CZECH TECHNICAL UNIVERSITY IN PRAGUE

Faculty of Nuclear Sciences and Physical Engineering

Department of Physics

Research project

Analysis of b-jets in pp collisions at 8 TeV

Bc. Lukáš Kramárik

Supervisor: Mgr. Jaroslav Bielčík, Ph.D.

Consultant: Gyulnara Eyyubova, Ph.D.

Prague, 2015

Prehlásenie

Prehlasujem, že som svoju výskumnú úlohu vypracoval samostatne a použil som len literatúru a publikácie uvedené v priloženom zozname.

Nemám závažný dôvod proti použitiu tohoto školského diela v zmysle §60 Zákona č.121/1200Sb., o autorskom práve, o právach súvisiacich s autorským právom a o zmene niektorých zákonov (autorský zákon).

V Prahe dňa

Názov práce: **Analýza b-jetov v pp zrážkach pri 8 TeV**
Autor: Lukáš Kramárik

Odbor: Jaderné inženýrství
Druh práce: Výzkumný úkol

Vedúci práce: Mgr. Jaroslav Bielčík, Ph.D.
Katedra fyziky, Fakulta jaderná a fyzikálně inženýrská,
České vysoké učení technické v Praze.

Konzultant: Gyulnara Eyyubova, Ph.D.
Katedra fyziky, Fakulta jaderná a fyzikálně inženýrská,
České vysoké učení technické v Praze.

Abstrakt:

Kvarkovo-gluónová plazma je stav horúcej a hustej jadrovej hmoty, ktorého existencia sa predpokladá tesne po Veľkom Tresku a ktorý môže byť vytvorený pri zrážkach jadier ťažkých iónov. Jedným z detektorov, schopných detegovať takéto zrážky, je ALICE na urýchľovači LHC v CERNe. Kvarkovo-gluónová plazma môže byť študovaná pomocou partónov, ktoré prechádzali touto hmotou. Tieto hadronizujú a vytvárajú jety, kolimované spŕšky hadrónov. Táto práca sa sústredí na jety pochádzajúce z b quarku, b-jety. V tejto práci sú ukázané vlastnosti b-jetov v pp zrážkach a prebiehajúce štúdie b-jetov v p-Pb zrážkach na experimente ALICE.

Kľúčové slová: kvarkovo-gluónová plazma, jet, b-jet, ALICE, LHC

Title: **Analysis of b-jets in pp collisions at 8 TeV**
Author: Lukáš Kramárik

Abstract:

Quark-gluon plasma is a state of hot and dense nuclear matter, which probably existed right after Big Bang and which could be created in collisions of heavy nuclei. One of detectors able to detect these collisions is ALICE at LHC in CERN. Quark-gluon plasma can be studied by observation of partons, which were passing this matter. These can hadronize to form jets, collimated shower of hadrons. The aim of this work is study of jets originating from b quark, b-jets. In this work, b-jet properties in pp collisions are shown, as well as ongoing study of b-jets in p-Pb collisions at ALICE experiment.

Keywords: quark-gluon plasma, jet, b-jet, ALICE, LHC

Acknowledgements

I would like to give thanks to Mgr. Jaroslav Bielčík, Ph.D. and to Gyulnara Eyyubova, Ph.D. for their advices, corrections and patience during not only typing this work, but also during supervision of my analysis.

Contents

Introduction

Quark-gluon plasma (QGP) is new state of hot and dense nuclear matter, that most likely existed shortly after Bing Bang. In this matter, quarks are deconfined and quasi free. The properties of QGP are studied in laboratories with experiments, that are able to detect collisions of heavy ions. These experiments are e.g. STAR at RHIC (Relativistic Heavy Ion Collider) in BNL (Brookhaven National Laboratory), ATLAS (A Toroidal LHC ApparatuS), CMS (Compact Muon Solenoid) and ALICE (A Large Ion Collider Experiment) at LHC (Large Hadron Collider) at CERN (Conseil Européen pour la Recherche Nucléaire). These experiments have large volume of detectors, that are constructed to detect large number of particles created in collisions of heavy ions very quickly and with the highest possible efficiency. In this work, I will focus mostly on ALICE experiment, that study proton-proton (pp) collisions, collisions of lead ions (Pb-Pb) and collisions of protons with lead ions (p-Pb).

Studies of b-jets are presented in this thesis. Via their measurements, b-quark production can be determined. Studies of b-jets in heavy-ion collisions could also investigate color and mass dependence of parton energy loss in the quark-gluon plasma.

The main goal of this thesis is the analysis of properties of b-jets in pp collisions at $\sqrt{s} = 7$ TeV. This study compares results from ALICE simulations with results from other experiments, it also tests properties of jet finding and b-tagging algorithms. I also made a study of b-jets in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV.

In first chapter, short introduction in the world of elementary particles is shown. Standard Model is described, as well as quark-gluon plasma and its main properties. Phase diagram of nuclear matter is also shown.

ALICE detector is described in second chapter. This detector consist of 18 subdetectors, only three of them are described in more detail: ITS (Inner Tracking System), TPC (Time Projection Chamber) and EMCal (ElectroMagnetic Calorimeter).

In next chapter, jets, collimated sprays of hadrons, are described. It is focused mainly on jets in heavy ions collisions, because these are investigated in my analysis. Also the most used jet-finding algorithm, anti-k_T is described. At the end of this chapter, two heavy flavor tagging algorithms, which are used at ALICE, are presented.

In fourth chapter, main steps of b-jet analysis from CMS experiment is described. This analysis is described, since CMS first measured b-jets in heavy ions collisions. At ALICE experiment, similar analysis is currently beeing studied.

Last two chapters shows results of analysis of b-jets in pp and p-Pb collisions. First analysis is focused on properties of b-jet tagging and jet-finding algorithms in pp collisions. Results from Pythia simulations are in agreement with results from CMS experiments, and results of jet properties are reasonable.

Finally, steps of actually running analysis of p-Pb data at ALICE are shown. We focus mainly on properties of b-jets and its secondary vertices. We are trying to fit distributions of different properties of these vertices in simulated data, and to use results of this fit in data. Fits by functions were tested, as well as template fits of data.

Chapter 1

Introduction to quark-gluon plasma studies

1.1 Natural units

In the field of elementary nuclear and particle physics, natural units are used. It is based on universal physical constants, such as speed of light c, elementary charge e, reduced Planck constant $\hbar$ and Boltzmann constant k_B. In this system, these constant are fixed: $\hbar = c = k_B = 1$. All physical quantities are expressed in terms of energy E. Quantities, with appropriate natural units and with conversion to SI units system are in Table 1.1. Usually, constants $c, \hbar$ and k_B are not written in units.

Quantity	**Natural units**	**Conversion**	
energy	E	1 eV	$= 1.602\ 177 \cdot 10^{-19}$ J
momentum	E	1 eV$/c$	$= 5.344\ 286 \cdot 10^{-28}$ kg·m/s
mass	E	1 eV$/c^2$	$= 1.782\ 662 \cdot 10^{-36}$ kg
temperature	E	1 eV$/k_B$	$= 11\ 604.522\ 1(67)$ K
time	$1/E$	1 $\hbar$/eV	$= 6.582\ 119 \cdot 10^{-16}$ s
length	$1/E$	1 $\hbar c$/eV	$= 1.973\ 27 \cdot 10^{-7}$ m
velocity	none	1	$= c = 2.997\ 924 \cdot 10^{8}$ m/s

Table 1.1: Natural units of different quantities and their conversion to SI units. Data taken from Ref. [1].

1.2 Standard Model

Currently the most used theory to study elementary particles and interactions is the Standard Model. It was theoretically predicted in 1970s and it was already experimentally confirmed. This theory includes elementary particles classified in three groups (leptons: electron, muon, tauon and corresponding neutrinos; bosons: W^+, W^-, Z^0, gluon, photon, Higgs boson; and quarks: u, d, s, c, b, t), elementary antiparticles and three fundamental forces. These forces are weak interaction, strong interaction and electromagnetic interaction. Each of these forces is intermediated by elementary particles called bosons.

One of boson is photon and it is exchange particle in electromagnetic interaction.

Photon is particle with no mass or charge, so it propagates freely and photons do not interact between each other. This is why the range of electromagnetic interaction is theoretically infinite. Every particle could have positive, negative or zero charge. Charged particles are influenced by magnetic field, it can be observed as curvature of their trajectories. Photon is a energy quantum, that can be radiated in different processes, where particles loss energy or dramatically changes their flight directions (Compton scattering, Bremsstrahlung).

In case of weak interaction, intermediate particles are W^+, W^- and Z^0 bosons. Because of their relatively large masses ($M_{W^\pm} = 80.385 \pm 0.015$ GeV/c^2 [2] and $M_{Z^0} = 91.1876 \pm 0.0021$ GeV/c^2 [2]), the range of this interactions is quite short, approximately 1000 times smaller than dimension of nucleus ($\sim$1 fm). Actually, in case of low-energy collision, their range is considered as negligible. In interactions, where $W^\pm$ are participated, the particle transformations can occur. However, these transformations should conserve number of leptons and also other symmetries. Examples of processes, where one lepton is transformed to another, are β-decays.

Another interaction described by Standard Model is the strong interaction. This interaction is mediated by gluons, carrying so called color charge. Other elementary particles, that have this color charge are quarks. Three different color charges of quarks are red, green and blue, for antiquarks it is antired, antigreen and antiblue. As gluons connect quarks, they can have different combinations of color and anticolor (for example red-antired, green-antiblue). One of consequence of gluon color charge is that gluons interact between each other.

One of the most important property of strong interaction is asymptotic freedom. When two quarks are binded by some gluons, energy of this binding rises with their mutual distance. For small distance, quarks are quasi free. If energy is provided to this binding, quarks are drawn appart from each other. If the provided energy rises, in some moment it is more energetically profitable to create new pair of quark and antiquark from vacuum. Matter, in which this dependency occurs, is called confined nuclear matter. However, in case of very large energy density or temperature ($\approx$170 MeV), this binding is "melted" and quarks and gluons are free. This kind of matter is called deconfined nuclear matter, and example of it is quark-gluon plasma.

In Standard Model, six different quarks are described: u, d, s, c, b, t. Their main properties are in Table 1.2. Standard Model also describe antiquarks ($\bar{u}$, $\bar{d}$, $\bar{s}\ldots$), particles of antimatter, thus have same mass, but opposite charge, quantum spin, baryon number and other properties as quarks that compose matter.

q	m [MeV/c^2]	Q [e]
u	$2.3^{+0.7}_{-0.5}$	$^2/_3$
d	$4.8^{+0.5}_{-0.3}$	-$^1/_3$
s	95 ± 5	-$^1/_3$
c	1275 ± 25	$^2/_3$
b	4180 ± 30	-$^1/_3$
t	173210 ± 710	$^2/_3$

Table 1.2: Main properties of quarks q, their mass m and electric charge Q. Data taken from Ref. [2].

In confined nuclear matter, quarks are binded by gluons to form different hadrons. All hadrons are composed in such way, that they have no color charge (so they are "white"). In case of mesons, hadrons composed from two quarks, this is done by binding of quark with color and antiquark with corresponding anticolor. In case of baryons, composed from three quarks, these should have red, green and blue color charge. For antibaryons, quarks are replaced by antiquarks and color is replaced by anticolor, so antiquarks should have antired, antigreen and antiblue color charge.

Another group of elementary particles in Standard Model are leptons: electron e^- , muon μ^-, tau particle τ^- and three corresponding neutrinos (ν_e, ν_μ, ν_τ). As in other cases, for every particle there is antiparticle. Leptons do not carry color charge, so they are not influenced by strong interaction. Every lepton family (lepton and corresponding neutrino) have its own lepton number, that should be conserved in all typer of interaction.

$\boldsymbol{l}$	$\boldsymbol{m}$ **[MeV/$\boldsymbol{c^2}$]**	$\boldsymbol{Q}$ **[e]**
e^-	$0.510998928 \pm 0.000000011$	-1
ν_e	<2	0
μ^-	$105.6583715 \pm 0.0000035$	-1
ν_μ	<2	0
τ^-	1776.82 ± 0.16	-1
ν_τ	<2	0

Table 1.3: Main properties of leptons l, their mass m and electric charge Q. Data taken from Ref. [2].

The last particle from Standard Model, experimentally approved in 2012, is Higgs boson. Its mass is $M_{H^0} = 125.7 \pm 0.4$ GeV$/c^2$. It composes Higgs field, in consequence of which $W^\pm$ and Z gain mass. This explains differences of masses of $W^\pm$, Z bosons and photon. Higgs boson is massive scalar particle. Its mass is free parameter in model and it can give a hint on relevance of Standard Model theory or another theories, such as Supersymmetry.

Despite the fact that Standard Model describes confinement of quarks in mesons or baryons, particles composed from 5 quarks and antiquarks, pentaquarks, have been observed. The last observation is from LHCb experiment in CERN. They were observed in decays of Λ_b^0 baryons (more details in Ref. [3]). Pentaquarks named $\mathrm{P_c}(4450)^+$ and $\mathrm{P_c}(4380)^+$ were intermadiate states in these decays observed in combinations of spectra of final products, J/ψ, proton and kaon (Fig. 1.1). These pentaquarks are thus formed of d, c, $\bar{c}$ and two u quarks. Quarks in pentaquarks can be tightly bound, or they can form a meson-baryon molecule. More studies are needed to distinguish between these two options.

One of problems in Standard Model is, that it does not describe gravitational force. Between two elementary particles, this one is too small to be currently measured. It is expected to be intermediated by particle called graviton. Graviton is probably massless particle, propagating by speed of light (same as photon) with spin 2.

1.3 Quark-gluon plasma

As already mentioned, example of deconfined nuclear matter is quark-gluon plasma. It is a new state of hot and dense nuclear matter, that is expected to exist during Big Bang. This matter could be actually created in accelerator experiments, by colliding nuclei of heavy ions and thus creating large temperatures and densities of matter.

Astrophysical arguments for Big Bang, and maybe presence of QGP are:

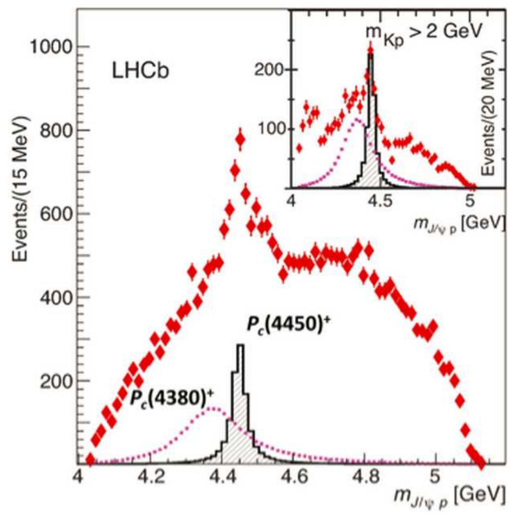


Figure 1.1: The mass spectrum of J/ψ and proton combinations from $\Lambda_b^0 \to J/\psi K^- p$ decays. The data are shown as red diamonds. The predicted contributions from the $P_c(4450)^+$ and $P_c(4380)^+$ states are indicated in the purple and black distributions, respectively. Inset: the mass of J/ψ and proton combinations for a restricted range of the Kaon and proton mass, where the contribution of the wider $P_c(4380)^+$ state is more pronounced. Taken from Ref. [3].

- Cosmic Microwave Background (CMB). It is isotropic radiation in universe, and it corresponds to black body spectrum of temperature 2.73 K. This confirms theory of creation of atoms and hadronisation after Big Bang. It was observed in 1965 by A. Penzias and R. Wilson and they received Nobel Prize for this in 1978.
- Observed ratio of primordial helium to total mass of baryons, 0.25. This is matching to theoretical value from theory of primordial fusion of nuclei.
- Hubble's law (1929). It describes expansion of universe from observations of mutual distances between Earth and galaxies.

