LC-TOOL-2007-010
October 5, 2007

π^0 Reconstruction within the full simulation framework

Philippe Gris

Laboratoire de Physique Corpusculaire
IN2P3/CNRS
Clermont-Ferrand

Abstract

Neutral pions represent an important part of the visible energy of hadronic events and a proper π^0 reconstruction may be of interest for Particle-Flow and detector optimisation. This document describes studies related to neutral pions reconstruction within the full simulation of the LDC detector at ILC. After a careful calibration of the EM sub-detector, the impact of constrained fits on cluster pairs coming from single π^0 is evaluated. A method to reconstruct π^0's in physics events is then exposed. This method has been implemented in a C++ `MarlinReco` processor.

1 Introduction

Neutral pions represent an important part of the particle content in hadronic events. In $t\bar{t}$ or hZ events at $\sqrt{s}$=500 GeV, around 20% of the visible energy is due to π^0's and most of the photons in an event come from their decays. Owing to the low energy spectrum of these photons (about 55% of the photons in $hZ \rightarrow b\bar{b}\nu\bar{\nu}$ at $\sqrt{s}$=500 GeV have an energy lower than one GeV), the influence of π^0's on Particle-Flow and on detector optimisation may have to be considered. τ-lepton identification may also benefit from proper π^0 reconstruction in particular in the hadronic decay modes. All these arguments lead to the conclusion that it may be important not only to reconstruct and identify photons but also to go further and to identify π^0's . The strategy may be staged in three steps: photon reconstruction, cluster pairing to form π^0 candidates and then perform constrained fits to recover from mismeasurements and fluctuations.

Studies aiming at reconstructing π^0's within the full simulation framework are exposed in this note. The software context is developped in a first part. Details on the calibration of the electromagnetic (EM) calorimeter are given on a second part. The constrained fit method applied on single π^0's is described in a third section. An approach to reconstruct π^0's in physics events is exposed in a fourth part. The method developped has been implemented in a C++ based class. The main feature of this processor are precised in a fifth part before concluding.

2 Framework of the study

The steps of the full simulation chain were the following: event generation has been done with `Pythia v6.321`[1]; the `GEANT4` based full Monte Carlo simulation has been performed with `Mokka v06.02`[2] and the reconstruction with `Marlin v00-09-05` [3] and `MarlinReco v00-02` [4]. Analyses were performed with `Root v5.10.00` [6].

The detector model used is LDC00[2]. The present study relies mostly on the electromagnetic calorimeter whose main characteristics are given in table 1.

	Barrel	Endcap
Symmetry	8	2
R_{inner}(mm)	1704	299 (R_{out}=1899)
outer z or inner z (mm)	2729	2829
Si-W	30 layers 3.9 mm (absorber: 1.4 mm) 10 layers 6.7 mm (absorber: 4.2 mm) 24 X_0	

Table 1: Main characteristics of the LDC00 EM calorimeter as given by `Mokka`

The studies were perfomed with events that were comprised of single monoenergetic particles or of more realistic configurations: single π^0 , single photon or $hZ \rightarrow b\bar{b}\nu\bar{\nu}$ at $\sqrt{s}$=500 GeV were considered.

3 EM calorimeter calibration

The EM calorimeter has to be calibrated before addressing the π^0's reconstruction.

3.1 The calibration method

Single photons generated with `Mokka` and reconstructed with `Marlin` were used to calibrate the EM calorimeter. A wide energy spectrum was considered: 0.25, 0.30, 0.35, 0.40, 0.50, 1, 2, 4, 10,

25, 50 GeV. For each energy, the angular coverage was: $0 \leq \theta \leq \pi/2$ (step: 0.1) and $0 \leq \Phi \leq 2\pi$ (step: $2\pi/16$ to take the detector geometry into account).
The goal of the calibration procedure was to evaluate α_0 and β_0 in the following formula:

$$E_\gamma^{clus} = \alpha_0(E_1^{30} + \beta_0 E_{31}^{40}) \quad (1)$$

where E_i^j is the sum of the raw energies deposited in the silicon cells of the i-to-j layers. All the hits were used (no clustering algorithm applied).
The β_0 parameter is firstly evaluated by minimizing $\sigma(E_1^{30} + \beta E_{31}^{40})/ < E_1^{30} + \beta E_{31}^{40} >$ where the mean and sigma come from a fit by a gaussian distribution (see figure 1). α_0 is then given by the mean value of $E_\gamma^{truth}/(E_1^{30} + \beta_0 E_{31}^{40})$ where E_γ^{truth} is the Monte-Carlo truth energy of the generated photon.

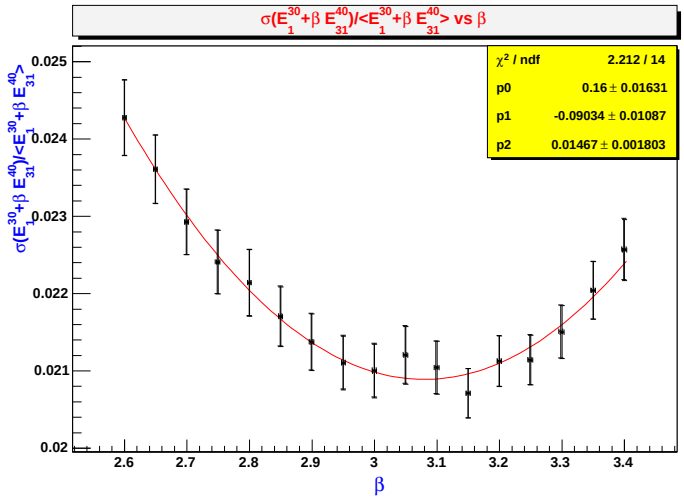

Figure 1: $\sigma(E_1^{30} + \beta E_{31}^{40})/ < E_1^{30} + \beta E_{31}^{40} >$ as a function of β for E_γ^{truth}=50 GeV and θ=0.2 . For each value of the βparameter, the distribution $E_1^{30} + \beta E_{31}^{40}$ is fitted by a gaussian.

