A Minimum Spanning Tree clustering algorithm for the Fermi LAT Calorimeter

Fermi Gamma-ray Space Telescope

INFN Istituto Nazionale di Fisica Nucleare

C. Sgrò[1] and L. Baldini[1]
on behalf of the Fermi Large Area Telescope Collaboration
[1]Istituto Nazionale di Fisica Nucleare, sezione di Pisa

Summary: A clustering algorithm for the Fermi LAT calorimeter. The main goal is to separate the energy deposited by genuine γ-rays from the accidental on-orbit pile-up activity.

The event reconstruction developed for the Fermi Large Area Telescope before the launch, and currently in use with minor modifications, does not feature a calorimeter clustering stage. In a low occupancy environment such as the one in which the telescope operates this approach proved to be adequate to support the science analysis of the first two years. However, it became clear in the first months of operation that some clustering algorithm was necessary to recover a loss of effective area caused by the signal from pile up and accidental coincidences in the detector.

We present a clustering algorithm for the Large Area Telescope calorimeter based on the concept of a Minimum Spanning Tree: for each event the tree spanning the entire crystal collection is first constructed and then the edges exceeding an adjustable energy-dependent threshold are removed. Tests performed on flight data and Monte Carlo simulations show that our approach is effective in separating the genuine gamma-ray signal from the leftovers of accidental coincidences.

Minimum Spanning Tree clustering

The Fermi LAT calorimeter [1, 2] is composed of 1536 CsI(Tl) crystals, each providing a 3-dimensional hit position, arranged in a hodoscopic (not projective) configuration.

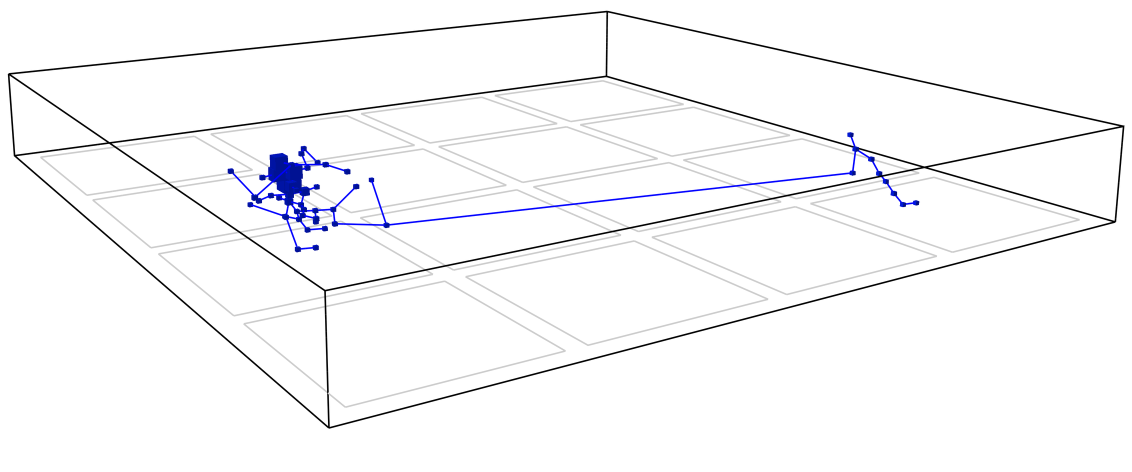


Figure 1: *MST for a 1.6 GeV simulated γ-ray event overlaid with on-orbit instrumental pile-up.*

