

CLUSTER ANALYSIS OF HIGH MULTIPLICITY NUCLEUS–NUCLEUS COLLISIONS

A. DĄBROWSKA, R. HOŁYŃSKI, M. SZARSKA, A. TRZUPEK
W. WOLTER, B. WOSIEK AND K. WOŹNIAK

H. Niewodniczański Institute of Nuclear Physics
Kawiry 26a, 30-055 Kraków, Poland

(Received July 6, 2000)

Analysis of the clustering properties in the multiparticle final states produced in central collisions of Pb nuclei with the Ag/Br target at 158 A GeV is presented. It is mainly focused on investigation of different effects which influence the results of the search for high density phase space regions. A comparison of different clustering procedures is performed with varied cuts on the cluster size and on the number of particles per cluster. We also discuss the dependence of the obtained results on the shape of single particle distributions and on the multiplicity of produced particles.

PACS numbers: 25.75.Gz, 12.38.Mh

1. Introduction

High multiplicity final hadronic states produced in relativistic heavy ion collisions are expected to show some cluster structures. These structures can be due to the production of Quark Gluon Plasma droplets [1], to the formation of Disoriented Chiral Condensate [2] or to the production of jets or mini-jets [1, 3]. Looking for clustering effects is also an efficient method to characterize nucleus–nucleus interactions on an event-by-event basis.

In this paper we present the analysis of collisions of lead nuclei with the Ag/Br emulsion targets at 158 A GeV. The experimental data were obtained by the KLM Collaboration from the EMU13 CERN experiment. A standard nuclear emulsion stack technique, used to record and measure collision events, provides a 4π solid angle coverage for produced charged particles. The measurements are, however, limited to particle emission angles. The analyzed sample consists of central collisions which constitute about 9% of all events selected as those with highest multiplicities. A typical total charged particle multiplicity is of the order of 900 per event. More experimental details can be found elsewhere [4, 5].

In our previous paper [5] we presented different methods of analysis of individual collision events, among which the search for cluster-like objects in the two-dimensional pseudorapidity–azimuth phase space was also discussed. In this paper we focus only on this latter problem, and we extend the previous analysis in order to account for different factors which may influence results of the search for highly occupied phase space domains. The main purpose of this paper is to provide a suitable tool for studying clustering properties in heavy ion collisions at the SPS energies as well as at much higher RHIC energies.

Particularly, we want to check the dependence of the results on the cluster definition. Three different clustering procedures are applied. One of these methods is the same as used recently in [6, 7], where the clustering phenomena were searched for in some of our Pb–Ag/Br collision events as well as in Monte Carlo generated events. The discussion of different methods, and cuts on the cluster size as well as on the minimal number of particles in a cluster are included in Section 2. A high particle density environment, which is specific for heavy ion collisions, makes the search for cluster structures difficult and sensitive to different effects. We include in Section 3 the study of the dependence on the shape of single particle distributions, and in Section 4 the dependence on the total number of particles produced in a collision. Finally, the Section 5 contains summary and concluding remarks.

2. Clustering procedures

The clustering procedure is based on distances measured between pairs of particles. In the two-dimensional pseudorapidity ($\eta = -\ln \tan \theta/2$)–azimuth (φ) phase space we define as a distance between two particles i and k :

$$R^2 = (\delta\eta)^2 + (\delta\varphi)^2, \quad (1)$$

where $\delta\eta = \eta_i - \eta_k$ and $\delta\varphi = \varphi_i - \varphi_k$. The value of R^2 is left as a parameter, $\varepsilon = R^2$, and the results for different ε values are shown. In the following the ε parameter will be referred to as the 'cluster size', although in fact it reflects only spacings between particles (in η – φ space) contained in the cluster. A real cluster size depends on the procedure used to build a cluster and for certain clustering procedures this quantity can not be well defined. The normalized cluster size, defined as $\varepsilon/\varepsilon_0$ is also used, where

$$\varepsilon_0 = (\Delta\eta)^2 + (\Delta\varphi)^2, \quad (2)$$

with $\Delta\eta$ and $\Delta\varphi$ denoting a full analyzed range correspondingly in η and φ space. In this analysis the η range covers 6 units of pseudorapidity (in the laboratory frame from 0 to 6) and $\Delta\varphi=\pi$, with the φ coverage from 0 to 2π .

