

Dissertation
submitted to the
Combined Faculty of Mathematics, Engineering and Natural Sciences
of Heidelberg University, Germany
for the degree of
Doctor of Natural Sciences

Put forward by
Martin Kroesen
Born in Solingen, Germany
Oral examination on February 15th, 2023

Reconstruction of Photon Conversions using the ALICE Transition Radiation Detector

Referees: Prof. Dr. Klaus Reygers
Prof. Dr. Norbert Herrmann

Abstract

Photons emitted in heavy-ion collisions are an important probe to study the quark–gluon plasma. In this thesis, the standard reconstruction of photon conversions into electron–positron pairs in the ALICE experiment using mainly the Time Projection Chamber (TPC) is expanded by using the Transition Radiation Detector (TRD). Because of the material budget in front and within the TRD, many conversions take place in its vicinity. To reconstruct these conversions, a stand-alone track reconstruction was implemented. The photon reconstruction was also made possible by a newly developed kinematic fit that uses automatic differentiation. This offers improved parameter estimation compared to other kinematic reconstruction methods. Using simulated and real proton–lead collision data at center-of-mass energies per nucleon–nucleon pair of $\sqrt{s_{\text{NN}}} = 5.02$ TeV, the functionality of the approach is proven by reconstructing neutral pions in the decay channel $\pi^0 \rightarrow \gamma\gamma$ via the photon conversion method. By reconstructing both photons in the TRD, as well as using a mixed mode with one photon in the TRD and one photon in the TPC, the neutral pion statistics could be improved by about 75% compared to using only photons in the TPC. The photons reconstructed in the TRD could be important for LHC Run 3, as the new continuous readout may cause difficulties for photon reconstruction in the TPC.

In addition, an outlook is given on the detection of other interactions in the material with the TRD using graph neural networks. In particular, it is shown that the TRD can be used to search for sexaquarks and to identify inelastic interactions of light antinuclei.

Zusammenfassung

Bei Schwerionenkollisionen emittierte Photonen sind eine wichtige Probe, um das Quark-Gluon-Plasma zu untersuchen. In dieser Dissertation wird die Standardrekonstruktion von Photonkonversionen in Elektron-Positron-Paare im ALICE-Experiment mittels Zeitprojektionskammer (TPC) um eine Rekonstruktion mittels des Übergangsstrahlungsdetektors (TRD) erweitert. Aufgrund des vorhandenen Materials finden viele Konversionen in der unmittelbaren Umgebung des TRDs statt. Um diese Photonen zu rekonstruieren, wurde eine unabhängige Spurrekonstruktion implementiert. Ermöglicht wurde die Photonrekonstruktion zudem durch einen neu entwickelten kinematischen Fit, der sich des Verfahrens des automatischen Differenzierens bedient. Dieser bietet im Vergleich zu anderen kinematischen Rekonstruktionen eine verbesserte Parameterschätzung. Anhand von simulierten und echten Proton-Blei-Kollisionsdaten bei einer Schwerpunktsenergie pro Nukleon-Nukleon-Paar von $\sqrt{s_{\text{NN}}} = 5.02$ TeV wird die Funktionalität durch eine Rekonstruktion neutraler Pionen im Zerfallskanal $\pi^0 \rightarrow \gamma\gamma$ mittels Photonkonversionsmethode gezeigt. Durch die Rekonstruktionen von beiden Photonen im TRD, als auch mittels eines gemischten Modus mit TRD und TPC, konnte die Statistik der neutralen Pionen um 75% gegenüber einer reinen TPC-Rekonstruktion erhöht werden. Die im TRD rekonstruierten Photonen könnten im LHC Run 3 von besonderer Bedeutung sein, da das kontinuierliche Auslesen der TPC Schwierigkeiten bei der Photonrekonstruktion verursachen könnte.

Zudem wird ein Ausblick auf den Nachweis anderer Interaktionen im Material mit dem TRD durch Graph Neural Networks gegeben. Insbesondere wird gezeigt, dass mit dem TRD Sexaquarks gesucht werden können und sich unelastische Interaktionen von leichten Antikernen identifizieren lassen.

Contents

1	Introduction	5
2	Physics Motivation	9
2.1	Quantum Chromodynamics and the Quark–Gluon Plasma	9
2.2	Heavy-Ion Collisions	12
2.3	Photons in Heavy-Ion Collisions	13
2.4	Other Probes	15
2.5	Inelastic Interactions of Light Nuclei	16
2.6	Sexaquark	16
3	Experimental Setup	19
3.1	The Large Hadron Collider	19
3.2	A Large Ion Collider Experiment	21
3.3	Transition Radiation Detector	24
4	Charged Particle Tracking	29
4.1	The Track Model	29
4.2	Interaction with Material	31
4.3	Kalman Filter	33
4.4	Helix Fits	36
4.5	Track Finding and Track Seeding	36
5	Reconstruction of Photon Conversions	39
5.1	The Physics of Photon Conversions	39
5.2	The Kalman Filter Particle Package	42
5.3	Reconstruction of Photon Conversions using the ALICE TPC	43
6	Dataflow and Differentiable Programming	49
7	Stand-Alone Tracking with the TRD	53
7.1	Tracklet Reconstruction	53
7.2	Data Set	55
7.3	Challenges of TRD Stand-Alone Tracking	57
7.3.1	Material Budget	58
7.3.2	Noise Tracklets	58

Contents

7.3.3	Empty Layers	59
7.3.4	Shared Clusters	60
7.3.5	Z Pad Resolution	60
7.3.6	Pad Tilting	62
7.4	Helix Fit	63
7.4.1	Fit Algorithm	63
7.4.2	Advantages	65
7.5	Combinatorial Search	67
7.6	Tracking Output Evaluation	69
7.7	Monte Carlo Performance	71
7.7.1	Matching	72
7.7.2	Efficiency	72
7.7.3	Momentum Resolution	74
7.7.4	Vertex Resolution	75
8	Synchronous Fitting of Daughter Tracks from Photon Conversions	77
8.1	Concept	77
8.2	Technical Implementation	79
8.2.1	Stage 2	79
8.2.2	Stage 1	81
8.2.3	Regularization	82
8.2.4	Pad Tilting	82
8.2.5	Multiple Scattering	83
8.3	Syncfitter versus KFParticle	83
8.4	Performance	84
8.5	Detector Effects	85
8.5.1	Bremsstrahlung	86
8.5.2	Pad Tilting	86
9	Measurement of Photons from Conversions in the TRD	89
9.1	Preselection	89
9.2	Post Cleaning	90
9.3	Performance using Full Monte Carlo Simulation	92
9.3.1	Conversion Vertices	92
9.3.2	Efficiency of reconstructable candidates	94
9.3.3	Purity	95
9.3.4	Transverse Momentum Resolution	97
9.4	Performance using real p-Pb data at $\sqrt{s_{NN}} = 5.02$ TeV	99
9.5	π^0 Reconstruction	102

10 Identification of Inelastic Interactions of Light Antinuclei and the Sexaquark using Graph Neural Networks	107
10.1 Introduction to Graph Neural Networks	107
10.1.1 Graphs	108
10.1.2 Neural Networks	108
10.1.3 Graph Neural Networks	110
10.2 Identification of Inelastic Interactions of Light Antinuclei	112
10.2.1 Dataset	113
10.2.2 Model and Workflow	113
10.2.3 Performance	115
10.3 Identification of the Sexaquark	117
10.3.1 Dataset	118
10.3.2 Model and Workflow	119
10.3.3 Performance	120
11 Summary and Outlook	123
A Syncfitter Momentum Resolution Estimation	127
B Pion Invariant Mass Fits	129
Bibliography	131

1 Introduction

Among the four large experiments at CERN's Large Hadron Collider (LHC) [1], A Large Ion Collider Experiment (ALICE) [2] is the dedicated heavy-ion experiment. The main research interest of ALICE is the state of strongly interacting matter at a high energy density, the quark–gluon plasma (QGP). This state of matter is created via high energy nucleus–nucleus collisions. An insight into the dynamics of these collisions is provided by the emerging particles. To detect them, a detector system is provided that is capable of delivering track finding and parameter estimation in a high multiplicity environment especially in the MeV/ c to GeV/ c transverse momentum range and also particle identification. An important component of this detector system is the Transition Radiation Detector (TRD) [3], which delivers electron triggering and identification capabilities and also contributes to the full tracking and calibration of the central barrel.

So-called direct photons give us an undisturbed insight into the quark–gluon plasma. Direct photons are defined as those that do not originate from particle decays. Besides using calorimeters, these are reconstructed using the Photon Conversion Method (PCM) via electron–positron pairs which originate from photon conversions in the detector material. Compared to the calorimeter measurements, the photon conversion method provides a higher energy resolution in the region of interest, i.e. also in the region $1 \lesssim p_T \lesssim 3$ GeV/ c , in which thermal photons are expected to give a significant contribution [4]. A huge statistics is needed for this analysis since the background from the two-photon decays of neutral pions and η mesons is large and the reconstruction efficiency via the photon conversion method relatively low.

The TRD, which is located at a radial distance from the beamline of about 3 m, contains a material budget with a mean radiation length of $\langle X/X_0 \rangle \approx 25\%$. In addition there are also mechanical structures with a significant material budget right in front of the TRD. In front of the TPC, the mean radiation length of the material budget is below 15%. Naturally, this leads to a high conversion probability and thus to many signals in the TRD from the electron–positron pairs created in photon conversions. A detection of these conversions can increase the total photon efficiency. It also serves as a cross-check, and it provides a natural insight into the distribution of the material. The material distribution is also relevant for analyses that use the time-of-flight measurement from behind the TRD. Furthermore, the recent upgrades of the

1 Introduction

ALICE Time Projection Chamber (TPC) [5] with continuous readout deliver uncertainties for reconstruction capabilities in the future. A clear assignment of the TPC tracks to certain events is not possible anymore, which enhances the combinatorics to find candidates for photon conversions. The electron and positron candidates from fake photon conversions may be created in different events. In addition, if requiring hits in the Inner Tracking System (ITS) to reduce this effect, only photon conversions happening in the first ITS layers can be considered. In these layers the material budget is significantly reduced and thus also the conversion probability.

Up to now, the TRD is only used in conjunction with the other detector systems, i.e. the TPC. Thus, the current tracking is not capable of finding the conversion electrons and positrons from the material within and right before the TRD. In this work, a full stand-alone tracking procedure with this detector is presented that is capable of enhancing the photon reconstruction capabilities of ALICE. The maximal limits of the photon reconstruction efficiency are tested by using a calibration procedure that is not limited by the readout capacities of the TRD, i.e. the currently used online tracking. Lower momenta can be achieved by also allowing a larger inclination of the track segments called tracklets. Thus, the photon analysis using the TRD presented in this thesis is a feasibility study.

The development of tracking methods in this work was highly influenced by modern computational approaches. In recent years, great progress in the field of so-called Artificial Intelligence raised a lot of attention [6]. Artificial Neural Networks revolutionized among other things image and language processing. One driving force behind these advances is dataflow programming, where the program is interpreted as a directed graph of the data flowing between the operations. Graphic cards provide the hardware for efficient calculation of vectorized programs. Besides that, differentiable programming played a crucial role. Complex minimization procedures can now be employed highly efficient with lower effort than ever before and is highly stable. In a recent white paper, Tommaso Dorigo et al. argue that ‘the most relevant recent advancement for particle detector development is the rise of differentiable programming’ [7]. Sophisticated software development tools such as TensorFlow [8] and PyTorch [9] accelerate this process as they allow the developer to focus on the model design.

These side products of modern neural networks are exploited in this work for the track finding and parameter estimation procedure. A combinatorial track search in combination with helix fits is presented. For the reconstruction of photon conversions, the exploitation of differential programming capabilities is driven even further. A mathematical model of the photon conversion topology is embedded into

a differentiable framework, and thus a very precise fit is performed. In contrast to common frameworks for kinematic fitting, fewer approximations need to be made, which shows up in improvements of the fit quality.

Photons are not the only interesting particles that interact with the material resulting in daughter particles. In the collisions studied by ALICE, particles are produced that are not only interesting to deliver more insights into the QGP. Light nuclei for instance interact with the material. Thereby the cross section is an important measurement [10]. This cross section is important to get insights into the generation of nuclei in the galaxy that might arise from dark matter. A first measurement of the absorption of antihelium-3 in ALICE is presented in [10]. Here, the disappearance of the antinuclei in the TRD is calculated by using the surrounding detectors, but not the signal in the TRD explicitly. Besides that also dark matter candidates like the hypothetical Sexaquark (S) might be generated directly in the collisions within ALICE [11]. These also might interact with the material within and in front of the TRD. A first search for the \bar{S} with CMS is presented in [12]. The reconstruction suffers from a low efficiency that might be improved with ALICE, i.e. also using the signals of the TRD. Thus, these two examples are further interesting interactions to be detected with the TRD. A first detection approach for these two particle types is presented in this work.

Due to the large amount of fragmentation to daughter tracks it often happens that the nuclei and the sexaquark interactions do not result in a sufficient number of tracklets to be reconstructed with the stand-alone algorithm. The local spatial structure of the relation of the tracklets can be formulated as graphs. Analog to the image classification with convolutional neural networks, graph neural networks can efficiently operate on these structures [13]. First reconstruction approaches are presented in this work.

2 Physics Motivation

The underlying experiment, the ALICE experiment, is the dedicated heavy-ion collision experiment at the LHC. Thereby, quantum chromodynamics and especially the quark–gluon plasma and its thermodynamic properties are the main interest. The theoretical basics of these topics are introduced within this chapter. A focus is put on photons as a probe, as they are of particular interest within this work. Besides studying the quark–gluon plasma, ALICE is also well equipped to study nuclei and exotic particle yields. As subsidiary topics within this work, nuclear interaction and a proposed exotic particle state, the sexaquark, are also introduced in this chapter.

2.1 Quantum Chromodynamics and the Quark–Gluon Plasma

Until in 1963 Gell-Mann and Zweig found a first ordering principle [14][15], a variety of different particles were found from cosmic rays and in accelerator experiments. This mess was known as particle zoo. With the eightfold way Murray Gell-Mann and Yuval Ne’eman proposed a first quark model to describe the structure and some particle properties in 1961 [16]. Consequently, Gell-Mann and Zweig combined the particle mess to a model for strongly interacting particles later called quarks. Initially, Zweig called them aces. Counter-intuitively, free quarks with fractional electrical charge could not be found and serious symmetry issues for baryons appeared [17]. The baryons require the assumption that the wave function is totally symmetric under interchange of quark spin and flavor quantum numbers, contrary to the expectation that quarks are fermions with spin $1/2$. Han and Nambu as well as Greenberg and Gell-Mann introduced a new quantum number, the color charge, that solved the given issues [18]. Inspired by the Yang-Mills Lagrangian, a new $SU(3)$ global symmetry was introduced. This theory is today called Quantum Chromodynamics (QCD).

QCD is a gauge field theory and the eight traceless generators can be associated with eight colored gluons, the corresponding exchange particle of the given force [17]. While the quarks are spin $1/2$ fermions, the gluons are spin 1 bosons. The non-Abelian QCD Lagrangian includes a self coupling between the gluons. As every Yang-Mills Lagrangian, QCD represents an asymptotically free theory. Asymptotic

2 Physics Motivation

free means that in the limit of very small distances between the quarks the quark-quark interaction vanishes. This was supported by electron-proton and neutrino-proton scattering experiments. In the deep inelastic scattering regime so-called parton scattering was observed. These partons are interpreted as quarks nowadays.

With the observation of the J/Ψ [19][20] a new quark flavor was found, which was postulated earlier by the GIM mechanism (or Glashow-Iliopoulos-Maiani mechanism) [21]. In the following the standard quark model was completed with three families of fermions and six flavors of quarks belonging to the three families and the gauge bosons of the electroweak and strong forces. Another interesting observation of the existence of the J/Ψ is the similarity of its spectrum to the positronium (e^+e^-) spectrum.

In contrast to Quantum Electrodynamics (QED) the calculation of observables in the context of QCD is more complex. In the high energy regime, the coupling constant becomes small and perturbative methods are feasible. This type of calculations is called perturbative QCD (pQCD). In the regime of strong coupling an approximation scheme introduced by Wilson leads to a statistical mechanical system. These methods are called lattice QCD (lQCD) [22][23]. For example, mass predictions of quark-anti quark pairs or the proton are made by lQCD methods. It is noteworthy that the gluons only couple to the quarks and themselves.

Experimentally, free quarks were not observed. As a consequence, a concept that does not allow unconfined color charges like quarks was introduced. This is concept is known as confinement. It is a natural property of the QCD Lagrangian. The confinement of the quarks can be described by the phenomenological quark-antiquark potential:

$$V_{q\bar{q}} = -\frac{4}{3} \frac{\alpha_s(r)}{r} + k \cdot r. \quad (2.1)$$

The distance between the quarks q and \bar{q} is given by r , α_s is the coupling of the strong interaction and k a proportionality factor known as "string tension". The first term describes the quark-antiquark interaction at very small distances for pointlike particles. The second term is an effective description for long distances due to the self-coupling of the gluons. An infinite amount of energy is necessary to separate the particles. This potential successfully describes the masses of bound quark-antiquark states, e.g. J/Ψ . The effective coupling constant of QCD α_s cannot be defined for on-shell quarks, because they do not physically exist. Thus, the coupling constant depends on an arbitrary renormalization point. This is removed using a mass scale Λ and then α_s can be written as [24][25]

$$\alpha_s(Q^2) = \frac{12\pi}{(33 - 2N_f) \cdot \ln(Q^2/\Lambda^2)}, \quad (2.2)$$

2.1 Quantum Chromodynamics and the Quark–Gluon Plasma

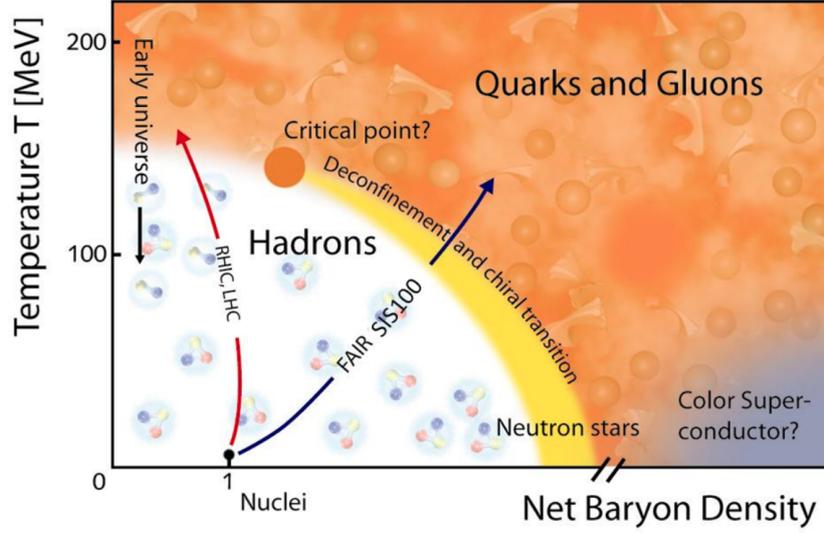


Figure 2.1: The QCD phase diagram with the temperature versus the net baryon density. At the LHC the high temperature region at low net baryon densities is explored. High baryon densities at low temperatures may occur in neutron stars. Figure taken from [30].

where N_f is the number of quark flavors ($N_f \leq 6$) and Q the momentum transfer. In the limit of infinite momentum transfer the coupling vanishes.

Very high temperatures and/or high net baryon densities lead to a state of matter where confinement and chiral symmetry breaking are no longer present [26][27][28]. Different assumption can motivate this state of matter: On the one hand, it can be assumed that if two or more hadrons are very close to each other, i.e. in the order of their size, the quarks cannot be uniquely assigned to one of the hadrons. On the other hand, if a temperature is reached above the mass of the pion, the lightest hadron, new hadrons are created. This was first mentioned by Hagedorn [29]. It should be mentioned, that this does not imply an upper limit of the temperature. This new state of matter, which is interpreted as a soup of quark and gluons, is known as Quark–Gluon Plasma (QGP). It has an increase of degrees of freedom arising from the different quantum states, e.g. flavors, colors, spins, and polarization states of quarks and gluons. Due to their low mass, the dominant quark flavors are up, down and strange quarks, described by the so-called (2+1)-flavor QCD [28]. Other known flavors like the charm quark also appear in the QGP due to hard partonic scattering processes.

In Fig. 2.1 the QCD phase diagram is shown. The QGP is understood in the

2 Physics Motivation

framework of a thermodynamic system. Thus, it is described by the temperature and net baryon density. As can be seen in the figure, at very low net baryon densities, the transition temperature is close to the pion mass. This is in agreement with theoretical predictions of lattice QCD, that predict the temperature as $T_C = 155 \pm 1(\text{stat.}) \pm 8(\text{syst.}) \text{ MeV}$ [31]. With an increase of the net baryon density, the deconfined state of matter also occurs at lower temperatures. At the Large Hadron Collider (LHC) [1] as well as the Relativistic Heavy Ion Collider (RHIC) [32][33][34][35] the low baryon density limit is explored. At the Facility for Antiproton and Ion Research (FAIR) [36], which is currently under construction, it is aimed to reach higher baryon densities. The transition at vanishing net baryon density is not of first or second order, but it is a continuous crossover [37]. Moreover, the QGP is relevant for the astronomical interpretation of the development of the universe, as it is expected that the universe developed through this kind of state of matter within the first fractions of a second. The experimental access is described in the next section.

2.2 Heavy-Ion Collisions

Heavy-ion collisions have been found to be an appropriate tool to access the QGP experimentally. First heavy-ion collisions with several GeV were performed at the Brookhaven National Laboratory (BNL) with the accelerator AGS and at CERN with the SPS. Clear indications for a QGP were then found with the follow-up accelerators RHIC with $\sqrt{s_{\text{NN}}} = 200 \text{ GeV}$ using gold ions and with the LHC at TeV center-of-mass energies using lead ions. These accelerators are also the first colliders used for heavy-ion collisions. At the corresponding energies the lead ions are contracted due to the high relativistic gamma factor to flat "pancakes" perpendicular to the beam line. Thus, very high energy densities are generated. The collision overlap varies and thus the collisions can be referred to different centrality classes.

The space-time evolution of a collision is shown in Fig. 2.2. While the pancakes collide, in a very short thermalization phase ($\tau_0 < 1 \text{ fm}/c$) QCD processes constitute an equilibrated QGP. As it is spatially not constrained, an expansion takes place that is determined by the inner pressure given by the collision geometry and the temperature. At the LHC, the lifetime of the QGP is estimated to be about $10 \text{ fm}/c$ [39]. After reaching a certain pseudocritical temperature a continuous crossover region leads to the hadronization. For a short time period the hadrons establish a hadron gas before a kinetic freeze-out happens and observable particles leave the interaction region.

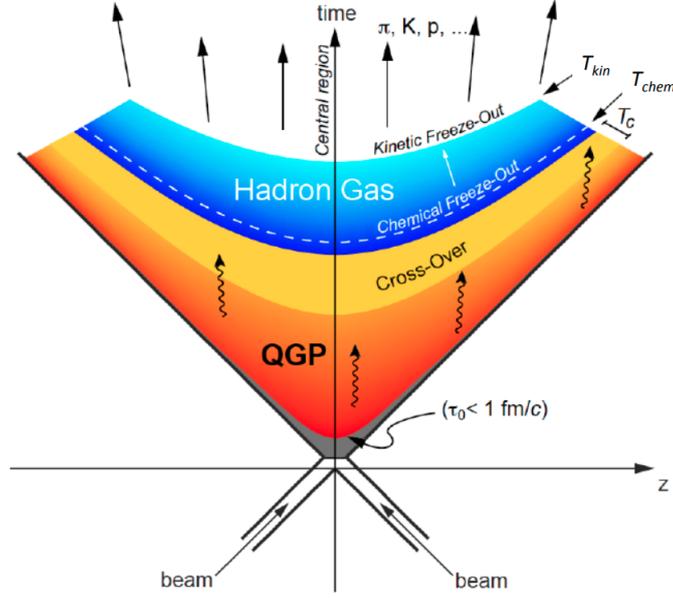


Figure 2.2: The space-time evolution of a heavy-ion collision is visualized in the light cone. The different stages consist of the thermalization, the QGP, the cross-over region and finally the hadron gas which kinetically freezes out to the measurable particles. Figure taken from [38].

2.3 Photons in Heavy-Ion Collisions

As the reconstruction of photons is the main topic within this work, a theoretical insight into photon production in heavy-ion collisions is given as a first example of a probe of the dynamics of the collision. The electromagnetic nature of the photon implies that there is no strong coupling. The electromagnetic coupling constant α_{em} is much smaller than the (running) coupling constant of the strong interaction α_s . This leads to a mean free path length that is long enough to leave the plasma without significant interaction [41]. The photon production thus gives a unique insight into the evolution of the collision.

The photons from heavy-ion collisions are structured by production processes according to the evolution. In the following the main sources of direct photons are specified [42]. Prompt photons are produced in initial interactions of the partons from the colliding nuclei. These photons are also called pQCD photons as the production can be calculated via perturbative QCD calculations [43][44]. Before reaching thermodynamic equilibrium photons are also expected to be produced: the pre-equilibrium photons [45]. Then the thermal phase begins. Here, a hydrodynam-

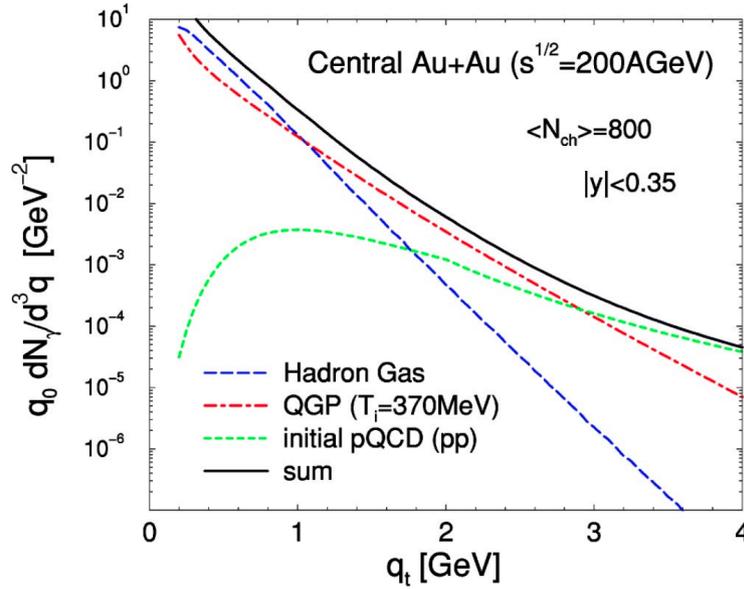


Figure 2.3: The invariant photon yield as a function of the transverse momentum. With the models, different parts of the spectrum can be assigned to different stages of the time evolution of the heavy-ion collision process. Figure taken from [40].

ical description of the evolution is usual [40]. Photons produced during this phase reflect the temperature of the medium and they are called thermal photons.

In reality, these direct photons are overlaid by decay photons, for example from neutral pion and eta decays. Given that the yield of all photons γ_{inc} (inclusive photons) is measured, the yield of the direct photons is defined by

$$\gamma_{\text{dir}} = \gamma_{\text{inc}} - \gamma_{\text{dec}}, \quad (2.3)$$

where γ_{dir} is the direct photons yield and γ_{dec} is the decay photons yield. Experimentally one extracts the double ratio R_γ defined as

$$R_\gamma = \frac{\gamma_{\text{inc}}}{\gamma_{\text{dec}}} = \frac{(\gamma_{\text{inc}}/\pi^0)_{\text{meas}}}{(\gamma_{\text{dec}}/\pi^0)_{\text{calc}}} \quad (2.4)$$

in order to minimize systematic uncertainties. It follows

$$\gamma_{\text{dir}} = \left(1 - \frac{1}{R_\gamma}\right) \gamma_{\text{inc}}. \quad (2.5)$$

The via a Monte Carlo approach calculated decay photon spectrum γ_{calc} is based on the measured π^0 spectrum. With this ratio an access to the direct photon yield is

given.

An example of the expected direct photon spectrum in Au–Au collisions is shown in Fig. 2.3. Here, the invariant yield is plotted versus the transverse momentum. The yield as given by hydrodynamical calculations for the hadron gas and QGP as well as pQCD photons are separately shown. At low transverse momenta, the hadron gas dominates, in the range between 1 GeV/ c and 3 GeV/ c the thermal photons dominate, and at high p_T most photons are created by the initial pQCD phase as expected. For the hydrodynamical calculations it is worth to mention that the expansion of the fireball leads to a blueshifted photon spectrum which has to be considered for the slope of the photon spectrum. A further observable is the elliptic flow [46]. By using the Fourier series of the spectrum and taking the second order coefficient v_2 from $dN/d\phi \propto N(1 + 2v_2 \cos(2\phi))$ a description of the ellipticity is provided. Measurements show a v_2 that is unequal to zero. Thus, the observation of v_2 also delivers insights into the dynamics. Hydrodynamical models also show a v_2 unequal to zero.

A simultaneous description of the spectrum and elliptic flow by the models is still not perfectly given. This tension is referred to as the photon puzzle.

2.4 Other Probes

As the neutral pions mainly decay to two photons, a reconstruction for the direct photon measurement is necessary. But besides providing a background for the photon measurement, they are also interesting as a probe by themselves.

Usually, the yield in AA collisions N_{AA} and in pp collision are compared using the nuclear modification factor R_{AA} . This factor is defined by [47]

$$R_{AA} = \frac{d^2 N_{AA}/dp_T dy}{\langle T_{AA} \rangle d^2 \sigma_{pp}/dp_T dy}, \quad (2.6)$$

where σ_{pp} is the cross section in pp collisions and $\langle T_{AA} \rangle$ is the nuclear overlap function. The nuclear overlap would scale the pp collision to an AA collision if it was an independent superposition of pp collisions. Thus, a value of R_{AA} below 1 indicates a suppression in the yield induced by nuclear effects. This modification factor is also considered for other hadronic spectra.

There are also many further observables of heavy-ion collisions including soft probes originating from processes with low momentum transfer and hard probes from initial hard parton scattering.

2.5 Inelastic Interactions of Light Nuclei

Besides studying the quark–gluon plasma, the given collisions also provide particle yields of light nuclei and antinuclei such as protons, deuterons or helium-3 and the corresponding antiparticles [10].

These nuclei are also produced by processes in the universe. For a better understanding of these processes, precise measurements of the nuclear interaction cross sections are needed. Heavy-ion collisions and the corresponding reference systems provide such particle yields.

Dark matter candidates that may exist in the galaxy are expected to produce antinuclei in annihilation processes that peak at low energies with $E_{\text{kin}} \lesssim 1 \text{ GeV}/A$ [48]. As the spectrum of produced antinuclei from cosmic rays is much larger, a nearly background free signal of antinuclei yields from dark matter can be expected in the GeV range.

These generated antinuclei travel a long path through our galaxy and thus many nuclear interactions appear. For the corresponding models the cross section for nuclear interactions play a crucial role, as it reduces the yield that can be measured close to earth. Thus, it is important to have precise measurements of the inelastic cross sections of these particles.

2.6 Sexaquark

Besides producing nuclei, collisions at the LHC could also produce exotic particles. A candidate is the Sexaquark (S), which was proposed by G. Ferrar in [49]. Heavy-ion collisions are also a possible good production environment [11]. The sexaquark is a compact bound state with the quark content $uudds$. A recent overview is given in [11]. Other multiple quark states were already observed, for instance the pentaquark state with $ccuud$. In contrast to the sexaquark the found states involve no pure quark or antiquark combinations.

Already in 1976 R. Jaffe proposed the existence of an H-dibaryon [50], a composition of two weakly bound uds hyperons. Despite searches with many experiments no evidence yet exists.

The sexaquark has the following properties: spin 0, flavor singlet, parity even, not charged, baryon number $B = 2$, strangeness $S = -2$. Lattice QCD is not capable of predicting the mass precisely nowadays and empirical assumption provide a

large range for it. For effective stability a mass below $2.054 \text{ GeV}/c^2$ is required [11]. The stability of the sexaquark corresponds to the possibility to be a dark matter candidate. This increases the relevance for the search for it.

3 Experimental Setup

In the following chapter, the experimental setup is described. First, the Large Hadron Collider (LHC) is introduced which provides the collision beams. The experiment of interest is the ALICE experiment. The most important detectors of ALICE are explained before describing the transition radiation detector in more detail.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [1] is a circular particle collider which has a circumference of 26.7 km and it is placed approximately 100 m under ground. It is the world's largest accelerator and provides the highest energies. The LHC is placed in the former Large Electron–Positron Collider (LEP) tunnel within the CERN accelerator complex. It is located close to Geneva and crosses the border of France and Switzerland. A sketch including its interaction points with the experiments and its pre-accelerators is shown in Fig. 3.1. It provides proton beams with an injection that starts at the LINAC 2 and lead beams with an injection starting at the LINAC 3. The lead ions are stripped during two stages until they are fully ionized. The pre-accelerator chain ending with the Super Proton Synchrotron (SPS) allows injection energies of 450 GeV for proton beams and 177 GeV per nucleon for lead beams. The LHC itself provides proton–proton beams with a center-of-mass energy of up to 14 TeV. The energy of the injected $^{208}_{82}\text{Pb}^{82+}$ ions is scaled corresponding to the charge per mass. The peak luminosity of proton–proton collisions is $10^{34} \text{ cm}^{-2}\text{s}^{-1}$, which is also the highest value provided from any existing accelerator.

In total, eight straight sections in the LHC exist, either for experiments or beam tuning. The points are named in numerical order from Point 1 to Point 8. At four of these places big experiments are located: ATLAS [51], CMS [52], LHCb [53] and ALICE. At Point 2 the experiment of interest is placed: ALICE. This is the only dedicated heavy–ion research experiment at the LHC. But at the other facilities also heavy-ion programs exist.

3 Experimental Setup

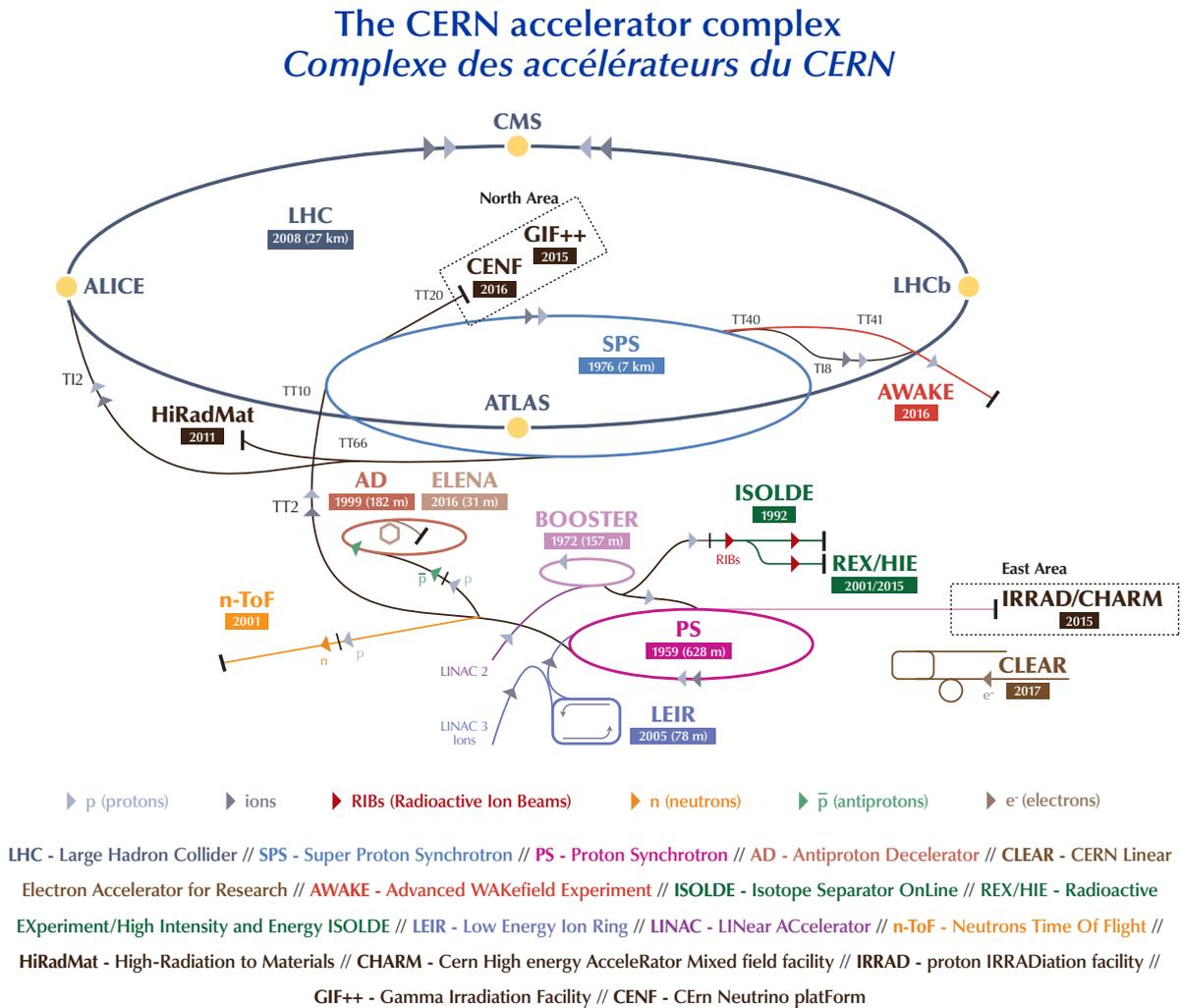


Figure 3.1: A sketch of CERN’s whole accelerator complex, including the LHC [54]. The position of ALICE is indicated as well.

3.2 A Large Ion Collider Experiment

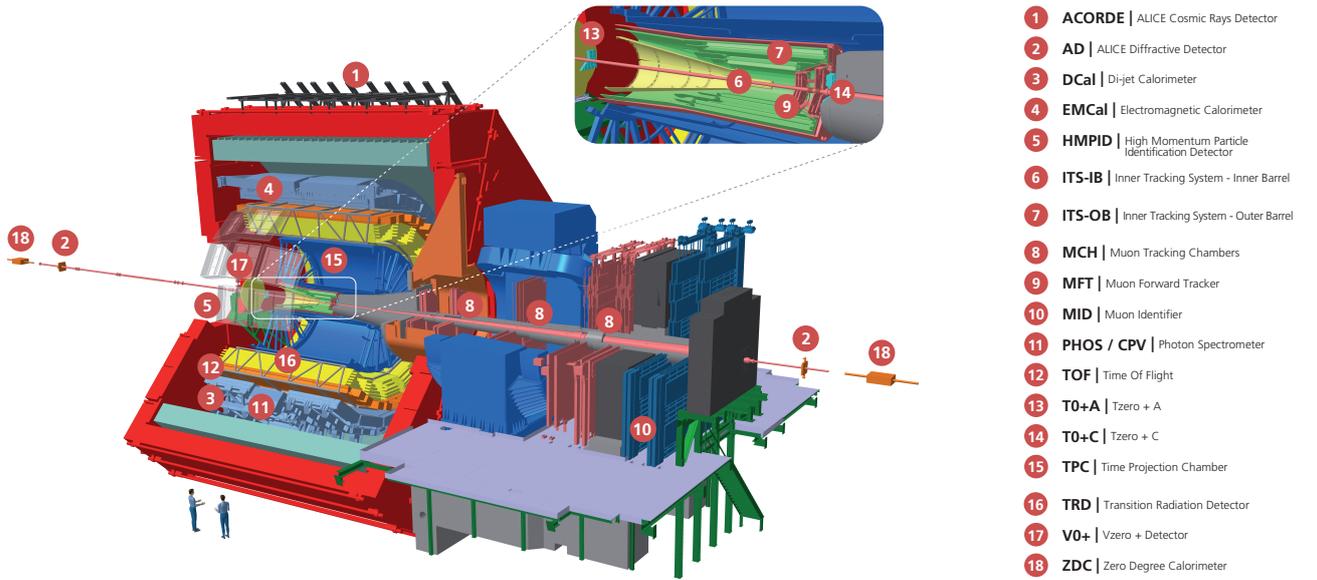


Figure 3.2: A sketch of the ALICE detector system as it is given in Run 3 including the central barrel detectors as well as the forward muon arm. Figure taken from [55].

