Czech Technical University in Prague

Faculty of Nuclear Sciences and Physical Engineering

Bachelor's Thesis

Study of Identification of Jets Containing Heavy Quark

Prague, 2008 **Author: Michal Vajzer**

Title: Study of identification of jets containing heavy quark
Author: Michal Vajzer
Specialization: Nuclear Engeneering
Sort of project: Bachelor's Thesis
Supervisor: Mgr. Jaroslav Bielčík, Ph.D., Katedra fyziky, FJFI, ČVUT v Praze.
Abstract:

Jets have become important source of information in study of events at high energies, where number of particles produced and observed is immense. In study of quark-gluon plasma, hot nuclear medium created from energetic collisions of heavy ions, jets will be used for example to measure energy loss of partons, when travelling through this medium. In theoretical works about QGP an eminent role plays difference between gluons, light and heavy quarks. This work deals with methods applied in reconstruction of these jets, mainly conebased algorithm, k_T clustering algorithm and gives information about algorithm used in jet-search at ALICE. Furthermore it discusses possible ways of identification of heavy quarks utilising their unique properties.

Key words:
jet reconstruction, quark tagging, heavy quark, heavy ion collision, Pythia

Názov práce: Štúdium identifikácie jetov častíc obsahujúcich ťažký kvark
Autor: Michal Vajzer
Abstrakt:

Jety sa stali dôležitým zdrojom informácii pri štúdiu vysokoenergetických zrážok, kde počty produkovaných a pozorovaných častíc je obrovský. Pri štúdiu kvark-gluónovej plazmy, horúceho jadrového prostredia vytvoreného energetickými jadro-jadrovými zrážkami, sa jety použijú k zisťovaniu energetických strát partónov prechádzajúcich týmto horúcim prostredím. V teoretických prácach o kvark-gluónovej plazme hrajú rozdiely medzi gluónmi, ľahkými a ťažkými kvarkami významnú úlohu. Táto práca sa zaoberá metódami používanými pri rekonštrukcii týchto jetov, hlavne kužeľovo založenými algoritmami, k_T zhlukujúcim algoritmom a zahŕňa informácie o algoritme pre hľadanie jetov na ALICE. Taktiež diskutuje možne cesty identifikácie ťažkých kvarkov využitím ich význačných vlastností.

Kľúčové slová:
rekonštrukcia jetov, identifikácia kvarkov, ťažký kvark, jadro-jadrová zrážka, Pythia

Acknowledgment
I would like to thank my supervisor, Jaroslav Bielčík, for his patience and invaluable help during creation of this work. My thanks as well belongs to my family for their support up until now.

Prehlásenie

Prehlasujem, že som svoju bakalársku prácu vypracoval samostatne a použil som iba podklady (literatúru, projekty, SW atd.) uvedené v priloženom zozname.

Nemám závažný dôvod proti použitiu tohto školského diela v zmysle §60 Zákona č.121/2000 Sb., o práve autorskom, o právach súvisiacich s právom autorským a o zmene niektorých zákonov (autorský zákon).

V Prahe dňa 30. júna 2008

podpis

Table of Contents

List of Figures

List of Tables

Introduction

Development of Large Hadron Collider in CERN at Geneva opens new possible fields of study in high energy nuclear physics, accessible only at very high collision energies. In all fields, jets come as important observable. This includes search for Higg's boson and it's decays into $b\bar{b}$ and $t\bar{t}$ pair, where in hadronic decays multi-jet production study is important, as well as in search for R-parity violations in SUSY. Signatures of jets are also used in testing of perturbative QCD at high energy regimes, constrainment and measurement of coupling constant and momentum distributions in hadrons with greater precision.

Among research fields, identification of heavy quarks is necessary, when study of quark-gluon plasma (QGP) is involved. QGP is hot state of nuclear matter, in which quarks and gluons are proposed to exist in deconfined state. This state of matter was naturally present after Big Bang, before quarks formed hadrons. We are able to create it by energetic collisions of heavy nuclei.

Study of jets coming from heavy quarks may give us information about colour and flavour dependence of parton energy loss and nuclear modification of partons' fragmentation functions.

Aim of this work is to discuss methods used in identification of jets originating from heavy quark.

Chapter 1

Particle physics and jets

1.1 Standard Model

Standard model is unifying theory trying to explain phenomena of particle physics in terms of properties and interactions of basic particles by utilization of quantum field theory.

We know three types of fundamental particles, these are leptons, quarks and gauge bosons. These particles interact through weak, electromagnetic, strong and gravitational interaction. Basic properties of these interactions is summarised in Table 1.1, where gravitational interaction is excluded.

Interaction	El.-mag.	Weak	Strong
Acts on	El. charge	Flavour	Colour charge
Gauge Boson (G.B.)	γ	Z^0; $W^{\pm}$	g
Mass of G.B. $[GeV/c^2]$	0	91.2; 80.4	0
Charge of G.B.	$\emptyset$	electric	colour
Relative strength	10^{-2}	10^{-9}	1

Table 1.1: Table of fundamental interactions and their basic properties.

Leptons are fermions not experiencing strong force. They come in three generations, in Table 1.2 summarising their masses and charges, they are separated into a different rows. Overally there are six leptons and for every one of them there exists one antiparticle. Every generation shows conservation of specific lepton number in every reaction. This is *electron number* for first generation containing electron, electron neutrino and their antiparticles. For second generation it is *muon number*

and *tau number* for third. Specific lepton number is calculated by formula (1.1), where *N(Y)* is number of particles Y and x from the formula may be either e, μ or τ.

$$L_x = N(x^-) - N(x^+) + N(\nu_x) - N(\overline{\nu_x}) \tag{1.1}$$

Concerning stability of leptons, only ones that decay are muon and tau. Muon decays purely leptonically, on the other hand, most of tau's the decay modes involve hadrons. Neutrinos are also least reacting particles from lepton family.

LEPTONS [spin = 1/2]					
Name and symbol		Mass [Gev/c^2]	El.charge [C]	Antiparticle and symbol	
electron	[e]	0.000511	-e	positron	[e^+]
electron neutrino	[ν_e]	$< 10^{-8}$	0	electron antineutrino	[$\overline{\nu_e}$]
muon	[μ]	0.106	-e	antimuon	[μ^+]
muon neutrino	[ν_μ]	< 0.0002	0	muon antineutrino	[$\overline{\nu_\mu}$]
tau	[τ]	1.7771	-e	antitau	[τ^+]
tau neutrino	[ν_τ]	< 0.02	0	tau antineutrino	[$\overline{\nu_\tau}$]

Table 1.2: Table of leptons with their basic properties.

Quarks are constituent fermions feeling every type of interaction. Quarks are bound either in triplets, create baryons, or are bound in pair with other anti-quark, creating mesons, both of which fall into category called hadrons. Basic properties of quarks are summarised in Table 1.3.

QUARKS [spin = 1/2]					
Generation	Name and symbol		Mass [Gev/c^2]	El.charge [C]	Flavour
1st	up	[u]	0.003	2/3e	$\emptyset$
1st	down	[d]	0.006	-1/3e	$\emptyset$
2nd	charm	[c]	1.3	2/3e	C = +1
2nd	strange	[s]	0.1	-1/3e	S = -1
3rd	top	[t]	175	2/3e	T = +1
3rd	bottom	[b]	4.3	-1/3e	B = -1

Table 1.3: Table of quarks with their basic properties.

In nuclear reactions, similarly to leptons, are conserved certain values. Every hadron has *baryon number* that is zero for mesons, one for baryons and minus one for anti-baryons. Of course we assume baryon number equal to zero for leptons as well.

Thanks to the observation of baryons like Δ^{++} with charge +2e, quark content {uuu} and fact that quarks are fermions, Pauli's exclusion principle implies existence of another quantum number, which was designated *colour*. Quark's colour may be of three states: *red*, *blue* and *green*. All other fermions in Standard model are assumed colourless.

Quantum field theory resulting from requiring Standard model Langrangian to be invariant to $SU(3)_{color}$ transformations is called Quantum Chromodynamics (QCD). In order to maintain invariance, it was necessary to introduce bi-coloured, massless gauge bosons called gluons, carriers of strong interaction. Because of properties of SU(3) symmetries of QCD, gluons interact with one another. Properties of colour field lead to two features of QCD different from QED: confinement and asymptotic freedom.

Confinement refers to fact that quarks have never been observed as free particles, in other words, no free colour-charged particle exist. This is consequence of long-distance behaviour of QCD coupling constant. Supposedly in mesons, if we try to separate quarks strength of force between quarks increases with separation. So to free quark is required infinite energy. But before that, when energy in colour field between quarks is sufficient, quark-antiquark pair would be created in vacuum, thus resulting with two mesons.

Asymptotic freedom refers to short-distance regime in QCD, when coupling constant gets weaker with decreasing distances, i.e. at large energy scales. At large enough energy scale, quarks behave mostly like free particles and are treated with perturbative method. This regime is experimentally explored by collisions of hadrons at high energies.

1.2 Evolution of jet

Jets are observable secondary particles in highly collimated form produced in hard-scattering collisions at high enough $\sqrt{s}$, energy at centre of mass, when there is high momentum transfered. QCD definition of jet states that jet is a consecutive cascade of partons from primary parton. Mainly in collisions of nuclei they are difficult to separate from underlying event. For the proper analysis of jets, it is necessary to understand particular steps taken in their creation, from hard-scattering, process in which are created partons with high transverse momentum p_T, through fragmentation and hadronisation processes. All these processes are depicted in the Figure 1.1.

First step in analysis and further modeling of collision of A and B producing C & D , is to calculate cross-section of this process. It is schematically shown as formula 1.2.

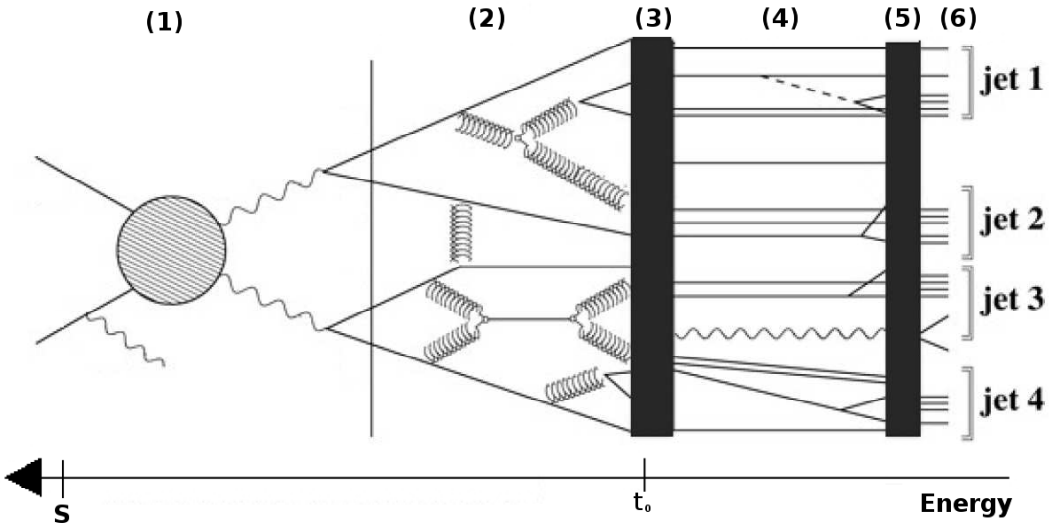


Figure 1.1: Schematic depiction of jet evolution from hard-scattering with initial state radiation (1), followed by fragmentation, i.e by parton showering (2) up to energy scale t_0 when hadronisation occurs (3). Created hadrons undergo decay (4). Final state particles are measured in detector (5) and ultimately jets are reconstructed (6). This scheme was taken over from [1].

$$\sigma(AB \to CD) = \sum_{a,b} f_a^A(x_a) f_b^B(x_b) \otimes \hat{\sigma}(ab \to cd) \otimes F_c^C(z_c) F_d^D(z_d) \tag{1.2}$$

Here $\hat{\sigma}(ab \to cd)$ is parton level cross-section for a and b to produce c and d, $f_i^j(x_i)$ is parton distribution function for parton i in beam particle j with fraction z_i of beam particle's momentum. $F_k^l(x_k)$ is fragmentation function, it parametrises likelihood of producing hadron l with fraction z_k of momentum of final state parton k. For example, in collisions of two protons, i.e. A and B are protons, each to be composed of 3 valence quarks and gluon field, thus $f_u^p(x_u)$ is probability to find u-quark with fraction of protons momentum equal to x_u. On the other hand, fragmentation function characterises what will be created from products of scattering. Together with knowledge of parton level cross-section for all possible interaction, we are able to find likelihood of interaction of pp producing back-to-back jets originating from B-hadrons.

In beam hadrons, we cannot calculate way in which partons evolve, but we are able to parametrise they behaviour with set of distribution functions for every type of parton, to be found with given part of hadron's momentum. These distributions are called *parton distribution functions* and are found by global fit on data taken from experiments sensitive to different x. Their study is necessary, because at large

energy scale, not only constituent quarks have fraction of hadron's momentum, but even the gluon field carries non-negligible momentum. Secondly, they are used in monte-carlo simulations to calculate cross-sections as seen in schematic formula 1.2.

The process of a single quark or gluon generating a jet of particles in a high-energy QCD interaction is referred to as *fragmentation*. It involves transition through energy regime, that is untreatable by perturbative methods, therefore we may divide whole process into two stages, perturbative and non-perturbative.

In the first one, emission of gluons from quarks and splitting of gluons into $q\bar{q}$ pairs is included. Second part is parametrised by number of phenomenological models divided into two categories. Independent fragmentation approximates each final state parton separately, on the other hand string/cluster fragmentation uses colour flow information to account for correlations between partons. The second stage is referred to as *hadronisation*.

