Dissertation
submitted to the
Combined Faculties for Natural Sciences and Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Put forward by
Marten Ole Schmidt (M.Sc. Physics)
Born in Rendsburg, Germany
Oral examination on July 15$^{\text{th}}$, 2020

Space point calibration of the ALICE TPC with track residuals

Referees: Prof. Dr. Klaus Reygers
Prof. Dr. Hans-Christian Schultz-Coulon

Abstract

In the upcoming LHC Run 3 the upgraded Time Projection Chamber (TPC) of the ALICE experiment will record Pb–Pb collisions in a continuous readout mode at an interaction rate up to 50 kHz. These conditions will lead to the accumulation of space charge in the detector volume which in turn induces distortions of the electron drift lines of several centimeters that fluctuate in time. This work describes the correction of these distortions via a calibration procedure that uses the information of the Inner Tracking System (ITS), which is located inside, and the Transition Radiation Detector (TRD) and the Time-Of-Flight system (TOF), located around the TPC, respectively.

The required online tracking algorithm for the TRD, which is based on a Kalman filter, is the main result of this work. The procedure matches extrapolated ITS-TPC tracks to TRD space points utilizing GPUs. The new online tracking algorithm has a performance comparable to the one of the offline tracking algorithm used in the Run 1 and 2 for tracks with transverse momenta above 1.5 GeV/c, while it fulfills the computing speed requirements for Run 3. The second part of this work describes the extraction of time-averaged TPC cluster residuals with respect to interpolated ITS-TRD-TOF tracks, in order to create a map of space-charge distortions. Regular updates of the correction map compensate for changes in the TPC conditions. The map is applied in the final reconstruction of the data.

Kurzfassung

Im bevorstehenden Run 3 des LHC wird die aufgerüstete Zeitprojektionskammer (TPC) des ALICE Experimentes Blei-Blei-Kollisionen mit einer Rate von bis zu 50 kHz in einem kontinuierlichen Auslesemodus aufzeichnen. Unter diesen Bedingungen sammelt sich Raumladung im Detektorvolumen an, die Verzerrungen der Driftspuren der Elektronen von mehreren Zentimetern bedingt. Diese Arbeit verwendet die Informationen der umliegenden Detektoren, um die auftretenden Verzerrungen zu korrigieren. Dieses sind auf der inneren Seite der TPC das Innere Spurrekonstruktionssystem (ITS) und auf der äußeren Seite der Übergangsstrahlungsdetektor (TRD) und das Flugzeitmessungssystem (TOF).

Der in dieser Arbeit entwickelte online Algorithmus zur Spurrekonstruktion im TRD basiert auf dem Kalman Filter und kann auf Grafikkarten ausgeführt werden. Teilchenspuren rekonstruiert in ITS und TPC werden in den TRD extrapoliert und mit den dort gemessenen Raumpunkten verknüpft. Die Ergebnisse des Algorithmus sind vergleichbar mit der offline Rekonstruktion in Run 1 und 2 für Teilchenspuren mit einem Tansversalimpuls oberhalb von 1.5 GeV/c. Gleichzeitig erfüllt der Algorithmus die Anforderungen an die Rechengeschwindigkeit für Run 3. Der zweite Teil dieser Arbeit beschreibt die Berechnung der zeitlich gemittelten Abstände zwischen gemessenen Positionen innerhalb der TPC und interpolierten Positionen zwischen den umliegenden Detektoren ITS, TRD und TOF. Basierend auf den berechneten Abständen wird in regelmäßigen Intervallen eine Karte von Raumladungsverzerrungen erstellt. Diese Karte wird in der finalen Datenrekonstruktion angewendet.

Contents

1. Introduction **1**

2. Upgrade of the ALICE experiment **5**
2.1. Physics motivation . . . 5
2.2. Detectors and upgrades . . . 6
2.2.1. Inner Tracking System . . . 8
2.2.2. Time Projection Chamber . . . 9
2.2.3. Transition Radiation Detector . . . 11
2.2.4. Time-of-Flight Detector . . . 12
2.2.5. Other upgrades . . . 12
2.3. New online-offline computing system . . . 13
2.3.1. Data flow in Run 3 and 4 . . . 14
2.3.2. Time frame building . . . 16
2.3.3. Synchronous and asynchronous data processing . . . 17

3. Space-charge distortions in the TPC **19**
3.1. Observations in Run 2 . . . 19
3.2. Expectations for Run 3 . . . 23
3.3. Calibration procedure in Run 2 . . . 24
3.4. Performance . . . 30

4. Global track reconstruction **33**
4.1. Current offline reconstruction . . . 33
4.2. HLT reconstruction . . . 34
4.3. Reconstruction in Run 3 . . . 35
4.3.1. Conditions in Run 3 . . . 35
4.3.2. Synchronous processing of time frames . . . 36
4.3.3. Asynchronous processing . . . 38
4.4. Mathematical basics . . . 38
4.4.1. The ALICE track model . . . 39
4.4.2. Track covariance matrix . . . 42
4.4.3. Kalman filter . . . 43

5. Online tracking with the TRD **47**
5.1. Detector overview . . . 47
5.2. Processing in the Front-End Electronics . . . 49
5.2.1. The online tracklets . . . 49
5.2.2. Comparison to offline tracklets . . . 55

5.3. Tracking algorithm . . . 56
5.3.1. Conversion of tracklets to space points . . . 57
5.3.2. Matching procedure . . . 58
5.3.3. Tracklet covariance . . . 61
5.3.4. Configuration and quality assurance . . . 65
5.3.5. Comparing Monte Carlo and real data . . . 68
5.4. Performance . . . 69
5.4.1. Efficiency and purity . . . 69
5.4.2. Computing speed . . . 71
5.5. Development and implementation details . . . 72
5.5.1. Prototyping in the High-Level-Trigger . . . 72
5.5.2. Implementation in **O²** . . . 74
5.5.3. Running on GPUs . . . 75

6. TPC calibration with track residuals **77**
6.1. Collection of TPC cluster residuals . . . 78
6.2. Cleaning of the input data . . . 80
6.3. Conversion to local residuals . . . 84
6.4. Determination of the residuals per voxel . . . 84
6.4.1. Voxel validation . . . 87
6.4.2. Smoothing the residuals . . . 87
6.5. Voxel dispersion . . . 90
6.6. Comparison to AliRoot version . . . 91
6.7. Discussion . . . 93

7. Summary and outlook **95**

A. TRD tracklet parameterizations **99**

Bibliography **101**

1. Introduction

Since the beginning of the last century, particle detectors have played an increasingly important role not only in science, but in everyday life as well. In their famous gold foil experiment in 1909 Geiger and Marsden used a scintillating screen made out of zinc sulphide to detect α-particles that had passed through a thin gold foil [1]. The flashes induced by the α-particles had to be registered *by eye* by the experimenters. From the angular distribution of the scattered α-particles, Rutherford deduced that atoms consist of a small positively charged core surrounded by a large cloud of electrons [2].

Today, various particle detector technologies exist for the most diverse applications, from gaseous detectors like the Geiger-Müller tube to image sensors for example in cameras or computed tomography scanners. The signals are almost always digitized as early as possible and further on processed by computers. Depending on their field of application the detectors are optimized for a given rate, high resolution, radiation hardness, low cost or a combination of the above.

A major driving factor for innovation in particle detectors is the great development of particle accelerators. They provide collisions which can be studied under well-controlled conditions. One of the first accelerator experiments was performed by Cockcroft and Walton in 1932 [3]. Protons were accelerated to 500 keV using a DC generator and shot onto a Lithium target in order to study disintegration processes [4]. In their experimental setup the uncertainty in the composition of the beam prevented Cockcroft and Walton to make quantitative statements about the absolute number of disintegrations they observed. Significant progress has been made since then and by now accelerators are not only used in particle or nuclear physics experiments, but amongst others also for medical applications as for example radiotherapy or the production of short lived radionuclides which are needed for diagnostics in nuclear medicine [5].

The highest energy achieved in accelerators today is 6.5 TeV for beams of protons at the Large Hadron Collider (LHC) at CERN. In addition to the energy the luminosity is an important characteristic for an accelerator because it determines the collision rate. The LHC also holds the record in instantaneous luminosity for proton-proton (pp) collisions with $2.06 \times 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ [6] which is more than twice its design value [7]. A higher luminosity enables detailed measurements also for rare processes, but it places great demands on the detectors as well, especially in terms of their speed. In a collider the luminosity $\mathscr{L}$ is defined as

$$\mathscr{L} = \frac{n\ N^2\ f}{4\pi\sigma_x\sigma_y} \tag{1.1}$$

with N describing the number of particles per bunch, n the number of bunches and f the revolution frequency. The denominator corresponds to the cross section of the beams assuming a gaussian beam profile with widths σ_x and σ_y.

Currently the LHC is shut down for its second major upgrade campaign and maintenance work, called Long Shutdown 2 (LS2), until spring 2021. The main part of the upgrade program of the LHC itself is the renewal of the injector chain in view of the High-Luminosity LHC (HL-LHC). The HL-LHC is scheduled to start only after the third shut down period LS3 in 2026 with the luminosity increasing by a factor 5 to 7 [8]. At the same time the four main LHC experiments are undergoing upgrades to exploit the full potential of the accelerator. While the two general purpose experiments A Toroidal LHC ApparatuS [9] (ATLAS) and the Compact Muon Solenoid [10] (CMS) are going to replace major parts of their detectors only in LS3 in view of the HL-LHC, the two more specialized experiments LHC beauty [11] (LHCb) and A Large Ion Collider Experiment [12] (ALICE) are now undergoing major upgrade programs.

ALICE [13] is the dedicated heavy-ion experiment at the LHC, designed to study physics of strongly interacting matter at extreme energy densities and temperatures. The highest energy densities and temperatures in the laboratory are reached in ultra relativistic heavy-ion collisions. In dedicated running periods the LHC provides besides pp also lead lead (Pb–Pb) and proton lead (p–Pb) collisions which are of special interest to ALICE. After LS2 the luminosity will increase in Pb–Pb collisions as well. During the last data taking period, commonly denoted as Run 2, the luminosity in Pb–Pb for ALICE was leveled to about $1 \times 10^{27}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, which is the LHC design value. For the upcoming third data taking period (Run 3) the luminosity will increase by a factor of about six. Already in 2018 the LHC delivered a peak luminosity of about $6 \times 10^{27}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ to ATLAS and CMS, but it decreased during the course of the fill due to burn off (the loss of nuclei in collisions). In contrast the luminosity for ALICE will be mostly leveled to that value in the future. Therefore the integrated luminosity per fill will reach a new record level. The increase in luminosity translates to an increase in interaction rate for ALICE from currently 8 kHz to 50 kHz for Pb–Pb. This implies manifold changes in the requirements for the experiment in comparison to the previous LHC running periods.

The present work describes a novel calibration technique for the ALICE Time Projection Chamber (TPC), which is essential to maintain the current reconstruction performance also after the upgrade. A major challenge arising from operating the TPC at such high rates is the accumulation of space charge inside the detector volume. In the amplification region a large amount of ions is produced which partially drift back into the active region distorting the drifting electrons from subsequent collisions. The space-charge distortions can reach up to 20 cm in radial direction [14] and have to be corrected to obtain the intrinsic track resolution of the TPC of a few hundred µm.

Outline of this thesis

The following Chapter 2 discusses the physics objectives that led to the ALICE upgrade project and gives a short overview of the detector modifications and the new computing framework. In Chapter 3 the calibration procedure to mitigate the effect of space-charge distortions in the TPC for Run 2 will be described. It is the basis for the future calibration procedure which is the main topic of this thesis. A major part of this procedure is the implementation of a new tracking algorithm for the Transition Radiation Detector (TRD)

which is presented in Chapter 5 after the global reconstruction strategy for ALICE is introduced in Chapter 4. Subsequently the new calibration procedure and its performance are presented in Chapter 6. Summary and outlook constitute Chapter 7.

2. Upgrade of the ALICE experiment

2.1. Physics motivation

The ALICE experiment is designed to carry out comprehensive studies of Quantum ChromoDynamics (QCD), the gauge theory of the strong interaction. At very high temperatures and densities lattice QCD calculations predict a new state of matter, called Quark-Gluon Plasma (QGP), where quarks and gluons are no longer bound in color neutral objects and whose properties are very different from ordinary hadronic matter [15]. In the early universe, about 10 µs after the big bang, the conditions allowed for the existence of a QGP [16]. However, astronomical observations from this phase of the universe are not available. The earliest directly detected signal, the cosmic microwave background [17], dates back to a few hundred thousand years after the big bang. Therefore, the only possibility to study the QGP is its recreation in the laboratory. ALICE records high energy nucleus-nucleus collisions provided by the LHC, where both energy and density are sufficient for the creation of a QGP.

A direct observation of the QGP is not feasible, because of its short lifetime of about 10 fm/c. Instead, appropriate probes shall reveal its properties such as temperature, equation of state or viscosity and demonstrate the involved phase transitions experimentally.

After LS2 ALICE will focus on rare probes. Two examples for such probes are briefly introduced here to motivate the work carried out in this thesis.

Low-mass dilepton measurements allow one to determine the space-time evolution of the QGP, because electromagnetic radiation is produced during all stages of the collision and because the leptons do not interact strongly with the medium. Spectra of the invariant mass and the transverse momentum (p_{T}) of low-mass dileptons provide information on the temperature of the system. They also complement the measurement of direct photons for transverse momenta above 1 GeV/c, if dilepton masses down to about 200 MeV/c^2 are detected [15]. The measurement of electron-positron pairs in the corresponding p_{T} range is experimentally challenging, because of the large background from decays involving two leptons in the final state, for example π^0 Dalitz decays, or photon conversions in the detector material. It requires excellent particle identification capabilities and precise distance-of-closest-approach estimates to first identify the electrons and second distinguish between electrons originating from the primary vertex or from secondary vertices or conversions. The large combinatorial background prevents the use of online triggers and requires to collect a large data sample with minimum bias, because of the low production rate of thermal dileptons, which is suppressed by a factor $\propto \alpha_{\mathrm{em}}^2$ [15]. Especially the new inner tracker [18] with very low material budget (see Section 2.2.1)

and special data taking periods with a reduced magnetic field enable these measurements which were not possible with the original detector setup [15].

Heavy-flavor production, i.e. the production of charm and beauty quarks, also represent an important probe for the parameters of the strongly interacting matter. For example the thermalization of heavy quarks in the QGP can be tested by measuring the elliptic flow (v_2), a non-uniform variation in azimuthal direction of energy and momentum of the particles produced in the collisions [19, 20]. ALICE has measured the elliptic flow for different D mesons in non-central Pb–Pb collisions in the p_{T} range of 2 to 6 GeV/c that indicate their participation in the collective flow [21]. Especially at low transverse momenta around 1 GeV/c a large v_2 is predicted by theoretical models, a p_{T} range which becomes accessible only with the upgraded ALICE detectors [15]. Furthermore, the measurement of the nuclear modification factor R_{AA} with improved precision for heavy particles especially at low transverse momentum will permit to quantify the mass dependence of energy loss in the medium and thus allow to discriminate between energy loss models [22]. The R_{AA} is defined as the ratio of particle production in nucleus–nucleus collisions to particle production in pp collisions scaled by the average number of colliding nucleons. Therefore the measurement of heavy-flavor production needs to be performed in pp collisions by ALICE as well.

Low-mass dileptons and heavy-flavor production are only two examples for the motivation of the upgrade of the ALICE detector. They are mentioned here, because they will both strongly benefit from the improved tracking and vertexing capabilities and from the increased data sample of minimum bias events. Traditional online trigger techniques are not applicable to these complex probes. Hence, ALICE will profit significantly from the increased luminosity delivered by the LHC after LS2, by collecting as many events as possible with minimum bias.

2.2. Detectors and upgrades

The ALICE experiment [24, 25] is located at interaction point 2 of the LHC. Overall, it is designed as a multi-purpose detector for physics at midrapidity and the forward region. Its detector setup is shown in Figure 2.1. The central barrel of ALICE is enclosed by the large magnet inherited from the L3 experiment at LEP which provides a solenoidal magnetic field of 0.5 T along the direction of the beam. The detectors inside the central barrel cover at least the pseudorapidity range $|\eta| \lesssim 0.9$. In forward direction a muon spectrometer behind an absorber covers the pseudorapidity range of $-4 < \eta < -2.5$. It consists of a warm dipole magnet with an integrated horizontal field of 3 T · m, six tracking stations and additional trigger stations behind further shielding. The global coordinate system is a right-handed Cartesian system with its origin at the nominal interaction point. The z-axis is aligned with the beam and pointing away from the muon arm, the y-axis pointing upwards and the x-axis points towards the center of the LHC

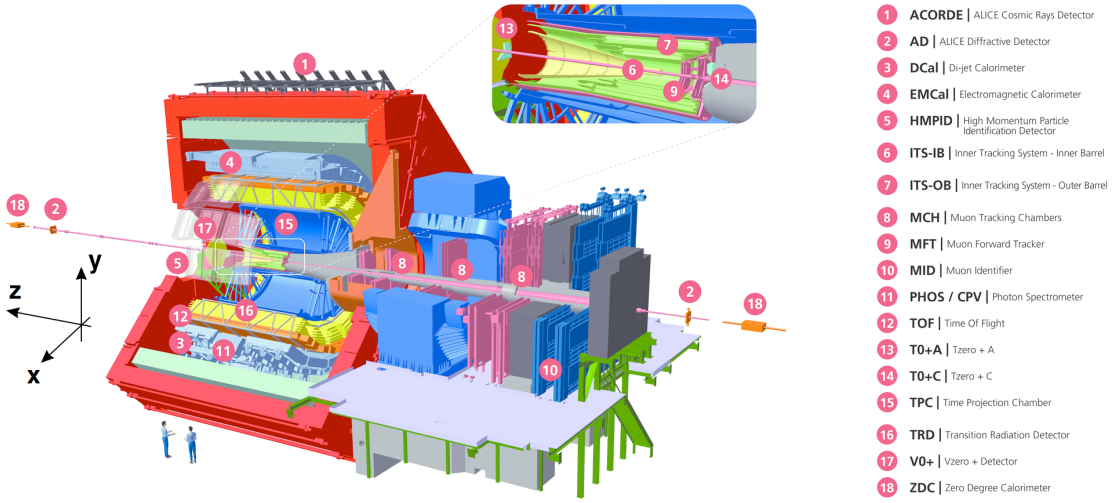


Figure 2.1.: Overview of the upgraded ALICE apparatus. The central barrel is enclosed by a large solenoidal magnet (red). A muon spectrometer is installed in forward direction. The global coordinate system is indicated on the left. Adapted from [23].

ring. In the following, first the detector setup during Run 1 and 2 will be described and subsequently the relevant upgrades for Run 3.

The silicon-based Inner Tracking System (ITS) is installed closest to the interaction region. It consists of two layers respectively of pixel, drift and strip detectors [26]. The ITS is surrounded by the large Time Projection Chamber (TPC), the main tracking system of ALICE [27]. It is a hollow cylinder with an inner and outer radius of 0.85 m and 2.47 m and a length of 5 m. The central high-voltage electrode is at a potential of 100 kV and splits the active volume into two drift regions. The drift time for electrons from the central electrode to the end plates is about 95 µs at the given electrical field of 400 V/cm. The TPC employs Multi-Wire Proportional Chambers (MWPCs) with cathode pad readout. The azimuthal readout plane is segmented into 18 trapezoidal sectors, each covering 20°.

Next to the TPC, following its sector segmentation in the azimuthal plane, the Transition Radiation Detector (TRD) [28] is installed. It contributes to the electron identification capabilities of ALICE, the global tracking and during Runs 1 and 2 it provided various triggers for high transverse momentum processes as well as for leptonic decays and light nuclei [29]. The TRD is described in more detail in Chapter 5 where its new tracking algorithm is presented. The TRD is enclosed by the Time-Of-Flight System (TOF) at a radius of 3.7 m which provides a timing resolution better than 100 ps at an occupancy of 10% employing multi-gap resistive plate chambers [30]. Time of flight measurements allow the mass of high-energy particles to be determined, if their momentum is known and their speed is significantly below the speed of light. With a pad size of 3.5 x 2.5 cm^2 the TOF also contributes to the central barrel tracking with space points at large radii.

The central barrel acceptance is equipped partly with three different calorimeters.

The ElectroMagnetic CALorimeter (EMCAL) provides fast triggers on high-energy jets, measurements of jet energy, high momentum photons, neutral pions and electrons [31]. Together with the Di-Jet Calorimeter (DCal) EMCal forms a two-arm electromagnetic calorimeter [32]. The PHOton Spectrometer (PHOS) composed of $PbWO_4$ crystals is designed to measure 4-momenta of photons with high resolution [33].

A High Momentum Particle IDentification (HMPID) detector identifies charged particles employing seven ring imaging Cherenkov counters [34].

Several smaller detectors are used for event characterization and to provide fast interaction triggers, namely the T0, V0 and the Zero Degree Calorimeter (ZDC) which is installed $\sim 115\,\mathrm{m}$ away on both sides from the interaction region close to the beam pipe [35].

Upgrades

In order to exploit the full potential of the LHC ALICE is undergoing a major upgrade program. The interaction rate for Pb–Pb collisions is going to increase by a factor of about 6. The high multiplicities and hence very low signal-over-background ratio for rare probes especially at low transverse momenta render a dedicated high-level trigger unrealistic. Hence, the overall strategy of ALICE is to record as many events as possible with minimum bias and to compress the data as much as possible before it is written to tape permanently. This is achieved by reading out the detectors in a continuous mode instead of a triggered one. The combination of continuous readout and higher luminosity translate to a total increase of a factor 100 for the number of accumulated events in a given time interval.

The following subsections outline the upgrades of the detectors relevant for this thesis. The working principle for the TRD is explained in detail in Chapter 5.

2.2.1. Inner Tracking System

The ITS will be completely replaced by a new 7 layer silicon detector employing Monolithic Active Pixel Sensors (MAPS) technology [18]. Three layers are arranged in the inner barrel at radii 2.3 cm, 3.2 cm and 3.9 cm from the interaction point. The other four layers are grouped together in the outer barrel. They are at radii between 19.6 and 39.3 cm. The pseudorapidity coverage is $|\eta| \leq 2.0$ for the inner barrel and $|\eta| \leq 1.3$ for the outer barrel for tracks originating from the nominal interaction point at $z = 0$. Compared to the old ITS the innermost layer is much closer to the luminous region. Furthermore, the material budget is significantly reduced which in turn reduces multiple scattering effects. These changes improve the impact parameter resolution by a factor of about 3. This enables for example the measurement of the charm baryon Λ_c with its decay length ($c\tau$) of only 60 µm. With the old ITS the resolution was not sufficient to distinguish between secondary vertices so close to the interaction point in the relevant transverse momentum region between 1 and 2 GeV/c. Another drawback of the old ITS was the limitation of its readout rate to 1 kHz. The new ITS will be readout continuously and is designed for an interaction rate of 400 kHz (100 kHz) for pp (Pb–Pb) collisions. Due to

Characteristics	ITS 1	ITS 2
# of layers	6	7
technology	pixel (2), drift (2), strip (2)	pixel (7)
r of innermost layer	4 cm	2.3 cm
min. pixel size	50 µm($r\varphi$) × 300 µm(z)	30 µm × 30 µm
material budget (all layers combined)	6.5% X_0	1.1% X_0
max. readout rate	1 kHz	unlimited (continuous)

Table 2.1.: Comparison of the old ITS 1 with the upgraded ITS 2.

the integration time of about 30 µs higher interaction rates lead to difficulties to assign the detector signals to the correct interaction vertex. More details on the reconstruction with the new ITS will be given in Section 4.3.

The new ITS will not measure the energy loss of the traversing particles, but employ a binary readout instead. Possibilities to use the size of the charge clusters, which are reconstructed from the detector signal, for the identification of light hyper-nuclei are under investigation [18]. Table 2.1 compares the most relevant parameters for the old and the new ITS.

2.2.2. Time Projection Chamber

In order to motivate the upgrade of the ALICE TPC, its working principle is briefly introduced followed by an explanation of the upgrade.

Working principle

Charged particles traversing the TPC leave a trail of ionization inside the detector gas. The TPC is operated with neon or argon as a counting gas and CO_2 is added as a quencher. The gas mixture was changed several times during Run 1 and 2 and is discussed in more detail in Section 3.1. The geometry of the TPC is shown in Figure 2.2. Due to the electrical drift field parallel to the beam axis the electrons move towards the readout chambers located at the end caps while the positive ions move towards the central electrode where they are absorbed.

The central electrode splits the TPC in two drift volumes. They are referred to as A-side (closer to ATLAS) and C-side (closer to CMS), respectively. The readout chambers for each of the $2 \cdot 18$ sectors are divided into an Inner Readout Chamber (IROC) and an Outer Readout Chamber (OROC). Here the electrons are amplified and the induced signal is sampled at a high frequency in the front-end electronics. The readout pads have a varying size of about 0.75 to 1.5 cm in row direction (corresponding to the radial direction r) and 0.4 to 0.6 cm in column direction (corresponding to $r\varphi$). For each pad row *clusters* within the deposited charge are reconstructed in two dimensions. In the native TPC coordinates these are pad column number and time bin with uncertainties based on the cluster shape. The time t is measured with respect to the collision time.

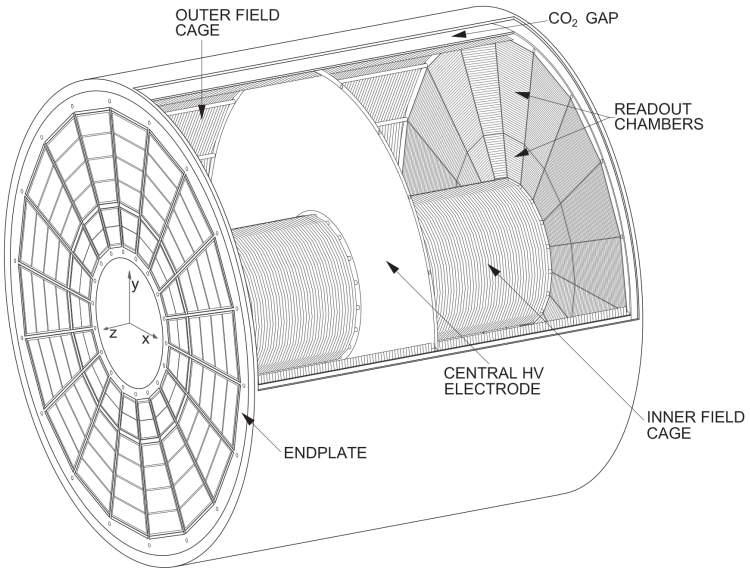


Figure 2.2.: Schematic drawing of the ALICE TPC, taken from [14]. The lines at the readout chambers indicate the pad-rows. The segmentation into IROCs and OROCs is visible.

With the known electron drift velocity v_d it can be converted into the cluster z position via

$$z_{\text{cls}} = \pm z_{\text{max}} \mp v_d \cdot t, \tag{2.1}$$

with the maximum drift length $z_{\text{max}} = 250\,\text{cm}$. Clusters reconstructed on the A-side have a positive z position. In a continuous readout mode the time of the corresponding collision for a cluster is not a priori known. The changes in the reconstruction algorithm will be explained in Section 4.3.

Taken together, for an ideal track this yields a 2D space point at each pad row. The pad row in turn yields the third dimension. In addition, Particle IDentification (PID) information is available based on the deposited charge. This is not discussed further as it is irrelevant for the correction of the space-charge distortions described in this work.

Upgrade

The data taking rate of the TPC was limited in Run 1 and 2 by two factors. The restricted bandwidth from the front-end electronics to the Data AcQuisition (DAQ) on the one hand lead to an upper limit of $\mathcal{O}(500\,\text{Hz})$ for recording minimum bias Pb–Pb interactions [15]. On the other hand, even without bandwidth constraints the readout rate of the TPC would be limited to about 3.5 kHz, far below the target rate of 50 kHz in Run 3. The reason for this is the following: The MWPC-based readout chambers have to be shielded from the drift region with a gating grid to prevent ions created in the avalanche process around the anode wires from drifting back into the active region. Since the drift velocity of ions in the TPC is about a factor 5000 slower than the drift velocity of electrons [24], the gating grid needs to be closed for about 180 µs after it

opened for a full drift time to absorb the created ions. More details are given in the following Chapter 3.

The upgraded TPC employs Gas Electron Multiplier (GEM) foils for gas amplification [14]. In contrast to MWPCs, most of the ions created in the avalanche process inside the holes of the GEM foils end up on the foil itself and do not drift back into the active detector region. The ion backflow, defined as the ratio of cathode to anode current, can be as low as 0.25% for GEM stuctures [36]. The ALICE TPC in Run 3 will have an ion backflow of about 1% (more details are given in Section 3.2). Note that the ion backflow incorporates the contribution of ions created in the avalanche process. Therefore, a continuous operation without a gating grid is possible. At the time ALICE was designed GEM technology was not ready to be adopted for such a large detector project. But significant R&D efforts for future high-rate drift detectors for example at the International Linear Collider (ILD) [37] or at the Facility for Antiproton and Ion Research (FAIR) [38] have pushed the GEM technology to a state where it is feasible to be employed in ALICE. In an extensive optimization process a stack of four GEM foils with different hole sizes and pitches was found to provide the best performance in terms of effective gas gain, ion back-flow and stability [14]. While the readout rate of the upgraded TPC will significantly increase, the energy resolution will be preserved by careful tuning of the individual GEM voltage settings.

The continuous operation obviously has strong implications on the data rate and the reconstruction algorithm. Both will be discussed in more detail in the dedicated Sections 2.3.3 and 4.3, respectively. The space-charge distortions and the calibration procedure are explained in the following Chapter 3.

2.2.3. Transition Radiation Detector

During Run 2 the readout rate of the Transition Radiation Detector (TRD) in Pb–Pb collisions was limited to a few kHz [39]. An increase in readout rate can be achieved by reducing the data volume shipped from the detector to the DAQ system. Instead of transmitting full zero-suppressed ADC data the TRD will transmit only tracklets, which are short track segments assembled in the detector Front-End Electronics (FEE). The processing in the FEE is described in Section 5.2. As the FEE does not support multi-event buffers, events occurring while the detector is busy will be lost. For an interaction rate $R_{\mathrm{int}} = 50\,\mathrm{kHz}$ the readout efficiency can be estimated assuming a busy time of $t_{\mathrm{busy}} = 6\,\mu\mathrm{s}$ and a Poisson distribution for the events:

$$P(k) = \frac{\lambda^k e^{-\lambda}}{k!}. \tag{2.2}$$

The average number of events which occur during the dead time is $\lambda = t_{\mathrm{busy}} R_{\mathrm{int}} = 0.3$. The readout efficiency is given by

$$\epsilon = \frac{\sum_{k=0}^{\infty} P(k)}{\sum_{k=0}^{\infty} (k+1) P(k)} = \frac{1}{1+\lambda} = 0.77. \tag{2.3}$$

Reaching higher readout rates would require the complete disassembly of the TRD, replacement of the FEE mounted on top of each chamber and subsequently the reassembly

and reinstallation of the detector. With the limited amount of time and resources available during LS2 this is not feasible, but also not necessary. It has been shown that the impact of having information from the TRD only for a fraction of all events on analysis is very limited [39]. Furthermore for the space point calibration of the TPC a subset of all events is sufficient as it will be shown in Chapter 6.

In addition to a change in the readout scheme a large repair campaign was conducted which was originally not foreseen for the long shutdown: 9 out of 18 sectors were extracted to repair the HV circuits of several chambers. These could not be operated at nominal conditions due to broken capacitors. Although all used capacitors were thoroughly tested before they were installed, some of the 4.7 nF and 2.2 nF capacitors of the circuits were failing presumably due to aging affects. They were removed to allow again for nominal operation of the affected chambers. In order to extract the capacitors the sectors did not need to be dismantled, because a minimally invasive procedure was found where the sectors were opened just at the location of the HV filter board via milling.

As the capacitors broke at random times (the limited statistics only allowed for a rather imprecise estimate of their lifetime) and they cannot be removed in short technical stops, all of the 4.7 nF and 2.2 nF capacitors were removed for the extracted sectors. The drawback of the operation without the capacitors is a coupling of all the readout pads of one chamber leading to a common mode effect. A signal generated on one readout pad leads to the inverse signal integrated over all other pads. This in turn leads to a shift of the baseline estimated to be between 1 and 2 ADC counts for events with the highest multiplicity [40]. For the tracking this effect will be negligible, but for the optimal $\mathrm{d}E/\mathrm{d}x$ measurement it needs to be corrected for in the reconstruction. The TRD rework was concluded in December 2019.

2.2.4. Time-of-Flight Detector

Based on extrapolations from previous data taking periods rates of $60\,\mathrm{Hz\,cm^{-2}}$ are expected for the TOF detector at 3.7 m from the interaction point in Pb–Pb collisions after LS2. Test beam results indicate that the detector can handle these rates without loss in performance [39]. The target readout rate of 200 kHz can be reached without modifications to the present hardware. Hence, TOF will not receive major modifications and will remain, as the TRD, to be readout in a triggered mode. A continuous readout is not feasible both for hardware and budget reasons.

2.2.5. Other upgrades

Other upgrades include:

- a new Beryllium beam pipe with a smaller radius (29 mm $\mapsto$ 17.2 mm) [15],
- a Fast Interaction Trigger (FIT) detector which combines Cherenkov arrays (T0+) and scintillator rings (V0+) to provide a fast minimum bias trigger for detectors with triggered readout and information on the geometry of the collision [39],

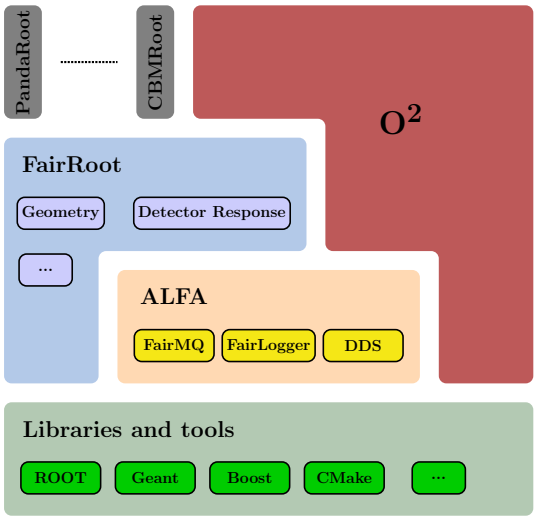


Figure 2.3.: Overview of the O^2 software ecosystem. Adapted from [42].

- new readout for all detectors except EMCal, PHOS and HMPID via the Common Readout Unit (CRU, see Section 2.3.1) [39],
- and a Muon Forward Tracker (MFT) installed in front of the absorber to match muons found in the forward arm [41]. It utilizes the same silicon chips as the new ITS.

More details can be found in the ALICE upgrade letter of intent [15] and in the respective referenced upgrade technical design reports.

2.3. New online-offline computing system

The major upgrade of the ALICE experiment described above requires fundamental changes in the data processing approach. The software developed for the online and offline systems during Run 1 and Run 2 cannot be adopted to meet the requirements. Instead, a completely new software framework is being developed. It will combine the functionalities of the DAQ, the High-Level-Trigger (HLT) and the offline systems. Because it combines the previously separated online and offline frameworks it is named O^2.

The greatest challenges will be to handle the increased data volume and rate and to provide the computing power to process these data. But also coping with space-charge distortions of the order of 15 cm for space points in the TPC to reach a target precision of $\sim 500\,\mu m$ poses a major computing challenge. Already in the last data taking period space-charge distortions were affecting the TPC as described in Chapter 3. A calibration method to mitigate the effect was developed and is the basis for the calibration presented in Chapter 6.

The O^2 software framework depicted in Figure 2.3 is not only based on general libraries and tools such as ROOT [43], CMake [44] and Boost [45], but also on two other frameworks called ALFA and FairRoot. ALFA is a common project of ALICE and FAIR. Its main features are FairMQ, providing a message-queue based data transport over different protocols, DDS, a dynamic deployment system allowing a dynamic distribution of different user processes on any resource management system or a laptop, and FairLogger, a lightweight and fast logging library. FairRoot provides base classes for geometry, detector response and parameter database.

