GEM TN-91-33

Showershape and Preradiator Study for Higgs $\rightarrow \gamma + \gamma$ Background

Hiroaki Yamamoto
California Institute of Technology

March 28, 1992

Abstract:

In order to observe a clear signal of low mass Higgs with mass around 80 to 140 GeV, it is very important to reject fake isolated γs. In this note, the showershape distributions and preradiator signals in the BaF_2 are studied to reduce these fate isolated γs.

Showershape and Preradiator Study for Higgs $\rightarrow \gamma + \gamma$ Background

Hiroaki Yamamoto
Lauritsen Laboratory
California Institute of Technology
Pasadena, CA 91125

March 28, 1992

Abstract

In order to observe a clear signal of low mass Higgs with mass around 80 to 140 GeV, it is very important to reject fake isolated γs. In this note, the showershape distributions and preradiator signals in the BaF_2 are studied to reduce these fake isolated γs.

1 Overview

This note summarizes the preliminary results of the study to reject fake isolated γs using the shower shape and the preradiator information in the BaF_2 calorimeter, where "fake isolated γ"s mean multiple almost collinear γs which pass photon isolation criteria using the transverse energy distributions. Hereafter, "isolated γ candidates" mean real single γs and "fake isolated γ"s.

As has been extensively studies by Ren-Yuan Zhu and others [R-Y Zhu, CALT-68-1777, and references there in], the QCD background against the low mass Higgs (80-140 GeV) signal, Higgs $\rightarrow \gamma+\gamma$, is very large. The isolation cut by requiring small E_t deposits around the isolated γ candidates effectively reduces these fake γs. But still the reducible backgrounds is around the same size as the irreducible backgrounds. To increase the significance of the signal, it is very important to reduce these reducible backgrounds as much as possible.

2 Monte Carlo

Pythia 55 is used for the physics event generator. The rejection was studied using three sets of event samples, in addition to the comparison between single γs and π^0, η and K_s^0. These three sets of events are the ones generated by Ren-Yuan Zhu, 1) jet-jet events, 2) γ-jet events and 3) $t\bar{t}$ events. One important thing to note is, around 50 % of the fake isolated γs contain more than 2 γs. (At least, Pythia 55 predicts that way.) Isolated γ candidates are selected using the method proposed by Ren-Yuan Zhu, requiring the ΣE_t in a cone with radius R being less than a given value. Typical values used are : $\Sigma E_t(max) = 5, 10, 15, 20$ GeV and dR = 0.3, 0.45, 0.6, 0.75.

BaF_2 calorimeter proposed by Ren-Yuan Zhu and others [R-Y Zhu, CALT-68-1767 and references there in] are used in this analysis. The Monte Carlo simulation is done by GEANT, using the package CALO developed by A. Gurtu. This package implements the BaF_2 EM calorimeter, consisted of 11x11 BaF_2 crystals. The front of each crystal is 3 cm x 3 cm and the end is 5 cm x 5 cm, with the length of 50 cm, corresponding to 25 X_0.

When simulating the detector, each γ, π^0 or isolated γ candidates from the event samples are rotated to hit the crystal at the center of the 11x11 crystals. The direction is randomly chosen so that hits distribute uniformly in the 3 cm x 3 cm crystal front end. The z position of the event vertex was generated assuming a Gaussian distribution with $\sigma = 5cm$. To simulate this effect, the hit position determined in the first step is shifted by the displacement of the event vertex, by keeping the direction vector of each γ. The net effect is, each crystal is not aligned to point to the real event vertex. This will be the simulation for the worst case. The effect will be smaller for crystals in the forward region.

One layer preradiator is placed at 10 cm from the front of the crystal, corresponding to $5X_0$, the dimension being 3.4 cm x 3.4 cm.

3 Shower Shape Analysis

3.1 Method

This showershape analysis uses energy distributions in the 5x5 crystal, the center of which is the crystal which has the largest energy deposit among the 11x11 crystals. The basic idea is, the spread of the energy distribution is smaller for the single γ hit than the multi γ hit, and the distribution is circular if there is one hit, but will be oval if there are multiple hits.

The energy distributions in the 5x5 crystals are fit by an oval shape Gaussian distribution to find the spread in the minor and major axes. If there were no event vertex smearing, both spreads, along the major and minor axes, help to distinguish the single and multi γ hits. But the event vertex smearing of the size of 5 cm along the beam

direction makes the spread in the major axis (most likely along the bean direction) less useful. So only the spread along the minor axis is used to distinguish the single and multi γ hits.

The shape analysis goes as follows. In the following, $e_{i,j}$ means the transverse energy deposit in the cell (i,j), where i and j = -2 to 2, in the 5x5 crystals around the center of the hit.

1. The center of the 5x5 crystals is the one which has the largest transverse energy deposit.

2. If the cells which have the 2nd and 3rd largest energy deposits are not in 3x3 cells around the center, this candidate is rejected.

3. Fit the projection of the $e_{i,j}$ to x and y by Gaussian to find the center and σ. These 4 values are used as the initial values for the following fit.

4. Fit the 5x5 energy deposits using the following oval shape, with 6 parameters.
   $f(x,y) = A * exp(-\frac{X^2}{2*\sigma_{min}} - \frac{Y^2}{2*\sigma_{max}})$

   $X^2 = cos(\theta)^2 * (x - x0)^2 + sin(\theta)^2 * (y - y0)^2$
   $Y^2 = sin(\theta)^2 * (x - x0)^2 + cos(\theta)^2 * (y - y0)^2$

   where
   A : overall normalization
   σ_{min} and σ_{max} : standard deviations of the oval shape ($\sigma_{min} < \sigma_{max}$)
   θ : angle between the original xy axis and the principal axis of the oval.
   x0, y0 : center of the oval

5. Reject if $\sigma_{min} < 0.0165$ to keep 95% of the real single γ Reject if $\sigma_{min} < 0.0160$ to keep 90% of the real single γ. These two values are determined by 40 and 80 GeV γs simulated by GEANT. The efficiency for the lower energy γ is slightly lower. For example, the efficiencies for the 20 GeV γ are several % less in both cases. See Fig. 2 for more details, which are explained below.

Technical comments about this fit are in order.