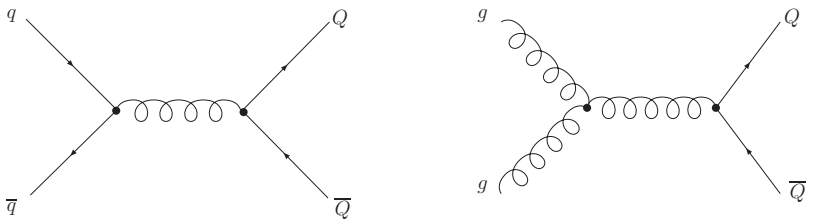
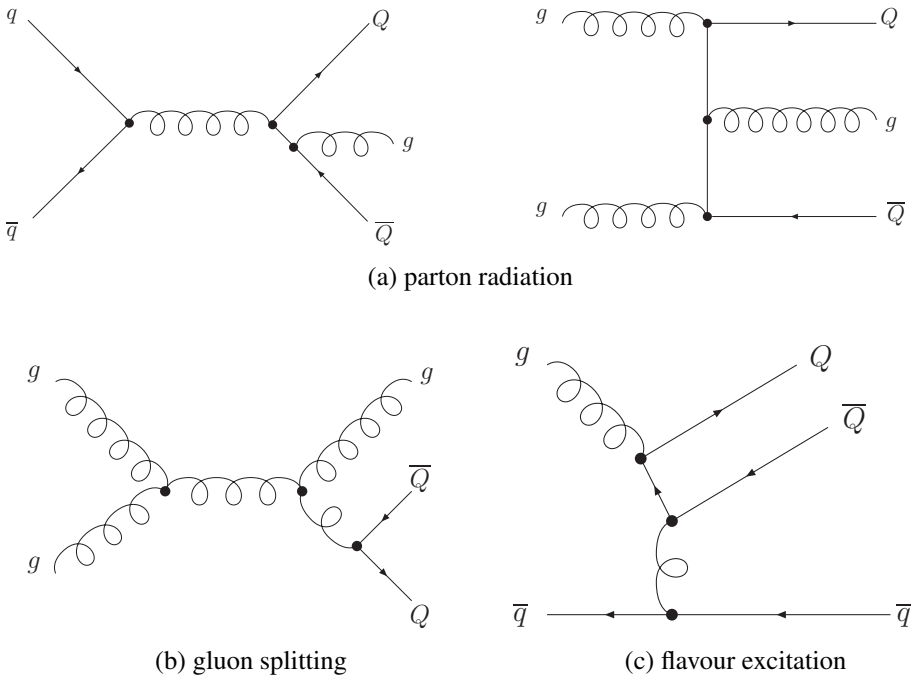


Figure 1.2: Feynman diagrams of leading order processes.

To calculate cross-section of final partons in simulations, we use leading order (LO) calculations to estimate $\hat{\sigma}(ab \rightarrow Q\bar{Q})$. This LO calculation is shown in the Figure 1.2 depicting direct creation processes from quark-antiquark pair and pair of gluons.

Results of this calculations are insufficient and need to be corrected. This is done in two ways. First uses *next-to-leading order*(NLO) calculations contributing to cross-section of final partons, including processes shown in the Figure 1.3, like gluon or quark radiation in Figure 1.3a, as well as gluon splitting in Figure 1.3b and flavour excitation in Figure 1.3c.

An alternate to NLO calculations is *parton shower model*. It is utilised as basis for many Monte Carlo programs. Key element to this alternative is showering algorithm. Parton shower model deals with collision according to chronological order, firstly one parton, taking role in 2-to-2 collision, in incoming beam particles' showers from each beam are taken, generating two final state partons. Final state partons' showers are estimated with addition to remaining partons of the beams. Finally, confinement effects of strong interaction take place and force colour partons



(a) parton radiation

(b) gluon splitting

(c) flavour excitation

Figure 1.3: Next-to-leading order processes.

into colourless hadrons.

1.3 Particle physics experiments

For a search for new particles, new data and for testing new theories in particle physics, a lot of resources have been invested into building a large experimental apparatus accelerating particles to high energies so that parton interaction may take place. In this section, aim and results of several large experiments are discussed.

First questions dealing with ordinary hadronic matter, what is its limit, when separate hadrons do not regain their identity and what are limits of confinement, i.e. whether quarks can be freed from hadrons, were answered by Brookhaven's AGS and RHIC experiments, as well by CERN's SPS.

Evidence for deconfinement gave SPS, fix target heavy ion experiment colliding ions of lead with $\sqrt{s_{NN}} = 17$ GeV, with enhanced production of strange and multi-strange particles with respect to extrapolations from proton-proton collisions. Another evidence was suppression of J/Ψ meson with respect to extrapolation from proton-proton collisions. These two facts were predicted in 80's.

RHIC, with its $\sqrt{s_{NN}}$ = 200 GeV for central Au-Au collisions, has nuclear density of matter well above 1 GeV/fm^3 necessary to produce QGP. There is increased cross-section of hard processes producing energetic partons, allowing to form new observables. PHENIX and STAR experiments measured high-p_T particle production, discovering that it is suppressed in central collisions compared to proton-proton collisions. Also it was observed that back-to-back correlation of high-p_T hadrons disappears with increasing centrality of a collision, leading to absorption of 1 or 2 jets. This was not observed in collisions of deuteron-gold and is generally interpreted as a consequence of parton energy loss in dense nuclear matter.*

Role of LHC is to study strongly interacting medium in conditions of high density and temperature. With its $\sqrt{s_{Au-Au}}$ = 5.5 TeV and $\sqrt{s_{pp}}$ = 14 TeV, energy density of QGP, ϵ, produced is high enough to allow us to study its properties systematically. Determination of ϵ is done by $\frac{dN_{ch}}{dy}$ in central collisions, i.e. number of charged particles produced per unit of rapidity.

Deconfinement produced in LHC is much more similar to theoretical point of view. Large $\sqrt{s}$ results in large ϵ, thus giving large initial temperature that extend lifetime and volume of deconfined medium,freeze-out temperature is around 170 MeV. Large number of gluons favours energy and momentum exchange, thus thermal equilibrium of medium is reached sooner. Overally, these lead to hotter, larger and long-living drops of QCD plasma. Comparison of parameters in several experiments is in Table 1.4.

Parameters		SPS	RHIC	LHC
$\sqrt{s_{NN}}$	[GeV]	17	200	5500
$\frac{dN_{gluon}}{dy}$		≃450	≃1200	≃5000
$\frac{dN_{charged}}{dy}$		400	650	≃3000
initial T	[MeV]	200	350	>600
ϵ	[GeV/fm^3]	3	25	120
freeze-out V	[fm^3]	$\sim 10^3$	$\sim 10^4$	$\sim 10^5$
life-time	[fm/c]	<2	2-4	>10

Table 1.4: Comparison of parameters characterising nucleus-nucleus collisions at different energy regimes adopted from [2].

At LHC, *high-density parton distributions* are expected to dominate parton production, as well as *hard processes* shall significantly contribute to total nucleus-nucleus cross-section, thus it will be qualitatively new regime to study. Low-x

*More about results from first years of RHIC program have been summarised in *"White papers"*, [12]-[15].

range accessible, x stands for Bjorken's x, will be valuable, because charm and bottom production cross-sections at LHC are significantly affected by parton dynamics in small-x regime. For a comparison, values of x accessible for heavy quarks for different energy regimes are in the Table 1.5.

Machine	SPS	RHIC	LHC	LHC
System	Pb-Pb	Au-Au	Pb-Pb	pp
$\sqrt{s_{NN}}$	17GeV	200GeV	5.5TeV	14TeV
$c\bar{c}$	$x \simeq 10^{-1}$	$x \simeq 10^{-2}$	$x \simeq 4 \cdot 10^{-4}$	$x \simeq 2 \cdot 10^{-4}$
$b\bar{b}$	-	-	$x \simeq 2 \cdot 10^{-3}$	$x \simeq 6 \cdot 10^{-4}$

Table 1.5: Bjorken x values correspondent to charm and bottom production at central rapidity and $p_t \to 0$, taken over from [5].

Charm and bottom production allows us to investigate mechanism of heavy quark production, propagation and at low p, hadronisation in hot, dense medium formed in high energy nucleus-nucleus collisions, it gives information about parton densities. Cross-section of charm and bottom quark production are needed for further study, as mentioned before, they are significantly affected by parton dynamics in small-x regions. Reason, why are heavy partons good probe of QGP medium is that they are produced in early stage of collision in primary partonic scattering thus their production is unaffected by medium. Their cross-section may be calculated with perturbative approach and they are expected to have significant energy losses thus giving information about mechanisms of energy loss.

Cross-section of hard process is affected by two types of effects, initial state and final state. Initial state effects of cross-section is *nuclear shadowing*, which is modification of PDF in nucleus due to gluon recombination at small x, leading to depletion of parton density with respect to proton densities. It affects cross-section in a way dependant on size and energy of nucleus, but not on the medium. On the other hand, final state effects are induced by medium. They won't be correlated to initial state effects and depend on properties of medium, thus giving information about it.

Chapter 2

Jet reconstruction

Jets are defined as cascades of partons emitted from initial parton followed by fragmentation. Reconstruction of jets by itself has varying level of difficulty depending on type of collision. For proton-proton, or proton-antiproton, it is much simpler than for collisions of nuclei, where underlying event plays significant role, because number of charged particles per unit of rapidity is expected to be several thousands.

In case of heavy ion collision, jet structure modifications are used to study properties of medium. Initial partons traveling through dense colour medium, where fragmentation function is changed, are expected to loose energy through collisional loss and induced gluon radiation, significant in QGP. Therefore jets are quenched. This jet quenching is proposed to probe properties of hot dense nuclear matter.

2.1 Proton-proton collisions

From standpoint of calorimeters, detectors measuring energy deposited by particle flying through it, jet is shower of large number of tracks seemingly close together in space with energy deposition spread over a similar area as tracks. It is convenient to look on energy deposited in energy towers in calorimeters as in Figure 2.1c, where calorimetric towers surrounded by circle represent energy deposited by particles from jet. Grouping together energy towers, i.e. sum of energy lost by particles in all layers of given calorimeter at given coordinates, with high energy lying in cone defined by azimuthal angle and pseudo-rapidity, is base of cone algorithm. Another type of algorithm is k_T algorithm. This groups together towers nearby in momentum space.

2.1.1 Cone algorithm

There exist two main types of cone algorithms, seed and seedless. Both types utilise cone in pseudo-rapidity and azimuthal angle. Segmentation of calorimeter in this system of coordinates is depicted on the Figure 2.1a, where θ is polar angle from beam direction. Schematic distribution of transverse energy, in space defined by η-ϕ, detected in detector is depicted in the Figure 2.1c. Here are possible jets marked with circle surrounding them. Top right jet is the most significant, but other jets may or may not be reconstructed, depending on method applied, i.e. what parameters and particular algorithm is used.

Energy towers inside of cone of radius, defined as $R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, i.e. circle in η-ϕ of detector from coordinates of centre, add up to total energy of jet, similarly it is valid for transverse energies, defined by formula 2.1,

$$E_T^J = \sum_i E_T^i sin(\theta_i) \tag{2.1}$$

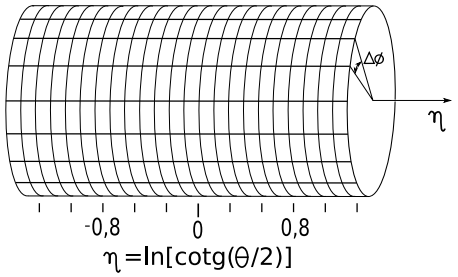
where i goes through all energy towers of cone. Weighted coordinates are defined as $\eta^J = \sum_i \frac{E_T^i \eta^i}{E_T^J}$ for pseudo-rapidity and for azimuthal angle it is defined as $\phi^J = \sum_i \frac{E_T^i \phi^i}{E_T^J}$. Slightly different approach may be applied by defining four-vector of momenta, by formula 2.2,

$$P^J = \sum_i P^{tow} \tag{2.2}$$

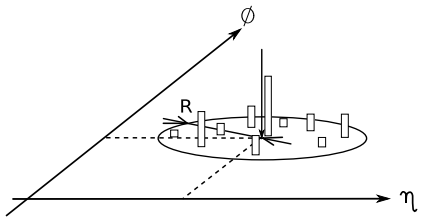
where $P^{tow} = (p_x, p_y, p_z, E)$.

These weighted coordinates define direction of the jet and are shown by arrow pointing on it in Figure 2.1b in η-ϕ. The same cone is in Figure 2.2 shown as cone containing tracks of detected particles. Vertex of this cone is assumed primary vertex, point of collision of incoming beam hadrons. As shown, some particles may originate from this primary vertex, some from secondary vertices inside this cone and some from decays of particles that did not originally belong inside the cone. Arrow signifying direction of jet is shown as originating from primary vertex and passing through centre of cone's base.

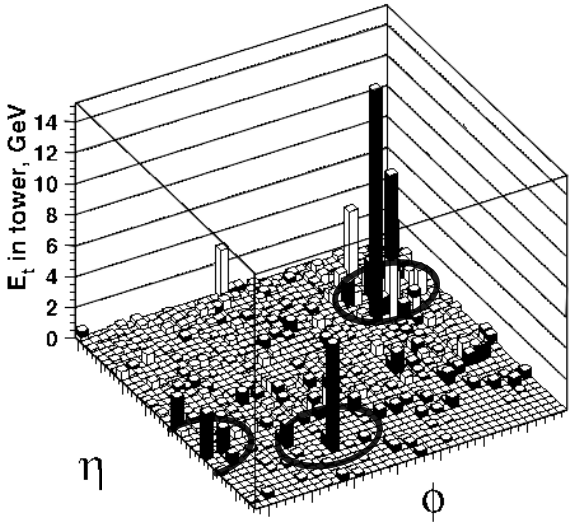
Seeded algorithms begin with selection of energy towers that may serve as iterative centres of cones. Depending on used algorithm, there may be threshold on energy of seed tower. This is to reduce noise caused by underlying events. Afterwards, all towers in radius around seed tower are linked together into protojets and weighted coordinates are calculated and used as new centre of cone. To reduce inclusion of background, cut on energy towers included into jet may be set. This is done until jet is stable, i.e. until small change in centre of jet results in small change in it's properties. Again, to prevent noise, threshold on energy of jet may be set.



(a) segmentation of detector



(b) cone in η-ϕ



(c) example of energy towers in detector, where jets are denouted by circles

Figure 2.1: η-ϕ representations for calorimeter and jets.

