Eva Sicking

Multiplicity Dependence of Two-Particle Angular Correlations in Proton-Proton Collisions Measured with ALICE at the LHC

— 2012 —

Experimentelle Physik

Multiplicity Dependence of Two-Particle Angular Correlations in Proton-Proton Collisions Measured with ALICE at the LHC

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften im Fachbereich Physik
der Mathematisch-Naturwissenschaftlichen Fakultät
der Westfälischen Wilhelms-Universität Münster

vorgelegt von
Eva Sicking
aus Vreden

— 2012 —

Dekan:	Prof. Dr. Tilmann Kuhn
Erster Gutachter:	Prof. Dr. Johannes P. Wessels
Zweiter Gutachter:	Dr. Andreas Morsch

Tag der Disputation:	26.09.2012
Tag der Promotion:	26.09.2012

Contents

Introduction

The Large Hadron Collider (LHC) produces collisions of protons and lead ions at so far unprecedented ultra-relativistic collision energies. The LHC is located at the European Organization for Nuclear Research (CERN) in Geneva. It began operating in November 2009. The high-energy particle collisions at the LHC are studied by six collider experiments. By recording the properties of the collision products, the experiments try to explore the subatomic structure, the properties of the fundamental forces and the elementary particles, as well as the properties of nuclear matter. Using different detector concepts, the collider experiments cover a broad range of high-energy physics research topics.

One of the LHC experiments is ALICE (A Large Ion Collider Experiment). ALICE's main goal is to study the quark gluon plasma (QGP) by means of heavy-ion collisions. The QGP is a state of nuclear matter which is assumed to have existed shortly after the big bang. ALICE also studies proton-proton collisions to provide important reference measurements for the heavy-ion measurements, but also as part of a unique proton-proton physics program. In particular, high multiplicity proton-proton collisions are an interesting field to study particle production mechanisms. The high sensitivity to low-momentum particles in ALICE due to its low material budget and low magnetic field allows ALICE to make a unique contribution to studies of the particle production mechanisms in proton-proton collision at the LHC.

Motivated by the composite structure of the proton, the role of multiple parton interactions in high-energy proton-proton collisions has been discussed for many years. In one proton-proton collision, several pairs of partons can collide with each other, with each parton-parton interaction giving rise to production of final state particles.
At LHC collision energies, multiple parton interactions are considered to play an important part in the production of particles. Multiple parton interactions significantly affect many physics observables, such as the charged particle multiplicity. Furthermore, they are assumed to have a sizable impact on the background of many discovery channels at the LHC. Hence, a good understanding of multiple parton interactions is a precondition for many physics analyses at the LHC.

In the past, the physics of multiple parton interactions has been studied mainly by exploring the shape of the charged particle multiplicity distributions and multi-jet events. Analyses of the inclusive charged particle multiplicity include particles down to the lowest accessible momenta; however, the information about the correlation in the particle production is not obtained in this approach. Jet analyses provide information about the correlation in the particle production; however, only those jets are analyzed which have energies significantly larger than the underlying event.

In this thesis, a new, combined analysis approach is presented that comprises both the particle production at low momenta as well as the information about the correlation in the particle production. For this purpose, an analysis of charged particle multiplicity distributions is combined with a jet analysis to a new, powerful tool for the study of jet fragmentation and multiple parton interactions. In order to include jets down to the lowest possible jet energies, the jet analysis is performed using two-particle angular correlations instead of jet reconstruction on an event-by-event basis. Observing the strength of the correlation as a function of the charged particle multiplicity reveals jet fragmentation properties as well as the contribution of jets to the overall charged particle multiplicity. Furthermore, the correlation in the particle productions allows the characterization of the underlying multiple parton interactions.

This thesis is structured as follows: Chapter 1 introduces the concept of multiple parton interactions and reviews theoretical models as well as experimental observations of multiple parton interactions.
Chapters 2 and 3 briefly describe the Large Hadron Collider and the ALICE experiment. Chapter 4 describes the data sets which are used for the data analysis in combination with the setup specifications of the LHC and the ALICE experiment during recording of the data. Also, the specification of the event and track selections used in the data analysis is given. For the data set selection, a quality assurance analysis is developed.
In Chapter 5, the analysis method of two-particle angular correlations is presented. A correction procedure for the proton-proton collision data recorded by ALICE is developed in Chapter 6. The correction accounts for all relevant detector effects, including the tracking efficiency of the detector and the contamination of the track sample with tracks coming from secondary particles. In Chapter 7, the systematic uncertainties related to the measurement are discussed and evaluated.
The results of the two-particle angular correlation analysis for the three center-of-mass energies ($\sqrt{s} = 0.9$, 2.76, and 7.0 TeV) are presented and discussed in Chapter 8 including comparisons to theoretical model descriptions. The results are discussed in the framework of multiple parton interactions in proton-proton collisions.

1. Theoretical Background

In this chapter, the theoretical background of multiple parton interactions (MPI) in high-energy proton-proton is introduced. First, the basic concepts of the standard model of particle physics are described in Section 1.1. Then, the importance of the study of multiple parton interactions is motivated (Section 1.2).
Signatures of multiple parton interactions which have been studied in the past are summarized in Section 1.3 followed by the description of theoretical models of multiple parton interactions in Section 1.4. Also, implications of multiple parton interactions resulting in proton-proton collisions events of high charged particle multiplicities are discussed (Section 1.5). Finally, Section 1.6 introduces the Monte Carlo generators based on multiple parton interactions that are used in the comparison to real collision data.

1.1. The Standard Model of Particle Physics

The standard model of particle physics (SM) is a theory of the electro-magnetic interaction, the weak interaction, and the strong interaction. These three interactions together with the gravitation (not included in the SM) represent all known fundamental interactions between the subatomic particles, the quarks and the leptons. In the SM, each force is mediated by force carriers, so-called gauge bosons of spin 1. The interactions are characterized in Table 1.1.

Force	Gauge boson(s)	Applies to
strong	8 gluons g	quarks
electro-magnetic	photon γ	charged particles
weak	$\mathrm{W}^{\pm}$, Z^0	quarks, leptons

Table 1.1.: The fundamental forces of the standard model of particle physics [PDG10].

The strong interaction between color-charged particles is mediated by eight gluons (g) that carry combinations of color charges (cf. Section 1.1.1). The electro-magnetic interaction between charged particles is mediated by photons (γ). The weak interaction between particles of different flavors is mediated by the massive particles W^+, W^-, and Z^0. Similarly, gravitation could possibly be mediated by hypothetical gravitons. At the time of writing, however, this particle has not yet been observed.
Besides the gauge bosons of the interactions, the SM comprises the quarks, the leptons, and the Higgs boson. To the present knowledge, these particles are elementary particles.

All quarks and leptons have spin $\frac{1}{2}$, hence, they belong to the group of fermions [PDG10]. The quarks and leptons are subdivided into three so-called generations, each containing two quarks or two leptons respectively. Table 1.2 lists all known quarks and leptons together with their electrical charge and mass.

	Quarks			Leptons		
Gen.	Symbol	Charge	Mass	Symbol	Charge	Mass
1	u	2/3 e	1.7 - 3.3 MeV/c^2	e^-	-e	0.511 MeV/c^2
	d	-1/3 e	4.1 - 5.8 MeV/c^2	ν_e	0	$\leq$ 2 eV/c^2
2	c	2/3 e	1.18 - 1.34 GeV/c^2	μ^-	-e	105 MeV/c^2
	s	-1/3 e	80 - 130 MeV/c^2	ν_μ	0	$\leq$ 0.19 MeV/c^2
3	t	2/3 e	169.8-174.2 GeV/c^2	τ^-	-e	1.78 GeV/c^2
	b	-1/3 e	4.1 - 4.4 GeV/c^2	ν_τ	0	$\leq$ 18.2 MeV/c^2

Table 1.2.: Table of the elementary particles and their masses as described in the standard model of particle physics [PDG10].

The six different quarks are called up (u), down (d), strange (s), charm (c), top (t), and bottom (b). For each quark (q), there is also an anti-quark ($\overline{q}$) of same mass, opposite electric charge, and opposite baryon number.

Composite particles made up of quarks are called hadrons. Hadrons are subdivided into mesons and baryons. Mesons consist of two valence quarks, i. e. a quark-anti-quark pair. Baryons comprise three valence-quarks. The anti-symmetry of the baryon wave function is maintained by the so-called color degree of freedom. This color charge is the charge related to the strong interaction. It has three possible values. Quarks carry such a color charge, usually called (anti-)red, (anti-)blue, or (anti-)green. Either three quarks of red, blue, and green color or two quarks of color and anti-color build color neutral objects. Quarks are always grouped together in color-neutral hadrons of two or three quarks. Isolated color charge has never been observed in nature as well as objects that constitute of more than three valence quarks [PDG10]. In addition to the color charge, quarks carry an electric charge of $-\frac{1}{3}\,e$ or $\frac{2}{3}\,e$.

The group of leptons comprises the electron (e), the muon (μ), and the tau (τ), each carrying an electric charge of $-e$, and the neutral neutrinos of the same generations, namely the electron neutrino (ν_e), the muon neutrino (ν_μ), and the tau neutrino (ν_τ). For each of the six described leptons, there are also anti-leptons of same mass, opposite charge, and opposite lepton number.

The first generation of particles of the SM contains the lightest and most stable particles.

With each generation, the particles are heavier and less stable. The stable particles present in the universe are only made up of particles of the first generation, e. g. hadrons like protons (uud), neutrons (udd), and electrons. Heavier elementary particles decay quickly into more stable particles, e. g. heavy quarks like s, c, t, and b decay weakly into u and d.
Furthermore, the standard model predicts the existence of the Higgs boson [Hig64, EB64, GHK64]. The Higgs boson is assumed to be massive however its mass is not predicted by the SM. The Higgs theory is an important building block of the standard model giving an explanation why some of the elementary particles are massive.
At the time of writing, clear signs of a new boson at approximately $126\,\mathrm{GeV}/c^2$ are observed at the Large Hadron Collider. One of the analyzed decay channels of the Higgs boson is the decay into two photons ($H \rightarrow \gamma + \gamma$). An invariant mass spectrum of two photons measured at the LHC is shown in Figure 1.1.
It is assumed that this particle is identical with the SM Higgs boson, however further data analysis is needed before the new boson can be positively identified as Higgs boson beyond doubt.

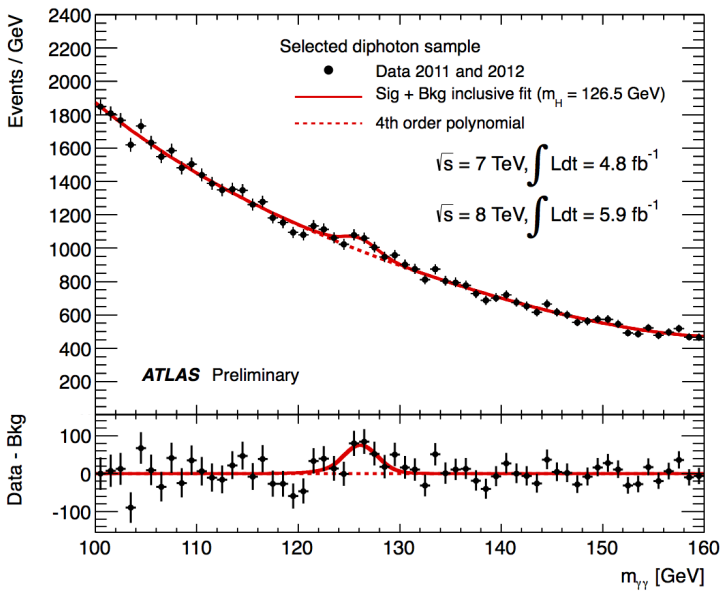

Figure 1.1.: The invariant mass spectrum of two photons measured at the LHC in comparison to a background model fit [Fab12]. The spectrum reveals a peak at approximately 126 GeV.

Even though, the SM is very successful in describing the physics of the elementary particles including three of the know interactions, several questions remain unsolved [Ell02]. First of all, the SM does not include a theory of gravitation. The Higgs boson mass is not predictable in the SM despite its fundamental role. In addition to the free Higgs mass, the SM comprises several free parameters such as fermion Yukawa coupling matrices, the coupling constants, and the vacuum expectation value of the Higgs field, all of which can not be derived from first principles within the SM. The SM does not include dark matter or dark energy. The neutrino oscillations that indicate a neutrino mass larger than zero are not explained within the SM. Also, in the SM, there is no explanation for the rather large asymmetry between matter and anti-matter as it is observed in our universe. Finally, the SM has a so-called hierarchy problem. From radiative corrections, a rather large Higgs mass is expected related to the vast scale difference of the weak and the gravitational interaction. The strength of the gravitation is by a factor 10^{-32} weaker than the weak interaction.
Extensions to the standard model as well as new theories have been proposed in order to solve the afore mentioned problems. For example, there are super-symmetric theories (SUSY) (an introduction and an overview can be found in [Mar97]), in which each fermion and each boson of the standard model has a super-symmetric partner particle. Another theory, the so-called String theory [Sch82, Gre84] incorporates aspects of SUSY and allows to combine quantum mechanics with general relativity.
In the following section, one part of the standard model, Quantum Chromodynamics (QCD), is discussed in more detail. QCD is the theory of the strong interaction between color charged particles.

1.1.1. Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the quantum theory of the strong interaction. It describes the strong interaction between the quarks and gluons. The strong interaction is mediated by the gauge bosons of the strong interaction, the gluons. In total, eight different gluons exist. Each gluon carries a combination of color and anti-color charges. This is in contrast to Quantum Electrodynamic (QED), in which the corresponding gauge boson, the photon, does not carry any electric charge. The color charge of the gluons results in a self-interaction between gluons.
The strength of the strong interactions is described by the coupling constant of the strong interaction α_s, which is given as a function of a renormalization scale μ_R^2 [tH74]. For the estimation of α_s, μ_R^2 can be chosen close to the scale of the momentum transfer Q of a given process. It is found that α_s strongly depends on the momentum transfer. For large

momentum transfers and $Q^2/\Lambda^2 \ll 1$, α_s can be approximated perturbatively in leading order [Roe96, Bet09] with

$$\alpha_s(Q^2) \approx \frac{12\pi}{(33 - 2 \cdot N_\mathrm{f}) \cdot \ln \frac{Q^2}{\Lambda^2}}. \tag{1.1}$$

Here, Λ is the QCD scale parameter with an approximate value of $\Lambda \approx 200\,\mathrm{GeV}$ [Roe96, Bet09]. It represents the scale at which the perturbative coupling would diverge, if extrapolated outside the perturbative domain. N_f is the number of active flavors. Figure 1.2 shows the energy dependence of α_s as a function of the energy scale Q.

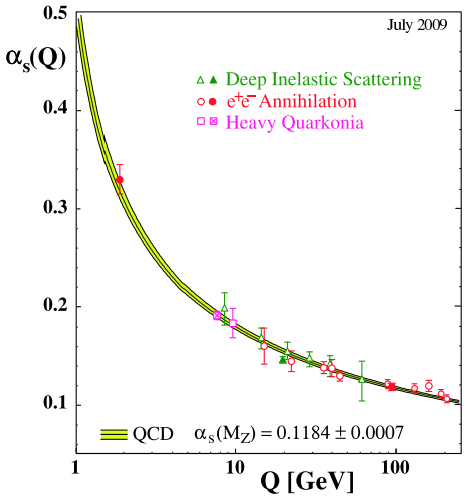

Figure 1.2.: Summary of measurements of α_s as a function of the respective energy scale Q [Bet09].

With increasing momentum transfers (correspondingly at short distances) the coupling becomes weaker. At asymptotically large momentum transfers, α_s approaches zero. This means that the quarks can behave as if they are free. This feature of the strong interaction is called asymptotic freedom.

In contrast, at low momentum transfers, the coupling increases. For low momentum transfers and $Q \approx \Lambda$, α_s becomes large and the perturbative approach is not valid anymore. The energy scale dependence of α_s is presented in Figure 1.2.

A phenomenological potential of the strong force between two quarks as a function of the distance r [Per00] is given as

$$V_s(r) = -\frac{4}{3}\frac{\alpha_s \hbar c}{r} + kr. \tag{1.2}$$

In contrast to the potential of the QED, the QCD potential between e. g. two quarks does not vanish for large distances r but grows linearly with r.

The implications of the fact that two quarks can not be separated easily can be discussed

in an intuitive picture. If the quarks are pulled apart from each other, the energy between the quarks increases until it is energetically more favorable to generate a new pair of quark and anti-quark out of the energy of the string between the quarks. This results in two quark pairs, each pair with a shorter distances r. This effect explains why quarks and gluons have not yet been observed as free particles. The effect that two quarks can not be separated easily is called color confinement.

QCD in High-Energy Collisions

The structure of the strongly interacting proton has been explored in deep inelastic scattering experiments of electrons (e) and protons (p) [ESW96]. The results of the ep scattering experiments suggested that the electrons were undergoing elastic scattering-off of constituent objects inside the proton. The proton behaves like a collection of weakly interacting, charged, point-like constituents called partons (quarks and gluons) [GM64, Fey69]. The fraction of the hadron energy carried by the parton is given by the Bjorken scaling variable x [Bjo69].
Parton Distribution Functions (PDF) can be understood as the probability distribution of partons to carry a momentum fraction x of the hadron in a high-energy collision with momentum transfer Q^2. In the example of the proton, beside the valence quarks (up, up, and down) also gluons and sea quarks carry fractions of the proton momentum. Sea quarks are quark pairs generated in quark-antiquark pair production. Gluons dominate at low x.
PDFs are parametrizations of experimental measurements of deep inelastic scatterings. PDFs are a necessary input for all theory predictions at hadron colliders [Pla11]. Figure 1.3 shows as example the CTEQ5L-PDF [L$^+$00] which is the default PDF used in the Pythia event generator [SMS06, SMS08] (cf. Section 1.6.1).

The QCD factorization theorem states that for some processes the decomposition of the matrix elements into independent building-blocks is allowed. In this way, any QCD cross-section can be separated into two parts: the process-dependent short-distance parton cross-sections which are calculable in perturbative QCD and the universal long-distance contribution, the PDFs. Thus, using this two blocks found and parametrized from data can be used for predictions of some other processes.
The perturbative approximation of the QCD has been very successful in describing hard quark, anti-quark and gluon scatterings with large momentum transfers. Perturbative QCD (pQCD) is established as the main tool for studying hard interactions of quarks and gluons. So-called soft process at low small transverse momenta that are the dominant processes of hadron collisions, can not be easily described by QCD, as the perturbative

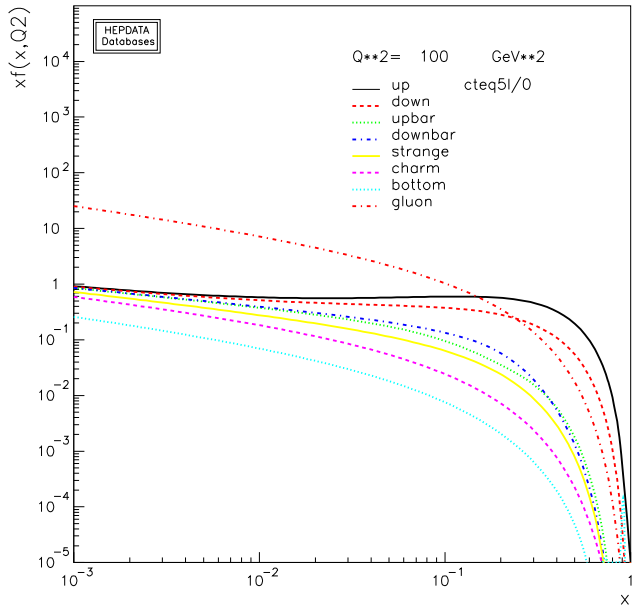

Figure 1.3.: The CTEQ5L parton distribution function [L$^+$00] illustrating basic features of PDFs. At low x, the gluons dominate. At high x, the quarks dominate. In the high x range, the up quark contribution is twice as high as the down quark contribution. This ratio corresponds to the number of the constituent quarks of the proton (u,u,d). The plot is generated using an online PDF plotting and calculation tool [Dur12].

approach is not valid in this region. Therefore, other phenomenological models need to be used to describe the soft processes.

1.2. Motivation for the Study of Multiple Parton Interactions

Hadrons traveling at increasing velocities close to the velocity of light appear to be made up of an increasing number of point like partons [Fey69, Bjo69]. This can be observed due to their de Broglie wavelength that becomes much smaller than the size of the hadron at these energies. In high-energy hadron collisions, the inner hadron structure can be resolved.

Following this observation, high-energy proton-proton (pp) collisions can be interpreted as collisions of two "bunches of partons". Hence, it is possible that multiple distinct pairs of partons collide in a single high-energy proton-proton collision. This process is called "multiple parton interaction" (MPI). Figure 1.4 shows a schematic illustration of a proton-proton collision in which two partons collides with each other.

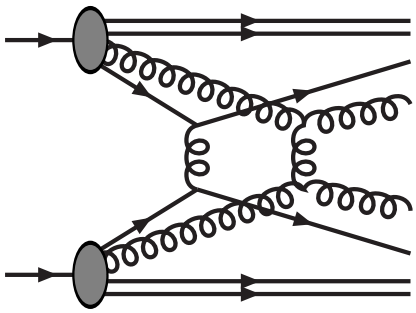

Figure 1.4.: Schematic illustration of a proton-proton collision with two $2 \rightarrow 2$ perturbative interactions [SS04]. The circles display the protons. The lines display the partons, while straight line indicate constituent quarks and the curled lines represent gluons.

With increasing center-of-mass energy in high-energy hadron collisions, multiple parton interactions become increasingly important [Wal04, SS04, BF$^+$10]. It is predicted that MPI at LHC energies have significant impact. Integrated observables like the charged particle multiplicity, transverse momentum spectra, the underlying event, as well as multi-jet events depend on the number of MPI. Hence, the MPI may represent a significant background to many discovery channels, e. g. the channels of the Higgs particle or particles of the super-symmetric extension of the standard model [SS04, BF$^+$10]. For a proper understanding of the collision event background it is important to understand the mechanisms and the cross-sections of multiple parton interactions in high-energy collisions events.

1.3. Signatures of Multiple Parton Interactions

In the following section, the most important experimental signatures of double and multi-parton interactions and their historical discovery are discussed.
First, charged particle multiplicity distributions and their scaling with the collision energy are discussed in Section 1.3.1. In Section 1.3.2, multi-jet events are presented. Finally, the growth of the total cross-section with the center-of-mass energy (Section 1.3.3) and the growth of the J/Ψ production as function of the charged particle multiplicity (Section 1.3.4) in pp collisions are discussed.

As part of the presented analyses, often a classification of the processes and the collision events into “soft”, “semi-hard”, and “hard” is applied. It has to be mentioned, that even though the naming schemes sound similar, slightly different definitions of these terms are used, which can lead to confusion.
In general, hard interactions include transverse momentum transfers above an afore

defined transverse momentum threshold. Hence, events containing hard interactions should be called hard events. On the other hand, soft interactions are defined by having only small momentum transfers. Soft events should be those events that do not contain hard interactions. The name "semi-hard" is often used for interaction of transverse momentum transfers just below the hard threshold.
In all examples, in which the definition of soft and hard *events* is used in the following, the reader is asked to interpret the soft fraction of events as those events coming from single parton interactions. The so-called hard events should be read as events including more than one parton interactions. Whether the single or the multiple parton interactions are hard or soft interactions is not specified. Single parton interactions can be both, soft and hard. This includes the possibility that all parton interactions in an event with multiple parton interactions are soft.

1.3.1. Charged Particle Multiplicity Distributions

Koba-Nielsen-Olesen Scaling Violations

At low collision energies, the mean number of particles $\langle n \rangle$ produced in a high-energy particle collision was found to rise logarithmically with the center-of-mass energy $\sqrt{s}$ (cf. Appendix A).

$$\langle n \rangle \propto \ln \sqrt{s} \; . \tag{1.3}$$

This follows from Feynman scaling of particle production in which the cross section is assumed to be a function of x and p_{T}. Based on Feynman scaling, it was suggested that the evolution of the charged particle multiplicity distributions $P(n)$ as a function of the center-of-mass energy of the colliding system can be described by the Koba-Nielsen-Olesen scaling (KNO) [KNO72]. In the KNO scaling, the charged particle multiplicity distribution $P(n)$ is expressed by the following formula

$$P(n) = \frac{1}{\langle n \rangle} \cdot \psi \left(\frac{n}{\langle n \rangle} \right) + \mathcal{O} \left(\frac{1}{\langle n \rangle} \right) . \tag{1.4}$$

Here, $\psi(z = n/\langle n \rangle)$ is an energy independent function. In the approximation of neglecting the corrections of the second term of Equation 1.4, $\mathcal{O}\left(1/\langle n \rangle\right)$, the charged particle multiplicity distributions in KNO variables ($n \rightarrow n/\langle n \rangle$ and $P(n) \rightarrow P(n)\langle n \rangle$) at all center-of-mass energies fall on top of each other, if KNO scaling is fulfilled.

The left panel of Figure 1.5 shows that the charged particle multiplicity distribution in KNO variables for proton-proton collision energies below $\sqrt{s} = 62.2\,\mathrm{GeV}$ lie on top of each other [B$^+$84]. The KNO scaling is fulfilled for these center-of-mass energies.
Hadron-hadron collisions at larger center-of-mass energies violate the KNO scaling [UA585, UA589, AAB$^+$98] (right panel of Figure 1.5). The violation of the KNO scaling is interpreted as an onset of double (or multi) parton-parton collisions and mini-jet produc-

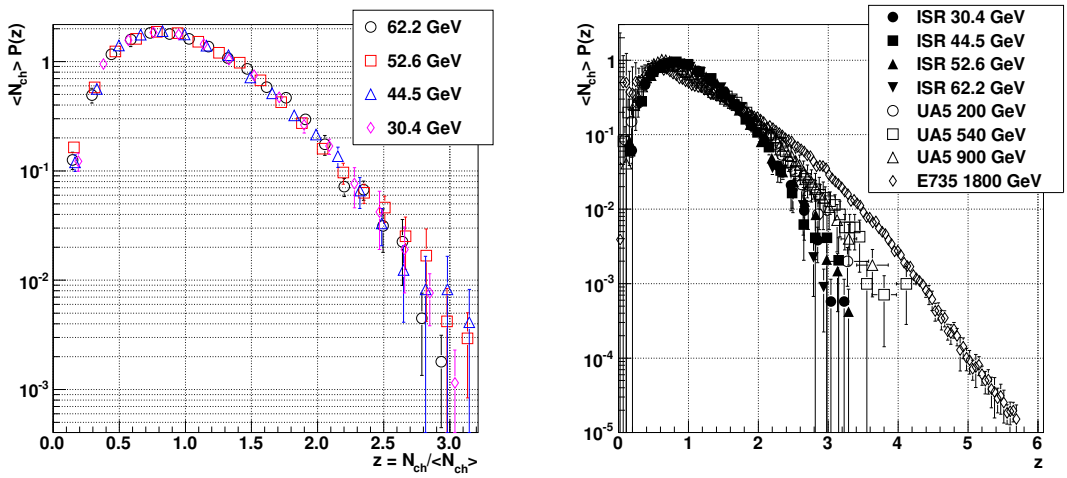

Figure 1.5.: Left panel: Charged particle multiplicity distribution of non-single diffractive (NSD) collision events in full phase space in KNO variables at energies up to $\sqrt{s}$ = 62.2 GeV fulfill the KNO scaling [B$^+$84, GOR10]. Right panel: Charged particle multiplicity distributions of NSD events in full phase space in KNO variables for energies above $\sqrt{s}$ = 200 GeV do not fulfill KNO scaling [B$^+$84, UA585, UA589, AAB$^+$98, GOR10]. For example, the charged particle multiplicity distribution measured at 1800 GeV (◇) deviates at high charged particle multiplicities from the other distributions.

tion [AFS87, UA291, CDF93, CDF97b, CDF97a, Wal04]. The effect is observed to become stronger with increasing center-of-mass energy of the colliding hadrons [AAB$^+$98, MW99].

Negative Binomial Distributions

In order to study the nature of KNO scaling violation further, the charged particle multiplicity distributions are parametrized with a negative binomial distribution (NBD). The NBD is given by

$$P^{\mathrm{NBD}}_{p,k}(n) = \begin{pmatrix} n+k-1 \\ n \end{pmatrix} (1-p)^n \, p^k. \tag{1.5}$$

For processes of success probabilities p, $P^{\mathrm{NBD}}_{p,k}(n)$ gives the probability for n failures and $k-1$ successes in any order before the k'th success.
A common representation of the NBD used to describe charged particle multiplicity distributions can be derived by relating p to the mean of the charged particle multiplicity distribution $\langle n \rangle$ by

$$p^{-1} = 1 + \langle n \rangle / k. \tag{1.6}$$

Hence, Equation 1.5 can be transformed to

$$P^{\mathrm{NBD}}_{\langle n \rangle,k}(n) = \begin{pmatrix} n+k-1 \\ n \end{pmatrix} \left(\frac{\langle n \rangle / k}{1 + \langle n \rangle / k} \right)^n \frac{1}{(1 + \langle n \rangle / k)^k}. \tag{1.7}$$

The fact that charged particle multiplicity distributions can be described by NBD distributions can be derived using the clan model [Eks85, GVH86, GVH88] introduced in Section 1.4.2.
NBDs have been used to describe charged particle multiplicity distributions measured in proton-proton collisions at e. g. $\sqrt{s}$ = 200 and 900 GeV [UA585, UA586]. However, it has been observed that a single NBD deviates from the charged particle multiplicity distribution if the charged particle multiplicity is measured in the full phase space. In this case, a shoulder structure in the charged particle multiplicity distribution appears.
Charged particle multiplicity distributions measured in pp collisions at $\sqrt{s}$ = 900 GeV have been described with better agreement by a combination of two independent NBDs [Fug89]. The sum of the two NBDs can be expressed as

$$P(n) = \alpha_1 \cdot P^{\mathrm{NBD}}_{\langle n\rangle_1,k_1}(n) + (1-\alpha_1)\cdot P^{\mathrm{NBD}}_{\langle n\rangle_2,k_2}(n), \tag{1.8}$$

each of the NBDs with independent k_i and $\langle n\rangle_i$ parameters ($i = 1, 2$). The two-NBD approach gives a much better description of the charged particle multiplicity distribution (right panel of Figure 1.6) than the single-NBD approach (left panel of Figure 1.6).

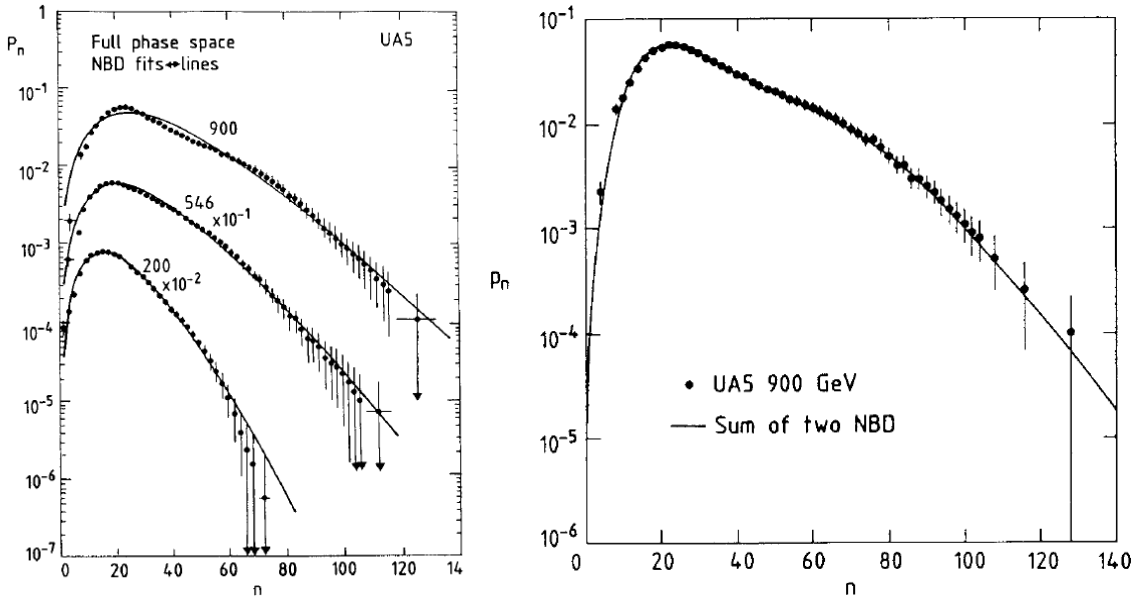

Figure 1.6.: Left panel: The best fitted single NBDs compared to the charged particle multiplicity distributions in full phase space at $\sqrt{s}$ = 200, 546, and 900 GeV [Fug89]. Right panel: Best fit of the sum of two NBDs compared to the charged particle multiplicity distribution in full phase space at $\sqrt{s}$ = 900 GeV [Fug89].

The fact that two NBD components describe the data with a better agreement can be interpreted as an indication that more than one source or process of particle production in

high-energy collisions exists [Gho12]. However, it has to be mentioned that a fit function with an increased number of parameters most commonly gives a better description of a data distribution than a fit function with a lower number of parameters.
The two sources of the particle production are further interpreted as "soft" ("1") and "semi-hard" ("2") processes [GVH86]. While the "soft" fraction of the charged particle multiplicity distribution fulfills KNO scaling, the "semi-hard" fraction violates it. The $\langle n \rangle_{\mathrm{semi-hard}}$ is larger than $\langle n \rangle_{\mathrm{soft}}$.

Decomposition of Charged Particle Multiplicity Distribution

As discussed in Section 1.3.1, the KNO scaling of charged particle multiplicity distributions is violated at energies above $\sqrt{s} = 62.2\,\mathrm{GeV}$ (maximum energy at the Intersecting Storage Ring ISR [Sch68]), below this energy, the KNO scaling is fulfilled. In the left panel of Figure 1.7, the charged particle multiplicity distributions measured at energies above ISR energies are presented [AAB$^+$98]. The solid line in this figure represents a fit of ISR data in KNO variables. The charged particle multiplicities have been scaled by $\langle n_1 \rangle$ for all data sets to $x = n/\langle n_1 \rangle$. Here, n_1 is the average charged particle multiplicity of that part of each charged particle multiplicity distribution that follows KNO scaling. All displayed cross-sections $\mathrm{d}\sigma/\mathrm{d}x$ have been normalized to the maximum value.

The left panel of Figure 1.7 shows that a fraction of the charged particle multiplicity distributions at all center-of-mass energies fulfills KNO scaling. This fraction is interpreted as being initiated in those proton-proton collisions that obey single parton collisions only [AAB$^+$98, Wal01, Wal04]. The average charged particle multiplicity of single parton collisions $\langle n_1 \rangle$ rises with the center-of-mass energy (plot not shown) [Wal04].
The fraction of the charged particle multiplicity distributions presented in the left panel of Figure 1.7, which does not fulfill the KNO scaling, shows a peak at approximately $x = n/\langle n_1 \rangle = 2$. This fraction is interpreted as mainly being due to double parton interactions, to a lesser extent also due to higher order interactions [AAB$^+$98, Wal01, Wal04]. The threshold of the onset of double parton interactions is estimated to be located between $\sqrt{s} = 100$ and $200\,\mathrm{GeV}$ [Wal04]. It is also estimated that the threshold for triple parton interactions is slightly below $\sqrt{s} = 546\,\mathrm{GeV}$ [Wal04].
The right panel of Figure 1.7 shows the decomposition of a charged particle multiplicity distribution measured at $\sqrt{s} = 1.8\,\mathrm{TeV}$. The decomposition of the charged particle multiplicity distribution is performed in two steps. First, the part of the charged particle multiplicity distribution that follows KNO scaling is subtracted. For the data set at $\sqrt{s} = 546\,\mathrm{GeV}$, the remaining fraction of the charged particle multiplicity distribution is assumed to be solely due to double parton interactions. Hence, the shape of this distribution can be used for the second subtraction step applied to charged particle multiplicity

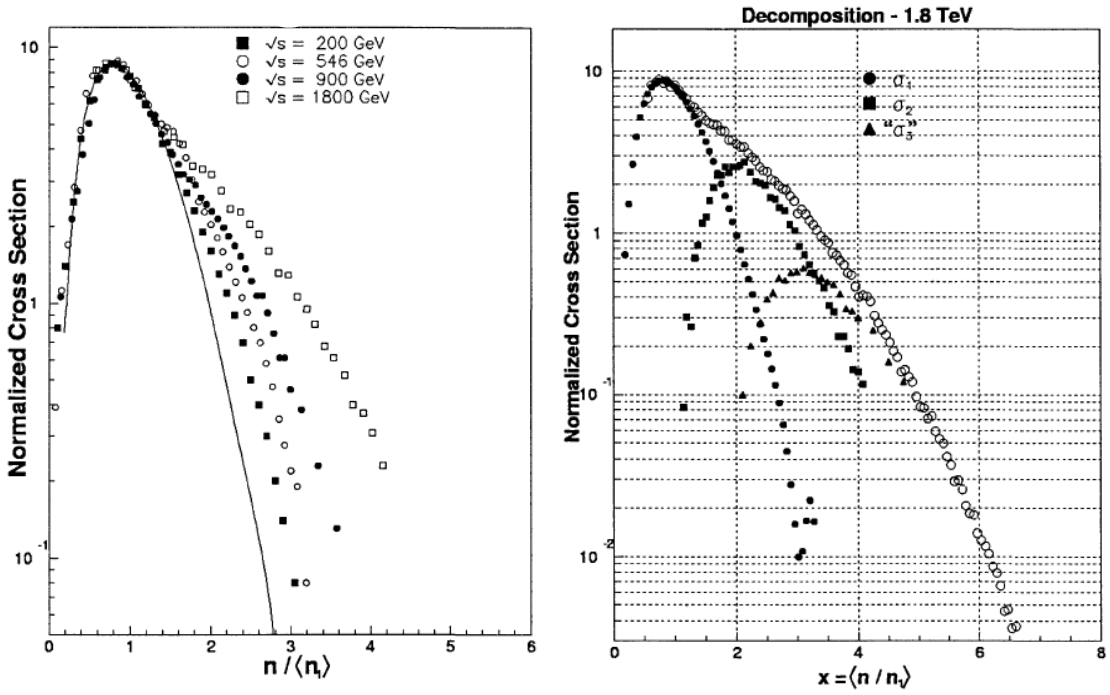

Figure 1.7.: Left panel: A comparison of charged particle distributions at different center-of-mass energies. The solid line is the KNO distribution represents a fit of ISR data in KNO variables. The distributions have been normalized to the maximum value of $d\sigma/dx$ with $x = n/\langle n_1 \rangle$ [AAB+98]. Right panel: The decomposition of the charged particle multiplicity distribution at $\sqrt{s} = 1.8\,\text{TeV}$. The charged particle multiplicity generated in pp collisions of 1, 2, and 3 parton-parton interactions is shown [Wal01, Wal04].

distributions of higher collisions energies.

The decomposition of the charged particle multiplicity distribution reveals that single, double, and triple parton interactions contribute to the charged particle multiplicity distribution at $\sqrt{s} = 1.8\,\text{TeV}$ (right panel of Figure 1.7). The curve of the triple parton interaction may also contain multi-parton interactions of even higher orders ($n_{\text{parton}-\text{parton}} > 3$) [Wal01, Wal04]. The study leads to the assumption, that multi-parton interactions become increasingly important with increasing collision energy [Wal01, Wal04]. Is is observed that the events with single parton interactions are probably identical at all collision energies.

1.3.2. Multi-Jet Events

When quarks and gluons are produced in hard (high-p_{T}) parton scatterings, they fragment and hadronize to collimated sprays of energetic hadrons which are called jets [Fey69, Bjo69, BP69, Sal10]. The measurements of jets have been used to explore the physics of the underlying hard-scattering processes in high-energy collisions [UA282, Axi82a, Axi82b, UA183, CER83, CDF89, CMS11b]. Collisions with momentum

balanced back-to-back jets (di-jets) [UA183, UA284, CDF90] have been observed. Such a pair of jets is interpreted to arise from a single hard scattering. High-p_{T} two-particle correlations in azimuthal angle and pseudorapidity have also been studied in order to analyze jet properties without the necessity of jet reconstruction [UA182a].

It has been proposed that the simultaneous interaction of two (or more) pairs of quarks manifests itself in multi-jet events with pair-wise p_{T}-balanced jets [LPS75, LP78, GHS80, Hum83, HO85, GHM86, Man89]. Accordingly, experimental evidences of multi-jet events are interpreted as due to multi-parton interaction.
First experimental evidence of double parton scattering in multi-jet events has been reported in the studies of 4-jet events in proton-proton collisions at $\sqrt{s} = 63\,\mathrm{GeV}$ [AFS87]. Here, pair-wise balanced 4-jet events have been observed. This was interpreted as being due to double-parton scatterings (DPS). The process-independent effective cross-section was estimated as $\sigma_{\mathrm{DPS,\,eff}} \approx 5\,\mathrm{mb}$. Even though, the kinematic features of these events could be described by DPS, standard QCD processes such as double bremsstrahlung (DB) could describe them, too.
Multi-jet events in $\mathrm{p\bar{p}}$ collisions have been studied at $\sqrt{s} = 630\,\mathrm{GeV}$ [UA291]. An upper limit of the DPS cross-section was estimated with an value of $\sigma_{\mathrm{DPS}} < 0.82\,\mathrm{nb}$ at 95 % confidence level (C.L.), which translates to an effective cross-section of $\sigma_{\mathrm{DPS,\,eff}} > 8.3\,\mathrm{mb}$ at 95 % C.L.
Double parton interactions in $\mathrm{p\bar{p}}$ collisions have also been studied at $\sqrt{s} = 1.8\,\mathrm{TeV}$ [CDF93, CDF97b, CDF97a]. The study was performed using both four jet events [CDF93, BF11] and events with three jets and one photon or multiple photons from neutral meson decay in jet fragmentation which approximately mimic a single photon [CDF97b, CDF97a]. A sizable DPS fraction has been measured in all approaches. The most precise value of the effective cross-section is $\sigma_{\mathrm{DPS,\,eff}} > 14.5 \pm 1.7^{+1.7}_{-2.3}\,\mathrm{mb}$ measured in $\mathrm{p\bar{p}}$ at $\sqrt{s} = 1.8$ TeV [CDF97a]. Hence, the double parton scattering accounts for approximately a quarter of the inelastic cross section measured at this collision energy.
Taking into account the different experimental setups, the double parton scattering cross-sections estimated in multi-jet events are in agreement with each other [CDF97a].

1.3.3. Growth of the Total Cross-Section

The total cross-section σ_{tot} of proton-proton collisions as a function of the center-of-mass energy is presented in Figure 1.8. The cross-sections have been estimated at fixed target experiments, at collider experiments, and, for the highest collision energies, using cosmic rays [Gia07, TOT11].

For center-of-mass energies above 100 GeV, the total cross-section of hadron-hadron collisions increases [TOT11]. The origin of this increase is not yet clear. Many QCD inspired

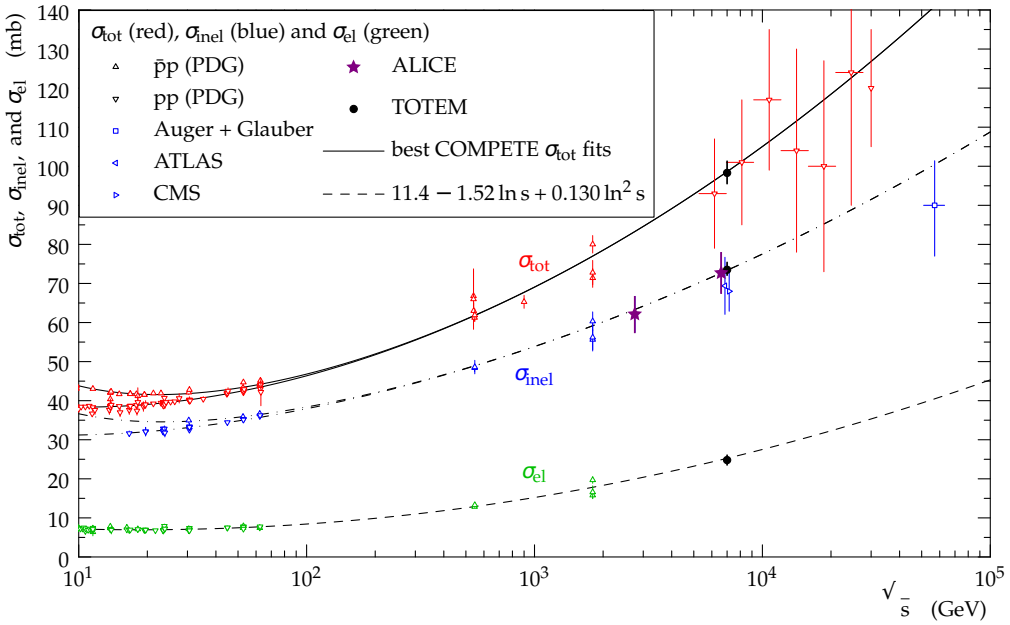

Figure 1.8.: Total cross-section as a function of the collision energy [TOT11].

models have incorporated the rise of the cross-section as due to an increase of the number of minijets and "semi-hard" parton interactions [PS86, Wal01, SS04] (cf. Section 1.4.1).
In the already introduced decomposition of the charged particle multiplicity distribution (cf. Section 1.3.1), cross-sections are derived for single, double, and triple parton collisions per proton-proton collisions [AAB+98, Wal01, Wal04]. These cross-sections are presented in the left panel of Figure 1.9. The triple parton cross-section σ_3 may contain also contributions from higher orders ($\sigma_3 = \sigma_{\geqq 3}$).
While the single parton cross-section σ_1 is almost independent from the center-of-mass energy, the double σ_2 and triple and higher order parton cross-section σ_3 increase with the energy as visible in the left panel of Figure 1.9. The double parton cross-section estimated in this decomposition is in good agreement with the DPS cross-section estimated in multi-jet events [CDF97a].
The right panel of Figure 1.9 shows the correlation between the multi-parton cross-section $\sigma_2 + \sigma_3$ and the non-single-diffractive cross-section σ_{NSD} [Wal04]. The multi-parton cross-section account for the increase of the total cross-section [AAB+98, Wal01, Wal04]. This is visible in the linear dependence between $\sigma_2 + \sigma_3$ and σ_{NSD}.

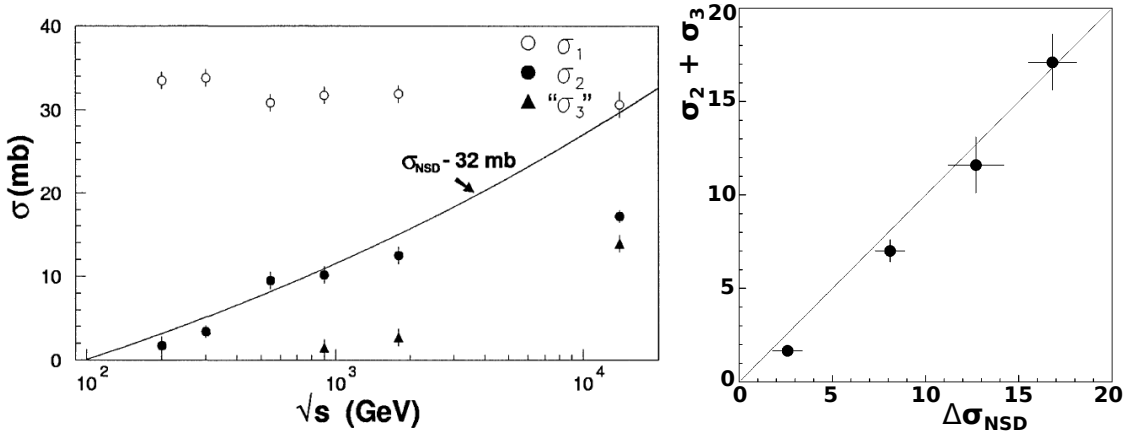

Figure 1.9.: Left panel: The cross-sections for single, double, triple parton interactions as a function of the center-of-mass energy [AAB+98, Wal01, Wal04]. The triple parton interactions may include also higher order interactions. Right panel: Increase of the multi-parton cross-section $\sigma_2 + \sigma_3$ against the increase of non single diffractive cross-section σ_{NSD} [Wal04].

1.3.4. Growth of the J/Ψ Yield

In order to test whether multiple parton interactions have an impact not only on soft processes related to the soft particle production which dominate the charged particle multiplicity but also on hard processes, it has been proposed to study the correlation between the production of heavy particles and the charged particle multiplicity. As one possible test, it has been proposed to study the correlation between the charmonium production and charged particle multiplicity [PGdC11].
The production of the J/Ψ mesons ($c\bar{c}$, $m_{J/\Psi} = 3.096\,\mathrm{GeV}/c^2$) as a function of the charged particle multiplicity has been studied in proton-proton collisions at $\sqrt{s} = 7.0\,\mathrm{TeV}$ [ALI12c]. The J/Ψ particles have been reconstructed using the invariant mass distribution of electron-position pairs measured in the central rapidity region ($|y| < 0.9$) as well as of muon pairs ($\mu^+\mu^-$) measured in the forward region ($2.5 < y < 4$). Both measurement show independently that the J/Ψ yield grows approximately linearly as a function of the charged particle multiplicity as shown in Figure 1.10.

A possible interpretation of these results is that the increase of the J/Ψ yield as a function of the charged particle multiplicity is due to multiple parton interactions. With increasing charged particle multiplicity which presumably results in the selection of events with a high number of multiple parton interactions, the yield of the J/Ψ increases [ALI12c].

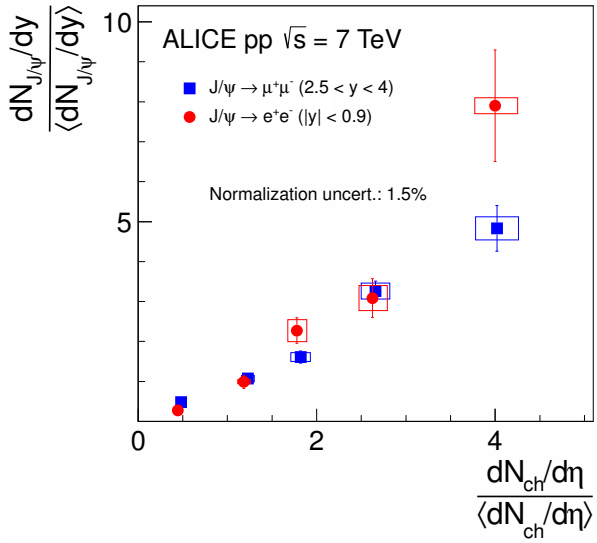

Figure 1.10.: J/Ψ yield $\mathrm{d}N_{J/\Psi}/\mathrm{d}y$ as a function of the charged particle multiplicity density at mid-rapidity $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$ measured at $\sqrt{s} = 7.0\,\mathrm{TeV}$. The figure shows the results at forward rapidities ($J/\Psi \to \mu^+\mu^-$, $2.5 < y < 4$) and at mid-rapidity ($J/\Psi \to e^+e^-$,$|y| < 0.9$). Both values are normalized to the corresponding value for minimum bias pp collisions [ALI12c].

1.4. Models for Multiple Parton Interactions

Several models give quantitative and qualitative descriptions of multiple parton interactions in pp collisions. The most common models are presented in the following.

1.4.1. pQCD Inspired Model

In pQCD inspired models like implemented in the Pythia event generator [SvZ87, SS04] (cf. Section 1.6.1), parton-parton interactions in high-energy collisions are assumed to be describable by pQCD. This assumption is valid for large transverse momentum values ($p_{\mathrm{T}} \geq$ few GeV/c). For the low-p_{T} region, however, a regularization to correct the emerging divergence in the cross-section needs to be introduced.
The differential perturbative QCD cross-section as a function of p_{T}^2 for a $2 \to 2$ parton interactions is given by [SS04]

$$\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}^2} = \sum_{i,j,k} \int \mathrm{d}x_1 \int \mathrm{d}x_2 \int \mathrm{d}\hat{t} \times f_A^i(x_1, Q^2) f_B^j(x_2, Q^2) \frac{\mathrm{d}\hat{\sigma}_{i,j}^k}{\mathrm{d}\hat{t}} \delta\left(p_{\mathrm{T}}^2 - \frac{\hat{t}\hat{u}}{\hat{s}}\right). \tag{1.9}$$

Here, $\hat{\sigma}_{i,j}^k$ is the cross-section for the hard scattering for the k'th sub-process between the incoming partons i and j. $\hat{s}$, $\hat{t}$ and $\hat{u}$ are the Mandelstam variables described in Appendix A. $f_A^i(x_1, Q^2)$ and $f_B^j(x_2, Q^2)$ are the parton distribution functions of the incoming hadrons

giving the probability to find a parton $i(j)$ with an energy fraction $x_1(x_2)$ of the incoming hadron $A(B)$ at an energy scale Q^2 with

$$Q^2 = p_{\mathrm{T}}^2 = \frac{\hat{t}\hat{u}}{\hat{s}}. \tag{1.10}$$

The integrated interaction cross-section above any $p_{\mathrm{T,\,min}}$ is written as [SS04]

$$\sigma_{\mathrm{int}}(p_{\mathrm{T,\,min}}) = \int_{p_{\mathrm{T,\,min}}}^{s/4} \frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}} \mathrm{d}p_{\mathrm{T}}. \tag{1.11}$$

There are two drawbacks to this approach. The first problem is that at low p_{T} low x are probed ($x \sim |p_{\mathrm{T}}|/\sqrt{s}$) while the PDFs rise at low x steeply (compare Figure 1.3 on page 9). Hence, the integrated cross-section

$$\int_{p_{\mathrm{T,\,min}}}^{s/4} \frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}} \mathrm{d}p_{\mathrm{T}} \propto \frac{1}{p_{\mathrm{T,\,min}}^2} \tag{1.12}$$

diverges at $p_{\mathrm{T,\,min}} \to 0$ [SS04].
This problem is avoided by introducing a cut-off parameter $p_{\mathrm{T,\,min}}$ which can be interpreted as inversely related to the color screening length of the hadron.
The second problem is that the integrated cross-section at $p_{\mathrm{T}} \lessapprox$ few GeV/c exceeds the total cross-section. This can be solved by introducing the concept of multiple parton interactions. Two or more independent hard parton-parton interactions can happen in the same pp collision. The number of parton interactions in this approach is estimated by

$$n_{\mathrm{parton-parton}}(p_{\mathrm{T,\,min}}) = \frac{\sigma_{\mathrm{int}}(p_{\mathrm{T,\,min}})}{\sigma_{\mathrm{non-diffractive}}}. \tag{1.13}$$

In a simple approach, the parton interactions take place independently from each other and can be described by a Poisson distribution [SvZ87]. In a more complex approach, the nonzero size of the hadron is taken into account. Here, the number of parton interactions depends strongly on the overlap of the hadrons quantified by the impact parameter b.

1.4.2. Clan Model

In the clan model, particle production in high-energy collisions is described in terms of cascades of particles in a two-step process. Mother particles (step 1) produced directly in the collisions cascade into daughter particles (step 2) by decay and fragmentation. All particles produced together with one mother are called “clan” or “cluster” [Eks85, GVH86, GVH88]. In the most simple model, clans are produced independently and the number of clans n_c follows a Poisson distribution

$$P_{\mathrm{clan}}(n_c, \mu_c) = \frac{\mu_c^n e^{-\mu_c}}{n_c!}. \tag{1.14}$$

Here, $\mu_c = \langle n_c \rangle$ is the average number of clans produced in a collision.
The number of final state particles n_p produced per clan is described by a probability distribution $F_p(n_p)$. The probability p to produce one extra particle per clan is assumed to be proportional to the current number of particles per cluster. It then can be derived that

$$F_{\mathrm{p}}(n_p) = F_p(1) \cdot \frac{p^{n_p-1}}{n_p}, \qquad n_p \geq 1. \tag{1.15}$$

Per definition, each clan contains at least one final state particle, hence $F_p(0) = 0$. The two distributions P_{clan} and F_p can be combined to describe the charged particle multiplicity distribution $P(n)$

$$P(n) = \sum_{n_c=1}^{n} P_{\mathrm{clan}}(n_c, \langle n_c \rangle) \sum^{*} F_p(n_{p,1}) \cdot F_p(n_{p,2}) ... F_p(n_{p,n_c}), \tag{1.16}$$

in which the second sum includes only those combinations of $n_{p,i}$ in which $n = \sum_{i=1}^{n_c} n_{p,i}$ is true [GVH86, GVH88, GU05].
The characteristic parameters of the clan model are the average number of clans $\langle n_c \rangle$ and average number of particles per clan $\langle n_p \rangle$. The description of charged particle multiplicity distributions in the clan model, can be translated into the charged particle multiplicity distribution of NBD distributions described in Section 1.3.1. The NBD parameters k and $\langle n \rangle$ and the clan parameters are related in the following way:

$$\langle n_c \rangle = k \cdot \ln\left(1 + \frac{\langle n \rangle}{k}\right), \tag{1.17}$$

$$\langle n_p \rangle = \frac{\langle n \rangle}{\langle n_c \rangle}. \tag{1.18}$$

For pp collisions measured at the LHC at center-of-mass energies of $\sqrt{s}$ = 0.9, 2.36, and 7.0 TeV [CMS11a], charged particle multiplicity distributions have been evaluated in terms of the clan model using NBD fits [DJD$^+$11, Gho12]. While in [DJD$^+$11] only a single-NBD fit was used, in [Gho12] two NBDs were used to perform the fit allowing to investigate the development of a "soft" and a "semihard" component of the charged particle multiplicity distribution as a function of the collision energy. The NBD parameter $1/k$ as well as the clan parameters, number of clans $\langle n_c \rangle$ and number of particles per clan $\langle n_p \rangle$, are shown for the soft and the semi-hard component in Figure 1.11.

The parameters of the "soft" component show almost no center-of-mass energy dependence while the parameters of the "semihard" components vary strongly as a function of $\sqrt{s}$. $1/k$ and $\langle n_p \rangle$ increase while $\langle n_c \rangle$ decreases with $\sqrt{s}$ for the "semihard" component. The observations are compatible with a KNO violation of the hard component and a KNO scaling of the soft component. This "soft" component is also interpreted as being due to collision events of single parton collision events only, the "semihard" component as due

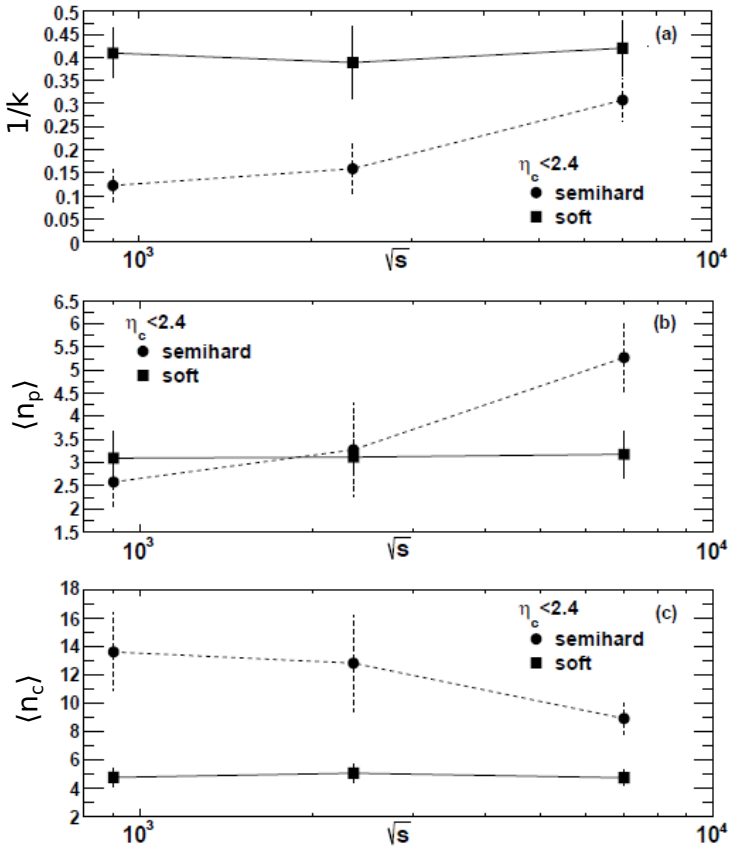

Figure 1.11.: Energy dependence of the NBD parameter $1/k$ (top panel) and the clan parameters, the avergae number of clans per collision $\langle n_c \rangle$ (mid panel) and the average number of final state particles per clan $\langle n_p \rangle$ (bottom panel), for the "soft" and the "semihard" component of the charged particle multiplicity distribution [Gho12]. The analyzed charged particle multiplicity distributions are measured including charged particles of a pseudorapidity range of $|\eta| < 2.4$ [CMS11a]. The lines in the plots are drawn to guide the eye.

multiple parton interactions [Wal01, Wal04].
It is discussed whether the clan concept can be related to real physical objects whose partonic partners might be QCD parton showers and thus give an insight into multiple parton interactions [GU05].

1.4.3. Independent Pair Parton Interaction Model

Within the framework of the independent pair parton interaction model (IPPI), high-energy hadron-hadron collisions are interpreted as a set of independent binary parton collisions [DN04, DN11]. Each parton-parton collision gives rise to a charged particle multiplicity distribution described by a negative binomial distribution $P^{\mathrm{NBD}}_{\langle n\rangle,k}(n)$ (cf. Section 1.3.1). The sum of the NBDs is expressed by

$$\begin{aligned} P_{m,k}(n) &= \sum_{j=1}^{j_{\max}} w_j P^{\mathrm{NBD}}_{m_j,k_j}(n) \\ &= \sum_{j=1}^{j_{\max}} w_j \sum_{(n_p)} \prod_{p=1}^{j} P^{\mathrm{NBD}}_{m,k}(n_p). \end{aligned} \tag{1.19}$$

Here, $P^{\mathrm{NBD}}_{m,k}(n)$ is the probability of creating n particles, $m = \langle n\rangle$ is the mean number of particles produced per parton interaction, $j_{\max}$ is the number of active parton pairs, w_j is the probability of the j'th pair to be active, and n_p is the number of particles produced by the p'th parton pair.

The charged particle multiplicity distributions measured in pp collisions at $\sqrt{s}$ = 0.9, 2.36 and 7.0 TeV [CMS11a] have been analyzed in terms of the IPPI framework [DN11]. For this purpose, the charged particle multiplicity distributions are fitted with Equation 1.19. Figure 1.12 shows the IPPI parameters $j_{\max}$ and m obtained as a function of the center-of-mass energy. The number of active parton pairs $j_{\max}$ is presented in the top panel of Figure 1.12, and the mean charged particle multiplicity per single parton interaction m is presented in the bottom panel. The data points are compared to results of the quark gluon string model (QGSM) described in the next section.

With increasing center-of-mass energy, the number of parton-parton interactions $j_{\max}$ estimated within the IPPI framework increases. Also, the mean number of particle produced in each single parton interaction m increases with the center-of-mass energy.

Within the IPPI framework, events with up to 6 parton-parton collisions are found to contribute to the charged particle multiplicity distribution measured in pp collisions at $\sqrt{s}$ = 7.0 TeV [DN11].

1.4.4. Quark Gluon String Model and the Dual Parton Model

In hadronic collision, soft interactions of low momentum transfers are dominant [KP10]. Hence, it is discussed, whether models completely based on soft processes are more suitable to describe hadronic collisions rather than pQCD motivated models.

Regge theory based on pomeron exchange phenomena has been shown to be able to describe soft processes in high-energy collisions [KP10]. Pomerons are associated with cylinder-type diagrams of large N topological expansions of QCD. At very high collision

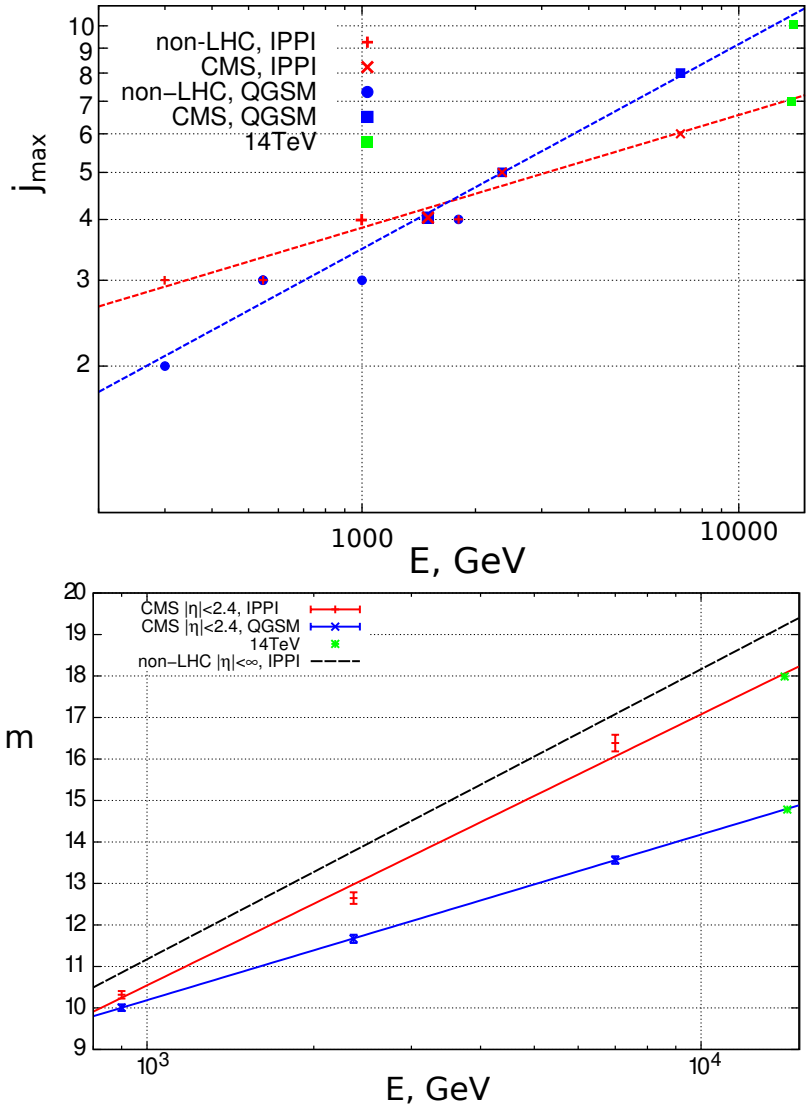

Figure 1.12.: Top panel: Number of parton interactions $j_{\max}$. Bottom panel: Mean charged particle multiplicity m generated in a single parton interaction [DN11]. Both, the results of the IPPI model (Section 1.4.3) and the QGSM model (Section 1.4.4) are presented. Extrapolations of both models for pp collisions at $\sqrt{s} = 14\,\text{TeV}$ have been performed.

energies, many terms of the expansion need to be taken into account. This corresponds to multi-pomerons exchanges [KP10].

The quark gluon string model (QGSM) [Kai82] and the dual parton model (DPM) [C$^+$94] are based on reggeon and pomeron exchanges and combine non-perturbative topology expansion of QCD with the theoretical concepts of duality [CPT75, CR76, CR78], Gribov's reggeon field theory [Gri68, GM69], and the parton structure of hadrons [Fey69, Bjo69].

In Figure 1.12, the IPPI model parameters discussed in the previous section are compared to results of the multi-pomeron exchange model of QGSM. Each pomeron exchange is

here related to single pair parton interactions. The mean charged particle multiplicity produced in a single pair parton interaction in the QGSM model is lower as compared to the IPPI model for all analyzed collision energies. The number of single pair parton interactions is lower at low center-of-mass energies and slightly larger at large center-of-mass energies as compared to the IPPI model [DN11]. Within the QGSM model, events with up to 8 parton-parton collisions are found to contribute to the charged particle multiplicity distribution measured in pp collisions at $\sqrt{s}$ = 7.0 TeV [DN11].
Even though the IPPI and QGSM show slight disagreements, both models agree in the qualitative conclusion that the number of parton interaction as well as the charged particle multiplicity per parton interaction increase with the center-of-mass energy [DN11].

1.5. High-Multiplicity Proton-Proton Collision

Proton-proton collisions of high charged particle multiplicities at $\sqrt{s}$ = 7.0 TeV show charged particle multiplicities very similar to peripheral heavy-ion (Cu-Cu) collision at $\sqrt{s_{\mathrm{NN}}}$ = 200 GeV [PHO11]. This observation motivates the question whether pp collision can also show collective effects as observed in heavy-ion collisions. It could be possible that a state called quark gluon plasma (QGP) is formed in these pp collisions of high charged particle multiplicity [WKP11]. The QGP is a state of nuclear matter which is assumed to have existed shortly after the big bang and which has been generated and studied in heavy-ion collisions [YHM05].
A possible indication of collective effects in pp collisions is the near side "ridge" phenomenon measured in two-particle angular correlations[1] of collision events with high charged particle multiplicities. Two measured two-particle correlations as a function of the difference in pseudorapidity $\Delta\eta$ and the azimuthal angle $\Delta\phi$[2] are presented in Figure 1.13 for both heavy-ion collisions at $\sqrt{s_{\mathrm{NN}}}$ = 200 GeV [STA10] and for high-multiplicity pp collisions at $\sqrt{s}$ = 7.0 TeV [CMS10].
Both figures show besides a near side peak around $\Delta\phi = 0$ and $\Delta\eta = 0$ an enhancement of the near side correlated yield over many units of $\Delta\eta$, referred to as ridge. For the heavy-ion collisions, the near side ridge has been interpreted as related to hydrodynamical evolution of matter [WKP11]. The ridge structure measured in pp collisions of high charged particle multiplicities could not be reproduced by common event generators as, for example, Pythia [SMS06, SMS08].

[1] The concept of two-particle angular correlations is described in detail in Chapter 5.
[2] In this thesis, the azimuthal angle is denoted as φ. The two presented analyses however use the symbol ϕ [STA09, CMS10].

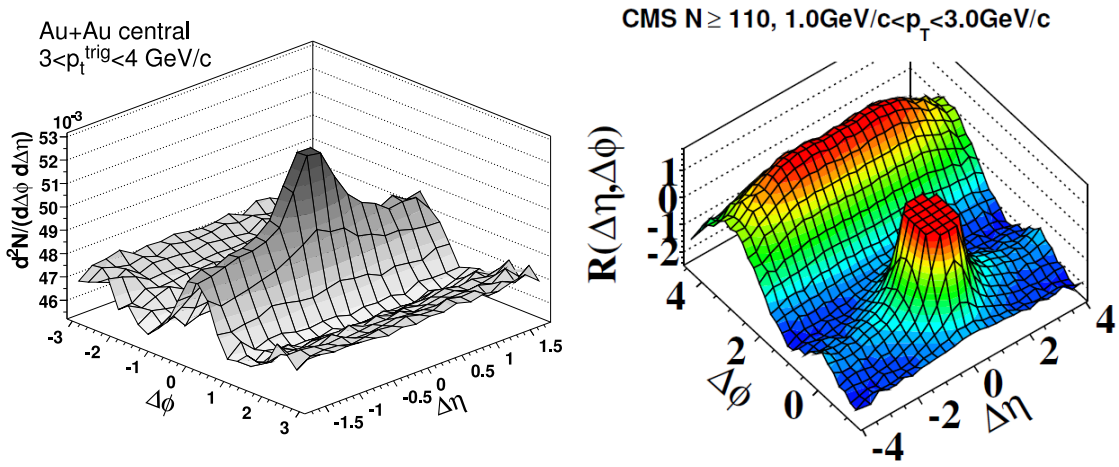

Figure 1.13.: Left panel: Charged di-hadron distribution for central gold-gold collisions at $\sqrt{s_{\mathrm{NN}}} = 200\,\mathrm{GeV}$, $3 < p_{\mathrm{T,\,trig}} < 4\,\mathrm{GeV}/c$ and $p_{\mathrm{T,\,assoc}} < p_{\mathrm{T,\,trig}}$ [STA09] Similar observations have been made in [PHO10a, PHO10b]. Right panel: Two-particle correlation function R for charged particle of $1 < p_{\mathrm{T}} < 3\,\mathrm{GeV}/c$ measured at $\sqrt{s} = 7.0\,\mathrm{TeV}$ in pp collisions of high charged particle multiplicity ($N_{\mathrm{offline\ tracks}} > 110$). The sharp near side peak from jet correlations is cut off in order to better illustrate the structure outside that region [CMS10]. Note that both figures have different ranges in $\Delta\phi$.

1.6. Simulation of Multiple Parton Interactions in Event Generators

The simulation of high-energy collisions can be performed using Monte Carlo (MC) event generators. Based on a (pseudo)random number generator, MC event generators simulate the production of particles according to the probability of the underlying physics models. Common MC generators used to simulate high-energy proton-proton collision are for example Pythia6 [SvZ87, SMS06], Pythia8 [SMS08], and Phojet [ERR95]. In the following, the physics models implemented in Pythia and Phojet are briefly described.

