

**Online selection of short-lived particles  
on many-core computer architectures  
in the CBM experiment at FAIR**

Dissertation  
for attaining the Ph.D. degree  
of Natural Science

submitted to the Faculty of Computer Science and Mathematics  
of the Johann Wolfgang Goethe University  
in Frankfurt am Main, Germany

by  
Maksym Zyzak  
from Yuzhnoukrainsk, Ukraine

Frankfurt am Main 2015  
(D 30)



accepted by the Faculty of Computer Science and Mathematics of the  
Johann Wolfgang Goethe University as a dissertation.

Dean: Prof. Dr. Uwe Brinkschulte

Expert assessor: Prof. Dr. Ivan Kisel  
Prof. Dr. Volker Lindenstruth

Date of disputation: 7 July 2016



# Abstract

Modern experiments in heavy ion collisions operate with huge data rates that can not be fully stored on the currently available storage devices. Therefore the data flow should be reduced by selecting those collisions that potentially carry the information of the physics interest. The future CBM experiment will have no simple criteria for selecting such collisions and requires the full online reconstruction of the collision topology including reconstruction of short-lived particles.

In this work the KF Particle Finder package for online reconstruction and selection of short-lived particles is proposed and developed. It reconstructs more than 70 decays, covering signals from all the physics cases of the CBM experiment: strange particles, strange resonances, hypernuclei, low mass vector mesons, charmonium, and open-charm particles.

The package is based on the Kalman filter method providing a full set of the particle parameters together with their errors including position, momentum, mass, energy, lifetime, etc. It shows a high quality of the reconstructed particles, high efficiencies, and high signal to background ratios.

The KF Particle Finder is extremely fast for achieving the reconstruction speed of 1.5 ms per minimum-bias AuAu collision at 25 AGeV beam energy on single CPU core. It is fully vectorized and parallelized and shows a strong linear scalability on the many-core architectures of up to 80 cores. It also scales within the First Level Event Selection package on the many-core clusters up to 3200 cores.

The developed KF Particle Finder package is a universal platform for short-lived particle reconstruction, physics analysis and online selection.

# Kurzfassung

Moderne Experimente der Hochenergiephysik mit Schwerionenstrahlen müssen so hohe Datenraten verarbeiten, dass die Daten nicht mehr vollständig mit den derzeit verfügbaren Speicher-Technologien aufgezeichnet werden können. Es ist daher erforderlich, den Datenfluss auf die potentiell interessanten physikalischen Ereignisse zu reduzieren. Dem zukünftigen CBM Experiment stehen keine einfachen Kriterien zur Auswahl solcher Ereignisse zur Verfügung, da viele der gesuchten Observablen sehr selten erzeugt werden und in einen hohen Untergrund produzierter Teilchen eingebettet sind. Die Strategie zur Selektion bestimmter Ereignisse beruht daher auf einer vollständigen Rekonstruktion der Kollisionstopologie in Echtzeit einschließlich der Rekonstruktion instabiler Teilchen.

Im Rahmen dieser Arbeit wurde das KF Particle Finder Paket entwickelt zur Online-Rekonstruktion von Schwerionenreaktionen mit hohen Multiplizitäten inklusive der Selektion kurzlebiger Teilchen. Das Paket rekonstruiert über 70 Zerfälle von seltsamen Teilchen, seltsamen Resonanzen, Hyperkernen, leichte Vektormesonen, Charmonium und Open-Charm-Teilchen, alles Signale, die sensitive sind auf die Eigenschaften dichter Kernmaterie, wie sie in Schwerionenreaktionen erzeugt wird.

Das Paket basiert auf der Kalman-Filter-Methode, die einen vollständigen Satz der Teilchenparameter und ihrer Fehler liefert, inklusive Position, Impuls, Masse, Energie, Lebensdauer usw. Das Paket rekonstruiert die Teilchen mit hoher Qualität, mit hoher Effizienz und einem großen Signal-zu-Untergrund-Verhältnis.

Das KF Particle Finder Paket ist extrem schnell und erreicht Rekonstruktionsgeschwindigkeiten von 1,5 ms pro Kollision zweier Goldkerne bei einer Strahlenergie von 25 AGeV auf nur einem CPU-Core. Das Paket ist vollständig vektorisiert und parallelisiert, die Anzahl der verarbeiteten Ereignisse skaliert mit der Anzahl der Cores auf "many-core"-Rechnerarchitekturen bis zu 80 Cores. Die Skalierbarkeit ist auch erfüllt, wenn das Paket innerhalb des First-level Event Selector Pakets auf many-core-Clustern bis zu 3200 Cores benutzt wird.

Das KF Particle Finder Paket bietet eine universelle Plattform zur Rekonstruktion kurzlebiger Teilchen, zur Physikanalyse und zur Online-Selektion von Ereignissen.

# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>The CBM experiment</b>	<b>10</b>
2.1	The future FAIR facility . . . . .	11
2.2	Physics program of CBM . . . . .	12
2.3	CBM detector system . . . . .	13
2.4	First Level Event Selection (FLES) . . . . .	23
<b>3</b>	<b>Many-core computer architectures and tools for parallel programming</b>	<b>26</b>
3.1	Topology of CPU . . . . .	27
3.2	Intel Xeon Phi coprocessor . . . . .	30
3.3	GPU architecture . . . . .	31
3.4	Approaches for parallel programming . . . . .	33
3.5	Tools for vectorization . . . . .	35
3.6	Tools for parallelisation . . . . .	37
3.7	Universal tools for vectorization and parallelization . . . . .	40
<b>4</b>	<b>Reconstruction of charged particles trajectories</b>	<b>43</b>
4.1	The Kalman filter method . . . . .	44
4.2	Kalman filter for track parameters estimation . . . . .	49
4.3	The matrix inversion problem . . . . .	54
4.4	Kalman filter based smoother . . . . .	55
4.5	Deterministic Annealing Filter . . . . .	58
4.6	KF track fit on many-core systems . . . . .	60

<b>5</b>	<b>Reconstruction of short-lived particles with KF Particle package</b>	<b>63</b>
5.1	The KF Particle package for the short-lived particle reconstruction	65
5.2	Functionality of KF Particle . . . . .	68
5.3	Development of KF Particle . . . . .	69
5.3.1	Point of the closest approach between a particle and a vertex	70
5.3.2	Point of the closest approach between two particles . . . . .	73
5.3.3	Nonlinear mass constraint . . . . .	80
<b>6</b>	<b>KF Particle Finder — a common platform for online selection and physics analysis</b>	<b>86</b>
6.1	Search of short-lived particles . . . . .	87
6.2	Physics analysis with KF Particle Finder . . . . .	93
6.2.1	Strange and multi-strange particles . . . . .	95
6.2.2	Strange and multi-strange resonances . . . . .	98
6.2.3	Hypernuclei and heavy multi-strange objects . . . . .	100
6.2.4	Open charm particles . . . . .	102
6.2.5	Low mass vector mesons . . . . .	104
6.2.6	Charmonium . . . . .	106
6.2.7	$\gamma$ conversion and $\pi^0$ . . . . .	108
6.3	Optimization of the package with respect to speed . . . . .	110
6.4	KF Particle Finder on many-core systems . . . . .	111
6.5	KF Particle Finder in the FLES package . . . . .	112
<b>7</b>	<b>Summary</b>	<b>116</b>
<b>A</b>	<b>Properties of the Kalman filter method on an example of the straight line</b>	<b>119</b>
	<b>Bibliography</b>	<b>142</b>
	<b>Zusammenfassung</b>	<b>153</b>

# Chapter 1

## Introduction

The structure of matter was one of the most exciting questions since elementary particles were discovered in the beginning of the last century. What are the smallest indivisible structure elements? How do they interact with each other? Digging deeper and deeper into this problem scientists discovered that atom consists of a nucleus and electrons. Then a large number of subatomic particles were discovered including nucleons, hyperons and mesons. By the middle of the 20<sup>th</sup> century there were found so many of such particles, that the need of their systematization appeared. In their works Gell-Mann [1, 2] and Zweig [3] introduced quarks for systematization and classification of hadrons.

According to the present knowledge there are six quarks (see Fig. 1.1), which are the smallest strongly interacting particles: up (u), down (d), strange (s), charm (c), bottom (b) and top (t). Together with six leptons they form a list of fundamental spin  $\frac{1}{2}$  particles — fermions. There are three charged leptons: electron (e), muon ( $\mu$ ) and taon ( $\tau$ ). Each charged lepton has a corresponding flavour of a neutral lepton — neutrino ( $\nu$ ). All elementary particles have antiparticle partners. Antiparticle has the same mass and spin as a particle, but opposite attributes such as electric charge or color charge. Within the Standard Model of particle physics [4] the whole variety of existing matter is build from this small list of fermions.

The Standard Model also comprises interaction between particles. There are four basic types of interaction or field: strong, weak, electromagnetic and gravitational. Only quarks interact strongly, while other interactions are common

	three generations of matter			gauge bosons	Higgs boson
quarks	<b>u</b> M = 0.0023 GeV/c <sup>2</sup> q = 2/3	<b>c</b> M = 1.275 GeV/c <sup>2</sup> q = 2/3	<b>t</b> M = 160 GeV/c <sup>2</sup> q = 2/3	<b>g</b> M = 0 q = 0	<b>H</b> M = 125.7 GeV/c <sup>2</sup> q = 0
	<b>d</b> M = 0.0048 GeV/c <sup>2</sup> q = -1/3	<b>s</b> M = 0.095 GeV/c <sup>2</sup> q = -1/3	<b>b</b> M = 4.18 GeV/c <sup>2</sup> q = -1/3	<b>γ</b> M = 0 q = 0	
	<b>e</b> M = 0.511 MeV/c <sup>2</sup> q = -1	<b>μ</b> M = 105.6 MeV/c <sup>2</sup> q = -1	<b>τ</b> M = 1776.8 MeV/c <sup>2</sup> q = -1	<b>Z</b> M = 91.2 GeV/c <sup>2</sup> q = 0	
leptons	<b>ν<sub>e</sub></b> M < 2·10 <sup>-6</sup> MeV/c <sup>2</sup> q = 0	<b>ν<sub>μ</sub></b> M < 190·10 <sup>-3</sup> MeV/c <sup>2</sup> q = 0	<b>ν<sub>τ</sub></b> M < 18.2 MeV/c <sup>2</sup> q = 0	<b>W<sup>±</sup></b> M = 80.4 GeV/c <sup>2</sup> q = ±1	

Figure 1.1: Schematic depiction of particles and boson mediators in the Standard model of particle physics. Parameters of the particles are taken from PDG 2014 [5]. Charge is given in terms of the elementary charge  $e$ .

for all types of particles. In the Standard Model the interaction is mediated by bosons. The strong interaction is mediated by eight massless gluons ( $g$ ), weak — by the  $W^\pm$  and  $Z$  bosons, electromagnetic — by the photon ( $\gamma$ ). The gravitational interaction is the weakest in scales of the particle physics and usually can be neglected. It is supposedly mediated by spin 2 boson called graviton, however gravitons are not yet observed. The mass of the particles is generated by the scalar field of Higgs bosons. The scheme of the particles and bosons of the Standard Model is illustrated in Fig. 1.1.

At normal nuclear densities and temperatures quarks together with gluons (bosons carrying the strong interaction) are bound into hadrons: baryons, including proton and neutron, that can be thought of as composed of three quarks, and mesons, which contain a quark and an antiquark. Being constructed of quarks, groups of hadrons form multiplets [6], that manifest symmetries of the Quantum Chromodynamics (QCD) [7] — the theory of strong interactions. For example, the lowest  $SU(3)$  baryon and meson multiplets together with the quark structure of the corresponding particles are shown in Fig. 1.2. The approximate

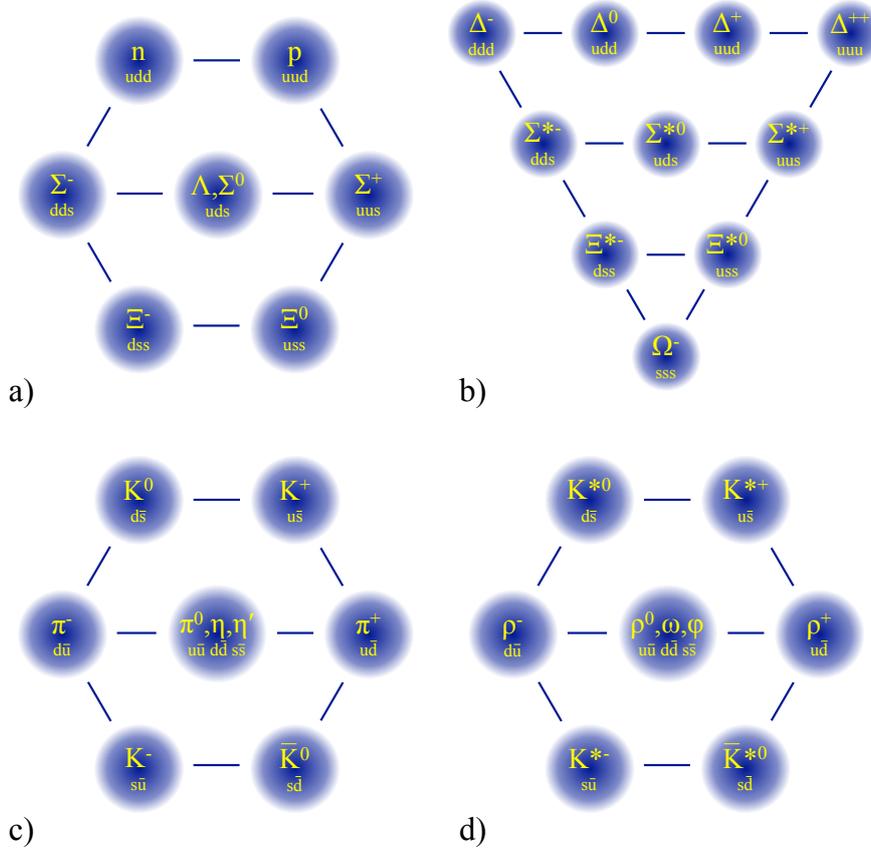


Figure 1.2: The lowest SU(3) baryon and meson multiplets together with the quark structure of particles: the baryon octet (a) with spin  $S = \frac{1}{2}$ , the baryon decuplet (b) with  $S = \frac{3}{2}$ , the pseudo-scalar nonet (c) with  $S = 0$  and the vector nonet (d) with  $S = 1$ .

SU(3) symmetry of these multiplets is the manifestation of the approximate flavor symmetry between the up, down and strange quarks.

According to the Standard Model at normal conditions quarks and gluons cannot exist as free particles. This effect is known as confinement [8]. At higher temperatures and densities according to the QCD they can form completely new states of matter like Quark-Gluon Plasma (QGP), when the quarks and gluons are deconfined and can move freely in the region, where QGP exists, on the scales larger than the size of hadrons (about  $10^{-15}$ m). It is believed that the matter existed in such a state at the first stages of the universe formation. The nuclear matter under the extreme pressures and temperatures may exist also nowadays in the interior of compact stellar objects like, for instance, neutron stars.

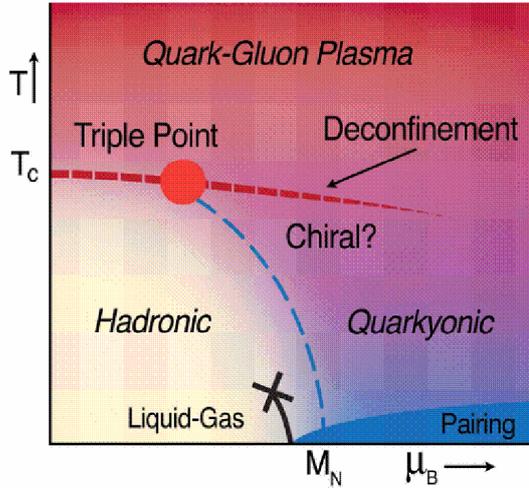


Figure 1.3: A possible form of the QCD phase diagram [10].

The QCD equation of state is dominated by rather different degrees of freedom depending on the values for the temperature and the chemical potentials, that makes a structure of the QCD phase diagram quite complicated [9, 10]. A possible sketch of the QCD phase diagram of temperature ( $T$ ) versus baryon chemical potential ( $\mu_B$ ) is shown in Fig. 1.3. At low densities, the nucleons (i.e. protons and neutrons) behave like a gas.

With increasing temperature and density the nucleons are excited into baryon resonances, which subsequently decay into pions and nucleons. This mixture of nucleons, baryonic resonances and mesons is called hadronic matter (represented by the yellow area in Fig. 1.3). In this region of the phase diagram hadrons are the relevant degrees of freedom. At very high temperatures the hadrons melt into quarks and gluons forming QGP (shown in red). In QGP degrees of freedom are associated with quarks and gluons. In the region of small baryon chemical potential lattice QCD calculations predict a smooth cross-over between these two phases [11, 12]. According to the present-day understanding at large baryon chemical potentials a first-order phase transition takes place [13]. The phase transition line starts with a second order phase transition point — critical endpoint. Current estimation for the critical temperature is 155 – 165 MeV [14, 15]. In highly compressed cold nuclear matter (shown in blue), as it may exist in the interior of neutron stars, the baryons also dissolve into quarks and gluons. The critical density at which this transition occurs, however, is not known.

The only way to investigate nuclear matter at high pressures and temperatures in the laboratory are experiments studying collisions of heavy ions. In order to explore different regions of the QCD phase diagram several experiments were started since 1980s (see Table 1.1): from low energetic collisions of few AGeV

(GeV per nucleon) at AGS<sup>1</sup> [16] and SIS<sup>2</sup> [17] accelerators up to hundreds AGeV at RHIC<sup>3</sup> [18] and even ATeV at LHC<sup>4</sup> [19]. To study the intermediate energy region with a rich structure of the phase diagram, as it is now predicted, new FAIR<sup>5</sup> [20] and NICA<sup>6</sup> [21] facilities are under development.

Accelerator	Laboratory	Type of experiments	Ions (heaviest)	Top energy
SIS 18	GSI, Germany	fixed target	U	2 AGeV
AGS	BNL, USA	fixed target	Au	14.5 AGeV
SIS 100*	FAIR, Germany	fixed target	U	11 AGeV
SIS 300*	FAIR, Germany	fixed target	U	35 AGeV
NICA*	JINR, Russia	collider	Au	11 AGeV
SPS <sup>7</sup>	CERN, Switzerland	fixed target	Pb	158 AGeV
RHIC	BNL, USA	collider	U	200 AGeV
LHC	CERN, Switzerland	collider	Pb	5.52 ATeV

Table 1.1: A list of already running and future (marked with \*) accelerators for studying heavy ion collisions. For accelerators with the fixed target experiments the energy is given in the laboratory system, for collider type — center-of-mass collision energy is given.

Searching for rare signals modern experiments in physics of heavy ion collisions have to process petabytes of data. Thus the dedicated computer farms equipped with the latest processor architectures are used by each experiment. For example, the CBM experiment will operate with a farm, which is currently estimated to have a compute power equivalent to 60 000 modern CPU cores. Since more than 10 years the microprocessor industry tends to introduce technologies for parallel computation rather than increasing computational speed of each unit. Therefore reconstruction algorithms should be parallel and should scale with the increasing number of compute elements.

Since short-lived particles are one of the most important probes for analyzing properties of the created hot and dense matter, the present thesis is aimed to develop a method for short-lived particle reconstruction and selection, that can utilize the whole potential of modern many-core computer architectures. This

<sup>1</sup>Alternating Gradient Synchrotron, Brookhaven National Laboratory (BNL), USA

<sup>2</sup>SchwerIonenSynchrotron, GSI Helmholtzzentrum für Schwerionenforschung, Germany

<sup>3</sup>Relativistic Heavy Ion Collider, BNL, USA

<sup>4</sup>Large Hadron Collider, Conseil Européen pour la Recherche Nucléaire (CERN), Switzerland

<sup>5</sup>Facility for Antiproton and Ion Research, GSI, Germany

<sup>6</sup>Nuclotron-based Ion Collider fAcility, Joint Institute for Nuclear Research (JINR), Russia

<sup>7</sup>Super Proton Synchrotron, CERN, Switzerland

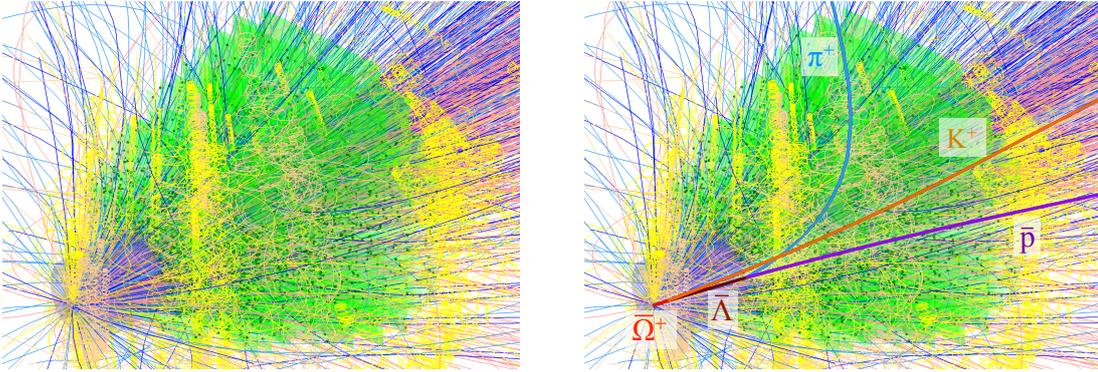
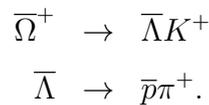


Figure 1.4: The simulated central Au-Au collision at 25 AGeV energy in the CBM experiment with about 1000 charged particles (left) with a highlighted  $\bar{\Omega}^+$  decay (right). Different colors correspond to different types of particles.  $\bar{\Omega}^+$  decays into  $\bar{\Lambda}$  and  $K^+$  with further decay of  $\bar{\Lambda}$  to antiproton ( $\bar{p}$ ) and  $\pi^+$ . One of such decay is expected per about 1000 central collisions.

method is developing to be run for online event (collision) reconstruction and selection in the future Compressed Baryonic Matter (CBM) experiment [22] at the FAIR facility. Fig. 1.4 illustrates a simulated Au-Au collision at 25 AGeV energy in the CBM experiment with a highlighted  $\bar{\Omega}^+$  decay:



The method should reconstruct and recognize a one signal particle per thousands or even millions of such collisions (see Table 1.2), therefore both speed and efficiency are crucial and can not be compromised. Being intrinsically parallel tasks of particle search and reconstruction perfectly suite modern computer architectures. The developed method is based on the Kalman filter [23, 24] because of it's properties of high speed and mathematical correctness of the obtained results.

The thesis is structured as follows. Chapter 2 describes the CBM experiment, its physics goals, challenges and detector systems. Modern architectures of microprocessors are discussed in Chapter 3. Methods and tools for utilization of their advantages are also explained there together with illustrative examples. Application of the Kalman filter method to reconstruction of charged particles trajectories is given in Chapter 4. Being the key method for the current work

Particle	$\Omega^-$	$\bar{\Omega}^+$	$D^0$	$\bar{D}^0$	$\rho$	$J/\psi$
Decay channel	$\Lambda K^-$	$\bar{\Lambda} K^+$	$K^- \pi^+$	$K^+ \pi^-$	$e^+ e^-$	$e^+ e^-$
Multiplicity per collision	$4.7 \cdot 10^{-2}$	$3.5 \cdot 10^{-3}$	$3.8 \cdot 10^{-5}$	$1.2 \cdot 10^{-4}$	23	$1.9 \cdot 10^{-5}$
Branching ratio	67.8%	67.8%	3.9%	3.9%	$4.7 \cdot 10^{-5}$	5.97%

Table 1.2: Multiplicities (by HSD model [25]) and branching ratios [5] of the CBM observables on example of  $\Omega^-$  and  $\bar{\Omega}^+$  for multi-strange hyperons,  $D^0$  and  $\bar{D}^0$  for open charm,  $\rho$  for low mass vector mesons and  $J/\psi$  in a central Au-Au collision at 25 AGeV energy. On top of the low multiplicity and branching ratio the efficiency of reconstruction for such signals is typically on the few percent level.

the Kalman filter and its main properties are also described there together with implementation of the algorithm for the CBM experiment and its modifications. Chapter 5 describes the KF Particle package [26, 27] for short-lived particle reconstruction and its modifications and improvements. Based on KF Particle the KF Particle Finder package for online reconstruction and selection was developed. It is described in Chapter 6 together with its implementation for many-core computer architectures. The capabilities of the package for the physics analysis together with obtained results are discussed here as well. Chapter 7 summarizes the work of the thesis.

# Chapter 2

## The CBM experiment

One of the main goals of heavy ions experiments is exploration of the QCD phase diagram (see Fig. 1.3). Present accelerators and experiments mainly cover low (SIS18) and high (RHIC, LHC) energy regions. At low energies the QGP phase can not be formed and only the region of hadron gas can be studied. In collisions at very high energies large number of particles is produced and they are balanced with respect to their matter-antimatter content, the baryon chemical potential is relatively small in such matter. Therefore such experiments explore the baryon-poor but very hot region of the phase diagram, where a cross-over transition from quark-gluon plasma to hadron gas is expected.

To explore the QCD phase diagram in the region of the expected first-order phase transition and the critical endpoint collisions, that leads to a much higher baryon chemical potential at relatively moderate temperatures, should be studied. Several experimental programs are devoted to the exploration of this region: the beam energy scan (BES) at RHIC [28], the upgraded NA49 detector (NA61) at SPS using light and medium size ion beams [29] and the future dedicated NICA facility [21] in JINR.

However, due to luminosity or detector limitations these experiments can not measure rare observables with very low production cross sections and are constrained to the investigation of abundantly produced particles. In contrast, the Compressed Baryonic Matter (CBM) experiment at the Facility for Antiproton and Ion Research (FAIR) [20] in Darmstadt is designed for precision measurements of much wider set of observables including extremely rare charmed particles

together with the comprehensive study of their properties.

## 2.1 The future FAIR facility

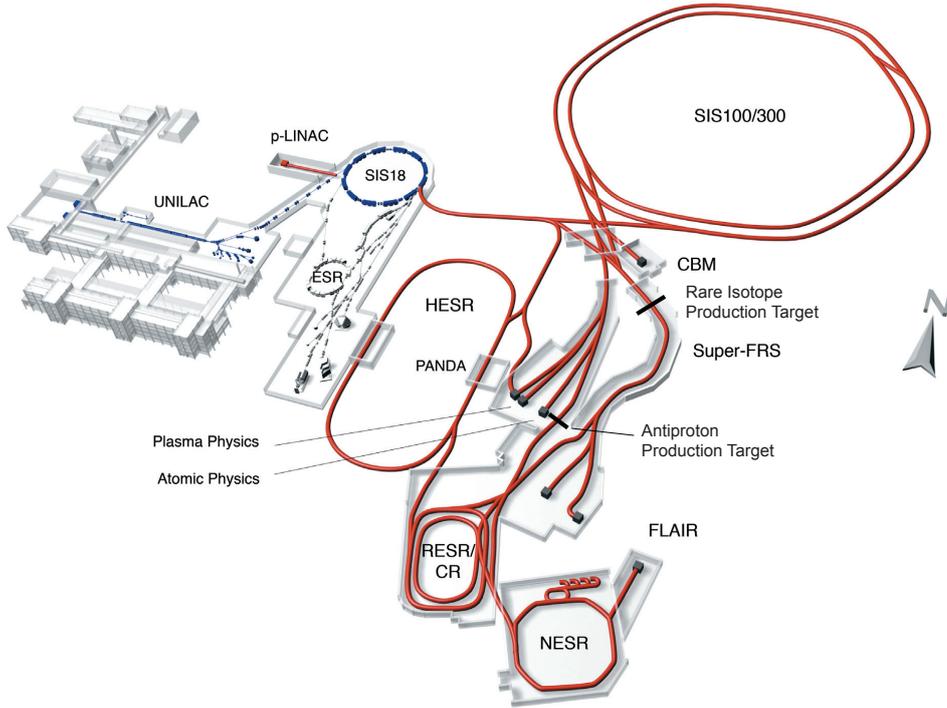


Figure 2.1: Layout of the Facility for Antiproton and Ion Research (FAIR) [20]. Blue lines represents existing GSI facilities, the new accelerator complex is shown with red. The superconducting synchrotrons SIS100 or SIS300 will provide ion beams for the CBM experiment. They also will provide secondary beams of antiprotons or rare nuclear isotopes from a production target to the fragment separator (super FRS) and storage rings, such as the HESR for the antiproton research program.

The future FAIR facility will provide unique research opportunities in the fields of nuclear, hadron, atomic and plasma physics [20]. The layout of FAIR is presented in Fig. 2.1. Ion beams for the CBM experiment will be delivered in the beginning by the SIS100 synchrotron and later, after upgrade of magnetic rigidity from 100 to 300 Tm, by SIS300. The minimal available ion beam energy will be about 2 AGeV. The top energies of the SIS100 synchrotron will be up to 29 GeV for protons and up to 11 AGeV for Au ions. In case of the SIS300 synchrotron the top energy is expected to be up to 90 GeV for protons and 35 AGeV for Au

ions. Due to the slow extraction from the synchrotrons the beam at the CBM cave will have a continuous structure, that allows to increase the collision rate since the beam particles will be uniformly distributed in time. The intensities of the beam are expected to reach  $10^9$  Au ions per second.

## 2.2 Physics program of CBM

The physics program of the CBM experiment assumes a comprehensive study of the phase diagram of strongly interacting matter and a study of the equation of state of the matter at extremely high baryonic density [22]. The research program comprises a complex scan of observables, beam energies and collision systems.

At the first stage the experiment will operate with the beams provided by the SIS100 accelerator, later upgraded to SIS300. The experiment will address the following fundamental questions [22, 30]:

1. phase transition at high baryon chemical potential and search for the critical endpoint;
2. investigation of the equation of state of nuclear matter at densities similar to the densities in the core of neutron stars and determination of the degrees of freedom at these densities;
3. investigation of hypernuclei and heavy multi-strange objects, if such objects exist;
4. investigation of the properties of hadrons in dense baryonic matter and the possible modification of their properties;
5. investigation of charm production at threshold beam energies and investigation of the charm properties in dense baryonic matter.

The most promising observables to answer these questions are:

1. particles containing strange or charm quarks (for example, strange hyperons —  $\Lambda$ ,  $\Xi$ ,  $\Omega$ , charmed D mesons,  $J/\psi$ );
2. low mass vector mesons decaying into dilepton channel ( $\rho, \omega, \phi \rightarrow e^+e^-$ ,  $\rho, \omega, \phi \rightarrow \mu^+\mu^-$ );
3. the excitation functions of yields, spectra, and collective flow of these particles;

4. the in-medium mass distribution of low mass vector mesons;
5. event-by-event fluctuations of conserved quantities like baryons, strangeness, net-charge etc. in heavy-ion collisions with high precision as a function of beam energy.

The experimental goal is to measure these rare probes with unprecedented precision in spite of the very low particle yields.

## 2.3 CBM detector system

The CBM experiment is being designed for studying multiplicity, distribution in phase space and flows of all interesting particles, including very rare particles. The experiment requires extreme rates of the beam collisions with the target (up to  $10^7$  Hz) with a multiplicity of produced charged particles up to 1000 per collision. Under these conditions it is impossible to use slow detectors, such as, for instance, Time-Projection Chamber (TPC). The experiment requires from the detectors to be fast and radiation hard. Also the research program of CBM requires from the detector systems to provide efficient lepton identification, a high-resolution determination of decay vertices and high-speed trigger and data acquisition systems.

The experiment assumes two options of the operating scenario:

- with a setup for electron measurements;
- with a setup for muon measurements.

The electron and muon setups are shown in Fig. 2.2. Such approach combines the advantages of both methods and guarantees reliable results as in the end the data sets should agree to each other inspite of the very different background sources. The full list of sub-detector systems for both setups is:

1. Micro-Vertex Detector (MVD);
2. Silicon Tracking System (STS);
3. Ring Imaging CHerenkov detector (RICH);
4. Muon Chambers (MuCh);
5. Transition Radiation Detector (TRD);

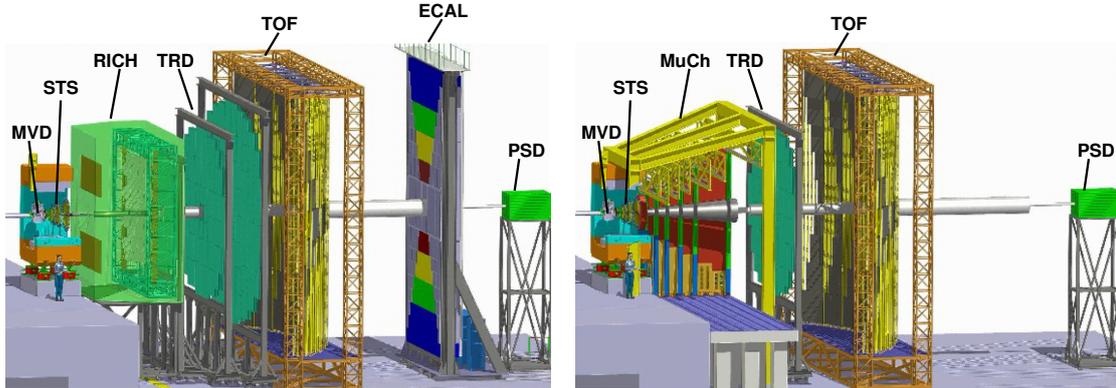


Figure 2.2: The layout of the CBM detector setup [22]. Options for electron measurements (left) and muon measurements (right) are presented. The electron setup includes the dipole magnet with a large acceptance, the Micro-Vertex Detector (MVD), the Silicon Tracking System (STS), the Ring Imaging CHerenkov detector (RICH), the Transition Radiation Detector (TRD), the Time Of Flight detector (TOF), the Electromagnetic CALorimeter (ECAL) and the Projectile Spectator Detector (PSD). In case of the muon setup RICH, TRD and ECAL are replaced with the Muon Chambers (MuCh).

6. Time Of Flight detector (TOF);
7. Electromagnetic CALorimeter (ECAL);
8. Projectile Spectator Detector (PSD).

The sort of a particle	MVD	STS	RICH	MUCH	TRD	TOF	ECAL	PSD
$\pi, K, p$		x	(x)		(x)	x		x
Hyperons		x			(x)	(x)		x
Open charm	x	x	(x)		(x)	x		x
electrons	x	x	x		x	x		x
muons		x		x		(x)		x
gamma							x	x
gamma through $e^\pm$ conversion	x	x	x		x	x		x

Table 2.1: Observables and detectors for their registration. Detectors marked as x are those, which are necessary for particle registration. Detectors marked as (x) can be used to suppress background.

According to the tasks, the electron setup includes all the detectors except MuCh. The muon option assumes operation of the experiment with all detectors except ECAL and RICH detectors, which are devoted for electron identification. Each setup will use different detectors for registration of certain particle types.

Table 2.1 summarizes which detectors will be used for registration of different particles.

### 2.3.1 Micro-Vertex Detector (MVD)

To achieve a track position resolution of several microns in the target region the CBM experiment will use the MVD detector [22, 31]. It will be located at a distance from 5 to 20 cm downstream the target. The main task of the detector is to provide a high separation level of short-lived particles from the collision point, for instance, particles containing charm quark with lifetime of about 50–300  $\mu\text{m}/c$ . Also, MVD will help to reduce background in dielectron decay channels of low mass vector mesons.

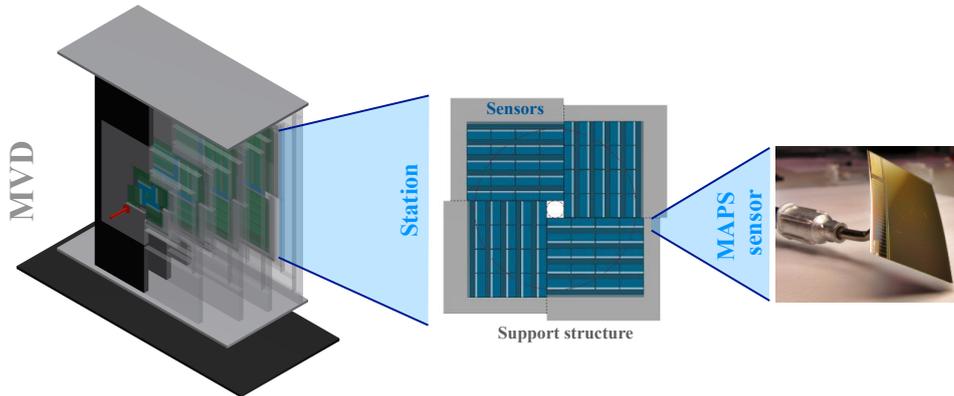


Figure 2.3: The MVD detector of the CBM experiment [31, 32]. MVD will be built from MAPS sensors grouped into 2–4 stations. To reduce the amount of material MVD will be placed in vacuum. The detector will be mounted inside the dipole magnet.

The detector will be built from thin pixel detector modules based on the MAPS (Monolithic Active Pixel Sensors) technology. Pixel size will be between  $25 \times 25 \mu\text{m}^2$  to  $40 \times 40 \mu\text{m}^2$ . MAPS technology allows to obtain very thin detectors with high spatial resolution (few microns). It is expected that the detector will be composed of 2 to 4 stations. Thickness of stations will be several hundred microns. To reduce the amount of material in the active volume, the detector will be placed in vacuum. Fig. 2.3 illustrates the detector geometry option with four stations located at a distance of 5, 10, 15 and 20 cm from the target.

### 2.3.2 Silicon Tracking System (STS)

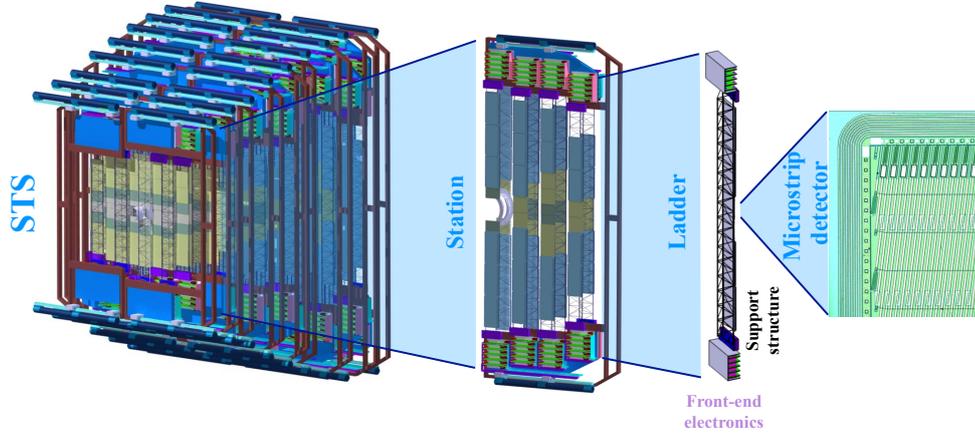


Figure 2.4: The STS detector of the CBM experiment [33]. STS will be build from silicon microstrip detector modules grouped into lightweight vertical ladders with carbon fibers as the support structure. Ladders will form eight stations. The STS detector will be mounted inside the dipole magnet.

The main tracking detector of the CBM experiment is the Silicon Tracking System (STS) [33]. The task of the STS detector is reconstruction of the particle trajectory and determination of the particle momentum. The detector will be placed in the magnetic field produced by the dipole magnet with 1 Tm bending power [34]. In the magnetic field trajectories of charged particles are bended and thus the momentum can be reconstructed. For successful reconstruction of particles of the physics interest, relative momentum resolution  $\Delta p/p = 1\%$  is required. Such resolution can be achieved only in case of thin detector modules with low multiple scattering of charged particles. Therefore the detector is planned to be build up from the silicon micro-strip detector modules on lightweight ladder-like mechanical supports. The detector modules will be grouped into eight stations placed in the direction of the beam on distances from 30 to 100 cm from the target (see Fig. 2.4). To minimize the amount of material in the active volume, the electronics of the detector will be placed at the periphery of the stations. Signals from detector modules will be transported to the electronics through multi-line micro-cables. The STS detector will work with loads up to 700 charged particles per collision and in order to achieve the designed momentum resolution the modules are being designed to have following parameters:

- length of strips — from 2 to 6 cm;
- thickness of the module — 300  $\mu\text{m}$ ;
- a stereo angle between strips — from  $7^\circ$  to  $15^\circ$ ;
- a strip pitch — about 58  $\mu\text{m}$ ;
- total amount of material together with cables — 400–800  $\mu\text{m}$  in the silicon equivalent.

### 2.3.3 Ring Imaging CHerenkov detector (RICH)

The main task of the RICH detector [35] is identification of electrons and reduction of the number of background pions in the reconstruction of short-lived particles decaying by the dielectron channel. The idea of RICH is based on detection of the radiation emitted by a particle moving in a medium with a speed greater than the speed of light in that medium. This effect is called Cherenkov effect. The photons are radiated at a fixed angle, which is determined by the properties of the medium and the speed of the particle. Thus, these photons form a cone, which gives a ring on the photodetector plane (see Fig. 2.5, right). The emission angle of the cone is determined by the speed of a particle. Thus, having the same momentum particles with different masses forms ring with different radius.

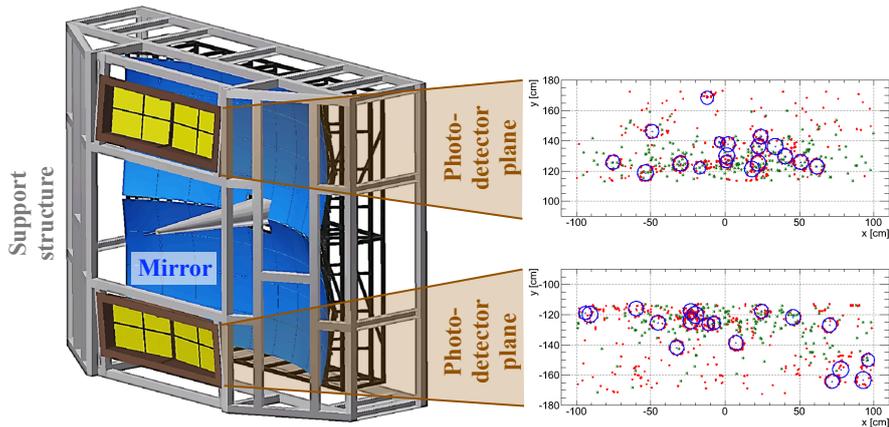


Figure 2.5: The RICH geometry including mirrors, two photodetector planes and the support structure (left) and rings produced in a simulated AuAu collision at 25 AGeV by the Cherenkov light at the photodetector planes (right) [35].

The proposed design for the RICH detector of CBM includes (Fig. 2.5, left)

1.7 m container filled with  $\text{CO}_2$ , glass mirrors of 6 mm thickness with a reflective  $\text{MgF}_2$  coating and two planes of photodetectors. With such parameters it will be possible to separate electrons from other particles in the range of momenta up to 10 GeV/c. Having the smallest mass all electrons, even with low momenta, produced in the collision emit the Cherenkov light, while pions start to irradiate at about 5 GeV/c momenta. The designed level of the pion suppression factor of about 500 is foreseen.

### 2.3.4 Muon Chambers (MuCh) detector

One of the focuses of the CBM experiment is measurement of rare particles decaying into muons like low mass vector mesons and  $J/\psi$ . Reconstruction of such particles requires an accurate muon identification. Being minimum ionizing particles in a wide momentum range from 100 MeV/c to 100 GeV/c (see Fig. 2.6), muons can penetrate several meters of iron unlike all other particles: as electrons, they do not interact strongly but have large enough mass not to be significantly scattered by the electromagnetic field inside the material.

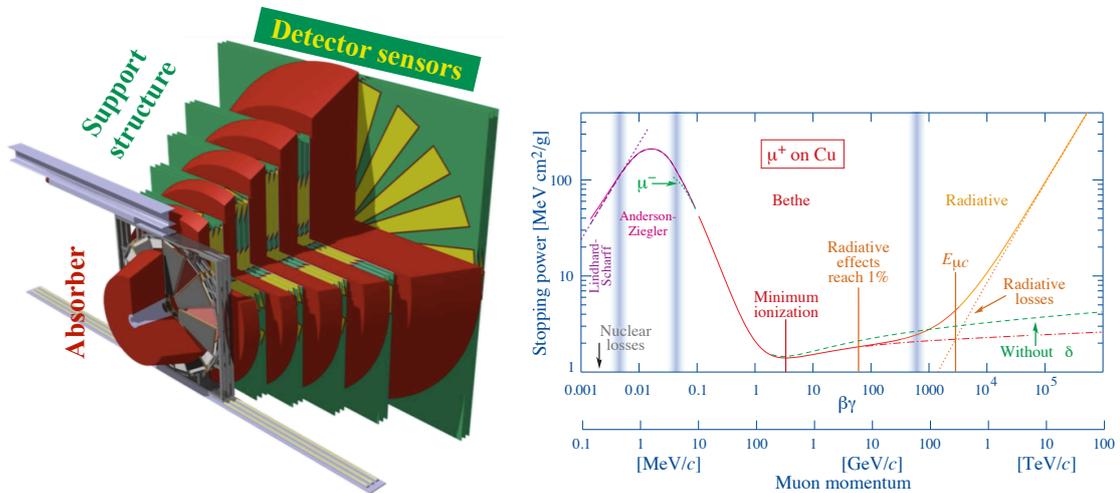


Figure 2.6: The MuCh geometry (left) including absorbers (red) and gaseous detectors of charged particles (yellow) together with the support structure (green) [36] and muon stopping power depending on the momentum (right) [5].