Depending on approach, after inclusion of energy tower, they may be removed from list of unused towers, thus preventing problem of overlapping cones. In the opposite case, this is dealt by summing up towers in region of overlap. When ratio of this shared energy and overall energy of energetically smaller jet is more than preset value, cones are merged as in Figure 2.3b and new iteration of centre begins. Otherwise cones are split and energy towers are assigned to jet whose centre is

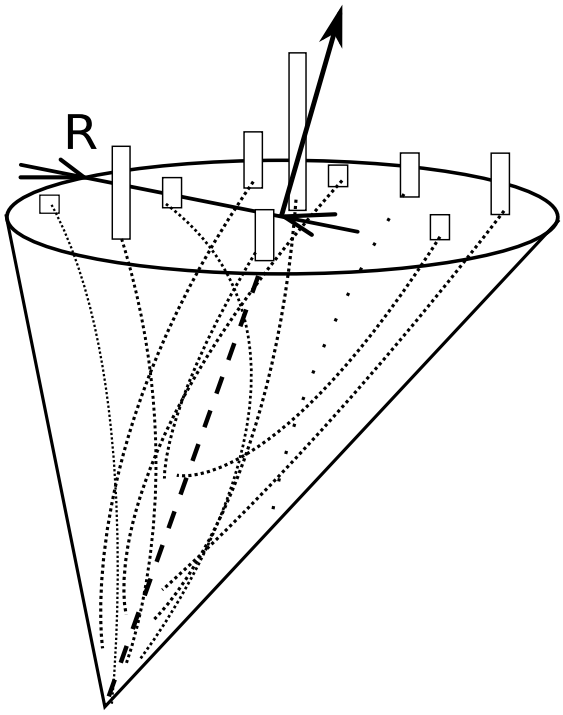
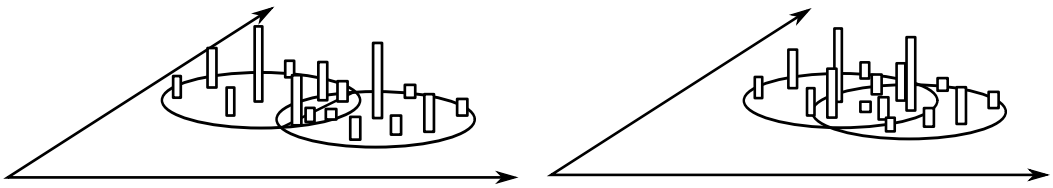


Figure 2.2: Energy towers in cone and possible track matching.

closer in η-ϕ system of coordinates, as shown in Figure 2.3a. All thresholds may be in either transverse momentum or transverse energy, depending on source of information being reconstructed particles' tracks or calorimetric towers.



(a) splitted jets (b) merged jets

Figure 2.3: Splitting and merging of cones in η-ϕ system of coordinates.

Also we may have option of *ratcheting* in iterative process of finding stable cone. When calculating cone of new centre, this makes previously merged energy

towers stay together, i.e. it may result in merging of energy towers outside of final cone to it.

Comparison between algorithm using four-vector of momentum and one utilising η-ϕ system of coordinates, *Mid-Point* and *JetClu* algorithms respectively, is discussed in more details in [1]. Both these algorithms were used at CDF in $p\overline{p}$ collisions.

Cone algorithm that does not use seed towers as trial centres of cone, thus name seedless algorithm, goes through all the calorimetric towers. For each, it makes trial cone and then finds weighted coordinates. If distance between trial tower and weighted centre is small enough, cone is iterated, otherwise next tower is taken as trial centre. After iteration, before accepting cone, it is checked whether this cone wasn't found in previous cone iteration. Surviving stable cones are protojets, and undergo process of splitting or merging as mentioned before with seeded algorithms.

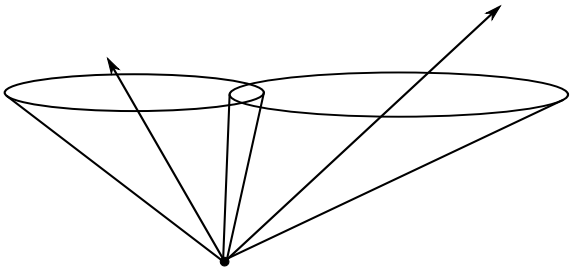
Between theoretical properties of ideal algorithm belongs also infrared safety to soft radiation, collinear safety, boundary stability, order independence, detector independence, etc. Infrared safety indicates, that when we have two reconstructed cones close together as shown in Figure 2.4a, soft radiation of parton does not change original configuration. In Figure 2.4b detected radiation leads to change in reconstruction of event into single cone from two original cones.

Problem of collinear safety is connected to the fact that energy of collinear particles or radiation does not have to be deposited into same energy tower and thus algorithm fails to produce seed tower, as seen in Figure 2.5a. On the other hand, detecting several particles in same tower may bias overall reconstruction and algorithm may result in production of fake jet, as in Figure 2.5b.

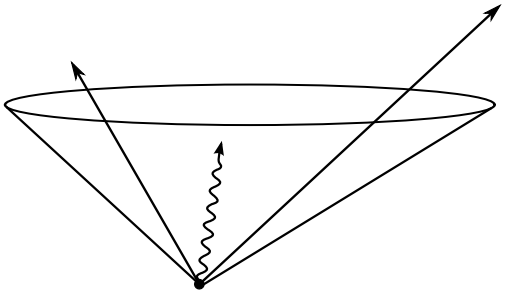
Cone algorithms are thus sensitive to threshold of transverse energy of seed towers. Possible removal, or improvement, of this may be done by ordering seed towers according to particles that are creating them. The seeds produced by hard particles are iterated first, those of softer origin are iterated later, if not already included in reconstructed cone.

For above mentioned Mid-Point algorithm, the first one based on creating midpoints between towers merged in cone, has reduced sensitivity to collinear and infrared radiation compared to seeded cone algorithm. On the other hand, JetClu is infrared and collinearly unsafe.

Boundary stability is property of cone, characterising stability of jets properties against inclusion of energy towers on it's border. Similarly to stability of jet, boundary stability may be viewed as stability to small change in cone's radius, i.e. whether with small change in radius, change in properties of jets will be small. Reconstructed cone has boundary instability, when for example there is large energy tower on it's border.

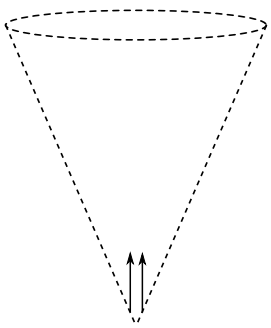


(a) initial configuration without soft radiation

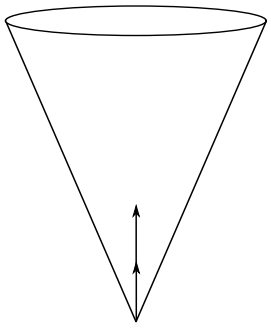


(b) infrared unsafe configuration

Figure 2.4: Infrared unsafety of clustering algorithm, soft gluon radiation between two jets causes them to be reconstructed as one.



(a) one particle in two towers, no jet reconstructed



(b) two particles inside one tower, jet is reconstructed

Figure 2.5: Problem of collinear safety where in case *a)*, energy of one particle is divided into two towers, both having energy below threshold for reconstruction of jet. In case *b)*, two particles deposit energy in same calorimetric tower,thus total energy stored is above threshold.

Order independence signifies same number of jets at parton, particle and detector level. Detector independence, is independence of performance on segmentation, resolution or energy response. For example, it is necessary to firstly calibrate detector's response on type of particle detected. It is different for leptons and hadrons in electromagnetic calorimeter. Also it is good to pay attention to boundary of detector, where part of the jet may be completely lost.

2.1.2 k_T algorithm

Another type of algorithm is k_T algorithm, clustering jetfinding algorithm used in proton - proton, or anti-proton, collisions. It uses four-vector of $(E, \vec{k})$, where $\vec{k} = \frac{E}{c}(sin\theta cos\phi, sin\theta sin\phi, cos\theta)$, θ is polar angle and ϕ is azimuthal angle. Direction of z-axis is parallel to the incoming beam direction. From kinematic point of view is more convenient to use instead of θ pseudorapidity, defined as $\eta = \frac{1}{2} ln \frac{E+c \cdot k_z}{E-c \cdot k_z}$. Transverse component of $\vec{k}$ is defined as $k_T = \sqrt{k_x^2 + k_y^2}$.

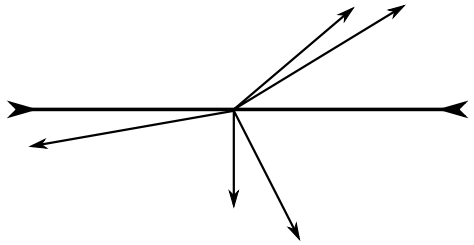
This algorithm combines particles in order of increasing transverse component of $\vec{k}$. Variable that is crucial for this merging is $d_{ij} = min\{k_{Ti}^2, k_{Tj}^2\} R_{ij}^2$, where $R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ is square of distance between two particles on surface of cylinder defined by η and ϕ as seen on Figure 2.1a.

General algorithm is as follows. For every pair of particles, d_{ij} is defined. Also for every parton i is defined distance from beam as $d_{iB} \equiv k_{Ti}^2$. Then, the smallest of d_{ij}, d_{iB} is taken. Unless it is second mentioned, pair of particles i and j is merged into l, for which new energy and $\vec{k}$ is calculated. $\mathrm{E}_l = \mathrm{E}_i + \mathrm{E}_j$ and $\vec{k_l} = \vec{k_i} + \vec{k_j}$. Otherwise, i is final state jet and is removed from list. This process is repeated, also with calculating d_{il} for newly created particle, until no particles are left in list. This process is schematically shown on simple example in Figure 2.6 Alternatively, a cut on value of d is set, so that only objects with certain value of d_{ij}, d_{iB} and thus speeding up the algorithm.

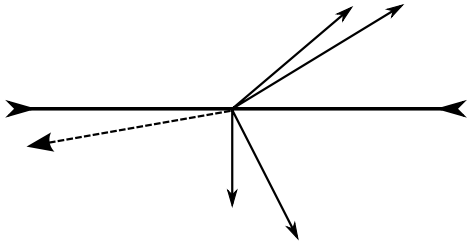
This general algorithm may be changed so that for example energy of newly created particle is weighted by its parents, or variable d could be rescaled by factor $\sim \frac{1}{D^2}$, where $D \leq 1$, and final jets are separated by $\Delta R \geq D$.

Every particle is assigned to unique jet, therefore there is no need for merging or splitting as was in cone algorithm. This algorithm is infrared safe and collinear safe, because every particle is treated individually. Disadvantage of this algorithm is of computing nature, it is of order $\sim N^3$, but may be modified, for example as FastJet, to order $\sim NlnN$, where N is number of particles. More can be found in [3].

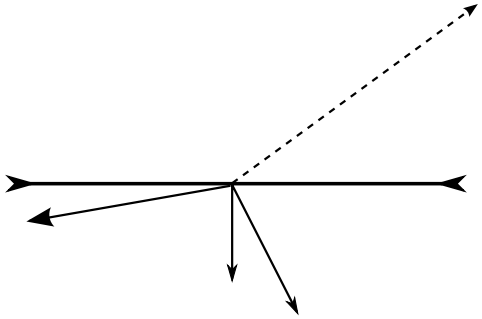
Similar algorithms to this one are DURHAM and JADE, working with energies instead of $\vec{k}$.



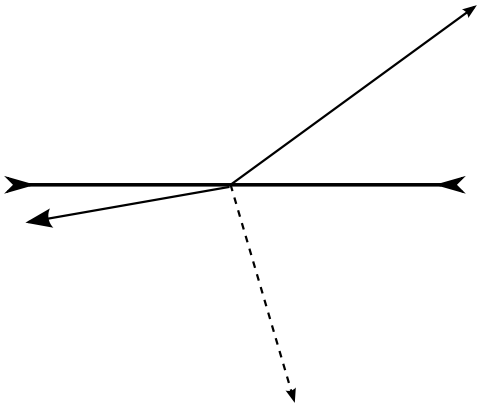
(a) initial configuration of particles detected in given event, before any iteration



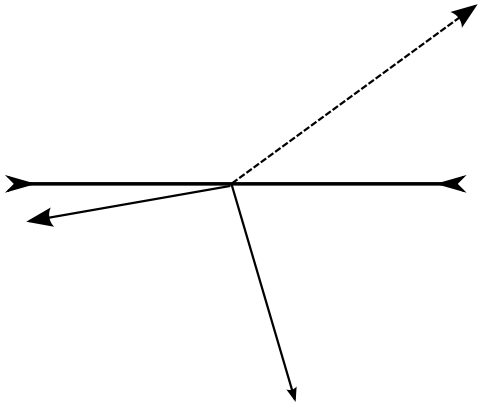
(b) first jet found, because of smallest d_{iB} between particle and incoming beam from set of all d_{ij}, d_{iB}



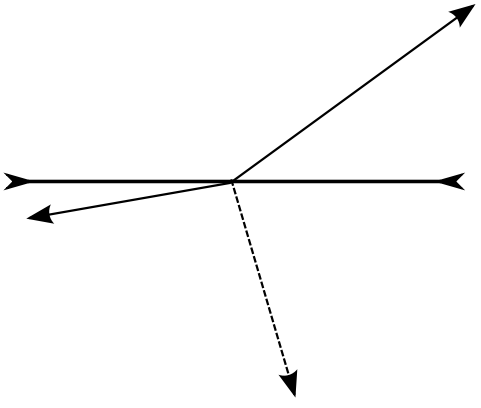
(c) upper two particles are merged due to having smallest d_{ij} from all d_{ij}, d_{iB}



(d) lower two particles are merged for same reason as in case *c)*



(e) smallest value is for d_{iB}, thus another jet is found



(f) final particle of the list has to be final jet

Figure 2.6: Schematics showing iterative process of finding jet using k_T algorithm, where small arrows signify particle's vector $\vec{k}$, large arrow already found jet and reversed arrows show incoming beam's paths. Dashed lines show $\vec{k}$ that undergone iteration most recently. Iteration proceeds alphabetically, from *a)* to *f)*.

