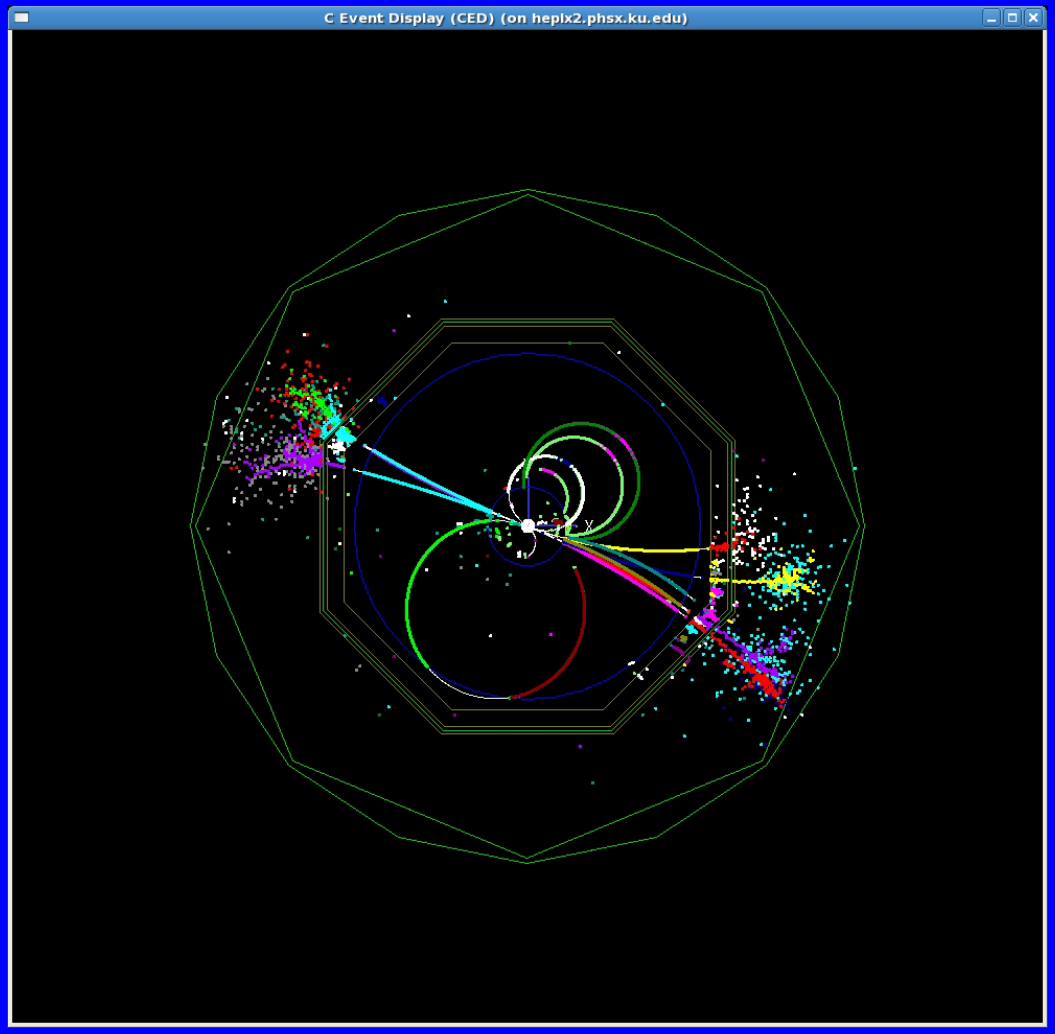
Event-Specific Hadronic Event Reconstruction



Graham W. Wilson, University of Kansas

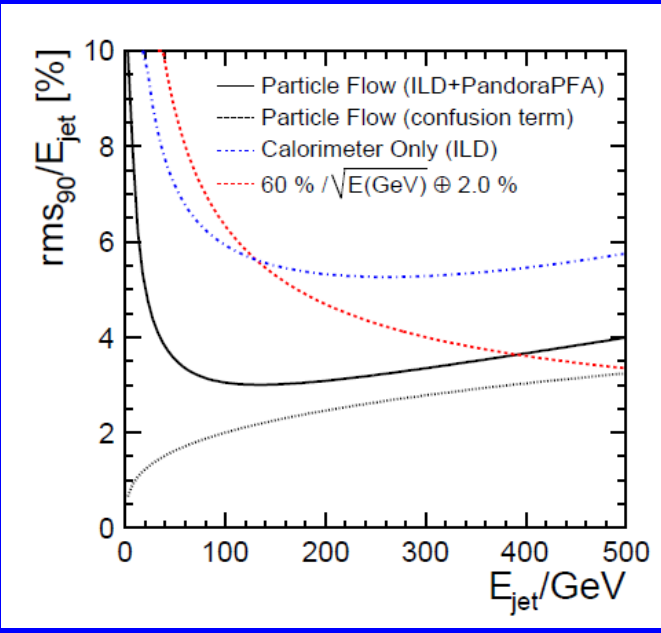
Outline

- Introduction
- Some plots on event-specific energy resolution
  - From ILD optimization talk last May
- Study of event-specific energy response function using multiply resimulated generator-level events

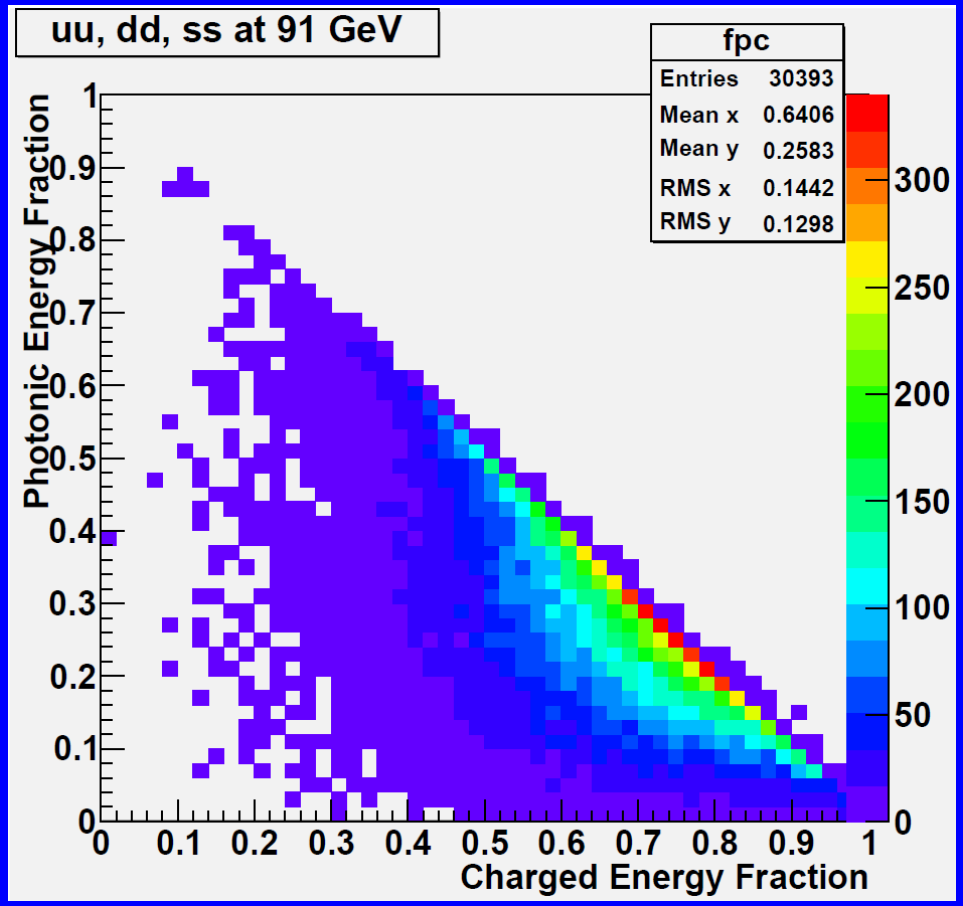
Introduction

- Over-arching goal: Understand and fully exploit particle-flow based reconstruction of hadronic jets.
- We have some characterization of the AVERAGE behaviour and estimates of the major component error sources based on both empirical fits to B,R,E dependence and by using MC truth info to "perfect" some components.
  - "intrinsic resolution", "confusion", leakage, "constant term".
- But estimating the error per jet or per event is at an early stage.

- Today, I will present an approach targeted at understanding one main thing.
- What is the visible energy response function for individual events ?
  - Is it Gaussian ?
  - Is there a significant energy bias event-to-event ?
  - Does the event-specific energy resolution vary significantly from event-to-event ?
  - Can we correct/exploit this and understand it ?



