Software Compensation / Weighting in the CALICE AHCAL

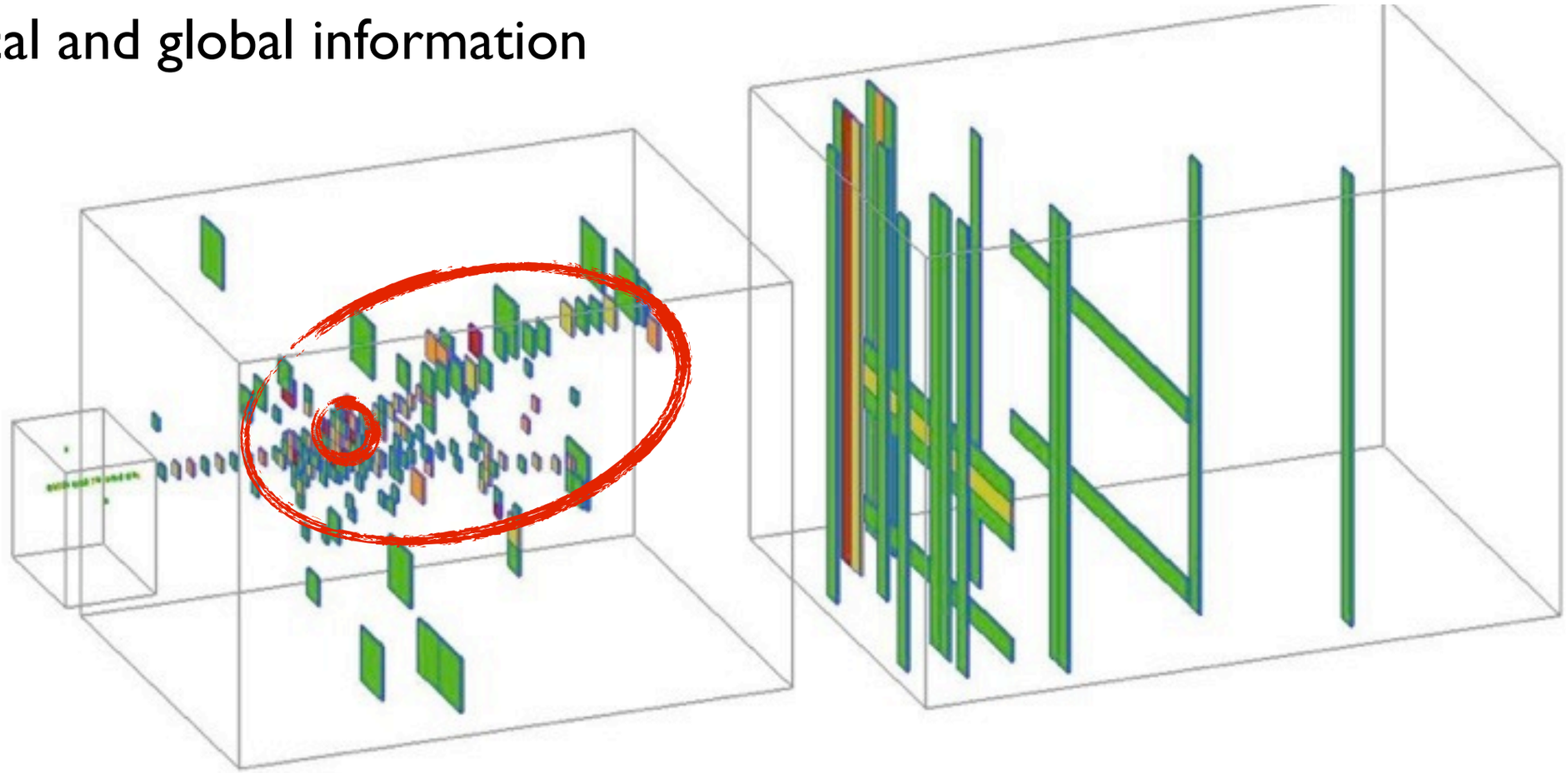
Frank Simon
MPI for Physics & Excellence Cluster 'Universe'
Munich, Germany

for the CALICE Collaboration

CALICE: Software Compensation Studies

Detailed images of hadronic showers: Resolution on the level of
1 X_0 longitudinally, 1 R_M laterally, Non-compensating calorimeter ($e/\pi \sim 1.2$)

⇨ Excellent conditions for software compensation / weighting techniques using local and global information



Techniques investigated so far:

Local compensation: Weights applied on the cell level - Data driven weights

Global compensation: One weight per shower - Data driven & MC weights

Local Software Compensation
Data Driven

Local Compensation - Technique

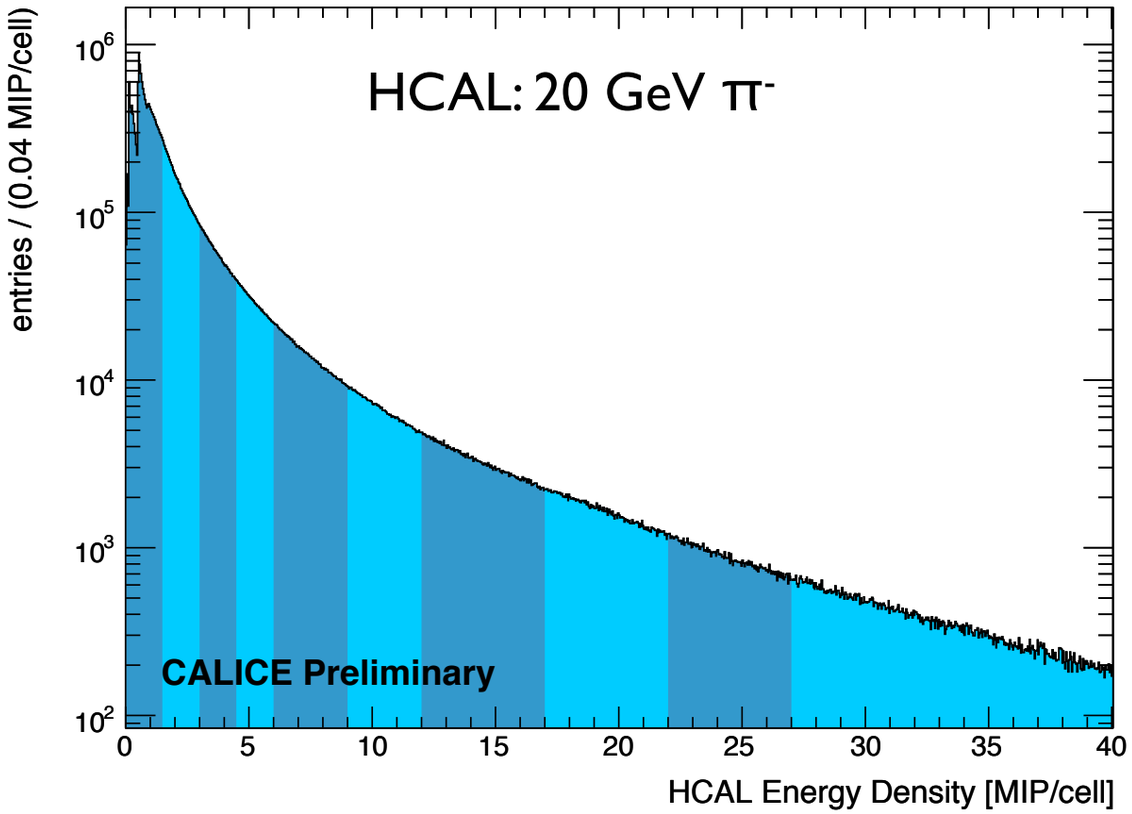
- Underlying assumption: Electromagnetic sub-showers tend to have a higher local energy density than hadronic ones
  - Apply weights based on the energy content of each cell

Local Compensation - Technique

- Underlying assumption: Electromagnetic sub-showers tend to have a higher local energy density than hadronic ones
  - Apply weights based on the energy content of each cell

- Group cells into bins of energy density (amplitude / volume), apply weight by bin
- Weights determined from data with a minimization technique:

$$\chi^2 = \sum_{events} \left(\sum_i E_i \omega_i - E_{beam} \right)^2$$



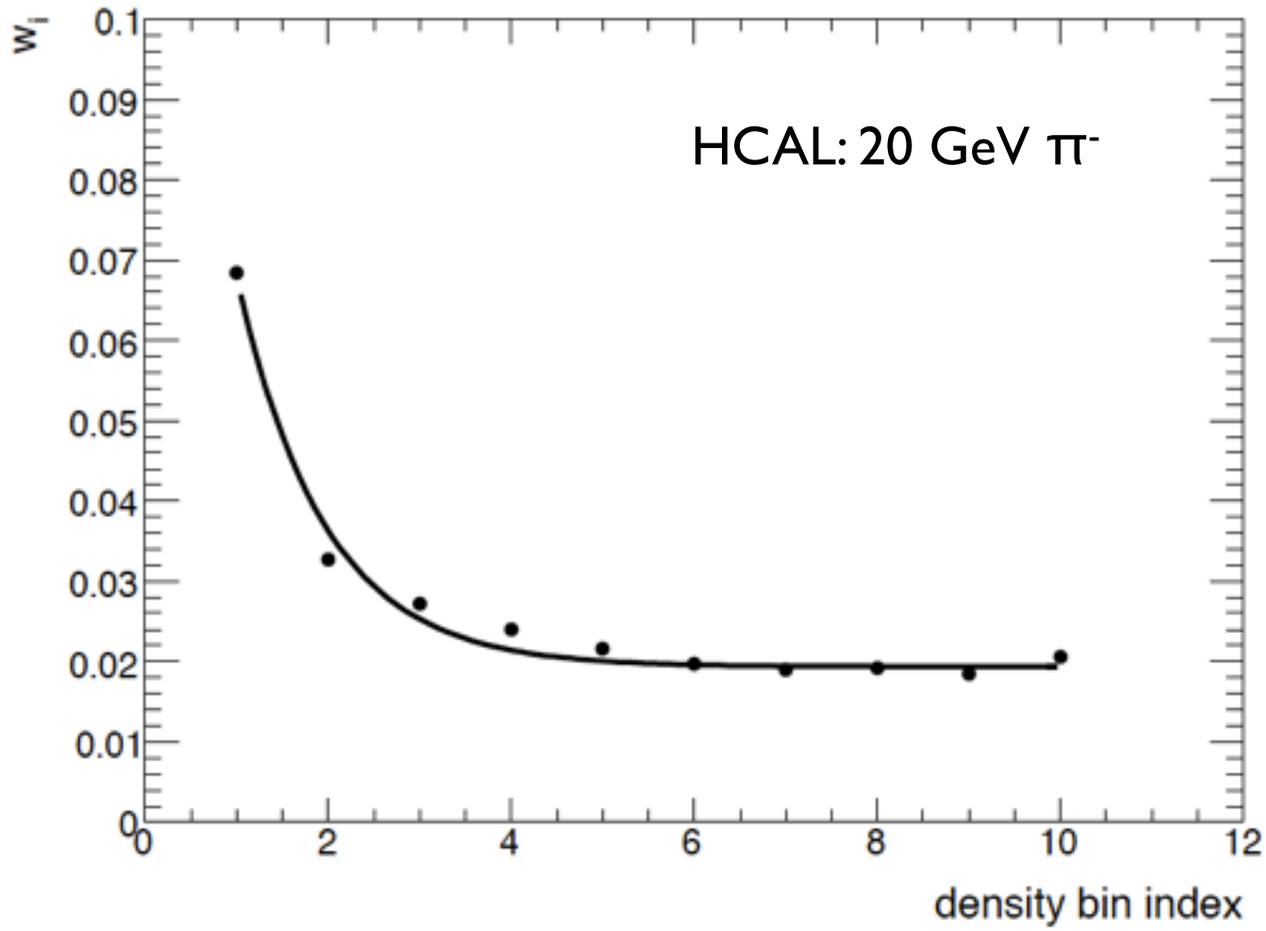
Local Compensation - Technique

- Underlying assumption: Electromagnetic sub-showers tend to have a higher local energy density than hadronic ones
  - Apply weights based on the energy content of each cell

- Group cells into bins of energy density (amplitude / volume), apply weight by bin
- Weights determined from data with a minimization technique:

$$\chi^2 = \sum_{events} \left(\sum_i E_i \omega_i - E_{beam} \right)^2$$

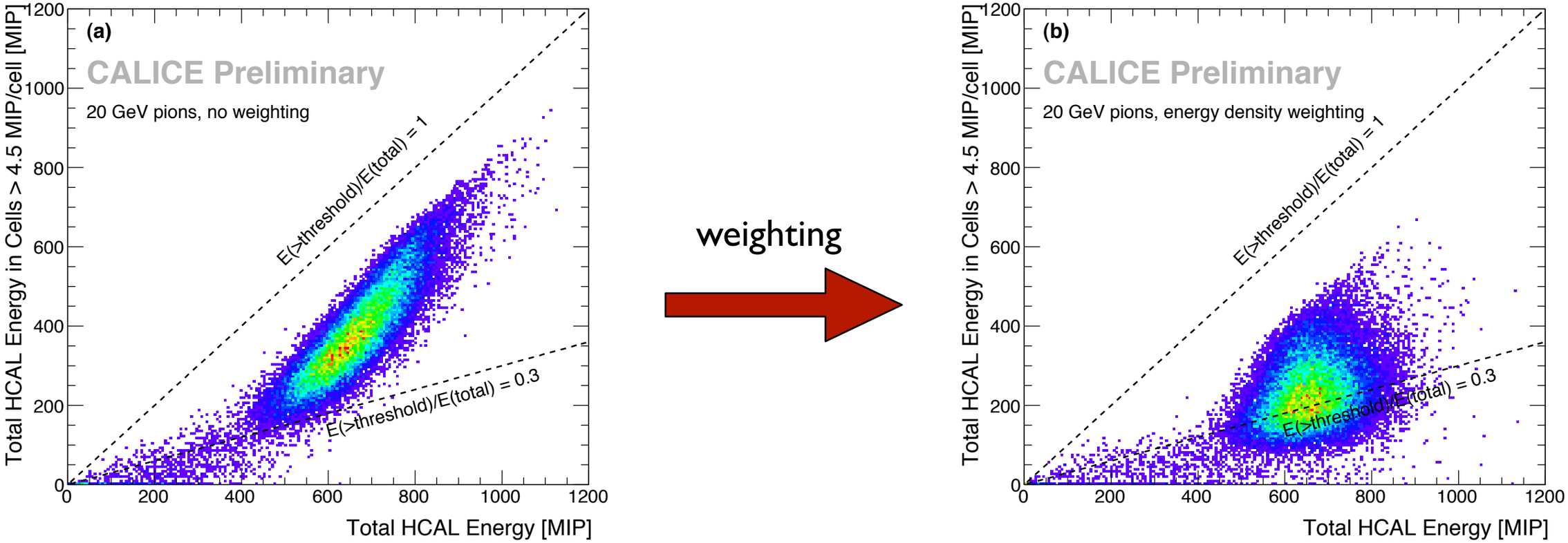
- Weights are energy dependent - Weights are parametrized by an analytic function with energy dependent parameters - allows energy reconstruction without prior knowledge of particle energy