The basis of O^2 is the Data Processing Layer (DPL). DPL devices can be specified with particular input and output and they wrap certain algorithms. Workflows are subsequently automatically created for a specified group of DPL devices by matching of their inputs and outputs. Thus, to create a workflow for a local reconstruction of one detector one would need to specify only the required devices, for example a digit reader, a clusterizer, a tracker and a file writer in arbitrary order. The framework takes care of the workflow topology, data transfer and load balancing without any further intervention required from the user. The general concept of the DPL is described in [46].

The high flexibility of the software framework allows the utilization of heterogeneous computing resources such as compute nodes equipped with hardware accelerators, for example Graphics Processing Units (GPUs). The TRD tracking for instance can run both on CPUs and GPUs as described in Section 5.5.3.

2.3.1. Data flow in Run 3 and 4

The new data flow model for ALICE is depicted in Figure 2.4. A little more than 20k new optical fibers were pulled to transmit 3.5 TB/s raw data from the detectors in the cavern to Counting Room 1 (CR1). CR1 houses the First Level Processors (FLP), around 200 dual-socket Dell R740 servers equipped with up to three Common Readout Units (CRUs) each. The CRU provides the interface between the front-end electronics of the detectors, the Central Trigger Processor (CTP), which distributes the trigger and timing information, the Detector Control System (DCS) and the computing farm [47].

The CRU was initially developed by the LHCb collaboration under the name PCIe40 [48]. It is equipped with 48 bidirectional optical links and an Intel Arria10 FPGA, which offers computing power for low-level processing of incoming data, such as base-line correction or zero-suppression. The data transfer from the CRUs to the FLPs is done via the PCIe interface. A dedicated FLP collects data from the DCS, since calibrations for the detectors are affected by their running conditions.

The data from all different detectors and from the DCS are aggregated on the Event Processing Nodes (EPNs). The EPNs are located on the surface at Point 2 in CR0. The most promising server candidate for the EPN farm, a Supermicro A+ server, is currently being tested. Around 250 servers fully equipped with 8 GPUs each are required to allow for the processing of the incoming data in real time. The interconnection between the FLPs and EPNs is realized as a high performance Infiniband network which is capable of sustaining a high throughput traffic of 625 GB/s [42]. In the synchronous stage the EPNs perform mainly standalone processing tasks for the detectors, e.g. track finding for

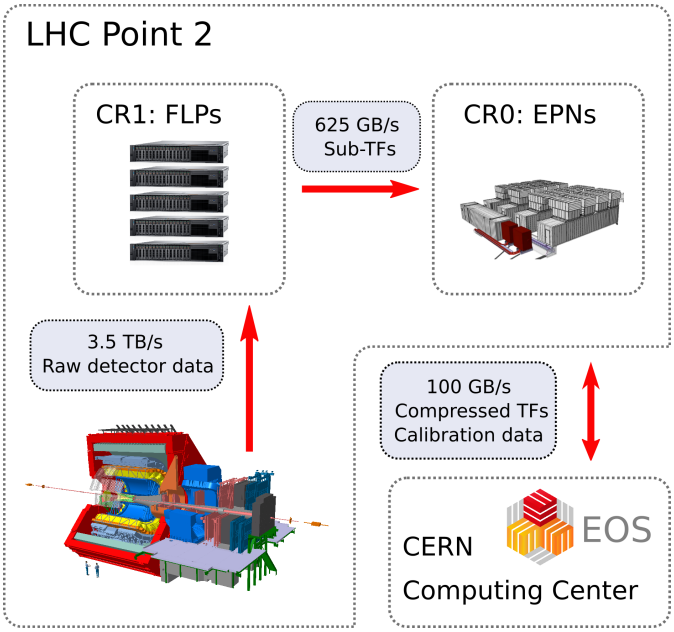


Figure 2.4.: The diagram shows the data flow from the experiment to the counting rooms CR1 (equipped with FLPs) and CR0 (equipped with EPNs) and to permanent storage at the Tier centers.

the ITS and TPC both for data reduction and calibration purposes. As the TPC will take data in a continuous mode and the clusters cannot be assigned to a certain bunch crossing a priori, the data are not processed on a per collision basis, but in the form of Time Frames (TFs), which contain the readout data of all detectors for a given time span. The TF building is described in more detail in Section 2.3.2. Compressed data from this step are stored on a local disk buffer with 60 PB storage space. The output of the calibration tasks is accumulated on a dedicated Condition and Calibration Data Base (CCDB) server. Subsequently, global procedures, e.g. ITS-TPC track matching and TRD tracking using ITS-TPC matched tracks as seeds are performed. The data processing on the EPNs is divided into tasks which run synchronously and asynchronously to the data taking, respectively. The former tasks enable an optimal online compression of the data to reduce the necessary size for the disk buffer and the latter utilize the computing farm in periods where there is no beam for computationally more complex tasks. The layout is further discussed in Section 2.3.3, since it is relevant for the TRD tracking described in Chapter 5.

The global reconstruction enables the final calibration and lastly the event extraction. The output of this last step are Analysis Object Data (AOD) which are distributed over different analysis facilities to facilitate a timely service for the analyzers to study the data.

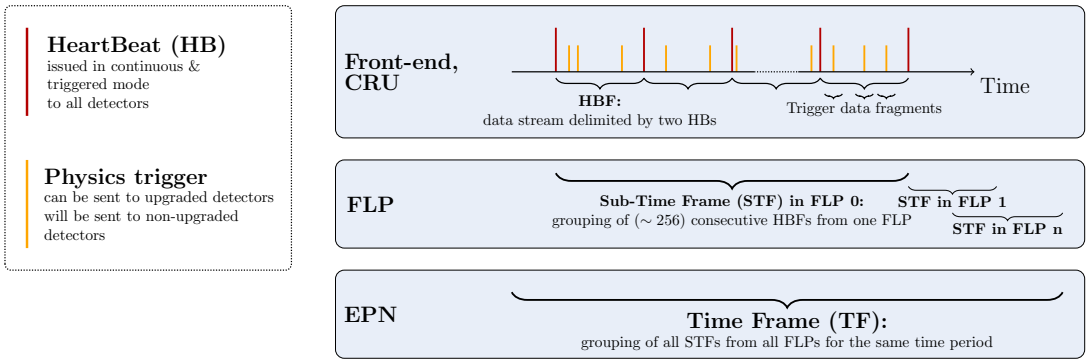


Figure 2.5.: Depiction of the time frame building with continuous and triggered readout.

2.3.2. Time frame building

The data stream from the detectors is delimited by HeartBeats (HBs) which are issued by the Central Trigger Processor (CTP) to all detectors independent of their readout mode (continuous or triggered). They provide the timing information necessary to align data from different detectors in the EPNs. A HB is issued once per LHC orbit which corresponds to 89.4 µs. Two consecutive HBs enclose a HeartBeat Frame (HBF), the smallest coherent chunk of data. HBFs can be dropped if the quality is bad. A configurable number of HBFs are grouped to Sub-Time Frames (STFs) in the FLPs. The technical design report [42] states 256 HBFs per STF which would correspond to a duration of ~ 23 ms, but this could be reduced to alleviate the processing. The STFs from all FLPs for the same time interval are send to a single EPN. The EPN combines all the information and builds the full Time Frame (TF). The TF building scheme is depicted in Figure 2.5.

The EPNs do not receive the data for consecutive TFs. Instead, the data are distributed in a round robin way. The information for the last TPC drift time contained in a TF is incomplete, since some clusters can be contained in the following TF. The TPC drift time is approximately 100 µs. This leads to a data loss below 0.5% for the nominal length of a TF. Hence the optimal length of a TF is a trade-off between data taking efficiency and buffer size. The lost data could in principle be recovered by duplicating the TPC data of the first drift time in a TF an appending it to the previous TF. The increase in processing time and storage size would scale linearly with the recovered statistics. The disadvantage is the possibility to reconstruct the same track twice, if the vertex is near the time frame border.

The upper limit for the processing time $t_{\text{proc}}^{\text{lim}}$ of a TF on an EPN in the synchronous mode is given by

$$t_{\text{proc}}^{\text{lim}} = t_{\text{TF}} \cdot N_{\text{EPNs}} \approx 46\,\text{s} \tag{2.4}$$

for the nominal length of a TF $t_{\text{TF}} = 23$ ms and 2000 EPNs. The total number of EPNs is not necessarily equal to the number of servers which constitute the EPN farm. In case one server is equipped with 8 GPUs one server will host 8 virtual EPNs.

In addition to HBs the CTP sends physics triggers to the detectors which are not read

out in a continuous mode like the TRD and optionally to the other detectors as well. In case of the TRD the physics triggers are used to assign the TRD signal to the TPC clusters of the same interaction. After matching of TPC and ITS track segments the time of the corresponding bunch crossing is known. If a TRD trigger was issued for this bunch crossing the information of the TRD can be matched to the ITS-TPC tracks.

2.3.3. Synchronous and asynchronous data processing

Data processing in ALICE for Run 3 will be divided into synchronous (online) processing during data taking and asynchronous (offline) processing during periods without beam in the LHC. The whole system is optimized for a maximum reduction of permanently stored data and best possible utilization of the on-site server farm in CR0.

The largest contributor to the data volume in ALICE is the TPC with a share above 90% [42]. The TPC data will be compressed in several steps. The first step is the conversion of the charge collected on the readout pads into clusters. The second step is again a lossy compression which is based on one of the two following strategies: Either identified background, for example looping tracks or track segments with a transverse momentum below 10 MeV/c, is rejected or only the cluster assigned to or in close proximity to identified signal tracks are kept. The last step will consist of a lossless entropy compression, which will also be applied to the data of the other detectors.

Since both strategies for the second lossy compression step of the TPC data require the full TPC tracking, it must run synchronously and thus defines the hardware requirements of the EPNs. Parallelization over TPC hits and tracks enables the application of fast parallel computing architectures like GPUs [49].

A small fraction of the events will be used in the synchronous processing for the TPC calibration with track residuals described in Chapter 6. The calibration requires vertexing and tracking with the ITS, matching between ITS and TPC tracks and subsequently TRD tracking and matching to the TOF.

In the asynchronous phase, the full reconstruction with the final calibration is performed for all detectors. Ideally, use of the available GPUs should be made as well, since they will otherwise be idle. Furthermore, future grid sites may deploy GPUs which should naturally be utilized.

3. Space-charge distortions in the TPC

Distortions of the drift lines inside the ALICE TPC caused by space charge are not a novel phenomenon. They were anticipated, simulated and analyzed already at the time the ALICE experiment was designed in the nineties [24]. The 3 to 4 orders of magnitude smaller drift velocity of ions compared to electrons leads to a permanent presence of positive charge inside the large TPC volume. Additional electrical fields induced by these charges affect the drift of electrons from successive collisions towards the readout chambers inside the TPC.

It was concluded during the design phase of the experiment that through the usage of a gating grid the distortions in Pb–Pb collisions should stay within an acceptable limit of $\sim 150\,\mu\text{m}$ in radial direction while they were well within the overall resolution in azimuthal direction [25]. The calculations were based on the LHC design luminosity and assuming a Ne-CO_2 (90-10) gas mixture and a charged particle rapidity density of $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}y = 8000$. The gating grid separates the readout chambers from the drift region and opens only after a collision has occurred for as long as the primary electrons drift, typically $\sim 95\,\mu\text{s}$. Afterwards the gating grid has to be kept closed for a clearing time τ_{clear} such that the ions created inside the readout chamber in the amplification process cannot enter the drift region.

During Run 1 the space-charge distortions in the TPC indeed did not have a significant effect on the reconstruction performance. However, at the start of Run 2 in 2015 much larger distortions than anticipated were observed which will be discussed in Section 3.1. And in Run 3 even larger distortions are expected due to the higher interaction rate and the absence of the gating grid. They have to be corrected via a calibration procedure which is introduced in this chapter.

3.1. Observations in Run 2

In the first year of Run 1 the TPC was operated with a Ne-CO_2-N_2 (90-10-5)[1] gas mixture. Nitrogen was added as an additional quencher to the originally foreseen Ne-CO_2 (90-10) gas mixture [27] for better protection against trips of the HV power supplies. A trip denotes a controlled ramp down of the voltage when the power supply detects a current or charge exceeding a set threshold. From 2011 until 2013 the nitrogen was removed and the TPC was operated with the originally foreseen gas mixture, because it enabled a higher gain with lower anode voltage. In addition, some damage to the front-end electronics was observed in 2010 which could not be excluded to be caused by the nitrogen [50].

[1] The indicated ratios represent units of volume.

Gas	μ_{ion} $(\frac{\text{cm}^2}{\text{Vs}})$	X_{rad} (m)	$\sigma_{\text{p}}(10^{-20}$ cm$^{-2})$
Ne	4.1	322	43.3
Ar	1.5	110	90.3

Table 3.1.: Relevant properties of the counting gases employed in the TPC for estimating the space-charge distortions. The values for the ion mobility are taken from [51] and the values for the radiation length and minimal primary ionization cross-section σ_{p} from [52]. σ_{p} is given for a minimum ionizing particle.

With the Ne-CO_2 mixture the TPC was rather stable and HV trips were very rarely observed. But at highest luminosities during the p–Pb data taking of 2013 some chambers occasionally showed self-sustained currents induced by radiation. Even though the chambers did not trip the data was not usable due to fluctuating voltages. The chambers had to be turned off for several hours for recovery. This was the main motivation to switch to Ar-CO_2 (88-12) for Run 2, since the glow discharge probability for argon is much less compared to neon. Additionally, twice the number of primary electrons N_{p} are created in argon compared to neon, so the TPC can be operated at a factor 2 lower gain without impairing the PID capabilities. The disadvantages of argon are its slower ion mobility μ_{ion} leading to a longer τ_{clear}, its shorter radiation length X_{rad} leading to multiple scattering and lastly the higher number of primary electrons also produces a larger amount of positive charge inside the detector volume which cannot be avoided with the gating grid. The above mentioned values are compared in Table 3.1 for both gases.

Still the advantages of argon outweigh its disadvantages: the lower radiation length was not considered problematic. The minimal efficiency loss due to the increased τ_{clear} is compensated by not having chambers with self-sustained currents which need to be recovered and the PID performance could even be improved. The distortion from space charge were expected to scale by a factor which is proportional to the expected increase in the space charge density ρ_{SC}, given by

$$\frac{\rho_{\text{SC}}^{\text{Ar}}}{\rho_{\text{SC}}^{\text{Ne}}} = \frac{\sigma_{\text{p}}^{\text{Ar}}\mu_{\text{ion}}^{\text{Ne}}}{\sigma_{\text{p}}^{\text{Ne}}\mu_{\text{ion}}^{\text{Ar}}} \approx 5. \tag{3.1}$$

While the simple scaling turned out to be valid for the bulk of the TPC, distortions which were larger by a factor of 50 to 100 were observed for a few small regions at the beginning of Run 2 in 2015. The small regions comprise six regions close to sector edges of some of the IROCs and one region in the center of one of the OROCs.

At first the reason for such large distortions close to the edges of certain IROCs was not known, while for the OROC a gating grid wire whose connection to ground was cut (a *floating* wire) was suspected. Extensive tests including interaction and readout rate scans, runs with special gating grid settings and runs where the anode HV was switched off in every other IROC revealed that [53]:

1) Distortions in the OROC indeed originate from a floating gating grid wire which leads to a local transparency of the gating grid.

2) Distortions in the IROCs are only observed for some sectors and for them always at the same side of the chamber. The observations indicate columns of ions drifting from the readout chambers towards the central electrode in these regions.

3) Distortions show a strong dependence on the interaction rate.

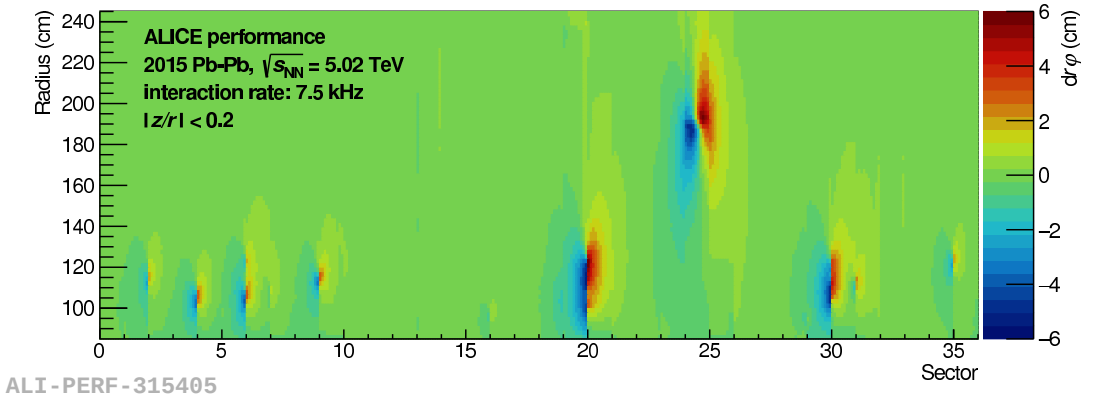
The analysis of Run 1 data showed that the local distortions were already visible in the neon gas mixture, but their magnitude was much smaller. For this reason they were not observed during Run 1 using the standard quality assurance procedure. Unfortunately, at the same time the large space-charge distortions were observed first with the new argon mixture, also other incidents occurred. A broken valve in the gas system lead to an injection from the backup system which might have been contaminated and a strong gain loss which recovered over time was also observed. Thus it was not clear if solely the change to argon was responsible for the large distortions. Therefore the most stable gas mixture Ne-CO_2-N_2 (90-10-5) was injected back into the TPC for 2016−2017. In the meantime the damage to the front-end electronics observed in 2010 was found to be caused by beam background [53] which was strongly diminished in Run 2. With neon the distortions were again of a comparable size as in Run 1 excluding that any permanent damage had been done to the detector either in the long shutdown between Run 1 and Run 2 or in 2015 due to the other incidents.

Distortion mitigation

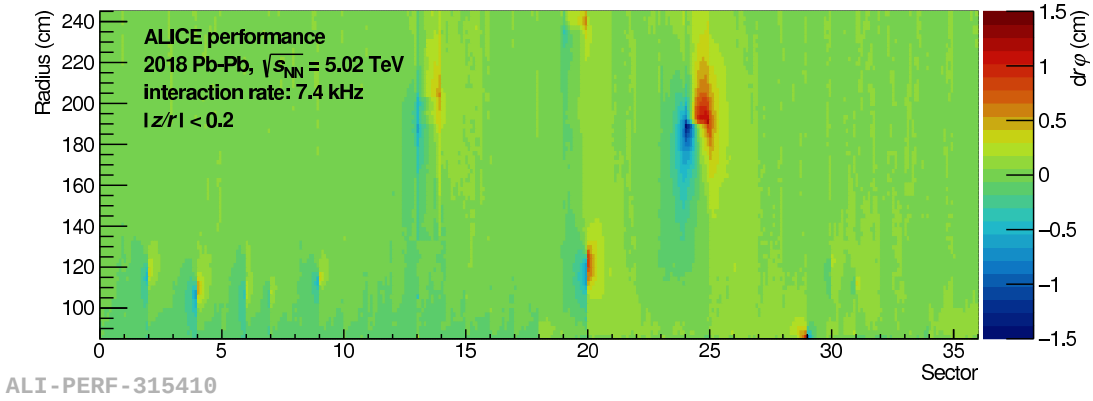
Mitigation of the space-charge distortions in Run 2 was achieved by special cover electrode voltage settings to minimize the charge transfer to and from gaps between chambers. This minimized the distortions originating from the sides of the IROCs. At the location of the floating wire in the OROC the gain was reduced such that a lower number of positive ions is created and in addition the gating grid voltage for the wires surrounding the floating wire was increased to reduce the gating grid transparency. With these measures the TPC could be operated with argon again in 2018, which was important, because at the end of 2018 one month of Pb–Pb data taking was performed. The probability to suffer from HV trips would have been much higher with neon than with argon. How well the distortions were mitigated can be seen when comparing Figure 3.1a with Figure 3.1b. A reduction of the distortions in the hot spots by a factor of about 4 was achieved with the special voltage settings in 2018 compared to 2015.

Origin of the distortions

The results of the extensive tests suggested the origin of the distortions in the affected IROCs to be always at the same side of the respective sectors. It was suspected that insufficient insulation of some wires in this area could cause strong electric fields leading to amplification and therefore columns of ions drifting back from in between the readout chambers. After the TPC was brought to the surface at the beginning of LS2 indeed tips of anode wires sticking out of the ledges were found on all affected chambers near the expected radii.



(a)



(b)

Figure 3.1.: Measured space-charge distortions in $r\varphi$ near the central electrode are shown for Pb–Pb runs with high interaction rates in 2015 (upper plot) and 2018 (lower plot). For both runs the TPC was filled with argon. Note the different z-axis scales. Figures taken from [54].

3.2. Expectations for Run 3

The estimates for space-charge distortions in Run 3 were already done before the large distortions were observed in Run 2 and were based on data collected during Run 1. But since the large local distortions in Run 2 could be traced back to mechanical imperfections of the readout chambers and the chambers are entirely replaced, the estimates are not invalidated.

In order to keep the drift distortions due to space charge well below 10 cm in the majority of the drift volume the Ion BackFlow (IBF), defined as

$$\mathrm{IBF} = \frac{1+\epsilon}{G_{\mathrm{eff}}}, \tag{3.2}$$

is required to be below 1%. Because the effective gas gain G_{eff} is required to be ~ 2000 for a signal-to-noise ratio of 20 [14], the number of ions flowing back into the drift region per incoming electron (ϵ) needs to be below 20.

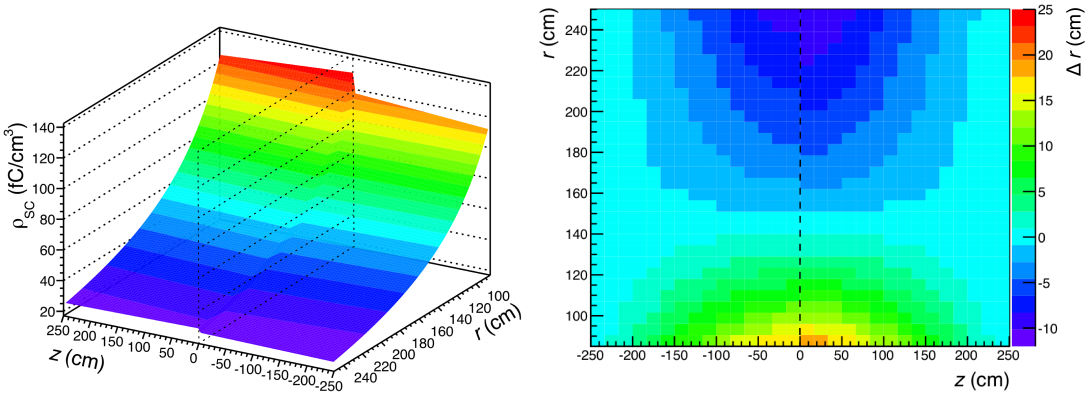
The average drift time for ions in the TPC is $t_d^{\mathrm{ion}} = 160\,\mathrm{ms}$. Hence, for an interaction rate $R_{\mathrm{int}} = 50\,\mathrm{kHz}$ on average ions from 8000 collisions contribute to the space charge inside the TPC at any given moment. The number of ions created in a single collision can be calculated using parameterized charged particle density distributions, the average number of ionization processes for particles traversing the TPC and the given IBF of 1% as introduced in Equation (3.2). This leads to an estimate for the space charge distribution in the TPC $\rho(r, \varphi, z)$ which is depicted in Figure 3.2a. The largest charge density of about $\rho_{\mathrm{SC}} = 120\,\mathrm{fC\,cm^{-3}}$ develops at inner radii close to $r = 85\,\mathrm{cm}$ in the vicinity of the central electrode located at $|z| = 0\,\mathrm{cm}$. The effective distortions of the electrical drift field in cylindrical coordinates ΔE_r, ΔE_φ and ΔE_z are found numerically with a relaxation method [55]. After the drift field distortions are obtained the resulting distortions of the drift paths of the electrons Δr, $\Delta\varphi$ and Δz up to 2nd order are calculated by solving the Langevin equation analytically [56]:

$$m\frac{\mathrm{d}\vec{u}}{\mathrm{d}t} = qe\vec{E} + qe(\vec{u}\times\vec{B}) - K\vec{u}, \tag{3.3}$$

where m is the mass of the particle, $\vec{u}$ its drift velocity, q its charge, e the elementary charge and K the coefficient of friction. The distortions in r are plotted exemplary in Figure 3.2b.

Although the absolute magnitude of the distortions increases compared to the distortions observed in Run 2 (even compared to the large local distortions observed in 2015), the distortion gradients over the TPC volume are expected to be much lower. This is beneficial both for the reconstruction and for the calibration with track residuals presented in the next section. In particular this applies to the smoothing step.

An important factor for the reconstruction is not only the average space-charge distortion, but also its fluctuation. The effect was simulated using superimposed collision data from Run 1 with random positions along z [14]. While the calibration with track residuals introduced in the next section corrects only for the average space-charge distortion and shifts the mean of the TPC cluster residuals to zero, the large RMS remains. The



(a) **(b)**

Figure 3.2.: The average space charge density in the TPC volume is shown in **(a)** for a Ne-CO_2-N_2 gas mixture (90-10-5), $R_{\text{int}} = 50\,\text{kHz}$ and $\epsilon = 20$. Azimuthal symmetry is assumed which is a simplification due to the segmentation of the detector in 18 sectors. The resulting radial distortions in cm are shown in **(b)**. They reach up to 19 cm at the inner radius close to the central electrode, indicated by the dashed lines at $z = 0\,\text{cm}$. Both figures adapted from [14].

fluctuations will be accounted for using the digital currents measured on the readout pads with a frequency of about 1 kHz. The additional calibration of the fluctuations requires the history of the currents at the end plates over the last ion drift time $t_d^{\text{ion}} = 160\,\text{ms}$, which extends to the previous time frames. Neural networks will probably be utilized to realize this complex calibration procedure [14].

3.3. Calibration procedure in Run 2

The appearance of large space-charge distortions in the TPC already during LHC Run 2 required the implementation of a calibration procedure [53] which was originally foreseen only for Run 3 and beyond. Otherwise reconstruction of the data with optimal precision would not have been possible. This section describes the calibration procedure employed in Run 2 in more detail, since it is the basis for the calibration algorithm described in Chapter 6. An important difference is that while the calibration in Run 2 was performed entirely offline, in Run 3 a large part of the calibration is done in the synchronous phase with concurrent performance.

The space point calibration of the TPC uses information from the surrounding detectors, ITS on the inside and TRD and TOF on the outside. The surrounding detectors are not affected by space-charge distortions. Therefore, they can provide reference cluster positions for global tracks inside the TPC.

The calibration is illustrated in Figure 3.3. It consists of the following steps which are all described in more detail in dedicated paragraphs below.

1) Track seeding and following in the TPC is done with relaxed tolerances.

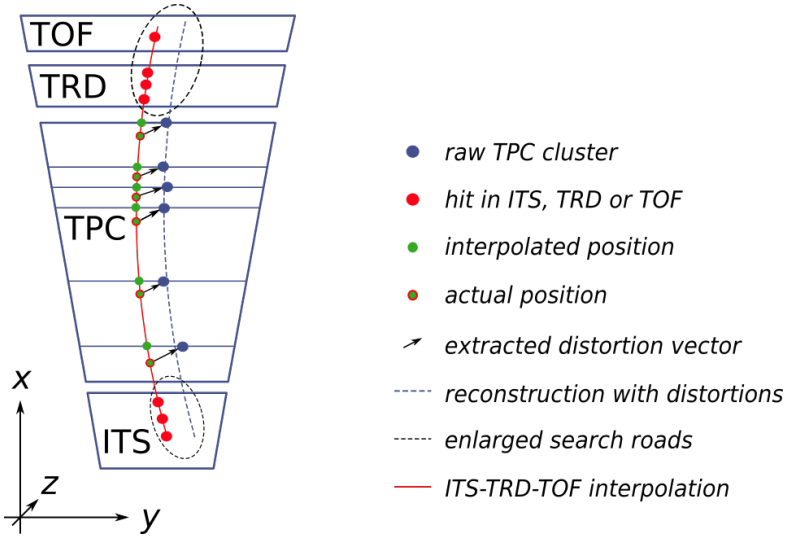


Figure 3.3.: Sketch of the TPC space point calibration procedure. The distorted TPC clusters (blue dots) are compared to reference positions (green dots) obtained via track interpolation (red solid line) between ITS, TRD and TOF. The TPC clusters contain y and z coordinates at a given pad row. Since the position measurement is distorted in all three dimensions, the true cluster position (green dot with red border) is not necessarily on the pad row.

2) The tracks are matched to the ITS on one side and to the TRD and TOF on the other, also with relaxed tolerances.

3) A refit is performed for the track neglecting the information from the TPC, i.e. using only the ITS-TRD-TOF information.

4) The residuals between the distorted TPC clusters and the ITS-TRD-TOF tracks are collected for sub-volumes of the TPC.

5) For each sub-volume a vector representing the distortion in x, y and z is calculated.

6) A polynomial parameterization is created from an interpolation of fitted values of the distortions in sub-volumes with a carefully tuned kernel smoother, which still tries to preserve real discontinuities in the distortions.

1) Track seeding and following

The global reconstruction in ALICE for Run 1 and 2 starts at the outer radius of the TPC, as the track multiplicity is lowest there. After seeds are found the track following is done in inwards direction. A combination of two reference distortion maps is already applied in this step to increase the reconstruction efficiency. One map, $\vec{\Delta}^{\mathrm{IR}^{\mathrm{ref}}_{\mathrm{low}}}$, is created from data taking periods with low interaction rate and accounts for static distortions for example due to misalignment or the inhomogeneous $\vec{E}$-field. A second map, $\vec{\Delta}^{\mathrm{IR}^{\mathrm{ref}}_{\mathrm{high}}}$, accounts for the interaction rate dependence of the space-charge distortions. It is obtained from

data taking periods with a high interaction rate where largest space-charge distortions are present. The static distortions are subtracted from this map and the remaining distortions are scaled according to the current interaction rate $\mathrm{IR}_{\mathrm{curr}}$. Thus, the applied distortion map is calculated according to

$$\vec{\Delta}^{\mathrm{IR}_{\mathrm{curr}}} = (\vec{\Delta}^{\mathrm{IR}^{\mathrm{ref}}_{\mathrm{high}}} - \vec{\Delta}^{\mathrm{IR}^{\mathrm{ref}}_{\mathrm{low}}}) \cdot \frac{\mathrm{IR}_{\mathrm{curr}}}{\mathrm{IR}^{\mathrm{ref}}_{\mathrm{high}}} + \vec{\Delta}^{\mathrm{IR}^{\mathrm{ref}}_{\mathrm{low}}}. \tag{3.4}$$

Still, in the presence of large space-charge distortions the TPC clusters can be $\mathcal{O}(\mathrm{cm})$ away from the helix parameterization of the trajectory in the $\vec{B}$-field. In order to still be efficient, the tolerances for the residuals have to be enlarged by an additional 3.5 cm. Furthermore, the TPC cluster uncertainties are artificially increased by 2 cm for the reconstruction [53]. This in turn leads to a higher number of fake hits attached to the tracks, especially in Pb–Pb collisions with high track density. Additional filtering and outlier rejection is therefore necessary. Also, events with a too large number of primary-like tracks are not taken into account for the calibration.

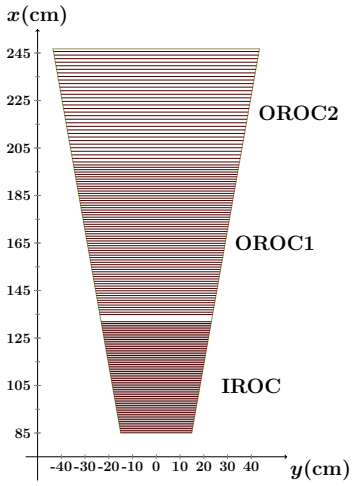
2) Matching to the surrounding detectors

The TPC tracks can have an offset with respect to the clusters reconstructed in the ITS, TRD or TOF detectors. This is not primarily because of the space-charge distortions, but due to the fact that in the very first reconstruction the drift velocity of the TPC is only known with a precision of about 1% before the matching.

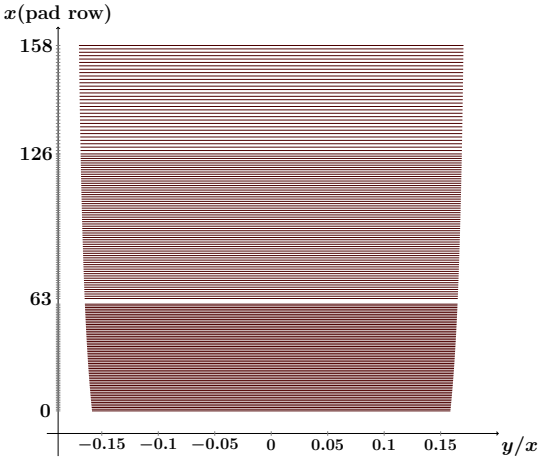
The matching windows are proportional to the uncertainties of the track parameters. Hence, the artificially increased TPC cluster uncertainties lead to a larger search road, but a constant value is added in addition nevertheless. Matching clusters are identified first in the ITS and subsequently in the outward direction in the TRD and TOF, as it is also done in the standard ALICE tracking procedure described in Chapter 4.

3) ITS-TRD-TOF refit

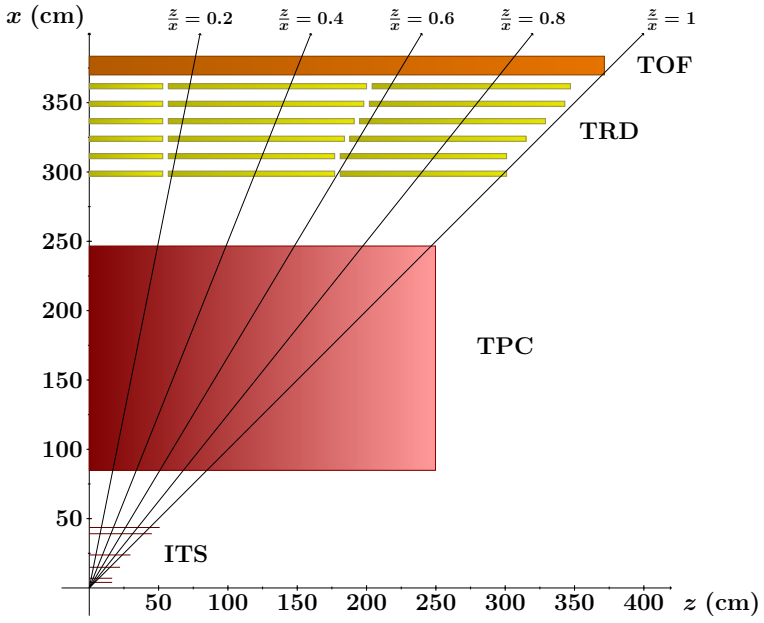
To obtain a reference position for the TPC clusters, the tracking information from the TPC needs to be neglected and a refit using only the information from the ITS, TRD and TOF has to be performed. Technically, two independent refits are done. One starts at the outermost layer in which a matching cluster has been found, either in the TRD or in TOF. Another one starts at the innermost layer of the ITS which has a match. The refits are done by updating the track with all available points either in the TOF and TRD or in the ITS. The refitted tracks are farther propagated inwards and outwards, respectively. At each pad row in which a TPC cluster has been previously associated to the track a weighted average from both refits serves as reference position and the residuals in sector y and z coordinates are stored. The sector coordinate system is a right-handed system with the x-axis being along the bisector of the sector pointing outwards in radial direction, the y-axis parallel to the pad rows and the z-axis along the beam line.



(a) Azimuthal cross section for a TPC sector.

