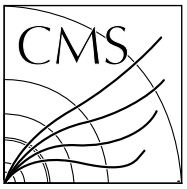
Available on CMS information server

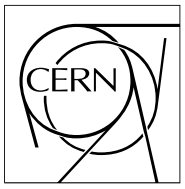
CMS NOTE 2006/050



The Compact Muon Solenoid Experiment

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland



An algorithm for jet identification and reconstruction in a densly populated calorimetric system

O. Kodolova

Moscow State University

I. Vardanyan

Moscow State University

A. Nikitenko[a)]

Imperial College, London, UK

A. Oulianov

ITEP, Moscow

Abstract

A jet reconstruction algorithm is developed for event with a high particle density in the calorimetric system. The performance of the reconstruction of hard QCD jets with initial parton energies 50-300 GeV is studied in central Pb–Pb collisions with a modified cone jet finder which includes an algorithm for event-by-event background subtraction. The heavy ion background is simulated using the HIJING Monte-Carlo generator with $dN_{\rm ch}/dy|_{y=0} = 5000$. Results on the achieved jet reconstruction efficiency, purity, energy and spatial resolution are presented.

[a)] on leave Institute of Experimental and Theoretical Physics, Russia

1 Introduction

In future heavy ion collisions experiments at the LHC, the medium-induced energy losses of a hard partonic jet propagating through the dense matter can provide information on the quark-gluon plasma. In order to observe these losses in the processes $\mathrm{jet} + \mathrm{jet}$ [1, 2], $\gamma + \mathrm{jet}$ [3], $Z + \mathrm{jet}$ [4] and by measuring the leading particle momenta in a jet [5, 6] one should also measure the energy and spatial position of jets with good resolution. The main challenge in the jet recognition in heavy ion collisions is to distinguish them from the "false" jet background arising from fluctuations of the transverse energy flow from the large multiplicity of secondary particles in the event [7]. Predictions vary from 2000 to 8000 charged particles per unit of rapidity in central Pb–Pb collisions at the LHC. Under these circumstances, the reconstruction of "true" QCD jets resulting from hard parton-parton scattering is very important. The definition of an object like a jet is difficult in pp-collisions. In particular, the energy and spatial resolution of jets are sensitive to the parameters of the jet finding algorithm [8].

An original jet finding algorithm was developed for reconstructing hard jets (E_T of order 100 GeV) in heavy ion collisions [9, 10, 11, 12], where particles produced in a typical collision deposit a transverse energy E_T of up to 10 GeV in every calorimeter tower. The algorithm allows subtraction of the background energy due to the underlying event energy flow and identification of the hard jets on an event by event basis. In this study, the jet finding is done with the background subtraction algorithm currently implemented in the CMS reconstruction software.

2 Detector description

A characteristic feature of the CMS detector is its large superconducting solenoid delivering an axial magnetic field of 4 T. The hadron and electromagnetic calorimeters are located inside the coil and cover the pseudorapidity range $|\eta| < 3$ [13, 14]. The forward calorimeter is situated outside the coil and cover the pseudorapidity range $3 < |\eta| < 5$ [14].

The following segmentation of calorimeter towers, including both electromagnetic (ECAL) and hadronic (HCAL) parts, is used. In the barrel and most of the endcap part of the HCAL, the size of the calorimeter towers is $\Delta\eta = 0.0870$ by $\Delta\varphi = 2\pi/72 \simeq 0.0873$. At high η in the HCAL endcap ($1.74 < \eta < 3.0$), the towers are larger in η and double size in φ.

The detailed description of the CMS detector is used in simulation [15].

3 Background subtraction algorithm

A modified jet finding algorithm has been developed to search for "jet-like" clusters above the average transverse energy flow and to subtract this background due to the underlying event.

The algorithm is the flavour of the "Noise/Pedestal subtraction". The subtraction is the iterative procedure. Firstly, it is considered, that there is not known whether there are jets or not and the mean value of energy and mean dispersion in cell are calculated for all η-rings, i.e. the Pedestal as a function of η (P(η)) is determined. Than the value of Pedestal function (P(η)) is subtracted from the all cells and jets are found from the remaining non-empty cells. After jets are found, the value of the Pedestal function is recalculated using cells outside jets area. The parameter of algorithm is the cut on the E_T^{jets}, found after Pedestal subtraction. It is 30 GeV for Pb–Pb event and 10 GeV for high-lumi pile-up.