- Cells with low energy content are assigned a higher weight than cells with high content

Energy Reconstruction with Local Compensation

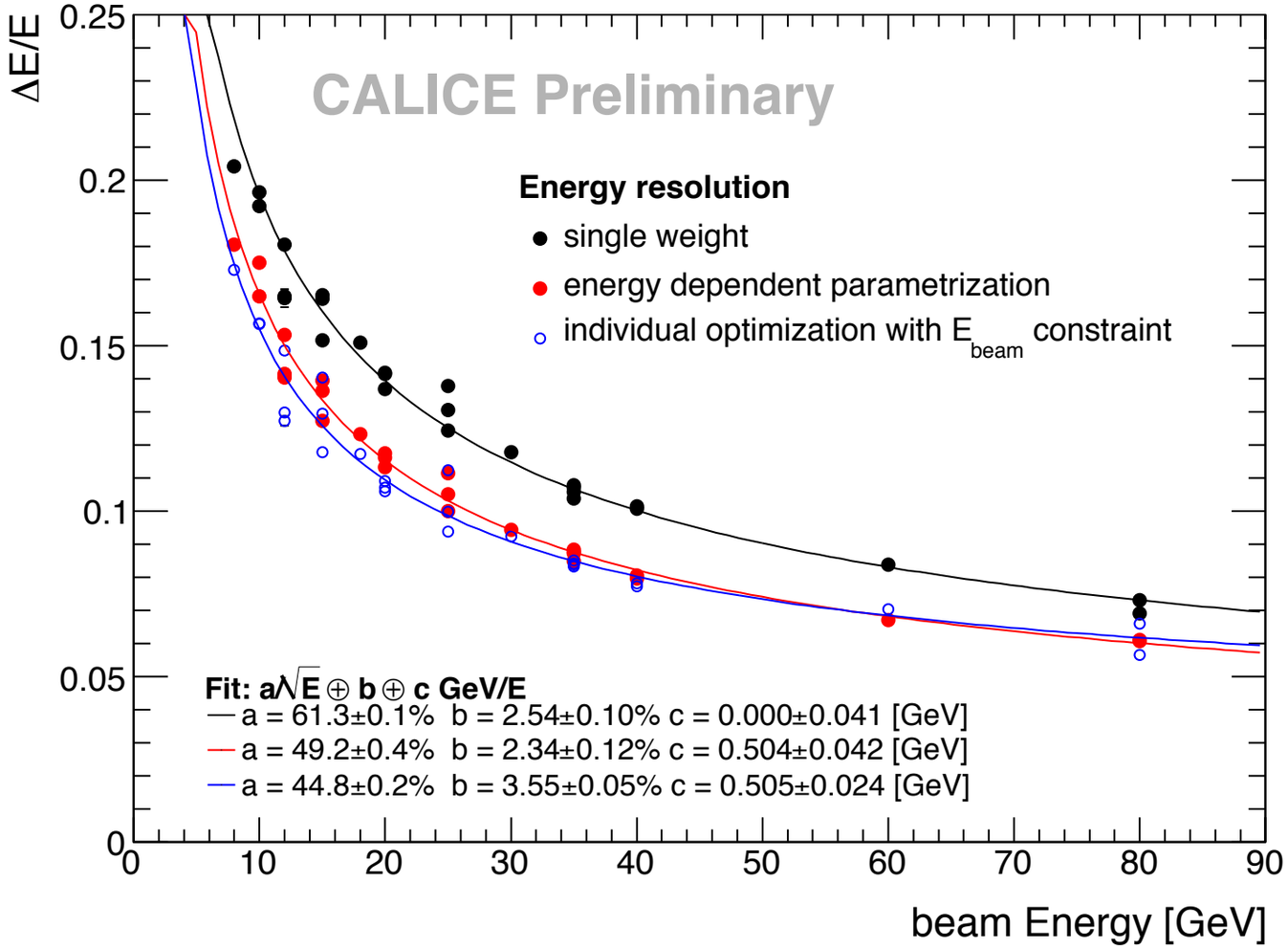
- Effect of local compensation: Reduction of contribution of high density cells, improvement of energy reconstruction



- Generalization of concept to complete CALICE setup: Weights in ECAL, HCAL and TCMT

Local Compensation: Energy Resolution

- Preliminary Results - No temperature corrections applied, accounts for run-to-run spread at fixed energy
- Energy resolution with parametrized weights comparable to "idealized" local compensation using point-by-point minimization



Complete CALICE Setup (ECAL, HCAL, TCMT):

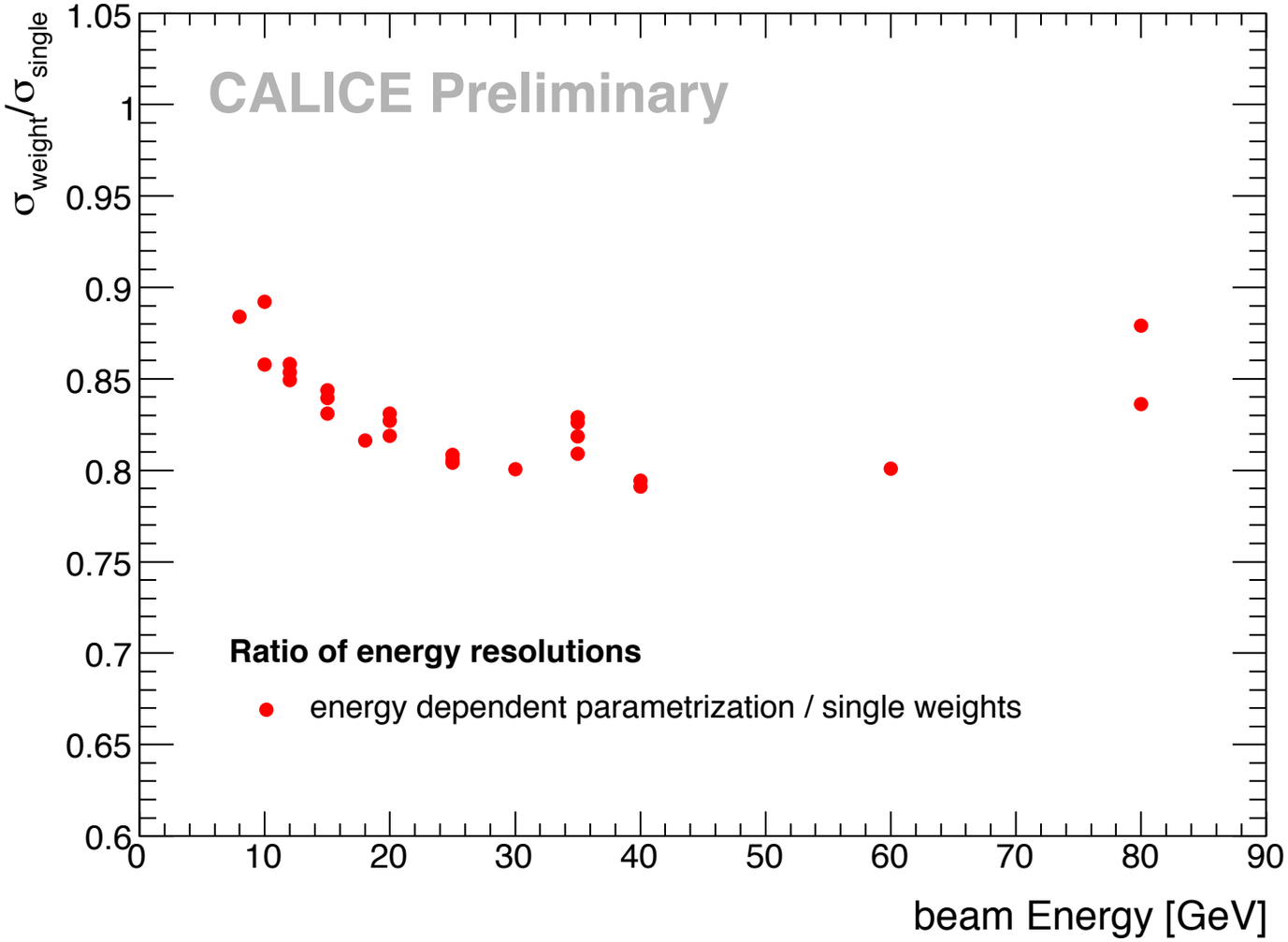
Resolution improved by ~ 18%

Linearity improved: better than 3% over all energy range, dominated by temperature

Resolution given by Gaussian sigma / mean of a fit to the distribution within 1.5σ of peak

Local Compensation: Energy Resolution

- Preliminary Results - No temperature corrections applied, accounts for run-to-run spread at fixed energy
- Energy resolution with parametrized weights comparable to "idealized" local compensation using point-by-point minimization



Complete CALICE Setup (ECAL, HCAL, TCMT):

Resolution improved by ~ 18%

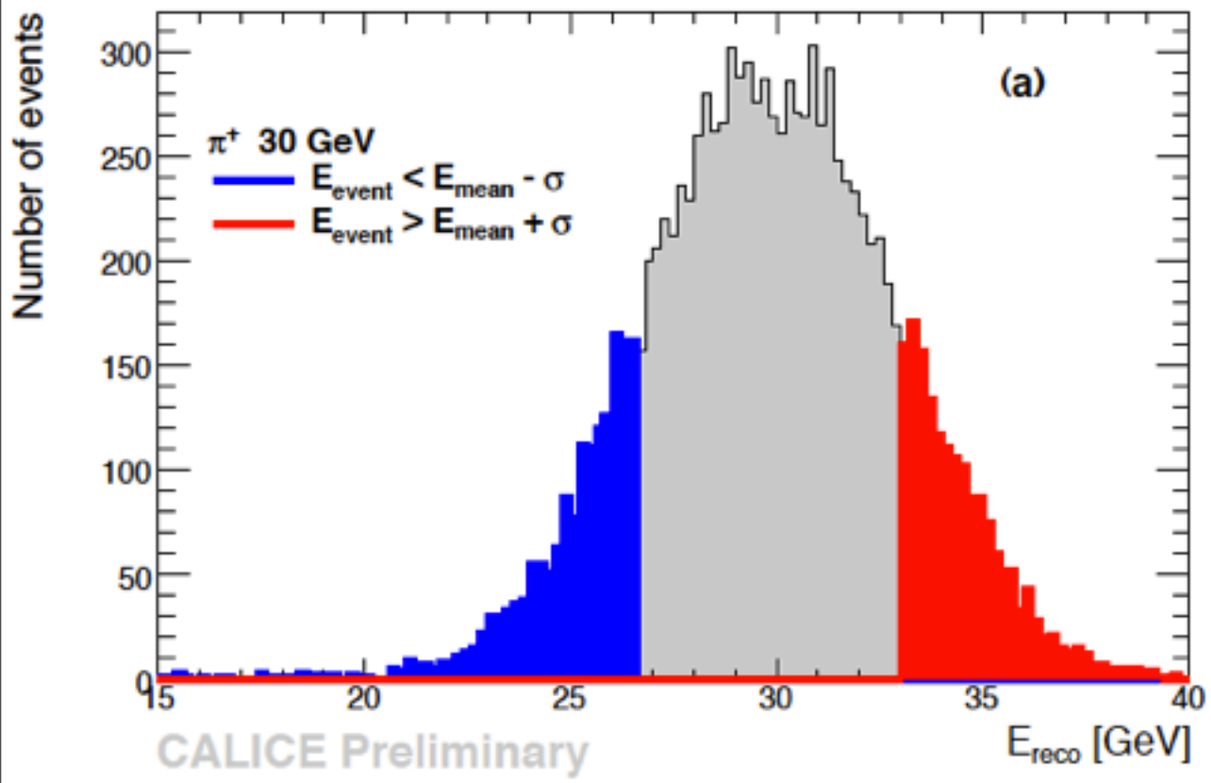
Linearity improved: better than 3% over all energy range, dominated by temperature

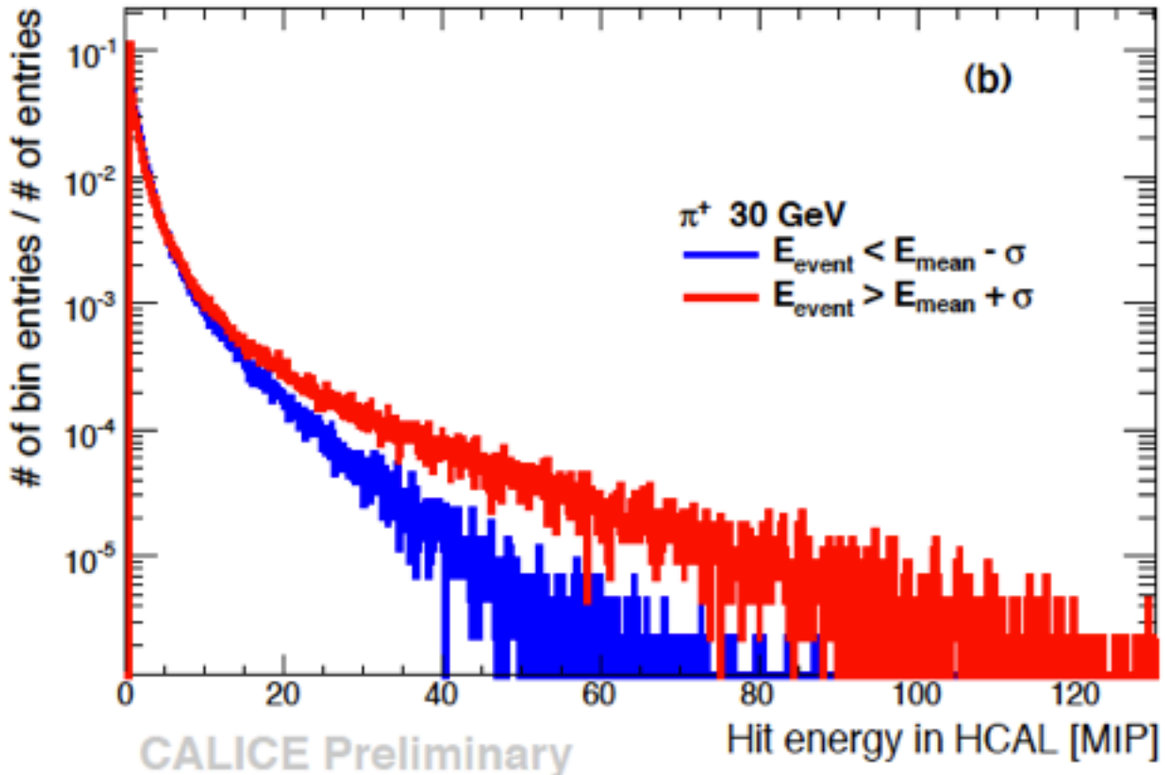
Resolution given by Gaussian sigma / mean of a fit to the distribution within 1.5σ of peak

Global Software Compensation
Data Driven

Global Compensation - The Technique

- Same starting point: Electromagnetic subshowers tend to have higher energy cells than hadronic components
  - For a given particle energy, events that reconstruct too low in energy have few high density cells, events that reconstruct too high have many