Event-Specific Energy Resolution



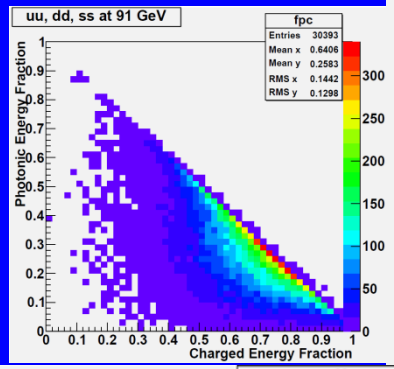
- We expect the overall particle-flow based energy resolution to depend strongly on the relative energy fractions in charged hadrons, neutral hadrons and photons in each event.
- Particularly at low energy where the intrinsic resolution is supposedly dominant and confusion is not so important.
- The overall rms_{90} corresponds to 25.2%/√E per 45 GeV jet.

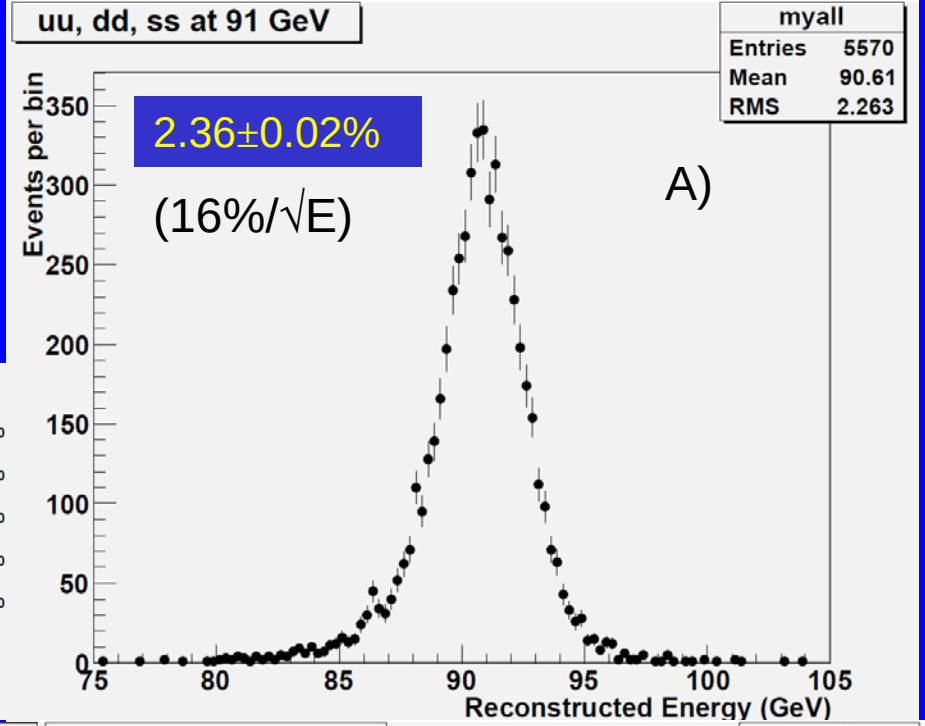
Here, the events are divided into 3 distinct classes based on reconstructed neutral hadron energy.

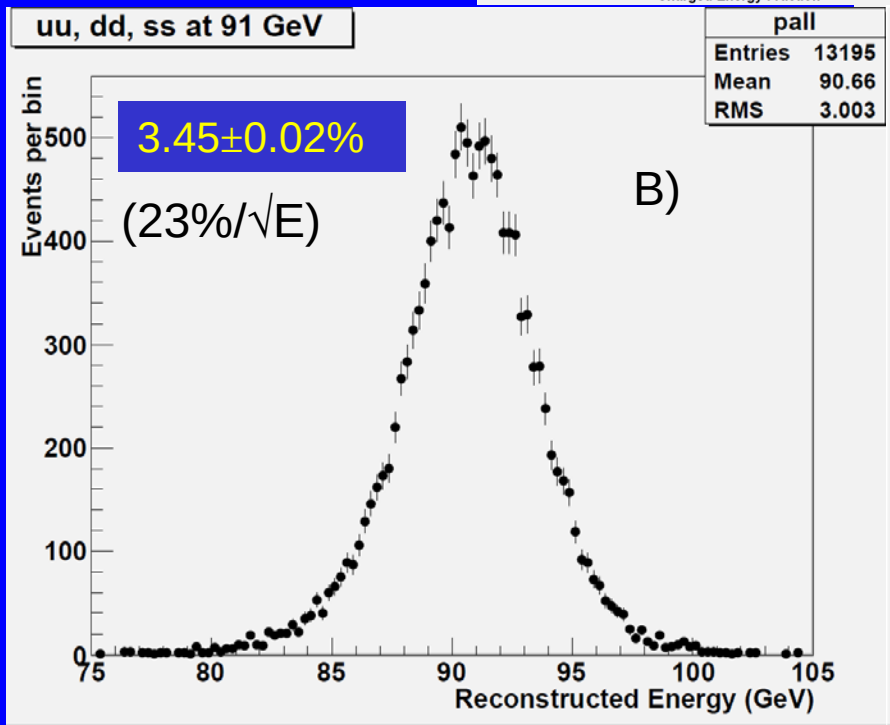
A) ENH < 2 GeV
B) 2 < ENH < 10 GeV
C) ENH > 10 GeV.

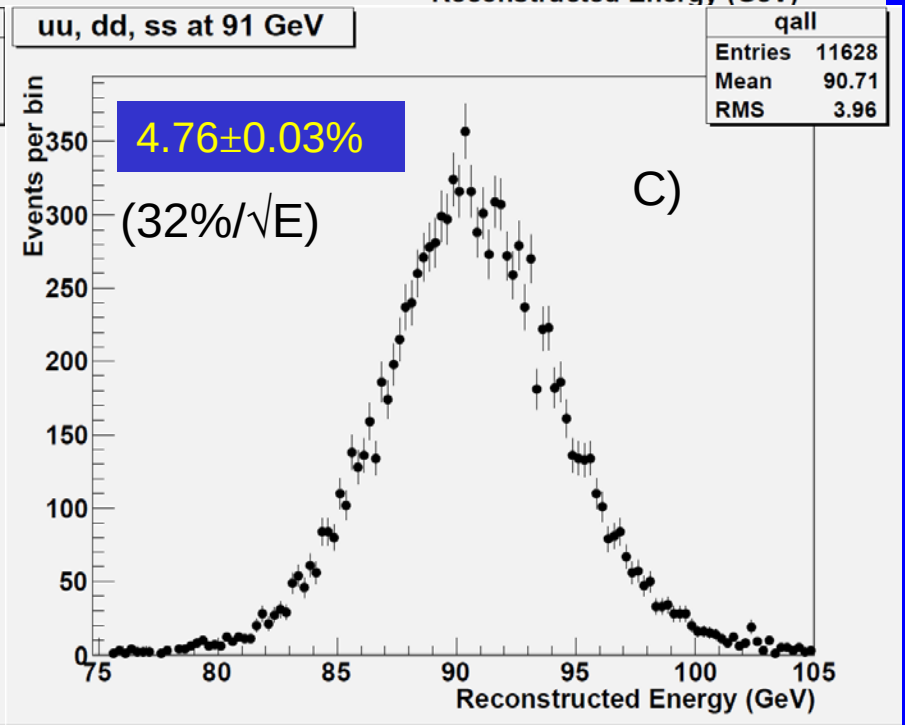
Significant differences based on MEASURED quantity: can be exploited immediately. Rather Gaussian distributions.