1.6.1. Pythia

Pythia [SMS06, SMS08] is a general purpose event generator for high-energy particle physics reactions. It is widely used for the simulation of high-energy collisions of various combinations of hadrons and leptons. Pythia combines pQCD for hard interactions with pQCD inspired models for the description of soft hadronic interactions which have been introduced in Section 1.4.1. Pythia includes models for initial and final-state parton showers, multiple parton-parton interactions, beam remnants, string fragmentation, and particle decays.

Two main Pythia versions are available Pythia6.4 [SMS06] and Pythia8.1 [SMS08].

Pythia8, written in C++, is the successor of Pythia6 (written in Fortran). The basic concepts of both Pythia versions are very similar, however, Pythia8 introduces some conceptional changes, e. g. the final state radiation has been combined with the initial state radiation and the multiple parton interactions [CS11]. Both Pythia versions incorporate various tunable parameters, e. g. the p_{T}-cut-off parameter introduced in Section 1.4.1. Sets of parameters are combined to predefined "tunes" which are optimized to reproduce specific measurements. In the following, the tunes Pythia6.4 Perugia-0 [Ska10] and Perugia-2011 [Ska10] as well as Pythia8 4C [CS11] are introduced. These tunes are used in Chapter 8 for the comparison to real collision data.

Pythia6.4 Tune Perugia-0

Both discussed Perugia tunes are based on the Pythia6 tune S0(A). The parameters of S0(A) have been tuned to a good description of fundamental minimum bias results obtained at hadron colliders of energies up to $\sqrt{s} = 1.96\,\mathrm{TeV}$. From this starting point, the hadron collider parameters have been re-tuned to describe results of more advanced data analyses. This includes the charged particle multiplicity distribution, the transverse momentum distribution, the correlation between the average transverse momentum and the charged particle multiplicity, the underlying event, and the forward-backward correlation strength [Ska10].

Pythia6.4 Tune Perugia-2011

In comparison to the Perugia-0 tune, the Perugia-2011 tune has been further tuned to early minimum-bias and underlying event data measured at the LHC at energies from $\sqrt{s} = 0.9\,\mathrm{TeV}$ to $7.0\,\mathrm{TeV}$. The modifications relevant for this thesis are a faster scaling of the charged particle multiplicities with collision energy and a slightly larger underlying event. Furthermore, a slightly softer fragmentation function has been implemented [Ska10].

Pythia8 Tune 4C

The 4C tune is the latest tune of Pythia8. The tune focuses on the description of early LHC data, e. g. charged particle multiplicity distribution, the transverse momentum distribution, the correlation between the average transverse momentum and the charged particle multiplicity, and the underlying event. For collision energies below LHC energies, Pythia8 4C produces in general too large particle yields [CS11].

1.6.2. Phojet

Phojet (version 1.12) is a minimum bias event generator for hadron pp, γp, and $\gamma\gamma$ interactions [Eng95, ERR95, LSB$^+$95]. For the simulation of soft processes, Phojet uses the Dual

Parton Model (DPM) and the Quark Gluon String model (QGSM) (cf. Section 1.4.4). The simulation of hard processes is based on pQCD. The soft and the hard part of the description are combined with the use of unitarity considerations [C$^+$94] leading to the possibility to have multiple soft and hard interactions in one collision event [LSB$^+$95]. During the event simulation, the main focus of Phojet is on the soft component of the hadron collision. Hard processes are added in a later step of the event simulation.
In addition to soft and hard interactions, Phojet also incorporates initial state radiations and final state radiation as Pythia. In contrast to Pythia, Phojet also comprises the simulation of central diffractive events based on double pomeron exchange.
Phojet has been used to successfully describe experimental data measured at collision energies up to $\sqrt{s}$ = 1.8 TeV. Unlike Pythia, Phojet has not been optimized for LHC data.

1.7. Summary

In this chapter, the concept of multiple parton interactions in high-energy proton-proton has been introduced. A review of signatures of multiple parton interactions which have been studied in the past has been given. This includes distributions of the charged particle multiplicity, multi-jet events, the increase of the total cross-section as a function of the collision energy, as well as the increase of the J/Ψ production as a function of the charged particle multiplicity. All observables indicate the existence of multiple parton interactions in pp collisions. Also, it has been observed that the fraction of pp collisions showing multiple parton interactions increases with increasing center-of-mass energy. The review of the signatures of multiple parton interactions has been completed by a review of theoretical models that describe the physics of multiple parton interactions. Monte Carlo generators based on theoretical model descriptions of the multiple parton interactions have been introduced. These models are used in the comparison to real collision data in Chapter 8.

2. The Large Hadron Collider

The Large Hadron Collider (LHC) [LHC95, BCL$^+$04a, BCL$^+$04b, BCM$^+$04, EB08] is a superconducting hadron accelerator and collider at the European Organization for Nuclear Research (CERN) in Geneva, Switzerland. The LHC is designed to perform proton-proton (pp) collisions as well as heavy-ion collisions, i. e. lead-lead (Pb–Pb) collisions. The heavy-ion program also foresees p–Pb collisions.
At the time of writing, the LHC provides the highest collision energy and the highest luminosity ever achieved by a hadron accelerator.

2.1. LHC Design

The LHC is a synchrotron for two hadron beams traveling in opposite directions. It re-uses the 26.7 km-long tunnel originally hosting the Large Electron Positron collider[1] (LEP) [EB08]. The tunnel is placed at a depth of 45 m to 170 m below the earth's surface which itself has a slope from the Jura mountains to the Lake Geneva. The plane of the accelerator ring is also inclined by 1.4% sloping towards Lake Geneva reflecting the earth topology and allowing for a connection to the CERN accelerator complex and a cost-saving construction. The tunnel has eight straight sections and eight arcs (Figure 2.1).
At four of the straight sections, the LHC experiments are located (cf. Section 2.4). The remaining straight sections are used for the beam injection, the acceleration of the beam and compensating for energy losses, the cleaning of the beam, and the beam dumping.
The acceleration of the proton bunches is performed by superconducting radio frequency (RF) cavities. Besides acceleration, the cavities keep the particle bunches tight. This ensures a high luminosity at the collision points [EB08]. In total, 8 cavities are used per beam direction. The cavities generate an accelerating field of up to 5 MV/m at a frequency of 400 MHz. They are operated at 4.5 K.
In total, the LHC requires 9600 magnets of different types (dipoles, quadrupoles, sextupoles, octupoles, decapoles, etc.) including 1232 main dipoles of each 14 m. Most of the main dipole magnets are placed within the 8 archs, however, there are also some dipoles used in the transfer tunnels connecting the LHC to the CERN accelerator complex. The dipole magnets are used to bend the beams on a circular trajectory. The operating temperature of these magnets is 1.9 K. The magnets that produced fields of higher order are used to focus and to de-focus the beams. Alternating focusing and defocussing results in an overall focused and stable beam. All magnets contribute to the optimization of the

[1] The Large Electron Positron collider was a particle accelerator built at CERN. It was in operation from 1989 to 2000. Its maximum center-of-mass energy was 209 GeV [LEP84].

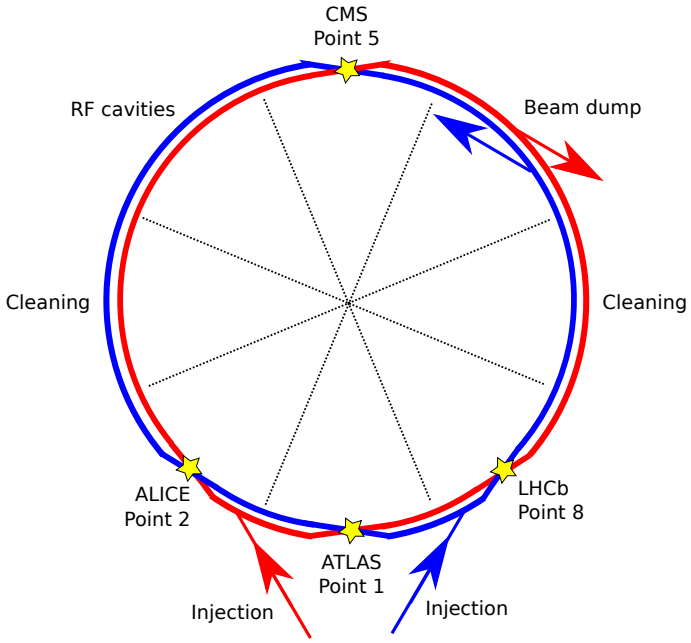

Figure 2.1.: Schematic layout of the Large Hadron Collider. The beam directions are marked in red for clockwise and blue for counter-clockwise. The positions of the four interaction points and the four major LHC experiments are marked by yellow stars. The LHC is subdivided into eight octants indicated by the black, dashed lines. The sectors are counted clockwise starting at the interaction point of ATLAS. In sector 3 and 7, the beam is cleaned. In sector 4, RF cavities accelerate the particle bunches or compensate for energy loss. In sector 6, the beam can be dumped in case of a system abort. The figure is derived from [EB08].

particle trajectories within the LHC [EB08].

Operated at design specifications, the LHC dipole magnets will produce a magnetic dipole field of 8.33 T. This will allow for an energy per proton of 7.0 TeV corresponding to pp collisions at a center-of-mass energy of $\sqrt{s} = 14.0\,\text{TeV}$. A fully accelerated lead ion will have an energy of 1150 TeV which corresponds to 2.76 TeV per nucleon. Hence, Pb–Pb collisions will have a center-of-mass energy per nucleon pair of $\sqrt{s_{\text{NN}}} = 5.52\,\text{TeV}$.

Besides the energy of the collision, the performance of the collider is also characterized by the luminosity $\mathcal{L}$. For an intersecting storage ring with two equal beams of the size $\sigma_{1,(x,y,s)} = \sigma_{2,(x,y,s)}$, the luminosity can be calculated as

$$\mathcal{L} = f \cdot n_b \cdot \frac{N_2 \cdot N_1}{4\pi\sigma_x\sigma_y}. \tag{2.1}$$

with the orbital frequency f, the number of bunches in one beam of the storage ring n_b, the number of particles per bunch for beam 1 N_1 and beam 2 N_2, and the cross section of the beams $4\pi\sigma_x\sigma_y$. At the design performance, the LHC stores 2808 bunches of protons per beam direction, each comprising 2×10^{14} protons. For lead-lead collisions, it is 592

ion bunches per beam direction, each comprising 5×10^7 lead-ions. The luminosity design goals of the LHC are $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ for pp collisions and $\mathcal{L} = 10^{27}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ for Pb–Pb collisions. At high luminosities, there is a non-negligible probability of more than one collision per bunch crossing. This effect is called pileup.

2.2. CERN Accelerator Complex

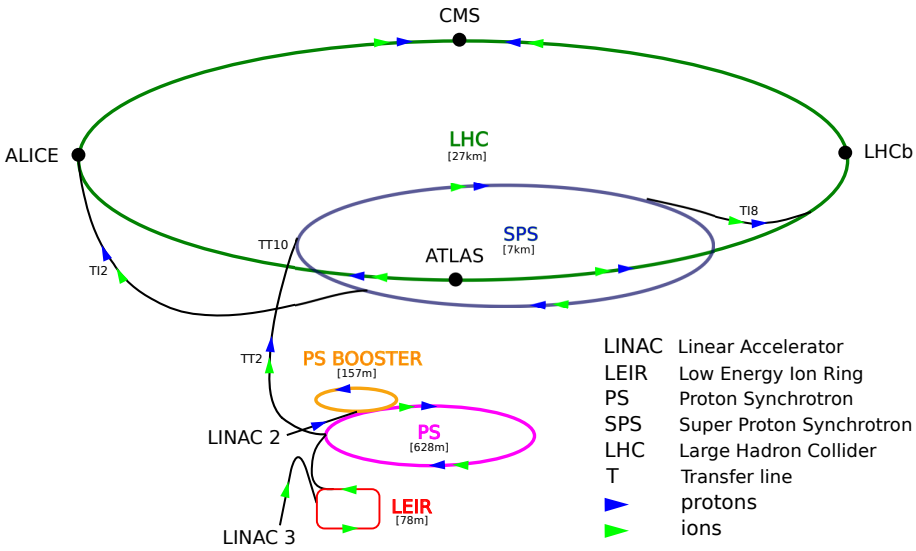

Figure 2.2.: The LHC injection complex. Note that CERN has further accelerators which are not shown is this figure. For each accelerator, its circumference is quoted. The figure is based on [EB08].

Before protons or lead-ions enter the LHC and before they reach their final energy, they pass through a chain of smaller accelerators of the CERN accelerator complex [BBB+00] shown in Figure 2.2.

Protons are produced by stripping off electrons from hydrogen atoms. These protons enter the Linear Accelerator LINAC2. After being accelerated to an energy of 50 MeV, they are injected into the PS Booster (PSB). Here they reach 1.4 GeV. In the following, the Proton Synchrotron (PS) and the subsequent Super Proton Synchrotron (SPS) accelerate the protons to 25 GeV and then to 450 GeV. From the SPS, the proton bunches are injected into the LHC where they can be accelerated to their final energy of 7.0 TeV per proton (at the time of writing, up to 4.0 TeV per proton).

Lead ions are generated by stripping off all 82 electrons from a lead atom [ABB+93]: To do so, a piece of lead (99.57% pure ^{208}Pb) is heated in a micro oven to about 500 °C. This allows a small number of lead atoms to vaporize. By mixing this vapor with oxygen and ionizing it with an electron current, differently charged states are produced. Out of the spectrum of states, a magnetic spectrometer selects Pb^{29+} ions. The Pb^{29+} ions are

injected into the linear accelerator LINAC3 [HKSV03]. After leaving LINAC3, the ions pass a carbon foil which strips off most of the remaining electrons. Out of these ions, another spectrometer selects Pb^{54+} ions. The Pb^{54+} ions are then filled into the Low Energy Ion Ring (LEIR) [Cha04] where they are cooled by electron cooling [Bud67] and accumulated to bunches. The resulting ion bunches are then accelerated in the PS. After a final stripping foil, the fully ionized Pb^{82+} are selected and injected into the SPS. From here, they finally reach the LHC. In the LHC, the lead-ions can be accelerated to maximal energy per nucleon of 2.76 TeV (at the time of writing, up to 1.38 TeV per nucleon).

2.3. LHC History, Status, and Outlook

The LHC project was approved by the CERN Council in 1994 [EB08] after an early planning and discussion stage reaching back to 1984. The construction works started in 2001. In June 2008, the first proton bunch injection tests were performed. The first circulation of the proton beams within the LHC were achieved on September 10, 2008. Unfortunately, nine days later, a magnet quench incident caused by a faulty electrical connection between two LHC superconducting magnets occurred. It was followed by a helium gas explosion in which several superconducting magnets and their mountings were damaged.
After more than one year of repair works including the installation of an upgraded protection system, proton bunches were circulating again in the LHC on November 20, 2009. Only three days later, all LHC experiments recorded first proton-proton collisions. The left panel of Figure 2.3 shows the first proton-proton collision candidate shown by the event display in the ALICE control room on September 23, 2009 recorded at a center-of-mass energy of $\sqrt{s} = 0.9$ TeV. The right panel of Figure 2.3 shows one of the first Pb–Pb collisions recorded by ALICE (first collision on November 7, 2010) at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV.

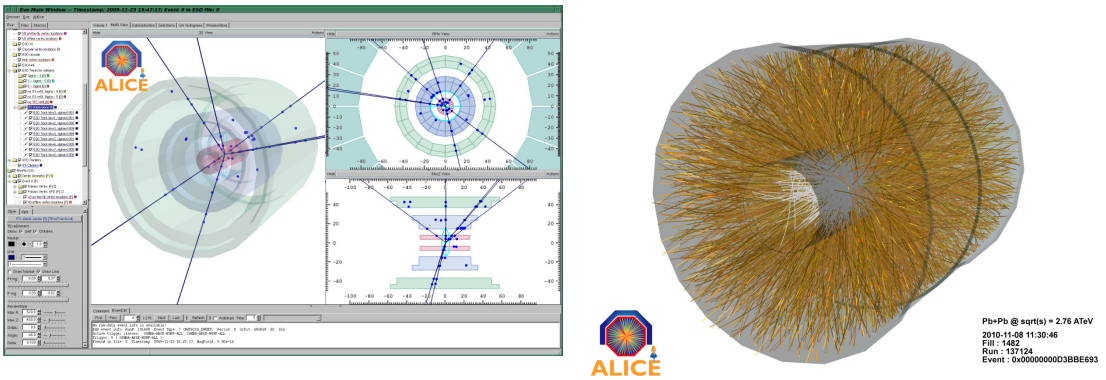

Figure 2.3.: Left panel: The first proton-proton collision candidate shown by the event display in the ALICE control room on September 23, 2009 [ALI10a]. Right panel: One of the first collisions of lead ions recorded by the ALICE detector on November 8, 2010.

At the time of writing, the LHC has performed proton-proton collisions at center-of-mass energies of $\sqrt{s} = 0.9$, 2.36, 2.76, 7.0 and 8.0 TeV and lead–lead collisions at the center-of-mass energy per nucleon pair of $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV.
The current center-of-mass energy reached by the LHC in pp collisions is 4 times higher compared to the center-of-mass energy reached by the former most powerful accelerator, the Tevatron,[2] [Wil77, TeV82] a proton-antiproton accelerator and collider with $\sqrt{s} = 1.96$ TeV.
For Pb–Pb collisions, the increase in collision energy is even higher: the RHIC[3] [HLO03] accelerator and collider is performing heavy-ion collisions at $\sqrt{s_{\mathrm{NN}}} = 0.2$ TeV, which is almost 14 times less than so far reached by the LHC in Pb–Pb collisions.
Figure 2.4 presents the total integrated luminosity recorded by the 4 LHC experiments during stable beams for 7.0 TeV center-of-mass energy in 2010. Due to a different detector configuration optimized for heavy-ion collisions, ALICE has asked for a lower luminosity and lower pileup compared to the other experiments. The left panel of Figure 2.4 shows a slower rise in integrated luminosity delivered to ALICE since July 2010 compared to CMS, ATLAS, and LHCb.
In December of 2010 and 2011, heavy-ion collisions were performed (right panel of Figure 2.4). Here, only ALICE, ATLAS, and CMS were recording data.
For December 2012, a p–Pb run is scheduled. As of 2013, a long technical shutdown is planned meanwhile the accelerator will be upgraded. This will allow for an increase of the magnetic field in the dipole magnets and hence, an increase in collision energy to the LHC design values of $\sqrt{s} = 14.0$ TeV for pp collisions and $\sqrt{s_{\mathrm{NN}}} = 5.52$ TeV for Pb–Pb collisions.

2.4. Experiments at the LHC

There are seven experiments at the LHC, six of which are recording collision data and one of which is still under construction.

ALICE (A Large Ion Collider Experiment) [ALI08] is the dedicated heavy-ion experiment at the LHC built for the study of the quark gluon plasma (QGP) generated in heavy-ion collisions. ALICE is also studying proton-proton collisions (cf. Chapter 3).

[2]The Tevatron is a particle accelerator and collider for protons and antiprotons at the Fermi National Accelerator Laboratory (Fermilab), close to Batavia, Illinois, US. Its name is derived from its center-of-mass energy range: the Tera Electron Volt. It began operation in 1987. It was decommissioned in 2011.
[3]The Relativistic Heavy Ion Collider (RHIC) is a heavy-ion collider and a spin-polarized proton collider. RHIC is located at the Brookhaven National Laboratory (BNL) in Upton, New York, US. It is in operation since 2000.

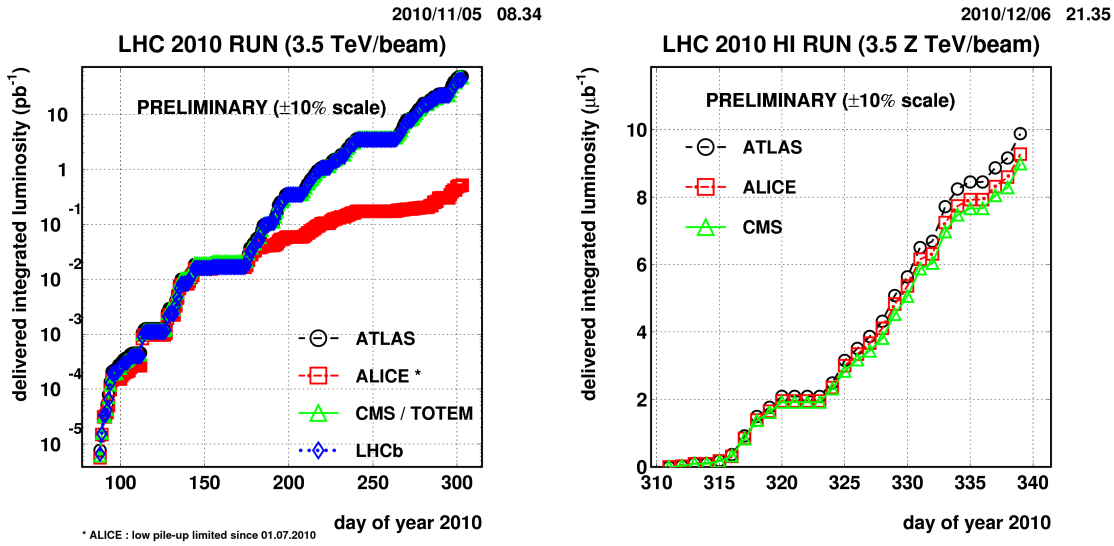

Figure 2.4.: Total integrated luminosity recorded by the LHC experiments in 2010 [GMFL11]. Left panel: Proton-Proton collisions at $\sqrt{s} = 7.0\,\text{TeV}$. Right panel: Lead-lead collisions at $\sqrt{s_{\text{NN}}} = 2.76\,\text{TeV}$.

ATLAS (A Toroidal Lhc ApparatuS) [ATL08] and CMS (Compact Muon Solenoid) [CMS08] are two general-purpose experiments designed to cover a broad spectrum of physics topics. The main goals of ATLAS and CMS are the search for the Higgs boson [Hig64, EB64, GHK64], extra dimensions, and particles that constitute dark matter such as particles predicted by the super-symmetric extension of the standard model (SUSY) [Mar97]. Heavy-ion collisions and the study of the QGP are also in the scope of the two experiments setups. While ATLAS and CMS have a similar research program, they use different technical setups and designs in order to complement each other and to give the possibility of cross checks and reassurance in case of discoveries.

LHCb (Large Hadron Collider beauty experiment) [LHC08a] looks into charm and beauty particle decays for possible CP violating processes in order to explain the large matter to anti-matter asymmetry observed in the universe.

LHCf (Large Hadron Collider forward) [LHC08b] studies particles generated in proton-proton collisions that have trajectories almost in line with the beam axis. Looking into the particle production in this forward direction might help to understand the origin of ultra-high-energy cosmic rays. LHCf consists out of two detector parts installed at a distance of 140 m at either sides of the ATLAS experiment.

TOTEM (TOTal Elastic and diffractive cross section Measurement) [TOT08] studies, like LHCf, the particles which are produced in the forward direction of the proton-proton collision. TOTEM's aim is the measurement of the total proton-proton cross section, the elastic scattering and diffractive dissociation at the LHC. Furthermore, it monitors the LHC's luminosity with high accuracy. TOTEM consists of two sub-detectors placed at either sides of the CMS experiment close to the CMS end-caps.

MoEDAL (Monopole and Exotics Detector at the LHC) [MoE09] was approved in 2009 as the seventh experiment at the LHC. The experiment is planned to be installed during the long shutdown of the LHC starting at the end of 2012. It will be placed close to the beam interaction point of the LHCb experiment. Its aim is the direct search for magnetic monopoles or dyons (particles with both electric and magnetic charges), and other exotic, highly ionizing stable and pseudo-stable massive particles produced at the LHC. MoEDAL will consist of an array of approximately 400 nuclear track detectors (NTD). Each NTD consisting of a 10-layer stack of plastic that will show damages if highly ionizing particles have crossed it.

3. The ALICE Detector

ALICE (A Large Ion Collider Experiment [ALI08]) is one of the seven experiments at the Large Hadron Collider. ALICE is designed to study heavy-ion collisions but it also has a rich proton-proton (pp) program. ALICE's heavy-ion physics program includes collisions of lead ions (Pb–Pb), asymmetric proton–ions, and possibly deuteron–ion collisions.
Since the start-up of the LHC in November 2009, ALICE has recorded data of pp collisions at the center-of-mass energies of $\sqrt{s} = 0.9, 2.36, 2.76, 7.0$, and 8.0 TeV and Pb–Pb collisions at the center-of-mass energy per nucleon pair of $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV.

3.1. Overview

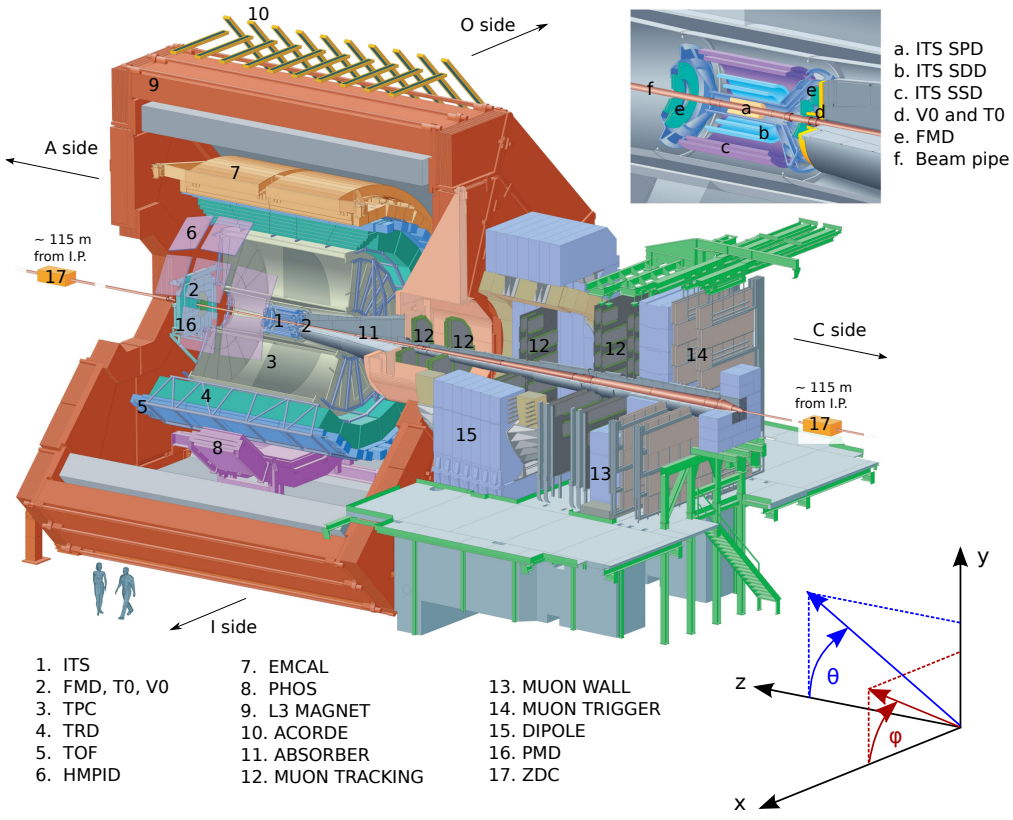

Figure 3.1.: Schematic view of the ALICE detector [ALI08]. On the upper right side of the Figure, a detailed view of the innermost part of the central barrel is shown. The global coordinate system and the nomenclature of the ALICE coordinate system (cf. Appendix B) is indicated.

The ALICE experiment is located in St. Genis-Pouilly, France, at the so-called Point 2 of the LHC (cf. Figure 2.1 on page 30). Figure 3.1 shows a schematic view of the ALICE detector.

ALICE is build around one of the four interaction points of the two hadron beams and their beam axis. The interaction point at Point 2 is located at a depth of 44 m underground. It is placed in the center of the experiment and surrounded by an 800μm-thick beryllium beam pipe. The outer diameter of the pipe is 6 cm.

In total, ALICE is 16 m each in height and width (x and y direction in Figure 3.1). ALICE's total length along the beam axis (z-direction) is 26 m. The experiment weighs approximately 10,000 t.

ALICE is composed out of the so-called central barrel, as well as several detectors in forward direction which measure particle multiplicities of particles emitted in the forward direction, and a forward muon spectrometer on one side of the central barrel. The detector components are described in detail in [ALI08].

The central barrel detectors are contained within a solenoid magnet, the L3 magnet, which was originally built for the L3 experiment [L3 90] at the Large Electron Position collider (LEP) [LEP84]. The L3 magnet consists of an octagonal steel yoke with an aluminum coil and can be closed by pole cap doors. The magnet is operated at room temperature. With an operating current of 30 kA, it has a nominal flux density of 0.50 T with field variations of only up to 2 % in the inside detector volume.

The detectors in the central barrel are designed in a cylindrical structure around the beryllium beam pipe: The innermost six layers build the so-called Inner Tracking System (ITS) [ALI99b]. It comprises three different types of silicon detectors: The Silicon Pixel Detector (SPD), the Silicon Drift Detector (SDD), and the Silicon Strip Detector (SSD).

Surrounding the ITS detector at increasing radii with respect to the beam pipe, the Time Projection Chamber (TPC) [ALI00], the Transition Radiation Detector (TRD) [ALI01] and the Time of Flight detector (TOF) [ALI02] are located. In addition to these detectors, which all have full azimuthal coverage, the central barrel hosts three additional detectors with limited azimuthal coverage: The High Momentum Particle Identification Detector (HMPID) [ALI98], which is a ring imaging Cherenkov detector, and two electromagnetic calorimeters, the Electro-Magnetic Calorimeter (EMCAL) [ALI10b] and the Photon Spectrometer (PHOS) [ALI99c]. The EMCAL is a Pb-scintillator sampling calorimeter. PHOS is built using high density scintillating lead tungstate crystals.

The ALICE Cosmic Ray Detector (ACORDE) is placed on top of the L3 magnet. It is an array of plastic scintillator counters, which provides a trigger signal for the tracking detectors in case of a cosmic shower.

The forward detectors include the Forward Multiplicity Detector (FMD), the VZERO (multiplicity and vertex) detector and the T0 (time and vertex) detector [ALI04], the Zero De-

gree Calorimeter (ZDC) [ALI99d] and the Photon Multiplicity Detector (PMD) [ALI99e]. The forward muon spectrometer comprises a particle absorber that allows only for muons to traverse it, a dipole magnet, trigger chambers, and tracking chambers [ALI99a].
The combination of all sub-detectors of ALICE allows to measure charged and neutral particles in a wide momentum range as well as to identify the particle species.
In the following sections, the detectors that have been used in the presented data analysis, VZERO, ITS, and TPC, are described in more detail. The VZERO detector as well as the ITS detector are used for the minimum bias trigger. The ITS and the TPC detector are used for the measurement of trajectories of charged particles.

3.2. VZERO Detector

The VZERO scintillator hodoscopes [ALI04] consists of two arrays of scintillator counters. One of these arrays (called VZERO-C) is placed in front of the muon absorber, 90 cm from the nominal interaction point along the beam axis, and the second one (called VZERO-A) is placed at the opposite side at 340 cm from the nominal interaction point. Both parts of the VZERO are partitioned in 32 counters arranged in two rings each. The VZERO-A covers the pseudorapidity range of $2.8 < \eta < 5.1$ and VZERO-C covers $-1.7 > \eta > -3.7$. Figure 3.2 shows an online event display of the Inner Tracking System (cylinders) and the two VZERO detector arrays (in forward direction) indicating the position of the sensitive areas of the two detectors.

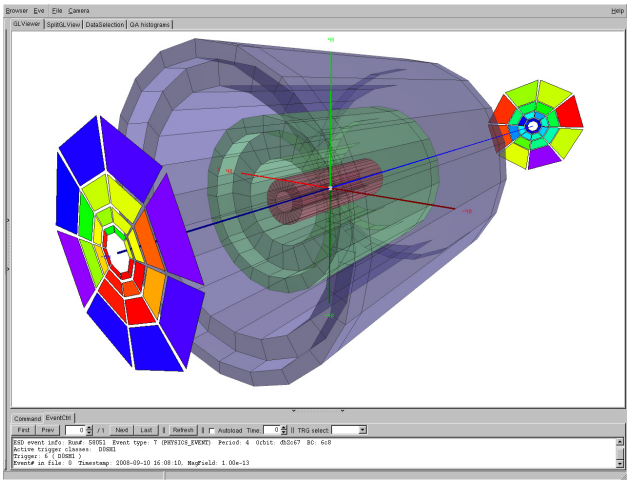

Figure 3.2.: An on-line event display showing the Inner Tracking System (cylinders) and the VZERO detector (in forward direction) while the first LHC beams arrive on 10 September 2008 [ALI12d].

The VZERO detector provides different kinds of triggers: it contributes to the minimum bias trigger for the central barrel detectors. It is also used to validate the di-muon trigger

during proton-proton collisions. During heavy-ion collisions, the VZERO detector allows for centrality resolving trigger decisions. In addition, the VZERO detector distinguishes between beam-beam collisions and beam-gas collisions and it is used as an luminosity monitor.

3.3. Inner Tracking System

The Inner Tracking System (ITS) [ALI99b] is built out of three different silicon detectors, each of which contributes with two detector layers. The six layers are arranged around the beam pipe at radii between 3.9 cm and 43.0 cm with respect to the nominal beam axis. The ITS measures track properties of charged particles with transverse momenta down to 35 MeV/c and it is crucial for the reconstruction of the primary vertex positions. Furthermore, the first two detector layers contribute to the minimum bias trigger and the tracklet reconstruction (cf. Section 3.3.1). The subsequent four detector layers contribute to the particle identification.
Altogether, the six layers cover roughly the pseudorapidity region of $|\eta| < 0.9$ for tracks which come from collisions that are placed within $\pm 1\sigma_z$ of the beam-beam interaction diamond with respect to the nominal interaction point in the middle of the experiment. The ITS detector has full azimuthal coverage[1].

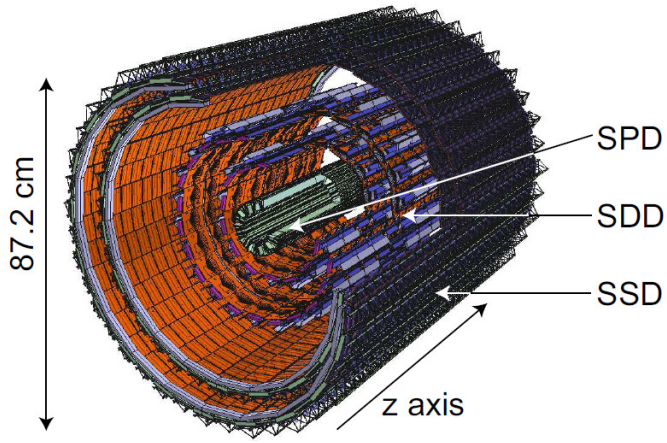

Figure 3.3.: Schematic view of the Inner Tracking System surrounding the beam pipe [GO09].

3.3.1. Silicon Pixel Detector

The Silicon Pixel Detector (SPD) constitutes the two innermost layers of the ALICE central barrel detectors. It is composed out of two cylindrical detector layers placed at

[1] Over time, some SPD half-staves at $\varphi = 2$, 4, 5, and 6 rad. needed to be switched off due to damage of the modules. This results in a reduced ITS reconstruction efficiency in these φ-regions.

radii of $r_{\mathrm{SPD1}} = 3.9\,\mathrm{cm}$ and $r_{\mathrm{SPD2}} = 7.6\,\mathrm{cm}$ from the nominal beam axis. Both SPD layers have a length of 33 cm in beam direction (z-direction).
The SPD is based on hybrid silicon pixels consisting of a 2-dimensional matrix of reversed-biased silicon detector diodes. In total, the SPD has 15.7 million read-out channels. The detector combined with its read-out electronics chips on top accounts for only $250\,\mu\mathrm{m}$ thickness per SPD layer.
The SPD covers a pseudorapidity range of $|\eta| < 1.98$ with its first layer and $|\eta| < 1.4$ with its second layer for collisions at the nominal interaction point. The effective η-acceptance is larger due to the longitudinally spread beam-beam interaction region.
The SPD contributes to the minimum bias trigger of ALICE. To this end, it generates a pulse when one or more of the silicon pixels are hit by a charged particle.
Out of the SPD hits, so-called tracklets are built in the data reconstruction. Tracklets are the most basic type of tracks measured with the ALICE detector.
The SPD is crucial in the reconstruction of the primary collision vertex and the secondary vertex of weak decays of strange, charm, or beauty particles.

3.3.2. Silicon Drift Detector

The Silicon Drift Detector (SDD) constitutes the two intermediate layers of the ITS detector. The layers are placed at radii of $r_{\mathrm{SDD1}} = 14.9\,\mathrm{cm}$ and $r_{\mathrm{SDD2}} = 23.8\,\mathrm{cm}$ on average. The first SDD layer has a length of 44.4 cm in z-direction, the second layer has a length of 59.4 cm.
The SDD consists of homogeneous high-resistivity $300\,\mu\mathrm{m}$-thick n-type silicon wafer. With its 133,000 channels, the SDD is able to resolve tracks also at high multiplicities. The SDD contributes not only to the tracking of charged particles but also to the measurement of the energy loss $\mathrm{d}E/\mathrm{d}x$ of the traversing particles allowing for a particle identification even for particles with low momenta.

3.3.3. Silicon Strip Detector

The Silicon Strip Detector (SSD) constitutes the two outer layers of the ITS detector. The layers are placed at radii of $r_{\mathrm{SSD1}} = 39.1\,\mathrm{cm}$ and $r_{\mathrm{SSD2}} = 43.6\,\mathrm{cm}$ on average. The first SSD layer has a length of 90.2 cm in z-direction, the second layer has a length of 101.6 cm. The two layers are built up of double sided Silicon Strip Detectors mounted on lightweight carbon-fiber support structures. The SSD has 2.7 million read-out channels.
The SSD is optimized for the matching of tracks between the ITS and the Time Projection Chamber (Section 3.4). Therefore, the SSD provides two-dimensional information of the track position. Just as the SDD, the SSD provides energy loss information used in the particle identification.

3.4. Time Projection Chamber

The Time Projection Chamber (TPC) [ALI00] is the most important tracking detector in the ALICE central barrel. Figure 3.4 shows a schematic view of the TPC.

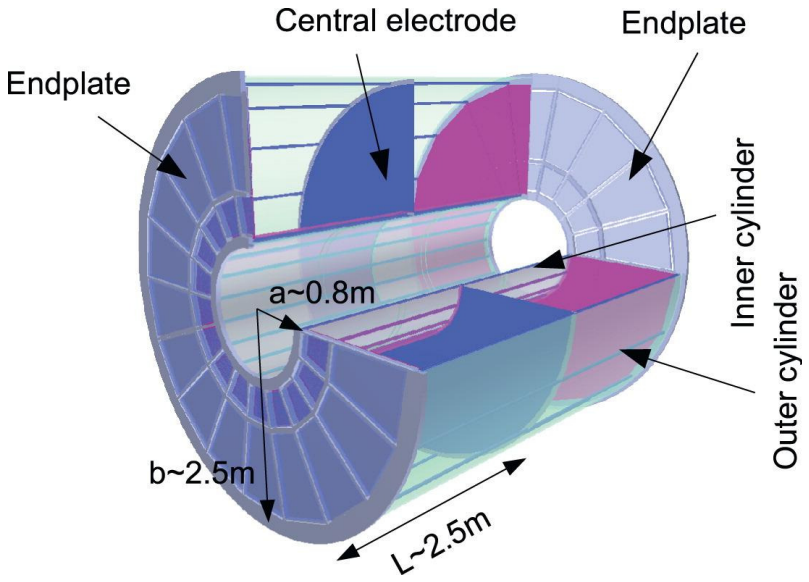

Figure 3.4.: Schematic representation of the ALICE Time Projection Chamber [RSR11].

In co-operation with the other central barrel detectors, it is used to perform charged-particle momentum measurements with a two-track separation that is good enough to cope also extreme particle densities. The TPC tracks can be used to determine the collision vertex. Furthermore, the TPC contributes to the particle identification. This is done based on energy loss measurements.

The TPC surrounds the ITS. It is located at radii of $r = 85\,\text{cm}$ to $r = 250\,\text{cm}$ with respect to the nominal beam axis. The TPC has an overall length of 500 cm in beam direction.

The TPC is a cylindrical gas detector with a cylindrical field cage. High-voltage is applied between the central membrane (central electrode) and the outer detector walls (end-plates). The gas volume is filled with $90\,\text{m}^3$ of $Ne/CO_2/N_2$ at a mixing ratio of (90/10/5). This mixture has been optimized for the drift velocity, low electron diffusion, and a low radiation length.

The end plates host multi-wire proportional chambers (MWPC) with fine grained cathode-pad read-out chips mounted into 18 trapezoidal sectors. Each sector covers 20° of the full azimuthal acceptance. Along the radius r, the sectors are separated into 2 parts. The parts are called the inner and the outer read-out chambers (IROC and OROC).

The TPC acceptance in terms of pseudorapidity is $|\eta| < 0.9$ for tracks with full radial track length and up to $|\eta| < 1.5$ for particles with reduced track length. The TPC has a full azimuthal coverage except for small dead zones (2°) between the 18 TPC sectors,

where very straight tracks of high-momentum particles can be lost. The TPC is designed to measure tracks of charged particles with transverse momenta from 0.1 GeV/c up to 100 GeV/c with a good transverse momentum resolution (cf. Figure 3.5).

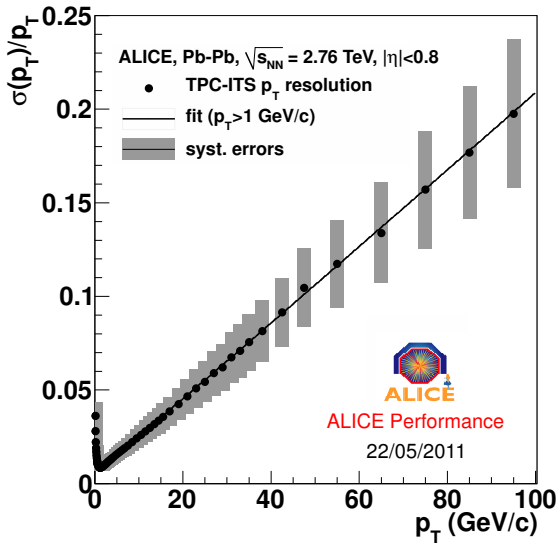

Figure 3.5.: Transverse momentum resolution for ALICE TPC+ITS combined tracking estimated in Pb–Pb collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76$ TeV.

3.5. The ALICE Computing Tools

ALICE uses a single offline software framework called AliRoot [BBC+03] for simulation, reconstruction, detector calibration and alignment, visualization, and data analysis. The framework is itself based on the widely used data analysis framework ROOT [BR97]. The AliRoot framework was commenced in 1998 and is under constant development.

AliRoot includes uniform interfaces to various event generators and transport Monte Carlo programs. As such, it allows for the seamless integration and the comparison of the event generators for proton-proton and nucleus-nucleus collisions, e. g. Pythia6 [SMS06], Pythia8 [SMS08], Phojet [ERR95], Therminator [KTBF06], Epos [WLP05], Herwig [CKM+02], Hijing [WG92], and Dpmjet [RER00], as well as the transport Monte Carlo programs Geant3 [BBM+87], Geant4 [Gea03, Gea06], and Fluka [FFRS93].

AliRoot is complemented by AliEn (AliCE Environment) [SAB+03]. AliEn manages the distributed computing infrastructure of ALICE. It follows the "grid paradigm" [FKT01, Fos02, BBB+05] in order to provide a uniform layer for data storage and data processing on top of the heterogeneous and loosely coupled computing resources of the university groups and research institutes that participate in ALICE.

4. Data, Event, and Track Selection

In this chapter, the data, event, and track selections used in the data analysis are described. Section 4.1 specifies the proton-proton collision data sets and Section 4.2 specifies the simulated data sets. Sections 4.3 and 4.4 describe the event and track selections applied in the data analysis.
In the final Section 4.5, a quality assessment method for collision data is introduced which has been necessary for the selection decision of the data sets used in this thesis. In turn, the assessment method has become part of the ALICE wide data quality assurance framework.

4.1. ALICE Proton-Proton Collision Data Selection

The data analysis is performed based on three proton-proton collision data samples recorded at the center-of-mass energies of $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV. The selection of the data sets is based on the quality assessment performed in the ALICE wide quality assurance framework which estimates the quality of all recorded data sets. The ALICE quality assurance framework has been extended in the course of this thesis as described in Section 4.5. The central quality assessment is based on the following conditions:
The data sets need to be identified as physics data set. This excludes, for instance, data sets which have been recorded during luminosity scans of the LHC. Furthermore, the quality of the calibration and reconstruction needs to be classified as good. For this purpose, the calibration and the reconstruction of the data sets are repeated until no more signs of mis-calibration or malfunctioning of the reconstruction are found in the properties of the reconstructed events and tracks. This usually takes between two and three iterations.
As the data analysis of this thesis is based on tracks measured by the two inner tracking detectors, the Inner Tracking System (ITS, cf. Section 3.3) and the Time Projection Chamber (TPC, cf. Section 3.4), the performance of the ITS and the TPC has to be evaluated, too.
Based on the afore mentioned criteria, the following data sets have been selected.

4.1.1. Proton-Proton Collision Data at $\sqrt{s} = 0.9$ TeV

At $\sqrt{s} = 0.9$ TeV, ALICE has recorded 7 million minimum bias (cf. Section 4.3.1) proton-proton collision events of good quality. This corresponds to 87 % of all recorded minimum bias data at this collision energy. The data has been collected in May 2010 and belongs to the ALICE data taking period LHC10c. The corresponding run numbers are listed in Table C.1 of Appendix C. The calibration and reconstruction of the data has been iterated

three times.
During data taking, the magnetic field of the L3 magnet (cf. Section 3.1) has been set to $B = +5\,\mathrm{kG}$. The mean number of interactions per bunch crossings denoted as μ in these runs varies between low values of $\mu = 0.0133$ and 0.0319. Hence, the number of pileup events in this data sample is negligible.

4.1.2. Proton-Proton Collision Data at $\sqrt{s} = 2.76\,\mathrm{TeV}$

At $\sqrt{s} = 2.76\,\mathrm{TeV}$, ALICE has recorded 34 million minimum bias proton-proton collision events of good quality. This corresponds to 86 % of the recorded minimum bias data at this collision energy. The data has been collected in March 2011 and belongs to the ALICE data taking period LHC11a. The run numbers are listed in Table C.2 of Appendix C. Results from the pass 2 reconstruction are used in the analysis. The magnetic field of the L3 magnet was set to $B = -5\,\mathrm{kG}$. This is the same strength as in the data set recorded at $\sqrt{s} = 0.9\,\mathrm{TeV}$ but with opposite direction. With μ between 0.03 and 0.05, pileup is negligible for these data sets, too.
In the data taking period corresponding to the data sets at $\sqrt{s} = 2.76\,\mathrm{TeV}$, a fast detector read-out has been commissioned in which the comparably slow Silicon Drift Detector (SDD) has been excluded from data taking. A significant fraction of the events in this period has been recorded using the fast read-out. However, in approximately 20 million of the collected events, the SDD has been included in the read-out partition. Only the fraction of events which has been recorded using also the SDD is considered in the data analysis in order to have the same detector set-up for all data sets. The remaining fraction of the events corresponds to 51 % of all minimum bias events at this collision energy.

4.1.3. Proton-Proton Collision Data at $\sqrt{s} = 7.0\,\mathrm{TeV}$

ALICE has recorded 270 million minimum bias proton-proton collision events at $\sqrt{s} = 7.0\,\mathrm{TeV}$ of good quality. The data has been collected in the four data taking periods LHC10b, LHC10c, LHC10d, and LHC10e recorded between April 2010 to August 2010. The run numbers for all data taking periods are listed in Table C.3 of Appendix C. In the data taking period LHC10b, 25 million minimum bias events of good quality have been collected. During these runs, the magnetic field of the L3 magnet has been set to both +5kG and $B = -5\,\mathrm{kG}$. The mean number of interactions varies between $\mu = 0.004$ and 0.03. The calibration and reconstruction has been iterated three times (pass 3).
68 million events of good quality have been collected in LHC10c ($B = +5\,\mathrm{kG}$, $\mu = 0.01 - 0.16$, pass 3), 107 million events in LHC10d ($B = -5\,\mathrm{kG}$, $\mu = 0.004 - 0.03$, pass 2) and in LHC10e, 72 million events have been collected ($B = -5\,\mathrm{kG}$, $\mu = 0.02 - 0.07$, pass 2). In total, this corresponds to 27 % of the recorded proton-proton collision data at this collision energy recorded in 2010.

The impact of pileup events on the data analysis results for the data recorded at $\sqrt{s} = 7.0\,\text{TeV}$ is studied with a special data sample recorded during the LHC10d period. The mean number of interactions μ in these runs varies between $\mu = 0.475$ and 2.03. These runs are not included in the final analysis results because the pileup events could influence the analysis results as described in Section 7.10.

4.2. Monte Carlo Proton-Proton Collisions Data

For the data analysis, Monte Carlo simulations of proton-proton collisions are used in two ways:

1. Monte Carlo data after full detector simulation and reconstruction are used to determine the detector performance. This includes the trigger efficiency, the vertex reconstruction efficiency, the tracking efficiency as well as the contamination of the track sample with tracks from secondary particles after quality cuts. The detector configurations used in the detector simulation correspond to real detector configurations present during data taking. So-called "anchor runs" deploy the same detector configuration as present during the real data runs.
2. The analysis results of real collision data are compared to analysis results of Monte Carlo generator results.

4.2.1. Standard Simulations

Different models of soft particle production and multiple parton interactions are used in the comparison to real data. The models are Pythia6.4 [SMS06] (tune Perugia-0 [Ska10] and Perugia-2011 [Ska10]), Pythia8 [SMS08] (tune 4C [CS11]), and Phojet (version 1.12) [ERR95] (cf. Section 1.6.1 and 1.6.2).
As input for full detector simulations, only Pythia6 Perugia-0 and Phojet events are used. These two event generators differ substantially from each other. This allows to study the impact of the event generator choice on the data correction procedure (cf. Section 7.2).
The Monte Carlo data sets used in the comparison to ALICE data and in the correction procedure are listed in the Tables C.4, C.5, and C.6 in Appendix C.

4.2.2. Special Geant4 Transport Simulation

The detector simulation is mainly performed using the transport Monte Carlos Geant3 [BBM$^+$87]. A special simulation using the successor of Geant3, Geant4 [Gea03, Gea06], has been performed which is used to estimate the systematic uncertainty generated by the choice of the transport Monte Carlo (cf. Section 7.4).
In contrast to Geant3 (written in Fortran), Geant4 is a simulation "toolkit" written in

C++. Geant4 incorporates an increased set of physics processes. While Geant3 has essentially only a single setup, Geant4 users have to select among various "physics lists". These physics lists optimize the transport simulation for various energy regimes and physics applications (such as particle physics, nuclear physics, accelerator design, space engineering, and medical physics) and for different trade-offs of simulation speed and simulation accuracy.
In the course of this thesis, the integration of Geant4 into the ALICE simulation framework has been validated and optimized in order to use the transport Monte Carlo Geant4 in addition to ALICE's default transport Monte Carlo Geant3 [HDG+11]. In particular, a default physics lists has been selected based on the performance of the simulation tests and the accuracy.

4.2.3. Special Therminator Event Simulation

In order to estimate the contribution of short lived particles to the correlation strength studied in the two-particle angular correlation analysis, events generated by the Therminator event generator [KTBF06] are analyzed (cf. Section 5.4.2). Therminator is used, because it does not include the simulation of hard parton-parton collisions, jets, and their fragmentation.
Therminator is a Monte Carlo generator which has been developed for the study of particle production in ultra-relativistic heavy-ion collisions. It incorporates thermal models of particle production with single freeze-out, i. e. the chemical and the thermal freeze-out occur simultaneously [KTBF06]. Therminator simulations comprise all particles of the Particle Data Table [PDG10] including the simulation of the decay of unstable particles.
Even though Therminator is an event generator primarily used for the simulation of heavy-ion collisions, it can be used to study effects that occur also in other colliding systems (e. g. proton-proton collisions) such as the decay of unstable particles. In order to use Therminator for non-heavy-ion collisions, effects that occur only in heavy-ion collisions need to be switched off (for example, collective effects such as the elliptic flow observed in peripheral heavy-ion collisions [Oll92, STA01]). In addition, the longitudinal and the transverse collision system size, the system temperature, and the chemical potential used to describe the heavy-ion collisions need to be chosen such that the transverse momentum spectra as well as the charged particle multiplicity distributions are of the same magnitude as observed in the studied proton-proton collisions.

4.3. Event Selection

The event sample used in the analysis is composed of events with at least one track with a transverse momentum of $p_{\mathrm{T}} > 0.2\,\mathrm{GeV}/c$ in the acceptance regions of the Inner Tracking System (ITS) and the Time Projection Detector (TPC) of $|\eta| < 0.9$ selected by the minimum bias trigger. In addition, the events have to containing exactly one reconstructed vertex of good quality.

4.3.1. Trigger Selection

The analysis uses events selected by the ALICE minimum bias trigger. Minimum bias triggers are triggers of the lowest selection criteria designed to select preferably all inelastic proton-proton interactions. Hence, these triggers have the lowest possible bias on the triggered event sample in comparison to all inelastic interactions.
The ALICE minimum bias trigger used for this analysis is defined by the following requirements. At least one charged particle needs to be detected in either the SPD or in one of the two VZERO detectors VZERO-A and VZERO-C in coincidence with signals from the two BPTX beam pick-up counters indicating the presence of two passing proton bunches [ALI10a, GO09]. This translates to the requirement of at least one charged particle anywhere in the central eight units of pseudorapidity in coincidence with a proton bunch crossing.
As a natural consequence, the trigger efficiency increases with the charged particle multiplicity. The trigger efficiency is estimated in Section 6.6.
The described minimum bias trigger allows to reject beam-gas events over beam-beam events by cutting on the timing information of the VZERO detectors.

4.3.2. Vertex Selection

Out of the minimum bias events, only those events are selected which have exactly one reconstructed vertex of good quality. The quality of the vertex is ensured by the requirement that at least one track is used to reconstruct the vertex. Furthermore, the reconstructed, longitudinal vertex position needs to be within $|z_{\mathrm{vertex}}| < 10\,\mathrm{cm}$ with respect to the nominal interaction point at $z = 0\,\mathrm{cm}$ along the z-axis (the ALICE coordinate system is described in Appendix B). The choice of events with vertices in this central region of z assures that most of the tracks of $|\eta| < 0.9$ fall into the ITS-TPC acceptance [ALI10a, ALI10c]. In addition, in this range the simulated vertex position agrees with the real data vertex position. The reconstructed vertex positions in z-direction for the 3 different center-of-mass energies are presented in the left panel of Figure 4.1. The right panel shows the excellent agreement of the reconstructed vertex position of real data and simulated data.

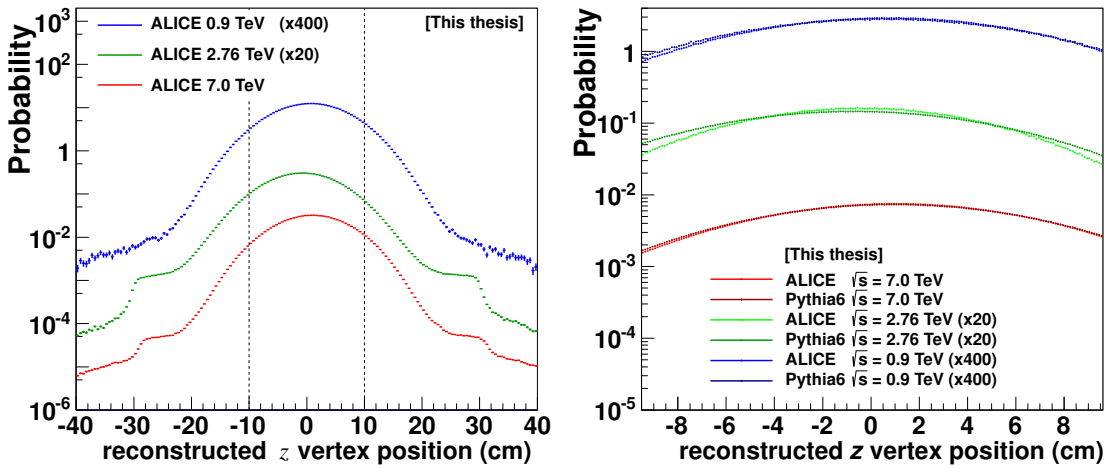

Figure 4.1.: Left panel: Longitudinal, reconstructed vertex position measured at $\sqrt{s} = 0.9,\ 2.76,$ and 7.0 TeV scaled for visibility. Right panel: Comparison of the vertex z-position in real data and simulated data. The simulation is performed with Pythia6 Perugia-0. It is visible that the mean vertex position is not identical with $z = 0\,\text{cm}$ but slightly shifted.

Events with more than one reconstructed vertex (pileup events) either would need to be resolved and analyzed as events with pileup vertices or they would need to be rejected from the analysis. For simplicity, pileup events are rejected from the analysis using a dedicated selection software. Furthermore, all tracks with a distance of closest approach in z-direction to the primary vertex of $\text{DCA}_z > 2\,\text{cm}$ are rejected from the analysis (see the track selection in Section 4.4. The DCA is introduced in Appendix D).
The impact of the remaining pileup events after the pileup rejection is estimated and discussed in Section 7.10.
The impact of the vertex reconstruction efficiency on the data analysis results is discussed in Section 7.6.

4.4. Track Selection

In this analysis, only tracks of charged primary particles are considered. The definition of primary particles includes all prompt particles produced in the collision and all decay products (with $c\tau < 10\,\text{mm}$[1]), except products from weak decays of strange particles (with $c\tau > 10\,\text{mm}$) [ALI10a].
The tracks of the charged particles are measured by the two main tracking detectors of the ALICE central barrel, the ITS and the TPC. The track selection criteria are based on the "ITS-TPC track cuts 2010" which have been developed for data recorded in the data taking period 2010. These track cuts have been used before in ALICE data analyses, e. g. [ALI10c]. The cut selection is motivated by a high reconstruction probability of tracks for primary particles in combination with a strong rejection of tracks from secondary particles. An additional optimization of these track cuts has been performed in order to account for the reduced reconstruction efficiency of tracks in the region of switch-off SPD half-staves [ALI11f, Val12], resulting in the "optimized ITS-TPC track cuts 2010".
The optimized ITS-TPC track cuts 2010 represent the default track selection in this data analysis. They are given by the following selection criteria:

- A successful refit procedure of the ITS and TPC reconstruction during the global tracking procedure[2]. This cut selects tracks which have a high precision in the track parameters.
- At least three hits in the ITS per track, one of which needs to be located in the first three ITS layers. This cut reduces the contamination of the track sample with tracks from secondary particles, which are produced in the interaction with the ITS material or in the decay of strange particles at displaced secondary vertices. While the standard ITS-TPC 2010 track cuts request a hit in the first two ITS layers, the optimized cuts allow a hit in one of the first three ITS layers. This modification increases the reconstruction probability of tracks which traverse detector regions of switched-off SPD half-staves as shown in the left panel of Figure 4.2.
- At least 70 reconstructed TPC clusters (out of 159 pad rows in the TPC). A minimum number of clusters per track assures a good definition of the track found in the track fitting procedure.

[1] c is the speed of light and τ is the mean life time of the particle produced in the collision. In the example process of $K^0_L \rightarrow \pi + \pi$, with a $c\tau = 15.34\,\text{m}$, the mother is classified as the primary particle and the daughters are classified as the secondary particles. In the example of $D^0 \rightarrow K^+ + \pi^-$ with a $c\tau = 122.9\,\mu\text{m}$, the daughter particles are considered as primary particles.

[2] The reconstruction of tracks of charged particles in the ALICE detector is based on a Kalman filter technique. The track finding in the central barrel starts in the TPC. Then, tracks found in the TPC are propagated to the outer layer of the ITS followed by the propagation to the inner layer of the ITS. From here, the track finding is repeated in outwards direction. The track finding continues beyond the TPC including the outer detectors. A final refit procedure from the outer detector layers inwards is applied to optimize the track parameters close to the primary vertex [Hri06].

- A maximal χ^2 per TPC cluster of 4. The χ^2 per TPC cluster determines the quality of the fit between the track and the contributing clusters. Fake high-p_{T} tracks are significantly suppressed.
- No tracks with a kink topology indicating a particle decay.
- A p_{T}-dependent DCA_{xy} cut corresponding to 7σ of the input track sample ($\mathrm{DCA}_{xy,\ max} \approx 0.2\,\mathrm{cm}$). This cut assures that the tracks originate in the primary vertex or rather close to it. Hence, it selects tracks from primary particles and rejects tracks from secondary particles.
- A distance of closest approach in z-direction DCA_z of maximal 2 cm (the DCA is introduced in Appendix D). Similarly to the DCA_{xy} cut, this cut improves the selection of primary particles above secondary particles. In addition, the cut removes tracks coming from displaced pileup vertices.
- The tracks need to be placed within the combined ITS-TPC acceptance of $|\eta| < 0.9$ and $p_{\mathrm{T}} > 0.2\,\mathrm{GeV}/c$. This cut assures an optimal detector coverage of the track and hence a good track quality.

The left panel of Figure 4.2 displays the increase of the reconstruction efficiency between the default ITS-TPC track cuts and the optimized ITS-TPC track cuts as function of the azimuthal angle φ. The dips in φ correspond to the areas of the switched-off SPD half-staves.

The distribution of the azimuthal angle φ for tracks accepted by the optimized ITS-TPC track selection is presented in the right panel of Figure 4.2 for real data and Pythia6 Perugia-0 simulations at $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV. The simulations reproduce the real data fairly well. The damage of the SPD is largest for the data set at $\sqrt{s}$ = 2.76 TeV, the data set which has been recorded at last.

For the study of the systematic uncertainty related to the choice of the track cuts, also a variety of other tracks cuts samples is used (cf. Section 7.5). For this purpose, the nominal cuts, the optimized ITS-TPC track cuts (1), are tightened (2) and relaxed (3) within reason. In addition, so-called TPC-only track cuts (4) are used.

TPC-only tracks cuts are widely used in correlation analyses performed within the ALICE collaboration [ALI12f, ALI11b, ALI12b]. For the reconstruction of TPC-only tracks, mainly the information of the TPC detector is used. The advantage of the TPC-only track cuts over the discussed ITS-TPC track cuts is that they provide a very flat azimuthal acceptance. However, TPC-only track cuts have a larger contamination of the track sample with tracks from secondary particles.

The four track cut samples are summarized in Table 4.1.

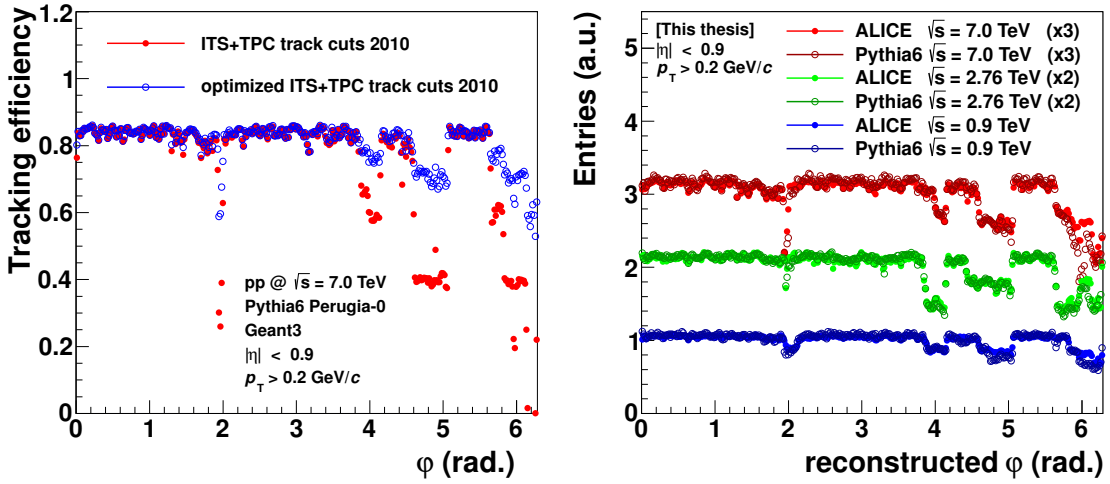

Figure 4.2.: Left panel: Tracking efficiency for the ITS-TPC track cuts 2010 and the optimized ITS-TPC track cuts 2010. Right panel: Reconstructed azimuthal angle φ for tracks of the optimized ITS-TPC track cuts for real data and Pythia6 Perugia-0 simulations at $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV.

	"Default"	"Tight"	"Loose"	"TPC-only"
ITS refit	yes	yes	yes	no
TPC refit	yes	yes	yes	yes
ITS-hits (1 to 6)	(1,2,3)+any	(1,2,3)+any	(1,2)+any	-
χ^2/NDF in TPC	4	4	4	4
$\mathrm{DCA}_{xy,\ \mathrm{max}}$[cm]	0.2*	0.1*	2.4	2.4
$\mathrm{DCA}_{z,\ \mathrm{max}}$[cm]	2	2	3.2	3.2
Vertex 2D†	no	no	no	yes
Kinks	no	no	no	no
$N_{\mathrm{min,\ TPC\ cluster}}$	70	80	70	70

Table 4.1.: The track selections used in the estimation of the systematic uncertainty of the track selection choice estimated in Section 7.5. As default track selection, the optimized ITS-TPC track cuts are used.
*In the default and in the tight track selection, the DCA_{xy} cut is based on the p_{T}-dependent DCA_{xy} cut corresponding to 7σ of of the input tracks sample. In the tight track selection, an additional limitation of $|\mathrm{DCA}_{xy}| < 0.1$ cm is applied.
†The requirement of a successful vertex finding of the two-dimensional vertex finding algorithm is only applied in the TPC-only track selection.

4.5. Quality Assessments

In this section, the central quality assurance (QA) of the collision data is described and extended by a quality assurance analysis for the two inner tracking detectors.
The general ALICE data quality assurance framework is described in Section 4.5.2. In Section 4.5.3, a new QA data analysis focusing on reconstructed signals of the central tracking detectors, the Inner Tracking System (ITS) and the Time Projection Chamber (TPC), is presented. In Section 4.5.4, this quality assurance analysis is used to demonstrate that all data sets listed in Section 4.1 are of good quality. In Section 4.5.5, the quality evolution within the presented data taking periods is monitored.

4.5.1. Motivation

Before processing recorded collision data, the data quality needs to be assessed. Common examples of data deficiency tests cover the following points.

1. Mis-configurations during the collision data taking need to be identified before reconstructing the raw collision data.
2. Problems in the reconstruction algorithms, the detector calibration, and detector alignment need to be identified and repaired.
3. Data with malfunctioning or blind detector regions, e. g. disabled detector modules, need to be discovered. These data sets need to be specially treated or rejected in the data analysis if the sub-detector information is needed.

Without these checks, physics signals could appear distorted leading to a wrong interpretation of data analysis results. Early data quality assurance also helps to minimize the waste of computing resources.

4.5.2. Global ALICE Quality Assurance Scheme

A first data quality assurance is already done during data taking by a dedicated data quality monitoring system (DQM) [vH$^+$11]. The DQM system retrieves information of online configurations during the data taking allowing for a direct feedback on the data quality.
Once recorded and accepted as good collision data, the raw data signals are calibrated and reconstructed using the AliRoot framework for simulation, reconstruction, and analysis [ALI05]. Immediately after the data calibration and reconstruction, specially designed data analyses are used to monitor the reconstructed quantities measured with the sub-detectors of ALICE. These analyses are combined in the so-called quality assurance analysis train (QA-train). The QA-train is an application of the general ALICE analysis framework [Ghe08].
The output of the QA-train allows to classify the data sets in terms of the performance

of each sub-detector separately. If the quality of the reconstructed data is not sufficiently good but it is likely to gain better quality in another iteration of calibration and reconstruction, another chain of calibration, reconstruction, and quality assurance is performed on the data sets.
The classification of the data sets is summarized in the dedicated ALICE Run Condition Table [ALI12a] hosted by the MonALISA framework [LNV+08]. From here, the data sets can be selected by the physics analysis groups depending on the requirements of the specific data analyses.

4.5.3. Central Tracking Quality Assurance Focusing on ITS and TPC

This section describes a new analysis for the QA-train designed to assess the performance of the central tracking detectors ITS and TPC: Properties of both reconstructed events and tracks are monitored including stand-alone tracks of the ITS and TPC detectors as well as combined tracks of ITS and TPC.
The performance assessment is based on the observation of the following event properties:

- The reconstructed position of the primary vertex in x, y, and z-direction of
  - the primary vertex estimated using SPD information,
  - the primary vertex estimated using ITS-TPC tracks,
  - the primary vertex estimated using TPC-only tracks.
- The number of accepted tracks after quality cuts.

For reconstructed tracks of accepted events after vertex quality cuts, the following observables are analyzed:

- Pseudorapidity η,
- azimuthal angle φ,
- transverse momentum p_{T},
- distance of closest approach in xy-direction DCA_{xy},
- distance of closest approach in z-direction DCA_{z}.

A powerful tool to detect displacement or distortions of the detector as well as biases in the reconstruction software is the analysis of the signed distance of closest approach to the event vertex of positive and negative tracks separately and the yield ratios of positive and negative tracks [vL06]. Any experimental bias with fixed direction in space results in opposite biases for tracks of opposite charge (or opposite pseudorapidity). Correlations between the above listed track and event properties are monitored, too.
Deviations from nominal track efficiencies as a function of the pseudorapidity η and the azimuthal angle φ uncover switched-off detector modules or indicate deviations from nominal data taking conditions.
As an example for the identification of malfunctioning detector components, Figure 4.3 shows the correlation of the track pseudorapidity η and its azimuthal angle φ for two data

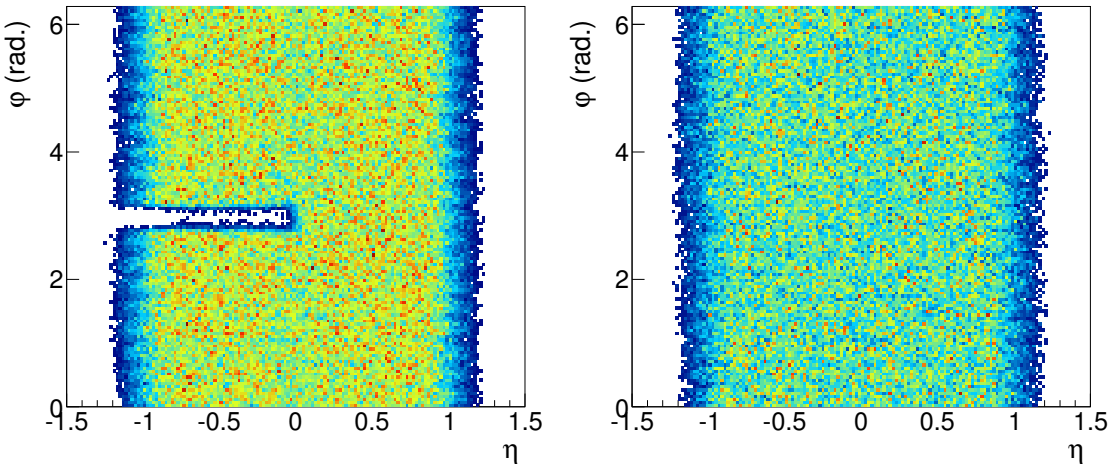

Figure 4.3.: Correlation between reconstructed pseudorapidity η and azimuthal angle φ of TPC-only tracks measured in lead-lead collisions at $\sqrt{s_{\mathrm{NN}}} = 2.76\,\mathrm{TeV}$. The left panel shows data where the TPC outer read-out chamber C08 has not been read out during data taking (data set LHC11h-169506). The right panel shows the same distribution of a data set measured when all TPC modules were well functioning (data set LHC11h-167807).

sets. One of the data sets (left panel of Figure 4.3) has been recorded while one if the TPC read-out sectors was switched off. This results in a reduced number of reconstructed TPC-only tracks at $\eta < 0$ and $\varphi \approx 3$[3]. Both data sets show the central membrane of the TPC (at $\eta = 0$) and the gaps between the TPC sectors (18 divisions in φ) by a slightly reduced number of reconstructed tracks in these areas.
The afore mentioned properties of the reconstructed events and tracks are analyzed for all data sets separately. Also, the evolution of the average properties within a given data taking period is under investigation in order to check the stability of the calibration and reconstruction.

4.5.4. Detailed Quality Analysis of Used Data Sets

In this section, track and event properties for the data sets described in Section 4.1 are validated using the new central tracking quality assurance. For the event and track selection, the selection criteria defined in Section 4.3 and 4.4 are used. Thus, a decent quality level of the data sets used for this thesis is ensured.
The reconstructed position of the primary vertex in x, y, and z-direction measured at $\sqrt{s} = 0.9,\ 2.76,\ 7.0\,\mathrm{TeV}$ is shown in Figure 4.4. The position is measured with respect to the nominal interaction point placed in the center of the ALICE central barrel. The vertex

[3]The azimuthal angle φ has the unit radians (rad.). For a better readability, in the following φ is presented without this unit.

