FERMILAB FN-0709

Energy Flow in CMS Calorimetry

Dan Green

September, 2001

Operated by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the United States Department of Energy

Introduction

The Compact Muon Solenoid (CMS) is one of the two general purpose detectors being built to perform experiments at the CERN Large Hadron Collider (LHC). The calorimetry in CMS consists of an electromagnetic compartment of PbWO4 crystals [1]. Thus, there is only a single longitudinal segment which contains the photons and electrons. However, there is a very fine transverse segmentation. A crystal tower subtends a $d\eta \times d\varphi \sim 0.0174 \times 0.0174$ element of solid angle. This segmentation corresponds roughly to the Moliere radius in the crystal.

The hadronic compartment is also a single longitudinal segment [2]. The transverse segmentation is five times coarser, 0.087 x 0.087. Indeed, it is well matched to the physical size of an hadronic shower.

In this note, we examine the possibility of performing particle identification using the longitudinal and transverse segmentation of the CMS calorimetry. In addition, we consider a mixed measurement for jets, consisting of a mixture of neutral particle calorimeter clusters and charged particle momenta measured in the CMS magnetic tracking system. This hybrid, or "energy flow" set of measures should be a better measure of the kinematics of a jet because the tracking error is much smaller than the calorimetric energy error for low momentum charged particles. At issue is whether the jet errors are dominated by energy measurement errors or by other mechanisms, such as gluon radiation.

Transverse Shape in ECAL

The electromagnetic calorimeter, ECAL, has a fine transverse segmentation, as mentioned above. However, there is no longitudinal segmentation, so that the detector is blind to the details of the shower development in depth. Test beam data with a setup close to that proposed fro CMS has been taken [3]. In particular, a 7 x 7 array of crystals was exposed to beams of pions and electrons. The data for four single pions of 100 GeV beam momentum are shown in Fig. 1. Note that the pions which interact in ECAL typically deposit energy in at least two crystals.

The analogous deposits for 100 GeV electrons are shown for four individual events in Fig.2. Note that the showers are, indeed, narrower than the pion showers. However, it is clear that the electron can always have an impact point near to the crystal boundary. Therefore, electrons depositing energy in two crystals are not rare.

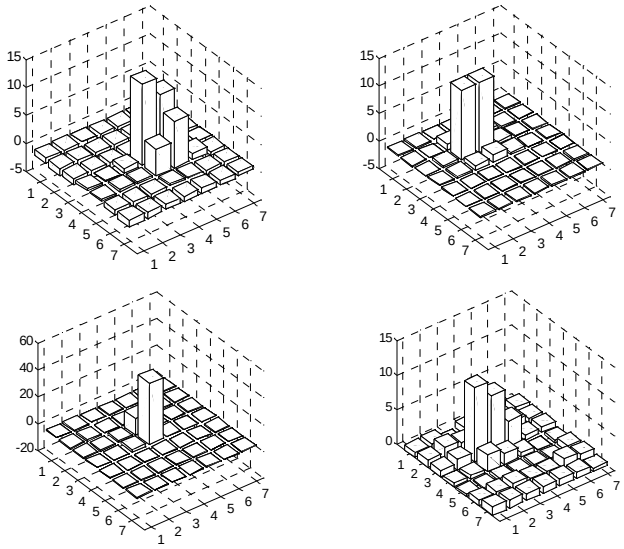

Figure 1: Energy deposition in a 7 x 7 crystal ECAL array for 100 GeV pions which have interacted in ECAL.

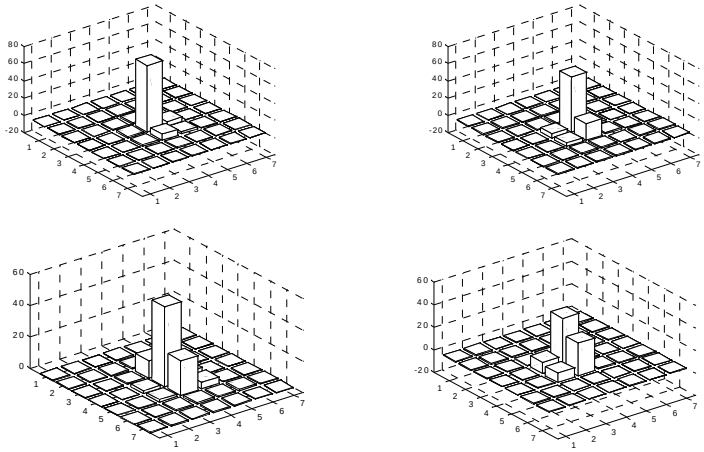

Figure 2: Energy deposition in a 7 x 7 crystal ECAL array for 100 GeV electrons.