2.1.3 Comparison of k_T and cone algorithms

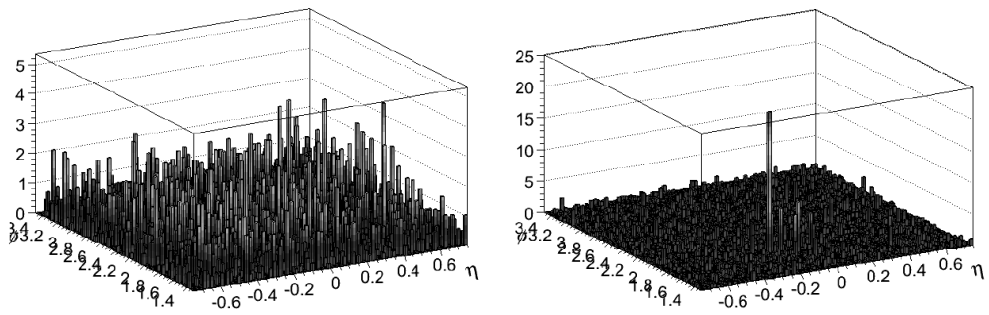
The greatest hindrance of k_T algorithm's application is its computing time. Unmodified version has to pass all the particles from the event, thus for central heavy-ion collisions producing large number of particles, is this algorithm inefficient from point of view of computing time, whereas cone-based algorithm presented, deals only with particles passing quality cuts. This holds true for seeded algorithm. Seedless one may face similar problem as k_T algorithm. In both cases applying a threshold for accepted energy towers reduces number of particles used in iteration. On the other hand, k_T algorithm directly creates new "particle" by adding up $\vec{k}$ of two particles fulfilling required conditions. This eliminates stability problems of jets that are faced, when cone algorithms are applied.

2.2 Heavy ion collisions

Due to multiplicity of particles in event containing heavy ions, reconstruction of jets becomes much more complicated. Not only number of particles processed in algorithm increases significantly, but also number of particles coming from underlying event grows drastically. This is depicted in Figure 2.7, where in case of Figure 2.7a we have energy towers of both signal, information we want to reconstruct looking as in Figure 2.7b, and background. Due to multiplicity of particles, k_T algorithm becomes really slow. Thus k_T is modified to create its faster versions, like FASTJET for ALICE group, so that it searches for the nearest particle according to certain algorithm. In process of production for JETAN module, jet finder module in development in AliRoot, is seedless infrared and collinear safe cone algorithm, SISCONE. Currently, JETAN includes UA1 based iterative cone algorithm optimised for heavy-ion collisions at LHC. It includes hadronic correction for double counts of charged hadrons in EMCal as well as in TPC, tools for background subtraction. Removal of energy from electrons that would be counted twice is in developement.

2.2.1 HIJA

Heavy Ion Jet Algorithm is cone algorithm, used in ALICE heavy ion experiment. Use of algorithms is meant to get rid of unnecessary background and find stable jet cones. In typical cone used, up to 2 TeV is expected to be from underlying event, therefore cones of smaller radii, R, are used. Input energies for algorithm are from charged tracks' transverse momenta and energy cells' transverse energy deposited in electromagnetic calorimeters. Also, cut on transverse momentum of charged tracks is set to p_T^{cut} = 2 GeV/c, which removes 98% of background particles on average in simulations.



(a) before algorithm (b) event without background

Figure 2.7: Particles form heavy ion collisions in η-ϕ. In case of *a)* both signal and background and in case of *b)* only information reconstructed from case *a)*, possible jet, is seen.

First step in algorithm is to estimate background energy in calorimeters, thus all energy cells are sorted according to decreasing transverse energy present, E_T^{cell}. Then, for at least two iterations and until change in E_T^{bg}, i.e. background energy per cell, is smaller than preset threshold value, following cycle is repeated.

1. List of jets is cleared at first and all cells are marked as being outside of jet.
2. From cell with highest E_T^i, if $E_T^i - E_T^{bg} > E_T^{seed}$, where E_T^{seed} is set threshold transverse energy for seed of a jet, and it is marked as outside of jet, then this cell is set as seed of a jet and reconstruction of cone begins.
    - Center of jet is set to coordinates of seed cell.
    - Using all cells in radius R_c from center of jet. New weighted centre is calculated from $(E_T^i - E_T^{bg})$. Repeat this process until change in position of center in one iteration is no more than one cell. This jet-candidate is stored and all cells inside cone are marked as inside of a jet.
3. Recalculation of estimated background energy using information from cells outside of jets follows.
4. For every jet, calculate energy by summing energies and of all cells in cone and subtract energy of a background. If final energy of jet is greater than energy threshold E_T^{jet}, then jet is found.

While measured jet energy increases only for small R, the background energy is proportional to R^2. However, the problems are event-by-event fluctuations, that

limit reconstruction and energy resolution, defined by 2.3. Energy resolution consists of own jet resolution and fluctuations in the background.

$$\frac{\sigma(E_T)}{E_T} \lesssim \frac{1}{E_T}\sqrt{\sigma(E_T^{jet})^2 + \sigma(E_T^{BG} - E_T^{\hat{B}}G)^2} \tag{2.3}$$

These fluctuations of the background originate mainly from fluctuations of impact parameter and dependence of background in cone for this type is quadratically proportional to cone radius and can be corrected by background subtraction dependant on impact parameter. Another fluctuation is poissonian fluctuation of uncorrelated particles, background is linearly proportional to cone radius. This are region-to-region fluctuations that are mostly eliminated by p_T cut. Third type are correlated particles from common source that increases region-to-region fluctuations.

As seen from Figure 2.8, where resolution is dependant on cone radius, the most effective radius is in region of 0.3 - 0.4, where resolution has the minimum. Also resolution decreases as jets of higher energies are more collimated, and they are easier to differentiate from background.

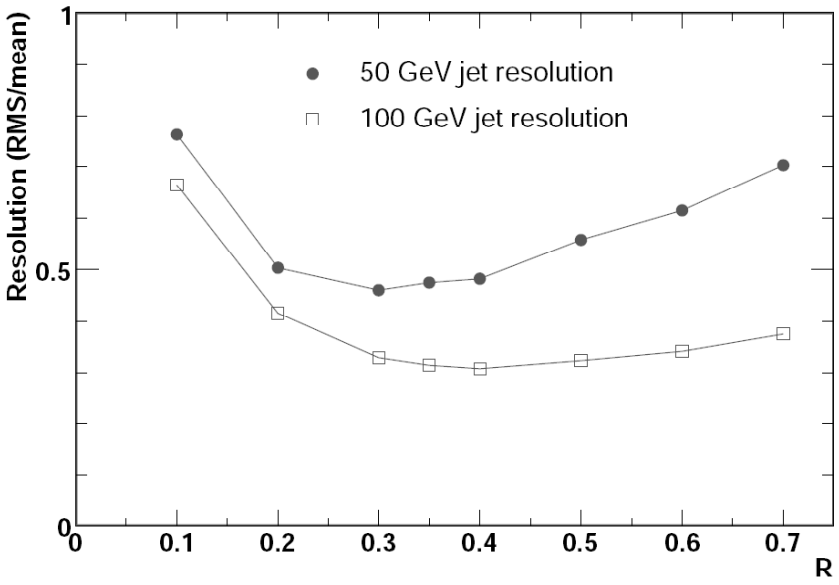


Figure 2.8: Dependence of resolution on size of cone radius, R, calculated for 50 GeV jets, circles, and for 100 GeV jets, squares. This figure is taken over from [7].

To increase reconstruction efficiency of jets, cones of smaller radius reconstructed close together could be merged with algorithm similar to k_T.

2.3 Jet Energy Correction

There are many factors that cause change in energy deposited in detectors compared to original energy of a jet, such as variation in detector response with pseudo-rapidity, non-linearity in detector response or origin of jet. This was studied for example at CDF [8].

First factor mentioned, variation in detector's response with pseudo-rapidity, is corrected by *relative jet energy correction*. This is done by process of *di-jet balancing*, when E_T of two jets in $2 \rightarrow 2$ process should be equal and independent of η. Correction is done by comparison of transverse energy inside pseudo-rapidity region of $0.2 < |\eta| < 0.6$ and that outside of this region. Ratio is taken as correction factor.

Absolute jet energy correction accounts for non-linearity in energy response and energy loss in regions of calorimeter that cannot detect, e.g. area between two segments of detector.

Correction for Jet Origin is applied to account for difference in energy of a hadron jet and energy reconstructed in calorimeter. Hadron jets are divided according to presence of B-hadron around jet axis to *b-jets*, to *c-jet* if b-hadron is not present, but hadron containing c-quark is. If it is neither b nor c-jet, but contains gluon, then it is gluon jet . Unless it does not fall into any previously mentioned category, it is categorised as jet originating from u, d or s-quark. Correction factor then depends on category to which given jet belongs. For example, [8] states that correction factor in proton-antiproton collisions, at CDF's center of mass energies, is for jets having energies between 25 to 55 GeV roughly 1 for uds-jets, 1.068 for gluon jets, 1.098 for c-jets and 1.16 for b-jets.

Chapter 3

Tagging

For present research, it is necessary to know, if and what kind of heavy quark was present in jet or what was its position. Process, in which we identify presence of particle is called *tagging*.

3.1 Lifetime tagging

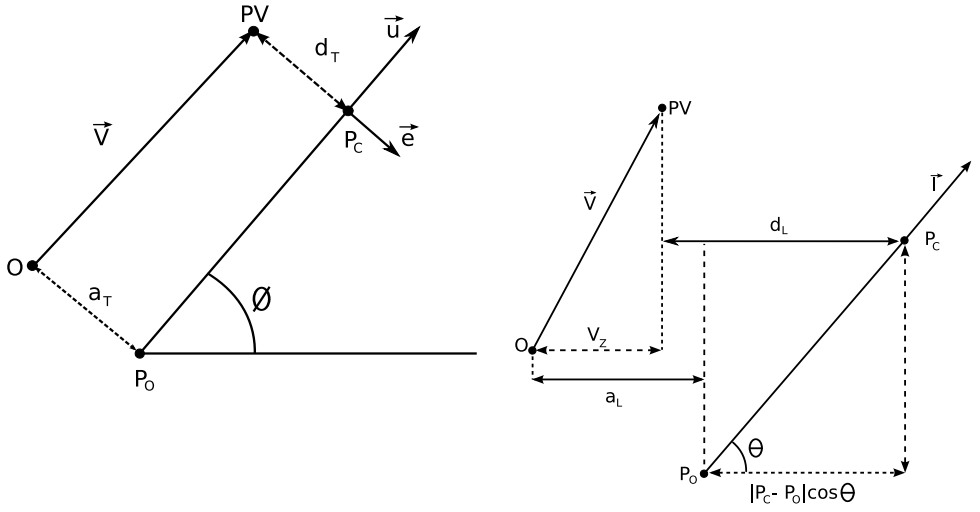
For B-hadrons, the most significant property is long lifetime, e.g. in case of $B^{\pm}$ it is $\tau = 1.638 \pm 0.011\text{ps}$*. From point of primary interaction, they may travel up to several hundreds of micrometers before they decay. This long lifetime is therefore used in tagging and it's most notable use is in secondary vertex tagging. B-hadrons, i.e. products in which are b-quarks fragmented may play important role in search for Higg's boson decaying through t-channel, while t-quarks may decay to b-quarks, or decay modes involving b-quarks in other way.

3.1.1 Impact Parameter

Impact parameter is the minimal distance of track from primary vertex of interaction. Long living particles decay into particles with tracks having large impact parameters. We may define separately two impact parameters, one in R-z plane, defined in Figure 3.1b, and one in R-ϕ plane according Figure 3.1a. One reason for their separate calculation is that they are calculated with different precision another is that discriminating power of longitudinal impact parameter is much smaller than power of the transverse one.

Impact parameter in these pictures is designated d_T for transverse impact parameter and d_L for longitudinal one. They are shortest distances between particle's track, with direction of unit vector $\overrightarrow{u}$ and passing through points designed P_O and

*Adopted from [6].



(a) IP in transverse, R-ϕ, plane (b) IP in longitudinal, R-z, plane

Figure 3.1: Impact Parameter's geometrical depiction in transverse, *a)*, and longitudinal, *b)* plane.

P_C, meaning point of closest approach to origin, *0*, and point of the closest approach to primary vertex, *PV*. It is good to note that in longitudinal case, only difference in projection to beam direction path, z-axis, is taken into consideration for impact parameter and also a_L. a_L and a_T are distances of the closest approach of particle's track to origin.

In the Figure 3.1 are also $\overrightarrow{V}$ and V_Z standing for position vector of primary vertex and it's z-component respectively. θ and ϕ are azimuthal and polar angles and $\overrightarrow{e}$ is unit vector perpendicular to direction of particle's path.

Impact parameters are calculated according to following formulae 3.1 and 3.3:

$$d_T = a_T - (\overrightarrow{e} \cdot \overrightarrow{V}) \tag{3.1}$$

$$d_L = a_L + \cot\phi(\overrightarrow{l} \cdot \overrightarrow{V}) - V_Z \tag{3.2}$$

$$= a_L - (\overrightarrow{l} \cdot \overrightarrow{V}) \tag{3.3}$$

where

$$\overrightarrow{u} = \cos\theta, \sin\theta, 0 \tag{3.4}$$

$$\overrightarrow{e} = \sin\theta, -\cos\theta, 0 \tag{3.5}$$

$$\overrightarrow{l} = -\cot\phi\cos\theta, -\cot\phi\sin\theta, 1 \tag{3.6}$$

$$|\overrightarrow{P_O P_C}| = (\overrightarrow{l} \cdot \overrightarrow{V})\frac{1}{\sin\phi} \tag{3.7}$$

As a measure of quality of found impact parameter, *significance* is defined as

$$S_T = \frac{d_T}{\sigma_T} \tag{3.8}$$

$$S_L = \frac{d_L}{\sigma_L} \tag{3.9}$$

Tracks decaying from longliving particles have large impact parameters exceeding σ.

Sign of a impact parameter, and also of significance, is assigned according fact whether track crosses jet's axis before primary vertex or after as shown in the Figure 3.2. Negative value removes possibility of secondary vertex being track's place of origin.