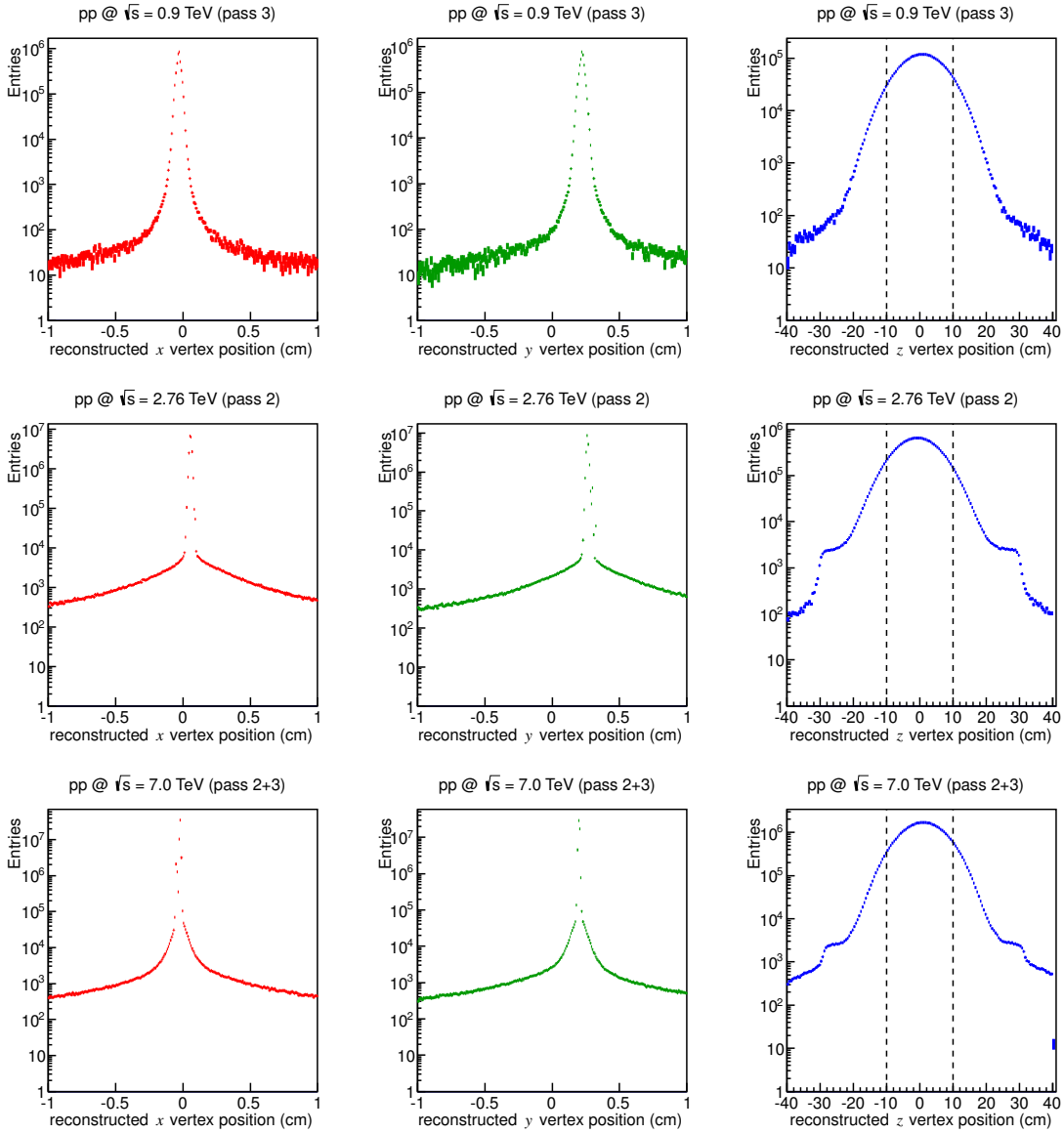

Figure 4.4.: Reconstructed vertex position in x, y and z-direction measured at $\sqrt{s} = 0.9\,\text{TeV}$ (top panels), 2.76 TeV (panels in the middle), and 7.0 TeV (bottom panels).

position averaged over all measured events should be distributed around a single mean interaction point. The z-distribution within the accepted area of $|z_{\text{vertex}}| < 10\,\text{cm}$ should be describable by a Gaussian function. Only properties of events of $|z_{\text{vertex}}| < 10\,\text{cm}$ from the nominal interaction point are considered in further QA and analysis steps.

Figure 4.4 shows that the reconstructed vertex positions of all data sets are sharply distributed around a single mean interaction point in x, y, and z per collision energy. The positions are not exactly centered at the nominal interaction point $x = y = z = 0\,\text{cm}$, however, they come close to it. Slight deviations of the interaction region from the

nominal interaction point are expected because the accuracy in the positioning of the bunch crossing region is limited. Problems in the reconstruction as well as vertices of parasitic collisions would result in broad vertex distributions and extra peaks at displaced values of x, y, and z. These effects are not visible in the data sets at hand.

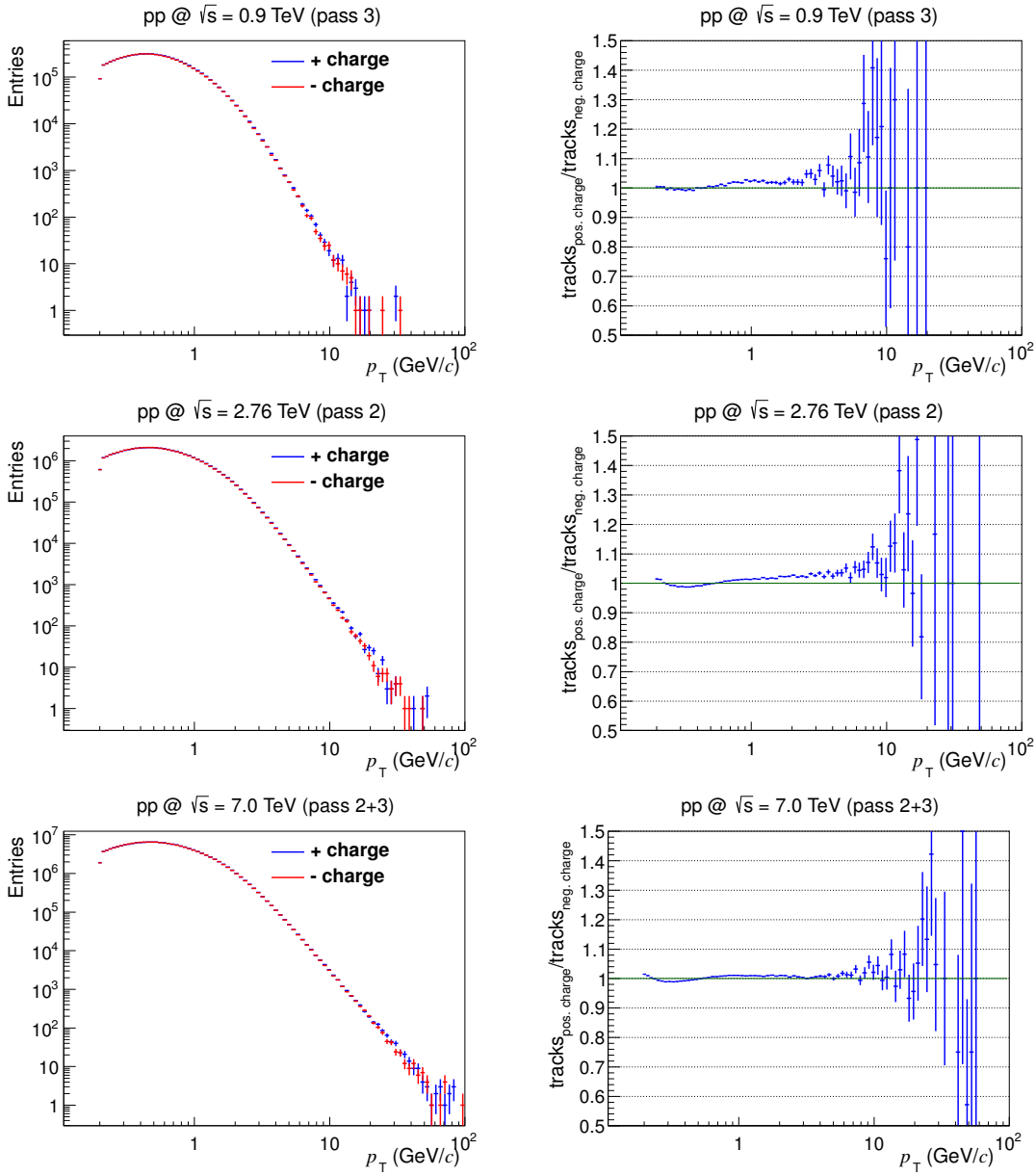

Figure 4.5.: Reconstructed transverse momentum distribution separated for positive and negative charged tracks (panels at the left side) and the ratio of positive to negative charged tracks (panels at the right side). The upper panels present results measured at $\sqrt{s} = 0.9\,\text{TeV}$. The panels in the middle present results measured at $\sqrt{s} = 2.76\,\text{TeV}$. The panels at the bottom present results measured at $\sqrt{s} = 7.0\,\text{TeV}$.

The transverse momentum p_{T} distribution is studied for tracks with positive charge and negative charge separately. In case of problems in the detector calibration or reconstruction, the ratio of positive and negative tracks differs clearly from unity. In first approximation, the ratio of the p_{T}-spectra of positive and negative charged particles are in good agreements for all center-of-mass energies (Figure 4.5). A slight deviation of the ratio from unity is expected because of the difference in the reconstruction efficiency for positive and negative particles. In terms of the ratio of positive and negative tracks, the data sets at hand fulfill the quality requirements.

Furthermore, the distribution of the distance of closest approach of the tracks to the primary vertex is studied. For well calibrated and reconstructed tracks after track selection cuts, the mean value of the DCA distribution is centered around zero. The distribution should have only one peak and it should be symmetric when averaging over all charges.

The panels at the left side of Figure 4.6 show the DCA in xy-direction and the panels at the right side show the DCA in z-direction. The DCA distributions are centered at zero and they have a smooth shape without distortions. The data sets at hand do not show problems in the reconstruction which manifest in distortions of the DCA distributions.

The reconstruction efficiency of a particle depends on the number of hits in the detector associated to the track. If any module of a sub-detector is switched off or is running at a gain below the nominal value, the reconstruction efficiency has a lower value in the affected region. Therefore, the average $\eta - \varphi$–distribution of tracks per event uncovers parts of the detectors which have not been working at nominal conditions.

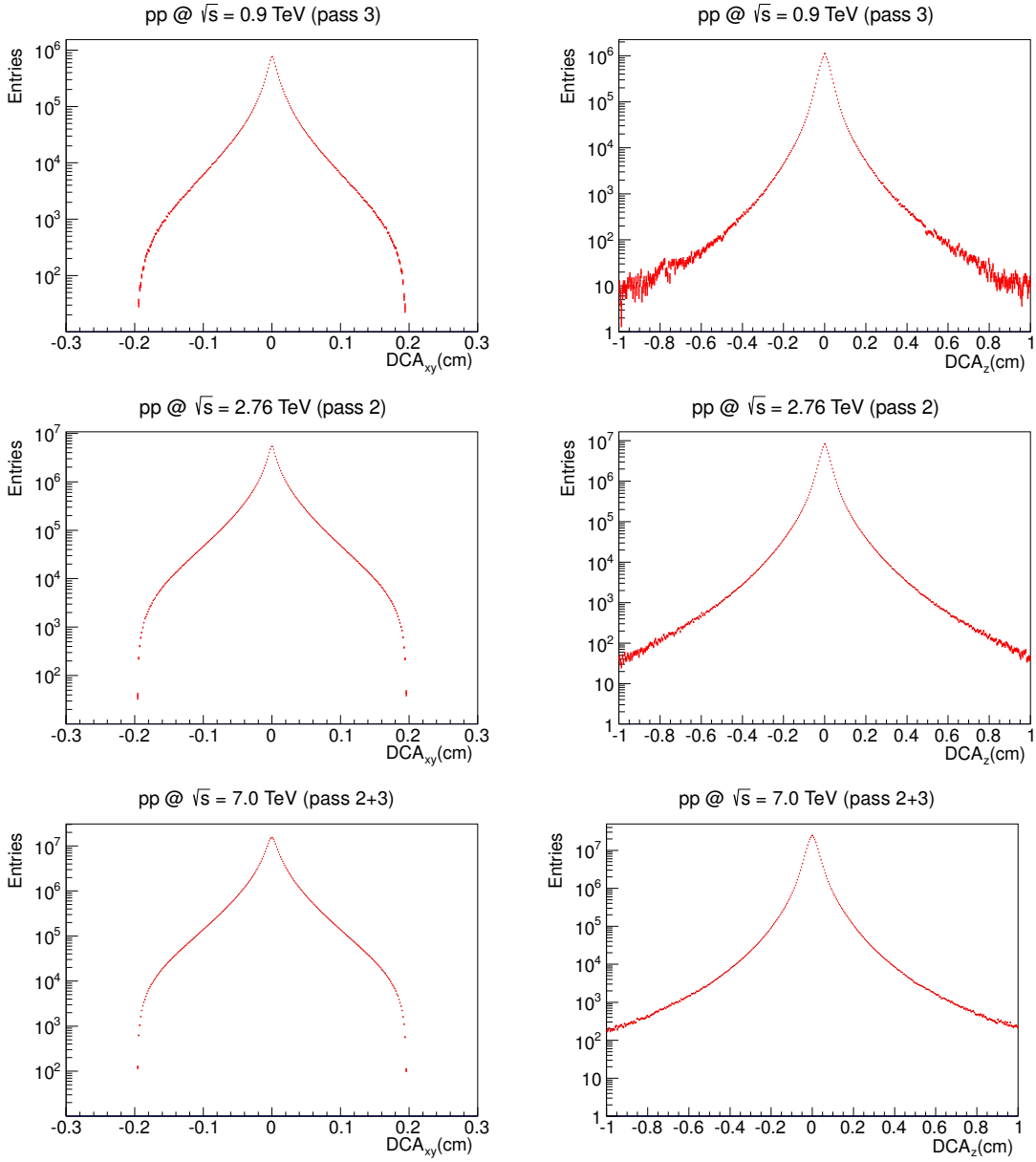

Figure 4.6.: The distance of closest approach (DCA) of the tracks to the primary collision vertex. The panels at the left side show the DCA_{xy} in xy-direction. The panels the right side show the DCA_z in z-direction. The upper panels present results measured at $\sqrt{s}$ = 0.9 TeV. The panels in the middle present results measured at $\sqrt{s}$ = 2.76 TeV. The bottom panels present results measured at $\sqrt{s}$ = 7.0 TeV.

The reconstructed $\eta - \varphi$ track distributions of the three center-of mass energies are shown in Figure 4.7. As mentioned before, over time, the number of operable SPD half-staves have decreased over time. The effect of the switched-off half staves at $\varphi \approx 5$ is visible for all data sets. While the data set at $\sqrt{s}$ = 0.9 TeV recorded in May 2010 exhibits only small gaps, the data set at $\sqrt{s}$ = 2.76 TeV recorded in March 2011 shows an increased

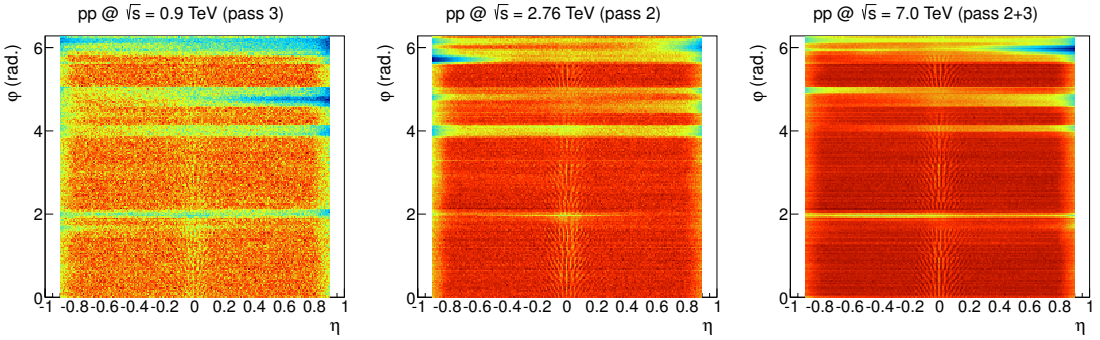

Figure 4.7.: Correlation between the reconstructed track properties pseudorapidity η and azimuthal angle φ. The panel at the left side shows track properties measured at $\sqrt{s} = 0.9\,\text{TeV}$. The panel in the middle shows track properties measured at $\sqrt{s} = 2.76\,\text{TeV}$. The panel at the right side shows track properties measured at $\sqrt{s} = 7.0\,\text{TeV}$.

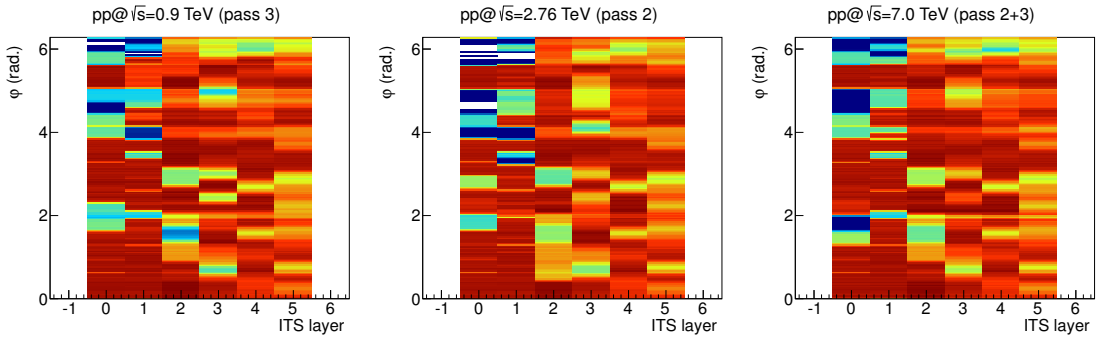

Figure 4.8.: Reconstructed azimuthal angle φ correlated combined with the position of the associated ITS hits per track (0,1=SPD, 2,3=SDD, 4,5=SSD). The panel at the left side shows results measured at $\sqrt{s} = 0.9\,\text{TeV}$. The panel in the middle shows results measured at $\sqrt{s} = 2.76\,\text{TeV}$. The panel at the right side show results measured at $\sqrt{s} = 7.0\,\text{TeV}$.

number of gaps. The damage in the SPD can not be rewound. Hence, Figure 4.7 displays the nominal detector conditions. If additional detector modules were switched off as shown in Figure 4.3, additional structures would be visible in the $\eta - \varphi$-distribution. This is not the case for the data sets at hand.

In addition to the $\eta - \varphi$–distributions of the tracks, it is useful to monitor the correlation between the track direction (η or φ) and the position of the associated hits per track within the different layers of the Inner Tracking System. Layer 0 and 1 represent the Silicon Pixel Detector (SPD), layer 2 and 3 are the Silicon Drift Detector (SDD), and layer 4 and 5 are the Silicon Strip Detector (SSD).
The correlation between the azimuthal angle φ of the track and the associated hits per ITS layer are shown in Figure 4.8. The switched-off half-staves in the SPD are clearly visible

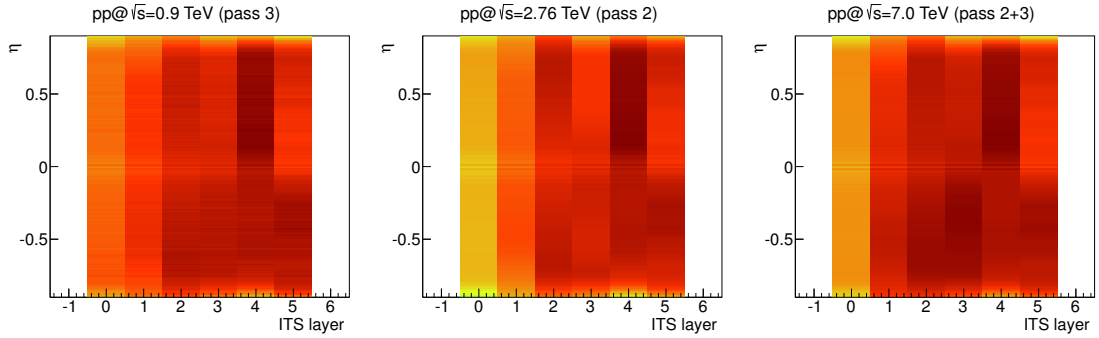

Figure 4.9.: Reconstructed pseudorapidity η correlated with the position of the associated hits in the ITS layers (0,1=SPD, 2,3=SDD, 4,5=SSD). The panel at the left side show results measured at $\sqrt{s} = 0.9\,\text{TeV}$. The panel in the middle show results measured at $\sqrt{s} = 2.76\,\text{TeV}$. The panel at the right side show results measured at $\sqrt{s} = 7.0\,\text{TeV}$.

as blue or white areas. Also, it can be seen that the number of missing modules in the SPD increased with time. Besides the known issues of the switched-off SPD half-staves, no additional dead areas in the ITS detector are visible. All three sub-detectors of the ITS have been running at nominal conditions during the data taking for the data sets at hand.

The correlation between the pseudorapidity η of the track and the associated hits per ITS layer are shown in Figure 4.9. For all ITS layers, the hits are equally distributed in η. Only a reduced number of reconstructed tracks is visible at mid-rapidity (TPC central membrane) and at high η. In both regions, the probability of the track reconstruction is reduced. In these regions, the tracks need to cross the central-membrane or the end-plates of the TPC instead of the sensitive detector area. Less clusters per track are produced causing a reduced reconstruction probability. Also, the detector coverage of tracks of large η is limited. A fraction of the trajectory can fall outside the ITS-TPC acceptance.

In summary, the detectors have been running at nominal conditions during the data taking of the data sets at hand. The presented quality assurance checks revealed no problems in the calibration or the reconstruction.

4.5.5. Evolution of Average Track and Event Properties

By monitoring the evolution of event and track properties for all data sets recorded over a long period, the stability of the detector performance and its calibration can be tested. Furthermore, it is possible to identify outliers in terms of average track and event properties, which are most likely a result of corrupted data sets.
In the following, the track and event property evolution is exemplarily shown for the data sets recorded at $\sqrt{s} = 7.0\,\text{TeV}$. The evolution of the event and track properties for the three collision energies $\sqrt{s} = 0.9$, 2.76, and 7.0 TeV are shown in Appendix E.

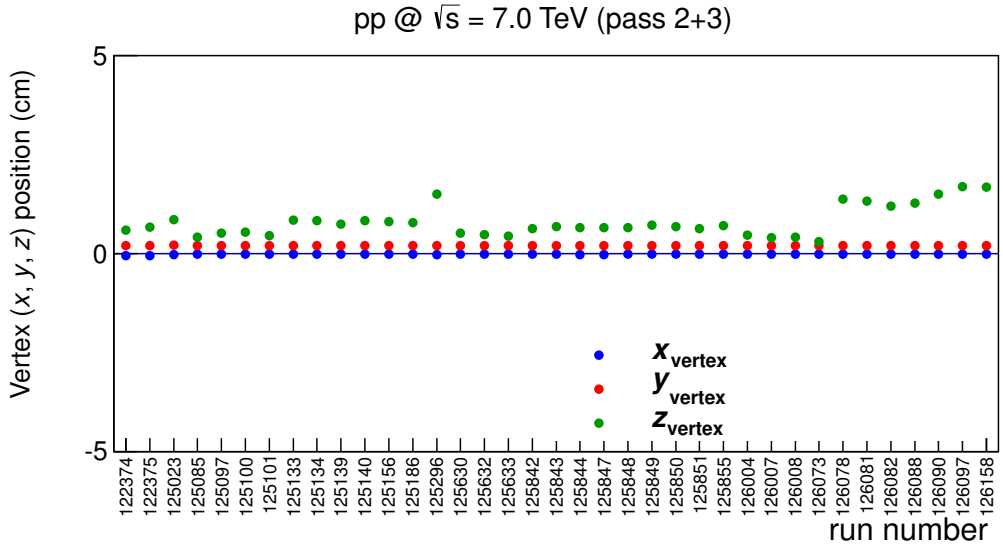

Figure 4.10.: Vertex position in x, y, and z-direction measured at $\sqrt{s} = 7.0\,\text{TeV}$. The list of the run numbers is shown in Appendix C.

The mean of the reconstructed primary vertex position in x, y, and z-direction is shown in Figure 4.10. The x and y-position are comparably stable. The z-position varies due to slightly different configurations provided by the LHC in the different LHC fills. Still, the mean position of the vertex is always placed in the central region of ALICE with $|z_{\text{vertex}}| < 10$ cm.

In Figure 4.11, the mean values of the DCA_{xy} distribution per run are presented. In addition, the same values for sub-sets of the tracks are shown: The DCA for positive and negative charged tracks and for tracks of positive and negative pseudorapidity. The mean DCA values are stable for all runs.

Figure 4.12 presents the evolution of the mean values of DCA_z for all runs. The evolution of the DCA_z values is stable for all runs. However, a shift to positive values of DCA_z

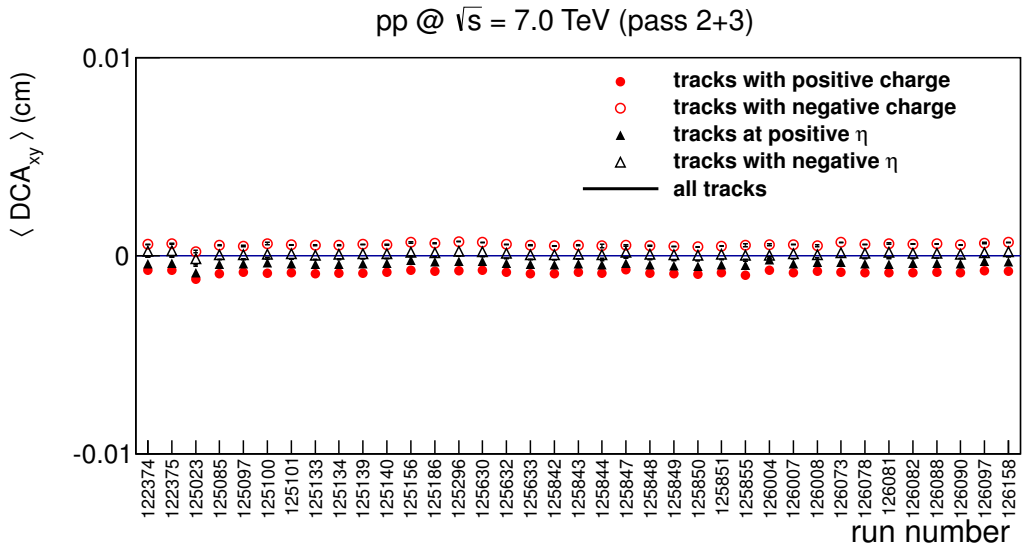

Figure 4.11.: The mean values of the DCA_{xy}-distribution per run measured at $\sqrt{s} = 7.0\,\mathrm{TeV}$.

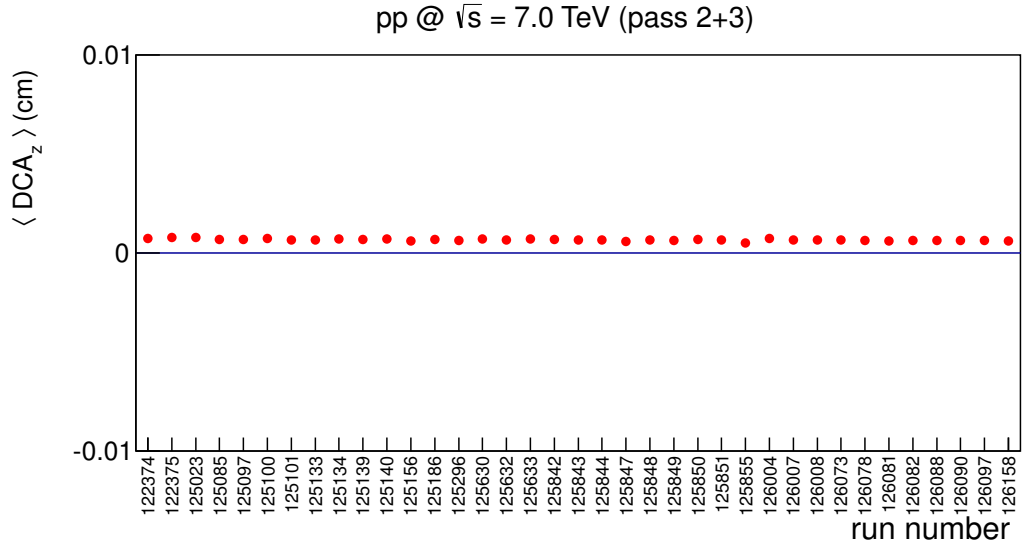

Figure 4.12.: The mean values of the DCA_z-distribution per run measured at $\sqrt{s} = 7.0\,\mathrm{TeV}$.

is visible for all runs. The shift can be explained by the shift of the vertex position in z direction, which itself is shifted to positive values, too.

Figure 4.13 shows the average ratio of positive and negative tracks. Only tracks with the transverse momentum of $1 < p_{\mathrm{T}} < 5\,\mathrm{GeV}/c$ are considered in this ratio in order to guarantee a stable fit result. The ratios are slightly above 1 but stable for all runs within the statical error bars. The ratio is above 1 due to the slight difference in the reconstruction probabilities for tracks of positive and negative charged particles.

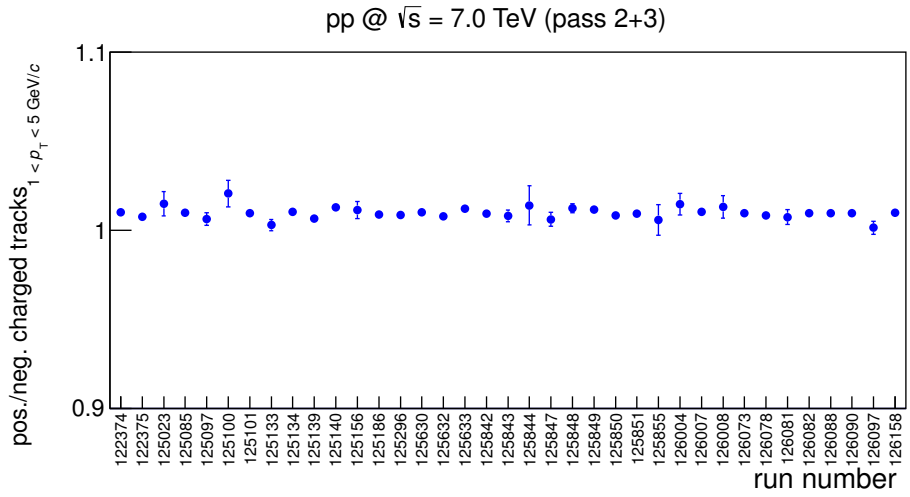

Figure 4.13.: The average ratio of tracks with positive charge to tracks with negative charge per run measured at $\sqrt{s}$ = 7.0 TeV. Only tracks with a reconstructed transverse momentum of $1.0 < p_{\mathrm{T}} < 5.0\,\mathrm{GeV}/c$ are shown. This allows to estimate a stable fit result.

4.5.6. Conclusion

In this section, a quality assurance analysis for the two inner tracking detectors ITS and TPC has been presented. The presented QA analysis has become part of the ALICE wide quality assurance framework for the central estimation of the quality of all reconstructed data sets.

For the data sets used in this thesis, it could be shown that the calibration and the reconstruction has been performed satisfactorily and that all relevant detector components of the Inner Tracking System and the Time Projection Chamber have been operating properly. This additional validation has been necessary as the data set selection relies heavily on the performance of the central tracking system.

4.6. Summary

In this chapter, the input for the data analysis presented in this thesis has been discussed. The input is a selection of proton-proton collision data sets at three different center-of-mass energies, $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV. The data sets have been selected based on quality estimations of the ALICE wide quality assurance framework with a focus on the performance of the central tracking system. The proton-proton collision event selection is based on a minimum bias trigger and a vertex selection including a rejection of events with pileup vertices. Optimized ITS-TPC track cuts have been introduced for the selection of the tracks.

5. Analysis Method

High-p_{T} two-particle angular correlation analyses have been used for the study of jets and the underlying hard scattering process for some time, e. g. [UA182a, CMS10].
In the following, a data analysis method based on two-particle angular correlations is introduced. The data analysis investigates the charged particle multiplicity dependence of two-particle angular correlations in proton-proton collisions. The measurement aims for a better understanding of the jet fragmentation, the contribution of jets to the overall charged particle multiplicity, and the nature of multiple parton interactions in proton-proton collisions.
In Section 5.1, the algorithm of the two-particle angular correlation analysis is introduced. The direct observable of the two-particle correlation, the azimuthal correlation, is discussed in Section 5.2. The signal extraction method using a fit function and the resulting derived observables of the azimuthal correlation are described in Section 5.3. In the subsequent Section 5.4, the impact of combinatorics in high-multiplicity jets as well as the impact of decays of short lived particles on the two-particle correlation is estimated. In Section 5.5, the feasibility to access the number of multiple parton interactions (MPI) in proton-proton collisions is discussed using Pythia simulations [SMS06, SMS08], as Pythia is an event generator that incorporates the physics of MPI.

5.1. Description of the Data Analysis Algorithm

In the two-particle correlation data analysis, pairs of particles of the same collision event are built. The particle pairs are formed by trigger particles defined by the transverse momentum threshold $p_{\mathrm{T}} > p_{\mathrm{T,\,trig}}$ and associated particles defined by $p_{\mathrm{T}} > p_{\mathrm{T,\,assoc}}$. Trigger particles can act as associated particles to other trigger particles as well. Particles with a transverse momentum below the trigger particle momentum threshold but above the associated particle momentum threshold ($p_{\mathrm{T,\,assoc}} < p_{\mathrm{T}} < p_{\mathrm{T,\,trig}}$) act as associated particles only. Particles with a transverse momentum below the associated transverse momentum threshold ($p_{\mathrm{T}} < p_{\mathrm{T,\,assoc}}$) do not enter the particle correlation at all. The data analysis is performed for particles measured in the central region of the ALICE detector given by the pseudorapidity range $|\eta| < 0.9$.
Each pair of particles is considered only once. Two particles are combined to a particle pair only if both particles belong to the same collision event.

For an event with N_{trig} trigger particles and $N_{\mathrm{pure\,assoc}}$ purely associated particles of $p_{\mathrm{T,\,assoc}} < p_{\mathrm{T}} < p_{\mathrm{T,\,trig}}$, the following number of pairs can be built

$$N_{\mathrm{pair}} = \frac{N_{\mathrm{trig}} \cdot (N_{\mathrm{trig}} - 1)}{2} + N_{\mathrm{trig}} \cdot N_{\mathrm{pure\,assoc}}. \tag{5.1}$$

These pairs are formed for each collision event and averaged over all analyzed collision events. Events without any track in the ALICE acceptance are discarded from the analysis. From the remaining events, three distributions are considered:

$$\begin{aligned} &N_{\mathrm{pair}}(\Delta\varphi,\ \Delta\eta,\ p_{\mathrm{T,\,trig}},\ p_{\mathrm{T,\,assoc}},\ f_{\mathrm{sign}},\ N_{\mathrm{ch}}),\\ &N_{\mathrm{trig}}(p_{\mathrm{T}},\ \eta,\ N_{\mathrm{ch}}),\\ &N_{\mathrm{event}}(z_{\mathrm{vertex}},\ N_{\mathrm{ch}}). \end{aligned} \tag{5.2}$$

$\Delta\varphi$ represents the difference in terms of the azimuthal angle between the trigger and the associated particles, $\Delta\varphi = \varphi_{\mathrm{trig}} - \varphi_{\mathrm{assoc}}$, $\Delta\eta$ is the difference in the pseudorapidity, $\Delta\eta = \eta_{\mathrm{trig}} - \eta_{\mathrm{assoc}}$.
f_{sign} indicates whether the particles of the pair have the same charge or opposite charge.
The charged particle multiplicity $N_{\mathrm{ch}} = N_{\mathrm{charged},\,|\eta|<0.9,\ p_{\mathrm{T}}>0.2\,\mathrm{GeV}/c}$ is given by all accepted tracks defined by the track cuts introduced in Section 4.4 and located in the central acceptance region of ALICE given by $|\eta| < 0.9$ and $p_{\mathrm{T}} > 0.2\,\mathrm{GeV}/c$.
For trigger particles, the transverse momentum p_{T}, the pseudorapidity η and the charged particle multiplicity are considered. For the characterization of the collision event, the position of the primary vertex in z-direction, z_{vertex}, and also the charged particle multiplicity are taken into account. The pair, trigger, and event properties are used, on one hand, for the characterization of the data allowing to define various pair, trigger, and event classes. Hence, a comparison between analysis results between those different classes is possible. On the other hand, the pair, trigger, and event information is needed during the data correction procedure discussed in Chapter 6. For example, the correction for the tracking efficiency is very sensitive on the transverse momentum of the particles. The multi-dimensional approach of the analysis is realized by sparse histograms[1].
All observables of the presented two-particle correlation analysis are derived from the basic associated per-trigger yield as a function of the azimuthal angle

$$\frac{dN}{d\Delta\varphi} = \frac{1}{N_{\mathrm{trig}}} \frac{dN_{\mathrm{assoc}}}{d\Delta\varphi}. \tag{5.3}$$

This associated per-trigger yield can be computed from the distributions given in Equation 5.2 for various configurations in N_{charged}, $p_{\mathrm{T,\,trig}}$, $p_{\mathrm{T,\,assoc}}$, f_{sign} and $\Delta\eta$. Thereby, the analysis allows to access the per-trigger associated yield and its dependence on the given pair properties, trigger particle properties, and event properties.

[1] THnSparse of the data analysis framework ROOT [BR97]

5.1.1. Definition of Regions

As the difference in the azimuthal angle is periodic ($\Delta\varphi = 0 = 2\pi$), the $\Delta\varphi$-range is limited to the essential range of 2π. The $\Delta\varphi$-limits are chosen to be $[-\pi/2, 3\pi/2]$ in order to provide a good visibility of the correlation patterns peaked around 0 and π. Pairs with $\Delta\varphi$-values outside that range are redistributed to corresponding values at $\Delta\varphi_{\text{trans}} = \Delta\varphi_{\text{orig}} \pm 2\pi$.
In the following analysis of the $\Delta\varphi$-distributions, the hemisphere of the trigger particle in $\Delta\varphi$ is referred to as the "near side" of the trigger particle. The opposite side in $\Delta\varphi$ of the trigger particle is called "away side". The edges of the regions are given by the following limits:

- near side: $-\frac{\pi}{2} < \Delta\varphi < \frac{\pi}{2}$,
- away side: $\frac{\pi}{2} < \Delta\varphi < \frac{3\pi}{2}$.

5.2. Direct Correlation Observable

A typical per-trigger yield as a function of the difference in the azimuthal angle $\Delta\varphi$ for Pythia6 Perugia-0 events at $\sqrt{s} = 7.0\,\text{TeV}$ is shown in Figure 5.1. The threshold for trigger and associated particles is chosen to be $p_{\text{T, trig}} > 0.7\,\text{GeV}/c$ and $p_{\text{T, assoc}} > 0.4\,\text{GeV}/c$, respectively. The per-trigger yield is integrated over all charged particle multiplicities, it includes pairs of all analyzed $\Delta\eta$, and pairs of both like-signed charges and unlike-signed charges.

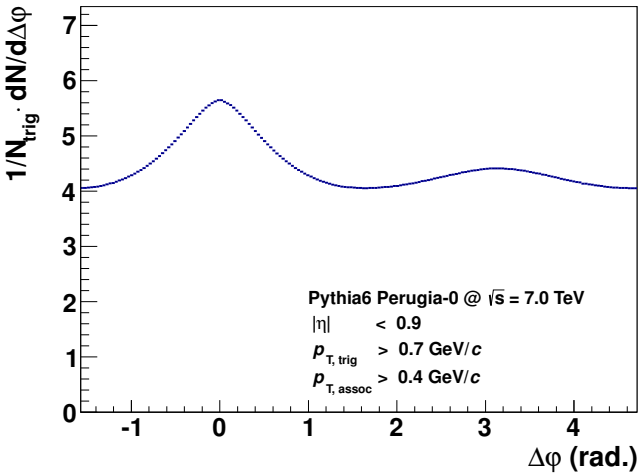

Figure 5.1.: Example per-trigger yield as a function of $\Delta\varphi$ for Pythia6 Perugia-0 simulations of proton-proton collisions at $\sqrt{s} = 7.0\,\text{TeV}$ and $p_{\text{T, trig}} > 0.7\,\text{GeV}/c$ and $p_{\text{T, assoc}} > 0.4\,\text{GeV}/c$.

The multi-dimensional approach of the analysis allows to build the presented $\Delta\varphi$-distribution separated for 100 bins in N_{ch}, 40 bins in $p_{\mathrm{T,\,trig}}$, 40 bins in $p_{\mathrm{T,\,assoc}}$, 10 bins in $\Delta\eta$, and 2 bins for same charge and opposite charge of the particle pairs. The number of bins per dimension is limited by reasonable memory consumption during the analysis.
Due to the limited number of analyzed events, not all bins of the sparse histograms are filled. Nevertheless, at least 10,000 $\Delta\varphi$-distribution for various combinations of pair, event, and track properties can be build.
When comparing the complete $\Delta\varphi$-correlation pattern between the various combinations of pair, trigger, and event properties, a huge amount of information has to be handled at a given time.
A first step in the management of the vast information is to fix almost all pair, trigger, and event properties. Then, the evolution of the $\Delta\varphi$-distribution as a function of only one single pair, trigger, or event property (e. g. the event charged particle multiplicity) can be studied.

5.3. Derived Correlation Observables

A further simplification in the comparison of the $\Delta\varphi$-distributions is the reduction of the $\Delta\varphi$-distribution into its components.

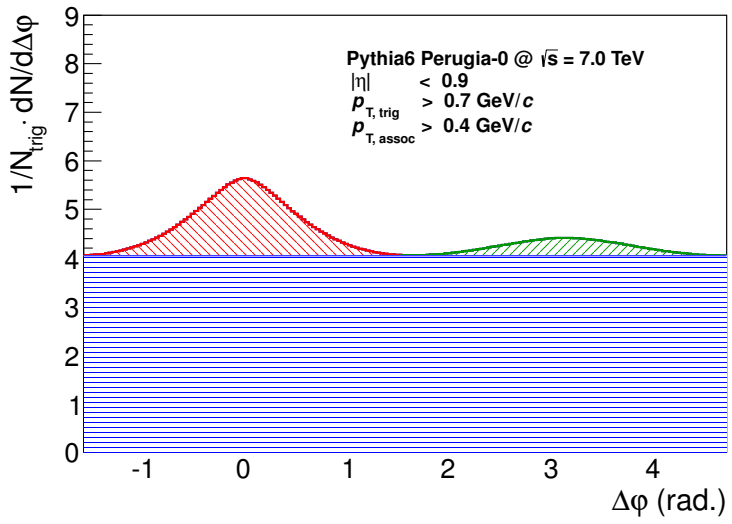

Figure 5.2.: Geometrical contributions to the example per-trigger yield as a function of $\Delta\varphi$ with $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ and $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$ measured in Pythia6 Perugia-0 events at $\sqrt{s} = 7.0\,\mathrm{TeV}$. Combinatorial background (blue), near side peak dominated by associated yield of jet (red), away side peak dominated by associated yield of corresponding backward jets (green).

In a simplified view, the $\Delta\varphi$-distribution of the two-particle correlation as given in Figure 5.2 can be characterized by the following, most prominent properties:

- the average per-trigger yield in the isotropically distributed, combinatorial background, denoted by $\langle N_{\text{isotrop}} \rangle$,
- the average per-trigger yield of the associated tracks at the near side and the away side, denoted by $\langle N_{\text{assoc, near side}} \rangle$ and $\langle N_{\text{assoc, away side}} \rangle$, respectively, both reduced by the isotropically distributed background.

Furthermore, the width of the $\Delta\varphi$-distribution of the near and away side peak can be used to characterize the distribution. Also, the relation between all described observables can be analyzed.

5.3.1. Description of the Correlation using a Fit Function

A fit function is used to extract the properties of the correlation. The integral of the near side peak and the away side peak is dominated by the central parts of the peaks at $\Delta\varphi \approx 0$ and $\Delta\varphi \approx \pi$, respectively. Thus, minor shape deviations at the side of the peak regions at $\Delta\varphi \approx -\pi/, \pi/2$ and $3\pi/2$ are in principle considered acceptable[2]. In the following, the fit function is introduced step by step.

The $\Delta\varphi$-distribution can roughly be approximated by a constant and two Gaussian functions. One Gaussian function is centered at the near side peak ($\Delta\varphi = 0$) and another Gaussian function is centered at the away side peak ($\Delta\varphi = \pi$). The offset of the $\Delta\varphi$-distribution is given by a constant function.

As the $\Delta\varphi$-distribution is a periodically continuing distribution with $\Delta\varphi = 0 = 2\pi$, the fit function has to be periodically continuing as well. Hence, copies of the described Gaussian functions need to be added at integer multiples of π, $\Delta\varphi \pm n2\pi$, $n \in \mathrm{Z}$. For example, the near side peak is located at $\Delta\varphi = 0$ and also at $\Delta\varphi = \pm 2\pi$, $\pm 4\pi$ and $\pm 6\pi$, the away side peak is located at $\Delta\varphi = \pi$ and also $\Delta\varphi = -\pi$, $\pm 3\pi$ and $\pm 5\pi$.

The periodically continuation of the fit function to values below $\Delta\varphi < 0$ and above $\Delta\varphi > 2\pi$ has to be performed even though the pair $\Delta\varphi$-values from outside the displayed $\Delta\varphi$-range are redistributed to values between $-\pi/2$ and $3/2\pi$. Otherwise, it could happen during the fitting procedure that the fit function tries to approach 0 for values outside the $\Delta\varphi$-acceptance.

As an approximation, the periodical continuation can be achieved by using only a single additional term per Gaussian functions. The periodical continuation of the Gaussian functions are located at $\Delta\varphi \pm 2\pi$ with respect to the first versions.

Furthermore, it has been found that one Gaussian function does not fully describe the near side peak at $\Delta\varphi = 0$. The near side peak exhibits in comparison to a single Gaussian

[2] In Section 7.1 it is shown that the fit function introduced in the section provides an excellent description of the $\Delta\varphi$-distribution for all analyzed data sets.

function a narrow inner peak and a broad outer peak. Hence, the integral of the near side peak is described better when using two independent near side Gaussian functions at $\Delta\varphi = 0$, one for the narrow contribution and one for the broad contribution. The two near side Gaussian functions should have slightly different widths, however both Gaussian functions are centered at $\Delta\varphi = 0$.
It has been found that the away side peak at $\Delta\varphi = \pi$ is sufficiently described by one Gaussian function only.
This approximation results in the following new fit function,

$$\begin{aligned} f(\Delta\varphi) = C &+ A_1 \exp\left(-\frac{\Delta\varphi^2}{2\cdot\sigma_1^2}\right) &&+ A_1 \exp\left(-\frac{(\Delta\varphi - 2\pi)^2}{2\cdot\sigma_1^2}\right) \\ &+ A_2 \exp\left(-\frac{\Delta\varphi^2}{2\cdot\sigma_2^2}\right) &&+ A_2 \exp\left(-\frac{(\Delta\varphi - 2\pi)^2}{2\cdot\sigma_2^2}\right) \\ &+ A_3 \exp\left(-\frac{(\Delta\varphi - \pi)^2}{2\cdot\sigma_3^2}\right) &&+ A_3 \exp\left(-\frac{(\Delta\varphi + \pi)^2}{2\cdot\sigma_3^2}\right). \end{aligned} \tag{5.4}$$

$f(\Delta\varphi)$ combines a constant C and Gaussian functions, where the A_i represent the amplitude and the σ_i represent the width of the Gaussian functions. The mean position of the single Gaussian functions are fixed to $\Delta\varphi = 0$ and $\Delta\varphi = 2\pi$ for the near side Gaussian functions and $\Delta\varphi = \pi$ and $\Delta\varphi = -\pi$ for the away side Gaussian functions.
In order to reduce the free parameters of the fit function, the constant C is substituted by the integrated yield of the $\Delta\varphi$-distribution reduced by the integrals of the Gaussian functions.
As a further simplification of the fit function, the allowed range of the Gaussian functions is limited to $[-\pi/2, \pi/2]$ for the near side and $[\pi/2, 3\pi/2]$ for the away side. For the shifted Gaussian functions, the fit ranges are shifted accordingly to $[-\pi/2 \pm 2\pi, \pi/2 \pm 2\pi]$ for the shifted near side and $[\pi/2 \pm 2\pi, 3\pi/2 \pm 2\pi]$ for the shifted away side. The function can be fitted with a stable result when using suitable start values and limits for the remaining parameters.
Figure 5.3 shows an example $\Delta\varphi$-distribution and its description by the fit function given in Equation 5.4. Here, the four components of the fit, the summed fit result as well as the real data distribution are presented. The fit and the data $\Delta\varphi$-distribution lay on top of each other. The agreement is further studied in Section 7.1.

5.3.2. Signal Extraction

Using the parameters of the fit function (Equation 5.4), three correlation observables can be derived. These describe (1) the integrated yield in the average combinatorial background of the correlation, (2) the integrated near side yield, and (3) the integrated

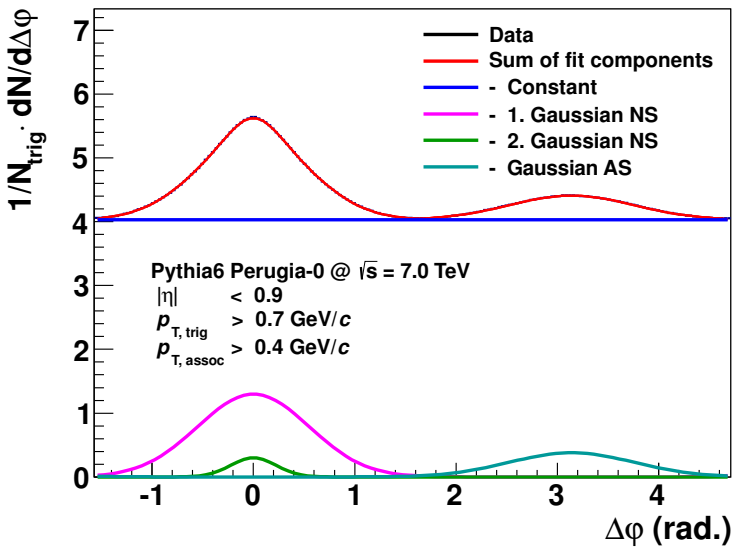

Figure 5.3.: Fit function describing the example per-trigger yield as a function of $\Delta\varphi$ with $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ and $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$ measured in Pythia6 Perugia-0 events at $\sqrt{s} = 7.0\,\mathrm{TeV}$. The black line represents the measured $\Delta\varphi$-distribution. The read line represents the sum of all contributions to the fit function. The green and the magenta line represent the Gaussian functions of the near side. The line in cyan represents the away side Gaussian function. The dark blue line represents the constant.

away side yield normalized to the number of trigger particles. Furthermore, the average number of trigger particles per event can be measured. The number of uncorrelated sources of particle production can be derived from these observables. It will be shown that this observable is a measure of the number of multiple parton interactions.
The formal description of the average per-trigger yield in the combinatorial background is

$$\langle N_{\mathrm{isotrop}} \rangle = \frac{1}{N_{\mathrm{trigger}}} \cdot C. \tag{5.5}$$

The average per-trigger associated yield of the near side above the combinatorial background can be described by the integral of the near side Gaussian functions

$$\langle N_{\mathrm{assoc,\,near\ side}} \rangle = \frac{\sqrt{2\pi}}{N_{\mathrm{trigger}}}(A_1 \cdot \sigma_1 + A_2 \cdot \sigma_2), \tag{5.6}$$

and the average per-trigger associated yield of the away side above the combinatorial background is given by the integral of the away side Gaussian function

$$\langle N_{\mathrm{assoc,\,away\ side}} \rangle = \frac{\sqrt{2\pi}}{N_{\mathrm{trigger}}}(A_3 \cdot \sigma_3). \tag{5.7}$$

The average number of trigger particles per event can be computed as

$$\langle N_{\mathrm{trigger}} \rangle = \frac{N_{\mathrm{trigger}}}{N_{\mathrm{events}}}. \tag{5.8}$$

Using these observables, it is possible to extract the number of uncorrelated sources of particle production which corresponds to the number of multiple parton interactions. This relation is explained in the following.
The number of sources of particle production can be estimated using the number of *correlated* trigger particles in comparison to *all* trigger particles. In one collision event, the average correlated faction of the trigger particles is given by the trigger particle defining the event axis for the correlation itself, the associated trigger particles at the near side and the associated trigger particles at the away side, both above the combinatorial background. Due to the correlation, it can be concluded that on average all of these trigger particles come from the same particle production process. For example, in events with 20 trigger particles, each trigger particle could have in average 2 associated trigger particles at the near side and in average 2 trigger particles at the away side above the combinatorial background. Together, 5 (1+2+2) trigger particles are correlated to each other. In average, $20/5 = 4$ independent sources of particle production have produced those trigger particles. In general, the number of sources of particle production is given by the ratio of the average number of trigger particles to the average number of correlated trigger particles. In the following, this observable is called number of uncorrelated seeds. The formal description is

$$\langle N_{\text{uncorrelated seeds}} \rangle = \frac{\langle N_{\text{trigger}} \rangle}{\langle N_{\text{trigger, correlated}} \rangle} = \frac{\langle N_{\text{trigger}} \rangle}{\langle 1 + N_{\text{assoc, near+away, } p_{\text{T}} > p_{\text{T, trig}}} \rangle} \quad (5.9)$$

Here, $\langle N_{\text{assoc, near+away, } p_{\text{T}} > p_{\text{T, trig}}} \rangle$ is given by the sum of the results of $\langle N_{\text{assoc, near side}} \rangle$ (Equation 5.6) and $\langle N_{\text{assoc, near side}} \rangle$ (Equation 5.7) obtained by using $p_{\text{T, assoc}} = p_{\text{T, trig}}$.
The observables of Equation 5.5 to 5.9 as a function of the charged particle multiplicity N_{ch} are presented in Figure 5.4. The results are obtained using Phojet simulations at $\sqrt{s} = 7.0\,\text{TeV}$ using $p_{\text{T, trig}} > 0.7$ and $p_{\text{T, assoc}} > 0.7\,\text{GeV}/c$.

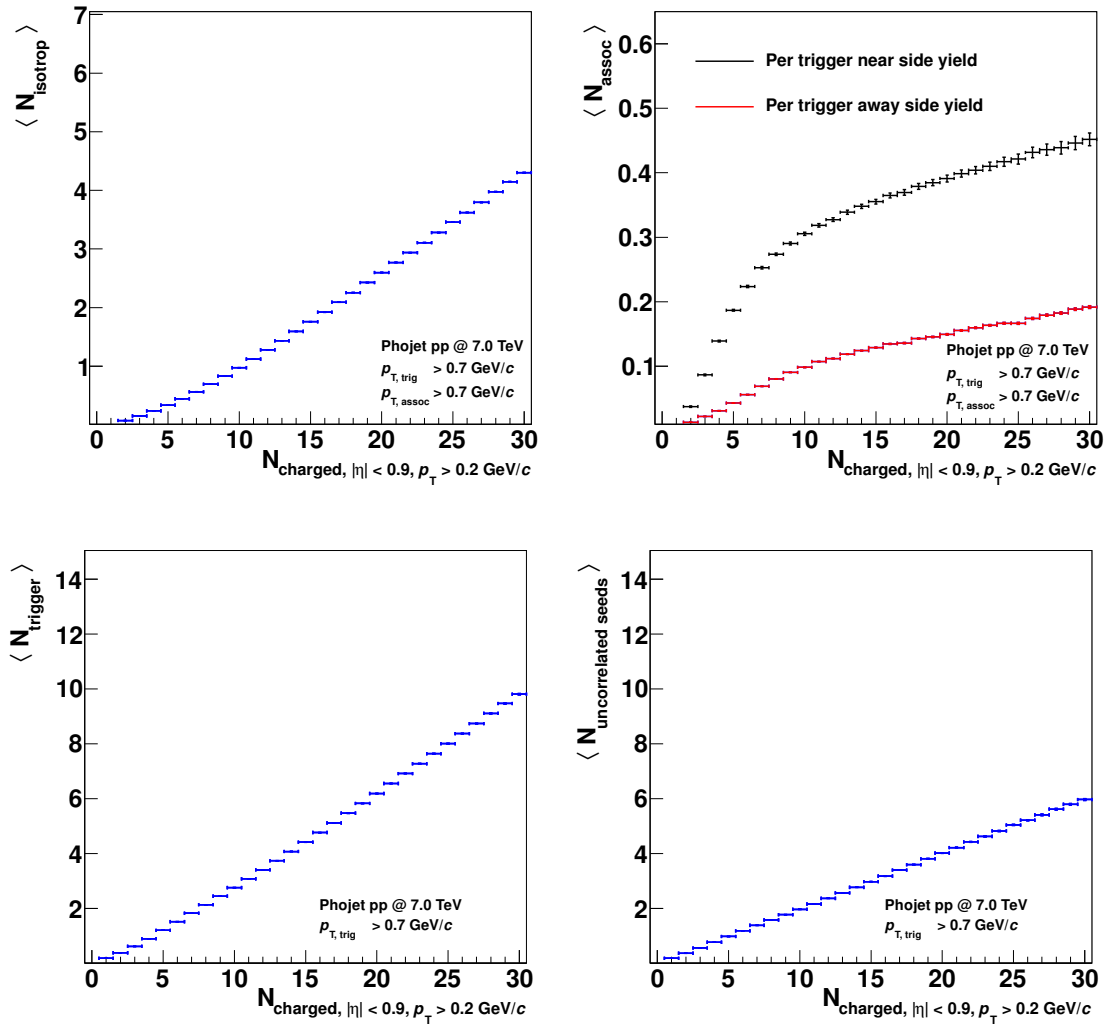

Figure 5.4.: Correlation observables as a function of the charged particle multiplicity. Top row, left panel: Average per-trigger yield in the combinatorial background. Top row, right panel: Average per-trigger associated yield for the near and the away side. Bottom row, left panel: Average number of trigger particles per collision event. Bottom row, right panel: Average number of uncorrelated seeds per collision event.

5.4. Consideration of Biases in Two-Particle Correlations

A key contribution of the correlation data analysis is the estimation of the average number of particles produced per hard process (jet multiplicity). In the following section, it is shown that under certain conditions the correlation observable $\langle N_{\mathrm{assoc}} \rangle$ is identical with the average number of associated charged particles per hard process. This identity is not trivial.

For instance, in an auto-correlation where each particle is used as trigger particle, jets

with a high number of particles per jet (high multiplicity jets) contribute stronger to the correlation as low-multiplicity jets. A bias towards high-multiplicity jets is the result. The estimated average number of associated particles will always be shifted to higher values.
Another example are correlation analyses with respect to the leading particle of the event. The leading particle of a collision event is the particle with the highest value in p_{T}. Events and jets of high multiplicity have due to the high number of particles most likely a higher $p_{\mathrm{T,\,leading}}$ than low multiplicity events. In a correlation measurement which is performed with respect to the leading particle, a bias to the hard momentum scale is introduced when analyzing collision events separated for the charged particle multiplicity.
In the following Section 5.4.1, the conditions are described under which the bias to high-multiplicity jets and the bias to the hard momentum scale can be suppressed. The impact of correlated decay products of short lived particles to the per-trigger near side yield is discussed in Section 5.4.2.

5.4.1. Impact of Auto-Correlations of High-Multiplicity Jets

In the following, the intra-jet charged particle multiplicity is denoted by the probability distribution $P(N)$. For an a priori unknown intra-jet multiplicity distribution $P(N)$, the per-trigger auto-correlation of trigger particles is given by

$$\frac{1}{\langle N\rangle}\frac{\langle N(N-1)\rangle}{2} = \frac{1}{2}\left(\frac{\langle N^2\rangle}{\langle N\rangle} - 1\right). \tag{5.10}$$

For steadily falling multiplicity distributions $P(N)$ and small mean number of particles per jet $\langle N\rangle$, Equation 5.10 can be transformed in good approximation to

$$\frac{1}{2}\left(\frac{\langle N^2\rangle}{\langle N\rangle} - 1\right) \rightarrow \frac{\langle N\rangle}{1-P(0)} - 1, \tag{5.11}$$

which becomes equal to $\langle N\rangle$ with the trigger condition -1. Therefore, it is true that in the two-particle correlation analysis for jets with steeply falling multiplicity distributions and low mean number of trigger particles, the mean number of associated particles per-trigger particle can be measured.
The validity of Equation 5.11 is tested for known steadily falling functions, such as the geometric series, a Poisson distribution, and the log-series.

Example: Geometrical Series

The geometrical series is given by

$$P_{\mathrm{geo}}(N) = (1-q)q^N. \tag{5.12}$$

Here, N represents the jet multiplicity and q is a free parameter. The mean value of $P_{\mathrm{Geo}}(N)$ is given by

$$\langle N\rangle = \frac{q}{1-q}. \tag{5.13}$$

The average value of N^2 of the geometrical series can be written as

$$\langle N^2 \rangle = 2\langle N \rangle^2. \tag{5.14}$$

Hence, the left side of Equation 5.11 can be transformed to

$$\frac{1}{2}\left(\frac{\langle N^2 \rangle}{\langle N \rangle} - 1\right) = \langle N \rangle, \tag{5.15}$$

and the right side of Equation 5.11 can be transformed to

$$\frac{\langle N \rangle}{1 - P(0)} - 1 = \langle N \rangle. \tag{5.16}$$

For the geometrical series, Equation 5.11 is exact.

Example: Poisson Distribution

The Poisson distribution has the following parametrization:

$$P_{\text{Poisson}}(N) = \frac{\mu^N e^{-\mu}}{N!}. \tag{5.17}$$

The mean value of the Poisson distributions is given by μ

$$\langle N \rangle = \mu, \tag{5.18}$$

and the $\langle N^2 \rangle$ is given by

$$\langle N^2 \rangle = \mu^2 + \mu. \tag{5.19}$$

The left side of Equation 5.11 can therefore be transformed to

$$\frac{1}{2}\left(\frac{\langle N^2 \rangle}{\langle N \rangle} - 1\right) = \frac{\mu}{2}. \tag{5.20}$$

The right side of Equation 5.11 can be approximated by

$$\frac{\langle N \rangle}{1 - P(0)} - 1 = \frac{\mu}{1 - e^{-\mu}} - 1 = \frac{\mu}{2} - \frac{\mu^2}{6} + ... \tag{5.21}$$

Thus, to first order, Equation 5.11 is fulfilled for a Poisson distribution.

Example: Log-Series

The log-series is given by

$$P_{\log}(N) = -\frac{1}{\ln(1-p)}\frac{p^N}{N}. \tag{5.22}$$

p is a free parameter of the log-series. The mean value of the log-series is

$$\langle N \rangle = -\frac{1}{\ln(1-p)}\frac{p}{1-p}, \tag{5.23}$$

and the $\langle N^2 \rangle$ is given by

$$\langle N^2 \rangle = \frac{\langle N \rangle}{1-p}. \tag{5.24}$$

Hence, the left side of Equation 5.11 can be transformed to

$$\frac{1}{2}\left(\frac{\langle N^2 \rangle}{\langle N \rangle} - 1\right) = \frac{p}{2(1-p)}. \tag{5.25}$$

while its right side can be approximated by

$$\frac{\langle N \rangle}{1-P(0)} - 1 = \frac{p}{2(1-p)} + \frac{p^2}{3(1-p)} + ... \tag{5.26}$$

In the first order approximation, Equation 5.11 is true for the log-series.

Working Hypothesis

An important observable of the data analysis is the number of uncorrelated seeds $\langle N_{\text{uncorrelated seeds}} \rangle$ (Equation 5.9). As discussed, the number of uncorrelated seeds is assumed to be related to the number of multiple parton interactions.
The validity of Equation 5.9 is only satisfied if the number of the associated particles at the near and away side, $\langle N_{\text{assoc, near side}} \rangle$ and $\langle N_{\text{assoc, away side}} \rangle$, calculated in Equation 5.6 and 5.7, is computed correctly. In the computation of the associated per-trigger yield, a bias to high multiplicity jets needs to be monitored carefully.
In comparison to low multiplicity jets, high multiplicity jets contribute stronger to the correlation in which all particles are used for the correlation. For example, a jet with 5 trigger particles contributes to a simplified auto-correlation which can be computed as

$$N_{\text{pair, auto}} = \frac{N_{\text{trig}} \cdot (N_{\text{trig}} - 1)}{2}, \tag{5.27}$$

with $N_{\text{pair,5 trigger}} = (5 \cdot 4)/2 = 10$ while a jet with 2 trigger particles contributes only with $N_{\text{pair,2 trigger}} = (2 \cdot 1)/2 = 1$. When using all particles per jet for the correlation despite the particle's transverse momentum, high multiplicity jets could dominate the correlation function and the associated yields, $\langle N_{\text{assoc, near side}} \rangle$ and $\langle N_{\text{assoc, away side}} \rangle$, could be overestimated.
In contrast, it could be possible to reduce the bias of high multiplicity jets by applying a transverse momentum threshold for particles that enter the correlation. For this, the trigger threshold $p_{\text{T, trig}}$ used in the correlation has to be chosen such, that no measurable bias to high multiplicity jets is introduced.
The near and the away side yield presented in Figure 5.4 have been estimated using the momentum threshold $p_{\text{T, trig}}$ and $p_{\text{T, assoc}}$ chosen to $0.7\,\text{GeV}/c$. The sum of the per-trigger near side and away side yield is small - below 1 associated particle - for all charged particle multiplicity bins. Therefore, it is assumed, that at least the momentum threshold

of $p_{\mathrm{T}} > 0.7\,\mathrm{GeV}/c$ and all higher values are sufficient to reduce the bias towards high multiplicity jets.
Furthermore, it is assumed that the multiplicity $P(N)$ of particles per jet is a steadily falling function.
In Section 5.5 the number of uncorrelated seeds is computed for Pythia6 simulations. Pythia simulations are based on the principle of multiple parton interactions (cf. Section 1.6.1). The scaling between the measured number of uncorrelated seeds in Pythia and its underlying number of multiple parton interactions is studied in order to verify the data analysis approach as well as the validity of the discussed assumption that the biases are negligible.

5.4.2. Impact of Decay Products of Short Lived Particles

Short lived particles have a very short decay length. Due to their short decay lengths, the secondary vertex, at which the particles decay, cannot be distinguished from the primary collision vertex. Therefore, the decay products are reconstructed as primary particles.
The decay products of short lived particles are correlated with each other. Most likely, they are close to each other in terms of the azimuthal angle φ and the pseudorapidity η. Hence, they can generate a non-jet-like contribution to the azimuthal correlation measured in two-particle angular correlations. In this section, the impact of decays of short lived particles on the correlation analysis is investigated and quantified using the Therminator event generator [KTBF06] (cf. Section 4.2).
The Monte Carlo event generator Therminator generates hadronic events using thermal models for the particle production [KTBF06]. It is originally designed to study the particle production and the space-time evolution of heavy-ion collisions, but it can also be used to study the aspect of particle decays in other colliding systems e. g. proton-proton collisions.
Unlike Pythia, Therminator does not incorporate hard parton-parton collisions, jets, and their fragmentation in the particle production process. Hence, all correlated final state particles generated with the Therminator setup used for this study originate from particle decays. Therefore, Therminator events can be used to estimate the impact of the decays of short lived particles on the correlation strength measured in two-particle angular correlations.

The multiplicity distribution and the transverse momentum spectrum for Pythia6 Perugia-0 and Therminator simulations are shown in Figure 5.5. The spectra differ strongly. When comparing the associated yield between Pythia6 Perugia-0 and Therminator, the trigger thresholds $p_{\mathrm{T,\,trig}}$ and $p_{\mathrm{T,\,assoc}}$ used in the analysis of the two event generators need to be adjusted. For example, the cuts can be chosen such that the yield in the isotropically distributed background is of the same scale.

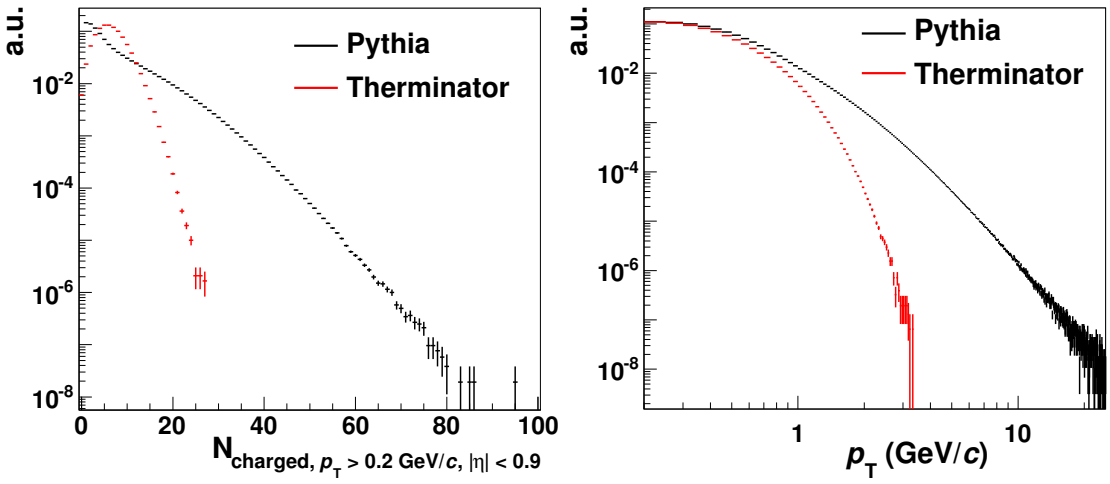

Figure 5.5.: Left panel: Charged particle multiplicity distribution obtained in Therminator and Pythia6 Perugia-0 simulations. Right panel: Transverse momentum spectrum obtained in Therminator and Pythia6 Perugia-0 simulations. The distributions are scaled for the best visibility. The y-axis have arbitrary units (a.u.). The Pythia6 Perugia-0 data are simulated for the center-of-mass energy of $\sqrt{s}$ = 7.0 TeV.

The Therminator results are analyzed for several trigger thresholds $p_{\mathrm{T,\,trig}}$ and $p_{\mathrm{T,\,assoc}}$ as presented in Figure 5.6. When choosing the trigger threshold to values between $p_{\mathrm{T,\,trig}}$ = 0.35 and 0.40 GeV/c, the yield of $\langle N_{\mathrm{isotrop}} \rangle$ is in the same range as for the Pythia6 Perugia-0 simulation at $p_{\mathrm{T,\,trig}} = 0.70\,\mathrm{GeV}/c$.

The per-trigger near side yield estimated for Pythia6 Perugia-0 and Therminator events is presented in Figure 5.7. Compared to the Pythia6 results, Therminator generates a lower near side yield and no away side yield (plot not shown). The Therminator near side yield corresponds to about 10 % of the near side yield measured in Pythia6 simulations. Thus, the contribution of decay products from short lived particles is sizable, though still approximately a magnitude smaller than the jet fragmentation contribution measured in Pythia.

As part of the correction procedure, a two-track correction is introduced for the correction of two-track effects including the contribution of particle decays to the near-side yield (cf. Section 6.3).

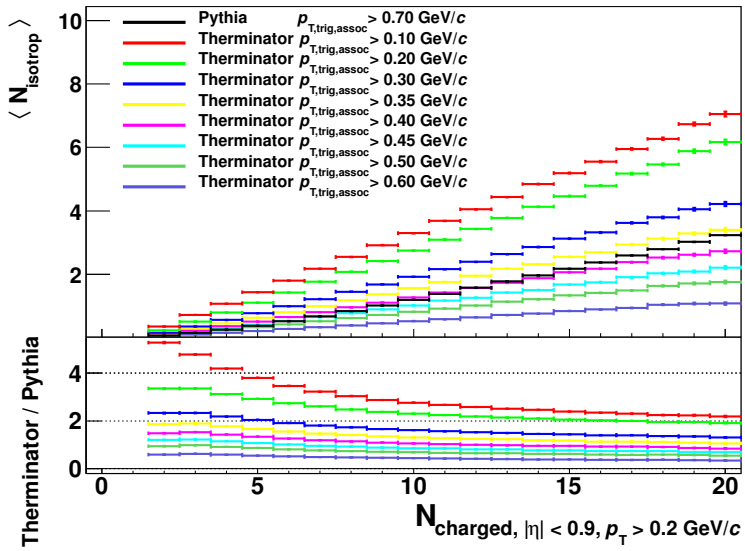

Figure 5.6.: Integrated yield in the isotropically distributed background per trigger particle measured for Therminator and Pythia6 Perugia-0 simulations. For $p_{\mathrm{T,\,trig}}$ and $p_{\mathrm{T,\,assoc}} > 0.35$ or $0.40\,\mathrm{GeV}/c$, the Therminator simulation gives the same per-trigger yield in the combinatorial background, $\langle N_{\mathrm{isotrop}} \rangle$, as obtained at $p_{\mathrm{T,\,trig}}$ and $p_{\mathrm{T,\,assoc}} > 0.70\,\mathrm{GeV}/c$ in the Pythia6 Perugia-0 simulations.

