# Development of Reconstruction Methods by CALICE

**Frank Simon**
**Max-Planck-Institute for Physics**

*CLIC Workshop*
*CERN, January 2019*

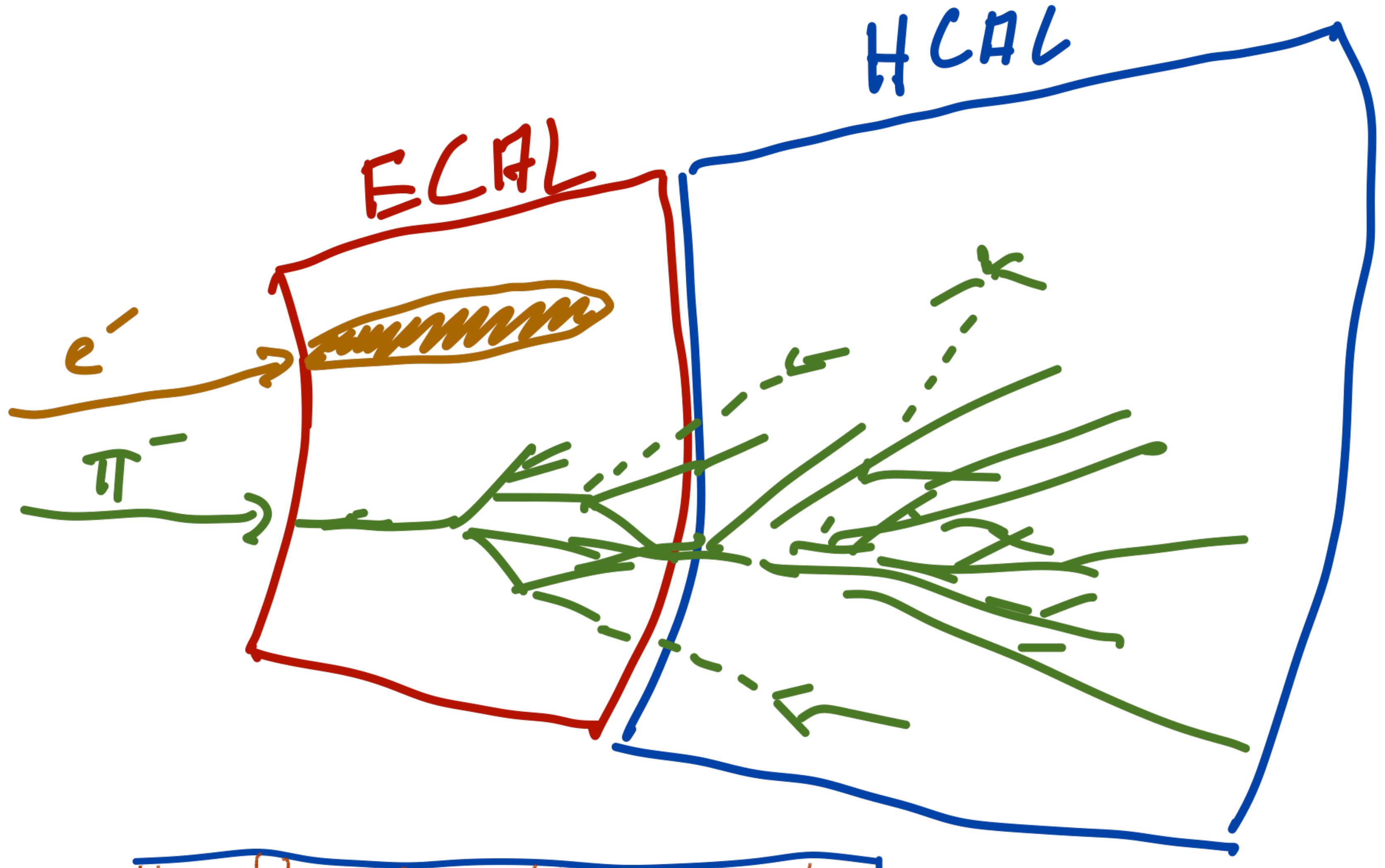
Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

# Outline

- Introduction: Energy reconstruction in calorimeters
- Software compensation: Improving hadronic energy reconstruction
- Towards more complex techniques
- Outlook

# Introduction

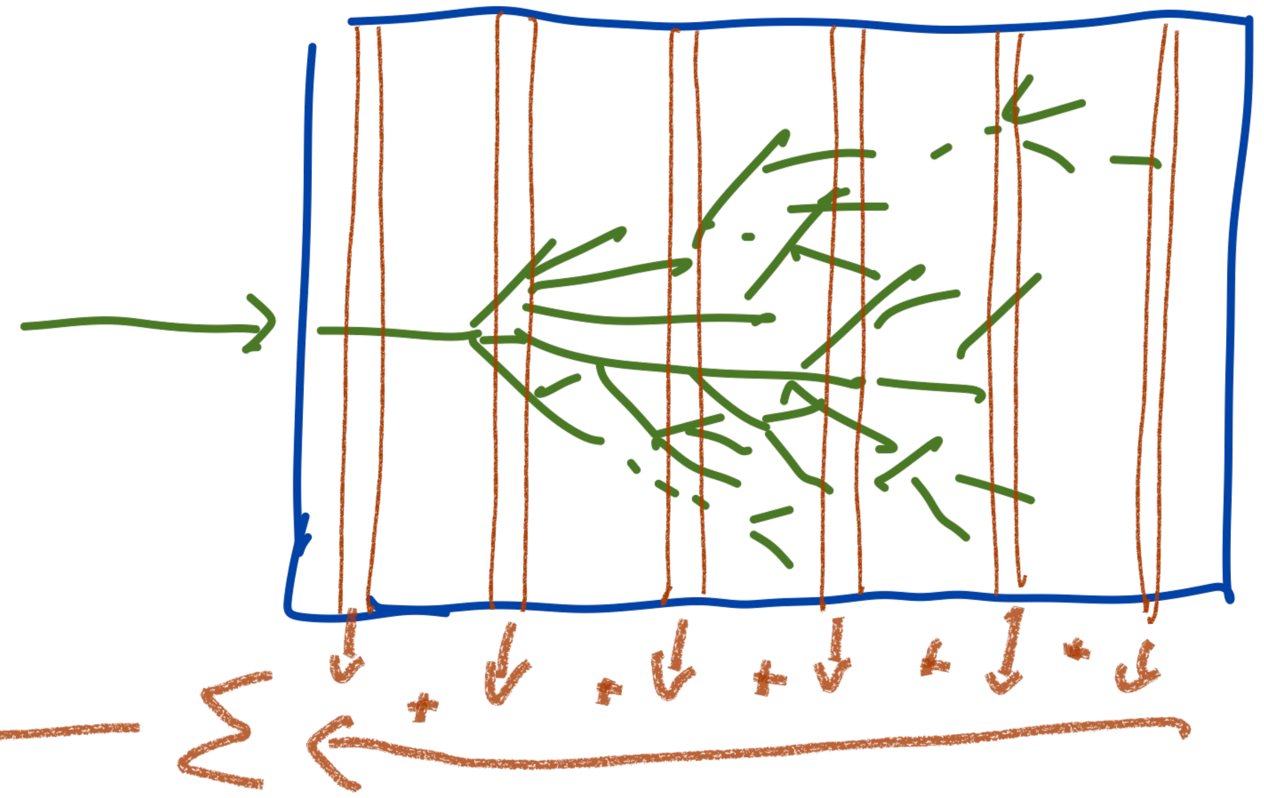
*Energy Reconstruction in Calorimeters*




- Energy reconstruction is the key task of HEP calorimeter systems - for electromagnetic and hadronic particles

- The (somewhat naive) assumption: signals seen in active detector elements is a energy- and particle type - independent fraction of the particle energy:

$$E_{reco} = c_{calib} \times \sum E_{cluster/tower/layer/hit}$$



- In practice: Particle-type and possible energy dependent “calibration” of conversion of visible energy to particle energy a minimal requirement - *with more sophistication possible & useful*

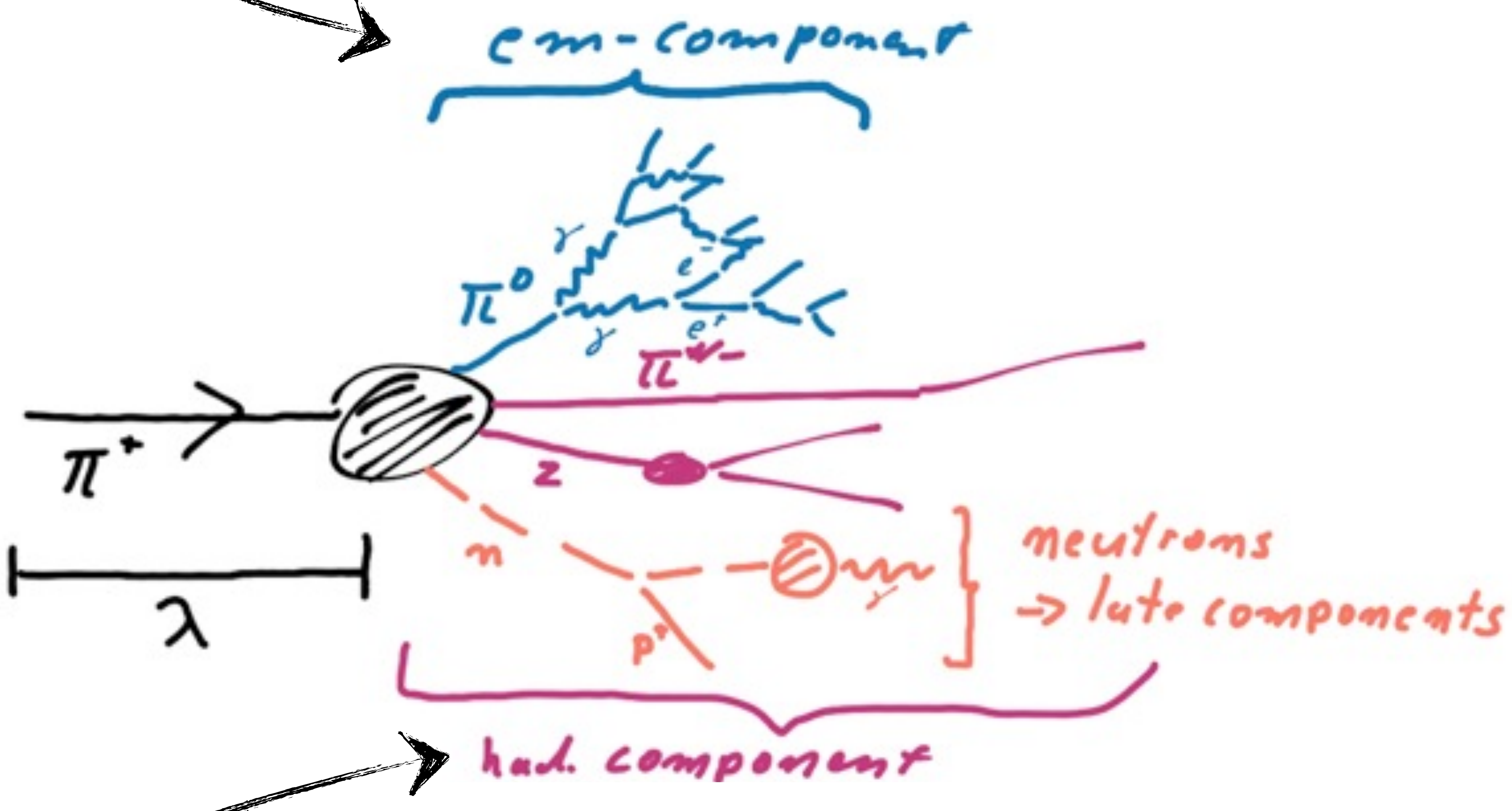
# Limitations of Energy Resolution

*… and Handles to improve it*

- The energy resolution for hadronic showers typically is relatively poor:

prompt energy depositions only

active elements see a ~ constant fraction of shower energy




“invisible” energy due to binding energy losses

delayed & displaced energy deposits due to neutrons

…

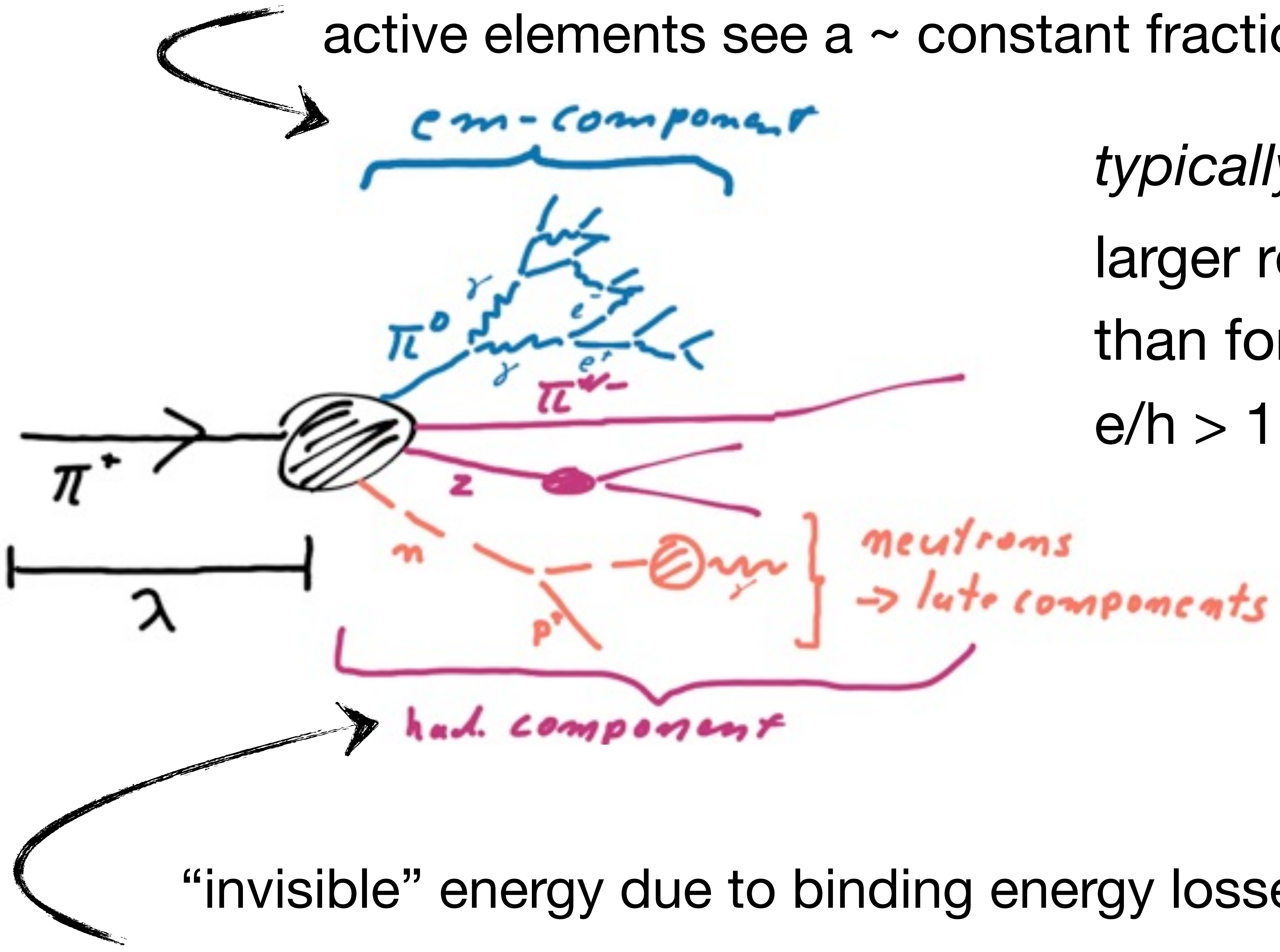
# Limitations of Energy Resolution

*… and Handles to improve it*

- The energy resolution for hadronic showers typically is relatively poor:

prompt energy depositions only

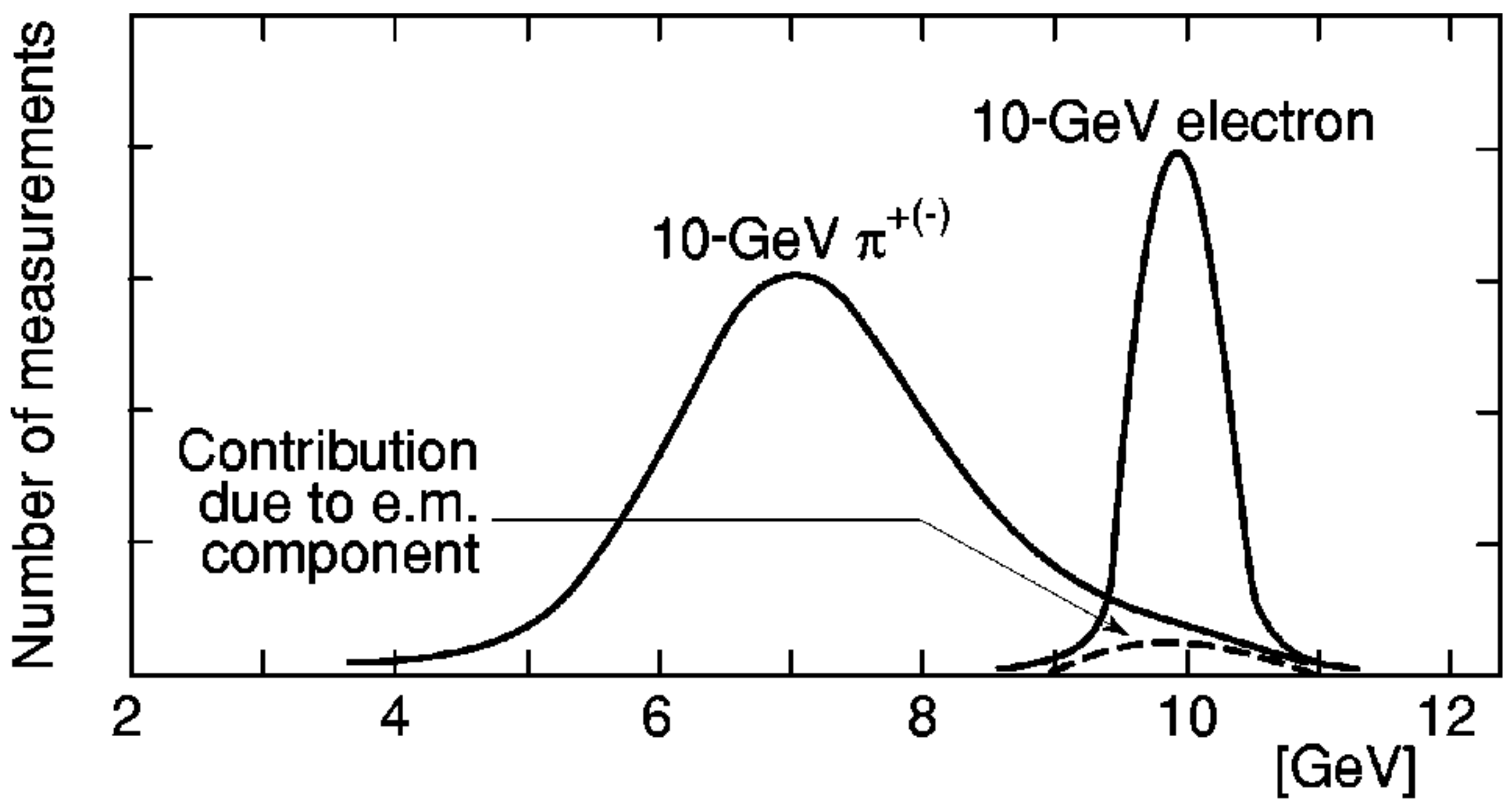
active elements see a ~ constant fraction of shower energy



*typically*:

larger response for em showers than for hadronic showers:

e/h > 1 => non-compensating




Signal (in energy units) obtained for a 10 GeV energy deposit

"invisible" energy due to binding energy losses

delayed & displaced energy deposits due to neutrons

…

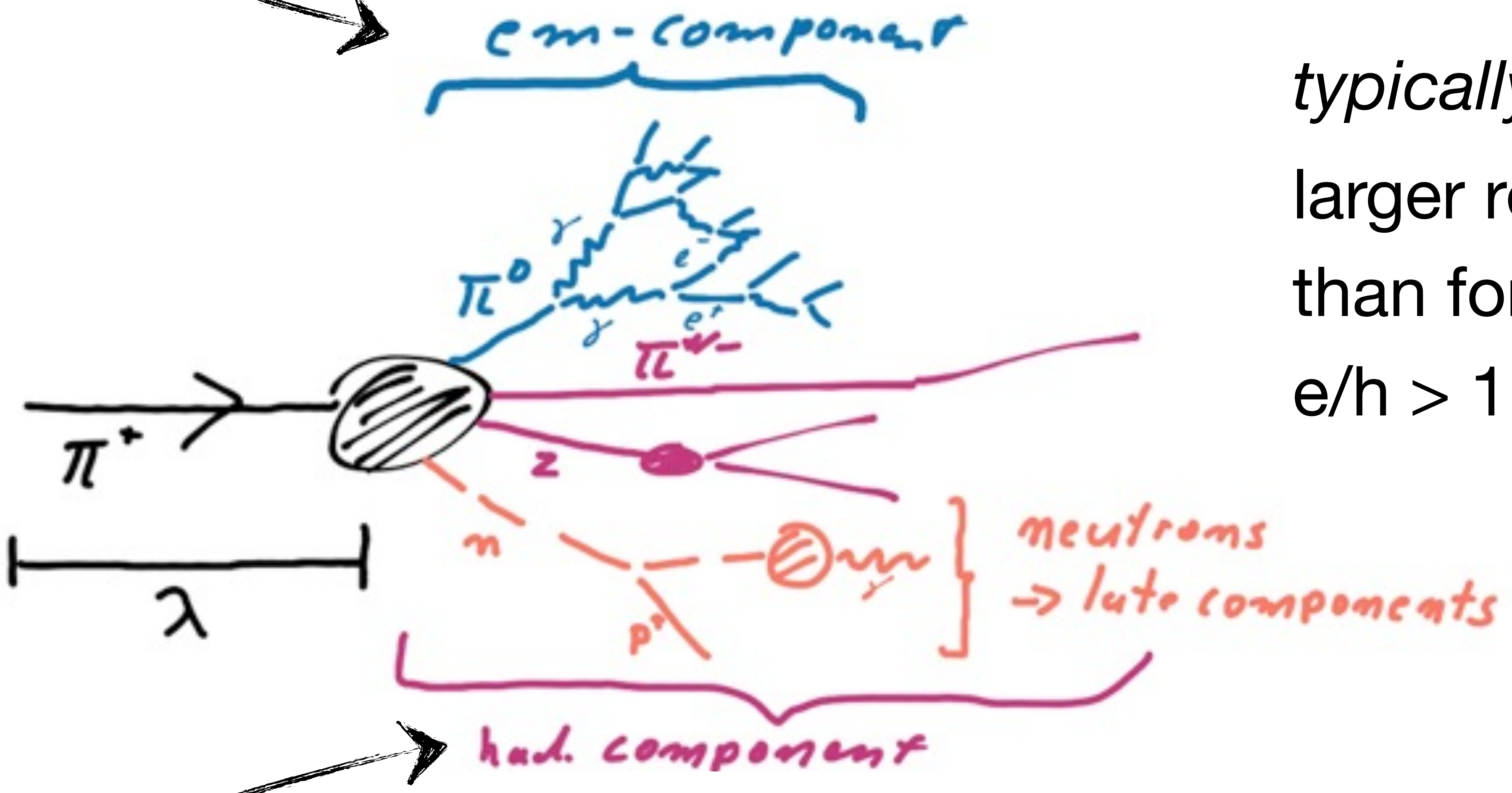
# Limitations of Energy Resolution

*… and Handles to improve it*

- The energy resolution for hadronic showers typically is relatively poor:

prompt energy depositions only

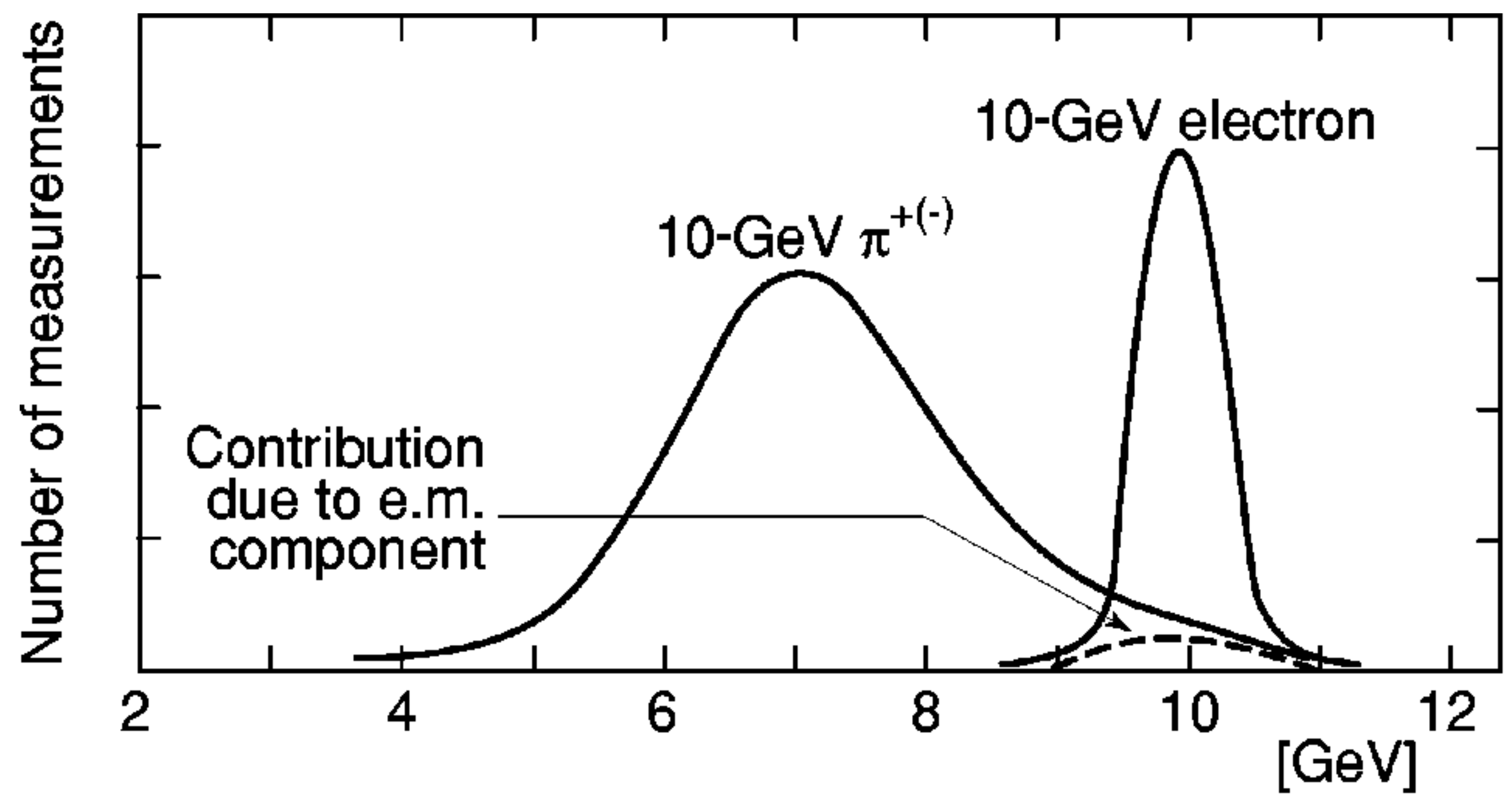
active elements see a ~ constant fraction of shower energy




*typically*:

larger response for em showers than for hadronic showers:

e/h > 1 => non-compensating




Signal (in energy units) obtained for a 10 GeV energy deposit

"invisible" energy due to binding energy losses

delayed & displaced energy deposits due to neutrons

…

The path to a better energy resolution:

➯ Compensating calorimeters: Highest potential provided by *Dual Readout*

➯ *Software compensation* / offline weighting: Shower-by-shower energy corrections, profits from high granularity