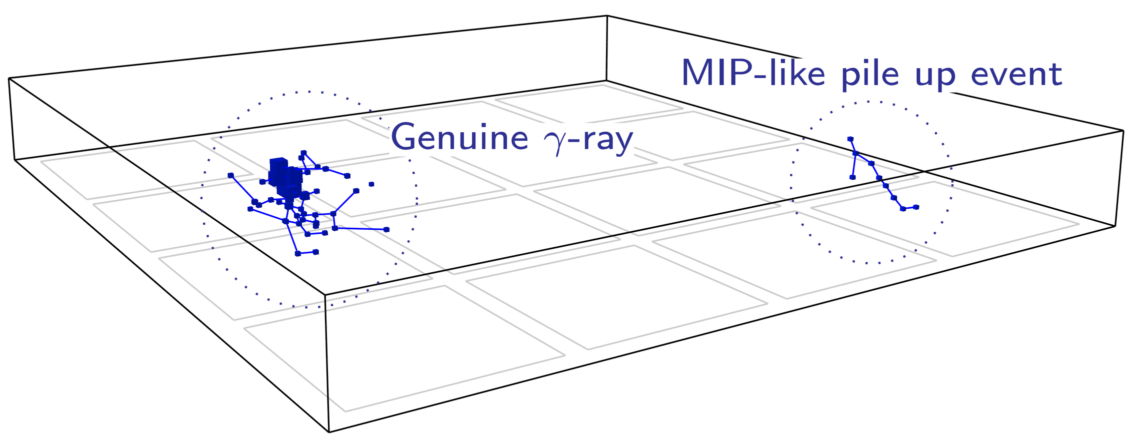


Figure 2: *Sub-trees after the clustering stage.*

The Minimum Spanning Tree (MST) is the shortest tree with no loops connecting all the nodes of a graph.

- Well studied combinatorial problem with efficient algorithms for solution.
- Univocally determined if the lengths of the edges are all different.
- The Euclidean distance (used here) is the simplest choice for the metric.

MST clustering in a nutshell:

- Construct the MST for the entire collection of calorimeter hits (Fig. 1).
- Remove every single edge whose length exceeds a pre-defined, energy-dependent threshold (Fig. 2).
- Group together hits belonging to the same sub-tree.

The choice of the threshold

The threshold needs to be relaxed at low energy, where the showers are sparse:

- γ-ray events below ≈ 1 GeV deposit a sensible fraction of their energy in the tracker.
- Shower topology heavily depends on the conversion point in the tracker.
- Artificial splitting is detrimental for energy resolution and background rejection.

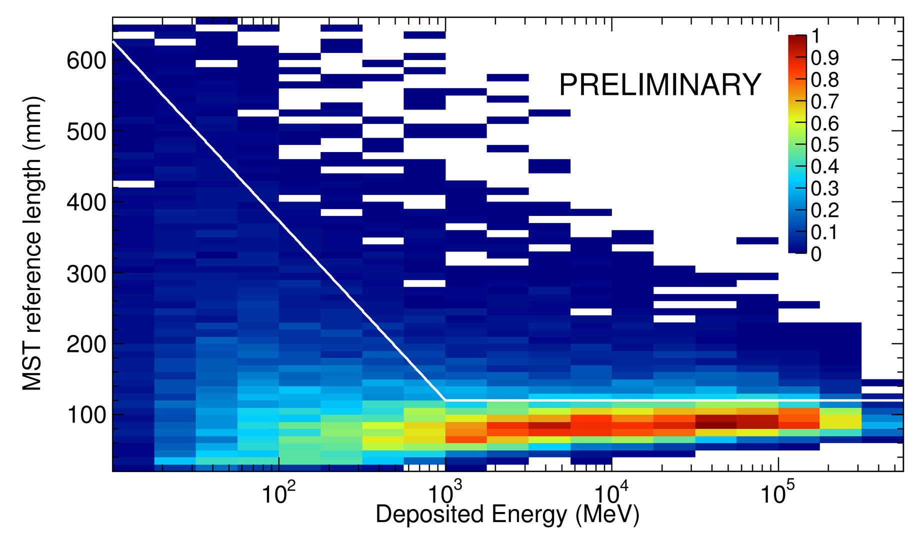
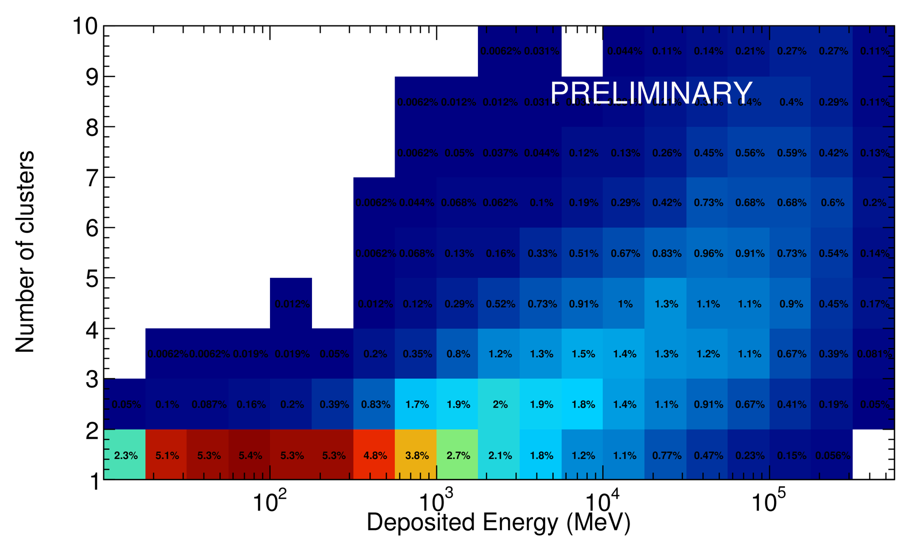


