TB2003 Topological Clustering Algorithms for Pions in Position 4L

Louise Heelan

July 9, 2008

1 Introduction

This document presents a first look at eight different topological clustering algorithms for pions in the 4L position. The results presented in this document are post-athena and post-merger. The standard beam-cleaning and beam envelope cuts, and halfling timing cuts have been applied (as in the final 4L analysis). This analysis was carried out in Athena 13.0.40. The pion energy distributions are shown on the electromagnetic (em) scale, and are fit to a 6-parameter double Gaussian (vs the now standard 4-parameter double Gaussian, as with each clustering configuration the input constraining parameters would need to be changed slightly). For comparison, when possible, results when clustering with an 16 cm cylinder cluster are presented with the same cuts as above (this will result in slightly different numbers from those published in the final paper due to the lack of halfling time cuts and 6 versus 4 parameter double Gaussian fit). Recall that the energy resolution is worse using the em scale, as the flat weights are designed to minimize the energy resolution.

1.1 Seed, Neighbor, Cell Noise Sigma Cuts

In the topological clustering algorithm clusters are formed based on the cell energy over noise significance. In the default algorithm, a cluster is seeded in the hadronic calorimeters by a cell if it has an energy over noise significance greater than four. A 4σ cut implies that statisically only 0.006 % of clusters will be seeded by a noise and not true signal energy. A cluster is expanded upon by examining the cell energy significance of neighboring cells. Neighboring cells in the FCal are cells that share a common edge of the same FCal module (typically the 4 cells above, below, to the left and right, and the 4 cells on the corners for a total of 8 neighboring cells), and cells that are located in the previous or next module that cover the same η and ϕ region. By default for the hadronic calorimeters a neighbor cell energy significance of two is required to expand the cluster. The last configurable option is the cell energy significance, these are neighboring cells that are included by default in the cluster if they neighbor a cell that is already included in the cluster. By default, this value for the cell energy significance is zero. The FCal is considered a hadronic calorimeter in the Athena software, and as such has a default topological clustering configuration of: seed/neighbor/cell = $4\sigma/2\sigma/0\sigma$, which will be referred to as t420.

In the test beam 2003 we use noise files generated on a run-by-run and cell-by-cell basis to identify the noise in a given cell for a given run.

2 Energy Distributions and Linearity

Sample energy distributions for topological clustering t420 is shown in Figure 1. Similar distributions and fit exist for the other topological clustering configurations. One concern with the topological clustering technique was the noise subtraction, since the topocluster clusters a different number of cells each event, the worry was that the noise distribution (summing the energy from pedestal runs of the same cells that are clustered in the physics data) would not be centered at zero, or would have some non-Gaussian shape. The noise distributions for t420 are shown in Figure2, where there are similar distributions for the other configuraitons. The distributions appear to be Gaussian, and well behaved like the cylinder clustering technique. Therefore, in calculating the energy resolution the noise has been subtracted in quadrature from the signal rms.

The linearity is shown in Figure 4 for the eight topological clustering algorithms. As the clustering requirement gets more stringent less energy is clustered, resulting in a degraded linearity. Note that the linearity for the 10 GeV point is around 0.45, and below the scale of this figure. Remember that these figures are at the em scale, hence the large difference from one.

The reconstructed noise increases as a function of beam energy for the topological clustering, I guess due to the fact that at higher beam energy more cells are included in the cluster, giving way to larger fluctuations. For cylinder clustering the number of cells clustered is approximately constant, and the reconstructed noise as a function of beam energy is relatively flat. The reconstructed noise distributions for t420 and c16 are shown in Figure 3. It is interesting to note that the topological clustering does do a better job at suppressing the noise.