The MuCh detector will be placed just after the STS tracking system. The experimental challenge for muon measurements in heavy ion experiments at the CBM energies is identification of low-energy muons at high multiplicities of

charged particles. To cope with this problem the muon detector of the CBM experiment [36] will have the hadron absorber divided into several parts, each part of the absorber will be followed by three detecting stations for charged particles registration (see Fig. 2.6, left). The first absorber is planned to be build from carbon, all others — from iron. To reduce the muon background that is formed due to decays of mesons by the muon channel, the system must be compact. The detector consists of six hadron absorbers with total thickness of about 250 cm. For registration of charged particles gaseous tracking chambers located in triplets behind each iron slab will be used.

### 2.3.5 Transition Radiation Detector (TRD)

The TRD detector is based on the effect, that a charged particle emits the radiation when passing through a boundary between two different media. This happens because of the different electromagnetic properties of the media. The intensity of the radiation depends on the energy of the particle and its type.

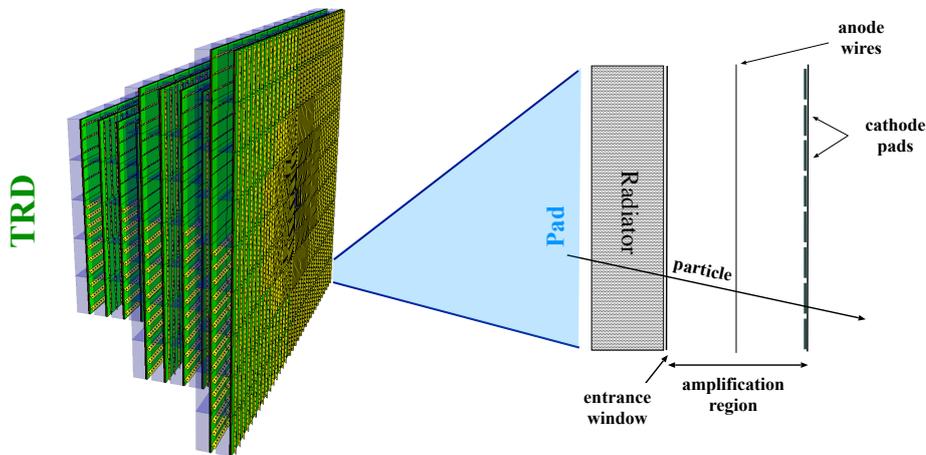


Figure 2.7: The TRD geometry consisting of rectangular pads for transition radiation detection and particles tracking [31] (left) and the scheme of a pad (right) [37]. When crossing the radiator of a pad, particle produces transition radiation, which is registered by the gaseous detector.

TRD will be used to identify electrons with momentum larger than 1.5 GeV/c and separate electron and pion tracks. This will help to reduce the background of short-lived particles, such as low mass vector mesons and charmonium, decaying

by the electron channel.

The TRD detector will consist of three stations located at distances from 5 m to 9.5 m from the target. The stations will consist of three layers each. The detector is divided into rectangular modules (pads) with a resolution of 300–500  $\mu\text{m}$  across and 3–30 mm along the pad [22]. To ensure an accurate measurement of both coordinates every second layer of the sensors will be rotated by  $90^\circ$  with respect to the previous one. Such structure of the detector will also allow reconstruction of charged particles tracks and extrapolation of them to the downstream detectors TOF and ECAL.

### 2.3.6 Time Of Flight detector (TOF)

In order to identify particles of different types we can use the fact, that having the same momentum particles with different masses have different speed. Therefore the mass of a particle can be determined by measuring speed together with momentum and, as a result, particle can be identified.

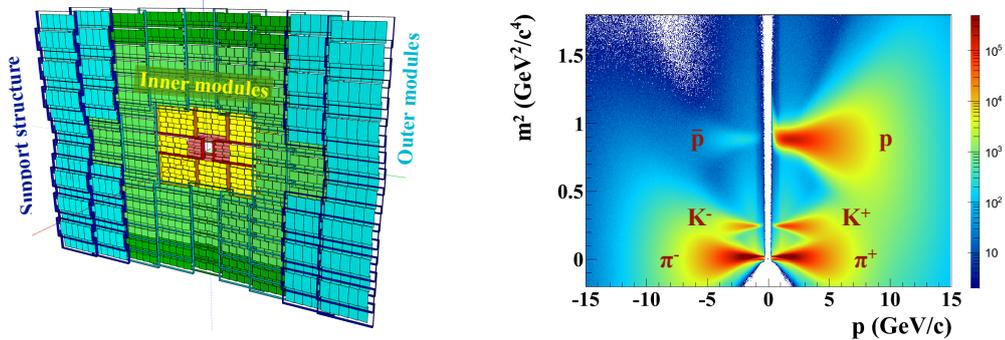


Figure 2.8: The TOF geometry (left) including gaseous detectors, grouped into modules, and the support structure [38] with (right) the distribution of the mass squared ( $m^2$ ) versus momentum obtained for positively charged hadrons — pions ( $\pi^+$ ), kaons ( $K^+$ ) and protons ( $p$ ) and their antiparticles; negative sign of momentum stands for negative particles [39].

The main task of the TOF detector [38] is identification of hadrons: pions, kaons and protons. The geometry of the detector together with its performance on the simulated data are shown in Fig. 2.8. TOF will be built from the Resistive Plate Chambers (RPC). RPCs consist of parallel plates made of the high resistivity material and separated by a gas volume. A high voltage is applied to the

external surfaces of the plates. A charged particle ionizes the gas and the electric field amplifies this ionization by an electron avalanche. The avalanche is stopped by the resistive plates and induces a fast signal on the pickup electrodes.

The detector will be placed on the distance of 10 m from the target. The active area of the detector will be about 120 m<sup>2</sup>. The time resolution will be better than 80 ps.

### 2.3.7 Electromagnetic CALorimeter (ECAL)

The task of ECAL in CBM is measuring of fast primary photons and neutral mesons decaying by a photon channel. Interacting with the material of the calorimeter, a photon produces a particle shower, which consists of electrons, positrons and secondary photons (see Fig. 2.9, right). In such a way the energy of the initial photon is deposited in ECAL, where it is collected and measured.

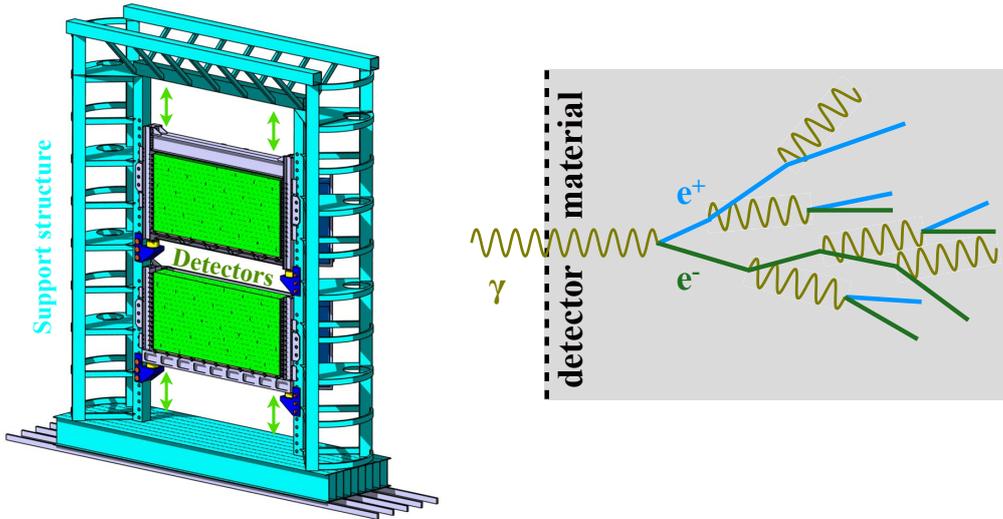


Figure 2.9: The ECAL geometry (left) including moving sections with about 1000 modules, which consist of absorbers followed by the scintillator detectors [40] and a scheme of the electromagnetic shower produced in the detector (right).

ECAL for CBM will consist of absorbers followed by scintillator detectors (see Fig. 2.9, left). The calorimeter will be composed of modules which consist of 140 layers of 1 mm lead and 1 mm scintillator, with cell size of 6x6 cm<sup>2</sup> [22, 40]. Similar ECAL detectors built by the same technology operate already in the experiments at RHIC and LHC.

### 2.3.8 Projectile Spectator Detector (PSD)

The PSD detector [41] will be used to determine the collision centrality and the reaction plane, that is necessary to measure collective properties of a collision, like anisotropy in particles distributions or event-by-event fluctuation of particle yields. The PSD detector is a calorimeter, which will consist of 44 individual modules with a transverse size of  $20 \times 20 \text{ cm}^2$ , each module includes 60 lead-scintillator sandwiches (see Fig. 2.10). The detector will measure the number of non-interacting nucleons (spectators) from a projectile nucleus in the collision.

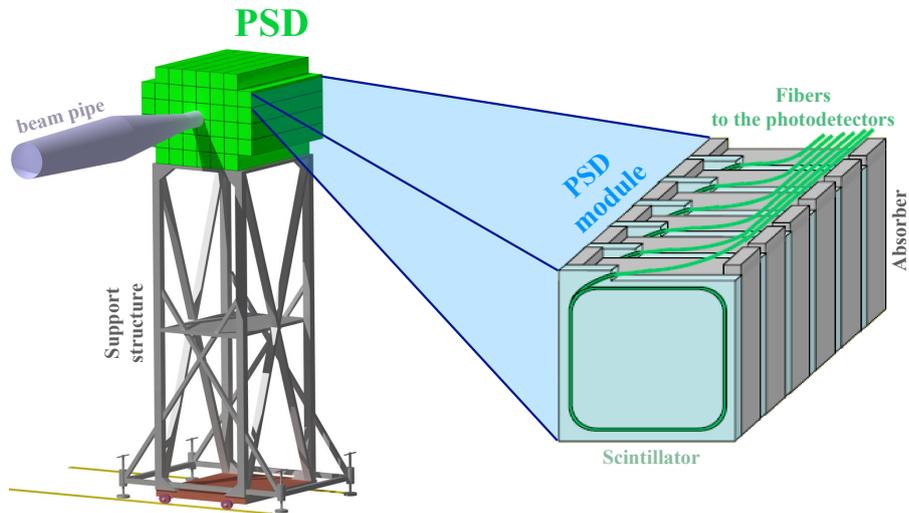


Figure 2.10: The PSD geometry (left) and the design of a module consisting of consecutive absorber and scintillator layers (right) [41].

### 2.3.9 Data AcQuisition system (DAQ)

The physics program of CBM requires measurement of rare signals with high statistics. Therefore the experiment is being designed to operate at extremely high collisions frequencies — up to 10 MHz [22]. Based on the possible recording speed of the modern storage media of about 1 GB/s and the data flow of 40 kB from gold-gold collisions at energies of 25 GeV, it is possible to store the data with a collisions frequency of only 25 kHz. Therefore, measurements with the collision frequency of 10 MHz require algorithms for online event selection that can reduce the number of background collisions, i.e. those that do not contain

interesting signals, in 400 or more times.

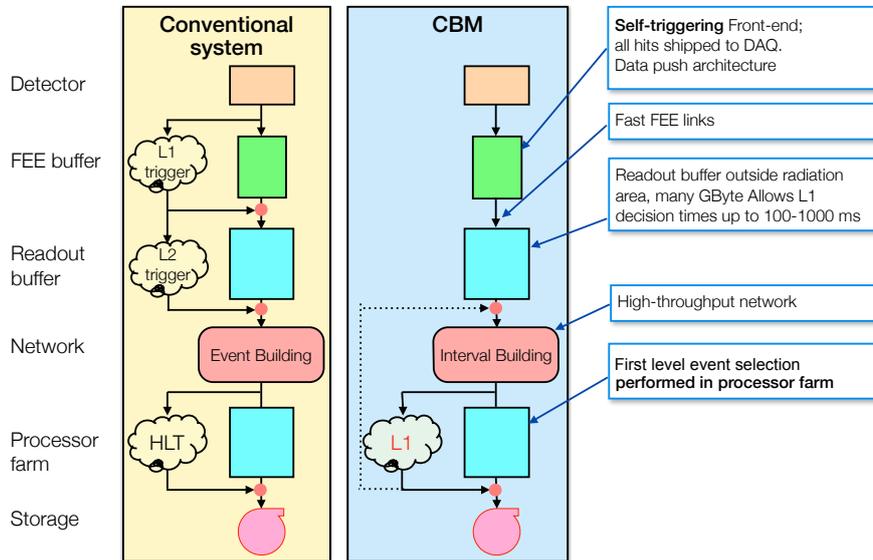


Figure 2.11: The CBM data acquisition model in comparison to the conventional approach [42]. Usually the input data passes several trigger levels being gradually suppressed. Due to its running scenario, the CBM experiment will not have such triggers. The level one trigger (L1) in case of CBM will be already a High Level Trigger (HLT) running on the First Level Event Selection (FLES). Therefore the data should be transmitted to FLES through the high throughput links.

The CBM experiment will not have a clear signature based on the raw data from the detectors for selection of interesting collisions. This makes impossible to build low level triggers, which are commonly used by other experiments in high energy physics and physics of heavy ions. In contrast to the conventional approach, CBM will rely on the high-level trigger only (see Fig. 2.11): the data will be collected from the detectors and transmitted to the First Level Event Selection algorithms running on many-core computer farm.

## 2.4 First Level Event Selection (FLES)

For the most important signals of the CBM physics case there are no simple criteria for collisions triggering and selection. Therefore the full reconstruction of the collision is required already at the selection stage and will be run in the online mode [22]. The basis for the online event selection will be fast algorithms

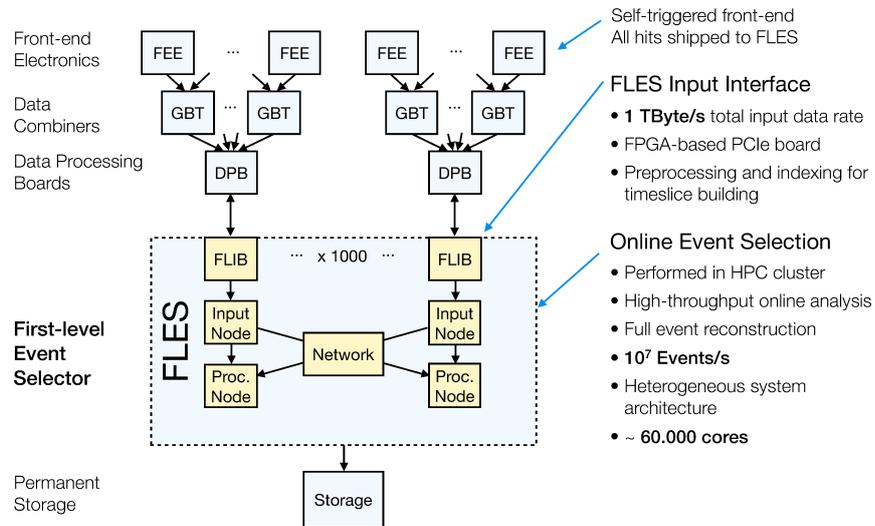


Figure 2.12: The FLES system for the CBM experiment [43]: the data collected from the detectors will be transported to the dedicated computer farm, where it will be processed, selected and stored.

for reconstruction of hits, tracks and short-lived particles running on a dedicated computer farm. Based on the reconstructed short-lived particles the collisions will be triggered and stored on tape. Fig. 2.12 illustrates the scheme of FLES and the data transfer.



Figure 2.13: The L-CSC supercomputer at GSI [44], which has been ranked as the most energy-efficient supercomputer.

The farm for the online event selection of the CBM experiment will consist of nodes equipped with modern many-core central processing units (CPU) and graphic cards with the total power of about 60 000 CPU cores equivalent. Thus, the farm is expected to be a massively parallel heterogeneous system. The algo-

gorithms for the FLES reconstruction, including short-lived particle reconstruction, should be parallelized and optimized for such a system.

Several prototypes for the FLES farm were already designed and built by the group of Prof. Dr. Lindenstruth at the Frankfurt Institute for Advanced Studies (FIAS). Having a large number of computer nodes, such cluster should be not only fast, but also energy efficient. This is taking into account in the supercomputer development at FIAS and the most recent L-CSC supercomputer at GSI (see Fig. 2.13) [44] is ranked as the world's most energy-efficient supercomputer. With a computing power of 5.27 GFLOPS per watt, it has set a new world record for energy efficiency among supercomputers.

# Chapter 3

## Many-core computer architectures and tools for parallel programming

Modern experiments in the high-energy physics require the highest possible speed of reconstruction because of high interaction and data rates. Thus, the power of existing computational servers should be utilized in the most efficient way.

At present, the development of high-performance servers tend to increase the number of processors per server, the number of cores per processor and the width of SIMD vectors [45] in the core. Each server contains one or more CPUs and several accelerators — graphic cards or coprocessors. All processors are connected with each other, with the main memory and with the peripherals by the motherboard — the main printed circuit board of the server. The data flow between CPU, memory and peripherals is managed by the chipset, which is a part of the motherboard, accelerators are connected to the motherboard by the PCI Express bus. All the processors are many-core devices, so a single server is already a many-core heterogeneous system.

Since nowadays computers are mainly parallel machines, reconstruction algorithms, including reconstruction of short-lived particles (see Chapter 6), should be developed as parallel algorithms. It is important that the speed of implemented algorithms should grow proportionally to the number of cores per processing unit and number of elements per register of the processing unit. For the code vec-

torization and parallelization different tools exist, that will be described in the current Chapter.

### 3.1 Topology of CPU

Modern servers are equipped mainly with two or more Central Processing Units (CPU), therefore their topology is complex and should be well understood when developing a program. Moreover, being a many-core device, already a single CPU has a complex structure. For example, a block-diagram of the Intel Xeon E7-4860 CPU is shown in Fig. 3.1.

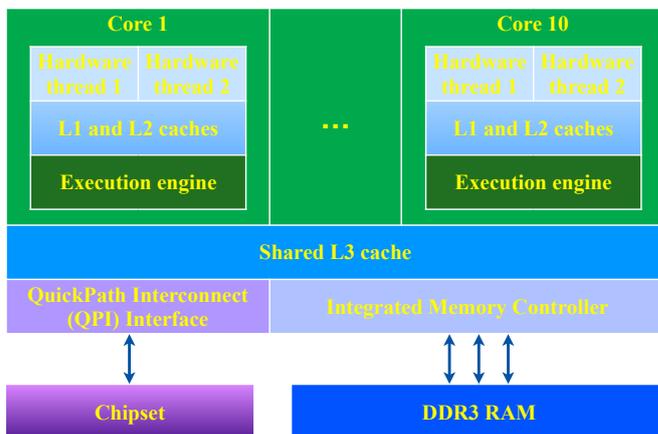


Figure 3.1: Block-diagram of the Intel Xeon E7-4860 CPU with 10 cores.



Figure 3.2: Typical access speed to different memory layers on modern CPUs.

The CPU contains 10 physical cores at 2.27 GHz and supports Intel Hyper-Threading (HT) technology [46], thus, architecturally each core consists of two logical ones. Each logical core has a full set of registers: data, segment, control, debug, and most of the machine specific registers. Also it has its own Advanced Programmable Interrupt Controller (APIC). This allows to run on each physical core two hardware threads simultaneously. Unlike the multiple processor systems with completely independent processor units, the logical cores share some resources: caches, execution engine, including Arithmetic Logic Unit (ALU), Floating Point Unit (FPU), and Vector Processing Unit (VPU), memory and

chipset interfaces. For example, the lxir075 server at GSI equipped with four Intel Xeon E7-4860 processors can operate in total with 80 threads in parallel.

Nowadays processing speed is growing faster, then memory speed. Thus, the memory issue is essential for the overall performance. In order to speedup calculations, memory on modern servers is organized in several levels with different access speed (see Fig. 3.2). The main memory of the server (Random-Access Memory, RAM) has the biggest size, usually several GB or even more than hundred GB, but has the lowest access speed. The cache memory is a fast memory located directly on the CPU chip. The amount of cache is usually considerably smaller, then the main memory: several KB of Level 1 (L1), several hundreds KB of Level 2 (L2), and several MB of Level 3 (L3) cache. For instance, each core of the Intel Xeon E7-4860 CPU has 32 KB of L1 cache for instructions, 32 KB of L1 cache for data and 256 KB of L2 cache for data and instructions. In addition, each CPU contains 24 MB of L3 cache memory, which is shared among the cores. The cache is one-two orders of magnitude faster then the main memory, thus, the frequently used data can be kept in the cache, reducing the time for the memory access. All caches of the CPU are usually coherent, that guaranties that all copies of the memory cell will have exactly the same value in all memory levels, therefore a software developer should not add to the program additional synchronizations.

The cache structure is of the particular importance when designing software for CPU: reducing the amount of memory used by the program in order to fit the cache can speed up the program by one-two orders of magnitude. It is also important to minimize the number of read operations from the main memory. Once the data is read to the cache, it should be used by all parts of the program, where it is needed, to avoid repeated accessing to RAM.

The structure of the main memory is also optimized with respect to the access speed. If many cores operate with the same memory, which has limited number of ports for reading and writing, the processor will stall while waiting for data to arrive. One of the solution is to build a Non-Uniform Memory Access (NUMA) system. Modern servers are usually based on the NUMA architecture, like the lxir075 server. The main memory of such systems is distributed among CPUs. For example, in lxir075 16 GB of RAM is attached to each CPU (see Fig. 3.3).

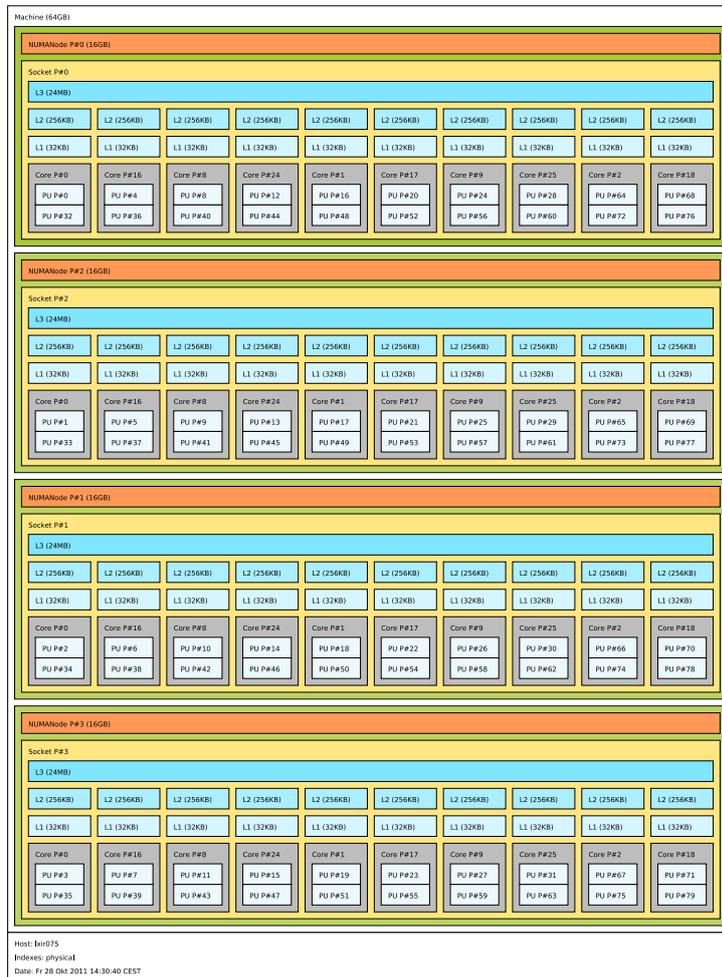


Figure 3.3: The structure of cores, cache and RAM on the lxir075 server at GSI. Each physical core contains two logical cores, 32 KB of L1 and 256 KB of L2 cache memory. Each CPU operates with 24 MB of L3 cache common for all 20 logical cores and 16 GB of RAM.

Access to the memory local for the CPU is considerably faster than to the remote memory. In NUMA systems CPUs are interconnected with each other, that allows access to the remote main memory. For instance, in Intel Xeon CPUs the QuickPath Interconnect (QPI) technology [47] is used for such connection. When operating with NUMA servers it is important to keep the data, which is used by a current CPU, in the correspondent part of memory in order to utilize the advantage of the memory architecture.

### 3.2 Intel Xeon Phi coprocessor

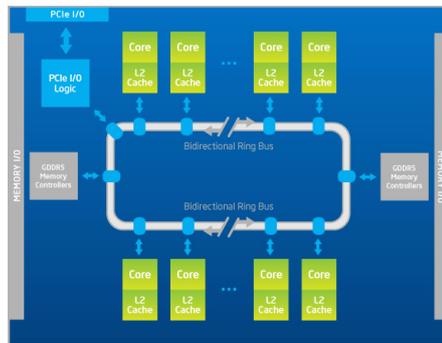


Figure 3.4: Block-diagram of the Intel Xeon Phi architecture: all cores are connected through the bus with each other and with the DDR5 memory.

The idea of the Intel Xeon Phi coprocessor [48] is to provide a highly parallel computing platform, which at the same time allows to work with standard for CPU programming languages and tools (like OpenMP, Intel TBB, Intel Cilk). The KNC (Knights Corner) card is based on the x86-compatible architecture, that provides a possibility to develop applications, which can be run both on a CPU and a coprocessor. Such approach can

greatly reduce the complexity of software development.

The KNC coprocessor is build using 22 nm manufacturing technology and provides up to 61 cores with a total performance of 1.2 TFLOPS (Tera Floating-point Operations Per Second). It contains upto 16 GB of GDDR5 memory. Intel Xeon Phi operates with larger than CPU SIMD vectors: 16 floats can be packed together in a vector. The HT technology of the card is also more aggressive: four threads can be run simultaneously on each physical core, which gives in total up to 244 threads. VPU of each core has 128 of 512-bit vector registers divided among the threads, thus providing 32 entries per thread. In addition, there are eight 16-bit mask registers per thread, which are part of the vector register file. The mask can be used to mark active elements of the SIMD-vector during calculations. Similar to CPU the core contains 32 KB of L1 cache for data and 32 KB for instructions and 0.5 MB of L2 cache both for data and instructions. The cores are connected through a bidirectional ring bus (see Fig. 3.4). The core can look up the L2 caches of other cores for the needed data in order to speed up the data access. Communication between the cores and the main memory is organized through eight memory controllers distributed along the bus in order to optimize the data throughput.

The next generation of Intel Xeon Phi (KNL, see Fig. 3.5) is built using 14 nm

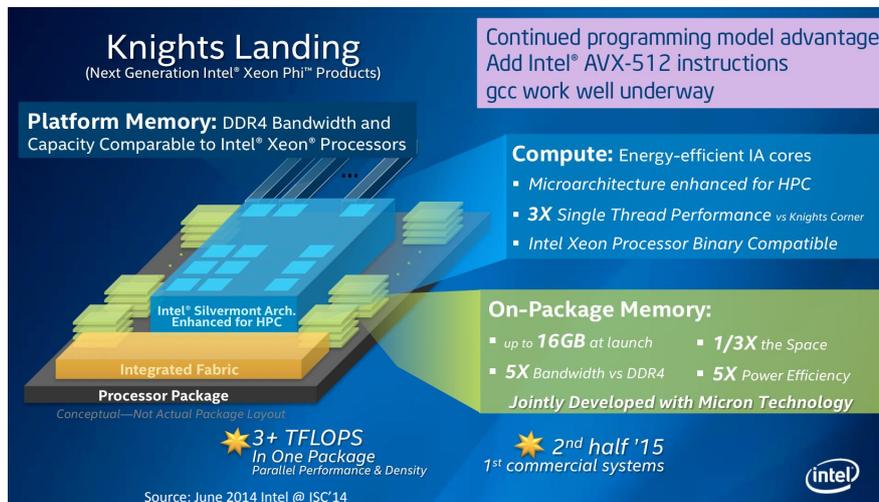


Figure 3.5: The next Intel Xeon Phi product — Knights Landing (KNL). It offers increased number of cores together with increased memory bandwidth, that allows to achieve more than 3 TFLOPS of the computing power [49].

manufacturing technology. It is produced in two options: a standalone bootable processors is available in addition to the PCIe coprocessor version. The number of cores and memory bandwidth is increased. Vectorization is implemented using the AVX-512 instruction set , binaries compiled for the Intel Xeon Phi and Intel Xeon CPUs are compatible.

### 3.3 GPU architecture

Modern servers are usually equipped not only with CPUs, but also with high performance Graphics Processing Units (GPUs). This provides significant increase of the total performance of a server.

The architecture of GPU differs from the CPU. GPU contains more cores with less local memory. Let us consider the GPU architecture on an example of Nvidia GTX 480 [50], that is used for calculations in the current work. This card contains 480 cores, which have a fully pipelined integer ALU and FPU. The cores are grouped into streaming multiprocessors 32 cores each sharing the buses, schedulers and memory resources. The structure of the streaming multiprocessor of Nvidia GTX 480 is given in Fig. 3.6, left. Modern GPUs provide the L1 data and instruction cache shared by all cores from the streaming multiprocessor. Like



Figure 3.6: Structure of streaming multiprocessors of the Nvidia GTX 480 GPU based on the Fermi architecture (left) [50] and the Nvidia GTX 980 GPU based on the Maxwell architecture (right) [51].

in CPU, caching allows to minimize the number of the main memory access and, thus, to speed up the code. On top of this, the L2 cache accessible by all streaming multiprocessors is added, that allows better memory optimization. Nvidia GTX 480 has 1.2 GB of the main memory with a bandwidth of 133.9 GB/s.

Current development of GPUs tends to increase a total number of cores, a number of cores per streaming multiprocessor, memory size including caches and the memory bandwidth, that improves the overall performance. For example, one of the most recent GPUs, Nvidia GTX 980 (see Fig. 3.6, right) based on the Maxwell architecture [51], has 2048 cores divided into streaming multiprocessors 128 cores each. It has 4 GB main memory with 224.3 GB/s bandwidth.

## 3.4 Approaches for parallel programming

According to Flynn's taxonomy, a classification proposed by M. Flynn, four basic classes of computer architectures are considered based upon the number of concurrent instruction and data streams:

- **SISD** (Single Instruction, Single Data stream). As follows from the name, SISD operates with only one stream of instructions. The data, which is used in the calculations in this case, is a scalar data, i.e. a one single value. This approach is used, for instance, when only one core of a CPU is used with scalar data types, such as `int`, `float` or `double` in C++.
- **SIMD** (Single Instruction, Multiple Data streams) means execution of the same instruction on a set of data (a vector) simultaneously. Because in calculations vectors of data are used, implementation of this approach is called vectorization or data parallelism. In CPUs SIMD approach is applied, when the MMX, SSE<sup>1</sup> or AVX<sup>2</sup> instruction sets are used, in Intel Xeon Phi — the IMIC<sup>3</sup> instruction set. These sets differ in possible instructions, their hardware implementation and in a size of registers (data vectors). MMX and SSE instructions operate with a dedicated registers of size 128 bit, which allow to pack 4 values in a single precision (floats), AVX instructions — with registers of size 256 bit (8 floats), IMIC — with registers of size 512 bit (16 floats). In GPUs a group of cores (a compute unit or a streaming multiprocessor), each running its own thread, executes the same instruction on a set of data, therefore this approach is usually called SIMT (Single Instruction Multiple Threads), when working with GPUs.
- **MISD** (Multiple Instruction, Single Data stream). This class includes architectures, where many functional units perform different operations on the same data. It is used, for example, in fault-tolerant computers executing the same instructions redundantly in order to detect and mask errors. Such approach was used in the Space Shuttle program of NASA in flight control computers.

---

<sup>1</sup>Streaming SIMD Extensions

<sup>2</sup>Advanced Vector Extensions

<sup>3</sup>Initial Many Core Instructions

- **MIMD** (Multiple Instruction, Multiple Data streams). At any time, in MIMD architectures different instructions may be executed by different processors on different data. Having several independent cores, most processors are MIMD systems. Computer with two or more processors are another example of MIMD machines. Utilization of this approach in a program code is called parallelization or task parallelism.

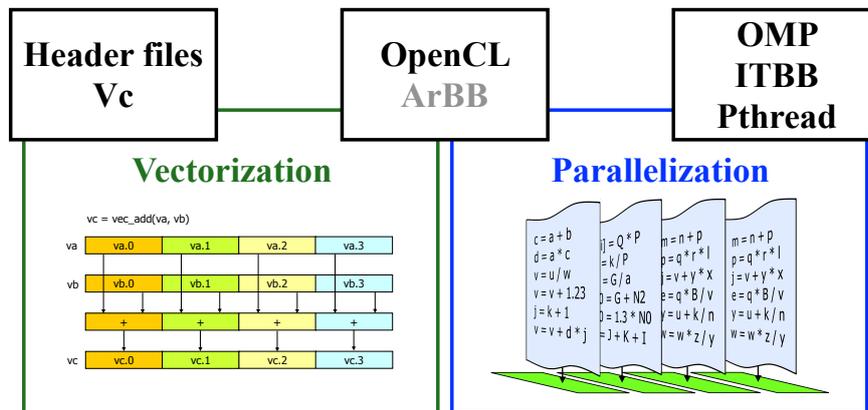


Figure 3.7: Tools for code vectorization and parallelization.

For utilization of SIMD and MIMD approaches in a program code different tools and methods are used (see Fig. 3.7). Vectorization of the code is done, for instance, with:

- the header files overloading the operators (Header files) [52];
- the Vc (Vector classes) library [53, 54];
- the OpenCL (Open Computing Language) framework [55];
- Intel ArBB (Array Building Blocks) [56].

For parallelization of the program code between cores of CPU a number of frameworks and libraries exists, for example:

- OpenMP (Open Multi-Processing) [57];
- Intel TBB (Threading Building Blocks) [58];
- the Pthreads (or POSIX threads) library [59];
- OpenCL;
- Intel ArBB.

## 3.5 Tools for vectorization

Since most of modern CPUs contain a SIMD unit (VPU), vectorized code makes better use of the already existing hardware capabilities and allows to reach significant acceleration of the code. However, writing a code using SIMD instructions directly is a complicated task. For example, a simple code to calculate a first-order polynomial function, which looks clear in a scalar mode:

```
float y = a*x + b;
```

with SIMD instructions:

```
__m128 y = _mm_add_ps(_mm_mul_ps(a,x),b);
```

looks rather difficult.

### Header files

The simplest approach for development of a vectorized code in a high-level scalar-like mode is a construction of the header files overloading all the basic arithmetic and logic operations using SIMD instructions, which makes the code compact and easy readable [52, 60]. The code above for calculation of the first-order polynomial function using the header files:

```
fvec y = a*x + b;
```

with overloaded operators

```
friend fvec operator+(const fvec &a, const fvec &b) {  
    return _mm_add_ps(a,b); }  
friend fvec operator*(const fvec &a, const fvec &b) {  
    return _mm_mul_ps(a,b); }
```

recovers the scalar form.

For different CPU architectures the instruction set as well as the size of a vector can be different. Due to the simplicity of implementation, the header files approach is flexible with respect to different CPUs. When running on a new architecture it allows the program code to remain without changes, only the appropriate header file should be included. Also, a header file with a scalar implementation exists, that allows to debug a program in a usual scalar mode.

## Vector classes (Vc) library

As an extension of the header files idea the Vc library [53, 54] for code vectorization was developed. The Vc library is a convenient and accurate tool to use SSE, AVX and IMIC instructions. Like the header files, it provides access to basic arithmetic and logical operations and functions, contains a scalar implementation for debugging. The example code for the calculation of a polynomial function using Vc is also simple and intuitive:

```
float_v y = a*x + b;
```

In addition to the operations, which use the entire vector (so called vertical operations) Vc implements operations inside the vector (horizontal operations), which allow, for example, summation or sort of the vector elements.

To implement the interfaces for conditional operations, all functions can optionally take a mask as an argument. For example, to replace all the negative elements of the vector “y” from the previous example with zeros, the code must be changed accordingly:

```
float_m mask = y<0.f;  
y(mask) = 0.f;
```

The library allows to work with a misaligned memory to store the data into a vector or read from it, implementing functions “gather” and “scatter”. This makes possible, for example, to fill a vector, using the elements of different parts of an array using the indices of the required elements.

When compiling a code the Vc library automatically detects the platform on which it is running and selects the appropriate instructions, which makes the code written with Vc easily portable between platforms. The library allows to achieve maximum possible speedup of the code using SIMD instructions. For example, speedup factors for basic mathematical operations: addition, subtraction, multiplication and division are shown in Fig. 3.8.

However, it is worth to notice that the Vc library is a complex tool, which is based on other software packages, such as binutils and a compiler. Therefore for its correct operation installation of corresponding versions of these programs is required. In case of testing a simple code, which does not use the benefits of Vc,

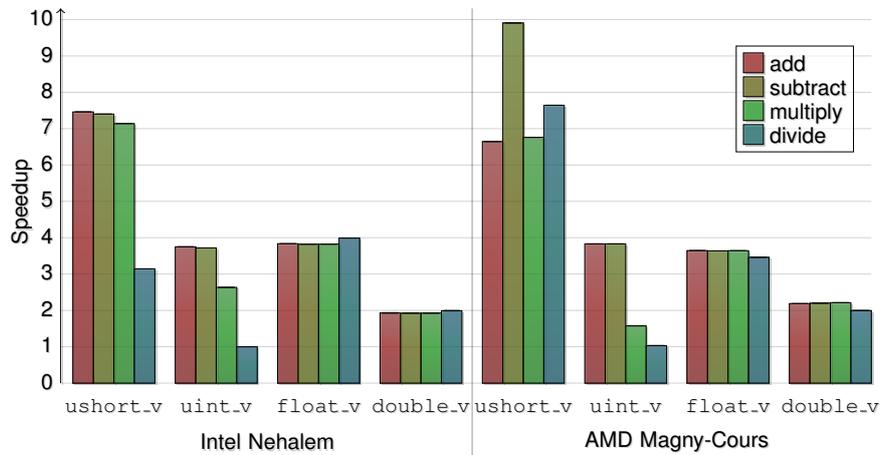


Figure 3.8: Speedup of the arithmetic operations with Vc SSE types vs. scalar types for unsigned short (`ushort_v`), unsigned integer (`uint_v`), float (`float_v`) and double (`double_v`) types on Intel Nehalem and AMD Magny-Core architectures [53].

the header file is a more flexible approach, that requires less efforts to move from one platform to another.

## 3.6 Tools for parallelisation

Here those tools for parallelization of the tasks between cores of CPU are considered, that were used for the current thesis.

### Open Multi-Processing (OpenMP)

OpenMP [57] is an API (Application Programming Interface), which consists of a set of compiler directives, library routines and environment variables that are used to develop multithreaded applications on multiprocessor systems with a shared memory. It defines a simple interface that allows to parallelize tasks between cores of a CPU.

The programming model of OpenMP is the fork-join parallelism (see Fig. 3.9): the master thread is created, when a program starts and when the calculation should be parallelized additional threads are created and the task is distributed between them. Nested parallelism is possible within the OpenMP programming model. The parts of the code, which should be executed in parallel, are marked

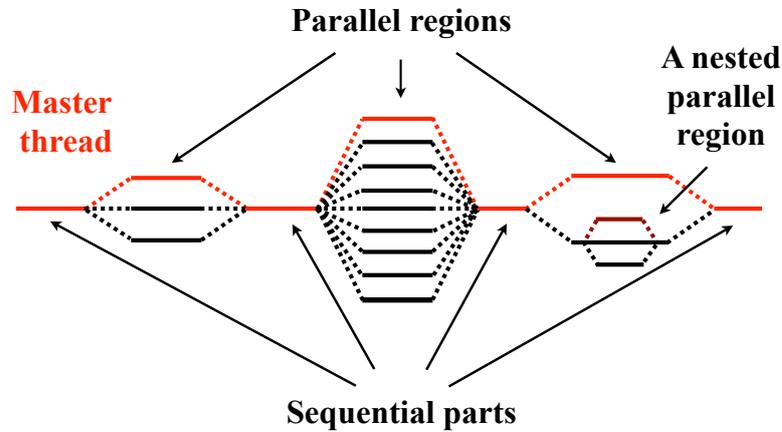


Figure 3.9: Illustration of the OpenMP programming model — the fork-join parallelism.

with the directives of the compiler preprocessor:

```
#pragma omp parallel
```

Threads in the OpenMP programming model are communicated by sharing variables. Each variable can be private or shared. The shared variables can be accessed by any thread. If the variable is declared as private a local copy of it is created by each thread.

Most of the constructs in OpenMP are compiler directives:

```
#pragma omp <construct> [clause [clause]...]
```

The implementation of programs using OpenMP constructs allow to keep the code unchangeable with respect to the single core version. OpenMP clauses allow to specify attributes of a construct. In order to enable the OpenMP directives an appropriate flag should be added during the compilation. OpenMP includes such constructs, for example, as `omp for`, `omp reduction(op:list)`, `omp num_threads()`, `omp scheduler [clause]`, etc. These directives allow to control the parallel regions, to collect the data from different threads, to synchronize threads, to set a scheduler, etc.

In order to control the race conditions the framework contains a set of tools for thread synchronization including high level directives, like `omp critical`, `omp atomic`, `omp barrier`, and a set of simple lock functionality for a low level control.

OpenMP contains a set of runtime library routines, which allow, for instance, setting and checking the number of threads, checking the maximum number of threads, getting the number of a current thread, checking the number of cores in a computer, controlling the nested parallelism, operating with the simple lock functionality etc.

## Intel Threading Building Blocks (ITBB)

Intel Threading Building Blocks [58] is a C++ template library, which offers a rich functionality allowing parallel execution of a big number of different types of problems. In order to abstract access to the multiple processors the groups of operations is treated as tasks. An ITBB program creates, synchronizes and destroys graphs of dependent tasks. Tasks are then executed respecting graph dependencies. As OpenMP, ITBB supports nested parallelism, so larger parallel components can be build from smaller parallel components.

The library interfaces mostly employ generic programming, in which interfaces are defined by requirements on types and not specific types. This allows to customize components to a specific needs of the application.

Like OpenMP, ITBB can be used for loops parallelization with `parallel_for`. In addition to the standard schedulers, ITBB allows to create a custom scheduler to control the task size and their distribution among threads. The library also includes such functions as `parallel_reduce`, `parallel_scan`, `parallel_while`, `parallel_sort`, etc. Thread synchronization is done by atomic operations and mutual exclusion.

## Pthreads library

In NUMA systems memory is distributed among CPUs and the access speed to the local memory is considerably faster, than to the remote memory. If during the run-time a thread migrates from one CPU to another, the already allocated memory becomes remote for it and a program performance degrades. In order to avoid such situation each thread should be bound to a certain core or a group of cores of the CPU by setting a core affinity. For this purpose the Pthreads library [59] is used in current work.

Pthreads (or POSIX threads) are defined as a set of C language programming types and procedure calls, implemented with a `pthread.h` header file and a thread library. In the modern systems this library is usually a part of another libraries, such as `libc`. With the Pthreads library threads can be created and manipulated manually. Being a low level library, Pthreads allows to highly optimize the code.

### 3.7 Universal tools for vectorization and parallelization

With already described tools code developers are still dependent on the low-level mechanisms of the parallelism and hardware architectures. Several attempts were made to develop a common language, which supports both ways of the code parallelization and combines their advantages and automatically optimizes the code for the running platform. Here we consider two of them: Intel ArBB and OpenCL frameworks.

#### Intel Array Building Blocks (ArBB)

The Intel ArBB software [56] is an experimental language developed by Intel. It consists of a standard C++ library and powerful runtime-dynamic compiler. The whole available SIMD functionality is provided by this software. ArBB provides the following benefits:

- programs written with ArBB are scalable across cores and vector units (SIMD) allowing them to productively harness available CPU resources;
- integrates into C++ applications as a standard library with no compiler-specific extensions for compatibility with standard compilers and IDEs; it offers powerful language constructs to quickly express complex large data structures for vector parallel workloads;
- balancing performance and safety, ArBB utilizes a managed environment to prevent parallel programming bugs such as data races and deadlocks.

The code example with ArBB is as easily understandable, as with the previous tools:

```
dense<f32> y = a*x + b;  
dense<boolean> mask = y<0;  
y = select( mask, y, 0 );
```

In addition to vectorization, ArBB can automatically parallelize computations between cores of CPU. No additional modifications of the code is needed for that.

During development of Intel ArBB we participated in investigation of its applicability to the high energy physics experiments in collaboration with Intel [61].

## Open Computing Language (OpenCL)

OpenCL [55] is an open standard for general purpose parallel programming across CPUs, GPUs and other processors. It allows to write the universal code, which can be run both on the CPU and GPU giving software developers portable and efficient access to the power of the heterogeneous processing platforms. OpenCL supports a wide range of applications through a low-level, high-performance, portable abstraction. OpenCL consists of an API for coordinating parallel computation across heterogeneous processors and a cross-platform programming language with a well-specified computation environment. The OpenCL standard:

- supports both data- and task-based parallel programming models;
- utilizes a subset of ISO C99 with extensions for parallelism;
- defines consistent numerical requirements based on IEEE 754;
- defines a configuration profile for handheld and embedded devices Efficiently interoperates with OpenGL, OpenGL ES and other graphics APIs.

