon tail. For the sake of fit stability we decide to fix this parameter. In this section we estimate the influence of this choice on the fitted yield, by comparing different reasonable values of  $\lambda_{K,i}$ .

In the kinematic range where the kaon and proton peaks are well-separated, the fits where  $\lambda_{K,i}$  is allowed to vary yield values in the range  $0.008 < \lambda_{K,i} < 0.010$ , for all  $(p_T, \eta)$  bins. Fig. 4.7 shows a comparison between the fitted pion, kaon, and proton yields under the assumption  $\lambda_{K,i} = 0.008$ ,  $\lambda_{K,i} = 0.010$ , and  $\lambda_{K,i}$  as a free parameter, to the case where  $\lambda_{K,i} = 0.009$ . From this figure it is clear that for  $p_T < 3.0$  GeV/ $c$ , the true value of  $\lambda_{K,i}$  is fairly constant. It is most likely that the observed deviations between the

---

<sup>7</sup>Here we assume the uncertainty at  $p_T$  beyond the region where the TOF SA fit converges to be the same as the uncertainty at the upper edge of that region.



**Figure 4.7.:** Ratio of raw identified particle yield with  $\lambda_{\text{TOF}}$  free or fixed, and the default choice ( $\lambda_{\text{TOF}} = 0.009$ ). This ratio is calculated as a function of  $p_T$  and  $\eta$ , and projected onto the  $p_T$  axis. The error bars reflect the RMS of the ratio of the different  $\eta$  bins at the same  $p_T$ . Collision system: Pb–Pb.

fits where  $\lambda_{K,i}$  is and is not fixed are due to the overlapping of the kaon and proton peak. For this reason we decide to fix  $\lambda_{K,i} = 0.009$  for the entire  $p_T$  range, and assign a 0 – 3% uncertainty on the kaon yield, and a 0 – 2% uncertainty on the proton yield, depending on  $p_T$ .

#### 4.4.5. Contamination from Leptons

From the detector response in Figs. 3.3 and 3.4 it is clear that while the TOF has very limited separation power for pions and electrons, the TPC separates them very well. In our 2D fit procedure all electron contamination is cut out by the definition of our fit region.

Muon contamination is more difficult to estimate, since neither TOF nor TPC has good separation power for muons and pions. It is certainly not feasible to include a muon peak to our fit model with no further assumptions on its parameters. A pragmatic approach is to add a highly restricted muon peak to the fit function, where the “shape” parameters ( $\sigma$ ,  $\lambda$ ) of the muon peak are fixed to those from the pion peak, obtained in a previous fit. The position of the muon peak is also fixed to the position of the pion peak, but appropriately shifted, to compensate for the small mass difference between the pion and the muon.

Unfortunately even this minimal approach does not produce reliable fit results, confirming the very limited separation power. On the other hand,

the  $\mu/\pi$  fraction is small, and since muons typically arise from weak decays, a large fraction is effectively cut out by the strict DCA requirement on the tracks (see Sect. 4.6 and Appendix A).

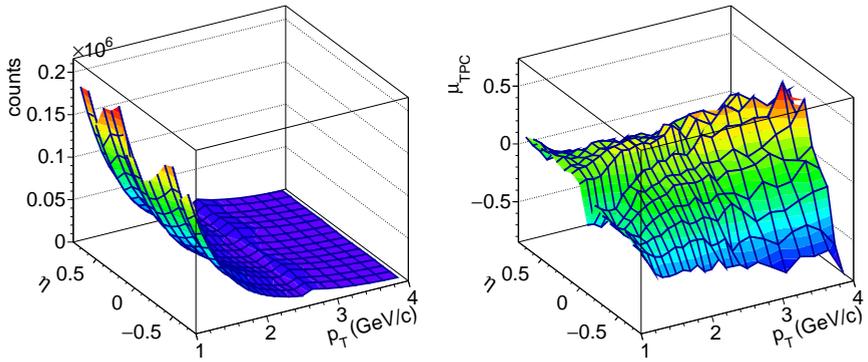
Following other physics analyses in ALICE, the muon contamination in the  $p_T$  region of interest is estimated to be less than 1%, and only present in the lower end of our  $p_T$  range. We do not apply a correction for the lepton contribution, as the effect is negligible compared to other effects.

## 4.5. Fit Results

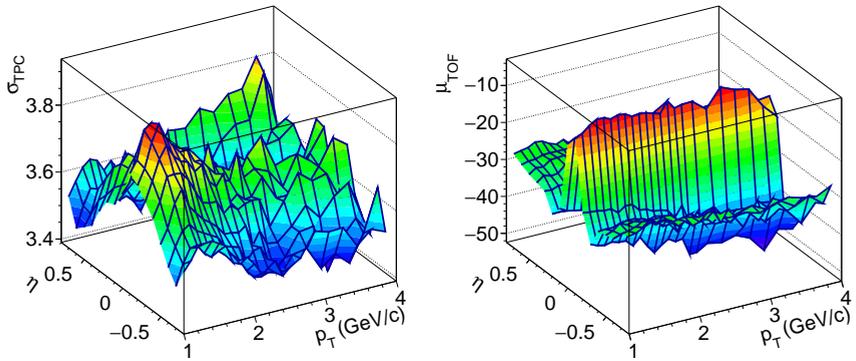
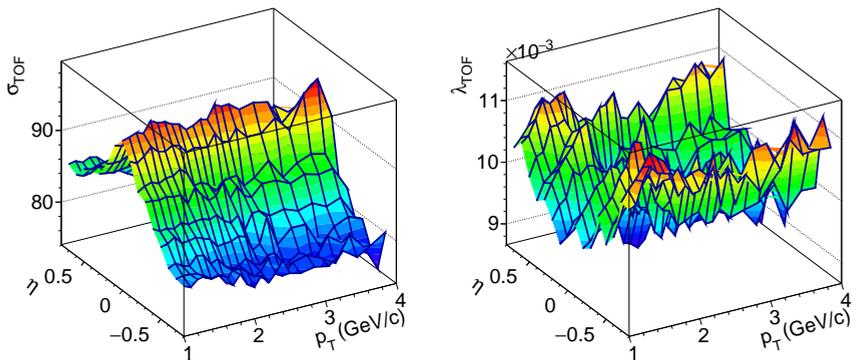
In Fig. 4.8 we present some results from the fits performed on the p-Pb dataset, in the kinematic range  $1.0 < p_T < 4.0$  GeV/ $c$ ,  $-0.8 < \eta < 0.8$ . As the fit results for the other collision systems are comparable and also the fit results from the different  $(T_i, X_i)$ -planes are also comparable, we only show the results from the fit in the  $(T_\pi, X_\pi)$ -plane. As mentioned in the beginning of this chapter, each fit has up to seventeen parameters, however only the peak that is centered is expected to be unaffected by the “smearing” due to the dependence of the expected distance between the peaks on the kinematic variables. Moreover the off-centered peaks are not used for determining a particle yield, and are only included to take into account their influence on the centered peak. For this reason Fig. 4.8 only shows the parameters belonging to the pion peak.

In order to be sensitive to the dependence of the fit parameters on  $p_T$  and  $\eta$ , the data is divided in narrow bins in both coordinates. On the other hand, as each bin corresponds to an independent fit, it is also important to avoid choosing them too small, so that all bins contain sufficient data for a stable fit. In  $\eta$  we use a constant bin size of 0.1, while the  $p_T$  dimension is divided in 22 bins with increasing bin size, from 0.1 GeV/ $c$  at low  $p_T$  to 0.3 GeV/ $c$  at high  $p_T$ .

The fit results in Fig. 4.8 possess several notable features. First of all, both  $\mu_{\text{TOF}}$  and  $\mu_{\text{TPC}}$  show a statistically significant  $\eta$ -dependence. The fact that these features persist in  $p_T$  shows that it is unlikely that they are artifacts from the fit procedure, instead they probably point to a small mistake in the detector calibration. It is instructive to compare these discrepancies to the typical detector resolution, i.e.,  $\mu_{\text{TPC}} \sim 0.1 \sigma_{\text{TPC}}$ ,  $\mu_{\text{TOF}} \sim 0.3 \sigma_{\text{TOF}}$ . While our analysis should not be affected by these discrepancies, these numbers show that one should be cautious when performing an identification procedure based on a cut on the PID signals.



(a) Pion Raw Counts

(b) Pion  $\mu_{\text{TPC}}$ (c) Pion  $\sigma_{\text{TPC}}$ (d) Pion  $\mu_{\text{TOF}}$ (e) Pion  $\sigma_{\text{TOF}}$ (f) Pion  $\lambda_{\text{TOF}}$ 

**Figure 4.8.:** Fit parameters for the pion peak, from the fit of the p-Pb data in the  $(T_\pi, X_\pi)$ -plane.

We also notice a significant  $\eta$  and  $p_T$  dependence, both in  $\sigma_{\text{TOF}}$  and  $\sigma_{\text{TPC}}$ . These dependencies are however not unexpected. For the various reasons discussed in Sects. 4.1 and 4.2 both the TPC and TOF are expected to have a better resolution at larger momenta and pseudorapidity.

## 4.6. Correction for Secondary Particles

Particles that are produced directly at the collision are called *primary particles*. Other particles, originating from weak decays and from the interaction of particles with the detector material, are called *secondary particles*. The most important weak-decay processes producing charged pions and (anti-)protons are:  $K_S^0 \rightarrow \pi^+ + \pi^-$ ,  $\Lambda \rightarrow p + \pi^-$  and  $\bar{\Lambda} \rightarrow \bar{p} + \pi^+$ . Charged kaons are rarely produced through weak decays, and can be ignored. Secondary particles from material interaction are mostly produced at very low  $p_T$ . In the range  $p_T > 1.0$  GeV/c, the only particle species with a non-negligible ( $\lesssim 3\%$ ) contribution from material interactions are the protons<sup>8</sup>.

As only the primary particles are sensitive to the mechanism of hadronization, we need a method to separate them from the secondary particles. Primary particles distinguish themselves from the secondary particles by the small distance of closest approach (DCA) of their (extrapolated) reconstructed tracks to the primary interaction vertex. The method that we describe here has also been used by various other analyses performed by members of the ALICE collaboration.

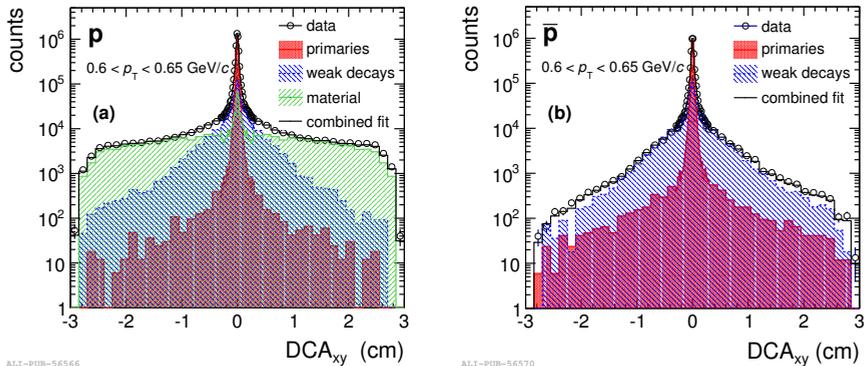
The first step in the procedure is to use a MC simulation to generate probability distributions of the radial component of the DCA vector ( $r_{\text{DCA}}$ )<sup>9</sup> for primary particles, weak-decay products and remnants of material interaction. These distributions are generated for six types of particles:  $\pi^+$ ,  $\pi^-$ ,  $K^+$ ,  $K^-$ ,  $p$  and  $\bar{p}$ .

The second step is to select a pure sample of those six particle types from the experimental data, in our implementation we do this by selecting particles from a  $1\sigma$ -region around the origin in the  $(T_i, X_i)$  plane, where  $i$  corresponds to the particle species, and  $\sigma$  is the detector resolution taken from the detector calibration<sup>10</sup>. The fraction of secondary particles decreases

<sup>8</sup>Not anti-protons.

<sup>9</sup>Defined as the vector pointing from the primary interaction vertex to the (extrapolated) reconstructed track at its point of closest approach.

<sup>10</sup>In light of the results presented in Sect. 4.5, we could have made this more accurate by using the fitted values of  $\mu$  and  $\sigma$ . On the other hand, it is the purity of the sample much more than its size that is important for determining particle fractions.



**Figure 4.9.:** Example of template fits to the radial component of the DCA distribution for protons (a) and anti-protons (b) in 0 – 5% central Pb–Pb collisions. The total fit is shown as a black line, while the colored lines represent the individual templates for primary particles, weak decay products and particles from material interaction. Figure taken from Ref. [94].

rapidly as a function of  $p_T$ , and the  $1\sigma$  requirement gives a sufficiently pure sample in the  $p_T$  region of interest. From this pure sample one can extract the  $r_{\text{DCA}}$  distributions, as a function of  $p_T$ .

Finally, the templates that are generated in the first step are fitted to the  $r_{\text{DCA}}$  distributions taken from data, using the model:

$$\frac{1}{N_{\text{tot},i}} \frac{dN_i}{dr_{\text{DCA}}} = f_P D_P(r_{\text{DCA}}) + f_W D_W(r_{\text{DCA}}) + f_M D_M(r_{\text{DCA}}), \quad (4.14)$$

where  $i$  is one of the six particle types, the subscripts  $P$ ,  $W$  and  $M$  stand for “primary”, “weak” and “material” respectively,  $D$  represents the normalized templates from the first step, and the fractions  $f_{P,W,M}$  are the fit parameters.

In Fig. 4.9 we show an example of a template DCA fit for (anti-)protons. From this figure it is clear that each template has a very distinct shape, which constitutes the working principle of a template fit. It also clearly shows that material interactions are responsible for some of the proton production, but that no anti-protons are produced in this way.

## 4.7. Summary

In this chapter we introduced a fitting procedure which uses both the TOF and TPC signal, with the purpose of measuring the pion, kaon, and proton yield. This procedure is optimized to give stable results for pp, p-Pb, and Pb-Pb data, in the range  $1.0 < p_T < 4.0$  GeV/ $c$ . The fit parameters extracted from these fits generally depend on both  $p_T$  and  $\eta$ . While the dependence of the detector resolution is expected, the dependence of the peak position likely point to a small error in the calibration. A clear advantage of our method over methods based on a cut on the detector signal is that it is not sensitive to these residual calibration discrepancies, and that the results are not influenced by contamination from other particles.



## Chapter 5.

# Identified Associated Yield

The final aim of our work is to study the hadrochemistry of the yield associated with a high- $p_T$  trigger particle. In the first half of this chapter we introduce the general methods used to obtain the associated yield and to correct it for the finite two-particle efficiency and acceptance<sup>1</sup>. Next, we outline the method we used to separate the associated pions, kaons, and protons. The final analysis of the identified associated distributions are discussed in Chapter 6.

## 5.1. Associated Particle Distributions

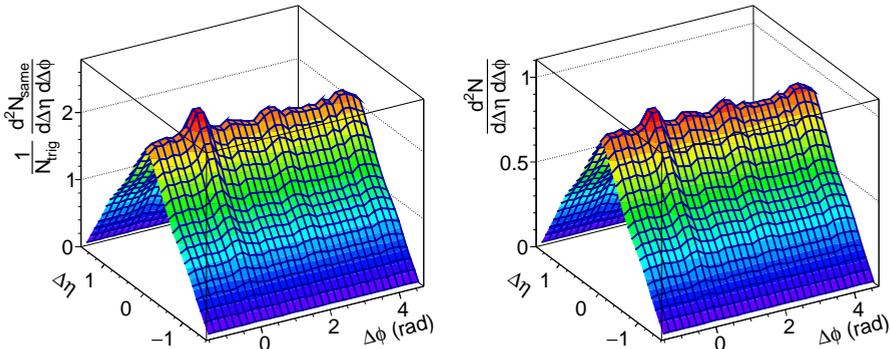
In a two-particle correlation analysis one studies the per-trigger yield associated with some trigger particle, i.e., one measures:

$$\frac{1}{N_{\text{trig}}} \frac{d^2 N_{\text{assoc}}}{d\Delta\phi d\Delta\eta}, \quad (5.1)$$

where  $\Delta\phi \equiv \phi_{\text{trig}} - \phi_{\text{assoc}}$ ,  $\Delta\eta \equiv \eta_{\text{trig}} - \eta_{\text{assoc}}$ , and  $N_{\text{trig}}$  is the number of trigger particles in the analyzed sample. Notice that by normalizing this quantity with the number of trigger particles, the observable becomes independent of the sample size. In this analysis we choose the trigger particle to be of high- $p_T$ . Since high- $p_T$  particles are often part of a highly energetic jet, studying their associated yield is similar to studying the contents of jets.

---

<sup>1</sup>Acceptance refers to the physical extent of the sensitive region of the detectors that we use, while the efficiency denotes the probability of detecting a particle within that sensitive region.



**Figure 5.1.:** Example of a two-particle correlation in central (0-10%) Pb–Pb collisions. Kinematic ranges:  $4.0 < p_{T,\text{trig}} < 10.0$  GeV/ $c$ , and  $1.2 < p_{T,\text{assoc}} < 1.3$  GeV/ $c$ . Left panel: raw associated yield distribution, right panel: mixed event distribution (scaled to be unity at the origin).

On the first panel of Fig. 5.1 we show an example of the particle yield in the range  $1.2 < p_{T,\text{assoc}} < 1.3$  GeV/ $c$ , associated with a high- $p_T$  trigger particle ( $4.0 < p_{T,\text{trig}} < 10.0$  GeV/ $c$ ), in central (0-10%) Pb–Pb collisions. First of all, notice the triangular shape of this distribution in the  $\Delta\eta$ -direction. This shape is not physical, but reflects the limited two-particle acceptance, caused by the requirement that all tracks fall within the range  $|\eta| < \eta_{\text{max}} = 0.8$ . Notice also that there are some “bumps” and “dips” in the  $\Delta\phi$ -direction. These occur when there are at least two places of increased (decreased) detector efficiency along the azimuthal angle  $\phi$ , changing the efficiency of measuring a pair of particles at an angular difference equal to the angular difference between the places of increased (decreased) detector efficiency<sup>2</sup>.

In order to study the physical features of the associated yield distribution, one has to remove all features that are due to the limited two-particle efficiency and acceptance by applying a correction. This can be done by using the *two-particle* or *pair efficiency*, which is calculated by taking the cross-correlation of the trigger particle efficiency and the associated particle

<sup>2</sup>The roughness in  $\Delta\phi$  does not reflect the tracking efficiency of ALICE, but is due to the more stringent track cuts that we require in our analysis (see Appendix A).

efficiency, denoted by  $\epsilon_{\text{trig}}(\phi, \eta)$  and  $\epsilon_{\text{assoc}}(\phi, \eta)$  respectively, i.e.:

$$\begin{aligned} \epsilon_{\text{pair}}(\Delta\phi, \Delta\eta) &= \frac{A(\Delta\eta)^{-1}}{4\pi\eta_{\text{max}}} \int_0^{2\pi} d\phi \int_{\eta_1}^{\eta_2} d\eta \epsilon_{\text{assoc}}(\phi, \eta) \epsilon_{\text{trig}}(\phi + \Delta\phi, \eta + \Delta\eta) \\ &= \frac{A(\Delta\eta)^{-1}}{4\pi\eta_{\text{max}}} (\epsilon_{\text{assoc}} \star \epsilon_{\text{trig}})(\Delta\phi, \Delta\eta), \end{aligned} \quad (5.2)$$

where  $\eta_{\text{max}}$  indicates the size of the sensitive region of the detector, the integration boundaries are defined as:  $\eta_1 \equiv \max\{-\eta_{\text{max}}, -\eta_{\text{max}} - \Delta\eta\}$  and  $\eta_2 \equiv \min\{\eta_{\text{max}}, \eta_{\text{max}} - \Delta\eta\}$ , and the inverse normalization factor is given by:  $A(\Delta\eta) = (\Delta\eta_{\text{max}} - |\Delta\eta|)/(\Delta\eta_{\text{max}})$ , where  $\Delta\eta_{\text{max}} = 2\eta_{\text{max}}$ . On the second line of Eq. (5.2) it is implied that  $\epsilon_{\text{assoc}}$  and  $\epsilon_{\text{trig}}$  drop to zero at  $\eta = \pm\eta_{\text{max}}$ , so that the integration boundaries  $\eta_{1,2}$  are automatically enforced. Note that the function  $A(\Delta\eta)$  removes the triangular shape, i.e., Eq. (5.2) corrects only for two-particle efficiency, not for two-particle acceptance. As discussed in Sect. 6.7, the single-particle efficiencies necessary to calculate the two-particle efficiency can be determined by means of a MC simulation.

Since the observable in Eq. (5.1) is normalized with the total number of triggers, the full efficiency correction matrix is given by:  $\epsilon_{\text{pair}}(\Delta\phi, \Delta\eta)/\langle\epsilon_{\text{trig}}\rangle$ , where the  $\langle\cdot\rangle$  denotes the average over  $\phi$  and  $\eta$ . Note that if either  $\epsilon_{\text{trig}}$  or  $\epsilon_{\text{assoc}}$  does not depend on  $\phi$  and  $\eta$ , then the two-particle efficiency reduces to:  $\epsilon_{\text{pair}} = \langle\epsilon_{\text{trig}}\rangle \langle\epsilon_{\text{assoc}}\rangle$ , and the full efficiency matrix reduces to a constant factor of  $\langle\epsilon_{\text{assoc}}\rangle$ . Unfortunately neither  $\epsilon_{\text{trig}}$  nor  $\epsilon_{\text{assoc}}$  are sufficiently constant to make this approximation, however, as shown in Eq. (5.4), a somewhat weaker version of this approximation can be used.

There are a number of disadvantages to calculating the pair-efficiency using Eq. (5.2). First of all, the process of propagating MC particles through the detector simulation is computationally expensive, and the amount of available MC events is therefore limited<sup>3</sup>. This causes a large relative statistical uncertainty between the calculated two-particle efficiencies in different  $(\Delta\eta, \Delta\phi)$  bins. Another concern is that this method is relatively sensitive to errors in the description of the virtual detector in the MC simulation.

An alternative is to use the so-called mixed-event technique. It relies on the principle that the distribution of the particle yield from one event associated with trigger particles from another (uncorrelated) event is sensitive to the limited two-particle efficiency and acceptance, but does not contain any two-particle correlations by construction. The resulting distribution is called

---

<sup>3</sup>Especially in Pb–Pb, where the amount of available MC events is far less than the amount of recorded events.

the *mixed-event distribution*, as opposed to the associated yield distribution, which, in contrast, is often referred to as the *same-event distribution*. An example of a mixed-event distribution is shown on the right panel of Fig. 5.1. One major advantage of the mixed-event technique, is that by correlating triggers from one event to particles from many other events, the statistical fluctuations in the mixed-event distribution can be made extremely small, thus avoiding additional (relative) statistical fluctuations in the corrected associated yield distribution.