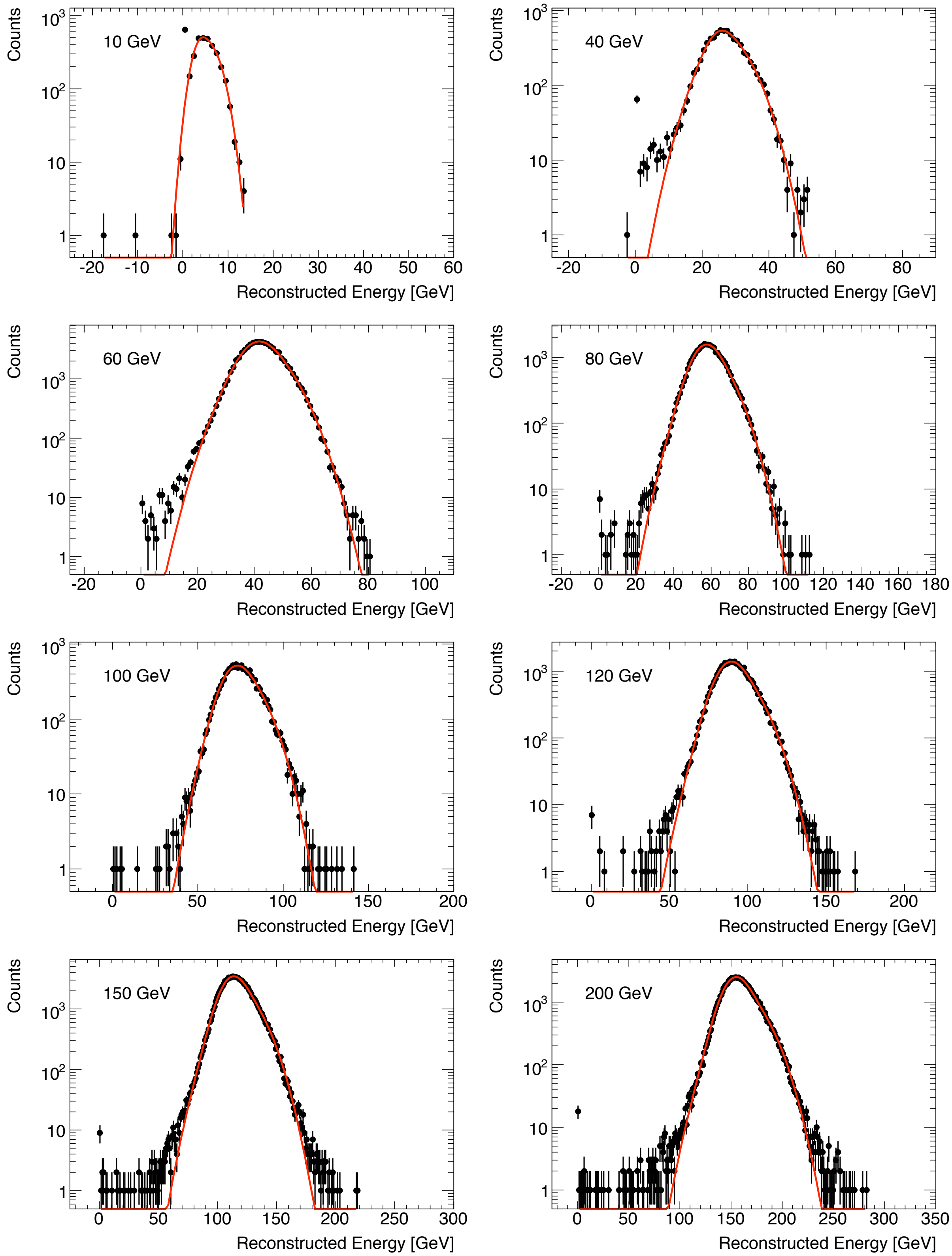


Figure 1: Pion energy distributions (em scale) when the data are clustered using topological clustering configuration 4/2/0. The double Gaussian fit results are superimposed.