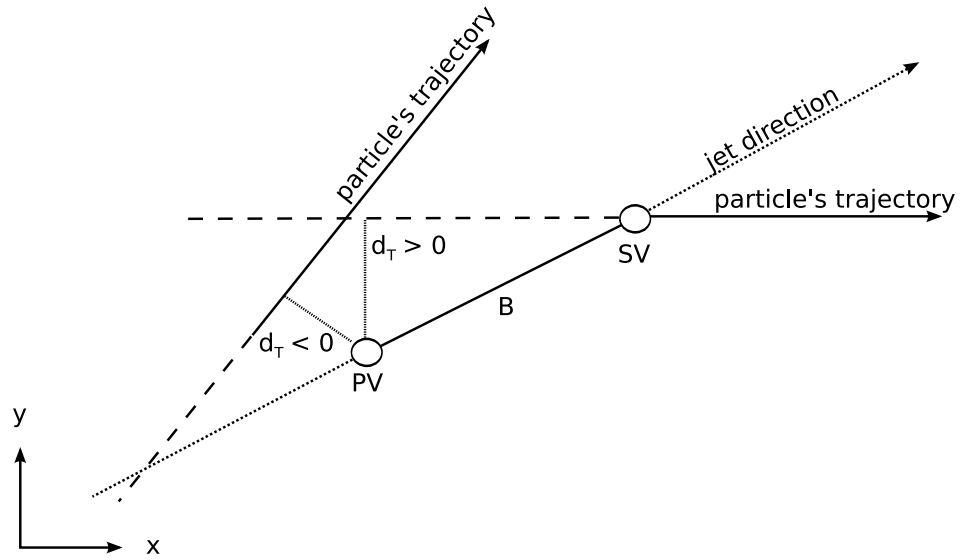


Figure 3.2: Depiction of particle's trajectory with negative and positive sign assigned to impact parameter.

Use of impact parameter, or significance, of track is based on estimation of probability of track originating, for example, from b-quark according to formula 3.10, in which $f_i(S_j)$ is probability of a track with significance of S_j to originate from i-quark.

$$W_i = \frac{f_b(S_i)}{f_{uds}(S_i)} \tag{3.10}$$

Discriminating variable, X, i.e. variable, on which cuts for selection are applied and which is used in combination with other discriminating variables for tagging, is derived according to formula 3.11, where $W = \Pi_i W_i$ and i cycles through all tracks.

$$X = \frac{W}{1 + W} \tag{3.11}$$

3.1.2 Primary vertex

Since primary vertex is used in calculation of secondary vertex, and may be used in calculation of B-hadron's flight path, it may be convenient to calculate its position. This is done by constraining tracks with small impact parameter to a common vertex. For a primary vertex, it is necessary to use at least 4-5 tracks of good enough quality. If there is not enough tracks, 3-4 are taken, but with more strict quality condition and cuts on impact parameter.

As a starting estimate, or estimate when there is not enough tracks to reconstruct primary vertex, is taken average position from previous events. When calculating position, tracks contributing mostly to χ^2, statistical function of best fit, in which values of parameters may float within their errors to find a best fit, are removed.

Firstly, from all tracks $\chi^2(N_{tr})$ is calculated. Then for every trajectory is calculated $\chi^2_i(N_{tr} - 1)$ without this trajectory. If difference $\chi^2(N_{tr}) - \chi^2_i(N_{tr} - 1)$ is greater than preset threshold, trajectory i is rejected from primary vertex iteration.

3.1.3 Secondary vertex

Secondary vertex is reconstructed vertex displaced from primary one. It may be significantly helpful in tagging long living particles, such as b-hadrons.

Secondary vertex algorithms search for displaced vertexes inside jets. Starting with jet having the highest transverse energy, it tries to combine tracks from *track list* to form displaced vertex.

Track list lists tracks available for vertex creation. They may be listed in several jets, but may be used only for one vertex. To be listed, they have to be reconstructed in 3D in tracking detector and pass quality cuts placed on them. Also, tracks are removed from track list, if they are considered to originate from decay of $K_S \rightarrow \pi\pi$ or $\Lambda \rightarrow \pi p$. This requires pairs of tracks of sufficient significance and invariant mass of pair, to be $K_S \pm 0.01\ GeV/c^2$ or $\Lambda \pm 0.006\ GeV/c^2$, fit to common vertex with $\chi^2(tracks)$ passing quality cut. Each track in *track list* is corrected for energy loss and multiple scattering.

Tracks in list are ordered according to number of hits, transverse momenta and significance. In first attempt to reconstruct secondary vertex, at least three tracks are required. Firstly, all tracks with significance great enough are taken as seed tracks

trying to reconstruct vertex. Any track with χ^2 contribution big enough is removed from vertex creation process, similarly as in primary vertex reconstruction.

Final reconstructed vertex has to be reconstructed from track with high enough transverse momentum and fulfill several conditions on quality, such as $\chi^2/d.o.f.$, invariant mass of reconstructed vertex smaller than preset cut, significance of displacement of secondary vertex from primary one to be greater than set cut, as well as displacement to be smaller than given value.

If in first attempt secondary vertex is not found, second attempt is made, in which only two tracks are necessary for reconstruction, but cuts on their impact parameter's significances, transverse momenta, quality of reconstruction are tightened, while cuts on reconstructed vertex are nearly same. [†]

3.2 Lepton identification

Another useful property of B-hadron decay is their high branching ratio into leptons. B-hadron decay into lepton is direct, $b \to l$, when it decays in way that

$$b \to W^{-*} X$$

$$W^{-*} \to l^- \bar{\nu}_l$$

Branching ratio of this decay is $Br(b \to l^-) = 10.70 \pm 0.22\%$.

Second possible decay channel is cascade decay, $b \to c \to l$, when

$$b \to W^{-*} c$$

$$c \to l^+ \nu_l X$$

which has branching ratio of $Br(b \to c \to l^+) = 8.02 \pm 0.19\%$.

Third type of decay is called "wrong sign" cascade decay, in which

$$b \to W^{-*} X$$

$$W^{-*} \to q\bar{c}$$

$$\bar{c} \to l^- \bar{\nu}_l Y$$

and it's branching ratio is $Br(b \to \bar{c} \to l^-) = 1.62^{+0.44}_{-0.36}\%$.

To sum it up, the branching ratio of producing at least one lepton is $19.3 \pm 0.5\%$ and chance of getting two leptons from either direct or cascade decay channel is approximately 1%.[‡]

[†] Details can be found in [10]

[‡] All values of branching ratios involving leptons are taken from [9]

Electron identification is typically based on matching electromagnetic shower in calorimeter with associated track in tracking system. Electrons with high transverse momenta, sufficient impact parameters and good quality, assured by sufficient number of hits in detector layers, are taken into consideration. Also candidate track has to pass cuts on energy and if possible ratio between energy deposited in hadronic and electromagnetic calorimeters. Tracks are extrapolated to calorimeters and distance between extrapolation and shower centroids has to be small enough.

Showers between transverse and longitudinal profiles are compared with showers from test beam electrons. For electron to be accepted, χ^2 comparison has to be smaller in both cases than given values. Also measure of difference between observed sharing of energy deposited between towers in detector and expectation from real shower, measured by *lateral shower sharing*, has to be small enough. High value indicates that shower spread is wide.

For muons, good quality is ensured by hits in muon detectors, and usual cuts on momenta are higher than those for electrons. Muons pass through calorimeter without significant energy loss. Also on both leptons is set requirement of angle between given lepton and secondary vertex track, i.e. track of particle from primary to secondary vertex, to be greater than 2° and smaller than right angle. Another requirement is set for invariant mass of lepton and track consistent with B-decay, opposite charge track forming vertex with lepton in question.

Discriminating variables in lepton tagging are discussed at the end of section 3.5 on page 34.

3.3 Jet probability algorithm

This type of algorithm assigns probability to all tracks coming from secondary vertex or tracks included in jet's cone. This probability depends on impact parameter of a track, greater parameter is the more probable that track originates in secondary vertex. For jet tag, only impact parameters with positive value are taken and for given track appropriate parametrisation is used to to find probability that track from primary vertex has larger significance of impact parameter. These probabilities for each track are used to calculate probability for whole jet.

Similar tagging is discussed in section 3.5 on page 33.

3.4 Flavour tagging techniques

When it is necessary to determine content of neutral meson, either if it is created by heavy quark, B^0 or its antiparticle, $\overline{B^0}$. There are several methods for this. First one is *same-side tagging* method utilizing information from creation of given meson.

Second one is *opposite-side tagging* method and utilizes information from "other side", i.e. if heavy quark was created in pair with its anti-particle, we use information about this second particle's baryon. Another method is determination of quark content from shape of jet.

3.4.1 Jet shapes

Shape of jet is given by parton emission from primary parton described by fragmentation models. It depends on transverse momentum of particles in jet, showing how transverse momentum is distributed inside the cone with increasing distance from centre of the cone. A non-negligible role in internal structure of jet, and therefore it's shape as well, has underlying event. As underlying event changes with energy of incoming colliding beams, it is necessary to take into consideration that shapes may vary according to the change in these energies.

Shape of jet provides information about flavour of initial parton, but it may also help with fragmentation models and understanding of hadronisation, that are necessary for proper comparison of simulation results and observations from detector.

There are two types of jet shapes, integral and differential. First one is function of cone radius defined as portion of total transverse momentum of concentric cone inside with radius *r*. It is defined by equation 3.12 and depicted in the Figure 3.3, where subcone of radius r < R and same centre is drawn representing portion of cones area that is taken into consideration, when calculating transverse momentum of subcone of given radius.

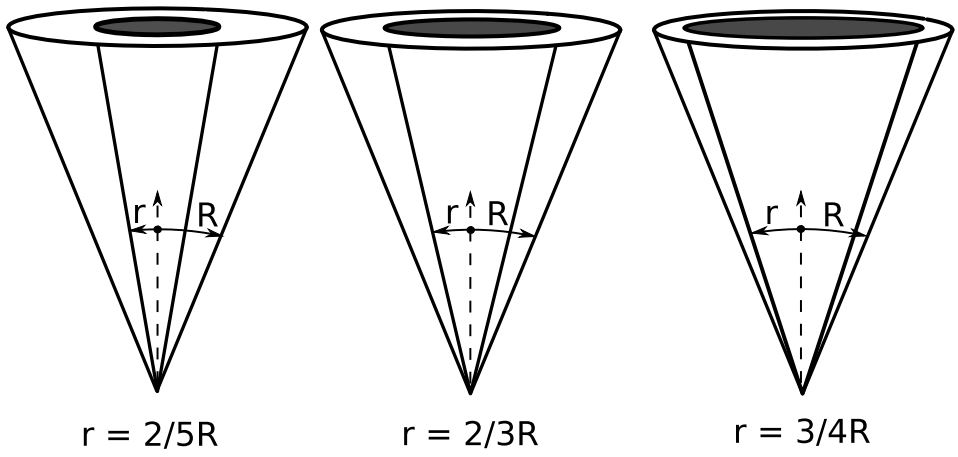


Figure 3.3: Representation of integrated jet shape as sum of transverse momenta inside subcone, dark region of the base of cone, with given radius.

$$\Psi(r) = \frac{1}{N_{jets}} \sum_{jets} \frac{p_T(0,r)}{p_T(0,R)} \tag{3.12}$$

Here, N_{jets} is number of jet taken into account and function $p_T(0,r)$ is sum of transverse momenta of all particles within cone with same centre as cone, in whose shape we are interested in, but with radius equal to r. Therefore $p_T(0,R)$ is sum of momenta of all particles in jet whose shape we are looking for. Jet shape is in fact average taken from several jets of same properties in question, for example jets being b-jets, that's why we sum over several jets. $p_T(0,R)$ and N_{jets} normalize Ψ to 1, i.e. $\Psi(r = R) = 1$ and for representation of integral shape of just one jet, both constant before sum, as well as sum are omitted from definition, because sum goes through just one jet and therefore $N_{jets} = 1$.

Second type is defined as increase in transverse momentum in small change of radius. It is defined by equation 3.13 and is depicted in the Figure 3.4. Δr is small increment in radius, it defines area between two cones with difference of radii equal to Δr, in which sum of momenta is calculated and used or differential jet shape.

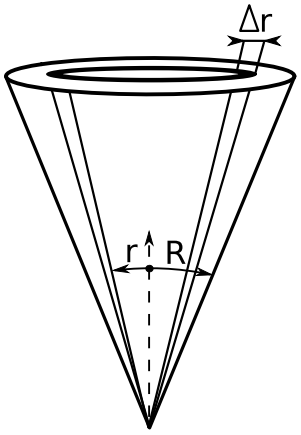


Figure 3.4: Representation of step taken in calculation of differential jet shape, where transverse momenta of area of increment Δr innermost cone with radius r are taken to be after normalisation $\rho(r)$.

$$\rho(r) = \frac{d\psi}{dr} = \frac{1}{N_{jets}} \lim_{\Delta r \to 0} \sum_{jets} \frac{p_T(0, r+\Delta r) - p_T(0,r)}{p_T(0,R)\Delta r} \tag{3.13}$$

According to nature of this type of shape, small increments in total transverse momentum in small area defined by increment in radius, it must give $\Psi(r) = \int_0^r \rho(y)dy$

and for $r = R$ it has to be equal to 1, according to normalisation of $\Psi(R)$. In formula 3.13 is $\lim_{\Delta r \to 0}$ only theoretical depiction of principle, because we are unable to measure difference in radii more accurately than the size of energy cells in calorimeters of detector. Δr thus won't be smaller than this size.

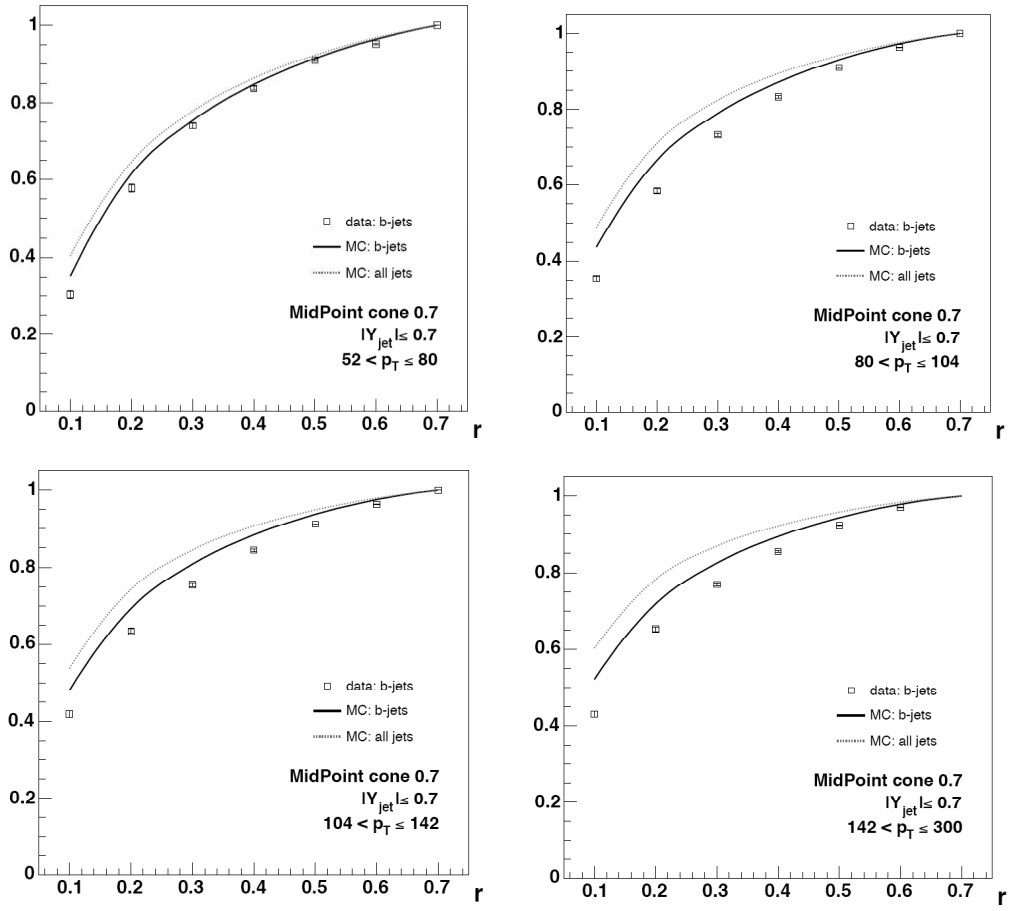


Figure 3.5: Jet shapes for different cuts on particles' transverse momentum taken over from [1]. MC simulations were done by Pythia tune A and data taken from CDF run II.