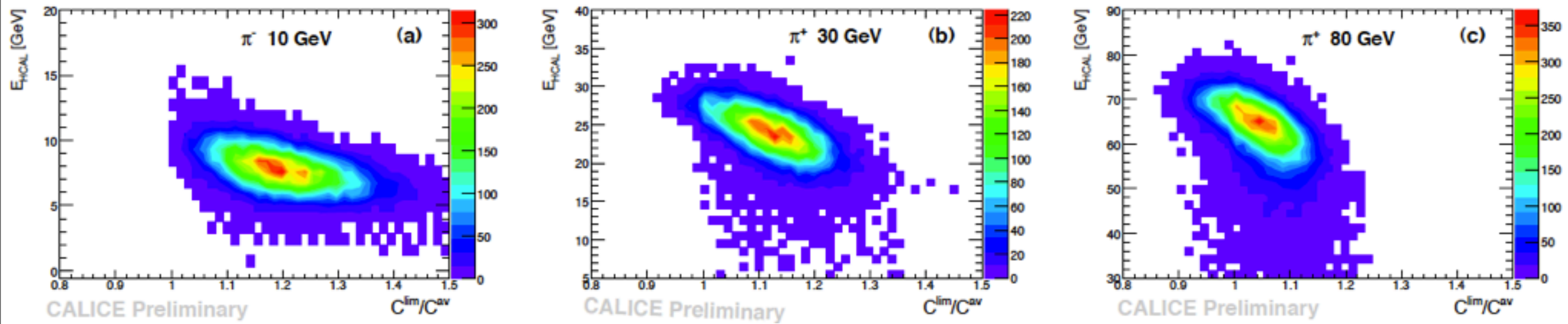




- Introduction of two new variables: C_{lim}, the probability to find a hit with an energy above a certain threshold (5.5 MIP); C_{av}, the probability to find a hit above the average hit energy (provides information on energy dependence)

Energy Reconstruction with Global Compensation

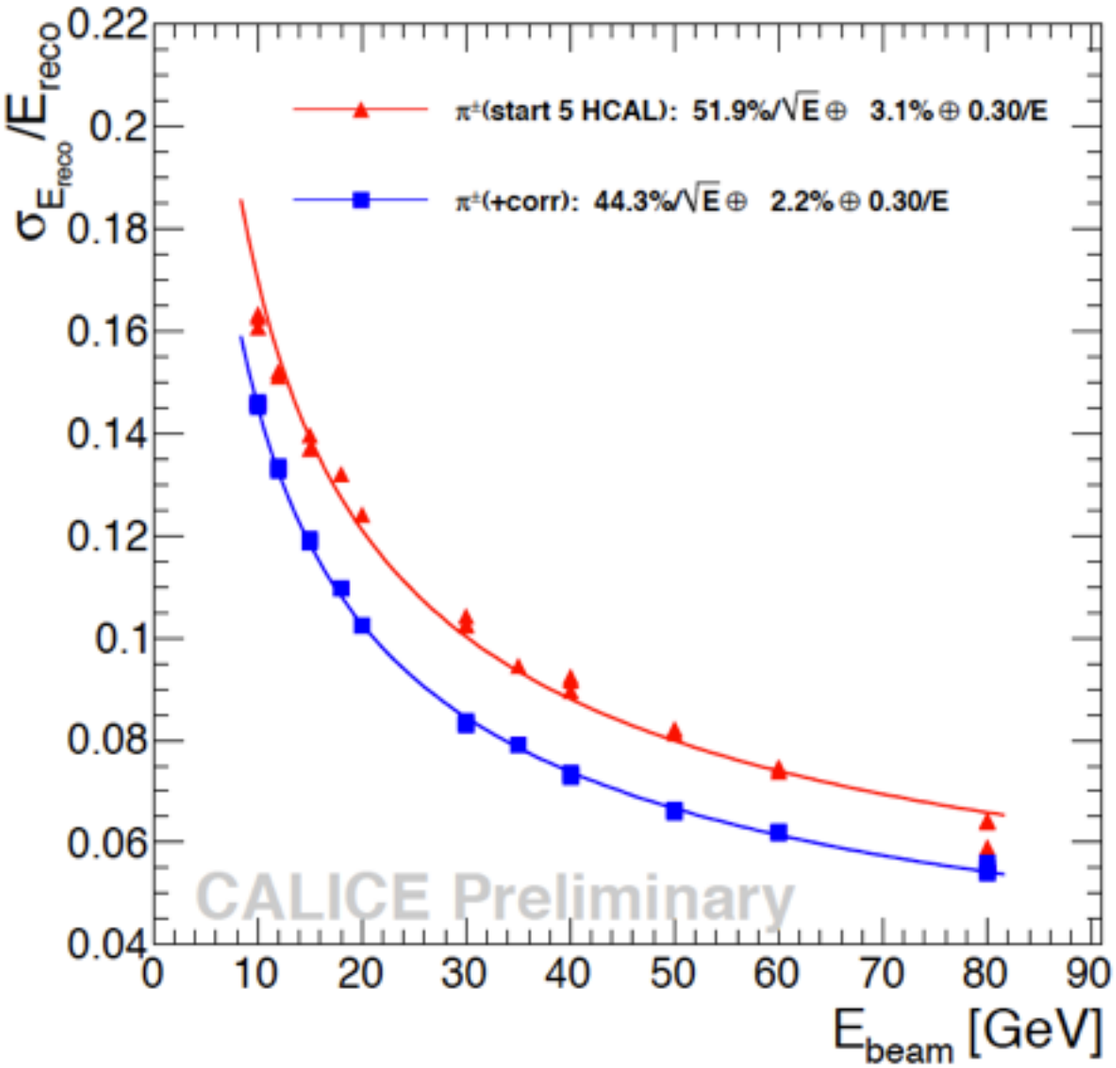
- Consider only events with identified shower start in the HCAL (require MIP-like track in ECAL, no requirement on shower containment)
- Exploit correlation between reconstructed energy and C_{lim}/C_{av}

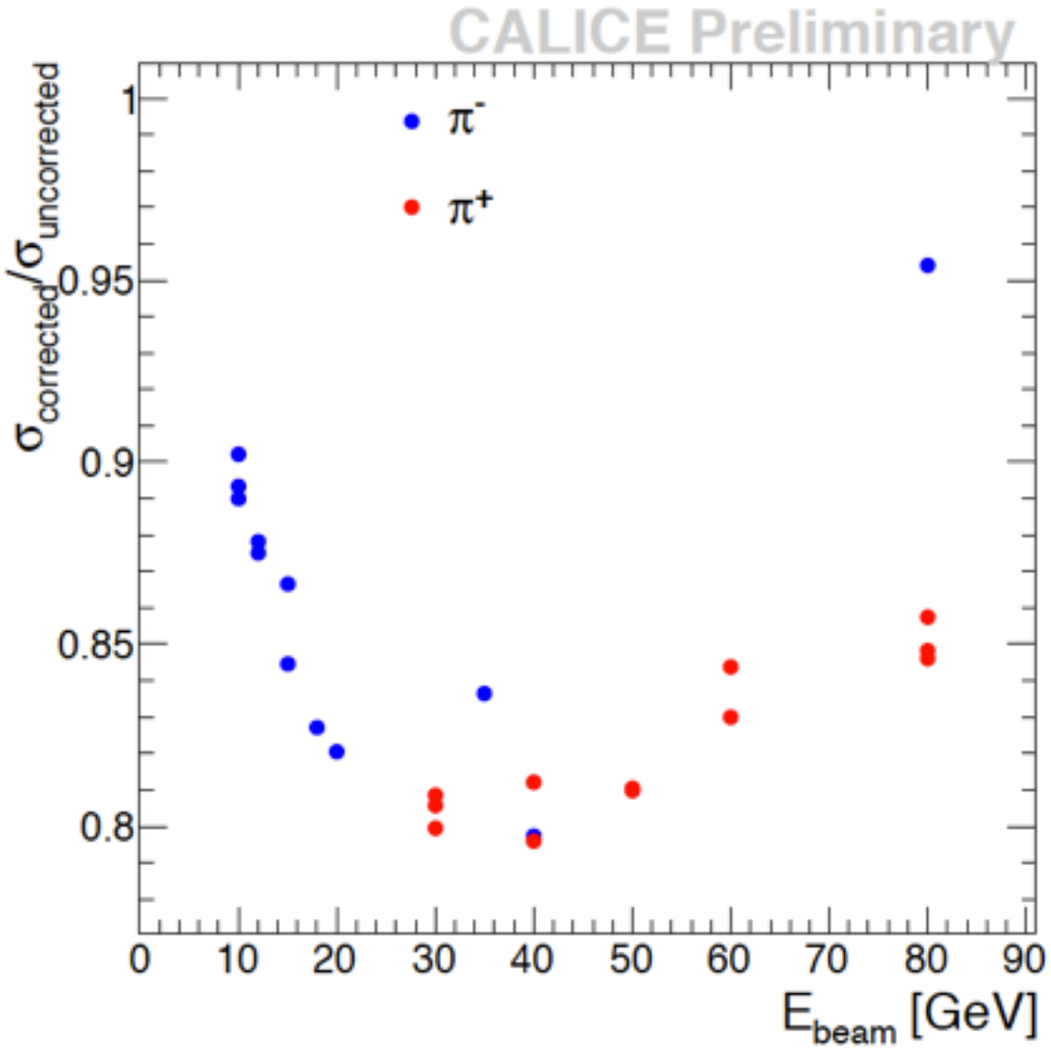


- Reconstructed corrected HCAL energy: $E_i^{HCALcor} = E_i^{HCAL} \times \frac{C_i^{lim}}{C_i^{av}}$
- Include TCMT for full shower energy: $E_i^{SH} = E_i^{HCALcor} + E_i^{TCMT}$
- Energy dependence of correlations breaks linearity, needs an additional correction: $E_i^{SHcor} = E_i^{SH}(a_1 + a_2 E_i^{SH} + a_3 (E_i^{SH})^2)$

Global Compensation: Energy Resolution & Linearity

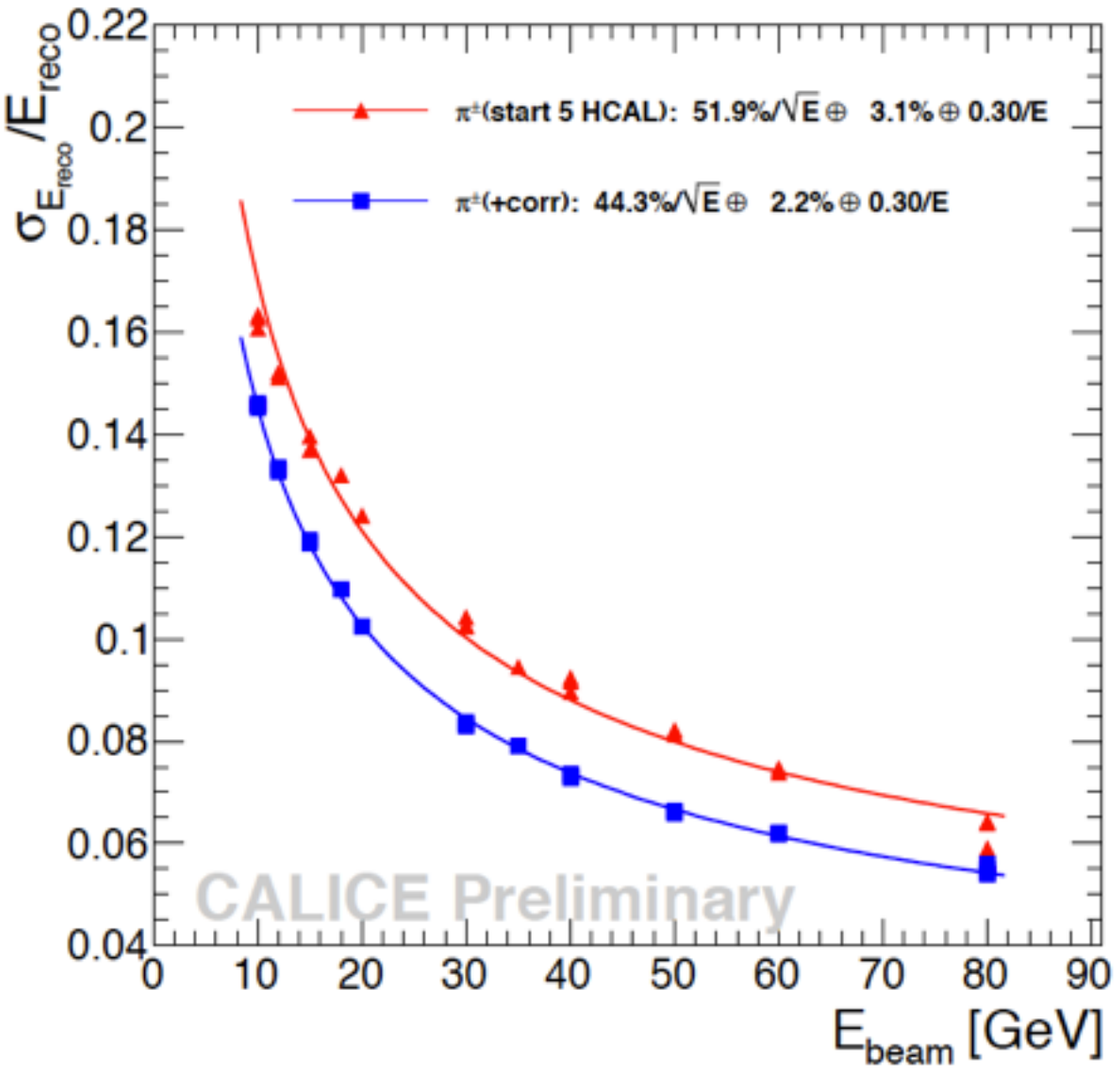
- Significant improvement of energy resolution: ~ 15%,
  up to 20% around 20 - 50 GeV (showers starting in first 5 layers of HCAL)

