# Forward-Backward Multiplicity Correlations for Proton-Proton Collisions with ALICE at LHC-CERN

**August Koppers**

Thesis Presented for Degree of Bachelor in Science
Project Duration: 2 Months

Supervisor: Peter Christiansen
Co-Supervisor: Alice Ohlson


LUND
UNIVERSITY

Department of Physics
Division of Particle Physics
Lund University
Sweden
May 2022

**Abstract**

This thesis presents the study of two-particle pseudorapidity correlations in pp-collisions at $\sqrt{s} = 13$ TeV using the ALICE detector at LHC, CERN. The correlation function was calculated as a function of multiplicity and was repeated for several multiplicity classes and different charge combinations in the forward-backward and central regions of the detector, respectively. The results reveal an enhancement of particle pairs being emitted in the same direction and strong autocorrelations using the mid-multiplicity estimator at low multiplicities. Comparing the results with PYTHIA simulations, similar results are obtained for both multiplicity estimators. This reveals that PYTHIA simulations are able to accurately reproduce these results, notably, also the autocorrelations are reproduced.

**Acknowledgments**

First, I would like to thank my supervisor Peter Christiansen for introducing this project idea to me, helping me during this project, and guiding me through the process of writing this thesis, making it better than I could ever have hoped. I would also like to thank my co-supervisor Alice Ohlson and the ALICE team at Lund for ideas and feedback, and Omar Vazquez Rueda who helped get started with ROOT.

# Contents

# Acronyms

**ALICE** A Large Ion Collider Experiment.

**ATLAS** A Toroidal LHC Apparatus.

**CERN** Conseil Européen pour la Recherche Nucléaire.

**CM** Center of Mass.

**CMS** Compact Muon Solenoid.

**DCA** Distance of Closest Approach from trajectory to the primary vertex.

**ITS** Inner Tracking System.

**LHC** Large Hadron Collider.

**LHCb** LHC-beauty.

**MC** Monte Carlo.

**ME** Mixed Event.

**QCD** Quantum Chromodynamics.

**QED** Quantum Electrodynamics.

**QGP** Quark-Gluon Plasma.

**SDD** Silicon Drift Detector.

**SE** Same Event.

**SM** Standard Model.

**SPD** Silicon Pixel Detector.

**SSD** Silicon Strip Detector.

**TPC** Time-Projection Chamber.

# 1 Introduction

During the very first moments of the universe, after the Big Bang, it is believed that an entirely different phase of matter was present. This phase was so dense that quarks, the building blocks of all matter, were no longer confined to the fundamental particles we observe every day. In this state of matter, quarks are deconfined (more or less free to move on their own) and is known as the Quark-Gluon Plasma (QGP). It is crucial to analyze this phase to understand how the universe has evolved and came to be what we see today.

To study the QGP one might expect measurements on astronomic scales to be necessary, but instead we can study this on Earth by analyzing some of the smallest particles we know. By accelerating protons or heavy ions, such as lead, to speeds close to the speed of light, and letting them collide, the QGP has been replicated on Earth. One still unanswered question about this phase is if a QGP is formed during smaller scale collisions, such as proton-proton (pp) collisions.

In this thesis, I will analyze two-particle pseudorapidity correlations, with pseudorapidity defined in section 3.1. This correlation function is a function of multiplicity, the number of recorded particles, and analyzing this will give insight into the dynamics of small-scale collisions. This is then compared with PYTHIA (see section 5.5.1) simulations to receive a better understanding of the limitations and advantages of the "Lund String model" used in PYTHIA.

## 1.1 Aim of Thesis

This thesis aims to reproduce parts of the analysis done by ATLAS [1], but with the ALICE detector described in section 4. Namely, two-particle correlation in pseudorapidity will be studied for different charge combinations. Additionally, the analysis will be done for the forward-backward (V0M) and central ($N_{\text{tracklets}}$) multiplicity estimators and will be compared with simulations done in PYTHIA. As the ATLAS analysis was only done for the central multiplicity estimator and most of ALICE's multiplicity dependent results are obtained for the forward estimator this will be of particular interest to understand possible biases introduced by using different multiplicity estimators.

# 2 The Standard Model

The Standard Model (SM) was developed throughout the 20th century and has since been the fundamental theory of particle physics [2]. Here a brief introduction to the SM is given, presenting the elementary particles and forces with a focus on the strong force.

## 2.1 Elementary Particles of the SM

In the SM there are three classes of particles: quarks, leptons, and bosons, where the fermionic quarks and leptons constitute all observable matter, and the bosons are the force carriers. There are six quarks: up (u), down (d), charmed (c), strange (s), top (t), and bottom (b). There are also six leptons: electron ($e^-$), muon ($\mu^-$), tau ($\tau^-$); with respective corresponding neutrino: electron-neutrino ($\nu_e$), muon-neutrino ($\nu_\mu$), and tau-neutrino ($\nu_\tau$). Each quark and lepton also have a corresponding antiparticle with an opposite charge and spin.

The quarks and leptons, described above, can further be classified into three generations. Each generation consists of particles with identical properties except for the mass, which increases with each generation [3]. This classification of generations, together with all particles and bosons, excluding corresponding antiparticles, can be seen in Fig. 1.

Finally, the five known bosons are the photon ($\gamma$), the Z-boson ($Z^0$), the W-boson ($W^{+/-}$), the gluon ($g$), and the Higgs-boson ($H$). Where the photon and gluon are the force carriers for the electromagnetic and strong force respectively, and the Z- and W-bosons are the force carriers for the weak force. The Higgs boson is responsible for giving elementary particles their mass through the Higgs mechanism. [4]

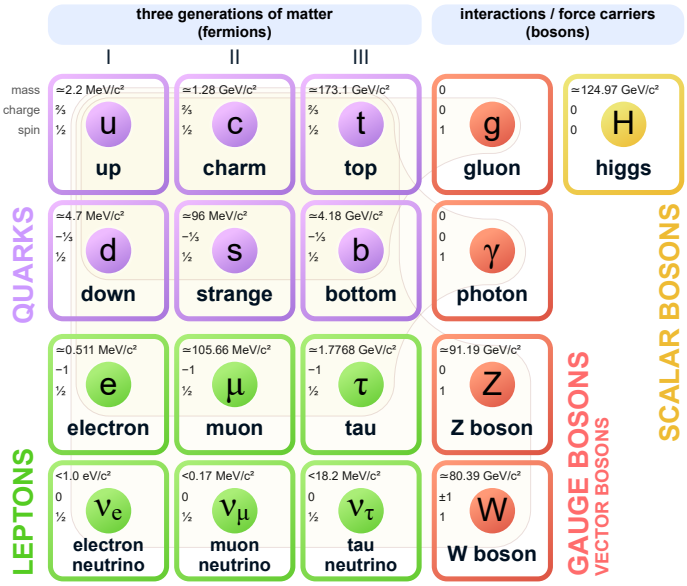



Figure 1: Shows the elementary particles of the standard model. The fermions, quarks, and leptons are divided into generations I, II, and III and shown as colons from left to right. The figure is taken from [5].

## 2.2 Fundamental Forces in the SM

The four known forces of nature are electromagnetism, the strong force, the weak force, and gravity. As a result of the elementary particles having so little mass, gravity hardly affects them. Therefore, gravity is effectively neglected in the SM, leaving us with three forces. The electromagnetic interactions are described by Quantum Electrodynamics (QED), acting on electrically charged particles. The weak force acts on all fermions and has been unified with the electromagnetic force into the electroweak force [3].

### 2.2.1 Strong Interactions

The strong force acts on all quarks and is mediated through the gluon. The strong interaction acts on "color-charged" particles. There are three color charges: red, green, and blue together with respective "anti-color". All quarks and gluons carry color charge, and their interaction is predicted by Quantum Chromodynamics (QCD).

Quarks are never observed alone, but in combinations with other quarks and anti-quarks such that the quarks together are "colorless". That is, all matter built out of quarks,

known as hadrons, consists of a combination of colors that mix to white or no color. For instance, a hadron consisting of the colors: red, blue, and green (or corresponding anti-colors) mix to white, forming a hadron known as a baryon (or anti-baryon). This quark triplet can be denoted as $(qqq)$ or $(\bar{q}\bar{q}\bar{q})$ for baryons and anti-baryons respectively. Having a quark-antiquark pair $(q\bar{q})$, the quarks must be of the opposite color, for example blue + anti-blue, for the hadron to be colorless, forming a meson [4].

### 2.2.2 Asymptotic Freedom and Jets

An important property of the strong force is its asymptotic freedom resulting in jet formations. Contrary to the electromagnetic force, which strength decreases with distance, the strong force grows stronger the further the separation between quarks is. Consequently, quarks are deconfined (move relatively freely) within hadrons, known as asymptotic freedom, but cannot be separated very much, making quarks confined inside hadrons.

This also means that when colliding two hadrons a high-energy gluon may be exchanged, knocking out the constituent partons (quarks and gluons). As these "free" partons have color but all individual particles must be colorless, sprays of particles known as jets are created [3]. This process can be modeled using the Lund String model described in section 3.3.

# 3 Introduction to High-Energy Physics

The smaller the system that one wants to study, the higher the energies are needed, making it natural to work with high energies in particle physics. In this section, common variables used in high-energy physics are first presented, followed by a brief presentation of the quark-gluon plasma and an introduction to the Lund string model.