The distribution of the energy weighted transverse radius for electrons and pions is shown in Fig. 3. Clearly, there is some discrimination between pions and electrons. However, incisive particle identification cannot be made. A clean particle identification requires additional data, such as a fine grained shower maximum detector or much more information of the development of the shower in depth.

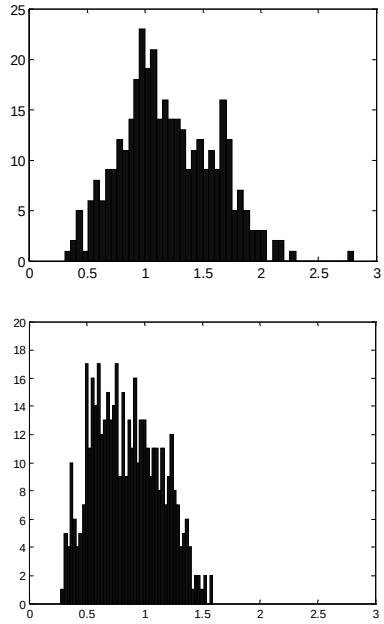

Figure 3: Distribution of the energy weighted transverse distance of energy deposits for pions and electrons.

Calorimeter Clusters

In order to explore the energy flow concept, a sample of "Z" with 120 GeV mass was generated. Initial and final state radiation was turned on in the PYTHIA Monte Carlo program. The resulting final state particles were passed through the CMS detector simulation. The output which was available was the energy deposited in the towers of the calorimetry and the generator level charged particle tracks at the production vertex. The first task was to combine contiguous calorimeter hits into "clusters" and to look at the longitudinal information available for that cluster.

The mean number of hits with transverse energy above 0.1 GeV in the electromagnetic compartment is 324, while the hadronic compartment has 344 hits on average. Since there are roughly 5000 hadronic towers in the CMS barrel plus endcap (spanning $|\eta| < 3$ with size 0.087 x 0.087), the occupation of the calorimeter is fairly sparse.

The calorimeter hits were clustered into groups of towers. The clustering began in the electromagnetic calorimetry, ECAL, as it had the finest granularity. Hits were ordered in transverse energy, Et. The largest Et tower served as a seed for the cluster. Additional hits in the ECAL were added to the cluster in a 3 x 3 ECAL array centered on the seed tower. All hits above 0.1 GeV were accepted. The ECAL clustering was terminated when the seed tower transverse energy fell below 0.2 GeV. All towers that participate in a cluster are removed from further clustering.

The information in depth was used to sort the ECAL clusters into two categories. Energy in the hadronic calorimeter compartment (HCAL) behind the ECAL seed was added in a 3 x 3 HCAL tower array if the found HCAL transverse energy is greater than 30% of the ECAL cluster energy. This value was picked using test beam data [3] for pions interacting in the ECAL compartment. If there was no HCAL deposit, then the cluster was considered to be a neutral particle, flag = 0. If there was a match, the cluster was considered to be a charged pion which interacted in the ECAL compartment.

When the ECAL energy is exhausted, the remaining HCAL energy is ordered in Et. The seed in HCAL has adjacent 3 x 3 HCAL towers clustered similarly to the ECAL clustering. The process is terminated when the seed tower Et falls below 0.2 GeV.

The result of this clustering is a cluster energy, Et weighted centroid in (η,φ), and a particle flag. The flag is 0 if the ECAL cluster has no associated HCAL energy. It is 1 if there is associated HCAL energy of at least 30% of the ECAL energy. Finally, the flag is 2 if there is only HCAL energy without sufficient associated ECAL energy. This last category is assumed to be a charged pion which interacted only in the HCAL.

There are 241 calorimeter clusters on average. The reduction with respect to calorimeter tower hits is due to the large fraction of ECAL events with Et slightly above 0.1 GeV which fall below the seed threshold. As noted below, there is a rough agreement between the number of clusters and the number of charged tracks.

