

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE  
Fakulta Jaderná a Fyzikálně Inženýrská

# DIPLOMOVÁ PRÁCE

2008

Bc. Radek Šmakal

**CZECH TECHNICAL UNIVERSITY IN PRAGUE**

**Faculty of Nuclear Sciences and Physical Engineering**

Department of Physics

**Study of ultra-relativistic nuclear collisions by means of  
multi-resolution methods**

**(Vyžití metod s proměnným rozlišením při studiu  
ultra-relativistických jaderných srážek)**

Master's Thesis



**Author:** Bc. Radek Šmakal

**Supervisor:** doc. RNDr. Vojtěch Petráček CSc.

**Academic year:** 2007/2008

## **Study of ultra-relativistic nuclear collisions by means of multi-resolution methods (summary)**

This master's degree project offers an introduction into multiscale analysis and its possible application in jet studies at heavy-ion collision experiments. It also summarizes jet-finding methods and compares their advantages and disadvantages. Particular implementation of the multiscale algorithm is described and tested on two jet samples, obtained from Monte-Carlo generator, with random generated 'soft' background.

*Keywords: multiresolution, multiscale, jet*

## **Využití metod s proměnným rozlišením při studiu ultra-relativistických jaderných srážek (abstrakt)**

Tato diplomová práce nabízí úvod do problematiky multiškálové analýzy a zabývá se jejím možným využitím při studiu jetů na experimentech s kolizemi těžkých iontů. Také shrnuje metody pro hledání jetů a srovnává jejich výhody a nevýhody. V práci je popsána konkrétní implementace multiškálového algoritmu a testována na dvou vzorcích jetů, získaných Monte-Carlo generátorem, s náhodnými částicemi simulujícími 'měkké' pozadí ve srážkách těžkých iontů.

*Klíčová slova: multirezoluční, multiškálová, jet*

## **Prohlášení**

Prohlašuji, že jsem svou diplomovou práci vypracoval samostatně a použil jsem pouze podklady (literaturu, projekty, SW atd.) uvedené v příloženém seznamu.

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

V Praze dne 9.5.2008

Radek Šmakal

## **Acknowledgements:**

First of all, I am very grateful to doc. RNDr. Vojtěch Petráček CSc. for his patience, professional assistance and invaluable remarks. I would like to thank Bc. Tomáš Valenta for his font Atomino and Jiří Král for the language and programmer remarks.

# Table of contents

<b>1</b>	<b>MRA</b>	<b>7</b>
<b>2</b>	<b>Jets</b>	<b>9</b>
2.1	Jets in hadron-hadron collision . . . . .	9
2.2	Jets in nucleus-nucleus collision . . . . .	10
2.2.1	Suppression of high- $p_T$ hadrons . . . . .	11
2.2.2	Two-particle azimuthal correlations . . . . .	13
2.3	Jets in heavy-ion collisions at the LHC . . . . .	13
<b>3</b>	<b>Jet-finding algorithms</b>	<b>15</b>
3.1	Cone algorithms . . . . .	17
3.2	$k_T$ algorithms . . . . .	18
<b>4</b>	<b>MRA jet-finder</b>	<b>21</b>
4.1	Principles of MRA jet-finder . . . . .	21
4.2	Cauchy-Lorentz distribution . . . . .	25
4.3	Boundary conditions . . . . .	25
4.4	Threshold . . . . .	25
4.5	Domain detection . . . . .	27
4.6	Test jets . . . . .	27
4.7	Background . . . . .	28
4.8	Testing . . . . .	28
4.8.1	Background subtraction . . . . .	30
4.8.2	Reconstruction precision . . . . .	30
4.8.3	Cleanness and Efficiency . . . . .	31
4.9	Test results . . . . .	36

# Introduction

In high-energy heavy-ion collisions, hard or semi hard parton scatterings in the initial stage may result in a large amount of jet production. Knowledge of the jet characteristics may greatly improve our understanding of the processes which taking part during the collision. Although there is a significant number of jet-finding algorithms (for example  $k_T$  and cone algorithms), they are of limited use in the high-multiplicity environment such as in heavy-ion collisions. Main aim of this thesis is to explore a different approach to the jet-finding, based on general principles of the multi-resolution analysis [1]. This approach might be better for case of heavy-ion collisions than other methods. We can observe jets as fluctuations, regions with higher number of particles and higher transverse momentum. We can look at these fluctuations at different levels of resolution and find the resolution, which will highlight them (which gives the strongest signal in corresponding region).

Method used in this thesis should be also applicable in other fields of high-energy physics, for example in search for signatures of the disoriented chiral condensate (event-by-event fluctuations of the charged/neutral ratio of produced pions) [13] and in search of droplets of quark-gluon plasma [14].

The introductory part of this master's degree work discusses multi-resolution analysis and gives mathematical background needed for understanding of this approach.

The second part studies jets, both jets in hadron-hadron collision and in nucleus-nucleus collisions. This part includes definition of the jet and reasons for study of jets. It also overviews methods for jet studies used at the RHIC and possible methods which can be used at the LHC.

The third part deals with existing jet-finding algorithms (both  $k_T$  and cone jet-finders), describes several basic requirements for jet-finders and compares their speed and complexity. This part includes definition of the jet algorithm and the recombination scheme.

The closing fourth part explains principles of a new, MRA-based jet-finding method and describes particular implementation of the algorithm which has been used for testing. Test jets sample, course of testing and test results are also described in this part.

The path to the new jet-finding algorithm is covered by thorns. It is path full of obstacles (both theoretical and programming), dead ends, and long-time development but it is an opportunity to create "something new".

# Chapter 1

## MRA

Multi-Resolution Analysis (MRA, also known as MultiScale Approximation - MSA) is a group of methods used in signal processing, image data compression and functional analysis [6, 3]. MRA is often connected with concept of wavelet analysis [7, 8] (or ondelette transformation, original name comes from French), but we will not use wavelets in this thesis (the wavelet analysis is using orthogonal functions - wavelets, our jet-finding concept is based on Cauchy-Lorentz functions, which are not orthogonal). MRA has been developed from theory of microlocal analysis (used in theory of differential equations) and from pyramidal algorithms [5, 2].

Lets have space  $L^2(R)$  - space of square integrable functions. We will define sequence of resolution, indexed by integers, so that all details of signal (or investigated function) at scales lower then  $2^{-j}$  are suppressed at resolution  $j$  [9]. Subspace of functions, which contains information about signal up to scale  $2^{-j}$  is identified as  $V_j$ . MRA contains analysis of function to the system of subspaces  $V_j$ .

First requirement is involvement of  $V_j$  to all higher subspaces, thus

$$\dots \subset V_0 \subset V_1 \subset \dots \subset V_n \subset V_{n+1} \subset \dots \subset L_2(R) \quad (1.1)$$

It is called 'nested subspaces'.

Lets identify approximation of function  $f(t)$  at level  $j$  as  $f_j(t)$ . Then evidently  $f_j(t) \in V_j$ . Difference between  $f_{j+1}(t)$  and  $f_j(t)$  is information about details at level  $2^{-(j+1)}$ , lets identify  $d_j(t)$ .

Then

$$f_{j+1}(t) = f_j(t) + d_j(t) \quad (1.2)$$

We can analyse similarly our subspace and get:

$$V_{j+1} = V_j \oplus W_j \quad (1.3)$$

,where  $W_j$  is called detail space at resolution level  $j$  and is orthogonal to  $V_j$ . By repeating this analysis of space  $V$  we get:

$$V_{j+1} = W_j \oplus V_j = W_j \oplus W_{j-1} \oplus V_{j-1} = \dots = W_j \oplus W_{j-1} \oplus W_{j-2} \oplus \dots \oplus W_{j-J} \oplus V_{j-J} \quad (1.4)$$

It is worth of note that any two detail spaces with different resolution are orthogonal and also that detail space  $W_j$  is orthogonal to the space  $V_k$  only if  $k < j$ .

Second requirement of MRA is including all square integrable functions into the finest resolution and also that there is only zero function at the coarsest level.

As we are approaching to coarser and coarser resolution more details is cutted off and in the limit  $j \rightarrow -\infty$  will left only constant function (must be null due to square integrability). In the second extreme there is added more and more details to the infinity resolution, so we cover whole space of square integrable functions.

Third requirement is a dilation invariance. It should be expressed as

$$f(t) \in V_j \iff f(2t) \in V_{j+1} \quad (1.5)$$

Fourth requirement is translation invariance (if  $f(t)$  is from space  $V_0$ , then  $f(t - k)$  is also from space  $V_0$ ,  $k \in Z$ )

Fifth requirement is existence of a function  $\phi$  with a property that its translations are orthonormal basis for  $V_0$ . This function is called the scale function.

Now we will summarise all requirements to the formal definition of MRA

MRA of space  $L^2(R)$  is sequence of nested subspaces  $\{V_j\}_{j \in Z}$  with properties:

- 1)  $\dots \subset V_{-1} \subset V_0 \subset V_1 \subset \dots \subset L^2(R)$
- 2)  $\bigcap V_j = 0, \overline{\bigcup V_j} = L^2(R)$
- 3)  $f(t) \in V_j \iff f(2t) \in V_{j+1}$
- 4)  $f(t) \in V_0 \implies f(t - k) \in V_0$
- 5)  $\exists$  function  $\phi(t)$ , called scale function;  $\phi(t - k)$  is orthonormal basis of  $V_0$

# Chapter 2

## Jets

### 2.1 Jets in hadron-hadron collision

When two high-energy hadrons collide, one of four kinds of scattering processes occur: elastic, diffractive, soft inelastic or hard inelastic.

In our case, there are important only inelastic processes - one or both participating hadrons decay. Soft inelastic scattering induce only small momentum transfer. It is described by virtual hadron exchange (Regge theory [15]) and comprise the biggest part of the total cross section.

Hard inelastic scattering (which can produce a jet) is plotted in Fig. 2.1 [10]

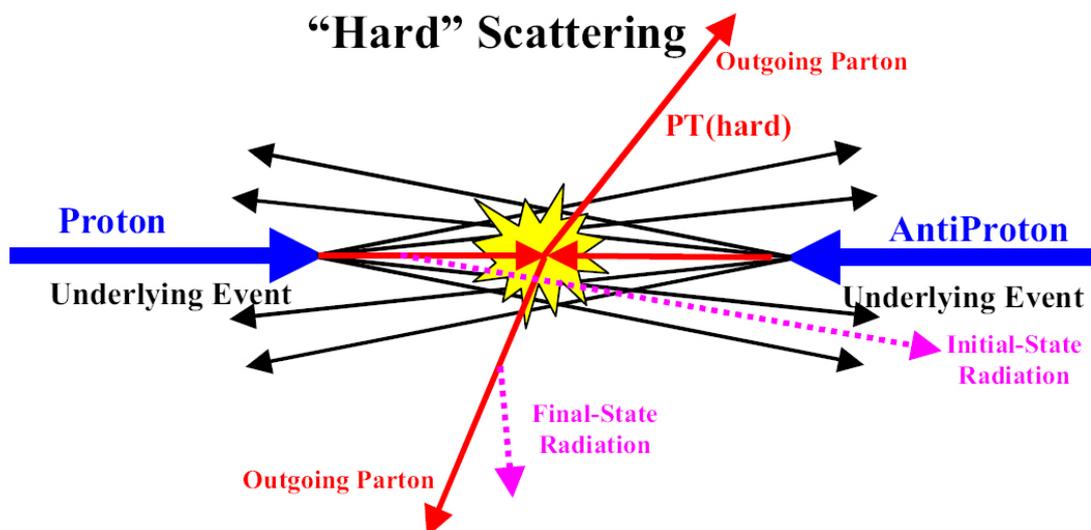


Figure 2.1: Illustration of the proton-antiproton collision with 'hard' 2-to-2 parton scattering. The resulting event contains particles that originate from the two outgoing partons (plus initial and final-state radiation) and particles that come from the breakup of the protons and antiprotons (beam remnants)

Partons in hard inelastic scattering interact directly. Hadrons decay and relatively large amount of particles is produced. Initial partons from hard subprocess are developing due to quark and gluon radiation and then fragment to the so-called JET (see Fig. 2.2) [20]. Hadrons in the jet have small transverse momenta with respect to the direction of the par-

ent parton and sum of the longitudinal momenta with respect to the direction of the parent parton is approximately equal to momentum of the parent parton. Thus, jets can be thought as the 'fingerprints' of the underlying partons. However, although we would like to associate number of the final hadrons with the jet from a single scattered parton, such mapping cannot be, in principle, precise. Jet is not intrinsically well-defined. The partons (quarks and gluons) carry color charge and are massless in the theoretical calculation. On the other hand, a jet of hadrons has no color charge and large invariant mass. Jets must arise from the coherent, collaborative activities of at least two partons - jet can not be the residue of a single parton. So jets are a little ambiguous objects and we need to treat them in such a way that these ambiguities do not play an important role.

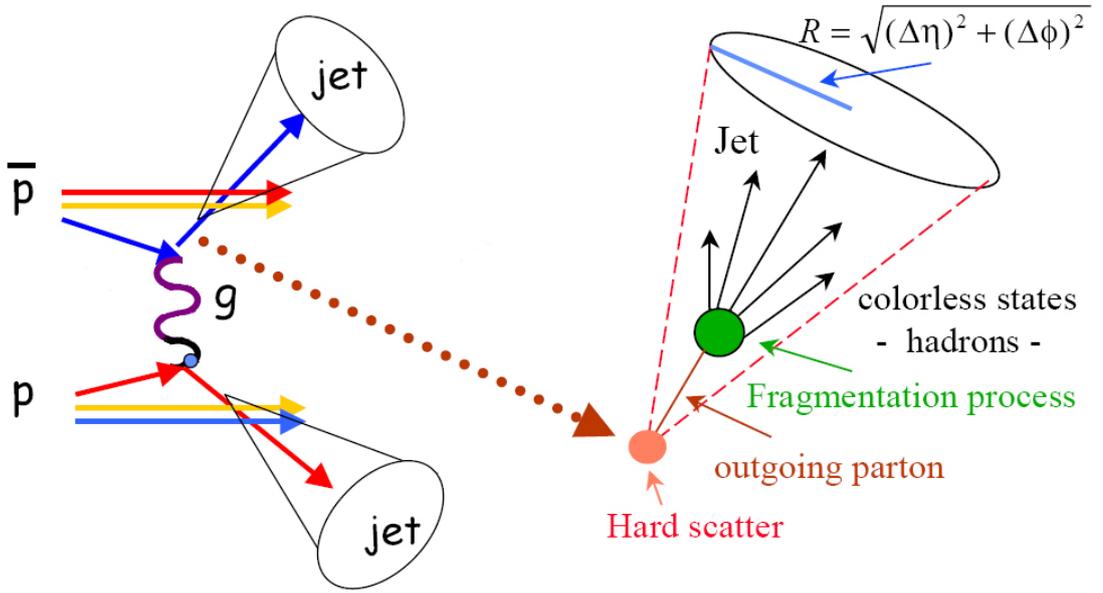


Figure 2.2: Formation of two jets in hard inelastic scattering.

Hard inelastic scattering induce a big momentum transfers and allows us to probe the inner structure of hadrons.  $\alpha_s$  (QCD coupling constant) can be in this regime small (smaller than 0.3) due to asymptotic freedom and therefore it is possible to use perturbative methods for description.

Measurement of cross sections and other jet properties could be used for testing predictions of perturbative QCD, looking for deviations from the Standard Model at very short distances (e.g. due to compositeness) and also for improvement of parton distribution functions at large distances. Finally, one can study the jets in detail in order to differentiate quark and gluon induced jets [11][10]. Gluon jets have higher multiplicity ( $R_{ch} \equiv \frac{\langle n_{ch}(gluon) \rangle}{\langle n_{ch}(quark) \rangle} = 1.19 \pm 0.04(stat) \pm 0.02(syst)$  [12]), softer fragmentation and are broader in  $(\eta, \varphi)$  at the same  $E_T$  as quark jets (see Fig. 2.3).