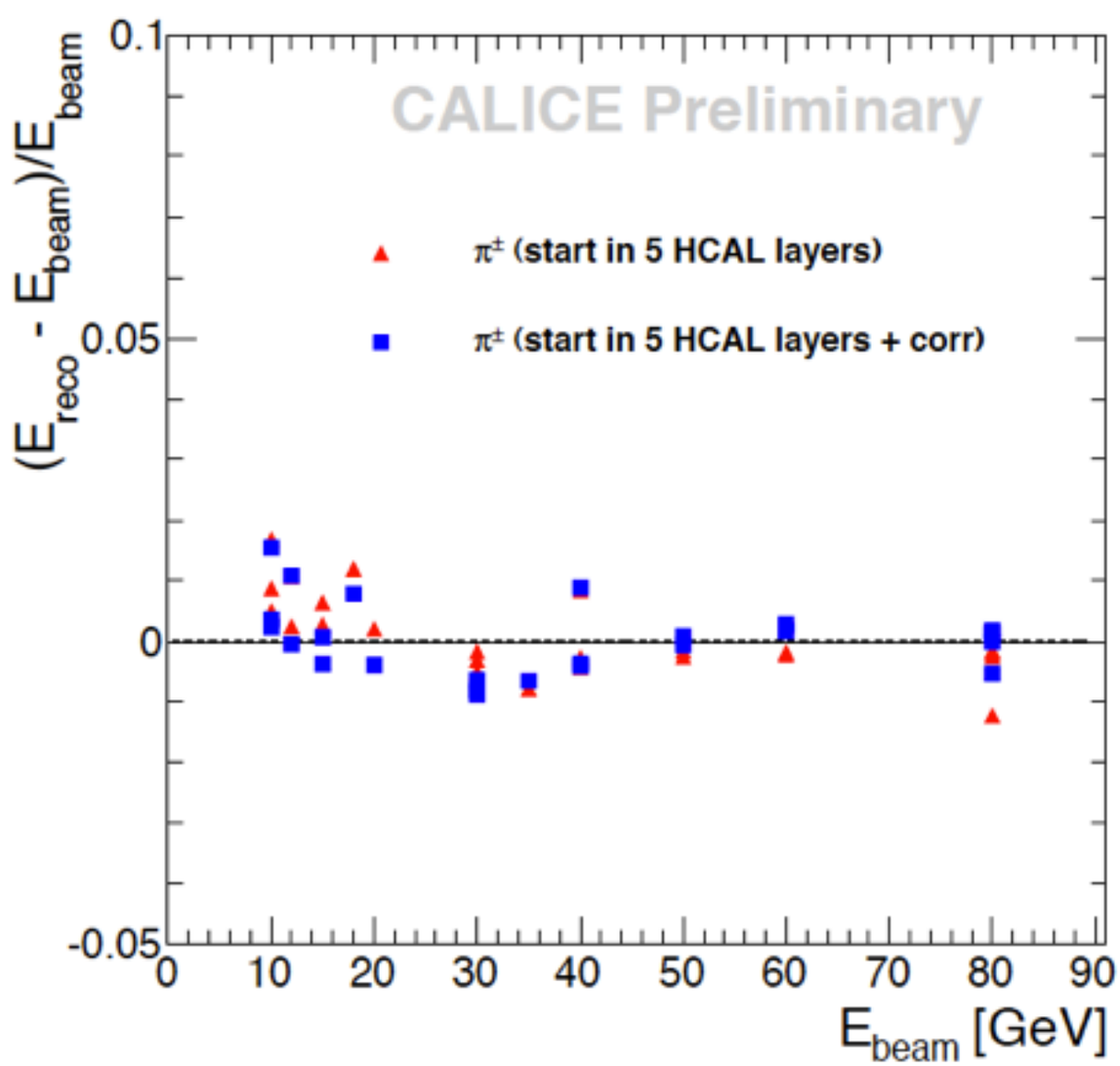




Global Compensation: Energy Resolution & Linearity

- Significant improvement of energy resolution: ~ 15%,
  up to 20% around 20 - 50 GeV (showers starting in first 5 layers of HCAL)



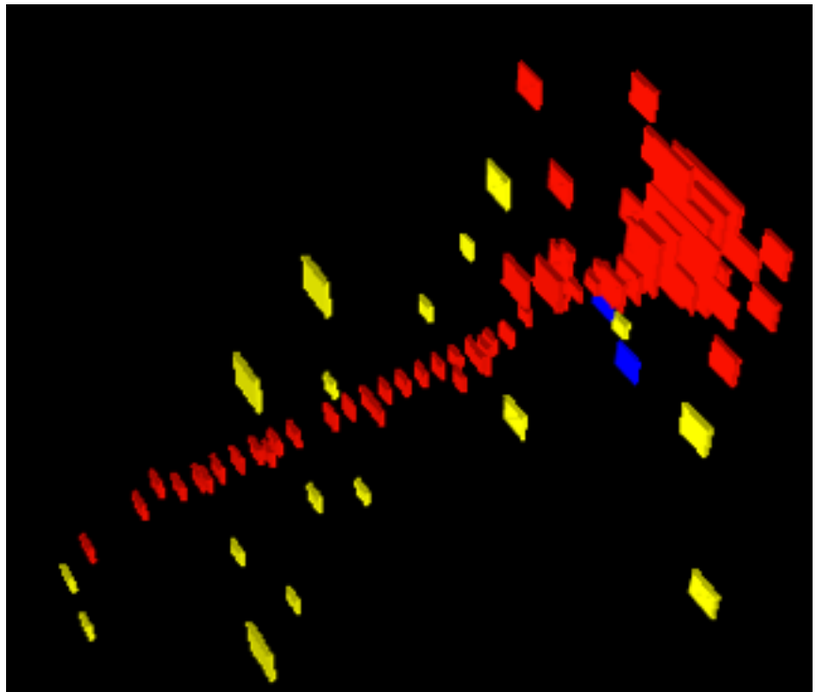


- Excellent linearity: typically within 1%, small remaining run to run fluctuations

Global Software Compensation
MC Trained Neural Network

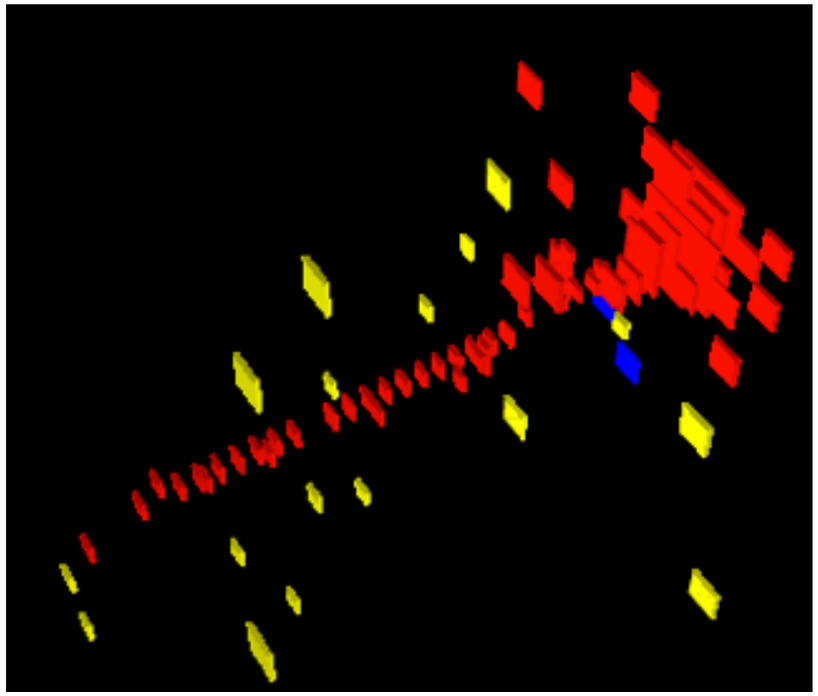
Global Compensation with NN - The Technique

- Simple cluster finding performed to select shower (using events with a track in the ECAL, no other containment requirements)
  - Includes clustering in TCMT
- Determine cluster parameters:
  energy, length, width, volume,
  energy in last 5 HCAL layers, TCMT energy



Global Compensation with NN - The Technique

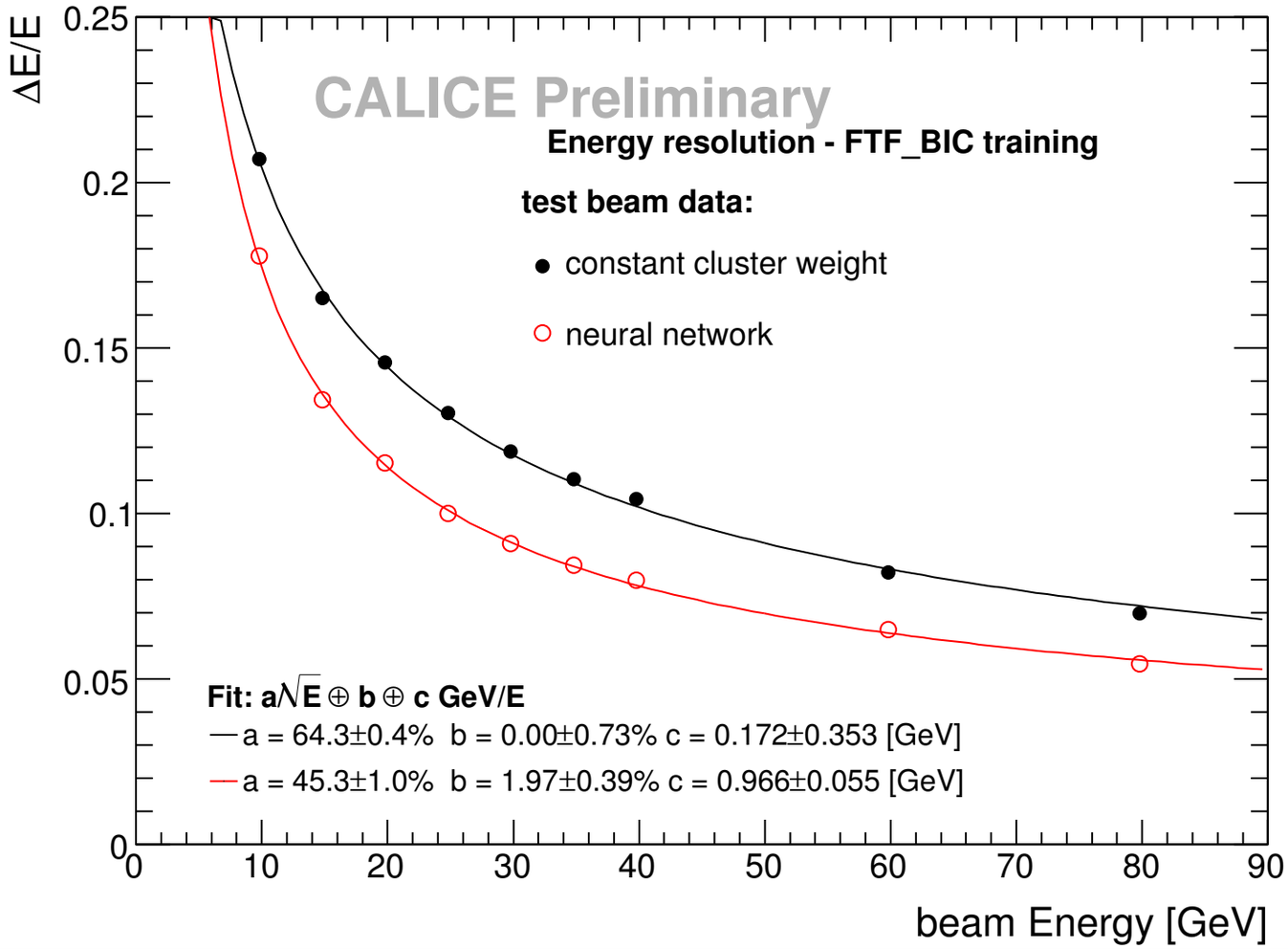
- Simple cluster finding performed to select shower (using events with a track in the ECAL, no other containment requirements)
  - Includes clustering in TCMT
- Determine cluster parameters:
  energy, length, width, volume,
  energy in last 5 HCAL layers, TCMT energy



- Neural Network (TMVA), trained with simulations
  - Used a quasi-continuous energy distribution: 200 k events in steps of 0.1 GeV, in the range from 5 - 105 GeV
    - Crucial to avoid artifacts from discrete beam energies which influences training
  - Weight expression in training to assign higher weight to lower energy events
    - needed to obtain good performance at low energies, otherwise absolute resolution is optimized by NN

Global Compensation with NN: Results

- MC gives a reasonably good description of shower parameters, but:
  An energy dependent difference of global reconstructed energy between data and MC was observed
  - Correction to MC energy scale applied (multiplicative factor given by a smooth function, depending on reconstructed energy)



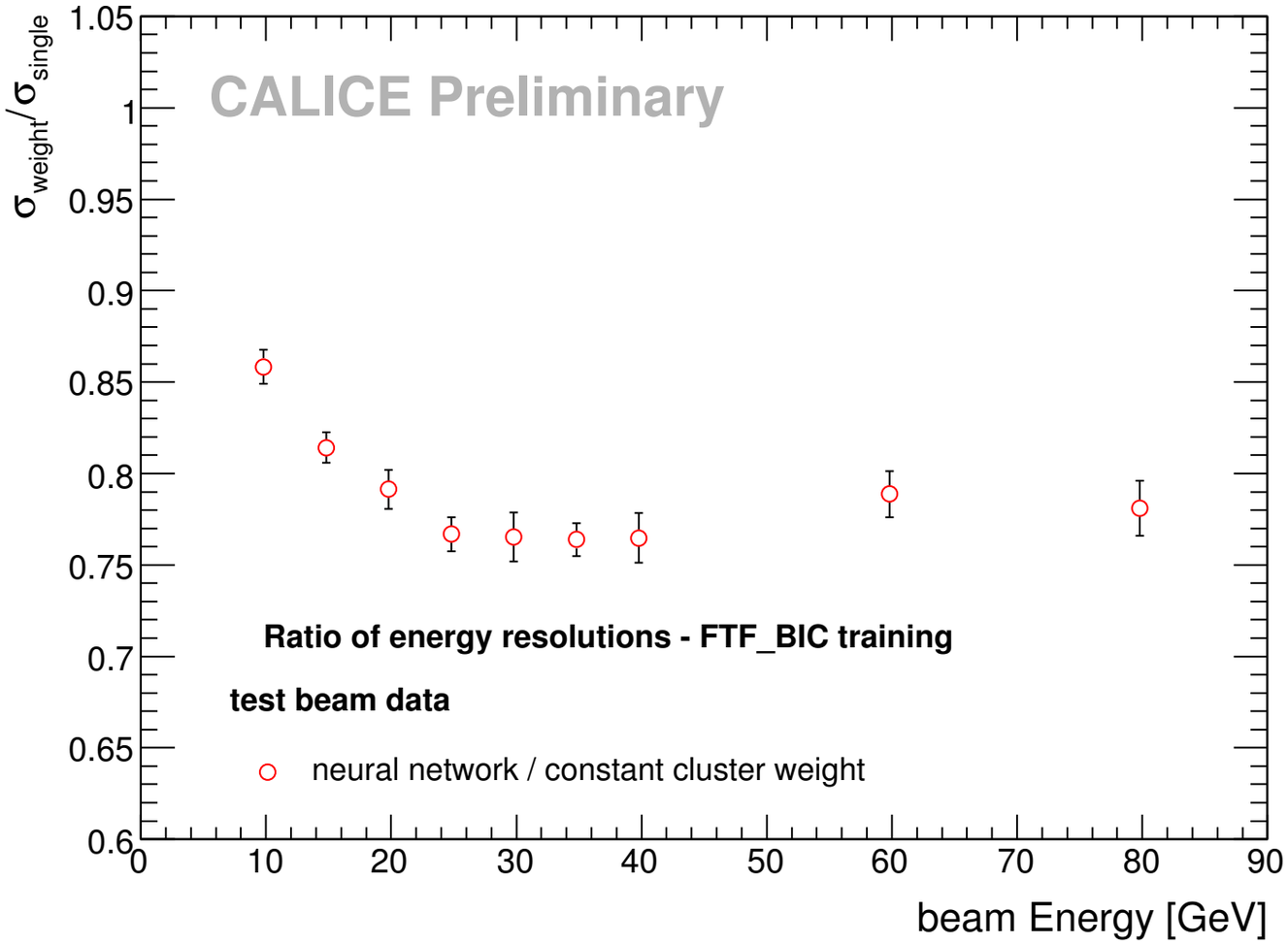
HCAL+TCMT
NN trained with FTF_BIC

Resolution improved by ~25%
(~15% at 10 GeV, ~20% at 15 GeV)

