CZECH TECHNICAL UNIVERSITY IN PRAGUE

Faculty of Nuclear Sciences and Physical Engineering

Department of Physics

Research project

Reconstruction of strange hadrons in collisions of nuclei at RHIC

Author:	Bc. Jakub Kubát
Supervisor:	RNDr. Petr Chaloupka, Ph.D.
Consultant:	Ing. Michal Kocan
Year:	2019

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ
V PRAZE
Fakulta jaderná a fyzikálně inženýrská

Katedra Fyziky
Obor: Experimentální jaderná a částicová fyzika

Rekonstrukce podivných hadronů v jádro-jaderných srážkách na urychlovači RHIC

VÝZKUMNÝ ÚKOL

Vypracoval:	Bc. Jakub Kubát
Vedoucí práce:	RNDr. Petr Chaloupka, Ph.D.
Konzultant:	Ing. Michal Kocan
Rok:	2019

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
FAKULTA JADERNÁ A FYZIKÁLNĚ INŽENÝRSKÁ
PRAHA 1 - STARÉ MĚSTO, BŘEHOVÁ 7 - PSČ 115 19

Katedra: fyziky **Akademický rok:** 2018/2019

VÝZKUMNÝ ÚKOL

Student: Bc. Jakub Kubát

Studijní program: Aplikace přírodních věd

Obor: Experimentální jaderná a částicová fyzika

Vedoucí úkolu: RNDr. Petr Chaloupka, Ph.D.

Název úkolu (česky/anglicky):

Rekonstrukce podivných hadronů v jádro-jaderných srážkách na urychlovači RHIC

Reconstruction of strange hadrons in nucleus-nucleus collisions at RHIC collider

Pokyny pro vypracování:

Cílem práce je rekonstrukce podivných částic, především lambda baryonu, které jsou produkovány v jádro-jaderných srážkách na urychlovači RHIC (Relativistic Heavy Ion Collider) a detekovány experimentem STAR. Student se zaměří na využití a optimalizaci výběru sekundárních rozpadových vrcholů, které pocházejí z rozpadů podivných částic, s použitím nově vyvinutého systému „KFParticle finder". Výběrová kritéria budou optimalizována s pomocí nástrojů strojového učení.

Práce bude vypracována v anglickém jazyce.
Součástí zadání výzkumného úkolu je jeho uložení na webové stránky katedry fyziky.

Literatura:

1. K. Yagi, et al., Quark-Gluon Plasma, Cambridge University Press, 2005
2. I. Kisel, J. Phys. Conf. Ser. 1070, 012015 (2018)
3. T. O. Ablyazimov, et al., Phys. Part. Nucl. Lett. 12, 423-427 (2015)
4. S. Shi, J. Phys. Conf. Ser. 779, 012008 (2017)

Datum zadání: 22.10.2018

Datum odevzdání: 28.06.2019

....

vedoucí katedry

Prohlášení

Prohlašuji, že jsem svůj Výzkumný úkol vypracoval samostatně a použil jsem pouze podklady (literaturu, projekty, SW atd.) uvedenou v přiloženém seznamu.

Nemám závažný důvod proti užití tohoto školního díla ve smyslu § 60 Zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Praze dne ....

Bc. Jakub Kubát

Acknowledgments

I would like to express my gratitude to my supervisor, Dr. Chaloupka, for guiding me through the analysis presented in this work, for his time, valuable advices and shared experience. Next, I would like to acknowledge my consultant, Ing. Michal Kocan, for the time he gave me and his help with software employed in this research project. Also, I would like to show gratitude to Doc. Jaroslav Bielčík, Dr. Pavol Federič and Dr. Maksym Zyzak for all consultations they gave me. Last but not least, I am grateful to my friends and family for all their support.

Bc. Jakub Kubát

Název práce:
Rekonstrukce podivných hadronů v jádro-jaderných srážkách na urychlovači RHIC

Autor: Bc. Jakub Kubát

Studijní program: Aplikace přírodních věd
Obor: Experimentální jaderná a částicová fyzika
Druh práce: Výzkumný úkol

Vedoucí práce: RNDr. Petr Chaloupka, Ph.D., Katedra Fyziky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze
Konzultant: Ing. Michal Kocan, Katedra Fyziky, Fakulta jaderná a fyzikálně inženýrská, České vysoké učení technické v Praze

Abstrakt: KF Particle Finder je balíček napsaný v C++ pro rychlou a efektivní rekonstrukci krátce žijících částic. Přestože byl původně vytvořen pro účely online rekonstrukce na experimentu CBM v laboratoři FAIR, byl úspěšné implementován také na experimentu STAR v Brookhavenské Národní Laboratoři. V této práci byl balíček KF Particle Finder použit pro analýzu dat ze sážek zlatých jader při různých energiích naměřených během stále probíhající druhé fáze programu Beam Energy Scan na STAR. Diskutováno je potenciální vylepšení signifikance signálu podivného baryonu Λ v porovnání se standardním postupem v analýze, zejména pro nízké příčné hybnosti. KF Particle Finder byl též kombinován s klasifikační technikou vylepšených rozhodovacích stromů pro další vylepšení signifikance.

Klíčová slova: RHIC, STAR, kvark-gluonové plazma, Beam Energy Scan, KF Particle Finder

Title: **Reconstruction of strange hadrons in collisions of nuclei at RHIC**

Author: Bc. Jakub Kubát

Abstract: KF Particle Finder is a C++ package developed for fast and effective reconstruction of short-lived particles. Although it was initially created for the purposes of online reconstruction at CBM experiment in FAIR, it was successfully implemented at STAR in Brookhaven National Laboratory. The KF Particle Finder package was used in this work for analysing data from Au+Au collisions at various energies collected during ongoing second phase of Beam Energy Scan program at STAR. Possible improvement of signal significance of strange Λ baryon with respect to conventional analysis, especially in low p_T region, is discussed. KF Particle Finder was also combined with Boosted Decision Trees classifier for further improvement of significance.

Key words: RHIC, STAR, quark-gluon plasma, Beam Energy Scan, KF Particle Finder

Contents

Introduction

The standard model of elementary particles encapsulates our current understanding of building blocks of the universe and interactions that govern them. The three fundamental forces the standard model considers are strong, weak and electromagnetic interactions. The fourth important force, gravity, is missing from the picture and in standard model gravitational interaction between particles is actually neglected. In the description of standard model fundamental interactions are mediated by force-carrying particles which belong to the family of bosons.

The ordinary matter around us consists of atoms which have a positively charged nucleus in the center and negatively charged electrons around. The force which binds them together is the electromagnetic one. While electrons are considered elementary, it is not the case with protons and neutrons inside the nucleus. Subnuclear structure of protons was probed in 1968 with deep inelastic experiments at the Stanford Linear Accelerator Complex. At the time point-like particles contained in protons were called partons, but later they were identified with *up* and *down* quarks whose existence was predicted independently by Murray Gell-Mann [1] and George Zweig [2] in 1964. Although every day matter we encounter consist from already mentioned light quarks named *up* and *down*, there are actually three generations of quarks counting with total of six flavours.

Quarks interact with each other through strong interaction which is mediated by gluons. The strong force is described by Quantum Chromodynamics (QCD) which assigns a color charge to every strongly-interacting object. In contrast to photons, gluons are also charged with color. Under standard circumstances quarks and gluons are confined in color neutral particles called hadrons which proton and neutron are the most stable representatives of. Therefore, quarks or gluons are never observed directly. There are two groups of hadrons, baryons consist of three valence quarks with all possible color charges, i.e. red, green and blue, while mesons are formed by pair of quark and anti-quark.

The Quantum Chromodynamics predicts that the state of deconfined quarks and gluons can actually exist. Under extremely high temperatures and energy densities are expected to become asymptotically free and able to move inside a hot and dense medium quasi-freely on distances larger than the size of hadrons. This medium is called quark-gluon plasma (QGP) and has been intensely studied in heavy ion collisions at Relativistic Heavy Ion Collider [3] and Large Hadron Collider [4]. As for a hypothetical phase of a QCD matter it is important to study phase transition between gas of hadrons and the partonic degrees of freedom if there is any. Current understanding of QCD matter phase diagram is that there is a smooth crossover between hadron gas and QGP at low baryon densities and high temperature, while at higher densities the transition is of first order which means that there should also be a critical point in the diagram where the transition changes its nature.

The Beam Energy Scan program is dedicated to study of QCD phase diagram and phase transition to QGP at experiment STAR. In order to map out the QCD phase diagram, physicists need to measure heavy ion collisions at various collision energies. That is why the latest development in collider technology was focused not only on further increasing the collision energy as it was in the past, but also on improving luminosity at lower energies to compensate for small cross sections, i.e. probabilities, of researched phenomena.

The Compressed Baryonic Matter (CBM) experiment at future Facility for Antiproton and Ion Research (FAIR) in GSI, Darmstadt is another planned experiment meant to investigate QCD phase diagram [5]. The data rate promised to be delivered at FAIR is so high that new solutions for online reconstruction of particles had to be found. This motivated the development of KF Particle Finder package which employs Kalman Filter method for the purpose of reconstruction of short-lived particles [6, 7]. Besides its speed, KF Particle Finder brings also whole new approach to reconstruction of particles whose effectiveness, especially for particles with low transverse momentum, is subject of this work. KF Particle Finder was used with data collected at STAR during ongoing second phase of Beam Energy Scan for reconstruction of strange baryon Λ. The data were produced by Tracking Focus Group and made available for express physics analysis at STAR. KF Particle Finder is also combined with machine learning methods in order to investigate improvement in signal significance.

In the first chapter experimental setup of Relativistic Heavy Ion Collider (RHIC) and Solenoidal Tracker at RHIC (STAR) is described. Second chapter deals with Beam Energy Scan program and also fixed target program which complements BES with data from the lowest energy collisions. Strangeness as a probe of quark-gluon plasma is briefly discussed in this chapter too. In the third chapter KF Particle and KF Particle Finder packages are introduced. The next chapter looks at Toolkit for Multivariate Analysis with focus on Boosted Decision Trees and the fifth chapter present authors analysis of data from collisions of gold nuclei at BES energies with KF Particle Finder and TMVA. The work finishes with summary of what has been presented.

Chapter 1

Experimental setup

1.1 Relativistic Heavy Ion Collider

The Relativistic Heavy Ion Collider is particle accelerator built in Brookhaven National Laboratory (BNL) for the purpose of colliding heavy ions and protons at high luminosities in order to study transition of nuclear matter into quark-qluon plasma [8]. The construction of the facility was finished in 1999 and since than four experiments have operated there. These are BRAHMS, PHENIX, PHOBOS and STAR. Currently, only STAR is active because PHOBOS and BRAHMS were decommissioned in 2005 and 2006 respectively and PHENIX is being modernized into new experiment called sPHENIX [9]. In Fig. 1.1 the whole RHIC accelerator complex is depicted.

Before ions are injected into RHIC they are already accelerated in a complex of smaller accelerators. From the EBIS (Electron Beam Ion Source) they go through linear accelerator LINAC into Booster synchrotron. Booster divides ion bunches into six buckets, accelerates them to 95 MeV and injects them into AGS (Alternating Gradient Synchrotron) in four cycles. The AGS separates the total of 24 buckets and forms final four which are also accelerated to 10.8 GeV and sent to RHIC. The RHIC collider itself consists of two separate hexagonal storage rings titled "Blue" and "Yellow". The Blue ring accelerates ions clockwise and the Yellow ring counter-clockwise. Rings intersect each other at six locations, although only at four of them experiments are located. Beams of ions are steered by magnetic fields of superconductive dipole magnets at each vertex of the ring and the focusing is done with quadrupole magnets. The magnets have to be cooled down to temperature below 4,5 K [8].

The maximum beam energy RHIC was designed to reach is 100 GeV per nucleon. This allows the research of heavy ion collisions at center of mass energy $\sqrt{s_{NN}} = 200$ GeV. However, it is possible to run at lower energies which is important for the Beam Energy Scan program at STAR, since it is aimed to study the QGP phase diagram and therefore requires to explore systems with different baryon chemical potential [10]. The main difficulty of running at low energies is to focus the beam and deal with decreasing luminosity. This is solved for example with electron cooling which enables to go as low as $\sqrt{s_{NN}} = 7.7$ GeV with collisions energy [11]. For the purpose of improving luminosity at very low energies RHIC can also run in fixed target mode. In this mode only one beam circulates in one of the storage rings and collider operators are able to deflect it so it hits a gold target placed in a beam pipe near the STAR experiment. By including fixed target program to BES collisions down to $\sqrt{s_{NN}} = 3.0$ GeV can be studied [12].

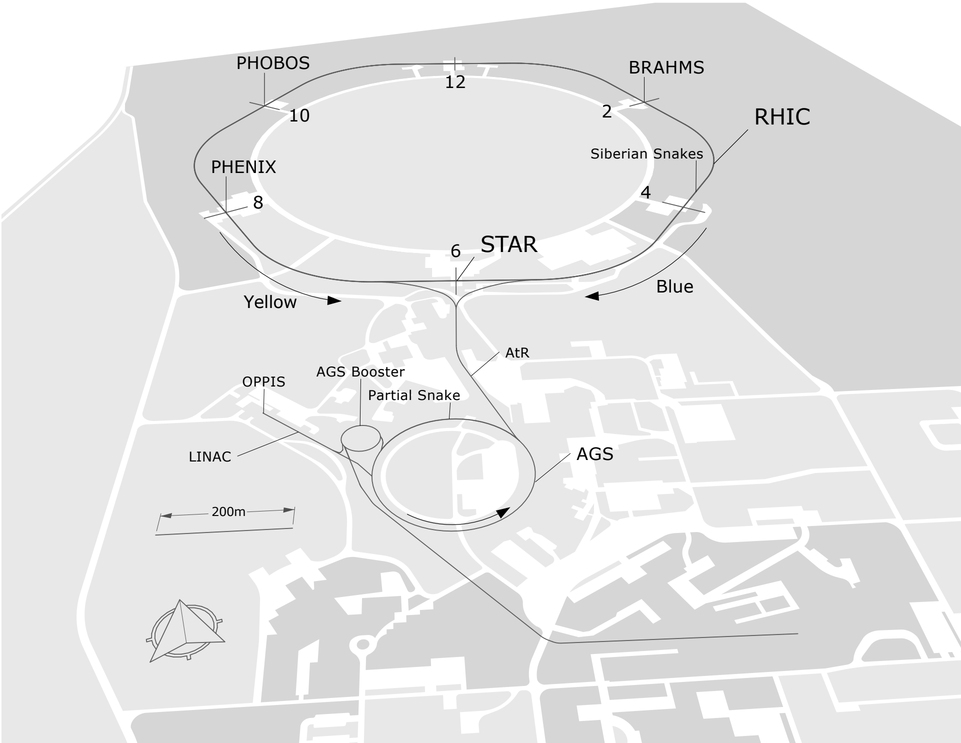

Fig. 1.1: Sketch of the RHIC accelerator complex. Taken from [13].