Before starting the clusterization procedure, the ordering of particles in an event was randomized to eliminate some possible sequences of particles introduced during the measurement stage. Then, the first particle is selected as a center of the first cluster and around this particle we are building a cluster. This first step of the cluster formation is common for all studied clustering procedures.

In our previous paper [5] we used a rather restrictive clustering criterion, requiring that all particles forming a cluster should be at distances smaller than ε with respect to the first particle. After creating a first cluster, the same procedure is repeated for all remaining particles. We will refer to this type of clusters as star-like clusters. We may relax a clustering criterion, by accepting a consecutive particle, $i + 1$, if its distance to the particle i is smaller than ε . Objects formed in this way will be called snake-like clusters. Finally, the most general and least restrictive criterion is used to form tree-like structures (see Fig. 1 for a schematic drawing of possible cluster objects). In this case we require only that each particle belonging to a given cluster should be at the distance smaller than ε with respect to at least one other particle from this cluster. In this procedure, we form a first level of the

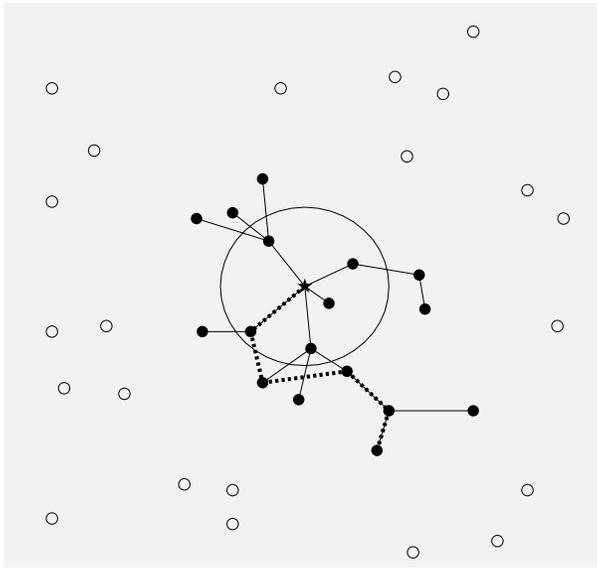


Fig. 1. A schematic illustration of different cluster structures. The first particle is marked by a star. Connected filled points (particles) represent a tree cluster. Dotted line connects particles forming a snake cluster. Particles forming a star cluster are contained inside a large circle. For clarity only a single tree, snake and star clusters are marked.

tree (a branch) exactly like in the star method. Then all particles from this branch are searched for companions within a distance smaller than ε , and the procedure is continued for all consecutive branches of a tree. For small ε values the number of branches will be small, while for large ε practically all particles must form a single large cluster. The above description of the clustering methods obviously indicates that for less restrictive criteria (tree-like or snake-like clusters) we will form fewer clusters but with larger number of particles contained in the cluster than for more demanding star clusters definition. It should be pointed out that the star and snake methods depend on the ordering of particles in an event. This dependence, however, does not obscure differences between the three algorithms. We have chosen the random ordering of particles since in this case the star method provides clusters with largest particle multiplicities.

Table I lists the number of clusters, N_{cl} and the average multiplicity of particles per cluster, $\langle m \rangle$, for the three methods and for a range of ε values. The smallest ε value corresponds to the limit set by our measuring resolution in η and φ . Clusters which contain at least two particles, $m_0=2$, are counted. Later on this cut will be changed. All variables quoted in Table I are averaged over all central collision events. As expected, for a star method we observe more but smaller clusters as compared to the results obtained from the other algorithms. This holds in a wide range of ε values. In the case of the tree method even for cluster sizes about 30 times smaller than the total $\Delta\eta \times \Delta\varphi$ phase space, all produced particles form a single large cluster. The same effects are also observed on the basis of individual events. Significant differences obviously should also show up in the distributions of

TABLE I

Number of clusters and average cluster multiplicities for different ε values and for the three cluster definitions.