It is important to stress that the mixed-event distribution does not necessarily contain the same information as the two-particle efficiency. The reason for this is that the mixed-event distribution is sensitive to any feature in the single-particle associated and trigger distribution, regardless of whether it is due to efficiency or due to a true physical effect. For example, if there is more particle production at mid-rapidity than at large rapidity, then there are in fact more pairs with a small rapidity separation, compared to pairs with a large rapidity separation, also when the two members of the pair are taken from uncorrelated events. Hence, the mixed-event distribution would be affected, while the two-particle efficiency remains unchanged. Within the ALICE acceptance, the true-single particle distributions are to a good approximation constant<sup>4</sup>, and in this limit the mixed-event technique is equivalent to the two-particle efficiency. In this work we exclusively use the mixed-event technique for the correction of associated yield distributions, i.e., we use:

$$\epsilon_{\text{pair}}(\Delta\phi, \Delta\eta) = \gamma \text{ME}(\Delta\phi, \Delta\eta)/A(\Delta\eta), \quad (5.3)$$

where  $\text{ME} = \frac{d^2 N_{\text{mixed}}}{d\Delta\phi d\Delta\eta}$ . The proportionality factor  $\gamma$  is given by:

$$\gamma = \frac{\int A(\Delta\eta) \epsilon_{\text{pair}}(\Delta\phi, \Delta\eta) d\Delta\phi d\Delta\eta}{\int \text{ME}(\Delta\phi, \Delta\eta) d\Delta\phi d\Delta\eta} = \frac{\langle \epsilon_{\text{assoc}}(\phi, \eta) \rangle \langle \epsilon_{\text{trig}}(\phi, \eta) \rangle}{2 \langle \text{ME}(\Delta\phi, \Delta\eta) \rangle}, \quad (5.4)$$

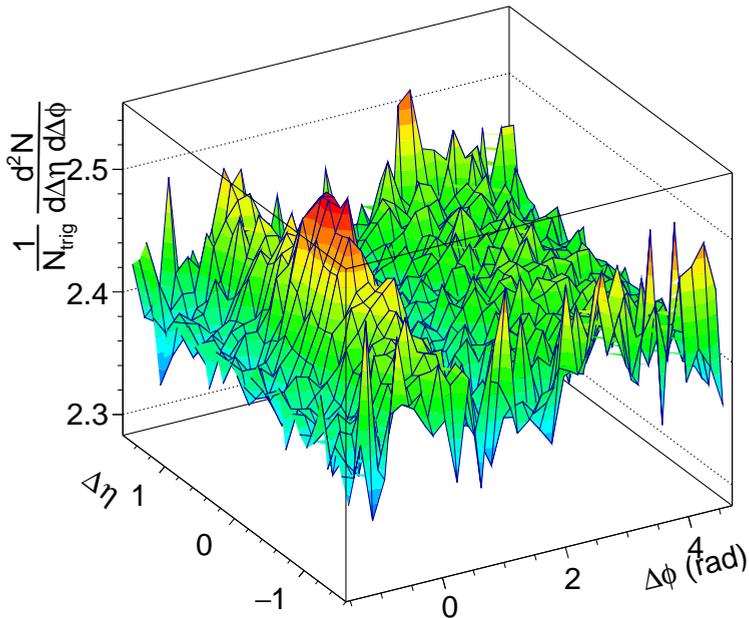
where we used that  $A(\Delta\eta) \epsilon_{\text{pair}}(\Delta\phi, \Delta\eta) = \epsilon_{\text{assoc}} \star \epsilon_{\text{trig}}$ , and the fact that, due to the identity  $\mathcal{F}\{f \star g\}(k) = (\mathcal{F}\{f\}(k))^* \cdot \mathcal{F}\{g\}(k)$ , where  $\mathcal{F}$  denotes the Fourier transform, an integral over a cross-correlation  $f \star g$  is equal to the product of the integrals over the functions  $f$  and  $g$  individually<sup>5</sup>.

<sup>4</sup>At intermediate- $p_T$  this is a good approximation.

<sup>5</sup>This follows from the identity evaluated at  $k = 0$ .

The fully corrected two-particle yield is given by:

$$\begin{aligned} \frac{1}{N_{\text{trig}}} \frac{d^2 N}{d\Delta\phi d\Delta\eta} &= \frac{1}{A(\Delta\eta)} \frac{\langle \epsilon_{\text{trig}} \rangle}{\epsilon_{\text{pair}}(\Delta\phi, \Delta\eta)} \frac{1}{N_{\text{trig}}^{\text{raw}}} \frac{d^2 N^{\text{raw}}}{d\Delta\phi d\Delta\eta} \\ &= \frac{1}{\langle \epsilon_{\text{assoc}} \rangle} \frac{2 \langle \text{ME} \rangle}{\text{ME}(\Delta\phi, \Delta\eta)} \frac{1}{N_{\text{trig}}^{\text{raw}}} \frac{d^2 N^{\text{raw}}}{d\Delta\phi d\Delta\eta}. \end{aligned} \quad (5.5)$$



**Figure 5.2.:** Example of an associated yield distribution in central (0-10%) Pb–Pb collisions, corrected using a mixed-event distribution. No correction for the single-particle efficiency  $\langle \epsilon_{\text{assoc}} \rangle$  is made. Kinematic ranges:  $4.0 < p_{\text{T,trig}} < 10.0$  GeV/ $c$ , and  $1.2 < p_{\text{T,assoc}} < 1.3$  GeV/ $c$ .

In Fig. 5.2 we show the associated yield from the left panel of Fig. 5.1, corrected with the mixed-event distribution from the right panel of that same figure. In this figure, the jet-peak at the origin is clearly visible, on top of a ridge-structure, extending in  $\Delta\eta$ . Also at  $\Delta\phi = \pi$  we see a ridge-structure, which is somewhat broader than the one  $\Delta\phi = 0$ . The physics behind these

features is introduced in Chapters 1 and 2, and in Chapter 6 we discuss methods to quantify them.

Some words of caution are in order when one uses a mixed-event procedure. Other analyses show that at low  $p_{T,\text{assoc}}$  it is important to perform mixed-event correction in a narrow bins in centrality and in the  $z$ -coordinate of the primary vertex ( $z_{\text{vtx}}$ ), i.e.:

$$\frac{d^2 N}{d\Delta\phi d\Delta\eta} \propto \sum_{i \in \text{bins}} \frac{\text{SE}_i}{\text{ME}_i}, \quad (5.6)$$

where SE represents the same-event distribution, and *bins* stands for bins in centrality and  $z_{\text{vtx}}$ . If one does not do this, the mixed-event corrected associated yield distribution becomes slightly concave up at large  $\Delta\eta$ , an effect caused by the small acceptance difference between events with a primary vertex at a different location.

In Sect. 5.2 we discuss a method to identify the same and mixed-event distributions. This method relies on having a sufficient number of particles per  $(\Delta\phi, \Delta\eta)$  bin. The mixed-event correction as described in Eq. (5.6) requires us to first identify the same and mixed-event distributions for every centrality and  $z_{\text{vtx}}$  bin separately, which is not feasible given the size of the current datasets. For this reason we perform our mixed-event correction by taking the ratio of the sums instead of the sum of the ratio, i.e.:

$$\frac{d^2 N}{d\Delta\phi d\Delta\eta} \propto \frac{\sum_{i \in \text{bins}} \text{SE}_i}{\sum_{j \in \text{bins}} \text{ME}_j}. \quad (5.7)$$

The expected impact of using Eq. (5.7) instead of Eq. (5.6) is largest for the associated yield in Pb–Pb collisions at low  $p_{T,\text{assoc}}$ , since there the jet-peak is relatively small compared to the background. The corrected associated yield in Fig. 5.2 shows that the height of the jet-peak in Pb–Pb at low  $p_{T,\text{assoc}}$  is typically of the order of a percent of the background. In the same  $p_{T,\text{assoc}}$  bin, the difference between using Eq. (5.7) and Eq. (5.6) is found to be three orders of magnitude smaller, and can therefore be safely neglected.

Finally, when creating a mixed-event distribution, one typically wants to mix the same trigger and associated particles multiple times, so that the statistical uncertainty on the mixed-event distribution can be made much smaller than the statistical uncertainty on the same-event distribution. In practice, when analyzing a dataset one creates a so-called “trigger pool”, which keeps track of trigger particles from previous events, in order to correlate

them to the associated particles of the current event. In our implementation for example, we store trigger particles of 50 events when analyzing central Pb–Pb data, and 150 events when analyzing p–Pb or pp data. The main reason for this difference is that central Pb–Pb events typically contain many more trigger particles than p–Pb or pp events.

It is important to make sure that we do not use one trigger more often than the other. In order to achieve this, we only begin mixing events when the trigger pool is full, i.e., for every centrality and vertex bin we have triggers from 50 (150) events stored. We then record the event at which we start mixing, and after the analysis is completed, we go back to the first event and mix the events up to the point that event mixing started.

## 5.2. Identifying Associated Yield

To create identified associated yield distributions<sup>6</sup>, we start by filling three five-dimensional histograms, one for each particle species  $i \in \{\pi, K, p\}$ :

$$N_i(p_T, \Delta\eta, \Delta\phi, T_i, X_i), \quad (5.8)$$

where  $\Delta\eta$  and  $\Delta\phi$  are taken with respect to a high- $p_T$  trigger, and  $X_i$  and  $T_i$  were defined before, in Eq. (4.1) and Eq. (4.4) respectively. The five-dimensional histogram essentially contains a  $(T_i, X_i)$  plane for every  $(p_T, \Delta\eta, \Delta\phi)$  bin, from which the identified particle yield can be extracted using a fit.

In Chapter 4 we described how fits of  $(T_i, X_i)$  distribution can be used to obtain the identified inclusive yield. Identifying the yield associated with a high- $p_T$  trigger particle through a fitting procedure is more involved for a number of reasons, most importantly:

- Many events do not contain any high- $p_T$  trigger particle, effectively reducing the dataset.
- In bins at large  $\Delta\eta$  the particle yield is very small due to “triangular” acceptance in  $\Delta\eta$ , see Fig. 5.1.
- A single bin in  $\Delta\eta$  can contain tracks of a large interval in  $\eta$  (see Fig. 5.3). In Sects. 4.1 and 4.2 we showed that the detector response varies significantly as a function of  $\eta$ , hence the PID signal in a  $\Delta\eta$

---

<sup>6</sup>For the identified mixed-event distribution the procedure is analogous.

bin is “smeared out”, and the model in Eq. (4.9) no longer provides a sufficiently accurate description of the data.

In this section we show how the results from Chapter 4 can be used to construct template distributions for the pion, kaon, and proton peaks in the  $(T_i, X_i)$ -plane, for  $i \in \{\pi, K, p\}$ , and every  $(p_T, \Delta\eta)$  bin, designed to match the smearing caused by the non-trivial relation between  $\eta$  and  $\Delta\eta$ <sup>7</sup>. For the TOF-mismatched yield a template can be generated in the same way. Note that the fit that remains to be done for every  $(p_T, \Delta\eta, \Delta\phi)$  bin only has a four-dimensional parameter space, each one being the scale factor for one of the templates. This immediately addresses also the first two points on the list above, as a low dimensional template fit is much less likely to fail compared to the seventeen-dimensional fit from Chapter 4, especially when the histogram contains only a small number of entries. Moreover, the template fit converges much faster than the seventeen-dimensional fit. This is a clear advantage, as the number of independent fits per collision system is around 7,500.

### 5.2.1. Fit Templates

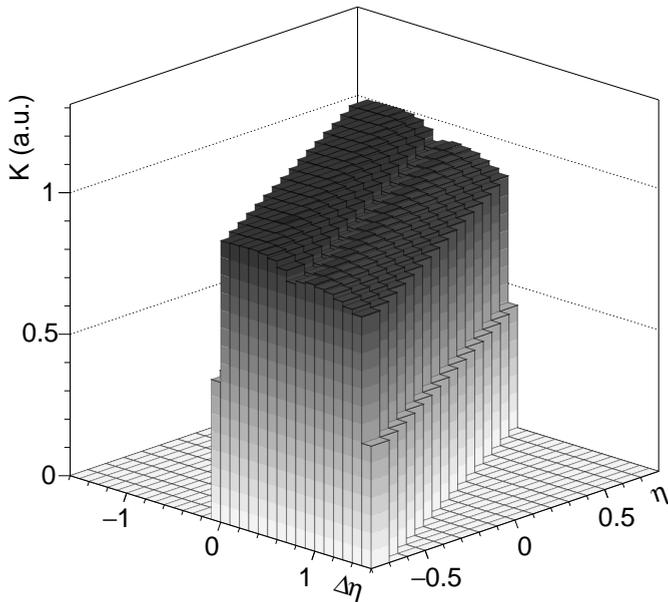
The principle behind the fit templates is that the detector response as a function of  $\Delta\eta$  can be approximated by an appropriately weighted sum over the model functions (see Eq. (4.9)), fitted to the inclusive data binned in  $\eta$ , i.e.,

$$P_{ij}^{\text{assoc}}(T_j, X_j; \Delta\eta) = \sum_{\eta} K(\eta, \Delta\eta) P_{ij}^{\text{incl}}(T_j, X_j; \eta), \quad (5.9)$$

where  $P_{ij}^{\text{incl}}$  denotes the inclusive model functions fitted to the inclusive data for the peak of species  $i \in \{\pi, K, p\}$  in the  $(T_j, X_j)$ -plane,  $K(\eta, \Delta\eta)$  is a matrix containing the weight factors, and  $P_{ij}^{\text{assoc}}$  is the fit template.

The first step in generating the fit templates is to determine the matrix  $K(\eta, \Delta\eta)$ , shown in Fig. 5.3. If the detector efficiency would be independent of  $\eta$  and  $\phi$ , a simple analytic expression would suffice, i.e.,  $-\eta_{\text{max}} - \Delta\eta < \eta_{\text{assoc}} < \eta_{\text{max}} - \Delta\eta$ , where  $\eta_{\text{max}} = 0.8$  in our analysis. However, since not all parts of the detector are equally likely to register a trigger particle, we generated the matrix  $K$  by building a random pair distribution with trigger particles from the dataset with random values of  $\eta$  in the range

<sup>7</sup>Notice that the shape of the template is unaffected by  $\Delta\phi$ .



**Figure 5.3.:** The matrix  $K(\eta, \Delta\eta)$ , which is proportional to the probability distribution of finding an associated particle at  $\eta$ , as part of a pair with a separation  $\Delta\eta$ , while only taking into account the detector efficiency of the trigger particles. The efficiency of the associated particles is taken into account by explicitly using the fit results  $P_{ij}^{\text{incl}}$  in Eq. (5.9).

$-\eta_{\text{max}} < \eta < \eta_{\text{max}}$ . Notice from Fig. 5.3 that slightly more trigger particles are measured in the range  $\eta < 0$ , causing a small deviation compared to the ideal case<sup>8</sup>. The efficiency of measuring the associated particles is not taken into account by the matrix  $K$ . This can, however, be remedied by using the actual (unnormalized) fit results for  $P_{ij}^{\text{incl}}$ . This ensures that possible differences between the efficiency of different particle species are correctly propagated to the fit templates.

<sup>8</sup>This can most clearly be seen at  $\Delta\eta = 0$ , which corresponds exactly to the likelihood of observing a trigger particle at  $\eta$ .

### 5.2.2. Fit Procedure

Once the fit templates are generated according to Eq. (5.9), all  $(T_i, X_i)$  projections corresponding to single  $(p_T, \Delta\eta, \Delta\phi)$  bins in the 5 dimensional matrix (Eq. (5.8)) can be fitted using the following model:

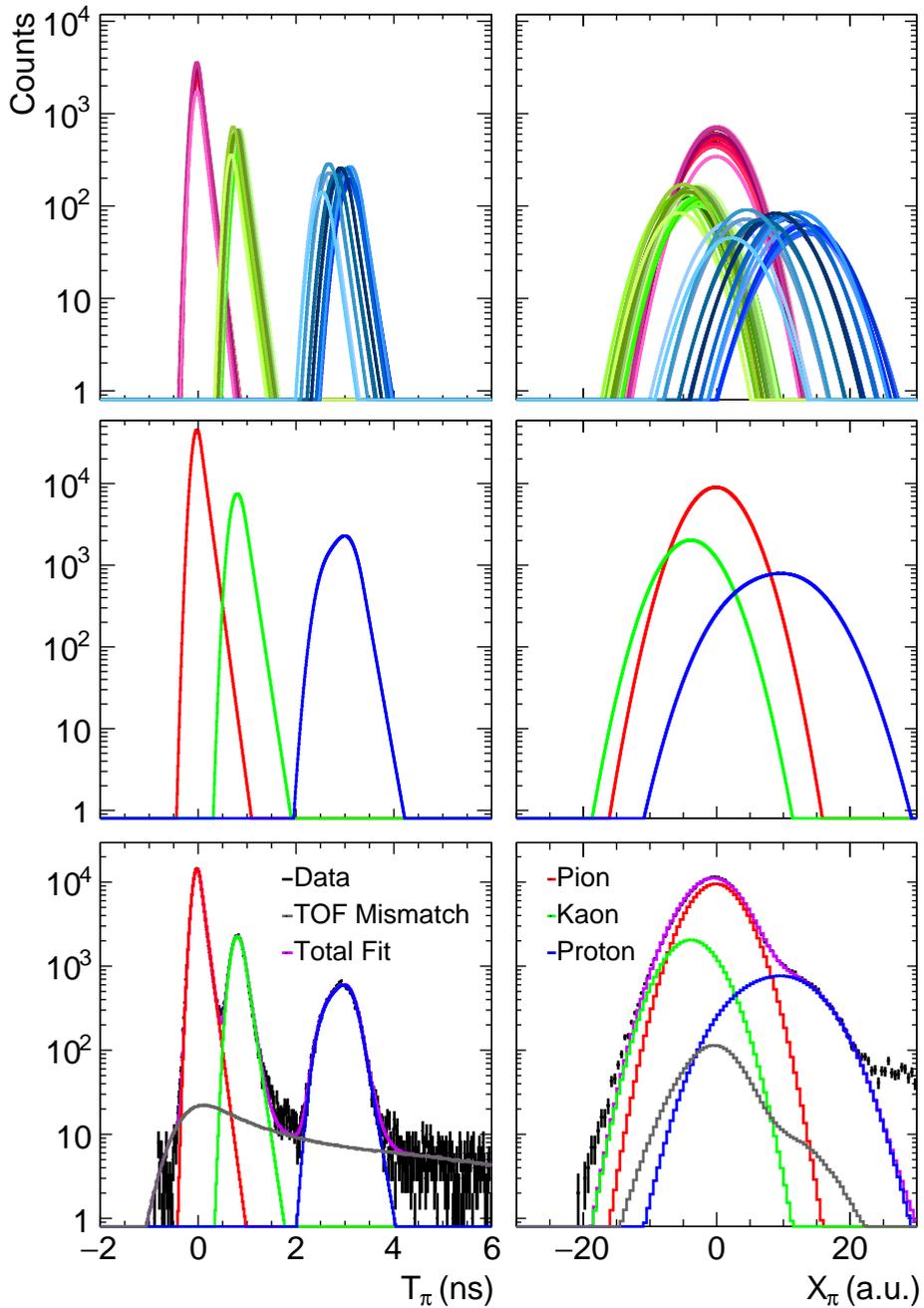
$$\frac{d^2 N_{\text{tot}}}{dT_j dX_j}(p_T, \Delta\eta, \Delta\phi) = \sum_i N_{ij} P_{ij}^{\text{assoc}}(p_T, \Delta\eta), \quad (5.10)$$

where the  $N_{ij}$  are the fit parameters and index  $i$  is assumed to run over all included particle species, as well as the mismatch template.

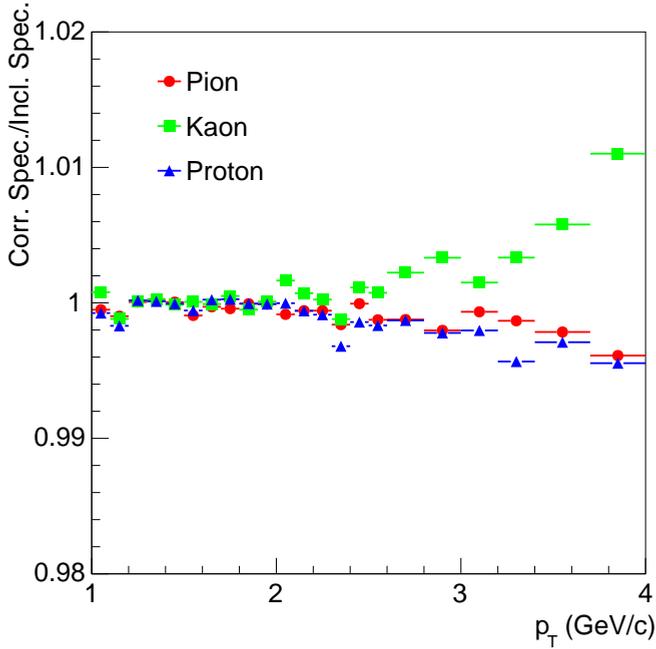
From this point on, the fit procedure is essentially the same as the one that we described in Sect. 4.3. First, the number of mismatched tracks is determined, in a region far away from the peaks. After that, the optimal values for the remaining parameters  $N_{ij}$  are found by minimizing the negative log-likelihood assuming Poisson-distributed data in every bin, and in a restricted region  $R$  that includes all peaks (see Eq. (4.11)).

In Fig. 5.4 we show an example of how a template is built up from different fit results. This example shows that the smeared-out peaks are well-described by our templates. We wish to emphasize that we use this fitting procedure on the same and mixed-event distributions separately, and the mixed-event correction is only done afterwards. It is essential that the order of these steps is not reversed, because the corrected data is not Poisson distributed. Especially at large  $\Delta\eta$ , where the number of entries is typically small, this becomes an issue.

In pp events, high- $p_T$  trigger particles are rare, and the number of particles per event is small compared to p-Pb and Pb-Pb events, which, especially at large  $\Delta\eta$  and/or high  $p_T$  leads to relatively sparsely filled  $(T_i, X_i)$  histograms. In these cases it can happen that the region where the number of mismatches is determined is (almost) empty. In this case, the large uncertainty on the number of mismatches significantly influences the quality of the entire fit. For this reason, we demand that there are at least ten entries in the region where the mismatches are determined. For the histograms where this condition is not met, we assume that the fraction of mismatches is the same as in the mixed-event fit of the same  $(p_T, \Delta\eta, \Delta\phi)$  bin. This is a viable solution, as the amount of data in the mixed-event distribution is about two orders of magnitude larger than in the same-event distribution.



**Figure 5.4.:** Template decomposition of one  $(\Delta\phi, \Delta\eta)$  bin of an associated yield distribution. Top row: Different tints represent the fits at different  $\eta$ , middle row: templates (not normalized), bottom row: template fits. Collision system: p-Pb, kinematic range:  $1.2 < p_{T, \text{assoc}} < 1.3$  GeV/ $c$ , and  $0.0 < \Delta\eta < 0.2$ .



**Figure 5.5.:** The ratio of identified spectra, measured with the template method and integrated over  $(\Delta\phi, \Delta\eta)$ , and the identified spectra measured with the inclusive method (see Chap. 4). Collision system: Pb–Pb.

### 5.3. Systematic Uncertainty

To assess the performance of our template fit method, we applied it on a dataset generated by a Monte Carlo simulation, where the detector response is simulated using GEANT3 (see Sect. 6.7). The number of particles determined using the template PID method shows no significant difference with the number of generated particles ( $\pi$ ,  $K$ ,  $p$ ). There are, however, two reasons why this first check is not sufficient in our opinion. First of all, the available MC dataset of Pb–Pb events is much smaller than the real dataset, limiting the  $p_T$  range in which the check can be performed properly. Second, the simulated PID response differs from the real PID response. The shape of the “mismatch band” is significantly different, and the separation power in MC is typically larger than in reality.

To further assess the performance of the template PID method, we compared the template PID method directly with the inclusive PID method from Chapter 4, by performing the following procedure:

- Fill the  $N_i(p_T, \Delta\eta, \Delta\phi, T_i, X_i)$  as usual.
- Fill a histogram  $N_i(p_T, \eta, T_i, X_i)$  with exactly the same tracks. This means that if an event contains multiple trigger tracks, then the associated tracks are filled multiple times.
- Perform the template PID method on the first histogram and the inclusive PID method (Chapter 4) on the second, and calculate the ratio of the particle yields.