rms90 values for $|\cos\theta| < 0.7$ and 45 GeV jets

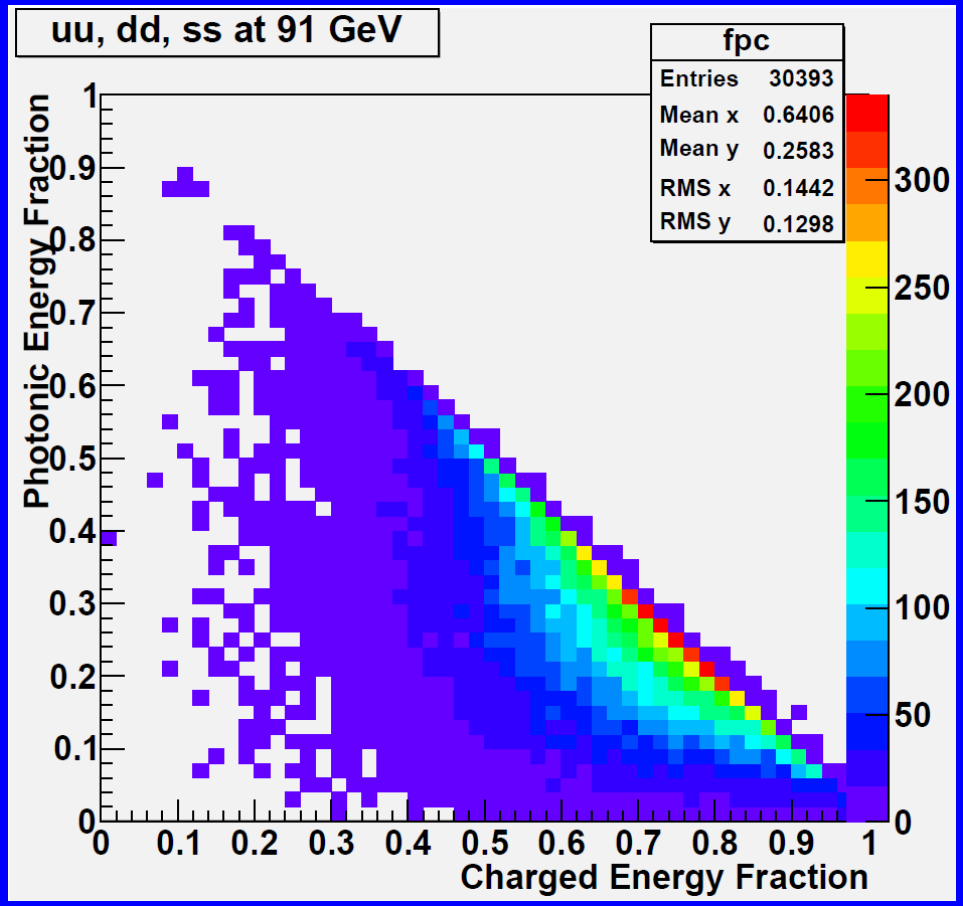








Event-Specific Energy Resolution



- We expect the overall particle-flow based energy resolution to depend strongly on the relative energy fractions in charged hadrons, neutral hadrons and photons in each event.
- Particularly at low energy where the intrinsic resolution is supposedly dominant.
- **Indeed the energy resolution does vary significantly.**
  - **And it is feasible to start classifying events/jets into resolution categories.**

Event-Specific Study

- Today, concentrate on the results from a study where individual generator level events (light-quark Z decays) are passed through the full detector simulation (Mokka) again and again with different random number seeds.
- These re-simulated individual events sample specific points in the (fγ,fch) generator-level Dalitz plot.
  - The generator level information is the same for each re-simulation including things like the decay kinematics of short-lived particles (π^0, K^0_S, Λ).
  - The stochastic components of the energy response is different
    - Calorimeter response
    - But also, interaction points of stable particles: photon conversion, hadron interaction, decay in flight, V0 decay

Event-Specific Study

- Use two MC data-sets.
- Each with 100,000 Z0 simulations.
- A) 100 generator events resimulated 1000 times.
- B) 1000 generator events resimulated 100 times.
- NO cuts on cosθT – full solid angle acceptance.

- A) gives a more detailed view of the individual event response function but for a small statistics sample of events.
- B) more representative sample for drawing conclusions but with higher statistical error in each event.

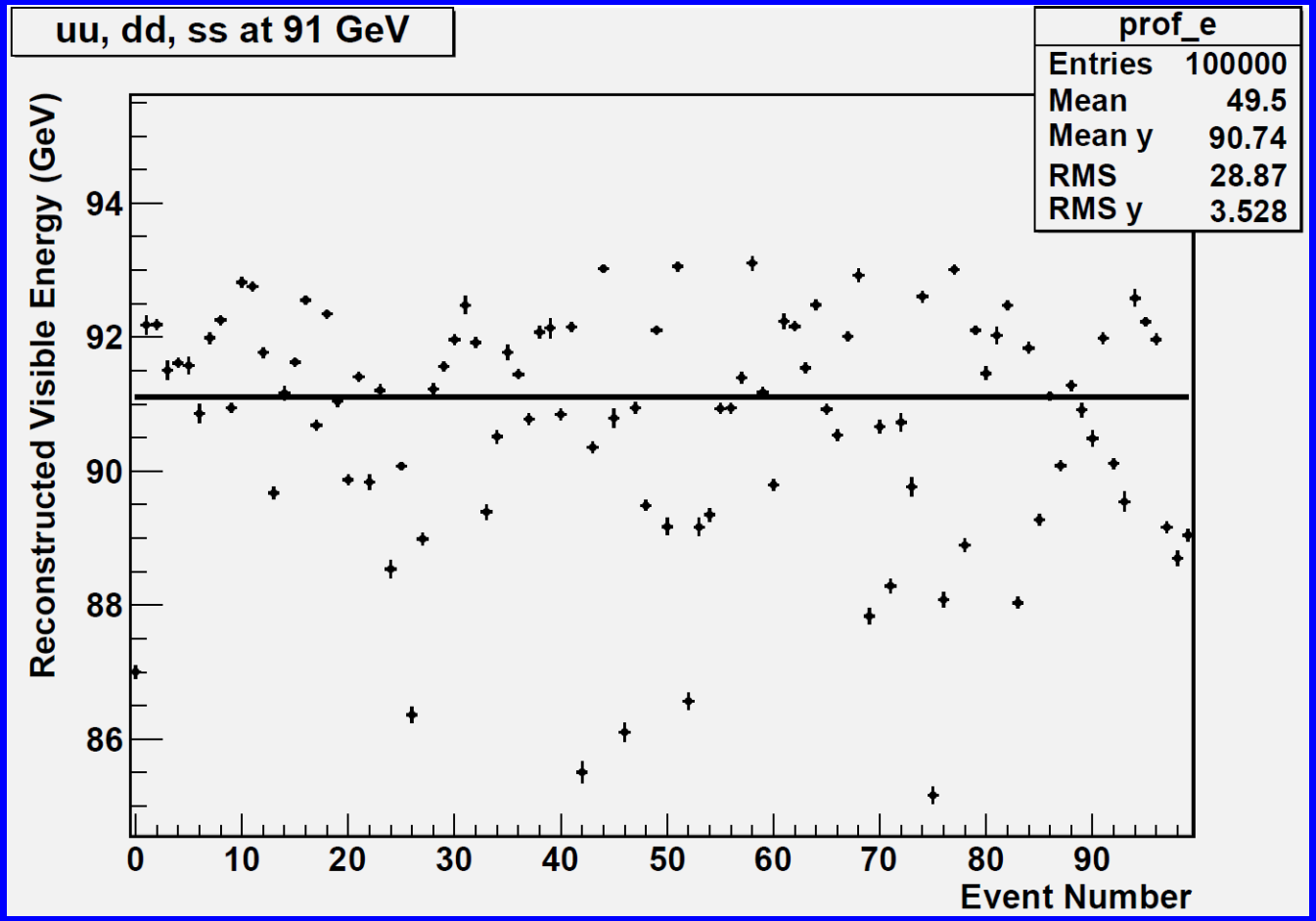
In each case, simulated events are then reconstructed using various versions of Marlin and PandoraPFANew.

Versions

- cambridge_uds.stdhep files
  - Sample A, events 0-99 from file 000
  - Sample B, events 0-999 from file 010
- ilcsoft v01-09 Mokka, ILD00 detector
- reconstruction using the following tags
  - MarlinReco v00-18-02
  - PandoraPFANew tag-1.28
  - MarlinPandora tag-1.13

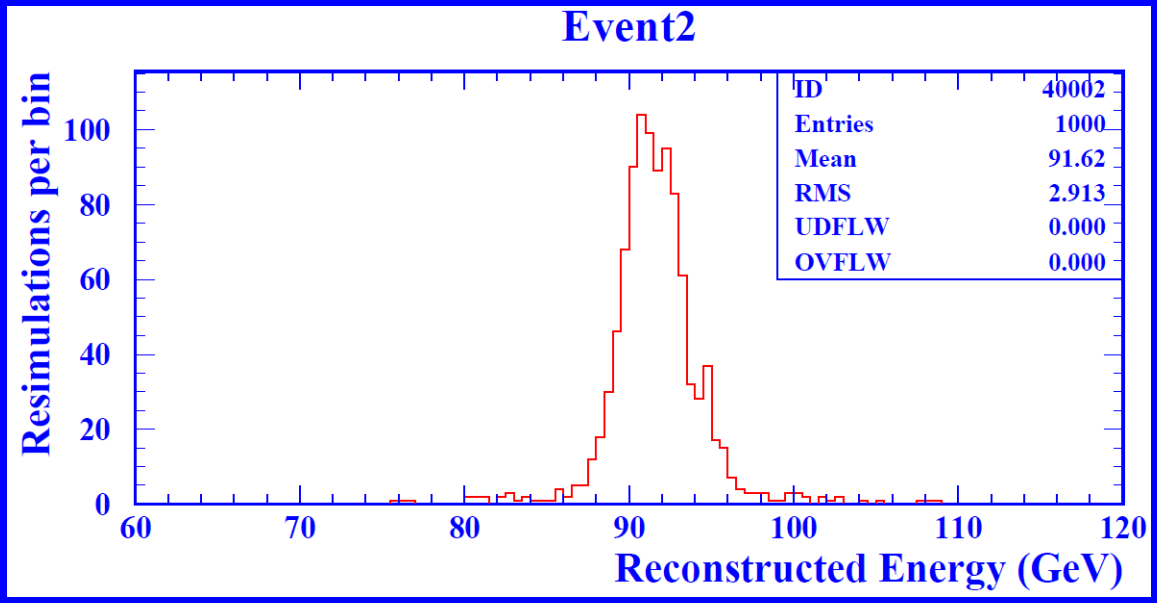
Sample A.