β_0 values obtained with this procedure are given on figure 2 for E_γ^{truth}=50 GeV and for different values of θ. A good agreement between central and endcap values is observed. A linear fit gives $\beta_0^{fit} = 3.0 \pm 0.1$. α_0 is then taken as being equal to the mean value of $E_\gamma^{truth}/(E_1^{30} + \beta_0^{fit} E_{31}^{40})$. The result is given on figure 3 for E_γ^{truth}=10 GeV. The fit values are α_0^{fit}=27.62±0.03 and α_0^{fit}=28.83±0.05 for the central and the endcap respectively.

3.2 Energy and Angle Resolution

Once the EM calorimeter is calibrated, the performance in terms of energy and angular resolution as well as on linearity may be evaluated. All the hits coming from the monoenergetic photons generated in `Mokka` were considered to form a cluster. The energy was obtained from relation (1) using fitted values α_0^{fit} and β_0^{fit}. The cluster energy was then fitted by a gaussian distribution so as to estimate the energy resolution:

$$\frac{\sigma(E)}{E} = C \oplus \frac{S}{\sqrt{E}} \quad (2)$$

where C and S are the constant term and sampling terms respectively; E and $\sigma(E)$ are the gaussian fit values. The linearity is given by the comparison between E and E_γ^{truth}.

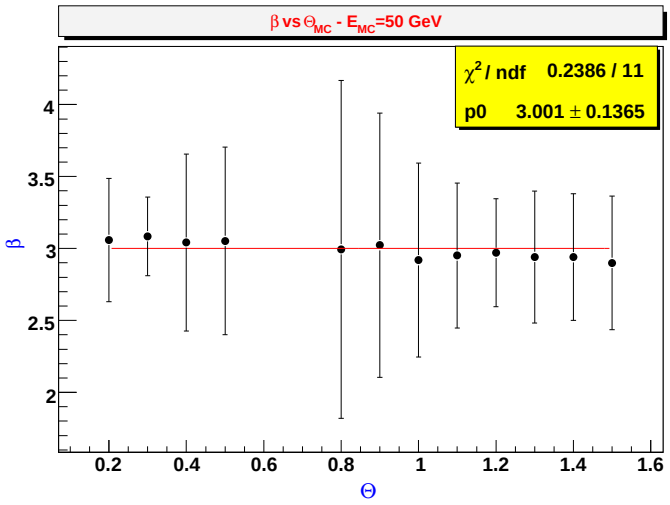

Figure 2: β_0 as a function of θ for E_γ^{truth}=50 GeV.

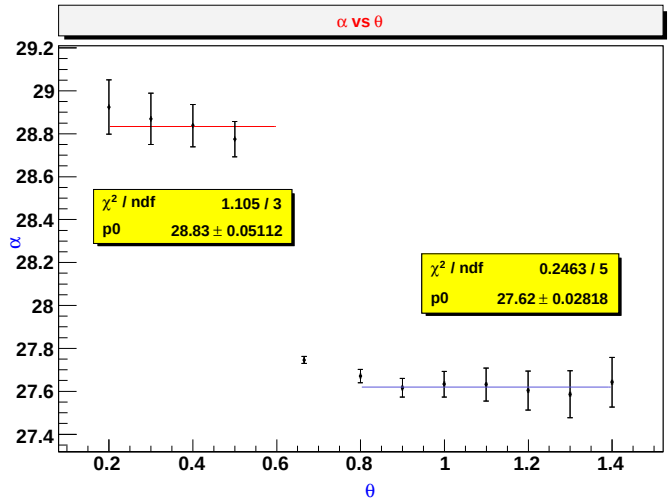

Figure 3: α_0 as a function of θ for E_γ^{truth}=10 GeV. Error bars were obtained from a one σ variation of β_0

Energy resolution and linearity plots are given on figure 4 for the barrel(left) and endcap(right) EM calorimeters. The sampling term is around 12% and the constant term lies between one and two percent. The non-linearity is rather low: it does not exceed 3% (5%) at very low energies (ie around 250 MeV) for the barrel (endcap) part.

Angular resolutions were estimated in the same way: for a given energy and a given (θ,Φ), reconstructed θ and ϕ were fitted by gaussian distributions and the results obtained are shown on figure 5. σ_θ depends on θ and on E whereas σ_Φ depends only on Φ. Two regimes can be indentified for θ and ϕ resolution depending on the value of the energy (bigger or lower than 1.6 GeV): the slope of the resolution is higher at low energies because of the structure of the