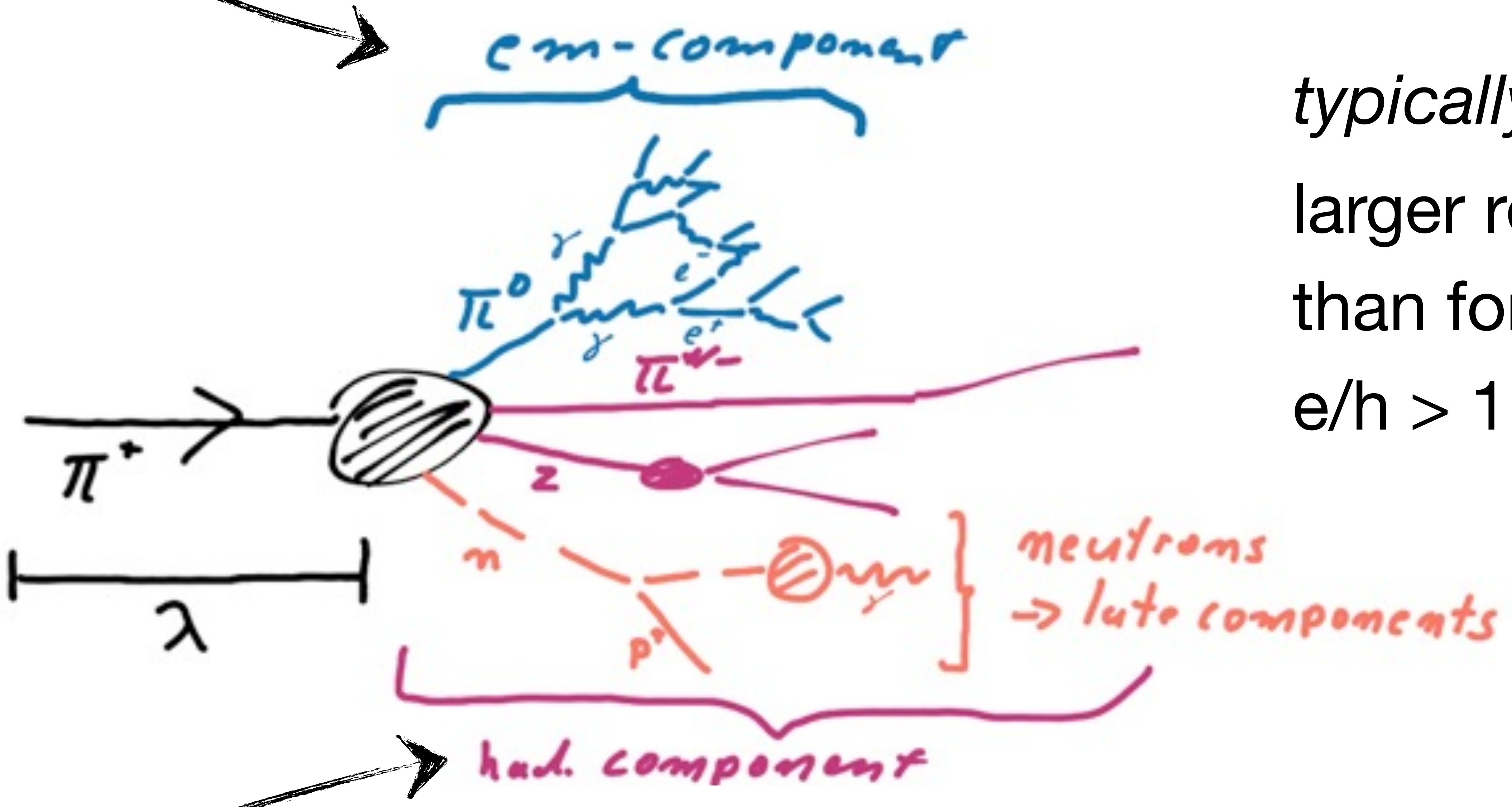
# Limitations of Energy Resolution

*… and Handles to improve it*

- The energy resolution for hadronic showers typically is relatively poor:

prompt energy depositions only

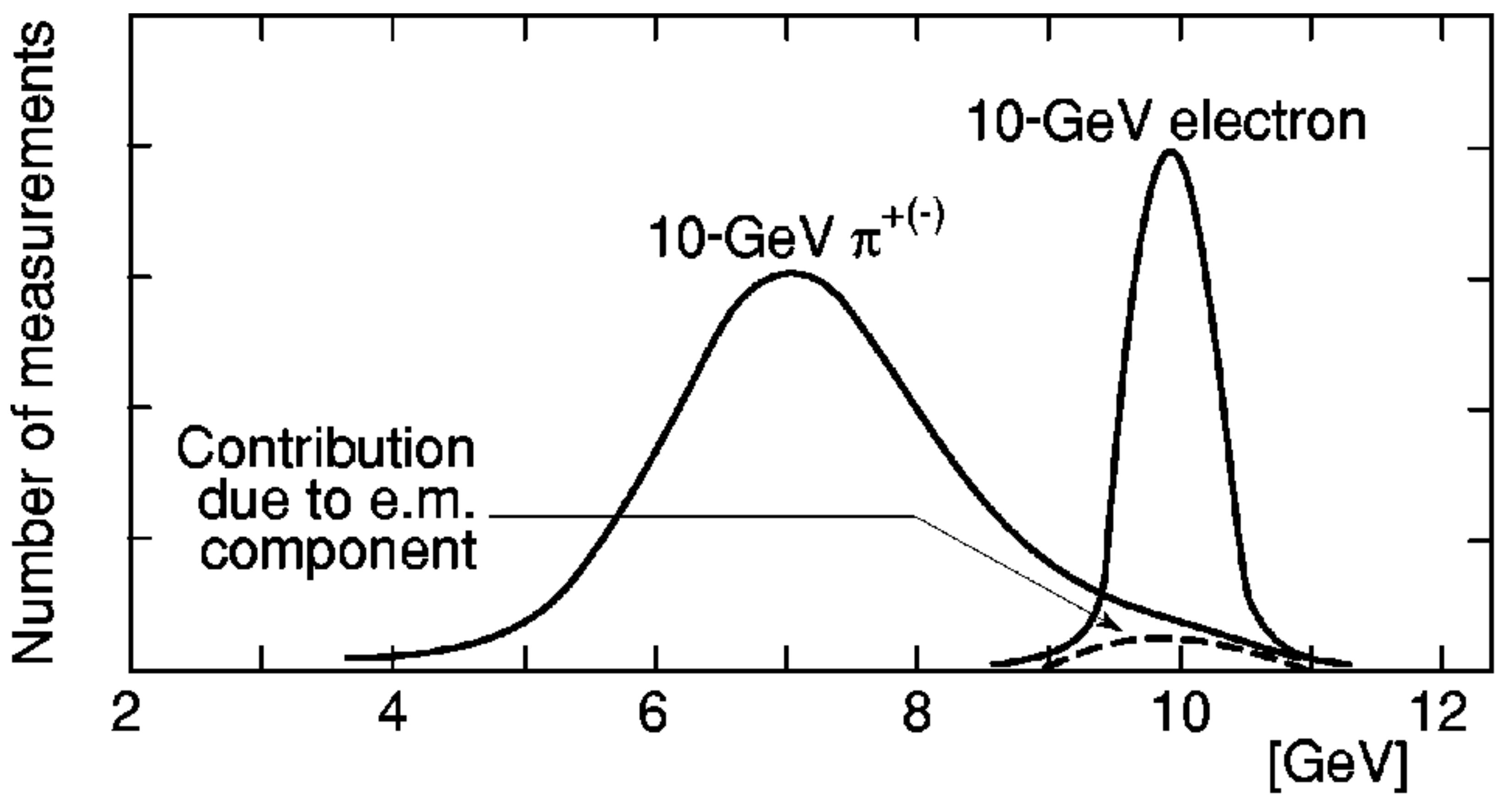
active elements see a ~ constant fraction of shower energy




*typically*:

larger response for em showers than for hadronic showers:

e/h > 1 => non-compensating




"invisible" energy due to binding energy losses

delayed & displaced energy deposits due to neutrons

…

The path to a better energy resolution:

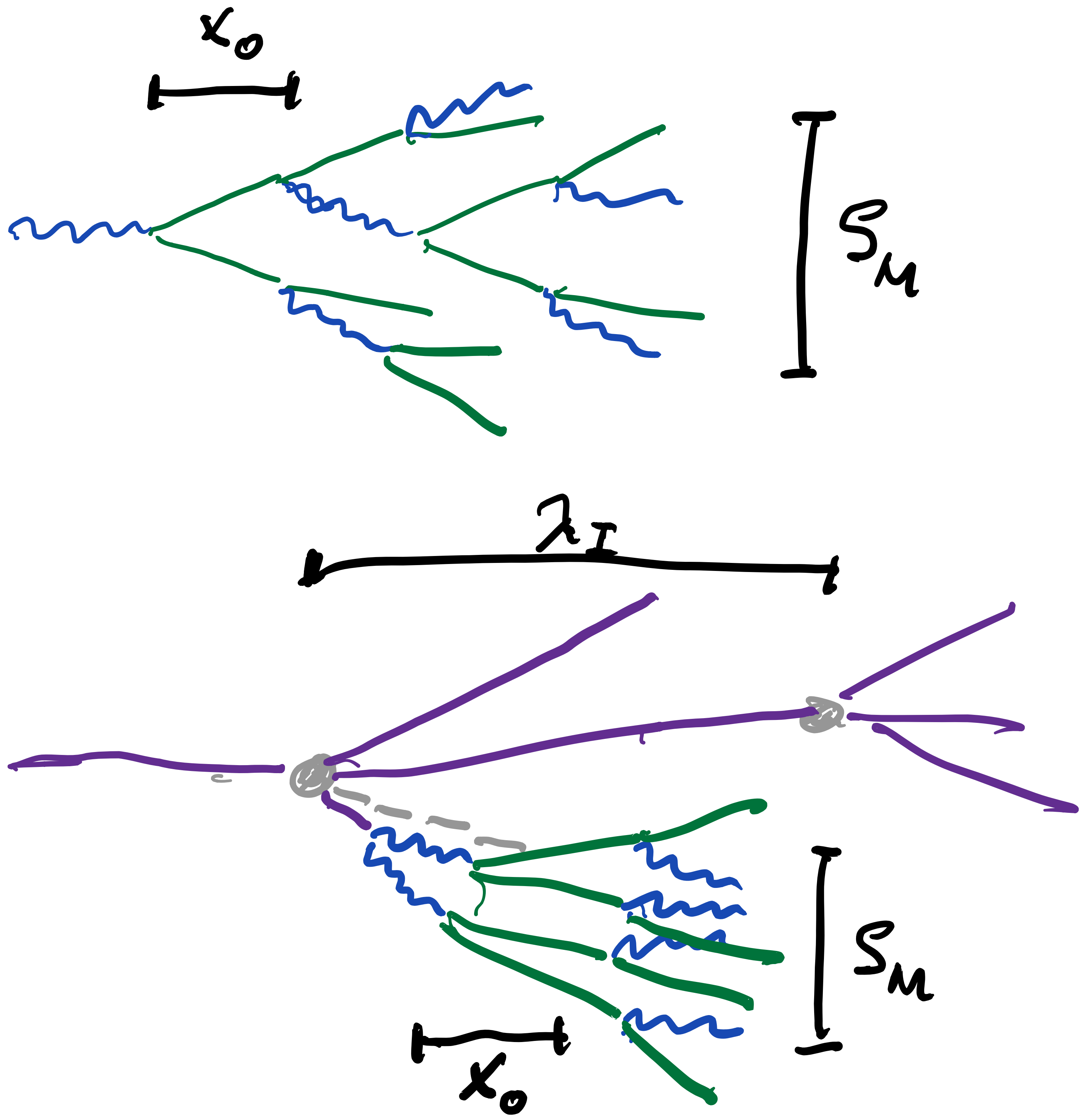
⇨ Compensating calorimeters: Highest potential provided by *Dual Readout*

⇨ *Software compensation* / offline weighting: Shower-by-shower energy corrections, profits from high granularity

# CALICE Calorimeters

*Granularity & Prototypes*

- Granularity motivated by shower physics:



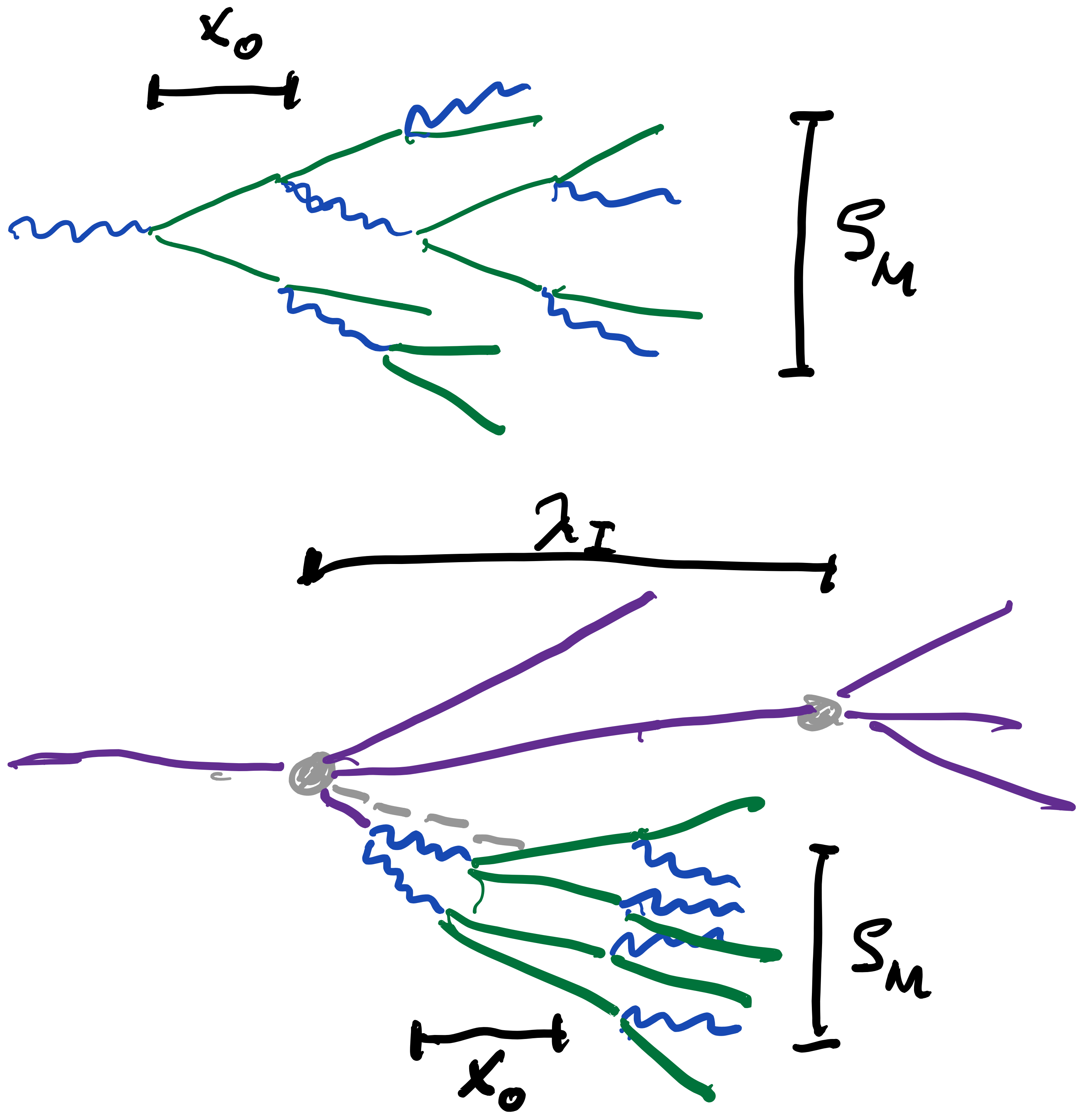

Calorimeter voxel size given by

**$X_0$, $\rho_M$ =>** ~ *(5 mm)$^3$ - (30 mm)$^3$*

# CALICE Calorimeters

*Granularity & Prototypes*

- Granularity motivated by shower physics:



Calorimeter voxel size given by
**$X_0$, $\rho_M$ =>** ~ *(5 mm)$^3$ - (30 mm)$^3$*

*Consequences for the Calorimeter Systems:*

⇨ *O* $10^{7-8}$ cells in HCAL, $10^8$ cells in ECAL for typical detector systems!
(compared to a few 10k - 100k for current LHC detectors)

⇨ fully integrated electronics needed

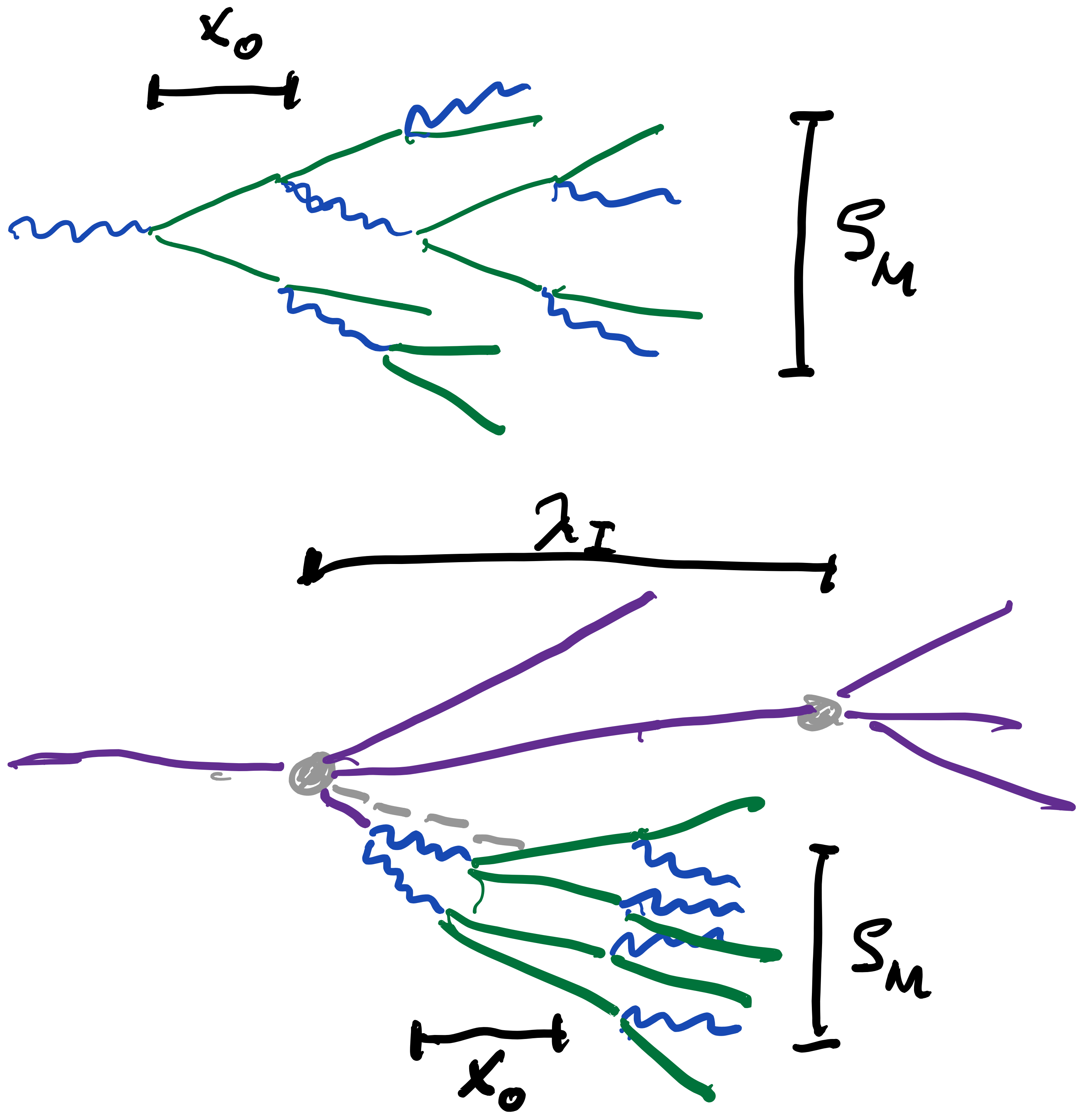
⇨ requires active elements that support high granularity and large channel counts

⇨ need technical solutions amenable to mass production & automatisation

# CALICE Calorimeters

*Granularity & Prototypes*

- Granularity motivated by shower physics:




Calorimeter voxel size given by
**$X_0$, $\rho_M$ =>** ~ *(5 mm)$^3$ - (30 mm)$^3$*

*Consequences for the Calorimeter Systems:*

➯ *O* $10^{7-8}$ cells in HCAL, $10^8$ cells in ECAL for typical detector systems!
(compared to a few 10k - 100k for current LHC detectors)

➯ fully integrated electronics needed

➯ requires active elements that support high granularity and large channel counts

➯ need technical solutions amenable to mass production & automatisation

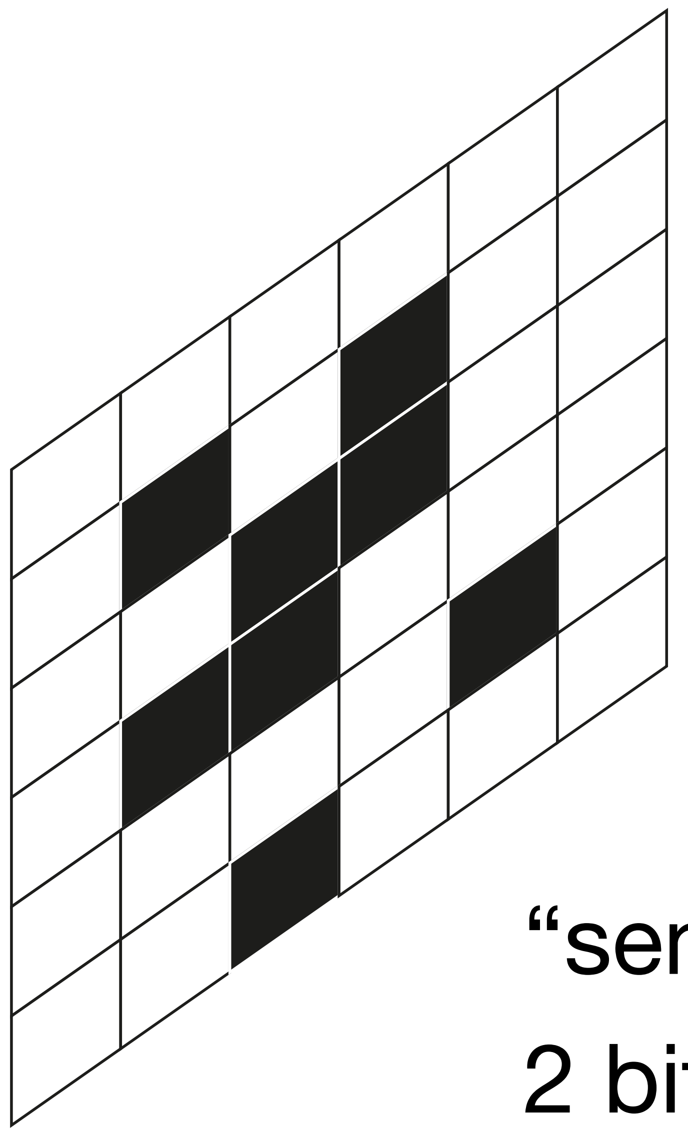
Developed and studied in CALICE
Principles, performance, technological feasibility and scalability demonstrated in the last 12 years

# CALICE Calorimeters

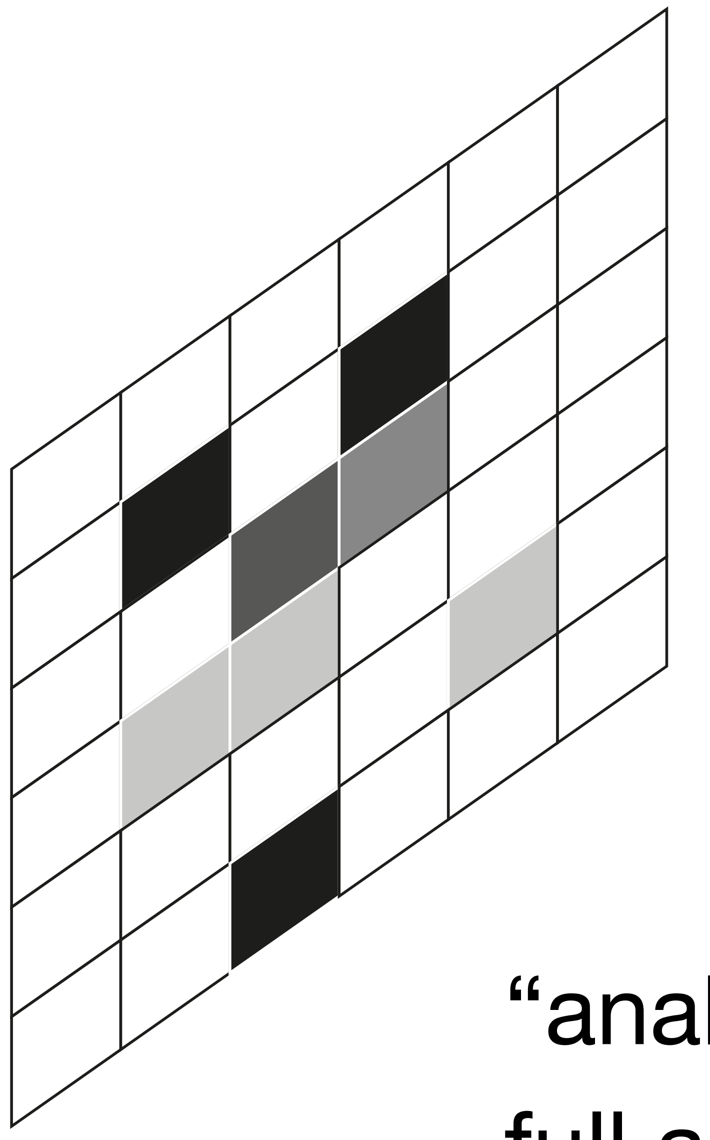
*Readout Schemes*

- Depending on active detector technology and granularity, different readout schemes are used:

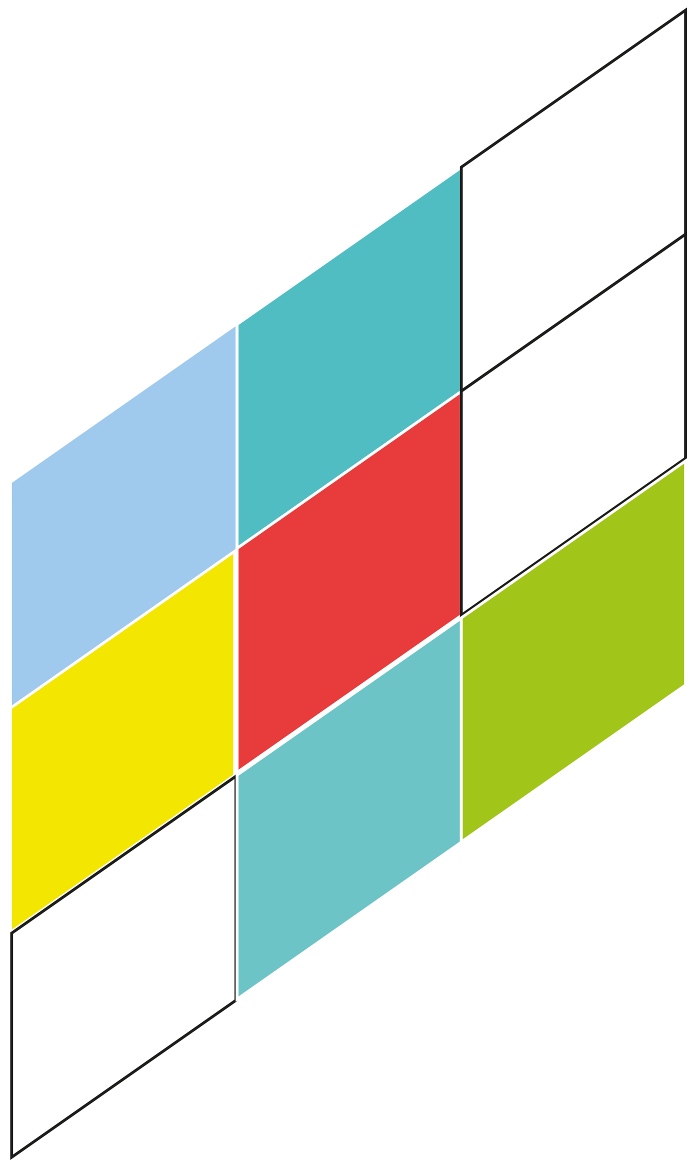
“digital”
counting hit cells



“semi-digital”
2 bits per cell:
one, a few, many particles



“analog”
full analog information
typically 14+ bits



| Active elements: | gas detectors | gas detectors | Plastic scintillator elements |
|---|---|---|---|
| | Silicon pixel detectors | | Silicon pad detectors |

# Semi-Digital Energy Reconstruction in CALICE

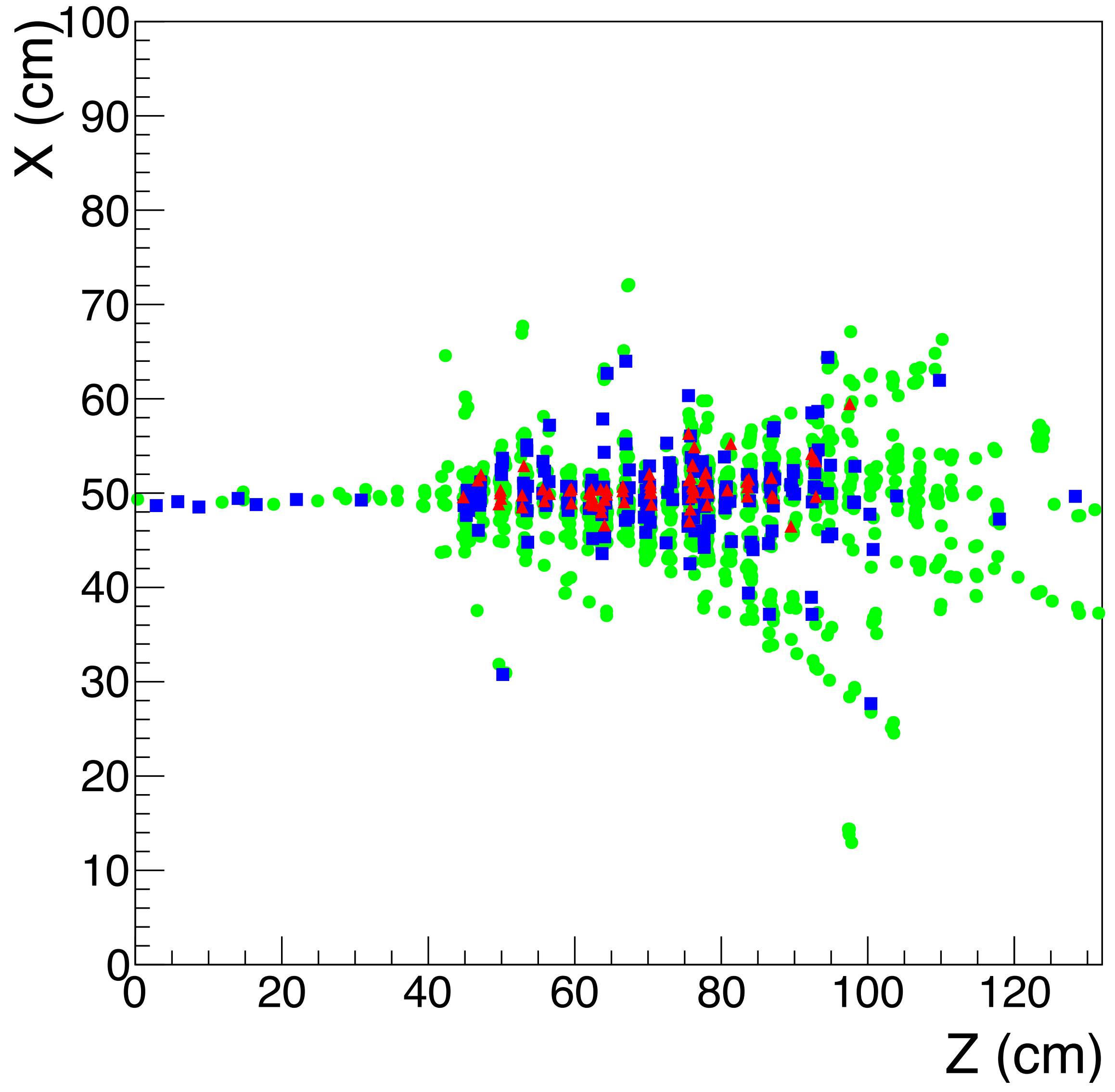
*Performance of the RPC SDHCAL*

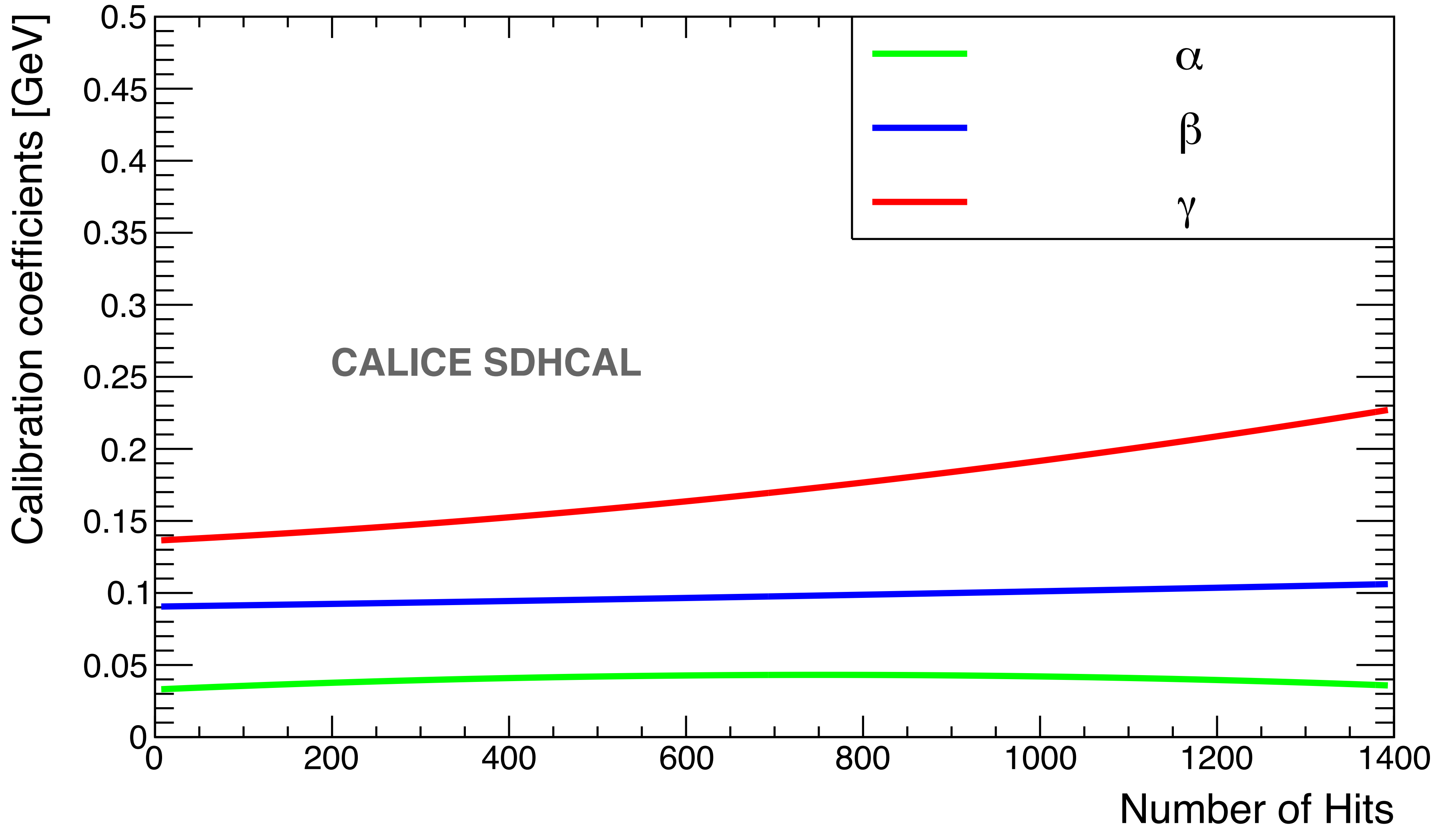
- Exploitation of semi-digital readout requires energy (hit) - dependent calibration factors for each of the three thresholds

$$E_{\text{reco}} = \alpha N_1 + \beta N_2 + \gamma N_3$$

parameters determined by optimising energy resolution

$$\chi^2 = \sum_{i=1}^{N} \frac{(E^i_{\text{beam}} - E^i_{\text{reco}})^2}{\sigma_i^2}$$