1.2 Solenoidal Tracker at RHIC

The experiment STAR (Solenoidal Tracker at RHIC) was constructed for the purpose of study of heavy ion collisions and quark-gluon plasma. For this reason it was designed to record high multiplicity events and cover full azimuth within pseudorapidity region $|\eta| < 1$. The detection system consists of Time Projection Chamber inside a solenoidal magnet to enable tracking, momentum analysis and particle identification via $\mathrm{d}E/\mathrm{d}x$. An Time of Flight detector for particle identification at higher momenta surrounds TPC. Outside of magnet there is also Barrel Electro-Magnetic Calorimeter for triggering and distinguishing electrons from hadrons [14].

Throughout the years STAR subsystems were upgraded, some new like Muon Telescope Detector (MTD) were added and some were already removed. In Fig. 1.2

reader can see a diagram of STAR and its subdetectors from year 2016. Note that there is Heavy Flavor Tracker (HFT) present, which was added to STAR in 2014 and remained in use until 2016 [15]. Most recent upgrades of STAR include inner Time Projection Chamber (iTPC) and endcap Time of Flight (eToF) which are now already completely built into tracking and time of flight measurement. Another planned upgrade denoted simply as STAR forward upgrade is meant to combine tracking system with electromagnetic and hadronic calorimeters in order to cover region $2.5 < \eta < 4.5$ in pseudo-rapidity [16]. Below two components of STAR will be described in more detail, namely Time Projection Chamber and Time of Flight detector.

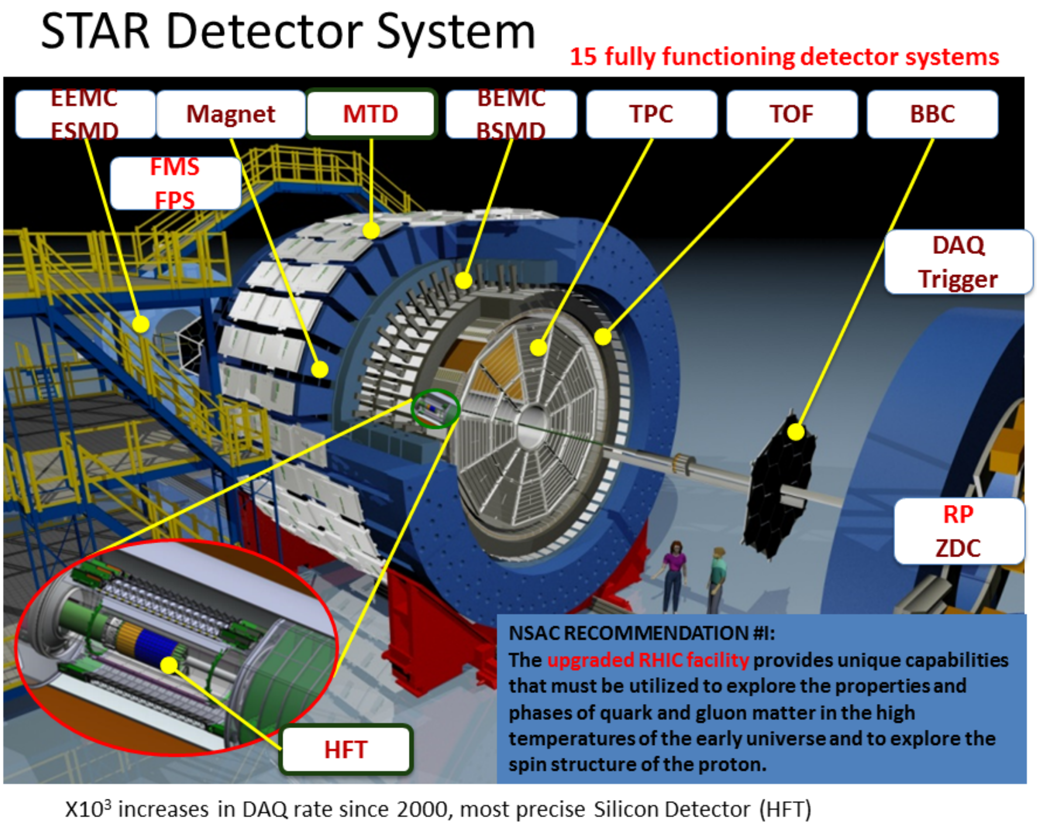

Fig. 1.2: STAR detector configuration from 2016. Taken from [17].

1.2.1 Time Projection Chamber

The Time Projection Chamber is a main STAR tool used for tracking of particles. It is positioned inside magnetic field of a solenoidal magnet with induction of $B = 0.5$ T and besides recording of tracks also measures their momenta and allows to determine the energy loss $\mathrm{d}E/\mathrm{d}x$ which helps to identify charged particles. Acceptance of the TPC covers $|\eta| < 1.8$ and it is possible to measure particle momenta in range from 100 MeV$/c$ to 30 GeV$/c$ [18].

The chamber itself is 4.2 m long and 4 m in diameter. It is filled with P10 gas (10% methane, 90% argon). The gas gets ionized by charged particles passing through volume of the chamber and releases electrons which drift up to 2.1 m to the readout end caps thanks to the present uniform electric field of $\approx$ 135 V/cm. The readout system consists of Multi-Wire Proportional Chambers with readout pads [18].

The inner Time Projection Chamber (iTPC) upgrade was introduced to STAR in 2019 for the purposes of Beam Energy Scan II and its physics goals, i.e. study of phase diagram of nuclear matter. The idea of the upgrade is to increase the segmentation on the inner padplane and renew the inner sector wire chambers. This should lead to improvement of tracking at small angles relative to beam line and hence expansion of acceptance of the chamber to $|\eta| \leq 1.5$ [19]. The momentum and $\mathrm{d}E/\mathrm{d}x$ resolution as well as acceptance at low momenta should also benefit from this update as is discussed further in the analysis section of this work.

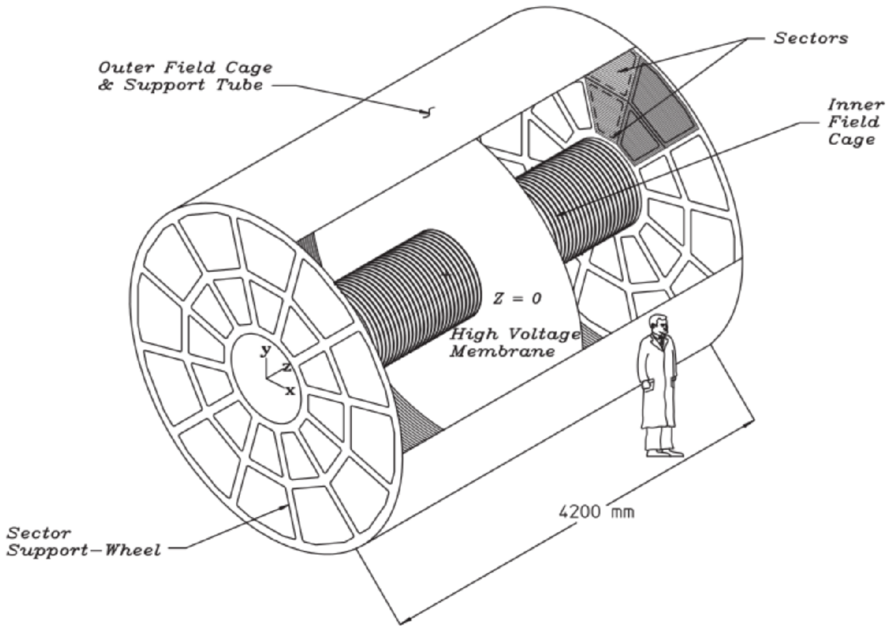

Fig. 1.3: Schematic depiction of STAR TPC. Taken from [18].

1.2.2 Time of Flight detector

The barrel Time of Flight detector is based on Multi-gap Resistive Plate Chamber Technology (MRPC) [20]. It was built in order to improve particle identification for particles with momenta up to 3 GeV/c. MRPC modules are arranged in 120 segments around a cylinder which covers $|\eta| < 0.9$ in pseudo-rapidity. They take the time when particle passes through the detector, while Vertex Position Detector (VPD) sets start time of the flight.

Endcap Time of Flight (eToF) was similarly to iTPC installed at STAR for run 19 in order to further enhance PID. It incorporates technology that will be used at future project CBM in FAIR facility which has been adapted for STAR.

Chapter 2

Beam Energy Scan

This chapter will discuss physics motivation behind Beam Energy Scan program. Next, fixed target program that is crucial for obtaining data from heavy ion collisions at low energies will be described. Finally, strangeness as a probe of quark-gluon plasma and its properties will be explored.

2.1 QCD Phase Diagram

The main idea of Beam Energy Scan is to investigate the phase diagram of strongly interacting matter. The current understanding of nuclear matter is illustrated in Fig. 2.1. The diagram depicts relation between temperature and the so called baryon chemical potential μ_B which quantifies balance between matter and anti-matter. Finite temperature lattice QCD calculations suggest that the transition from hadron gas to the state of deconfined quarks and gluons takes form of crossover at vanishing baryon chemical potential and temperature around $T_c = 154 \pm 9$ MeV [21]. That means the transition is not accompanied by any discontinuities. On the other hand, based on QCD calculations it is believed that at lower temperature and higher chemical potential the transition is of first order [22]. If correct this would mean that there has to be a critical point at which the first order phase transition changes to crossover.

When establishing signatures of presence of QGP in heavy ion collisions, STAR investigated mainly collisions of gold nuclei at center of mass energy $\sqrt{s_{NN}} = 200$ GeV and found no evidence of first order transition. At this energy or at even higher energies at LHC the baryon chemical potential is very low since matter and anti-matter are created equally. So in order to map out the QCD phase diagram one has to go lower with collision energy. This is why STAR already collected data at energies from $\sqrt{s_{NN}} = 200$ to 7.7 GeV during first phase of Beam Energy Scan in 2010 and 2011 [23], although with very low statistics. The second phase of BES introduces upgrades to both STAR (iTPC, eToF, EPD) and RHIC (electron cooling) and will add statistics from collisions at $\sqrt{s_{NN}} = 7.7$, 9.1, 11.5, 14.6 and 19.6 GeV while the linked fixed target program will supply data from collisions at $\sqrt{s_{NN}} = 3.0$, 3.2, 3.5, 3.9 and 4.5 GeV [24]. The goals of BES include searching for the predicted first order phase transition and the critical point, investigate expected turn-off of QGP signatures and look for evidence of chiral symmetry restoration.

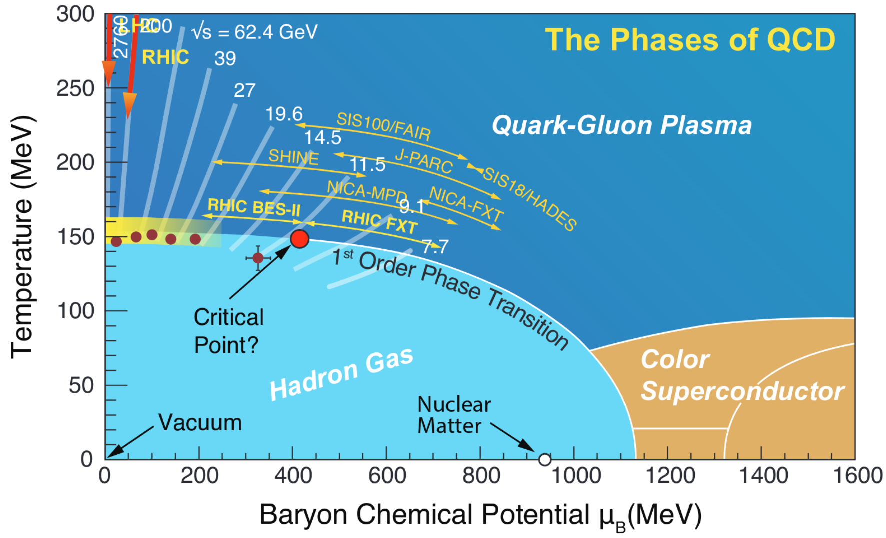

Fig. 2.1: Illustration of phase diagram of nuclear matter. Taken from [10].

2.2 Fixed Target Experiment Setup

As mentioned in previous section fixed target program is meant to complement the collection of data from collisions with low center of mass energies and thus high baryon chemical potential. It allows to go as low as $\sqrt{s_{NN}} = 3.0$ GeV with collision energy which corresponds to $\mu_B = 721$ MeV [24]. When running in fixed target mode there is only one beam circulating in RHIC which is subsequently deflected by RHIC operators to hit a fixed gold target placed inside beam pipe 200 cm from the center of TPC. The energy of the beam is actually the same as for one of the BES-II settings but the center of mass energy in fixed target mode is lower. So for example the 14.6 GeV collisions were measured with two beams having 7.3 GeV each and the same beam energy was used for fixed target mode to get 3.9 GeV collisions.