The algorithm is described step-by-step here.

- As a first step, the average tower transverse energy $\overline{E_T^{\mathrm{tower}}(\eta)}$ and its dispersion $\sigma_T^{\mathrm{tower}}(\eta)$ are calculated event by event at each η ring over all towers in the barrel and endcap calorimeters. The dispersion is defined as $\sigma_T^{\mathrm{tower}}(\eta) = \sqrt{\overline{(E_T^{\mathrm{tower}}(\eta))^2} - (\overline{E_T^{\mathrm{tower}}(\eta)})^2}$.
- In a next step, all tower energies are recalculated as $E_T^{\mathrm{tower}*} = E_T^{\mathrm{tower}} - \overline{E_T^{\mathrm{tower}}(\eta)} - k \cdot \sigma_T^{\mathrm{tower}}(\eta)$, where E_T^{tower} is the original transverse energy in the tower. If the value of the transverse tower energy after subtraction becomes negative, it is set to zero.
- Using the corrected tower energy, jets are found with an iterative cone algorithm as described in [16] .
- Then the average tower energies and dispersions are recalculated again using only towers outside of the jets, the original tower energies are used in this calculation.

- The tower energies are recalculated again as $E_T^{\mathrm{tower}*} = E_T^{\mathrm{tower}} - \overline{E_T^{\mathrm{tower}}(\eta)} - k \cdot \sigma_T^{\mathrm{tower}}(\eta)$. If the tower energy after subtraction becomes negative, it is set to zero.
- Then the iterative cone algorithm is used again to reconstruct the final jets from the towers with the new energies.

The different methods of background subtraction can be considered to reduce the effects of background events on the jet reconstruction. In all these methods, the average tower energy offset due to background (pileup, noise) contribution is evaluated event by event as a function of the tower pseudorapidity. This offset is then subtracted from the tower energies. Methods of background subtraction differ in the way the tower energy offsets are calculated and subtracted.

- *Method 1*. The offset is calculated as the average tower energy in the ETA ring and then subtracted from the tower energies. All the towers (including those with negative energies) are then used to reconstruct jets.
- *Method 2*. The RMS of the tower energies in each tower ring is also calculated. The offset is calculated and subtracted as in "Method 1", but a tower is suppressed if its energy after subtraction falls below two RMS.
- *Method 3*. The tower energy offset is calculated as

$$\mathrm{Offset} = \mathrm{Mean} + k \cdot \mathrm{RMS}$$

of the energies of the towers in the ring. A tower is suppressed if it has negative energy after sutraction. The factor k is meant to compensate the positive bias in the jet reconstructed energy due to suppression of towers with negative energies. The value $k = 1$ is chosen so as to eliminate the dependence of the reconstructed jet energy and energy resolution on event multiplicity in Pb–Pb collisions and luminosity in pp interaction. Another words, $k = 1$ is more optimal for jet reconstruction in events with large background.

The method (3) with a factor $k = 1$ is used in this study to compensate the positive bias in the reconstructed jet energy due to the suppression of towers with negative energy. This scheme gives an approximation for the reconstructed jet energy in Pb–Pb which is close to the jet energy obtained in events without background.

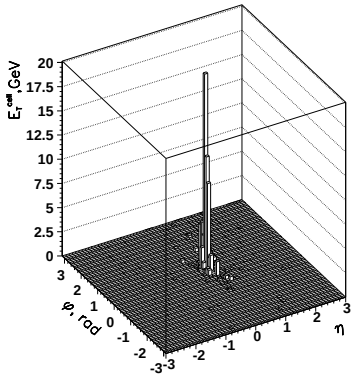


Figure 1: A 100 GeV jet in event without background after the background subtraction procedure.

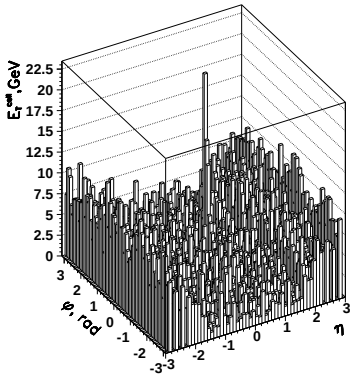


Figure 2: A 100 GeV jet is superimposed on a Pb–Pb event.