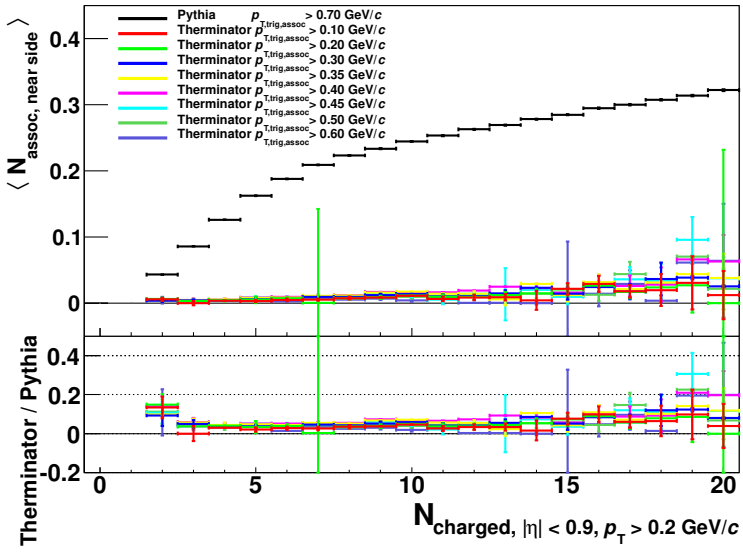

Figure 5.7.: Per-trigger near side yield estimated in Therminator and Pythia6 Perugia-0 simulations. For all transverse momentum thresholds, the per-trigger near side yield is much smaller in the Therminator simulation set than in the Pythia6 Perugia-0 simulation.

5.5. Derivation of the Number of Multiple Parton Interactions

A key observable of the data analysis approach is the estimation of the number of multiple parton interactions N_{MPI} in proton-proton collisions. In the following, the dependence of the measured number of uncorrelated seeds $N_{\text{uncorrelated seeds}}$ on N_{MPI} for Pythia6 [SMS06] simulations will be presented. Pythia6 is used as an example of Monte Carlo generators which simulate proton-proton collisions including the principle of multiple parton interactions.

5.5.1. Multiple Parton Interactions in Pythia

Within Pythia6, the number of multiple parton interactions, N_{MPI}, is defined as the number of hard or semi-hard scatterings that occurred in a collision event in the multiple interaction scenario [SMS06]. The scale of the (semi-)hard scatterings in Pythia is set to $p_{\mathrm{T,\,hard}} = 0.8\,\mathrm{GeV}/c$.
In Pythia6 simulations, there is a strong dependence between the charged particle multiplicity generated per event and the number of multiple parton interactions of the event, N_{MPI}. The number of multiple parton interactions N_{MPI} and the correlation between N_{MPI} and N_{ch} are shown in Figure 5.8 for Pythia6 simulations using the Perugia-2011 tune [Ska10].

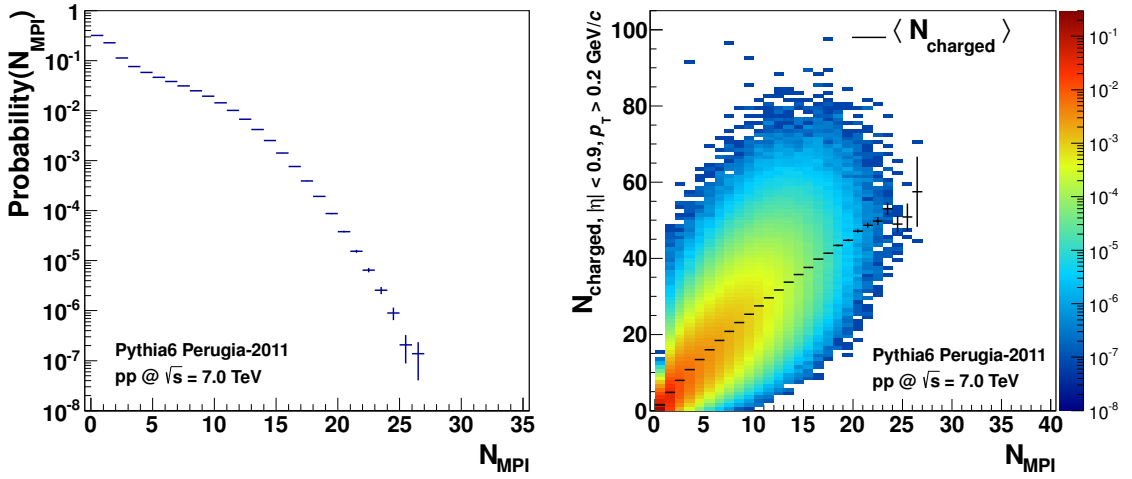

Figure 5.8.: Left panel: Number of multiple parton interactions N_{MPI} in Pythia6 Perugia-2011 simulations. Right panel: Correlation between the number of multiple parton interactions N_{MPI} and the charged particle multiplicity in Pythia6 Perugia-2011 simulations. The multiplicity is given by all charged primary particles measured with $p_{\mathrm{T}} > 0.2\,\mathrm{GeV}/c$ and $|\eta| < 1.0$.

The mean charged particle multiplicity in Pythia6 Perugia-2011 simulations grows steadily with the number of multiple parton interactions. For a single bin in N_{MPI}, however, a large range of charged particle multiplicities is generated. For instance, at $N_{\mathrm{MPI}} = 20$, events with charged particle multiplicities between $15 \lessapprox N_{\mathrm{ch}} \lessapprox 90$ are generated. Extremely high multiplicities are reached only by selecting events with high N_{MPI}, where each MPI has produced a significantly large number of particles. For these high multiplicities, the dependence of N_{MPI} on N_{ch} grows faster than linearly.

5.5.2. Measurement of Number of Multiple Parton Interactions

In this section, the feasibility to access the number of multiple parton interactions by means of the discussed two-particle correlation data analysis is explored.

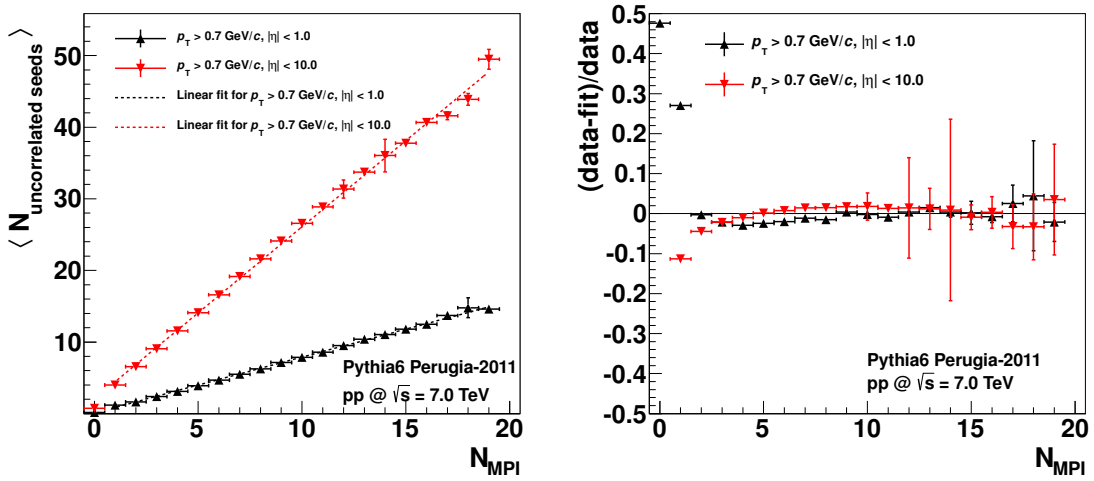

Figure 5.9.: Linear dependence between $N_{\mathrm{uncorrelated\ seeds}}$ and N_{MPI} in Pythia6 Perugia-2011 simulations. Left panel: $N_{\mathrm{uncorrelated\ seeds}}$ versus N_{MPI} combined with a linear fit function. Right panel: Residual of the data and the fit distribtution.

Figure 5.9 shows the number of uncorrelated seeds $N_{\mathrm{uncorrelated\ seeds}}$ measured in Pythia6 Perugia-2011 simulations with a trigger threshold of $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ as a function of the number of multiple parton interactions N_{MPI}. The dependence of $N_{\mathrm{uncorrelated\ seeds}}$ and N_{MPI} is described by a linear function given by the dashed linear fits in the left panel of Figure 5.9. The residuals of the data distribution and the linear fit functions are presented in the right panel of Figure 5.9.

When using particles measured in the inner acceptance of ALICE, $|\eta| < 1.0$, only a small fraction of all produced tracks is taken into account. By enlarging the η-acceptance such that the pseudorapidity of all produced particles is covered ($|\eta| < 10.0$), the correlations between the N_{MPI} and all generated primary particles can be measured. The linear dependence between the $N_{\mathrm{uncorrelated\ seeds}}$ and the N_{MPI} is equally good for both

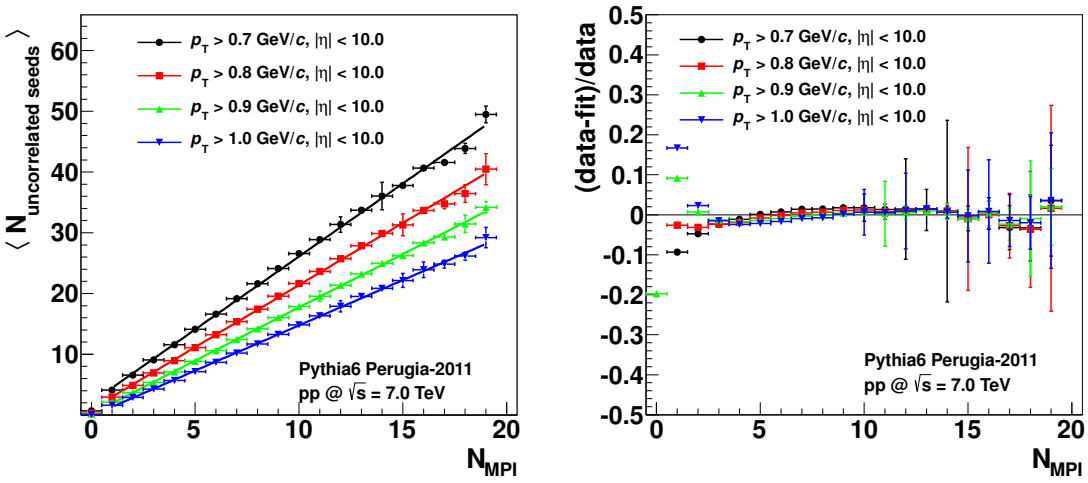

Figure 5.10.: Linear dependence between $N_{\text{uncorrelated seeds}}$ and N_{MPI} in Pythia simulations. Left panel: $N_{\text{uncorrelated seeds}}$ versus N_{MPI} combined with a linear fit function. Right panel: Residual of data and fit distribution. The linear dependence is shown for different $p_{\text{T, trig}}$-cuts.

η-acceptances. The quality of the fit can be tested with the χ^2/NDF test, which gives equal values for both η-acceptances.

Figure 5.10 shows the $N_{\text{uncorrelated seeds}}$ measured for different cuts in $p_{\text{T, trig}}$. The cut in $p_{\text{T, trig}}$ should be close to the scale of the (semi-)hard scattering $p_{\text{T, hard}}$ in order to guarantee an agreement between $N_{\text{uncorrelated seeds}}$ and N_{MPI}. A linear dependence is given for all presented values of $p_{\text{T, trig}}$. With an increasing deviation between $p_{\text{T, trig}}$ and $p_{\text{T, hard}}$, the deviations from a linear dependence of N_{MPI} and $N_{\text{uncorrelated seeds}}$ becomes more significant.

5.5.3. Conclusion

It was shown that for Pythia6 Perugia-2011 simulations the number of uncorrelated seeds, $N_{\text{uncorrelated seeds}}$, measured at $p_{\text{T, trig}} > 0.7\,\text{GeV}/c$ and with different acceptances in pseudorapidity ($|\eta| < 10$ and $|\eta| < 1$) is proportional to the number of multiple parton interactions, N_{MPI}. At low N_{MPI}, however, the scaling slightly deviates from a proportional dependence.

Based on this finding, the $N_{\text{uncorrelated seeds}}$ measured by ALICE in proton-proton collisions can be used as a measure of the number of multiple parton interactions.

5.6. Summary

In this chapter, the two-particle angular correlation analysis has been introduced. One direct observable of the two-particle correlation is the azimuthal correlation between two particles generated in the same proton-proton collision normalized by the number of trigger particles. This distribution is called per-trigger yield as a function of the difference in the azimuthal angle $\Delta\varphi$.
For the selection of the tracks that enter the two-particle correlation, two transverse momentum thresholds are used, $p_{\mathrm{T,\,trig}}$ for the trigger particles and $p_{\mathrm{T,\,assoc}}$ for the associated particles.
The azimuthal correlation reveals a near side peak at $\Delta\varphi = 0$, an away side peak at $\Delta\varphi = \pi$, and a constant background. The near side peak represents particles coming mainly from the fragmentation of jets. The away side peak represents particles from the fragmentation of recoiling jets. The constant background contains particles coming from uncorrelated particle production processes building the combinatorial background of the two-particle correlation. The correlation is measured separately for different charged particle multiplicities and for various transverse momentum thresholds used to define the track sample building the correlation.
A fit function has been introduced for the decomposition of the azimuthal correlation. The fit function allows for a simplified comparison of properties of the azimuthal correlation between various event classes without the need to compare the complete, complex azimuthal distributions to each other. The fit function is based on a combination of Gaussian functions for the near and the away side and a constant for the combinatorial background. With the help of the fit function, five observables can be derived. The per-trigger yield in the combinatorial background, the per-trigger near side yield, and the per-trigger away side yield. In addition, the average number of trigger particles is estimated. These observables are combined to the number of uncorrelated seeds, a measure of the number of uncorrelated sources of particle production.
The impact of combinatorics in auto-correlations at high multiplicities has been studied. It is assumed that the impact of combinatorics on the correlation can be limited, if the transverse momentum thresholds used for the selection of trigger particle in the correlation are chosen to large enough values. A minimum value of $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ is used in the following data analysis. The estimation is based on the probable assumption that the intra-jet multiplicity distribution is a steeply falling function.
Also, the impact of decays of short lived particles on the two-particle correlation has been estimated. In simulations based on the Therminator event generator, it is found that products of particle decays can contaminate the per-trigger yield at the near side. However, the yield coming from these particle decays has been found to be a magnitude smaller that the yield coming from jet fragmentation. A correction for the contribution of tracks from

particle decays to the azimuthal correlation is introduced in Section 6.3. This correction is applied as part of the full correction procedure introduced in Chapter 6.
Based on Pythia simulations, it has been shown that the uncorrelated seeds measured within the ALICE acceptance rises linearly as a function of the number of multiple parton interactions. Hence, the number of uncorrelated seeds is used in the following as a measure of the number of multiple parton interactions.

6. Correction Procedure

In this chapter, a correction procedure for the presented two-particle angular correlation analysis of proton-proton collision data is described. The procedure aims at correcting for all relevant experimental biases on the analysis results.
The correction is structured as follows: The sample of reconstructed and accepted tracks (cf. Section 4.4) needs to be corrected for the contamination of tracks from secondary particles. Furthermore, the inefficiency of the reconstruction of tracks from primary particle needs to be corrected. A data driven contamination correction is applied taking into account the fact that the Monte Carlo (MC) event generators used in the correction procedure significantly underestimate the production of strange particles [ALI11c, ALI11d]. Hence, the MC generators also underestimate the contamination from secondary particles in the reconstructed track samples. Both, the contamination corrections as well as the reconstruction efficiency correction are evaluated depending on the particle properties pseudorapidity and transverse momentum. In addition to these standard track-to-particle corrections based on single particle properties, two-track corrections also need to be applied. These two-track corrections take into account the fact that some effects cause correlated particle yield modifications, e. g. track splitting or merging during the reconstruction and the decay of particles into several correlated secondary particles. The charged particle multiplicity is corrected using an unfolding procedure based on detector response matrices. As the last steps of the correction, the impact of the vertex reconstruction efficiency (cf. Section 4.3.2) and the impact of the minimum bias trigger efficiency (cf. Section 4.3.1) are corrected.

In summary, the correction procedure is subdivided into 7 main steps (0-6). All steps and the modifications between the steps are presented in Table 6.1. On the event level, step 6 corresponds to collision events that are detected by the minimum bias trigger and which have a reconstructed and accepted vertex. On the track level, all reconstructed tracks are considered that pass the track selection cuts which are optimized for the selection of tracks from primary particles. Step 0 corresponds to events of the $\mathrm{INEL} > 0_{|\eta|<0.9}$ event class and primary charged particles. Each correction step is introduced in the following sections. The impact of each correction step on the event and track (particle) distribution is evaluated. For MC data with full detector simulation, all correction steps are computed. For MC data without detector simulation, only step 0 is computed. For real collision data, all correction steps are computed, starting with the measured results in step 6 and correcting each bias backwards in step 5 to step 0. All correction steps except the step 6b are based on correction maps obtained by MC data with full detector simulations. However, the detector response used in the MC events is mod-

Step	Events			Tracks
	Trigger	Vertex	Multiplicity	
0	all	all	N_{prim}	Prim
1	**off-line trigger**	all	N_{prim}	Prim
2	off-line trigger	**rec. vertex**	N_{prim}	Prim
3	off-line trigger	rec. vertex	$N_{\mathbf{rec}}$	Prim
3b				**Two-track+detector**
4	off-line trigger	rec. vertex	N_{rec}	$\mathbf{Prim_{rec}}$
5	off-line trigger	rec. vertex	N_{rec}	$\mathbf{(Prim + Sec)_{rec}}$
6	off-line trigger	rec. vertex	N_{rec}	$\mathbf{Tracks_{rec}}$
6b			**Strangeness**	**Strangeness**

Table 6.1.: Correction steps used in the correction procedure of the two-particle correlation analysis. Measured data is corrected from step 6 back to step 0. In each correction step, one more efficiency is considered. The changes between step N to step N+1 are indicated in bold font. ("rec"= reconstructed, "prim"=primary particles, "sec"=secondary particles, "strangeness" = data driven strangeness correction.)

eled using the real detector configuration. Step 6b is computed as a data driven correction.

For an infinite number of MC events, the correction from step N to the next step $N-1$ is the ratio of the results of the two steps. Using only a limited number of MC events, poorly populated regions of the correction maps, e. g. a track map at high transverse momenta, need to be filled with values extrapolated from fully populated regions, e. g. intermediate transverse momenta.
The Monte Carlo based correction procedure is validated by comparing full detector simulation results after correction to the corresponding true Monte Carlo data distributions (cf. Section 7.2).

6.1. Correction of Contamination From Secondary Particles

6.1.1. Correction Based on Full Detector Simulations

The track contamination describes the fraction of all reconstructed and accepted tracks that is produced by secondary particles ("secondaries"). Whether tracks are accepted or not is based on the track selection described in Section 4.4. Accepted tracks coming from secondary particles are mostly produced by particle decays of hadrons containing strange

quarks, conversions of photons, and hadronic interactions in the detector material. The amount of the contamination in the track sample depends predominantly on the track properties transverse momentum p_{T} (and the pseudorapidity η). The contamination of the track sample as well as the inverse value, the contamination correction, are estimated using full detector simulations of collision events from Monte Carlo generators. The contamination correction factor is defined as the ratio of the number of reconstructed tracks coming from primary particles to the number of all reconstructed tracks. The contamination correction estimated with Pythia6 Perugia-0 simulations at the collision energies $\sqrt{s} = 0.9\,\mathrm{TeV}$, $2.76\,\mathrm{TeV}$ and $7.0\,\mathrm{TeV}$ are presented in Figure 6.1 as projections of the 2-dimensional $\eta - p_{\mathrm{T}}$-correction map on the η-axis and the p_{T}-axis.

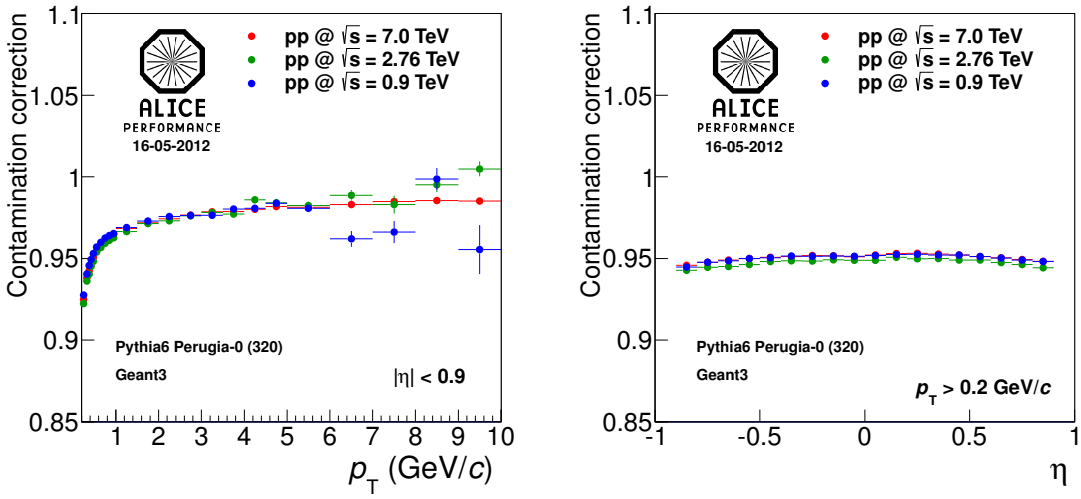

Figure 6.1.: Contamination correction [Sic12]. Left panel: Contamination correction versus track transverse momentum ($|\eta| < 0.9$). Right panel: Contamination correction versus track pseudorapidity ($p_{\mathrm{T}} > 0.2$ GeV/c). Both figures are projections of the 2-dimensional $\eta - p_{\mathrm{T}}$-correction map. The values are estimated with a full detector simulation of Pythia6 Perugia-0 events at $\sqrt{s} = 0.9\,\mathrm{TeV}$, $2.76\,\mathrm{TeV}$ and $7.0\,\mathrm{TeV}$ using the detector configurations corresponding to the ALICE data taking period LHC10c, LHC10b-c-d-e, and LHC11a (cf. Chapter 4). Only statistical uncertainties are presented. The uncertainties in the x-direction represent the bin width. Systematic uncertainties are discussed in Section 7.

The contamination estimated with Pythia6 Perugia-0 simulations is higher at low transverse momenta as compared to high transverse momenta. The contamination is almost constant for all values of pseudorapidity.

Having only a limited number of simulated data at high transverse momenta ($p_{\mathrm{T}} > 5$ GeV/c) the number of generated particles is too low for a precise computation of the contamination correction. The resulting correction would be dominated by fluctuations. In order to avoid an influence of these fluctuations on the correction, contamination

correction values at lower transverse momentum values are fitted and extrapolated to higher transverse momenta. The extrapolated values are then used for the correction at high transverse momenta.
Integrated over all transverse momenta and pseudorapidities, the contamination correction for the presented phase space is approximately 6%. This is the second largest contribution to the total track-to-particle correction.
The detector simulation used to estimate the corrections maps is performed using detector configurations corresponding to the ALICE data taking periods LHC10c (0.9 TeV), LHC11a (2.76 TeV) and LHC10b -c -d -e (7.0 TeV) (cf. Chapter 4). These detector configurations are used for all following correction steps.
Figure 6.1 and all other figures in this chapter include only statistical uncertainties. The systematic uncertainties of the corrections are discussed in Section 7. The uncertainties in the x-direction represent in all figures the bin width.

6.1.2. Data Driven Strangeness Correction

ALICE has measured the production of hadrons containing strange quarks at central rapidity in proton-proton collision at $\sqrt{s} = 0.9\,\text{TeV}$ [ALI11c, ALI11d]. It is shown that the strange particle yields are strongly underestimated by the tested Monte Carlo generators including the generators Phojet and Pythia6 Perugia-0.
The underestimated production of particles with strange quarks in the Monte Carlo generators leads to an insufficient correction of the contamination in ALICE data when using correction maps of these event generators. The discrepancy between data and MC contamination is corrected with a data driven approach.

As an example, Figure 6.2 presents the K^0_S yield in proton-proton collisions at $\sqrt{s} = 900\,\text{GeV}$ measured by ALICE [ALI11d] in comparison to MC predictions. Here, the yield predicted by Pythia6 Perugia-0 is much lower as in the ALICE data. The contamination from secondaries in Pythia6 Perugia-0 correction maps would be underestimated. The underestimation of the contamination in the Pythia6 Perugia-0 correction maps is corrected by weighting all secondaries from strange mothers by the factors by which their production is underestimated. The values are extracted from Figure 6.2 and the corresponding figures for other hadrons with strange quarks (plots not shown) [ALI11c, ALI11d]. The same values are applied for all center-of-mass energies. The uncertainty of the extrapolation to higher center-of-mass energies is taken into account in the systematic studies described in Chapter 7.

Figure 6.3 shows the data driven correction for contamination from strange particles. The correction values show a dependence on the transverse momentum of the particles. The correction value is almost constant as a function of the pseudorapidity.

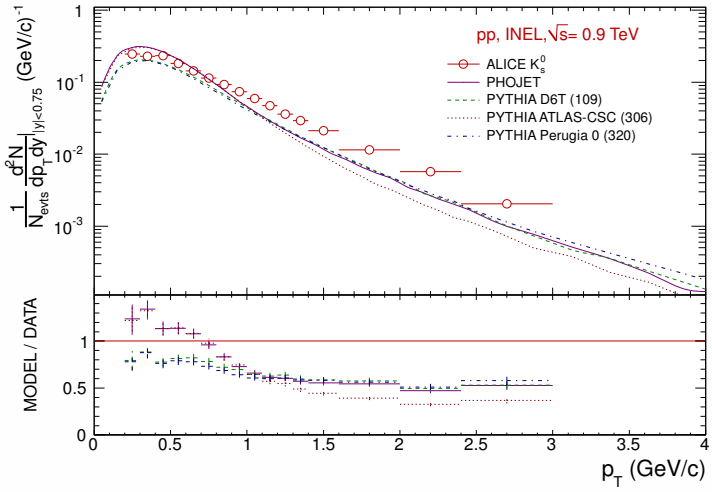

Figure 6.2.: Comparison of the transverse momentum differential yield for the K^0_S particles for inelastic proton-proton collisions at $\sqrt{s} = 0.9\,\text{TeV}$ with Phojet and Pythia6 tunes 109, 306 and 320 [ALI11d].

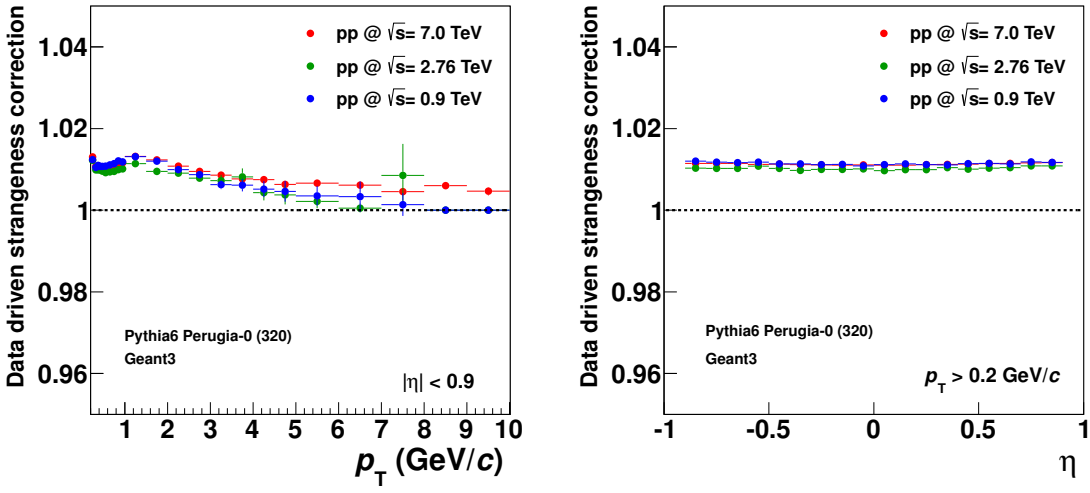

Figure 6.3.: Data driven strangeness correction applied on Pythia6 Perugia-0 events. Left panel: Data driven strangeness correction versus track transverse momentum ($|\eta| < 0.9$). Right panel: Data driven strangeness correction versus track pseudorapidity ($p_\text{T} > 0.2$ GeV/c). Both figures are projections of the 2-dimensional $\eta - p_\text{T}$-correction map. The values are estimated with a full detector simulation of Pythia6 Perugia-0 events at $\sqrt{s} = 0.9\,\text{TeV}$, 2.76 TeV and 7.0 TeV.

The correction is applied when correcting ALICE data. The correction weights are used only as a function of the transverse momentum of the particles. For high transverse momenta, extrapolated values from intermediate transverse momenta are applied.
Integrated over all transverse momenta and pseudorapidities, the data driven strangeness

correction for the presented phase space is approximately 1% while the standard contamination correction has an integrated value of approximately 6 %. The data driven contamination correction represents the third largest contribution to the total track-to-particle correction.
The correction needs to be applied as single track correction. Also, a strangeness correction needs to be applied during the correction of the charged particle multiplicity. This correction is discussed in Section 6.4.

6.2. Track Reconstruction Efficiency Correction

The track reconstruction efficiency characterizes the fraction of all primary particles that is found during the reconstruction and that also passes the ITS-TPC track selection cuts described in Section 4.4. The track reconstruction efficiency depends strongly on the track properties transverse momentum p_{T} and pseudorapidity η. The reconstruction efficiencies computed in Pythia6 Perugia-0 simulations are presented in Figure 6.4.

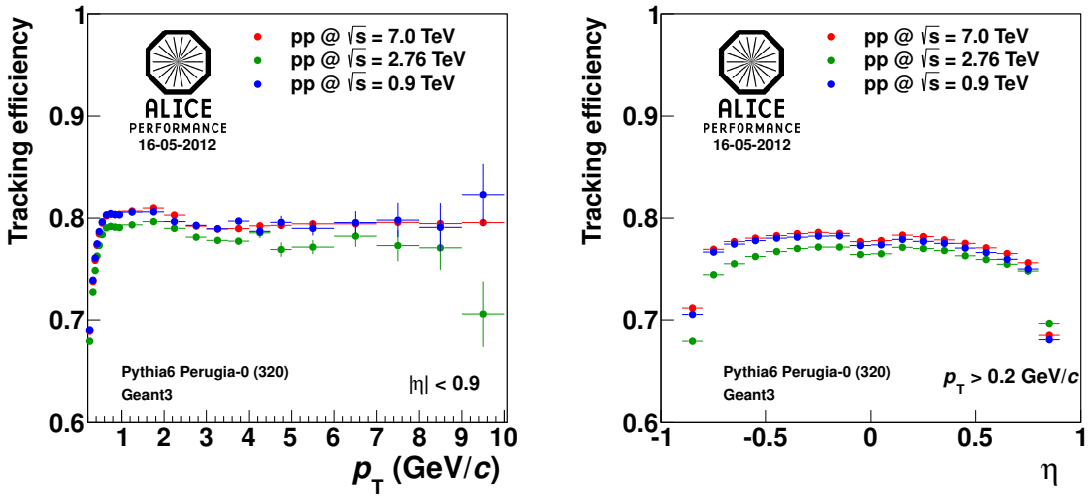

Figure 6.4.: Reconstruction efficiency for primary particles. Left panel: Reconstruction efficiency versus particle transverse momentum ($|\eta| < 0.9$). Right panel: Reconstruction efficiency versus particle pseudorapidity ($p_{\mathrm{T}} > 0.2$ GeV/c). Both figures are projections of the 2-dimensional $\eta - p_{\mathrm{T}}$-correction map. The values are estimated with a full detector simulation of Pythia6 Perugia-0 events at $\sqrt{s} = 0.9$ TeV, 2.76 TeV and 7.0 TeV [Sic12].

Whether a track of a particle is reconstructed or not depends on the length of the track within the TPC. As a first approximation, the TPC track length scales with the number of crossed pad rows of the TPC read-out chambers. For tracks above a p_{T}-threshold, the

number of crossed pad rows is correlated to the number of clusters which are associated to the track. Either a minimum number of crossed pad rows or a minimum number of clusters needs to be associated to a track to be accepted and considered in the analysis. Charged tracks are bent in the magnetic field. A minimum p_{T} of 150 MeV/c is required to reach the TPC. With increasing transverse momentum, the track length within the TPC increases for tracks around mid-rapidity. These tracks are more likely to be reconstructed and accepted by the track quality cuts. Therefore, an increase of the track reconstruction efficiency with increasing transverse momenta p_{T} is observed in Figure 6.4.
At $p_{\mathrm{T}} = 3$ GeV/c, the tracks are almost straight. A fraction of these tracks might fall into blind areas between the TPC sectors. The TPC subdivided in 18 sectors in the azimuthal angle. Between each sector, an angle of $\varphi = 2°$ is not covered by sensitive detector material. In total, 10 % of the full azimuth is not covered. The reconstruction efficiency decreases again for high momenta tracks falling into these blind zones.
At central pseudorapidity, more tracks are lost as compared to intermediate $|\eta|$. This is caused by the central membrane of the TPC, which is placed at mid-rapidity. Tracks that pass the membrane are more likely to be stopped compared to those only traversing the gas in the TPC due to interaction with comparably dense material. For these trajectories through the central membrane, also the track length in the sensitive detector area of the TPC is reduced by the fraction of the path within the central membrane.
For the data set measured at $\sqrt{s} = 2.76$ TeV, an increased number of sensors mounted on the half-staves of the Silicon Pixel Detector (SPD) is switched off. Therefore, the reconstruction efficiency is lower for this data set than in the data sets measured at $\sqrt{s} = 0.9$ and 7.0 TeV.
As pointed out in Section 6.1, the estimated tracking efficiency correction calculated by simulations is, at high transverse momenta, limited by the sample size. In order to avoid an influence of statistical fluctuations on the correction, track reconstruction efficiencies at intermediate transverse momenta are fitted and extrapolated to higher transverse momenta.
Integrated over all transverse momenta and pseudorapidities, the correction of the tracking efficiency is approximately 22%. The tracking efficiency correction represents the largest contribution to the total track-to-particle correction.

6.3. Correction of Two-Track Effects and Detector Effects

Two-track effects such as track splitting, track merging, the decay of particles, hadronic interactions with the material, and gamma conversion (discussed in [ALI11f] and [ALI11a]) must be taken into account. Also, detector effects such as gaps in the φ-acceptance which result in a none-flat $\Delta\varphi$-distribution of reconstructed tracks need to be corrected. In this section, the correction of such two-particle and detector effects is discussed.

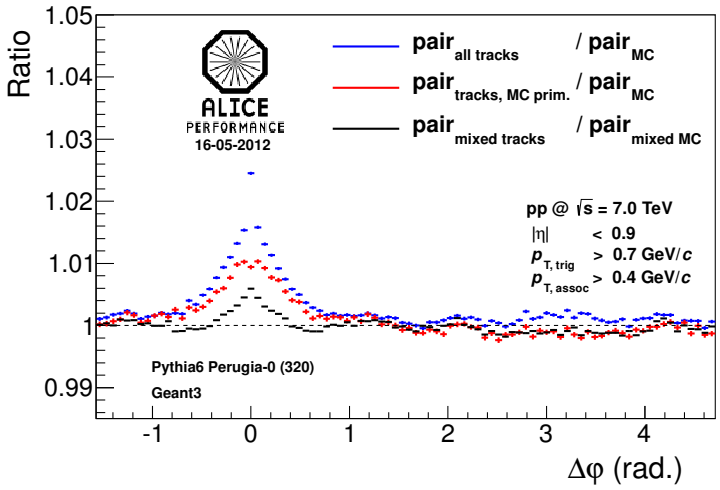

Figure 6.5.: Ratio between the track pair distribution of reconstructed and corrected tracks using single track corrections and the pair distribution of MC primary particles ($p_{\mathrm{T,\ trig}} > 0.7$ GeV/c and $p_{\mathrm{T,\ assoc}} > 0.4$ GeV/c at central rapidity) as a function of the difference in azimuthal angle $\varphi_{\mathrm{track1}} - \varphi_{\mathrm{track2}} = \Delta\varphi$. The simulations have been performed for Pythia6 Perugia-0 events at $\sqrt{s}$ = 7.0 TeV [Sic12].

Figure 6.5 shows the ratio $\mathrm{pair}_{\mathrm{corrected\ tracks}}/\mathrm{pair}_{\mathrm{MC\ particle}}$ for a full detector simulation of Pythia6 Perugia-0 events at $\sqrt{s}$ = 7.0 TeV. The numerator is given by the $\Delta\varphi$-distribution of reconstructed tracks after the single track correction of contamination and reconstruction efficiency. The denominator is given by the $\Delta\varphi$-distribution of the corresponding Monte Carlo particles. The ratio is presented for all reconstructed tracks, for reconstructed tracks of primary particles only, and for mixed tracks of different events.
For tracks pairs with an azimuthal angle close to each other $\varphi_{\mathrm{track\,1}} - \varphi_{\mathrm{track\,2}} = \Delta\varphi \approx 0$ and tracks with opposite azimuthal angles, $\Delta\varphi \approx \pi$, more pairs are found in the reconstructed and single track corrected sample than in the Monte Carlo primary sample. The ratio is slightly above 1.
The enhanced number of pairs is due to several contributions. One contribution is given by tracks coming from decay products. Particles may decay within the detector and produce daughter particles. The daughter particles can be secondary or primary particles depending on the life time of the mother. The primary daughter particles are accepted in the analysis. Also, a small fraction of these secondary particles can pass the track cuts. The azimuthal angles of the daughter particles of these processes are correlated to each other. If the mother particles have transverse momenta larger than zero, the daughter particles are boosted in the same direction and cause particle correlations at $\Delta\varphi \approx 0$. If the mother particles decay at rest, they cause the daughter particles to travel in opposite direction, $\Delta\varphi \approx \pi$.

Following the same principle, gamma conversions can cause correlated secondary particles. Hadronic interactions of particles with the detector material can cause particle correlations at $\Delta\varphi \approx 0$ as particle from the detector material can be kicked out of the material in the direction of the causing particle.
Another contribution is given by split tracks related to an imperfect reconstruction procedure. It has to be noted, that the effect of track merging can partially compensated the track splitting effect.

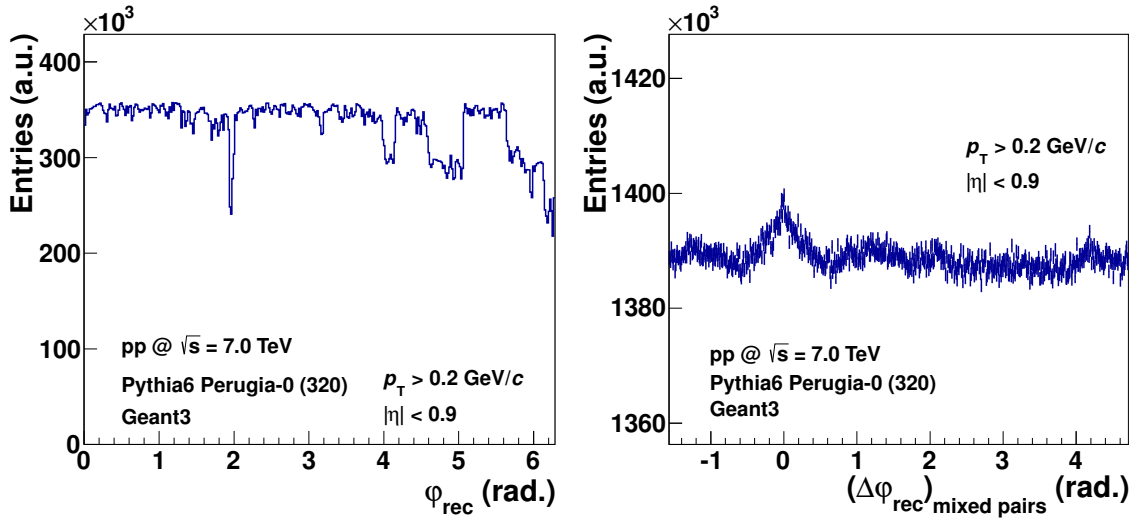

Figure 6.6.: Right panel: Distribution of the reconstructed azimuthal angle φ for Pythia6 Perugia-0 simulations at $\sqrt{s} = 7.0\,\text{TeV}$. Due to switched-off SPD modules, the reconstruction efficiency is lower at various φ-values. Right panel: Mixed pair distribution as a function of $\Delta\varphi$ generated by folding the reconstructed φ distributions presented in the left panel with itself. Small structures appear in the $\Delta\varphi$ distribution due to the non-flat φ distribution. Note that the 0 of the y-axis is suppressed for a better visibility of the small effect.

The last contribution is caused by a non-flat φ-acceptance resulting in a non-flat $\Delta\varphi$-distribution: The left panel of Figure 6.6 shows the reconstructed φ-distribution. While the φ of Monte Carlo primary particles is flat (not shown), the reconstructed φ-acceptance shows dips at $\varphi = 2, 4, 5$, and 6. This is due to a reduced reconstruction efficiency in these regions due to several switched-off SPD modules. By folding the φ distribution with itself, the $\Delta\varphi$ distribution is obtained (right panel of Figure 6.6). Due to the non-flat φ-acceptance, small structures in $\Delta\varphi$ are generated, too.
In order to correct all three two-track and detector effects, an additional $\Delta\varphi$-correction step needs to be performed. The correction taking into account all three effects is below 2.6 % for all track pair bins in $\Delta\varphi$ shown in Figure 6.5. The correction integrated over all $\Delta\varphi$-bins for the presented selection cuts in transverse momentum and pseudorapidity has a value of

0.5 %. This correction is very small in comparison to the correction of the contamination and the tracking efficiency. Of all relevant corrections, the two-track correction presents the smallest contribution to the total track-to-particle correction. With increasing transverse momentum, the two-track and detector effect correction factor decreases even more (plot not shown).

6.4. Charged Particle Multiplicity Correction

In the previous sections, the correction of the presented two-particle correlation observables is performed on a track basis. In the final presentation of the analysis results, however, the observables are presented as a function of the charged particle multiplicity[1]. This section describes the correction of the multiplicity on an event basis as already demonstrated in [Lüt10, ALI10c].
In the following, the two-particle correlation observable is called $O(N_{\mathrm{ch}})$. For example, an observable is the per-trigger near side yield $\langle N_{\mathrm{assoc,\,near\ side}}\rangle(N_{\mathrm{ch}})$. At this part of the correction chain, O is already track-wise corrected, $O = O_{\mathrm{corr}}$. When correcting the multiplicity from N_{rec} to N_{corr}, also the correction from $O(N_{\mathrm{rec}})$ to $O(N_{\mathrm{corr}}) \equiv O(N_{\mathrm{prim}})$ is performed. This correction step is achieved by a weighting procedure of O with a correlation matrix between the number of reconstructed tracks and the number of primary particles, $R(N_{\mathrm{prim}}, N_{\mathrm{rec}})$. The correlation matrix represents the relation between the true multiplicity and the reconstructed multiplicity (defined by the track cuts introduced in Section 4.4) obtained from Monte Carlo simulations. Figure 6.7 shows such correlation matrices determined for the event generator Pythia6 Perugia-0 at the center-of-mass energies $\sqrt{s}$ = 0.9, 2.76 and 7.0 TeV.

Figure 6.7 shows that the number of reconstructed tracks increases in average linearly with the number of primary charged particles. The scaling factor for all three center-of-mass energies is about $N_{\mathrm{rec}}/N_{\mathrm{prim}} = 0.8$. Two competing effects affect the number of the reconstructed tracks: (1) Due to the limited reconstruction efficiency, the number of the reconstructed tracks is in average lower as compared to the value of the true primary multiplicity. (2) The contamination of the reconstructed tracks by tracks coming from secondary particles increases the average number of reconstructed tracks slightly.
One bin in charged primary particle multiplicity contributes to a wide range of reconstructed track multiplicities. The widths of the $P_{N_{\mathrm{prim}}}(N_{\mathrm{rec}})$-distribution per true multiplicity bin are on average described by Poisson distributions.

[1] In the following, the charged particle multiplicity is call multiplicity as an abbreviation.

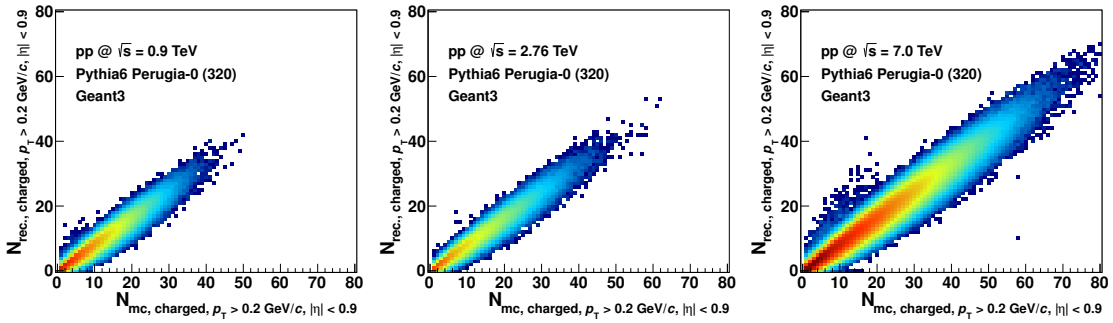

Figure 6.7.: Correlation matrix between the number of reconstructed tracks and the number of primary particles. Left panel: Correction matrix at $\sqrt{s} = 0.9\,\text{TeV}$. Mid panel: Correction matrix at $\sqrt{s} = 2.76\,\text{TeV}$. Right panel: Correction matrix at $\sqrt{s} = 7.0\,\text{TeV}$. The detector response is estimated by a full detector simulation of Pythia6 Perugia-0. Note that different numbers of input events are available for the three center-of-mass energies.

6.4.1. Normalized and Extrapolated Correlation Matrix

As a first step of the correction, the correlation matrix is normalized. For a given column in N_{prim}, the sum of all bins in N_{rec} is normalized to 1 according to

$$\forall N_{\text{prim}} : \sum_{N_{\text{rec}}} R_1(N_{\text{prim}}, N_{\text{rec}}) = 1. \tag{6.1}$$

The resulting matrix $R_1(N_{\text{prim}}, N_{\text{rec}})$ contains the probability distribution describing that an event with the primary charged particle multiplicity N_{prim} is measured at an reconstructed track multiplicity N_{rec}. A normalized correlation matrix is presented in the left panel of Figure 6.9.

Due to the limited number of events, only few entries are available for high multiplicities. For an unbiased correction also at high multiplicities, the correlation matrix needs to be extended. This is done by fitting the distribution $N_{\text{prim}}(N_{\text{rec}})$ for each bin in N_{rec} (or vice versa). The resulting distributions for each bin in N_{rec} can be described by Gaussian functions. The mean value and the width of the Gaussian functions are extracted using a fit function. The width scales as $\sigma \propto \sqrt{N}$, hence $\sigma^2 \propto N$. For a full detector simulation of Pythia6 Perugia-0 events at $\sqrt{s} = 7.0\,\text{TeV}$, the mean value $\langle N_{\text{prim}} \rangle$ and the standard deviation of the Gaussian functions $\sigma_{N_{\text{prim}}}$ and $\sigma^2_{N_{\text{prim}}}$ together with a linear fit as a function of the reconstructed multiplicity are shown in Figure 6.8.

In a further step, the mean value $\langle N_{\text{prim}} \rangle(N_{\text{rec}})$ and the standard deviation of the Gaussian functions $\sigma_{N_{\text{prim}}}(N_{\text{rec}})$ can be parametrized, e. g. using linear functions for $\langle N_{\text{prim}} \rangle(N_{\text{rec}})$ and $\sigma^2_{N_{\text{prim}}}(N_{\text{rec}})$. The resulting parametrizations are then used to generate extrapolated Gaussian functions. These are used as input for the extension of the normalized correlation

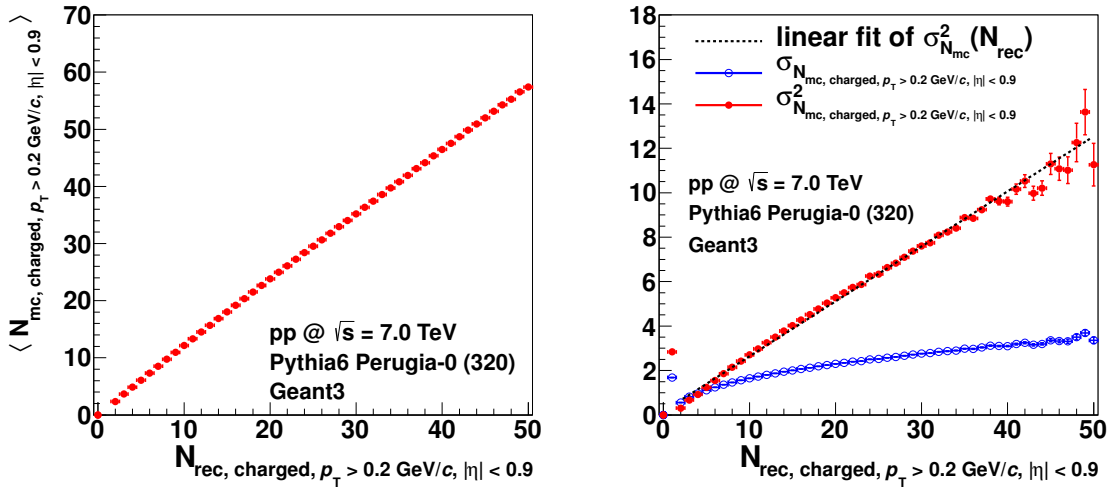

Figure 6.8.: Properties of the correlation matrix presented in the right panel of Figure 6.7: Left panel: Mean value of the true multiplicity per reconstructed multiplicity bin. Right panel: σ and σ^2 of the true multiplicity distribution per reconstructed multiplicity bin. The detector response matrix is estimated in a full detector simulation of Pythia6 Perugia-0 events at $\sqrt{s} = 7.0\,\text{TeV}$.

matrix at high multiplicities. An extended, normalized correlation matrix is presented in the right panel of Figure 6.9.

6.4.2. Weighting Procedure using an Extrapolated Correlation Matrix

The transformation of the observable distribution $O(N_{\text{rec}})$ into $O(N_{\text{corr}})$ is performed using the following equation:

$$O(N_{\text{corr}}) = \sum_{N_{\text{rec}}} O(N_{\text{rec}}) \cdot R_1(N_{\text{prim}}, N_{\text{rec}}) \equiv O(N_{\text{prim}}). \tag{6.2}$$

The O for a given true multiplicity bin N_{prim} is therefore given by a weighted sum of $O(N_{\text{rec}})$ for all bins in the reconstructed multiplicity N_{rec} which correspond to N_{prim}. This results in the corrected $O(N_{\text{corr}}) \equiv O(N_{\text{prim}})$.

Here, it has to be pointed out, that the results after multiplicity correction $O(N_{\text{corr}})$ is always a weighted combination of results of several reconstructed multiplicity bins. The observables measured in a fixed reconstructed multiplicity $O(N_{\text{rec}})$ bin contribute to a large range of corrected multiplicity bins. Hence, the observables of neighboring corrected-multiplicity bins are correlated with each other as they contain information from the same reconstructed multiplicity bins. This correlation of the data points of neighboring multiplicity bins is also the reason why the corrected results are more smooth as a function of the multiplicity as the uncorrected data points.

For the correction of ALICE data, the multiplicity correction needs to include also a

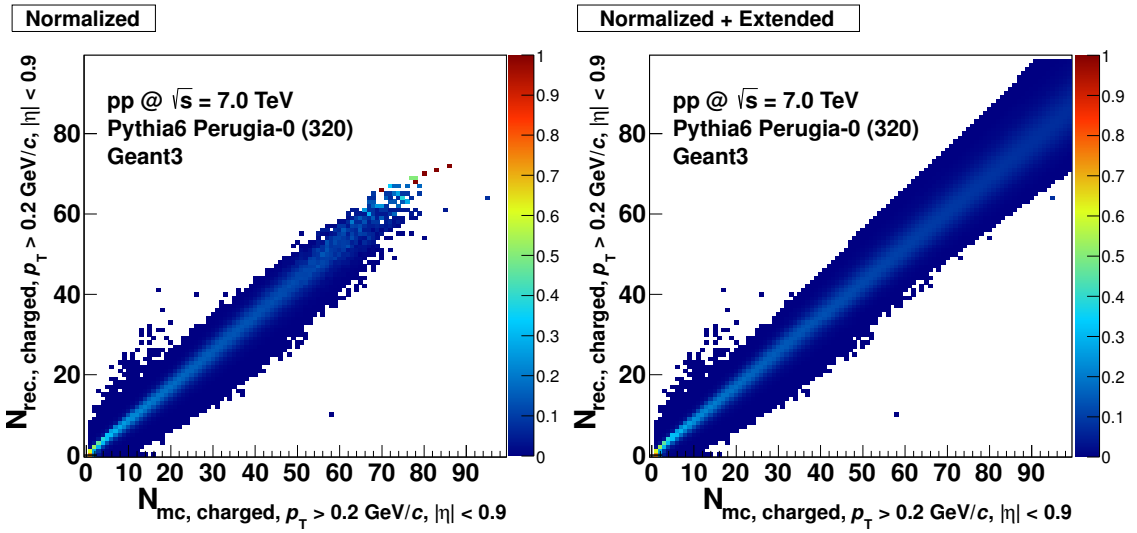

Figure 6.9.: Normalization and extension of a correlation matrix. The original distribution is presented in the right panel of Figure 6.7: Left panel: Correlation matrix with normalized rows. Right panel: Normalized and extended correlation matrix. The extension is performed for all multiplicities above a threshold which depends on the number of simulated events. Here, the threshold is chosen to $N_{\mathrm{ch}} = 50$. Input for the extension are Gaussian distributions with extrapolated $\langle N_{\mathrm{prim}} \rangle$ and $\sigma_{\langle N_{\mathrm{prim}} \rangle}$ from Figure 6.8.

data driven strangeness correction as introduced in Section 6.1.2. Therefore, we do not use the standard correlation matrix of $R(N_{\mathrm{prim}}, N_{\mathrm{rec}})$ for the correction, but the $R(N_{\mathrm{prim}}, N_{\mathrm{rec+strange}})$.

6.5. Vertex Reconstruction Efficiency Correction

The vertex reconstruction efficiency describes the difference in the number of all triggered events compared to the number of triggered events with a reconstructed and accepted vertex. In the analysis, only those events are accepted which have only a single reconstructed primary vertex that fulfills the quality selection cuts as described in Chapter 4.3. The efficiency is evaluated using a full detector simulation of simulated proton-proton events using the event generators Pythia6 Perugia-0 and Phojet. Figure 6.10 shows the resulting vertex reconstruction efficiencies for $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV.

For low multiplicity events ($N_{\mathrm{ch}} = 1$), the vertex reconstruction efficiency is lowest and has a minimum value of 72 % to 80 % depending on the collision energy. At high multiplicities above $N_{\mathrm{ch}} > 10$, the efficiency is consistent with unity. The vertex reconstruction efficiency estimated in Phojet simulations and Pythia6 Perugia-0 simulations are in good overall agreement (right panel of Figure 6.10). The vertex reconstruction efficiency represents the largest event correction.

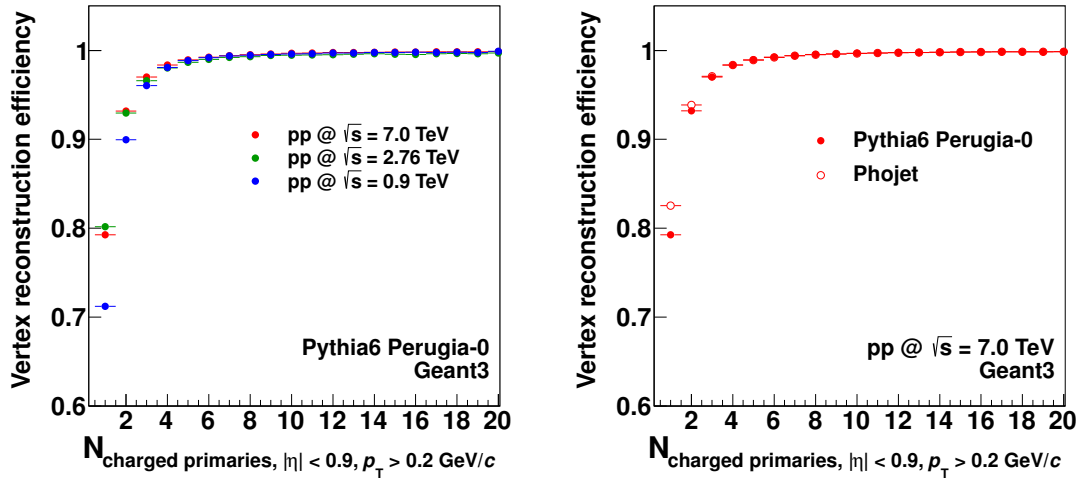

Figure 6.10.: Left panel: The vertex reconstruction efficiency for $\sqrt{s} = 0.9$, 2.76, and 7.0 TeV estimated with Pythia6 Perugia-0. Right panel: Comparison of the vertex reconstruction efficiency estimated with Pythia6 Perugia-0 and Phojet $\sqrt{s} = 7.0$ TeV.

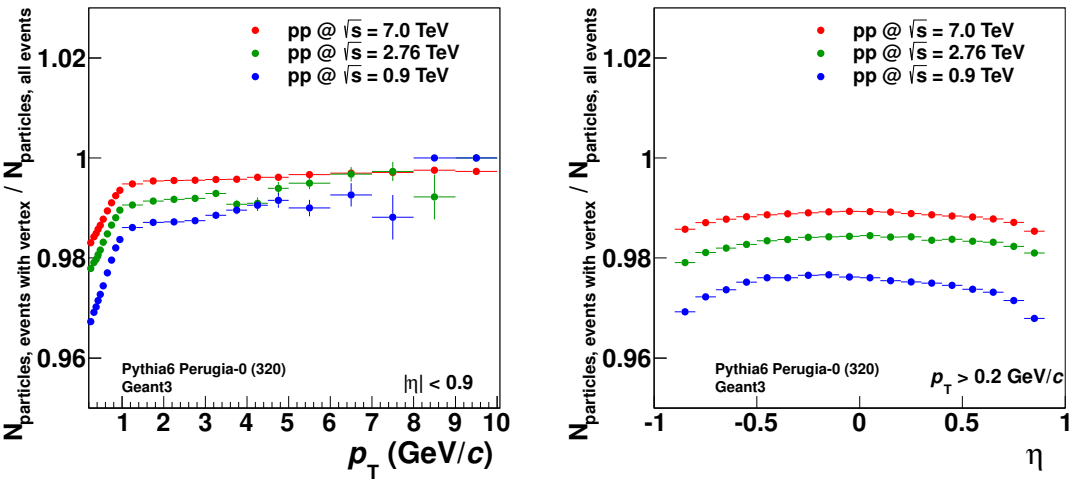

Figure 6.11.: Ratio between the number of particles of all triggered events and the number of particles of triggered events with a reconstructed vertex for $\sqrt{s} = 0.9$, 2.76, and 7.0 TeV estimated with Pythia6 Perugia-0. The distributions are integrated over all charged particle multiplicities ($N_{\mathrm{ch}} > 0$). Left panel: Transverse momentum dependence. Right panel: Pseudorapidity dependence.

The difference between the event samples collected at different center-of-mass energies has two reasons. On one hand, the transverse momentum spectra obtained at the different energies are slightly different. The higher the collision energy, the harder is the transverse momentum spectrum. On the other hand, the SPD suffered from increasing damage over time which results in a decreasing SPD vertex reconstruction probability. The 0.9 TeV-

data have been recorded in May 2010, the 7.0 TeV-data have been recorded between April and August 2010, and the 2.76 TeV-data have been recorded in March 2011.
The vertex reconstruction efficiency also has an impact on the total number of particles entering the data sample. Figure 6.11 illustrates the bias introduced by the vertex reconstruction efficiency in a ratio between the number of particles of triggered events with reconstructed vertices and the number of particles of all triggered events. The ratio is presented as a function of the transverse momentum and the pseudorapidity. Integrated over all transverse momenta and pseudorapidities, the particle related correction of the vertex reconstruction efficiency is 1 % to 3 % depending on the collision energy. From Figure 6.11 it can be seen that, due to the vertex reconstruction efficiency, predominantly events are selected which contain charged particles at high transverse momenta as well as at central rapidity. The effect increases with decreasing collision energy.

6.6. Trigger Efficiency Correction

The correction of the minimum bias trigger efficiency takes into account the fact that the number of triggered events is only a subset of the events of the INEL $> 0_{|\eta|<0.9}$ event class. The trigger bias needs to be corrected. The trigger efficiency can be estimated using a full detector simulation of MC events. Figure 6.12 shows the minimum bias trigger efficiency for Pythia6 Perugia-0 and Phojet simulations at $\sqrt{s} = 0.9$, 2.76, and 7.0 TeV as a function of the charged particle multiplicity.

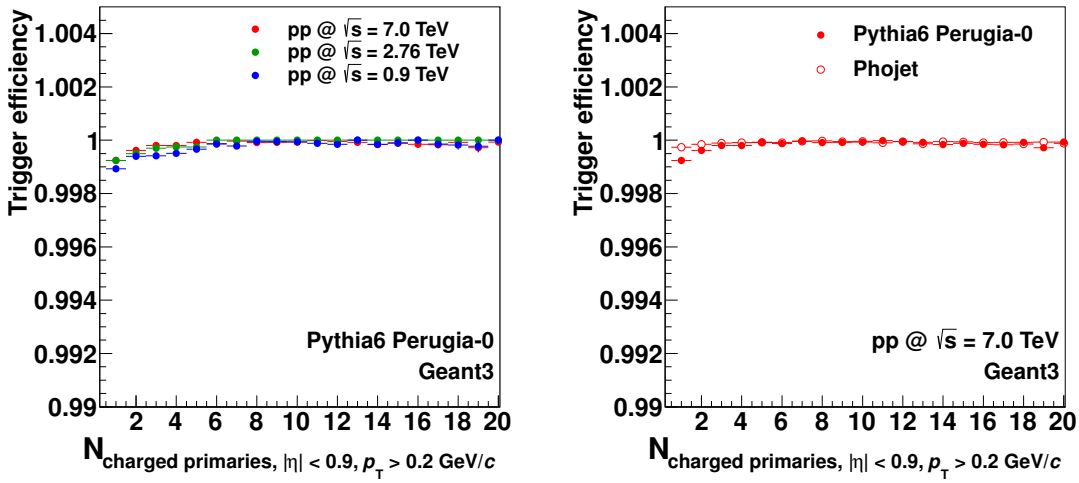

Figure 6.12.: Left panel: Trigger efficiency for $\sqrt{s} = 0.9$, 2.76 and 7.0 TeV estimated in a full detector simulation of Pythia6 Perugia-0 events. Right panel: Trigger efficiency for $\sqrt{s} = 7.0$ TeV estimated in a full detector simulation of Pythia6 Perugia-0 and Phojet events.

Figure 6.12 shows that the trigger bias only has an impact on low multiplicity events. The efficiency approaches unity with increasing multiplicity. The difference between triggered and complete INEL $> 0_{|\eta|<0.9}$ event sample is below 0.5 % for the first multiplicity bin at $\sqrt{s} = 900$ GeV and below 0.1 % at $\sqrt{s}$ =7.0 TeV.
The impact on the track distribution visible in the ratio $N_{\text{prim, triggered events}}/N_{\text{prim, all events}}$ is also very small and negligible (plot not shown). In comparison to the systematic uncertainties, which are introduced in the following Chapter 7, the impact of the trigger efficiency on the final correlation results is negligible.

6.7. Additional Corrections

6.7.1. Pileup Events

At high luminosities, the data sample of the triggered collision events with accepted vertex can be polluted by so-called pileup events, that are not detected by the pileup rejection implemented in the event selection framework. Pileup events are events in which more than one pair of protons collide with each other. Since the pileup probability in ALICE was low during the time of data taking of the analyzed data sets (visible in interaction probability values μ listed in Section 4.1), no correction for the contamination of pileup events is applied in the following. Nevertheless, the influence of pileup events is estimated by analyzing a high-pileup data samples and a low-pileup data sample. The impact of the pile-up events is accounted for in terms of a systematic uncertainty of the data analysis results. This study is presented as part of the study of the systematic uncertainties in Section 7.10.

6.7.2. Beam-Gas and Beam-Halo Events

The sample of the triggered events with accepted vertex can be contaminated by so-called beam-gas or beam-halo events in which the protons do not collide with each other but with other particles within the beam pipe.
It is shown in [ALI10a], that a substantial fraction of the beam-gas events can be rejected by the offline event selection procedure based on signals measured with the VZERO detector (introduced in Section 3.2). This selection is applied in the data analysis. The influence of beam-gas and beam-halo events after the offline event selection is found to be negligible for the analyzed data samples [ALI11f, ALI10c, ALI12e, ALI12g].

6.7.3. Cosmic Rays

Tracks of cosmic rays result in two-particle angular correlation structures with a sharp peak in the azimuthal angle distribution at $\Delta\varphi = \pi$. No peak in this $\Delta\varphi$-region has been

observed after the track cut selection in neither of the data sets.
Also other studies within the ALICE collaboration ([ALI10a, ALI10c, ALI12g]) have shown that the contamination of the track samples used in the data analysis by tracks coming from cosmic rays can be suppressed by the simple track selection cuts. The remaining contamination after the track selection cuts is found to be negligible.

6.8. Summary

In this chapter, a correction procedure for the presented two-particle correlation analysis has been introduced. The correction takes into account corrections on a track basis and on an event basis. The track correction includes the correction for the contamination of the track sample with tracks from secondary particles, the limited tracking efficiency of primary particles, as well as two-track and detector effects. The correction on the event basis takes into account the correction of the charged particle multiplicity correction, the vertex reconstruction efficiency correction and the trigger efficiency correction.
The corrections have been estimated for event and track cuts defined in Section 4.3 and Section 4.4. Here, the corrections are determined only with one choice of event and track selection cuts. The systematic uncertainty of the choice of selection cuts on the final corrected results is further studied in Chapter 7 (Sections 7.6 and 7.5).
The most important part of the single track correction is given by the tracking efficiency correction. The correction for the tracking efficiency is approximately 22 % integrated over all transverse momenta and pseudorapidities. The contamination correction is the second largest part of the track correction with approximately 6 % correction integrated over all transverse momenta and pseudorapidities. The data driven strangeness corrections increased the contamination correction by further 1 %. The two-track and detector effects account for an integrated correction of only 0.5 %.
The correction of the charged multiplicity also incorporates the combined factors of the tracking efficiency and the contamination correction stored in the correlation matrix $R(N_{\mathrm{prim}}, N_{\mathrm{rec}})$, as the multiplicity is defined by the number of accepted tracks or primary charged particles respectively. Here, it has to noted that the multiplicity correction of the final observables $O(N_{\mathrm{ch}})$ causes a correlation between the multiplicity-corrected data points in neighboring multiplicity bins. Hence, the multiplicity correction smooths the data point distribution as a function of the corrected multiplicity.
The vertex reconstruction efficiency influences only events of low charged particle multiplicities. At the lowest multiplicity bin $N_{\mathrm{ch}} = 1$, up to 15 % of the events are lost. However at multiplicities above $N_{\mathrm{ch}} = 5$, the vertex reconstruction efficiency is in agreement with unity. The vertex reconstruction efficiency has a small integrated impact (2 %) on the number of tracks entering the analysis, too. The trigger efficiency bias can be neglected as it is very small (maximal value of 0.1 % at multiplicities below only) in comparison to

all other effects.
The validity of the correction procedure is tested in Chapter 7. The correction procedure is use to correct ALICE data. Fully corrected ALICE results are presented in Chapter 8.

7. Systematic Uncertainties

In this chapter, the systematic uncertainties of the analysis results are discussed. The observables of the data analysis are shortly recalled for this purpose: The two-particle correlation analysis results include the per-trigger yield as a function of the difference in the azimuthal angle $1/N_{\mathrm{trig}} \cdot dN/d\Delta\varphi$ between trigger particles and associated particles (cf. Chapter 5). The two-particle correlations are performed as a function of the charged particle multiplicity. Two transverse momentum thresholds are used, $p_{\mathrm{T,\,trig}}$ for trigger particles and $p_{\mathrm{T,\,assoc}}$ for associated particles. In addition to the per-trigger yield, derived observables of the per-trigger yield are studied. For this purpose, the per-trigger yield is decomposed into the integrated yield of the isotropically distributed background $\langle N_{\mathrm{isotrop}} \rangle$, the integrated yield of the near side peak $\langle N_{\mathrm{assoc,\,near\ side}} \rangle$, and the away side peak $\langle N_{\mathrm{assoc,\,away\ side}} \rangle$. The average number of trigger particles per event $\langle N_{\mathrm{trigger}} \rangle$ and the number of uncorrelated seed $\langle N_{\mathrm{uncorrelated\,seeds}} \rangle$ are measured, too. For the decomposition of the per-trigger yield, a fit function is used. All observables as well as the fit function have been introduced in detail in Chapter 5.

As a first step of the estimation of the systematic uncertainty, the operational reliability of the fit function used for the decomposition of the per-trigger yield is studied (Section 7.1). The resulting systematic uncertainties are only applied to the derived observables $\langle N_{\mathrm{assoc,\,near\ side}} \rangle$, $\langle N_{\mathrm{assoc,\,away\ side}} \rangle$, $\langle N_{\mathrm{isotrop}} \rangle$, and $\langle N_{\mathrm{uncorrelated\,seeds}} \rangle$. All other systematic uncertainties are estimated for all discussed observables.

In Section 7.2, the disagreement between the true results and the corrected results has been studied in a validation of the correction procedure that is introduced in Chapter 6. The dependence of the results on the event generator (Section 7.3) and the transport Monte Carlo (Section 7.4) used in the correction procedure is studied, too. Furthermore, the effect of the variation of the track and event selection on the final results is under investigation (Section 7.5 and 7.6). Modifications of the correction maps due to uncertainties in the particle composition of the event generators are discussed in Section 7.7. The limited knowledge of the detector in terms of the detector efficiency (Section 7.8) and the material budget (Section 7.9) are studied as well as the impact of pileup events (Section 7.10). In Section 7.11 and Section 7.12, two sources of the systematical uncertainties of the data driven strangeness correction are discussed.

The systematic uncertainties are estimated for all measured observables separated for all three center-of-mass energies. In the following, the effect of the systematic uncertainties is exemplarily studied for one of the observables, the per-trigger near side yield, $\langle N_{\mathrm{assoc,\,near\ side}} \rangle$, measured in pp collisions at $\sqrt{s} = 7.0\,\mathrm{TeV}$.

7.1. Per-Trigger Yield Measurement based on a Fit Function

As already introduced in Section 5.3.1, a fit function is used to measured the yield and the yield contributions of the per-trigger yield $1/N_{\mathrm{trig}} \cdot dN/d\Delta\varphi$. Using the fit, the properties of the $\Delta\varphi$-distribution can be expressed by the parameters of the fit. These are the amplitudes $A_{1,2,3}$ and the widths $\sigma_{1,2,3}$ of the Gaussian functions for the near and away side, and the constant C for the combinatorial background of the correlation. These parameters are then used to estimate the yield of the $\Delta\varphi$-distribution separated into the per-trigger near side yield, the per-trigger away side yield and the per-trigger yield in the combinatorial background.
The analysis approach relies on a good agreement between the per-trigger yield and its fit as well as a good stability of the fit result.
In the following, the quality of the description of the $\Delta\varphi$-distribution by the fit is discussed, as well as how strongly potential deviations between the data distribution and the fit influences the accuracy of the signal extraction and the final results. Also, the stability of the fit is quantified in this section.