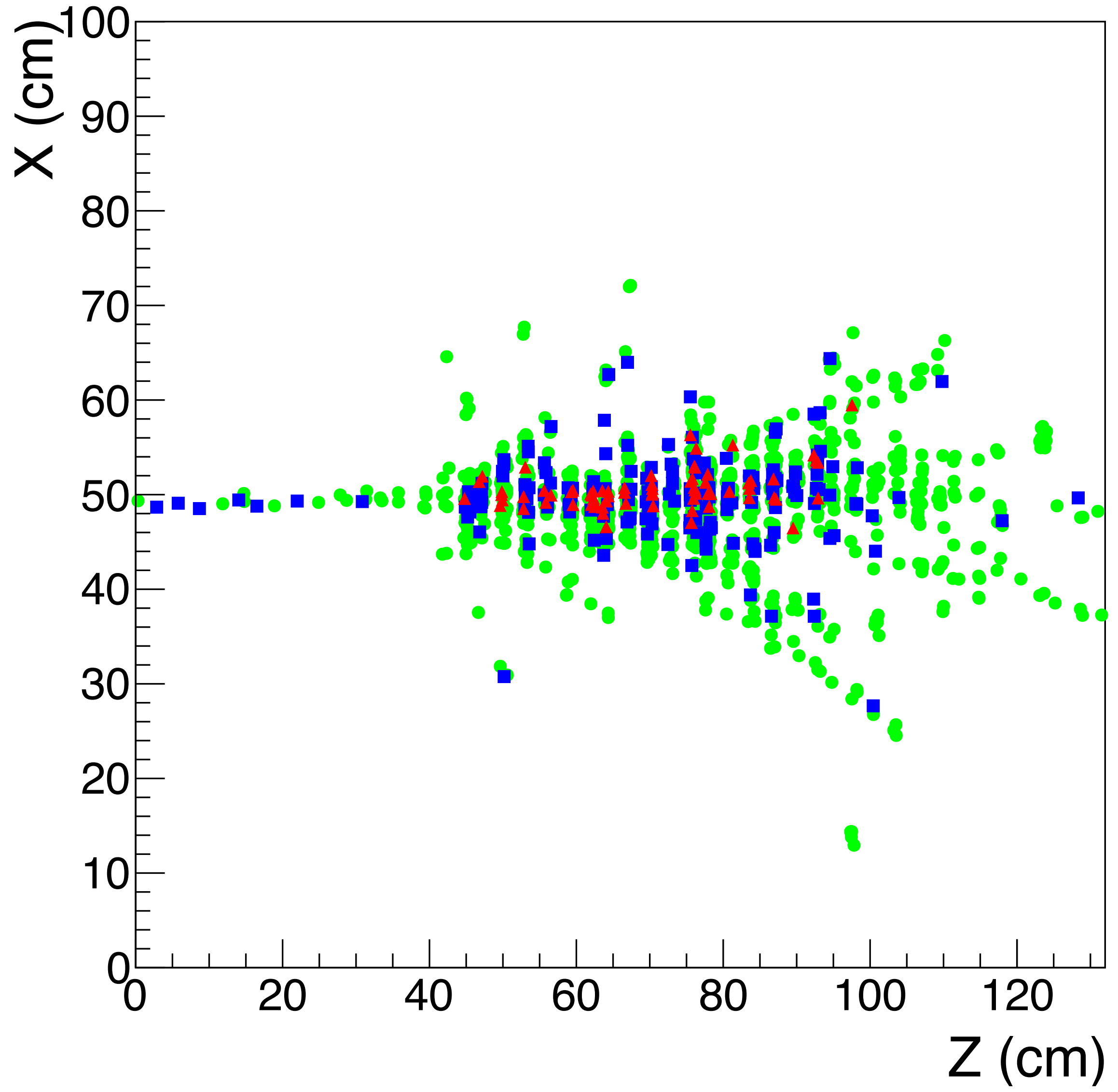

JINST 11 P04001 (2016)

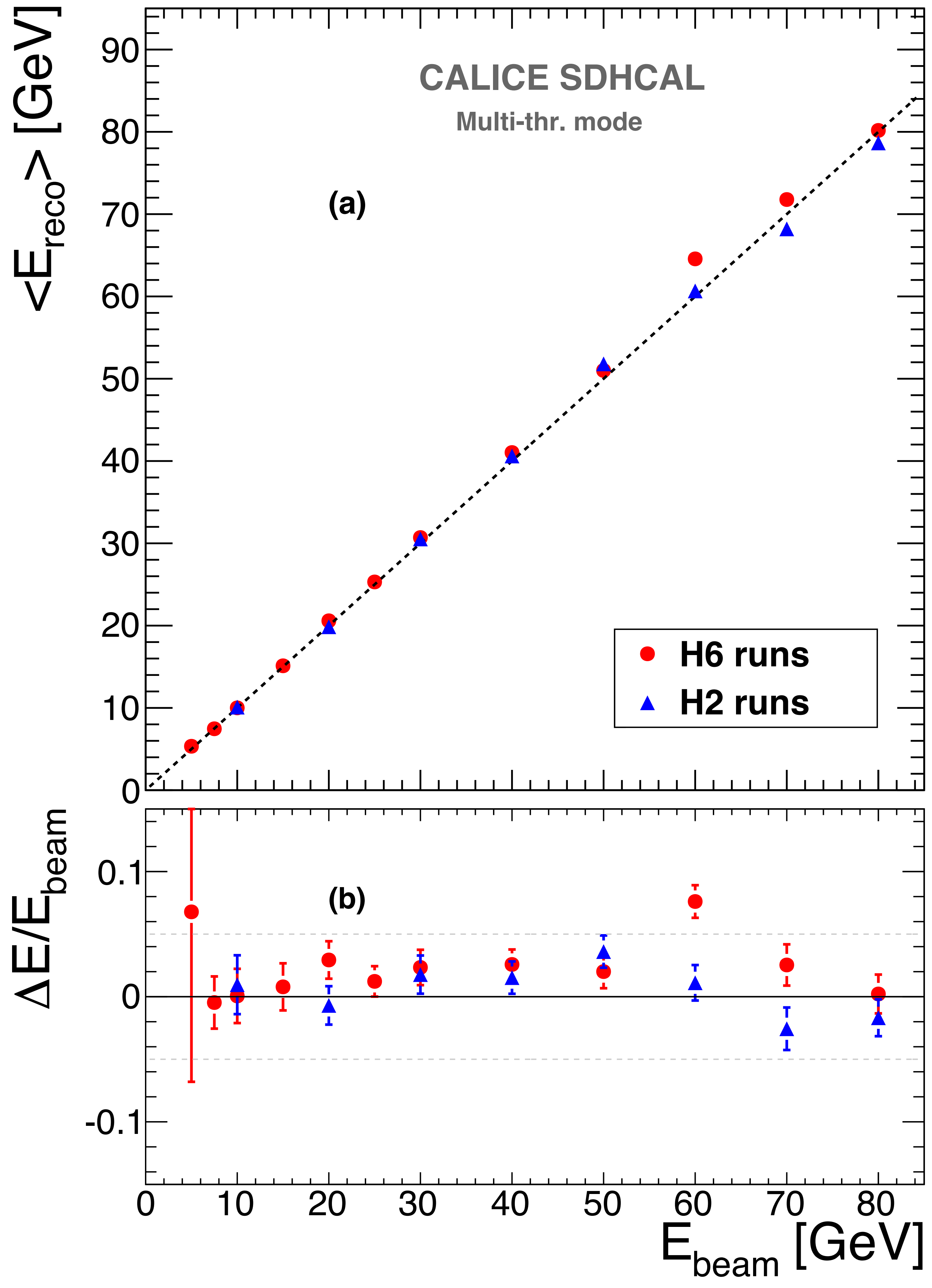
- Exploitation of semi-digital readout requires energy (hit) - dependent calibration factors for each of the three thresholds

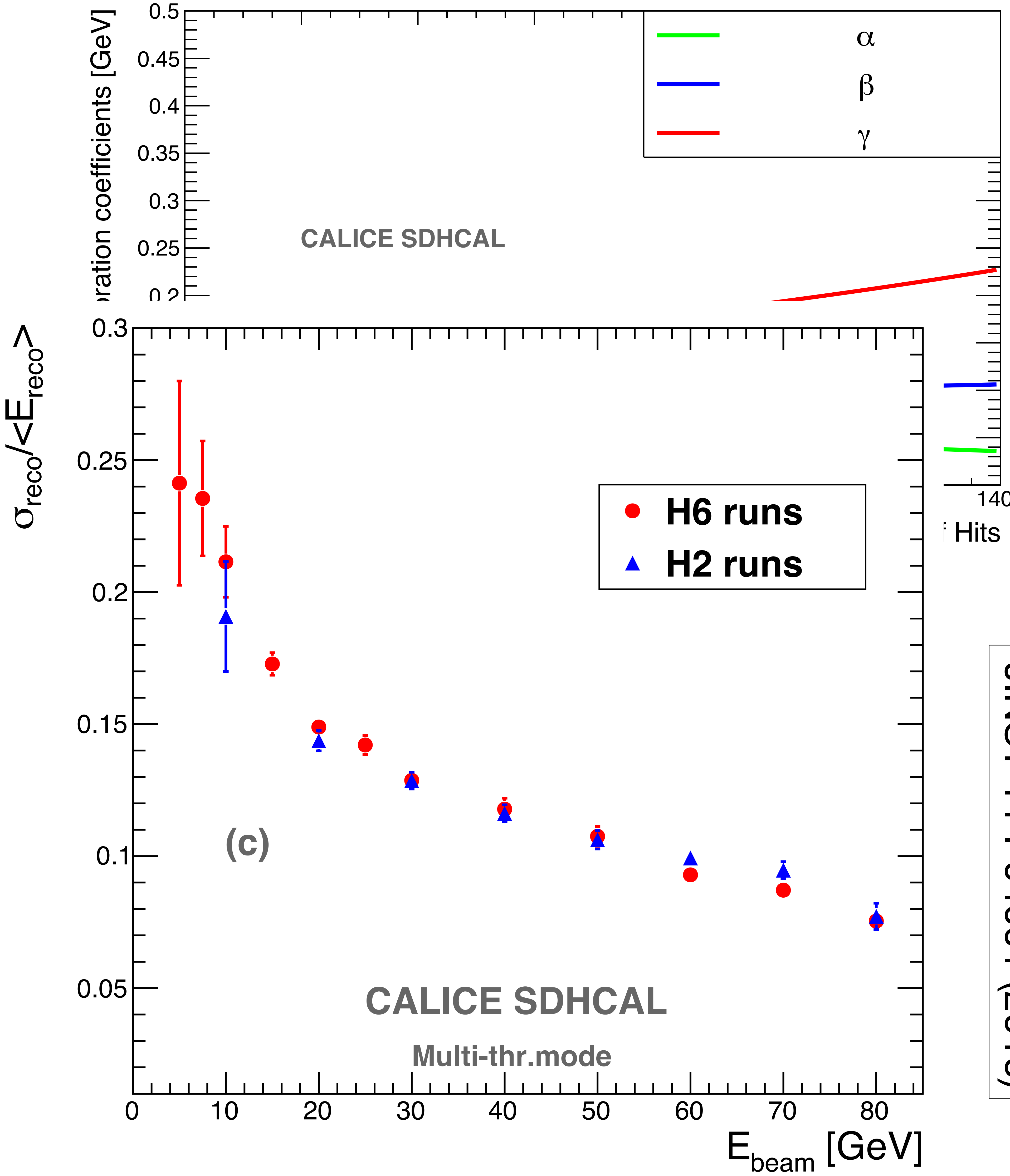
$$E_{\mathrm{reco}} = \alpha N_1 + \beta N_2 + \gamma N_3$$

parameters determined by optimising energy resolution

$$\chi^2 = \sum_{i=1}^{N} \frac{(E^i_{\mathrm{beam}} - E^i_{\mathrm{reco}})^2}{\sigma_i^2}$$




JINST 11 P04001 (2016)

# Software Compensation in CALICE

*Different Techniques to improve the Energy Resolution with Analog Readout*

- Full analog energy information in each cell of the AHCAL provides different handles to implement energy reconstruction techniques. Two main strategies for software compensation studied:

JINST 7 P09017 (2012)

# Software Compensation in CALICE

*Different Techniques to improve the Energy Resolution with Analog Readout*

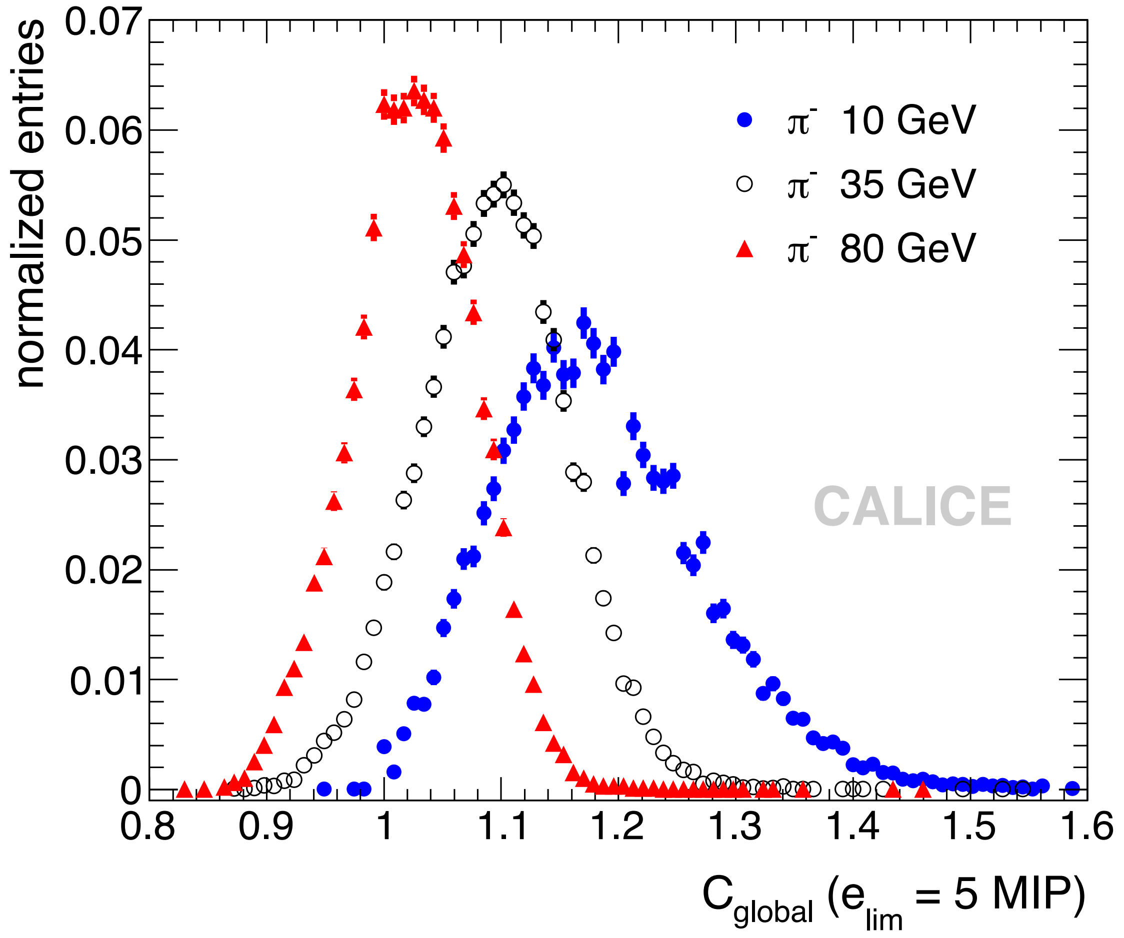
- Full analog energy information in each cell of the AHCAL provides different handles to implement energy reconstruction techniques. Two main strategies for software compensation studied:

## *Global*

- Event-by-event correction of energy sum with a shower-dependent *global* factor

correction based on $c_{global}$, given by

$$c_{global} = \frac{N_{hits}(E_{hit} < e_{lim})}{N_{hits}(E_{hit} < \langle E_{hit} \rangle)}$$




with an additional energy dependence of the correction factor

$e_{lim}$ = 5 MIP

# Software Compensation in CALICE

*Different Techniques to improve the Energy Resolution with Analog Readout*

JINST 7 P09017 (2012)

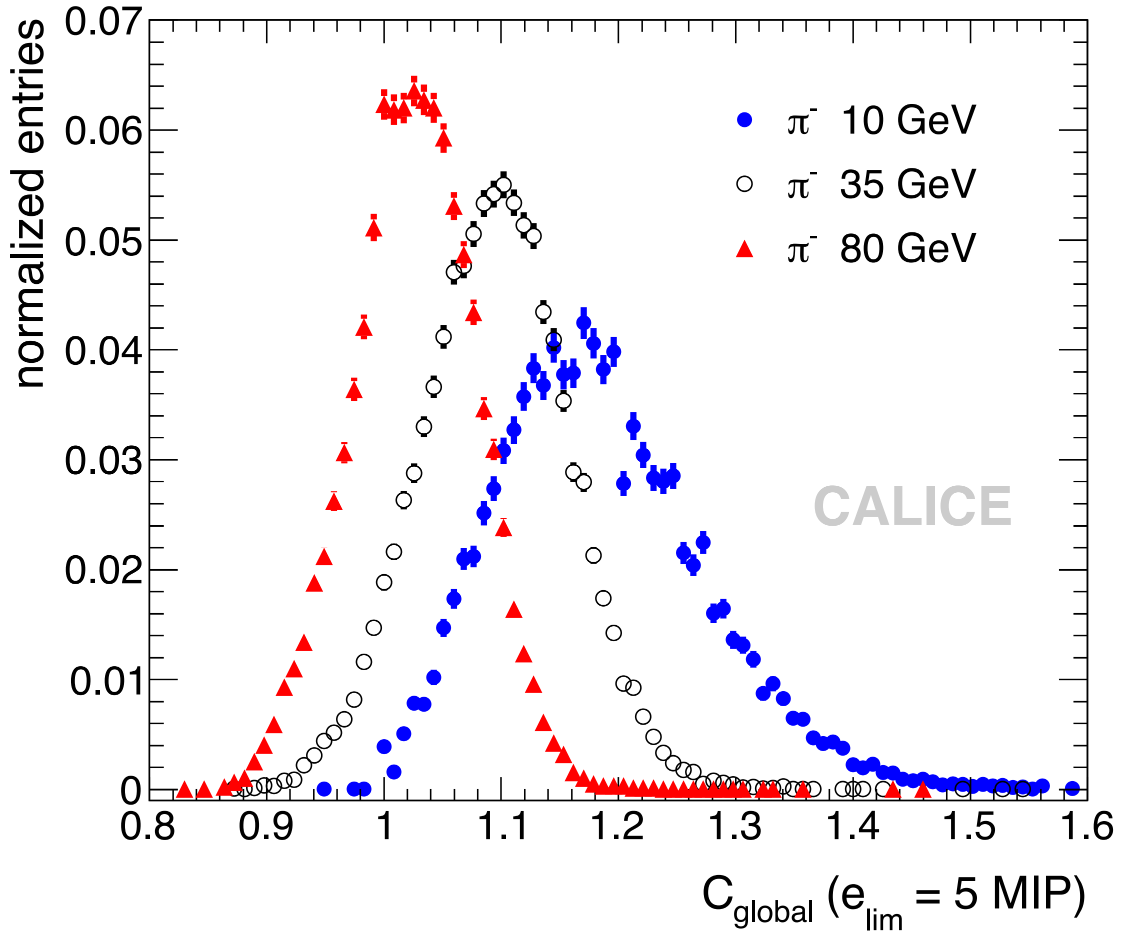
- Full analog energy information in each cell of the AHCAL provides different handles to implement energy reconstruction techniques. Two main strategies for software compensation studied:

***Global***

- Event-by-event correction of energy sum with a shower-dependent *global* factor

correction based on $c_{global}$, given by

$$c_{global} = \frac{N_{hits}(E_{hit} < e_{lim})}{N_{hits}(E_{hit} < \langle E_{hit} \rangle)}$$

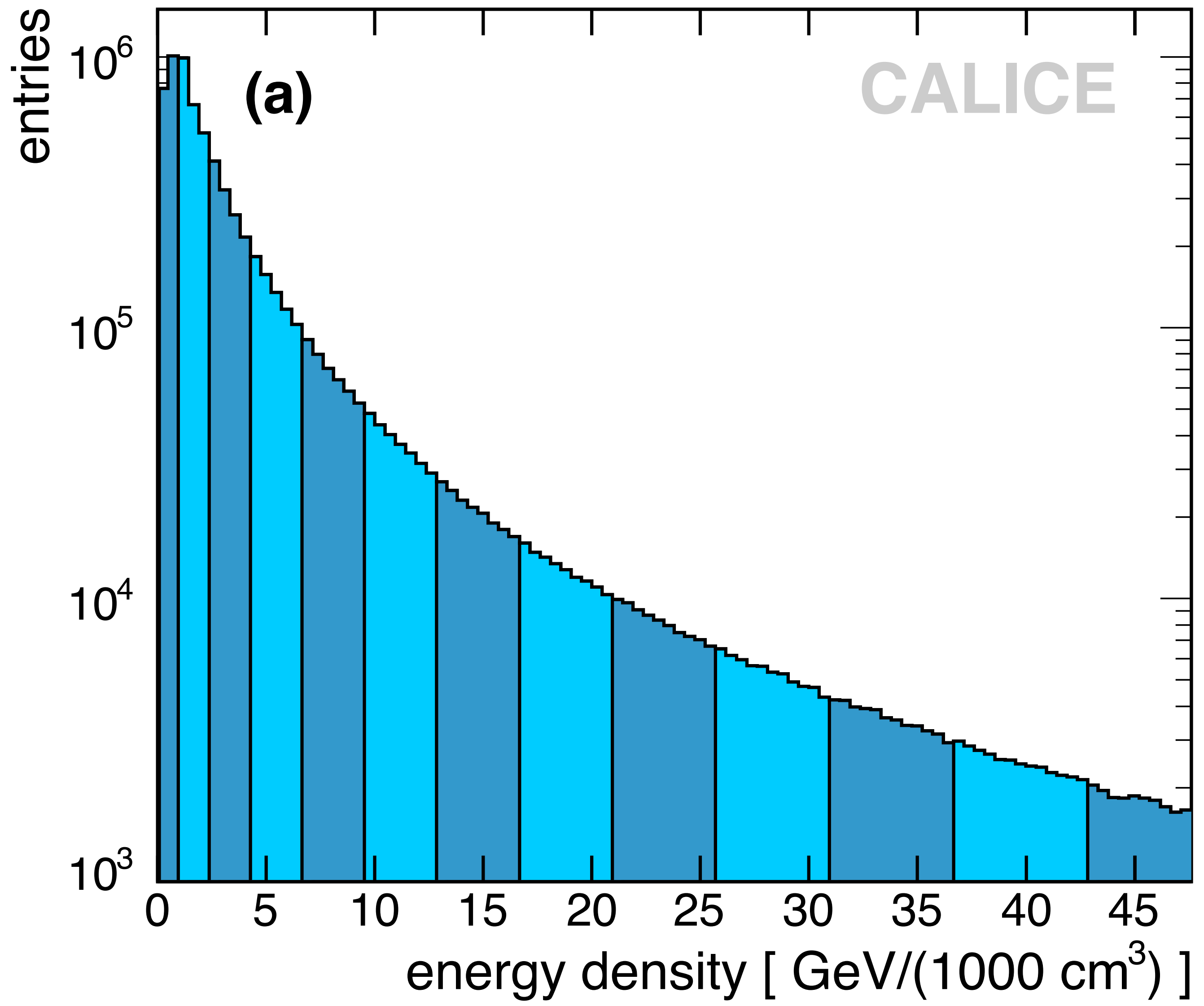



with an additional energy dependence of the correction factor

$e_{lim}$ = 5 MIP

***Local***

- Cell-by-cell correction of energy with energy-density dependent *weights*




additional parametrisation for energy dependence of the weights, separate weights for each energy-density bin

# Software Compensation in CALICE

*Different Techniques to improve the Energy Resolution with Analog Readout*

JINST 7 P09017 (2012)

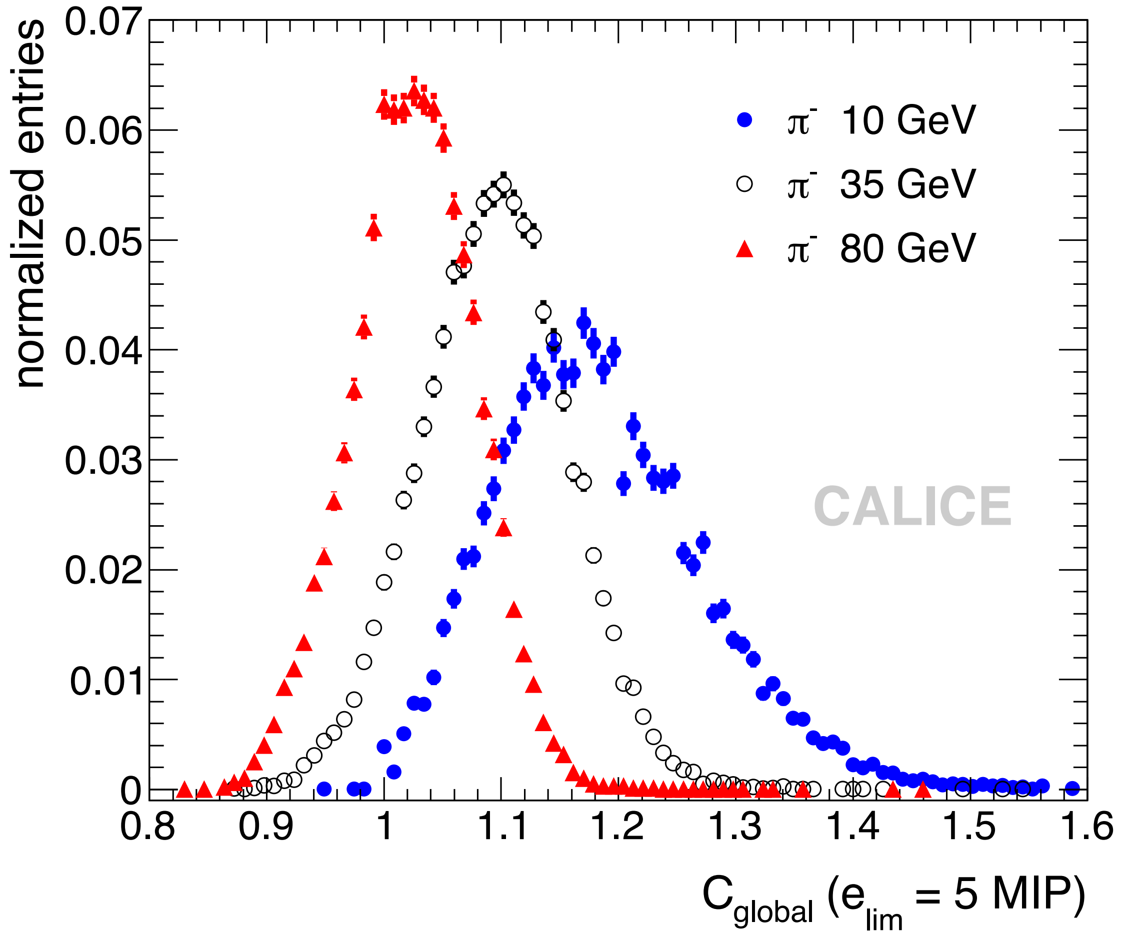
- Full analog energy information in each cell of the AHCAL provides different handles to implement energy reconstruction techniques. Two main strategies for software compensation studied:

### *Global*

- Event-by-event correction of energy sum with a shower-dependent *global* factor

correction based on $c_{global}$, given by

$$c_{global} = \frac{N_{hits}(E_{hit} < e_{lim})}{N_{hits}(E_{hit} < \langle E_{hit} \rangle)}$$