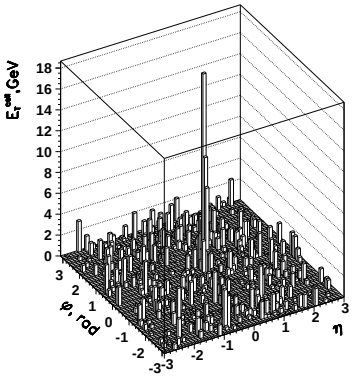


Figure 3: A 100 GeV jet superimposed on a Pb–Pb event after background subtraction procedure.

Figures 1 - 5 illustrate the jet reconstruction with the background subtraction algorithm. First, a 100 GeV jet is reconstructed in the event without background (Fig. 1, 4). The same jet is then superimposed on a Pb–Pb event (Fig. 2) and reconstructed from an event with background (Fig. 3, 5).

Figs. 4 and 5 are essentially the same as Figs. 1 and 3 but for clarity only towers inside the cone of reconstructed jet are shown. In this example, the jet is reconstructed in a cone with $R = 0.5$. There are ~ 100 calorimeter towers inside the 0.5 cone. After the background subtraction procedure the number of towers in the jet with positive energy is equal ~ 30.

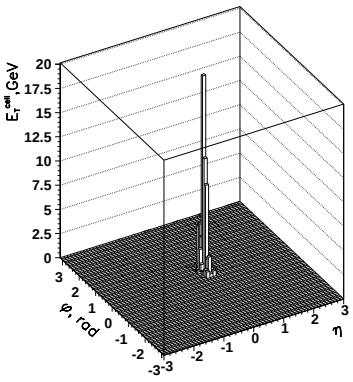


Figure 4: The same as in Fig. 1 with only the towers inside the cone of reconstructed jet

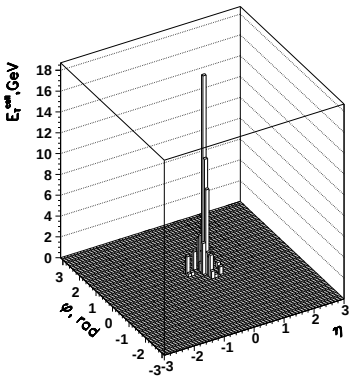


Figure 5: The same as in Fig. 3 with only the towers inside the cone of reconstructed jet.

4 Event generation and simulation

Samples of signal QCD dijet events in different intervals of the initial parton transverse monentum $\hat{p}_T$: 50-60 GeV/c, 70-80 GeV/c, 90-105 GeV/c, 120-130 GeV/c, 200-210 GeV/c, 300-310 GeV/c are generated as pp collisions with PYTHIA 6.158 [17]. The highest E_T jet is found first at the generator level as a cluster of generated particles (stable particles, including muons and neutrinos) from the hard $\hat{p}_T$ interaction inside a cone of $R = 0.5$ around the leading particle. In the following, this jet is referred to as a particle jet. A single particle jet is taken and the particles belonging to this jet are passed through detector simulation. Other particles in the QCD event are ignored.

Central Pb–Pb events at LHC energy $\sqrt{s_{NN}} = 5.5$ TeV are simulated using the default setting (quenching on) of HIJING Monte-Carlo generator [18] for nucleus-nucleus interactions with $dN_{\rm ch}/dy|_{y=0} = 5000$. The signal event sample (QCD jets) and background event sample (central Pb–Pb, HIJING) are processed separately event by event through the CMS detector (CMSIM125 [15]). Then each simulated signal event is superimposed on a background event to form a heavy ion event and digitized with CMS reconstruction software (ORCA6 [19]).

5 Pure background heavy ion events

At first, pure background heavy ion events are studied, i.e. only HIJING events without hard collision signal events added to the event.

The energy flow, $< E_T^{\rm tower}(\eta) >$, is defined as the transverse energy per tower, where the transverse energy is averaged over all towers around the whole azimuth at a given $\eta^{\rm tower}$. The energy flow as a function of $\eta^{\rm tower}$ is shown in Fig. 6 for ECAL and HCAL calorimeters separately. Most of the energy is reconstructed in the ECAL. The transverse energy flow shows a strong η dependence in the endcap region ($|\eta| > 1.6$), but not so much in the barrel. Fig. 7 shows the same as Fig. 6, but the average transverse energy (summed in φ) is normalised to one unit of pseudorapidity. The amount of fluctuations in the background $< \sigma_T^{\rm tower}(\eta) >= \sqrt{< D_T^{\rm tower}(\eta) >}$ in the calorimeter towers (ECAL + HCAL) is shown in Fig. 8 where $D_T^{\rm tower}(\eta)$ is defined as the variance of transverse energy per tower $D_T^{\rm tower}(\eta) = \overline{(E_T^{\rm tower}(\eta))^2} - (\overline{E_T^{\rm tower}(\eta)})^2$ calculated over all towers around the whole azimuth at a given $\eta^{\rm tower}$. The average values of E_T per tower (ECAL + HCAL) for central Pb–Pb collisions are: 1.7 ± 0.9 GeV in the barrel and 4.8 ± 5.0 GeV in the endcap.