7.1.1. Reproduction of Signal Distribution

A discrepancy between the $\Delta\varphi$-fit function and the measured $\Delta\varphi$-distribution (called data in the following) can result in an over- or an underestimation of the fit-extracted yields. The agreement between the data distribution $d(\Delta\varphi, N_{\mathrm{ch}})$ and its fit $f(\Delta\varphi, N_{\mathrm{ch}})$ can be monitored by residua. The data distribution is given by

$$d(\Delta\varphi, N_{\mathrm{ch}}) = \frac{1}{N_{\mathrm{trig}}(N_{\mathrm{ch}})} \frac{d^2 N_{\mathrm{assoc}}}{d\Delta\varphi dN_{\mathrm{ch}}}, \tag{7.1}$$

and the residuum distribution is defined as

$$R(\Delta\varphi, N_{\mathrm{ch}}) = \frac{d(\Delta\varphi, N_{\mathrm{ch}}) - f(\Delta\varphi, N_{\mathrm{ch}})}{\sigma_d(\Delta\varphi, N_{\mathrm{ch}})}. \tag{7.2}$$

Here, the $\sigma_d(\Delta\varphi, N_{\mathrm{ch}})$ is the uncertainty of the data points. The residua can be estimated for each charged particle multiplicity bin and for different ranges in the transverse momentum separately.
An example $\Delta\varphi$-distribution described by the fit function is presented in the left panel of Figure 7.1. The residuum between the data and the fit distribution for many bins in the charged particle multiplicity is shown in the right panel of Figure 7.1.
The fit quality is good if the residuum is distributed around $R \approx 0$ for all values in $\Delta\varphi$. Furthermore, the normalized width of the residuum data points should be $\sigma_R \approx 1$. A deviation from $R = 0$ corresponds to an under- or overestimation of the data distribution

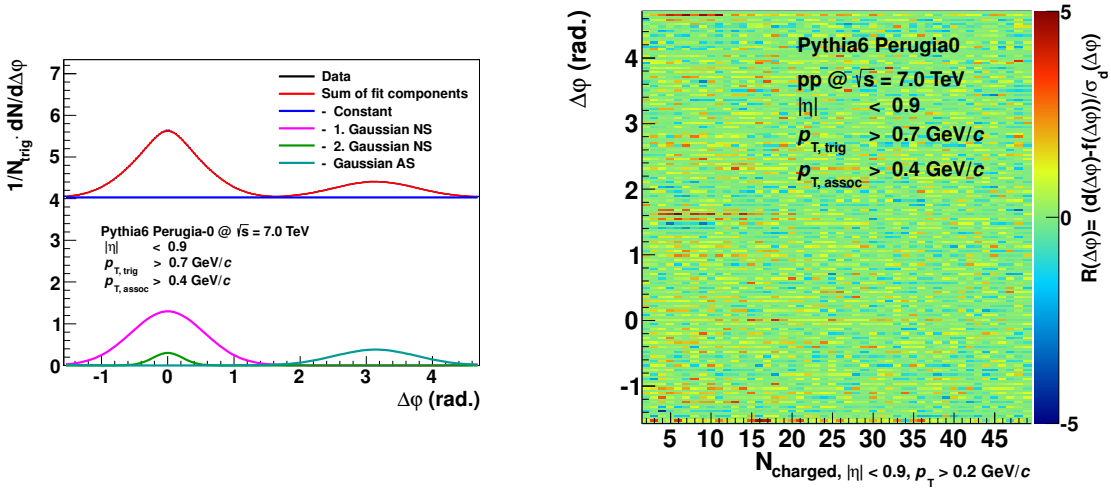

Figure 7.1.: Left panel: $\Delta\varphi$-distribution described by the fit function. The distribution is obtained using Pythia6 Perugia-0 simulations of proton-proton collisions at $\sqrt{s} = 7.0\,\text{TeV}$ for $p_{\text{T, trig}} > 0.7\,\text{GeV}/c$ and $p_{\text{T, assoc}} > 0.4\,\text{GeV}/c$. Right panel: Residua estimated for each bin in $\Delta\varphi$ and as a function the charged particle multiplicity.

with the fit. If σ_R is larger or smaller than 1, the error bars of the data points are either over- or underestimated.

For all presented bins in $\Delta\varphi$ and charged particle multiplicity, the deviations fluctuate roughly around $R = 0$ (green range of the color palette) as visible in the left panel of Figure 7.1. The deviations in R do not show any significant dependence on $\Delta\varphi$. Only a small range at $\Delta\varphi \approx \pi/2$ and low charged particle multiplicities shows slight underestimations. From Figure 7.1, the spread of the residuum values for all bins in $\Delta\varphi$, σ_R, can not be extracted easily by eye. The σ_R is further studied in Figure 7.2.

In the left panel of Figure 7.2, the R distribution for all $\Delta\varphi$ bins of Figure 7.1 are presented as a function of the charged particle multiplicity. The mean value of the residuum,$\langle R\rangle$, and the spread of the residuum, σ_R, are extracted by fitting the R distributions in each charged particle multiplicity bin with the Gaussian fit functions. The extracted properties $\langle R\rangle$ and σ_R are shown in the right panel of Figure 7.2. While at high charged particle multiplicities, the fit properties approach the optimal values of $\langle R\rangle = 0$ and $\sigma_R = 1$, slight deviations at low charged particle multiplicities are observed. It can be concluded from these results that the $\Delta\varphi$-fit slightly underestimates the $\Delta\varphi$-data distribution at low charged particle multiplicities ($\langle R\rangle > 1$), and the uncertainties of the $\Delta\varphi$-distribution seem to be slightly underestimated ($\sigma_R > 1$). The resulting impact of the fit-disagreement for the yield extraction is further discussed in Figure 7.4.

Besides the fit quality test using residuals, it is common to test the fit quality further by the χ^2/NDF-test (NDF = number of degrees of freedom). The χ^2/NDF-test probes

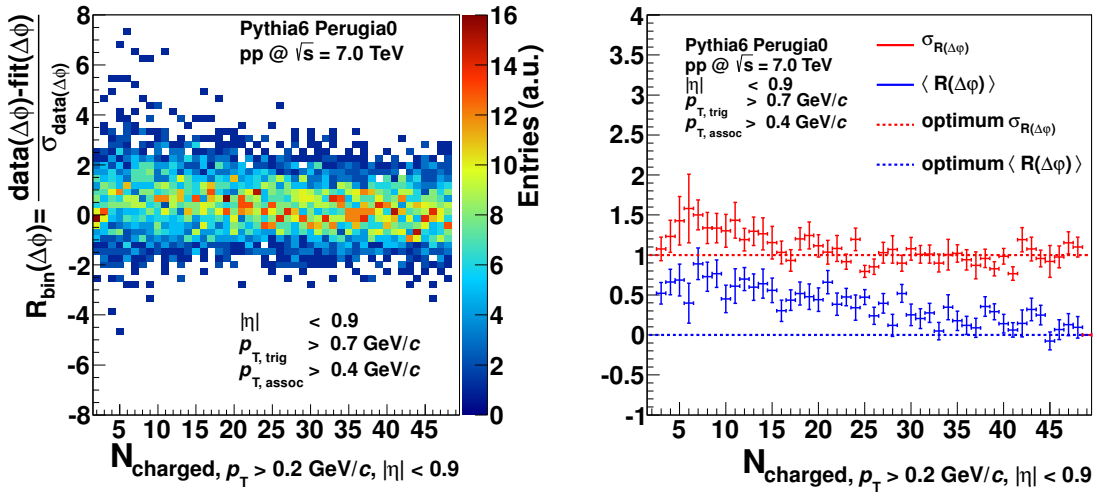

Figure 7.2.: Left panel: Values of $R(\Delta\varphi)$ as a function of the charged particle multiplicity integrated over all bins in $\Delta\varphi$. Right panel: Mean value $\langle R \rangle$ and width σ_R estimated with Gaussian fit functions for the distribution in the right panel of Figure 7.2 separated for the charged particle multiplicity. Both figures are based on results obtained by Pythia6 Perugia-0 events at $\sqrt{s} = 7.0\,\text{TeV}$. ALICE data and Phojet data give similar results.

how well a given data distribution can be described by a model. χ^2 is a measure for the disagreement between fit f and the data distribution d summed for all data points M. χ^2 is given by

$$\chi^2 = \sum_{i=1}^{M} \frac{(d_i - f_i)^2}{\sigma_{d_i}^2}. \tag{7.3}$$

By dividing the χ^2 by the number of degrees of freedom NDF$= M$ of the fit, the goodness of the fit can be quantified. The model is able to described the data distribution properly if the χ^2/NDF is close to 1. Figure 7.3 presents the χ^2/NDF for the fit results estimated as a function of charged particle multiplicity for Pythia6 Perugia-0 simulations. It is visible that at high charged particle multiplicities the agreement between fit and data distribution is slightly better than at low charged particle multiplicities. The χ^2/NDF for all charged particle multiplicities is never larger than 1.5. The fit describes the data distribution with fairly good agreement.

For the estimation of the per-trigger yield and the per-trigger yield components, the integral of the fit function is used. The estimation of the integral of the measured distribution using the fit function integral is only valid, if data and fit distribution agree with each other. Figure 7.4 shows that the integral of the data distribution and the integral of the fit are in excellent agreement for all discussed data sets. The effective disagreement between the data-yield and the fit-yield separated for near and away

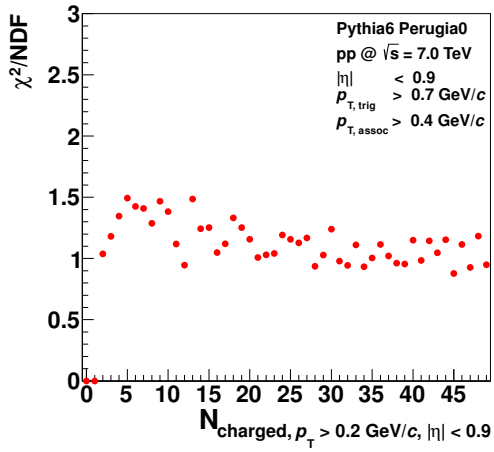

Figure 7.3.: χ^2/NDF as a measure of the fit quality as a function of the charged particle multiplicity. The figure is based on results obtained by Pythia6 Perugia-0 events at $\sqrt{s} = 7\,\text{TeV}$. ALICE data and Phojet data give similar results.

side is extremely low. In each charged particle multiplicity bin, the fit only slightly underestimates the near and away side yields, the maximum deviation is always below 0.4 %, and with increasing charged particle multiplicity, the disagreement decreases.
In comparison to the other systematic uncertainties of the measurement discussed in the subsequent sections, the here discussed systematic uncertainty of the disagreement between the $\Delta\varphi$-fit function and the $\Delta\varphi$-data distribution is extremely small.

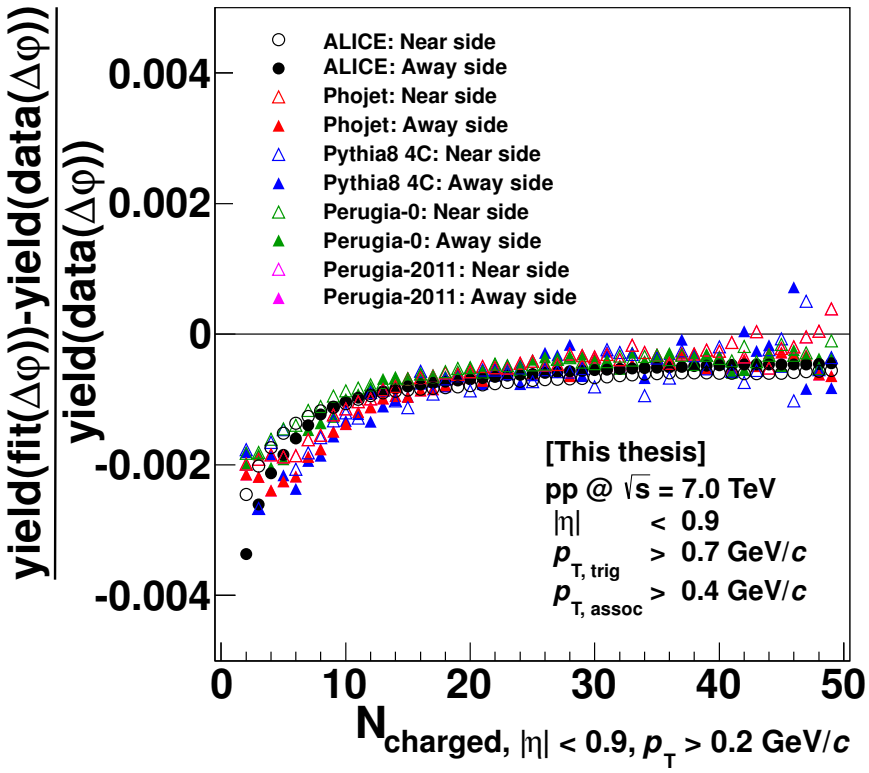

Figure 7.4.: Disagreement between $\Delta\varphi$-distribution and fit function averaged for the near side ($[-\pi/2, \pi/2]$) and the away side ($[\pi/2, 3\pi/2]$). The figure is based on ALICE data, simulations of Phojet, Pythia8 4C, Pythia6 Perugia-0, and Pythia6 Perugia-2011 at $\sqrt{s} = 7\,\text{TeV}$.

7.1.2. Stability of the Fit Results

In this section, the stability of the fit result of the per-trigger yield is studied. First, the stability of the fit is tested by modifying the fitted distributions slightly by adding combinatorial background to the per-trigger yield. Then, a data driven method using a special case of event mixing is used to validate the signal extraction method. Finally, the minimum number of entries per data distribution that is needed for a stable, reliable fit result is estimated.

Addition of Random Background to the $\Delta\varphi$-Distribution

As a cross check of the yield measurement using the fit function, the isotropically distributed part of a $\Delta\varphi$-distribution is modified. The isotropically distributed fraction of the $\Delta\varphi$-distribution is referred to as background in the following. When modifying the background of the $\Delta\varphi$-distribution, the fit function should only be modified in terms of its constant contribution representing the combinatorial background. The yields of the near side and the away side should remain unmodified. Hence, also the fit-extracted per-trigger near side yield and away side yield should remain unmodified.

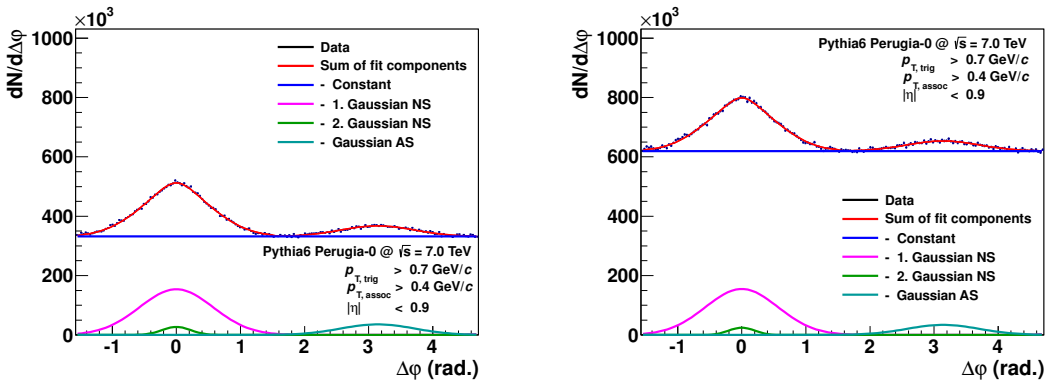

Figure 7.5.: Left panel: $\Delta\varphi$-distribution described by a fit function. The distribution is obtained using Pythia6 Perugia-0 at $\sqrt{s}$ = 7.0 TeV for events with $N_{\mathrm{ch}} = 10$. Right panel: The same distribution as in the left panel, but combined with isotropically distributed background that is generated by a random number generator. In order to generate the distribution of the left panel, 2 million additional, randomly distributed entries are added to the distribution of the right panel.

Figure 7.5 shows the same $\Delta\varphi$-distribution in the left and on the right panel, in the left panel before, and in the right panel after a modification of the constant fraction. The initial distribution is generated with Pythia6 Perugia-0 events at a charged particle multiplicity of $N_{\mathrm{ch}} = 10$. The $\Delta\varphi$-distribution is modified by adding randomly distributed entries over

all bins[1]. The modification is performed several times, each time (step) the same amount of addition background is added.

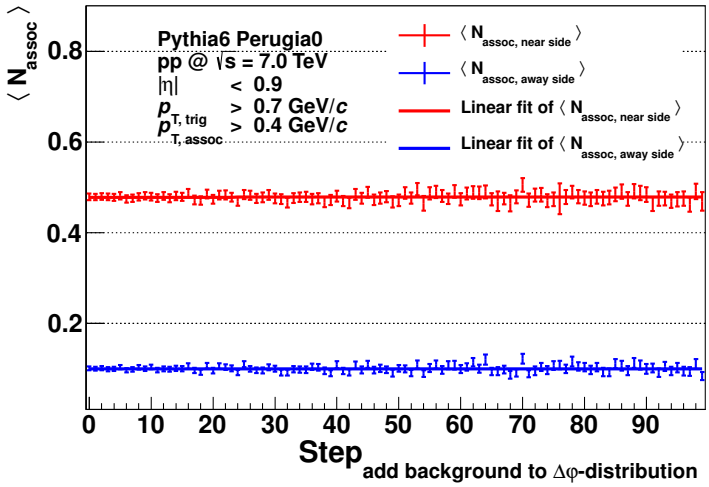

Figure 7.6.: Measured per-trigger near side yield and away side yield of the $\Delta\varphi$-distribution presented in Figure 7.5. With each step along the x-axis, the constant fraction of the input distribution has been increased. As expected, the same near side yield and away side yield is measured when modifying only the combinatorial background of the per-trigger yield.

Figure 7.6 shows the extracted per-trigger yield at the near and the away side as a function of an increasing amount of randomly added background. Each step along the x-axis represents additional 100,000 entries summed to the previous $\Delta\varphi$-distribution. The estimated yield at the near and away side stays constant with increasing background. The constancy is tested by fitting the slope of the distribution with a linear function. The slopes of the linear functions agree with zero within the uncertainties.
The measured near and away side yields stay constant when modifying the constant fraction of the pair distribution. However, a slight increase in the uncertainty given by the error bars of the fit results can be observed with increasing contribution to the constant fraction. This is expected and due to the fact that with decreasing signal over background ratio S/B also the significance α of the signal decreases. The significance is given as

$$\alpha \propto \frac{S}{\sqrt{S+B}}. \tag{7.4}$$

Here, S is the near and away side yield and B is the yield in the combinatorial background.

[1] As random number generator, the TRandom3 class of the analysis framework ROOT [BR97] is used.

Combination of Events

In this section, a data-driven consistency check of the yield measurement method is presented. For this purpose, combined events are built by superimposing several events to one event. For simplicity, the combined events are built of events of a fixed charged particle multiplicity N_{orig} only. Events with the combined multiplicities N_{combined} of $2 \cdot N_{\mathrm{orig}}$, $3 \cdot N_{\mathrm{orig}}$, ... $n \cdot N_{\mathrm{orig}}$ are found.
The average properties of the per-trigger near and away side yields measured in these artificial, combined events should be the same as those measured in the original events. The combinatorial background of the per-trigger yield as well as the number of trigger particles should increase with the number of combined events.
Figure 7.7 shows the per-trigger near side and away side yield estimated for events with $N_{\mathrm{orig}} = 15$ charged particles, and for events which are build out of several $N_{\mathrm{original}} = 15$ events, eg. $N_{\mathrm{combined}} = 2 \cdot N_{\mathrm{orig}} = 30$, $N_{\mathrm{combined}} = 3 \cdot N_{\mathrm{orig}} = 45$, and so forth. Per N_{combined}-bin, each input event is used only once. This means that the number of combined events in the n-th bin is by $1/n$ lower as in the first bin. In order to exclude an additional influence of different numbers of entries per bin on the fit results, here, exactly the same number of events per N_{combined}–bin are used.

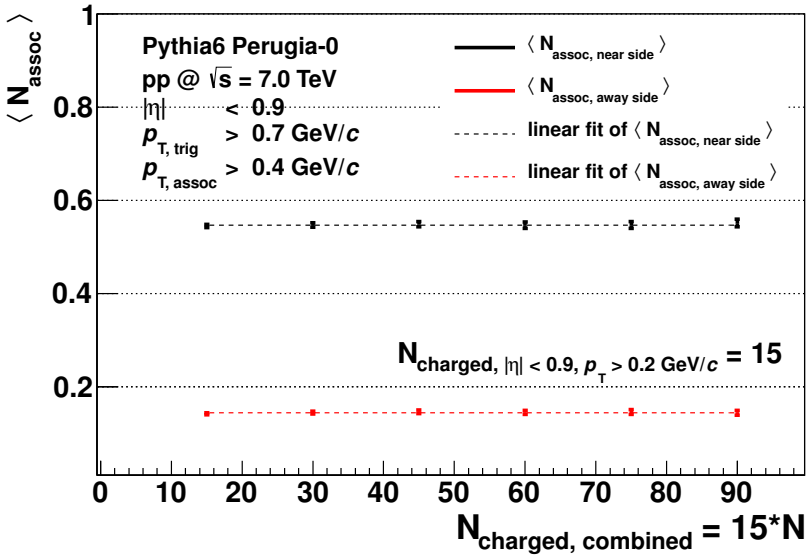

Figure 7.7.: Measured associated near and away side yield per trigger particle. The x-axis represents the number of particles per artificial combined event. Only events with a charged particle multiplicity of $N_{\mathrm{ch}} = 15$ are used in this exercise as input. The subsequent bins in N_{combined} are built of combinations of events with $N_{\mathrm{ch}} = 15$. For all N_{combined}–bins, the same number of combined events is used. As expected, for all charge particle multiplicity bins N_{combined}, the same near side yield and away side yield is measured.

The measured per-trigger near and away side yields $\langle N_{\text{assoc, near side}} \rangle$ and $\langle N_{\text{assoc, away side}} \rangle$ stay constant for the standard input events and all combined events. The slope of the $\langle N_{\text{assoc}} \rangle$ evolution is compatible with zero within the uncertainties of the slope.

Minimum Number of Events Per Bin

If the number of events and pairs entering the $\Delta\varphi$-distribution are too low, the fit result is likely to become unstable. Hence, it is necessary to determine the minimal number of events or pair-entries per histogram required to get a stable and reliable fit result.
In order to probe the stability of the fit results, a given $\Delta\varphi$-distribution can be modified slightly by adding additional, isotropically distributed entries. As discussed in the previous section, additional, limited background does not influence the near and away side fit results given by the Gaussian functions. The variation of the fit results between the steps of added background can then be interpreted as a measure of the fit stability.
For the estimation of the fit stability, the fit of the $\Delta\varphi$-distribution is repeated 100 times. Each time, 0.1 % of the initial number of pairs building the $\Delta\varphi$-distribution is added in terms of additional, randomly distributed background.

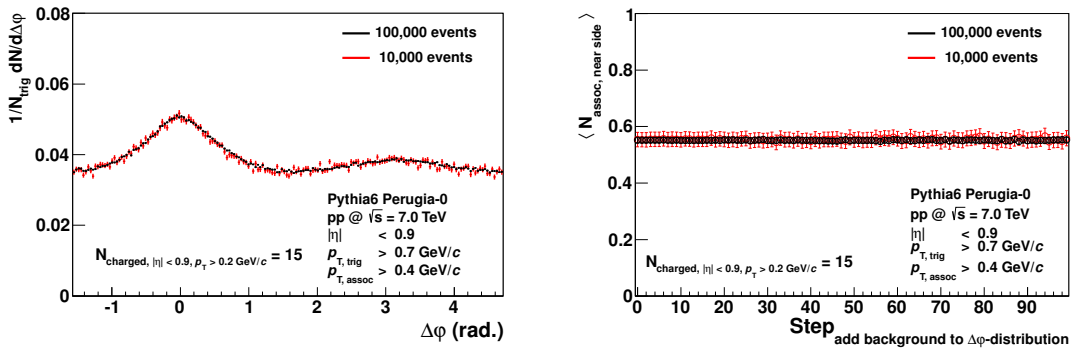

Figure 7.8.: Left panel: per-trigger yield as a function of $\Delta\varphi$ measured for events with a charged particle multiplicity of $N_{\text{ch}} = 15$. The measurement is performed with 100,000 events (black) and 10,000 events (red) respectively. Right panel: The per-trigger near side yield estimated using 100,000 and 10,000 events. The x-axis represents the amount of added background per step which corresponds to 0.1 % of the initial entries of the $\Delta\varphi$-distribution.

The left panel of Figure 7.8 shows the $\Delta\varphi$-distribution normalized to the number of trigger particles at a charged particle multiplicity of $N_{\text{ch}} = 15$. One distribution is estimated using 10,000 events, the other one is estimated using 100,000 events (these data points have smaller statistical uncertainties). The bin entries of the per-trigger yield fluctuate stronger the smaller the number of input events is. The right panel of Figure 7.8 shows the measured per-trigger near side yield estimated using the already described fit function.

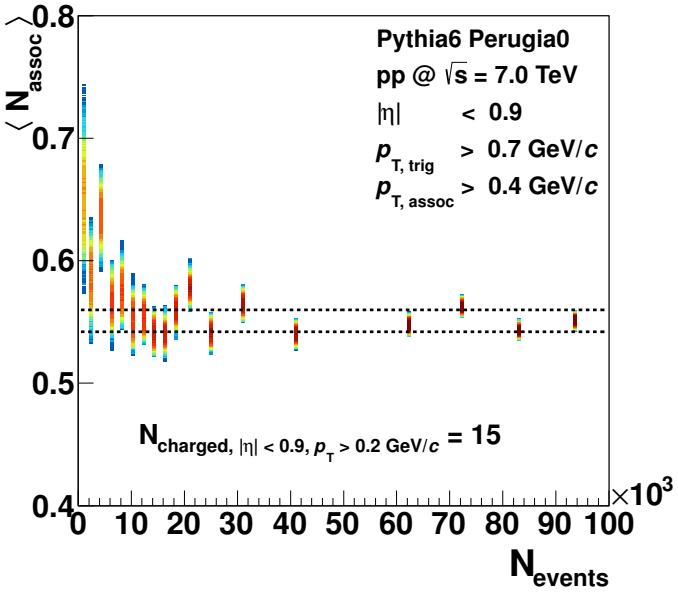

Figure 7.9.: Measured per-trigger near side yield using data samples of different sizes. The x-axes represents the number of events entering the two-particle correlation. The result estimated with the highest number of events is the reference value and interpreted as close to the true value. The reference value and its uncertainty is presented by the black, dashed lines. Only events with a charged particle multiplicity of $N_{\mathrm{ch}} = 15$ are used in this study.

In order to find the minimum number of events needed for a stable fit result, the number of events used for the estimation is reduced gradually. With each data set of a reduced number of input events, the near side yield is estimated exemplarily. The measured per-trigger near side yield for an example charged particle multiplicity of $N_{\mathrm{ch}} = 15$ is shown in Figure 7.9. The x-axis represents the number of input events. The uncertainty of the fit results is given by the fluctuation of the fit result when modifying the pair distribution slightly by adding additional combinatorial background as described in the previous section. For each data sample, the near side yield is estimated 100 times.
Using comparably few events in the presented example, the uncertainty of the extracted associated yield is larger than estimated using many events. In the extreme case of very few events, the estimated near side yield can even be in disagreement within the uncertainties to the value found using a high number of events. Based on Figure 7.9, it can be concluded that for a charged particle multiplicity of $N_{\mathrm{ch}} = 15$, at least 10,000 pairs are needed to get proper fit results.
The analyzed minimum bias event samples consist predominately of events with low charged particle multiplicities. Hence, for the analyzed data and MC data samples of at least 10,000,000 events per sample, the limit in the number of events applies only to

high charged particle multiplicities, where comparably few events are collected. At charged particle multiplicities above the presented $N_{\mathrm{ch}} > 15$, the ratio of the near and the away side yield to the yield in the combinatorial background is stable in the covered charged particle multiplicity range. This ratio represents the signal over background ratio in terms of the fit function. If this ratio had changed dramatically by orders of magnitude within the covered charged particle multiplicity range, it would have been necessary to adjust the minimum number of events as a function of the charged particle multiplicity in order to guarantee stable fit results for all charged particle multiplicities. As this is not the case, the minimum number of events of 10,000 can be chosen as limit for all analyzed charged particle multiplicities.

Influence of the Bin Width in the $\Delta\varphi$-Distribution

In the analysis and its correction procedure, memory intensive multi-dimensional sparse histograms are used to save particle-pair-, particle-, and event-properties. Each axis in the histogram is limited in the number bins. This is a technical limitation necessary to keep the amount of memory used to store the histograms during processing under control. For the same reason, data structures that operate on exact values[2] instead of values in bins, like in histograms, can not be used.
As a result, the pair property $\Delta\varphi$ is only available in a binned histogram with a limited resolution. In the following, the influence of the bin width on the final fit result is tested. First, the resolution of the $\Delta\varphi$-distribution is modified. Then, the fit function is used to extract the per-trigger yield. By comparing the results of the fit function between the different bin-size settings, the minimum resolution of the $\Delta\varphi$-distribution is estimated.

The left panel of Figure 7.10 shows the $\Delta\varphi$-distribution per trigger particle measured for events with an example charged particle multiplicity of $N_{\mathrm{ch}} = 10$. The same $\Delta\varphi$-distribution is presented using histograms with an increasing bin size. The histograms are not normalized to the bin size for a better visibility. The right panel of Figure 7.10 shows the extracted per-trigger near side yield as a function of the charged particle multiplicity. The results obtained with a histogram of a very fine binning of 720 bins is chosen as reference.
Using a small number of bins ($N_{\mathrm{bin}} \leqq 45$) for the $\Delta\varphi$-distribution, the estimated near side yield is comparably low. Compared to results estimated with a higher number bins, the results are 1 % lower for all charged particle multiplicity bins. This behavior can be explained by the fact that the minimum of the distribution between the near side and the away side peak cannot be resolved anymore when using too large bins. The combinatorial background is overestimated and therefore the integral of the near side

[2] TTree or TNtuple of the data analysis framework ROOT [BR97]

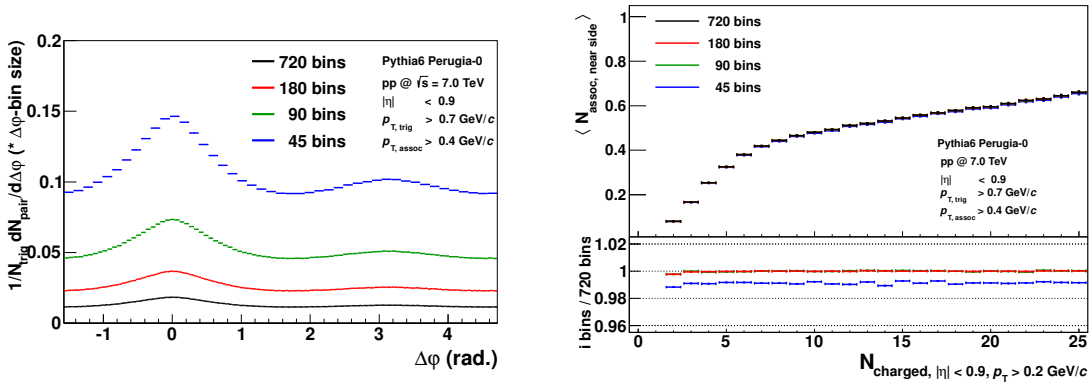

Figure 7.10.: Left panel: Per-trigger yield as a function of $\Delta\varphi$ measured for events with an example charged particle multiplicity of $N_{\mathrm{ch}} = 10$. Right panel: Measured per-trigger near side yield extracted from histograms of varying bin size in $\Delta\varphi$. Also, histograms of 360, 240, 144, 120, 102, 80, 72, 65, 60 55, 51, and 48 bins have been tested in this study (plot not shown).

peak is underestimated. The same is valid for the integral of the away side peak.
Using bin numbers between 720 and 90, the estimated near side yield is in agreement with each other. The deviation between the results estimated using between 720 and 90 bins is only 0.2 % for the first charged particle multiplicity bin and below 0.1 % for all other bins in charged particle multiplicity.
As a reasonable trade-off between accuracy and memory consumption, the present analysis uses 90 bins for the $\Delta\varphi$-axis.
In comparison to other systematic uncertainties discussed in the subsequent sections, the systematic uncertainty of the measurement due to the bin-width choice is very small.

7.2. Non-Closure in the Monte Carlo Correction

Using a full detector simulation of Monte Carlo events, detector effects and their influence on the final results can be estimated. The correction chain for all detector effects is described in Chapter 6. When applying the estimated corrections of the detector effects to the reconstructed Monte Carlo data, the true Monte Carlo results should be recuperate from the reconstructed signals, and the corrected data should again represent the properties of primary charged particles. The remaining discrepancy between the true charged primary results and the fully corrected results should be small or non-existent for a good unfolding procedure of the detector effects. Any remaining discrepancy is referred to as "non-closure" of the Monte Carlo correction.

Both panels of Figure 7.11 show the per-trigger near side yield. The near side yield is estimated for the true Monte Carlo results as well as for the reconstructed Monte Carlo results after correction of all detector effects. The left panel shows the per-trigger near

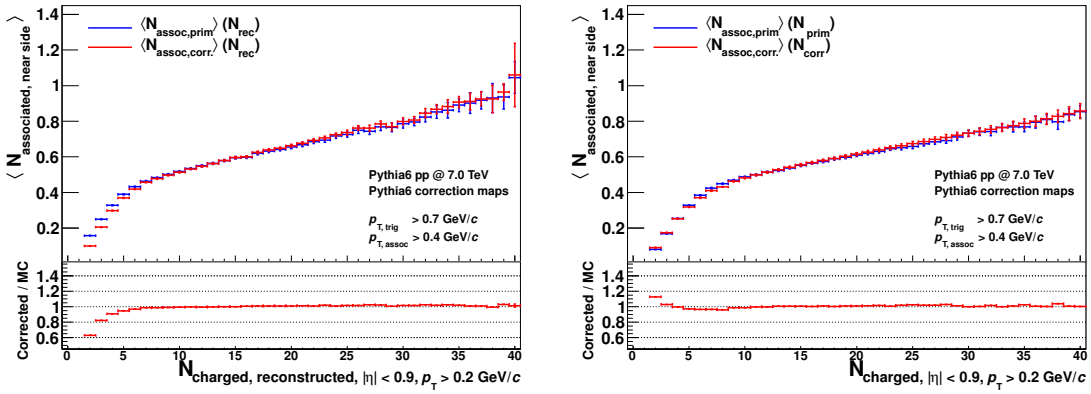

Figure 7.11.: Left panel: Per-trigger near side yield estimated for MC primary particles and reconstructed and fully corrected data as a function of the reconstructed multiplicity. Right panel: Per-trigger near side yield estimated for MC primary particles and reconstructed and fully corrected data as a function of the true charged particle multiplicity. The events are generated using Pythia6 Perugia-0 at $\sqrt{s}$ = 7.0 TeV. The correction is performed using Pythia6 Perugia-0 correction maps obtained offline by analyzing a full detector simulation of Pythia6 Perugia-0 data at $\sqrt{s}$ = 7.0 TeV.

side yield as a function of the reconstructed multiplicity of tracks coming from charged particles. The right panel of Figure 7.11 shows the per-trigger near side yield as a function of the corrected charged particle multiplicity after unfolding of the charged particle multiplicity. The bottom fractions of both figures show the ratio of the true and the corrected results. The analyzed data set is generated in Pythia6 Perugia-0 simulations performed at the center-of-mass energy of $\sqrt{s}$ = 7.0 TeV.

Before the multiplicity correction (left panel of Figure 7.11), the true and the corrected results deviate stronger from each other than after the multiplicity correction (right panel of Figure 7.11). This is in agreement with the fact that indeed the multiplicity integrated correction gives the same particle yields for corrected tracks and Monte Carlo primary particles, however, the true and the corrected yields separated into different reconstructed multiplicity bins still do not agree with each other. Only after multiplicity correction, this disagreement is solved.

The true MC results and the reconstructed and corrected results after unfolding of the charged particle multiplicity are in good agreement (cf. right panel of Figure 7.11). The ratio of the two results is close to unity. Only at charged particle multiplicities smaller than $N_{\mathrm{ch}} < 10$, a deviation between the MC results and the corrected results can be observed. First, the corrected result is slightly higher as the true result ($N_{\mathrm{ch}} < 5$), then the corrected result slightly undershoots the true result ($5 < N_{\mathrm{ch}} < 10$), and then the two results are consistent with each other ($N_{\mathrm{ch}} > 10$).

At the lowest charged particle multiplicity bin ($N_{\mathrm{ch}} = 2$), the deviation has its highest value of approximately 12 %. This is due to the fact that a lost particle or an additionally

measured particle at low charged particle multiplicities causes a larger relative variation of the particle yield as compared to high charged particle multiplicities.
With an uncertainty of 12 % for the first bin in the charged particle multiplicity, this systematic uncertainty is one of the dominant contributions of the total systematic uncertainty below $N_{\mathrm{ch}} < 5$. However, the focus of the analysis is on the high multiplicity range, where the systematic uncertainty decreases to small values. At larger charged particle multiplicities, other sources of systematic uncertainties dominate.

7.3. Event Generator

The correction maps used in the correction procedure are based on detector simulations of events generated by Monte Carlo generators. In this section, the influence of the choice of the event generator used to obtain the correction maps is studied.
For this purpose, reconstructed data of the Monte Carlo generator Phojet is corrected using either correction maps obtained with Phojet itself or using correction maps obtained with Pythia6 Perugia-0 simulations. The disagreement between the corrected results of the different correction maps give the systematic uncertainty of the correction procedure on the choice of the event generator.

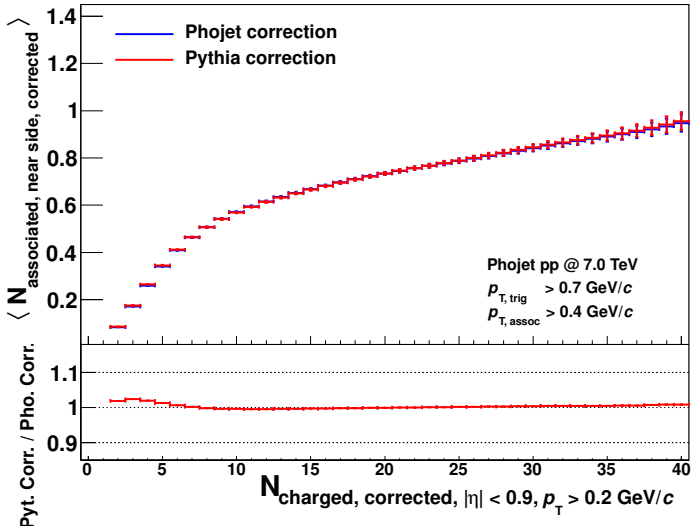

Figure 7.12.: Measured associated near side yield ($p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$) per trigger particle ($p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$) estimated with Phojet Monte Carlo data and corrected data. For the correction, Pythia6 Perugia-0 correction maps or Phojet correction maps are used. The pp collision data and the corresponding correction maps are generated at a center-of-mass energy of $\sqrt{s} = 7.0\,\mathrm{TeV}$.

The discrepancy between true and corrected results at low charged particle multiplicities is slightly higher when using correction maps obtained by another event generator. Figure

7.12 shows the corrected observables obtained with correction maps of different event generators. The difference between corrected results using maps of two different event generators is small and for none of the charged particle multiplicity bins higher than 2 %. For all $N_{\mathrm{ch}} > 7$, the corrected results estimated with the different correction maps are in agreement within the statistical uncertainties.
In comparison to the dominant systematic uncertainty of the correction procedure at low charged particle multiplicities (cf. Section 7.2), the event generator dependence contributes with a comparably small systematic uncertainty to the total systematic uncertainty.

7.4. Particle Transport Monte Carlo

After simulating the collision event with an event generator, the detector response is simulated using a particle transport Monte Carlo program. By simulating the interaction of the collision products with the detector material including the sensitive volumes as well as the support structures, detector efficiencies can be estimated. The estimated detector effects, such as tracking efficiency and the contamination of the reconstructed tracks by tracks from secondary particles, are extracted from these simulations and they are applied as a correction to the measured data. By default, ALICE simulations are performed using the transport Monte Carlo Geant3 [BBM$^+$87].
In order to estimate the dependence of the final, corrected results on the choice of the particle transport Monte Carlo, a sample of proton-proton collision events is simulated using Geant4 [Gea03, Gea06] as an alternative, independently developed transport Monte Carlo (cf. Section 4.2). The Geant4 simulation is used to compute an alternative correction map. In this test, the event generator Pythia6 Perugia-0 is used as event generator.
Figure 7.13 presents the associated near side yield per trigger particle estimated with Pythia6 Perugia-0. For the detector response, Geant3 and Geant4 are used. The correction maps estimated with one transport MC are applied to its corresponding data set as well as to the data set of the other transport MC and vice versa.
For all combinations, the corrected results are in good agreement with each other. The results agree within 1 % for all charged particle multiplicities. This systematic uncertainty is small compared to other systematic uncertainties discussed in the subsequent sections.

7.5. Track Selection

The track selection is performed with a fixed set of default track cuts introduced in Section 4.4. In order to estimate the impact of the choice of the track selection cuts on the final results, the tracks cuts used in the data analysis are varied (cf. Table 4.1 on Page 53). The variation of the final corrected results gives the systematic uncertainty of the choice of the track selection.

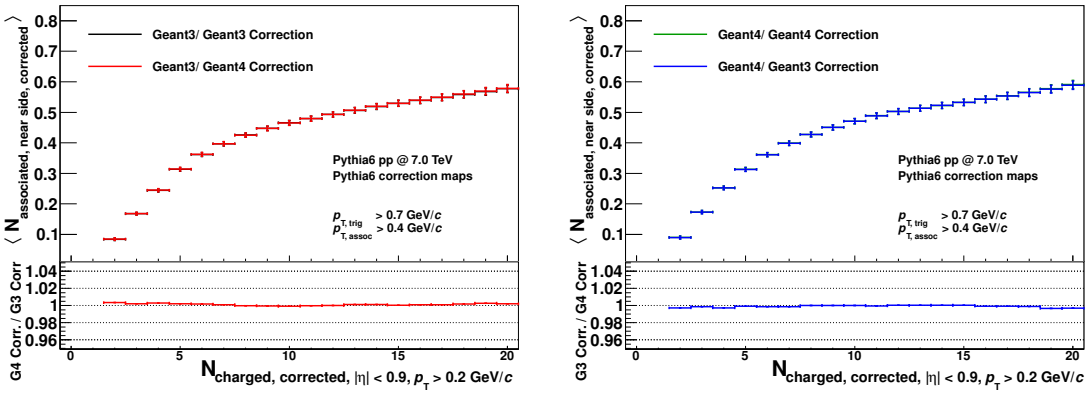

Figure 7.13.: Corrected associated near side yield ($p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$) per trigger particle ($p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$) estimated with Pythia6 Perugia-0. For the detector response used to create the correction maps, Geant3 and Geant4 used. The pp collision data and the corresponding correction maps are generated at a center-of-mass energy of $\sqrt{s} = 7.0\,\mathrm{TeV}$.

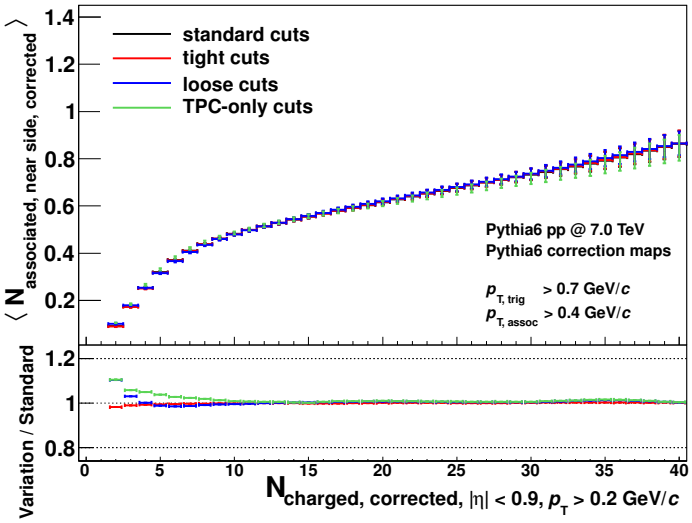

Figure 7.14.: Corrected associated near side yield ($p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$) per trigger particle ($p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$) estimated with Pythia6 Perugia-0. The corrected results are presented for four different sets of track cuts. The pp collision data and the corresponding correction maps are generated at a center-of-mass energy of $\sqrt{s} = 7.0\,\mathrm{TeV}$.

The associated near side yield per trigger particle estimated for Pythia6 Perugia-0 simulations is shown Figure 7.14. The full analysis and correction chain is performed using the four different sets of track selection cuts.
In the first charged particle multiplicity bin, the results show a difference of 11 % between the track sample accepted by the default track cuts ("optimized ITS-TPC track cuts") and

the track sample accepted by comparably more loose track cuts, the "loose" cuts and the "TPC-only" track cuts. An agreement within $\pm 1\,\%$ between all data samples is obtained for charged particle multiplicities above $N_{\mathrm{ch}} \geq 10$.
The systematic uncertainty of the track cut choice is the second largest contribution to the overall systematic uncertainty at low charged particle multiplicities. Only, the systematic uncertainty of the correction procedure with a value of $12\,\%$ for the first charged particle multiplicity bin is larger. The systematic uncertainty of $1\,\%$ for all high charged particle multiplicities is small in comparison to contributions of other systematic uncertainties discussed in the subsequent sections.

7.6. Event Selection

In the data analysis, a minimum bias collision event is selected by the event quality cuts, if, for example, the reconstructed event vertex has at least one contributing track (cf. Section 4.3.2). In order to estimate the influence of this event selection on the final results, the analysis and correction chain is repeated selecting events having vertices with at least two associated tracks.
This test selection is expected to have an impact only at low charged particle multiplicities as the vertices of high charged particle multiplicity events have in most cases much higher number of contributing tracks.
Figure 7.15 presents the influence of the different vertex quality cuts on the per-trigger near side yield. For the lowest charged particle multiplicity $N_{\mathrm{ch}} = 2$, the largest relative difference in the final results of $2\,\%$ is observed. The difference decreases with increasing charged particle multiplicity. For events with $N_{\mathrm{ch}} > 10$, the results from the two approaches agree.
The systematic uncertainty of the event selections is small in comparison to the systematic uncertainties of the correction procedure and the track selection dependence.

7.7. Particle Composition

The reconstruction efficiency of a particle depends strongly on the particle type. This is due to e. g. different decay lengths of the particles or different probabilities of the particles to be absorbed by the detector material. If the particle composition given in a Monte Carlo model does not reflect the particle composition of the real collision data, the reconstruction efficiency and contamination correction based on this model does not accurately correct the efficiency and contamination of real collision data. The study of the impact of the disagreement in the particle composition on the final results is described in this section.
ALICE has measured the production of pions, kaons and protons in proton-proton

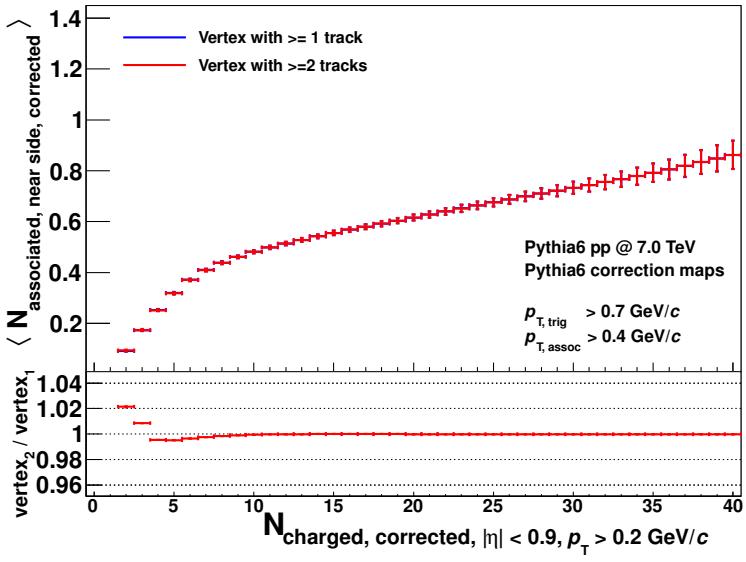

Figure 7.15.: Difference between corrected results when applying different vertex quality cuts. (1) Vertices with at least one contributing track are accepted. (2) Only events with at least two contributing tracks are accepted. Presented is the per-trigger near side yield measured in Pythia6 Perugia-0 events at $\sqrt{s}$ = 7.0 TeV.

collisions [ALI11c] as well as the production of hadrons including strange quarks [ALI11d]. The results show that some particle yields as measured by ALICE are poorly described by current Monte Carlo models. The ALICE analysis shows that the maximal deviation of the momentum integrated particle yield per particle type is about 30 %.
In Section 6.1.2, the effect of an underestimate strangeness yield (K, Λ) in the Monte Carlo generators on the correction procedure has been studied. It has been used to correct the known discrepancies of the strange particle yields in MC and real data. In this section, the effect is studied once more in order to estimate the systematic uncertainty of the variation of all particle yields.
The influence of a varied particle composition on the correction maps and the final corrected results is studied in the following by varying the particle yields of pions, protons, and kaons in the MC models by ±30 %. The same approach has already been deployed in other analyses within the ALICE experiment [Val12, ALI11f].

Figure 7.16 shows the reconstruction efficiency of tracks from primary particles, the contamination of the reconstructed track sample with tracks from secondary particles, and the sum of both values.
The dependence of the reconstruction efficiency on the transverse momentum has been discussed in detail in Section 6.2. In summary, with increasing momentum of the primary

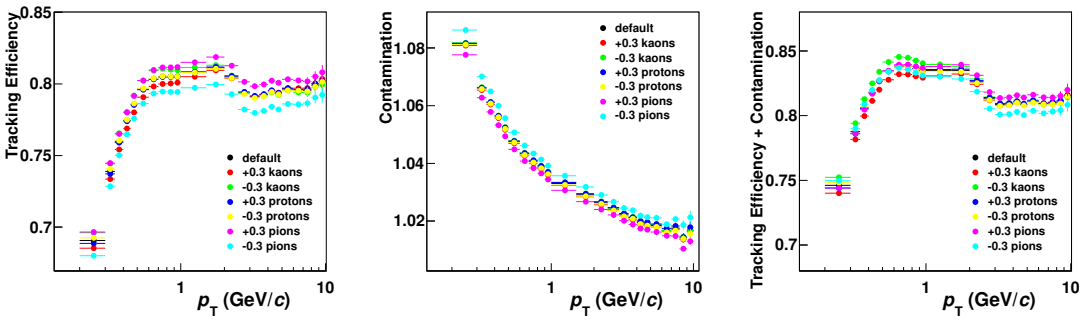

Figure 7.16.: Left panel: Reconstruction efficiency for primary particles estimated for data sets of varied particle composition. Mid-panel: Contamination of reconstructed track sample with track from secondary particles estimated for data sets of varied particle composition. Right panel: Combination of reconstruction efficiency and contamination. The values are estimated with a full detector simulation of Pythia6 Perugia-0 events at $\sqrt{s}$ = 7.0 TeV combined with a weighting of the modified particle species.

charged particles, the number of associated TPC clusters increases in average. A minimum number of 70 TPC clusters is requested to accept the tracks with the track selection cuts. Hence, the reconstruction efficiency for low momentum particles increases with p_{T}. At $p_{\mathrm{T}} > 3\,\mathrm{GeV}/c$, the charged particles have almost straight tracks. A fraction of the tracks can fall into one of the eighteen 2°-wide blind zones between the eighteen TPC sectors. Hence, the reconstruction efficiency decreases slightly in this transverse momentum range. When combining the reconstruction efficiency and the contamination correction, the effect on the combined correction in the low momentum region is largest when changing the kaon yield. For the high momentum region ($p_{\mathrm{T}} > 2\,\mathrm{GeV}/c$), the impact of the pion yield modification is largest. As the per-trigger yield studied in this data analysis is estimated including low momenta ($p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ and $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$), the impact of the kaon yield modification influences the analysis results strongest.

Figure 7.17 shows the impact of the kaon yield modification in the correction maps on the per-trigger near side yield. For all charged particle multiplicities, the impact is approximately 2 %. This systematic uncertainty is the second largest contributions to the total systematic uncertainty at high charged particle multiplicities. However, the systematic uncertainty of the detector efficiency discussed in the next section is much larger.

As in Section 6.1.2 the impact of the discrepancy of the kaon yield between MC models and real data has already been corrected, this systematic uncertainty does not necessarily need to be applied to the data. In the following, this systematic uncertainty is applied anyway as it is comparably small.

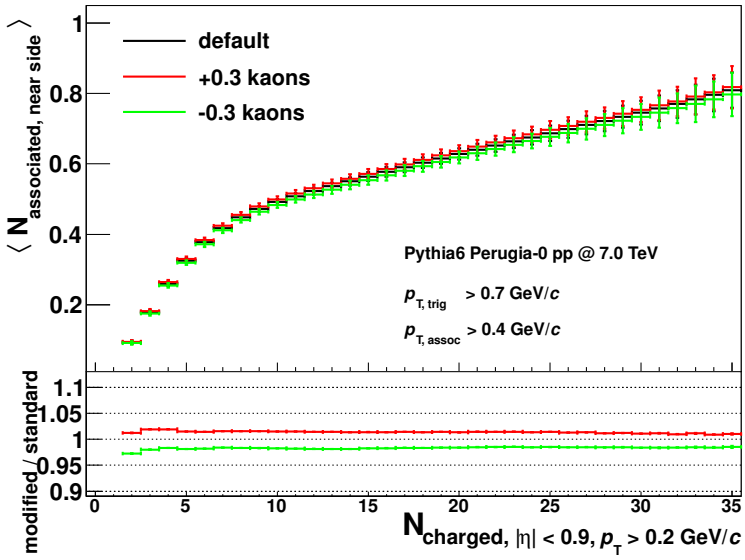

Figure 7.17.: Per-trigger near side yield for a Pythia6 Perugia-0 data sets at $\sqrt{s}$ = 7.0 TeV. The data analysis is performed for the (1) default setting, (2) 30 % more kaons, (3) 30 % less kaons as in the default setting. Also, the correction was performed with the default Pythia6 Perugia-0 particle composition (1) as well as with a modified relative particle composition of charged kaons (2,3).

7.8. Detector Efficiency

The precision in estimating the tracking efficiency using a full detector simulations depends on the accuracy of the detector description used for the detector simulation. In the presented data analysis, the Inner Tracking System (ITS) and the Time Projection Chamber (TPC) are used for the track measurement. The uncertainty of the ITS-TPC efficiency has been estimated by comparing the TPC-to-ITS and ITS-to-TPC track matching efficiencies in Monte Carlo data and ALICE data [ALI10c, ALI11e].
For proton-proton collision data sets measured at $\sqrt{s}$ =0.9 TeV, the detector efficiency uncertainties have the following transverse momentum dependent values ([ALI10c, ALI11e])

$$1.\,\% + 6.\,\% \text{ for } p_{\mathrm{T}} < 0.20\,\mathrm{GeV}/c,$$
$$1.\,\% + 3.\,\% \text{ for } 0.20 < p_{\mathrm{T}} < 0.25\,\mathrm{GeV}/c,$$
$$1.\,\% + 2.\,\% \text{ for } p_{\mathrm{T}} > 0.25\,\mathrm{GeV}/c,$$

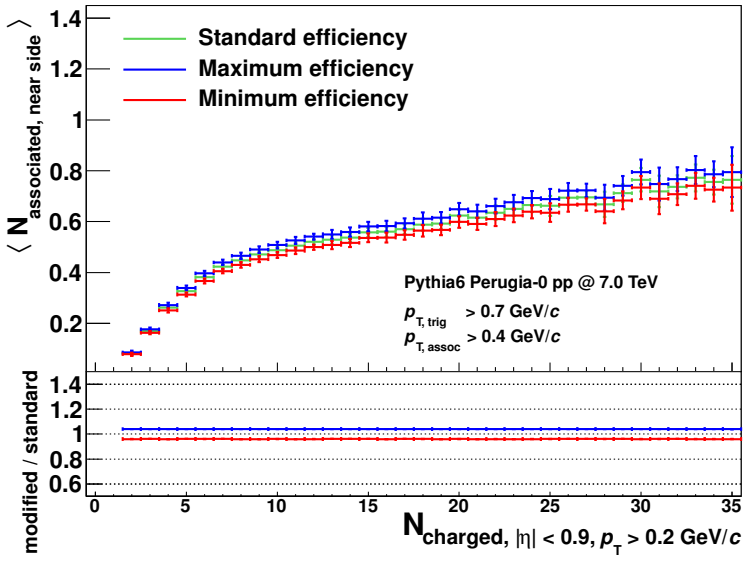

Figure 7.18.: Per-trigger near side yield for Pythia6 Perugia-0 events at $\sqrt{s}$ = 7.0 TeV. The data analysis is performed using the default detector efficiency, the lowest possible efficiency and the highest possible efficiency.

and for $\sqrt{s}$ =2.76 and 7.0 TeV, the detector efficiency uncertainties are

$$1.\,\% + 8.\,\% \text{ for } p_{\mathrm{T}} < 0.20\,\mathrm{GeV}/c,$$
$$1.\,\% + 5.\,\% \text{ for } 0.20 < p_{\mathrm{T}} < 0.25\,\mathrm{GeV}/c,$$
$$1.\,\% + 3.\,\% \text{ for } 0.25 < p_{\mathrm{T}} < 0.30\,\mathrm{GeV}/c,$$
$$1.\,\% + 3.\,\% \text{ for } 0.30 < p_{\mathrm{T}} < 1.00\,\mathrm{GeV}/c,$$
$$2.\,\% + 2.\,\% \text{ for } p_{\mathrm{T}} > 1.00\,\mathrm{GeV}/c.$$

The efficiency does not vary as a function of the pseudorapidity for central pseudorapidities. For the estimation of the impact of the detector efficiency uncertainty on the correlation observables, the transverse momentum dependent detector efficiency is varied within the systematic uncertainties.

Figure 7.18 shows the per-trigger near side yield for Pythia6 Perugia-0 simulations at $\sqrt{s}$ = 7.0 TeV. The impact of the variation of the detector efficiency on the observable is almost constant for all charged particle multiplicity bins. The impact on the per-trigger near side yield at $\sqrt{s}$ = 7.0 TeV is $\pm 4.1\,\%$.

The systematic uncertainty of the detector efficiency represents the largest contribution to the total systematic uncertainty for all charged particle multiplicities. Only at low multiplicities, the contributions of the correction procedure and the track cut dependence have larger values.

7.9. Material Budget

The uncertainty of the material budget in the central region of ITS and TPC has been studied in [ALI10c, ALI11e] and the uncertainty of the material budget has been estimated to have the same impact at all center-of-mass energies, $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV. The derived transverse momentum dependence for the reconstruction of the uncertainty is given by

$$1.5\,\% \text{ for } p_{\mathrm{T}} < 0.20\,\mathrm{GeV}/c,$$
$$0.5\,\% \text{ for } p_{\mathrm{T}} < 0.75\,\mathrm{GeV}/c,$$
$$0.2\,\% \text{ for } p_{\mathrm{T}} > 0.75\,\mathrm{GeV}/c.$$

The uncertainty is insensitive to the pseudorapidity.
The uncertainty of the material budget can be translated into an uncertainty of the reconstruction efficiency of charged primary particles. In order to estimate the impact of the material budget uncertainty on the final analysis results, the analysis is performed using a variation of the reconstruction efficiencies defined by the uncertainty interval of the material budget.

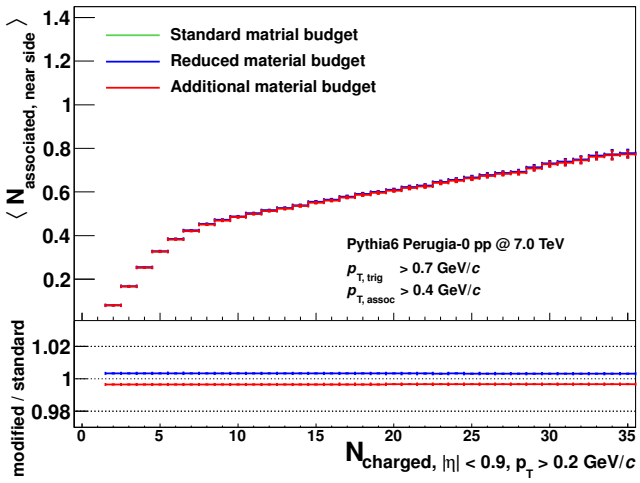

Figure 7.19.: Per-trigger near side yield for Pythia6 Perugia-0 events at $\sqrt{s}$ = 7.0 TeV. The data analysis is performed with the standard material budget, with enhanced material budget, and reduced material budget.

Figure 7.19 shows the maximum influence of the material budget uncertainty on the per-trigger near side yield. The influence of the material budget uncertainty is 0.4 % for low charged particle multiplicities with a slight, linear decrease towards high multiplicities. At very high multiplicities, the uncertainty approaches 0.3 %. This uncertainty is very

small in comparison the systematic uncertainty of the already discussed detector efficiency uncertainty, which is 10 times larger at all charged particle multiplicities (cf. Section 7.8).

7.10. Pileup Events

Pileup events are events with more than one proton-proton collision per bunch crossing. They need to be considered as source of systematic uncertainties as they could influence the average per-trigger yield in the following way. Events with pileup vertices can be misidentified as events with a single proton-proton collision of a higher charged particle multiplicity if the two (or more) pileup vertices are that close to each other that they can not be resolved from each other. The per-trigger yield measured in this fake high-multiplicity event is the average of the per-trigger yield of the two (or more) low-multiplicity events. The per-trigger near side yield grows with the charged particle multiplicity. If the accepted high-multiplicity events are contaminated by pileup events, the measured per-trigger near side yield measured at high charged particle multiplicities can be underestimated.

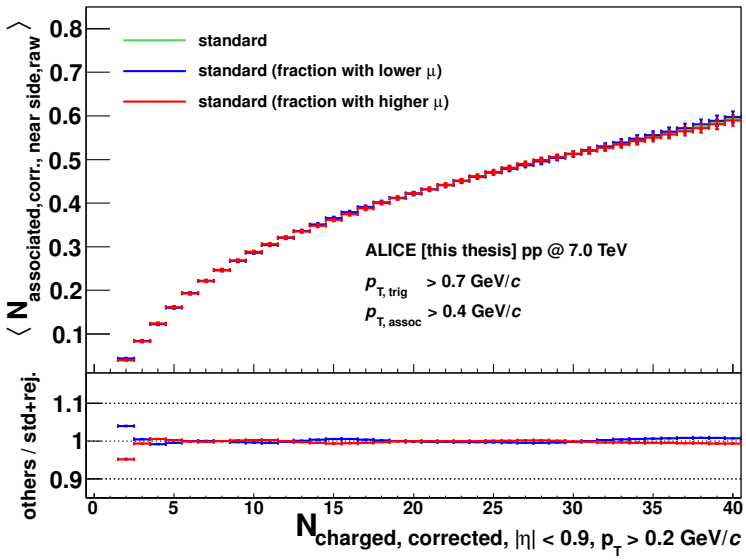

Figure 7.20.: Associated near side yield for ALICE data sets at $\sqrt{s}$ = 7.0 TeV. By comparing the results at different ranges in the mean number of interactions per bunch crossing μ, the influence of pileup events on the measurement is tested.

The influence of pileup events is estimated by comparing the analysis results of data sets recorded at different ranges of the mean number of interactions per bunch crossing denoted as μ. For this purpose, the standard data set at $\sqrt{s}$ = 7.0 TeV is divided into a fraction

with relatively low μ and a fraction of relatively high μ.
The resulting variation of the measured per-trigger near side yield is shown in Figure 7.20. The results of both data sub-sets are in good agreement with each other. Only for low charged particle multiplicities, the results differ. The disagreement to the standard data set is never higher than 5 %. The systematic uncertainty of pileup events is only applied to the results measured at $\sqrt{s}$ = 7.0 TeV. At $\sqrt{s}$ = 0.9 and 2.76 TeV, the pileup probability can be neglected due to very low values of the mean number of interactions μ.
The uncertainty related to pileup is small in comparison to other systematic uncertainties, such as the uncertainty related to the correction procedure and the track selection.

7.11. Correction of Strangeness Yields

The uncertainty of the strangeness yield estimation has already been covered partially as part of the study of the particle composition uncertainty discussed in Section 7.7. Here, the uncertainty has been estimated by the impact of yield modifications of charged kaons. Additional particles which contain strange quarks and which are produced in comparably high number in pp collisions are K^0, Λ, and ϕ. However, the yield of the ϕ particle production is well reproduced by all tested Monte Carlo generators [ALI11d]. Hence, solely the uncertainty of the K^0 and Λ yields needs to be considered in this estimation.
In other studies of the ALICE collaboration, it has been shown that the uncertainty of the strangeness correction is of the order of 0-2.3 % of the total particle yield depending on the transverse momentum of the charged particles [Val12, ALI11f]. Here, the same transverse momentum range has been used as in the data analysis of this thesis. The uncertainty of the charged kaon yield has been included in this uncertainty.
The uncertainty of the charged kaon yield itself estimated in the study of the particle composition uncertainty (cf. Section 7.7) is evaluated to be of approximately 2 %. The data driven strangeness correction itself estimated in Section 6.1.2 has a value of approximately 1 %.
It is concluded that the remaining uncertainty for the strangeness correction uncertainty in the analysis of this thesis is comparably small. Given that the combination of all other discussed systematic uncertainties is rather large compared to the expected systematic uncertainty of the strangeness yield correction, this fraction of the systematic uncertainty is neglected.

7.12. Extrapolation Uncertainty of Strangeness Correction

ALICE has measured the production of particles that include strange quarks in pp collisions at $\sqrt{s}$ = 0.9 TeV [ALI11d, ALI11c]. As part of this data analysis, it has been measured that the Monte Carlo generators Pythia6 Perugia-0 and Phojet used during

the correction of the ALICE data underestimate the production of particles including strange quarks. Hence, both MC also underestimate the contamination correction for tracks which come from secondary particles. This problem has been discussed in detail in Section 6.1.2.
The strangeness correction needs to be applied to all ALICE pp data sets including data at $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV. At the time of writing, ALICE has estimated the strange particle yield only for pp collisions at 0.9 TeV. Hence, the correction estimated for $\sqrt{s}$ = 0.9 TeV has also been applied to the data sets at $\sqrt{s}$ = 2.76, and 7.0 TeV. In the following, the systematic uncertainty of this extrapolation of the correction values from 0.9 TeV to higher center-of-mass energies is estimated. For this purpose, strange particle yields measured by the CMS experiment at $\sqrt{s}$ = 0.9 and 7.0 TeV are studied [CMS11c]. The CMS experiment has studied the particle yields of K^0_S, Λ, and Ξ^- for $\sqrt{s}$ = 0.9 and 7.0 TeV in comparison to Monte Carlo predictions [CMS11c]. It is found that the disagreement between the CMS results and the Pythia6 Perugia-0 predictions changes slightly between $\sqrt{s}$ = 0.9 and 7.0 TeV. The corresponding correction factors (Perugia − 0/CMS data) for the particles including strange quarks are extracted for the two collision energies and are used as weights for the strange particles found in ALICE simulations. The resulting strangeness correction factor (as already estimated in Section 6.1.2) are presented in the left panel of Figure 7.21 in combination with the ALICE strangeness correction factors found using the results measured at 0.9 TeV. It is visible that the results of ALICE and CMS at $\sqrt{s}$ = 0.9 TeV do not fully agree with each other. This is presumably due to the fact, that CMS and ALICE cover different acceptances in pseudorapidity. Also, slightly different event classes are used in the two data analyses. Nevertheless, all estimated correction factors are very small ($\approx$ 1 %) resulting in a small absolute difference between the correction factors.

The influence of the strangeness correction factors obtained with data at different center-of-mass energies is estimated using the correction values obtained by the CMS results. The right panel of Figure 7.21 shows the modification of the corrected per-trigger near side yield when using the two different strangeness correction factors. Only at low charged particle multiplicities any influence is visible. Here, the maximal deviation is of the order of 2 %. The estimated systematic uncertainty of the strangeness correction extrapolation is applied to the ALICE data at 7.0 TeV as well as 2.76 TeV, however the systematic of the extrapolation is expected to be slightly smaller for the 2.76 TeV data. For 2.76 TeV, no analysis results of strange particle yields are available at the time of writing.
The systematic uncertainty related to this extrapolation is small in comparison to other systematic uncertainties.

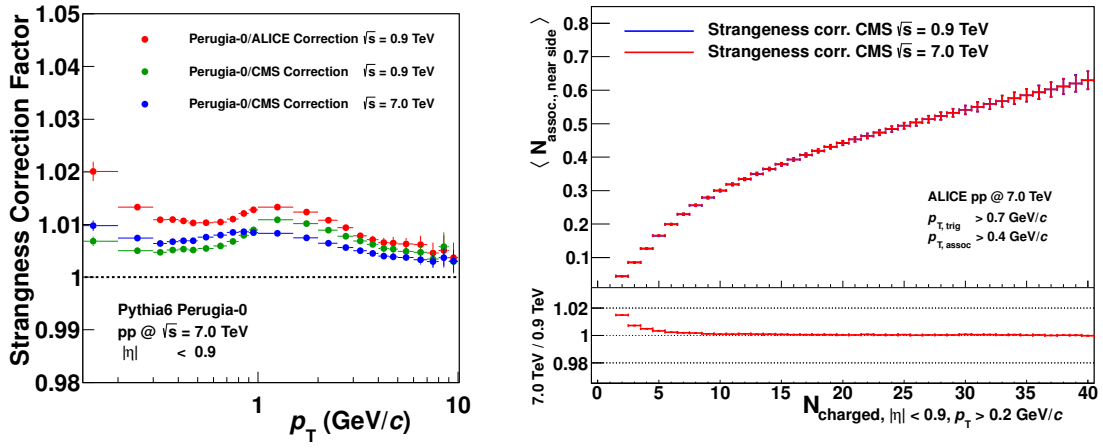

Figure 7.21.: Left panel: Strangeness correction factor obtained for Pythia6 Perugia-0 simulations at $\sqrt{s}$ = 7.0 TeV based on measurements of strange particle yields by ALICE and CMS at 0.9 and 7.0 TeV. Right panel: Fully corrected per-trigger near side yield of ALICE data measured at $\sqrt{s}$ = 7.0 TeV obtained with data driven correction factors based on strange particle yields measured by CMS at $\sqrt{s}$ = 0.9 and 7.0 TeV.

7.13. Summary

All systematic uncertainties discussed in this chapter are combined quadratically to the total systematic error of the measurement. In the region of very high charged particle multiplicities, the number of simulated or measured events is very low as the charged particle multiplicity distribution $P(N_{\mathrm{ch}})$ is a steeply falling distribution. Hence, the computation of the systematic uncertainties is limited to low and intermediate N_{ch}. For high charged particle multiplicities, extrapolated values from intermediate charged particle multiplicities are applied. Partially, the systematic uncertainty at high charged particle multiplicities are dominated by statistical uncertainties. The systematic uncertainty is computed for each observable and each center-of-mass energy separately.

The impact of many systematic uncertainties is biggest for the first bin of the charged particle multiplicity. To loose or to gain particles in this bin has the highest relative impact because only few particles are present. The first bins are $N_{\mathrm{ch}} = 2$ for the per-trigger yields (at least two particle are needed to perform a two-particle correlation) and $N_{\mathrm{ch}} = 1$ for the observables N_{trig} and $N_{\mathrm{uncorrelated\ seeds}}$. For high charged particle multiplicities above $N_{\mathrm{ch}} > 2 \cdot \langle N \rangle$, the systematic uncertainties are almost constant. At very high charged particle multiplicities above $N_{\mathrm{ch}} > 4 \cdot \langle N \rangle$, the systematic uncertainties increase again due to statistical fluctuations related to the low number of events.

As example of the total systematic uncertainty for all observables, the total systematic uncertainty of the per-trigger near side yield as a function of the charged particle multiplic-

ity is presented in Figure 7.22 for pp collisions at $\sqrt{s}$ = 7.0 TeV. The detector efficiency uncertainty represents the largest contribution to the total systematic uncertainty for all charged particle multiplicities. The second largest contribution is given by the particle composition uncertainty. At low charged particle multiplicities, however, the uncertainty of the correction procedure and the track cut dependence represent the largest contributions to the total systematic uncertainty.

The total systematic uncertainty of the per-trigger near side yield reveals a small bump at $N_{\mathrm{ch}} \approx 8$. This structure is due to the structure of the systematic uncertainty of the correction procedure. While for the lowest charged particle multiplicities, the corrected results are slightly higher as the true MC results ($N_{\mathrm{ch}} < 5$), the corrected results slightly undershoot the true results at $5 < N_{\mathrm{ch}} < 10$. At higher multiplicities, the true and the corrected results are again consistent with each other ($N_{\mathrm{ch}} > 10$).