3.2 A Large Ion Collider Experiment

ALICE is an acronym for A Large Ion Collider Experiment. It is designed to focus on studies of the properties of the color deconfined state of strongly interacting matter called Quark–Gluon Plasma (QGP) created in heavy-ion collisions. These studies require a reconstruction of events with high multiplicities and a good resolution down to low momenta ($p_T \approx 150 \text{ MeV}/c$) where many particles in heavy-ion collisions are produced. These requirements are fulfilled by ALICE [2].

The central barrel detectors surrounding the nominal collision point include the Inner Tracking System (ITS), the Time Projection Chamber (TPC), the Transition Radiation Detector (TRD) and the Time-Of-Flight (TOF) detector. They are shown in Fig. 3.2 together with the other detectors. The solenoid magnet provides a 0.5 T magnetic field parallel to the beam axis. At high pseudorapidities smaller detectors are placed that are used for global event characterization and triggering. Furthermore, the EMCal/DCal (Di-Jet Calorimeter) and PHOS (Photon Spectrometer) are used for photon reconstruction. The relevant detectors for this work are explained in the following. Information regarding the remaining detectors can be found in [56] as given for Run 2 and information regarding upgrades of the readout

3 Experimental Setup

and trigger system can be found in [57].

The **ITS** [58] is the detector closest to the interaction point. Until the upgrade for Run 3 it consisted of six layers of silicon detectors with three different technologies placed at radii between 3.9 cm and 43.0 cm. The first two layers constitute the Silicon Pixel Detector (SPD). Operating closest to the interaction point it provides high precision and granularity which is important for determining primary and secondary vertices. The intermediate layers of the ITS form the Silicon Drift Detector (SDD) which combine a very good multi-track reconstruction capability with dE/dx information. The last two layers, the Silicon Strip Detector (SSD), also provide dE/dx information. The energy loss signals can be used for particle identification (PID) and provide information for low-momentum particles.

Recently, the ITS was upgraded for Run 3 [59]. The different mentioned technologies were replaced by Monolithic Active Pixel Sensors (MAPS). The new ITS consists of seven layers. Three of them are close to the vertex at 2.3 cm, 3.2 cm and 3.9 cm. The inner layers have a reduction of the material budget to 0.35% X_0 compared to 1.14% X_0 before. This reduces the conversion probability of photons, which makes the standard PCM measurement more difficult. The other four layers are between 19.6 cm and 39.3 cm. The new ITS is closer to the primary vertex, has less material budget and allows a higher readout rate equipped for a continuous readout. There will be no energy loss information anymore.

The **TPC** [60] surrounds the ITS and has an overall volume of nearly 90 m³. A sketch of the TPC and the measured signals are shown in Fig. 3.3. A crossing particle traverses and ionizes the gas in the environment of the magnetic field. The ionization charges drift under the influence of an electric field to the endplates of the TPC. There, the arrival point and time can be precisely measured and so the complete trajectory be reconstructed. Until the upgrades of Run 3 the readout was based on Multi-Wire Proportional Chambers (MWPC). These MWPC require a gating grid due to the backfloating ions which limits the readout frequency. Thus, for Run 3 the MWPC were replaced by Gas Electron Multipliers (GEMs) that allow a readout of 50 kHz for Pb–Pb collisions [61]. The GEMs provide intrinsic ion blocking capabilities. An ion backflow of about 1% still remains [62].

The design of the TPC delivers the capabilities to be the main device for tracking of charged particles and particle identification via the measurement of specific ionization energy loss in the central barrel. For PID, the mass dependence of the energy loss given by the Bethe Bloch formula is exploited. More on the Bethe Bloch formula can be found in Chapter 4. The mass dependency is also shown in Fig. 3.3. The resolution of the specific energy loss measurement is about 5% [56].

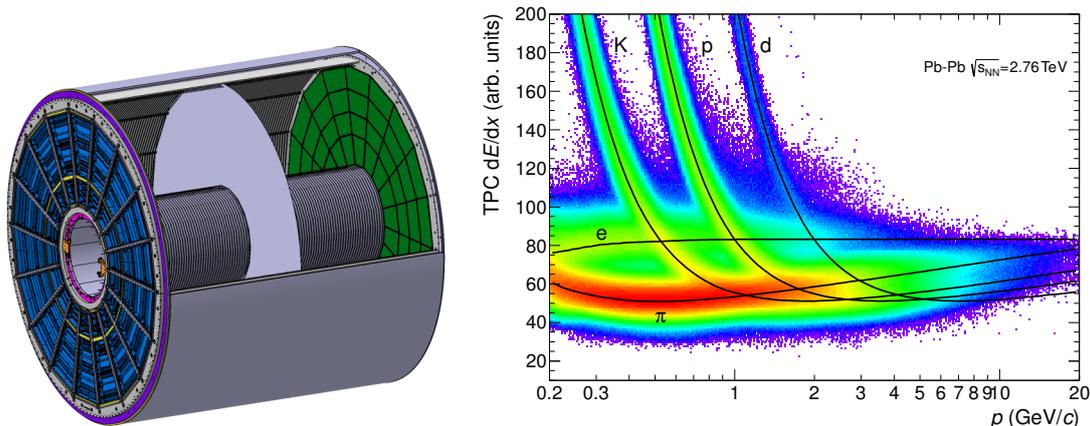


Figure 3.3: Left: The central electrode, the inner and outer field cage and the end plates with readout chambers form the TPC [5]. Right: The TPC provides particle identification via the measurement of the dE/dx information of crossing particles [56].

The **TRD** [3] provides enhanced electron identification by exploiting specific ionization energy loss and transition radiation. In addition, it can be used as a trigger to enhance rare probes that e.g. include an electron in the decay. The TRD standalone capabilities for the reconstruction of photon conversions and other material interactions are discussed in this work. Since the emphasis of this work lies on the TRD, it will be explained in detail in the next section.

The **TOF** [63] covers the momentum region from 0.5 to 2.5 GeV/ c for pion/kaon separation and up to 4 GeV/ c for pion/proton separation via time-of-flight measurements using Multi-gap Resistive Plate Chambers (MRPCs). This provides particle identification in the crossing regions of the deposited charge in the TPC for electrons and kaons or protons, respectively. The overall TOF resolution is 80 ps for pions with a momentum around 1 GeV/ c .

In Run 2 on each side of the detector a scintillation counter called **V0** was placed [2]. They are used for triggering as well as collision centrality estimation. The Cherenkov detector **T0** was located on both sides, allowing a time resolution of 20-25 ps in Pb-Pb collisions and around 40 ps in pp collisions [56]. This served as the start time for the TOF measurement. In Run 3, a Fast Interaction Trigger (**FIT**) detector combines Cherenkov arrays (T0+) and scintillator rings (V0+) as fast minimum bias trigger [57]. Furthermore, there is the Zero Degree calorimeter (**ZDC**) consisting of two tungsten-quartz neutron and two brass-quartz proton calorimeters also in both

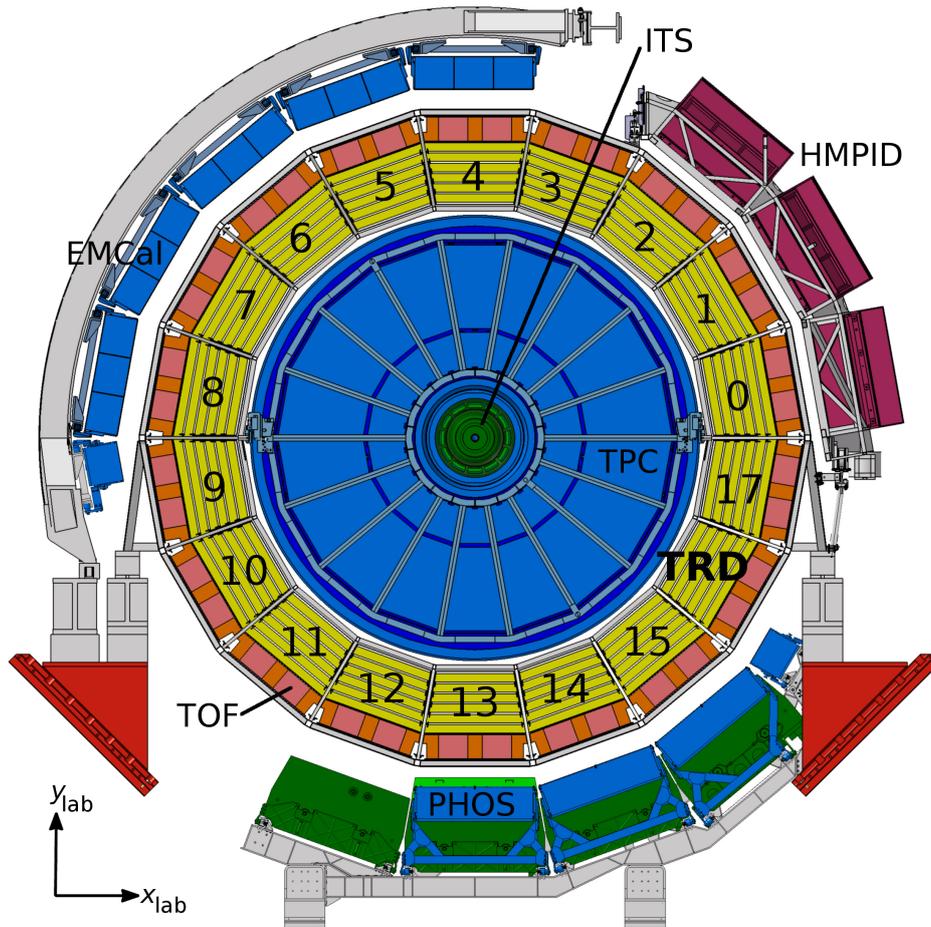


Figure 3.4: The TRD cross section with its 18 supermodules (yellow) surrounding the TPC. Figure taken from [3].

sides in beam direction to detect spectator nucleons [56].

3.3 Transition Radiation Detector

Details concerning the TRD can be found in [3]. The TRD was built to improve the electron identification at relatively large momenta ($p > 1 \text{ GeV}/c$) where the separation via specific ionization energy loss in the TPC worsens. The specific energy loss dE/dx via charge deposition is described by the Bethe-Bloch formula. In the relativistic rise a weak mass dependence at fixed momenta due to the dependence of $\beta\gamma = p/(mc)$ complicates particle identification. For highly relativistic particles ($\gamma \gtrsim 1000$) [56] transition radiation (TR) can be exploited. TR was predicted by V. Ginzburg and I. Frank in 1946 [64]. Highly relativistic particles crossing a

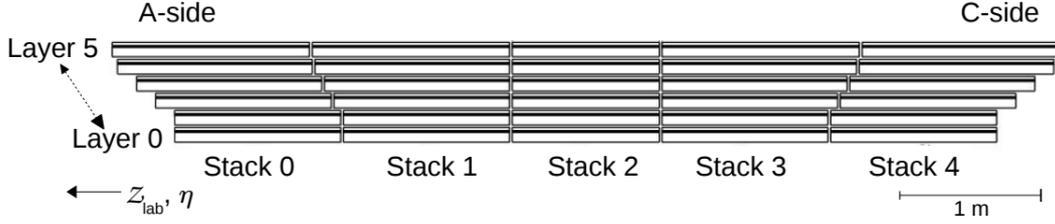


Figure 3.5: One supermodule consists of five stacks with six read-out chambers each. Figure taken from [3].

boundary surface between two media with different dielectric constants emit extremely forward peaked photons [65]. Since the TR only depends on $\gamma = E/(mc^2)$ pion/electron discrimination is supported at the desired momenta, i.e. at momenta larger than $1 \text{ GeV}/c$. On average, for each electron with a momentum above $1 \text{ GeV}/c$, at least one TR photon is produced. For pions the onset of TR is above $100 \text{ GeV}/c$ and so not appearing at relevant momenta at ALICE. Additionally, the TRD can contribute to trigger on rare probes such as electrons and jets as well as tracking. Stand-alone tracking capabilities are presented in this work.

The TRD is positioned in the central barrel behind the TPC at radii of 2.90 m to 3.68 m from the nominal interaction point. It consists of 18 supermodules with five stacks and each stack with six layers, which cover the full azimuth range and a pseudorapidity range of $-0.84 < \eta < 0.84$. To reduce the material budget in front of the PHOS detector, the middle stacks of three supermodules are left out (sector 13–15). This results in a total of 522 individual readout chambers. The TRD cross section with its 18 supermodules is shown in the context of the surrounding detector components in Fig. 3.4. The structure of one supermodule is shown in Fig. 3.5.

A schematic cross section of a TRD chamber as can be seen in Fig. 3.6 consists of the radiator, the drift region and the amplification region. The active area per chamber varies from $0.90 \text{ m} \times 1.06 \text{ m}$ to $1.13 \text{ m} \times 1.43 \text{ m}$ ($x \times z$). The thickness is about 12 cm including radiator, electronics and cooling.

Since the intensity of TR is of the order of the fine structure constant ($\alpha = 1/137$) per boundary crossing, many boundaries are necessary to produce an appropriate signal. In the ALICE TRD this is reached via a fiber/foam sandwich radiator with a total depth of 4.7 cm. The material choice is due to reasons of mechanical stability and space. For a fast enough particle ($\gamma \gtrsim 1000$) this results in a mean production of approximately 1.45 TR photons [67].

3 Experimental Setup

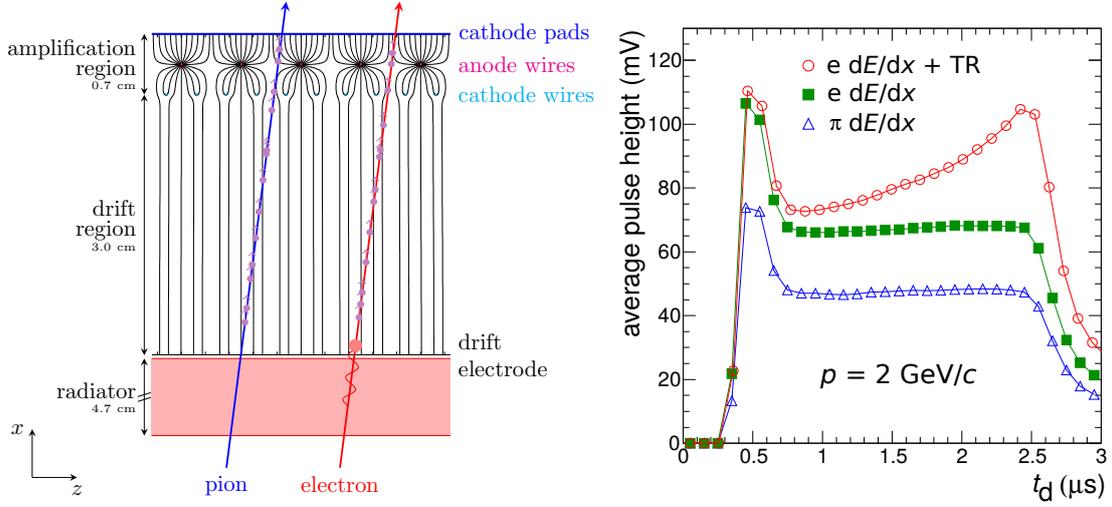


Figure 3.6: Left: A TRD chamber consists of the radiator, the drift region and the amplification region [3]. Right: The TRD signal is shown as a function of time. Here, the average signals for electrons and pions with momenta of $2 \text{ GeV}/c$ are shown. At small times a clear peak due to the amplification region can be seen. A second peak appears only for electrons due to TR contributions [66].

The following drift region of 3 cm is filled with a Xe-CO₂ gas mixture (85:15). The forward peaked TR photons in the X-ray regime are absorbed by xenon in less than a centimeter on average [3]. CO₂ is chosen as a good quenching gas. The TR photons, which are fully absorbed, and the crossing particles ionize the gas. The ionization can be measured in temporal dependence with a granularity of 100 ns. For that, the drift under the influence of a constant electric field is exploited. In the amplification region a strong electric field provokes avalanches and allows a clear pad readout induced in the MWPC. The length of the amplification region is 0.7 cm.

The temporal signal is shown in Fig. 3.6 as an average. A clear difference between the charge deposition of electrons and pions can be seen. The peak at small times is due to charges drifting from both sides in the amplification region. This peak is followed by charges from the drift region. A clear difference between the charge deposition without TR can be seen because of the different specific energy loss due to the different masses. By including the contribution of TR, a second peak in the electron signal at larger times appears. The charges arrive later since the TR photons are absorbed right after the radiator.

In Fig. 3.7 a sketch of the TRD cathode pad plane can be seen. Manufactured

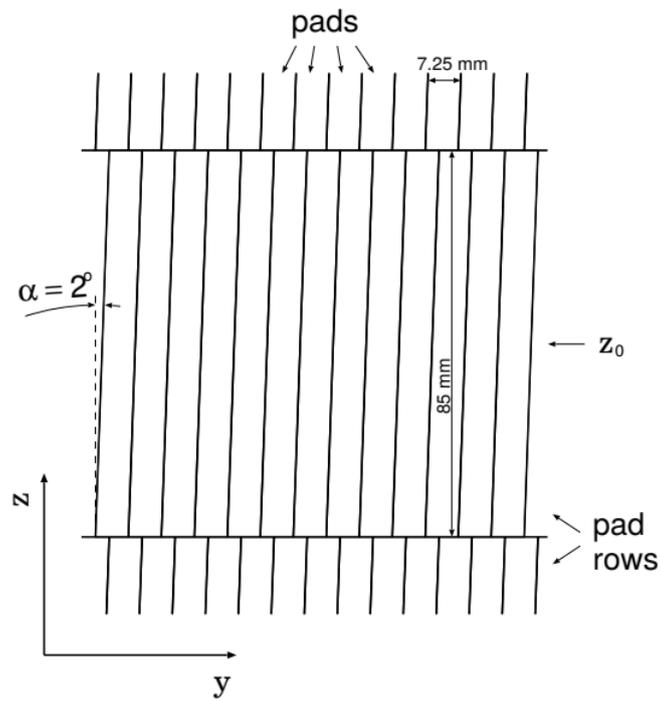


Figure 3.7: A sketch of the TRD pad plane. The pads are much larger in z direction than in y direction. The pads are tilted to still obtain a better z resolution. Figure taken from [3].

3 *Experimental Setup*

from thin printed circuit boards, they are stabilized by a light honeycomb and carbon fiber sandwich. The design goal of a deviation from planarity of $150\ \mu\text{m}$ is achieved by almost all chambers. The plane is segmented into 12 or 16 pads along z and 144 pads in $r\phi$ with an area of $0.635\ \text{cm} \times 7.5\ \text{cm}$ to $0.785\ \text{cm} \times 9\ \text{cm}$ ($y \times z$ in the local system). To diminish the effect of the large z pads to the z resolution the pads are tilted alternating from layer to layer by 2° in the yz plane. There is still a dead zone given in $r\phi$ from 12% to 14% and in z on average of 2.5%. Multi-Chip Modules (MCMs) are connected directly to 18 pads of the cathode plane. Analog to Digital Converters (ADCs) return a signal of 100 ns time bins leading to the input for the cluster finder and tracklet calibration procedure.

4 Charged Particle Tracking

To motivate the algorithms for the TRD stand-alone tracking, this chapter gives a general overview over the most important tracking basics. A more detailed insight can be found in [68]. The motion of a charged particle in a magnetic field is described. The interaction of the particles with matter influences the motion of the particle, and thus it is also an important topic of charged particle tracking. In general, the motion of the particle is described by a helix which has to be fitted to selected clusters produced in the detectors to obtain the track parameters. A brief introduction to helix fits is given. The standard method used in ALICE for parameter estimation is the Kalman filter [69]. Track seeds that are necessary for this procedure are given using the cellular automaton [70]. The Kalman filter and the corresponding seeding procedure is also explained.

4.1 The Track Model

The motion of a particle with an electric charge in a magnetic field is governed by the Lorentz force [71], given by

$$\vec{F} = q(\vec{E} + \vec{v} \times \vec{B}). \quad (4.1)$$

Here, q is the signed charge number, \vec{B} the magnetic field and \vec{v} the velocity of the particle. This equation is also valid for a relativistic motion with $p = \beta\gamma m$. Moreover, the absence of material is assumed. In case there is no electric field, this leads to the differential equation

$$\frac{d\vec{p}}{dt} = q\vec{v} \times \vec{B}. \quad (4.2)$$

From this equation it can be obtained that the absolute value of the momentum is not time dependent. This is shown by the calculation

$$\frac{d\vec{p}^2}{dt} = 2\vec{p} \frac{d\vec{p}}{dt} = \frac{2}{m\gamma} q\vec{p}(\vec{p} \times \vec{B}) = 0 \quad (4.3)$$

where the relativistic momentum $\vec{p} = m\gamma\vec{v}$ was inserted, with m the rest mass of the particle and γ the Lorentz factor. A typical choice for the direction of a magnetic field is $\vec{B} = B\vec{e}_z$, such that the field points along the beam axis. The magnetic field

4 Charged Particle Tracking

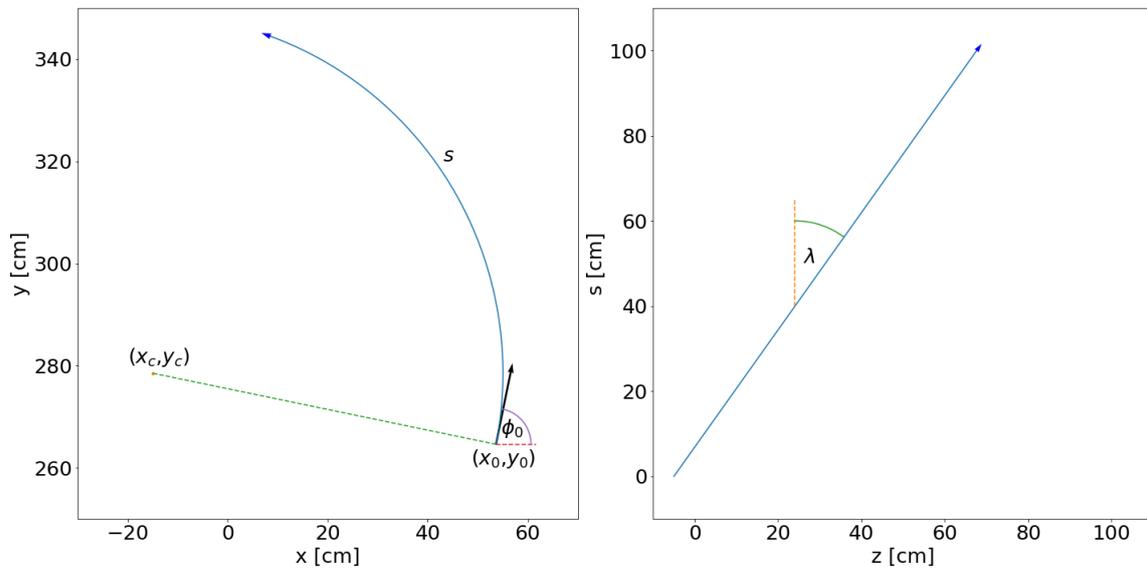


Figure 4.1: Left: An example of a trajectory of a charged particle in the xy plane perpendicular to the direction of the magnetic field. Right: The definition of λ is given by the linear dependence of the motion in the zs plane. The parameter s is a running parameter along the arc of the two-dimensional circle.

is created as homogenous as possible. B is the absolute value of the magnetic field. By inserting this relation into the differential equation Eq. 4.2 and using the fact that the momentum is conserved, it follows that the momentum p_z in z direction and the transversal momentum p_T do not change in time. A helical ansatz solves the differential equation. The motion of the charged particle using $p_T = p \cos \lambda$ and $p_z = p \sin \lambda$ can be described by

$$p_x = p_T \cos(Cs + \phi_0) \quad (4.4a)$$

$$p_y = p_T \sin(Cs + \phi_0) \quad (4.4b)$$

$$p_z = p_T \tan \lambda, \quad (4.4c)$$

where s is the relative arc length of the two-dimensional circular part of the helix. It is a natural running parameter along the helix. The curvature C is defined as the inverse of the radius of this circle $|C| = 1/R$. The parametrization

$$x(s) = x_c + \sin(Cs + \phi_0)/C \quad (4.5a)$$

$$y(s) = y_c - \cos(Cs + \phi_0)/C \quad (4.5b)$$

$$z(s) = z_0 + s \tan \lambda \quad (4.5c)$$

is in agreement with the upper one. Here, x_c and y_c are the coordinates of the center of the circle and z_0 is the corresponding z coordinate at $s = 0$. The centers might be replaced by

$$x_c = x_0 - \sin(\phi_0)/C \quad (4.6a)$$

$$y_c = y_0 + \cos(\phi_0)/C. \quad (4.6b)$$

A visualization of the given parameters is shown in Fig. 4.1.

4.2 Interaction with Material

The track model as given in the previous section describes the motion of a charged particle in a homogeneous magnetic field in the absence of material. In real detectors, the motion of the particles is influenced by material. One source of a change of the movement caused by material is the energy loss due to ionization or excitation. The mean energy loss of charged heavy particles crossing material is described by the Bethe-Bloch formula [72]

$$\left\langle -\frac{dE}{dx} \right\rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 W_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \quad (4.7)$$

4 Charged Particle Tracking

with a precision of a few percent for $0.1 \leq \beta\gamma \leq 1000$. As it is the mass stopping power, the units are $[\text{MeV g}^{-1}\text{cm}^2]$ of the mean energy loss. To obtain the energy loss per unit length, the formula has to be multiplied with the mass density ρ . K is a coefficient of proportionality with $K = 0.307075 \text{ MeV mol}^{-1} \text{ cm}^2$, z the charge number of the incident particle, Z the atomic number of the absorber, A the atomic mass of the absorber, W_{max} the maximum possible energy transfer to an electron in a single collision, I the mean exciting energy and $\delta(\beta\gamma)$ the density correction. In general, a small mass dependence of W_{max} can be neglected and the function only depends on β . This β dependence enables the particle identification capabilities if measuring the momentum besides the energy loss.

Besides very small effects of Møller scattering, Bhabha scattering and positron annihilation, the ionization energy loss described before dominates for low energy electrons/positrons. For most materials, bremsstrahlung starts to dominate at a few tens of MeV. For heavier particles this is not relevant at typical ALICE momenta. For copper, the fraction of ionization and bremsstrahlung of electrons to the total energy loss is shown in Fig. 4.2. Close to a nucleus an electron/positron can radiate a photon ($e + N \rightarrow \gamma + e + N$). Compared to the energy loss by ionization this is a very fluctuating process. The probability density function can be described by the Bethe-Heitler formula [73]

$$f(z) = \frac{(-\ln(z))^{c-1}}{\Gamma(c)} \quad (4.8)$$

with $c = t/\ln(2)$, $z = E_f/E_i$, $t = X/X_0$. E_f is the energy of the electron before the radiation and E_i after the radiation. The length within the material crossed by the particle is X , and X_0 is defined as the characteristic radiation length in a certain material. The radiation length is the mean distance where $1/e$ of the energy of the electron is lost by bremsstrahlung in units of $[\text{g cm}^{-2}]$. The linear behavior can be seen in the formula, as there is no energy dependence for the relative energy loss in the probability density function $f(z)$.

The third important interaction leading to distortions, which needs to be added to the ideal track model from before, is multiple scattering. These deflections mainly come from Coulomb scattering [72]. It consists of many small angle deflections, which can be described by a Gaussian approximation due to the central limit theorem. This approximation is valid for the central 98% of the angular distribution and from Lynch & Dahl [74] the rms width of the angle is given by

$$\theta = \frac{13.6 \text{ MeV}}{\beta c p} z \sqrt{\frac{X}{X_0}} \left[1 + 0.038 \ln \left(\frac{X z^2}{X_0 \beta^2} \right) \right], \quad (4.9)$$

where z is the charge number of the incident particle and X/X_0 is the length passed by the particle in terms of the radiation length as defined before. Nevertheless, some

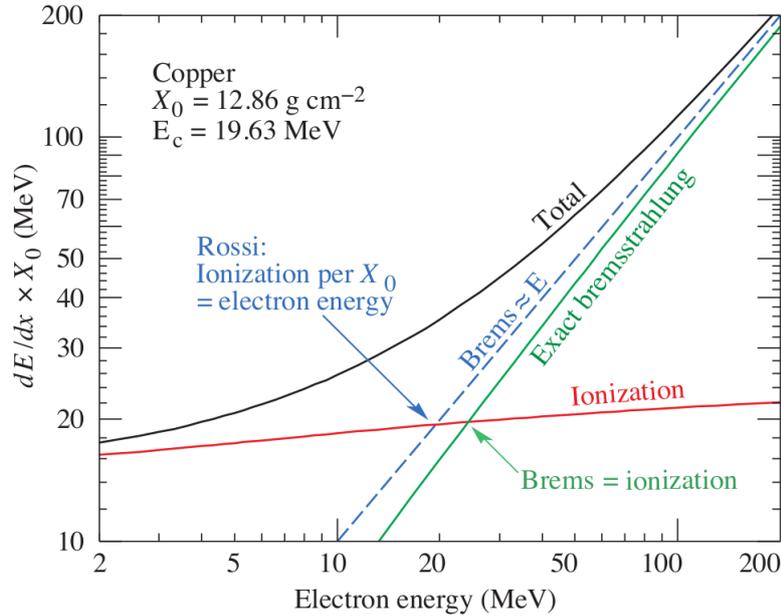


Figure 4.2: The energy loss of electrons crossing copper. Bremsstrahlung is already dominant compared to ionization at energies much lower than it is possible to measure with the TRD. Figure taken from [72].

"hard" scatters produce non-Gaussian tails.

4.3 Kalman Filter

The Kalman filter, named after Rudolf Emil Kálmán, who published one of the first descriptions of the algorithm in 1960 [75], is the standard method for particle tracking in ALICE. Details about the implementation in ALICE can be found in [69]. The given algorithm is a recursive parameter estimator/fit. Besides the parameters also the uncertainty of the parameters are an output of the algorithm. An initial estimate (seed) is required.

Having an optimal initial estimate and measurements with Gaussian uncertainties, the Kalman filter is an optimal estimator for linear models. In this case, the result is equivalent to a least squares fit. The given algorithm with a linearized model by using a local Taylor series expansion is called the extended Kalman filter.

One of the first and very famous applications was the use within the Apollo navigation computer by NASA Ames Research Center [76]. NASA Ames was faced with a non-linear problem and contributed to the non-linear formulation of the extended

4 Charged Particle Tracking

Kalman filter.

For the application of charged particle tracking in a magnetic field the Kalman filter provides a generalization for a linearized helix fit. The requirement for an optimal tracking in ALICE are besides the track parameter estimation also a very good selection of the space points which belong to a track. Being recursive, the Kalman filter provides a combination of both. It also provides a natural implementation of external noise for instance from material effects.

To start with the mathematical description [69] of the linear Kalman filter, a functional model description is needed. This is given by a state vector \vec{x}_k which changes by the time-like parameter t_k . The function representing the variation from state t_{k-1} to t_k given a certain noise $\vec{\epsilon}_k$ is then

$$\vec{x}_k = F_k \vec{x}_{k-1} + \vec{\epsilon}_k. \quad (4.10)$$

The linear operator F_k represents the linear transport. This description provides a recursive propagation of the fit function including the uncertainty from external disturbance. The full uncertainties of the parameters are described by its covariance matrix C_k . The measurements \vec{m}_k are a projection H_k of the state x_k . The transformation from the model parameter to the observation m_k is then defined by

$$\vec{m}_k = H_k \vec{x}_k + \vec{\delta}_k, \quad (4.11)$$

with δ_k being the uncertainty of the measurement. The observations also have a covariance matrix V_k . In the generalization of the extended Kalman filter, in Eq. 4.10 the propagation operation $F_k \vec{x}_{k-1}$ is substituted by a functional representation $f_k(\vec{x}_{k-1})$.

The process of the Kalman filtering now starts with a measurement or the seed, which is extrapolated to the current state. The goal is to update the parameters from t_{k-1} to t_k . In the following $\vec{x}_{k-1|k-1}$ is the unpropagated state, $\vec{x}_{k|k-1}$ the propagated but not updated state and $\vec{x}_{k|k}$ the final updated state. For the covariance matrix the same convention is used. The extrapolation then applies as follows:

$$\vec{x}_{k|k-1} = F_k \vec{x}_{k-1|k-1} \quad (4.12a)$$

$$C_{k|k-1} = F_k C_{k-1|k-1} F_k^T + Q_k. \quad (4.12b)$$

Q_k is the contribution of external noise to the covariance matrix. Given the measurement \vec{m}_k at the current state, the fit parameters $\vec{x}_{k|k-1}$ and the covariance matrix $C_{k|k-1}$ can now be updated. To do that, the Kalman gain K_k is introduced and

Extrapolation:	Update:
<ul style="list-style-type: none"> • extrapolate track parameters: $\vec{x}_{k k-1} = F_k \vec{x}_{k-1 k-1}$ • extrapolate covariance matrix $C_{k k-1} = F_k C_{k-1 k-1} F_k^T + Q_k$ 	<ul style="list-style-type: none"> • update track parameters: $\vec{x}_{k k} = \vec{x}_{k k-1} + K_k (\vec{m}_k - H_k \vec{x}_{k k-1})$ • update covariance matrix $C_{k k} = C_{k k-1} + K_k H_k C_{k k-1}$ • Kalman gain: $K_k = C_{k k-1} H_k^T (V_k + H_k C_{k k-1} H_k^T)^{-1}$

Figure 4.3: A summary of the most important formulas of the Kalman filter [75]. The algorithm consists of an extrapolation step and the parameter updates. The meaning of the variables is described in the text.

applies as follows:

$$\vec{x}_{k|k} = \vec{x}_{k|k-1} + K_k (\vec{m}_k - H_k \vec{x}_{k|k-1}) \quad (4.13a)$$

$$C_{k|k} = C_{k|k-1} + K_k H_k C_{k|k-1} \quad (4.13b)$$

Here, the Kalman gain is given by

$$K_k = C_{k|k-1} H_k^T (V_k + H_k C_{k|k-1} H_k^T)^{-1}. \quad (4.14)$$

Via

$$\chi_k^2 = (\vec{r}_k)^T (R_k)^{-1} \vec{r}_k \quad (4.15)$$

with the residual $\vec{r}_k = \vec{m}_k - H_k \vec{x}_{k|k}$ and $R_k = V_k - H_k C_{k|k} H_k^T$ finally the χ^2 increment can be calculated to check the quality of the fit. This is equivalent to the χ^2 of the least squared fit in the linear case. The most important formulas are summarized in Fig. 4.3.

In concrete applications within ALICE, a local coordinate system is typically chosen, where the running time-like parameter is in radial direction perpendicular to the detector layers. Multiple scattering and energy loss described by the Bethe-Bloch formula can be implemented by adding a corresponding noise term Q_k to the covariance matrix of the measurements. Details can be found in [77]. Noise from bremsstrahlung in the case of electrons is in general not Gaussian and thus the Kalman filter is not accurate. A possibility to get more accurate results is to apply a modification of the Kalman filter called the Gaussian Sum Filter [78].

4.4 Helix Fits

The three-dimensional helix parameter estimation is in general a non-linear problem. In the case of the Kalman filter these circumstances are simplified by using a linearization of the propagation of the covariance matrix.

A common ansatz to solve the parameter estimation is to separate the helix into a circle fit and a straight line fit with the arc length s of the circle as parameter in direction of the magnetic field. For the circle parameter the following expression has to be minimized:

$$\chi^2 = \sum_i \left(\sqrt{(x_i - x_c)^2 + (y_i - y_c)^2} - R \right)^2. \quad (4.16)$$

This minimization procedure results in estimates for the center x_c and y_c of the circle and the radius R . It should be remarked that with this parameterization there is no continuous transition between positive and negative curvatures/charges. The radius and the center are highly correlated. The given minimization is not analytical solvable.

Instead, a common approach is to solve the minimization of the following $\tilde{\chi}^2$:

$$\tilde{\chi}^2 = \sum_i [(x_i - x_c)^2 + (y_i - y_c)^2 - R^2]^2. \quad (4.17)$$

This approximation provides a system of linear equations which can be solved analytically and the approximation holds as long as the residuals are much smaller than the radius of the circle. Karimäki proposed a parameterization based on this assumption which provides linear errors [79]. This parameterization is also smooth in the transition from positive to negative curvatures and vice versa. When applying this fit for studies for this work it was noted that the structure of this model with the double square nature leads to issues with outliers, that are heavily weighted.

4.5 Track Finding and Track Seeding

It is distinguished between track finding and track parameter estimation. The aim of track finding is to associate the detector clusters to tracks. A special property of the Kalman filter is that after having a proper seed it can also contribute to track finding. In the case of the Kalman filter the track finding is thus also necessary at least in a part of the detector to select seed clusters. From the seed clusters, proper start parameters for the state and for the covariance matrix need to be estimated.

A typical approach for track finding or track seeding with respect to the Kalman filter is a cellular automaton based algorithm [70][68]. The cellular automaton is also

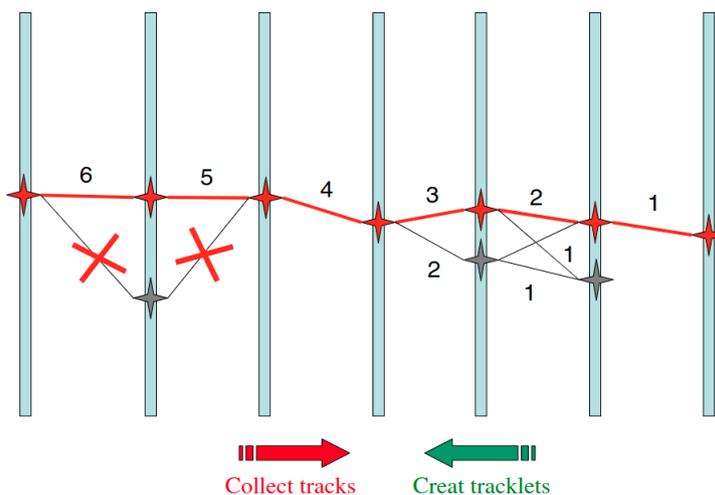


Figure 4.4: The working principle of the cellular automaton is depicted. The cellular automaton creates the tracklets and assigns state numbers, which are then used to collect the tracks. Figure taken from [70].

used in ALICE [80]. A cellular automaton consists of a graph with nodes and edges. An introduction to graphs is given later in Sec. 10.1.1 in the context of graph neural networks. From two clusters in neighboring detector layers, tracklets are built. These tracklets are short track segments described by a position and direction vector or two hits, respectively, and called cells. These are represented by the nodes of the graph. To construct the graph, a neighbor discovery is necessary. Sometimes it is also allowed to skip detector layers. In common cases, the graph is a directed graph from the inner layers to the outer layers. The cells are called neighbors if they share a hit. Different cuts depending on the detector geometry are applied for accepting the tracklets, for example cuts on the angles.