The fit used in step 4 is not a good fit from the theoretical point of view. The reasons are

1. the distribution in each cell is not taken into account and

2. the real distribution is a sum of 2 Gaussians with narrow spread and wide spread.

Because of this, the σ_{min} comes out to be 2 to 3 times larger than the real spread and the tails are not fit well. The reason current method is chosen is

1. it works phenomenologically and

2. the more complicated fit has too many parameters and the fitted result does not help to distinguish single and multi γs as well as the current method.

3.2 Results

Fig.1 shows the differences in σ_{min} distributions of a single and multi-γ hits. Fig 1-a is the distribution for the 40 GeV single γ hits, and 1-b is the one for single γs in $pp \rightarrow \gamma + jet + X$ events. Fig.1-c is the one for 40 GeV π^0s, and 1-d is the one for multi-γs in $pp \rightarrow \gamma + jet + X$ events. In short, 1-a and 1-b are signals and 1-c and 1-d are backgrounds.

Fig. 1-e (1-f) compares the efficiencies between 1-a and 1-c (1-b and 1-d) as a function of the σ_{min} cut values. In these plots, solid lines show the efficiencies for the single γ hits and dashed line for the multi γ hits. The efficiency is the fraction of hits which pass the σ_{min} cut. E.g., in Fig.1-f, if you choose $\sigma_{min} < 0.016$, around 90% of the the single γs' in γ - jets events (solid line) are kept, but only 50% of the fake isolated γs (dashed line) are accepted.

As is seen from Fig. 1-e, this method does not work well to separate γs and π^0s with high momenta. But the multi γ hits in the real events, like γ-jet events, contain multi-γs (≥ 3) and this method (or any method) works better than the case to reject single π^0 background with the same energy.

Fig. 2 summarizes the comparison of γ and π^0 efficiencies. Fig. 2-a, 2-b, 2-c and 2-d compare the efficiencies for 12 GeV, 20 GeV, 40 GeV and 80 GeV γ and π^0. As is clear from these plots, the cut works better for low momentum hits, and does not work well for the high momentum hits.

3.3 $d\overline{\Theta}$ cut : rough estimation in the analysis

Shower simulation by GEANT is time consuming. To roughly estimate the shower shape cut in the analysis, $d\overline{\Theta}$ cut was proposed by R-Y Zhu. $d\overline{\Theta}$ is defined for the isolated γ candidates as

$$d\overline{\Theta} = \frac{\Sigma \theta_i \cdot E_i}{\Sigma E_i}$$

where θ_i is the angle between the direction of the ith γ and the direction of the vector sum of the all γs, and E_i is the energy of the ith γ.

Fig. 3 shows the correspondence between the σ_{min} cut and $d\overline{\Theta}$ cut for 20 GeV and 40 GeV π^0s. Fig. 3-a and 3-d are $d\overline{\Theta}$ distributions of 20 GeV and 40 GeV π^0s, and Fig. 3-b and 3-e are the fractions of π^0s which pass the shower shape cut with $\sigma_{min} < 0.0165$, and Fig. 3-c and 3-f are those after the cut $\sigma_{min} < 0.0160$.

Fig. 4 shows the same for those isolated γ candidates in the $pp \rightarrow \gamma + jet$ and $pp \rightarrow t\bar{t} + X \rightarrow lepton + \gamma + \gamma$ events. Fig. 4-a is the $d\overline{\Theta}$ distribution in $pp \rightarrow \gamma + jet + X$

events, and 4-d in $pp \rightarrow t\bar{t} + X \rightarrow lepton + \gamma + \gamma + X$ events. Negative $d\overline{\Theta}$ entries represent single γ hits. Fig.4-b and 4-e are the fractions after the cut $\sigma_{min} < 0.0165$ for each process and 4-c and 4-f after the cut $\sigma_{min} < 0.0160$.

The rejection by the shower shape cut depends on various factors, like the number of γs or the energies of individual γs in the candidates, but roughly speaking, the rejection by $d\overline{\Theta} < 0.05$ corresponds to the cut by $\sigma_{min} < 0.016$. The comparison of these two cuts applied to jet-jet, jet-γ and the $t\bar{t}$ events shows reasonable agreement, and, for the rough estimation, $d\overline{\Theta}$ cut can be used.

4 Preradiator

4.1 Method

One layer of preradiator is placed at 10 cm ($5X_0$) from the front of BaF_2 crystal, 3.4cm x 3.4cm dimension. Two cases were studied.

1. 1mm x 1mm cells
2. 1mm spacing strips crossing each other diagonally

The detector simulation was done with the event z vertex position at 0 and the momentum perpendicular to the beam axis. The effect of the event vertex smearing does not change the result very much.

First, we discuss the cell type preradiator, case 1. The active cell distribution is used to distinguish the single and multi-γ hits, where a active cell is the one through which $\geq N$ (called threshold hereafter) charged particles with momentum larger than 2 MeV pass.

We cannot use the height information, like the energy deposit or number of charged tracks in one cell, and the distribution of the active cells is the only way to distinguish between single γ and multi-γ hits. The difference is, for the single γ hit, the distribution is smaller in radius, grouped in one cluster, and looks like circle, compared to the case for multi-γ hits, which have larger distributions, can have separated clusters and looks like oval.

The method used in this study is as follows:

1. Join two active cells which are placed side-by-side.
2. Repeat step 1 to form a clusters
3. Repeat 1 and 2 until all cells are assigned to one of the clusters.
4. Take the cluster which contains the largest number of active cells (C1) as the signal of the γ.

For examples, in Fig. 5, all A are joined to form one cluster and all B are joined to form another. The bottom-right A and top-left B are not joined, because they are not side-by-side.

Some quantities useful to distinguish hits by one γ and multiple γs are

1. The distance between the center of the all active cells and the center of the largest cluster (C1), where the center is calculated as

   $x_0 = \frac{\Sigma x_i}{N}$
   $y_0 = \frac{\Sigma y_i}{N}$

   where x_i (y_i) is the coordinate of each active cell and N is the total number of active cells.