HIJING generator includes soft and hard part of heavy ion event both. For hard part it includes multiple mini-jets production (many $2 \to 2$ semi-hard scatterings). Thus HIJING takes into account mini-jets background for high-pt jets. However there is no overlap between high-pt PYTHIA jet pair and HIJING multi-jets due to different scale of Pt (and therefore different scale of corresponding cross sections). The probability to have hard scattering in HIJING with, say $\hat{p}_T$=70 GeV or 100 GeV, is much less than 1.

All jets in pure background events, including real mini-jets from HIJING are considered as fake jets in comparison with hard jets from signal events. The dependence of the transverse energy $E_T^{\rm jet}$ vs $\eta^{\rm jet}$ of fake jets extracted from pure heavy ion events without background subtraction is shown on two-dimensional plot in Fig. 9. For the main part of fake jets with $E_T^{\rm jet} \sim 300$ GeV, pseudorapidity $|\eta|^{\rm jet}$ is ~ 3. The correlation between transverse

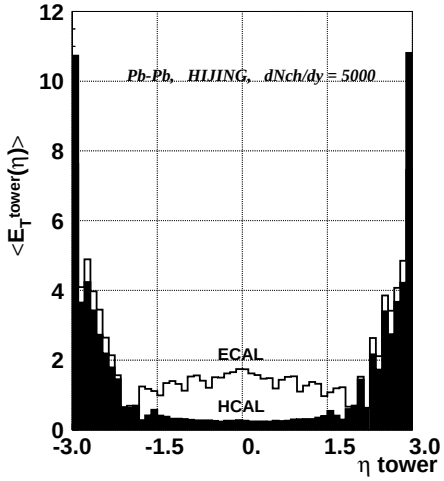


Figure 6: Dependence of transverse energy flow $< E_T^{tower}(\eta) >$, averaged over φ, on pseudorapidity $\eta^{\rm tower}$. The open histogram shows the reconstructed energy in ECAL, the closed histogram shows the reconstructed energy in HCAL.

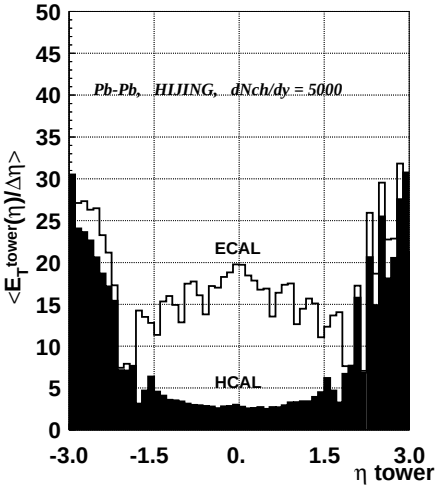


Figure 7: The same as in Fig. 6, but the transverse energy is normalised to one unit of pseudorapidity.

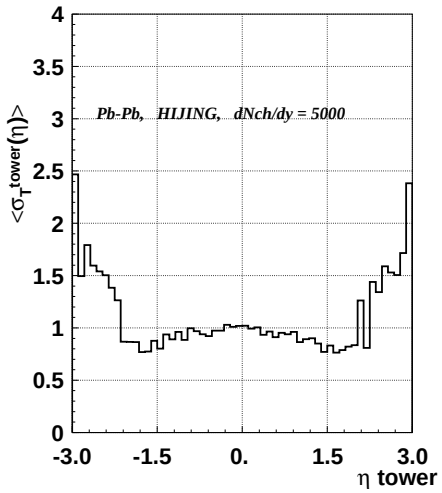


Figure 8: Dependence on pseudorapidity $\eta^{\rm tower}$ of the dispersion $< \sigma_T^{\rm tower}(\eta) >$ of the transverse energy in towers (ECAL + HCAL).