Nuclear matter have its own phase diagram, that can be seen in Fig. 1.2. It shows state of matter for different temperatures and barychemical potential. Barychemical potential μ_B is energy needed to add one baryon to system. As we can see, confined nuclear matter (hadrons) exists for temperatures below 170 MeV for $\mu_B = 0$ MeV. For barychemical potential bigger then 1200 MeV and temperatures down to approximately 100 MeV, there exists state called color superconductor.

There exist two different phase transition from quark-gluon plasma to hadrons. For low $\mu_B < 350$ MeV (dashed line in Fig. 1.2) it is cross over transition. It is rather fast change of states, that can not be described by some derivations of thermodynamical variables. For higher μ_B, there is first order transition. Between cross over transition and first order transition, for 200 MeV$< \mu_B <$ 500 MeV, there is critical point (E in Fig. 1.2). Search of this point is one of goals for heavy-ions physics, since its precise position in phase diagram is still unknown.

In the evolution of the heavy-ion collisions, different phases are distinguished. Schematic description of this evolution is shown in Fig. 1.3 In first phase, there are many inelastic collisions between nuclei. After this phases, system become nearly stable, and in this state

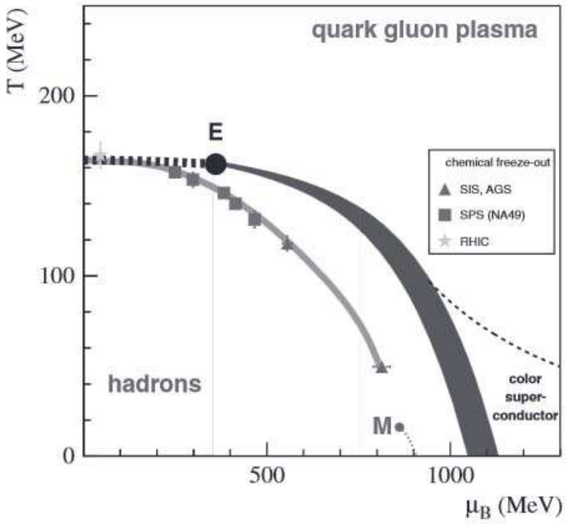


Figure 1.2: Phase diagram of nuclear matter, in space of baryochemical potential μ_B and temperature T. Solid line shows transition betweeen partons and hadrons and it ends in critical point E, dashed line shows cross-over transition. Bottom solid line shows barychemical freeze-out.Taken from Ref. [4].

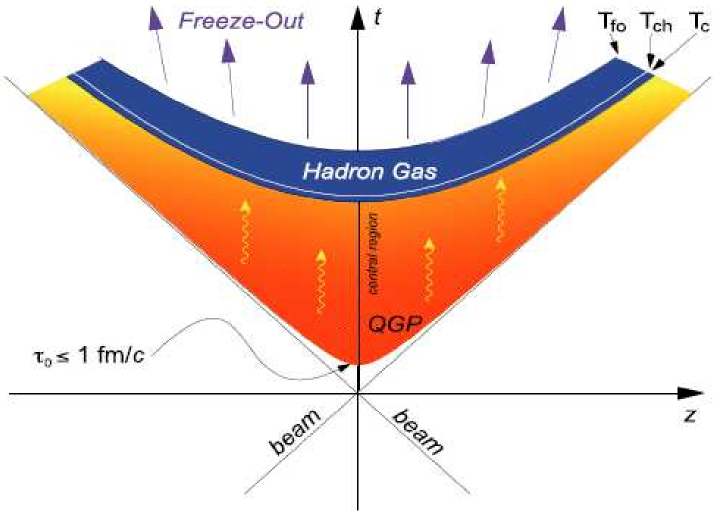


Figure 1.3: Description of heavy-ion collisions in one space (z) and one time (t) dimension, light cone. Showed the evolution of these collisions: critical temperature of phase transition T_C, temperature of hadrochemical freeze-out T_{ch} and temperature of thermal freeze-out T_{fo}. Taken from Ref. [4].

quark-gluon plasma could exist. Because of expansion of the system, temperature is decreasing. When it drops to critical temperature T_C, confined nuclear matter start to be formed in process called hadronisation. For $\mu_B = 0$ MeV, this temperature is expected to be around 170 MeV. When temperature drops below temperature of hadrochemical freeze-out T_{ch}, hadron gas is present. In this medium, inelastic collisions still occur. System continues in its expansion and below some T_{fo}, there are no inelastic collisions betweeen hadrons. This is point of thermal freeze-out, and for $\mu_B = 0$ MeV it occurs at temperature close to critical temperature. For higher μ_B, it occurs at temperature around 10-20 MeV smaller than critical temperature.

Chapter 2

ALICE detector

ALICE (A Large Ion Collider Experiment, Fig. 2.1) is one of the four experiments situated at LHC (Large Hadron Collider) at CERN (Conseil Européen pour la Recherche Nucléaire). Other experiments at LHC are ATLAS (A Toroidal LHC ApparatuS), CMS (Compact Muon Solenoid) and LHCb.

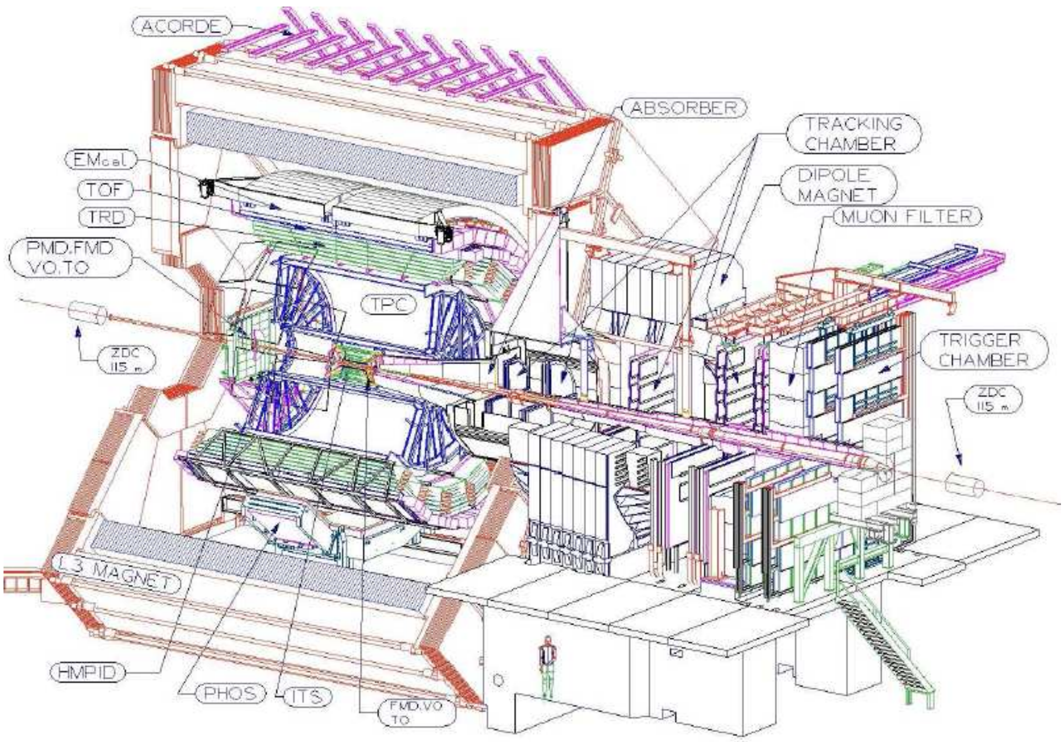


Figure 2.1: General schematic view of ALICE detector and its components. Taken from Ref. [5].

LHC is 27 km long circuit, that works in different phases. Phase called Run is the working phase, when collisions and measurements occur. During phase called Long Shutdown (LS) there are no collisions, and LHC is being upgraded or repaired. In Run 1 (2010-2013), there were collisions of protons (pp) at maximal energies of collisions $\sqrt{s} = 7$ TeV[1], Pb nuclei (Pb-Pb) at maximum energy per colliding nucleon pair $\sqrt{s_{NN}} = 2.76$ TeV

[1] $\sqrt{s}$ is total energy in CMS of colliding particles, it is also invariant mass of the CMS. $\sqrt{s}$ =

and protons with Pb nuclei (p-Pb) at $\sqrt{s_{NN}} = 5.02$ TeV. In Run 2, that begun in 2015, collision systems are the same and energies are planned to be approximately double for heavy ions and $\sqrt{s} = 13$ TeV for pp collisions.

The goal of ALICE detector (Fig. 2.1) is to study different phases of nuclear matter, such as hot nuclear matter (QGP) or cold nuclear matter, and the phase transition between QGP and hadronic matter. This detector consists of 18 different subdetector systems, that are specialized for detection of low energy particles and jets with very high energy, momentum and space resolution. Central barrel of detector is enclosed in solenoid. Detectors in central barrel, from closest to the beam pipe to furthest one are Inner Tracking System (ITS), Time-Projection Chamber (TPC), Time-of-Flight (TOF), Ring Imaging Cherenkov (RICH), High Momentum Particle IDentification (HMPID), Transition Radiation Detector (TRD), ElectroMagnetic Calorimeter (EMCal) and PHOton Spectrometer (PHOS). In a forward beam direction there are systems for muon detection, partially enclosed in dipole magnet. Zero Degree Calorimeter (ZDC) measures amount of nucleons, that did not participate in collisions and it is situated 116 from interaction point. Some of subdetectors, that are the most important for our studies, will be introduced, as they were designed before LS1 upgrades. Data and informations in this chapter were taken from Ref. [5].

2.1 Inner Tracking System (ITS)

ITS is used to detect position of primary vertex (places, where collision occured), secondary vertices (places, where heavy hadrons decayed) and to track particles with low transverse momentum $p_T < 200$ MeV/c [2]. It is situated as close to beam pipe as possible and it covers pseudorapidity interval $\eta < 0.9$ and full azimuthal angle. It consists of 6 layers of detectors, as it can be seen in Fig. 2.2. Two innermost layers are SPD (Silicon Pixel Detector), next two are SDD (silicon drift detectors) and two outermost are SSD (Silicon Strip Detector).

In Table 2.1 are shown different properties of all layers of ITS. SPD is used for reconstruction of position of primary vertex and for measurements of impact parameters of tracks, coming from heavy flavour decays. SDD measures mainly energy loss of particles, that aids in further particle identification in ITS. Finally, SSD is used for matching track with signals from other detectors, mainly TPC. It provides also information about energy loss of particles.

2.2 Time-Projection Chamber (TPC)

Detector situated around ITS is TPC. It is filled by 90 m^3 of $Ne/CO_2/N_2$ (90/10/5). In this drift gas, signals from charged particles are transported on either side of central electrode to the end plates. At each end plate, there are multi-wire proportional chambers.

TPC is the main tracking detector that offers measurements of momentum of charged particles, particle identification and helps with determination of vertex position. For fully reconstructed tracks (with signals also in ITS, TRD, TOF) it has coverage $|\eta| < 0.9$ and for reduced tracks (reconstructed with lower resolution) $|\eta| < 1.5$. It covers full azimuth angle. Momentum range, that could be detected, is from 0.1 GeV/c to 100 GeV/c.

$\sqrt{m_1^2 + m_2^2 + 2E_{proj}m_2}$, where $m_{1,2}$ are masses of colliding particles and E_{proj} is their energy.

[2] Transverse momentum p_T is value of momentum in 2D space perpendicular to the direction of colliding particles (or to the beam direction).

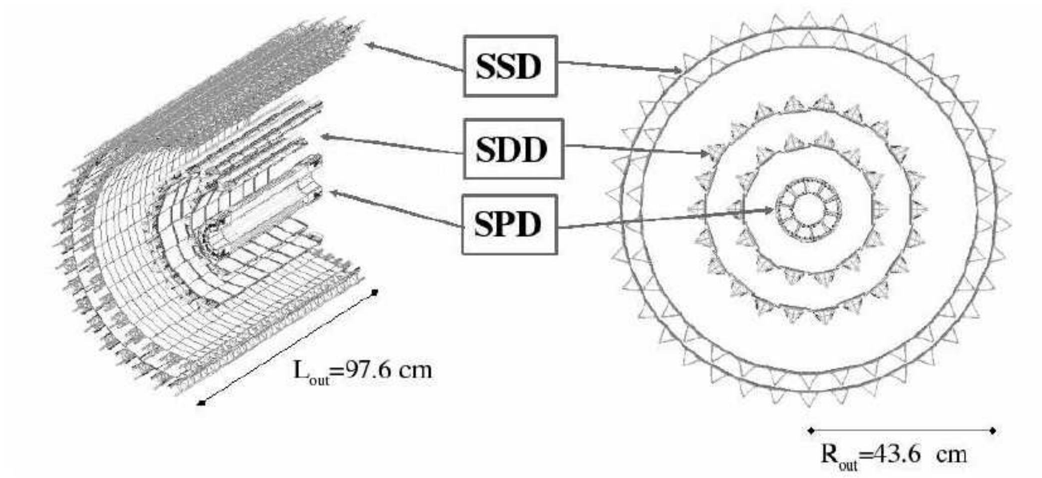


Figure 2.2: Layers of ITS detector. Taken from Ref. [5].

Layer	Type	Position r [cm]	Position $\pm z$ [cm]	Resolution $r\phi$ [μm]	Resolution z [μm]
1	pixel	3,9	14,1	12	100
2		7,6	14,1		
3	drift	15,0	22,2	35	25
4		23,9	29,7		
5	strip	38,0	43,1	20	820
6		43,0	48,9		

Table 2.1: Properties of layers of ITS detector: type of detection system, its distance from beam pipe r, length along beam pipe from center of ITS to both sides $\pm z$, its resolution in $r\phi$ space (perpendicular to beam pipe) and in beam direction z. Data taken from Ref. [5].

Resulting position resolution is from 800 μm in outer radius to 1100 μm in inner radius of TPC. In beam direction it is from 1100 μm to 1250 μm. Energy loss resolution for isolated tracks is around 5%, depending on multiplicity of tracks in collision. Energy loss in detector and associated momentum of passing particle can be used for particle identification. Performance of particle identification in TPC is showed in plot in Fig. 2.3.

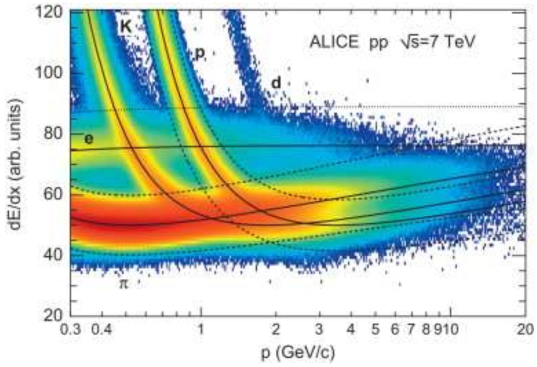


Figure 2.3: Energy loss, dE/dx spectrum versus momentum in the ALICE TPC from pp collisions at $\sqrt{s} = 7$ TeV. Taken from Ref. [6].

2.3 Electromagnetic calorimeter (EMCal)

Main purpose of EMCal is measurement of jet quenching in heavy ions collisions over the large kinematic range, also for high p_T. It is able to measure neutral energy of jets, thanks to this it can measure full jets. It is situated right under ALICE magnet, so around 4.5 m from interaction point. It covers $|\eta| < 0.7$ and azimuthal angle interval with size $\Delta\phi = 107°$. In azimuth, its situated opposite to PHOS (PHoton Spectrometer). Position of EMCal in central barrel is shown in Fig. 2.4.