2. The number of clusters with active cells larger than 1.

4.2 Threshold : Number of charged tracks in active cells

The main purpose of preradiator is to reject high momentum fake isolated γs. If you set the threshold too low, like N = 2, there are too many active cells by the fructuation of the tail of the single γ hit. The threshold should be set to see the peak of each γ, not the tail.

For example, active cell pattern line Fig. 6-a is not rare for 80 GeV γ with N = 2. If you set N to 4, this pattern changes to the one in Fig.6-b.

Fig. 6-c is the active cell distribution of a 80 GeV π^0 with N = 2. If you choose N = 4, the pattern looks more reasonable, two separated clusters, at around the right place as is shown in Fig.6-d.

Because of this, you cannot set the threshold too low to distinguish one γ and multi γ hits. But if you set the threshold to high, low momentum γs cannot show any signal in the preradiator, and some kinds of multi-γs cannot be recognized, like low momentum π^0s or high momentum π^0 with one γ carrying low momentum. For example, if you set N = 10, it is very inefficient to reject 20 GeV π^0s.

Fig. 7-1 and 7-2 are the distributions of the distance between the center of the all active cells and the center of the largest cluster. Fig.7-1 is the case with N = 2, and 7-2 is the case with N = 6. In each figure, 4 distributions are compared, 20 GeV and 80 GeV γ and 20 GeV and 80GeV π^0.

As is clear from Fig.7-1, with N = 2, it is easy to separate 20 GeV γ and π^0, but is difficult to separate 80 GeV ones. But with N = 6, as is shown in Fig.7-2, it is easier to separate 80 GeV γ and π^0 than to distinguish 20 GeV γ and π^0.

4.3 1mm spacing strip

The preradiator using strips with 1 mm spacing was also studied. The strips go across orthogonally each other. When one uses strips, instead of cells, the same kind of problem occurs. I.e., it is very difficult to obtain high background rejection rates for wide momentum range.

The gamma selection using the strip information, in stead of cell information, goes as follows

1. Combine neibouring active strips in each direction (called cluster).
2. If there are more than 1 cluster with more than one strip in either direction, reject the candidate.
3. If the largest cluster in a direction contains only one strip, and if there are more than 1 cluster in that direction, reject the candidate.
4. Take the largest clusters in each direction.
5. Calculate the distance between the overall center and the center of the largest cluster.
6. Require that the quadrature sum of these distances in each direction is small.

4.4 Performance

The optimal threshold for the strip is around the twice of the optimal threshold for the cell. The best value has not been found yet, but in the following calculation, N = 4 for the 1mmx1mm cell and N=8 for the 1mm spacing strips are used. The cut on the distance between the overall center and the largest cluster center is required to be less than 1.5mm for both type.

The following table summarizes the results after applying these cuts.

Table 1: Acceptance for single γ and multi γ particles

Energy (GeV)	γ	π^0	η	K_s^0
20	90/93	33 / 67		
40	92 / 93	36 / 50		
80	92 / 99	53 / 74	22 / 26	12 / 16
120	98 / 94	64 / 86		
160	95 / 96	70 / 90	30 / 44	16 / 43

This table shows the fraction that each type of particle pass the single γ selection criteria at various energies. The results for η and K_s^0 are shown in addition to a single γ and π_0, to see the efficiencies for fake isolated γs with multiple gammas (≥ 3). Note that $K_s^0 \rightarrow 4\gamma s$ and $\eta \rightarrow 2$ or 6 γs. The two numbers, A/B, are the acceptances when the preradiator with 1mmx1mm cell is used and when the one with the 1mm spacing strips are used.

As the momentum goes up higher, the π^0 rejection becomes worse, because the opening angle between the two γs becomes smaller. E.g., the opening angle of γs coming from the 160 GeV π^0 decay is 1.6mr, which corresponds to 1.4 mm separation on the preradiator plane.

In the energy region around 80 GeV to 160 GeV, η and K_s^0 can be rejected very well, because

1. there are multiple γs, causing wider spreads, and
2. the energy of individual γ is high enough to show the signal of each γ in the preradiator.

5 Summary

In this note, the shower shape distribution and preradiator signals are studied to reject fake isolated γs. The reduction of fake isolated γ is very important to enhance the significance of the low mass Higgs particle signal, Higgs $\rightarrow \gamma + \gamma$.

The shower shape can help to reduce the fake isolated γs in the medium momentum range, around 40 GeV and below. When the momentum goes up, the distinction between the single and and multiple γs become difficult. Fortunately, the fake isolated γs in the QCD events contain multiple γs, these backgrounds can be rejected by 50% on average, although the rejection of high momentum π^0 is almost impossible.

The preradiator is better suited to reject the high momentum fake isolated γs. In this note, two kinds of one layer preradiator were studied, one with 1mm x 1mm cells, the other with 1 mm spacing strips. Unfortunately, high momentum π^0s are very difficult to reject using one layer preradiator. But, again, the real QCD events contains fair amount of fake isolated γs with multiple γs, around 50% of background rejection in the high momentum range is possible.

The analysis in this note has been applied in the study of the Higgs signal [R-Y Zhu, CALT-68-1777] and the result shows that the irreducible background against the Higgs signal can be reduced to tolerable size, around $10 - 20\%$ of the irreducible background.

Acknowledgements

I would like to thank B. Barsih for his financial support and giving me the opotunity to work at Caltech, and R-Y Zhu for his continuos supports and discussions.

