Investigation of event-shape classifiers for proton-proton collisions with the ALICE experiment

Tilde Bonnevier Wallstedt

Thesis submitted for the degree of Master of Science
60 hp

Supervisor: Alice Ohlson

Department of Physics: Division of Particle and Nuclear Physics

May 2024

LUNDS UNIVERSITET

Abstract

This project uses proton-proton collision data from the ALICE experiment at CERN to investigate the differences between three ways to classify events, known as *event shape classifiers*. The event shape classifiers that were investigated are R_T, spherocity and flattenicity.

The motivation behind this is the consistent use of these classifiers to select events when investigating the heavy-ion like behavior observed in high-multiplicity proton-proton collisions. A comprehensive understanding of the types of events that are being selected by these classifiers is essential, and providing the community with such an understanding is the goal of this project.

Two-particle angular correlation functions were created in classifier bins to investigate the event shape. While all three classifiers can select more isotropic events (dominated by the underlying event, where QGP-like effects are more likely to be found) and more jet-like events, the jet-like effects are present for all event shape classes.

The results indicate several differences between the event selection being done by the three classifiers. These differences include the bias towards events with an average of 90 degrees between the tracks for higher spherocities, and a bias against events where two or more tracks belong to the same flattenicity cell for lower flattenicities. Flattenicity appears to select the most isotropic events out of the three classifiers.

Popular science summary

The universe consists of elementary particles that come together to form what we recognize as ordinary matter. You may know that atoms are made up of protons and neutrons. However, these are not the smallest constituents we know of. Protons and neutrons are what we call *hadrons*, and they are made up of *quarks* and *gluons*. The gluons act as a glue, keeping the quarks together inside the hadrons. In particle accelerators, such as the Large Hadron Collider (LHC) in Switzerland, protons and heavy ions can be accelerated and collided with the goal of ripping them apart and studying the products.

One example of interesting research that can be done in such colliders is quark-gluon plasma (QGP) research. The QGP is a state of matter that we believe the universe consisted of just after the Big Bang, and it can also exist inside neutron stars. The QGP is the result of highly energetic environments, where there is so much energy that the quarks and gluons are not as bound to each other as they normally are, and they exist semi-freely in a "soup" of quarks and gluons. Due to the high energies when colliding heavy ions such as Pb (lead), it is possible to, for a brief moment, create a QGP at a collider like the LHC. One can then study the way that particles behave after the collision to extract information about the QGP.

A QGP should not be able to be created inside proton-proton collisions due to the smallness of these systems. However, scientists have observed several signatures that were previously assigned to be signatures of the QGP, in a certain type of proton-proton collisions known as "high-multiplicity" proton-proton collisions. High multiplicity means that a lot of particles are created in the collision.

A hot topic in the field of heavy-ion physics is understanding why we see these QGP-like signatures in such high-multiplicity proton-proton collisions. Recently scientists have begun to explore the effect that the *event shape* may have on this behavior. The event shape refers to how the particles after the collision are spread out. There are three major classifiers that are currently being used to select different shapes of collisions (so-called *event shape classifiers*): R_T, spherocity and flattenicity. The problem is that people are using these classifiers to classify collisions, but there has been no study on what the difference between the collisions that are being selected by these classifiers is, or what types of collisions are being selected by each of them.

The goal of this project is to make a study of the types of collisions that are being selected by each of the classifiers, and how they differ from each other. This was done by analyzing proton-proton collision data from one of the experiments at the LHC: the ALICE (A Large Ion Collider Experiment) experiment, which detects the products of both heavy-ion collisions and proton-proton collisions. This is an important contribution to the ongoing research on how event shapes are correlated with the heavy-ion like behavior in high multiplicity proton-proton collisions.

Contents

1 Theory **3**
- 1.1 The standard model of particle physics . . . 3
  - 1.1.1 The leptons . . . 4
  - 1.1.2 The quarks . . . 4
  - 1.1.3 The gauge bosons . . . 4
  - 1.1.4 The scalar bosons . . . 5
- 1.2 Heavy-ion physics . . . 5
  - 1.2.1 Quark-gluon plasma . . . 5
- 1.3 Heavy-ion sign. in pp collisions . . . 6
  - 1.3.1 Collective flow . . . 6
  - 1.3.2 Strangeness enhancement . . . 8
  - 1.3.3 Suggested explanations . . . 9
- 1.4 The ALICE experiment . . . 10
  - 1.4.1 TPC . . . 11
  - 1.4.2 ITS . . . 11
  - 1.4.3 V0 . . . 11
  - 1.4.4 Coordinates . . . 11
- 1.5 Event shape classifiers . . . 12
  - 1.5.1 R_T . . . 13
  - 1.5.2 Spherocity . . . 14
  - 1.5.3 Flattenicity . . . 15
- 1.6 Project aim . . . 16

2 Method **17**
- 2.1 Track and event selection . . . 17
- 2.2 Two-part. ang. correlation functions . . . 18
  - 2.2.1 Purpose and interpretation of the correlation function . . . 18
  - 2.2.2 Definition . . . 18
- 2.3 Efficiency corrections . . . 21
  - 2.3.1 ZYAM subtraction . . . 22
- 2.4 Evaluation of systematic uncertainties . . . 23
- 2.5 Summary . . . 25

3 Results and discussion **27**
3.1 R_T . . . 27
3.1.1 Configuration 1: p_T inclusive . . . 29
3.1.2 Configuration 2: p_T bins . . . 30
3.1.3 Configuration 3: w.r.t. R_T trigger track . . . 32
3.1.4 Configuration 4: within the transverse region . . . 34
3.2 Spherocity . . . 35
3.2.1 Configuration 1: p_T inclusive . . . 37
3.2.2 Configuration 2: p_T bins . . . 38
3.2.3 Configuration 3: w.r.t. S_0 unit vector . . . 40
3.3 Flattenicity . . . 41
3.3.1 Configuration 1: p_T inclusive . . . 43
3.3.2 Configuration 2: p_T bins . . . 44
3.4 Comparisons . . . 46
3.5 Conclusion and outlook . . . 49

A p_T-weighted spherocity **54**

B p_T-weighted flattenicity **56**

Chapter 1

Theory

1.1 The standard model of particle physics

The standard model of particle physics is a theoretical model that attempts to describe the fundamental particles of the universe and their interactions. The particles that make up the standard model (figure 1.1) are divided into four groups: the *leptons*, the *quarks*, the *gauge bosons* and the *scalar bosons*. Each of the groups will be described briefly below.

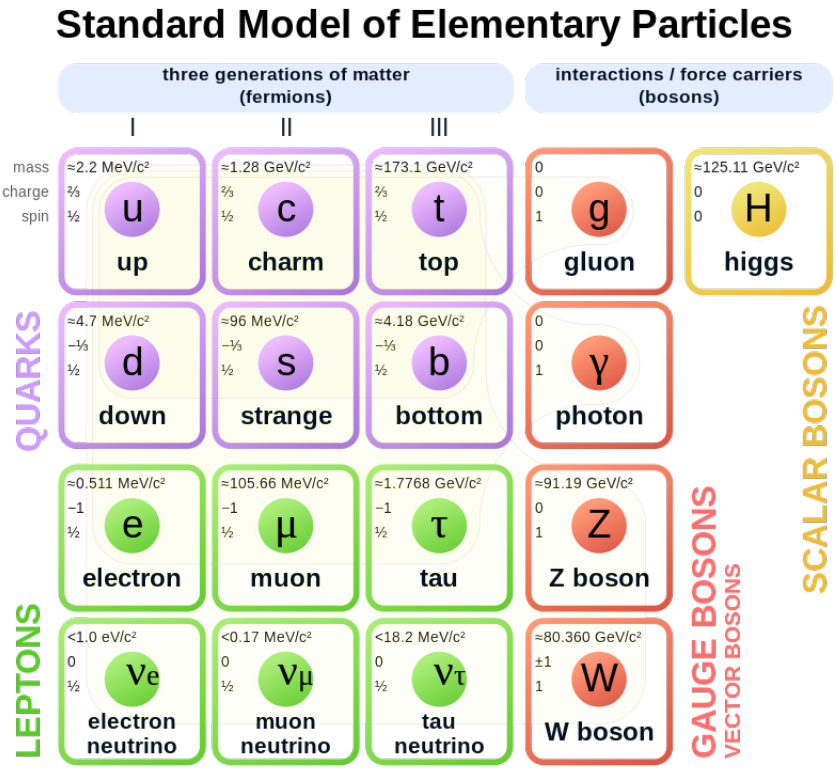


Figure 1.1: The standard model of particle physics. Image taken from [1]

1.1.1 The leptons

The leptons are half-integer spin particles that are divided into six particles in three generations. The particles are the electron e^-, the muon μ^-, the tau lepton τ^-, and their respective neutrinos ν_e, ν_μ and ν_τ and their antiparticles. The generations are each lepton and its neutrino. The leptons can interact weakly and the charged leptons can also interact electromagnetically (see section 1.1.3). The electron, muon and tau lepton have an electric charge, whereas their respective neutrinos are neutral. The charged leptons interact with the Higgs field and thus have a mass. The neutrinos are predicted by the standard model to be massless, but experimental evidence shows that they have a nonzero mass. This contradiction is one of the many phenomena that point to physics beyond the standard model.

1.1.2 The quarks

Quarks are half-integer spin particles with varying masses. There are six flavors and three generations of quarks. Each generation consists of two quarks: one with electrical charge $+2/3$ and one with electrical charge -1/3. The first generation consists of the up ($+2/3$) and the down (-1/3) quarks. They are the lightest generation and make up regular matter (protons and neutrons). The second generation consists of the charm ($+2/3$) and strange (-1/3) quarks, and finally the third generation consists of the top ($+2/3$) and the bottom (-1/3) quarks. The top quark is by far the heaviest quark and decays immediately. All the quarks have a respective anti-quark.

Quarks possess a property called color charge; a quark has a color (red, green, blue) and an anti-quark has an anti-color (anti-red, anti-green, anti-blue). Quarks form the hadrons, which are what matter is made from. Hadrons can either consist of a quark and an anti-quark (known as a *meson*) or three quarks or three anti-quarks (known as a *baryon*). A hadron is required to be "white", i.e. if the hadron is a meson the quarks must have opposite color (a color and its anti-color) and if it is a baryon its quarks must have the colors red, green and blue, or anti-red, anti-green and anti-blue.

The quarks are bound together by the strong force, which is mediated by the *gluons*, described in section 1.1.3, and may not escape each other.

1.1.3 The gauge bosons

The gauge bosons are integer-spin particles, and the group consists of the force mediators of three of the four fundamental forces of nature: the weak force with its mediators the *W and Z bosons*, the strong force with its mediator the *gluon* and the electromagnetic force with its mediator the *photon*. A mediator for the gravitational force has yet to be found.

The weak force, mediated by the massive W and Z bosons, is responsible for radioactive decay. The W boson can be either positively or negatively charged (W^+ and W^-). The Z boson is electrically neutral.

The strong force is responsible for binding together quarks into hadrons, by exchange of the massless gluons. The strong force has a property known as *confinement* which means that no

color charged particles can exists freely. Another property called *asymptotic freedom* means that the strength of the strong force decreases as the energy scale increases (and the length scale decreases).

If there is such an excess of energy (such as in a particle collider) that quarks attempt to escape each other, colored particles are formed around these quarks in order to maintain color confinement. This process may continue if enough energy is present, forming more and more colored particles. Eventually, when the energy decreases, the quarks will recombine into hadrons. The result is a collimated spray of particles in a cone shape, known as a *jet*. The theory of the strong interaction is called quantum chromodynamics (QCD).

The strength of the electromagnetic force decreases with distance and has an infinite range.

1.1.4 The scalar bosons

Only one particle in this category has been discovered: the *Higgs boson*, experimentally discovered in 2012 [2]. The Higgs boson is the mediator of the Higgs field. Interactions with the Higgs field give particles mass, so massless particles like the photon or the gluon do not interact with the Higgs field. The Higgs boson has spin 0, and as it has a mass it interacts with itself (known as *self-coupling*).

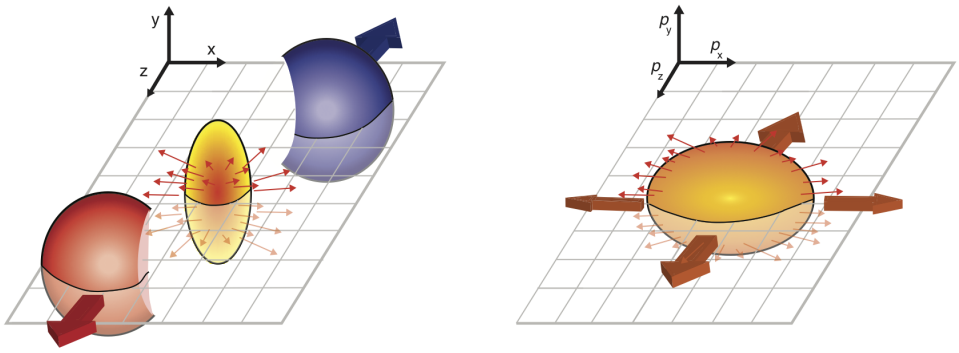
1.2 Heavy-ion physics

Heavy-ion physics refers to the study of colliding heavy nuclei and the effects that come from such collisions. Heavy ions are being collided at for instance the LHC (Large Hadron Collider) at CERN, where Pb ions are being collided and the ALICE detector (A Large Ion Collider Experiment) (see section 1.4) is used to study the resulting particles.

1.2.1 Quark-gluon plasma

In the year 2000, experimental results from CERN's heavy-ion program revealed behavior that could be explained by the presence of a state of matter known as the *quark-gluon plasma* (QGP) that can be created in heavy-ion collisions, whose existence had been suggested already in the 1970s. At extremely high temperatures or densities, quarks and gluons (which are normally bound together by the strong force) become quasi-free. This can be explained by the phenomenon known as asymptotic freedom, explained briefly in section 1.1.3. As the energy scale increases, the strength of the strong force decreases. Thus, if enough energy is present, the quarks and gluons may form a medium where they are quasi-free. When studying this state of matter, it was discovered that it behaves like a liquid, as the particles move like one medium rather than individual particles and a flow coefficient may be extracted by studying the movement of this medium. [3]

In a heavy-ion collision, due to the non-negligible size of the colliding nuclei the overlap region leads to pressure gradients creating a flow in the azimuthal plane, which is illustrated in figure 1.2. This azimuthal flow has been perceived as a signature of the QGP, and is sensitive to properties of the medium such as viscosity.



(a) Spatial coordinates (b) Momentum coordinates

Figure 1.2: Illustration of the elliptic flow in heavy-ion collisions. (a) shows the flow in spatial coordinates, where the direction of the flow results from the shape of the overlap region. (b) shows the flow in momentum coordinates after the collision. Figure credit: B. Hippolyte

1.3 Heavy-ion signatures in proton-proton collisions

The past years of research in the heavy-ion physics community have revealed that signatures of heavy-ion collisions (which typically were assigned to be signatures of the strongly interacting medium, the QGP, see section 1.2.1) can be observed in high-multiplicity proton-proton (pp) collisions [4]. High-multiplicity refers to a large amount of particles being produced in the collision. These signatures include two phenomena referred to as *collective flow* and *strangeness enhancement*.

1.3.1 Collective flow

As described in section 1.2.1, the azimuthal flow created in heavy-ion collisions was perceived as a signature of the QGP. A signature of this collective flow is the so-called *ridge structure* that appears in two-particle angular correlation functions in $\Delta\eta$ (ϕ and η are the coordinates used in the detector, see section 1.4.4), see figure 1.3 below. It can be seen that there is a correlation in $\Delta\phi$ between particles that are far away from each other in η (large $\Delta\eta$) around $\Delta\phi = 0$.

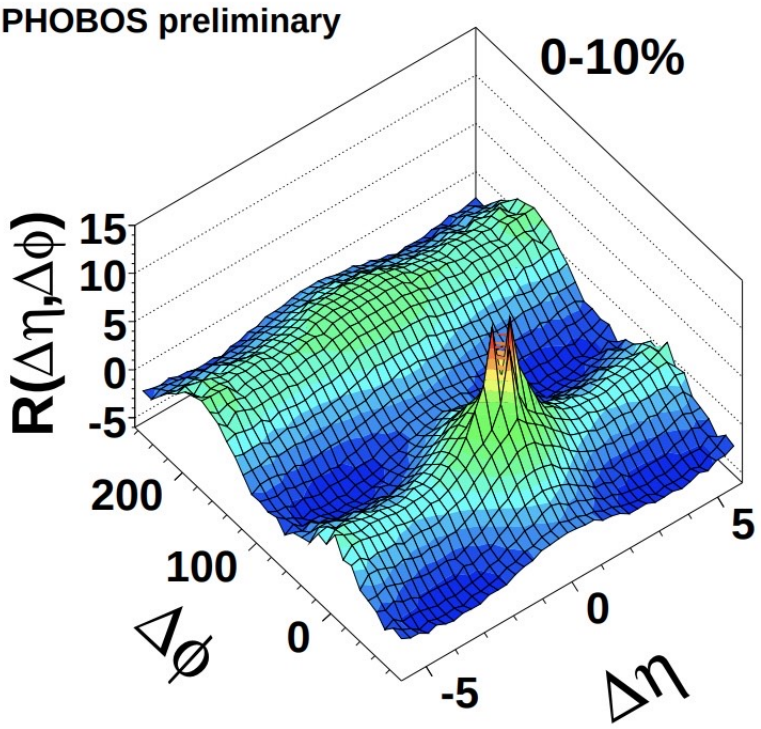


Figure 1.3: The ridge structure in two-particle angular correlation functions in Au-Au collisions. Image taken from [5].

In 2010, the CMS experiment presented the first observation of such a ridge structure in high-multiplicity pp collisions, indicating a collective flow in small collision systems [6] (see figure 1.4). This challenges the idea that most general-purpose QCD-inspired Monte Carlo (MC) generators (e.g. PYTHIA, see [7]) are based on: the fact that pp collisions can be described as sums of independent parton-parton collisions [8].

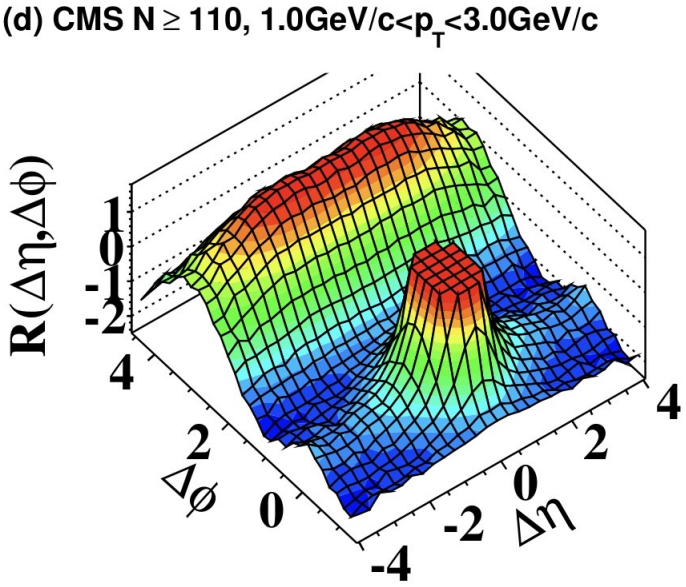


Figure 1.4: The ridge structure in two-particle angular correlation functions in high-multiplicity pp collisions found by the CMS experiment. Image taken from [6].

1.3.2 Strangeness enhancement

In heavy-ion collisions, one signature of the QGP is a phenomenon known as strangeness enhancement, which refers to the increased ratio of strange particles (hadrons containing strange quarks) to pions in heavy-ion collisions compared to pp collisions. In 2016 the ALICE collaboration published a paper presenting the first observation of such a strangeness enhancement in high-multiplicity pp collisions (see figure 1.5), providing the community with yet another heavy-ion feature that can be observed in pp collisions [9].