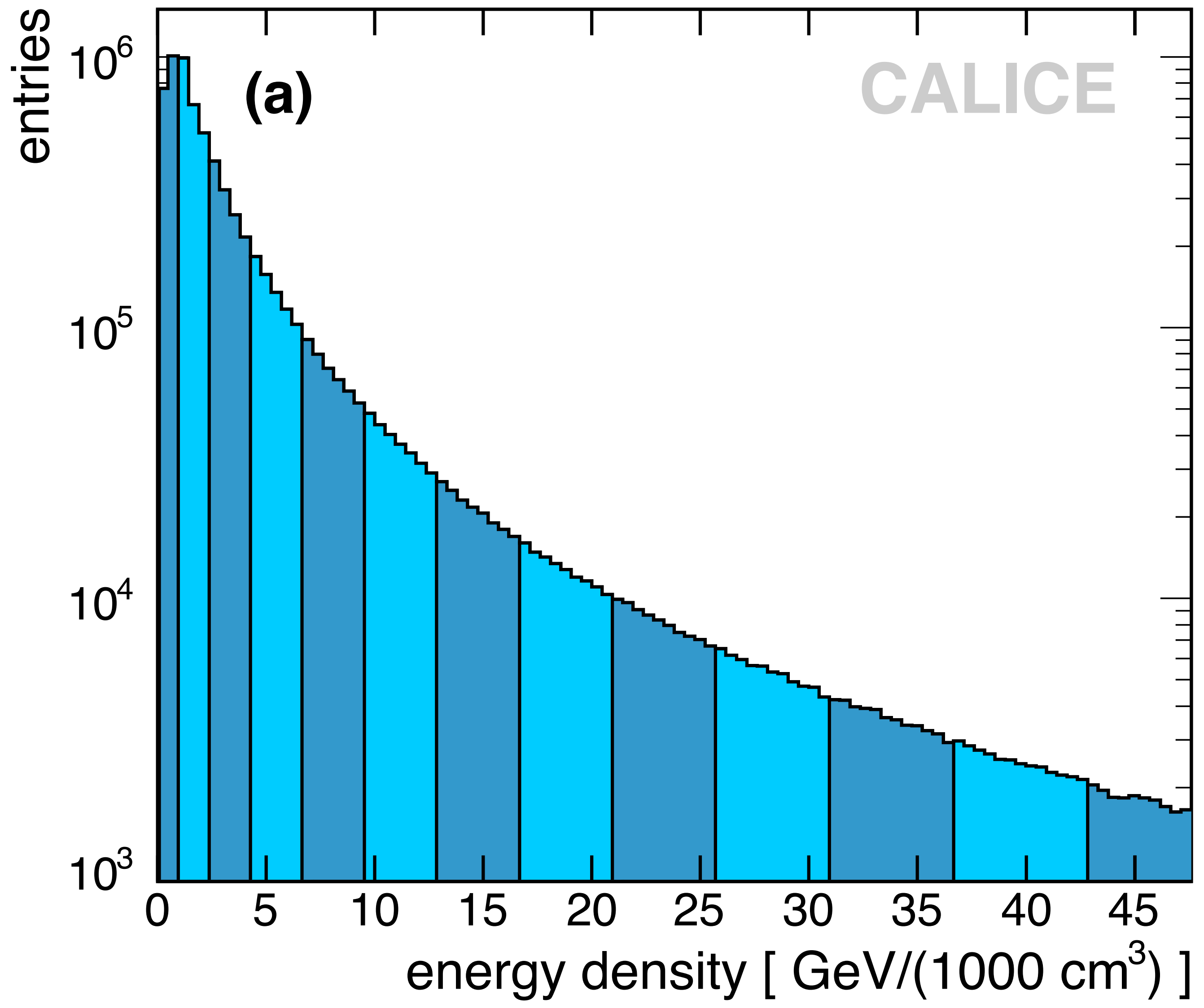



with an additional energy dependence of the correction factor

$e_{lim}$ = 5 MIP

### *Local*

- Cell-by-cell correction of energy with energy-density dependent *weights*

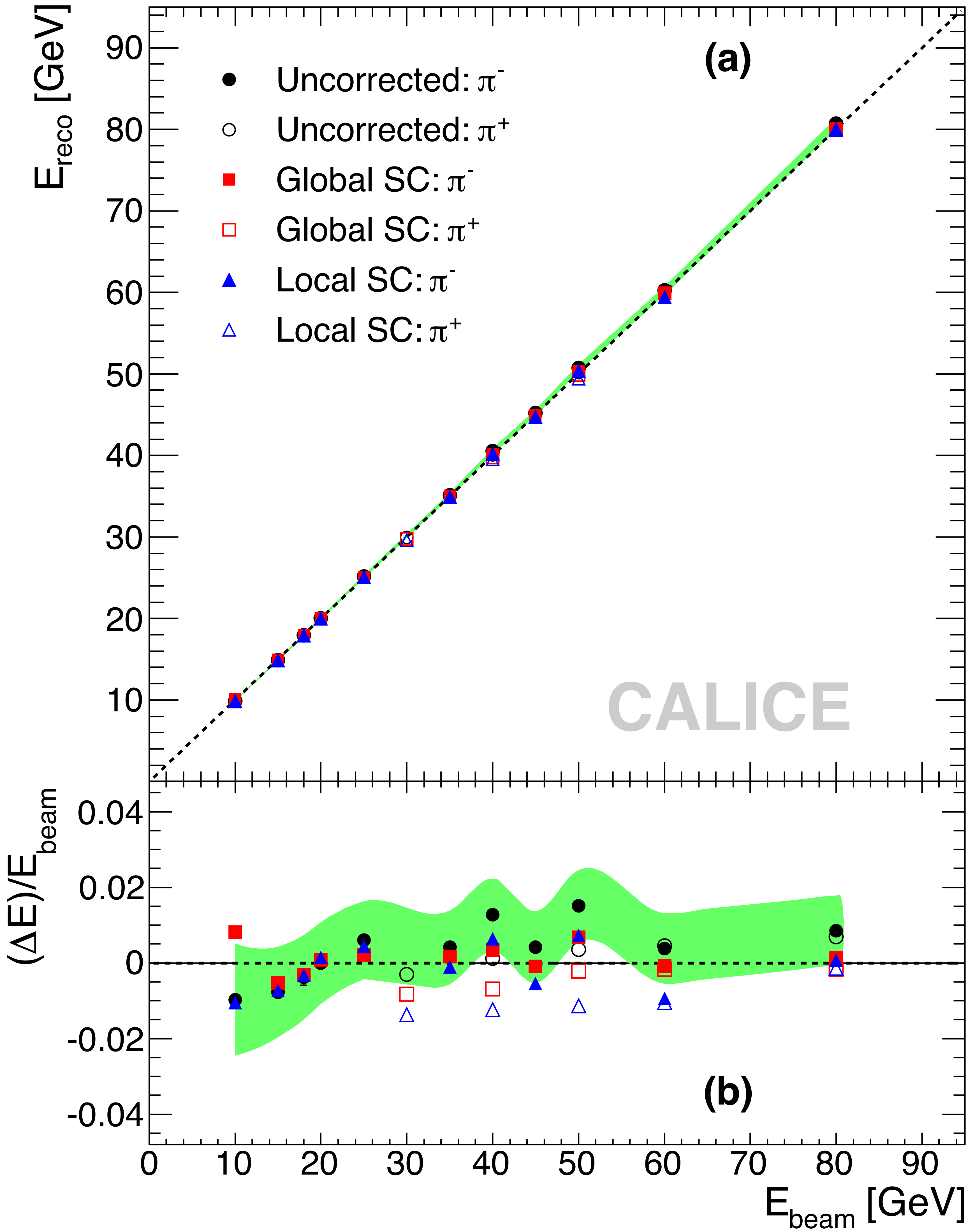


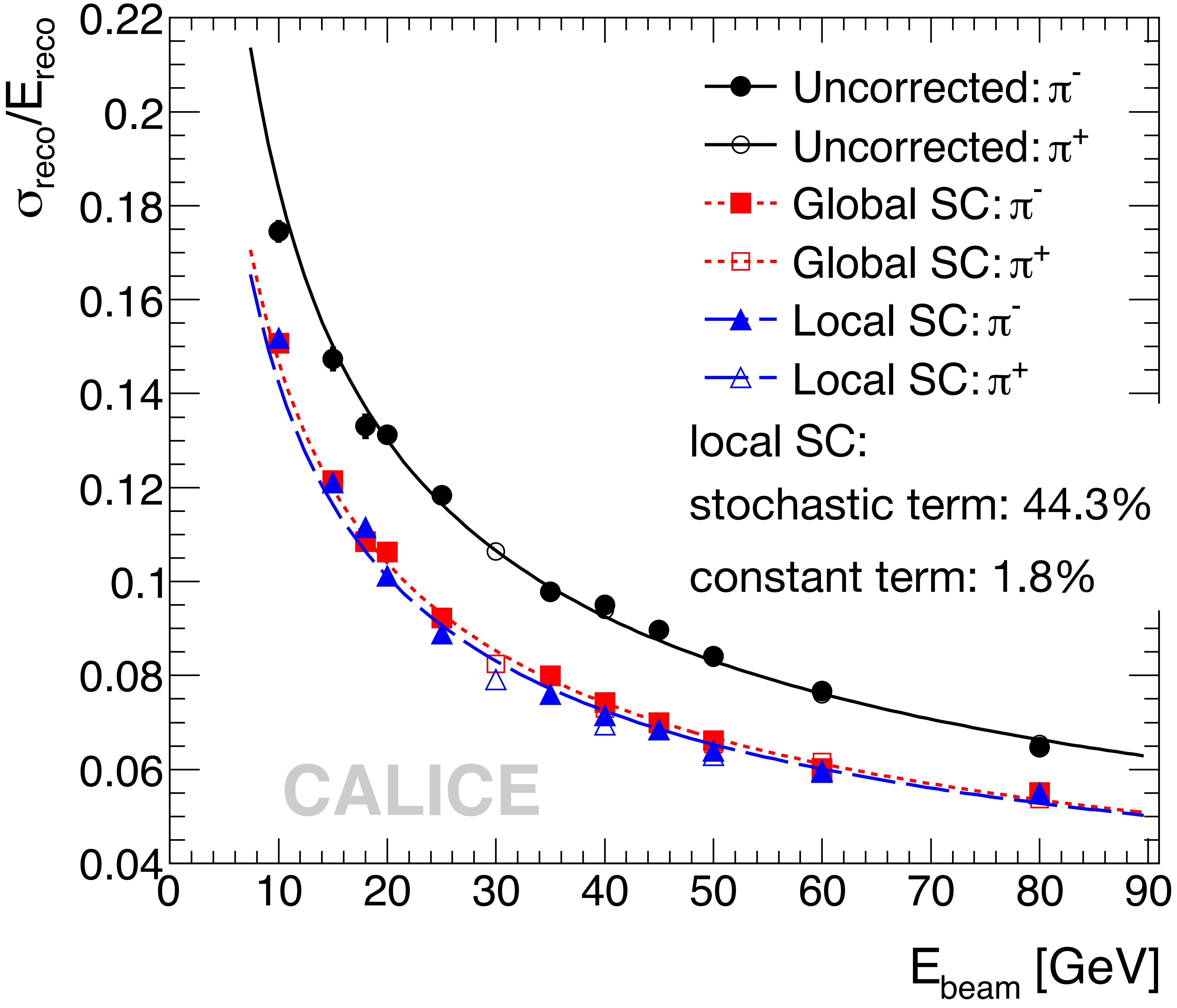

additional parametrisation for energy dependence of the weights, separate weights for each energy-density bin

For both: Parameters / weights determined by $\chi^2$ minimisation of energy resolution

# Software Compensation in the AHCAL

*Comparing Local and Global SC*



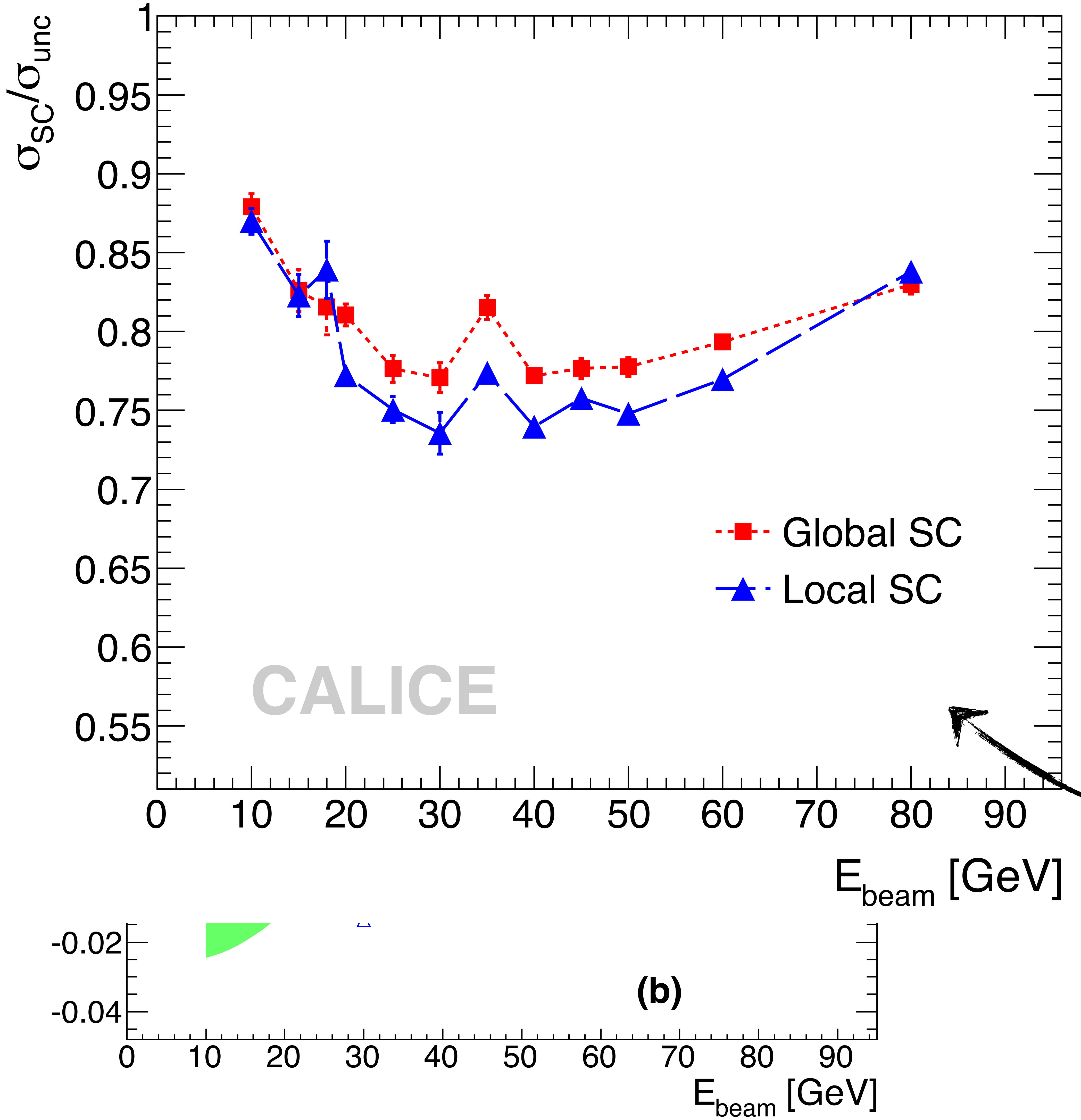



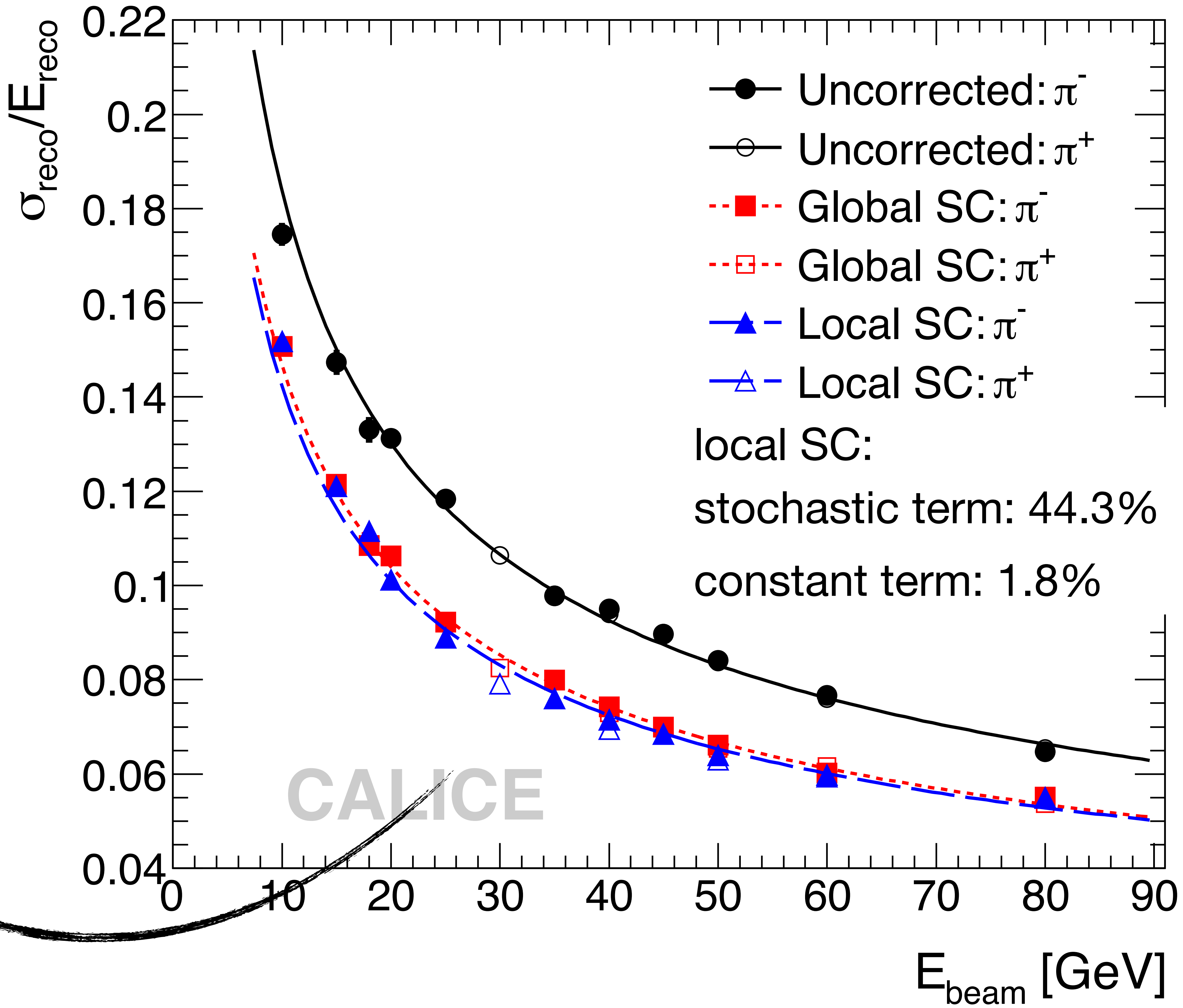

JINST 7 P09017 (2012)

- Substantial improvement of energy resolution with SC

# Software Compensation in the AHCAL

*Comparing Local and Global SC*




JINST 7 P09017 (2012)

- Substantial improvement of energy resolution with SC

  Local SC slightly better improvement - in excess of 25% in intermediate energy range

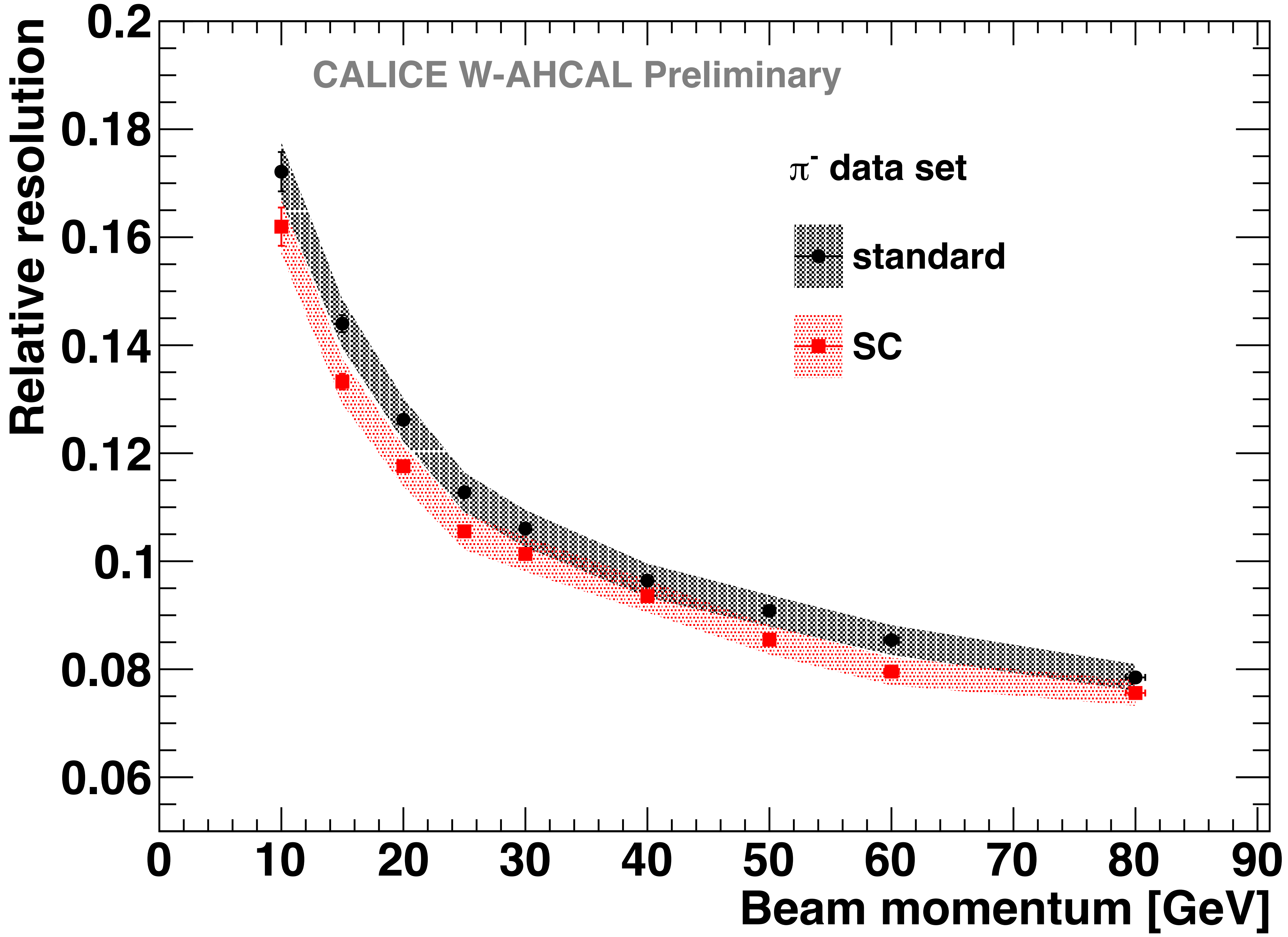
# Software Compensation in the W-AHCAL

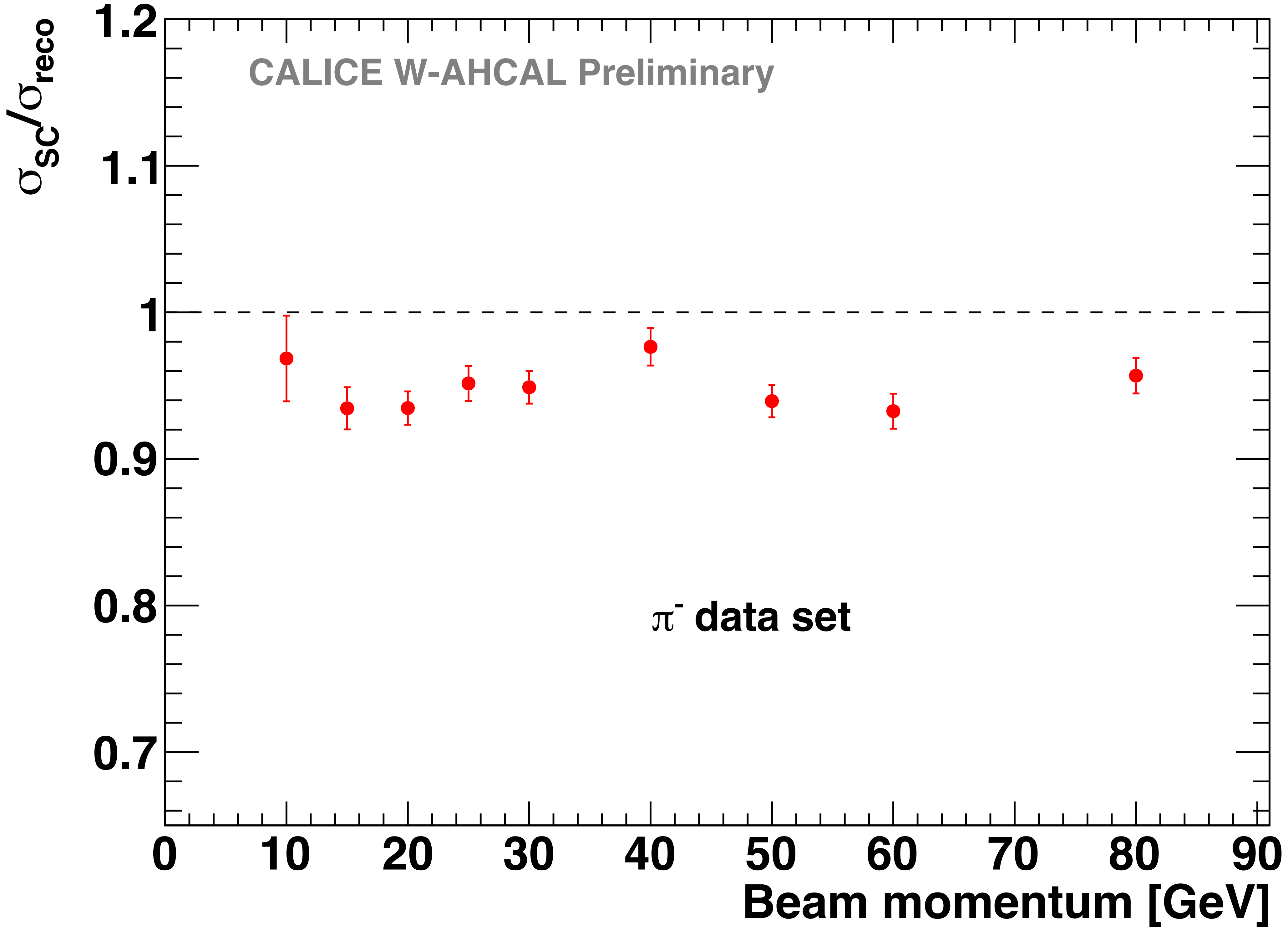
*Global Software Compensation*

- The CALICE W-AHCAL is close to compensating - leaves little handle for software compensation techniques

➡ Tested with global SC (local SC in progress)

| | Fe-AHCAL | W-AHCAL | ratio (Fe/W) |
|---|---|---|---|
| Number of layers | 38 | 38* | 1 |
| Total depth $[\lambda_I]$ | 5.2 | 4.9 | $\sim$1.06 |
| Layer depth [mm] | 31.73 | 24.73 | $\sim$1.28 |
| Layer depth $[\lambda_I]$ | 0.137 | 0.129 | $\sim$1.06 |
| Layer depth $[X_0]$ | 1.24 | 2.80 | $\sim$0.44 |

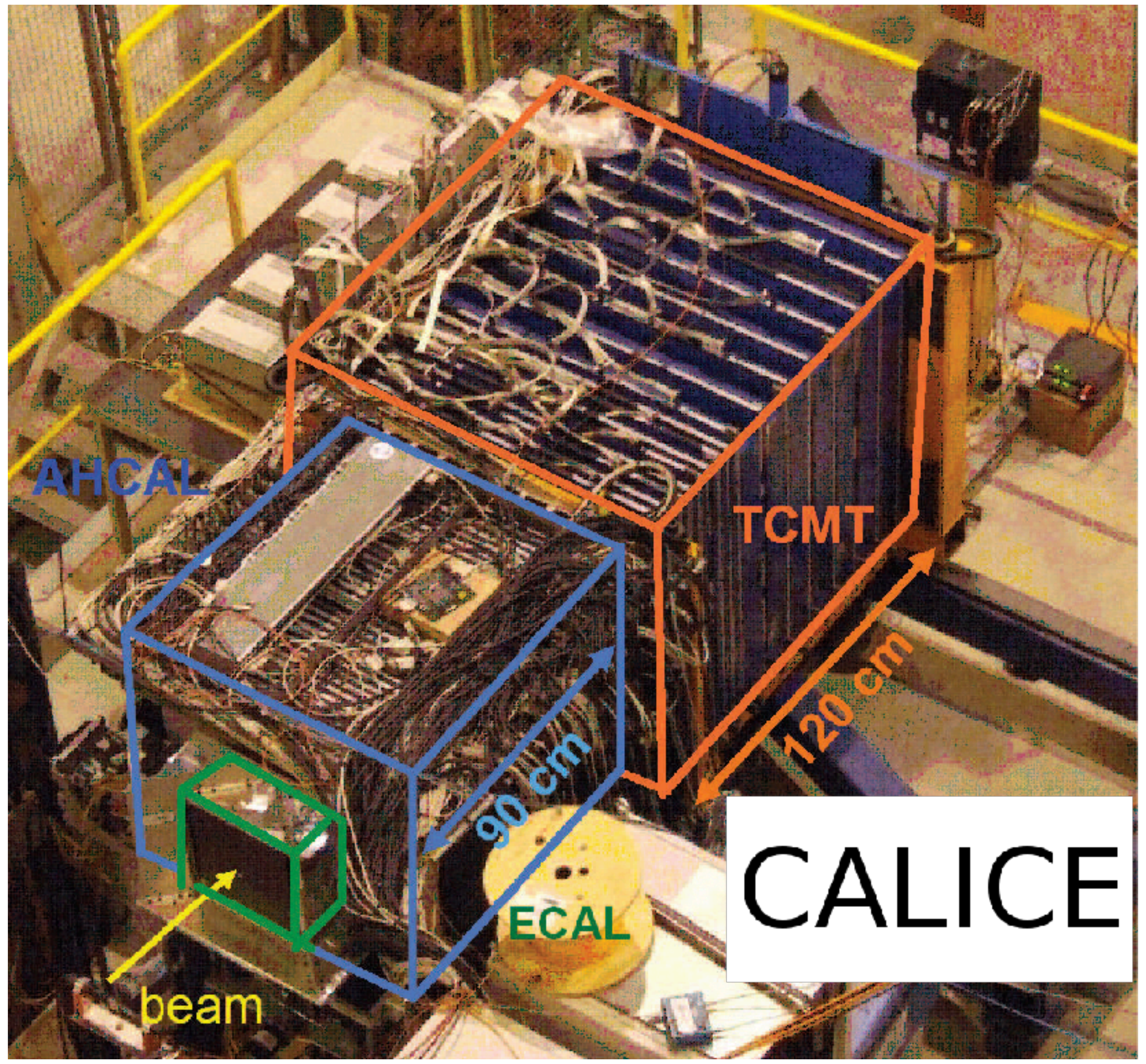



CAN-062

# Extension to Combined ECAL/HCAL Systems

*One Example: SiW ECAL + Scintillator / Fe HCAL*




- Studying energy resolution in a “real-world” setting: A combined system of SiW ECAL, Scintillator/FE HCAL, Tail Catcher
  - A combination of non-compensating systems with different active and absorber materials and varying longitudinal sampling
- Local software compensation extended by subsystem-dependen binning and weight parameter