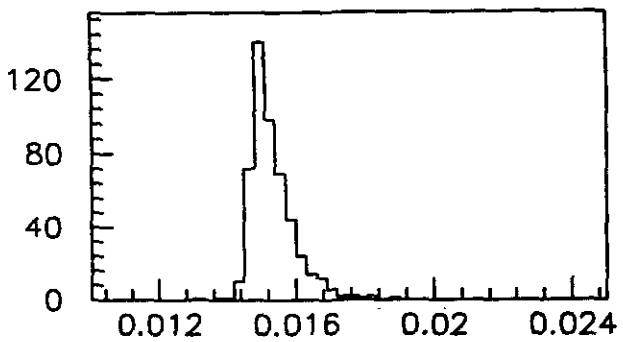

Fig.1-a 40 GeV γ

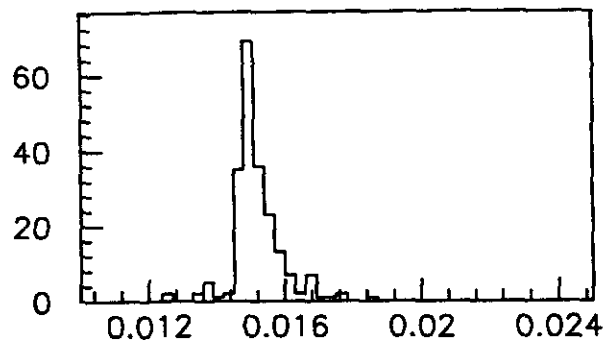

Fig.1-b real γ in γ+jet event

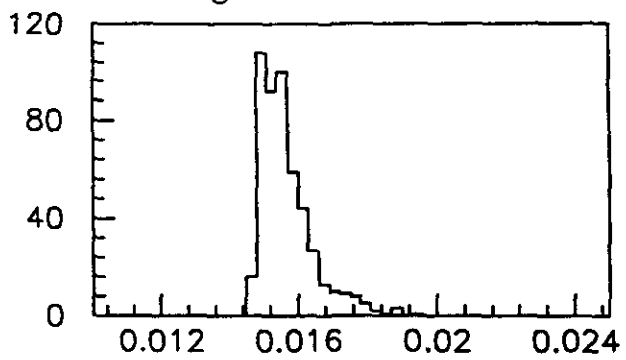

Fig.1-c 40 GeV π^0

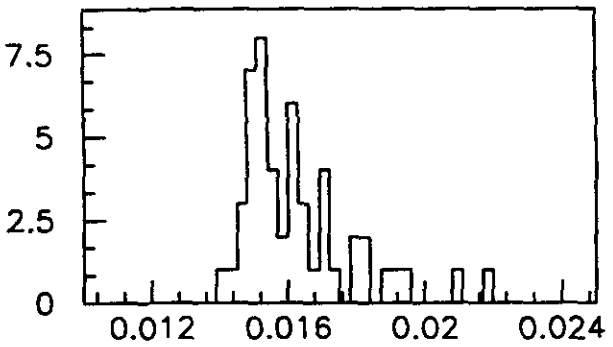

Fig.1-d fake isolated γ in γ+jet

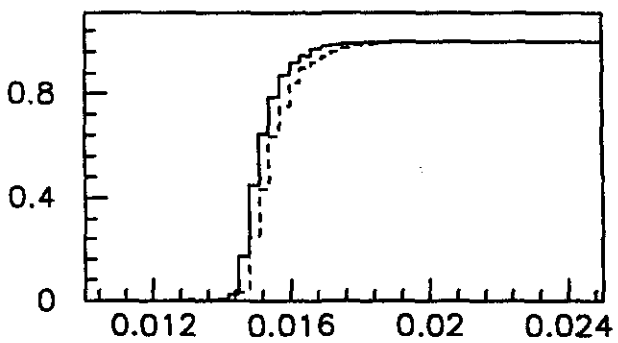

Fig.1-e Efficiency for γ(solid) and π^0(dotted)

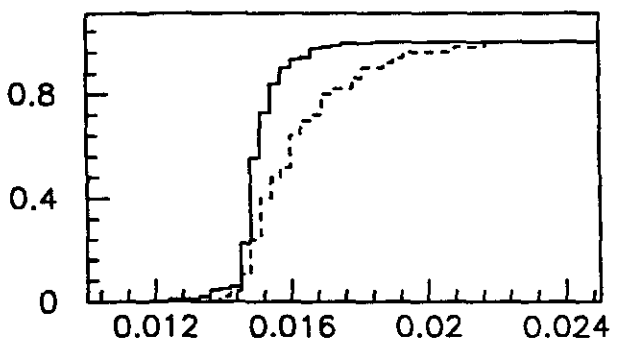

Fig.1-f Efficiency for single(solid) and fake γ in γ+jet events

Fig. 1 σ_{min} Distribution : single γ vs multi γs

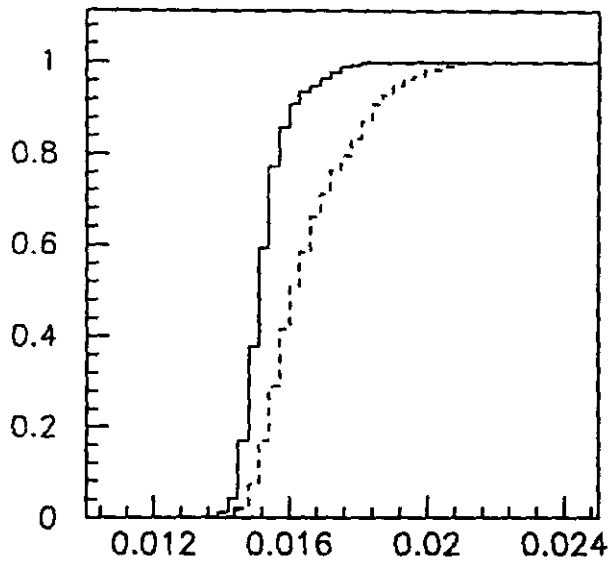

Fig.2-a 12 GeV γ(solid) vs π^0(dotted)

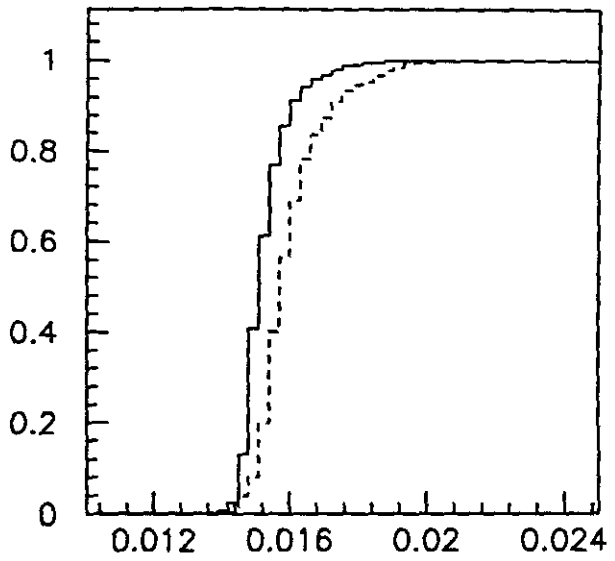

Fig.2-b 20 GeV γ(solid) vs π^0(dotted)