EMCal is Pb-scintillator, that is segmented into 12 288 towers of $6.0 \times 6.0 \times 24.6$ cm^3, all directed to interaction point. Every tower contains alternating layers of Pb (thickness 1.44 mm) and polystyrene base scintillators (BASF143E + 1.5%pTP + 0.04%POPOP, thickness 1.76 mm). Every tower is optically isolated.

Resulting energy resolution is $15\%/\sqrt{E} \oplus 2\%$ [7] for jet measurements and $12\%/\sqrt{E} \oplus 1,7\%$ [7] for measurements of electrons and photons. Position of electromagnetic showers is measured with precision of 1.5 mm + 5.3 mm/$\sqrt{E_{deposit}}$ [7], that is nearly the same in all directions.

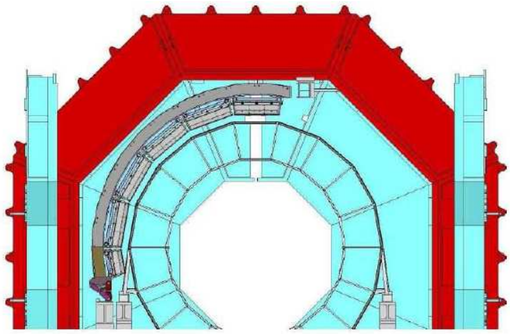


Figure 2.4: Position of EMCal in central barrel of ALICE detector. Taken from Ref. [5].

Chapter 3

Jets in collisions

Jets are defined as collimated sprays of hadrons, produced via fragmentation of high-energy quarks or gluons. After hadronisation, some heavy hadrons can be formed, especially if some heavy quark fragmented to form a jet. These heavy hadrons decay to light hadrons, which can be also contained in jet cones.

Different types of jets are considered, depending on partons (quark or gluon) from which they were formed (fragmented): usdg-jets (mother particle could be one of light quarks u, s, d or gluon), c-jets (mother particle is c quark) and b-jets (mother particle is b quark). After hadronisation, this particle is contained in relevant hadron, so for example, if mother b quark fragments, after hadronisation, there are few light hadrons and B hadron. The last one decays to lighter hadrons, which are detected.

3.1 Motivation for jet studies

Jets are considered as one of the most important probe of the partonic medium, that can be created in collisions of heavy ions. In case of pp collisions, jet production can be quite satisfyingly predicted by pQCD calculations and vice versa, so pQCD calculations can be improved by jet measurements. Jets can be also used to study hadronisation and hard scattering. Collisions of protons are usually used as reference of measurements in p-Pb or Pb-Pb collisions. In these cases, measured jet productions could be suppressed, mainly for central collisions.

This suppression could be expressed by nuclear modification factor R_{AA}, usually defined as

$$R_{AA} = \frac{1}{N_{coll}} \frac{Y(AA)}{Y(pp)}, \tag{3.1}$$

where $Y(AA)$ and $Y(pp)$ are particle yields in heavy ions and pp collisions (usually in some specific range of momentum or pseudorapidity) and N_{coll} is number of binary collisions of nucleons in heavy ions collisions, that depends on centrality of collision. If $R_{AA} < 1$, production is suppressed, in case $R_{AA} > 1$, production is enhanced. Measured R_{AA} for jets in Pb-Pb collision at $\sqrt{s_{NN}} = 2.76$ TeV is shown in Fig. 3.1, taken from Ref. [8]. We can observe strong suppression for jets. R_{AA} slowly rise with higher transverse momentum of jet.

By comparing jet production in pp and Pb-Pb collisions, the properties of produced medium can be studied. These are mainly temperature of the medium, its shape or energy

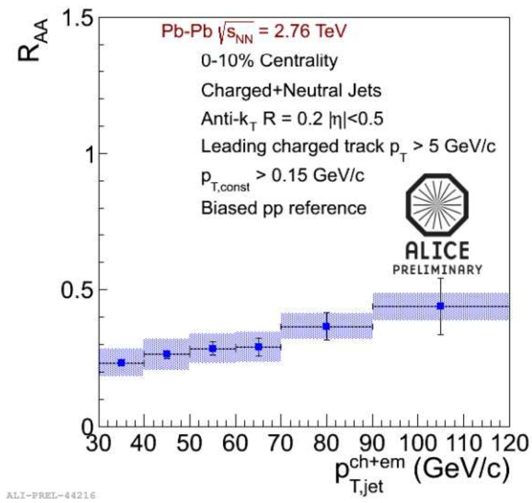


Figure 3.1: Nuclear modification factor R_{AA} of fully reconstructed jets (reconstructed using anti-k$_T R < 0.2$) and requiring a high p_T leading track, $p_T > 5$ GeV/c) in the 0-10% centrality bin, for transverse momentum of jet 30 GeV/c $< p_{T,jet}^{ch+em} <$ 120 GeV/c measured by ALICE experiment. Taken from Ref. [8].

loss of particles in it. The last one is expected to be different for gluons, light quarks and heavy quarks (more details in Ref. [9]). For b quarks, it is expected to be smaller than for c-quarks, which is smaller than for light quarks or gluons. This energy loss dependence on particle mass is one of our motivation to study heavy flavour jets. Because of large mass of b quark, it is expected to be created right after the collision, in hard scattering. So b quarks experience the full evolution of the system, that makes them an excellent probe of medium properties. They enable us to study redistribution of lost energy of quarks in medium or possible modification of b quark fragmentation in medium.

3.2 Jet-finding algorithms

Jet reconstruction occur in space defined by pseudorapidity η and azimuthal angle ϕ. Coordinates of axis of jet in this space are

$$\eta = \sum_i \frac{E_T^i \eta^i}{E_T^J}, \phi = \sum_i \frac{E_T^i \phi^i}{E_T^J}, \tag{3.2}$$

where E_T^J is total transverse energy of jet, η^i and ϕ^i are coordinates of particles in jet and E_T^i are energies of these particles. Transverse energies in definitions could be replaced by transverse momentum p_T.

After tracks and energies of particles in event are reconstructed, different algorithms can be used to find and reconstruct jets. There are different requirements on these algortihms, the most important are:

- **Infrared safety:** soft particle should not change number and properties of reconstructed jet.
- **Collinear safety:** in sense of reconstructed clusters, two particles with low energy or mass, propagating close to each other, should not be mismatched as one more

energetic particle and vice verse. Analysis of energetic cluster have serious influance on efficiency of jet reconstruction.

- **Order independence:** after reconstructions, resulting jets should be same in parton, hadron and detecter level. This could be tested in MC simulations of collisions.
- **Independence on detector geometry and granularity.**
- **Maximum jet-finding efficiency vs. CPU time.**

3.2.1 Cone algorithms

Cone algorithms firstly search the most energetic particles in η-ϕ space of event. These clusters should have larger energy than set up threshold value, if it is, they are tagged as "seeds". After that, all particles, which distance from seed is smaller as threshold value of R ($R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$) are inherited in a jet cone. Then, all particles in actual jet cone are considered to find new seed. It is a weighted center of actual jet cone. This new seed is used to construct new jet cone, in a same way as before. Process repeats, till some stability of shape of jet or jet axis cone is achieved.

The problem is, that these algorithms are not usually collinear and infrared safe.

3.2.2 Clustering algorithms

Another group of algorithms for jet reconstruction is clustering algorithms. They are based on finding some kind of weighted distance beetwen particles (i, j) defined as

$$d_{ij} = \min(k_{Ti}^p, k_{Tj}^p)\frac{\Delta_{ij}^2}{D^2}, \tag{3.3}$$

where parameter p define influence of transverse momentum of particle vs. its geometrical properties, parameter D assures minimal distance between reconstructed jets,

$$\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2, \tag{3.4}$$

and k_{Ti} is defined for every particle i as

$$k_{Ti} = \frac{E_i}{c}\sin\theta_i, \tag{3.5}$$

where θ_i is azimuthal angle of particle and E_i is its energy. Then distance between beam and particle i is defined as

$$d_{iB} = k_{Ti}^p. \tag{3.6}$$

For event, all d_{iB} and d_{ij} are calculated. From these numbers, the smallest one is chosen. If it is some of d_{ij}, so particles i and j are merged and later considered as one particle, for which energies, momentum and distances should be recalculated. If it is some of d_{iB}, particle i is tagged as jet (in later steps, more particles are "hidden" in this one). This steps are repeated, till all particles are part of jets.

Algorithm, that used $p = 2$ is called k$_T$ algorithm, for $p = 0$ it is called C/A (Cambridge/Aachen) algorithm and for $p = -2$ it is anti-k$_T$. The last is actually the most used one, it was used in our analysis, too.

3.3 Heavy flavour jets

Secondary vertices are places, where heavier hadrons decay to some daughter particles, so multiple tracks can be coming out from them. These tracks are specific for their large displacement from primary vertex (interaction point), this is also how these secondary vertices are found in data. The invariant mass of secondary vertex is

$$m_{inv}^2 = (\sum_i E_i)^2 - \|\sum_i \overrightarrow{p_i}\|^2, \tag{3.7}$$

where E_i are energies of particles coming out from secondary vertex and p_i are their momentum. Distance of secondary vertex from primary vertex, L_{xy}, depends on decay length of meson decaying in this vertex.

As already mentioned, b-jets contains B hadrons, mainly B mesons. They are heavy, so also secondary vertices created after they decay have large invariant mass (≈ 5 GeV/c^2). Lifetime of B meson is large, their flight distance is ≈ 500 μm. Fig. 3.2 shows the geometrical schema of b-jet.

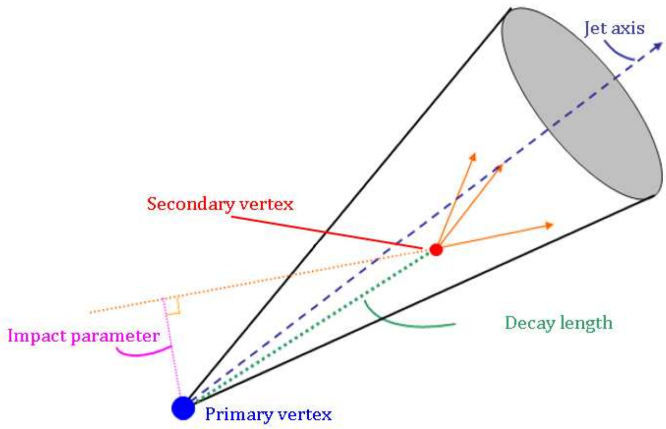


Figure 3.2: Geometry of jet, showing jet axis, decay length of secondary vertex (L_{xy}) and impact parameter of track (vertex).

3.4 Tagging of b-jets

Measurements of b-jets in heavy ions collisions were already successfully done at CMS experiment, see Ref. [10]. Algorithms used for b-tagging exploit from B meson or b-jet properties described before. Simple Secondary Vertex (SSV) algorithm uses properties of secondary vertices in jet, whereas Track Counting (TC) algorithms uses properties of tracks in a jet (e.g. large displacement). Performance of every algorithm in these algorithms is expressed by b-tagging efficiency and its udsg-jets and c-jets mis-tagging efficiencies. Efficiency of tagging for a given jet flavor is ratio of number of tagged jets of a given flavor to all jets of a given flavor. Mis-tagging efficiency is defined similarl. Its name differs, because it refers to jets, we do not want to tag (other kind of jets as searched). To estimate value of efficiencies, simulation data are be used, because the number of real jets is needed. Efficiencies strongly depend on selected values of discriminators.

Example of binding of b-tagging and mis-tagging efficiencies for different values of discriminators and for different b-tagging algorithms is shown in Fig. 3.3. It can be observed, that for all algorithms, higher b-tagging efficiency means also higher mis-tag efficiencies. This is natural property of discriminators, our goal is to set up them in a such way, that c-jets and usdg-jets are rejected, but b-tagging efficiency is high enough.

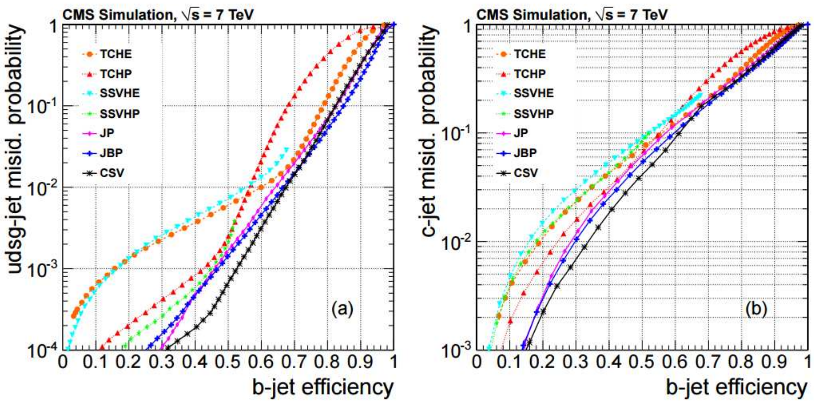


Figure 3.3: On the left usdg-jet and on the right c-jet misidentification probabilities as functions of the b-jet efficiency for different b-tagging algorithms at CMS and for different values of their discriminators, for simulation of pp collisions at $\sqrt{s} = 7$ TeV. Taken from Ref. [10].

3.4.1 Simply Secondary Vertex algorithm

Firstly, all secondary vertices in event are found. This is done with two (high efficiency - SSVHE) or three (high purity - SSVHP) prolongated tracks, depending on analysis strategy or data statistics. For b-tagging, different properties of these secondary vertices could be used, mostly it is their invariant mass or distance from primary vertex. Mostly, from all secondary vertices in jet, we choose only secondary vertex with the furthest distance from primary vertex and use its properties. At first, some cuts are applied on distributions of variables to construct (b-)tagged sample, that is expected to have reduced number of light jets. Efficiency of this cut has significant influence in further b-tagging efficiency. After that, another cut is applied on this b-tagged sample. All secondary vertices passing it, are considered as b-jets.

Example of discriminator is the variable called signed flight distance

$$L_{xy} = L'_{xy} \times \mathrm{sign}(L_{xy} \cdot p_{T,jet}), \tag{3.8}$$

where L'_{xy} is value of flight distance of secondary vertex (shown in Fig. 3.2) and sign of scalar product is positive, when SV is in same direction from primary vertex as transverse momentum of jet $p_{T,jet}$, and negative, when it in opposite direction. Another variable, that can be used is signed flight distance significance, SL_{xy}. It is defined as

$$SL_{xy} = \frac{L_{xy}}{\sigma_{L_{xy}}} \tag{3.9}$$

where L_{xy} is value of signed flight distance of secondary vertex and $\sigma_{L_{xy}}$ is its error.

3.4.2 Track Counting algorithm

As a first step, impact parameter of every track in jet d_0 is calculated, then it is projected along jet axis. Then, these are ordered in decreasing order. The third value is used as discriminator and it is compared with some threshold value. If it is bigger than threshold, jet is tagged as b-jet. Actually at ALICE, typical threshold value is $\approx 100\ \mu$m. With this value, achieved b-tagging efficiency is ≈ 0.1.

Current results from this algorithm for ALICE MC data are shown in Fig. 3.4. In right column, we can observe comparison of b-tagging efficiency and mis-tagging eficiencies for c-jets and usdg-jets. In tagging algorithms, our goal is to suppress mis-tagging efficiency in comparison with b-tagging efficiency. As we can see in Fig. 3.4, usdg-jet mis-tagging efficiency is suppressed by factor 10 to 100, but c-jet mis-tagging is by factor only around 10. Even for small values of transverse momentum of jet, ratio of c-jet mis-tagging to b-tagging eficiency is around 0.2, so in data 20% of jets tagged as b-jets could be c-jets. We should also be conscious of b-tagging efficiency of 0.1, so with this setup, we are actually able to find only 10% of real b-jets.