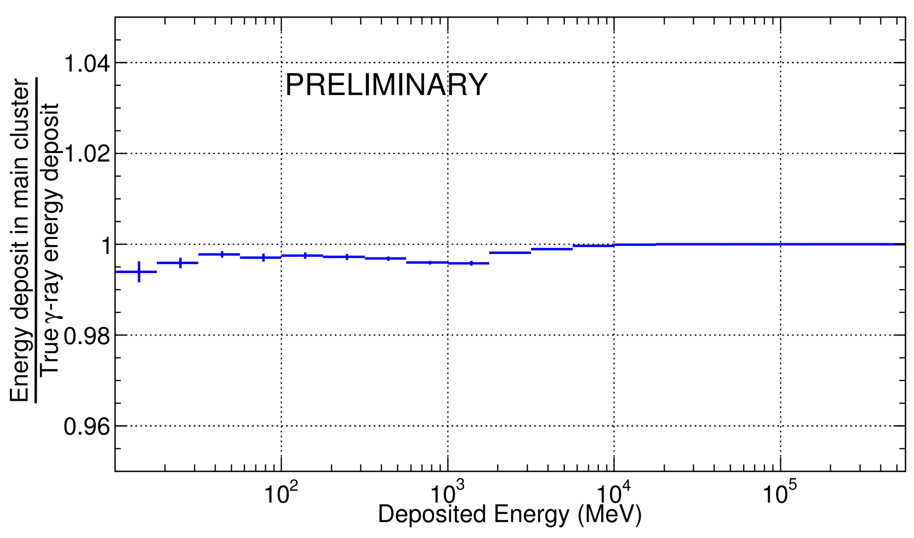
Figure 3: *MST threshold as a function of deposited energy (white line), compared with the reference length (average $+2\sigma$ of MST lengths) of simulated γ-rays.*

Tests without instrumental pile-up



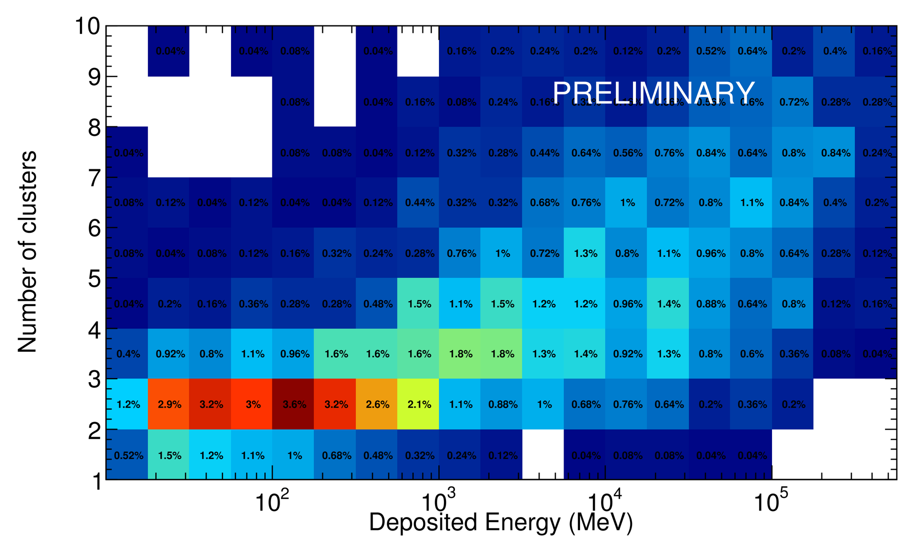
Simulated γ-ray events with no instrumental pile-up (ideally we don't want to split them):

- At low energy we produce one single cluster for the vast majority of the events.
- We do produce multiple clusters at high energy but. . .



