Nuclear Modification of charged-particles production at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV in ALICE

Master thesis

Institut für Kernphysik Frankfurt

by **Youssef El Mard Bouziani**

Fachbereich Physik
der Goethe-Universität
Frankfurt am Main

April 2021

Erstgutachter: Prof. Dr. Henner Büsching
Zweitgutachter: Prof. Dr. Harald Appelshäuser

Contents

0 Abstract

1 Introduction

1.1 The Structure of Matter

Throughout the twentieth century, successive crucial advances in the field of particle physics led to a revolution of the world-view on the structure of matter and the fundamental forces that govern the universe. Following these developments, it became necessary to consolidate the acquired knowledge into a single theory: the Standard Model, which marked by the end of the 1970s the foundation of modern particle physics. This systematization provided from a theoretical point of view subsequently confirmed experimentally a comprehensive description of the nature of elementary particles and the fundamental interactions they mediate: the electromagnetic, the weak and the strong interaction.

All matter in the universe is composed of indivisible particles that lack spatial extension and substructure, the so-called elementary particles. The first particles to be proposed as elementary were the atoms. This picture changed in the early 20th century after it was proved that atoms are made up of a nucleus constituted by protons and neutrons, around which electrons orbit. In the subsequent decades, an extensive number of other subatomic particles were discovered, raising the question among physicists of whether all of them were in fact structureless. The Standard Model sorted out this confusion, historically known as particle zoo, describing the existence of a low number of elementary particles subdivided into two types named quarks and leptons that build all known particles, including protons and neutrons. In addition, the Standard Model clarified that the interactions between quarks and leptons occur by means of an exchange of force-carrier particles, also considered elementary.

In order to classify the elementary particles, it is required to consider the immutable and distinguishable physical properties each of them possess. In this regard, the most fundamental distinction between the building blocks of matter and the force-carrier particles lies in the intrinsic angular momentum, known generally as spin, all particles carry, whether or not they are elementary. Particles of half-integer spin in units of the Planck's constant $\hbar$ are named fermions

and present an antisymmetric wave function, whereas those of integer spin are known as bosons and their wave function describes a symmetric behavior. In this context, quarks and leptons are considered elementary fermions since their spin is equal to $\pm 1/2$, while the force-carriers, which have a spin 1, are referred to as gauge bosons.

The Standard Model states the existence of six quarks and six leptons, summarized in Figure 1.1. Both groups of elementary fermions are arranged according to the mass in three doublets named generations. Each pair of quarks or leptons have a mass one or two orders of magnitude lower than the pair of the next generation. Furthermore, the electrical charge is another feature used to sort elementary fermions. The triplets composed of elementary fermions from different generations have the same electrical charge, namely $+2/3$ and $-1/3$ for the case of the quarks and 0 and -1 for the one of the leptons. The Standard Model predicts also the occurrence in Nature of antiquarks and antileptons, elementary fermions with identical mass as their counterparts but opposite electrical charge and spin.

The six species or flavours of quarks known at the present day are up u, down d, charm c, strange s), top t and bottom b. At this point, it should be emphasized that quarks could not yet be observed existing on their own, but only building bound states named hadrons. These composite particles, detected for the first time during the era of the particle zoo, come in two types: hadrons built from three different quarks are called baryons, while those constituted of a quark-antiquark pair mesons. For instance, the quark composition of protons and neutrons, respectively *uud* and *udd*, allows to identify them as baryons. All known hadrons, excluding the proton, decay after a given period of time. It should be also noted that since the spin is an additive property, baryons are considered fermionic particles and mesons, by contrast, bosonic particles.

The quark composition of a hadron determines intrinsic facets of the particle, which are associated with additive quantities named quark quantum numbers: the well-known electrical charge, a parametrization of the number of quarks called baryon number B and other parametrization of the second- and third-generation flavours. However, the valence quarks, i.e. the quarks that determine the identity of a hadron, aren't the only components. In fact, a hadron should be understood as a more complex and dynamic structure that contains, aside from the valence quarks, an amalgam of quark-antiquark pairs named sea quarks in a constant process of creation and annihilation as well as a large number of gluons, the force-carrier of the strong interaction, which will be discussed later. This picture is consistent with the fact that the sum of the masses of the valence quarks of a given hadron doesn't coincide with the observed mass of the hadron,

Figure 1.1: (Das Bild ist nicht fertig!) Elementary particles of the Standard Model.

being this significantly larger.

As already mentioned, six flavours of leptons are known at this time: three electrically charged leptons, the electron e, the muon μ and the tau τ, and the corresponding electrically neutral leptons that complete each generation, the electron neutrino ν_e, the muon neutrino ν_μ and the tau neutrino ν_τ. In contrast to quarks, leptons don't arrange themselves to form composite particles, with the big exception of the electron shell of an atom, and can exist on their own.

According to our actual understanding, all observed phenomenon at subatomic level, such as the formation of composite particles or the particle decays, manifest themselves by means of the existence of three fundamental interactions: the electromagnetic, the weak and the strong interaction. In modern particle physics, each of these interactions is explained as an exchange of quanta of elementary bosons between the elementary particles that exert the force. In order to experience a given fundamental force, an elementary particle must carry the corresponding charge. Particles are charged in different magnitudes, which determines the intensity with they create or undergo the force.

The force-carrier particle of the electromagnetic interaction is named photon γ, a massless boson that mediates the attraction and repulsion forces between all electrically charged particles. In the case of the weak interaction, which all elementary fermions can experience, there are three massive mediators: one electrically neutral, namely the Z boson, and two electrically charged, the $W^\pm$ bosons. Finally, the strong interaction is mediated by the already introduced gluon g.

The charge of the strong interaction is called color, but it should not be mistaken for the concept generally known and described by the optical physics. Quarks are the only elementary fermions that have color charge and therefore gluons can couple to them. In addition to quarks, gluons carry also color what allows them to couple to themselves, a particularity only present in the strong interaction. The basic processes of the strong interaction are described by a theory named quantum chromodynamics (QCD). Since some of these processes represent a key knowledge for the understanding of this thesis, this theory will be treated in the coming section more exhaustive.

1.2 Quantum Chromodynamics

1.3 Phase diagram and Quark-Gluon-Plasma

1.4 Physics of particle collisions (pp, A-A col.)

1.5 (charged particle) p_T spectra and nuclear modification factor

In order to understand R_{AA} explain pseudorapidity

1.6 Monte Carlo simulations

Comparison of spectra with HIJING models

2 The ALICE Experiment at CERN

2.1 The LHC

2.2 ALICE

2.2.1 ITS

2.2.2 TPC

explain clusters

2.2.3 Track reconstruction

Explain track variables

2.2.4 V0 detectors

Centrality determination

3 Analysis

The Quark-Gluon plasma has remained since its theoretical proposal forty years ago as a subject of huge interest in the high-energy physics given that it could offer answers to fundamental questions about the structure of matter. An example of the significance of this research field is the ALICE experiment at the LHC, which is dedicated to the study of this state of matter that is believed to occur in ultra-relativistic particle collisions.

As exposed in the first chapter, the physics of strongly interacting matter can be investigated by means of the nuclear modification factor R_{AA} which compares the yield of inclusive charged particles in AA collisions and a corresponding reference measurement in pp collisions. In this chapter, an analysis of transverse momentum distributions in Pb-Pb and in pp collisions at a center-of-mass energy $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV measured with ALICE is presented. The analysis includes the approach used for the data selection as well as the corrections required to reduce the detector effects that have an impact on the p_{T} distributions. Once the corrected p_{T} distributions are determined and following the calculation of the nuclear modification factors, the results will be discussed based on a comparison with a previous analogous ALICE publication (cite paper here).

3.1 Data sample

This thesis will make use of pp and Pb-Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ TeV recorded by ALICE. The particle collisions took place in the years 2017 and 2018 respectively. In the pp data sets, the interaction rate was 50 kHz, whereas in the Pb-Pb ones, 7.5 kHz, it was much lower. Furthermore, for each data period there is at least one MC production anchored to which simulates the exact same conditions and technical specifications of the detector response at the time of the data taking. The events of the MC productions were generated by PYTHIA in the case of pp collisions and by HIJING as regards the Pb-Pb collisions while the corresponding detector response was simulated by GEANT3. The limited computing time available for these simulations allows to cover only a percentage of the total statistics. This information as well as an overview of the number of

System	statistics (M)	MC generator	MC statistics
pp	560 (FAST) 318 (CENT)	PYTHIA + GEANT3	25%
Pb-Pb	110	HIJING + GEANT3	2%

Table 3.1: Overview of the used data periods, the approximate number of events contained in them and the MC productions they are associated with including the percentage of simulated statistics.

events contained in each data set are summarized in Table 3.1. The number of events was measured after the event selection depicted in the following section of the chapter.