## 2.2 Jets in nucleus-nucleus collision

Jets in heavy ion collisions are strongly interacting with constituents of the medium where are created. Partons (which we observe indirectly as clusters of hadrons - jets) are losing

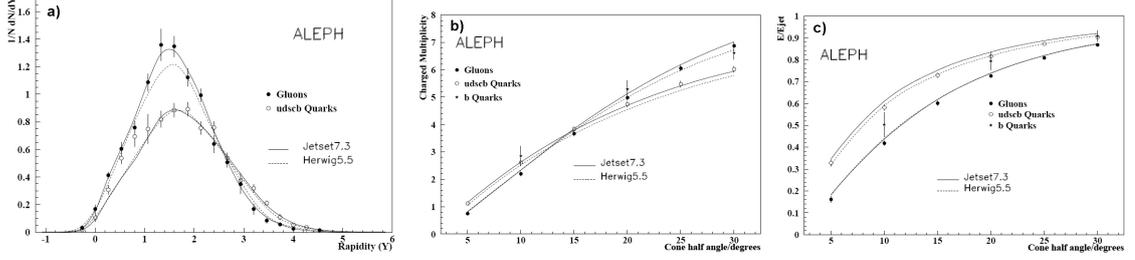


Figure 2.3: Rapidity distributions for natural flavour mix quark and gluon jets (a), integrated charged multiplicity (b) and integrated energy fraction within successive cones (c).

energy due to bremsstrahlung and due to rescattering of partons on medium constituents. Intensity of rescattering strongly increases with rising temperature, hot and dense nuclear originating in high-energy heavy-ion collisions can cause significant energy losses [13]. Deconfined matter generates even higher energy loss compared to hot and dense hadronic medium [21].

The direct reconstruction of jets in nucleus-nucleus collisions at RHIC energies is very difficult due to the presence of a large background of soft partons, but it can be changed at LHC, where significant amount of high- $p_T$  jets (with  $p_T$  over 10GeV/c) will be produced. Although the direct reconstruction of jets at RHIC is, in practice, not possible, indirect methods for study interactions between parton and nuclear matter have been used. Such methods are for example observation of the suppression of high- $p_T$  hadrons and the analysis of two-particle azimuthal correlations.

### 2.2.1 Suppression of high- $p_T$ hadrons

Suppression of high- $p_T$  spectra has been observed in central Au+Au collisions at RHIC [22]. This should be expressed in terms of a nuclear modification factor  $R_{AA}(p_T)$ :

$$R_{AA} = \frac{d^2 N^{AA}/dp_T d\eta}{\langle N_{binary} \rangle d^2 N^{NN}/dp_T d\eta} \quad (2.1)$$

as can be seen in Fig. 2.4.  $d^2 N^{AA}/dp_T d\eta$  in Eq. 2.1 is particle  $p_T$  distribution in A+A collisions,  $d^2 N^{NN}/dp_T d\eta$  is particle  $p_T$  distribution in hadron-hadron collisions and  $\langle N_{binary} \rangle$  is the average number of binary nucleon-nucleon collisions corresponding to a given centrality. At high transverse momenta and without the effect of medium, it is only simple superposition of nucleon-nucleon collisions.  $R_{AA} \approx 1$  is expected at high  $p_T$ ,  $R_{AA} < 1$  for  $p_T < 2\text{GeV}/c$ . We have to mention here the Cronin effect [23, 24] which cause  $R_{AA} > 1$  for  $p_T > 2\text{GeV}$  in low-energy AA collisions due to a multiple parton scattering.

However,  $R_{AA} < 1$  was observed at high transverse momenta for central collisions and  $R_{AA} \approx 1$  for more peripheral collisions [25]. The energy loss is proportional to square of length of color medium transversed. This mechanism reduce  $p_T$  of leading partons and therefore reduce  $p_T$  of leading particles in the jets.

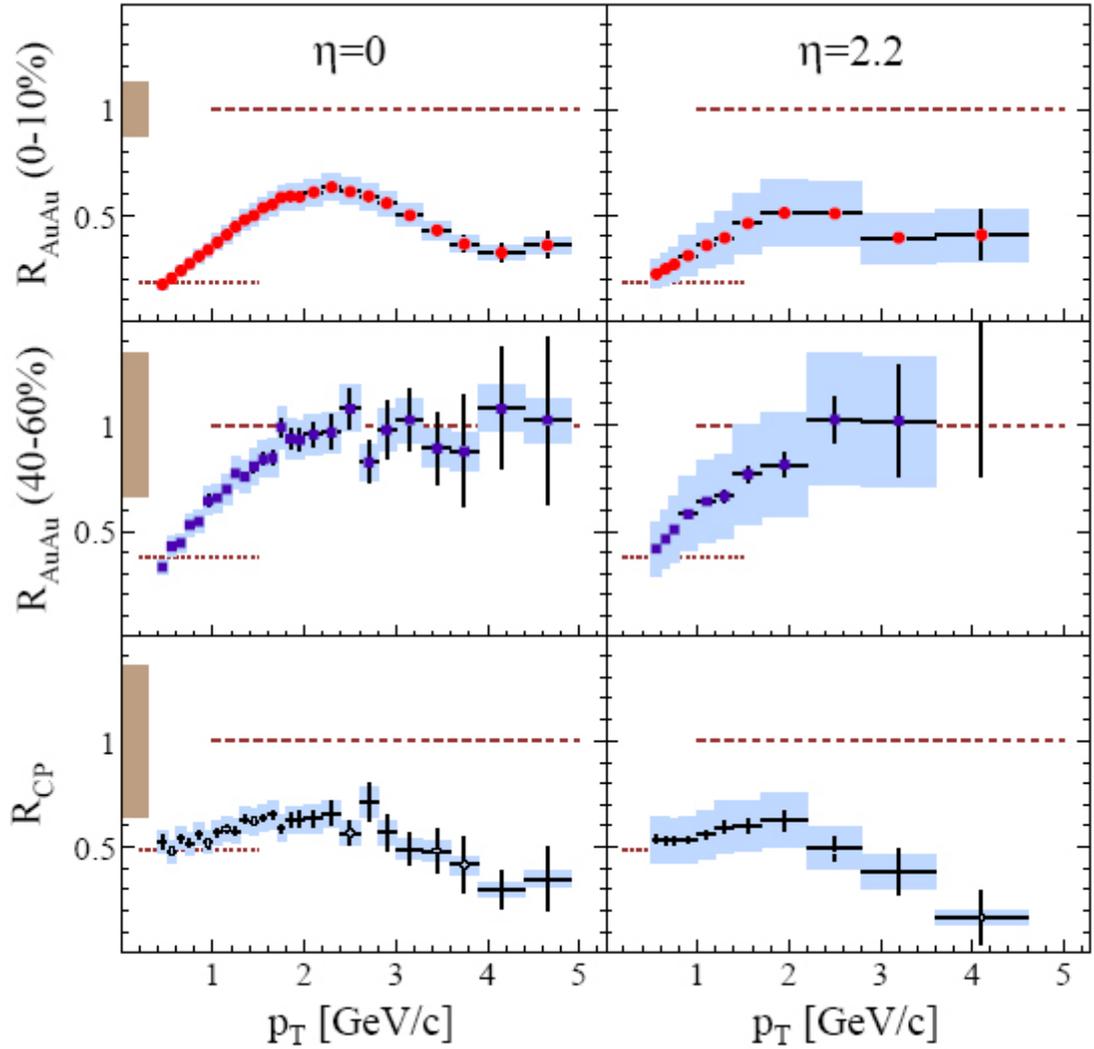


Figure 2.4: Nuclear modification factors  $R_{AuAu}$  measured by BRAHMS for central (top row) and semi-peripheral (middle row) Au+Au collisions at midrapidity (left) and forward pseudorapidity (right). Strong suppression of the high  $p_T$  component above  $p_T > 2\text{GeV}/c$  is visible at both rapidities. The lower row shows the factor  $R_{cp}$ , i.e. the ratio of the  $R_{AuAu}$  for central and peripheral collisions. This ratio has the property of being independent of the p+p reference spectrum.

## 2.2.2 Two-particle azimuthal correlations

Inspired by the discovery of the high- $p_T$  suppression, detailed studies of the jet structure have begun. STAR and PHENIX studied jets by means of two-particle azimuthal correlations. Azimuthal angular correlations of charged particles (with  $p_T > 2\text{GeV}/c$ ) relative to particles with  $p_T > 4\text{GeV}/c$  (trigger hadron) in 200GeV Au+Au collisions are compared in Fig. 2.5 [26, 27, 28]. Nearside peak ( $\Delta\varphi = 0$ ) is very similar in Au+Au, d+Au and p+p col-

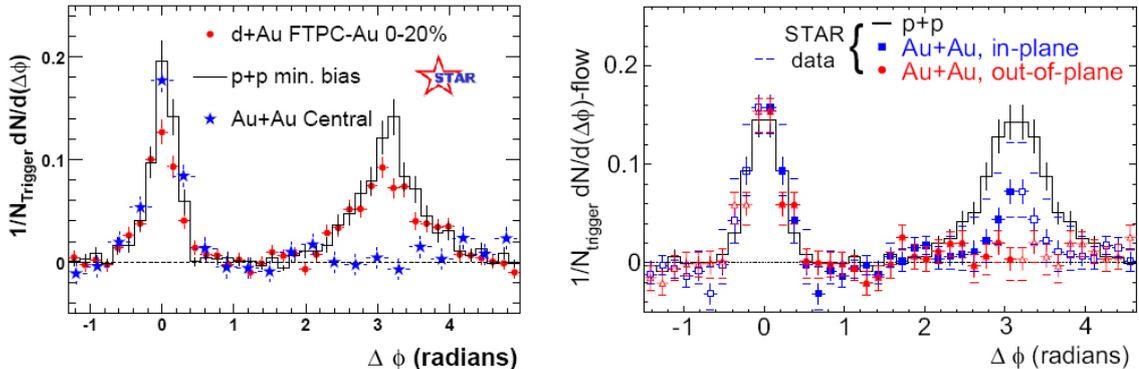


Figure 2.5: Dihadron azimuthal correlations at high  $p_T$ . Left panel shows correlations for p+p, central d+Au and central Au+Au collisions (background is subtracted) from STAR. Right panel shows the background-subtracted high  $p_T$  dihadron correlation for different orientations of the trigger hadron relative to the Au+Au reaction plane

lisions. This is typical for a jet produced by a parton fragmentation process. Awayside peak ( $\Delta\varphi = \pi$ ) disappears in central Au+Au collisions compared to d+Au and p+p collisions. Because the effect is not observed in the central d+Au collisions and peripheral Au+Au collisions, suppression has some relation to interactions between parton and hot and dense nuclear matter. This effect is called jet quenching.

## 2.3 Jets in heavy-ion collisions at the LHC

There are two fundamentally new features in central Pb-Pb collisions at the LHC, as compared to jet physics at RHIC. The multi-jet production per event is not restricted to the minijet region ( $E_T < 2\text{GeV}$ ) but extends to about 20GeV and jet rates are high at energies at which jets can be distinguished from the background energy of the underlying event. Hence, event-by-event reconstruction of jets will be possible [29].

High sensitivity to the medium properties is expected from studies of modifications of the the reconstructed jets structure - for example decrease of the number of particles carrying a high fraction ( $z$ ) of the jet energy, an increase of the number of low-energy particles with low  $z$  values and also broadening of the distribution of jet-particle momenta perpendicular to the jet axis ( $j_T$ ).

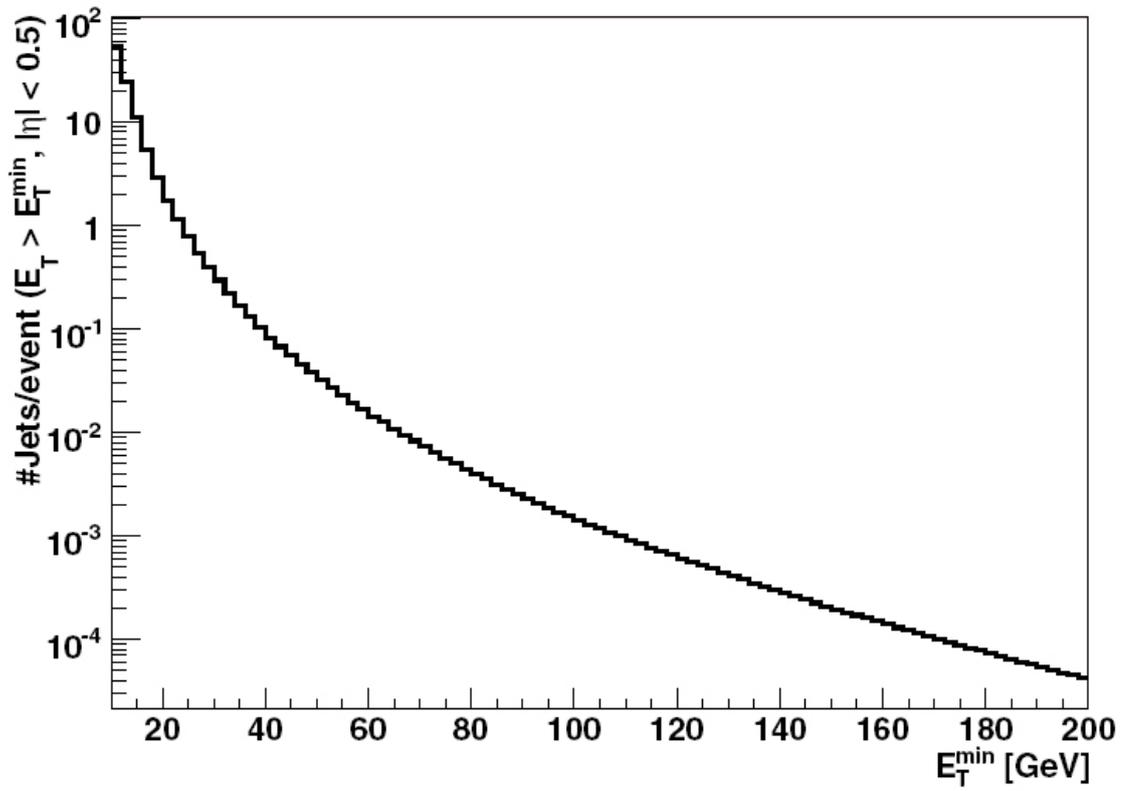


Figure 2.6: Average number of jets with  $E_T > E_T^{\min}$  and  $|\eta| < 0.5$  per event in the 10% most central Pb-Pb collisions.

# Chapter 3

## Jet-finding algorithms

Jets are important objects for quantitative understanding of the underlying strong-interaction theory, QCD, underlying the observed p+p or nuclear collisions. Role of the jet algorithm is to assign clusters of particles (or calorimeter towers and hadrons at experimental level, partons in QCD calculations) to the jets, so that kinematic variables of the jets (e.g. momenta) can be related to the corresponding properties of partons in hard scattering process. Jets allow us 'to see' partons (or at least their fingerprints) in hadronic final states.

We can choose a set of particles in one event (usually set of particles, which are emitted close to each other in the angle) and combine their momenta into momentum of the jet. This process is called JET ALGORITHM and the rule for momentum combination is called RECOMBINATION SCHEME. These two steps are logically different. For example, we could use one set of kinematic variables for assigning particles to the jet (jet algorithm) and then construct another set of kinematic variables to characterize identified jet (recombination scheme).

Number of different jet algorithms have been developed. Each algorithm has to fulfill several basic requirements [17], for example infrared (Fig. 3.1) and collinear stability (Fig. 3.2 and Fig. 3.3).