Resolution given by
Gaussian sigma / mean of a fit to the distribution within 1.5σ of peak

Global Compensation with NN: Results

- MC gives a reasonably good description of shower parameters, but:
  An energy dependent difference of global reconstructed energy between data and MC was observed
  - Correction to MC energy scale applied (multiplicative factor given by a smooth function, depending on reconstructed energy)

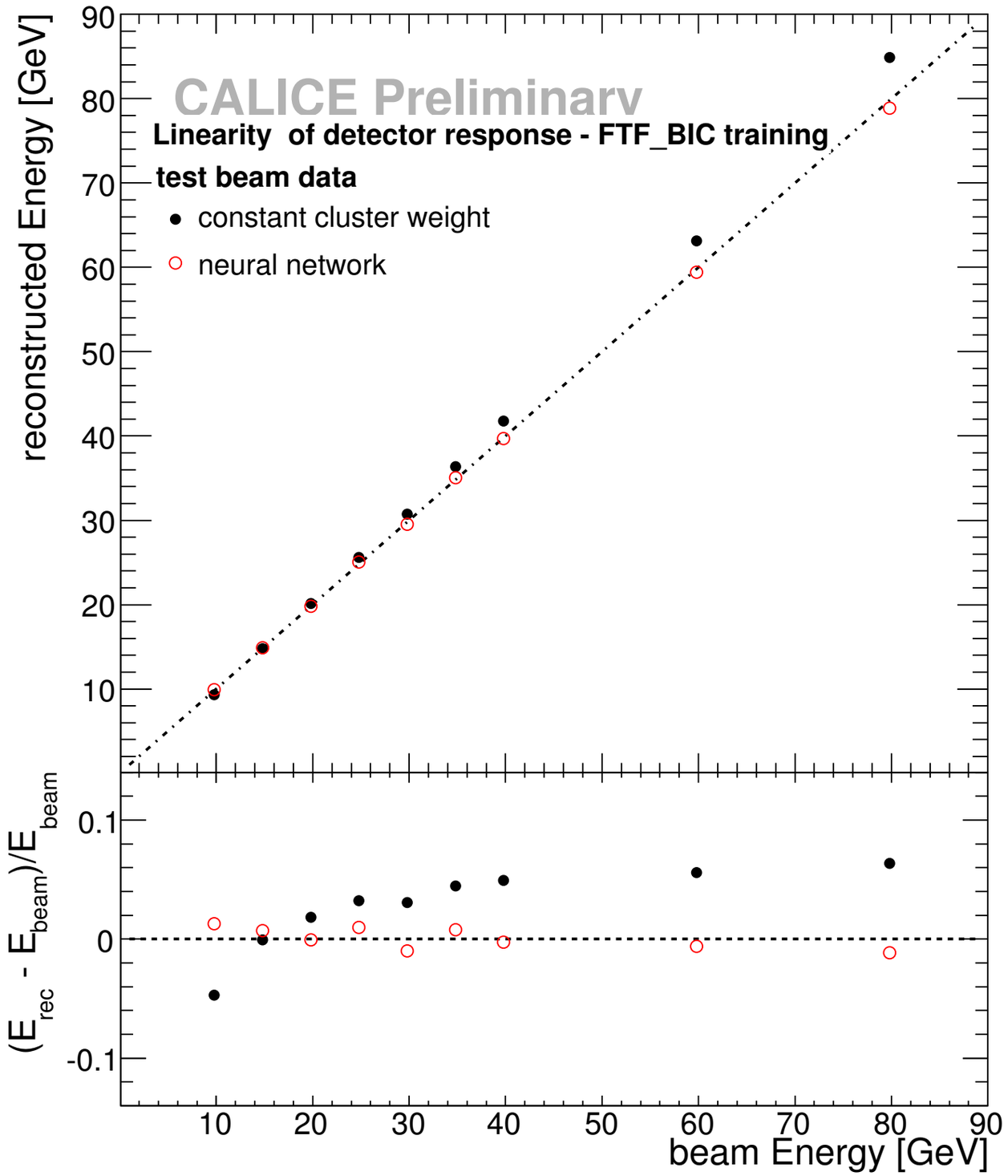


HCAL+TCMT
NN trained with FTF_BIC

Resolution improved by ~25%
(~15% at 10 GeV, ~20% at 15 GeV)

Resolution given by
Gaussian sigma / mean of a fit to the distribution within 1.5σ of peak

Software Compensation: Linearity of Response



- Unweighted reconstruction shows typical non-linear behavior: Increased response at high energies
- Software compensation recovers linearity within < 2% from 10 to 80 GeV

Summary / Outlook

- The CALICE AHCAL offers excellent possibilities to test software compensation / weighting methods on different scales of the hadronic cascade

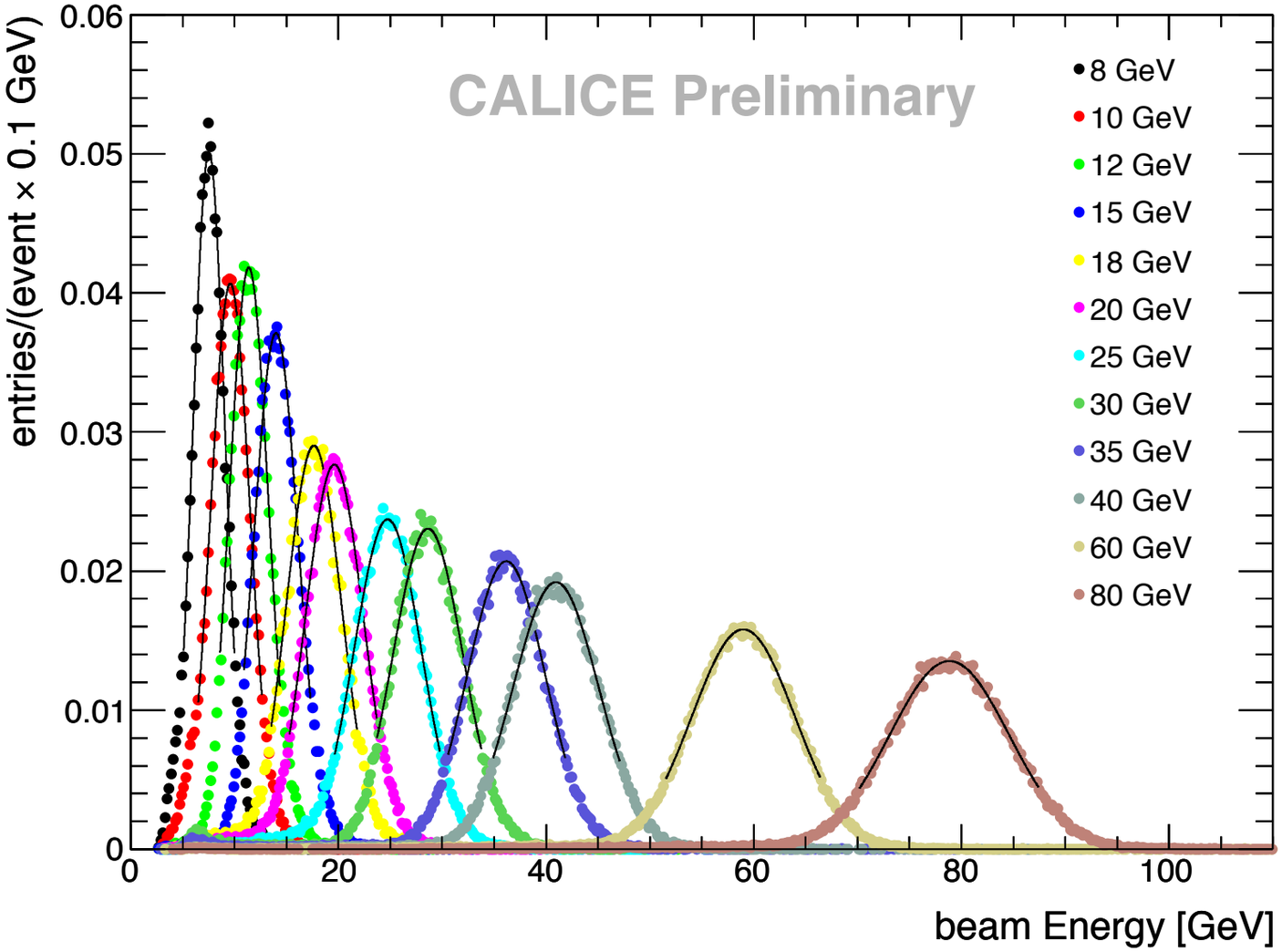
- Data-driven methods and one method based on simulations have been successfully applied - Improvements in energy resolution of 15 - 25% are achieved, resulting in stochastic terms of < ~ 50%/Sqrt(E)
  - Discrepancies between methods partially due to different level of event selection / data quality
  - Success of Neural Network weighting demonstrates already good global agreement of data and simulations

- Further steps & ideas
  - Extend MC-driven methods to more local scales
  - Explore potential of compensation methods in simulation studies (em vs had components)
  - Explicit analysis of shower substructure - Identification of em / had subshowers

Backup

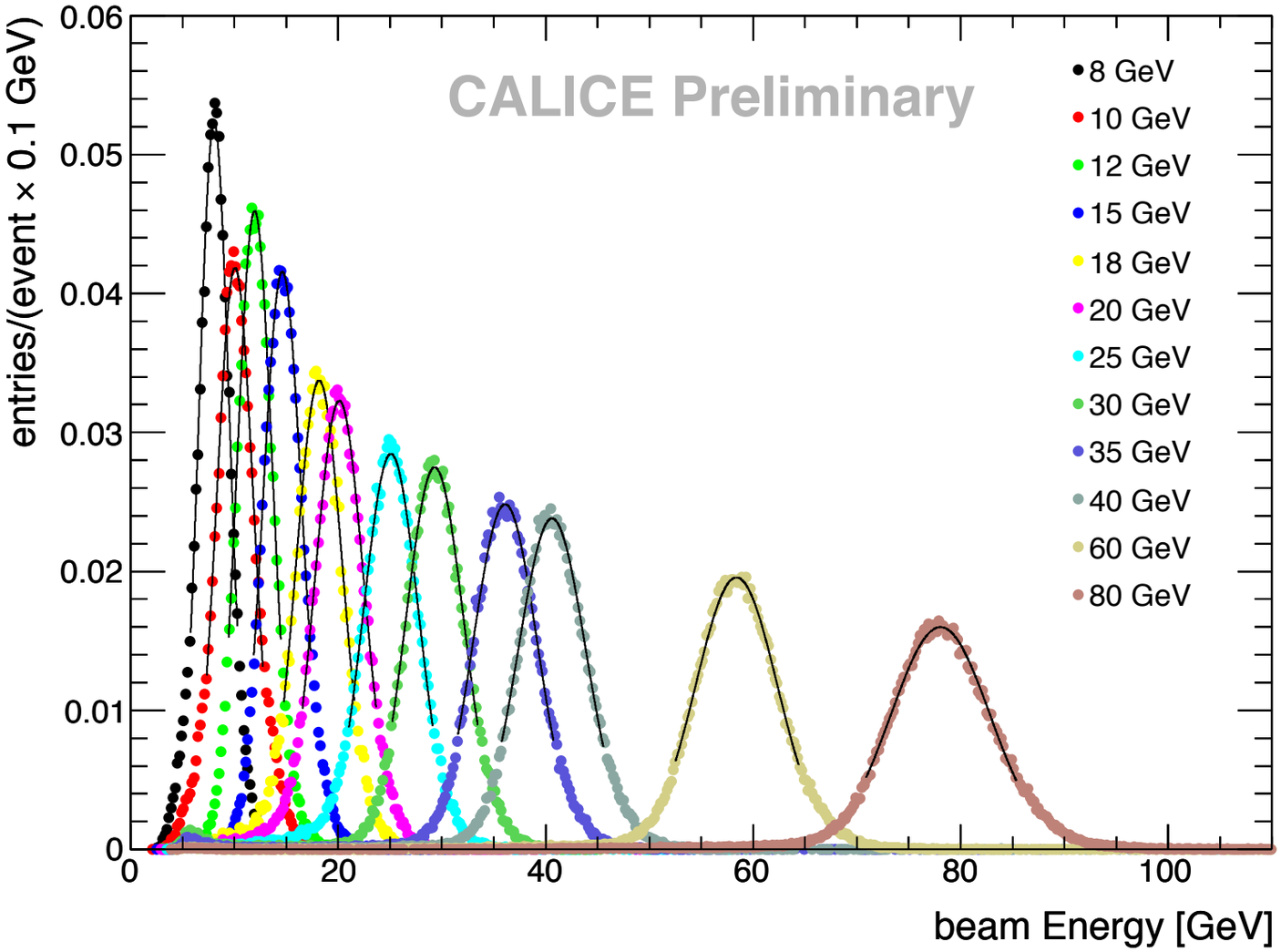
Energy Distributions for Local Compensation

- Two reconstruction methods:
  - Simple reconstruction w/o weighting: one calibration factor (MIP to GeV) per subdetector (ECAL, HCAL, TCMT)



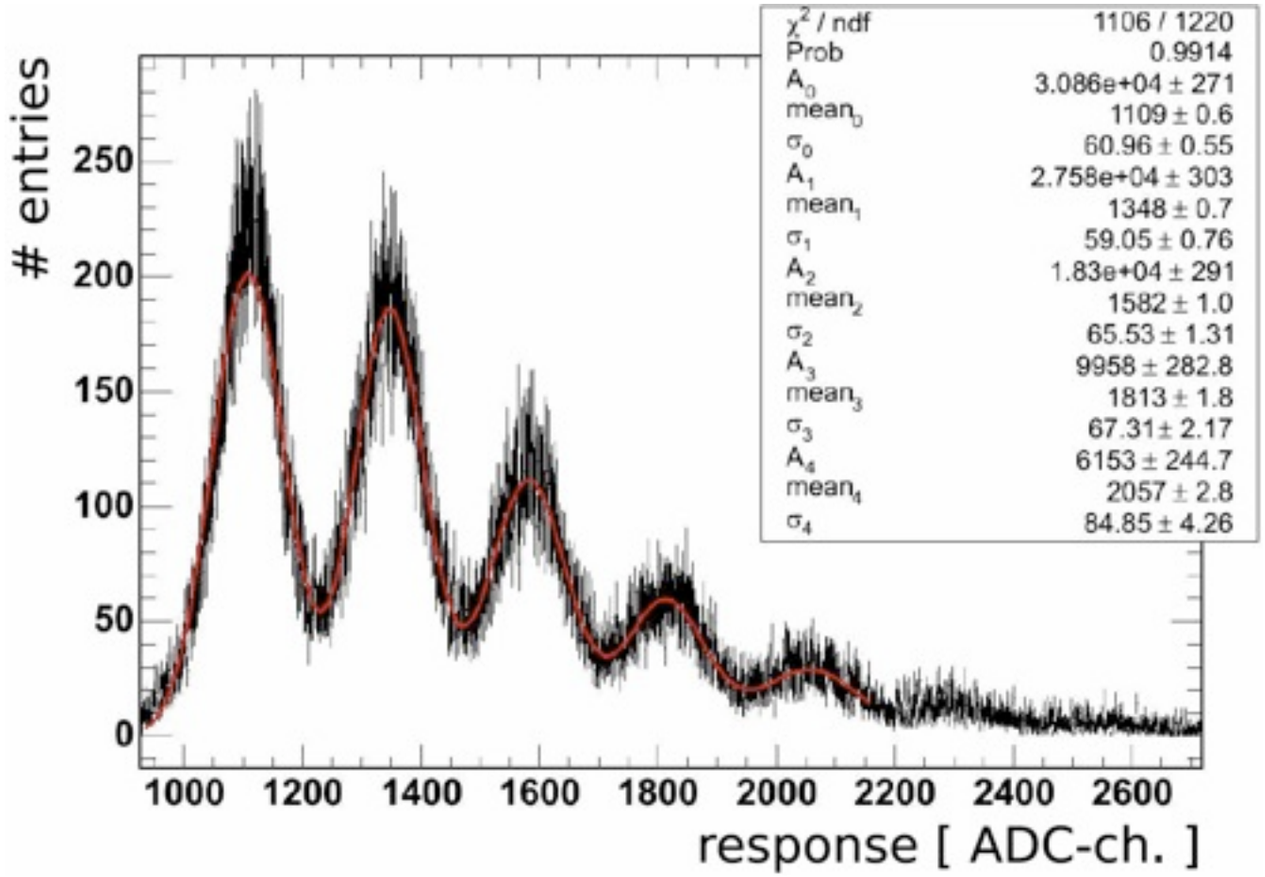
Energy Distributions for Local Compensation