*ECAL (30 layers):*

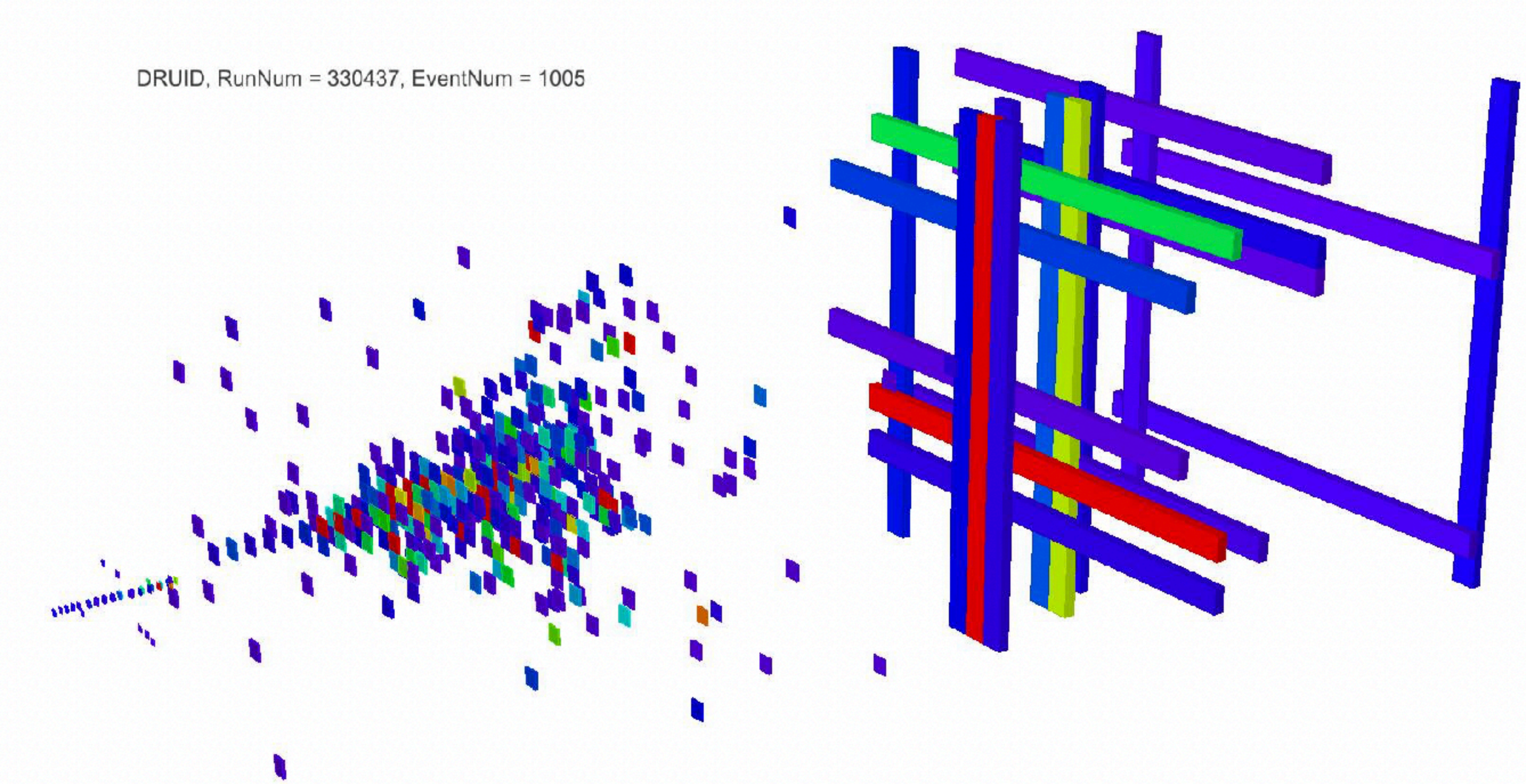
Absorber: W; 1.4 mm, 2.8 mm, 4.2 mm

Active: Si; 525 μm

*HCAL (38 layers) / TCMT (8+8 layers):*

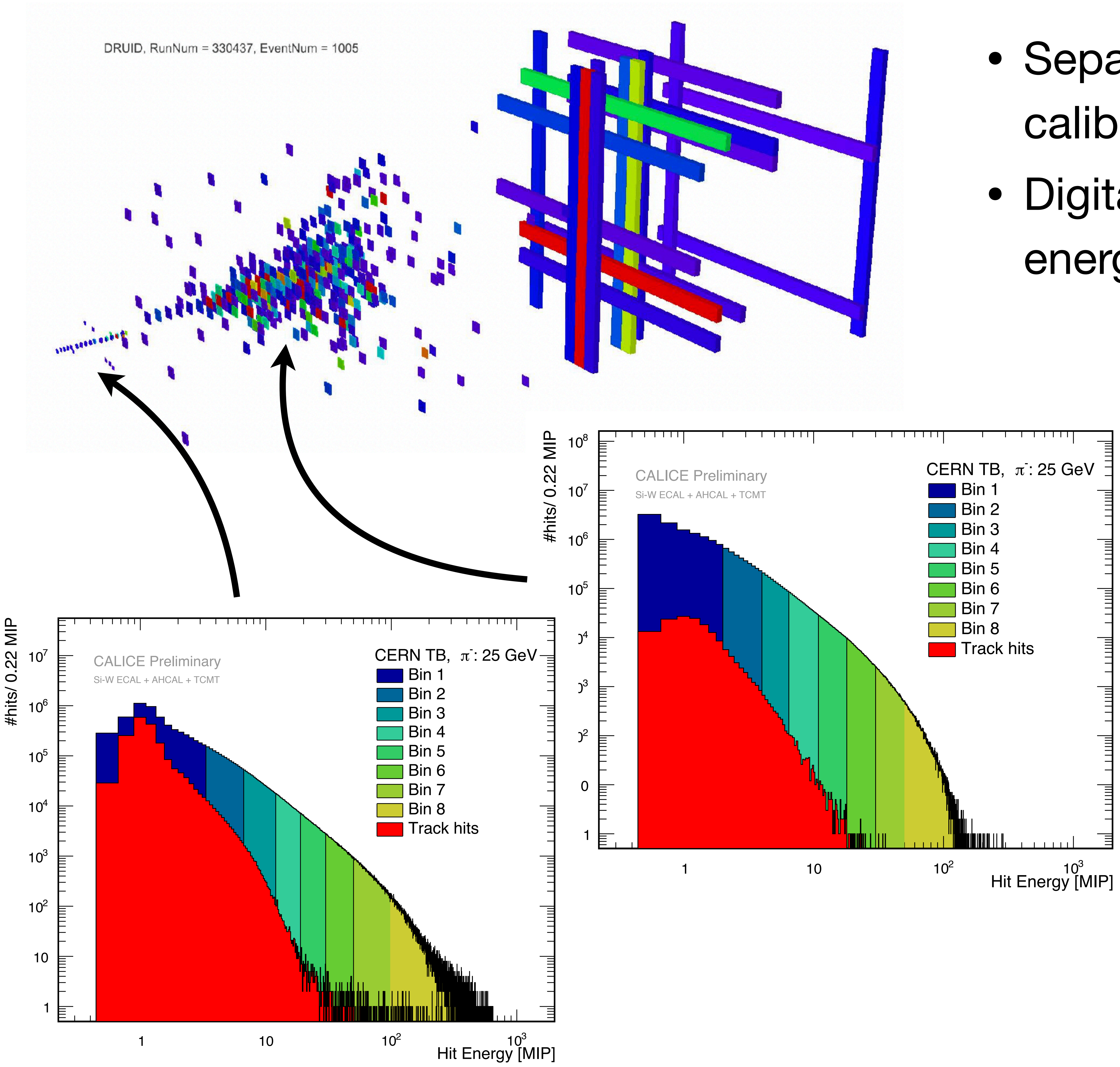
Absorber: Steel; ~ 21 mm (including cassettes)

Active: Plastic scintillator; 5 mm



# Local Software Compensation in Combined Systems

*The Implementation*

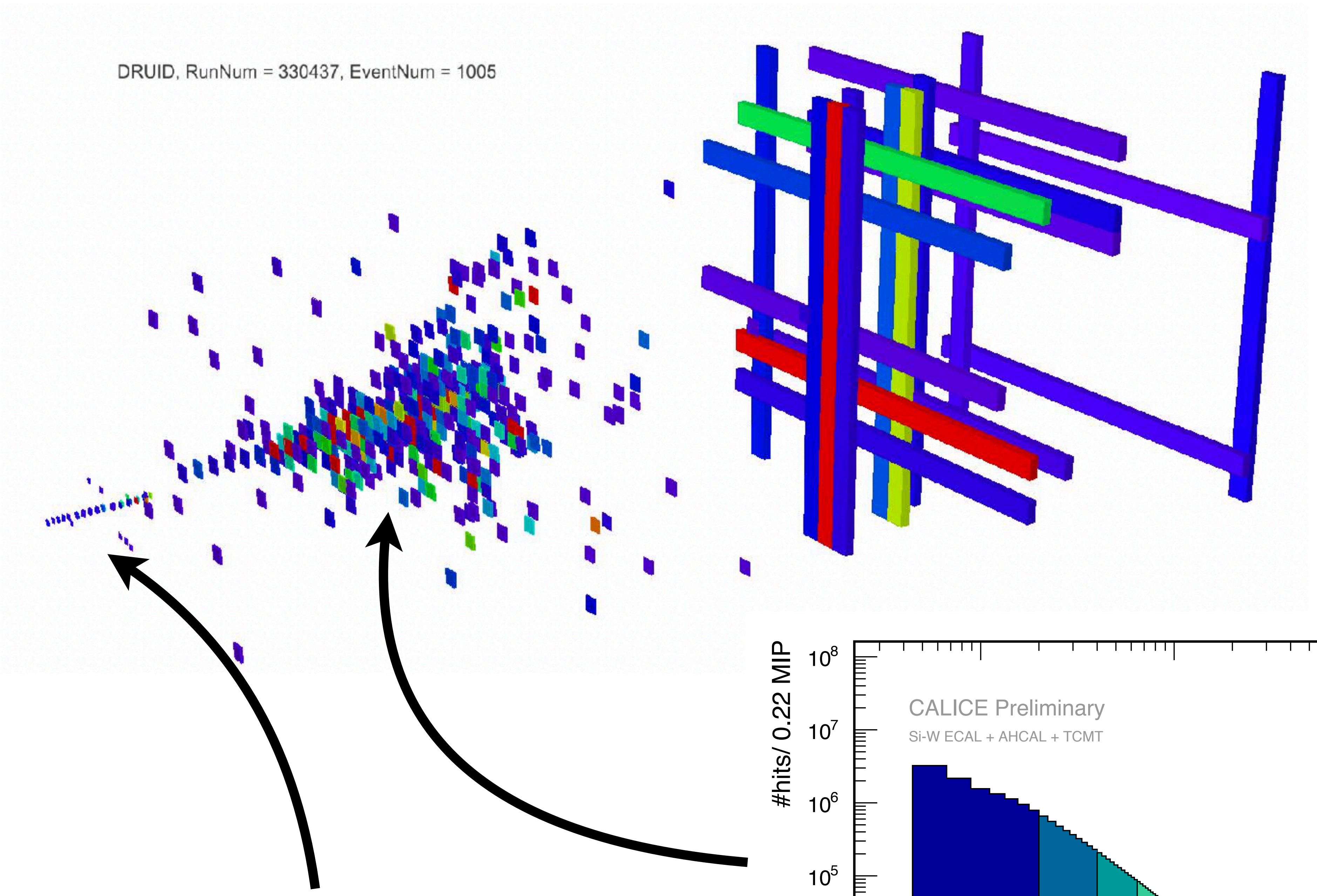



- Separate treatment of incoming track up to first interaction: calibration factor different than that for showers
- Digital weighting for first two bins: Slight advantage for energy resolution due to suppression of Landau fluctuations

CAN-058

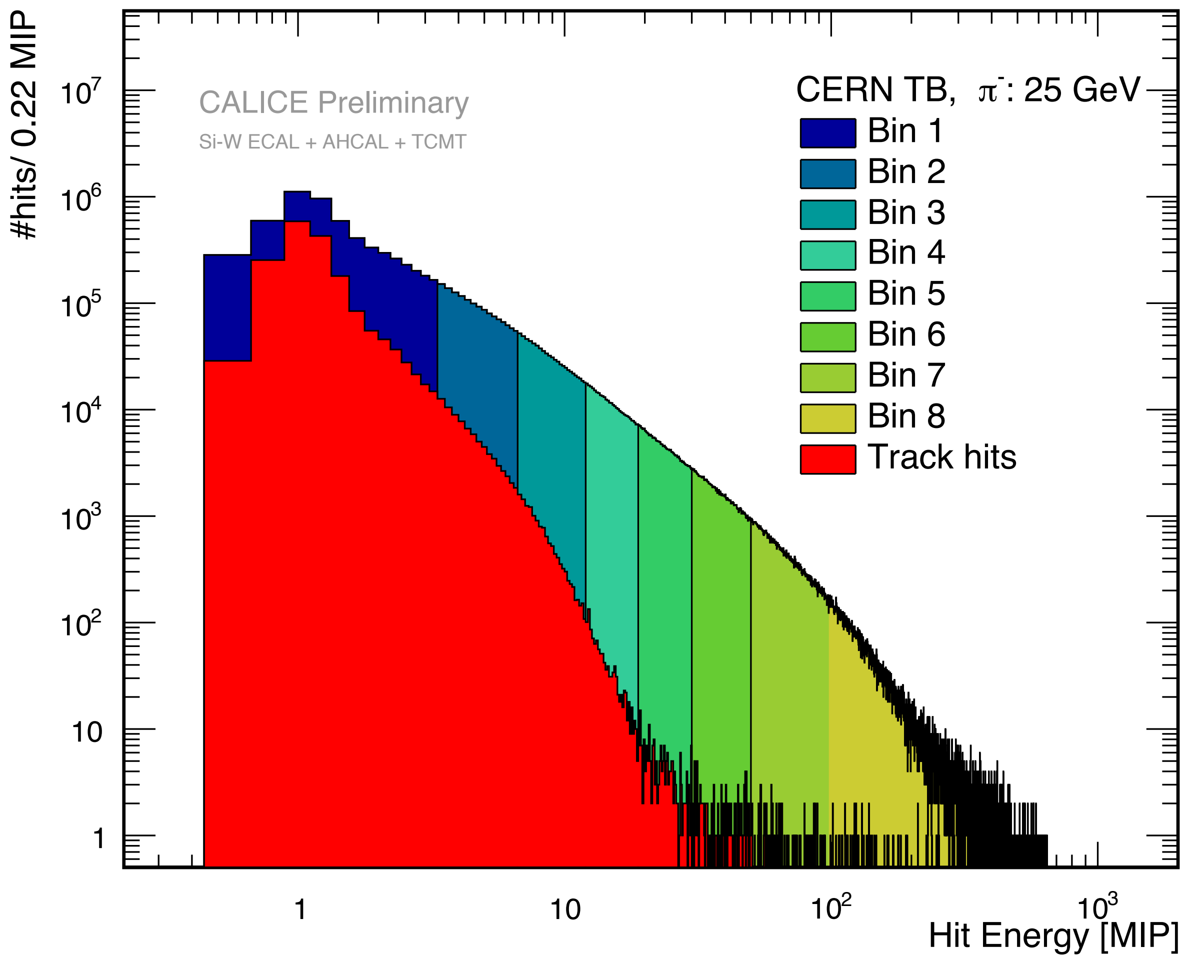
# Local Software Compensation in Combined Systems

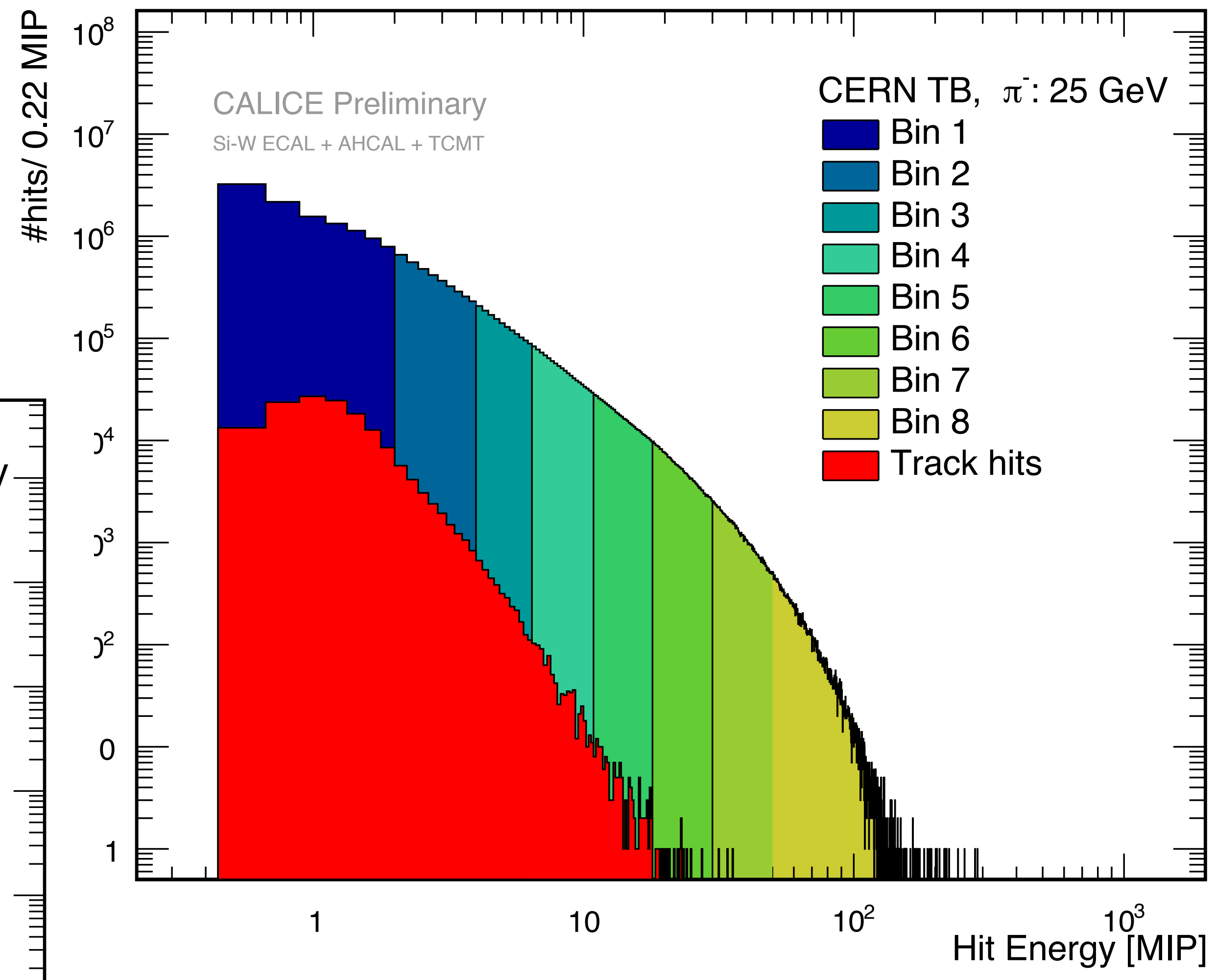
*The Implementation*



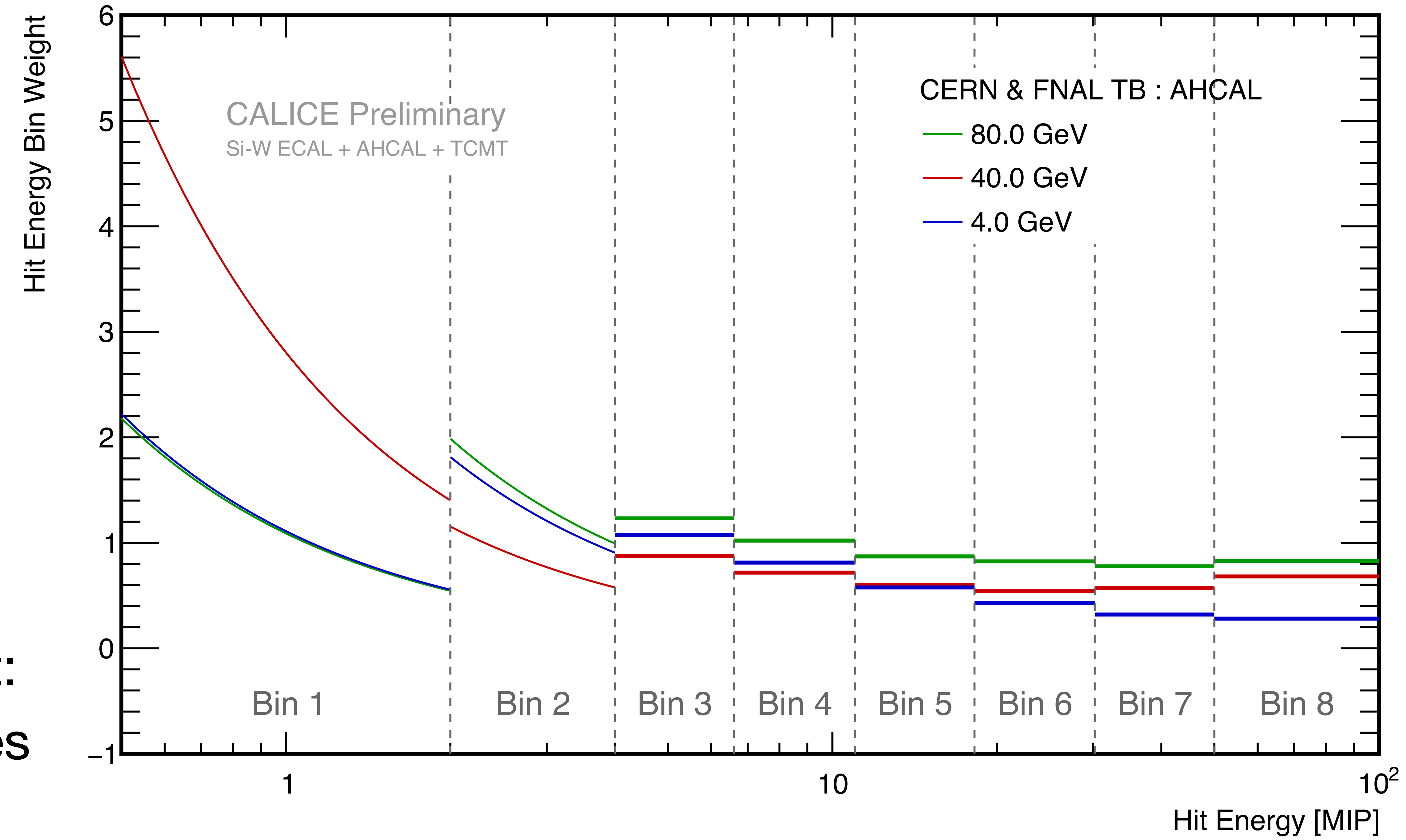

- Separate treatment of incoming track up to first interaction: calibration factor different than that for showers
- Digital weighting for first two bins: Slight advantage for energy resolution due to suppression of Landau fluctuations

CAN-058



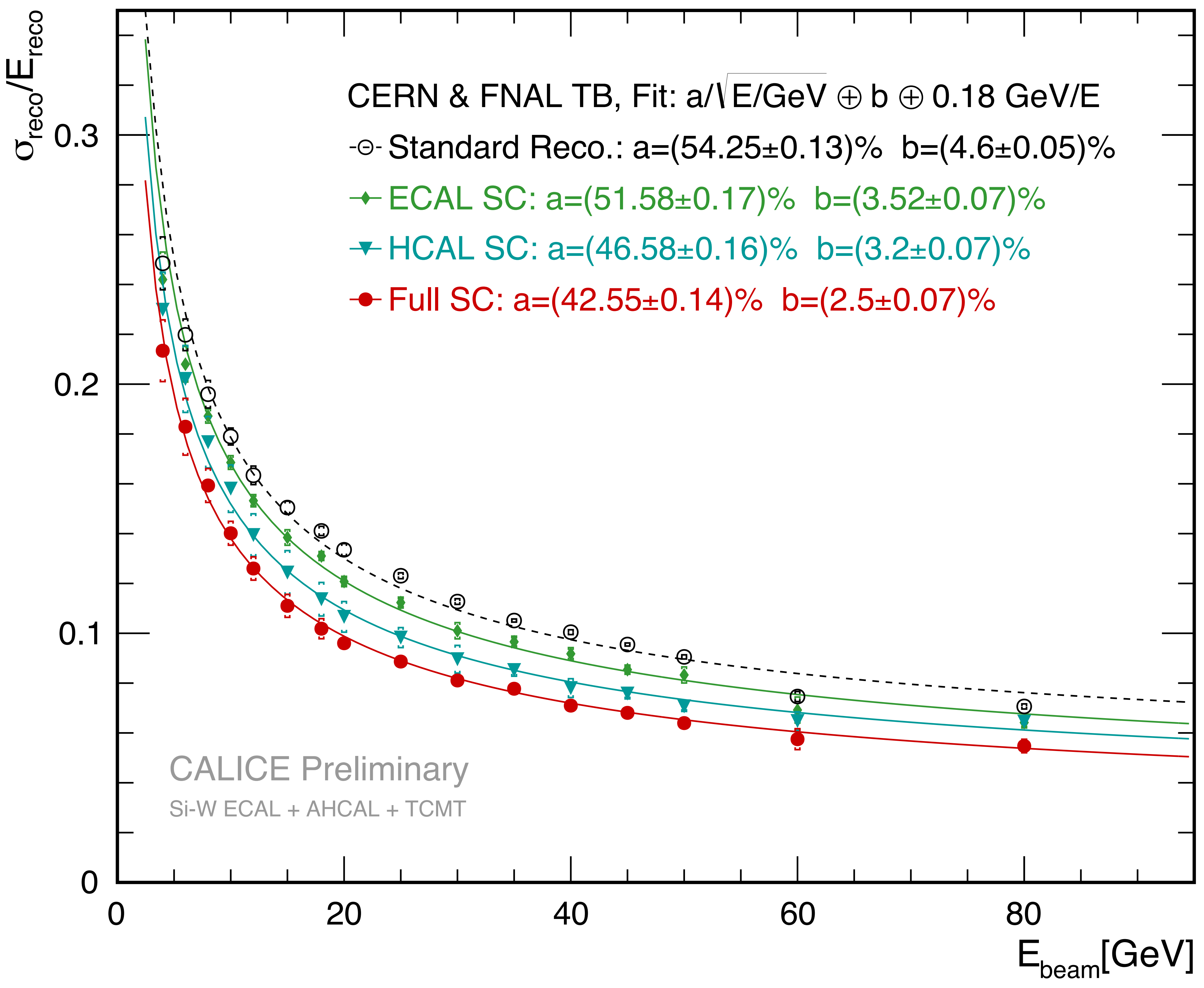

weights are energy dependent: overall shower density changes with energy!



# Local Software Compensation in Combined Systems

*Resulting Performance*

CAN-058



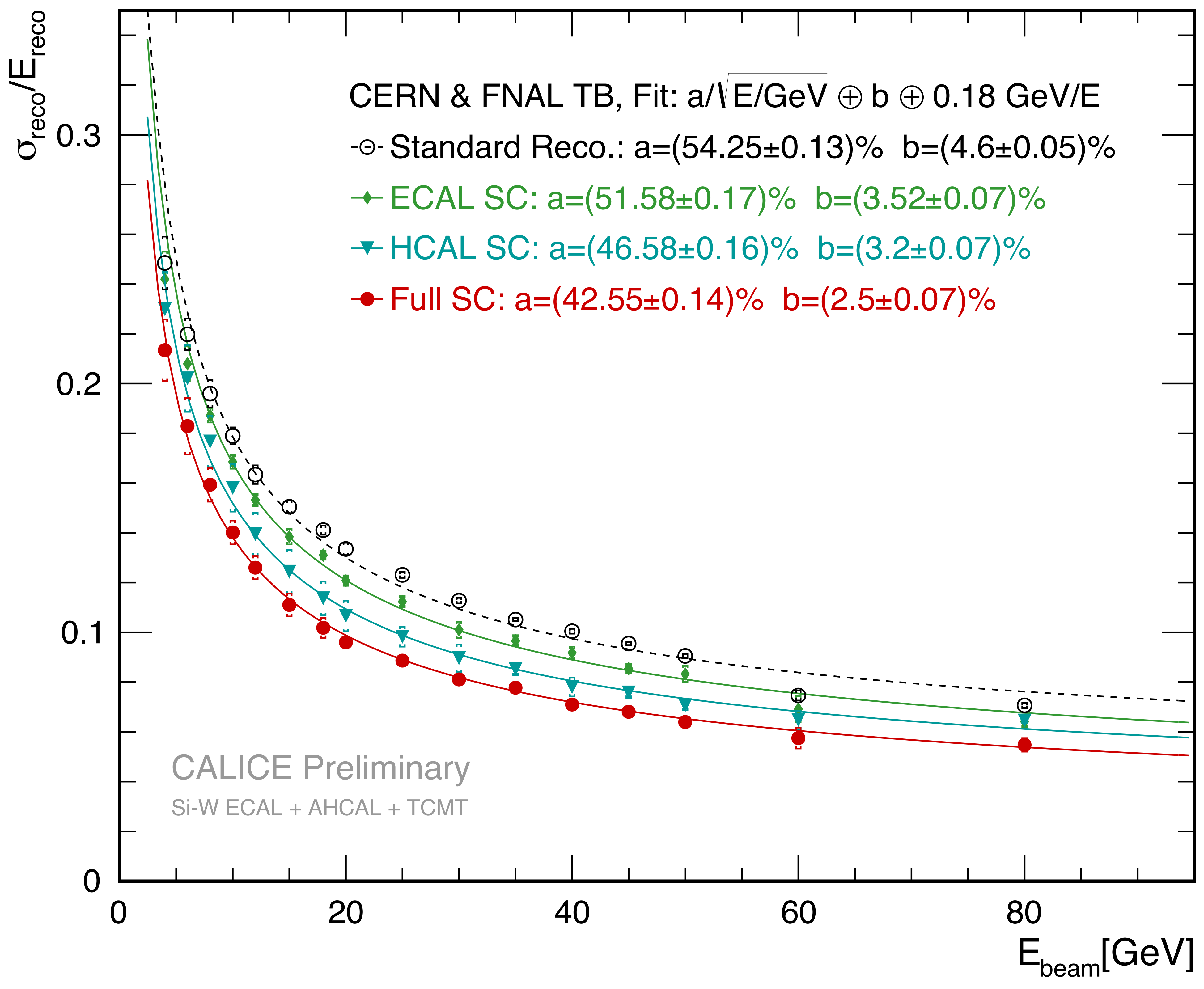

- Substantial improvement in energy resolution:
  - SC in ECAL alone up to 8% improvement
  - SC in HCAL alone up to 23% improvement
  - Full SC up to 30% improvement, for a stochastic term of 42.5% and a constant term of 2.5%