The actual track search starts by assigning a counter to each cell, initialized with 1. Then, every counter is updated by 1 if the counter of the inner neighbor has the same value. This is called a time step. This procedure is repeated in every subsequent time step until no state changes. This procedure allows parallel execution. On the basis of the counter values, starting from the highest values, the tracks can be constructed by appending neighbors with a counter value decremented by 1. Depending on the track model, finally a χ^2 -criterion can be applied for the final track selection after fitting the track candidates. The working principle is depicted in Fig. 4.4. This algorithm profits from a high parallelizability and the local search suppresses excessive combinatorics.

5 Reconstruction of Photon Conversions

A common tool to detect photons is the reconstruction of the electron–positron pairs from photon conversions, which is also the only way by utilizing the tracking detectors of ALICE. In this thesis, an important topic is the reconstruction of photon conversions in the ALICE TRD. This can complement the reconstruction of photon conversions in the ALICE TPC. This chapter gives an introduction to this topic.

To reconstruct photon conversions, the topological properties of the daughter pairs are exploited. Thus, it is important to introduce the physics of photon conversion, which is explained in the first section. To exploit the physics on the level of tracking, in ALICE the so-called Kalman Filter Particle package (KFParticle) is used. This enables constraints concerning for instance the invariant mass or a vertex constraint. Section 2 is meant to give an insight to this method. Finally in the last section of this chapter, a short overview of the general procedure of the photon conversion method using the ALICE TPC is given.

5.1 The Physics of Photon Conversions

To reconstruct photon conversions, physical and kinematic constraints of the daughter tracks are exploited. In the following, the physical and kinematic properties are described.

In Fig. 5.1 the total photon cross section is shown in carbon as an example for a light material and lead as an example for a heavy material. The three dominant processes over different energy scales are the photo effect at low energies below 100 keV, the Compton effect at the order of 1 MeV, and at higher energies the cross section for pair creation starts to rise quickly. Pair creation is the process in which a photon converts to an electron–positron pair. Pair creation was first observed by Carl David Anderson in 1933 in the context of the discovery of the positron [81]. Due to energy momentum conservation, this process is only possible in the vicinity of a nucleus transferring a bit of its energy. The energy E_γ of the photon must exceed the threshold for the creation given by [82]

$$E_\gamma \geq 2m_e c^2 + 2 \frac{m_e^2 c^2}{m_{\text{nucleus}}}. \quad (5.1)$$

5 Reconstruction of Photon Conversions

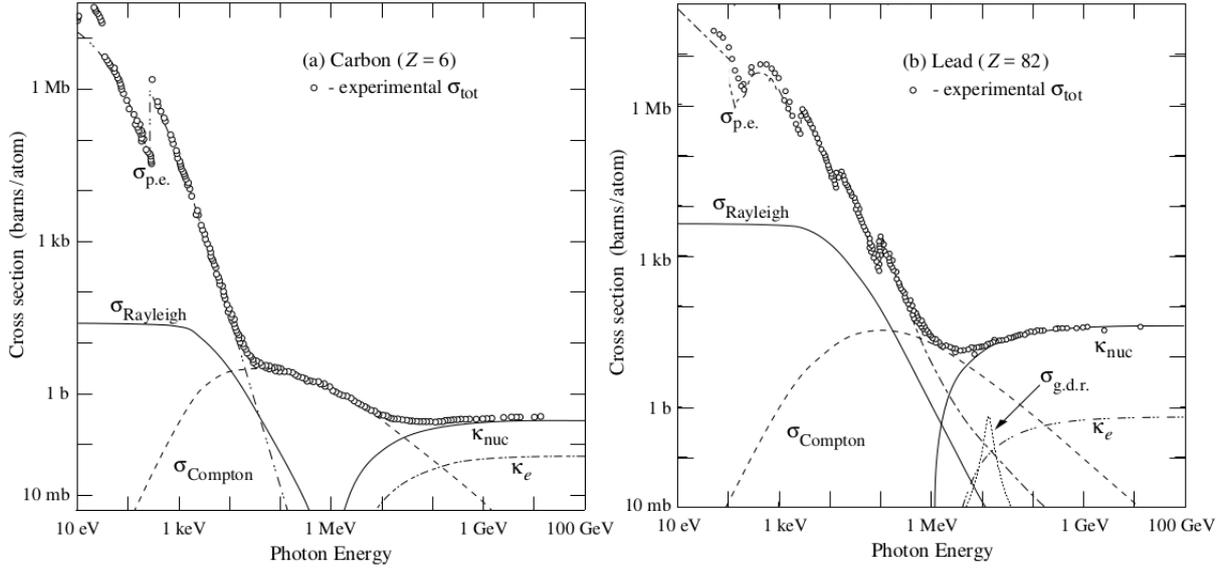


Figure 5.1: Total photon cross section in carbon (left) and lead (right) as a function of the photon energy. The different contributions are indicated by: $\sigma_{\text{p.e.}}$: atomic photoelectric effect, σ_{Rayleigh} : Rayleigh scattering – atom neither ionized nor excited, σ_{Compton} : Compton scattering, κ_{nuc} : pair production in nuclear field, κ_e : pair production in electron field, $\sigma_{\text{g.d.r.}}$: photonuclear interactions (giant dipole resonance). Figure taken from [72] and modified.

As the mass of the electron is much smaller than the mass of the nucleus, the second term can be neglected. At the energies measured in ALICE starting above 10 MeV this is already fulfilled and the pair creation is the dominant effect. Pair creation can also happen in the vicinity of an electron field, but is strongly suppressed, as can be seen in Fig. 5.1.

The kinematics of the pair creation is described by the Feynman diagrams shown in Fig. 5.2. Let ϵ be defined as the ratio of the photon energy to the electron mass:

$$\epsilon = \frac{E_\gamma}{m_e c^2}. \quad (5.2)$$

Then, in the limit $\epsilon \gg 1/(\alpha Z^{1/3})$, where complete screening by atomic electrons of the nuclear charge appears, the cross section for pair production σ_{pair} is [82][83]

$$\sigma_{\text{pair}} = 4\alpha r_e^2 Z^2 \left(\frac{7}{9} \ln \frac{183}{Z^{1/3}} - \frac{1}{54} \right), \quad (5.3)$$

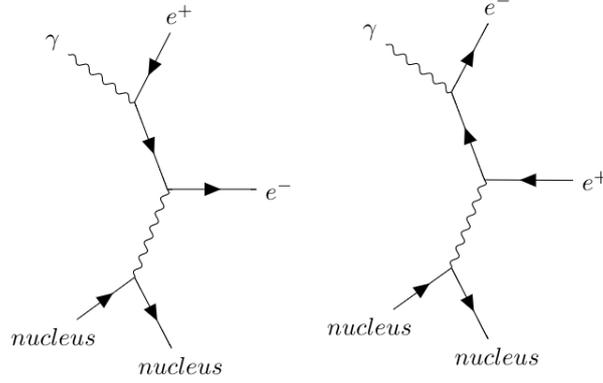


Figure 5.2: The two Feynman diagrams for pair creation.

where α is the fine-structure constant, Z is the atomic number and r_e is the classical electron radius. For the asymptotic case the term $\frac{1}{54}$ can be neglected, and it follows

$$\sigma_{\text{pair}} \approx \frac{7}{9} 4\alpha r_e^2 Z^2 \ln \frac{183}{Z^{1/3}} \approx \frac{7}{9} \cdot \frac{A}{N_A} \cdot \frac{1}{X_0} \quad (5.4)$$

with X_0 being the radiation length introduced in the previous chapter and N_A is the Avogadro number. This equation is accurate within a few percent down to 1 GeV [72]. The photon conversion probability is thus given by [83]

$$P = 1 - e^{-\frac{7}{9} \frac{X}{X_0}}. \quad (5.5)$$

For the understanding of the kinematics of a photon conversion two observables are of interest. The first one is the angular distribution of the electron or positron with respect to the photon. An overview is given in [73]. The mean angle is of the order of $\theta \approx m_e c^2 / E_0$. E_0 is the energy of the created electron/positron. This angle vanishes at large energies. The explicit probability distribution is approximately given by

$$\Phi(\theta_0) d\theta_0 = \frac{\theta_0 d\theta_0}{\Theta^2 + \theta_0^2}, \quad \Theta = \frac{m_e c^2}{E_0}. \quad (5.6)$$

This formula is not valid for energies of the order of mc^2 . An approximation for the angle between the electron and positron is given in [84]:

$$\Theta_{+-} \approx 1.6 \text{ MeV} / E_\gamma. \quad (5.7)$$

In addition, there is the asymmetry given by the energy fraction $x = E_1/(E_1 + E_2)$, where E_1 is the energy of the electron/positron with the lower energy and E_2 the one with the higher energy. A calculation of the cross section of the positron or electron $\sigma(E_+) = \sigma(E_-)$ can be found in [73] and [85]. From this cross section the ratio can be obtained. At photon energies of 100 MeV the distribution of the energy fraction is nearly flat and for higher energies it starts to peak at the edges.

5.2 The Kalman Filter Particle Package

In the photon conversion method, the fact is exploited that the photons convert to an electron–positron pair with a negligible energy loss during the transition. These charged particles can make signals in the tracking detectors and then be reconstructed. In the simplest approach, the closest distance between the tracks is considered as an estimation of the conversion vertex and the momentum of the photon can be estimated using the daughter momenta.

Due to further topological and physical constraints, more information can be used for the reconstruction of the photon kinematics as shown in the previous section. In case of the photon conversion of a primary photon, these constraints are the common conversion vertex, an invariant mass constraint of zero and the primary vertex as production vertex of the photon.

A commonly used method in ALICE that was developed for short-lived particles but also for photon conversion is the Kalman Filter Particle package or short KF-Particle package. A precise insight into the state-of-the-art application is given in [86] and [87]. Beyond the value for photon conversion reconstruction, decay chains can also be reconstructed.

The principle behind the KFParticle package is to use the Kalman filter introduced in Chapter 4 for mother particle reconstruction. Instead of reconstructing tracks from position coordinates, mother tracks are updated by recursively adding daughter tracks treated as measurements. Also, physical and topological constraints are handled as measurements. Again, the starting point is a state vector and its covariance matrix. The canonical choice is given by

$$\mathbf{r} = \{x, y, z, p_x, p_y, p_z, E, s\} \quad (5.8)$$

with the position and momentum, as well as the energy E and $s = l/p$ the length of the particle trajectory l normalized to its momentum. After an initial guess of the state vector, Kalman updates using the state vectors of daughters tracks and the corresponding covariance matrices deliver a common estimate of the state vector including the fit quality. Thereby, a common vertex is constrained. Mother and

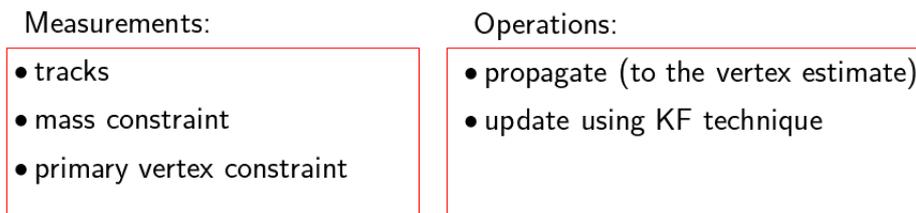


Figure 5.3: The KFParticle package consists of the Kalman filter operating on track level. Other constraints can be added as additional pseudo-measurements. Updates are performed at a vertex estimate, which is then a vertex constraint. Input and state are of the same parametrization.

daughter tracks are handled with the same set of parameters. After the reconstruction of the mother particle, a transport along the trajectory parameter including the covariance matrix can be performed. This allows a precise reconstruction of decay chains. Moreover, particles can also be removed afterwards. Due to the architecture of the implementation, the KFParticle package is a high-performance solution for vertex and mother particle kinematics reconstruction. The procedure is summarized in Fig. 5.3.

An important property of the KFParticle package is the handling of constraints. Constraints are fitted as additional measurements. The "Lagrange multipliers" technique is a standard tool to implement hard constraints to fits. This is not feasible in the Kalman filter approach. Instead, soft constraints are added as measurements, i.e. regularization terms. Thereby, it is considered that the hard constraint is a particular case of the soft constraint. Besides a primary vertex constraint also a non-linear mass constraint can be applied [87]. Especially in the case of mass constraints, a linearization of energy momentum conservation leads to tails of the mass distribution. Thus, the solution used in the KFParticle package for this particular case includes an iterative ansatz of the Lagrangian multiplier problem.

5.3 Reconstruction of Photon Conversions using the ALICE TPC

In this section, it is shown how a typical photon signal extraction proceeds in ALICE for photon conversions that take place mainly in the ITS and also in the TPC. As

5 Reconstruction of Photon Conversions

an example, a Pb–Pb analysis at $\sqrt{s_{\text{NN}}} = 2.76$ TeV is shown following reference [88]. For this application, the excellent tracking and PID capabilities of the ALICE TPC are exploited.

The photon extraction procedure can be divided into four steps:

- candidate selection
- track and V0 cuts
- PID cuts
- photon cuts

The most important cuts are listed and explained in this section.

First, candidates are searched via the so-called V0 finder. The online V0 finder, which runs during the reconstruction, searches for two unlike-sign tracks with a vertex that is at least 1 mm away from the primary vertex and a minimal distance between the tracks of less than 1.5 cm. A cut on the pointing angle is then performed. This is the angle between the vector pointing from the primary vertex to the vertex estimate and the common momentum of the track pair. The cosine of the pointing angle must be greater than 0.9.

The output of the V0 finder are not only photon candidates but also K_s^0 , Λ and $\bar{\Lambda}$. The tracks are refitted using the KFParticle method introduced in the previous section. A common χ^2 is thus given as selection quality criterion. Moreover, a recalculation of the vertex takes place where the electron and positron are considered to have parallel momenta at their origin, as the photon is massless.

Besides other track quality criteria, it is required that the photon conversion takes place within a conversion radius of $5 \text{ cm} < R_{\text{conv}} < 180 \text{ cm}$ and $|z_{\text{conv}}| < 240 \text{ cm}$ with a transverse momentum of $p_{\text{T}} > 0.05 \text{ GeV}/c$ of the daughter tracks and a pseudorapidity $|\eta| < 0.9$ for the daughter tracks as well as the mother track. In addition, a cone selection is applied that removes tracks that lie outside the fiducial region. These cuts allow a signal within the TPC.

For the TPC PID an electron signal of the form of $-3 < N_{\sigma_e} < 5$ is required. N_{σ_e} is the deviation to the mean dE/dx in terms of standard deviations for electrons. Additionally, pions are rejected for momenta of $0.4 \text{ GeV}/c < p_e < 2 \text{ GeV}/c$ via $N_{\sigma_\pi} > 3$ and for $p_e > 2 \text{ GeV}/c$ via $N_{\sigma_\pi} > 1$. If available, additional TOF PID is applied via $-5 < N_{\sigma_e} < 5$. The TPC signal is shown in Fig. 5.4, before and after all selection cuts are applied.

5.3 Reconstruction of Photon Conversions using the ALICE TPC

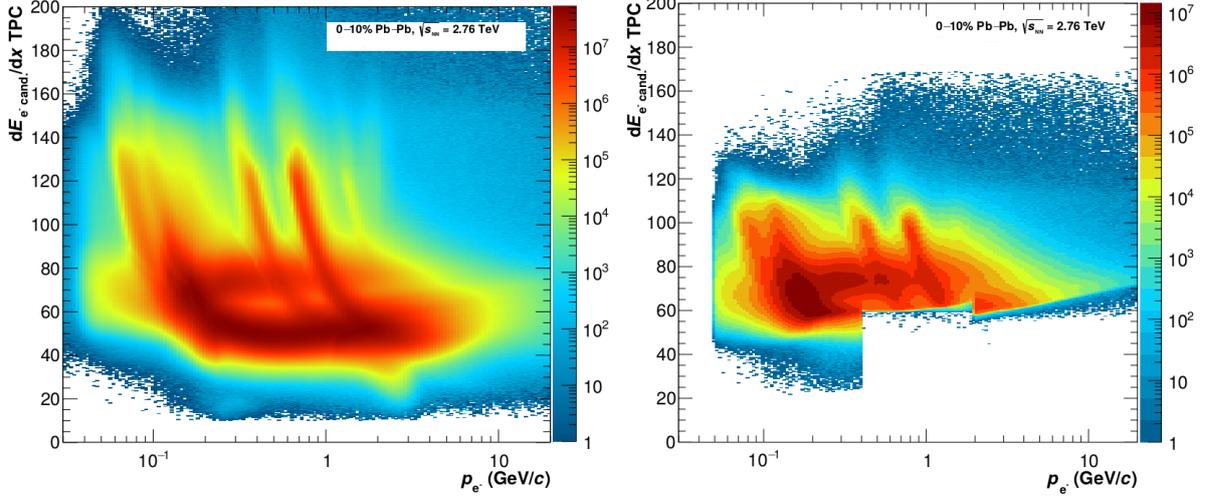


Figure 5.4: The TPC signal left before and right after all selection cuts are applied for real Pb–Pb data at $\sqrt{s_{NN}} = 2.76$ TeV. Especially the PID rejection cuts are clearly visible. Figure taken from [88].

To benefit from the topological constraints given by the kinematics of a photon conversion different variables are introduced. The longitudinal momentum asymmetry α is given by

$$\alpha = \frac{p_1^{e^+} - p_1^{e^-}}{p_1^{e^+} + p_1^{e^-}}, \quad (5.9)$$

where p_1 is the component of the electron/positron momentum pointing vector in the direction of the photon. The transverse component is then given by $q_T = p_e \cdot \sin(\theta_{V^0,e})$. The plotted distribution given by q_T versus α is called Armenteros-Podolanski plot [89]. The distribution is commonly used for the separation of different V0 candidates. For the photon conversions, α is symmetric, as the daughter particles have the same mass and q_T is nearly zero. In [88], a two-dimensional elliptic cut is applied on this distribution via

$$\left(\frac{\alpha_{V^0}}{\alpha_{V^0,\max}} \right)^2 + \left(\frac{q_T}{q_{T,\max}} \right)^2 < 1, \quad (5.10)$$

where $\alpha_{V^0,\max} = 0.95$ and $q_{T,\max} = 0.05$ GeV/ c . The Armenteros-Podolanski plot as given in [88] before and after all selection cuts are applied is shown in Fig. 5.5.

An appropriate fit quality is guaranteed via a cut on the photon χ_γ^2/ndf of 20 given by the KFParticle package.

5 Reconstruction of Photon Conversions

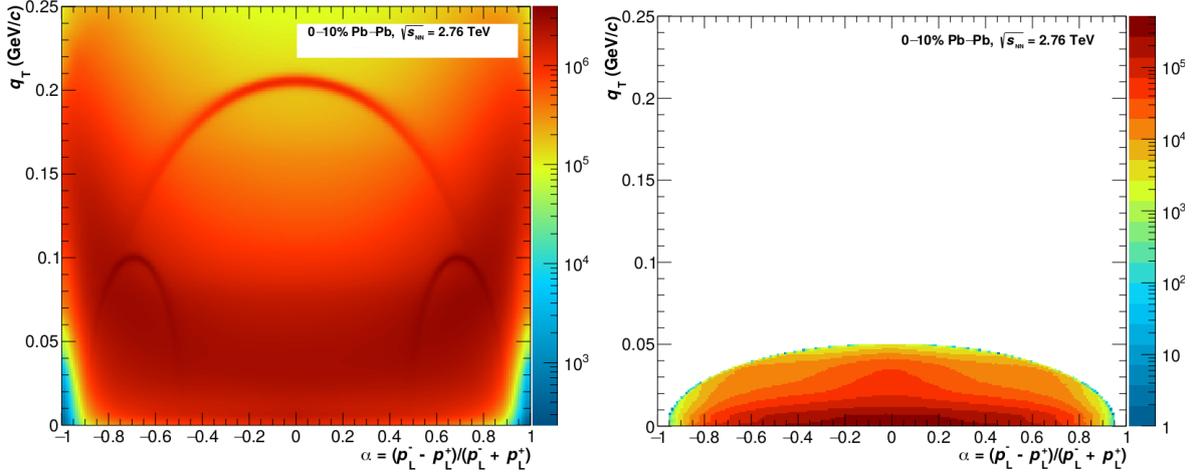


Figure 5.5: The Armenteros-Podolanski plot before and after applying all cuts, i.e. an elliptic cut on the given distribution. The variables are described in the text. Figure taken from [88].

The opening angle ξ_{pair} is defined as the angle between the electron and positron momentum vectors at the conversion point:

$$\xi_{\text{pair}} = \arccos \left(\frac{\vec{p}_{e^-} \cdot \vec{p}_{e^+}}{|\vec{p}_{e^-}| \cdot |\vec{p}_{e^+}|} \right). \quad (5.11)$$

The angle between the xy plane where the magnetic field affects the bending and the plane in which the opening happens is defined by

$$\psi_{\text{pair}} = \arcsin \left(\frac{\Delta\theta}{\xi_{\text{pair}}} \right), \quad (5.12)$$

where $\Delta\theta = \theta_{e^-} - \theta_{e^+}$ is defined as the difference between the polar angles of the electron and positron. The definition of ψ_{pair} is illustrated in Fig. 5.6. This angle is cut at $\psi_{\text{pair,cut}} = 0.1$ rad.

In addition, this cut is put in relation with the fit quality χ^2 using a two-dimensional cut

$$|\psi_{\text{pair}}| < \frac{-\psi_{\text{pair,cut}}}{\chi_{\gamma,\text{cut}}^2 \cdot \chi_{\gamma}^2} + \psi_{\text{pair,cut}} \quad (5.13)$$

with the cut values given before.

5.3 Reconstruction of Photon Conversions using the ALICE TPC

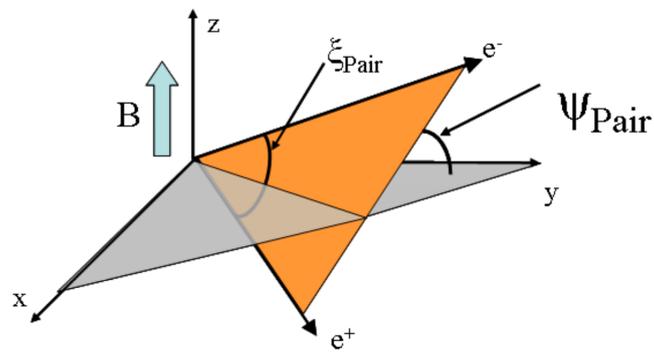


Figure 5.6: The angle ψ_{pair} is the angle between the plane defined by the momentum vectors of the electron and of the positron and the xy plane. Figure taken from [90].

6 Dataflow and Differentiable Programming

In the following chapters, frameworks and methods for machine learning are used in the context of track and particle reconstruction and the estimation of corresponding parameters. In this chapter, the main concepts of some of these tools are presented. More concretely, the tracking and the photon conversion fit were implemented using modern machine learning tools.

In recent years, there have been big steps forward in the area of machine learning, especially in the subcategory of deep learning [6]. These advances were mainly driven by the capability of deep neural networks to automatically discover representations from raw data in the areas of image, video, speech, text and audio processing. These raw data are handled as a data flow via array-based programming, as the execution of neural networks is based on multidimensional array operations (called tensor operations within TensorFlow [8]). For instance, matrix multiplications and corresponding linear algebra operations provide a high parallelizability and thus a very efficient computation on multiple Central Processing Units (CPUs) and Graphics Processing Units (GPUs). Recently, there is also a growing market for special devices for this purpose such as Tensor Processing Units (TPUs) [91].

For the optimization of deep neural networks a stable minimization of a loss function with respect to a large number of free parameters is necessary. To accelerate the research, many machine learning frameworks were developed for the implementation and deployment for large scale machine learning models. The most popular ones are TensorFlow [8] and PyTorch [9]. These frameworks provide the possibility for flexible design and experimentation with new models and research while delivering a high parallelization with low effort. The models are optimized for the corresponding devices (system resources), which are automatically allocated. For the stability of high dimensional parameter minimization of both frameworks the automatic differentiation feature is crucial. Automatic differentiation exploits the chain rule and the fact that computational operations are reduced to combined basic operations to give derivatives with working precision, i.e. not a numerical approximation is used. In this work, the implementation of the TRD stand-alone tracking is based on TensorFlow and therefore the following description refers mainly to functionalities given by TensorFlow. An implementation using PyTorch would also be possible. The

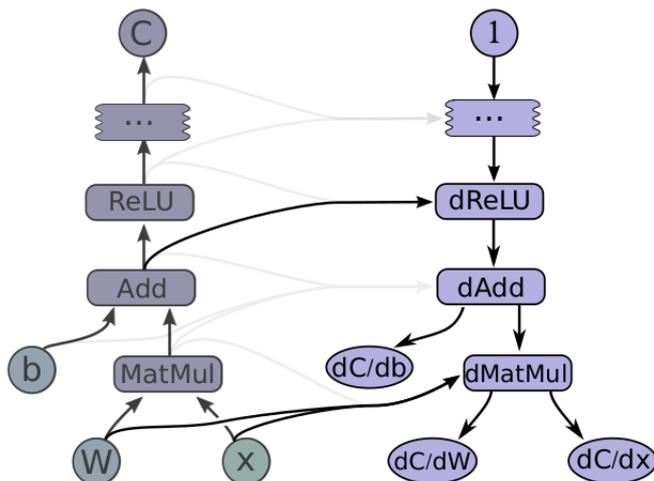


Figure 6.1: An example for a computational graph with the corresponding reverse graph for gradient calculation. The gray lines indicate variables that are not used in the backward graph. Figure taken from [8].

application of Graph Neural Networks in Chapter 10 is based on PyTorch, as with PyTorch Geometric [92] a very convenient framework for these purposes is provided, that is also well documented. In Chapter 10, the theoretical background of Graph Neural Networks is then also introduced.

Based on the dataflow programming paradigm in TensorFlow, computational graphs in the form of directed graphs are produced. These graphs result in the big advantages of TensorFlow compared to conventional programming methods.

On the one hand, there is the very efficient and stable gradient computation via automatic differentiation, short AutoDiff. Other forms of gradient computation are numerical differentiation and symbolic differentiation. AutoDiff is more accurate and less compute intense than numerical differentiation [93]. The accuracy is of the order of machine precision. In contrast to symbolic differentiation, it delivers efficient runtime calculations. In TensorFlow, procedures for the derivative are registered for primitive differentiable operations. Now, for each tensor gradient tensors are computed by extending the computational graph [8]. After first creating the computational graph in forward direction, it backtracks again and for each node the partial derivatives are calculated using the chain rule. This procedure is visualized in Fig. 6.1. The goal is not to calculate a formula as in symbolic regression, but an efficient and accurate computation.

On the other hand, computational graphs deliver the opportunity to (automatically) split the execution of the program onto a broad variety of multi-device settings in the context of a parallelized execution. The generation of the graph and the execution of the graph are different steps, and can be done in different environments. For instance, one can create a model in a python environment and the execution can be performed using the C++ API. For the following implementations this functionality was used for the coupling of ROOT environments with TensorFlow models. Therefore, the rapid prototype advantages of for instance python and jupyter notebooks were exploited in this thesis. Worth to mention is also the TensorFlow profiler, which can be used to precisely analyze where time consumption of the models happen and where optimization can give improvements.

7 Stand-Alone Tracking with the TRD

In this chapter, the algorithm and performance of the TRD stand-alone tracking is shown. The TPC and the other detectors are not used in this approach. Thus, the reconstruction is completely independent. Particles that are produced in the outer region of the TPC and in front of and within the TRD can be observed with this approach. In the following chapters, the photon conversion will be used as main application. A new TRD tracklet reconstruction and calibration was developed for Run 3 [94]. A variation of this reconstruction is used within this work for the stand-alone tracking. Here, not only TRD tracklets along the TPC tracks are stored, as in Run 2, but all possible signals are measured. Lower momenta can be achieved by also allowing a larger inclination of the tracklets compared to the new standard tracklet reconstruction approach for Run 3. Using the modified approach, the full potential of the tracking procedure can be shown. This reconstruction is introduced in the first section. Then the properties of the used data set are shown. The challenges of the tracking are explained, and the algorithm is motivated. The following sections focus on track finding and parameter estimation. Finally, the performance is discussed.

7.1 Tracklet Reconstruction

As an input to the stand-alone tracking, the TRD tracklets are used with their position and direction vectors. Several steps are performed to obtain this information from the charge deposition generated by the charged particles which ionize the gas in the detector. The charge deposition is stored as ADC counts in 24 time bins. In LHC Run 3, TRD tracklets are generated online based on a new calibration. A prototype of this new calibration is applied for the data that is used within this work. The following description of the algorithm is based on [94]. Three important steps are performed: First, the spatial position of the ADC digits is calibrated for a shift due to the Lorentz angle and an inaccurate drift time. Then the relevant clusters are calculated using a cluster finder. Finally, a straight line fit of the clusters is applied that delivers the position and direction. In Fig. 7.1 the raw clusters and the corrected clusters are shown for the different time bins.

The TRD chambers are aligned in a way that the magnetic field is perpendicular to the electric field in the chamber. This leads to a drift of the ionized charges that is

7 Stand-Alone Tracking with the TRD

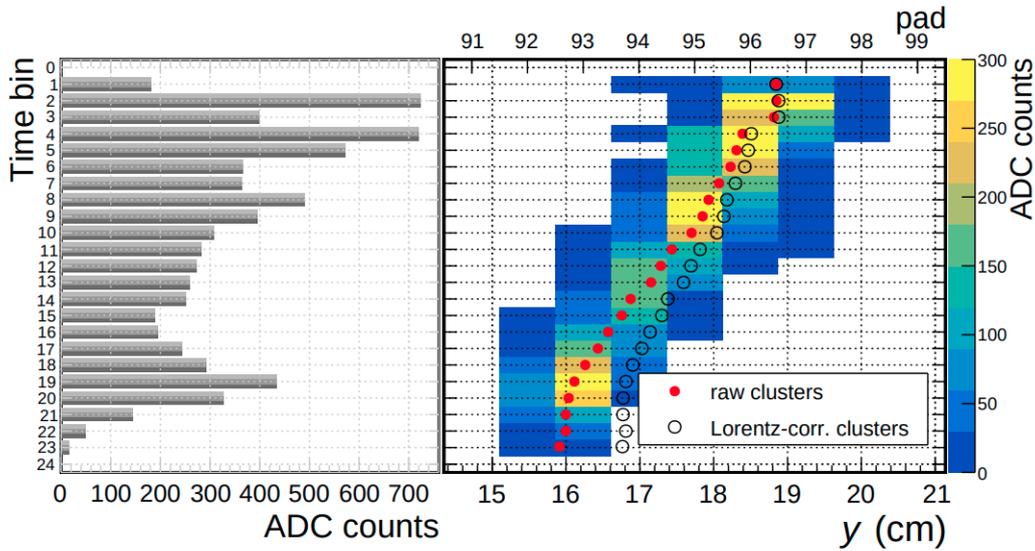


Figure 7.1: Left: The ADC counts for 24 time bins. Right: The raw cluster positions and the cluster positions after the correction are shown for the different time bins. Figure taken from [3].

not directly in the direction of the anode wire, but under a certain angle called the Lorentz angle. This affects the tracklet position and direction. The second effect is the drift time that may vary due to gas pressure, gas mixture fluctuations and/or high voltage changes. The slower drift of the backflowing ions also leads to an effective change of the drift velocity. By using a certain fit model, both effects are entangled and corrected for. This is done using default values as a starting point and the TPC tracks that are prolonged to the TRD as a reference.

The next step is the cluster finding. Starting point is the maximum ADC count in a chamber for each of the 24 time bins. The ADC counts in an interval of ± 3 pads in each column in y direction and ± 1 pads in each row in z direction are added. This procedure is performed for all ADC counts above a certain threshold. The position of the clusters is weighted by the ADC counts subtracted by the baseline.

These clusters are the foundation for the final fit. Up to 24 clusters from the different time bins are connected by a straight line fit. The straight line fit based on the three-dimensional orthogonal distance regression is performed using the so-called singular value decomposition (SVD). The intersection of the straight line with the anode wire plane is calculated. This position together with the direction vector defines the final tracklet.

7.2 Data Set

The stand-alone tracking was developed and tested using different data sets:

- simplified GEANT4 simulation of a TRD-like detector
- simple toy simulation of the TRD including detector effects
- 300k events p–Pb MC data at $\sqrt{s_{\text{NN}}} = 5.02$ TeV with additional pions injected
- 10M events p–Pb real data at $\sqrt{s_{\text{NN}}} = 5.02$ TeV

The purpose of the simplified GEANT4 simulation is to visualize the interaction with the material and obtain a clear impression of how the kinematics looks like. For this purpose, a simplified setup was used by taking six silicon layers instead of the gas chambers. The distance between the layers is equal to the real one and also the material budget was implemented in a realistic way. In addition, a few centimeters before the first layer, a lead plate was placed to induce additional photon conversions. Now, neutral pions and other particles can be injected and with the GEANT GUI the interaction and topological hits can be visualized. In addition, the exact hit position is saved in a file. Examples of photon conversions as given by this simulation are shown in Fig. 7.2. A goal of the tracking is to find and reconstruct such electron–positron pairs from photon conversion with a high accuracy.

With the experience of the GEANT simulation for test purposes of the tracking algorithm, a toy simulation was developed that includes the most relevant detector effects of the electrons: multiple scattering and bremsstrahlung. Also resolution effects and the pad tilting of the TRD can be tested. With this simulation, all these effects can be analyzed separately. This separation is especially important for development of tracking, debugging, performance tests and cross checks.

Roughly 300k p–Pb MC data were simulated at $\sqrt{s_{\text{NN}}} = 5.02$ TeV using the HIJING event generator [95] to test the functionality and obtain corresponding performance plots. For later parts of this work, in addition to the particles coming from the p–Pb collision, 50 extra neutral pions were added to each event. The transverse momentum of these injected neutral pions lies between 1 and 5 GeV/ c as in this region there are naturally fewer pions in the production distribution from the collisions due to the steeply falling pion transverse momentum spectrum. The η range is between -0.85 and 0.85. Two drawbacks should be noted here. On the one hand, with the given TRD calibration the real label of the corresponding tracks to the tracklets is not propagated through the reconstruction steps and thus not available on the tracklet level. Thus, an additional matching was performed of the MC tracks with the tracklets. Here, it can happen, that a track is lost due to the interaction

7 Stand-Alone Tracking with the TRD

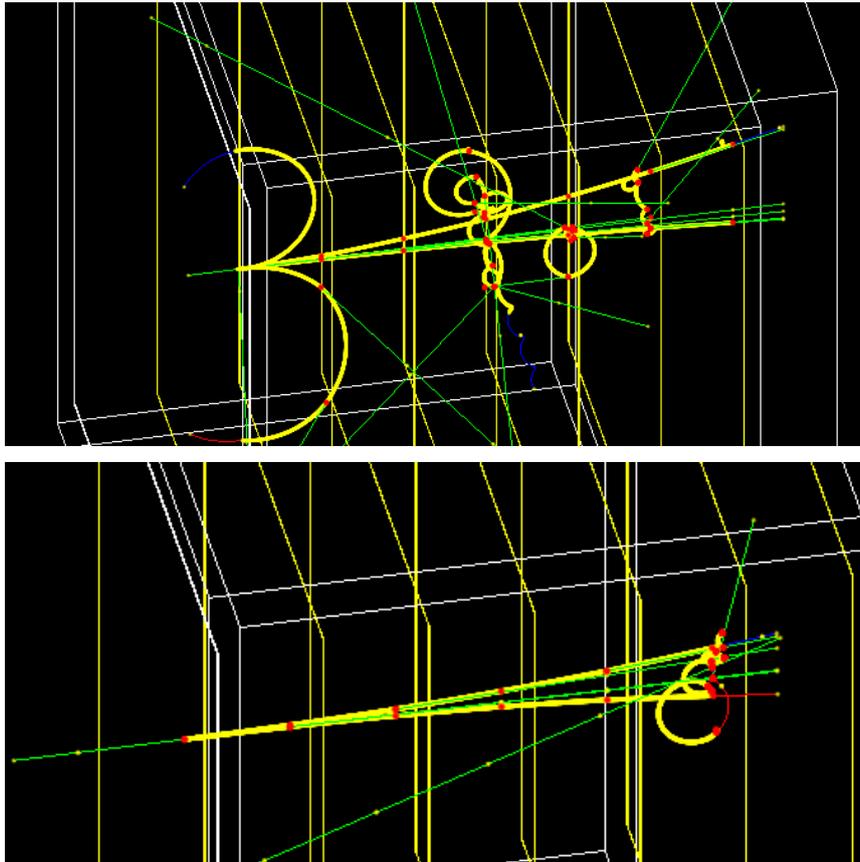


Figure 7.2: Two photon conversions from a simplified GEANT4 TRD simulation. Photons are indicated by green lines, while the charged tracks are shown in yellow. The hits are indicated by the red points. Many photons from bremsstrahlung can be seen. A goal of the tracking is to find and reconstruct the electron/positron tracks.

with the material and not matched. Also fake tracklets might be assigned. In the corresponding sections, more information regarding this procedure will be provided. On the other hand, the calibration was applied to real data and there might be a difference in the drift time and thus the MC data is most probable calibrated worse than the real data. But, in contrast to the real data, the alignment of the chambers in the MC data is perfect.

The tracking was also tested using 10M real p–Pb events at $\sqrt{s_{\text{NN}}} = 5.02$ TeV at the beamtime LHC16q. To finally achieve a good performance with these data is the most important goal. An event display of a full p–Pb event is shown in Fig. 7.3.