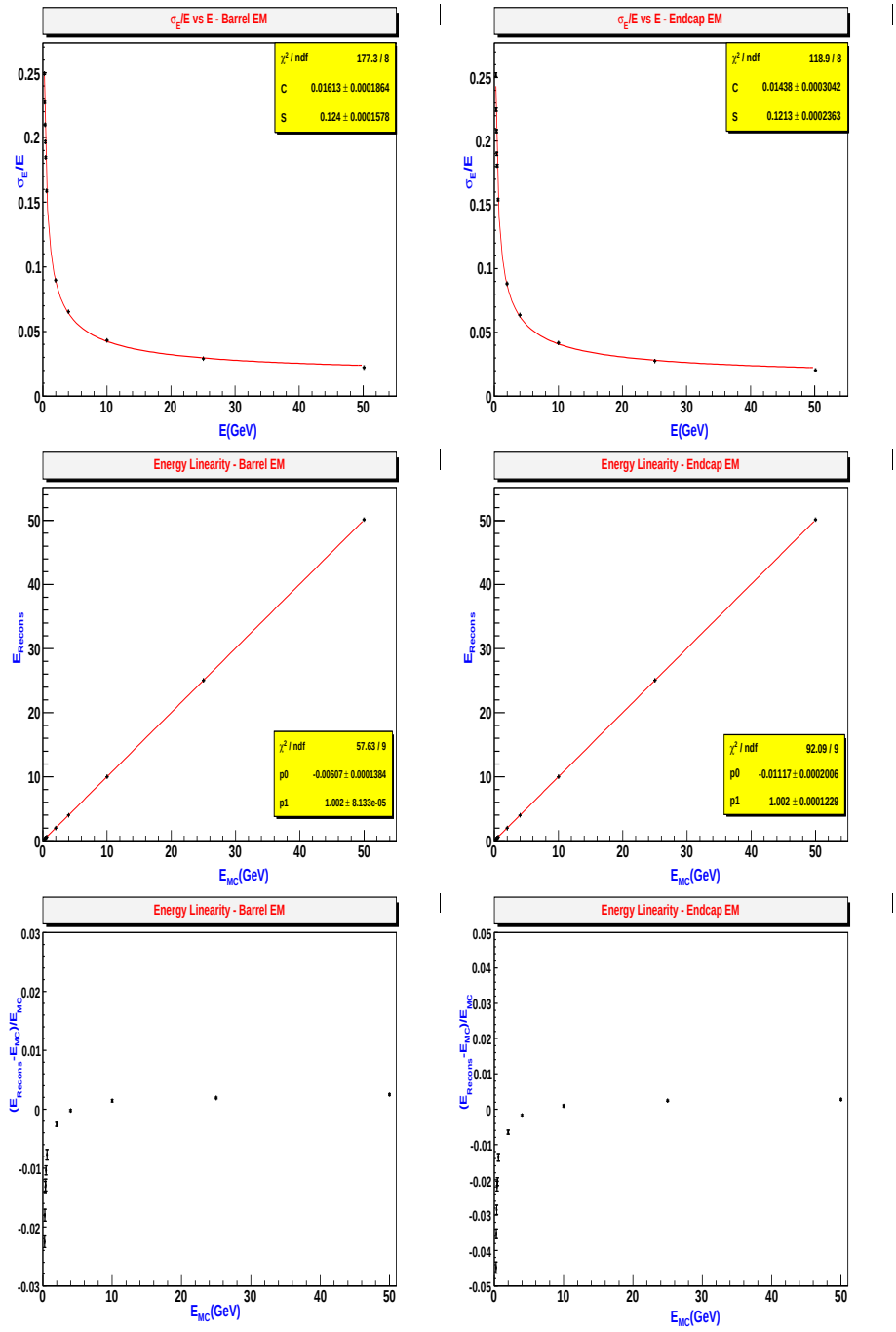

Figure 4: Energy resolution and linearity plots for the barrel (left) and endcap (right) EM calorimeter after calibration: σ_E/E vs E (top), E_{recons} vs E_{MC} (middle) (E_{recons}-E_{MC})/E_{MC} vs E_{MC} (bottom) are displayed.

calorimeter (the last ten layers are three times wider than the first thirties). Nonetheless the fluctuations are higher at low energies and the angular resolution is then worse. σ_θ does not exceed 2.2 mrad (1.4 mrad) and σ_Φ is lower than 2.5 mrad (4 mrad) for the barrel (endcap) part of the EM calorimeter.

The energy and angular resolutions obtained are summarized in table 2.