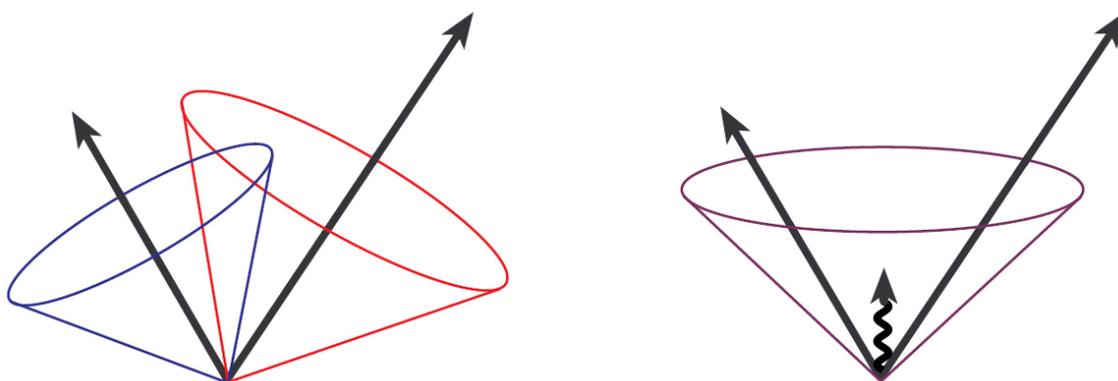


Figure 3.1: Illustration of the infrared instability. Two jets are recognised as a one due to soft radiation.

We will distinguish two basic groups of jet algorithms - recombination (cluster) algorithms and cone algorithms. Recombination algorithms are often called ' $k_T$ ' algorithms because of one successful algorithm from this group, developed back in 1991. Both groups are based on assumption of 'closeness' of hadrons to each other. The definition of cone algorithms is

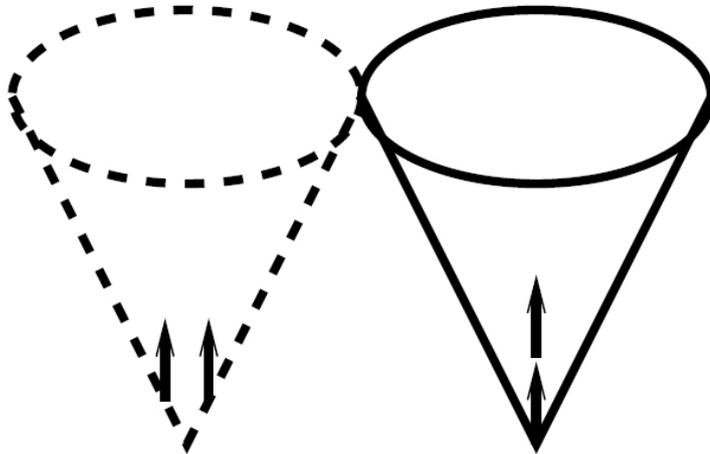


Figure 3.2: Illustration of the collinear instability. In the left case jet-finding failed due to splitting of the energy into two detector cells. Energy in each cell is then lower than certain threshold and cells are not able to create seeds.

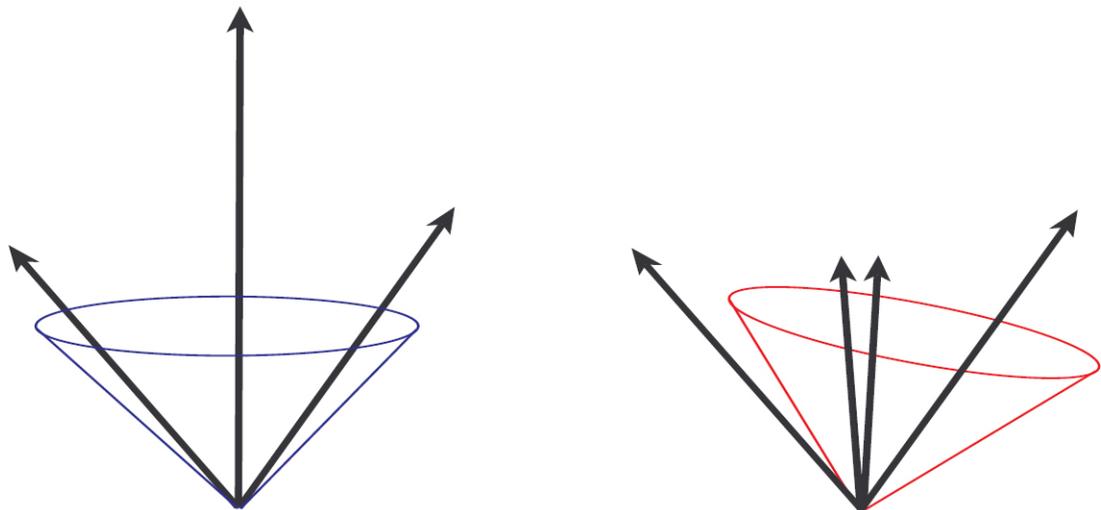


Figure 3.3: Illustration of another collinear problem. In the right case the seed placed higher in the list has been used for jet reconstruction and the left particle has been excluded from the jet.

based on 'closeness' in the real space (angles), while recombination algorithms are based on 'closeness' in momentum space.

### 3.1 Cone algorithms

Historically, cone algorithm was developed for hadron-hadron collision experiments (SNOW-MASS [16]).

Cone jet with radius  $R$  is composed of particles (we assume no bending in magnetic field) which are lying in area  $A = \pi R^2$  in the  $\eta \times \varphi$  space ( $\eta$  is a pseudorapidity, defined  $\eta = -\ln \tan \frac{\theta}{2}$ ;  $\varphi$  is the azimuthal angle,  $\theta$  is the polar angle). Coincidence of the cone axis with the jet direction is required. Jet direction is defined as  $E_T$  weighted center of mass of the particles within the cone. We are searching for all such 'stable' cones to define the jet content of an event. However, in order to save computing time, one is searching for those cones only about the most energetic particles in the event (they are called seeds) - usually of the kinetic energy over few hundred MeV. One can calculate  $E_T$ -weighted center of mass for the particles in each seed cone and then are these centers of mass used as centers for new cones.

$$R \equiv \sqrt{(\varphi_i - \varphi_0)^2 + (\eta_i - \eta_0)^2} \quad (3.1)$$

,where  $\varphi_0, \eta_0$  is center of the cone and  $\varphi_i, \eta_i$  are coordinates of partons, centers of calorimeter towers, etc.

Parton or energy in the calorimeter towers is associated with the jet, if  $R \leq R_0$

$$\eta_{jet} = \frac{1}{E_T} \sum_{i \in R \leq R_0} E_{Ti} \eta_i \quad (3.2)$$

$$\varphi_{jet} = \frac{1}{E_T} \sum_{i \in R \leq R_0} E_{Ti} \varphi_i \quad (3.3)$$

Iteration continue for each cone until the cone axis coincides with the computed  $E_T$  weighted center of mass (aligning of centers is called 'stable' cone). There is no exact rule for determination of cone size  $R_0$ . Simulations[15,16] of jet fragmentation for  $E_T$  over 20GeV suggest values in region 0.4 - 1.0, where effects of the hadronization and the influence of underlying events are minimized[10,11]. Optimum for  $R_0$  is close to 0.7. This is value preferred in the most of cases, but some measurements are using different values[3,12].

However, things are not so simple, a single particle may belong to several cones. Therefore, cone finding algorithm has to involve procedure to specify how to split or merge overlapping cones [30]. For example, two clusters should be separated, if some local minimum is found between two energy peaks[9].

Good algorithm has to include splitting/merging (S/M) rule for protojets (stable cones which are candidates for jets) with overlapping cones. One approach is to find all stable cones, then go through a list of protojets (going from protojets with highest  $E_T$  to the protojets with lower  $E_T$ s) and to check the list after each S/M procedure. If a protojet has no overlap with other protojets, it becomes a jet and is not affected by next S/M processes. S/M decision is typically based on fraction of  $E_T$  shared with protojets with lower  $E_T$ . For example, two

protojets with more than 50%  $E_T$  shared are merged to a one jet, others are splitted and particle (calorimeter tower, parton) is assigned to the nearest protojet in  $\eta \times \varphi$  space.

Another problem arise from sensitivity of cone algorithms to a soft radiation. Imagine two seed partons that will just fit inside of a single cone, but at opposite sides. Standard Snowmass-type algorithm will reconstruct two jets from these two partons. A soft gluon could be radiated between these two partons and serve as a seed. Single jet, with both partons inside, will be identified. Recent algorithms, dealing with this problem, were developed, e.g. Improved Legacy Cone Algorithm [17].

Demonstration of typical cone jet algorithm follows:

- (1) creation of seed (3-vector) from the direction of input particle (possibility of implementig way to reducing number of seeds and thus increase of speed)
- (2) for each seed  $s$  create cone in  $\eta \times \varphi$  space with the radius  $R$ , so that particle  $p$  with  $(\eta_s - \eta_p)^2 + (\varphi_s - \varphi_p)^2 < R^2$  is defined inside of the cone
- (3) recombination of each particle in this cone into a jet
- (4) creation of new cone around the axis of the jet and repeating of step (3). If the axis of the new jet is collinear with previous axis, the jet is stable and is added to the list of meta-jets. If axis is not collinear, whole process is repeated until the stable jet is found or until maximal number of iterations is reached.
- (5) repeating steps (2)-(4) with a new set of seeds between each pair of jet  $i, j$  (which were found in previous step), if  $i$  and  $j$  are between 1 and 2 radiuses of cone. This is because of infrared safety  
if  $R^2 < (\eta_i - \eta_j)^2 + (\varphi_i - \varphi_j)^2 < (2R)^2$   
then  $\eta_s = \frac{\eta_i + \eta_j}{2}$   $\varphi_s = \frac{\varphi_i + \varphi_j}{2}$
- (6) every jet with  $p_T$  less then defined parameter (usually around 5GeV/c) is removed from the list
- (7) if summary  $p_T$  of particles shared with higher- $p_T$  jet is higher then defined ratio, jet is removed from the list
- (8) for each particle, which belongs to more than one jet - remove this particle from all jets except the one closest to the particle, i.e. with jet with smallest  $\Delta(\eta)^2 - \Delta(\varphi)^2$
- (9) repeat step 6

## 3.2 $k_T$ algorithms

This class of jet algorithms is inspired by QCD and was developed originally for  $e^+e^-$  collisions.  $k_T$  algorithms merge pairs of particles one by one in the direction of growing transverse momentum. They contain parameter D (or R in some articles), which controls end of merging and characterize approximate size of the jet (similar to R in cone jet algorithms). Every particle is assigned to a unique jet, so there are not problems emerging in cone jet algorithms, where one particle could belongs to several cones. Because of the design of  $k_T$  algorithms, they are really infrared and collinear safe to all orders of QCD calculations. Until recently, they were consuming too much time for computing ( $N^3$  complexity) and thus have little use in hadron-hadron collisions. One of solutions is a preclustering step, where is number of particles significantly reduced before the  $k_T$  algorithm is used. In hadron-hadron collisions,

there were also problems with energy subtraction from spectator fragments and the pile-up from multiple hadron interactions because of irregular shape of  $k_T$  jets.

Several implementations of  $k_T$  jet-finding algorithm exist, here are two examples:

$k_T$  jet finder (1991)

$$d_{ij} = \min(k_{Ti}^2, k_{Tj}^2) \Delta R_{ij}^2, \quad d_{iB} = k_{Ti}^2 R^2$$

Cambridge/Aachen (1998)

$$d_{ij} = \Delta R_{ij}^2, \quad d_{iB} = R^2$$

,where  $i \neq j$ ,  $\Delta R_{ij}^2 = \Delta y_{ij}^2 + \Delta \varphi_{ij}^2$  and  $R$  is the merging parameter.

Now we will go through the  $k_T$  algorithm step-by-step. At the beginning, we have two lists - list of preclusters (usually, the preclustering step is added to reduce number of input particles because of slow speed of the  $k_T$  algorithm) and an empty list of jets.

- (1) construct  $d_{iB}$  and  $d_{ij}$  for all preclusters  $i, j$  ( $\mathcal{O}(N^2)$ , done once)
- (2) find minimum of  $d_{iB}$  and  $d_{ij}$  and define this minimum as  $d_{MIN}$  ( $\mathcal{O}(N^2)$ , done  $N$  times)
- (3) if  $d_{MIN}$  is  $d_{iB}$ ,  $i$  is not mergable, is removed from precluster list and moved to jet list ( $\mathcal{O}(1)$ , done  $N$  times)
- (4) if  $d_{MIN}$  is a  $d_{ij}$ , remove  $i$  and  $j$  from precluster list and replace them with new merged precluster ( $E_{ij}, \vec{p}_{ij}$ ) defined:
 
$$E_{ij} = E_i + E_j$$

$$\vec{p}_{ij} = \vec{p}_i + \vec{p}_j$$
 ( $\mathcal{O}(N)$ , done  $N$  times)
- (5) go to the step (2), if any precluster remains

step (2) dominates, requiring  $\mathcal{O}(N^2 \times N = N^3)$  operations

$\mathcal{O}(N^3)$  is very bad for high-energy heavy-ion experiments - one event containing 50 000 particles (towers, etc.) would be computed in more than one day. Because of this, new generation of  $k_T$  jet finding algorithms was developed, using techniques as Voronoi diagrams [18] with the Delaunay triangulation [19]. First such algorithm is called FastJet [31]. With complexity  $\mathcal{O}(N^2) - \mathcal{O}(N \ln N)$  is faster than any other  $k_T$  algorithm and is as fast as cone algorithms.

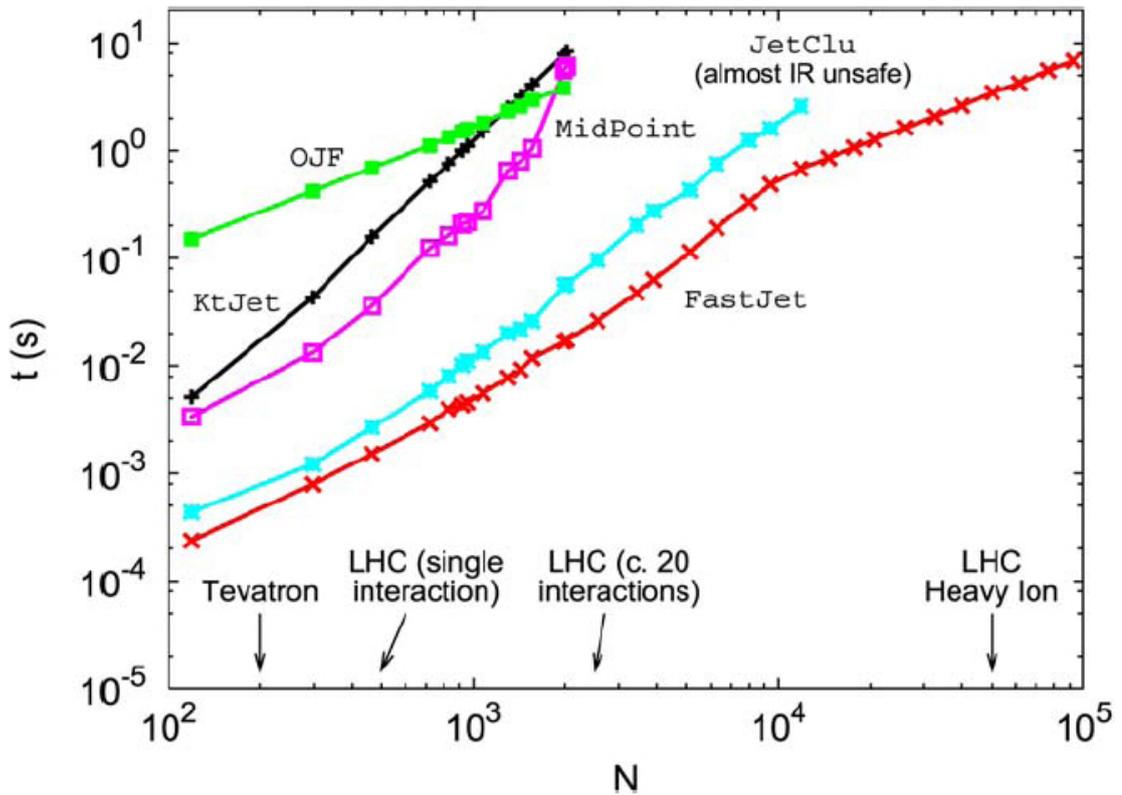


Figure 3.4: The running times of various jet-finders versus the number of initial particles. JetClu is a widely-used cone jet-finder, however it is not infrared stable. MidPoint is an infrared safe cone jet finder. For both code and parameters from CDF have been used. The optimal jet finder (OJF) has been run with  $\Omega_{cut} = 0.15$  and a maximum of 8 jets, so as to produce a final state similar to that returned by the  $k_T$  and cone jet-finders and to limit its run time.