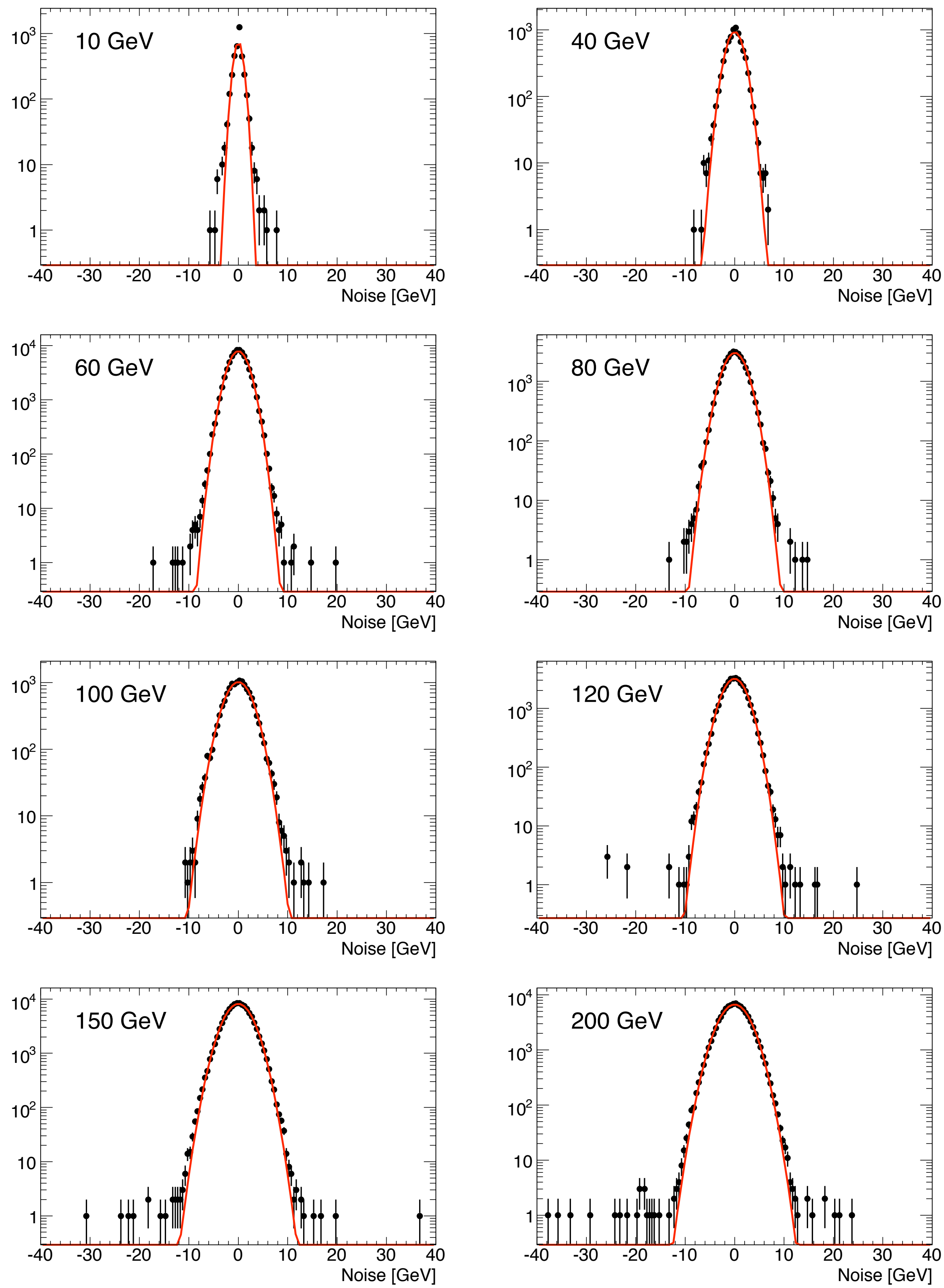
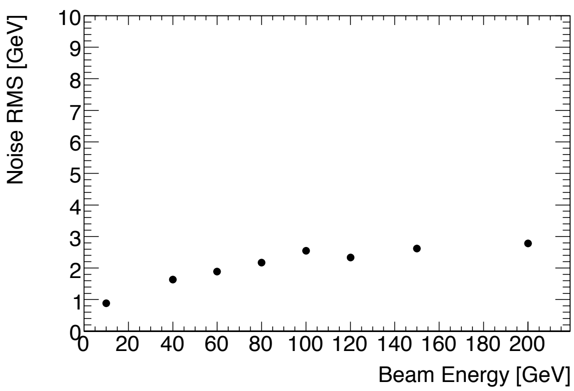
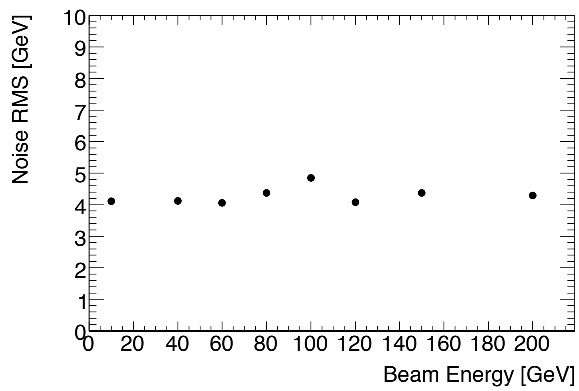