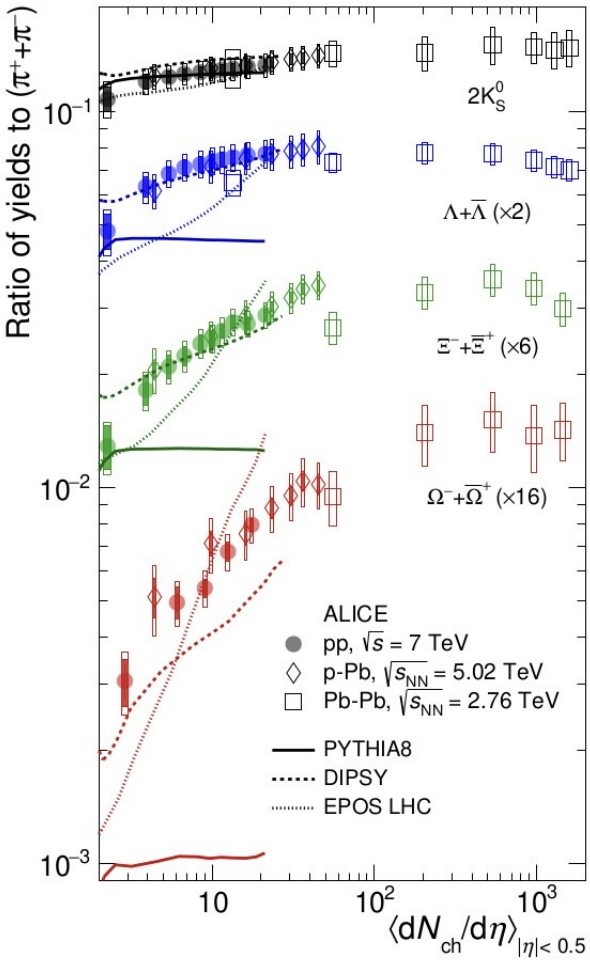


Figure 1.5: The ratio of the yields of K^S_0, Λ, Ξ and Ω to pions as a function of multiplicity for pp, p-Pb and Pb-Pb measured by ALICE and compared with models. Image taken from [9]

1.3.3 Suggested explanations

Multi-parton interactions (MPIs) have been suggested as a possible explanation of some of the flow-like behavior observed in high-multiplicity pp collisions. MPI refers to the phenomenon where large parton densities result in an enhanced probability of multiple partons from each proton interacting with each other simultaneously [10]. The higher the multiplicity, the more likely it is for MPIs to occur. MPIs can lead to a similar effect in particle momentum spectra to radial flow [11].

PYTHIA is the main Monte Carlo model for describing pp collisions [7]. The effect of radial flow can be observed when introducing color reconnection. In order to obtain elliptic flow, PYTHIA has introduced a concept called "shoving" which refers to partons being significantly influenced by the neighbouring partons due to the high densities after high-multiplicity pp collisions [12]. Shoving could affect the final distribution of particles, and lead to features in two-particle correlation functions reminiscent of elliptic flow. The

rope hadronization model that PYTHIA 8 has implemented is able to describe the strangeness enhancement in high-multiplicity pp collisions [13] [11].

Furthermore, there are QGP-inspired models that use hydrodynamics to describe the collisions, such as SONIC [14] and MUSIC [15]. These models can describe some of the flow-like behavior but are not able to describe strangeness enhancement.

While there are several suggestions for phenomena that could explain the heavy-ion like behavior observed in high-multiplicity pp collisions, no model can explain both phenomena at the same time, or give a full quantitative description of the data.

1.4 The ALICE experiment

ALICE is one of the four major experiments at the LHC. The other three are ATLAS (A Toroidal LHC Apparatus), CMS (Compact Muon Solenoid) and LHCb (LHC-beauty). The ALICE detector is dedicated to detecting the products of heavy-ion collisions in order to investigate the properties of the QGP (see section 1.2.1). The detector has several detector systems, including tracking, particle identification systems, calorimeters and collision characterization systems. In this project, data from the TPC, ITS and V0 is utilized, and these systems will be described briefly below. The data used in this thesis are from LHC run 2 from 2018. Some of the detector subsystems have since been upgraded. However, in this thesis the instrumentation will be discussed as it was at the time of the data taking.

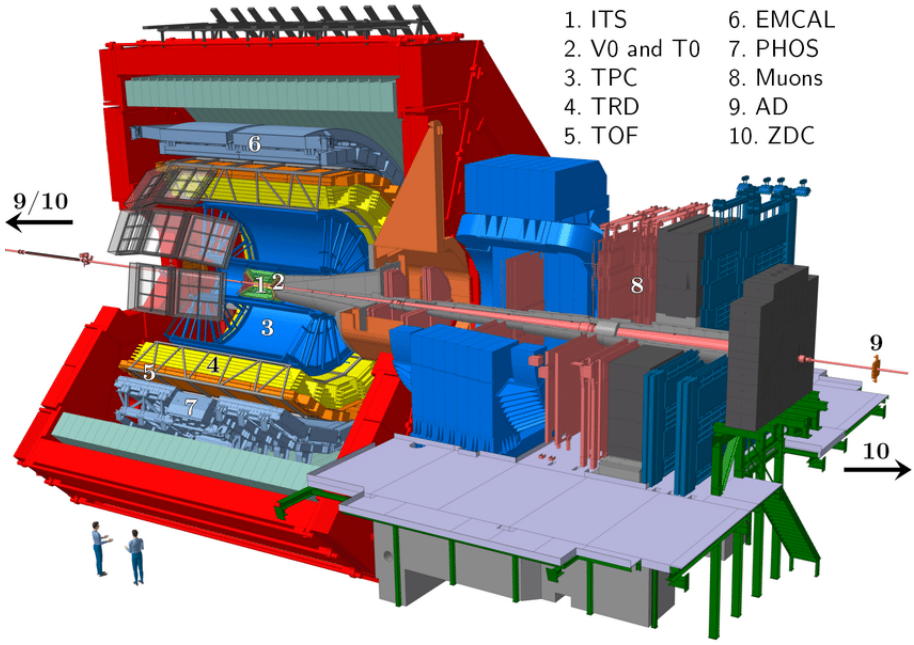


Figure 1.6: The ALICE detector. Image taken from [16]

1.4.1 TPC

The ALICE Time Projection Chamber (TPC) is designed to detect the charged particles that are produced in Pb-Pb collisions, even in very high-multiplicity events. A TPC is a particle detector that utilizes the ionization of gas atoms to determine the trajectory of a particle. When a charged particle traverses the gas, it will ionize the gas atoms and the electrons will drift toward the read-out chambers which consist of multi-wire proportional chambers (MWPC) at the ends of the cylinder. By measuring the time it takes for an electron to reach the MWPC one can determine the position along the beam axis at which the ionization took place. The radial coordinate is determined by wires in the MWPC along the azimuthal direction. The read-out chambers of the ALICE TPC are segmented into 18 parts azimuthally. The ALICE TPC has a full azimuthal acceptance, and a pseudorapidity acceptance of $|\eta| < 0.9$. When tracks are reconstructed in the TPC, the number of clusters deposited in the TPC are utilized. Normally somewhere between 70 and 100 clusters are required to be deposited in the TPC for good-quality track reconstruction. [17]

1.4.2 ITS

The Inner Tracking System (ITS) has several purposes, such as determining the primary vertex of a collision, reconstructing secondary vertices and identifying low-momentum particles. It complements the TPC tracking with its excellent spatial resolution. It is the inner-most subsystem of the central barrel, and it consists of six silicon detectors: two layers of silicon pixel detector, two layers of silicon drift detector and two layers of silicon strip detector (ordered from closest to the beamline to furthest away). The different subsystems of the ITS have different acceptances, varying from $|\eta| < 2$ to $|\eta| < 0.9$. [17]

1.4.3 V0

The V0 forward detector consists of two subsystems: V0A and V0C, which are located on the beam axis on opposite sides of the collision point. They consist of scintillator detectors, and they are able to determine the centrality of a collision by measuring the energy deposited in these detector systems, since this energy is directly proportional to the centrality of the collision. The centrality refers to the amount of overlap the two colliding nuclei have. It is typically defined as a percentile where a small percentage corresponds to a large overlap region and a large percentage corresponds to a small overlap region. The V0A has an acceptance of $2.8 < \eta < 5.1$, and the V0C has an acceptance of $-3.7 < \eta < -1.7$. [17]

1.4.4 Coordinates

Figure 1.7 displays the coordinates used in the ALICE detector. The relevant coordinates for this analysis are the azimuthal angle ϕ and the pseudorapidity η. ϕ can be seen in figure 1.7. η is related to the polar angle θ by the following relation:

$$\eta = -ln\left(tan\frac{\theta}{2}\right) \tag{1.1}$$

The η value of a track is 0 when it is perpendicular to the beam axis, and it goes to ∞ when it is parallel to the beam axis.

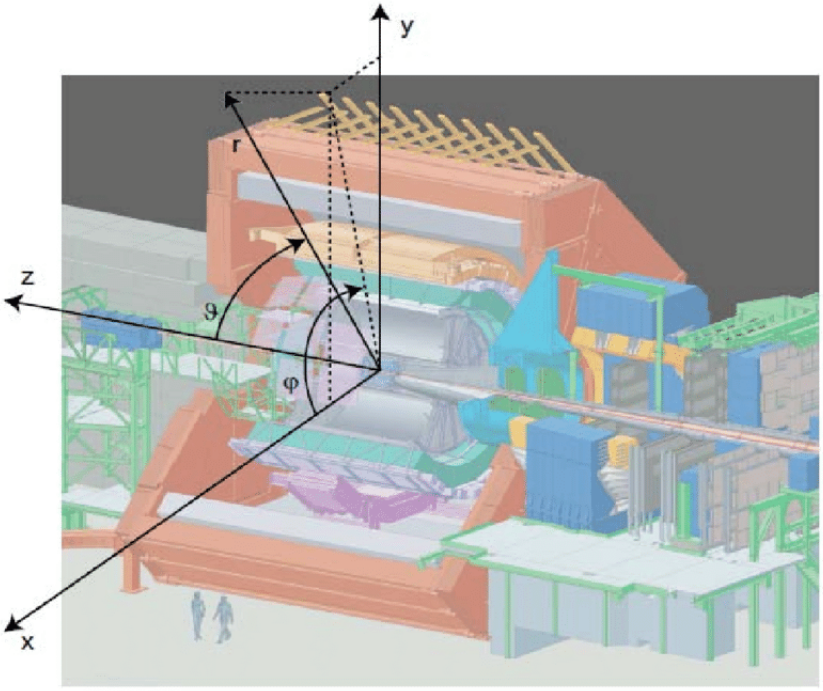


Figure 1.7: The ALICE coordinate system. Image taken from [18]

The ALICE detector detects particles within the full azimuthal range $\phi = (0, 2\pi)$, and within the pseudorapidity range $\eta = (-0.8, 0.8)$. Another relevant variable is the *z vertex* of an event, denoted as v_z. This refers to the position along the z axis (along the beamline) where the collision takes place, and it is defined as the distance (in cm) from the center of the detector.

1.5 Event shape classifiers

Perturbative quantum chromodynamics (pQCD) is reasonably good at describing high-p_T particle production (hard processes), whereas it fails at describing low-p_T particle production (soft processes), and it is in this low-p_T region one usually discovers new effects in pp collisions [19]. In order to investigate the effect of the soft processes and how they relate to the heavy-ion like signatures that have been observed in high-multiplicity pp collisions (see section 1.3), one needs to be able to distinguish between events dominated by soft processes and events dominated by hard processes.

One way to select soft processes is by doing the event selection with a classifier based on the *shape* of the event: *event shape classifiers*. The shape refers to the momentum distribution of an event. A number of different event-selection methods have been developed in order to do this. In this project, three such classification methods are discussed and investigated: R_T, *spherocity* and *flattenicity*.

1.5.1 R_T

The *underlying event* (UE) is defined as the particles in an event that do not originate from the initial hard scattering of the event, nor can be traced back to this scattering somehow [4]. The UE consists of relatively low-p_T tracks, and it exists outside (and underneath) the core of jets. It is not possible to determine with certainty whether or not a track originates from the initial hard scattering. An alternative approach to experimentally measure the UE is to measure it geometrically, i.e. define a region transverse to the axis of a leading particle or jet and define this region as the UE region [20], see figure 1.8.

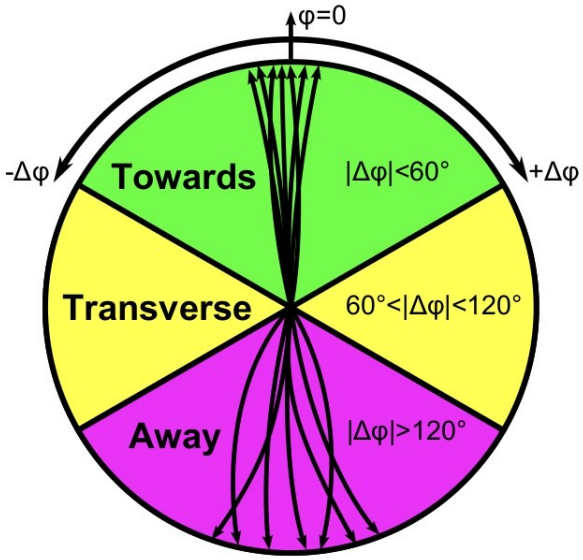


Figure 1.8: Illustration of the transverse, towards and away regions, used for the R_T measurements. Image taken from [20].

R_T stands for relative transverse activity, and it is an event shape classifier that measures the activity in this transverse region. The definition of R_T is:

$$R_T = \frac{N_T}{< N_T >} \tag{1.2}$$

where N_T is the number of particles in the transverse region for the event, and $< N_T >$ is the average number of particles in the transverse region for all events [4]. R_T measures the relative multiplicity in the transverse region. For events dominated by jet production, this number is expected to be relatively low, whereas the number is expected to be higher for UE-dominated events. A value of 1 indicates an average value, a value below 1 indicates a jetty event, and a value above 1 indicates a more isotropic event.

It has been shown that the number of multi-parton interactions (MPI) is strongly correlated with the UE activity. Therefore R_T as a classifier allows for examining the potential impact that MPI have on particle production [21].

Previous studies

Ref. [20] is the first mention of the event shape classifier R_T. In this paper R_T was proposed as a way to probe the UE activity, and it was suggested to investigate whether very low UE levels exhibit behavior more consistent with models than the high UE levels. This was suggested because pQCD should be able to describe jet-like events well, so if R_T selects events with non-perturbative physics the models should not work as well.

Ref. [4] studied particle spectra for pions, kaons and protons for varying values of R_T. It was concluded that models are able to describe the particle production in low-R_T events well, but fail at describing the production at higher R_T. This confirms the fact that one can reveal new features of the UE by measuring particle production as a function of R_T.

1.5.2 Spherocity

The final-state azimuthal topology is expected to reflect whether an event is dominated by hard or soft processes [8]. As explained in section 1.5, we are interested in probing the soft processes in order to examine the heavy-ion like behavior in high-multiplicity pp collisions.

Spherocity (S_0) is an event shape classifier that has been proposed as a way to measure this final-state azimuthal topology. It measures how spherical the momentum distribution of an event is, and it is defined, for the unit vector $\hat{\mathbf{n}}$ that minimizes the ratio, as:

$$S_0 = \frac{\pi^2}{4} \min \left(\frac{\sum_i |\vec{p_{T,i}} \times \hat{\mathbf{n}}|}{\sum_i \vec{p_{T,i}}} \right)^2 \tag{1.3}$$

[19], where the sums are performed over all tracks in the event.

This value is expected to go to zero for jetty events and to one for isotropic events.

Another way to define the spherocity is by using the *unweighted spherocity*, according to the following equation:

$$S_0^{p_T=1} = \frac{\pi^2}{4} \min \left(\frac{\sum_i |\vec{p_{T,i}} \times \hat{\mathbf{n}}|}{N_{trks}} \right)^2 \tag{1.4}$$

where the particles are not weighted by their transverse momentum. This is the version of the spherocity used in this thesis. The spherocity was calculated by looping over 100 unit vectors between $\phi = 0$ and 2π to find the unit vector that minimizes the ratio in equation 1.4.

Previous studies

In ref. [19] spherocity was used as a tool to select isotropic events. The goal of this study was to investigate the importance of jets in high-multiplicity pp collisions and how they contribute to the charged-particle production at low p_T. The study concluded that isotropic events are well described by models whereas the average p_T is overestimated for jet-like events.

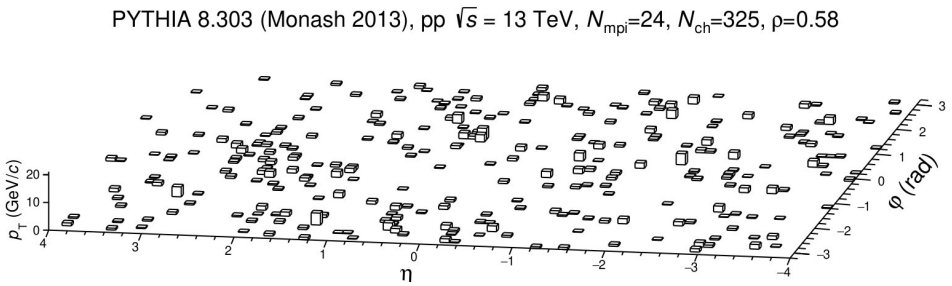
1.5.3 Flattenicity

The flattenicity, ρ, is an event shape classifier that attempts to distinguish between jetty events and so-called *hedgehog events*, i.e. events with an almost isotropic distribution of low-p_T particles. The flattenicity of an event is calculated by splitting the entire phase-space (ϕ-η space) into 80 elementary cells and calculating the total p_T in each cell. The definition is as follows:

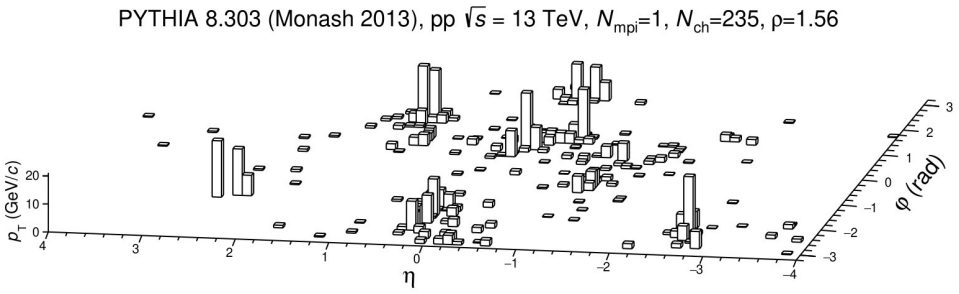
$$\rho = \frac{\sqrt{\sum_i (p_T^{cell,i} - < p_T^{cell} >)^2 / N_{cell}}}{< p_T^{cell} >} \tag{1.5}$$

where $p_T^{cell,i}$ is the total p_T in cell i, N_{cell} is the total number of cells, and $< p_T^{cell} >$ is the average cell p_T over all cells in the event [11]. Flattenicity goes to zero for isotropic events and to one for more jet-like events. In this thesis, 80 cells were used: 10 in ϕ and 8 in η.

Figure 1.9 illustrates the idea of hedgehog and multi-jet events.



(a) Illustration of hedgehog events in the ϕ-η plane



(b) Illustration of multi-jet events in the ϕ-η plane

Figure 1.9: Illustration of hedgehog vs multi-jet events in ϕ-η plane. Images taken from [22]

Previous studies

In ref. [11] flattenicity as a tool to select isotropic events was explored, and the PYTHIA 8 predictions for the p_T spectra of light- and heavy-flavored hadrons as a function of flattenicity was investigated. It was concluded that flattenicity does not have a significant dependence on the size of the cells in the expression (equation 1.5). Events selected by flattenicity were compared with events selected by the V0M multiplicity, and it was concluded that flattenicity has a higher tendency of selecting several low-p_T parton-parton

scatterings than the V0M multiplicity method. Moreover, the p_T spectra of the high-multiplicity pp collisions selected by flattenicity were softer than the p_T spectra of the collisions selected by the V0M multiplicity.