Before the launch of fully-fledged fixed target program there had to be several tests done [25]. First, as a proof of principle analysis of collisions between ions in beam halo and aluminium nuclei in the vacuum pipe was carried out. Fig. 2.2 shows vertices of these collisions. After that in 2014 a thin gold foil target was installed in the beam pipe on the west edge of the TPC about 210 cm from the center. The target was however illuminated only with gold beam halo again. In Fig. 2.4 one can see detector setup from Run14 during which the fixed target test was done. Fig. 2.3 is a photo of the gold foil target and its support structure. Fig. 2.5 shows vertices of collisions between gold beam halo and gold nuclei in target.

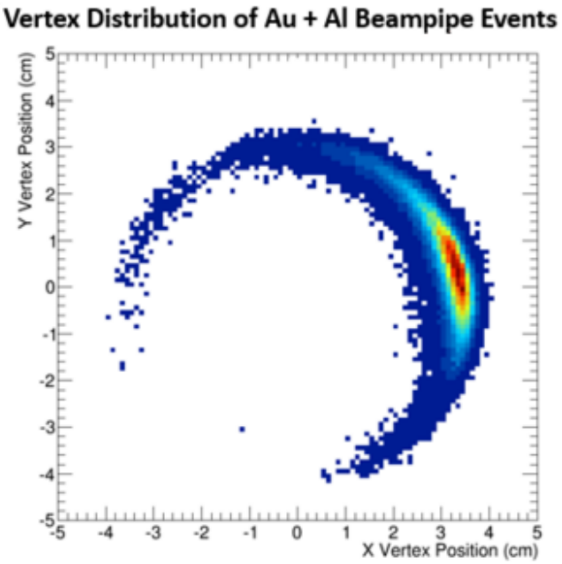

Fig. 2.2: Vertices of collisions between gold beam halo and aluminium nuclei in beam pipe. Taken from [25].

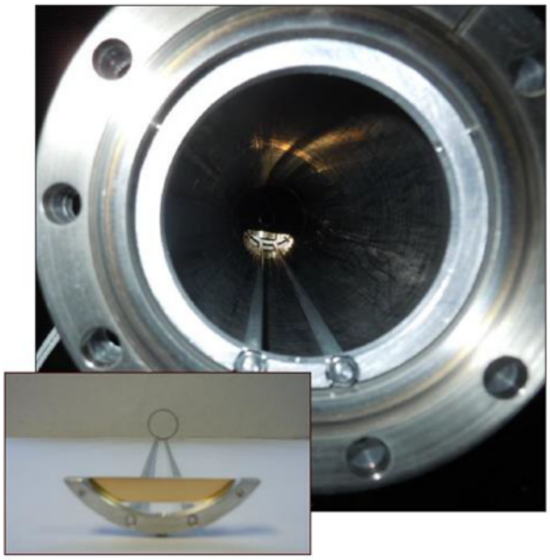

Fig. 2.3: Photo of gold foil target and its support structure. Taken from [25].

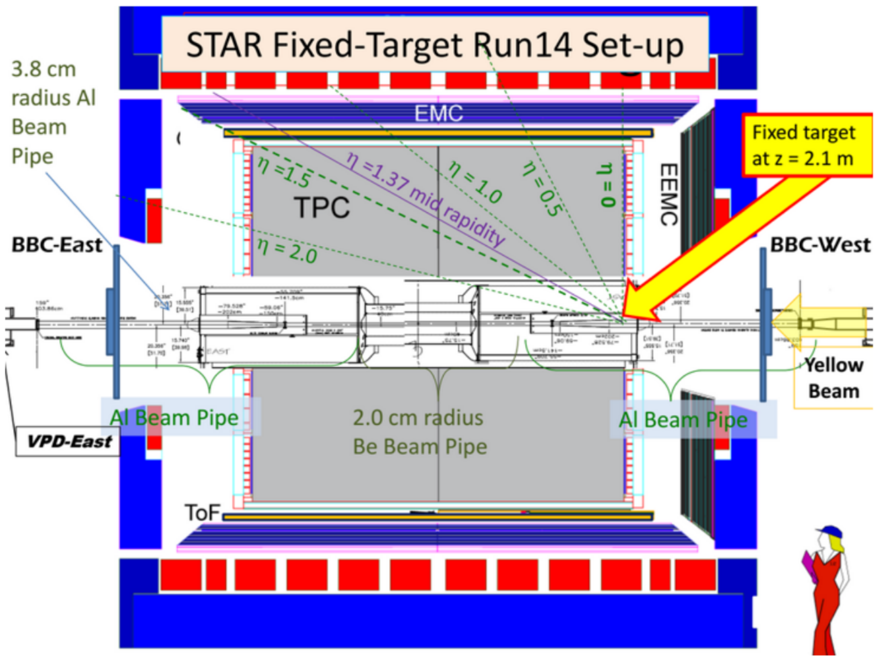

Fig. 2.4: Diagram of the fixed target test run detector setup. Taken from [25].

2.3 Strangeness in Heavy Ion Collisions

Strangeness has played a critical role in establishing quark-gluon plasma signatures in heavy-ion collisions. The increase in production of strange and multi-strange hadrons in heavy ion collisions with respect to p+A or p+p interactions was one of the first proposed signatures of quark-gluon plasma formation [26] and is considered a "smoking-gun" for confirming QGP presence in heavy ion collisions.

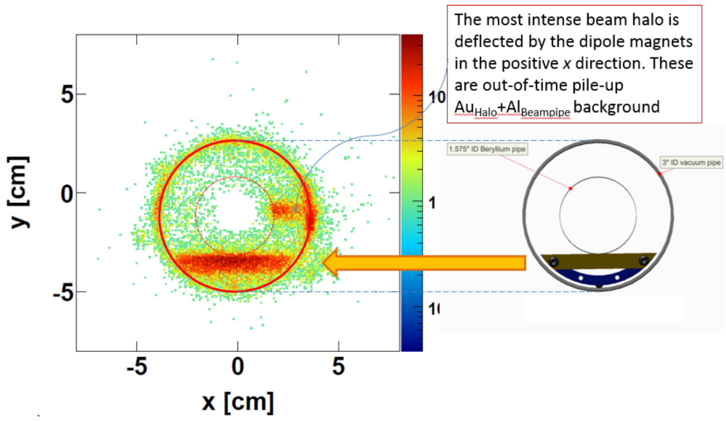

Fig. 2.5: Vertices of collisions between gold beam halo and gold nuclei in target. Taken from [25].

Since in quark-gluon plasma pairs of strange quarks can be produced directly, i.e. not confined in strange hadrons, it costs less energy to create them. Later, in the hadronization stage these strange quarks bind with light flavour quarks to form strange hadrons and thus it is argued that, thanks to the presence of QGP, the production of strange hadrons is enhanced [27, 28]. Since the production of strange particles is important probe for the search of the nuclear matter phase boundaries, the analysis of strangeness is crucial for Beam Energy Scan.

If one denotes light quarks with $q = u, d$, the stable strange particles that can be considered for reconstruction are following:

$$\phi(s\bar{s}), K(q\bar{s}), \bar{K}(\bar{q}s), \Lambda(qqs), \bar{\Lambda}(\bar{q}\bar{q}\bar{s}), \Xi(qss), \bar{\Xi}(\bar{q}\bar{s}\bar{s}), \Omega(sss), \bar{\Omega}(\bar{s}\bar{s}\bar{s}).$$

Multi-strange hadrons such as Ω or Ξ are subject to cascading decay in which neutral Λ baryon is produced which is invisible to TPC. Therefore, when analysing Λ particles it is usual to divide them into primary and secondary. Secondary Λs can be subsequently combined with for example π^- in order to reconstruct Ξ^- as is illustrated in Fig. 2.6.

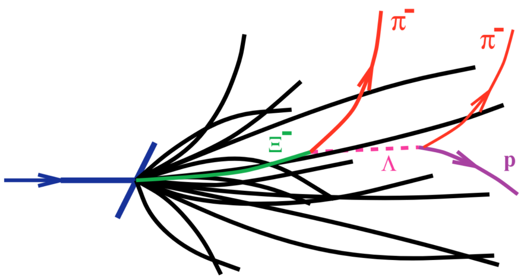

Fig. 2.6: Illustration of Ξ^- cascade. Taken from [27].

Chapter 3

KF Particle framework

In this chapter KF Particle package and KF Particle Finder will be described. These C++ packages offer means to reconstruct short-lived particles from parameters of their daughters, i.e. stable particles like protons and electrons or particles that live long enough to reach detectors of particle physics experiments.

3.1 KF Particle

The KF Particle is a package of C++ libraries developed at FIAS (Frankfurt Institute for Advanced Studies) initially for experiments CBM and ALICE [6]. It is based on Kalman filter (KF) algorithm and its purpose is to expand upon Kalman filter mathematics in order to offer means of fast and effective reconstruction of tracks and particles in High Energy Physics experiments. The development was motivated by the ever increasing amount of data collected at collider experiments and the need to process them at high rates. For example, the CBM experiment which is currently under construction at FAIR facility in GSI (Darmstadt, Germany) will have to analyse up to 10^7 events per second online since there is no way to store this amount of data on any physical media [29].

The package describes both mother and daughter particles with a state vector

$$\vec{r} = (x, y, z, p_x, p_y, p_z, E, s), \tag{3.1}$$

where (x, y, z) are space coordinates, (p_x, p_y, p_z) are components of momentum, E is particle energy and $s = l/p$ is the time between production and decay point of the particle measured in a distance normalized on the momentum. The covariance matrix than includes information about accuracy of the parameter estimation and allows to calculate χ^2 criteria that characterize quality of reconstruction. In the reconstruction of particles these criteria can be later used for background rejection. This is the main difference from conventional analysis which does not take errors of daughter tracks into account in any way, although they affect the quality of reconstruction of the mother particle. Other important properties of KF Particle package include

- independence of experiment geometry
- ability to reconstruct decay chains
- ability to run on computers with SIMD architecture [7].

The independence of experiment geometry is one of the reasons why both KF Particle and KF Particle Finder were successfully implemented at STAR and also why it is simple to switch between collider mode and fixed target experiments as is demonstrated further in this work. The ability to reconstruct decay chains stems from equal description of mother and daughter particles since any mother particle can enter the process of reconstruction again as a daughter of heavier particle. The reason for employment of parallel programming and Simple Instruction Multiple Data architectures comes again from the aim to make the reconstruction as fast as possible in order to use it for online reconstruction in high intensity and high multiplicity environments such as CBM will be [30]. It was shown that the speed of reconstruction with KFPF scales linearly with the number of CPUs [31]. The benefits of parallel computing were not explored in this work, however.

3.2 KF Particle Finder

The KF Particle Finder is another C++ package which employs methods of working with particle state vectors and correlation matrices defined in KF Particle [7]. It offers an algorithm to reconstruct particles in wide range of decay channels and comes with an interface that enables user to study their properties.

The reconstruction of short-lived particles begins with providing tracks of charged particles detected in the experiment as an input. In the case of STAR the data come typically in the form of MuDst.root or PicoDst.root format. The input should also include the correlation matrix with information about track uncertainties. At first, KF Particle Finder classifies tracks into primary and secondary, i.e. tracks that either do or do not come from primary vertex - the position in Time Projection Chamber where the collision of ions happened. For this purpose, χ^2_{prim} criterion is calculated

$$\chi^2_{prim} = \Delta\vec{r}^T \left(C_{track} + C_{PV}\right)^{-1} \Delta\vec{r}, \tag{3.2}$$

where $\Delta\vec{r}$ is difference between the track and the primary vertex position and C_{track} and C_{PV} are covariance matrices of the track and primary vertex respectively. The criterion is basically distance between the track and primary vertex normalized on the total error. Under the assumption of particle parameters following Gaussian distribution and χ^2_{prim} following χ^2 distribution, the criterion represents probability of the trajectory intersecting primary vertex within uncertainties, i.e. classifying as primary. So for example if the criterion value is higher than $\chi^2_{prim} > 18.6$ there is an probability of 0.01% that the track is primary and this is the value that is used in the KFPF code by default to divide tracks into primary and secondary.

Other probability criteria are constructed in similar fashion and can be used in the analysis for background rejection:

- χ^2_{fit}/NDF criterion characterizes probability of daughter trajectories intersecting within their uncertainties,
- χ^2_{topo}/NDF characterizes whether the mother particle comes from the primary vertex region and therefore it is used to divide mother particles into primary and secondary,
- $l/\Delta l$ is a distance from the primary vertex to the decay point (decay length) normalized on its error.

The two main classes that any user should focus on when using KF Particle Finder for analysis are `StKFParticleAnalysisMaker` and `StKFParticleInterface`, however in order to really control the package one needs to go through the whole code, since there are still some hard-coded cuts.

The `StKFParticleAnalysisMaker` is a standard STAR maker with `Init()`, `Make()` and `Finish()` functions. Here the reconstruction of particles is initiated and it is where user can go through all reconstructed particles and work with them, e.g. store them into ntuples.

`StKFPartcileInterface` controls processing of events. In the method `ProcessEvent()` one can setup their own event-selection cuts or track quality cuts on track parameters. There are also methods for particle cuts defined. These can be called in macro that runs the analysis once the instance of `StKFParticleInterface` is created. In this way one can apply cuts not only on criteria such as χ^2_{fit} or $l/\Delta l$ but also on maximum distance between particles and so on, although for example the cut on χ^2_{topo} is hard-coded and can be modified in class `KFParticleTopoReconstructor`. In the analysis macro user is also able to specify the decay to be studied. Currently, KF Particle Finder offers over 70 decay channels that can be analysed as is demonstrated in Fig. 3.1 [7].