The average cluster transverse energy summed over all flags is 0.94 GeV, with 0.2 GeV in the ECAL and 0.75 in the HCAL. This scale of energy is that of a typical single particle from a minimum bias event. Since the calorimetry covers $|\eta| < 5$, the number of clusters corresponds to a particle density of 2.4 per unit of rapidity. The number of towers is 0.4 in the ECAL and 1.15 in the HCAL, again summed over all flag types. The distribution of the number of HCAL towers in all clusters is shown in Fig. 4.

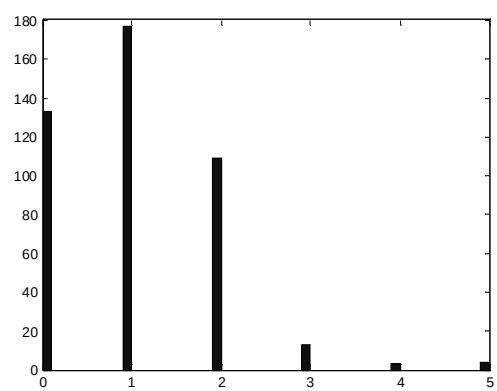

Figure 4: Distribution of the number of HCAL towers in a cluster for all clusters. The maximum number in the 3 x 3 allowed array is 9. Clusters with zero HCAL towers are neutral particles (flag = 0).

The clustering procedure was used to organize the "raw" calorimeter tower hits. The resulting Et of all clusters in one Z(120) event as a function of the cluster φ is shown in Fig. 5. The two jet structure resulting from the Z(120) → qq decay is very evident.

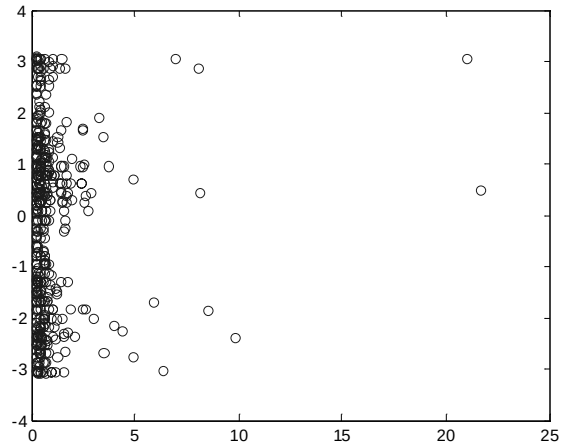

Figure 5: Calorimeter clusters – one event with Et of the cluster plotted against the azimuthal angle of the cluster. The back to back dφ ~ π dijet structure is clearly seen.

Track – Cluster Matches

The tracks used in this study were not actual CMS tracks but the PYTHIA generated charged particles. Given that the tracking momentum resolution is much better than the calorimeter energy resolution, this approximation should not alter the qualitative conclusions which are made.

For events with a produced Z of 120 GeV mass, there are 233 charged tracks on average, with a standard deviation of 144. The primary vertex is on the coordinate origin with a standard deviation in z of 5.2 cm. There are 32% of the tracks which never reach the barrel calorimetry , i.e. loopers.

These tracks are swum in a uniform axial magnetic field to a radius corresponding to the front face of the HCAL. The event by event z position of the vertex is used in the swim. Only tracks with $|\eta| < 3$ and transverse momentum, Pt, greater than 1 GeV are considered. These cuts approximate the angular coverage of the CMS tracker and the lowest transverse momentum not curled up by the magnetic field (loopers). The "looper" transverse momentum to strike ECAL is 0.78 GeV.

Deviations of the swum tracks from the calorimeter clusters are calculated taking the cluster error to be 0.087 (HCAL tower size) for both φ and η. Matching in energy assumes a dominant calorimetric error with a 100% stochastic term and a 5% constant term [3]. Tracks that are within the calorimetric resolution (tower size) in η and φ at the HCAL radius are first merged and their momenta added when considering a match to a calorimeter tower. This procedure

allows for matches in the dense core of jets where charged tracks overlap within a calorimeter tower.