The examined pp data taking contain two separated data subsets that differ according to the readout configuration: the first data subset corresponds to collisions measured excluding the information of the SDD layers, which in this thesis will be referred to as *FAST*. The information is excluded whenever the SDD is inoperable at the time of a collision due to the dead time of this detector. In contrast, the collisions in the second data subset were reconstructed including the information of the SDD layers. These collisions are labeled as *CENT*. Furthermore, the events of the *CENT* subset were reconstructed also without the SDD information allowing to study the impact of the absence of the SDD layers during the tracking. To this purpose, a comparison of the same observable extracted from *CENT* with SDD and from *CENT* without SDD can be undertaken. As shown in Table 3.1, the *FAST* subset contains almost three times as many events as *CENT*. In order to be able to use all the available data, the equivalence of the *FAST* and the *CENT* without SDD subsets will be demonstrated and, therefore, the feasibility of combining them. In the big picture, this increase of the statistics should lead to an improvement of the precision of the charged-particle p_T spectra in pp collisions, specially in the high p_T region, which in turn could also enhance the precision of the R_{AA} distributions. In this thesis, to determine the viability of this merging, the final corrected p_T spectra obtained from the *CENT* with and without SDD subsets will be compared at the end of the chapter.
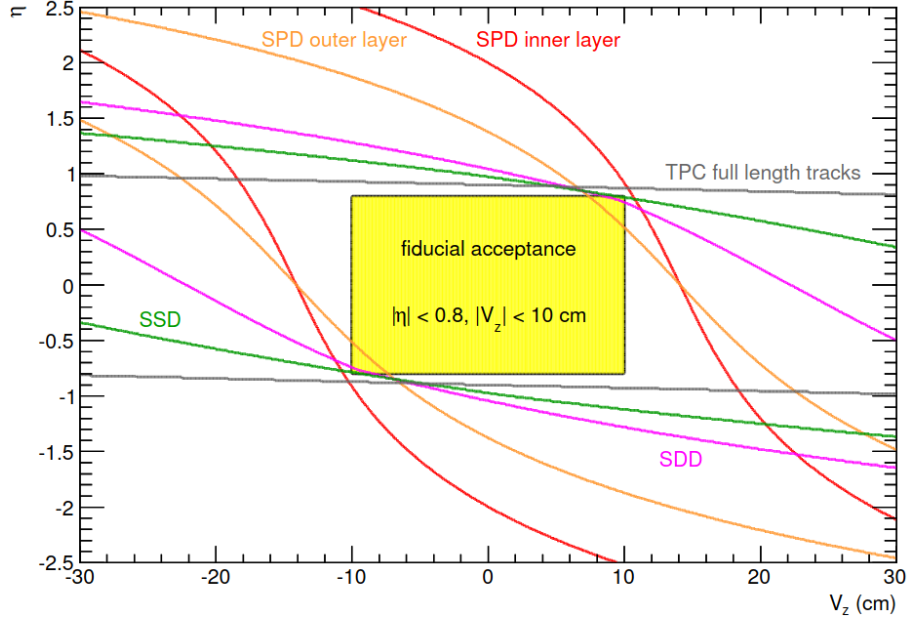
Figure 3.1: Pseudorapidity acceptance of the SPD, the SDD, the SSD and the TPC shown as function of the position of the primary vertex along the *z*-axis. The fiducial acceptance represents the region where most tracks can be reconstructed (cite mknichel).

3.2 Trigger and event selection

3.2.1 Trigger

All studied events must meet quality criteria in a selection procedure called trigger. During the data taking, the trigger makes a decision in real time based on physical requirements whether a collision candidate is worth to be stored or not. The ALICE experiment focuses its research activities on the investigation of strongly interacting matter an inelastic collisions are of particular importance. The trigger decision is also motivated by the large detector readout time and a limited data storage capacity which represent obstacles that prevent the recording of all events that take place.

The trigger configuration which imposes the smallest bias and thus accepts the majority of the overall inelastic cross section is called minimum bias (MB). For the data sets used in this thesis, a MB trigger that requires signals in the V0A and V0C detectors in coincidence with the bunch crossing (MB$_{and}$) was established as it offers more advantages for collision systems with large luminosities (cite "Performance of ALICE", s. 22).

3.2.2 Selection of events

In an offline stage of the event selection, collision candidates that satisfy the trigger condition undergo a selection process in order to reject the contamination caused by poor quality events. A part of this contamination corresponds to background events, a particle-flux caused by the interaction between the beams and the matter in the machine. The beam background is discriminated mainly by means of the signal arrival time of the events in the V0 modules. The arrival time of background events is shorter than the one of a collision produced in the nominal interaction point. This difference is exploited to exclude background events from the data sample. Another source of contamination are the so-called pile-up events, multiple collisions which occur within the same bunch crossing. In addition, events produced in a bunch crossing different from the one that triggered the data recording are also considered pile-up. An event is identified as pile-up and thereby removed from the data sample when multiple primary vertices are reconstructed using the information from the SPD in a single event. The measured charged-particle production was required to have a pseudorapidity within a range $|\eta| < 0.8$. The pseudorapidity acceptance exhibits a dependence on the position of the primary vertex V_z along the beam axis z as illustrated in Figure 3.1 for the ITS subdetectors and the TPC. Here, it is observed that the influence of this correlation on the detectors increases with the radius of these and hence the TPC is the detector least affected. In the middle of the plane, marked in yellow, a region in which most of the tracks are covered arises. Therefore, only events with a position of the primary vertices along the beam axis of $|V_z| < 10$ cm is selected.

3.3 Track selection

Primary charged particles traverse the TPC-ITS detector system and leave a track that can be reconstructed by means of the procedure described in (cite section). The resulting tracks must fulfill several requirements which aim for a selection of good resolution tracks as well as a reduction of the contamination from secondary particles present in the data sample. To this purpose, predefined track selection criteria or track cuts are used in the analysis. These correspond to variations of track variables related to the quality of the tracks. In addition, the cuts on the *DCA* are implemented to eliminate the contamination by secondary particles.

Over the years, several analyses on the charged-particle production in ALICE, such as (cite theses), have refined the variations and, as result, the standard track

Track variable	Track cut
Selection of primaries	
DCA_z	≤ 2 cm
DCA_{xy}	$\leq 7\sigma$
ITS selection	
at least one hit in the SPD	required
ITS refit	required
χ^2 per ITS cluster	≤ 36
TPC selection	
TPC refit	required
χ^2 per TPC cluster (pp collisions)	≤ 4
χ^2 per TPC cluster (Pb-Pb collisions)	≤ 2.5
fraction of shared TPC clusters	≤ 0.4
ratio of crossed rows over findable clusters	≥ 0.8
geometric length in the active TPC area	$130, 3, 1.5, 0.85, 0.7$
TPC-ITS selection	
χ^2 TPC constrained track vs. global track	≤ 36

Table 3.2: Standard track selection criteria used for analysis of primary charged particles in ALICE.

selection listed in Table 3.2 has been established. The track cuts are implemented in both collision systems pp and Pb-Pb for tracks measured in the kinematic range $p_T > 0.15$ GeV/c and $|\eta| < 0.8$, in data and in the corresponding MC productions.

3.4 Corrections

After the implementation of the described event and track selection, the uncorrected p_T dependent distributions of the produced charged particles in pp and Pb-Pb are given in Figure 3.2. In the case of heavy-ion collisions, the p_T distributions are divided into nine centrality classes. The vertical error bars represent the statistical uncertainties, which increase with the transverse momentum since
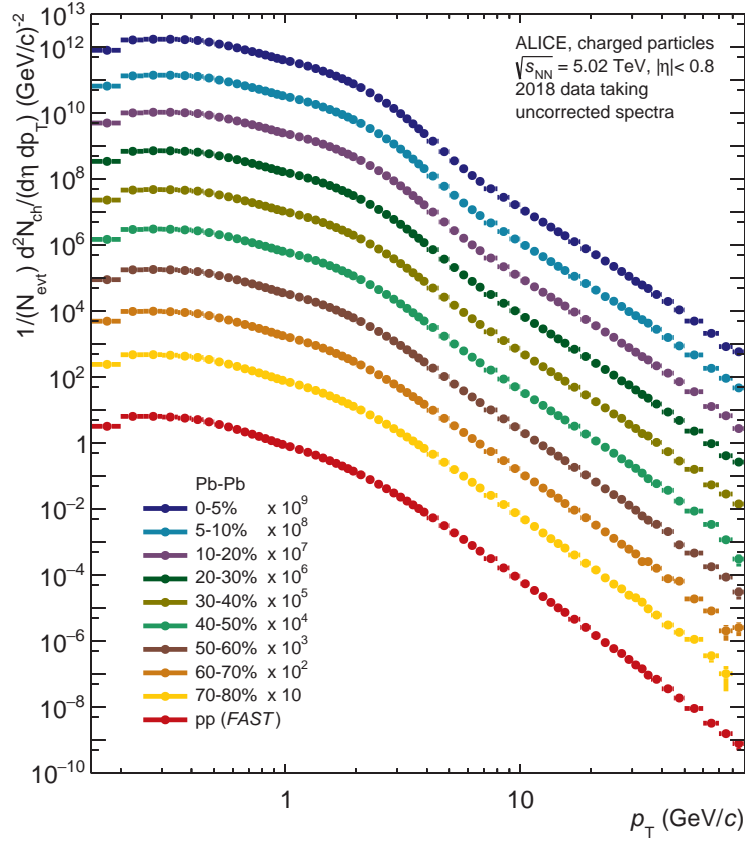
Figure 3.2: Uncorrected transverse momentum distributions of primary charged particles produced in pp and Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV. The latter are divided into nine centrality classes. The p_T spectra of Pb-Pb collisions are scaled for a better visibility.

the high p_T regions of the spectra are less populated by tracks.

However, several effects related to the detector response as well as limitations of the above-described selection criteria to reject the contamination by secondaries must be corrected. To this end, most of the corrections are implemented using information provided by MC simulations, while data-driven approaches are also applied for some others.

Two groups of corrections can be defined according to the level at which they operate:

- corrections which affect the tracks and, therefore, posses a p_T dependent behavior. This group consequently conditions the shape of the spectrum.

- corrections which intervene at event-level and, as a result, influence the shift of the spectrum.

In the following sections, the corrections applied on the p_T spectra and the analysis strategies used for their implementation will be discussed.

3.4.1 Overall normalization

Trigger ϵ_{Trig}	Vertex reconstruction ϵ_{Vz}
75.25%	98.48%

Table 3.3: Trigger and event vertex reconstruction efficiencies for pp collisions at $\sqrt{s} = 5.02$ TeV recorded with ALICE in 2017.

The p_{T} distributions are normalized with the total number of inelastic events N_{INEL}, as in Equation (cite formula). This value is approximated by using the number of events $N_{\text{ev}}^{\text{rec}}$ that passed the selection criteria. Nevertheless, some inelastic events could not be measured because a lack of triggering or vertex reconstruction. These two phenomena are corrected by means of the so-called trigger efficiency ϵ_{Trig} and the vertex reconstruction ϵ_{Vz} efficiency:

$$N_{\text{ev}}^{\text{rec}} = N_{\text{INEL}} \cdot \epsilon_{\text{Trig}} \cdot \epsilon_{\text{Vz}} \tag{3.1}$$

In heavy-ion collisions, both trigger and vertex reconstruction efficiency have been observed to be unity in the studied centrality interval 0-80% by previous works through MC simulations (cite mknichel/Pb-Pb papers). However, the bias is particularly notable in pp collisions. The determination of the respective efficiencies in pp collisions, described bellow, is therefore required.