The platform model for OpenCL (see Fig. 3.10) consists of a host connected to one or more OpenCL devices. An OpenCL device is divided into one or more compute units which are further divided into one or more processing elements. The OpenCL application submits commands from the host to execute computations on the processing elements within a device. An OpenCL application runs on a host according to the models native to the host platform. The processing elements within a compute unit execute a single stream of instructions as SIMD units or each processing element maintains its own program counter.

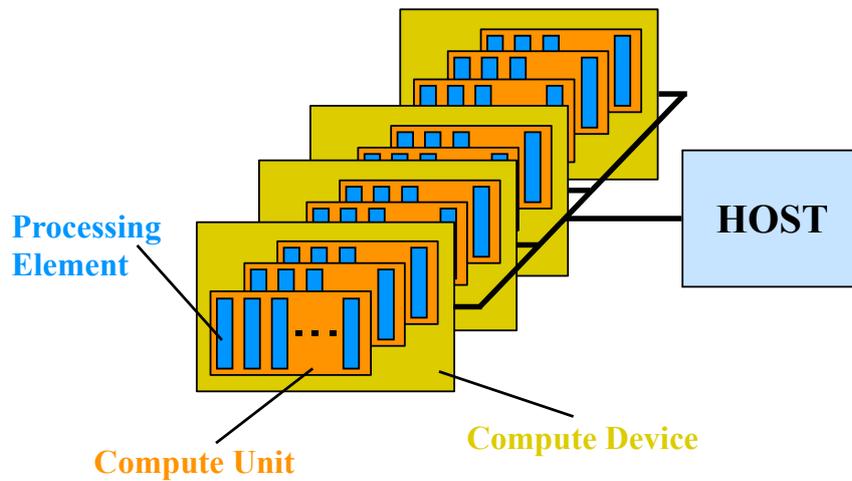


Figure 3.10: Platform model for OpenCL: one host plus one or more compute devices each with one or more compute units each with one or more processing elements [55].

Execution of an OpenCL program occurs in two parts: kernels, that are executed on one or more OpenCL devices, and a host program, that is executed on the host. The host program defines the context for the kernels and manages their execution. Also it defines a queue of the tasks and runs corresponding kernels according to the queue. The data is divided into work-groups. Each work-group receives a part of the data array and executes it concurrently on the processing elements of a single compute unit.

All the tools described in the current Chapter are used in the thesis for parallelization and speedup of particle reconstruction algorithms described in Chapters 4–6 and for their scalability investigation.

# Chapter 4

## Reconstruction of charged particles trajectories

Modern experiments in high energy physics operate at high track densities. Moreover, extremely rare signals with multiplicity down to  $10^{-7}$  per collision are of particular interest. To find such rare signals terabytes of data should be processed. Therefore reconstruction of collisions in high energy physics requires both high accuracy and high speed. One of the main tasks in the event reconstruction is reconstruction of particle trajectories (tracks). The accuracy of track parameters defines a quality of reconstructed rare signals and a probability to detect these signals.

The Kalman filter (KF) method [23, 24, 52, 62] is a recursive method for analysis of linear discrete dynamic systems described by a vector of parameters, which is called state vector. It allows to find an optimal estimation of the particle tracks parameters and to achieve the highest possible accuracy.

In many cases the Kalman filter is a core of event reconstruction in high energy physics due to its properties [62]:

- optimal estimation of the parameters;
- locality with respect to measurements: measurements are added one by one independently;
- nonlinearities of the particle trajectory in the general case of the nonhomogeneous magnetic field can be taken into account in a convenient way: the particle is transported between measurements taking into account all

nonlinear effects and its trajectory is linearized only in the neighborhood of each measurement;

- material of a detector crossed by the particle is added in the neighborhood of each measurement, which simplifies the procedure and allows to take into account all effects of the material with a minimal set of assumptions;
- operation with matrices of a small size (up to  $5 \times 5$ );
- number of calculation operations is proportional to the number of measurements.

## 4.1 The Kalman filter method

Linear discrete dynamical systems are completely described by the vector of parameters  $\mathbf{r}^t$ , which is called the state vector. The Kalman filter method is developed as a linear recursive method for parameters estimation of the desired state vector  $\mathbf{r}^t$  according to known measurements. The estimation  $\mathbf{r}$  can be found with some error  $\boldsymbol{\xi} = \mathbf{r} - \mathbf{r}^t$ . To estimate this error the covariance matrix (or matrix of errors)  $C \equiv \langle \boldsymbol{\xi} \cdot \boldsymbol{\xi}^T \rangle$  is introduced and calculated by the Kalman filter. The Kalman filter starts with an arbitrary initial approximation of the parameters  $\mathbf{r} = \mathbf{r}_0$  and refines the state vector consecutively adding one measurement after the other. The optimum value is attained after the addition of the last measurement.

In general, the measurement is a vector  $\mathbf{m}_k^t$ ,  $k = 1 \dots n$ , which depends on the state vector  $\mathbf{r}^t$ . Here  $n$  is a number of measurements,  $k$  is an index of the measurement. The dimension of the measurement should not coincide with the dimension of the state vector necessarily. The measurement is assumed to be known with an error  $\boldsymbol{\eta}_k$ :

$$\mathbf{m}_k = \mathbf{m}_k^t + \boldsymbol{\eta}_k. \quad (4.1)$$

The Kalman filter method requires that the covariance matrix of each measurement  $V_k \equiv \langle \boldsymbol{\eta}_k \cdot \boldsymbol{\eta}_k^T \rangle$  should be known.

The Kalman filter method assumes that  $\mathbf{m}_k^t$  depends linearly on  $\mathbf{r}^t$ :

$$\mathbf{m}_k^t = H_k \mathbf{r}^t. \quad (4.2)$$

The matrix  $H_k$  is called the model of measurement.

The state vector  $\mathbf{r}^t$  of the dynamic discrete system can change from one measurement to another. For instance, in the high energy physics experiments the measurements are done by the different subdetectors and are separated in space. Let us denote the state vector  $\mathbf{r}^t$  in conditions of the  $k$ -th measurement  $\mathbf{r}_k^t \equiv \mathbf{r}^t(k)$ .

The evolution of the linear system between the current measurement  $\mathbf{m}_k$  and the next  $\mathbf{m}_{k+1}$  is described by the linear equation:

$$\mathbf{r}_{k+1}^t = F_k \mathbf{r}_k^t + \boldsymbol{\nu}_{k+1}, \quad (4.3)$$

where  $F_k$  is a known matrix of the system evolution,  $\boldsymbol{\nu}_{k+1}$  is a random process noise. The noise can be caused, for example, by the multiple scattering of a charged particle passing the detector material. The covariance matrix of the process noise  $Q_k \equiv \langle \boldsymbol{\nu}_k \cdot \boldsymbol{\nu}_k^T \rangle$  should be known in order to estimate the noise correctly during the parameters estimation.

It is assumed that the errors of the measurements are unbiased and do not correlate with each other. Also the process noise is assumed to be unbiased and uncorrelated with the measurement errors:

$$\begin{aligned} \langle \boldsymbol{\eta}_k \cdot \boldsymbol{\eta}_j \rangle &= 0, & k = 1 \dots n, & j = 1 \dots n \\ \langle \boldsymbol{\nu}_k \cdot \boldsymbol{\nu}_j \rangle &= 0, & k = 1 \dots n, & j = 1 \dots n \\ \langle \boldsymbol{\eta}_k \cdot \boldsymbol{\nu}_j \rangle &= 0, & k = 1 \dots n, & j = 1 \dots n \\ \langle \boldsymbol{\eta}_k \rangle &= \boldsymbol{\nu}_l = \mathbf{0}. \end{aligned} \quad (4.4)$$

The conventional Kalman filter method includes several steps and can be formulated as follows (the method is illustrated in Fig. 4.1):

1. Initialization step. The initial approximation of the parameters  $\mathbf{r}_0$  together with its covariance matrix  $C_0$  should be set before filtering. The initial approximation  $\mathbf{r}_0$  can be chosen arbitrary. For example, a least squares approximation with certain simplifications can be used. If the initial es-

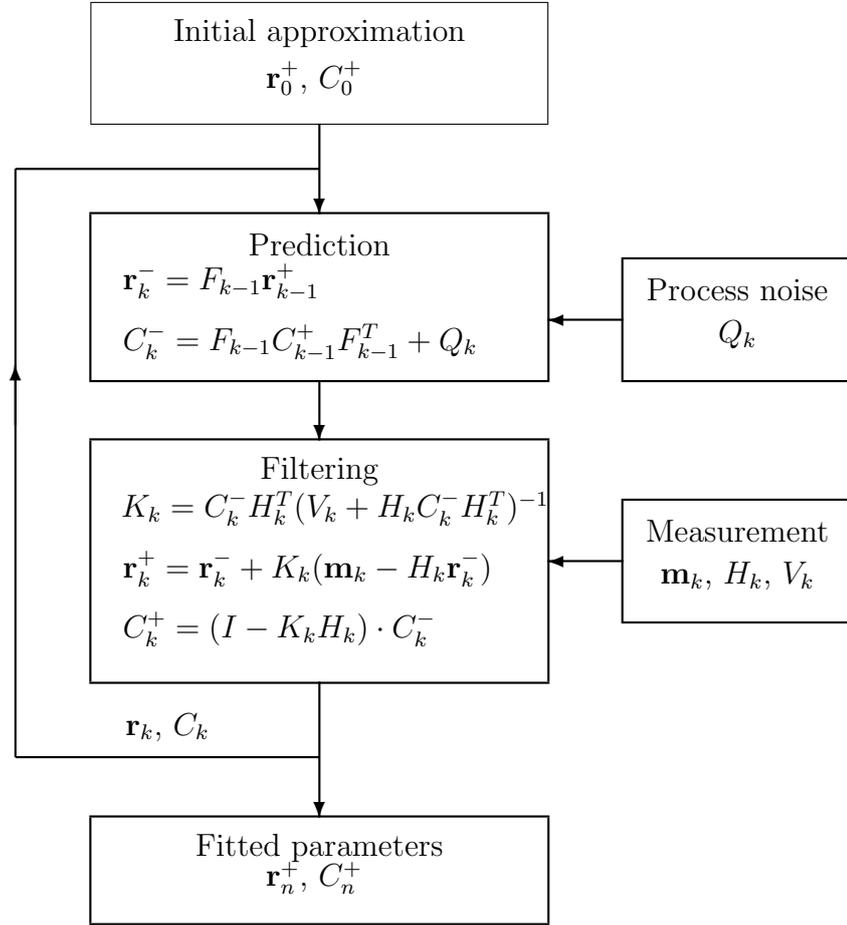


Figure 4.1: Block diagram representation of the Kalman filter method [52].

timization is not known the  $\mathbf{r}_0$  could be set to zero. The initial covariance matrix  $C_0$  is chosen as a diagonal matrix with infinite diagonal values. Such choice of the covariance matrix allows to minimize the influence of the initial approximation on the final optimum estimation.

2. Prediction step. Before addition of each next measurement  $\mathbf{m}_k$  the estimation of the state vector obtained at the previous filtration step  $\mathbf{r}_{k-1}^+$  together with its covariance matrix  $C_{k-1}^+$  should be known within the conditions of this measurement. For this the estimation should be extrapolated to the point of measurement  $\mathbf{m}_k$ . Also the process noise  $\boldsymbol{\nu}_k$  is taken into account at this step. The estimation of the state vector  $\mathbf{r}_k^-$  and its covariance matrix

$C_k^-$  after extrapolation are:

$$\mathbf{r}_k^- = F_{k-1} \mathbf{r}_{k-1}^+, \quad (4.5a)$$

$$C_k^- = F_{k-1} C_{k-1}^+ F_{k-1}^T + Q_k, \quad (4.5b)$$

In the first iteration it is assumed that  $\mathbf{r}_0^+ = \mathbf{r}_0$  and  $C_0^+ = C_0$ .

3. Filtration step. At this step the estimation  $\mathbf{r}_k^-$  is updated with the new measurement  $\mathbf{m}_k$  that gives the optimum estimation of the state vector  $\mathbf{r}_k^+$  according to first  $k$  measurements.

In the beginning the residual  $\boldsymbol{\zeta}_k$  between the measurement  $\mathbf{m}_k$  and the state vector estimation after the prediction step  $\mathbf{r}_k^-$  is calculated together with the inverse covariance matrix of the residual  $S_k$ :

$$\boldsymbol{\zeta}_k = \mathbf{m}_k - H_k \mathbf{r}_k^-, \quad (4.6a)$$

$$S_k = (V_k + H_k C_k^- H_k^T)^{-1}. \quad (4.6b)$$

The matrix  $S_k$  is also called a weighting matrix.

Based on the calculated weighting matrix a gain matrix  $K_k$  is calculated, that is proportional to the current estimation of the state vector covariance matrix  $C_k^-$  and the matrix  $S_k$ :

$$K_k = C_k^- H_k^T S_k. \quad (4.6c)$$

The gain matrix  $K_k$  is an optimal coefficient matrix, which determines the contribution of the residual  $\boldsymbol{\zeta}_k$  into a new estimation  $\mathbf{r}_k^+$  of the state vector:

$$\mathbf{r}_k^+ = \mathbf{r}_k^- + K_k \boldsymbol{\zeta}_k. \quad (4.6d)$$

Note, that in case of the zero residual  $\zeta_k$  or the zero gain matrix  $K_k$  the estimation of the state vector remains the same. Thus, measurements with the larger weight will correct the estimation of the state vector stronger. The correction will be negligible if the estimation of the state vector is known with a high accuracy and has a small covariance matrix.

The covariance matrix estimation  $C_k^+$  of the refined state vector  $\mathbf{r}_k^+$  is calculated by the formula:

$$C_k^+ = (I - K_k H_k) \cdot C_k^- . \quad (4.6e)$$

The reliability of the fit by the Kalman filter method is characterized by the  $\chi^2$ -deviation. The  $\chi^2$ -deviation is calculated iteratively based on the residual  $\zeta_k$  and the weighting matrix  $S_k$ . Note, that the residual  $\zeta$  and the  $\chi^2$ -deviation are important statistics. Their distributions are well known and therefore can be used for the characterization of a fit quality. The total  $\chi^2$ -deviation of the estimation  $\mathbf{r}_k^+$  from first  $k$  measurements is:

$$\chi_k^2 = \chi_{k-1}^2 + \zeta_k^T S_k \zeta_k . \quad (4.6f)$$

The last two steps are repeated iteratively for each measurement  $\langle \mathbf{m}_k \rangle$ . The desired optimum estimation  $\mathbf{r}_n^+$  of the state vector together with its covariance matrix  $C_n^+$  are obtained after filtration of the last measurement.

If the estimated measurement  $\langle \mathbf{m}_k^t \rangle$  does not depend linearly on the state vector  $\mathbf{r}^t$  the dependency  $\mathbf{m}_k^t(\mathbf{r}^t) = \mathbf{h}_k(\mathbf{r}^t)$  should be linearized by the Taylor expansion:

$$\begin{aligned} \mathbf{m}_k^t(\mathbf{r}_k^t) &= \mathbf{h}_k(\mathbf{r}_k^t) \approx \\ &\approx \mathbf{h}_k(\mathbf{r}_k^{lin_m}) + H_k(\mathbf{r}_k^t - \mathbf{r}_k^{lin_m}), \\ H_{k_{ij}} &= \left. \frac{\partial \mathbf{h}_{k_i}(\mathbf{x})}{\partial x_j} \right|_{\mathbf{x}=\mathbf{r}_k^{lin_m}} , \end{aligned} \quad (4.7)$$

here  $H_{k_{ij}}$  is a matrix element  $(i, j)$  of the matrix  $H_k$ ,  $\mathbf{h}_{k_i}$  is an element  $i$  of the

vector  $\mathbf{h}_k$ ,  $x_j$  is an element  $j$  of the vector  $\mathbf{x}$ . The vector  $\mathbf{r}_k^{lin_m}$  is a point of the expansion, it is chosen close to the possible real value of the vector  $\mathbf{r}_k^t$ . For example,  $\mathbf{r}_k^{lin_m}$  can be set as  $\mathbf{r}_k^-$ .

The extrapolation formula  $\langle \mathbf{r}_{k+1}^t \rangle = \mathbf{f}_k(\mathbf{r}_k^t)$  should be linearized in a similar way if it is nonlinear:

$$\begin{aligned} \langle \mathbf{r}_{k+1}^t \rangle &= \mathbf{f}_k(\mathbf{r}_k^t) \approx \\ &\approx \mathbf{f}_k(\mathbf{r}_k^{lin_p}) + F_k(\mathbf{r}_k^t - \mathbf{r}_k^{lin_p}), \\ F_{k_{ij}} &= \left. \frac{\partial \mathbf{f}_{k_i}(\mathbf{x})}{\partial x_j} \right|_{\mathbf{x}=\mathbf{r}_k^{lin_p}}. \end{aligned} \quad (4.8)$$

Here the point of expansion  $\mathbf{r}_k^{lin_p}$  can be chosen, for instance, as  $\mathbf{r}_k^+$ .

In case of the track fitting by the Kalman filter method the state vector  $\mathbf{r}_k$  is a vector of the track parameters, the extrapolation matrix  $F_k$  describes the motion of the charged particle in the nonhomogeneous magnetic field from one detector to another, the covariance matrix of the process noise describes the multiple scattering of the charge particle on the detector material.

## 4.2 Kalman filter for track parameters estimation

One of the most important applications of the Kalman filter method in high energy physics is parameters reconstruction of charged particles tracks. Crossing the material of detectors, charged particles interact with it. The points of interaction registered by the detector are measurements of the charged particle trajectory. According to these measurements the state vector of the track parameters is estimated using the Kalman filter method, that defines the trajectory of the particle.

Since the CBM experiment is a fixed target experiment, trajectories of the charged particles are close to the straight line<sup>1</sup>. A straight line is described in a

---

<sup>1</sup>Properties of the Kalman filter track fit applied to the CBM experiment in absence of the magnetic field, when tracks are straight lines, are investigated in Appendix A. Current investigation is a further update of the work [63] for a case of the CBM track state vector.

natural way by the point on the line and the slope at this point. In the presence of magnetic field the slope is not a constant value and varies from point to point according to the equation of motion:

$$dt_x/dz = c t_r (q/p) ( t_y(B_z + t_x B_x) - (1 + t_x^2)B_y), \quad (4.9a)$$

$$dt_y/dz = c t_r (q/p) (-t_x(B_z + t_y B_y) + (1 + t_y^2)B_x), \quad (4.9b)$$

$$t_r = \sqrt{t_x^2 + t_y^2 + 1}, \quad (4.9c)$$

where  $x$  and  $y$  are coordinates of the track at a given point with a fixed  $z$  coordinate ( $Z$ -axis is directed along the axis of the spectrometer in the beam direction,  $x$  and  $y$  are coordinates in the plane perpendicular to the axis of the spectrometer);  $t_x = dx/dz$  and  $t_y = dy/dz$  are slopes of the track in  $XZ$  and  $YZ$  planes respectively;  $q/p$  is an inverse momentum signed with the charge of the particle;  $B_x$ ,  $B_y$  and  $B_z$  are the components of the magnetic field at a given point;  $c$  is the speed of light. Thus, a complete and sufficient set of track parameters, that describes the state of the particle is:

$$\mathbf{r} = \{ x, y, t_x, t_y, q/p \}. \quad (4.10)$$

To apply the Kalman filter method for estimation of the track parameters an initial approximation  $\mathbf{r}_0$ , the matrix of the measurement model  $H$  and the extrapolation matrix  $F$  should be set.

For the initial approximation of the track parameters the least squares method is applied in assumption of the one-component magnetic field and multiple scattering in the material of the detector stations is neglected. This approach allows to estimate the parameters with sufficient accuracy for fast convergence of the algorithm [52].

In the CBM experiment tracks are measured by the MVD and STS detectors (see Chapter 2). MVD is a pixel detector, each pixel measures two independent coordinates  $x$  and  $y$ , so the measurement in MVD can be represented as a set of two independent one-dimensional measurements. The STS detector consists of doublesided microstrip modules, strips on each side measure position of the track independently, therefore STS provides two independent one-dimensional

measurements. This allows the data from both detectors to be represented in the same way. It is worth to notice, that in case of a one-dimensional measurement the covariance matrix degenerates into a scalar, that allows to speedup calculations and improve numerical stability.

Measurements depend linearly on the position of the particle in the plane of the detector, so the model of measurement has a form:

$$H_k = \{ \cos(\alpha_k), \sin(\alpha_k), 0, 0, 0 \}, \quad (4.11)$$

where in STS  $\alpha_k$  is an angle between a strip and the  $Y$ -axis, in MVD  $\alpha_k = 0$  for  $x$ -coordinate and  $90^\circ$  for  $y$ -coordinate.

The extrapolation matrix  $F_k$  describes the motion of a particle in the nonhomogeneous magnetic field  $\mathbf{B}$ . The trajectory of a particle is determined by the differential equations (4.9a–4.9c), which can be expressed in terms of the track parameters.

The obtained equations can not be solved exactly, since  $\mathbf{B}$  depends on the coordinates in a complicated way. Therefore it is necessary to use numerical methods for the solution of nonlinear differential equations. For the CBM experiment two methods are implemented for finding the parameters of the track in a point of  $(k + 1)$ -th measurement: the standard fourth order Runge-Kutta method [64] and analytic formula based on a Taylor series expansion [65]. The fourth order Runge-Kutta method has an accuracy of the order of  $O(qcB\Delta z/p)^5$ . Since the distance between stations is small enough ( $qcB\Delta z/p \approx 0.01$ ), the track parameters at the station of the next measurement can be found in one step of the Runge-Kutta method. The speed and accuracy of the analytic formula are determined by the number of terms in the Taylor series expansion. By choosing this number a compromise between accuracy and calculation time is achieved.

Thus, both methods allow to describe the evolution of the system:  $\langle \mathbf{r}_{k+1}^t \rangle = \mathbf{f}_k(\mathbf{r}_k^t)$ . This dependence is linearized by the formula (4.8) and the required extrapolation matrix  $F_k$  is obtained. The vector  $\mathbf{r}_k^{lin} = \{ x_k^+, y_k^+, t_{x\ k}^+, t_{y\ k}^+, q/p_0 \}$  is used as a linearization point. Here  $x_k^+, y_k^+, t_{x\ k}^+, t_{y\ k}^+$  are the track parameters after a previous filtration step and  $q/p_0$  is a value of the  $q/p$  parameter obtained at the stage of the state vector initialization.

For more efficient utilization of the CPU cache memory and acceleration of the algorithm approximation of the full magnetic field map is required. The magnetic field of the CBM experiment is smooth and, as a result, can be locally described by a polynomial function [52]. It has been shown that it is sufficient to use the polynomials of the fifth order to approximate the field in the plane of the stations. Since the reconstruction of the track needs only the value of the magnetic field along the trajectory of the particle the field between stations is described by a parabola with coefficients calculated using the three closest measurements of the track. It has been shown that such a simplified representation of the field does not affect the quality of the tracks reconstruction, but allows to achieve acceleration of the algorithm of 50 times [52].

The covariance matrix of the process noise  $Q_k$  describes multiple scattering of particles in the detector material. The tracking detectors of the CBM experiment will operate with thin modules (about 300  $\mu\text{m}$ ). In this case, scattering in the detector material occurs at small angles. At the same time, the amount of the material is big enough for the particle to have about 20 acts of scattering, which allows to describe the distribution of the total deviation angle of the particle by the Gaussian function according to the central limit theorem. A width of the deviation angle distribution can be described by the formula [66]:

$$\sigma(\theta) = \frac{13.6\text{MeV}}{\beta p} q \sqrt{S/X_0} [1 + 0.038 \ln(S/X_0)], \quad (4.12)$$

where  $p$  is a momentum of the particle,  $q$  is a charge of the particle in terms of the elementary charge,  $S$  is a thickness of the detector material, which is crossed by the particle,  $X_0$  is a radiation length of the material. Based on these assumptions the matrix  $Q_k$  can be obtained similarly to [67]:

$$Q_k = \sigma_k^2(\theta) \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & (t_x^2 + 1)t_r^2 & t_x t_y t_r^2 & 0 \\ 0 & 0 & t_x t_y t_r^2 & (t_y^2 + 1)t_r^2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (4.13)$$

Energy losses of the particle in the detector material are not taken into account as they are small comparing to other effects.

Usually Kalman filter implementation is based on calculations in double precision. Moving to single precision calculations allows to use more effectively the SIMD units of CPU and its cache memory, therefore in the package for track reconstruction in the tracking and vertexing detectors of the CBM experiment implementation in single precision is considered as a basic one. In this case, numerical rounding errors are increased by 8 orders of magnitude and part of the tracks can be fitted incorrectly.

To improve the stability of the fitting procedure following modification of the algorithm were implemented [52]: the accurate initial estimation of track parameters described above, stabilization of the calculation of the covariance matrix at the filtration step, a special procedure to account the first measurement. At the filtration step the numerical divergence can appear in case when the error of the extrapolated parameters exceeds a few times the error of the measurement. It was found that the direct comparison of the errors before the filtration step and neglecting the errors which are more then four times larger then the mean measurement error keeps the algorithm stable. A similar problem occurs when the first measurement is added, where the parameter errors are infinite and are set as large positive numbers. For this particular case, the filtration formulas can be simplified analytically and calculations with large numbers can be avoided. Note, that the implemented changes also positively influence the calculations in double precision.

Tracks in the CBM experiment are found by the Cellular Automaton (CA) method [68]. In presence of noise measurements in the detector, the track finder can pick them to a track, especially in case of low-momentum secondary tracks. Since the Kalman filter method assumes, that all measurements belong to the track, it treats all of them exactly in the same way. Thus, track parameters in case of the presence of noise hits will be distorted. In order to reduce an impact of noise hits on the track parameters the Deterministic Annealing Filter (DAF) can be used, which is based on the Kalman filter smoother. DAF introduces weights for measurements, that are changed independently. Gradually decreasing weights for noise hits, the method fits the track based on the correct measurements.

### 4.3 The matrix inversion problem

The algorithms based on the Kalman filter need the covariance matrix inversion. Since this procedure is common for several algorithms, let us describe it first.

All the reconstruction algorithms of the CBM experiment are intended to be implemented in single precision for efficient calculations with the SIMD units of CPU or for calculations with GPU. That's why numerical stability of the algorithms, including the method for the matrix inversion, is crucial. Since covariance matrix is symmetric, it is possible to apply the Cholesky decomposition [64]. This method has been chosen because of its numerical stability: the error of the decomposition is stable and does not depend on the pivoting.

If the matrix  $A = \{a_{ij}\}$  is symmetric and positively defined, it could be decomposed with the Cholecky method:

$$A = U^T U. \quad (4.14)$$

$U$  is a lower or upper triangular matrix. In the present work the upper case has been chosen. During the fit the diagonal elements of the covariance matrix can become negative. In order to apply the algorithm for a not positively defined matrix  $A$ , it is necessary to introduce an additional diagonal matrix  $D = \{d_{ii}\}$  into the decomposition:

$$A = U^T D U. \quad (4.15)$$

We have defined the diagonal elements of the matrix  $D$  to be equal  $\pm 1$ . With such definition the division by the element  $d_{ii}$  in further calculations is replaced by multiplication, that speeds up the algorithm and increases stability of the method. The formulas for the matrices  $U$  and  $D$  are derived directly from their definition:

$$\begin{aligned} u_{ii}'^2 &= a_{ii} - \sum_{k=0}^i u_{ki}^2 d_{kk}, \\ d_{ii} &= \text{sgn}(u_{ii}'^2), \\ u_{ii} &= \sqrt{d_{ii} u_{ii}'^2}, \\ u_{ij} &= \frac{d_{ii}}{u_{ii}} \left( a_{ij} - \sum_{k=0}^i u_{ki} u_{kj} d_{kk} \right), j = (i + 1) \dots 4. \end{aligned} \quad (4.16)$$

The next step is to invert the obtained matrices. As follows from (4.15) and from the definition of the matrix  $D$ :

$$A^{-1} = U^{-1}D(U^{-1})^T, \quad (4.17)$$

where the  $5 \times 5$  matrix  $U^{-1} = \{\hat{u}_{ij}\}$  is calculated by formulas:

$$\begin{aligned} \hat{u}_{ii} &= \frac{1}{u_{ii}}, i = 0 \dots 4, \\ \hat{u}_{ii+1} &= -u_{ii+1}\hat{u}_{ii}\hat{u}_{i+1i+1}, i = 0 \dots 3 \\ \hat{u}_{ii+2} &= \hat{u}_{ii+1}u_{ii}\hat{u}_{i+1i+2} - u_{ii+2}\hat{u}_{ii}\hat{u}_{i+2i+2}, i = 0 \dots 2 \\ \hat{u}_{ii+3} &= \hat{u}_{ii+2}u_{ii+2}\hat{u}_{i+2i+3} - u_{ii+3}\hat{u}_{ii}\hat{u}_{i+3i+3} - \\ &\quad \hat{u}_{ii+1}u_{i+1i+1}(\hat{u}_{i+1i+2}u_{i+2i+2}\hat{u}_{i+2i+3} - \hat{u}_{i+1i+3}), i = 0 \dots 1 \\ \hat{u}_{04} &= \hat{u}_{02}u_{22}\hat{u}_{24} - u_{04}\hat{u}_{00}\hat{u}_{44} + \hat{u}_{01}u_{11}(\hat{u}_{14} - \hat{u}_{13}u_{33}\hat{u}_{34} - \\ &\quad \hat{u}_{12}u_{22}\hat{u}_{24}) + \hat{u}_{34}u_{33}(\hat{u}_{03} - \hat{u}_{23}u_{22}(\hat{u}_{02} - \hat{u}_{01}u_{11}\hat{u}_{12})). \end{aligned}$$

The proposed method is implemented for inversion of matrices in all routines based on the Kalman filter described further. It provides numerically stable results and, due to the small number of operations, high speed of computation.

## 4.4 Kalman filter based smoother

Sometimes it is necessary to obtain an estimation of the track parameters at the position of each hit. For example, this problem occurs when the alignment of the detector system is needed, when deciding whether measurements in the track are noisy or not, when applying sophisticated algorithms for finding the parameters of the track, for example, the Deterministic Annealing Filter (see section 4.5). To find an optimal estimation  $\mathbf{r}_i$  of the desired vector  $\mathbf{r}_i^t$  in each measurement point  $\mathbf{m}_i$  according to all measurements or all other measurements, except the current, a KF smoother is used.

The standard approaches of the Kalman filter based smoother extrapolate the optimal estimate of the state vector of the system obtained in the last measurement point to the desired measurement point  $\mathbf{m}_i$ . This procedure can lead the

covariance matrix to become negatively defined through the numerical instability and numerical rounding errors [62]. Therefore an alternative approach of the smoother is implemented that is based on two Kalman filters that run in the forward and backward directions. This method obtains two sets of optimal estimations of the track parameters  $\mathbf{r}_i^t$  at a given position according to the first  $i$  and last  $(n - i)$  measurements and merges them. If the estimation of the track parameters is needed, that is based on all measurements except of the measurement at a given measurement, then estimation of the forward Kalman filter is based on the first  $(i - 1)$  measurements. Estimations obtained with forward and backward Kalman filters are merged also by the Kalman filter equations. Since both sets of parameters are taken already at the same point only the filtration step of the Kalman filter method is applied. One estimate is considered as a measurement and added to another one according to the formulas (4.6a–4.6e). In this case the matrix of the measurement model  $H$  is equal to the unit matrix. Formulas for the filtration step of the Kalman filter thus are simplified:

$$\begin{aligned}
 K &= V_1(V_1 + V_2)^{-1}, \\
 \mathbf{r} &= \mathbf{m}_1 + K(\mathbf{m}_2 - \mathbf{m}_1), \\
 C &= V_1 - KV_1, \\
 \chi^2 &= (\mathbf{m}_2 - \mathbf{m}_1)^T(V_1 + V_2)^{-1}(\mathbf{m}_2 - \mathbf{m}_1),
 \end{aligned}
 \tag{4.18}$$

where  $\mathbf{m}_1$  and  $V_1$  are the estimation of the parameters and its covariance matrix of the forward filter,  $\mathbf{m}_2$  and  $V_2$  are the estimation of the parameters and its covariance matrix of the backward filter.

In order to characterize the quality of the smoother residuals and pulls have been calculated. Residuals of the track parameters, for instance, of the  $x$ -coordinate of the track, are defined as:

$$\rho_x = x_{\text{reco}} - x_{\text{mc}},
 \tag{4.19}$$

where  $x_{\text{reco}}$  is a reconstructed value of the  $x$ -coordinate,  $x_{\text{mc}}$  is a true Monte-Carlo value of the coordinate. For characterization of the reliability of the filter and obtained estimation of the parameters a normalized to the covariance matrix residual (pull) is introduced. The are defined by the formula:

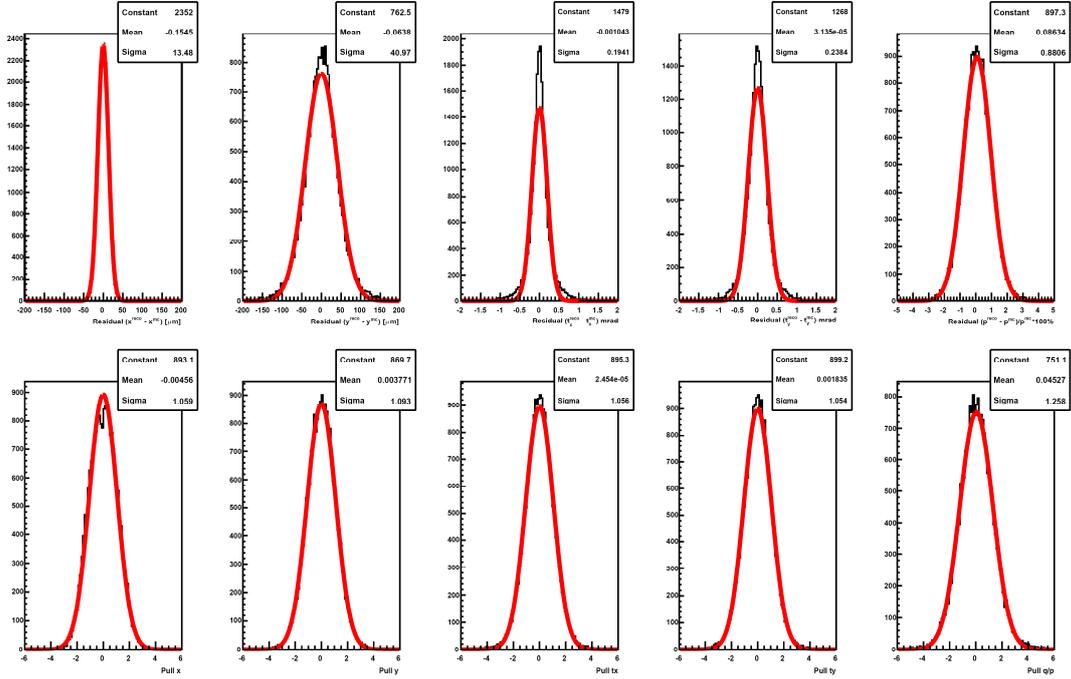


Figure 4.2: Distribution of the residuals and pulls for the track parameters obtained by the smoother at the first station of the STS detector of the CBM experiment.

Residuals										
Detector	MVD		STS							
Station	1	2	1	2	3	4	5	6	7	8
$x, \mu\text{m}$	4.89	4.75	13.5	12.3	13.1	14.2	14.5	15.2	14.1	15.0
$y, \mu\text{m}$	4.97	4.79	41.0	42.8	43.7	42.7	51.6	55.8	59.2	68.1
$t_x, 10^{-3}$	0.38	0.14	0.19	0.26	0.25	0.20	0.20	0.19	0.21	0.55
$t_y, 10^{-3}$	0.38	0.13	0.24	0.34	0.36	0.33	0.33	0.30	0.34	0.68
Resolution										
$\frac{(p_{mc} - p_r)}{p_{mc}}, \%$	0.88	0.88	0.88	0.88	0.88	0.88	0.89	0.88	0.87	0.87
Pulls										
$x$	1.01	1.00	1.06	1.09	1.09	1.08	1.07	1.05	1.06	1.02
$y$	1.02	1.00	1.09	1.09	1.09	1.10	1.11	1.11	1.06	1.06
$t_x$	1.08	1.04	1.06	1.14	1.16	1.25	1.23	1.29	1.42	1.13
$t_y$	1.09	1.02	1.05	1.14	1.16	1.17	1.18	1.21	1.34	1.18
$q/p$	1.26	1.26	1.26	1.26	1.25	1.25	1.25	1.24	1.22	1.21

Table 4.1: The standard deviations of the residuals and pulls distributions for the track parameters obtained at the point of each measurement of the track by the Kalman filter based smoother for the tracking system of the CBM experiment.

$$P_x = \rho_x / \sigma_x, \quad (4.20)$$

where  $\sigma_x$  is an estimation of the error of the parameter  $x$  obtained from the covariance matrix, i.e. a square root of the corresponding diagonal element of the covariance matrix. In the ideal case pull should be unbiased and distributed according to the Gaussian distribution with the standard deviation  $\sigma = 1$ .

For the evaluation of the KF smoother a geometry with two MVD and eight STS stations is used. Residuals and pulls of the state vector are calculated at the position of each measurement. All distributions are unbiased (see Fig. 4.2), the standard deviation of all distributions are given in Table 4.1. The smoother shows high quality of reconstruction: residuals at all points are in a good agreement with the errors of hits, the momentum resolution is about 1%, that fulfill the CBM requirements, all pulls have the standard deviation close to one. The pull distribution of the  $q/p$  parameter shows the maximum deviation from  $\sigma = 1$ , since momentum is the most sensitive parameter to approximations done in the material treatment and in the extrapolation within the nonhomogeneous magnetic field.

## 4.5 Deterministic Annealing Filter

The Deterministic Annealing Filter (DAF) was developed for the optimal track trajectory and parameters finding [69]. In case of the CBM experiment, when trajectories of charged particles are already found by the CA method and their parameters are estimated by the Kalman filter, the Deterministic Annealing Filter will be used to reduce the impact of noise measurements, matched to a track, on the final estimation of the track parameters.

The idea of DAF is to introduce a “temperature” depending weight to each hit and to perform several filtration iterations gradually “cooling” the system, i.e. decreasing the value of the temperature. At first, the Kalman filter based smoother with the same weights for all measurements predicts the track position at every layer of the detector. Based on these predictions the probability that a

hit belongs to the track and the weight of a hit are calculated. Decreasing the temperature the influence of noise hits is suppressed with each next iteration. On each next iteration weights are recalculated using smoothed track parameters from the previous iteration.

The Kalman filter mathematics has been modified in order to include weights of hits. During the iteration in DAF the same hit weights are used both for the forward and backward filters in the smoother. Weights are calculated by the formula:

$$p = \frac{1}{1 + \exp((\chi^2 - \chi_{\text{cut}}^2)/(2T))}, \quad (4.21)$$

where  $p$  is a weight of the hit on the station;  $T$  is the temperature of the current iteration, the values of  $T = 9, 4, 1$  and  $0.1$  are used in the algorithm;  $\chi^2$  is a squared distance between the hit and a smoothed track position normalized on the hit and track parameters errors;  $\chi_{\text{cut}}^2$  is a threshold, which determines the region, where the hit is accepted (the value of  $\chi_{\text{cut}} = 4$  has been used, that provides for the normal distribution the probability of 99,995% to attach a correct hit).

Station		unshifted	1 $\sigma_{\text{hit}}$	2 $\sigma_{\text{hit}}$	3 $\sigma_{\text{hit}}$
MVD	1	0.4	0.4	0.4	0.4
	2	0.7	0.7	0.7	0.7
STS	1	0.4	0.4	0.4	0.4
	2	0.4	0.4	0.4	0.4
	3	0.4	0.6	0.5	0.5
	4	0.5	61.2	95.6	100
	5	0.5	1.3	0.8	0.6
	6	0.5	0.6	0.5	0.5
	7	0.5	0.5	0.5	0.5
	8	0.1	0.1	0.1	0.1

Table 4.2: Percentage of rejected hits depending on the distance from the shifted hit to its Monte-Carlo position on the 4<sup>th</sup> STS station.

The evaluation of the algorithm have been performed on the server with two Xeon X5550 processors at 2.7 GHz and 8 MB L3 cache, one core has been used. A setup of two stations of the MVD and eight stations of the STS detectors has been used. For tests 20000 high-momentum primary tracks, reconstructed with the CA track finder, have been used. The tracks could have on average up to

0.5% incorrectly attached hits.

To test the algorithm noise measurements are simulated displacing hits on the 4<sup>th</sup> STS station by a certain amount of the hit errors ( $\sigma_x = 17 \mu\text{m}$ ,  $\sigma_y = 91 \mu\text{m}$ ) with respect to its Monte-Carlo position. The percentage of rejected hits has been calculated on all stations of the detectors. The obtained results are given in Table 4.2. As follows from the weight definition 4.21, the probability to reject a hit depends on its displacement from the track trajectory normalized on the error, that is described by pulls of the track. The pulls obtained by the smoother are distributed according to the unit normal distribution (see Fig. 4.2), therefore rejection probability should be also described by the normal distribution, that is proved by the obtained results. The execution time of DAF is 14  $\mu\text{s}$  per track.

## 4.6 KF track fit on many-core systems

Modern processing units have many-core architectures. Current development tends to increase further the number of cores per processing unit. It is of the particular importance that with this growth the performance of the programs should grow up proportionally. Therefore the scalability with respect to the number of cores is one of the most important parameters of the reconstruction algorithms.

In order to utilize the potential of modern computer architectures, reconstruction of charged particles parameters with the Kalman filter is implemented using all tools described in Chapter 3. Fig. 4.3 shows scalability of the KF track fit for different CPUs using Vc for vectorization together with OpenMP for parallelization or OpenCL for both vectorization and parallelization. In case of OpenMP the order of cores usage is set manually: at first, first physical core is filled, then the next cores from the same CPU are filled, then the next CPU is filled in the same order. With an Intel HT technology a stair-like structure appears, since running the second thread on the same physical core give about 30% of speedup (see section 3.1). In case of OpenCL it is impossible to control the thread to core affinity and the system order is used, therefore scalability picture looks different (green line in Fig. 4.3): first 32 points correspond to a one thread running on first 32 physical cores; next 32 points — gradually, on each of these cores a second

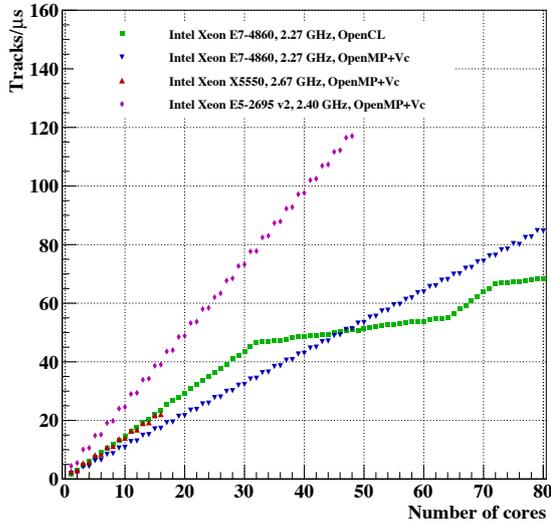


Figure 4.3: Scalability of the Kalman filter based track fit on many-core CPUs using Vc for vectorization together with OpenMP for parallelization or OpenCL for both vectorization and parallelization. The stair-like structure is caused by the Intel Hyper-Threading technology.

thread is run; then a one thread is run on still free 8 cores, and in the last part a second thread is run on these 8 cores. All implementations shows a strong linear behavior, which proves, that with increasing number of cores in CPU the performance does not degrade and the maximum speed is achieved.