In a paper by the ALICE collaboration (see [21]), the flattenicity was calculated in the V0 detector to avoid auto-correlation. In that paper a version of flattenicity called *charged-particle flattenicity* was used, which counts the number of charged particles rather than the p_T in the following way:

$$\rho = \frac{\sqrt{\sum_i (N_{ch}^{cell,i} - < N_{ch}^{cell} >)^2 / N_{cell}^2}}{< N_{ch}^{cell} >} \tag{1.6}$$

In that paper the Q_{pp} ratio was measured as a function of flattenicity as well as p_T spectra in different flattenicity classes. It was concluded that these observations are well described by PYTHIA 8 with color reconnection, which suggests that pp collisions cannot be described as a sum of independent parton-parton scatterings. Furthermore, the particle ratios as a function of flattenicity were measured and it was found that the kaon-to-pion ratio increased with multiplicity which suggests that flattenicity could be useful as a tool to select isotropic high-multiplicity pp collisions.

In this thesis the definition of flattenicity given by equation 1.6 will be used, and the flattenicity will be calculated in the TPC due to limitations of available data.

1.6 Project aim

The goal of this project is to investigate the three event shape classifiers discussed in section 1.5: R_T, spherocity and flattenicity. Recently these three classifiers have been used to investigate the heavy-ion like behavior observed in high-multiplicity pp collisions. However, there is a knowledge gap in what types of events are being selected by these classifiers and how they differ. Understanding the events that they select is crucial in order to fully interpret the results and use classifiers, and contributing to that understanding is the goal of this project.

In order to compare the effect of selecting events with respect to these three classifiers, two-dimensional two-particle angular correlation functions (defined is section 2.2.2) in $\Delta\phi$ and $\Delta\eta$ were created and compared. This was done using 2018 pp collision data from the ALICE detector. The correlation functions were made in four configuration for R_T, three configurations for spherocity, and two configurations for flattenicity.

Chapter 2

Method

This thesis utilized ROOT, CERN's software framework for data handling (see [23]), to perform the data analysis and to visualize the results.

2.1 Track and event selection

For this analysis, data from the ALICE detector was used, from pp collisions collected during LHC Run 2 in 2018. Pileup events (i.e. when multiple collisions occur within the read-out time) were discarded, events were required to have a reconstructed vertex, and the vertex was required to be within $-10 < v_z < 10$, i.e. the collisions were required to take place within ± 10 cm from the center of the detector.

A multiplicity cut was implemented for the spherocity and flattenicity analysis, only selecting the top 10 % multiplicity events. This was done since neither the spherocity nor flattenicity definitions include any constraint on the multiplicity, and due to the way that these observables are defined, there is an inherent multiplicity dependence of the spherocity and flattenicity, biasing higher spherocities (lower flattenicities) towards higher multiplicity events. The multiplicity cut ensures that any dependence spherocity and flattenicity have on the multiplicity gets removed, and it is the high-multiplicity events that are of interest in the studies that have been performed using spherocity or flattenicity to select events. The multiplicity was measured in the V0 of the ALICE detector (see 1.4).

Tracks were reconstructed in the TPC and ITS. The η of a track was required to be within $-0.8 < \eta < 0.8$. The standard p_T range for this analysis was 0.2 GeV/c $< p_T <$ 5 GeV/c. However, for the leading tracks in the R_T analysis the p_T was required to be within 5 GeV/c $< p_T <$ 20 GeV/c.

A DCA cut was implemented in order to decrease the number of *secondary tracks*. A secondary track is when the track is not a direct product of the primary interaction, but the result of a weak decay. Secondary tracks stand in contrast to *primary tracks*, which come from the primary interaction. DCA stands for distance of closest approach, and this distance determines how close a track has to be to the primary collision vertex in order to be accepted as a part of an event. A loose DCA cut results in a high efficiency but low accuracy as some secondary particles will be included. A tight DCA cut results in a lower efficiency

since some of the tracks will not meet the criteria even though they were a part of the events, but higher accuracy, as fewer secondaries will be included.

2.2 Two-particle angular correlation functions

2.2.1 Purpose and interpretation of the correlation function

The method used in this project to investigate the difference between the three event shape classifiers introduced in section 1.5 is to create *two-particle angular correlation functions*. The purpose of a correlation function is to illustrate how particles correlate with each other, with the goal of creating a representation of the relative angle between tracks over a lot of events. This can give information about the types and shapes of events that are being selected, which is what this thesis aims to investigate.

Correlation functions were made, selecting events using R_T, spherocity and flattenicity. By comparing, for instance, the appearance of the correlation function for events with a high spherocity and the correlation function for events with a high R_T, conclusions can be made regarding how spherocity and R_T differ in their event selection.

Two-particle angular correlation functions use the angular differences $\Delta\phi = \phi_1 - \phi_2$ and $\Delta\eta = \eta_1 - \eta_2$, where ϕ_1 and η_1 are the ϕ and η (see section 1.4.4 for definitions and figure 1.3 for an example correlation function) of one track in an event, and ϕ_2 and η_2 are the ϕ and η of another track in that event. A signature of jets in such correlation functions is a near-side peak at $(\Delta\phi, \Delta\eta) = (0, 0)$ and an away-side peak at $\Delta\phi = \pi$. The near-side peak represents particles from the fragmentation of the same jet, whereas the away-side peak represents particles in back-to-back jets. As mentioned in section 1.3, a signature of elliptic flow in correlation functions is the double ridge structure in $\Delta\phi$.

2.2.2 Definition

The definition of the correlation function used in this work is the following:

$$C(\Delta\phi, \Delta\eta) = \frac{S(\Delta\phi, \Delta\eta)}{M(\Delta\phi, \Delta\eta)} \tag{2.1}$$

where S stands for same-event correlations and M stands for mixed-event correlations. The same-event correlations are defined as:

$$S(\Delta\phi, \Delta\eta) = \frac{1}{N_{trigger}} \frac{d^2 N_{pairs}^{same}}{d\Delta\phi d\Delta\eta} \tag{2.2}$$

where $N_{trigger}$ is the number of trigger tracks (for this analysis always the higher p_T track unless otherwise stated), and N_{pairs}^{same} is the number of trigger-associate particle pairs. The mixed-event correlations are defined as:

$$M(\Delta\phi, \Delta\eta) = \frac{1}{\alpha} \frac{d^2 N_{pairs}^{mixed}}{d\Delta\phi d\Delta\eta} \tag{2.3}$$

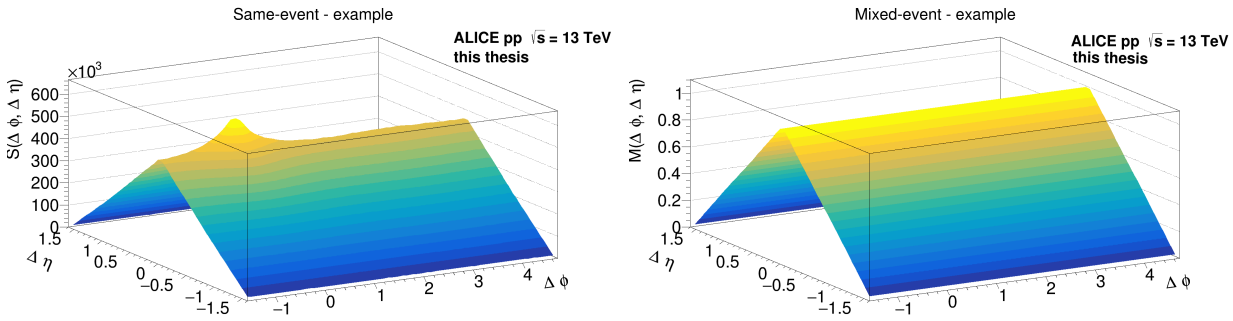
where α normalizes the histogram such that the value at $(\Delta\phi, \Delta\eta) = (0, 0)$ is 1.

Same-event correlations are correlations between tracks that belong to the same event, and mixed-event correlations are the result of correlating tracks from one events with tracks from another event. The same-event correlations are the ones that contain physical information, since correlating particles from different events will not reveal any physics. The reason for dividing the same-event correlations with the mixed-event correlations is due to a phenomenon known as *acceptance effects*, which is explained below.

Acceptance effects

The ALICE detector only detects particles within the pseudorapidity region $\eta = (-0.8, 0.8)$, known as the *acceptance region*. Thus, if the pseudorapidity difference between two particles $(\Delta\eta)$ is high, there is a larger probability that one of the particles is outside this acceptance region. The smaller the $\Delta\eta$, the larger the probability that both of the particles fall within the acceptance region. This will create a triangular appearance in $\Delta\eta$ with a peak at 0, which one wishes to eliminate in order to obtain a purely physical analysis, independent of such detector effects.

When doing correlations between tracks that belong to different events, there will be no physical correlation between these tracks. The only correlation will come from this acceptance effect. Thus, in order to remove the effect from the same-event correlations, one can divide with the mixed-event correlations. Figures 2.1 and 2.2 illustrate this triangular appearance, and how it is removed after the mixed-event division.



(a) Same-event (b) Mixed-event

Figure 2.1: Examples of same-event and mixed-event two-particle angular correlation functions. S_0-inclusive and p_T-inclusive.

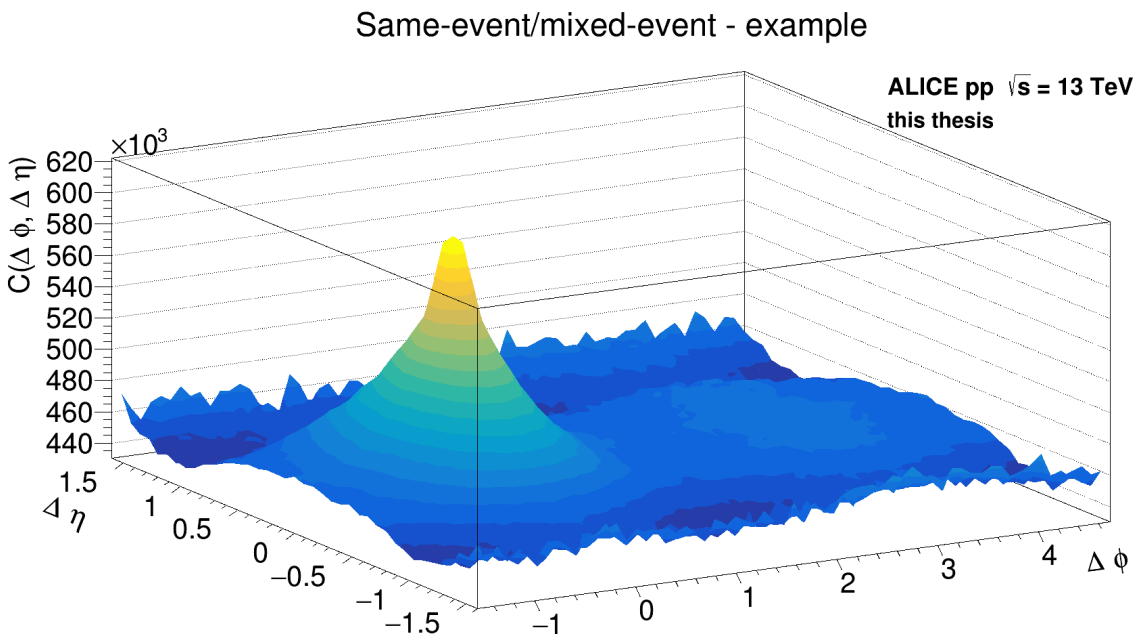


Figure 2.2: Example of two-particle angular correlation function after mixed-event division. S_0-inclusive and p_T-inclusive.

v_z corrections

The detector acceptance effects described in section 2.2.2 have slightly different appearances depending on where in the detector the collision takes place, i.e. depending on the v_z. For instance, the triangular appearance might be slightly shifted towards one side for events with a large v_z. In order for the mixed-event division to completely remove the detector effects from the same-event correlations, the mixed-events one divide with are required to have that same bias. Therefore, the correlations were made in v_z *pools*; each same-event should be divided with a mixed-event that is similar in v_z. Each v_z pool has the size of 2 cm, ranging from -10 to 10 (10 pools in total). Correlations were made in different pools depending on the v_z, and the division was performed for events in the same v_z pool and finally averaged together:

$$C(\Delta\phi, \Delta\eta) = \frac{1}{i} \sum_i \frac{S_i(\Delta\phi, \Delta\eta)}{M_i(\Delta\phi, \Delta\eta)} \tag{2.4}$$

where i denotes the v_z pool, ranging from 1 to 10.

To account for the fact that there might be more or fewer events in some of the pools, one can scale the correlations by the number of trigger tracks in that pool, and then divide the sum with the total number of trigger tracks in order to weight the correlation function correctly:

$$C(\Delta\phi, \Delta\eta) = \frac{1}{N_{trig,tot}} \sum_i \frac{S_i(\Delta\phi, \Delta\eta)}{M_i(\Delta\phi, \Delta\eta)} N_{trig,i} \tag{2.5}$$

This can be simplified as:

$$C(\Delta\phi, \Delta\eta) = \frac{1}{N_{trig,tot}} \sum_i \frac{S_i(\Delta\phi, \Delta\eta)}{M_i(\Delta\phi, \Delta\eta)} \tag{2.6}$$

where $M(\Delta\phi, \Delta\eta)$ is normalized such that the value at $(\Delta\phi, \Delta\eta) = (0,0)$ is 1 and $S_i(\Delta\phi, \Delta\eta)$ is no longer normalized with a factor $\frac{1}{N_{trig,i}}$. This simplification can be made due to the fact that the expression for the same-event correlation already includes a factor of $\frac{1}{N_{trig,i}}$. Therefore, when removing this scaling factor from the same-event correlations these factors cancel each other out. This definition of the correlation function, equation 2.6, is the definition that was used throughout this project.

2.3 Efficiency corrections

The detector is not able to detect all particles that are being created in a collision. The track reconstruction efficiency of the detector is defined as the ratio of reconstructed to produced particles:

$$\epsilon = \frac{N^{recon.}_{particles}}{N^{prod.}_{particles}} \tag{2.7}$$

This efficiency can depend on a number of properties of the track or event: the p_T of the particle, the η of the particle and the v_z of the event. In order for the correlation functions to be as accurate as possible, each track should be scaled by $\frac{1}{\epsilon}$ to account for track reconstruction efficiency.

The track reconstruction efficiencies were corrected for by using MC data generated by PYTHIA, a MC particle physics event generator (see ref. [7]). There were two types of MC data available: the pure generator level MC particles, i.e. the information about the particles that were generated by PYTHIA, as well as the reconstructed tracks after they had gone through a simulation of the ALICE detector. This way one is able to compute the ratio from equation 2.7.

The efficiencies were calculated in three dimensions, where each specific configuration of p_T, η and v_z was given an efficiency, and the filling of the histograms was done accordingly. Figure 2.3 shows the efficiency versus the p_T.

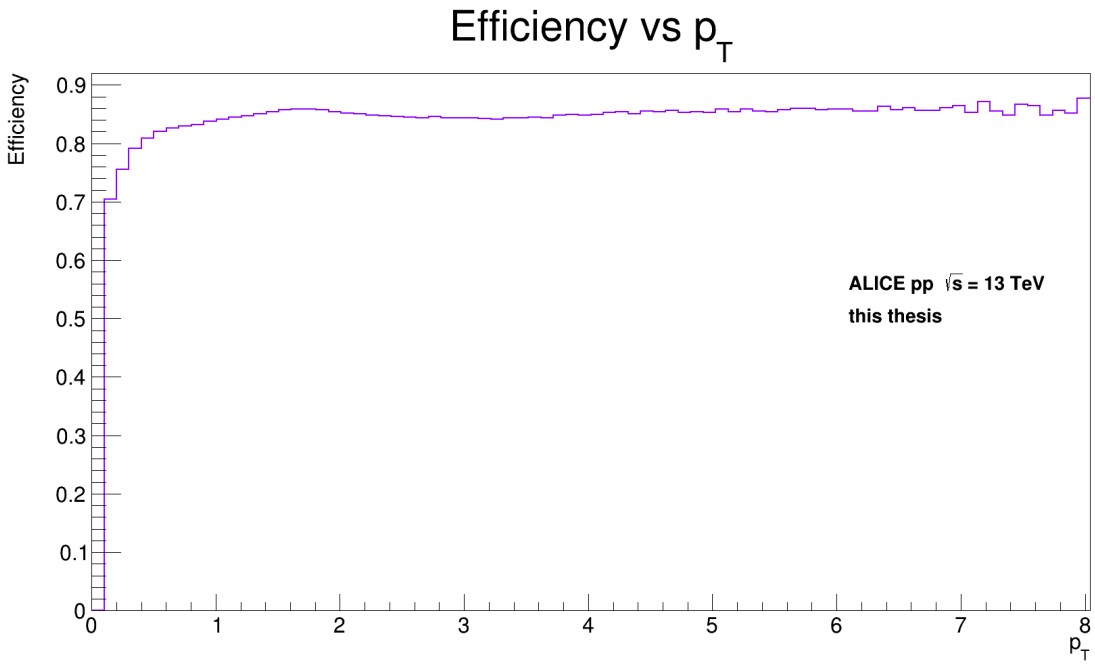
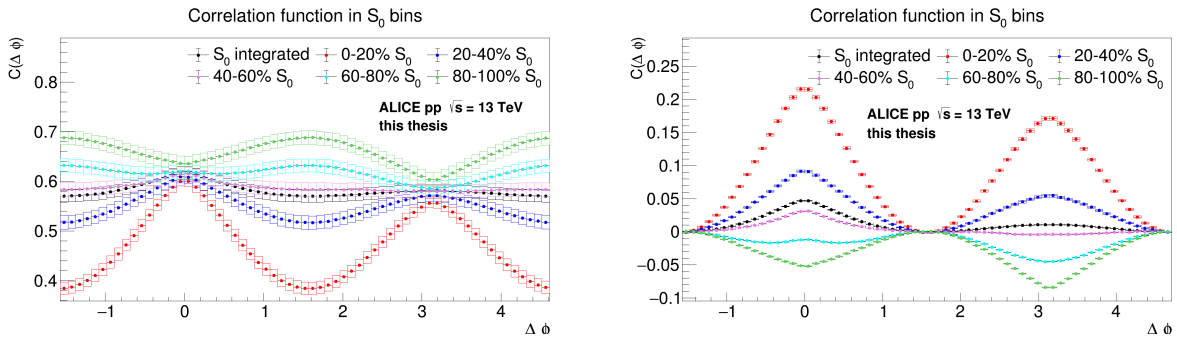


Figure 2.3: Efficiency vs p_T.

2.3.1 ZYAM subtraction

Different correlation functions might reach very different values due to a varying amount of UE, which varies depending on the multiplicity. However, it is the jet shapes one wishes to isolate. Therefore, in order to increase comparability between different correlation functions, *ZYAM subtraction* was performed on the $\Delta\phi$ projections. ZYAM stands for zero yield at minimum. In the case of this analysis it meant taking three histogram points around $\Delta\phi = \frac{\pi}{2}$ and three histogram points around $\Delta\phi = \frac{3\pi}{2}$, computing the average of the histogram values for these six points, and subtracting this average from each bin in the histogram. The goal is to compare the shapes of the correlation functions, and in order to do this it is required to remove the background from each correlation function. This is what is being done by the ZYAM subtraction. Figure 2.4 below shows a correlation function before and after ZYAM subtraction.