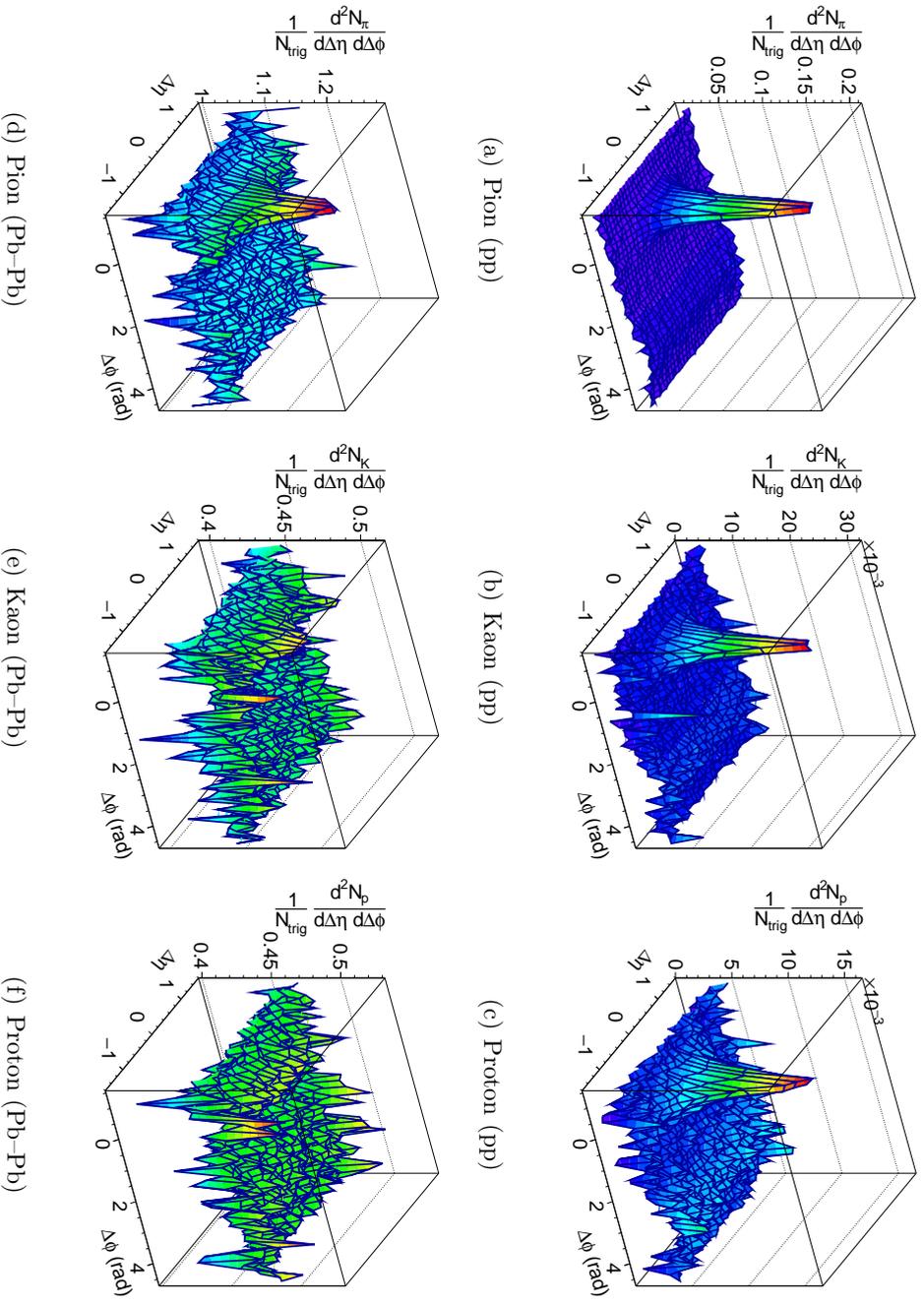
The results of this study are shown in Fig. 5.5. As the agreement of this comparison is very good, we assume that the systematic uncertainties of the inclusive PID method carry over to the template PID method.

## 5.4. Fit Results

In this analysis we created both same-event and mixed-event associated yield distributions, for pions, kaons, and protons in twenty-two different  $p_T$  bins, for three different collision systems. In Fig. 5.6 we show the fit results for pions, kaons, and protons, in the range  $1.8 < p_T < 2.2$  GeV/ $c^9$ . The associated yield distributions of Pb–Pb collisions distinguish themselves by their large uncorrelated background, and as a consequence, a relatively small signal (jet-peak) over background ( $S/B$ ) ratio. In pp collisions, the  $S/B$  ratio is generally large, due to a very small uncorrelated background. In the pion associated yield of Pb–Pb, the previously discussed correlated background (“ridges” at  $\Delta\phi = 0$  and  $\Delta\phi = \pi$ ) is clearly visible. For the other particle species these features are largely obscured by the statistical uncertainty on the data points. The associated yield distributions in p–Pb (not shown) are qualitatively between the pp and Pb–Pb distributions. The background density is larger than in pp, but much smaller than in Pb–Pb, and also the  $S/B$  ratio is between that of the pp and Pb–Pb associated distributions.

---

<sup>9</sup>This is actually the sum over four independent fit results, as the original binning had a granularity of  $\Delta p_T = 0.1$  GeV/ $c$ .



**Figure 5.6.:** Associated yield distribution of pions, kaons, and protons, in the range  $1.8 < p_{T,\text{assoc}} < 2.2$  GeV/ $c$ , for pp and Pb-Pb.

## 5.5. Summary

In this chapter we showed that by means of a template fit method, the yield associated with a high- $p_T$  trigger particle can be measured in an accurate and efficient way. This method explicitly uses the inclusive fit results, which were discussed in Chapter 4. Although the template fit method is more involved than a simple cut on the detector signal, it clearly has a number of advantages. Similar to the inclusive PID method, these include an independence of the detector calibration, no misidentified particles, and a better separation power at high  $p_T$ .



A dark, textured rectangular background with a white, stylized symbol resembling a vertical dollar sign with a horizontal bar, positioned on the right side. The text 'Chapter 6.' is written in white on the left side.

## Chapter 6.

# Associated Yield Analysis

We now turn to the final part of our analysis, which is the analysis of the identified associated yield distributions. The most straightforward way to analyze these distributions would be to fit them to a very general model describing both the near- and away-side peak structures and the flow background. This approach was, however, unsuccessful, since a model that is sufficiently general to describe all the typical features of an associated yield distribution has so much freedom that the resulting fits were typically unstable. We therefore chose to start by analyzing projections of the associated yield on the  $\Delta\eta$ -axis (Sect. 6.1) and the  $\Delta\phi$ -axis (Sect. 6.2) independently.

An interesting observation that we made from the  $\Delta\eta$  and  $\Delta\phi$  analyses was that the “width” of the jet peak decreases as a function of  $p_T$  and as a function of particle mass. On the other hand, we did not see evidence for a significantly different jet-peak shape in the different collision systems. In principle, the  $\Delta\eta$  and  $\Delta\phi$  analyses provide us with an estimate of the per-trigger jet-peak yield. We found, however, that these estimates were not very reliable, for reasons that we elaborate upon in Sects. 6.1 and 6.2.

In Sects. 6.3 and 6.4 we describe a different method to determine the per-trigger yield in the jet peak, using the background model that we obtained from the  $\Delta\phi$  analysis and the estimate of the peak width from both the  $\Delta\phi$  and  $\Delta\eta$  analyses. This method entails subtracting the background from the  $(\Delta\eta, \Delta\phi)$ -distributions, and simply counting the remainder in a circular region around the origin. The radius of the circular region that we use is proportional to the jet-peak “width”, determined from the independent  $\Delta\eta$  and  $\Delta\phi$  analyses.

The associated yield distributions of the different collision systems generally do not have the same prominent features. For example, in pp events there

is no flow background, and the distribution can be completely characterized by the near-side jet peak at  $(\Delta\eta, \Delta\phi) = (0, 0)$ , and the away-side jet “ridge” at  $\Delta\phi = \pi$ . In central Pb–Pb events on the other hand, the flow background dominates at low  $p_T$ , and the away-side jet is absent. In an attempt to avoid large differences in the systematic uncertainty on the results between collision systems, we made the analysis procedure as similar as possible.

In Sect. 6.5 we describe a study that we performed to assess the robustness of the method described in Sects. 6.3 and 6.4, by applying it to data which was generated by a toy model. After that, in Sect. 6.6, we outline how one can correct the identified associated yield for secondary particles. In Sects. 6.7 and 6.8 we describe how we corrected the identified associated yield for detector efficiency and we estimate the size of the most important sources of systematic uncertainty. Finally, in Sect. 6.9 the results are shown and we discuss the most interesting observations in some detail.

## 6.1. A Model for the $\Delta\eta$ Projection

We start by analyzing the  $\Delta\eta$  projections of the near-side ( $-\frac{\pi}{2} < \Delta\phi < \frac{\pi}{2}$ ). The advantage of studying the  $\Delta\eta$  projection is that we do not need to consider flow, as this is assumed to be independent of  $\Delta\eta$ . The drawback is that due to the approximately triangular acceptance in  $\Delta\eta$  (see Sect. 5.1), the statistical uncertainty at large  $\Delta\eta$  is large.

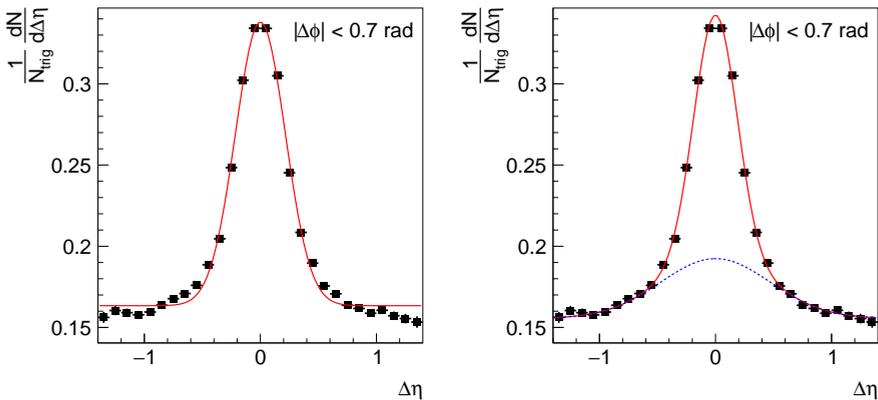
The simplest model that we tried to fit to the  $\Delta\eta$  projection was a single Gaussian on top of a constant background. As can be seen on the first panel of Fig. 6.1, this model does not possess sufficient freedom, and, as can be seen on the second panel of the same figure, we achieved much better fit results by using a double Gaussian on top of a constant background, i.e.,

$$\frac{dN}{d\Delta\eta} = C + N_{\text{jet}} \left[ \frac{f}{\sqrt{2\pi}\sigma_1} \exp\left\{-\frac{\Delta\eta^2}{2\sigma_1^2}\right\} + \frac{1-f}{\sqrt{2\pi}\sigma_2} \exp\left\{-\frac{\Delta\eta^2}{2\sigma_2^2}\right\} \right], \quad (6.1)$$

where  $C$ ,  $N_{\text{jet}}$ ,  $f$ ,  $\sigma_1$  and  $\sigma_2$  are fit parameters. Note that we use the fit parameters  $N_{\text{jet}}$  and  $f$  instead of, for example,  $N_{\text{jet},1}$  and  $N_{\text{jet},2}$ . The reason for this is that by using this parameterization, the statistical uncertainty on  $N_{\text{jet}}$  can immediately be determined, and we do not explicitly need to consider the anti-correlation between  $N_{\text{jet},1}$  and  $N_{\text{jet},2}$ .

In the example of Fig. 6.1, we observe that the jet peak can be characterized as a “narrow” Gaussian peak on top of a “broad” Gaussian background. This qualitative description turns out to apply to the jet peaks of pions,

kaons, and protons alike, irrespective of the collision system or  $p_T$  range<sup>1</sup>. As is shown in a moment, the width of the complete jet-peak structure in  $\Delta\eta$  typically decreases with increasing  $p_T$  or decreasing particle species mass. When we compare the single- and double-Gaussian fit, we find that the



**Figure 6.1.:** Fits of the  $\Delta\eta$ -projection of the pion associated yield distribution in p-Pb collisions with a single Gaussian (left), and a double Gaussian (right). The kinematic range is:  $1.8 < p_{T,\text{assoc}} < 2.2$  GeV/c.

width from the single-Gaussian fit falls between the two widths found in the double-Gaussian fit, but is most comparable to the width of the narrowest Gaussian. In case of the example in Fig. 6.1 we found for the single-Gaussian fit:  $\sigma \approx 0.22$ , while for the double-Gaussian fit we found:  $\sigma_1 \approx 0.18$ , and:  $\sigma_2 \approx 0.41$ .

The example shown in Fig. 6.1 is relatively easy to fit, since pions are the most abundant species, and they have the most narrow jet-peak structure, so that even the broad background structure generously falls within the limited  $\Delta\eta$ -acceptance. In many other instances however, the projections were much harder to fit, and the statistical uncertainties on the fit parameters became prohibitively large.

<sup>1</sup>Another parameterization that gives a reasonable description of the data is a properly normalized t-distribution on top of a constant background, with a scaled dependent parameter  $\Delta\eta$ , i.e.,  $dN/d\Delta\eta = A \cdot T(B\Delta\eta, C) + D$ , where  $T(x, \nu)$  represents the t-distribution and  $A - D$  are fit parameters. Unfortunately, our fits did not produce a smooth dependence of the fit parameters on  $p_T$ .

## 6.2. A Model for the $\Delta\phi$ Projection

We now turn to the analysis of the  $\Delta\phi$  projection of the distribution of associated particles. A clear advantage of studying the  $\Delta\phi$  projection over the  $\Delta\eta$  projection is that all data points have a comparable statistical uncertainty. Furthermore, the  $\Delta\phi$  analysis should give a better estimate of the background yield than the  $\Delta\eta$  analysis, since it does not rely only on the data at large  $\Delta\eta$ . On the other hand, the model that we use to describe the data is more complicated due to the presence of a significant away side jet peak at  $\Delta\phi \sim \pi$ , and a flow background in case of p-Pb and Pb-Pb.

In its most general form, the model that we use is the following<sup>2</sup>:

$$\begin{aligned} \frac{dN}{d\Delta\phi} = & \left\langle \frac{dN_{\text{bgd}}}{d\Delta\phi} \right\rangle (1 + 2 \sum_n V_{n\Delta} \cos n\Delta\phi) \\ & + N_{\text{jet,NS}} \left[ \frac{f}{\sqrt{2\pi}\sigma_{\text{NS1}}} \exp \left\{ -\frac{\Delta\phi^2}{2\sigma_{\text{NS1}}^2} \right\} + \frac{1-f}{\sqrt{2\pi}\sigma_{\text{NS2}}} \exp \left\{ -\frac{\Delta\phi^2}{2\sigma_{\text{NS2}}^2} \right\} \right] \\ & + \frac{N_{\text{jet,AS}}}{\sqrt{2\pi}\sigma_{\text{AS}}} \exp \left\{ -\frac{(\Delta\phi-\pi)^2}{2\sigma_{\text{AS}}^2} \right\}, \end{aligned} \quad (6.2)$$

where  $V_{n\Delta}$  are the pair-anisotropy coefficients, which have been shown to be a good description of long-range  $\Delta\eta$  correlations in Pb-Pb [16]. The subscript ‘‘NS’’ refers to the near-side jet peak at  $\Delta\phi \sim 0$ , and the subscript ‘‘AS’’ refers to the away-side peak. As in Eq. (6.1), we use the fit parameters  $N_{\text{jet,NS}}$  and  $f$  for the NS jet.

Our first attempt was to directly fit the model in Eq. (6.2) to the data. Dependent on the collision system, we fixed several parameters to zero. In case of pp we neglected any anisotropies in the background, i.e., we fixed  $V_{n\Delta} = 0$ , but left all other parameters free. For p-Pb it was shown that the only significant contribution to the azimuthal anisotropy came from the second Fourier coefficient [95], hence we fixed  $V_{n\Delta}$  to zero for all  $n$ , except for  $n = 2$ . Finally, in case of central Pb-Pb collisions we allowed  $V_{n\Delta}$  to vary for  $n = 1 - 4$ . The anisotropy coefficients for  $n > 4$  are small compared to the lower orders, and are therefore neglected. The away-side jet is hard to distinguish from (a combination of) lower-order anisotropy coefficients, and its influence is so small that we fixed  $N_{\text{jet,AS}} = 0$ .

As in the analysis of the  $\Delta\eta$  projection, one of the main complications with this approach was that while the double-Gaussian description of the NS peak is clearly superior to the single-Gaussian description, fits with the double-Gaussian model tend to be unstable. Another complication was that

<sup>2</sup>In the actual implementation of the model it is important to take the implied periodicity in  $\Delta\phi$  into account.

the statistical uncertainties on the proton and kaon distributions in Pb–Pb were often too large to give reliable fit results.

### 6.3. Jet-Peak Yield Measurement

Since we found that for a substantial part of the data a direct fit to the  $(\Delta\eta, \Delta\phi)$ -plane, as well as to its projections, did not provide a viable option for measuring the jet-peak yield, we devised an alternative method. In the introduction of this chapter we already briefly mentioned this approach, and in this section we work it out further.

Using this method one measures the near-side jet-peak yield  $N_{\text{jet},i}$ <sup>3</sup>, the average background density  $d^2N_{\text{bgd},i}/d\Delta\phi d\Delta\eta$ , the anisotropy coefficients  $V_{n\Delta}$ , and the near-side peak widths  $\sigma_{\Delta\eta}$  and  $\sigma_{\Delta\phi}$ . The following steps outline the method:

1. Remove all bins outside the range  $-1.4 < \Delta\eta < 1.4$ . The statistical uncertainty on the data outside this range is extremely large. In fact, for pp data at high  $p_{\text{T}}$ , these bins often had no data at all.
2. Project the  $(\Delta\eta, \Delta\phi)$ -distribution onto the  $\Delta\eta$ -axis, and fit the resulting distribution with a constant plus a Gaussian (see the first panel of Fig. 6.1). The fitted standard deviation,  $\sigma_{\Delta\eta,i}(p_{\text{T}})$ , where  $i \in \{\pi, K, p\}$ , serves as a  $p_{\text{T}}$  and species-dependent estimate of the jet-peak width.
3. Project the  $(\Delta\eta, \Delta\phi)$ -distribution onto the  $\Delta\phi$ -axis, but ignore all bins with a bin center obeying the following inequality:  $\sqrt{\Delta\eta^2 + \Delta\phi^2} \leq \min\{\alpha\sigma_{\Delta\eta,i}(p_{\text{T}}), 1\}$ . The implications of various choices for the value of the factor  $\alpha$  are discussed later in this section. The maximum radius of one is chosen so that there are still sufficient data points left at  $\Delta\phi \sim 0$  to determine the background (see next step).
4. Under the assumption that the jet peak does not influence the projection from the previous step, we can determine the constant background, as well as the anisotropy coefficients by applying a discrete cosine transform (DCT). The relation between the DCT coefficients and the anisotropy coefficients is described in Appendix C. In case of pp and p-Pb, only the data in the range  $-\frac{\pi}{2} < \Delta\phi < \frac{\pi}{2}$  can be used, as for these systems the away-side jet is assumed to dominate the correlation

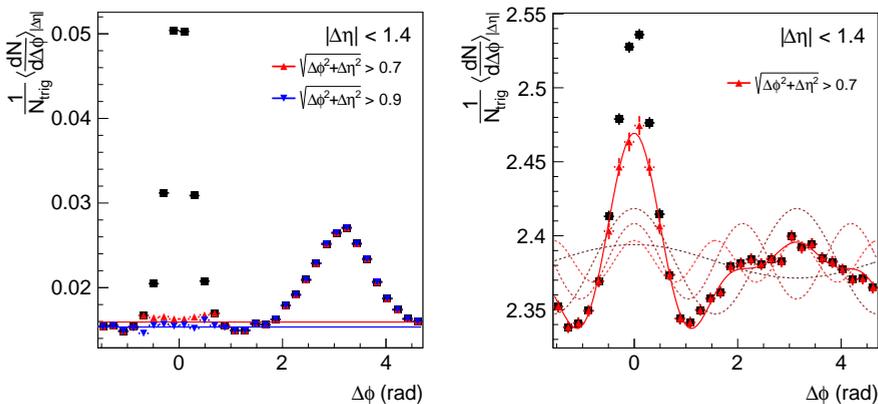
---

<sup>3</sup>The subscript “NS” for near-side will from here on be omitted.

structure in the range  $\frac{\pi}{2} < \Delta\phi < \frac{3\pi}{2}$ . Two examples of this step are shown in Fig. 6.2. Additional examples are shown in Appendix D.

5. We now create a background-subtracted  $\Delta\phi$ -projection of the associated yield by using the results from the previous step. We do this by subtracting a constant (pp), a constant plus second harmonic (p-Pb), or a constant plus the first four harmonics (Pb-Pb). The resulting distribution is then fitted with a constant plus a single Gaussian at  $\Delta\phi = 0$  and at  $\Delta\phi = \pi$  (pp and p-Pb), or only a single Gaussian at  $\Delta\phi = 0$  (Pb-Pb). From the Gaussian fitted to the NS peak we find  $\sigma_{\Delta\phi,i}$ .
6. From the analysis of both pp and p-Pb datasets we find that the fitted standard deviation of the  $\Delta\phi$  and  $\Delta\eta$  projections are approximately equal, i.e.,  $\sigma_{\Delta\eta,i}(p_T) \approx \sigma_{\Delta\phi,i}(p_T)$  (see Fig. 6.3). Furthermore, the peak widths seem to agree very well among these two systems. The peak widths measured from the Pb-Pb dataset are compatible as well, however, for all but the pion measurement, the statistical uncertainty is very large. Based on this observation, we define  $\sigma_{\text{jet},i} \equiv \text{Avg}\{\sigma_{\Delta\eta}^{\text{pp}}, \sigma_{\Delta\phi}^{\text{pp}}, \sigma_{\Delta\eta}^{\text{p-Pb}}, \sigma_{\Delta\phi}^{\text{p-Pb}}\}$ . As can be seen in Fig. 6.3, the function  $\sigma_{\text{jet},i}(p_T)$  is very well described by a polynomial of order two. At this point we repeat steps 3-5, but use the polynomial fit of  $\sigma_{\text{jet},i}(p_T)$  as an estimate for the peak width instead of the measurement of  $\sigma_{\Delta\eta,i}(p_T)$ , thereby improving the measurement of the background parameters. After that, we continue with measuring the jet-peak yield in the next step.
7. The jet-peak yield is measured by taking a sum over all bins with a bin center within the disc  $\sqrt{\Delta\eta^2 + \Delta\phi^2} = \min\{\alpha\sigma_{\text{jet},i}(p_T), 1\}$ , subtracting from each bin the background from the previous steps, evaluated in the center of that bin. The statistical uncertainties on the pair anisotropy coefficients as well as on the constant background are propagated to the peak-yield measurement by adding them in quadrature, properly scaled by a factor  $\partial N/\partial\sigma_{V_{n\Delta}}$ . The way to determine these statistical uncertainties from the data is described in Appendix C.

In the following we refer to the set of bins that are used to estimate the background parameters ( $\langle dN_{\text{bgd}}/d\Delta\phi \rangle$  and  $V_{n\Delta}$ ) as the *background region*, and to the set of bins that are used for counting the number of particles in the jet peak as the *jet-peak region*.

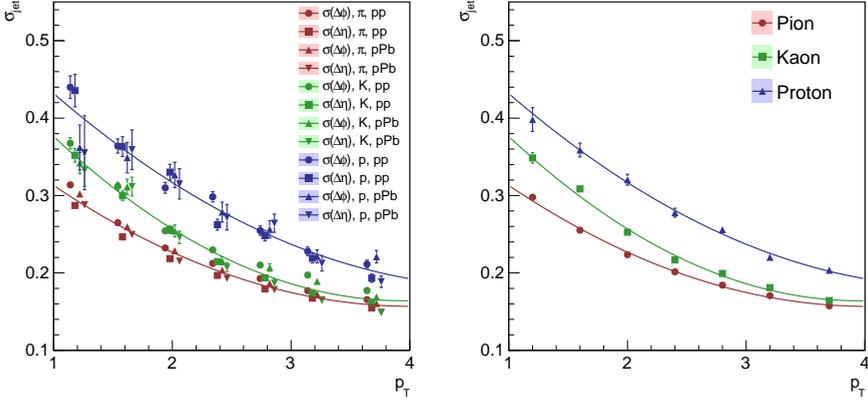


**Figure 6.2.:** Examples of background measurements in pion associated yield, in pp (left panel) and Pb–Pb (right panel). The total background ( $\alpha = 3$ ) is indicated with a solid red line in both panels, and the anisotropy coefficients for  $1 \leq n \leq 4$  are shown as dashed lines. As a comparison, in the left panel the background measurement with  $\alpha = 4$  is shown in blue. The kinematic range is:  $1.8 < p_{T,\text{assoc}} < 2.2 \text{ GeV}/c$ .

From the example of the background estimate of typical pp data shown in Fig. 6.2, it is clear that when  $\alpha = 3$  the measurement of the uncorrelated background (solid red line) is biased. As expected, this bias is significantly reduced when  $\alpha = 4$ . In Sect. 6.4 we outline the by that should be considered when choosing a value for  $\alpha$  and in Sect. 6.5 we discuss how the associated bias influences the peak-yield and background measurement.