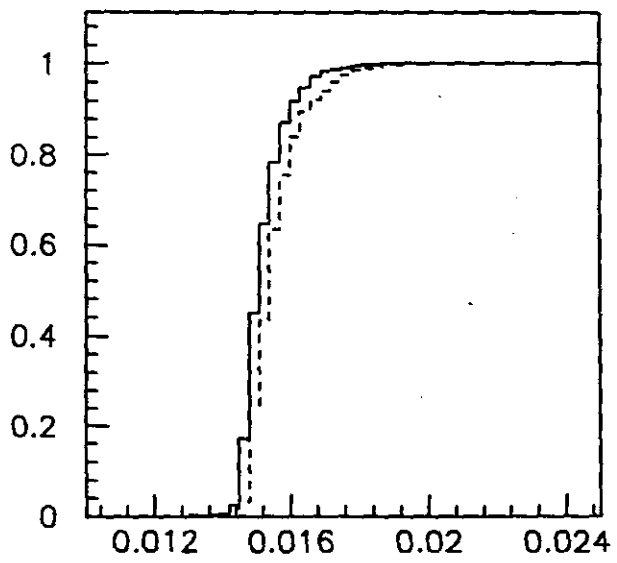

Fig.2-c 40 GeV γ(solid) vs π^0(dotted)

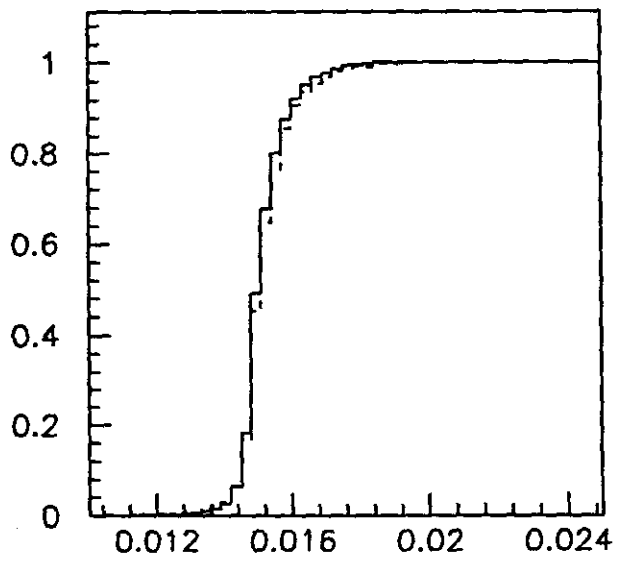

Fig.2-d 80 GeV γ(solid) vs π^0(dotted)