Trigger efficiency

The MB$_{\text{and}}$ condition of the ALICE trigger system is able to measure a significant percentage of the total inelastic cross section, yet this measurement is limited due to the efficiency of the trigger ϵ_{Trig}. This connects the visible and the total inelastic cross section as follows:

$$\epsilon_{\text{Trig}} = \frac{\sigma_{\text{vis}}}{\sigma_{\text{INEL}}} \tag{3.2}$$

For each data period, the visible cross section is measured by determining the luminosity of the experiment, a quantity that evaluates the number of collisions produced in the detector and which is measured in events per time per area. To this purpose, a technique called van-der-Meer scan is used, in which the two beams are displaced with respect to each other in the transverse directions and the rate of interactions is monitored as a function of this displacement. The ratio of the head-on rate to the luminosity corresponds then to the desired visible cross section (cite S. vdM). In the studied pp collisions collected in 2017

at $\sqrt{s} = 5.02$ TeV, the visible cross section amounts to $\sigma_{\text{vis}} = 50.87 \pm 1.07$ mb (cite value).

The total inelastic cross section σ_{INEL} has not been measured in pp collisions at $\sqrt{s} = 5.02$ TeV. However, a Monte Carlo Glauber model is able to predict by means of a data-driven parametrization the value of σ_{INEL} as function of the center-of-mass energy (cite MC Glauber paper). Here, experimental values from several collaborations are utilized as input. According to the prediction, the total inelastic cross section at $\sqrt{s} = 5.02$ TeV is $\sigma_{\text{INEL}} = 67.6 \pm 0.6$ mb (cite value). The resulting trigger efficiency is shown in Table 3.3.

Vertex reconstruction efficiency

The vertex reconstruction efficiency ϵ_{Vz} can be defined as follows:

$$\epsilon_{\text{Vz}} = \frac{N_{\text{Vtx}}}{N_{\text{Trig}}} \tag{3.3}$$

where N_{Vtx} is the number of events for which a vertex was reconstructed and N_{Trig} is the total number of triggered events. Both values were determined from data disregarding the event selection. The resulting efficiency can be found once again in Table 3.3.

Signal loss correction

As already seen, a fraction of the inelastic events is missing due to the trigger and the vertex reconstruction efficiency and it must be thus restored. However, this correction has to be combined with a complementary correction at track-level in order to reestablish the tracks from events wrongly rejected by the event selection. This is achieved using the MC production to compare the distribution of primary tracks generated from inelastic events within the vertex cut $|V_z| < 10$ cm before and after they undergo the event selection:

$$\text{S. L.} = \frac{(\mathrm{d}N/\mathrm{d}p_{\text{T}})_{\text{Generated}}^{\text{After event selection}}}{(\mathrm{d}N/\mathrm{d}p_{\text{T}})_{\text{Generated}}^{\text{True INEL>0, }|V_z|<10\text{ cm}}} \tag{3.4}$$

As result, the p_{T} dependent distribution of the signal loss is shown in Figure 3.3. The effect is more pronounced at low p_{T}, where it amounts to about 3%. The signal loss decreases until becoming negligible above 10 GeV/c. The required correction is implemented by scaling the p_{T} spectrum in each p_{T} interval with the inverse of the corresponding value of the signal loss.
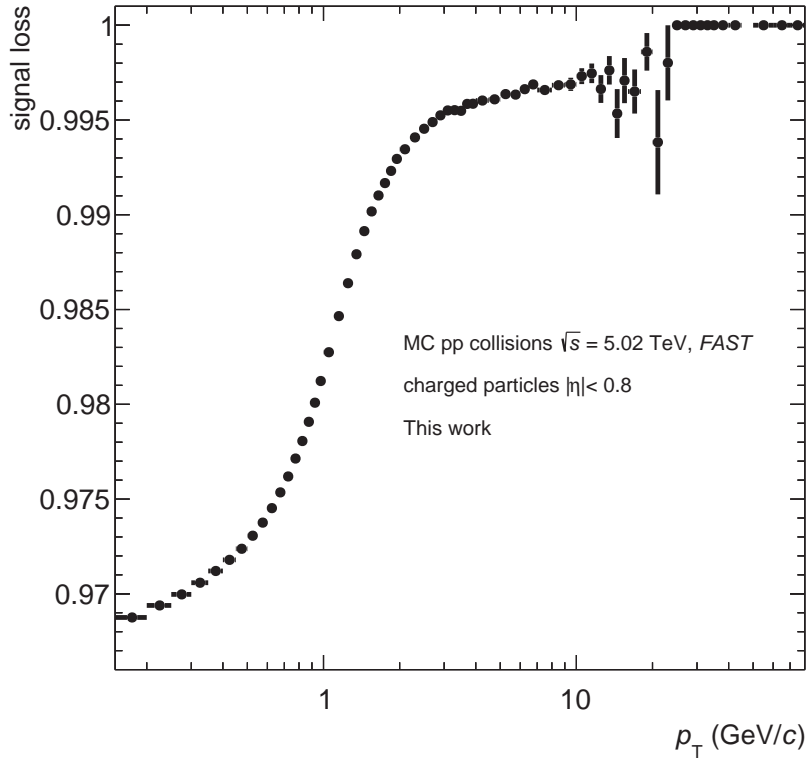
Figure 3.3: Signal loss for pp collisions.

3.4.2 Tracking efficiency and acceptance

The TPC-ITS detector system is able to measure most of the tracks that cross the detector volume. However, the reconstruction algorithm is sometimes unable to detect a track or to reconstruct it correctly which then causes its discard by the track selection. In fact, tracks are measured with a probability which depends on their kinematic properties. Moreover, the acceptance coverage of the detector, defined by the kinematic range of $0 \leq \phi \leq 2\pi$, $|\eta| < 0.8$ and $p_T > 0.15$ GeV/c, affects also the measurement of charged particles.

Both of these two phenomena, the tracking efficiency and the acceptance, can be summarized in an overall efficiency ϵ for the reconstruction of tracks which depends on the kinematic variables: the transverse momentum p_T , the pseudorapidity η, the position of primary vertex along the beam axis V_z and the multiplicity class N_{ch}. The implementation of the corresponding correction will thus affect the shape of the p_T distributions. The p_T dependent efficiency is calculated using the information gathered by the MC production as the ratio of reconstructed primary tracks that survive the track selection N_{rec}^{MC} to generated
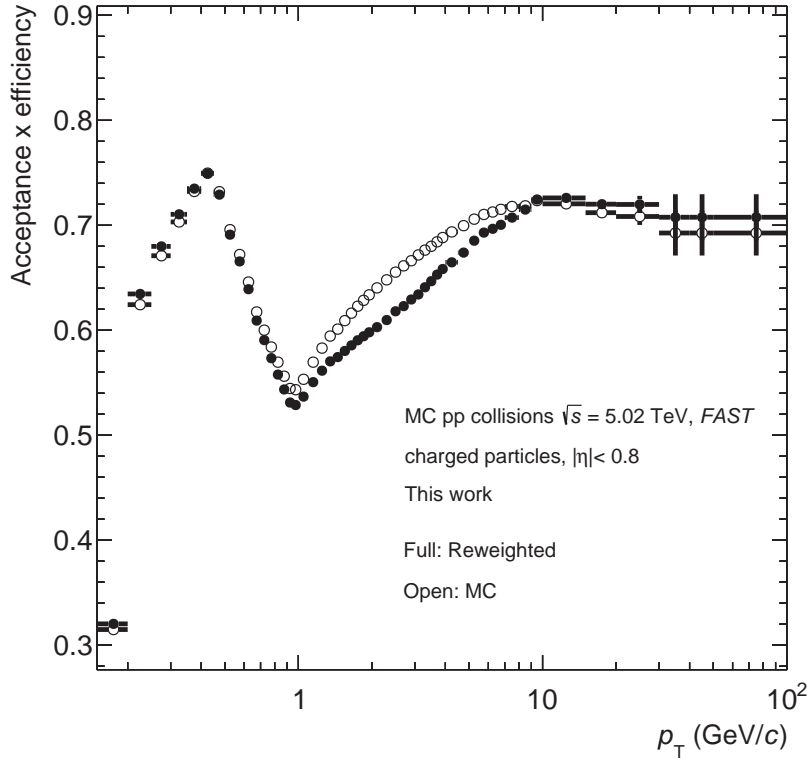
Figure 3.4: Acceptance and tracking efficiency as function of the transverse momentum for pp collisions obtained from MC simulations.

primary charged-particles $N_{\text{gen}}^{\text{MC}}$:

$$\epsilon(p_{\text{T}}) = \frac{N_{\text{prim,rec}}^{\text{MC}}(p_{\text{T}})}{N_{\text{prim,gen}}^{\text{MC}}(p_{\text{T}})} \tag{3.5}$$

For the correction, the inverse of the overall efficiency is applied as a multiplicative factor on the p_{T} spectra.

The overall efficiency is also dependent on the type of particle measured. This leads to the implementation of a complementary correction, later on explained, that scales the overall efficiency by means of the relative abundance of each particle type (cite Patricks thesis).