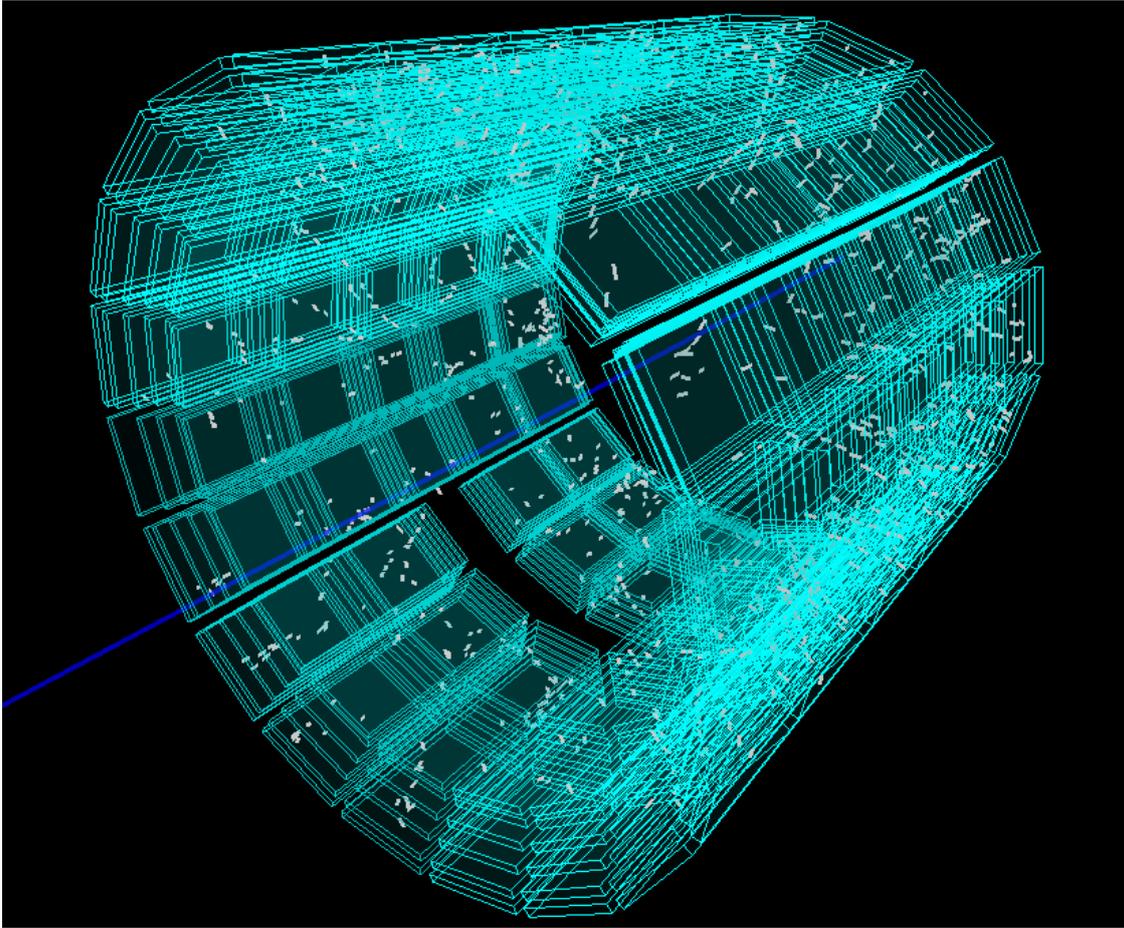


Figure 7.3: All tracklets of a real real p–Pb event in an event display. The TRD chambers are indicated by the boxes.

7.3 Challenges of TRD Stand-Alone Tracking

To motivate the tracking algorithm, it should be clear which challenges for the tracking should be kept in mind. All considerations are made with respect to reconstructing photon conversions in front of and within the TRD, and other tracks that are generated far away from the interaction point. It was shown, that only about 6% of the tracklets generated by the standard TRD online tracking are produced by primary tracks in [62]. With the new tracklet reconstruction algorithm used in this work, it is expected to be similar. In the case of primary tracks, a track seeding using the TPC tracks should be favored and used as in [62]. In this work, tracklets from particles that do not produce TPC tracks should also be considered.

7 Stand-Alone Tracking with the TRD

In this section, the focus is on the main challenges of the stand-alone tracking:

- the material budget
- noise tracklets
- empty layers
- shared clusters
- z pad resolution
- pad tilting

7.3.1 Material Budget

The reason for the large number of tracklets generated by non-primary tracks is the increased material budget in front of and within the TRD. The distribution of the radiation length as a function of the radial distance to the interaction point is shown in Fig. 7.4. The increased amount of material in front of and within the TRD is clearly visible. Four categories of material sum up to a total of 2.85% for one TRD chamber in units of radiation length X/X_0 : the radiator (0.69%), chamber gas and amplification region (0.21%), the pad plane (0.77%) and the electronics (1.18%) [3]. This leads to different effects like bremsstrahlung and multiple scattering that have a significant effect on the tracking.

7.3.2 Noise Tracklets

Under the given circumstances a lot of tracklets cannot be assigned to a certain track. As mentioned before, only 6% of the tracklets are generated by primary tracks. The noise tracklets are generated due to the high material budget, for instance from delta electrons or other knockout particles as well as electrons and positrons from conversions. Also detector noise might play a role. Other sources for noise tracklets are lost tracks due to multiple scattering with large angles or bremsstrahlung. All these tracklets cannot be assigned to a certain track, either because they do not belong to a track or there is not enough information to reconstruct the track. These tracklets disturb the track reconstruction process and fake matches cannot be completely avoided at high fake densities, for instance by curling tracks. As a consequence, the tracking algorithm must be robust against outliers.

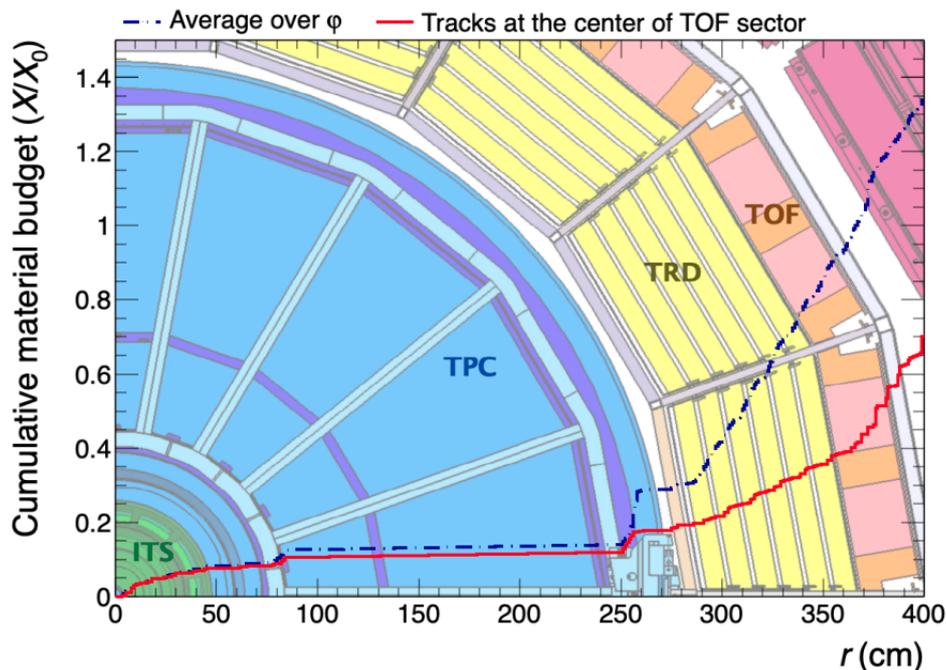


Figure 7.4: The material in terms of the radiation length X/X_0 as a function of the radial distance to the interaction point for tracks in the xy plane. The center of a TRD sector is equal to the center of a TOF sector. The large contribution to the material budget in front of and within the TRD is clearly visible. Figure taken from [96].

7.3.3 Empty Layers

A particle might not produce tracklets in all six layers. This leads to empty layers. If only tracks with six tracklets are considered, the reconstruction efficiency suffers and a lot of reconstruction potential would be lost.

There are different reasons for empty layers. Tracks might get lost at the edges of a chamber or due to large bremsstrahlung, multiple scattering with large angles, and some particles are annihilated. Also some photon conversions occur within the TRD such that the most inner layers naturally have no signal. Some chambers in the middle of the track might not have a signal. There are different reasons for this. For instance, a chamber could have a defect. Thus, another requirement is that the tracking algorithm should work for all combinations of quartets spread over the six layers in a row.

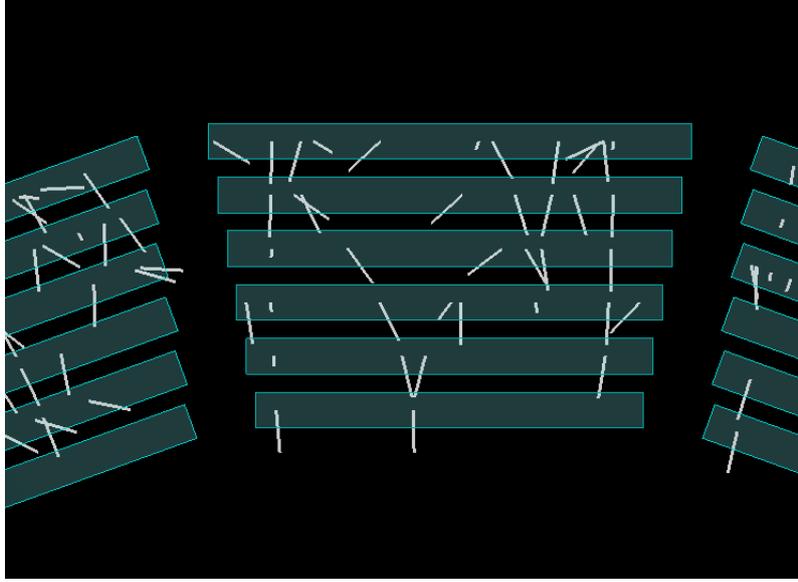


Figure 7.5: In the middle of the shown sector the tracklets of a typical photon candidate from the p–Pb MC simulation can be seen in the xy plane. The tracking algorithm must be capable of reconstructing the electron and positron created by the photon conversion.

7.3.4 Shared Clusters

Shared clusters of multiple tracks are produced within the TRD in case two tracks are very close. For instance, these are generated by photon conversions, where the electron and positron especially at higher momenta cross the detector very close to each other if these conversions happen in a TRD layer. In this case, the clusters are spread over more pads and thus, the position resolution as well as the direction resolution of the shared tracklets deteriorates. This leads to "natural" outliers. Also the tracking has to be flexible and a tracklet might be assigned to more than one track. The tracking algorithm should respect non-fake tracklets e.g. from other particles nearby. A shared cluster as given by an electron–positron pair from a photon conversion can be seen in Fig. 7.5.

7.3.5 Z Pad Resolution

The resolution of the pads in z direction is significantly lower than the resolution in y direction. The pads have sizes of $(0.515 - 0.965) \times (7.5 - 9.0) \text{ cm}^2$ ($r\phi \times z$). The resolution given by the pad size is $l_{\text{pad}}/\sqrt{12}$. This corresponds to a resolution in y direction of around 0.2 cm and in z direction of around 2 cm. In reality, the resolution depends on the inclination angle and is calculated for y via a fit over

7.3 Challenges of TRD Stand-Alone Tracking

many measurements given by the time dependent signal in several time bins. Also the charge sharing over adjacent pads comes into play. In case of the y direction, this leads to effective values up to 0.04 cm for an optimal inclination angle in the direction of the Lorentz angle [62]. In the case of the online tracking, the MCMs are only connected to one pad row and thus clusters in z that are spread over more than one pad cannot be found. The z resolution is therefore limited to the size of the pad. A spread over more than one z pad is considered in the reconstruction used for the studies shown in this work, but are expected to have a low impact. By using a pad tilting, the z resolution is improved after accounting for this effect. More details can be found in the next subsection. In the later course of this thesis it is shown via a toy simulation that this pad tilting improves the z resolution by a factor of two.

The large z pad resolution has different consequences to be considered for the stand-alone tracking. First of all, the z component of the direction vectors contain little information, only direction vectors in the xy plane are useful. The second point is that for a full helix fit the weights that determine the relation between the residuals in y and z must be set appropriately. The position estimate in z is also not as good as in the xy plane.

The impact of the size of the z resolution on $\tan \lambda$ and the z component of the momentum can be estimated using the solution of an analytical linear fit. For the assumption of a track with infinite momentum the curvature is zero and a straight line fit can be performed as a function of the length/arc length in the xy plane. The slope of the track is then given by $\tan \lambda$. The uncertainty of the slope parameter $\tan \lambda$ is given by

$$\Delta \tan \lambda = \frac{\Delta z}{\sqrt{n(\langle s^2 \rangle - \langle s \rangle^2)}}, \quad (7.1)$$

where Δz is the resolution in z direction, n is the number of measurements and s is the length/arc lengths in the xy plane. This relation is in general also approximately correct for bent tracks where the momentum is not infinite. A large variance of the arc length of the measurements leads to a lower resolution of the inclination due to the larger lever arm. The resolution of the momentum in z direction can then be found using $p_z = p_T \cdot \tan \lambda$ as

$$\Delta p_z = \sqrt{(\Delta p_T \cdot \tan \lambda)^2 + (\Delta \tan \lambda \cdot p_T)^2}. \quad (7.2)$$

Typical values for the uncertainty of $\tan \lambda$ lie between 0.03 and 0.044. The relative p_T resolution above 10% at the given momenta as shown later for the stand-alone tracking is thus dominant for the total momentum.

Another effect of the low resolution in z direction is the impact on possible multiple

7 Stand-Alone Tracking with the TRD

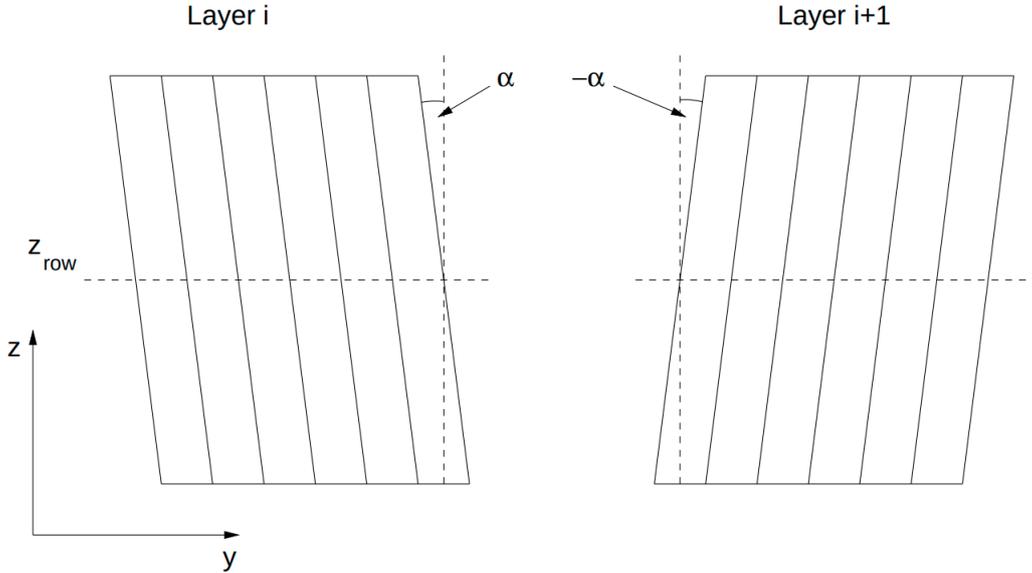


Figure 7.6: The pads of the TRD are tilted alternating from layer to layer as indicated in this sketch. Figure taken from [97].

scattering corrections. To exploit the effect of multiple scattering in the tracking, the fact is used that the angle in y direction and the angle in z direction are correlated. If it is not possible to resolve the multiple scattering angle in z direction, it is not possible to correct for multiple scattering. Arguing via information theory, two variables are added to the measurement, so two measurements are needed to correct for this. Otherwise it just disturbs the measurement. In the case of the TRD a scattering angle of roughly 1° is obtained, if assuming a very low momentum electron with $100 \text{ MeV}/c$ for a mean radiation length of 2.28% (one layer), calculated via Eq. 4.9 in Ch. 4. For an arc length of the helix of 15 cm between two layers, a mean angle of roughly 4° is obtained when considering a z resolution of 1 cm . Thus, the possible uncertainty of a multiple scattering angle measurement between two TRD layers is much larger than the real scattering angle.

7.3.6 Pad Tilting

To tackle the low z resolution, the pads are tilted [3][97]. A visualization is shown in Fig. 7.6. The tilt is $\pm 2^\circ$ in the yz plane with respect to the z axis alternating the sign with each layer. This very small tilt decreases the y resolution only marginally while improving the z resolution. By introducing the corresponding correlations to the uncertainties, an improvement of the z resolution is possible. The covariance

matrix for the measurements is then given by [62]

$$\mathbf{Cov}_{\text{TRD}} = \mathbf{R} \begin{pmatrix} \sigma_y^2 & 0 \\ 0 & \sigma_z^2 \end{pmatrix} \mathbf{R}^T. \quad (7.3)$$

A rotation \mathbf{R} is applied via

$$\mathbf{R} = \begin{pmatrix} \cos \beta_t & -\sin \beta_t \\ \sin \beta_t & \cos \beta_t \end{pmatrix}, \quad (7.4)$$

where β_t is the tilting angle. An alternative to implementing the tilt to the covariance matrix is a change of the coordinate system at the corresponding measurement position. By changing to a local rotated coordinate system with the pad tilt an uncorrelated measurement can be performed.

7.4 Helix Fit

In this section, it is shown why the least squares helix fit is chosen for the track parameter estimation. The gold standard for charged particle tracking is the Kalman filter. Nevertheless, the Kalman filter is not used in the presented approach. This section is about the track parameter estimation. The track finding is described in the next section.

The goal is to achieve a maximum efficiency in the track finding besides an optimal parameter estimation. This means that with empty layers in between, a seeding algorithm for the Kalman filter must be implemented that already iterates through nearly all possible combinations, as only a maximum of six layers is available. A usable seeding needs at least three tracklets. Thus, the track finding advantages of the Kalman filter are not useful here. A combinatorial search must be implemented anyway. The quality of the track parameter estimation is discussed after the fit itself is introduced and then compared to other parameter estimation methods.

7.4.1 Fit Algorithm

For the track fitting, a real helix fit is performed in several stages in the track finding. In the context of the implementation, the TensorFlow API with its autodiff and vectorization capabilities was used. Thus, no linearization is needed and the gradients are calculated automatically. This allows an easy feasibility study and fast and flexible prototyping.

7 Stand-Alone Tracking with the TRD

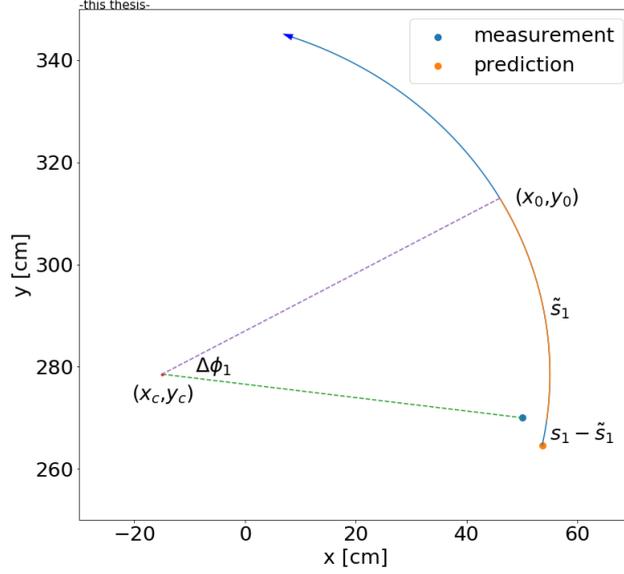


Figure 7.7: A sketch of the calculation of the residuals, i.e. the L^2 distance between measurement and prediction. The value of s_i converges to \tilde{s}_i .

The starting point is the track model given in Ch. 4. The implementation starts with the parametrization

$$x(s) = x_0 - \frac{\sin(\phi_0)}{C} + \frac{\sin(\phi_0 + Cs)}{C} \quad (7.5a)$$

$$y(s) = y_0 + \frac{\cos(\phi_0)}{C} - \frac{\cos(\phi_0 + Cs)}{C} \quad (7.5b)$$

$$z(s) = z_0 + \tan \lambda \cdot s \quad (7.5c)$$

with x_0 , y_0 and z_0 being a point on the helix and s a running parameter along the arc of the circle. For a curvature of $C = 0$ the singularity is removable and thus a continuous transition from positive to negative charges is possible. In the implementation the $1/C$ dependence was removed via trigonometric identities. The center of the circle of the helix in the xy plane are given by

$$x_c = x_0 - \frac{\sin \phi_0}{C} \quad (7.6a)$$

$$y_c = y_0 + \frac{\cos \phi_0}{C}. \quad (7.6b)$$

As there is freedom in the choice of x_0 and y_0 , now s_i is set such that $\sum s_i = 0$ holds, where i is the number of the layer. The points x_0 and y_0 are close to the center of gravity of the hits. Since the center (x_c, y_c) is correlated with the curvature C , the

chosen parametrization has the advantage that (x_0, y_0) is nearly uncorrelated to C . The iterative algorithm starts with an initial guess for x_0, y_0, z_0, ϕ_0, C and $\tan \lambda$. The following residuals are calculated:

$$\text{res}_{x,i} = x_i - x(s_i) \quad (7.7a)$$

$$\text{res}_{y,i} = y_i - y(s_i) \quad (7.7b)$$

$$\text{res}_{z,i} = z_i - z(s_i). \quad (7.7c)$$

The local running parameter s_i along the arc is calculated via $\Delta\phi_i/C$:

$$\tilde{s}_i = \frac{\arctan2(C(y_c - y_i), C(x_i - x_c)) - \phi_0}{C}. \quad (7.8)$$

This parameter is then normalized to $s_i = \tilde{s}_i - \sum \tilde{s}_i$ to fulfill $\sum s_i = 0$. The value of s_i converges to \tilde{s}_i . In this way, the chosen points on the circle stabilize the fit. These values for s_i are used to calculate the residuals. The residuals are in the xy plane and not only in y direction, but the deviation is sufficiently small. The residuals are weighted by rough estimates of the resolution. The full χ^2 is given by the sum of the weighted squares of the residuals. This can now be minimized using gradient descent, resulting in a stable set of parameter estimates. A sketch of the calculation of the residuals is shown in Fig. 7.7.

This fit is converging within very few optimizations steps. The amount of time spent for the fitting is much less than for the track finding itself. In contrast to the Kalman filter, this procedure is not very sensitive to the seeding of the parameters and a seeding of the covariance matrix is not needed at all. The fit results are equal to the Kalman filter in case there is no material and no corresponding adjustment of the covariance matrix.

7.4.2 Advantages

The primary goal of the fit procedure is to have a tool for tracklet association to track candidates. A high precision fit for electron/positron pairs from photon conversions is presented later. Here, also an increase of the purity can be reached. The found parameters are then used as initial guess for the precision fit.

A commonly used choice for helix fits is the Karimäki fit as shown in Ch. 4. It provides an analytical solution, i.e. it is non-iterative. But it was noticed that the use of the square in Eq. 4.17 leads to a higher sensitivity to outliers. The consequence is a systematic bias towards higher momenta. In the previous section it was already discussed that a certain robustness towards outliers is a requirement for the tracking. In addition, the used helix fit is a full three-dimensional fit that uses

7 Stand-Alone Tracking with the TRD

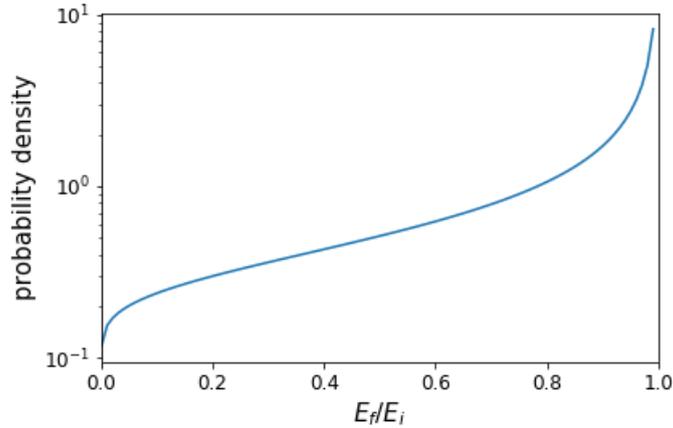


Figure 7.8: The probability density for the relative energy loss according to the Bethe Heitler formula (Eq. 4.8) at $\langle X/X_0 \rangle = 25\%$. E_i is the initial energy and E_f the energy after emitting bremsstrahlung.

additional information about the curvature given from the z resolution.

A big advantage of the Kalman filter is the uncomplicated handling of material corrections. Due to the following circumstances, these corrections are not straightforward. On the one hand, there is multiple scattering. The extra information from the scatter angles cannot be decomposed due to the large z pads as discussed before. Thus it only adds a smearing to the covariance matrix. In case of electrons, which is the most important observable in this work, multiple scattering is a minor effect. On the other hand, for energy loss corrections a particle identification guess would be necessary. The fact that electrons/positrons from photon conversions are in main interest in this work, one must be aware of bremsstrahlung. As shown in Ch. 4, over the full range of momenta, this is the dominant effect. Fig. 7.8 shows the distribution of the relative energy loss for a material thickness of $\langle X/X_0 \rangle = 25\%$. This corresponds to the full material budget in the TRD. Electrons and positrons might be created already in front of the TRD and even lose more energy. An integration shows an energy loss of more than 5% for 2/3 of the electrons/positrons. There is no correction that can be applied to the Kalman filter to be added to the covariance matrix as this is not a Gaussian effect. In general there is an adaption of the Kalman filter that is capable to handle these kind of uncertainties, the Gaussian Sum Filter [78]. But the given approach provides a sufficient track fit quality while having a low complexity.

The chosen approach is sufficient for the candidate finding. The corresponding

electrons/positrons from photons are then cleaned and fitted later with a higher precision, and then also the bremsstrahlung is taken into account. Thereby, the purity is also increased. Moreover, the influence of the pad tilting is then used to increase the quality of the parameter estimation.

7.5 Combinatorial Search

The input to the tracking are tracklets with a position and a direction vector from up to six TRD layers. The tracking has to face the following challenges: Empty layers, noise tracklets, limited resolution of direction vectors, large z pad size and high material budget.

To face the different challenges of the TRD stand-alone tracking, a purely combinatorial tracking was found to be optimal. To be more precise, this combinatorial tracking is a hierarchical search. A sketch of the track finding procedure is given in Fig. 7.9. The direction vectors in the xy plane of the tracklets are used to find pairs of tracklets. These pairs are used to build quartets, which are then combined to quintets and hexads. It is obvious that the runtime behavior is proportional to the number of relevant combinations, i.e. quadratic, and there is a strong filtering effect due to the hierarchical nature of the algorithm.

The track finding and fitting is separated, even though in some steps in between fits are used as selection criterion. A cellular automaton as an alternative, as used in other detectors, could be used in a similar manner, but is more complex. Combinatorial pairs would also be constructed. The chosen combinatorial track search is suited for a high efficiency, depending on the demands. Shared clusters, necessary for the reconstruction of photon conversions, can be handled. The final tracks must include at least four tracklets, which are necessary for the fit to have enough degrees of freedom, as the resolution of the direction vector is not sufficient.

The track finding consists of the following steps which will be explained in more detail in the further course of this section: pair finding, quartet finding, quintets and hexads, cleaning hexads, cleaning quintets, mixed cleaning and cleaning quartets. The calculations are performed in the local coordinate system of each sector.

Tracklet pairs (T_i, T_j) are searched within one sector. They should lie either in the same stack or a neighboring one. The layer of T_j is required to be further out than the layer of T_i . The maximum distance between the tracklets is $\Delta_l \times 30$ cm, where Δ_l is the number of layers between the given tracklets. The direction vectors \vec{v}_i and \vec{v}_j are then used in the following angle condition: $\cos(i, j) \equiv \frac{\vec{v}_i + \vec{v}_j}{|\vec{v}_i + \vec{v}_j|} \cdot \frac{\vec{v}_{ij}}{|\vec{v}_{ij}|} < \text{cut}$, with the

7 Stand-Alone Tracking with the TRD

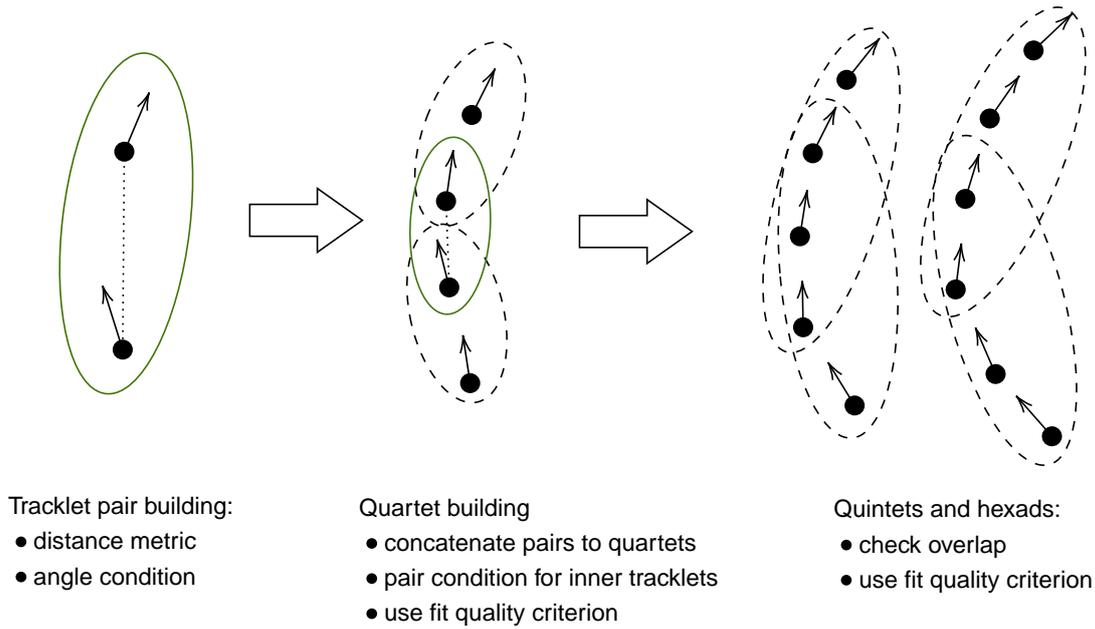


Figure 7.9: The different hierarchical steps of the track finding. After quartets, quintets and hexads are constructed, these are cleaned in case they are similar and overlapping, as described in the text.

cut set to 0.9. The vector \vec{v}_{ij} is the mean direction between the position of tracklet i and tracklet j , which is equal to the pointing vector from the position of tracklet i to the position of tracklet j . This is motivated by the fact that the normalized mean direction of two direction vectors on a circle is equal to the normalized connecting direction vector of the two points.

Quartet finding starts with a combination of all tracklet pairs (T_i, T_j) and (T_k, T_l) . The criteria for pairs are applied to (T_j, T_k) and (T_i, T_l) . These candidates are fitted with a helix fit, then a maximum curvature and a z constraint are applied. The result is a sample of many quartets with any possible overlaps.

Quintets and hexads are constructed by combining two quintets $(T_{i,a}, T_{j,a}, T_{k,a}, T_{l,a})$ and $(T_{i,b}, T_{j,b}, T_{k,b}, T_{l,b})$ with $T_{k,a} = T_{i,b}$ and $T_{l,a} = T_{j,b}$ for hexads, for quintets it applies $T_{j,a} = T_{i,b}$, $T_{k,a} = T_{j,b}$ and $T_{l,a} = T_{k,b}$.

Now there is a large number of overlapping track candidates. So a **cleaning of the quintets and hexads** is applied and only allows overlap in the first layers. All

candidates are fitted again and χ_{xy}^2 and χ_z^2 cuts as well as a curvature constraint are applied.

A **mixed cleaning** is implemented to reject all quintets that are part of the hexads, quartets that are part of the hexads and quartets that are part of the quintets. In case the fit of a hexad is much better than a fit of a quintet, the hexad is rejected due to a possible outlier tracklet.

The **quartet cleaning** is then done in the same way as for the quintets and quartets.

The result is a sample of track candidates with four to six tracklets, the associated track parameter estimates and fit quality parameters. The cuts are chosen such that the efficiency is sufficient high ($\geq 95\%$). The purity may suffer, but all possible photon candidates should be stored. A higher purity is achieved by using more precise photon fits later in the case of the conversion candidates.

7.6 Tracking Output Evaluation

Due to the limited MC information that is provided by the given data sets, it makes sense to first obtain a general impression of the performance of the tracking. The MC data are limited as there are no labels assigned to the tracklets and only the track parameters at the production vertex and the production vertex itself is given as MC truth information. These parameters change when crossing the material, and this local information is not provided. Thus, a purity cannot exactly be verified. A more detailed analysis of the information given by the MC data is shown in the next section. Nevertheless, some impression is given here for two variables for the p-Pb MC data with added neutral pions and the real p-Pb data, both at $\sqrt{s_{NN}} = 5.02$ TeV. The focus is on the number of tracks per event and the number of tracklets per track.

The first variable is the number of tracks per event found by the stand-alone tracking in comparison to the TPC tracks. This is shown in Fig. 7.10 in a two-dimensional histogram for MC and real data, respectively. The number of found TRD tracks is much lower than the number of TPC tracks from the standard track reconstruction. The injected neutral pions for the MC data lead to a minimum number of tracks, which shifts the distribution to more TRD tracks.

The acceptance of the TRD is reduced by the space between the supermodules and also by faulty chambers. This reduces the number of tracks in the TRD. Besides that, primary low p_T tracks do not reach the TRD. In the real data, many high multiplicity TPC events have even lower TRD multiplicities due to pile-up in the TPC. This effect is not present for MC data. In reality, some TRD chambers

7 Stand-Alone Tracking with the TRD

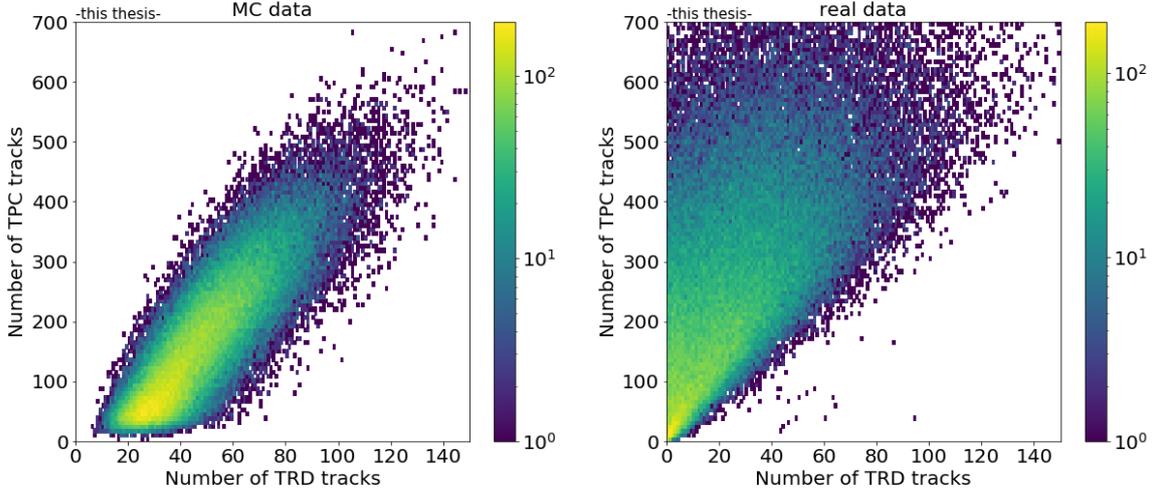


Figure 7.10: Left: Number of TPC tracks versus number of TRD tracks per event for p–Pb Monte Carlo data with added π^0 s. The TRD tracks are found by the TRD stand-alone tracking. Right: The same for real p–Pb data. In general, there are differences between the real and the MC data, as described in the text.

are broken for a certain period of time and the calibration is not always working properly. The ratio between TPC tracks and TRD tracks for the MC data is roughly 30% and for the recorded data only 10%. The added neutral pions in the MC data result in a large number of photons. The electrons/positrons from the photon conversions are quite well reconstructed. This can be seen, as there is a similar number of electron–positron pairs produced within the ITS before the TPC, as in the material before and within the TRD, and a shift of the number of tracks arising from these particles is more significant in the TRD than in the TPC. Low p_T and heavy particles experience a lot of multiple scattering in the TRD which is not the case for the electrons from conversion.

Figure 7.11 shows the number of quartets, quintets and hexads per event. Whereas there are more quintets than hexads, the number decreases for the quartets. This pattern is the same for real and MC data. Another property of the implemented tracking is that there are tracks that are counted twice. This happens if two tracklets are close to each other and the requirement for the direction vector is satisfied. These tracks will be cleaned for the conversion finding later, where a more precise fit is applied and so the choice which candidate is more appropriate can be made.

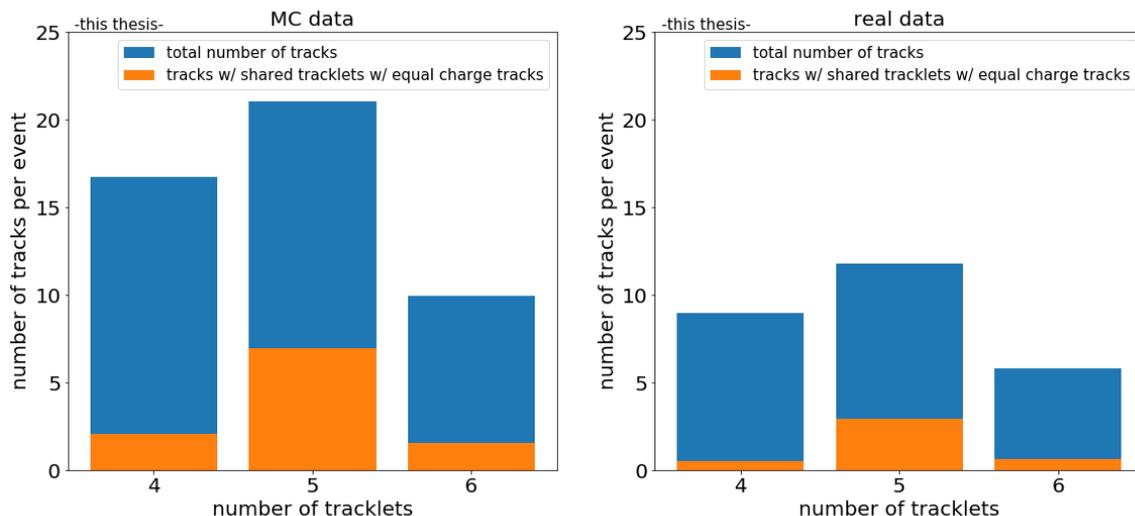


Figure 7.11: The number of tracks per event with a certain number of tracklets for the MC data with additional π^0 on the left and on the right for the real p–Pb data. The number of tracks which share a cluster with another track of the same charge is shown with the orange part of the bar.

Up to this stage this leads to a certain acceptable impurity. In Fig. 7.11 the number of tracks that share a tracklet with another track with the same charge is shown.

7.7 Monte Carlo Performance

For a more quantitative impression of the performance of the stand-alone tracking the given MC data are analyzed. Special attention is given to the secondary particles, which are of special interest in this work, i.e. electrons/positrons that might arise from photon conversions.

As mentioned before, the exact association between the tracklets and the MC tracks is not possible with the given data set. Only the production vertex, the momentum information at the production vertex and the particle species are given as MC truth information. Thus, to obtain performance plots, a matching is performed in two steps. First, the tracklets are assigned to the MC tracks. Then, the tracklets found by the tracking are compared to those from the MC matching. While propagating through the detector, the local momentum changes. Thus, especially for primary particles with a long path through the material in front of the TRD, the momentum might significantly differ from the original momentum at the production vertex. As

7 Stand-Alone Tracking with the TRD

the focus is on secondaries that are created closer to the TRD the impact of this effect decreases. But still, for instance particles with a lot of multiple scattering or electrons with a large bremsstrahlung contribution might be lost. Nevertheless, up to a certain level a clear impression of the functionality is given.