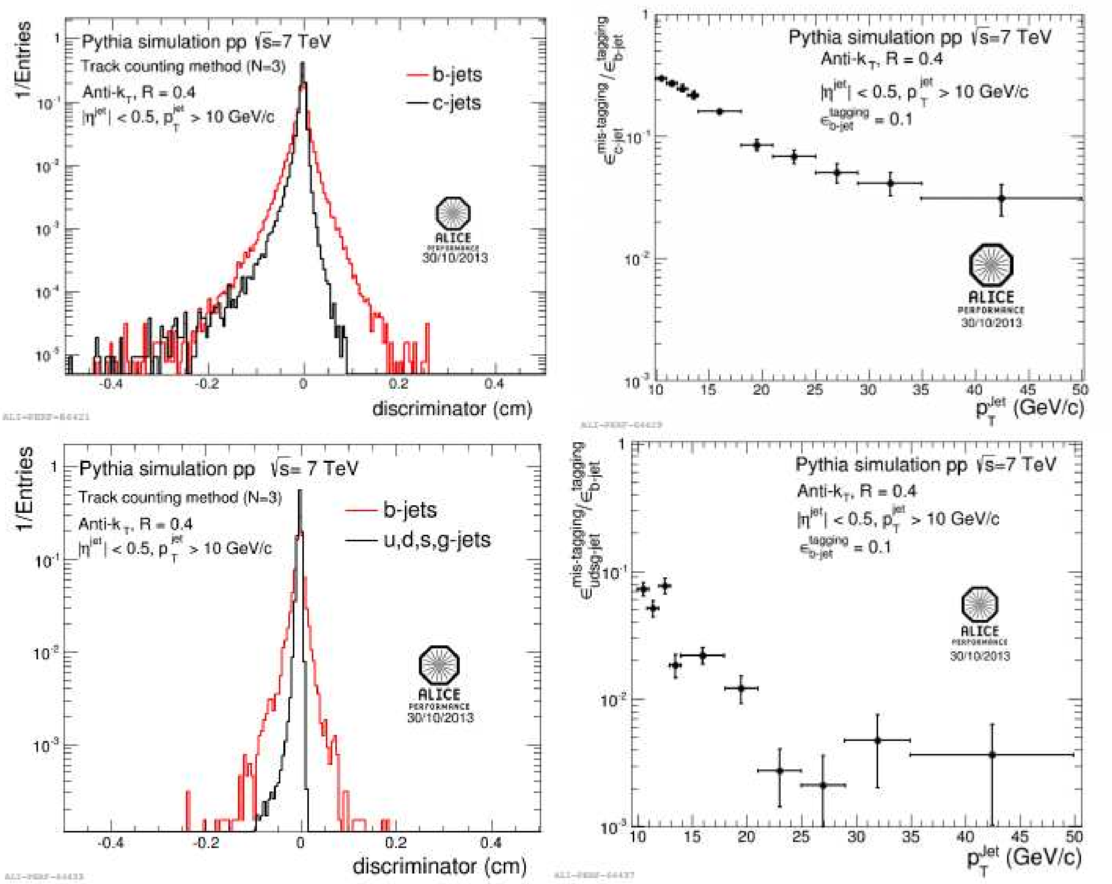


Figure 3.4: Current performance of Track Counting algorithm at ALICE, for PYTHIA simulation of pp collisions at $\sqrt{s} = 7$ TeV. First line compares b-jets and c-jets, second line b-jets and usdg-jets. In left column, distributions of impact parameter with respect to jet axis of the third most displaced track in jet (discriminator), in right ratio of other flavour mis-tagging efficiencies and b-tagging efficiency. Threshold value for third most displaces track in Track Counting was 100μm. Taken from Ref. [13].

This algorithm could also work with signed impact parameter of track, that is defined analogously as for secondary vertex (Eq. 3.9).

Chapter 4

Results of b-jet studies from other experiments

As already mentioned, b-tagging was successfully done at CMS experiment. I would like to present analysis as it was made in Ref. [11] , because one of our goals is to reproduce steps in this analysis. This article shows measurements of b-jet fraction (ratio of b-jets to inclusive jets) in pp colisions at $\sqrt{s} = 2.76$ TeV and Pb-Pb collisions at $\sqrt{s_{NN}} = 2.76$ TeV, and compare two different algorithms for b-tagging, Simply Secondary Vertex and Jet Probability algorithms and results from Pythia simulations. Method using secondary vertices is described in more details.

For SSVHE algorithm, impact parameter significance variable was used as discriminator. Jet Probability (JP) algorithm calculates probability, that one track comes from primary vertex, and then calulate this probablity for jet, using all tracks in it. The performance of discriminators for this two b-tagging algorithms is shown in Fig. 4.1. We can observe b-jet efficiency vs. usdg-jet and c-jet mis-tag efficiencies in pp and Pb-Pb collisions. All these informations were calulated using MC data from Pythia and Hydjet simulations. The red cross shows value of SSVHE discriminator used in analysis. For this value, mis-tag efficiencies are smaller for pp collisions and also b-jet efficiency is slightly higher for pp collisions. For Pb-Pb collisions, resulting b-tagging efficiency is around 45% with rejection of c-jets by factor 10 and usdg-jets by factor 100.

Fig. 4.2 shows distribution of secondary vertex mass in data tagged by SSVHE discriminator (so it is not distribution of all secondary vertices, but only these, that have bigger probability to be in b-jet). Unbinned maximum likelihood template fit was used to fit data, and resulting contributions of b-jets, c-jets and usdg-jets are shown. For this fit, templates for different kinds of jets were constructed using MC data. In fitting, relative contributions of different kinds of jets may vary, but their shapes are fixed. The resulting precision of fit, χ^2/NDF is also shown in Fig 4.2. It is small enough, so fits were considered as successful. We can observe, that for higher secondary vertex mass the contribution from b-jets is also bigger.

From distributions, b-jet purity can be extracted. It is defined as fraction of true b-jets in SSVHE-tagged sample (Fig. 4.2). Resulting distribution of b-jet purity for different transverse momentum of jet is shown in left in Fig. 4.3. There is a comparison of b-jet purity of SSVHE-tagged sample in MC (Pythia and Hydjet) data and from template fits in data. These two are consistent in full range of transverse momentum. Also b-tagging efficiency was calculated from simulations and this was compared for SSVHE method and for Jet Probability algorithm. As it can be observed in Fig. 4.3 right, two b-taggers have

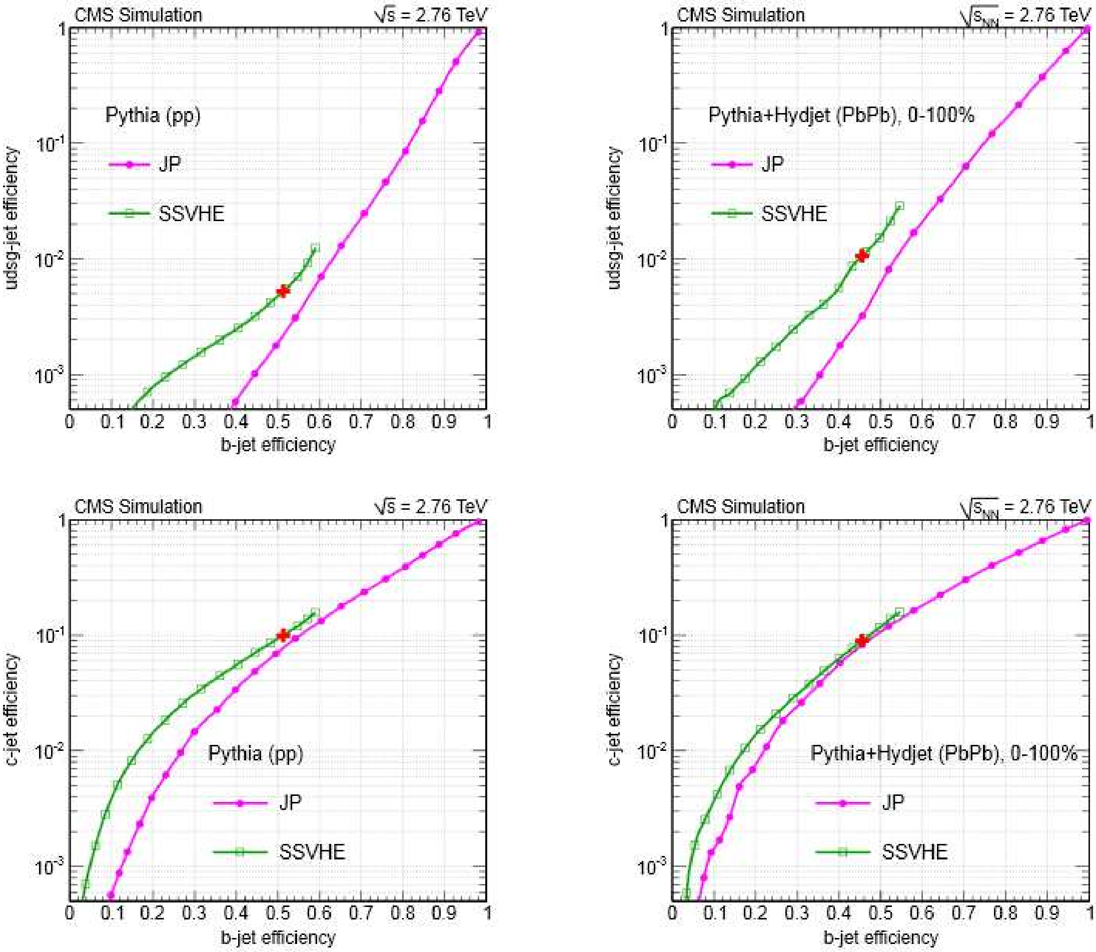


Figure 4.1: The b-jet tagging efficiency vs. the light jet (top) and charm jet (bottom) mis-tag efficiencies for simulated pp events from PYTHIA (left) and simulated Pb-Pb events from PYTHIA embedded in HYDJET (right) for the SSVHE and JP discriminators. The red cross marks the working point of the SSVHE discriminator used in this analysis. Taken from Ref. [11].

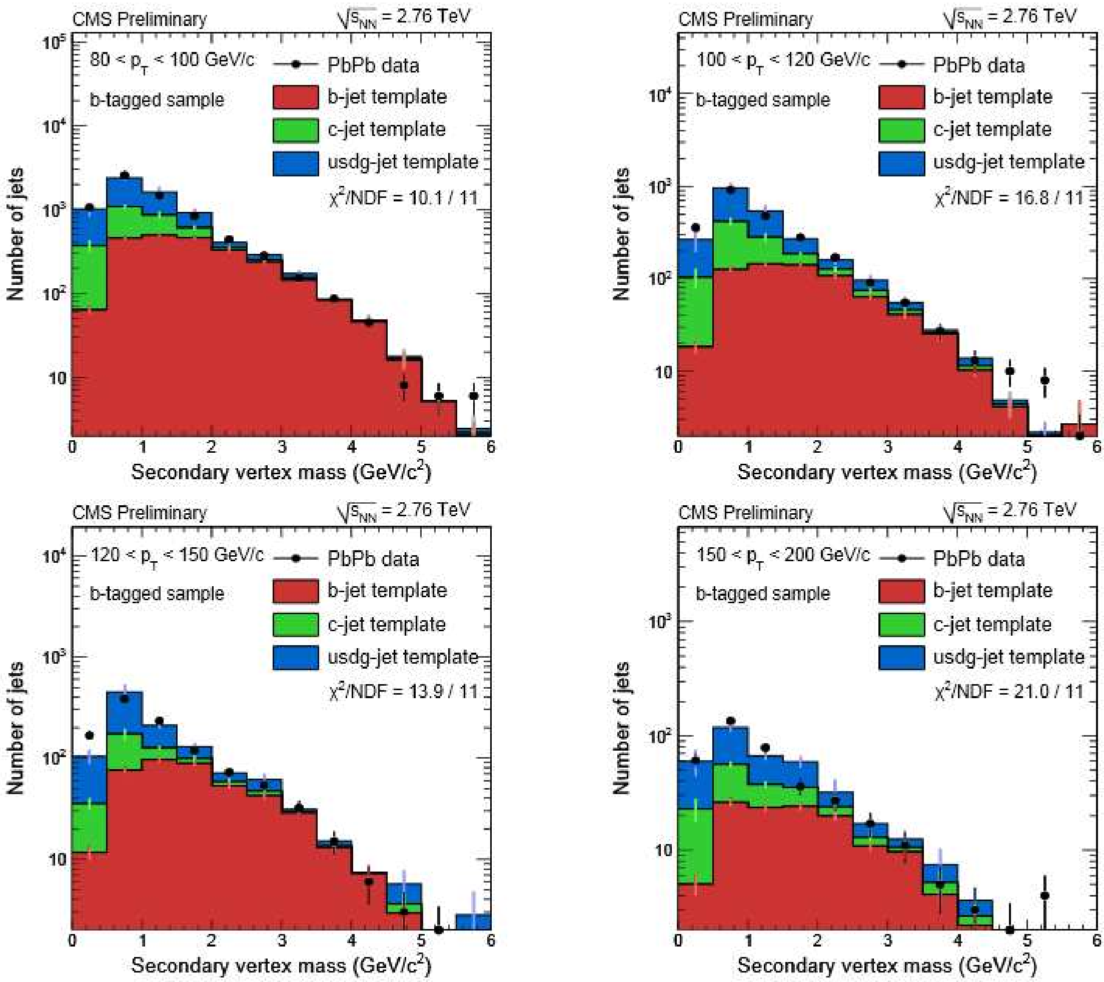


Figure 4.2: Template fits to the SV mass distributions in Pb-Pb collisions, after tagging with the SSVHE discriminator. Several p_T ranges are shown as indicated on the figures. The colored lines represent the statistical uncertainties on the MC templates. Taken from Ref. [11].

nearly same efficiency around 45% for transverse momentum $80 < p_T < 200$ GeV/c. Small difference between them is used in calculation of systematic uncertainty.

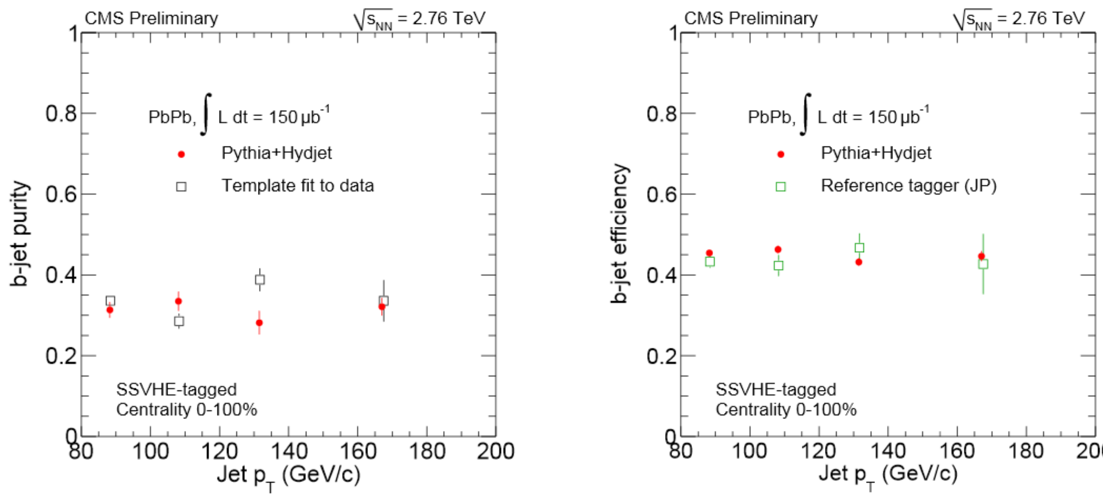


Figure 4.3: Left: The b-jet purity extracted from template fits to the SV mass distributions, compared to the input simulation. Right: The tagging efficiency of the SSVHE discriminator from simulation and from the reference tagger method. Taken from Ref. [11].