ε	N_{cl}			$\langle m \rangle$		
	tree	snake	star	tree	snake	star
$\varepsilon_0/2^3$	1.0	1.53 ± 0.08	4.77 ± 0.10	905 ± 15	681 ± 36	193.6 ± 4.9
$\varepsilon_0/2^4$	1.0	2.72 ± 0.11	8.70 ± 0.09	905 ± 15	362 ± 15	104.5 ± 2.0
$\varepsilon_0/2^5$	1.0	5.13 ± 0.15	16.25 ± 0.15	905 ± 15	184.9 ± 6.6	55.79 ± 0.95
$\varepsilon_0/2^6$	1.02 ± 0.02	10.45 ± 0.27	30.06 ± 0.22	896 ± 16	89.55 ± 3.05	30.06 ± 0.48
$\varepsilon_0/2^7$	1.26 ± 0.07	20.79 ± 0.41	53.02 ± 0.30	788 ± 30	43.98 ± 1.17	16.94 ± 0.25
$\varepsilon_0/2^8$	2.40 ± 0.17	41.19 ± 0.58	90.89 ± 0.53	482 ± 34	21.71 ± 0.47	9.79 ± 0.14
$\varepsilon_0/2^9$	8.04 ± 0.38	75.70 ± 0.76	145.64 ± 1.00	131.1 ± 9.1	11.46 ± 0.21	5.96 ± 0.08
$\varepsilon_0/2^{10}$	42.98 ± 1.30	134.92 ± 1.16	207.17 ± 2.18	21.72 ± 0.99	6.09 ± 0.09	3.94 ± 0.04
$\varepsilon_0/2^{11}$	143.62 ± 1.27	198.13 ± 2.27	241.32 ± 3.81	5.42 ± 0.12	3.67 ± 0.04	2.95 ± 0.02
$\varepsilon_0/2^{12}$	189.42 ± 3.34	207.42 ± 4.05	224.38 ± 4.84	3.15 ± 0.03	2.75 ± 0.02	2.47 ± 0.01

the number of particles per cluster, m . These are depicted in Fig. 2 for the ε value of $\varepsilon_0/2^{10}$. The long large multiplicity tails are clearly visible for the tree and snake type of clusters. It should be pointed out that in [6,7] the multiplicity distributions of particles in clusters were investigated by using the tree clustering algorithm, and the analysis was carried out down to ε values much smaller than the limit set by the experimental resolution. Therefore, their results cannot be directly compared to our previously published data, where the star procedure was applied¹.

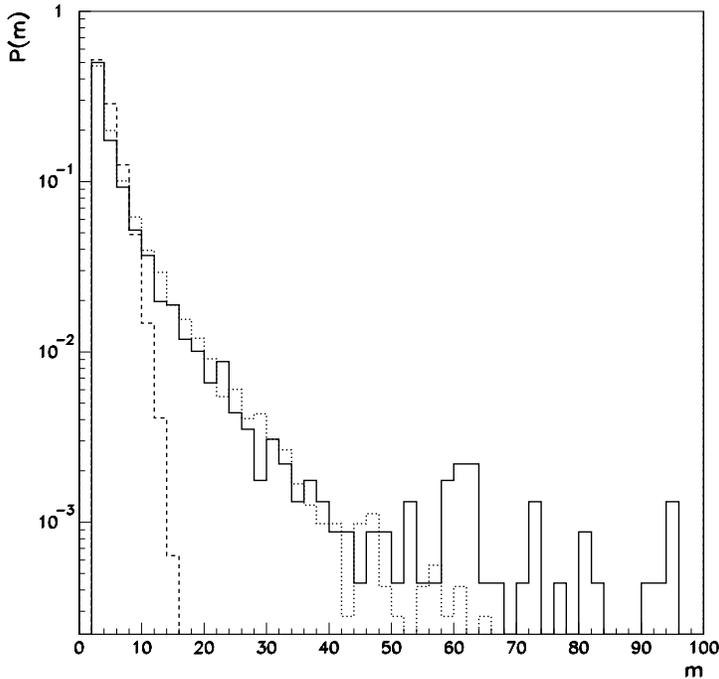


Fig. 2. Comparison of the cluster multiplicity distributions for tree clusters (solid), snake clusters (dotted) and star clusters (dashed) for a fixed cluster size $\varepsilon = \varepsilon_0/2^{10}$.

It was proposed in [7] that one can efficiently differentiate between individual collision events by studying the entropy, which is defined for a given event as:

$$S(\varepsilon) = - \sum_k p_k \ln p_k, \tag{3}$$

where $p_k = m_k / \sum_k m_k$ denotes the probability of finding a particle in the k -th cluster, and summations run over all N_{cl} clusters recorded in an event.