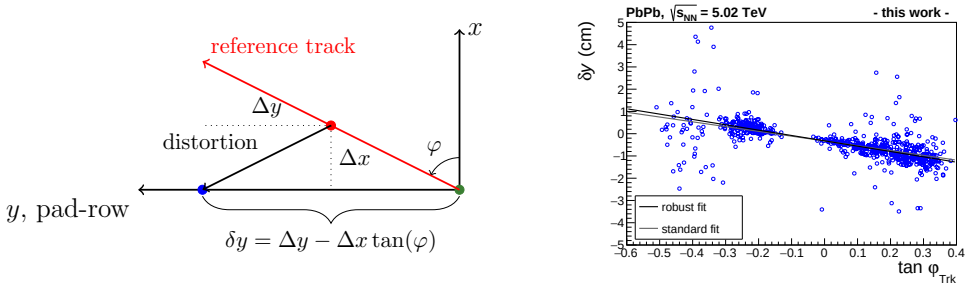


(b) Binning in y/x



(c) Binning in z/x

Figure 3.4.: Depiction of the TPC voxel binning. In radial direction the binning is done per pad row (a). It is not uniform, as the pad row pitch varies between IROC (0.75 cm), OROC1 (1 cm) and OROC2 (1.5 cm). The lines are drawn at the center of each pad row. In pad direction the binning is done in y/x and not in y, to obtain a more uniform bin size (b). Each depicted pad row is divided into 15 equal bins. The bins in y/x are not entirely uniform because of the fixed size dead zone between sectors of 1.5 cm. In (c) the binning in longitudinal direction is visible. Tracks with $z/x > 1$ are not used for the calibration. Only the active areas of ITS, TPC, TRD and TOF are drawn. All sketches are drawn to scale.



(a) **(b)**

Figure 3.5.: The left plot illustrates the fact that the measured distortion δy at the pad row is a convolution of a radial distortion Δx and distortion in azimuth Δy. The coloring scheme is inherited from Figure 3.3. The right plot shows the measured distortions δy as a function of the track inclination for a voxel with large distortions.

4) TPC voxel binning

The distribution of space charge is not uniform inside the TPC as it is described in Section 3.1. In addition, the calibration with track residuals can only serve as a *time-averaged* correction for the cluster position. Otherwise, if the track interpolation between the ITS, TRD and TOF would yield the ideal track position inside the TPC volume, the tracking information from the TPC could simply be ignored for the final refit. For the aforementioned reasons two requirements arise: the residuals need to be collected separately for different volume elements of the TPC (called voxels) and a minimum amount of global tracks is required to determine the average distortions with sufficient precision. The number of tracks depends naturally on the chosen size of the voxels. For the space charge distortion calibration in Run 2 the standard voxel binning for each of the TPC sectors is chosen as follows:

- There are 159 bins in radial direction, one bin for each pad row.
- 15 bins are chosen in the transverse plane for the ratio of y/x.
- 5 bins split the ratio of z/x in the longitudinal plane.

Hence, the 36 sectors of the TPC are altogether divided into 429300 voxels. The cross section for a TPC sector and the voxel binning is visualized in Figure 3.4. The space charge distortion calibration is performed for time intervals of 40 minutes to collect a sufficient number of tracks, around 1000 tracks per voxel or accordingly about 1-2M global tracks in total in the Pb–Pb data taking periods in 2015 and 2018.

5) Extraction of 3D distortions

The reference positions obtained as explained in the paragraph above do not correspond to the true cluster positions without distortions. As the radial position for the clusters

is given by the pad row on which these are detected, the distortions in x cannot be measured directly. Nevertheless, they affect both the distortions measured in y and z, called δy and δz.

Because δy and δz are collected per voxel, the effective distortions Δx, Δy and Δz can be disentangled. The procedure is shown for Δy and Δx in Figure 3.5a. Since the same TPC region is probed by tracks with different inclination angles φ with respect to the sector axis, a *robust* fit of δy as a function of $\tan\varphi_{\text{trk}}$, as is shown exemplary for one voxel in Figure 3.5b, yields Δy and Δx. The robust fit includes a truncation of the data sample and successively the absolute deviations to the straight line parameterization are minimized instead of a simple least squares minimization, in order to be less susceptible to outliers. The exact fit procedure is described in Section 6.4, since it is done in an analogous manner in Run 3.

A similar procedure is performed to disentangle the distortions in x and z. Tracks that probe the same TPC voxel in a good approximation have the same dip angle λ corresponding to the track inclination in the longitudinal xz plane. Hence, Δz can be extracted from the measured δz via

$$\delta z = \Delta z - \Delta x \tan\lambda, \tag{3.5}$$

with Δx known from the previous step.

The dispersion σ_D of the points in Δy with respect to the fitted values are calculated for quality assurance purposes and they also provide an estimate of the distortion fluctuations. The measured σ_D is not determined only by the distortion fluctuations, but it contains the finite resolutions of the TPC clusters and the track interpolations as well. The real fluctuations σ_{real} are thus given by the quadratic difference of the measured distortions and the dispersion measured for a data taking period with a very low interaction rate $\sigma_{D,0}$ where space-charge distortions are absent:

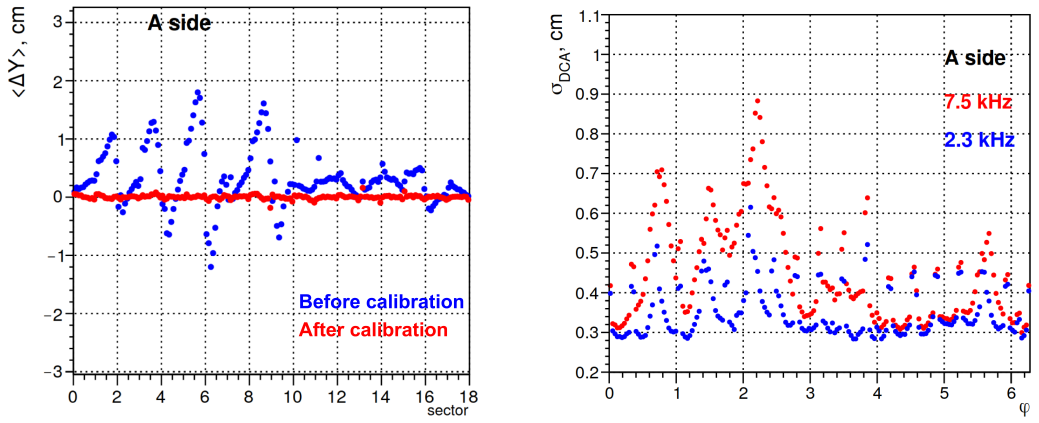
$$\sigma^2_{\text{real}} = \sigma^2_D - \sigma^2_{D,0}. \tag{3.6}$$

For the processing of Run 2 data the fluctuations were accounted for via adding an additional uncertainty to the TPC cluster positions based on the dispersion of the points inside that voxel. Measuring the currents of the readout pads with high granularity in time as it is foreseen for Run 3 was not possible.

6) Smoothing

Parts of the TPC volume are not covered by reference ITS-TPC-TOF tracks. Permanent holes in the detector acceptance of the TRD and TOF detectors to reduce the material budget in front of the PHOS calorimeter and temporary holes which can occur in all reference detectors lead to voxels without reference tracks. In addition, for some voxels the number of traversing tracks is very low and therefore the extracted distortions can fluctuate significantly.

To recover empty voxels or voxels with an insufficient amount of reference tracks a kernel smoother is applied: an interpolation is attempted from neighboring voxels within



(a) (b)

Figure 3.6.: The left plot shows the mean values for the DCA in y for primary tracks before (blue) and after (red) applying the distortion calibration for one side of the TPC. The right plot shows the TPC track resolutions at different interaction rates as a function of the azimuthal angle φ which is approximately propotional to the sector number for the same side of the TPC. While the DCA bias can be corrected for, the space charge fluctuations deteriorate the track resolution especially in the regions of large distortions. Both figures adapted from [53].

the same sector. Different kernels can be chosen for this step. In Run 2 Epanechnikov and Gaussian kernels were used. The kernel defines the weight of the contributions from neighboring voxels to the interpolation or extrapolation point.

The smoothed values for the distortions in all three dimensions and for the dispersion in y are subsequently parameterized by Chebyshev polynomials to allow for a fast query of Δx, Δy, Δz and σ_D as a function of pad row, y/x and z/x [53]. The obtained Chebyshev parameterization is afterwards used in the subsequent reconstruction of the full data which was recorded in the given time period.

3.4. Performance

As a result of the calibration a map of vectors is obtained which contain the distortions in all three dimensions for each voxel:

$$\vec{r}_{\text{cls}} = (x_{\text{ro}},\ y_{\text{ro}},\ z_{\text{ro}}) + \vec{\Delta}(x_{\text{ro}},\ y_{\text{ro}},\ z_{\text{ro}}). \tag{3.7}$$

Note that this map contains not only the distortions due to space charge, but also static distortions as for example inhomogenities of the magnetic field, the $\vec{E} \times \vec{B}$ effect or mechanical imperfections of the field cage:

$$\vec{\Delta}^{\text{tot}} = \vec{\Delta}^{\text{static}} + \vec{\Delta}^{\text{SC}}. \tag{3.8}$$

As mentioned above calibration maps created from runs with a low interaction rate yield the static distortions $\vec{\Delta}^{\text{static}}$ and serve as a reference map. The distortions due to space charge $\vec{\Delta}^{\text{SC}}$ are obtained by subtracting the static from the measured distortions.

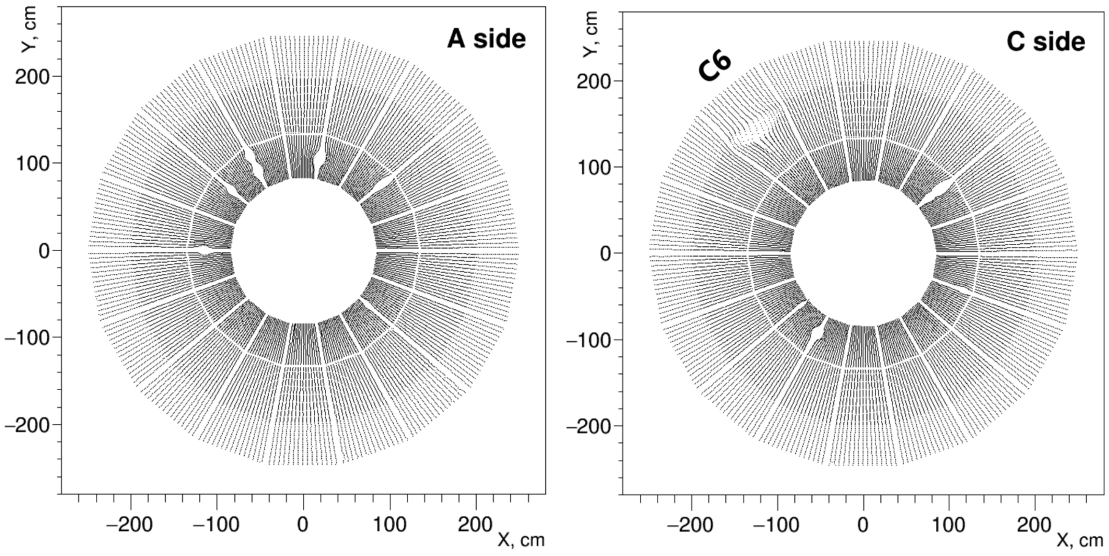


Figure 3.7.: Modified acceptance topology for the TPC close to the central electrode under the presence of large space-charge distortions. The dead zone between the effected IROCs is increased. Furthermore the floating gating grid wire in the OROC of sector C6 influences the acceptance. Adapted from [53].

Applying the map in the reconstruction minimizes the bias introduced by space-charge distortions on the measured Distance of Closest Approach (DCA) from the trajectory to the primary vertex as it is shown in Figure 3.6a. The resolution of the TPC tracks however remains degraded as it is shown in Figure 3.6b. To improve it a better time granularity for the distortion measurement is needed which will be available in Run 3 when the measurement of the currents of the readout pads allows for additional maps with a granularity of about 5 ms [14].

Effect on Monte Carlo simulation

Monte Carlo (MC) simulations are an effective tool for example for determining the detection sensitivity to different physics processes. In order for this to work the simulated and the real detector response need to be consistent. Now even if the above described calibration of the space-charge distortions were perfect an effect on the reconstruction efficiency would remain nevertheless. On the one hand the fluctuations lead to an increased measurement uncertainty for the TPC clusters. On the other hand columns of ions which move from the readout chambers to the central electrode in between the TPC sectors effectively enlarge the dead zone of the detector. The drift electrons which are liberated by traversing charged particles close to a sector boundary end up in between two sectors. The resulting acceptance topology is depicted in Figure 3.7. This effect is strongest for clusters which are produced in the vicinity of the central electrode and decreases for clusters at large $|z|$.

To account for the effect of the distortions in simulations the following procedure is applied. The correction maps obtained from the steps described in Section 3.3 are inverted to obtain distortion maps. These distortion maps are used to distort the simulated cluster positions inside the TPC. The fluctuations are simulated by generating a random number R from a Gaussian distribution with zero mean and a width of 1 for each event and increasing the time-averaged distortions of all hits by

$$R \cdot \sigma(x, y, z), \tag{3.9}$$

where σ is the estimate of the fluctuation at the given point. Therefore, a full correlation is assumed for the fluctuations inside the whole TPC. For the simulation of Run 2 data this simplification is appropiate, since the tracks traverse at most one of the regions with large distortions and therefore the introduced correlations between the different hot spots are not important [53].

Afterwards the reconstruction of the simulated data is done with the same correction maps which were obtained from real data and applied in its reconstruction.

4. Global track reconstruction

This chapter gives an overview of the global tracking procedure of ALICE with special emphasis on the Kalman filter technique [57], since it is also employed in the online TRD tracking algorithm described in the successive chapter. In this work global reconstruction refers to the reconstruction of particle tracks in the central barrel detectors. The reconstruction of data from the muon spectrometer and from cosmic data is described elsewhere [13, 58].

At the moment three different approaches coexist for the global reconstruction:

1) The current offline reconstruction is employed in official productions with data collected during Run 1 and 2 and in corresponding simulations.

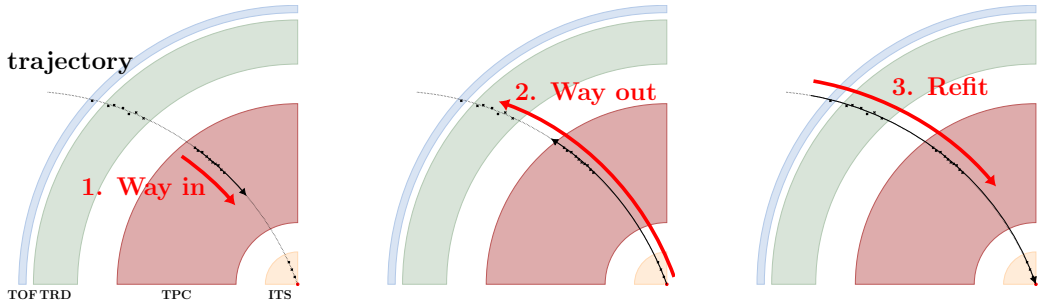
2) The online reconstruction in the HLT was running during data taking in Run 1 and 2 and served mostly for testing and developing faster reconstruction algorithms with similar performance as compared to the offline reconstruction in view of Run 3.

3) The global reconstruction for Run 3 within the O^2 framework is currently under development and will allow one to cope with continuous readout instead of an event based readout.

Most relevant for the work described in this thesis is clearly the third point. Hence, it will be discussed in greater detail below. Since the detector simulation for the TRD in the O^2 framework is not yet ready for operation, the development for the TRD tracking is based entirely on data from Run 1 and 2. Additionally, for reasons explained in the next chapter, the TRD tracking is implemented in a way such that it can accept data both from the offline and the HLT framework. Therefore, also the first and the second approach will briefly be explained.

4.1. Current offline reconstruction

A detailed overview of the current ALICE offline reconstruction including the resulting performance is given in [13]. First, clusterization is done in all detectors individually. A preliminary vertex is determined from clusters in the two innermost ITS layers. With the help of the preliminary vertex position, track seeding is performed at the outer TPC radius. Now an inward-outward-inward scheme is employed for all track candidates utilizing the Kalman filter technique. The scheme is illustrated in Figure 4.1. The seeds are propagated inwards to the inner wall of the TPC and the track parameters are



(a) Seeding and inward fit **(b)** Outward fit **(c)** Final refit

Figure 4.1.: The global reconstruction approach used in the offline processing of Run 1 and Run 2 data is shown. After track seeding at the outer radius of the TPC an inward-outward-inward scheme is employed.

updated along the way if matching clusters are found within a variable search road which is adjusted according to the event multiplicity and the track parameter uncertainties.

From now on the tracks become seeds for the ITS tracking. In multiple passes the best matching ITS clusters are identified and added to the tracks. Afterwards the tracks are propagated to the point of closest approach to the preliminary vertex and the back propagation starts. During the back propagation the tracks are refitted using the clusters found in the previous stage. If possible, the tracks are prolongated into the outer detectors, the TRD, TOF and the calorimeters. Lastly, the tracks are propagated inwards and a final refit with all available clusters is performed.

For all reconstructed tracks the track parameters are stored at different steps. Particularly the track parameters at the outer wall of the TPC before they are extrapolated to the outer detectors are stored in the Event Summary Data (ESD). These track parameters are used as seeds in the development phase of the TRD tracking.

4.2. HLT reconstruction

In contrast to the offline reconstruction, which is in principle not bound to specific processing time constraints, the HLT has to cope with the incoming data in real time. The HLT makes use of heterogeneous computing architectures and various levels of parallelization to increase its computing power. The TPC cluster finder runs on FPGAs, the TPC reconstruction supports Single Instruction, Multiple Data (SIMD) vector instructions and can run both on CPUs and GPUs.

Unlike the offline reconstruction the HLT applies a cellular automaton [59] for the seeding procedure within the whole TPC volume which does not rely on a preliminary vertex estimate from the ITS. For the subsequent track following again the Kalman filter is used as in the offline reconstruction. Various other differences to the offline reconstruction exist, but they are more of a technical nature and not relevant for the work carried out in this thesis. One exception are the different data types implemented

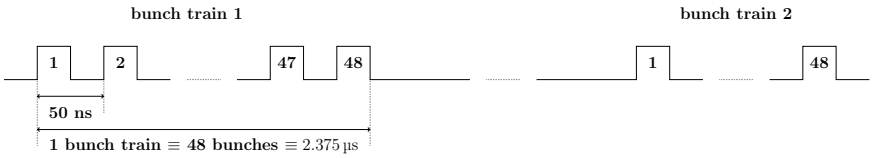


Figure 4.2.: Possible bunch train structure for Pb-ions in LHC Run 3. Figure adapted from [14].

in the offline reconstruction and in the HLT. This necessitates an additional interface layer to enable the processing of both tracks from offline and from the HLT by the TRD tracker. More details will be given in Section 5.5.2.

The resulting performance of the HLT reconstruction is compatible with offline in terms of efficiency with a slightly worse resolution [60]. It is however about 300 times faster than the offline reconstruction. The speedup is composed of a factor 20 when both reconstructions are performed on a single CPU core and an additional factor of 15, when the work is offloaded to the AMD S9000 GPUs of the HLT farm [61]. A more recent GPU yields an even greater speedup factor. Overall the HLT is able to process data at an input rate of 48 GB/s which corresponds to the highest input rate anticipated for Run 1 and 2.

4.3. Reconstruction in Run 3

Global reconstruction in the upgraded online-offline computing system O^2 is significantly different from the reconstruction approaches used in LHC Run 1 and 2. On the one hand it will be divided into a synchronous and an asynchronous stage as described in Section 2.3.3. On the other hand the continuous readout implies fundamental changes especially to the track reconstruction in the TPC.

4.3.1. Conditions in Run 3

Before going into more detail, the expected conditions in the TPC for Run 3 shall be estimated. For this a simplified LHC filling scheme for Pb ions is assumed as it is done in [14]. It consists of 12 equally spaced bunch trains with 48 bunches each, resulting in 576 bunches per ring. With a bunch spacing as indicated in Figure 4.2 each train corresponds to a length of 2.375 µs. The gap between two trains has a length of 5.1 µs. Hence, the instantaneous bunch crossing rate of $\frac{1}{50\,\mathrm{ns}} = 20\,\mathrm{MHz}$ translates to an average bunch crossing rate of 6.35 MHz. The probability for a collision to occur in a single bunch crossing can now be determined considering an average collision rate $R_{\mathrm{int}} = 50\,\mathrm{kHz}$:

$$\mu = \frac{R_{\mathrm{int}}}{6.35\,\mathrm{MHz}} = 0.0079. \tag{4.1}$$

Assuming a Poisson distribution with the expectation of $\mu \cdot 48$ collisions per bunch train crossing the probability for at least one collision or more than one collision per bunch

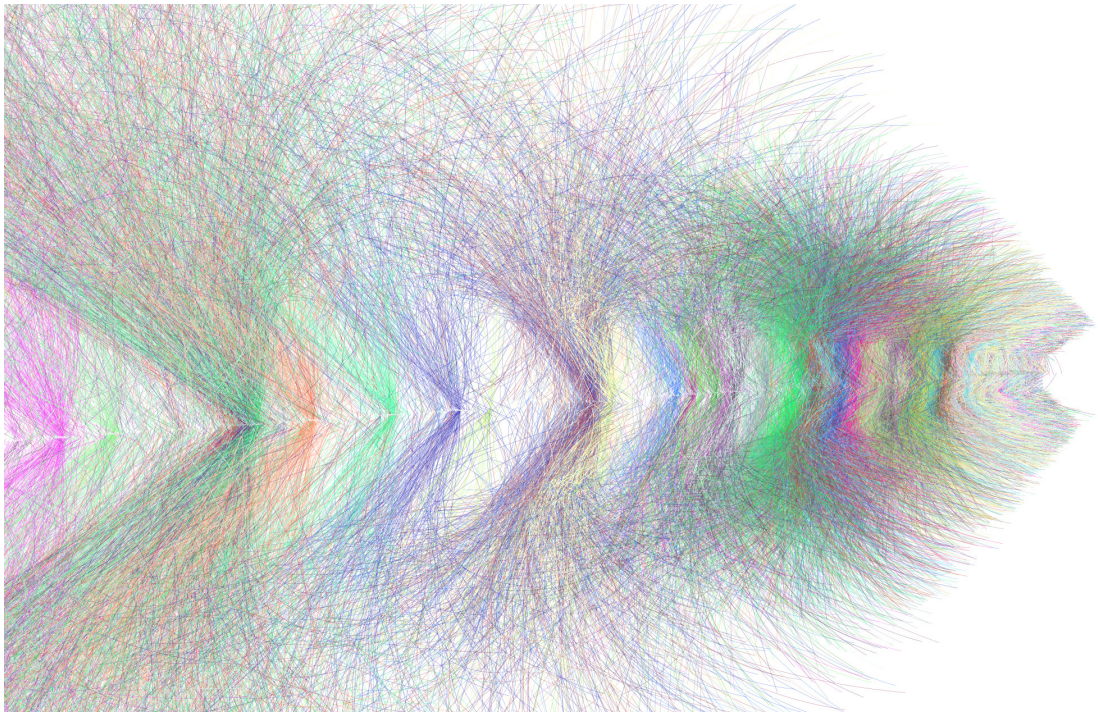


Figure 4.3.: Visualization of a 2 ms time frame with Pb–Pb collisions at 50 kHz. All tracks reconstructed in two opposing TPC sectors and propagated to the vertex are shown with different colors for different interactions [62].

train crossing is 31.6% and 5.6%, respectively. Furthermore, if a collision occurs, the probability for another collision within the same bunch train crossing is ∼17.7%. What these numbers imply for the TPC is illustrated in Figure 4.3. During one electron drift time the tracks from typically 4 to 5 collisions are present in the detector. Therefore, the tracking will be done for full time frames instead of single collisions [49]. Otherwise the data volume would be dominated by pile-up.

4.3.2. Synchronous processing of time frames

The synchronous processing is mainly driven by the need to compress the TPC data, which requires the full TPC tracking as mentioned in Section 2.3.3. The ITS tracking will be performed only partially at this stage, to allow for the space point calibration of the TPC on the one hand and for quality control on the other hand. In addition, the FIT and ZDC reconstructions are done at the synchronous stage providing interaction times and event planes.

The TPC tracking is fully performed on the GPUs of the EPNs. It starts with clusterization and subsequently track seeds are found using the cellular automaton: Triplets of hits forming a straight line in adjacent pad rows are found and compatible triplets are concatenated to form seeds. After the seeding, track following is performed using

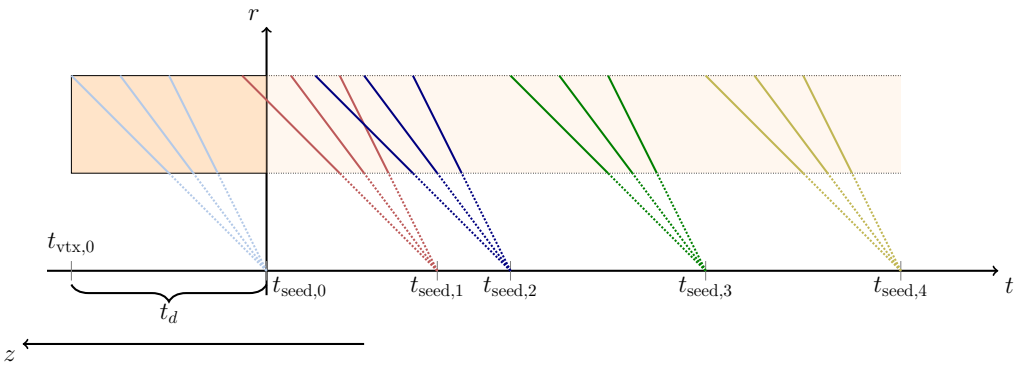


Figure 4.4.: The TPC tracking is sketched for the A-side of the TPC. In the first step the reconstruction is not limited to the detector volume, as the interaction time t_{vtx} is unknown. In the depicted example an interaction occurred at $t_{\mathrm{vtx},0}$. If the clock was started at the same time ($t_{\mathrm{vtx},0} = 0$), then the extrapolated $t_{\mathrm{seed},0}$ would be equal to one full drift time t_d. Only after the interaction time has been determined from the track seeds, the tracks are shifted in z into the actual detector volume.

the Kalman filter. All these steps are done using only a relative z coordinate, since the association of clusters to a certain interaction is a priori not known. Only an initial cluster calibration is performed assuming the cluster belongs to a track within $|\eta| = 0.45$ to maximize the efficiency of the seeding step. The η assumption yields a hypothetical z position with a maximum error of about half a drift length for tracks within the acceptance of the full TPC ($|\eta| < 0.9$). The hypothetical z position allows for an ad-hoc correction of the space-charge distortion using a scaled average distortion map.

Usually the z position of the clusters in the TPC is determined by the drift time t_d of the clusters to the end plates (see Equation (2.1)). In a continuous mode, instead of the drift time, the clusters are assigned the time

$$t' = t_d + t_{\mathrm{vtx}}, \tag{4.2}$$

which is measured with respect to the beginning of the time frame. Now the time of the interaction t_{vtx} needs to be determined to obtain the actual drift time t_d. With the help of the track seeds, assuming they belong to primary tracks, t_{vtx} can be estimated. The seeds are extrapolated in the x, y, t space to the interaction region ($x = y = 0$). The vertex time is then given by the extrapolated time at $x = y = 0$ minus one additional drift time:

$$t_{\mathrm{vtx}} \approx t_{\mathrm{seed}}(x = y = 0) - \frac{z_{\mathrm{max}}}{v_d}. \tag{4.3}$$

The procedure is illustrated in Figure 4.4.

Note that the vertex time is only an approximation, since the interaction does not necessarily occur at the nominal interaction point $z_{\mathrm{lab}} = 0$, but has a width of about $\sigma_z = 7\,\mathrm{cm}$ [14]. As a matter of fact the vertex spread helps at a later stage since tracks which cross the central electrode of the TPC and which are reconstructed both on the A-side and on the C-side help to determine the TPC drift velocity v_d with high accuracy.

After t_{vtx} is determined tracks with a compatible t_{vtx} are grouped together as they are assumed to originate from the same interaction. The vertex spread σ_z corresponds to a drift time of about 2.7 µs. Subsequently the estimate for t_{vtx} can be refined via a weighted average from all t_{vtx} estimates within one group of tracks. The weight is given by the uncertainty in z of the corresponding extrapolated TPC track.

Parallel to the TPC tracking ITS standalone tracking is performed. For all tracks found in one ITS readout frame the TPC tracks with a matching t_{vtx} are selected and matched, refining again the z position of the TPC tracks to $\sigma_{z,\text{trk}} \approx 100\,\text{ns} \cdot v_d$.

After a refit the ITS-TPC tracks are prolonged into the outer detectors TRD and TOF where they are matched to the clusters which belong to the same interaction based on the trigger time stamps which are available for the TRD and TOF.

4.3.3. Asynchronous processing

The tracking in the synchronous stage works well for primary tracks, but not so well for deep secondary tracks, which produce only a few clusters in the outer ITS layers. Furthermore, there can be two collisions in one ITS readout frame which has the length of $\tau^{\text{ITS}}_{\text{ro-frame}} \approx 2$ to 20 µs. In the asynchronous phase the full ITS tracking is done for all collisions. Multiple passes are executed with different constraints for example for incomplete or off-vertex tracks. Also the different interaction times for the collisions in a given time frame provided by the FIT are used at this stage. They serve as constraints when matching the remaining ITS clusters to TPC tracks. Further improvements for the track reconstruction of deep secondary particles are currently under study. The results of the final reconstruction are stored in the Analysis Object Data (AOD) which is distributed over different analysis facilities within the Worldwide LHC Computing Grid [63].

4.4. Mathematical basics

As described above the Kalman filter technique is utilized in various places in the global reconstruction in ALICE. Also the online TRD tracking will rely on it. This section provides the mathematical basis for the Kalman filter at the level of detail which is required to understand the TRD tracking algorithm. For more specific descriptions of subjects which do not affect the work carried out in this thesis references will be provided in given places. Before the filter itself will be discussed, the ALICE track model is introduced.

Trajectories of charged particles in a magnetic field

In a static magnetic field $\vec{B}$ a charged particle is affected by the Lorentz force [64]

$$\vec{F}_{\text{L}} = q\,\vec{v} \times \vec{B}, \tag{4.4}$$

where q is the signed charge of the particle and $\vec{v}$ its velocity. If there is no electrical field and furthermore material effects and bremsstrahlung are ignored, then the equation

of motion can be deduced from Equation (4.4) to be

$$\frac{\mathrm{d}\vec{p}}{\mathrm{d}t} = q\,\vec{v} \times \vec{B} \tag{4.5}$$

with the momentum $\vec{p} = \gamma m \vec{v}$, where m is the particle's rest mass, and the time t. Since energy loss is ignored ($\mathrm{d}m_{\mathrm{rel}}/\mathrm{d}t = 0$) and with $\vec{v} = \frac{\mathrm{d}\vec{r}}{\mathrm{d}t}$ this becomes

$$m\gamma \frac{\mathrm{d}\vec{v}}{\mathrm{d}t} = q\,\vec{v} \times \vec{B} \tag{4.6}$$

$$\Leftrightarrow m\gamma \frac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2} = q \frac{\mathrm{d}\vec{r}}{\mathrm{d}t} \times \vec{B}. \tag{4.7}$$

When the time is replaced by the path length along the trajectory $\mathrm{d}s = v\,\mathrm{d}t$ the equation of motion can be rewritten in the form of geomtrical quantities only:

$$m\gamma v^2 \frac{\mathrm{d}^2\vec{r}}{\mathrm{d}s^2} = qv \frac{\mathrm{d}\vec{r}}{\mathrm{d}s} \times \vec{B} \tag{4.8}$$

$$\Leftrightarrow \frac{\mathrm{d}^2\vec{r}}{\mathrm{d}s^2} = \frac{q}{p} \frac{\mathrm{d}\vec{r}}{\mathrm{d}s} \times \vec{B} \tag{4.9}$$

The magnetic field in ALICE is approximately homogeneous and only along the beam direction in z: $\vec{B} = (0, 0, B_z)$. When inserted into Equation (4.9) one thus obtains an array of differential equations which is solved by a helix with an axis parallel to z and a circle in the xy plane [65].

Using s as a running parameter, the trajectory is described by

$$x(s) = x_0 + R[\cos(\varphi_0 + hs\cos(\lambda/R)) - \cos(\varphi_0)] \tag{4.10}$$

$$y(s) = y_0 + R[\sin(\varphi_0 + hs\cos(\lambda/R)) - \sin(\varphi_0)] \tag{4.11}$$

$$z(s) = z_0 + s\sin\lambda \tag{4.12}$$

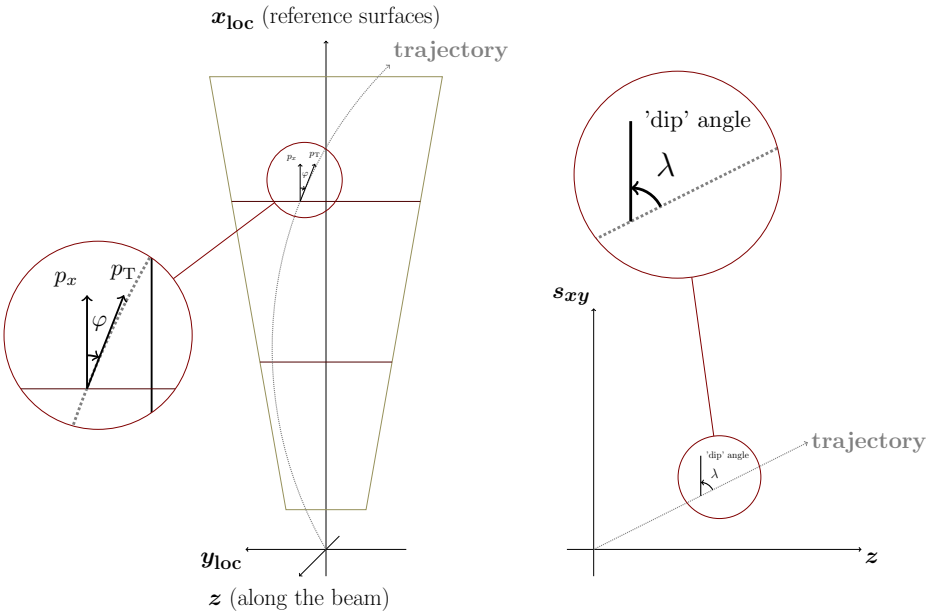
i.e. by a starting point x_0, y_0, z_0, two angles φ_0 and λ giving the direction at the point s and the radius R. The sense of rotation of the trajectory projected on the transverse xy plane is given by $h = \mathrm{sgn}(qB_z)$. Note that five parameters are sufficient to uniquely describe the helix on a reference surface, because

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \quad \Rightarrow \quad (\mathrm{d}x\,/\,\mathrm{d}s)^2 + (\mathrm{d}y\,/\,\mathrm{d}s)^2 + (\mathrm{d}z\,/\,\mathrm{d}s)^2 \equiv 1 \tag{4.13}$$

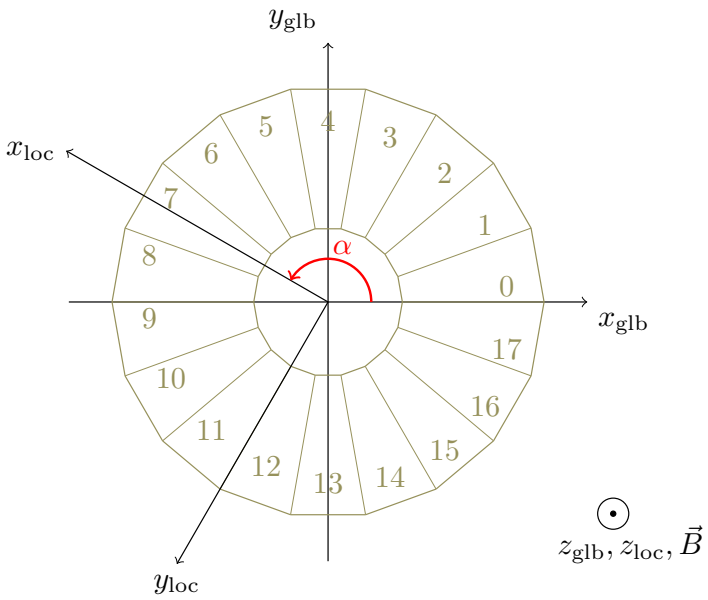
and therefore the six integration constants obtained when integrating Equation (4.9) are not independent.