In Fig. 6.3 we show the result of the jet-peak width measurements. We observe that for all  $p_T$  bins  $\sigma_{\text{jet},\pi} < \sigma_{\text{jet},K} < \sigma_{\text{jet},p}$ . It can also be seen that in the  $p_T$  range of this analysis, the maximum jet-peak width (protons at low  $p_T$ ) is around 0.4. The average jet-peak width measurements are well described by a second order polynomial, i.e.:  $\sigma_{\text{jet}} = a_0 + a_1 p_T + a_2 p_T^2$ , where the coefficients  $a_i$  are given in Table 6.1.

The stacked bar charts in Fig. 6.4 are meant to give an impression of the jet-peak yield measurement. In this figure, the background is represented by the blue bars, with the jet peak stacked on top in red. This figure is created as follows: all the bars outside of the jet-peak region belong to the background and are colored blue, while the bars within the jet-peak region



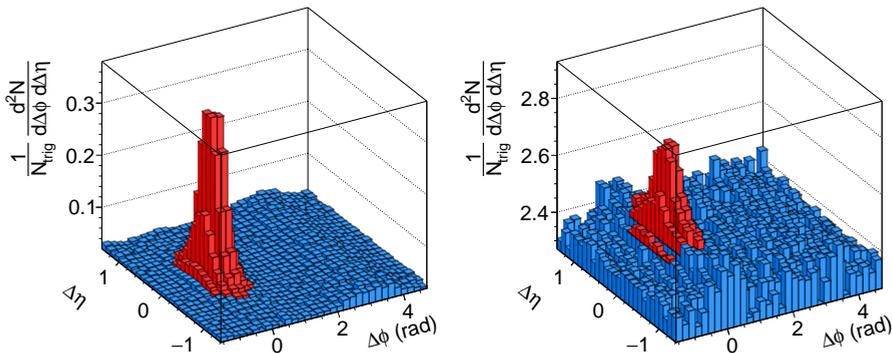
**Figure 6.3.:** Peak widths determined with a Gaussian fit of  $dN/d\Delta\phi$  and  $dN/d\Delta\eta$  for pp and p-Pb (left panel), and the average of the peak width measurements (right panel). The solid line is a polynomial fit of order two through the averaged points. The “clusters of points” in the left panel are all measured at the same  $p_T$ , however, the points are spread out for clarity.

	$a_0$	$a_1$	$a_2$
$\pi$	$0.433 \pm 0.003$	$-0.137 \pm 0.002$	$(0.171 \pm 0.005) \cdot 10^{-1}$
$K$	$0.544 \pm 0.014$	$-0.193 \pm 0.010$	$(0.244 \pm 0.019) \cdot 10^{-1}$
$p$	$0.582 \pm 0.030$	$-0.168 \pm 0.022$	$(0.176 \pm 0.041) \cdot 10^{-1}$

**Table 6.1.:** Fit parameters of the polynomial fit of  $\sigma_{\text{jet}}$ .

are blue up to the value of the background function (see Fig. 6.2) evaluated at the bin center and red beyond that value.

We wish to stress that these figures do not give a completely faithful representation of the jet-peak yield measurement, as it does not explicitly show bins with a negative contribution to the jet-peak yield. Due to statistical fluctuations it is possible for the background function to exceed the total content in the bin, in which case no red bar is shown. These bins, however, do contribute to the peak-yield measurement, as these statistical fluctuations are “balanced out” by upward statistical fluctuations in other bins.



**Figure 6.4.:** Examples of peak yield measurements in pp (left panel) and Pb–Pb (right panel). The blue bars are background, while the red bars stacked on top constitute the jet-peak yield. The kinematic range is:  $1.8 < p_{T,\text{assoc}} < 2.2 \text{ GeV}/c$ .

## 6.4. Choice of Peak Region in $(\Delta\eta, \Delta\phi)$

In the previous section we proposed a procedure for measuring the jet-peak yields, without clearly motivating some seemingly ad-hoc choices for the size and definition of the jet peak and background regions. In this section we motivate these choices by discussing their advantages over some reasonable alternative approaches.

Choosing an optimal value for the scale factor  $\alpha$  is not trivial, as both small and large values for  $\alpha$  come with distinct advantages and disadvantages. A clear advantage of choosing  $\alpha$  small ( $\sim 1$ ) is that one can obtain a small statistical uncertainty on the jet-peak yield measurement, because the bins close to the center of the peak deviate the most from the background model. On the other hand, since we do not exactly know the optimal model for the jet peak (or at least we cannot reliably determine its parameters), we cannot simply take a small  $\alpha$  and extrapolate the total jet-peak yield. The obvious advantage of choosing  $\alpha$  large is that the measurement is guaranteed to count the yield in the entire jet-peak structure, while being independent of the specifics of the shape. This, however, comes at the price of a decreased significance on the jet-peak yield measurement, since the bins in the tail of the jet peak on average only deviate very little from the background.

Put differently: by choosing  $\alpha$  large, the jet-peak yield measurement entails subtracting two large numbers with a small difference compared to the size of those numbers. Additional practical limits on how large one can choose  $\alpha$  are given by the limited two-particle acceptance of the ALICE detector ( $\Delta\eta < 1.6$ ) and by the constraint that sufficient data at  $\Delta\phi \sim 0$  needs to be “reserved” to perform the measurement of the background parameters.

In the procedure described in Sect. 6.3 we used all bins in a circle around the origin to measure the jet peak, and all the remaining bins to measure the background. While it is clear that bins which are part of the jet-peak region should not be part of the background region, it is an obvious choice to use the compliment of the jet-peak region. Ideally one would like to determine the background as far away from the jet-peak structure as possible, so for small  $\alpha$  our approach does not seem to be a good option. Not only would this exclude a sizable part of the tail of the jet peak from the jet-peak region, it would in fact include it in the background region, resulting in a less reliable measurement of the background parameters<sup>4</sup>. Note that this also induces an additional bias on the jet-peak yield measurement, as the subtracted background is higher than the true background.

An alternative approach that could potentially solve this problem is to use two scale factors:  $\alpha_{\text{jet}}$  and  $\alpha_{\text{bgd}}$ , such that  $\alpha_{\text{jet}} < \alpha_{\text{bgd}}$ , in which case we would define the jet-peak region as all bins with a bin center in the disc  $R < \alpha_{\text{jet}}$ , ignore bins in the ring  $\alpha_{\text{jet}} \leq R < \alpha_{\text{bgd}}$ , and to define the background region as  $R > \alpha_{\text{bgd}}$ , where  $R \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2}$ . We finally decided not to use this approach, since the spatial extent of the jet-peak structure of protons at low- $p_{\text{T}}$  makes it impossible to find a value for  $\alpha_{\text{bgd}}$  that is larger than  $\alpha_{\text{jet}}$ , yet leaves sufficient data to perform a reliable measurement of the background parameters.

Another alternative would be to choose  $\alpha$  to be an increasing function of  $p_{\text{T}}$  and/or a decreasing function of particle mass. Since the width of the jet-peak structure decreases with higher  $p_{\text{T}}$  and smaller particle mass (see Fig. 6.3), this is in principle a viable option. As discussed previously, this approach would lead to a decreased statistical significance at higher  $p_{\text{T}}$ , however, there is an even more important reason to disfavor this option. While it would largely eliminate the bias at high  $p_{\text{T}}$ , it would not cure the problem at low  $p_{\text{T}}$ , and as a consequence it would result in a  $p_{\text{T}}$  (and mass) dependent bias in the measured particle spectra. In our opinion, the gains

---

<sup>4</sup>For example, the observed  $\langle dN_{\text{bgd}}/d\Delta\phi \rangle$  will be larger than its true value if part of the jet-peak tail is included in the background measurement.

of a non-constant  $\alpha$  are not worth the price of losing homogeneity of the analysis by introducing procedural changes as function of mass and  $p_T$ .

It is clear from the previous discussion that the choice of jet peak and background region should be the result of a practical compromise to suit the various extremes encountered in our analysis (large/small  $S/B$  ratio, data sample, peak width, etc.). The key features that we want to achieve in this analysis are: stability by avoiding the use of a numerical minimization procedure, and homogeneity by choosing a “one size fits all” strategy, i.e., not using procedural changes in the extremes of the analysis.

After comparing several values for  $\alpha$ , we found that choosing  $\alpha = 3.0$  leaves sufficient data outside of jet-peak region to estimate the anisotropy coefficients with a reasonable statistical uncertainty, while still counting true jet-peak yield with high efficiency. As discussed in Sect. 6.5, if the true jet-peak structure would be exactly a Gaussian, close to 99% of the jet-peak structure would fall within the  $3\sigma$  region, and the bias on the background measurement due to including part of the jet-peak tail in the background region should be minimal. In reality, the this number is smaller due to the jet peak being only approximately Gaussian, however, the bias on the jet-peak yield and background is expected to be independent of particle species, so that the measurement of particle ratios is unbiased.

## 6.5. A Toy-Model Study

In this section we discuss a toy-model study, designed to test the robustness of the procedure described in Sect. 6.3. The aim of this study is to apply our analysis procedure on a large number ( $\sim 10^3$ ) of random associated-yield distributions, generated with parameters that are known *a priori*, in order to assess whether the estimators in our procedure are biased. Formally, an estimator  $\hat{\theta}$  for a model parameter  $\theta$ , is said to be *biased* if  $E[\hat{\theta}] \neq \theta$ , where the operator  $E[\cdot]$  stands for the expectation value [96]. If an estimator  $\hat{\theta}$  is biased, then the bias  $b$  is defined as:  $b_{\hat{\theta}} = E[\hat{\theta}] - \theta$ . By increasing the size of the set of generated associated-yield distributions, one can in principle determine the expectation value of an estimator (and hence its bias) to arbitrary precision. Another interesting quantity is the mean squared error (MSE) of the estimator, which is defined as:  $\text{MSE} = E[(\hat{\theta} - \theta)^2]$ <sup>5</sup>. Using the MSE one obtains a measure of the relevance of the bias with respect to the statistical uncertainty on that parameter.

<sup>5</sup>The MSE is related to the variance:  $\text{MSE} = V[\hat{\theta}] - b^2$ .

For this study to be meaningful, it is important that the randomly generated distributions resemble the experimental distributions as closely as possible. We randomly generate an associated-yield distribution of the background particles using the “background part” of Eq. (6.2) as probability density function (PDF), i.e.:

$$\frac{d^2 N_{\text{bgd}}}{d\Delta\phi d\Delta\eta} \propto A(\Delta\eta) \left( 1 + 2 \sum_n V_{n\Delta} \cos n\Delta\phi \right), \quad (6.3)$$

and similarly, to generate an associated-yield distribution of the jet-peak particles using a two-dimensional Gaussian as PDF:

$$\frac{d^2 N_{\text{jet peak}}}{d\Delta\phi d\Delta\eta} \propto A(\Delta\eta) \exp \left\{ -\frac{(\Delta\phi^2 + \Delta\eta^2)}{2\sigma_{\text{jet peak}}^2} \right\}, \quad (6.4)$$

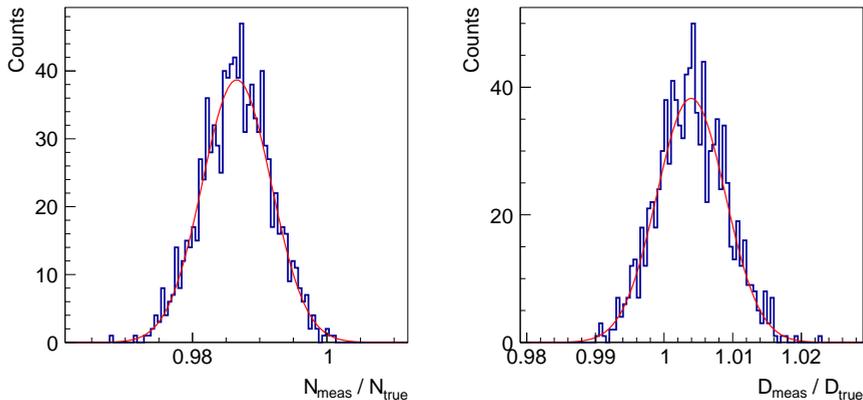
where  $A(\Delta\eta) = (\Delta\eta_{\text{max}} - |\Delta\eta|)/(\Delta\eta_{\text{max}})$ , and  $\Delta\eta_{\text{max}} = 1.6$ . Note that we did not use the more realistic double-Gaussian PDF for the jet-peak structure, as we cannot reliably determine its true parameters. The function  $A(\Delta\eta)$  is included to model the pair-acceptance of the detector. Its main function is to obtain an associated-yield distribution with an increasing statistical uncertainty towards larger  $\Delta\eta$ <sup>6</sup>.

The first step in generating the associated distributions is to scale the PDFs in Eqs. (6.3) and (6.4) such that they represent the expected number of counts per bin. After that, the associated distributions can be generated by drawing for each bin a random Poisson-distributed variable, with an expectation value given by the scaled PDFs. Then, the associated yield distributions have to be corrected for “mixed events”, which is done by dividing it by the function  $A(\Delta\eta)$ . Finally, we normalize the generated histograms with the number of “trigger particles”, apply the analysis described in Sect. 6.3 on the sum of the jet peak and background distributions, and compute the expectation value for each estimator.

Since we wish to get an estimate of the bias under realistic circumstances, the study is set up as follows: for every measured associated yield distribution in our analysis we generate  $10^3$  random associated yield distributions with the

---

<sup>6</sup>The model assumes a detector with perfect single-particle efficiency, however, this approximation should not have a noticeable influence on our results.



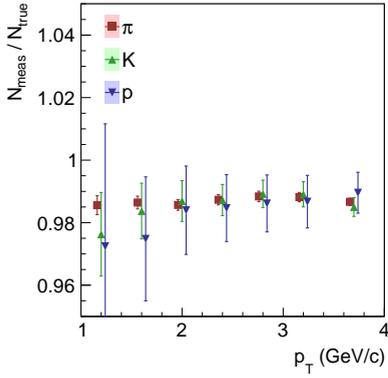
**Figure 6.5.:** Gaussian fits to data generated with the toy-model and analyzed with the method described in Sect. 6.3. The generated associated yield distributions resemble the associated pion distribution from p-Pb collisions, at  $2.2 < p_T < 2.6$  GeV/ $c$ . Left panel: measured jet-peak yield over true jet-peak yield, right panel: measured background density over true background density.

same parameters (flow coefficients, jet-peak width, etc.). From this sample, we calculate the bias and MSE for every measured observable<sup>7</sup>.

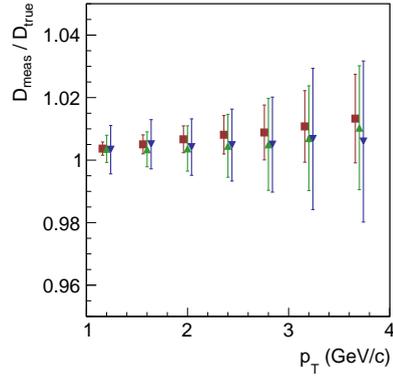
As an example, in Fig. 6.5 we show the ratio of the measured jet-peak yield and the true jet-peak yield, as well as the ratio between the measured background density and the true background density. Notice that, as expected, the choice of  $\alpha = 3$  induces a small bias in the background density observable, which occurs because the background does include some of the tail of the jet peak. Also, the jet-peak yield observable is biased; there is a bias of about  $\sim 1\%$  due to the jet region containing  $3\sigma$  of the jet peak, and there is a small additional bias due to the subtraction of an overestimated background density.

In Fig. 6.6 the results of the study are shown as a function of  $p_T$ . Note that while the bias on the jet-peak yield measurements is nearly equal in magnitude for the different collision systems, for pions it is significantly larger than the statistical uncertainty, while for protons the statistical uncertainty is larger than the bias. Furthermore, in jet-dominated events (pp) the

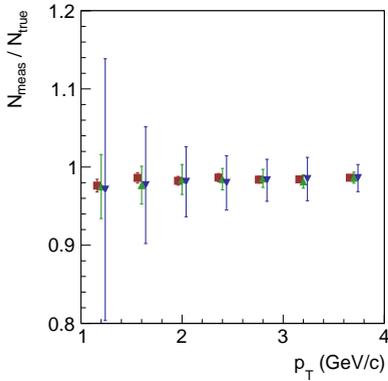
<sup>7</sup>For the MSE to be meaningful, it is important that the statistical uncertainties on the generated points are comparable to the data points.



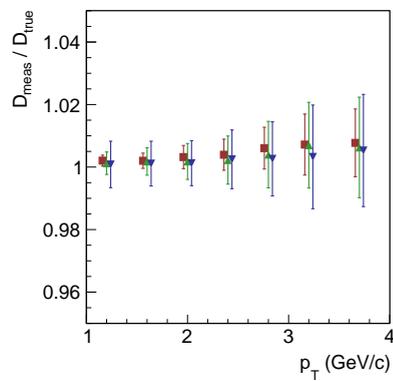
(a) pp Jet-Peak Yield



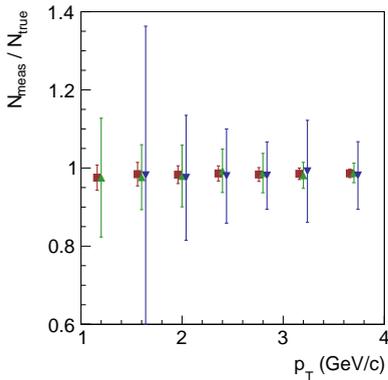
(b) pp Background Density



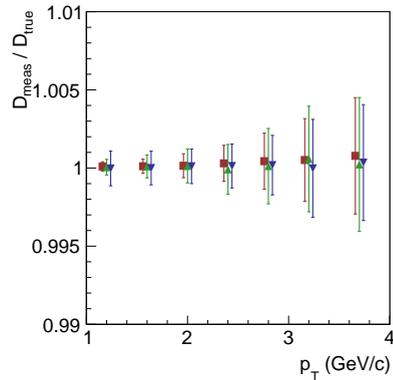
(c) p-Pb Jet-Peak Yield



(d) p-Pb Background Density



(e) Pb-Pb Jet-Peak Yield



(f) Pb-Pb Background Density

**Figure 6.6.:** Bias on the jet-peak yield and background density, determined by applying our measurement technique (Sect. 6.3) on data generated by the toy model and comparing the results to the true values. The error bars represent the MSE.

ratio of the bias to the statistical uncertainty is generally larger than in background-dominated events (central Pb–Pb). As expected, the bias is to a good approximation the same for pions, kaons, and protons, except in the first two  $p_T$  bins for protons and the first  $p_T$  bin for kaons, where the bias is  $\sim 1\%$  larger. This is due to the fact that the jet region has a radius of one, since for aforementioned bins have  $\alpha\sigma_{\text{jet}} > 1$  (see step 3 in Sect. 6.3). There is also a bias of about  $\sim 1\%$  on the background-density estimator in pp events. For the other systems the bias becomes negligible due to the fact that the events are background-dominated.

The results of the study show that while there is a bias on the estimators for the jet and background yield, they do not depend significantly on the particle species, hence the particle ratios that are presented in Sect. 6.9 are not expected to be biased. The jet-peak and background density spectra that are also presented in Sect. 6.9 are expected to be somewhat biased, however, we also decide not to correct for this. The main reason for this is that it is not possible to exactly estimate the size of the bias, as we do not have sufficient knowledge at present about the exact shape of the jet peak. A justified criticism that one can have is that this decision makes the results presented in Sect. 6.9 dependent on the details of the method that we used to perform the measurement. This is true, however, we advocate that it is worse to impose a correction for the bias based on a model we are not certain about. Crucially, the method that we use to obtain the results is well-defined, which makes the results reproducible.

## 6.6. Correction for Secondary Particles

In Sect. 4.6, we presented a method to determine the fraction of particles coming from weak decays and material interactions, using fits of the DCA distribution. Unfortunately, this method cannot be used to determine these fractions in the jet peak and the background separately, as it is not possible to create DCA distributions for the jet peak. While there is no reason to assume that the fraction of secondaries from material interactions is different in the jet peak and in the background, it is not *a priori* clear that the same holds for the secondaries from weak decays. In fact, if the difference in the hadrochemistry of the jet peak and the background extends to the strange sector<sup>8</sup>, then the relative importance between different weak-decay channels is altered as well.

---

<sup>8</sup>For example, a smaller  $\Lambda/K_S^0$  ratio in the jet peak.

The final results of our analysis are obtained from tracks with a tight cut on the DCA (see Appendix A), which removes most of the secondary particles. To study whether the remaining fraction of secondary particles is significantly different in the jet peak and the background, we also measured the jet-peak and background spectra using a loose DCA cut, removing almost no secondary particles. Notice that the dataset with the tight DCA cut is a subset of the dataset with the loose DCA cut. Similar to Sect. 4.6, where we used a MC simulation to generate template DCA distributions, a MC simulation can be used to determine the fraction of primary particles and weak-decay particles that are selected by the loose DCA cut, but also the tight one. Typical survival rates of primary particles are as high as 98 – 99%, while for secondary particles from weak decays it is approximately 15 – 25%, both increasing towards higher  $p_T$ . Under the assumption that the fraction of secondaries from material interactions is known, the measurement with the loose and tight DCA cuts, supplemented with these survival fractions, is sufficient to determine the fraction of primaries and secondaries in both jet peak and background.

To show this explicitly, consider a yield measurement  $N$  using the loose DCA cut<sup>9</sup>, and the same yield measurement  $N'$  using a tight DCA cut. Reusing the notation from Sect. 4.6, the sum of the fractions of primary particles, secondaries from weak decays, and secondaries from material interaction add up to one for both loose and tight DCA cuts, i.e.,  $f_P + f_W + f_M = f'_P + f'_W + f'_M = 1$ , where the prime again indicates the tight DCA cut. Denoting the MC survival fractions of primary particles and weak decay products by  $S_{P,W}$ , one can write:  $S_{P(W)}f_{P(W)}N = f'_{P(W)}N'$ . Assuming that the fractions of secondaries from material interactions  $f_M$  and  $f'_M$  are known, the fraction of primary particles after the tight DCA cut is given by:

$$f'_P = \frac{S_P(1 - f'_M) - S_P S_W \frac{N}{N'}(1 - f_M)}{S_P - S_W}. \quad (6.5)$$

Notice that the primary fraction in Eq. (6.5) depends linearly on the ratio of the measurements with loose and tight DCA cuts. A possible difference in the primary fraction in jet peak and background can therefore be written as

---

<sup>9</sup>The argument does not depend on particle species,  $p_T$  range, or whether it is a jet-peak or background yield. If the yield is associated with a trigger particle, as is the case here, then it is assumed to be normalized by the number of trigger particles.

a difference in these ratios, i.e.:

$$f'_{\text{P,jet}} - f'_{\text{P,bgd}} = \frac{S_{\text{P}}S_{\text{W}}}{S_{\text{P}} - S_{\text{W}}}(1 - f_{\text{M}}) \left[ \frac{N_{\text{jet}}}{N'_{\text{jet}}} - \frac{N_{\text{bgd}}}{N'_{\text{bgd}}} \right]. \quad (6.6)$$

For pions in pp and p-Pb collisions we find a statistically significant difference in the ratios of  $\sim 2\%$ , i.e., in the jet peak, the tight DCA cut removes a larger fraction of particles than in the background. For protons we observe differences of a similar magnitude, but with much less significance. In Pb-Pb collisions, we see no significant difference for any particle species, however, the statistical uncertainty on the jet-peak measurements are very large.

Using the typical values for the survival rates and the material fraction of protons ( $f_{\text{M}} \lesssim 2 - 3\%$ )<sup>10</sup>, we find that the pre-factor in Eq. (6.5) is approximately 0.2 – 0.3. This means that the difference in the primary fraction when using the tight DCA cut is  $\lesssim 1\%$ , and can safely be neglected.

## 6.7. Efficiency Correction

As described in Eq. (5.5), the mixed-event distribution can be used to correct for two-particle acceptance and two-particle efficiency effects. In this section we discuss the remaining overall efficiency factor, the average single-particle efficiency of the associated particles, i.e.,  $\langle \epsilon_{\text{assoc}} \rangle$ .