Implementation of the KF track fit with different tools allows to run it not only on CPUs, but also on other architectures including Intel Xeon Phi and GPUs (see Fig. 4.4). The scalability obtained on the Intel Xeon Phi with up to four threads running simultaneously on each core is similar to CPU due to the similar architectures. In order to run the KF track fit on graphics cards OpenCL is used. OpenCL controls the load of the GPU and distributes working groups between compute units (or streaming multiprocessors). It is important that the program scales linearly within each compute unit with respect to the number of tasks in the group. Fig. 4.4 shows that the KF track fit scales linearly on the graphics cards up to the number of cores in one compute unit (for Nvidia GTX480 — 32, for AMD Radeon HD 7970 — 64). Then a drop appears, because when first 32 (for Nvidia) or 64 (for AMD) tasks are processed, only one task is left and all other cores of the compute unit are idle. Increasing the number of tasks in the group further the speed reaches the maximum with the number of tasks dividable by the number of cores in the compute unit. Due to an overhead in tasks distribution the maximum performance is reached when the number of tasks in the group is

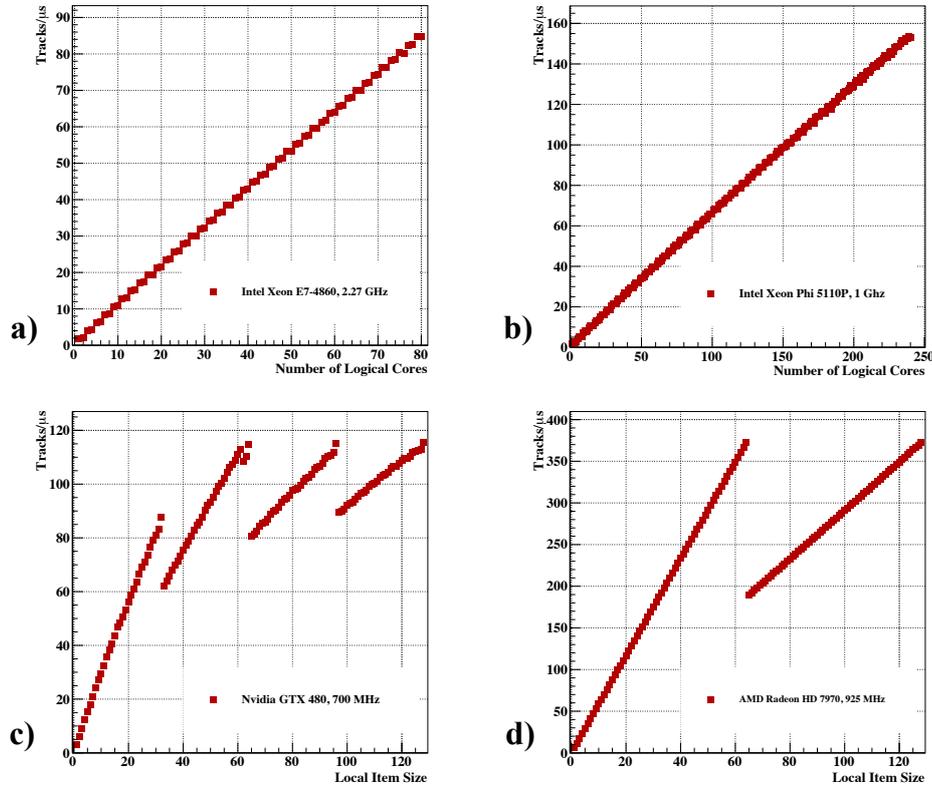


Figure 4.4: Scalability of the Kalman filter based track fit on different many-core processors: **a)** CPU (on an example of Intel Xeon E7-4860), **b)** Intel Xeon Phi, **c)** Nvidia GPU (on an example of Nvidia GTX 480) and **d)** ATI GPU (on an example of AMD Radeon HD 7970).

two-three times more than the number of cores in the compute unit.

The code running on the Xeon Phi 5110P (2 TFLOPS peak performance in single precision), Nvidia GTX480 (1.35 TFLOPS) and AMD Radeon HD 7970 (3.79 TFLOPS) shows 1.8, 1.4 and 4.4 times higher speed compared to four Intel Xeon E7-4860 CPUs (0.72 TFLOPS in total). The Kalman filter fit uses intensively load/store operations, therefore CPU benefits from a large amount of the fast on-chip memory comparing to the graphic cards and Xeon Phi. Also, all these architectures use different compilers with different optimization strategies.

Summarizing, implemented methods for track fit based on the Kalman filter assure high quality of the tracks parameters and their covariance matrices, that are used as an input for reconstruction of short-lived particles. The KF track fit is fast and show a strong linear scalability on many-core systems.

# Chapter 5

## Reconstruction of short-lived particles with KF Particle package

Particles, which are produced in a collision, can be divided into two groups: long-lived particles and short-lived particles. Long-lived particles have a lifetime large enough to cross the tracking detector system of the experiment and to be registered directly. This group includes stable particles, like electrons and positrons  $e^\pm$ , protons  $p^\pm$ , and particles with a large decay length  $c\tau$ , like muons  $\mu^\pm$ ,  $\pi^\pm$  and  $K^\pm$  mesons. Short-lived are those particles, which decay before or short within the tracking system and can be registered only indirectly. Short-lived particles include strange hyperons ( $\Lambda$ ,  $\Xi$ ,  $\Omega$ ), low mass vector mesons ( $\rho$ ,  $\omega$ ,  $\phi$ ), charmed particles (D mesons,  $J/\psi$ ), hypernuclei ( ${}^3_\Lambda\text{H}$ ,  ${}^4_\Lambda\text{H}$ ,  ${}^4_\Lambda\text{He}$ ), that are of the particular importance for the CBM physics program. Such particles can be reconstructed only indirectly by their decay products.

One of the most commonly used methods for reconstruction of short-lived particles is direct calculation of mother particle parameters based on parameters of daughter particles produced in a decay, for instance  $\Lambda \rightarrow p\pi^-$ . For this a point of the closest approach between daughter particles trajectories is found by numerical methods and parameters of the particles are extrapolated to this point. Then the obtained momentum and energy of daughters are summed up. Despite its simplicity, this method has several shortcomings:

- usually the information about the errors of daughter particles trajectories is not taken into account, that affects the reconstruction accuracy of the mother particle;
- not possible to calculate statistical criteria, which can be used for a background rejection;
- usually the methods for finding a point of the closest approach in a nonhomogeneous magnetic field have a low speed.

Another commonly used method is reconstruction of short-lived particles by the secondary vertex (decay point) fit with the Kalman filter method [70, 71]. The Kalman filter method itself is used to estimate a position of a decay vertex and the momentum of each daughter particle at this reconstructed vertex. The state vector is chosen to contain three coordinates of the vertex, three components of the mother particle momentum and three components of the momentum of each daughter particle:

$$r = \{x, y, z, p_x, p_y, p_z, p_x^{d1}, p_y^{d1}, p_z^{d1}, \dots, p_x^{dn}, p_y^{dn}, p_z^{dn}\}.$$

The covariance matrix of such a state vector allows to take into account correlations between parameters of all daughter particles with the mother particle and between each other. It is worth to notice, that the covariance matrix has a large size in case of three or more daughter particles, that requires more memory to be allocated.

A similar method for reconstruction of secondary vertices was implemented in the CBM experiment [72]. But because of its limitations:

- lack of flexibility with respect to the number of daughter particles;
- work with matrices of a large size, as a consequence — low speed and numerical instabilities;

it was replaced with the KF Particle package.

Existing implementations of these two methods do not use the advantages of parallel computer architectures, including the SIMD instruction set or multithreading, that leads to a low efficiency of the computational resources utilization.

## 5.1 The KF Particle package for the short-lived particle reconstruction

The KF Particle package [26, 27] has been developed for the complete reconstruction of short-lived particles with their momentum, energy, mass, lifetime, decay length, rapidity, etc. The package has following properties:

- is based on the Kalman filter mathematics;
- daughter and mother particles are described with the same set of parameters and are treated in exactly the same way;
- the package is geometry independent;
- daughter particles are added to the mother particle absolutely independently from each other;
- the package allows reconstruction of decay chains;
- implemented in single precision;
- fully SIMDized.

The package is based on the Kalman filter method with the state vector of the particle, that includes eight parameters: three coordinates of the particle ( $x, y, z$ ), three components of its momentum ( $p_x, p_y, p_z$ ), the energy of the particle ( $E$ ) and in case, when a production point of a particle is known, the time between production and decay points measured in a distance normalized on the momentum ( $s = l/p$ ):

$$\mathbf{r} = \{x, y, z, p_x, p_y, p_z, E, s\}. \quad (5.1)$$

The parameter  $s$  appears in the transport equations as a small parameter. These parameters are usually used for description of real physical objects and they do not depend on geometry of an experiment and its operational conditions. With this state vector KF Particle allows the full reconstruction of a particle: determination of the decay vertex, momentum and energy. The covariance matrix is estimated together with the state vector, that allows to obtain not only parameters but also their accuracy and calculate the value of the  $\chi^2$ -criterion, i.e. estimate the quality of reconstruction.

A measurement  $\mathbf{m}_k$  of the state vector in KF Particle is a state vector of a daughter particle. Since momentum and energy of daughter particles are summed up and only the vertex is fitted, the model of measurement, which is in the standard case defined by the equation 4.2, should be extended:

$$G_k \mathbf{m}^t = H_k \mathbf{r}_k^t, \quad (5.2)$$

where  $\mathbf{m}_k^t$  is a true value of the measurement,  $\mathbf{r}_k^t$  is a true value of the state vector of the mother particle. In this case the standard filtration equations (4.6a)–(4.6f) are modified as follows:

$$\begin{aligned} K_k &= C_k^- H_k^T (H_k C_k^- H_k^T + G_k V_k G_k^T)^{-1}, \\ \zeta_k &= G_k \mathbf{m}_k - H_k \mathbf{r}_k^-, \\ \mathbf{r}_k^+ &= \mathbf{r}_k^- + K_k \zeta_k, \\ C_k^+ &= (I - K_k H_k) C_k^-, \\ \chi_k^2 &= \chi_{k-1}^2 + \zeta_k^T (H_k C_k^- H_k^T + G_k V_k G_k^T)^{-1} \zeta_k. \end{aligned} \quad (5.3)$$

Note, that in case of the standard definition of the model of measurement ( $G_k$  is a unity matrix) equations (5.3) coincide with the standard Kalman filter equations (4.6a)–(4.6f).

Equations (5.3) give an optimum estimation of the state vector  $\mathbf{r}_k$  based on all previous measurements and a current measurement  $\mathbf{m}_k$ . But  $\mathbf{m}_k$  is a random variable and its optimal value  $\mathbf{m}_k^f$  can be also obtained with respect to all previous measurements. In case of short-lived particle reconstruction this means, that parameters of a  $k$ -th daughter particle can be improved with respect to the vertex position obtained on the previous  $(k-1)$ -th step, since it is assumed, that all daughter particles are produced at the same decay point. Equations, which gives the  $\mathbf{m}_k^f$  estimation are:

$$\begin{aligned} K_k^m &= V_k G_k^T (H_k C_k^- H_k^T + G_k V_k G_k^T)^{-1}, \\ \zeta_k &= G_k \mathbf{m}_k - H_k \mathbf{r}_k^-, \\ \mathbf{m}_k^f &= \mathbf{m}_k - K_k^m \zeta_k, \\ V_k^f &= (I - K_k^m G_k) \cdot V_k, \\ D_k^f &= K_k^m H_k C_k^-, \end{aligned} \quad (5.4)$$

where  $K_k^m$  is a gain matrix of a measurement,  $V_k^f$  is a covariance matrix of the desired optimum estimation  $\mathbf{m}_k^f$  of the measurement  $\mathbf{m}_k$ ,  $D_k^f$  is a covariance matrix between vectors  $\mathbf{m}_k^f$  and  $\mathbf{r}_k$ :

$$D_k(i, j) = \text{cov}(\mathbf{m}_k^f(i), \mathbf{r}_k(j)). \quad (5.5)$$

Taking into account equations (5.4) for the optimal estimation  $\mathbf{m}_k^f$  of the measurement, equations (5.3) are modified as follows:

$$\begin{aligned} K_k^f &= C_k^- H_k^T (H_k C_k^- H_k^T)^{-1}, \\ \zeta_k^f &= G_k \mathbf{m}_k^f - H_k \mathbf{r}_k^-, \\ \mathbf{r}_k^+ &= \mathbf{r}_k^- + K_k^f \zeta_k^f, \\ C_k^+ &= C_k^- - K_k^f (H_k C_k^- H_k^T - G_k V_k^f G_k^T) K_k^{fT}, \\ \chi_k^2 &= \chi_{k-1}^2 + \zeta_k^{fT} (H_k C_k^- H_k^T - G_k V_k^f G_k^T)^{-1} \zeta_k^f. \end{aligned} \quad (5.6)$$

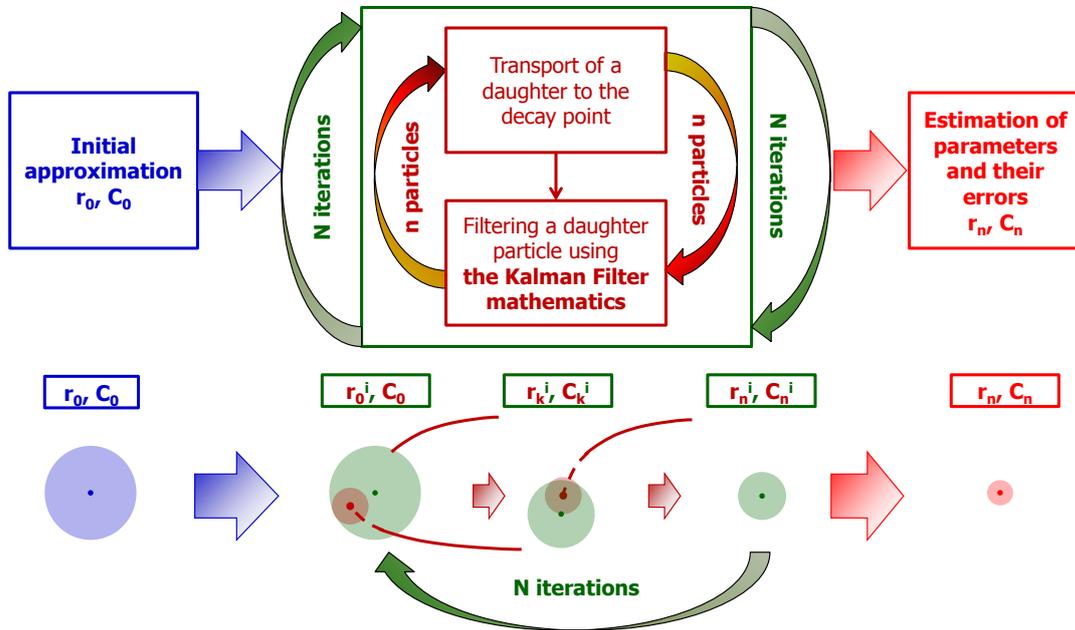


Figure 5.1: The block-diagram of short-lived particle reconstruction with the KF Particle package.

During reconstruction of short-lived particles with the KF Particle method a scheme similar to the standard Kalman filter method is used:

- initialization of the parameters;
- extrapolation of a measurement (daughter particle) to the point of the closest approach with a current estimation of the decay point;
- correction of the mother particle on a daughter particle according to the formulas (5.4) and (5.6).

The block-diagram of short-lived particle reconstruction with KF Particle is illustrated in Fig. 5.1.

## 5.2 Functionality of KF Particle

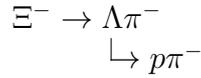
The KF Particle package provides the rich functionality for the users, including:

- construction of the mother particle from daughter particles;
- addition and subtraction of the daughter particle to (from) the mother particle together with the corresponding operators += and -=;
- access to the physics parameters of the particle together with their errors: mass, momentum, decay length, lifetime, rapidity, etc.;
- transport of the particle to an arbitrary point, to the decay and production points, to another particle, to a vertex, on the certain distance;
- calculation of a distance from a particle to a point, to another particle, to a vertex;
- calculation of a deviation in terms of  $\chi^2$  from a point, from another particle, from a vertex;
- calculation of the angle between two particles;
- the package allows to set constraints on the particle state vector on mass of the particle and on the production point of the particle.

The mass constraint allows to improve a mass and momentum information of the found particle, which is important, for instance, in decay chain reconstruction. The production point constraint helps to determine, whether the particle is coming from a certain point, for example, from the collision point, that is also important for decay chain reconstruction.

Despite the complexity of the mathematics, functionality of the package is simple and intuitive. For example, the simplest code for reconstruction of the  $\Xi^-$

decay chain



with KF Particle is easily understandable:

```
// Convert tracks into KF Particle objects using the PDG code
// (identifier of a particle specie introduced by PDG)
KFParticle pion1(kfptracks[0], -211); //pi-, PDG code is -211
KFParticle proton(kfptracks[1], 2212); //proton, PDG code is 2212
KFParticle pion2(kfptracks[2], -211); //pi-, PDG code is -211

// Construct Lambda-candidate
KFParticle Lambda;
const KFParticle* LambdaDaughters[2] = { &proton, &pion1 };
Lambda.Construct(LambdaDaughters, 2);
// Set a mass constraint on Lambda
Lambda.SetNonlinearMassConstraint(1.115683);

// Reconstruct Xi-
KFParticle Xi;
const KFParticle* XiDaughters[2] = { &Lambda, &pion2 };
Xi.Construct(XiDaughters, 2);
```

## 5.3 Development of KF Particle

The mathematics of KF Particle have to solve nonlinear problems, like extrapolation of charged particles in a nonhomogeneous magnetic field or setting the mass constraint. Such problems are usually solved with a number of approximations or by numerical methods, that can affect either reconstruction quality or speed of the algorithm.

For online reconstruction and selection of short-lived particles the CBM experiment requires the highest possible efficiency and speed. Therefore new approaches for such nonlinear problems, like search for a point of the closest approach and the mass constraint, have been developed.

### 5.3.1 Point of the closest approach between a particle and a vertex

A task of finding a point of the closest approach between a particle and a vertex is a common problem in short-lived particle reconstruction. For example, it appears, when a particle should be extrapolated to the primary vertex (the collision point) in order to determine whether this particle was produced in the collision directly or in a decay of another particle. Since the magnetic field in the CBM experiment is nonhomogeneous, the problem can be solved only by the approximate methods.

Let us consider the case of a particle motion in a constant one-component magnetic field  $B_z$ <sup>1</sup>. The equation of motion in the KF Particle package are parametrised with a parameter  $s = l/p$ , where  $l$  is a distance of propagation and  $p$  is a total momentum of the particle. Equations of motion of the particle with charge  $q$  assuming a constant one-component magnetic field  $B_z$  can be written in such parametrisation as:

$$\begin{aligned}
 x &= \left(x_0 + \frac{p_{y0}}{qB_z}\right) + \frac{p_{x0}}{qB_z} \sin(qB_z s) - \frac{p_{y0}}{qB_z} \cos(qB_z s), \\
 y &= \left(y_0 - \frac{p_{x0}}{qB_z}\right) + \frac{p_{x0}}{qB_z} \cos(qB_z s) + \frac{p_{y0}}{qB_z} \sin(qB_z s), \\
 z &= z_0 + sp_{z0}, \\
 p_x &= \frac{p_{x0}}{qB_z} - \frac{p_{x0}}{qB_z} \cos(qB_z s) + \frac{p_{y0}}{qB_z} \sin(qB_z s), \\
 p_y &= \frac{p_{x0}}{qB_z} - \frac{p_{x0}}{qB_z} \sin(qB_z s) - \frac{p_{y0}}{qB_z} \cos(qB_z s), \\
 p_z &= p_{z0},
 \end{aligned} \tag{5.7}$$

where  $\{x_0, y_0, z_0\}$  is an initial position of the particle,  $\{p_{x0}, p_{y0}, p_{z0}\}$  are initial components of it's momentum.

In order to find a point of the closest approach to a vertex with coordinates  $\{x_v, y_v, z_v\}$  a distance between a vertex and a particle should be minimized with respect to the parameter  $s$ :

$$R = \sqrt{(x_v - x)^2 + (y_v - y)^2 + (z_v - z)^2}$$

---

<sup>1</sup>The method was initially developed for the ALICE experiment with the main  $B_z$  component of the magnetic field.

$$\frac{dR}{ds} = 0,$$

which gives an equation:

$$(x_v - x)\frac{dx}{ds} + (y_v - y)\frac{dy}{ds} + (z_v - z)\frac{dz}{ds} = 0. \quad (5.8)$$

Derivatives  $\frac{dx}{ds}$ ,  $\frac{dy}{ds}$  and  $\frac{dz}{ds}$  are obtained by differentiating equations (5.7):

$$\begin{aligned} \frac{dx}{ds} &= p_{x0} \cos(qB_z s) + p_{y0} \sin(qB_z s), \\ \frac{dy}{ds} &= -p_{x0} \sin(qB_z s) + p_{y0} \cos(qB_z s), \\ \frac{dz}{ds} &= p_{z0}. \end{aligned} \quad (5.9)$$

Substituting equations (5.7) and (5.9) into (5.8) the equation on the parameter  $s$  is obtained:

$$\begin{aligned} &[(x_v - x_0)p_{y0} - (y_v - y_0)p_{x0} - p_t^2/(qB_z)] \sin(qB_z s) + \\ &[(x_v - x_0)p_{x0} + (y_v - y_0)p_{y0}] \cos(qB_z s) + (z_v - z_0)p_{z0} - p_{z0}^2 s = 0. \end{aligned} \quad (5.10)$$

The equation (5.10) is a transcendental equation and can not be solved explicitly. However, taking into account that for the majority of charged particles a point of the closest approach in the transverse  $XY$  plane, perpendicular to the field  $B_z$ , is close to a point of the closest approach in  $3D$  space, an approximate solution can be found. At first, the exact solution in the  $XY$  plane is found. For this  $z$ -component in equation (5.10) should be neglected:

$$\begin{aligned} a &= (x_v - x_0)p_{x0} + (y_v - y_0)p_{y0}, \\ b &= (x_v - x_0)p_{y0} - (y_v - y_0)p_{x0} - p_t^2/(qB_z), \\ a \cos(qB_z s_r) + b \sin(qB_z s_r) &= 0. \end{aligned} \quad (5.11)$$

A solution of the equation (5.11) is

$$s_r = \frac{1}{qB_z} \arctan \left( -\frac{(x_v - x_0)p_{x0} + (y_v - y_0)p_{y0}}{(x_v - x_0)p_{y0} - (y_v - y_0)p_{x0} - p_t^2/(qB_z)} \right). \quad (5.12)$$

The total solution in 3D space can be rewritten as a sum of a transverse ( $s_r$ ) and longitudinal ( $s_z$ ) terms:

$$s = s_r + s_z. \quad (5.13)$$

Substituting the expression (5.13) into the equation (5.10) and taking into account, that  $s_r$  is a solution of the equation (5.11), we obtain an equation on  $s_z$ :

$$\sin(qB_z s_z)[b \cos(qB_z s_r) - a \sin(qB_z s_r)] - p_{z0}^2 s_z + (z_v - z_0)p_{z0} - p_{z0}^2 s_r = 0. \quad (5.14)$$

Since  $s_z$  is assumed to be small,  $\sin(qB_z s_z) \approx qB_z s_z$  and the equation (5.14) can be significantly simplified. The solution for  $s_z$  in this case is:

$$s_z = \frac{(p_{z0} s_r - (z_v - z_0))p_{z0}}{qB_z [b \cos(qB_z s_r) - a \sin(qB_z s_r)] - p_{z0}^2}. \quad (5.15)$$

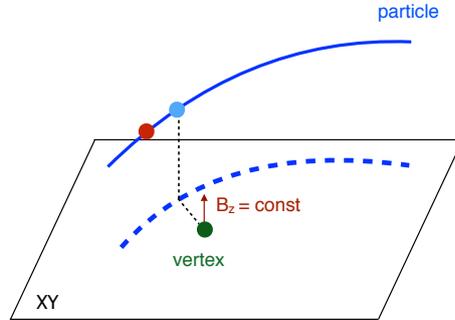


Figure 5.2: Determination of the point of the closest approach between a particle and a vertex assuming constant one-component magnetic field  $B_z$ : at first, a point in  $XY$  plane is found (the blue point), then a small correction on  $z$  coordinate is added (the red point).

The described procedure is illustrated by Fig. 5.2.

When the value of  $s$  is found by (5.13), the second iteration is performed in order to increase the accuracy. At first, parameters of the particle are transported on the value  $s$  according to equations of motion (5.7). Assuming, that the already

found solution is close to the exact value, a correction on the second iteration  $\Delta s$  should be small. When  $\Delta s \rightarrow 0$  we can write, that  $\sin(qB_z\Delta s) \approx qB_z\Delta s$  and  $\cos(qB_z\Delta s) \approx 1$ . Under such assumptions the equation (5.10) can be rewritten:

$$\begin{aligned} & [(x_v - x_{02})p_{y02} - (y_v - y_{02})p_{x02} - p_t^2/(qB_z)] \sin(qB_z\Delta s) + \\ & [(x_v - x_{02})p_{x02} + (y_v - y_{02})p_{y02}] \cos(qB_z\Delta s) + \\ & (z_v - z_{02})p_{z02} \cos(qB_z\Delta s) - p_{z02}^2/(qB_z) \sin(qB_z\Delta s) = 0, \end{aligned} \quad (5.16)$$

where index 2 stands for the propagated parameters. Collecting all terms near  $\cos(qB_z\Delta s)$  and  $\sin(qB_z\Delta s)$  we obtain the correction:

$$\Delta s = \frac{1}{qB_z} \arctan \left( \frac{(x_v - x_{02})p_{x02} + (y_v - y_{02})p_{y02} + (z_v - z_{02})p_{z02}}{(x_v - x_{02})p_{2y} - (y_v - y_{02})p_{x02} - p^2/(qB_z)} \right). \quad (5.17)$$

The total value of the parameter  $s$  after two iterations is calculated by the formula:

$$s = s_r + s_z + \Delta s. \quad (5.18)$$

In the CBM experiment the main component of the magnetic field is  $B_y$ . In the region of the STS detector two other components contribute to the total value of the magnetic field at the less then 10% level. In addition, the absolute value of the magnetic field within STS changes on the less then 10% level. Therefore an assumption of the constant one-component magnetic field works well. Since the method was initially developed for the ALICE experiment with the main component of the magnetic field  $B_z$ , the coordinate system is rotated such that  $B_y \rightarrow B_z$  and the described mathematics is applied for CBM.

### 5.3.2 Point of the closest approach between two particles

The problem of finding a point of the closest approach between two particles appears when searching for the initial approximation of the short-lived particles decay point. An accurate initial approximation allows to increase a numerical stability of the Kalman filter method, especially in single precision calculations, allows to improve quality of the reconstructed parameters. It also allows to reduce

a number of the Kalman filter iterations and in such a way significantly (two-three times) speed up the algorithm.

Assuming constant one-component magnetic field  $B_z$  equations of motion of each particle are described by formulas (5.7). Such assumptions are valid in case of the smooth behavior of the magnetic field and small extrapolation distances. Parameters of the first and second particles will be further denoted with indices 1 and 2 respectively. It is worth to notice, that the system is symmetric with respect to indices, since particles are absolutely equivalent, therefore their permutation does not change the solution.

The point of the closest approach for two particles can be found minimizing the distance between their trajectories:

$$R = \sqrt{(x_1(s_1) - x_2(s_2))^2 + (y_1(s_1) - y_2(s_2))^2 + (z_1(s_1) - z_2(s_2))^2}$$

$$\frac{dR}{ds_1} = 0,$$

$$\frac{dR}{ds_2} = 0,$$

which gives a system of equations on parameters  $s_1$  and  $s_2$  for the first and second particles respectively:

$$\begin{cases} (x_1 - x_2) \frac{dx_1}{ds_1} + (y_1 - y_2) \frac{dy_1}{ds_1} + (z_1 - z_2) \frac{dz_1}{ds_1} = 0 \\ (x_1 - x_2) \frac{dx_2}{ds_2} + (y_1 - y_2) \frac{dy_2}{ds_2} + (z_1 - z_2) \frac{dz_2}{ds_2} = 0 \end{cases} \quad (5.19)$$

Similarly to the case of a particle and a vertex, it is possible to find an exact solution in the transverse  $XY$  plane and then add a small correction on  $z$  coordinate. The total value of parameter  $s$  for both particles will again consist of two terms in this case: transverse  $s_r$  and longitudinal  $s_z$ :

$$\begin{cases} s_1 = s_{r1} + s_{z1} \\ s_2 = s_{r2} + s_{z2} \end{cases} \quad (5.20)$$

At first, terms of the equations (5.19) containing  $z$  coordinate are not taken into account, that gives a system of equations on the transverse components  $s_{r1}$  and

$s_{r2}$ :

$$\begin{cases} (x_1 - x_2) \frac{dx_1}{ds_1} + (y_1 - y_2) \frac{dy_1}{ds_1} = 0 \\ (x_1 - x_2) \frac{dx_2}{ds_2} + (y_1 - y_2) \frac{dy_2}{ds_2} = 0 \end{cases} \quad (5.21)$$

The solution of the system is rather complex, therefore for simplicity, let us introduce notations:

$$\begin{aligned} \Delta x_0 &= x_{01} - x_{02}, & \Delta y_0 &= y_{01} - y_{02} \\ P_{12} &= p_{x1}p_{x2} + p_{y1}p_{y2}, \\ k_{11} &= q_2 B_z (\Delta x_0 p_{x01} + \Delta y_0 p_{y01}) - (p_{x01}p_{y02} - p_{x02}p_{y01}), \\ k_{21} &= q_1 B_z [q_2 B_z (\Delta x_0 p_{y1} - \Delta y_0 p_{x1}) - P_{12}] + q_2 B_z p_{t1}^2, \\ k_{12} &= q_1 B_z (\Delta x_0 p_{x02} + \Delta y_0 p_{y02}) - (p_{x01}p_{y02} - p_{x02}p_{y01}), \\ k_{22} &= q_2 B_z [q_1 B_z (\Delta x_0 p_{y2} - \Delta y_0 p_{x2}) - P_{12}] - q_1 B_z p_{t2}^2, \\ k_p &= q_2 B_z (\Delta x_0 p_{y1} - \Delta y_0 p_{x1}) - q_1 B_z (\Delta x_0 p_{y2} - \Delta y_0 p_{x2}) - P_{12}, \\ k_d &= 0.5(\Delta x_0^2 + \Delta y_0^2) q_1 q_2 B_z^2 + k_p, \\ c_1 &= -(q_1 k_d + q_2 p_{t1}^2) B_z, \\ c_2 &= (q_2 k_d + q_1 p_{t2}^2) B_z, \\ d_1 &= \sqrt{p_{t1}^2 p_{t2}^2 - k_d^2}, \\ d_2 &= \sqrt{p_{t1}^2 p_{t2}^2 - k_d^2}. \end{aligned} \quad (5.22)$$

The solution of equations (5.21) gives two points of the minimum distance between two particle trajectories. Using notations (5.22) these solutions for curved trajectories ( $qB_z \neq 0$ ) can be written as:

$$\begin{cases} s_{r1}^1 = \frac{1}{q_1 B_z} \arctan \left( \frac{q_1 B_z (k_{11} c_1 + k_{21} d_1)}{q_1^2 B_z^2 k_{11} d_1 - k_{21} c_1} \right) \\ s_{r2}^1 = \frac{1}{q_2 B_z} \arctan \left( \frac{q_2 B_z (k_{12} c_2 + k_{22} d_2)}{q_2^2 B_z^2 k_{12} d_2 - k_{22} c_2} \right) \end{cases} \quad (5.23)$$

and

$$\begin{cases} s_{r1}^2 = \frac{1}{q_1 B_z} \arctan \left( \frac{q_1 B_z (k_{11} c_1 - k_{21} d_1)}{-q_1^2 B_z^2 k_{11} d_1 - k_{21} c_1} \right) \\ s_{r2}^2 = \frac{1}{q_2 B_z} \arctan \left( \frac{q_2 B_z (k_{12} c_2 - k_{22} d_2)}{-q_2^2 B_z^2 k_{12} d_2 - k_{22} c_2} \right) \end{cases} \quad (5.24)$$

In case if one of the trajectories is a straight line ( $qB_z = 0$ ), a corresponding equation in the solution (5.23) should be replaced by equations:

$$\begin{cases} s_{r1}^1 = -\frac{k_{11} c_1 + k_{21} d_1}{k_{21} c_1}, & q_1 B_z = 0 \\ s_{r2}^1 = -\frac{k_{12} c_2 + k_{22} d_2}{k_{22} c_2}, & q_2 B_z = 0 \end{cases} \quad (5.25)$$

and in the solution (5.24) — by equations:

$$\begin{cases} s_{r1}^2 = -\frac{k_{11} c_1 - k_{21} d_1}{k_{21} c_1}, & q_1 B_z = 0 \\ s_{r2}^2 = -\frac{k_{12} c_2 - k_{22} d_2}{k_{22} c_2}, & q_2 B_z = 0 \end{cases} \quad (5.26)$$

The approximation to the decay point should be selected among these two solutions. At first, both of them are checked to be well separated within the errors. For this parameters of the tracks together with their covariance matrices are transported on the values

$$\begin{cases} s_{r1} = (s_{r1}^1 + s_{r1}^2)/2 \\ s_{r2} = (s_{r2}^1 + s_{r2}^2)/2 \end{cases} \quad (5.27)$$

If trajectories of particles are overlapping within  $5\sigma$  errors at the obtained points, solutions (5.23) and (5.24) are considered to be equivalent and for further calculations values (5.27) are selected. If parameters are well separated within the errors, particles are transported using both solutions (5.23) and (5.24), then at each point a distance between particles is calculated taking into account all three coordinates and a point with a smaller distance is selected for further calculations.

Substituting formulas (5.20), where  $s_{r1}$  and  $s_{r2}$  are selected values, into equations (5.19) and using expressions for derivatives (5.9) a system of equations on

$s_{z1}$  and  $s_{z2}$  is obtained:

$$\left\{ \begin{array}{l} \frac{1}{q_2 B_z} \left[ -\frac{1}{q_1 B_z} a_{r1} \sin(q_1 B_z s_{z1}) + \right. \\ (\cos(q_2 B_z s_{z2}) \sin(q_1 B_z s_{z1}) - \cos(q_1 B_z s_{z1}) \sin(q_2 B_z s_{z2})) b_r + \\ (\cos(q_1 B_z s_{z1}) - \cos(q_2 B_z s_{z2}) \cos(q_1 B_z s_{z1}) - \sin(q_1 B_z s_{z1}) \sin(q_2 B_z s_{z2})) c_r \left. \right] + \\ c_z p_{z01} + p_{z01}^2 s_{z1} - p_{z01} p_{z02} s_{z2} = 0 \\ \frac{1}{q_1 B_z} \left[ -\frac{1}{q_2 B_z} a_{r2} \sin(q_2 B_z s_{z2}) + \right. \\ (\cos(q_2 B_z s_{z2}) \sin(q_1 B_z s_{z1}) - \cos(q_1 B_z s_{z1}) \sin(q_2 B_z s_{z2})) b_r + \\ (\cos(q_2 B_z s_{z2}) - \cos(q_2 B_z s_{z2}) \cos(q_1 B_z s_{z1}) - \sin(q_1 B_z s_{z1}) \sin(q_2 B_z s_{z2})) c_r \left. \right] + \\ c_z p_{z02} + p_{z01} p_{z02} s_{z1} - p_{z02}^2 s_{z2} = 0, \end{array} \right. \quad (5.28)$$

where

$$\begin{aligned} x_{r1} &= \sin(q_1 B_z s_{r1}) p_{x01} - \cos(q_1 B_z s_{r1}) p_{y01}, \\ y_{r1} &= \cos(q_1 B_z s_{r1}) p_{x01} + \sin(q_1 B_z s_{r1}) p_{y01}, \\ x_{r2} &= \sin(q_2 B_z s_{r2}) p_{x02} - \cos(q_2 B_z s_{r2}) p_{y02}, \\ y_{r2} &= \cos(q_2 B_z s_{r2}) p_{x02} + \sin(q_2 B_z s_{r2}) p_{y02}, \\ \Delta \bar{x}_0 &= \Delta x_0 B_z^2 q_1 q_2 + B_z q_2 p_{y01} - B_z q_1 p_{y02}, \\ \Delta \bar{y}_0 &= \Delta y_0 B_z^2 q_1 q_2 - B_z q_2 p_{x01} + B_z q_1 p_{x02}, \\ a_{r1} &= \Delta \bar{x}_0 x_{r1} + \Delta \bar{y}_0 y_{r1}, \\ a_{r2} &= \Delta \bar{x}_0 x_{r2} + \Delta \bar{y}_0 y_{r2}, \\ b_r &= x_{r1} x_{r2} + y_{r1} y_{r2}, \\ c_r &= x_{r2} y_{r1} - x_{r1} y_{r2}, \\ c_z &= (z_{01} - z_{02}) + s_{r1} p_{z01} - s_{r2} p_{z02}. \end{aligned} \quad (5.29)$$

Since  $s_z$  is assumed to be small, these equations can be expanded into the Taylor series neglecting terms smaller than  $o(s_z)$ . Under such assumption  $\sin(q_{1,2} B_z s_{z1,2}) \approx q_{1,2} B_z s_{z1,2}$ ,  $\cos(q_{1,2} B_z s_{z1,2}) \approx 1$  and  $\sin(q_1 B_z s_{z1}) \sin(q_1 B_z s_{z1}) \approx 0$  and the system (5.28) transforms into a system of linear equations:

$$\left\{ \begin{array}{l} (-a_{r1} + q_1 B_z b_r + q_2 B_z p_{z01}^2) s_{z1} - q_2 B_z (b_r + p_{z01} p_{z02}) s_{z2} = -q_2 B_z c_z p_{z01} \\ q_1 B_z (b_r - p_{z01} p_{z02}) s_{z1} + (a_{r2} - q_2 B_z b_r - q_1 B_z p_{z02}^2) s_{z2} = -q_1 B_z c_z p_{z02}. \end{array} \right. \quad (5.30)$$

The solution of the equations (5.30) is:

$$\begin{aligned} s_{z1} &= -c_z(q_2 B_z p_{z01} k_{z22} - q_1 B_z p_{z02} k_{z12})/\Delta, \\ s_{z2} &= -c_z(q_1 B_z p_{z02} k_{z11} - q_2 B_z p_{z01} k_{z21})/\Delta, \end{aligned} \quad (5.31)$$

where

$$\begin{aligned} k_{z11} &= -a_{r1} + q_1 B_z b_r + q_2 B_z p_{z01}^2, \\ k_{z12} &= -q_2 B_z (b_r + p_{z01} p_{z02}), \\ k_{z21} &= q_1 B_z (b_r + p_{z01} p_{z02}), \\ k_{z22} &= -a_{r2} - q_2 B_z b_r - q_1 B_z p_{z02}^2, \\ \Delta &= k_{z11} k_{z22} - k_{z12} k_{z21} \end{aligned} \quad (5.32)$$

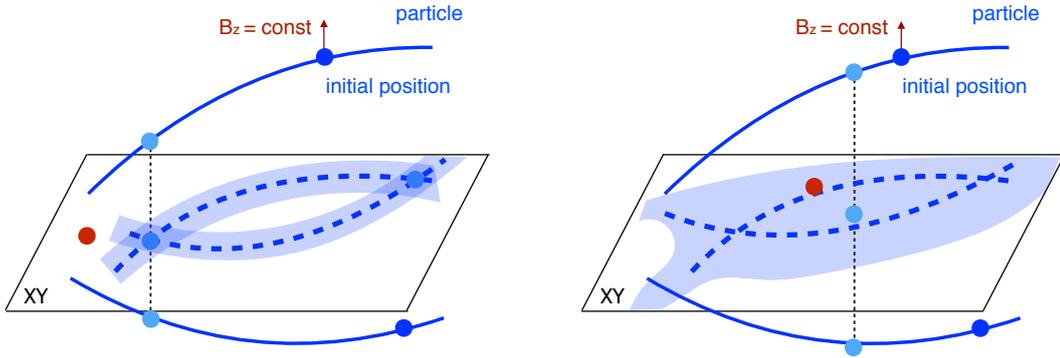


Figure 5.3: Determination of the point of the closest approach between two particles assuming constant one-component magnetic field  $B_z$ . At first, two points in  $XY$  plane are found. If the points are well separated within  $5\sigma$  errors, then the point where two particles are closer taking into account the  $z$  coordinate is selected. If within the errors points are not separated, the middle point is selected. Selected point is shown in light-blue. Then this point is corrected on  $z$  coordinate. The corrected point is shown in red.

The final approximation for the point of the closest approach is calculated by (5.20). The whole procedure is illustrated on Fig. 5.3.

Like in the previous section, the mathematics can be applied to the CBM experiment rotating the coordinate system  $B_y \rightarrow B_z$  under the assumption, that a point of the closest approach is close to the initial positions of daughter particles:

if a mother particle decays between the stations the distance should be less than 10 cm, if it decays in the target region — about 30 cm.

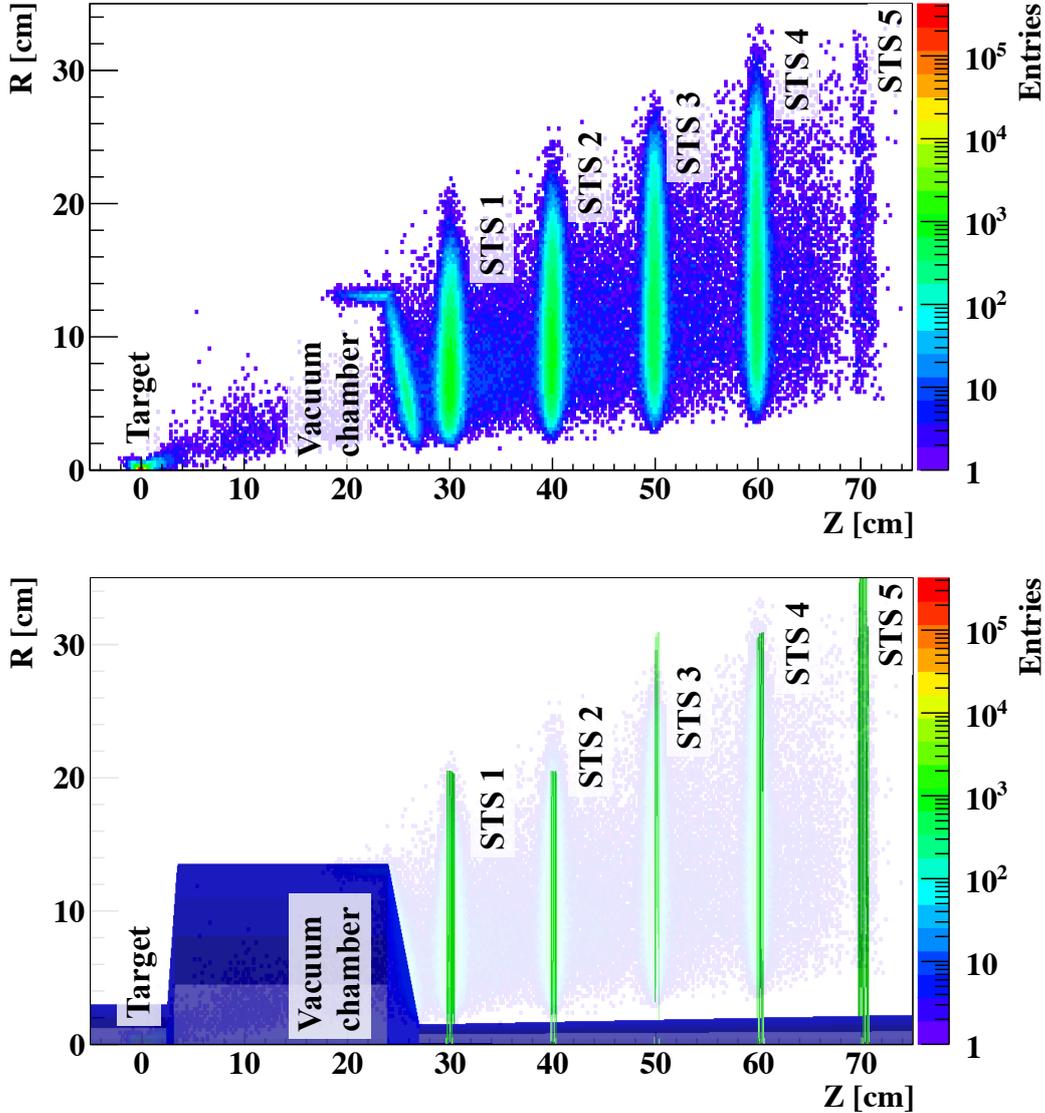


Figure 5.4: Determination of the distance of the closest approach between two particles at an example of  $\gamma \rightarrow e^+e^-$  conversion reconstruction. For illustration a distribution of reconstructed conversion points of true  $\gamma$  in  $ZR$  coordinates is shown (top), where  $R = \sqrt{x^2 + y^2}$ . From the comparison with the scheme of the detector geometry, drawn over the reconstructed  $\gamma$ -position (bottom), the material structure of the detector is clearly seen: the target, the vacuum chamber and five stations of STS. Due to the acceptance of the tracking system the material is not seen fully, but partially.

To evaluate the method  $\gamma$ -particles reconstruction through  $e^+e^-$  conversion

was investigated using the STS detector only. The  $\gamma$  was chosen since electrons are the most sensitive particles to the introduced changes: they have the smallest mass and therefore the biggest curvature of tracks. Another reason why  $\gamma$  conversion is sensitive to the quality of the initial vertex position approximation is that electron and positron trajectories are parallel in the conversion point. Therefore trajectories of  $e^+$  and  $e^-$  are crossing within the errors along several cm in  $z$ -direction, which gives several local minima for the Kalman filter algorithm and the final result of the vertex fit can converge to one of these local minima instead of the global one.

The electron pairs from true  $\gamma$ -conversions were selected using Monte Carlo information. The electron identification was also done using a Monte Carlo information. For illustration a distribution of reconstructed points of conversion in  $ZR$  coordinates is shown in Fig. 5.4, where  $R = \sqrt{x^2 + y^2}$ . The developed method for the point of the closest approach determination shows a high spatial resolution of the vertex reconstruction: a position of the target about 0 cm, five stations of the STS detector at 30, 40, 50, 60 and 70 cm (further stations are not seen since the track reconstruction algorithm requires hits on at least four consecutive stations) and the vacuum chamber material at about 25 cm are clearly seen, that allows to perform  $\gamma$ -tomography of the detector.

### 5.3.3 Nonlinear mass constraint

The mass constraint in case of the state vector (5.1) is a particularly nonlinear problem. The condition on the parameters of the state vector in case of the mass constraint can be written as:

$$f(\mathbf{r}) = E^2 - p_x^2 - p_y^2 - p_z^2 - m_0^2 = 0, \quad (5.33)$$

where  $m_0$  is a value of the mass to be set. In case of linearisation of such condition, that is a usual procedure for the mass constraint, the desired effect on the state vector can not be reached: together with a peak at  $m_0$  there will be tails around. The only solution which will guarantee the exact constraint on the particle mass is a nonlinear constraint.