➯ The bulk of the improvement is achieved in the AHCAL

# Local Software Compensation in Combined Systems

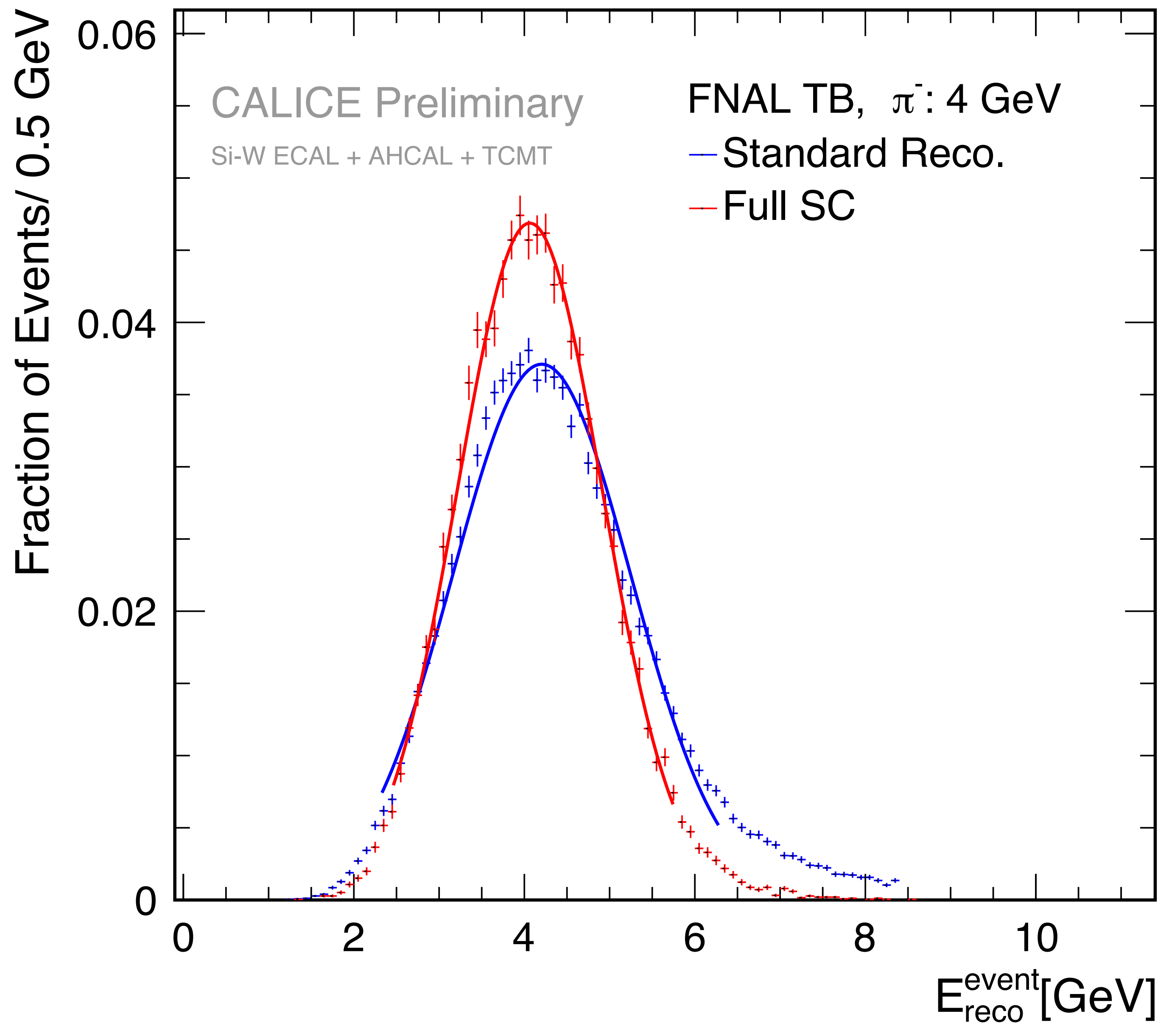
*Resulting Performance*

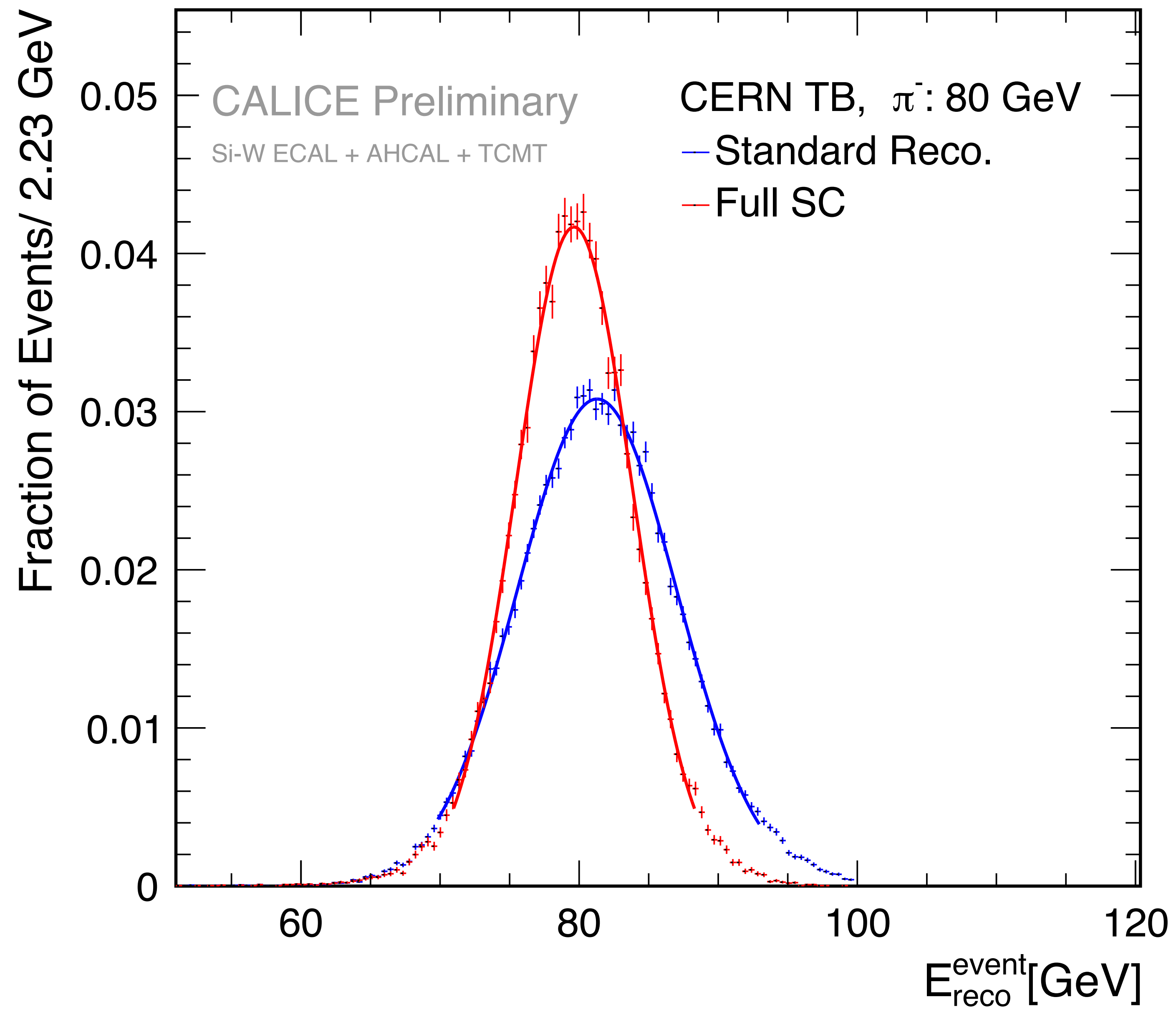
CAN-058




- Substantial improvement in energy resolution:
  - SC in ECAL alone up to 8% improvement
  - SC in HCAL alone up to 23% improvement
  - Full SC up to 30% improvement, for a stochastic term of 42.5% and a constant term of 2.5%

➡ The bulk of the improvement is achieved in the AHCAL



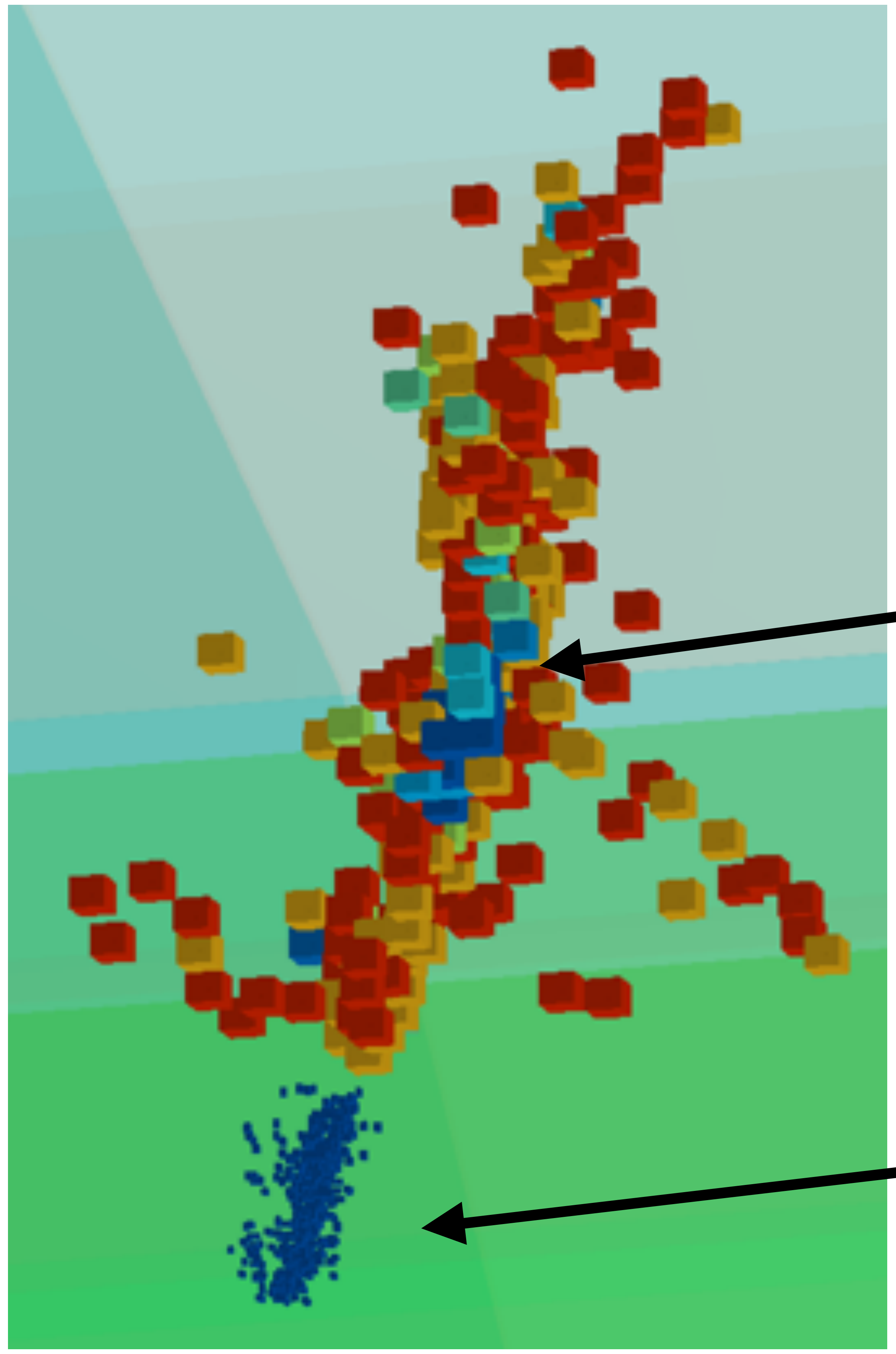




- Software compensation also reduces tails and asymmetries in the energy distribution, in particular at lower beam energies

# Combining Software Compensation with Particle Flow

*Local Software Compensation in PandoraPFA*

- Particle flow algorithms make use of calorimeter energy at two main points
  - Track - calorimeter cluster matching, and iterative reclustering
  - Energy of neutral particles



transfer software compensation algorithm and training strategies from CALICE to full ILD detector simulations

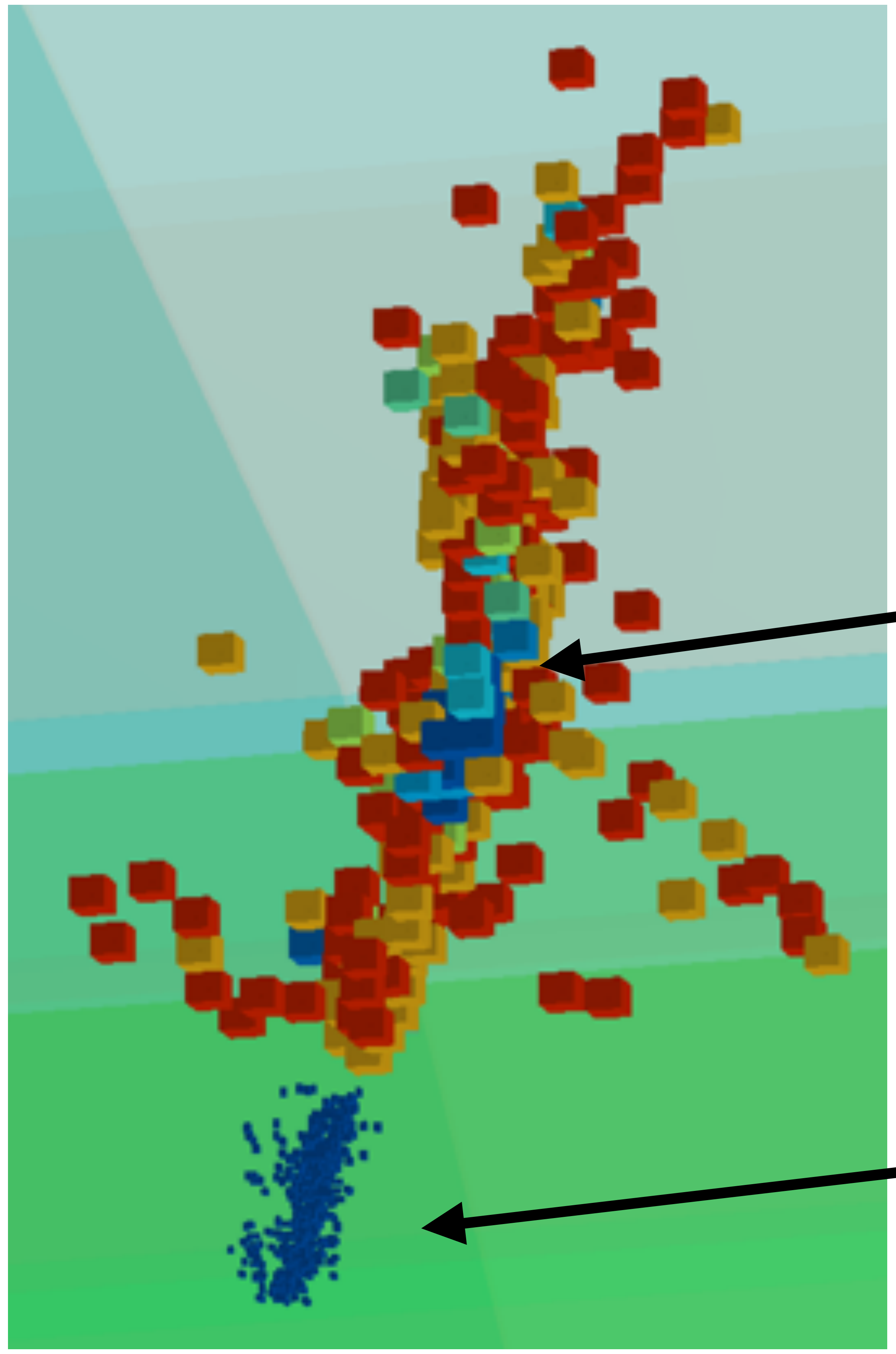
em sub showers (in shower core) weighted less than hadronic periphery

ECAL not yet included: standard reconstruction used

# Combining Software Compensation with Particle Flow

*Local Software Compensation in PandoraPFA*

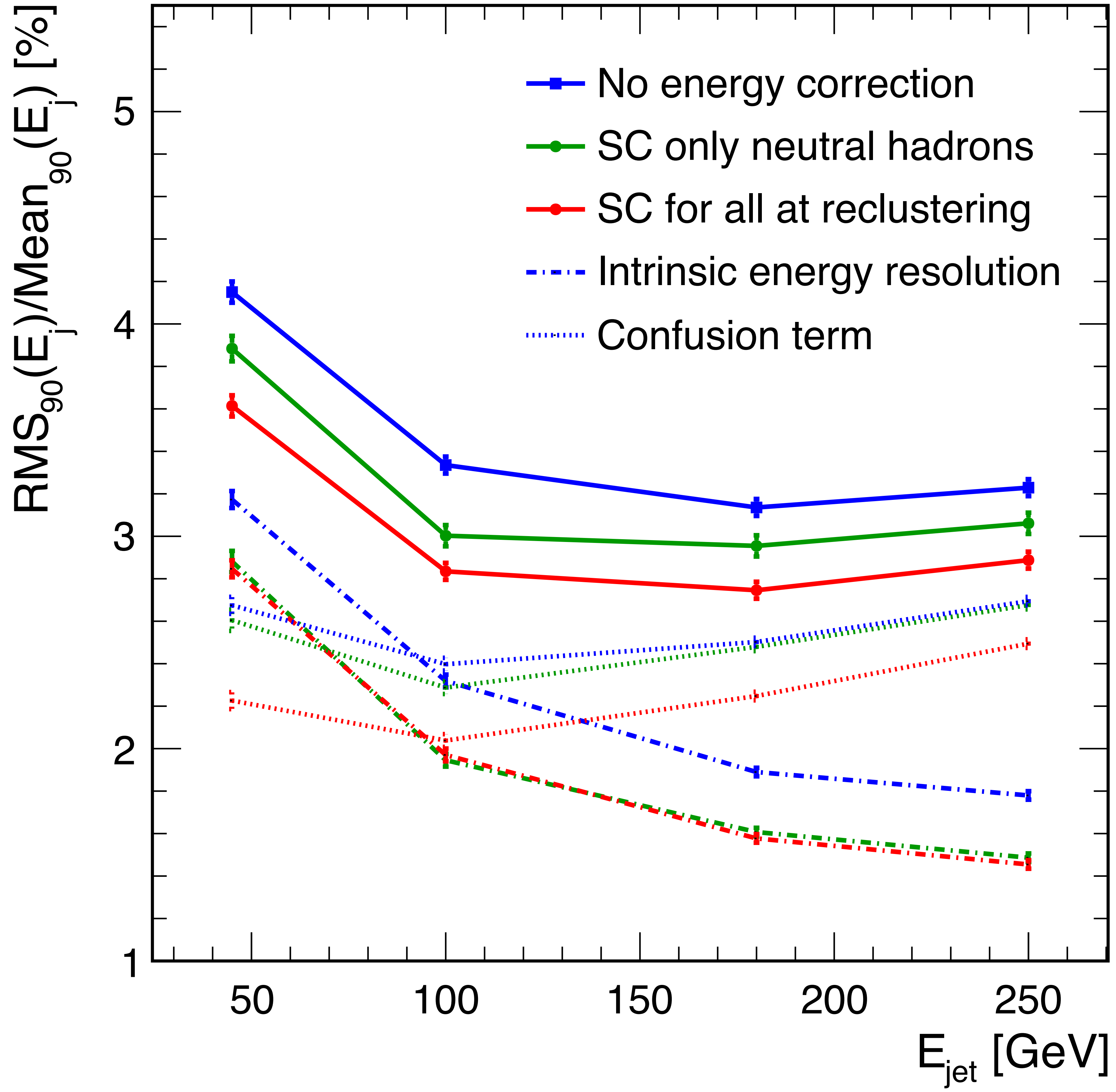
- Particle flow algorithms make use of calorimeter energy at two main points
  - Track - calorimeter cluster matching, and iterative reclustering
  - Energy of neutral particles



transfer software compensation algorithm and training strategies from CALICE to full ILD detector simulations

em sub showers (in shower core) weighted less than hadronic periphery

ECAL not yet included: standard reconstruction used

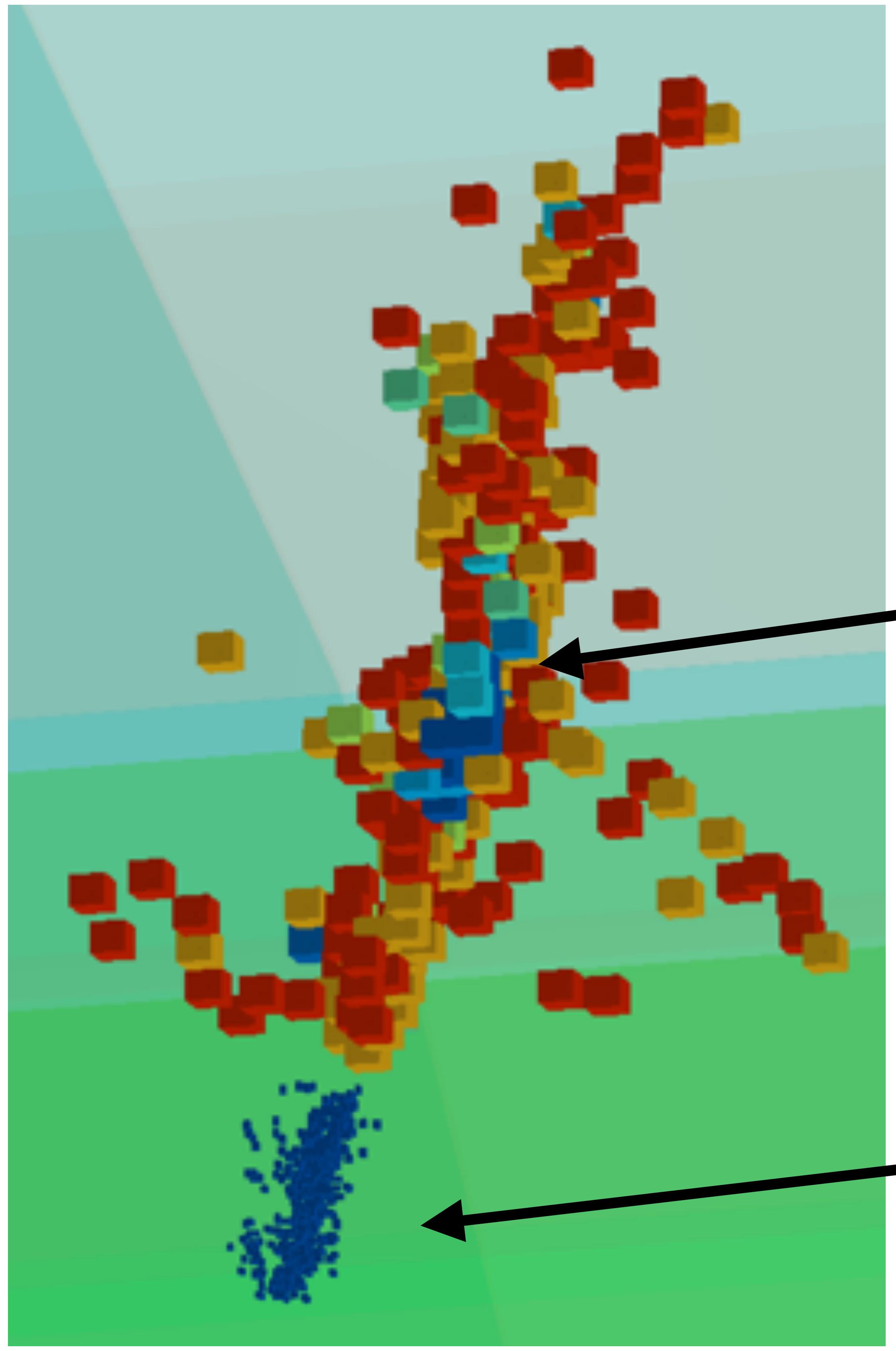



EPJ C77, 698 (2017)

# Combining Software Compensation with Particle Flow

*Local Software Compensation in PandoraPFA*

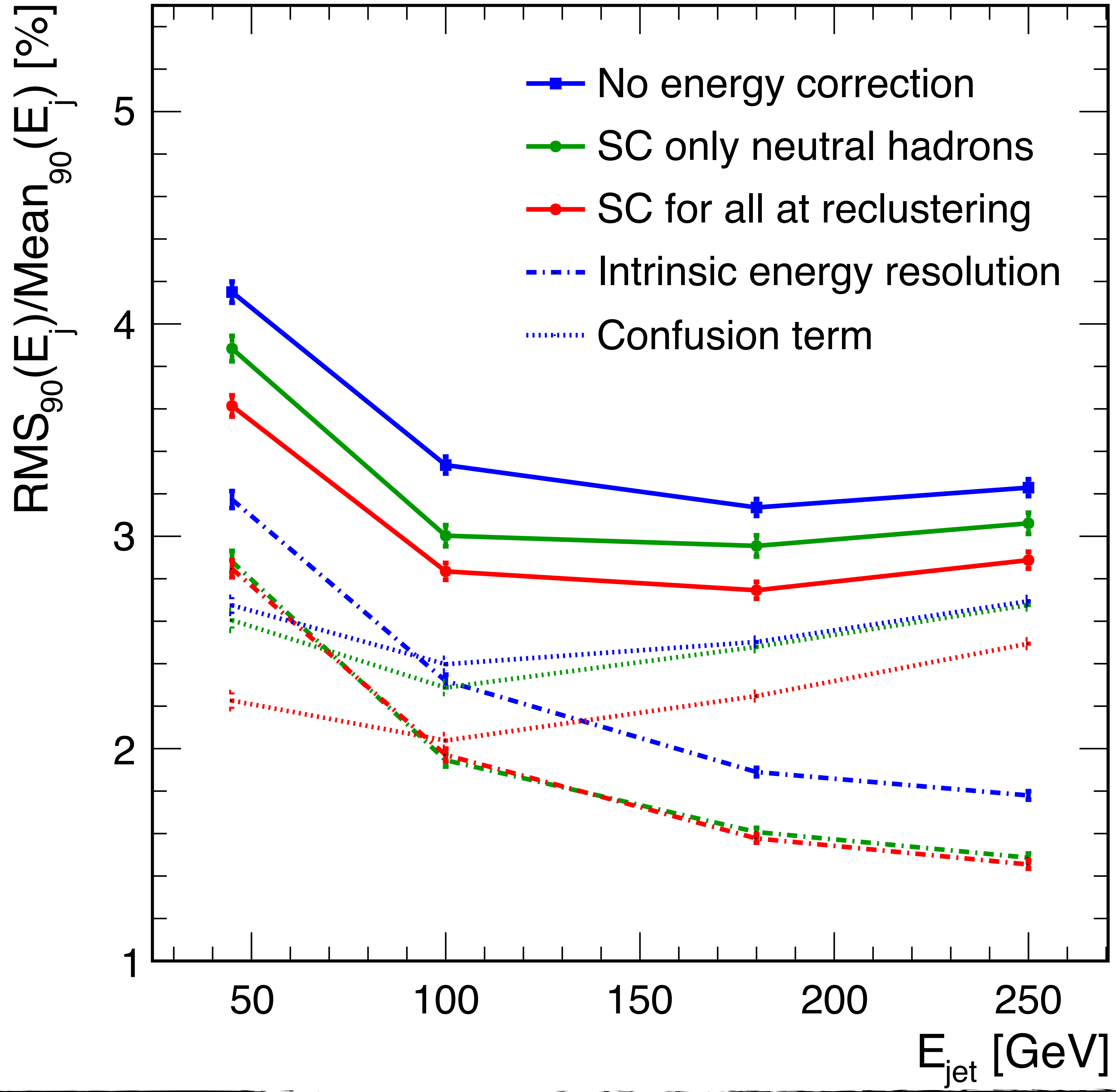
- Particle flow algorithms make use of calorimeter energy at two main points
  - Track - calorimeter cluster matching, and iterative reclustering
  - Energy of neutral particles



transfer software compensation algorithm and training strategies from CALICE to full ILD detector simulations

em sub showers (in shower core) weighted less than hadronic periphery

ECAL not yet included: standard reconstruction used




EPJ C77, 698 (2017)

Now the standard energy reconstruction method for CLIC and ILD physics and performance studies

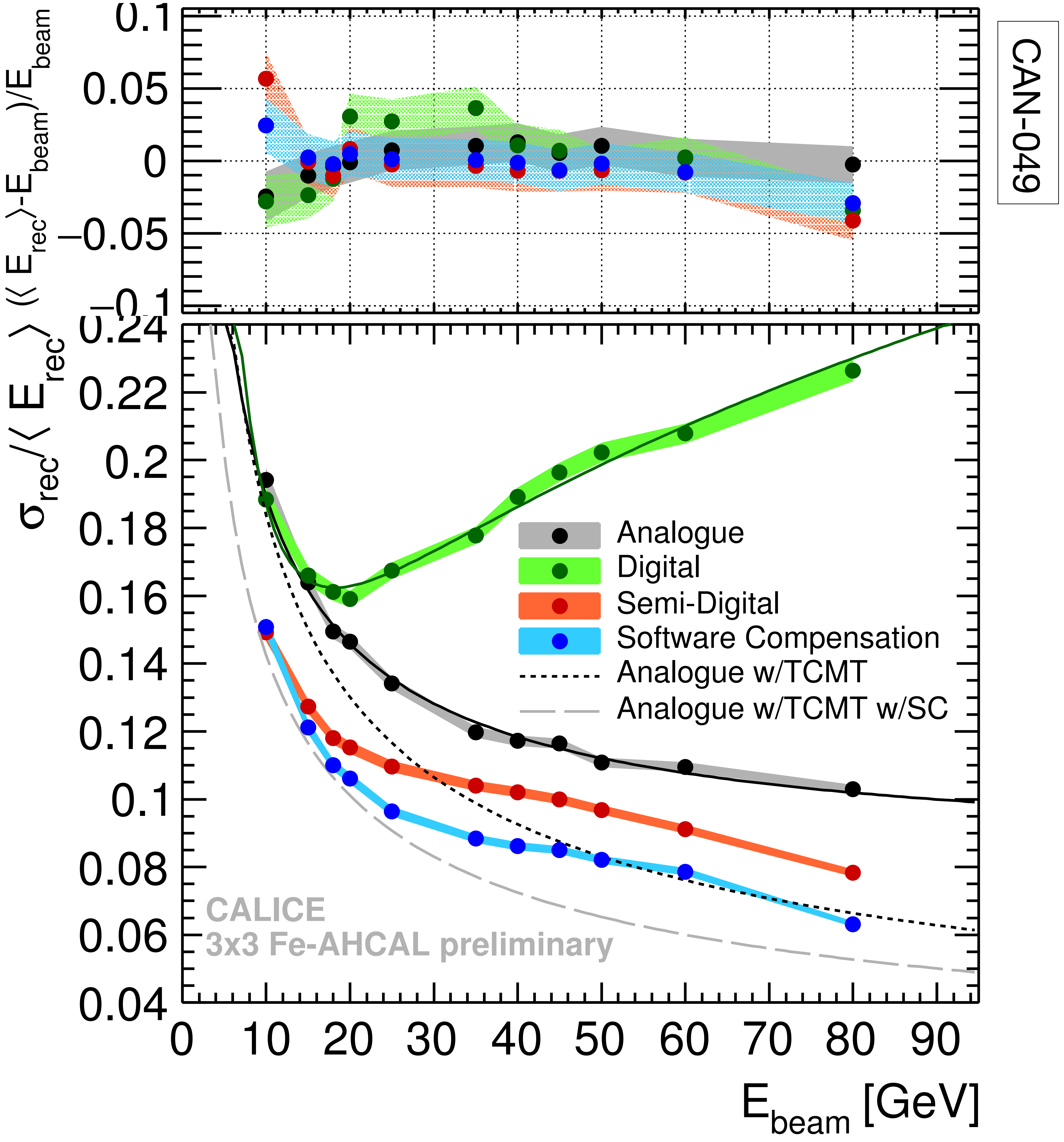
# Energy Reconstruction & Readout Schemes

*Understanding Resolution Impact of Granularity & Readout Technology*

- CALICE hadron calorimeters use different schemes for energy reconstruction - depending on readout technology:
  - *scintillator*: analog & software compensation
  - *gas*: digital (1 bit), semi-digital (2 bit)

N.B.: Semi-digital reconstruction and software compensation are related: both use optimised hit or energy dependent weighting factors

- Different schemes tested on AHCAL data (3 x 3 cm² granularity)