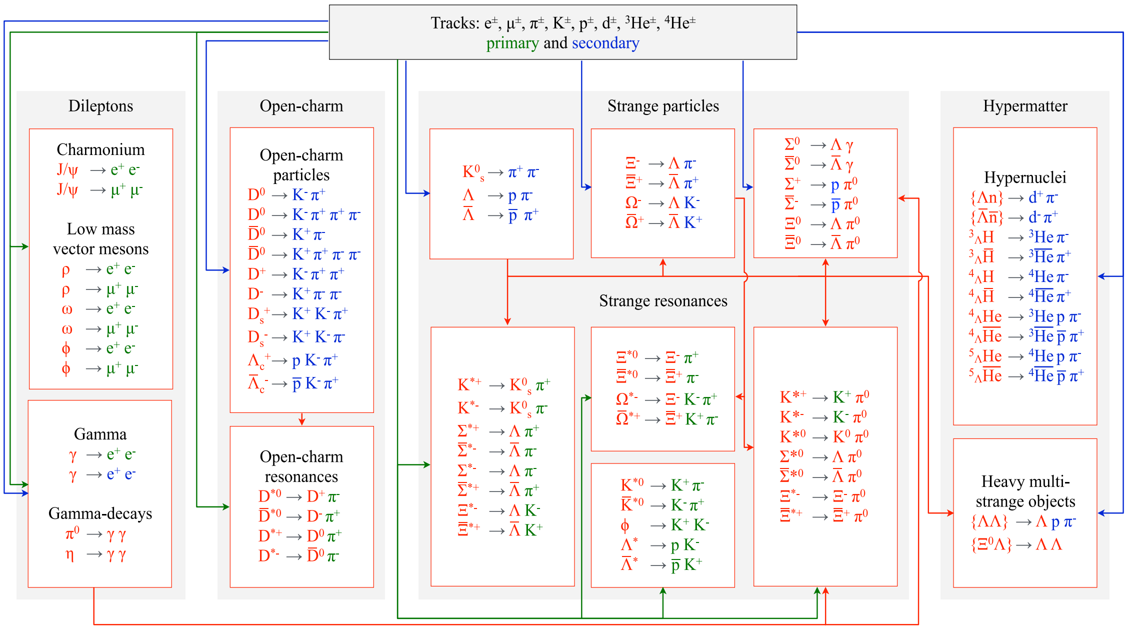

Fig. 3.1: List of decays available for analysis with KF Particle Finder. Taken from [7].

Chapter 4

Toolkit for Multivariate Analysis and Boosted Decision Trees

4.1 Toolkit for Multivariate Analysis

Multivariate analysis and machine learning techniques are becoming increasingly more popular with statisticians in science and industry. This includes also high-energy physics where it is required to identify rare signals in immense background [32]. The idea of multivariate analysis is to combine all input variables into single output one which is then typically used for classifying signal and background events.

ROOT integrated Toolkit for Multivariate Analysis (TMVA) was designed specifically for HEP, however it is not restricted for use in any other field. It offers wide selection of machine learning methods for both classification and regression implemented with C++. Besides large variety of multivariate classification algorithms TMVA comes with user interface through which classification via all available algorithms can be carried out simultaneously. The training and testing is controlled via C++ macro and there is even GUI available for displaying performance evaluation plots. The TMVA package includes following methods [33]:

- Neural networks
  - Deep networks
  - Multilayer perceptron
- Boosted/Bagged decision trees
- Function discriminant analysis (FDA)
- Linear discriminant analysis (H-Matrix, Fisher and linear (LD) discriminants)
- Multidimensional probability density estimation (PDE - range-search approach and PDE-Foam)
- Multidimensional k-nearest neighbour method

- Predictive learning via rule ensembles (RuleFit)
- Projective likelihood estimation (PDE approach)
- Rectangular cut optimisation
- Support Vector Machine (SVM)

4.2 Boosted Decision Trees

Together with neural networks Boosted Decision Trees (BDT) belong to the most popular machine learning classifiers in high-energy physics [32]. For example, they were employed in analysis of Higgs boson at CMS in CERN [34].

To grow a single decision tree one starts at the root node and splits input data with respect to value of a training variable into two subsets. By repeating the process recursively for subsets more nodes are formed until a whole decision tree is built as is ilustrated in Fig 4.1. The splitting variable is chosen at each node to maximize separation between signal and background. The separation may be quantified by various separation criteria, e.g. Gini Index or statistical significance. The splitting of nodes is repeated until some stopping condition is met, e.g. maximum depth of the tree. End-nodes of a decision tree are called leaf nodes and these are classified as background or signal depending on the majority of training events that end up in the node.

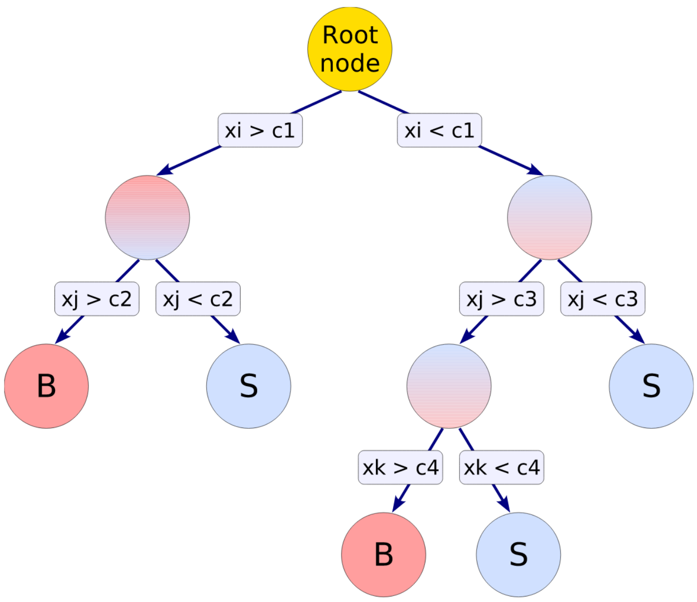

Fig. 4.1: Schematic view of a decision tree. Taken from [35].

The boosting of a decision tree means growing a large set of trees - a forest. Such a forest typically consists of a large number of shallow trees since the boosting in general is based on combining a large amount of weak learners (classifiers) into effective single one. The final classifier is given by weighted average of individual decision trees. In the case of TMVA the output of BDT is called BDT response value. It is in range from -1 to +1 and characterizes how is the particular event "background-like" or "signal-like". When compared to a single decision tree, boosting is better in performance and it is also able to stabilize the decision response with respect to fluctuations in the training sample. There are several ways to introduce boosting to decision trees. The approach used in the analysis of this work is called Adaptive Boosting (AdaBoost). In the process of growing a forest with adaptive boosting, events that were misclassified during the training of a decision tree are given a higher event weight in the training of the following tree [35].

One of the disadvantages of decision trees is susceptibility to overtraining. Overtraining occurs when MVA follows statistical fluctuations of input data. The performance of overtrained classifier will not be reproducible on independent training data. To avoid overtraining trees can be pruned. In the process of pruning trees are cut back from bottom after they have reached maximum size. This is meant to remove statistically insignificant splits (nodes) in the data. To check for overtraining user has the option of plotting superimposed distributions of BDT response value for training and testing events in GUI of TMVA. If the distributions overlap there is most probably no overtraining present, but if they do not, the classifier has probably learnt fluctuations in training sample and is therefore overtrained. In Fig 4.2 examples of two overtraining checks from the analysis of Au+Au collisions at 27 GeV per nucleon from this work are plotted. The left plot comes from training of BDT in transverse momentum region $p_T = 0.1 - 0.2$ GeV$/c$ and the right one from $p_T = 0.8 - 1.0$ GeV$/c$. One can observe that for higher p_T the distributions of BDT response value are identical, but for low p_T there is a small difference between training and test samples of the signal. This is mainly because of lower statistics obtained from the simulation at low p_T.

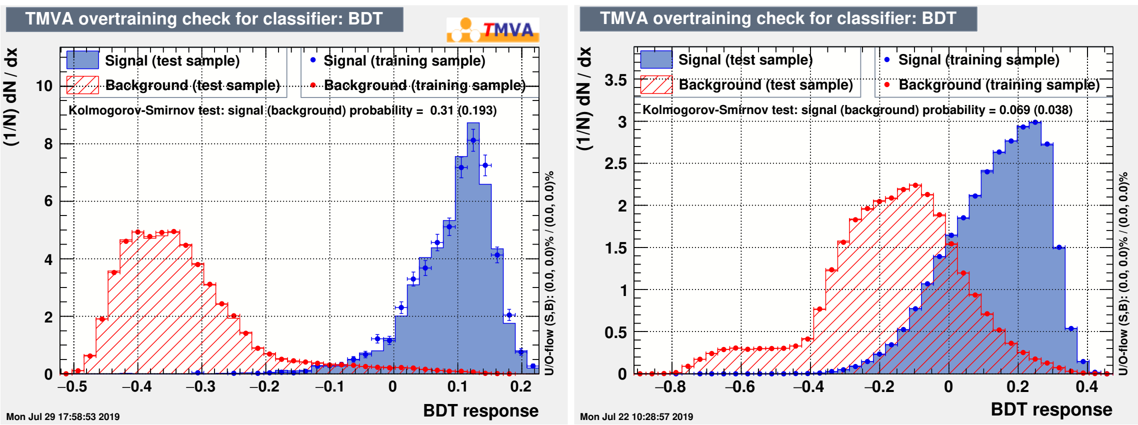

Fig. 4.2: Distributions of BDT response value for signal and background from training and testing samples in two transverse momentum bins: $p_T = 0.1 - 0.2$ GeV$/c$ (left) and $p_T = 0.8 - 1.0$ GeV$/c$ (right). Data come from Au+Au collisions at $\sqrt{s_{NN}} = 27$ GeV.

Chapter 5

Analysis

In this chapter the results from reconstruction of strange baryon Λ in heavy ion collisions done with both KF Particle Finder (KFPF) and Boosted Decision Trees (BDT) will be described. The decay channel $\Lambda \longrightarrow p + \pi^-$ was investigated and the data from Au+Au collisions were measured at experiment STAR. This includes collisions in collider mode at energies of $\sqrt{s_{NN}} = 27$ GeV, 14.6 GeV and fixed target collisions with $\sqrt{s_{NN}} = 3.9$ GeV.

5.1 Motivation

The motivation behind employing KF Particle Finder and subsequently Boosted Decision Trees will be briefly discussed in this section. Reasons for reconstructing strange particles and studying their properties in collisions with lower center of mass energies were already examined in chapter 2 about Beam Energy Scan.

As explained in chapter 3 about KF Particle Finder, the package takes into account errors of tracks measured by Time Projection Chamber. These errors are used to calculate χ^2-based statistical criteria which serve to reject background in reconstruction of particles. Concerning effectiveness of reconstruction of particles, the approach rooted in study of probabilities may turn out to surpass conventional analysis which employs topological cuts that do not take uncertainties into account at all. The effect might manifest itself especially when dealing with particles with low transverse momentum, since their tracks are sharply curved in the magnetic field of the detector and for this reason the fit of the trajectories comes with significant errors. These errors stem mainly from the fact that the helix model does not take into account changes of track curvature due to the energy loss.

It is worth mentioning that the data from collisions at 14.6 GeV were collected during 2019 which introduces iTPC upgrade of Time Projection Chamber into the tracking. This upgrade was meant to further enhance momentum resolution, energy loss resolution and improve acceptance at high rapidity or low momentum. Therefore, the aim of this work was to test the performance of the upgrade by analysing the data from express production with KFPF.

The main idea behind application of the KFPF on the data from the fixed target collisions was to test one of the prominent attributes of the package, specifically its independency of the experiment geometry. The data also come from the express production of 2019 run so there was no other analysis of these data done at STAR yet.

5.2 Au+Au collisions at $\sqrt{s_{NN}} = 27$ GeV

5.2.1 Dataset and event selection

The data from Au+Au collisions at $\sqrt{s_{NN}} = 27$ GeV were collected by experiment STAR during 2018. Production code-named P19ib of picoDst format was done using library SL19b. Total number of minimum bias events sampled in this analysis is 610M. From these minimum bias events, 370M were selected for further analysis. The event selection was based on the position of primary vertex of each collision. It was required that the primary vertex is less than 70 cm away from the center of the Time Projection Chamber, i.e. $|V_z| < 70$ cm. Other criteria for event selection were imposed using KFPF method `CleanLowPVTrackEvents`. This method demands at least 10 % of tracks to classify as primary and requires decent precision of primary vertex reconstruction in transverse plane, specifically $\sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} \leq 0.45$. Fig. 5.1 shows the distribution of z coordinate of primary vertex before event selection on the left, the plot on the right demonstrates distribution post selection.

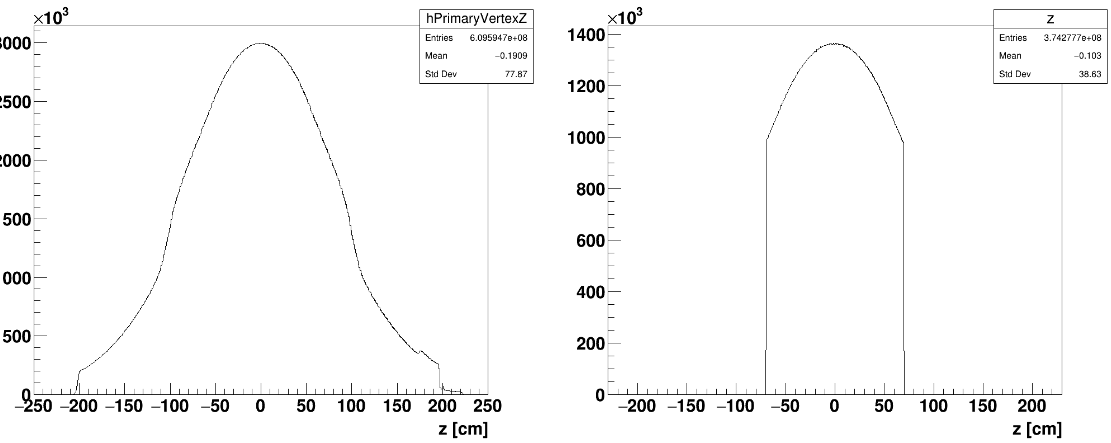

Fig. 5.1: Distribution of z coordinate of primary vertex in minimum bias events (left) from $\sqrt{s_{NN}} = 27$ GeV dataset. Distribution of z coordinate of primary vertex after event selection (right).