Figure 2: Pion noise energy distributions (em scale) when the data are clustered using topological clustering configuration 4/2/0. The single Gaussian fit results are superimposed.

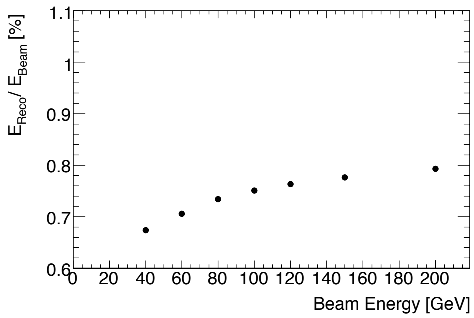


(a) t420

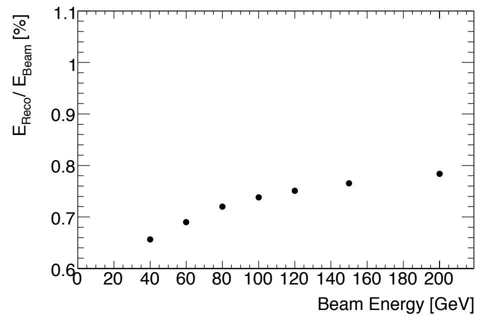


(b) c16

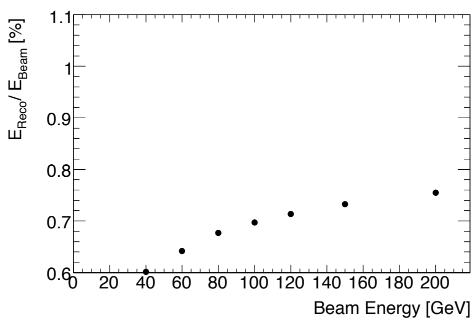
Figure 3: Pion reconstructed noise distributions (em scale) as a function of beam energy for topological clustering t420 (left) and the 16 cm cylinder cluster (right).