In Figure 3.4, the distribution of the overall efficiency as function of p_{T} is shown for pp collisions. Here, it can be seen the characteristic p_{T}-shape also observed in all previous studies of the charged-particle production regardless of the energy or the collision system. In the presented thesis, the efficiency amounts to a value constrained between 54% and 75% for the entire p_{T} range. At low p_{T}, below 4 GeV/c, there is a rapid growth of the efficiency which can be explained by the large curvature that the tracks in this p_{T}-range experience due to the in-
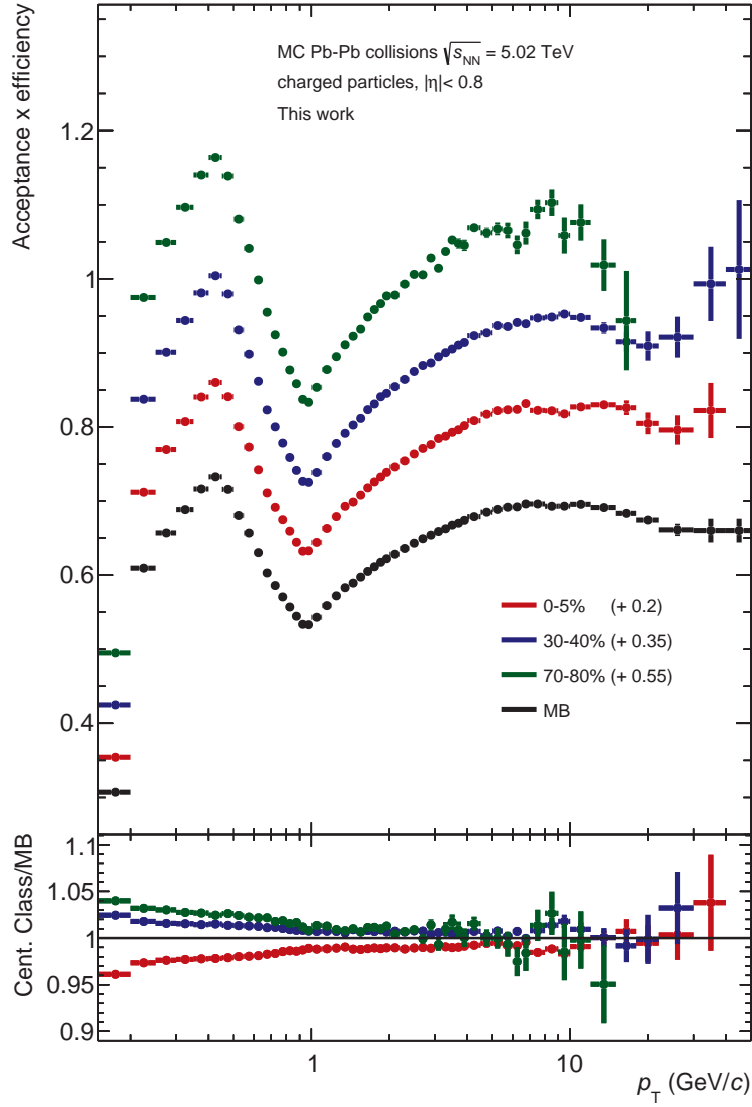
Figure 3.5: Tracking efficiency as function of the transverse momentum for MB, central, semi-central and peripheral Pb-Pb collisions obtained from MC simulations. The efficiencies are scaled with an additive factor for a better overview. In the bottom panel, the ratio of the individual efficiencies to the MB efficiency is shown. Data points with very large statistical deviations are excluded.

fluence of the magnetic field. Such a curvature reduces the uncertainty of the track reconstruction increasing, as result, the overall efficiency. After reaching the maximum, the curve drops quickly and hits its minimum around 1 GeV/c. This slump is caused mainly by the track length cut listed in Table 3.2. The influence of this requirement on the selection process rises within this particular p_T-range given that the corresponding tracks tend to cross the TPC boundaries reducing hence their probability to be selected. Once this effect fades, the distri-

Centrality class	Scaling factor
0-5%	0.980
5-10%	0.995
10-20%	1.000
20-30%	1.004
30-40%	1.007
40-50%	1.009
50-60%	1.010
60-70%	1.011
70-80%	1.010

Table 3.4: Scaling factors obtained from the parametrization of the ratios in Figure 3.5.

bution increases asymptotically due to the acceptance restrictions of the detector up to reaching a plateau. The overall efficiency is assumed to be constant for $p_T \geq 40$ GeV/c in order to reduce statistical fluctuations.

This same behavior is also observed in Pb-Pb collisions analysed in this thesis as illustrated in Figure 3.5, which shows the overall efficiency in dependency of p_T for MB, central ($0 - 5\%$), semi-central ($30 - 40\%$) and peripheral ($70 - 80\%$) collisions. The tracking efficiency exhibits a dependency on the centrality class as it decreases with this quantity. As also observed in this figure, the mid and high p_T ranges of the centrality-dependent efficiencies are dominated by very large statistical fluctuations which would contribute to increase the statistical uncertainties of the later calculated p_T spectra. To avoid that, the ratios of the individual efficiencies to the MB efficiency, illustrated in the bottom panel, are parametrized from a p_T of 1.5 GeV/c with a constant function. The obtained factors, summarized in Table 3.4, are used to scale the MB tracking efficiency. As result, The distributions obtained using this approach are represented in 3.6 with open marks.

Dependency on the particle type

Among the primary charged particles that compound the studied sample, a sharp disparity in terms of the lifetime is noticeable. For instance, the pions ($\pi^{\pm}$), the most frequent particle type produced in particle collisions, can travel in average 7.8 m within their lifetime, whereas the sigma baryons ($\Sigma^{\pm}$) have a
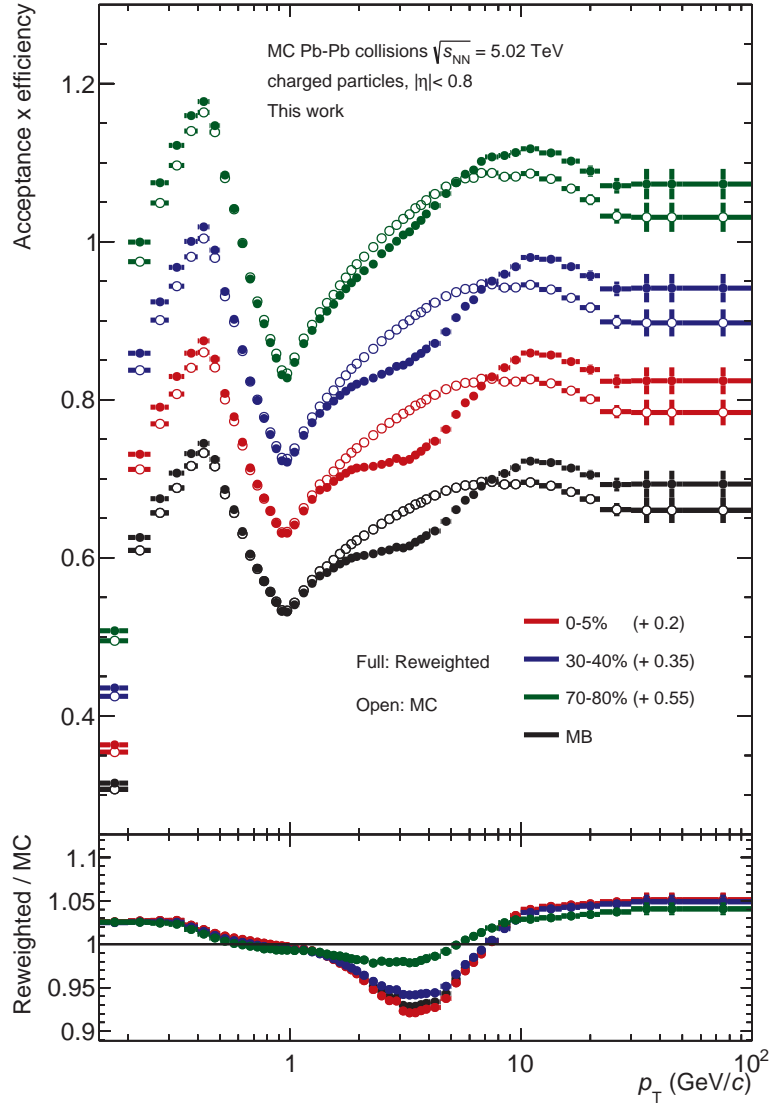
Figure 3.6: Tracking efficiency as function of the transverse momentum for MB, central, semi-central and peripheral Pb-Pb collisions obtained from the scaling of the distributions in Figure 3.5 using the values listed in Table 3.4. The distributions are scaled with an additive factor for a better overview.

decay length of only 2.4 cm (cite pdg booklet). As consequence, some particles are more likely to be rejected by the track length requirement than others, which creates a dependency of the tracking efficiency on the particle type.

The usage of MC event generators, such as PYTHIA or HIJING, allows to obtain the tracking efficiencies for different particle species as shown in the left panel of Figure 3.7. Here, results from an analysis of the particle dependent efficiency for pp collisions at $\sqrt{s} = 5.02$ TeV are presented (cite paper). As seen in the plot, the dependency of the efficiency on the particle type is strong, specially
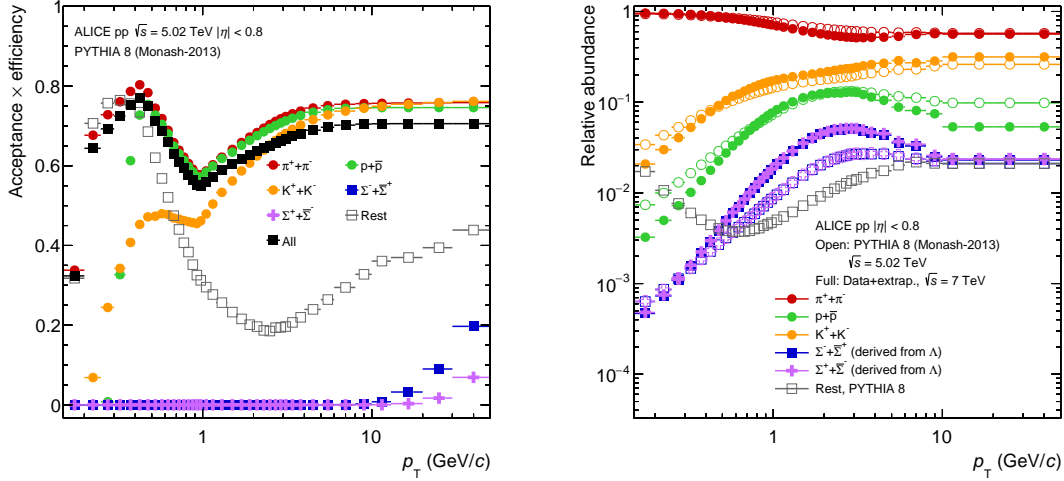
Figure 3.7: **Left:** Tracking efficiencies as function of the transverse momentum for different particle species obtained with a PYTHIA (Monash 13) simulation of pp collisions at $\sqrt{s} = 5.02$ TeV. **Right:** Relative particle abundances in dependency of the transverse momentum in data (full symbols, $\sqrt{s} = 7$ TeV) and in MC (open symbols, $\sqrt{s} = 5.02$ TeV) (cite paper). The MC predictions are shown for comparison purposes.

below 1 GeV/c. In Pb-Pb collisions, similar effects are observed as can be seen in Appendix A.1. Because of these similarities, the correction procedure is only shown for pp collisions, although an analogous procedure is also implemented for Pb-Pb collisions.