(a) Correlation function in S_0 bins without ZYAM subtraction

(b) Correlation function in S_0 bins with ZYAM subtraction

Figure 2.4: Correlation function in S_0 bins before and after ZYAM subtraction, top 10 % V0M multiplicity

2.4 Evaluation of systematic uncertainties

Systematic uncertainties were examined in three categories: systematic uncertainties on the events by examining the effect of changing the v_z range, systematic uncertainties on the tracks by examining the effect of changing the minimum number of clusters required in the TPC (see section 1.4.1), and systematic uncertainties on the correlation functions by examining the effect of changing the number of data points used to define the ZYAM subtraction.

Figures 2.5, 2.6 and 2.7 show plots with the default correlation function in the 40-60 % spherocity bin, with the correlation function with some variable changed (v_z range, number of required TPC clusters, or number of ZYAM points) on top. Below each such plot the ratio is shown. For the uncertainties on the ratio, it is assumed that the uncertainties on the default and the variation are fully correlated. This is not completely accurate, but it is more accurate than the completely uncorrelated uncertainties that ROOT automatically assumes.

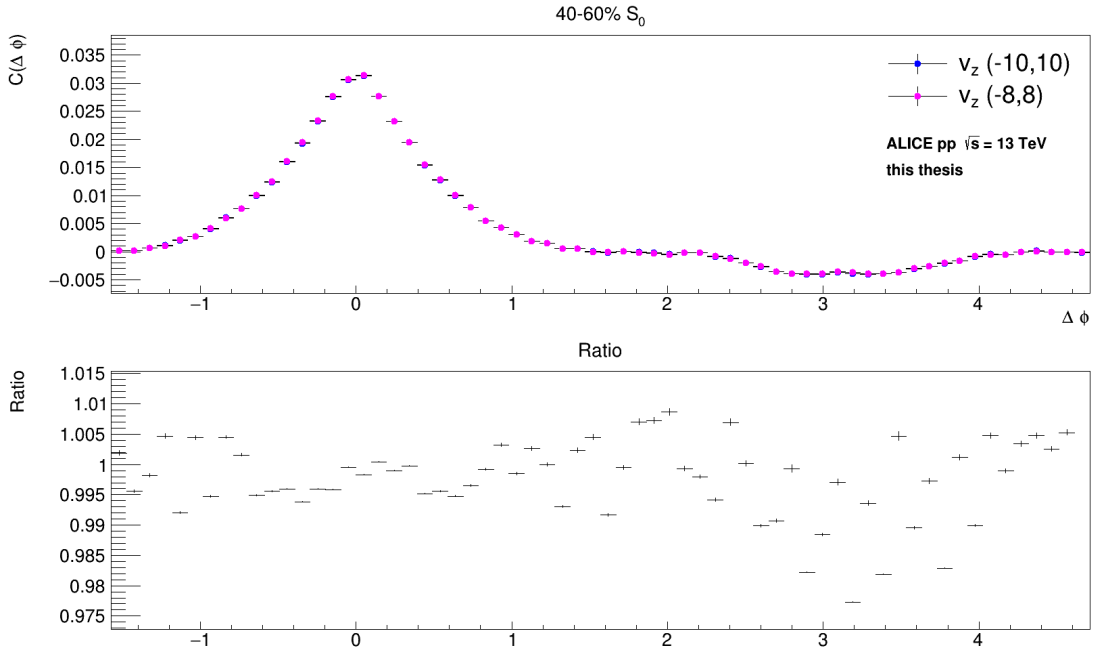


Figure 2.5: $\Delta\phi$ projection of correlation function for 40-60 % S_0, top 10 % V0M multiplicity. $v_z = (-10, 10)$ vs $v_z = (-8, 8)$. Bottom plot shows the ratio.

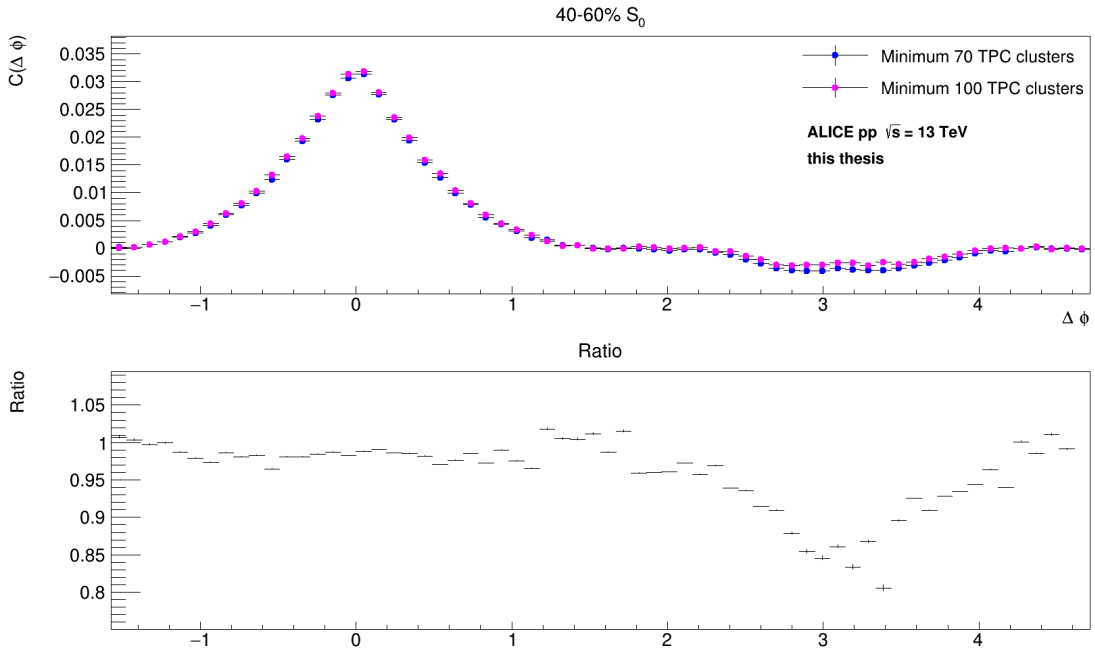


Figure 2.6: $\Delta\phi$ projection of correlation function for 40-60 % S_0, top 10 % V0M multiplicity. Minimum 70 TPC clusters vs minimum 100 TPC clusters. Bottom plot shows the ratio.

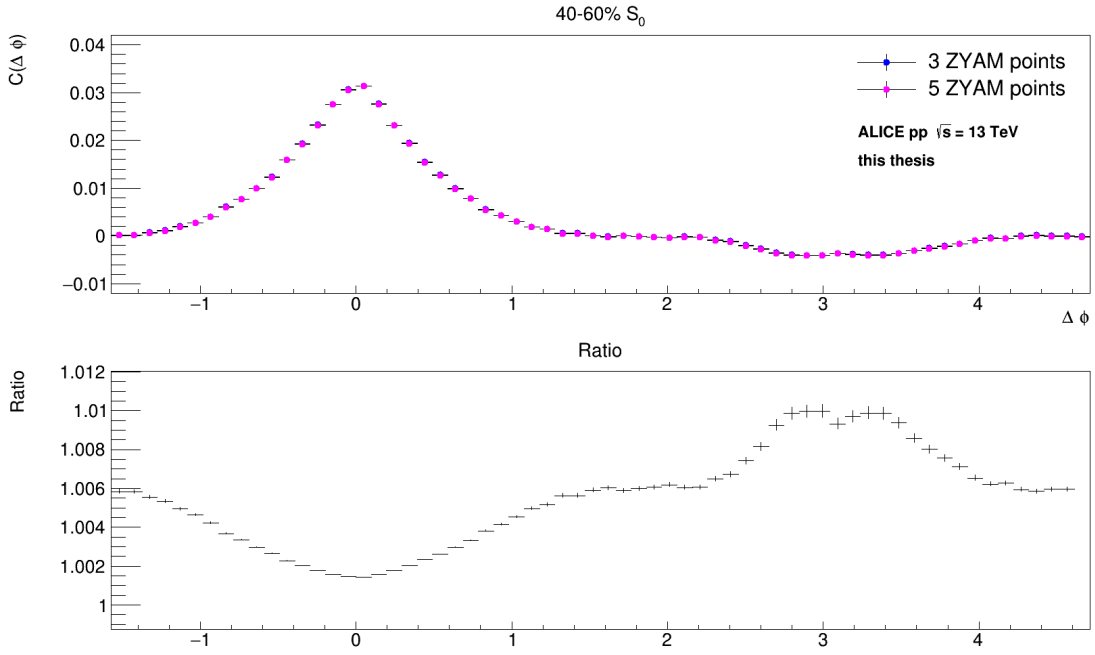


Figure 2.7: $\Delta\phi$ projection of correlation function for 40-60 % S_0, top 10 % V0M multiplicity. 3 ZYAM points vs 5 ZYAM points. Bottom plot shows the ratio.

Systematic error bars were manually added bin-by-bin symmetrically to each correlation function according to:

$$\sigma_{tot} = \sqrt{\sigma^2_{events} + \sigma^2_{tracks} + \sigma^2_{corrs}} \tag{2.8}$$

where σ_{events} is the difference between the two correlation functions when changing the v_z range, σ_{tracks} is the difference when changing the number of required TPC clusters, and σ_{corrs} is the difference when changing the number of points when doing the ZYAM subtraction.

2.5 Summary

For each of the three classifiers, correlation functions (according to equation 2.6) were plotted in percentile bins of the value of each classifier: 0-20 %, 20-40 %, 40-60 %, 60-80 % and 80-100 % on the same canvas. ZYAM subtraction was performed to increase comparability between the shapes of the different classifier bins, and systematic uncertainties were taken into account on three levels: the events, the tracks and the correlation functions. The R_T analysis was performed without any cut on the multiplicity, whereas the spherocity and flattenicity analysis were performed using a cut on the multiplicity, selecting only the top 10 % V0 multiplicities.

For all three classifiers, correlation functions were made in a p_T inclusive configuration ($0.2 < p_T^{assoc.} < p_T^{trig.} < 5$ GeV/c) as well as two configurations in p_T bins: high-p_T trigger track (3-4 GeV/c) vs low-p_T associate (0.5-1 GeV/c) and high-p_T trigger track vs high-p_T associate. The reason for doing correlations in p_T bins is that when looking at the p_T inclusive correlations, it is biased towards low-p_T trigger track versus low-p_T associate, since

this is the most common configuration. Selecting specific p_T bins allows for probing the hard and soft components of the jet.

For R_T, two additional configurations were made: correlations between tracks and the R_T trigger track (the leading track with $p_T > 5$ GeV/c) for that event, and correlations between tracks within the transverse region. For spherocity, one additional configuration was made: correlations between tracks and the unit vector used to calculate the spherocity for that event.

Chapter 3

Results and discussion

In this chapter the results are presented and discussed, first for R_T, then for spherocity and finally for flattenicity. The configurations described in section 2.5 are presented as well as the distributions for each classifier. Furthermore, some statistics are presented for each classifier in tables. This is followed by a summarized comparison between the classifiers.

3.1 R_T

Figure 3.1 shows the R_T distribution. It is discrete because N_T is an integer (an event can only have a discrete number of tracks in the transverse region). The average value is 1 but the most common value is below 1.

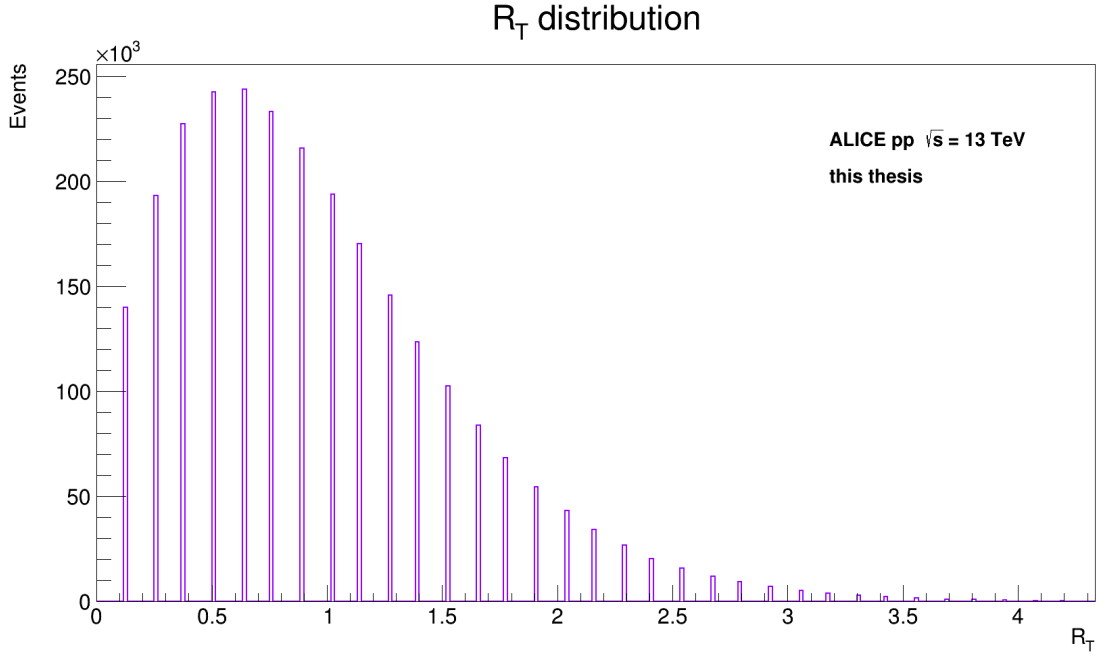


Figure 3.1: R_T distribution. Multiplicity-integrated

Table 3.1 displays some statistics for R_T. The meanings of $N_{T,max}$ and $N_{T,min}$ are explained

in section 3.1.3. The R_T bins are displayed in 3.1a, and 3.1b displays some average values. It can be seen that the average number of tracks per event (19.115) is significantly higher than the average number of tracks per events for the multiplicity-integrated S_0 and ρ configurations (see tables 3.3 and 3.5), indicating that requiring a $p_T > 5$ GeV/c leading track biases R_T towards higher-multiplicity events.

Percent bin	R_T
20 %	0.355
40 %	0.710
60 %	1.065
80 %	1.420

(a) The value of the R_T bin edges

Average number of tracks per event	19.115
$< N_T >$	5.634
Average p_T for R_T trigger track	6.680
Average $N_{T,max}$	4.127
Average $N_{T,min}$	1.665

(b) Table of average values related to R_T

Table 3.1: R_T statistics

3.1.1 Configuration 1: p_T inclusive

Figure 3.2 shows the p_T-inclusive correlation function for different R_T bins. It can be seen that both the near-side and the away-side peaks are significantly pronounced for the lowest R_T bin (0-20 %). As the R_T increases the events become less jet-like, and the peaks become less pronounced. This can be explained by the fact that when requiring a lot of tracks in the transverse region, i.e. perpendicular to the jet axis (higher R_T) the UE dominates, whereas when requiring few tracks in the transverse region (lower R_T) jet-like events dominate.

For the two highest R_T bins the away-side has almost disappeared completely. The reason for this is unclear, but it could be due to jet quenching effects, which is an effect that has not yet been observed in small systems. Moreover, it could be explained by kinematics, for example if the away-side jet falls outside of the detector or if three-jet events are being selected. To investigate this one could make the same measurements with some model (for example PYTHIA); if the effect originates from jet quenching it would not be seen in PYTHIA, and if it originates from the kinematics it would be seen in PYTHIA.

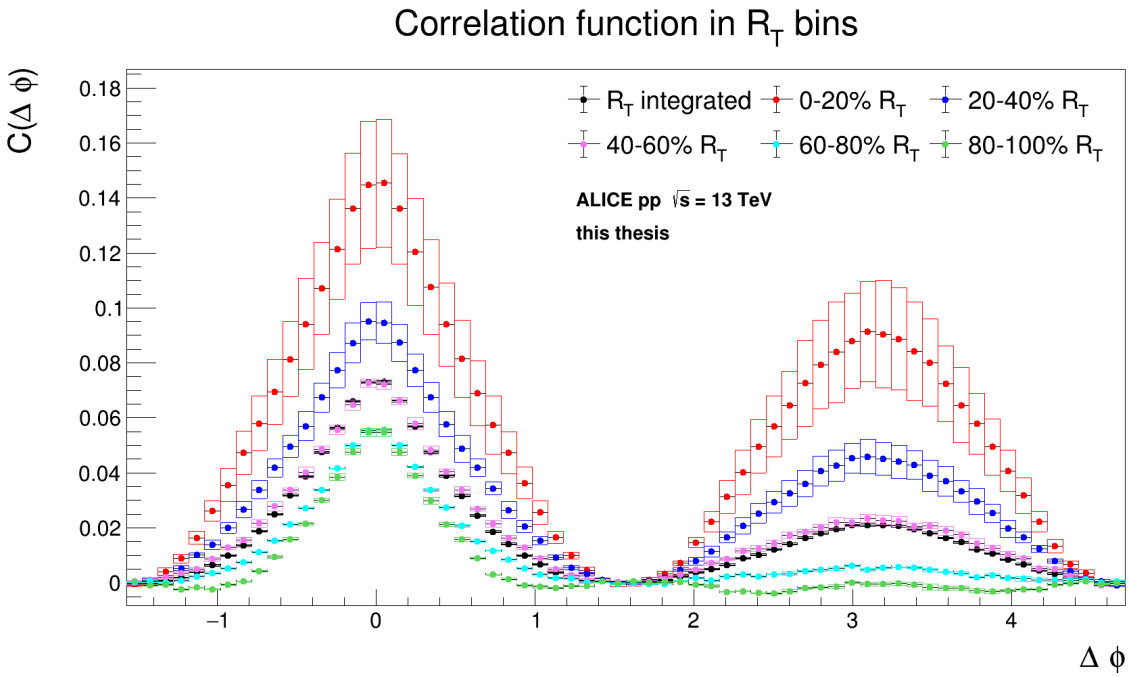


Figure 3.2: $\Delta\phi$ projection of p_T inclusive correlation function in R_T bins. Multiplicity-integrated. With ZYAM subtraction.

3.1.2 Configuration 2: p_T bins

Figure 3.3 shows the correlation function for a high-p_T (3-4 GeV/c) trigger track and a low-p_T (0.5-1 GeV/c) associate. No clear qualitative differences from the p_T-inclusive correlation function can be seen except the decreased statistics.

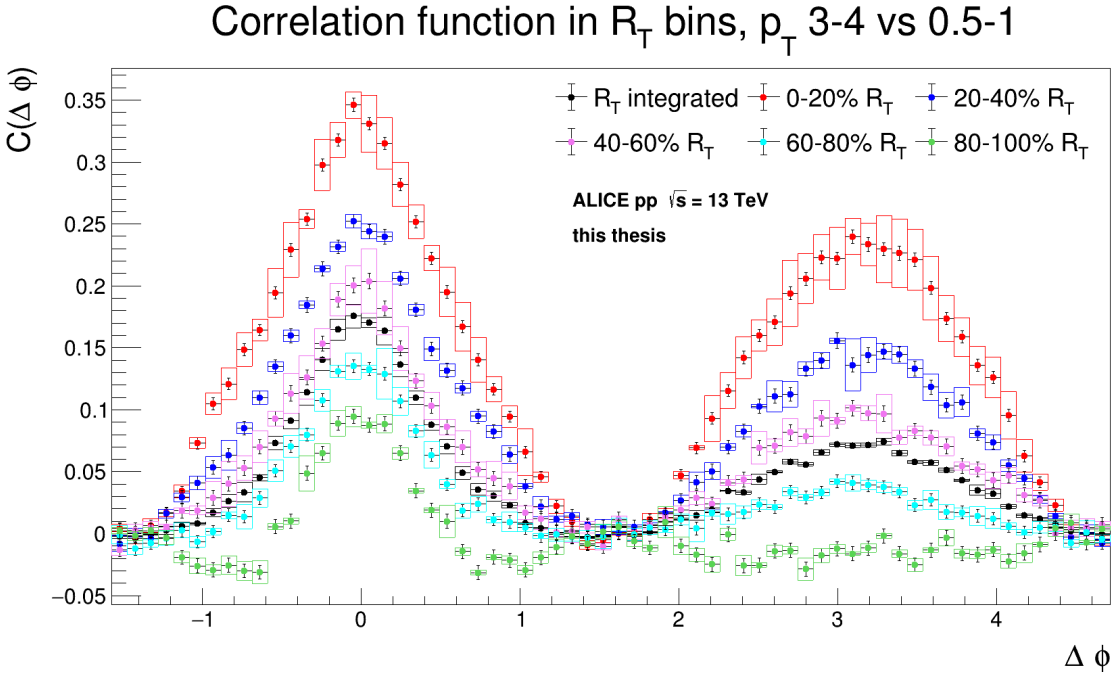


Figure 3.3: $\Delta\phi$ projection of p_T 3-4 GeV/c vs 0.5-1 GeV/c correlation function in R_T bins. Multiplicity-integrated. With ZYAM subtraction.