From shown figures and results, b-jet fraction for every p_T range can be calculated as

$$b_{frac} = \frac{N_{\text{jets}}^{\text{tagged}}}{N_{\text{jets}}} \frac{P}{\epsilon}, \tag{4.1}$$

where $N_{\text{jets}}^{\text{tagged}}$ is number of tagged jets by SSVHE, N_{jets} is number of all jets, P is the purity and ϵ is efficiency of algorithm, both plotted in Fig. 4.3. The left panel in Fig. 4.4 shows resulting b-jet fraction for 0-100% centrality Pb-Pb collisions as a function of jet p_T. This fraction is $2.9 - 3.5\%$ without significant dependence on jet transverse momentum. The values from Pythia and Hydjet are also shown, and they are slightly smaller than results from data, but within error these two are consistent. The right panel in Fig. 4.4 shows same comparison for pp collisions. Results from Pythia are consistent with results from data, and b-jet fraction is again around 3%, with small drop for $80 < p_T < 100$ GeV/c.

Centrality dependence of b-jet fraction is shown on left in Fig. 4.5. It is also around 3% with no dependence on centrality within uncertainties. The right panel in Fig. 4.5 shows the ratio of the b-jet fraction in $0 - 100\%$ central Pb-Pb to pp collisions. In the lowest p_T bin, it is around 1.6, but with a very large uncertainty. In the other bins, it is consistent with unity. The b-jet nuclear modification factor R_{AA} can be calculated as the product of the ratio of b-jet fraction in Pb-Pb and pp collisions (shown in the right panel in Fig. 4.5), and the inclusive jet R_{AA} (0.50 ± 0.01 (stat.) $\pm$ 0.06 (syst.) [12]) and has a value of 0.48 $\pm$ 0.09 (stat.) $\pm$ 0.18 (syst.), for $100 < p_{T,jet} < 120$ GeV/c.

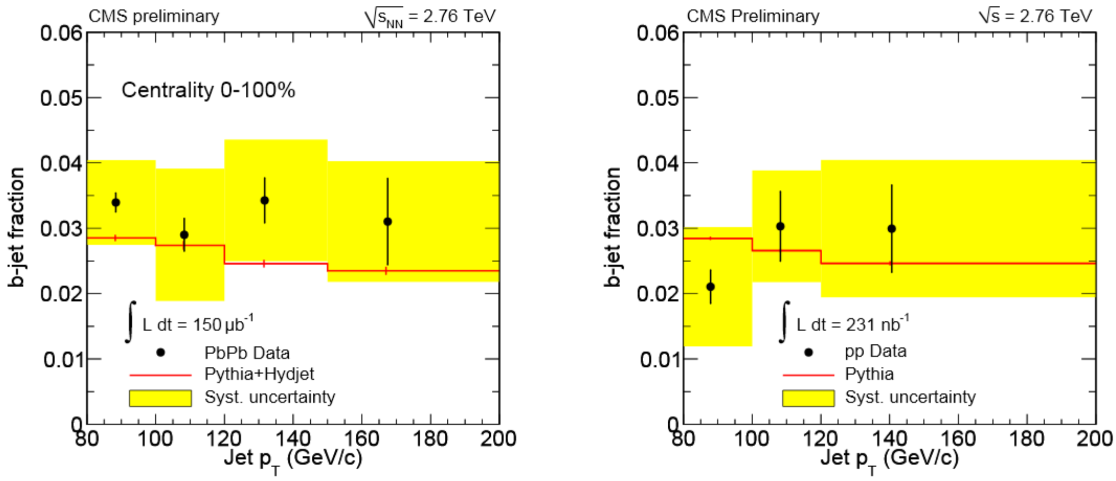


Figure 4.4: The b-jet to inclusive jet ratio in 0-100% Pb-Pb collisions (left) and pp collisions (right) as a function of jet transverse momentum p_T compared to PYTHIA embedded in HYDJET (Pb-Pb) and PYTHIA (pp). Data and MC have not been corrected for bin migration effects from finite jet resolution. Taken from Ref. [11].

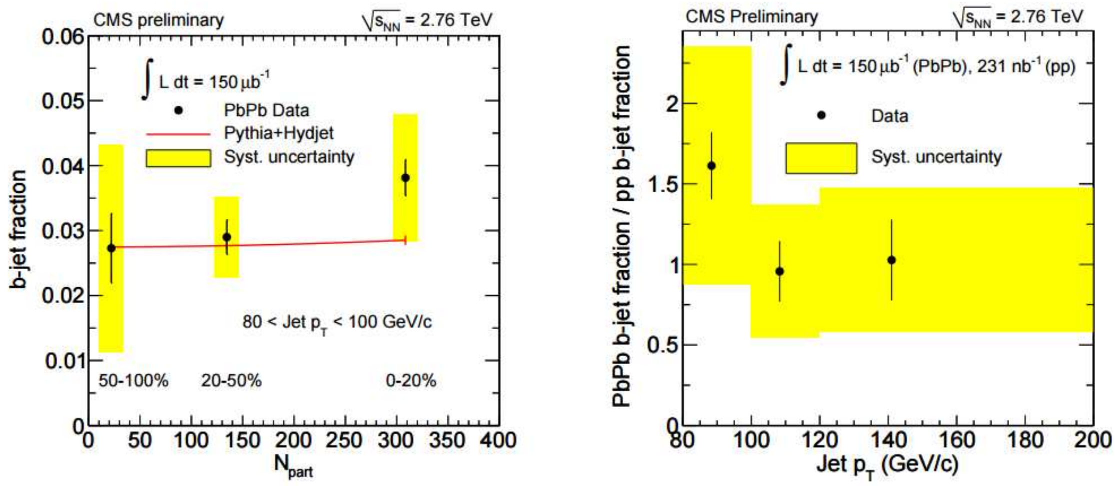


Figure 4.5: The b-jet to inclusive jet ratio for $80 < p_T < 100$ GeV/c as a function of Pb-Pb collision centrality (left) and the ratio of the b-jet fraction in Pb-Pb to the b-jet fraction in pp as a function of jet p_T (right). Data and MC have not been corrected for bin migration effects from finite jet resolution.Taken from Ref. [11].

Chapter 5

Study of b-jets in pp collisions from Monte Carlo data

In the following, the analysis of data that I have performed will be described. In this chapter MC productions LHC13d9 (879 000 events analysed) and LHC13d14 (2 108 800 events analysed) were used. These are simulations of pp collisions at $\sqrt{s} = 7$ TeV from Pythia generator [14]. For this analysis, I used Aliroot software with Fastjet package [15], used to find jets in collisions. Detector response was done with Geant package [16].

Before analysing b-jet, finding of all inclusive jets is needed. This was done using anti-k$_T$ algorithm from Fastjet package. Different values of paramater R were tested: 0.2, 0.4 and 0.7. The output of this jet-finding algorithm is array of all jets, array of all final state particles (detected in detector) in jet for every jet and array of all MC particles (these are not usually in final state, this information is available only in MC data) in jet for every jet. The last one was used for my next step, b-jet tagging.

5.1 Algorithm of HF-jet identification in MC

Two different MC b-tagging method were tested.. They use MC information about particles, so it could not be used in data.

One algorithm search for mother particle in MC array of particles in jet, for every particle in every jet. If mother particle is b quark, jet containing it is tagged as b-jet.

By default, ALICE use algorithm, that search for mother B hadrons. Schema of algorithm is as follows:

1. For every jet, look in its array of MC particles. For every particle, do as follows:
2. Search for its mother particle (particle, which decayed to produce this particle). Find its PID.
3. If mother particle is b or c hadron:
   (a) Calculate its distance dR from a jet axis, in η-ϕ space.
   (b) If d$R <$ dR_{thr} (different dR_{thr} tested):
      i. If mother particle is b hadron $\Longrightarrow$ jet is tagged as b-jet, stop loop over MC particles in jet.
      ii. If mother particle is c hadron $\Longrightarrow$ save it and its properties to different array, continue to search for b hadron (step 2).

4. When the loop over all MC particles in jet is ended and it has not been tagged as b-jet, look in array where c-particles are saved (from step 3(b)ii). If it is not empty $\Longrightarrow$ tag jet as c-jet.
5. If jet is not tagged as b-jet or c-jet, tagg it as light jet.

5.2 Properties of b-jets

In order to study how many B hadrons are used for b-tagging (their $\mathrm{d}R < \mathrm{d}R_{thr}$), I plotted their $\mathrm{d}R$ distribution (Fig. 5.1). For this, I searched for B hadrons in every event. After I found one, I calculated its distance from jet axis of every jet in event and plotted the shortest one. We can observe, that this distribution has maximum around $\mathrm{d}R = 0.25$ and then it drops nearly exponentially. In ALICE, for b-tagging $\mathrm{d}R_{thr} < 0.7$ is mostly used. As we can see in Fig. 5.1, this method use approximately 80% of all B hadrons. Other B hadrons are not used for b-tagging.

In following study, anti-k$_T$ with $R < 0.4$ and in HF-jets tagging $\mathrm{d}R_{thr} < 0.7$ were used.

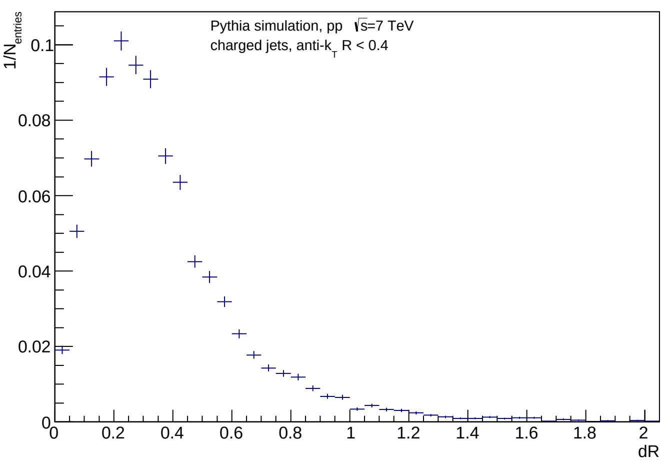
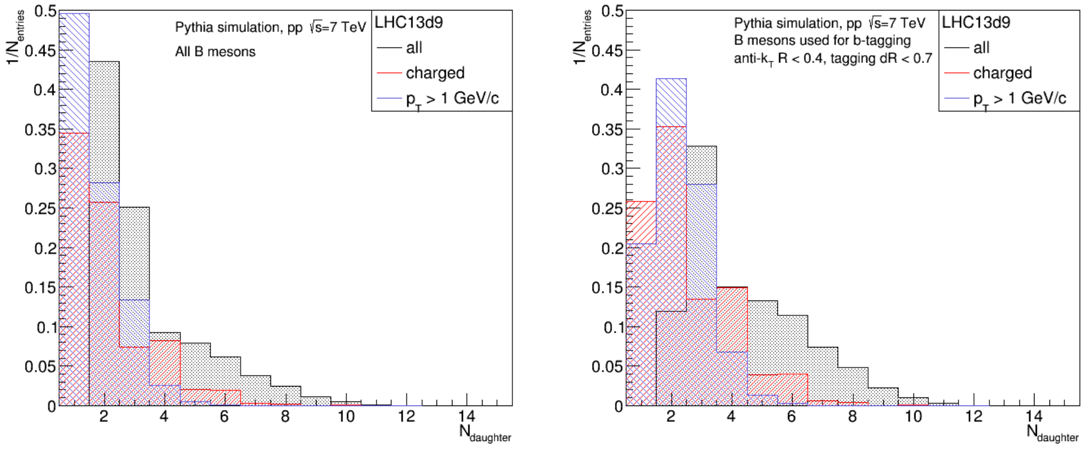


Figure 5.1: Distance of B meson from closest jet $\mathrm{d}R$ in MC pp collisions. In MC b-tagging, B mesons with $\mathrm{d}R < 0.7$ are usually used.

Another attribute of B mesons, that can be studied, is number of daughter particles created after they decay, shown in Fig. 5.2. Different type of daughter particles were compared, as well as different mother B mesons. In my analysis, I compared distributions of number of all daughter particles, charged ones and those with $p_T > 1$ GeV/c. I compared these distributions for all B mesons created in collisions (Fig. 5.2a) and for B meson that were used for b-jet tagging ($\mathrm{d}R$ from any jet was smaller than $\mathrm{d}R_{thr}$ used in analysis, Fig. 5.2b). We can observe, that distribution for B mesons, used for b-jet tagging, is “shifted” to higher number of daughter particles. Mean number of charged particle from all B decays is ≈2, for B mesons used for b-tagging it is ≈2.5. Same results are observed for number of daughter particles with $p_T > 1$ GeV/c.

Comparison of b-jet’s transverse momentum $p_{T,\mathrm{jet}}$ and transverse momentum of B meson in it $p_{T,\mathrm{meson}}$ is shown in correlation histogram Fig. 5.3. In spite of obvious small statistics, we can see that $p_{\mathrm{T,meson}} \approx 1.5 p_{\mathrm{T,jet}}$. It is in accordance with present results from



(a) (b)

Figure 5.2: Number of all daughter particles (blue), charged (red) and with $p_T > 1$ GeV/c (green) from B meson decays in MC pp collisions at $\sqrt{s} = 7$ TeV, (a) for all B mesons created in collisions; (b) for B mesons in tagged b-jets.

ALICE experiment [13].

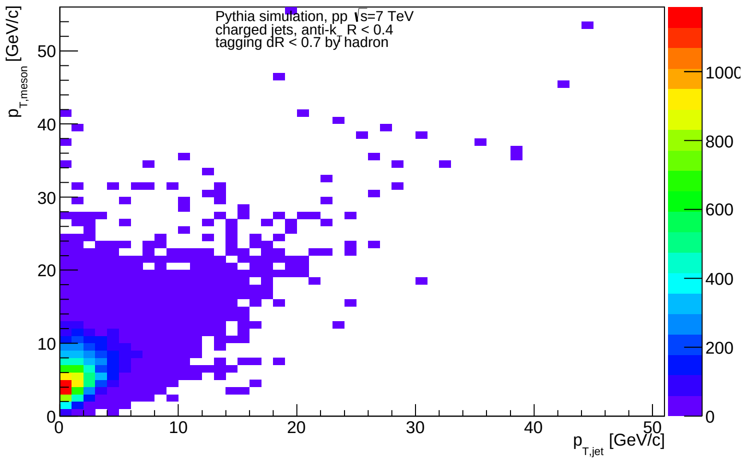
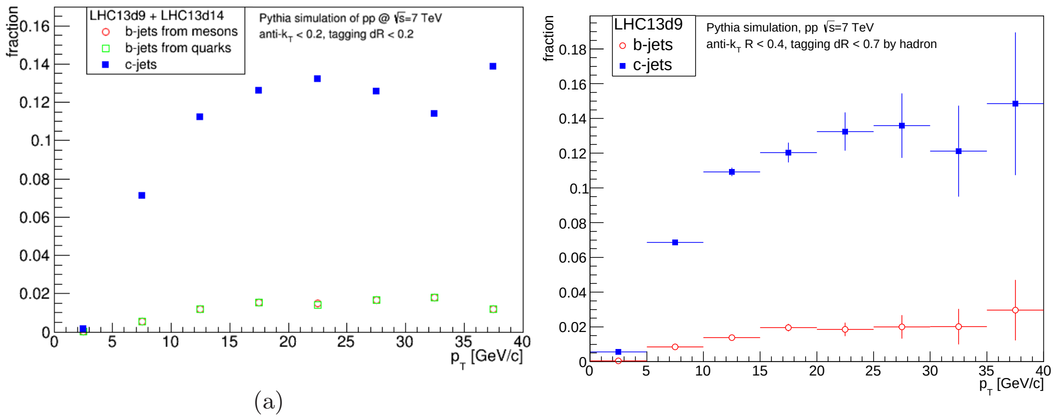


Figure 5.3: Correlation of B-meson transverse momentum $p_{T,\mathrm{meson}}$ and according b-jet transverse momentum $p_{T,\mathrm{jet}}$, MC pp collisions at $\sqrt{s} = 7$ TeV.