Considering the particle type dependency of the efficiency, the abundance of each particle plays an important role and must therefore be reflected in the correction. However, MC event generators have limitations in the description of the particle production, specially in the one regarding strange particles (cite 38,39 paper). As consequence, the efficiency is considerably affected by this underestimation and a correction by means of a reweighting of the efficiency with the relative abundance of each particle type is implemented. The right plot of Figure 3.7 shows the relative abundances measured in pp collisions recorded by ALICE at $\sqrt{s} = 7$ TeV. This center-of-mass energy will be used since no significant energy dependence has been observed experimentally and because of the lack of experimental data at the energy studied in this work (cite Patrick). These distributions were calculated using a data-driven procedure detailed in the cited analysis.

In summary, the overall efficiency for inclusive particles should be understood as a weighted superposition of the individual efficiencies for each particle type with the relative particle abundances. After applying this reweighting on the distributions represented in Figures 3.4 and 3.6 with open symbols, the corrected
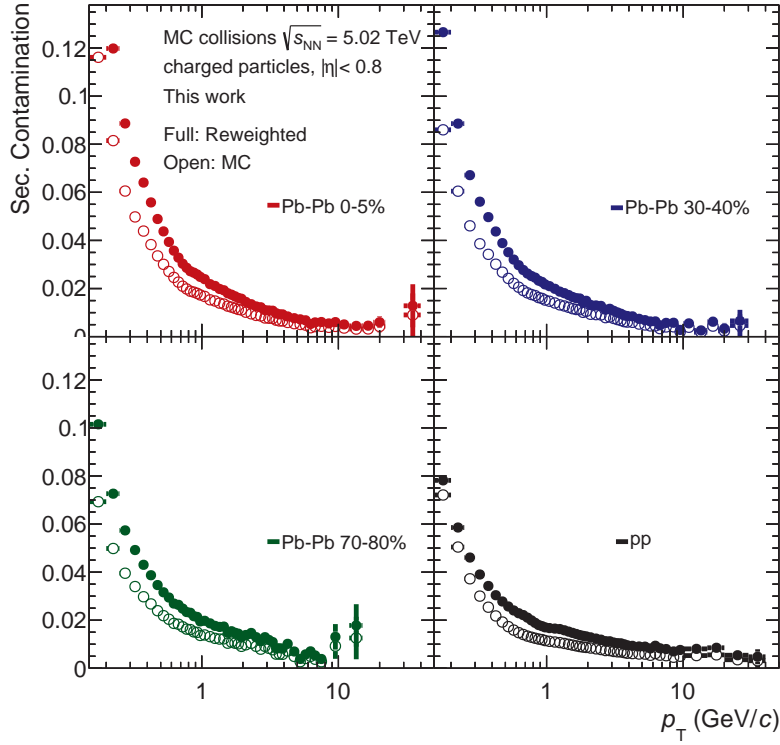
Figure 3.8: Contamination with secondary particles as function of the transverse momentum obtained from MC simulations for pp collisions as well as central, semi-central and peripheral Pb-Pb collisions. Open markers represent the secondary contamination obtained from pure MC, while the full ones correspond to the corrected contamination.

tracking efficiencies for inclusive particle are obtained and shown in these same figures with full markers. These distributions are utilized ultimately to correct the p_T distributions.

3.4.3 Contamination by secondary particles

Despite the selection procedure described in Section 3.3, which aims to reduce the contamination of the measured track sample with secondary particles, a small amount of these particles persists in it. This remnant originates from weak decays of kaons, Λ baryons and muons as well as by interactions in the detector material. The remaining fraction of secondary particles has to be subtracted from the p_T distributions to ensure purity of the data.

The estimation of contamination by secondaries is carried out using the MC productions anchored to the analysed data periods. In Figure 3.8, the result of these estimations can be seen with open markers for pp collisions as well as for MB,
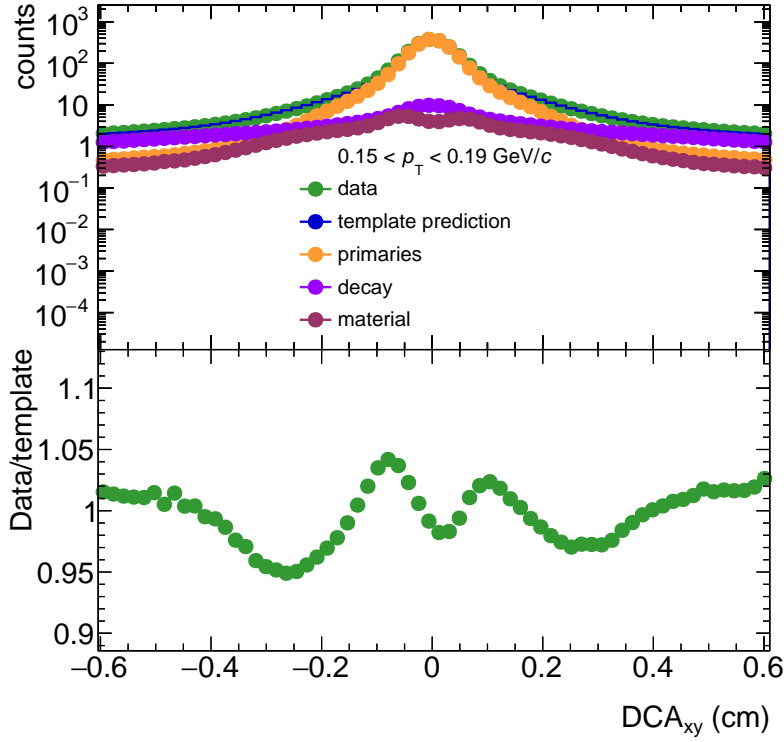
Figure 3.9: DCA_{xy} distributions for data and MC simulations for pp collisions at $\sqrt{}$ 5.02 TeV for the p_T range $0.15 \leq p_T$ 0.19 GeV/c. The MC predictions correspond to primaries, secondaries from decay and secondaries from interactions in the detector material. In the bottom panel, the ratio of distribution for data to the fit resulting from the MC templates is shown.

central $(0 - 5\%)$, semi-central $(30 - 40\%)$ and peripheral $(70 - 80\%)$ Pb-Pb collisions. In general, a contamination of around 10% at low p_T can be observed. The distributions then fall monotonously with increasing p_T up to bellow 1% in the high p_T region. These observations are consistent with the assumption that secondaries are inclined to carry small momenta since they correspond essentially to fractions of the momenta of the mother particles. A dependency of the contamination on the centrality is observed in Pb-Pb collisions. At low p_T, the contamination in central collisions $(0 - 5\%)$ amounts to almost twice as much as the contamination in the most peripheral $(70 - 80\%)$ collisions. This is a result of a suppression of the yield of strange particles observed in systems with low energy densities (cite here). The dependence on the centrality dissipates in the high p_T range.

Secondary scaling

As stated in the previous section, MC productions are affected by an underestimation of the yield of strange particles. This limitation distorts the true contamination by secondaries which is not entirely reflected in Figure 3.8. For this reason, an additional correction that scales the contamination according to the relative abundances of primary and secondary charged particles is required.

The corresponding scaling must therefore focus on a property which distinguishes between primaries and secondaries. In this regard, the distance of closest approach (DCA) offers the possibility for a clear discrimination between these two groups, given that secondaries present larger values than primaries since they origin in a later stage. This implies that the respective DCA_{xy} distributions should describe distinct shapes. The differences should be particular noticeable in the region of the tails, where the largest DCA_{xy} values are located. In this regard, the track selection must be modified in order to obtain a statistically significant result. In particular, the new track selection dispenses with the cuts on the DCA_{xy} and on the $\chi^2_{TPC\text{-}ITS}$, both of which suppress the yield of secondaries. This modification of the track selection leads to an amount of secondaries sufficiently large to extract scaling factors for a few intervals at low p_T in the range $0.15 \leq p_T \leq 1.5 \, GeV/c$.

In the upper panel of Figure 3.9, the DCA_{xy} distribution for pp collisions recorded by ALICE is illustrated for the p_T interval $0.15 \leq p_T \leq 0.2 \, GeV/c$. In the same figure, the corresponding MC predictions of the DCA_{xy} in pp collisions for primaries, secondaries from decays and secondaries originated from interactions in the detector material are represented. Primaries present a shape with a prominent peak at 0 cm that far exceeds the small peak that characterize both distributions for secondaries. The differences are also noticeable in the tails, where secondaries, as expected, show much broader curves. The data distribution clearly resembles the shape for primary particles since these dominate the particle production. These observations can be extended for the rest of p_T intervals and also for the case of Pb-Pb collisions as it can be seen in the corresponding Figures in the Appendix (cite sec. in App.).