Mean for 1000 resimulations.



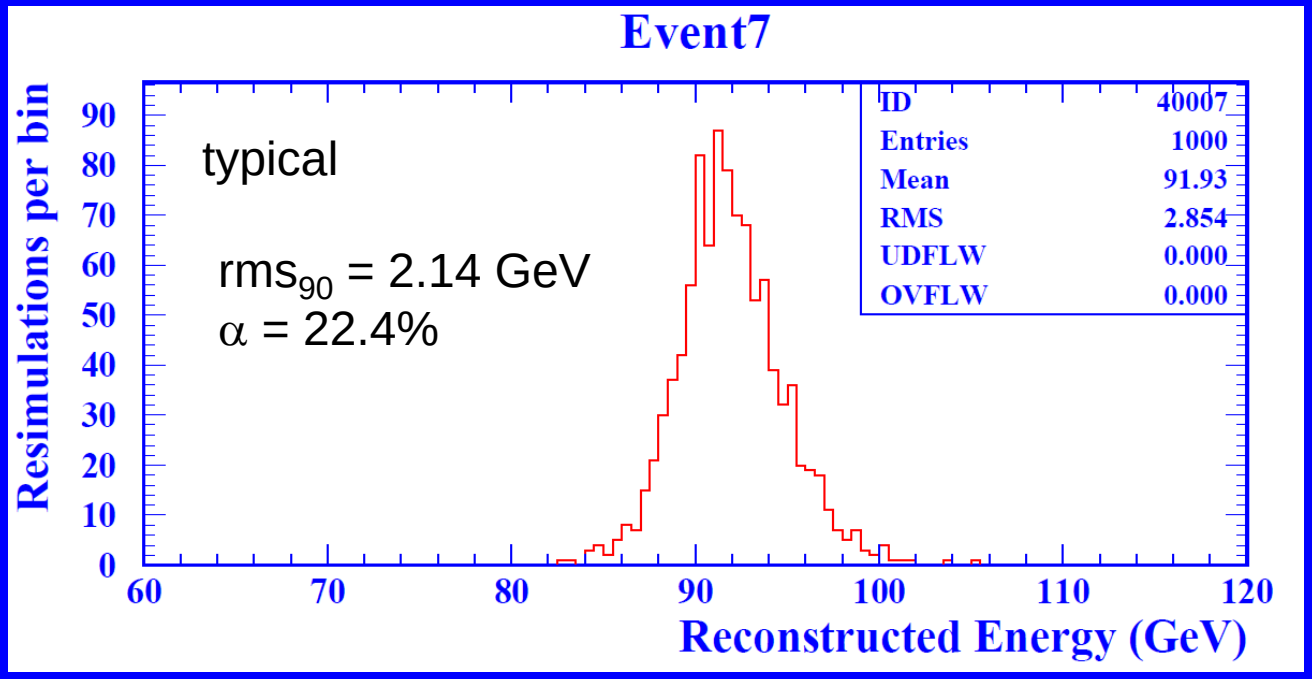
Significant variation in the mean reconstructed energy.