4.4.1. The ALICE track model

The track fitting in ALICE is performed in the sector coordinate system. As described in Section 3.3 it is a right-handed system with x-axis along the bisector of the sector pointing outwards in radial direction, the y-axis lying inside the detection planes and z-axis along the beam line. Out of the detectors relevant for this thesis the TPC, TRD



(a) Trajectories projected in the transverse (left) and longitudinal plane (right).



(b) Global ALICE and the local coordinate system for sector 7.

Figure 4.5.: The ALICE track parameters $P_3 = \sin\varphi$ and $P_4 = \tan\lambda$ are illustrated in **(a)**. In **(b)** the global and local coordinate systems and the sector numbering scheme are shown.

and TOF are segmented in the transverse plane in 18 sectors covering 20 degrees each. The detection planes are arranged tangentially to circles with fixed radii r_i. In the sector coordinate system each measurement yields y_i and z_i at a perpendicular distance x_i from the nominal interaction point. The detection planes are therefore the natural choice for the reference surfaces. At a given detection plane x the track is defined by the coordinates y_x, z_x, the inclination of the track in the transverse plane φ_x, the slope ('dip') angle $\lambda_x = \arcsin\left(\frac{\mathrm{d}z}{\mathrm{d}s}\right)$ and the signed radius r_H of the helix in the transverse plane which is proportional to the transverse momentum p_T. In conclusion, tracks in ALICE are parameterized as

$$\vec{P}_\mathrm{trk} = \begin{pmatrix} y \\ z \\ \sin(\varphi) \\ \tan(\lambda) \\ q/p_\mathrm{T} \end{pmatrix} \tag{4.14}$$

at given x. The fifth parameter contains the charge q of the particle and the inverse of the transverse momentum as the errors on $1/p_\mathrm{T}$ are Gaussian whereas the errors on p_T are not [66]. For an illustration see Figure 4.5a. This parameterization facilitates the required computations for the reconstruction, since most of the detectors provide measurements of y and z at a given detection plane plane x. The compatibility of these measurements with the track can easily be estimated. This will be discussed in more detail when the Kalman filter is introduced in Section 4.4.3.

An integral part of the track fit is the propagation of the track parameters from one detection plane to the next. Utilizing the simplifying assumptions made above (no interaction with the material) the components $P_4 = \tan\lambda$ and $P_5 = q/p_\mathrm{T}$ are constant and only the derivatives of y_1, z_1, and $\sin(\varphi_1)$ at $x = x_1$ with respect to y_0, z_0, $sin(\varphi_0)$ at $x = x_0$, $\tan\lambda$ and q/p_T have to be computed. The calculations can be found in [66]. In reality the propagation is much more complicated, because the following effects need to be included:

- The particles lose energy when they traverse matter through electromagnetic interactions. In addition, bremsstrahlung might not be negligible for high-energy electrons deflected by the magnetic field.
- Multiple scattering has a non deterministic impact on the flight direction of the particles.
- Elastic nuclear scattering can occur.
- The magnetic field is not entirely homogeneous and the components B_x and B_y are not exactly zero.
- The electrical fields present in some of the detectors affect the trajectories as well.

All these effects are considered in the propagation step in the ALICE reconstruction, some in a rather pragmatic way and some more rigorously.

Tracks can of course cross the borders between neighboring sectors. In that case the track parameters have to be rotated into the respective new local sector coordinate system. Due to the choice of the track parameters this is a simple rotation around the z-axis

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{4.15}$$

and the track momentum azimuthal angle becomes

$$\varphi' = \varphi - \alpha. \tag{4.16}$$

The other parameters z, $\tan\lambda$ and q/p_T remain constant under rotation. Also the covariance matrix (see next section) needs to be updated, but again only the entries associated with y and $\sin\varphi$ are affected. Figure 4.5b shows the global coordinate system and the local coordinate system exemplary for sector 7, which is rotated by $\alpha = 10° + 7 \cdot 20° = 150°$ with respect to the global coordinate system.

4.4.2. Track covariance matrix

The uncertainties and correlations of the track parameters are summarized in the covariance matrix of the tracks. The covariance between two variables X and Y is defined as [67]:

$$\mathrm{Cov}(X,Y) = \langle (X - \langle X\rangle)(Y - \langle Y\rangle)\rangle \overset{\text{Steiner}}{=} \langle XY\rangle - \langle X\rangle\langle Y\rangle, \tag{4.17}$$

where $\langle X\rangle$ denotes the expectation value for the variable X. The covariance matrix is a generalization of the variance:

$$\mathrm{Cov}(X,X) = \langle X^2\rangle - \langle X\rangle^2 = \mathrm{Var}(X) = \sigma_X^2 \tag{4.18}$$

and is furthermore symmetric: $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$.

The covariance is a measure of the correlation of X and Y:

$$\begin{aligned} \mathrm{Cov}(X,Y) > 0 \quad &\rightarrow \quad X,Y \text{ correlated} \\ \mathrm{Cov}(X,Y) = 0 \quad &\rightarrow \quad X,Y \text{ uncorrelated} \\ \mathrm{Cov}(X,Y) < 0 \quad &\rightarrow \quad X,Y \text{ anticorrelated} \end{aligned}$$

Since the correlation does not yield a specific quantitative measure for the dependence of two variables it is convenient to introduce the correlation coefficient [67]

$$\rho(X,Y) = \frac{\mathrm{Cov(X,Y)}}{\sigma_X \sigma_Y}, \tag{4.19}$$

which can take a value between -1 (full anticorrelation) and 1 (full correlation).

For the 5-dimensional track parameter vector $\vec{P}$ the covariance matrix $\mathbf{Cov}_\mathrm{trk}$ is a 5×5 matrix with the variances $\sigma_{P_i}^2$ of the track parameters on the diagonal and their

correlations σ_{P_i,P_j} stored in the off-diagonal elements:

$$\mathbf{Cov}_{\text{trk}} = \begin{pmatrix} \sigma_y^2 & & & & \\ \sigma_{z,\,y} & \sigma_z^2 & & & \\ \sigma_{\sin(\varphi),\,y} & \sigma_{\sin(\varphi),\,z} & \sigma^2_{\sin(\varphi)} & & \\ \sigma_{\tan(\lambda),\,y} & \sigma_{\tan(\lambda),\,z} & \sigma_{\tan(\lambda),\,\sin(\varphi)} & \sigma^2_{\tan(\lambda)} & \\ \sigma_{q/p_{\text{T}},\,y} & \sigma_{q/p_{\text{T}},\,z} & \sigma_{q/p_{\text{T}},\,\sin(\varphi)} & \sigma_{q/p_{\text{T}},\,\tan(\lambda)} & \sigma^2_{q/p_{\text{T}}} \end{pmatrix}. \tag{4.20}$$

By definition the covariance matrix is symmetric (hence only the lower triangle is depicted here for better visibility) and positive semi-definite.

4.4.3. Kalman filter

After the relevant coordinate systems, the track model, its covariance matrix and the propagation from one reference surface to the next were introduced, the track fitting employing the Kalman filter is described here. In its essence the Kalman filter [57] is a data processing algorithm which provides estimates for the states of dynamic systems via weighted averages of *measurements* and *predictions*. It has numerous historic applications (most famous is its use in the navigation computer of the Apollo space capsule [68]) and it is still widely used today, 60 years after its first publication, for example in global positioning system receivers or for smoothing the output from laptop trackpads [69]. Under the name progressive fit the Kalman filter entered into the high-energy physics community where it was first applied in the data analysis software of the DELPHI collaboration [70].

Extensive descriptions of the Kalman filter can be found for example in [65] and [71]. Its application for ALICE is described in detail in [72]. In the following the relevant formulas are introduced and put in the context of track fitting in ALICE. The state at time t, or, in the case of track fitting the track parameter $\vec{p}_x$ at the reference surface x, is assumed to evolve from the track parameter at the prior reference surface $\vec{p}_{x-1}$ via

$$\vec{p}_x = \mathbf{F}_x \vec{p}_{x-1} + \vec{w}_x, \tag{4.21}$$

where $\mathbf{F}_x$ is the state transition matrix which in this case propagates the track parameters from plane $x-1$ to plane x. The variable $\vec{w}_x$ contains the process noise term for each parameter of the state vector $\vec{p}_x$. Its mean is assumed to be zero $\langle \vec{w}_x \rangle = 0$ and its covariance matrix is denoted $\mathbf{Q}_x$. It contains random disturbances, mainly multiple scattering, and should not be confused with $\mathbf{Cov}_{\text{trk}}$. In the literature Equation (4.21) is referred to as system equation. Measurements $\vec{m}_x$ at plane x are given by the measurement equation:

$$\vec{m}_x = \mathbf{H}_x \vec{p}_x + \vec{v}_x, \tag{4.22}$$

where the matrix $\mathbf{H}_x$ transforms the state vector parameters into the measurement domain, because usually not all parameters are measured by one detector. For example the TRD tracking performs updates of the track parameters via measurements of y and z only. The measurement noise is given by $\vec{v}_x$ which is like the process noise assumed to be a zero mean Gaussian with covariance $\mathbf{R}_x$.

The true state $\vec{p}_x$ cannot directly be observed. The Kalman filter now provides an algorithm to determine an estimate $\hat{\vec{p}}_x$ by combining the *prediction* from models of the system and noisy *measurements* of certain parameters. In the following a notation is used where $\hat{\vec{p}}_{x|x-1}$ denotes the estimate of state $\vec{p}$ at the plane x including all measurements up to $x-1$. The estimates are optimal, if all noise is Gaussian, because in that case the Kalman filter minimizes the mean square error of the estimated parameters [57]. The prediction for a state and its covariance at plane x are now given by

$$\hat{\vec{p}}_{x|x-1} = \mathbf{F}_x \hat{\vec{p}}_{x-1|x-1} \quad \text{and} \tag{4.23}$$

$$\mathbf{Cov}_{x|x-1} = \mathbf{F}_x \mathbf{Cov}_{x-1|x-1} \mathbf{F}_x^\intercal + \mathbf{Q}_x, \tag{4.24}$$

respectively. The updated estimated state and covariance at x including the measurement $\vec{m}_x$ is now given by

$$\hat{\vec{p}}_{x|x} = \hat{\vec{p}}_{x|x-1} + \mathbf{K}_x(\vec{m}_x - \mathbf{H}_x \hat{\vec{p}}_{x|x-1}) \quad \text{and} \tag{4.25}$$

$$\mathbf{Cov}_{x|x} = \mathbf{Cov}_{x|x-1} - \mathbf{K}_x \mathbf{H}_x \mathbf{Cov}_{x|x-1}, \tag{4.26}$$

which are referred to as measurement update equations. They comprise the Kalman gain

$$\mathbf{K}_x = \mathbf{Cov}_{x|x-1} \mathbf{H}_x^\intercal (\mathbf{H}_x \mathbf{Cov}_{x|x-1} \mathbf{H}_x^\intercal + \mathbf{R}_x)^{-1}. \tag{4.27}$$

In the following these equations and their implications are discussed in more detail. For their derivation from first principle see for example [69] where a simple 1-dimensional problem statement is employed for a didactic explanation.

For the predictions only the state and its covariance at a prior plane are required. Furthermore the random disturbance contained in $\mathbf{Q}_x$ and of course the propagation formalism in $\mathbf{F}_x$ need to be provided. The filtered state $\hat{\vec{p}}_{x|x}$ is now given by adding the prediction and the residual of the measurement ($\equiv \vec{m}_x - \mathbf{H}_x \hat{\vec{p}}_{x|x-1}$) which is multiplied by the Kalman gain. The Kalman gain weights the prediction and the measurement based on their respective covariances and in addition can be seen as an amplifier which defines how strong the prediction is corrected due to the measurement [65].

Note that the states are described by probability density functions (pdfs). In the ideal case both are Gaussian pdfs where the given state is the mean of the distribution. To combine the pdfs they are multiplied and since the product of two Gaussian functions is again Gaussian the complexity of the Kalman filter does not grow. Again this is shown in detail in [69]. Furthermore the Kalman filter is an iterative procedure. In contrast to a global least squares fit a new measurement can be included without the need of re-evaluating all previous steps as they are contained in $\hat{\vec{p}}_{x-1|x-1}$ and $\mathbf{Cov}_{x-1|x-1}$. The working principle of the Kalman filter is illustrated in Figure 4.6.

The full information from the filter is only available at the last filtering step. Hence, the best parameter estimation at every point requires a smoothing step, where the filtered estimates are updated with the full information.

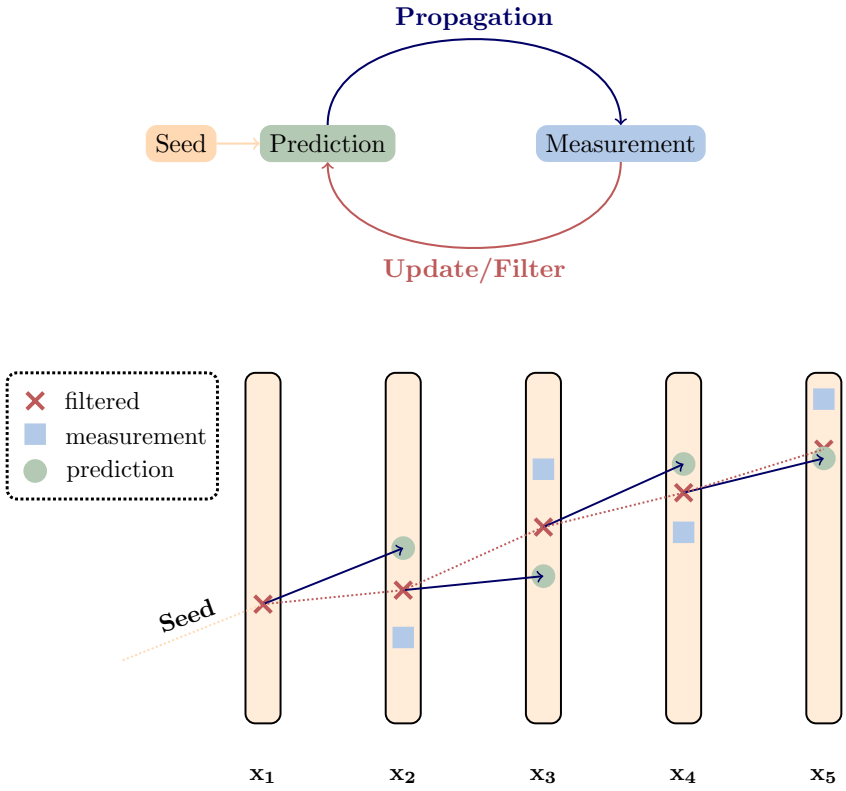


Figure 4.6.: Working principle of the Kalman filter. An initial prediction in the depicted case at x_1 is provided by the seed. It is propagated to the following layer x_2 where it yields a prediction. In the filtering step a weighted average between prediction and measurement is performed. The filtered state yields a prediction for the following layer and so on.

An important test for the filter is the χ^2. For a measurement at plane x it is calculated using the residual $r_{x|x-1} = \vec{m}_x - \mathbf{H}_x \hat{\vec{p}}_{x|x-1}$ and its covariance $\vec{v}_x = \mathbf{H}_x \mathbf{Cov}_{x|x-1} \mathbf{H}_x^\intercal$. In case only y and z are measured this results in the following calculation

$$\begin{aligned}\chi^2 &= \begin{pmatrix}\Delta y & \Delta z\end{pmatrix} \begin{pmatrix}\sigma_y^2 & \sigma_{yz} \\ \sigma_{yz} & \sigma_z^2\end{pmatrix}^{-1} \begin{pmatrix}\Delta y \\ \Delta z\end{pmatrix} \\ &= (\Delta y^2 \sigma_z^2 - 2\Delta y \Delta z \sigma_{yz} + \Delta z^2 \sigma_y^2)/(\sigma_y^2 \sigma_z^2 - \sigma_{yz}^2),\end{aligned} \tag{4.28}$$

where $\Delta y = y_p - y_m$ and $\Delta z = z_p - z_m$ denote the residuals in y and z, respectively, with p and m denoting the predicted and the measured value, respectively. The entries of the covariance matrix are simply the added values from the covariance matrix of the measurement and the prediction. This calculation can be done very fast, which is essential, because depending on the search road around the track multiple measurements might need to be compared and this comparison is done based on their χ^2 with respect to the track. In the following chapter the Kalman filter is applied to find the best matching TRD track points with respect to given tracks and update the track parameters with TRD information.

5. Online tracking with the TRD

The Transition Radiation Detector (TRD) was originally designed to provide electron identification and triggering capabilities for high transverse momentum processes [28]. Because ALICE will be operated in a continuous readout mode in the future, the hardware triggers of the TRD will become obsolete. Instead, the focus is shifted towards online tracking for the space point calibration of the TPC while the particle identification capabilities are to be maintained.

This chapter describes the online tracking procedure for the TRD which is developed in this work. The tracking algorithm uses ITS-TPC tracks prolongated into the TRD as seeds for a Kalman filter approach. It relies on track segments reconstructed online in the TRD Front-End Electronics (FEE), called tracklets. They have so far exclusively been used to generate trigger decisions [29] and are thus optimized for a fast standalone track finding procedure for tracks above $2\,\mathrm{GeV}/c$ [73]. As described in Section 2.2.3 the TRD readout will be limited to these tracklets instead of the full zero-suppressed ADC data to allow for a higher readout rate. In order to incorporate the tracklets into the global tracking, a detailed analysis of their reconstruction performance also at transverse momenta below $2\,\mathrm{GeV}/c$ is required. The resolution of the tracklets needs to be parameterized in order to perform a Kalman update of the track parameters.

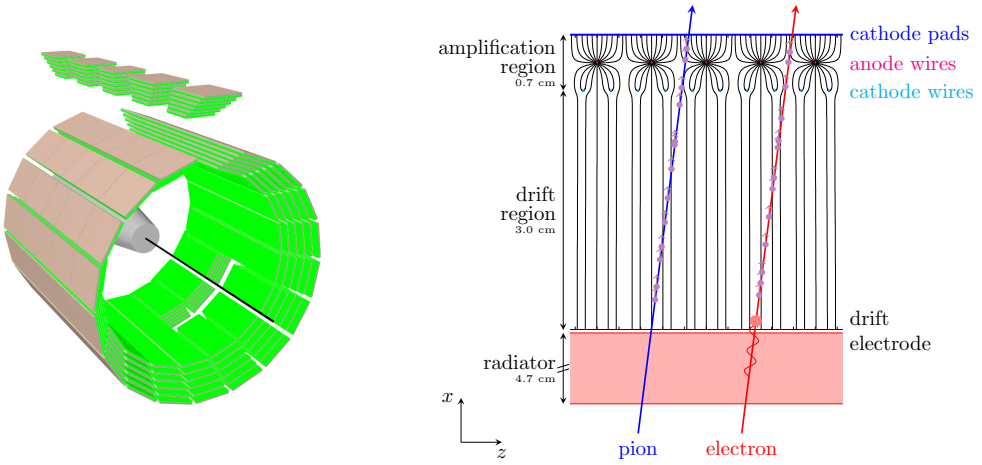
Before the newly developed tracking algorithm is presented, an overview on the TRD and its layout is given. Also the calculation of the online tracklets in the FEE is briefly described, in order to clarify their limitations.

The computing speed requirements necessitate different optimizations not only for the algorithm itself, but also for the involved data types. For example to enable one to utilize GPUs for the tracking procedure the utilized memory is managed manually. The explanation of the development process from the first prototype in the High-Level-Trigger (HLT) for Run 2 to the implementation in O^2 for Run 3 and beyond, including technical aspects of the execution on GPUs, concludes this chapter.

The tracking algorithm was presented at the *24th International Conference on Computing in High Energy & Nuclear Physics* in Adelaide and supplemental material has been made public in [74].

5.1. Detector overview

Transition Radiation (TR) is produced by charged particles crossing the boundary surface between two media with different refractive indices. The effect was predicted by Ginzburg and Frank in 1945 [75]. Since TR becomes relevant only for Lorentz factors $\gamma \geq 800$ and the Lorentz factor in turn is dependent on the particle mass via $\gamma = {}^{E}/_{mc^2}$, TR is produced



(a) Overall structure **(b)** Chamber cross section

Figure 5.1.: The overall structure of the TRD is shown on the left, taken from [73]. A hole is visible in the lower right, because the middle stacks for sectors 13, 14 and 15 were not installed to reduce the material budget in front of the PHOS detector. In total there are 522 individual chambers of which a cross section is shown on the right, taken from [29].

exclusively by electrons at momenta relevant for ALICE. Therefore, TR can be exploited to add additional electron/pion separation from 1 GeV/c up to high momenta where separation via $\mathrm{d}E/\mathrm{d}x$ or time-of-flight measurements is difficult. Since the probability to produce TR at a single boundary crossing is of the order of the fine structure constant $\alpha = 1/137$ and therefore rather low, detectors typically employ radiators composed of a stack of $\mathcal{O}(100)$ thin foils or unstructured material where many boundaries are crossed by traversing particles.

The ALICE TRD [29] consists of 522 chambers arranged in 18 sectors in azimuth, five stacks in longitudinal direction and six layers in radial direction, 2.90 m to 3.68 m away from the interaction point. It covers the full azimuth and the pseudorapidity range $|\eta| < 0.84$. The geometry of the detector is depicted in Figure 5.1a. Each chamber employs a radiator made out of Rohacell foam and polypropylene fibre mats. The TR photons which are in the X-ray regime and the charged particles traversing the TRD are detected by Multi-Wire Proportional Chambers (MWPCs) filled with Xe-CO_2 (85-15). A cross section of a TRD chamber is shown in Figure 5.1b. A 3 cm long drift region precedes the MWPC in the same gas volume and enables the measurement of not only a space point, but also the direction of the traversing particles. Furthermore, transition radiation which is produced by traversing electrons inside the radiator is absorbed at the beginning of the drift region. The extracted temporal information allows one to decouple the transition radiation signal from the specific ionization energy loss of the particle inside the chamber. The drift region is delimited from the 0.7 cm long amplification region by a cathode wire plane. The applied drift field of 700 V/cm accounts for a drift velocity of 1.56 cm/µs at nominal operating conditions. The signal is read out at the cathode pads

with sizes of $(0.515 - 0.965) \times (7.5 - 9.0)$ cm^2 ($r\varphi \times z$). The segmentation in $r\varphi$ enables charge sharing over typically three adjacent pads. In z direction the pads are tilted by $\pm 2°$ with the sign alternating from layer to layer, to improve the resolution in z during global tracking with only marginal deterioration in $r\varphi$. A chamber has 144 pads in $r\varphi$ direction and 16 pads along z direction (12 for the central stack 2). The total number of readout pads is $1.15 \cdot 10^6$. The active area of a chamber varies between $0.9 \times 1.06\,\mathrm{m}^2$ and $1.13 \times 1.43\,\mathrm{m}^2$ with their size increasing with larger distance to the interaction region in r and $|z|$. In longitudinal direction the gap between the readout pads of adjacent chambers from different stacks is 4 cm wide. The gaps in $r\varphi$ direction between adjacent sectors are given by the difference between the position of the outermost pad and the outer border at $R_{\mathrm{layer}} \cdot \tan(10°)$. The dead zone in $r\varphi$ varies from about 14% in the innermost layer to 12% in the outermost layer, while it is on average 2.5% in z. All in all the active area of the TRD is about 85% of the total detector surface.

5.2. Processing in the Front-End Electronics

Originally, the TRD was also built to add a level-1 trigger decision to ALICE. During Run 2 the latency of the level-1 trigger was 8.2 µs with respect to the interaction, much less than the drift time of the TPC of about 100 µs. Therefore, events which were not of interest physics-wise could be discarded very early before the readout of the TPC was finished. Hence, the limited bandwidth was saved for events which contained a specific physics signature, e.g. high momentum electrons or jets. In order to arrive at a trigger decision within the given narrow time window the data processing of the TRD makes use of parallel computing as much as possible and is divided into two stages. The first processing stage is performed in the FEE of each chamber where local track segments are calculated simultaneously. During Run 1 and 2 the tracklets were sent to the Global Tracking Unit (GTU) which in the second stage combined the tracklets of the six layers of a given stack to form TRD-only tracks [73, 76]. Based on these tracks trigger decisions were generated and sent to the Central Trigger Processor (CTP) of ALICE. In Run 3 the GTU will be decommissioned. The local tracking in the FEE will change marginally and the tracklets will be sent to the FLPs which forward them to the EPNs. At the EPNs the subsequent matching of the tracklets with global tracks from ITS and TPC will be performed.

5.2.1. The online tracklets

The first step of the TRD data processing is performed by the FEE consisting of Multi-Chip Modules (MCMs) mounted directly on top of the readout chambers. Each MCM is connected to 18 pads from a single pad row and shares in addition the edge channels between adjacent MCMs to avoid inefficiencies at the boundaries. Two custom chips are integrated in the MCMs, a PreAmplifier ShAper (PASA) and a TRAcklet Processor (TRAP), respectively. The PASA amplifies the very small positive signal induced on the readout pads corresponding to about 10^5 electrons [77] and converts it to a voltage on a pedestal value. Subsequently, the voltage is converted into a digital signal by 10-bit

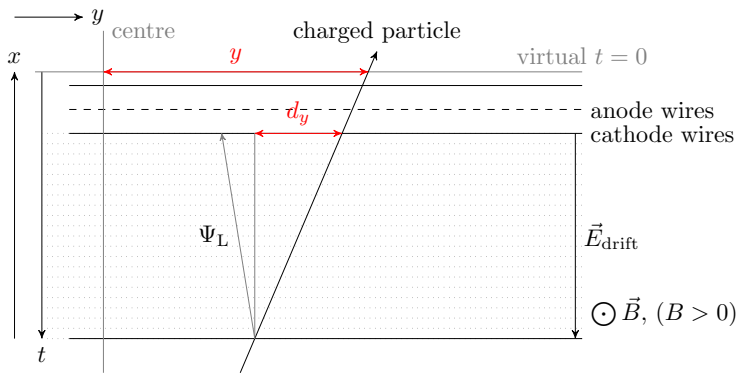


Figure 5.2.: Online tracklet reconstruction in the FEE. The deflection d_y over the drift length of 3 cm and the offset y with respect to the chamber center are encoded in the tracklet word. The z axis points in the same direction as $\vec{B}$. Figure taken from [29].

Analog to Digital Converters (ADCs) at a sampling frequency of 10 MHz, i.e. in time bins of 100 ns. At a nominal drift time of about 2 µs this corresponds in the ideal case to 20 measuring points for a particle traversing the detector. The ADC outputs are first extended by two binary digits to avoid rounding errors. Afterwards the signal is fed into a configurable digital filter chain. During production data taking in Run 1 and 2 two filters were active: the pedestal filter equilibrates the pedestal of the signal to a configurable value and the gain filter accounts for gain variations from pad to pad caused by mechanical imperfections or by the electronics themselves. All available filters are listed in [78]. The gain variations are measured in dedicated runs typically at the beginning of a new year. Metastable Krypton is injected into the gas system and the decay signal is used to calibrate the gain of each pad [79].

The next step is performed by a hardware preprocessor which searches timebin-wise for clusters as the charge values Q_i arrive. The index i corresponds to the pad column. Two conditions must be fulfilled for a cluster in a given channel:

1) it must be a local maximum : $Q_{i-1} \leq Q_i > Q_{i+1}$ and

2) a configurable charge threshold Q_{thr} must be reached: $Q_{i-1} + Q_i + Q_{i+1} \geq Q_{\text{thr}}$.

The center of gravity of the induced signal on three adjacent pads is calculated for each cluster and corrected with a value from a look-up table based on the pad response function.

Sums for a linear fit and two charge sums are accumulated for each pad timebin-wise based on the cluster properties. It should be noted that all calculations start as soon as the data arrive, so the cluster finding and the accumulation of the fit sums begin with the first time bin (at the outer radius of the chamber inside the amplification region) while the electrons from the inner radius of the drift region are still drifting. After the fit sums are accumulated up to four tracklet candidates are identified either on single pads or on two neighboring pads. A tracklet candidate has to fulfill two conditions: the number

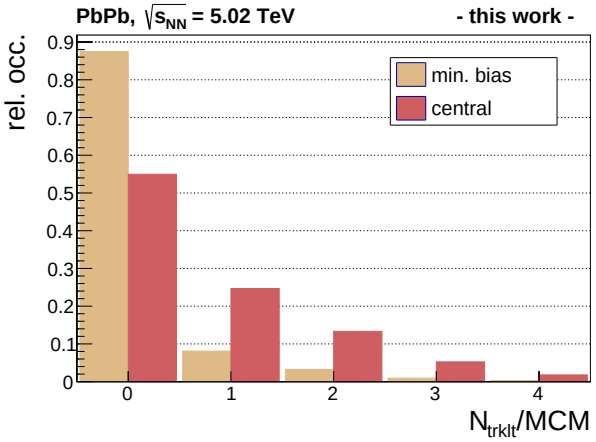


Figure 5.3.: The average number of tracklets per MCM is plotted for both minimum bias Pb–Pb and most central events. Even for central events only 1.8% of the MCMs send four tracklets.

of clusters on a single pad N_i must exceed a threshold and the accumulated number of clusters on the pad and its neighbor must exceed another threshold:

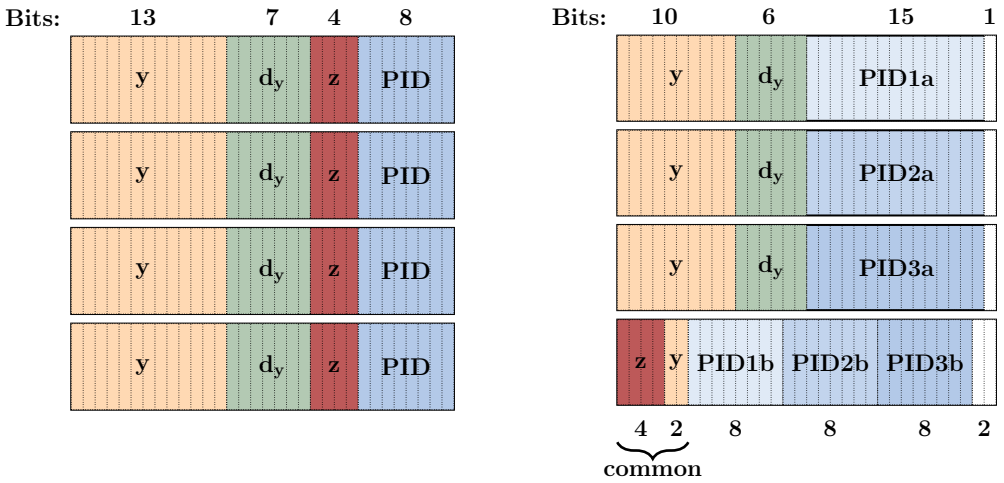
$$N_i \geq N_{\mathrm{CL}}^{\mathrm{thr}} \qquad \text{and} \qquad N_i + N_{i+1} \geq N_{\mathrm{CT}}^{\mathrm{thr}}. \tag{5.1}$$

Both thresholds $N_{\mathrm{CL}}^{\mathrm{thr}}$ and $N_{\mathrm{CT}}^{\mathrm{thr}}$ are configurable. In case of more than four tracklet candidates those with the highest number of clusters are chosen. In the final step four CPUs comprised in the TRAP calculate the tracklets individually: The fit sums of the i-th channel are merged with its neighbor and the transverse offset y in units of a pad width and the slope b in units of pads per timebin are calculated. The y position is given at a virtual radial position corresponding to the first timebin, slightly above the anode wire plane. The slope b is translated into a deflection d_y over the drift length $l_{\mathrm{drift}} = 3\,\mathrm{cm}$. Both fit parameters are shown in Figure 5.2. The deflection d_y has to be corrected in the FEE for two effects. Since the chambers reside in a magnetic field perpendicular to the drift field, the electrons are not drifting towards the cathode wire plane on the shortest path, but on a path tilted by the Lorentz angle Ψ_{L}. While clusters close to the anode wire plane are hardly affected, clusters at the inner chamber radius will reach the pad rows shifted by an offset of

$$d_{\mathrm{Lorentz}} = -\tan(\Psi_{\mathrm{L}}) \cdot l_{\mathrm{drift}}. \tag{5.2}$$

Furthermore the tilted pads are corrected with the help of a primary vertex assumption [76]. The result is another additive constant d_{tilt} that can be precalculated for each MCM. The sum of the two corrections $d_{\mathrm{Lorentz}} + d_{\mathrm{tilt}}$ is added to the measured deflection for each tracklet.

The longitudinal z position is given by the pad row to which the MCM is connected. The two charge sums are translated into a PID value with the help of a configurable lookup-table. For each tracklet the position information is encoded together with the PID value into a 32-bit word.

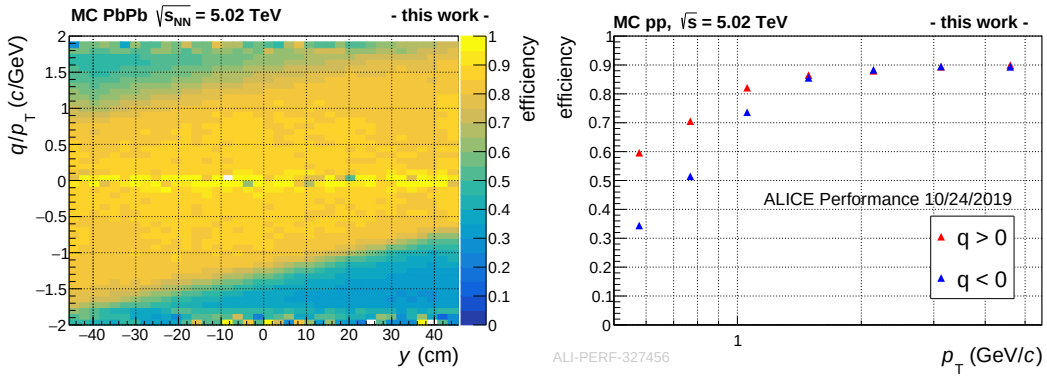


(a) Run 1 and 2 **(b)** Run 3

Figure 5.4.: Comparison of the old data format with up to four tracklet words send per MCM **(a)** and the new data format limited to three tracklets per MCM with 23 bits available for the PID information for each tracklet **(b)**. The bits are ordered from MSB on the left to LSB on the right. The white fields correspond to markers used to distinguish between a tracklet (bit set to `1`), the common information and extra PID information (bits set to `10`) and a tracklet end marker (`0x10001000`).

In the tracklet-only readout mode employed in Run 3 the available bandwidth is limited to four tracklets per trigger signal, i.e. 128 bits. At the time the TRD was designed rapidity densities of charged particles up to $\mathrm{d}N_{ch}/\mathrm{d}y \sim 8000$ where estimated in central Pb–Pb collisions [28] which led to the integration of four CPUs in the TRAPs. Since the actual track multiplicity is much lower than anticipated, namely $\langle \mathrm{d}N_{ch}/\mathrm{d}y \rangle \approx 2000$ for central Pb–Pb collisions at the full LHC collision energy of $\sqrt{s_{\mathrm{NN}}} = 5.5\,\mathrm{TeV}$ [14], the probability to have more than three tracklets in a single MCM is very low, as it is shown in Figure 5.3.