On the basis of the MC data three quantities are discussed: The efficiency with respect to MC p_T , the reconstructed p_T compared MC p_T and the closest distance to the production vertex in the xy plane dca_{xy} , which is especially important for photon candidate finding.

The purity of the tracking is not shown, as on the one hand, tracks might produce tracklets, for which the path changed so much that the tracklets cannot be assigned to the MC track at the vertex anymore. On the other hand, as shown in the previous section, there might be more than one track candidate corresponding to one MC track. The tracking aims to find candidates for the photon conversion fit, where a more precise fit gives more precise selection abilities.

7.7.1 Matching

To obtain performance plots, first the MC tracks and the TRD tracklets have to be matched. A tracklet is matched to a track if three criteria are fulfilled:

- the distance in xy must be smaller than 3 cm
- the distance in z smaller than 8 cm
- the cosine of the angle between the direction of the track and the tracklet direction must be larger than 0.92

These criteria already tolerate a certain amount of energy loss and multiple scattering. Heavier particles and/or particles with a lower momentum are nevertheless more likely to not fulfill the criteria. Tracks with at least four assigned tracklets are taken into account as tracks that can be found by the tracking, as this is the minimum number of tracklets required by the tracking. If more than one tracklet fulfills the conditions, the one that is closer in the xy plane is taken.

A track from the tracking is matched to one of these MC tracks in case it shares at least three tracklets, as only in this case enough information is given to determine the structure of the helix.

7.7.2 Efficiency

The number of MC tracks that are matched to at least four tracklets that share at least three tracklets with a stand-alone tracking track is the baseline for the presumed reconstructable efficiency. Certainly this number must be seen in view of

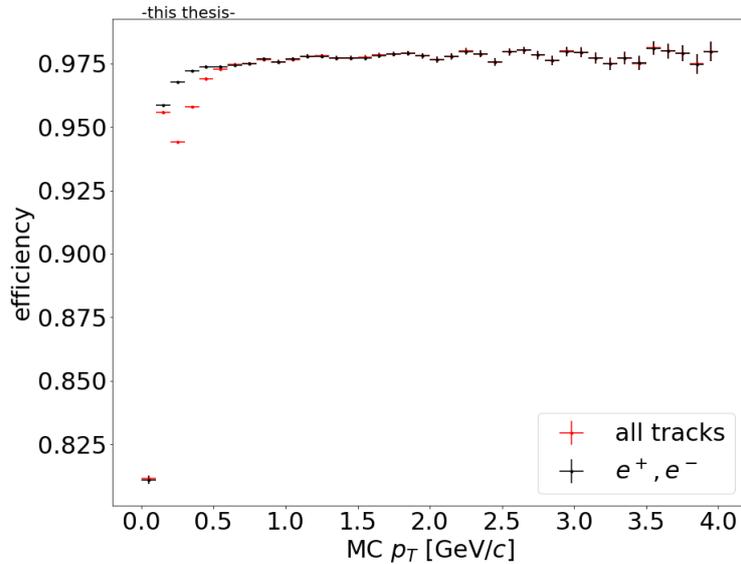


Figure 7.12: The reconstructable tracking efficiency (see Sec. 7.7.2) for all particle species as well as for electrons that are produced at vertices in radial direction between 250 cm and 330 cm. At very low p_T as well as at p_T larger than 0.6 GeV/ c nearly all tracks are electrons. Thus, in these regions the efficiency is equal for both shown samples. As other particles than electrons/positrons have a higher multiple scattering contribution the performance for them deteriorates at low p_T .

the limitation of the MC tracks to tracklet matching. Some reconstructable tracks might be lost, others might be fakes.

The resulting efficiency is shown as a function of transverse momentum in Fig. 7.12 for all particle species and for electrons/positrons. Only particles that are produced between 250 cm and 330 cm in radial direction are considered for the calculation. Primary particles are more likely to be fakes as mentioned before. The difference in decrease of the efficiency of the samples at low momenta is due to the particle composition where the other particles than electrons/positrons have a higher multiple scattering contribution, and thus the efficiency is lower. The efficiency starts at above 80% and quickly saturates at above 97% for the electrons/positrons at transverse momenta of about 300 MeV/ c and for all species at about 500 MeV/ c . Especially, the electrons/positrons that are important for the photon conversion studies are reconstructed very well. A part of the inefficiency may also come from fake MC matches.

7 Stand-Alone Tracking with the TRD

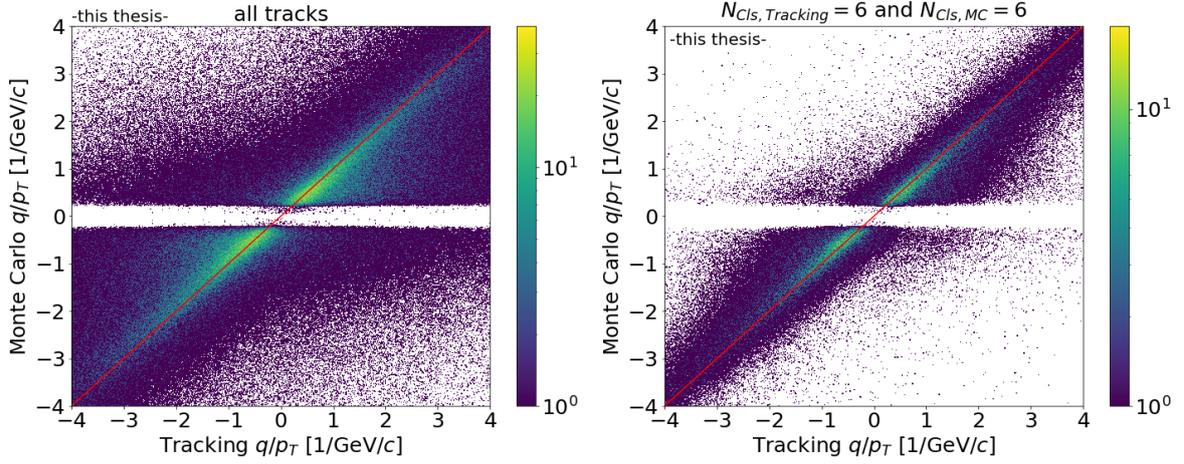


Figure 7.13: The inverse transverse momentum multiplied with the charge number q/p_T of the matched tracks from the tracking is shown versus the original q/p_T of the Monte Carlo information on the left side for all matched tracks and on the right side for the tracks with six tracklets in the Monte Carlo matching as well as the tracking. At the red line, Monte Carlo q/p_T and tracking q/p_T are equal.

7.7.3 Momentum Resolution

In Fig. 7.13 the tracking q/p_T , where q is the sign of the charge, versus the MC q/p_T is shown in two-dimensional histograms for all matched tracks and for the optimal case where the MC tracks have six tracklets assigned and also the tracks from the stand-alone tracking have six tracklets assigned. As expected, the TRD stand-alone tracking does not provide a very good p_T resolution, but the purpose of the procedure is the selection of the tracks. A rough estimate leads to a relative uncertainty clearly larger than 10%, increasing to higher values with increasing momenta due to the smaller lever arm. This momentum resolution was roughly confirmed by a toy simulation. The histogram delivers two other important insights into the tracking. The first one is that there is no clear shift of the mean. Thus, at this point it is sufficient to neglect the energy loss. The second point is that the tail from bremsstrahlung is clearly visible, especially for low q/p_T . The final fit for the photon conversion candidates delivers a more precise p_T resolution. Due to the large lever arm, this then provides an acceptable resolution that a stand-alone tracking of single tracks cannot provide without using the primary vertex in addition to the tracklets.

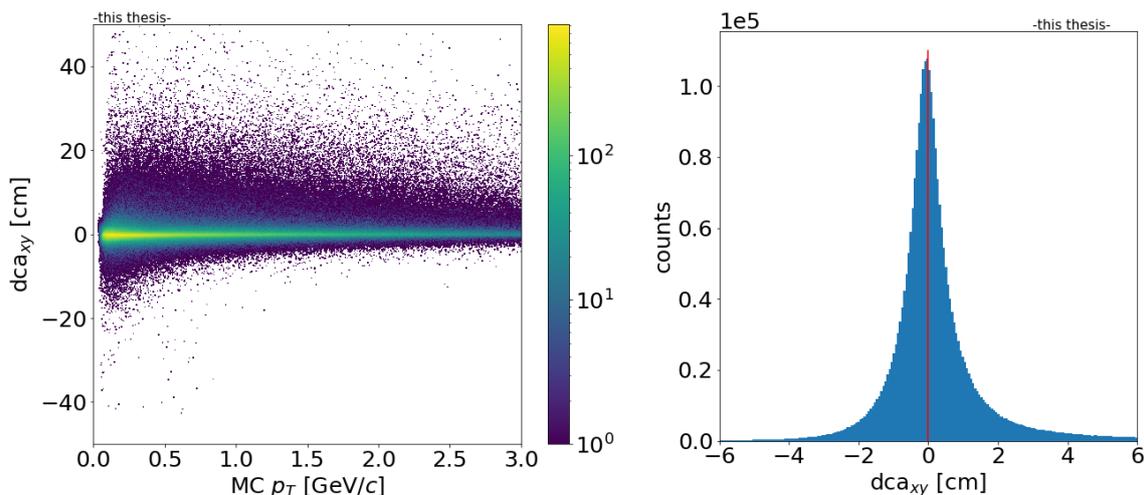


Figure 7.14: The minimal distance in the xy plane of the track to the production vertex dca_{xy} given by the Monte Carlo information is shown, on the left side in a two-dimensional histogram as a function of p_T , and on the right side, as a projection for the full p_T range. The red line indicates a perfect match. In case of values of dca_{xy} larger than zero, the production vertex is outside of the circle. Only those tracks with a production radius larger than 250 cm are shown.

7.7.4 Vertex Resolution

More important for the reconstruction of the photon conversions than the p_T resolution is the distance of the found track to the production vertex. In the following, for this quantity a signed value is used, and a negative sign means that the distance between the vertex and the center of the circle is smaller than the radius of the circle. A cut on this quantity is the main selection criterion for photon conversion candidates. In that case a smearing appears, as the vertex is not known and, instead of the distance between a point and track, the distance of closest approach between two tracks is calculated. A decreasing p_T resolution for large momenta has a smaller impact as the tracks get more straight. The distance in the xy plane is shown in Fig. 7.14 in a two-dimensional histogram as a function of p_T , and as a one-dimensional histogram. In the two-dimensional histogram it can be seen that the distance to the vertex is centered over the whole p_T range and the structure does not change. From the one-dimensional histogram it can be seen that the resolution is of the order of a centimeter. A tail from bremsstrahlung can also be seen. This allows a clear selection of candidates for photon conversion finding.

8 Synchronous Fitting of Daughter Tracks from Photon Conversions

In this chapter, the newly developed method of full synchronous photon conversion daughter track fitting using the conversion topology is introduced. In the following, this fit is called syncfit. The syncfitter operates on a preselection of electron–positron pair candidates based on the TRD stand-alone tracking and its helix parameter estimates shown before.

The aim is to identify electron–positron pairs from photon conversions in the TRD and to measure all necessary kinematic variables of primary photons. Before showing the performance on a full Monte Carlo simulation and real data in the next chapter, the concept and mathematical formulation is presented and evaluated using a toy simulation. This has the advantage that all detector effects can be analyzed separately and no distortions by a non-optimal track selection are included. This provides a detailed understanding of the syncfit and its performance. Thereby, this new method is also compared to the standard method for photon conversion daughter fitting, the KFParticle package [86][87]. Using the topological (geometrical/kinematical) constraints, a significant resolution gain of the photon momentum can be obtained compared to using the single fits and combine them. This is a consequence of the reduction of degrees of freedom within the fit.

8.1 Concept

In this section, the basic idea of the syncfit is described. The syncfitter simultaneously implements a classifier and regressor for photon conversion candidates. In this sense this algorithm provides a full synchronous fit of preselected TRD tracklets generated by daughter tracks of photon conversion candidates, i.e. the electron and positron track, simultaneously. Essential topological constraints given by the underlying physics of a conversion from a primary photon are implemented. This implies that a photon hypothesis test based on the fit quality can be used as selection criterion. In this context, a primary photon is defined as a photon originating at the primary vertex, which means its point of creation is indistinguishable close to the primary vertex.

As many of the criteria as possible given by the topology of the photon conver-

8 Synchronous Fitting of Daughter Tracks from Photon Conversions

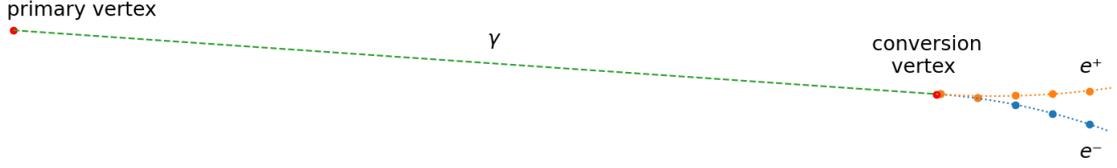


Figure 8.1: The full topology of a photon conversion is fitted via the syncfit, including the primary vertex, the conversion vertex as common point on the helices of the daughter tracks, and a vanishing opening angle is assumed. The long lever arm strongly constraints the direction of flight.

sion should be fulfilled. The following assumptions are taken into account:

- the primary vertex is known with infinite precision compared to the resolution given by the TRD
- the opening angle of the electron–positron pair can be assumed to be zero
- the conversion radius R in the xy plane is in a specified range $[R_{\min}, R_{\max}]$
- the daughter tracks cross the outermost TRD layer

By using the fact that the daughter tracks have a vanishingly small opening angle the constraints $\phi_{e^+} = \phi_{e^-}$ and $\lambda_{e^+} = \lambda_{e^-}$ are defined. These are the angles as defined by the helices. In addition, these angles are fixed by the common vertex: the conversion vertex. Then, only the curvatures C_{e^-} and C_{e^+} remain. So in total, the conversion vertex and the transverse momenta are the five free parameters. These are for instance given by the conversion vertex $(R \cos(\phi), R \sin(\phi), R \tan \lambda)$ and the transverse momenta of the daughter track represented by the curvatures C_{e^+} and C_{e^-} . Thus, it is obvious that the conversion vertex constrains the angles. A sketch of the topology is shown in Fig. 8.1. In contrast, two uncorrelated free helices have twelve parameters. The information gain by the syncfitter is the reason for the improvement of the momentum resolution and the reason why uncorrelated fake tracks are assigned a bad fit quality and can be rejected.

The algorithm proceeds in the following steps: The starting point is an initial guess based on the helix fits (seeds). The actual syncfitter is split into two parts: A stable fit that does not use the exact layer crossing positions. The residuals are minimized in the xy plane, i.e. not only in the y plane, and in the z plane separately. Starting at the conversion vertex, both daughter tracks are propagated iteratively from layer

to layer leading to certain predictions of the track in the local system at the layer crossing position. In the second step, the final syncfit is fed with the results of the first step as a seed, but with the exact layer crossing position and thus, with residuals only in the local y and z plane. Only in the second step, energy loss and pad tilting are included. Thereby a loose selection of good tracks is already applied, requiring a certain χ^2 criterion after the first step. Stabilizing regularizations are included in both steps. All derivatives of the complete chains of detailed formulas are provided by the TensorFlow autodiff functionality. Based on these derivatives, a stable minimization process of the residuals is performed. The full algorithm is automatically parallelized, i.e. some parts are vectorized.

8.2 Technical Implementation

As mentioned before, candidates consisting of two helices with four to six tracklets each, are fed to the syncfitter. Shared clusters are allowed. The syncfitter consists of two stages, which are described in parallel within this section. First, the syncfitter stage two is described in detail, as this one is more precise and the first stage is a less constrained version of the second one. A special emphasis is placed on regularization, which stabilizes the fit, and the pad tilting. The energy loss correction is explained in passing. The algorithm provides a quadratic loss function including regularization that is minimized iteratively. A summary of the different steps is shown as a sketch in Fig. 8.2.

8.2.1 Stage 2

Starting at the proposed conversion vertex (seed or result of the previous iteration), each track is transported from layer to layer in the xy plane after being transported to the first layer after the conversion vertex. This happens in the local coordinate system of the respective TRD sector. The distance between the layers Δx_n is known exactly. The transport is given by

$$\sin(\phi_{n+1}) = C_n \Delta x_n + \sin(\phi_n) \quad (8.1a)$$

$$y_{n+1} = y_n + \cos(\phi_n)/C_n - \cos(\phi_{n+1})/C_n \quad (8.1b)$$

$$s_{n+1} = (\phi_{n+1} - \phi_n)/C_n \quad (8.1c)$$

where y_n is the coordinate in the plane of the TRD layers. ϕ_n is the local angle of the momentum in the xy plane. C_n is the local curvature that might change during the propagation process due to energy loss. With

$$\cos(\phi_n) - \cos(\phi_{n+1}) = \frac{\cos^2(\phi_n) - \cos^2(\phi_{n+1})}{\cos(\phi_n) + \cos(\phi_{n+1})} \quad (8.2)$$

Input:	Stage 1:	Stage 2:
<ul style="list-style-type: none"> • use reasonable candidates • curvature from helices • vertex: mean y on first measured layer 	<ul style="list-style-type: none"> • propagate simultaneously to all layers • use residuals in xy plane • use regularization • delivers candidate selection for stage 2 	<ul style="list-style-type: none"> • propagate from layer to layer • residuals in y plane • use regularization • includes pad tilting and bremsstrahlung

Figure 8.2: The syncfitter consists of two stages. Starting with single helix candidates as input, reasonable seeds for stage 1 are generated. Stage 1 is then necessary to deliver stable seeds for stage 2, to avoid computational issues. Thereby, some corrections for higher precision are neglected and bad candidates are removed afterwards.

Eq. 8.1a and Eq. 8.1b are combined to

$$y_{n+1} = y_n + \frac{C_n \Delta x_n^2 + 2\Delta x_n \sin(\phi_n)}{\cos(\phi_n) + \cos(\phi_{n+1})}. \quad (8.3)$$

This equation is numerical more stable due to the eliminated $1/C$ dependence. To avoid singularities also for the arc length s_{n+1} a Taylor series up to order C^2 was calculated via SymPy [98] and implemented. In this context it might be noted that the circle center coordinates are implicitly changed by the choice of the parametrization in the case energy loss occurs. In this parametrization only the momentum is modified. Note that if $|C_n \Delta x_n + \sin(\phi_n)| > 1$ the arcus sinus is undefined. To avoid this, regularization terms are included. After each transport, the curvature is increased by a factor of $e^{f\Delta x}$ to correct for the mean energy loss by bremsstrahlung.

The results are the predictions y_n in each layer n . With these predictions, the residuals in y are calculated. They are given by

$$\text{res}_{y,n} = y_n - y_n^{\text{meas}}. \quad (8.4)$$

If there is no signal in a certain layer (given by dummy values of $x = -999$ and $y = -999$), the residual $\text{res}_{y,n}$ is multiplied by a weight of $w = 0$. This conserves the

input shape and with that also the parallelization capability of the implementation using TensorFlow.

With these residuals, the χ_y^2 is calculated. On the z axis $\tan \lambda$ and the z -coordinate of the vertex z_0 are calculated by using

$$z(s) = z_0 + s \tan \lambda \quad (8.5)$$

where s is the arc length in the xy plane. Here, the primary vertex assumption is used to calculate z_0 according to

$$z_0 = \tan \lambda \sqrt{x_0^2 + y_0^2} + z_0^{\text{prim}}. \quad (8.6)$$

z_0^{prim} is the z coordinate of the primary vertex. Hence, the residuals in z are given by

$$\text{res}_{z,n} = w_n \cdot (z_n - z_n^{\text{meas}}) \quad (8.7)$$

where $z_n = z(s_n)$ with the arc length s_n is calculated at each layer n in the xy plane.

By minimizing the full χ^2 given by

$$\chi^2 = \chi_{xy}^2 + \chi_z^2 + \chi_{\text{reg}}^2, \quad (8.8)$$

including also the regularization χ_{reg}^2 , the minimization takes place for the global χ^2 and the y and z residuals are optimized due to the chain rule applied by TensorFlow. For further comparisons the χ^2 is corrected by the weights w_n .

The minimization process is stabilized by the autodiff functionality of TensorFlow. The autodiff mechanism is capable of propagating the derivatives through the complete transport chain. Thus, this functionality enables the implementation of the full concept of the syncfit.

8.2.2 Stage 1

In case syncfitter stage 2 is seeded with bad starting parameters, issues could appear due to values in undefined regions. Syncfitter stage 1 solves this issue and has the functionality to produce more precise seeds and fit parameters from two helix fits as input for stage 2. Stage 2 then optimizes the fit by the mentioned detector effects. Thus, energy loss and pad tilting is neglected in stage 1.

There is only one global transport from the vertex to all layers to increase the parallelizability. The neglected energy loss correction enables this transport mechanism. To avoid the undefined arguments of the arcsin due to more fakes and worse

seeds, the residuals in the xy plane and not in the y plane are calculated. These residuals are given by the shortest distance between the circle and measurement as already used for helix fits in the tracking. This has the advantage that circles that do not reach the last layer still have a residuum.

The following seeds are calculated from the two single helices as input to stage 1: The curvatures are taken from the helix fits. The seeding x coordinate of the vertex is the closest measured x value to the primary vertex, and the y coordinate is the mean of the two closest measured y values of the two tracks.

8.2.3 Regularization

There are some fake photons or bad seeds in the preselection that lead to exceptions in the syncfitter (not invertible error, divergences). If the exception handling is done individually, the parallelization is broken. Therefore, a regularization loss is added to the full χ^2 . These extra constraints help the fit to converge (faster) and to avoid exceptions. The full loss of these fakes takes high values which allow one to identify them and increases the classification properties of the syncfitter.

A continuous cut-off of the following constraints is applied using the Relu function, which is commonly applied in neural networks as activation function. It is a linear function for values larger than zero and elsewhere zero. The first of these regularization constraints force the conversion vertex to be more than 60 cm away from the primary vertex. The radii of the circles are set to more than 50 cm, corresponding to a transverse momentum of at least 75 MeV/ c . $|\sin(\phi)|$ must be smaller than unity such that the track reaches all layers. Finally, the full χ^2 is minimized with respect to the parameters.

8.2.4 Pad Tilting

The pads in the TRD are tilted by the tilt angle α_{tilt} by two degrees with alternating sign from layer to layer. Thus, the residuals need to be modified for optimal fitting. The residuals are corrected by a rotation

$$y_{\text{res,tilt}} = \cos(\alpha_{\text{tilt}})y_{\text{res}} - \sin(\pm\alpha_{\text{tilt}})z_{\text{res}} \quad (8.9a)$$

$$z_{\text{res,tilt}} = \sin(\pm\alpha_{\text{tilt}})y_{\text{res}} + \cos(\alpha_{\text{tilt}})z_{\text{res}}. \quad (8.9b)$$

The sign within the sine is alternating from layer to layer.

8.2.5 Multiple Scattering

As mentioned in Ch. 7, it is not possible to resolve multiple scattering in the ALICE TRD due to the limited z resolution. Thus, a multiple scattering correction is not included in the syncfit. But in general, an extra loss term that exploits the angle could be included. Based on [99] an extra loss of the form

$$\chi_{\text{MS}}^2 = \sum_i \left(\frac{\Theta_i^2}{\sigma_{\Theta,i}^2} + \frac{\Phi_i^2}{\sigma_{\Phi,i}^2} \right) \quad (8.10)$$

can be added. Here, Φ and Θ are the azimuthal and polar multiple scattering angles and i is the index of the clusters/tracklets. The derivatives of the angles would again be automatically propagated via autodiff. This extra term was tested. It adds complexity, but had no effect on the result.

It should be mentioned, that, in the case that multiple scattering dominates the global momentum and local direction, parameters of a helix fit can be determined by only using the multiple scattering uncertainties. This is done in [99].

8.3 Syncfitter versus KFParticle

In this section, the difference between the syncfit method and the KFParticle package [86][87] is described qualitatively. The KFParticle package is used as the standard method to include the photon conversion topology to the fits of the daughter tracks in ALICE. The main difference is that the KFParticle package is based on tracks, i.e. not on the underlying tracklets directly, including their covariance matrices. In contrast, the syncfitter operates on the full sample of individual tracklets. Thus, the correlation between the tracklets are conserved as input to the calculation.

The KFParticle package is based on simple matrix operations using the Kalman filter technique. Thus, less complex computational operations need to be performed compared to the syncfit approach. However, due to modern programming methods with TensorFlow, the syncfit is reasonable fast for the given amount of data and stable. Thereby, a more complex model in the syncfitter exploits the underlying physics better and reflects the full information of the tracklets.

An important difference, which leads to more precise results, is the handling of the topology. Within the syncfit the number of free parameters is reduced to only five, as described before. Here, the topology is not considered as an external constraint, but as an intrinsic part of the model. In the case of the KFParticle package, the fit results from the helices of the individual tracks are used as input including their covariance matrices and used as measurements for the photon parameters.

Here, the helices are the measurements. The topological constraints do not lead to a reduction of the number of free parameters, but they are handled as extra (pseudo) measurements. The Kalman filter is based on Gaussian measurements. In contrast, the obvious complex correlations between the tracklets play a dominant role in the syncfit. This has the following implications: While shared clusters or outliers disturb the single fits heavily within the KFParticle, they play a smaller role for the syncfit. The estimation from the syncfitter of both angles λ and ϕ , of the photon kinematics, is excellent. For example, the azimuthal angle ϕ is restricted from both sides by the innermost clusters in the TRD and the long lever arm to the primary vertex. This directly causes a direction measurement of the electron and positron at the conversion vertex. This direction measurement stabilizes the transverse momentum measurement, especially also at higher momenta, where the small lever arm in the TRD implies a decreasing momentum resolution.

8.4 Performance

To obtain a first unbiased impression of the performance of the syncfitter before exploring the complex TRD data, a simple toy simulation is set up. In the toy simulation an electron–positron pair is drawn at a conversion vertex of (270 ± 2) cm without an opening angle. The transverse momenta of each electron/positron are independently drawn from a uniform distribution between 0.15 GeV/ c up to 2.15 GeV/ c , leading to a photon transverse momentum p_T up to and above 4 GeV/ c . For simplicity, the angles ϕ and λ are set to zero. Signals are generated in all six layers at the real positions of the TRD with a Gaussian resolution of 0.04 cm in y and $8/\sqrt{12}$ cm in z . This corresponds to the optimal y resolution with an inclination angle in the direction of the Lorentz angle. Thus, it is assumed that the z resolution is dominated by the pad size. Multiple scattering is included according to the mean material budget of a real TRD chamber. Bremsstrahlung and pad tilting are neglected in this first step to study the raw performance of the syncfit in nearly ideal circumstances.

In Fig. 8.3 the RMS of the photon parameters p_T , ϕ and $\tan \lambda$ given by the syncfitter are shown as a function of the transverse momentum in comparison to the performance of the KFParticle package. For the KFParticle package, the input tracks are fitted with the Kalman filter as implemented in AliRoot with the same seeding algorithm as implemented there. Thus, this process is fully decoupled from the new algorithms. The syncfitter is seeded with the standard helix fits as described before. The excellent resolution of the syncfitter in ϕ and $\tan \lambda$ can be clearly seen, and it outperforms the KFParticle package by far. Moreover, while with the KFParticle package the small lever arm for large momenta in the TRD is noticeable by a

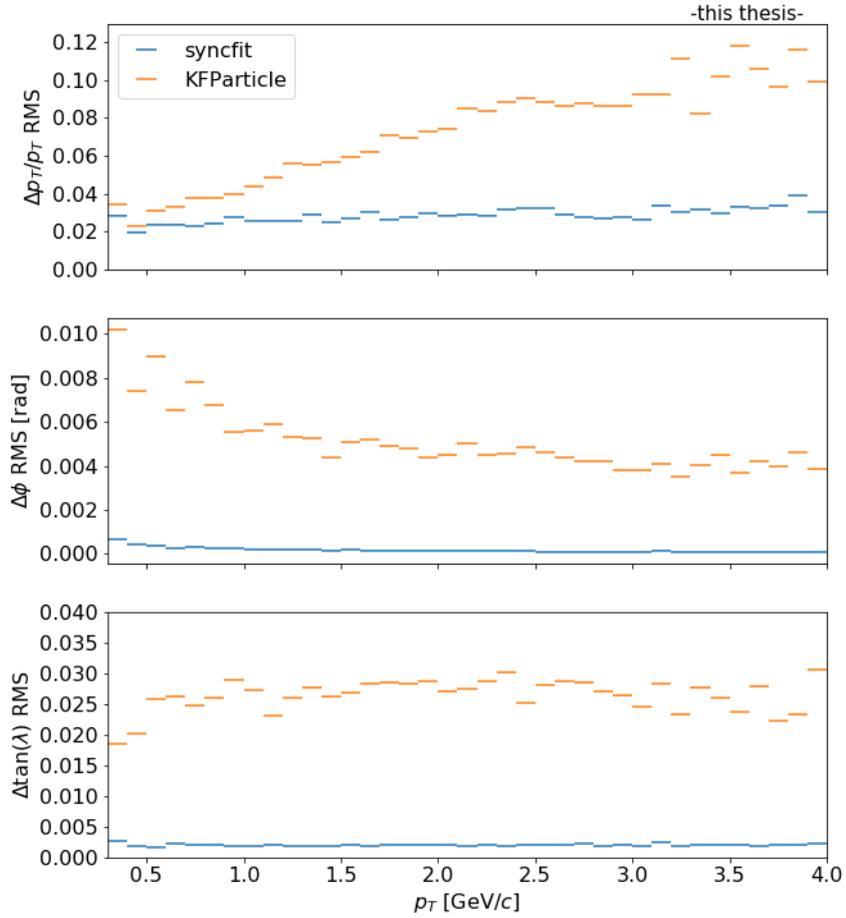


Figure 8.3: Syncfitter and KFPARTICLE performance for electron–positron pairs using a TRD toy simulation with only Gaussian errors, no pad tilting and no bremsstrahlung.

nearly linear decrease of the relative transverse momentum resolution from 2% at low p_T up to over 10% for 4 GeV/ c , the syncfitter stabilizes the resolution also at higher momenta as described in the previous section and the effect of the lever arm is mitigated.

8.5 Detector Effects

In this section, the two corrections for detector effects that are applied in the syncfitter stage two are discussed, i.e. bremsstrahlung and pad tilting. Again, toy simulations are used.

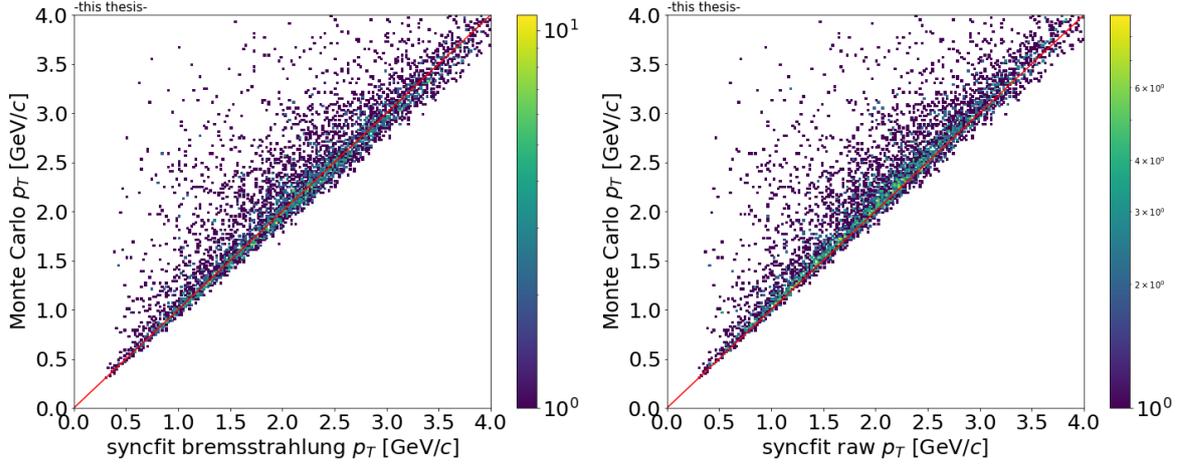


Figure 8.4: Scatter plot of the fitted transverse momentum using the syncfitter versus the true transverse momentum given by a toy simulation that includes multiple scattering and bremsstrahlung. Left: Syncfitter including the bremsstrahlung correction. Right: Syncfitter without the bremsstrahlung correction. The red line indicates the perfect fit to which the syncfit is centered after correcting for bremsstrahlung.

8.5.1 Bremsstrahlung

To illustrate the correction for bremsstrahlung, the previous toy simulation is extended by bremsstrahlung. In Fig. 8.4, the output data of the toy simulation is shown that is the result of fits with the syncfitter including this correction and without this correction. Shown are the true (drawn) transverse momenta versus the fitted transverse momenta in a two-dimensional histogram. It is apparent from the figure that the correction works well and centers the momenta to the right position. This also leads to a less biased χ^2 in the selection of the photon candidates. Nevertheless, a large tail remains towards lower transverse momenta.

8.5.2 Pad Tilting

Moreover, in Fig. 8.5 the effect of the pad tilting correction is shown. The toy simulation is extended by tilted pads with an alternating angle of $\pm 2^\circ$. Bremsstrahlung and multiple scattering are not considered in this test scenario. Now, $\tan \lambda$ is drawn between -0.5 and 0.5 , to study the impact on the z resolution. The distance between the fitted position in y and z and the real simulated true position of the track is calculated. An improvement of the resolution of a factor of two is visible for both dimensions. This was also confirmed using a loose fit, which is not shown. In the z

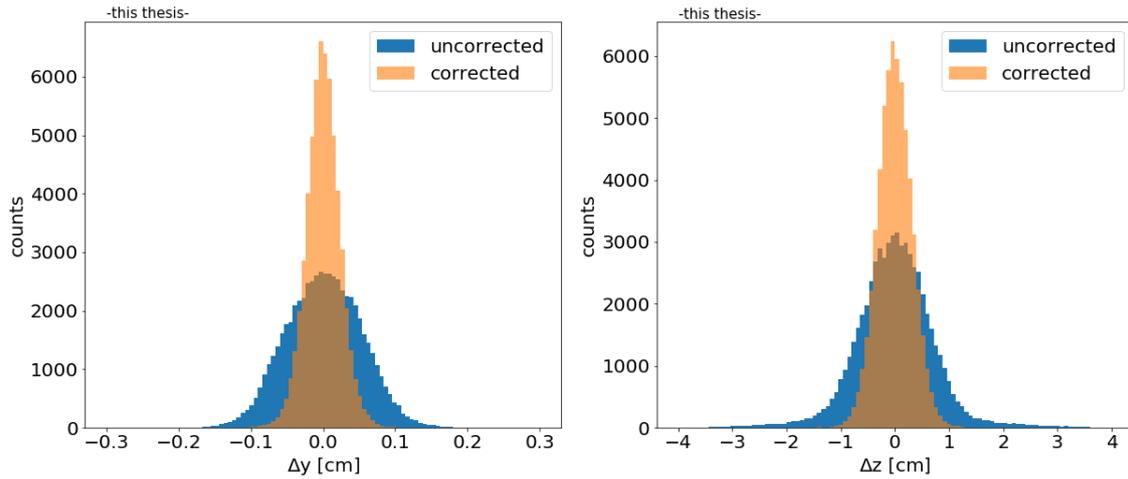


Figure 8.5: The distance between the fitted position and the true MC position left in y direction and right in z direction are shown for a toy simulation of the TRD including pad tilting for a syncfit that includes a pad tilt correction and one that does not include it. The distances are roughly halved compared to before for y and z .

direction, this corresponds to the improvement given by applying a pad tilting and in y this corresponds to removing the distortion given by the tilting. As the syncfitter uses the primary vertex, the $\tan \lambda$ resolution is excellent also without tilting, and thus the tilting is not that important when using this kind of fit. But this plot shows clearly that if tilting is used, a correction has a substantial impact on the resolution of the transverse momentum, as the y resolution is improved by a factor of two.

9 Measurement of Photons from Conversions in the TRD

In this chapter, the potential of photon reconstruction in the ALICE TRD is examined. In the previous chapters, it was shown that classical fits and the Kalman filter in combination with the KFParticle package have disadvantages compared to the syncfit. The workflow using the syncfitter is shown and afterwards the performance using the full ALICE Monte Carlo simulation as well as real data is evaluated.

The following approach is applied: First, the tracking algorithm as shown in Ch. 7 is applied. Then a loose preselection leads to photon candidates that are fed to the syncfitter. Based on quality criteria, photons are identified. Due to ambiguities, a post cleaning is necessary based on fit quality measures.

Finally, the performance is evaluated using the efficiency and purity, as well as the resolution of the track parameters as criteria. Moreover, the quality of the syncfitter is investigated using as data driven measure the π^0 invariant mass reconstruction in real data. Independent of using the π^0 invariant mass as quality measure, this is in general an important physics observable in ALICE.

9.1 Preselection

The first step is the preselection of the photon candidates using the stand-alone tracking helix fits and the associated tracklets. This selection is chosen to be very loose to achieve a high selection efficiency, but as tight as needed to significantly reduce runtime. The purity is increased afterwards by the classification capabilities of the syncfitter. Here, the quality measure is more precise than in the individual particle fits.

As the resolution of the pads in the transverse plane is much higher than the resolution of the pads in z direction, first a simple selection criterion in the xy plane is applied. It is based on the circle parameter estimates of the helices. The cuts are based on the ideal concept where the circles touch and there is a vanishing opening angle between the daughter pairs. To implement this, all unlike sign pairs of found tracks in each event are combined. First, the distance between the circles Δ_{circles} is

9 Measurement of Photons from Conversions in the TRD

calculated. This distance is defined as the absolute difference between the distance of the two centers and the sum of the two radii:

$$\Delta_{\text{circles}} = \left| \Delta_{\text{centers}} - \frac{1}{|C_{e^+}|} - \frac{1}{|C_{e^-}|} \right|. \quad (9.1)$$

C_{e^+} and C_{e^-} are the curvatures of the positron e^+ and the electron e^- candidates. At this hypothetical conversion vertex in the xy plane also the z -distance $\Delta_{\text{circles},z}$ between the helices is calculated. Additionally, a cut on the conversion radius at this point is applied. The conversion radius is required to lie within a certain interval, where it is expected to have good TRD reconstruction capabilities, and where a measurement in the TPC is not possible anymore. The conversion vertex should also be in a certain distance to the mean position of the associated tracklets. This rejects candidates for which the conversion vertex guess is for instance on the opposite side of the detector. A guess for $\tan \lambda$ is calculated from the hypothetical conversion vertex. It is compared to the $\tan \lambda_{e^+}$ and $\tan \lambda_{e^-}$ as a further criterion to ensure the right flight direction.

Some helix fits are disturbed by shared clusters and at large momenta the p_T resolution is not sufficient for the selection. Thus, besides the described selection, further criteria are applied. An additional selection includes pairs of helices with tracklets that are closer than 20 cm in the three-dimensional distance. At larger momenta sometimes an insufficient transverse momentum resolution leads to a switch of the sign of the curvature. To tackle this problem, for transverse momenta larger than 1 GeV/ c , the selection criteria are also applied with a switch of the sign.

The input to the syncfit is then given by two helices with opposite sign and between four and six tracklets each. Shared cluster are very common. The direction vector of the tracklets is neglected in the following as the resolution is low and the complexity of the fit algorithm would be increased. Nevertheless, it can make sense to investigate the effect of this additional information in the future.