5.3 Fraction of HF-jets in pp collisions

Fraction of HF jets is defined as number of HF-jets (b-jets or c-jets) divided by number of inclusive jets. It was study in CMS experiment for jets with $p_{T,\mathrm{jet}} > 80$ GeV/c [11]. However, in ALICE, the goal is to study jets with smaller $p_{T,\mathrm{jet}}$. I studied b-jet fraction for $0 < p_{T,\mathrm{jet}} < 40$ GeV/c. Results are in Fig. 5.4. I compared b-fraction results from different options in b-tagging algorithm. For b-fractions in Fig. 5.4a, I used tagging by mother B meson (red circles) and by mother b quark (green squares). For c-tagging I used only hadrons with c-quark. In all cases, I used anti-k$_T$ with $R < 0.2$ and in HF-jets tagging d$R_{thr} < 0.2$ (HF particle inside b-jet). As we can see, b-jet fraction is slowly rising to approx. 2 % and results are same for b-tagging by mother meson and mother quark. For these options, c-jet fraction is slowly rising to approx. 12 %. In Fig. 5.4b, there are results of HF-jet fractions for anti-k$_T$ < 0.4 and d$R_{thr} < 0.7$. In this case, statistical errors are also plotted. Results are close to results in Fig. 5.4a.



(a)

(b)

Figure 5.4: Fraction of b-jets and c-jets in MC pp collisions at $\sqrt{s} = 7$ TeV, with different MC b-tagging and jet-finding methods, (a) c-jets tagged by mother meson, b-jets tagged by mother meson and quark, distance between jet and mother particle < 0.2, anti-k$_T$ $R < 0.2$; (b) c-jets tagged by mother meson, b-jets tagged by mother meson, distance between jet and mother particle < 0.7, anti-k$_T$ $R < 0.4$.

In CMS experiment, b-jet fraction around 2 % for $p_{T,\mathrm{jet}} < 40$ GeV/c was measured [17]. From our results of analysis of PYTHIA data we can see, that PYTHIA gives approximately reasonable fraction of b-jets.

Chapter 6

Performance of the ALICE seondary vertex b-tagging algorithm

6.1 Template fitting of p-Pb data

In this section, fitting of distributions of different variables is described. I have done this analysis in collaboration with Physical Working Group - Heavy Flavour Correlations in Jets (PWG-HFCJ) at ALICE experiment and results were presented in meetings of this group. Goal is to study distributions of different variables describing properties of b-jets in MC data and then use this study to fit real data from p-Pb collisions. That can help to extract b-jet purity in data, which can be used, for example, to study b-jet fraction in data, as it was done at CMS experiment (Ref. [11], chapter 4).

In my analysis, I worked with these MC p-Pb data from MonALISA Repository for ALICE :

- LHC13b4_fix: Jet-Jet Pythia6 (Perugia 2011), repeat of LHC13b4, 0.5T, LHC13b anchors, 9 961 800 events;
- LHC14g3b: production with light quark sample in p_t hard bins and Hijing underlying event, LHC13bcde anchors, 61 555 500 events.

Energy of collisions was $\sqrt{s_{NN}} = 5.02$ TeV. I also worked with sample of minimum-bias p-Pb data recorded in Run 1. Data (MC or real) firstly passed analysis to find all secondary vertices and their properties. For this, official group's code was used. It is built in Aliroot since version vAN-20150409. This vertex-finder code use Simply secondary vertex algorithm with 3 prolongated tracks. In case of MC data, it could be determined, from which particle decay secondary vertices are produced, so they are tagged accordingly (b, c, usdg). There are different variables, that results from this code.

6.1.1 Distribution of variables

In Fig. 6.1, there is distribution of distance of most displaced vertex in jet from primary vertex (collision area) L_{xy} for b-jets, c-jets and usdg-jets. Distributions are results of analysis of LHC13b4_fix data. Different shapes of three distributions can be observed. All three distributions are falling exponentially, but the gradient differs. Number of usdg-jets is always the highest, bud this distribution is falling quickly. Gradients of distributions of b-jets and c-jets are falling slower, the smallest gradient is for b-jets. For values of $L_{xy} \gtrsim 0.3$ cm, the number of b-jets is even higher than for c-jets.

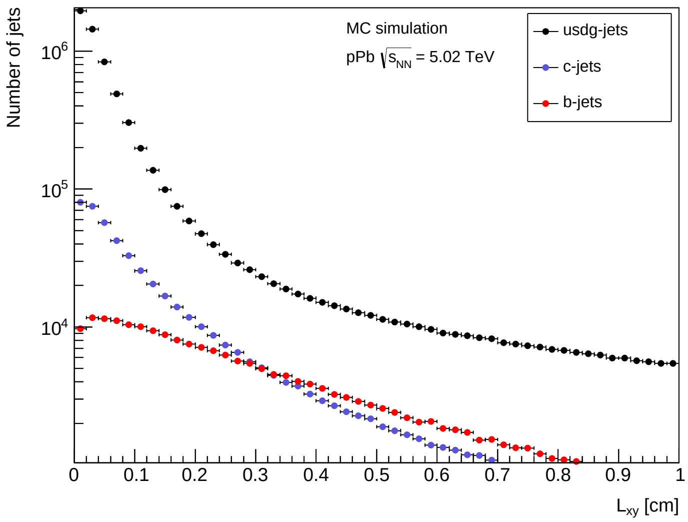


Figure 6.1: Distance of most displaced vertex in jet from primary vertex L_{xy} in MC p-Pb collisions for b-jets (red), c-jets (black) and usdg-jets (black).

Distribution of signed impact parameter is shown in Fig. 6.2. As already mentiones, signed impact paramater SL_{xy} is defined as value of impact parameter of furthest secondary vertex in jet from primary vertex L_{xy} divided by its error $\sigma_{L_{xy}}$, $SL_{xy} = L_{xy}/\sigma_{L_{xy}}$. We can observe nearly same behavour of shapes of three distributions (b-jets, c-jets and usdg-jets) as for distributions of L_{xy} in Fig. 6.1. Number of b-jets is bigger than number of c-jets for $SL_{xy} \gtrsim 10$.

These shape particularities are also observed for other variables. In our analysis we work only with one vertex by jet, we use the properties (variables) of the farthest one. They can be used for tagging in real data, where there is only one distribution for all jets (so sum of distributions for b-jets, c-jets and usdg-jets). It could be done by fitting distributions in MC, and use sum of these fits to fit distributions in data. So, if from MC normalized distribution for b-jets f_b, c-jets f_c and usdg-jets f_{usdg} are resulting, normalized data distribution is fitted by

$$f_{all} = p_b f_b + p_c f_c + p_{usdg} f_{usdg}, \tag{6.1}$$

where p_b, p_c and p_{usdg} are purities of b-jets, c-jets and usdg-jets in data. These are results of this fit. In case, we would like to fit non-normalized data, we used

$$N_{all} f_{all} = N_b f_b + N_c f_c + N_{usdg} f_{usdg}, \tag{6.2}$$

where N_{all}, N_b, N_c and N_{usdg} are numbers of all jets, b-jets, c-jets and usdg-jets. Last three are results of this fit. In my analysis I used mainly case in Eq. 6.1.

Next step is to make high-purity sample, and it is common for MC and for data. Motivation for this is to make shapes of distributions even more distinguishable/different. This sample is usually called (b-)tagged. In our cases it is done by rejecting jets with vertices, which are (with some efficiency) expected to come from light decays, or which are reconstructed with big error. This rejection is realised by making cut on some properties

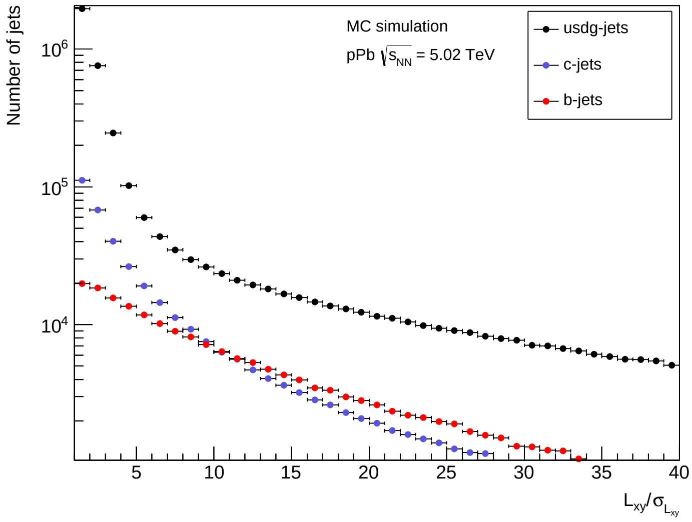


Figure 6.2: Distance of most displaced vertex in jet from primary vertex, divided by its error $L_{xy}/\sigma_{L_{xy}}$ in MC p-Pb collisions.

of secondary vertices. In this case, I made cuts on dispersion of vertex (σ_{vtx}) and on signed impact parameter of vertex (decay lenght of particle decaying in this vertex). By rejecting vertices with small SL_{xy}, vertices reconstructed with big error or coming from short-lived particles are eliminated. In next steps, different these three combinations of cuts were compared, from the loosest one to the most strict:

- $\sigma_{vtx} < 0.04$, $SL_{xy} > 5$,
- $\sigma_{vtx} < 0.02$, $SL_{xy} > 5$,
- $\sigma_{vtx} < 0.02$, $SL_{xy} > 10$.

6.1.2 Fitting by functions

I was trying to fit different MC tagged distributions (cuts $\sigma_{vtx} < 0.04$, $SL_{xy} > 5$, $\sigma_{vtx} < 0.02$, $SL_{xy} > 5$, $\sigma_{vtx} < 0.02$, $SL_{xy} > 10$) of secondary vertex invariant mass by probability distribution functions. For this, MC data from LHC13b4_fix were used. Different distribution functions were tested. Three of them are showed in Fig. 6.3: Exponential, Novosibirsk and Bukin probability distribution functions. Distributions were fit in different ranges, these were tuned to assure convergence of fit and to minimize χ^2/NDF of fit. Results of χ^2/NDF for this three functions, and fitted ranges are in Tab. 6.1.

In Fig. 6.3a, there is exponential fit of distribution after its maximum. These fits where done with good precision, as we can see in Tab. 6.1, but if we want to fit data using Eq. 6.1, it is not feasible. Problem is that three exponential functions have nearly same slopes. Motivation to fit also a "peak" (grow) in distributions came from this problem.

Fit by Novosibirsk distribution is in Fig. 6.3b. This distribution is defined as

$$P_{Nov}(x) = \exp\left(\left(-\frac{1}{2}\right)^{\frac{(\ln q)^2}{\Lambda^2}+\Lambda^2}\right), \tag{6.3}$$

		Exp.	Nov.	Bukin
Range	**min.**	1.0	0.0	0.5
	max.	5.0	5.0	5.0
$\frac{\chi^2}{\mathbf{NDF}}$	**b-jets**	4.23	8.34	3.48
	c-jets	39.23	16.29	15.87
	usdg-jets	0.89	9.19	6.76

Table 6.1: Precisions of fits by exponential (exp.), Novosibirsk (nov.) and Bukin distributions, in specified ranges for b-jets, c-jets and usdg-jets distributions.

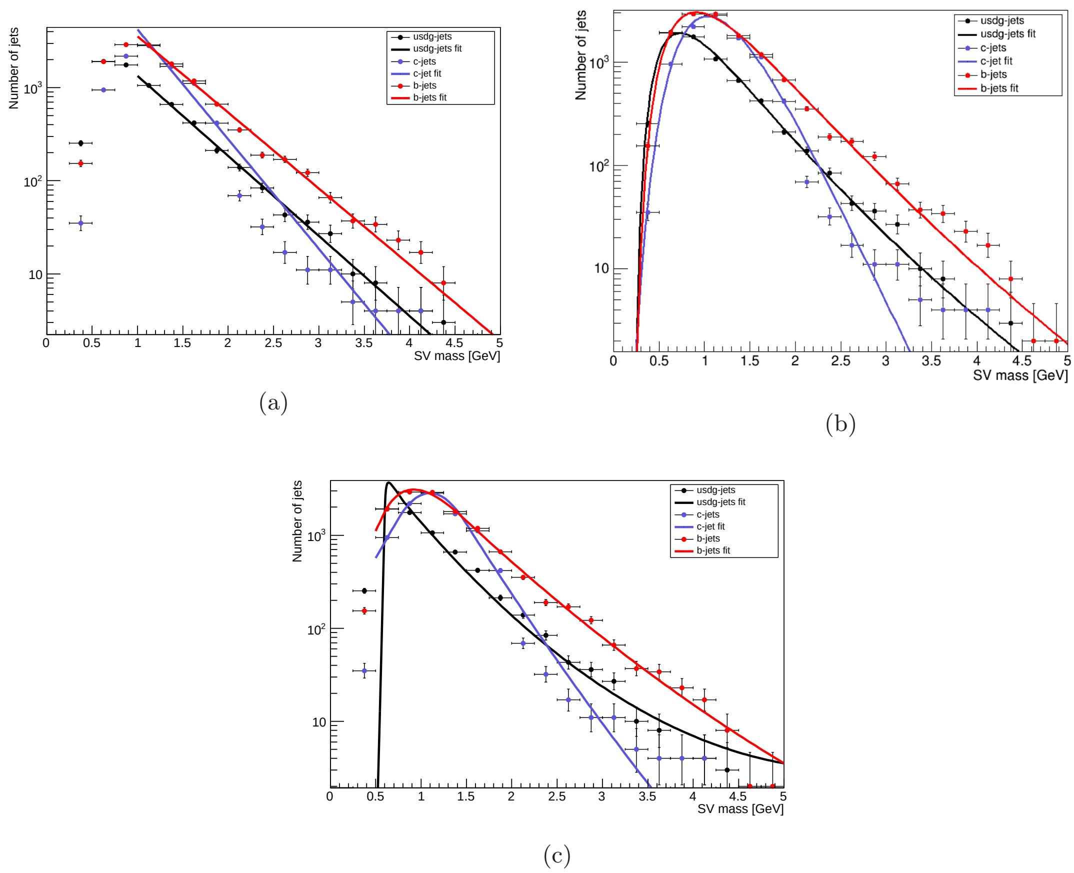


Figure 6.3: Fits to b-tagged SV mass distributions in MC p-Pb collisions: (a) exponential PDF, 0.8 - 5 GeV; (b) Novosibirsk PDF, 0.8 - 5 GeV; (c) Bukin PDF, 0.5 - 5 GeV. Functions resulting from fits compared with real distributions in MC.

where

$$q = 1 + \frac{\Lambda(x - x_0)}{\sigma} \frac{\sinh \Lambda\sqrt{\ln 4}}{\Lambda\sqrt{\ln 4}}, \tag{6.4}$$

x_0 is parameter describing peak position, σ is width of this peak and Λ is parameter describing tail of distribution.