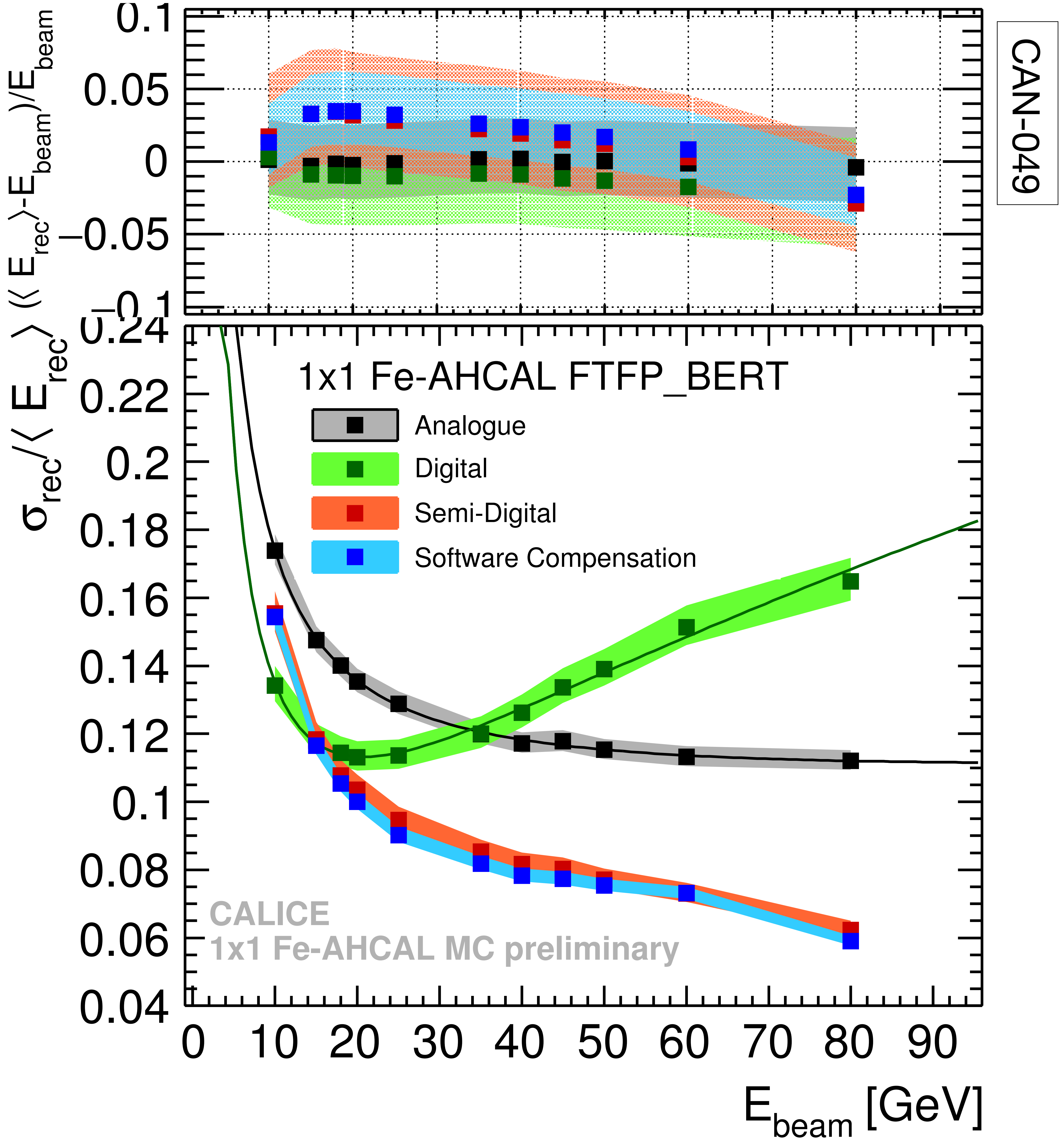
# Energy Reconstruction & Readout Schemes

*Understanding Resolution Impact of Granularity & Readout Technology*

- CALICE hadron calorimeters use different schemes for energy reconstruction - depending on readout technology:
  - *scintillator*: analog & software compensation
  - *gas*: digital (1 bit), semi-digital (2 bit)

N.B.: Semi-digital reconstruction and software compensation are related: both use optimised hit or energy dependent weighting factors

- Different schemes tested on AHCAL data (3 x 3 cm² granularity)
- Simulations used to study 1 x 1 cm² granularity (scintillator)
  - Digital & fine granularity best at low energy: Suppression of fluctuations
  - SC & semi-digital comparable
    NB: Sampling fraction matters: Semi-digital reconstruction in RPCs does not reach the same resolution



# Towards Multivariate Techniques

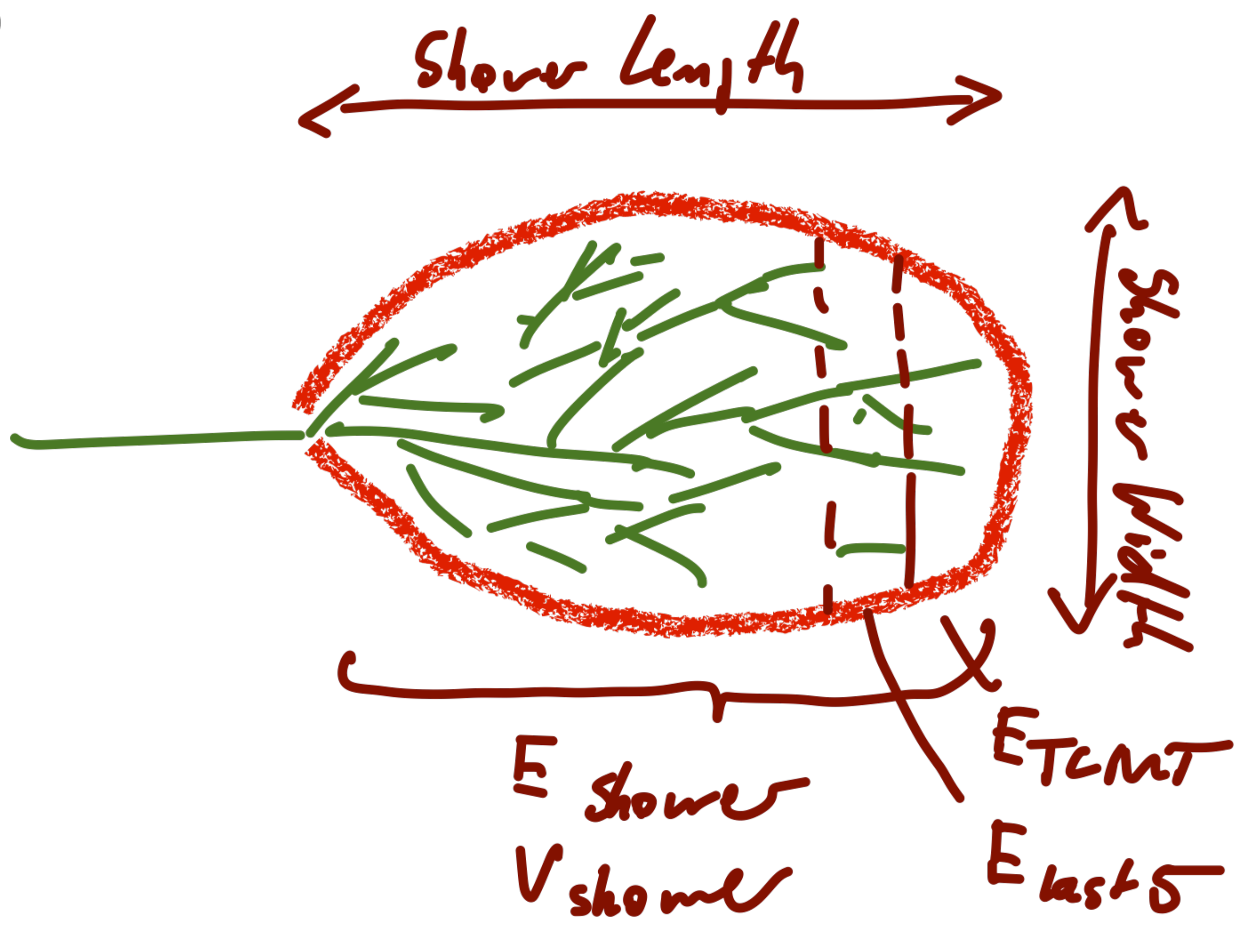
*Initial Studies with Neural Networks*

CAN-021

- Performed with AHCAL physics prototype in 2010 (K. Seidel, FS)

  *The strategy:*

  - Use "simple" clustering to define a set of shower variables
  - Train a neural network on MC data (NB: quasi-continuous energy distribution to avoid bias)
  - Apply NN to data (requires additional energy correction to account for differences between data and MC)




Shower variables used in NN:
Global observables

# Towards Multivariate Techniques

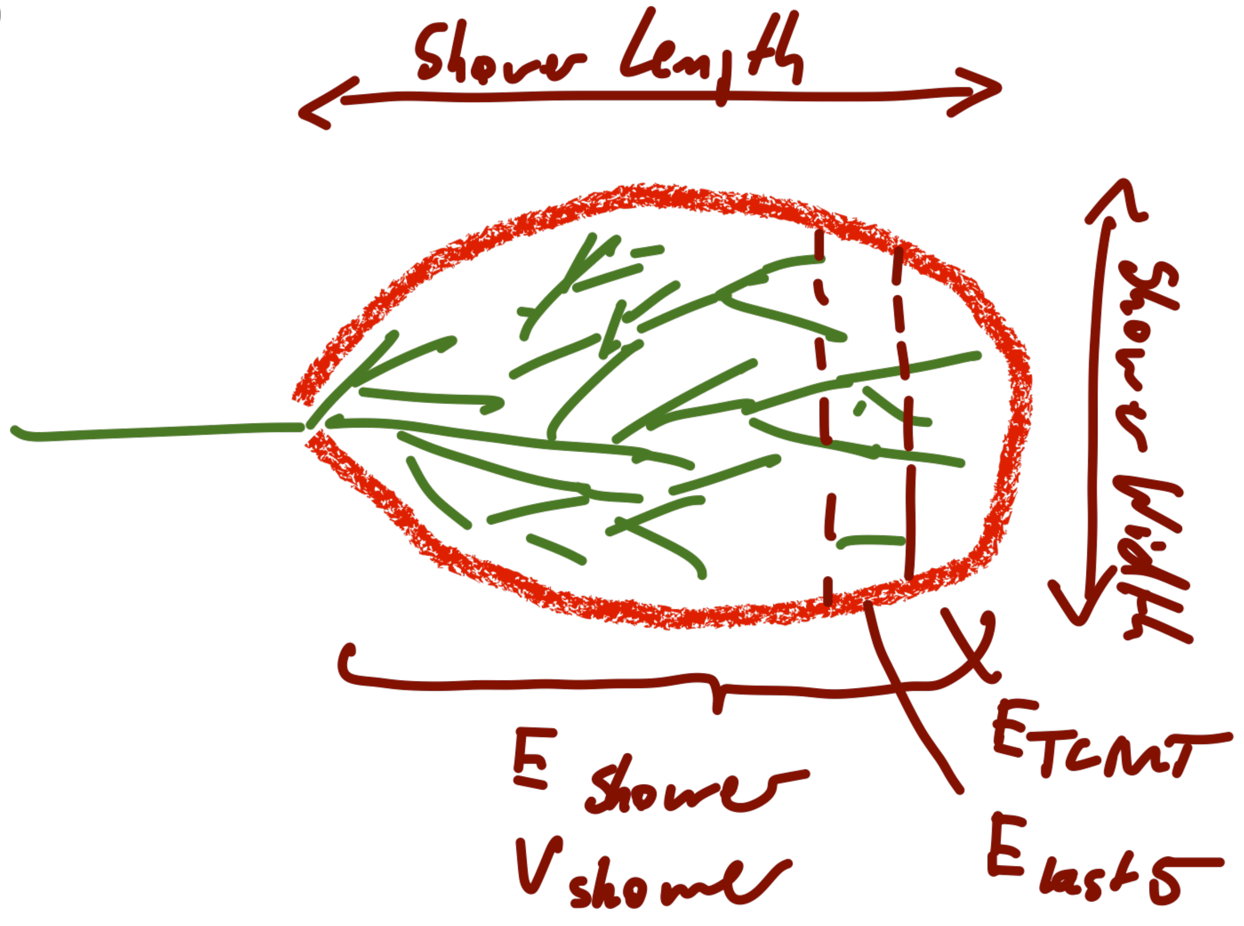
*Initial Studies with Neural Networks*

CAN-021

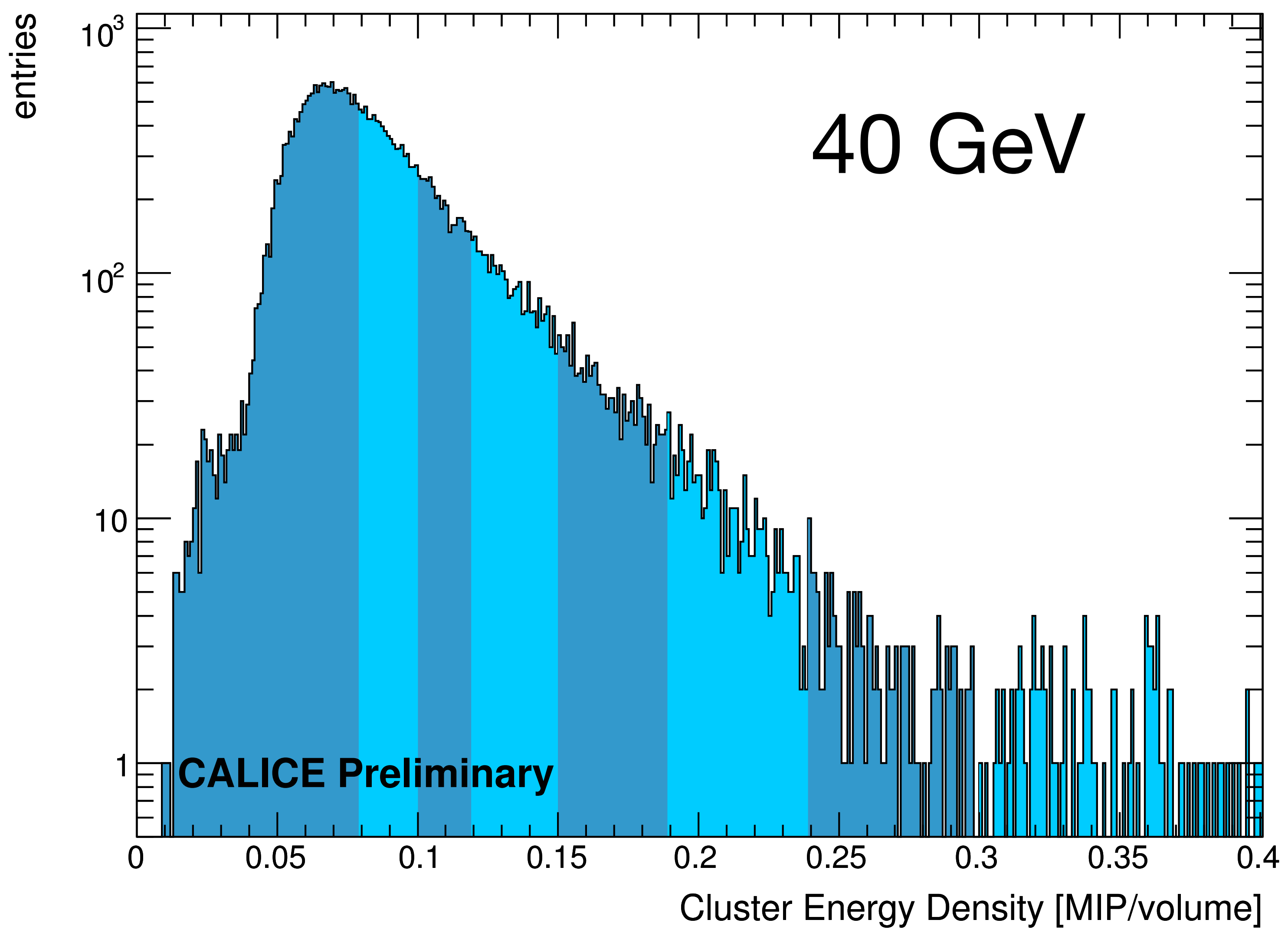
- Performed with AHCAL physics prototype in 2010 (K. Seidel, FS)

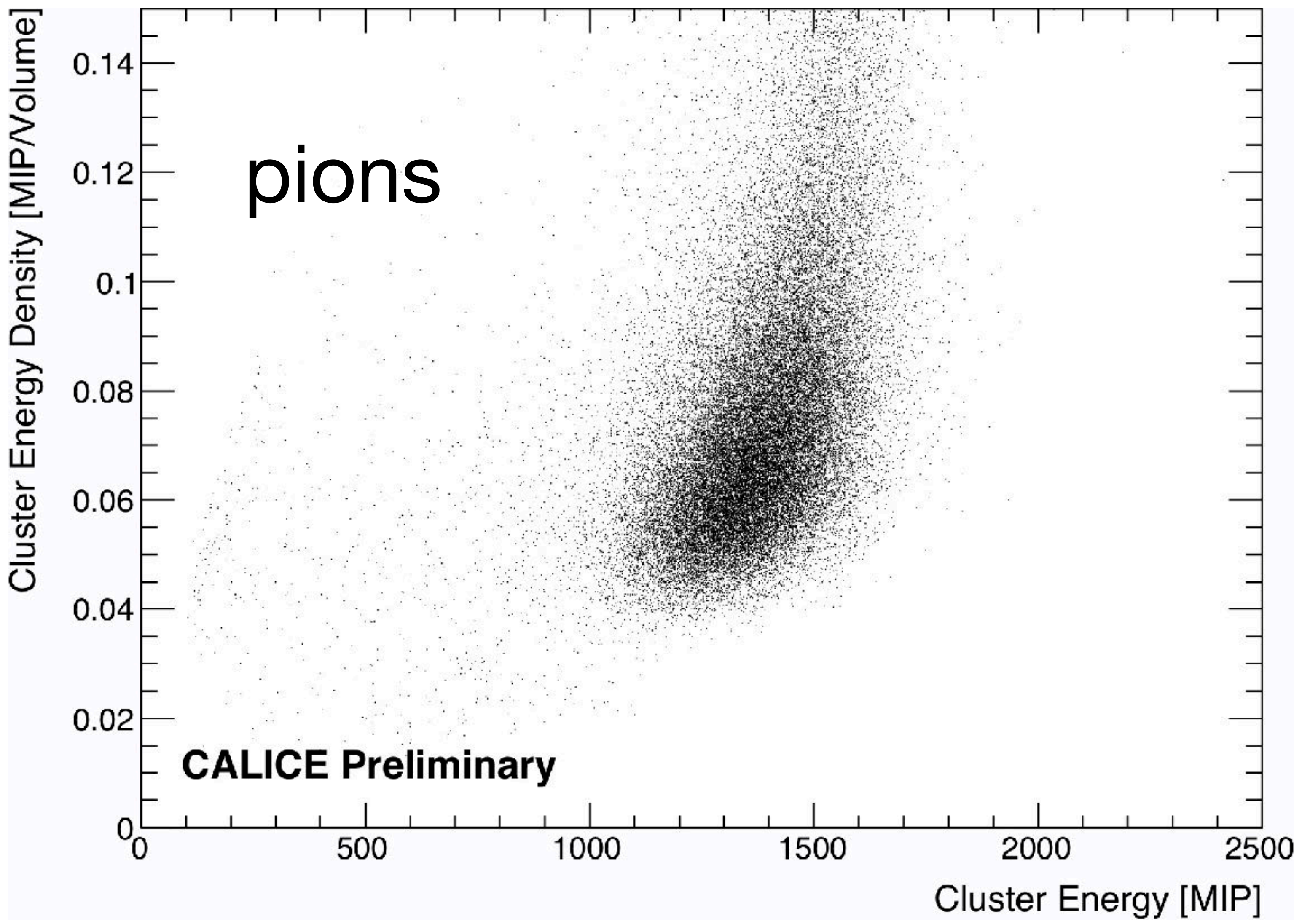
  *The strategy:*

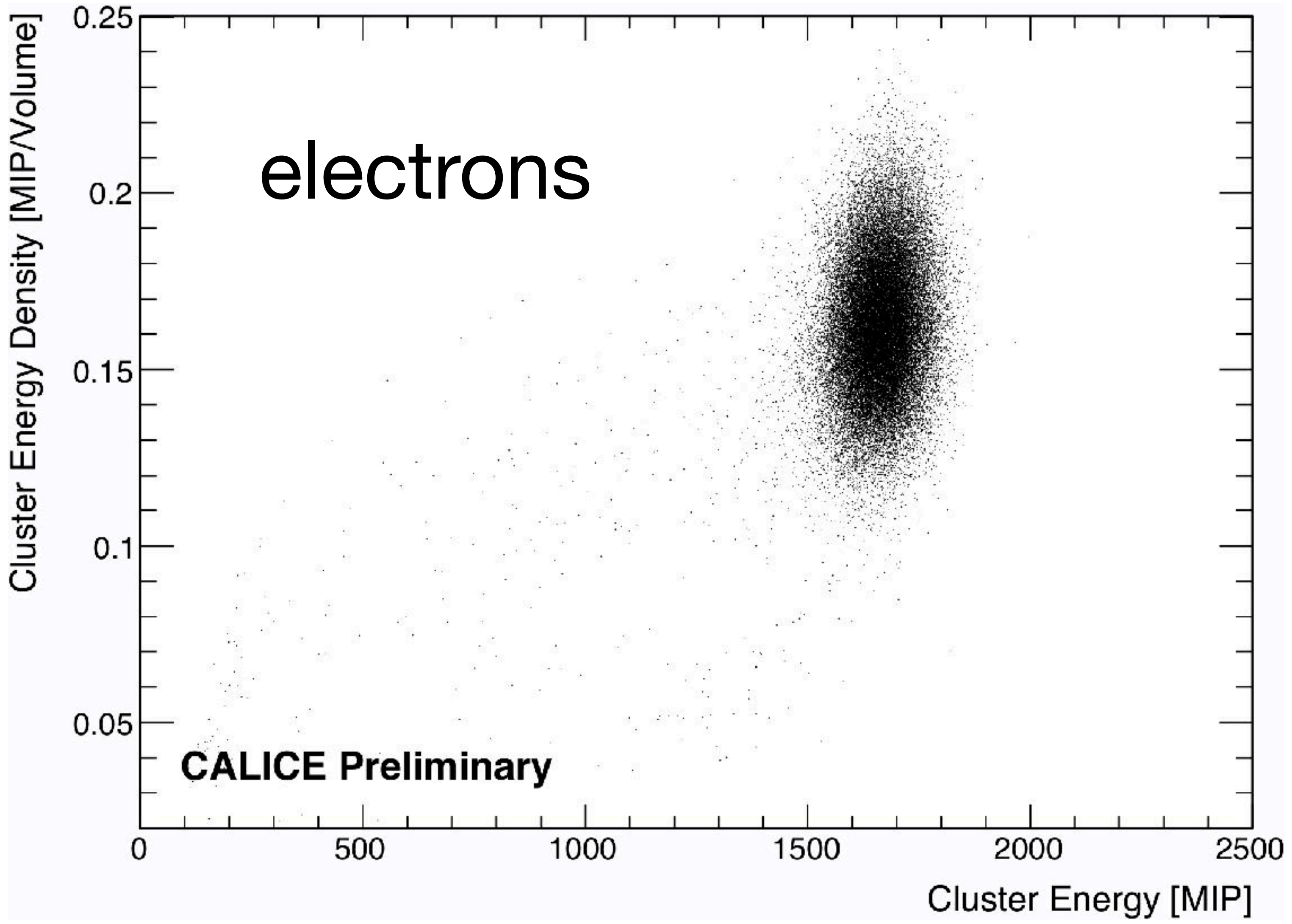
  - Use “simple” clustering to define a set of shower variables
  - Train a neural network on MC data (NB: quasi-continuous energy distribution to avoid bias)
  - Apply NN to data (requires additional energy correction to account for differences between data and MC)




Shower variables used in NN:
Global observables



# Towards Multivariate Techniques

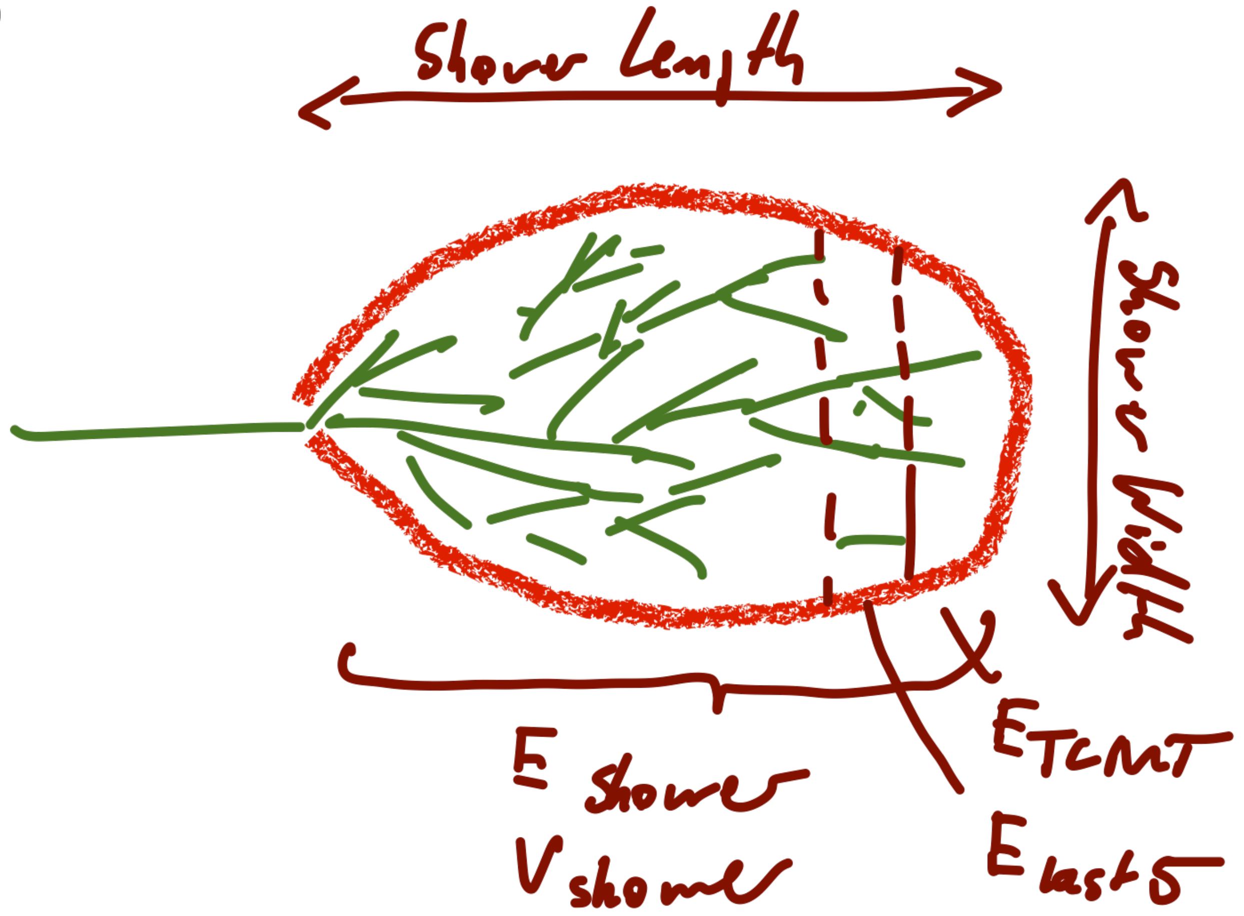
*Initial Studies with Neural Networks*

CAN-021

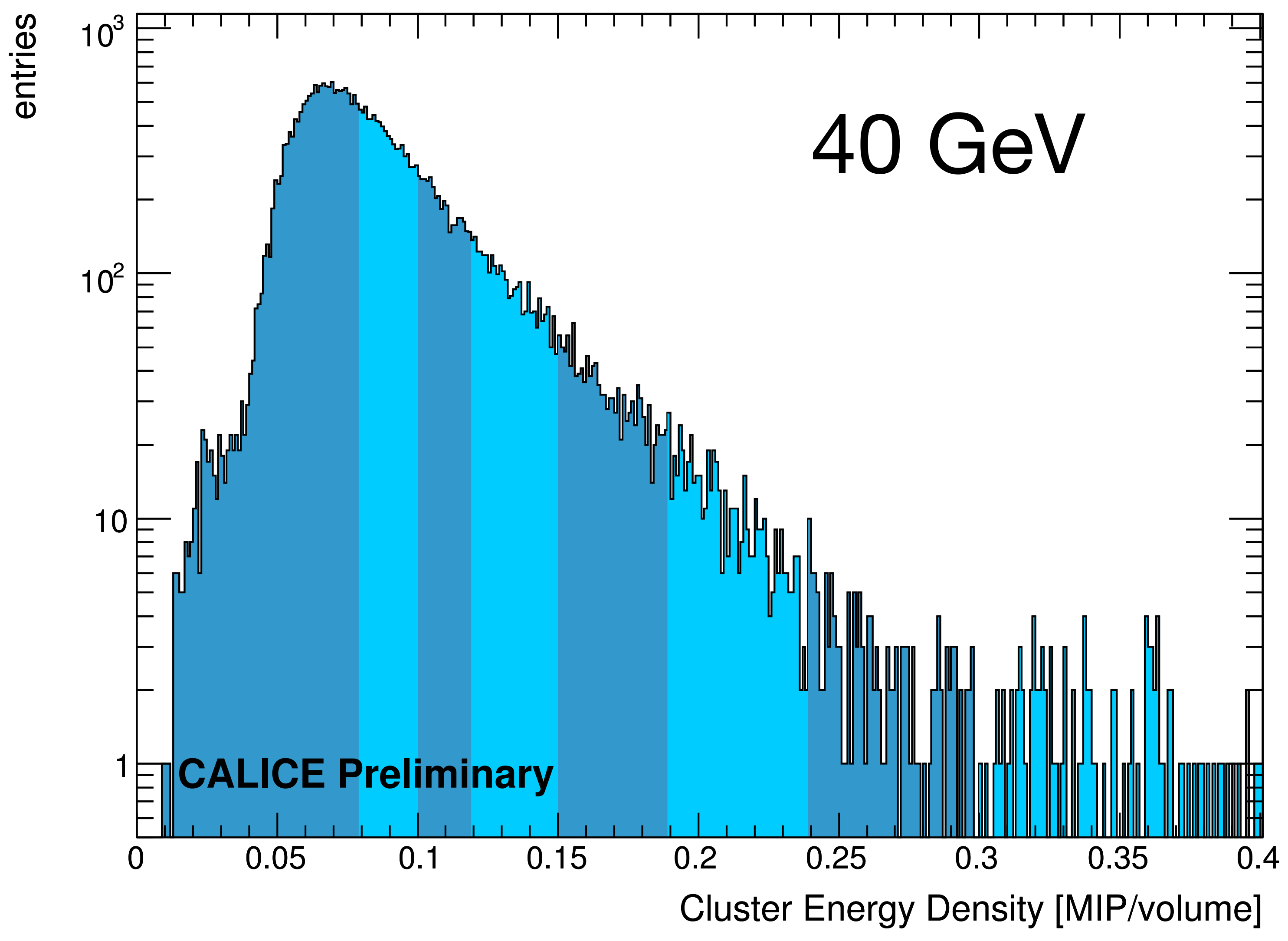
- Performed with AHCAL physics prototype in 2010 (K. Seidel, FS)