## 3.1 Variables of High-Energy Physics

During relativistic high-energy collisions, it can be useful to define new variables to simplify the description of these events. Here rapidity, pseudorapidity, and center of mass-energy will be described.

Rapidity, y, is defined as

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \tag{1}$$

where $E$ is the energy and $p_z$ is the longitudinal (along the beam-axis) momentum of the particle. It is often convenient to use rapidity rather than momentum as differences in rapidity are Lorentz-invariant under boosts along the beam-axis.

If the particle emitted makes an angle $\theta$ with the beam axis, then $p_z = p\cos\theta$, where $p$ is the magnitude of the total momentum. Using that $E = \sqrt{m^2 + p^2}$ and assuming that the transverse momentum $p_T \gg m$, we see that Eq. (1) becomes:

$$y = \frac{1}{2} \ln \frac{\sqrt{m^2 + p^2} + p\cos\theta}{\sqrt{m^2 + p^2} - p\cos\theta} \approx \frac{1}{2} \ln \frac{p + p\cos\theta}{p - p\cos\theta} = -\ln\tan\theta/2 \tag{2}$$

where in the last step the half-angle formula for the tangent was used.

The last expression in Eq. (2) is the definition of pseudorapidity,

$$\eta \equiv -\ln\tan\theta/2 \tag{3}$$

That is, for large transverse momentum the rapidity converges to the pseudorapidity. Additionally, pseudorapidity is often easier to work with than rapidity, as it is independent of energy and momentum making it natural to use in particle detectors where we do not always know the energy of particles [6].

Interpreting pseudorapidity, $\eta = 0$ corresponds to a perpendicular angle with the beam axis, and if $\eta \to \pm\infty$ corresponds to an angle along the positive and negative beam axis respectively. A schematic of the correspondence between $\eta$ and $\theta$ can be seen in Fig. 2.

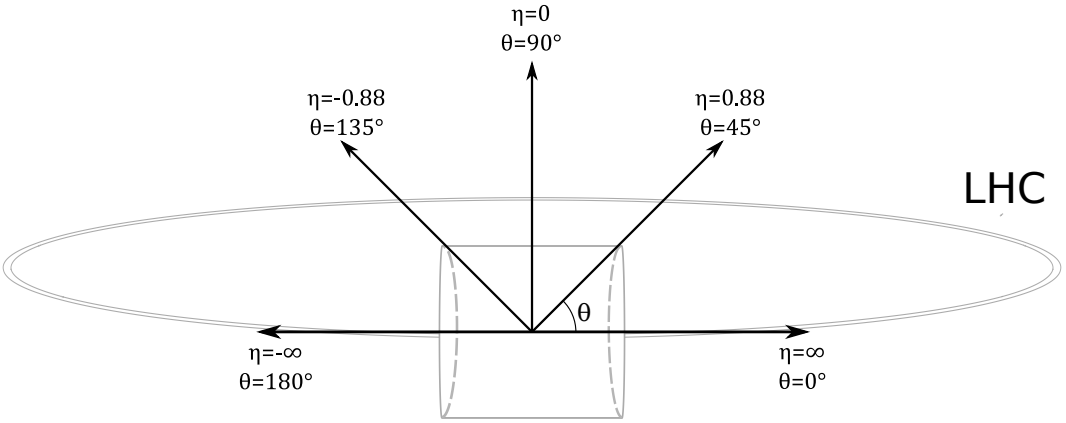



Figure 2: Illustration of the relation between $\eta$-values and their corresponding angle made with the beam-axis, $\theta$. The figure is inspired by [7].

Considering two particles, 1 and 2, one defines the Lorentz-invariant variable $s \equiv (p_1 + p_2)^2$, where $p$ denotes the four-momentum. $s$ is known as one of the Mandelstam variables, in the Center of Mass (CM) frame $s = (E_1 + E_2)^2 - (\bar{p_1} + \bar{p_2})^2 = (E_{\text{CM}})^2$, where $\bar{p}$ denotes the three-momentum and $E_{\text{CM}}$ is the CM-energy. Hence, $\sqrt{s}$ is the CM-energy of the system. This is the standard description of energy in particle colliders, considering the two beams having four-momentum $p_1$ and $p_2$. Typical CM-energies reached in particle accelerators are in orders of several TeV in magnitude [8]. In this thesis pp-collisions at $\sqrt{s} = 13$ TeV will be studied.

## 3.2 The Quark-Gluon Plasma

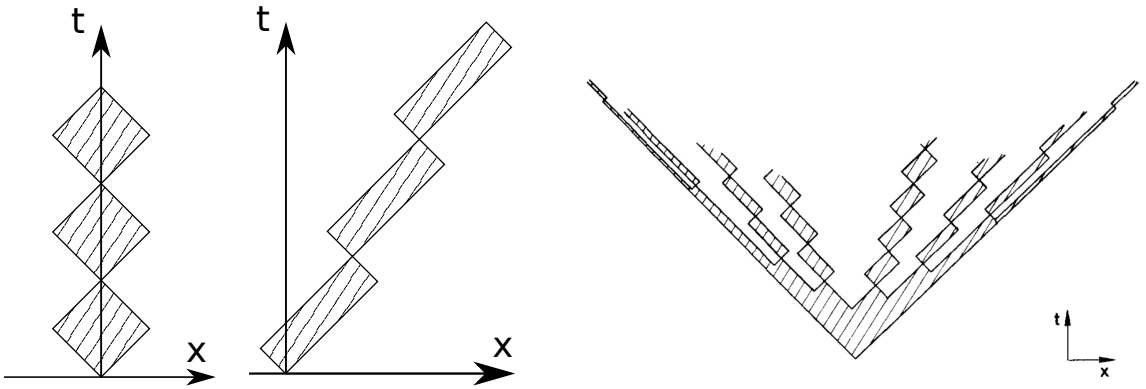
During high energy collisions, the energy-density of hadronic matter becomes large enough resulting in the strong force being less dominant, as the strong force is weaker at short distances. We therefore approach asymptotic freedom at high densities, resulting in quarks moving quasi-freely over extended volumes, known as deconfinement. At the same time, such high energies are reached that new $q\bar{q}$ pairs can be produced, further blurring the boundaries between different hadrons, resulting in a "quark soup" known as Quark-Gluon Plasma (QGP) [9].

There are several observables for QGP, such as strangeness enhancement and jet quenching, which are discussed in more detail in [8]. This has mainly been observed in collisions involving heavy ions, such as lead, and some experiments insinuate that this is also formed in smaller scale collisions, such as proton-proton collisions [10].

### 3.3 Lund String Model

The Lund string model is a statistical model that describes the formation of hadrons after a collision event, also known as hadronization. The model is based on the string-like properties of quark confinement. At large distances, the QCD interaction can be modeled by a linear string-like potential $V(r) \propto \kappa r$ where $\kappa$ is the string constant. As color charges in the initial interaction move apart, the distance between color charges will be large enough for it to be energetically favorable for the string to break, forming a quark anti-quark pair ($q\bar{q}$). Considering a hadron to be such a string, when the string breaks or fragments, new hadrons are produced [8].

As an example, consider a stable $q\bar{q}$-pair (meson). One can visualize the movement or excitation of the quarks inside the mesons with a space-time diagram, see Fig. 3a. Performing a Lorentz boost on this results in Fig. 3b. The fragmentation of the string is modeled as a tunneling probability, meaning it exponentially increases with the string length. This fragmentation process of multiple produced Lorentz boosted mesons can be seen in Fig. 3c.




(a) (b) (c)

Figure 3: Space-time diagram of two bound quarks. (a) shows a bound meson exhibiting a "yo-yo"-like movement in one dimension. Performing a Lorentz boost on (a) results in (b). Multiple fragmentations and meson-formations can be seen in (c). The figures are from [11].

Notably, as discussed in [12] and [13], iterating the fragmentation model one receives a flat plateau in rapidity, meaning that the hadron density produced by a single string is almost constant in rapidity, as seen in Fig. 4.

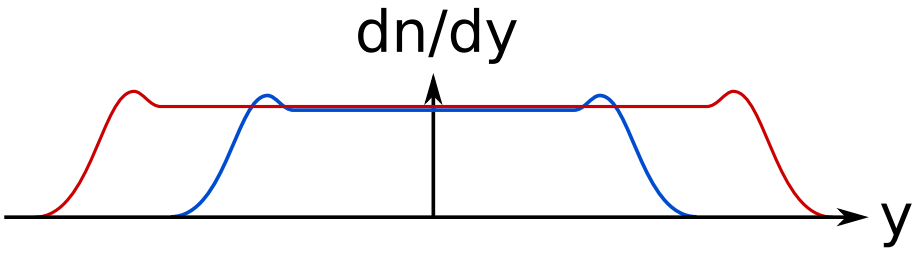



Figure 4: The produced plateau in hadron density is made by iterating the fragmentation model. The plateau is shown with some endpoint corrections. Here, $n$ denotes the number of particles, and $y$ the rapidity. The blue and red curves show this for lower and higher energies respectively. The figure is from [14].

# 4 The ALICE Experiment at the LHC

The ALICE detector is one of four main detectors at the Large Hadron Collider (LHC) at Conseil Européen pour la Recherche Nucléaire (CERN), Geneva [15]. The LHC is a circular particle accelerator, with a 27 km circumference, consisting of two beam-lines. The beams of accelerated particles are bent and focused by superconducting dipole and quadruple magnets respectively. The maximum CM-energy that was available at the LHC was $\sqrt{s} = 13$ TeV [7].