# Chapter 4

## MRA jet-finder

### 4.1 Principles of MRA jet-finder

Vojtěch Petráček proposed a method for local analysis with variable resolution, capable of detection of localised domains (areas with different density distribution of charged and neutral particles) [1]. This method looks promising also in other areas, e.g. in searching for droplets (spinodal instability of QGP fireball [14]). The aim of this thesis is to test, if this method is applicable also for jet detection and reconstruction.

Detection method is based on application of a two-dimensional Cauchy-Lorentz distribution function [32]. Input particles are approximated by  $p_T$  weighted Cauchy-Lorentz function and all functions are summed up into amplitude function.

Let  $L(\eta, \varphi, \epsilon_\eta, \epsilon_\varphi)$  is two-dimensional Cauchy-Lorentz function, where  $\epsilon_\eta, \epsilon_\varphi$  are resolutions of  $\eta, \varphi$  respectively.

Regions with higher particle density (and with higher  $p_T$  per particle) are then highlighted and can be isolated when the amplitude of summed Cauchy-Lorentz functions exceed the certain threshold. The basic idea is to get optimal resolution (reflected in the  $\gamma$  parameter of the Cauchy-Lorentz function - see section Lorentz) for identification of the jet - a low frequency object made from individual particles. It is possible to make amplitude function at different resolution scales and make differences between them:

$$F(\eta, \varphi, \epsilon_\eta^1, \epsilon_\varphi^1, \epsilon_\eta^2, \epsilon_\varphi^2) = L(\eta, \varphi, \epsilon_\eta^2, \epsilon_\varphi^2) - L(\eta, \varphi, \epsilon_\eta^1, \epsilon_\varphi^1) \quad (4.1)$$

, where  $\epsilon_\eta^1 > \epsilon_\eta^2$  and  $\epsilon_\varphi^1 > \epsilon_\varphi^2$   
when we write:

$$L(\eta, \varphi, \epsilon_\eta^1, \epsilon_\varphi^1) = L(\eta, \varphi, \epsilon_\eta^2, \epsilon_\varphi^2) + F(\eta, \varphi, \epsilon_\eta^1, \epsilon_\varphi^1, \epsilon_\eta^2, \epsilon_\varphi^2) \quad (4.2)$$

we get well-known formula Eq. 1.2 from chapter (MRA) - analysis of function to a row of detail functions. In this sense,  $L$  is equivalent of the mother function and  $F$  is corresponding to the father function from wavelet analysis and MRA. We can use function  $F$  instead of  $L$  for jet-finding, in this case the mean value is zero and effects with characteristic size larger than coarser resolution scale  $\epsilon^1$  are isolated - as is for example elliptic flow.

But this method is not exactly MRA, it uses only basic principles of MRA, because the Cauchy-Lorentz functions we are using are not orthogonal.

If we normalize  $\eta \times \varphi$  space  $((-\pi, +\pi) \times (-\eta_0, +\eta_0) \rightarrow (0, 1) \times (0, 1))$ , we have to also transform particle coordinates:

$$x_{hit} = \frac{\varphi + \pi}{2\pi} \quad (4.3)$$

$$y_{hit} = \frac{\eta + \eta_0}{2\eta_0} \quad (4.4)$$

Then we discretize this normalized space (Cauchy-Lorentz function is continuous, but all computations of this function have to be done in discretized coordinates). We used the same segmentation for both coordinates. If we want to split each coordinate into  $s$  pieces, we get  $(s \times s)$  segments of our normalized  $(\eta \times \varphi)$  space. We should identify this segments by two integers:  $i = 0, 1, \dots, s - 1$ ,  $j = 0, 1, \dots, s - 1$  and centres of this segments will have coordinates  $[i + 0.5, j + 0.5]$

We will calculate amplitude function in each of segments. Particle  $k$  contributes to amplitude in segment  $[i, j]$  by:

$$\frac{p_T[k]}{\gamma(1 + \frac{((i+0.5)dx - x_{hit}[k])^2 + ((j+0.5)dy - y_{hit}[k])^2}{\gamma^2})} \quad (4.5)$$

,where  $dx$  and  $dy$  is a size of step in normalized  $\varphi$  and  $\eta$  coordinate respectively,  $\gamma$  (see sect Lorentz) is a scale parameter of the Cauchy-Lorentz function (half-width at half-maximum) and it is an average size of step in normalized coordinates ( $\gamma = 0.5(dx + dy)$ ), in our case  $\gamma = dx = dy$ .

We chose  $dx = dy = \frac{1}{s}$ . We can identify the parameter  $s$  as an inverse resolution, so we will speak about division  $s = 10$ ,  $s = 80$ , etc. These values of parameter  $s$  are not random, they are the lowest and the highest resolution parameters we tested. As we can see, we discretized normalized  $\eta \times \varphi$  space by this parameter - this is because at low resolution (low values of parameter  $s$ ) we have smooth, slowly fluctuating function  $L$  and we may afford discretization of normalized  $\eta \times \varphi$  into smaller number of segments. On the other hand, at high values of parameter  $s$  is the function  $L$  fast fluctuating and we need finer discretization of normalized  $\eta \times \varphi$  space. In the limit  $s \rightarrow \infty$  these fluctuations goes to delta-functions and individual particles can be seen.

Computational procedure consist of a cycle over all particles and for each particle is calculated contribution to the function  $L$  according to equation (ta pred chvili) for corresponding segments of normalized  $\eta \times \varphi$  space. We do not need to put contributions of Cauchy-Lorentz function of one particle to all segments, because Cauchy-Lorentz function is decreasing rapidly. We calculated contributions of amplitude function only for segment to which particle belongs and for three segments in each direction (thus the amplitude is stored in a block 7x7 segments). When we extended this to four segments in each direction, resulting amplitude function for whole event changed approximately about 1% and this can be neglected.

Value of function  $L$  for scale parameter  $s$  in the segment  $[i, j]$  can be written therefore as:

$$L(i, j, s) = \sum_{k \in O} \frac{s \cdot p_T[k]}{1 + (i + 0.5 - s \cdot x_{hit}[k])^2 + (j + 0.5 - s \cdot y_{hit}[k])^2} \quad (4.6)$$

,where  $O$  is a set of particles lying in segments  $(i - 3, i + 3) \times (j - 3, j + 3)$ .

Because of possibility to write core part of this algorithm as one cycle over all input particles, this MRA based method is possibly the fastest jet-finding mechanism ever, with complexity of  $\mathcal{O}(N)$  (however, there are other important aspects to test).

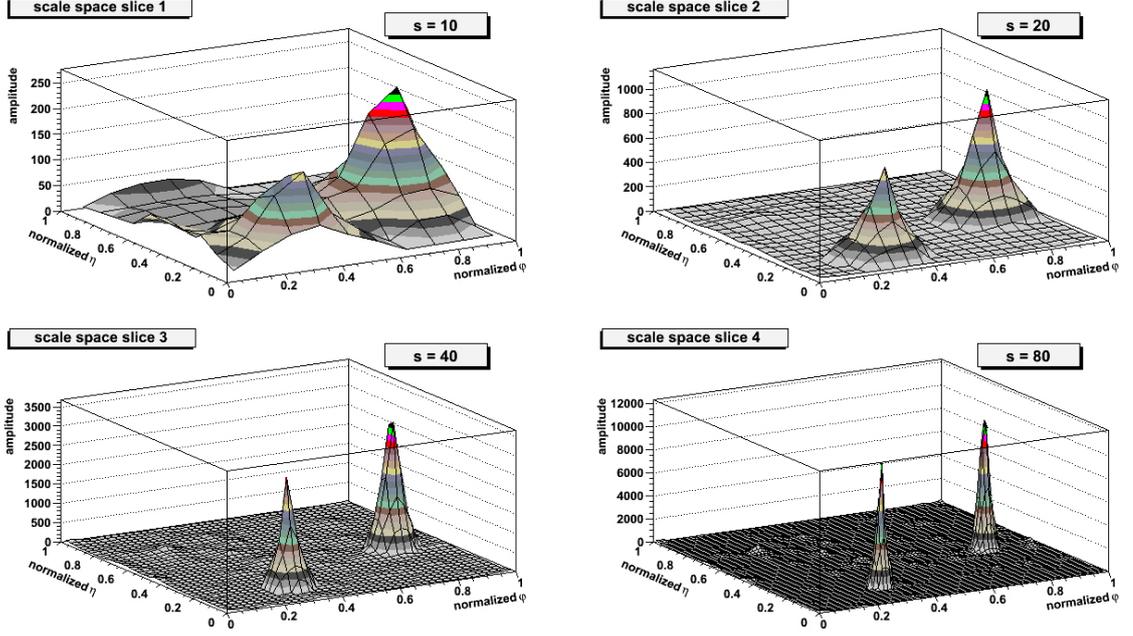


Figure 4.1: Example of the  $L$  function for di-jet event with average jet  $p_T = 37\text{GeV}/c$ .

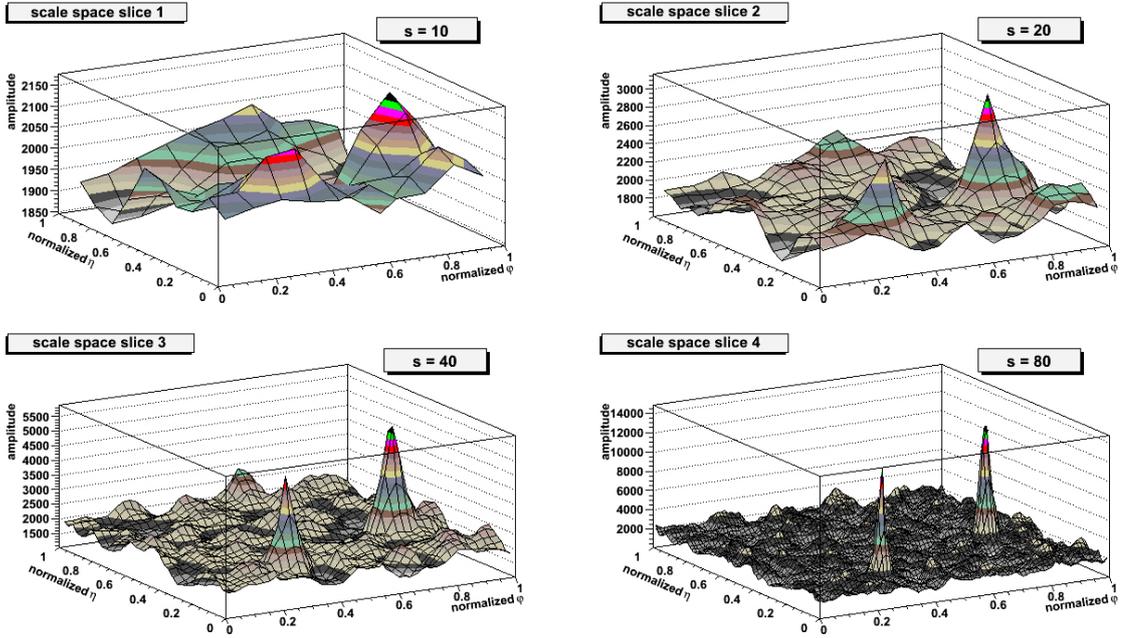


Figure 4.2: Example of the  $L$  function for di-jet event with average jet  $p_T = 37\text{GeV}/c$  with random background  $dN/d\eta_{ch} = 2000$ , evenly distributed in normalized  $\eta, \phi$  coordinates and exponentially ( $Exp(0.5)$ ) distributed in  $p_T$ .

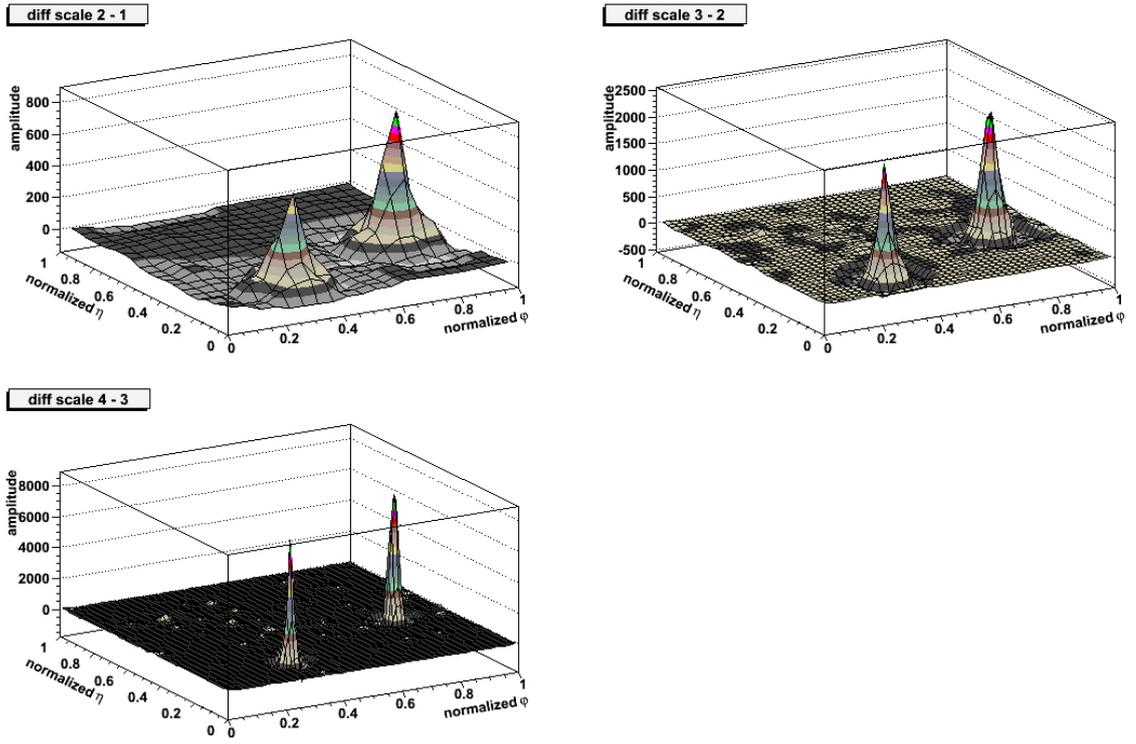


Figure 4.3: Example of the  $F$  function for di-jet event shown in Fig. 4.1.