As is standard in analyses using particle identification, the single-particle efficiency is written as the product of three types of efficiencies, which are listed below. In the description of the different types of efficiencies we often refer to the track cuts, which are detailed in Appendix A.

- *Tracking Efficiency* - The probability that a particle's track is successfully reconstructed, and that it passes the standard track cuts. The tracking efficiency is denoted by:  $\epsilon_{\text{tracking}}$ .
- *PID Efficiency* - The probability that a reconstructed track that passes the standard track cuts also has a usable signal from both TOF and TPC detectors. This means, for example, that the track needs to have a sufficiently large number of specific energy-loss measurements along the track. Furthermore, the track needs to be matched with a hit in

<sup>10</sup>For protons the material fraction is much larger than for the other particle species.

the TOF detector, essentially this means that a TOF hit needs to be found which is sufficiently close to the reconstructed track. The PID efficiency is denoted by  $\epsilon_{\text{PID}}$ .

- *TOF Good-Matching Efficiency* - The probability that a reconstructed track with usable signals from TOF and TPC is matched with the correct TOF hit. While the PID efficiency includes the probability for finding a match with a TOF hit, this match is not guaranteed to be correct. The TOF good-matching efficiency is related to the number of mismatches and is most important for events with a high track density, i.e., for central Pb–Pb events<sup>11</sup>. The TOF matching efficiency is denoted by:  $\epsilon_{\text{TOFGoodMatch}}$ .

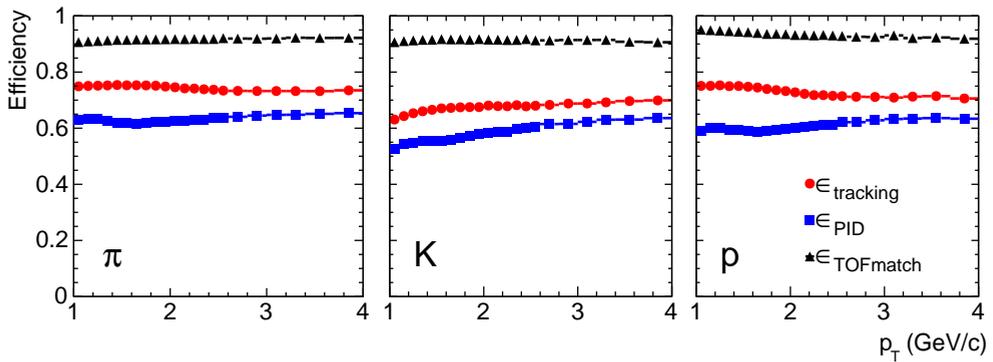
These different efficiencies can be calculated by means of a Monte Carlo simulation, and a virtual model of the ALICE detector. First, an event is generated using an event generator such as PYTHIA for pp events [97], DPMJet for p-Pb events [98], or Hijing for Pb–Pb events [99]. The generated particles are subsequently propagated through the virtual ALICE detector, and the interactions of the particles with the detector material are modeled by dedicated simulation software, which in our case is GEANT3 [93]. This way a virtual “measurement” of the generated event is created, and the track reconstruction algorithms can be applied. The tracking efficiency is calculated by taking the ratio of the number of reconstructed tracks and the number of generated particles. Similarly, the PID efficiency and TOF matching efficiency can be determined from the reconstructed tracks and the criteria listed above. The efficiencies are calculated for primary particles only, as secondary particles are removed from our final results.

The reliability of the efficiency factors depends on how close the resemblance is between the response of the true detector and the detector model. One known issue with the MC runs used for determining the efficiencies for pp and Pb–Pb (see Appendix B) is a problem with GEANT3, which overestimates the annihilation cross-section of anti-protons, and to a much lesser extent that of anti-kaons, with the detector material [100]. For this reason the produced efficiencies are too small, and a correction factor needs to be applied. This correction factor is determined using the FLUKA simulation package [101].

In Fig. 6.7 we show the different efficiency factors for central Pb–Pb collisions as a function of  $p_{\text{T}}$ .

---

<sup>11</sup>The TOF matching efficiency is only applied in the Pb–Pb analysis.



**Figure 6.7.:** Tracking, PID, and TOF matching efficiency for central Pb–Pb collisions.

## 6.8. Systematic Uncertainties

Several sources of systematic uncertainty on the jet-peak and bulk density spectra are due to imperfections in the model of the ALICE detector, used to determine the efficiencies. These are common to all (PID) analyses in ALICE, which is why we use the same uncertainties, determined during previous analyses. The relevant sources are: tracking efficiency, TOF matching efficiency,<sup>12</sup> and secondary correction. The relatively large uncertainty on the TOF matching efficiency is due to the presence of the Transition Radiation Detector (TRD), which constitutes a significant amount of material in between the TPC and the TOF detector. As described in Ref. [94], the TRD was not fully installed in 2010, hence the associated systematic uncertainty can be estimated by explicitly comparing identified spectra in regions with and without TRD.

As discussed in Sects. 4.4 and 5.3, the most important sources of systematic uncertainty coming from the TOF/TPC fit method are due to an imperfect model of the detector response. Three different checks were performed to assess the extent of the uncertainty: the dependence of the fitted yield on the choice of fit region, the comparison with TOF SA fits (in Tables 6.2 and 6.3 these two sources are collectively referred to as “Fit Function”), and the effect of fixing the tail-parameter  $\lambda_{K|\pi}$ . Finally, as discussed in Sect. 6.5, the method for determining the jet-peak yield and background density is

<sup>12</sup>Note that in this analysis the TOF matching efficiency is included in the PID efficiency. It is not the TOF good-matching efficiency.

$p_T$ range (GeV/c)	$\pi^\pm$		$K^\pm$		$p$ and $\bar{p}$	
	1	4	1	4	1	4
Tracking Eff. [95, 102]	3%	3%	3%	3%	3%	3%
ITS-TPC Matching Eff. [102]	2%	2%	2%	2%	2%	2%
TOF Matching Eff. [94, 102]	3%	3%	6%	6%	4%	4%
Secondary Correction [94]	1%	1%	negl.	negl.	2%	1%
Fit Function (pp) <sup>13</sup>	1%	2%	2%	5% <sup>14</sup>	2%	5%
Fit Function (p-Pb) <sup>13</sup>	1%	2%	2%	5% <sup>14</sup>	2%	2%
Fit Function (Pb-Pb) <sup>13</sup>	1%	2%	1%	2%	1%	1%
Fixing Tail Parameter (Ch. 4)	0%	0%	0%	3%	0%	2%
Choice of $\alpha$	N/A	N/A	N/A	N/A	N/A	N/A
Total (pp)	5%	5%	7.5%	9%	6%	7.5%
Total (p-Pb)	5%	5%	7.5%	9%	6%	6%
Total (Pb-Pb)	5%	5%	7%	8%	6%	6%

**Table 6.2.:** Main sources of systematic uncertainty on jet-peak and background density spectra.

expected to be slightly biased. As we do not have sufficient information about the shape of the jet peak to exactly estimate the size of the bias, we have decided to present the results without correcting for this bias. By doing this we essentially adopted a working definition of the “jet peak” and “background density” to be that what is measured by our method, and for this reason no systematic uncertainty is assigned. The main sources of systematic uncertainty in our measurement are collected in Table 6.2.

The systematic uncertainties that are discussed so far apply to the jet-peak yield and the background density. For the ratio of the yield of two different particle species, some of these uncertainties can reasonably be assumed to be correlated (tracking efficiency, matching efficiencies), and therefore the overall effect on the ratio is reduced. For this reason, the systematic uncertainty on the tracking efficiency and both matching efficiencies is taken to be the absolute difference of the two particle species involved, with a minimum of

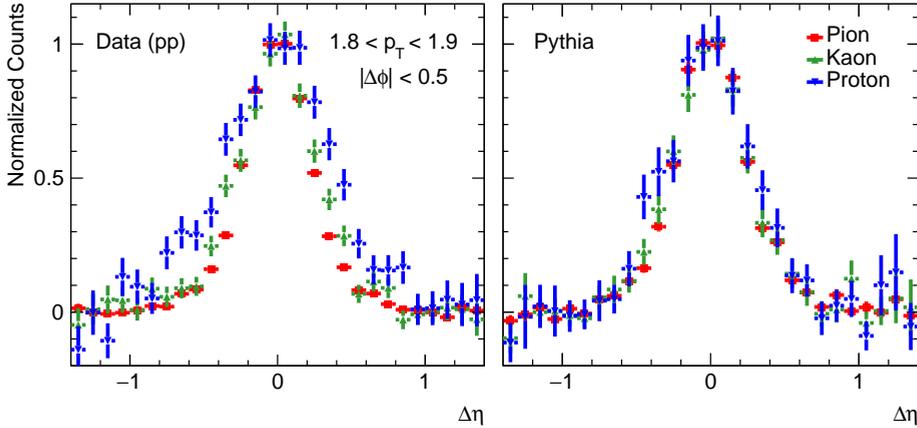
<sup>13</sup>Based on Sects. 4.4.2 and 4.4.3.

<sup>14</sup>Maximum value attained at  $p_T = 2.0$  GeV/c.

$p_T$ range (GeV/c)	$\frac{K^+ + K^-}{\pi^+ + \pi^-}$		$\frac{p + \bar{p}}{\pi^+ + \pi^-}$	
	1	4	1	4
Tracking Eff.	1%	1%	1%	1%
ITS-TPC Matching Eff.	1%	1%	1%	1%
TOF Matching Eff.	3%	3%	1%	1%
Secondary Correction	1%	1%	1%	1%
Fit Function (pp) <sup>13</sup>	2%	7% <sup>14</sup>	2%	5%
Fit Function (p-Pb) <sup>13</sup>	2%	7% <sup>14</sup>	1.5%	3%
Fit Function (Pb-Pb) <sup>13</sup>	2%	3%	1.5%	3%
Fixing Tail Parameter	0%	3%	0%	2%
Choice of $\alpha$	N/A	N/A	N/A	N/A
Total (pp)	4%	8.5%	3.0%	5.5%
Total (p-Pb)	4%	8.5%	2.5%	4.0%
Total (Pb-Pb)	4%	5.5%	2.5%	4.0%

**Table 6.3.:** Main sources of systematic uncertainty on particle ratios.

1% per effect. This minimum of 1% is introduced because while the efficiency correction and matching efficiencies are similar from species to species, they are not exactly the same (see Sect. 6.7). For the systematic uncertainty due to the secondary correction, we use the uncertainty from the pions. In case of the  $K/\pi$  ratio, the reason is that the amount of secondary kaons is negligible and in case of the  $p/\pi$  ratio, the reason is that the most important process that produces secondary protons is lambda decay, which produces pions in equal amounts (see Sect. 4.6). The uncertainty due to the choice of fit region is only partially correlated (see Sect. 4.4.2), while the uncertainty from the TOF SA comparison is fully anti-correlated for pions and kaons (see Sect. 4.4.3). The appropriate values for the ratios are evaluated by taking the ratio of the histograms in the different panels of Figs. 4.5 and 4.6. Finally, since fixing the TOF tail parameter  $\lambda_{K|i}$  has a negligible effect on the pion measurement, the systematic uncertainty on the  $K/\pi$  and  $p/\pi$  ratio are simply the uncertainties on the kaon and proton spectra respectively. The uncertainties that are used for the measurement of the identified ratios are summarized in Table 6.3.



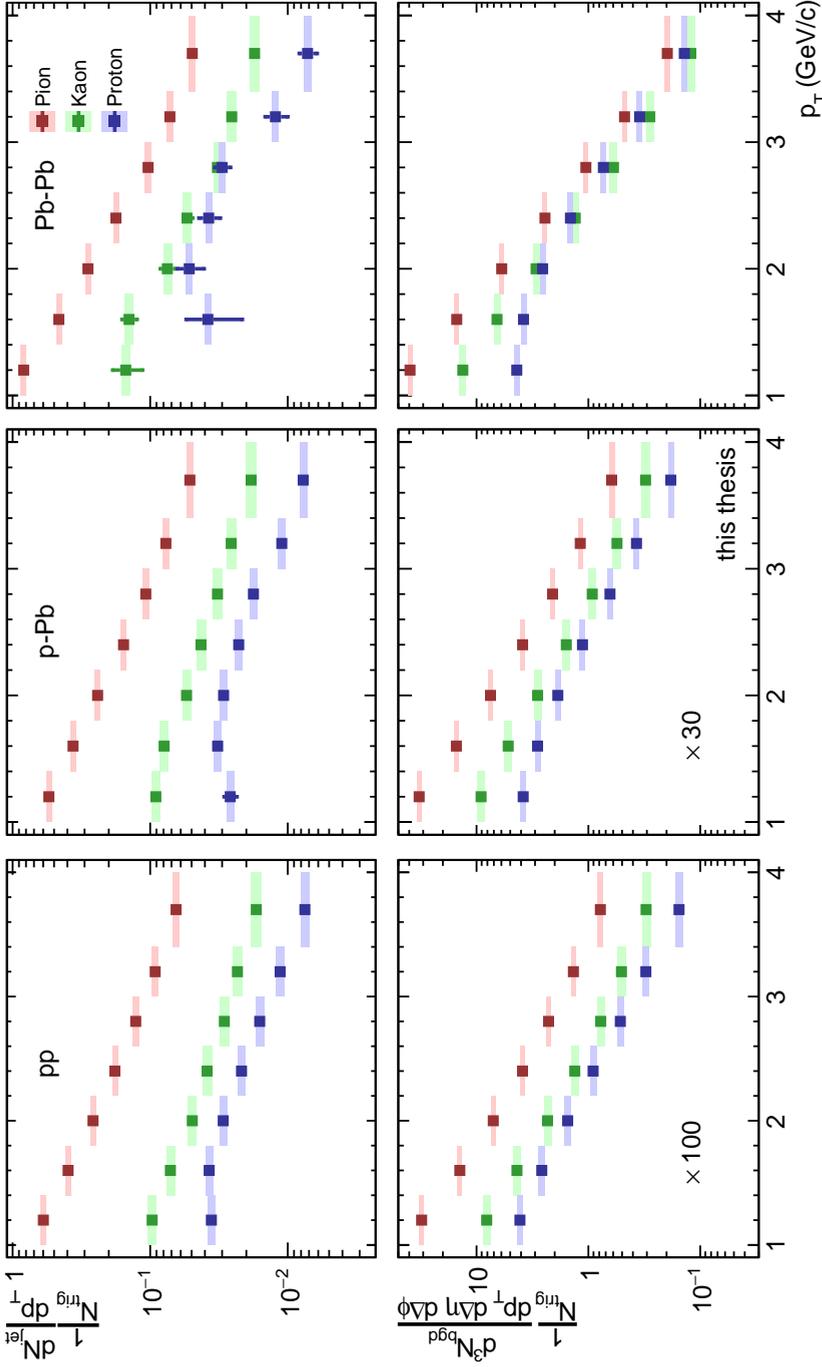
**Figure 6.8.:** A comparison of the normalized jet-peak shapes as a function of  $\Delta\eta$  as measured in pp collisions (left panel) and as calculated by the PYTHIA model (right panel). The background level is set to 0 by subtracting the average of the yield in the bins  $|\Delta\eta| > 1.0$  and after that the histogram is scaled by the inverse of the average yield in the two central bins.

## 6.9. Results

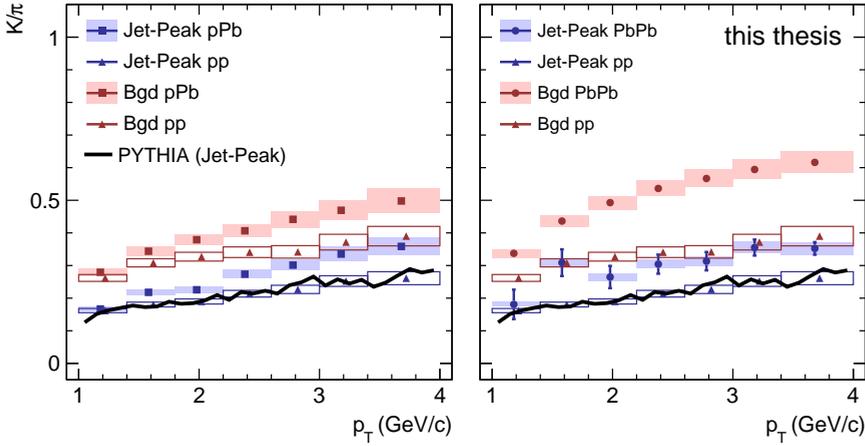
The first interesting observation that we made was that the width of the jet peak (both in  $\Delta\eta$  and  $\Delta\phi$ ) turns out to increase with larger mass and smaller  $p_T$  (see Fig. 6.3). In Fig. 6.8 we show a qualitative comparison between the (normalized) peak shapes in a narrow  $p_T$  interval in pp collisions and in PYTHIA (Perugia 11 tune, using the CTEQ5L parton distribution functions). The left panel of this figure clearly shows that peak width increases with the mass of the particle, i.e.,  $\sigma_{\text{jet},\pi} < \sigma_{\text{jet},K} < \sigma_{\text{jet},p}$ , while the right panel shows that this feature does not seem to be reproduced by PYTHIA.

A possible explanation of the widening of the proton peak would be that the majority of protons are produced by gluon jets [103, 104], and that gluon jets are generally wider than quark jets [105]. However, if this were the explanation, then we would also see a similar widening of the proton peak in the PYTHIA results. We therefore conclude that the mass-dependence is not a natural consequence of the fragmentation dynamics in PYTHIA.

We also noted that the jet-peak shape appeared to be non-Gaussian (see Fig. 6.1). One possibility is that the non-Gaussianity is caused by the fact



**Figure 6.9.:** Identified jet-peak and average background yield per trigger in pp, p-Pb, and Pb-Pb collisions. The error bars represent the statistical uncertainty on the measurement, while the lighter colored solid boxes represent the systematic uncertainty. The proton jet-peak yield in the lowest  $p_T$  bin of Pb-Pb could not be observed, as the signal-to-noise ratio for this measurement is too large, given the current dataset.

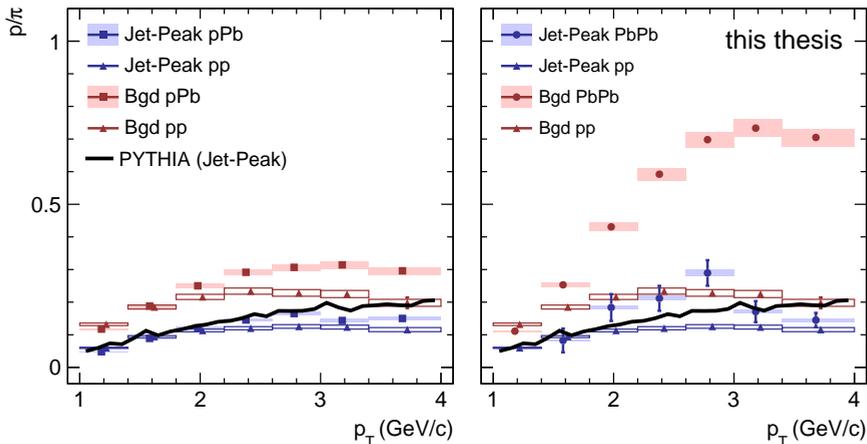


**Figure 6.10.:** The  $K/\pi$  ratio in jet peak (blue) and background (red), in pp, p-Pb, and Pb-Pb collisions. The error bars represent the statistical uncertainty, while the boxes represent the systematic uncertainty. On the left panel we make a comparison between pp and p-Pb collisions, while on the right panel we compare pp and Pb-Pb collisions. Both panels include a PYTHIA reference at  $\sqrt{s} = 7$  TeV.

that we integrate over a reasonably large  $p_T$  interval, and that the width of the jet peak is  $p_T$ -dependent. To check if this could cause an effect of the same magnitude as observed in the data we performed a simulation of a toy model, using a reasonable parameterization of the  $p_T$  spectrum, assuming a Gaussian shape of the peak width, and taking  $\sigma(p_T)$  from a second order polynomial with the parameters as denoted in Table 6.1. This way we simulated a jet peak integrated over a  $p_T$  range comparable to the bin sizes used in our analysis. The non-Gaussianity of the simulated peaks was in general negligible, and we conclude that there must be a different underlying cause.

The identified jet-peak and average background spectra for pp, p-Pb, and Pb-Pb collisions are shown in Fig. 6.9. Notice that while the jet-peak spectra for each collision system are of the same order of magnitude, the background density spectra differ by one or two orders of magnitude.

To study the hadrochemical composition of the jet peak and the background, we show the  $K/\pi$  ratio in Fig. 6.10 and the  $p/\pi$  ratio in Fig. 6.11, together with a PYTHIA reference at the same center of mass energy as the pp collisions ( $\sqrt{s} = 7$  TeV). To make sure that the PYTHIA results can



**Figure 6.11.:** The  $p/\pi$  ratio in jet peak (blue) and background (red), in pp, p-Pb, and Pb-Pb collisions. The error bars represent the statistical uncertainty, while the boxes represent the systematic uncertainty. On the left panel we make a comparison between pp and p-Pb collisions, while on the right panel we compare pp and Pb-Pb collisions. Both panels include a PYTHIA reference at  $\sqrt{s} = 7$  TeV.

be compared with the measurements, we have performed exactly the same analysis on the generated dataset as on the experimental dataset.

There are several interesting observations one can make from Figs. 6.10 and 6.11. One observation is that within each collision system, the  $K/\pi$  and  $p/\pi$  ratio are significantly larger in the background than in the jet peak, most prominently so for the  $p/\pi$  ratio in Pb-Pb collisions. In central Pb-Pb collisions, the large  $p/\pi$  ratio of the background at intermediate  $p_T$  is also seen in the inclusive spectra, and is known as the baryon anomaly (see Sect. 2.7). For  $p_T \lesssim 2.0$  GeV/c, the baryon anomaly is explained by collective flow effects, while for higher  $p_T$  it can be understood in terms of a recombination model, where the dominant mechanism of hadronization is the recombination of (mostly thermal) partons. The fact that the  $p/\pi$  ratio in the jet peak does not show a similar dramatic increase suggests that the mechanism of hadron production is not the same as in the background, and is likely for a large part determined by the dynamics of jet fragmentation. In pp events on the other hand, the production of all hadrons is generally modeled by independent fragmentation of color strings between the partons in the parton shower and the partons in the beam remnants [97]. From this

it is not immediately obvious why there is a difference in the particle ratios in jet peak and bulk in pp collisions.

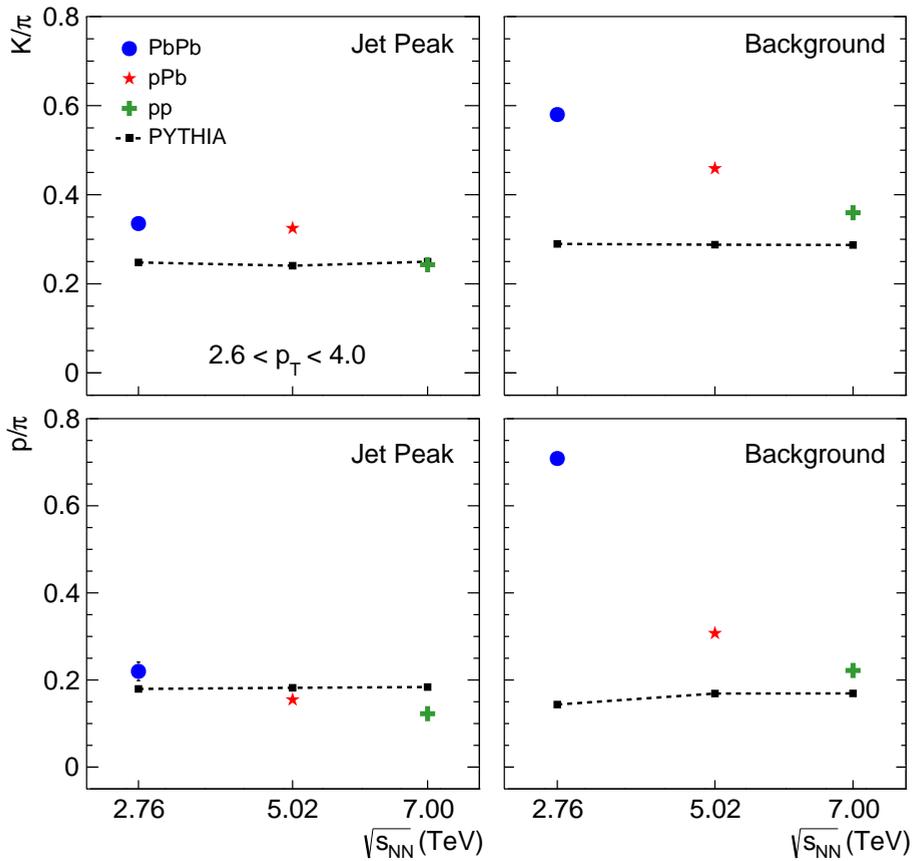
Another interesting comparison can be made between the particle ratios in the jet peak among the three collision systems. It has been suggested that the particle ratios in the different systems may be similar due to the so-called *surface-bias*, i.e., the fact that the high- $p_T$  trigger particles are more likely to originate from a parton shower that was initiated from a hard partonic collision that took place close to the surface of the hot QCD medium. This would imply that the fragmenting partons may have had very little interaction with the medium, suggesting that the resulting jet should be similar in composition to a jet created in a pp collision. We observe, however, that the  $p/\pi$  ratio in the jet peak in Pb–Pb collisions is larger than the same ratio in pp and p-Pb collisions, which suggests that there is some influence from the hot QCD medium on the hadronization process. Similarly, we observe that the  $K/\pi$  ratio in the jet peak is, both in Pb–Pb and p-Pb collisions, significantly larger than in pp collisions. It is somewhat surprising that we also observe a difference between pp and p-Pb collisions. This observation adds an additional piece of information to the discussion that pp and p-Pb events are not as similar as initially expected.