The four main experiments at LHC are ATLAS (A Toroidal LHC Apparatus), CMS (Compact Muon Solenoid), LHCb (LHC-beauty), and ALICE (A Large Ion Collider Experiment). ATLAS and CMS are general-detection experiments, while LHCb and ALICE are more specialized. LHCb explores hadrons containing the bottom, also known as beauty, quark, while ALICE is optimized for the study of heavy-ion collisions and the QGP [15].

The remainder of this section will describe some of the most relevant parts of the ALICE detector used in the data collected for this thesis. A more detailed description can be found in [16].

## 4.1 ALICE

The ALICE experiment specializes in heavy-ion collisions and, hence, is also the detector capable of recording the highest multiplicities, number of recorded particles, at CERN. A schematic picture of the ALICE detector and all its subdetectors can be seen in Fig. 5.

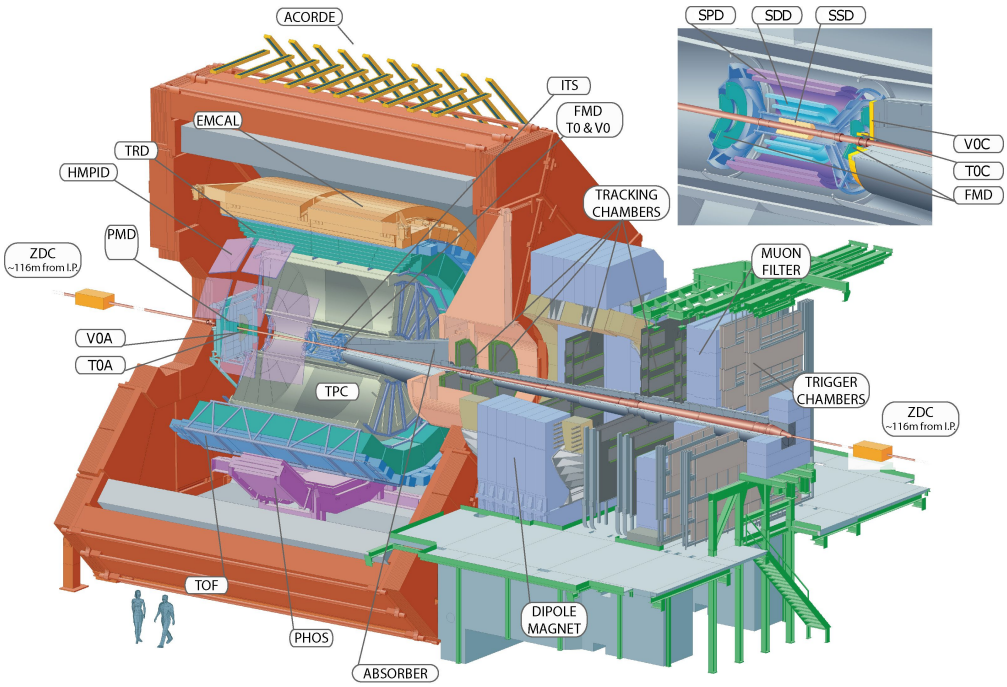



Figure 5: The ALICE detector and many of its subdetectors. Notably, the detectors discussed in this section can be seen. The ITS is at the center of the detector and an enlarged image of this can be seen in the top right corner showing the SPD, SDD and SSD more clearly. The TPC can be seen as the second innermost detector and the V0A and V0C can be seen on either side of the interaction point. The modified figure is taken from [17].

## 4.2 Inner Tracking System (ITS)

The Inner Tracking System (ITS) is the innermost part of the ALICE detector. The central barrel consists of six layers of silicon-based semiconductor detectors, located at radii between 4 and 43 cm and together covering a pseudorapidity-region of $|\eta| < 0.8$.

These silicon-based detectors operate by having a semiconductor run in reverse bias, creating a depletion region. Charged particles passing through will excite electrons, creating a detectable current, proportional to the energy of the detected particle [18].

The first two innermost layers consist of arrays of hybrid Silicon Pixel Detector (SPD), covering a pseudorapidity-region of $|\eta| < 2.0$ and $|\eta| < 1.4$, respectively, and are in this analysis mainly used to determine the primary vertex position of tracks. Of the remaining four semiconductor layers in the ITS, the first two are known as the Silicon Drift Detector (SDD), while the last two, where the track density is expected to be the lowest have double-sided Silicon micro-Strip Detectors, are the SSD. The SDD and SSD have similar functions and mainly detect the energy loss per distance, $dE/dx$, of particles along with finding the track segments closest to the primary vertex together with the SPD.

## 4.3 Time Projection Chamber (TPC)

The Time-Projection Chamber (TPC) is the second innermost detector system and is the main tracking detector in the central region of ALICE. It is essentially a long drift chamber. In particular, the ALICE TPC has both electric and magnetic fields, with the latter being produced by the L3 magnet encapsulating the whole central detector. The TPC in ALICE consists of a large container, $|\eta| < 0.8$, containing a gas mixture of Ne, $CO_2$ and $N_2$.

When a charged particle enters the gas, it will be ionized. Respective parts of the ion pair are accelerated by the electric field, resulting in the released electrons drifting towards the TPC endplates where they and their position are measured using wire chambers. Using the recorded orthogonal projection of the position and the time taken to reach the endplates of the produced ions, tracks can accurately be reconstructed in three dimensions. This allows the transverse momentum, $p_T$, to be measured from the radius of curvature the particle makes in the magnetic field. Together with the recorded 3D direction of the particle, the total momentum can be determined. The TPC is also used to detect the energy loss per distance and, together with other central detectors, properties such as vertex position and particle identification. [19, 20, 16].

## 4.4 Multiplicity Estimators

To identify the activity of collision events, multiplicity estimators are used. The multiplicity can be estimated either by separate detectors or by using the same detectors that reconstruct the collision tracks. Two multiplicity estimators measuring the activity of different parts of the detector will be discussed here: the V0M and the $N_{\text{tracklets}}$ estimators.

The multiplicity recorded by these estimators is commonly shown in terms of multiplicity percentiles. For example, a multiplicity of 0-1% corresponds to the event having the highest 1% recorded multiplicity. This is illustrated in Fig. 6, which shows a typical multiplicity curve with different multiplicity percentiles.

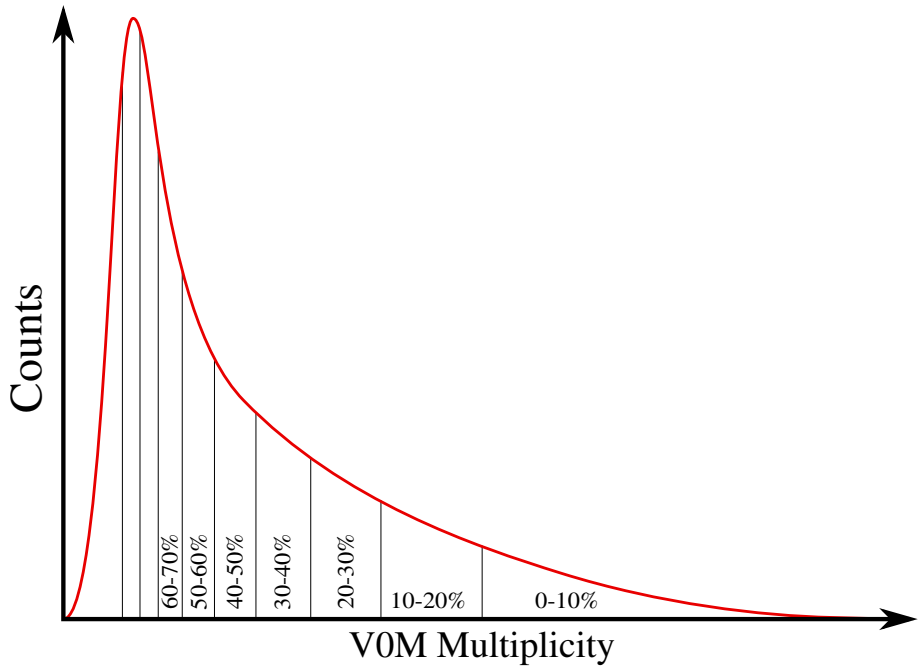



Figure 6: Shows a typical graph of the number of events and the corresponding multiplicity recorded by the V0M estimator for pp collisions. This is usually expressed in terms of multiplicity percentiles, shown as the different intervals under the graph. Each interval contains events with the top x% multiplicity in the corresponding interval. The figure is taken from [21].

### 4.4.1 V0M estimator

The V0A and V0C are two forward multiplicity estimators. These are scintillator counters found on each side of the ALICE interaction point, covering a pseudorapidity $2.8 < \eta < 5.1$ for the V0A, and $-3.7 < \eta < -1.7$ for the V0C estimator.

The scintillator counter functions on the principle that charged particles moving through the medium inside the scintillator will excite this medium. Then, when the medium is de-exited photons will be emitted, which are then guided to a photomultiplier system, where the signal is amplified, which then is recorded and used to count the number of particles that have passed through the detector. [19]. The sum of the V0A and V0C is referred to as the V0M, which will be used in this work as the forward-multiplicity estimator.