Therefore, the readout can be limited to three tracklets only with more space inside each tracklet word for information on the deposited charge. The PID capabilities of the TRD depend on this information [80]. The exact data format is still under study, but the information required for the tracking is fixed and will in principle be identical to what was available before. Slight optimizations are implemented to avoid sending redundant information: the z coordinate, which occupied 4 bits, is the same for all tracklets coming from a single MCM, because it is determined by the pad row the MCM is connected to. Also the y coordinate can be stored with a reduced number of bits with competitive precision. While y was stored in 13 bits with a granularity of $\frac{1}{160\,\mu\mathrm{m}}$, i.e. the range covered the full chamber width, the natural unit of pad widths will be used in the future and a common offset with respect to the center of the chamber will be added to each tracklet from the same MCM. The old and new data format are compared in Figure 5.4.



(a) **(b)**

Figure 5.5.: The left plot shows the reconstruction efficiency for online tracklets as a function of transverse momentum and position. A slight asymmetry is visible due to the Lorentz angle. The right plot shows the reconstruction efficiency as a function of transverse momemtum for particles traversing the active volume of the TRD.

In the past the information relevant to the GTU was the absolute position of a track segment with respect to the TRD detector stack in order to reconstruct global tracks. Therefore, it was beneficial to encode the absolute transverse position directly in the tracklets to reduce the processing time in the GTU. For Run 3, this additional complexity can be handled by the CRU or the FLP which converts the packed tracklet data into a 64-bit data structure. This data structure contains the same position information as the tracklets in Run 1 and 2, however with more information on the PID and the information in which chamber the tracklet was reconstructed. Hence, the online tracklets from Run 1 and 2 can be used for the development of the tracking algorithm for Run 3.

An important constraint arises from the fact that tracklets are found on maximum two neighboring pads. This implies a cut on the maximum deflection over the drift length of a chamber which in turn leads to a position dependent lower bound in p_{T} for primary particles. The effect is illustrated in Figure 5.5a. The reconstruction efficiency of the online tracklets as a function of the transverse momentum is shown in Figure 5.5b separately for positively and negatively charged tracks. Above $1.5\,\mathrm{GeV}/c$ the efficiency is approximately flat at 90%. For lower p_{T} the efficiency decreases. The fact that this effect is more strongly pronounced for negatively charged tracks is an artifact of the tracklet fit in the FEE.

The deflection d_y corresponds to the slope of the tracklet in the transverse plane, normalized to the length of the drift region $l_{\mathrm{drift}} = 3\,\mathrm{cm}$. Hence, it can directly be converted into the inclination in azimuth via

$$\varphi = \arctan\left(\frac{d_y}{l_{\mathrm{drift}}}\right). \tag{5.3}$$

It was observed before that the shift of the tracklet deflection with respect to the Monte Carlo truth is not symmetric [73]. The shift was understood as a combination of the

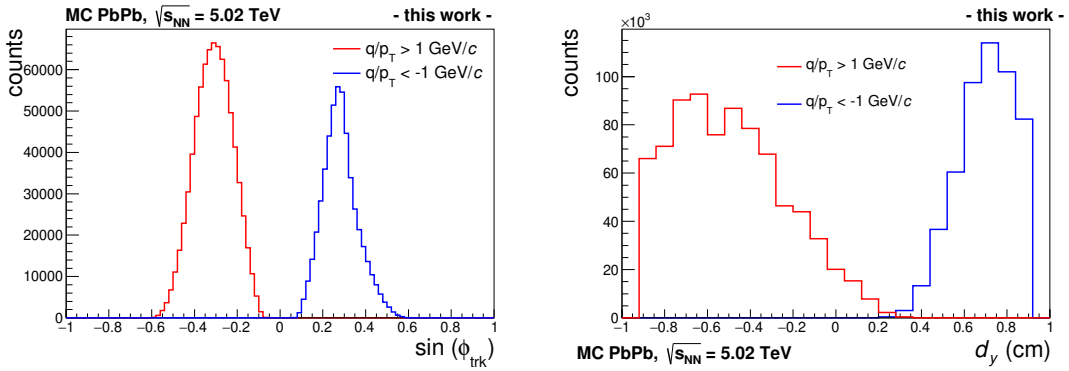
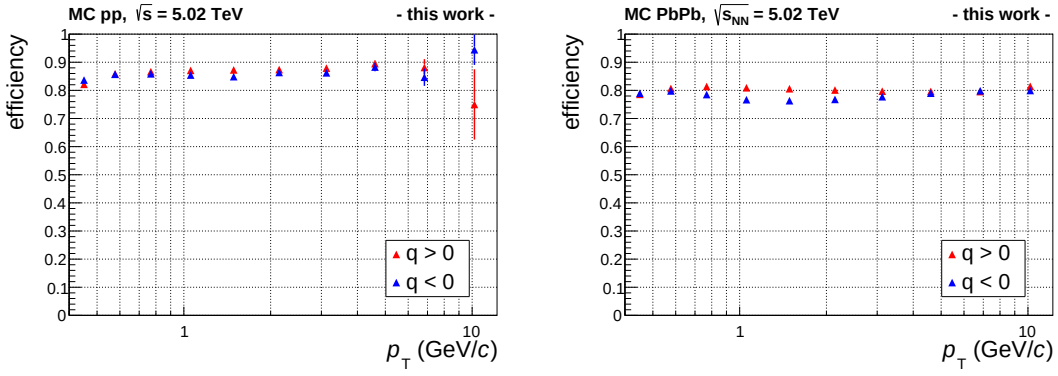


Figure 5.6.: The left plot shows the distribution of the azimuthal track angle in the TRD for primary-like tracks with $p_T < 1\,\text{GeV}$. The right plot shows the distribution of the deflection of the online tracklets which belong to the tracks. The tracklet deflection should be symmetric as well.

Lorentz drift and ion tails. Although the Lorentz drift is corrected via an additive constant in the FEE (see Equation (5.2)), the ion tails bias the reconstruction towards smaller values. This in turn leads to an overcorrection of the Lorentz drift. This was not a concern for the TRD triggers, as the affected range $p_T \lesssim 1.5\,\text{GeV}/c$ is not in their range of interest.

Below around $1\,\text{GeV}/c$ this effect becomes significant. Figure 5.6a shows the azimuthal angle $\sin\varphi$ determined by the ITS and the TPC for primary-like tracks with $p_T < 1\,\text{GeV}/c$, separately for positive and negative tracks. The distributions are symmetric. The absolute number of positive tracks is higher due to knockout from the detector material. Figure 5.6b shows the deflection of the tracklets which belong to these tracks. Tracklets from negatively charged tracks are biased towards larger deflections. Thus they exceed the valid range for the deflection from -8.96 to $8.96\,\text{mm}$ for lower track angles than the tracklets from positively charged tracks. A new calibration procedure is currently under development to mitigate this effect.

However, even with the new calibration the efficiency will still decrease below $1\,\text{GeV}/c$, because of the limitation to two neighboring pads. This is not ideal for the space point calibration of the TPC, because as it is shown in Figure 3.5b tracks with variable inclination are required in order to disentangle radial and $r\varphi$ distortions. Studies are currently ongoing whether the tracklet calculation in the MCMs can be extended to three neighboring pads. Obviously this would increase the number of tracklet candidates for each MCM and the possible efficiency loss in central Pb–Pb collisions needs to be reevaluated.



(a) pp **(b)** Pb–Pb

Figure 5.7.: Offline tracklet efficiency for pp collisions (left) and Pb–Pb collisions (right). The efficiency is defined as the fraction of reconstructed offline tracklets for particles which produce a hit in the detector.

5.2.2. Comparison to offline tracklets

The offline TRD reconstruction in Run 1 and 2 is based on the raw detector signal [29]. After clusterization seeding tracks from the TPC are propagated through the TRD. In each layer clusters are assigned to the tracks and from all assigned clusters offline tracklets are built. The efficiency for finding an offline tracklet if a track traverses the active detector volume and if it produces a hit in the detector is shown in Figure 5.7. In contrast to the efficiency for the online tracklets depicted in Figure 5.5b the offline tracklet efficiency is almost flat over the full p_{T} range. This is because the offline tracklets are not limited to a fixed number of neighboring pads. Also there is almost no dependence on the charge, since the correction of the Lorentz angle, which is performed offline, is optimized for the full p_{T} range.

Comparing the absolute efficiencies is not straightforward, because the offline tracklet efficiency comprises the tracking efficiency of the TPC and the matching efficiency. The online tracklets on the other hand depend solely on information from the TRD itself. Nevertheless, the comparison is required to determine the performance of the new tracking algorithm, which is entirely tracklet-based. Since the efficiency of the online tracklets deteriorates significantly below 1 GeV/c, the matching between tracks reconstructed in the inner detectors and TRD online tracklets will also deteriorate below 1 GeV/c.

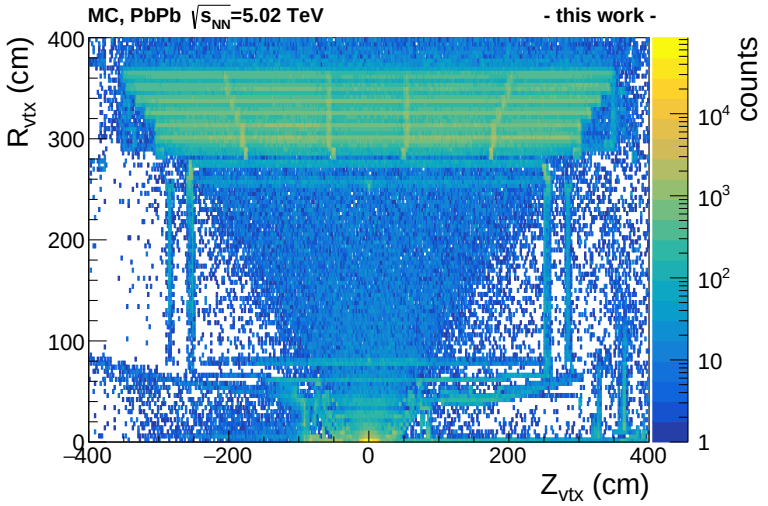


Figure 5.8.: Production vertices of particles generating a tracklet in the TRD. A large fraction of the tracklets (about 90%) is attributed to tracks originating from conversion processes in the material or from knockout.

5.3. Tracking algorithm

The Run 3 TRD tracking is based on tracks reconstructed in the inner detectors ITS and TPC. These tracks serve as seeds for the Kalman filter algorithm. Prolongated into the TRD they provide a predicted position to which the reconstructed online tracklets can be compared and thus matching tracklets can be identified. A standalone tracking procedure using only TRD information is much more difficult, as the support material between TPC and TRD and the material budget of the TRD itself lead to a large amount of conversion and knockout particles. From MC studies the fraction of TRD online tracklets belonging to primary-like particles, which are produced within a radial distance from the interaction point $R_{\text{vtx}} < 2\,\text{cm}$ and $|z_{\text{vtx}}| < 3\,\text{cm}$ in the longitudinal direction, is estimated to be as low as about 6%. This is illustrated in Figure 5.8 where the production vertex for particles creating a tracklet in the TRD is shown. In fact this was an issue for the TRD triggers during Run 1 where a large background was attributed to conversions at large radii [81].

As it was shown in Section 4.4.3, the track fitting in the TRD requires the knowledge of

a) a mathematical track model to approximate the particle trajectories,

b) a characterization of the track propagation through the detector taking into account the inhomogeneity of the magnetic field, the interaction with the detector material, etc.,

c) the detector characteristics, i.e. its geometrical layout including possible misalignment

d) and the tracklet coordinates in the sector coordinate system including their uncertainties.

The track model was introduced in Section 4.4.1. To recapitulate: the trajectories are described by helices which are defined by five parameters at a reference surface. The reference surfaces are parallel to the readout pad planes of the TRD. The parameters are the y and z coordinate for the position, the track inclination in azimuth $\sin\varphi$, the dip angle $\tan\lambda$ and the inverse of the transverse momentum multiplied by the charge q/p_{T}. The track propagation is described in [72]. Since the TRD tracking algorithm uses the general propagation functions provided by AliRoot, the HLT framework and O^2, respectively, they are not discussed in detail here. The propagation incorporates effects from multiple scattering, energy loss and inhomogeneities of the magnetic field. In the following the treatment of misalignment and the coordinate transformation is discussed.

5.3.1. Conversion of tracklets to space points

The TRD tracklets provide measurements in y, z and the deflection $d_y \propto \sin\varphi$ at the nominal radii of the pad planes possibly with a small offset due to the virtual time t_0 at which the straight line fit is evaluated (see above Figure 5.2). The tracklets do not contain information on either possible misalignment or on the quality of the straight line fit. The global alignment in ALICE is based on the millepede algorithm [82]. The millepede algorithm is a minimization procedure of the measurement residuals for each track. This minimization is done not only with respect to the global parameters representing alignment and calibration, but also with respect to the parameters of each track. Global tracks from collision and from cosmic data are employed to determine the 6 parameters for each alignable volume (3 translations + 3 rotations). For the TRD each chamber represents an alignable volume in the reconstruction code[1]. The last global alignment measurement was performed at the beginning of Run 2 in 2015 and the obtained parameters were used throughout Run 2. At the beginning of Run 3 again an alignment campaign is required after all detectors are installed in their final positions.

In AliRoot the `TGeoHMatrix` class is used for transformations from the chamber coordinates contained in the tracklets $\vec{x}_{\mathrm{det}} = (x_{\mathrm{det}}, y_{\mathrm{det}}, z_{\mathrm{det}})$ into the sector tracking coordinates $\vec{x}_{\mathrm{loc}} = (x_{\mathrm{loc}}, y_{\mathrm{loc}}, z_{\mathrm{loc}})$:

$$\vec{x}_{\mathrm{loc}} = \vec{T}_i + \mathbf{R}_i \vec{x}_{\mathrm{det}}. \tag{5.4}$$

$\vec{T}_i$ and $\mathbf{R}_i$ represent the translation vector and rotation matrix, respectively, for chamber i. Without misalignment $\mathbf{R}_i \equiv \mathbf{I}$, so only a translation is performed. The chamber coordinates are given by

$$\vec{x}_{\mathrm{det}} = \begin{pmatrix} x_{\mathrm{anode}} \\ y_{\mathrm{trklt}} \\ z_{\mathrm{trklt}} \end{pmatrix}, \tag{5.5}$$

[1] There are in total about 27k alignable volumes for the full ALICE geometry.

with the fixed radial offset $x_{\text{anode}} = 2.22\,\text{cm}$ and the z position equal to the center of the pad row on which the tracklet was reconstructed. The measurement in y does not require additional transformations.

Initially the measurement uncertainties of the tracklet positions are set to

$$\sigma_y = 1\,\text{mm} \quad \text{and} \tag{5.6}$$

$$\sigma_z = \frac{l_{\text{pad}}}{\sqrt{12}}, \tag{5.7}$$

where l_{pad} is the length of the pad in longitudinal direction. Due to the tilted pads the uncertainties in y and z are correlated ($\sigma_{yz} \neq 0$). Hence, the covariance matrix is given by

$$\begin{aligned} \mathbf{Cov}_{\text{TRD}} &= \mathbf{R} \begin{pmatrix} \sigma_y^2 & 0 \\ 0 & \sigma_z^2 \end{pmatrix} \mathbf{R}^\intercal \\ &= \cos^2(\beta_t) \begin{pmatrix} \sigma_y^2 + \tan^2(\beta_t)\sigma_z^2 & \tan(\beta_t)(\sigma_z^2 - \sigma_y^2) \\ \tan(\beta_t)(\sigma_z^2 - \sigma_y^2) & \sigma_z^2 + \tan^2(\beta_t)\sigma_y^2 \end{pmatrix}, \end{aligned} \tag{5.8}$$

where $\mathbf{R}$ is the rotation matrix accounting for the tilting angle β_t:

$$\mathbf{R} = \begin{pmatrix} \cos(\beta_t) & -\sin(\beta_t) \\ \sin(\beta_t) & \cos(\beta_t) \end{pmatrix}. \tag{5.9}$$

As will be shown in Section 5.3.3, the measurement uncertainties for the tracklets depend on the associated track. Therefore, the covariance matrix has to be recalculated for each tracklet on-the-fly during the matching procedure. At this point the tracklets are converted into space points which are defined in the same coordinate system as the tracks and which have default estimates for their accuracy. All space points together with their deflection d_y of a given event or time frame are stored in an internal data structure. In the following technically only these space points are used for the tracking algorithm, but they are still referred to as tracklets.

5.3.2. Matching procedure

In the same way as the space points the ITS-TPC matched tracks of the respective collision(s) are stored in an internal data structure. The parameters of these input tracks are defined at the inner TRD radius, i.e. at about 290 cm. The matching procedure described in the following is performed for each track individually. Therefore, parallelization can easily be achieved over the given tracks.

First, the track is propagated to the average radius of the first TRD layer at $r_0 \approx 300\,\text{cm}$. If it crossed the boundary between two sectors in this step, the track is rotated into the coordinate system of the new sector. Note that after the rotation another propagation is required, since the rotation naturally affects both x and y (see Equation (4.15)). This is repeated until the track lies inside a sector at the correct radius r_i for a given layer i, i.e. until

$$|y_{\text{trk}}| \leq y^i_{\text{max}} = r_i \cdot \tan(10°). \tag{5.10}$$

Due to rounding errors in rare cases a track can lie between two sectors such that endless rotation and successive propagation steps are performed. During the commissioning phase in the HLT this created occasional backpressure resulting in dropped events. Since these cases are so rare they were not observed during the development phase. In order to avoid this, tracks which do not lie in one sector after more than two rotation and propagation steps are discarded.

Now a search road in the yz plane is defined. In y it is based on the uncertainties of the track and the default tracklet uncertainty:

$$w_y = N_\sigma^y \sqrt{\sigma_{y,\mathrm{trk}}^2 + \sigma_{y,\mathrm{trklt}}^2} + w_y^{\mathrm{extra}}, \tag{5.11}$$

with a configurable extension w_y^{extra} which is set by default to 2 cm and $N_\sigma^y = 7$. In z it is simply defined as twice the longest pad length $l_{\mathrm{pad}}^{\mathrm{max}} = 9\,\mathrm{cm}$,

$$w_z = 2\ l_{\mathrm{pad}}^{\mathrm{max}}, \tag{5.12}$$

because the measurement uncertainty of the track $\sigma_{z,\mathrm{trk}}$ is much smaller compared to the measurement uncertainty of the tracklet $\sigma_{z,\mathrm{trklt}}$. The search roads are optimized such that there is no efficiency loss expected from the implied selection and at the same time they are as small as possible to avoid unnecessary comparisons between tracklets and tracks which would increase the computing time.

Based on the search road and on the position of the track, the chamber in which the track ends up and possibly neighboring chambers which need to be searched for tracklets are determined. At maximum four chambers are taken into account, if the track is close to the chamber boundaries both in y and z. For each chamber in the search road the track is propagated to the average radius of that chamber, which can deviate from r_i due to misalignment. Tracklets which are in the correct chamber but outside of the search road are ignored. For each remaining tracklet the track parameters are extrapolated linearly and therefore fast to the exact tracklet radius. Subsequently the compatibility of the tracklet with respect to the track can be determined after two corrections.

1. Tilted pad correction

As described above the readout pads of the TRD are tilted by $\beta_t = \pm 2^\circ$ with the sign alternating from layer to layer. The z position of the track is used to correct the y position of the tracklet:

$$y' = y + (z_{\mathrm{row}} - z_{\mathrm{trk}}) \cdot \tan(\beta_t) = y + \Delta z \cdot \tan(\beta_t). \tag{5.13}$$

This correction is only applied, if $\Delta z < l_{\mathrm{pad}}$ and if $\sigma_{z,\mathrm{trk}} < \frac{l_{\mathrm{pad}}}{\sqrt{12}}$. Otherwise the raw tracklet position is taken, because a correction would not be meaningful.

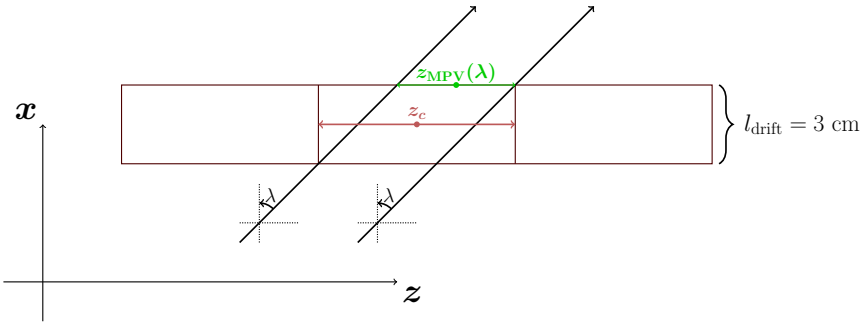
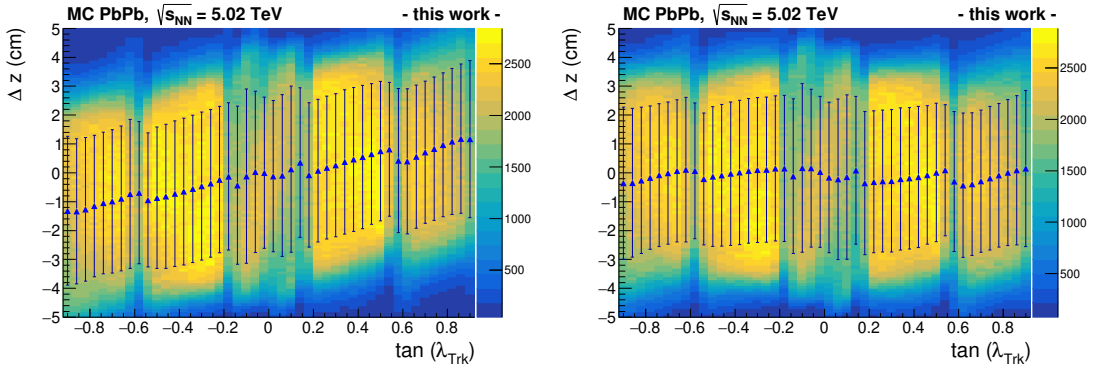


Figure 5.9.: The most probable value for the z position of the tracklet z_{MPV} depends on the dip angle λ and is not necessarily at the center of the pad.



(a) Before correction **(b)** After correction

Figure 5.10.: The mean z position is shifted from the center of the pad rows for tracklets which are found in the outer TRD stacks ($|\tan(\lambda_{\mathrm{trk}})| > 0.2$). The error bars correspond to the RMS of the distributions. The correction described in the text mitigates this effect.

2. Average z correction

The z coordinate of the tracklet is defined at the center of the pad row on which it is found. This is a good approximation for tracks with $\eta \simeq 0$, but it leads to a systematic shift for inclined tracks. This effect is illustrated in Figure 5.9. Therefore, a correction is added to the z_c position of the tracklets based on the dip angle λ of the track:

$$z_{\mathrm{MPV}}(\lambda_{\mathrm{trk}}) = z_c + f_z \cdot \tan(\lambda_{\mathrm{trk}}). \tag{5.14}$$

The correction factor $f_z = 1.4\,\mathrm{cm}$ is chosen, because it shifts the most probable value for Δz to zero as it is shown as a function of $\tan\lambda_{\mathrm{trk}}$ in Figure 5.10a and Figure 5.10b. Inserting $\frac{l_{\mathrm{drift}}}{2} = 1.5\,\mathrm{cm}$ leads to a larger bias for Δz.

Matching

If the corrected tracklet position still lies within the search road around the track, the χ^2 value for a possible update of the track with this tracklet can be calculated employing Equation (4.28). This requires the knowledge of the tracklet covariance which is described in detail in the following Section 5.3.3. The tracklet deflection does not enter into the χ^2 calculation. As it will be shown below the angular resolution of the tracklets is more than two orders of magnitude worse compared to the angular resolution of the ITS-TPC tracks, due to their long lever arm. As the χ^2 prediction weighs the measurements with their uncertainties, the impact of the tracklet angle is negligible. Nevertheless, the deflection can be used as an exclusion criterion to reduce the fake tracklet attachment.

All tracklets for which the χ^2 of a possible update is below a configurable threshold and for which the deflection does not deviate too much from the track inclination are considered for matching. In addition, the possibility of no available tracklet is considered with a fixed penalty χ^2_{penalty}. After the χ^2 is calculated for all available tracklets the update candidates are sorted according to their χ^2. If a best matching tracklet is found with

$$\chi^2 < \chi^2_{\text{penalty}} \tag{5.15}$$

the track is extrapolated to this tracklet and the track parameters are updated with the measurement in y and z according to Equation (4.25) and Equation (4.26). The estimated χ^2 is added to the global track χ^2. In case the track points into the dead zone in a given layer and no matching tracklet can be identified there is no penalty added to the total χ^2 of the track.

For each layer the best N_{hypo} possibilities based on their χ^2 are taken into account. This includes the possibility of no available update. An array of tracks is filled for each possible update and these tracks become the seeds for the matching algorithm in the successive layer where the matching procedure is repeated. To limit the increase in combinatorics from layer to layer only the best $N^{\text{max}}_{\text{hypo}}$ tracks are propagated further. In the last layer finally the best hypothesis is stored as resulting track.

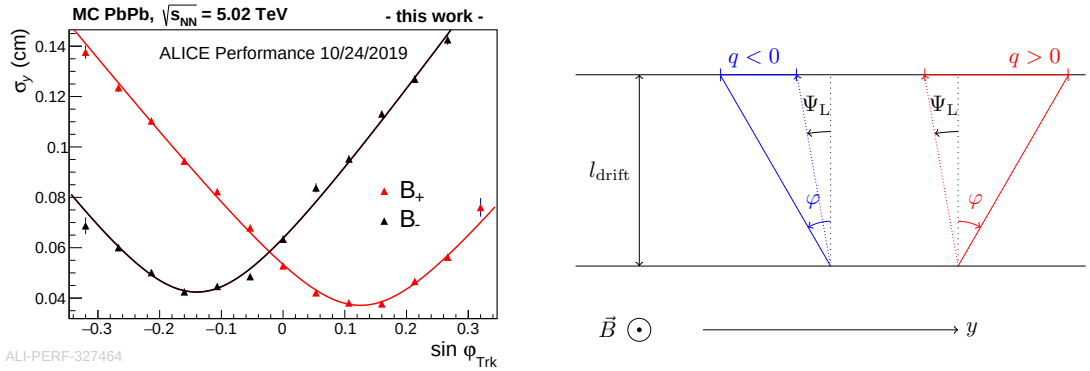
5.3.3. Tracklet covariance

Resolution in y

The TRD chambers are operated in a magnetic field perpendicular to the direction of the anode wires. Because of the $\vec{E} \times \vec{B}$ effect the resolution of the tracklets depends on the inclination of the corresponding track in the transverse plane. According to [52] the measurement variance for a track that traverses an MWPC at an angle φ in the transverse plane is given by:

$$\langle y^2 \rangle - \langle y \rangle^2 = \frac{\sigma^2}{\cos^2 \varphi} + \frac{b^2}{12} (\tan \varphi - \tan \Psi_{\text{L}})^2, \tag{5.16}$$

where the width σ accounts for diffusion and b is the wire pitch which is 5 mm for the TRD. Because the position is not determined by the sense wire, but by the cathode



(a) **(b)**

Figure 5.11.: Position resolution of the online tracklets in y (a). Due to the $\vec{E} \times \vec{B}$ effect the clusters of positively charged particles are spread over more pads on average, as indicated in (b). This effect is independent of the polarity of the magnetic field, as a change of the polarity would also change the sign of the Lorentz angle Ψ_L. The sketch corresponds to the B_+ case where the magnetic field points in the same direction as z_{glb} (away from the muon spectrometer).

pad plane, the actual variance is more complex. Furthermore, in case of the tracklets not the individual cluster position, but the offset of a straight line fit determines the y position. Nevertheless, it can be deduced from Equation (5.16) that the optimal precision is expected for tracks which traverse the chamber under the Lorentz angle ($\varphi_{trk} = \Psi_L$).

The effective position resolution for the tracklets is based on a parameterization which is obtained from simulations and confirmed in real data in Section 5.3.5. During the matching procedure the track positions at the matched tracklets are stored without updating the track parameters with the TRD information. For the inner layers $i = 1, 2, 3, 4$ the tracklet residuals r_y^i are calculated with respect to an interpolation from the surrounding layers:

$$r_y^i = \Delta y_i - \frac{\Delta y_{i-1} + \Delta y_{i+1}}{2}, \tag{5.17}$$

with $\Delta y_i = y_{trklt,i} - y_{trk,i}$. Thus, the uncertainties connected to the track extrapolation and to the track parameters themselves are mitigated. For collisions with a rather low track multiplicity the residuals are collected for all tracks which have a matching tracklet in each layer. In a higher multiplicity environment the position resolution of the tracklets deteriorates, as multiple tracks are more likely to contribute to the same tracklet [83]. The residuals can now be plotted with respect to the inclination angle of the tracks. A Gaussian fit is performed for the collected residuals in each $\sin\varphi_{trk}$ bin. The resulting resolution for both magnetic field polarities $B_z = \pm 0.5\,\mathrm{T}$ is plotted in Figure 5.11a. The optimal resolution of about 400 µm is obtained for tracks at an angle of about $\varphi_{trk} \approx \pm 8°$, which corresponds to the Lorentz angle for the applied magnetic field, gas mixture and drift field [29]. For strongly inclined tracks the resolution deteriorates to about 1.4 mm for tracks belonging to positively charged particles and to about 800 µm for tracks belonging to negatively charged particles. For negatively charge particles the Lorentz effect on average focuses the clusters along the wire direction while they are drifting towards the

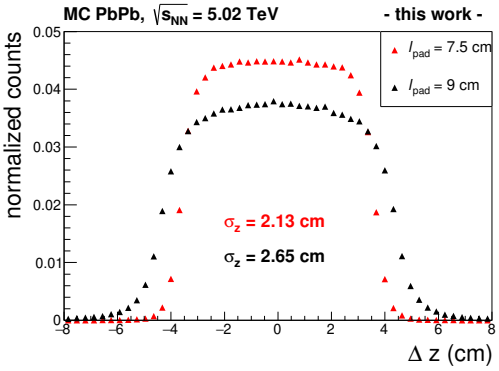


Figure 5.12.: The two histograms show the residuals of the tracklets with respect to matched tracks in z. The red points correspond to tracklets found on the inner pads which have a length of 7.5 cm and the black histogram corresponds to tracklets on the outer pads with a length of 9 cm. The resolution in z is given by the length of the pad on which the tracklet is reconstructed.

amplification region, while for positively charged particles it has the opposite effect (see Figure 5.11b).

As shown in Figure 5.11a the resulting resolution in y is well described by the following fit function

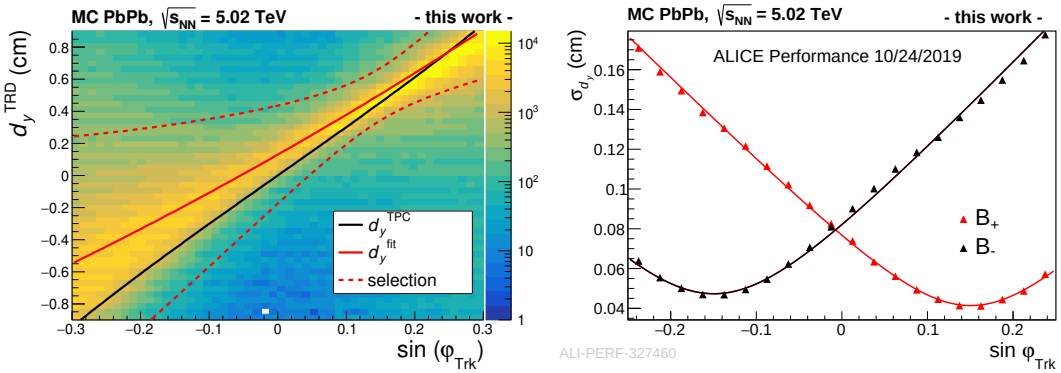
$$\sigma_y(\sin\varphi_{\mathrm{trk}}) = \sqrt{a_y^2 + c_y^2(\sin\varphi_{\mathrm{trk}} - b_y)^2}. \tag{5.18}$$

In this parameterization a_y corresponds to the optimal resolution σ_y^{min}, $b_y \equiv \sin\Psi_{\mathrm{L}}$ and c_y is a scaling factor. The fit parameters for all possible magnetic field polarities are stored in the configuration of the TRD tracking algorithm. Thus, for each track the expected resolution in y can be deduced from Equation (5.18) during the matching procedure.

Resolution in z

As explained above, the tracklet resolution in z is given by the length of the pad on which the tracklet is reconstructed (see Equation (5.7)). The residuals in z are plotted in Figure 5.12 separately for tracklets reconstructed on inner pads ($l_{\mathrm{pad}} = 7.5\,\mathrm{cm}$) and outer pads ($l_{\mathrm{pad}} = 9\,\mathrm{cm}$). In this case the uncertainty due to the track extrapolation is not corrected for, yielding a slightly wider distribution than $\frac{l_{\mathrm{pad}}}{\sqrt{12}}$.

Inclined tracks can induce tracklets on two neighboring pad rows, if they produce enough clusters on both pads. In the former offline reconstruction these tracklets could be identified during the tracklet reconstruction. The z coordinate for these tracklets could thus be determined more precisely. In the FEE row-crossing tracklets cannot be identified, as the MCMs are connected to pads from a single row only. During the matching procedure to ITS-TPC tracks the neighboring pad row could be searched for an additional matching tracklet. For the space point calibration of the TPC this is not required, but it could be an extension to the TRD tracking at the asynchronous stage to



(a) (b)

Figure 5.13.: The left plot shows the tracklet deflection d_y as a function of $\sin\varphi$ for the matched track. The direct conversion d_y^{TPC} does not describe the data as well as the second order polynomial parameterization d_y^{fit} described in the text. The right plot shows the resolution for the tracklet deflection which is proportional to the angular resolution of the TRD for both polarities of the magnetic field $B_{\pm} = \pm 0.5\,\mathrm{T}$.

improve the PID performance of the TRD. This is relevant because the tracklets from a row-crossing track carry only a fraction of the total deposited charge.

Resolution in d_y

Due to the asymmetry of the reconstructed tracklet deflection described above, the inclination measured with the TPC, φ_{trk}, cannot directly be converted into a deflection, d_y^{TPC}, via Equation (5.3). Figure 5.13a shows the deflection of the tracklets d_y^{TRD} as a function of the track inclination $\sin\varphi_{\mathrm{trk}}$. A parabolic fit function

$$d_y^{\mathrm{fit}} = a + b\sin\varphi_{\mathrm{trk}} + c(\sin\varphi_{\mathrm{trk}})^2 \tag{5.19}$$

describes the distribution much better. Also fake matches are included in the figure which can be reduced by the indicated selection on the angular pull which is described later.

The resolution of the deflection σ_{d_y} is given by the variance of a Gaussian fit of the residuals

$$r_{d_y} = d_y^{\mathrm{TRD}} - d_y^{\mathrm{fit}} \tag{5.20}$$

accumulated for a given track inclination. The results are plotted in Figure 5.13b, again for both magnetic field polarities. The same fit function as in Equation (5.18) is used to parameterize $\sigma_{d_y}(\sin\varphi_{\mathrm{trk}})$ and the fit parameters are stored in the tracking algorithm as well. The optimal resolution $\sigma_{d_y} \approx 400\,\mu\mathrm{m}$ corresponding to $\sigma_{\varphi} = 0.76°$ is much worse compared to the angular resolution of the extrapolated ITS-TPC tracks with their much longer lever arm. For them the resolution is $\langle\sigma_{\varphi}^{\mathrm{trk}}\rangle \approx 0.0034°$, i.e. about a factor of 200 better at the innermost TRD layer for $p_{\mathrm{T}} \approx 1$ to $2\,\mathrm{GeV}/c$.

Parameter	Default value
χ^2_{penalty}	12
χ^2_{max}	15
$N^{\text{max}}_{\text{hypotheses}}$	1
$p_{\text{T,min}}$	0.5 GeV/c
η_{max}	0.84
w_z	18 cm
w_y^{extra}	2 cm
x_{offset}	0.1 cm
$N^{\text{max}}_{\text{ly,missing}}$	6
N_σ^y	7
f_z	1.4

Table 5.1.: Overview of the configuration parameters for the TRD tracking algorithm. The default values are the same for pp and Pb–Pb collisions.