Two possible mechanisms through which the QCD medium can affect the hadronization of the jet were introduced in Chapter 2, namely: recombination of thermal and shower partons (Sect. 2.7) and medium-induced modification of the splitting functions (Sect. 2.8). In order to investigate the origin of the presented observations, a more thorough study using these models is warranted.

It should be noted that the center-of-mass energies of the three systems under study are not the same<sup>15</sup>. For this reason it cannot be ruled out that the differences between the three systems are due to the difference in center-of-mass energy. In order to investigate this, we ran PYTHIA simulations at  $\sqrt{s} = 2.76, 5.02$  and 7 TeV, and in Fig. 6.12 we compare the particle ratios in jet peak and background from the PYTHIA simulation to the particle ratios measured in pp, p-Pb, and Pb–Pb collisions. From this figure it is clear that the particle ratios from PYTHIA, both in the jet peak and the background, show very little dependence on  $\sqrt{s_{NN}}$ , while a clear trend as a function of energy or system size is seen in the data. For this reason, we conclude that

---

<sup>15</sup>ALICE has also taken pp data at  $\sqrt{s} = 2.76$  TeV, however, this dataset is too small to perform the analysis presented in this work.



**Figure 6.12.:** The  $K/\pi$  and  $p/\pi$  ratios in jet peak and background in one single  $p_{T,assoc}$  bin, as a function of  $\sqrt{s_{NN}}$ . We compare results from pp, p-Pb, and Pb-Pb collisions in one figure, and PYTHIA results are included in the same figure as reference. The size of the statistical uncertainty is smaller than the marker size for all but the  $p/\pi$  ratio in the jet peak in Pb-Pb collisions.

the observed differences in the particle ratios between the collision systems are most likely due to the difference in system size.

## 6.10. Summary

In this chapter we discussed several methods that can be employed to study the near-side structure of the (identified) yield associated with a high- $p_T$  trigger particle. We presented results on several observables: the width of the jet peak, the jet-peak yield and the average particle density of the underlying event.

Given the size of the dataset that is presently available, numerically fitting the associate yield distribution in  $(\Delta\eta, \Delta\phi)$  with an appropriate model does not yield sufficiently stable results for all particle species and  $p_T$  bins to be generally used, unless an oversimplified model of the jet-peak shape is used. To quantify the background density and the number of particles in the jet peak, the preferred method is a combination of a discrete Fourier transform, and a simple counting method. This method is stable by construction, and is shown to have a minimal bias.

One interesting observation that could be made from fitting the jet peaks, projected either onto the  $\Delta\eta$  or  $\Delta\phi$  axis and using a simplified model for the jet peak, is that the width of the jet peak is larger for heavier particles and smaller at higher  $p_T$ . Furthermore, the width is independent of the collision system within the present uncertainties.

Finally, the  $p/\pi$  and  $K/\pi$  ratios in the jet peak and background for pp, p-Pb, and Pb-Pb collisions are presented as a function of  $p_T$ . In all collision systems we found that the  $p/\pi$  and  $K/\pi$  ratios are significantly smaller in the jet peak compared to the underlying event. We furthermore found that at intermediate  $p_T$  ( $2.6 < p_T < 4.0$  GeV/ $c$ ), both  $p/\pi$  and  $K/\pi$  ratios in the background are smaller in pp than in p-Pb collisions and also smaller in p-Pb than in Pb-Pb collisions. The particle ratios in the jet peak show a similar trend, except the  $K/\pi$  ratio, which is similar in p-Pb and Pb-Pb.

In addition, PYTHIA calculations show that the center of mass energy dependence of the  $p/\pi$  and  $K/\pi$  ratio in jet peak and background are negligible, which indicates that the observed effect is related to the difference in system size.



## Chapter 7.

# Discussion and Outlook

In this work we have developed a new and powerful particle identification strategy, based on fitting the combined signal of two independent detector systems, the TOF ( $\Delta t$ ) and the specific energy loss ( $dE/dx$ ) measured by the TPC. While there are several complications to this method, most notably the large number of fit parameters, at intermediate  $p_T$  ( $1.0 < p_T < 4.0$  GeV/ $c$ ) it has a much better separation power than methods based on a single detector signal. The  $p_T$  range of the TOF-TPC method could be extended to higher  $p_T$ , however, this most likely requires making some assumptions on the parameters of the fit function. One way to do this would be to parameterize some parameters as a function of  $p_T$ .

As the shape of the TOF and TPC detector response depends on different kinematic parameters, including  $\eta$ , the response as a function of  $\Delta\eta$  is “smeared out”. In this work we showed that the response as a function of  $\Delta\eta$  can effectively be modeled by fit-templates, constructed by taking a linear combination of the model functions fitted to the inclusive data for different values of  $\eta$ . Fitting (TOF, TPC) distributions with templates reduces the number of fit parameters to the number of particle species in the analysis, plus one for the TOF mismatched particles. The decreased dimensionality of the parameter space is in particular useful for fits of the associated yield distribution with a high- $p_T$  trigger, as the high- $p_T$  trigger requirement greatly reduces the size of the dataset. This high- $p_T$  trigger cut is especially stringent in pp collisions. In this work we have used the template fit method exclusively to perform fits in the (TOF, TPC) plane, however, the same strategy can in principle be applied to any type of PID method that uses a fit procedure.

Using the TOF-TPC template fit method we performed a measurement of the identified ( $\pi$ ,  $K$ ,  $p$ ) yield associated with an unidentified high- $p_T$  trigger

particle for pp, central p-Pb, and central Pb-Pb collisions. Using Gaussian fits of the near-side jet-peak structure, we observed that the width of the jet peak is larger for heavier particles and smaller towards larger  $p_T$ . Using a discrete Fourier transform we determined the background density and the flow coefficients as a function of  $p_T$  particle species. The identified jet-peak yield was measured by first subtracting the background from the data, and subsequently counting the remainder in a circle of three times the expected width of the jet-peak structure, which was determined by the Gaussian fits. Using the resulting identified spectra, we studied the hadrochemistry of the jet peak and background.

In all three systems, we observe that the  $K/\pi$  and  $p/\pi$  ratios are much larger in the background than in the jet peak. The difference is largest in Pb-Pb collisions. These observations indicate that the mechanism of hadronization is different in the jet peak and the background, however, further theoretical studies of this effect are warranted. The composition of the jet peaks in the three different systems also turns out to be somewhat different, however, the difference is not as large as in the background. Comparisons with PYTHIA results suggest that the observed differences between the collision systems are indeed due to the difference in system size, and not due to a difference in the center-of-mass energy.

The difference between the  $p/\pi$  ratio in the jet peak and underlying event in central Pb-Pb collisions was presented at the 5th Hard Probes conference in 2012 [106], and the current work extends on this study by improving the particle identification method, including kaons in the analysis, and by performing the same analysis on p-Pb, and pp collisions. Furthermore, it provides a preliminary study of the near-side jet-peak shape for identified particles, a topic that was previously only studied for unidentified associated particles [107].

The results presented in this work open up a number of possibilities for further investigation. There are several extensions that are possible with the currently available datasets. For example, one can measure the  $\Lambda/K_0$  ratio in jet peak and background; this study is currently underway.

As mentioned in Sect. 6.9, it is possible that there is a surface-bias, i.e., jets containing a high- $p_T$  trigger particle are more likely to originate from a parton-parton collision that took place near the surface of the hot QCD medium. This could explain why the difference in the particle ratios from system to system is relatively small, since the parton at the origin of the jet containing the high- $p_T$  trigger particle may have had little interaction with the hot QCD medium. One way to study this is by measuring the particle

ratios in the away-side jet peak as a function of collision centrality. If the high- $p_T$  trigger requirement indeed causes a bias towards probing near-side jets originating from partons with a very short path-length through the medium, then conversely, the parton at the origin of the away-side jet must have traveled a much longer distance through the medium. This distance should, however, decrease as a function of centrality, so if there is indeed a surface-bias, then one would expect to see a centrality dependence of the hadrochemistry in the jet peak.

For pp collisions, measuring the particle ratios in the away-side jet peak is a relatively straightforward extension of the current analysis, and also for p-Pb collisions one can likely separate the away-side ridge into a part due to collective flow and a part due to the away-side jet peak. For Pb-Pb collisions, the situation is more complicated because collective flow dominates the away-side structure. One possibility is to start by analyzing peripheral Pb-Pb collisions, and work towards more central collisions, up to the point where the analysis of the away-side peak is no longer feasible.

We used the fact that PYTHIA calculations showed no center-of-mass energy dependence of the particle ratios in jet peak and background as an important piece of evidence for the conclusion that the differences in particle ratios between collision systems are related to differences in the size of the collision system. The fact that the particle ratios measured in the inclusive yield of pp and  $\bar{p}p$  collisions are nearly independent of center-of-mass energy for  $\sqrt{s} \gtrsim 1$  TeV [102] leads us to believe that the same will hold for the particle ratios in jet peak and background. Nonetheless, our conclusion would be stronger if we could confirm this experimentally. For this, a pp data sample would be required of similar size ( $\sim 200$ M events) at  $\sqrt{s} = 2.76$  TeV or  $\sqrt{s} = 5.02$  TeV.

Towards lower  $p_T$ , the most important limiting factor in the study of the jet-peak structure is the increasing width of the jet peak, combined with the ALICE detector acceptance of  $|\Delta\eta| < 1.6$ . Especially for protons we expect that for this reason the jet-peak structure will be difficult to study at  $p_{T,\text{assoc}} < 1.0$  GeV/ $c$ . Given the size of the current dataset, the lower limit at which we could study the proton jet peak was  $p_{T,\text{assoc}} \sim 1.4$  GeV/ $c$ , however, we expect that by increasing the size of the dataset this can be extended down to the point where the acceptance becomes the limiting factor. Towards higher  $p_T$ , the main difficulty with the current analysis is the particle identification. Similar to most PID methods, the separation power of the combined fit method decreases with increasing  $p_T$ . Furthermore, the number

of available associated particles becomes smaller at larger  $p_T$ . With a larger dataset the current analysis can likely be extended to  $p_{T,\text{assoc}} \sim 5 - 6 \text{ GeV}/c$ .

Furthermore, with a larger dataset, several extensions to the current study become possible. One option is to perform the study with trigger particles at higher  $p_T$ , and make a comparison with the current results. Furthermore, for Pb–Pb collisions, a larger dataset would allow us to examine the shape of the jet-peak structure in more detail.

# Appendix A.

## Event and Track Cuts

### Event Cuts

Only those events are used of which the primary vertex is reconstructed at most 7cm away from the center of the detector, in the direction of the beam-axis. For p-Pb and Pb-Pb we also demand that the event belongs to the 10% most central events.

### Standard Track Cuts (2010)

The standard track cuts (2010) on the ITS and TPC parameters, which we used for the pp and Pb-Pb datasets, are shown in the table below.

---

<b>ITS</b>	Require clusters in ITS: At least one hit in the SPD Maximum $\chi^2$ per cluster: 36 Require ITS refit: YES $z_{\text{DCA}} < 2.0$ cm
<b>TPC</b>	Minimum number of clusters: 70 Maximum $\chi^2$ per cluster: 4 Require TPC refit: YES Accept Kink Daughters: NO

---

## Standard Track Cuts (2011)

The p-Pb dataset was collected in 2011, and the standard track cuts on the ITS and TPC are slightly altered from the year before. Instead of the requirement of 70 TPC clusters, the track is required to have crossed at least 70 rows in the TPC. Also, the number of crossed rows has to be more than 80% of the number of findable clusters<sup>1</sup>.

## Standard Track Cuts, Loose DCA

The standard cuts are used, however, with a less strict DCA cut:

$$\frac{\sqrt{x_{\text{DCA}}^2 + y_{\text{DCA}}^2}}{2.4 \text{ cm}} + \frac{z_{\text{DCA}}}{3.2 \text{ cm}} < 1. \quad (\text{A.1})$$

## Standard Track Cuts, Tight DCA

The standard cuts are used, however, with a very strict DCA cut:

$$\sqrt{x_{\text{DCA}}^2 + y_{\text{DCA}}^2} < \left( 0.0182 + \frac{0.0350}{p_{\text{T}}^\nu} \right) \text{ cm}, \quad (\text{A.2})$$

where  $\nu = 1.01$  for the 2010 datasets from 2010 and  $\nu = 1.1$  for the datasets from 2011. Additionally, the  $\chi^2$  of the global track constrained to the primary vertex cannot be larger than 36. These two cuts remove a large fraction of secondary particles.

## PID Specific cuts

For the associated tracks the following PID flags are required:

---

<b>TOF</b>	A valid interaction time measurement ( $t_{\text{int}}$ ). A hit in the TOF detector, assigned to the global track.
<b>TPC</b>	At least sixty $dE/dx$ measurements per track.

---

<sup>1</sup>The number of findable clusters decreases for large  $\eta$ , as well as for tracks that intersect a boundary between the readout chambers in the TPC.

# Appendix B.

## Data Sample

pp

- **LHC10d (AOD135)** -  $N_{\text{MB}} = 110\text{M}$

126437, 126432, 126425, 126424, 126422, 126409, 126408, 126407,  
126406, 126405, 126404, 126403, 126359, 126352, 126351, 126285,  
126284, 126283, 126168, 126160, 126158, 126097, 126090, 126088,  
126082, 126081, 126078, 126073, 126008, 126007, 126004, 125855,  
125851, 125850, 125849, 125848, 125847, 125844, 125843, 125842,  
125633, 125632, 125630, 125628, 125296, 125134, 125101, 125100,  
125097, 125085, 125023, 124751, 122375, 122374.

- **LHC10e (AOD135)** -  $N_{\text{MB}} = 95\text{M}$

130840, 130834, 130799, 130798, 130795, 130793, 130704, 130696,  
130519, 130517, 130480, 130356, 130354, 130343, 130342, 130179,  
130178, 130172, 130158, 130157, 130149, 129983, 129961, 129960,  
129959, 129744, 129742, 129738, 129736, 129735, 129729, 129726,  
129725, 129723, 129667, 129666, 129659, 129654, 129653, 129652,  
129650, 129647, 129641, 129639, 129599, 129587, 129586, 129540,  
129528, 129527, 129523, 129520, 129514, 129513, 129512, 128913,  
128855, 128853, 128843, 128836, 128835, 128833, 128824, 128823,  
128820, 128778, 128777, 128678, 128677, 128615, 128611, 128609,  
128605, 128582, 128507, 128504, 128503, 128495, 128494, 128486,  
128483, 128452, 128366, 128260, 128192, 128191, 128189, 128186,  
128185, 127942, 127941, 127940, 127937, 127936, 127935, 127933,  
127822, 127718, 127714, 127712.

- **LHC10f6a, MC (AOD161)** -  $N_{\text{MB}} = 120\text{M}$   
 125855, 125851, 125850, 125849, 125848, 125847, 125101, 125100,  
 125097, 125085, 125023, 126097, 126090, 126008, 126007, 126004,  
 125844, 125843, 125842, 125633, 125632, 125630, 125186, 125156,  
 125140, 125139, 125134, 125133, 122375, 126158, 126088, 126082,  
 126081, 126078, 126073, 125296, 122374.

## p-Pb

- **LHC13c, pass2 (AOD154)** -  $N_{\text{MB}} = 11.6\text{M}$  (0 – 20%)  
 195529, 195531, 195566, 195567, 195568, 195592, 195593, 195596,  
 195633, 195635, 195644, 195673, 195675, 195677.
- **LHC13b2\_efix, MC (AOD158)** -  $N_{\text{MB}} = 29\text{M}$  (0 – 20%)  
 195344, 195346, 195351, 195389, 195390, 195391, 195478, 195479,  
 195480, 195481, 195482, 195483, 195529, 195531, 195566, 195567,  
 195568, 195592, 195593, 195596, 195633, 195635, 195644, 195673,  
 195675, 195677.

## Pb-Pb

- **LHC10h, pass2 (AOD160)** -  $N_{\text{MB}} = 1.37\text{M}$  (0 – 10%)  
 139510, 139507, 139505, 139503, 139465, 139438, 139437, 139360,  
 139329, 139328, 139314, 139310, 139309, 139173, 139107, 139105,  
 139038, 139037, 139036, 139029, 139028, 138872, 138871, 138870,  
 138837, 138732, 138730, 138666, 138662, 138653, 138652, 138638,  
 138624, 138621, 138583, 138582, 138579, 138578, 138534, 138469,  
 138442, 138439, 138438, 138396, 138364, 138275, 138225, 138201,  
 138197, 138192, 138190, 137848, 137844, 137752, 137751, 137724,  
 137722, 137718, 137704, 137693, 137692, 137691, 137686, 137685,  
 137639, 137638, 137608, 137595, 137549, 137544, 137541, 137539,  
 137443, 137441, 137440, 137439, 137434, 137432, 137431, 137430,  
 137366, 137243, 137236, 137235, 137232, 137231, 137230, 137162,  
 137161, 137135.
- **LHC11a\_10a\_bis, MC (AOD143)** -  $N_{\text{MB}} = 80\text{k}$  (0 – 10%)  
 same runs as LHC10h.

# Appendix C.

## The Discrete Cosine Transform

The discrete cosine transform (DCT) can be used to determine the Fourier cosine coefficients of a discrete data set. There are several prescriptions on how to discretize the continuous formulation of the cosine transform, however in this work we will only use the most common one, DCT-II.

Suppose, one period of a signal has been measured on  $N$  equally spaced intervals, then the DCT-II transform  $X_n$  is defined as:

$$X_n = \sum_{k=0}^{N-1} x_k \cos \left[ \frac{\pi}{N} \left( k + \frac{1}{2} \right) n \right], \quad (\text{C.1})$$

where  $x_k$  are the measurements. Treating these measurements as uncorrelated, we can write the statistical uncertainty  $\sigma_{X_n}$  as:

$$\sigma_{X_n} = \sqrt{\sum_{k=0}^{N-1} \sigma_{x_k}^2 \cos^2 \left[ \frac{\pi}{N} \left( k + \frac{1}{2} \right) n \right]}. \quad (\text{C.2})$$

Now we turn to the application of the DCT to determine the flow coefficients  $v_{n\Delta}$ . In Eq. (6.2) we defined these coefficients as follows:

$$\frac{dN}{d\Delta\phi} = \left\langle \frac{dN}{d\Delta\phi} \right\rangle \left( 1 + 2 \sum_{n=1}^{\infty} V_{n\Delta} \cos n\Delta\phi \right). \quad (\text{C.3})$$

Using the DCT, the average yield can be extracted as follows:

$$\left\langle \frac{dN}{d\Delta\phi} \right\rangle = \frac{1}{\pi} \int_0^\pi \frac{dN}{d\Delta\phi} d\Delta\phi \simeq \frac{1}{N} \sum_{k=0}^{N-1} \left( \frac{dN}{d\Delta\phi} \right)_k = \frac{X_0}{N}, \quad (\text{C.4})$$

where:  $\left( \frac{dN}{d\Delta\phi} \right)_k = \frac{dN}{d\Delta\phi}(\Delta\phi_k)$ , and  $\Delta\phi_k = \frac{\pi}{N}(k + \frac{1}{2})$ .

The flow coefficients can be written as:

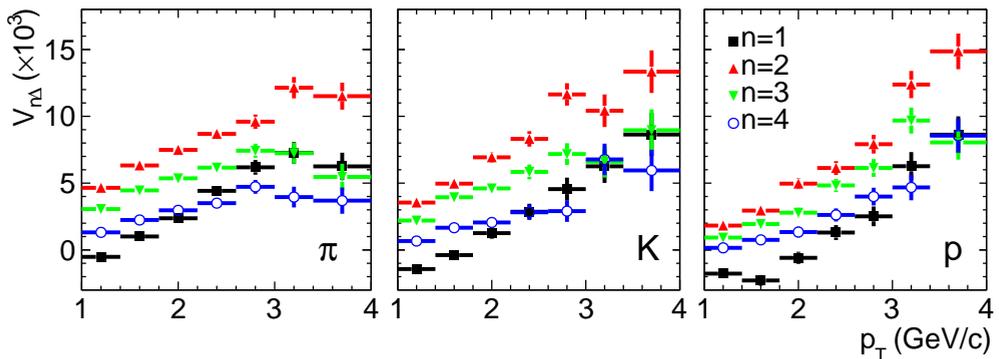
$$\begin{aligned} V_{n\Delta} &= \frac{1}{\pi \left\langle \frac{dN}{d\Delta\phi} \right\rangle} \int_0^\pi \frac{dN}{d\Delta\phi} \cos(n\Delta\phi) d\Delta\phi \\ &\simeq \frac{1}{X_0} \sum_{k=0}^{N-1} \left( \frac{dN}{d\Delta\phi} \right)_k \cos \left[ \frac{\pi}{N} \left( k + \frac{1}{2} \right) n \right] = \frac{X_n}{X_0}, \quad (\text{C.5}) \end{aligned}$$

where we used the identity  $\int_0^\pi \cos(nx) \cos(mx) dx = \pi/2$  for  $n = m \neq 0$ .

# Appendix D.

## Associated Yield Projections

In this appendix we provide some additional figures, illustrating the way we model the background of the associated yield distribution in central Pb–Pb collisions. In Fig. D.1 we show the values of  $V_{n\Delta}$  with  $n = 1..4$  for pions, kaons, and protons.