### 4.4.2 $N_{\text{tracklets}}$ estimator

The $N_{\text{tracklets}}$ estimator is used to measure the multiplicity, or size, of a collision in the central region of the detector. Using the two SPD detectors of the ITS, one can construct a tracklet, a short two-point track segment between two SPD-hits. $N_{\text{tracklets}}$ is then the count of the number of reconstructed tracklets using the SPD in the pseudorapidity region $|\eta| < 0.8$ and gives an estimate of the multiplicity in the central region of the detector.

# 5 Analysis Method

The analysis resembles that made by ATLAS [1] but with proton-proton (pp) collisions using the ALICE detector. Mainly two-particle correlation studies will be investigated, which is explained in this section. This analysis was repeated for different multiplicity classes, multiplicity estimators, and charge combinations. The result was also compared with PYTHIA simulations. The following analysis was executed using *ROOT* and *AliROOT*, two offline data processing software, standard for ALICE analysis [22].

## 5.1 Two Particle Correlations

Forward-backward correlations can be studied by the two-particle correlation function, as done in [1, 23], and is defined as:

$$C(\eta_1, \eta_2) = \frac{\langle N(\eta_1)N(\eta_2)\rangle}{\langle N(\eta_1)\rangle \langle N(\eta_2)\rangle} \tag{4}$$

where $N(\eta)$ is the multiplicity density distribution for a single event and $\langle\cdot\rangle$ denote the event average. This is a function of pseudorapidity and returns a distribution conveying the likelihood of different particle pairs being produced. The correlation function above is then normalized so that the average value of $C(\eta_1, \eta_2)$ is one. Hence, a value of one corresponds to two independent tracks, and larger or smaller values correspond to the probability of the track-pair being seen is larger or smaller, respectively.

We expect that particles that end up correlated and are close in momentum-space are the end products of the same early-time microscopic processes, resulting in this distribution being sensitive to the early-time dynamics of the system. In terms of the Lund string model, see section 3.3, we expect that correlated particles are produced from the same string, resulting in a similar hadron density-distribution as seen in Fig. 4.

Given collision data, the pseudorapidity of all measured hadrons in an event is known. Their average distribution can be displayed in a histogram, as in Fig. 7a. Combining all pairs of particles, one obtains the average pseudorapidity distribution for all measured particle pairs. Hence, the point $(\eta_1, \eta_2) = (0, 0)$, in such distribution, corresponds to both particles being emitted perpendicularly to the beam axis, and a point $(0, \infty)$ corresponds to one particle being emitted perpendicularly and the other along the beam axis. A schematic example of this type of distribution is shown in Fig. 7b and will be referred to as a Same Event (SE) distribution.

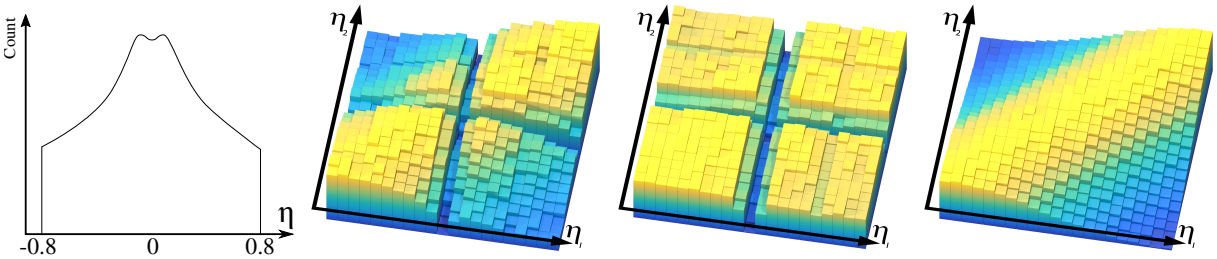
## 5.2 Event Mixing

In contrast to the SE distribution, described above, which only uses track combinations from the same event, a new distribution is created of mixed pairs from two different events and will be called the Mixed Event (ME) distribution. By combining tracks from two separate events, one is sure that any non-trivial correlation must be introduced by the detector. The ME distribution is hence used to correct detector inefficiencies, an example of a schematic ME distribution can be seen in Fig. 7c.

Denoting the SE distribution as $S(\eta_1, \eta_2)$ and the ME distribution as $B(\eta_1, \eta_2)$, we have $S(\eta_1, \eta_2) \propto \langle N(\eta_1)N(\eta_2)\rangle$ and $B(\eta_1, \eta_2) \propto \langle N(\eta_1)\rangle \langle N(\eta_2)\rangle$, as described in [1]. Since the SE and ME distributions are proportional to the numerator and denominator of Eq. (4) respectively, we get that:

$$C(\eta_1, \eta_2) = \frac{S(\eta_1, \eta_2)}{B(\eta_1, \eta_2)} \tag{5}$$

with the proportionality constant canceling since $C(\eta_1, \eta_2)$ is normalized. An example of such a distribution can be seen in Fig. 7d. In this distribution, particles with similar pseudorapidity, $\eta$, are referred to as short-range correlations and particles with different $\eta_1$ and $\eta_2$ are referred to as long-range correlations.




(a) Single event $\eta$-spectrum.

(b) SE-distribution.

(c) ME-distribution.

(d) Two particle correlation-distribution.

Figure 7: Several different schematic pseudorapidity distributions. The left-most spectrum shows how $\eta$ in a single event. (b), (c) and (d) show the SE, ME, and final correlation distributions respectively, using multiple events.

## 5.3 Event and Track Selection

To ensure that the pseudorapidity distributions are not dependent on other variables, we impose event and track selection conditions that limit possible biases.

The first event selection criteria is for the recorded event not to be a pileup. That is to ensure that each accepted event does not have recorded tracks that overlap with other events or background events [24]. Secondly, we impose the criteria that a primary vertex was successfully reconstructed. And lastly, we need the position of the reconstructed vertex along the beam-line, $z_{\text{vtx}}$, to be within 10 cm from the center of the detector, $|z_{\text{vtx}}| < 10$ cm, to ensure that the detector response is homogeneous.

For the track selection, ALICE ITSTPC2011 track cuts were used, discussed in [25]. These cuts are optimized to reduce contamination from secondary tracks and give the best momentum resolution. This is obtained by requiring that track segments in the ITS and TPC are matched and that at least one SPD hit is recorded.

A separate selection criteria to filter out particles not created during the collision, that may be products of material interactions or weak decays, was performed. The criteria relates $p_T$ and the transverse DCA component (the minimum distance between the reconstructed track and the primary vertex), denoted by $|\text{DCA}_{xy}|$. The criteria, as discussed in [25], requires tracks to fulfill:

$$|\text{DCA}_{xy}| < 0.0105 + \frac{0.0350}{[p_T(\text{GeV}/c)]^{1.1}} \tag{6}$$

## 5.4 Multiplicity Classes and Charge Combinations

For the extraction of the correlation function, Eq. (5), the data-set was divided into groups with different multiplicity classes, multiplicity estimators, and charge combinations, each of which the analysis was repeated for. Defining three multiplicity classes with the following intervals: 0-10%, 10-30%, and 30-100%. The $S(\eta_1, \eta_2)$ and $B(\eta_1, \eta_2)$ distributions were calculated for respective multiplicity intervals. In other words, all tracks within the corresponding distribution are within the multiplicity interval. Additionally, this was also repeated with the V0M and $N_{\text{tracklets}}$ multiplicity estimators respectively, defined in section 4.4, as well as for the charge combinations: same charge, ++/- -, and opposite charge, +-/-+.

## 5.5 Monte Carlo Simulations

Monte Carlo (MC) simulations are particularly useful for studying the properties of complex systems or systems with non-analytic solutions. The core concept of MC methods is to make use of randomness by iterating a model, repeatedly creating samples with different variables. The samples are then used to gain information about the entire system. However, it is not necessary for the system one wants to study to be inherently probabilistic for MC methods to be applicable. A common example of the use of MC is for the computation of multivariable integrals, which is done by sampling a region with points and identifying the fraction of points inside the integral region, which then is used to estimate the volume of this space [26].

### 5.5.1 PYTHIA

During high-energy particle collisions, many new particles are created in processes, briefly discussed in section 2.2.2, resulting in models describing all these particles becoming more convoluted. This makes it difficult and sometimes impossible to compare models and experiments. One way of resolving this problem is by introducing event generators [13]. One MC event generator commonly used in particle physics is PYTHIA, used to generate collision events that resemble real data. The MC methods are used to introduce some randomness, representing quantum mechanical variability [27]. The simulation model PYTHIA utilizes is the Lund String model [28], described in section 3.3. PYTHIA simulations will therefore resemble real collision data, making direct comparisons between the two useful.

In this thesis, PYTHIA[1] was used to simulate $10^6$ events of inelastic non-diffraction pp-collisions, with CM-energy $\sqrt{s} = 13$ TeV and considering particles with a lifetime, $\tau \geq \frac{1\text{cm}}{c}$, as primary particles where $c$ is the speed of light. The pseudorapidity and charge of all tracks within the same $\eta$ range as the ALICE detector ($|\eta| < 0.8$) was recorded to simulate ALICE collision events. Each event multiplicity in the forward-backward and central regions was also recorded, corresponding to the V0M and $N_{\text{tracklets}}$ estimators. The multiplicity was then converted to multiplicity percentiles, and the same analysis that was performed for the ALICE data, as described in sections 5.1 and 5.2, was repeated for the PYTHIA simulations.