A match between the calorimeter cluster and the track is defined to exist if $|d\eta|$ is less than 1 (deviation normalized to HCAL tower size) if $|d\varphi|$ is less than 2 and if the difference between track momentum and calorimeter cluster energy, normalized to calorimeter resolution, is between –1 and 3. The offset is due to the observation that the calorimeter energies are systematically lower than the track momenta. In the case of multiple matches, the match with the lowest "chi squared" is taken.

The difference in η between tracks and clusters is shown in Fig. 6 for matches which pass the cuts mentioned above. Note that the match improves as the cluster Et increases as we expect due to reduced bending in the magnetic field and improved multiple scattering. There is clearly a band of "accidental" matches which occur at low Et. A similar situation obtains for matches in φ except the cut limits are wider because this is the bend plane in CMS.

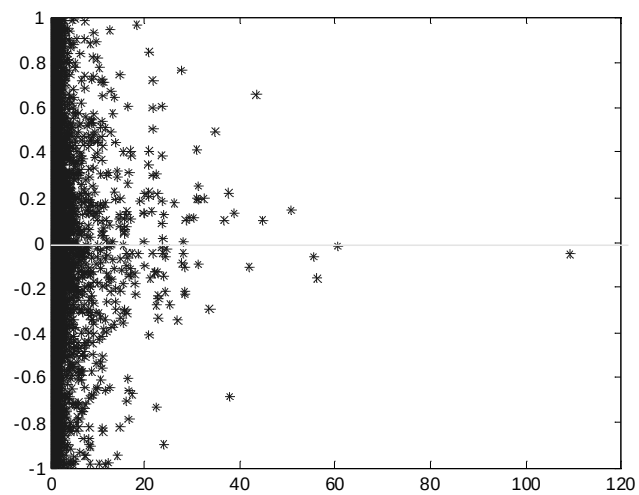

Figure 6: Match in pseudorapidity between charged tracks and calorimeter clusters (flag > 0) as a function of cluster Et.

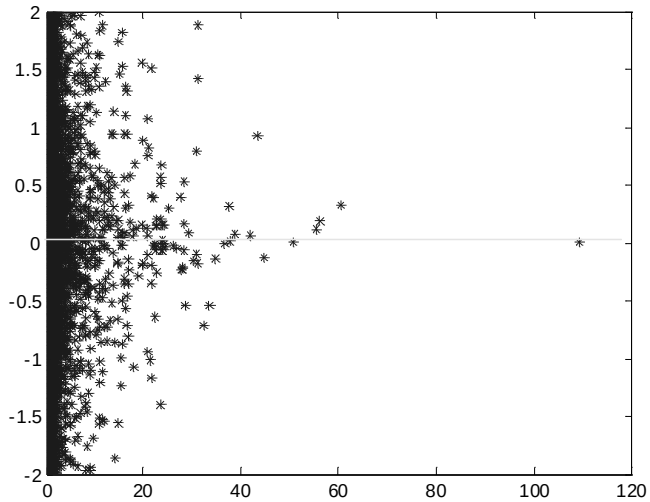

Figure 7: Match in azimuthal angle between charged tracks and calorimeter clusters (flag > 0) as a function of cluster Et.

The match of track momentum to cluster energy is shown in Fig. 8. Note that the calorimetry is systematically under calibrated. Therefore, when clusters are replaced by tracks, we expect the jet Et will be increased. Note also that the error scale set by the calorimeter energy resolution is reasonable. A cut of +- 2 standard deviations about a mean of one is efficient. The CMS calorimeter is nonlinear. This behavior is seen in Fig. 8 in that very high momenta are well matched, while lower momenta are underestimated by the calorimetry. Note that these energy matches are not "accidentals". They survive increasingly hard cuts on matching in η and φ.

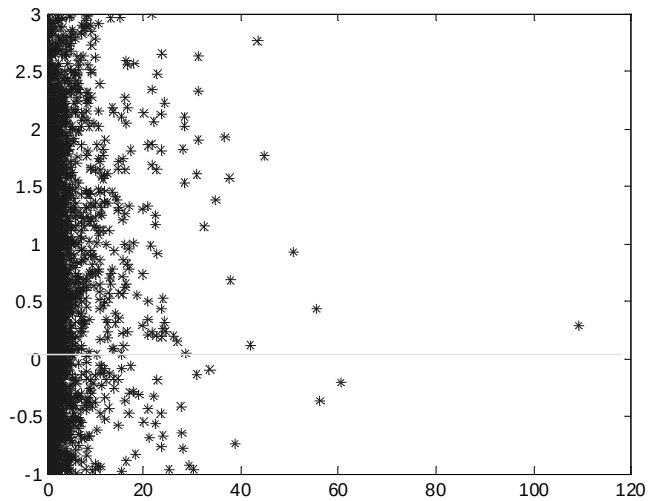

Figure 8: Match in energy between charged tracks and calorimeter clusters (flag > 0) as a function of cluster Et.