Fig. 2 γ and π^0 efficiencies

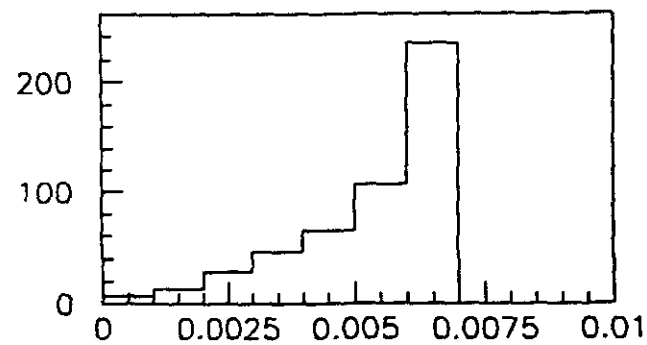

Fig.3-a $d\overline{\Theta}$ distribution 20 GeV π^0

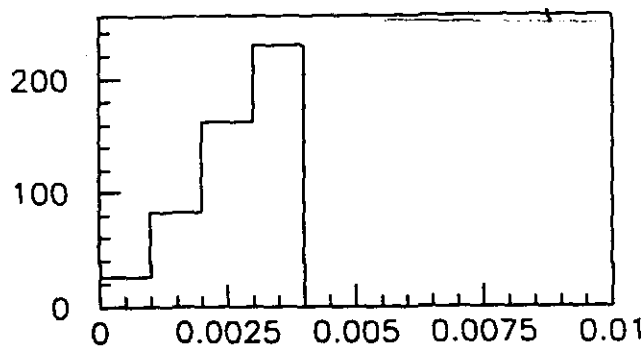

Fig.3-d $d\overline{\Theta}$ distribution 40 GeV π^0

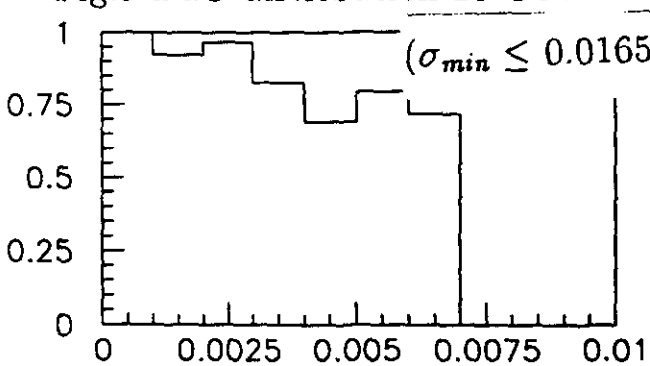

Fig.3-b Efficiency 20 GeV π^0

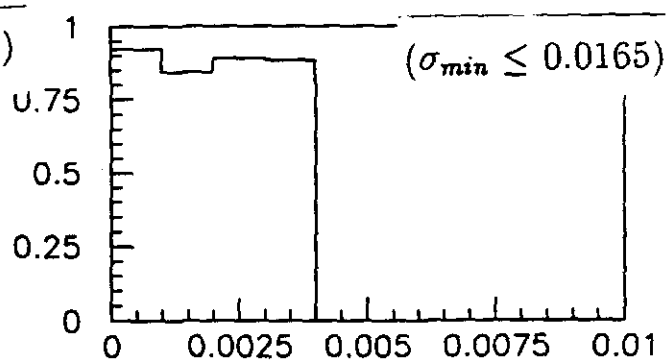

Fig.3-e Efficiency 40 GeV π^0

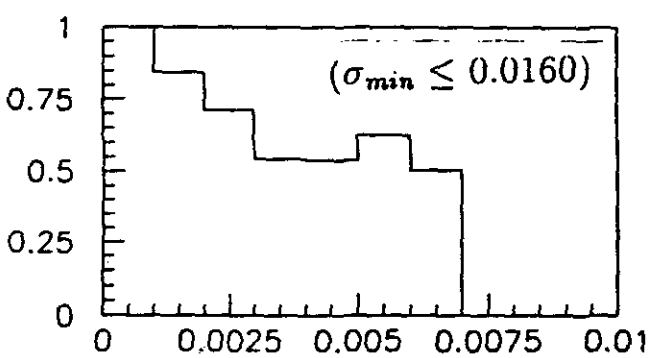

Fig.3-c Efficiency 20 GeV π^0

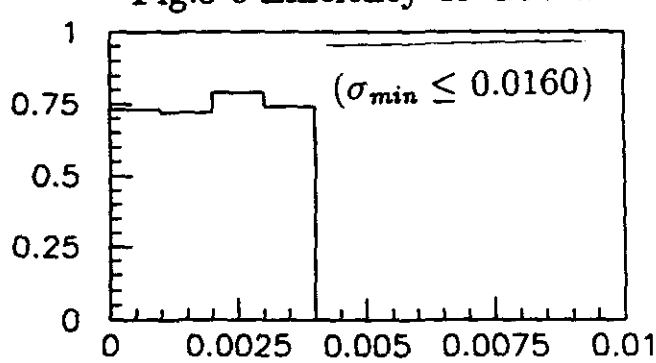

Fig.3-f Efficiency 40 GeV π^0

Fig. 3 $d\overline{\Theta}$ distribution of π^0

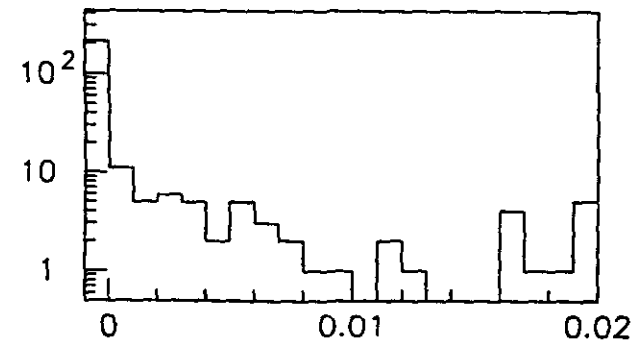

Fig.4-a Distribution in γ+jet events

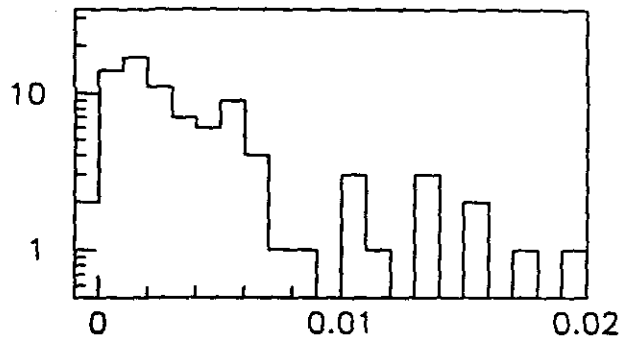

Fig.4-d Distribution in $t+\bar{t}$ events

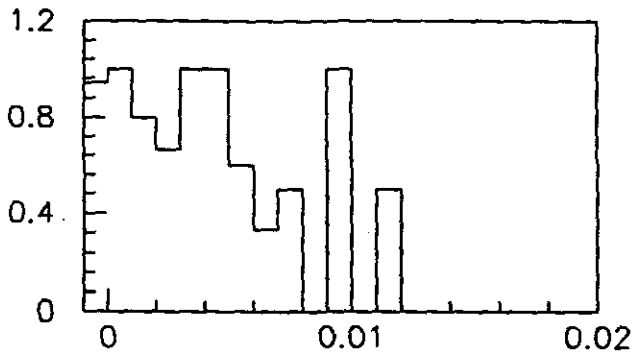

Fig.4-b Efficiency ($\sigma_{min} \leq 0.0165$)

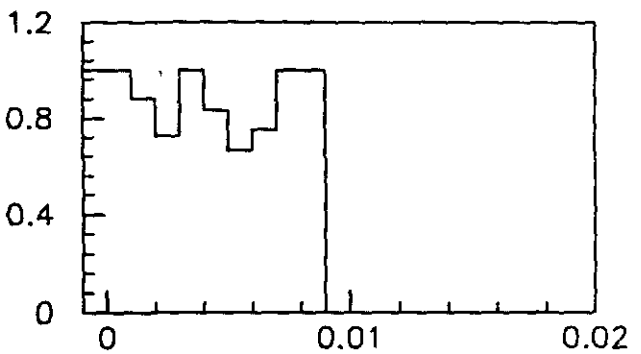

Fig.4-e Efficiency ($\sigma_{min} \leq 0.0165$)

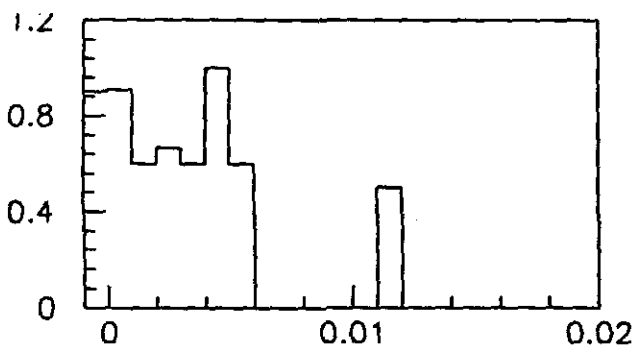

Fig.4-c Efficiency ($\sigma_{min} \leq 0.0160$)