The method developed in the current section is based on [73]. Using matrix

notation the condition (5.33) can be rewritten:

$$f(\mathbf{r}) = \mathbf{r}^T M \mathbf{r} - m_0^2 = 0 \quad (5.34)$$

with a state vector  $\mathbf{r}$  defined by (5.1) without the last parameter  $s$  and a matrix  $M$ :

$$M = \begin{pmatrix} \hat{O} & \hat{O} & 0 \\ \hat{O} & -I & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (5.35)$$

where  $\hat{O}$  is the null  $3 \times 3$  matrix and  $I$  is the unity  $3 \times 3$  matrix.

The Lagrangian for the constraint will be:

$$L = (\hat{\mathbf{r}} - H\mathbf{r})^T (\hat{\mathbf{r}} - H\mathbf{r}) + \lambda f(\hat{\mathbf{r}}), \quad (5.36)$$

where  $\mathbf{r}$  is an initial state vector,  $\hat{\mathbf{r}}$  is a final state vector with a mass constraint set on. Minimizing the Lagrangian we obtain an expression for an updated state vector  $\hat{\mathbf{r}}$ :

$$\frac{\partial L}{\partial \hat{\mathbf{r}}} = 0 \Rightarrow -H^T \mathbf{r} + (H^T H + \lambda M) \hat{\mathbf{r}} = 0 \Rightarrow \hat{\mathbf{r}} = (H^T H + \lambda M)^{-1} H^T \mathbf{r} \quad (5.37)$$

Since the form of the state vector will not change, the model of measurement is the unity matrix:  $H = I$ . Then the expression (5.37) can be significantly simplified:

$$\hat{\mathbf{r}} = (I + \lambda M)^{-1} \mathbf{r} \quad (5.38)$$

Since a constrained state vector should satisfy condition (5.34), we can obtain a matrix equation on  $\lambda$ :

$$\mathbf{r}^T((I + \lambda M)^{-1})^T M(I + \lambda M)^{-1} \mathbf{r} - m_0^2 = 0. \quad (5.39)$$

Taking into account, that the matrix  $M$  is symmetric, we can write:

$$\mathbf{r}^T(I + \lambda M)^{-1} M(I + \lambda M)^{-1} \mathbf{r} - m_0^2 = 0, \quad (5.40)$$

where

$$(I + \lambda M)^{-1} = \begin{pmatrix} \frac{1}{1-\lambda} & 0 & 0 & 0 \\ 0 & \frac{1}{1-\lambda} & 0 & 0 \\ 0 & 0 & \frac{1}{1-\lambda} & 0 \\ 0 & 0 & 0 & \frac{1}{1+\lambda} \end{pmatrix} \quad (5.41)$$

Substituting the expression for the matrix  $M$  into the equation (5.40), we obtain the equation on  $\lambda$ :

$$\begin{cases} -m_0^2 \lambda^4 + a \lambda^2 + b \lambda + c = 0, \\ a = (E^2 - p^2 + 2m_0^2), \\ b = -2(E^2 + p^2), \\ c = (E^2 - p^2 - m_0^2), \end{cases} \quad (5.42)$$

where  $p^2 = p_x^2 + p_y^2 + p_z^2$ .

In general, the equation (5.42) can not be solved exactly, but it can be solved numerically, for instance, by the Newton's method. The root of the equation is found iteratively with a root on an iteration  $n$  calculated based on the previous iteration:

$$\lambda_n = \lambda_{n-1} - \frac{f(\lambda_{n-1})}{f'(\lambda_{n-1})}, \quad (5.43)$$

where, as follows from (5.42):

$$\begin{aligned} f &= -m_0^2 \lambda^4 + a\lambda^2 + b\lambda + c, \\ f' &= \frac{df}{d\lambda} = -4m_0^2 \lambda^3 + 2a\lambda + b. \end{aligned} \quad (5.44)$$

The mass correction is assumed to be small, therefore as an initial value we can take a smallest positive root of the quadratic equation obtained from the equation (5.42) neglecting a  $\lambda^4$  term.

When the value of  $\lambda$  is found, parameters of the state vector (5.1) should be corrected according to the formula (5.38). The covariance matrix also should be corrected:

$$\begin{aligned} \hat{C} &= J^T C J, \\ J &= \frac{\partial \hat{r}}{\partial r}, \end{aligned} \quad (5.45)$$

where  $\hat{C}$  is a corrected covariance matrix,  $C$  is an initial covariance matrix. A Jacobian matrix  $J$  has  $J_{xx} = J_{yy} = J_{zz} = 1$  and

$$\begin{aligned} J_{p_i p_i} &= \frac{\partial \hat{p}_i}{\partial p_i} = \frac{1}{1-\lambda} + \frac{p_i}{(1-\lambda)^2} \frac{\partial \lambda}{\partial p_i}, \\ J_{p_i p_j} &= \frac{\partial \hat{p}_i}{\partial p_j} = \frac{p_i}{(1-\lambda)^2} \frac{\partial \lambda}{\partial p_j}, \quad i \neq j \\ J_{p_i E} &= \frac{\partial \hat{p}_i}{\partial E} = \frac{p_i}{(1-\lambda)^2} \frac{\partial \lambda}{\partial E}, \\ J_{E p_j} &= \frac{\partial \hat{E}}{\partial p_j} = -\frac{E}{(1+\lambda)^2} \frac{\partial \lambda}{\partial p_j}, \\ J_{EE} &= \frac{\partial \hat{E}}{\partial E} = \frac{1}{1+\lambda} - \frac{E}{(1+\lambda)^2} \frac{\partial \lambda}{\partial E}, \quad i, j = x, y, z. \end{aligned} \quad (5.46)$$

with all other elements equal to zero. The partial derivatives can be calculated as:

$$\begin{aligned} \frac{\partial \lambda}{\partial p_i} &= \frac{\partial f}{\partial p_i} / \frac{\partial f}{\partial \lambda}, \quad i = x, y, z \\ \frac{\partial \lambda}{\partial E} &= \frac{\partial f}{\partial E} / \frac{\partial f}{\partial \lambda}, \end{aligned} \quad (5.47)$$

where  $\frac{\partial f}{\partial p_i}$  and  $\frac{\partial f}{\partial E}$  are obtained differentiating the equation (5.42):

$$\begin{aligned}\frac{\partial f}{\partial p_i} &= -2p_i(1 + \lambda)^2, \quad i = x, y, z \\ \frac{\partial f}{\partial E} &= 2E(1 - \lambda)^2\end{aligned}\tag{5.48}$$

and  $\frac{\partial f}{\partial \lambda}$  defined by the equation (5.44)

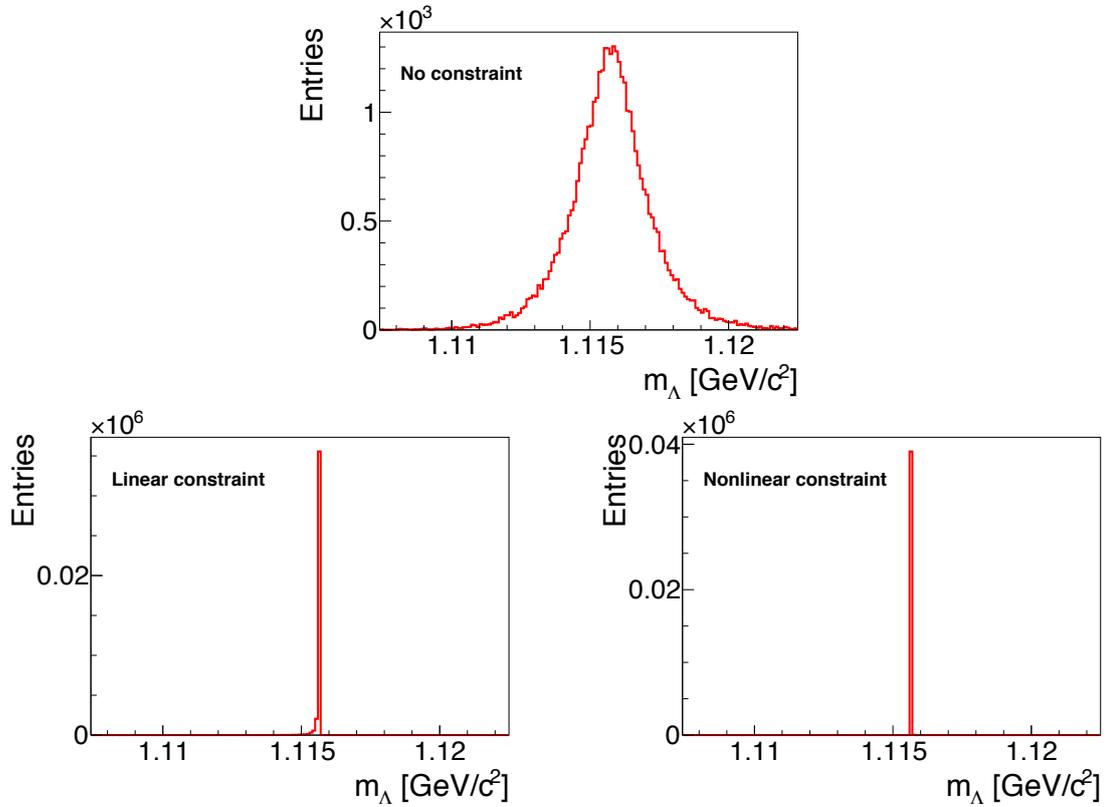
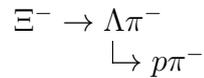


Figure 5.5: Comparison of linear and nonlinear mass constraints. Linear and nonlinear constraints are set on the reconstructed mass of  $\Lambda$  hyperons. The linear constraint can not guarantee the exact value of the constrained parameter: the distribution has a tail in the region of lower mass values. The nonlinear approach sets the exact value.

The mass constraint can be useful, for instance, in reconstruction of decay chains, like



In this case a  $\Lambda$ -candidate has a mass distributed according to the Gaussian law that introduces additional uncertainty to the mass of the reconstructed  $\Xi^-$  hyperon. However, setting a mass constraint on the  $\Lambda$ -candidate just before constructing the  $\Xi^-$ -candidate will improve a mass resolution of  $\Xi^-$ . And a better quality of the constraint provides a better mass resolution.

A difference between linear and nonlinear approaches can be illustrated at an example of  $\Lambda$  hyperon (see Fig. 5.5). Since the detector can not provide an infinite spatial and momentum resolution for daughter particles, the mass of the reconstructed  $\Lambda$  hyperon form a peak with a Gaussian shape. When setting the linear constraint on the reconstructed mass, the peak becomes much sharper. However, it is still has a tail in the lower mass region. When setting the nonlinear constraint described in the current section, the mass is set exactly to the needed value.

The rich functionality of the KF Particle package is perfectly suited for the reconstruction of decay trees. Thus, it is possible to build the package for the full event reconstruction based on the KF Particle mathematics.

## Chapter 6

# KF Particle Finder — a common platform for online selection and physics analysis

Today the most interesting physics is hidden in the properties of short-lived particles, that can be reconstructed only through their decay products — daughter particles. Short-lived particles that have a very small production probability or small branching ratio of the channel, which is suitable for registration, are of the particular interest. Some signals are expected to have production yields of the order  $10^{-5}$ – $10^{-7}$  particles per collision. Because of this a statistically significant result can be obtained only in case of operation with the high rates of collisions. This raises the problem of processing and storage of the data, since the data flow is so large that it can not be fully stored on tape. Therefore only events, which can potentially contain interesting particles, are stored.

In case of the CBM experiment there are no simple criteria for selecting of such events. Therefore the experiment requires the full reconstruction of events including reconstruction of short-lived particles already at the selection stage. Since the reconstruction programs should work in real time mode, their speed is a critical factor. Also, the algorithms should be efficient, not to lose events of the experiment interest.

In the present Chapter a fast and efficient algorithm for finding, reconstruction and selection of short-lived particles is described. In order to maximize the speed

and to utilize all possible resources of modern CPUs the software package, which implements the algorithm, is based on SIMD instructions. The implemented algorithm has been parallelized between cores of the CPU and demonstrates linear scalability on many-core servers with respect to the number of cores.

## 6.1 Search of short-lived particles

The method for reconstruction of short-lived particles as an input information requires charged particles registered by the tracking system of the experiment together with their PID hypothesis. In the CBM experiment trajectories of charged particles (tracks) are found by the Cellular Automaton based track finding [68]. The PID hypothesis is provided by the PID detectors. For investigating the properties of the method the PID hypothesis taken from the Monte Carlo information can be used as well.

At first, all tracks are divided into two groups: secondary tracks, which do not overlap within the errors with the primary vertex (the collision point), and primary tracks, i.e. those that are produced within the errors at the primary vertex. Secondary tracks are produced in decays of short-lived particles, that have sufficient lifetime to move away from the primary vertex. They are: strange particles ( $K_s^0$ -mesons and  $\Lambda$ -hyperons), multi-strange hyperons ( $\Xi^\pm$  and  $\Omega^\pm$ ), charmed particles ( $D^0$ ,  $D^\pm$ ,  $D_s^\pm$  and  $\Lambda_c^\pm$ ), and hypernuclei. Tracks of the particles produced at the detector material are mainly included in the set of secondary tracks, that increases background for the decays reconstruction. Primary tracks are tracks of those particles, which are produced directly in the collision of the beam with a target. Tracks of particles produced in a decay of resonances are also considered as primary tracks, since they are produced directly at the point of the primary vertex due to a small lifetime of resonances and can not be distinguished from tracks produced directly in the collision.

In order to distinguish between tracks of secondary and primary particles a  $\chi_{prim}^2$  criterion is calculated, that is a distance of the closest approach between a track and a primary vertex normalized on their total error:

$$\chi_{prim}^2 = \Delta \mathbf{r}^T (C_{track} + C_{PV})^{-1} \Delta \mathbf{r},$$

where  $\Delta \mathbf{r}$  is difference between the track and the primary vertex position,  $C_{track}$  is a covariance matrix of a track,  $C_{PV}$  is a covariance matrix of a primary vertex.

The  $\chi_{prim}^2$  criterion defines the probability, that the track intersects the primary vertex within the errors, under the assumptions that parameters of tracks are distributed according to the Gaussian law and  $\chi_{prim}^2$  is distributed according to the  $\chi^2$ -distribution with the number of degrees of freedom (NDF) equal 2. When the Kalman filter is used for track parameters estimation these two assumptions are fulfilled. The NDF for  $\chi_{prim}^2$  is equal 2 since only two coordinates of a track out of three are independent, since the tracking systems provide measurements with one coordinate fixed. In case of the CBM experiment the  $z$  coordinate is fixed. If the probability is smaller than 0.01% ( $\chi_{prim}^2 > 18.4$ ), the track is considered as the secondary one, otherwise — as the primary track. This means, that the true primary track can be wrongly assigned as the secondary one with the probability of 0.01%.

In order to calculate  $\chi_{prim}^2$  criterion all tracks are extrapolated to the primary vertex region. Parameters of those tracks, that have been assigned as primary, are stored at the primary vertex position. Since extrapolation of track parameters is one of the most time consuming procedures, this allows to save time and speed up the program on the stage of resonances reconstruction.

On top of this, all tracks are divided into positive and negative. Thus, four groups are formed:

- secondary positive tracks;
- secondary negative tracks;
- primary positive tracks;
- primary negative tracks.

Tracks in each of these groups are then sorted according to the PID hypothesis. Since pions and protons are the most abundant particles and  $\Lambda$  has a high signal to background ratio, tracks, which does not have the PID information, are assumed to be pions for all decays reconstruction and both pions and protons for the  $\Lambda$  decay reconstruction. Such assumption allows to recover the inefficiency of the PID detectors and does not decrease the signal to background ratio significantly even for rare decays.

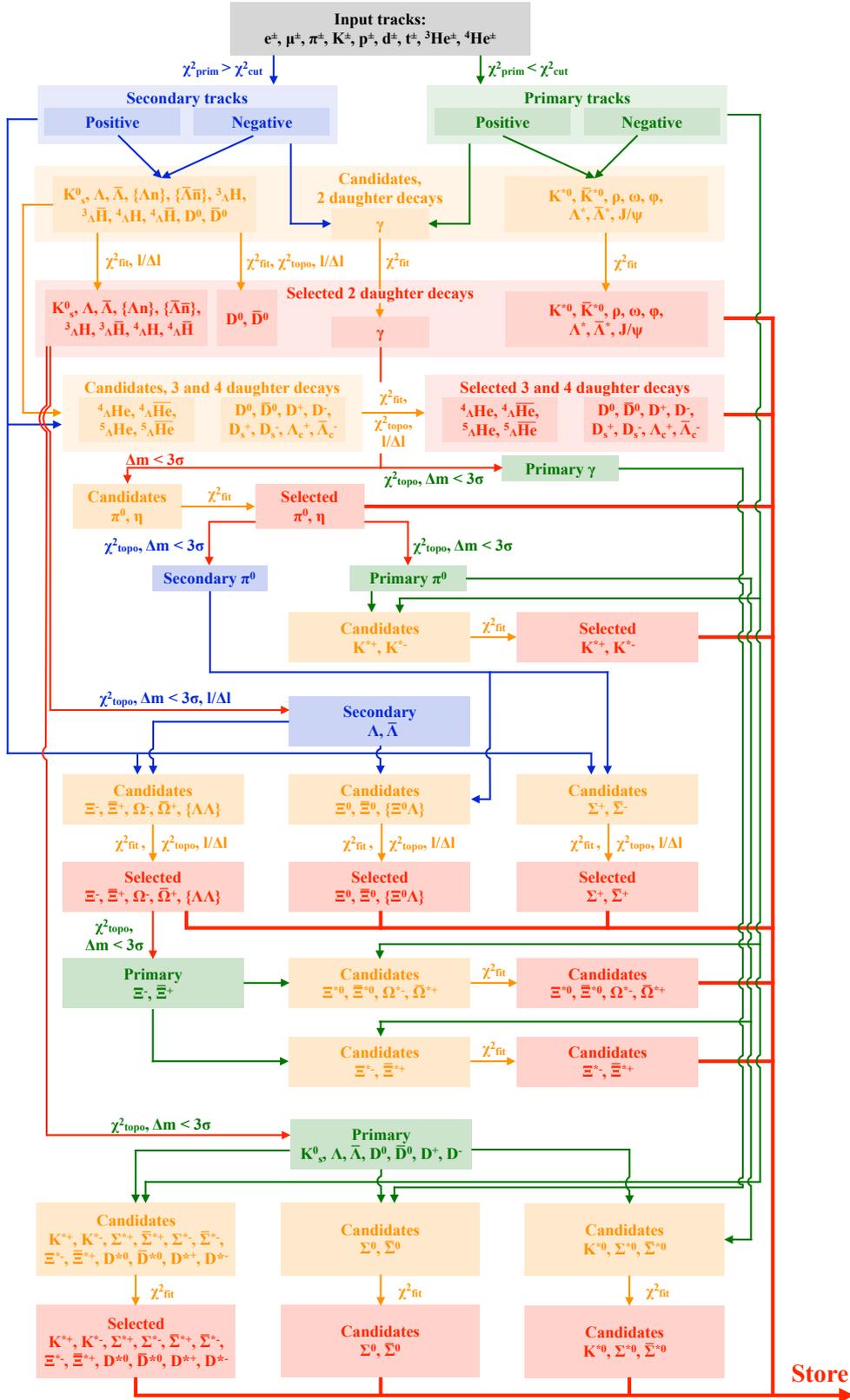


Figure 6.1: Block-diagram of the reconstruction method of the KF Particle Finder package. Cuts on  $\chi^2_{\text{fit}}$  and  $\chi^2_{\text{topo}}$  are given in terms of NDF.

When all tracks are sorted, short-lived mother particles are reconstructed using KF Particle out of them according to the block-diagram shown in Fig. 6.1. At first, particles with two daughters in the decay channels are reconstructed:

- Mother particles, which have a lifetime large enough to be separated from the primary vertex (from  $D$  mesons with  $c\tau$  of hundreds  $\mu\text{m}$  to strange and multi-strange hyperons with  $c\tau$  of several cm), are constructed from secondary tracks: one positive and one negative.

In case of charmed particles, which have a small production yield, the background should be strongly suppressed, therefore tracks should fulfill additional requirements: the cut on  $\chi_{prim}^2$  criterion should be tighter, that allows to decrease the probability to use for reconstruction primary tracks wrongly assigned as secondary, and the transverse momentum of the track ( $p_t = \sqrt{p_x^2 + p_y^2}$ ) should be larger, then 200 MeV/c, that is caused by the decay properties of heavy charmed particles. Since open charm particles have a decay length of hundreds  $\mu\text{m}$ , they decay in the neighbourhood of the primary vertex, therefore the distance to the primary vertex should be smaller than the cut value.

- Mother particles, that decay at the primary vertex position, including resonances, low mass vector mesons and charmonium, are constructed combining one positive primary track and one negative primary track.

The additional background suppression is needed for dilepton channels of low mass vector mesons ( $\rho$ ,  $\omega$ , and  $\phi$ ) and  $J/\psi$  because of their small branching ratio and production yield respectively. Therefore tracks in case of low mass vector mesons are checked to have  $p_t > 200$  MeV/c and in case of  $J/\psi$   $p_t > 1$  GeV/c, that is caused by the decay kinematics.

- For reconstruction of  $\gamma \rightarrow e^+e^-$  conversion all electrons are combined with all positrons, since conversion can take place at any point of the detector.

The tracks are combined according to their PID hypothesis into the particle candidates decaying by the corresponding channels.

Besides real physical signals arbitrary combinations of tracks can produce a candidate, that does not correspond to the real particle. Such combinations form the background for the physical signals. Particle candidates constructed from

the tracks with incorrectly assigned PID hypothesis are another source of the background. While arbitrary combinations can be rejected with cuts on the fit quality and decay topology, the second type of the background can be reduced only by the correct particle identification.

In order to suppress the background a set of cuts is applied to the candidates:

- a cut on the  $\chi_{fit}^2/\text{NDF}$  criterion calculated by the KF Particle mathematics in the candidate fit, that characterizes whether trajectories of daughter particles intersect within their errors;
- a cut on the  $\chi_{topo}^2/\text{NDF}$ , which is a cut on  $\chi^2$  criterion of the candidate with a production point constraint set to the primary vertex; this criterion characterizes whether the particle is produced in the primary vertex region: the smaller the obtained value — the larger probability, that trajectory of the particle intersects the primary vertex within the errors;
- a cut on the distance from the decay point of the candidate to the primary vertex normalized on the error  $l/\Delta l$ , that shows how far the candidate is constructed from the primary vertex in terms of  $\sigma$  assuming the Gaussian distribution.

The cuts for different groups of decays are optimized with respect to the particle yields, branching ratios of the decay channel under investigation, and decay properties.

Strange particles ( $K_s^0$ ,  $\Lambda$ ,  $\bar{\Lambda}$ ) and hypernuclei ( $\{\Lambda n\}$ ,  ${}_{\Lambda}\text{H}$ ) have a large lifetime ( $c\tau$  is in the order of several cm), that allows to separate them well from the primary vertex, therefore only two cuts are applied:  $\chi_{fit}^2/\text{NDF}$ , which selects correctly fitted candidates, and  $l/\Delta l$ , which selects candidates constructed more than several  $\sigma$  away from the primary vertex region (cuts are optimized with respect to the signal to background ratio [74, 75, 76]).

$D^0$  mesons have  $c\tau = 123 \mu\text{m}$ , that still allows to separate them from the primary vertex, but in order to assure sufficient signal to background ratio taking into account small production yields tighter cuts are applied on  $\chi_{fit}^2/\text{NDF}$  and  $l/\Delta l$ , and in addition the cut on  $\chi_{topo}^2/\text{NDF}$  is set, since  $D$  mesons are expected to be produced only in the collision, i.e. should point to the primary vertex.

Resonances ( $K^{*0}$  and  $\Lambda^*$ ), low mass vector mesons ( $\rho$ ,  $\omega$ ,  $\phi$ ) and  $J/\psi$  are

produced and decay at the primary vertex position, therefore only the cut on the fit quality  $\chi_{fit}^2/\text{NDF}$  can be applied [77, 78].

The same cut is applied for  $\gamma$  reconstruction, since conversion can happen at any point of the detector.

The particle candidates, which pass all cuts, are stored on tape.

For reconstruction of particle trees, that appear in decay of multi-strange hyperons, secondary  $\Lambda$  and  $\bar{\Lambda}$  are selected. The particle is considered as secondary, if  $\chi_{topo}^2/\text{NDF} > \chi_{cut}^2$ . Also, the mass is checked to be consistent with the table value within  $3\sigma$  ( $\Delta m < 3\sigma$ ), that according to the Gaussian distribution defines the probability of 0.3% to reject correct signal, and the candidate should be well separated from the primary vertex ( $1/\Delta l > 1/\Delta l_{cut}$ ). For reconstruction of resonances primary  $\gamma$ ,  $K_s^0$ ,  $\Lambda$ ,  $\bar{\Lambda}$ ,  $D^0$ ,  $\bar{D}^0$ ,  $D^+$ ,  $D^-$  are selected, that satisfy conditions  $\chi_{topo}^2/\text{NDF} < \chi_{cut}^2$  and  $\Delta m < 3\sigma$  [74, 75, 76].

Since particles with different number of decay products have daughters of the same type in the decay channels, for reconstruction of three and four daughter decays, including hypernuclei ( ${}_{\Lambda}\text{He}$ ),  $D$  mesons and  $\Lambda_c^+$ , candidates obtained on the stage of two daughter decays reconstruction are combined with secondary tracks. Again, for reconstruction of charmed particles the tighter cut on  $\chi_{prim}^2$  is set. The cuts on  $\chi_{fit}^2/\text{NDF}$ ,  $1/\Delta l$ , and  $\chi_{topo}^2/\text{NDF}$  are applied to the obtained candidates, and if they are fulfilled, the candidate is stored.

In order to reconstruct  $\pi^0$  particles those  $\gamma$  candidates are combined with each other, that have the mass consistent with the table value within  $3\sigma$ . Since  $\pi^0$  are produced both directly in the collision and in the decay of other particles, only the cut on  $\chi_{fit}^2/\text{NDF}$  is applied [79]. If it is satisfied, the candidate is stored. For further analysis the secondary and primary  $\pi^0$  are selected: the candidate is considered as secondary, if it satisfies conditions on  $\chi_{topo}^2/\text{NDF} > \chi_{cut}^2$  and  $\Delta m < 3\sigma$ , and primary if  $\chi_{topo}^2/\text{NDF} < \chi_{cut}^2$  and  $\Delta m < 3\sigma$ .

Combining primary  $\pi^0$  with primary tracks ( $K^+$  and  $K^-$ ) the candidates for  $K^{*+}$  and  $K^{*-}$  are constructed, which are stored if they pass the fit goodness condition  $\chi_{fit}^2/\text{NDF}$ .

Secondary  $\pi^0$  and tracks ( $p$  and  $\bar{p}$ ) form  $\Sigma^+$  and  $\Sigma^-$ , which are produced directly in the collision and decay far from the primary vertex ( $c\tau=2.4$  cm), therefore they are checked to fulfill cuts on  $\chi_{fit}^2/\text{NDF}$ ,  $1/\Delta l$ , and  $\chi_{topo}^2/\text{NDF}$ .

For reconstruction of  $\Xi^-$ ,  $\bar{\Xi}^+$ ,  $\Omega^-$ ,  $\bar{\Omega}^+$  hyperons and hypothetical  $\{\Lambda\Lambda\}$  dybarion the already selected secondary  $\Lambda$  and  $\bar{\Lambda}$  are combined with secondary tracks. Since these particles have large lifetime (their  $c\tau$  are of the order of several cm) and are expected to be produced directly in the collision, they should pass the cuts on  $\chi_{fit}^2/\text{NDF}$ ,  $1/\Delta l$ , and  $\chi_{topo}^2/\text{NDF}$  [74, 75, 76]. Out of these candidates primary  $\Xi^-$  and  $\bar{\Xi}^+$  are selected according to the conditions  $\chi_{topo}^2/\text{NDF} < \chi_{cut}^2$  and  $\Delta m < 3\sigma$ . The obtained primary  $\Xi^-$  and  $\bar{\Xi}^+$  are combined with primary tracks and  $\pi^0$  into  $\Xi^*$  and  $\Omega^*$ , which should satisfy condition  $\chi_{fit}^2/\text{NDF} < \chi_{cut}^2$ . Also, secondary  $\Lambda$  are combined with each other into hypothetical  $\{\Xi^0\Lambda\}$  dybarion, and  $\Xi^{*0}$ ,  $\bar{\Xi}^{*0}$  are reconstructed combining secondary  $\Lambda$  and  $\bar{\Lambda}$  with secondary  $\pi^0$ ; these secondary particle candidates are checked to fulfill cuts on  $\chi_{fit}^2/\text{NDF}$ ,  $1/\Delta l$ , and  $\chi_{topo}^2/\text{NDF}$ . All the candidates, which passed the cuts, are stored for further analysis.

For reconstruction of resonances ( $K^*$ ,  $\Sigma^*$ ,  $\Xi^*$ ,  $D^*$ ) and  $\Sigma^0$  the selected primary  $K_s^0$ ,  $\Lambda$ ,  $\bar{\Lambda}$ ,  $D^0$ ,  $\bar{D}^0$ ,  $D^+$ , and  $D^-$  are combined with primary tracks,  $\pi^0$ , and  $\gamma$ . Since all these decays take place at the primary vertex region, they are stored if cut on  $\chi_{fit}^2/\text{NDF}$  is fulfilled [80, 81].

Currently, the KF Particle Finder includes more than 70 decays (see Fig. 6.2). The method is developed to reconstruct signals from all the physics cases of the CBM experiment, including strange particles, strange resonances, hypernuclei, low mass vector mesons, charmonium and open-charm particles. This makes the KF Particle Finder a universal platform for short-lived particle reconstruction and physics analysis.

## 6.2 Physics analysis with KF Particle Finder

The physics processes, which take place in the collisions of heavy ions, can be investigated only through the properties of particles produced in the collision. Therefore all possible species of the particles should be found. The KF Particle Finder package provides extended opportunities for analysis of the collision physics by reconstructing a wide range of decays.

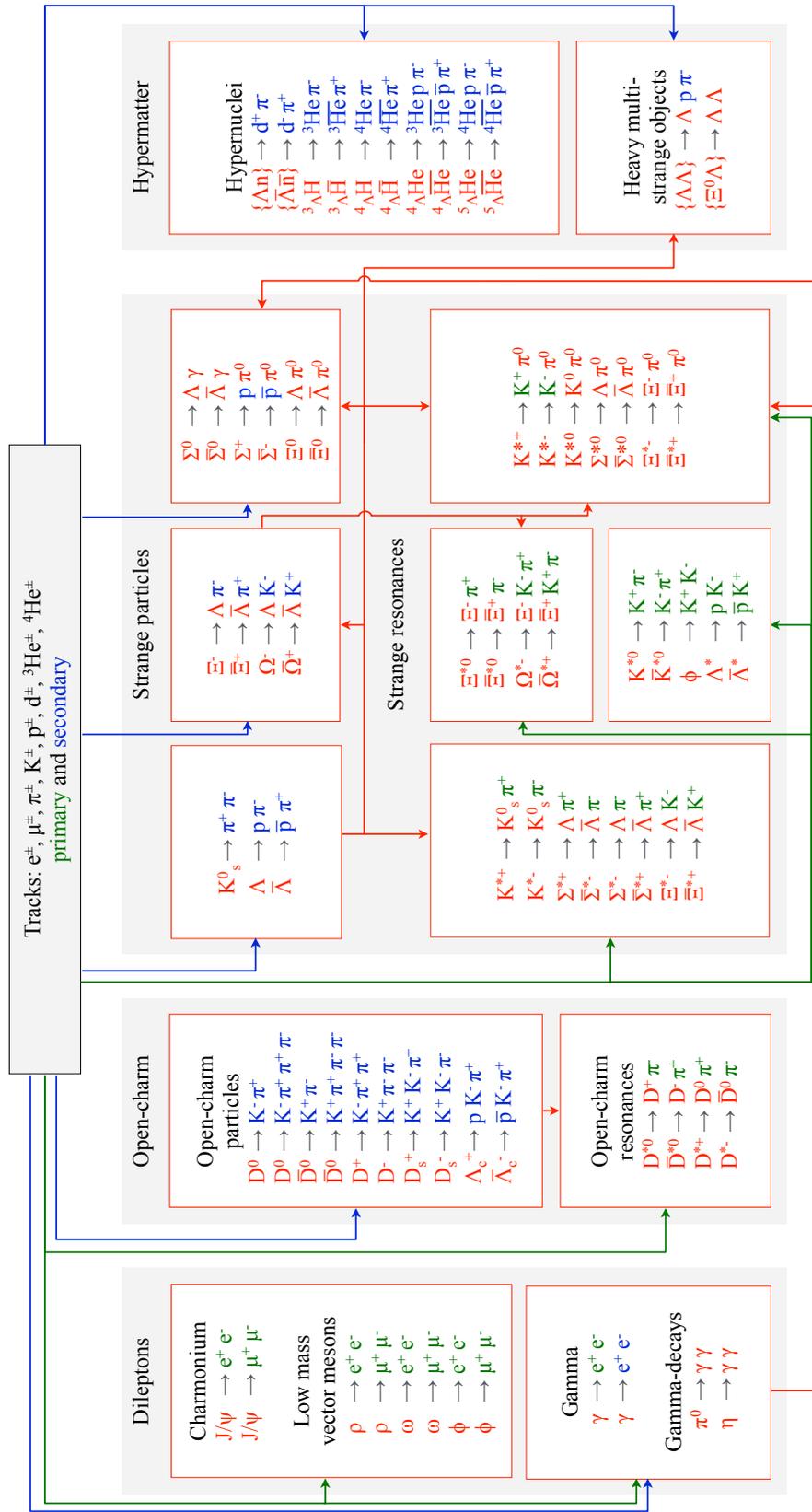


Figure 6.2: Block-diagram of the KF Particle Finder package for finding, reconstruction and selection of rare short-lived particles.

### 6.2.1 Strange and multi-strange particles

In the collision of heavy ions new quarks are produced including strange  $s$  and  $\bar{s}$  quarks. According to the theoretical predictions enhanced strangeness production can be an indicator for creating deconfined matter, or QGP, in ultrarelativistic heavy-ion collisions [82]. At top SPS energies it was observed that the multiplicity of strange hyperons per participating nucleon increases with the mass of the collision system [83]. Moreover, the effect of enhancement increases with the number of strange quarks in the hyperon. According to the current knowledge strange and multi-strange hyperons can be produced in sequential collisions involving kaons and  $\Lambda$ s, and, therefore, are sensitive to the density in the collision system. This sensitivity is expected to increase towards lower beam energies close to or even below the production threshold [22, 31].

Strangeness production is one the main focuses of the CBM experiment. Having relatively small mass of  $95 \text{ MeV}/c^2$  strange quarks are abundantly produced at CBM energies. Thus, they are perfectly suited for investigation of hot and dense matter. The decay channels of the strange particles, search of which is implemented in the KF Particle Finder, are listed together with the decay parameters in Table 6.1.

Particle	Antiparticle	Branching ratio, %	$c\tau$ , cm
$K_s^0 \rightarrow \pi^+ \pi^-$	—	69.2	2.7
$\Lambda \rightarrow p \pi^-$	$\bar{\Lambda} \rightarrow \bar{p} \pi^-$	63.9	7.9
$\Sigma^+ \rightarrow p \pi^0$	$\bar{\Sigma}^- \rightarrow \bar{p} \pi^0$	51.6	2.4
$\Sigma^0 \rightarrow \Lambda \gamma$	$\bar{\Sigma}^0 \rightarrow \bar{\Lambda} \gamma$	100.0	$2.2 \times 10^{-13}$
$\Xi^0 \rightarrow \Lambda \pi^0$	$\bar{\Xi}^0 \rightarrow \bar{\Lambda} \pi^0$	99.5	8.7
$\Xi^- \rightarrow \Lambda \pi^-$	$\bar{\Xi}^+ \rightarrow \bar{\Lambda} \pi^+$	99.9	4.9
$\Omega^- \rightarrow \Lambda K^-$	$\bar{\Omega}^+ \rightarrow \bar{\Lambda} K^+$	67.8	2.5

Table 6.1: Parameters of strange particles and antiparticles included in the KF Particle Finder reconstruction scheme: the implemented channel, its branching ratio, and the lifetime ( $c\tau$ ) of a particle [5].

The fit quality of the reconstructed strange particles is shown in Fig. 6.3 at the example of  $\Lambda$  hyperon: distributions of residuals ( $\rho$ ), pulls,  $\chi^2$  criterion, and prob (probability for a certain  $\chi^2$  and NDF) are presented. Residuals and pulls,

which are obtained at the decay point of the particles, are shown for  $x$ ,  $p_x$ , and  $E$ , other parameters demonstrate the same behavior. For track reconstruction the STS detector is used. The distributions prove the high quality of the fit:

- All distributions of residuals and pulls are unbiased.
- The width of residuals is in a good agreement with the detector resolution.
- The width of pulls is close to the expected value of one, the pulls of momentum and energy are wider due to the higher sensitivity to the approximation in the detector material and the magnetic field treatment.
- The  $\chi^2$  and corresponding prob distributions have the expected shape, prob is close to the flat uniform distribution from 0 to 1. Sharp edges represent the cut on  $\chi^2/\text{NDF}$  for selection of particle candidates.

Particle	$\varepsilon_{method}, \%$	$\varepsilon_{4\pi}, \%$	S/B	$\sigma, \text{MeV}/c^2$
$K_s^0$	71.5	24.9	2.5	3.6
$\Lambda$	65.4	27.0	8.2	1.6
$\bar{\Lambda}$	66.5	17.0	4.1	1.5
$\Xi^-$	48.0	12.8	16.4	2.0
$\bar{\Xi}^+$	44.5	6.7	21.8	1.7
$\Omega^-$	45.4	5.5	91.9	2.1

Table 6.2: Efficiencies of the strange particles reconstruction, widths of the signal peaks, and S/B ratios for 5M central AuAu UrQMD events at 10 AGeV using TOF PID.

The high quality of the parameters together with  $c\tau$  of several cm (see Table 6.1) allow reconstruction of strange particles both with the high signal to background (S/B) ratio and efficiency (see Table 6.2). Mass spectra of  $K_s^0$ ,  $\Lambda$ ,  $\bar{\Lambda}$ ,  $\Xi^-$ ,  $\bar{\Xi}^+$ , and  $\Omega^-$  obtained by the KF Particle Finder for 5M central AuAu UrQMD [84, 85] events at 10 AGeV using TOF PID [75] are shown in Fig. 6.4. The S/B ratio is calculated in  $\pm 2\sigma$  of the Gaussian fit of the signal peak, the background was approximated by the second order polynomial function. Two types of efficiency are given: the efficiency of the method ( $\varepsilon_{method}$ ) itself, which represents the efficiency of the secondary particles selection and cuts on the particle candidates, and the total efficiency ( $\varepsilon_{4\pi}$ ) in the  $4\pi$  solid angle, which additionally takes into account the acceptance of the detector, the inefficiencies of track reconstruction and PID algorithms, and therefore several times smaller.

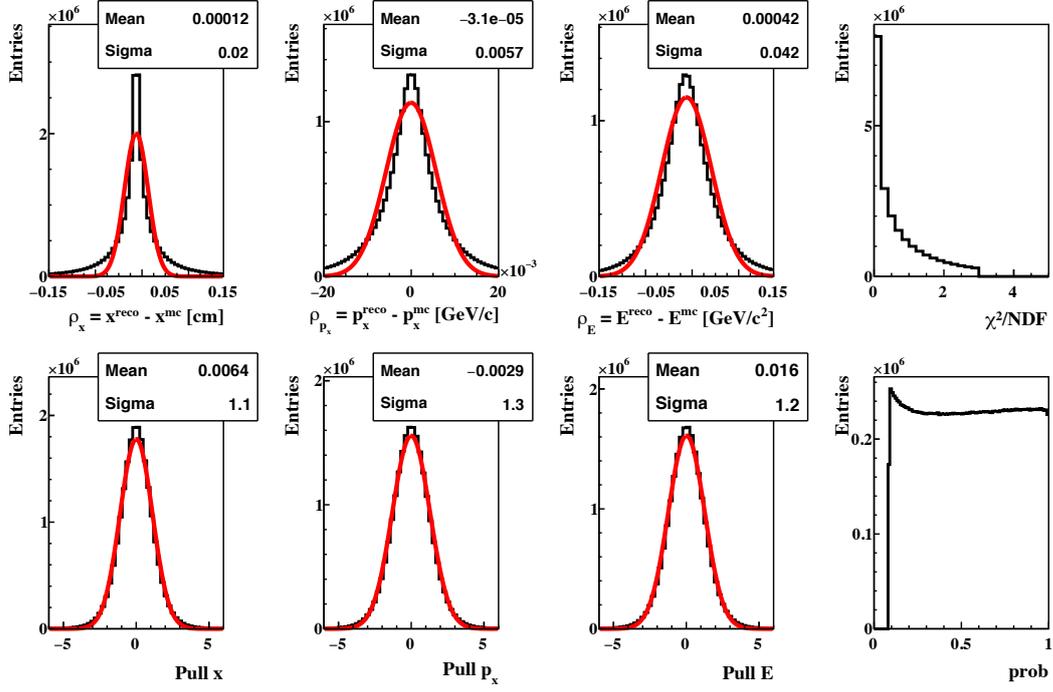


Figure 6.3: Fit quality of the reconstructed strange particles at the example of  $\Lambda$ . Distributions of residuals  $\rho$ , pulls,  $\chi^2$  criterion, and prob (probability for a certain  $\chi^2$  and NDF) demonstrate the high quality of the reconstructed parameters.

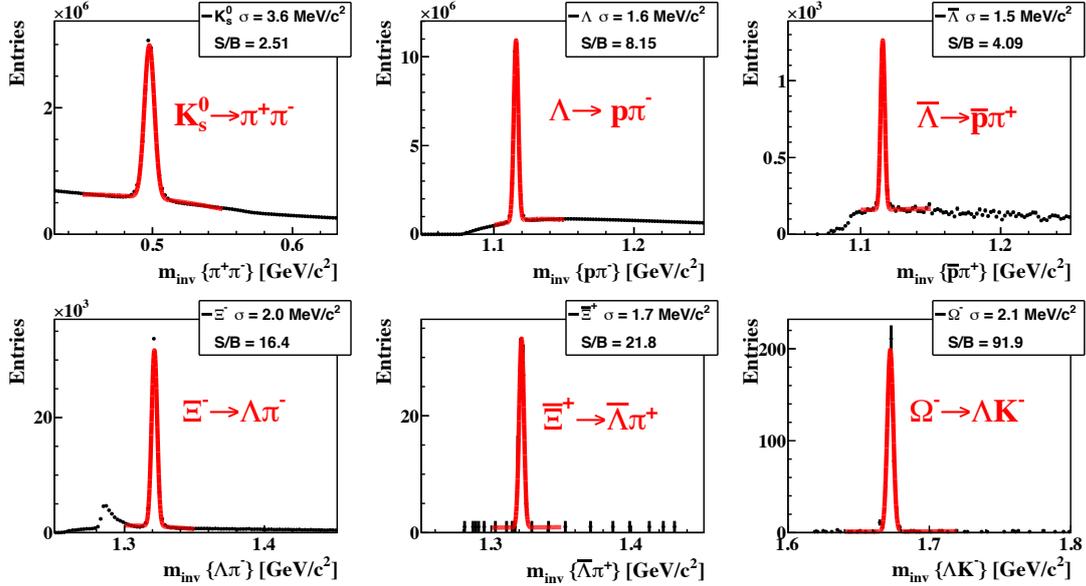


Figure 6.4: Mass spectra of  $K_s^0$ ,  $\Lambda$ ,  $\bar{\Lambda}$ ,  $\Xi^-$ ,  $\Xi^+$ , and  $\Omega^-$  obtained by the KF Particle Finder package together with their signal to background (S/B) ratios for 5M central AuAu UrQMD [84, 85] events at 10 AGeV using TOF PID [75].

## 6.2.2 Strange and multi-strange resonances

Reconstruction of strange and multi-strange resonances (see Table 6.3) completes the picture of strangeness production. Decaying in the medium, resonances allow to investigate properties of the dense baryonic matter. However, with strongly decaying resonances only late stages of the collision can be studied due to final-state absorption [22, 31].