In order to correct the underestimation of strangeness, a data-driven approach that fits the data distribution with a linear combination of the MC templates is applied. This method takes into consideration both data and MC statistical uncertainties for a more accurate template prediction (cite TFractionFitter). The goal of the template fit is to predict the fraction of primaries and secondaries in the data sample for each p_T range. The lower panel of Figure 3.9 shows the ratio of the distribution for data to the result of the template fit. The template fits de-
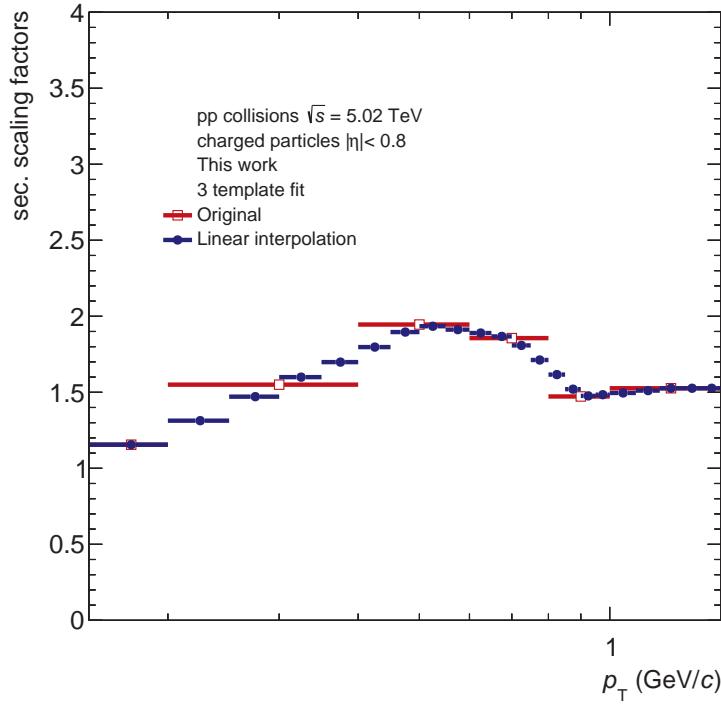
Figure 3.10: Scaling factors used in pp collisions for the correction of the secondary contamination before and after a linear interpolation.

termine the fraction of secondaries in data ρ_{Data} which can be used to calculate the factor c_{sec} that scales the contamination in the given p_{T} with:

$$c_{\text{sec}} = \frac{\rho_{\text{Data}}}{\rho_{\text{MC}}} \tag{3.6}$$

where ρ_{MC} represents the fraction of secondaries in the MC sample. In Figure 3.10, the resulting scaling factors in pp collisions are shown as example. To be able to scale the secondary contamination in all p_{T} intervals, the missing scaling factors are constructed by a linear interpolation as given in this same figure. For $p_{\text{T}} > 1.5$ GeV/c, the scale factor of the last considered p_{T} interval is used since the amount of secondaries varies barely beyond this limit. The corrected contamination by secondaries is shown in Figure 3.8 with full markers.

Moreover, the intervals of the DCA_{xy} distributions as well as the p_{T} intervals for which the fit is performed are adjusted by optimizing the deviation of the data points from the fit as function of the DCA as shown in Figure 3.15. Here, it can be seen that the template prediction is in good agreement with the the data points in the tails, the region of interest in this approach.

To ensure the goodness of the fit, a Pearson's chi-squared test was performed with a significance level of 0.05 (cite Karl Pearson paper). In this approach, the
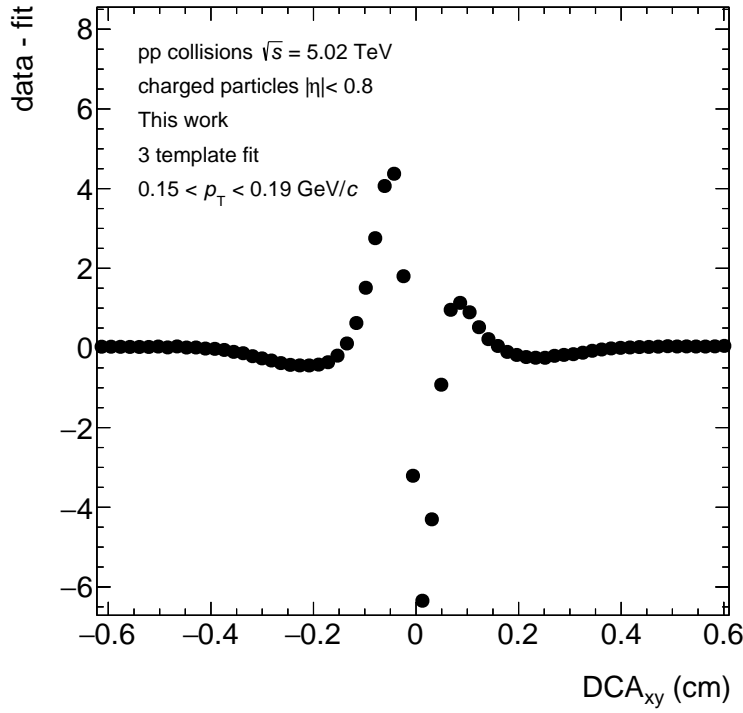
Figure 3.11: Deviation of the template prediction from the data points as function of the DCA in pp collisions for the p_T interval $0.15 \leq p_T \, 0.19 \, \text{GeV}/c$.

test statistic ratio χ^2 is computed:

$$\chi^2 = \sum_i \frac{(\text{DCA}_{\text{data},i} - \text{DCA}_{\text{fit},i})^2}{\text{DCA}_{\text{fit},i}} \tag{3.7}$$

Following this, the critical χ^2 value is calculated by means of the chi-squared probability distribution function determined by the the number of degrees of freedom in the fit and the significance level. According to the chi-squared test, the evaluated fit is rejected whenever the test statistic χ^2 exceeds the critical χ^2 value. All fits used for the calculation of the scaling factors are statistically significant according to this test.

3.4.4 Transverse momentum resolution

As discussed in Section (cite section), the trajectory of a charged particle traveling through the TPC is characterized by a helical form caused by the prevailing magnetic field. This phenomenon is conditioned by the occurrence of the Lorentz force which connects the radius r of the track curvature with the trans-
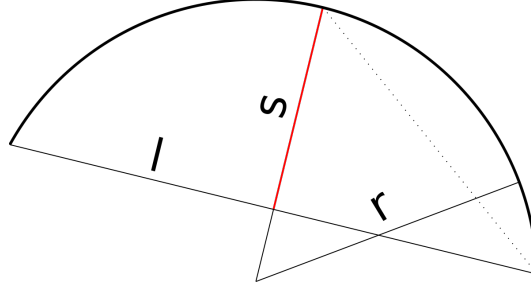
25

Figure 3.12: Visualization of the radius r, the sagitta s and the chord l of a circular arc (cite wikipedia).

verse momentum p_T of the particle:

$$r = \frac{p_T}{e \cdot B} \tag{3.8}$$

where B is the magnetic flux density perpendicular to the particle of charge e. In practice, the measurement of this radius is demanding since even particles with low transverse momenta present radii that far exceed the outer radius of the TPC. For this reason, the radius is substituted by another quantity which parametrizes the track curvature and which can be measured within the detector volume. This is the so-called sagitta, an arc parameter which turns inverse proportional to the radius of the track for the extreme case of a chord $L << r$ spanning the base of the arc (cite here):

$$s = \frac{l^2}{8r} \tag{3.9}$$

Therefore, we obtain for B given in T, L in m and p_T in GeV/c

$$s = \frac{0.3e \cdot B \cdot l^2}{8p_T} \tag{3.10}$$

As stated in Section (cite tracking), tracks are characterized by means of several parameters stored together with the corresponding uncertainties in the so-called covariance matrix. The inverse of the transverse momentum $1/p_T$ is one of the track parameters provided by the covariance matrix.

The sagitta is measured with an uncertainty due to the spatial resolution of the detectors which in turn implies that the resulting inverse of the transverse momentum is affected by an uncertainty $\sigma(1/p_T)$. It can be then shown that the relative transverse momentum resolution $\sigma(p_T)/p_T$ can be approximated by (cite knichel):

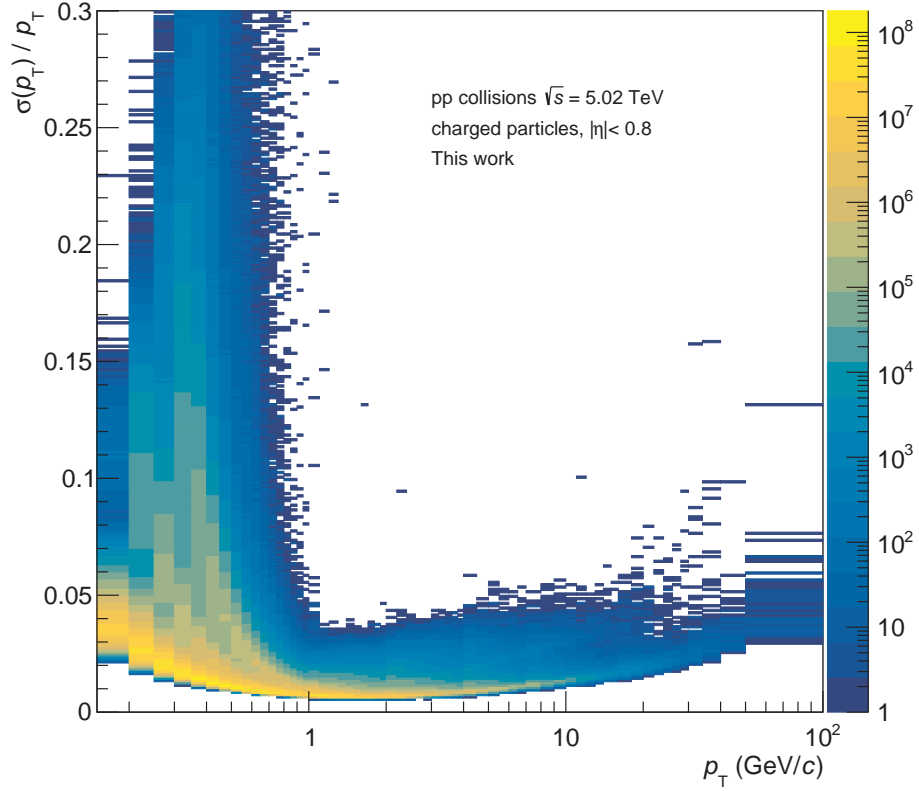$$\frac{\sigma(p_T)}{p_T} \approx p_T \cdot \sigma(1/p_T) \tag{3.11}$$

Figure 3.13: Transverse momentum resolution as function of the transverse momentum for pp collisions at $\sqrt{s} = 5.02$ TeV recorded in ALICE in 2017.