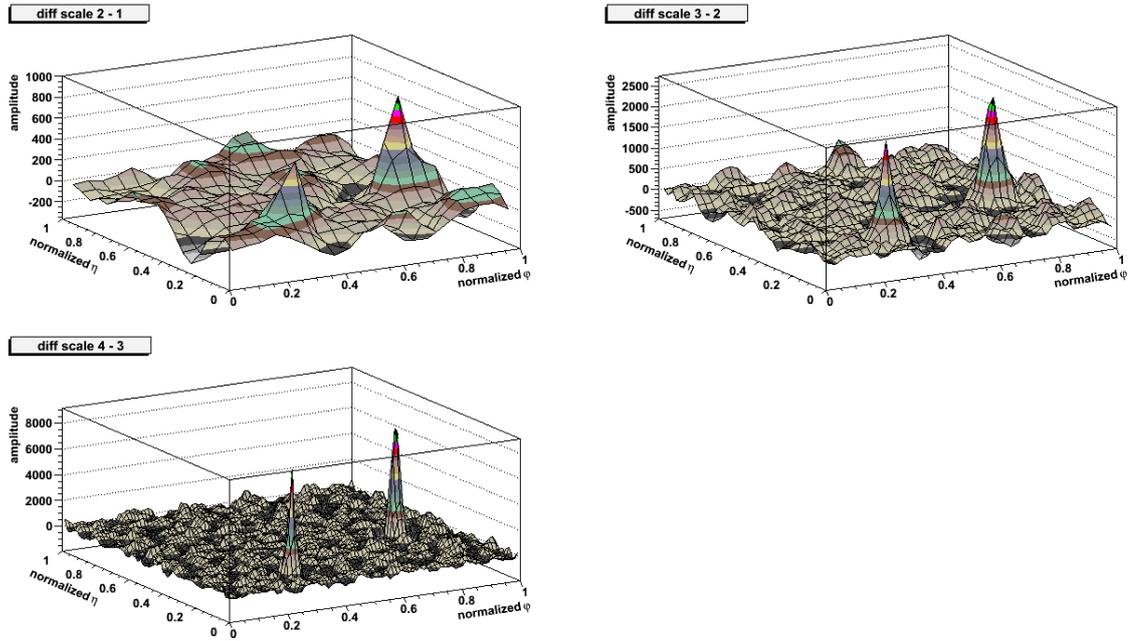


Figure 4.4: Example of the  $F$  function for di-jet event shown in Fig. 4.2.

## 4.2 Cauchy-Lorentz distribution

The Cauchy-Lorentz distribution, named after Augustin Cauchy and Hendrik Lorentz, is a continuous probability distribution. It is known also as the Breit-Wigner distribution. It is widely used in physics - it is the solution of the differential equation describing forced resonance and has its use for example in spectroscopy (homogeneous broadening of spectral lines) and particle physics (non-interfering cross-section of particle resonant states). The Cauchy-Lorentz distribution is symmetrical and so-called 'heavy-tailed' (it means that a high proportion of the population is comprised of extreme values) - a normal (Gaussian) distribution decreases much faster in the tails. The fact the Cauchy-Lorentz function is 'heavy-tailed' can be advantage for description of long-range fluctuations in the MRA based jet-finder, the Cauchy-Lorentz function can correlate far hits in the  $\eta \times \varphi$  plane. The Cauchy-Lorentz function can be seen as a special case of the Levy Skew alpha-stable distribution (with parameters  $\alpha = 1$  and  $\beta = 0$  with scale parameter  $c$  and shift parameter  $\mu$ )

Probability density function of the Cauchy-Lorentz distribution:

$$f(x, x_0, \gamma) = \frac{1}{\pi\gamma[1 + (\frac{x-x_0}{\gamma})^2]} = \frac{1}{\pi} \left[ \frac{\gamma}{(x - x_0)^2 + \gamma^2} \right] \quad (4.7)$$

,where  $x_0$  is the location parameter, specifying the location of the peak;  $\gamma$  is the scale parameter, which define the half-width at half-maximum.

The Cauchy-Lorentz distribution has no mean, variance of other moments defined. Its mode and median are equal to  $x_0$ .

## 4.3 Boundary conditions

As mentioned in the introduction to this chapter, we are representing particles (with  $[\eta, \varphi, p_T]$  coordinates) by  $p_T$  weighted two-dimensional Cauchy-Lorentz function. Then we are filling the corresponding amplitudes into the two-dimensional array (coordinates of cells are corresponding to the normalized and segmented  $\eta, \varphi$  coordinates, while the value stored in the cell is corresponding amplitude of summed contributions of individual particles). We have to be careful with boundary conditions, when we are filling this array (and also when we are identifying jets in this array). We have several choices how to handle the boundary conditions[4]:

$$\begin{aligned} c(k + N) &= c(N - k) \text{ (mirror)} \\ c(k + N) &= c(k) \text{ (periodicity)} \\ c((k + N)) &= c(N) \text{ (continuity)} \end{aligned}$$

We used cyclic (periodic with periode  $N = 2\pi$ ) boundaries in the normalized  $\varphi$  coordinate, as one could expect. More difficult matter is the boundary in the  $\eta$  direction. We used so far only cyclic boundaries.

## 4.4 Threshold

Basic idea of our MRA based jet-finder is to make a function  $L$  and then observe, where this function exceeds certain threshold. So, the resolution is not only important parameter of

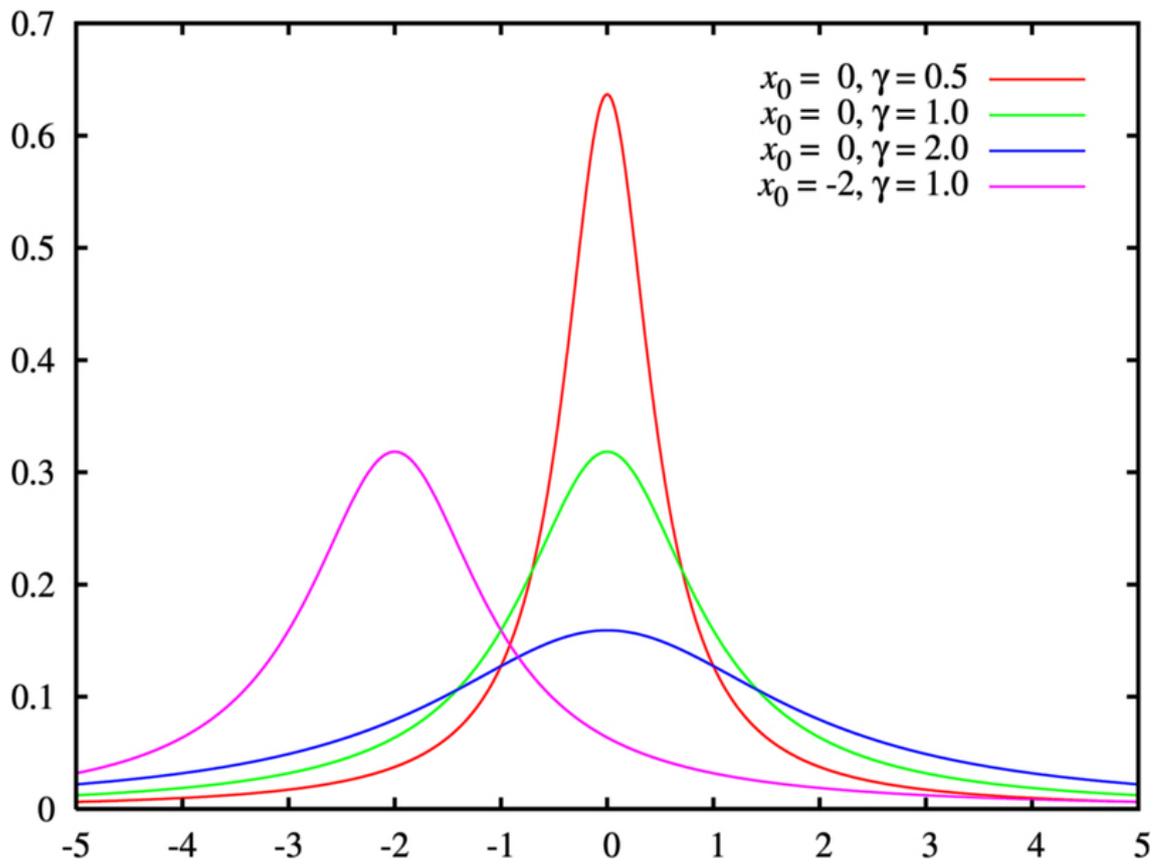


Figure 4.5: Probability density function for Cauchy-Lorentz distribution. The green line is the standard Cauchy distribution.

this jet-finder, but also the definition of the threshold. When we were calculating amplitude contributions of individual particles, we also summed them up into one variable, let's mark it  $\Sigma$ . Then we calculated the global mean value of the function  $L$  in the whole normalized  $\eta \times \varphi$  space as

$$mean = \frac{\Sigma}{s^2} \quad (4.8)$$

, where  $s$  is the scale parameter. After then, it is possible to evaluate threshold as follows:

$$threshold = mean + 3 \frac{s}{10} \sqrt{mean} \quad (4.9)$$

When the amplitude stored in the cell exceeds this threshold, the cell is marked as a candidate for place, where the jet can be.

There are many possibilities how to define the threshold. In the ideal case, we want to make mean value only from background particles. But this is not, of course, possible. Cells of the amplitude array, where are main contributions of jet particles have much higher value than others. We know (viz testovani), that jet-candidates (areas with amplitude over threshold, where jets should be) in one event occupies only few percent of the total area of normalized  $\eta \times \varphi$  space. We can compute then the mean value from 90% cells with lower amplitude, due to  $p_T$  weight this corresponds mostly to the soft  $p_T$  background.

## 4.5 Domain detection

We know, which cells are candidates for jets (cells of amplitude array over the threshold), but they can be isolated fluctuations of the  $L$  function. Jet will be seen as a group of such cells, creating a continuous area of approximately circular shape. We implemented algorithm, which is finding neighbouring cells over threshold.

Algorithm description: amplitude array is searched and each passed cell is marked (thus it will not be processed again). If the cell has been marked as a jet candidate, algorithm will search in 8 surrounding cells (i.e. cells which have common edge or corner with the current cell) for other candidates. These other candidate cells are joined to one jet together with the previous cell and algorithm will search also in their neighbourhood. When no candidate cells remain in surroundings of the jet, algorithm will continue in searching the amplitude array for another jet until whole amplitude array is searched.

This simple algorithm can be modified, for example only cells with common edge can be merged together to form a jet. It may reduce situations where close background fluctuations were added to the jet.

Random fluctuations should be also grouped together, but probability of creating larger areas is very small (rapid exponential decrease). We can base our selection criterion on this property - we can mark particles in the area as a jet if the area is sufficiently large and also if the total sum of amplitudes stored in this area is sufficiently large.

## 4.6 Test jets

So far, the jet-finding algorithm was tested on two files containing di-jets. The di-jet events were generated by Pythia 8.100 Monte Carlo generator [34]. Pythia 8 is still developing, but it is written in the C++ language. Pythia 6.4 [33] is much more stable, but has not

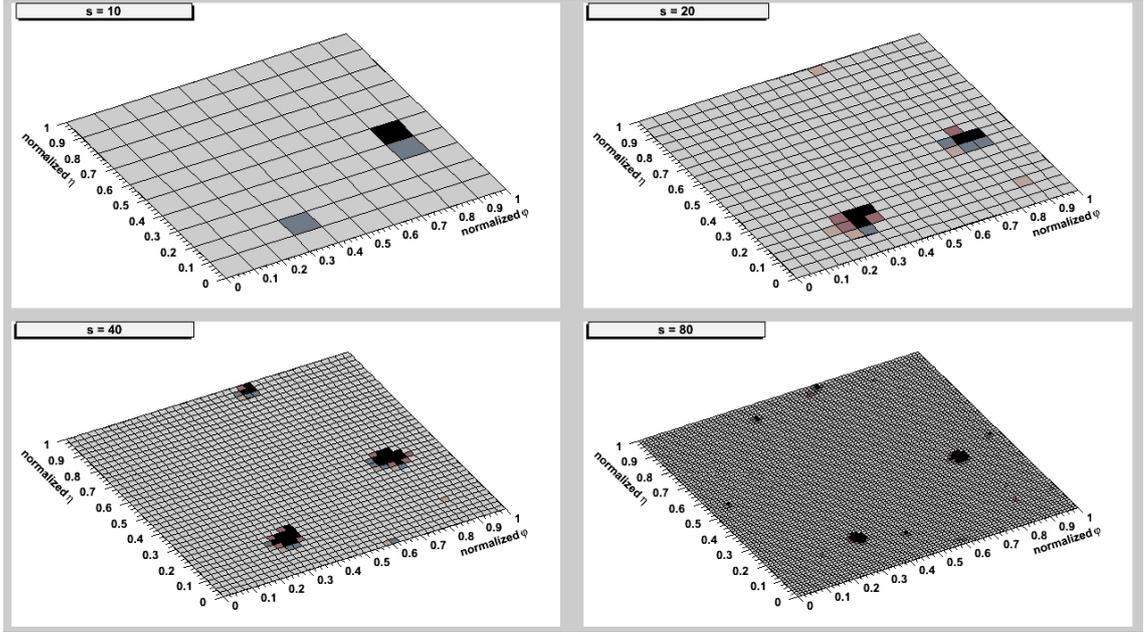


Figure 4.6: Example of the di-jet event shown in Fig. 4.2, only cells with amplitude greater than threshold shown.

C++ interface for its jet finders (ClusterJet and CellJet), so there is no simple way how to call this jet finders from a ROOT script. These interfaces are important for generating a jet sample, because in the Pythia Event List one can not see directly, which particle belongs to which jet. Parameters of the collisions were:

```
HardQCD:all = on
PhaseSpace:pTHatMin = 100.0 //we also used parameter 20.0
pythia.init(2212, 2212, 14000.0)
```

It is the group of p+p collisions at centre-of-mass energy 14TeV with all hard QCD processes allowed and with the minimum invariant  $p_T = 100\text{GeV}/c$  ( $20\text{GeV}/c$  respectively) We have used jet-finding algorithm CellJet implemented in Pythia with default parameters ( $R = 0.7$ ) and saved to the root file only final charged particles. Cone algorithm CellJet has its own limitations and determination of center of generated jet is not accurate at low  $E_T$ .

## 4.7 Background

Only the random background has been used at this current stage of testing of the jet-finding algorithm. We have used ROOT 5.16/00 default random generator (TRandom3 object) based on Mersenne-Twister generator [35]. Background particles have even (Rndm()) distribution in normalized  $\eta$ ,  $\varphi$  coordinates and exponential (Exp(0.5)) distribution in  $p_T$  coordinate. This generated background reproduce the 'soft' background of heavy ion collision.

## 4.8 Testing

Jet finder based on MRA principles has been implemented as a script for ROOT 5.16/00 and tested with di-jets from Pythia 8.100 generator (see section 4.6) merged with random

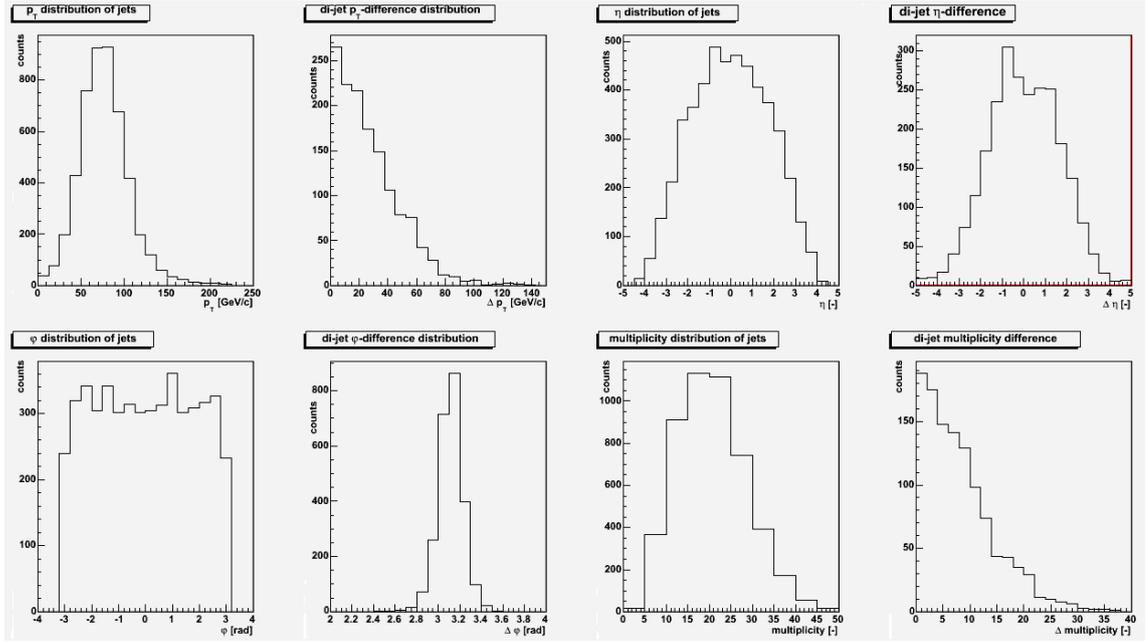


Figure 4.7: Properties of the jet sample generated for  $\hat{p}_{Tmin} = 100$ .