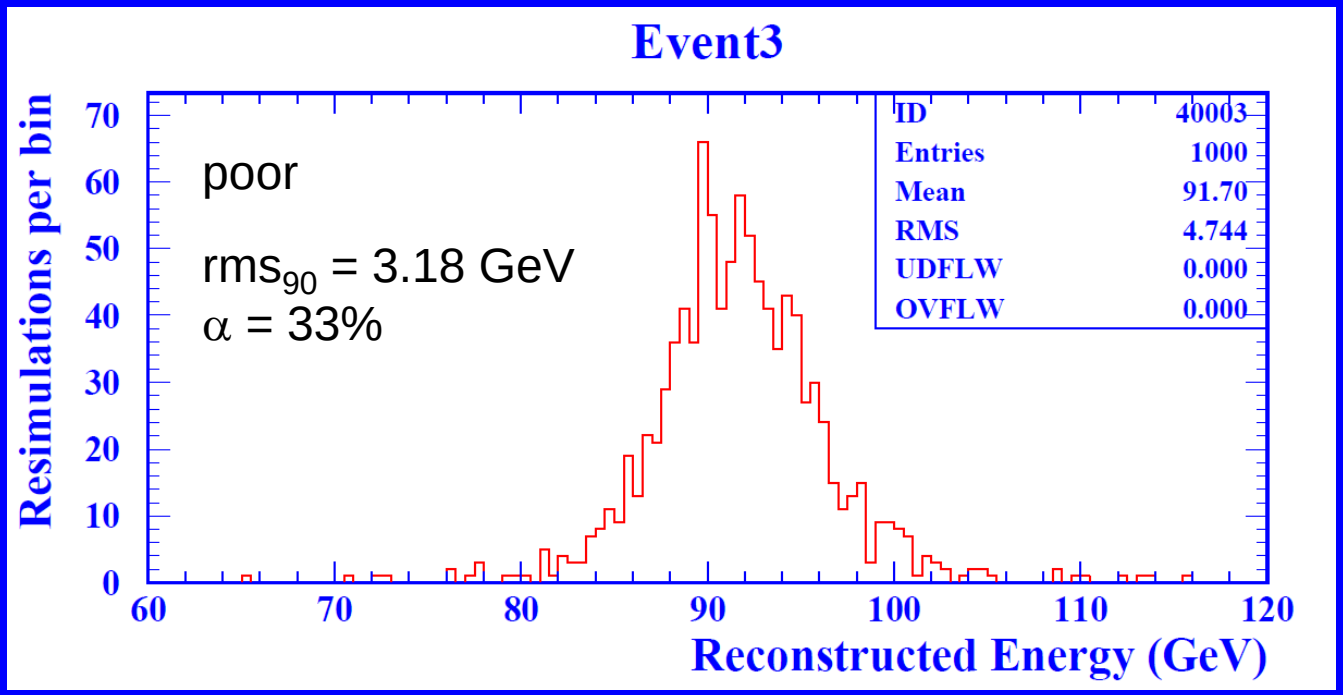
Resolution and Mean Estimates

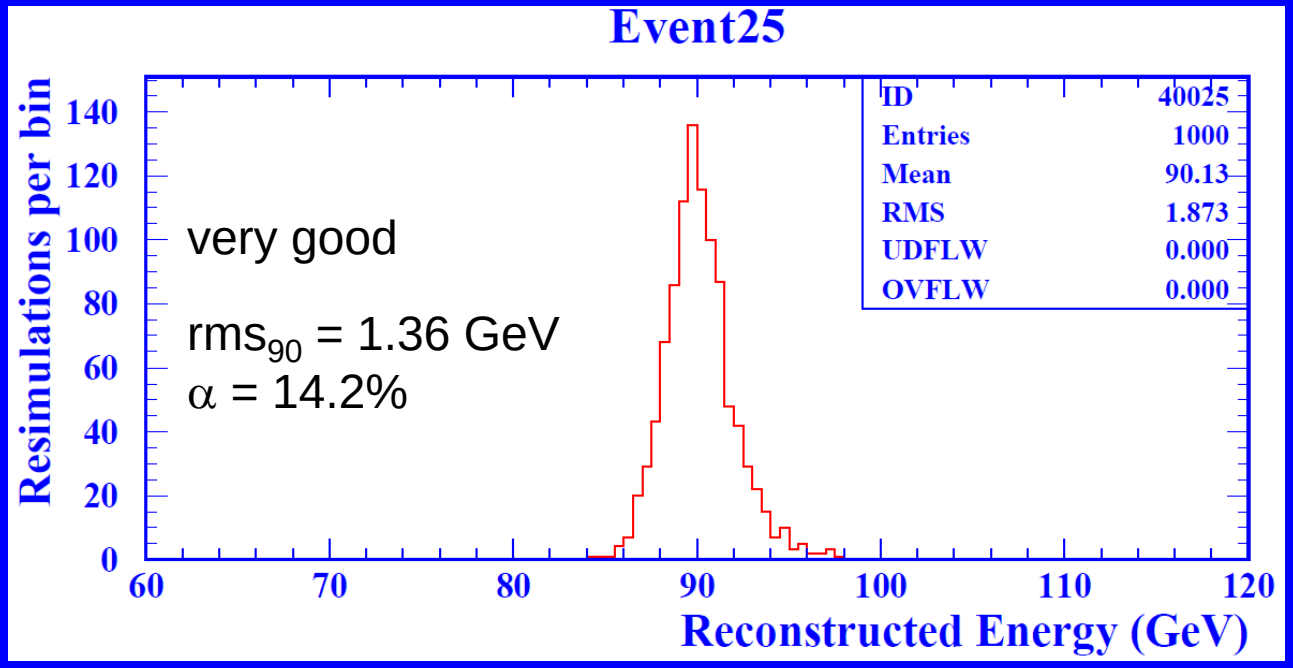


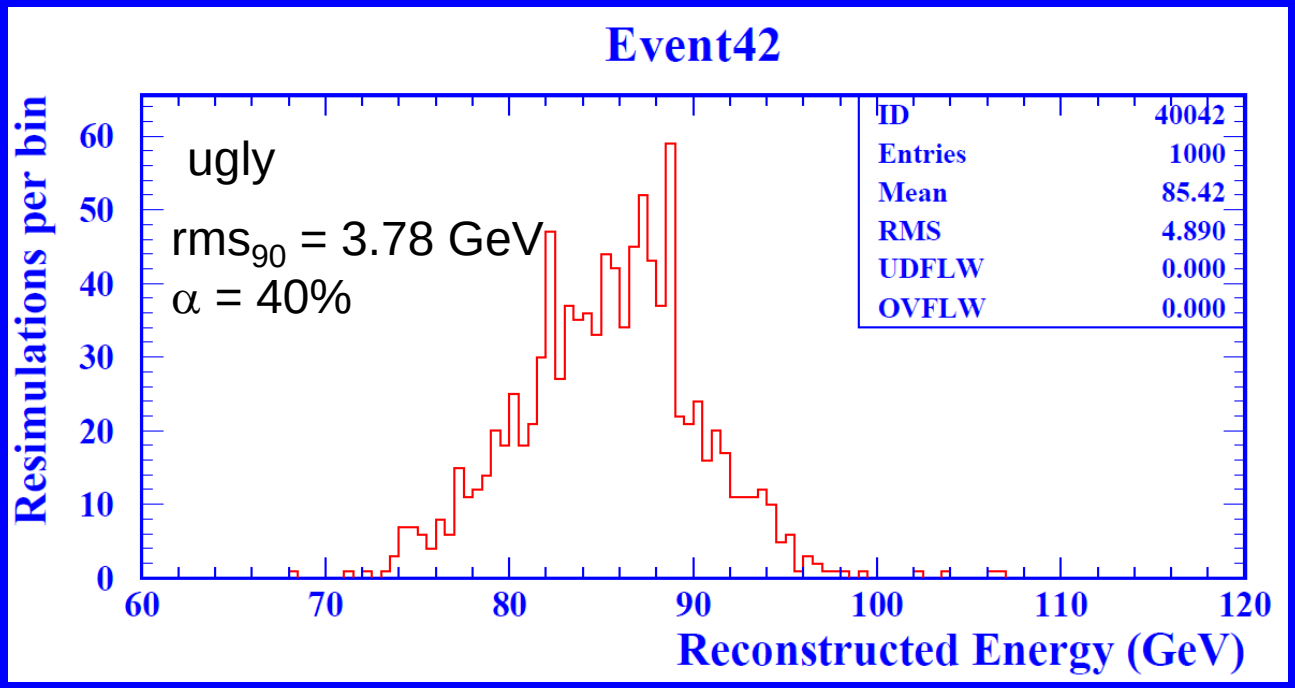
- In order to obtain robust resolution estimates I have followed an approach similar to rms_{90}.
- But have relatively low statistics per generator level event.
  - So have used a symmetrically trimmed rms using the central 90% of simulated events, where the event measurements are ordered and the lower and upper 5% tails are discarded.
  - Have used the same approach for the event-specific bias calibration: using a trimmed mean using the central 90%.

Examples

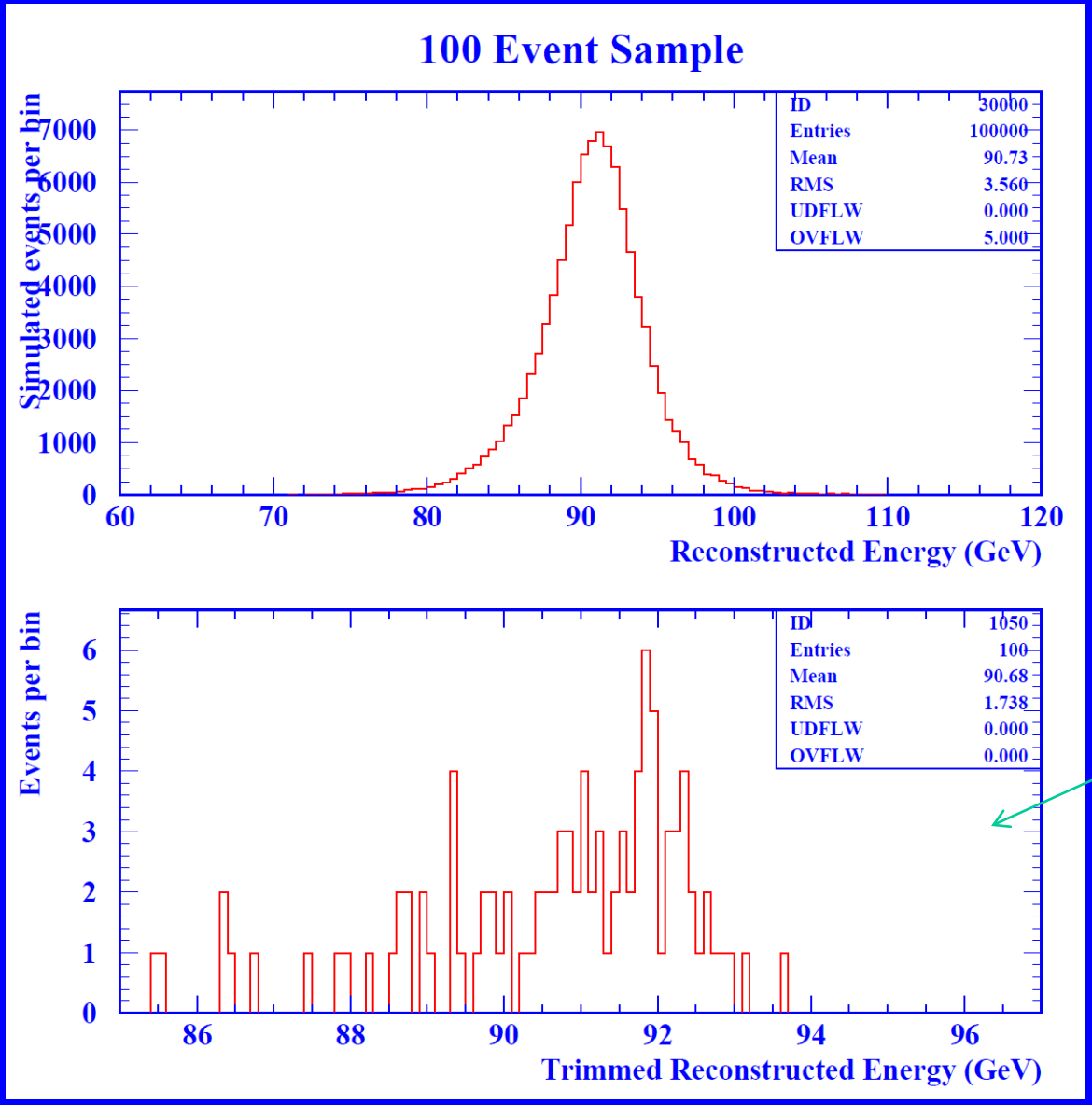

Event7
typical
rms90 = 2.14 GeV
α = 22.4%
ID 40007
Entries 1000
Mean 91.93
RMS 2.854
UDFLW 0.000
OVFLW 0.000
Resimulations per bin
Reconstructed Energy (GeV)


Event3
poor
rms90 = 3.18 GeV
α = 33%
ID 40003
Entries 1000
Mean 91.70
RMS 4.744
UDFLW 0.000
OVFLW 0.000
Resimulations per bin
Reconstructed Energy (GeV)


Event25
Resimulations per bin
Reconstructed Energy (GeV)
very good
rms90 = 1.36 GeV
α = 14.2%
ID 40025
Entries 1000
Mean 90.13
RMS 1.873
UDFLW 0.000
OVFLW 0.000


Event42
Resimulations per bin
Reconstructed Energy (GeV)
ugly
rms90 = 3.78 GeV
α = 40%
ID 40042
Entries 1000
Mean 85.42
RMS 4.890
UDFLW 0.000
OVFLW 0.000