9.2 Post Cleaning

After the syncfit is performed, different selection criteria for reasonable tracks are applied. Instead of a single cut on the full χ^2 that is minimized during the fitting process, three cuts on the constituents of this value are adjusted. The constituents are the χ^2 in y direction, the one in z direction and a regularization term. The use of multiple cuts has the advantage that the regularization term due to a large prefactor does not dominate the selection. Also a clearer selection is possible. The raw output distributions from the syncfit of the different χ^2 values including the

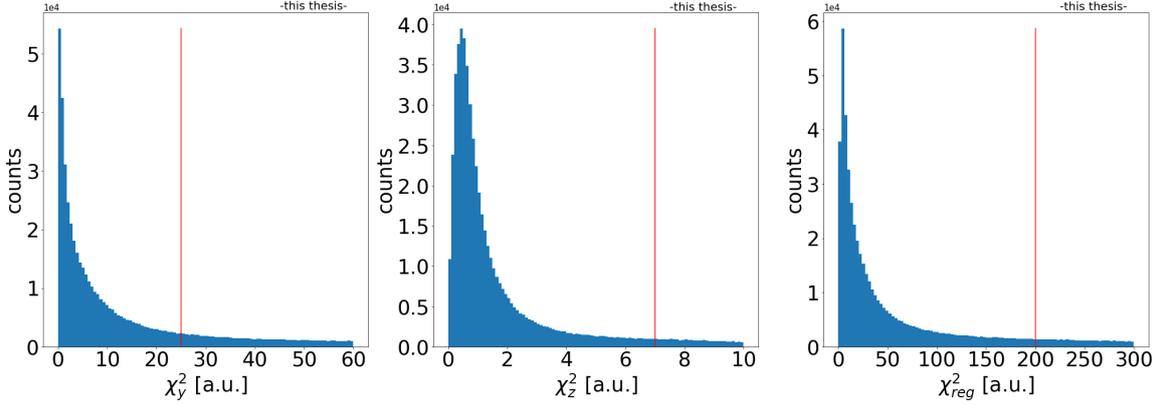


Figure 9.1: The χ^2 from the syncfit consists of a part in the local y direction and z direction, as well as a regularization term. For the final photon selection, single cuts on the different parts are applied. The distributions including the cut values (red lines) are shown.

cuts are shown in Fig. 9.1. The choice of the cut values are adjusted by hand to find a good trade-off between efficiency and purity. The purity and efficiency are shown later in this chapter.

In the next step, further quality criteria are applied. On the one hand, the conversion radius is required to be between 200 cm and 320 cm, such that a reasonable measurable signal is within the TRD and within the range where the syncfitter works best. On the other hand, the daughter tracks are required to have at least three unique tracklets assigned beside the shared ones, to ensure enough information for a reasonable helix. Furthermore, these unique tracklets are required to have a distance of less than 1.8 cm in the xy plane to the output helix from the syncfitter and the cosine of the angle between the tracklet direction and the fitted helix should be larger than 0.97.

By allowing shared tracklets within the track finding, it sometimes happens that two electron/positron candidates or more are combined with another photon daughter candidate. To avoid these cases and other similar ambiguities, shared tracklets between two photon candidates are not allowed. For these cases the candidate with the better χ_y^2 is chosen.

9.3 Performance using Full Monte Carlo Simulation

Within this section, the identified photons are analyzed and compared to the MC truth information. It should be noted, that as before for the MC information the tracklets are not assigned to the tracks, but only the vertex position and the momenta at the vertex, as well as the production vertex, are known. Thus, a matching procedure is described not only for the matching between the MC tracks and the identified tracks, but also between the MC tracks and the measured tracklets. This procedure differs from the single track matching described before. Now a full photon MC matching is implemented.

9.3.1 Conversion Vertices

An important measure of photon conversions are the conversion vertices, as they give an insight into the material budget distribution. Furthermore, it is a check for the proper functioning of the algorithm by a comparison of the MC truth vertices with reconstructed MC vertices. Many conversions should be observed in areas with a lot of material.

In Fig. 9.2 the vertices are shown in the xy plane and the z - R_{conv} plane. Shown are only conversion vertices with a reconstructable vertex in the given R_{conv} and z range. In the xy plane, it is obvious that photons are mainly reconstructed in a short distance before and within the first two layers of TRD. Photons converting further out cannot generate enough TRD tracklets to be reconstructed with sufficient quality. While in the MC data most conversion vertices are visible between the eighteen supermodules, these cannot be reconstructed with the given algorithm as they do not generate enough signals within the TRD. In the z - R_{conv} plane the main sources of material for conversion are clearly visible. Seven clusters from support structures of which six are symmetric around $z = 0$ can be seen. These can be assigned to certain positions also for the MC truth information. The one closest to the primary vertex is at $z = 0$ at the outer end of the central high voltage electrode of the TPC. Two clusters are generated at the most outer distance of the end caps of the TPC. Four clusters appear between the stacks of the TRD support structures.

Four different components can also be seen in the R_{conv} histograms in the lowest row of Fig. 9.2. Tests have shown that a Gaussian smearing of 2 cm brings the MC truth and measured histogram in the relevant range together and the double peak structures are merged to single peaks. The peaks at the outer field cage can be reconstructed very well.

9.3 Performance using Full Monte Carlo Simulation

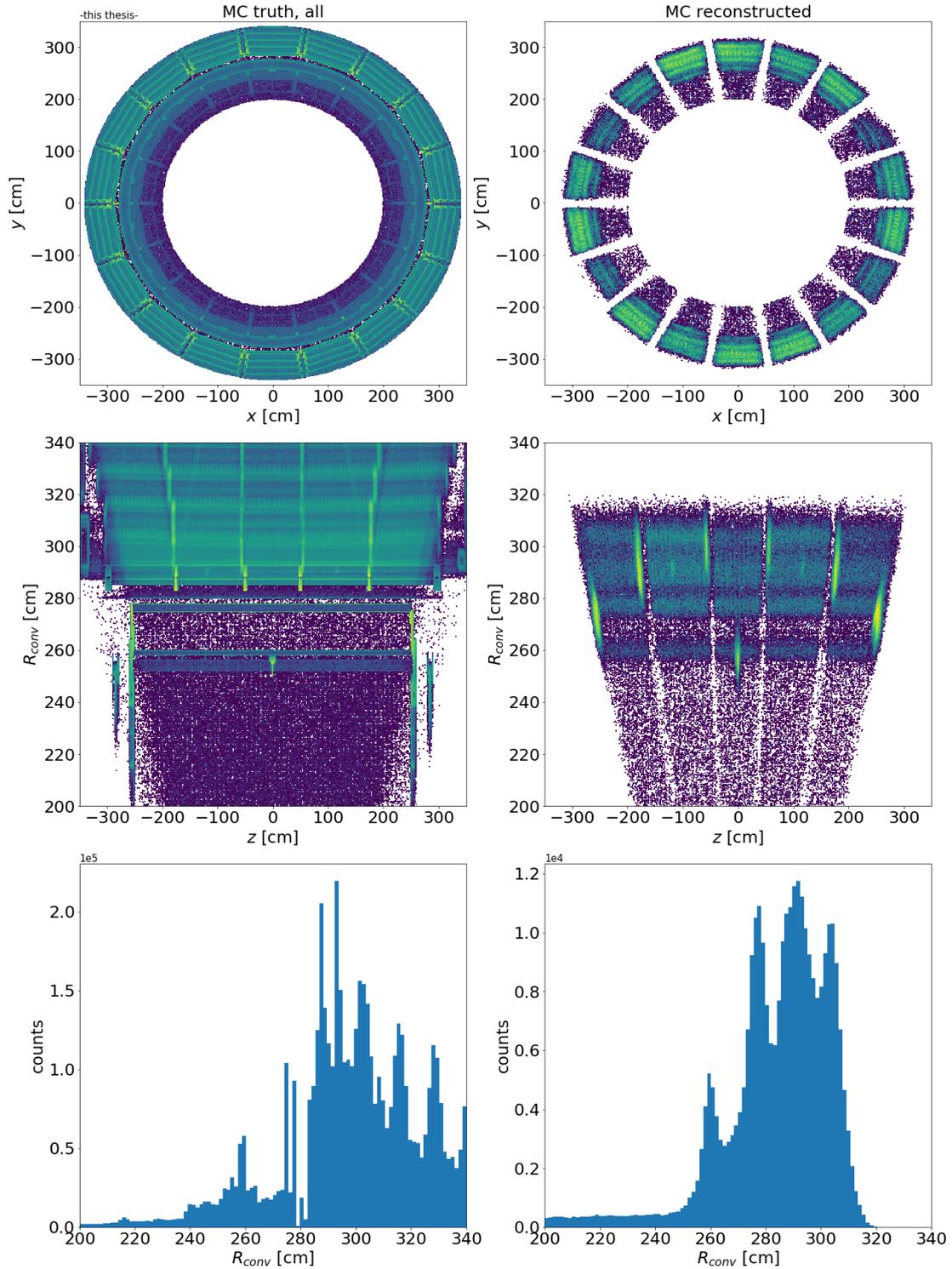


Figure 9.2: In the upper two rows the conversion vertices are shown, left the truth information for all MC photons in the given window and right the reconstructed vertices. In the lowest row, histograms with the distribution of the conversion radii are shown.

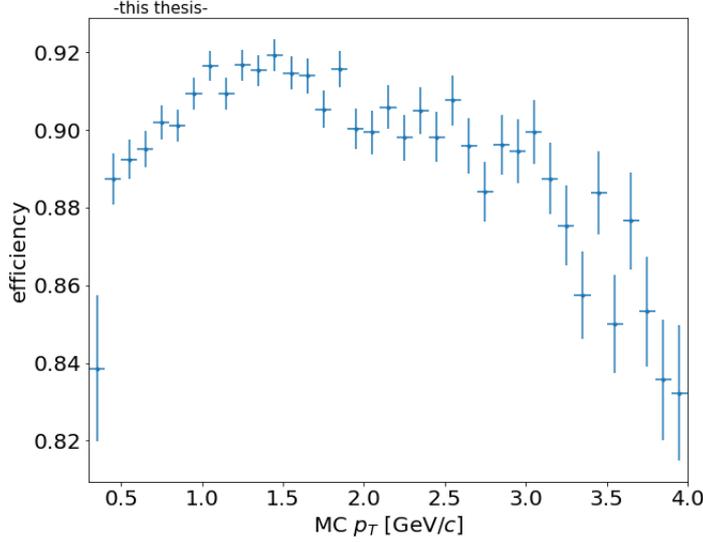


Figure 9.3: The reconstructable efficiency as defined in the text. The efficiency deteriorates at low momenta due to multiple scattering and at high momenta due to the spatial resolution.

9.3.2 Efficiency of reconstructable candidates

As mentioned before, only the momenta of the MC information at the true conversion vertex are available. Thus, a matching procedure also for the MC tracks and the measured tracklets has to be performed to obtain an efficiency measure. The aim is to obtain the tracks that are likely reconstructable and from this, the relative number of reconstructed tracks as a function of transverse momentum.

The matching procedure of the MC tracks and the tracklets for the photon conversion daughters is as follows: The shortest distance between the tracks and the tracklets in the xy plane is required to be less than 3 cm and in z less than 7.5 cm. These values are larger than the detector resolution to consider for detector effects. The cosine of the angle between the direction of the track and the tracklet direction at the tracklet position should be larger than 0.98 corresponding to a maximum of roughly 11.5° . Every track is required to have at least four unique tracklets and a maximum of two missing tracklets. The conversion vertex is required to lie in the range $250 \text{ cm} \leq R_{\text{conv}} \leq 320 \text{ cm}$. Furthermore, the radial position of the conversion point needs to be smaller than the radial position of the assigned tracklets.

The next step is the matching of the tracklets that are assigned to MC tracks with

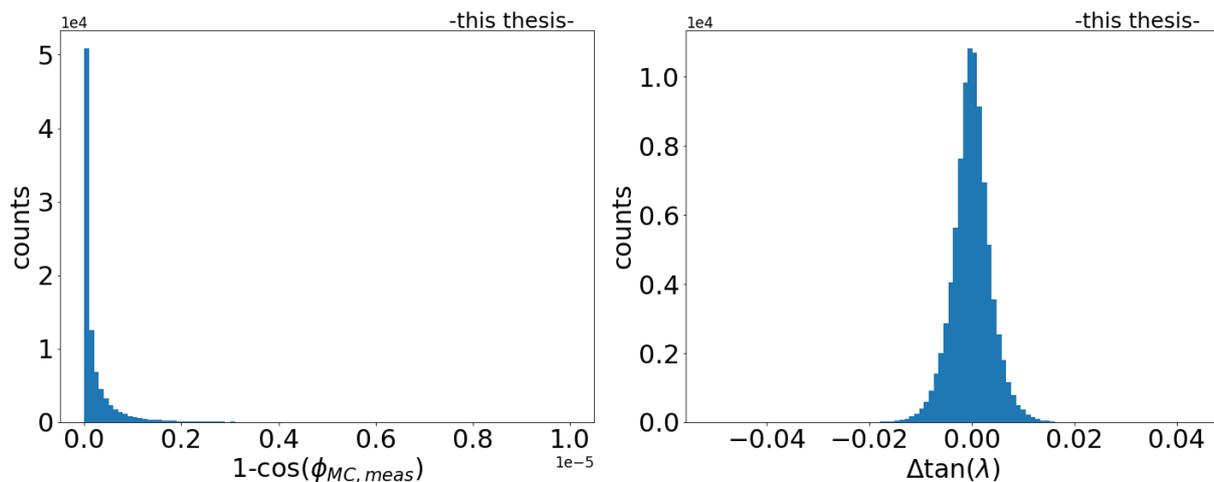


Figure 9.4: The distribution of the angles between the MC true and the measured information is very narrow in both directions of the coordinate system, but especially in ϕ direction.

the tracklets found by the reconstruction algorithm. Tracks that have at least two tracklets for each daughter track in common with the syncfitter tracks are counted as matched.

The resulting efficiency plot is shown in Fig 9.3 as function of the transverse momentum. Over the whole range an efficiency of roughly 90% can be achieved. It increases at low momenta up to higher momenta, where multiple scattering plays a role and decreases for large momenta, where the tracking resolution deteriorates. Still, fake matches in the MC matching procedure with the tracklets cannot be excluded, and the efficiency might be better than shown.

9.3.3 Purity

While examining the data, it was found that many tracks that are not matched via the given matching procedure for the efficiency probably are still photons. As mentioned before, the tracklets are not perfectly assigned to the MC tracks due to multiple scattering and bremsstrahlung. But as shown in this section, the purity is nevertheless good. For the estimation of the purity a different appropriate matching procedure is applied.

The toy MC simulation already showed that the angular resolution of the syncfit is excellent. For the so far used matched candidates, the resolution in ϕ and λ is shown in Fig. 9.4. Here, the cosine of the azimuth angle difference $\cos(\phi_{MC, meas})$ is

9 Measurement of Photons from Conversions in the TRD

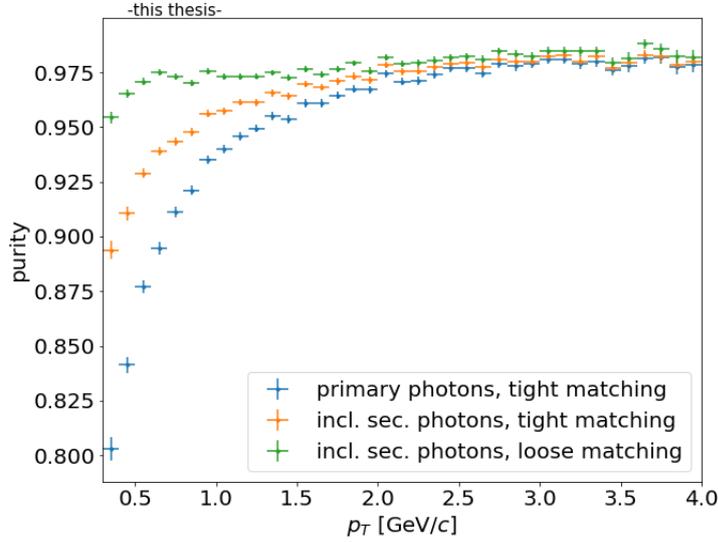


Figure 9.5: The purity for different matching criteria. A clear selection of primary photons compared to looser selections with also secondary photons and looser matching criteria show that secondary photons are the dominating source of contamination.

defined by the normalized scalar product of the conversion vertices in xy :

$$\cos(\phi_{\text{MC,meas}}) = \frac{x_{\text{conv,MC}}x_{\text{conv,meas}} + y_{\text{conv,MC}}y_{\text{conv,meas}}}{R_{\text{conv,MC}}R_{\text{conv,meas}}}. \quad (9.2)$$

For demonstration, $1 - \cos(\phi_{\text{MC,meas}})$ is plotted. The values in the distribution correspond to angles smaller than 0.1° . Also the difference between $\tan \lambda$ for measured and MC tracks is shown in this figure, which is very small as well. As in the MC data only a handful of conversion take place per event, a selection using these criteria is sufficient. Thus, for calculating the purity, the MC matching with the tracklets is neglected and only a vertex matching of the MC vertices with cuts on the angles given in Fig. 9.4 is performed.

In Fig. 9.5 the resulting purity is shown as a function of p_T for different scenarios. The first selection requires a production of the photon at the primary vertex and selection of $\cos(\phi_{\text{MC,meas}}) > 0.99999$ as well as $\Delta \tan \lambda < 0.03$. For these primary photons, it can be seen that starting at 80% purity at momenta of 0.3 GeV/ c the purity rapidly increases up to 1.5 GeV/ c to above 95%. The same tight matching including secondary particles already shows a large increase for the whole range to nearly at least 90%. If now also the selection cut is loosened, over the whole

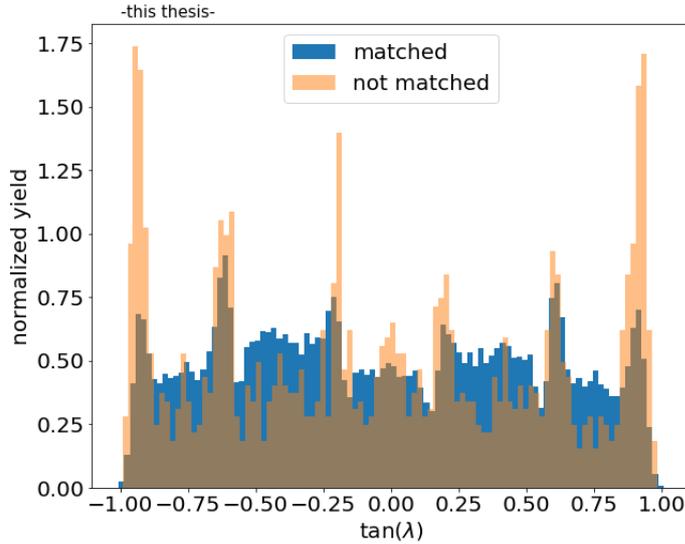


Figure 9.6: The $\tan \lambda$ distribution of the vertices for matched tracks and unmatched tracks for transverse momenta smaller than $0.8 \text{ GeV}/c$. A comparison with Fig. 9.2 shows that the contamination comes still from those areas with a high material budget. Thus, also most of the leftover contamination tracks probably originate from secondary photons.

range more than 95% can be achieved. Many secondary photons appear due to bremsstrahlung. These cannot be distinguished physically from primary tracks if they point to the primary vertex within given detector resolution. These appear more often at low momenta as the low momenta bremsstrahlung contribution is more probable.

In Fig. 9.6 the normalized $\tan(\lambda)$ distribution of the unmatched MC candidates after applying the loose cuts and for all matched candidates for transverse momenta smaller than $0.8 \text{ GeV}/c$ is plotted. It becomes obvious that the remaining impurity also mostly originates from the areas with a high material budget as shown in Fig. 9.2, more than for the matched candidates. In these areas, the probability to generate bremsstrahlung is larger. Thus, the purity is even higher than calculated if also secondary photons are considered, which unfortunately cannot all be excluded.

9.3.4 Transverse Momentum Resolution

With the tight selection via the vertex matching, it is now possible to estimate the transverse momentum resolution for primary photons. In Fig. 9.7 on the left side the

9 Measurement of Photons from Conversions in the TRD

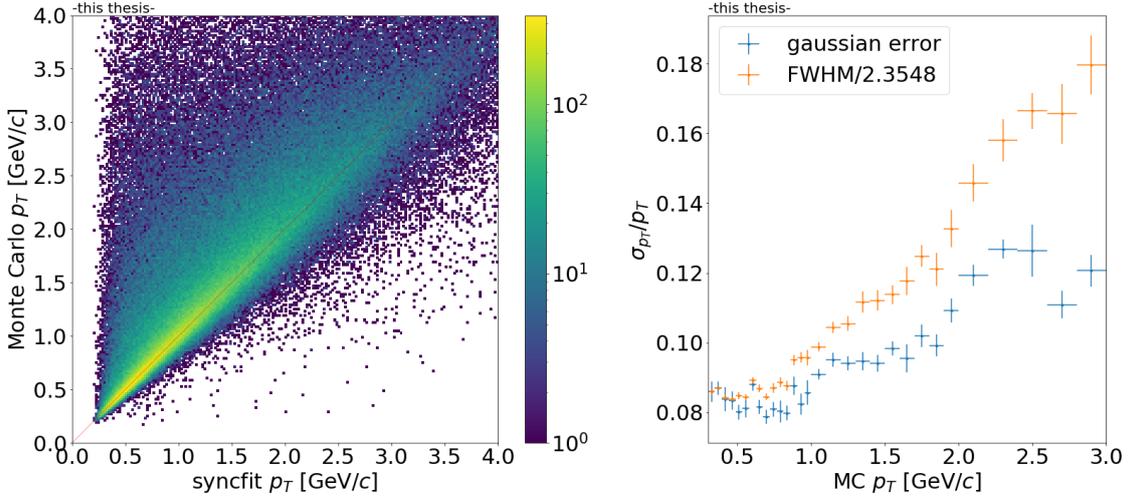


Figure 9.7: Left: The MC p_T versus the measured p_T . Even though the distribution is well centered, a tail from bremsstrahlung remains. Right: From different slices from the distribution left, the resolution is calculated. The crystal ball function is fitted to the different histograms. To account for the tail, not only the Gaussian sigma is shown, but also the FWHM/2.3548, to emphasize the effect from bremsstrahlung.

MC p_T is shown versus the reconstructed p_T from the syncfitter in a logarithmic two-dimensional histogram. It is well centered, but the large tail from bremsstrahlung remains towards smaller measured momenta, which is especially visible due to the logarithmic scale. A tail from outliers also appears at large momenta, as here the previously mentioned lever arm plays a role. In that case, only a very small segment of the full circle of the motion of the particle is subtended by the area of the TRD.

To obtain a quantitative description of the resolution, projected slices of the two-dimensional histogram are fitted using a crystal ball function as given in the python framework SciPy [100]. This function accounts for lossy processes, as given by bremsstrahlung. In particular, the tail of the bremsstrahlung can be described by this distribution. It has the form

$$f(x; \alpha, n, \bar{x}, \sigma) = N \cdot \begin{cases} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right) & \text{for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A \cdot \left(B - \frac{x-\bar{x}}{\sigma}\right)^{-n} & \text{for } \frac{x-\bar{x}}{\sigma} \leq -\alpha \end{cases} \quad (9.3)$$

9.4 Performance using real p–Pb data at $\sqrt{s_{NN}} = 5.02$ TeV

with

$$A = \left(\frac{n}{|\alpha|} \right)^n \cdot \exp \left(-\frac{|\alpha|^2}{2} \right) \quad (9.4a)$$

$$B = \frac{n}{|\alpha|} - |\alpha|, \quad (9.4b)$$

where N is a normalization factor. The distribution consists of a Gaussian part and a tail towards low values.

The corresponding fits of the transverse momentum distribution can be found in appendix A. In Fig. 9.7 in the right plot, the resulting standard deviation from the Gaussian part, as well as the full width half maximum (FWHM) divided by $2\sqrt{2\ln 2} \approx 2.3548$ are plotted for the different fitted transverse momentum intervals. The FWHM is used to account for the tail and see the corresponding effect. The chosen factor is the ratio between the FWHM and the standard deviation of a normal distribution. The fits are only shown up to momenta of 3 GeV/ c , as for higher momenta the tail towards large momenta is not described by the fit function anymore. The tail from bremsstrahlung becomes more significant for large momenta. For small momenta, tracks with a lot of bremsstrahlung are more probable during track reconstruction. Up to more than 1 GeV/ c , the resolution is better than 10% and quite stable, even though the segment of the circle becomes very small.

9.4 Performance using real p–Pb data at $\sqrt{s_{NN}} = 5.02$ TeV

The next step is to verify the functionality of the syncfitter using real data. Roughly ten million p–Pb events at $\sqrt{s_{NN}} = 5.02$ TeV from the 2016 run are analyzed. The same photon selection cuts as for the MC data are applied. About 1.6 million photons are found and reconstructed. The corresponding transverse momentum distribution is shown in Fig. 9.8 with a log scale on the y -axis. The rise of the distribution starts at 200 MeV/ c and a significant number of photons can be found starting at 250 MeV/ c .

9 Measurement of Photons from Conversions in the TRD

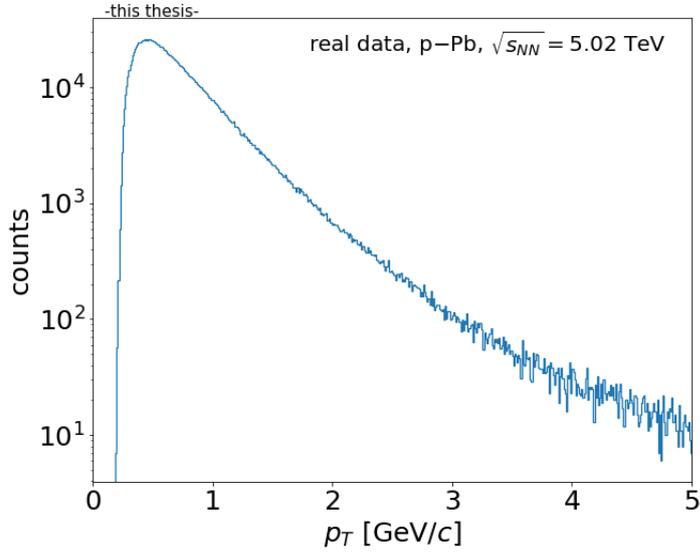


Figure 9.8: The transverse momentum distribution for all identified photons with a log scale on the y -axis. The lower boundary to measure photons from conversions in the TRD are transverse momenta of 200 MeV/ c .

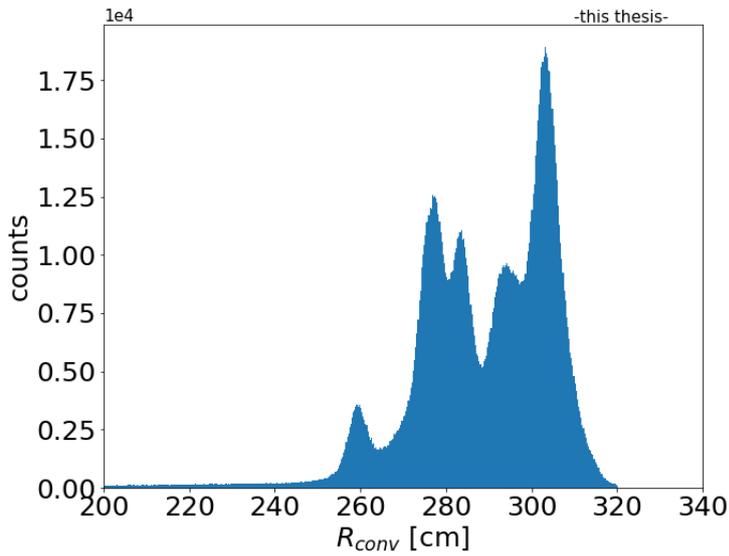


Figure 9.10: The distribution of the conversion radii for real measured photons. The most outer peak is more pronounced as there are more tracks with lower transverse momenta.

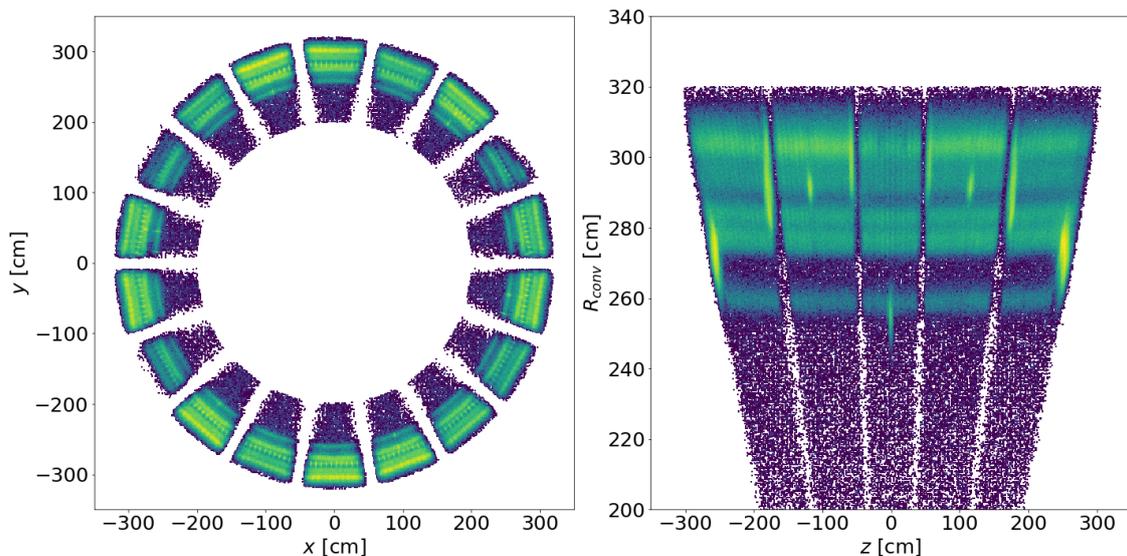


Figure 9.9: Photon conversion vertices for all identified photons left in the xy plane and right in the z - R_{conv} plane. The density is plotted in a logarithmic scale.

As for the MC data, in Fig. 9.9 the conversion vertices are also shown in logarithmic two-dimensional histograms in the xy plane as well as in the z - R_{conv} plane. Compared to the MC reconstructed vertices in Fig. 9.2 photons from the support structures are less pronounced. Especially in the TRD layers, a lot more conversions take place. By having a larger relative number of low p_T photons than in the MC simulation, later conversion are more probable to be measured. These effects can also be seen in the one-dimensional histogram of the conversion radii.

9.5 π^0 Reconstruction

With the reconstructed photons it is possible to reconstruct the neutral pion and eta meson. The π^0 decays to two photons with a branching ratio of $(98.823 \pm 0.024)\%$ and the η decays to two photons with a ratio of $(39.41 \pm 0.20)\%$. Energy and momentum conservation delivers the invariant mass via

$$M_{\gamma\gamma} = \sqrt{2E_{\gamma_1}E_{\gamma_2}(1 - \cos\theta_{12})} \quad (9.5)$$

with the energy E_{γ_1} and E_{γ_2} of each of the two photons and the angle θ_{12} between the momenta of both photons.

For the real data the two photon invariant mass spectrum is shown in Fig. 9.5 in the upper plot. For all two photon combinations the invariant mass is calculated with Eq. 9.5 and shown as a histogram. The well known mass of the neutral pion as well as the mass of the eta meson are indicated by red lines. The π^0 peak is clearly centered at the right position, but a tail from bremsstrahlung remains. This provides a further cross check that shows that the photon reconstruction works properly. There is not enough statistics to see the η meson. The lower plot of Fig. 9.5 shows the invariant mass vs. the transverse momentum. A lower transverse momentum limit of roughly $0.7 \text{ MeV}/c$ can be achieved for the reconstructed neutral pions.

Furthermore, the photons reconstructed with only the TRD can now be combined with photons reconstructed with the TPC and thus increase the full π^0 reconstruction efficiency. To do so, the TPC photons are reconstructed based on the TPC tracks with a simple photon finder. This simple photon finder uses the preselection criteria with adjusted cuts that were used as input for the syncfit before. This selection is based on the distance between two helices with opposite charge sign. This selection does not deliver a very high purity, but it is efficient and thus the integration of the π^0 peak gives a reasonable estimate of the number of reconstructable particles. The invariant mass for two photons from the TPC is shown in the upper panel of Fig. 9.12. Due to the higher momentum resolution of the TPC the peak is of course narrower. Also, there is a higher photon efficiency such that the peak of the η meson can already be seen with the available 10 million events.

In the lower panel of Fig. 9.12 the combined invariant mass of the TPC and TRD photons is shown. As expected, the width of the peak is of a magnitude between TPC only and TRD only. But also with this combination a small eta peak is visible.

To obtain a quantitative estimate of the increase of the π^0 efficiency, the invariant mass peaks are fitted together with the background. The following function was

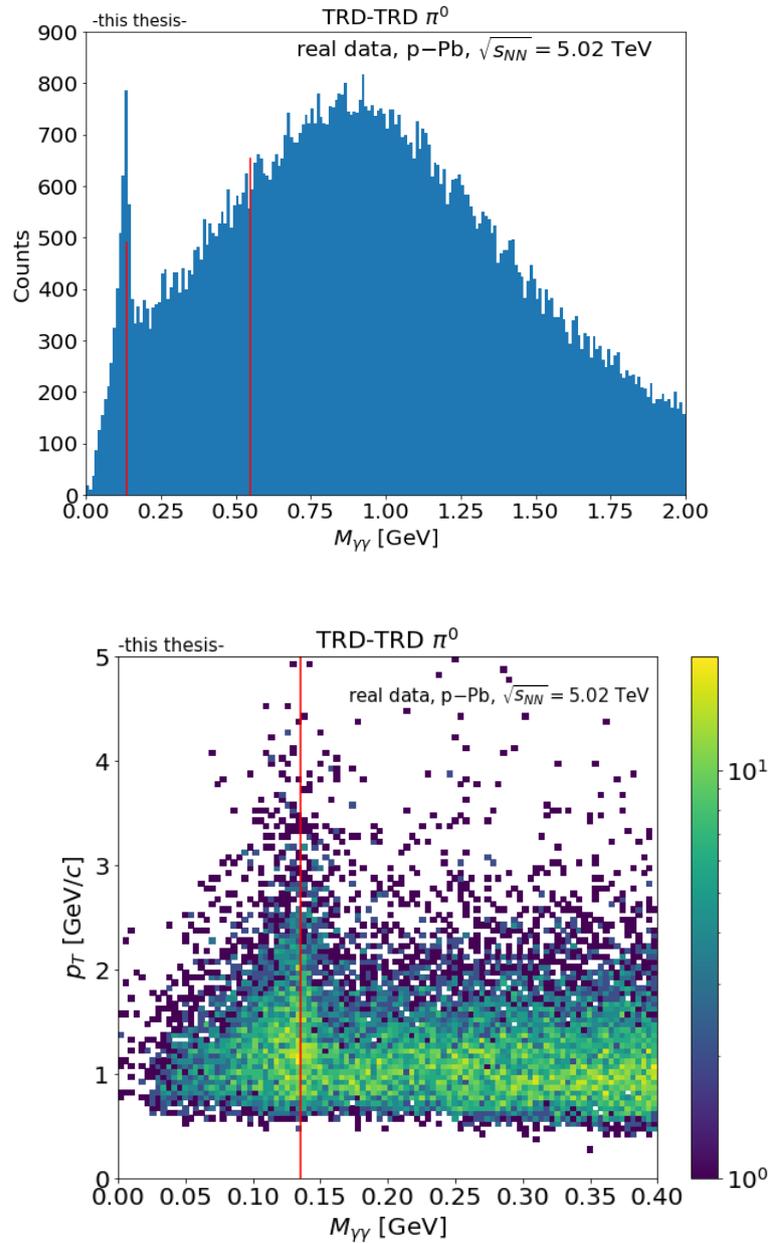


Figure 9.11: Top: The invariant mass spectrum of π^0 candidates from two photons reconstructed only with the TRD. The red lines indicate the values for the invariant mass of the π^0 as well as the η meson. A clear π^0 signal can be seen. Bottom: The invariant mass is shown vs. the transverse momentum. A lower limit of roughly 0.7 GeV/ c can be achieved.

9 Measurement of Photons from Conversions in the TRD

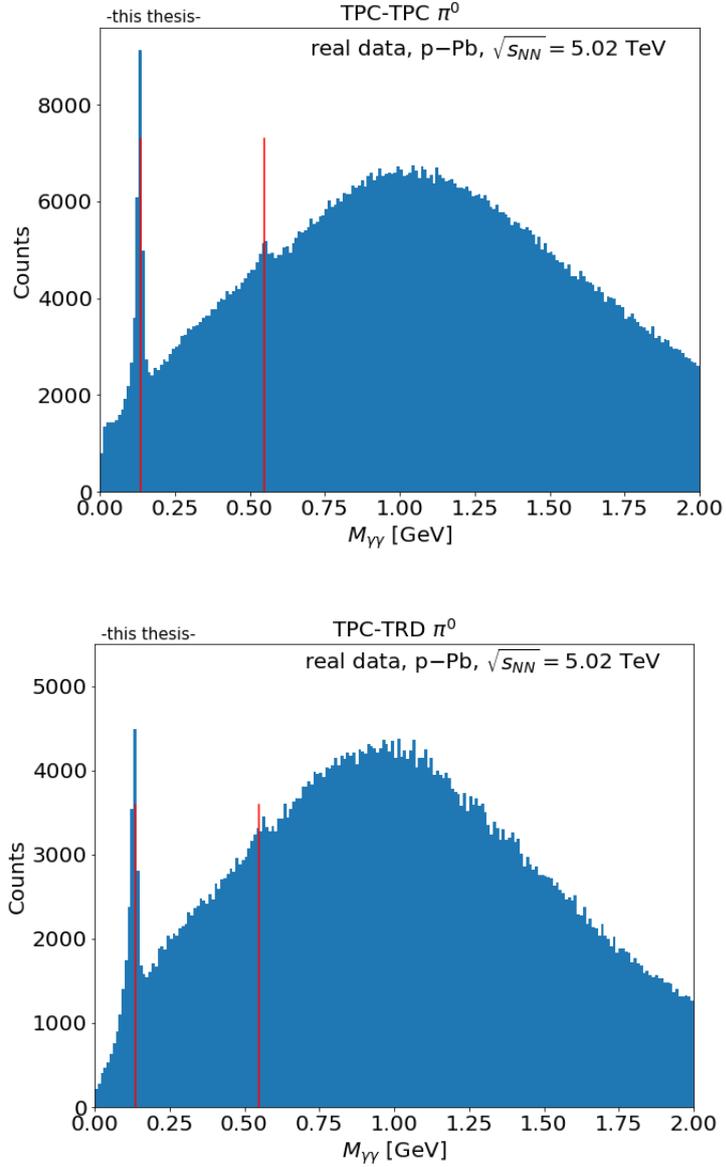


Figure 9.12: Top: The invariant mass for π^0 candidates from two photons reconstructed with the TPC. Bottom: The invariant mass for π^0 candidates with one photon reconstructed in the TRD and the other in the TPC. In both cases, a clear π^0 signal and a small η signal can be seen.

chosen, similar as in [101]

$$f(M_{\gamma\gamma}) = A \cdot (g(M_{\gamma\gamma}) + e(M_{\gamma\gamma})) + r(M_{\gamma\gamma}) \quad (9.6a)$$

$$g(M_{\gamma\gamma}) = \exp\left(-\frac{1}{2}\left(\frac{M_{\gamma\gamma} - \mu}{\sigma}\right)^2\right) \quad (9.6b)$$

$$e(M_{\gamma\gamma}) = \exp\left(\frac{M_{\gamma\gamma} - \mu}{\lambda}\right) \cdot (1 - g(M_{\gamma\gamma})) \cdot (\mu - M_{\gamma\gamma}) \quad (9.6c)$$

$$r(M_{\gamma\gamma}) = a + bM_{\gamma\gamma} + cM_{\gamma\gamma}^2. \quad (9.6d)$$

In total, there are seven free parameters ($A, \mu, \sigma, \lambda, a, b, c$). Thus, the background is described with a polynomial of third order and the signal with a Gaussian and an exponential function accounting for the tail from bremsstrahlung. The fits are shown in appendix B. Integrating each signal delivers 2239 TRD-TRD π^0 s, 16636 TPC-TPC π^0 s and 10315 TRD-TPC π^0 s. Thus, there is an increase from only TPC-TPC π^0 of about 75% when including both other combinations.