[1] PYTHIA version 8.307

### 5.6 Data Fits

To be able to easily compare the main aspects of different correlation functions, fits to $C(\eta_1, \eta_2)$ was performed. Defining the two new parameters: $\eta_+ \equiv \eta_1 + \eta_2$ and $\eta_- \equiv \eta_1 - \eta_2$, we define the following fit function:

$$F(\eta_+, \eta_-) = f(\eta_+) \cdot g(\eta_-) + C \tag{7}$$

with $f(\eta_+)$ being an even polynomial along $\eta_+$:

$$f(\eta_+) = A + B\eta_+^2 \tag{8}$$

and $g(\eta_-)$ being a Gaussian along the $\eta_-$:

$$g(\eta_-) = \exp\left[-\frac{1}{2}\left(\frac{\eta_-}{\sigma}\right)^2\right] \tag{9}$$

and A, B, C, and $\sigma$ are constants determined during the fitting process.

Hence, Eq. (7) corresponds to a function that is a polynomial along the bottom left to top right diagonal and a Gaussian function, with zero mean, along the other diagonal in the $C(\eta_1, \eta_2)$ histogram. Small $\eta_-$ corresponds to short-range correlations, and large values correspond to long-range correlations.

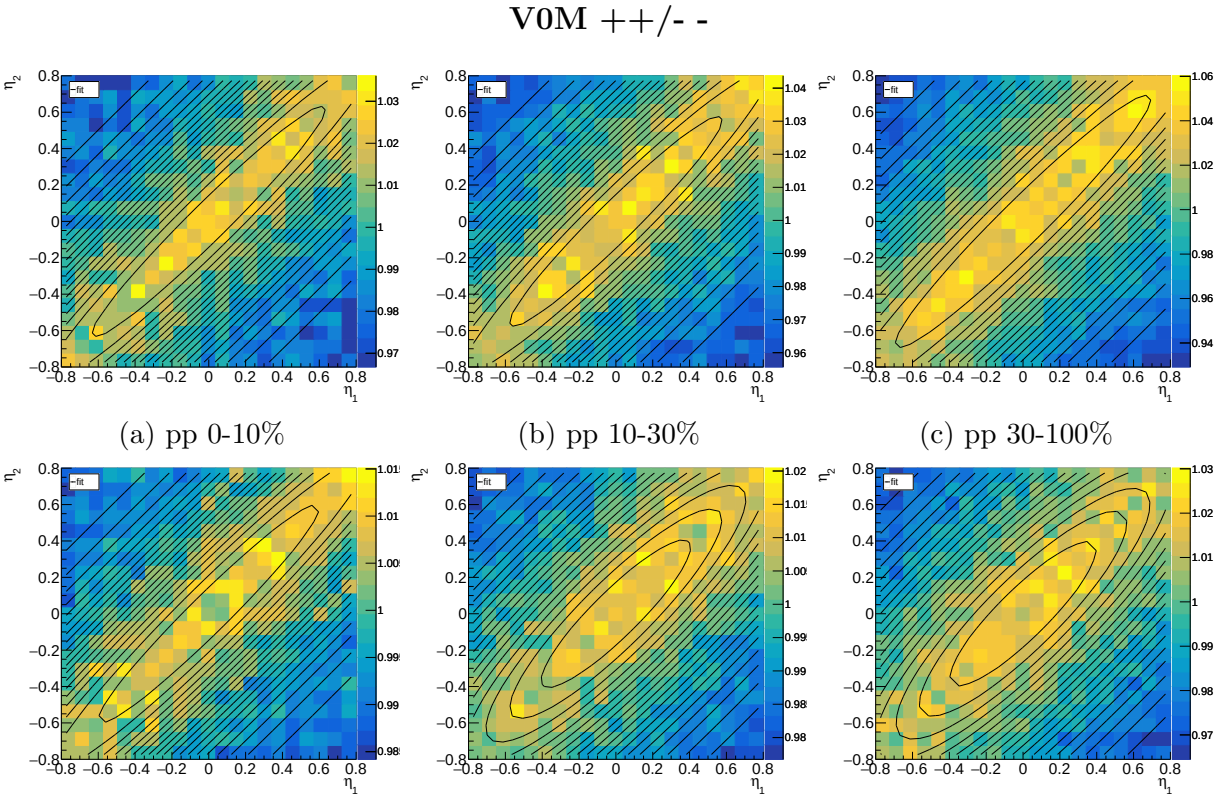
# 6 Results

Several measurements of the two-particle pseudorapidity correlation function was done for pp-collisions at $\sqrt{s} = 13$ TeV for $6 \cdot 10^5$ events, as described previously, in section 5. The analysis was repeated for three multiplicity classes (0-10%, 10-30%, and 30-100%), and for both the V0M and $N_{\text{tracklets}}$ multiplicity estimators, described in section 4.4. Additionally, same and opposite charge combinations were considered and are presented in the two succeeding sections, 6.1 and 6.2. The results are compared with PYTHIA simulations, presented in section 5.5. Fit functions were fit, described in section 5.6, to respective distribution and displayed as contour plots in the results presented below. Respective component functions of Eq. (7) can be seen in section 6.3.

## 6.1 Same Charge Combinations (++/- -)

In this section, two-particle correlations by pseudorapidity are presented, requiring the two particles to have the same charge, ++/- -. Fig. 8 shows a comparison of the three multiplicity classes for the V0M estimator, with ALICE and the simulated data in the top and bottom row respectively. Fig. 9 also compares the three multiplicity classes for real and PYTHIA data respectively, but using the $N_{\text{tracklets}}$ estimator. The results are shown together with a contour plot $F(\eta_+, \eta_-)$.

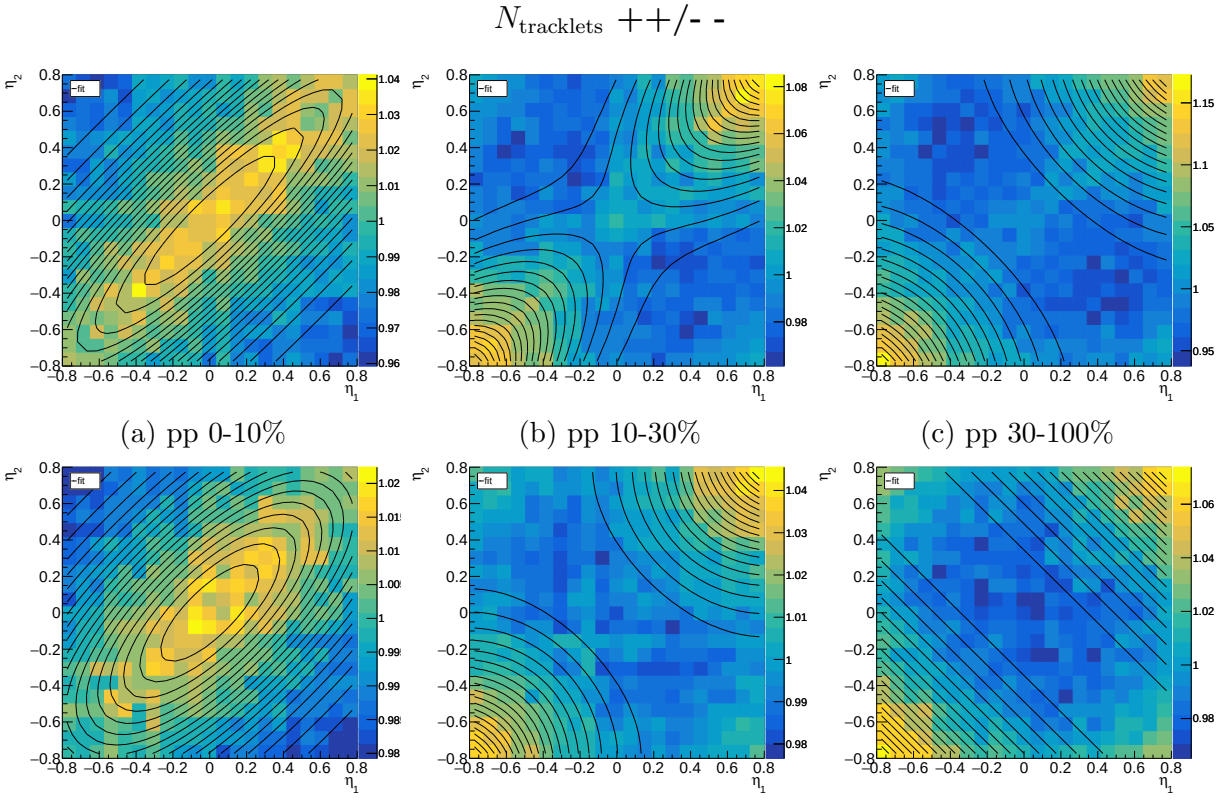
A ridge-like structure is observed in all correlation distributions with the V0M estimator, as seen in Fig. 8, and in the highest multiplicity class with the $N_{\text{tracklets}}$ estimator, as seen in the first column in Fig. 9. A decrease in correlations in the middle of the detector for lower multiplicities is seen in the second and third columns in Fig. 9. PYTHIA accurately predicts the shape of all correlation distributions, as seen in the bottom row of each figure.