If a match exists in all three kinematic variables, the calorimeter cluster is replaced by the matching track. A new flag is defined to be = 3 if this replacement is made. Tracks which would never reach the barrel, “loopers”, are added back into the “energy flow” list and a flag = 4 is set. The end result is an energy flow list consisting of calorimeter clusters (flag = 0, 1, or 2), matched (merged) tracks (flag = 3), and tracks which would not register in the calorimetry (flag = 4).

Jet Properties

The data set was 120 GeV “Z” bosons decaying into light quarks with initial state and final state radiation turned on. The data used was either calorimeter clusters or a hybrid energy flow set consisting of unmatched clusters plus matched tracks and unobserved track loopers.

The largest Et cluster served as the seed for the first jet. The next highest Et cluster a distance > 2R from the first seed was defined to be the seed of the second jet. All clusters within a cone of radius R about the two seeds were taken to be fragments of the jet. The vector sum of all the fragments defined the momentum of the jet. The clusters were assumed to be massless. A plot of the Et of the first seed cluster is shown in Fig. 9.

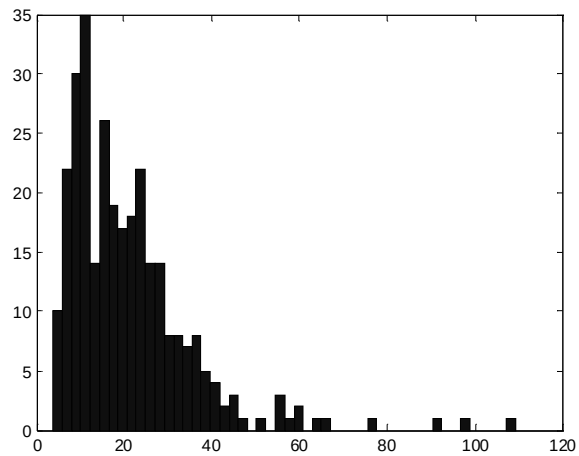

Figure 9: Transverse energy of the first seed cluster found by the jet algorithm. Note that the leading cluster is not soft, but has an Et ~ 20 GeV.

The number of clusters found within a cone of R = 0.8 is shown in Fig 10. Note that there are ~ 300 distinct HCAL towers within this cone, so the cluster density even within the jet is not very high.

The transverse energy of the first found jet is shown in Fig. 11. Note the prominent Jacobean peak at ~ half the Z(120) mass.

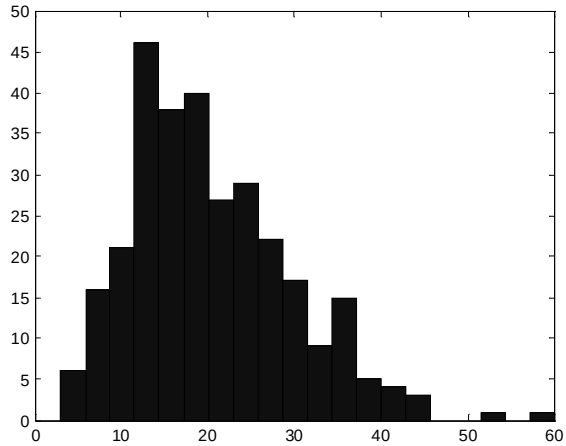

Figure 10: Number of calorimeter clusters found within a cone of radius R = 0.8 for the first found jet. The mean number is ~ 23.