In Table 7.1, the values of the systematic uncertainties are shown for the three studied center-of-mass energies, $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV. For simplicity, the systematic uncertainties for all charged particle multiplicities are presented for only two bins. The first charged particle multiplicity bin is $N_{\mathrm{ch}} = 2$, and the second bin is chosen at intermediate charged particle multiplicity $2 \cdot < N_{\mathrm{CMS}} >$.

In summary, all relevant systematic uncertainties have been studied and evaluated in this section. The systematic uncertainties have been presented as example only for one observable, the per-trigger near side yield, $\langle N_{\mathrm{assoc,\,near\ side}} \rangle$. However, the estimation of the systematic uncertainties has been performed also for all other observables independently. The combined systematic uncertainty is applied to all analysis results presented in Chapter 8.

	Systematic uncertainty of the per-trigger near-side yield					
	$\sqrt{s}$ = 0.9 TeV		$\sqrt{s}$ = 2.76 TeV		$\sqrt{s}$ = 7.0 TeV	
	$N = 2$	$N = 2 \cdot \langle N \rangle$	$N = 2$	$N = 2 \cdot \langle N \rangle$	$N = 2$	$N = 2 \cdot \langle N \rangle$
Signal extraction	0.3 %	0.1 %	0.2 %	0.1 %	0.5 %	0.1 %
Bin width	0.2 %	0.1 %	0.2 %	0.1 %	0.2 %	0.1 %
Correction procedure	1.9%	0.9 %	5.0 %	3.0 %	12.8 %	1.2 %
Event generator	1.1 %	1.8 %	1.8 %	2.0 %	1.9 %	0.1 %
Transport Monte Carlo	value at 7 TeV	value at 7 TeV	value at 7 TeV	value at 7 TeV	0.3 %	0.1 %
Track cut	15.0 %	2.5 %	16.9 %	2.3 %	10.6 %	2.0 %
Vertex cut	2.7 %	0.5 %	1.5 %	-	2.1 %	-
Detector efficiency	3.0 %	3.0 %	4.1 %	4.1 %	4.1 %	4.1 %
Material budget	0.4 %	0.3 %	0.4 %	0.3 %	0.4 %	0.3 %
MC Particle composition	2.0 %	1.0 %	2.1 %	1.3 %	2.0 %	1.5 %
Pileup	-	-	-	-	5.0 %	1.0 %
Strangeness corr.	-	-	-	-	-	-
Strangeness corr. extrapolation	-	-	2 %	-	2 %	-

Table 7.1.: Systematic uncertainties for the per-trigger near side yield ($p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$, $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$) as example for all observables. The systematic uncertainties of the other observables have similar values.

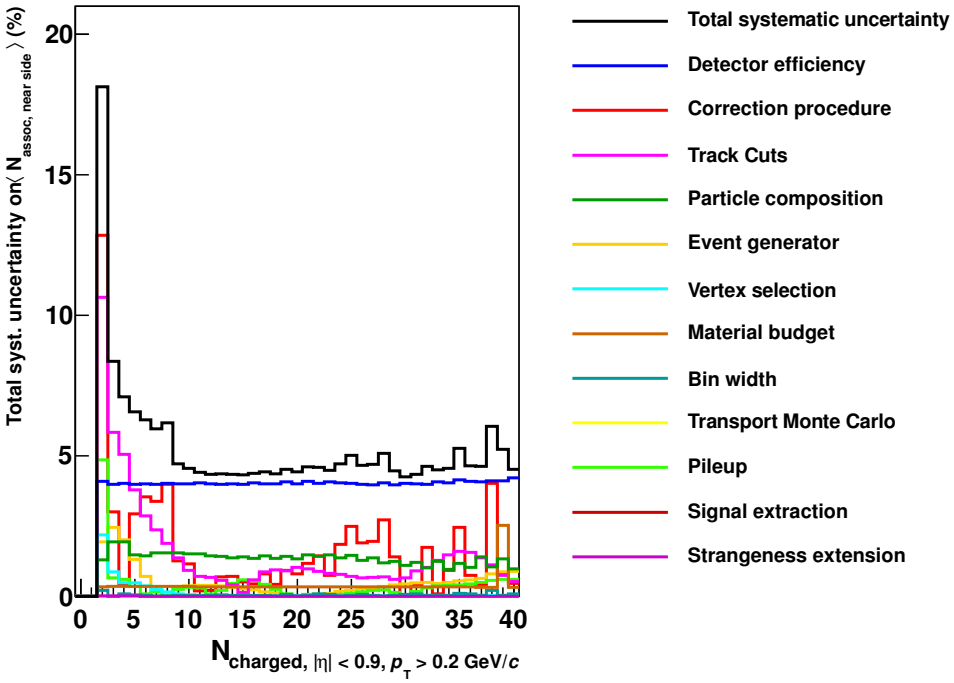

Figure 7.22.: Systematic uncertainty of the per-trigger near side yield measured at $\sqrt{s}$ = 7.0 TeV and its contributions. The detector efficiency uncertainty is the largest contribution to the total systematic uncertainty for almost all charged particle multiplicities. At low charged particle multiplicities, the track selection uncertainty and the uncertainty of the correction procedure represent the largest contributions.

8. Results

In this chapter, the results of the two-particle angular correlation analysis is presented. For the collision data recorded by ALICE, the correction procedure described in Chapter 6 is applied on the measured events and tracks. The corrected track and event properties represent the properties of charged primary particles and INEL $> 0_{|\eta|<0.9}$ events. For collision data generated by Monte Carlo event generators, also the properties of primary charged particles and events of the event class INEL $> 0_{|\eta|<0.9}$ are analyzed.
The uncertainties of the data points are subdivided into the statistical uncertainty and the systematic uncertainty.
The statistical uncertainty is given by the uncertainty of the signal extraction procedure using the fit function of Equation 5.4 on page 72. In the result figures, the statistical uncertainties are shown as error bars connected to the data points in the color of the data points.
The total systematic uncertainty is the combination of all systematic uncertainties discussed in Chapter 7, e. g. the ITS-TPC reconstruction efficiency uncertainty, the uncertainty of the correction procedure, and the track cut dependence uncertainty. In the result figures, the total systematic uncertainty is shown by boxes around the data points. The systematic uncertainty can be subdivided into a correlated fraction, which is of the same relative height for all data points, and an uncorrelated fraction. For a better visibility, the fraction of the systematic uncertainty that is correlated for all data points is shown by single boxes at the left side of the figures instead of presenting the uncertainty at each data point. The error bars in x-direction represent the bin widths.
In Section 8.1, the per-trigger yield as a function of $\Delta\varphi$, also called the azimuthal correlation, is presented for a few example charged particle multiplicities for ALICE data at $\sqrt{s}$ = 0.9, 2.76 and 7.0 TeV in comparison to Monte Carlo predictions. Also, a comparison of the azimuthal correlation for ALICE data measured at the three different center-of-mass energies to each other is presented.
In Section 8.2, the derived observables of the azimuthal correlation are shown. These are the near side yield, the away side yield, the yield in the combinatorial background, the average number of trigger particles, and the number of uncorrelated seeds. The derived observables are presented for ALICE data in direct comparison to Monte Carlo predictions for all charged particle multiplicities. Here, also a comparison of results obtained at different collision energies is discussed. Also, the dependence of the derived observables of the azimuthal correlation on the choice of the transverse momentum threshold is discussed.
Finally, in Section 8.3, the observable $\langle N_{\text{uncorrelated seeds}} \rangle$ is tested for a linear dependence on the charged particle multiplicity.

8.1. Per-Trigger Yield as a Function of $\Delta\varphi$

In this section, the total per-trigger yield as a function of $\Delta\varphi$ estimated in two-particle angular correlations is presented. In Section 8.1.1, the ALICE data are compared to MC predictions. The comparison is performed for the center-of-mass energies $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV separately. In Section 8.1.2, the ALICE results obtained at the three different collisions energies are directly compared to each other.

8.1.1. Comparison of ALICE Results to Model Predictions

In this section, the total per-trigger yield as a function of $\Delta\varphi$ estimated in two-particle angular correlations of ALICE data is presented in comparison to Monte Carlo predictions. The distributions are presented for six example bins in the charged particle multiplicity for each of the center-of-mass energies $\sqrt{s}$ = 7.0, 2.76, and 0.9 TeV starting with the highest collision energy. The decomposed properties of the azimuthal correlation for all charged particle multiplicities are presented in Section 8.2.1. The ALICE results are compared to Phojet (version 1.12), Pythia8 tune 4C, Pythia6 tune Perugia-0, and Pythia6 tune Perugia-2011.

The upper left panel of Figure 8.1 shows the per-trigger yield as a function of $\Delta\varphi$ measured at $\sqrt{s}$ = 7.0 TeV for events with a charged particle multiplicity of $N_{\mathrm{ch}} = 5$. As momentum thresholds used in the correlation, $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ and $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$ are selected.

The azimuthal correlations of ALICE data and MC data reveal a near side peak at $\Delta\varphi = 0$ corresponding mainly to particles from the fragmentation of jets, and an away side yield at $\Delta\varphi = \pi$ corresponding to particles from the fragmentation of recoiling jets. The near side peak and the away side peak lie on top of a constant background representing particles generated in uncorrelated particle production processes.

The minima of the distributions located at approximately $\Delta\varphi = \pm\pi/2$ measured in ALICE data and MC data are roughly at the same heights. However, all Monte Carlo generators show a slightly lower minimum than the ALICE results. Pythia6 Perugia-2011 underestimates the minimum less than the other MC models. This is due to the fact that in this Pythia6 tune the underlying event has been increased compared to older Pythia6 tunes based on input from the LHC at $\sqrt{s}$ = 7.0 TeV.

Phojet, Pythia6 Perugia-0, and Pythia8 4C overestimate the near side peak, while they underestimate the away side peak. However, when taking into account the different constant backgrounds in ALICE data and in the Monte Carlo predictions, also the integrated yield of the away side peak is slightly overestimated by Phojet and Pythia8 4C.

In contrast to the other models, Pythia6 Perugia-2011 reproduces the near side peak within the systematic uncertainties of the ALICE results. However, integrated over all $\Delta\varphi$-bins in the near side peak and taking into account the slightly lower constant background, the

integrated yield of the near side peak is slightly overestimated, too. Also, the away side peak is described well by Perugia-2011, however, for many bins in $\Delta\varphi$ the peak is slightly underestimated resulting in an underestimation of the integrated yield of the away side peak.
With increasing charged particle multiplicity (subsequent panels of Figure 8.1) and therefore with increasing integrated per-trigger yield, the relative differences between the ALICE results and the Monte Carlo predictions shrink. Still, none of the Monte Carlo generators achieves a sufficient description of all properties of the azimuthal correlation altogether. In comparison to the other event generators, Pythia6 Perugia-2011 gives the best description of all correlation properties. Only the integrated yield of the away side peak is better described by Pythia6 Perugia-0.
Figure 8.2 and Figure 8.3 present the per-trigger yield as a function of $\Delta\varphi$ for the center-of-mass energy $\sqrt{s}$ = 2.76 and 0.9 TeV, respectively. For the lower center-of-mass energies, the overall agreement between ALICE data and Monte Carlo predictions is better as compared to $\sqrt{s}$ = 7.0 TeV. This is expected as all Monte Carlo generators have been tuned to measurements at energies more close to $\sqrt{s}$ = 2.76 and 0.9 TeV. Only Pythia6 Perugia-2011 has been tuned using input from LHC data at $\sqrt{s}$ = 7.0 TeV.

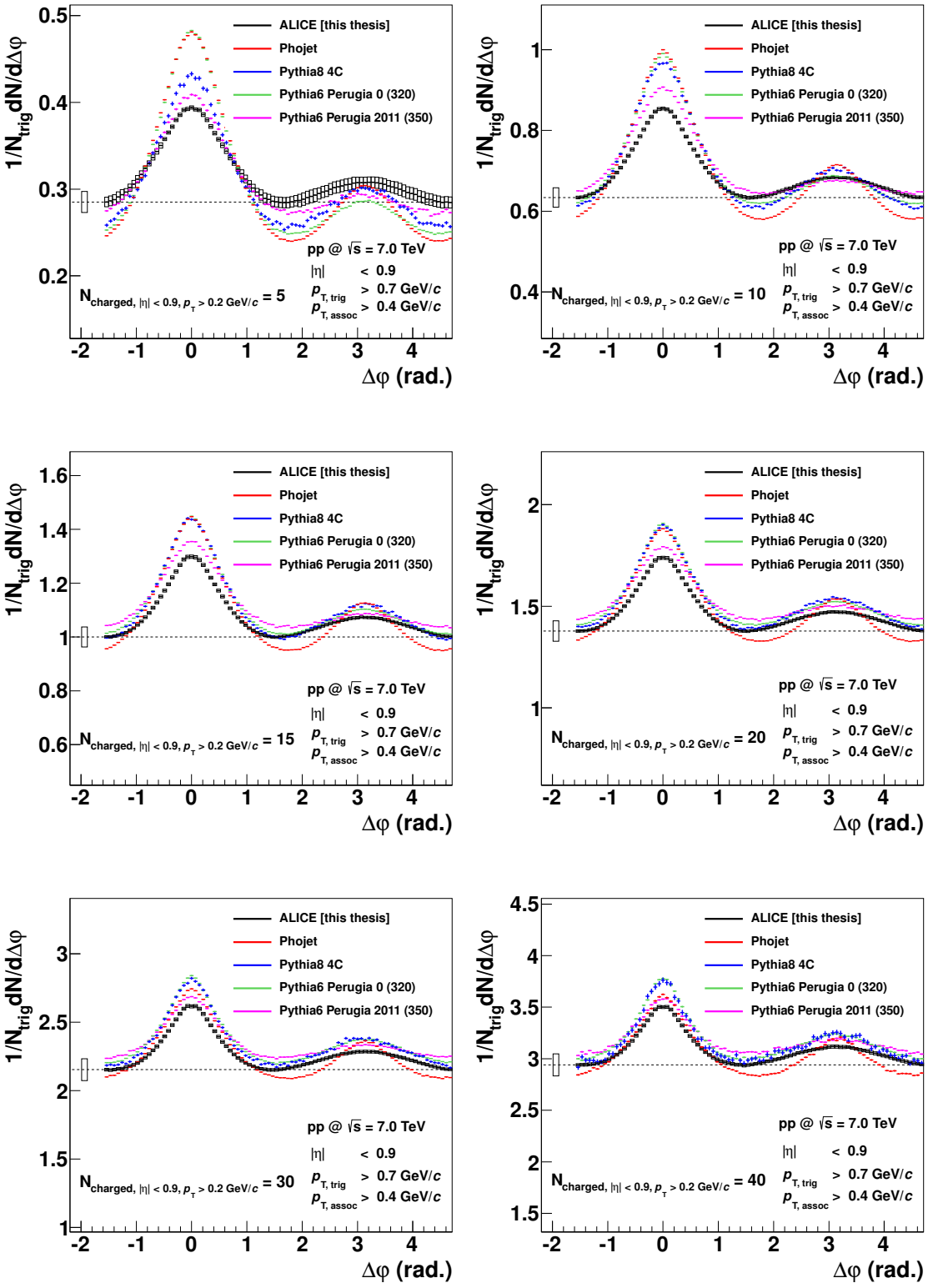

Figure 8.1.: The per-trigger yield distribution as a function of $\Delta\varphi$ for the charged particle multiplicity bins $N_{\mathrm{ch}} = 5$, 10, 15, 20, 30, and 40 at $\sqrt{s} = 7.0\,\mathrm{TeV}$. The fraction of the systematic uncertainty that has the same relative height for all $\Delta\varphi$-bins is presented as a box at the left side of all data points. The box has the absolute height corresponding to the leftmost data point at $\Delta\varphi = -\pi/2$ and it has to be scaled for all other data points according to their heights.

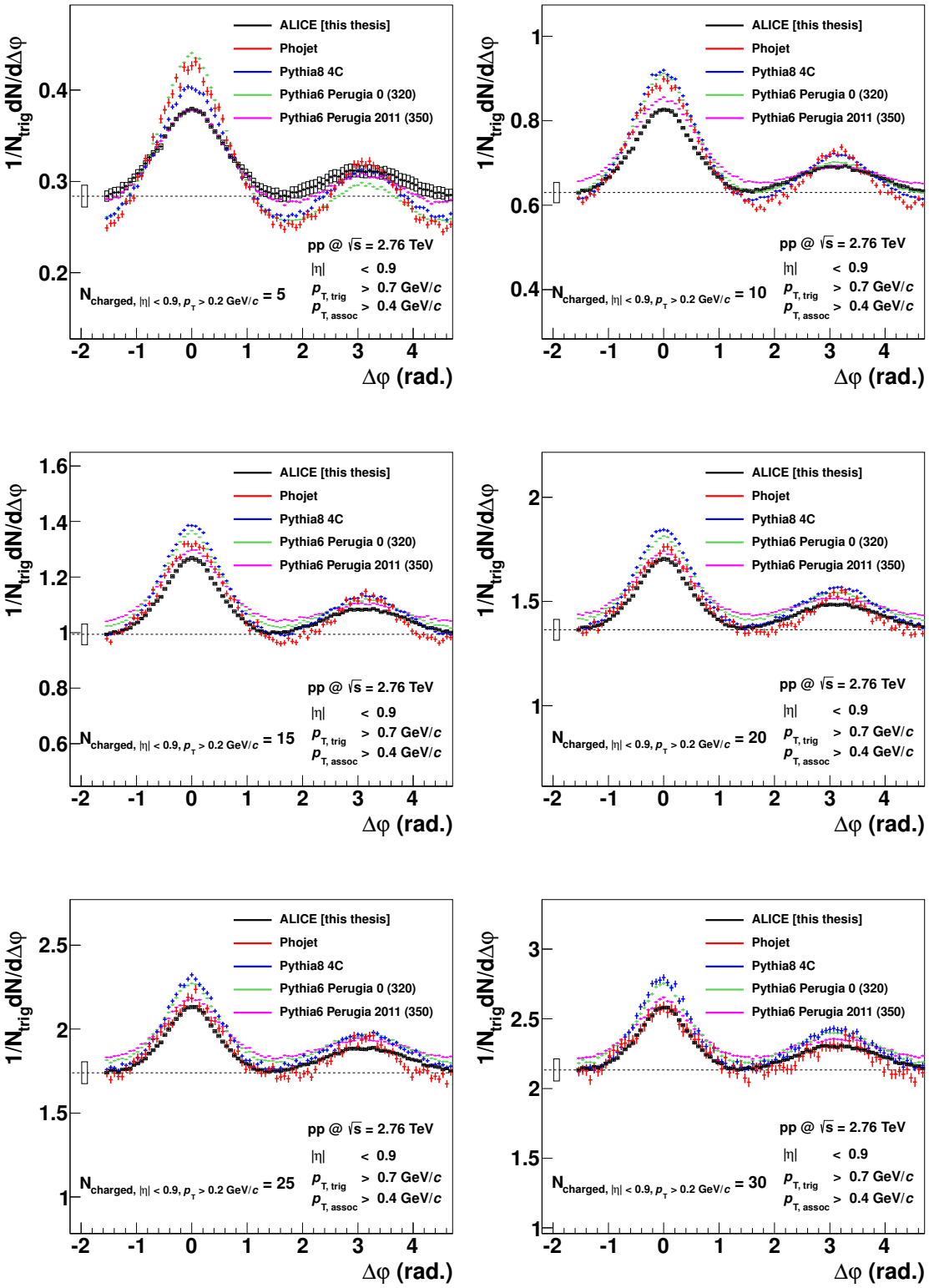

Figure 8.2.: The per-trigger yield distribution as a function of $\Delta\varphi$ for the charged particle multiplicity bins $N_{\mathrm{ch}} = 5$, 10, 15, 20, 25, and 30 at $\sqrt{s} = 2.76\,\mathrm{TeV}$. The fraction of the systematic uncertainty that has the same relative height for all $\Delta\varphi$-bins is presented as a box at the left side of all data points. The box has the absolute height corresponding to the leftmost data point at $\Delta\varphi = -\pi/2$ and it has to be scaled for all other data points according to their heights.

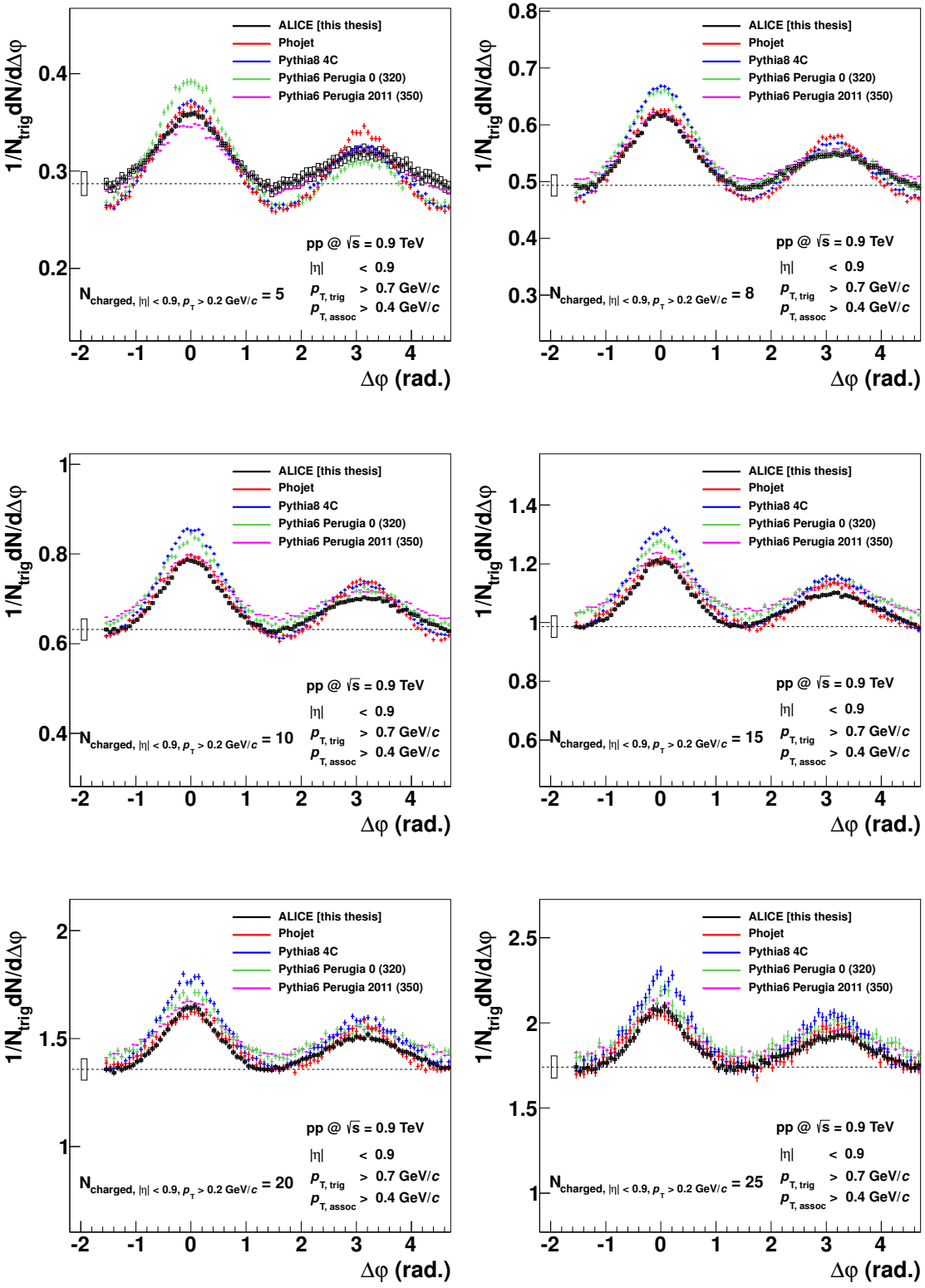

Figure 8.3.: The per-trigger yield distribution as a function of $\Delta\varphi$ for the charged particle multiplicity bins $N_{\mathrm{ch}} = 5$, 8, 10, 15, 20, and 25 at $\sqrt{s} = 0.9\,\mathrm{TeV}$. The fraction of the systematic uncertainty that has the same relative height for all $\Delta\varphi$-bins is presented as a box at the left side of all data points. The box has the absolute height corresponding to the leftmost data point at $\Delta\varphi = -\pi/2$ and it has to be scaled for all other data points according to their heights.

8.1.2. Comparison of ALICE Results at Different Center-of-Mass Energies

In this section, the per-trigger yield as a function of $\Delta\varphi$ estimated in two-particle angular correlations of ALICE data is presented for three different center-of-mass energies.
The distributions are presented for four example values of the charged particle multiplicity.
The properties of the decomposed azimuthal correlations for all analyzed charged particle multiplicities are presented in Section 8.2.2.

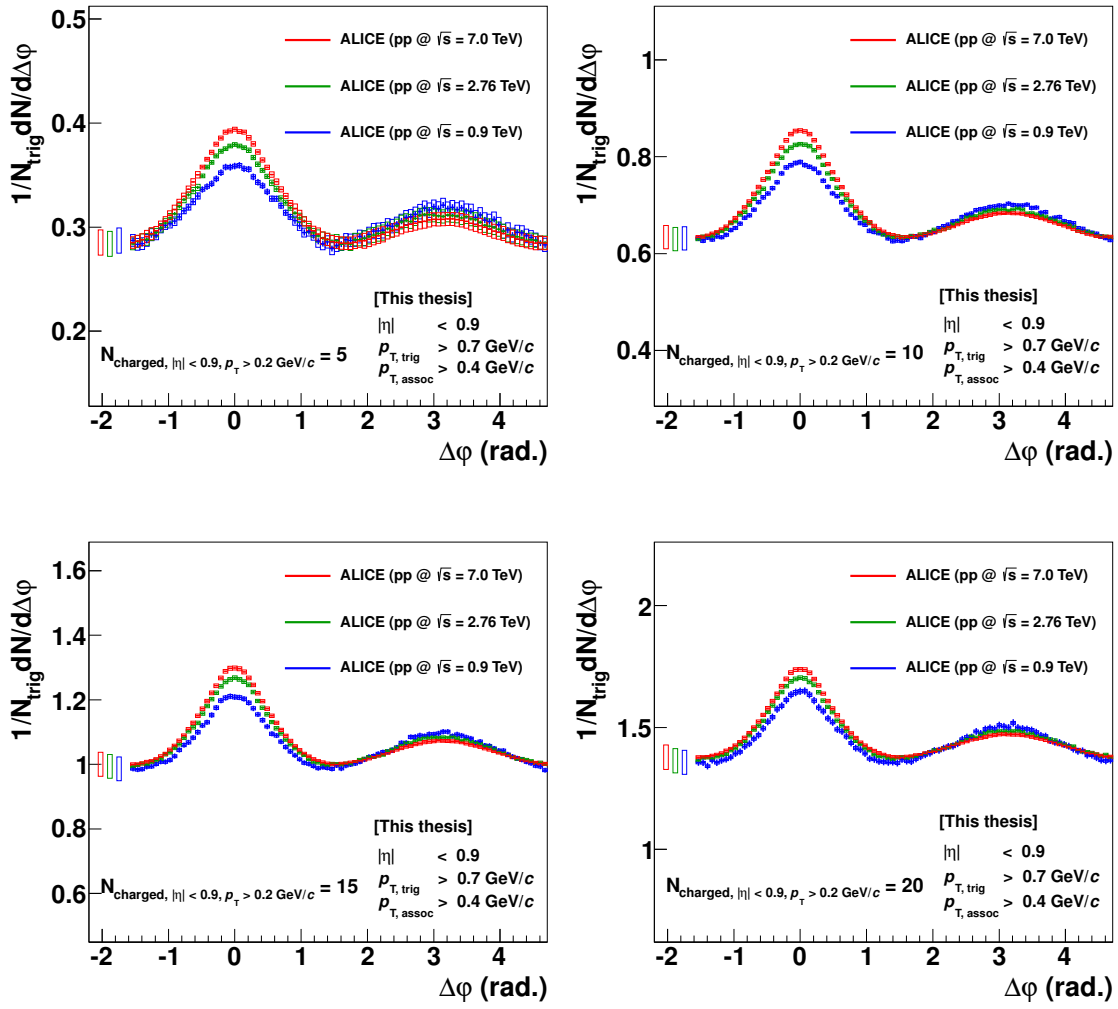

Figure 8.4.: The per-trigger yield distribution as a function of $\Delta\varphi$ for the charged particle multiplicity bins $N_{\mathrm{ch}} = 5$, 10, 15, 20, at $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV. The fraction of the systematic uncertainty that has the same relative height for all $\Delta\varphi$-bins is presented as a box at the left side of all data points. The box has the absolute height corresponding to the leftmost data point at $\Delta\varphi = -\pi/2$ and it has to be scaled for all other data points according to their heights.

Figure 8.4 shows the per-trigger yield as a function of $\Delta\varphi$ for the charged particle multiplicity bins $N_{\mathrm{ch}} = 5$, 10, 15, and 20. As momentum thresholds used in the two-particle correlation, $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ and $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$ are selected.
The per-trigger yields measured at a fixed charged particle multiplicity have slightly modified properties for the three different collision energies. Only the isotropically distributed background, $\langle N_{\mathrm{isotrop}} \rangle$, approximately described by the yield below the yield minimum at $\Delta\varphi = \pm\pi/2$, is about identical for all three center-of-mass energies. The near side peak increases as a function of the center-of-mass energy while the away side peak decreases. The increasing near side peak can be interpreted as a rise in the charged particle multiplicity per jet with increasing center-of-mass energy. The decreasing away side peak can be explained by the fact, that at higher center-of-mass energies it is more likely that recoiling jets fall out of the ALICE acceptance.

8.2. Integrated Per-Trigger Yield

In this section, the derived observables of the azimuthal correlations are presented. This includes the near side yield, the away side yield, the yield in the combinatorial background, the average number of trigger particles, and the number of uncorrelated seeds. All observables are presented as a function of the charged particle multiplicity. The analysis results are obtained using transverse momentum cuts of $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ and $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$ as well as $p_{\mathrm{T,\,assoc}} > 0.7\,\mathrm{GeV}/c$. The ALICE results are compared to Phojet (version 1.12), Pythia8 tune 4C, Pythia6 tune Perugia-0, and Pythia6 tune Perugia-2011.
In Section 8.2.1, the ALICE results are directly compared to the MC predictions. The comparison is performed for data sets measured at the center-of-mass energies $\sqrt{s}$ = 7.0, 2.76, and 0.9 TeV starting with the highest collision energy. In Section 8.2.2, the ALICE results obtained at the three different collisions energies are directly compared to each other. Also, the MC predictions for the three collision energies are compared to each other. This comparison reveals the collision energy scaling in ALICE data and MC data.

8.2.1. Comparison of ALICE Results to Model Predictions

In this section, the decomposed properties of the azimuthal correlations are presented for ALICE data in comparison to Monte Carlo predictions. First the results at $\sqrt{s}$ = 7.0 TeV are discussed.
The top row of Figure 8.5 shows the per-trigger near side yield above the combinatorial background, $\langle N_{\mathrm{assoc,\,near\ side}}\rangle$, at $\sqrt{s}$ = 7.0 TeV as a function of the charged particle multiplicity for $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ and $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$ in the left panel and for $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ and $p_{\mathrm{T,\,assoc}} > 0.7\,\mathrm{GeV}/c$ in the right panel. The per-trigger near side yield rises as a function of the charged particle multiplicity for ALICE data and all Monte Carlo generators. At charged particle multiplicities below $N_{\mathrm{ch}} = 20$ and for both $p_{\mathrm{T,\,assoc}}$-thresholds, the near side yield, $\langle N_{\mathrm{assoc,\,near\ side}}\rangle$, is overestimated by 50 % to 100% by all event generators except Pythia6 Perugia-2011. Perguia6 Perugia-2011 overestimates the $\langle N_{\mathrm{assoc,\,near\ side}}\rangle$ by less than 40%. At charged particle multiplicities above $N_{\mathrm{ch}} = 20$, Pythia6 Perugia-2011 agrees with the ALICE results. In this multiplicity region, all other event generators overestimate the per-trigger near side yield by some 50%.
The bottom row of Figure 8.5 shows the per-trigger away side yield above the combinatorial background, $\langle N_{\mathrm{assoc,\,away\ side}}\rangle$, at $\sqrt{s}$ = 7.0 TeV as a function of the charged particle multiplicity for $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ and $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$ in the left panel and for $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ and $p_{\mathrm{T,\,assoc}} > 0.7\,\mathrm{GeV}/c$ in the right panel. Just as for the near side yield, the per-trigger away side yield rises as a function of the charged particle multiplicity. For $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$, Pythia6 Perugia-0 gives the best description of the ALICE results, while at $p_{\mathrm{T,\,assoc}} > 0.7\,\mathrm{GeV}/c$, Pythia6 Perugia-0 slightly overestimates the per-

trigger away side yield. Pythia6 Perugia-2011 underestimates the per-trigger away side yield by up to 50%. Phojet and Pythia8 4C both overestimate it by 50 % to 80%.
The top row of Figure 8.6 shows the per-trigger yield in the combinatorial background, $\langle N_{\mathrm{isotrop}} \rangle$, at $\sqrt{s} = 7.0\,\mathrm{TeV}$ as a function of the charged particle multiplicity for $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ and $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$ in the left panel and for $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ and $p_{\mathrm{T,\,assoc}} > 0.7\,\mathrm{GeV}/c$ in the right panel. All Monte Carlo generators agree within the systematic uncertainties with the ALICE results for $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$. For $p_{\mathrm{T,\,assoc}} > 0.7\,\mathrm{GeV}/c$, only Phojet underestimates the yield in the constant background $\langle N_{\mathrm{isotrop}} \rangle$.
The bottom left panel of Figure 8.6 shows the average number of trigger particles, $\langle N_{\mathrm{trigger}} \rangle$. The increase of the average number of trigger particles with the charged particle multiplicity is slightly steeper than a linear increase. This is due to the known increase of the mean transverse momentum per event, $\langle p_{\mathrm{T}} \rangle$, with the charged particle multiplicity [UA182b, ALI10c] which is also reproduced by all event generators qualitatively. The bottom right panel of Figure 8.6 shows the number of uncorrelated seeds, $\langle N_{\mathrm{uncorrelated\,seeds}} \rangle$. For low and intermediate charged particle multiplicities, $\langle N_{\mathrm{uncorrelated\,seeds}} \rangle$ rises almost linearly with the charged particle multiplicity. At high charged particle multiplicities, however, the rise of $\langle N_{\mathrm{uncorrelated\,seeds}} \rangle$ seems to level off from a linear dependence. This is further discussed in the Section 8.3.
Both $\langle N_{\mathrm{trigger}} \rangle$ and $\langle N_{\mathrm{uncorrelated\,seeds}} \rangle$ are fairly well described by all Pythia tunes, while Phojet underestimates the ALICE results.
Pythia6 Perugia-2011 has been optimized based on minimum bias measurements at the LHC at $\sqrt{s} = 7.0\,\mathrm{TeV}$. Therefore, an improved agreement between ALICE data at $\sqrt{s} = 7.0\,\mathrm{TeV}$ and the tune Perugia-2011 in comparison to the other tunes, which do not incorporate tuning with data up to $\sqrt{s} = 7.0\,\mathrm{TeV}$, is expected.
Figures 8.7 to 8.10 show the same per-trigger yield properties measured at $\sqrt{s} = 2.76\,\mathrm{TeV}$ and $\sqrt{s} = 0.9\,\mathrm{TeV}$. The relations of the per-trigger yield properties between ALICE data and Monte Carlo predictions are similar to the behavior discussed for the results at $\sqrt{s} = 7.0\,\mathrm{TeV}$. With decreasing center-of-mass energy, the overall agreement between the ALICE results and the Monte Carlo predictions increases.

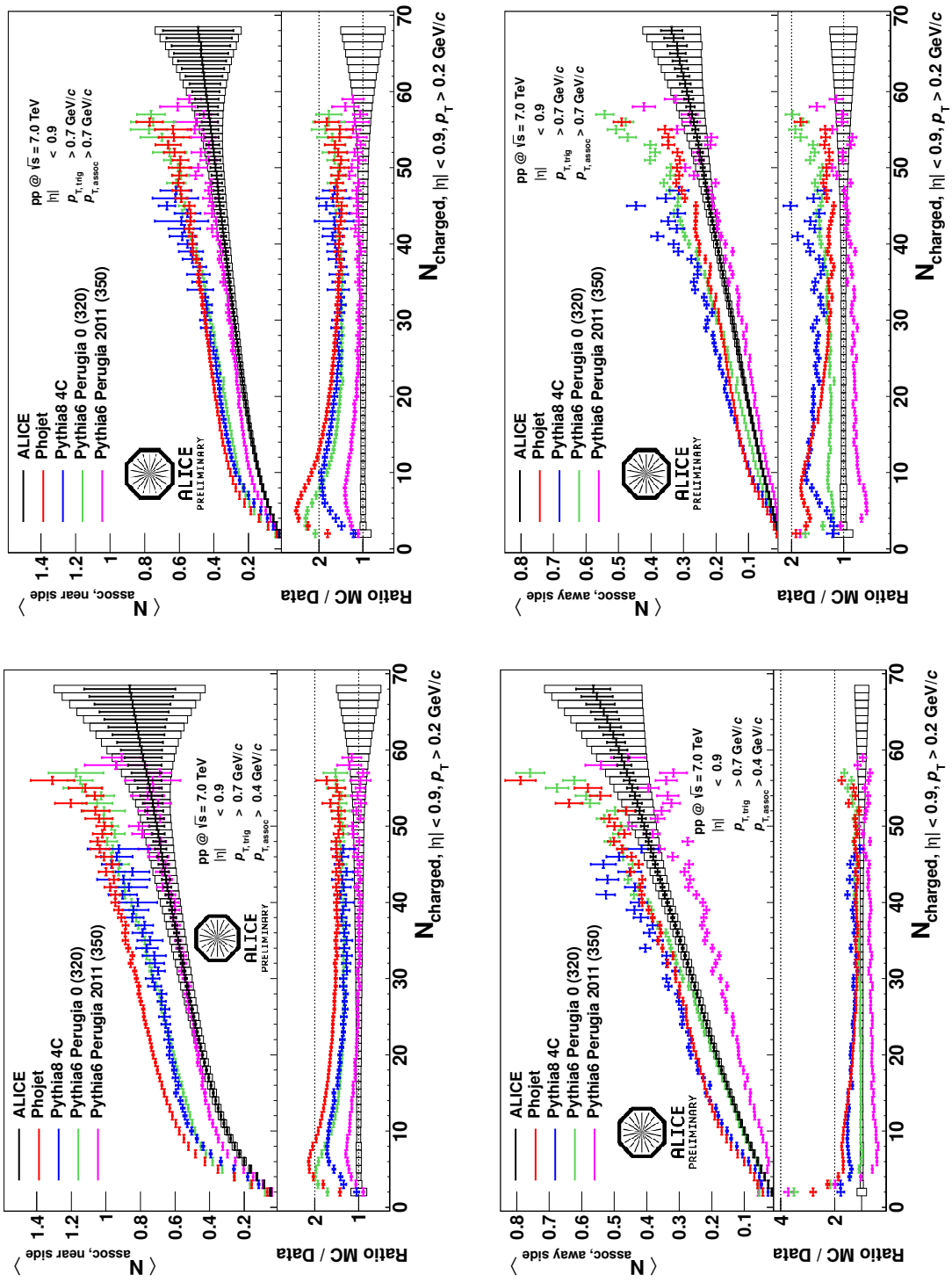

Figure 8.5.: Per-trigger near side yield (top row), per-trigger away side yield (bottom row) measured at $\sqrt{s}$ = 7.0 TeV [Sic12].

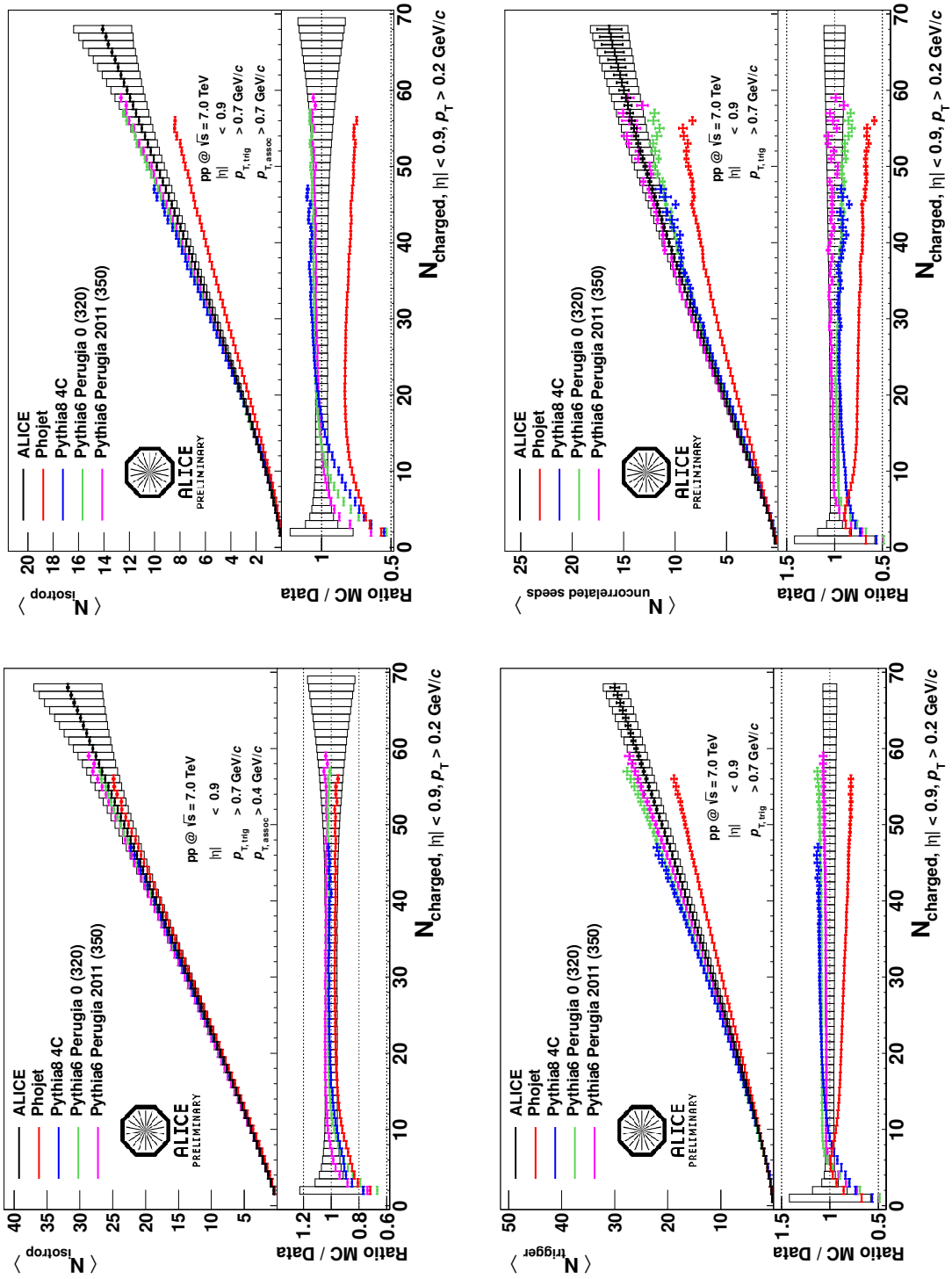

Figure 8.6.: Per-trigger yield in the combinatorial background (top row), average number of trigger particles (bottom row, left panel), and number of uncorrelated seeds (bottom row, right panel) measured at $\sqrt{s}$ = 7.0 TeV [Sic12].

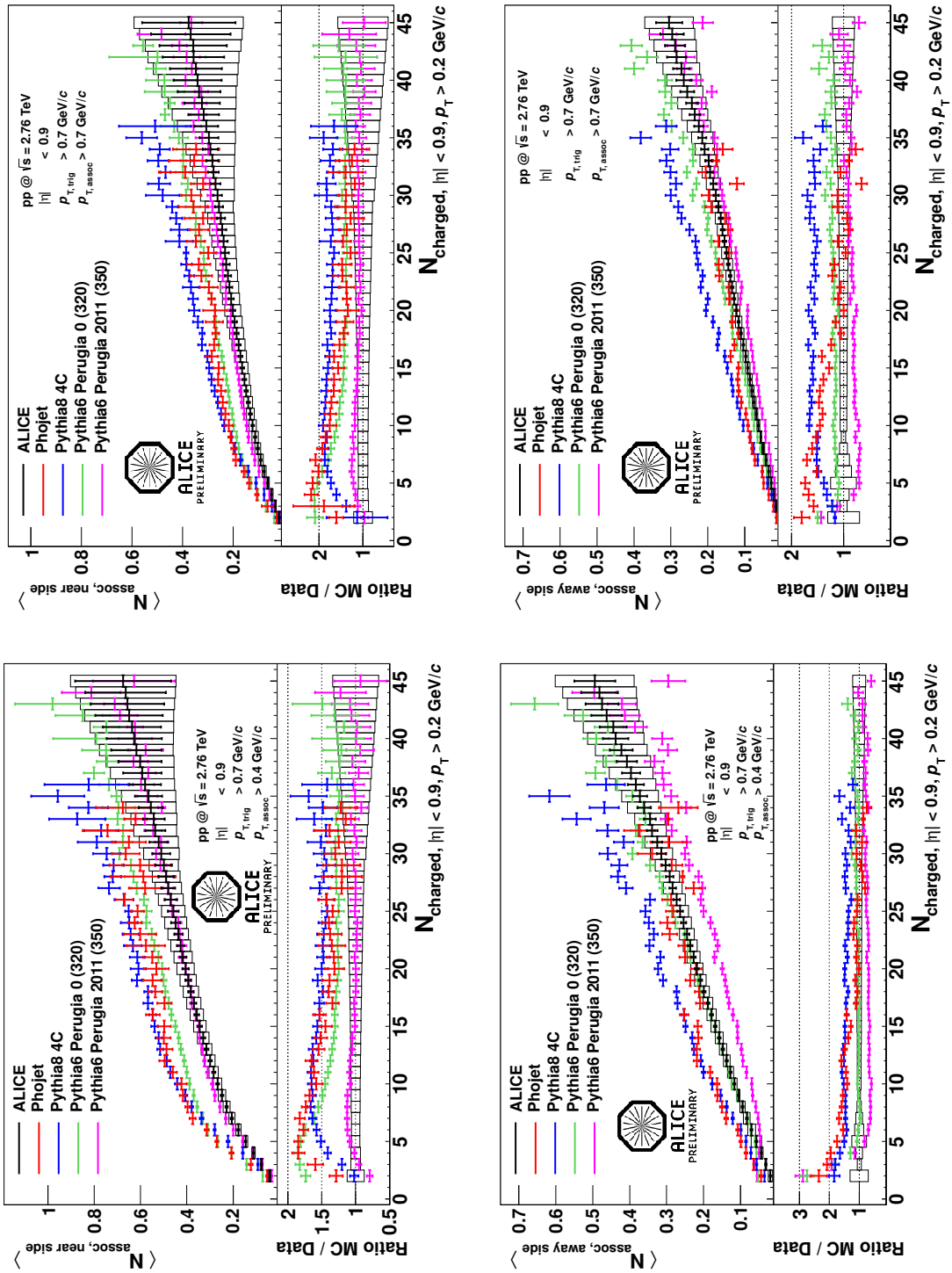

Figure 8.7.: Per-trigger near side yield (top row), per-trigger away side yield (bottom row) measured at $\sqrt{s}$ = 2.76 TeV [Sic12].

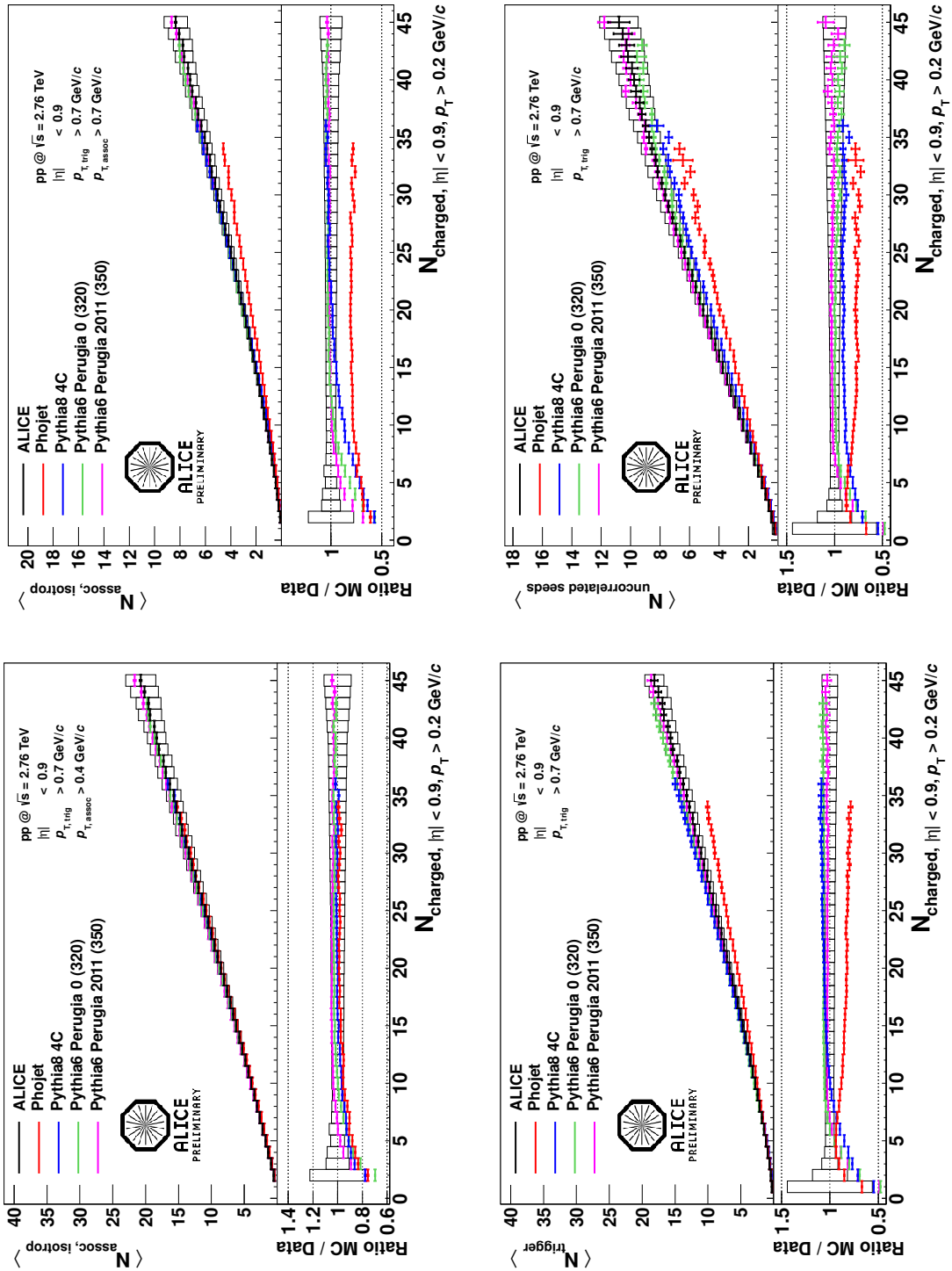

Figure 8.8.: Per-trigger yield in the combinatorial background (top row), average number of trigger particles (bottom row, left panel), and number of uncorrelated seeds (bottom row, right panel) measured at $\sqrt{s}$ = 2.76 TeV [Sic12].

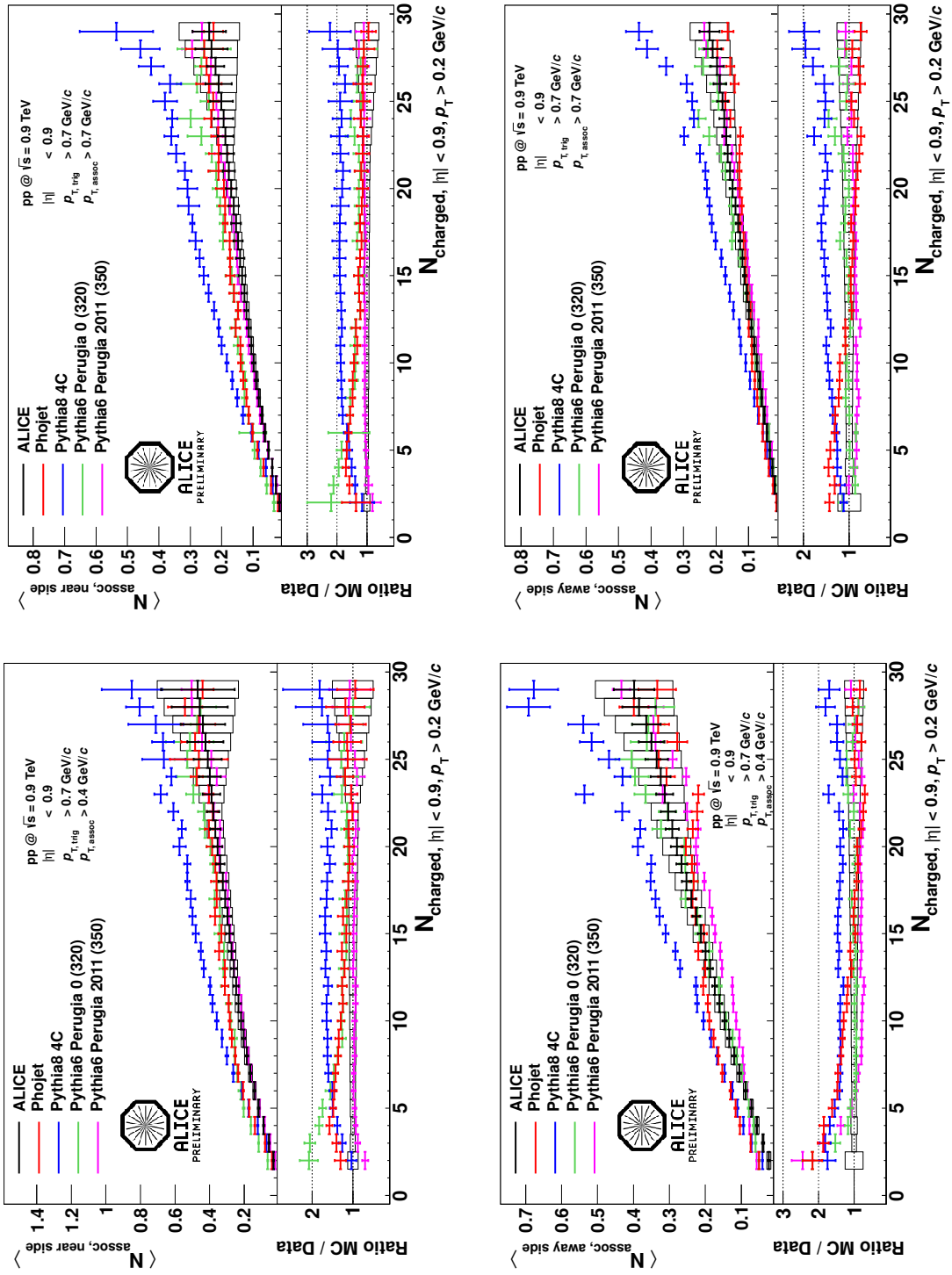

Figure 8.9.: Per-trigger near side yield (top row), per-trigger away side yield (bottom row) measured at $\sqrt{s} = 0.9\,\text{TeV}$ [Sic12].

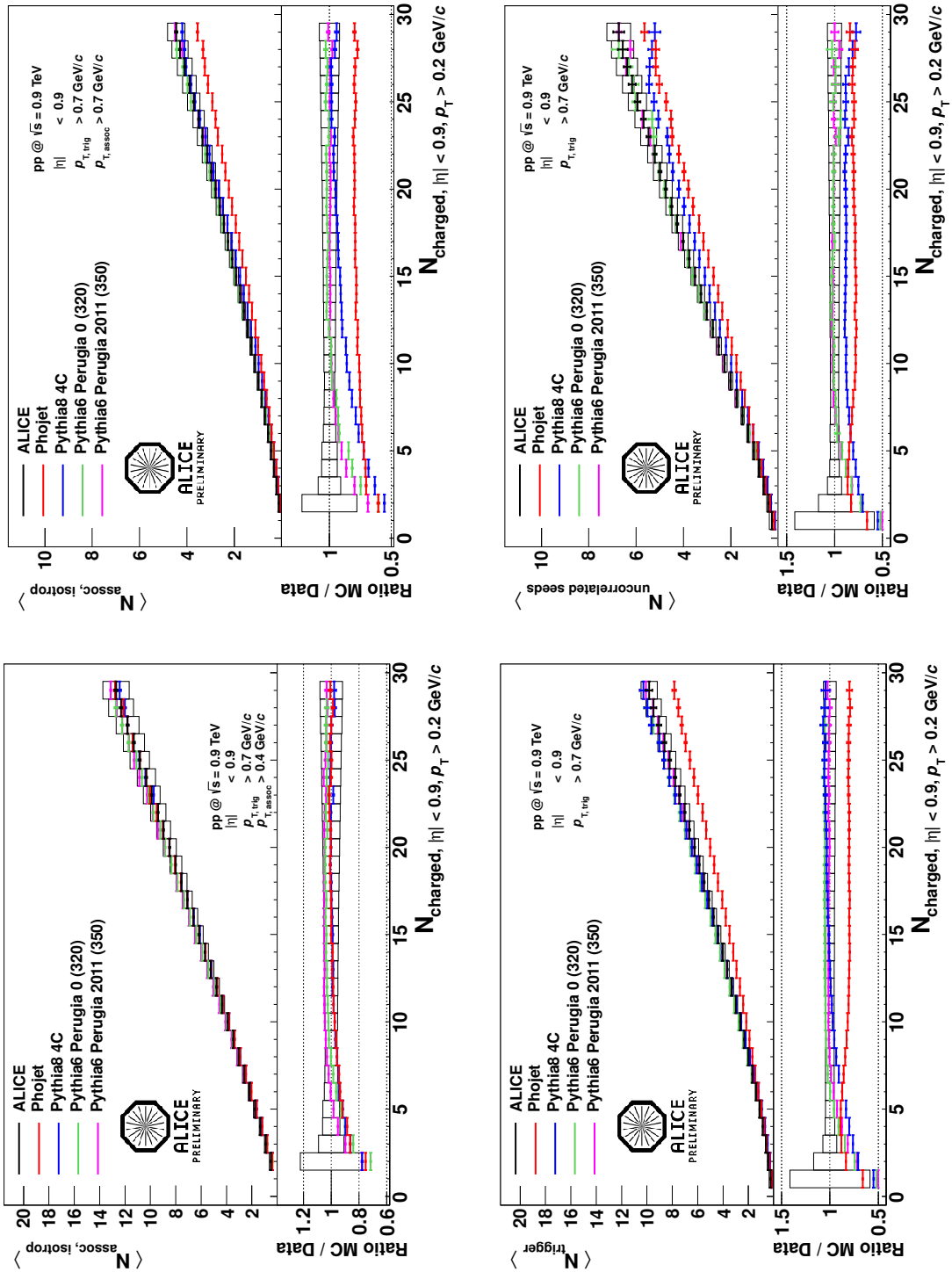

Figure 8.10.: Per-trigger yield in the combinatorial background (top row), average number of trigger particles (bottom row, left panel), and number of uncorrelated seeds (bottom row, right panel) measured at $\sqrt{s}$ = 0.9 TeV [Sic12].

8.2.2. Comparison of ALICE Results at Different Center-of-Mass Energies

In this section, the center-of-mass energy dependence of the two-particle angular correlations is studied for ALICE data as well as for Monte Carlo generator predictions.
Here, the same data points are shown as already shown in the previous section. In the previous section, the ALICE results are compared to MC predictions for only one collision energy each. In this section, in contrast, the three ALICE results measured at $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV are solely compared to each other. In the same manner, the MC predictions estimated for the three collision energies are compared to each other. In these comparisons, it is tested whether the observables are modified as a function of the center-of-mass energy and whether the MC generators reproduce the energy scaling observed in the ALICE results.
The analysis is performed using transverse momentum thresholds of $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ and $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$.
Each of the Figures 8.11 to 8.15 is divided into 5 panels. The top left panels show the ALICE results, the top right panels show the Phojet results, the mid left panels show Pythia6 Perugia-0 results (abbreviated by "Perugia-0"), the mid right panels show Pythia6 Perugia-2011 results (abbreviated by "Perugia-2011"), and the bottom panels show Pythia8 4C results. In each panel, the results of the three center-of-mass energies are presented.
In Figure 8.11, the per-trigger near side yield above the combinatorial background, $\langle N_{\mathrm{assoc,\,near\ side}} \rangle$, and its dependence on the center-of-mass energy is presented. With increasing center-of-mass energy, the near side yield increases in all fixed charged particle multiplicity bins. This can be interpreted as an increase of the intra-jet charged particle multiplicity with the collision energy. The increase of the near side yield is reproduced by all event generators. Phojet shows the largest increase of the near side yield with the center-of-mass energy. Pythia6 Perugia-2011 gives the best description of the ALICE data for all center-of-mass energies.
Figure 8.12 shows the per-trigger away side yield above the combinatorial background, $\langle N_{\mathrm{assoc,\,away\ side}} \rangle$, and its dependence on the center-of-mass energy. The away side yield measured with ALICE data decreases with increasing center-of-mass energy. This is due to the fact, that the probability of recoiling jets being located outside the ALICE acceptance increases with the center-of-mass energy. While all Pythia tunes show a similar decrease as a function of the collision energy, Phojet does not show a clear center-of-mass energy dependence. While all presented away side yields rise as a function of the charged particle multiplicity, the away side yield for Pythia6 Perugia-2011 reveals a minimum at $N_{\mathrm{ch}} \approx 6$ for $\sqrt{s}$ = 2.76 and 7.0 TeV. Pythia6 Perugia-0 gives the best description of $\langle N_{\mathrm{assoc,\,away\ side}} \rangle$ measured in ALICE for all discussed center-of-mass energies.

Figure 8.13 shows the per-trigger yield in the isotropically distributed background, $\langle N_{\mathrm{isotrop}} \rangle$, and its dependence on the center-of-mass energy. $\langle N_{\mathrm{isotrop}} \rangle$ appears to be independent from the center-of-mass energy within the statistical uncertainties. The yield and the center-of-mass energy dependence is reproduced well by all event generators.
In Figure 8.14, the average number of trigger particles per event, $\langle N_{\mathrm{trigger}} \rangle$, given by all particles with $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ is presented. The increase of the $\langle N_{\mathrm{trigger}} \rangle$ distributions as a function of the charged particle multiplicity is slightly steeper than linear. This is due to the afore mentioned increase in $\langle p_{\mathrm{T}} \rangle$ with the multiplicity. In the ALICE results, the average number of trigger particles in a fixed multiplicity bin increases with the center-of-mass energy. This behavior is reproduced by all event generators. While all Pythia tunes slightly overestimate the ALICE results of $\langle N_{\mathrm{trigger}} \rangle$ for all multiplicity bins, Phojet underestimates the ALICE results.
Figure 8.15 shows the number of uncorrelated seeds, $\langle N_{\mathrm{uncorrelated\,seeds}} \rangle$, for the different center-of-mass energies. For low and intermediate charged particle multiplicities, the number of uncorrelated seeds grows approximately linear with the charged particle multiplicity. At high charged particle multiplicities, the increase seems to fall below a linear dependence. The $\langle N_{\mathrm{uncorrelated\,seeds}} \rangle$ measured in ALICE for a fixed bin in the charged particle multiplicity grow slightly with increasing center-of-mass energies. The Pythia6 tunes Perugia-0 and Perugia-2011 agree with the ALICE results within the systematic uncertainties. Pythia8 4C slightly underestimates the ALICE results for all multiplicity bins and center-of-mass energies. Phojet underestimates the ALICE results most. In contrast to the ALICE results and the Pythia predictions, Phojet does not show any dependence of the number of uncorrelated seeds on the center-of-mass energy.

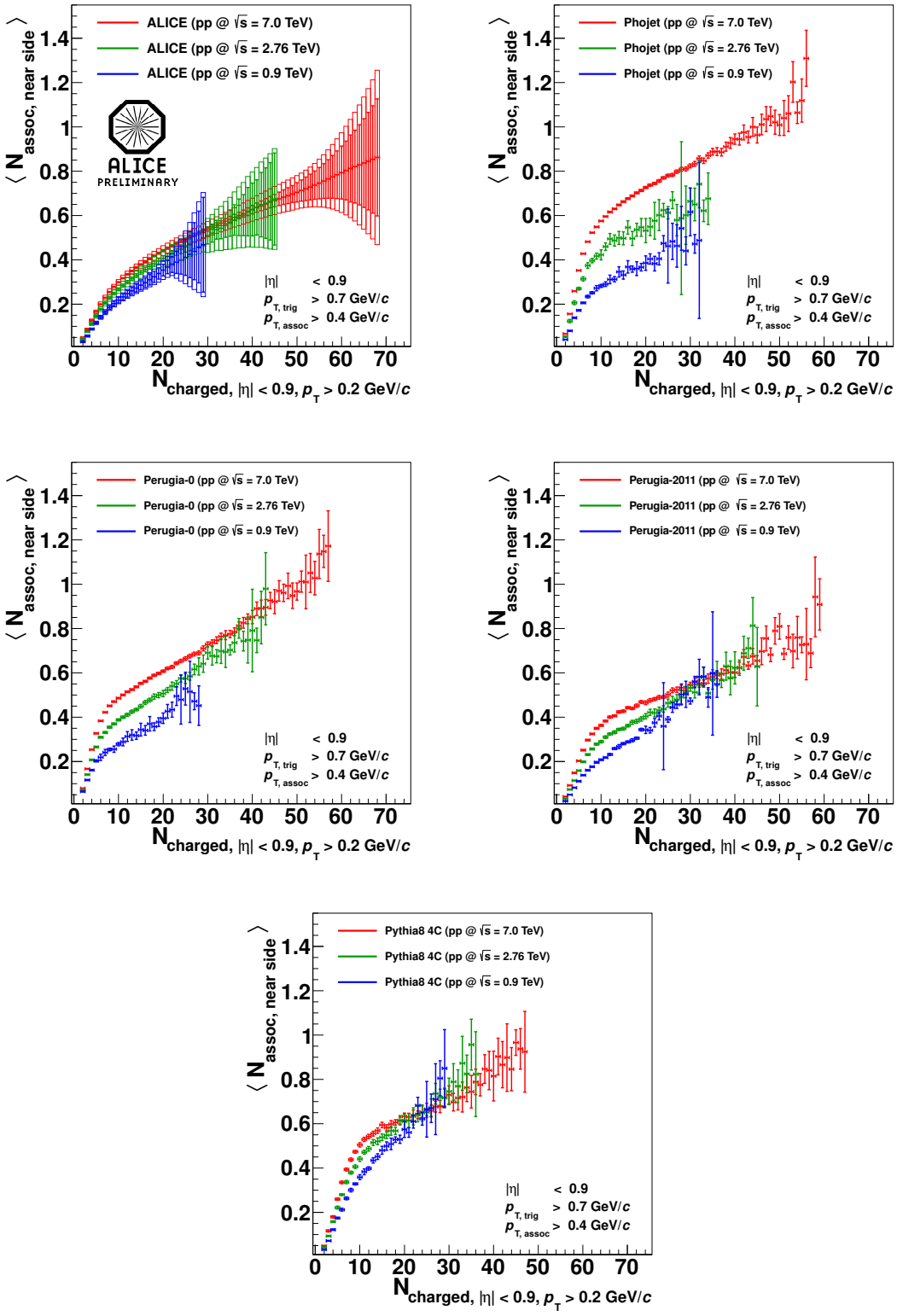

Figure 8.11.: Per-trigger near side yield as a function of the charged particle multiplicity measured for $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV [Sic12].

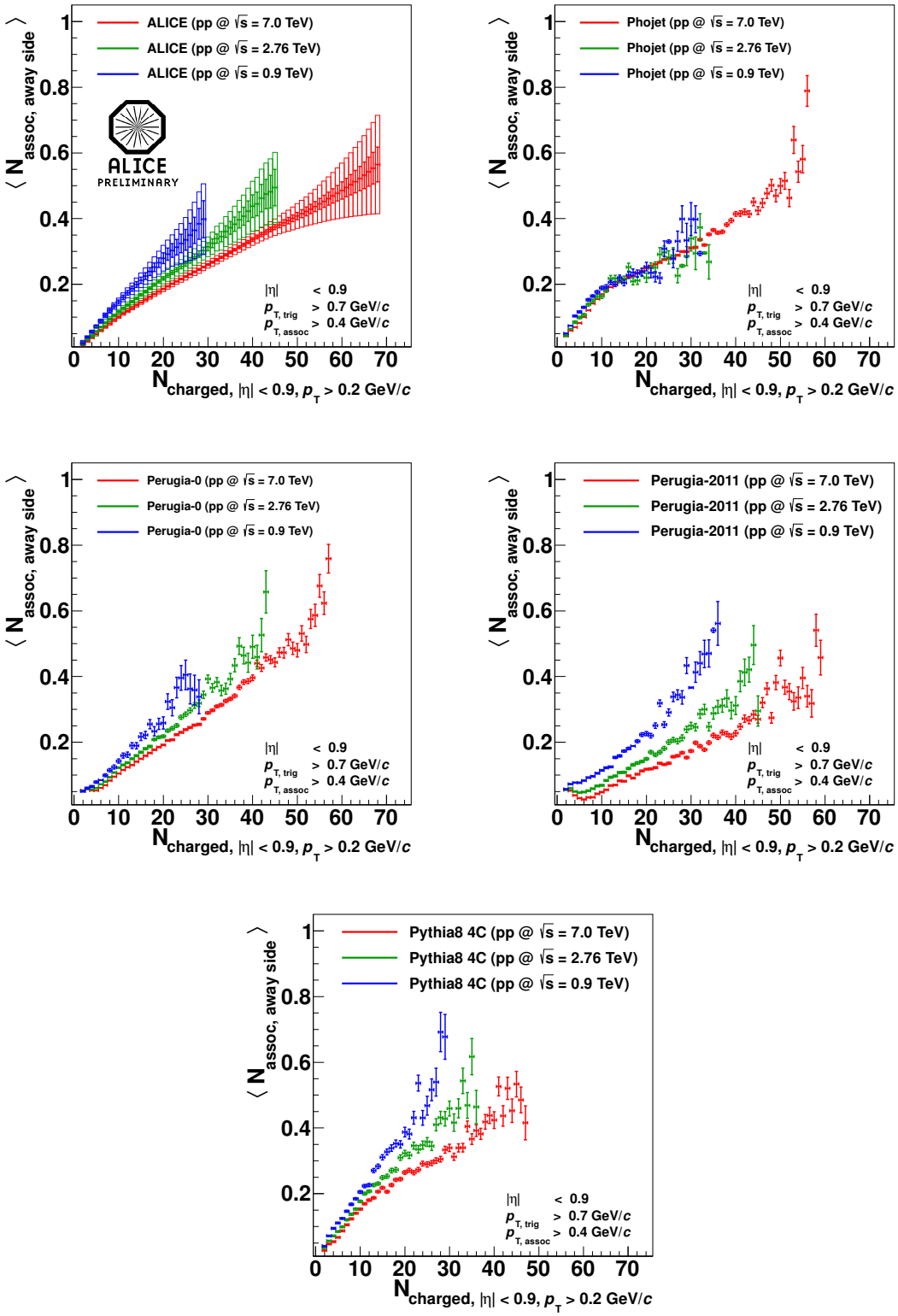

Figure 8.12.: Per-trigger away side yield as a function of the charged particle multiplicity measured for $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV [Sic12].