10 Identification of Inelastic Interactions of Light Antinuclei and the Sexaquark using Graph Neural Networks

Besides photon conversions also other interaction in the material in front of and within the TRD can be detected and analyzed. As in Ch. 2 presented, interesting candidates are the identification of inelastic interactions of light antinuclei as well as the identification of the exotic sexaquark state.

These interactions were analyzed with MC data and it was found that not enough tracklets of daughter tracks exist in most cases to use the stand-alone tracking approach. In recent years, graph neural network (GNN) gained attention as an alternative to the established tracking methods. An overview can be found in [13]. In this chapter, a first feasibility study for the identification of interactions with the TRD using GNNs for the two mentioned physics topics is presented.

10.1 Introduction to Graph Neural Networks

Especially for image processing tasks, neural networks were a great success in the last years. Convolutional neural networks played a crucial role in this context [6]. Detector signals can often be considered as three-dimensional images. Unfortunately, the corresponding detector data is very sparse. But the sparse data points can be put into a mathematical relation by formulating them as graphs. Neighborhood relations, i.e. metrics, can be exploited using a graph formalism.

To show why graph neural networks (GNNs) are a proper tool to operate as an alternative to the given tracking algorithm, graphs, neural networks and then finally graph neural networks are introduced within this section. First, graphs in general are introduced and it is shown why they are a mathematical model of raw data for particle tracking. Then, the basics of neural networks are introduced, before it is finally shown how they can operate on graph structured data.

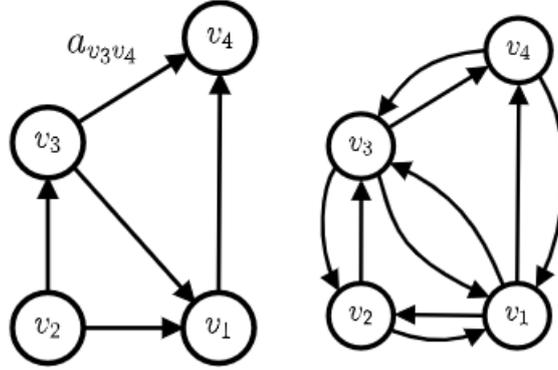


Figure 10.1: Examples of a directed graph (left) and an undirected graph (right).
Figure taken from [102].

10.1.1 Graphs

A graph is constructed from data points, the nodes, which are generally unordered. The relation of the nodes to the other nodes is given by edges. In our use cases each node is assigned a certain number of features and also the edges can be assigned features. Global graph attributes might also exist. Let N^v be the number of nodes, or also called vertices, and N^e be the number of edges. Not all vertices are connected with each other. The connection, i.e. the given edges, can be defined by the adjacency matrix. The dimension of this matrix is $N^v \times N^v$, and it is typically sparse as most nodes are not connected. The edges between the nodes are either directed or undirected. An illustration is shown in Fig. 10.1. Alternatives to the adjacency matrix are the incidence matrix with the dimension $N^v \times N^e$, where the connection between the nodes and the edges is described, and the coordinate list with the dimension $2 \times N^v$ as a list of all connected nodes.

Now, it is trivial that the tracks in a detector can be expressed as graphs. In this context, the detector clusters might be the nodes with the position coordinates as the features and the connections, given by the track of the particle, are the edges.

10.1.2 Neural Networks

In this section, modern neural networks are described in the context of classification and regression applications. Only supervised learning is considered in this work, i.e. labeled data is used to optimize the models.

In general, the neural networks are fed with a certain number of input features. These input features are propagated through several layers, that are connected by

weight matrices. Only neurons of consecutive layers are connected, which reduces the number of parameters and the complexity. The output is finally given by the output neurons, that are either a real number assigned to different classes or certain regression values. After a normalization, the output of the different classes can be interpreted as a probability. Thus, efficiencies can be defined. In case of a binary classification, where signal and background are defined, the signal efficiency can be determined as a function of the background efficiency. Such a plot is known as Receiver Operating Characteristic (ROC) curve.

The fitting of the neural network to the labeled data is performed by minimizing a certain loss function with respect to the weights of the neural network. This complex minimization procedure using gradient descent variants can nowadays be performed by the tools provided for differentiable programming as shown in Ch. 6. For classification tasks the cross entropy loss turned out to be a good choice, for regression one could for instance use the squared error loss, depending on the given circumstances (e.g. Gaussian errors).

In the following, different building blocks that are relevant within this work are introduced: dense layers, convolutional layers and the pooling layers. On simple tabular data, a wide variety of neural networks can be constructed with only dense layers. The other shown building block, the convolutional layers and pooling layer, led to the big success of large scale models for image classification, starting with the AlexNet [103]. This illustrates the basic functionality of modern neural networks.

The **dense layers** are the simplest building block used in modern neural networks. Here, each neuron of the previous layer is connected with each neuron from the subsequent layer. To describe non-linear behavior, the linearity is broken by applying a non-linear activation function on top of the given output. Let \vec{x}_l be the input layer l . The output of the dense layer $l + 1$ is provided by

$$\vec{x}_{l+1} = \phi(W_l \vec{x}_l + \vec{b}_l), \quad (10.1)$$

where W_l is the weight matrix which is optimized during the optimization procedure and \vec{b}_l are learnable bias weights. A typical choice for the non-linear activation function ϕ is the ReLU function ($\text{ReLU}(x) = \max(x, 0)$).

Convolutional layers led to a big breakthrough in the field of image classification. These perform especially well for data structured like images (e.g. $n \times m$ pixel with neighborhood relation), as they profit from weight sharing and thus the number of parameters is reduced drastically, and the network training is stabilized. Instead of connecting the previous layer with all neurons from the subsequent layer, certain windows are chosen, e.g. 3×3 , and on each of the given two-dimensional windows of

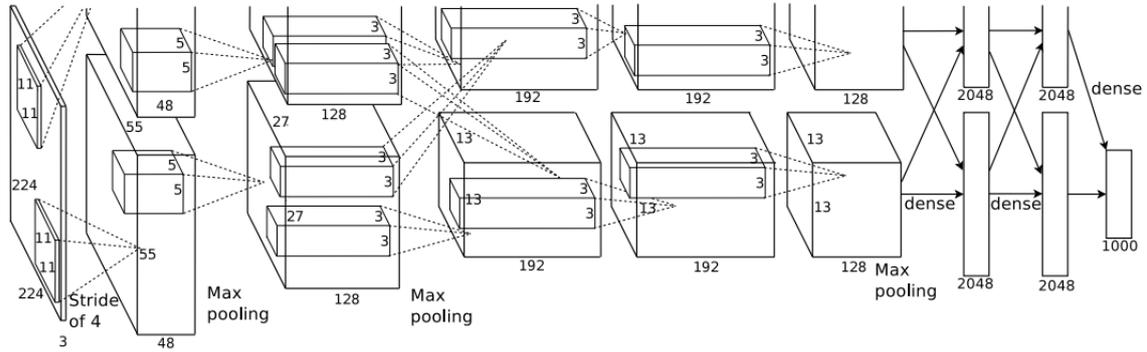


Figure 10.2: The AlexNet consists of convolutional layers, pooling layer and dense layers. Figure taken from [103].

the input the same weights are used. New pixels with a definable number of features are generated. Thus, local features are learned. Later, these are combined to global features by further operations.

A **pooling layer** reduces the total size of image-like data. The output values within certain window sizes are condensed to a single value by applying simple operations like using the max value or the average value.

In Fig. 10.2 AlexNet as used in [103] is shown as an example of a full convolutional network. The different building blocks are used consecutively and recurring.

10.1.3 Graph Neural Networks

The previously explained convolutional layers, which operate on image-like data, where the coordinates are given in a canonical way, can now be naturally generalized to graph structured data. As mentioned before, a more abstract relation between the feature vectors of the different nodes is given by the adjacency matrix. It should be mentioned that the input shape defined by the number of nodes might vary within the input data, i.e. the number of input nodes and edges varies from input to input.

In this section, two variants of graph neural networks that operate on the given data structure, i.e. with each node vector having the same size and no edge features, are given. The first one are the graph convolutional networks as proposed in [104] and the second one are the graph attention networks which have first been proposed in [105]. The most important operations of these networks, the graph convolution

and graph attention, preserve the number of nodes after each layer and result in each layer being a graph with a new number of features. Every layer corresponds to a message passing to its neighbors.

The graph convolution is defined as [104]:

$$H^{(l+1)} = \sigma \left(\tilde{D}^{\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)} \right). \quad (10.2)$$

Here, the adjacency matrix A including self-connections is given by \tilde{A} . The degree matrix of the corrected adjacency matrix is given by $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$. The weights of the given layer are called $W^{(l)}$, and $H^{(l)}$ is the matrix for activations in layer l , for instance the input nodes are given by $H^{(0)}$. The activation function σ might again be the ReLU function. The degree matrix \tilde{D} normalizes the input. It is split to ensure the symmetry of the normalization.

For the graph attention layer [105] the first step is to calculate the attention coefficients e_{ij} from the features \vec{h} for each node i with its neighbors j

$$e_{ij} = a \left(W \vec{h}_i, W \vec{h}_j \right), \quad (10.3)$$

where W is a weight matrix and a is a single-layer feedforward network. The network a is parametrized by a learnable weight vector \vec{a} with a LeakyReLU ($\max(x, \beta x)$, with β the slope parameter) such that the output is a real number:

$$e_{ij} = \text{LeakyReLU} \left(\vec{a}^T \left[W \vec{h}_i || W \vec{h}_j \right] \right) \quad (10.4)$$

The symbol $||$ represents a concatenation. The attention coefficients are then normalized via a softmax function for all neighbors k of each node:

$$\alpha_{ij} = \text{softmax}_j(e_{ij}) = \frac{\exp(e_{ij})}{\sum_k \exp(e_{ik})} \quad (10.5)$$

The final node features are then calculated via

$$\vec{h}_i^{(l+1)} = \sigma \left(\sum \alpha_{ij} W \vec{h}_j^l \right) \quad (10.6)$$

by applying an activation function σ . An extension of this mechanism is the application of several attention coefficients in parallel, called multi-head attention. An illustration of the algorithm is shown in Fig. 10.3.

On the basis of the shown algorithm, different kinds of regression and classification tasks can be performed: node classification, edge classification and graph classification. For the graph classification, that is of interest in this work, a global pooling needs to take place after several steps of message passing.

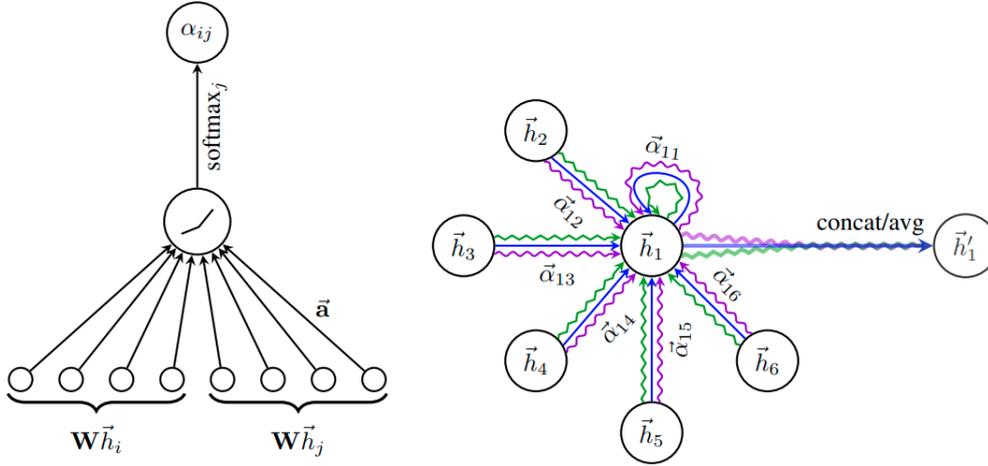


Figure 10.3: Left: First, the attention coefficients are calculated. Right: Then the message passing is performed. In this case, multi-head attention with three heads is shown. Figure taken from [105].

10.2 Identification of Inelastic Interactions of Light Antinuclei

In this section, the results from the bachelor's thesis from Max Hammermann 'Identification of inelastic interactions of light antinuclei in the Transition Radiation Detector in ALICE using Graph Neural Networks' [106] are summarized. This thesis provides a first example of the identification of nuclear interactions in the TRD with GNNs. A brief physics motivation is given in Sec. 2.5.

As the tracks that are generated by the nuclear interaction are not well-structured and many fragments exist with less than four tracklets, it is not feasible to use a straightforward track finder as for the photon conversion analysis. In this Monte Carlo study, nuclear interactions of charged antinuclei are analyzed. These provide a higher cross section than the corresponding positively charged nuclei and are especially interesting for dark matter searches. Due to the nuclei being charged, TPC tracks might exist. The TPC tracks are extrapolated into the TRD and along the tracks TRD tracklets are collected that serve as an input to the GNNs. Now, the goal is to classify whether a nuclear interaction of these antinuclei takes place or not. Within this study, a perfect TPC track matching is assumed.

10.2.1 Dataset

For this study, an MC simulation with 257,200 p–Pb events at center-of-mass energies of $\sqrt{s_{\text{NN}}} = 5.02$ TeV were generated. The previous analyzed data could not be used, as it was necessary to use GEANT4 [107]. In GEANT3 nuclear interaction are not considered. In addition, in each event antinuclei are injected: one antideuteron (d), one antitriton (t) and one antihelium-3 (^3He). This is an additional benefit of the new simulation and provides enough statistics for these studies.

The target volume to identify interactions is set to a radial distance from 270 cm to 345.5 cm. This includes most of the interactions right in front and within the TRD that are well measurable. Tracks with a daughter creation before these 270 cm are rejected. Furthermore, to ensure a sufficient data quality, the MC antinucleus momentum is larger than $0.3 \text{ GeV}/c$ and $|\eta| < 0.84$ to reach the TRD. Now, signal events are required to have at least two daughter tracks with a momentum larger than $40 \text{ MeV}/c$.

The following values serve as input to the final training procedure: All tracks that are closer than 35 cm to the TPC tracks are used. This selection rejects a sufficient number of noise tracklets. At least seven tracklets are required, such that a reasonable graph can be constructed. These tracklets now serve as nodes for the input graph to the GNN. Besides the position and direction of the tracklets, the three-dimensional distance to the TPC track was calculated as an additional node feature.

In Fig. 10.4 and Fig. 10.5 examples of selected signal and background events are shown. The fragmentation of the different daughter tracks is clearly visible. An attribute of the background events is the fact that often the track of the charged nucleus generates tracklets. This is a reason why it makes sense to use also the distance to the TPC track as additional node feature.

10.2.2 Model and Workflow

Details of the optimization of the GNN can be found in [106]. For the graph construction a simple approach is to use a k-nearest-neighbor (kNN) graph. Five neighbors were found to be a good choice. In the following, Graph Attention (GAT) layers and GraphConv layers are used as they gave the best results. GraphConv, as proposed in [108], is a slight adaption of the graph convolution as shown before. Here, weight matrices are applied to the neighborhood. The general design of the full network (the same for both variants) is shown in Fig. 10.6. The optimization procedure by varying the architecture led to a design of three message passing layers. Furthermore, a readout multilayer perceptron (MLP) (a dense layer as explained

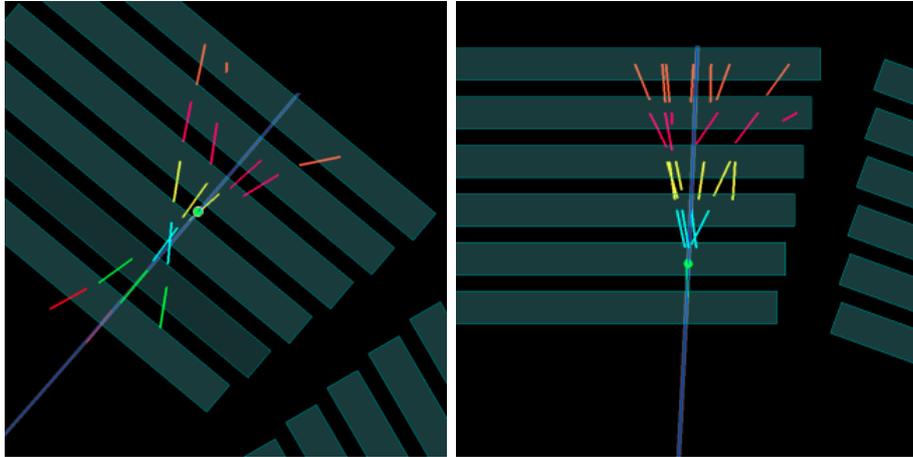


Figure 10.4: Two examples of input signals: a ${}^3\bar{\text{He}}$ (left) and $\bar{\text{d}}$ (right). The TPC track and the interaction vertex are also shown. Figures taken from [106].

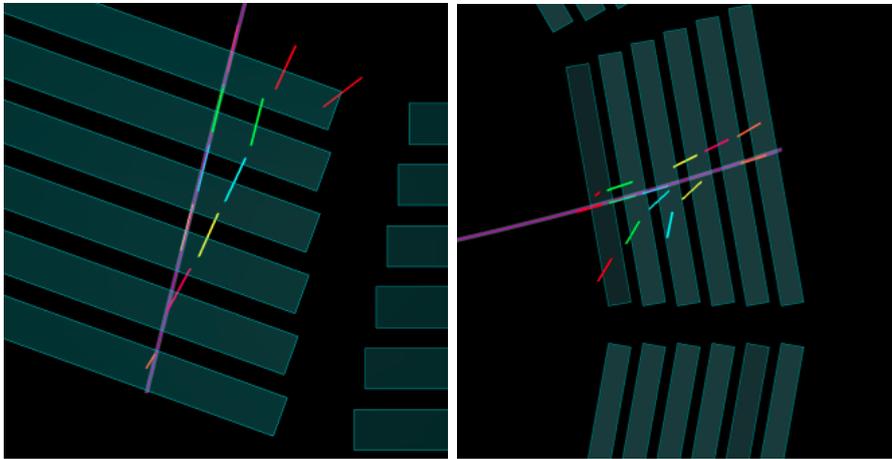


Figure 10.5: Two examples for background tracks from antideuterons. The track of the particles can be seen and thus it might be identified by the GNN. Figures taken from [106].

10.2 Identification of Inelastic Interactions of Light Antinuclei

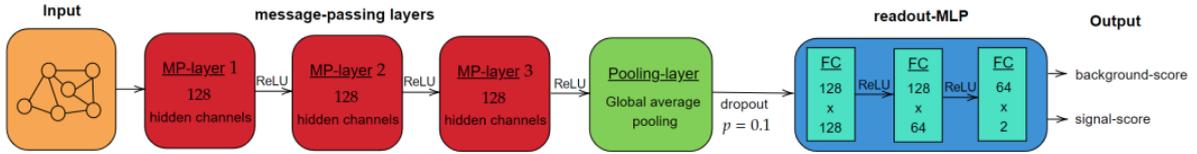


Figure 10.6: The design of the full GNNs as used in the shown analysis. Figure taken from [106].

in the section before) after the pooling was crucial for a good performance. After a training over 200 epochs (with a learning rate adjusted between the epochs) no significant improvement was found.

As a benchmark a random forest was chosen as reference using the implementation from scikit learn [109]. In the underlying work also a simple cut-based model was constructed for comparison, which is by construction disadvantaged compared to the random forest and thus not shown here. Due to the complex and variable data structure, simpler input features need to be constructed for the random forest. The following features were chosen:

- distances of the seven nearest tracklet position coordinates to the TPC track
- mean distance of all tracklet positions to the TPC track
- number of assigned tracklets

The hyperparameters were optimized with a cross validation procedure.

10.2.3 Performance

The two variants of the GNNs are now compared to the random forest. In Fig. 10.7 the ROC (receiver operating characteristic) curves, which show the signal efficiency vs. the background efficiency, as well as the precision-recall curves, where the purity is shown as a function of the signal efficiency. With the GNNs as well as the random forest a clear separation of the signal and background is possible. Thereby, the GNN performs better than the random forest model. Both GNN variants perform similarly well.

In Fig. 10.8 the purity is shown for all three models as a function of the transverse momentum for three different signal efficiencies. The models perform well over the full range of transverse momenta. With an efficiency of 70% a purity clearly above

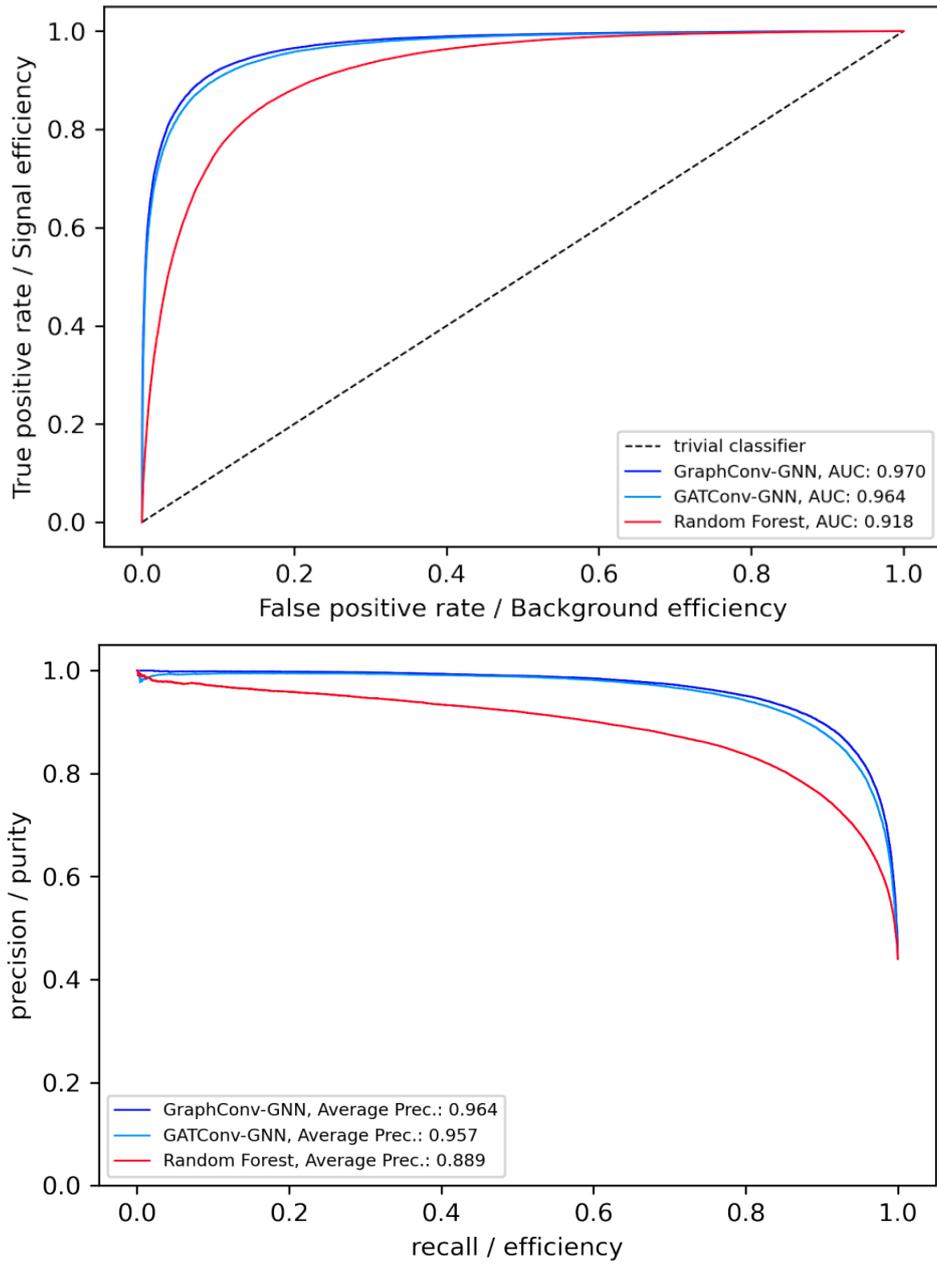


Figure 10.7: The signal efficiency vs. the background efficiency (upper plot) as well as the purity vs. the signal efficiency (lower plot) are shown for the two GNNs and the random forest. Figures taken from [106].

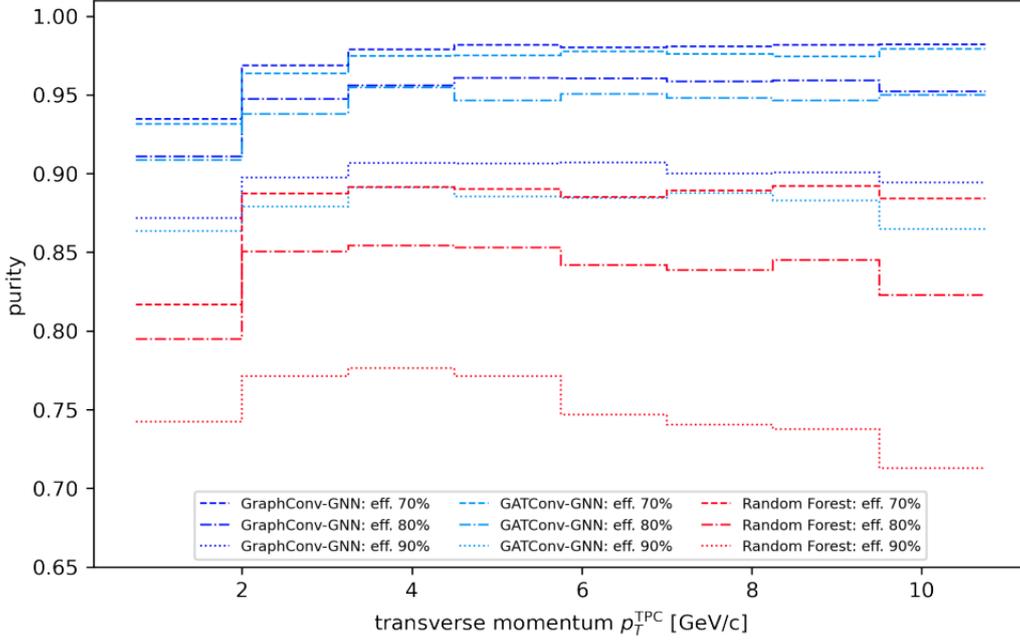


Figure 10.8: The purity is shown as a function of the transverse momentum for different signal efficiencies for the three models. Figure taken from [106].

90% can be achieved for GNNs. They perform better than the random forest over the full p_T range.

The GNNs provide a better exploitation of the underlying structures, probably due to the access to the full raw input information. But it should be mentioned that the training of the GNNs took order of days while the random forest can be trained within minutes. In [106] also the optimal operating point to set the background efficiency as function of the signal efficiency is discussed.

10.3 Identification of the Sexaquark

As a second application of the identification of particle interactions in the material with the TRD using GNNs, a simple sexaquark search is presented. A brief physics motivation is given in Sec. 2.6. The nuclear interaction of antisexaquarks with neutrons likely involves $V0$ decays, i.e. the Λ and/or K_S^0 . For reasons of simplicity only the channel $\bar{S} + n \rightarrow \bar{\Lambda} + K_S^0$ is considered. This channel is characterized by two $V0$'s in the final state with a clear topological structure. In contrast to the method in the previous section, no charged track that can be identified can exist as a TPC

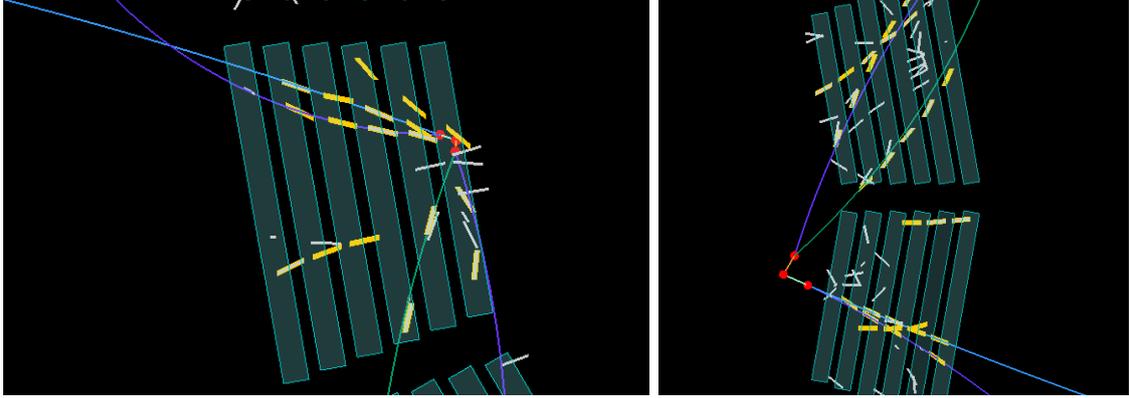


Figure 10.9: Two examples of signals from interactions of sexaquarks in the TRD. The TRD tracklets in the six TRD layers as well as the true MC tracks are shown.

track. Thus, a simple approach is to use the full existing background, i.e. all TRD sectors, and search for the interactions. This leads to a large amount of background compared to signal data and the goal is to achieve a very high background suppression. Another drawback is that the time for applying the GNN to the data is much longer compared to having a simple preselection. In this work, only Monte Carlo data is considered.

In Fig. 10.9 two examples of signals from the sexaquark in the TRD are shown. The Monte Carlo tracks with the parameters at the interaction vertex are also shown. As can be seen, the signal often spreads over more than one module.

10.3.1 Dataset

For this analysis a simulation was set up with roughly 200,000 p–Pb events anchored to the same runs as before [110]. In each event one sexaquark was injected. The mass is assumed to be $m_{\bar{s}} = 1.8 \text{ GeV}/c^2$. The transverse momentum p_T is flat and the upper limit is $5 \text{ GeV}/c$. The distribution of the rapidity y is flat with $|y| < 0.8$. Each sexaquark interaction happens in the given channel in a radial distance to the vertex between 260 cm and 340 cm such that enough signals can be measured in the TRD.

To obtain clean training data a few selection criteria are applied. Only decays without π^0 s are allowed. Thus, only $\bar{\Lambda}^0 \rightarrow \bar{p} + \pi^+$ (63.9%) and $K_S^0 \rightarrow \pi^+ \pi^-$ (69.2%) are considered, which reduces the statistics to 44.2%.

As in the photon analysis, again the tracklet labels are unknown. As the statistics are limited and the GNN is very sensitive to the amount of data, a loose selection is applied to obtain as much data as possible. A cone selection of signal tracklets is applied to label the data. For this selection, the (straight) $\bar{\Lambda}^0$ and K_S^0 MC tracks are prolonged. Tracklets nearby are selected by the following rule: if the distance of the tracklets to the MC track is within the cone around the prolonged MC track

$$d = 10 \text{ cm} + 0.3l + 0.005l^2/\text{cm}, \quad (10.7)$$

tracklets are accepted. Here, l is a running parameter along the MC track. The sexaquark interaction vertex is at $l = 0$ cm. For both of the two cones more than five accepted tracklets are required. This loose selection is necessary to obtain enough data, as there might be many tracklets generated by other particles (e.g. knockout particles).

If a sexaquark is accepted as signal, all tracklets from the corresponding sector to which the sexaquark flight direction vector points are stored as signal data. In addition, also the tracklets in the neighboring sectors are stored, as many sexaquark daughter tracks cross more than one sector.

In total, this leads to roughly 26,000 signals as input for the following analysis. All remaining sectors including the neighboring sectors are used as background candidates. The background is required to have at least five tracklets, which is the mathematical minimum for the signal criteria in case both sexaquark daughter tracklet candidates completely overlap. The signal candidates after the selection have a mean of 59 tracklets and the background candidates a mean of 39 tracklets.

10.3.2 Model and Workflow

To train the model, one half of the data is used as training sample and the other half as test sample. The coordinates are transformed into the local coordinate system of the central sector. The position and direction vectors are used as node features for the input graph and additionally the direction vector in the xy plane is added, as here the direction resolution is cleaner due to the large z pads of the TRD. As in the previous section, the five-nearest-neighbor graph is constructed. The GNN architecture as given in the previous section is varied around the found hyperparameters and the parameter from before seem to be also a good choice for this application. The training was performed in two stages. First, a random sample of the background with same size as the signal is used. For the second stage only the background candidates that are assigned a high probability to be a signal are taken into account with the same number as the signal candidates. This does not only accelerate the training, but also leads to a better performance.

10.3.3 Performance

In Fig. 10.10 the ROC curves for the graph attention model as well as the Graph-Conv model are shown. Thereby, the data is used as given for the GNN training, i.e. after the previously mentioned selection. Thus, some background is already suppressed before. The result after both training stages is shown. The background efficiency is shown with a linear scale as well as with a logarithmic scale. Again, both kinds of GNNs perform similarly. With a signal efficiency of around 80% it is possible to reject 99% of the background and for signal efficiencies of 20% it is possible to reject 99.9% of the background. But still, in such an analysis there is much more background than signal. In general, this procedure leads to promising results.

Classification dependencies with respect to the momentum or position were checked and were not observed. During the training it was seen that the performance is still very sensitive to the amount of data. In Fig. 10.11 a with a high probability correctly classified signal and a wrong classified signal are shown. In Fig. 10.12 a background candidate with a high signal probability given by the output of the GNN is shown. From these examples it can be already seen that the quality of the data can be improved. With more statistics it is possible to select for signals with a higher quality which would lead to a better classification performance. This cannot be done with the given data, because the given selection provides only 26,000 true signals. Also without a better label selection, the GNNs were sensitive to the amount of data and more data will still deliver better results.

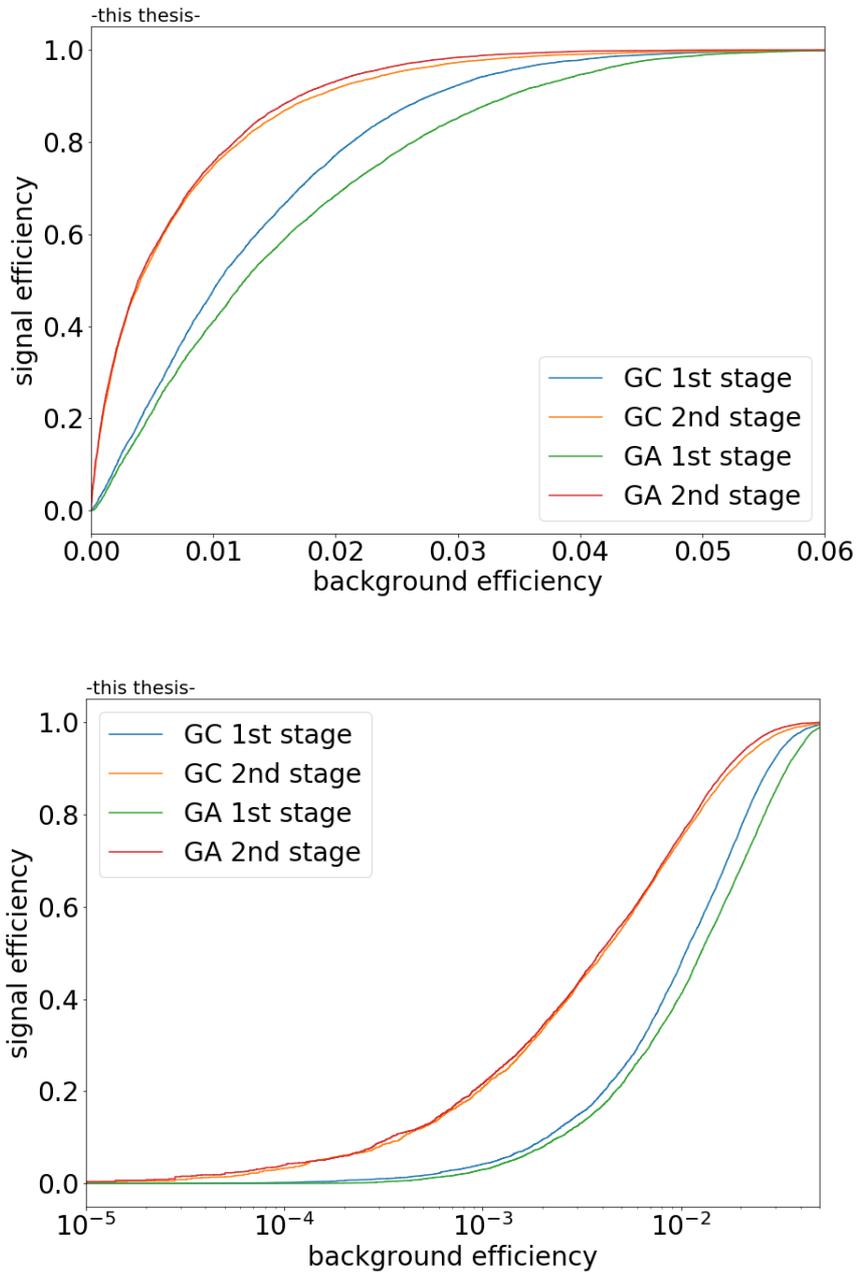


Figure 10.10: The signal efficiency is shown as a function of the background efficiency for the two variants of the GNN, the GraphConv (GC) and the graph attention (GA) based networks. On the upper plot a linear scale and in the lower plot a logarithmic scale is used. A more efficient selection of the background (2nd stage) leads a significant better performance.