Particle	Antiparticle	Branching ratio, %	Full width $\Gamma$ , MeV/c <sup>2</sup>
$K^+(892) \rightarrow K_s^0 \pi^+$	$K^-(892) \rightarrow K_s^0 \pi^-$	33.3	50.8
$K^+(892) \rightarrow K^+ \pi^0$	$K^-(892) \rightarrow K^- \pi^0$	33.3	50.8
$K^0(892) \rightarrow K^+ \pi^-$	$\bar{K}^0(892) \rightarrow K^- \pi^+$	33.3	47.4
$K^0(892) \rightarrow K_s^0 \pi^0$	$\bar{K}^0(892) \rightarrow K_s^0 \pi^0$	33.3	47.4
$\Lambda(1520) \rightarrow p K^-$	$\bar{\Lambda}(1520) \rightarrow \bar{p} K^-$	22.4	15.6
$\Sigma^+(1385) \rightarrow \Lambda \pi^+$	$\bar{\Sigma}^-(1385) \rightarrow \bar{\Lambda} \pi^-$	87.0	36.0
$\Sigma^0(1385) \rightarrow \Lambda \pi^0$	$\bar{\Sigma}^0(1385) \rightarrow \bar{\Lambda} \pi^0$	87.0	36.0
$\Sigma^-(1385) \rightarrow \Lambda \pi^-$	$\bar{\Sigma}^+(1385) \rightarrow \bar{\Lambda} \pi^+$	87.0	39.4
$\Xi^0(1530) \rightarrow \Xi^- \pi^+$	$\bar{\Xi}^0(1530) \rightarrow \bar{\Xi}^+ \pi^-$	66.6	9.1
$\Xi^-(1530) \rightarrow \Xi^- \pi^0$	$\bar{\Xi}^+(1530) \rightarrow \bar{\Xi}^+ \pi^0$	33.3	9.9
$\Xi^-(1820) \rightarrow \Lambda K^-$	$\bar{\Xi}^+(1820) \rightarrow \bar{\Lambda} K^+$	—	24.0
$\Omega^-(2250) \rightarrow \Xi^- K^- \pi^+$	$\bar{\Omega}^+(2250) \rightarrow \bar{\Xi}^+ K^+ \pi^-$	—	55.0

Table 6.3: Parameters of strange resonances and their antiparticles included in the KF Particle Finder reconstruction scheme: the implemented channel, its branching ratio, and the full width ( $\Gamma$ ) of the resonance [5]. The most abundant resonances decaying by the selected channels are shown.

The fit quality of the reconstructed strange resonances is shown in Fig. 6.5 at the example of  $\Sigma^+(1385)$  with STS used for track reconstruction. The quality is as high as for the strange particles. Mass spectra of  $\Sigma^+(1385)$ ,  $\Sigma^-(1385)$ , and  $\Xi^0(1530)$  obtained by the KF Particle Finder for 5M central AuAu UrQMD events at 10 AGeV using TOF PID are shown in Fig. 6.6. Corresponding efficiencies, S/B ratios, and widths of the peaks are given in Table 6.4. Since resonances decay directly at the collision point, the decay topology can not be used for the background suppression and the S/B ratio is several orders smaller comparing to the strange particles, however the reconstruction efficiency is similar. The reconstructed width of the peaks is in agreement with the UrQMD model.

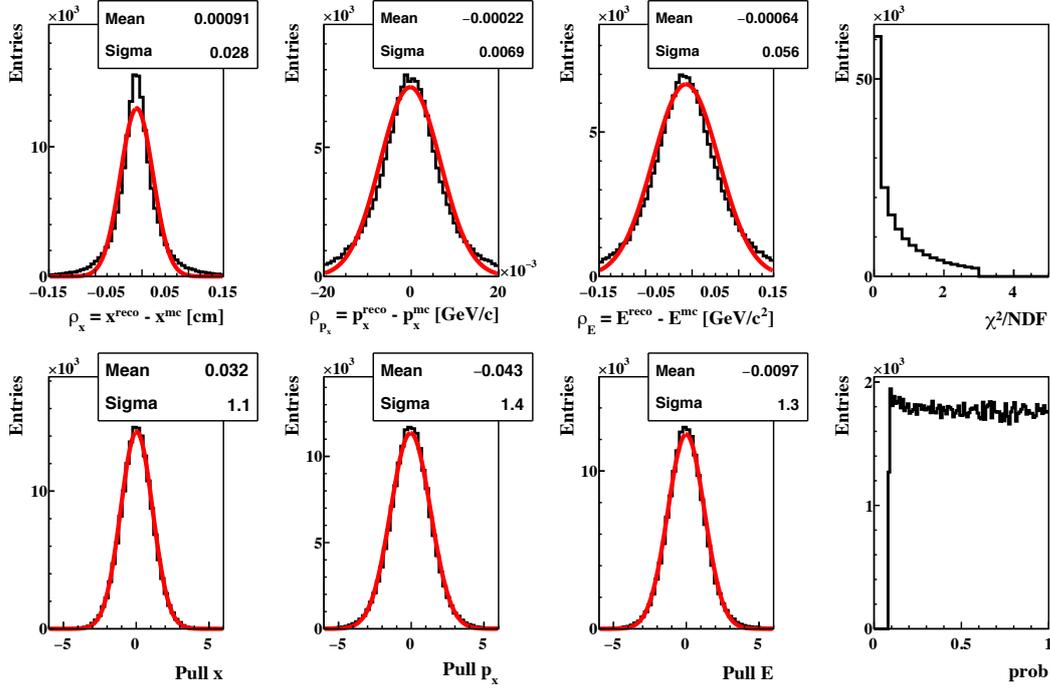


Figure 6.5: Fit quality of the reconstructed strange resonances at the example of  $\Sigma^+(1385)$ . Distributions of residuals  $\rho$ , pulls,  $\chi^2$  criterion, and prob (probability for a certain  $\chi^2$  and NDF) demonstrate the high quality of the reconstructed parameters.

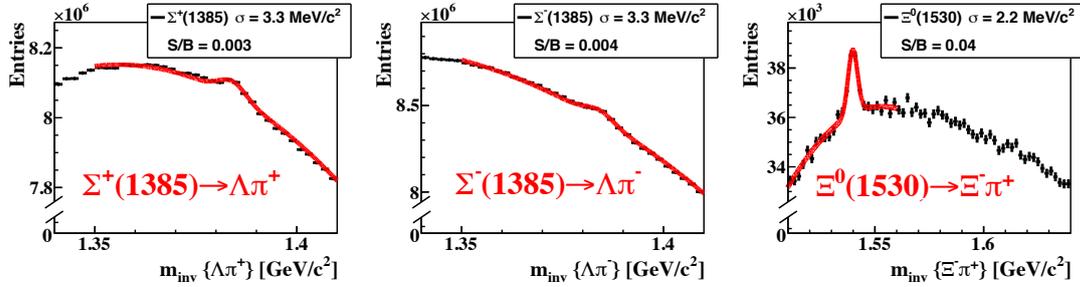


Figure 6.6: Mass spectra of  $\Sigma^+(1385)$ ,  $\Sigma^-(1385)$ , and  $\Xi^0(1530)$  obtained by the KF Particle Finder package together with their signal to background (S/B) ratios for 5M central AuAu UrQMD [84, 85] events at 10 AGeV using TOF PID.

Particle	$\varepsilon_{method}, \%$	$\varepsilon_{4\pi}, \%$	S/B	$\sigma, \text{MeV}/c^2$
$\Sigma^+(1385)$	57.8	11.9	0.003	3.3
$\Sigma^-(1385)$	58.3	11.9	0.004	3.3
$\Xi^0(1530)$	49.4	7.5	0.040	2.2

Table 6.4: Efficiencies of the strange resonances reconstruction, widths of the signal peaks, and S/B ratios for 5M central AuAu UrQMD events at 10 AGeV using TOF PID.

### 6.2.3 Hypernuclei and heavy multi-strange objects

Theoretical models predict single and double hypernuclei, and heavy multi-strange short-lived objects to be produced in heavy-ion collisions with the maximum yield at SIS100 energies [86, 87]. Investigation of these objects will shed light on the hyperon-nucleon and hyperon-hyperon interactions, which are important for the nuclear equation-of-state at high densities and low temperatures. Such conditions are believed to exist at cores of neutron stars [22, 31]. Particles added to the KF Particle Finder are listed in Table 6.5.

Particle	Antiparticle	Branching ratio, %	$c\tau$ , cm
$\{\Lambda n\} \rightarrow d^+ \pi^-$	$\{\overline{\Lambda} \overline{n}\} \rightarrow d^- \pi^+$	25 <sup>1</sup>	11.1
${}^3_{\Lambda}H \rightarrow {}^3He \pi^-$	${}^3_{\overline{\Lambda}}\overline{H} \rightarrow {}^3\overline{He} \pi^+$	24	5.5
${}^4_{\Lambda}H \rightarrow {}^4He \pi^-$	${}^4_{\overline{\Lambda}}\overline{H} \rightarrow {}^4\overline{He} \pi^+$	50	5.4
${}^4_{\Lambda}He \rightarrow {}^3He p \pi^-$	${}^4_{\overline{\Lambda}}\overline{He} \rightarrow {}^3\overline{He} p \pi^+$	34	4.5
${}^5_{\Lambda}He \rightarrow {}^4He p \pi^-$	${}^5_{\overline{\Lambda}}\overline{He} \rightarrow {}^4\overline{He} p \pi^+$	43	4.2
$\{\Lambda\Lambda\} \rightarrow \Lambda p \pi^-$	—	unknown	unknown
$\{\Xi^0\Lambda\} \rightarrow \Lambda\Lambda$	—	unknown	unknown

Table 6.5: Parameters of hypernuclei and heavy multi-strange objects included in the KF Particle Finder reconstruction scheme: the implemented channel, its branching ratio, and the lifetime ( $c\tau$ ) of a particle [88, 89, 90, 91, 92, 93, 94].

The fit quality of these particles is shown in Fig. 6.7 at the example of  ${}^3_{\Lambda}H$  with STS used for track reconstruction. The quality is as high as for the previous groups of particles. Mass spectra of  $\{\Lambda n\}$ ,  ${}^3_{\Lambda}H$ , and  ${}^4_{\Lambda}He$  obtained by the KF Particle Finder for 5M simulated central AuAu events at 10 AGeV using TOF PID are shown in Fig. 6.8. For background simulation UrQMD events are used. Since the UrQMD model does not produce hypernuclei, the signal is generated and reconstructed separately and added to the background spectra with the theoretically predicted multiplicities [87] and branching ratios listed in Table 6.5.

Corresponding efficiencies, S/B ratios, and widths of the peaks are given in Table 6.6. The large lifetime ( $c\tau$  of several cm) together with a low multiplicity of produced background nuclei fragments ( $d$ ,  $t$ ,  $He$ ) allow a strong suppression of the background for hypernuclei keeping the reconstruction efficiency high.

<sup>1</sup> The estimation is based on [92, 93, 94]

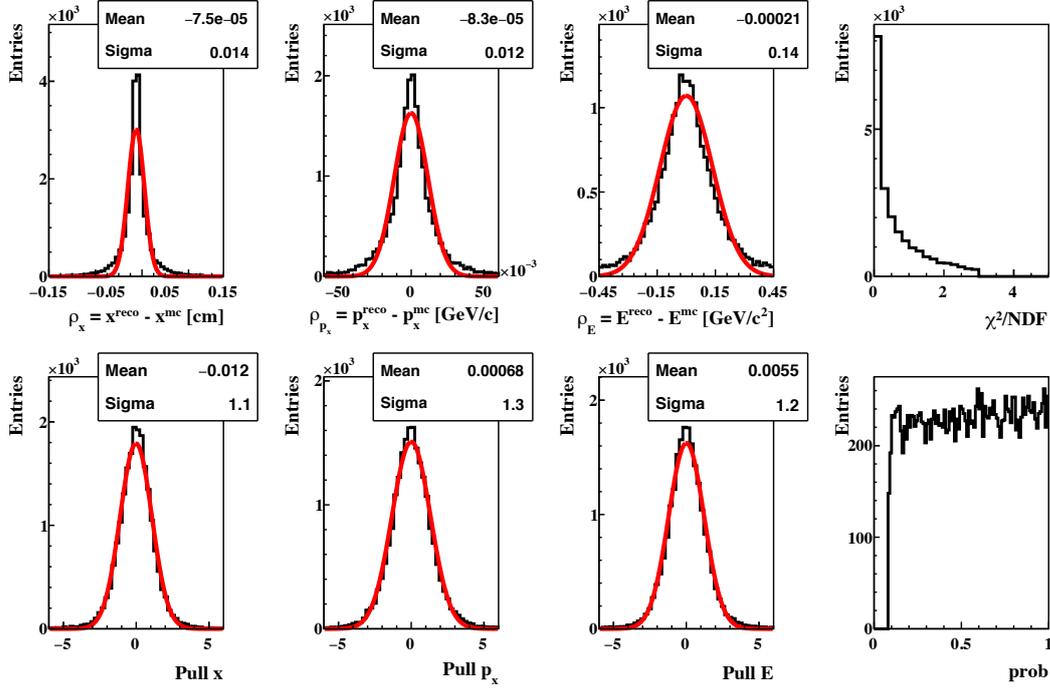


Figure 6.7: Fit quality of the reconstructed hypernuclei at the example of  ${}^3_{\Lambda}H$ . Distributions of residuals  $\rho$ , pulls,  $\chi^2$  criterion, and prob (probability for a certain  $\chi^2$  and NDF) demonstrate the high quality of the reconstructed parameters.

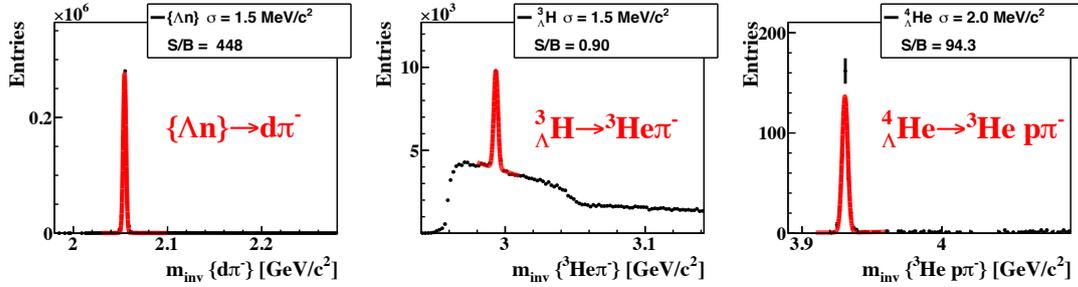


Figure 6.8: Mass spectra of  $\{\Lambda n\}$ ,  ${}^3_{\Lambda}H$ , and  ${}^4_{\Lambda}He$  hypernuclei obtained by the KF Particle Finder package together with their signal to background (S/B) ratios for 5M simulated central AuAu events at 10 AGeV using TOF PID.

Particle	$\varepsilon_{method}, \%$	$\varepsilon_{4\pi}, \%$	S/B	$\sigma, \text{MeV}/c^2$
$\{\Lambda n\}$	49.5	21.9	450	1.5
${}^3_{\Lambda}H$	44.7	16.5	0.9	1.5
${}^4_{\Lambda}He$	38.0	10.8	94.3	2.0

Table 6.6: Efficiencies of the hypernuclei reconstruction, widths of the signal peaks, and S/B ratios for 5M simulated central AuAu events at 10 AGeV using TOF PID.

## 6.2.4 Open charm particles

At the CBM energies charm quarks are expected to be produced in primary hard scattering processes, that allows to study the early phase of the collision. Depending on the interaction with the medium, they hadronize into  $D$  mesons, charmed baryons, or charmonium. At SIS100 energies charm production mechanisms at the threshold energies will be studied and investigation of charm production in collisions of medium size nuclei and proton-nucleus will be possible. At SIS300 measurements of charm in collisions of heavy nuclei will be performed [22, 31]. Particles added to the KF Particle Finder are listed in Table 6.7.

Particle	Antiparticle	Branching ratio, %	$c\tau$ , $\mu\text{m}$ / $\Gamma$ , $\text{MeV}/c^2$
$D^+ \rightarrow K^- \pi^+ \pi^+$	$D^- \rightarrow K^+ \pi^- \pi^-$	9.1	311.8 $\mu\text{m}$
$D^0 \rightarrow K^- \pi^+$	$D^- \rightarrow K^+ \pi^-$	3.9	122.9 $\mu\text{m}$
$D^0 \rightarrow K^- \pi^+ \pi^+ \pi^-$	$\bar{D}^0 \rightarrow K^+ \pi^+ \pi^- \pi^-$	8.1	122.9 $\mu\text{m}$
$D_s^+ \rightarrow K^+ K^- \pi^+$	$D_s^- \rightarrow K^+ K^- \pi^-$	5.4	149.9 $\mu\text{m}$
$\Lambda_c^+ \rightarrow p K^- \pi^+$	$\bar{\Lambda}_c^+ \rightarrow \bar{p} K^+ \pi^-$	5.0	59.9 $\mu\text{m}$
$D^+(2010) \rightarrow D^0 \pi^+$	$D^-(2010) \rightarrow \bar{D}^0 \pi^-$	67.7	0.083 $\text{MeV}/c^2$
$D^0(2460) \rightarrow D^+ \pi^-$	$\bar{D}^0(2460) \rightarrow D^- \pi^+$	—	49.0 $\text{MeV}/c^2$

Table 6.7: Parameters of open charm particles, antiparticles, and resonances included in the KF Particle Finder reconstruction scheme: the implemented channel, its branching ratio, and the lifetime ( $c\tau$ ) or the full width ( $\Gamma$ ) of a particle [5].

The high fit quality of the open charm particles is illustrated in Fig. 6.9 at the example of  $\bar{D}^0$  with STS and MVD used for track reconstruction.

The mass spectra of  $D^0$  and  $\bar{D}^0$  obtained by the KF Particle Finder for  $10^{11}$  central NiNi events at 15 AGeV using TOF PID are shown in Fig. 6.10. For simulation of the background the mixed event technique is used: tracks produced in different UrQMD collisions are mixed and combined with each other, that gives the number of arbitrary combinations growing quadratically with the number of events. Such technique allows to estimate the background for the several orders higher statistics than a usual event-by-event analysis. Corresponding efficiencies, S/B ratios, and widths of the peaks are given in Table 6.8. The small lifetime ( $c\tau$  of hundreds  $\mu\text{m}$ ) together with a low multiplicities of open charm particles require tighter cuts that leads to a smaller reconstruction efficiency.

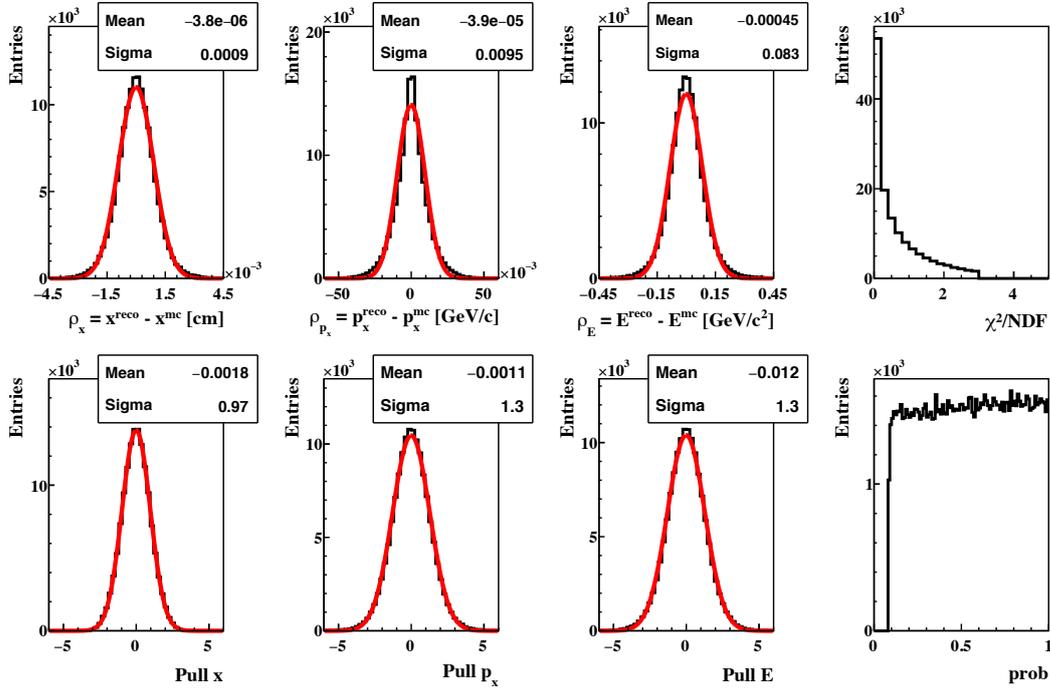


Figure 6.9: Fit quality of the reconstructed open charm particles at the example of  $\bar{D}^0$ . Distributions of residuals  $\rho$ , pulls,  $\chi^2$  criterion, and prob (probability for a certain  $\chi^2$  and NDF) demonstrate the high quality of the reconstructed parameters.

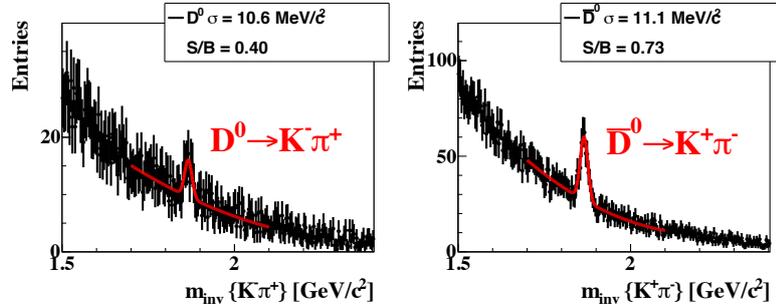


Figure 6.10: Mass spectra of  $D^0$  and  $\bar{D}^0$  mesons obtained by the KF Particle Finder package together with their signal to background (S/B) ratios for  $10^{11}$  simulated central NiNi events at 15 AGeV using TOF PID. The mixed event technique is used for the background estimation.

Particle	$\varepsilon_{method}, \%$	$\varepsilon_{4\pi}, \%$	S/B	$\sigma, \text{MeV}/c^2$
$D^0$	7.6	1.7	0.4	10.6
$\bar{D}^0$	7.6	1.7	0.7	11.1

Table 6.8: Efficiency of  $D^0$  and  $\bar{D}^0$  reconstruction, widths of the signal peaks, and S/B for  $10^{11}$  simulated central NiNi events at 15 AGeV using TOF PID.

## 6.2.5 Low mass vector mesons

The low mass vector mesons (LMVM) decay directly in the fireball because of their extremely low lifetime (see Table 6.9). On top of this, such particles can decay by the dilepton channels. Leptons do not interact strongly, therefore they pass through the baryonic matter that is formed during the collision without interaction and carry out exactly the information about the parent particle and its parameters. Thus, low mass vector mesons are perfectly suited for investigation of not only the late stages of the fireball evolution but also its very early stages. In case of the  $\phi$ -meson it is also possible to study a hadron decay channel  $\phi \rightarrow K^+K^-$  that carry out the information on the late stages of the collision. Decays of LMVM added to the KF Particle Finder are listed in Table 6.9.

Particle	Branching ratio	Full width $\Gamma$ , MeV/ $c^2$	$c\tau$ , fm
$\rho \rightarrow e^+e^-$	$4.7 \times 10^{-5}$	149.1	1.3
$\rho \rightarrow \mu^+\mu^-$	$4.6 \times 10^{-5}$	149.1	1.3
$\omega \rightarrow e^+e^-$	$7.3 \times 10^{-5}$	8.5	23.2
$\omega \rightarrow \mu^+\mu^-$	$9.0 \times 10^{-5}$	8.5	23.2
$\phi \rightarrow e^+e^-$	$3.0 \times 10^{-4}$	4.3	46.3
$\phi \rightarrow \mu^+\mu^-$	$2.9 \times 10^{-4}$	4.3	46.3
$\phi \rightarrow K^+K^-$	49%	4.3	46.3

Table 6.9: Parameters of the low mass vector mesons included in the KF Particle Finder reconstruction scheme: the implemented channel, its branching ratio, the full width ( $\Gamma$ ), and the lifetime ( $c\tau$ ) of a particle [5].

The high fit quality of the low mass vector mesons is illustrated in Fig. 6.11 at the example of  $\phi \rightarrow \mu^+\mu^-$  with STS used for track reconstruction.

The mass spectra of LMVM decaying by the dimuon channel obtained by the KF Particle Finder for  $10^{11}$  central AuAu events at 25 AGeV using MuCh PID and of  $\phi \rightarrow K^+K^-$  for 5M central AuAu UrQMD events at 10 AGeV using TOF PID are shown in Fig. 6.12. For simulation of the background in the dimuon channel the mixed event technique is used. Corresponding efficiencies, S/B ratios, and widths of the peaks are given in Table 6.10. The low branching ratio of dilepton channels require tight cuts on the PID stage that leads to a smaller reconstruction efficiency.

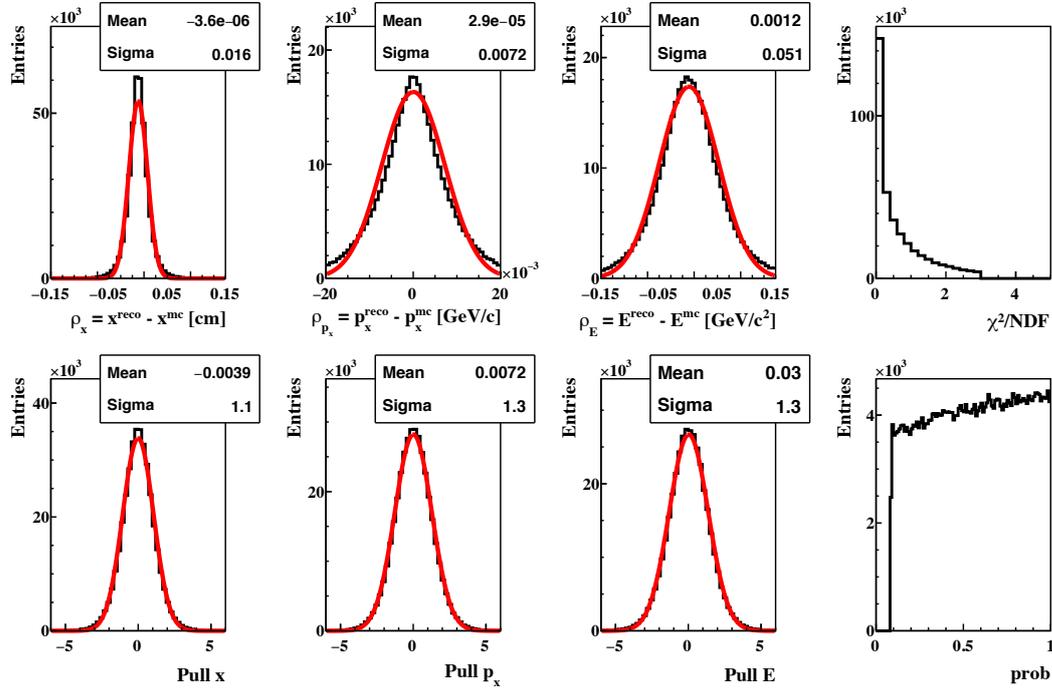


Figure 6.11: Fit quality of the reconstructed LMVM at the example of  $\phi \rightarrow \mu^+\mu^-$ . Distributions of residuals  $\rho$ , pulls,  $\chi^2$  criterion, and prob demonstrate the high quality of the reconstructed parameters. Irregularities in the prob distribution are due to some features in the current reconstruction procedure.

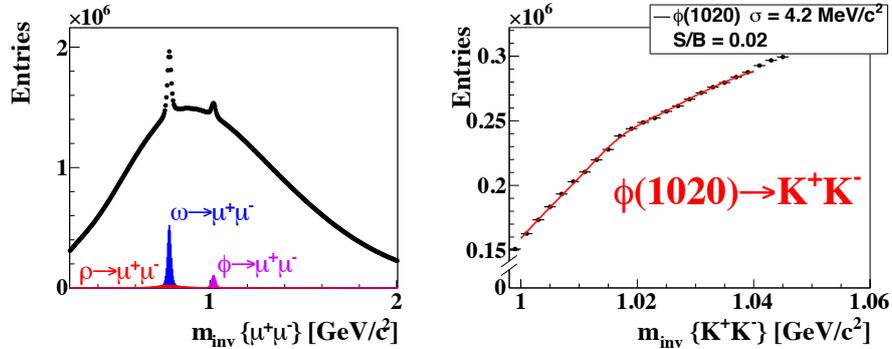


Figure 6.12: Mass spectra of LMVM together with their signal to background (S/B) ratios: left — in dimuon channel for  $10^{11}$  simulated central AuAu events at 25 AGeV using MuCh PID and mixed event technique, right — 5M simulated central AuAu UrQMD events at 10 AGeV using TOF PID.

Particle	$\varepsilon_{method}, \%$	$\varepsilon_{4\pi}, \%$	S/B	$\sigma, \text{MeV}/c^2$
$\rho \rightarrow \mu^+\mu^-$	45.4	1.2	0.009	149.0
$\omega \rightarrow \mu^+\mu^-$	42.1	1.1	0.2	8.5
$\phi \rightarrow \mu^+\mu^-$	48.5	2.1	0.04	4.2
$\phi \rightarrow K^+K^-$	85.6	15.7	0.02	4.2

Table 6.10: Efficiency of the LMVM reconstruction, width of the signal peak, and S/B.

### 6.2.6 Charmonium

One of the possibility for the  $c\bar{c}$  quark pair to hadronize is charmonium creation. The probability of the charmonium creation strongly depends on the properties of the baryonic medium formed in the collision, that makes it well suited for the QCD phase diagram exploration. It was predicted that in presence of the quark-gluon plasma phase due to color screening of the heavy quark potential the production of charmonium should be suppressed [95]. Charmonium suppression was first observed in central Pb+Pb collisions at 158 AGeV [96], and then also found in experiments at RHIC [97] and LHC [98].

As open charm, charmonium will be measured at SIS100 energies close to the production threshold for the first time. The measurements in collisions of medium size nuclei and in proton-nucleus collisions will be possible. SIS300 will allow the study of charmonium production in collisions of heavy nuclei [22, 31]. Charmonium decays added to the KF Particle Finder are listed in Table 6.11.

Particle	Branching ratio, %	Full width $\Gamma$ , keV/c <sup>2</sup>
$J/\psi \rightarrow e^+e^-$	6.0	92.9
$J/\psi \rightarrow \mu^+\mu^-$	6.0	92.9

Table 6.11: Parameters of the  $J/\psi$  decays included in the KF Particle Finder reconstruction scheme: the implemented channel, its branching ratio, and the full width ( $\Gamma$ ) of a particle [5].

The high fit quality of the charmonium is illustrated in Fig. 6.13 at the example of  $J/\psi \rightarrow \mu^+\mu^-$  with STS used for track reconstruction. The prob distribution is not flat due to some features in the current implementation of the hits reconstruction procedure. This effect is the most notable in case of high  $p_t$  daughter tracks, that is exactly the case of  $J/\psi$ .

The mass spectrum of  $J/\psi$  decaying by the dimuon channel obtained by the KF Particle Finder for  $10^{11}$  central AuAu events at 25 AGeV using MuCh PID is shown in Fig. 6.14. For simulation of the background the mixed event technique is used. Corresponding efficiency, S/B ratio, and the width of the peak are given in Table 6.12. The extremely low multiplicity of  $J/\psi$  requires tight cuts on the PID stage that leads to a smaller reconstruction efficiency.

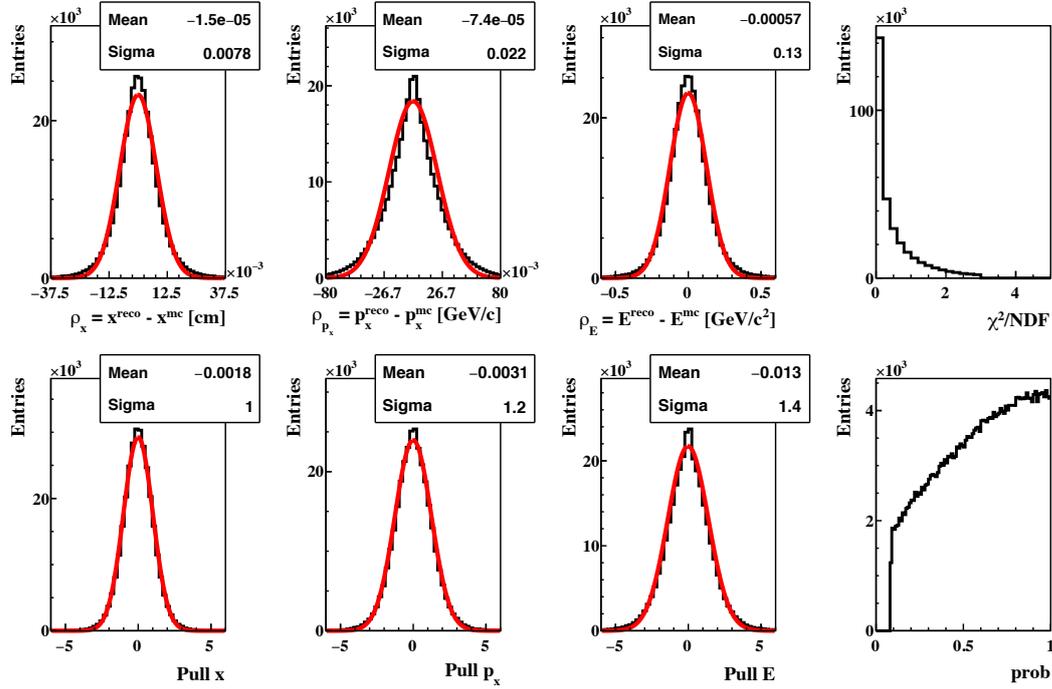


Figure 6.13: Fit quality of the reconstructed charmonium at the example of  $J/\psi \rightarrow \mu^+\mu^-$ . Distributions of residuals  $\rho$ , pulls,  $\chi^2$  criterion, and prob demonstrate the high quality of the reconstructed parameters. Irregularities in the prob distribution are due to some features in the current implementation of the hits reconstruction procedure.

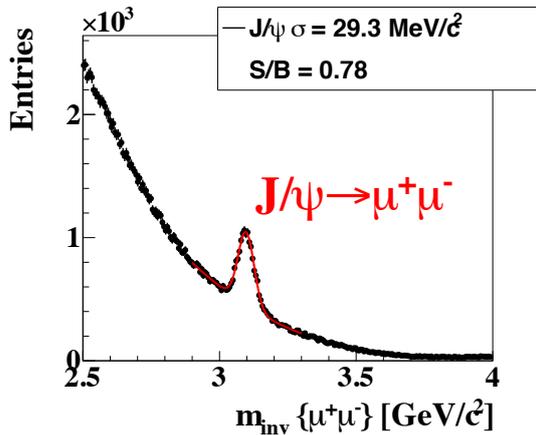


Figure 6.14: Mass spectra of  $J/\psi$  in dimuon channel reconstructed by the KF Particle Finder together with the signal to background (S/B) ratio for  $10^{11}$  simulated central AuAu events at 25 AGeV using MuCh PID. For the background simulation the mixed event technique is used.

Particle	$\varepsilon_{method}, \%$	$\varepsilon_{4\pi}, \%$	S/B	$\sigma, \text{MeV}/c^2$
$J/\psi \rightarrow \mu^+\mu^-$	15.1	6.9	0.78	29.3

Table 6.12: Efficiency of the  $J/\psi$  reconstruction, the width of the signal peak, and S/B for  $10^{11}$  simulated central AuAu events at 25 AGeV using MuCh PID.

### 6.2.7 $\gamma$ conversion and $\pi^0$

The main sources of the background electrons for the dileptons studies are  $\gamma$  conversion and Dalitz decays of  $\pi^0$  and  $\eta$  mesons. Because of the low branching ratios and multiplicities of the dileptons decays, background suppression is an essential issue for their reconstruction. Thus,  $\gamma$  conversion have to be carefully studied. Moreover, reconstruction of  $\gamma$  is important for studying the fast photons, that carry out the information of the physical processes taking place directly in the collision, and for a detailed analysis of the detector material.

The  $\gamma \rightarrow e^+e^-$  conversion can be reconstructed in the tracking system if it happens before or in the tracking system itself and both  $e^+$  and  $e^-$  are registered. The  $\pi^0$  and  $\eta$  mesons can be reconstructed in the tracking system as well by the full  $\gamma$  conversion method, when both  $\gamma$  are converted and found.

Particle	Branching ratio, %	$c\tau$ , nm
$\pi^0 \rightarrow \gamma\gamma$	98.8	25.50
$\eta \rightarrow \gamma\gamma$	39.4	0.15

Table 6.13: Parameters of  $\pi^0$  and  $\eta$  decays included in the KF Particle Finder reconstruction scheme: the implemented channel, its branching ratio, and the lifetime ( $c\tau$ ) of a particle [5].

The fit quality of these decays is illustrated in Fig. 6.15 at the example of  $\pi^0 \rightarrow \gamma\gamma$  with STS used for  $e^+$  and  $e^-$  tracks reconstruction. Since electrons are light particles, their energy losses are significant while passing the detector material. This leads to the momentum underestimation and, as a result, wider pulls. Also, the prob distribution is slightly affected.

The mass spectrum of reconstructed  $\gamma$  conversion and  $\pi^0$  reconstructed by the full  $\gamma$  conversion method obtained by the KF Particle Finder for 5M simulated central AuAu UrQMD events at 25 AGeV using RICH, TRD and TOF PID is shown in Fig. 6.16. Corresponding efficiencies, S/B ratios, and widths of the peaks are given in Table 6.14. The probability of  $\gamma$  to convert on the material of the target and the STS detector is about 6%. Thus, the reconstruction efficiency of the  $\gamma$  conversion and especially  $\pi^0$  is extremely low. However, the large abundance allows to perform a comprehensive study of these particles.

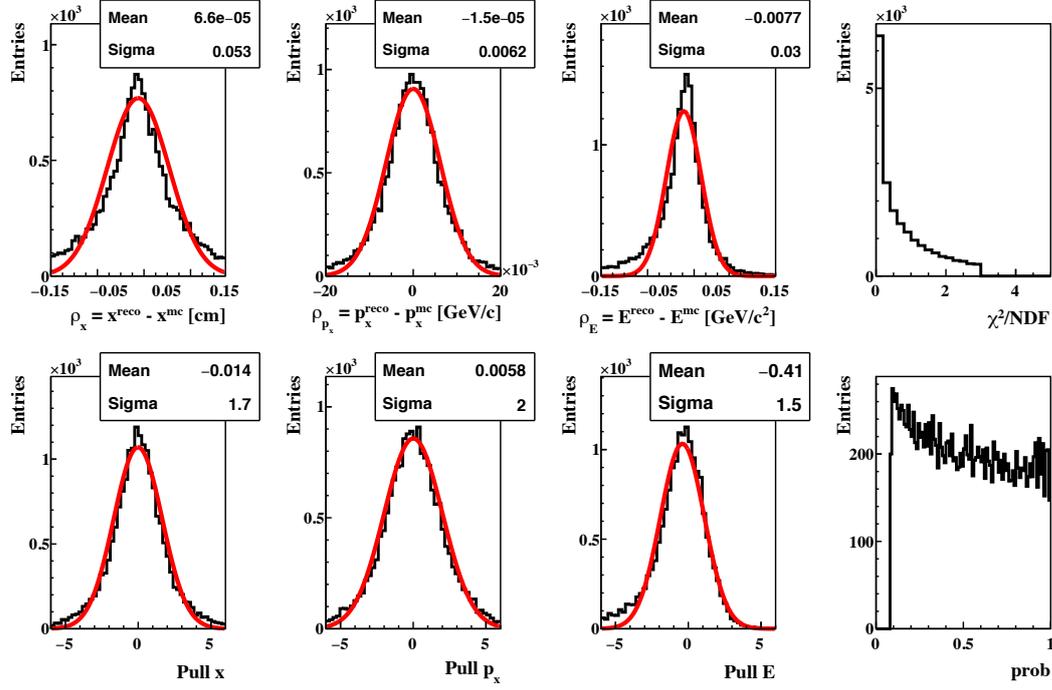


Figure 6.15: Fit quality of the  $\pi^0$  particles reconstructed by the full  $\gamma$  conversion method. Distributions of residuals  $\rho$ , pulls,  $\chi^2$  criterion, and prob (probability for a certain  $\chi^2$  and NDF) demonstrate the high quality of the reconstructed parameters.

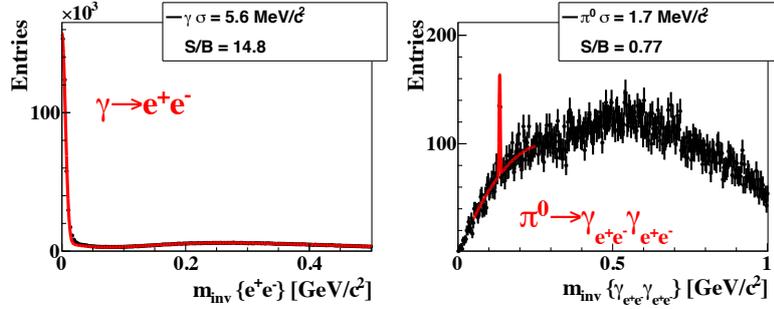


Figure 6.16: Mass spectra of  $\gamma$  conversion and  $\pi^0$  (from  $\gamma$ -conversion) reconstructed by the KF Particle Finder together with the signal to background (S/B) ratio for 5M simulated central AuAu UrQMD events at 25 AGeV using RICH, TRD and TOF PID.

Particle	$\varepsilon_{method}$	$\varepsilon_{4\pi}$	S/B	$\sigma$ , MeV/c <sup>2</sup>
$\gamma$	57.1	2.5e-4	14.8	5.6
$\pi^0$	21.4	6.5e-8	0.8	1.7

Table 6.14: Efficiency of  $\gamma$  conversion and  $\pi^0$  reconstruction, widths of the signal peaks, and S/B ratios for 5M simulated central AuAu UrQMD events at 25 AGeV using RICH, TRD and TOF PID.

## 6.3 Optimization of the package with respect to speed

The KF Particle Finder package is designed for the online reconstruction and selection of short-lived particles. Thus, the speed of the algorithm is as important, as its efficiency and reconstruction quality. For the speed optimization the KF Particle Finder next steps were done:

- the package is fully vectorized;
- all operations are performed in single precision;
- the full magnetic field map is approximated and excluded from the calculations;
- a fast and numerically stable matrix inversion algorithm is used in all procedures;
- fast algorithms for particles transport and search of the points of the closest approach are implemented.

To achieve the maximum speed of the reconstruction algorithm it is completely based on the SIMD instructions. For more efficient utilization of SIMD instructions, all operations are performed in single precision. In addition, operation in single precision allows to store a larger amount of data in the cache memory and, thus, reduce the number of accesses operations to the main memory that helps to achieve even higher acceleration of the program. The KF Particle package, the core of the KF Particle Finder, was also implemented in single precision and vectorized.

For more efficient utilization of SIMD instructions the full map of the magnetic field was excluded from the calculations [52]. To do this, each of the three components of the field is approximated along the trajectory of the daughter particles with a parabola, which depends on the  $z$  coordinate. The magnetic field in the CBM experiment has a smooth behaviour without sharp changes, therefore a parabola approximation has a sufficient accuracy for the problem under consideration. During reconstruction of a short-lived the average magnetic field at the decay point of the particle, at the primary vertex and at the point between them is calculated using the field approximation of the daughter particles. The

coefficients of the parabolas are calculated from obtained values.

The accuracy of the reconstructed particles should remain high in single precision. Thus, fast and numerically stable algorithms should be applied. The most unstable part of the KF Particle mathematics is the covariance matrices inversion. Since the covariance matrix is symmetric, it is possible to invert it by the Cholesky decomposition, which is both fast and stable (see Section 4.3): the error of the decomposition is stable and does not depend on the pivoting.

To speed up the transport of particles in a magnetic field the package uses a simplified analytic formula for extrapolation of the particle parameters [65]. The formula is obtained by decomposition of the equations of motion of a charged particle in a magnetic field in a Taylor series. Depending on the required accuracy the needed number of its terms are chosen, that allows to compromise between the speed and accuracy.

The search of the point of the closest approach are also optimized with respect to speed (see Sections 5.3.1–5.3.2). For this the point of the closest approach in the  $XZ$  plane is found at first and then corrected in space taking into account  $z$  coordinate by the Taylor expansion.

The optimization allows to reach a high computational speed of the KF Particle Finder of 1.5 ms per minimum bias and 10.5 ms per central AuAu collision at 25 AGeV beam energy.

## 6.4 KF Particle Finder on many-core systems

The KF Particle Finder is designed to perform the real-time online analysis on the dedicated FLES farm equipped with many-core processors. Therefore, it should scale with respect to the number of cores of the many-core servers, so that with the further increase of the number of cores on the newer architectures the performance of the package will not degrade and will grow proportionally.

For the scalability measurement groups of 1000 minimum bias AuAu collision at 25 AGeV beam energy were formed per each thread and distributed among the cores. The threads running an instance of the KF Particle Finder each were created with the Intel TBB library. The measurement is done on the NUMA server, therefore in order to prevent thread migration and the performance degradation

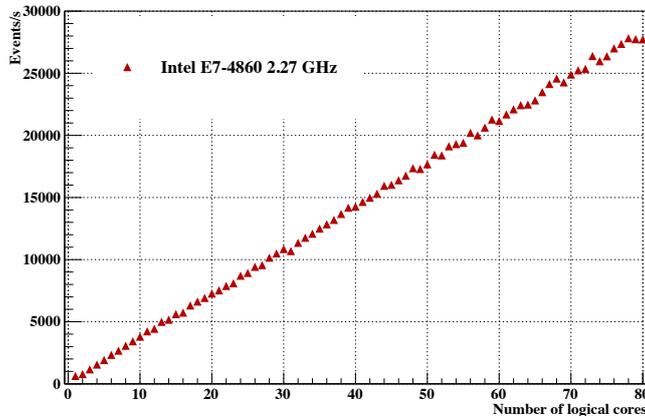


Figure 6.17: Scalability of the KF Particle Finder package on the many-core server lxir075.gsi.de equipped with four Intel E7-4860 (2.27 GHz) CPUs.

the thread to core affinity is set using the Pthreads library. The order of cores usage is set manually: at first, the first physical core is filled, then the next cores from the same CPU are filled, then the next CPU is filled in the same order.