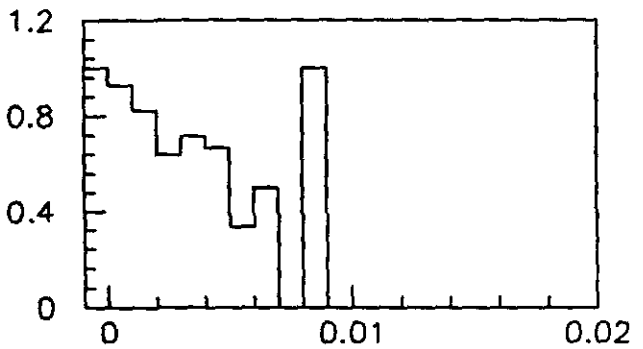

Fig.4-f Efficiency ($\sigma_{min} \leq 0.0160$)

Fig. 4 $d\overline{\Theta}$ distribution of fake isolated γs in QCD events

```
AAAAAA
 AAAAA
       BBBBBB
       BBBBB
```

Fig. 5 An example of cell clustering

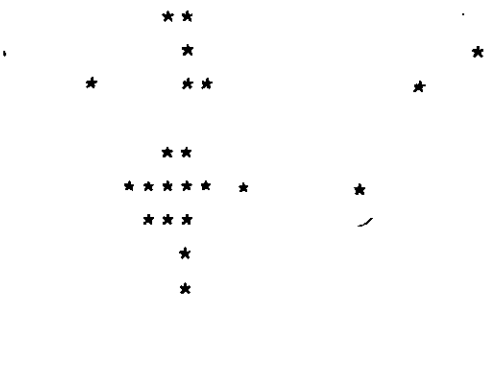

Fig.6-a 80 GeV γ with N = 2

Fig.6-b 80 GeV γ with N = 4

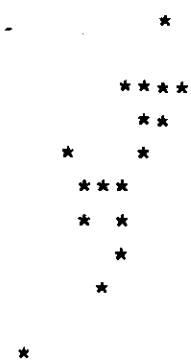

Fig.6-c 80 GeV π^0 with N = 2

Fig.6-d 80 GeV π^0 with N = 4

Fig. 6 Examples of active cell distribution of γ and π^0

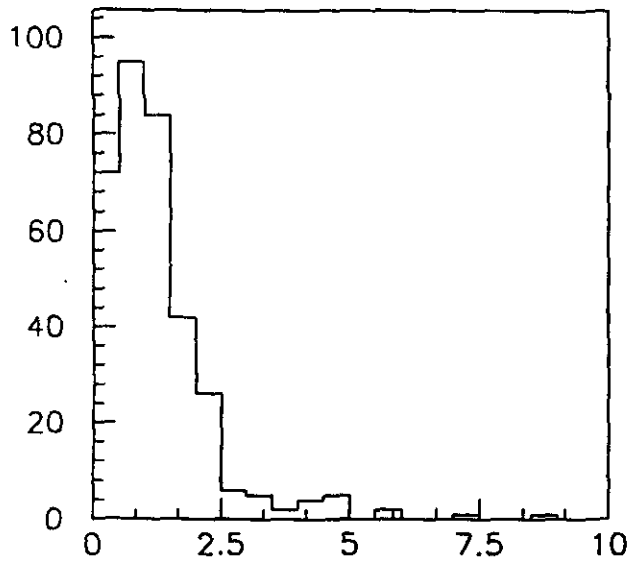

Fig.7-1-a 20 GeV γ N = 2

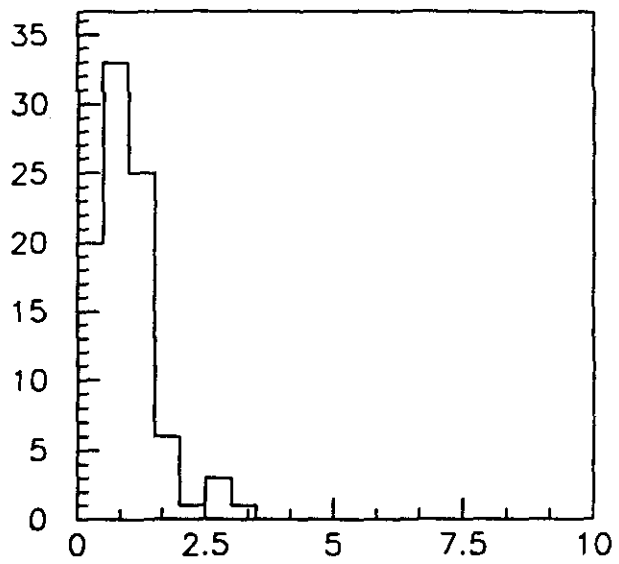