- Two reconstruction methods:
  - Simple reconstruction w/o weighting: one calibration factor (MIP to GeV) per subdetector (ECAL, HCAL, TCMT)
  - Energy density dependent weighting, parametrized energy dependence



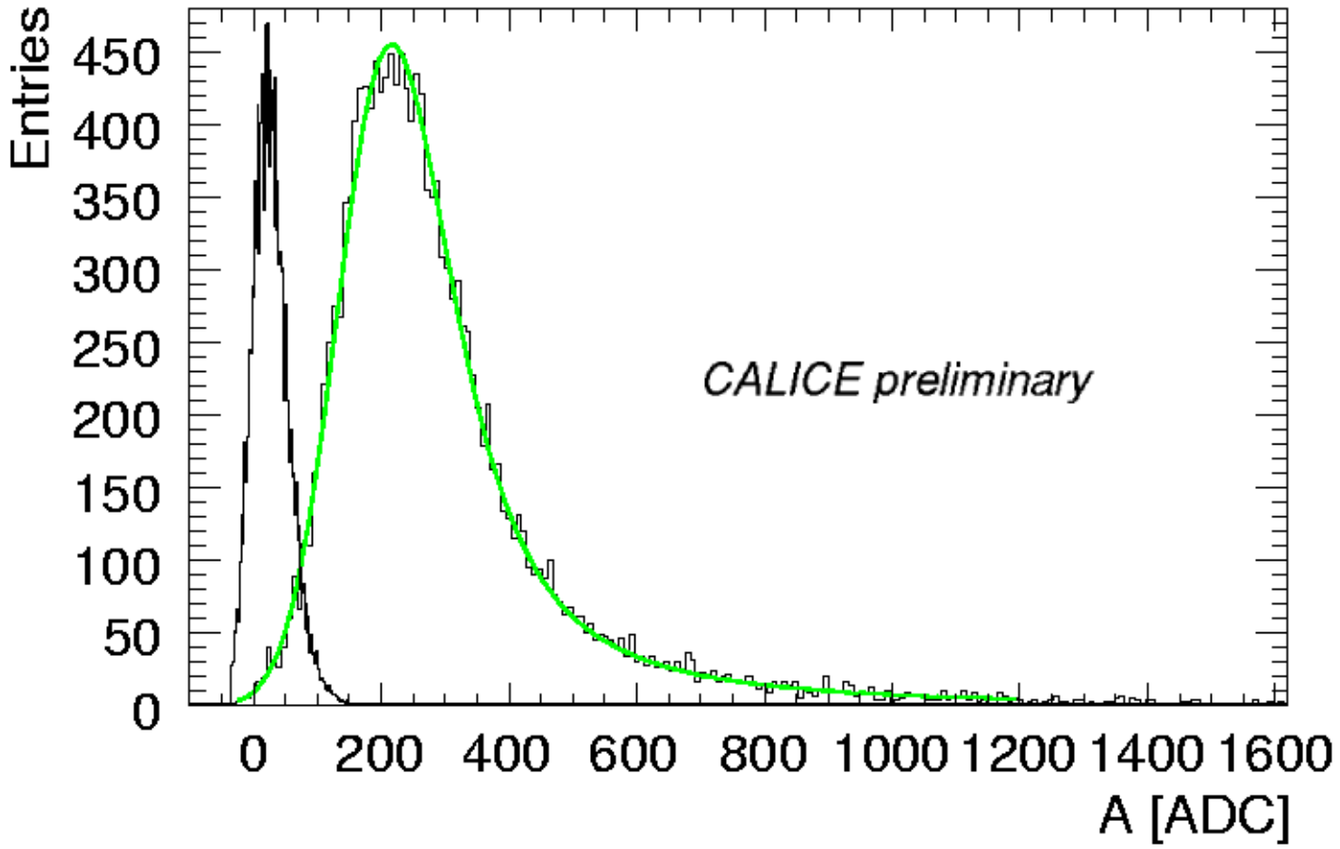
Analog HCAL: Calibration

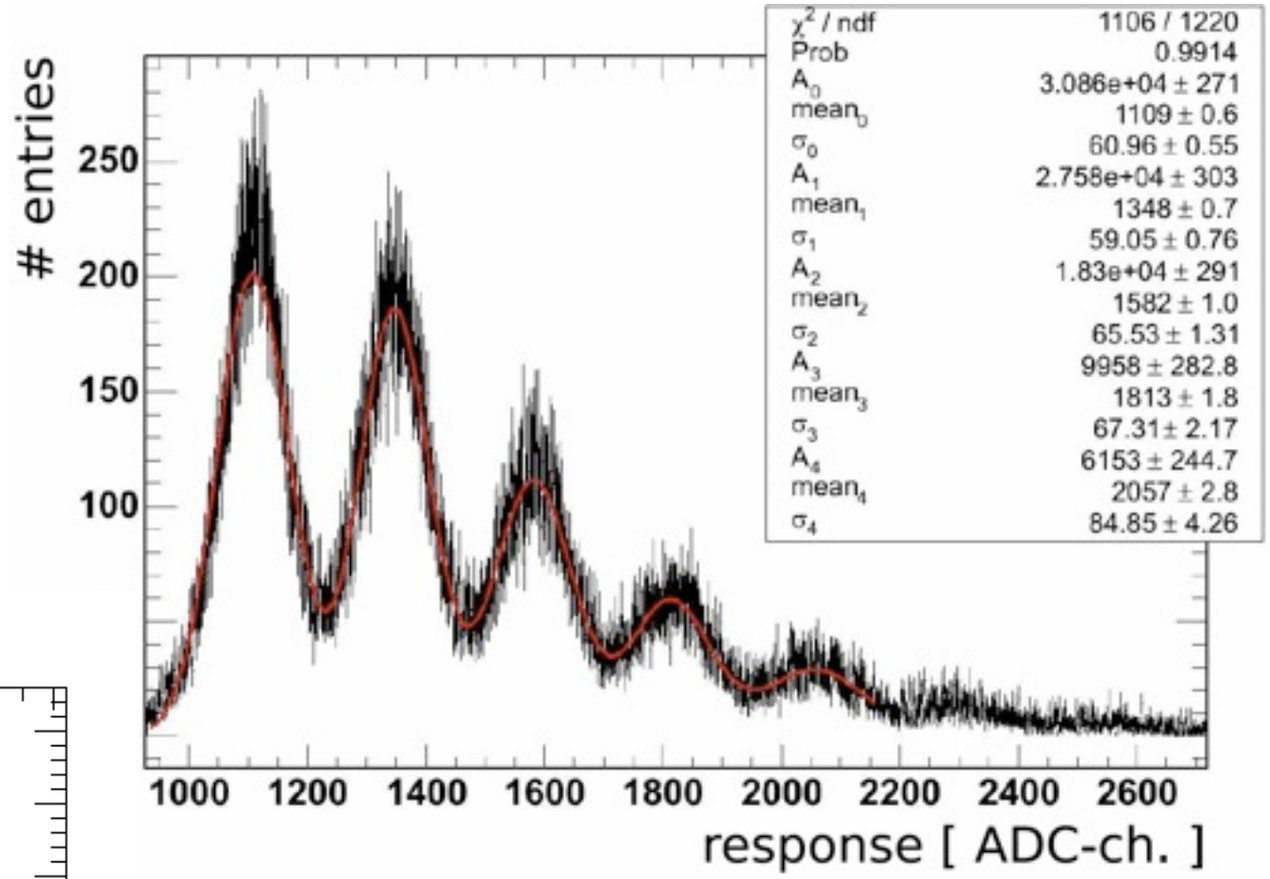
- Auto-calibration of SiPM gain:
  Individual photons can be resolved
  - Low-intensity LED light coupled into each detector cell
  - high gain setting of front-end electronics



Analog HCAL: Calibration

- Auto-calibration of SiPM gain:
  Individual photons can be resolved
  - Low-intensity LED light coupled into each detector cell
  - high gain setting of front-end electronics

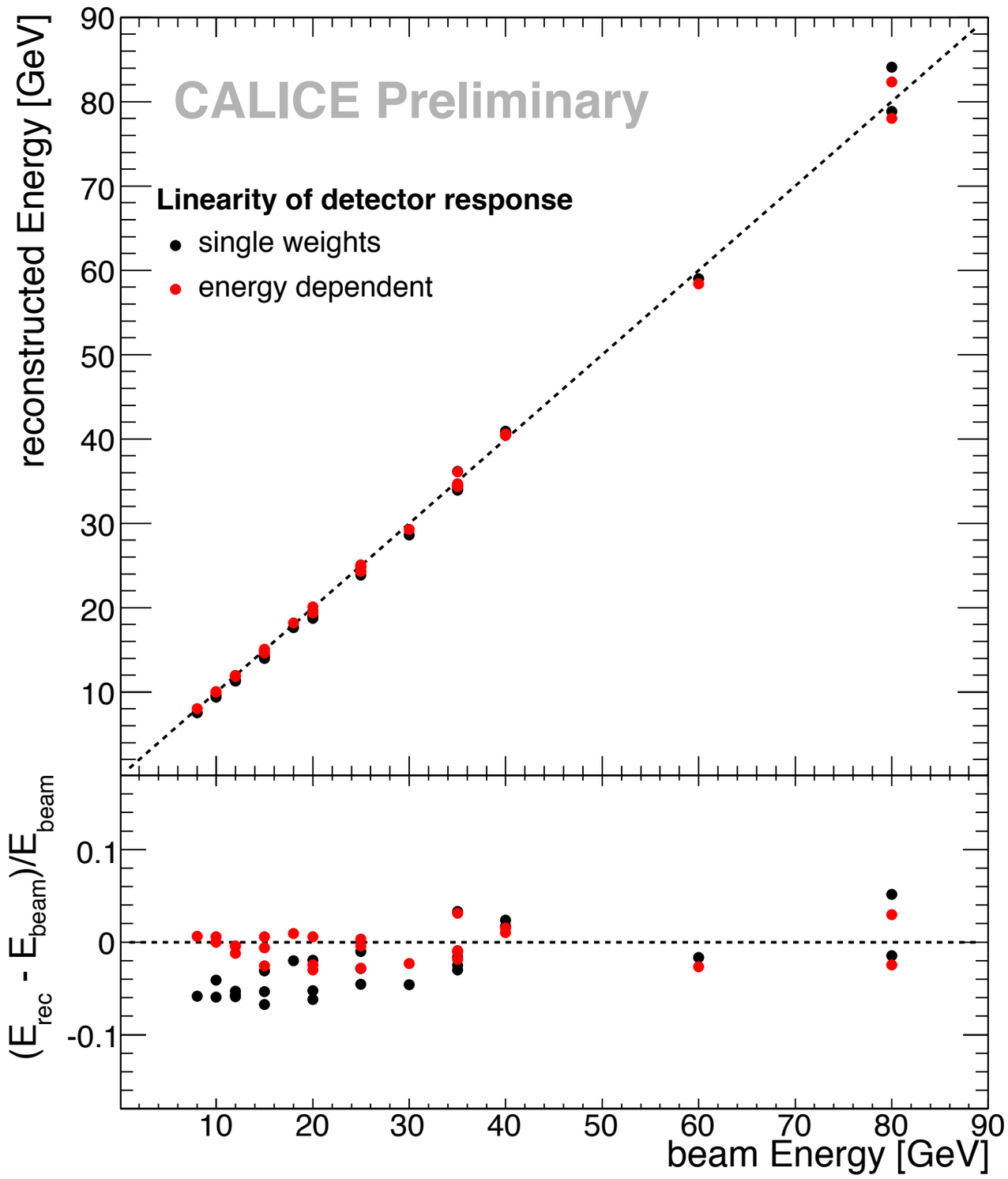




- MIP-Calibration with Muons
  - Complete detector illuminated with high energy muons
  - equalization of response of all cells by matching the MPV position