In Figure 3.13, the distribution of $\sigma(p_T)/p_T$ is shown as function of p_T for the analysed pp collisions. For a more clear representation, the mean value $\langle\sigma(p_T)/p_T\rangle$ is calculated in each p_T interval of this distribution as well as for its analogous in MB Pb-Pb collisions. As result, the mean relative p_T resolution as function of p_T is obtained as shown in Figure 3.14. At low p_T, the relative p_T-resolutions are influenced mainly by the multiple scattering of the charged particles in the detector material. In this region, the results in pp and Pb-Pb as well as among the different centrality classes resemble each other. At $p_T = 0.15$ GeV/c, the resolution is on average around 3.7%. It is also worth to point out the three structures present within this p_T region and visible in Figure 3.13. These can be explained by the influence of the multiple scattering on the track parameters. Such an influence is dependent on the particle mass and the observed distinct trends of the resolution correspond, in decreasing order of uncertainty, to protons, kaons and pions.

The effect of the multiple scattering dissipates gradually until the distribution hits the minimum around 1.5 GeV/c with a p_T resolution of approximately 0.7% in pp and 0.8% in Pb-Pb. At high p_T, charged particles experience a less pro-
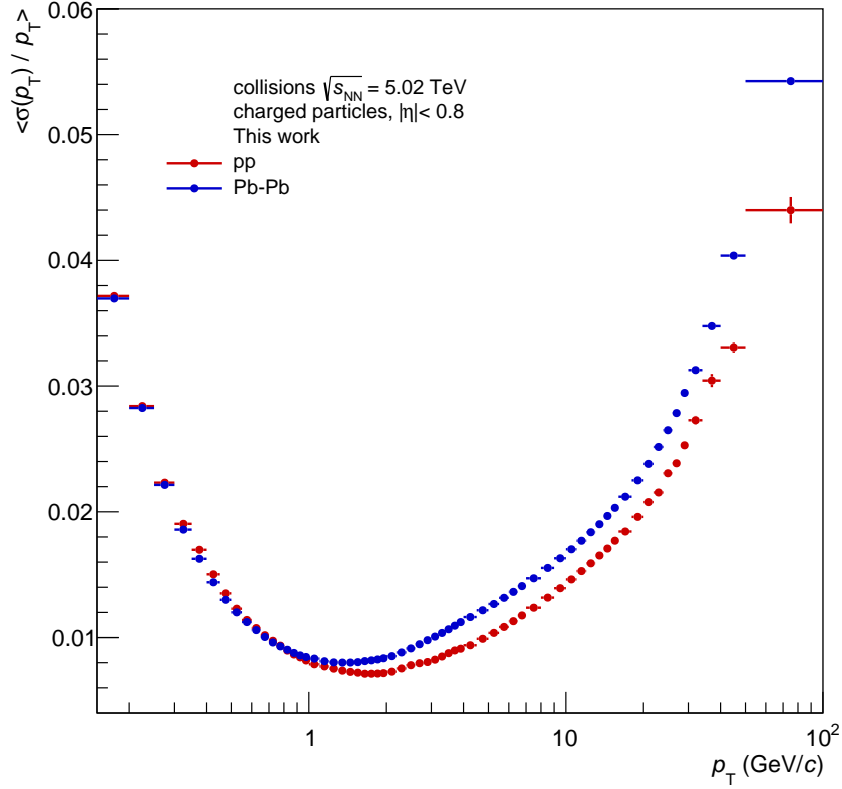
Figure 3.14: The mean relative transverse momentum resolutions are obtained as function of the transverse momentum for pp and MB Pb-Pb collisions.

nounced curvature according to Equation 3.8, which causes a deterioration of the spatial resolution. As consequence, the uncertainty grows linearly reaching values of 4.4% in pp and 5.4% in Pb-Pb.

Since the measurement of p_T is smeared due to the resolution, the p_T spectra must be accordingly corrected. While the smearing has a minor effect on the p_T spectra at low p_T given the smooth slope of the distributions, it becomes more pronounced in the high p_T region, where the p_T spectra present a steeply falling shape. The corresponding correction is thus applied in the range $7 \leq p_T \leq 100$ GeV/c by means of a p_T-dependent factor c_{res} calculated as follows:

$$c_{res}(p_T) = \frac{(dN/dp_T)_{smeared}}{(dN/dp_T)_{true}} \tag{3.12}$$

where the numerator corresponds to the measured p_T distribution, while the denominator is a p_T distribution free of smearing. As starting point, the p_T distributions from the previous ALICE publication, depicted in Section (cite theo. section), are used given that they offer a good approximation of the true p_T distributions. As consequence, correction factors are calculated for pp collisions
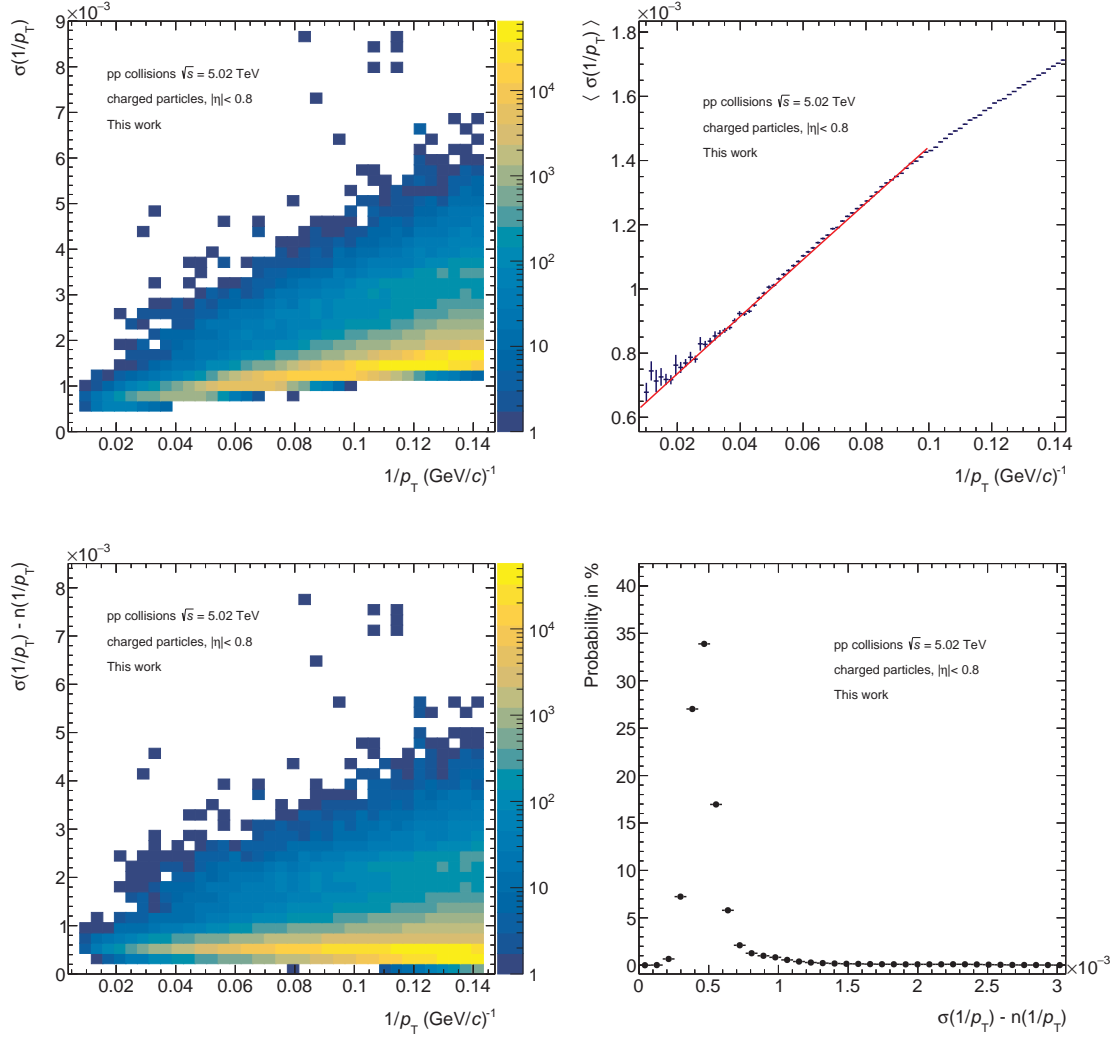
Figure 3.15: **Top.** Left: Two-dimensional distribution of the uncertainty on the inverse of the transverse momentum $\sigma(1/p_{\mathrm{T}})$ as function of $1/p_{\mathrm{T}}$. Right: Quadratic polynomial parametrization $n(1/p_{\mathrm{T}})$ of the mean uncertainty $\langle\sigma(1/p_{\mathrm{T}})\rangle$ as function of $1/p_{\mathrm{T}}$. **Bottom.** Left: Two-dimensional distribution of the scaled uncertainty $\sigma(1/p_{\mathrm{T}})$ as function of $1/p_{\mathrm{T}}$. Right: Probability function of $\sigma(1/p_{\mathrm{T}})$ obtained after projecting the scaled uncertainty into the y-axis.

and for each centrality class in Pb-Pb. For the calculation of $c_{\mathrm{res}}(p_{\mathrm{T}})$, these p_{T} distributions are smeared in such a way as to emulate the effect of the p_{T} resolution. The details of the used approach are outlined in the following.