energy $E_T^{\rm jet}$ and pseudorapidity $\eta^{\rm jet}$ of fake jets in Fig. 9 can be explained by the procedure of jet finding with the iterative cone algorithm used here. The endcap towers at high pseudorapidity have the largest transverse size in the calorimeter and due to the large energy flow in this region, these towers tend to have the largest transverse energy. Thus the iterative cone jet finding starts from these towers as seeds, and jets are first reconstructed in this region. Because of larger transverse energy flow in this region these jets collect higher transverse energy than jets in the other parts of the calorimeter. This explanation of the pseudorapidity dependence of the transverse energy of fake jets in Fig. 9 is supported by the pseudorapidity distribution of fake jets in Fig. 10 extracted from pure heavy ion events without background subtraction, where most of fake jets are in the endcap region. After the subtraction procedure for pure background events, most of the fake jets have energy less than 30 GeV and the pseudorapidity distribution of jets from HIJING becomes flat (Fig. 11).

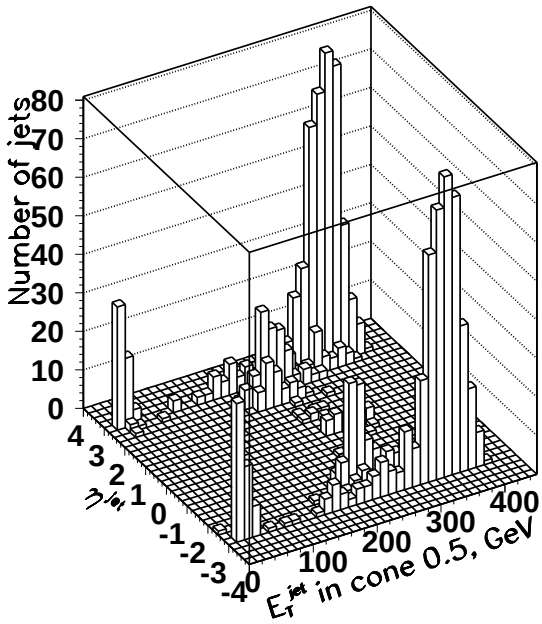


Figure 9: The correlation between transverse energy $E_T^{\rm jet}$ and pseudorapidity $\eta^{\rm jet}$ of fake jets extracted from pure heavy ion events without background subtraction.

6 Performance of the jet reconstruction in heavy ion events

Jet reconstruction is studied using the cone background subtraction algorithm in events without and with superimposed background from central Pb–Pb collisions. Jets are reconstructed in cone of radius $R = 0.5$. The treshold on tower seed is 1 GeV. The threshold on tower energy is 0.5 GeV. The threshold on reconstructed jet energy is chosen 30 GeV. Only one jet per event with the largest transverse energy is used for futher analysis.

6.1 Jet energy scale

The correlation between the reconstructed $E_T^{\rm reco}$ and the generated $E_T^{\rm MC}$ transverse energy of jets in a cone of $R = 0.5$ in Pb–Pb events and in jet events without background is shown in Fig. 12 (barrel) and Fig. 13 (endcap). The points in this plots are the mean energy of the reconstructed jets, and the error bars are the dispersion of jet energy distribution. On average, the measured jet energy in Pb–Pb collisions is the same as that in jet events without background. In other words, the background subtraction algorithm gives a reconstructed jet energy which is approximate the same in the events without background and with background (Pb–Pb). The mean values of the reconstructed jet energy in events with background are slightly lower then the points in events without background at $E_T^{\rm MC} > 200$ GeV and higher at $E_T^{\rm MC} < 100$ GeV. The values are within error bars, but there is an indication of the slightly different slopes of the $E_T^{\rm reco}$ vs $E_T^{\rm MC}$ in events with and without background.

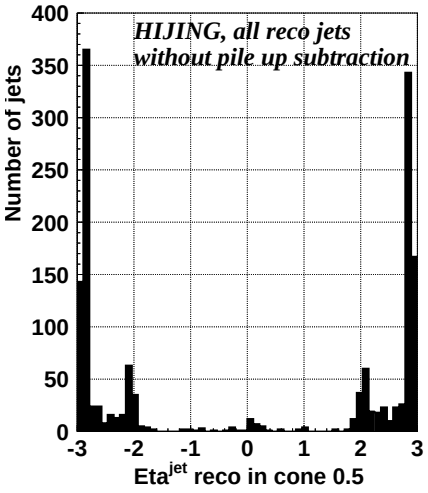


Figure 10: The pseudorapidity distribution of the fake jets extracted from pure heavy ion events (central Pb–Pb) in the region $|\eta| \leq 3.0$ without background subtraction.