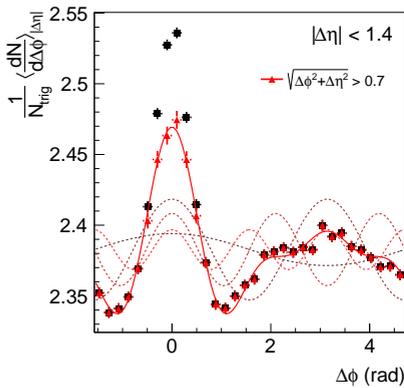
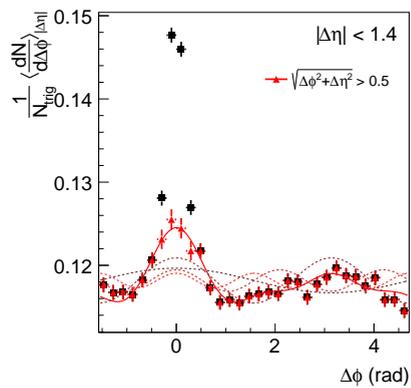
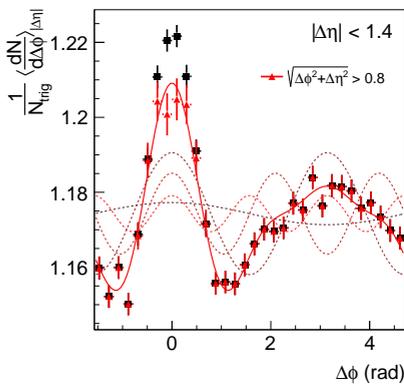
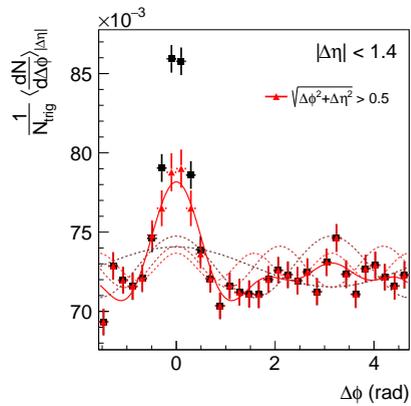
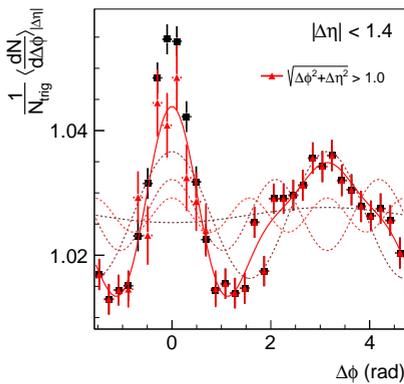
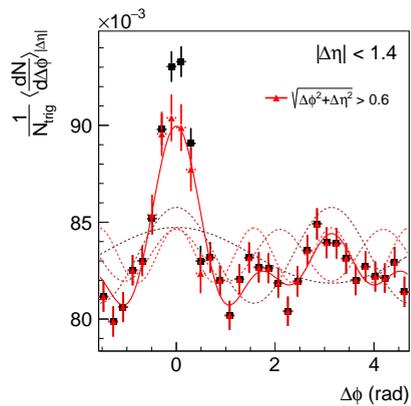


**Figure D.1.:**  $V_{n\Delta}(p_T)$  with  $n = 1..4$  for pions, kaons, and protons. These values are determined using a DCT (see Eq. (C.5)) on the identified yield associated with a high- $p_T$  trigger particle, with the near-side peak removed (see Sect. 6.3).

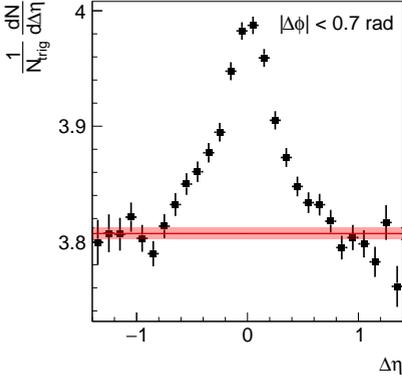
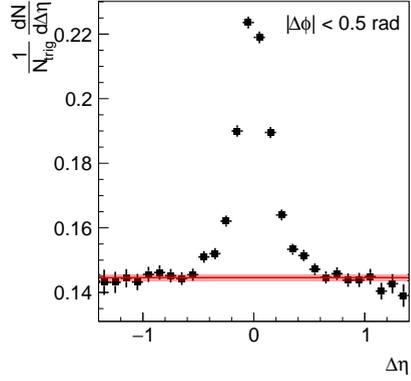
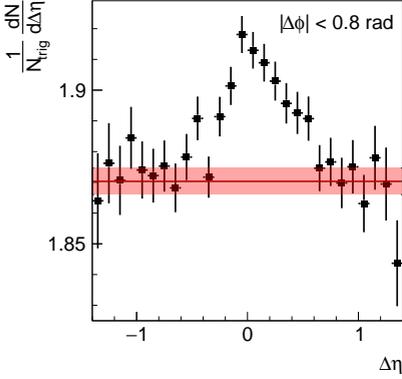
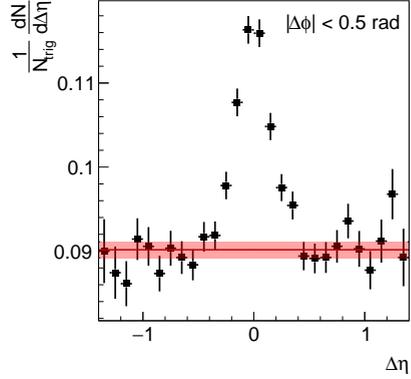
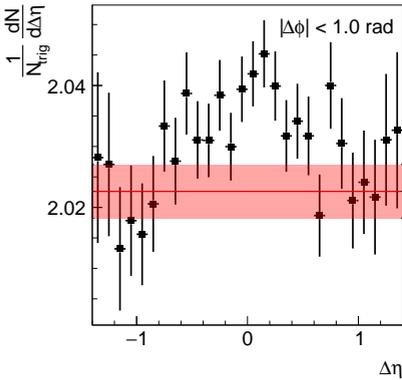
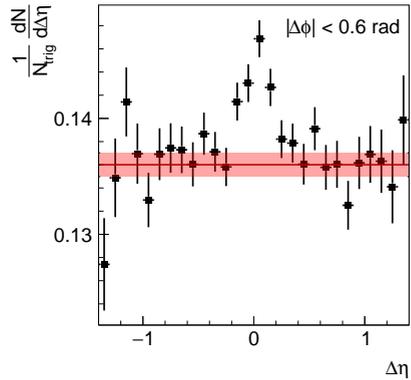
In Fig. D.2 we show examples of the decomposition of the background, used to measure the values in Fig. D.1. For each particle species ( $\pi$ ,  $K$ ,  $p$ ), we show a low- $p_T$  bin and a high- $p_T$  bin. From these figures one can clearly see the difference in signal strength between low and high  $p_T$  and among the

different particle species. Furthermore, it can be seen that the first four DCT coefficients give a good description of the background.

Finally, in Fig. D.3, we show examples of the projection of the associated yield onto the  $\Delta\eta$  axis, using the same  $p_T$  bins as in Fig. D.2. The range in  $\Delta\phi$  is chosen to correspond to the size of the jet-peak region (see Sect. 6.3). From these figures, we can clearly see that the jet peaks are much wider at low  $p_T$  compared to high  $p_T$ . At high  $p_T$ , the jet peaks become sufficiently narrow to see that the background at large  $\Delta\eta$  becomes flat.

(a) Pions,  $1.8 < p_T < 2.2$  GeV/c(b) Pions,  $3.4 < p_T < 4.0$  GeV/c(c) Kaons,  $1.8 < p_T < 2.2$  GeV/c(d) Kaons,  $3.4 < p_T < 4.0$  GeV/c(e) Protons,  $1.8 < p_T < 2.2$  GeV/c(f) Protons,  $3.4 < p_T < 4.0$  GeV/c

**Figure D.2.:** Decomposition of the identified yield associated with a high- $p_T$  trigger particle, using the first four terms of the DCT expansion (see Eq. (C.5)).

(a) Pions,  $1.8 < p_T < 2.2$  GeV/ $c$ (b) Pions,  $3.4 < p_T < 4.0$  GeV/ $c$ (c) Kaons,  $1.8 < p_T < 2.2$  GeV/ $c$ (d) Kaons,  $3.4 < p_T < 4.0$  GeV/ $c$ (e) Protons,  $1.8 < p_T < 2.2$  GeV/ $c$ (f) Protons,  $3.4 < p_T < 4.0$  GeV/ $c$ 

**Figure D.3.:** Projection of the identified yield associated with a high- $p_T$  trigger particle on the  $\Delta\eta$  axis, together with the estimate of the background level. The statistical uncertainty on the estimated background is represented by the lighter red band.

# Bibliography

- [1] H. Geiger and E. Marsden. On a diffuse reflection of the  $\alpha$ -particles. *Proceedings of the Royal Society of London A*, 82:495–500, 1909.
- [2] J. Chadwick. Possible existence of a neutron. *Nature*, 129:312, February 1932.
- [3] J. Chadwick. Bakerian lecture - the neutron. *Proceedings of the Royal Society of London A*, 142, October 1933.
- [4] M. Gell-Mann. A schematic model of baryons and mesons. *Physics Letters*, 8(3):214–215, 1964.
- [5] G. Zweig. An  $SU_3$  Model for Strong Interaction Symmetry and its Breaking. Technical Report CERN-TH-412, CERN, February 1964.
- [6] G. Zweig. Memories of Murray and the Quark Model. *International Journal of Modern Physics A*, 25(20):3863–3877, 2010.
- [7] R. K. Ellis, W. J. Stirling, and B. R. Webber. *QCD and Collider Physics*. Cambridge University Press, 1996.
- [8] D. J. Gross and F. A. Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Physical Review Letters*, 30:1343–1346, June 1973.
- [9] H. D. Politzer. Reliable perturbative results for strong interactions? *Physical Review Letters*, 30:1346–1349, June 1973.
- [10] J. C. Collins and M. J. Perry. Superdense matter: Neutrons or asymptotically free quarks? *Physical Review Letters*, 34:1353–1356, May 1975.
- [11] G. S. Bali and K. Schilling. Static quark-antiquark potential: Scaling behavior and finite-size effects in  $SU(3)$  lattice gauge theory. *Physical*

- Review D*, 46:2636–2646, September 1992.
- [12] CMS Collaboration. Studies of jet quenching using isolated-photon+jet correlations in Pb–Pb and pp collisions at. *Physics Letters B*, 718(3):773–794, January 2013.
- [13] ATLAS Collaboration. Measurement of the jet radius and transverse momentum dependence of inclusive jet suppression in lead-lead collisions at with the ATLAS detector. *Physics Letters B*, 719(4-5):220–241, February 2013.
- [14] M. Cacciari, G. P. Salam, and G. Soyez. The anti- $k_t$  jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063, April 2008.
- [15] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber. Longitudinally-invariant  $k_\perp$ -clustering algorithms for hadron-hadron collisions. *Nuclear Physics B*, 406(1-2):187–224, September 1993.
- [16] ALICE Collaboration. Harmonic decomposition of two particle angular correlations in Pb–Pb collisions at. *Physics Letters B*, 708(3-5):249–264, February 2012.
- [17] STAR Collaboration. Disappearance of Back-To-Back High- $p_T$  Hadron Correlations in Central Au+Au Collisions at  $\sqrt{s_{NN}} = 200$  GeV. *Physical Review Letters*, 90(8):082302, February 2003.
- [18] STAR Collaboration. Long range rapidity correlations and jet production in high energy nuclear collisions. *Physical Review C*, 80(6):064912, December 2009.
- [19] STAR Collaboration. Distributions of Charged Hadrons Associated with High Transverse Momentum Particles in pp and Au+Au Collisions at  $\sqrt{s_{NN}} = 200$  GeV. *Physical Review Letters*, 95(15):152301, October 2005.
- [20] PHENIX Collaboration. System Size and Energy Dependence of Jet-Induced Hadron Pair Correlation Shapes in Cu+Cu and Au+Au Collisions at  $\sqrt{s_{NN}} = 200$  and 62.4 GeV. *Physical Review Letters*, 98(23):232302, June 2007.
- [21] PHENIX Collaboration. Dihadron azimuthal correlations in Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV. *Physical Review C*, 78(1):014901, July

- 2008.
- [22] PHOBOS Collaboration. High Transverse Momentum Triggered Correlations over a Large Pseudorapidity Acceptance in Au+Au Collisions at  $\sqrt{s_{\text{NN}}} = 200$  GeV. *Physical Review Letters*, 104(6):062301, February 2010.
- [23] PHOBOS Collaboration. System size dependence of cluster properties from two-particle angular correlations in Cu+Cu and Au+Au collisions at  $\sqrt{s_{\text{NN}}} = 200$  GeV. *Physical Review C*, 81(2):024904, February 2010.
- [24] CMS Collaboration. Centrality dependence of dihadron correlations and azimuthal anisotropy harmonics in Pb–Pb collisions at  $\sqrt{s_{\text{NN}}} = 2.76$  TeV. *The European Physical Journal C*, 72(5):2012, May 2012.
- [25] ALICE Collaboration. Particle-yield modification in jetlike azimuthal dihadron correlations in Pb–Pb collisions at  $\sqrt{s_{\text{NN}}} = 2.76$  TeV. *Physical Review Letters*, 108:1–11, March 2012.
- [26] S. Sapeta and U. A. Wiedemann. Jet hadrochemistry as a characteristic of jet quenching. *The European Physical Journal C*, 55(2):293–302, 2008.
- [27] W. Florkowski. *Phenomenology of Ultra-Relativistic Heavy-Ion Collisions*. World Scientific, 2010.
- [28] STAR Collaboration. Experimental and theoretical challenges in the search for the quark-gluon plasma: The STAR collaboration’s critical assessment of the evidence from RHIC collisions. *Nuclear Physics A*, 757(1-2):102–183, 2005.
- [29] J. D. Bjorken. Highly relativistic nucleus-nucleus collisions: The central rapidity region. *Physical Review D*, 27:140–151, January 1983.
- [30] R. Glauber. Cross Sections in Deuterium at High Energies. *Physical Review*, 100(1):242–248, October 1955.
- [31] ALICE Collaboration. Centrality determination of Pb–Pb collisions at  $\sqrt{s_{\text{NN}}} = 2.76$  TeV with ALICE. *Physical Review C*, 88:044909, October 2013.
- [32] U. Heinz and R. Snellings. Collective flow and viscosity in relativistic heavy-ion collisions. *Annual Review of Nuclear and Particle Science*,

- 63(1):123–151, 2013.
- [33] CMS Collaboration. Measurement of the Pseudorapidity and Centrality Dependence of the Transverse Energy Density in Pb–Pb Collisions at  $\sqrt{s_{\text{NN}}} = 2.76$  TeV. *Physical Review Letters*, 109:152303, October 2012.
- [34] E. V. Shuryak. Quantum chromodynamics and the theory of superdense matter. *Physics Reports*, 61(2):71–158, May 1980.
- [35] M. L. Miller, K. Reygers, S. J. Sanders, and P. Steinberg. Glauber Modeling in High-Energy Nuclear Collisions. *Annual Review of Nuclear and Particle Science*, 57:205–243, 2007.
- [36] A. Białas, M. Bleszyński, and W. Czyż. Multiplicity distributions in nucleus-nucleus collisions at high energies. *Nuclear Physics B*, 111:461–476, 1976.
- [37] R. D. Woods and D. S. Saxon. Diffuse surface optical model for nucleon-nuclei scattering. *Physical Review*, 95:577–578, July 1954.
- [38] W. Broniowski, M. Rybczyński, and P. Bożek. GLISSANDO: GLauber Initial-State Simulation AND mOre... *Computer Physics Communications*, 180(1):69–83, 2009.
- [39] P. Bożek and W. Broniowski. Transverse-momentum fluctuations in relativistic heavy-ion collisions from event-by-event viscous hydrodynamics. *Physical Review C*, 85:044910, April 2012.
- [40] A. Bilandzic, R. Snellings, and S. Voloshin. Flow analysis with cumulants: Direct calculations. *Physical Review C*, 83(4):044913, April 2011.
- [41] P. Braun-Munzinger and K. Redlich and J. Stachel. Particle Production in Heavy Ion Collisions. *arXiv:0304013 [nucl-th]*, 2003.
- [42] A. Andronic and P. Braun-Munzinger and K. Redlich and J. Stachel. The statistical model in Pb–Pb collisions at the LHC. *Nuclear Physics A*, 904–905(0):535c–538c, 2013.
- [43] J. Steinheimer, J. Aichelin, and M. Bleicher. Nonthermal  $p/\pi$  ratio at lhc as a consequence of hadronic final state interactions. *Physical review letters*, 110(4):042501, 2013.

- [44] F. Becattini, P. Castorina, J. Manninen, and H. Satz. The thermal production of strange and non-strange hadrons in  $e^+e^-$  collisions. *The European Physical Journal C*, 56(4):493–510, 2008.
- [45] F. Becattini and U. Heinz. Thermal hadron production in  $pp$  and  $p\bar{p}$  collisions. *Zeitschrift für Physik C Particles and Fields*, 76(2):269–286, 1997.
- [46] P. Castorina and H. Satz. Causality constraints on hadron production in high energy collisions. *International Journal of Modern Physics E*, 23(04):1450019, 2014.
- [47] R. N. Cahn. The eighteen arbitrary parameters of the standard model in your everyday life. *Rev. Mod. Phys.*, 68:951–959, July 1996.
- [48] S. W. Hawking. Particle creation by black holes. *Communications In Mathematical Physics*, 43(3):199–220, August 1975.
- [49] V. Trimble. Existence and Nature of Dark Matter in the Universe. *Annual Review of Astronomy and Astrophysics*, 25(1):425–472, September 1987.
- [50] S. Dimopoulos and H. Georgi. Softly broken supersymmetry and SU(5). *Nuclear Physics B*, 193(1):150–162, December 1981.
- [51] G. 't Hooft. Introduction To String Theory, 1994. Lecture Notes.
- [52] OPAL Collaboration. A study of jet production rates and a test of QCD on the  $Z^0$  resonance. *Physics Letters B*, 235(3-4):389–398, 1990.
- [53] J. Binnewies, B. A. Kniehl, and G. Kramer. Pion and kaon production in  $e^+e^-$  and ep collisions at next-to-leading order. *Physical Review D*, 52:4947–4960, November 1995.
- [54] B. A. Kniehl and G. Kramer and B. Pötter. Fragmentation functions for pions, kaons, and protons at next-to-leading order. *Nuclear Physics B*, 582(1-3):514–536, 2000.
- [55] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nuclear Physics B*, 126:298–318, 1977.
- [56] M. A. Stephanov. QCD phase diagram: an overview. December 2006.

- [57] K. G. Wilson. Confinement of quarks. *Physical Review D*, 10:2445–2459, October 1974.
- [58] K. G. Wilson. The origins of lattice gauge theory. *Nuclear Physics B - Proceedings Supplements*, 140:3–19, 2005. LATTICE 2004 Proceedings.
- [59] A. Bazavov, T. Bhattacharya, C. DeTar, et al. Equation of state in (2+1)-flavor QCD. *Physical Review D*, 90(9):094503, November 2014.
- [60] T. Nayak. Search for the QCD Critical Point at RHIC: First Experience with Au-Au Collisions at 9.2 GeV, 2008.
- [61] Y. Aoki, G. Endrodi, Z. Fodor, S. D. Katz, and K. K. Szabó. The order of the quantum chromodynamics transition predicted by the standard model of particle physics. *Nature*, 443(7112):675–8, October 2006.
- [62] G. Martinez. Advances in Quark Gluon Plasma. April 2013.
- [63] L. Kumar. Review of Recent Results from the RHIC Beam Energy Scan. *Modern Physics Letters A*, 28(36):16, November 2013.
- [64] ALICE Collaboration. Production of charged pions, kaons and protons at large transverse momenta in pp and Pb–Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV. *Physics Letters B*, 736:196–207, September 2014.
- [65] STAR Collaboration. Identified Baryon and Meson Distributions at Large Transverse Momenta from Au + Au Collisions at  $\sqrt{s_{NN}} = 200$  GeV. *Physical Review Letters*, 97:152301, October 2006.
- [66] PHENIX Collaboration. Scaling Properties of Proton and Antiproton Production in  $\sqrt{s_{NN}} = 200$  GeV Au + Au Collisions. *Physical Review Letters*, 91:172301, October 2003.
- [67] P. Bożek and I. Wyskiel-Piekarska. Particle spectra in Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV. *Physical Review C*, 85(6):064915, 2012.
- [68] R. J. Fries, B. Müller, C. Nonaka, and S. A. Bass. Hadron production in heavy ion collisions: Fragmentation and recombination from a dense parton phase. *Physical Review C*, 68(4):044902, 2003.
- [69] K. Werner, I. Karpenko, M. Bleicher, T. Pierog, and S. Porteboeuf-Houssais. Jets, bulk matter, and their interaction in heavy ion collisions at several TeV. *Physical Review C*, 85(6):064907, 2012.

- [70] R. C. Hwa and C. B. Yang. Scaling behavior at high  $p_T$  and the  $p/\pi$  ratio. *Physical Review C*, 67:034902, March 2003.
- [71] R. C. Hwa and C. B. Yang. Recombination of shower partons in fragmentation processes. *Physical Review C*, 70:024904, August 2004.
- [72] R. C. Hwa and C. B. Yang. Recombination of shower partons at high  $p_T$  in heavy-ion collisions. *Physical Review C*, 70:024905, August 2004.
- [73] R. C. Hwa. Hadron Correlations in Jets and Ridges Through parton Recombination. *arXiv:0904.2159 [nucl-th]*, 2009.
- [74] L. Zhu and R. C. Hwa. Effects of Shower Partons on Soft and Semi-hard hadrons Produced in Pb–Pb Collisions at  $\sqrt{s_{NN}} = 2.76$  TeV. *arXiv:1406.5733 [nucl-th]*, 2014.
- [75] R. C. Hwa and C. B. Yang. Hadron correlation in jets on the near and away sides of high- $p_T$  triggers in heavy-ion collisions. *Physical Review C*, 79:044908, April 2009.
- [76] PHENIX Collaboration. Formation of dense partonic matter in relativistic nucleus-nucleus collisions at RHIC: Experimental evaluation by the PHENIX Collaboration. *Nuclear Physics A*, 757(1-2):184–283, August 2005.
- [77] E. Shuryak. Why does the quark-gluon plasma at RHIC behave as a nearly ideal fluid? *Progress in Particle and Nuclear Physics*, 53(1):273–303, July 2004.
- [78] G. Policastro, D. T. Son, and A. O. Starinets. From AdS/CFT correspondence to hydrodynamics. *Journal of High Energy Physics*, 2002(09):043–043, September 2002.
- [79] J. L. Nagle. The letter “s” (and the sQGP). *European Physical Journal C*, 49:275–279, 2007.
- [80] CERN. LHC: the guide, 2009.
- [81] ALICE Collaboration, F. Carminati, P. Foka, et al. ALICE: Physics Performance Report, Volume I. *Journal of Physics G: Nuclear and Particle Physics*, 30(11):1517, 2004.
- [82] ALICE Collaboration, B. Alessandro, F. Antinori, et al. ALICE: Physics Performance Report, Volume II. *Journal of Physics G: Nuclear*

- and Particle Physics*, 32(10):1295, 2006.
- [83] ALICE Collaboration. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002–S08002, August 2008.
- [84] ALICE Collaboration. Performance of the ALICE experiment at the CERN LHC. *International Journal of Modern Physics A*, 29(24):1430044, September 2014.
- [85] K. A. Olive et al. Review of Particle Physics. *Chin.Phys.*, C38:090001, 2014.
- [86] C. Lippmann. Particle identification. *Nuclear Instruments and Methods in Physics Research A*, 666:148–172, 2012.
- [87] H. Bichsel. A method to improve tracking and particle identification in tpcs and silicon detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 562(1):154–197, 2006.
- [88] M. Bondila, V. A. Grigoriev, F. F. Guber, et al. ALICE T0 detector. *IEEE Transactions on Nuclear Science*, 52(5):1705–1711, 2005.
- [89] A. Akindinov, A. Alici, A. Agostinelli, et al. Performance of the ALICE Time-Of-Flight detector at the LHC. *European Physical Journal Plus*, 128:1–9, 2013.
- [90] A. Bhasin, D. Evans, G. T. Jones, et al. Minimum bias trigger for pp physics with ALICE. *Indian Journal of Physics*, 84(12):1733–1738, 2010.
- [91] M. Ivanov, I. Belikov, P. Hristov, and K. Šafařík. Track reconstruction in high density environment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 566(1):70–74, October 2006.
- [92] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 262(2-3):444–450, December 1987.
- [93] R. Brun, A. C. McPherson, P. Zancarini, M. Maire, and F. Bruyant. Geant 3. Technical report, CERN, 1987.