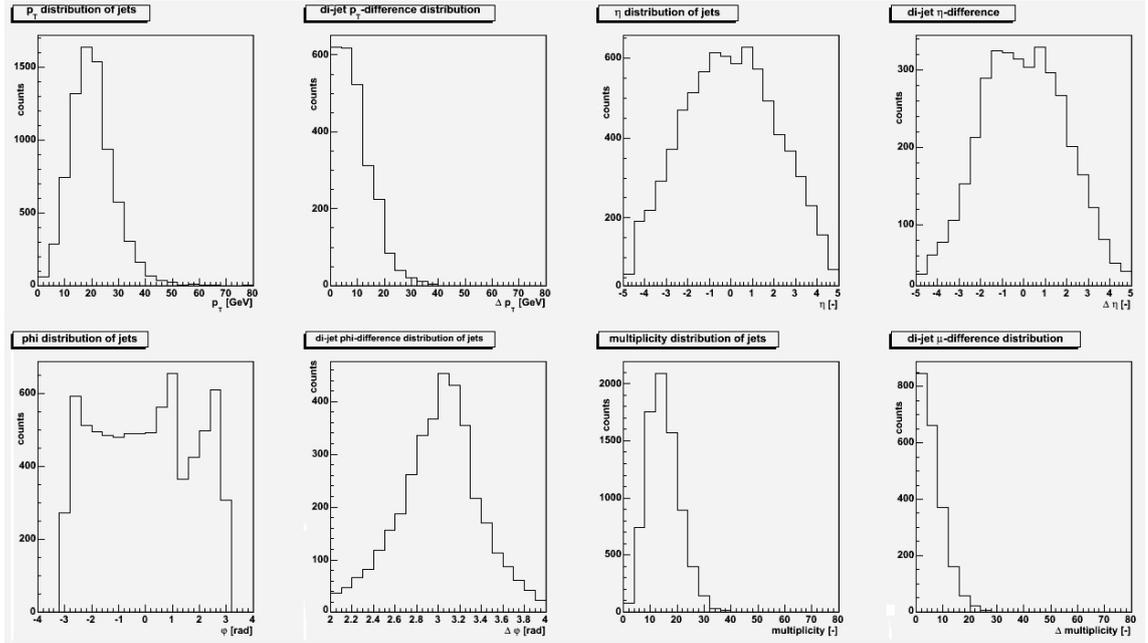


Figure 4.8: Properties of the jet sample generated for  $\hat{p}_{Tmin} = 20$ .

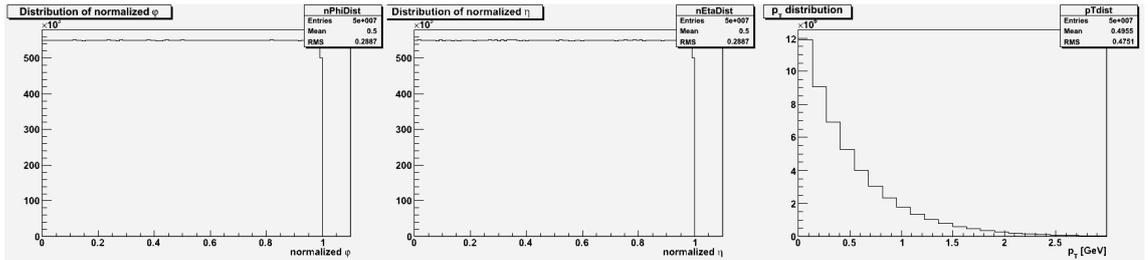


Figure 4.9: Properties of the generated background.

background (see section 4.7). Tests were performed with two jet samples (one containing 3866 di-jets with minimum  $\hat{p}_T = 20\text{GeV}/c$ , second containing 2471 di-jet events with minimum  $\hat{p}_T = 100\text{GeV}/c$ ).

As one could expect, at lower values of scale parameter ( $s < 20$ ) where domains are large, they represent significant fraction of the whole normalized  $\eta \times \varphi$  space. In some cases they merge together (in other words, two jets create one long domain treated as one jet). In some cases one jet of a di-jet event was not recognized because the signal was not able to exceed the threshold - this effect is significant for low-energy jets (their particles are spread in the large part of the  $\eta \times \varphi$  space). Also the cone jet algorithm, which we used for creating input file, has limited use in such low-energy domain.

At high resolutions ( $s > 40$ ) the MRA based jet algorithm loses its ability 'to see' jets as an object and particle structure starts appear (high, narrow peaks for individual particles). This results in state with a large number of small domains, many of them are only fluctuations of the background.

Algorithm has been tested for both di-jet files at 24 different resolutions ( $s = 16, 18, 22, \dots, 56$ ), so far only for small number of background particles (3000 background particles at  $\eta \pm 4$ ). Lower  $dN/d\eta$  does not matter, because more particles mean lower density fluctuations - it is proportional to  $\frac{1}{\sqrt{N}}$ .

This jet-finding method is still under development, where big and essential interventions in code are frequent and computation for one resolution can take several days for 2GHz processor with this number of background particles. However, tests for multiplicity 2000 particles per unit of pseudorapidity are in progress now. Because of statistical nature of this method, we can expect even better results.

More realistic background containing minijets, which will be copiously produced at LHC energies is, of course, the next important step in development of test procedures for validation of our MRA approach. We plan also tests on data from the STAR experiment.

Tests were performed so far only for di-jets, because we have jet-finding method, but yet no splitting/merging rule (as mentioned in chapter blabla) developed. This should be next step in the method development. If we put two jets too close, they merge into one domain (this effect is getting stronger as we go to the lower resolutions and lower jet energies) and we need some rule, which will split them into two domains (maybe based on finding the gap between two peaks of function  $L$  or based on typical domain shape).

### 4.8.1 Background subtraction

Background subtraction is needed for  $p_T$  reconstruction of jets. Currently we have adopted the following approach: When domains are recognized, mean  $p_T$  per unit of area is calculated from all particles which not belong to any domain. Then this value is multiplied by the area of the domain and this value is subtracted from the domain.

### 4.8.2 Reconstruction precision

In this test phase, two types of domains are distinguished - true jets, corresponding to one of the input jets and fake jets created due to background fluctuation. Domain is recognized as a true jet if the  $p_T$  weighted  $\eta, \varphi$  center of one of the input jets lies in that domain. When a domain (area in  $\eta \times \varphi$  with function  $L$  greater than certain threshold) is isolated and is

recognized as a true jet, we can calculate  $p_T$  weighted center and total  $p_T$  of this domain and compare it with the input jet.

In Fig. 4.10 are plotted mean values of differences in the  $\varphi$  coordinate ( $\varphi_{true-jet} - \varphi_{input-jet}$ ). The  $\varphi$  coordinate is not normalized, it is in radians.

In Fig. 4.11 is compared root mean square of the difference in the  $\varphi$  coordinate for all tested momenta. RMS is decreasing for increasing scale parameter. It is decreasing rapidly for  $s < 30$  and for greater  $s$  is the change of the RMS small. We want to minimize the  $s$ , because the number of fluctuations (fake jets) is decreasing with smaller  $s$  parameters. Therefore, scale parameters around the value of 30 can be seen as an optimal value (this trend is similar also in  $\eta$  coordinate and also in the  $p_T$  reconstruction). RMS converge to two different values, each for different input jet sample. Also note the inverse  $p_T$  ordering for each sample. It can be due to some phenomena in Pythia jet-finding procedure, because the same MRA-based algorithm was used for both samples. This matter is currently under investigation.

In Fig. 4.12 are shown mean values of differences in the  $\eta$  coordinate ( $\eta_{true-jet} - \eta_{input-jet}$ ). It is in the units of pseudorapidity. There is small systematic error, increasing with decreasing  $s$  parameter and with increasing jet  $p_T$ .

In Fig. 4.13 is compared RMS of differences in the pseudorapidity coordinate for all tested momenta. It is very similar to the RMS Fig. for  $\varphi$  coordinate.

Finally, in Fig. 4.14 are shown mean values RMS values of the  $\frac{\Delta p_T}{p_T}$  defined as:

$$\frac{\Delta p_T}{p_T} = \frac{p_T \text{ true-jet} - p_T \text{ input-jet}}{p_T \text{ input-jet}} \quad (4.10)$$

### 4.8.3 Cleanness and Efficiency

One of observables used for interpretation of jet-finding method is a domain size. Dependence of relative area of a domain (related to the total area of normalized  $\eta \times \varphi$  space) on jet  $p_T$  and scale variable  $s$  is shown in Fig. 4.15 for  $\hat{p}_T = 20$  sample and in Fig.4.16 for  $\hat{p}_T = 100$ . Fake domains are red, true domains (which correspond to input jets) are blue. Domain is recognized as a true jet if the  $p_T$  weighted  $\eta, \varphi$  center of one of the input jets lies in that domain.  $p_T$  dependence is clearly visible - the most of fake domains are small in area and have small  $p_T$ . The signal is purer for higher values of the  $s$  parameter (reverse trend for  $s > 60$  is expected due to decay of domains into individual particles).

Lets define CLEANNESS of the jet signal for scale parameter  $s$  and for jet  $p_T$  interval  $(p_{T1}, p_{T2})$  as

$$C(s, p_{T1}, p_{T2}) = \frac{\int_{x_0}^1 fake(s, p_{T1}, p_{T2}, x) dx}{\int_{x_0}^1 [true(s, p_{T1}, p_{T2}, x) + fake(s, p_{T1}, p_{T2}, x)] dx} \quad (4.11)$$

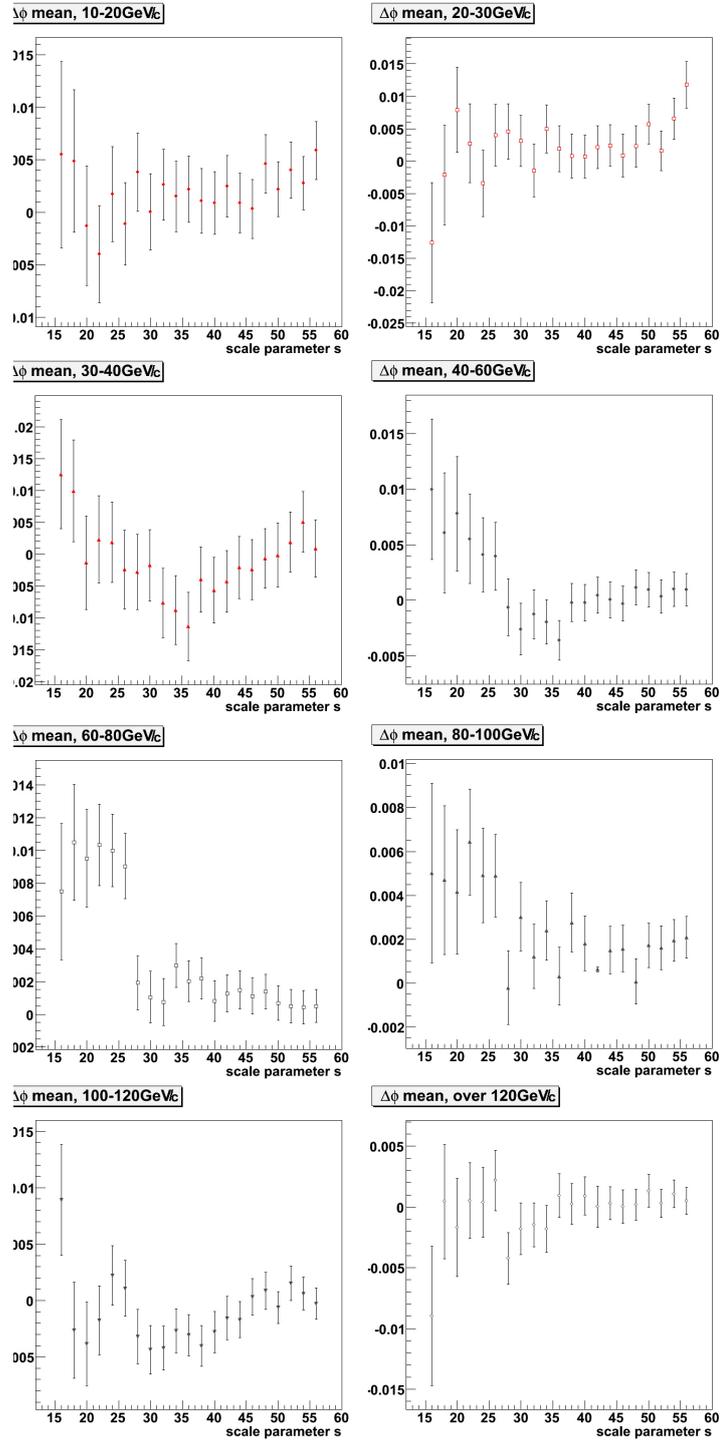


Figure 4.10: Mean value of the difference in the  $\varphi$  coordinate ( $\varphi_{\text{true-jet}} - \varphi_{\text{input-jet}}$ ). Units are radians.

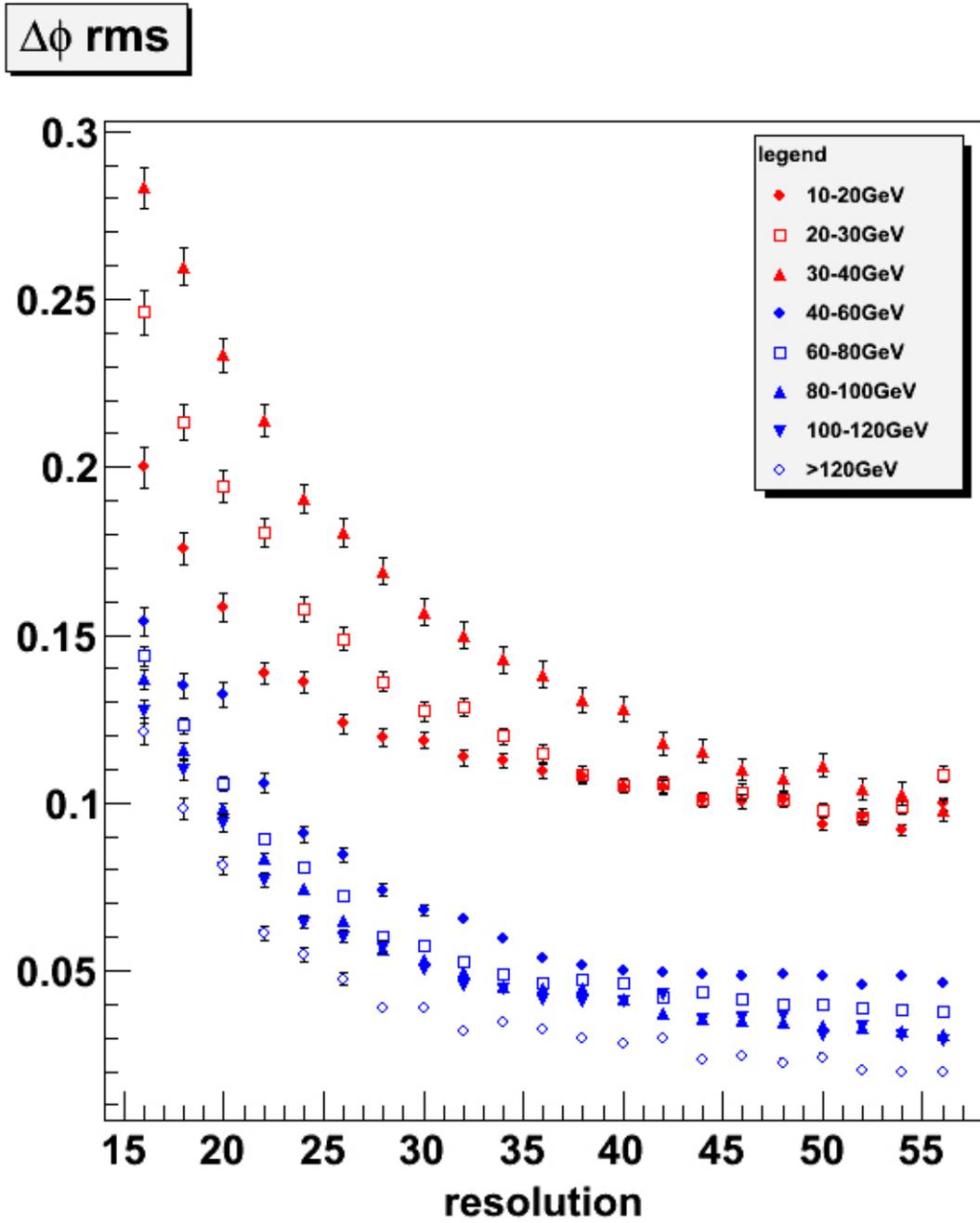


Figure 4.11: RMS for the difference in the  $\varphi$  coordinate ( $\varphi_{true-jet} - \varphi_{input-jet}$ ).