Sample A summary

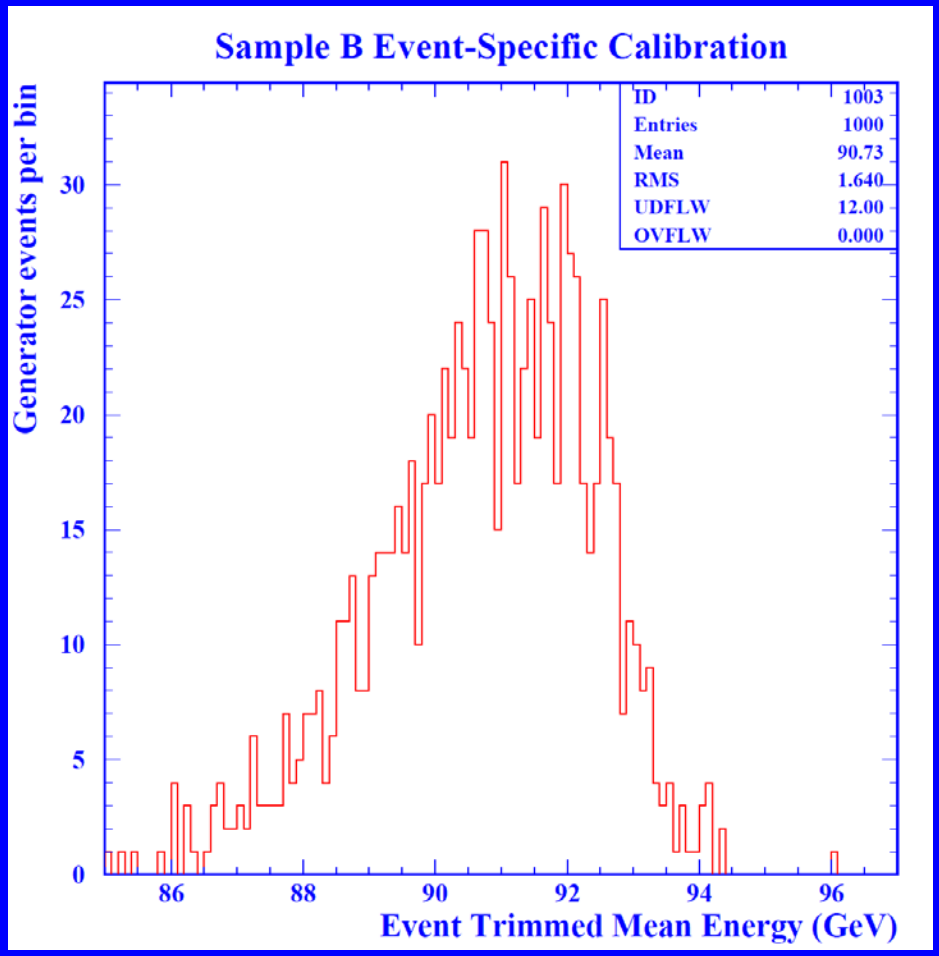


Significant contribution to overall resolution from event-specific bias.

Raw rms here is equivalent to 18%/√E.

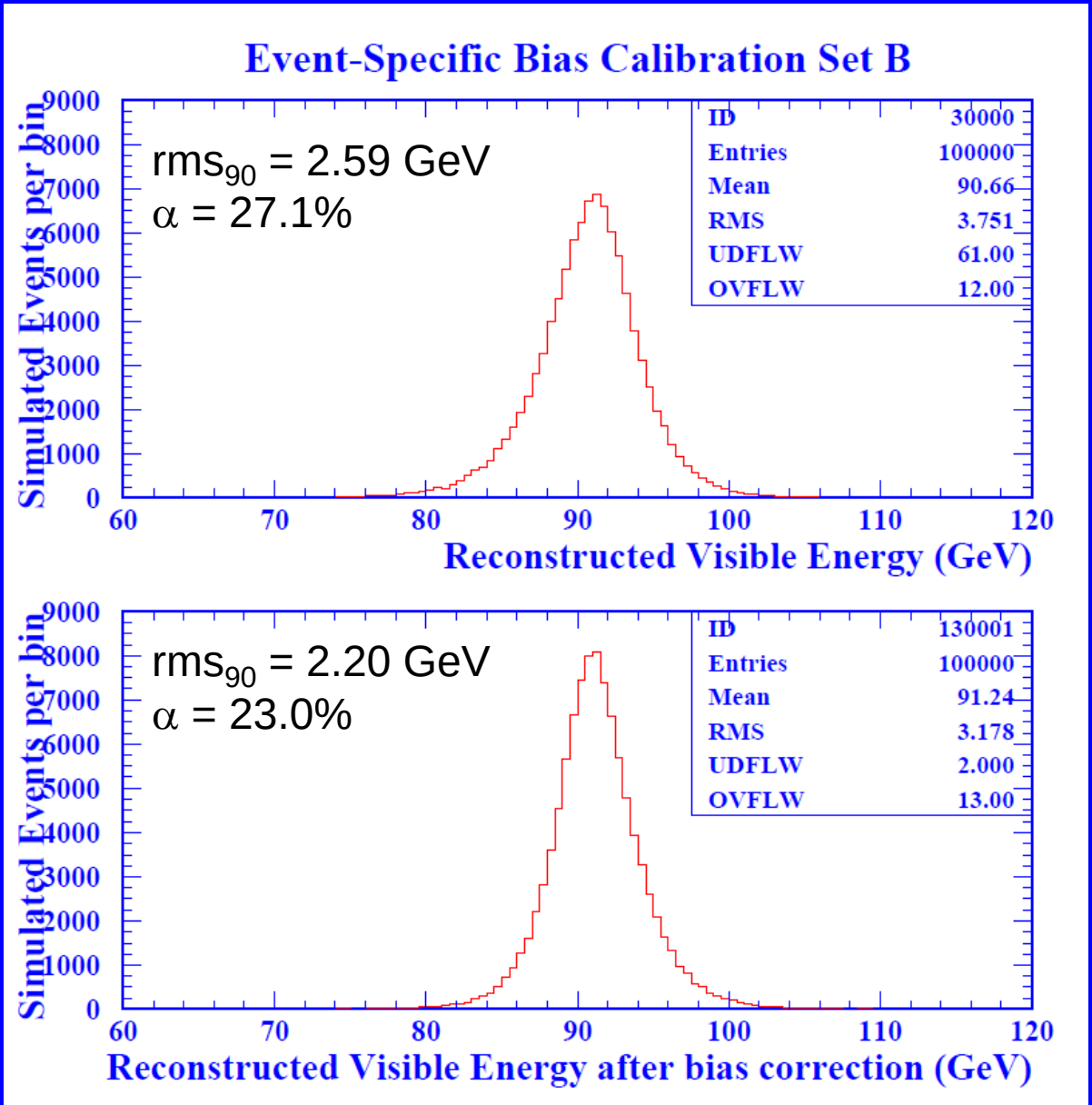
(Here use the symmetric 90% "trimmed mean" for each event to calibrate each event to 91.2 GeV)

Sample B event-specific bias



Raw rms equivalent to 17%/√E

Event-Specific Bias Correction



By correcting for the event-specific bias, the overall rms_{90} based resolution estimate is reduced by a factor of 0.847.