Figure 3.4 shows the correlation function for a high-p_T trigger track and a high-p_T associate. The order of the correlations for the different R_T bins remains. However, the difference between the low and high R_T bins is smaller due to the requirement of both tracks being high-p_T, shifting the selection towards more jet-like events. The narrowing of the near-side peak at $\Delta\phi = 0$ occurs due to conservation of momentum, which leads to particle production at high-p_T being dominated by particles being produced with small angles between them.

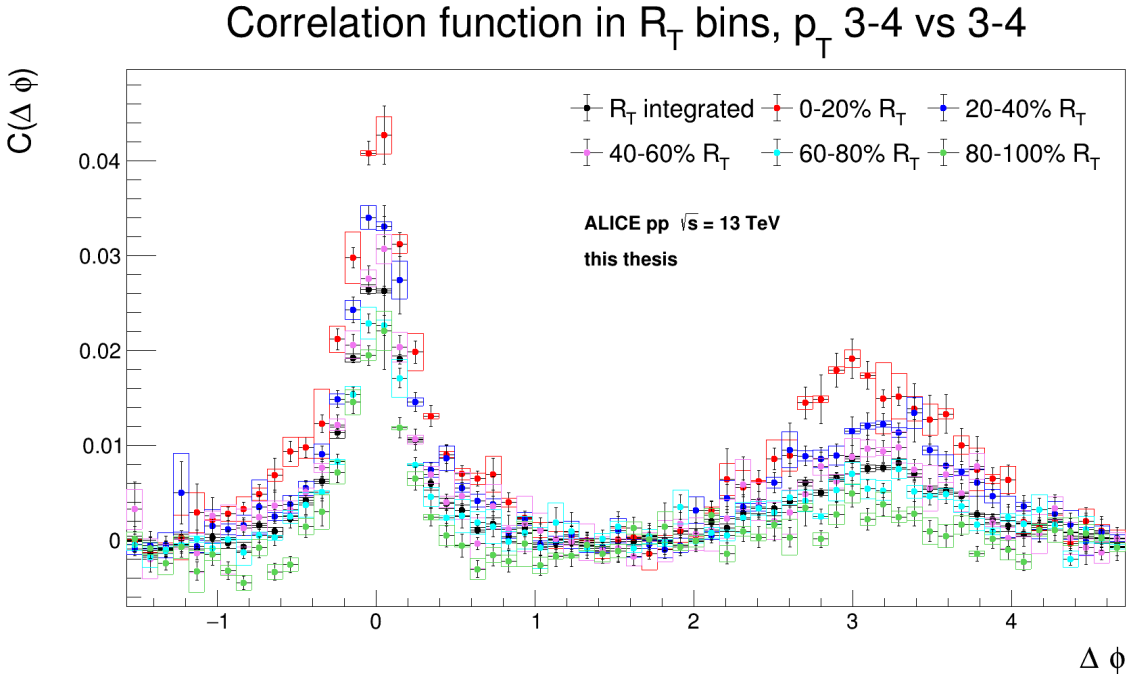


Figure 3.4: $\Delta\phi$ projection of p_T 3-4 GeV/c vs 3-4 GeV/c correlation function in R_T bins. Multiplicity-integrated. With ZYAM subtraction.

3.1.3 Configuration 3: w.r.t. R_T trigger track

Figure 3.5 shows the $\Delta\phi$ correlation function with respect to the $p_T > 5$ GeV/c leading track that is required in the R_T calculations. It can be seen that there are sharp steps at the edges of the transverse regions $\Delta\phi = (\frac{\pi}{3}, \frac{2\pi}{3})$ and $\Delta\phi = (\frac{4\pi}{3}, \frac{5\pi}{3})$, with a high track density in high-R_T events a low track density in low-R_T events. This can be explained by the fact that when selecting events with high R_T, events with a lot of tracks in that specific transverse region are being selected. The same argument is used for explaining the dips for the low-R_T events.

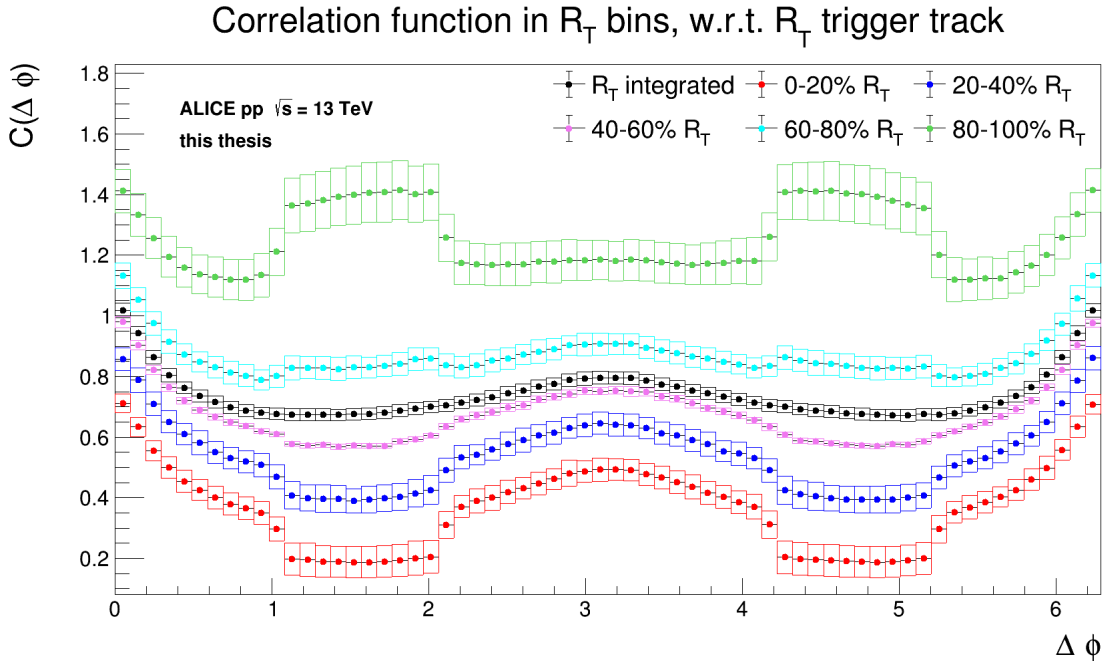


Figure 3.5: $\Delta\phi$ projection of correlation function in R_T bins, w.r.t. the R_T trigger track. Multiplicity-integrated. Without ZYAM subtraction.

It can be seen that there is a slope towards π in the transverse region, especially visible in the 80-100 % R_T bin. This slope might be explained in the following way: if a three-jet event occurs, i.e. a hard gluon is radiated from the jet, it could fragment and its products could make it inside the transverse region. The closer to the direction the jet axis, the higher the probability that the fragmentation products of the recoil gluon reached that area. Thus a slope appears favoring particles near that axis.

In order to further investigate this effect, correlation functions were made with respect to the R_T trigger track with a method known as $N_{T,min/max}$. In this method, one defines for each event the $N_{T,min}$ and $N_{T,max}$ regions. The $N_{T,min}$ region is the side of the transverse region that contains the least number of tracks, and the $N_{T,max}$ region is the side that contains the most tracks. The $N_{T,max}$ region will likely be the region that contains the products of the recoil gluons. The method used fills the histograms such that the $N_{T,max}$ region is the first transverse region ($\Delta\phi = (\frac{\pi}{3}, \frac{2\pi}{3})$), and the $N_{T,min}$ region is the second transverse region ($\Delta\phi = (\frac{4\pi}{3}, \frac{5\pi}{3})$). Figure 3.6 below displays this correlation function.

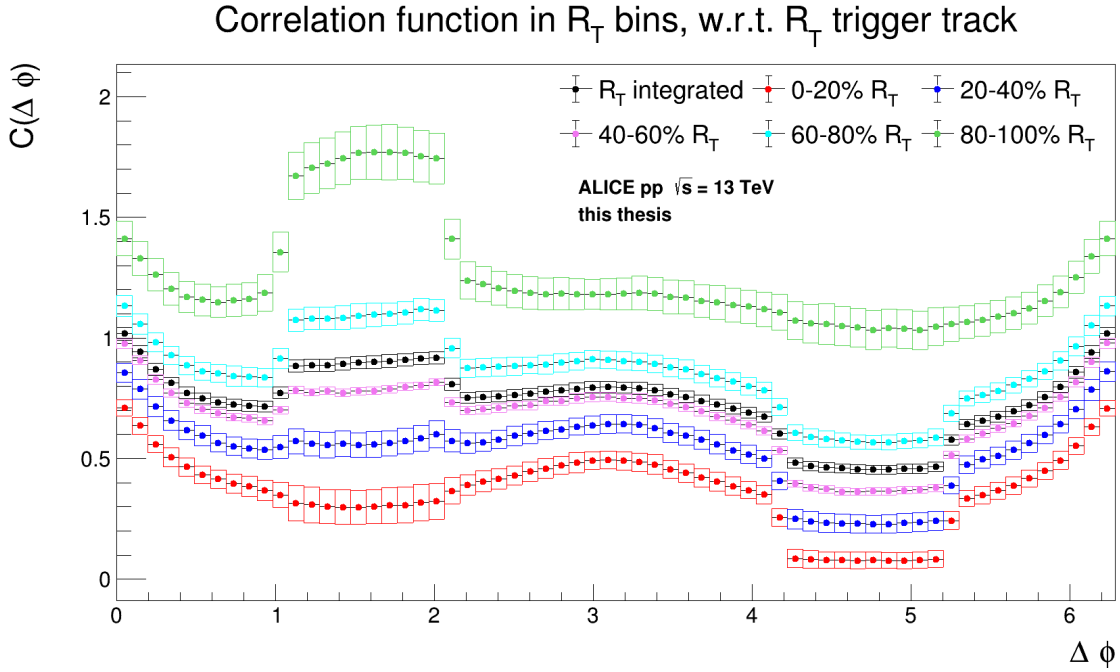


Figure 3.6: $\Delta\phi$ projection of correlation function in R_T bins, w.r.t. R_T trigger track, $N_{T,min/max}$ method. Multiplicity-integrated. Without ZYAM subtraction.

As expected, there are a lot more tracks in the first transverse region than the second one. It can be seen that there is a clear slope in the first transverse region, as expected, and this slope is not clearly visible in the second transverse region.

3.1.4 Configuration 4: within the transverse region

Figure 3.7 shows the correlation function when only correlating tracks in the transverse region. It can be seen that even though the correlations were exclusively made within the transverse region (the region perpendicular to the jet axis) there are still correlations (see the peak at $\Delta\phi = 0$, especially visible in the 80-100 % R_T bin). This suggests that the transverse region that R_T focuses on for its selection of isotropic events does not only contain the UE but also contains some jet components. The reason that there are areas with no correlations is because these are the regions where the angular difference $\Delta\phi$ is within a range such that it is not possible that both tracks are in the transverse region.

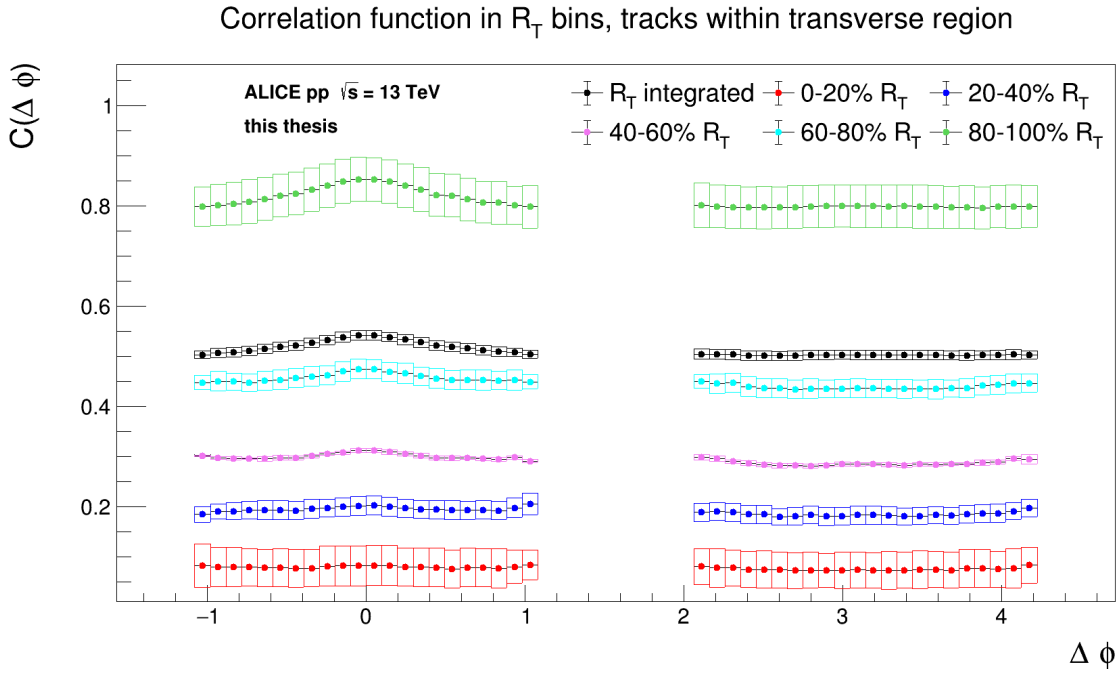


Figure 3.7: $\Delta\phi$ projection of correlation function for tracks within the transverse region, R_T inclusive. Multiplicity-integrated. Without ZYAM subtraction.

3.2 Spherocity

Note that all correlation functions, measured with the p_T-weighted spherocity given by equation 1.3, are displayed in Appendix A. Figure 3.8 shows the spherocity distribution. As expected, it can be observed that the values lie between 0 and 1, with significant drop-offs at the edges, displaying how highly spherical and highly non-spherical events are more uncommon than intermediate spherocities. This distribution is in line with previously measured distributions (see for example [8]).

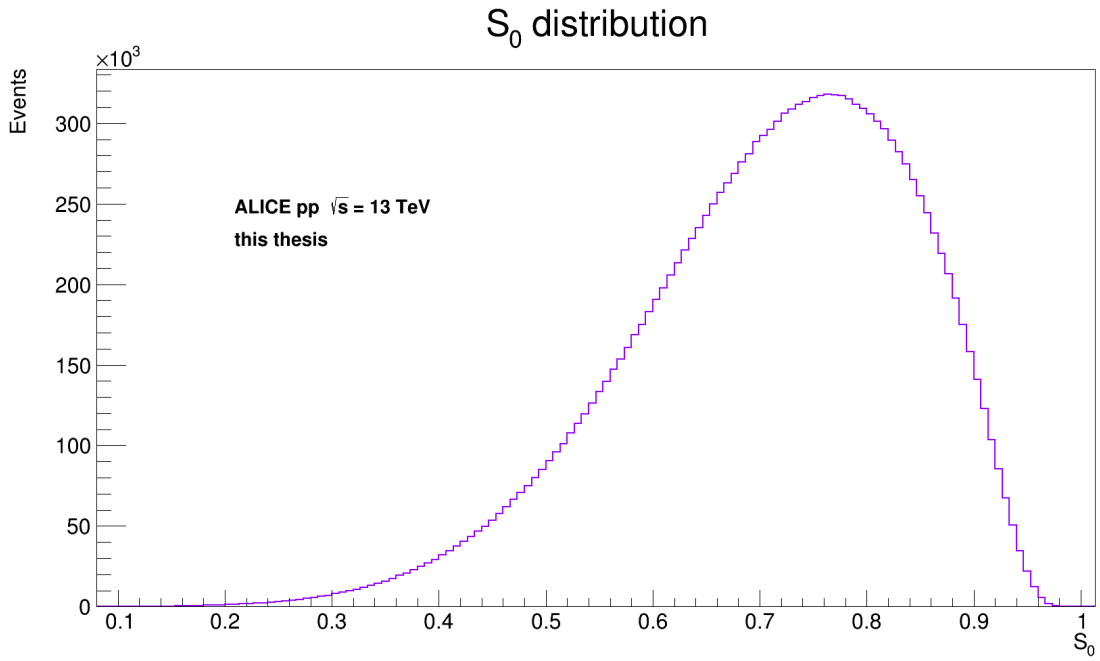


Figure 3.8: Spherocity distribution for top 10 % V0M multiplicities.

Tables 3.2a and 3.2b show the value of the spherocity in percentile bins with and without the V0M multiplicity cut, 20 % meaning the bottom 20 % and 80 % meaning the top 20 %. The values of the S_0 bin edges are higher when including a multiplicity cut than without the cut. This means that high-multiplicity events are in general more spherical than low-multiplicity events, which is expected. As mentioned in section 2.1, this is why a multiplicity cut was implemented: to be able to vary the spherocity without at the same time varying the multiplicity significantly.

Percent bin	S_0
20 %	0.096
40 %	0.292
60 %	0.450
80 %	0.606

(a) The value of S_0 percentile bin edges, multiplicity-integrated

Percent bin	S_0
20 %	0.432
40 %	0.550
60 %	0.646
80 %	0.743

(b) The value of S_0 percentile bin edges, top 10 % V0M multiplicity

Table 3.2: The value of S_0 percentile bin edges

Table 3.3 shows the average spherocity and the average number of tracks per event, with and without the multiplicity cut.

	Multiplicity-integrated	Top 10 % V0M multiplicity
Average S_0	0.366	0.583
Average number of tracks per event	7.836	76.64

Table 3.3: Table of average values related to S_0

3.2.1 Configuration 1: p_T inclusive

Figure 3.9 displays the $\Delta\phi$ projection of the p_T-inclusive correlation function in S_0 bins. It can be seen that both the near-side and away-side peaks are highly pronounced for the lowest spherocities, as these events are the most jet-like events. As the spherocity decreases, the peaks become less pronouced. For the most spherical events, a phenomenon can be observed where peaks appear at $\frac{\pi}{2}$ and $\frac{3\pi}{2}$. This phenomenon can be explained by the fact that when spherocity is attempting to select the most spherical events, it is selecting events where the tracks have an average angle of 90 degrees between them.