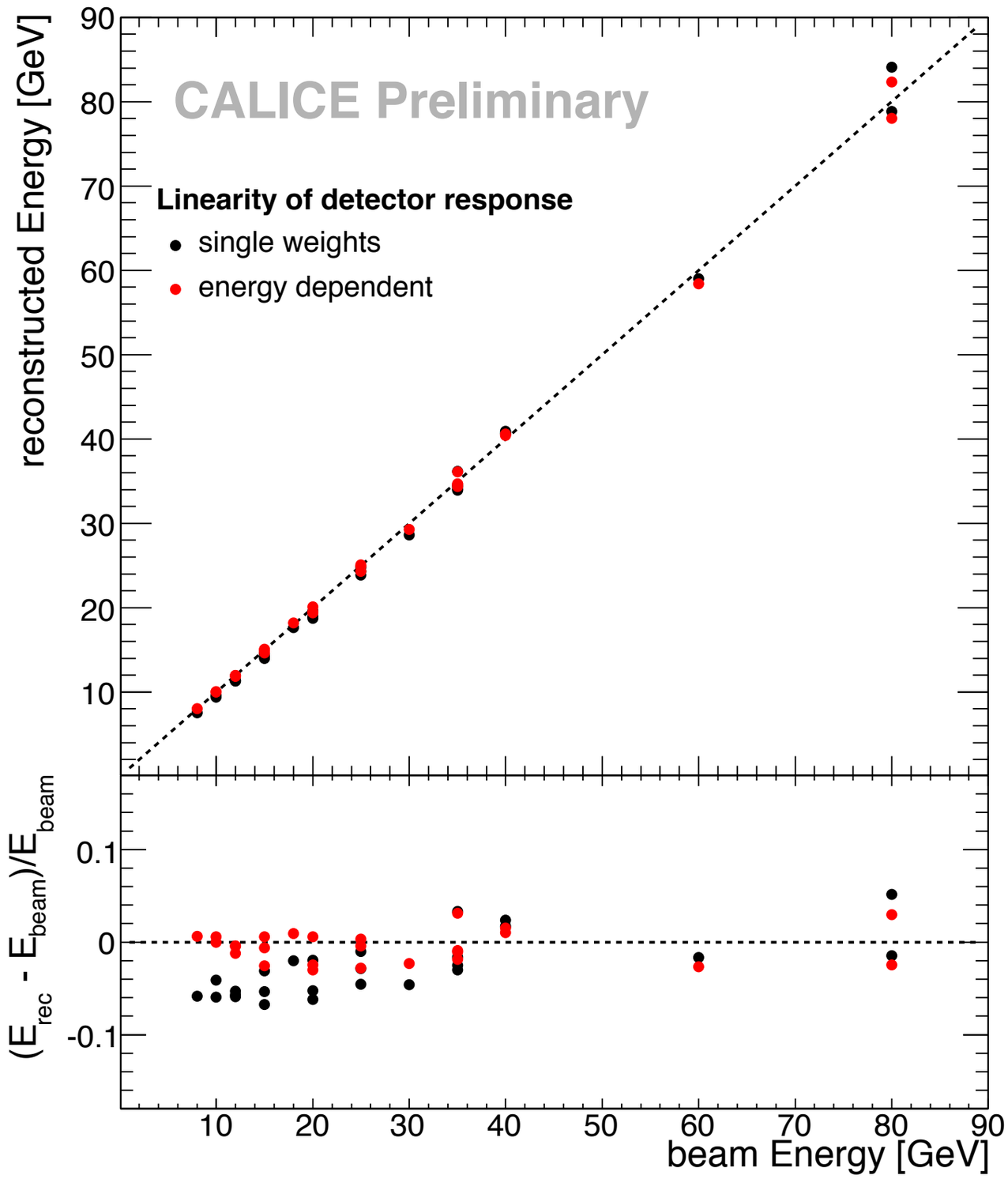
Linearity for Local Compensation

- Energy reconstructed with single conversion factors and with parametrized density dependent weighting
- Noise rejection: Isolated noise hits (and isolated neutrons) rejected in the analysis

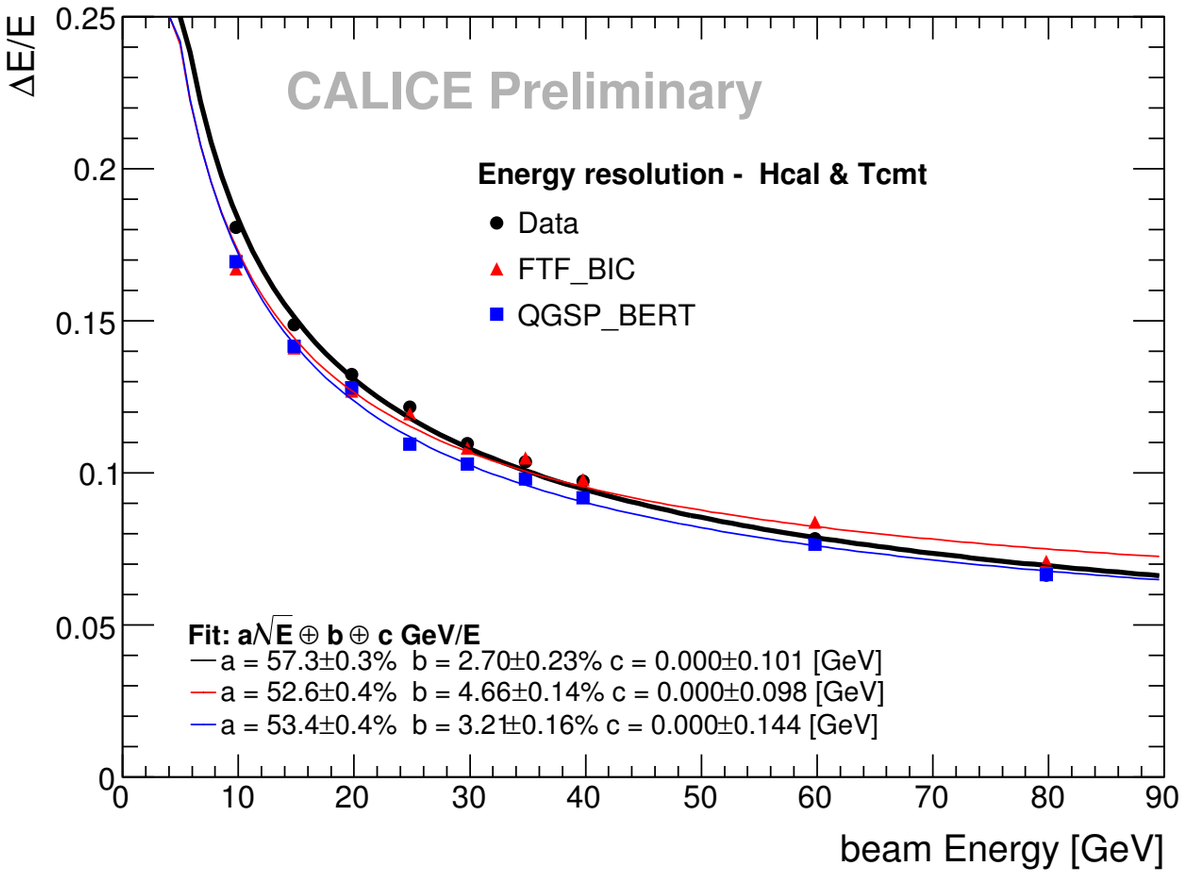


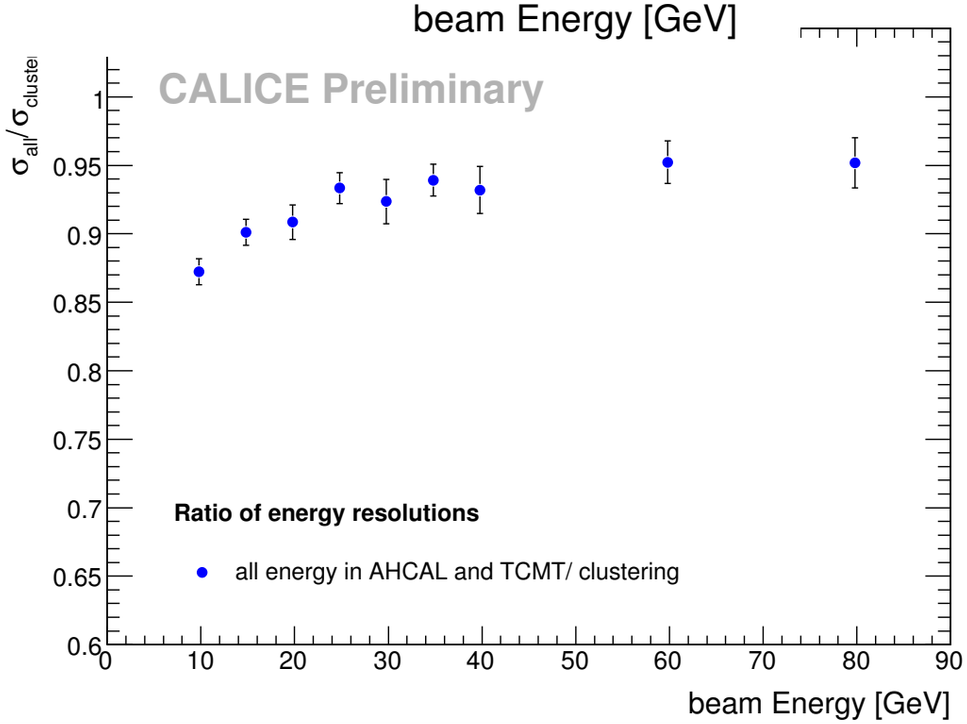
Linearity for Local Compensation

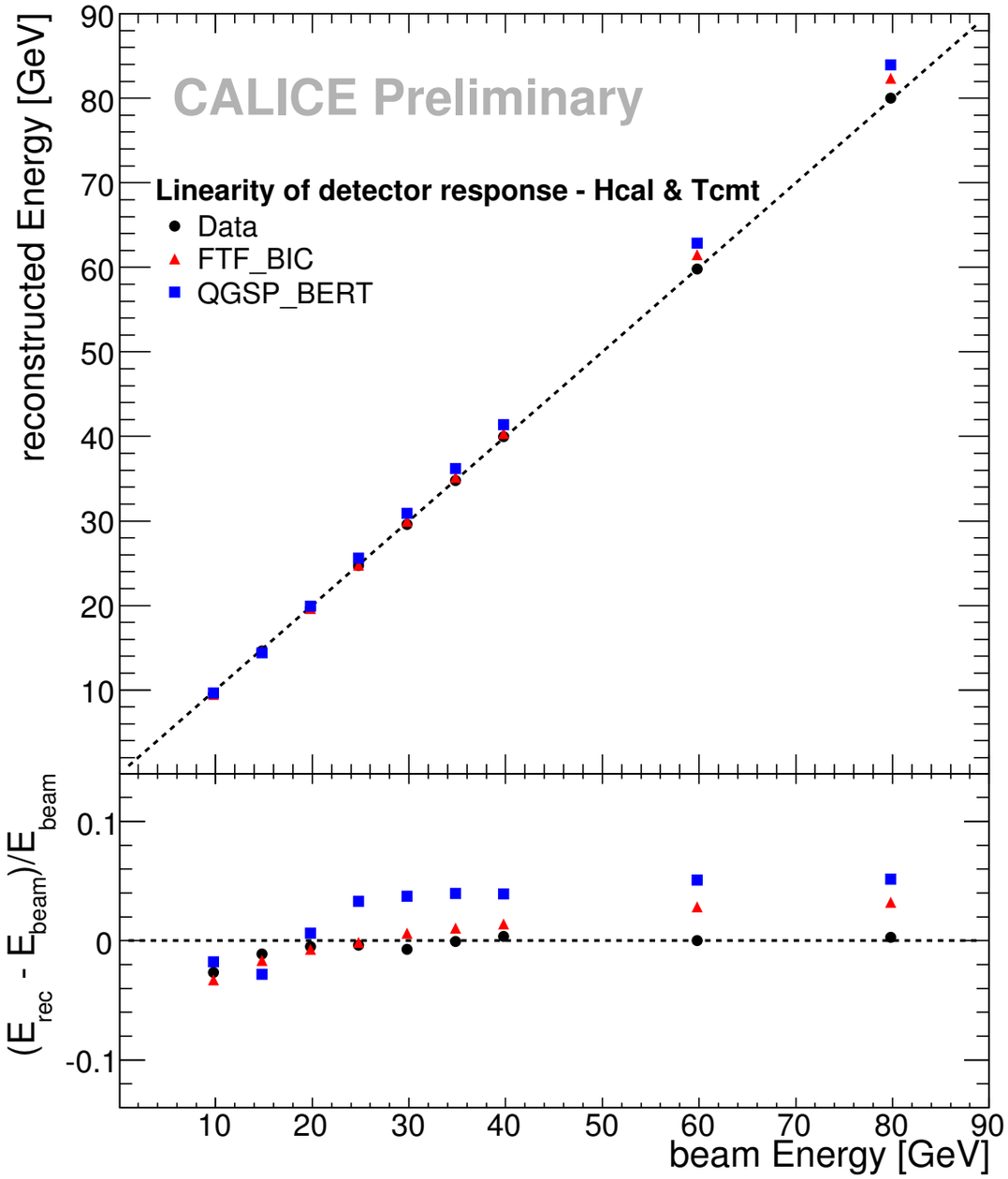
- Energy reconstructed with single conversion factors and with parametrized density dependent weighting
- Noise rejection: Isolated noise hits (and isolated neutrons) rejected in the analysis
- Weighting of cells according to their energy content improves linearity of the detector: better than 4% from 8 to 80 GeV
- Cell-by-cell temperature correction not yet included: leads to a run-to-run spread at a given energy



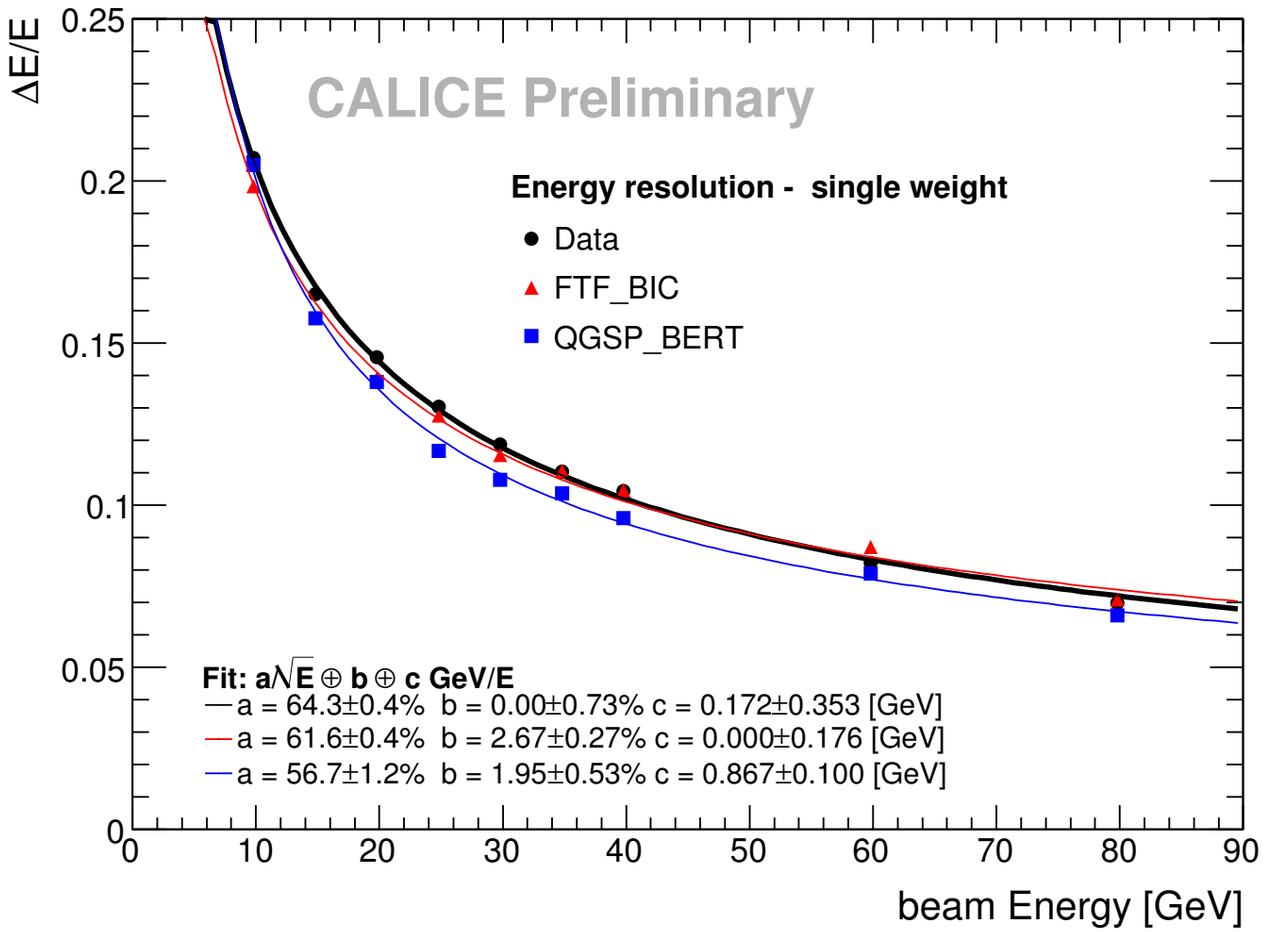
Resolution without Clustering: Data vs MC