¹ Furthermore, in our previous paper [5] we required that at least 5 particles should form a cluster whereas in [6,7] this threshold was set to 2.

Certainly this quantity is worth studying. In Fig. 3 we compare the dependence of the entropy, averaged over all events in the sample, on $\varepsilon/\varepsilon_0$ for the three discussed methods. One can see that the entropy is also quite sensitive to the procedure used to select cluster objects in the range of studied cluster sizes. It seems that the entropy of star-like clusters reaches the maximum in the range of studied cluster sizes. The maximal entropy calculated for the other two types of clusters seems to be reached at yet smaller cluster sizes than those shown in Fig. 3.

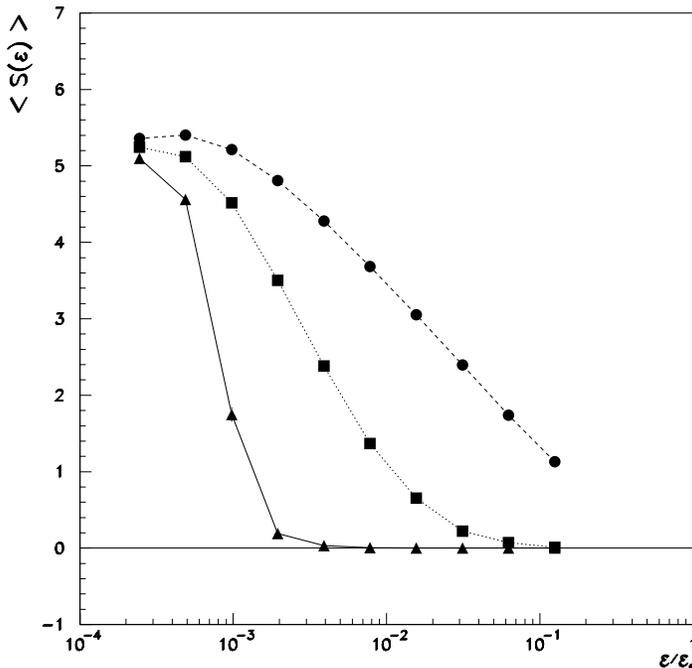


Fig. 3. Cluster entropy as a function of the normalized cluster size for the star (circles), snake (squares) and tree (triangles) clusters. Lines are drawn to guide the eye.

For nucleus–nucleus collisions it is reasonable to set a cut on the minimal number of particles per cluster at higher m_0 values in order to identify densely populated phase space regions in the overall large density environment. In Table II the values of cluster parameters are listed for different cuts on m : $m_0 = 2$, $m_0 = 4$, $m_0 = 8$ and $m_0 = 12$ for the value of $\varepsilon = \varepsilon_0/2^{10}$. Obviously, with increasing m_0 the number of clusters rapidly decreases, irrespectively of the method of analysis. However, this suppression of cluster production is the strongest for the most restrictive cluster definition, *i.e.* for star clusters.

TABLE II

Number of clusters and average cluster multiplicities for different m_0 cuts and for a fixed size of the cluster, $\varepsilon = \varepsilon_0/2^{10}$.

m_0	N_{cl}			$\langle m \rangle$		
	tree	snake	star	tree	snake	star
$m_0 = 2$	42.98 ± 1.30	134.92 ± 1.16	207.17 ± 2.18	21.27 ± 0.99	6.09 ± 0.09	3.94 ± 0.04
$m_0 = 4$	21.49 ± 0.78	70.53 ± 0.87	99.47 ± 2.23	42.17 ± 2.31	9.49 ± 0.16	5.58 ± 0.04
$m_0 = 8$	10.02 ± 0.45	30.08 ± 0.69	14.23 ± 0.87	88.45 ± 5.82	15.34 ± 0.27	9.07 ± 0.06
$m_0 = 12$	6.21 ± 0.37	16.40 ± 0.57	1.06 ± 0.17	159.43 ± 19.0	20.37 ± 0.36	13.27 ± 0.15

Clearly, the resulting clustering structure strongly depends on the way the clusters are formed. Particularly striking is the difference between the star procedure and the tree method. Fig. 4 shows the distribution of the number of branches, N_{BR} , per tree clusters of size $\varepsilon = \varepsilon_0/2^{10}$. This distribution is rather broad with the mean value of branches, $\langle N_{BR} \rangle = 5.74 \pm 0.15$, whereas for the star clusters there is always only a single branch developed from one particle. It is expected that these differences should disappear