Jet shapes are sensitive type of initial hard-scattered parton. According to results of [1], measurement of b-quark jet shapes is good method of checking monte-carlo modelling of ratio between flavour creation and gluon splitting in b-quark production. Also it is good test for fragmentation models, when we may measure distribution of particles, or tracks, with transverse momentum from certain interval and compare it to data measured in experiment.

Example of b-jet shapes are shown in Figure 3.5. Every subfigure represents shape of particles with given transverse momentum range. It is clearly seen that with increasing momentum particles tend to be closer to jet axis. Also simulation predicts having generally smaller relative transverse momentum, compared to transverse momenta of a whole jet, in concentric cones of smaller radius than in those of greater. Another shapes, from events generated using Pythia, are in section 4.2 on page 43.

3.4.2 Same-side tagging

As mentioned before, *same-side tagging method* utilizes additional information directly related to meson in question. One of the ways is determination of decay products of neutral meson. For example, neutral B-meson's, B^0, semi-leptonic decay may be used, positively charged lepton indicates indicates B-hadron, negative charge $\overline{B}$-hadron.

To use these information, it is necessary to first correctly identify lepton from decay. Several quality cuts may be placed on tracks and secondary vertex reconstructed. Also distance between secondary vertex projection along beam direction, denoted as z-axis, and triggered lepton shall be small enough. To differentiate trigger lepton from photon conversion electron cuts are placed on their separation in transverse plane and on angle between lepton and it's conversion partner. If possible 3D vertex fit ensures that track and its conversion partner originate from same vertex. Invariant mass between trigger lepton and secondary vertex track coming from B-decay shall be between 2 and 4 GeV/c^2.

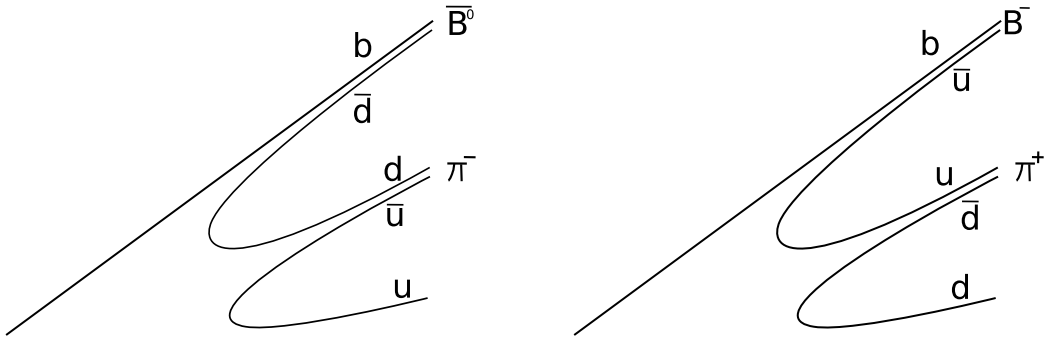
Other way, when direct information of hadron is not available, information from fragmentation may be used. For example, b-quark may fragment into $\overline{B^0}$, it is joint by $\overline{d}$ from $d\overline{d}$ pair. If remaining d-quark is bound with $\overline{u}$, they create π^- as depicted in the Figure 3.6a. Alternatively, as in Figure 3.6b, if correlated π^+ is produced with B^--meson, it is highly probable that b-quark fragmented to B^- by joining $\overline{u}$ that was created as $u\overline{u}$ pair. Remaining u-quark is bound with d-quark from $d\overline{d}$ pair.

Alternatively, with π^+ and π^- may be correlated B^0 and B^+ respectively.

3.4.3 Opposite-side tagging

Opposite-side tagging is method that uses information from B-hadron produced from antiparticle to b-quark in question, both from $b\overline{b}$ pair produced in flavour creation. This can be done by determining charge of lepton from semi-leptonic B-decay, thus determining if it was B-meson or anti-meson, and giving information about quark content. This tells us whether it was particle or antiparticle.

Information about "opposite side" b-quark may come from charge of kaon from decay of subsequent charm meson.



(a) fragmentation to $\overline{B^0}$

(b) fragmentation to B^0

Figure 3.6: Same-side tagging method utilizing information from fragmentation of b-quark

3.5 Discriminating variables

For tagging, combination of different variables may be used. One way of using them is applying set of cuts on these variables. Other one utilizes likelihood ratio method, which is very simple method separating signal from background.

A combined discriminating variable is defined as follows:

$$y = \frac{f^{bgd}(x_1, ..., x_n)}{f^{sig}(x_1, ..., x_n)} \tag{3.14}$$

where $f^{bgd}(x_1, ..., x_n)$ and $f^{sig}(x_1, ..., x_n)$ are probability density functions of background and signal of discriminating variables. We select events, whose $y < y_0$. However, utilization of multidimensional probability density functions is usually complicated, therefore distribution functions for each discriminating variable is used separately.

$$y = \prod_{i=1}^{n} \frac{f^{bgd}(x_i)}{f^{sig}(x_i)} = \prod_{i=1}^{n} y_i \tag{3.15}$$

This holds true for independent variables with small correlations.

All discriminating variables are calculated for every jet separately in an event. Every jet is categorised into one of three categories according to number of reconstructed secondary vertices, if any such vertex is reconstructed, jet belongs to a first category. Reconstructed vertex provides large number of discriminating variables and thus ensures clean selection. If no secondary vertex is reconstructed, tracks passing a cut on significance probability of a track to be from B-hadron are taken

to originate in it, and for a jet to belong to second category there has to at least two tracks fulfilling such a condition. Jets in third category do not have reconstructed secondary vertex nor sufficient number of such tracks, therefore having limited set of discriminating variables available for testing, like lifetime probability etc. For jets in second and of coarse in third group, it is much more probable for a background jet, for example c-jet, to pass cut and be mistagged.

Classification of jets in given type of a event may be simulated and known distributions of jets in those three categories can help in tagging, acting as another discriminating variable y_{categ}, defined by formula 3.16, where n^j_{categ} are normalised rates of production of jet in gicen category, such that $\sum n^j_{categ} = R_j$ and $R_b + R_c + R_q = 1$.[§]

$$y_{categ} = \frac{n^c_{categ}}{n^b_{categ}} \prod_i \frac{f^c_{i,categ}(x_i)}{f^b_{i,categ}(x_i)} + \frac{n^q_{categ}}{n^b_{categ}} \prod_i \frac{f^q_{i,categ}(x_i)}{f^b_{i,categ}(x_i)} \tag{3.16}$$

It may be useful to define variable $X_{jet} = -\log_{10} y_{categ}$, then similar variable for whole event is defined as sum of X_{jet}'s in the event, in some cases it is the sum of only two largest values in the event.

In b-tagging discriminating variables used are:

- Jet lifetime probability, calculated from tracks with large positive impact parameter included in jet. It is probability for tracks coming from primary vertex to have product of track probabilities exceed observed value. Tracks from secondary vertex shift lifetime probability to lower values, producing peak near 0,
- Mass of particles combined in secondary vertex,
- Transverse momentum of secondary vertex,
- Rapidity of each track from secondary vertex,
- Transverse momentum of identified lepton.

In tagging soft leptons from secondary vertex, as discriminating variable may be used

1. Variables using information from inner detector
    - Number of high energy hits along track,
    - Impact parameter,

[§]Here q stands for quarks u, d and s.

2. Variables using combined information from inner detector and electromagnetic calotimeter

    - Ratio between transverse energy deposited in electromagnetic calorimeter transverse momenta reconstructed in inner detector,
    - Core energy computed from computed in different strip-layers of detector,
    - Fraction of energy deposited in first longitudinal compartment of electromagnetic calorimeter centered around strip with maximal energy,
    - Fraction of energy deposited in third compartment of electromagnetic calorimeter, where largest deposit from π is expected,
    - Difference between track and shower position,
    - Shower isolation in electromagnetic calorimeter, measured from ration between energy deposited in (3x3) and (3x7) clusters of cells around energy cell with highest energy deposit.

Chapter 4

Pythia data

In this chapter the properties of jets are discussed. For generation of events, monte-carlo program Pythia was used.

4.1 Pythia

Pythia is programme for generation of high energy physics events from collisions of incoming particles, i.e. it generates sets of particles produced in interactions. Aim of this program is to offer as precise representation of properties of events as possible in large scale of interactions in and beyond Standard Model. Because physics may not be understood well enough, Pythia is based on combinations of several analytical QCD models. It also contains own research from areas like fragmentation and hadronisation, deep inelastic scattering and photon physics, supersymmetry, technicolor and extra dimensions and other.

For generation of an event, physical aspects of evolution are ordered as follows:

- Two initial particles are coming together, each characterised by set of PDF's for their partonic constituents,
- Initial shower is built up,
- One parton from each of two particles'showers is taken for hard process, where number of outgoing partons is made,
- Hard process may produce resonances that must be taken into consideration with hard process,
- Outgoing partons branch up, building up final state shower,
- Other semi-hard processes may occur between other partons of incoming hadrons,

- Due to QCD confinement, quarks and gluons are hadronised into colour-neutral hadrons,
- Many of produced hadrons are unstable and undergo a further decay.

For this thesis, I used Pythia 8.105. It yet works only for collisions of protons, proton and anti-proton, electron-positron and muon-antimuon collisions. This version uses default CTEQ 5L parton distribution functions, for hadronisation only Lund string fragmentation is utilised and particle data are taken from 2006 Particle Data Group tables.

4.2 Generated data

4.2.1 General informations

Pythia setting

Due to options of Pythia, I focused on proton-proton collisions at 14 TeV, that will be energy at center of mass for LHC, and then at $\sqrt{s} = 200$ GeV, what is currently energy at RHIC.

First of all, I generated 100 000 events of all hard QCD processes that include heavy-quark creation from both, gluons and from quarks, in proton-proton collisions (set by command `"HardQCD:all=on"`). Cone radius, for cone based inbuilt PYCELL algorithm used, was set to 0.4 and 0.7 for LHC energies and for 200 GeV energies was radius 0.7 and 1.0, as is typically used (e.g. [7]). Seed tower threshold for jet cone reconstruction in Pythia was set to 3 GeV for default setting of cells (50x32 cell in η-ϕ system of coordinates in range of $(-5; +5) \otimes (0; 2\pi)$). Minimal energy required to be inside a cone was 20 GeV. Invariant transverse momentum's minimum value for particles from hard-scattering was set to 20 GeV/c. For production of events used in comparison of general properties in Figure 4.1, 4.2 and 4.3.

Data

Comparing properties of jet from two collisional energies, we see in Figure 4.1b and 4.1a, that with increased collisional energy, sum of masses of particles inside a cone also increases. This is observed from shift of maximum in case of cone's radius equal to 0.7 (from 5 to 9 GeV). Logical consequence of increasing radius is increase in number of particles inside reconstructed jet. This is clearly seen from these figures (For 200 GeV collision, it is shift from 5 to 7 GeV from cone radius of 0.7 to 1.0. In case of 14 TeV collision, the shift is from 6 to 9 GeV with cone's radii 0.4 and 0.7). Range of masses inside cones for 200 GeV varies from 0 to 17 GeV

for $R = 0.7$ and from 0 to 36 GeV for $R = 1.0$. In case of 14 TeV events, mass ranges from 0 to 17 GeV for cone with $R = 0.4$ and between 0 and 47 GeV.

For transverse energy, at $\sqrt{s} = 200$ GeV shown in Figure 4.1c, jets are reconstructed having up to 40 GeV, whereas at 14 TeV, in Figure 4.1d, maximum is around 100 GeV. These values hold true for cases of both cone radii used and they have nearly same shape of distribution.

Jets are produced symmetrically around $\eta = 0$, where maximum is located. This is seen in Figures 4.1f and 4.1e, where weighted centres of jet cones are plotted into a histogram. Also it is obvious that with higher energies, the jets are produced further from central region, η is from $(-2.2, 2.2)$ for 200 GeV events and $(-5, 5)$ for 14 TeV.

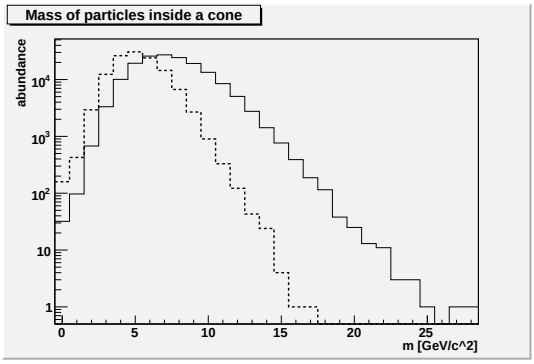
In collision at $\sqrt{s} = 200$ GeV with cone $R = 0.7$ there are at maximum 2 jets in one event or for $R = 1.0$ are 3 jets per event, at $\sqrt{s} = 14$ TeV 13 jets are produced for cone radius equal to 0.4. For radius 0.7, 18 jets are produced. Compared to 2 jets for 200 GeV event with same cone, it is significant increase. This is seen in Figures 4.2b and 4.2a, where number of jets per event is plotted into histogram's x-axis.