The package shows strong linear scalability up to 80 cores achieving the speed of 28000 event/s (see Fig. 6.17).

## 6.5 KF Particle Finder in the FLES package

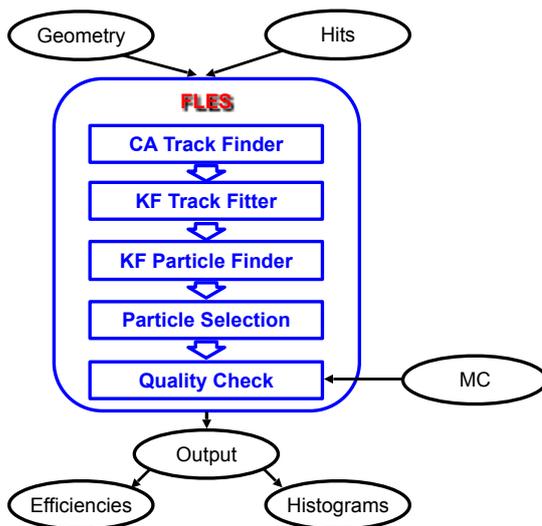


Figure 6.18: Block-diagram of the FLES package [99].

Due to its physics program, the CBM experiment will not have clear and simple criteria for the events selection based on the raw data from the detectors, therefore the usual triggering scheme will not be possible. Thus, the full reconstruction of the collision will be done in the online mode by the FLES package running on the dedicated many-core cluster. The KF Particle Finder will be used as a part of the FLES package for short-lived particle reconstruction and event selection based on the list of the reconstructed particles.

The FLES package of the CBM experiment [99] is intended to reconstruct the full topology of the event including tracks of charged particles and short lived particles. The first version of the package consists of the CA track finder, the Kalman filter track fitter and the KF Particle Finder package for short-lived particle reconstruction and events selection. The block-diagram of the FLES package is shown in Fig. 6.18.

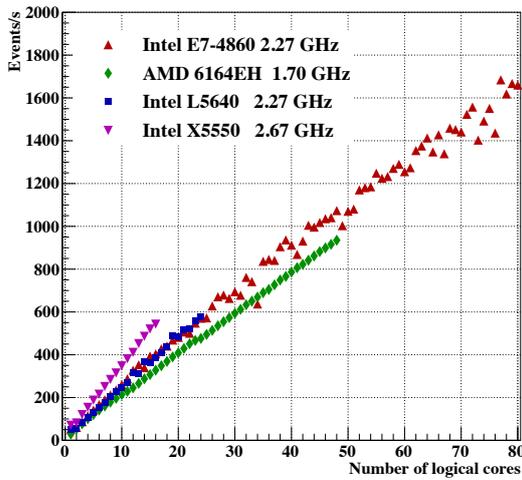


Figure 6.19: Scalability of the FLES package on different Intel and AMD many-core servers [99]. The FLES package shows strong linear scalability on all architectures.

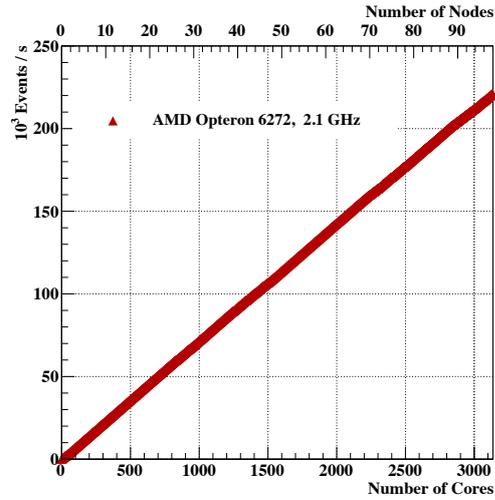


Figure 6.20: Scalability of the FLES package on FAIR-Russia HPC Cluster (ITEP, Moscow) [100]. The FLES package scales up to 3200 cores.

Operating on many-core clusters the FLES package should scale with respect to the number of cores. Its scalability was measured using four servers with Intel Xeon E7-4860, L5640 and X5550 processors and with AMD 6164EH processor. The AMD server has 4 processors with 12 physical cores each, in total 48 cores. All Intel processors have the Hyper-Threading technology, therefore each physical core can run two threads simultaneously. The most powerful Intel server has 4 processors with 10 physical cores each, that gives 80 logical cores in total.

The scalability measurement was performed similarly to the KF Particle Finder: groups of 1000 minimum bias AuAu collision at 25 AGeV beam energy were formed per each thread, threads were created using the Intel TBB library and the thread to core affinity was set using the Pthreads library. Fig. 6.19 shows

a strong scalability of the FLES package for all many-core systems achieving the reconstruction speed of 1700 events per second on the 80-cores server.

The scalability measurement was also performed on the many-core FAIR-Russia HPC Cluster (ITEP, Moscow) [100] with 100 nodes 32 cores each. Due to the structure of the cluster and limited access time, as the first step, input data was copied to each node beforehand and the measurement was started simultaneously on all nodes. As expected from such a simple parallel model, the FLES package shows the linear scalability there with a speed of  $2.2 \cdot 10^5$  events per second reached on 3200 cores of the cluster.

Thus, the KF Particle Finder is successfully included to the first version of the FLES package, which is fast, efficient and scalable.

Summarizing, being fast and efficient the KF Particle Finder will allow to collect spectra of short-lived particles with high signal to background ratios (see Fig. 6.21) in the online mode. The plots of the strange particles, hypermatter,  $\gamma$  and  $\pi^0$  represent the spectra for about 1 sec of data tacking with AuAu events at 10 AGeV beam energy, the plots of D-mesons — for about 2 weeks of data tacking with NiNi events at 15 AGeV beam energy, the plots of LMVM and  $J/\psi$  — for about 6 hours of data tacking with AuAu events at 25 AGeV.

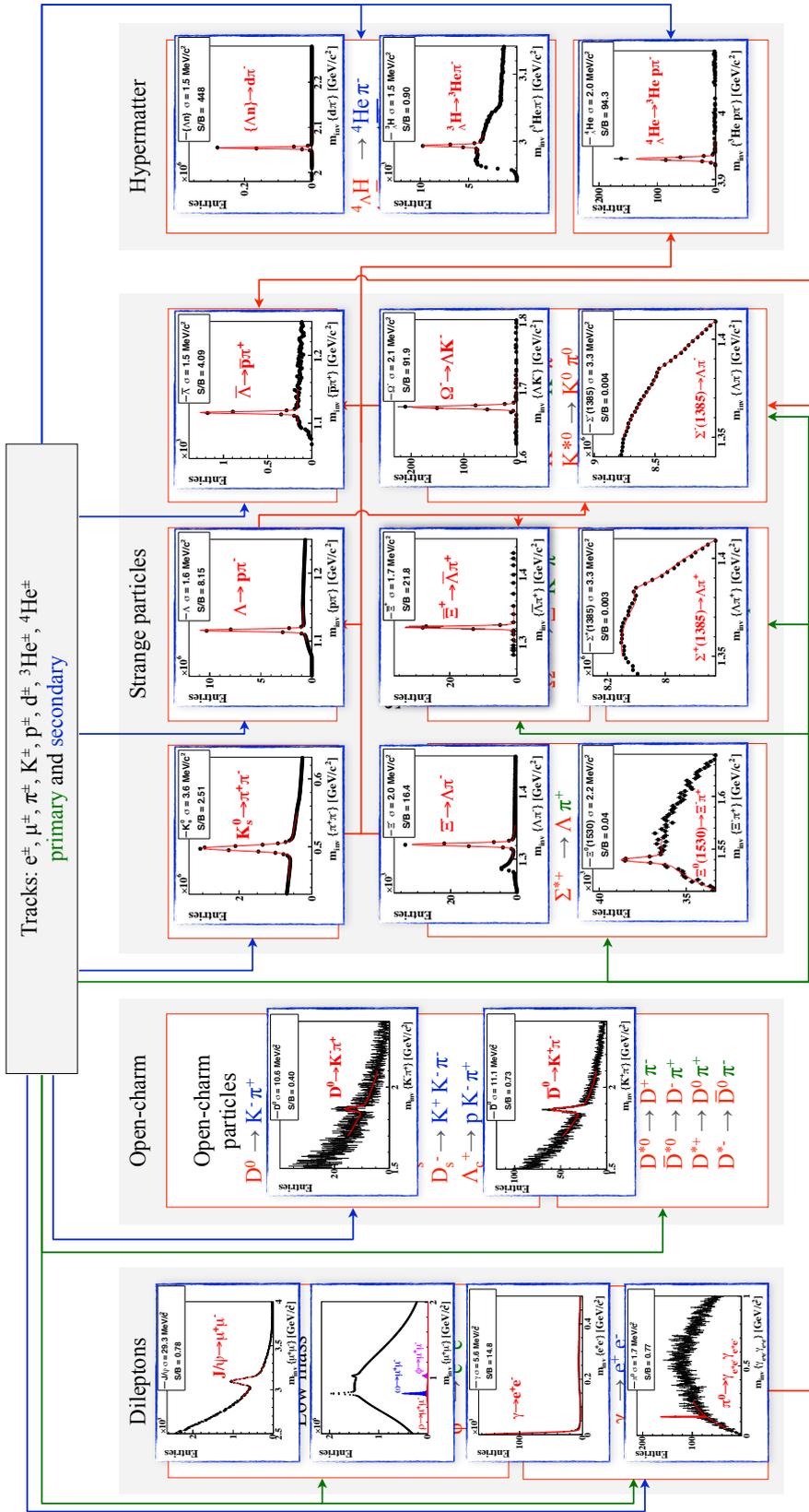


Figure 6.21: Block-diagram of the KF Particle Finder package with the mass distribution of the corresponding decays.

# Chapter 7

## Summary

Modern experiments in high energy physics and heavy ion experiments tend to increase the interaction rates and particle multiplicities per collision. In addition to the growing experimental complexity, there appears a necessity to study extremely rare probes that are sensitive to new physics and to conditions and features of matter existing during various stages of a nucleus-nucleus collision. High statistics measurements of rare probes require collection of a huge amount of data. However, currently available storage media do not allow to save the whole amount of the collected data. Thus, collisions, which potentially contain interesting information, should be selected online and stored.

The future CBM experiment will have no simple criteria for collision selection because of the complicated decay topology of particles like  $\Omega$  hyperons or  $D$  mesons. Therefore the collisions topology including short-lived particles have to be reconstructed online and events should be selected with fast algorithms running on a high-performance computing farm.

The thesis provides the following solution for the problem of online reconstruction and selection of short-lived particles.

First of all, an additional functionality was added to the set of tools for the Kalman-filter-based track fit. The reconstruction quality of short-lived particles strongly relies on the quality of the track fit, since parameters of the tracks and their covariance matrices are used as an input for the short-lived particle reconstruction.

The functionality of the KF Particle package was extended as well. KF Particle

is the core method for reconstruction of short-lived particles. The package is based on the Kalman filter mathematics. It provides a full set of the physics parameters of the particle along with their errors including position, momentum, mass, energy, lifetime, etc. The improvements provide a possibility to increase reconstruction quality of short-lived particles, especially for the decay trees, that are of particular importance for particles like strange hyperons, hypernuclei and resonances.

Based on these improved tools, the KF Particle Finder package for fast and efficient online search and selection of short-lived particles has been developed. It reconstructs more than 70 decays covering signals from all the physics cases of the CBM experiment: strange particles, strange resonances, hypernuclei, low mass vector mesons, charmonium and open-charm particles.

The KF Particle Finder is fast and efficient. All short-lived particles found by the package have high quality of the reconstructed parameters and thus, high efficiencies and signal to background ratios have been achieved.

The KF Particle Finder was successfully included into the first version of the FLES package for the CBM experiment, which will perform the full online event reconstruction.

The online running scenario requires from the developed packages to be both efficient and fast. Therefore in order to achieve a high speed of the reconstruction the developed packages are:

- heavily optimized and implemented in a single precision, therefore extremely fast;
- stabilized with respect to the single precision, that guaranties the high quality of the reconstruction;
- vectorized and parallelized;
- scalable on many-core processors and many-core clusters.

The KF Particle Finder is capable for further modifications, the list of the found particles can be extended in order to search for the already known or completely new, still unexplored particles. It provides an opportunity to reconstruct short-lived particles with one neutral daughter by the missing mass method. It perfectly suits for reconstruction of complicated cascade decays of hypernuclei

or hypothetical dibarion particles. Operating with a general model of a particle and being geometry independent, the package can be applied for other high energy physics experiments. The package is optimized with respect to memory consumption that makes it a perfect candidate for the hardware accelerators like GPGPUs and Intel Xeon Phi.

The KF Particle Finder package provides a universal platform for short-lived particle reconstruction, physics analysis and online selection.

# Appendix A

## Properties of the Kalman filter method on an example of the straight line

For a more complete understanding of the Kalman filter method, we will consider it first step by step on a simple example. For simplicity, the fit of a straight track in the plane is chosen. The example is a further update of the work [63] with a state vector appropriate for the CBM experiment. The case where stations are located along the detector axis  $x$  (see Fig. A.1) is considered. In this case, the stations measure  $y$ -coordinate at a given position  $x$ . For simplicity, we assume, that the detector stations are placed one from each other at equal distances  $L$ , the measurement errors of the  $y$ -coordinate at all stations are the same and equal to  $\sigma$ . Let us denote measurements of the  $y$  coordinates at each of the stations starting with the last  $y_1, y_2, y_3, y_4$  and  $y_5$ . Also, for simplicity of analysis, at first we consider the case of thin stations without multiple scattering of the charged particles on their material, and then consider how the presence of multiple scattering changes the track parameters and their covariance matrix. This approach is also useful for understanding of the Kalman filter mathematics. The problem, which should be solved in this example, is to find the estimation of the trajectory parameters in a region of a production point of the track IP. Therefore, the starting point for the Kalman filter method will be the last station.

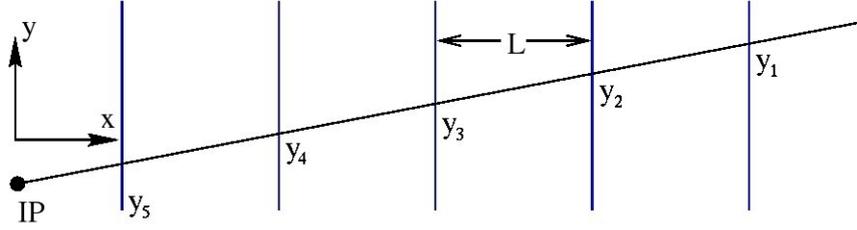


Figure A.1: A straight track on a plane. IP is a production point of the track. The line, that starts at the point IP, shows a trajectory of the track. Vertical lines represent a position of detector stations placed at equal distances  $L$  from each other.

Since the track is straight, the corresponding equation of motion is:

$$y = t_y x + b.$$

Let's choose the state vector of the system, that contains a value of the  $y$ -coordinate and a slope of the track  $t_y$  to the axis  $x$ , and construct according to this choice the covariance matrix:

$$\mathbf{r} = \begin{pmatrix} y \\ t_y \end{pmatrix}, \quad C = \begin{pmatrix} C_{yy} & C_{yt_y} \\ C_{yt_y} & C_{t_y t_y} \end{pmatrix}. \quad (\text{A.1})$$

Since the detector station measures only one coordinate  $y$ , the vector measurements degenerates into a scalar as well as the covariance matrix that corresponds to it:

$$\mathbf{m}_k = \{y_k^m\}, \quad V_k = \sigma^2.$$

According to (4.2) the model of measurement takes the form:

$$H_k = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

Let's simplify equations (4.6a)-(4.6e) for the considered example:

$$\begin{aligned}
S_k &= \frac{1}{\sigma^2 + C_{kyy}^-}, \\
K_k &= \frac{1}{\sigma^2 + C_{kyy}^-} \begin{pmatrix} C_{kyy}^- \\ C_{kyty}^- \end{pmatrix}, \\
\zeta_k &= y_k^m - y_k^-, \\
\mathbf{r}_k^+ &= \begin{pmatrix} y_k^- + \frac{C_{kyy}^-}{\sigma^2 + C_{kyy}^-} (y_k^m - y_k^-) \\ t_{y_k}^- + \frac{C_{kyty}^-}{\sigma^2 + C_{kyy}^-} (y_k^m - y_k^-) \end{pmatrix}, \\
C_k^+ &= \begin{pmatrix} C_{kyy}^- \left(1 - \frac{C_{kyy}^-}{\sigma^2 + C_{kyy}^-}\right) & C_{kyty}^- \left(1 - \frac{C_{kyy}^-}{\sigma^2 + C_{kyy}^-}\right) \\ C_{kyty}^- \left(1 - \frac{C_{kyy}^-}{\sigma^2 + C_{kyy}^-}\right) & C_{kyty}^- - \frac{C_{kyty}^{-2}}{\sigma^2 + C_{kyy}^-} \end{pmatrix}.
\end{aligned} \tag{A.2}$$

## Extrapolation of the state vector between measurements

Before adding each next measurement to the state vector, track parameters should be transported to the point of the measurement, i.e. to the corresponding station of the detector. Because stations are placed at a distance  $L$  from each other and measurements are taken into account in the direction opposite to the direction of the axis  $x$ , the position of track  $y_k^-$  at the next station just before adding the measurement  $y_k^m$  is calculated by formulas:

$$y_k^- = t_{y_{k-1}}^+ (x_{k-1} - L) + b = y_{k-1}^+ - t_{y_{k-1}}^+ L.$$

Since the slope of the track remains the same during extrapolation of the parameters, than

$$t_{y_k}^- = t_{y_{k-1}}^+.$$

For extrapolation of the covariance matrix formulas (4.5b) are used, but at this stage it is assumed that multiple scattering on the material of the detector station is absent, so the covariance matrix of the process noise  $Q_k = 0$ . By definition, the extrapolation matrix:

$$F_{k-1} = \begin{pmatrix} \frac{\partial y_k^-}{\partial y_{k-1}^+} & \frac{\partial y_k^-}{\partial t_{y_{k-1}}^+} \\ \frac{\partial t_{y_k}^-}{\partial y_{k-1}^+} & \frac{\partial t_{y_k}^-}{\partial t_{y_{k-1}}^+} \end{pmatrix} = \begin{pmatrix} 1 & -L \\ 0 & 1 \end{pmatrix}$$

Substituting the resulting expression in formulas (4.5a) and (4.5b) we will find formulas for extrapolation of the straight line on the plane in the absence of multiple scattering on the material of the detector station:

$$\begin{aligned} r_k^- &= \begin{pmatrix} y_{k-1}^+ - t_{y_{k-1}}^+ L, \\ t_{y_{k-1}}^+ \end{pmatrix}, \\ C_k^- &= \begin{pmatrix} C_{k-1yy}^+ - 2LC_{k-1yt_y}^+ + L^2C_{k_{t_y t_y}}^+ & C_{k-1yt_y}^+ - LC_{k-1t_y t_y}^+ \\ C_{k-1yt_y}^+ & C_{k-1t_y t_y}^+ \end{pmatrix}. \end{aligned} \quad (\text{A.3})$$

## Initialization of the parameters

Let's assume, that we have an estimation of track parameters obtained, for example, with the least squares method. The covariance matrix is set as a diagonal matrix, in which the diagonal elements tend to infinity:

$$\mathbf{r}_0 = \begin{pmatrix} y_0 \\ t_{y_0} \end{pmatrix} \quad C_0 = \begin{pmatrix} C_{0yy} & 0 \\ 0 & C_{0t_y t_y} \end{pmatrix}. \quad (\text{A.4})$$

## Filtration of the first measurement

Before the first measurement addition no other calculations are needed and  $\mathbf{r}_1^- = \mathbf{r}_0$  and  $C_1^- = C_0$ . Note, that  $C_{0yy}$  tends to infinity, so the expression  $\frac{C_{0yy}^-}{\sigma^2 + C_{0yy}^-}$  introduces uncertainty that leads to numerical instability in calculations. To get rid of the uncertainty, we decompose this expression in a Taylor series, and since  $C_{0yy}$  tends to infinity, we consider only terms of the first degree of smallness:

$$\frac{C_{0yy}}{\sigma^2 + C_{0yy}} = \frac{1}{1 + \frac{\sigma^2}{C_{0yy}}} \approx 1 - \frac{\sigma^2}{C_{0yy}}. \quad (\text{A.5})$$

Taking into account the expression (A.5) and the fact, that  $\frac{\sigma^2}{C_{0_{yy}}} \approx 0$ , formulas (A.2) give an expression for the state vector and its covariance matrix after adding the first measurement:

$$\begin{aligned} \mathbf{r}_1^+ &= \begin{pmatrix} y_1 \\ t_{y_0} \end{pmatrix}, \\ C_1^+ &= \begin{pmatrix} \sigma^2 & 0 \\ 0 & C_{0_{t_y t_y}} \end{pmatrix}. \end{aligned} \quad (\text{A.6})$$

The analysis of obtained expressions for the covariance matrix shows, that already after the first measurement the position of the track is known with a high accuracy, but the information about the slope of the track is almost absent. Interesting is the fact, that already at this step the dependence on the initial value of  $C_{0_{yy}}$  is lost. Thus, the choice of the initial covariance matrix with large diagonal elements allows to estimate the system parameters based only on measurements themselves instead of the initial approximation.

## Extrapolation to the point of second measurement

Before filtering the next measurement let's find the estimation of track parameters and their covariance matrix at the corresponding detector station putting formulas (A.6) in formulas (A.3):

$$\begin{aligned} \mathbf{r}_2^- &= \begin{pmatrix} y_1 - t_{y_0} L \\ t_{y_0} \end{pmatrix}, \\ C_2^- &= \begin{pmatrix} \sigma^2 + L^2 C_{0_{t_y t_y}} & -L C_{0_{t_y t_y}} \\ -L C_{0_{t_y t_y}} & C_{0_{t_y t_y}} \end{pmatrix}. \end{aligned} \quad (\text{A.7})$$

Note, that after extrapolation the correlation coefficients of the parameters contains large initial values. The error of the track position is increased again, since in extrapolation value of the slope with an infinite accuracy is used.

## Filtration of the second measurement

When estimation of track parameters at the position of the the second station is obtained, a corresponding measurement should be added using again formulas (A.2):

$$\begin{aligned}
 y_2^+ &= y_1 - t_{y_0}L + \frac{\sigma^2 + L^2C_{0t_yt_y}}{2\sigma^2 + L^2C_{0t_yt_y}}(y_2 - y_1 + t_{y_0}L), \\
 t_{y_2}^+ &= t_{y_0} - \frac{LC_{0t_yt_y}}{2\sigma^2 + L^2C_{0t_yt_y}}(y_2 - y_1 + t_{y_0}L), \\
 C_{2y_2}^+ &= (\sigma^2 + L^2C_{0t_yt_y}) \left(1 - \frac{\sigma^2 + L^2C_{0t_yt_y}}{2\sigma^2 + L^2C_{0t_yt_y}}\right) = \frac{(\sigma^2 + L^2C_{0t_yt_y})}{2\sigma^2 + L^2C_{0t_yt_y}}\sigma^2 \\
 C_{2y_2t_y}^+ &= -LC_{0t_yt_y} \left(1 - \frac{\sigma^2 + L^2C_{0t_yt_y}}{2\sigma^2 + L^2C_{0t_yt_y}}\right) = -\frac{LC_{0t_yt_y}}{2\sigma^2 + L^2C_{0t_yt_y}}\sigma^2 \\
 C_{2t_yt_y}^+ &= C_{0t_yt_y} - \frac{L^2C_{0t_yt_y}^2}{2\sigma^2 + L^2C_{0t_yt_y}} = \frac{2\sigma^2C_{0t_yt_y}}{2\sigma^2 + L^2C_{0t_yt_y}}.
 \end{aligned}$$

Since  $C_{0t_yt_y}$  tends to infinity some expressions can be simplified by expanding in a Taylor series:

$$\begin{aligned}
 \frac{\sigma^2 + L^2C_{0t_yt_y}}{2\sigma^2 + L^2C_{0t_yt_y}} &\approx \frac{L^2C_{0t_yt_y}}{L^2C_{0t_yt_y}} = 1, \\
 \frac{LC_{0t_yt_y}}{2\sigma^2 + L^2C_{0t_yt_y}} &\approx \frac{LC_{0t_yt_y}}{L^2C_{0t_yt_y}} = \frac{1}{L}, \\
 \frac{C_{0t_yt_y}}{2\sigma^2 + L^2C_{0t_yt_y}} &\approx \frac{C_{0t_yt_y}}{L^2C_{0t_yt_y}} = \frac{1}{L^2}.
 \end{aligned}$$

Then the state vector of the track and its covariance matrix take a simple form:

$$\begin{aligned} \mathbf{r}_2^+ &= \begin{pmatrix} y_2 \\ \frac{y_1 - y_2}{L} \end{pmatrix}, \\ C_2^+ &= \begin{pmatrix} \sigma^2 & -\frac{\sigma^2}{L} \\ \sigma^2 & \frac{2\sigma^2}{L^2} \end{pmatrix}. \end{aligned} \tag{A.8}$$

Note, that after the second measurement dependence on the initial approximation disappeared completely. This is due to the properties of the Kalman filter and the symmetry of the problem.

### **Extrapolation of the track to the third measurement and filtration of the measurement**

For track extrapolation to the third station we substitute formulas (A.8) into formulas (A.3):

$$\begin{aligned} \mathbf{r}_3^- &= \begin{pmatrix} 2y_2 - y_1 \\ \frac{y_1 - y_2}{L} \end{pmatrix}, \\ C_3^- &= \begin{pmatrix} 5\sigma^2 & -\frac{3\sigma^2}{L} \\ 3\sigma^2 & \frac{2\sigma^2}{L^2} \end{pmatrix}. \end{aligned} \tag{A.9}$$

Let's analyze obtained result. In extrapolation the slope of the track does not change, therefore corresponding element of the covariance matrix also remains unchanged. The error of the track position estimation increases again, because the slope, which is used for its calculation, is also known with some accuracy. But this increase is smaller, than on the previous step, since now the error of the slope is completely determined by measurements and not by the infinite initial approximation.

Now let us add the third measurement by putting formulas (A.9) into (A.2):

$$\begin{aligned} \mathbf{r}_3^+ &= \begin{pmatrix} \frac{2y_2 - y_1 + 5y_3}{\frac{6}{2L} \frac{y_1 - y_3}{2L}} \end{pmatrix}, \\ C_3^+ &= \begin{pmatrix} \frac{5}{6}\sigma^2 & -\frac{\sigma^2}{2L} \\ \frac{\sigma^2}{-2L} & \frac{\sigma^2}{2L^2} \end{pmatrix}. \end{aligned} \tag{A.10}$$

As one can see, the covariance matrix decreases again. And all the elements are smaller comparing to the case of the second measurement. Also the estimation of the track slope does not contain an information about the measurement on the second station, however, this is not a law but just a consequence of the symmetry in the considered problem. With further calculations the contribution of the second measurement will return to this expression.

### Extrapolation of the track to the fourth measurement and filtration of the measurement

For extrapolation of the track to the fourth station the formula (A.10) is substituted into (A.3):

$$\begin{aligned} \mathbf{r}_4^- &= \begin{pmatrix} \frac{y_2 - 2y_1 + 4y_3}{\frac{3}{2L} \frac{y_1 - y_3}{2L}} \end{pmatrix}, \\ C_4^- &= \begin{pmatrix} \frac{7}{3}\sigma^2 & -\frac{\sigma^2}{L} \\ \frac{\sigma^2}{-L} & \frac{\sigma^2}{2L^2} \end{pmatrix}. \end{aligned} \tag{A.11}$$

In order to add the fourth measurement formulas (A.11) are substituted into (A.2):

$$\mathbf{r}_4^+ = \begin{pmatrix} \frac{-2y_1 + y_2 + 4y_3 + 7y_4}{\frac{10}{10L} \frac{3y_1 + y_2 - y_3 - 3y_4}{10L}} \end{pmatrix},$$

$$C_4^+ = \begin{pmatrix} \frac{7}{10}\sigma^2 & -\frac{3\sigma^2}{10L} \\ \frac{3\sigma^2}{10L} & \frac{\sigma^2}{5L^2} \end{pmatrix}. \quad (\text{A.12})$$

Adding the fourth measurement shows nothing new at this stage, but this step will be helpful further for illustration how multiple scattering impacts parameters of the track and their covariance matrix. Next measurements does not provide a new information and we will stop on four measurements.

Note, that expansion in a Taylor series of expressions, that contain infinite diagonal elements of the initial covariance matrix, in this example is done only for clear explanations of the properties of the state vector and its covariance matrix. Calculations in real applications use the exact formulas, so the state vector and covariance matrix depend although slightly on the initial approximation. Therefore, the initial selection of the covariance matrix is extremely important.

## **Influence of multiple scattering on track parameters and their covariance matrix**

Now consider the case, when the station of the detector have a finite thickness. For simplicity, we will consider them thin. This situation is quite common for the high energy physics experiments and it is exactly the case of the tracking detectors of the CBM experiment. In this case multiple scattering gives a contribution only to the errors of the track slope estimation and does not affect the accuracy of the estimated track position. In our simple example it does not affect nondiagonal elements of the covariance matrix. We assume, that all stations contribute with the same error in the estimation of the track slope and this contribution equal to  $\delta$ .

To take into account the material during extrapolation of the state vector between the stations the formula (4.5b) is used. The matrix  $Q_k$  corresponds to the covariance matrix of the process noise, which is caused by multiple scattering in the material, and has the form:

$$Q_k = \begin{pmatrix} 0 & 0 \\ 0 & \delta^2 \end{pmatrix}.$$

This formalism is based on the assumption that the state vector and its covariance matrix is first extrapolated to each next station after filtration of a measurement, and then the covariance matrix is modified to take into account multiple scattering. Thus, it is assumed that the material of the detector must be considered just before filtering a measurement, i.e. measurement of the track coordinates is done by the outer side of the station with respect to the direction of fit. Often, however, a situation arises, when the coordinates are measured by the whole working volume of the detector and not by only one side. In this case, the method is approximate, and adding the noise matrix  $Q_k$  before filtration or after is a subject of an agreement.

Since the error  $\delta$  caused by multiple scattering is a small value comparing to infinite element of the covariance matrix  $C_{0_{t_y t_y}}$ , the material does not change formulas (A.6)-(A.7) and, as a consequence, formulas (A.8). The first formulas, which will be affected by the process noise, are the formulas of extrapolation to the third measurement (A.9):

$$\mathbf{r}_3^- = \begin{pmatrix} 2y_2 - y_1 \\ \frac{y_1 - y_2}{L} \end{pmatrix},$$

$$C_3^- = \begin{pmatrix} 5\sigma^2 & -\frac{3\sigma^2}{L} \\ -\frac{3\sigma^2}{L} & \frac{2\sigma^2}{L^2} + \delta^2 \end{pmatrix}.$$

Accordingly, parameters and their covariance matrix after filtration of the third measurement will be changed:

$$\mathbf{r}_3^+ = \begin{pmatrix} \frac{2y_2 - y_1 + 5y_3}{6} \\ \frac{y_1 - y_3}{2L} \end{pmatrix},$$

$$C_3^+ = \begin{pmatrix} \frac{5}{6}\sigma^2 & -\frac{\sigma^2}{2L} \\ -\frac{\sigma^2}{2L} & \frac{\sigma^2}{2L^2} + \delta^2 \end{pmatrix}. \tag{A.13}$$

At this stage, the result is obvious: multiple scattering influences only the error of the slope estimation. Now check the changes, which will occur after

extrapolation to the fourth measurement:

$$\mathbf{r}_4^- = \begin{pmatrix} \frac{y_2 - 2y_1 + 4y_3}{3} \\ \frac{y_1 - y_3}{2L} \end{pmatrix},$$

$$C_4^- = \begin{pmatrix} \frac{7}{3}\sigma^2 + \delta^2 L^2 & -\frac{\sigma^2}{L} - \delta^2 L \\ -\frac{\sigma^2}{L} - \delta^2 L & \frac{\sigma^2}{2L^2} + 2\delta^2 \end{pmatrix}.$$

The covariance matrix is increased with respect to the case without multiple scattering: the errors of the parameters are increased as well as the correlation coefficients. Note, that the error of the track position estimation now also contains a term caused by multiple scattering. However, the state vector is not yet affected. Let's add the fourth measurement:

$$\mathbf{r}_4^+ = \begin{pmatrix} \frac{y_2 - 2y_1 + 4y_3}{3} + \frac{\frac{7}{3}\sigma^2 + \delta^2 L^2}{\frac{10}{3}\sigma^2 + \delta^2 L^2} \left( y_4 - \frac{y_2 - 2y_1 + 4y_3}{3} \right) \\ \frac{y_1 - y_3}{2L} - \frac{\sigma^2 + \delta^2 L^2}{\left( \frac{10}{3}\sigma^2 + \delta^2 L^2 \right) L} \left( y_4 - \frac{y_2 - 2y_1 + 4y_3}{3} \right) \end{pmatrix}, \quad (\text{A.14})$$

$$C_4^+ = \begin{pmatrix} \frac{\frac{7}{3}\sigma^2 + \delta^2 L^2}{\frac{10}{3}\sigma^2 + \delta^2 L^2} \sigma^2 & -\frac{\sigma^2}{L} \cdot \frac{\sigma^2 + \delta^2 L^2}{\frac{10}{3}\sigma^2 + \delta^2 L^2} \\ -\frac{\sigma^2}{L} \cdot \frac{\sigma^2 + \delta^2 L^2}{\frac{10}{3}\sigma^2 + \delta^2 L^2} & \frac{\sigma^2}{L^2} \cdot \frac{\frac{2}{3}\sigma^2 + \frac{31}{6}\delta^2 L^2 + \delta^4 L^4 / \sigma^2}{\frac{10}{3}\sigma^2 + \delta^2 L^2} \end{pmatrix}.$$

The obtained formulas are complicated and are difficult to analyze, therefore let us consider several examples with respect to the  $\delta$  value. In our estimations we will use values similar to parameters of the STS detector of the CBM experiment:  $\sigma \approx 17 \mu\text{m}$ ,  $L = 10 \text{ cm}$ .

First, we consider the case of pions with momentum of 10 GeV/c. For them  $\delta \cdot L \approx 0.4\sigma$  and terms, that correspond to multiple scattering, can be neglected. Formulas in this case will not differ from the case without multiple

scattering (A.12):

$$\begin{aligned} \mathbf{r}_4^+ &= \begin{pmatrix} -0.2y_1 + 0.1y_2 + 0.4y_3 + 0.7y_4 \\ \frac{1}{L}(0.3y_1 + 0.1y_2 - 0.1y_3 - 0.3y_4) \end{pmatrix}, \\ C_4^+ &= \begin{pmatrix} 0.7\sigma^2 & -0.3\frac{\sigma^2}{L} \\ -0.3\frac{\sigma^2}{L} & 0.2\frac{\sigma^2}{L^2} \end{pmatrix}. \end{aligned} \quad (\text{A.15})$$

Now consider the case, when  $\delta \cdot L = \sigma$ , that corresponds to the case of pions with momentum of 5 GeV. In this case, the expression (A.14) is greatly simplified:

$$\begin{aligned} \mathbf{r}_4^+ &= \begin{pmatrix} -0.15y_1 + 0.08y_2 + 0.31y_3 + 0.77y_4 \\ \frac{1}{L}(0.19y_1 + 0.15y_2 + 0.12y_3 - 0.46y_4) \end{pmatrix}, \\ C_4^+ &= \begin{pmatrix} 0.77\sigma^2 & -0.46\frac{\sigma^2}{L} \\ -0.46\frac{\sigma^2}{L} & 1.58\frac{\sigma^2}{L^2} \end{pmatrix}. \end{aligned} \quad (\text{A.16})$$

Let's compare the obtained results with the expressions from the previous case (A.15). First, note, that the state vector is changed: the contribution of the fourth measurement  $y_4$  into estimation of the track position is increased, the estimation of the track slope become asymmetric with respect to the measurements. Values of the covariance matrix elements are again increased, when multiple scattering is present. One result of this behavior is that the next measurement will make a smaller contribution to the estimation of the track slope.

Now consider the case of pions with momentum of 1 GeV. In this case  $\delta \cdot L \approx 4\sigma$  and expressions (A.14) are changed as follows:

$$\begin{aligned} \mathbf{r}_4^+ &= \begin{pmatrix} -0.03y_1 + 0.02y_2 + 0.07y_3 + 0.95y_4 \\ \frac{1}{L}(-0.09y_1 + 0.29y_2 + 0.67y_3 - 0.88y_4) \end{pmatrix}, \\ C_4^+ &= \begin{pmatrix} 0.95\sigma^2 & -0.88\frac{\sigma^2}{L} \\ -0.88\frac{\sigma^2}{L} & 17.5\frac{\sigma^2}{L^2} \end{pmatrix}. \end{aligned} \quad (\text{A.17})$$

As a result of the increased error caused by the multiple scattering, the track

position is mainly determined by the position of the fourth measurement. The contribution of the first measurement in the track slope is decreased several times. The contribution of the second and especially the third and fourth measurements is increased. Thus, the track parameters practically do not depend on the first measurement and the most sensitive to the last two. The covariance matrix is also increased: the error of the slope estimation is increased by one order of magnitude compared to the case of pions with momentum of 5 GeV and two orders of magnitude compared to the case without multiple scattering.

# List of Figures

1.1	Schematic depiction of particles and boson mediators in the Standard model of particle physics. Parameters of the particles are taken from PDG 2014 [5]. Charge is given in terms of the elementary charge $e$ . . . . .	4
1.2	The lowest SU(3) baryon and meson multiplets together with the quark structure of particles: the baryon octet (a) with spin $S = \frac{1}{2}$ , the baryon decuplet (b) with $S = \frac{3}{2}$ , the pseudo-scalar nonet (c) with $S = 0$ and the vector nonet (d) with $S = 1$ . . . . .	5
1.3	A possible form of the QCD phase diagram [10]. . . . .	6
1.4	The simulated central Au-Au collision at 25 AGeV energy in the CBM experiment with about 1000 charged particles (left) with a highlighted $\bar{\Omega}^+$ decay (right). Different colors correspond to different types of particles. $\bar{\Omega}^+$ decays into $\bar{\Lambda}$ and $K^+$ with further decay of $\bar{\Lambda}$ to antiproton ( $\bar{p}$ ) and $\pi^+$ . One of such decay is expected per about 1000 central collisions. . . . .	8
2.1	Layout of the Facility for Antiproton and Ion Research (FAIR) [20]. Blue lines represents existing GSI facilities, the new accelerator complex is shown with red. The superconducting synchrotrons SIS100 or SIS300 will provide ion beams for the CBM experiment. They also will provide secondary beams of antiprotons or rare nuclear isotopes from a production target to the fragment separator (super FRS) and storage rings, such as the HESR for the antiproton research program. . . . .	11

- 
- 2.2 The layout of the CBM detector setup [22]. Options for electron measurements (left) and muon measurements (right) are presented. The electron setup includes the dipole magnet with a large acceptance, the Micro-Vertex Detector (MVD), the Silicon Tracking System (STS), the Ring Imaging CHerenkov detector (RICH), the Transition Radiation Detector (TRD), the Time Of Flight detector (TOF), the Electromagnetic CALorimeter (ECAL) and the Projectile Spectator Detector (PSD). In case of the muon setup RICH, TRD and ECAL are replaced with the Muon Chambers (MuCh). 14
- 2.3 The MVD detector of the CBM experiment [31, 32]. MVD will be built from MAPS sensors grouped into 2–4 stations. To reduce the amount of material MVD will be placed in vacuum. The detector will be mounted inside the dipole magnet. . . . . 15
- 2.4 The STS detector of the CBM experiment [33]. STS will be build from silicon microstrip detector modules grouped into lightweight vertical ladders with carbon fibers as the support structure. Ladders will form eight stations. The STS detector will be mounted inside the dipole magnet. . . . . 16
- 2.5 The RICH geometry including mirrors, two photodetector planes and the support structure (left) and rings produced in a simulated AuAu collision at 25 AGeV by the Cherenkov light at the photodetector planes (right) [35]. . . . . 17
- 2.6 The MuCh geometry (left) including absorbers (red) and gaseous detectors of charged particles (yellow) together with the support structure (green) [36] and muon stopping power depending on the momentum (right) [5]. . . . . 18
- 2.7 The TRD geometry consisting of rectangular pads for transition radiation detection and particles tracking [31] (left) and the scheme of a pad (right) [37]. When crossing the radiator of a pad, particle produces transition radiation, which is registered by the gaseous detector. . . . . 19

2.8	The TOF geometry (left) including gaseous detectors, grouped into modules, and the support structure [38] with (right) the distribution of the mass squared ( $m^2$ ) versus momentum obtained for positively charged hadrons — pions ( $\pi^+$ ), kaons ( $K^+$ ) and protons (p) and their antiparticles; negative sign of momentum stands for negative particles [39]. . . . .	20
2.9	The ECAL geometry (left) including moving sections with about 1000 modules, which consist of absorbers followed by the scintillator detectors [40] and a scheme of the electromagnetic shower produced in the detector (right). . . . .	21
2.10	The PSD geometry (left) and the design of a module consisting of consecutive absorber and scintillator layers (right) [41]. . . . .	22
2.11	The CBM data acquisition model in comparison to the conventional approach [42]. Usually the input data passes several trigger levels being gradually suppressed. Due to its running scenario, the CBM experiment will not have such triggers. The level one trigger (L1) in case of CBM will be already a High Level Trigger (HLT) running on the First Level Event Selection (FLES). Therefore the data should be transmitted to FLES through the high throughput links. . . . .	23
2.12	The FLES system for the CBM experiment [43]: the data collected from the detectors will be transported to the dedicated computer farm, where it will be processed, selected and stored. . . . .	24
2.13	The L-CSC supercomputer at GSI [44], which has been ranked as the most energy-efficient supercomputer. . . . .	24
3.1	Block-diagram of the Intel Xeon E7-4860 CPU with 10 cores. . . .	27
3.2	Typical access speed to different memory layers on modern CPUs.	27
3.3	The structure of cores, cache and RAM on the lxir075 server at GSI. Each physical core contains two logical cores, 32 KB of L1 and 256 KB of L2 cache memory. Each CPU operates with 24 MB of L3 cache common for all 20 logical cores and 16 GB of RAM. .	29

---

3.4	Block-diagram of the Intel Xeon Phi architecture: all cores are connected through the bus with each other and with the DDR5 memory. . . . .	30
3.5	The next Intel Xeon Phi product — Knights Landing (KNL). It offers increased number of cores together with increased memory bandwidth, that allows to achieve more than 3 TFLOPS of the computing power [49]. . . . .	31
3.6	Structure of streaming multiprocessors of the Nvidia GTX 480 GPU based on the Fermi architecture (left) [50] and the Nvidia GTX 980 GPU based on the Maxwell architecture (right) [51].	32
3.7	Tools for code vectorization and parallelization. . . . .	34
3.8	Speedup of the arithmetic operations with Vc SSE types vs. scalar types for <code>unsigned short</code> ( <code>ushort_v</code> ), <code>unsigned integer</code> ( <code>uint_v</code> ), <code>float</code> ( <code>float_v</code> ) and <code>double</code> ( <code>double_v</code> ) types on Intel Nehalem and AMD Magny-Core architectures [53]. . . . .	37
3.9	Illustration of the OpenMP programming model — the fork-join parallelism. . . . .	38
3.10	Platform model for OpenCL: one host plus one or more compute devices each with one or more compute units each with one or more processing elements [55]. . . . .	42
4.1	Block diagram representation of the Kalman filter method [52]. . . . .	46
4.2	Distribution of the residuals and pulls for the track parameters obtained by the smoother at the first station of the STS detector of the CBM experiment. . . . .	57
4.3	Scalability of the Kalman filter based track fit on many-core CPUs using Vc for vectorization together with OpenMP for parallelization or OpenCL for both vectorization and parallelization. The stair-like structure is caused by the Intel Hyper-Threading technology. . . . .	61

- 
- 4.4 Scalability of the Kalman filter based track fit on different many-core processors: **a)** CPU (on an example of Intel Xeon E7-4860), **b)** Intel Xeon Phi, **c)** Nvidia GPU (on an example of Nvidia GTX 480) and **d)** ATI GPU (on an example of AMD Radeon HD 7970). 62
- 5.1 The block-diagram of short-lived particle reconstruction with the KF Particle package. . . . . 67
- 5.2 Determination of the point of the closest approach between a particle and a vertex assuming constant one-component magnetic field  $B_z$ : at first, a point in  $XY$  plane is found (the blue point), then a small correction on  $z$  coordinate is added (the red point). . . . . 72
- 5.3 Determination of the point of the closest approach between two particles assuming constant one-component magnetic field  $B_z$ . At first, two points in  $XY$  plane are found. If the points are well separated within  $5\sigma$  errors, then the point where two particles are closer taking into account the  $z$  coordinate is selected. If within the errors points are not separated, the middle point is selected. Selected point is shown in light-blue. Then this point is corrected on  $z$  coordinate. The corrected point is shown in red. . . . . 78
- 5.4 Determination of the distance of the closest approach between two particles at an example of  $\gamma \rightarrow e^+e^-$  conversion reconstruction. For illustration a distribution of reconstructed conversion points of true  $\gamma$  in  $ZR$  coordinates is shown (top), where  $R = \sqrt{x^2 + y^2}$ . From the comparison with the scheme of the detector geometry, drawn over the reconstructed  $\gamma$ -position (bottom), the material structure of the detector is clearly seen: the target, the vacuum chamber and five stations of STS. Due to the acceptance of the tracking system the material is not seen fully, but partially. . . . . 79
- 5.5 Comparison of linear and nonlinear mass constraints. Linear and nonlinear constraints are set on the reconstructed mass of  $\Lambda$  hyperons. The linear constraint can not guarantee the exact value of the constrained parameter: the distribution has a tail in the region of lower mass values. The nonlinear approach sets the exact value. 84