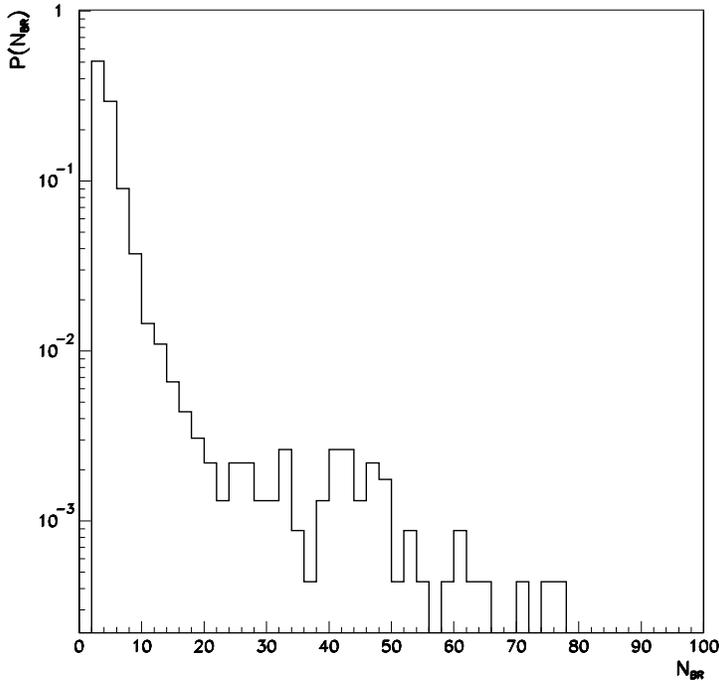


Fig. 4. Distribution of the number of branches, N_{BR} , for tree-like clusters for $\varepsilon = \varepsilon_0/2^{10}$.

for very small ε values. However, the analysis of experimental data must be limited to cluster sizes that are not smaller than the resolution in the two-track separation.

3. Dependence on the shape of single particle distributions

The non-uniform shape of single particle distributions may also affect the search for cluster-like objects. For our data the measured φ distribution is uniform, but the $dN/d\eta$ distribution has a Gaussian shape with σ of 1.5 [4]. To check whether the shape of pseudorapidity distribution influences the results we performed the cluster analysis for the two samples of events with randomly generated tracks. In the both samples the total event multiplicities are the same and match those in measured events. For the first sample (Sample I) both particles azimuths and pseudorapidities are randomly generated according to the uniform distributions. Thus for this sample we preserve only the correlations due to the average particle density (*i.e.* total number of particles contained in the $\Delta\eta \times \Delta\varphi$ phase space). For the other sample (Sample II) particles azimuths are generated in the same way as in the Sample I, but their pseudorapidities are randomly distributed according to a Gaussian distribution of mean $\langle\eta\rangle = 3.4$ and variance $\sigma = 1.5$. This distribution reproduces the experimentally measured $dN/d\eta$ spectrum. In Fig. 5 we plot the ratio of the average cluster multiplicities, $\langle m \rangle_{\text{II}}/\langle m \rangle_{\text{I}}$ (open circles) as a function of the normalized cluster size for the snake (Fig. 5(a)) and star (Fig. 5(b)) clusters. One can see that for the non-uniform $dN/d\eta$ distribution stronger clustering of particles is present. We systematically observe larger cluster multiplicities for the Sample II as in the case of uniformly distributed particles (Sample I) for the both snake and star procedures. Furthermore, for the smallest cluster size ($\varepsilon = \varepsilon_0/2^{12}$), the number of clusters is also larger for the Sample II than for the Sample I. Still larger differences are observed by assuming a narrower $dN/d\eta$ distribution².

For the tree method, which provides quite extended objects, we see differences between the results obtained for the Sample II and the Sample I, but only for small cluster sizes ($\varepsilon/\varepsilon_0 < 0.02$). They show up either in the increased number of clusters or in the increased cluster multiplicities for the Sample II as compared to the Sample I.

² This was checked for the Gaussian $dN/d\eta$ distribution with $\sigma = 1.3$, which was found for the Pb–Ag/Br collisions generated from the Venus Monte Carlo Model [4, 8].