The Bukin function is given by

$$P(x; x_p, \sigma_p, \xi, \rho) = A_p \exp\left(\frac{\xi\sqrt{\xi^2+1}(x - x_1)\sqrt{2\ln 2}}{\sigma_p(\sqrt{\xi^2+1} - \xi)^2 \ln\left(\sqrt{\xi^2+1} + \xi\right)} + \rho\left(\frac{x - x_i}{x_p - x_i}\right)^2 - \ln 2 \right), \tag{6.5}$$

where $\rho = \rho_1$ and $x_i = x_1$ for $x < x_1$ and $\rho = \rho_2$ and $x_i = x_2$ for $x \geq x_2$ and

$$x_{1,2} = x_p + \sigma_p\sqrt{2\ln 2}\left(\frac{\xi}{\sqrt{\xi+1}} \mp 1\right), \tag{6.6}$$

parameter x_p is the peak position, σ_p is the width (FWHM/2.35) and ξ is an asymmetry parameter. Fit of SV mass by this function is shown in Fig. 6.3c.

In all cases, results of fits by functions were not plausible. If the well-fitting functions would be found, it would probably be only some empirical decription of shape of distribution, without any physical reasons. Also search for a function, that decribes physics behind these distributions, is hadrened by fact, that we are fitting tagged samples.

6.1.3 Template fits

Fitting of distributions by analytical functions was not succesful, so we decided to try template fittig method. In this method, shapes of distributions for different jet flavours from MC are taken exactly just as they are, and then real data are fitted as in Eq. 6.1. As shape of MC distributions are taken exactly, we need to suppres ideally all statistical fluctuations (they could modify fitting of data), for this big data samples are needed.

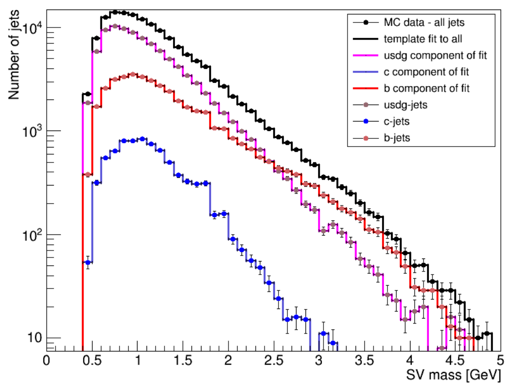
Firstly, we wanted to test, if this method could even work in our case. For this, instead of real data we used MC data, that were only sum of b-jets, c-jets and usdg-jets distributions, taken also as templates (f_b, f_c and f_{usdg} in Eq. 6.1) in this case. At first sight, it could be obvious, that these method should result in purities with good precisions (also compared to "real" purities from MC), but experiences from fitting by functions, especially close slopes of tails of ditributions and resulting disability to fit data, were suasive enough to make to this test.

For this, MC data LHC14g3b were used. We used all three already mentioned cuts to construct b-tagged samples. Results of fits are different jet purities (Eq. 6.1), we compared these with real ones (as MC data were fitted). This comparison is in Tab. 6.2 and fits are shown in Fig. 6.4. As it could be observed, purities are nearly same and errors of fits are very small. Also shape of real and fitted distributions are nearly same. This approves, that template fitting method could work and could be precise.

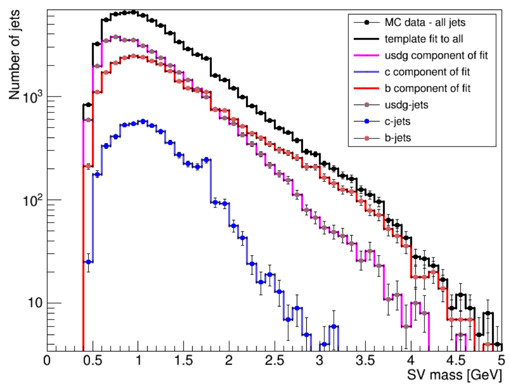
After template fitting method was approved as working in some circumstances, it was also applied on data, so templates (shapes) of MC b-jets, c-jets and usdg-jets distributions were used to fit real data. As MC data, LHC14g3b was taken and for real data minimum bias p-Pb collision were taken. Results of this for three different tagged samples are in Tab. 6.3 and in Fig. 6.5, where, especially for more strict cuts, an obvious lack of data is observed. Actual minimum bias data sample have small statistics, especially for the most

		$\sigma_{vtx} < 0.04$ $SL_{xy} > 5$	$\sigma_{vtx} < 0.02$ $SL_{xy} > 5$	$\sigma_{vtx} < 0.02$ $SL_{xy} > 10$
b-jets	**fit**	0.289 ± 0.007	0.41 ± 0.01	0.70 ± 0.01
	real	0.2898	0.4132	0.7029
c-jets	**fit**	0.06 ± 0.01	0.07 ± 0.01	0.08 ± 0.01
	real	0.0551	0.0746	0.0823
usdg-jets	**fit**	0.66 ± 0.01	0.51 ± 0.01	0.21 ± 0.02
	real	0.6550	0.5121	0.2147

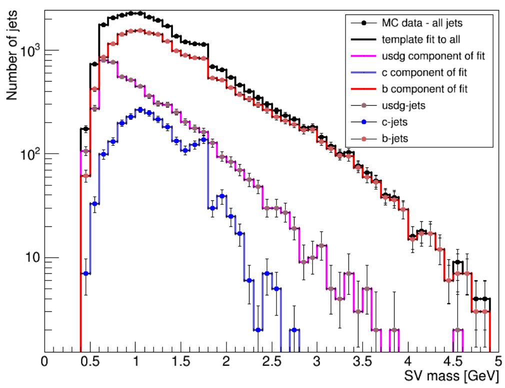
Table 6.2: Purity of b-jets, c-jets and usdg-jets resulting from template fitting of different tagged samples of secondary vertex mass distributions compared to input values of these purities, in MC data from p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV.



(a)



(b)



(c)

Figure 6.4: Template fits to different b-tagged SV mass distributions in MC p-Pb collisions: (a) $L_{xy}/\sigma_{L_{xy}} > 5, \sigma_{\rm vtx} < 0.04$; (b) $L_{xy}/\sigma_{L_{xy}} > 5, \sigma_{\rm vtx} < 0.02$; (c) $L_{xy}/\sigma_{L_{xy}} > 10, \sigma_{\rm vtx} < 0.02$. Functions resulting from fits compared with real distributions in MC.

strict cut, that was used to construct b-tagged sample ($L_{xy}/\sigma_{L_{xy}} > 10, \sigma_{\rm vtx} < 0.02$), that is also partially motivation to use looser cuts on date to construct b-tagged samples.

For this results, there is no data or other results to make comparisons, so we could only work with some preliminary expectations. As it could be seen in Tab. 6.3, errors of resulting purities are small and χ^2/NDF of fits are close to 1. In comparison with Tab. 6.2, b-jet purities are sligthly higher and stay nearly constant for all three cuts. For more strict cuts, purities of c-jets are higher and purities of usdg-jets are smaller, what is in accordance with our expectations (more strict cuts on b-tagged samples reject usdg-jets). In all cases, it is good and positive, that this method converged with such small errors, and that values of resulting purities are changing (for different cuts) approximately same as we expected.

	$\sigma_{vtx} < 0.04$ $SL_{xy} > 5$	$\sigma_{vtx} < 0.02$ $SL_{xy} > 5$	$\sigma_{vtx} < 0.02$ $SL_{xy} > 10$
b-jets	0.46 ± 0.06	0.44 ± 0.08	0.49 ± 0.09
c-jets	0.36 ± 0.08	0.45 ± 0.09	0.48 ± 0.09
usdg-jets	0.18 ± 0.09	0.11 ± 0.12	0.03 ± 0.13
$\boldsymbol{\chi^2}$**/NDF**	0.78	0.97	0.70

Table 6.3: Purities of b-jets, c-jets and usdg-jets, and precision of fit $\frac{\chi^2}{\rm NDF}$ resulting from template fitting of different tagged samples of secondary vertex mass distributions in data from p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV.

Next step could test, if different MC templates results in same b-jet purities. After that, two or more dimensional fits could be tested also. More dimensions means more variable (properties) of secondary vertices (jets), that are used in same time, to results in one value of b-jet purity.

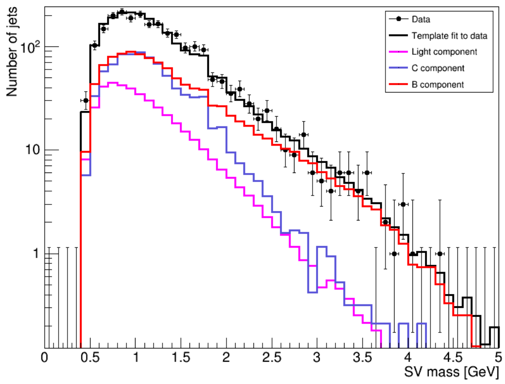
6.2 Secondary vertex algorithm performance in p-Pb collisions

Results showed in this section were presented in form of poster in International Conference on New Frontiers in Physics 2015. Poster, that was created by me and Gyulnara Eyyubova, Ph.D, can be seen in Appendix A.

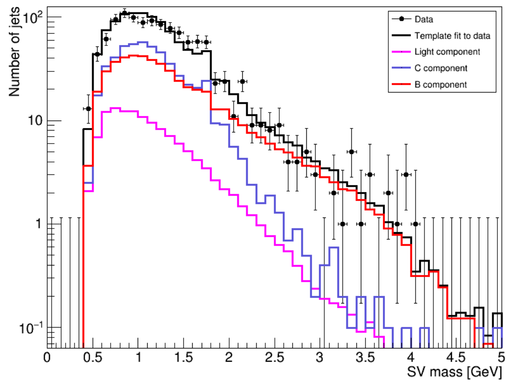
Studies presented in this section are MC based: p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV were generated with PYTHIA and HIJING. Jets were reconstructed with anti-k_T algorithm, $R = 0.4$. Secondary vertices were reconstructed using 3 prolongated tracks, that should have $p_{T,trac} > 1$ GeV/c.

As already described, tagging efficiency of algorithm is influenced by cuts on different dicriminating variables. To study different tagging and mistagging efficiencies of Secondary vertex algorithm, signed flight distance significance SL_{xy} was varied. Results of this study are in Fig. 6.6. In Fig. 6.6a, b-tagging efficiency vs. mistagging rate for different cuts on SL_{xy} and for constant cut on dispersion of vertex $\sigma_{vtx} < 0.02$ cm is shown. It can be observed, that for higher b-jet efficiencies also mistagging rate of c-jets and usdg-jets is rising. Mistagging rate of c-jets is always higher than mistagging rate of usdg-jets and it varies from 2% up to 20%.

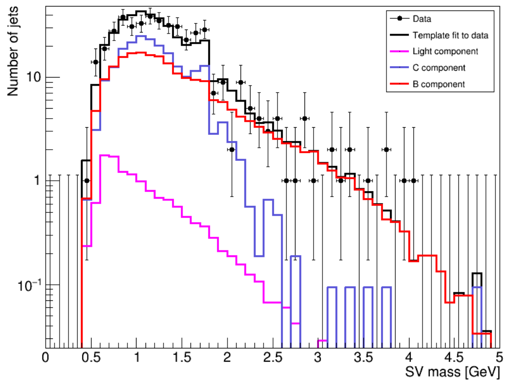
One of cuts in Fig. 6.6a, $SL_{xy} > 10$ was chosen to plot efficiencies vs. transverse momentum of jet. This is showed in Fig. 6.6b. Efficiency of b-tagging is around 20% for



(a)

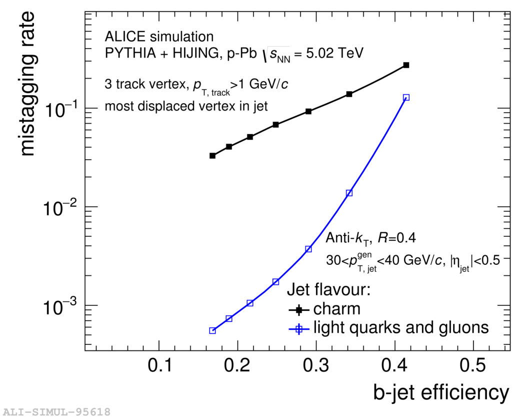


(b)

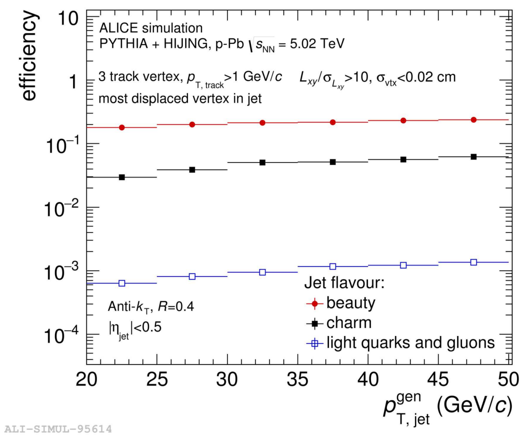


(c)

Figure 6.5: Template fits to different b-tagged SV mass distributions in data from p-Pb collisions: (a) $L_{xy}/\sigma_{L_{xy}} > 5, \sigma_{\text{vtx}} < 0.04$; (b) $L_{xy}/\sigma_{L_{xy}} > 5, \sigma_{\text{vtx}} < 0.02$; (c) $L_{xy}/\sigma_{L_{xy}} > 10, \sigma_{\text{vtx}} < 0.02$.

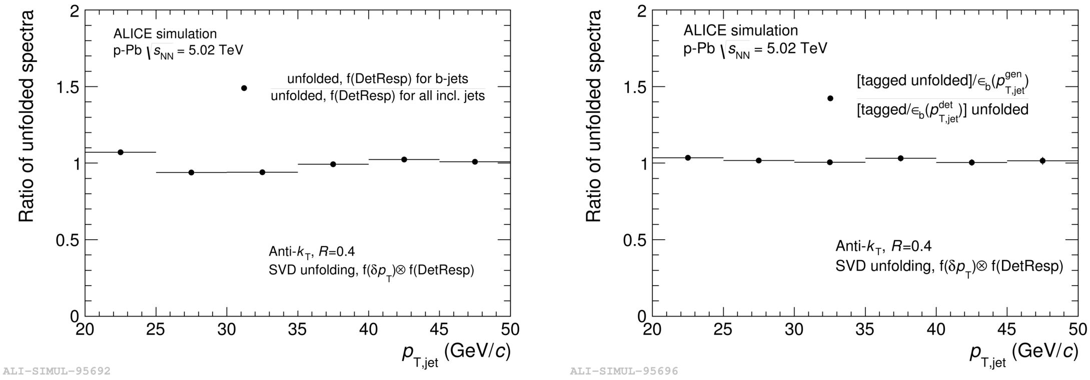


(a)



(b)

Figure 6.6: (a) Efficiency vs. mistagging rate of SV tagging with different cuts for $30 < p_T < 40$ GeV$/c$. The cut on dispersion of the SV vertex is fixed, $\sigma < 0.02$ cm, the cut on signed flight distance significance SL_{xy} is variated. (b) Efficiency of tagging algorithm as a function of jet transverse momentum $p_{T,jet}^{gen}$, for cuts $\sigma < 0.02$ cm, $SL_{xy} > 10$.



(a) (b)

Figure 6.7: (a)Comparison of SVD unfolding of b-jet spectrum with 2 matrices: detector matrix for all inclusive jets and for b-jets. Both matrices are combined with background fluctuation matrix from MC. Ratio of two unfolded results. (b) SVD unfolding for tagged spectrum with combined matrix for 2 scenarios: the measured tagged b-jet spectrum is first unfolded and then corrected for efficiency as a function of p_T^{gen}; the measured tagged b-jet spectrum is first corrected for efficiency as a function of p_T^{rec} and then unfolded. Ratio of two unfolded results.

full displayed $p_{T,jet}$ range, while mistagging rate of c-jets is slowly rising from 3% to 7%. Mistagging efficiency of usdg-jets is suppressed enough, its maximum value is smaller than 0.1%.