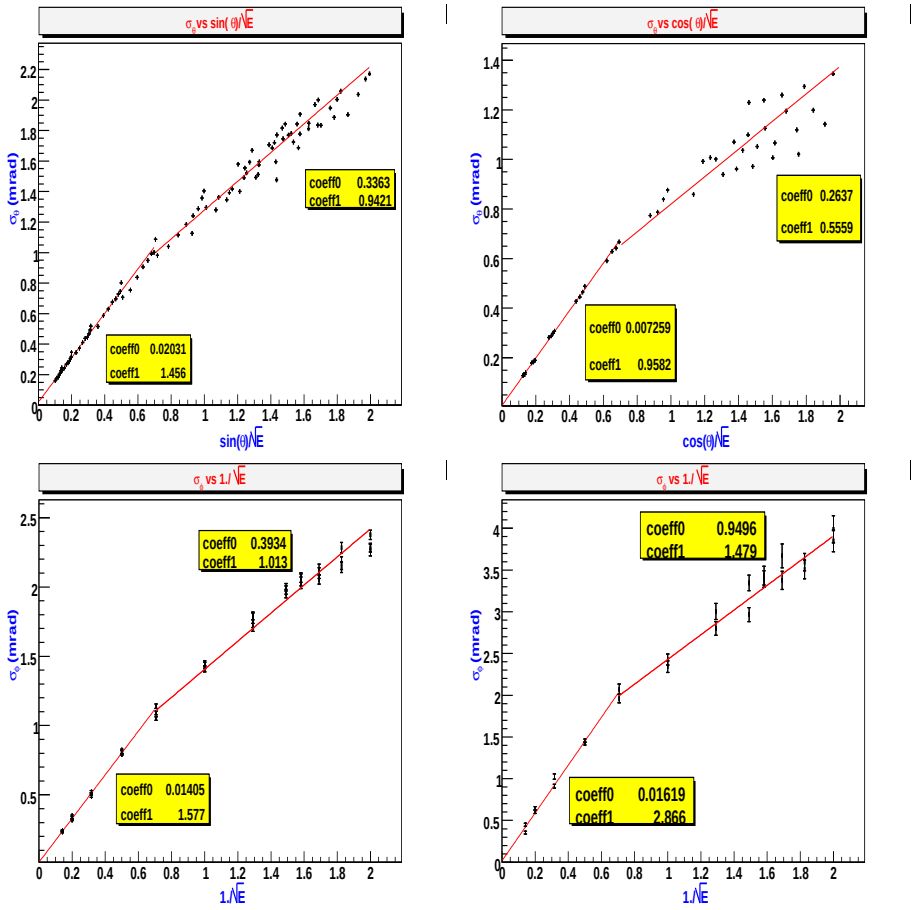

Figure 5: Angular resolution for the barrel (left) and endcap (right) EM calorimeters: σ_θ vs $sin\theta/\sqrt{E}$ (top left), σ_Φ vs $1/\sqrt{E}$ (bottom left), σ_θ vs $cos\theta/\sqrt{E}$ (top right) and σ_Φ vs $1/\sqrt{E}$ (bottom right) are displayed

	Barrel		Endcap	
Energy resolution	$\frac{\sigma_E}{E} = 1.6\% \oplus \frac{12.4\%}{\sqrt{E}}$		$\frac{\sigma_E}{E} = 1.4\% \oplus \frac{12.1\%}{\sqrt{E}}$	
Non linearity	$< 2.5\%$		$< 5\%$	
	E<1.5 GeV	E>1.5 GeV	E<1.5 GeV	E>1.5 GeV
θ resolution (mrad)	$\sigma_\theta = 0.34 + \frac{0.94 sin\theta}{\sqrt{E}}$	$\sigma_\theta = 0.02 + \frac{1.4 sin\theta}{\sqrt{E}}$	$\sigma_\theta = 0.26 + \frac{0.55 cos\theta}{\sqrt{E}}$	$\sigma_\theta = 0.007 + \frac{0.96 cos\theta}{\sqrt{E}}$
Φ resolution (mrad)	$\sigma_\Phi = 0.39 + \frac{1.01}{\sqrt{E}}$	$\sigma_\Phi = 0.01 + \frac{1.56}{\sqrt{E}}$	$\sigma_\Phi = 0.94 + \frac{1.5}{\sqrt{E}}$	$\sigma_\Phi = 0.02 + \frac{2.9}{\sqrt{E}}$

Table 2: LDC00 EM calorimeter resolutions after calibration

4 π^0 reconstruction

To study neutral pion reconstruction within the full simulation framework, single π^0 have been generated with `Mokka` at the following energies: 0.4, 0.6, 0.7, 1.2, 4.9, 12.4, 29.5 GeV and for similar values in (θ, Φ) as indicated in part 3. A clustering has been performed with the `TrackWiseClustering` processor [5]. Events with exactly two clusters were selected. The reconstructed invariant mass evaluated from the two clusters is shown on figure 6. It has to be noticed that the peak width depends on the π^0 energy.

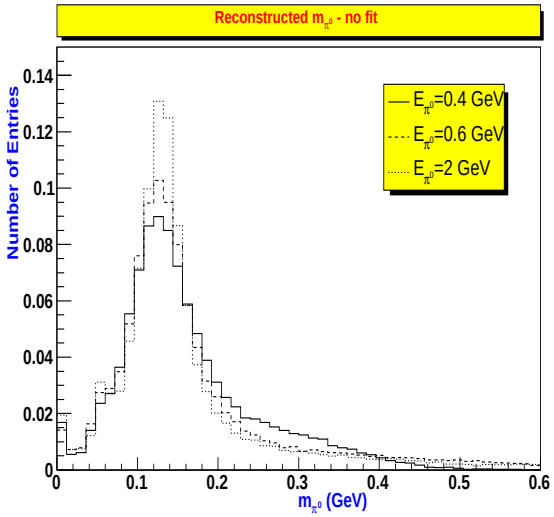

Figure 6: Reconstructed dicluster mass before fit for π^0 energies of 0.4, 0.6 and 2.4 GeV

To correct for mismeasurements and fluctuations, a constrained fit has been applied on cluster pairs with the minimization of the following χ^2:

$$\chi^2 = \sum_{i=1}^{2} \frac{(E_i^{cl} - E_i)^2}{\sigma_{E_i}^2} + \sum_{i=1}^{2} \frac{(\theta_i^{cl} - \theta_i)^2}{\sigma_{\theta_i}^2} + \sum_{i=1}^{2} \frac{(\Phi_i^{cl} - \Phi_i)^2}{\sigma_{\Phi_i}^2} \tag{3}$$

where $E_i^{cl}, \theta_i^{cl}, \Phi_i^{cl}$ are the energy and angles of the i-th cluster. E_i, θ_i, Φ_i are the corrected energy and angles coming out of the fit procedure. The values of the variances $\sigma_{E_i}^2, \sigma_{\theta_i}^2, \sigma_{\Phi_i}^2$ were set to the estimation of table 2. Only events (so-called fitted events hereafter) fulfilling the following conditions were selected:

- the fit has converged;
- the number of iteration of the fit, `niter` , is less or equal than 4;
- the χ^2 of the fit is less or equal than 3.9.

The impact of the fit on the energy resolution is shown on figure 7 where the relative energy difference (wrt MC truth π^0 energy) as a function of the MC truth π^0 energy is displayed. The gain is spectacular at very low energies (few hundred MeV) where the resolution is ten times better after the fit. The resolution is furthermore quite stable after the fit and of the order of 3%.