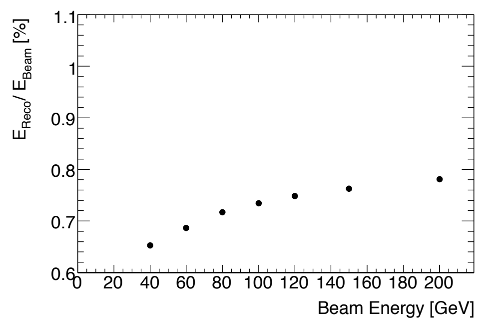
(a) t420



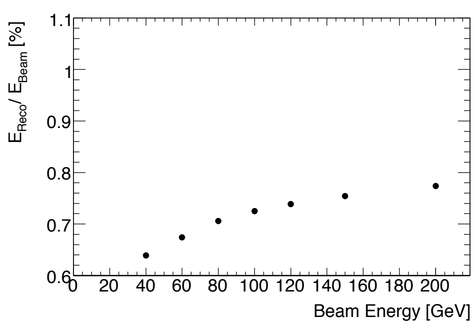
(b) t421



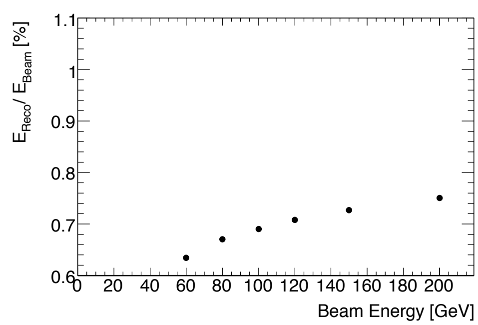
(c) t422



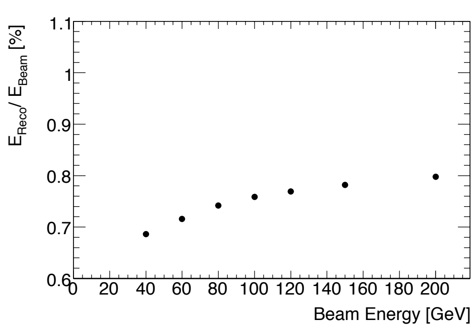
(d) t430



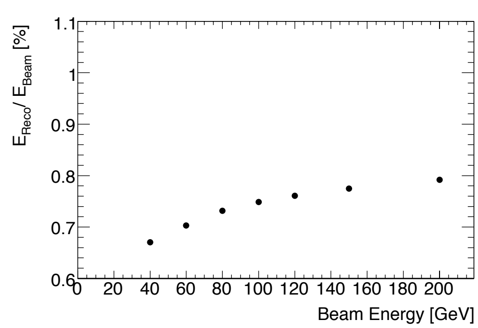
(e) t431



(f) t432



(g) t320



(h) t520

Figure 4: Pion linearity: reconstructed energy versus beam energy, for the various topological clustering algorithms.

3 Energy Resolution

The (noise subtracted) energy resolution points used for all energies for each clustering configuration are shown in Table 1. These points are fit to the noise subtracted energy resolution function:

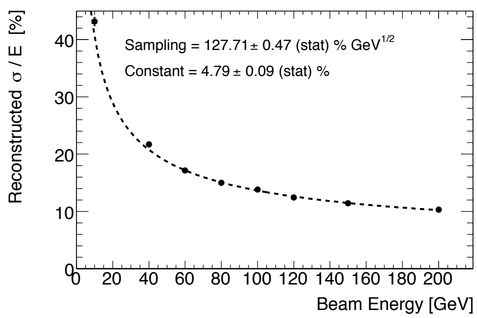
$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus b \tag{1}$$

where a is the sampling term and b is the constant term. The resulting fits for a and b are shown on the respective energy resolution figures for each clustering type, shown in Figure 5.

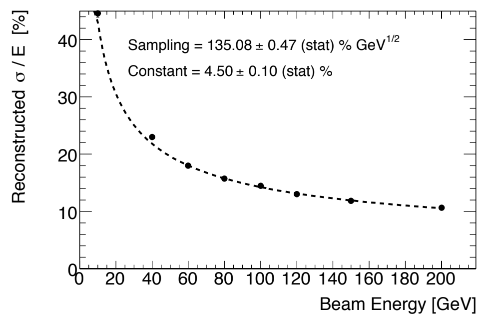
	10 GeV	40 GeV	60 GeV	80 GeV	100 GeV	120 GeV	150 GeV	200 GeV
t420	43.179	21.708	17.163	14.988	13.825	12.449	11.419	10.292
t421	44.594	23.016	17.978	15.710	14.468	12.998	11.841	10.631
t422	50.642	26.218	20.358	17.543	16.163	14.484	13.117	11.643
t430	43.460	23.385	18.337	15.884	14.603	13.208	12.038	10.775
t431	45.642	24.391	18.972	16.456	15.105	13.591	12.365	11.033
t432	50.559	26.929	20.963	18.027	16.581	14.835	13.423	11.876
t320	57.435	21.541	17.064	15.055	13.822	12.375	11.322	10.191
t520	39.454	21.978	17.405	15.132	13.965	12.607	11.501	10.357
c16	31.225	19.318	14.420	13.039	12.464	11.370	10.687	9.894

Table 1: Energy resolution (sigma/mean) [%] after noise has been subtracted in quadrature for the various topological seed/neighbor/cell configurations. For comparison the 16 cm cylinder clustering is shown, where the same cuts have been applied.