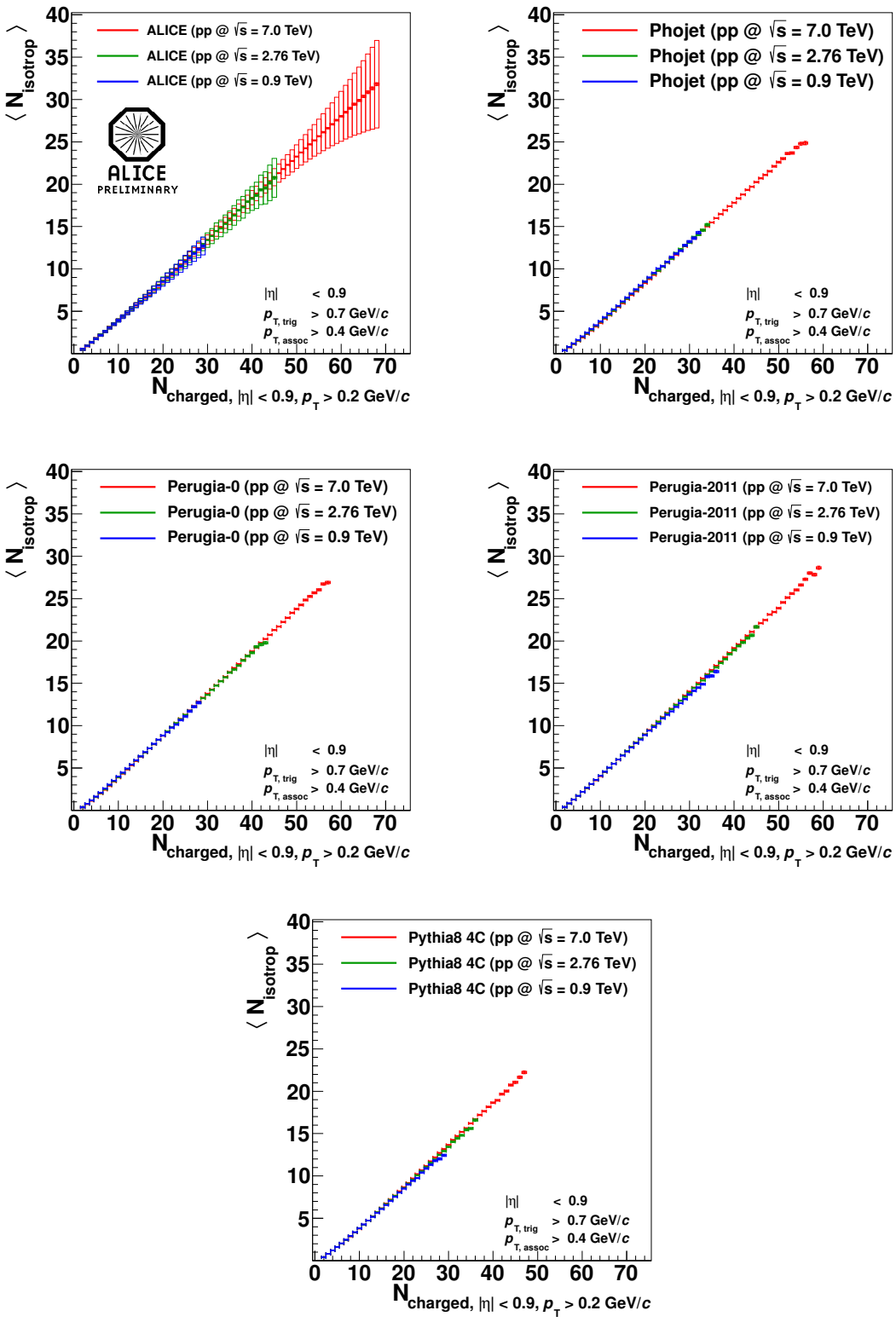

Figure 8.13.: Per-trigger yield in the combinatorial background as a function of the charged particle multiplicity measured for $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV [Sic12].

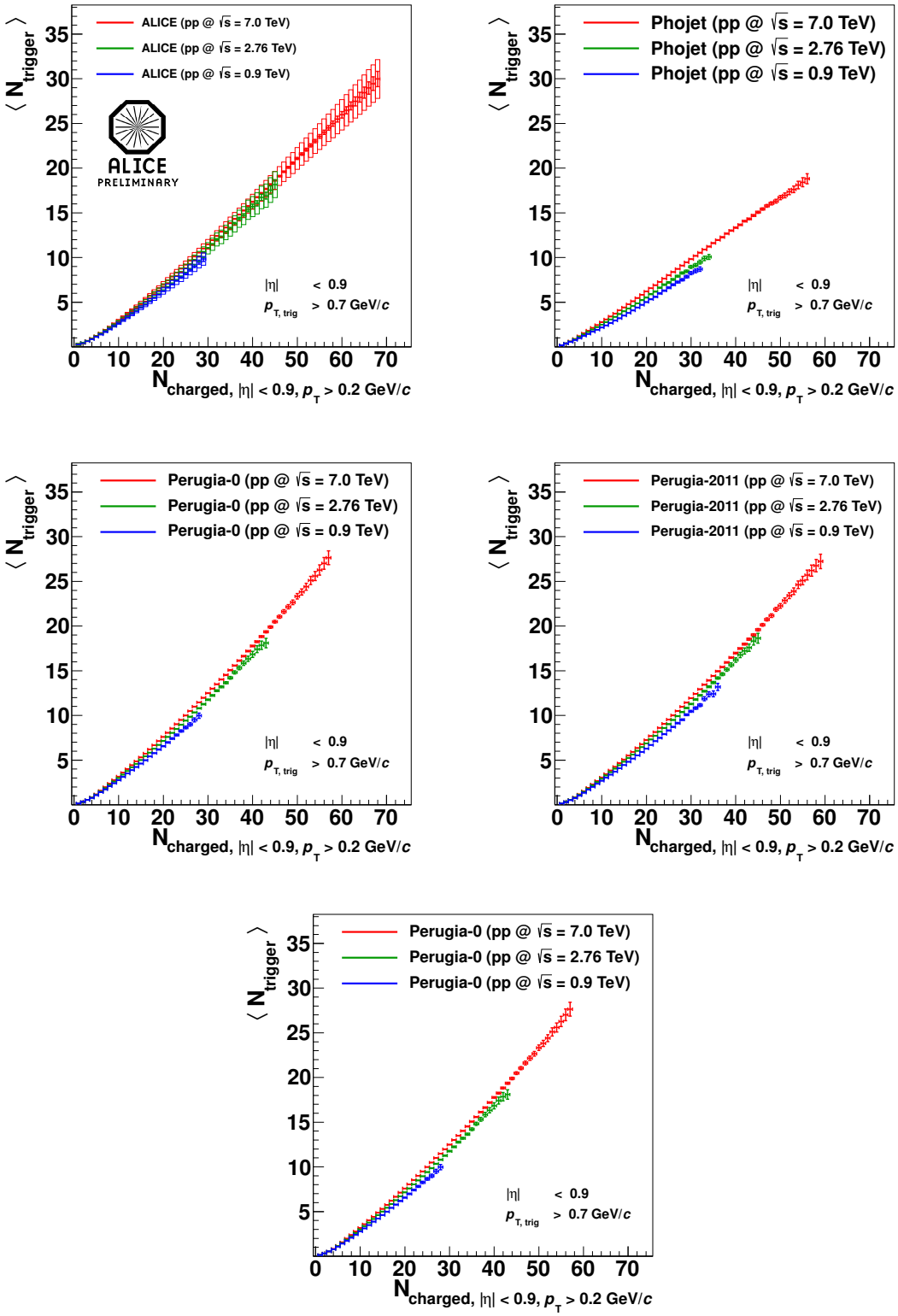

Figure 8.14.: Average number of trigger particles per event as a function of the charged particle multiplicity measured for $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV [Sic12].

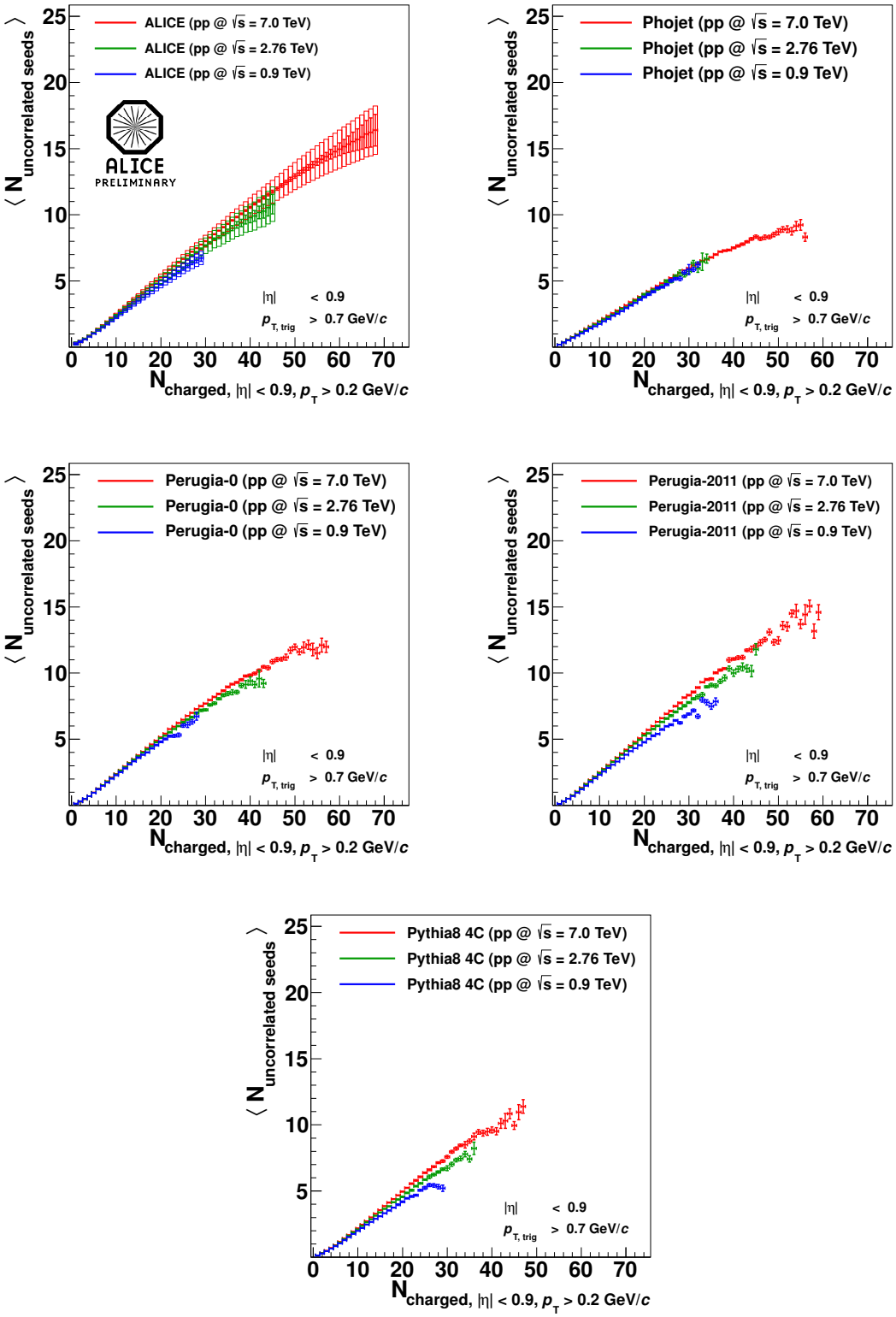

Figure 8.15.: Average number of uncorrelated seeds per event as a function of the charged particle multiplicity measured for $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV [Sic12].

8.2.3. Results at Different Transverse Momentum Thresholds

So far, the analysis results have been estimated using only one transverse momentum threshold for the trigger particles ($p_{\mathrm{T,\,trig}} = 0.7\,\mathrm{GeV}/c$) and two transverse momentum thresholds for associated particles ($p_{\mathrm{T,\,assoc}} = 0.4$ and $0.7\,\mathrm{GeV}/c$). In this section, the dependence on the transverse momentum thresholds is studied further. For simplicity, the transverse momentum thresholds for trigger particles and associated particles are set to the same value.

In the Figures 8.16 to 8.20, the correlation properties $\langle N_{\mathrm{assoc,\,near\ side}}\rangle$, $\langle N_{\mathrm{assoc,\,away\ side}}\rangle$, $\langle N_{\mathrm{isotrop}}\rangle$, $\langle N_{\mathrm{trigger}}\rangle$, and $\langle N_{\mathrm{uncorrelated\,seeds}}\rangle$ are presented as a function of the charged particle multiplicity and as a function of the transverse momentum $p_{\mathrm{T,\,trig}}$ and $p_{\mathrm{T,\,assoc}}$. For $p_{\mathrm{T,\,trig}}$ and $p_{\mathrm{T,\,assoc}}$, the bins $p_{\mathrm{T}} > 0.7$, 0.8, 0.9, 1.0, 1.5, 2.0, 2.5, and $3.0\,\mathrm{GeV}/c$ are used. For each bin in p_{T}, the charged particle multiplicity runs from $N_{\mathrm{ch}} = 0$ to 50.

The transverse momentum thresholds present the lower limit for the particle momentum. No upper limit in the transverse momentum is applied.

The axis of the transverse momentum threshold is shown at the upper edge of the figures. The multiplicity axes for each p_{T}-bin are shown at the lower edge of the figures. The multiplicity axes labels for the different ranges in p_{T} are shown by alternating color. The edge between two bins in p_{T} are marked by vertical dashed lines. For a new bin in p_{T}, the charged particle multiplicity starts again at $N_{\mathrm{ch}} = 0$.

Each of the Figures 8.16 to 8.20 is divided into two sub-figures. The top panel presents the observables itself for ALICE results and Monte Carlo generator predictions, the bottom panel shows the ratio between Monte Carlo predictions and the ALICE results.

The per-trigger near side yield presented in Figure 8.16 decreases with increasing transverse momentum threshold. All but one Monte Carlo generator overestimate the near side yield (between 100 % and 200 %). Pythia6 Perugia-2011 gives the best description (50 % at low N_{ch} and 10 % at high N_{ch}). The overall discrepancy is largest for $p_{\mathrm{T}} > 1.0\,\mathrm{GeV}/c$ to $1.5\,\mathrm{GeV}/c$. At higher transverse momenta of the analyzed p_{T}-range, the discrepancy decreases.

The per-trigger away side yield also decreases as a function of the transverse momentum threshold (Figure 8.17). All MC generators but Pythia6 Perugia-2011 overestimate the away side yield. Phojet and Pythia8 4C do not show any transverse momentum dependence in the relative overestimation of the yield. Pythia6 Perugia-0 shows an increasing discrepancy to the ALICE results with increasing transverse momentum (30 % to 100 %) for the analyzed p_{T}-range. Pythia6 Perugia-2011 underestimates the away side yield slightly at low transverse momenta, it agrees at $p_{\mathrm{T}} > 1.0\,\mathrm{GeV}/c$, and it overestimates the yield at high transverse momenta.

The per-trigger yield in the combinatorial background of the two-particle correlation shown in Figure 8.18 is described well at low transverse momenta by all MC generators but Pho-

jet. Here, Phojet underestimates the yield by up to 50 %. At high transverse momenta, however, Phojet approaches again the ALICE results. For transverse momenta above $p_{\mathrm{T}} > 1.0\,\mathrm{GeV}/c$, also Pythia6 Perugia-0 and Pythia8 4C start to deviate slightly from the ALICE results at high charged particle multiplicities (up to 30 %). Pythia6 Perugia-2011 has the best agreement with the ALICE results for all presented transverse momentum thresholds. Only at low N_{ch}, and here especially at high p_{T}, the yield is underestimated by Pythia6 Perugia-2011.

Figure 8.19 shows the average number of trigger particles as a function of the charged particle multiplicity for different lower limits in $p_{\mathrm{T,\,trig}}$. As the transverse momentum spectrum is steeply falling, the number of trigger particles decreases strongly with $p_{\mathrm{T,\,trig}}$. At low charged particle multiplicities, the Monte Carlo generators reproduce the ALICE results well (agreement within 10 %). Only Phojet underestimates the number of trigger particles. At high transverse momenta, Pythia6 Perugia-0 and Pythia8 4C do not reproduce the dependence on the charged particle multiplicity. The average number of trigger particle is overestimated by up to 50 % by Pythia6 Perugia-0 and Pythia8 4C. Pythia6 Perugia-2011 shows the best description of the average number of trigger particles for all transverse momentum thresholds and all charged particle multiplicities.

The number of uncorrelated seeds is presented in Figure 8.20. At low transverse momenta, all Pythia tunes give a good description of the ALICE results within the systematic uncertainties. Phojet underestimates the number of uncorrelated seeds by up to 40 %. With increasing transverse momenta, Pythia6 Perugia-0 and Pythia8 4C show increasing deviations from the ALICE results (up to 30 %). Pythia6 Perugia-2011 gives the best description of the average number of trigger particles for all transverse momentum cuts and all multiplicities.

In summary, none of the MC generators achieves an accurate description of all observables at all charged particle multiplicities and all transverse momentum thresholds. Pythia6 Perugia-2011 gives the best overall agreement with the ALICE results. This can be explained by the fact that this tune was optimized for minimum bias data measured at $\sqrt{s} = 7.0\,\mathrm{TeV}$. Still, there are large deviations between result this tune and the ALICE results. For example, the per-trigger near side yield is overestimated by up to 50 % over a large range of transverse momentum thresholds. The per-trigger away side yield is underestimated for low transverse momentum thresholds and overestimated for large transverse momentum thresholds by up to 50 %.

This disagreement between the ALIE results and model predictions can be used as a feedback to the theoretical models allowing to adjust the parametrization of the jet fragmentation down to low energies into the non-perturbative regime.

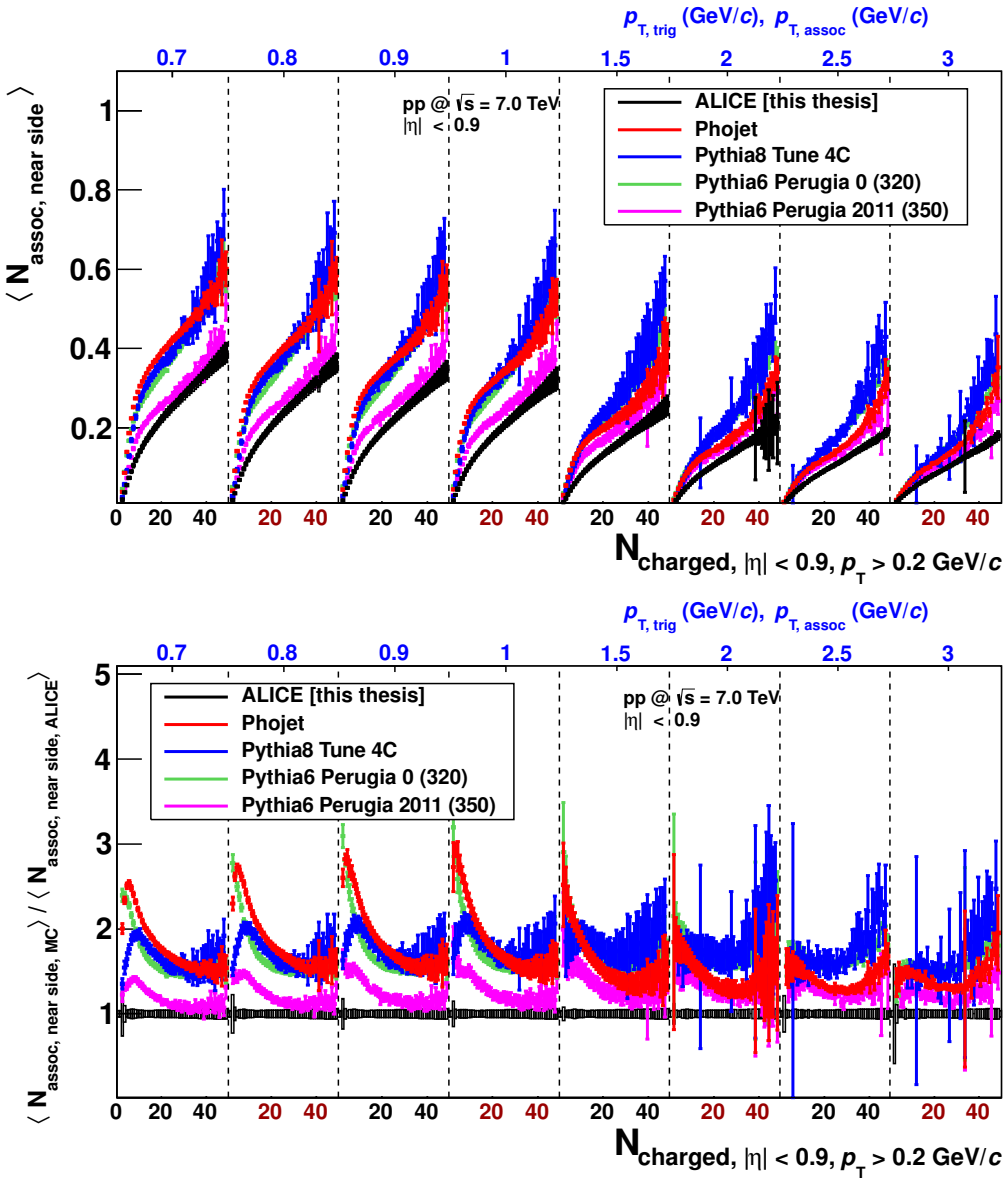

Figure 8.16.: Average per-trigger near side yield measured at $\sqrt{s} = 7.0\,\text{TeV}$ as a function of charged particle multiplicity N_{ch} and as a function of the transverse momentum threshold.

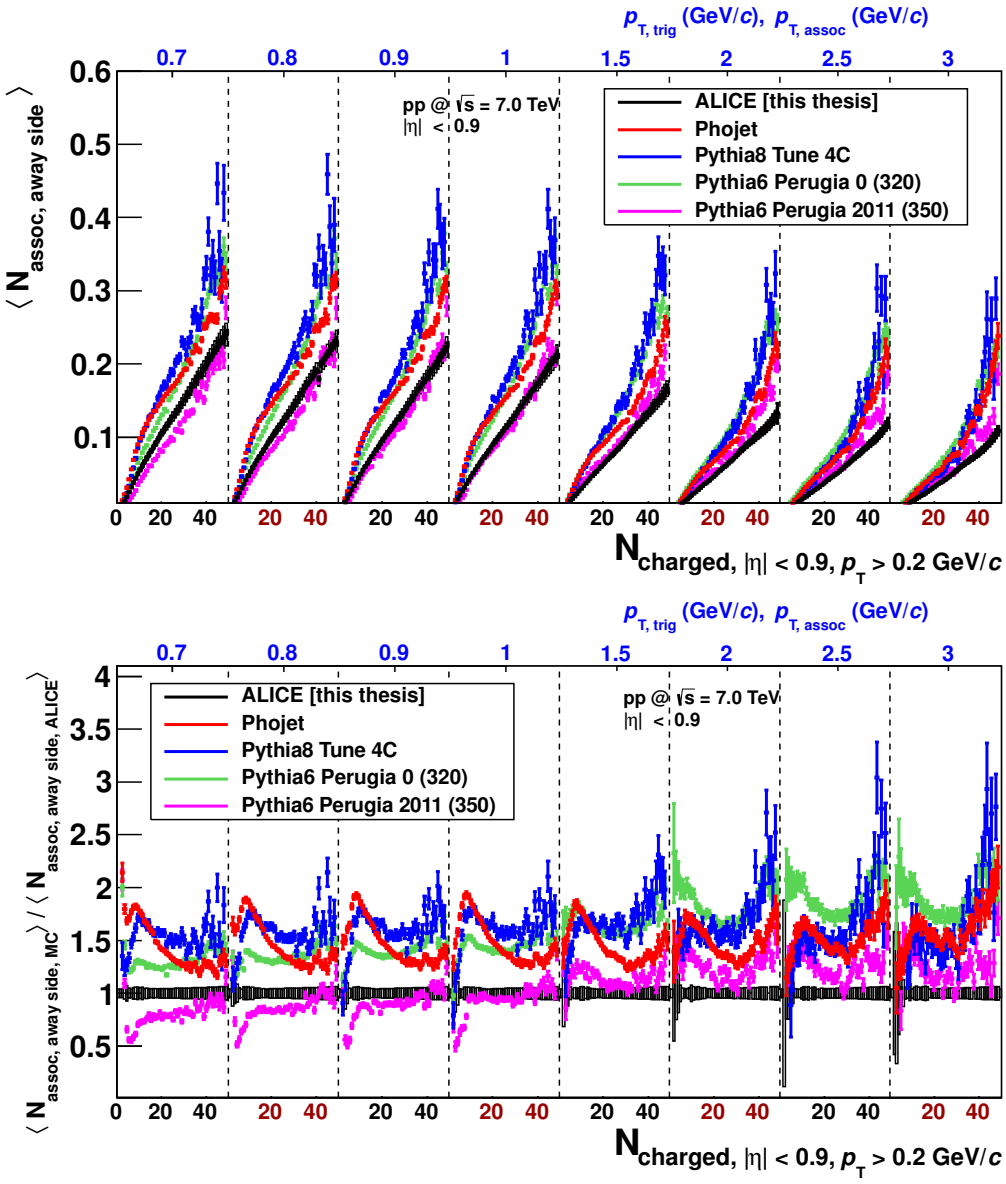

Figure 8.17.: Average per-trigger away side yield measured at $\sqrt{s} = 7.0\,\text{TeV}$ as a function of charged particle multiplicity N_{ch} and as a function of the transverse momentum threshold.

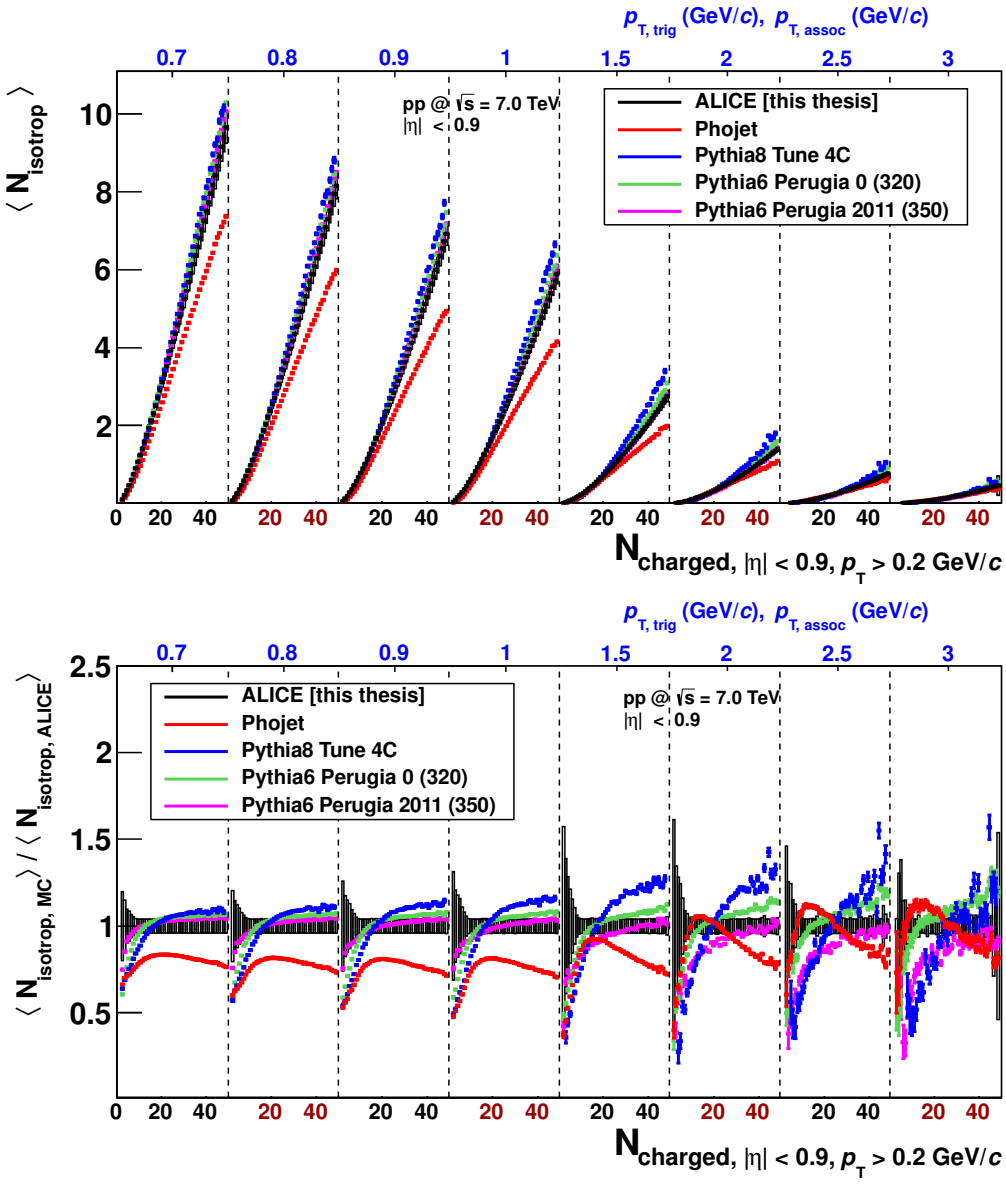

Figure 8.18.: Average per-trigger yield in isotropically distributed background measured at $\sqrt{s} = 7.0\,\text{TeV}$ as a function of charged particle multiplicity N_{ch} and as a function of the transverse momentum threshold.

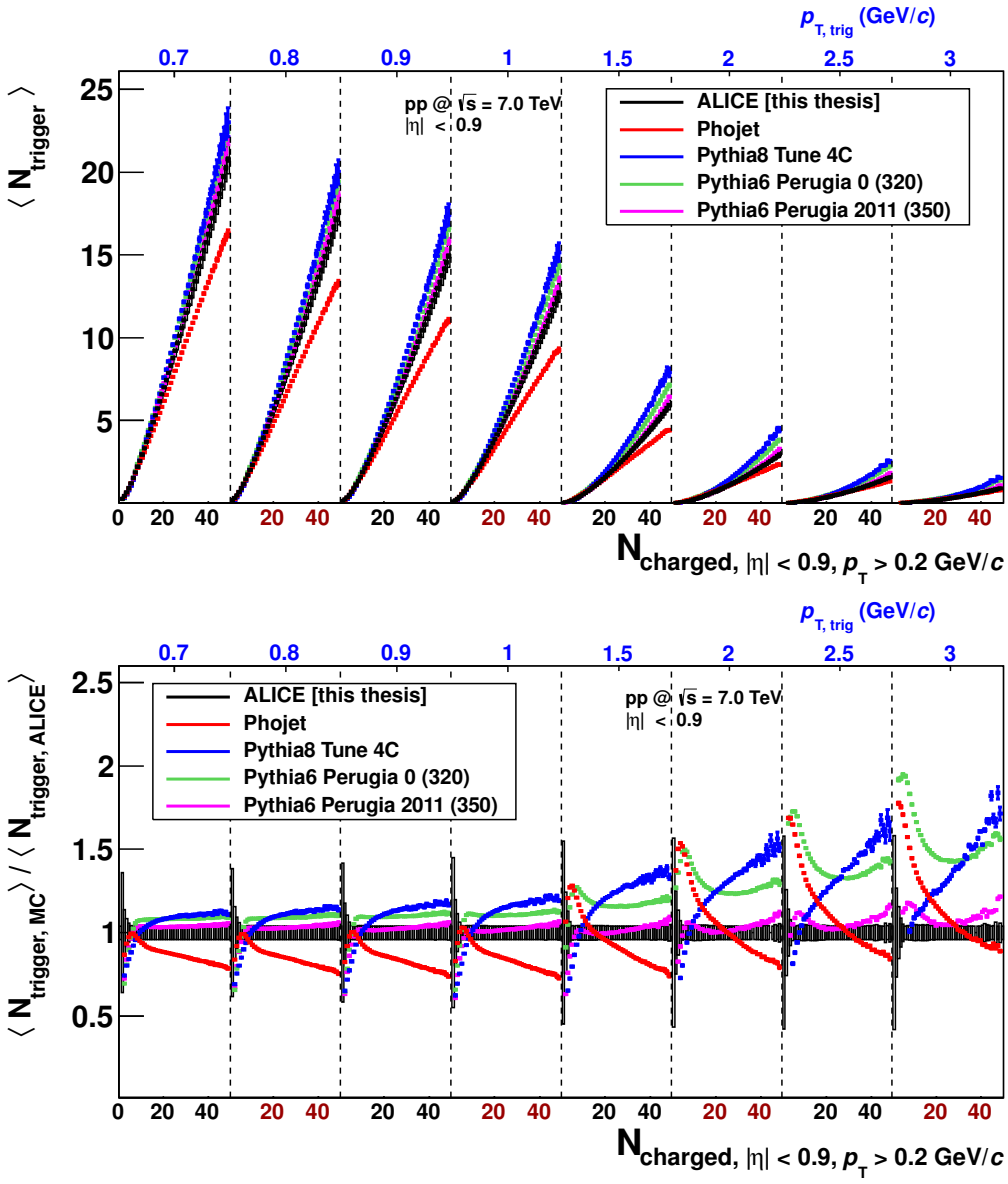

Figure 8.19.: Average number of trigger particles per event measured at $\sqrt{s} = 7.0\,\text{TeV}$ as a function of charged particle multiplicity N_{ch} and as a function of the transverse momentum threshold.

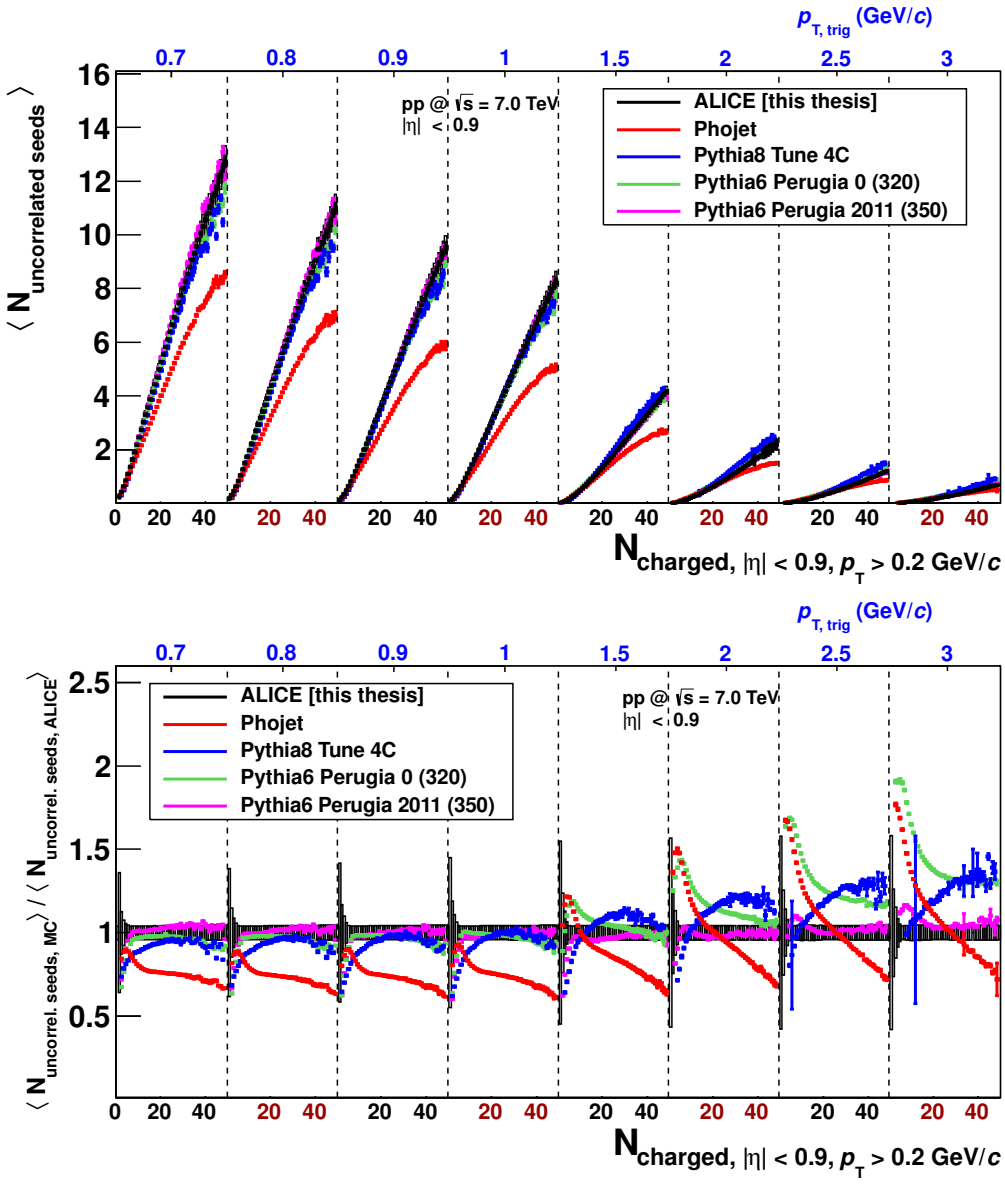

Figure 8.20.: Average number of uncorrelated seeds per event measured at $\sqrt{s} = 7.0\,\mathrm{TeV}$ as a function of charged particle multiplicity N_{ch} and as a function of the transverse momentum threshold.

8.3. Insight into Multiple Parton Interactions

Pythia simulations incorporate multiple parton interactions (MPI). In Pythia6, the number of MPI is defined as the number of (semi)-hard scatterings that have occurred in the current event [SMS06]. Here, hard interactions include transverse momentum transfers above a fixed transverse momentum threshold of a few GeV. Semi-hard interactions include transverse momentum transfers just below this hard threshold.
In Section 5.5, it has been demonstrated that for Pythia6 Perugia-2011 simulations the average number of uncorrelated seeds, $\langle N_{\text{uncorrelated seeds}} \rangle$, measured in the ALICE acceptance of $|\eta| < 0.9$ and using $p_{\text{T, trig}} > 0.7\,\text{GeV}/c$ scales almost linearly with the number of multiple parton interactions. This suggests that the $\langle N_{\text{uncorrelated seeds}} \rangle$ measured by ALICE in proton-proton collisions is proportional to the number of multiple parton interactions in proton-proton collision data. In the following, the dependence of the number of uncorrelated seeds on the charged particle multiplicity is investigated. Also, a simulation based test is performed in order to probe the feasibility of measuring a limit in the number of multiple parton interactions.

8.3.1. Limit in the Number of Multiple Parton Interactions

In the following, it is tested whether a sharp limit in the number of multiple parton interactions can be detected with the two-particle correlation data analysis. For this purpose, two Pythia6 Perugia-2011 simulations are performed. In the first simulation, the default Perugia-2011 settings are used. In the second simulation, an artificial limit in the number of multiple parton interactions of $N_{\text{MPI}} \leq 10$ is added to the default simulation setup.
The top left panel of Figure 8.21 shows the probability distribution $P(N_{\text{MPI}})$ for both simulations. As the $P(N_{\text{MPI}})$ is falling steeply and the number of events with $N_{\text{MPI}} > 10$ is very low, the data points of both simulations for $N_{\text{MPI}} \leq 10$ are almost on top of each other. While the simulation including the artificial N_{MPI}-limit covers the range $0 \leq N_{\text{MPI}} \leq 10$, the default simulation comprises N_{MPI} up to 25.
The top right panel of Figure 8.21 shows the charged particle multiplicity distribution $P(N_{\text{ch}})$ for charged particles in the ALICE acceptance for both simulations. The simulations reach charged particle multiplicities of up to $N_{\text{ch}} \approx 75$ and $N_{\text{ch}} \approx 90$, respectively.
The bottom left panel of Figure 8.21 shows the correlation between the number of multiple parton interactions and the charged particle multiplicities. Each bin in N_{MPI} corresponds to a large range of possible charged particle multiplicities. Vice versa, for each bin in charged particle multiplicity, a large range in N_{MPI} is covered.
The bottom right panel of Figure 8.21 shows the average number of multiple parton interactions as a function of the charged particle multiplicity for the two Pythia6 Perugia-2011 simulations. The average number of N_{MPI} in a fixed multiplicity bin is always dominated

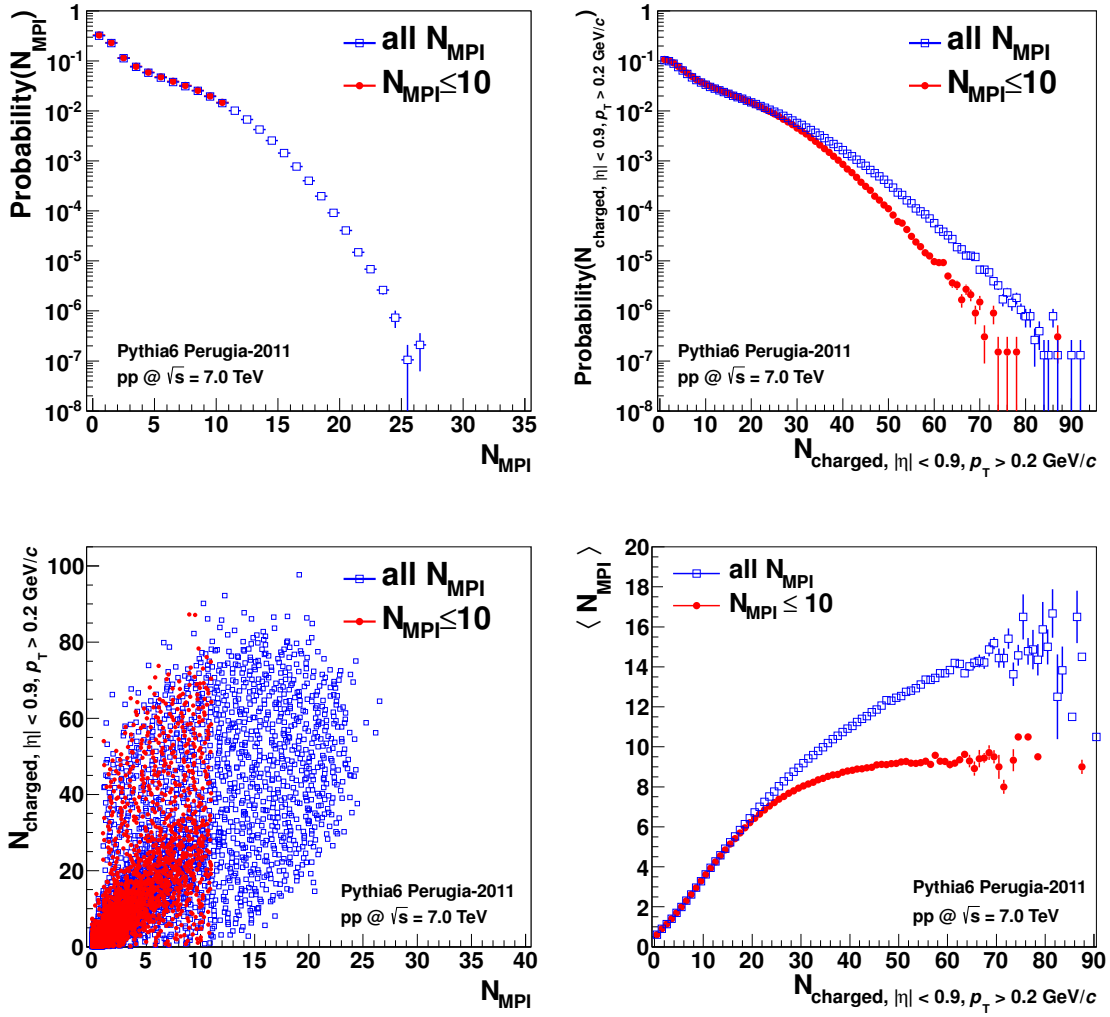

Figure 8.21.: Distributions measured in two Pythia6 Perugia-2011 simulations at $\sqrt{s} = 7.0\,\mathrm{GeV}/c$. The first production (blue) is performed with default Perugia-2011 parameters. The second production (red) is performed with an additional artificial limit in the number of multiple parton interactions ($N_{\mathrm{MPI}} \leq 10$). Top left panel: Distribution of the number of multiple parton interactions $P(N_{\mathrm{MPI}})$. Top right panel: Distribution of the charged particle multiplicity $P(N_{\mathrm{ch}})$. Bottom left panel: Correlation between the number of multiple parton interactions and the charged particle multiplicity. Bottom right panel: Average number of multiple parton interactions $\langle N_{\mathrm{MPI}} \rangle$ as a function of the charged particle multiplicity.

by low N_{MPI}, as the $P(N_{\mathrm{MPI}})$ distribution is falling steeply towards high N_{MPI}. The mean number of multiple parton interactions increases for both data sets as a function of the charged particle multiplicity. For the simulation using the artificially limited N_{MPI}, the average number of multiple parton interactions saturates above $N_{\mathrm{ch}} = 30$ where it approaches the limit of $N_{\mathrm{MPI}} = 10$. Also, in the simulation without the artificial limit in N_{MPI}, the $\langle N_{\mathrm{MPI}} \rangle$ slightly saturates at very high charged particle multiplicities.

In Section 5.5, the linear dependence between the number of multiple parton interactions and the number of uncorrelated seeds has has been demonstrated. Now, it is tested whether the limit in the number of multiple parton interactions can also be accessed by measuring the number of uncorrelated seeds.

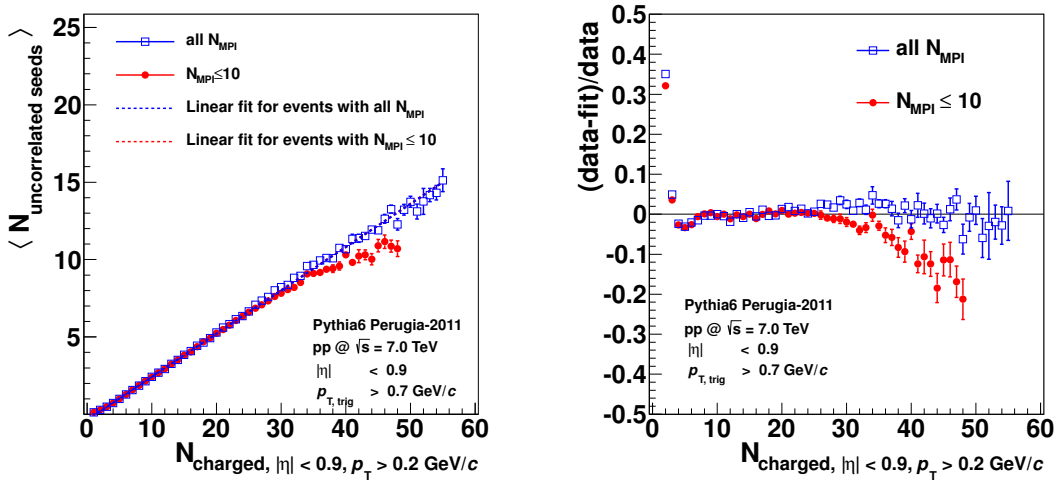

Figure 8.22.: Left panel: Number of uncorrelated seeds as a function of the charged particle multiplicity for Pythia6 Perugia-2011 simulations together with linear fit functions. One of the simulations is performed with a limit in the number of multiple parton interactions of $N_{\mathrm{MPI}} \leq 10$. Right panel: Residuals between the number of uncorrelated seeds as a function of the charged particle multiplicity and linear fit functions for the same Pythia6 Perugia-2011 simulations.

The left panel of Figure 8.22 shows the number of uncorrelated seeds as a function of the charged particle multiplicity in combination with linear fit functions for both Pythia6 Perugia-2011 simulations. The right panel of Figure 8.22 shows the residuals between the data distribution and the corresponding linear fit functions.
In both panels of Figure 8.22, it can be seen that the number of uncorrelated seeds estimated for the data set with the limited number of multiple parton interactions falls below the linear dependence at high charged particle multiplicities. This is in agreement with the limit in $\langle N_{\mathrm{MPI}} \rangle$ displayed in the bottom right panel of Figure 8.21.
For the simulation without the artificial limit in N_{MPI}, the bottom right panel of Figure 8.21 also shows a slight stagnation in the average number of multiple parton interactions, $\langle N_{\mathrm{MPI}} \rangle$, for high multiplicities. However, for this simulation almost all measured $\langle N_{\mathrm{uncorrelated\,seeds}} \rangle$ at high multiplicities agree within the uncertainties with a linear dependence. This can be explained in the following way. The estimation of the number of uncorrelated seeds is based on a fitting procedure (cf. Section 5.3.1). Hence, a minimum number of events is needed to obtain stable fit results (cf. Section 7.1.2). The charged particle multiplicity distribution is steeply falling. Given the number of events of the two

Pythia6 productions used for this simulation study, the multiplicity bins above $N_{\mathrm{ch}} = 50$ are populated with too few events to be accepted in the two-particle correlation analysis algorithm. Hence, the number of uncorrelated seeds can only be estimated up to $N_{\mathrm{ch}} \approx 50$. In the multiplicity range $0 < N_{\mathrm{ch}} < 50$, the distribution of $\langle N_{\mathrm{MPI}} \rangle$ only deviates slightly from a linear dependence on the charged particle multiplicity. In order to measure this soft limit in N_{MPI} for events with $N_{\mathrm{ch}} > 50$, more events at high multiplicities would be needed.
In summary, it is possible to detect sharp limits in the number of multiple parton interactions using the presented data analysis method. The more events are detected at high charged particle multiplicities, the more sensitive the analysis will be to softer limits in MPI.

8.3.2. Number of Multiple Parton Interactions in ALICE Data

In this section, the dependence between the number of multiple parton interactions and the charged particle multiplicity measured in pp collisions in ALICE is discussed.

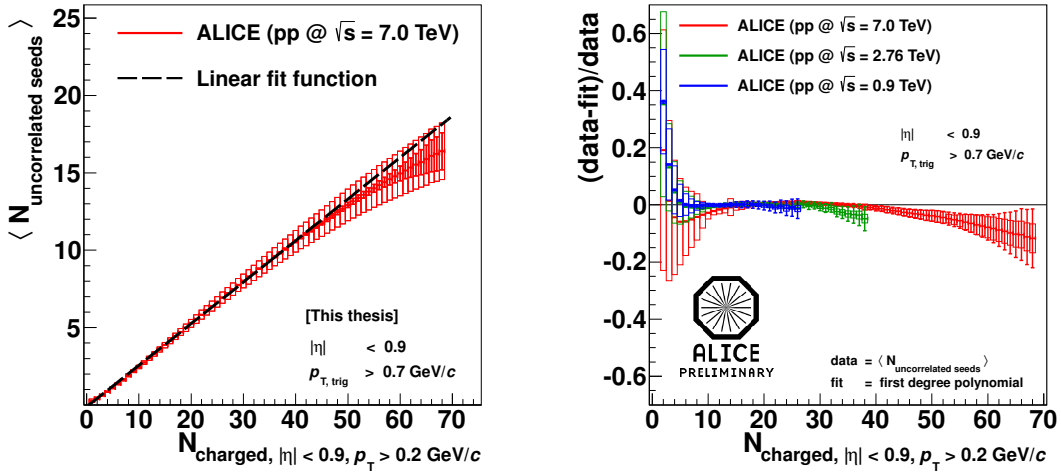

Figure 8.23.: Left panel: Number of uncorrelated seeds as a function of the charged particle multiplicity described with a linear fit function for $\sqrt{s} = 7.0\,\mathrm{TeV}$. Right panel: Residuals between the number of uncorrelated seeds as a function of the charged particle multiplicity and a linear fit function for ALICE data measured at $\sqrt{s} = 0.9$, 2.76, and 7.0 TeV [Sic12]. The same trend is visible in the $\langle N_{\mathrm{uncorrelated\ seeds}} \rangle$-distribution as a function of the reconstructed multiplicity (plot not shown).

The left panel of Figure 8.23 shows the number of uncorrelated seeds as a function of the charged particle multiplicity measured by ALICE in pp collision data at $\sqrt{s} = 7.0\,\mathrm{TeV}$. For low and intermediate charged particle multiplicities, the number of uncorrelated seeds rises almost linearly with the charged particle multiplicity. The linear rise of $\langle N_{\mathrm{uncorrelated\ seeds}} \rangle$ is demonstrated using a linear fit function. While at low and intermediate charged particle

multiplicities all data points agree with the linear fit function, at high charged particle multiplicities, the data points are systematically found below the linear fit function.
The right panel of Figure 8.23 shows the residuals ((data − fit)/data) between the data points and the linear fit functions for ALICE pp collision data measured at $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV.
The systematic uncertainty of the residuum in Figure 8.23 is much smaller than the systematic uncertainty of the $\langle N_{\text{uncorrelated seeds}} \rangle$ data points. An explanation for this is that correlated systematic uncertainties of all data point also cause a modification of the linear fit function. As many sources of the systematic uncertainties give correlated shifts of the data points in the same direction, the remaining systematic uncertainty on the residuum is smaller than for the $\langle N_{\text{uncorrelated seeds}} \rangle$ itself.
At all three collision energies and at correspondingly high charged particle multiplicities, there is a hint of a deviation from the linear dependence on the charged particle multiplicity. While each data point almost agrees within the combined error bars with the linear fit functions, many data points at high multiplicities are systematically below the fit function. Further analysis on high multiplicity data would be needed in order to validate the assumption.
A deviation from the linear dependence can be interpreted as an indication of a limit in the number of multiple parton interactions as discussed in Section 8.3.1.
For each center-of-mass energy, the hint of the deviation from the linear dependence is observed at a different value of the charged particle multiplicity. The higher the center-of-mass energy is, the higher is the value of the charged particle multiplicity at which the number of uncorrelated seeds deviate from a linear dependence.

8.4. Summary

In this chapter, the results of the two-particle angular correlation analysis of proton-proton collision data have been presented. In Section 8.1.1 and 8.1.2, the per-trigger yield as a function of $\Delta\varphi$ has been presented for ALICE data at $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV in comparison to Monte Carlo predictions as well as in a direct comparison of ALICE data for the three center-of-mass energies. The per-trigger yield as a function of $\Delta\varphi$ has been presented exemplarily for a few selected bins in charged particle multiplicity only.
In a more compact visualization shown in Section 8.2.1 and 8.2.2, the decomposed azimuthal correlation properties $\langle N_{\text{assoc, near side}} \rangle$, $\langle N_{\text{assoc, away side}} \rangle$, $\langle N_{\text{isotrop}} \rangle$, $\langle N_{\text{trigger}} \rangle$, and $\langle N_{\text{uncorrelated seeds}} \rangle$ have been presented for ALICE data and Monte Carlo generator predictions. The correlation properties have been studied for their dependence on the charged particle multiplicity as well as for the dependence on the center-of-mass energies. In Section 8.2.3, the transverse momentum dependence of the correlation results has been studied.

While Pythia6 Perugia-2011 gives the best overall agreement with the ALICE results, none of the presented MC generators achieves an accurate description of all observables at all charged particle multiplicities and all transverse momentum thresholds. These disagreements can be used to adjust the parametrization of the jet fragmentation in theoretical models at these low transverse momenta.
An analysis of the number of uncorrelated seeds $\langle N_{\text{uncorrelated seeds}} \rangle$ and its dependence on the charged particle multiplicity has been presented in Section 8.3. The analysis hints to a deviation from the linear dependence of $\langle N_{\text{uncorrelated seeds}} \rangle$ on the charged particle multiplicity at high charged particle multiplicities. A deviation from the linear dependence is consistent with a limit in the number of multiple parton interactions. This is suggested by simulation studies of Pythia6 Perugia-2011 events. With increasing center-of-mass energy, the limit in MPI seems to be shifted to larger charged particle multiplicities corresponding to larger number of multiple parton interactions.

Summary

With the start-up of the Large Hadron Collider (LHC), a new era of collider physics has begun. Since 2009, the LHC has continuously performed proton-proton (pp) collisions and collisions of heavy-ions (Pb–Pb). In comparison to previous collider facilities, the collision energies has been increased by a factor of 4 for the pp collisions and by a factor 14 for the heavy-ion collisions. A further doubling of the collision energy is foreseen within the next few years of operation.

In this thesis, characteristics of proton-proton collisions are studied that arise from the composite structure of the proton. Protons comprise multiple quarks and gluons, also denoted as partons. When protons are accelerated to velocities close to the velocity of light, their de Broglie wavelength becomes much smaller than the size of the proton and their inner structure is resolved. A natural consequence of the proton's substructure is that in high-energy proton-proton collisions, several pairs of partons can undergo scatterings (multiple parton scattering). Each parton-parton interaction can give rise to the production of final state particles. Hence, multiple parton interactions influence many physics observables, for instance the charged particle multiplicity. Due to the increase in the parton cross section, the effect becomes more important at higher collision energies. At LHC energies, this effect becomes significant.

In the past, the physics of multiple parton interactions has mainly been studied by exploring charged particle multiplicity distributions and multi-jet events. In this thesis, a new analysis technique has been presented that combines a charged particle multiplicity analysis with a jet analysis. The jet analysis is based on a two-particle angular correlation analysis.
The result of a two-particle angular correlation is a pair distribution of the difference in the azimuthal angle $\Delta\varphi$ and in the pseudorapidity $\Delta\eta$ for particle pairs normalized to the number of trigger particles. Two-particle angular correlations allow to study on an average basis whether particles generated in the same collision event are correlated to each other or not. For example, particles generated in the same jet are close to each other in φ and η, while particles coming from uncorrelated sources are independent in φ and η. Such angular correlation analyses with low transverse momentum thresholds allow to access jet properties on a statistical basis down to the lowest possible jet energies. In

contrast, jet reconstruction on an event-by-event basis would only identify those jets which have energies significantly larger than the underlying event. Furthermore, two-particle angular correlation analyses are insensitive to overlapping jets.

The main observable of the presented two-particle angular correlation analysis is the azimuthal correlation. This is the pair distribution as a function of the difference in the azimuthal angle $\Delta\varphi$ of the particle pairs normalized to the number of trigger particles, also referred to as per-trigger yield as function of $\Delta\varphi$. The azimuthal correlation reveals a near side peak at $\Delta\varphi = 0$ corresponding mainly to particles from the fragmentation of jets, and an away side yield at $\Delta\varphi = \pi$ corresponding to particles from the fragmentation of recoiling jets. The near side and the away side peak lie on top of a constant background representing particles generated in uncorrelated particle production processes.
For the decomposition of the azimuthal correlation, a fit function has been introduced. The fit function allows a simplified comparison of properties of the azimuthal correlation between various event classes without the need to compare the complete, complex azimuthal distributions directly to each other.
Five observables have been derived from the azimuthal correlation. They have been analyzed as a function of the charged particle multiplicity and separated for various ranges in transverse momentum. Three observables represent the decomposed per-trigger yield, the yield in the combinatorial background, the near side yield, and the away side yield. In addition, the average number of trigger particles is measured. Finally, the combination of these observables results in the number of uncorrelated seeds, a measure of the uncorrelated sources of particle production. It has been demonstrated that in Pythia the number of uncorrelated seeds is related to the number of multiple parton interactions.

The two-particle angular correlation analysis has been used to analyze LHC proton-proton collision data measured with the ALICE detector. The high sensitivity to low-momentum particles of ALICE due to its low material budget and low magnetic field allows ALICE to make a unique contribution to studies of the (soft) particle production mechanisms in proton-proton collision at the LHC. The analysis procedure uses tracks of charged particles measured with the ALICE central tracking detectors, the Inner Tracking System (ITS) and the Time Projection Chamber (TPC).

In the course of the data analysis, a quality assurance (QA) analysis of collision data with a focus on the two inner tracking detectors ITS and TPC has been developed. In addition to its application in the data selection of the presented data analysis, the QA analysis is used as part of the ALICE wide quality assurance framework for the classification of all recorded collision data sets.

For the comparison of simulated data an in-depth comparison of different transport Monte Carlos has been done. In order to use the transport Monte Carlo Geant4 in addition to ALICE's default transport Monte Carlo Geant3, the implementation of Geant4 in the ALICE simulation framework has been validated and optimized. First large scale ALICE simulations using Geant4 on the ALICE computing grid have been performed. By now, Geant4 has been established as the second official transport Monte Carlo used by the ALICE collaboration.

The analysis results of the collision data at three different center-of-mass energies, $\sqrt{s}$ = 0.9, 2.76, and 7.0 TeV, have been compared to theoretical model predictions. As theoretical models, Pythia6, Pythia8, and Phojet1.12 are used, all of which incorporate the process of multiple parton interactions.
While the dominant contribution of the per-trigger yield, the yield in the combinatorial background, is reproduced well by all event generators, the yields in the near side peak and the away side peak representing the jet fragmentation show significant deviations. For example, at $\sqrt{s}$ = 7.0 TeV, $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$, $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$, and low multiplicities, the per-trigger near side yield is overestimated by all models by between 40 % and 120 %.
The model predictions of the near and away side yield disagree with ALICE results not only at low transverse momenta, but over a large range of transverse momentum thresholds ($p_{\mathrm{T}} > 0.7\,\mathrm{GeV}/c$ to $p_{\mathrm{T}} > 3.0\,\mathrm{GeV}/c$). For example, the per-trigger near side yield at $\sqrt{s}$ = 7.0 TeV and at low multiplicities is overestimated by 40 % to 200 % in this transverse momentum range. The agreement increases for large charge particle multiplicities. Here, the disagreement is limited to at most 100 %, some models agree with the data.
The average number of trigger particles and the number of uncorrelated seeds are well reproduced by all model predictions for the low transverse momentum region. At large transverse momentum values, discrepancies of up to 50 % are visible.
The best description of all features of the two-particle angular correlation is given by Pythia6 Perugia-2011 which has been optimized to other underlying event observables at $\sqrt{s}$ = 7.0 TeV. In comparison to the other Pythia tunes, the Perugia-2011 tune incorporates an increased strength of the underlying event and a softer jet fragmentation function.
The disagreement between the model predictions and the real collision data is most pronounced for the highest collision energy, $\sqrt{s}$ = 7.0 TeV, and decreases with decreasing collision energy ($\sqrt{s}$ = 2.76 and 0.9 TeV). This is expected since the theoretical models have been optimized to measurements at collision energies close to these energies.
The disagreement between real data and model predictions can be used as a feedback

to the developers of theoretical models allowing to adjust the parametrization of the jet fragmentation down to low energies into the non-perturbative regime.

The number of multiple parton interactions has been studied using the number of uncorrelated seeds. It has been observed that the charged particle multiplicity rises linearly with the number of uncorrelated seed. This is expected as each parton-parton interaction can give rise to particle production. However, the linear dependence is violated at large charged particle multiplicities. Here, the rise of the number of uncorrelated seeds levels off. This would be consistent with the assumption that there is a limit in the number of multiple parton interactions in proton-proton collisions.

In summary, a new analysis method for the study of jet fragmentation and the number multiple parton interactions has been established. This method has been successfully used to gain a better understanding of the particle production in proton-proton collisions connected to jet fragmentation and multiple parton interactions. Based on these results, theoretical models can be optimized for LHC energies.

Zusammenfassung

Mit der Inbetriebnahme des Large Hadron Collider (LHC, großer Hadronenbeschleuniger) hat eine neue Ära der Hochenergiephysik begonnen. Seit 2009 werden in ihm Protonen und Bleikerne zur Kollision gebracht. Im Vergleich zu früheren Protonenbeschleunigern wurde am LHC die Kollisionsenergie um den Faktor vier und die Kollisionsenergie für Schwerionenkollisionen um den Faktor vierzehn erhöht. Für die kommenden Jahre des LHC-Betriebes ist eine weitere Verdopplung der Kollisionsenergie geplant.

In der vorliegenden Arbeit wurden Effekte in Proton-Proton-Kollisionen untersucht, die in der Substruktur des Protons begründet liegen. Protonen beinhalten mehrere Quarks und Gluonen, die gemeinsam auch Partonen genannt werden. Wenn Protonen zu Geschwindigkeiten nahe der Lichtgeschwindigkeit beschleunigt werden, wird die De-Broglie-Wellenlänge des Protons viel kleiner als das Proton selbst, so dass seine innere Struktur aufgelöst werden kann. Als natürliche Folge der Substruktur des Protons können in hochenergetischen Proton-Proton-Kollisionenen mehrere Paare von Partonen miteinander kollidieren. Dieser Effekt wird Multi-Parton-Interaktion genannt. Jede Parton-Parton-Interaktion kann zur Erzeugung von Endzustandsteilchen beitragen. Multi-Parton-Interaktionen können viele Messgrößen in Proton-Proton-Kollisionsexperimenten beeinflussen, etwa die Multiplizität geladener Teilchen. Aufgrund des Anstiegs des Parton-Parton Wirkungsquerschnittes gewinnt dieser Effekt bei ansteigenden Kollisionsenergien an Bedeutung. Bei LHC-Energien ist der Anteil aller Proton-Proton-Kollisionen, der Multi-Parton-Interaktionen vorweist, signifikant.

In der Vergangenheit wurden Multi-Parton-Interaktionen vornehmlich durch die Analyse von Multiplizitätsverteilung geladener Teilchen und durch Multi-Jet-Analysen untersucht. In der vorliegenden Arbeit wurde eine neue, kombinierte Methode eingeführt, mit der Multi-Parton-Interaktionen untersucht werden können. Diese Methode vereinigt eine Multiplizitätsanalyse mit einer Zwei-Teilchen-Winkelkorrelationsanalyse.
In Zwei-Teilchen-Winkelkorrelationen wird der Abstand zweier Teilchen in Bezug auf ihren azimuthalen Winkel $\Delta\varphi$ und ihrer Pseudorapidität $\Delta\eta$ vermessen. Zwei-Teilchenkorrelationen ermöglichen es, statistisch darüber Aufschluss zu geben, ob Teilchen, die in derselben Proton-Proton-Kollision erzeugt wurden, zueinander korreliert sind oder nicht. Teilchen, die im selben Jet erzeugt worden sind, haben im Mittel

kleine Abstände in Bezug auf φ und η, wohingegen Teilchen, die aus unkorrelierten Teilchenproduktionsmechanismen stammen, weitestgehend unabhängig in φ und η sind. Winkelkorrelationsanalysen, die Teilchen mit sehr niedrigen Impulsen einschließen, ermöglichen es, Jets auf einer statistischen Basis bis hin zu sehr geringen Jet-Energien zu analysieren. Demgegenüber erlauben Analysen, die auf Jet-Rekonstruktion basieren, nur die Analyse von Jets oberhalb einer weit höher gelegenen Energieschwelle, der Energie des sogenannten Underlying Events. Zwei-Teilchenkorrelationen sind außerdem im Vergleich zur Jetrekonstruktion unempfindlich gegenüber sich überlappenden Jets.

Die Hauptmessgröße der hier vorgestellten Zwei-Teilchen-Winkelkorrelation ist die Verteilung des Abstandes von zwei Teilchen in Bezug auf ihren azimuthalen Winkel $\Delta\varphi$ normiert auf die Anzahl der Triggerteilchen. Die azimuthale Korrelation beinhaltet drei Komponenten. Einen Peak in Richtung der Triggerteilchenrichtung (um $\Delta\varphi = 0$), einen Peak in die Gegenrichtung der Triggerteilchen (um $\Delta\varphi = \pi$) und einen konstanten Untergrund. Der Peak auf der Seite der Triggerteilchen beinhaltet hauptsächlich Teilchenpaare aus der Fragmentation von Jets. Der Peak in die Gegenrichtung repräsentiert Teilchenpaare, die aus der Fragmentation von Jets stammen, die in die Gegenrichtung emittiert wurden. Die Teilchenpaare, die den konstanten Untergrund bilden, kommen von Teilchenproduktionsmechanismen, die zu den Triggerteilchen unkorreliert sind.
Um die azimuthale Korrelationsverteilung auf ihre Anteile und Eigenschaften hin analysieren zu können, wurde in dieser Arbeit eine Fit-Funktion entwickelt. Die Fit-Funktion erlaubt es einzelne Anteile der Korrelationsverteilung zu extrahieren. So kann ein direkter Vergleich zwischen verschiedenen Proton-Proton-Kollisionsklassen durchgeführt werden, ohne die vollständigen, komplexen Korrelationsverteilungen miteinander vergleichen zu müssen.
Von der azimuthalen Korrelation wurden fünf Messgrößen abgeleitet. Diese Messgrößen wurden als Funktion der Multiplizität geladener Teilchen und für verschiedene Transversalimpulse gemessen. Drei dieser Messgrößen stellen das Integral der azimuthalen Korrelation dar, also das Integral des konstanten Untergrundes, das Integral des Peaks in Richtung der Triggerteilchen und das Integral des Peaks in Gegenrichtung der Triggerteilchen. Weiterhin wurde die mittlere Anzahl der Triggerteilchen pro Kollision gemessen. Von diesen Observablen kann dann die Anzahl der unkorrelierten Triggerteilchen abgeleitet werden. Diese Messgröße entspricht den unkorrelierten Quellen der Teilchenproduktion. Es wurde gezeigt, dass in Pythia-Simulationen die Messgröße der Anzahl der unkorrelierten Triggerteilchen mit der Anzahl der Multi-Partonen-Interaktionen in einer linearen Beziehung steht.

Die Zwei-Teilchen-Winkelkorrelationsanalyse wurde verwendet, um Proton-Proton-Kollisionsdaten zu analysieren, die mit dem LHC-Experiment ALICE aufgenommen

worden sind. Aufgrund seines vergleichbar geringen Magnetfeldes und seines geringen Material-Budgets kann das ALICE-Experiment auch Teilchen mit sehr geringen Transversalimpulsen vermessen. So können ALICE-Daten einen einzigartigen Beitrag zum Studium der (weichen) Teilchenproduktion in Proton-Proton-Kollisionen am LHC leisten. Die in der vorgelegten Arbeit präsentierte Datenanalyse verwendet Spuren geladener Teilchen, die mit den zwei innersten Spurendetektoren von ALICE in Proton-Proton-Kollisionen gemessen worden sind. Die Detektoren sind das Inner Tracking System (ITS, Inneres Spurenmessgerät) und die Time Projection Chamber (TPC, Spurendriftkammer).

Zusammen mit der Datenanalyse wurde eine Qualitätsanalyse von Kollisionsdaten entwickelt, die die Qualität von Kollisionsdaten in Bezug auf Detektorzustand, Kalibrations- und Rekonstruktionsqualität analysiert und klassifiziert. Diese Analyse wurde speziell für die Detektoren ITS und TPC konstruiert. Neben der Anwendung dieser Analyse zur Selektion der Datensätze für die vorgestellte Korrelationsanalyse wurde diese Qualitätsanalyse auch in die globale Qualitätsicherungsanalyse von Kollisionsdaten in ALICE aufgenommen.

In dieser Arbeit wurde zudem ein Beitrag zu Studien über die systematische Unsicherheit bezüglich der Auswahl des Transport-Monte-Carlo-Programms in ALICE-Simulationen geleistet. Um neben dem nominellen Transport-Monte-Carlo-Programm Geant3 auch den Nachfolger Geant4 verwenden zu können, wurde die Integration von Geant4 in das ALICE-Simulationssystem validiert und optimiert. Im Zuge der Validierung wurden erste Simulationen auf dem ALICE Computing Grid durchgeführt. Geant4 wurde damit als das zweite offizielle Transport-Monte-Carlo-Programm in ALICE etabliert.

Die Analyseergebnisse der Proton-Proton-Kollisionsdaten, die bei den drei verschiedenen Kollisionsenergien $\sqrt{s}$ = 0.9, 2.76 und 7.0 TeV gemessen worden sind, wurden mit Vorhersagen von theoretischen Modellen verglichen. Als theoretische Modelle wurden Pythia6, Pythia8 und Phojet1.12 verwendet. All diese Modelle beinhalten das Konzept der Multi-Parton-Interaktionen.
Während der konstante Untergrund der azimuthalen Korrelation von allen Monte-Carlo-Generatoren für alle Multipliziätsklassen gut reproduziert wird, weicht die Vorhersage der Verteilung in den Peaks in Triggerrichtung und in Gegenrichtung, welche die Jetfragmentation repräsentieren, stark von den ALICE-Ergebnissen ab. Zum Beispiel weichen die theoretischen Vorhersagen zu den Peakintegralen in Triggerrichtung und Gegenrichtung bei $\sqrt{s}$ = 7.0 TeV, $p_{\mathrm{T,\,trig}} > 0.7\,\mathrm{GeV}/c$ und $p_{\mathrm{T,\,assoc}} > 0.4\,\mathrm{GeV}/c$ mit 40 % bis 120 % von den in ALICE gemessenen Ergebnissen ab. Diese Diskrepanz ist nicht nur bei geringen Transversalimpulsen zu finden sondern erstreckt sich über den gesamten Bereich der vermessenen Transversalimpulse ($p_{\mathrm{T}} > 0.7\,\mathrm{GeV}/c$ bis $p_{\mathrm{T}} > 3.0\,\mathrm{GeV}/c$). Zum Beispiel wird

das Peakintegral in Triggerrichtung bei $\sqrt{s} = 7.0\,\text{TeV}$ und bei geringen Multiplizitäten von den theoretischen Modellen mit 40 % bis 200 % überschätzt. Die Vorhersagen der theoretischen Modelle nähern sich bei hohen Multiplizitäten den ALICE-Ergebnissen an. Hier überschätzen die theoretischen Modelle das Peakintegral in Triggerrichtung mit maximal 100 %; einige Modelle stimmen mit den ALICE-Ergebnissen überein.
Die Anzahl aller Triggerteilchen und die Anzahl unkorrelierter Triggerteilchen werden von allen Modellen bei niedrigen Transversalimpulsen gut beschrieben. Bei hohen Transversalimpulsen hingegen weichen die Modelle mit bis zu 50 % von den ALICE-Ergebnissen ab.
Die beste Übereinstimmung mit allen Eigenschaften der Zwei-Teilchen-Winkelkorrelation findet sich bei Pythia6 Perugia-2011. Pythia6 Perugia-2011 wurde bereits darauf optimiert, einige grundlegende Messgrößen von Proton-Proton-Kollisionen bei $\sqrt{s} = 7.0\,\text{TeV}$ zu reproduzieren. Im Vergleich zu älteren Pythia6-Versionen beinhaltet Perugia-2011 ein erhöhtes Underlying Event und eine weichere Jetframentationsfunktion.
Die Diskrepanzen zwischen den Vorhersagen der theoretischen Modelle und den ALICE-Ergebnisse ist bei $\sqrt{s} = 7.0\,\text{TeV}$ am größten. Wie zu erwarten sind die Diskrepanzen bei vergleichsweise geringeren Kollisionsenergien ($\sqrt{s} = 2.76$ and $0.9\,\text{TeV}$) vermindert, da die theoretischen Modelle, basierend auf Proton-Proton-Kollisionsdaten dieses Energiebereichs, optimiert worden sind.
Das Wissen um die Abweichung der Jetfragmentationsstärke kann verwendet werden, um Eventgeneratoren für den niederenergetischen Bereich zu optimieren, in dem keine störungstheoretische Behandlung der QCD möglich ist.