The fit efficiency (defined as the fraction of fitted events) shows a dependence on the reconstructed angle between the two clusters (see figure 8 left). This is due to the fact that the

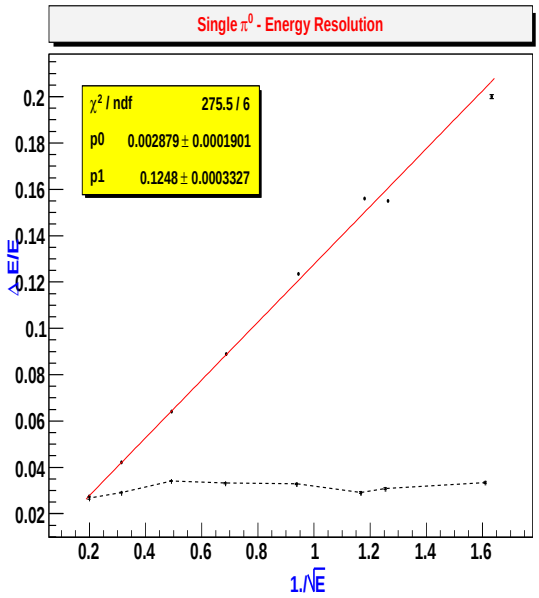

Figure 7: Energy resolution $\Delta E/E$ as a function of $1./\sqrt{E}$ for single π^0 before (full line) and after (dashed line) application of a constrained fit on the cluster pairs.

relative energy difference between the two photons increases with the opening angle between the two photons as it is shown on figure 8 (right) where the energy difference between the two photons divided by the energy sum (which is the π^0 energy) is displayed as a function of the opening angle between the two photons (MC-truth information were used). As a result of the Lorentz boost, the energy difference increases with the opening angle. This leads to two energy-imbalanced clusters at high angle. The lower energy cluster may then be subject to fluctuations and mismeasurements. This may explain why the fit efficiency decreases at high opening angle.

To use a more realistic π^0 energy spectrum, we selected single π^0 coming from $hZ \rightarrow b\bar{b}\nu\bar{\nu}$ events. Events were passed through the full simulation and a constrained fit was applied on events with two clusters (processor used for clustering: `TrackWiseClustering`). The event fraction having 0, 1, 2 and more than 3 clusters is 0.3%, 14.3%, 73.2% and 12.2% respectively. 87.4% of the two-cluster events were successfully fitted. The impact of the fit on the energy resolution of the π^0 is shown on figure 9: the fit reduces the energy uncertainty by more than a factor of two namely from 7.4% to 3.4%.

The two-cluster events that did not fit successfully (12.6% out of the total number of two-cluster events) were studied more precisely. For 4.1% of them, the fit did not converge. When the fit has converged, there are 70% of the events have a too high χ^2, 23% have more than four iterations, and 7% have both a too high χ^2 and too many iterations. These events have two close clusters (figure 10) typically separated by less than few Moliere radii. The clustering algorithm may be confused in that case and these events could be recovered thanks to a dedicated study.

5 Strategy to reconstruct π^0 in physics events

The number of π^0's produced in a given physics event can be quite large: for $hZ \rightarrow b\bar{b}\nu\bar{\nu}$ events at $\sqrt{s}$=500 GeV, 12 to 13 π^0 on average are produced leading to more than 20 clusters in the EM calorimeter. Cluster pairs have to be formed so as to reconstruct π^0's . A strategy is thus

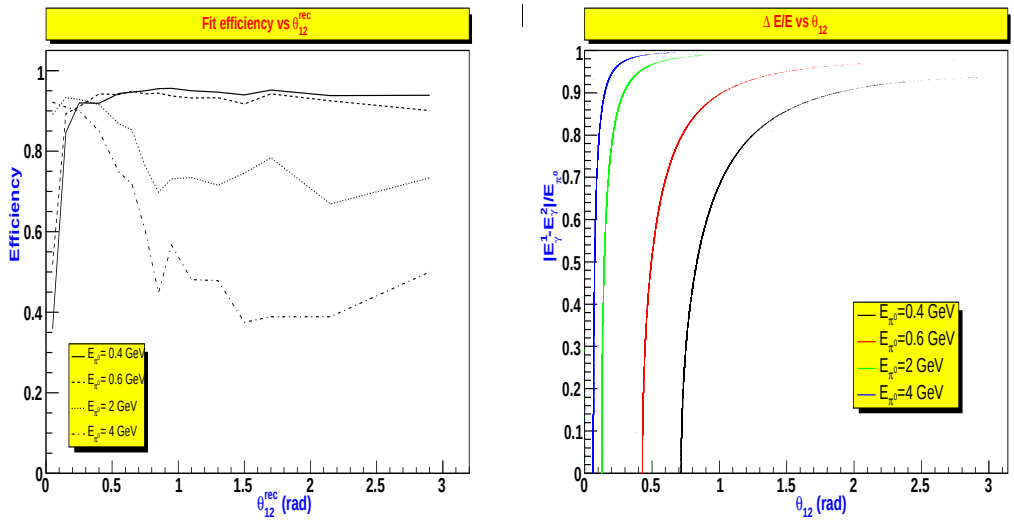

Figure 8: Fit efficiency as a function of the reconstructed angle between the two clusters (left) and relative energy difference between the two photons coming from a π^0 decay as a function of the angle between the two photons (right) for π^0 energies of 0.4, 0.6, 2 and 4 GeV.