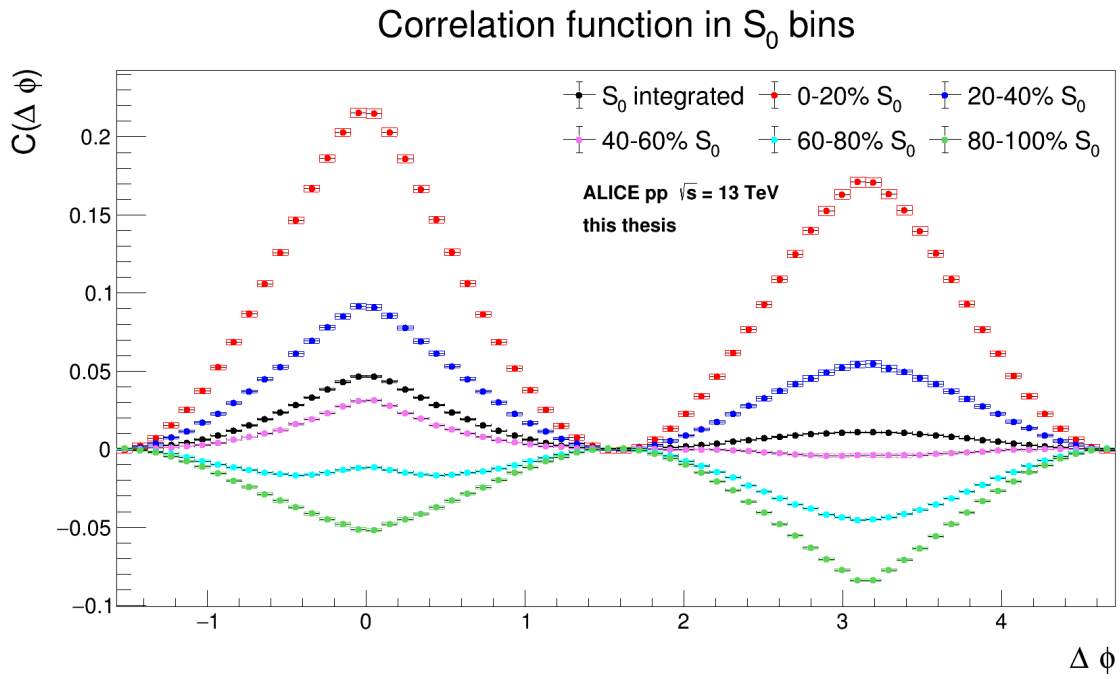


Figure 3.9: $\Delta\phi$ projection of p_T-inclusive correlation function in S_0 bins. Top 10 % V0M multiplicity. With ZYAM subtraction.

3.2.2 Configuration 2: p_T bins

Figure 3.10 shows the $\Delta\phi$ projection of the 3-4 GeV/c trigger p_T vs 0.5-1 GeV/c associate p_T correlation function in S_0 bins. This correlation function is very similar to the p_T-inclusive correlation function (figure 3.9), with a structure starting to appear for the most spherical events at $\Delta\phi = 0$.

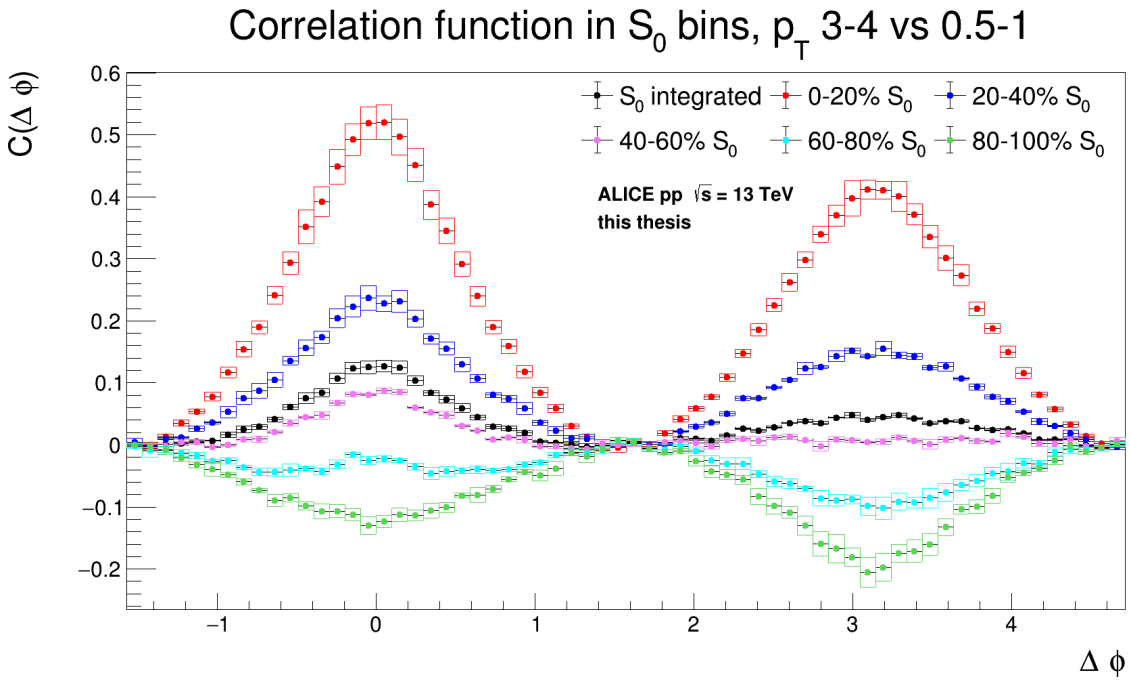


Figure 3.10: $\Delta\phi$ projection of p_T 3-4 GeV/c vs 0.5-1 GeV/c correlation function in S_0 bins. Top 10 % V0M multiplicity. With ZYAM subtraction.

Figure 3.11 shows the $\Delta\phi$ projection of the p_T 3-4 GeV/c vs 3-4 GeV/c correlation function in S_0 bins. This correlation function clearly displays sharper peaks than the previous ones. This is due to the fact that high-p_T particle production is dominated by particles being produced with a small angle between them due to conservation of momentum, which was also observed in the R_T correlation function, see figure 3.4. Furthermore, it can be seen that the structure at $\Delta\phi = 0$ for the most spherical events becomes even more pronounced. This can be explained by the fact that even highly spherical events can contain jets. When requiring both tracks to have high p_T, there are no longer peaks at $\Delta\phi = \frac{\pi}{2}$ and $\frac{3\pi}{2}$ for the highest S_0 bin, since this track selection is essentially requiring the presence of a jet.

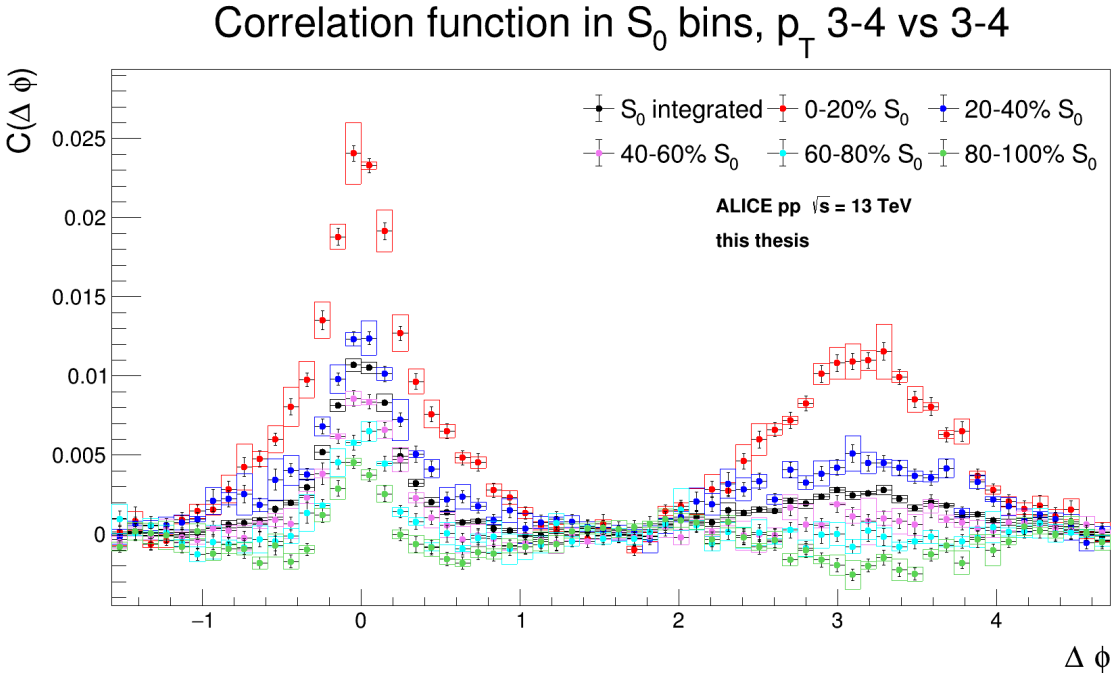


Figure 3.11: $\Delta\phi$ projection of p_T 3-4 GeV/c vs 3-4 GeV/c correlation function in S_0 bins. Top 10 % V0M multiplicity. With ZYAM subtraction.

3.2.3 Configuration 3: w.r.t. S_0 unit vector

Figure 3.12 shows the correlation function for correlations between a track and the unit vector used for the spherocity calculations, see equation 1.4. It can be seen that for all spherocity bins there is a peak at $\Delta\phi = 0$ and π. This peak is significantly broader for the lower spherocity bins, but it is still present for the higher spherocities, indicating some jet contribution from these events. The bins at 0 reach very high values. This indicates that the spherocity unit vector is always one of the tracks, which leads to an auto-correlation, making the bin at 0 very high.

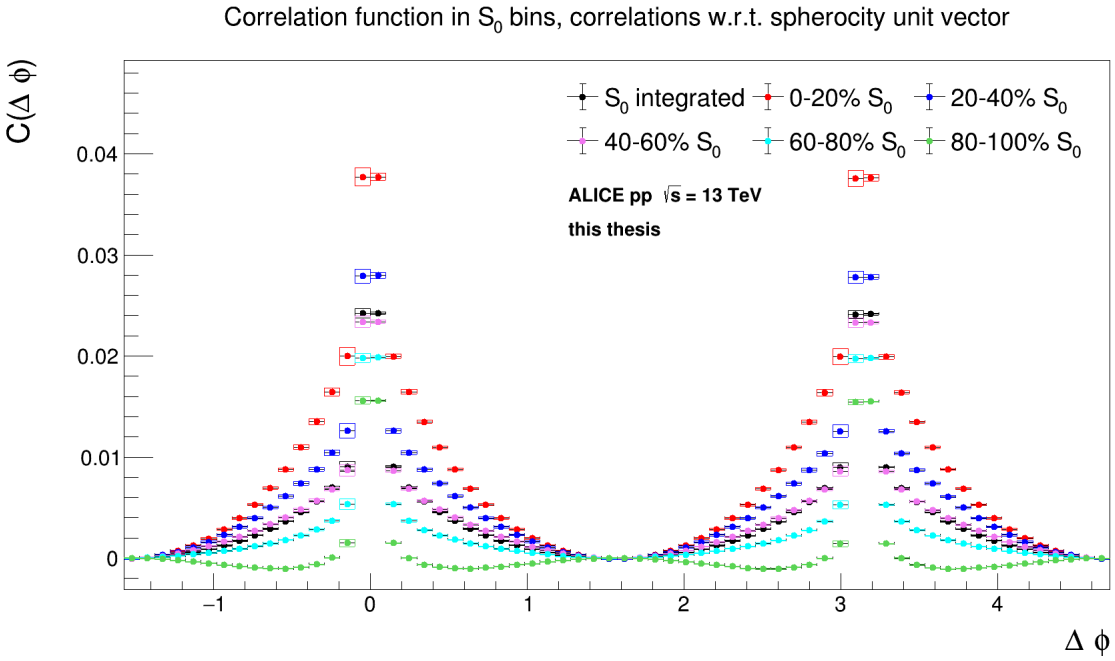


Figure 3.12: p_T-inclusive correlations between a track and the spherocity unit vector used to calculate the spherocity of that event. Top 10 % V0M multiplicity. With ZYAM subtraction.

3.3 Flattenicity

For this section, note that flattenicity scales as the inverse of R_T and S_0, i.e. high flattenicity corresponds to jet-like events whereas high R_T and S_0 corresponds to isotropic events. Note that all correlation functions, measured with the p_T-weighted flattenicity given by equation 1.5, are displayed in Appendix B.

Figure 3.13 shows the flattenicity distribution, with $1 - \rho$ instead of ρ to increase comparability with the spherocity distribution (figure 3.8). It is similar to the spherocity distribution; however, the range is smaller with values ranging from around 0.75 to around 0.95, whereas the spherocity values range from around 0.15 to just below 1. However, a direct comparison between flattenicity and spherocity can still be made by dividing the event sample into percentile bins.

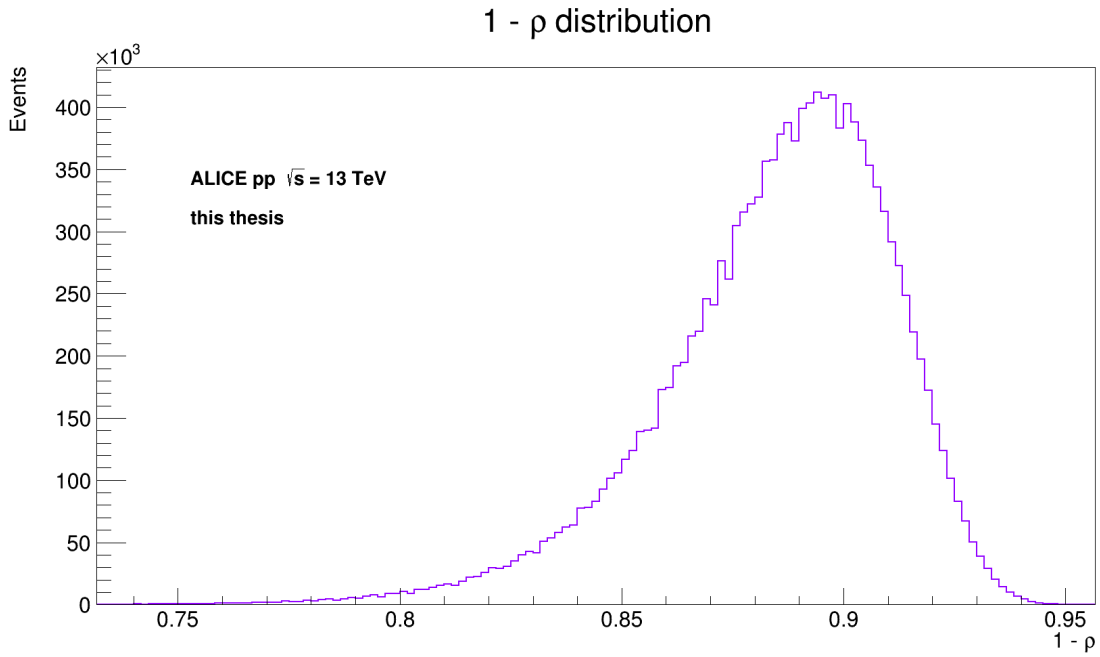


Figure 3.13: Flattenicity distribution for top 10 % V0M multiplicities.

Tables 3.4a and 3.4b below show the values of the flattenicity bins with and without the multiplicity cut. Similarly to spherocity, the events become more isotropic as the multiplicity increases.

Percent bin	ρ
20 %	0.140
40 %	0.180
60 %	0.233
80 %	0.335

(a) The value of ρ percentile bin edges, multiplicity-integrated

Percent bin	ρ
20 %	0.103
40 %	0.117
60 %	0.132
80 %	0.154

(b) The value of ρ percentile bin edges, top 10 % V0M multiplicity

Table 3.4: The value of ρ percentile bin edges

Table 3.5 below show the average flattenicity, average number of tracks per event, and the average value of some of the variables from the flattenicity equations, see equations 1.5 and 1.6, with and without multiplicity cut.

	Multiplicity-integrated	Top 10 % V0M multiplicity
Average ρ	0.248	0.131
Average number of tracks per event	7.836	76.64
Average $< p_{T,cell} >$	0.069	0.195
Average $< N_{ch}^{cell} >$	0.443	1.118

Table 3.5: Table of average values related to ρ

3.3.1 Configuration 1: p_T inclusive

Figure 3.14 shows the $\Delta\phi$ projection of the p_T inclusive correlation function in ρ bins. Similarly to R_T and S_0, the near-side peak at $\Delta\phi = 0$ becomes less and less pronounced the lower the flattenicity. However, there appears to be almost no flattenicity-dependence in the appearance of the away-side peak at $\Delta\phi = \pi$, unlike for R_T and S_0 where this peak decreased in size in a similar manner to the near-side peak.

Furthermore, for the lowest flattenicity bin (0-20 %), there is a clear dip around $\Delta\phi = 0$. This dip has a width of around the size of a cell in ϕ in the definition of the flattenicity (see equation 1.6). This can be explained by flattenicity being defined in a way that makes low flattenicity (highly isotropic events) biased against events where more than one track is in the same cell. When $\Delta\phi$ becomes smaller than the size of a cell there is a probability of the particles being in the same cell, hence the dip. As $\Delta\phi$ decreases this probability becomes larger and larger, and thus the dip becomes lower and lower.

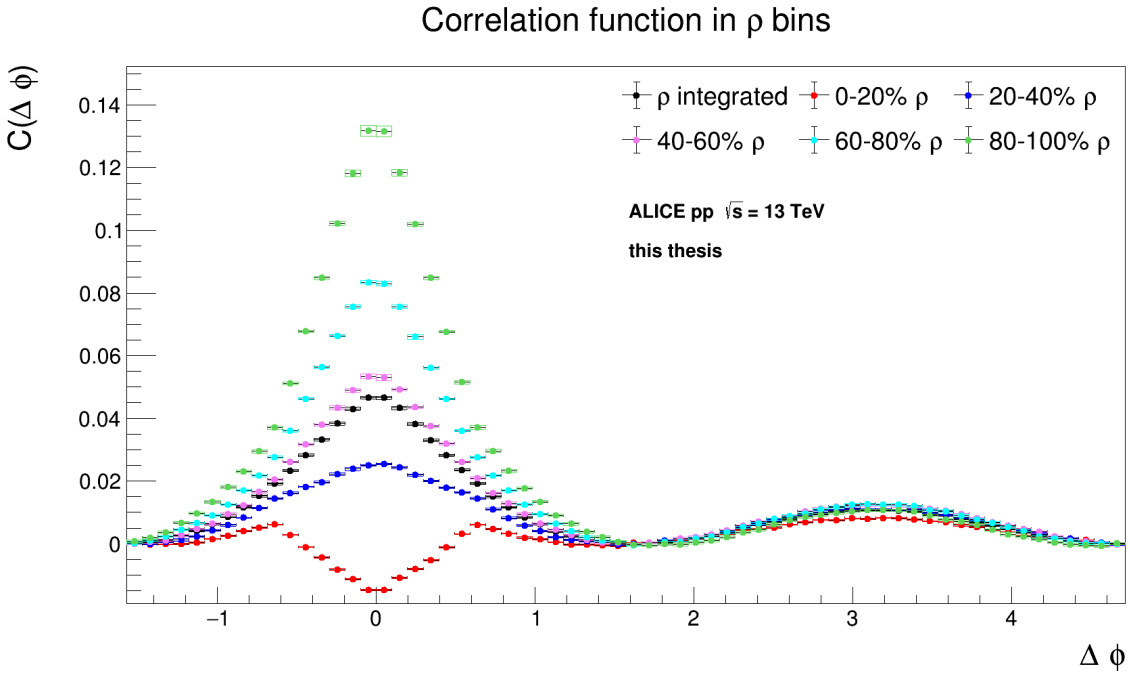


Figure 3.14: $\Delta\phi$ projection of p_T-inclusive correlation function in ρ bins. Top 10 % V0M multiplicity. With ZYAM subtraction.

3.3.2 Configuration 2: p_T bins

Figure 3.15 below shows the correlation function in ρ bins when correlating a high-p_T (3-4 GeV/c) particle with low-p_T (0.5-1 GeV/c) particle. The correlation function looks similar to the p_T inclusive correlation function, but one clear difference is the smoothing of the dip for the lowest flattenicity bin (0-20 %) at $\Delta\phi = 0$.