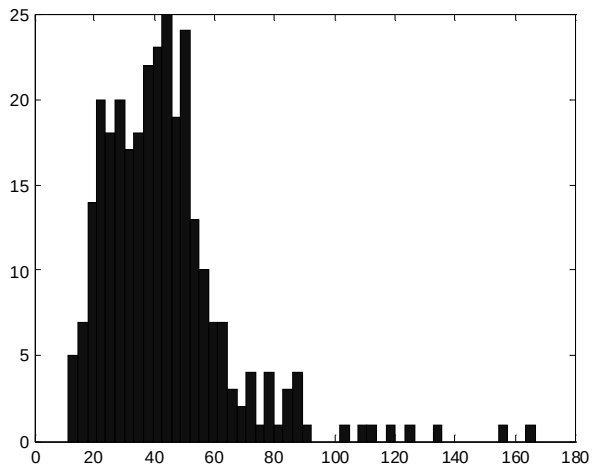

Figure 11: The transverse energy of the first found jet in Z(120) events.

The clusters are treated as massless particles which are summed to find the jet momentum vector. The mass of the jet was, then, non – zero. The distribution of mass for the first found jet is shown in Fig. 12. It is , typically, small on the scale of the parent mass of 120 GeV. We have assumed that this procedure does not add to the error in the dijet mass determination.

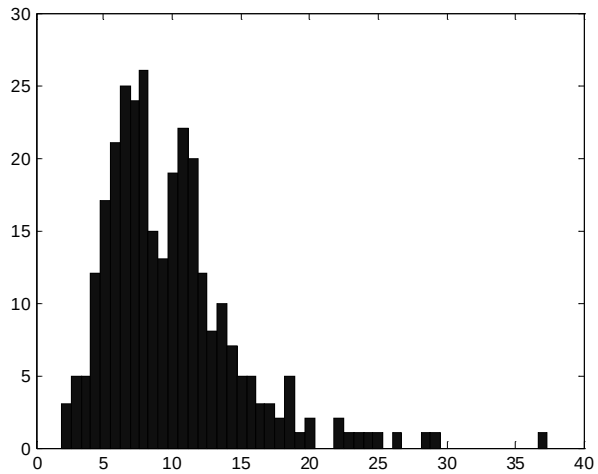

Figure 12: Distribution of jet mass for the first found jet in Z(120) events.

Dijet Mass

The dijet mass was calculated for the two largest transverse momentum clusters. The calculation was done both for pure calorimeter clusters and for the energy flow list consisting of clusters, matched charged tracks, and barrel loopers. The distributions shown here require that both jets have an axis within $|\eta| < 3$, roughly where tracking exists in CMS.

The event by event correlation between the mass calculated from clusters and from the energy flow list is shown in Fig. 13. Note that in most cases the latter calculation gives the larger mass. This fact is simply a reflection of the offset in the track momentum – calorimeter cluster energy match, Fig. 8.

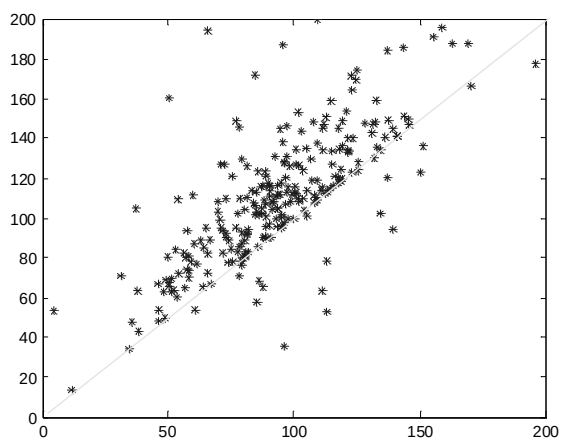

Figure 13: The correlation between dijet masses for Z(120) events calculated from calorimeter clusters and from clusters plus tracks (energy flow).

The distribution in dijet mass computed using only calorimeter clusters is shown in Fig. 14. The distribution calculated using all, the available information from calorimetry and tracking is shown in Fig. 15. Note that the latter mass is systematically larger than the former.

Gaussian fits were made to the histogrammed data. The mean mass is 84 GeV for the calorimeter clusters. The fractional mass error is 39%. For the case where matched tracks are used and loppers are added the mean mass is 108 GeV and the fractional mass error is 29 %. In both cases the Gauss Ian fit has a chi squared ~ 1, indicating an adequate fit. It appears that the energy flow technique improves the mass measurement. That indicated that the energy error on the charged hadrons in a jet makes a non-negligible contribution to the overall mass resolution of dijets.