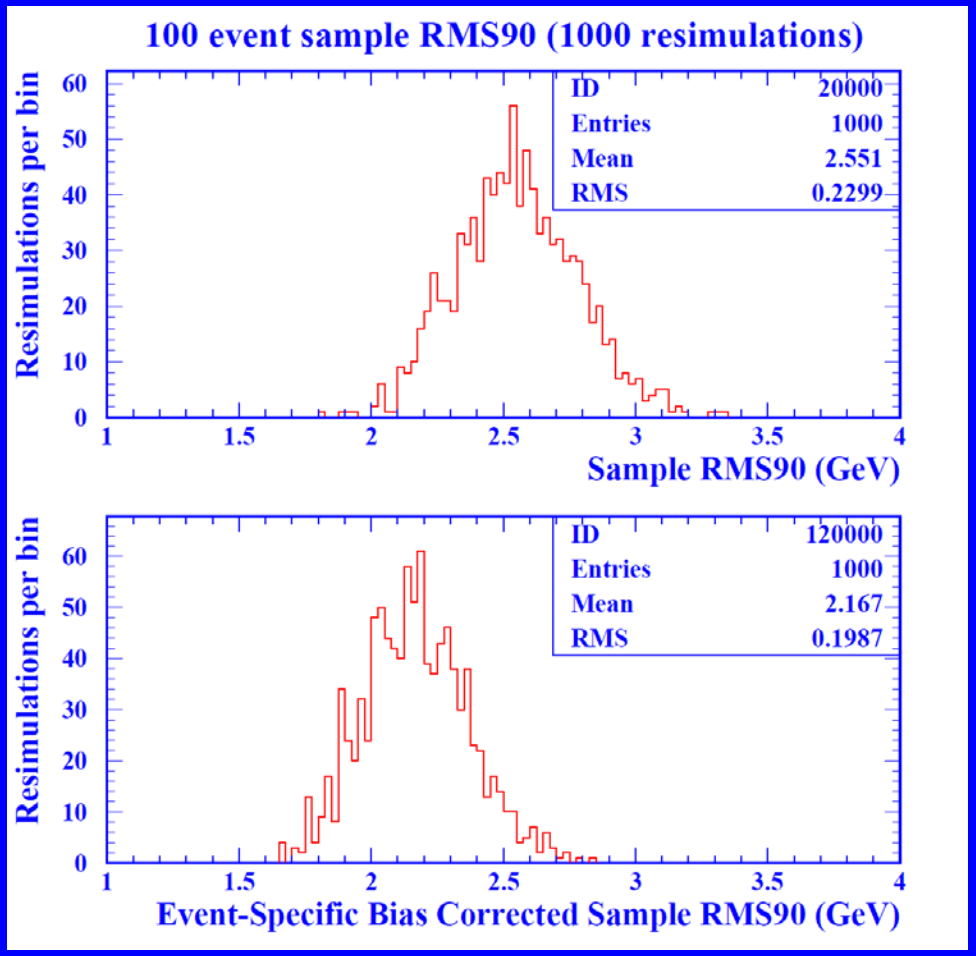
This corresponds to removing a component of resolution of around14.4%/√E.

Sample A

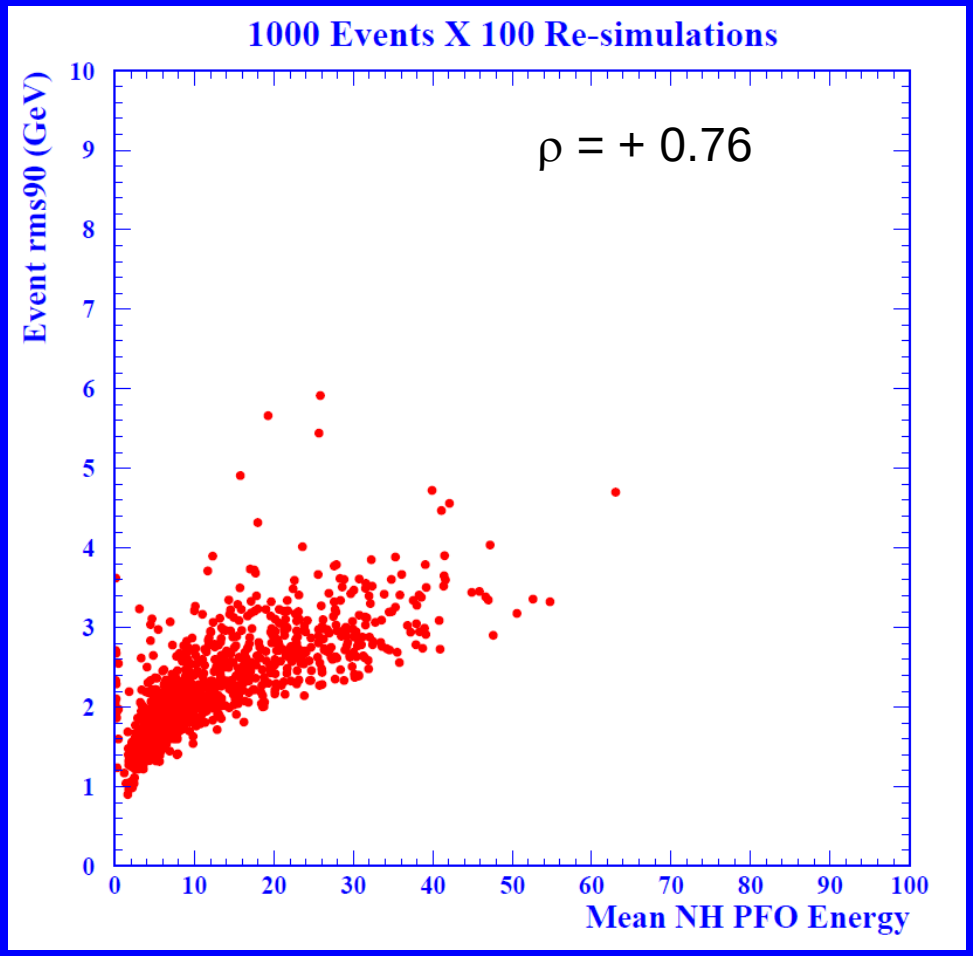
Independent cross-check.

Examine the sample rms_{90} for the 100 generator events for each of the 1000 resimulations.

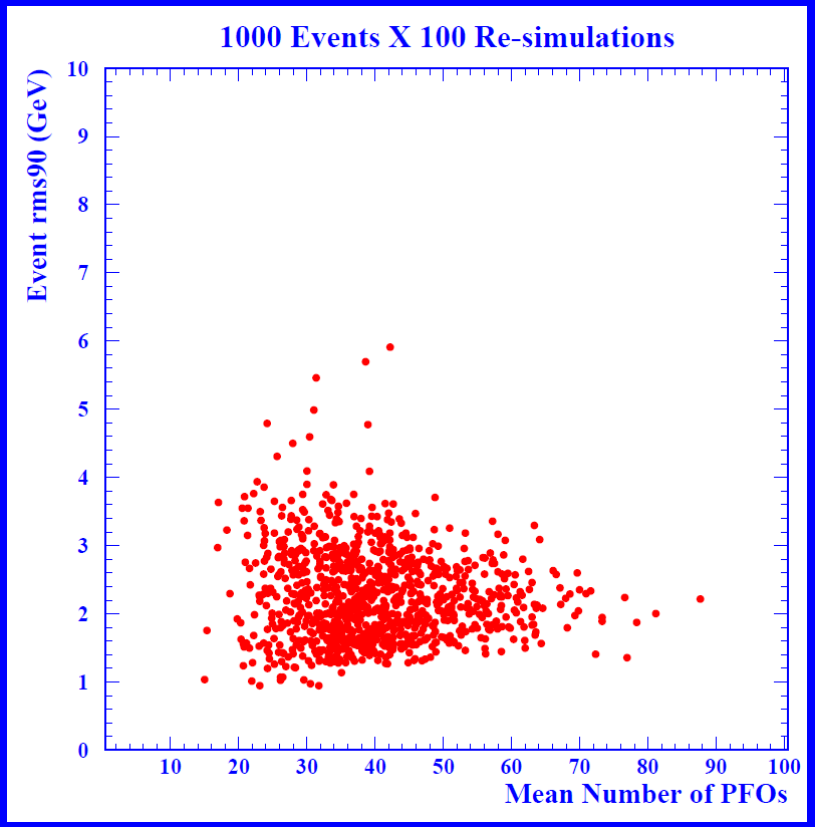
Again factor of 0.85 improvement after event-specific bias correction (note the statistical errors on rms_{90} are quite large with each 100 event ensemble).



rms_{90} correlations with event characteristics



Resolution **does** depend significantly on neutral hadron energy



Resolution does not seem to depend significantly on PFO multiplicity at this energy