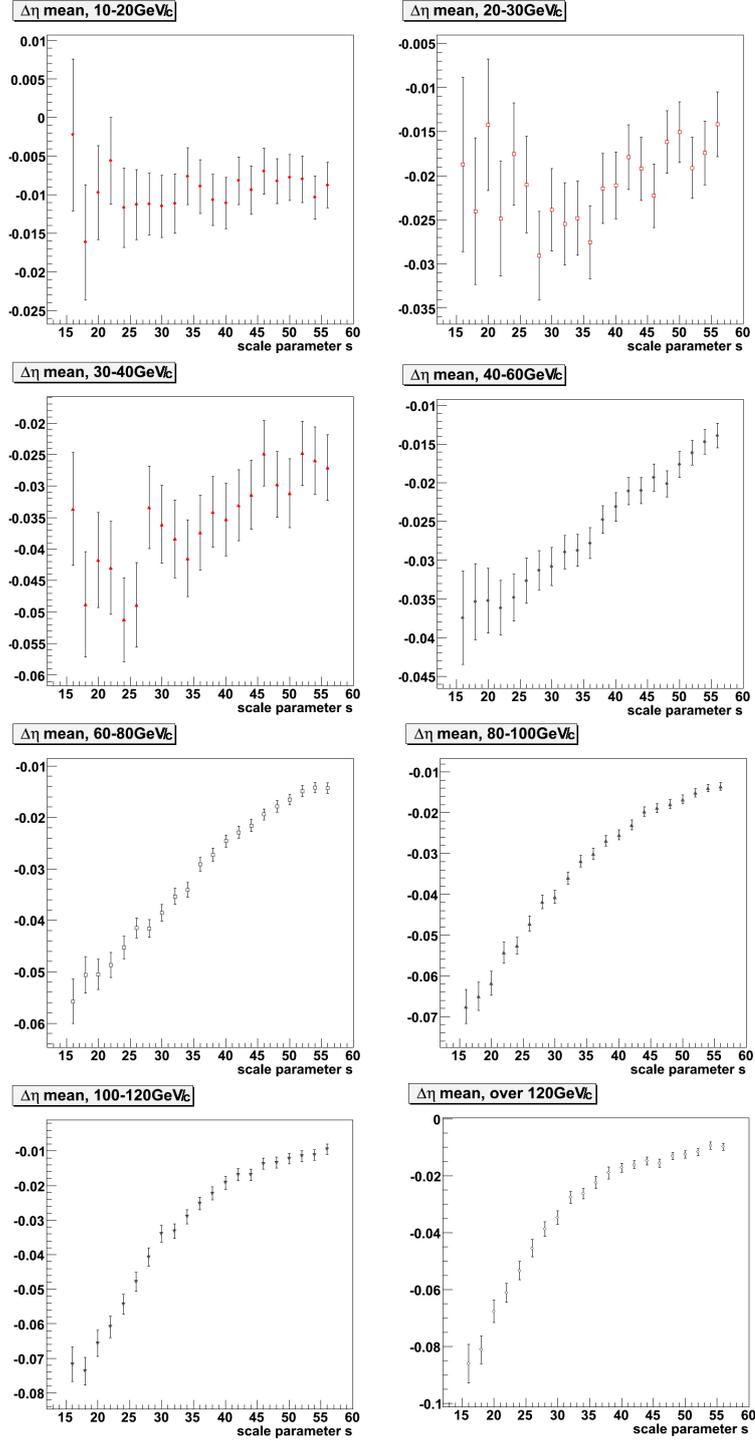


Figure 4.12: Mean value of the difference in the  $\eta$  coordinate ( $\eta_{true-jet} - \eta_{input-jet}$ ) in units of pseudorapidity.

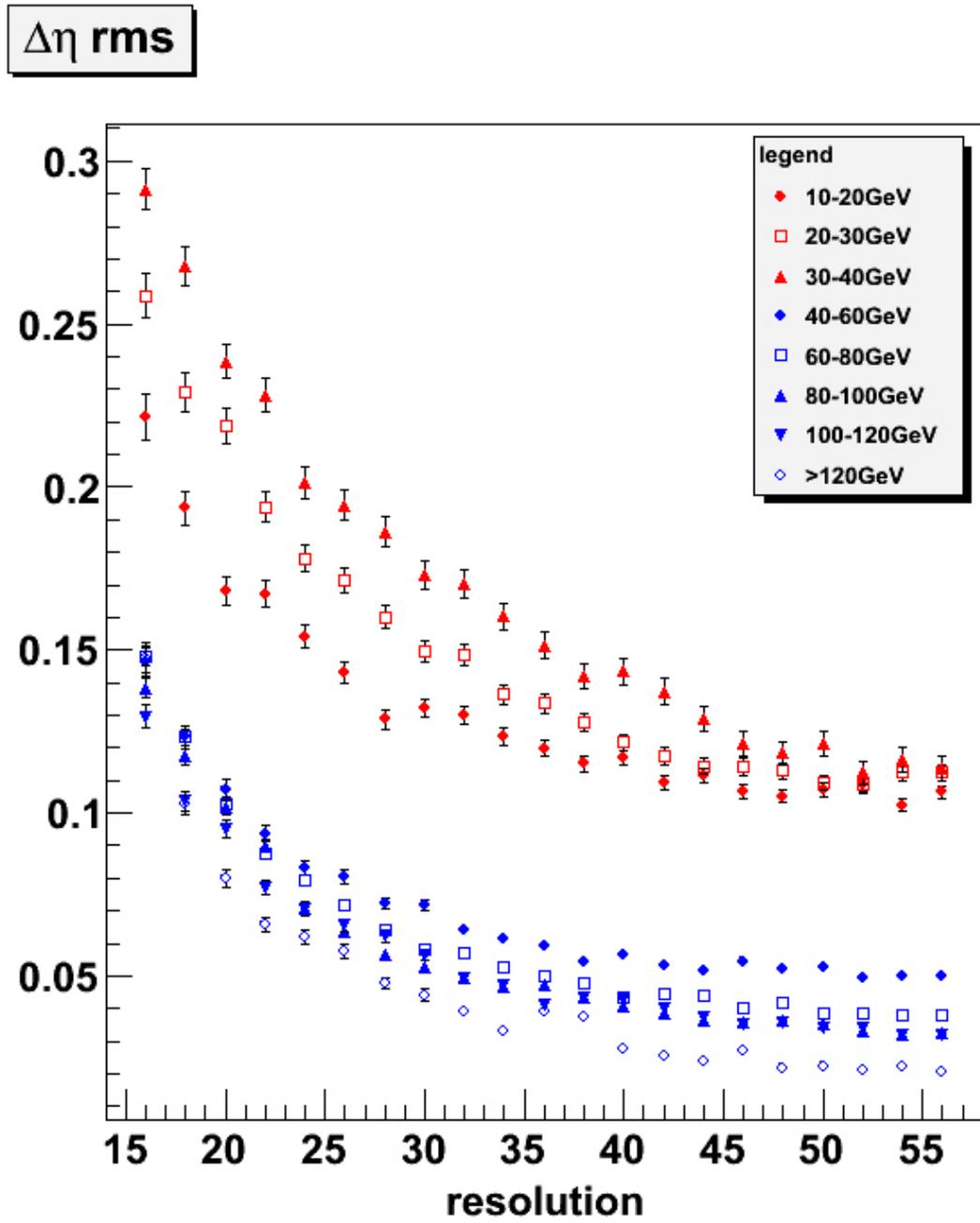


Figure 4.13: RMS for the difference in the  $\eta$  coordinate ( $\varphi_{true-jet} - \varphi_{input-jet}$ ).

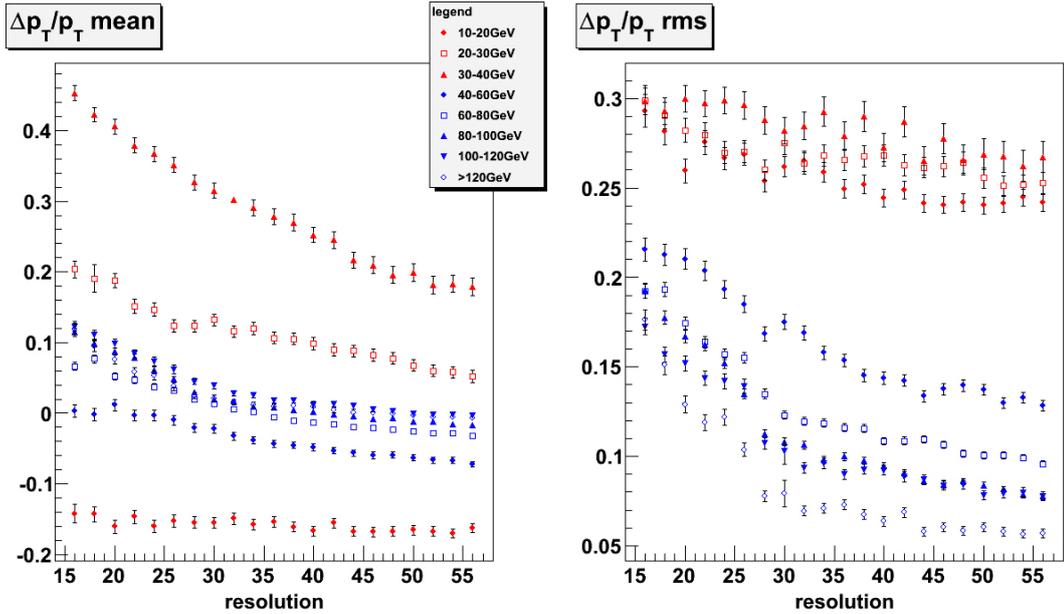


Figure 4.14: Precision of the  $p_T$  reconstruction

and EFFICIENCY of reconstruction as

$$\epsilon(s, p_{T1}, p_{T2}) = 1 - \frac{\int_0^{x_0} \text{true}(s, p_{T1}, p_{T2}) dx}{\int_0^1 \text{true}(s, p_{T1}, p_{T2}) dx} \quad (4.12)$$

, where  $x$  is relative area of a domain (related to the total area of normalized  $\eta \times \varphi$  space).  $x_0$  is the relative area of domain for which the  $p_T$  weighted number of fake domains is equal to the  $p_T$  weighted number of true domains.  $x_0$  has been given from histograms as are in Fig. 4.15 and

Cleanness of the jet signal express the fraction of the fake jets in the 'usable' part of the signal (where is the signal of true jets stronger than background-fluctuation based fake jets). In the best case we obtain  $C = 0$ , which means the completely clean signal with no fake jets. As can be seen in Fig. 4.15, the fake domains have relative small  $p_T$ . Therefore the fraction of fake domains is decreasing rapidly in higher  $p_T$  intervals. Cleanness makes approximately 50% for jets between 10-20GeV/c, 10 – 30% in the interval 20-30GeV/c, only several % between 30-40GeV/c and less then 1% for higher transverse momenta.

jeste sem podle obrazku napsat jak se to meni s rozlisenim

Dependence of cleanness on parameter  $s$  and jet  $p_T$  is shown in Fig. 4.17 Efficiency is equal to 1.0 for all domains with  $p_T$  over 20GeV/c and is summarized in Tab. for interval 10-20GeV/c.

## 4.9 Test results

Tests performed with particular implementation of MRA-based jet-finding algorithm looks promising. Jet-finders currently developing for high-energy heavy ion collisions have good parameters ( $C < 0.1$ ,  $\epsilon \approx 1$ ) for jets with  $E_T$  above 100GeV and have serious difficulties with jets below 60GeV ( $C > 0.5$ ,  $\epsilon < 0.9$ ) [osobniKomunikaceSpousta]. MRA-based

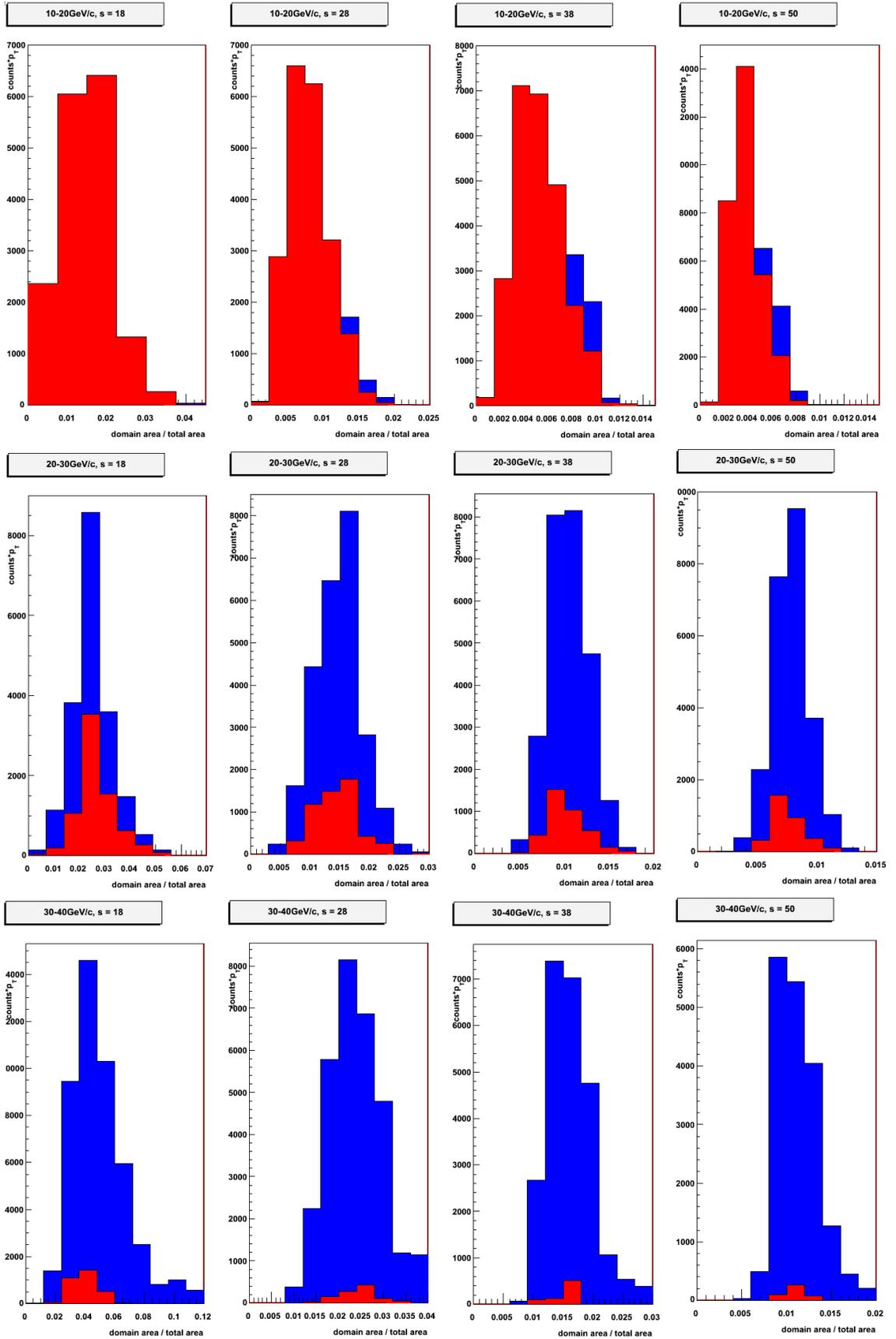


Figure 4.15: Relative area of a domain for  $\hat{p}_T = 20$  sample. Counts are  $p_T$ -weighted. Fake domains are red, true domains (which correspond to input jets) are blue.