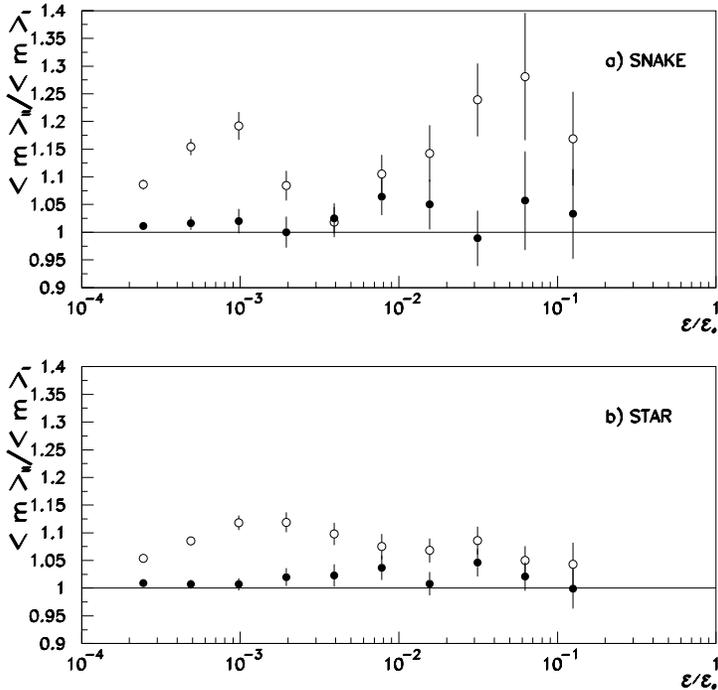


Fig. 5. Dependence of the ratio of cluster multiplicities for the two samples of events with randomly generated particles on the normalized cluster size for the snake (a) and star (b) method. Open circles denote the results of the analysis of normal η and φ variables, while filled circles show the same for the scaled η^* and φ^* variables. See text for more explanations.

Thus, it is clear that the clustering structures depend on the shape of the single particle distributions. In order to eliminate this dependence one can use scaled variables [9]. A scaled pseudorapidity, η^* , is defined as:

$$\eta^*(\eta) = \int_0^\eta \rho(\eta') d\eta' / \int_{\Delta\eta} \rho(\eta') d\eta', \quad (4)$$

where $\rho(\eta)$ is the density distribution averaged over all events in the sample. The $\rho(\eta^*)$ distribution ranges from 0 to 1 and is flat. For the sake of completeness we applied the same procedure for the azimuthal angle. The resulting two-dimensional distribution $d^2N/d\eta^*d\varphi^*$ is indeed flat, indicating that there is no sizable correlations between pseudorapidity and azimuth³.

³ By transforming original variables separately in pseudorapidity and azimuth we assumed that these two variables are independent.

TABLE III

Number of clusters and average cluster multiplicities for the two samples of events with different average total multiplicities, $\langle N \rangle$.

ε	sample $\langle N \rangle$	N_{cl}			$\langle m \rangle$		
		tree	snake	star	tree	snake	star
$\varepsilon_0/2^4$	796 ± 9	1.0	3.05 ± 0.23	9.20 ± 0.17	796 ± 9	314 ± 41	86.82 ± 1.24
	1023 ± 11	1.0	3.16 ± 0.19	9.37 ± 0.13	1023 ± 11	347 ± 22	109.67 ± 2.08
$\varepsilon_0/2^5$	796 ± 9	1.0	11.80 ± 0.47	31.25 ± 0.33	796 ± 9	69.89 ± 3.78	25.45 ± 0.34
	1023 ± 11	1.0	11.42 ± 0.53	32.42 ± 0.28	1023 ± 11	92.78 ± 4.19	31.52 ± 0.35
$\varepsilon_0/2^8$	796 ± 9	1.05 ± 0.05	42.65 ± 0.54	99.00 ± 0.84	795 ± 20	18.19 ± 0.23	7.91 ± 0.07
	1023 ± 11	1.0	45.05 ± 0.93	105.84 ± 0.56	1023 ± 11	22.46 ± 0.52	9.57 ± 0.10
$\varepsilon_0/2^{10}$	796 ± 9	77.35 ± 2.20	153.15 ± 1.62	210.20 ± 2.42	9.99 ± 0.39	4.63 ± 0.04	3.33 ± 0.02
	1023 ± 11	49.11 ± 1.85	163.05 ± 2.04	245.58 ± 2.64	20.97 ± 0.93	5.76 ± 0.06	3.80 ± 0.03
$\varepsilon_0/2^{12}$	796 ± 9	168.05 ± 3.45	173.30 ± 3.30	180.25 ± 3.30	2.68 ± 0.02	2.47 ± 0.01	2.31 ± 0.01
	1023 ± 11	227.53 ± 3.24	243.84 ± 3.41	258.58 ± 3.84	2.95 ± 0.02	2.62 ± 0.02	2.40 ± 0.01