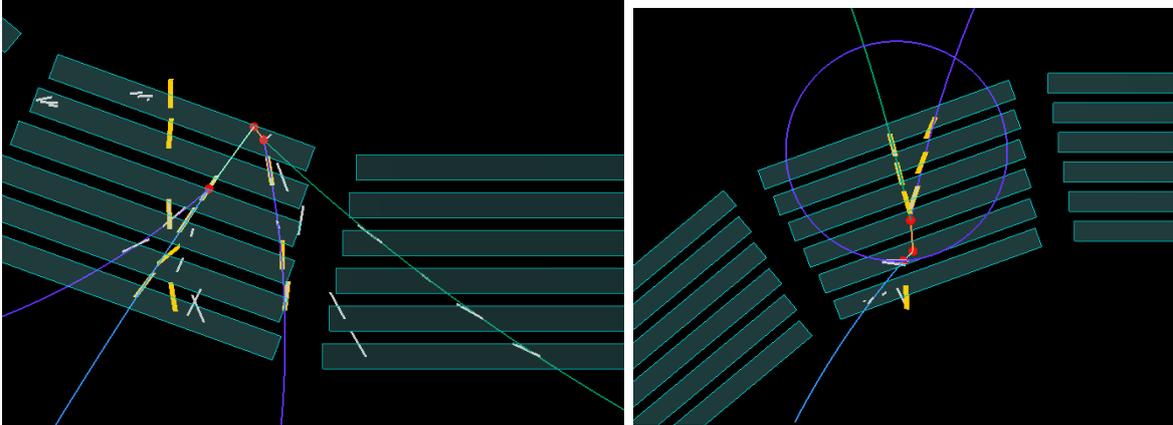


Figure 10.11: Left: a true positive sexaquark candidate. Right: a false negative sexaquark candidate.

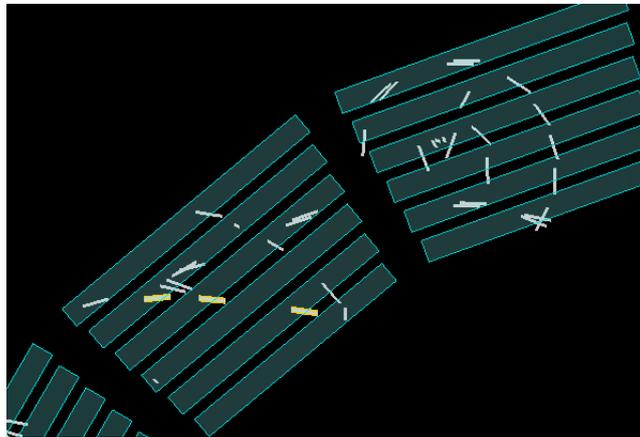


Figure 10.12: A false positive sexaquark candidate.

11 Summary and Outlook

The Transition Radiation Detector (TRD) [3] is constructed especially for electron triggering and particle identification. It also contributes to the tracking and calibration of the central barrel. In this work, it was shown that the TRD is also capable of reconstructing photon conversions from larger radii. In addition, it also can be used for the identification of inelastic interactions of light antinuclei as well as for the search for sexaquarks.

For the reconstruction of the photons different steps were performed. Track candidates for the conversion products, the electrons and positrons, were identified using a newly developed stand-alone tracking including track finding and track parameter estimation. A new photon fit procedure is then performed on top of the found candidates. It exploits most of the combined kinematic information of the conversion daughter tracks. It is more precise than the typical kinematic fit used in ALICE, the KFPARTICLE package [86][87]. The TRD is intentionally not constructed for stand-alone tracking. The precision of the combined photon daughter fit allows nevertheless a very good reconstruction. We call this fit syncfit. The advantages of this fit, especially of the precision in the parameter estimation, is shown first using toy data and then also with a full GEANT ALICE simulation (p-Pb at $\sqrt{s_{NN}} = 5.02$ TeV + injected π^0 s) as well as with real p-Pb data at $\sqrt{s_{NN}} = 5.02$ TeV. A clearly well centered neutral pion peak confirms the applicability to real measured data. A Pb-Pb collision system has not yet been considered. For the higher multiplicities a distinct modification of the tracking parameters would be necessary, i.e. smaller cells for the combinatorial track finding.

The capabilities of the full reconstruction chain of the photons for real data as well as for the data of the full GEANT simulation were shown with a reconstruction procedure of the track segments called tracklets created among other purposes for this work. It is not based on the currently implemented TRD online tracklet reconstruction, as the goal is to show the best possible result quality. The online tracklets are limited especially for lower momenta by the inclination. The reconstruction procedure used in this work does not provide the real tracklet labels, and thus the tracklet information is obtained via a matching with the Monte Carlo labels with the kinematic properties at the vertices afterwards. A propagation of the true labels could lead to better hyperparameter estimation for different steps within the

11 Summary and Outlook

track reconstruction.

The stand-alone tracking was built respecting the challenges given by the construction of the TRD and in view of being able to find as many electron/positron candidates as possible for the photon reconstruction. These electron/positron tracks suffer from bremsstrahlung. A combinatorial approach is chosen that is capable of handling noise tracklets, empty layers, shared clusters and also blurring effects from the material. The strategy thereby is similar to the cellular automaton. The fit precision is just as good as necessary to have a good selection for the final photon fits, but not the focus during the single track finding. The purity is later optimized for the photons, as there is more information accessible. Pad tilting is thus neglected at this stage. The efficiency for electrons/positrons starting at transverse momenta of 100 MeV/ c is above 95%. At smaller transverse momenta multiple scattering has a deteriorating influence.

For the photon reconstruction a full kinematic fit called syncfit was developed. Thereby the tracks of the positron and electron are fitted synchronously with its topological constraints to the tracklets within the TRD, i.e. not with tracks as it is done within the KFParticle package [86][87]. A primary vertex position and an opening angle of zero are constraint. Corrections for the mean bremsstrahlung and the pad tilting are included. The fit is technically implemented with TensorFlow which allows one to apply the complex relations from the kinematics without approximations. Standard tools for kinematic fitting such as KFParticle use linear approximations and soft constraints. It was shown via toy data that compared to these tools the syncfitter gives better results for all kinematic parameters and provides a good transverse momentum resolution also for higher transverse momenta. The high parallelizability of TensorFlow permits a reasonable runtime for these complex calculations, although the syncfit is much slower than KFParticle.

Applied to the full GEANT simulation, a photon efficiency that is close to 90% could be achieved for findable candidates. The purity of these data suffered at lower momenta especially from secondary bremsstrahlung arising from high momentum particles in the inner tracking system. If this contribution to impurities is neglected, the purity is clearly larger than 95%. The contribution from bremsstrahlung is distinctive in this simulation, as many electrons and positrons at large momenta are created by the many photons from neutral pions, that were injected to this simulation with a flat momentum distribution. A higher purity might also be achieved by using the particle identification properties of the detector. The lower limit for the reconstructable transverse momentum for the photons is 200 MeV/ c . Below, multiple scattering disturbs the reconstruction. Due to the relatively large size of the

pads in z direction of the TRD, the momentum conservation of multiple scattering cannot be exploited within the reconstruction.

For the first time, a tomographical map of the TRD material distribution is shown that is created from the conversion vertices. Contributions from support material structures are clearly visible. Photons from between the chambers do not give sufficient signals to be reconstructed. This improved knowledge about the material in the TRD and especially the difference between the Monte Carlo data and the real data can contribute to the matching with the time-of-flight detector positioned behind the TRD by a more accurate track propagation.

As a cross check neutral pions are reconstructed. The invariant mass of two photon combinations show a clear peak at the position of the neutral pion mass. The lower limit in the transverse momentum is roughly $700 \text{ MeV}/c$. By using the TRD in conjunction with the TPC, the yield of using only TPC-TPC pion candidates can be increased by 75% by adding TRD photons.

Furthermore, graph neural networks were studied for the identification of interactions in the material in a full Monte Carlo simulation. The interactions of antineutrons were detected using candidates selected along the TPC tracks of the respective particles. For the sexaquark, no such clear preselection was possible. A first approach was to use all chambers as input to the algorithm. In both cases, the neural networks give promising results. The most straightforward way to achieve even better results is to improve the statistics of the Monte Carlo data. Especially in the case of the sexaquark, a subsequent analysis of the found candidates could provide information about the kinematics and thus the basis for a hypothesis test. Inspiration for such an analysis could also be given by the syncfitter.

In conclusion, the syncfitter is a good tool for high precision photon fits. Candidates from nuclear interaction in the TRD can be found using graph neural networks. An implementation using the currently in Run 3 used online tracking is not clear. It would be possible to think of an algorithm that finds promising candidates and thus respects the online readout capacities in addition. Dataflow programming known from machine learning is suitable to improve the modeling of complex topological physical processes in particle detectors in a very efficient way. The intrinsic vectorization capabilities lead to a natural usage of modern parallelization devices such as GPUs which are necessary to handle the large amount of data in the future.

A Syncfitter Momentum Resolution Estimation

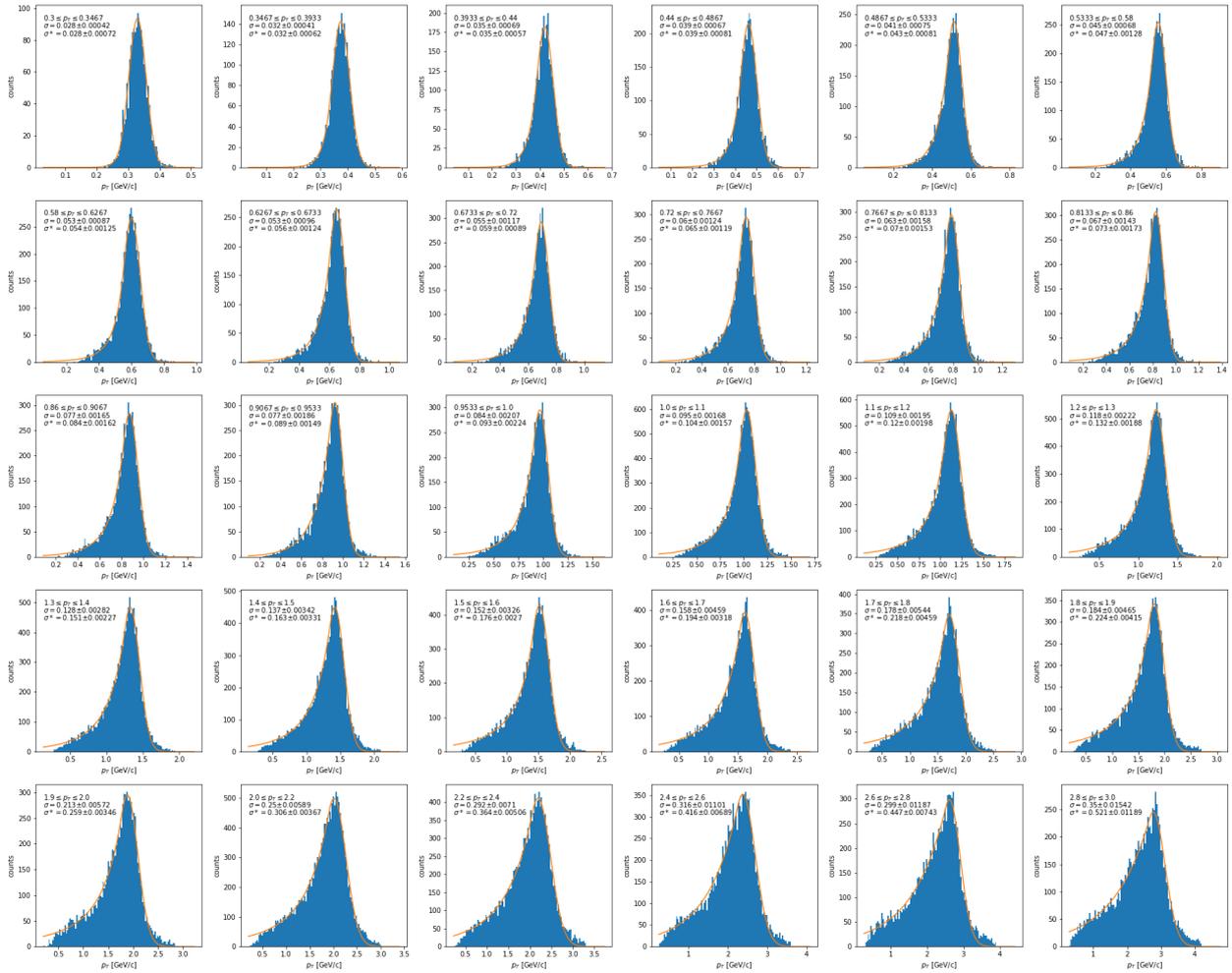


Figure A.1: The fits of the measured transverse momentum in different MC truth transverse momentum intervals for the resolution estimation in Ch. 9. The Gaussian standard deviation is given by σ and σ^* accounts for the tail as described in the text.

B Pion Invariant Mass Fits

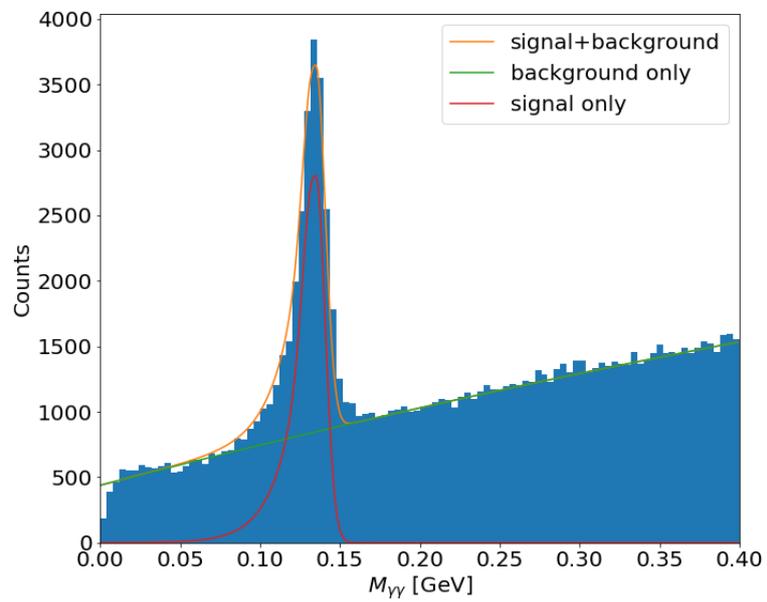


Figure B.1: Fit of the neutral pion invariant mass peak for TPC-TPC photon combinations.

B Pion Invariant Mass Fits

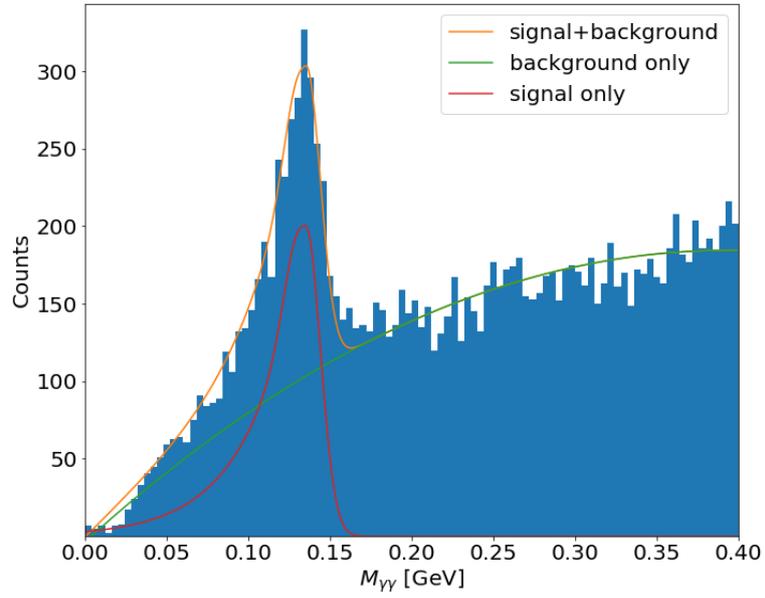


Figure B.2: Fit of the neutral pion invariant mass peak for TRD-TRD photon combinations.

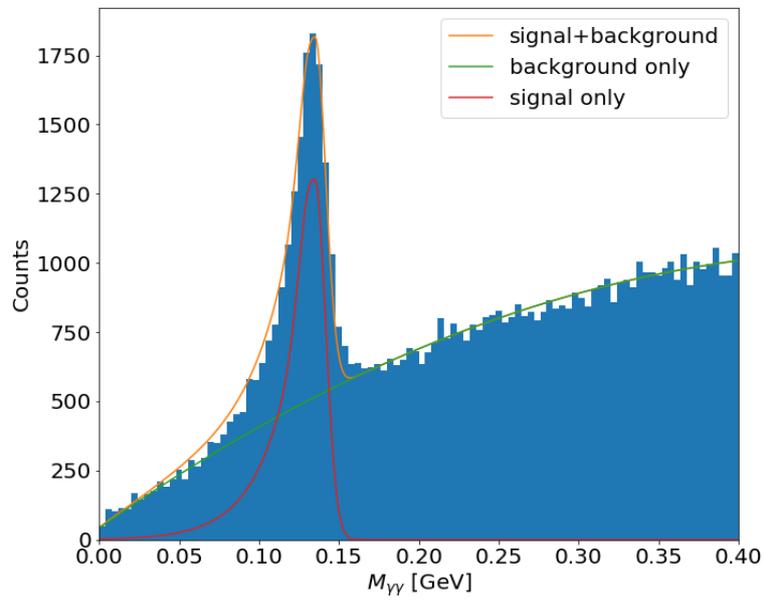


Figure B.3: Fit of the neutral pion invariant mass peak for TPC-TRD photon combinations.

Bibliography

- [1] L. Evans and P. Bryant. LHC Machine. *J. Instrum.*, 3(08):S08001, 2008. doi:10.1088/1748-0221/3/08/S08001.
- [2] K. Aamodt et al. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002, 2008. doi:10.1088/1748-0221/3/08/S08002.
- [3] S. Acharya et al. The ALICE Transition Radiation Detector: construction, operation, and performance. *Nucl. Instrum. Meth. A*, 881:88–127, 2018. doi:10.1016/j.nima.2017.09.028.
- [4] B. Abelev et al. Neutral pion production at midrapidity in pp and Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. *Eur. Phys. J. C*, 74(10):3108, 2014. doi:10.1140/epjc/s10052-014-3108-8.
- [5] J. Adolfsson et al. The upgrade of the ALICE TPC with GEMs and continuous readout. *Journal of Instrumentation*, 16(03):P03022, 2021. doi:10.1088/1748-0221/16/03/p03022.
- [6] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521:436–44, 2015. doi:10.1038/nature14539.
- [7] T. Dorigo et al. Toward the end-to-end optimization of particle physics instruments with differentiable programming: a white paper, 2022. doi:10.48550/ARXIV.2203.13818.
- [8] M. Abadi et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org. URL: <https://www.tensorflow.org/>.
- [9] A. Paszke et al. Pytorch: An imperative style, high-performance deep learning library, 2019. doi:10.48550/ARXIV.1912.01703.
- [10] S. Acharya et al. First measurement of the absorption of ${}^3\overline{\text{He}}$ nuclei in matter and impact on their propagation in the galaxy, 2022. doi:10.48550/ARXIV.2202.01549.
- [11] G. R. Farrar. A stable sexaquark: Overview and discovery strategies, 2022. doi:10.48550/ARXIV.2201.01334.

Bibliography

- [12] J. T. De Clercq. *The Upgraded Outer Tracker for the CMS Detector at the High Luminosity LHC, and Search for Composite Standard Model Dark Matter with CMS at the LHC*. Doctoral thesis, 2019. URL: <https://cds.cern.ch/record/2708026>.
- [13] J. Duarte and J.-R. Vlimant. Graph neural networks for particle tracking and reconstruction. In *Artificial Intelligence for High Energy Physics*, pages 387–436. 2022. doi:10.1142/9789811234033_0012.
- [14] M. Gell-Mann. A Schematic Model of Baryons and Mesons. *Phys. Lett.*, 8:214–215, 1964. doi:10.1016/S0031-9163(64)92001-3.
- [15] G. Zweig. *An SU(3) model for strong interaction symmetry and its breaking. Version 2*, pages 22–101. 1964. URL: <https://cds.cern.ch/record/570209>.
- [16] M. Gell-Mann. The Eightfold Way: A Theory of strong interaction symmetry. 1961. doi:10.2172/4008239.
- [17] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, 1995. Reading, USA: Addison-Wesley (1995) 842 p.
- [18] M. Y. Han and Y. Nambu. Three-triplet model with double SU(3) symmetry. *Phys. Rev.*, 139:B1006–B1010, 1965. doi:10.1103/PhysRev.139.B1006.
- [19] J. E. Augustin et al. Discovery of a narrow resonance in e^+e^- annihilation. *Phys. Rev. Lett.*, 33:1406–1408, 1974. doi:10.1103/PhysRevLett.33.1406.
- [20] J. J. Aubert et al. Experimental Observation of a Heavy Particle *J. Phys. Rev. Lett.*, 33:1404–1406, 1974. doi:10.1103/PhysRevLett.33.1404.
- [21] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak Interactions with Lepton-Hadron Symmetry. *Phys. Rev. D*, 2:1285–1292, 1970. doi:10.1103/PhysRevD.2.1285.
- [22] K. G. Wilson. Confinement of quarks. *Phys. Rev. D*, 10:2445–2459, 1974. doi:10.1103/PhysRevD.10.2445.
- [23] K. G. Wilson. The Origins of lattice gauge theory. *Nucl. Phys. B Proc. Suppl.*, 140:3–19, 2005. doi:10.1016/j.nuclphysbps.2004.11.271.
- [24] D. J. Gross and F. Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Phys. Rev. Lett.*, 30:1343–1346, 1973. doi:10.1103/PhysRevLett.30.1343.
- [25] H. D. Politzer. Reliable perturbative results for strong interactions? *Phys. Rev. Lett.*, 30:1346–1349, 1973. doi:10.1103/PhysRevLett.30.1346.

- [26] N. Cabibbo and G. Parisi. Exponential hadronic spectrum and quark liberation. *Physics Letters B*, 59(1):67–69, 1975. doi:[https://doi.org/10.1016/0370-2693\(75\)90158-6](https://doi.org/10.1016/0370-2693(75)90158-6).
- [27] E. V. Shuryak. Theory of Hadronic Plasma. *Sov. Phys. JETP*, 47:212–219, 1978.
- [28] A. Bazavov et al. Equation of state in (2+1)-flavor QCD. *Phys. Rev. D*, 90(9), 2014. doi:[10.1103/physrevd.90.094503](https://doi.org/10.1103/physrevd.90.094503).
- [29] R. Hagedorn. Statistical thermodynamics of strong interactions at high energies. *Nuovo Cimento, Suppl.*, 3:147–186, 1965. URL: <https://cds.cern.ch/record/346206>.
- [30] M. Durante et al. All the fun of the FAIR: fundamental physics at the facility for antiproton and ion research. *Physica Scripta*, 94(3):033001, 2019. doi:[10.1088/1402-4896/aaf93f](https://doi.org/10.1088/1402-4896/aaf93f).
- [31] T. Bhattacharya et al. Qcd phase transition with chiral quarks and physical quark masses. *Phys. Rev. Lett.*, 113:082001, 2014. doi:[10.1103/PhysRevLett.113.082001](https://doi.org/10.1103/PhysRevLett.113.082001).
- [32] J. Adams et al. Experimental and theoretical challenges in the search for the quark gluon plasma: The STAR collaboration’s critical assessment of the evidence from RHIC collisions. *Nucl. Phys. A*, 757(1-2):102–183, 2005. doi:[10.1016/j.nuclphysa.2005.03.085](https://doi.org/10.1016/j.nuclphysa.2005.03.085).
- [33] B. B. Back et al. The PHOBOS perspective on discoveries at RHIC. *Nucl. Phys. A*, 757:28–101, 2005. doi:[10.1016/j.nuclphysa.2005.03.084](https://doi.org/10.1016/j.nuclphysa.2005.03.084).
- [34] K. Adcox et al. Formation of dense partonic matter in relativistic nucleus-nucleus collisions at RHIC: Experimental evaluation by the PHENIX collaboration. *Nucl. Phys. A*, 757:184–283, 2005. doi:[10.1016/j.nuclphysa.2005.03.086](https://doi.org/10.1016/j.nuclphysa.2005.03.086).
- [35] I. Arsene et al. Quark gluon plasma and color glass condensate at RHIC? The Perspective from the BRAHMS experiment. *Nucl. Phys. A*, 757:1–27, 2005. doi:[10.1016/j.nuclphysa.2005.02.130](https://doi.org/10.1016/j.nuclphysa.2005.02.130).
- [36] T. Ablyazimov et al. Challenges in QCD matter physics –the scientific programme of the compressed baryonic matter experiment at FAIR. *The European Physical Journal A*, 53(3), 2017. doi:[10.1140/epja/i2017-12248-y](https://doi.org/10.1140/epja/i2017-12248-y).

Bibliography

- [37] Y. Aoki et al. The order of the quantum chromodynamics transition predicted by the standard model of particle physics. *Nature*, 443(7112):675–678, 2006. doi:10.1038/nature05120.
- [38] P. Braun-Munzinger and B. Dönigus. Loosely-bound objects produced in nuclear collisions at the lhc. *Nucl. Phys. A*, 987:144–201, 2019. doi:10.1016/j.nuclphysa.2019.02.006.
- [39] K. Aamodt et al. Two-pion Bose–Einstein correlations in central Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV. *Physics Letters B*, 696(4):328–337, 2011. doi:10.1016/j.physletb.2010.12.053.
- [40] S. Turbide, R. Rapp, and C. Gale. Hadronic production of thermal photons. *Phys. Rev. C*, 69:014903, 2004. doi:10.1103/PhysRevC.69.014903.
- [41] M. H. Thoma. Damping rate of a hard photon in a relativistic plasma. *Phys. Rev. D*, 51:862–865, 1995. doi:10.1103/PhysRevD.51.862.
- [42] J.-F. Paquet et al. Production of photons in relativistic heavy-ion collisions. *Phys. Rev. C*, 93(4), 2016. doi:10.1103/physrevc.93.044906.
- [43] P. Aurenche et al. Prompt photon production at large p_T in QCD beyond the leading order. *Physics Letters B*, 140(1):87–92, 1984. doi:10.1016/0370-2693(84)91053-0.
- [44] M. Klasen et al. How robust is a thermal photon interpretation of the ALICE low- p_T data? *Journal of High Energy Physics*, 2013(10), 2013. doi:10.1007/jhep10(2013)119.
- [45] J. Churchill, L. Yan, S. Jeon, and C. Gale. Electromagnetic radiation from the pre-hydrodynamics stage of the quark-gluon plasma. *Nucl. Phys. A*, 1005:121946, 2021. doi:10.1016/j.nuclphysa.2020.121946.
- [46] R. Snellings. Elliptic flow: a brief review. *New Journal of Physics*, 13(5):055008, 2011. doi:10.1088/1367-2630/13/5/055008.
- [47] F. Prino and R. Rapp. Open heavy flavor in QCD matter and in nuclear collisions. *Journal of Physics G: Nuclear and Particle Physics*, 43(9):093002, 2016. doi:10.1088/0954-3899/43/9/093002.
- [48] A. Ibarra and S. Wild. Prospects of antideuteron detection from dark matter annihilations or decays at AMS-02 and GAPS. *Journal of Cosmology and Astroparticle Physics*, 2013(02):021–021, 2013. doi:10.1088/1475-7516/2013/02/021.

- [49] G. R. Farrar. Stable sexaquark, 2018. doi:10.48550/ARXIV.1708.08951.
- [50] R. L. Jaffe. Perhaps a Stable Dihyperon. *Phys. Rev. Lett.*, 38:195–198, 1977. doi:10.1103/PhysRevLett.38.195.
- [51] G. Aad et al. The ATLAS experiment at the CERN large hadron collider. *Journal of Instrumentation*, 3(08):S08003–S08003, 2008. doi:10.1088/1748-0221/3/08/s08003.
- [52] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004–S08004, 2008. doi:10.1088/1748-0221/3/08/s08004.
- [53] A. A. Alves et al. The LHCb detector at the LHC. *Journal of Instrumentation*, 3(08):S08005–S08005, 2008. doi:10.1088/1748-0221/3/08/s08005.
- [54] E. Mobs. The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018. 2018. General Photo. URL: <https://cds.cern.ch/record/2636343>.
- [55] A. Tauro. ALICE Schematics. General Photo, 2017. URL: <https://cds.cern.ch/record/2263642>.
- [56] B. Abelev et al. Performance of the ALICE Experiment at the CERN LHC. *Int. J. Mod. Phys.*, A29:1430044, 2014. doi:10.1142/S0217751X14300440.
- [57] P Antonioli, A Kluge, and W Riegler. Upgrade of the ALICE Readout & Trigger System. Technical report, 2013. URL: <https://cds.cern.ch/record/1603472>.
- [58] G. Dellacasa et al. ALICE technical design report of the inner tracking system (ITS). 1999. URL: <https://cds.cern.ch/record/391175>.
- [59] B. Abelev et al. Technical Design Report for the Upgrade of the ALICE Inner Tracking System. *Journal of Physics G: Nuclear and Particle Physics*, 41(8):087002, 2014. doi:10.1088/0954-3899/41/8/087002.
- [60] J. Alme et al. The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events. *Nucl. Instr. Meth. Phys. Res. A*, 622:316–367, 2010. doi:10.1016/j.nima.2010.04.042.
- [61] B. Abelev et al. Upgrade of the ALICE Time Projection Chamber. Technical report, 2013. URL: <https://cds.cern.ch/record/1622286>.

Bibliography

- [62] M. O. Schmidt. *Space Point Calibration of the ALICE TPC with Track Residuals*. Doctoral thesis, University of Heidelberg, 2020. doi:[10.11588/heidok.00028663](https://doi.org/10.11588/heidok.00028663).
- [63] A. Akindinov et al. Performance of the ALICE Time-Of-Flight detector at the LHC. *Eur. Phys. J. Plus*, 128(4):44, 2013. doi:[10.1140/epjp/i2013-13044-x](https://doi.org/10.1140/epjp/i2013-13044-x).
- [64] V. L. Ginzburg and I. M. Frank. Radiation of a uniformly moving electron due to its transition from one medium into another. *J. Phys.(USSR)*, 9:353–362, 1945.
- [65] G.M. Garibian, L.A. Gevorgian, and C. Yang. The calculation of x-ray transition radiation generated in regular- and irregular-layered media. *Nucl. Instr. Meth. Phys. Res.*, 125(1):133 – 137, 1975. doi:[10.1016/0029-554X\(75\)90565-0](https://doi.org/10.1016/0029-554X(75)90565-0).
- [66] A. Andronic. Electron identification performance with ALICE TRD prototypes. *Nucl. Instr. Meth. Phys. Res. A*, 522:40–44, 2004. doi:[10.1016/j.nima.2004.01.015](https://doi.org/10.1016/j.nima.2004.01.015).
- [67] J. Klein. *Jet Physics with A Large Ion Collider Experiment at the Large Hadron Collider*. Doctoral thesis, University of Heidelberg, 2014. doi:[10.11588/heidok.00017710](https://doi.org/10.11588/heidok.00017710).
- [68] R. Frühwirth and A. Strandlie. Pattern recognition, tracking and vertex reconstruction in particle detectors. 2021. doi:[10.1007/978-3-030-65771-0](https://doi.org/10.1007/978-3-030-65771-0).
- [69] Y. Belikov, K. Safarik, and B. Batyunya. Kalman Filtering Application for Track Recognition and Reconstruction in ALICE Tracking System. Technical report, CERN, Geneva, 1997. URL: <http://cds.cern.ch/record/689414>.
- [70] I. Kisel. Event reconstruction in the CBM experiment. *Nuclear Instruments & Methods in Physics Research Section A-accelerators Spectrometers Detectors and Associated Equipment*, 566:85–88, 2006. doi:[10.1016/j.nima.2006.05.040](https://doi.org/10.1016/j.nima.2006.05.040).
- [71] J. D. Jackson. *Classical electrodynamics*. Wiley, New York, NY, 3rd ed. edition, 1999. URL: <http://cdsweb.cern.ch/record/490457>.
- [72] P. A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020. doi:[10.1093/ptep/ptaa104](https://doi.org/10.1093/ptep/ptaa104).

- [73] H. A. Bethe and W. H. Heitler. On the stopping of fast particles and on the creation of positive electrons. *Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences*, 146:83–112, 1934. doi:10.1098/rspa.1934.0140.
- [74] G. R. Lynch and O. I. Dahl. Approximations to multiple coulomb scattering. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 58(1):6–10, 1991. doi:10.1016/0168-583X(91)95671-Y.
- [75] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35–45, 1960. doi:10.1115/1.3662552.
- [76] G. L. Smith, S. F. Schmidt, and L. A. McGee. *Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity on Board a Circum-lunar Vehicle*. NASA technical report. National Aeronautics and Space Administration, 1962. URL: <https://ntrs.nasa.gov/citations/20190002215>.
- [77] J. H. Stiller. *Full kinematic reconstruction of charged B mesons with the upgraded Inner Tracking System of the ALICE Experiment*. Doctoral thesis, University of Heidelberg, 2016. doi:10.11588/heidok.00020710.
- [78] T. M. Atkinson. *Electron reconstruction with the ATLAS inner detector*. Doctoral thesis, The University of Melbourne, 2006. URL: <https://cds.cern.ch/record/1448253>.
- [79] V. Karimäki. Effective circle fitting for particle trajectories. *Nucl. Instrum. Meth. A*, 305:187–191, 1991. doi:10.1016/0168-9002(91)90533-V.
- [80] D. Rohr. Tracking performance in high multiplicities environment at ALICE, 2017. doi:10.48550/ARXIV.1709.00618.
- [81] C. D. Anderson. The positive electron. *Phys. Rev.*, 43:491–494, 1933. doi:10.1103/PhysRev.43.491.
- [82] C. Grupen and B. Shwartz. *Particle Detectors*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology. Cambridge University Press, 2008. doi:10.1017/CB09780511534966.
- [83] Y.-S. Tsai. Pair production and bremsstrahlung of charged leptons. *Rev. Mod. Phys.*, 46:815–851, 1974. doi:10.1103/RevModPhys.46.815.
- [84] H. Olsen. Opening angles of electron-positron pairs. *Phys. Rev.*, 131:406–415, 1963. doi:10.1103/PhysRev.131.406.

Bibliography

- [85] G. Baroni, A. Boesellino, Livio Scarsi, and G. Vanderhaeghe. An analysis of the electron pairs produced in nuclear emulsions by cosmic rays photons. *Il Nuovo Cimento (1943-1954)*, 10:1653–1661, 1953. doi:[10.1007/BF02781660](https://doi.org/10.1007/BF02781660).
- [86] S. Gorbunov. *On-line reconstruction algorithms for the CBM and ALICE experiments*. Doctoral thesis, Goethe University Frankfurt, 2013.
- [87] M. Zyzak. *Online selection of short-lived particles on many-core computer architectures in the CBM experiment at FAIR*. Doctoral thesis, Goethe University Frankfurt, 2016.
- [88] L. Leardini. *Measurement of neutral mesons and direct photons in Pb–Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV with the ALICE experiment at the LHC*. Doctoral thesis, University of Heidelberg, 2017. doi:[10.11588/heidok.00023756](https://doi.org/10.11588/heidok.00023756).
- [89] J. Podolanski and R. Armenteros. III. Analysis of V-events. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 45(360):13–30, 1954. doi:[10.1080/14786440108520416](https://doi.org/10.1080/14786440108520416).
- [90] T. Dahms. Measurement of photons via conversion pairs with the PHENIX experiment at RHIC. Master’s thesis, Stony Brook University, 2005.
- [91] N. P. Jouppi et al. In-datacenter performance analysis of a tensor processing unit, 2017. doi:[10.48550/ARXIV.1704.04760](https://doi.org/10.48550/ARXIV.1704.04760).
- [92] M. Fey and J. E. Lenssen. Fast Graph Representation Learning with PyTorch Geometric, 2019. doi:[10.48550/ARXIV.1903.02428](https://doi.org/10.48550/ARXIV.1903.02428).
- [93] A. G. Baydin et al. Automatic differentiation in machine learning: a survey, 2018. doi:[10.48550/ARXIV.1502.05767](https://doi.org/10.48550/ARXIV.1502.05767).
- [94] A. Schmäh. private communication.
- [95] X.-N. Wang and M. Gyulassy. HIJING: A Monte Carlo model for multiple jet production in pp, pA, and AA collisions. *Phys. Rev. D*, 44:3501–3516, 1991. doi:[10.1103/PhysRevD.44.3501](https://doi.org/10.1103/PhysRevD.44.3501).
- [96] S. Acharya et al. Measurement of the low-energy antideuteron inelastic cross section. *Phys. Rev. Lett.*, 125:162001, 2020. doi:[10.1103/PhysRevLett.125.162001](https://doi.org/10.1103/PhysRevLett.125.162001).
- [97] P. Cortese. *ALICE transition-radiation detector: Technical Design Report*. Technical design report. ALICE. CERN, Geneva, 2001. URL: <https://cds.cern.ch/record/519145>.

- [98] A. Meurer et al. Sympy: symbolic computing in python. *PeerJ Computer Science*, 3:e103, 2017. doi:[10.7717/peerj-cs.103](https://doi.org/10.7717/peerj-cs.103).
- [99] N. Berger et al. A new three-dimensional track fit with multiple scattering. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 844:135–140, 2017. doi:[10.1016/j.nima.2016.11.012](https://doi.org/10.1016/j.nima.2016.11.012).
- [100] P. Virtanen et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi:[10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2).
- [101] M. C. Danisch. *Measurement of neutral mesons and direct photons in Pb-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV with ALICE at the LHC*. Doctoral thesis, University of Heidelberg, 2021. doi:[10.11588/heidok.00031202](https://doi.org/10.11588/heidok.00031202).
- [102] D. Bacciu et al. A gentle introduction to deep learning for graphs. *Neural Networks*, 129:203–221, 2020. doi:[10.1016/j.neunet.2020.06.006](https://doi.org/10.1016/j.neunet.2020.06.006).
- [103] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012. URL: <https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>.
- [104] T. N. Kipf and M. Welling. Semi-Supervised Classification with Graph Convolutional Networks, 2016. doi:[10.48550/ARXIV.1609.02907](https://doi.org/10.48550/ARXIV.1609.02907).
- [105] P. Veličković et al. Graph attention networks, 2017. doi:[10.48550/ARXIV.1710.10903](https://doi.org/10.48550/ARXIV.1710.10903).
- [106] M. Hammermann. Identification of inelastic interactions of light antinuclei in the Transition Radiation Detector in ALICE using Graph Neural Networks. Bachelor’s thesis, University of Heidelberg, 2022.
- [107] S. Agostinelli et al. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003. doi:[10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [108] C. Morris et al. Weisfeiler and leman go neural: Higher-order graph neural networks, 2018. doi:[10.48550/ARXIV.1810.02244](https://doi.org/10.48550/ARXIV.1810.02244).

Bibliography

- [109] F. Pedregosa et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011. doi:[10.48550/ARXIV.1201.0490](https://doi.org/10.48550/ARXIV.1201.0490).
- [110] A. Bórquez. private communication.

Acknowledgement

First of all, I would like to thank Prof. Dr. Johanna Stachel for the opportunity to work on my project in her group. I am very grateful to my supervisor Prof. Dr. Klaus Reygers for the support, the time and the interest he invested into my research and also for providing me a lot of freedom in my daily work such that I could focus on things I really liked to learn. I would like to thank Prof. Dr. Norbert Herrmann for acting as a referee for this thesis. I am also very grateful to Dr. Alexander Schmah for all his support, ideas and everything I learned from him. I would like to thank Dr. Marian Ivanov for everything he taught me about detector physics and also for the discussions with him about statistical methods. I would like to thank Max Hammermann for the great cooperation for parts of this work. Thank you to the whole ALICE@PI group. Last but not least, a huge thank you to my friends and family, without whom none of this would have been possible anyway.