Similarly, in Figures 4.2d and 4.2c, we can see increase in number of particles per every jet of event. For 200 GeV, 40 particles, nearly 50 for cone radius 1.0, are at maximum inside a jet. The most abundant number of particles is 10 for $R = 0.7$ and 15 for 1.0. In case of 14 TeV, maximum number of particles per jet is 60, nearly 100 and most abundant number of particles is 15, 25 for cone of 0.4 and 0.7 respectively.

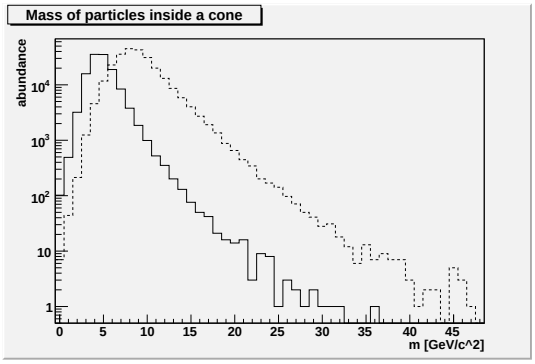
When selecting events containing solely two jets, we may observe in Figure 4.2f and Figure 4.2e distances of these two jets from each other. Distance of two jets is defined as $\Delta R = \sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2}$, where η_i is weighted η-coordinate of i-th jet, in this case only first or second and ϕ_i is weighted ϕ-coordinate of this jet. Maximum is located around value 3.14, which is significant for back-to-back jets indicating that

there was produced pair, particle-antiparticle, around which jets were formed. For collisions at lower center-of-mass energy, this peak is much more significant compared to collision at higher energy, where greater variations from 3.14 are much more abundant.

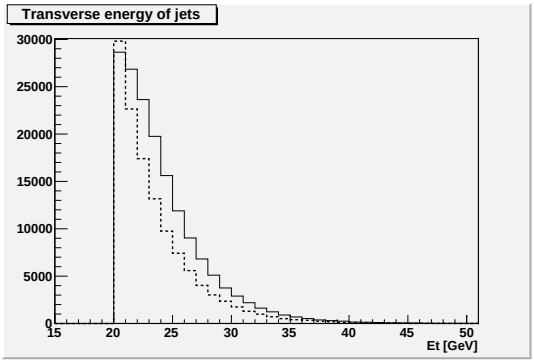
From point of view of particle content, Figures 4.3b and 4.3a, i.e. histograms of particles present inside every jet's cone, we see that jets are dominated by photons and pions in cases of both $\sqrt{s}$ energies. These are all final-state particles, that should make it to detector without decaying first.



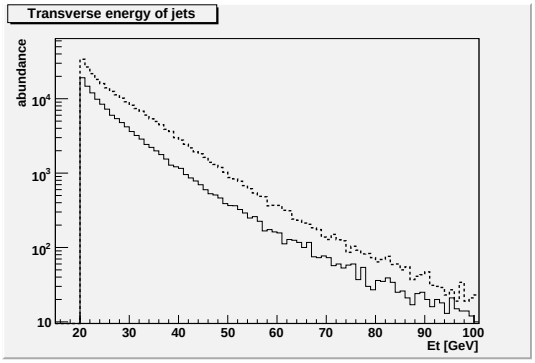
(a) recombined mass of jet's particles - 200 GeV



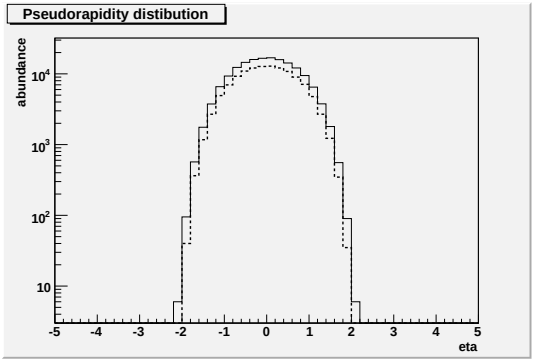
(b) recombined mass of jet's particles - 14 TeV



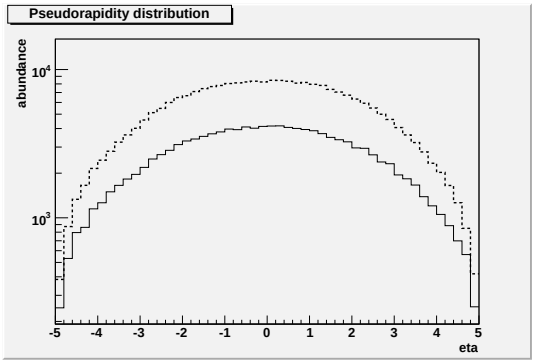
(c) transverse energy - 200 GeV



(d) transverse energy - 14 TeV

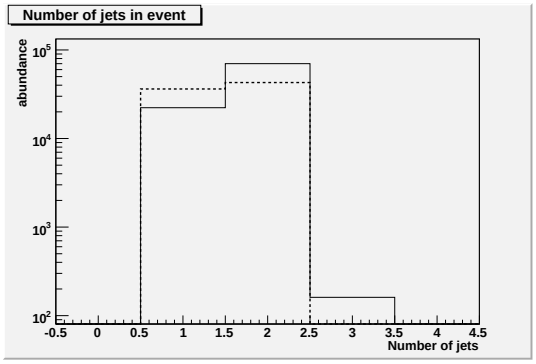


(e) η-coordinate of cone's center - 200 GeV

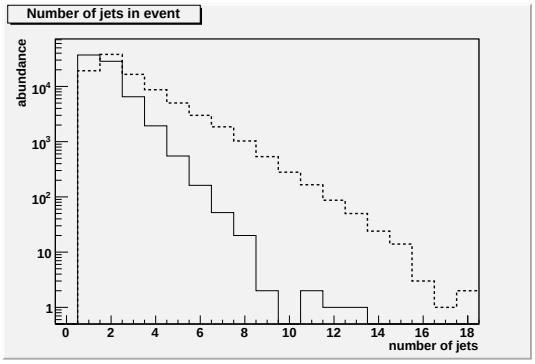


(f) η-coordinate of cone's center - 14 TeV

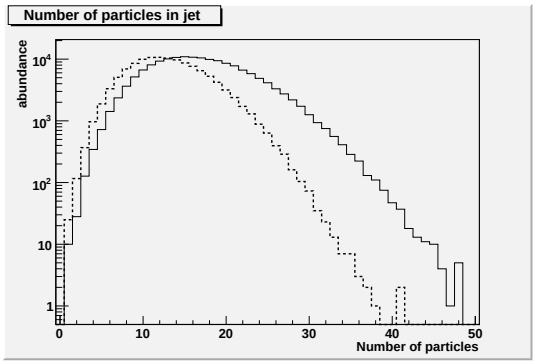
Figure 4.1: General information about events produced in proton-proton collisions, in left column with $\sqrt{s}$ = 200 GeV and in right column with $\sqrt{s}$ = 14 TeV. Dashed line stands for cone radius R = 0.7, while full line for R = 0.4 in 14 TeV events and R = 1.0 in 200 GeV events.



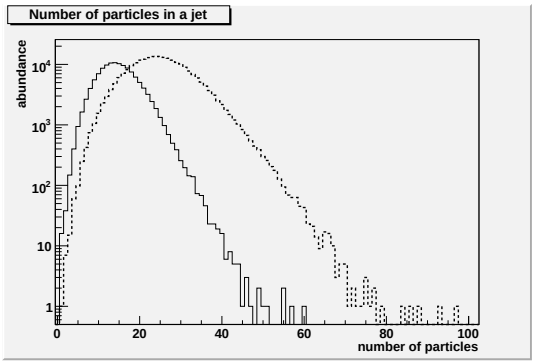
(a) Number of jets in one event - 200 GeV



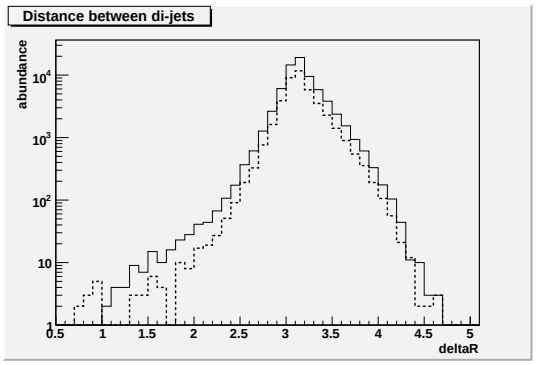
(b) Number of jets in one event - 14 TeV



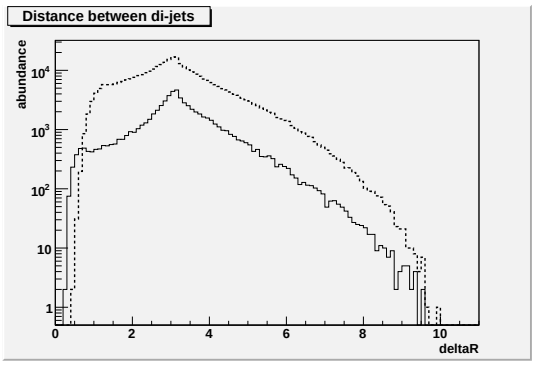
(c) Number of particles in one jet -200 GeV



(d) Number of particles in one jet -14 TeV

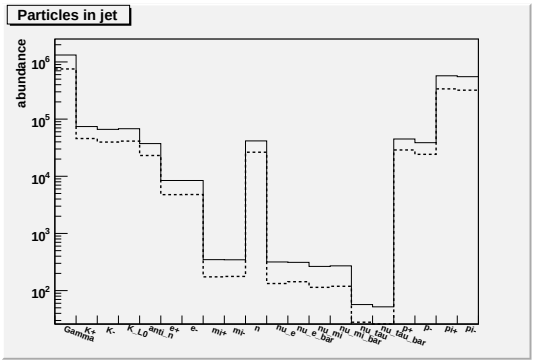


(e) Distances between di-jets - 200 GeV

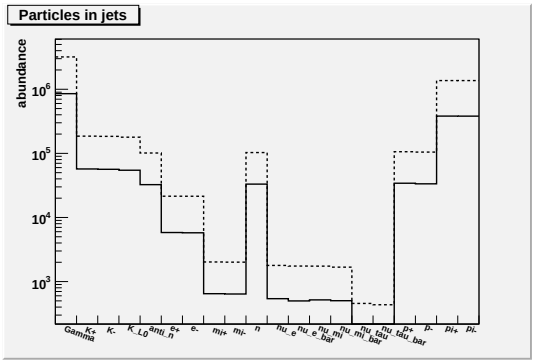


(f) Distances between di-jets - 14 TeV

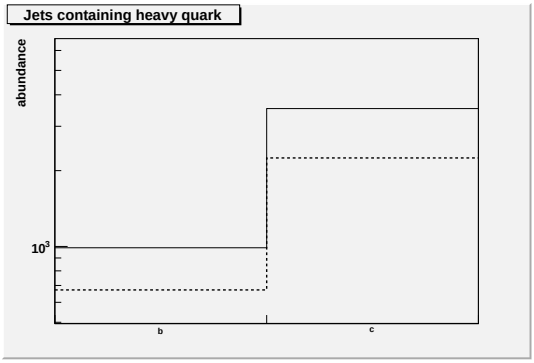
Figure 4.2: General information about events produced in proton-proton collisions, in left column with $\sqrt{s}$ = 200 GeV and in right column with $\sqrt{s}$ = 14 TeV. Dashed line stands for cone radius R = 0.7, while full line for R = 0.4 in 14 TeV events and R = 1.0 in 200 GeV events.



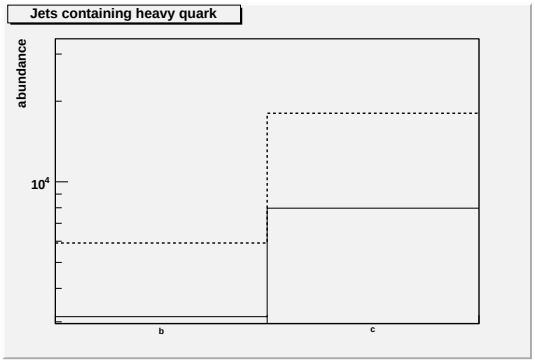
(a) Particles present inside jets - 200 GeV



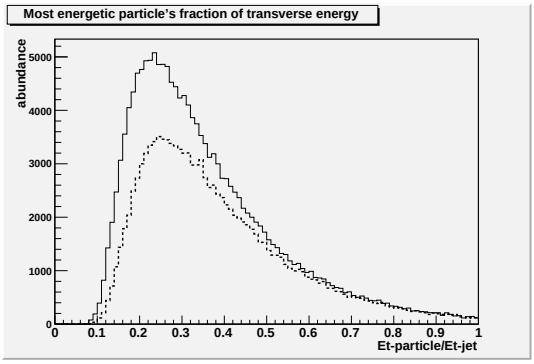
(b) Particles present inside jets - 14 TeV



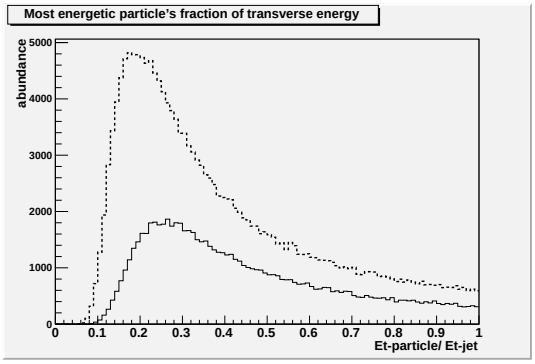
(c) Heavy-quark jet's origin - 200 GeV



(d) Heavy-quark jet's origin - 14 TeV



(e) Fraction of jet's transverse energy in particle with highest transverse energy - 200 GeV



(f) Fraction of jet's transverse energy in particle with highest transverse energy - 14 TeV

Figure 4.3: General information about events produced in proton-proton collisions, in left column with $\sqrt{s}$ = 200 GeV and in right column with $\sqrt{s}$ = 14 TeV. Dashed line stands for cone radius R = 0.7, while full line for R = 0.4 in 14 TeV events and R = 1.0 in 200 GeV events.