5.2.2 Analysis with default KFPF cuts

At first, in order to test the functionality of KF Particle Finder with custom changes in event-processing part of the code, the analysis of Λ was done with default particle cuts in KFPF. This means that cuts on various statistical criteria described in chapter 3 were set by constructors of relevant classes of the package as implemented by original developers. These values are tuned for CBM experiment and are explicitly listed in Tab. 5.1.

Cut value	Cut description
$\chi^2_{geom} < 10$	χ^2 of the track to the second daughter track
$l/dl > 5$	decay length normalized on the error
$l > 5$ cm	decay length
$\chi^2_{prim} > 18.6$	χ^2 of the track to primary vertex
$\chi^2_{topo} < 5$	χ^2 of the mother particle to primary vertex
$d_{max} < 1$ cm	maximum distance between daughter particles

Tab. 5.1: Default values of KFPF cuts.

The invariant mass spectra were plotted for all Λ particle candidates satisfying the criteria above in several transverse momentum bins with changing bin width in range from 0.1 to 6.0 GeV/c. There is no signal in transverse momentum bin from 0.0 to 0.1 GeV/c and this is caused by kinematics of the decay and the momentum accpetance of TPC. In the rest frame of Λ, its decay daughters, proton and pion, each carry momentum of around 101 MeV/c [36]. For stationary Λ, rest and lab frames are identical and thus daughter tracks of such Λ are out of acceptance of TPC since it is able to detect only pions with momentum $p_\pi > 150$ MeV/c [18]. For non-stationary Λ with low transverse momentum the situation gets more complicated, since one needs to boost mother particle into lab frame in order to calculate p_T of daughters, but the logic stays the same. This should be improved with iTPC upgrade [19].

The distribution of mass was fitted in fixed range $M_\Lambda = 1.096$ - 1.138 GeV/c^2 with double gaussian sitting on a polynomial background. There was no physics motivation behind the usage of double gaussian to describe the signal other than inspiration from other analyses done at STAR and reasonably successful fit of pure Λ signal simulation generated with STAR VMC, i.e. Monte Carlo simulation with detector effects done with GEANT. The polynomial background was of third order for $p_T < 0.4$ GeV/c and of second order for $p_T > 0.4$ GeV/c.

The broader of the two gaussians was selected to define the 3σ mass window in which the sum of signal and background was calculated as a sum of bin content in this region. The background was estimated by integrating the polynomial function and dividing the result by the bin width of the histogram. The resulting raw signal yield was than obtained by subtracting the background counts from the total sum of bin content.

In order to evaluate the effectiveness of the analysis with KFPF or to compare it with other methods, the significance of the signal was calculated as it is usually done in particle physics

$$\alpha = \frac{S}{\sqrt{S+B}}. \tag{5.1}$$

For the purpose of comparison, the significance was also recalculated for the total number of events after selection equal to 1M. Employing the assumption of signal and background counts being proportional to the number of events, the projected significance can be calculated as follows

$$\alpha_1 = \sqrt{\frac{N_1}{N_0}}\alpha_0. \tag{5.2}$$

As an example, plots from the lowest transverse momentum bins $p_T = 0.1 - 0.2$ GeV/c and $p_T = 0.2 - 0.4$ GeV/c, from the medium bin $p_T = 1.6 - 1.8$ GeV/c and from the high p_T bin $p_T = 5.0 - 6.0$ GeV/c are shown in Fig. 5.2. Note that in the first bin the signal was not successfully extracted, however there was an indication of a peak visible which further strengthened the motivation for employing machine learning techniques with ambition to reach lower p_T. Fig. 5.3 presents raw yield divided by bin width and significance dependence on transverse momentum. Both quantities are recalculated per 1M events post selection.

5.2.3 BDT training

The TMVA package offers broad selection of machine learning methods. For this particular analysis, Boosted Decison Trees were employed. The training was done using default BDT settings, i.e. adaptive boosting, number of trees $N_{tree} = 850$ with maximum depth equal to 3, number of cuts $N_{cuts} = 20$ and Gini index for separation in nodes. For more detailed description of the method see chapter 4.

To train the BDT, user has to supply a sample of signal and background. As a signal simulation of pure Λ signal generated with STAR VMC was used. The simulation lacks time of flight information because it was not included in the STAR VMC package and thus this variable was not used for training. There were 20 Λ particles in each generated event and the distribution of transverse momentum was thermal. Background sample was taken from the sideband region in the data, however the Λ candidates come from approximately 1/5 of full statistics. The width of the mass window that was cut from the background sample was decided after investigating width of the simulated signal peak. Both background and simulation were analysed with KFPF using looser cuts to give BDT space to operate in. The list of these criteria can be found in Tab. 5.2. As an output from KFPF analysis,

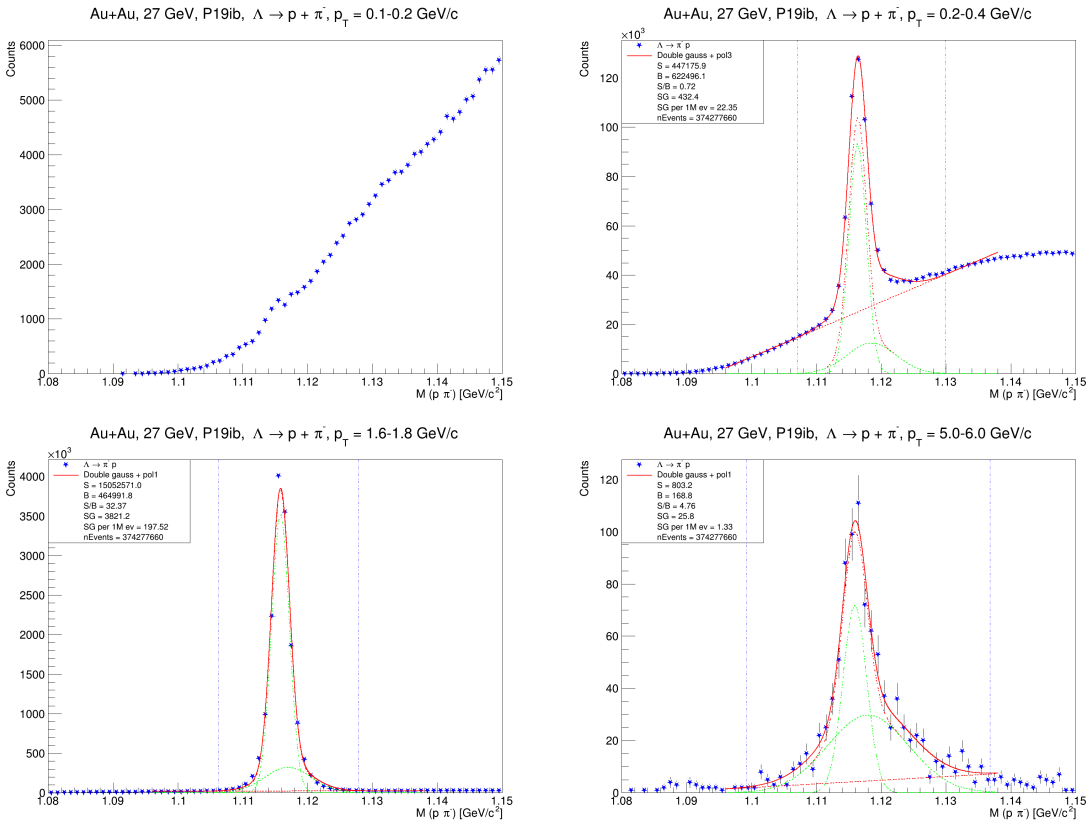

Fig. 5.2: Invariant mass spectra of Λ candidates in various transverse momentum bins.

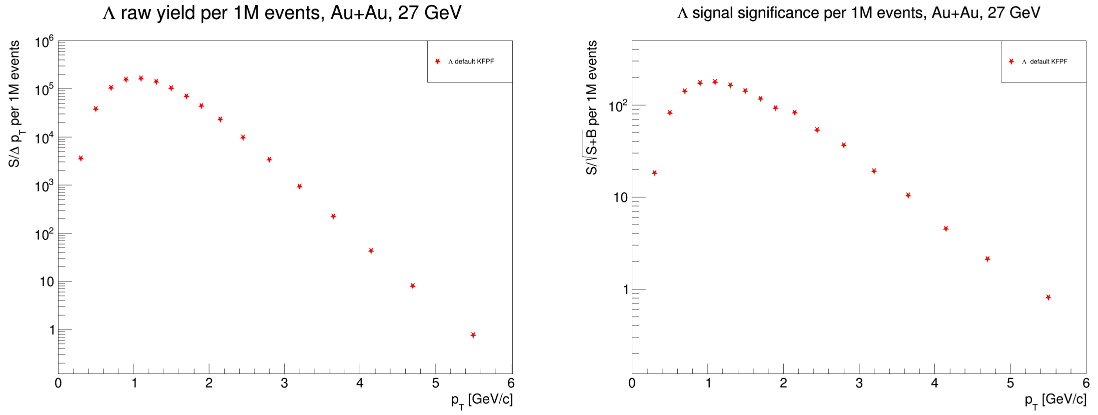

Fig. 5.3: Uncorrected Λ yield (left) and signal significance (right) dependence on transverse momentum. Both quantities are recalculated per 1M events post selection.

ntuples filled with values of training variables for each candidate were stored.

Since the training was completed in each transverse momentum bin separately, following plots serve as an example from bin with $p_T = 0.2 - 0.4$ GeV$/c$. The variables the BDT were trained on are presented in Tab. 5.3. Distributions of these

Cut value	Cut description
$\chi^2_{geom} < 14$	χ^2 of the track to the second daughter track
$l/dl > 3$	decay length normalized on the error
$l > 1$ cm	decay length
$\chi^2_{prim} > 3$	χ^2 of the track to primary vertex
$\chi^2_{topo} < 5$	χ^2 of the mother particle to primary vertex
$d_{max} < 1$ cm	maximum distance between daughter particles

Tab. 5.2: Selection criteria used in KFPF for obtaining samples for the training of Boosted Decision Trees.

quantities for both signal and background are plotted in Fig. 5.4. In Fig. 5.5 one can observe linear correlation coefficients of training variables for signal and background. The Fig. 5.6 contains the ROC curve which describes the relation between background rejection and signal efficiency (left plot). The curve follows expected trend which expresses the fact that it is not possible to train a classifier which is able to reject increasingly more background without sacrificing signal, i.e. loosing signal efficiency.

Variable code-name	Description
`pt_P`	p_T of daughter proton
`chi2Primary_P`	χ^2 of daughter proton track to PV
`pt_Pi`	p_T of daughter pion
`chi2Primary_Pi`	χ^2 of daughter pion track to PV
`Chi2NDF`	χ^2 of daughter track to another
`LdL`	normalized decay length
`Chi2Topo`	χ^2 of mother particle to PV

Tab. 5.3: Training variables used in classification with BDT.

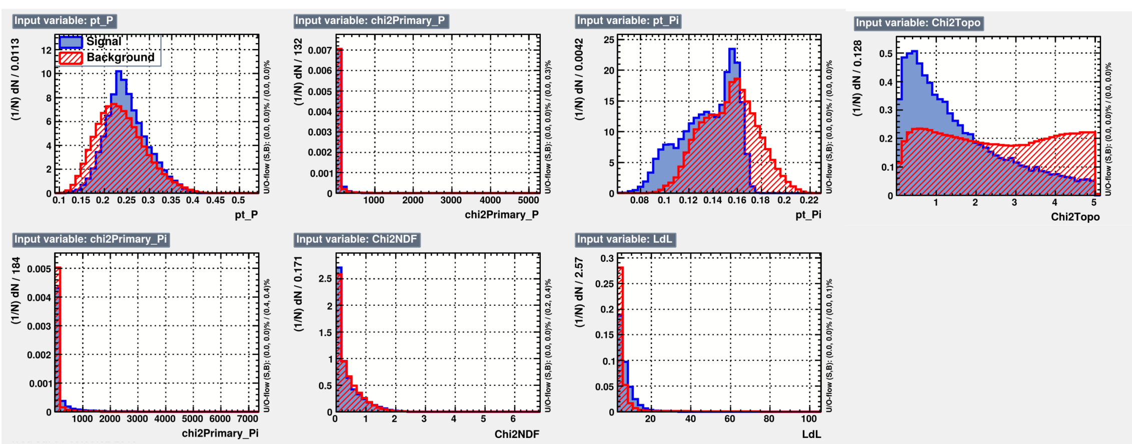

Fig. 5.4: Distributions of training variables from training sample.

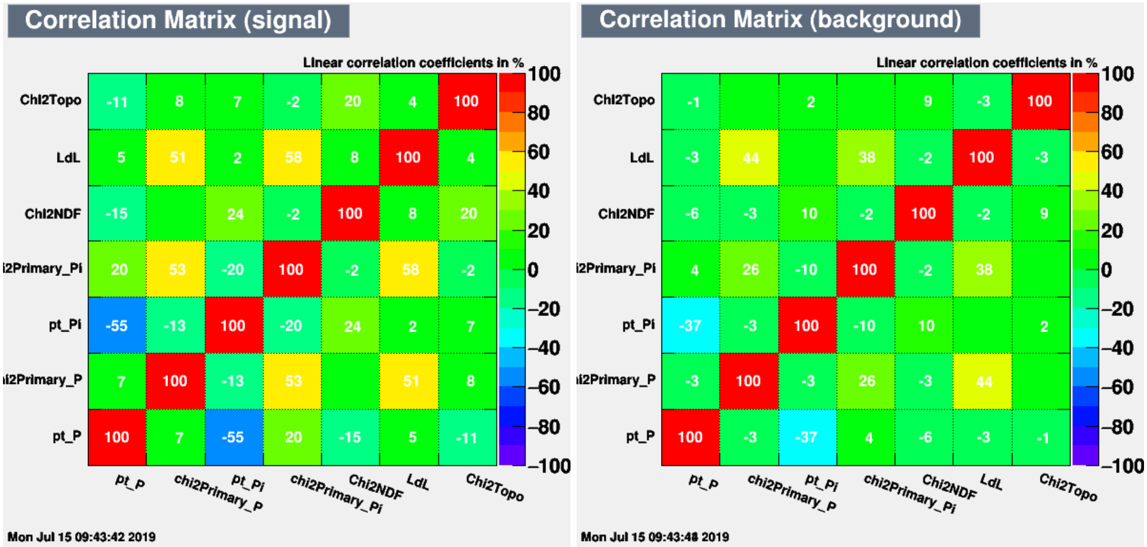

Fig. 5.5: Linear correlation coefficients of variables used in training of BDT for signal (left) and background (right).