Now we re-analyzed the previously used Sample II in terms of scaled variables. The results are shown in Fig. 5 as filled circles. It can be seen that the use of the scaled variables significantly reduces the differences due to the non-uniform shape of $dN/d\eta$ distribution. Therefore it is advisable to use the scaled variables, particularly when we want to compare data sets or collision events with different shapes of single particle spectra, *e.g.* our data with the Venus model simulations.

4. Dependence on the event multiplicity

It is obvious that the local density fluctuations must depend on the event multiplicities. Confining more particles in the finite $\Delta\eta \times \Delta\varphi$ space should give different clustering patterns than in the case of small event multiplicities. The question remains how large this difference in the event multiplicity should be to change the pattern. Presumably, in the high density regions, even a small difference in particle multiplicity may lead to different structures. In our sample of selected central events, the particle multiplicities vary from 740 up to 1120. We have selected the two sub-samples of events from our data set: the first with $N \leq 850$ and the second with $N \geq 950$, where N is the charged particle multiplicity in the analyzed $\Delta\eta$ and $\Delta\varphi$ range. The average multiplicities are correspondingly 796 and 1023. The results obtained for these two data sets are compared in Table III for several ε values. The analysis was performed using scaled variables and with $m_0 = 2$. For the snake and star procedures we obtain more clusters and larger cluster multiplicities for high multiplicity events than for the low multiplicity collisions. For the tree method the same effect is seen only for the smallest ε value. It should be noted that, even for the selected central collisions, the event to event multiplicity can likely vary by 20–30 %.

5. Summary

The search for cluster-like objects in the two-dimensional pseudorapidity–azimuth phase space was discussed. This study mainly concentrated on testing different clustering procedures as well as dependence on cuts on the cluster size and cluster multiplicity. We also checked the dependence on the shape of single particle distributions and on the overall event multiplicity. High multiplicity events of Pb collisions with the Ag/Br emulsion target at 158 A GeV were used in this analysis.

It was shown that three different analysis methods, which can be used to search for densely populated phase space domains, lead to different patterns of event cluster structures in a wide range of cluster sizes. Particularly pronounced are differences between the restrictive 'star' method and the other two methods, 'snake' and 'tree' in which more loose criteria are used. The effect of applying different cuts on the minimal number of particles per cluster was also investigated. We observe expected change of the event cluster properties when this cut was set at higher values, qualitatively independent of the method of analysis. Finally it was also shown that the effects due to the non-uniform shape of single particle distributions and different event multiplicities are important and should be eliminated by using variables uniformly distributed in the phase space and comparing the results for events with similar multiplicities.

Summarizing, it is evident that in the high density final hadronic states produced in heavy ion collisions, the search for clustering structures is sensitive to many different effects. It is recommended that in order to distinctly identify densely populated regions, one has to use a rather restrictive clustering algorithm like the 'star' method. The care should be put when comparing results for events with different multiplicities of produced hadrons and different shapes of single particle spectra. We are going to apply such thorough procedure in the analysis of central collisions of Pb nuclei with the Pb target at the energy of 158 A GeV. The particle densities produced in these collisions are higher than in the analyzed Pb collisions with lighter, Ag/Br target. It would be also interesting to perform such analysis for Au–Au collisions at the RHIC energy where one expects to produce still higher particle densities. In Fig. 6 we show the comparison of the star cluster analysis with $m_0 = 8$ for our sample of Pb–Ag/Br collisions and for simulated [11] head on Au–Au collisions at RHIC energy. The effect of the large particle density (the average density for simulated RHIC data is by a factor of about 5 larger than for Pb–Ag/Br measured events) is clearly seen. More clusters and higher cluster multiplicities are observed for large density data. The analysis of this rich cluster structure requires the application of a quite restrictive clustering procedure in order to identify interesting events. The full

solid angle coverage of the Multiplicity Array of the PHOBOS experiment at RHIC [10] will make the study of cluster effects possible as soon as the first measurements are available.