6.1	Block-diagram of the reconstruction method of the KF Particle Finder package. Cuts on $\chi_{fit}^2$ and $\chi_{topo}^2$ are given in terms of NDF.	89
6.2	Block-diagram of the KF Particle Finder package for finding, reconstruction and selection of rare short-lived particles. . . . .	94
6.3	Fit quality of the reconstructed strange particles at the example of $\Lambda$ . Distributions of residuals $\rho$ , pulls, $\chi^2$ criterion, and prob (probability for a certain $\chi^2$ and NDF) demonstrate the high quality of the reconstructed parameters. . . . .	97
6.4	Mass spectra of $K_s^0$ , $\Lambda$ , $\bar{\Lambda}$ , $\Xi^-$ , $\bar{\Xi}^+$ , and $\Omega^-$ obtained by the KF Particle Finder package together with their signal to background (S/B) ratios for 5M central AuAu UrQMD [84, 85] events at 10 AGeV using TOF PID [75]. . . . .	97
6.5	Fit quality of the reconstructed strange resonances at the example of $\Sigma^+(1385)$ . Distributions of residuals $\rho$ , pulls, $\chi^2$ criterion, and prob (probability for a certain $\chi^2$ and NDF) demonstrate the high quality of the reconstructed parameters. . . . .	99
6.6	Mass spectra of $\Sigma^+(1385)$ , $\Sigma^-(1385)$ , and $\Xi^0(1530)$ obtained by the KF Particle Finder package together with their signal to background (S/B) ratios for 5M central AuAu UrQMD [84, 85] events at 10 AGeV using TOF PID. . . . .	99
6.7	Fit quality of the reconstructed hypernuclei at the example of ${}^3_{\Lambda}H$ . Distributions of residuals $\rho$ , pulls, $\chi^2$ criterion, and prob (probability for a certain $\chi^2$ and NDF) demonstrate the high quality of the reconstructed parameters. . . . .	101
6.8	Mass spectra of $\{\Lambda n\}$ , ${}^3_{\Lambda}H$ , and ${}^4_{\Lambda}He$ hypernuclei obtained by the KF Particle Finder package together with their signal to background (S/B) ratios for 5M simulated central AuAu events at 10 AGeV using TOF PID. . . . .	101
6.9	Fit quality of the reconstructed open charm particles at the example of $\bar{D}^0$ . Distributions of residuals $\rho$ , pulls, $\chi^2$ criterion, and prob (probability for a certain $\chi^2$ and NDF) demonstrate the high quality of the reconstructed parameters. . . . .	103

- 
- 6.10 Mass spectra of  $D^0$  and  $\bar{D}^0$  mesons obtained by the KF Particle Finder package together with their signal to background (S/B) ratios for  $10^{11}$  simulated central NiNi events at 15 AGeV using TOF PID. The mixed event technique is used for the background estimation. . . . . 103
- 6.11 Fit quality of the reconstructed LMVM at the example of  $\phi \rightarrow \mu^+\mu^-$ . Distributions of residuals  $\rho$ , pulls,  $\chi^2$  criterion, and prob demonstrate the high quality of the reconstructed parameters. Irregularities in the prob distribution are due to some features in the current implementation of the hits reconstruction procedure. . . . 105
- 6.12 Mass spectra of LMVM together with their signal to background (S/B) ratios: left — in dimuon channel for  $10^{11}$  simulated central AuAu events at 25 AGeV using MuCh PID and mixed event technique, right — 5M simulated central AuAu UrQMD events at 10 AGeV using TOF PID. . . . . 105
- 6.13 Fit quality of the reconstructed charmonium at the example of  $J/\psi \rightarrow \mu^+\mu^-$ . Distributions of residuals  $\rho$ , pulls,  $\chi^2$  criterion, and prob demonstrate the high quality of the reconstructed parameters. Irregularities in the prob distribution are due to some features in the current implementation of the hits reconstruction procedure. . 107
- 6.14 Mass spectra of  $J/\psi$  in dimuon channel reconstructed by the KF Particle Finder together with the signal to background (S/B) ratio for  $10^{11}$  simulated central AuAu events at 25 AGeV using MuCh PID. For the background simulation the mixed event technique is used. . . . . 107
- 6.15 Fit quality of the  $\pi^0$  particles reconstructed by the full  $\gamma$  conversion method. Distributions of residuals  $\rho$ , pulls,  $\chi^2$  criterion, and prob (probability for a certain  $\chi^2$  and NDF) demonstrate the high quality of the reconstructed parameters. . . . . 109
- 6.16 Mass spectra of  $\gamma$  conversion and  $\pi^0$  (from  $\gamma$ -conversion) reconstructed by the KF Particle Finder together with the signal to background (S/B) ratio for 5M simulated central AuAu UrQMD events at 25 AGeV using RICH, TRD and TOF PID. . . . . 109

---

6.17 Scalability of the KF Particle Finder package on the many-core server lxir075.gsi.de equipped with four Intel E7-4860 (2.27 GHz) CPUs. . . . .	112
6.18 Block-diagram of the FLES package [99]. . . . .	112
6.19 Scalability of the FLES package on different Intel and AMD many-core servers [99]. The FLES package shows strong linear scalability on all architectures. . . . .	113
6.20 Scalability of the FLES package on FAIR-Russia HPC Cluster (ITEP, Moscow) [100]. The FLES package scales up to 3200 cores. . . . .	113
6.21 Block-diagram of the KF Particle Finder package with the mass distribution of the corresponding decays. . . . .	115
A.1 A straight track on a plane. IP is a production point of the track. The line, that starts at the point IP, shows a trajectory of the track. Vertical lines represent a position of detector stations placed at equal distances L from each other. . . . .	120
1 Blockdiagramm des KF Particle Finder Pakets zur Rekonstruktion und Auswahl der seltenen kurzlebigen Teilchen. . . . .	158

# List of Tables

1.1	A list of already running and future (marked with *) accelerators for studying heavy ion collisions. For accelerators with the fixed target experiments the energy is given in the laboratory system, for collider type — center-of-mass collision energy is given. . . .	7
1.2	Multiplicities (by HSD model [25]) and branching ratios [5] of the CBM observables on example of $\Omega^-$ and $\bar{\Omega}^+$ for multi-strange hyperons, $D^0$ and $\bar{D}^0$ for open charm, $\rho$ for low mass vector mesons and $J/\psi$ in a central Au-Au collision at 25 AGeV energy. On top of the low multiplicity and branching ratio the efficiency of reconstruction for such signals is typically on the few percent level.	9
2.1	Observables and detectors for their registration. Detectors marked as x are those, which are necessary for particle registration. Detectors marked as (x) can be used to suppress background. . . .	14
4.1	The standard deviations of the residuals and pulls distributions for the track parameters obtained at the point of each measurement of the track by the Kalman filter based smoother for the tracking system of the CBM experiment. . . . .	57
4.2	Percentage of rejected hits depending on the distance from the shifted hit to its Monte-Carlo position on the 4 <sup>th</sup> STS station. . .	59
6.1	Parameters of strange particles and antiparticles included in the KF Particle Finder reconstruction scheme: the implemented channel, its branching ratio, and the lifetime ( $c\tau$ ) of a particle [5]. . . .	95

6.2	Efficiencies of the strange particles reconstruction, widths of the signal peaks, and S/B ratios for 5M central AuAu UrQMD events at 10 AGeV using TOF PID. . . . .	96
6.3	Parameters of strange resonances and their antiparticles included in the KF Particle Finder reconstruction scheme: the implemented channel, its branching ratio, and the full width ( $\Gamma$ ) of the resonance [5]. The most abundant resonances decaying by the selected channels are shown. . . . .	98
6.4	Efficiencies of the strange resonances reconstruction, widths of the signal peaks, and S/B ratios for 5M central AuAu UrQMD events at 10 AGeV using TOF PID. . . . .	99
6.5	Parameters of hypernuclei and heavy multi-strange objects included in the KF Particle Finder reconstruction scheme: the implemented channel, its branching ratio, and the lifetime ( $c\tau$ ) of a particle [88, 89, 90, 91, 92, 93, 94]. . . . .	100
6.6	Efficiencies of the hypernuclei reconstruction, widths of the signal peaks, and S/B ratios for 5M simulated central AuAu events at 10 AGeV using TOF PID. . . . .	101
6.7	Parameters of open charm particles, antiparticles, and resonances included in the KF Particle Finder reconstruction scheme: the implemented channel, its branching ratio, and the lifetime ( $c\tau$ ) or the full width ( $\Gamma$ ) of a particle [5]. . . . .	102
6.8	Efficiency of $D^0$ and $\bar{D}^0$ reconstruction, widths of the signal peaks, and S/B for $10^{11}$ simulated central NiNi events at 15 AGeV using TOF PID. . . . .	103
6.9	Parameters of the low mass vector mesons included in the KF Particle Finder reconstruction scheme: the implemented channel, its branching ratio, the full width ( $\Gamma$ ), and the lifetime ( $c\tau$ ) of a particle [5]. . . . .	104
6.10	Efficiency of the LMVM reconstruction, width of the signal peak, and S/B. . . . .	105

6.11	Parameters of the $J/\psi$ decays included in the KF Particle Finder reconstruction scheme: the implemented channel, its branching ratio, and the full width ( $\Gamma$ ) of a particle [5]. . . . .	106
6.12	Efficiency of the $J/\psi$ reconstruction, the width of the signal peak, and S/B for $10^{11}$ simulated central AuAu events at 25 AGeV using MuCh PID. . . . .	107
6.13	Parameters of $\pi^0$ and $\eta$ decays included in the KF Particle Finder reconstruction scheme: the implemented channel, its branching ratio, and the lifetime ( $c\tau$ ) of a particle [5]. . . . .	108
6.14	Efficiency of $\gamma$ conversion and $\pi^0$ reconstruction, widths of the signal peaks, and S/B ratios for 5M simulated central AuAu UrQMD events at 25 AGeV using RICH, TRD and TOF PID. . . . .	109

# Bibliography

- [1] M. Gell-Mann, “The Eightfold Way: A Theory of strong interaction symmetry,” CTSL-20, TID-12608.
- [2] M. Gell-Mann, “A Schematic Model of Baryons and Mesons,” *Phys. Lett.* **8** (1964) 214.
- [3] G. Zweig, “An SU(3) model for strong interaction symmetry and its breaking. Version 1,” CERN-TH-401.
- [4] D. H. Perkins, “Introduction to high energy physics; 4th. ed,” Cambridge Univ. Press (2000) 426 p.
- [5] K. A. Olive *et al.* [Particle Data Group Collaboration], “Review of Particle Physics,” *Chin. Phys. C* **38** (2014) 090001.
- [6] M. Gell-Mann and Y. Ne’eman, “The Eightfold Way: A Review - With Collection of Reprints, *Frontiers in Physics*,” ed. D. Pines, W. A. Benjamin, Inc. New York - Amsterdam (1964) 168.
- [7] W. Greiner, S. Schramm, E. Stein, “Quantum chromodynamics; 2nd ed,” Springer Science and Business Media (2002) 551 p.
- [8] K. G. Wilson, “Confinement of Quarks,” *Phys. Rev. D* **10** (1974) 2445.
- [9] K. Fukushima and T. Hatsuda, “The phase diagram of dense QCD,” *Rept. Prog. Phys.* **74** (2011) 014001
- [10] A. Andronic *et al.*, “Hadron Production in Ultra-relativistic Nuclear Collisions: Quarkyonic Matter and a Triple Point in the Phase Diagram of QCD,” *Nucl. Phys. A* **837** (2010) 65

- 
- [11] Y. Aoki, G. Endrodi, Z. Fodor, S. D. Katz and K. K. Szabo, “The Order of the quantum chromodynamics transition predicted by the standard model of particle physics,” *Nature* **443** (2006) 675
- [12] F. Karsch, “Moments of charge fluctuations, pseudo-critical temperatures and freeze-out in heavy ion collisions,” *J. Phys. G* **38** (2011) 124098
- [13] Z. Fodor and S. D. Katz, “The Phase diagram of quantum chromodynamics,” arXiv:0908.3341 [hep-ph].
- [14] S. Borsanyi *et al.* [Wuppertal-Budapest Collaboration], “Is there still any  $T_c$  mystery in lattice QCD? Results with physical masses in the continuum limit III,” *JHEP* **1009** (2010) 073
- [15] A. Bazavov *et al.*, “The chiral and deconfinement aspects of the QCD transition,” *Phys. Rev. D* **85** (2012) 054503
- [16] R. A. Beth and C. Lasky, “The Brookhaven Alternating Gradient Synchrotron,” *Science* 5, Vol. 128, no. 3336, December 1958, Pages 1393-1401.
- [17] B. Franczak, K. Blasche and K. H. Reich, “A Variable Transition Energy Lattice for SIS 12/18,” *IEEE Trans. Nucl. Sci.* **30** (1983) 2120.
- [18] M. Harrison, T. Ludlam, S. Ozaki, RHIC project overview, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Volume 499, Issues 2-3, 1 March 2003, Pages 235-244.
- [19] G. Brianti, “Large Hadron Collider in the LEP Tunnel: A Feasibility Study of Possible Options,” *Conf. Proc. C* **8405141** (1984) 207.
- [20] “FAIR Baseline Technical Report,” GSI, Darmstadt (2006) 92 p.  
[www.fair-center.eu/fileadmin/fair/publications\\_FAIR/FAIR\\_BTR.1.pdf](http://www.fair-center.eu/fileadmin/fair/publications_FAIR/FAIR_BTR.1.pdf)
- [21] A. Kovalenko *et al.*, “NICA — Nuclotron-based Ion Collider Facility — a base for fundamental research and innovation development.”  
<http://nica.jinr.ru/files/Booklet-NICA.pdf>

- [22] B. Friman, C. Hohne, J. Knoll, S. Leupold, J. Randrup, R. Rapp and P. Senger, “The CBM physics book: Compressed baryonic matter in laboratory experiments,” *Lect. Notes Phys.* **814** (2011) pp. 980.
- [23] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Trans. ASME, Series D, J. Basic Eng.*, **82** (1960) 35.
- [24] D. Simon, “Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches,” John Wiley & Sons (2006) 526 p.
- [25] W. Ehehalt and W. Cassing, “Relativistic transport approach for nucleus nucleus collisions from SIS to SPS energies,” *Nucl. Phys. A* **602** (1996) 449.
- [26] S. Gorbunov and I. Kisel, “Reconstruction of decayed particles based on the Kalman filter,” CBM-SOFT-note-2007-003, 7 May 2007.
- [27] S. Gorbunov, “On-line reconstruction algorithms for the CBM and ALICE experiments,” Dissertation thesis, Goethe University of Frankfurt, 2012.
- [28] M. M. Aggarwal *et al.* [STAR Collaboration], “An Experimental Exploration of the QCD Phase Diagram: The Search for the Critical Point and the Onset of De-confinement,” arXiv:1007.2613 [nucl-ex].
- [29] A. Aduszkiewicz [NA61 Collaboration], “NA61/SHINE at the CERN SPS: Plans, status and first results,” *Acta Phys. Polon. B* **43**, 635 (2012), arXiv:1201.5879 [nucl-ex].
- [30] P. Senger and V. Friese, “The CBM Collaboration: Nuclear Matter Physics at SIS-100,” GSI, Darmstadt (2012) 18 p.
- [31] “CBM Progress Report 2014,” GSI, Darmstadt (2015) 159 p.
- [32] M. Koziel, “MVD Status: Integration,” 25<sup>th</sup> CBM Collaboration Meeting, Darmstadt, 20-24 April 2015.  
<https://indico.gsi.de/getFile.py/access?contribId=17&sessionId=9&resId=0&materialId=slides&confId=2960>
- [33] J. Heuser *et al.*, “Technical Design Report for the CBM Silicon Tracking System (STS),” GSI, Darmstadt (2013) 167 p.

- [34] A. Malakhov *et al.*, “Technical Design Report for the CBM Superconducting Dipole Magnet,” GSI, Darmstadt (2013) 80 p.
- [35] C. Höhne *et al.*, “Technical Design Report for the CBM Ring Imagine Cherenkov (RICH),” GSI, Darmstadt (2013) 201 p.
- [36] S. Chattopadhyay *et al.*, “Technical Design Report for the CBM Muon Chambers (MuCh),” GSI, Darmstadt (2014) 192 p.
- [37] “Compressed Baryonic Matter Experiment. Technical Status Report.” GSI, Darmstadt (2005) 406 p.
- [38] N. Herrmann *et al.*, “Technical Design Report for the CBM Time-of-Flight System (TOF),” GSI, Darmstadt (2014) 182 p.
- [39] I. Vassiliev and **M. Zyzak**, “Hypermatter and Hypernuclei search at SIS100 using ToF particle ID,” 25<sup>th</sup> CBM Collaboration Meeting, Darmstadt, 20-24 April 2015.  
<https://indico.gsi.de/getFile.py/access?contribId=27&sessionId=16&resId=1&materialId=slides&confId=2960>
- [40] I. Korolko, “CBM Calorimeter for SIS100 (TDR status),” 25<sup>th</sup> CBM Collaboration Meeting, Darmstadt, 20-24 April 2015.  
<https://indico.gsi.de/getFile.py/access?contribId=138&sessionId=29&resId=1&materialId=slides&confId=2960>
- [41] F. Guber *et al.*, “Technical Design Report for the CBM Projectile spectator detector (PSD),” GSI, Darmstadt (2014) 78 p.
- [42] D. Hutter, “The CBM FLES Input Interface,” DPG Spring Meeting 2015, Heidelberg, 23-27 March 2015.  
<https://www-alt.gsi.de/documents/DOC-2015-Apr-12-1.pdf>
- [43] J. de Cuveland, “The CBM First-level Event Selector (FLES): Overview and Status,” DPG Spring Meeting 2015, Heidelberg, 23-27 March 2015.  
<https://www-alt.gsi.de/documents/DOC-2015-Apr-8-1.pdf>

- [44] L-CSC Supercomputer.  
<http://www.fair-center.de/de/news-events/news-view/article/kooperation-von-fias-goethe-uni-und-gsi-auf-platz-1-der-green500.html>
- [45] Intel Corporation, “Intel<sup>®</sup> 64 and IA-32 Architectures Optimization Reference Manual,” (2014).
- [46] Intel Corporation, “Intel<sup>®</sup> Xeon<sup>™</sup> Processor Family for Servers with Hyper-Threading Technology,” Intel White Paper.
- [47] Intel Corporation: Intel QPI.  
<http://www.intel.com/content/www/us/en/io/quickpath-technology/quickpath-technology-general.html>
- [48] R. Rahman, “Intel<sup>®</sup> Xeon Phi<sup>™</sup> Coprocessor Architecture and Tools The Guide for Application Developers,” (2013) 232 p.
- [49] J. Reinders, “Knights Corner: Your Path to Knights Landing,” (2014).  
<https://software.intel.com/sites/default/files/managed/e9/b5/Knights-Corner-is-your-path-to-Knights-Landing.pdf>
- [50] Nvidia Corporation: Nvidia GTX 480.  
<http://de. GeForce.com/hardware/desktop-gpus/geforce-gtx-480/architecture>
- [51] M. Harris, “Maxwell: The Most Advanced CUDA GPU Ever Made.”  
<http://devblogs.nvidia.com/paralleforall/maxwell-most-advanced-cuda-gpu-ever-made/>
- [52] S. Gorbunov, U. Kebschull, I. Kisel, V. Lindenstruth and W.F.J. Müller, “Fast SIMDized Kalman filter based track fit,” *Comp. Phys. Comm.* **178** (2008) 374.
- [53] M. Kretz and V. Lindenstruth, “Vc: A C++ library for explicit vectorization,” *Softw: Pract. Exper.* **42** (2012) 1409.
- [54] M. Kretz: Vector Classes.  
<http://code.compeng.uni-frankfurt.de/projects/vc>

- 
- [55] Khronos Group: The OpenCL Specification, Version: 1.2 and Version: 2.1.  
<http://www.khronos.org>
- [56] Intel Corporation: Intel Array Building Blocks.  
<http://software.intel.com/en-us/articles/intel-array-building-blocks>
- [57] OpenMP Architecture Review Board: “The OpenMP<sup>®</sup> API specification for parallel programming.”  
<http://openmp.org>
- [58] Intel Corporation, “Intel Threading Building Blocks Reference Manual.”  
<http://threadingbuildingblocks.org>
- [59] POSIX Threads Programming.  
<https://computing.llnl.gov/tutorials/threads>
- [60] I. Ollmann, “AltiVec (a.k.a Velocity Engine) Tutorial”.  
<http://web-docs.gsi.de/~ikisel/reco/Systems/Altivec.pdf>
- [61] I. Kisel, I. Kulakov, H. Pabst, **M. Zyzak**, “Towards Parallel Track Reconstruction with Intel ArBB,” CBM Progress Report 2010, Darmstadt 2011, p. 70.
- [62] R. Frühwirth *et al.*, “Data Analysis Techniques for High-Energy Physics. Second Edition,” Cambridge Univ. Press (2000) 421 p.
- [63] R. K. Kutschke, “Example of a Kalman filter on a straight line,” Fermi National Accelerator Laboratory (1998).
- [64] W. H. Press *et al.*, “Numerical Recipes in C, Second Edition,” Cambridge Univ. Press (1992) 965 p.  
<http://www.nrbook.com/a/bookpdf.php>
- [65] S. Gorbunov and I. Kisel, “Analytic formula for track extrapolation in non-homogeneous magnetic field,” Nucl. Instrum. Meth. A **559** (2006) 148.
- [66] G. R. Lynch and O. I. Dahl, “Approximations to multiple Coulomb scattering,” Nucl. Instrum. Meth. B **58** (1991) 6.

- 
- [67] E. J. Wolin and L. L. Ho, "Covariance matrices for track fitting with the Kalman filter," Nucl. Instrum. Meth. A **329**, 493 (1993).
- [68] I. Kisel, "Event reconstruction in the CBM experiment," Nucl. Instrum. Meth. A **566** (2006) 85.
- [69] R. Frühwirth and A. Strandlie, "Track fitting with ambiguities and noise: A study of elastic tracking and nonlinear filters," Comput. Phys. Commun. **120** (1999) 2-3, 197.
- [70] R. Frühwirth, "Application of Kalman filtering to track and vertex fitting," Nucl. Instrum. Meth. A **262** (1987) 444.
- [71] W. Waltenberger, "RAVE: A detector-independent toolkit to reconstruct vertices," IEEE Trans. Nucl. Sci. **58** (2011) 434.
- [72] S. Gorbunov and I. Kisel, "Secondary vertex fit based on the Kalman filter," CBM-SOFT-note-2006-002, 14 September 2006.
- [73] Chun Yang, E. Blasch, "Kalman Filtering with Nonlinear State Constraints," IEEE Transactions on Aerospace and Electronic Systems, vol.45, no.1 (2009) 70.
- [74] I. Vassiliev, I. Kisel, I. Kulakov, A. Lebedev, and **M. Zyzak**, "Reconstruction of multi-strange baryons in the CBM experiment," CBM Progress Report 2012, Darmstadt 2013, p. 100.
- [75] I. Vassiliev, I. Kisel, I. Kulakov, and **M. Zyzak**, "Reconstruction of anti-hyperons in the CBM experiment," CBM Progress Report 2012, Darmstadt 2013, p. 101.
- [76] I. Vassiliev, I. Kisel, I. Kulakov, and **M. Zyzak**, "Multi-strange hyperon reconstruction with the CBM experiment," CBM Progress Report 2013, Darmstadt 2014, p. 106.
- [77] V. Vovchenko, I. Vassiliev, I. Kisel, and **M. Zyzak**, " $\phi$  meson production in Au+Au collisions and its reconstruction feasibility in the CBM experiment," CBM Progress Report 2014, Darmstadt 2015, p. 124.

- [78] O. Derenovskaya and I. Vassiliev, “ $J/\psi$  reconstruction in Au + Au collisions at 10A GeV,” CBM Progress Report 2012, Darmstadt 2013, p. 107.
- [79] **M. Zyzak**, I. Kisel, and I. Vassiliev, “ $\pi^0$  reconstruction through a  $\gamma$ -conversion method with KF Particle Finder in the CBM experiment,” CBM Progress Report 2014, Darmstadt 2015, p. 132
- [80] I. Vassiliev, I. Kisel, I. Kulakov, and **M. Zyzak**, “ $\Sigma^*$  decay reconstruction in the CBM experiment,” CBM Progress Report 2012, Darmstadt 2013, p. 102.
- [81] I. Vassiliev, I. Kisel, and V. Vovchenko, “ $D^{*+}$  decay reconstruction in C+C collisions at 25A GeV in the CBM experiment,” CBM Progress Report 2011, Darmstadt 2012, p. 83
- [82] J. Rafelski, B. Muller, “Strangeness production in the quark-gluon plasma,” Phys. Rev. Lett. **48** (1982) 1066.
- [83] F. Antinori *et al.* [WA97 Collaboration], “Production of strange and multistrange hadrons in nucleus nucleus collisions at the SPS,” Nucl. Phys. A **661** (1999) 130.
- [84] S. A. Bass, *et al.*, “Microscopic Models for Ultrarelativistic Heavy Ion Collisions,” Prog. Part. Nucl. Phys. **41** (1998) 225-370
- [85] M. Bleicher, *et al.*, “Relativistic Hadron-Hadron Collisions in the Ultra-Relativistic Quantum Molecular Dynamics Model,” J. Phys. G: Nucl. Part. Phys. **25** (1999) 1859-1896
- [86] J. Steinheimer *et al.*, “Strangeness fluctuations and MEMO production at FAIR,” Phys. Lett. B **676** (2009) 126
- [87] J. Steinheimer *et al.*, “Hypernuclei, dibaryon and antinuclei production in high energy heavy ion collisions: Thermal production versus Coalescence,” Phys. Lett. B **714** (2012) 85
- [88] R.J. Prem, P.H. Steinberg, “Lifetimes of Hypernuclei,  ${}_{\Lambda}H^3$ ,  ${}_{\Lambda}H^4$ ,  ${}_{\Lambda}He^5$ ,” Phys. Rev. **136** (1964) B1803

- [89] B. I. Abelev [STAR Collaboration], “Observation of an Antimatter Hypernucleus,” *Science* **328** (2010) 58
- [90] C. Rappold *et al.* [HypHI Collaboration], “Search for evidence of  ${}^3_{\Lambda}n$  by observing  $d + \pi^{-}$  and  $t + \pi^{-}$  final states in the reaction of  ${}^6\text{Li} + {}^{12}\text{C}$  at 2A GeV,” *Phys. Rev. C* **88** (2013) 4, 041001.
- [91] C. Rappold *et al.*, “On the measured lifetime of light hypernuclei  ${}^3_{\Lambda}H$  and  ${}^4_{\Lambda}H$ ,” *Phys. Lett. B* **728** (2014) 543.
- [92] C. Rappold *et al.*, “Hypernuclear production cross section in the reaction of  ${}^6\text{Li} + {}^{12}\text{C}$  at 2A GeV,” *Phys. Lett. B* **747** (2015) 129.
- [93] I. Kumagai-Fuse, S. Okabe and Y. Akaishi, “Evidence of central repulsion potential from pionic decay of s shell Lambda hypernuclei,” *Phys. Lett. B* **345** (1995) 386.
- [94] J. J. Szymanski *et al.*, “Nonleptonic weak decay of  ${}^5_{\Lambda}\text{He}$  and  ${}^{12}_{\Lambda}\text{C}$ ,” *Phys. Rev. C* **43** (1991) 849.
- [95] T. Matsui and H. Satz, “ $J/\psi$  Suppression by Quark-Gluon Plasma Formation,” *Phys. Lett. B* **178** (1986) 416.
- [96] M. C. Abreu *et al.* [NA50 Collaboration], “Anomalous J / psi suppression in Pb - Pb interactions at 158 GeV/c per nucleon,” *Phys. Lett. B* **410** (1997) 337.
- [97] A. Adare *et al.* [PHENIX Collaboration], “ $J/\psi$  suppression at forward rapidity in Au+Au collisions at  $\sqrt{s_{NN}} = 200$  GeV,” *Phys. Rev. C* **84** (2011) 054912
- [98] B. Abelev *et al.* [ALICE Collaboration], “ $J/\psi$  suppression at forward rapidity in Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV,” *Phys. Rev. Lett.* **109** (2012) 072301
- [99] I. Kisel, I. Kulakov, M. Zyzak, “Standalone First Level Event Selection Package for the CBM Experiment,” *IEEE Transactions on Nuclear Science*, vol. 60, No. 5, October 2013, p. 3703.

- [100] V. Akishina, I. Kisel, I. Kulakov and **M. Zyzak**, “FLES — first level event selection package for the CBM experiment,” Proceedings of the GPU Computing in High-Energy Physics Conference 2014, Pisa, Italy, September 10–12, 2014, DOI: 10.3204/DESY-PROC-2014-05 (2014) 23–29.

# Zusammenfassung

## Motivation

Moderne Experimente der Hochenergie-Schwerionenphysik werden auf immer höhere Wechselwirkungsraten und Teilchenmultiplizitäten optimiert. Dadurch eröffnet sich die Möglichkeit, neue Teilchen zu entdecken oder extrem seltene Proben zu untersuchen, welche besonders empfindlich sind auf die Eigenschaften heißer und komprimierter Kernmaterie. Messungen seltener Teilchen mit hoher Statistik erfordern die Verarbeitung sehr hoher Raten und großer Datenmengen. Allerdings erlauben die derzeit verfügbaren Speichermedien nicht, die gesamte Menge der gemessenen Daten aufzuzeichnen. Daher müssen Ereignisse, die potenziell interessante Information enthalten, online selektiert und gespeichert werden.

Das künftige “Compressed Baryonic Matter” (CBM) Experiment an der internationalen “Facility für Antiproton and Ion Research” (FAIR) hat zum Ziel, seltene Teilchen wie vielfach-seltene (Anti-)Hyperonen oder charmante Mesonen in Schwerionenstößen zu messen. Aufgrund der komplizierten Zerfallstopologie dieser Teilchen und der hohen Multiplizität der Reaktionen stehen keine einfachen Auswahlkriterien (“Trigger”) für interessante Ereignisse zur Verfügung. Um die Datenmengen zu reduzieren, müssen alle gemessenen Reaktionen online rekonstruiert werden, um potenziell interessante Ereignisse selektieren zu können. Dies erfolgt durch schnelle Algorithmen die auf einem Hochleistungs-Compute-Cluster (HPC) ausgeführt werden.

Moderne HPC Server enthalten Datenverarbeitungseinheiten mit many-core-Architekturen. Darüber hinaus zielt die aktuelle Entwicklung darauf ab, die Anzahl der Cores pro Verarbeitungseinheit weiter zu erhöhen. Um diese modernen Rechnerarchitekturen optimal ausnützen zu können, müssen die Rekon-

struktionsalgorithmen für die many-core-Prozessoren optimiert werden, sodass ihre Geschwindigkeit bzw. die Rate der verarbeiteten Ereignisse mit der Anzahl der Cores skaliert.

## Rekonstruktion der Spuren geladener Teilchen

Die Rekonstruktion der Reaktionen in Experimenten der Hochenergiephysik erfordert sowohl eine hohe Genauigkeit als auch hohe Geschwindigkeit der Algorithmen. Eine der Hauptaufgaben ist die Rekonstruktion der Teilchentrajektorien (Spuren). Die Genauigkeit der Track-Parameter definiert die Qualität der rekonstruierten seltenen Signale und damit die Wahrscheinlichkeit, diese Signale nachzuweisen.

Die Kalman-Filter (KF) Methode ist eine rekursive Methode zur Analyse von linearen, diskreten dynamischen Systemen, die durch einen Vektor von Parametern, den so genannten Zustandsvektor, beschrieben werden können. Die Methode ermöglicht es, eine optimale Abschätzung der Parameter der Teilchenspuren zu finden und die höchstmögliche Genauigkeit zu erzielen.

Das Kalman-Filter wird in vielen Experimenten der Hochenergiephysik zur Ereignisrekonstruktion verwendet aufgrund folgender Eigenschaften:

- optimale Abschätzung der Parameter;
- Lokalität in Bezug auf Messungen: die Messungen werden unabhängig voneinander eine nach der anderen addiert;
- Nichtlinearitäten der Teilchenbahnen im Fall eines inhomogenen Magnetfelds können in geeigneter Weise berücksichtigt werden: ein Teilchen wird zwischen den Messungen unter der Berücksichtigung aller nichtlinearen Effekte transportiert, und seine Flugbahn wird nur in der Nähe jeder Messung linearisiert;
- Das Material eines Detektors, das die Teilchen durchgedringen, wird in der Nähe jeder Messung berücksichtigt, was das Verfahren vereinfacht und die Berücksichtigung aller Materialeffekte mit einem minimalen Satz von Annahmen ermöglicht;
- Arbeit mit Matrizen von kleiner Größe (bis zu  $5 \times 5$ );

- Die Anzahl der Rechenoperationen ist proportional zur Anzahl der Messungen.

Die Zahl der existierenden Methoden zur Spurenanpassung auf Basis des Kalman-Filters für das CBM Experiment wurde erweitert: durch den “KF Glätter” zur Abschätzung der Track-Parameter an der Position jedes Hits, und durch den Deterministischen Annealing-Filter (DAF) zur Reduzierung der Auswirkung der Rausch-Signale auf die endgültige Abschätzung der Spurparameter.

Die implementierten Methoden erzielen eine hohe Rekonstruktionsqualität: Residuen (die Differenz zwischen dem rekonstruierten und wahren Wert) und Pulls (Residuen normiert auf die Kovarianzmatrix) sind frei von systematischen Fehlern und haben die erwarteten Breiten.

Um das Leistungspotenzial moderner Rechnerarchitekturen zu nutzen, verwendet die Rekonstruktion der Parameter geladener Teilchen mithilfe des Kalman-Filters Verfahren, Frameworks und Bibliotheken wie Header-Dateien mit SIMD-Operationen, Vc, OpenMP, Intel TBB, Pthreads, OpenCL und Intel ArBB. Die vektorisierte und parallelisierte KF Spurenanpassung weist eine hohe Skalierbarkeit auf many-core-CPUs und GPUs auf.

## Rekonstruktion kurzlebiger Teilchen mit dem KF Particle Paket

Teilchen, die in einer Kollision erzeugt werden, können in zwei Gruppen unterteilt werden: langlebige Teilchen und kurzlebige Teilchen. Langlebige Teilchen haben eine Lebensdauer, die ausreicht um das vollständige Detektorsystem des Experiments zu durchfliegen und mithilfe verschiedener Detektoren identifiziert werden zu können: Elektronen und Positronen  $e^\pm$ , Protonen  $p^\pm$ , Myonen  $\mu^\pm$ ,  $\pi^\pm$  und  $K^\pm$  Mesonen. Kurzlebige Teilchen hingegen zerfallen bereits vor oder innerhalb des Tracking-Systems und können nur indirekt nachgewiesen werden. Zu den kurzlebigen Teilchen gehören z.B. seltsame Hyperonen ( $\Lambda$ ,  $\Xi$ ,  $\Omega$ ), leichte Vektormesonen ( $\rho$ ,  $\omega$ ,  $\phi$ ), Charm-Teilchen (D-Mesonen,  $J/\psi$ ) und Hyperkerne ( ${}^3_\Lambda\text{H}$ ,  ${}^4_\Lambda\text{He}$ ). Diese Teilchen können nur indirekt über ihre Zerfallsprodukte rekonstruiert werden, sie sind aber von besonderem Interesse für das CBM Physikpro-

gramm.

Das KF Partice Paket wird im CBM Experiment für die vollständige Rekonstruktion kurzlebiger Teilchen verwendet, einschließlich der Bestimmung von Impuls, Energie, Masse, Lebensdauer, Zerfallslänge, Rapidität usw. Das KF Particle Paket basiert auf der Kalman-Filter-Methode, es ist unabhängig von der Geometrie, beschreibt Tochter- und Mutterteilchen mit dem gleichen Parametersatz und behandelt sie auf genau gleiche Weise.

Das KF Particle Paket löst nichtlineare Probleme, wie die Extrapolation geladener Teilchen in einem nicht-homogenen Magnetfeld, oder die Wahl der Masseneinschränkung. Solche Probleme werden in der Regel gelöst mit einer Reihe von Näherungswerten oder durch numerische Methoden, die entweder Rekonstruktionsqualität oder die Geschwindigkeit des Algorithmus beeinflussen können.

Für die Online-Rekonstruktion und Identifizierung der kurzlebigen Teilchen werden im CBM Experiment eine möglichst hohe Effizienz und Geschwindigkeit verlangt. Deshalb wurden neue Verfahren zur Suche nach dem Punkt der größten Annäherung und der Masseneinschränkung entwickelt. Durch die eingeführten Verbesserungen ist die Funktionalität des KF Particle-Pakets perfekt geeignet zur Rekonstruktion der Zerfallskaskaden instabiler Teilchen. Somit war es möglich, ein KF Particle Paket zur vollständigen Rekonstruktion der Ereignisse zu entwickeln.

## **KF Particle Finder — eine Plattform zur Online-Selektion und Physikanalyse**

Viele der kurzlebigen Teilchen, die nur durch ihre Zerfallsprodukte rekonstruiert werden können, sind als diagnostische Sonden der Eigenschaften komprimierter Kernmaterie von besonderem Interesse. Manche dieser Teilchen haben entweder eine sehr kleine Produktionswahrscheinlichkeit oder ein sehr kleines Verzweigungsverhältnis, die Ausbeute liegt oft in einer Größenordnung von  $10^{-5}$ – $10^{-7}$  Teilchen pro Kollision. Aus diesem Grund kann ein statistisch signifikantes Ergebnis nur durch sehr hohe Kollisionsraten erreicht werden. Dies wirft das Problem der Datenverarbeitung und Speicherung auf. Der Datenfluss ist üblicherweise

so groß, dass nicht alle Daten vollständig auf Band gespeichert werden können. Nur die Ereignisse, die möglicherweise interessante Teilchen enthalten, werden gespeichert.

Im Falle des CBM Experiments gibt aufgrund der komplizierten Zerfallstopologie vieler kurzlebiger Teilchen und des hohen Untergrunds an anderen im Stoß erzeugten Teilchen keine einfachen Kriterien für die Auswahl interessanter Ereignisse. Daher erfordert das Experiment die vollständige Rekonstruktion der Ereignisse einschließlich der Rekonstruktion der kurzlebigen Teilchen in Echtzeit, damit interessante Kandidaten selektiert und gespeichert werden können. Dies erfordert Rekonstruktionsprogramme, die mit extrem hoher Geschwindigkeit laufen, und hoch-effiziente Algorithmen, um die interessanten Ereignisse nicht zu verlieren.

Das KF Particle Finder Paket zur Rekonstruktion kurzlebiger Teilchen erfordert als Startinformation die rekonstruierten Trajektorien geladener Teilchen, zusammen mit ihrer PID-Hypothese (PID: Particle Identification). Aufgrund PID-Hypothese und der Zerfallstopologie werden die kurzlebigen Mutterteilchen aus den Tochterteilchen mit Hilfe des KF Particle Pakets rekonstruiert. Um den Untergrund zu unterdrücken, müssen die Kandidaten eine Reihe von Auswahlkriterien erfüllen. Die Auswahlkriterien für verschiedene Gruppen von Zerfällen sind in Bezug auf die Teilchenausbeuten, die Verzweigungsverhältnisse und die Zerfallseigenschaften optimiert.

Alle kurzlebigen Teilchen, die das KF Particle Finder Paket findet, weisen bezüglich ihrer rekonstruierten Parameter eine hohe Qualität auf: alle Verteilungen der Residuen und Pulls sind frei von systematischen Fehlern; die Breite der Residuen ist in guter Übereinstimmung mit der Detektorauflösung, die Breite der Pulls ist nahe dem erwarteten Wert von eins; die  $\chi^2$  und die entsprechenden Wahrscheinlichkeitsverteilungen haben die erwartete Form, die Wahrscheinlichkeitsverteilung entspricht der flachen gleichmäßigen Verteilung zwischen 0 und 1. Dadurch werden große Effizienzen und hohe Signal-zu-Untergrund-Verhältnisse erreicht.

Derzeit umfasst das KF Particle Finder Paket mehr als 70 Zerfälle (siehe Abb. 1). Das Paket erlaubt, alle wichtigen Observablen eines Schwerionenstoßes zu rekonstruieren, einschließlich der seltsamen Teilchen, der seltsamen Reso-

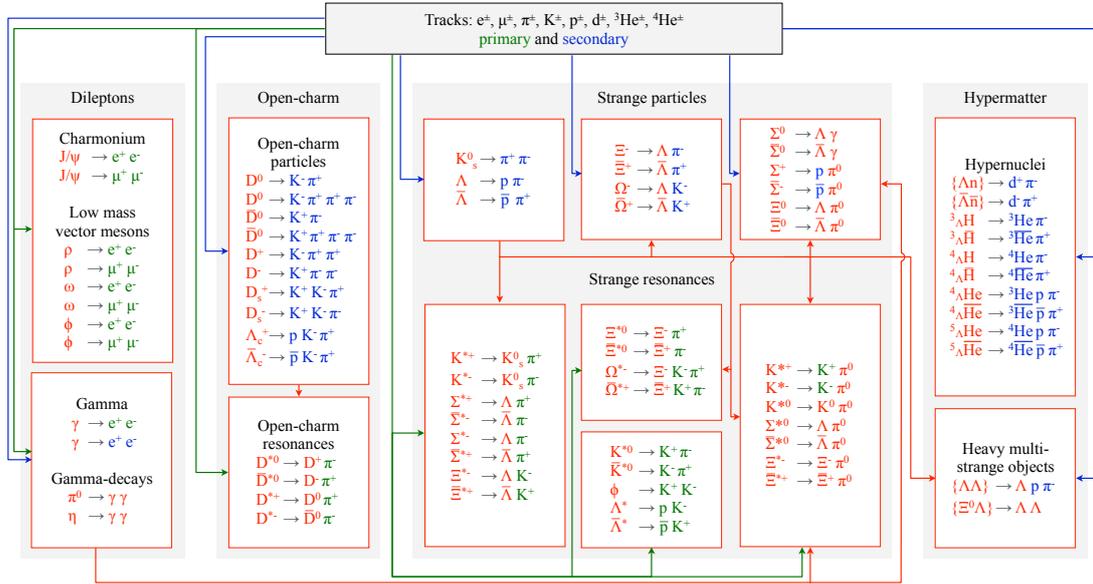


Abbildung 1: Blockdiagramm des KF Particle Finder Pakets zur Rekonstruktion und Auswahl der seltenen kurzlebigen Teilchen.

nanzen, der Hyperkerne, der leichten Vektormesonen, Charmonium und Open-Charm-Teilchen.

Das KF Particle Finder Paket wurde auf Geschwindigkeit optimiert. Es wurde in einfacher Präzision, die eine hohe Qualität der Rekonstruktion garantiert, implementiert und stabilisiert. Das Paket wurde vektorisiert und, um die Möglichkeiten der many-core-Prozessorarchitekturen zu nutzen, parallelisiert. Das Paket zeigt starke lineare Skalierbarkeit bis zu 80 Cores und verarbeitet bis zu 28000 Ereignisse pro Sekunde.

Das KF Particle Finder Paket wurde erfolgreich in die erste Version des “First Level Event Selection” (FLES) Pakets, das die vollständige Online-Ereignis-Rekonstruktion im CBM Experiment durchführen wird, integriert. Innerhalb des FLES Pakets läuft das KF Particle Finder Paket parallel auf einem many-core-Cluster mit 100 Knoten mit jeweils 32 Cores. Das FLES Paket erreicht die Geschwindigkeit von  $2.2 \cdot 10^5$  Ereignissen pro Sekunde auf 3200 Cores des Clusters.

Das KF Particle Finder Paket kann weiter modifiziert werden. Die Liste der rekonstruierten Teilchen kann erweitert werden, um nach bereits bekannten oder ganz neuen, noch unerforschten Teilchen zu suchen. Es bietet die Möglichkeit,

kurzlebige Teilchen mit einem neutralen Tochterteilchen mithilfe der Methode der fehlenden Masse zu rekonstruieren. Es eignet sich perfekt zur Rekonstruktion von komplizierten Kaskadenzerfällen von Hyperkernen oder von hypothetischen Dibaryon-Teilchen. Aufgrund seiner Unabhängigkeit von der Detektorgeometrie kann das Paket auch in anderen Hochenergiephysik-Experimenten verwendet werden. Das Paket wurde auch in Bezug auf den Speicherverbrauch optimiert, was es zu einem sehr geeigneten Kandidaten für die Hardware-Beschleuniger wie GPGPUs und Intel Xeon Phi macht.

Das KF Particle Finder Paket bietet eine universelle Plattform zur Rekonstruktion kurzlebiger Teilchen, zur Physikanalyse und zur Online-Selektion von Ereignissen.