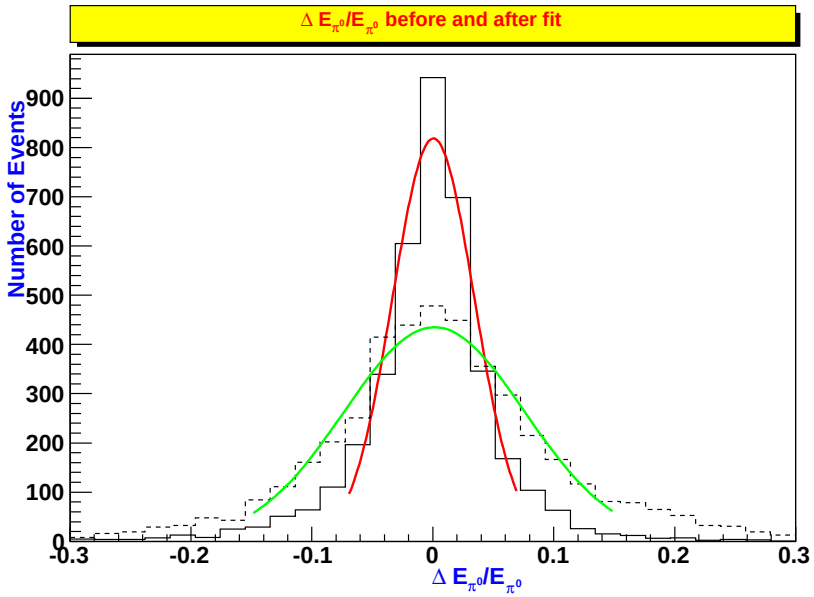

Figure 9: $\Delta E_{\pi^0}/E_{\pi^0}$ distribution for single π^0 coming from $hZ \to b\bar{b}\nu\bar{\nu}$ events without any fit (dashed histogram) and with the constrained fit applied (solid histogram). The width is divided by more than two after the fit (7.4% to 3.4%)

needed to perform a correct pairing (ie association of clusters coming from the same π^0) and to minimize fake pair production.

A study has been performed using MC-truth information of photons coming from π^0 decays

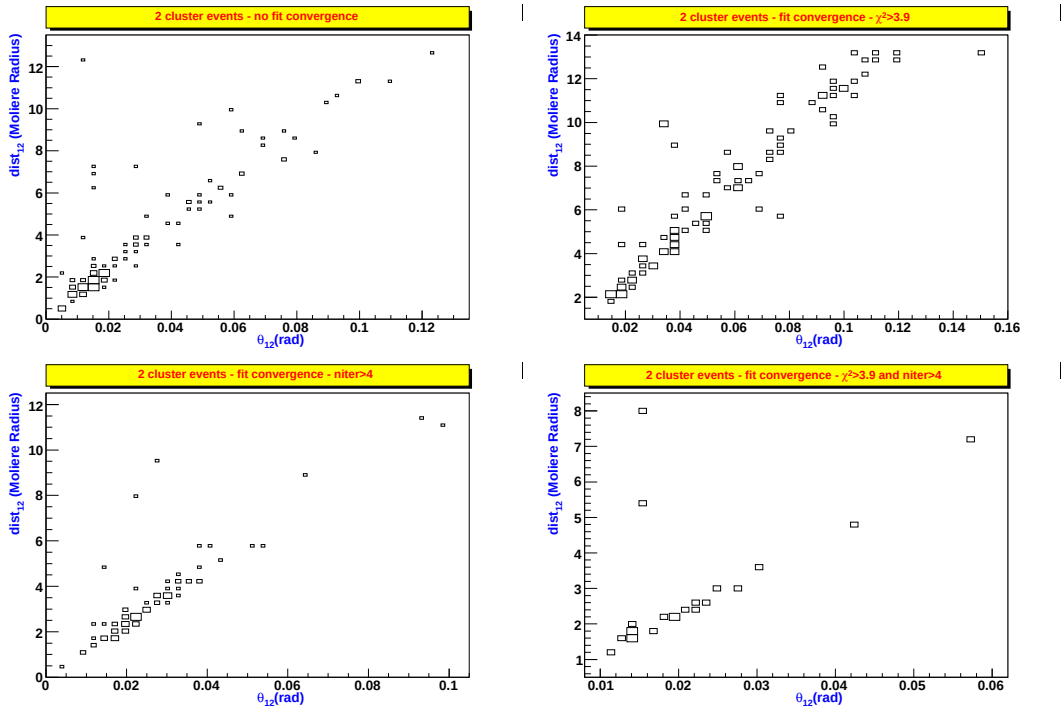

Figure 10: Two-cluster events from single π^0 coming from $hZ \to b\bar{b}\nu\bar{\nu}$ events that did not fit successfully. The plots show the distance at the EM calorimeter entrance of two clusters (in Moliere radius: R_M=15 mm) as a function of the angle between the two clusters. The four displays are : events for which the fit did not converged (top left), events for which the fit did converge but with χ^2 >3.9 (top right), events for which the fit did converge but with niter>4 (bottom left) and events for which the fit did converge but with χ^2 >3.9 and niter>4 (bottom right)

in $hZ \to b\bar{b}\nu\bar{\nu}$ events at $\sqrt{s}$=500 GeV and $hhZ \to b\bar{b}b\bar{b}\nu\bar{\nu}$ at $\sqrt{s}$=800 GeV. Energies and angles of the photons were smeared according to the resolutions given in table 2. The "clusters" (here defined as the smeared MC-truth photons) were then associated in the following way: starting with the highest energy objects, a probability depending on the mass of the pair, and/or the angle between the two objects (these two quantities depend on the energy of the pair) has been estimated. For a given cluster, many pairs may be formed (with all other clusters having a smaller energy than the considered cluster). The pair giving giving the highest probability is selected and the corresponding clusters are removed from the cluster list. The probability for the mass was taken from a gaussian distribution. The variances used are shown in the figure 11. The angle distribution was reconstructed from a toy MC. To quantify the quality of the procedure, we have used the ratio of the energy sum of pairs coming from π^0's (which we know from MC) divided by the total π^0 energy of the event.