6.2.1 Unfolding corrections

Described b-tagging algorithm and its properties is used for b-jet spectrum reconstruction. Important step in this reconstruction is jet unfolding. The measured jet spectrum $m(x')$ is a convolution of the true jet spectrum $t(x)$ and detector response function $A(x, x')$, obtained from MC simulations,

$$m(x') = \int \mathrm{d}x A(x, x')t(x). \tag{6.7}$$

In practice, spectra are histograms and detector response function is matrix (2D histogram). In this analysis, SVD (singular value decomposition) unfolding was used. Two different detector matrices were used: matrix used for all inclusive jets and for b-jets. Both matrices were also combined with background fluctuation matrix from MC. Comparison of unfolding results, done with these 2 matrices is in Fig. 6.7a. There can be seen ratio of them vs. transverse momentum of jet $20 < p_{T,jet} < 50$ GeV/c. Ratio is consistent with 1, so these 2 methods are equivalent and b-jets can be unfolded with matrix for all inclusive jets.

As already mentioned, jet spectra should be corrected also by efficiency of tagging algorithm. Two different scenarios of these corrections were tested. In first one, b-jet spectrum is first corrected for efficiencies and then unfolded, in other one b-jet spectrum is unfolded and then corrected for efficiencies. These scenarios are compared for $20 < p_{T,jet} < 50$ GeV/c, results are in Fig. 6.7b. This ratio is again constant with 1, so it was verified, that the correction order does not give significant differences in the resulting spectrum.

Conclusions

The aim of this work was to introduce b-jet studies, to show results of analysis of b-jets in pp collisions at $\sqrt{s} = 7$ TeV and some steps of actually running analysis of b-jets in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV.

I presented how jets could be detected and found at ALICE experiment and consequently tagged as b-jets. The most used jet-finding algorithm is anti-k_T, for b-tagging in data Simply Secondary Vertex, Track Counting and Jet Probability algorithms are used. MC tagging algorithm of b-jets searches for mother b-quark of B hadron within some distance from jet axis.

Last two chapters of this work showed my results. At first, it were different properties of MC b-jet tagging algorithms in pp collisions. I compared different values of jet cone radius in anti-k_T algorithm, as well as required distance between jet axis and mother b quark or B hadron in b-tagging algorithm. Finally, using these information, I showed that result of b-jet fraction in Pythia data is same as results from CMS experiment.

In the last chapter, I showed part of analysis of PWG-HFCJ group at ALICE experiment. Goal of this analysis is to tag b-jets, to study b-jet fraction and purity in p-Pb collisions. Different variables of jet properties were fitted. We tried to fit distributions of secondary vertex mass by probability distribution functions. This fitting was unsuccessful, so we tried template fitting method. Actual results show, that these method could be used in data.

Actually, template fitting method is tested in more details. We will try to test it also in data from pp collisions. Also other b-jet analysis are running simultaneously at ALICE, for example b-jet unfolding or b-tagging by high energy electrons. In my following studies, I will focus on studies of b-jet tagging efficiencies for different values of discriminators of tagging algorithms in pp collisions.

Appendix A

Poster presentation from International Conference on New Frontiers in Physics 2015

International Conference on New Frontiers in Physics 2015 was hosted in the Conference Center of the Orthodox Academy of Creta in Kolymbari. It started on August 23 and ended on August 30. 23 different posters were presented. One of them was "Performance of the ALICE secondary vertex b-tagging algorithm", that was created by me and Gyulnara Eyyubova, Ph.D.

To acces to full quality of poster, see in Ref. [18].

Performance of the ALICE secondary vertex b-tagging algorithm

Gyulnara Eyyubova[1,2], Lukáš Kramárik[1] on behalf of the ALICE collaboration

1) FNSPE, Czech Technical University in Prague

2) SINP MSU, Russia

4th International Conference on New Frontiers in Physics

23-30 August 2015, Crete, Greece

Motivation

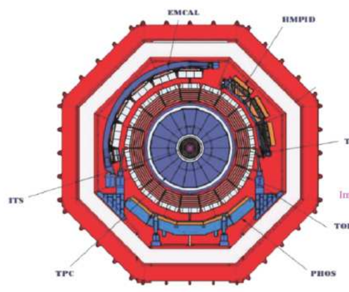
- Determine b-quark production via the measurements of beauty(b)-jets.
  Jets vs. heavy-flavor hadrons: access the kinematics of hard scattering in an unbiased way.
- Color and mass dependence of parton energy loss in the Quark-Gluon Plasma (QGP).

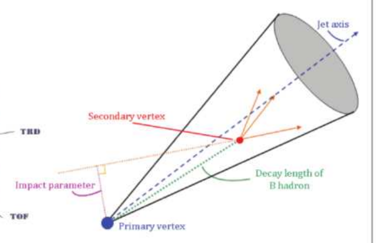
Analysis steps

- Jet-finding: jet reconstruction with charged tracks.
- Jet b-tagging: exploit long lifetime and large mass of beauty hadrons.
- Corrections: correction of jet transverse momentum p_T (or jet energy) for background and detector response (unfolding), corrections for b-tagging efficiency and charm/light jet contamination.
- *The studies presented in this poster are MC based.*

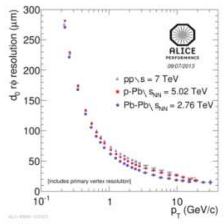
Secondary vertex (SV) tagging algorithm

Benefits from long lifetime ($c\tau$ ~ 500 μm) of beauty hadrons.

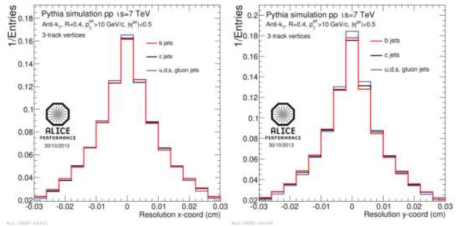




The Inner Tracking System (**ITS**) and Time Projection Chamber (**TPC**) are used for tracking and secondary vertex reconstruction. For all collision systems ALICE measures the track impact parameter d_0 with a resolution better than 70 μm for p_T > 1 GeV/c.



Track impact parameter resolution in the transverse plane vs p_T.



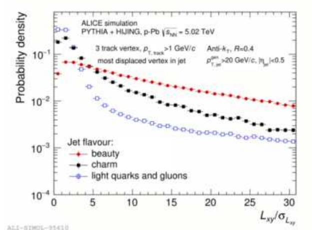
3-prong vertex resolution (in the xy plane).

- Jet reconstruction:
  FastJet Anti-k_T [1], R = 0.4
  $p_{T,\,track}$ > 0.15 GeV/c
- 3-prong SV are reconstructed using tracks in a jet $p_{T,\,prong}$ > 1 GeV/c
- Discriminating variables:
  significance of signed SV distance of flight L_{xy}/σ_{Lxy} in a transverse plane and SV dispersion σ_{vtx}.

$$L_{xy} = L'_{xy} \times \mathrm{sign}(L_{xy} \cdot p_{T,jet})$$

$$\sigma_{vtx} = \sqrt{d_1^2 + d_2^2 + d_3^2}$$

$d_{1,2,3}$ are the distances of tracks from SV.



Signed flight distance significance of SV. Jets of different flavor defined in MC based in presence of B- (D-) hadron.

[1] M. Cacciari and G. P. Salam, Phys. Lett. B 641 (2006) 57.

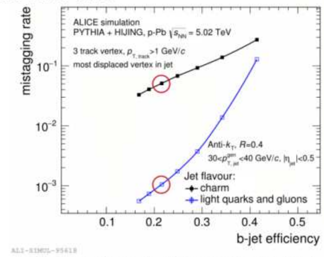
Secondary vertex algorithm performance in p-Pb collisions

$$\varepsilon_b(p_T) = \frac{dN_b^{tagged} / dp_T}{dN_b / dp_T}$$

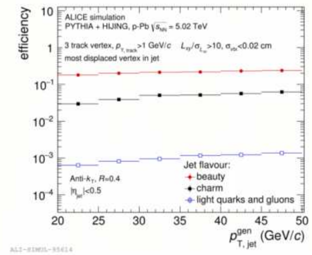
Tagging efficiency: the ratio of properly tagged jets vs all jets of a given flavor.
It is obtained via MC studies.

Cut variation → different tagging and mistagging effiiencies.

The goal is to find the working point with high purity and reasonably high efficiency at the same time.



Tagging efficiency and mistagging rate with different cuts on L_{xy}/σ_{Lxy} of SV for 30 < $p^{gen}_{T,\,jet}$ < 40 GeV/c.



(Mis)Tagging efficiencies of SV algorithm for L_{xy}/σ_{Lxy} > 10 cm and σ_{vtx} < 0.02 cm.

Unfolding corrections

$$m(x') = \int dx A(x;x')t(x)$$

The measured spectrum *m(x')* is a convolution of a true spectrum *t(x)* and detector responce function *A(x;x')*. The true jet spectrum is found via unfolding procedure.

The Singular Value Decomposition (SVD) unfolding was used [2].

- The constant background density ρ_{CMS} (calculated with CMS method [3]) is subtracted from $p_{T,\,jet}$:

$$\rho_{CMS} = \mathrm{median}\left\{\frac{p_{T,i}}{A_i}\right\} \cdot C$$

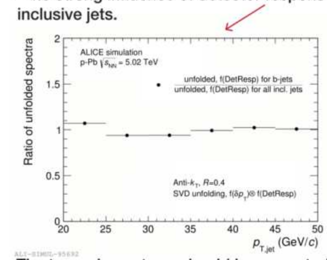
$p_{T,i}$ are soft clusters, found by FastJet k_T, A_i is the jet area, C is the correction factor for the empty clusters.

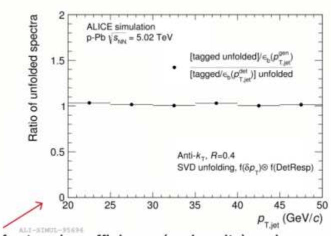
- The background fluctuations δp_T are calculated with MC (HIJING) via random cone method:

$$\delta p_T = \sum_i p_{T,i} - \rho \cdot A_{cone}$$

The sum is over track $p_{T,i}$ in a cone.

- MC b-jet p_T spectrum is unfolded with the combined matrix: detector response and background fluctuations matrix. Same results when unfolding with detector response matrix of inclusive and b-jets
  → no strong influence of detector response in different fragmentation patterns of b- and inclusive jets.





- The tagged spectrum should be corrected for tagging efficiency (and purity) and unfolded. It was verified with p-Pb simulations that the correction order does not give significant differences in the resuting spectrum.

[2] H. Hoecker, V.Kartverlishvili, Nucl. Instrum. Meth. A372 (1996) 469.

[3] CMS collaboration, JHEP 1208 (2012) 130.

Prospects with LHC/ALICE upgrades

- Upgrade of ITS (2018):
  improvement of the track impact parameter resolutions by a factor 3 (6) in the transverse (longitudinal) direction
  → better light flavor rejection in b-tagging analysis.
  [CERN-LHCC-2012-012, CERN-LHCC-2012-013]
- ALICE read-out and LHC upgrades (2018):
  higher integrated luminosities: ~10 pb^{-1} for pp collisions at $\sqrt{s}$ = 14 TeV and ~10 nb^{-1} for Pb-Pb collisions at $\sqrt{s_{NN}}$ = 5.5 TeV required by the ALICE upgrade program
  → high precision heavy-flavor results, posibility to study on-line b-jet tagging.

Summary and ongoing studies

- The tagging cuts for SV algorithms are optimized for keeping beauty efficiencies as large as possible and at the same time charm and light-jet contamination small.
- Corrections for background and background fluctuations as well as for detector response are implemented. It was found that the b-jet spectrum can be corrected with a detector response matrix for all inclusive jets.
- The order of corrections (tagging efficiency vs unfolding) gives compatible results.
- MC and data-driven estimation of tagging purity is under study.
- Study of track cuts and selection for SV reconstruction in order to obtain higher purity is ongoing.

european social fund in the czech republic · EUROPEAN UNION · MINISTRY OF EDUCATION, YOUTH AND SPORTS · OP Education for Competitiveness

INVESTMENTS IN EDUCATION DEVELOPMENT

Acknowledgment

This work was supported by the European social fund within the framework of realizing the project „Support of inter-sectoral mobility and quality enhancement of research teams at Czech Technical University in Prague", CZ.1.07/2.3.00/30.0034 and by Grant Agency of the Czech Technical University in Prague, grant No. SGS13/215/OHK4/35/14.

Bibliography

[1] L. Hsu, J. P. Hsu, *The physical basis of natural units and truly fundamental constants,* Eur. Phys. J. Plus (2012) 127:11.

[2] K.A. Olive et al. (Particle Data Group), *2014 Review of Particle Physics,* Chin. Phys. C, 38, 090001 (2014).

[3] R. Aaij et al. (LHCb collaboration), *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi K^- p$ decays,* Phys. Rev. Lett. 115, 072001 (2015).

[4] S. Sarkar et al., *Lecture notes in Physics 785,* The Physics of the Quark-Gluon Plasma, Springer (2010).

[5] The ALICE Collaboration et al, *The ALICE experiment at the CERN LHC,* 2008 JINST 3 S08002.

[6] W. Yu, for the ALICE TPC Collaboration, *Particle identification of the ALICE TPC via dE/dx,* Nuclear Instruments and Methods in Physics Research A 706 (2013) 55–58.

[7] R. Bellwied, *ALICE EMCal Physics Performance Report*, arXiv:1008.0413 (2010).

[8] A. SHABETAI, *Recent results on jet physics from ALICE at the LHC*, Journal of Physics: Conference Series 458 (2013) 012025.

[9] Y.L. Dokshitzer, D.E. Kharzeev, *Heavy quark colorimetry of QCD matter*, Phys. Lett. 519, 199 (2001).

[10] CMS Collaboration, *Identification of b-quark jets with the CMS experiment*, JINST 8 (2013) P04013.

[11] CMS Collaboration, *Measurement of the b-jet to inclusive jet ratio in PbPb and pp collisions at $\sqrt{s_{NN}} = 2.76$ TeV with the CMS detector*, CMS PAS HIN-12-003 (2012).

[12] CMS Collaboration, *"Nuclear modification factor of jets at high p_T in PbPb collisions $\sqrt{s_{NN}} = 2.76$ TeV with the CMS detector*, CMS PAS HIN-12-004 (2012).

[13] S. LaPointe on behalf of the ALICE collaboration, *Prospects for beauty-jet measurements with ALICE*, poster presentation, Quark Matter 2014, Darmstadt.

[14] T. Sjostrand, S. Mrenna, P. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP 0605 026 (2006).

[15] M. Cacciari, G. P. Salam, G. Soyez, *FastJet user manual*, Eur.Phys.J. C72 (2012) 1896.

[16] S. Agostinelli, et al., *Geant4 - a simulation toolkit*, Nucl.Instrum.Meth. A506 (2003) 250-303.

[17] CMS Collaboration, *Inclusive b-jet production in pp collisions at $\sqrt{s_{NN}} = 7$ TeV*, JHEP04(2012)084.

[18] G. Eyyubova, L. Kramárik on behalf of the ALICE collaboration, *Performance of the ALICE secondary vertex b-tagging algorithm*, poster presentation, ICNFP 2015, Crete (https://indico.cern.ch/event/344173/session/22/contribution/404).