Clustering of Hadronic Showers with a Structural Algorithm

M. J. Charles
The University of Iowa, Iowa City, IA 52242, USA

The internal structure of hadronic showers can be resolved in a high-granularity calorimeter. This structure is described in terms of simple components and an algorithm for reconstruction of hadronic clusters using these components is presented. Results from applying this algorithm to simulated hadronic Z-pole events in the SiD concept are discussed.

1. INTRODUCTION

The Silicon Detector (SiD) [1] is a proposed detector concept for the future International Linear Collider (ILC). The goal of the ILC is to make precision measurements which will shed light on a broad range of topics, including anticipated new physics in the Higgs and Supersymmetry sectors [2]. To achieve these physics goals, an ILC detector must have an excellent jet energy resolution of the order of $30\,\%/\sqrt{E}$ [3]. The SiD concept is based on the assumption that although this resolution cannot be attained with a conventional sampling calorimeter, it can be reached with a Particle Flow Algorithm (PFA) in which charged energy is measured in the central tracking system and neutral energy is measured in the calorimeters, provided that those calorimeters have sufficiently fine granularity to resolve individual final-state particles. In particular, separation of charged and neutral hadronic energy is critical. In this paper, a method to accomplish this is presented and evaluated.

2. DESCRIPTION OF THE STRUCTURAL ALGORITHM

Hadronic showers exhibit a great deal of variation in their structure compared to electromagnetic showers. The fine granularity of the SiD calorimeters allows much of this structure to be resolved. From studies of simulated single-particle and low-multiplicity events, it was found that the showers can generally be described in terms of the following components: (a) dense clumps, (b) track segments, (c) a halo of less dense hits following a hard interaction, and (d) displaced secondary fragments. By linking together components (a) and (b), the "skeletons" of hadronic showers are reconstructed; components (c) and (d) can then be added to recover the remaining hits. In this way, hadronic showers can, in principle, be reconstructed with high efficiency and purity even in dense environments.

Algorithms to identify components (a) and (b) were written, based on existing implementations of a Minimum Spanning Tree (MST) clusterer and a Minimum Ionising Particle (MIP) finder [4]. A selector is then required to determine whether a given pair of components were directly linked. A preliminary, cut-based implementation was presented at Snowmass; this has now been replaced by a likelihood selector using the following geometrical quantities: (i) the spatial separation between any pair of components, (ii) the distance of closest approach of a track to another component, (iii) the distance of two tracks to their point of closest approach, (iv) whether the point of closest approach is inside the calorimeter, and (v) how frequently additional hits are found when extrapolating a track to the point of closest approach[1]. Likelihood distributions produced with these variables in hadronic Z-pole events using the `org.lcsim` framework are shown in Figure 1.

After linking components (a) and (b) into skeletons, components (c) and (d) are identified and associated to the corresponding skeleton. For component (c), the MST algorithm is used to find hits close to the cluster structure.

[1]Used in the `hep.lcd` software framework but not yet in `org.lcsim` due to an incompatiblity in the geometry description.

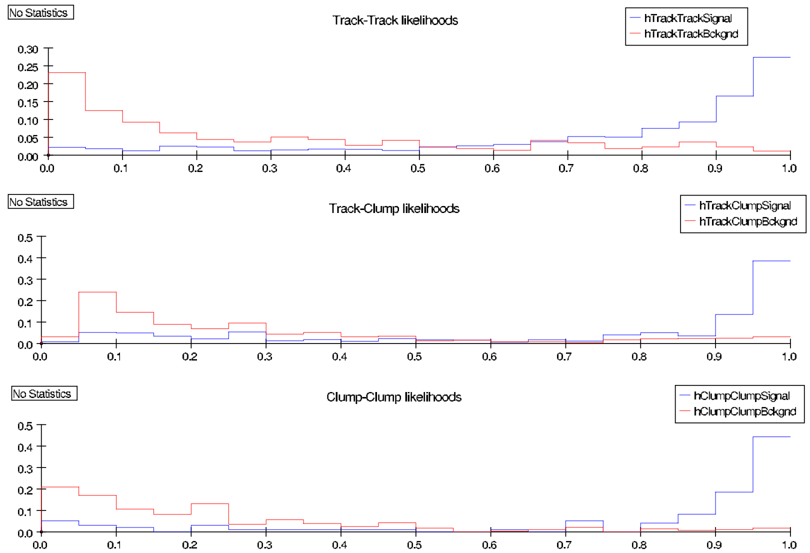

Figure 1: Likelihood distributions for links between components of hadronic clusters, obtained in a sample of hadronic Z-pole events.

The association is more complex for component (d), displaced fragments produced by neutral secondary particles. There are two problems to consider: first, whether a small cluster is really a secondary fragment, or whether it was produced by a neutral primary particle (e.g. a soft neutron); second, if it is a fragment, to which primary cluster it should be associated. In the current implementation, fragments are identified by a cut-based selector, requiring that (i) the cluster has few hits, (ii) the cluster does not contain any dense clumps or track segments, (iii) the principal axis of the cluster does not point towards the interaction point, and (iv) the cluster does not immediately follow a track found in the central tracking system. Clusters identified as fragments are then associated to the closest non-fragment. In the future, both the fragment identification and association algorithms will be replaced by more sophisticated methods.