Fig.7-1-b 80 GeV γ N = 2

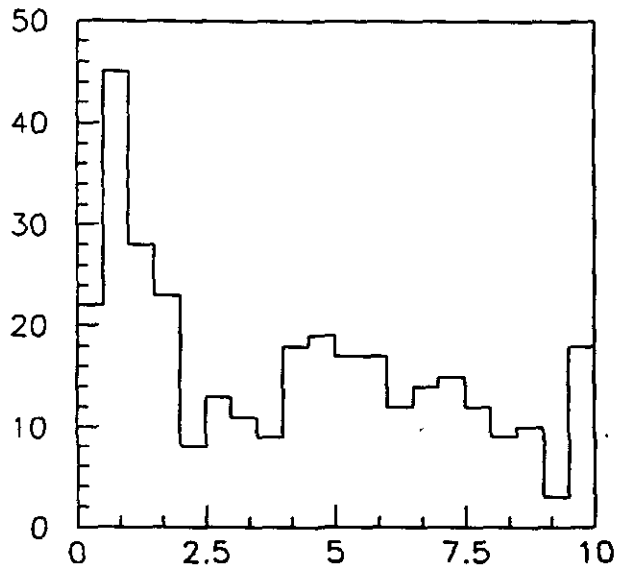

Fig.7-1-c 20 GeV π^0 N = 2

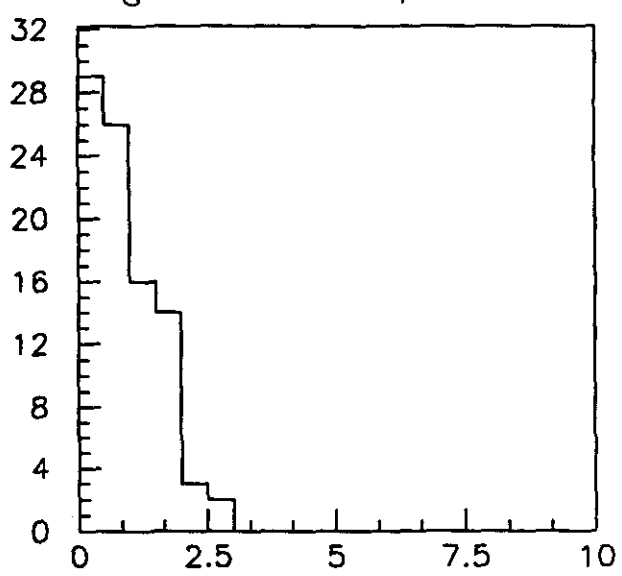

Fig.7-1-d 80 GeV π^0 N = 2

Fig. 7 Distance between the center of all active cells and the center of the largest cluster

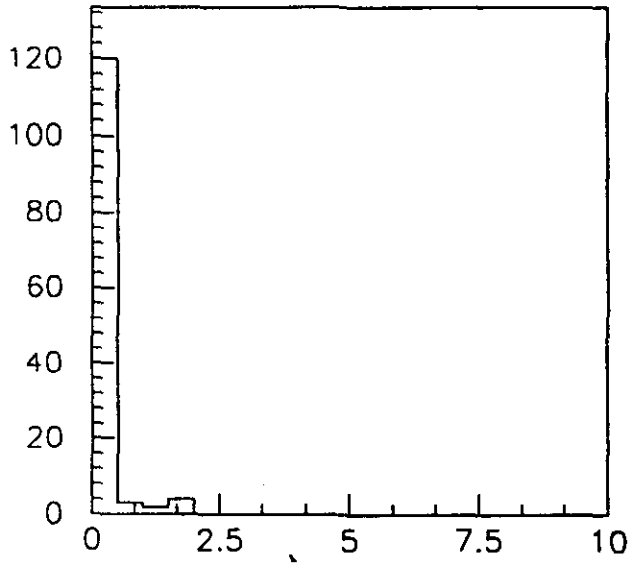

Fig.7-2-a 20 GeV γ N = 6

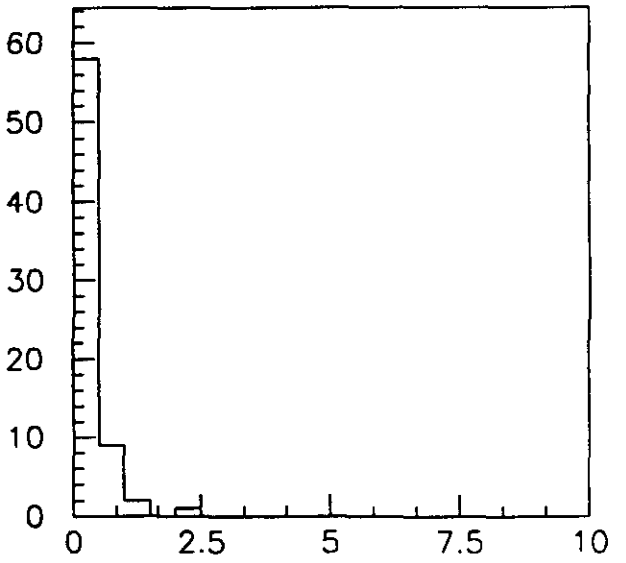

Fig.7-2-b 80 GeV γ N = 6

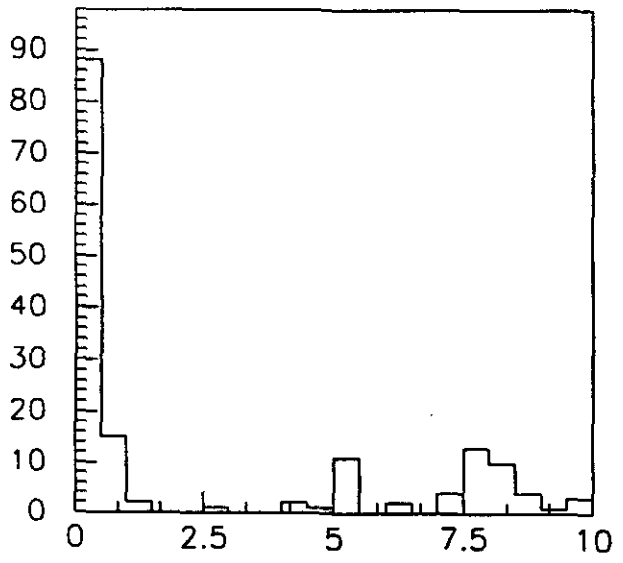

Fig.7-2-c 20 GeV π^0 N = 6

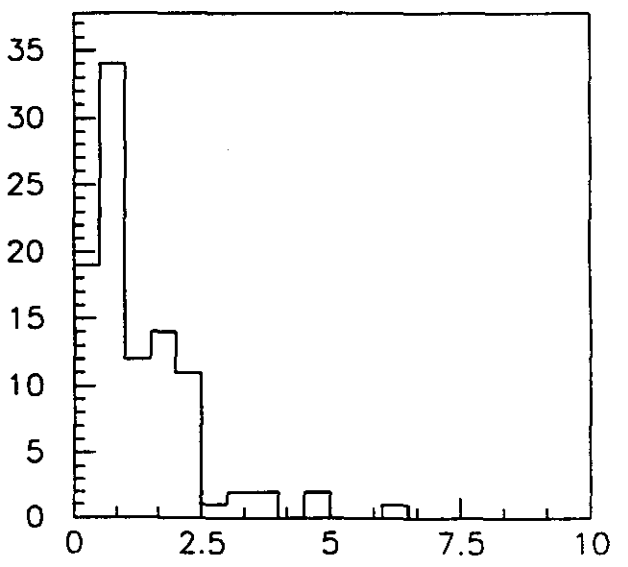

Fig.7-2-d 80 GeV π^0 N = 6

Fig. 7 Distance between the center of all active cells and the center of the largest cluster