Note that this study makes as complete a simulation of the CMS detector as is presently possible. It is of interest to compare the results of this study to the conclusions of earlier and simpler models of jet resolution.

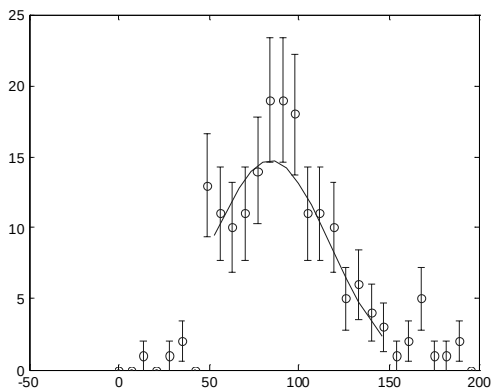

Figure 14: Mass distribution for Z(120) events using calorimeter clusters. The line is a Gaussian fit to the data over the mass range Z(50, 150) GeV.

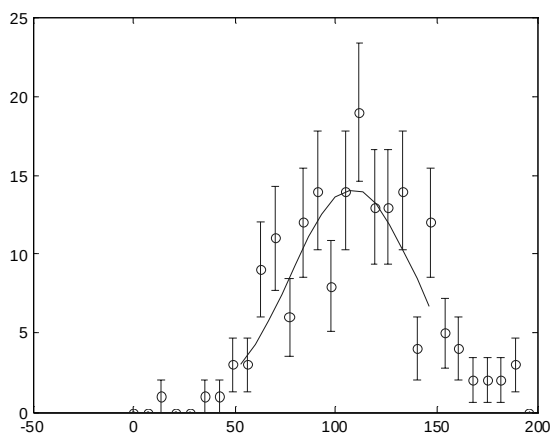

Figure 15: Mass distribution for Z(120) events using the energy flow list. The line is a Gaussian fit to the data over the mass range (50, 150) GeV.

Discussion

An early study [4] of the dijet mass resolution for W,Z masses at low transverse momentum revealed several competing effects. Using an assumed calorimetric resolution with a stochastic term of 60% and a constant term of 3% [2], the contribution of the energy error to the mass resolution was ~ 7.2 %. The contributions due to fragmentation and the underlying event depend on the cone size. For R = 0.6 the fragmentation error is 5.5 %, while the underlying event contributes 9.4 %. At a larger cone size, R = 1.0, the fragmentation outside the cone has a reduced 3.6 % contribution, while the fluctuations on the underlying event within the cone rise to 13.8%.

These conclusions were made in the absence of initial state or final state radiation. The classic early treatment of the emission of a single gluon by quarks [5] indicated that the problems caused by radiation could be severe. Consequently, a simple Monte Carlo program was written to look at gluon showers from quarks in the loading log approximation (LLA).

An Et = 60 GeV jet was generated and the effect of gluon showering could be isolated. The total quark plus gluon energy within a cone of radius R = 1.0 is shown in Fig. 16. Note that there is a substantial fraction of the events with observed energies < 50 GeV, indicating large effects due to radiation.

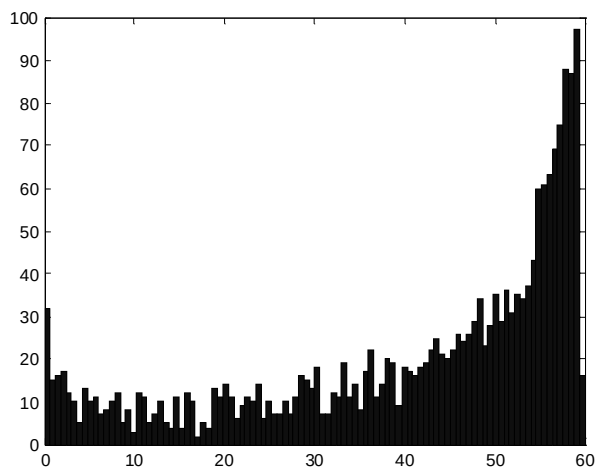

Figure 16: Distribution of the energy within a cone of radius R = 1.0 starting with a quark of energy 60 GeV and allowing it to shower gluons.