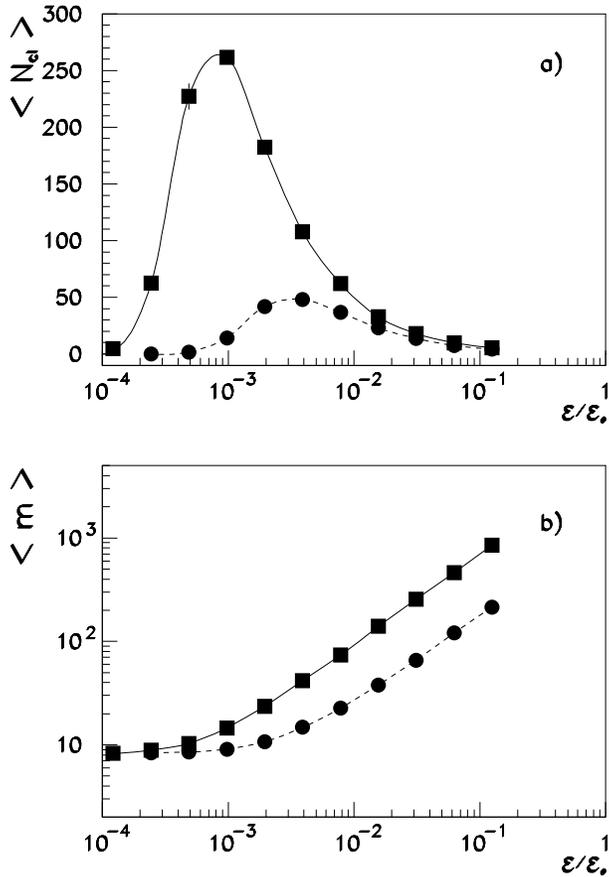


Fig. 6. Comparison of the star cluster analysis ($m_0 = 8$) for Pb–Ag/Br data at $\sqrt{s} = 17$ A GeV (circles) and simulated Au–Au collisions at $\sqrt{s} = 200$ A GeV (squares). Plots show the dependence of the average number of clusters (a) and average cluster multiplicity (b) on the normalized cluster size.

This work is partially supported by the Polish Committee for Scientific Research under Grants No. 2 P03B 05417 and 2 P03B 04916.

REFERENCES

- [1] J.-P. Blaizot, *Nucl. Phys.* **A661**, 3c (1999); S. Bass *et al.*, *Nucl. Phys.* **A661**, 205c (1999).
- [2] A.A. Anselm, M.G. Ryskin, *Phys. Lett.* **B266**, 482 (1991); J.-P. Blaizot, A. Krzywicki, *Phys. Rev.* **D46**, 246 (1992); J.D. Bjorken, *Int. J. Mod. Phys.* **A7**, 4189 (1992); T.K. Nayak *et al.*, *Nucl. Phys.* **A638**, 249c (1998).
- [3] X.-N. Wang, *Nucl. Phys.* **A661**, 609c (1999); M. Gyulassy, P. Levai, L. Vitev, *Nucl. Phys.* **A661**, 637c (1999); I. Sarcevic, *Nucl. Phys.* **A638**, 531c (1998).
- [4] A. Dąbrowska *et al.*, (KLM Collab.), *Nucl. Phys.* **A633**, 357 (1998).
- [5] M.L. Cherry *et al.*, (KLM Collab.), *Acta Phys. Pol.* **B29**, 2129 (1998).
- [6] K. Fiałkowski, R. Wit, *Acta Phys. Pol.* **B30**, 2759 (1999).
- [7] K. Fiałkowski, R. Wit, *Acta Phys. Pol.* **B31**, 681 (2000).
- [8] K. Werner, *Phys. Rep.* **232**, 87 (1995).
- [9] A. Białas, M. Gaździcki, *Phys. Lett.* **B252**, 483 (1990).
- [10] J.M. Katzy *et al.*, *Nucl. Phys.* **A661**, 690c (1999).
- [11] X.N. Wang, *Nucl. Phys.* **A661**, 210c (1999).