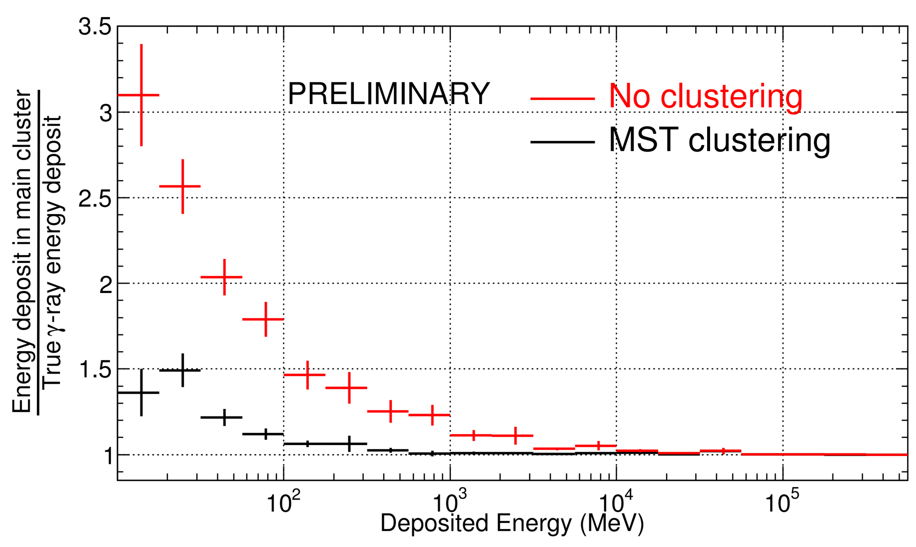
. . . most importantly: on average more than 99% of the deposited energy belongs to the main cluster across the whole energy range.

Tests with instrumental pile-up



We added on-orbit instrumental pile-up (from a dedicated periodic trigger) with at least 5 MeV activity in the calorimeter on top of simulated γ-ray events.

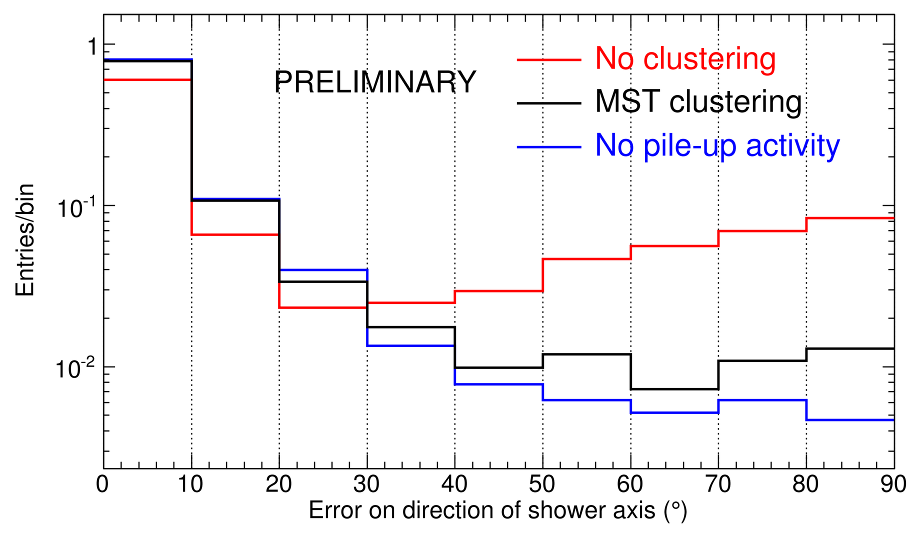
- We produce two or more clusters for most of the events.



Pile-up increases the energy deposited in the calorimeter:

- More problematic for small γ-ray energy deposition.
- The clustering removes most of the effect of the pile-up activity (caveat: the "main" cluster is selected using Monte Carlo information, see also [4]).

Reconstruction of shower axis



The direction of the shower axis, reconstructed through a moments analysis, serves as an input for the track finding and the background rejection stages.

- Due to the large footprint of the instrument, pile-up activity can flip the reconstructed shower axis by $\sim 90^\circ$ (events above 100 MeV with minimal quality cuts shown in the figure on the left).
- The clustering sensibly reduces this effect.

References

[1] Atwood, W. B. et al., *The Large Area Telescope on the Fermi Gamma-ray Space Telescope Mission*, ApJ, 697, 1071 (2009)

[2] J. Eric Grove and W. Neil Johnson, *The calorimeter of the Fermi Large Area Telescope*, Proc. SPIE 7732, 77320J (2010)

[3] G. Mavromanolakis, *Calorimeter clustering with minimal spanning trees*, arXiv:physics/0409039v2

[4] M. Pesce-Rollins et al., *Application of the Naive Bayes Classifier technique to the Fermi LAT Calorimeter data*, poster at this symposium