3. PERFORMANCE IN Z-POLE EVENTS

The performance of the algorithm in approximately 400 simulated hadronic Z-pole events at a center-of-mass energy of 91.0 GeV in the `SDFeb05_SciHcal` detector was studied in the `hep.lcd` framework. The structural algorithm described in the previous section was used to reconstruct clusters in the calorimeters. Charged tracks were then matched to the corresponding clusters; this was done with truth information in lieu of a full tracking algorithm. No cuts were made on the angular distribution of energy in the event, nor on the amount of neutrino energy. The default sampling fractions were used; this calibration is known to be somewhat flawed but should be good to within 5–10%. The event energy sum E was then computed as $E = E_{\text{charged}} + E_{\text{neutral}} + E_{\nu}$, where E_{charged} is the total charged energy in the event, E_{neutral} is measured as the total non-charged energy in the calorimeters, and E_{ν} is the total neutrino information in the event, taken from truth information. The energy of a charged cluster is determined from the momentum and mass hypothesis of the associated track in the central tracking system—for this study, these were taken from truth information.

Applying the clustering algorithm described in the previous section, the energy sum distribution shown in Figure 2a with RMS of 4.2 GeV was obtained. The measurement was then repeated when using "perfect" (cheating) pattern recognition; the distribution of event energy sums E is shown in Figure 2b and has an RMS of 3.5 GeV. As a crosscheck, the measurement was again repeated with the algorithm described in the previous section, but with fragment identification and association performed using truth information; the RMS was found to be 3.4 GeV,

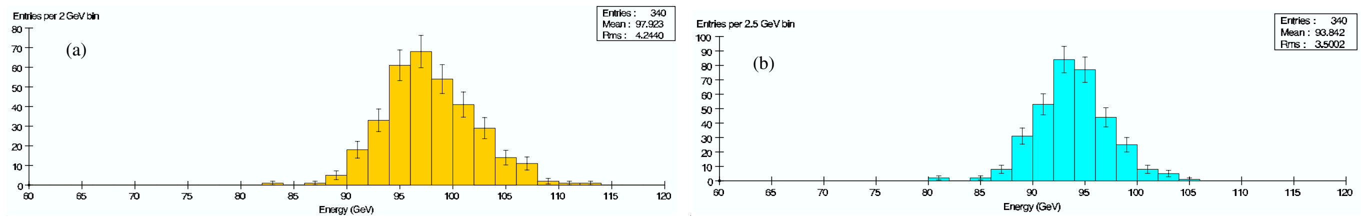

Figure 2: Energy sum plots for Z-pole events, showing (a) the reconstructed energy sums without cheating, and (b) the reconstructed energy sum for a “perfect” cheating PFA. The true energy sum is 91.0 GeV, but the correct sampling fractions were not available for this simulated detector; as a result, the overall energy scale is incorrect and the resolutions are worse than could be achieved with full calibration. The RMS values correspond to $37\%/\sqrt{E}$ for (a) and $45\%/\sqrt{E}$ for (b).

consistent with the perfect reconstruction within the experimental limitations.

These results suggest that the algorithm performs well at identifying and linking components (a), (b), and (c), but that the identification and association of secondary neutral fragments needs to be improved substantially. The difference in the means of Figure 2 (a) and (b) also indicates a related problem: if secondary fragments of charged clusters are not correctly associated, their energy may be double-counted (both as neutral energy and as part of the track momentum). Again, this implies that fragment association needs to be improved significantly.

4. CONCLUSIONS

The preliminary results in the `hep.lcd` framework are promising, but indicate the need for further improvements in the fragment association. The algorithm has since been rewritten for the `org.lcsim` framework and is currently being tested. A clustering algorithm for electromagnetic showers is also under development [5] and will be integrated when it is complete. With a reliable calibration, the performance can then be optimised towards the $30\,\%/\sqrt{E}$ goal.

Acknowledgments

I am grateful to Ron Cassell for his help and advice. This work has been supported by the Department of Energy under grant numbers DE-FG02-91ER40664 and 234151N.

References

[1] J. E. Brau, M. Breidenbach and Y. Fujii, “The silicon detector (SiD) and linear collider detector R&D in Asia and North America,” SLAC-PUB-11413 *Prepared for 4th ECFA / DESY Workshop on Physics and Detectors for a 90-GeV to 800-GeV Linear e+ e- Collider, Amsterdam, The Netherlands, 1-4 Apr 2003.*

[2] H. Murayama and M. E. Peskin, Ann. Rev. Nucl. Part. Sci. **46**, 533 (1996) [arXiv:hep-ex/9606003].

[3] J. Brau *et al.*, “Linear collider detector R & D,” *Prepared for International Workshop on Linear Colliders (LCWS 2002), Jeju Island, Korea, 26-30 Aug 2002.*

[4] W. F. Mader, “Mip Reconstruction Techniques And Minimum Spanning Tree Clustering,” SLAC-PUB-11359 *Contributed to 2005 International Linear Collider Workshop (LCWS 2005), Stanford, California, 18-22 Mar 2005.*

[5] N. Meyer, “Electromagnetic Shower Reconstruction for the Silicon Detector,” ALCPG1108, these proceedings.