The difference in number of jet containing c-quark to that containing b-quark, when generating events using all hard processes, Figure 4.3d and Figure 4.3c, is most probably due to different production cross-sections of c- and b-pair in initial hard scattering. Also abundance of heavy quarks is much greater for more energetic collisions signifying greater cross-section of heavy quark creation process. Difference between number of heavy quarks reconstructed using different cone's radii indicates, that efficiency of reconstructing jets containing heavy quark depends on this variable. Search for heavy quark's presence inside reconstructed cone was done by utilisation of concentric subcone of smaller radius. If heavy quark or B-hadron was found inside this subcone, cone is considered as containing heavy quark. If heavy quark or B-hadron would be found outside this subcone, but still inside original cone, this would give information only that cone was reconstructed near heavy quark, but it does not have to be responsible for its creation.

In Figure 4.3f and Figure 4.3e are histograms showing abundance of particles having highest transverse energy in jet, sorted according to the fraction of transverse energy of reconstructed cone. Most notable is fact that for particles with this fraction close to zero, there is non-negligible part of jets present. This corresponds to fact that there are several jets with relatively small number of particles, thus more probable is that they will have greater fraction of transverse energy. The most abundant fraction for all cases is located between 0.2 and 0.3. Also, the lowest possible fraction measured was around 0.08, the lower values were observed for greater cone radii, where more particles and higher energies were observed and for greater energies of collisions.

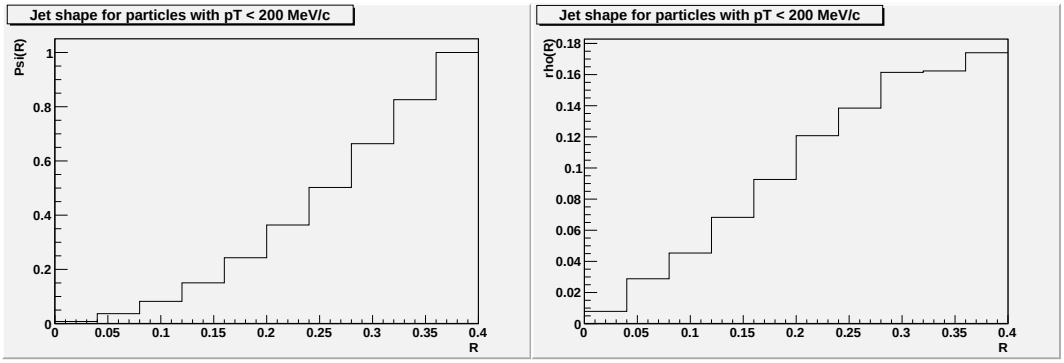
4.2.2 Jet shapes

Pythia setting

In a previous chapter jet shapes were mentioned. In Figure 4.4 and Figure 4.5 jet shapes were generated for b-jets identified by finding b-quark inside a subcone of R = 0.25 having same coordinates of centre. Events were generated with same conditions as in previous case of general information, except that minimal invariant transverse momenta was lowered to 10 GeV/c.

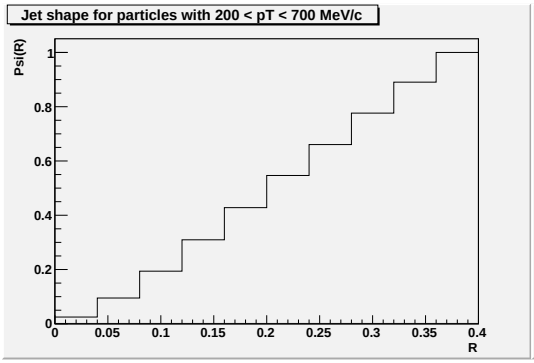
Data

In left column are integrated shapes, as explained showing what portion of total transverse momentum is located inside a subcone of given radius, and in right column are differential shapes. These show increase in transverse momentum in given layer of cone, ΔR.

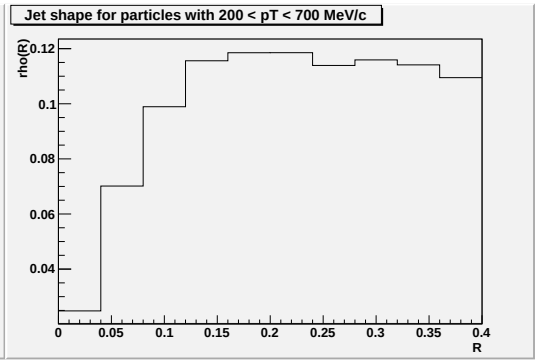


(a) integrated jet shape

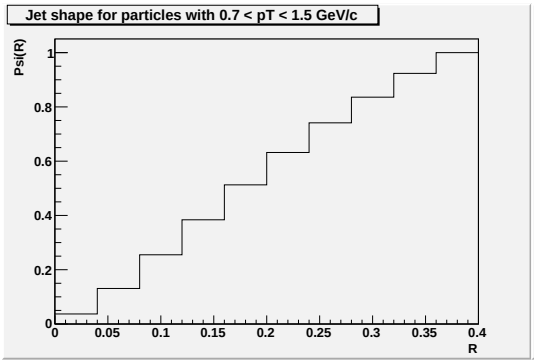
(b) differential jet shape



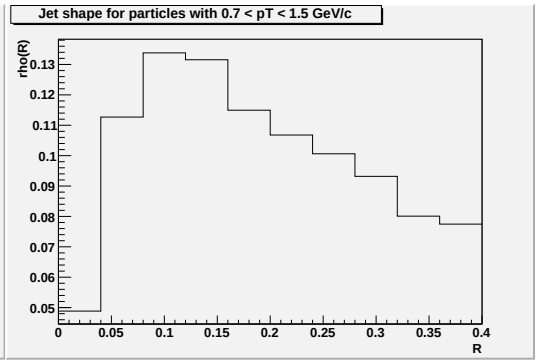
(c) integrated jet shape



(d) differential jet shape

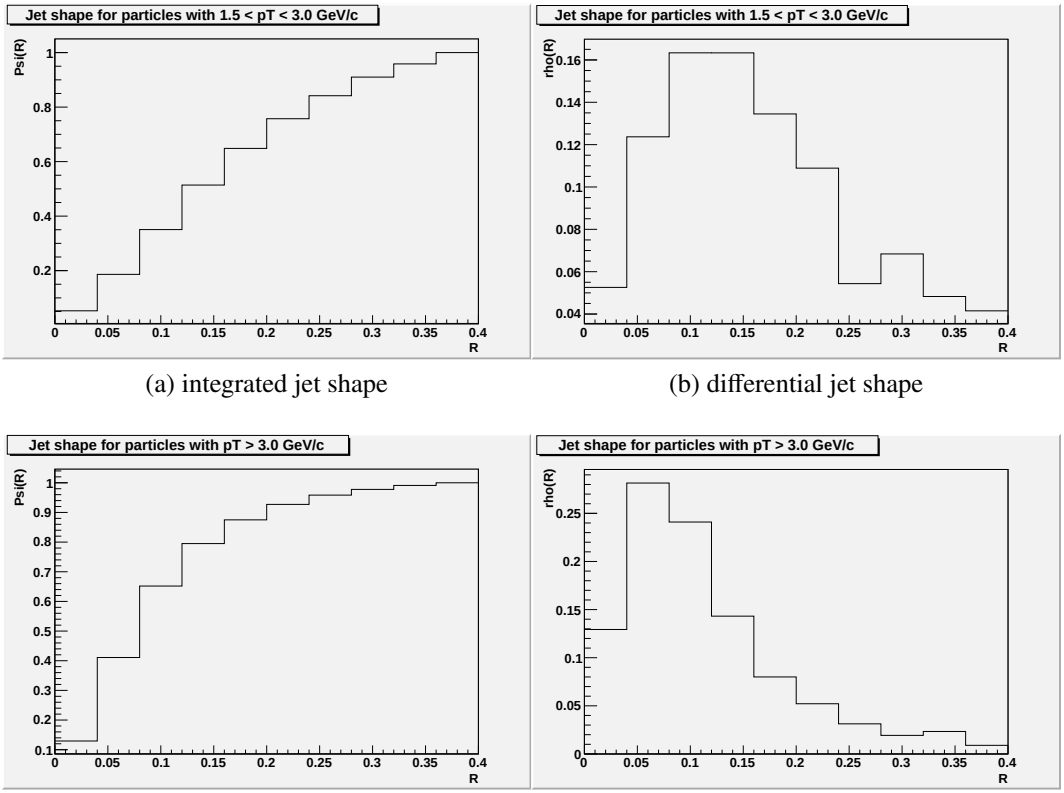


(e) integrated jet shape



(f) differential jet shape

Figure 4.4: Shapes for b-jets from 14 TeV events. Jets were found by default cone algorithm PYCELL in Pythia8 with radius R = 0.4. In left column are integrated shapes and in right corresponding differential jet shapes.



(a) integrated jet shape (b) differential jet shape

(c) integrated jet shape (d) differential jet shape

Figure 4.5: Shapes for b-jets from 14 TeV events. Jets were found by default cone algorithm PYCELL in Pythia8 with radius R = 0.4. In left column are integrated shapes and in right corresponding differential jet shapes.

Figures are divided according to transverse energy of a particle. First Figure 4.4a is made out of particles with transverse momentum smaller than 200 MeV/c. It shows nearly linear increase in transverse momentum and is expected to be dominated by background particles. As cut in transverse momentum increases, we observe that particles with greatest momentum are located mainly between subcones of radii 0.05 and 0.15. This is most notable in differential jet shapes in Figure 4.5 showing greatest increase in transverse momenta in this region. In heavy ion collisions jet shapes shall be modified due to energy loss inside hot medium.

Summary

Jets as tool in high energy particle physics for studying hot nuclear matter, in general cannot be fully reconstructed, because of our limited knowledge of origin of final state particles, whose tracks we observe or which leave energy deposited inside calorimeters. There will always be particles excluded from jets although belonging to cascade from original parton and, on the other hand, particles will be included inside jet although belonging to underlying event. Jet reconstruction is complex problem attempted to solve by different types of algorithms, whose two examples were presented and discussed in this work. For events with small multiplicities, k_T algorithm presents itself as more suitable choice, however with increasing number of particles modifications are necessary. In ALICE, besides FASTJET modification of k_T, cone algorithm is utilised with enhanced background subtraction and Seedless Infrared Safe CONE algorithm is being developed.

In second part, several possible ways of heavy quark tagging were discussed. This included tagging based on leptonic decays, information from impact parameter of secondary particles or secondary vertices and their reconstructed properties like mass and displacement from primary vertex. Ultimately the most effective way of tagging is by combination of different discriminating variables, whose diversity offers set of independent observables directly comparable with expectations for particles we are searching for. In LHC and ALICE, heavy quark studies allow us to investigate their production, propagation and hadronisation in hot dense nuclear matter. ALICE's Inner Tracking System will provide heavy quark identification by reconstruction of secondary vertices and measuring decay times. Also significant role will play identification of leptons from heavy quark decays.

This excellent particle identification shall serve for testing medium effects induced at parton splitting. Studies of nuclear modification will give insight into transport coefficient inside medium.

Bibliography

[1] A. Lister, *"Measurement of b-quark jet shapes at CDF"*.

[2] A. Dainese, *"Charm production and in-medium QCD energy loss in nucleus–nucleus collisions with ALICE",* [arXiv:nucl-ex/0311004].

[3] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, *"Longitudinally Invariant K_T Clustering Algorithms For Hadron Hadron Collisions",* Nucl. Phys. B **406** (1993) 187.

[4] M. Estienne, *"Jet reconstruction performance in p+p and Pb+Pb collisions in ALICE from full detector simulation",* ALICE Physics Week - Prague (2008).

[5] B. Alessandro *et al.* [ALICE Collaboration], *"ALICE: Physics performance report, volume II",* J. Phys. G **32** (2006) 1295.

[6] Particle Data Group, *"Review of particle physics 2006"*.

[7] S. L. Blyth *et al.*, *"A cone jet-finding algorithm for heavy-ion collisions at LHC energies",* J. Phys. G **34** (2007) 271 [arXiv:nucl-ex/0609023].

[8] A. Gajjar, *"Measurement of the inclusive b anti-b jet cross section at the Collider Detector at Fermilab",*

[9] P. Demin, *"Tagging b-jets with e^- and μ^- at CDF"*

[10] D. Acosta *et al.*, Phys. Rev. D **71** (2005) 052003.

[11] K. P. Lannon, PhD. Thesis, 2003.

[12] STAR Collaboration, *"Experimental and theoretical challenges in the search for the quark-gluon plasma: The STAR Collaboration's critical assessment of the evidence from RHIC collisions"*, Nuclear Physics AVolume 757, Issues 1-2, , First Three Years of Operation of RHIC, Pages 102-183.

[13] PHOBOS Collaboration, *"The PHOBOS perspective on discoveries at RHIC"*, Nuclear Physics AVolume 757, Issues 1-2, , First Three Years of Operation of RHIC, Pages 28-101.

[14] PHENIX Collaboration, *"Formation of dense partonic matter in relativistic nucleus-nucleus collisions at RHIC: Experimental evaluation by the PHENIX Collaboration"*, Nuclear Physics AVolume 757, Issues 1-2, , First Three Years of Operation of RHIC, Pages 184-283.

[15] BRAHMS Collaboration, *"Quark-gluon plasma and color glass condensate at RHIC: The perspective from the BRAHMS experiment"*, Nuclear Physics AVolume 757, Issues 1-2, , First Three Years of Operation of RHIC, Pages 1-27.

[16] T. Sjöstrand, S. Mrenna and P. Skands, *JHEP05 (2006) 026*.

[17] T. Sjöstrand, S. Mrenna and P. Skands, *arXiv:0710.3820*.

[18] DELPHI Collaboration, *"b-tagging in DELPHI at LEP"*, CERN–EP/2002-088.

[19] G. A. Giurgiu, *"B Flavor Tagging Calibration and Search for B(s) Oscillations in Semileptonic Decays with the CDF Detector at Fermilab"*.

[20] Yu.L. Dokshitzer and D.E. Kharzeev, Phys. Lett. B519 (2001) 199.

[21] D. Jeans [CDF Collaboration], *"B tagging at CDF"*.

[22] B.I. Abelev et al., Phys.Rev.Lett.98:192301,(2007), nucl-ex/0607012.

[23] D.Frederic (ATLAS), *"Soft electron identification and b-tagging with DC1 data"*, ATL-PHYS-2004-026 .

[24] G. P. Salam, M. Cacciari, *"Jet clustering in particle physics, via a dynamic nearest neighbour graph implemented with CGAL"*.