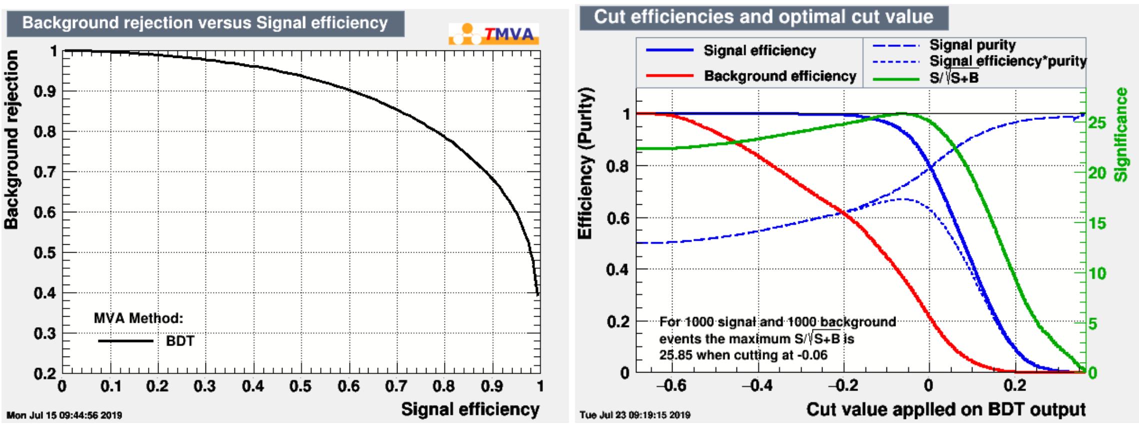

Fig. 5.6: ROC curve characterizing the relation between background rejection and signal efficiency of BDT (left), cut efficiencies and optimal BDT response value cut as calculated by TMVA (right).

The optimal cut on BDT response value (maximizing signal significance) depends on a number of signal and background events. In the plot on the right in Fig. 5.6 the optimal cut value is determined with assumption of number of signal and background events being equal to 1000 by default. Since the ratio of signal to background events supplied by the user to TMVA is artificial, one has to estimate the number of signal and background events in the data, for example in the same way it was done in this analysis with default KFPF cuts if one already has reconstructed a signal peak. The other option is to apply different values of BDT cut on the data, investigate the dependency of significance on the cut value and find the value which maximizes the signal significance. The second approach was selected for this analysis and the BDT were applied on a fraction of 3% of the full statistics to scan the significance. In this way the optimal cut value was found which was later fixed and applied on the whole dataset. This was meant to avoid bringing a bias to the analysis, but since the fraction of the data used for the training and for the significance scan is still included in the final statistics the approach might need more checking.

The plot of the significance scan in transverse momentum bin $p_T = 0.2 - 0.4$ GeV$/c$ is shown in Fig. 5.7 on the left. The right hand side plot displays dependency of signal to background ratio on BDT response value. It can be observed that with further increase of BDT value threshold after reaching maximum significance one can still notably improve the purity of the signal. The significance for $p_T = 0.2 - 0.4$ GeV$/c$ reached maximum with cut BDT $>$ -0.08 and the corresponding mass spectrum with double gaussian fit is plotted in Fig. 5.8. Note that significance per 1M events (34.75) is already higher than with default KFPF selection criteria (22.35). The final cut values for all transverse momentum bins are listed in Tab. 5.4 with the resulting value of signal to background ratio after application of BDT on whole dataset.

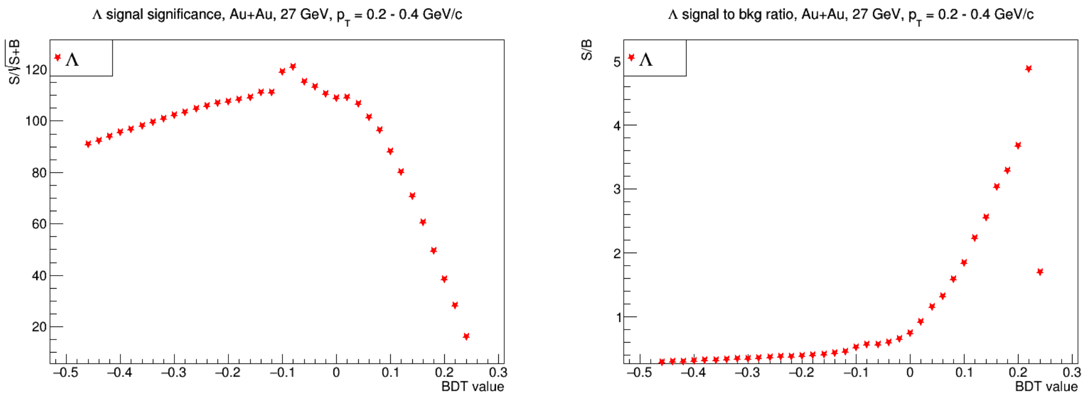

Fig. 5.7: Dependence of signal significance on BDT cut (left), signal to background ratio as a function of BDT cut (right).

p_T [GeV$/c$]	0.1 - 0.2	0.2 - 0.4	0.4 - 0.6	0.6 - 0.8	0.8 - 1.0	1.0 - 1.2
Min. BDT resp. val.	0.005	-0.070	-0.090	-0.110	-0.130	-0.130
Resulting S/B	0.58	0.72	2.00	3.08	3.67	4.87
p_T [GeV$/c$]	1.2 - 1.4	1.4 - 1.6	1.6 - 1.8	1.8 - 2.0	2.0 - 2.3	2.3 - 2.6
Min. BDT resp. val.	-0.110	-0.090	-0.110	-0.090	-0.110	-0.110
Resulting S/B	6.26	8.35	8.27	10.70	9.38	9.76
p_T [GeV$/c$]	2.6 - 3.0	3.0 - 3.4	3.4 - 3.9	3.9 - 4.4	4.4 - 5.0	5.0 - 6.0
Min. BDT resp. val.	-0.090	0.050	0.070	0.020	0.085	0.230
Resulting S/B	11.57	33.35	31.32	10.26	8.55	4.88

Tab. 5.4: BDT response cuts for all transverse momentum bins and resulting signal to background ratio S/B after application of BDT on whole dataset.

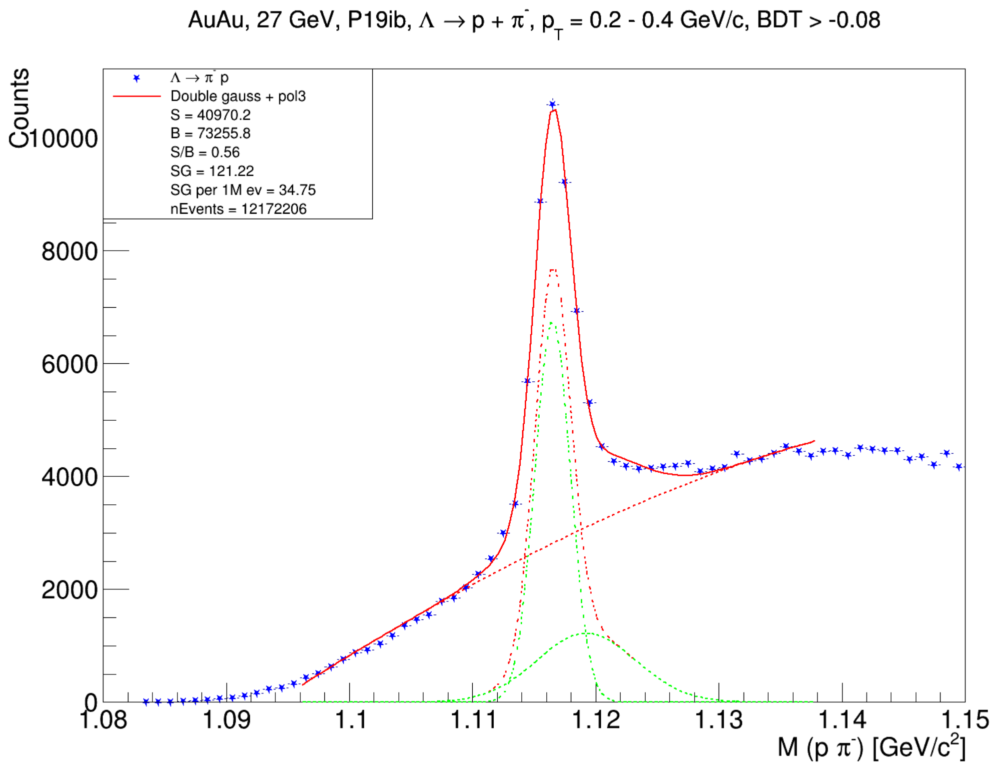

Fig. 5.8: Invariant mass plot for $p_T = 0.1 - 0.2$ GeV$/c$ with highest signal significance obtained by using cut on BDT response value $>$ -0.08.

5.2.4 BDT application

With the cut values found by means explored in previous section BDT were applied directly in KF Particle Finder to analyse the whole dataset. The invariant mass histograms were fitted in the same manner as before and also the yields and signal significance were calculated in the same way. By employing BDT it was now possible to reconstruct signal of Λ in region $p_T = 0.1 - 0.2$ GeV$/c$ as can be seen in Fig. 5.9. In Fig. 5.10 significance obtained with default KFPF cuts is compared to the one achieved with BDT. In general, there is an enhancement of up to 50% at low transverse momentum.

5.2.5 Comparison with conventional analysis

In this section results obtained with KFPF and BDT will be compared to conventional analysis which employs topological cuts to reject combinatorial background. For this purpose Analysis note which presents details of analysis of production of Λ in Au+Au collisions at 27 GeV (and other BES energies) was selected [37]. The analysis was done at STAR and the data were measured during phase one of Beam Energy Scan. There were few differences between the analyses which had to be taken into account in order to make the comparison meaningful. First of all, in this work

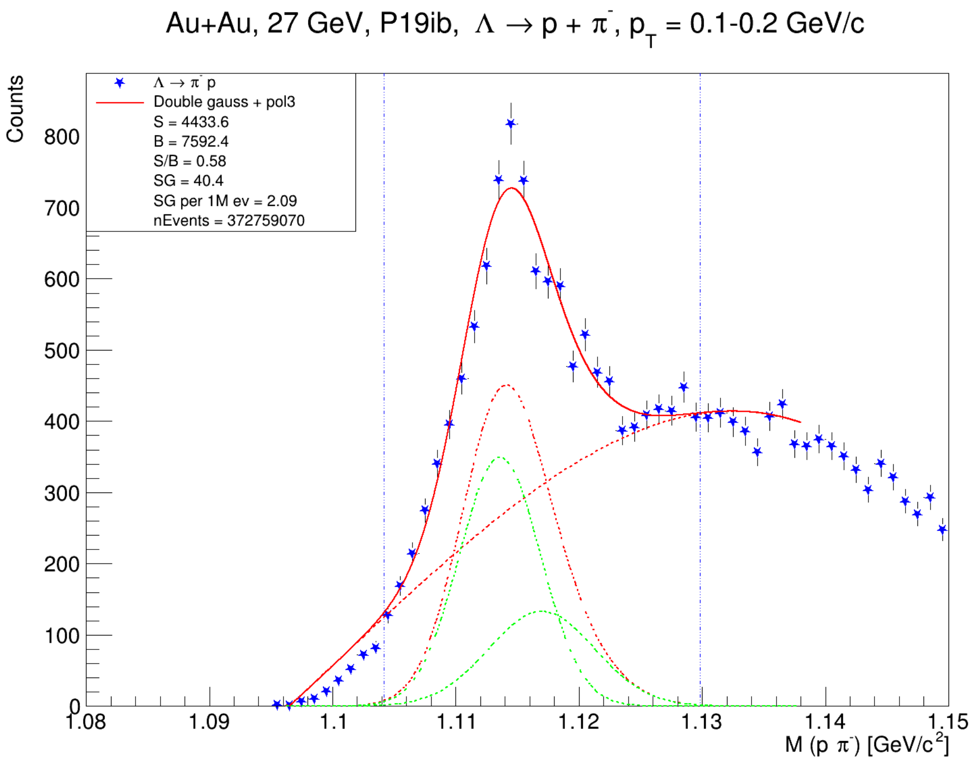

Fig. 5.9: Invariant mass plot for $p_T = 0.1 - 0.2$ GeV$/c$.

no study of centrality was done. Therefore, it was necessary to sum yields from all centrality bins in BES I analysis to obtain an estimate of integrated yield. Second, in BES I paper only Λ particles from mid-rapidity region were discussed, for this reason the same selection, i.e. $|y| < 0.5$, was carried out in this analysis too. Third, to estimate the total error δ of the yield parameter in BES I analysis following formula was applied

$$\delta N = \sqrt{\sum_i \delta N_i^2}, \tag{5.3}$$

where δN_i is the error of yield parameter in i-th bin. The total error was than used to calculate the significance α of the signal in each transverse momentum bin as follows

$$\alpha = \frac{N}{\delta N} \tag{5.4}$$

From the Fig. 5.11 one can observe that apparently the procedure of significance calculation presented above does not work from medium to high transverse momentum. For this reason, the comparison has to be taken with caution. However, the crucial conclusion is that conventional analysis does not reach under $p_T = 0.4$ GeV$/c$ and so it seems probable that even without BDT KF Particle Finder enhances significance in low transverse momentum region.