Die Anzahl der Multi-Parton-Interaktionen wurde anhand der Anzahl unkorrelierter Triggerteilchen abgeschätzt. Es konnte beobachtet werden, dass die Multiplizität geladener Teilchen mit der Anzahl der unkorrelierten Triggerteilchen linear anwächst. Den Erwartungen entsprechend ist zu beobachten, dass jede einzelne Parton-Parton-Interaktion Teilchen produzieren kann. Bei hohen Teilchenmultiplizitäten scheint die Linearität allerdings abzubrechen. Die Anzahl der unkorrelierten Triggerteilchen scheint hier weniger stark als linear mit der Multiplizität geladener Teilchen anzuwachsen. Diese Beobachtung ist konsistent mit der Annahme, dass es eine Grenze in der Anzahl der Multi-Parton-Interaktionen gibt.

In dieser Arbeit wurde eine neue Methode zur Untersuchung der Jetfragmentation und der Multi-Parton-Interaktionenen in Proton-Proton-Kollisionen entwickelt. Diese Methode wurde erfolgreich auf Proton-Proton-Kollisionsdaten am LHC angewendet und hat zu einem besseren Verständnis der Jetfragmentation und des Prozesses der Multi-Parton-Interaktionen beigetragen. Mit Hilfe der gewonnenen Analyseergebnisse können theoretische Modelle für den vom LHC abgedeckten Energiebereich angepasst und verbessert werden.

A. Kinematic Variables

Lorentz Transformations

The space-time vector x^μ of a point in a reference system S is given by the time t and 3 space dimensions x, y, and z

$$x^\mu = (x^0, x^1, x^2, x^3) = (c \cdot t, x, y, z). \tag{A.1}$$

Here, c is the velocity of light. In each inertial system, the velocity of light is the same. β is the ratio of the velocity to the speed of light c

$$\beta = \frac{v}{c}, \tag{A.2}$$

and γ is the Lorentz factor

$$\gamma = \frac{1}{\sqrt{1-\beta^2}}. \tag{A.3}$$

The space-time vector measured in another reference inertial system S' can be estimated in Lorentz transformations

$$\begin{aligned} x^{0'} &= \gamma(x^0 - \beta x^1), \\ x^{1'} &= \gamma(x^1 - \beta x^0), \\ x^{2'} &= x^2, \\ x^{3'} &= x^3. \end{aligned} \tag{A.4}$$

Energy and Momentum

A high-energy collision of two particles A and B can be described by the equation

$$A + B \to C + D. \tag{A.5}$$

C and D are the products of the collision.
A free particle can be described by its energy E and its momentum $\boldsymbol{p} = (p_\mathrm{x}, p_\mathrm{y}, p_\mathrm{z})$,

$$\boldsymbol{p} = \gamma \cdot m_0 \cdot \boldsymbol{v}, \tag{A.6}$$

$$E = \gamma \cdot m_0 \cdot c^2, \tag{A.7}$$

where m_0 is the rest mass of the particle, $\boldsymbol{v} = (v_\mathrm{x}, v_\mathrm{y}, v_\mathrm{z})$ is its velocity.
The energy E and the momentum $\boldsymbol{p}$ of the particle build together the energy-momentum vector $\boldsymbol{P}$ of the particle

$$\boldsymbol{P} = (E/c, \boldsymbol{p}) = (E/c, p_\mathrm{x}, p_\mathrm{y}, p_\mathrm{z}) \overset{c\equiv 1}{=} (E, p_\mathrm{x}, p_\mathrm{y}, p_\mathrm{z}), \tag{A.8}$$

using the common simplification of natural units $c = \hbar = 1$.
The energy conservation in the two-particle collision is given by

$$E_A + E_B = E_C + E_D, \tag{A.9}$$

and the momentum conservation can be described by

$$\boldsymbol{p}_A + \boldsymbol{p}_B = \boldsymbol{p}_C + \boldsymbol{p}_D. \tag{A.10}$$

For a free particle of the rest mass m_0, the energy E of the particle is

$$E = \sqrt{m_0^2 \cdot c^4 + \boldsymbol{p}^2 \cdot c^2} \overset{c\equiv 1}{=} \sqrt{m_0^2 + \boldsymbol{p}^2}. \tag{A.11}$$

The momentum $\boldsymbol{p}$ of a particle traveling in the coordinate system of ALICE (cf. Appendix B) can be divided into two parts: The fraction of the momentum that points in direction of the beam axis, the so-called longitudinal momentum p_{L}, and the fraction of the momentum perpendicular to beam axis, the so-called transverse momentum p_{T}.

$$p = |\boldsymbol{p}|, \tag{A.12}$$

$$p_{\mathrm{L}} = p_{\mathrm{z}} = p \cdot \cos\vartheta, \tag{A.13}$$

$$p_{\mathrm{T}} = \sqrt{p_{\mathrm{x}}^2 + p_{\mathrm{xy}}^2} = p \cdot \sin\vartheta. \tag{A.14}$$

Here, ϑ is the polar angle defined in the global coordinate system of ALICE.
The transverse momentum is invariant under Lorentz transformations. The longitudinal momentum, however, is variant under Lorentz transformations.

Rapidity and Pseudorapidity

The rapidity y is derived from the longitudinal momentum

$$y = \frac{1}{2} \ln\left(\frac{E + p_{\mathrm{L}}}{E - p_{\mathrm{L}}}\right). \tag{A.15}$$

In contrast to the longitudinal momentum, the rapidity is invariant under Lorentz transformations. The exact computation of the rapidity is complicated as the energy E is only known for identified particles. However, E can be substituted for the special case, where E is much larger than the rest mass of the particle m_0:

$$E = \sqrt{m_0^2 + \boldsymbol{p}^2} \overset{E\gg m_0}{=} p. \tag{A.16}$$

Replacing E with p, the defintion of the so-called pseudorapidity η is given

$$\eta = \frac{1}{2} \ln\left(\frac{p + p_{\mathrm{L}}}{p - p_{\mathrm{L}}}\right). \tag{A.17}$$

The pseudorapidity can be directly derived from the polar angle ϑ of the ALICE coordinate system

$$\eta = -\ln \tan \frac{\vartheta}{2}. \tag{A.18}$$

For ultra-relativistic particles with $E \gg m_0$, the pseudorapidity and the rapidity are equal.

Mandelstam Variables

In the collision described in Equation A.5, the so-called Mandelstam variables are defined as

$$s = (\boldsymbol{P}_A + \boldsymbol{P}_B)^2 = (\boldsymbol{P}_C + \boldsymbol{P}_D)^2, \tag{A.19}$$

$$t = (\boldsymbol{P}_A - \boldsymbol{P}_C)^2 = (\boldsymbol{P}_B - \boldsymbol{P}_D)^2, \tag{A.20}$$

$$u = (\boldsymbol{P}_A - \boldsymbol{P}_D)^2 = (\boldsymbol{P}_B - \boldsymbol{P}_C)^2. \tag{A.21}$$

Using E_i^* as the energy in the center-of-mass system of the colliding system of A and B, s can be transformed to

$$s = (E_A^* + E_B^*)^2. \tag{A.22}$$

Hence, $\sqrt{s}$ is the total energy of the colliding system of the two particles in their center-of-mass system. In heavy-ion collisions, the collision energy is commonly given by the center-of-mass energy of a nucleon-nucleon pair of the two ions called $\sqrt{s_{\mathrm{NN}}}$.
$\sqrt{t}$ can be interpreted as the momentum transfer of the collision.

B. The Global Coordinate System of ALICE

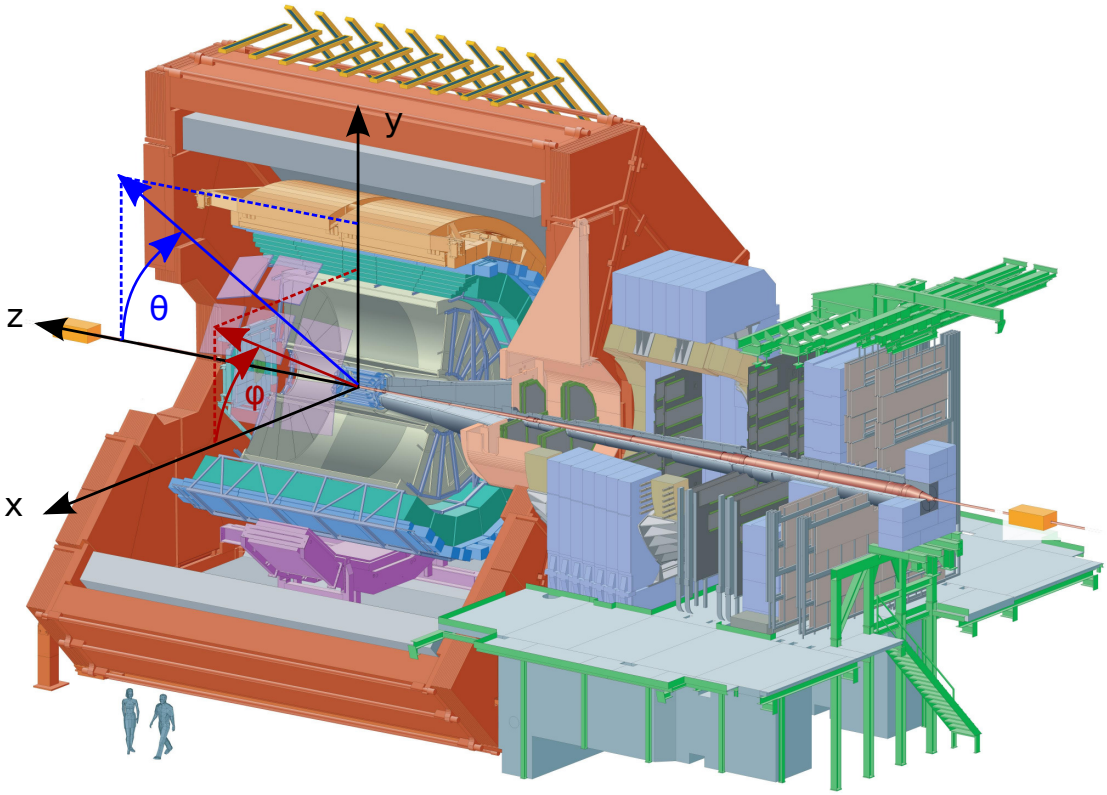

Figure B.1.: Schematic view of the ALICE detector [ALI08] (cf. Chapter 3) together with the global coordinate system of ALICE.

The global coordinate system of ALICE is shown in Figure B.1.

- The origin of the global coordinated system is given by the interaction point of the two beams. The interaction point is located at the mid point of the Inner Tracking System (cf. Section 3.3). ALICE can be described in a Cartesian coordinate system and in a cylindrical coordinate system.
- The z-axis is given by the beam direction. The muon arm is placed at negative z.
- The y-axis is perpendicular to the beam line and it is vertical. The positive y-axis points from the origin upwards.
- The x-axis is perpendicular to the beam direction and it is horizontal. The direction of positive x points from the center of ALICE to the center of the LHC.

- The Cartesian coordinates in the x-direction and the y-direction can be combined to a single cylindrical coordinate, the radius, $r = \sqrt{x^2 + y^2}$.
- The angle φ is the azimuthal angle. As the global coordinate system is a right-handed system, φ increases clockwise from the perspective at negative z. At the x-axis, φ is 0 and at the y-axis φ is $\pi/2$.
- The angle θ is the polar angle. At positive z, θ has the value 0. At the xy-plane, θ has the value $\pi/2$.
- Instead of the polar angle, often the Lorentz-invariant pseudorapidty η is used to describe the track direction (cf. Appendix A).

C. List of Analyzed Data Sets

In this appendix, the data samples used in the data analysis are listed. In the Tables C.1, C.2, and C.3 the ALICE data sets defined by the run numbers are shown. In the Tables C.4, C.5, and C.6, the corresponding Monte Carlo data sets are listed.

pp @ $\sqrt{s}$ = 0.9 TeV	
Data period	Accepted Data Sets (Run Numbers)
LHC10c (pass 3)	118506, 118507, 118512, 118518, 118556, 118558, 118560, 118561, 121039, 121040

Table C.1.: List of internal run numbers of the data sets measured at $\sqrt{s}$ = 0.9 TeV. The data sets have been collected in May 2010. The magnetic field of the L3 magnet has been set to $B = +5$ kG. The mean number of interactions per bunch crossings in these data sets varies between $\mu = 0.0133$ and 0.0319.

pp @ $\sqrt{s}$ = 2.76 TeV	
Data period	Accepted Data Sets (Run Numbers)
LHC11a (pass 2)	146746, 146747, 146748, 146801, 146802, 146803, 146804, 146805, 146806, 146807, 146817, 146824, 146856, 146858, 146859, 146860

Table C.2.: List of internal run numbers of the data sets measured at $\sqrt{s}$ = 2.76 TeV. The data sets have been collected in March 2011. The magnetic field of the L3 magnet has been set to $B = -5$ kG. The mean number of interactions per bunch crossing in these data sets varies between $\mu = 0.03$ and 0.05.

pp @ $\sqrt{s}$ = 7.0 TeV	
Data period	Accepted Data Sets (Run Numbers)
LHC10b (pass 3)	114931, 115186, 115193, 115393, 115401, 115414, 116102, 116288, 116402, 116403, 116562, 116571, 116574, 116643, 116645, 117048, 117050, 117052, 117053, 117059, 117060, 117063, 117092, 117099, 117109, 117112, 117116, 117220, 117222
LHC10c (pass 3)	119841, 119844, 119845, 119846, 119849, 119853, 119856, 119859, 119862, 120067, 120069, 120072, 120073, 120076, 120079, 120244, 120503, 120505, 120616, 120617, 120671, 120741, 120750, 120758, 120820, 120821, 120822, 120823, 120824, 120825, 120829
LHC10d (pass 2)	122374, 122375, 125023, 125085, 125097, 125100, 125101, 125133, 125134, 125139, 125140, 125156, 125186, 125296, 125630, 125632, 125633, 125842, 125843, 125844, 125847, 125848, 125849, 125850, 125851, 125855, 126004, 126007, 126008, 126073, 126078, 126081, 126082, 126088, 126090, 126097, 126158
LHC10e (pass 2)	127712, 127714, 127718, 127719, 128495, 128498, 128503, 128504, 128505, 128507, 128582, 128605, 128615, 128621, 128677, 128777, 128778, 128820, 128823, 128824, 128835, 128836, 128843, 128850, 128853, 128855, 128913, 129512, 129513, 129514, 129599, 129639, 129641, 129652, 129653, 129659, 129667, 130157, 130158, 130172, 130179, 130375, 130519, 130601, 130608, 130696, 130704, 130793, 130798, 130799, 130834, 130840

Table C.3.: List of internal run numbers of data sets measured at $\sqrt{s}$ = 7.0 TeV. The data sets have been recorded between April and August 2010. The magnetic field of the L3 magnet has been set to $B = -5$ kG as well as $B = +5$ kG. The mean number of interactions per bunch crossing in these data sets varies between $\mu = 0.004$ and 0.16.

pp @ $\sqrt{s}$ = 0.9 TeV				
Data period				Events
LHC10c (pass 3)		cf. Table C.1		7,430,000
MC period	Anchor	Generator	Tune	Events
LHC10e12	LHC10c	Pythia6	Perugia-0	7,309,200
LHC10e13	LHC10c	Phojet		7,322,400
-	MC only	Pythia8	4C	15,800,000
-	MC only	Pythia6	Perugia-2011	9,910,000

Table C.4.: ALICE data sets and Monte Carlo data sets used in the data analysis at $\sqrt{s}$ = 0.9 TeV and the number of events per data set.

pp @ $\sqrt{s}$ = 2.76 TeV				
Data period				Events
LHC11a (pass 2)		cf. Table C.2		20,000,000
MC period	Anchor	Generator	Tune	Events
LHC11b10a	LHC11a	Pythia6	Perugia-0	4,300,000
LHC11b12a	LHC11a	Pythia6	Perugia-0	2,942,000
LHC11e3a	LHC11a	Pythia6	Perugia-0	5,020,000
LHC11e3a_plus	LHC11a	Pythia6	Perugia-0	49,968,600
LHC11b12b	LHC11a	Phojet		3,398,400
-	MC only	Pythia8	4C	12,440,000
-	MC only	Pythia6	Perugia-2011	10,000,000

Table C.5.: ALICE data sets and Monte Carlo data sets used in the data analysis at $\sqrt{s}$ = 2.76 TeV and the number of events per data set.

pp @ $\sqrt{s}$ = 7.0 TeV				
Data period				Events
LHC10b (pass 3)		cf. Table C.3		25,000,000
LHC10c (pass 3)		”		67,000,000
LHC10d (pass 2)		”		107,000,000
LHC10e (pass 2)		”		71,000,000
MC period	Anchor	Generator	Tune	Events
LHC10d1	LHC10b	Pythia6	Perugia-0	25,000,000
LHC10d2	LHC10b	Phojet		27,000,000
LHC10d4	LHC10c	Pythia6	Perugia-0	61,000,000
LHC10d4a	LHC10c	Phojet		63,000,000
LHC10f6a	LHC10d	Pythia6	Perugia-0	103,000,000
LHC10f6	LHC10d	Phojet		103,000,000
LHC10e20	LHC10e	Pythia6	Perugia-0	3,600,000
LHC10e21	LHC10e	Phojet		3,800,000
-	MC only	Pythia8	4C	14,000,000
-	MC only	Pythia6	Perugia-2011	34,000,000

Table C.6.: ALICE data sets and Monte Carlo data sets used in the data analysis at $\sqrt{s}$ = 7.0 TeV and the number of events per data set.

D. Distance of Closest Approach

The distance of closest approach (DCA) of a track to the collision vertex is the minimal distance between the trajectory and the vertex position. The DCA is calculated by backwards propagation of the track in direction of the vertex and simultaneous searching for the minimal distance between track and vertex. Often, the DCA is also called impact parameter d. The principle of the DCA in xy-direction and z-direction is sketched in Figure D.1. Here, the same coordinate system definition is used as in the ALICE coordinate system. A magnetic field is pointing in z-direction just like it is generated in the L3 magnet of ALICE.

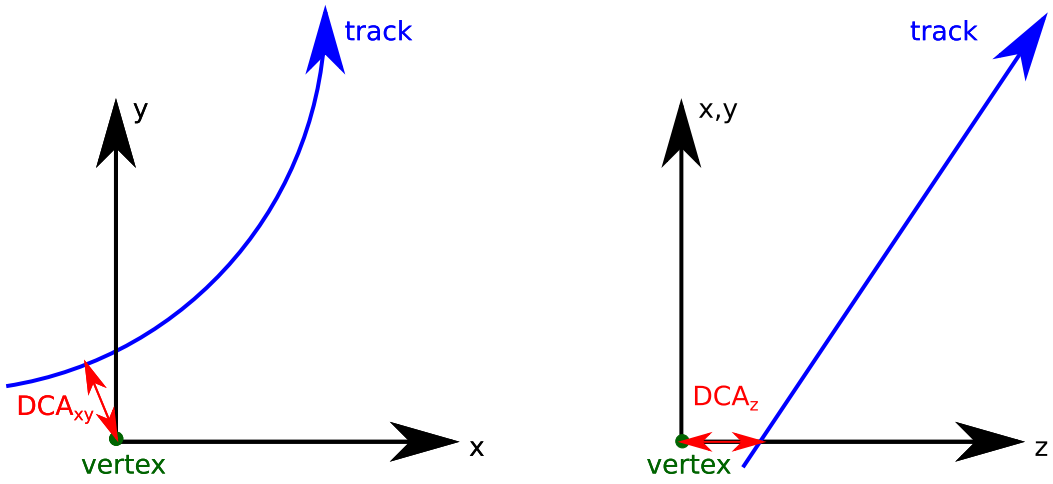

Figure D.1.: Left panel: Sketch of the distance of closest approach in xy-direction for a track of a charged particle. The particle follows a circular trajectory in the xy-plane due the Lorentz force through the magnetic field which points in the direction perpendicular to the xy-plane. Right panel: Sketch of the distance of closest approach in z-direction.

The trajectories of primary particles generated in the collision originate in the collision vertex. Due to a limited detector resolution, the reconstructed track DCA can differ slightly from zero. Trajectories of secondary particles often have a DCA which is different from zero. Secondary particles are generated by e. g. weak decays of strange hadrons, gamma conversions, or interactions with the material. The production of these particles happens at secondary vertices mostly not identical with the primary collision vertex. Applying a selection cut on the DCA of the tracks allows to select predominantly tracks from primary or secondary particles only.

E. Evolution of Track and Event Parameters

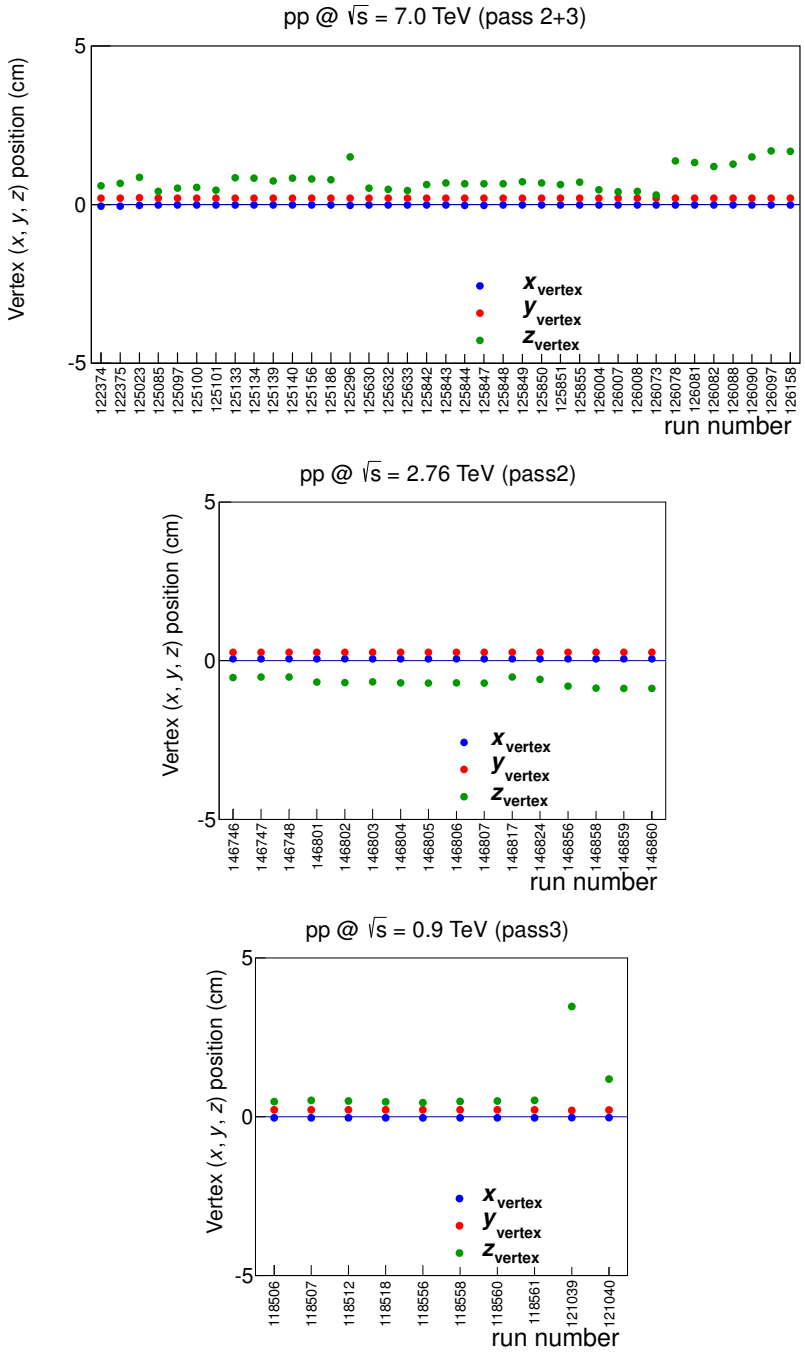

Figure E.1.: Vertex position in x, y, and z-direction measured at $\sqrt{s} = 0.9$, 2.76, and 7.0 TeV.

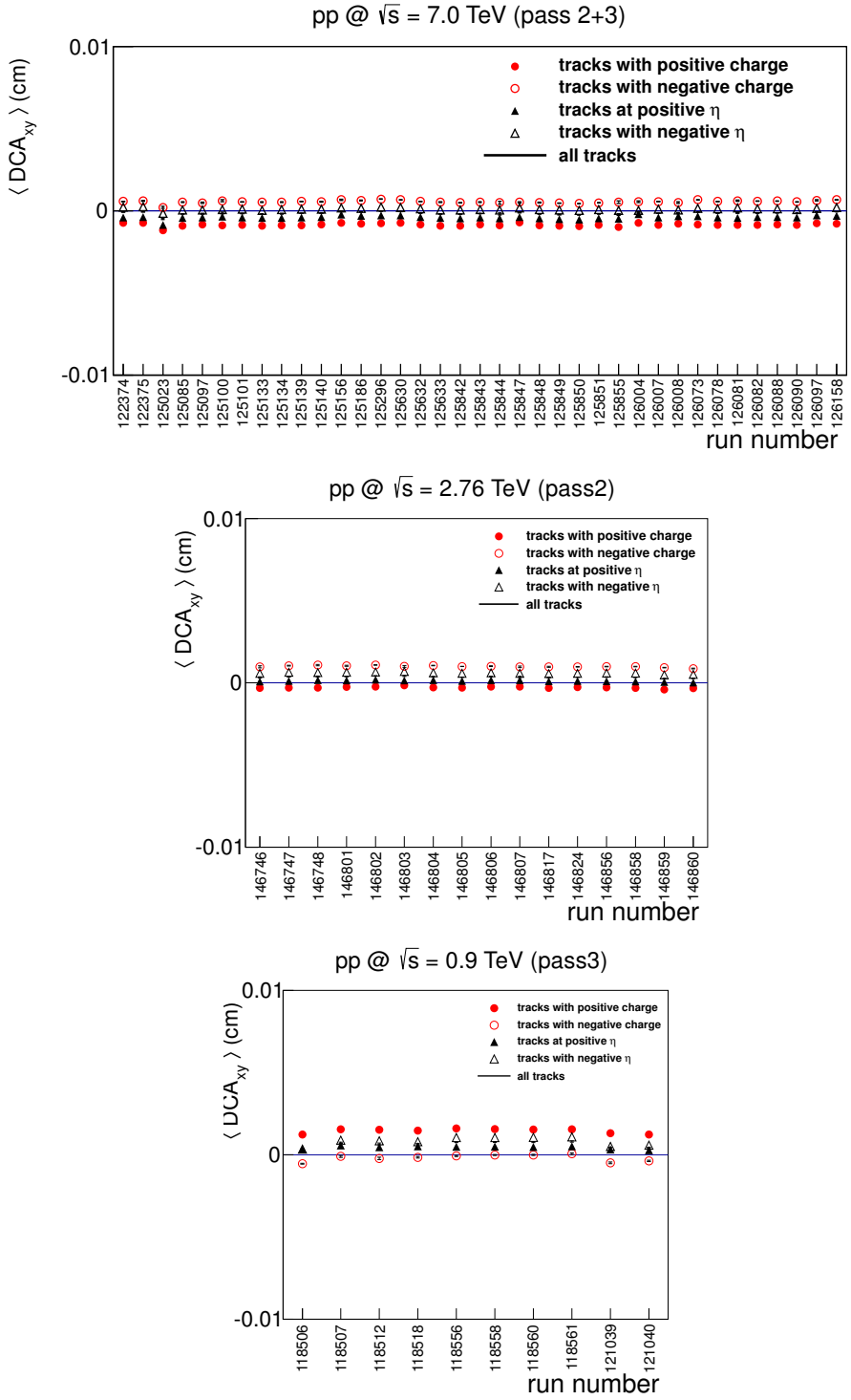

Figure E.2.: The mean values of the DCA_{xy} per run measured at $\sqrt{s} = 0.9$, 2.76, and 7.0 TeV.

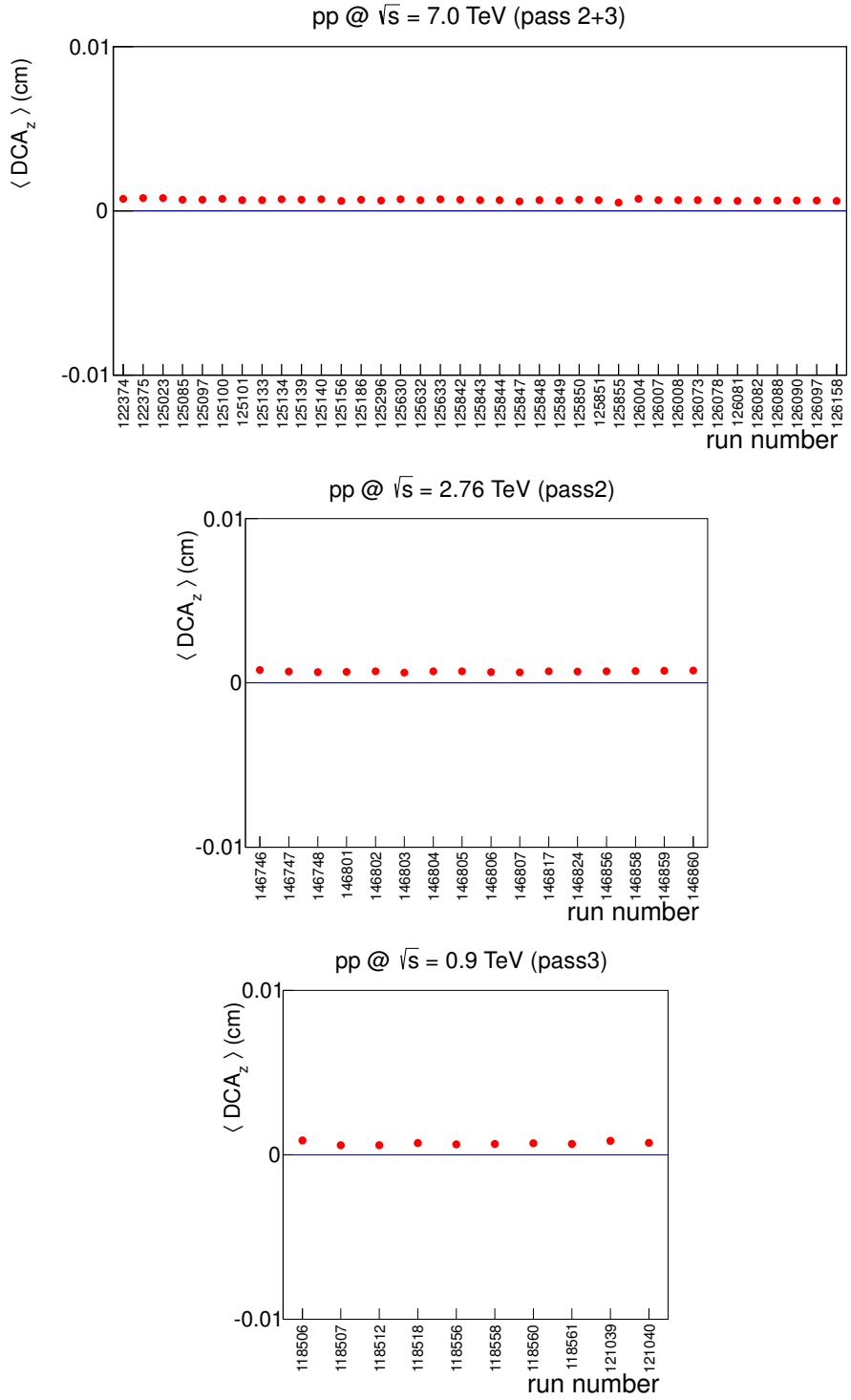

Figure E.3.: The mean values of the DCA_z per run measured at $\sqrt{s} = 0.9$, 2.76, and 7.0 TeV.

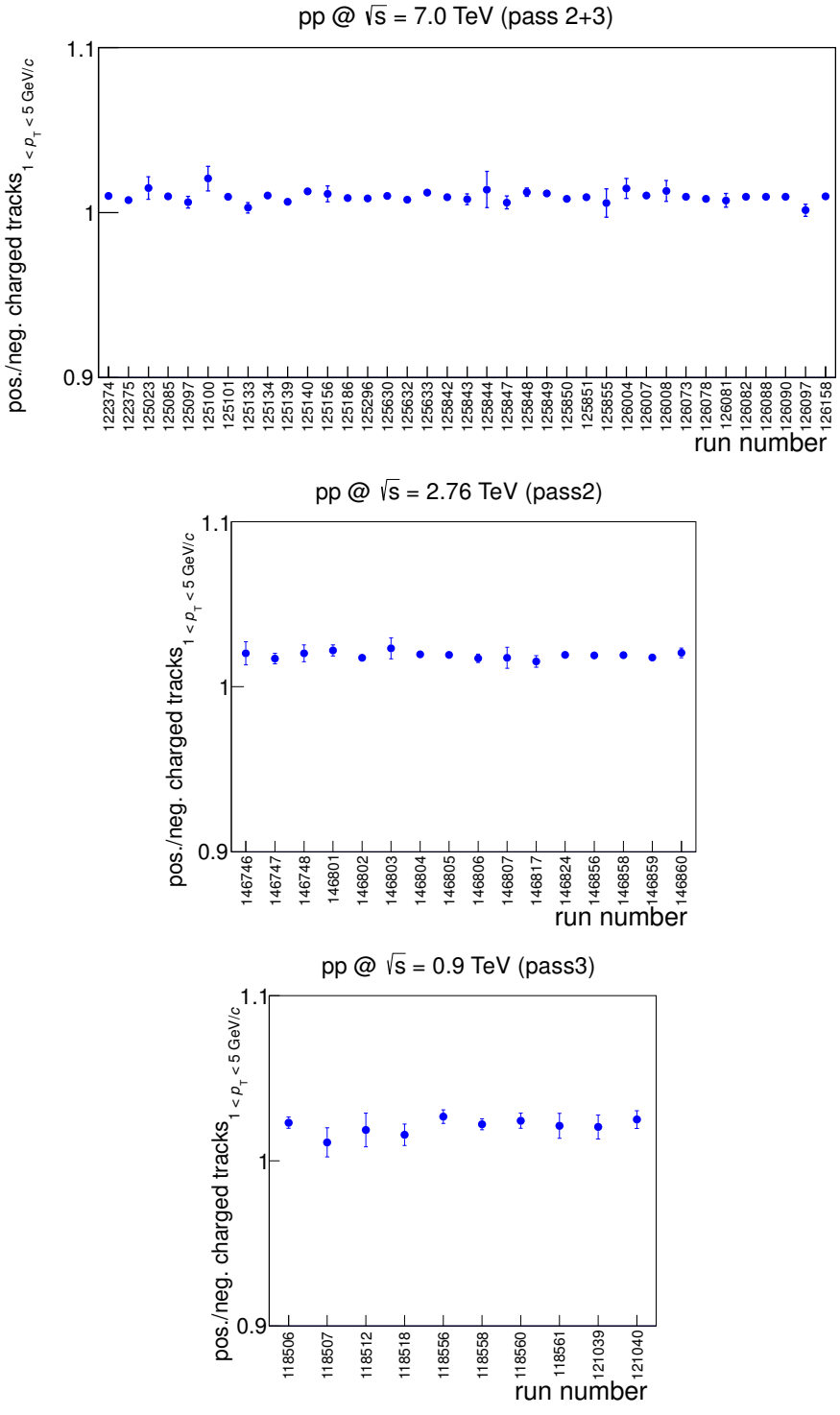

Figure E.4.: The average ratio of tracks of positive charge to tracks of negative charge per run measured at $\sqrt{s} = 0.9$, 2.76, and 7.0 TeV. Only tracks with a reconstructed transverse momentum of $1.0 < p_{\mathrm{T}} < 5.0$ GeV/c are shown.

Bibliography

[AAB+98] T. Alexopoulos, E. Anderson, N. Biswas, et al. *The Role of Double Parton Collisions in Soft Hadron Interactions.* Phys.Lett., **B435** pp. 453–457, 1998.

[ABB+93] N. Angert, M. Bourgarel, E. Brouzet, et al. *CERN Heavy Ion Facility Design Report*, 1993.

[AFS87] AFS Collaboration. *Double Parton Scattering in Proton-Proton Collions at* $\sqrt{s} = 63\,GeV$. Z.Phys., **C34** p. 163, 1987.

[ALI98] ALICE Collaboration. *ALICE Technical Design Report: Detector for High Momentum PID.* Tech. Rep. CERN-LHCC-98-19, 1998.

[ALI99a] ALICE Collaboration. *ALICE Technical Design Report of the Dimuon Forward Spectrometer.* Tech. Rep. CERN-LHCC-99-22, 1999.

[ALI99b] ALICE Collaboration. *ALICE Technical Design Report of the Inner Tracking System (ITS).* Tech. Rep. CERN-LHCC-99-12, 1999.

[ALI99c] ALICE Collaboration. *ALICE Technical Design Report of the Photon Spectrometer (PHOS).* Tech. Rep. CERN-LHCC-99-04, 1999.

[ALI99d] ALICE Collaboration. *ALICE Technical Design Report of the Zero Degree Calorimeter (ZDC).* Tech. Rep. CERN-LHCC-99-05, 1999.

[ALI99e] ALICE Collaboration. *ALICE Technical Design Report: Photon Multiplicity Detector (PMD).* Tech. Rep. CERN-LHCC-99-32, 1999.

[ALI00] ALICE Collaboration. *ALICE: Technical Design Report of the Time Projection Chamber.* Tech. Rep. CERN-LHCC-2000-001, 2000.

[ALI01] ALICE Collaboration. *ALICE Transition Radiation Detector: Technical Design Report.* Tech. Rep. CERN-LHCC-2001-021, 2001.

[ALI02] ALICE Collaboration. *ALICE: Addendum to the Technical Design Report of the Time of Flight System (TOF).* Tech. Rep. CERN-LHCC-2002-016, 2002.

[ALI04] ALICE Collaboration. *ALICE Technical Design Report on Forward Detectors: FMD, T0 and V0.* Tech. Rep. CERN-LHCC-2004-025, 2004.

[ALI05] *ALICE Technical Design Report of the Computing.* Tech. Rep. CERN-LHCC-2005-018, 2005.

[ALI08] ALICE Collaboration. *The ALICE Experiment at the CERN LHC.* JINST, **3** p. S08002, 2008.

[ALI10a] ALICE Collaboration. *First Proton-Proton Collisions at the LHC as Observed with the ALICE Detector: Measurement of the Charged Particle Pseudorapidity Density at $\sqrt{s}$ = 900 GeV*. Eur.Phys.J., **C65** pp. 111–125, 2010. arXiv:hep-ex/0911.5430.

[ALI10b] ALICE Collaboration. *Performance of Prototypes for the ALICE Electromagnetic Calorimeter*. Tech. rep., 2010. arXiv:physics.ins-det/0912.2005.

[ALI10c] ALICE Collaboration. *Transverse Momentum Spectra of Charged Particles in Proton-Proton Collisions at $\sqrt{s}$ = 900 GeV with ALICE at the LHC*. Phys.Lett., **B693** pp. 53–68, 2010. arXiv:hep-ex/1007.0719.

[ALI11a] ALICE Collaboration. *Femtoscopy of pp Collisions at $\sqrt{s}$ = 0.9 and 7 TeV at the LHC with Two-Pion Bose-Einstein Correlations*, 2011. arXiv:hep-ex/1101.3665.

[ALI11b] ALICE Collaboration. *Higher Harmonic Anisotropic Flow Measurements of Charged Particles in Pb–Pb Collisions at $\sqrt{s_{\mathrm{NN}}}$ = 2.76 TeV*. Phys.Rev.Lett., **107** p. 032301, 2011. arXiv:nucl-ex/1105.3865.

[ALI11c] ALICE Collaboration. *Production of Pions, Kaons and Protons in pp Collisions at $\sqrt{s}$ = 900 GeV with ALICE at the LHC*. Eur.Phys.J., **C71** p. 1655, 2011. arXiv:hep-ex/1101.4110.

[ALI11d] ALICE Collaboration. *Strange Particle Production in Proton-Proton Collisions at $\sqrt{s}$ = 0.9 TeV with ALICE at the LHC*. Eur.Phys.J., **C71** p. 1594, 2011. arXiv:hep-ex/1012.3257.

[ALI11e] ALICE Collaboration. *Suppression of Charged Particle Production at Large Transverse Momentum in Central Pb–Pb Collisions at $\sqrt{s_{\mathrm{NN}}}$ = 2.76 TeV*. Phys.Lett., **B696** pp. 30–39, 2011. arXiv:nucl-ex/1012.1004.

[ALI11f] ALICE Collaboration. *Underlying Event Measurements in pp Collisions at $\sqrt{s}$ = 0.9 and 7 TeV with the ALICE Experiment at the LHC*, 2011. arXiv:hep-ex/1112.2082.

[ALI12a] ALICE Collaboration. *ALICE Run Condition Table Hosted by the MonALISA Repository for ALICE*, 2012. URL http://alimonitor.cern.ch/configuration/.

[ALI12b] ALICE Collaboration. *Harmonic Decomposition of Two-Particle Angular Correlations in Pb–Pb Collisions at $\sqrt{s_{\mathrm{NN}}}$ = 2.76 TeV*. Phys.Lett., **B708** pp. 249–264, 2012. arXiv:nucl-ex/1109.2501.

[ALI12c] ALICE Collaboration. *J/ψ Production as a Function of Charged Particle Multiplicity in pp Collisions at $\sqrt{s}$ = 7.0 TeV*, 2012. arXiv:hep-ex/1202.2816.

[ALI12d] ALICE Collaboration. *A Large Ion Collider Experiment*, 2012. URL http://aliceinfo.cern.ch.

[ALI12e] ALICE Collaboration. *Measurement of Charm Production at Central Rapidity in Proton-Proton Collisions at $\sqrt{s}$ = 7 TeV*. JHEP, **1201** p. 128, 2012. arXiv:hep-ex/1111.1553.

[ALI12f] ALICE Collaboration. *Particle-Yield Modification in Jet-Like Azimuthal Di-Hadron Correlations in Pb–Pb Collisions at $\sqrt{s_{\mathrm{NN}}}$ = 2.76 TeV*. Phys.Rev.Lett., **108** p. 092301, 2012. arXiv:nucl-ex/1110.0121.

[ALI12g] ALICE Collaboration. *Transverse Sphericity of Primary Charged Particles in Minimum Bias Proton-Proton Collisions at $\sqrt{s}$ = 0.9, 2.76 and 7 TeV*, 2012. arXiv:hep-ex/1205.3963.

[ATL08] ATLAS Collaboration. *The ATLAS Experiment at the CERN Large Hadron Collider*. JINST, **3** p. S08003, 2008.

[Axi82a] Axial Field Spectrometer Collaboration. *Direct Evidence for the Emergence of Jets in Events Triggered on Large Transverse Energy in pp Collisions at $\sqrt{s}$ = 63 GeV*. Phys.Lett., **B118** p. 185, 1982.

[Axi82b] Axial Field Spectrometer Collaboration. *Further Evidence for Jet Production From The Charged Particles Pproduced in pp Collisions at $\sqrt{s}$ = 63 GeV Triggered on Large Transverse Energy*. Phys.Lett., **B118** p. 193, 1982.

[B+84] A. Breakstone et al. *Charged Multiplicity Distribution in Proton-Proton Interactions at ISR Energies*. Phys.Rev., **D30** p. 528, 1984.

[BBB+00] E. Benedikt, M., A. Blas, J. Borburgh, et al. *The PS Complex as Proton Pre-Injector for the LHC: Design and Implementation Report*, 2000.

[BBB+05] I. Bird, K. Bos, N. Brook, et al. *LHC Computing Grid. Technical Design Report*. Tech. Rep. CERN-LHCC-2005-024, 2005.

[BBC+03] R. Brun, P. Buncic, F. Carminati, et al. *Computing in ALICE*. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **502** pp. 339–346, 2003.

[BBM+87] R. Brun, F. Bruyant, M. Maire, et al. *GEANT3*. Tech. Rep. DD/EE/84-1, CERN, 1987.

[BCL+04a] O. S. Brüning, P. Collier, P. Lebrun, et al. *LHC Design Report. 1. The LHC Main Ring*. Tech. Rep. CERN-2004-003-V-1, 2004.

[BCL+04b] O. S. Brüning, P. Collier, P. Lebrun, et al. *LHC Design Report. 2. The LHC Infrastructure and General Services*. Tech. Rep. CERN-2004-003-V-2, 2004.

[BCM+04] M. Benedikt, P. Collier, V. Mertens, et al. *LHC Design Report. 3. The LHC Injector Chain*. Tech. Rep. CERN-2004-003-V-3, 2004.

[Bet09] S. Bethke. *The 2009 World Average of α_s*. Eur.Phys.J., **C64** pp. 689–703, 2009. arXiv:hep-ph/0908.1135.

[BF+10] P. Bartalini, L. Fano, et al. *1st International Workshop on Multiple Partonic Interactions at the LHC, Perugia, Italy, 2008 (Proceedings)*, 2010. arXiv:hep-ph/1003.4220.

[BF11] P. Bartalini and L. Fanò. *Multiple Parton Interactions Studies at CMS*, 2011. arXiv:hep-ex/1103.6201.

[Bjo69] J. Bjorken. *Asymptotic Sum Rules at Infinite Momentum.* Phys.Rev., **179** pp. 1547–1553, 1969.

[BP69] J. Bjorken and E. A. Paschos. *Inelastic Electron-Proton and Gamma-Proton Scattering, and the Structure of the Nucleon.* Phys.Rev., **185** pp. 1975–1982, 1969.

[BR97] R. Brun and F. Rademakers. *ROOT: An Object Oriented Data Analysis Framework.* Nucl.Instrum.Meth., **A389** pp. 81–86, 1997.

[Bud67] G. Budker. *An Effective Method of Damping Particle Oscillations in Proton and Antiproton Storage Rings.* Sov.Atom.Energ., **22** pp. 438–440, 1967.

[C+94] A. Capella et al. *Dual Parton Model.* Phys.Rept., **236** pp. 225–329, 1994.

[CDF89] CDF Collaboration. *Measurement of the Inclusive Jet Cross-Section in $\bar{p}p$ Collisions at $\sqrt{s}$ = 1.8 TeV.* Phys.Rev.Lett., **62** p. 613, 1989.

[CDF90] CDF Collaboration. *The Two Jet Invariant Mass Distribution at $\sqrt{s}$ = 1.8 TeV.* Phys.Rev., **D41** pp. 1722–1725, 1990.

[CDF93] CDF Collaboration. *Study of Four Jet Events and Evidence For Double Parton Interactions in $p\bar{p}$ Collisions at $\sqrt{s}$ = 1.8 TeV.* Phys.Rev., **D47** pp. 4857–4871, 1993.

[CDF97a] CDF Collaboration. *Double Parton Scattering in $\bar{p}p$ Collisions at $\sqrt{s}$ = 1.8 TeV.* Phys.Rev., **D56** pp. 3811–3832, 1997.

[CDF97b] CDF Collaboration. *Measurement of Double Parton Scattering in $\bar{p}p$ Collisions at $\sqrt{s}$ = 1.8 TeV.* Phys.Rev.Lett., **79** pp. 584–589, 1997.

[CER83] CERN-Oxford-Rockefeller Collaboration. *Observation of Jet Structure in High Transverse Energy Events at the CERN Intersecting Storage Rings.* Phys.Lett., **B126** p. 132, 1983.

[Cha04] M. Chanel. *LEIR: The Low Energy Ion Ring at CERN.* Nucl.Instrum.Meth., **A532** pp. 137–143, 2004.

[CKM+02] G. Corcella, I. Knowles, G. Marchesini, et al. *HERWIG 6.5 Release Note*, 2002. arXiv:hep-ph/0210213.

[CMS08] CMS Collaboration. *The CMS Experiment at the CERN LHC.* JINST, **3** p. S08004, 2008.

[CMS10] CMS Collaboration. *Observation of Long-Range Near-Side Angular Correlations in Proton-Proton Collisions at the LHC.* JHEP, **1009** p. 091, 2010. arXiv:hep-ex/1009.4122.

[CMS11a] CMS Collaboration. *Charged Particle Multiplicities in pp Interactions at $\sqrt{s}$ = 0.9, 2.36, and 7 TeV*. JHEP, **1101** p. 079, 2011. 1011.5531.

[CMS11b] CMS Collaboration. *Measurement of the Differential Dijet Production Cross Section in Proton-Proton Collisions at $\sqrt{s}$ = 7 TeV*. Phys.Lett., **B700** pp. 187–206, 2011. arXiv:hep-ex/1104.1693.

[CMS11c] CMS Collaboration. *Strange Particle Production in pp Collisions at $\sqrt{s}$ = 0.9 and 7 TeV*. JHEP, **1105** p. 064, 2011. arXiv:hep-ex/1102.4282.

[CPT75] H.-M. Chan, J. E. Paton, and S. T. Tsou. *Diffractive Scattering in the Dual Model*. Nucl.Phys., **B86** p. 479, 1975.

[CR76] G. Chew and C. Rosenzweig. *Asymptotic Planarity: An S Matrix Basis for the Okubo-Zweig-Iizuka Rule*. Nucl.Phys., **B104** p. 290, 1976.

[CR78] G. F. Chew and C. Rosenzweig. *Dual Topological Unitarization: An Ordered Approach to Hadron Theory*. Phys.Rept., **41** pp. 263–327, 1978.

[CS11] R. Corke and T. Sjostrand. *Interleaved Parton Showers and Tuning Prospects*. JHEP, **1103** p. 032, 2011. arXiv:hep-ph/1011.1759.

[DJD+11] S. Dutta, J. Jasvantlal, A. Dewanto, et al. *Extraction of Clan Model Parameters from Multiplicity Distributions Measured in Proton-Proton Collisions at LHC Energies*. Mod.Phys.Lett., **A26** pp. 2951–2961, 2011.

[DN04] I. Dremin and V. Nechitailo. *Independent Pair Parton Interactions-Model of Hadron Interactions*. Phys.Rev., **D70** p. 034005, 2004. arXiv:hep-ph/0402286.

[DN11] I. Dremin and V. Nechitailo. *Soft Multiple Parton Interactions as seen in Multiplicity Distributions at Tevatron and LHC*. Phys.Rev., **D84** p. 034026, 2011. arXiv:hep-ph/1106.4959.

[Dur12] Durham HepData Project. *Online PDF plotting and calculation*, 2012. URL http://hepdata.cedar.ac.uk/pdf/pdf3.html.

[EB64] F. Englert and R. Brout. *Broken Symmetry and the Mass of Gauge Vector Mesons*. Phys.Rev.Lett., **13** pp. 321–323, 1964.

[EB08] L. Evans and P. Bryant. *LHC Machine*. JINST, **3** p. S08001, 2008.

[Eks85] G. Ekspong. *On Scale Breaking in Multiplicities and a New Empirical Rule*. Conf.Proc., **C850609** pp. 309–320, 1985.

[Ell02] J. R. Ellis. *Limits of the Standard Model*, 2002. arXiv:hep-ph/0211168.

[Eng95] R. Engel. *Photoproduction within the Two Component Dual Parton Model. 1. Amplitudes and Cross-Sections*. Z.Phys., **C66** pp. 203–214, 1995.

[ERR95] R. Engel, J. Ranft, and S. Roesler. *Hard Diffraction in Hadron Hadron Interactions and in Photoproduction*. Phys.Rev., **D52**(SI-95-16) pp. 1459–1468, 1995. arXiv:hep-ph/9502319.

[ESW96] R. K. Ellis, W. J. Stirling, and B. Webber. *QCD and collider physics*. Camb.Monogr.Part.Phys.Nucl.Phys.Cosmol., **8** pp. 1–435, 1996.

[Fab12] Fabiola Gianotti on behalf of the ATLAS Collaboration. *Status of Standard Model Higgs Searches in ATLAS*. Presentation at the CERN Seminar on "Latest Update in the Search for the Higgs boson", 2012.

[Fey69] R. P. Feynman. *Very High-Energy Collisions of Hadrons*. Phys.Rev.Lett., **23** pp. 1415–1417, 1969.

[FFRS93] A. Fasso, A. Ferrari, J. Ranft, et al. *FLUKA: Present Status and Future Developments*. Conf.Proc., **C9309194** pp. 493–502, 1993.

[FKT01] I. Foster, C. Kesselman, and S. Tuecke. *The Anatomy of the Grid*. Int. Journal of High Performance Applications, **15**(3), 2001.

[Fos02] I. Foster. *What is the Grid? A Three Point Checklist*. GRID today, **1**(6) pp. 22–25, 2002.

[Fug89] C. Fuglesang. *UA5 Multiplicity Distribution And Fits Of Various Functions*, 1989.

[Gea03] Geant4 Collaboration. *Geant4: A Simulation Toolkit*. Nucl.Instrum.Meth., **A506** pp. 250–303, 2003.

[Gea06] Geant4 Collaboration. *Geant4 Developments and Applications*. IEEE Trans.Nucl.Sci., **53** p. 270, 2006.

[Ghe08] A. Gheata. *ALICE Analysis Framework*. PoS, **ACAT08** p. 028, 2008.

[GHK64] G. Guralnik, C. Hagen, and T. Kibble. *Global Conservation Laws and Massless Particles*. Phys.Rev.Lett., **13** pp. 585–587, 1964.

[GHM86] T. Gaisser, F. Halzen, and A. D. Martin. *Multiplicities In a QCD Motivated Description of Very High-Energy Particle Interactions*. Phys.Lett., **B166** p. 219, 1986.

[Gho12] P. Ghosh. *Negative Binomial Multiplicity Distribution in Proton-Proton Collisions in Limited Pseudorapidity Intervals at LHC up to* $\sqrt{s} = 7\,TeV$ *and the Clan model*, 2012. `arXiv:hep-ph/1202.4221`.

[GHS80] C. Goebel, F. Halzen, and D. Scott. *Double Drell-Yan Annihilations in Hadron Collisions: Novel Tests of the Constituent Picture*. Phys.Rev., **D22** p. 2789, 1980.

[Gia07] G. Giacomelli. *Rising Total Hadron-Hadron Cross-Sections*, 2007. `arXiv:hep-ex/0712.0906`.

[GM64] M. Gell-Mann. *A Schematic Model of Baryons and Mesons*. Phys.Lett., **8** pp. 214–215, 1964.

[GM69] V. Gribov and A. A. Migdal. *Properties of the Pomeranchuk Pole and the Branch Cuts Related to it at Low Momentum Transfer*. Sov.J.Nucl.Phys., **8** pp. 583–590, 1969.

[GMFL11] B. Gorini, E. Meschi, and M. Ferro-Luzzi. *LHC Programme Coordination web pages*, 2011. URL `http://lpc.web.cern.ch/lpc`.

[GO09] J. F. Grosse-Oetringhaus. *Measurement of the Charged-Particle Multiplicity in Proton-Proton Collisions with the ALICE Detector*. Ph.D. thesis, Institut für Kernpysik, Universität Münster, Germany, 2009.

[GOR10] J. F. Grosse-Oetringhaus and K. Reygers. *Charged-Particle Multiplicity in Proton-Proton Collisions*. J.Phys.G, **G37** p. 083001, 2010. `arXiv:hep-ex/0912.0023`.

[Gre84] M. B. Green. *Supersymmetrical Dual String Theories and Their Field Theory Limits: A Review*. Surveys High Energ.Phys., **3** pp. 127–160, 1984.

[Gri68] V. Gribov. *A Reggeon Diagram Technique*. Sov.Phys.JETP, **26** pp. 414–422, 1968.

[GU05] A. Giovannini and R. Ugoccioni. *Clan Structure Analysis and QCD Parton Showers in Multiparticle Dynamics: An Intriguing Dialog between Theory and Experiment*. Int.J.Mod.Phys., **A20** pp. 3897–4000, 2005. `arXiv:hep-ph/0405251`.

[GVH86] A. Giovannini and L. Van Hove. *Negative Binomial Multiplicity Distributions in High-Energy Hadron Collisions*. Z.Phys., **C30** p. 391, 1986.

[GVH88] A. Giovannini and L. Van Hove. *Negative Binomial Properties and Clan structure in Multiplicity Distributions*. Acta Phys.Polon., **B19** p. 495, 1988.

[HDG+11] I. Hřivnáčová, O. Datskova, A. Gheata, et al. *The ALICE Geant4 Simulation*. J.Phys.Conf.Ser., **331** p. 032016, 2011.

[Hig64] P. W. Higgs. *Broken Symmetries and the Masses of Gauge Bosons*. Phys.Rev.Lett., **13** pp. 508–509, 1964.

[HKSV03] C. Hill, D. Kuchler, R. Scrivens, et al. *ECR Ion Source and Linac3: Present Status and Future Work*. pp. 70–72, 2003.

[HLO03] M. Harrison, T. Ludlam, and S. Ozaki. *The Relativistic Heavy Ion Collider Project: RHIC and its Detectors*. Nucl.Instrum.Meth., **A499** pp. 235–880, 2003.

[HO85] B. Humpert and R. Odorico. *Multiparton Scattering and QCD Radiation as Sources of Four Jet Events*. Phys.Lett., **B154** p. 211, 1985.

[Hri06] P. Hristov. *The ALICE Offline Bible*. CERN, 2006.

[Hum83] B. Humpert. *Are There Multi-Quark Interactions?* Phys.Lett., **B131** p. 461, 1983.

[Kai82] A. Kaidalov. *The Quark-Gluon Structure of the Pomeron and the Rise of Inclusive Spectra at High-Energies*. Phys.Lett., **B116** p. 459, 1982.

[KNO72] Z. Koba, H. B. Nielsen, and P. Olesen. *Scaling of Multiplicity Distributions in High-Energy Hadron Collisions*. Nucl.Phys., **B40** pp. 317–334, 1972.

[KP10] A. Kaidalov and M. Poghosyan. *Predictions of Quark-Gluon String Model for pp at LHC*. Eur.Phys.J., **C67** pp. 397–404, 2010. arXiv:hep-ph/0910.2050.

[KTBF06] A. Kisiel, T. Taluc, W. Broniowski, et al. *THERMINATOR: THERMal heavy-IoN generATOR*. Comput.Phys.Commun., **174** pp. 669–687, 2006. arXiv:nucl-th/0504047.

[L+00] H. Lai et al. *Global QCD Analysis of Parton Structure of the Nucleon: CTEQ5 Parton Distributions*. Eur.Phys.J., **C12** pp. 375–392, 2000. arXiv:hep-ph/9903282.

[L3 90] L3 Collaboration. *The Construction of the L3 Experiment*. Nucl.Instrum.Meth., **A289** pp. 35–102, 1990.

[LEP84] LEP Collaboration. *LEP Design Report: Vol. 2. The LEP Main Ring*. Tech. Rep. CERN-LEP-84-01, 1984.

[LHC95] LHC Study Group. *The Large Hadron Collider: Conceptual design*. Tech. Rep. CERN-AC-95-05-LHC, 1995.

[LHC08a] LHCb Collaboration. *The LHCb Detector at the LHC*. JINST, **3** p. S08005, 2008.

[LHC08b] LHCf Collaboration. *The LHCf Detector at the CERN Large Hadron Collider*. JINST, **3** p. S08006, 2008.

[LNV+08] I. Legrand, H. Newman, R. Voicu, et al. *MonALISA: A Distributed Service System for Monitoring, Control and Global Optimization*. PoS, **ACAT08** p. 020, 2008.

[LP78] P. Landshoff and J. Polkinghorne. *Calorimeter Triggers for Hard Collisions*. Phys.Rev., **D18** p. 3344, 1978.

[LPS75] P. Landshoff, J. Polkinghorne, and D. Scott. *Production of Baryons with Large Transverse Momentum*. Phys.Rev., **D12** p. 3738, 1975.

[LSB+95] L. Lonnblad, M. H. Seymour, E. Boudinov, et al. *Gamma Gamma Event Generators*, 1995. arXiv:hep-ph/9512371.

[Lüt10] P. Lüttig. *Mittlerer Transversalimpuls in Proton-Proton-Kollisionen bei $\sqrt{s}$ = 900 GeV in ALICE*, 2010. Master thesis, Institut für Kernphysik, Johann Wolfgang Goethe-Universität Frankfurt, Germany, 2010.

[Man89] M. L. Mangano. *Four Jet Production at the Tevatron Collider*. Z.Phys., **C42** p. 331, 1989.

[Mar97] S. P. Martin. *A Supersymmetry Primer*, 1997. arXiv:hep-ph/9709356.

[MoE09] MoEDAL Collaboration. *Technical Design Report of the MoEDAL Experiment*. Tech. Rep. CERN-LHCC-2009-006, 2009.

[MW99] S. G. Matinyan and W. Walker. *Multiplicity Distribution and Mechanisms of the High-Energy Hadron Collisions*. Phys.Rev., **D59** p. 034022, 1999. arXiv:hep-ph/9801219.

[Oll92] J.-Y. Ollitrault. *Anisotropy as a Signature of Transverse Collective Flow.* Phys.Rev., **D46** pp. 229–245, 1992.

[PDG10] PDG - Particle Data Group. *Review of Particle Physics.* J.Phys.G, **G37** p. 075021, 2010.

[Per00] D. Perkins. *Introduction to High-Energy Physics.* Cambridge University Press, Cambridge, 2000.

[PGdC11] S. Porteboeuf and R. Granier de Cassagnac. *J/ψ Yield vs. Multiplicity in Proton-Proton Collisions at the LHC.* Nucl.Phys.Proc.Suppl., **214** pp. 181–184, 2011. arXiv:hep-ex/1012.0719.

[PHO10a] PHOBOS Collaboration. *High Transverse Momentum Triggered Correlations Over a Large Pseudorapidity Acceptance in Au+Au Collisions at $\sqrt{s_{\mathrm{NN}}} = 200\,GeV$.* Phys.Rev.Lett., **104** p. 062301, 2010. arXiv:nucl-ex/0903.2811.

[PHO10b] PHOBOS Collaboration. *System Size Dependence of Cluster Properties from Two-Particle Angular Correlations in Cu+Cu and Au+Au Collisions at $\sqrt{s_{\mathrm{NN}}} = 200\,GeV$.* Phys.Rev., **C81** p. 024904, 2010. arXiv:nucl-ex/0812.1172.

[PHO11] PHOBOS Collaboration. *PHOBOS Results on Charged Particle Multiplicity and Pseudorapidity Distributions in Au+Au, Cu+Cu, d+Au, and p+p Collisions at Ultra-Relativistic Energies.* Phys.Rev., **C83** p. 024913, 2011. arXiv:nucl-ex/1011.1940.

[Pla11] R. Placakyte. *Parton Distribution Functions*, 2011. arXiv:hep-ph/1111.5452.

[PS86] G. Pancheri and Y. Srivastava. *Low-p_{T} Jets and the Rise With Energy of the Inelastic Cross-Section.* Phys.Lett., **B182** pp. 199–207, 1986.

[RER00] S. Roesler, R. Engel, and J. Ranft. *The Monte Carlo Event Generator DPMJET-III.* pp. 1033–1038, 2000. arXiv:hep-ph/0012252.

[Roe96] B. Roe. *Particle Physics at the New Millennium.* Springer, New York, USA, 1996.

[RSR11] S. Rossegger, B. Schnizer, and W. Riegler. *Analytical Solutions for Space Charge Fields in TPC Drift Volumes.* Nucl.Instrum.Meth., **A632** pp. 52–58, 2011.

[SAB+03] P. Saiz, L. Aphecetche, P. Buncic, et al. *AliEn - ALICE Environment on the GRID.* Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, **502**(2-3) pp. 437–440, 2003.

[Sal10] G. P. Salam. *Towards Jetography.* Eur.Phys.J., **C67** pp. 637–686, 2010. arXiv:hep-ph/0906.1833.

[Sch68] W. Schnell. *The CERN Intersecting Storage Rings.* Tech. Rep. CERN-ISR-RF-68-29, 1968.

[Sch82] J. H. Schwarz. *Superstring Theory.* Phys.Rept., **89** pp. 223–322, 1982.

[Sic12] Sicking, E. on behalf of the ALICE Collaboration. *Multiplicity Dependence of Two-Particle Correlations in Proton-Proton Collisions measured with ALICE at the LHC.* Proceedings for the 24$^{\text{th}}$ Rencontres de Blois - Conference for Particle Physics and Cosmology, Blois, France, 2012.

[Ska10] P. Z. Skands. *Tuning Monte Carlo Generators: The Perugia Tunes.* Phys.Rev., **D82** p. 074018, 2010. Update from March 2011 including a set of tunes labeled "Perugia 2011", `arXiv:hep-ph/1005.3457v4`.

[SMS06] T. Sjostrand, S. Mrenna, and P. Z. Skands. *PYTHIA 6.4 Physics and Manual.* JHEP, **0605** p. 026, 2006. `arXiv:hep-ph/0603175`.

[SMS08] T. Sjostrand, S. Mrenna, and P. Z. Skands. *A Brief Introduction to PYTHIA 8.1.* Comput.Phys.Commun., **178** pp. 852–867, 2008. `arXiv:hep-ph/0710.3820`.

[SS04] T. Sjostrand and P. Z. Skands. *Multiple Interactions and the Structure of Beam Remnants.* JHEP, **0403** p. 053, 2004. `arXiv:hep-ph/0402078`.

[STA01] STAR Collaboration. *Elliptic Flow in Au+Au Collisions at $\sqrt{s_{\text{NN}}} = 130\,GeV$.* Phys.Rev.Lett., **86** pp. 402–407, 2001. `arXiv:nucl-ex/0009011`.

[STA09] STAR Collaboration. *Long Range Rapidity Correlations and Jet Production in High Energy Nuclear Collisions.* Phys.Rev., **C80** p. 064912, 2009. `arXiv:nucl-ex/0909.0191`.

[STA10] STAR Collaboration. *Three-Particle Coincidence of the Long Range Pseudorapidity Correlation in High Energy Nucleus-Nucleus Collisions.* Phys.Rev.Lett., **105** p. 022301, 2010. `arXiv:hep-ex/0912.3977`.

[SvZ87] T. Sjostrand and M. van Zijl. *A Multiple Interaction Model for the Event Structure in Hadron Collisions.* Phys.Rev., **D36** p. 2019, 1987.

[TeV82] TeVI Group. *Design Report Tevatron 1 Project.* Tech. Rep. FERMILAB-DESIGN-1982-01, 1982.

[tH74] G. 't Hooft. *A Planar Diagram Theory for Strong Interactions.* Nucl.Phys., **B72** p. 461, 1974.

[TOT08] TOTEM Collaboration. *The TOTEM Experiment at the CERN Large Hadron Collider.* JINST, **3** p. S08007, 2008.

[TOT11] TOTEM Collaboration. *First Measurement of the Total Proton-Proton Cross Section at the LHC Energy of $\sqrt{s} = 7.0\,TeV$.* Europhys.Lett., **96** p. 21002, 2011. `arXiv:hep-ex/1110.1395`.

[UA182a] UA1 Collaboration. *First Observation Of Correlations Between High Transverse Momentum Charged Particles in Events from the CERN $\bar{p}p$ Collider.* Phys.Lett., **B118** pp. 173–177, 1982.

[UA182b] UA1 Collaboration. *Transverse Momentum Spectra for Charged Particles at the CERN Proton anti-Proton Collider.* Phys.Lett., **B118** p. 167, 1982.

[UA183] UA1 Collaboration. *Observation of Jets in High Transverse Energy Events at the CERN $\bar{p}p$ Collider.* Phys.Lett., **B123** p. 115, 1983.

[UA282] UA2 Collaboration. *Observation of Very Large Transverse Momentum Jets at the CERN $\bar{p}p$ Collider.* Phys.Lett., **B118** pp. 203–210, 1982.

[UA284] UA2 Collaboration. *Measurement of Very Large Transverse Momentum Jet Production at the CERN $\bar{p}p$ Collider.* Phys.Lett., **B138** p. 430, 1984.

[UA291] UA2 Collaboration. *A Study of Multi - Jet Events at the CERN $\bar{p}p$ Collider and a Search for Double Parton Scattering.* Phys.Lett., **B268** pp. 145–154, 1991.

[UA585] UA5 Collaboration. *An Investigation of Multiplicity Distributions in Different Pseudorapidity Intervals in $\bar{p}p$ Reactions at a CMS Energy of 540 GeV.* Phys.Lett., **B160** p. 193, 1985.

[UA586] UA5 Collaboration. *Scaling Violations in Multiplicity Distributions at 200 GeV and 900 GeV.* Phys.Lett., **B167** pp. 476–480, 1986.

[UA589] UA5 Collaboration. *Charged Particle Multiplicity Distributions at 200 GeV and 900 GeV Center-Of-Mass Energy.* Z.Phys., **C43** p. 357, 1989.

[Val12] S. Vallero. *Study of the Underlying Event in pp Collisions with the ALICE detector at the LHC.* Ph.D. thesis, Physikalisches Institut, Universität Heidelberg, Germany, 2012.

[vH+11] B. von Haller et al. *The ALICE Data Quality Monitoring System.* J.Phys.Conf.Ser., **331** p. 022030, 2011.

[vL06] M. van Leeuwen. *Experience from High-p_{T} Tracking Analyses in STAR*, 2006. Private communication.

[Wal01] W. Walker. *Multiparton Collisions and Hadron-Nucleon Cross Sections.* Int.J.Mod.Phys., **A16S1A** pp. 184–186, 2001.

[Wal04] W. D. Walker. *Multiparton Interactions and Hadron Structure.* Phys.Rev., **D69** p. 034007, 2004.

[WG92] X.-N. Wang and M. Gyulassy. *A Systematic Study of Particle Production in p + p (anti-p) Collisions via the HIJING Model.* Phys.Rev., **D45** pp. 844–856, 1992.

[Wil77] R. R. Wilson. *The Tevatron.* Tech. Rep. FERMILAB-TM-0763, 1977.

[WKP11] K. Werner, I. Karpenko, and T. Pierog. *Hydrodynamical Evolution in Heavy Ion collisions and pp Scatterings at the LHC: Ridges in AA and pp Scattering.* Acta Phys.Polon.Supp., **4** pp. 629–634, 2011.

[WLP05] K. Werner, F. Liu, and T. Pierog. *Particle Production in Proton-Proton and Deuteron-Gold Collisions at RHIC.* J.Phys.G, **G31** pp. S985–S988, 2005. arXiv:hep-ph/0411329.

[YHM05] K. Yagi, T. Hatsuda, and Y. Miake. *Quark-Gluon Plasma: From Big Bang to Little Bang.* Camb.Monogr.Part.Phys.Nucl.Phys.Cosmol., **23** pp. 1–446, 2005.

Acknowledgements

I would like to thank Johannes P. Wessels for giving me the opportunity to perform my thesis work at CERN. For supporting my application at CERN, I thank Federico Carminati. Many thanks to Andreas Morsch for the interesting and diverse work I could perform in his team as well as for very fruitful discussions. Special thanks to Ivana Hřivnáčová for the pleasant collaboration during the work on the Geant4 simulations.
For the enjoyable teamwork and the nice working atmosphere, I would like to thank all members of the ALICE Offline group, especially Cinzia Luzzi, and the ALICE Quality Assurance task force.
For countless discussions and suggestions, I would like to thank Zaida Conesa Del Valle, Sara Vallero, Antonio Ortiz Velasquez, Davide Caffarri, Jan Fiete Große-Oetringhaus, Francesco Prino, Martin Poghosyan, Ruben Shahoyan, Jürgen Schukraft, Filip Krizek, Peter Skands, Arsen Hayrapetyan, Martin Vala, Adam Kisiel, Marco van Leeuwen, Christian Klein-Bösing, Steffen Schreiner, Andrea Dainese, Federico Carminati, Latchezar Betev, Olga Datskova, Andrei Gheata, Mihaela Gheata, Peter Hristov, Cvetan Cheshkov, Costin Grigoras, Alina Grigoras, Marian Ivanov, and Jacek Otwinowski.
Many thanks to Regina Kwee, Jakob Blomer, Christian Klein-Bösing, Stefan Flörchinger, Lee Barnby, Dhevan Gangadharan, Andrew Adare, Martin Poghosyan, and Willi Sicking for giving me corrections and suggestions during my thesis work.
Many thanks to my friends for the nice time we spent together in the last three years. For the emotional support I would like to thank Jakob and my family, Edith, Willi, Kerstin, and Annegret.