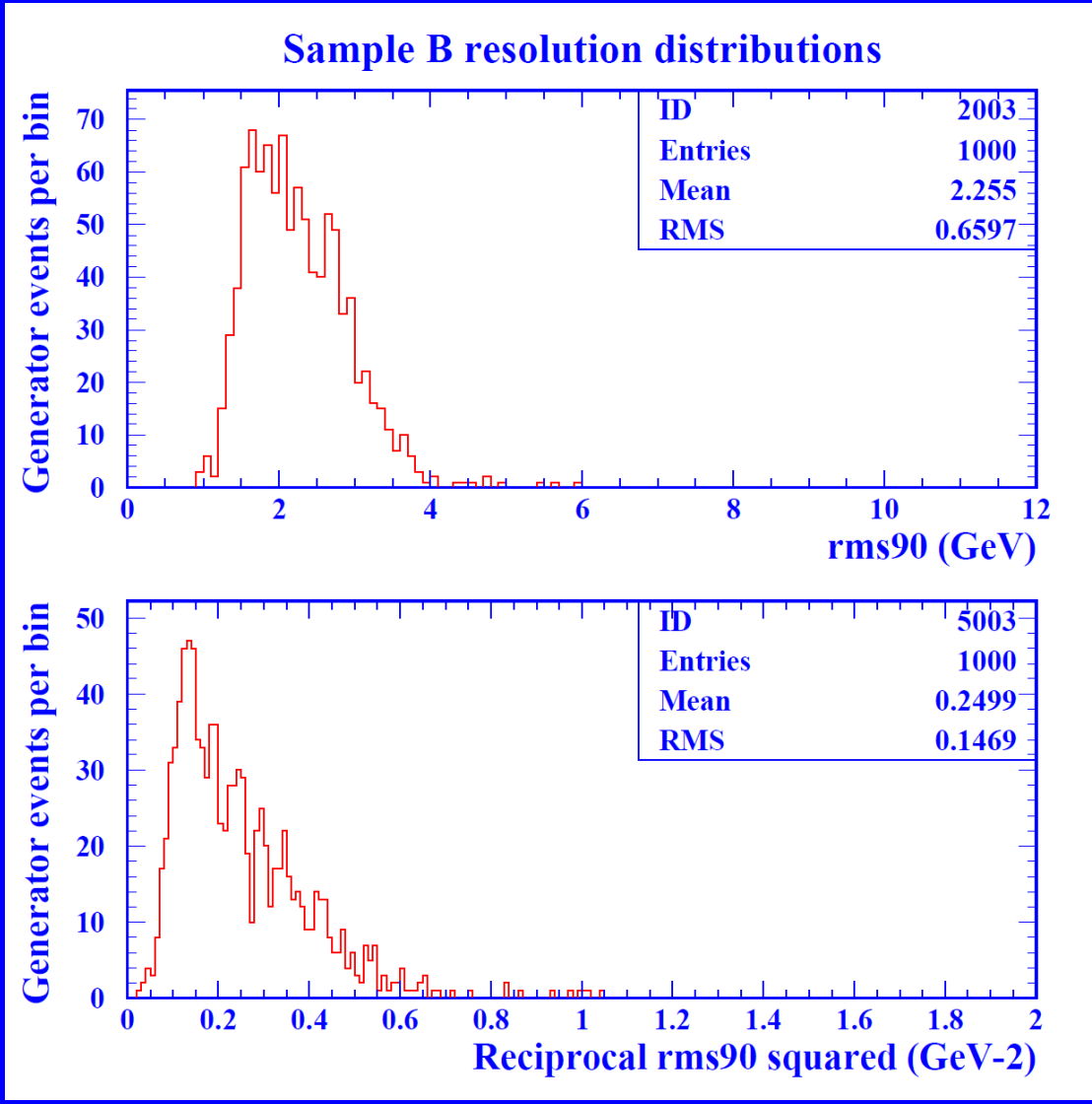
Event-Specific Resolution

1000 generator level events.

Measure the event-specific rms_{90} for the 100 resimulations (10% stat. errors ..)

Each event is quite different. This is in contrast to the situation with a classic calorimeter, and can potentially be exploited if we can estimate the error properly on an event-by-event basis.

If we form a weighted average, the ML estimate using Gaussian errors, is to weight with $1/\sigma_i^2$
So an effective resolution per event, σ^*, can be calculated from $\sqrt{(1/< 1/\sigma_i^2 >)}$ giving 2.00 GeV equivalent to α = 20.9% which is 0.89 of $<\sigma i>$



Observations on response function

- Is it Gaussian ?
  - No, but may often be a good enough approximation
- Is there a significant energy bias event-to-event ?
  - YES, and it is surprisingly big
  - Can improve resolution by a factor of 0.85 if we can correct it
- Does the event-specific energy resolution vary significantly from event-to-event ?
  - YES, but this was not too surprising
  - Can improve event resolution by another independent factor of 0.89 if we can correctly weight events
- Can we correct/exploit this and understand it ?
  - Open question. Needs more work.

What is going on ?

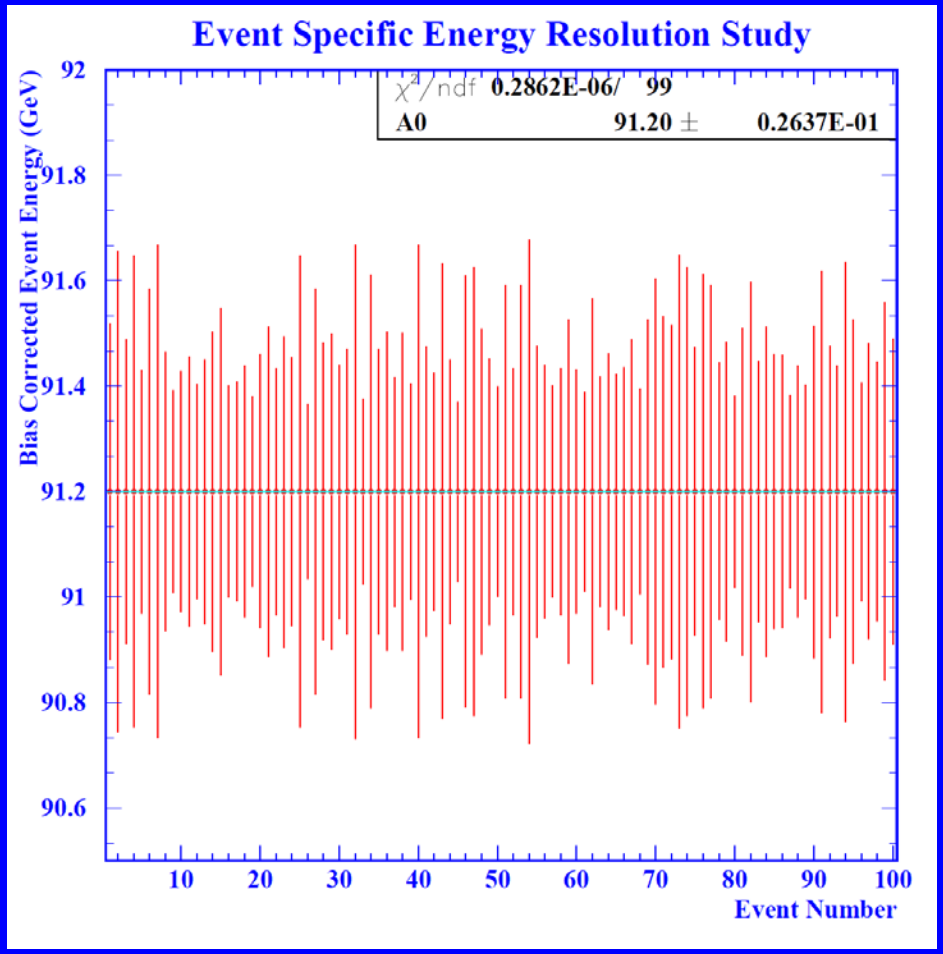
- 4-vectors consist of decayed short-lived particles. For example, if a prompt 4 GeV pi0 decays to a 3 GeV and 1 GeV photon, this happens the same in every event. Likewise for Kshort, Lambda.
- What does change is the decay length of secondary vertices, and the random interactions of the “stable” particles:
  - Photon conversions
  - Hadron interaction points
  - Decay in flight
  - Multiple scattering
- There may be some element of local response inhomogeneity, or particle loss, behind the event-specific bias.
- My naïve expectation was that the single event resolution is driven primarily by the Dalitz plot energy fractions and that confusion would be minimal. This does not seem to be the case. The typical single event resolution appears much bigger than the intrinsic resolution.
  - “Confusion” is a complicated function of not only the close by (pi- / n) 4-vectors but the specific interaction depths and hadronic shower characteristics.

Summary & Outlook

- A new way to look at evaluating the particle flow performance is presented.
- For 45 GeV jets, the ILD00 detector and recent PandoraPFANew the potential improvement factor from correcting event-specific bias and correctly weighting individual events is about 0.76.
  - Currently looking at event energy: $\sigma_E^2 = \sigma_{j1}^2 + \sigma_{j2}^2$
  - Eventually would like to apply to individual jets where the information can be used to improve physics performance: example jet pairing problem where typically one uses a χ^2 matching procedure.
- Plan:
  - Look at higher energy jets
  - Use specific events in on-going π^0 studies
- I expect that detector designs and algorithms which can minimize/correct event-specific energy bias and optimize event-specific energy resolution will have better overall physics performance.
- Finding the main drivers of event-specific performance, may be a promising way towards understanding and further improving particle-flow.

Backup Slides

Illustrating weighted mean



σ* is the same as what one does fitting a constant to data with errors.

Effective resolution

- Weighted mean: $$< x >= \sum_i w_i x_i \Big/ \sum_i w_i$$

Where $$w_i = \sigma_i^{-2}$$

- Leading to a variance on the weighted mean of $$\sigma^2 = 1 \Big/ \sum w_i$$

- The effective resolution, σ*, per event, is then given simply by σ* = 1 / √<wi >

See plots for all events at

- http://heplx3.phsx.ku.edu/~graham/Event_Specific/

- AllEvents.pdf (Sample A)
- AllEvents_1000.pdf (Sample B)