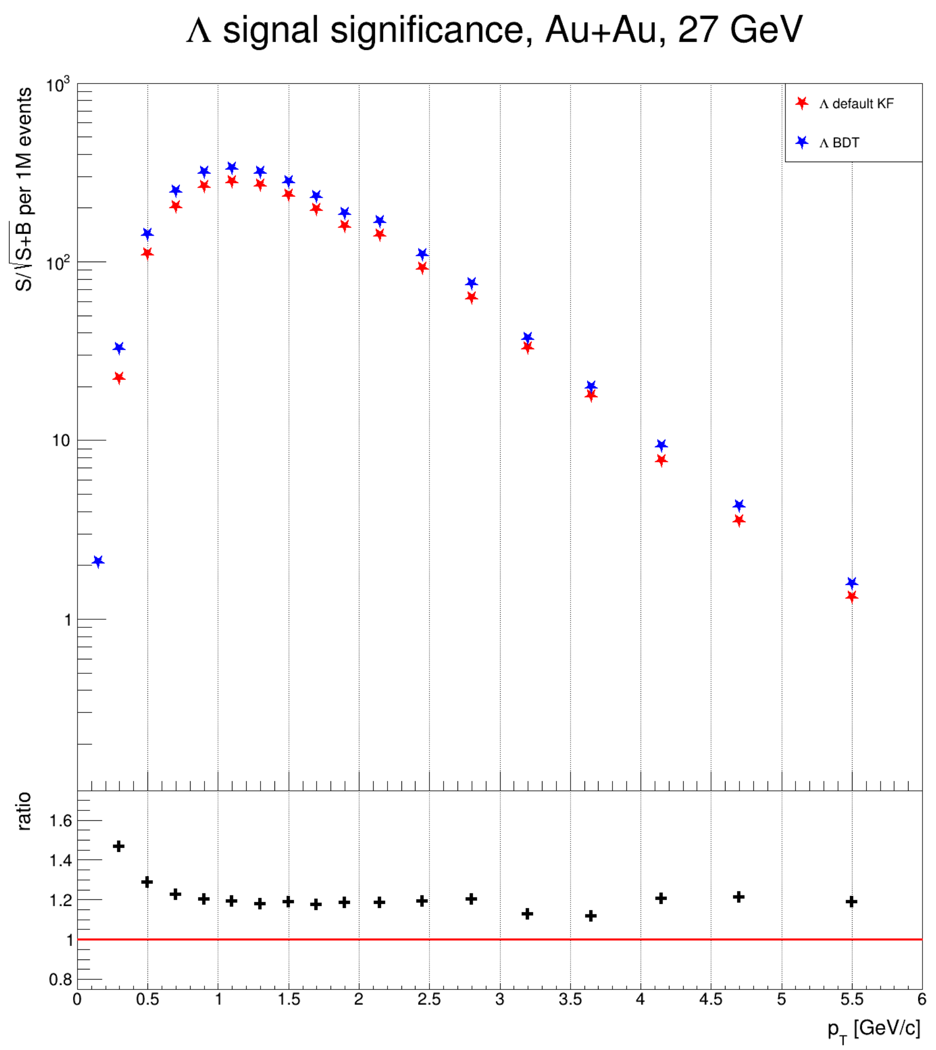

Fig. 5.10: Comparison of significance obtained with application of BDT to the significance achieved with default KFPF cuts.

5.3 Au+Au collisions at $\sqrt{s_{NN}} = 14.6$ GeV

5.3.1 Dataset and event selection

The data from Au+Au collisions at center of mass energy $\sqrt{s_{NN}} = 14.6$ GeV were taken at STAR experiment during run 19 which was aimed to measure at various low energy levels as a part of second phase of the Beam Energy Scan program. In the time of writing this work, the data were not yet produced officially in any form, although they were produced with TFG19 library by TFG (Tracking Focus Group) at STAR, which focuses on express production of measured data in order to give feedback about any tracking or collider issues as soon as possible. These data are than for some time available for express physics analyses in the form of PicoDsts.

The event selection was executed in the same way as in the 27 GeV case, i.e.

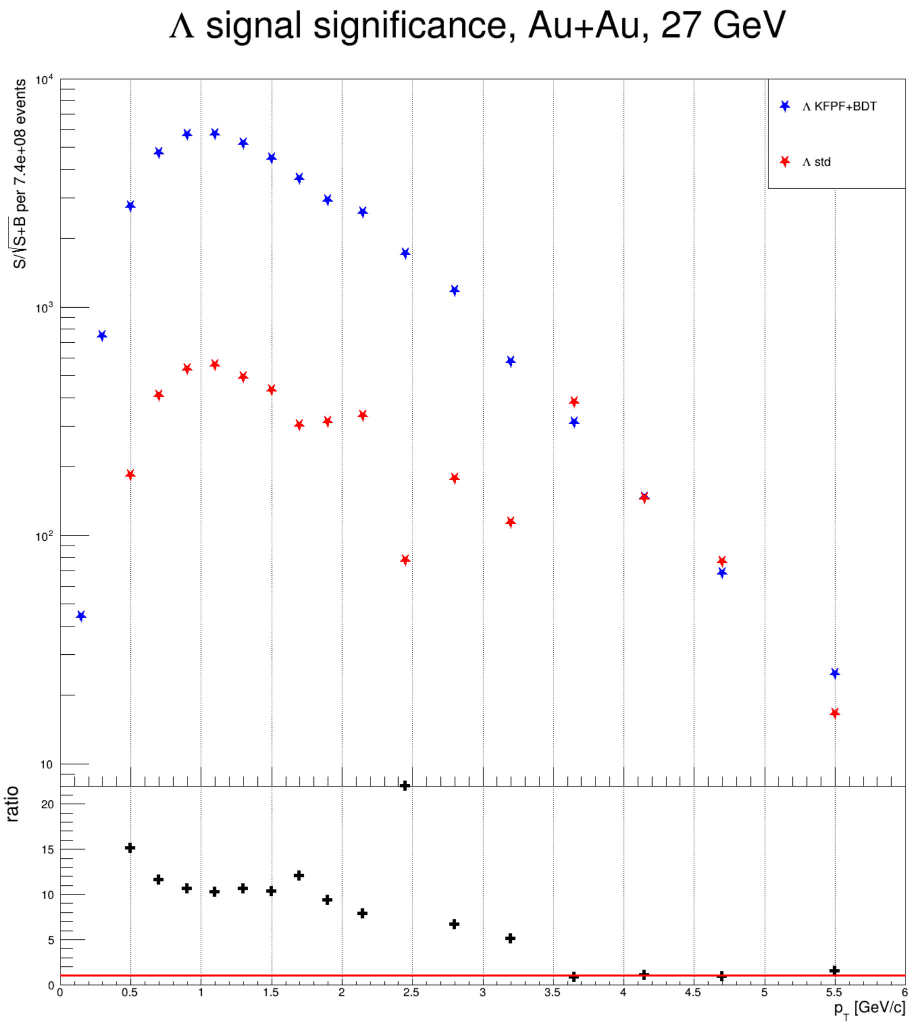

Fig. 5.11: Comparison of significance obtained with application of BDT to the significance achieved with conventional analysis.

selected events had at least 10% tracks classified as primary and satisfied conditions $|V_z| < 70$ cm and $\sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} \leq 0.45$. The total number of sampled minimum bias events counted 220M, from those 200M were selected.

5.3.2 Analysis with default KFPF cuts

In the same way as with the 27 GeV data default KFPF cuts were applied. However, the cut classifying tracks into primary and secondary was changed so that $\chi^2_{prim} > 5$. Also the soft ToF PID mode was turned on in KFPF which means that all hypothesis in 3σ window around predicted time of flight were accepted when identifying daughter particles. The rest of the reconstruction, including signal fitting and yield calculation, was done analogically to the analysis of 27 GeV collisions.

As an example, there are invariant mass histograms from two lowest transverse momentum bins plotted in the Fig. 5.12. The left plot proves that it is possible to reach practically zero transverse momentum even without inclusion of any machine learning methods. Fig. 5.17 presents the resulting uncorrected yields for all transverse momentum bins together with signal significance per 1M events.

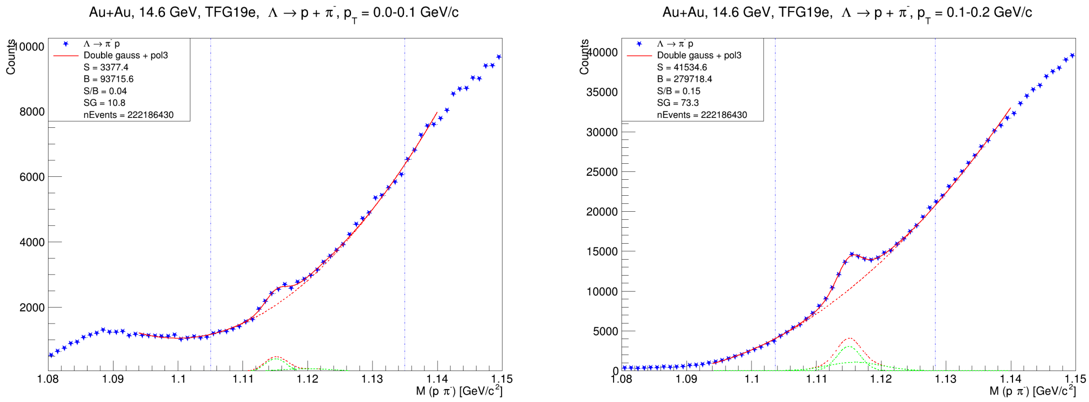

Fig. 5.12: Invariant mass histograms for $p_T = 0.0 - 0.1$ GeV$/c$ (left) and $p_T = 0.1 - 0.2$ GeV$/c$ (right).

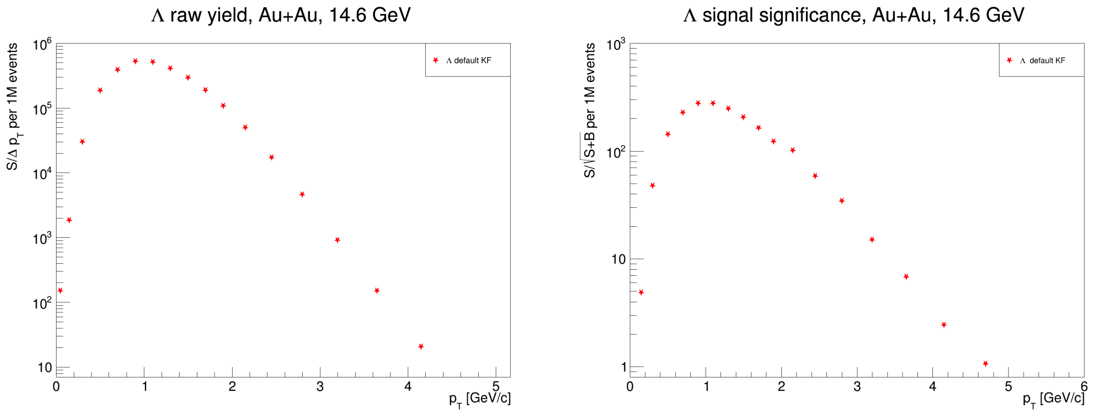

Fig. 5.13: Dependence of uncorrected Λ yield (left) and signal significance (right) on transverse momentum. Both quantities are recalculated per 1M events post selection.

5.4 Fixed target collisions at $\sqrt{s_{NN}} = 3.9$ GeV

5.4.1 Dataset and event selection

Fixed target data from collisions with center of mass energy $\sqrt{s_{NN}} = 3.9$ GeV come also from the express production of 2019 run measurements. In order to collect this data accelerator was basically in the same settings as for 14.6 GeV, but there was only one beam at energy 7.3 GeV which was deflected to hit the fixed gold target and produce a collision at 3.9 GeV. The total number of sampled minimum bias events in this analysis counted 2.8M and all of these were used for reconstruction of Λ baryons.

5.4.2 Analysis with default KFPF cuts

For the reconstruction of Λ particles the default KFPF cuts were employed again with soft ToF PID mode included. The calculation of yield and significance follows the same steps as before, although this time around the dependence on rapidity was explored as well. As an example plot from two transverse momentum bins and two plots from rapidity bins are plotted in Fig. 5.14 and Fig 5.15 respectively.

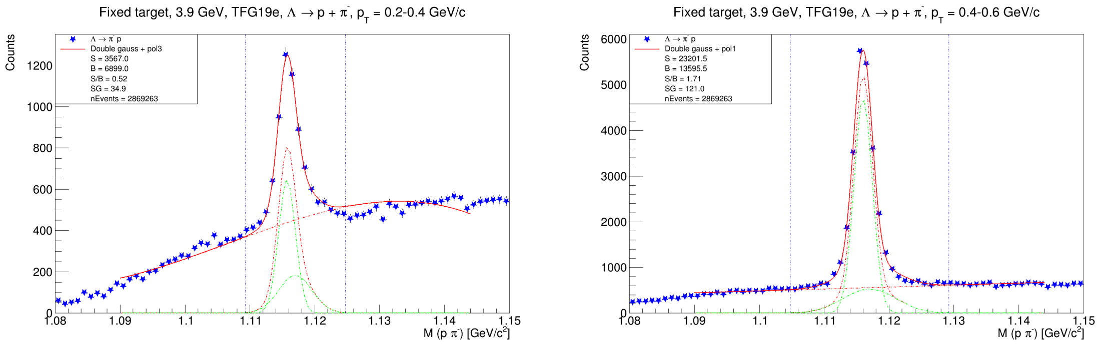

Fig. 5.14: Invariant mass histograms for $p_T = 0.2 - 0.4$ GeV$/c$ (left) and $p_T = 0.4 - 0.6$ GeV$/c$ (right).