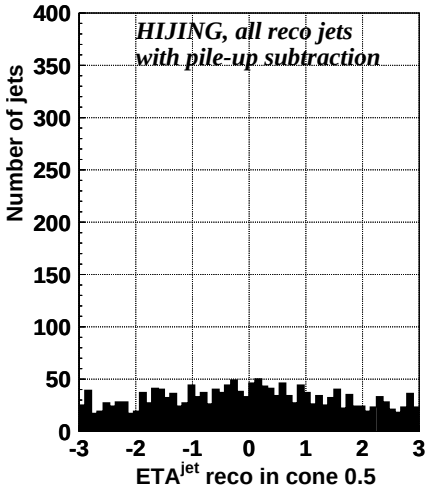


Figure 11: The pseudorapidity distribution of the fake jets extracted from pure heavy ion events (central Pb–Pb) in the region $|\eta| \leq 3.0$ with background subtraction.

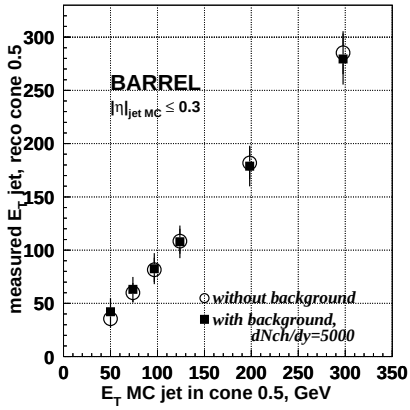


Figure 12: The correlation between the reconstructed and the generated jet transverse energies in Pb–Pb (full squares) and in jet events without background (open circles) in the barrel ($|\eta| < 0.3$).

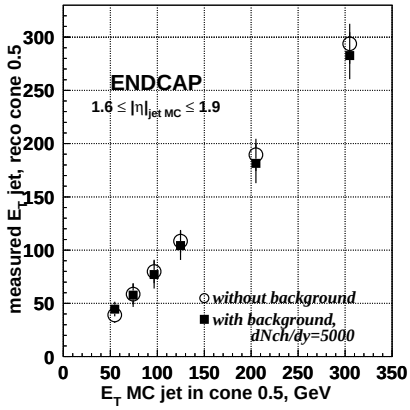


Figure 13: The correlation between the reconstructed and the generated jet transverse energies in Pb–Pb (full squares) and in jet events without background (open circles) in the endcap ($1.6 < |\eta| < 1.9$).

6.2 Jet energy resolution

The jet energy resolution is defined as $\sigma(E_T^{\rm reco}/E_T^{\rm MC})/\langle E_T^{\rm reco}/E_T^{\rm MC}\rangle$, where $E_T^{\rm reco}$ is the reconstructed transverse energy, and $E_T^{\rm MC}$ is the transverse energy of all generated particles inside the given cone of radius $R = 0.5$. For jets above E_T = 75 GeV, the jet energy resolution is degraded by a factor ~ 1.3 in high multiplicity central Pb–Pb collisions when compared to jets sample without background, as shown in Fig. 14 (barrel) and Fig. 15 (endcap).

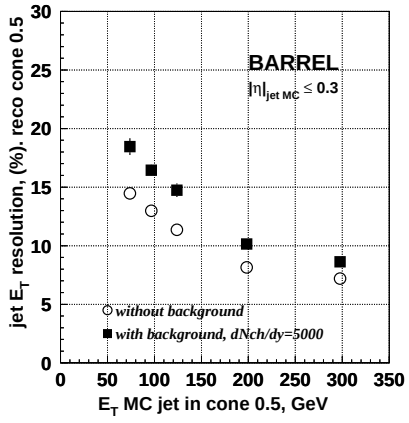


Figure 14: The jet energy resolution in Pb–Pb (full squares) and in jet events without background (open circles) in the barrel region of CMS ($|\eta| < 0.3$).

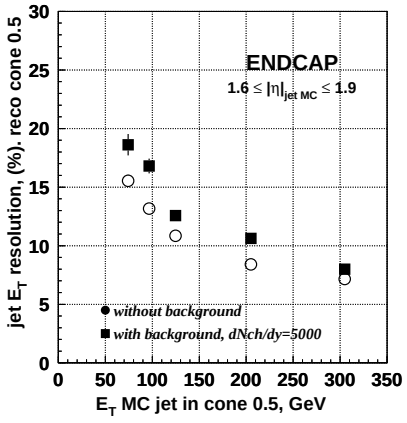


Figure 15: The jet energy resolution in Pb–Pb (full squares) and in jet events without background (open circles) in the endcap region of CMS ($1.6 < |\eta| < 1.9$).