(a) pp 0-10% (b) pp 10-30% (c) pp 30-100%

(d) PYTHIA 0-10% (e) PYTHIA 10-30% (f) PYTHIA 30-100%

Figure 8: Multiplicity correlations with V0M estimator and same charge combinations. The upper row shows the results obtained with ALICE data, and the second row displays the PYTHIA simulations. Each columns displays the multiplicity classes, from left to right: 0-10%, 10-30%, and 30-100%.




(a) pp 0-10% (b) pp 10-30% (c) pp 30-100%

(d) PYTHIA 0-10% (e) PYTHIA 10-30% (f) PYTHIA 30-100%

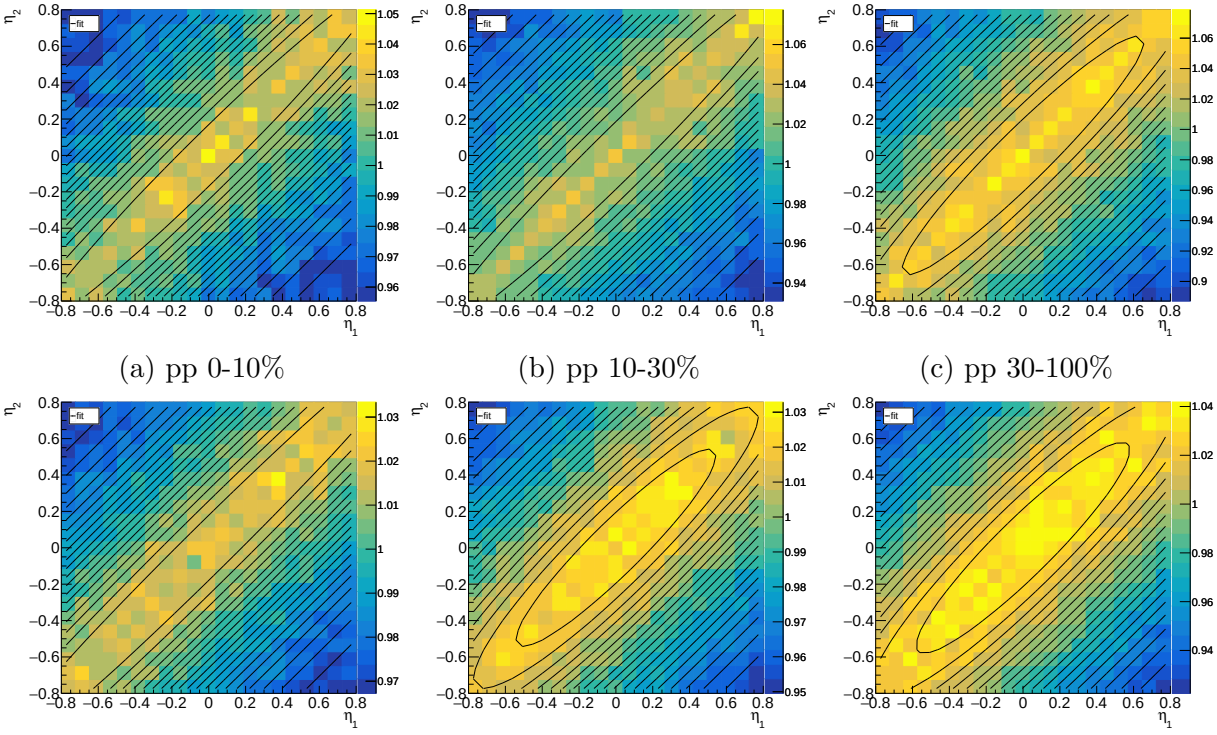
Figure 9: Multiplicity correlations with $N_{\text{tracklets}}$ estimator and same charge combinations. The upper row shows the results obtained with ALICE data, and the second row displays the PYTHIA simulations. Each column shows the multiplicity classes, from left to right: 0-10%, 10-30%, and 30-100%.

## 6.2 Opposite Charge Combinations (+-/-+)

In this section, two-particle correlations by pseudorapidity are presented, requiring the two particles to have the opposite charge, +-/-+. Fig. 10 shows the three multiplicity classes for the V0M estimator, with the top row displaying results obtained with ALICE data and the bottom row displaying the simulated data. Fig. 11 similarly compares the three multiplicity classes for real and PYTHIA data respectively, but using the $N_{\text{tracklets}}$ estimator.

A ridge-like structure is observed in all correlation distributions using the V0M estimator, as seen in Fig. 10, and in the 0-10% multiplicity class with the $N_{\text{tracklets}}$ estimator, as seen in the first column in Fig. 11. At lower multiplicities, the central multiplicity estimator displays less correlation in the middle of the detector due to autocorrelations, as seen in the second and third columns in Fig. 11. PYTHIA accurately predicts the shape of all correlation distributions, as seen in the bottom row of each figure.

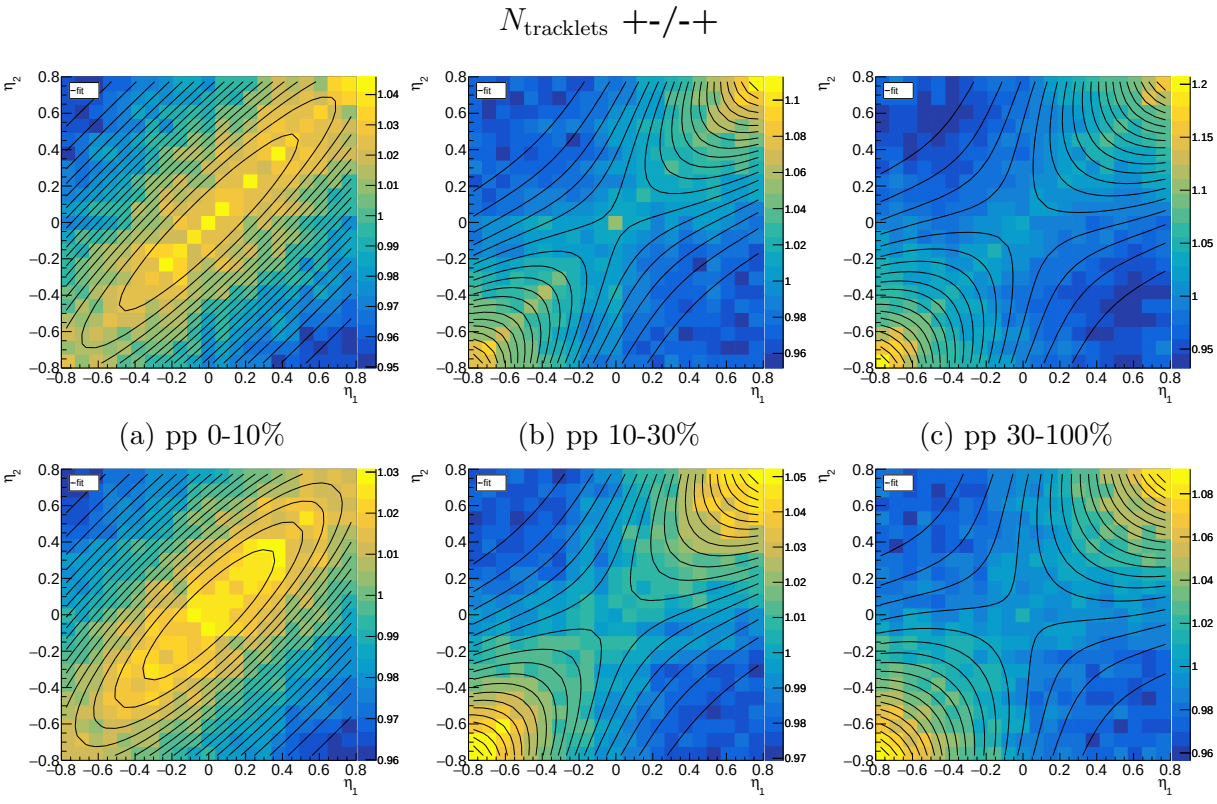
**V0M +-/-+**




(a) pp 0-10% (b) pp 10-30% (c) pp 30-100%

(d) PYTHIA 0-10% (e) PYTHIA 10-30% (f) PYTHIA 30-100%

Figure 10: Multiplicity correlations with V0M estimator and opposite charge combinations. The upper row shows the results obtained with ALICE data, and the second row displays the PYTHIA simulations. Each column shows the multiplicity classes, from left to right: 0-10%, 10-30%, and 30-100%.