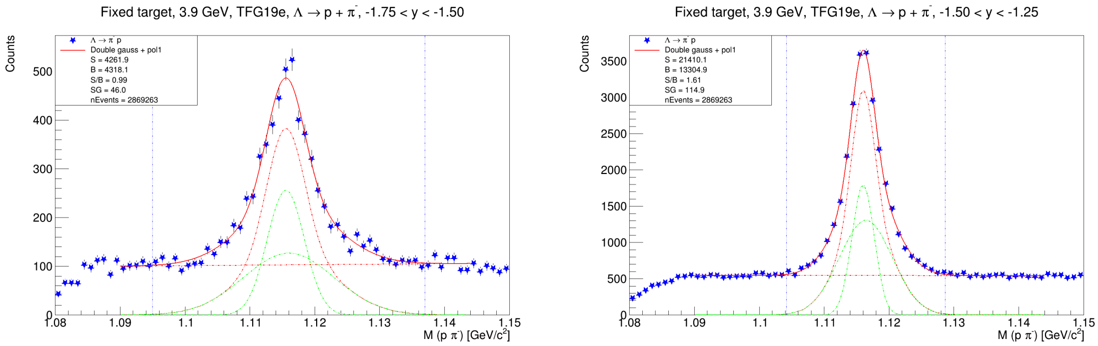

Fig. 5.15: Invariant mass histograms for $-1.75 < y < -1.50$ (left) and $-1.50 < y < -1.25$ (right).

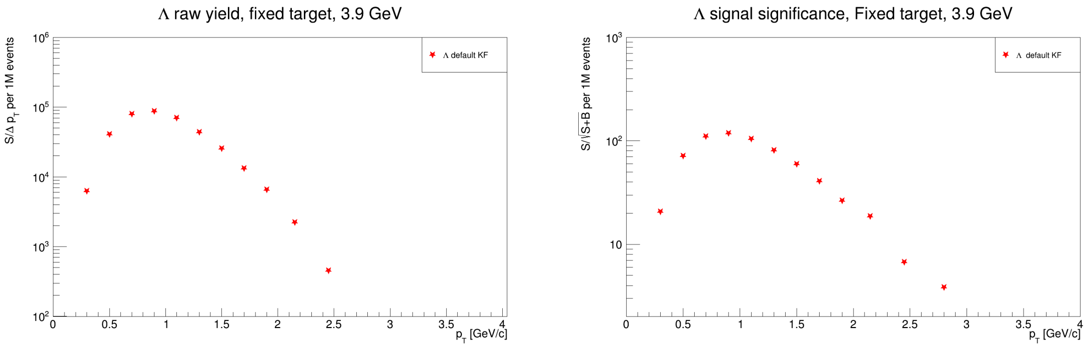

Fig. 5.16: Dependence of uncorrected Λ yield (left) and signal significance (right) on transverse momentum. Both quantities are recalculated per 1M events post selection.

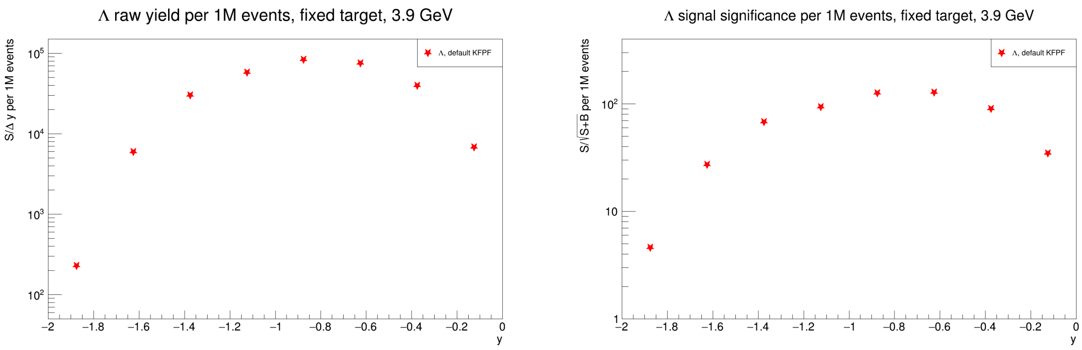

Fig. 5.17: Dependence of uncorrected Λ yield (left) and signal significance (right) on rapidity. Both quantities are recalculated per 1M events post selection.

Chapter 6

Summary

The aim of this research task was to get familiar with techniques of KF Particle Finder and investigate its possibilities in analysing data from Au+Au collisions from STAR, specifically in reconstruction of strange hadrons. In the first chapter experimental setup of Relativistic Heavy Ion Collider (RHIC) and Solenoidal Tracker at RHIC (STAR) was described. The second chapter introduced the Beam Energy Scan program and how it will try to map out the phase diagram of strongly-interacting matter. Also the experimental setup of fixed target mode at STAR was described there and strangeness as a probe of quark-gluon plasma was discussed. In the third chapter KF Particle and KF Particle Finder packages were introduced. In the fourth chapter Toolkit for Multivariate Analysis was described with focus on Boosted Decision Trees. The fifth chapter presented authors analysis of data from Au+Au collisions measured with STAR.

KF Particle Finder is a C++ package based on Kalman Filter mathematics. It allows to reconstruct short-lived particles and currently offers over 70 decay channels for analysis. In contrast to conventional analysis which employs topological cuts in order to reject background, KF Particle works with covariance matrix of tracks which carries full information about track parameters and their uncertainties as measured with Time Projection Chamber. Knowledge of uncertainties in tracking is used to calculate statistical criteria which can be used to estimate quality of reconstruction of both tracks and particles and can be used for efficient rejection of background. These criteria can also be given to any machine learning classifier as input variables for training. The analysis presented in this work investigated possible improvement of signal significance of Λ signal significance in Au+Au collisions at various BES energies with KF Particle Finder, especially at low transverse momentum, since low p_T tracks are significantly curved and the track fit comes with higher uncertainties. For data from collisions at 27 GeV KF Particle Finder is combined with Boosted Decision Trees in order to further improve significance.

It was found that for 27 GeV dataset it is possible with KF Particle to extract signal in transverse momentum bin $p_T = 0.2 - 0.4$ GeV/c which has not been done with conventional approach. Moreover by employing Boosted Decision Trees the analysis showed that one can go as low as $p_T = 0.1 - 0.2$ GeV/c and also significance at higher p_T by approximately 20% and at low p_T up to 50% with respect to default cuts on statistical criteria in KF Particle Finder. There is also significant improvement with respect to previous STAR results obtained by conventional approach to reconstruction.

The 14.6 GeV dataset was also analysed. This dataset comes from the express production of Tracking Focus Group with TFG19 library. The data were collected very recently during ongoing second phase of the Beam Energy Scan program at STAR. Since BES-II introduces iTPC upgarde of Time Projection Chamber to the tracking, it was expected that there will be significant improvement at low p_T. Indeed signal of Λ baryon was reconstructed even in transverse momentum bin $p_T = 0.0 - 0.1$ GeV/c. One could expect further enhancement of signal significance by employing machine learning techniques.

In order to investigate the geometrical independence of KF Particle Finder, data from fixed target mode collisions at $\sqrt{s_{NN}} = 3.9$ GeV/c were analysed as well within the same framework. The dependency of raw yield and signal significance on rapidity and transverse momentum was investigated. With experience obtained in analysis of 27 GeV dataset it was not difficult to switch to fixed target mode.

In the future more decays of strange particles including cascades are planned to be explored in different BES datasets with KF Particle Finder. In analysis of more complicated decays KF Particle may prove to be even more effective with respect to conventional approach. The research into express stream production of data from BES-II could provide useful for determining whether BES-II will have enough statistics to reach its physics goals. Author would also like to investigate possibilities of using machine learning techniques combined with KF Particle Finder in other datasets besides 27 GeV collisions.

Bibliography

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

[2] G. Zweig, "An SU(3) model for strong interaction symmetry and its breaking. Version 1," 1964.

[3] M. J. Tannenbaum, "Highlights from BNL-RHIC," *Subnucl. Ser.*, vol. 49, pp. 295–342, 2013.

[4] T. Nayak and B. Sinha, "Search and study of Quark Gluon Plasma at the CERN-LHC," pp. 131–144, 2009.

[5] P. Senger, "The CBM experiment at FAIR," *J. Phys. Conf. Ser.*, vol. 50, pp. 357–360, 2006.

[6] S. Gorbunov, "On-line reconstruction algorithms for the CBM and ALICE experiments," Ph.D. dissertation, Johann Wolfgang Goethe-Universität, 2013.

[7] M. Zyzak, "Online selection of short-lived particles on many-core computer architectures in the CBM experiment at FAIR," Ph.D. dissertation, Johann Wolfgang Goethe-Universität, 2015.

[8] M. Harrison, T. Ludlam, and S. Ozaki, "RHIC project overview," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 499, no. 2-3, pp. 235–244, 2003.

[9] A. Adare *et al.*, "An Upgrade Proposal from the PHENIX Collaboration," PHENIX Collaboration, Tech. Rep., 2015.

[10] D. Tlusty, "The RHIC Beam Energy Scan Phase II: Physics and Upgrades," October 2018, https://arxiv.org/pdf/1810.04767.pdf.

[11] W. Fischer, A. Drees, M. Blaskiewicz, A. Fedotov, H. Huang, C. Liu, M. Minty, and D. Raparia, "RHIC Collider Projections (FY 2019 – FY 2027)." [Online]. Available: https://www.rhichome.bnl.gov/RHIC/Runs/RhicProjections.pdf

[12] K. C. Meehan, "Fixed Target Collisions at STAR," *Nuclear Physics A*, vol. 956, no. 1, pp. 878–881, 2016.

[13] "Relativistic Heavy Ion Collider." [Online]. Available: https://drupal.star.bnl.gov/STAR/public/img/SketchUpSTAR

[14] M. Beddo *et al.*, "STAR: Conceptual design report for the Solenoidal Tracker

at RHIC," STAR Collaboration, Tech. Rep., 1992.

[15] H. Qui, "STAR heavy flavor tracker," *Nuclear Physics A*, vol. 931, no. 1, pp. 1141–1146, 2014.

[16] Y. Wang, "STAR upgrade program and future physics," *J. Phys. Conf. Ser.*, vol. 535, pp. 12–22, 2014.

[17] "STAR Detector System." [Online]. Available: https://drupal.star.bnl.gov/STAR/future

[18] M. Anderson, J. Berkovitz, W. Bettes *et al.*, "The STAR time projection chamber," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 499, no. 2-3, pp. 659–678, 2003.

[19] S. Collaboration, "Technical Design Report for the iTPC Upgrade," STAR Collaboration, Tech. Rep., 2015, [Online]. Available at: https://www.star.bnl.gov/public/tpc/tpc.html.

[20] P. Fachini, Z. Xu *et al.*, "Proposal for a Large Area Time of Flight System for STAR: STAR-TOF," STAR Collaboration, Tech. Rep., 2004, [Online]. Available at: https://www.star.bnl.gov/public/tof/.

[21] A. Bazavov, T. Bhattacharya, M. Cheng, C. DeTar, H.-T. Ding, S. Gottlieb, R. Gupta, P. Hegde, U. M. Heller, F. Karsch, E. Laermann, L. Levkova, S. Mukherjee, P. Petreczky, C. Schmidt, R. A. Soltz, W. Soeldner, R. Sugar, D. Toussaint, W. Unger, and P. Vranas, "Chiral and deconfinement aspects of the qcd transition," *Phys. Rev. D*, vol. 85, p. 054503, Mar 2012. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.85.054503

[22] S. Ejiri, "Canonical partition function and finite density phase transition in lattice QCD," *Phys. Rev.*, vol. D78, p. 074507, 2008.

[23] G. Odyniec, "The RHIC Beam Energy Scan program in STAR and what's next ..." *J. Phys. Conf. Ser.*, vol. 455, p. 012037, 2013.

[24] C. Yang, "The STAR beam energy scan phase II physics and upgrades," *Nucl. Phys.*, vol. A967, pp. 800–803, 2017.

[25] K. C. Meehan, "The fixed-target experiment at STAR," *J. Phys. Conf. Ser.*, vol. 742, no. 1, p. 012022, 2016.

[26] J. Rafelski and B. Müller, "Strangeness production in the quark-gluon plasma," *Phys. Rev. Lett.*, vol. 48, pp. 1066–1069, Apr 1982. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.48.1066

[27] J. Rafelski and J. Letessier, "Strangeness and quark gluon plasma," *J. Phys.*, vol. G30, pp. S1–S28, 2004.

[28] L. Adamczyk *et al.*, "Probing parton dynamics of QCD matter with Ω and ϕ production," *Phys. Rev.*, vol. C93, no. 2, p. 021903, 2016.

[29] T. O. Ablyazimov, M. V. Zyzak, V. V. Ivanov, and P. I. Kisel, "Kalman filter-based fast track reconstruction for charged particles in a Compressed Bary-

onic Matter experiment using parallel computing on a multicore server at the Laboratory of Information Technologies, Joint Institute for Nuclear Research," *Physics of Particles and Nuclei Letters*, vol. 12, no. 3, pp. 423–427, 2015.

[30] I. Kisel, G. Adam, J. Buša, and M. Hnatič, "New Approaches for Data Reconstruction and Analysis in the CBM Experiment," in *EPJ Web of Conferences*, 2016.

[31] T. O. Ablyazimov, M. V. Zyzak, V. V. Ivanov, and P. I. Kisel, "Kalman filter-based fast track reconstruction for charged particles in a Compressed Baryonic Matter experiment using parallel computing on a multicore server at the Laboratory of Information Technologies, Joint Institute for Nuclear Research," *Phys. Part. Nucl. Lett.*, vol. 12, no. 3, pp. 423–427, 2015.

[32] K. Albertsson *et al.*, "Machine Learning in High Energy Physics Community White Paper," *J. Phys. Conf. Ser.*, vol. 1085, no. 2, 2018.

[33] K. Albertsson, "TMVA Summary," 2018. [Online]. Available: https://drupal.star.bnl.gov/STAR/future

[34] S. Chatrchyan *et al.*, "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC," *Phys. Lett.*, vol. B716, pp. 30–61, 2012.

[35] A. Hocker *et al.*, "TMVA - Toolkit for Multivariate Data Analysis," 2007.

[36] M. Tanabashi *et al.*, "Review of Particle Physics," *Phys. Rev.*, vol. D98, no. 3, p. 030001, 2018.

[37] X. Zhu, "Recent results on strangeness production at STAR," *J. Phys. Conf. Ser.*, vol. 509, p. 012004, 2014.