The figure 12 shows clearly the importance of the reconstructed mass in our estimator. When the angle is also used, a slight improvement is observed: the mean value of the $E_{\pi^0}(goodpairs)/E_{\pi^0}(all)$ is 75% (rms:23%) when the mass is used. It increases to 78% (rms: 21%) when the mass and the angle information are combined. A fit may then be performed on pairs and the impact on the total EM energy evaluated. An illustration is given on the figure 13 where the RMS of the relative energy of the total π^0 energy is displayed as a function of the total π^0 MC energy is

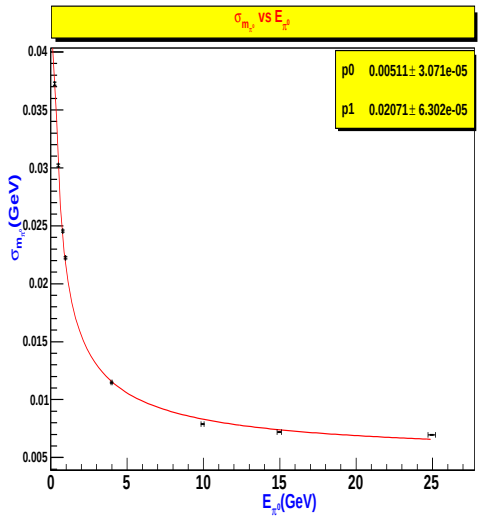

Figure 11: $\sigma_{m_{\pi^0}}$ vs E_{π^0} for smeared photons

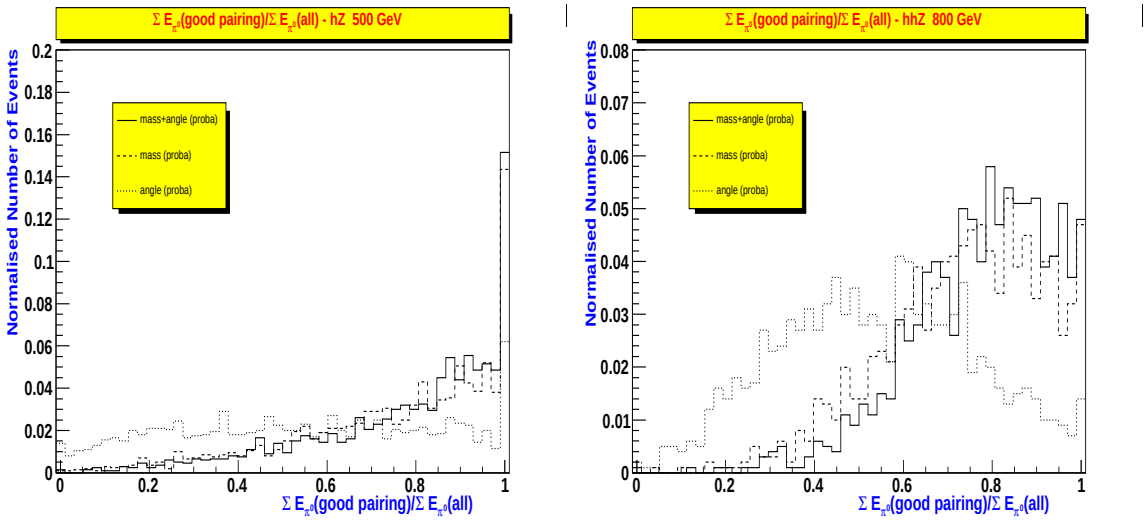

Figure 12: Distributions of the fraction of good (ie coming from π^0's decays) π^0 energy for hZ $\rightarrow$ b$\bar{\text{b}}\nu\bar{\nu}$ evsnta at $\sqrt{\text{s}}$=500 GeV and hhZ $\rightarrow$ b$\bar{\text{b}}$b$\bar{\text{b}}\nu\bar{\nu}$ events at $\sqrt{\text{s}}$=800 GeV. The resultas are given for three estimators using the mass, the opening angle and the mass+the opening angle of the cluster pairs.

displayed for three cases: when no fit is applied, when a fit is applied on clusters paired thanks to an estimator using the mass and the angle, and when a fit is applied on clusters coming from the same π^0 (no fakes at all). When the fit has not converged, the initial reconstructed energy was used. The gain of the fit is observed mainly at low energies. But it is also in this range that the fake rate seems to be the highest.

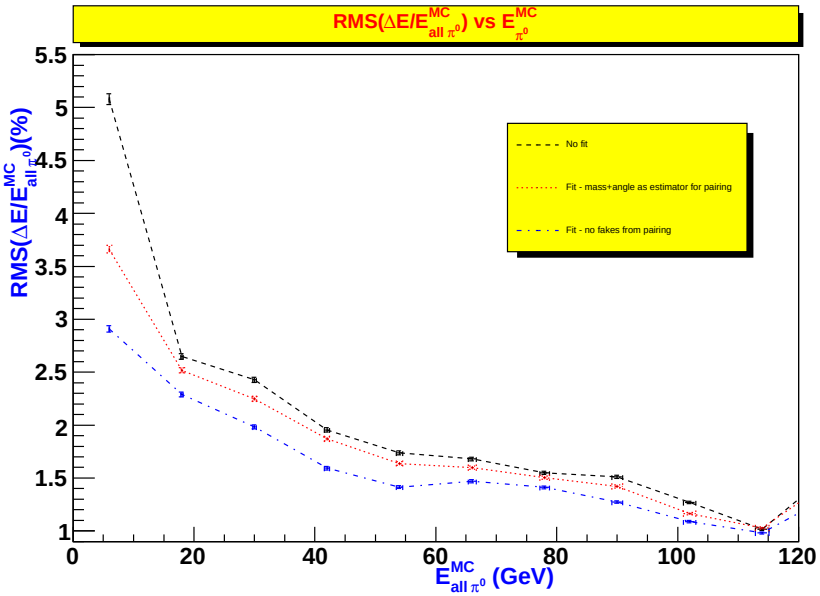

Figure 13: $RMS((E_{\pi^0} - E^{MC}_{\pi^0})/E^{MC}_{\pi^0}$ vs $E^{MC}_{\pi^0}$ for π^0's coming from hZ $\rightarrow$ b$\bar{\text{b}}\nu\bar{\nu}$ events at $\sqrt{\text{s}}$=500 GeV for three cases: without any fit (dashed line), with a fit of the clusters paired with a mass+angle estimator (dotted line) and with a fit of cluster pairs coming from π^0's (dashed-dotted line). When the fit has not converged, the initial reconstructed energy was used.