Calculation of the correction factor

The smearing is carried out through a simulation that uses as input a sample of filtered high p_{T} tracks that underwent the same track requirements summarized in Table 3.2. The parameter $\sigma(1/p_{\mathrm{T}})$ of these tracks is considered as function of $1/p_{\mathrm{T}}$ as shown in the top-left panel of Figure 3.15 for the case of pp colli-
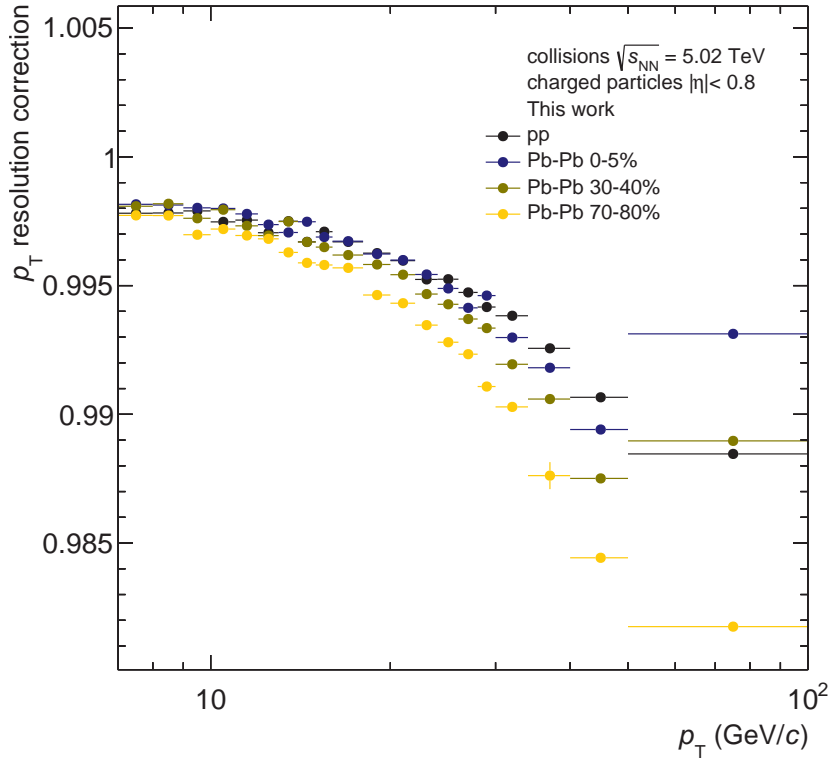
Figure 3.16: p_T resolution correction factor as function of p_T in pp and in three centrality classes of Pb-Pb collisions at $\sqrt{s} = 5.02$ TeV recorded in ALICE in 2017.

sions. This two-dimensional representation is converted to a distribution of the mean uncertainty $\langle\sigma(1/p_T)\rangle$ as function of $1/p_T$. In addition, this distribution is parametrized with a quadratic polynomial $n(1/p_T)$ which is then subtracted from the original uncertainty $\sigma(1/p_T)$ in order to reduce statistical fluctuations as seen in the top-right panel of Figure 3.15. The scaled $\sigma(1/p_T)$ distribution (see bottom left panel of Figure 3.15) is subsequently projected into the *y*-axis obtaining thereby a distribution that reveals the probability function of $\sigma(1/p_T)$ as shown in the bottom-right panel of Figure 3.15.

After determining the shape of $\sigma(1/p_T)$, the smearing is performed. In this process, tracks are simulated and distributed according to a power-law fit of the true p_T distributions in the range $7 \le p_T \le 100$ GeV/*c*. These are then transformed in $1/p_T$ distributions. Moreover, each $1/p_T$ value is smeared according to a Gaussian with a standard deviation generated randomly in respect of the probability function of $\sigma(1/p_T)$. This standard deviation is first adjusted with an additive factor extracted from an evaluation of the polynomial $n(1/p_T)$ as re-correction for the above mentioned suppression of statistical fluctuations (see bottom-left panel of Figure 3.15). Finally, the smeared $1/p_T$ distribution is re-converted in

a p_T spectrum.

The p_T resolution correction factors are calculated by means of the quotient of the original simulated p_T spectrum and the smeared one. In Figure 3.16, the resulting correction factors for the analyzed pp and for central, semi-central and peripheral Pb-Pb collisions are respectively illustrated. In pp collisions, the correction factor falls monotonically and the effect finds its maximum at the largest p_T interval with a percentage value of around 1.2%. On the other hand, the Pb-Pb results behave similarly as in the pp data collection. Nevertheless, a centrality dependence of the correction factors is observed which reflects that the impact of the correction on the p_T distributions increases with this quantity.

3.5 Implementation of the corrections

The analysis strategies presented in the previous sections aim to mitigate the effects derived from the detector response that distorts the data sample. As already mentioned, the track-level corrections affect the shape of the primary track distribution $N(p_T)$. The corrected primary track distribution $N_{corr}(p_T)$ can be defined as follows (cite mknichel):

$$
\begin{aligned}
N_{corr}(p_T) = {} & N(p_T) \times c_\epsilon(p_T) \cdot c_{par\ comp}(p_T) \\
& \times (1 - c_{cont}(p_T) \cdot c_{sec\ sca}(p_T)) \\
& \times c_{reso}(p_T) \times c_{sig\ loss}(p_T)
\end{aligned}
\tag{3.13}
$$

where c_i represent the contributions of each correction: $c_\epsilon(p_T)$ is the tracking efficiency, $c_{par\ comp}(p_T)$ the particle composition, $c_{cont}(p_T)$ the secondary contamination, $c_{sec\ sca}(p_T)$ the secondary scaling, $c_{reso}(p_T)$ the p_T resolution and $c_{sig\ loss}(p_T)$, which is only applied in pp collisions. Once the corrections are implemented, the p_T distributions can be calculated in respect of the definitions discussed in Section (cite theory section). According to the definition of the nuclear modification factor, the results in pp collisions are expressed in form of a differential cross section:

$$
\frac{d^2\sigma_{ch}^{pp}}{d\eta dp_T} = \sigma_{vis} \cdot \frac{1}{N_{INEL}} \cdot \frac{d^2 N_{corr}^{pp}}{d\eta dp_T}
\tag{3.14}
$$

where N_{INEL} corresponds to the total number of inelastic events and σ_{vis} the visible inelastic cross section as already explained in Section 3.4.1. In the case of Pb-Pb collisions, the physical quantity used for the determination of nuclear

Track parameter	Condition	Variations	
Selection of primaries			
DCA_z	$\leq$ 2 cm	1 cm	5 cm
DCA_{xy}	$\leq 7\sigma$	4σ	10σ
ITS selection			
at least one hit in the SPD	required	not required	
χ^2 per ITS cluster	$\leq$ 36	$\leq$ 25	$\leq$ 49
TPC selection			
χ^2 per TPC cluster (pp collisions)	$\leq$ 4.0	$\leq$ 3.0	$\leq$ 5.0
χ^2 per TPC cluster (Pb-Pb collisions)	$\leq$ 2.5	$\leq$ 1.7	$\leq$ 3.4
fraction of shared TPC clusters	$\leq$ 0.4	$\leq$ 0.2	$\leq$ 1.0
ratio of crossed rows over findable clusters	$\geq$ 0.8	$\geq$ 0.7	$\geq$ 0.9
geometric length (dead TPC area)	3 cm	2 cm	4 cm
geometric length (track length)	130	120	140
TPC-ITS selection			
χ^2 TPC constrained track vs. global track	$\leq$ 36	$\leq$ 25	$\leq$ 49

Table 3.5: Variations for the standard track selection criteria used for the calculation of the systematic uncertainties.

modification factors is the invariant yield and it is calculated as follows

$$\frac{\mathrm{d}^2 N_{\mathrm{ch}}^{\mathrm{Pb\text{-}Pb}}}{\mathrm{d}\eta\,\mathrm{d}p_{\mathrm{T}}} = \frac{1}{N_{\mathrm{INEL}}} \cdot \frac{\mathrm{d}^2 N_{\mathrm{corr}}^{\mathrm{Pb\text{-}Pb}}}{\mathrm{d}\eta\,\mathrm{d}p_{\mathrm{T}}} \tag{3.15}$$

3.6 Systematic uncertainties

The p_{T} distributions as well as the nuclear modification factors are affected by uncertainties product of inaccuracies in the measurement instruments or assumptions made during the development of the analysis approaches. These are the so-called systematic uncertainties. In this thesis, the source of systematics considered arises from the track selection criteria and from the secondary scaling correction. However, other sources of systematic uncertainties have been identified in a previous analysis of nuclear modification factors (cite paper). In

this chapter, these sources will be described and its impact on the p_T distributions approximated.

Each source of systematic uncertainty σ_i^{sys} contributes to the total systematic uncertainties $\sigma_{\text{tot}}^{\text{sys}}$ independently of the rest. For this reason, the total value of $\sigma_{\text{tot}}^{\text{sys}}$ is calculated via the root sum squared method:

$$\sigma_{\text{tot}}^{\text{sys}} = \sqrt{\sum_i \sigma_i^{\text{sys}}} \tag{3.16}$$

In the following sections, the approaches applied for the calculation of the individual contributions σ_i^{sys} are explained in detail.

3.6.1 Track selection

As in Section 3.3 explained, the track selection has been develop in the course of the years and validated through cross-checks. Nevertheless, the choice of the selection criteria, yet justified via these studies, is subjected to a systematic uncertainty. Each criterion contributes independently of the rest to the systematic uncertainties of the p_T spectra, and as a consequence, of the nuclear modification factors. For this reason, the total value of the systematic uncertainty $\sigma_{\text{sel,tot}}^{\text{sys}}$ results from the root sum squared over individual ones $\sigma_{\text{sel,i}}^{\text{sys}}$.

To determine the individual contributions, each nominal track requirement listed in Table 3.2 is varied twice within a range in which it is believed to be physically meaningful. The variations correspond to the maximum and the minimum of this range. The mentioned variations are listed in Table 3.5. Next, the calculation of fully corrected p_T distributions is repeated by performing one variation at a time. These p_T distributions are compared to the nominal one by means of a ratio, obtaining as result two ratios for each track cut, with the exception of the required hit in the SPD. Finally, the two ratios are evaluated in each p_T interval and the largest deviation from the nominal p_T spectrum between the both of them is assumed as systematic uncertainty in the corresponding p_T interval. Thereby, the systematic uncertainty corresponding to each track cut is presented as function of p_T as shown in Figure **??**. The total systematic uncertainty is calculated subsequently applying a root sum squared in each p_T interval among the individual systematic uncertainties. This same approach is also applied to determine the systematic uncertainties that affect the nuclear modification factors as seen in Figure **??**.

4 Results

Appendices

A Appendix

A.1 Tracking efficiency