- [94] ALICE Collaboration. Centrality dependence of  $\pi$ , K, and p production in Pb–Pb collisions at  $\sqrt{s_{\text{NN}}} = 2.76$  TeV. *Physical Review C*, 88(4):044910, October 2013.
- [95] ALICE Collaboration. Multiplicity dependence of jet-like two-particle correlation structures in p-Pb collisions at  $\sqrt{s_{\text{NN}}} = 5.02$  GeV. *Physics Letters B*, 741(0):38–50, 2015.
- [96] G. Cowan. *Statistical Data Analysis*. Oxford University Press, 1998.
- [97] T. Sjöstrand, S. Mrenna, and P. Skands. PYTHIA 6.4 physics and manual. *Journal of High Energy Physics*, 2006(05):026–026, May 2006.
- [98] S. Roesler, R. Engel, and J. Ranft. The Monte Carlo Event Generator DPMJET-III. In *Advanced Monte Carlo for Radiation Physics, Particle Transport Simulation and Applications*, pages 1033–1038. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.
- [99] M. Gyulassy and X. Wang. HIJING 1.0: A Monte Carlo program for parton and particle production in high energy hadronic and nuclear collisions. *Computer Physics Communications*, 83(2-3):307–331, December 1994.
- [100] ALICE Collaboration. Midrapidity Antiproton-to-Proton Ratio in pp Collisions at  $\sqrt{s} = 0.9$  and 7 TeV Measured by the ALICE Experiment. *Physical Review Letters*, 105(7):072002, 2010.
- [101] G. Battistoni, F. Cerutti, A. Fasso, et al. The fluka code: Description and benchmarking. In *Hadronic Shower Simulation Workshop (AIP Conference Proceedings Volume 896)*, volume 896, pages 31–49, 2007.
- [102] ALICE Collaboration. Measurement of pion, kaon and proton production in proton-proton collisions at  $\sqrt{s} = 7$  TeV. *The European Physical Journal C*, 75(5):226, May 2015.
- [103] T. Renk and K. J. Eskola. Proton-antiproton suppression in 200 AGeV Au–Au collisions. *Physical Review C*, 76(2):027901, August 2007.
- [104] S. Albino, B. A. Kniehl, and G. Kramer. Fragmentation functions for light charged hadrons with complete quark flavour separation. *Nuclear Physics B*, 725(1-2):181–206, October 2005.

- [105] J. Gallicchio and M. D. Schwartz. Quark and Gluon Tagging at the LHC. *Physical Review Letters*, 107(17):172001, June 2011.
- [106] M. Veldhoen.  $p/\pi$  Ratio in Di-Hadron Correlations. *Nuclear Physics A*, 910-911:306–309, 2013. Hard Probes 2012 5th International Conference on Hard and Electromagnetic Probes of High-Energy Nuclear Collisions.
- [107] J. Grosse-Oetringhaus. Hadron correlations measured with ALICE. *Nuclear Physics A*, 910-911:58–64, 2013. Hard Probes 2012 5th International Conference on Hard and Electromagnetic Probes of High-Energy Nuclear Collisions.

# Summary

The nucleus of any atom consists of neutrons, which carry no electrical charge, and positively charged protons. As protons electrically repel each other, there must be another force holding the atomic nucleus together as a stable entity. This force is called the *nuclear force*, and it is the energy stored in the nuclear force fields that is harnessed in nuclear fission and nuclear fusion. An interesting property of the nuclear force is that its strength becomes negligible at scales larger than the typical size of an atomic nucleus.

Today, the nuclear force is understood as an effective force, resulting from a more fundamental theory that describes the dynamics of the interior of protons and neutrons. In this theory, called *quantum chromodynamics (QCD)*, protons and neutrons are a bound state of several lighter particles, called *quarks*, held together by a fundamental force called the *strong force*. The mediator particles of the strong force are called *gluons* and the charge of the strong force, which is only carried by quarks and gluons, is called *color charge*. The proton and neutron are composed of quarks in such a way that they are themselves color neutral, which means that to first order they do not interact with other color-neutral objects through the strong force. However, because protons and neutrons have a non-trivial color charge distribution, a residual force exists between them, reminiscent of the van der Waals force. This residual force is nowadays identified with the nuclear force. While protons and neutrons are the most well-known bound states of quarks, many more occur in nature. These bound states are collectively called *hadrons*. In this work we are concerned with the three electrically charged hadrons that are most copiously produced in relativistic hadronic and nuclear collisions, namely the pion ( $\pi$ ), kaon ( $K$ ), and proton ( $p$ ).

Only color-neutral particles have been observed as free particles. In particular, no free quark or gluon has ever been observed, an effect known as *confinement*. Numerical studies of QCD thermodynamics suggest that at extreme temperatures quarks and gluons are no longer confined, instead QCD

undergoes a phase transition to a phase where quarks and gluons become deconfined, i.e., they can move over distance scales that are much larger than the typical size of a proton or neutron. This phase is known as the *Quark-Gluon Plasma (QGP)*.

The temperatures necessary to achieve the deconfinement phase transition are so high that a QGP does not occur naturally on Earth or even in our solar system. One way to create a physical system with a sufficiently high temperature in the laboratory is by colliding heavy ions, for example lead (Pb) or gold (Au), at extreme collision energies. The resulting QGP exists only very briefly, as it rapidly expands and cools directly after the collision. Once the system cools to a temperature below the temperature of the deconfinement phase transition the quarks and gluons turn into hadrons again. This process is also referred to as *hadronization*. The properties of the produced hadrons (identity, momentum, energy) can be determined using a particle detector, and observables based on these properties can be compared to model predictions or to observations from simpler particle collisions, for example, proton-proton collisions. From these comparisons one can infer information about the nature of the QGP.

The details of the mechanism driving hadronization in a cooling QGP are as of yet not known. In much simpler particle collisions, such as proton-proton collisions, relatively few quarks and gluons are produced, and the dominant mechanism of hadronization is *fragmentation*. The quarks and gluons radiate off other quarks and gluons, producing *showers*, which then hadronize into a collimated spray of hadrons, called a *jet*. The key feature here is that much of the information about the original quark or gluon (energy, momentum) is still contained in the jet it produces. On the other hand, when a QGP hadronizes, the density of quarks and gluons is much higher than in a proton-proton collision. It is therefore plausible that hadronization can take place through the recombination of thermalized quarks, or through a combination of fragmentation and recombination.

Measurements of hadron production ratios, in particular the  $p/\pi$  ratio, have shown to be much larger in central (head-on) Pb–Pb collisions than in pp collisions at intermediate transverse momentum<sup>1</sup> ( $1 < p_T < 8$  GeV/c). The  $p/\pi$  ratio in Pb–Pb collisions was found to be consistent with various recombination models.

---

<sup>1</sup>The transverse momentum  $p_T$  is the component of the momentum transverse to the beam pipe of the collider.

In this work we have extended on these previous results by investigating the hadron production ratios in the jet peaks and background of the yield associated with a high- $p_T$  trigger particle. This means that in every collision we collect the particles with a large transverse momentum ( $p_T > 4$  GeV/ $c$ ), and create a distribution of all the particles produced in the same collision with  $1 < p_T < 4$  GeV/ $c$  as a function of the difference in the azimuthal angle  $\Delta\phi$  and the pseudorapidity  $\Delta\eta$  with respect to the high- $p_T$  (trigger) particle<sup>2</sup>. As the trigger particle was likely part of a jet, the associated particle distribution has an (approximately Gaussian) peak structure at the origin, where the particle production is increased compared to the rest of the event. We investigated both the  $p/\pi$  and  $K/\pi$  ratios in this “jet peak” as well as in the background, that is, the region at large  $(\Delta\eta, \Delta\phi)$ . We performed this measurement for pp, central p-Pb, and central Pb-Pb collisions.

We developed several analysis techniques to make this measurement possible. The identification of the particles was done using two separate detector systems, namely the Time-of-Flight (TOF) detector and the Time Projection Chamber (TPC). The TOF detector is located far from the interaction point, and it measures  $\Delta t$ , the time a particle takes to travel from the interaction point to the TOF detector. The TPC consists of a cylindrical chamber filled with gas and it measures  $dE/dx$ , the energy that a charged particle loses per unit distance as it traverses the gas. Previously, fits to the individual TOF and TPC signals were used to identify particles, however, we performed a fit of the combined signal, thereby increasing the  $p_T$ -range over which a reliable identification could be achieved. The fit function, containing parameterizations for pions, kaons, and protons, consists of nineteen fit parameters.

To identify the yield associated with a high- $p_T$  trigger particle, we created a four-dimensional matrix  $(\Delta\eta, \Delta\phi, \Delta t, dE/dx)$ , and determined the number of pions, kaons, and protons in each  $(\Delta\eta, \Delta\phi)$  bin by performing a fit to the combined  $(\Delta t, dE/dx)$  signal for that bin. Since the detector response (expected value, resolution) is generally a function of  $p_T$  and  $\eta$ , the signal in a single  $(\Delta\eta, \Delta\phi)$  bin is “smeared out”. To deal with this we constructed fit-templates of the expected shapes of the response, based on the full nineteen-dimensional fits performed as a function of  $p_T$  and  $\eta$ . An additional advantage of using these fit templates was that the number of free parameters in the fit reduced to four, making the fit usable for much smaller datasets.

---

<sup>2</sup>Just like  $\Delta\phi$ ,  $\Delta\eta$  can be thought of as a measure of the angular distance to the trigger particle.

After we obtained the identified associated yields, we measured the yield in the jet peak by subtracting the background. The level of the background was determined at  $(\Delta\eta, \Delta\phi)$  greater than approximately three times the expected Gaussian width of the jet peak, which was found by performing fits to the  $\Delta\phi$  and  $\Delta\eta$  projections of the associated yield. In the case of pp collisions, the background is flat, however, in case of p-Pb and Pb-Pb collisions there is a dependence on  $\Delta\phi$  due to radial flow<sup>3</sup>. The radial flow was included in the description of the background by performing a discrete cosine transform.

We found that the Gaussian width of the jet peak decreases with  $p_T$  and increases with the mass of the particle. This behavior is not reproduced by calculations from the PYTHIA model. We also found that in pp, p-Pb and Pb-Pb collisions, both the  $p/\pi$  and  $K/\pi$  ratio are larger in the background than in the jet peak, the difference being the largest in the  $p/\pi$  ratio in Pb-Pb collisions. This suggests that in the jet peak of Pb-Pb collisions, fragmentation plays an important role in the production of hadrons. Additionally we found that there is a system-size dependence in the particle ratios in the jet peaks. This leads us to believe that, while fragmentation plays an important role in the hadron production in these jets, the fragmentation process is likely influenced by the presence of the QGP.

---

<sup>3</sup>Fluid behavior of the QGP.

# Samenvatting

De nucleus van een atoom bestaat uit neutronen, die elektrisch neutraal zijn, en positief geladen protonen. Omdat protonen elkaar elektrisch afstoten, moet er een andere kracht zijn die er voor zorgt dat het atoom niet uit elkaar valt. Deze kracht heet de *kernkracht*, en het is de energie in de kernkrachtvelden van een atoom die wordt gewonnen bij kernsplijting of kernfusie. Een interessante eigenschap van de kernkracht is dat de sterke van de kracht verwaarloosbaar klein wordt op schalen die groter zijn dan de typische grootte van een atoomkern.

Tegenwoordig wordt de kernkracht begrepen als een effectieve kracht, voortkomend uit een fundamentele theorie die beschrijft hoe protonen en neutronen van binnen werken. Binnen deze theorie, genaamd *quantum chromodynamica (QCD)*, worden protonen en neutronen beschreven als een gebonden toestand van verschillende lichtere deeltjes genaamd *quarks*, welke bij elkaar worden gehouden door een fundamentele kracht genaamd de *sterke kernkracht*. De deeltjes die deze kracht doorgeven worden *gluonen* genoemd, en de lading van de sterke kernkracht, die alleen gedragen wordt door quarks en gluonen heet *kleurlading*. De quarksamenstelling van protonen en neutronen zijn zodanig dat het proton en het neutron kleurneutraal zijn, dit betekent dat ze tot op eerste orde niet via de sterke kernkracht interacteren met andere kleurneutrale objecten. Echter, omdat de kleurlading binnenin een proton en neutron niet homogeen verdeeld is, bestaat er een restkracht tussen het proton en het neutron, vergelijkbaar met de van der Waals kracht. Deze restkracht is precies de kernkracht wat het atoom stabiel houdt. Protonen en neutronen zijn de meest bekende gebonden quarktoestanden, maar er blijken er veel meer voor te komen in de natuur. Een gemeenschappelijke naam voor deze gebonden toestanden is *hadronen*. In dit proefschrift houden we ons uitsluitend bezig met de drie elektrisch geladen hadronen, die het meest geproduceerd worden in een relativistische hadronische botsing. Dit zijn: het pion ( $\pi$ ), het kaon ( $K$ ) en het proton ( $p$ ).

Alle vrije deeltjes die ooit zijn geobserveerd blijken kleurneutraal te zijn. Een vrij quark is dus nooit waargenomen, dit effect wordt ook wel *confinement* genoemd. Numerieke berekeningen van QCD thermodynamica laten zien dat bij zeer extreme temperaturen quarks en gluonen zich wel als vrije deeltjes kunnen gedragen. Er vindt een faseovergang plaats van een fase waar quarks en gluonen “opgesloten” zitten in hadronen naar een fase waar ze zich kunnen voortbewegen over afstanden die veel groter zijn dan de grootte van een proton of neutron. Deze fase het *het Quark-Gluon Plasma (QGP)*.

De temperaturen die nodig zijn om een QGP te vormen zijn dusdanig hoog dat ze niet voorkomen op de aarde of zelfs in ons zonnestelsel. Een manier om in het laboratorium een fysisch systeem te creëren met een voldoende hoge temperatuur, is door zware ionen, bijvoorbeeld lood (Pb) of goud (Au), met extreem hoge kinetische energieën op elkaar te laten botsen. Het QGP wat in zo'n botsing gecreëerd wordt bestaat maar voor een zeer korte tijd, omdat het direct na de botsing uitdijt en afkoelt. Als het QGP voldoende is afgekoeld vindt er weer een faseovergang plaats, waarna de materie weer uit hadronen bestaat. Dit proces wordt ook wel *hadronisatie* genoemd. De eigenschappen van de geproduceerde hadronen (type, impuls, energie) kunnen bepaald worden door middel van een deeltjesdetector, en observabelen gebaseerd op deze eigenschappen kunnen vergeleken worden met modelvoorspellingen of met observaties die zijn gedaan bij simpelere botsingssystemen, zoals proton-proton botsingen. Met behulp van deze vergelijkingen kunnen we de eigenschappen van het QGP bepalen.

Het is nog niet bekend hoe de hadronisatie van het afkoelende QGP precies in zijn werk gaat. In simpelere botsingssystemen, zoals proton-proton botsingen, worden relatief weinig quarks en gluonen geproduceerd, en hadronisatie vindt plaats via *fragmentatie*. De quarks en gluonen stralen andere quarks en gluonen af, waardoor *showers* ontstaan. Deze showers hadroniseren in een gecollimeerde spray van hadronen, wat ook wel een *jet* wordt genoemd. Een belangrijke eigenschap van zo'n jet is dat verschillende eigenschappen van het originele quark of gluon (energie, impuls) behouden blijft in de jet. In een QGP is de dichtheid aan quarks en gluonen veel groter dan in een proton-proton botsing. Het is daarom plausibel dat de hadronisatie ook kan gebeuren door het samengaan van gethermaliseerde quarks (recombinatie), of door een combinatie van recombinatie en fragmentatie.

Metingen van de hadron productieratios, laten zien dat met name de  $p/\pi$  ratio in centrale Pb–Pb botsingen veel groter is dan in pp botsingen, in

het transversale impuls<sup>4</sup> interval  $1 < p_T < 8\text{GeV}/c$ . De  $p/\pi$  ratio in Pb–Pb botsingen bleek consistent met verschillende recombinitie modellen.

In dit proefschrift bouwen we voort op deze resultaten, door de hadron productieratios te bepalen in de jet piek en in de achtergrond van de deeltjesproductie geassocieerd aan een hoog- $p_T$  triggerdeeltje. Dit wil zeggen dat we voor elke botsing alle deeltjes met een groot transversaal impuls ( $p_T > 4\text{ GeV}/c$ ) verzamelen en een distributie creëren van alle deeltjes uit dezelfde botsing met  $1 < p_T < 4\text{ GeV}/c$ , als functie van de azimut  $\Delta\phi$  en de pseudorapiditeit  $\Delta\eta$  ten opzicht van het triggerdeeltje<sup>5</sup>. Omdat het triggerdeeltje waarschijnlijk onderdeel was van een jet, bevat de geassocieerde deeltjesdistributie een gepiekte structuur in de oorsprong, waar de productie groter is dan in de rest van de distributie. Wij hebben zowel de  $p/\pi$  en  $K/\pi$  ratio gemeten in zowel deze “jet piek” als in de achtergrond, dat wil zeggen, bij grote ( $\Delta\eta, \Delta\phi$ ). Deze meting hebben we uitgevoerd voor pp, centrale p-Pb en centrale Pb–Pb botsingen.

We hebben verschillende technieken ontwikkeld om deze meting mogelijk te maken. De identificatie van de deeltjes hebben we gedaan door gebruik te maken van twee verschillende detectorsystemen, de Time-of-Flight (TOF) en de Time Projection Chamber (TPC). De TOF detector bevindt zich ver van het botsingspunt, en het meet  $\Delta t$ , de reistijd die het deeltje van het botsingspunt naar de TOF detector. De TPC is een met gas gevuld cilindervormig vat, en het meet  $dE/dx$ , de energie die een geladen deeltje verliest per meter, als het door het gas heen vliegt. Voorgaande metingen hebben gebruik gemaakt van fits aan het TOF of TPC signaal om deeltjes te identificeren, wij hebben echter gebruik gemaakt van een fit aan het gecombineerde signaal, zodat deeltjesidentificatie tot op hogere  $p_T$  mogelijk werd. De fitfunctie die we hebben gebruikt bevat parameterisaties voor pionen, kaonen en protonen, en heeft 19 vrije parameters.

Om de geassocieerde deeltjesproductie te identificeren hebben we een 4-dimensionale matrix ( $\Delta\eta, \Delta\phi, \Delta t, dE/dx$ ) gecreëerd, en voor elke losse ( $\Delta\eta, \Delta\phi$ ) bin hebben we de hoeveelheid pionen, kaonen en protonen bepaald door een fit uit te voeren op het bijbehorende ( $\Delta t, dE/dx$ ) signaal. Omdat de detectorrespons (verwachtingswaarde, resolutie) afhangt van  $p_T$  en  $\eta$  is de respons in een ( $\Delta\eta, \Delta\phi$ ) bin “uitgesmeerd”. Om dit op te lossen hebben we fit-templates gemaakt van de verwachte vorm van de respons, gebaseerd op de resultaten van de volledige 19-dimensionale fit als functie van  $p_T$  en

<sup>4</sup>Het transversale impuls  $p_T$  is de component loodrecht op de botsingsas.

<sup>5</sup>Net zoals  $\Delta\phi$  kan  $\Delta\eta$  gezien worden als een maat voor de hoekafstand tot het triggerdeeltje.

$\eta$ . Een bijkomend voordeel van de fit-templates is dat de hoeveelheid vrije parameters in de fit tot vier kan worden teruggebracht.

Nadat we de geïdentificeerde geassocieerde deeltjesproductie hebben gemeten kunnen we de productie in de jet piek meten door de achtergrond af te trekken. Het achtergrondsniveau is bepaald in  $(\Delta\eta, \Delta\phi)$  groter dan ongeveer drie maal de verwachte jet piek breedte, bepaald door Gaussische fits uit te voeren op de  $\Delta\phi$  en  $\Delta\eta$  projecties van de geassocieerde deeltjesproductie. In pp botsingen is de achtergrond vlak, maar in p-Pb en Pb-Pb botsingen is er een afhankelijkheid van  $\Delta\phi$  door radiële stroming<sup>6</sup>. Deze radiële stroming is gemeten door het toepassen een discrete cosinus transformatie.

Uit onze metingen blijkt dat de breedte van de jet piek afneemt als functie van  $p_T$ , maar toeneemt als functie van de massa van het deeltje. Dit gedrag wordt niet voorspeld door berekeningen van het PYTHIA model. Ook bleek uit onze metingen dan in pp, p-Pb en Pb-Pb botsingen, zowel de  $p/\pi$  als de  $K/\pi$  ratio groter is in de achtergrond dan in de jet piek. Het grootste verschil is te zien in de  $p/\pi$  ratio in Pb-Pb botsingen. Dit suggereert dat in de jet piek van Pb-Pb botsingen fragmentatie een belangrijke rol speelt in de deeltjesproductie. Verder vonden we ook een afhankelijkheid van de systeem grootte op de deeltjesproductieratios in de jet pieken. Dit doet ons vermoeden dat het fragmentatie proces wel beïnvloed wordt door de aanwezigheid van het QGP.

---

<sup>6</sup>Dit is een gevolg van het vloeistofachtige gedrag van het QGP.

# Acknowledgements

Immediately after completing my master's degree in theoretical physics I started as a PhD student in the particle physics group in Utrecht. At that time there was a lot of excitement in the community, as the LHC was just starting up again and the first collisions at unprecedented energy were recorded.

Equipped with only some theoretical knowledge of the standard model, I was eager to learn how to actually interpret the experimental data that helps shape and improve our understanding of physics at the subatomic level. In all honesty, I was not very sure what the research would be like, and, as I quickly found out, the theoretical knowledge that I had was of relatively limited use. Fortunately, the subatomic physics group in Utrecht has a very friendly and open atmosphere, where people are always happy to help one another.

Especially my conversations with Marco have been invaluable for me throughout my PhD. To be a good particle physicist one has to be versed in a variety of relatively disjoint disciplines: software engineering, statistics, large-scale data analysis, detector hardware, theoretical particle physics, and, certainly not the least important: a solid grasp of the politics of a large international physics experiment. Most particle physicists that I have met put their main focus on a subset of the above list, however, Marco, somehow you have managed to excel across the board. Thank you for your patience and helpful advice on all aspects of my research.

Thomas, while relatively infrequent, the discussions that we had about my work provided a very welcome second opinion, identifying potential problems with the analysis that I had overlooked.

During my PhD I was the TA for the Subatomic Physics course taught by Raimond, for three consecutive years. It was a lot of fun; especially the CERN trips with the students were very memorable. Moreover, this course provided me with a lot of background knowledge of particle physics,

which I found to be very helpful for my general understanding of the field. Gradually, I developed some ideas on how aspects of the course could be modified. Raimond, you always were open to these ideas and you put a lot of faith in me by letting me find my way in implementing them. Thanks!

Thanks also to my successive officemates, Marek, You, and Redmer. You have all helped me on many occasions with practical problems that I ran into during my work. Arne, you were my loyal BND school roommate and I couldn't have wished for a better one. Thanks for the cold beer. Sarah, you joined the group just a few months after I started and I had a lot of fun hanging out with you. Thanks for the captain.

There are many more direct or indirect colleagues that I would like to thank for making my PhD a very pleasant time: Marta, Sandro, Andrea, Davide, Jacopo, Alberto, Emilia, Alessandro, Alis, Naghmeh, Deepa, Hans, Hongyan, Astrid, Panos, Andre, Philip, Leonardo, Jan Fiete, Darius, Annelies, Jasper, Chiara, Marco v. W., Daniel M., and Tim.

My parents, Jan and Ada, and my brother Niels, I would like to thank you for your support and the interest that you had in everything I did. Thanks also for the long strolls over de Hoge Veluwe and for always having an open door for me and my friends. Niels, thanks also for stopping by my office so regularly. To my friends, Ori, Jan, Mechteld, Daniel F., Nynke, Andrea, Antonio, and Floris, thanks for everything and for being great people. Special thanks to Sander for designing the artwork of this thesis, you did a marvelous job!

And finally my wonderful girlfriend Tara, meeting you has been the single most important thing that happened to me. I'm very grateful for all the wonderful times that we've had together, the traveling we did, the meals we cooked, the Star Trek we watched, and the English idioms you taught me. Without you, I'd be way up a fairly unhygienic creek. Thanks also to the Drwenski's for all the wonderful Christmas celebrations and for showing me the USA from an insider's perspective. Tara, thanks for moving to the Netherlands so that we could live together, and for adopting our cat Spinor, who's always up for belly rubs, even if it's the middle of the night.

# Curriculum Vitae

I was born on May 31st, 1985 in Zevenaar, The Netherlands. I attended high school in Duiven, and continued studying Applied Physics at the University of Twente in Enschede. After receiving my bachelor's degree in 2007, I started the master's program in Theoretical Physics at the University of Utrecht. In 2010 I obtained my master's degree in Theoretical Physics (*cum laude*), and started a PhD in Particle Physics, also at the University of Utrecht. This work is the result of the research that I did during my PhD. As of March 2015 I started working as a scientific software engineer and data scientist at VORtech, Delft.