(a) pp 0-10% (b) pp 10-30% (c) pp 30-100%

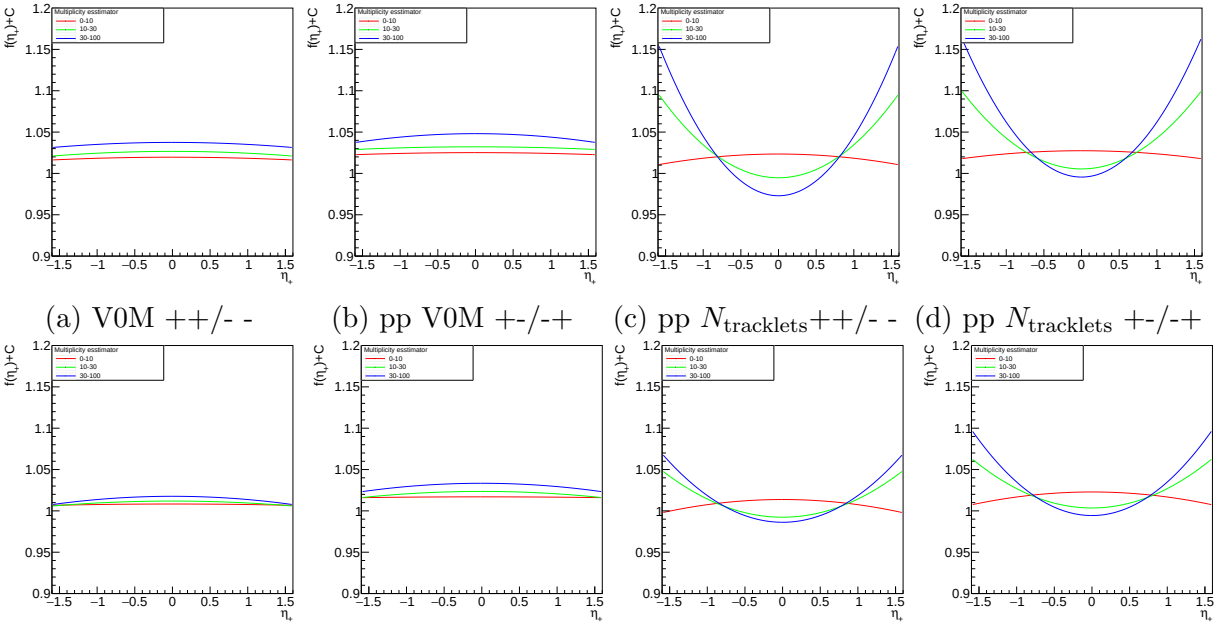
(d) PYTHIA 0-10% (e) PYTHIA 10-30% (f) PYTHIA 30-100%

Figure 11: Multiplicity correlations with $N_{\text{tracklets}}$ estimator and opposite charge combinations. The upper row displays the results obtained with ALICE data, and the second row displays the PYTHIA simulations. Each column shows the multiplicity classes, from left to right: 0-10%, 10-30%, and 30-100%.

## 6.3 Fit Functions

In this section parameters of the fit function, $F(\eta_+, \eta_-)$ (Eq. (7)) described in section 5.6, is presented. In Fig. 12 $F(\eta_+, \eta_-) = f(\eta_+) + C$ is shown for all charge combinations and multiplicity classes for the V0M and $N_{\text{tracklets}}$ estimators respectively. Fig. 13 compares the parameter $\sigma$ after fitting $F(\eta_+, \eta_-)$ to respective two-particle correlation distributions. $\sigma$ is compared for each multiplicity class, charge combination and multiplicity class for both the V0M and $N_{\text{tracklets}}$ estimators.

Both the amplitude and shape of the ridge are predicted well by PYTHIA, as seen by comparing the top and bottom rows in Fig. 12. A general decrease in ridge amplitude with higher multiplicities can also be noted. The standard deviation increases with lower multiplicities using the V0M estimator, as seen in Figs. 13a and 13b. The width of the Gaussian is more irregular for the $N_{\text{tracklets}}$ distributions and can be seen to diverge for the 30-100% multiplicity class for ++/- - combinations, see Fig. 13c.




(a) V0M ++/- -

(b) pp V0M +-/-+

(c) pp $N_{\text{tracklets}}$++/- -

(d) pp $N_{\text{tracklets}}$ +-/-+

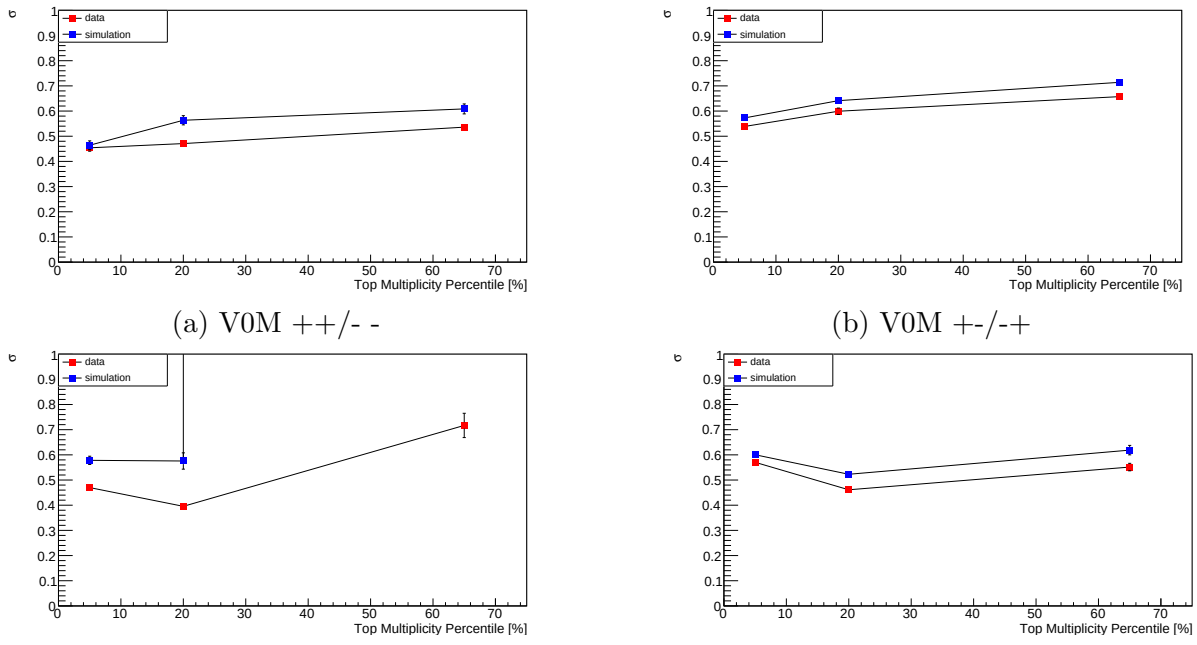
(e) PYTHIA V0M ++/- -

(f) PYTHIA V0M +-/-+

(g) PYTHIA $N_{\text{tracklets}}$ ++/- -

(h) PYTHIA $N_{\text{tracklets}}$ +-/-+

Figure 12: Respective graphs display $f(\eta_+)+C$ fitted to corresponding results presented in the previous two sections. The red, green, and blue graphs denote the multiplicity classes: 0-10%, 10-30%, and 30-100% respectively. This is shown for different multiplicity estimators, charge combinations, real data and PYTHIA simulations.




(a) V0M ++/- -

(b) V0M +-/-+

(c) $N_{\text{tracklets}}$ ++/- -

(d) $N_{\text{tracklets}}$ +-/-+

Figure 13: Shows how the fit-parameter $\sigma$ varies for the three multiplicity classes: 0-10%, 10-30%, and 30-100%, with each data point in the center of the respective interval. The red points show $\sigma$ when fitted to ALICE data, and the blue points for PYTHIA simulations. This is shown for the V0M and $N_{\text{tracklets}}$ multiplicity estimators, and for different charge combinations.

# 7 Discussion

Studying the results of the same and opposite charge combinations in Figs. 8 and 10; and Figs. 9 and 11, a stronger correlation around $\eta_- = 0$ is found for +-/-+ combinations.

This is best seen by considering the difference in ranges in these correlation distributions. Fig. 12 also shows a general increase along the ridge for opposite charge combinations. However, when comparing the standard deviation, as seen in Fig. 13, it seems that it is larger for +-/-+ combinations. This is due to the fact that we normalize all distributions, effectively decreasing the standard deviation of the larger ridge.

The fact that this result is observed is expected, due to charge conservation. When new quarks are produced the net charge must be zero, meaning quarks of different charge are always produced together ($q+\bar{q}$), therefore more short-range correlations are detected.

Comparing the results for the V0M and $N_{\text{tracklets}}$ multiplicity estimators, which can be seen in Figs. 8 and 9; Figs. 10 and 11; and in Fig. 12, a significant difference is noted. A constant correlation along the ridge, $\eta_- = 0$, is expected in accordance with the Lund string model. This corresponds to the flat plateau in rapidity discussed in section 3.3 and seen in Fig. 4.

This plateau is observed in all multiplicity classes using the V0M estimator, but the $N_{\text{tracklets}}$ estimator exhibits a quite different behavior. At lower multiplicities, 10-30% and 30-100%, the $N_{\text{tracklets}}$ pseudorapidity correlations see an enhancement for large and small $\eta_+$. In contrast to the V0M, the $N_{\text{tracklets}}$ estimator measures the multiplicity and tracks are recorded in the same region, making autocorrelations expected.

It appears that low $N_{\text{tracklets}}$ multiplicities are less driven by a low global multiplicity and more by local fluctuations. When selecting a low central multiplicity, we choose events where the color fields that produces particles dominantly covers the edges of the detector to achieve this. This reasoning explains the peaks around the edges, at $|\eta_+| \sim 1.6$, of the detector. However, when selecting on large multiplicity we instead select the events with the most recorded tracks in the central part of the detector, allowing jets to be recorded here, giving back the plateau-like ridge in the highest multiplicity class.

Now, considering how the standard deviation, $\sigma$, of the Gaussian fit-function, $F(\eta_+, \eta_-)$, varies with multiplicity, see Fig. 13. The V0M estimators exhibit a general decrease of standard deviation with multiplicity, following ATLAS [1]. However, the Gaussian width of the $N_{\text{tracklets}}$ estimator does not show such a pattern. Due to the narrow pseudorapidity range, the $N_{\text{tracklets}}$ correlation distribution cannot be fit to perfection, making the fit unrepresentative of the actual data.