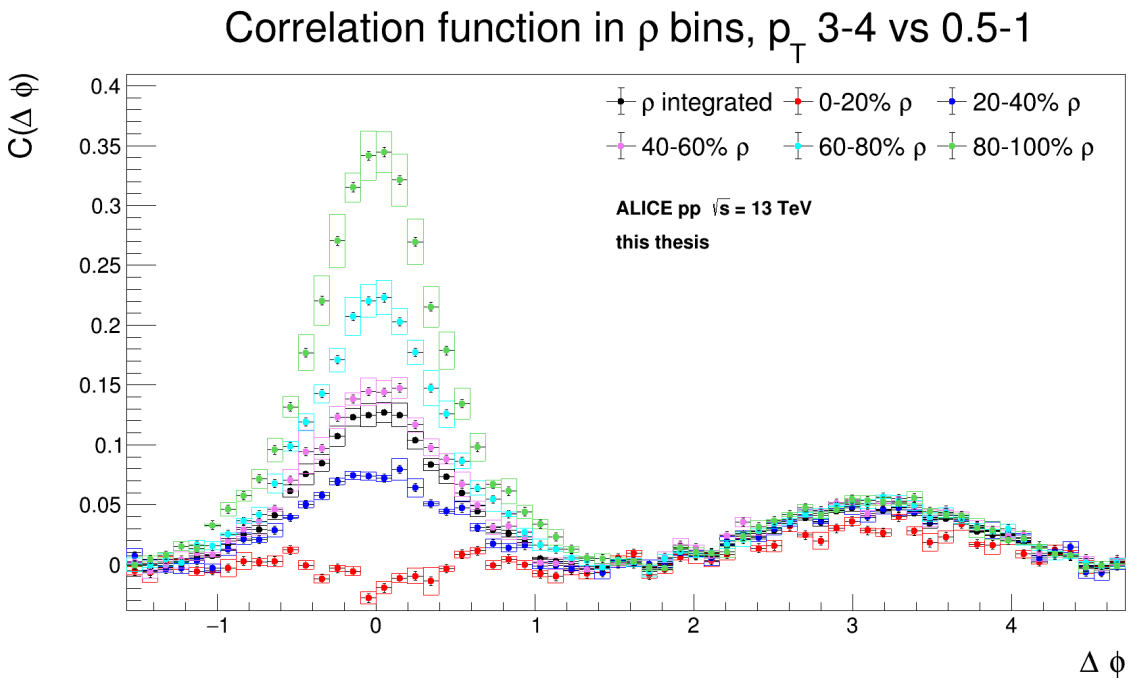


Figure 3.15: $\Delta\phi$ projection of p_T 3-4 GeV/c vs 0.5-1 GeV/c correlation function in ρ bins. Top 10 % V0M multiplicity. With ZYAM subtraction.

Figure 3.16 shows the correlation function in ρ bins when correlating a high-p_T particle with another high-p_T particle. In this correlation function the peaks at $\Delta\phi = 0$ have narrowed compared to the two previous correlation functions. This can be explained in the same way as it was explained for R_T and spherocity, see for example section 3.2.2.

Furthermore, the dip for the lowest flattenicity bin at $\Delta\phi = 0$ has disappeared completely. This can be explained similarly to how it was expalined why the high-p_T vs high-p_T correlation function in the S_0 analysis removed the peaks at $\Delta\phi = \frac{\pi}{2}$ and $\frac{3\pi}{2}$: this track selection is essentially requiring the presence of a jet.

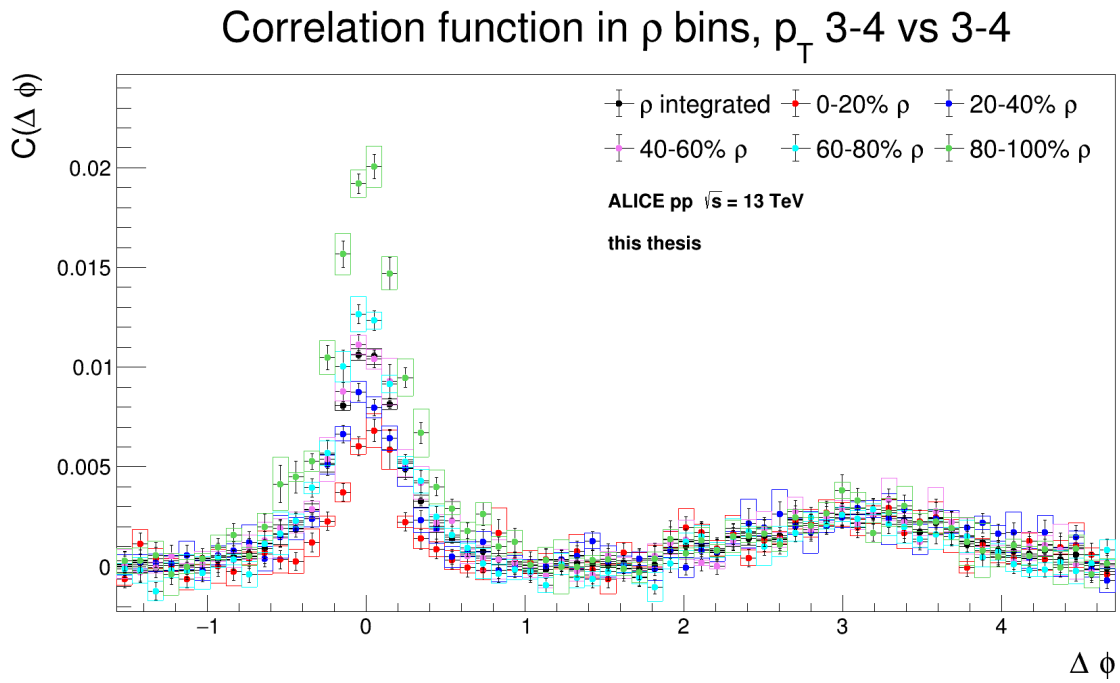


Figure 3.16: $\Delta\phi$ projection of p_T 3-4 GeV/c vs 3-4 GeV/c correlation function in ρ bins. Top 10 % V0M multiplicity. With ZYAM subtraction.

3.4 Comparisons

Below two figures are shown that display the correlation function for the same classifier bin for all three classifiers. Note that all the correlation functions below were measured with the V0M multiplicity cut. However, due to the requirement of the $p_T > 5$ GeV/c leading track in the R_T analysis, it is important to note that the event selection for R_T is not identical to the event selection for spherocity and flattenicity.

Figure 3.17 shows a comparison between the most isotropic bin for R_T, spherocity and flattenicity. It can be seen that a low flattenicity selects much flatter events than a high R_T and spherocity. Spherocity favors events where the tracks have an average angle of 90 degrees between them. R_T selects events with a low near-side peak and no away-side peak. While flattenicity selects flatter events than both of the other two classifiers, it is biased against events where two or more tracks are in the same flattenicity cell.

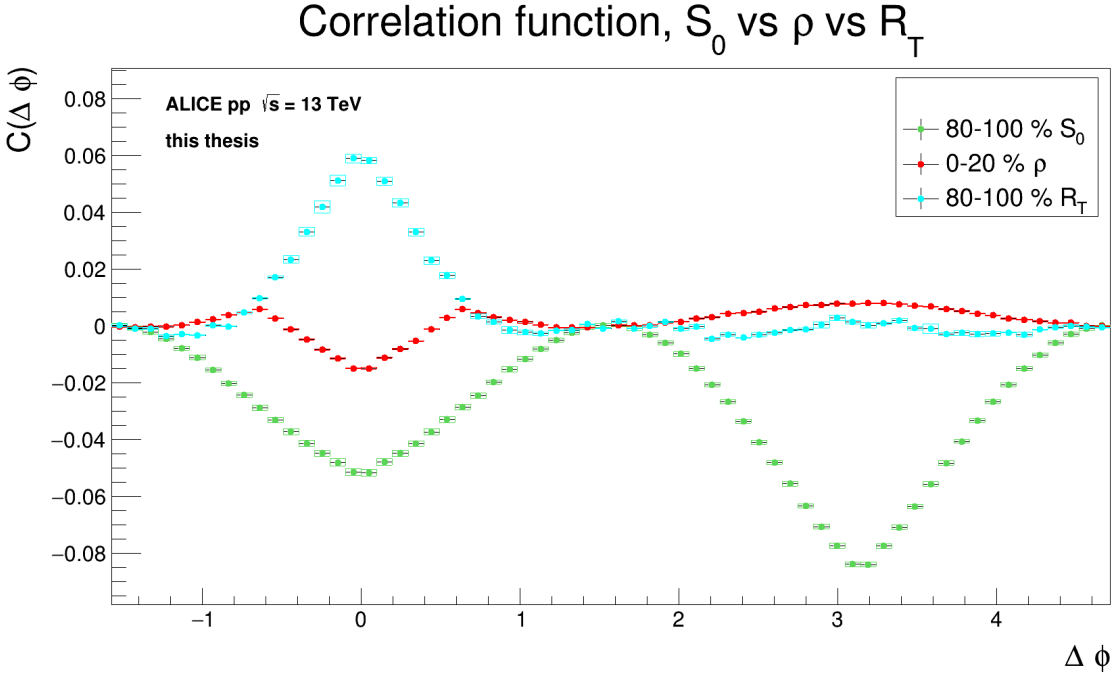


Figure 3.17: p_T-inclusive correlation functions for most isotropic R_T, S_0 and ρ bins. Top 10 % V0M multiplicity. With ZYAM subtraction.

Figure 3.18 shows a comparison between the most jet-like bin for R_T, spherocity and flattenicity. A low spherocity selects more jet-like events than both a low R_T and a high flattenicity. The 0-20 % spherocity bin has a clear near-side and away-side jet peak. The R_T also has a clear near-side and away-side peak but is lower than the spherocity. The 80-100 % flattenicity bin has a clear near-side peak which is significantly lower than the spherocity near-side peak, and a very small away-side peak.

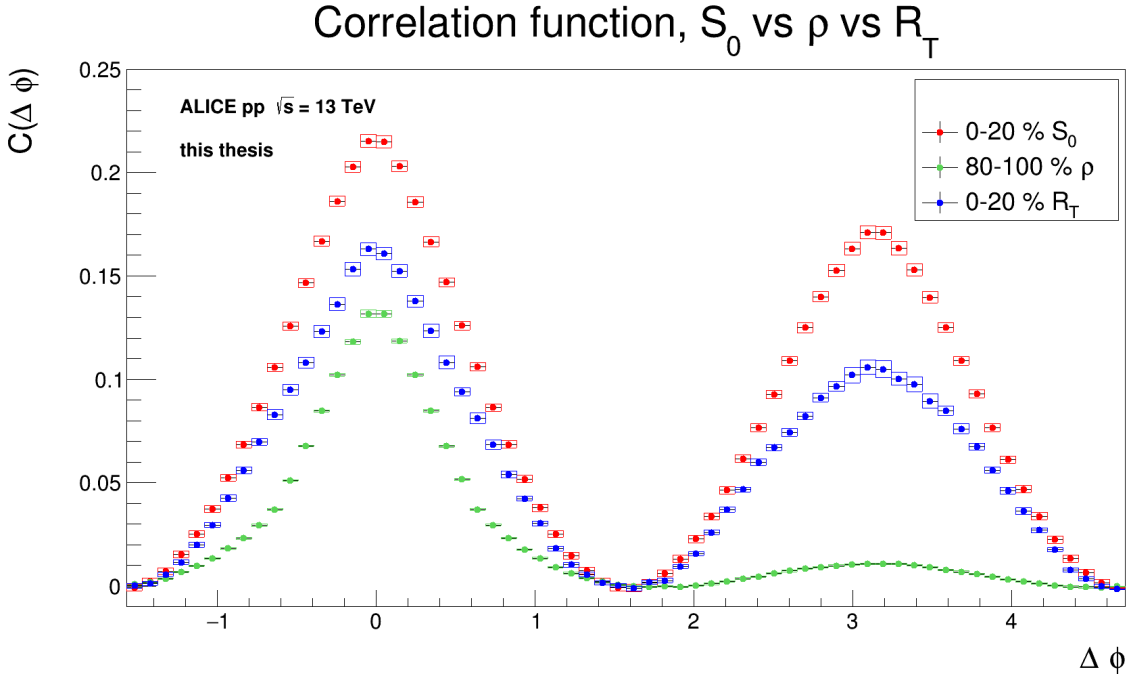


Figure 3.18: p_T-inclusive correlation functions for most jet-like R_T, S_0 and ρ bins. With ZYAM subtraction.

To conclude, when using the classifiers to select more isotropic events, flattenicity selects flatter events whereas spherocity selects events with a specific angle (average 90 degrees), and R_T selects events with a slight near-side peak. When using them to select jet-like events, spherocity selects highly jet-like events, R_T selects slightly less jet-like events than spherocity, and flattenicity selects events with a clear near-side peak but very low away-side peak.

Figure 3.19 shows a plot of flattenicity versus spherocity. It can be seen that they are clearly correlated, with a high spherocity indicating a low flattenicity and vice versa, as expected.

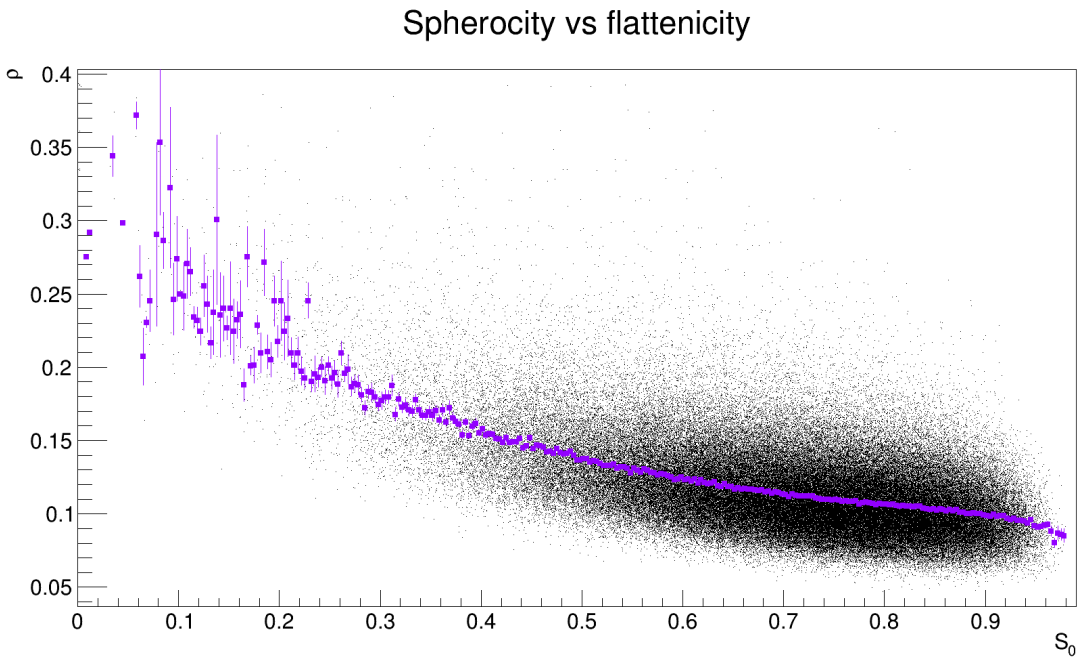


Figure 3.19: Flattenicity vs spherocity. The purple line displays the average correlation

3.5 Conclusion and outlook

Two-particle angular correlation functions were made for three event shape classifiers: R_T, spherocity and flattenicity. This contributes to the ongoing research in the heavy-ion physics community where one wishes to select isotropic events to study the reason behind the QGP-like behavior observed in high-multiplicity pp collisions. A clearer understanding of the differences between these three different methods of selecting events will be essential in order to fully analyze and interpret the results of using them.

R_T behaves mostly as expected, with the peaks at $\Delta\phi = 0$ and π becoming more and more pronounced when decreasing the R_T. When R_T attempts to select the most isotropic events, there is still a peak in the correlation function at $\Delta\phi = 0$. However, the away-side peak at $\Delta\phi = \pi$ disappears for this R_T bin. When doing correlations between tracks within the transverse region it was revealed that correlations remain, with a clear near-side peak at $\Delta\phi = 0$ for the higher R_T bins. This indicates that while R_T can select events with a more significant UE component, jet-like correlations remain.

Spherocity, similarly to R_T, makes the jet peaks more pronounced with lower spherocity. There are clear near-side and away-side peaks, decreasing in size as the spherocity increases. When reaching the lowest spherocity bins, peaks appear at $\Delta\phi = \frac{\pi}{2}$ and $\frac{3\pi}{2}$, due to spherocity selecting events where the tracks have a 90 degree angle between them.

Flattenicity, for all flattenicity bins, selects events with a very low away-side peak. The size of the away-side peak does not appear to be affected much by the flattenicity. Similarly to R_T and spherocity, flattenicity makes the near-side peak more pronounced with a higher flattenicity (analogous to lower R_T and spherocity). However, when selecting the lowest flattenicities, it is revealed that flattenicity biases against events where two or more tracks belong to the same cell.

When comparing spherocity and flattenicity it is concluded that flattenicity selects much flatter events than spherocity, both when looking at the least and most flat bins. Flattenicity appears to be able to select the most isotropic events out of the three investigated classifiers.

Future work in the context of this project includes using these results to further interpret previous measurements done using the three classifiers, as well as using them in future research when making measurements with respect to any of these classifiers. Furthermore, similar measurements to the ones done in this thesis could be done using models such as PYTHIA, and these model comparisons could further illuminate the physics of the selected events.

Acknowledgements

I would like to thank my supervisor, Alice Ohlson, for always encouraging me and being very supportive. I have learned so much during my projects with her, and I am very thankful to have had her as a supervisor for both my bachelor's and master's projects.

Thanks to Oliver Matonoha as well for helping me in the beginning of my project to get started with running over the trees on beast, and also helping me with the Monte Carlo data.

Finally, I would like to thank the ALICE group at Lund University for their support and for giving me their input on my work.

Bibliography

[1] Wikipedia. Standard Model — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Standard%20Model&oldid=1222953765, 2024. [Online; accessed 14-May-2024].

[2] The ATLAS collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29, September 2012. ISSN 0370-2693. doi: 10.1016/j.physletb.2012.08.020. URL http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[3] Roman Pasechnik and Michal Šumbera. Phenomenological review on quark–gluon plasma: Concepts vs. observations. *Universe*, 3(1):7, January 2017. ISSN 2218-1997. doi: 10.3390/universe3010007. URL http://dx.doi.org/10.3390/universe3010007.

[4] The ALICE collaboration. Production of pions, kaons, and protons as a function of the relative transverse activity classifier in pp collisions at $\sqrt{s} = 13$ tev. *Journal of High Energy Physics*, 2023(6), June 2023. ISSN 1029-8479. doi: 10.1007/jhep06(2023)027. URL http://dx.doi.org/10.1007/JHEP06(2023)027.

[5] The PHOBOS collaboration. System size dependence of two-particle angular correlations in p+p, cu+cu and au+au collisions. *Journal of Physics G: Nuclear and Particle Physics*, 35(10):104142, September 2008. ISSN 1361-6471. doi: 10.1088/0954-3899/35/10/104142. URL http://dx.doi.org/10.1088/0954-3899/35/10/104142.

[6] The CMS collaboration. Observation of long-range, near-side angular correlations in proton-proton collisions at the lhc. *Journal of High Energy Physics*, 2010(9), September 2010. ISSN 1029-8479. doi: 10.1007/jhep09(2010)091. URL http://dx.doi.org/10.1007/JHEP09(2010)091.

[7] Christian Bierlich and Smita Chakraborty et. al. A comprehensive guide to the physics and usage of pythia 8.3, 2022.

[8] ALICE Collaboration. Light-flavor particle production in high-multiplicity pp collisions at $\sqrt{s} = \mathbf{13}$ tev as a function of transverse spherocity, 2023.