6 `Pi0_Sniffer`: a processor to reconstruct π^0's within the full simulation framework

The method developped in part 5 has been implemented in a C++ based class. This processor, called `Pi0_Sniffer`, reconstructs π^0's within the full simulation framework. Its schematic diagram is shown on figure 14. The input is a collection of Clusters of the LCIO framework[7]. Pairs of clusters are then formed following the approach exposed in part 5. The choosen probability may depend on the dicluster mass, the dicluster opening angle, or on both (user's choice). A constrained fit is then applied on each pair and a ReconstructedParticle[7] is created. Its cluster components are made of the original clusters. If the fit has converged, the ReconstructedParticle four-vector is the sum of the fitted clusters four-vectors. The number of iterations and the χ^2 of the fit are stored in ParticleID. If the fit has not converged, the ReconstructedParticle four-vector is the sum of the original clusters four-vectors and the first word in ParticleID is set to -1. The output of the processor is a collection of ReconstructedParticles that may be used by other processors.

7 Conclusion and outlook

π^0's are a non negligible part of the particle content in physics events produced at ILC and it is important to reconstruct them. A study on various aspects of π^0 reconstruction was presented in this document. A precise calibration of the EM calorimeter has been performed. It was then shown that constrained fits on single π^0 greatly improves the energy resolution (from 7.4% to 3.4% for a π^0 spectrum coming from hZ $\rightarrow$ b$\bar{\text{b}}\nu\bar{\nu}$ events at $\sqrt{\text{s}}$=500 GeV) in particular at low energies (the resolution is ten times better for π^0 of few hundred MeV). A strategy to reconstruct

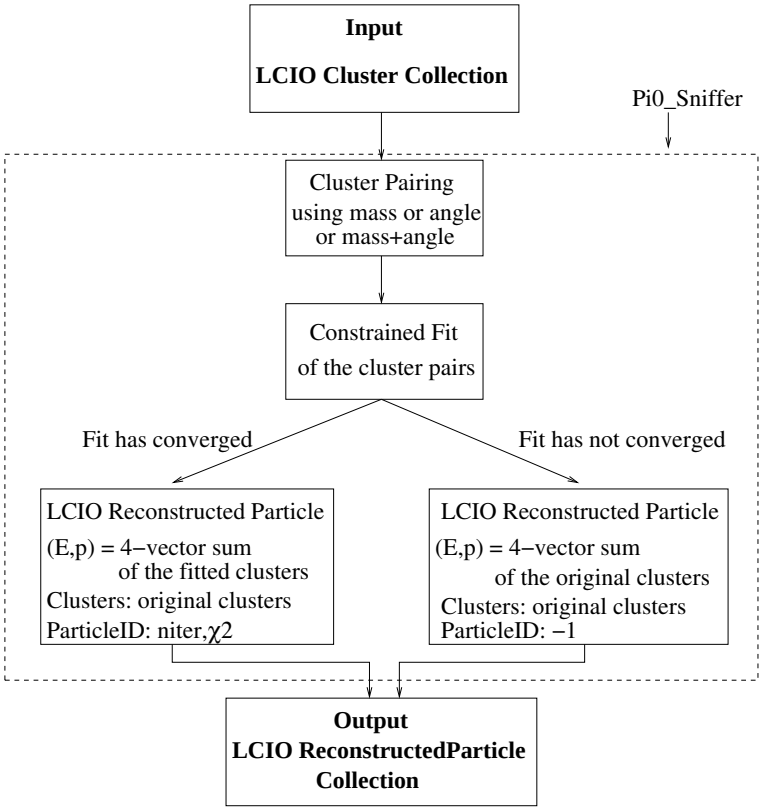

Figure 14: Schematic of the `Pi0_Sniffer` processor.

π^0's in physics events was then developped. Using an estimator based on the mass and the angle of cluster pairs, it was shown that 77% to 78% of the total π^0 energy in $hZ \to b\bar{b}\nu\bar{\nu}$ events at $\sqrt{s}$=500 GeV could be tagged properly. This approach was implemented in a C++ processor called `Pi0_Sniffer` that may be used in Marlin to reconstruct π^0's within the full simulation and reconstruction framework.

The strategy to reconstruct π^0's was developped on MC-smeared events. The next step is to use Particle-Flow and photon identification algorithms to perform π^0 reconstruction in the most realistic conditions. The impact of clustering algorithms has to be evaluated and the effect of π^0 reconstruction in jet clustering and pairing has to be estimated.

Acknowledgments

We are indebted to Pascal Gay for the reading of the manuscript.

References

[1] T.Sjostrand, S.Mrenna and P.Skands JHEP05 (2006) 026

[2] http://polywww.in2p3.fr/activites/physique/geant4/tesla/www/mokka/mokka.html

[3] http://ilcsoft.desy.de/marlin/

[4] http://www-flc.desy.de/ilcsoft/ilcsoftware/MarlinReco

[5] `TrackwiseClustering` in MARLIN presented by Dennis Martsch - ILC Software Workshop June 2005

[6] http://root.cern.ch

[7] http://lcio.desy.de/