Note that the standard deviation in the $N_{\text{tracklets}}$ distribution for same charge combinations in the multiplicity interval 30-100%, seen in Fig. 13c, becomes very large. This is due to this correlation distribution not containing any ridge-like structure, see Fig. 9f, making the standard deviation large.

The amplitude along the ridge, see Fig. 12, shows a general decrease with higher multiplicities, similar to the decrease of $\sigma$ with larger multiplicities. This is because, when increasing the multiplicity more strings are produced and hence more particles, but the number of particle pairs increases by the square of that, resulting in more uncorrelated pairs of particles from different strings and consequently lowering the correlations at higher multiplicities. Moreover, the normalization of the correlation function results in the lower multiplicity distributions, with more correlations, to receive a larger standard

deviation.

Analyzing the pp-collisions with PYTHIA simulations, see Figs. 12 and 13, a similar result between data and simulations is obtained. Notably, PYTHIA agrees with experiments for both the V0M and $N_{\text{tracklets}}$ estimator, even agreeing with the autocorrelations at low multiplicities.

## 7.1 Improvements and Outlook

Multiple improvements to the analysis and further studies are possible. One improvement could be to enforce additional event-mixing conditions for the ME distribution. For example, one could require the distance between the vertices of two events, which are mixed, to be within a certain distance. This is useful because when doing the SE mixing, the distance between two event-vertices is always zero, which is not necessarily the case in the ME distribution. This difference in the mixing of the SE and ME distributions might introduce some bias.

Another possible improvement could be to further correct for efficiency. As previously mentioned, the ME mixing will compensate for most detector inefficiencies, which means that efficiency has already been partially compensated for. Further efficiency-compensation of the ALICE detector, using MC simulations, would likely result in an almost uniform correction over pseudorapidity, which would cancel in the normalization step, making this correction less crucial.

Furthermore, a possible improvement to the analysis could be to perform a minimum bias comparison to see if the V0M estimator exhibits similar behavior as the $N_{\text{tracklets}}$ distributions at lower multiplicities with these additional constraints. Repeating the analysis for a larger data-set would also be of interest. In this thesis, only $6 \cdot 10^5$ events were used, and one could easily have used larger data-sets to reduce statistical errors.

Further investigation of what implications the results found in this thesis have for the ATLAS detector could be interesting. As mentioned in the introduction, the ATLAS detector only makes use of central multiplicity estimators, meaning that autocorrelations should be present. As PYTHIA simulations appear to predict autocorrelations well, one could investigate with PYTHIA if the ridge along $\eta_- = 0$ appears when considering larger pseudorapidity ranges when using the central multiplicity estimator. If so, this would explain why the ATLAS detector, with a larger pseudorapidity range, does not experience as strong autocorrelations as seen in this thesis with ALICE, because a smaller fraction of all detected strings would be at the edge of the detector.

Finally, comparing the resulting two-particle pseudorapidity distributions to that made with other simulation software, to gain insight into the advantages of each, could also be interesting. One such simulation software is EPOS, which in contrast to PYTHIA assumes a QGP formation [29].

# 8 Conclusion

This thesis aimed to study multiplicity correlations in pseudorapidity for three different multiplicity classes (0-10%, 10-30%, 30-100%), different charge combinations, and in the central and forward-backward region of ALICE, using the $N_{\text{tracklets}}$ and V0M estimators respectively. This was done for pp collisions at $\sqrt{s} = 13$ TeV. The results show strong autocorrelations with the $N_{\text{tracklets}}$ estimator at low multiplicities, and a ridge-like structure with a plateau at high multiplicity, which is also seen for all multiplicity classes with the V0M estimator. MC simulations were conducted using PYTHIA, showing similar results for all multiplicity classes, charge combinations, and multiplicity estimators. In particular, the PYTHIA results also accurately predicted the autocorrelations with the $N_{\text{tracklets}}$ estimator. Opposite charge combinations reveal an expected increase in short-range correlation due to charge conservation. Studying the standard deviation, perpendicular to the ridge, of the two-particle correlation distributions, no conclusions could be made with the $N_{\text{tracklets}}$ estimator, however, a general decrease in Gaussian width as multiplicity increases was noted for the V0M estimator. The accuracy of the PYTHIA simulations and its ability to predict autocorrelations seen in this thesis ensures its use as a high-energy physics simulation software.

# References


[1] Aaboud, M *et al.* (The ATLAS Collaboration), "Measurement of forward-backward multiplicity correlations in lead-lead, proton-lead, and proton-proton collisions with the ATLAS detector," *Phys. Rev. C*, vol. 95, p. 064914, Jun 2017.

[2] J. D. Wells, *Discovery Beyond the Standard Model of Elementary Particle Physics*. Cham: Springer, 2020.

[3] B. R. Martin and G. Shaw, *Particle physics*. Chichester: John Wiley & Sons, 3rd ed., 2013.

[4] E. Kennedy, *Particle Physics*. Rijeka: IntechOpen, 2012.

[5] Cush, Public domain, "Wikimedia commons." `https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg`. [Online; accessed: 2022-04-25].

[6] P. Langacker, *The standard model and beyond*. New York: Taylor & Francis, 2017.

[7] T. Schörner-Sadenius, *The Large Hadron Collider: Harvest of Run 1*. Springer, 2015.

[8] A. K. Chaudhuri, *A short course on relativistic heavy ion collisions*. Bristol: IOP Publishing, 2014.

[9] L. Yi, *Study of quark gluon plasma by particle correlations in heavy ion collisions*. New York: Springer, 2016.

[10] M. Aaboud and et.al, "Measurements of long-range azimuthal anisotropies and associated fourier coefficients for $pp$ collisions at $\sqrt{s} = 5.02$ and 13 tev and $p$+Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 5.02$ tev with the atlas detector," *Phys. Rev. C*, vol. 96, p. 024908, Aug 2017.

[11] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, "Parton fragmentation and string dynamics," *Physics Reports*, vol. 97, no. 2, pp. 31–145, 1983.

[12] B. Andersson, *The Lund Model*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge University Press, 1998.

[13] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual," *Journal of High Energy Physics*, vol. 2006, pp. 026–026, may 2006.

[14] T. Sjöstrand, "Old ideas in hadronization: The Lund String — a string that works —." `http://home.thep.lu.se/~torbjorn/talks/durham09.pdf`, 2009. [Online; accessed 2022-04-27].

[15] A. Datta, B. Mukhopadhyaya, and A. Raychaudhuri, *Physics at the large hadron collider*. Springer (India), 2010.

[16] The ALICE Collaboration. *et al*, "The ALICE experiment at the CERN LHC," *Journal of Instrumentation*, vol. 3, pp. S08002–S08002, aug 2008.

[17] Pcharito, Pcharito, CC BY-SA 3.0, "Wikimedia commons." https://commons.wikimedia.org/wiki/File:2012-Aug-02-ALICE_3D_v0_with_Text_(1)_2.jpg. [Online; accessed 2022-04-28].

[18] C. Ammerlaan, R. Rumphorst, and L. Koerts, "Particle identification by pulse shape discrimination in the p-i-n type semiconductor detector," *Nuclear Instruments and Methods*, vol. 22, pp. 189–200, 1963.

[19] R. Mann, *An introduction to particle physics and the standard model*. Taylor & Francis, 2010.

[20] H. J. Hilke, "Time projection chambers," *Reports on Progress in Physics*, vol. 73, p. 116201, oct 2010.

[21] K. Aamodt and et.al, "Centrality dependence of the charged-particle multiplicity density at midrapidity in pb-pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV," *Phys. Rev. Lett.*, vol. 106, p. 032301, Jan 2011.

[22] R. Brun, P. Buncic, F. Carminati, A. Morsch, F. Rademakers, and K. Safarik, "Computing in ALICE," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 502, no. 2, pp. 339–346, 2003.

[23] J. Jia, S. Radhakrishnan, and M. Zhou, "Forward-backward multiplicity fluctuation and longitudinal harmonics in high-energy nuclear collisions," *Physical Review C*, vol. 93, apr 2016.

[24] J. Adam, D. Adamová, and M. M. r. Aggarwal, "Charged-particle multiplicities in proton-proton collisions at $\sqrt{s} = 0.9$ to 8 TeV," *The European Physical Journal C*, vol. 77, no. 1, pp. 1–39, 2017.

[25] V. Vislavicius, *Identified Hadron Production as a Function of Event Multiplicity and Transverse Spherocity in pp Collisions at $\sqrt{s} = 7$ and 13 TeV with the ALICE Detector*. PhD thesis, Lund University, 2018.

[26] C. Lemieux, *Monte Carlo and Quasi-Monte Carlo Sampling*. New York: Springer, 1 ed., 2009.

[27] T. Sjöstrand, "The pythia event generator: Past, present and future," *Computer Physics Communications*, vol. 246, p. 106910, jan 2020.

[28] T. Sjöstrand, S. Mrenna, and P. Skands, "A brief introduction to pythia 8.1," *Computer Physics Communications*, vol. 178, no. 11, pp. 852–867, 2008.

[29] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko, and K. Werner, "EPOS LHC: Test of collective hadronization with data measured at the CERN large hadron collider," *Physical Review C*, vol. 92, sep 2015.