[9] The ALICE collaboration. Enhanced production of multi-strange hadrons in high-multiplicity proton–proton collisions. *Nature Physics*, 13(6):535–539, April 2017. ISSN 1745-2481. doi: 10.1038/nphys4111. URL http://dx.doi.org/10.1038/nphys4111.

[10] Torbjörn Sjöstrand and Maria van Zijl. A Multiple Interaction Model for the Event Structure in Hadron Collisions. *Phys. Rev. D*, 36:2019, 1987. doi: 10.1103/PhysRevD.36.2019.

[11] Antonio Oritz and Arvind Khuntia et. al. Unveiling the effects of multiple soft partonic interactions in pp collisions at $\sqrt{s} = 13.6$ tev using charged-particle flattenicity. *Physical Review D*, 107(7), April 2023. ISSN 2470-0029. doi: 10.1103/physrevd.107.076012. URL `http://dx.doi.org/10.1103/PhysRevD.107.076012`.

[12] Christian Bierlich, Gösta Gustafson, and Leif Lönnblad. A shoving model for collectivity in hadronic collisions. 12 2016.

[13] Christian Bierlich and Gösta Gustafson et. al. Effects of overlapping strings in pp collisions, 2015.

[14] Creation of quark–gluon plasma droplets with three distinct geometries. *Nature Physics*, 15(3):214–220, December 2018. ISSN 1745-2481. doi: 10.1038/s41567-018-0360-0. URL `http://dx.doi.org/10.1038/s41567-018-0360-0`.

[15] The ALICE collaboration. Investigations of anisotropic flow using multi-particle azimuthal correlations in pp, p-pb, xe-xe, and pb-pb collisions at the lhc. *Physical Review Letters*, 123(14), October 2019. ISSN 1079-7114. doi: 10.1103/physrevlett.123.142301. URL `http://dx.doi.org/10.1103/PhysRevLett.123.142301`.

[16] The ALICE collaboration. *Computer Physics Communications*, 242:25–48, September 2019. ISSN 0010-4655. doi: 10.1016/j.cpc.2019.04.011. URL `http://dx.doi.org/10.1016/j.cpc.2019.04.011`.

[17] The ALICE collaboration. The ALICE experiment at the CERN LHC. *JINST*, 3: S08002, 2008. doi: 10.1088/1748-0221/3/08/S08002.

[18] Costanza Cavicchioli. *Development and Commissioning of the Pixel Trigger System for the ALICE Experiment at the CERN Large Hadron Collider*. PhD thesis, Florence U., 2011.

[19] The ALICE collaboration. Charged-particle production as a function of multiplicity and transverse spherocity in pp collisions at $\sqrt{s} = 5.02$ and 13.6 tev. *The European Physical Journal C*, 79(10), October 2019. ISSN 1434-6052. doi: 10.1140/epjc/s10052-019-7350-y. URL `http://dx.doi.org/10.1140/epjc/s10052-019-7350-y`.

[20] T. Martin, P. Skands, and S. Farrington. Probing collective effects in hadronisation with the extremes of the underlying event. *The European Physical Journal C*, 76(5), May 2016. ISSN 1434-6052. doi: 10.1140/epjc/s10052-016-4135-4. URL `http://dx.doi.org/10.1140/epjc/s10052-016-4135-4`.

[21] The ALICE collaboration. Multi-parton interactions in pp collisions using charged-particle flattenicity with alice, 2023.

[22] Antonio Ortiz and Guy Paic. A look into the "hedgehog" events in pp collisions using a new event shape – flattenicity. *Suplemento de la Revista Mexicana de Física*, 3, 12 2022. doi: 10.31349/SuplRevMexFis.3.040911.

[23] I. Antcheva and M. et. al. Ballintijn. Root — a c++ framework for petabyte data storage, statistical analysis and visualization. *Computer Physics Communications*, 180 (12):2499–2512, December 2009. ISSN 0010-4655. doi: 10.1016/j.cpc.2009.08.005. URL `http://dx.doi.org/10.1016/j.cpc.2009.08.005`.

Appendix A

p_T-weighted spherocity

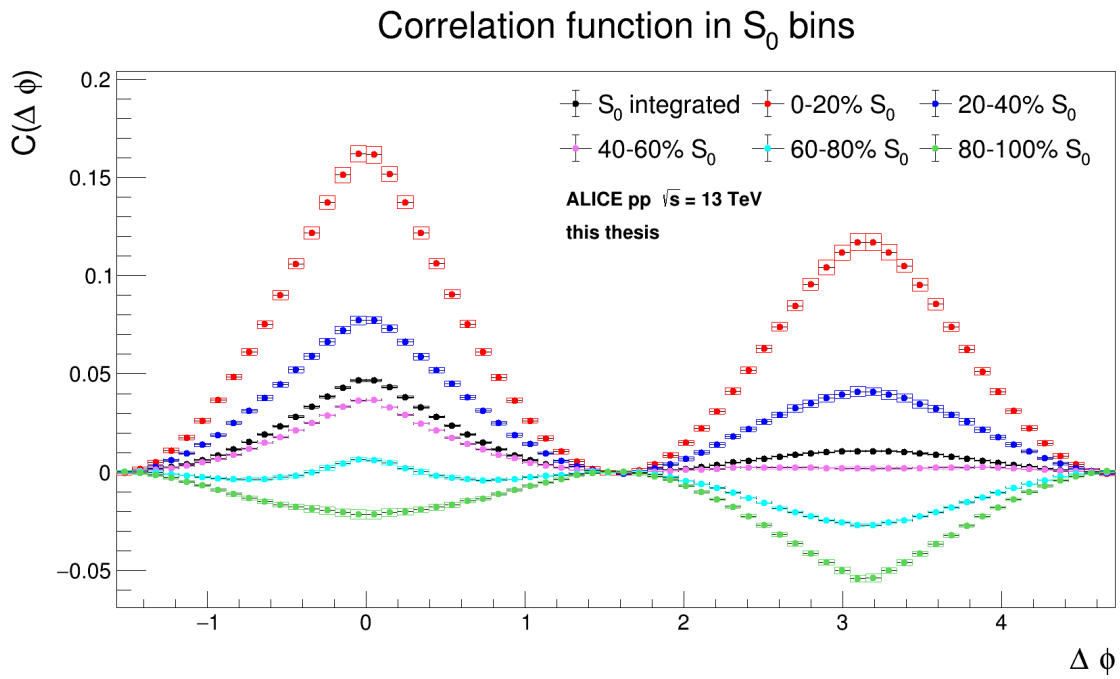


Figure A.1: $\Delta\phi$ projection of p_T-inclusive correlation function in S_0 bins, p_T-weighted spherocity. Top 10 % V0M multiplicity.

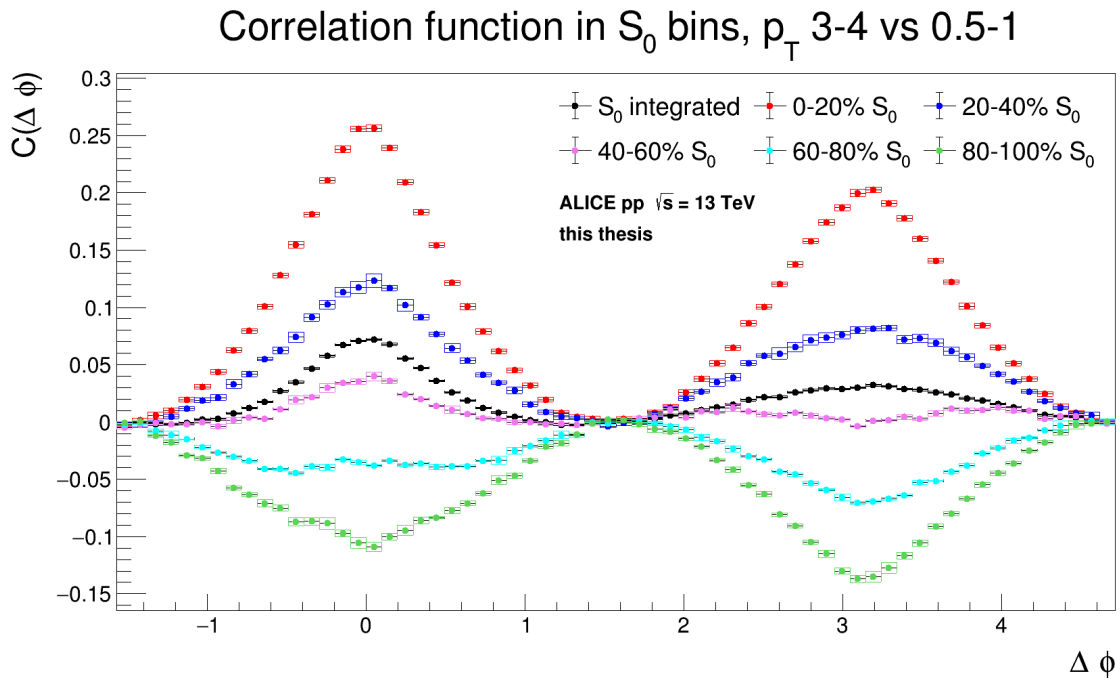


Figure A.2: $\Delta\phi$ projection of p_T 3-4 GeV/c vs 0.5-1 GeV/c correlation function in S_0 bins, p_T-weighted spherocity. Top 10 % V0M multiplicity.

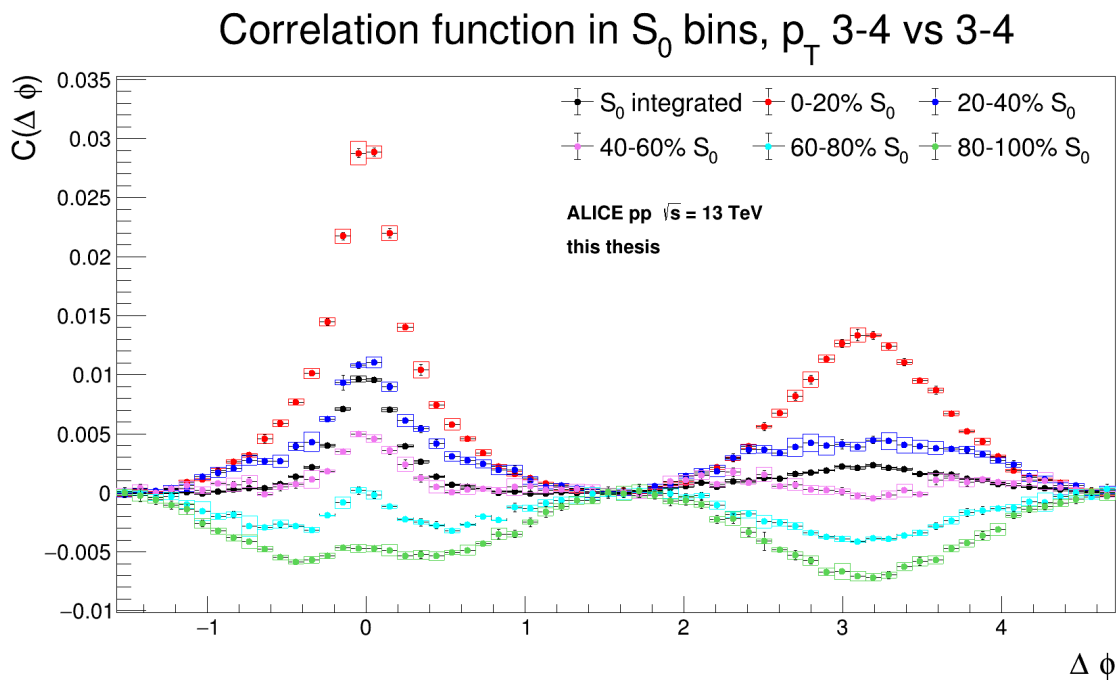


Figure A.3: $\Delta\phi$ projection of p_T 3-4 GeV/c vs 3-4 GeV/c correlation function in S_0 bins, p_T-weighted spherocity. Top 10 % V0M multiplicity.

Appendix B

p_T-weighted flattenicity

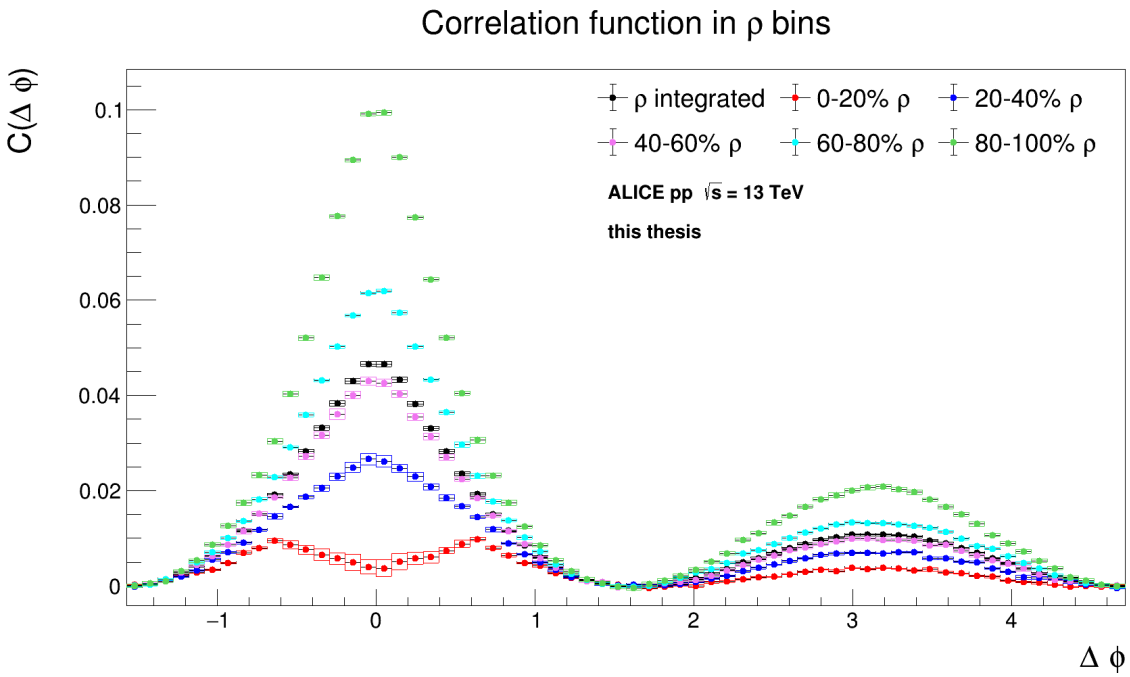


Figure B.1: $\Delta\phi$ projection of p_T-inclusive correlation function in ρ bins, p_T-weighted flattenicity. Top 10 % V0M multiplicity.

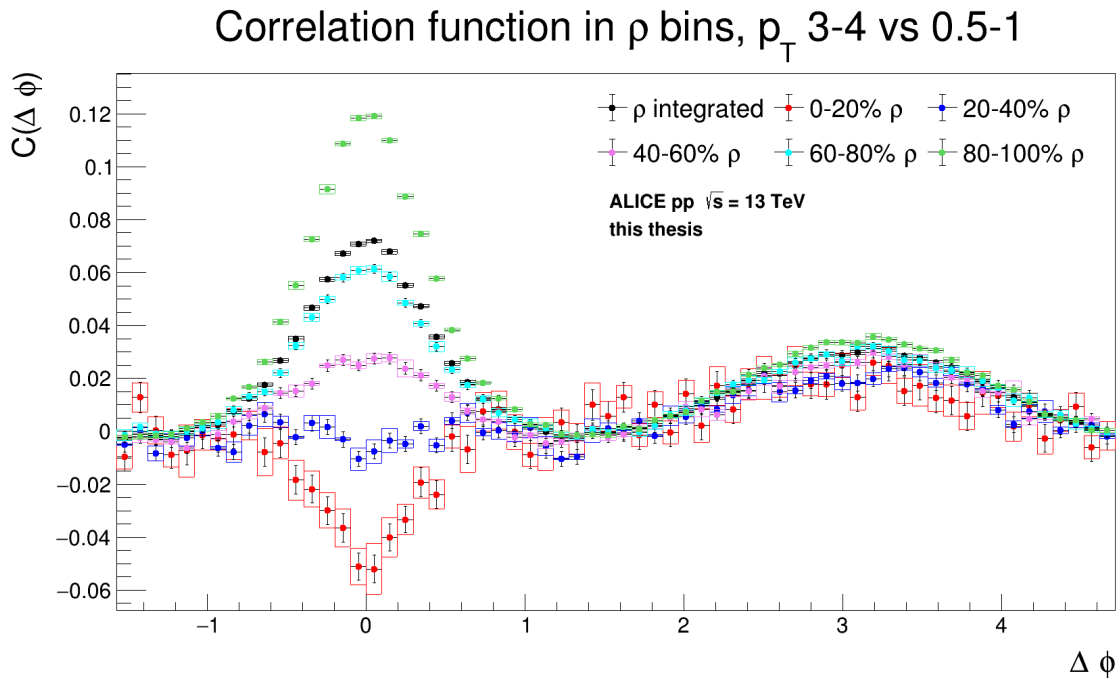


Figure B.2: $\Delta\phi$ projection of p_T 3-4 GeV/c vs 0.5-1 GeV/c correlation function in ρ bins, p_T-weighted flattenicity. Top 10 % V0M multiplicity.

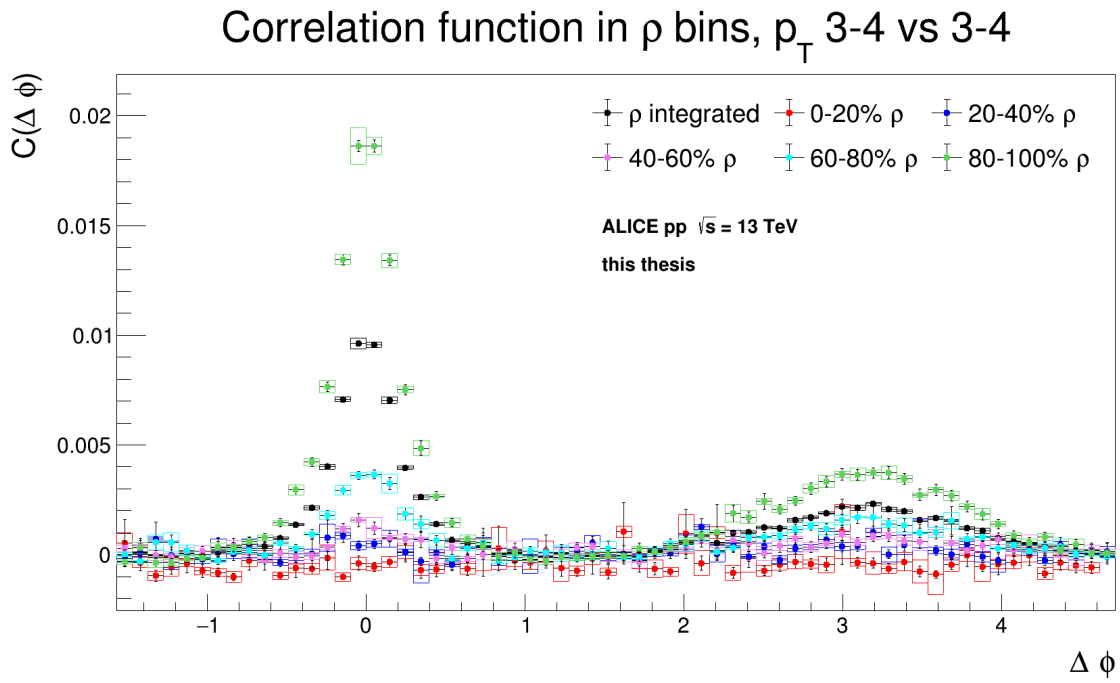


Figure B.3: $\Delta\phi$ projection of p_T 3-4 GeV/c vs 3-4 GeV/c correlation function in ρ bins, p_T-weighted flattenicity. Top 10 % V0M multiplicity.