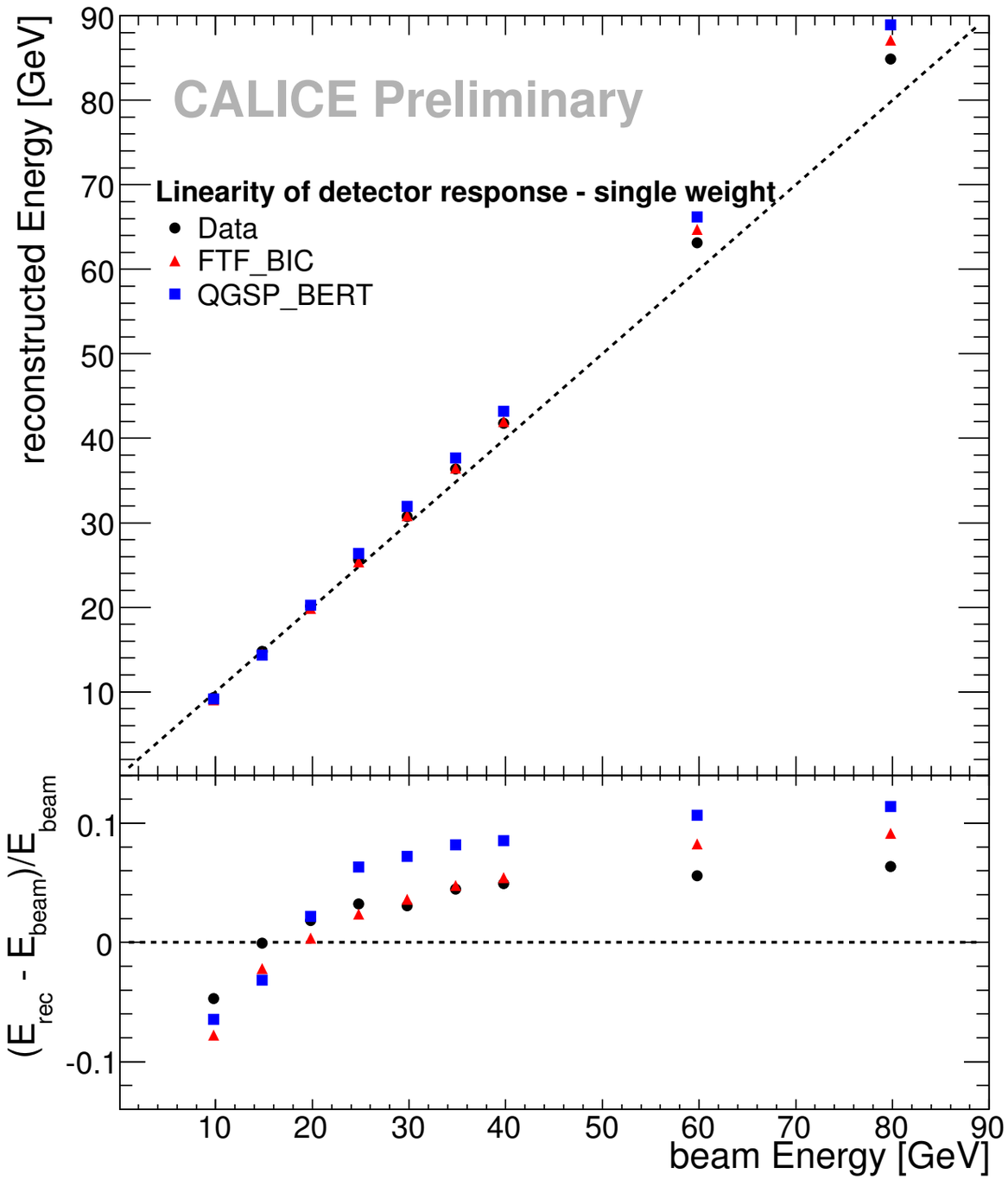




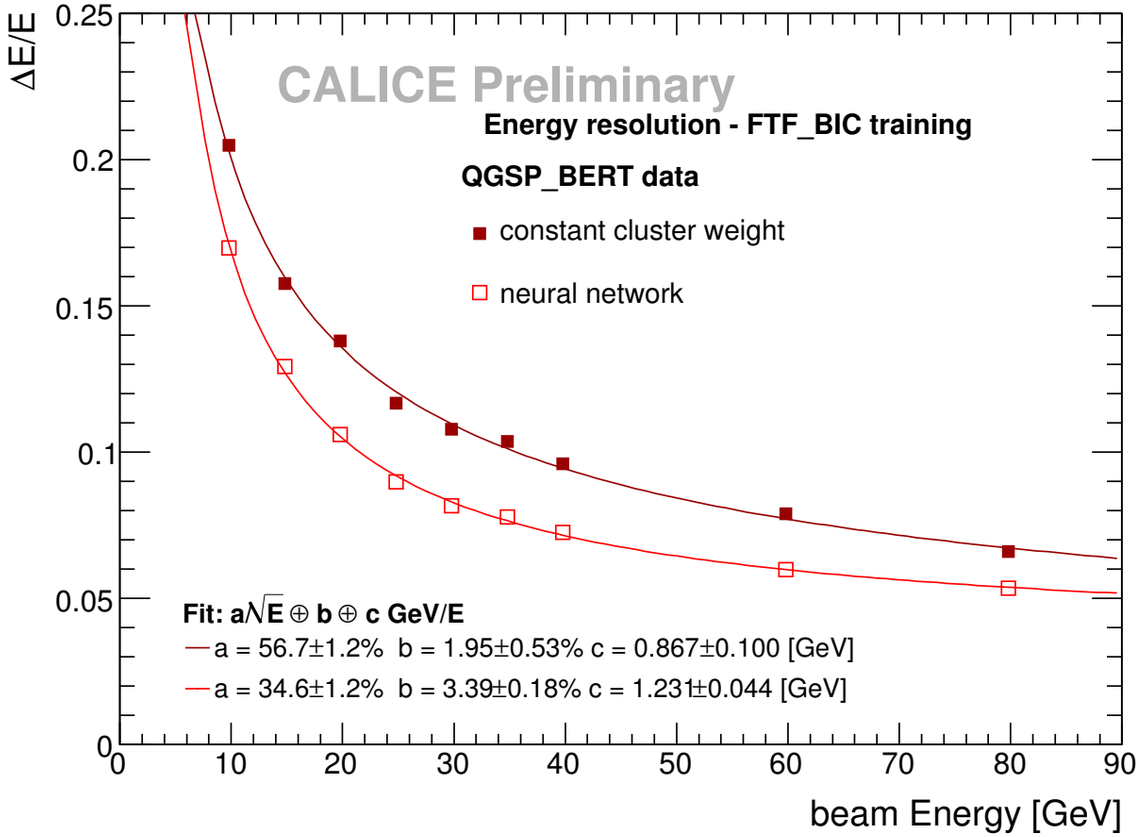


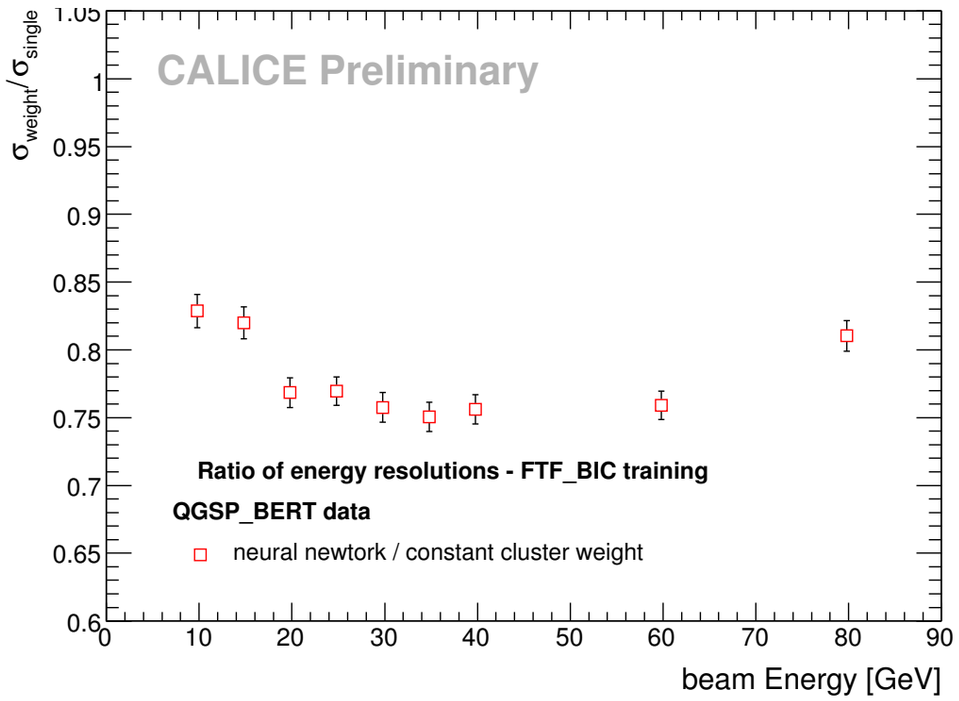
Resolution with Clustering: Data vs MC

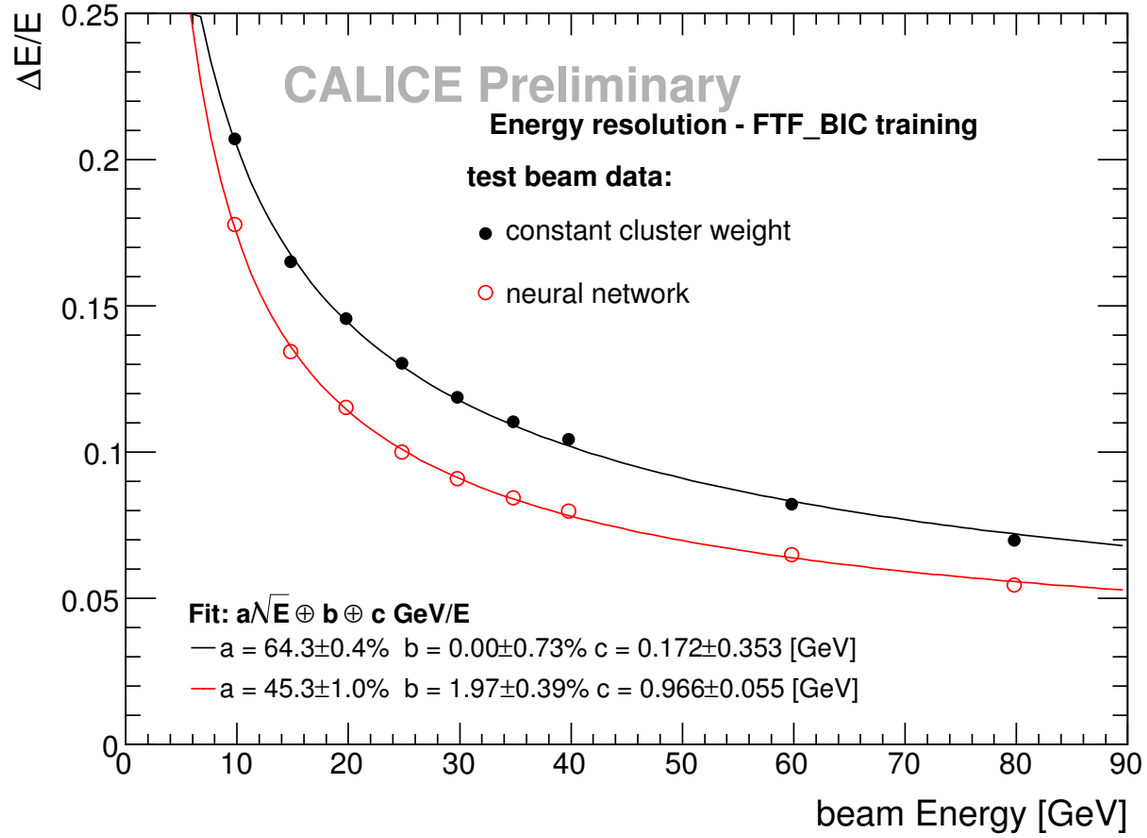


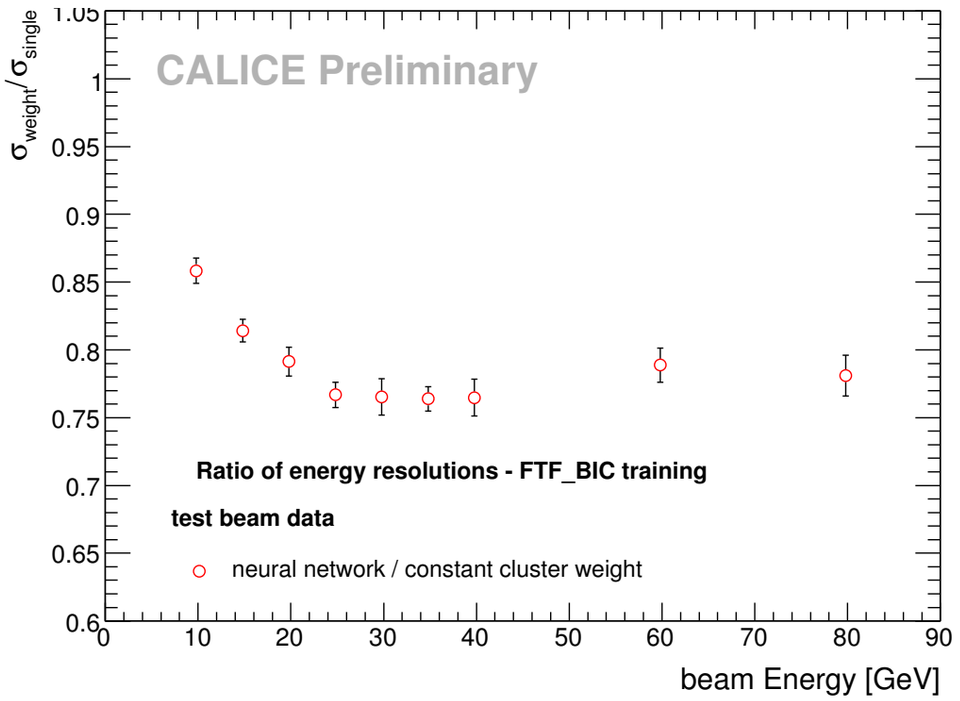


Global Compensation on MC









Cluster Variables used in NN