6.3 Jet spatial resolution

Since the azimuthal angle and the rapidity distributions of jets is of particular interest for jet quenching observables in the heavy ion collisions, spatial resolution is important. Figures 16 and 17 show the differences in pseudorapidity $\Delta\eta$ and azimuthal angle $\Delta\varphi$ between generated and reconstructed jets in events without and with Pb–Pb background for different generated jet energies in the barrel. For 100 GeV jets the η and φ resolutions are 0.028 and 0.032 correspondingly. In the endcap region, the η and φ resolution of jets is slightly better than in the barrel. The spatial resolution of jets is degraded in central Pb–Pb collisions in comparison with jet sample without background, but is still better than the η, φ size of one calorimeter tower (0.087×0.087). Thus the spatial position of a hard jet can be reconstructed in heavy ion collisions in CMS with high enough accuracy for the analysis of jet production as a function of azimuthal angle and pseudorapidity.

6.4 Efficiency and purity of jet reconstruction

Figures 18 and 19 show the efficiency and the purity of calorimetric jet reconstruction in central Pb–Pb collisions in the barrel and the endcap regions as a function of MC jet energy. Jets reconstructed in central Pb–Pb collisions with $E_T > 30$ GeV, that are within $\Delta R < 0.25$ around the direction of the generated MC jet are considered as true QCD jets. The efficiency of jet reconstruction in central Pb–Pb collisions is estimated as the fraction of events with such true QCD jets among all the generated events. The efficiency of finding a true QCD jet is $\sim 100\%$ already for $E_T = 75(100)$ GeV jets in the barrel (endcap) region.

The purity of the reconstructed jet sample is defined as the number of events with a true QCD jets divided by the number of events with at least one reconstructed jet (fake or real) with transverse energy above 30 GeV. Above E_T = 50 GeV (75 GeV) jets, the purity is $\sim 100\%$ for the barrel (endcap). Good purity is expected even for jets with $E_T > 30$ GeV, but with possibly lower efficiency. Thus for 50 GeV jets the efficiency is $\sim 61\%$, the purity is $\sim 98\%$ (barrel), since the threshold on jet E_T is 30 GeV and in pure background events the energy of the fake jets is less than 30 GeV.

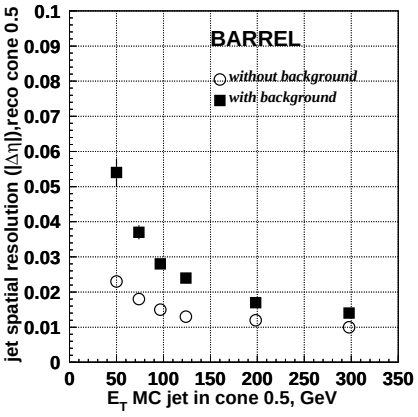


Figure 16: Energy dependence of jet pseudorapidity resolution in Pb–Pb (full squares) and in jet events without background (open circles) in the barrel ($|\eta| < 0.3$).

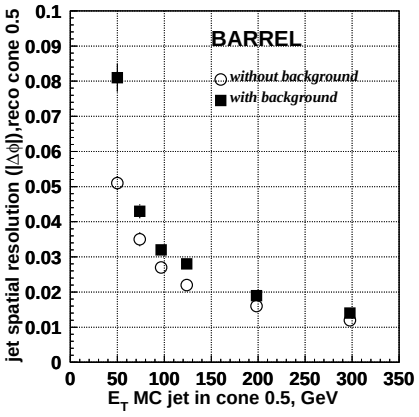


Figure 17: Energy dependence of jet azimuthal angle resolution in Pb–Pb (full squares) and in jet events without background (open circles) in the barrel ($|\eta| < 0.3$).

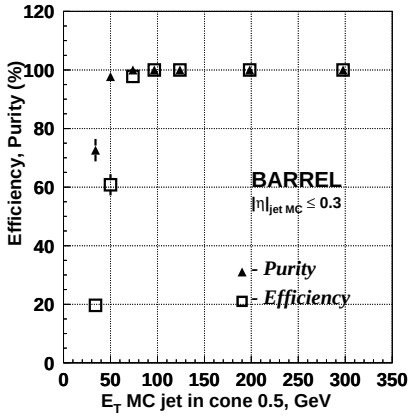


Figure 18: Energy dependence of efficiency (open squares) and purity (close triangle) of the jet reconstruction in Pb–Pb events in the barrel ($|\eta| < 1.4$).