This is the reason for neglecting d_y in the calculation of χ^2 and when updating the track parameters. But still the deflection is used as an exclusion criterion for tracklets, as it is mentioned above. The selection is based on the pull P which is given by:

$$P = \frac{d_y^{\text{trklt}} - d_y^{\text{fit}}}{\sigma_{d_y}(\sin\varphi_{\text{trk}})}, \tag{5.21}$$

where the deflection of the track is given by the fit function Equation (5.19). The default selection (dashed line in Figure 5.13a) on the pull of the deflection is $|P| \leq 4$. Without the selection especially at larger angles some tracklets would be matched to the tracks whose deflection disagrees strongly with the inclination of the track. Most of these tracklets belong to different tracks which are close to the extrapolated seed by chance. Applying the selection on the angular pull thus enhances the purity of the matching algorithm by about 2% in Pb–Pb.

5.3.4. Configuration and quality assurance

All configuration parameters for the tracking algorithm are presented in Table 5.1. Not included in the table are the fit parameters which are required to calculate the covariance of the tracklets. A comprehensive list including the tracklet error parameterizations for the different magnetic field settings can be found in the appendix in Table A.1. At the moment the default number of hypotheses which are kept per layer is 1. The reason for introducing the possibility to keep more than one hypothesis per layer is to avoid an efficiency loss in case a fake tracklet attachment impacts the track parameters such that a correct tracklet in the subsequent layer is not found anymore. These cases are so rare that the additional computing time in the synchronous reconstruction phase, where the TRD tracking is done only for peripheral events with a low track multiplicity, is not

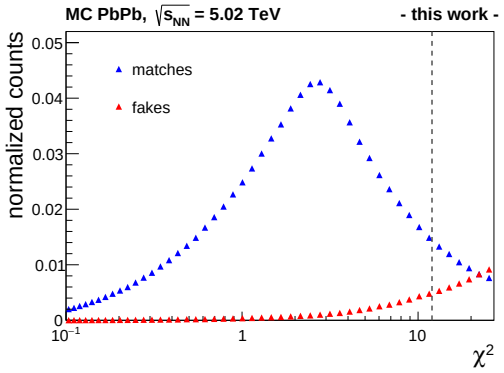


Figure 5.14.: The χ^2 distribution is shown both for matching and for fake tracklets. The curves are normalized to the total number of matches including fake matches. The dotted line is drawn at $\chi^2_{\text{penalty}} = 12$.

justified. However, in the asynchronous reconstruction phase it can improve the tracking efficiency especially for high multiplicity events.

The track following is abandoned in case no tracklet is found in $N^{\text{max}}_{\text{ly,missing}}$ consecutive layers. Only layers are counted in which the track points into the active area of the detector.

The selection on χ^2 is determined from Monte Carlo simulations and chosen such that the efficiency is as high as possible while the background is kept at a tolerable level. The χ^2 distributions for correctly and wrongly (based on their Monte Carlo labels) associated tracklets to tracks is shown in Figure 5.14. At $\chi^2 \approx 20$ the probability to attach a fake tracklet is higher than the probability to attach a matching tracklet. The most probable value for χ^2_{max} for a matching tracklet is about 2.5, slightly higher than the expected value of 2 for the two degrees of freedom, y and z. This is mainly caused by the non-Gaussian distribution of the tracklet residuals in z. The default value for χ^2_{max} is set to 15. Tracklets that are less compatible with the track parameters are not considered for matching. But since the penalty for not having a tracklet is set to $\chi^2_{\text{penalty}} = 12$ in case the best matching tracklet is not better than the applied penalty no update is performed in the given layer.

The configuration parameter x_{offset} accounts for the average radial shift for all TRD chambers. As mentioned above the y position of the tracklets is given as the offset of the straight line fit at the virtual t_0 position. The calibration parameter t_0 is determined for each chamber individually [84].

Next to t_0, also the drift velocity v_d, the Lorentz angle Ψ_{L} and the gain are calibrated on a per chamber basis. The chambers for which no satisfactory calibration can be obtained are masked for the data reconstruction in the final reconstruction passes and in the respective simulations [29]. Figure 5.15 shows the residuals in y for all chambers for Monte Carlo simulations. In the simulation the chambers are very well equilibrated as expected with a negligible dispersion. For real data the performance is shown in the following Section 5.3.5.

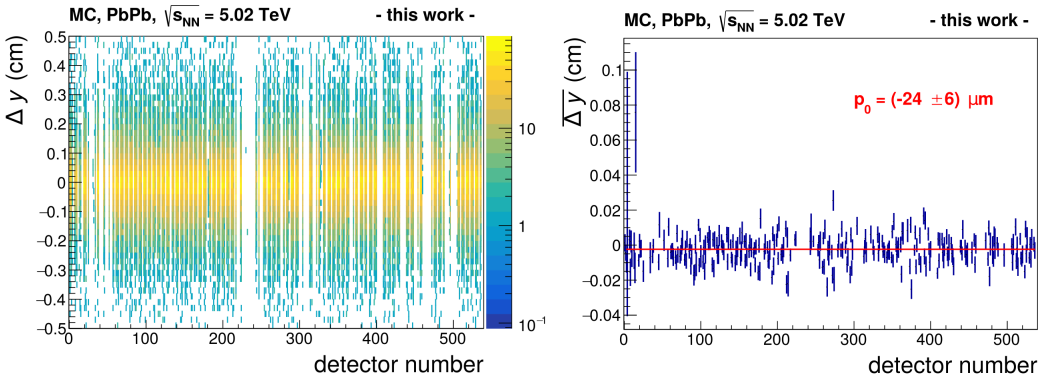


Figure 5.15.: The left plot shows the tracklet to track residuals in y as a function of the detector number. The residuals are determined as in Equation (5.17). The gaps correspond to the empty stacks in front of the PHOS spectrometer and in addition a few chambers have issues e.g. with the high-voltage (see Section 2.2.3) or with the FEE and thus send only little or no data. On the right the mean value for the residuals are shown which are close to zero with a very small dispersion. The two outlier detectors, 4 and 15, are surrounded by chambers which are either off or have a reduced drift field.

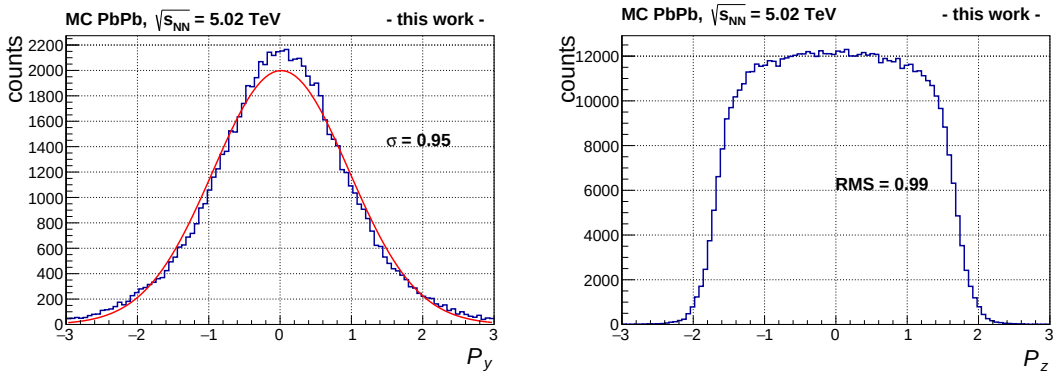


Figure 5.16.: Shown are the tracklet pulls for y on the left and for z on the right. Only tracklets with the correct Monte Carlo label are taken into account.

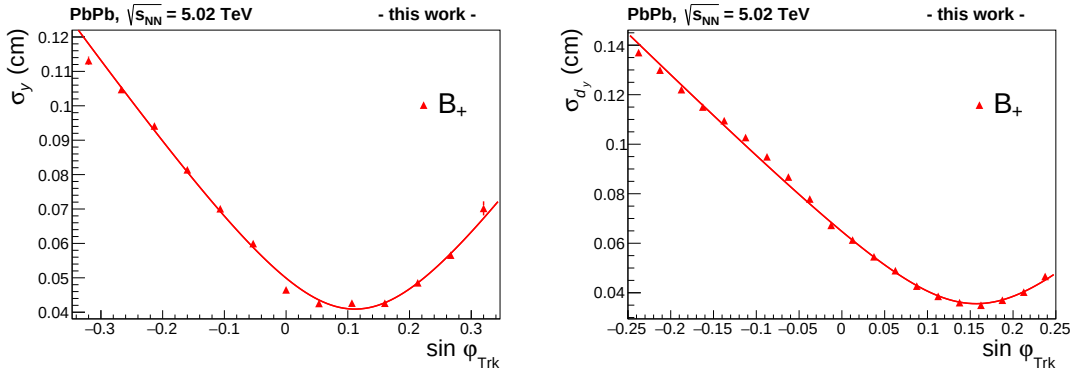


Figure 5.17.: Both position (left) and angular resolution (right) are shown for the online tracklets in real data. There is a very good agreement with the resolutions obtained in Monta Carlo simulations which are shown in Figure 5.11a and Figure 5.13b.

For quality assurance measures the pulls for the measurements in y and z should be monitored. If the resolutions were truly Gaussian shaped and correctly parameterized, the pulls, which are described by

$$P_y = \frac{y_{\text{trk}} - y_{\text{trklt}}}{\sigma_y^{\text{trk}} + \sigma_y^{\text{trklt}}} \tag{5.22}$$

and analogously for z, would again be a Gaussian with a width of 1. Larger (smaller) widths of the pulls indicate a too high (low) weight for the Kalman update and thus an underestimation (overestimation) of the measurement uncertainty. Figure 5.16 shows the pulls both in y and z for matching tracklets. In y the pulls are well described by a Gaussian function with a width close to 1. In z on the other hand the pulls have the same shape as the uncertainties in the tracklet position which clearly dominate the total uncertainty. Still the RMS for the distribution is very close to 1 confirming that the measurement uncertainties are neither under- nor overrated.

5.3.5. Comparing Monte Carlo and real data

For the development of a tracking algorithm the information contained in Monte Carlo simulations is very helpful. A label is assigned to each detector signal (hit) which allows for checks of for example whether a tracklet belongs to a given track or if it was created by a different particle or noise. This way the matching efficiency and especially the fake tracklet attachment can be easily determined. In the real data taking operation association between detector signals and particles is obviously not known a priori. However, the performance of the tracking algorithm needs to be determined in real data as well, because the simulation cannot describe all detector effects in a perfect way and in the end the tracking algorithm needs to handle real data.

Therefore, the tracklet resolutions in real data are shown in Figure 5.17. For real data there is of course no label information which indicates whether a tracklet belongs to

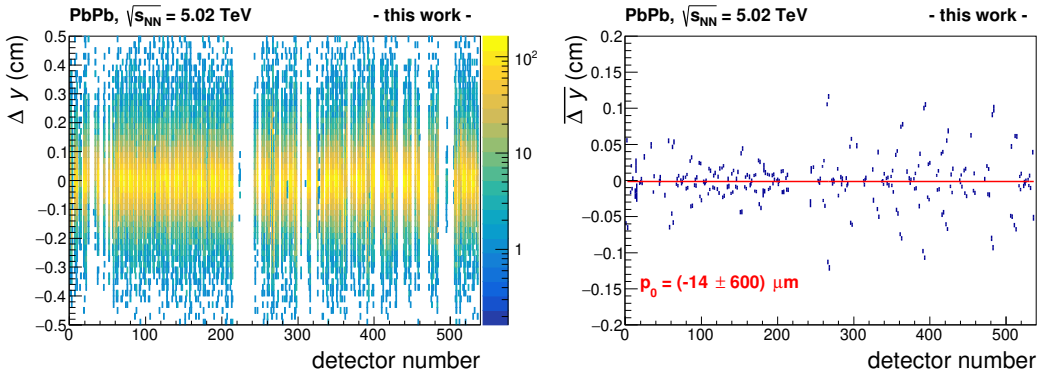


Figure 5.18.: The residuals in y as well as their mean are shown for each chamber individually as in Figure 5.15, but here for real data.

a track or not. By applying the same selection criteria as for Monte Carlo simulations, namely by considering only tracks with a tracklet in each layer and neglecting central collisions, a clean sample is obtained nevertheless. The resolutions in data and simulations agree very well.

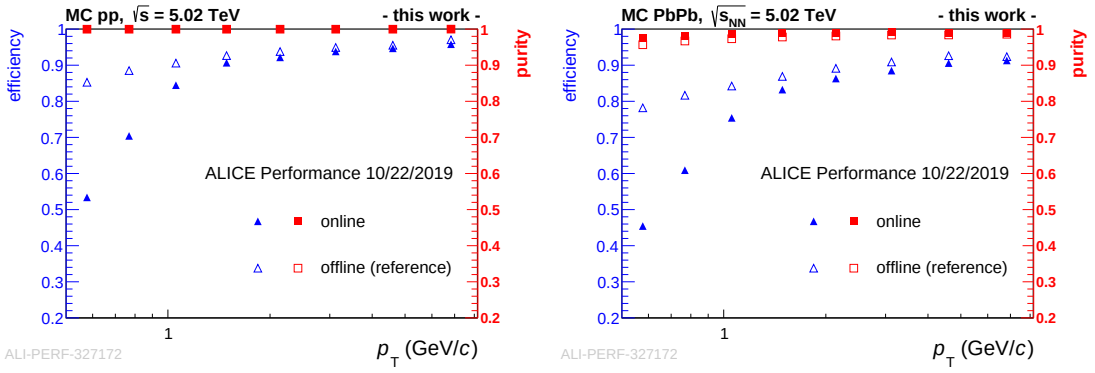
In a more differential check the residuals per chamber are depicted in Figure 5.18. As opposed to the simulation the dispersion for the individual chambers is much larger with a comparable magnitude to the position resolution of the tracklets, while the average value remains close to zero. A possible reason for the large dispersion is an inaccurate calibration of the chambers. The y position particularly depends on v_d and t_0, since both can introduce shifts in the radial direction. These shifts in r affect the measurement of y, especially for strongly inclined tracks.

5.4. Performance

5.4.1. Efficiency and purity

The efficiency and purity distributions for the TRD tracking are presented in Figure 5.19 for MC simulations of pp and Pb–Pb collisions. The efficiency is defined as the fraction of tracks which are matched to at least two *correct* (based on their MC label) tracklets, because a minimum of two tracklets is required for the TPC space point calibration. Therefore, tracks which do not point into the active detector region in at least two layers are not taken into account.

The online tracklets have up to three MC labels assigned to them. The labels are assigned in descending order based on the frequency of occurrence of the labels assigned to the contributing hits. For the offline tracklets the labels of all contributing clusters are collected, sorted and again the three most frequent labels are assigned to the tracklet. If a tracklet is matched to a track, but the labels of tracklet and track do not match, the tracklet can still be produced by a daughter particle (for example a TR photon or a δ-electron) of the mother which produced the track. These are counted separately from



(a) pp **(b)** Pb–Pb

Figure 5.19.: Efficiency and purity for the TRD tracking are shown as a function of transverse momentum of the seeding tracks for pp (a) and Pb–Pb (b).

the exact matches as related matches. However, these cases are rather rare. Integrated over the full p_{T} range the fraction of related matches is about 1.5% both for pp and Pb–Pb collisions. In the depicted efficiencies both exact and related matches are counted as correct matches.

Below $p_{\mathrm{T}} \approx 1.5\,\mathrm{GeV}/c$ the efficiency for the new TRD tracking algorithm deteriorates compared to the offline tracking which has access to the raw cluster data. The observed decrease in efficiency is expected given the efficiency for single online tracklets shown in Figure 5.5b. The fact that the tracklets in the FEE are limited to span over a maximum of two pads introduces a selection on the deflection inside the drift region. This in turn translates into a position dependent selection on the transverse momentum which is reflected in the observed smeared decrease of efficiency. Primary tracks with $p_{\mathrm{T}} < 300\,\mathrm{MeV}/c$ are neglected by the tracking algorithm, because they do not reach the TRD radius. At high p_{T} the new tracking algorithm yields results compatible to the offline tracking.

The purity is defined as the fraction of tracks with at least two attached tracklets which in addition do not have any fake tracklets attached. A fake tracklet is again identified based on its MC label. In pp collisions, the purity for both the cluster based and the tracklet based algorithm is practically 1, while it decreases to about 0.9 for low momentum tracks in Pb–Pb collisions where the track density is much higher. Overall the efficiency for both tracking algorithms is slightly higher in pp collisions compared to Pb–Pb. Due to the high track density in the detector multiple tracks can contribute to the same tracklet, significantly deteriorating the tracklet resolution [83]. In this case some tracklets can be missed due to the χ^2 selection.

For the TPC space point calibration a lower efficiency can be compensated by simply accumulating more collisions. The current estimate is that the global track reconstruction including the ITS, TPC, TRD and TOF detectors will be performed for a sub-sample of about 4% of all events in the synchronous phase. In Pb–Pb collisions this would allow for

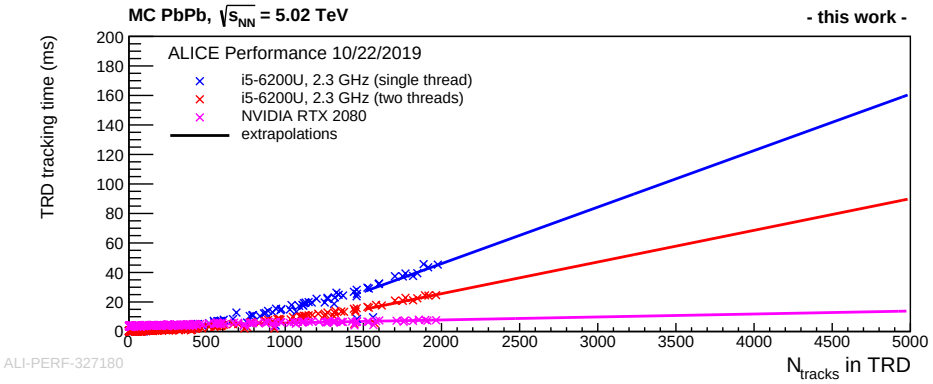


Figure 5.20.: Benchmark results for single-threaded, multi-threaded and GPU reconstructions.

distortion maps to be created for intervals of about 1 minute. The purity on the other hand is required to be as high as possible. Therefore, only peripheral events with a low multiplicity are considered when creating the distortion maps.

5.4.2. Computing speed

Since the TRD tracking will run synchronously to the data taking the processing time is very important. Parallelization can easily be achieved over the tracks and is implemented both on CPUs via OpenMP and on NVIDIA GPUs via CUDA. The processing time per event as a function of the number of tracks which reach the TRD is shown in Figure 5.20. The most central events contain about 2000 ITS-TPC tracks which reach the TRD. This number is not yet large enough to profit strongly from GPU utilization. The speedup for the CPU version when using two OpenMP threads compared to one thread on a dual core CPU is approximately 2, as it is expected in the ideal case for a program which can be parallelized fully.

As explained in Section 2.3.3 for Run 3 the data will not be processed on a per-event basis, but in time frames of 10 to 20 ms corresponding to about 500 to 1000 collisions. The number of tracks to be processed is therefore much higher and a large speedup on GPUs compared to CPUs is expected as can be deduced from the extrapolations shown in Figure 5.20.

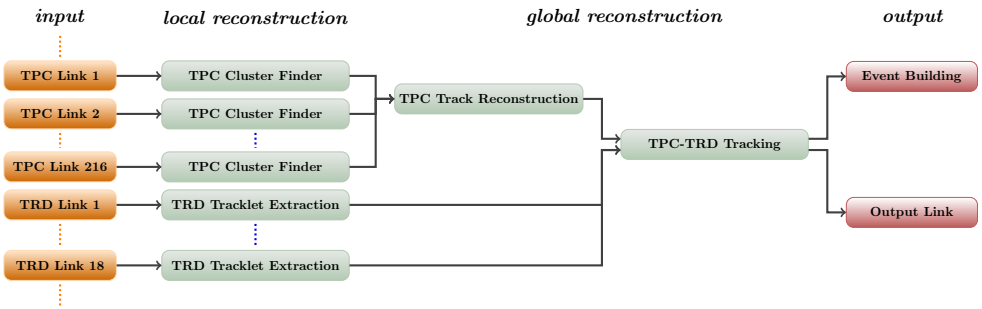


Figure 5.21.: The HLT is organized in components. In this simplified sketch only the components relevant to the TPC seeded TRD tracking are shown. All components in one column can run in parallel and the data volume is reduced when passed from one column to the next. The TRD tracking requires TPC tracks and TRD tracklets as input. The output can either be integrated into the event building or dumped into a ROOT file.

5.5. Development and implementation details

5.5.1. Prototyping in the High-Level-Trigger

At first the TRD tracking was developed in the framework of the HLT in AliRoot. This way it could already be deployed in the official reconstruction during data taking in Run 2. And as the data processing approach in the HLT serves as a prototype for the Data Processing Layer (DPL) in O^2, it is straightforward to port the code afterwards to O^2.

Data treatment within the HLT is done by individual processes which are referred to as components. The term component reflects the modular concept: each process has a well defined input and output and can be executed independently of other processes which act on different data [85]. All components relevant for the TRD tracking are shown in Figure 5.21. An important technical detail for the development of new components for the HLT is its data transport framework. In order to avoid intermediate copying all components are acting on shared memory. They exchange only small data description information, i.e. the data type, size and memory location, instead of the data itself. This approach however does not work for objects inheriting from the `TObject` class in ROOT. Those objects always have to be serialized and de-serialized, because they do not represent Plain Old Data (POD) structures. ROOT offers the possibility to write objects to a file directly using its own streamer, which is very convenient to write the data generated by all LHC detectors to tape, but it is very inefficient for a fast message based online data transfer between different processes. Therefore, new classes must be provided for the concerned data types, i.e. the online tracklets and the TRD tracks that do not inherit from `TObject` but offer all methods required for the tracking. The implementation is reviewed in Section 5.5.2.

The HLT [61] was a compute farm located at Point 2. It consisted of 180 compute nodes which received an exact copy of the detector data from the DAQ system. The HLT

was mainly used for data compression, i.e. the TPC cluster finding was performed on FPGAs, but it could also add trigger decisions to ALICE based on results from the global reconstruction. Each compute node was running an instance of the full reconstruction chain and the events were distributed in a round-robin fashion. Hence, the upper compute time per event depended on the input rate divided by the number of compute nodes. In addition to the components shown in Figure 5.21 many more components existed for the reconstruction of the other detectors, quality assurance and also for calibration tasks [61]. To the DAQ system the HLT looked like another detector. If all compute nodes were occupied and the buffers ran full, the HLT issued a busy signal and all subsequent events were dropped until there was enough space in the buffers again. Therefore, it was extremely important that the TRD tracking did not take too long in order to avoid data loss. Derived from the available resources and the requirements of the other components the available processing time for a single event was about 100 ms for pp and about 300 ms for Pb–Pb collisions.

Before the TRD tracking was enabled during data taking it was developed and tested offline. A centrally provided macro[2] can be used to re-run the HLT reconstruction locally for provided raw data. As the HLT is treated by the DAQ system just like a normal detector, the actual HLT output is already included in the raw data. It needs to be ignored when re-running the reconstruction. Furthermore, the processing chain can be specified to disable the components which are irrelevant for the TRD tracking and would otherwise take too much time.

During the development phase the offline tracking was considered a reference for the performance. But the output of the TRD tracking in the HLT could not be directly compared to the offline tracking for two reasons. First, the seeding tracks used in the HLT are different from the ones used in the offline reconstruction, because the reconstruction algorithms are not identical (compare Sections 4.1 and 4.2). Second, the TRD offline reconstruction is based on raw data: clusters are formed in close vicinity to extrapolated ITS-TPC tracks and these clusters are straight-line fitted to form offline tracklets. By definition these offline tracklets belong to a global track as opposed to the online tracklets which are calculated in the TRD FEE independently of any other detector. The resulting reconstruction efficiencies are not identical. The first point can be circumvented by running the TRD tracker as a standalone application with the same input tracks which are used in the offline reconstruction. These are the track parameters determined at the outer wall of the TPC before their extrapolation to the surrounding detectors. The different data formats, online calculated tracklets versus offline tracklets built from clusters around global seeding tracks, cannot be avoided and have to be taken into account. For this reason the online tracklets are characterized and compared to the offline tracklets in detail in Section 5.2.

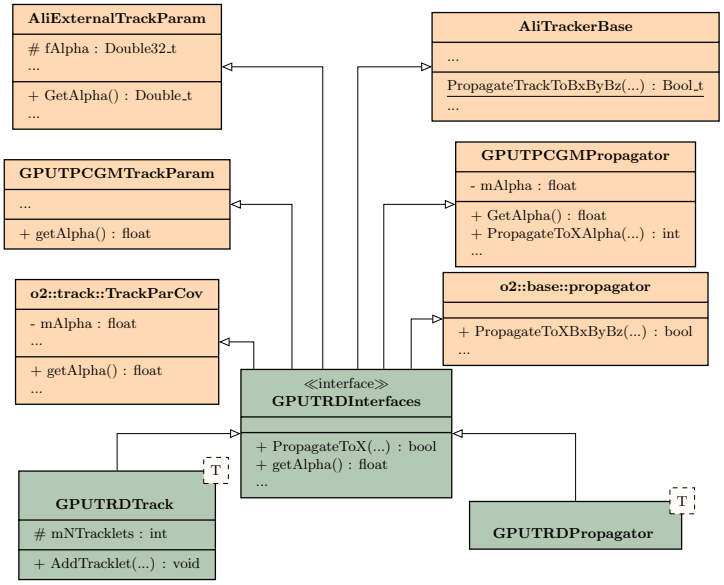


Figure 5.22.: UML class diagram for the main components of the TRD tracking algorithm. By the application of interfaces a single code base can be employed for AliRoot (represented by `AliExternalTrackParam` and `AliTrackerBase`), a standalone GPU framework (represented by `GPUTPCGMTrackParam` and `GPUTPCGMPropagator`) and for O^2 (represented by `o2::track::TrackParCov` and `o2::base::propagator`) simultaneously.

5.5.2. Implementation in O^2

The main parts of the TRD tracking algorithm are contained in the class `GPUTRDTracker`. These are the conversion of the online tracklets to space points and the subsequent track following through the detector with the matching between the respective tracks and space points. Between the different frameworks AliRoot, HLT and O^2 the logic of the tracking algorithm is identical. Thus, in the optimal case the same code can be used in all three frameworks, thereby simplifying the maintenance significantly.

The classes which represent the tracks and which contain the propagation and other methods required by the tracking algorithm do not have the same names in the different frameworks. Also the methods do not necessarily have the same prototype, i.e. their parameters or return type may vary. In order to avoid code duplication interfaces are created both for the track and for the propagation classes using C++ templates. Two new classes, `GPUTRDTrack` and `GPUTRDPropagator`, inherit from these interfaces which map the required methods from the different frameworks onto the required format for the TRD tracking algorithm. This way all conversions can be collected in a single interface header called `GPUTRDInterfaces` and the remaining implementation is deployable in all three frameworks without requiring any changes.

The interface for O^2 is implemented, but still requires further testing. At the moment the overall status of the TRD simulation in O^2 does not yet allow for extensive tests. The

[2] `$ALICE_ROOT/HLT/exa/recraw-local.C`

digitization is in place and a first version of the MCM simulator exists which assembles the tracklets from the digits. However, the trigger information required to map the tracklets to a specific interaction is not available and the new tracklet format still needs to be adapted to the one depicted in Figure 5.4. Currently the development work for the digitization and MCM simulator focuses on reproducing the Run 2 results obtained with the AliRoot framework. Afterwards the data formats will be adapted. At the moment the TRD tracking algorithm is operational in the O^2 framework with ITS-TPC matched tracks and tracklets from a single collision as input. However, in addition to the trigger information the tracklets also lack the MC labels. They are not yet propagated by the MCM simulator. Therefore, a fast feedback for the matching procedure is not yet available. As stated above, for the new TRD tracking algorithm the incorporation of the additional information only requires minor changes in the code without the need of a change of the actual algorithm.

5.5.3. Running on GPUs

In the synchronous phase the full TPC tracking will be performed on the GPUs of the EPN server farm [49]. In an optimistic scenario [86] in addition to the TPC tracking also the ITS tracking and subsequent matching between track segments from ITS and TPC is performed on GPUs. In this case it is beneficial to perform the TRD tracking on the GPUs as well. All input data for a full time frame is copied from the host onto the GPU and the full reconstruction is done without any intermediate data transfer. Frequent data transfers between *host* (here roughly referring to the CPU and the RAM) and *device* (GPU) via the PCI Express bus would otherwise deteriorate the achievable speedup on the GPU exploiting its parallelism with respect to the CPU version.

The TRD tracking is therefore implemented in a way such that it can be included in the ALICE GPU framework and run on GPUs. A benchmark with an NVIDIA RTX 2080 GPU is presented in Figure 5.20. The tracking algorithm does not necessarily utilize all available GPU resources in an optimal way. A GPU can be seen in a simplified way as a processor with thousands of cores. Large blocks of these cores perform the same operations on different data simultaneously. CPUs on the other hand have fewer cores, but they can perform independent operations which can be more complex. The CPU cores benefit from various optimization strategies (e.g. branch prediction) which do not exist on the GPU. During the track following each track is processed by an individual thread. On the GPU a thread corresponds to one of the many cores. Now the tracks can be stopped in different layers, leading to their respective threads being in a stalled state until the other tracks of this block have been processed. The TPC tracking implements a dynamic scheduling that periodically redistributes the remaining workload among all available GPU threads [61]. For the TRD such a scheduling system is not foreseen. The tracks which are considered to be matched to the TRD tracklets have a higher quality as the average tracks which are found in the TPC, because they have already been matched to ITS track segments and fulfill certain quality criteria. If the global χ^2 of the track is too large they are not propagated into the TRD. Therefore the fraction of abandoned tracks is much lower.

Another technical detail is the fact that the memory on the GPU will be managed manually [86]. At the beginning of the processing of a time frame the full memory of the GPU is reserved. During the processing of the time frame the memory is filled with non-persisting input data from one side and persistent data from the other. Blocks of memory from the middle are allocated as scratch space for algorithms. For the TRD tracking scratch space is allocated for example for the array of track hypotheses, the tracklet index arrays which are used to map the tracklets to the space points and for storing the average radii of each chamber. The amount of scratch space required for the TRD tracking algorithm is fixed and known beforehand. Therefore this memory is allocated when an instance of the `GPUTRDTracker` class is initialized. For the processing of all following events the objects will remain at the same memory location. The number of input tracks and tracklets on the other hand changes. For them new persistent memory is allocated, because after the full time frame is processed these tracks and the calculated space points need to be stored as they are used in the subsequent TPC calibration.

6. TPC calibration with track residuals

To correct for the large space-charge distortions inside the TPC in Run 3 a complex calibration procedure is under development. The calibration accounts both for the average distortions by evaluating the mean TPC cluster residuals in a similar way as it was done already in Run 2 and it accounts for the fluctuations exploiting the current measurements of the readout pads of the TPC with a high time granularity. This chapter describes the part of the calibration that was developed in this work, namely the extraction of the mean TPC cluster residuals for individual voxels of the TPC which are required to obtain the correction maps. In contrast to Run 2, the computation in Run 3 is performed in the synchronous stage and on much shorter time intervals. Because of the continuous readout in conjunction with a higher interaction rate, about one minute of data taking will provide the required statistics to obtain a correction map with sufficient precision. During Run 2 about 40 minutes of data taking were required for a single correction map. Consequently, this puts great demands on the computing capabilities.

In contrast to the TRD tracking, which was presented in the previous chapter, the algorithm which is described here is not new. Instead, it is adapted to the new data formats which are employed in the O^2 framework for Run 3 and which are needed to allow for the processing on the basis of time frames instead of a per-event basis. Furthermore, the software framework AliRoot did not fully support the modern C++ standard (C++11 or newer), because of backward compatibility issues with some of the hardware employed in the processing chain. Therefore, the original implementation made extensive use of raw pointers which should be avoided by modern C++ guidelines [87]. A lot of time was thus spend on introducing the use of smart pointers, Standard Template Library (STL) containers and other modern C++ features where it is appropriate. This chapter serves as extensive documentation of the implementation in O^2. The implementation in AliRoot is only briefly summarized within two pages of an ALICE internal note [53], which shall be made public in the near future.

In addition, the processing steps in O^2 are not managed by individual ROOT macros as it was done to a large extent in Run 1 and 2, but by the DPL (see Section 2.3). The TPC calibration is embedded into the DPL in the form of individual devices which can be assembled into a workflow. This enables for example to run the processing steps from the reconstruction in a chain, similar as it was done in the past in the HLT. Each processing step (device) acts directly on the output of its predecessor device without intermediate copying or even without the data leaving the main memory of the computer.

However, the underlying algorithms cannot simply be copied as well. Instead, each step was carefully checked for possible optimizations and during the porting procedure a few minor bugs in the original code were corrected. All steps described in the following were already introduced briefly in Section 3.3. Here, they are described in detail in the

way they are implemented in the software for Run 3.

6.1. Collection of TPC cluster residuals

The first step is the collection of the TPC cluster residuals with respect to the interpolation from the surrounding detectors. All TPC track segments which have an identified continuation in the ITS, TRD and TOF detectors are taken into account for this step. The algorithm is the same as in Run 2. The track segments reconstructed in the ITS are extrapolated through the TPC detector and the track position (y_{out}, z_{out}) is stored at every pad row where a TPC cluster is attached to the global track. Afterwards the ITS-only track is further propagated to the outermost available cluster, which can be either in the TRD or TOF detector. At this point the covariance matrix of the track is reset to remove any possible bias of the ITS track and subsequently a Kalman update of the track parameters is performed. Now the inward propagation starts. All available tracklets from the TRD are added to the track fit. Inside the TPC again the track positions (y_{in}, z_{in}) at each pad row, where a cluster was matched to the global track, are stored. The reference positions for the TPC clusters at each pad row can be calculated as the average of the positions obtained from the inward and outward extrapolations weighted by their uncertainties:

$$y_{\text{ref}} = \frac{y_{\text{out}}\ \sigma_{y_{\text{out}}}^{-2} + y_{\text{in}}\ \sigma_{y_{\text{in}}}^{-2}}{\sigma_{y_{\text{out}}}^{-2} + \sigma_{y_{\text{in}}}^{-2}}. \tag{6.1}$$

This is done analogously for the z reference positions. In addition to the positions also the two track angles are stored at each reference cluster. In principle also tracks without a continuation in the outer detectors TRD and TOF can provide reference positions for the TPC clusters. But due to the large extrapolation the position uncertainty grows significantly without a constraining measurement at a large radius as it is shown in Figure 6.1. At the moment the possibility to extract the residuals from ITS tracks alone is included in the software, but disabled by default.

As stated in Section 3.3, about 1000 cluster residuals are required per voxel for the creation of the space-charge distortion correction map. For about half a million voxels the reference position information alone (in double precision) would add up to about 60 GB which needs to be collected every minute. Compared to the total data stream of the ALICE experiment of about 100 GB/s the size is not that large, but to alleviate the processing the data should be made as compact as possible. The data members of the structure which is filled with the residuals is shown in Table 6.1. For example a small optimization is the storage of the distance Δ_{row} to the last row where a cluster is available instead of the absolute number of the pad row. As the global tracks have a TPC cluster in most of the pad rows this number will mostly be 1 which allows for a better entropy compression than the absolute pad row number.