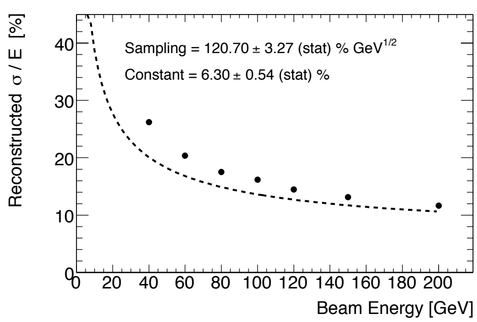
The cylinderical clustering give the best overall energy resolution at the range of energies from 10 GeV to 200 GeV. In general, the cylinderical clustering reconstructs a higher mean energy and rms than the topological clustering configurations. The fixed cylinder size most likely has less variation in the amount of energy clustered. The noise subtracted fits to the energy resolution function for the eight topological clustering algorithms are shown in Figure 5. For reference, the 16 cm cylinder cluster with the same cuts yields a fit for the sampling term of 99.2 % $\mathrm{GeV}^{1/2}$ and a constant term of 6.97 % [in the published paper these values are 94.2 % $\mathrm{GeV}^{1/2}$ and 7.5 %, respectively].



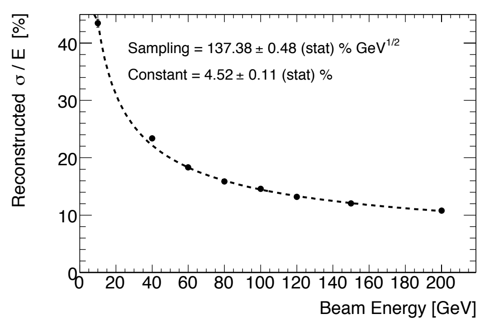
(a) t420



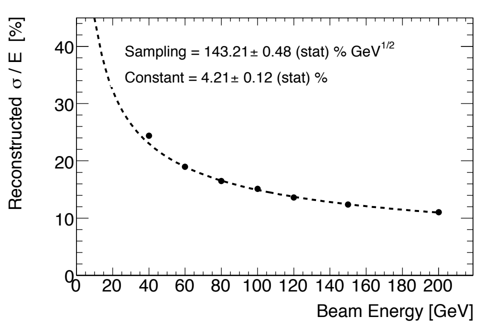
(b) t421



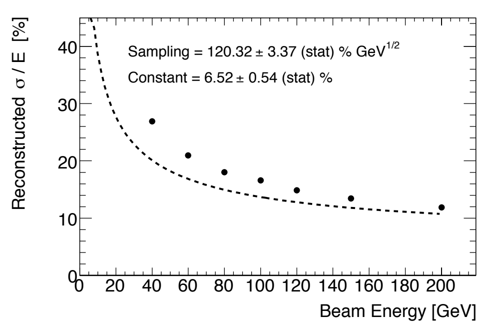
(c) t422



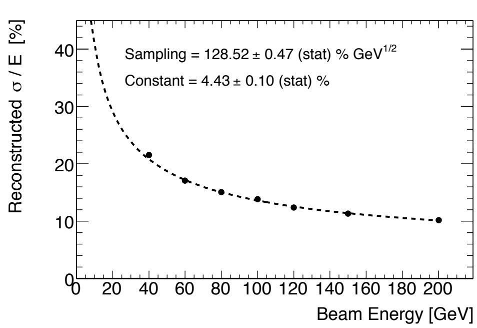
(d) t430



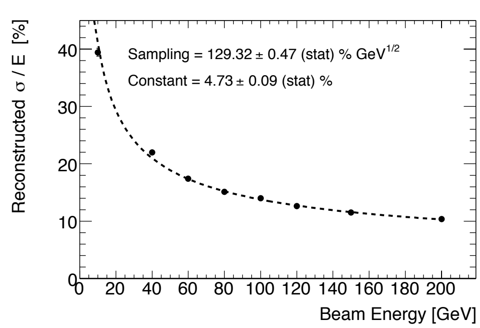
(e) t431



(f) t432



(g) t320



(h) t520

Figure 5: Pion energy resolution functions for the various topological clustering algorithms.