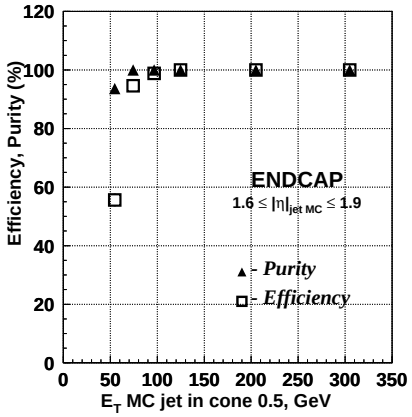


Figure 19: Energy dependence of efficiency (open squares) and purity (close triangle) of the jet reconstruction in Pb–Pb events in the endcap ($1.4 < |\eta| < 3.0$).

7 Conclusion

The possibility of reconstructing hard QCD jets with initial parton energies in the range 50-300 GeV is investigated in central Pb–Pb collisions (HIJING model, $dN_{\rm ch}/dy|_{y=0} = 5000$). The subtraction procedure allows the identification and measurement of jets in heavy ion collisions with very high efficiency and purity using only calorimeters. The performance of the algorithm is demonstrated with the CMS calorimetric system. On average, the measured jet energy in Pb–Pb collisions is the same as that in jet events without background. For jets above 75 GeV, the energy resolution of a jet reconstructed in the heavy ion environment is degraded by a factor ~ 1.3 as compared to jet events without background. The direction of hard jets in the heavy ion environment can be reconstructed with high accuracy, with resolutions in the η and φ smaller than the size of one calorimeter tower. The background subtraction algorithm can be used for jet reconstruction in heavy ion collisions as well as in high luminosity pp interactions.

8 Acknowledgments

The authors are thankful to L.Sarycheva, S.Kunori, I.Lokhtin, A.Snigirev, A.Demianov and A.Gribushin for useful discussions.

References

[1] M. Gyulassy and M. Plümer, Phys. Lett. **B 243** (1990) 432.

[2] I.P. Lokhtin and A.M. Snigirev, Eur. Phys. J. **C 16** (2000) 527.

[3] X.-N. Wang, Z. Huang, and I. Sarcevic, Phys. Rev. Lett. **231** (1996) 77; X.-N. Wang and Z. Huang, Phys. Rev. **C 55** (1997) 3047.

[4] V. Kartvelishvili, R. Kvatadze and R. Shanidze, Phys. Lett. **B 356** (1995) 589.

[5] M. Gyulassy and X.-N. Wang, Phys. Lett. **B 68** (1992) 1480.

[6] I.P. Lokhtin and A.M. Snigirev, Phys. Lett. **B 567** (2003) 39.

[7] N.A. Kruglov, I.P. Lokhtin, L.I. Sarycheva, and A.M. Snigirev, Z.Phys. C **76** (1997) 99.

[8] B. Flaugher and K. Meier, "A Compilation of Jet Finding Algorithms", Proc. of the Snowmass Summer Studies (1990), FERMILAB-CONF-90/248-E.

[9] N. Kruglov *et al.* CERN CMS TN/96-084.

[10] G. Baur *et al.* CERN CMS NOTE 1999/016.

[11] O. Kodolova *et al.* CERN CMS NOTE 1998/063.

[12] A. Accardi *et al.*, arXiv:hep-ph/0310274.

[13] CMS ECAL Technical Design Report, CERN/LHCC 97-33 (1997).

[14] CMS HCAL Technical Design Report, CERN/LHCC 97-31 (1997).

[15] CMSIM user's Guide at www.http://cmsdoc.cern.ch/swsi.html.

[16] "The Trigger and Data Acquisition Project", Volume II, Technical Design Report, CERN/LHCC 2002-26, CMS TDR 6.2 (2002).

[17] T. Sjöstrand *et al.*, Comput. Phys. Commun. **135** (2001) 238.

[18] M. Gyulassy and X.-N. Wang, Comput. Phys. Commun. **83**, 307 (1994).

[19] ORCA user's Guide at www.http://cmsdoc.cern.ch/swsi.html.