For each track the number of obtained cluster residuals and the index of the first residual which belongs to the track are stored together with additional information. For example the number of ITS cluster and TRD tracklets and the χ^2 values for the track

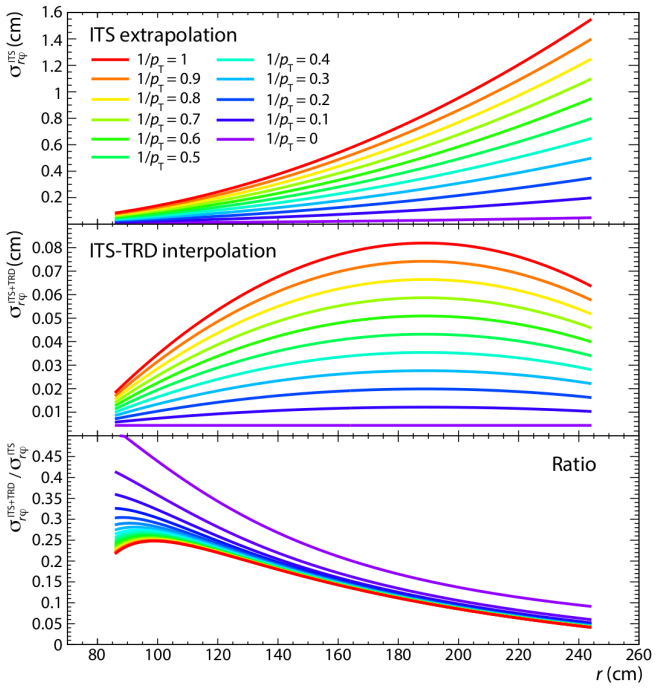


Figure 6.1.: Position resolution as a function of the radius r for ITS-only tracks extrapolated through the TPC volume (upper plot) and for track interpolations between the ITS and TRD detector (middle plot) for different transverse momenta of the tracks between 1 GeV/c and infinity. The bottom plot shows the ratio of the two. With constraining points at outer radii the interpolation is much better defined, even at the inner TPC radius close to the ITS. Taken from [14].

Data member	Type	Range
δy	`short`	−20 to 20 cm
δz	`short`	−20 to 20 cm
y_{ref}	`short`	−50 to 50 cm
z_{ref}	`short`	−300 to 300 cm
$\sin\varphi_{\mathrm{ref}}$	`short`	-1 to 1
$\tan\lambda_{\mathrm{ref}}$	`short`	-2 to 2
sector	`unsigned char`	0 to 17
Δ_{row}	`unsigned char`	1 to $N_{\mathrm{padrows}} - 1$

Table 6.1.: List of data members contained in the structure for a single measurement of a TPC cluster residual. All values are mapped to either `short` (16 bit) or `unsigned char` (8 bit) integer data types.

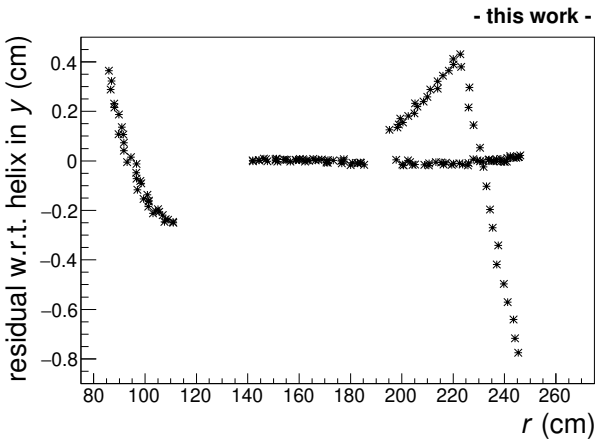


Figure 6.2.: The residuals of the interpolated track positions with respect to a helix fit in y are shown. The depicted track has multiple track segments attached which do not belong to the actual track.

segments in the individual detectors allows for additional selections before the residuals are partitioned in voxels.

6.2. Cleaning of the input data

Due to the increased search roads and inflated TPC cluster uncertainties even supposedly good tracks with a low reduced χ^2 can have a significant amount of fake clusters attached. Also the matching to the surrounding detectors can be wrong in some cases. Therefore, an additional cleaning procedure is required to reject tracks which have an overall bad quality or to reject individual fake clusters attached to otherwise good tracks.

Helix fit

In order to determine the quality of the interpolation, the reference track points y_{ref} and z_{ref} (i.e. the positions at the pad rows given by the ITS, TRD and TOF interpolation) are compared to a helix fit. The points are fitted by a circle in the transverse plane and subsequently by a straight line in the longitudinal plane. Tracks are marked as bad for which the maximum spread of the residuals of the points with respect to the helix fit either in y or z exceeds a configurable threshold. Note, that the residuals are only measured in y and z, since the x position is determined by the fixed pad row position. The thresholds for both directions are set to 0.3 cm by default. Before the circle fit is performed, the reference positions need to be transformed into the same coordinate system. They are defined in the sector coordinate system of their corresponding TPC cluster and for tracks which cross one or more sector boundaries the coordinate systems of the clusters differ.

Now a fast least squares fit is performed for the points in the xy plane. The algorithm

is adapted from the least squares circle fit of the SciPy package [88]. The function to be minimized is

$$F = \sum_i \left[(x_i - x_c)^2 + (y_i - y_c)^2 - r^2 \right]^2, \tag{6.2}$$

where x_c and y_c are the coordinates of the circle center. By first transforming the points to reduced coordinates

$$u_i = x_i - \bar{x}$$
$$v_i = y_i - \bar{y},$$

one arrives at a system of linear equations with only two unknowns:

$$u_c \sum_i u_i^2 + v_c \sum_i u_i v_i = \frac{1}{2} \left(\sum_i u_i^3 + \sum_i u_i v_i^2 \right)$$
$$u_c \sum_i u_i v_i + v_c \sum_i v_i^2 = \frac{1}{2} \left(\sum_i v_i^3 + \sum_i v_i u_i^2 \right).$$

All required sums can be calculated within a single loop and the reduced center coordinates u_c and v_c are easily determined. Finally the radius is given by

$$r^2 = u_c^2 + v_c^2 + \frac{\sum_i u_i^2 + \sum_i v_i^2}{N_{\text{points}}}. \tag{6.3}$$

Afterwards the center coordinates are transformed back into the sector coordinate system and the residuals in y for each point with respect to the circle fit are calculated. An example for a track which is rejected because of the results of the circular fit is shown in Figure 6.2.

The straight line fit in the longitudinal plane requires the path length in the transverse plane. The circular path (arc length, see Figure 6.3) between two consecutive points is approximated linearly by the chord length

$$l_c^2 = \delta x^2 + \delta y^2, \tag{6.4}$$

unless the curvature of the track or the circular path is too long, i.e. $l_c \cdot \frac{1}{r} > 0.05$. In that case the first two terms of the Taylor expansion for the exact formula for the arc length are used:

$$l_a = 2r \operatorname{asin}\left(\frac{l_c}{2r} \right) \approx l_c + \frac{l_c^3}{24 r^2}. \tag{6.5}$$

The Taylor expansion is employed to avoid the use of trigonometric functions in places of the code which are called very often and which should therefore be as fast as possible. It is also beneficial in case the TPC residual calibration is ported to the GPU at some point in the future. A simple straight line fit in the zs plane, where s is the estimated path length at each reference position, yields reference positions in the longitudinal plane for the helix fit. Thus, the residuals in z for each reference point with respect to the helix are determined.

In addition to the possibility to identify bad tracks based on the deviations of the reference points with respect to the helix, the circular fit also yields more precise estimates for q/p_{T} and for the track inclination at each pad row which are stored for later use.

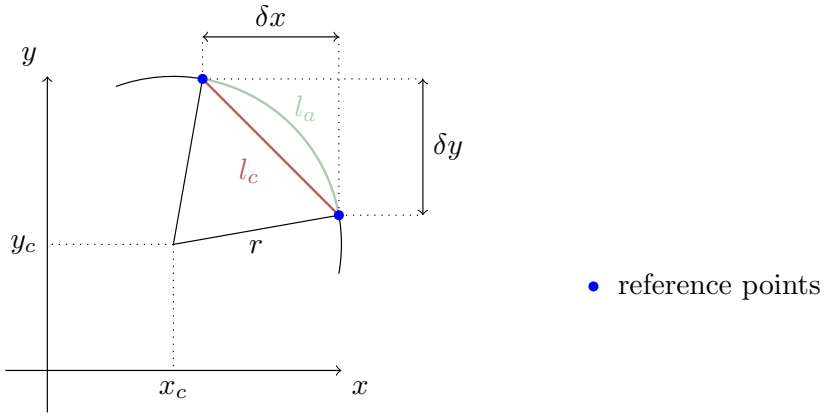


Figure 6.3.: This sketch shows the relevant variables for the arc length approximation. The figure is not drawn to scale. In reality the radius is much larger compared to the distances of two consecutive clusters δx and δy.

Outlier rejection and moving average filter

In the second to last validation step the distributions of the collected residuals δy and δz for each track are searched for outliers. This is done by a straight line approximation from the neighborhood of each point. For a given residual at the i-th pad row, δy_i, the residuals of the neighboring pad rows $i \pm 3$ are fitted by a straight line excluding δy_i. The straight line fits can be performed simultaneously for all points which lie in the same sector. The four required fit sums $\sum x_i$, $\sum x_i^2$, $\sum x_i \delta y_i$ and $\sum \delta y_i$ are first accumulated for all points in arrays. The array for $\sum x_i$ thus contains the following entries:

$$\sum_{i=1}^{1} x_i, \ \sum_{i=1}^{2} x_i, \ \ldots, \ \sum_{i=1}^{N} x_i. \tag{6.6}$$

Afterwards the fit sums required for a given point are extracted from the array excluding the point itself. For example the fit sums required at layer i are extracted via

$$\sum_{j=i-3}^{i+3} x_j \ (\text{for } j \neq i) \ = \ \sum_{j=1}^{i+3} x_j - \sum_{j=1}^{i-3} x_j - (\sum_{j=1}^{i} x_j - \sum_{j=1}^{i+1} x_j). \tag{6.7}$$

For the pad rows at the edges only the neighboring points in one direction are used. In total at least three neighbors are required, otherwise the local straight line fit is not possible and the point is marked as an outlier. To summarize, only two loops over the points are necessary to extract the deviations for all residuals δy_i with respect to their neighborhood $\delta \tilde{y}_i$ in a given sector.

For all validated points the absolute deviations $|\delta y_i - \delta \tilde{y}_i|$ are collected and the RMS is estimated for the distribution excluding 10% of the points which have the largest

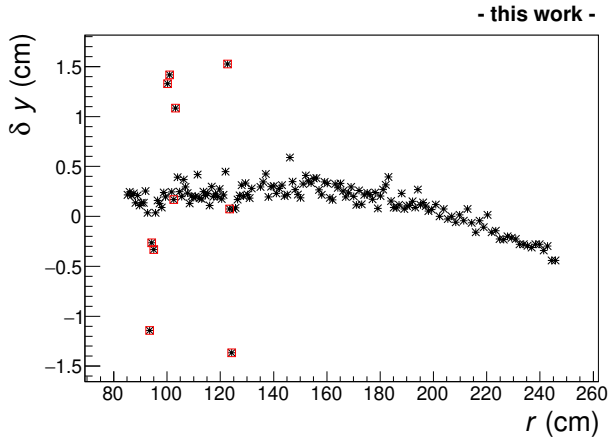


Figure 6.4.: Shown is the dispersion of the TPC cluster residuals with respect to the reference position from the ITS-TRD-TOF interpolation for a single track. This track is accepted, but a few of the measured cluster residuals are marked as outliers (red). The two outliers which seem to be consistent with the distribution are rejected because of their residuals in z which are not shown here.

deviation

$$\mathrm{RMS}_y = \sqrt{\frac{\sum_i (\delta y_i - \delta\tilde{y}_i)^2}{N_{\mathrm{frac}}}}, \tag{6.8}$$

where $N_{\mathrm{frac}} = 0.9 \cdot N$. The same is done for the δz residuals. Now points which are more than 5σ away from their local neighbors, i.e. for which the following comparison is true

$$\left(\frac{\delta y_i - \delta\tilde{y}_i}{\mathrm{RMS}_y}\right)^2 + \left(\frac{\delta z_i - \delta\tilde{z}_i}{\mathrm{RMS}_z}\right)^2 > 5^2 \tag{6.9}$$

are marked as outliers. An example for a track with a few outliers is depicted in Figure 6.4.

As a last check a long range moving average is determined for all N_{acc} points which remain valid up to this point. For every residual δy_i the surrounding points $i \pm 15$ are used to determine an average $\delta\bar{y}_i$. Now the RMS of the deviations of all points from their respective moving average is determined by

$$\mathrm{RMS}_{\mathrm{long}} = \sqrt{\frac{\sum_i (\delta y_i - \delta\bar{y}_i)^2}{N_{\mathrm{acc}}} - \left(\frac{\sum_i (\delta y_i - \delta\bar{y}_i)}{N_{\mathrm{acc}}}\right)^2}. \tag{6.10}$$

In case the RMS on the long range moving average is larger than 0.8 cm the whole track is rejected. These tracks are supposedly wrong ITS-TRD matches or TPC tracks made of track segments from different particles as in the example shown in Figure 6.2.

6.3. Conversion to local residuals

For every reference track the available cluster residuals are sorted into their respective voxels using the reference positions y_{ref} and z_{ref}, the pad row and the sector number. The voxel binning is not fixed. The default binning for Run 3 is similar to the one from Run 2 and depicted in Figure 6.5.

In contrast to Run 2 the ions which cause the distortions of the drift field inside the TPC are not accumulating in dedicated hot spots, but they are more evenly spread over the whole TPC volume. Therefore, the gradients in the space charge density distribution are expected to be smaller which in principle facilitates a coarser voxel binning. The implementation allows for both uniform and non-uniform bin sizes and a configurable total number of bins in all three dimensions. For example the binning in x can be non-uniform with one bin per pad row. In the z/x direction on the other hand a more granular binning might be required close to the central electrode where the magnitude of the distortions is larger due to the longer drift path of the electrons.

After the voxel binning is defined, containers for the local residuals are filled for each TPC sector. The local residuals contain only the compact information required to calculate the distortions in all three dimensions and the dispersion, namely δy, δz, $\tan\varphi_{\text{ref}}$ and a voxel identifier within the given sector. The reference angle $\tan\varphi_{\text{ref}}$ is obtained from the helix fit described in Section 6.2.

While the local residuals are filled also the Center Of Gravity (COG) is updated for each voxel whenever a residual is added to it. The COG $(\overline{y_{\text{ref}}},\ \overline{z_{\text{ref}}})$ is defined as the average of all reference positions that fall into the same voxel. It is required for the final smoothing step later.

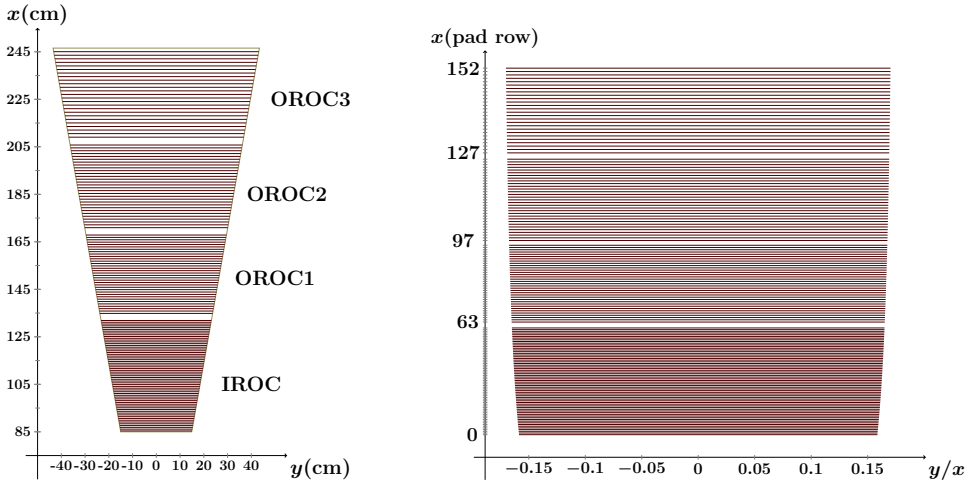
6.4. Determination of the residuals per voxel

Now the residuals are calculated for every voxel of a given sector as follows. The residuals in z are sorted and the resulting distribution is truncated to keep 70% of the data points. The points to reject are chosen from the largest and smallest residuals such that the RMS of the distribution is smallest. The mean value is stored as preliminary value for the distortion in z. The error on the RMS is stored as estimate for the dispersion.

Next and similar as in Run 2, a robust straight line fit of the residuals δy as a function of $\tan\varphi_{\text{ref}}$ is performed to obtain the deconvoluted residuals Δx and Δy (see Figure 3.5b) according to

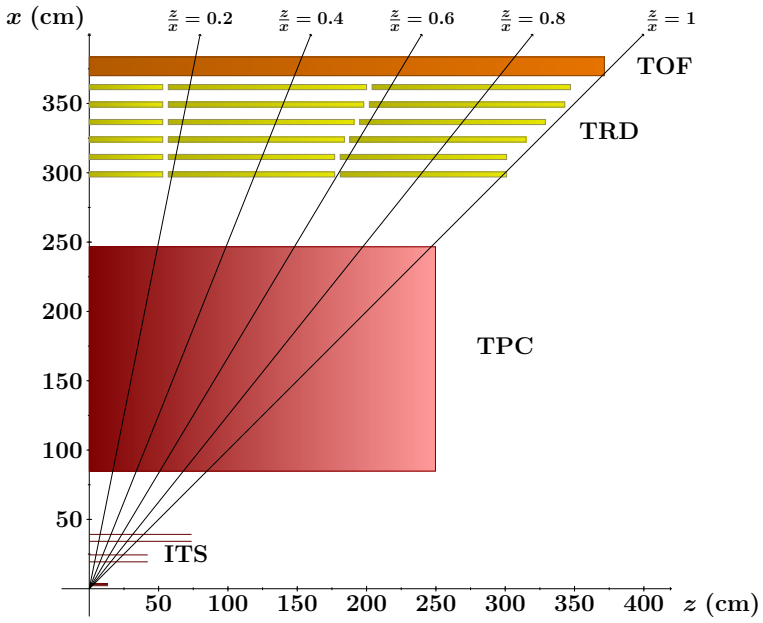
$$\delta y = \Delta y - \Delta x \tan\varphi_{\text{ref}}. \tag{6.11}$$

The δy distribution is truncated in the same way as the distribution in δz. Afterwards a first straight line fit is performed taking into account the remaining points $(\delta y_i,\ \tan\varphi_i)$ and minimizing their absolute deviations with respect to the straight line. In contrast to a linear least squares regression no analytical solving method exists for the minimization of the absolute deviations [89]. A fast numerical algorithm is adapted from [90]. The



(a) Azimuthal cross section for a TPC sector.

(b) Binning in y/x



(c) Binning in z/x

Figure 6.5.: Depiction of a possible TPC voxel binning for Run 3. With respect to Run 1 and 2 (see Figure 3.4) the layout of the pads changes slightly. There are now 153 instead of 159 pad rows grouped in four regions with different pitches. In addition, the inner ITS layers are moved closer to the beam pipe. Again all sketches are drawn to scale.

function to be minimized is

$$\sum_{i=0}^{N-1} |y_i - a - bx_i|, \tag{6.12}$$

where x_i corresponds to $\tan\varphi$ and y_i corresponds to the residuals δy. Above equation can be simplified by exploiting the fact that a median c_M of a set of numbers c_i minimizes the sum of the absolute deviations [90]

$$\sum_i |c_i - c_M|. \tag{6.13}$$

Therefore, the value of a that minimizes Equation (6.12) for a fixed value of b is given by

$$a = \text{median}\{y_i - bx_i\}. \tag{6.14}$$

The derivative of Equation (6.12) with respect to b is

$$\sum_{i=0}^{N-1} -x_i \ \text{sgn}(y_i - a - bx_i) \tag{6.15}$$

where $\text{sgn}(x)$ is the sign function which is defined as zero for $x = 0$. A minimum of Equation (6.12) is found by inserting a from Equation (6.14) in Equation (6.15) and setting the equation to zero. The resulting equation with a single unknown variable b can be solved numerically by bracketing and bisection. An initial guess for a and b is obtained from the solution of the least squares method. The standard deviation σ_b provides an estimate for the iteration step. Equation (6.15) is evaluated at b and $b + 3\sigma_b$. If the sign of Equation (6.15) is different for both guesses of b we have found 2 points b_1 and b_2 which include the minimum. After the bracketing step the minimum between these points is found by bisection. Equation (6.15) is evaluated at

$$b_c = b_1 + \frac{b_2 - b_1}{2} \tag{6.16}$$

and b_c replaces the one of the other two points which has the same sign. This is repeated until the abort criterion

$$|b_2 - b_1| < 0.01 \cdot \sigma_b \tag{6.17}$$

is reached. To evaluate Equation (6.15) efficiently a fast method to determine the median value is needed. In case less than 20 points remain for the voxel after the truncation a fast insertion sort is performed to find the median, otherwise the STL algorithm `std::nth_element` is used. In case at a later stage all computations to evaluate the voxel residuals should be moved to the GPU, which might be desirable in the future depending on the workload distribution on the EPNs, the code cannot rely on STL algorithms. Therefore, the `std::nth_element` algorithm has been implemented manually as well such that it is available for a possible later use. Again the algorithm has been adapted from [90].

The fit results of this first straight line fit are used to correct the residual distribution for its slope. The slope corrected distribution is obtained via

$$y_i^{\text{corr}} = y_i - (a + bx_i). \tag{6.18}$$

After the slope correction a robust estimate for the dispersion of the data points is given by the Median Absolute Deviation (MAD), which is defined as

$$\sigma_{\text{MAD}} = 1.4826 \cdot \text{median}\{|y_i^{\text{corr}} - \text{median}(y_i^{\text{corr}})|\}, \tag{6.19}$$

where 1.4826 is a scale factor for normally distributed data [91]. Now another truncation is performed for the residuals in y but in this case the expected dispersion σ_{MAD} is provided and the data sample is truncated such that its RMS matches σ_{MAD}. The only additional condition is that at least 50% of the points are kept. A second straight line fit is performed in the same way as the first one to the points (x_i, y_i) which remain after the second truncation.

The fit parameters a (corresponding to Δy) and b (corresponding to Δx) and their errors (these are the errors from the least squares estimate of the second fit) are stored in the structure holding the voxel results together with the mean distortion in z. A first estimate of the dispersion for the voxel in y is given by σ_{MAD}.

6.4.1. Voxel validation

The results obtained for a voxel might be considered bad for different reasons. A voxel might be masked due to too few entries. In addition, the fit results possibly do not meet the quality criteria. If the errors of the fit results of the robust straight line fit or the RMS of the residual distribution in z exceed configurable thresholds, the voxel is masked. Furthermore, the dispersion σ_{MAD} must not exceed a given threshold which is by default set to $\sigma_{\text{MAD}}^{\text{max}} = 1.1\,\text{cm}$. The pad rows, respectively bins in x, for which more than 50% of the voxels in y/x and z/x are masked, are marked as possibly bad rows. In case more than 40% of all rows are possibly bad the whole sector is masked. Otherwise sections of pad rows are identified which have more than four consecutive possibly bad rows or less than three consecutive good rows. These sections are fully masked to be less susceptible to outliers. In the following smoothing step an attempt is made to recover these voxels.

6.4.2. Smoothing the residuals

At this point the distortion maps for all individual voxels is available, i.e. for each voxel the following mean residuals are determined

$$\overrightarrow{\text{Res}} = \begin{pmatrix} \Delta x \\ \Delta y \\ \delta z \end{pmatrix}, \tag{6.20}$$

where δz is not yet corrected for the contribution from the radial distortion.

Some voxels can be empty due to too few entries or because the fit quality is bad. In addition, the residuals measured for the individual voxels can fluctuate strongly for a given sector. Before the distortion maps are parameterized by a polynomial interpolation which can be accessed fast during reconstruction of the full data, they are processed by a kernel smoother. This way outlier voxels are recovered while the structural features of the data are kept. A kernel smoother provides an estimate at an arbitrary point x as a polynomial interpolation from the measurements in the number of points (in this case voxels) neighboring this point x. The neighboring points are weighted by their distance to x. The function assigning the weight to the data points is called kernel K. Kernels are typically chosen to be unimodal and symmetric around zero. The kernel center is placed at the given data point. A popular choice is the Epanechnikov kernel [92]

$$K(u) = \begin{cases} \frac{3}{4}(1-u^2), & \text{if } |u| \leq 1 \\ 0, & \text{otherwise} \end{cases} \tag{6.21}$$

where u is the distance of the neighboring data point to the kernel center. The distance u is scaled by the bandwidth of the kernel which determines the size of the neighborhood that is taken into account for the smoothed estimate at x.

The smoothing is done for each TPC sector separately. Within the sector the smoothing is done for each voxel by evaluating the mean residuals of the neighboring voxels. The model used for the residuals interpolation is either linear or parabolic. In the following the linear model is discussed first and the necessary changes for the parabolic model are described later. For each mean residual Res_j the results of the neighboring voxels in all three directions x, y and z are taken into account. For the linear model the residual in dimension j is described by the function

$$f_j = \sum_i a + b_x d_{x_i} + b_y d_{y_i} + b_z d_{z_i}, \tag{6.22}$$

where a is the offset, i.e. the average residual at the center of the given voxel, $b_{x,y,z}$ are the slopes in all three dimensions of the linear model and d_{x_i,y_i,z_i} are the distances in the corresponding dimension. The sum over i is built for all neighboring points which are taken into account. The function to be minimized for each residual dimension j is

$$F = \sum_i (f_j - \text{Res}_{j_i})^2 w_i, \tag{6.23}$$

where w_i is the kernel weight which is determined by the kernel and the bandwidth. The calculation of w_i is described below. To find the minimum the derivatives of Equation (6.23) with respect to the offset and the three slopes needs to be set to zero. For example the derivative with respect to a is given by

$$\frac{\partial F}{\partial a} = \sum_i a w_i + b_x d_{x_i} w_i + b_y d_{y_i} w_i + b_z d_{z_i} w_i - \text{Res}_i w_i. \tag{6.24}$$

To arrive at a system of linear equations of the form

$$\mathbf{A}\vec{x} = \vec{b}, \tag{6.25}$$

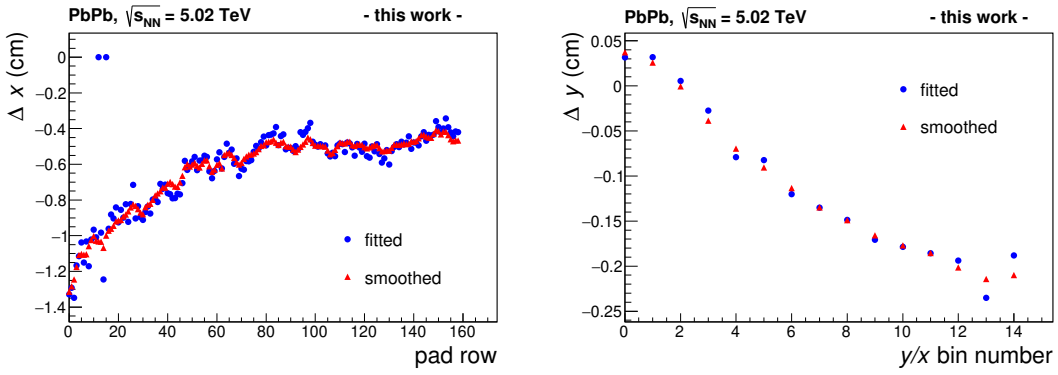


Figure 6.6.: Comparison between fitted and smoothed residuals. The left plot shows the residuals Δx for a single bin in y/x and z/x for sector 1A. Two voxels can be recovered by the smoothing step. The right plot shows the Δy residuals for fixed bins in x and z/x for the same sector.

where $\vec{x} = \begin{pmatrix} a & b_x & b_y & b_z \end{pmatrix}^\mathsf{T}$ holds the required variables, the coefficients for each element of $\vec{x}$ from every derivative are put in columns of a matrix. The remaining terms are moved to the right hand side into $\vec{b}$. The resulting equation looks as follows (the sum symbols are omitted for better visibility):

$$\begin{pmatrix} w_i & d_{x_i} w_i & d_{y_i} w_i & d_{z_i} w_i \\ d_{x_i} w_i & d_{x_i}^2 w_i & d_{x_i} d_{y_i} w_i & d_{x_i} d_{z_i} w_i \\ d_{y_i} w_i & d_{y_i} d_{x_i} w_i & d_{y_i}^2 w_i & d_{y_i} d_{z_i} w_i \\ d_{z_i} w_i & d_{x_i} d_{z_i} w_i & d_{y_i} d_{z_i} w_i & d_{z_i}^2 w_i \end{pmatrix} \cdot \begin{pmatrix} a \\ b_x \\ b_y \\ b_z \end{pmatrix} = \begin{pmatrix} \mathrm{Res}_i w_i \\ d_{x_i} \mathrm{Res}_i w_i \\ d_{y_i} \mathrm{Res}_i w_i \\ d_{z_i} \mathrm{Res}_i w_i \end{pmatrix} . \tag{6.26}$$

In the software all sums are calculated first. Afterwards the solution is found by Cholesky decomposition of $\mathbf{A}$ followed by forward and backward substitution. This last step is performed by the `TDecompChol` class of ROOT. An explanation of the algorithm can be found in many textbooks, for example [93]. Although $\mathbf{A}$ is positive-definite the Cholesky decomposition may fail due to rounding errors. In that case the size of the neighborhood around the voxel is increased and the procedure is repeated.

By default the residuals in z are smoothed by a linear function and the residuals in x and y by a parabolic function, because in z direction there are only 5 bins by default. The steps using a parabolic function are the same, only Equation (6.22) becomes

$$f_j = \sum_i a + b_x d_{x_i} + b_y d_{y_i} + b_z d_{z_i} + c_x d_{x_i}^2 + c_y d_{y_i}^2 + c_z d_{z_i}^2 \tag{6.27}$$

and in the end a set of seven instead of four linear equations has to be solved. The results of the smoothing are illustrated in Figure 6.6 where the residuals fitted for each voxel individually are compared to the results of the smoothing step.

So far the number of neighbors taken into account and how to determine their kernel weight has not been discussed. At least three (four) neighboring points are required

for kernel smoothing with a linear (parabolic) function. The bandwidth of the kernel is adjusted such that for the first try in solving Equation (6.25) only the minimum number of neighbors enters into the calculation. If no solution can be found or if some of the voxels in close vicinity are marked as bad, the effective bandwidth is increased before the first try such that more voxels enter into the calculation.

For the correction of the space-charge distortions in Run 2 the optimal bandwidths b_{w} have been found to be

$$\begin{pmatrix} b_{\mathrm{w}}^{x} \\ b_{\mathrm{w}}^{y} \\ b_{\mathrm{w}}^{z} \end{pmatrix} = \begin{pmatrix} 2.1 \\ 2.1 \\ 1.7 \end{pmatrix} \tag{6.28}$$

The effective kernel width is given by the bandwidth and the bin width w_{Bin} as

$$w_{\mathrm{eff}}^{i} = b_{\mathrm{w}}^{i} \cdot w_{\mathrm{Bin}}^{i}. \tag{6.29}$$

The effective kernel width can now be used to calculate the kernel weight for a given data point. First, the distance of the COG of the neighboring voxel to the given voxel, d_i is calculated for all three dimensions i and the distance is scaled with w_{eff}^{i}:

$$d_{w_i} = \frac{d_i}{w_{\mathrm{eff}}^{i}} \tag{6.30}$$

The weighted distance d_{w_i} (which has no unit) can now be inserted into Equation (6.21) to calculate the total weight

$$w = \prod_{i=x,y,z} \frac{3}{4}(1 - d_{w_i}^{2}). \tag{6.31}$$

If the neighborhood includes voxels at the edge of a TPC sector the effective kernel weight can be adapted to reduce the bias towards the side where more points are available.

6.5. Voxel dispersion

After smoothing the residuals, their dispersion in y can be determined more accurately for each individual voxel. For all collected residuals δy_i the angular dependence is removed and the smoothed residual in y, called $\Delta y_{\mathrm{smooth}}$, is subtracted

$$\delta y_i^{\mathrm{disp}} = \delta y_i - \Delta y_{\mathrm{smooth}} - \Delta x_{\mathrm{smooth}} \tan \varphi_i, \tag{6.32}$$

such that the points $\delta y_i^{\mathrm{disp}}$ are centered around zero and the dispersion is given by the resolution of the MAD of this distribution (see Equation (6.19)). Also the dispersions for each voxel are smoothed for the TPC sectors in the same way as it is done for the residuals.

Finally the contribution of the radial distortion can be subtracted from δz:

$$\Delta z = \delta z + \Delta x_{\mathrm{smooth}} \tan \lambda. \tag{6.33}$$

At this point the average residuals in all three dimensions and an estimate for the dispersion of the residuals in y is determined for every voxel of the TPC.

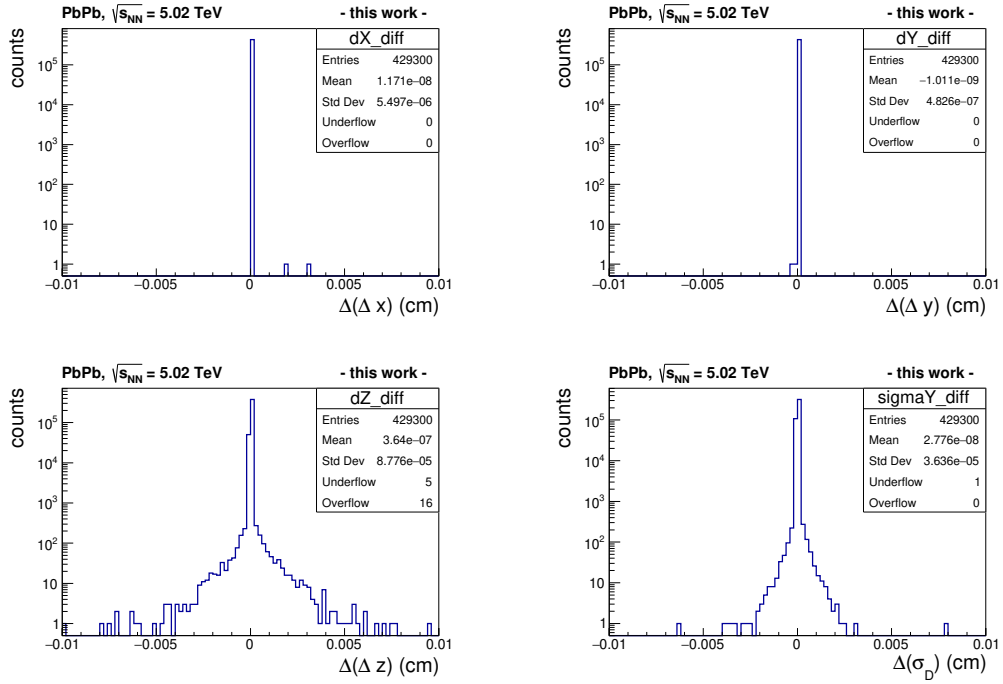


Figure 6.7.: Comparison of the results of the residuals and the dispersion before smoothing for each voxel between the AliRoot reference version and the newly developed O^2 version.

6.6. Comparison to AliRoot version

Currently the simulation framework in O^2 is not yet ready to produce a large scale data set which is required to create TPC space-charge distortion maps. The largest available data set in the new data format is a sample of 200 Pb–Pb events. In the summer of 2019 the simulation of 200 Pb–Pb events still required a dedicated server at CERN with 1 TB RAM. The memory requirements since then have been significantly reduced, but simulating 6000 consecutive timeframes, i.e. 3M Pb–Pb collisions is still not possible at the moment. The main limitation comes from processing steps which produce data types that require serialization for messaging. These messages quickly reach certain buffer limits. For example the limit for the `TBuffer` class which is used for serializing ROOT objects is 1 GB. Nevertheless these are technical details of the framework (which is still under development) that will be fixed in the near future.

The current simulation status allows one to test the extraction of the TPC cluster residuals by interpolating ITS-TOF tracks. As soon as the TRD simulation incorporates the bunch crossing counter which is required to assign a time to the tracklets, they can be matched with ITS-TPC tracks of the same bunch crossing. Then also the TRD information can be included in the interpolation. The algorithm does not require changes. The only difference is that there will be more points constraining the track parameters around the TPC improving the resolution of the interpolation.