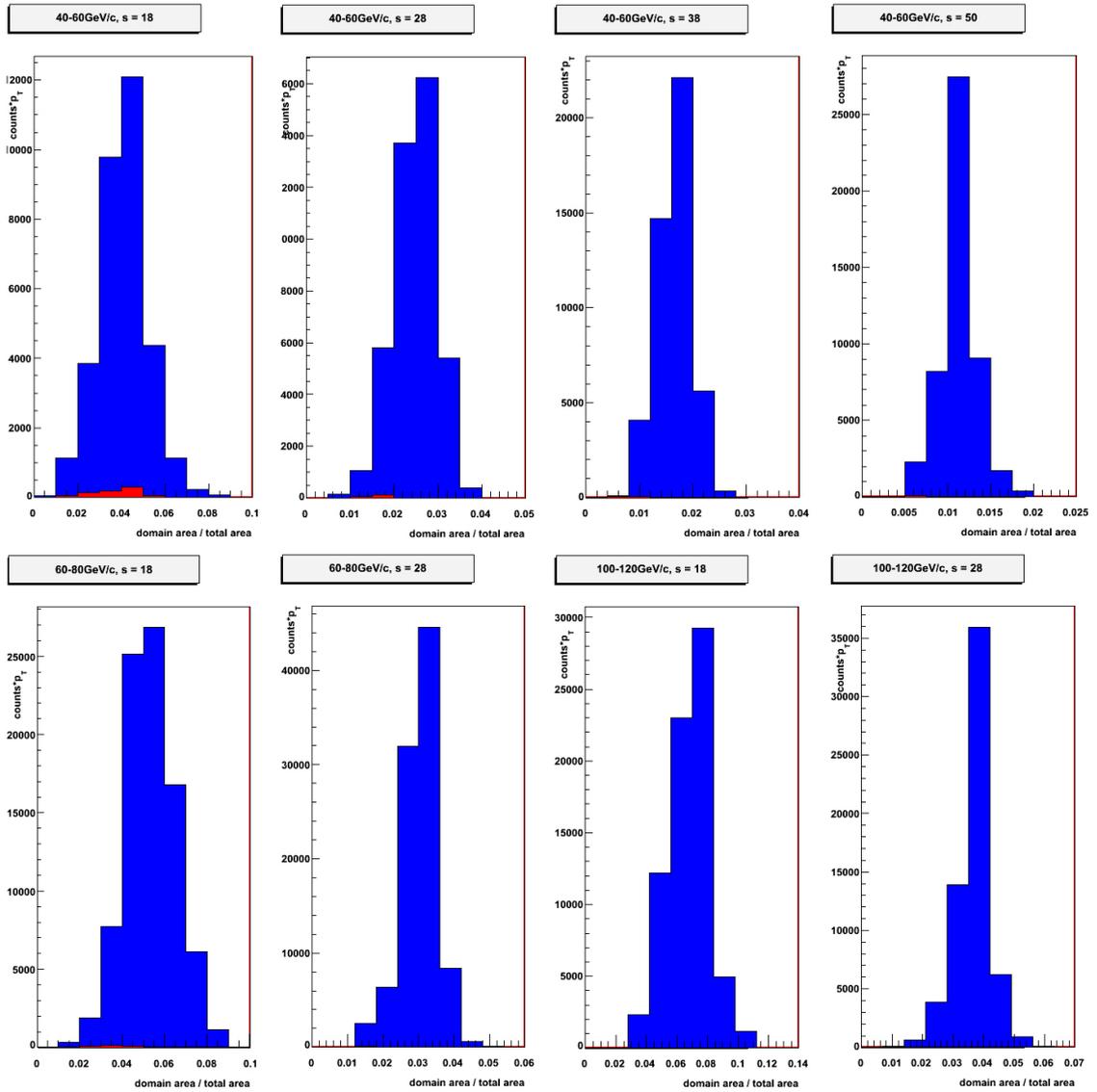


Figure 4.16: Relative area of a domain for  $p_T = 100$  sample. Counts are  $p_T$ -weighted. Fake domains are red, true domains (which correspond to input jets) are blue.

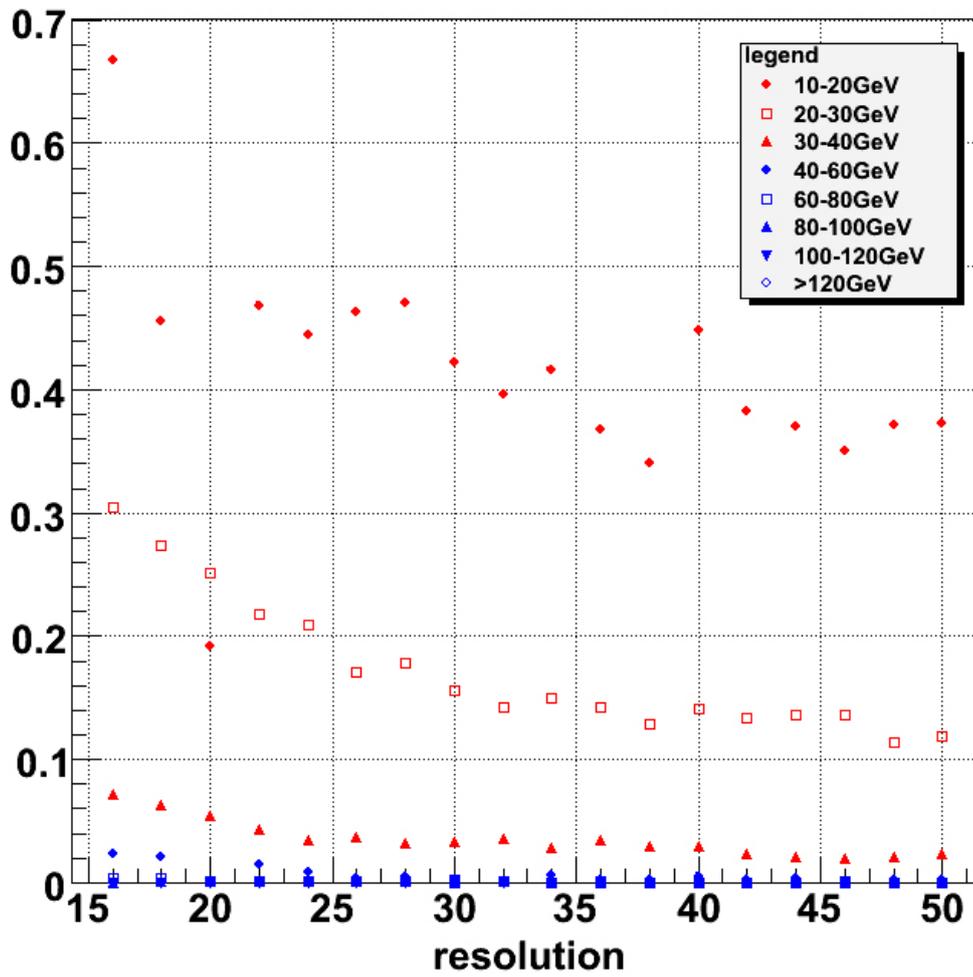
**C**

Figure 4.17: Cleanness of the MRA-based jet finder output

scale parameter $s$	16	18	20	22	24	26	28	30
$\epsilon$ (s,10GeV/c,20GeV/c)	0.935	0.064	0.018	0.201	0.128	0.170	0.350	0.193
scale parameter $s$	32	34	36	38	40	42	44	46
$\epsilon$ (s,10GeV/c,20GeV/c)	0.376	0.288	0.254	0.219	0.617	0.435	0.376	0.335
scale parameter $s$	48	50						
$\epsilon$ (s,10GeV/c,20GeV/c)	0.513	443						

Table 4.1:  $\epsilon$

method implies good je-finding abilities in the region above 40GeV/c (and with limitations even for  $p_T$  above 20GeV/c). However, more testing with more realistic background is recommended. Greater multiplicity of 'soft' background particles may not affect testing results of this MRA-based jet-finding method because of its statistical approach - more particles mean lower particle density relative fluctuations (which is proportional to  $\frac{1}{\sqrt{N}}$ ). New tests (with background  $dN/d\eta_{ch} = 2000$ ) implies very similar behaviour for jets with  $p_T$  above 30GeV/c, however, only small sample has been tested up to the present day.

Testing with minijets admixed to the background may worsen this results, it increases background particle density fluctuations and results in a harder  $p_T$  spectrum.

Examination of usefulness of jet-finding method based on multiscale approach was the main goal of performed tests. Another task was to explore dependence of the reconstruction precision on the scale parameter  $s$  and to determine the optimal value of this parameter. Preliminary tests done with scale parameters  $s = 10, 20, 40$  and  $80$  implied optimal scale parameter greater then  $20$  (effect of domain merging appeared at this value) and smaller then  $80$  (particle nature of the jet caused decay of the jet into many domains at this value of  $s$  parameter). More detailed examination has shown applicable region of scale parameter between  $30$  and  $56$ . Improvement of the  $\eta, \varphi, p_T$  reconstruction efficiency ends for the latter value of the scale parameter and worsening of the reconstruction precision is appears for greater scaling parameters, due to decay of jet domains into clusters and individual particles. This usable  $s$  parameter region can background dependent, therefore additional tests need to be done before implementation of the algorithm for particular experiment.

# Conclusions

The main aim of this work is to develop the MRA-based algorithm for jet detection in ultra-relativistic nuclear collisions and to test it on simulated data. Jet algorithm and recombination scheme have been developed and tested on two di-jet samples generated in Pythia 8.100 with random background. The Splitting/Merging rule is needed to be developed before the test of double jet resolution. Also the selection criterion for exclusion of fake jets has to be developed.

MRA-based jet-finding algorithm is, due to its nature, sufficiently insensitive to soft radiation and therefore it is infrared stable. It is also collinear stable, because it needs no seeds as starting points. Two particles close to each other are treated in similar way as one particle with higher  $p_T$  (this is controlled by scale parameter). Developed algorithm contains only one computing cycle through the input particles and therefore its speed can be compared with the fastest jet-finding algorithms.

Although the tests were performed with lower background than is expected at, for example, LHC energies, it may does not matter, as it was discussed in chapter... Test results are promising, reconstruction precision is comparable with other currently developing algorithms for high-energy heavy ion collisions.

Developed MRA-based jet-finding method can be further improved. It is possible to create more than one  $L$  function with different parameters  $s$  per event and to perform jet-finding for relevant  $F$  function. More strict criterion for merging relevant amplitude array cells into domains can be applicated (for example to merge only neighbouring cells with common edge, not with common corner). Also the threshold evaluation can be modified. Even if the MRA-based jet algorithm will have worse results in comparison with other jet-finding algorithms, it can be used, due to its speed, as a preclustering step for other methods.

Many of developed jet finding algorithms ended without any use. However, I hope development of this method will continue and it will be used someday for analysis of real experimental data.

# Bibliography

- [1] V. Petráček, Habilitační práce, České vysoké učení technické v Praze, Praha 2006.
- [2] A. Akansu, M. Smith, Subband and Wavelet Transforms, Kluwer Academic Publishers, USA 1996.
- [3] A. Rosenfeld et al., Multiresolution Image Processing and Analysis, Springer-Verlag, Berlín 1984.
- [4] J. Starck, F. Murtagh, A. Bijaoui, Image Processing and Data Analysis, Cambridge University Press, Cambridge 1998.
- [5] P. Brémaud, Mathematical principles of signal processing: fourier and wavelet analysis, Springer, New York 2002.
- [6] L. Debnath, Wavelet transforms and time-frequency signal analysis, Birkhäuser, Boston 2001.
- [7] R. Rao, A. Bopardikar, Wavelet Transforms: Introduction to theory and applications, Addison-Wesley, Reading 1998.
- [8] S. Mallat, A wavelet tour of signal processing, Academic Press, San Diego 1999.
- [9] <http://documents.wolfram.com/applications/wavelet/FundamentalsofWavelets/1.2.1.1>
- [10] R. Field, Underlying Event in Hard Scattering Processes, CDF/MIN-BIAS/PUBLIC/5746.
- [11] S. Ellis, Collider jets in perturbation theory, CERN-TH.6861/93.
- [12] The ALEPH Collaboration, Quark and Gluon Jet Properties in Symmetric Three-Jet Events, CERN-PPE/95-184.
- [13] B. Mohanty, J. Serreau, Disoriented Chiral Condensate: Theory and Experiment, arXiv:hep-ph/0504154v2.
- [14] C. Aguiar, E. Fraga, T. Kodama, Spinodal instability in the quark-gluon plasma, Braz. J. Phys. vol.34, São Paulo 2004.
- [15] A. Kaidalov, Regge poles in QCD, arXiv:hep-ph/0103011v1.
- [16] J. Huth et al., Proceedings 1990 Summer Study on High Energy Physics, ed. E. Berger. Singapore: World Scientific, 134 (1992).

- [17] G. Blazey et al., Run II Jet Physics, arXiv:hep-ph/0005012v2.
- [18] <http://mathworld.wolfram.com/VoronoiDiagram.html>
- [19] <http://mathworld.wolfram.com/DelaunayTriangulation.html>
- [20] N. Varelas, Lectures on Jet Physics, CTEQ Summer School 2000.
- [21] B. Zakharov, Parton Energy Loss in an Expanding Quark-Gluon Plasma: Radiative vs. Collisional, JETP Letters, 2007, Vol.86, No.7, pp.466-450, ISSN 0021-3640.
- [22] I. Arsene et al., BRAHMS collaboration, Phys. Rev. Lett **91**, 072305 (2003).
- [23] J. Cronin et al., Phys. Rev. **D11** 3105 (1975).
- [24] A. Accardi, arXiv:hep-ph/0212148v2.
- [25] J. Bjorken, Report No. Fermilab-Pub-82/59-THY (1982).
- [26] C. Adler et al., STAR collaboration, Phys. Rev. Lett. **90**, 082302 (2003).
- [27] J. Adams et al., STAR collaboration, Phys. Rev. Lett **91**, 072304 (2003).
- [28] J. Adams et al., STAR collaboration, Phys. Rev. Lett **93**, 252301 (2004).
- [29] B. Allesandro et al., ALICE collaboration, Physics Performance Report, Volume II, J. Phys. G: Nucl. Part. Phys. **32** 1882-1934.
- [30] B. Abbott et al., DØ collaboration, Fermilab-Pub-97-242-E (1997).
- [31] M. Cacciari, G. Salam, Dispelling the  $N^3$  myth for the  $k_t$  jef-finder, Phys. Lett. B 641 (2006) 57-61.
- [32] <http://mathworld.wolfram.com/CauchyDistribution.html>
- [33] Pythia 6.4 manual, hep-hp/0603175
- [34] Pythia 8.1 manual, <http://home.thep.lu.se/~torbjorn/pythiaaux/present.html>
- [35] ROOT User's Guide 5.16, <http://root.cern.ch>