Events were studied for containment of energy within a cone. In Fig. 17 is shown the fraction of 60 GeV Et jets which have a fraction of that energy within a cone of R = 1.0 and 0.6. For example, with R = 1.0, only ~ half the jets have 80% containment of the initial energy. This full shower result is somewhat worse than the simple estimate [5] based on single gluon emission.

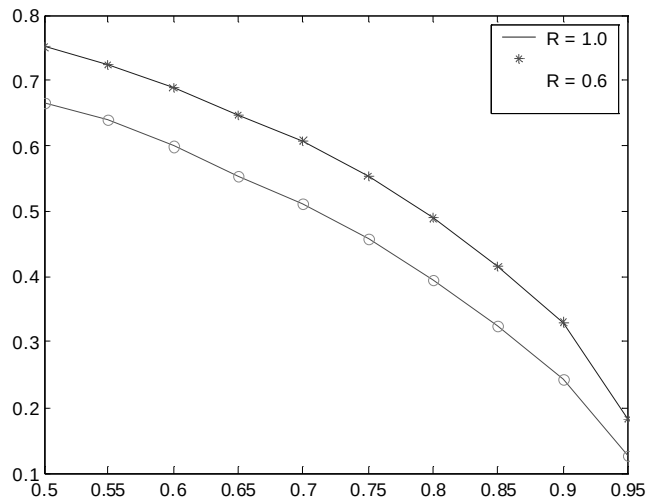

Figure 17: Fraction of jets where the initial energy is contained within a cone of R = 0.6 and 1.0 for different energy containment fractions.

The severity of the effect of radiation prompted a more complete study. This study has been made of the effects of initial state and final state radiation on the dijet mass resolution [7] for Z bosons. For a cone size of R = 0.7 the fractional mass error was 11% without radiation, rising to 19 % with radiation turned on. Unfolding in quadrature, radiation appears to contribute 15% to the mass error by itself. Note that in this study a toy detector was used.

The error without radiation compares well with the effects quoted above [4] folded in quadrature due to energy resolution, underlying event fluctuations, and fragmentation out of cone. These amount to 13% for R = 0.6. Folding in the effect of radiation, we expect a fractional mass resolution of 22% in the absence of the effects of the detector itself, such as inert material in the tracker, etc.

The Z mass resolution with and without radiation is shown in Fig. 18. The effect of radiation is to broaden the peak and induce a long, low mass, radiative tail into the mass distribution. The shape is reasonably consistent with that shown in Fig.15.

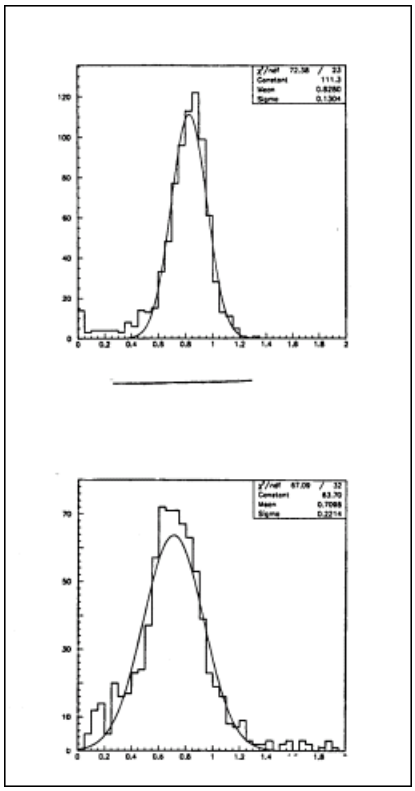

Figure 18: Z mass distribution with and without initial and final state radiation. Note that the distribution is the reconstructed mass divided by the generated mass. Radiation shifts the mean to lower values, broadens the peak, and induces a low mass radiative tail.

References

1. ECAL Technical Design Report, CERN/LHCC 97-33, December 1997
2. HCAL Technical Design Report, CERN/LHCC 97-31, June 20, 1997
3. V. Abramov et al., Nucl. Inst. Meth. In Physics Res., A 457 (2001) 75-100
4. A. Beretvas et al., Fermilab FN-626, October 1994
5. G. Sterman and S. Weinberg, Phys. Rev. Lett., 39, 23, 1977
6. D. Green, Fermilab, FN-666, February, 1998
7. K. Maeshima et al., CMS Note, 1998/026, February 28, 1998