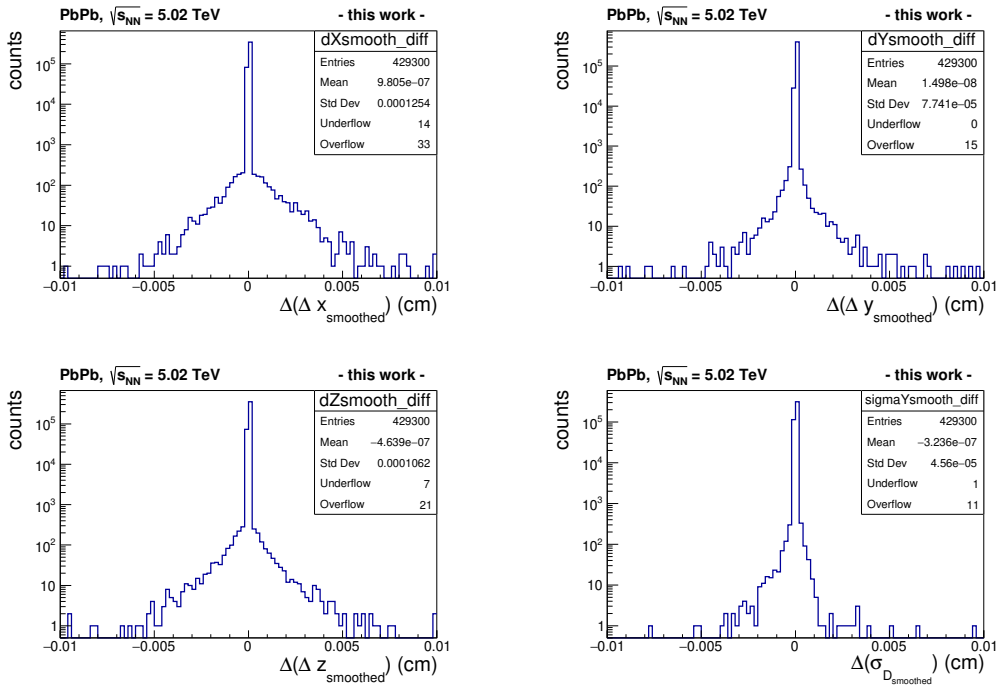


Figure 6.8.: Comparison of the results for the same data as in the previous Figure 6.7 after smoothing for each voxel.

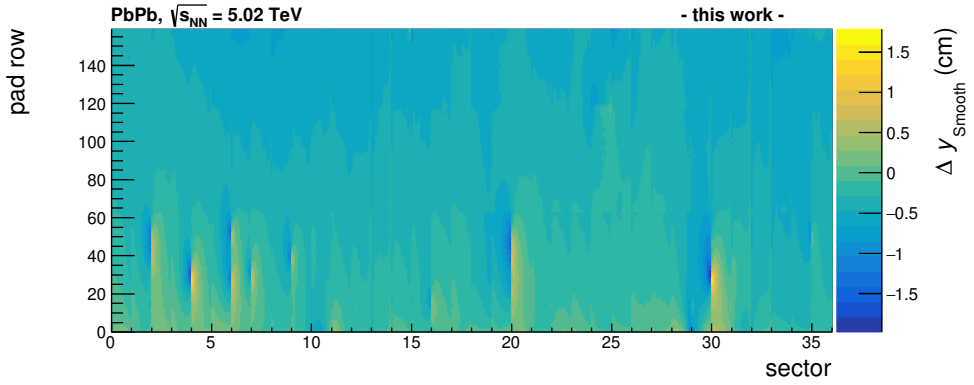


Figure 6.9.: Measured space-charge distortions in y (corresponds to distortion in $r\varphi$) for a Pb–Pb run of 2015 taken at an interaction rate of 2.3 kHz. Only the region close to the central electrode with $|z/x| < 0.2$ is shown as it is also done in Figure 3.1a.

To test the conversion to local residuals and the extraction of residuals and dispersion for each voxel, data from Run 2 is converted into the new data format and used as input. This allows to compare the results directly to the original implementation in AliRoot which is particularly helpful during the development phase. The input data which has been used is obtained from a 2015 Pb–Pb run taken at an intermediate interaction rate of 2.3 kHz.

The results of the voxel residuals obtained before smoothing are compared in Figure 6.7 for both the Run 2 and the newly developed O^2 version. They agree almost perfectly. A 100% match is not expected, since for example parts of the calculation are done with 32 bit floating point precision instead of 64 bit. Also tiny differences in the code can have a visible effect. For example the sum accumulated for an array in an iterative loop will depend on the order of the elements in the array, because of the limited floating point precision. All these effects have been mitigated to achieve the best agreement between the two implementations. The differences of the smoothed residuals are shown in Figure 6.8. Also here the results agree very well.

The measured space-charge distortions which are depicted in Figure 3.1a can be recreated within the new framework as it is shown in Figure 6.9. The hot spots at some of the IROC edges are visible. Because of the lower interaction rate, the magnitude of the distortions is lower and some features originating from static distortions are visible which were covered by the larger distortions in Figure 3.1a.

6.7. Discussion

A major part of the TPC space point calibration with track residuals has been implemented in the new O^2 framework for Run 3. However, before the smoothed residuals are used in the TPC track reconstruction they need to be parameterized by Chebyshev polynomials [94] which allow for a fast query of the distortions as a function of pad row,

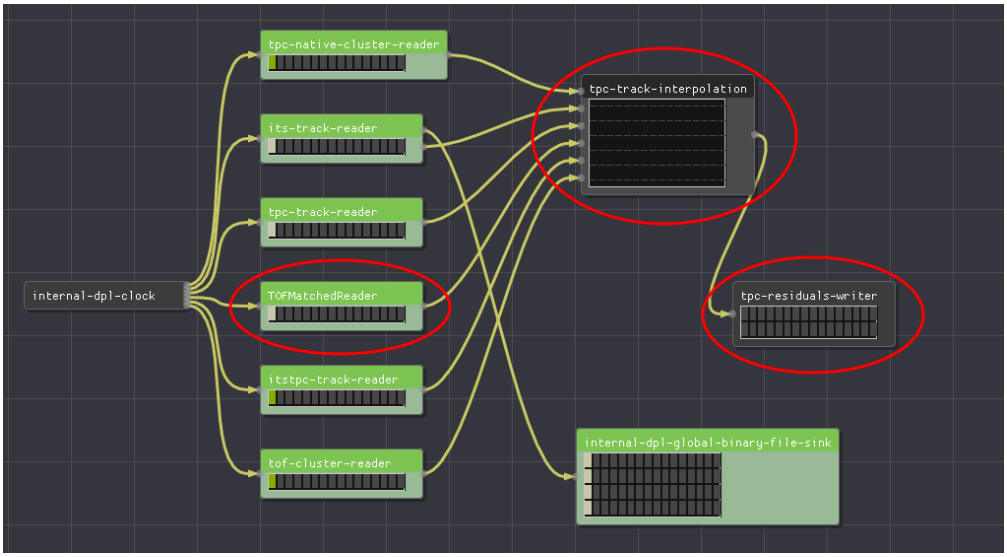


Figure 6.10.: Screenshot of the O^2 debug GUI for the track interpolation workflow. In the depicted case the input data (reconstructed track segments from different detectors, the TPC and TOF clusters and the matching information) is read from files and forwarded to the TPC track interpolation DPL device. The obtained residuals are written to a file. The DPL devices marked with red are developed within this work.

y/x and z/x. Furthermore, the parameterization is required for the inversion of the correction maps for the application in simulations which is described in Section 3.4.

All code which is ported is tested extensively. Also the extraction of the TPC cluster residuals with respect to the track interpolation is integrated into the DPL framework. A screenshot of the debug GUI of the O^2 framework is shown in Figure 6.10. The input in the depicted case is provided by individual reader devices which access the files created by the preceding reconstruction steps, for example the TPC cluster or the ITS-TPC matched tracks. In principle the framework allows one to concatenate workflows. It is possible to connect the track interpolation directly to the output of the preceding devices without the intermediate file writing and reading. However, at the moment the framework does not recognize in such cases that the writing and reading steps need to be dropped. This will be fixed in the near future by the core framework developers and then it is possible to start with a single command the full reconstruction. The topology of the workflow, i.e. the connections between the different devices and the order in which they are executed, is determined automatically by the framework.

7. Summary and outlook

In the upcoming LHC Run 3, starting in 2021, the upgraded ALICE experiment will record Pb–Pb collisions at an interaction rate up to 50 kHz, which is about a factor 7 higher than in Run 2. To record the largest possible number of collisions with minimum bias, several detectors including the TPC, the main tracking device of ALICE, will be operated in a continuous readout mode instead of a triggered one. This leads to a significant amount of ions being present in the TPC at any given moment. The ions cause distortions of the electron drift lines of several centimeters that fluctuate in time.

This work describes their correction via a calibration procedure using the information of the ITS detector, which is located inside, and the TRD and TOF detectors, located around the TPC, respectively. Since the ITS, TRD and TOF detectors are not affected by space-charge distortions, they can provide interpolated reference positions for the TPC clusters. For data taking intervals of about one minute the average residuals of the TPC clusters with respect to the reference positions will be extracted in small sub-volumes (voxels) of the TPC. Only peripheral collisions with a low track density are used for this step. In the final reconstruction the calculated map of residuals is applied to all collisions of the same time interval to correct the TPC cluster position bias.

The unexpected presence of large space-charge distortions in small regions of the TPC in Run 2 required to implement this calibration procedure, which was originally foreseen only for Run 3, already in Run 2. It serves as a reference for the implementation in the new online-offline computing framework of ALICE in Run 3, called O^2. There are manifold differences to be taken into account for the implementation in O^2, due to the switch to a continuous readout for most of the detectors on the one hand and due to a change of the computing paradigm on the other. While the TPC space point calibration was performed offline during Run 2, a large part of the procedure will be performed online, synchronously to the data taking, in Run 3.

To facilitate the highest possible readout rate the TRD will transmit only track segments reconstructed in the front-end electronics, called tracklets, instead of the full raw data as in Run 1 and 2. The development of the required online tracking algorithm for the TRD is the main part of this thesis. The algorithm employs a Kalman filter to match extrapolated ITS-TPC tracks to TRD tracklets. A detailed study of the resolution of the tracklets is necessary. The optimal resolution is found for tracks which traverse the TRD under the Lorentz angle. The transverse position of the tracklets is measured with a precision of about $\sigma_y = 400\,\mu\text{m}$, for the longitudinal direction it is around $\sigma_z = 2.4\,\text{cm}$ and the angular resolution is about $\sigma_\varphi = 0.76°$.

The matching performance is compatible with the Run 1 and 2 offline tracking algorithm, which has access to the raw cluster data, for tracks with $p_\text{T} > 1.5\,\text{GeV}/c$. At the same time it fulfills the computing speed requirements for Run 3. The processing time for

a central Pb–Pb collision is about 40 ms on a single CPU core and 8 ms on an NVIDIA RTX 2080 GPU. For the processing of full time frames of 500 to 1000 collisions a much larger speedup factor is expected. The CPU version of the algorithm developed in this thesis was already activated in the High-Level-Trigger during data taking in 2018 both in pp and in Pb–Pb collisions. Built on the ALICE GPU framework the algorithm is contained in a single source code file and supports wrappers to different APIs, i.e. there is no code duplication needed for the CPU and the GPU versions.

The second part of this work is the extraction of the TPC cluster residuals with respect to ITS-TRD-TOF tracks and the subsequent creation of maps of space-charge distortions. The implementation in O^2 follows the original Run 2 implementation in terms of the employed algorithms. Because the distortions in Run 2 appeared unexpectedly, the implementation in Run 2 was done under time pressure. No documentation of the code exists. The computation was split in different classes which exchanged their intermediate results via ROOT histograms and trees. A lot of redundant information was generated which required additional filtering steps. The space point calibration of the TPC was the most compute intense task of the offline reconstruction of the Pb–Pb data recorded in Run 2.

Chapter 6 describes the procedure which is within this work implemented in O^2 in detail. It follows modern C++ guidelines and reproduces the results of the original implementation. Since it is implemented as a Data Processing Layer device, it can be integrated into the global reconstruction workflow of O^2. The implementation of the required data structures as *Plain Old Data* data types enables an efficient data transfer between the individual processes. The simulation framework for Run 3 does not yet allow for the creation of a data sample large enough to extract the TPC cluster residuals on a per voxel basis. However, as soon as the framework is ready the developed ITS-TOF track interpolation can be performed on multiple time frames without requiring any code changes.

The track interpolation procedure is also ready to incorporate the TRD information. The simulation of the TRD front-end electronics in O^2, which is not part of this thesis, is delayed and not yet ready to produce tracklets with a reference to a certain bunch crossing. As soon as the references are provided, matching between ITS-TPC tracks and TRD tracklets can also be done for full time frames. This does not increase the algorithm complexity, but only the buffer size for the tracklets and the tracks. For single collisions, where the bunch crossing association is trivial, the new TRD tracking algorithm is already verified in the O^2 framework.

A possible optimization for the future is the tuning of the selection criteria of the TRD tracking algorithm for the asynchronous reconstruction phase. Currently, the tracking algorithm is optimized for purity, such that a clean sample of global tracks is obtained in the synchronous reconstruction which is used for the TPC calibration. In the asynchronous reconstruction the same tracking algorithm will run for all events. A longer processing time is affordable in this step and thus the track following with multiple hypotheses could be enabled.

In conclusion, the present work is one of the many building blocks which are required for a successful detector operation of ALICE beyond LS 2.

Acknowledgments

I thank Prof. Dr. Klaus Reygers and Prof. Dr. Johanna Stachel for their continuous support during my studies. I am thankful that I could work on the topic of charged particle tracking, which fascinated me since my master studies. I am glad that I was given the possibility to take many trips to CERN, not only to collaborate with the experts on site, but also for example to participate as a volunteer at the Open Days in 2019 or to support the hands-on detector work at Point 2, which I enjoyed very much. I thank Prof. Dr. Hans-Christian Schultz-Coulon to be available as second referee and particularly for his commitment to the HighRR Research Training Group.

Without the advice from Dr. David Rohr this work would not have been possible. Thank you for your patient explanations of all the programming related topics, but also for your encouragement to go to conferences and for your continuous availability for help. I heartily thank Dr. Ruben Shahoyan for his detailed explanation of the space-charge distortion correction framework in Run 2 and for always answering my questions clearly and extremely fast. I thank Dr. Marian Ivanov for sharing his expertise on charged particle tracking and especially for showing me how to work interactively with `TTree`s, which facilitated my performance analyses significantly.

For proofreading and discussions around my thesis I would like to thank PD Dr. Yvonne Pachmayer, Dr. David Rohr, Dr. Ruben Shahoyan Dr. Alexander Schmah and Dr. Sabrina Schecher.

I also want to say thank you to the whole ALICE@PI group, including their alumni Dr. Sebastian Klewin, Dr. Dennis Weiser and Dr. Hans Beck.

A. TRD tracklet parameterizations

As described in Section 5.3.3, the resolution of the tracklets transverse position y and deflection d_y are parameterized by a function of the form

$$\sigma(\sin\varphi_{\text{trk}}) = \sqrt{a^2 + c^2(\sin\varphi_{\text{trk}} - b)^2}. \tag{A.1}$$

The parameters a, b and c need to be stored for all possible magnetic field configurations. Furthermore, the conversion of the angle of the track $\sin\varphi_{\text{trk}}$ into a deflection over the length of the TRD drift region of 3 cm depends on the magnetic field configuration. It is parameterized by

$$d_y^{\text{trk}} = a + b\sin\varphi_{\text{trk}} + c(\sin\varphi_{\text{trk}})^2. \tag{A.2}$$

The following Table A.1 lists the parameters for both resolutions and the conversion $\sin\varphi_{\text{trk}} \to d_y^{\text{trk}}$ for the magnetic field configurations $B_z = \pm 0.2\,\text{T}, \pm 0.5\,\text{T}$. The parameterizations are obtained from the following runs:

- $B_z = -0.5\,\text{T}$: Run 246390 of LHC15o (Pb–Pb).
- $B_z = +0.5\,\text{T}$: Run 244340 of LHC15n (pp), Run 245353 of LHC15o (Pb–Pb).
- $B_z = -0.2\,\text{T}$: Run 271009 of LHC17g (pp).
- $B_z = +0.2\,\text{T}$: The parameterization is taken from run 271009 as well, but the opposite sign is assigned to the Lorentz angle parameters.

	$\mathbf{B_z} =$	**-0.2 T**	**+0.2 T**	**-0.5 T**	**+0.5 T**
σ_y	a^2 (cm^2)	$1.6 \cdot 10^{-3}$	$1.6 \cdot 10^{-3}$	$1.6 \cdot 10^{-3}$	$1.6 \cdot 10^{-3}$
	b	$1.43 \cdot 10^{-2}$	$-1.43 \cdot 10^{-2}$	-0.14	0.125
	c^2 (cm^2)	$4.55 \cdot 10^{-2}$	$4.55 \cdot 10^{-2}$	0.1156	0.0961
σ_{d_y}	a^2 (cm^2)	$1.225 \cdot 10^{-3}$	$1.225 \cdot 10^{-3}$	$2.209 \cdot 10^{-3}$	$1.681 \cdot 10^{-3}$
	b	$9.8 \cdot 10^{-3}$	$-9.8 \cdot 10^{-3}$	-0.15	0.15
	c^2 (cm^2)	$3.88 \cdot 10^{-2}$	$3.88 \cdot 10^{-2}$	0.2025	0.1849
d_y^{trk}	a (cm)	0.1	-0.1	-0.15	0.13
	b (cm)	1.89	1.89	2.34	2.43
	c (cm)	0.4	-0.4	0.56	-0.58

Table A.1.: Parameters for the tracklet parameterization stored in the TRD tracking algorithm.

Bibliography

[1] H. Geiger and E. Marsden. On a diffuse reflection of the α-particles. In: *Proceedings of the Royal Society A* 82.557 (1909), pp. 495–500.
DOI: 10.1098/rspa.1909.0054.

[2] E. Rutherford. The scattering of α and β particles by matter and the structure of the atom. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 21.125 (1911), pp. 669–688.
DOI: 10.1080/14786440508637080.

[3] J. D. Cockcroft and E. T. S. Walton. Experiments with high velocity positive ions. In: *Proceedings of the Royal Society A* 136.830 (1932), pp. 619–630.
DOI: 10.1098/rspa.1932.0107.

[4] J. D. Cockcroft and E. T. S. Walton. Experiments with high velocity positive ions. In: *Proceedings of the Royal Society A* 144.853 (1934), pp. 704–720.
DOI: 10.1098/rspa.1934.0078.

[5] F. Hinterberger. Physik der Teilchenbeschleuniger und Ionenoptik. German. Vol. 2. Springer, 2008.
DOI: 10.1007/978-3-662-09312-2.

[6] *Record luminosity: well done LHC.* Accessed: 2019-12-17.
URL: https://home.cern/news/news/accelerators/record-luminosity-well-done-lhc.

[7] L. Evans and P. Bryant. LHC Machine. In: *Journal of Instrumentation* 3.08 (2008), S08001–S08001.
DOI: 10.1088/1748-0221/3/08/s08001.

[8] G. Apollinari et al. High-Luminosity Large Hadron Collider (HL-LHC) : Preliminary Design Report. Tech. rep. CERN-2015-005. 2015.
DOI: 10.5170/CERN-2015-005.

[9] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. In: *Journal of Instrumentation* 3.08 (2008), S08003–S08003.
DOI: 10.1088/1748-0221/3/08/s08003.

[10] The CMS Collaboration. The CMS experiment at the CERN LHC. In: *Journal of Instrumentation* 3.08 (2008), S08004–S08004.
DOI: 10.1088/1748-0221/3/08/s08004.

[11] The LHCb Collaboration. The LHCb Detector at the LHC. In: *Journal of Instrumentation* 3.08 (2008), S08005–S08005.
DOI: 10.1088/1748-0221/3/08/s08005.

[12] The ALICE Collaboration. The ALICE experiment at the CERN LHC. In: *Journal of Instrumentation* 3.08 (2008), S08002.
DOI: 10.1088/1748-0221/3/08/S08002.

[13] The ALICE Collaboration. Performance of the ALICE experiment at the CERN LHC. In: *International Journal of Modern Physics A* 29.24 (2014), p. 1430044.
DOI: 10.1142/S0217751X14300440.

[14] The ALICE Collaboration. Upgrade of the ALICE Time Projection Chamber. Tech. rep. CERN-LHCC-2013-020; ALICE-TDR-016. 2014.
URL: https://cds.cern.ch/record/1622286.

[15] B. Abelev et al. Upgrade of the ALICE Experiment: Letter Of Intent. In: *Journal of Physics G: Nuclear and Particle Physics* 41.8 (2014), p. 087001.
DOI: 10.1088/0954-3899/41/8/087001.

[16] D. Boyanovsky. Phase transitions in the early and the present Universe: from the big bang to heavy ion collisions. In: *Proceedings of the International School of Astrophysics* (2001).
URL: https://arxiv.org/abs/hep-ph/0102120v2.

[17] A. A. Penzias and R. W. Wilson. A measurement of excess antenna temperature at 4080 Mc/s. In: *The Astrophysical Journal* 142 (1965), pp. 419–421.

[18] B. Abelev et al. Technical Design Report for the Upgrade of the ALICE Inner Tracking System. Tech. rep. CERN-LHCC-2013-024. ALICE-TDR-017. 2013.
DOI: 10.1088/0954-3899/41/8/087002.

[19] W. Reisdorf and H. G. Ritter. Collective flow in heavy-ion collisions. In: *Annual Review of Nuclear and Particle Science* 47.1 (1997), pp. 663–709.
DOI: 10.1146/annurev.nucl.47.1.663.

[20] K. Aamodt et al. Elliptic flow of charged particles in Pb-Pb collisions at 2.76 TeV. In: *Phys. Rev. Lett.* 105 (2010), pp. 252–302.
DOI: 10.1103/PhysRevLett.105.252302.

[21] The ALICE Collaboration. D Meson Elliptic Flow in Noncentral Pb–Pb Collisions at $\sqrt{s_{NN}} = 2.76$ TeV. In: *Phys. Rev. Lett.* 111.10 (2013), p. 102301.
DOI: 10.1103/PhysRevLett.111.102301.

[22] F. Prino and R. Rapp. Open heavy flavor in QCD matter and in nuclear collisions. In: *Journal of Physics G: Nuclear and Particle Physics* 43.9 (2016), p. 093002.
DOI: 10.1088/0954-3899/43/9/093002.

[23] A. Tauro. 3D ALICE Schematic Run 3 - with Description (2017).
URL: https://cds.cern.ch/record/2263642.

[24] N. Antoniou et al. Letter of Intent for A Large Ion Collider Experiment. In: CERN/LHCC (1993), pp. 93–16.
URL: http://cds.cern.ch/record/290825.

[25] The ALICE Collaboration. Technical proposal for A Large Ion Collider Experiment at the CERN LHC. LHC Tech. Proposal CERN-LHCC-95-71. (1995).
URL: http://cds.cern.ch/record/293391.

[26] The ALICE Collaboration. Technical Design Report of the Inner Tracking System. Tech. rep. CERN-LHCC-99-12. Geneva, 1999.
URL: https://cds.cern.ch/record/391175.

[27] The ALICE Collaboration. Technical Design Report of the Time Projection Chamber. Tech. rep. CERN-LHCC-2000-001. Geneva, 2000.
URL: https://cds.cern.ch/record/451098.

[28] The ALICE Collaboration. Technical Design Report of the Transition Radiation Detector. Tech. rep. Geneva, 2001.
URL: https://cds.cern.ch/record/519145.

[29] The ALICE Collaboration. The ALICE Transition Radiation Detector: construction, operation, and performance. In: *Nuclear Instruments and Methods in Physics Research Section A* 881 (2018), pp. 88–127.
DOI: 10.1016/j.nima.2017.09.028.

[30] The ALICE Collaboration. Technical Design Report of the Time of Flight System (TOF). Tech. rep. Geneva, 2000.
URL: https://cds.cern.ch/record/430132.

[31] J. Allen et al. Performance of prototypes for the ALICE electromagnetic calorimeter. In: *Nuclear Instruments and Methods in Physics Research Section A* 615.1 (2010), pp. 6–13.
DOI: 10.1016/j.nima.2009.12.061.

[32] J. Allen et al. ALICE DCal: An Addendum to the EMCal Technical Design Report Di-Jet and Hadron-Jet correlation measurements in ALICE. Tech. rep. CERN-LHCC-2010-011. ALICE-TDR-14-add-1. 2010.
URL: https://cds.cern.ch/record/1272952.

[33] The ALICE Collaboration. ALICE Photon Spectrometer (PHOS): Technical Design Report. Tech. rep. Geneva, 1999.
URL: https://cds.cern.ch/record/381432.

[34] F. Piuz et al. ALICE high-momentum particle identification: Technical Design Report. Tech. rep. Geneva, 1998.
URL: https://cds.cern.ch/record/381431.

[35] The ALICE Collaboration. Technical Design Report of the Zero Degree Calorimeter (ZDC). Tech. rep. Geneva, 1999.
URL: https://cds.cern.ch/record/381433.

[36] S. Blatt et al. Charge transfer of GEM structures in high magnetic fields. In: *Nuclear Physics B-Proceedings Supplements* 150 (2006), pp. 155–158.
DOI: 10.1016/j.nuclphysbps.2004.07.005.

[37] T. Abe et al. The International Large Detector: Letter of Intent. 2010.
DOI: 10.2172/975166.

[38] F. V. Böhmer, H. Angerer, et al. A Continuously Running High-Rate GEM-TPC for $\overline{\text{P}}$ANDA. In: *Nuclear Physics B-Proceedings Supplements* 215.1 (2011), pp. 278–280.
DOI: 10.1016/j.nuclphysbps.2011.04.030.

[39] The ALICE Collaboration. Upgrade of the ALICE Readout and Trigger System. Tech. rep. CERN-LHCC-2013-019. ALICE-TDR-015. 2013.
URL: https://cds.cern.ch/record/1603472.

[40] A. Schmah. private communication.

[41] The ALICE Collaboration. Technical Design Report for the Muon Forward Tracker. Tech. rep. CERN-LHCC-2015-001. ALICE-TDR-018. 2015.
URL: https://cds.cern.ch/record/1603472.

[42] The ALICE Collaboration. Technical Design Report for the Upgrade of the Online-Offline Computing System. Tech. rep. CERN-LHCC-2015-006; ALICE-TDR-019. 2015.
URL: https://cds.cern.ch/record/2011297.

[43] R. Brun and F. Rademakers. ROOT - An Object Oriented Data Analysis Framework. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389 (1997), pp. 81–86.
URL: https://root.cern.ch/.

[44] *CMake Software build system.*
URL: https://www.cmake.org/.

[45] *Boost C++ libraries.*
URL: https://www.boost.org/.

[46] G. Eulisse et al. Evolution of the ALICE Software Framework for Run 3. In: *EPJ Web of Conferences* 214 (Sept. 2019).
DOI: 10.1051/epjconf/201921405010.

[47] F. Costa, A. Kluge, P. Vande Vyvre, et al. The detector read-out in ALICE during Run 3 and 4. In: *Journal of Physics: Conference Series* 898 (Oct. 2017), p. 032011.
DOI: 10.1088/1742-6596/898/3/032011.

[48] The LHCb Collaboration. LHCb Trigger and Online Upgrade Technical Design Report. Tech. rep. CERN-LHCC-2014-016. LHCB-TDR-016. May 2014.
URL: http://cds.cern.ch/record/1701361.

[49] D. Rohr et al. Track Reconstruction in the ALICE TPC using GPUs for LHC Run 3. In: *4th International Workshop Connecting The Dots 2018 (CTD2018) Seattle, Washington, USA, March 20-22, 2018* (2018). arXiv: 1811.11481 [physics.ins-det].

[50] J. Wiechula. private communication.

[51] E. W. McDaniel and E. A. Mason. Mobility and diffusion of ions in gases. In: *Wiley series in plasma physics* (1973).
DOI: 10.1002/3527602852.

[52] W. Blum, W. Riegler, and L. Rolandi. Particle Detection with Drift Chambers. Vol. 2. Springer-Verlag, 2008.
DOI: 10.1007/978-3-540-76684-1.

[53] H. Appelshäuser et al. Space-charge distortion measurements and their calibration in the ALICE TPC. in preparation.
URL: https://alice-notes.web.cern.ch/node/863.

[54] E. Hellbär. The ALICE TPC: Optimization of the Performance in Run 2 and Developments for the Future. In: *Proceedings of Science, 7th Annual Conference on Large Hadron Collider Physics - LHCP2019* (2019).
DOI: 10.22323/1.350.0240.

[55] S. Rossegger and J. Thomas. Space-charge effects in the ALICE TPC: a comparison between expected ALICE performance and current results from the STAR TPC. Tech. rep. 2010.
URL: https://edms.cern.ch/document/1113087/1.

[56] M. Mager, S. Rossegger, and J. Thomas. The Langevin Equation expanded to 2nd order and comments on using the equation to correct for space point distortions in a TPC. Tech. rep. 2010.
URL: https://edms.cern.ch/document/1108138/1.

[57] R. E. Kalman. A new approach to linear filtering and prediction problems. In: *Journal of Basic Engineering* 82 (1960), pp. 35–45.
DOI: 10.1115/1.3662552.

[58] The ALICE Collaboration. Study of cosmic ray events with high muon multiplicity using the ALICE detector at the CERN Large Hadron Collider. In: *Journal of Cosmology and Astroparticle Physics* 2016.01 (2016), pp. 032–032.
DOI: 10.1088/1475-7516/2016/01/032.

[59] I. Kisel. Event reconstruction in the CBM experiment. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 566.1 (2006), pp. 85–88.
DOI: 10.1016/j.nima.2006.05.040.

[60] D. Rohr. Tracking performance in high multiplicities environment at ALICE. In: *arXiv preprint* (2017).
URL: https://arxiv.org/abs/1709.00618.

[61] The ALICE Collaboration. Real-time data processing in the ALICE High Level Trigger at the LHC. In: *Computer Physics Communications* 242 (2019), pp. 25–48.
DOI: 10.1016/j.cpc.2019.04.011.

[62] D. Rohr. private communication.

[63] I. Bird. Computing for the Large Hadron Collider. In: *Annual Review of Nuclear and Particle Science* 61 (2011), pp. 99–118.
DOI: 10.1146/annurev-nucl-102010-130059.

[64] W. Demtröder. Experimentalphysik 2. German. Vol. 2. Springer, 2005.
DOI: 10.1007/3-540-35146-9.

[65] R. Frühwirth et al. Data Analysis Techniques for High-Energy Physics. 2nd ed. Cambridge, UK: Cambridge University Press, 2000.

[66] P. Billoir. Error propagation in the helix track model. Tech. rep. 1987.
URL: https://cds.cern.ch/record/1330742/files/DELPHI-87-4.pdf.

[67] S. Brandt. Datenanalyse für Naturwissenschaftler und Ingenieure. German. 5th ed. Springer Spektrum, 2013.
DOI: 10.1007/978-3-642-37664-1.

[68] L. A. McGee and S. F. Schmidt. Discovery of the Kalman filter as a practical tool for aerospace and industry. Tech. rep. 1985.
URL: https://ntrs.nasa.gov/search.jsp?R=19860003843.

[69] R. Faragher. Understanding the basis of the Kalman Filter via a simple and intuitive derivation. In: *IEEE Signal processing magazine* 29.5 (2012), pp. 128–132.
DOI: 10.1109/MSP.2012.2203621.

[70] R. Frühwirth. Application of Kalman filtering to track and vertex fitting. In: *Nuclear Instruments and Methods in Physics Research Section A* 262.2-3 (1987), pp. 444–450.
DOI: 10.1016/0168-9002(87)90887-4.

[71] P. S. Maybeck. Stochastic models, Estimation, and Control Vol.1. Academic press, 1982.

[72] Y. Belikov et al. Kalman filtering application for track recognition and reconstruction in ALICE tracking system. Tech. rep. 1997.
URL: https://cds.cern.ch/record/689414/files/INT-1997-24.pdf.

[73] J. Klein. Jet Physics with A Large Ion Collider Experiment at the Large Hadron Collider. Doctoral thesis. Heidelberg, DE: University of Heidelberg, 2014.
URL: https://cds.cern.ch/record/1973326.

[74] M. O. Schmidt. Space-point calibration of the ALICE TPC with track residuals. In: *submitted to EPJ Web of Conferences* (2020).
URL: https://arxiv.org/abs/2003.03174.

[75] V. L. Ginzburg and I. M. Frank. Radiation of a uniformly moving electron due to its transition from one medium into another. In: *J. Phys.(USSR)* 9 (1945), pp. 353–362.

[76] J. de Cuveland. A Track Reconstructing Low-latency Trigger Processor for High-energy Physics. Doctoral thesis. Heidelberg, DE: University of Heidelberg, 2009.
URL: http://cds.cern.ch/record/1295509.

[77] J. Klein. Commissioning of and Preparations for Physics with the Transition Radiation Detector in A Large Ion Collider Experiment at CERN. Diploma thesis, Heidelberg University. 2008.

[78] V. Angelov. Design and performance of the ALICE TRD front-end electronics. In: *Nucl. Instrum. Meth.* A563 (2006), pp. 317–320.
DOI: 10.1016/j.nima.2006.02.169.

[79] J. Stiller. Full kinematic reconstruction of charged B mesons with the upgraded Inner Tracking System of the ALICE experiment. Doctoral thesis. Heidelberg, DE: University of Heidelberg, 2016.
URL: https://cds.cern.ch/record/2153601.

[80] H. Klingenmeyer. Tracklet-based particle identification with the ALICE TRD for LHC RUN 3. Master thesis. University of Heidelberg, 2017.

[81] M. O. Schmidt. Simulation and commissioning of the rejection of late conversions in the TRD triggers of ALICE. Master thesis. University of Heidelberg, 2016.

[82] V. Blobel and C. Kleinwort. A new method for the high-precision alignment of track detectors. In: *arXiv preprint* (2002).
URL: https://arxiv.org/abs/hep-ex/0208021.

[83] M. Ivanov et al. Track reconstruction in high density environment. In: *Nuclear Instruments and Methods in Physics Research Section A* 566.1 (2006), pp. 70–74.
DOI: 10.1016/j.nima.2006.05.029.

[84] R. Bailhache. Calibration of the ALICE Transition Radiation Detector and a study of Z^0 and heavy quark production in pp collisions at the LHC. Doctoral thesis. Darmstadt, DE: TU Darmstadt, 2009.
URL: https://cds.cern.ch/record/1320722.

[85] M. Richter. Development and Integration of on-line Data Analysis for the ALICE Experiment. Doctoral thesis. Bergen, NO: University of Bergen, 2009.
URL: https://cds.cern.ch/record/1331812.

[86] D. Rohr. GPU-based reconstruction and data compression at ALICE during LHC Run 3. In: *submitted to EPJ Web of Conferences* (2020).

[87] S. Meyers. Effective Modern C++. Vol. 1. O'Reilly, 2014.

[88] *SciPy Cookbook*. Accessed: 2020-04-06.
URL: https://scipy-cookbook.readthedocs.io/items/Least_Squares_Circle.html.

[89] G. Cowan. Statistical Data Analysis. Oxford university press, 1998.

[90] W. Press et al. Numerical recipes - The Art of Scientific Computing. Vol. 3. Cambridge University Press, 2007.

[91] D. Ruppert and D. Matteson. Statistics and Data Analysis for Financial Engineering. Vol. 2. Springer, 2011.
DOI: 10.1007/978-1-4419-7787-8.

[92] V. A. Epanechnikov. Non-parametric estimation of a multivariate probability density. In: *Theory of Probability & Its Applications* 14.1 (1969), pp. 153–158. DOI: 10.1137/1114019.

[93] M. Hanke-Bourgeois. Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens. German. Vol. 3. Vieweg+Teubner, 2002. DOI: 10.1007/978-3-8348-9309-3.

[94] J. C. Mason and D. C. Handscomb. Chebyshev polynomials. CRC press, 2002.