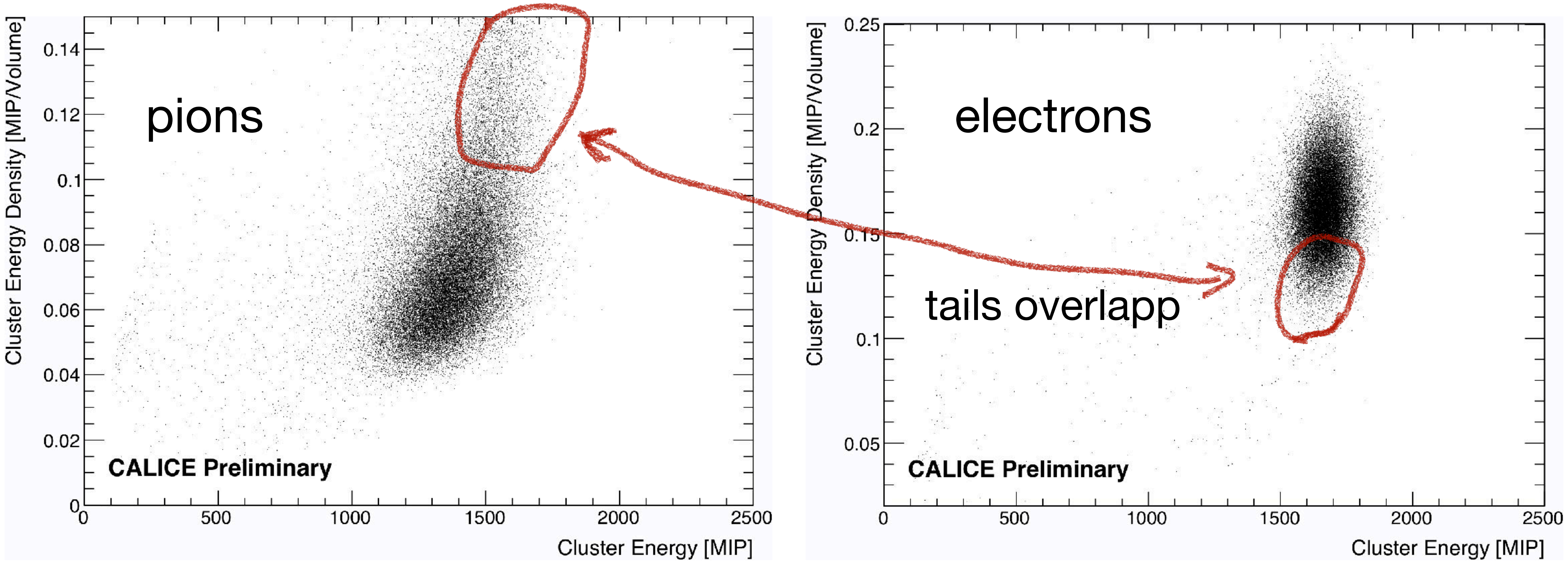
  *The strategy:*

  - Use “simple” clustering to define a set of shower variables
  - Train a neural network on MC data (NB: quasi-continuous energy distribution to avoid bias)
  - Apply NN to data (requires additional energy correction to account for differences between data and MC)




Shower variables used in NN: Global observables



# Towards Multivariate Techniques

*Initial Studies with Neural Networks*

CAN-021

- Performed with AHCAL physics prototype in 2010 (K. Seidel, FS)

  *The strategy:*

  - Use “simple” clustering to define a set of shower variables
  - Train a neural network on MC data (NB: quasi-continuous energy distribution to avoid bias)
  - Apply NN to data (requires additional energy correction to account for differences between data and MC)

Simple weighting using energy density only with parametrized weights from MC: ~ 15% improvement

# Towards Multivariate Techniques

*Initial Studies with Neural Networks*

CAN-021

- Performed with AHCAL physics prototype in 2010 (K. Seidel, FS)

  *The strategy:*

  - Use “simple” clustering to define a set of shower variables
  - Train a neural network on MC data (NB: quasi-continuous energy distribution to avoid bias)
  - Apply NN to data (requires additional energy correction to account for differences between data and MC)

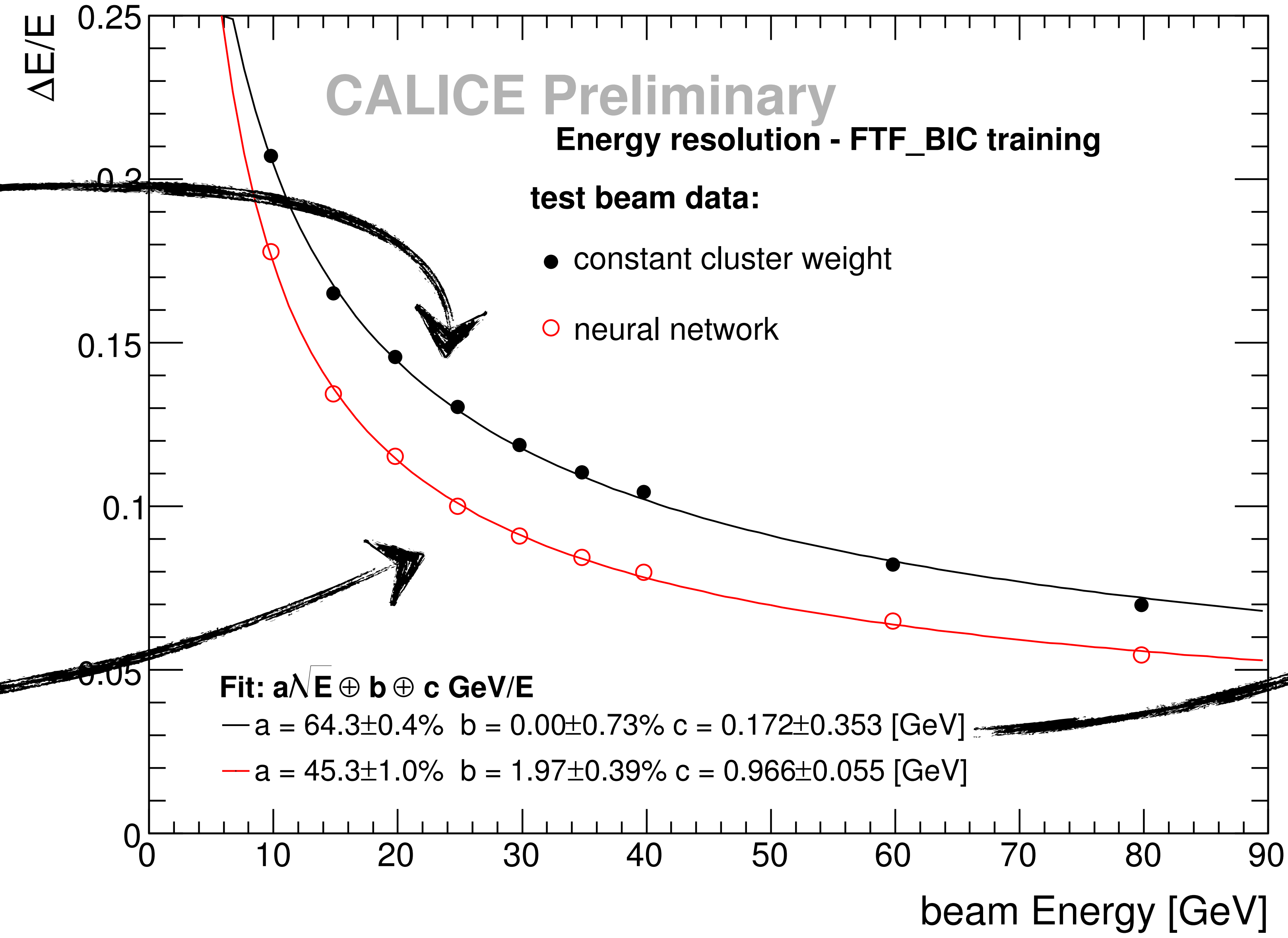
Simple weighting using energy density only with parametrized weights from MC: ~ 15% improvement

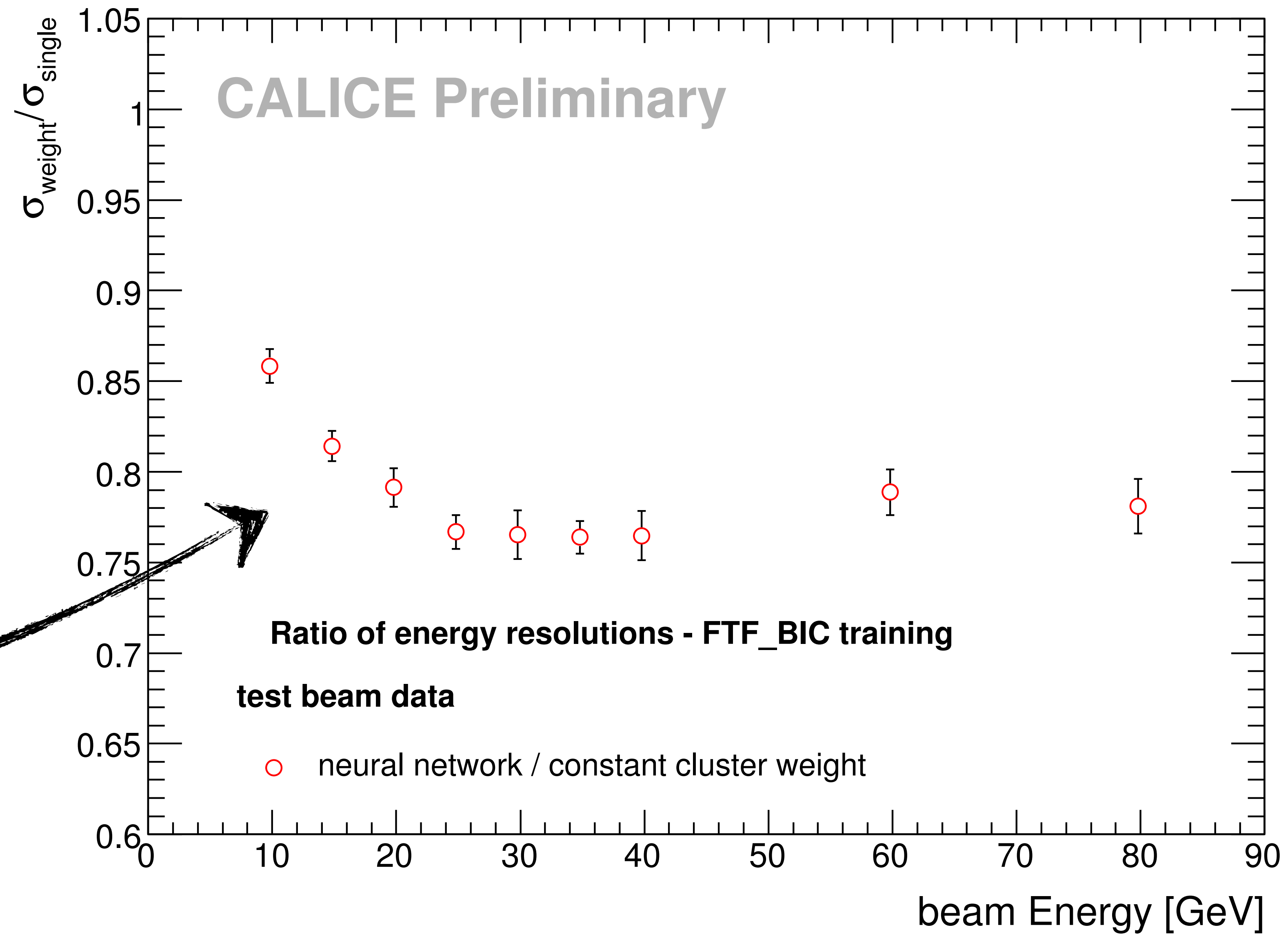
NN: up to 25% improvement

default resolution ~ 10-15% worse than other studies: clustering effects!

NN achieves performance comparable to other techniques, slightly larger improvement

⇨ *Substantial potential!*



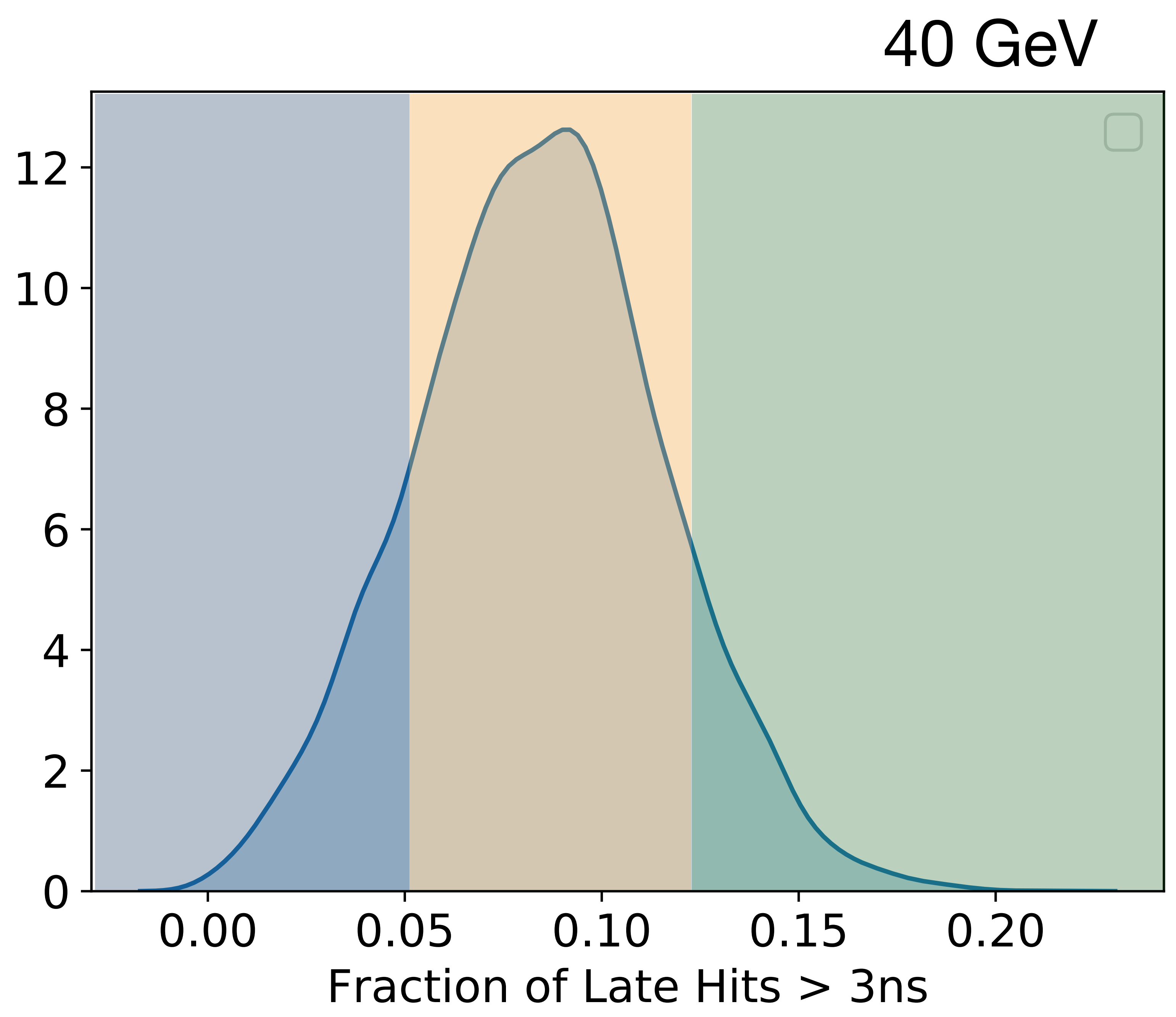
# Towards Multivariate Techniques

*Machine Learning & Timing Information*

C. Graf, work in progress

- New prototypes (and full detectors) will offer ns-level timing on the cell level
  - Obvious benefits for pattern recognition & background rejection
  - Benefits for energy resolution?

Simulation study for
AHCAL prototype



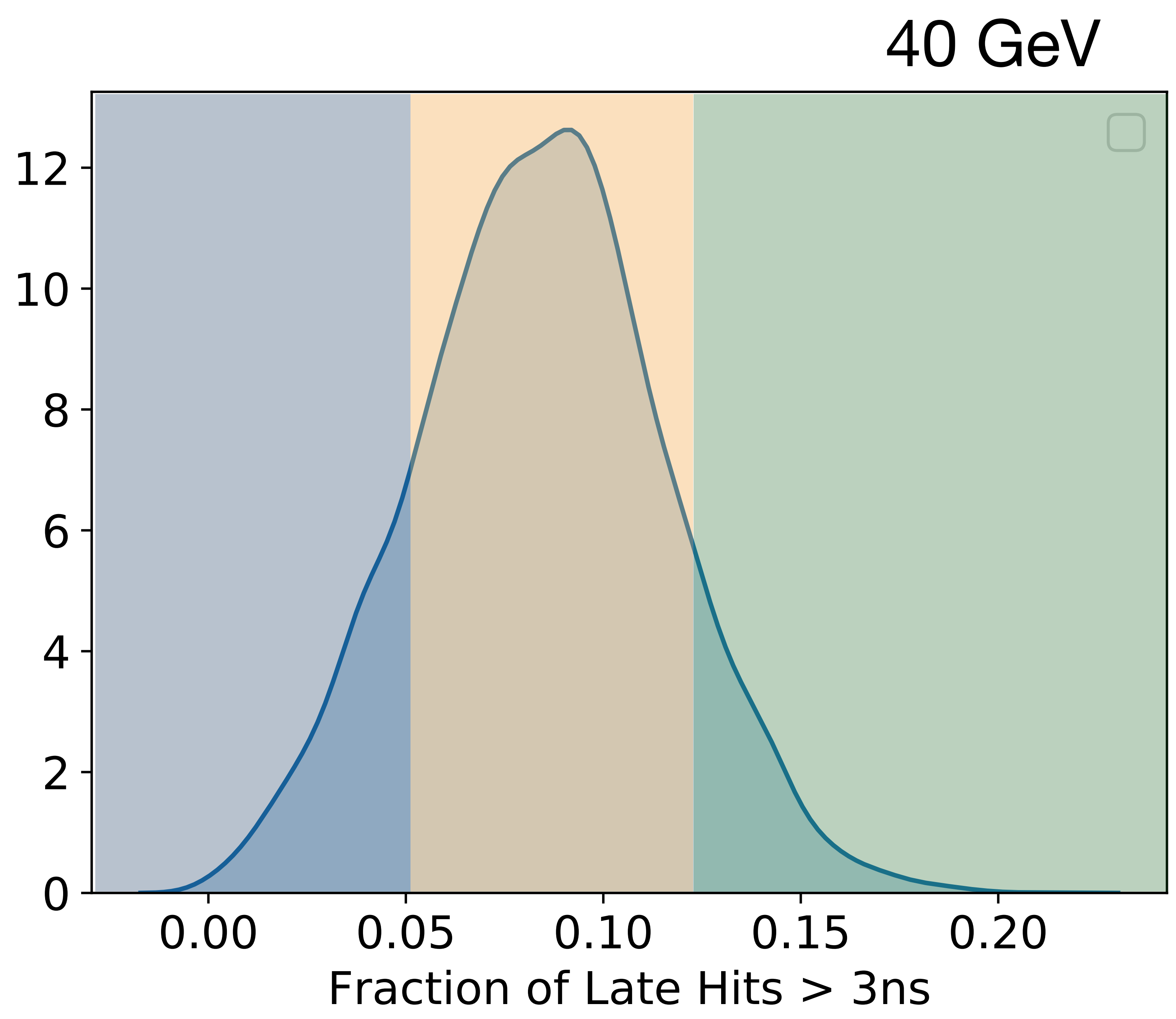
# Towards Multivariate Techniques

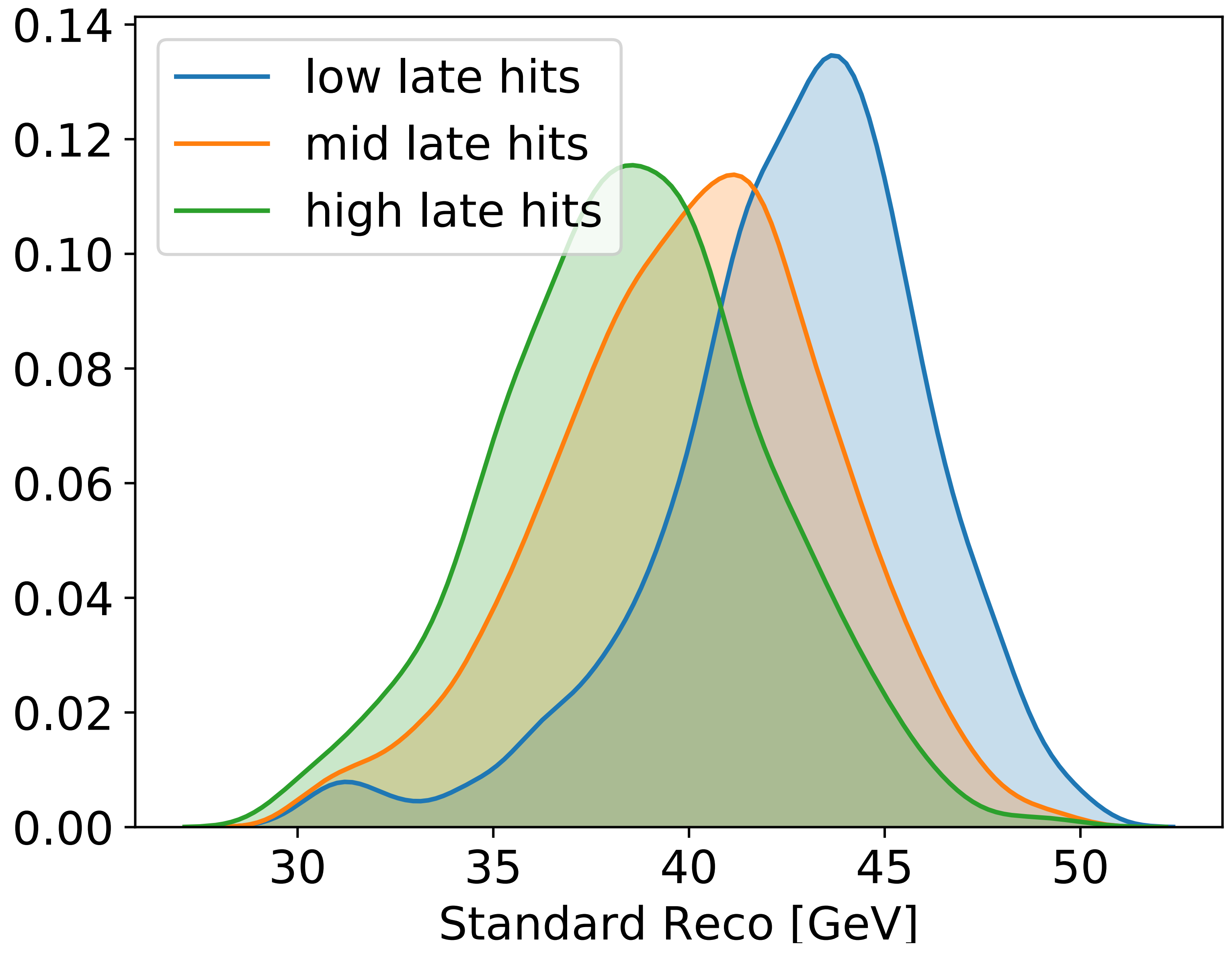
*Machine Learning & Timing Information*

C. Graf, work in progress

- New prototypes (and full detectors) will offer ns-level timing on the cell level
  - Obvious benefits for pattern recognition & background rejection
  - Benefits for energy resolution?

Simulation study for AHCAL prototype




fraction of late hits in shower as additional observable

Additional parameters, such as layer-wise energy in neural network

# Towards Multivariate Techniques

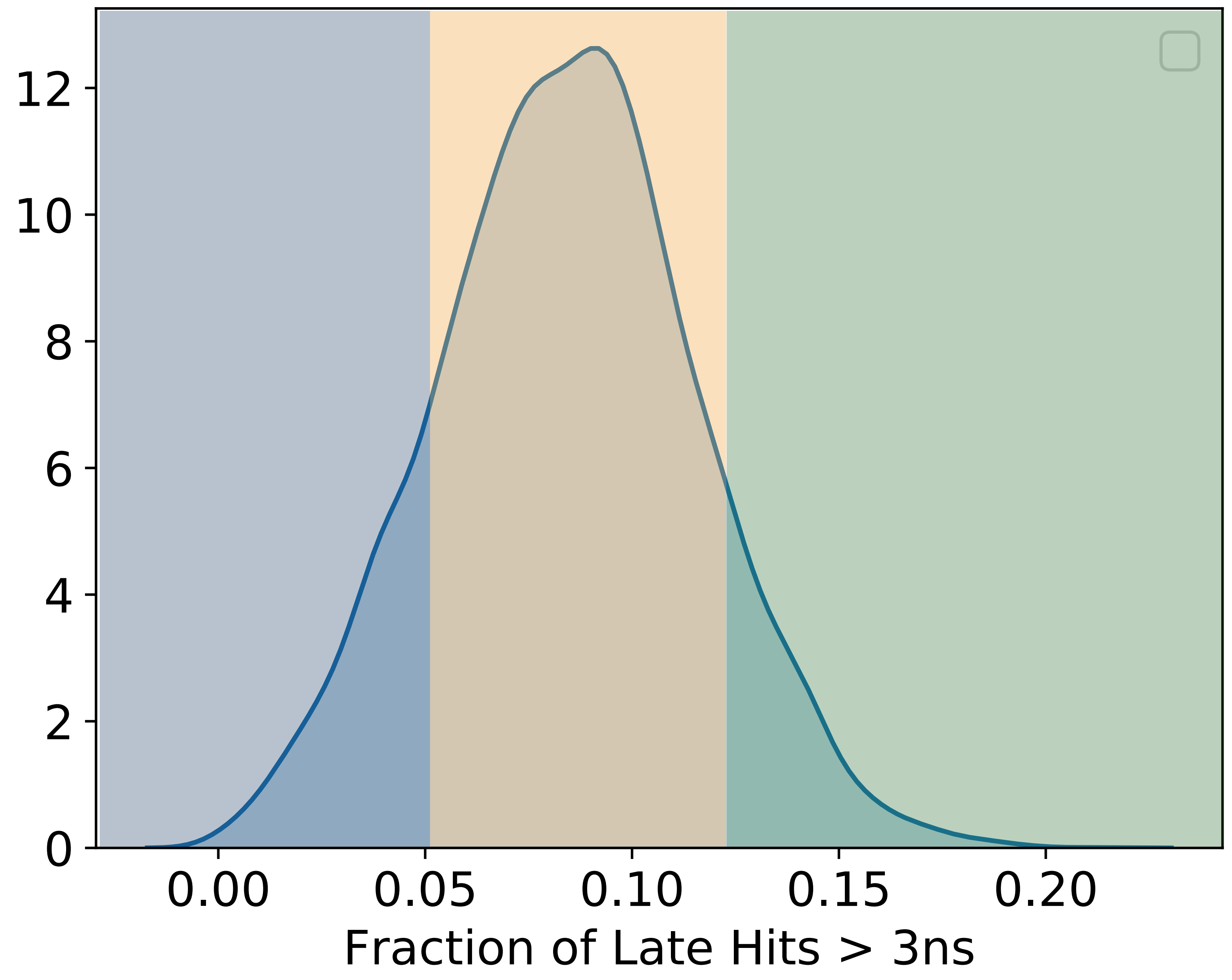
*Machine Learning & Timing Information*

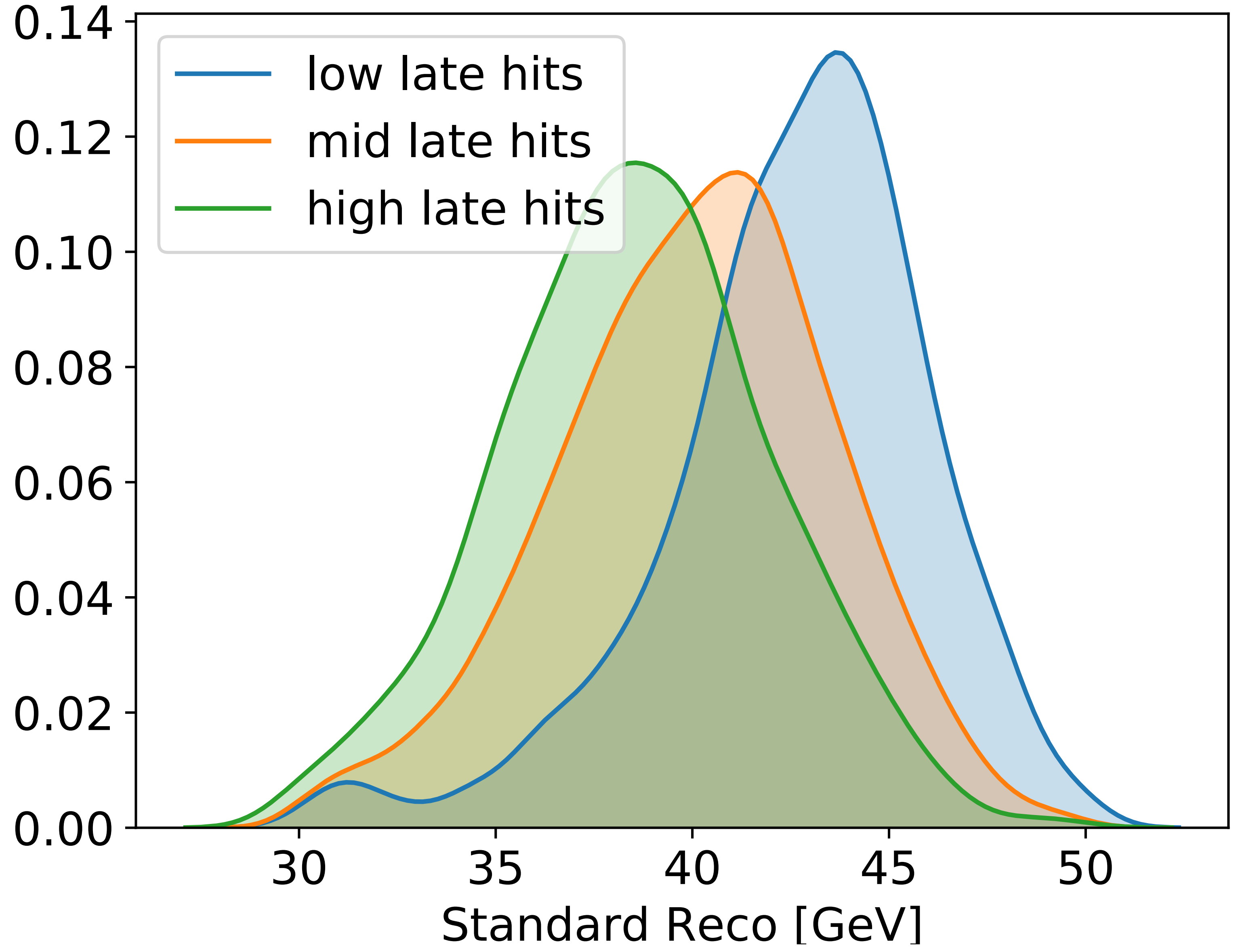
C. Graf, work in progress

- New prototypes (and full detectors) will offer ns-level timing on the cell level
    - Obvious benefits for pattern recognition & background rejection
    - Benefits for energy resolution?

Simulation study for AHCAL prototype

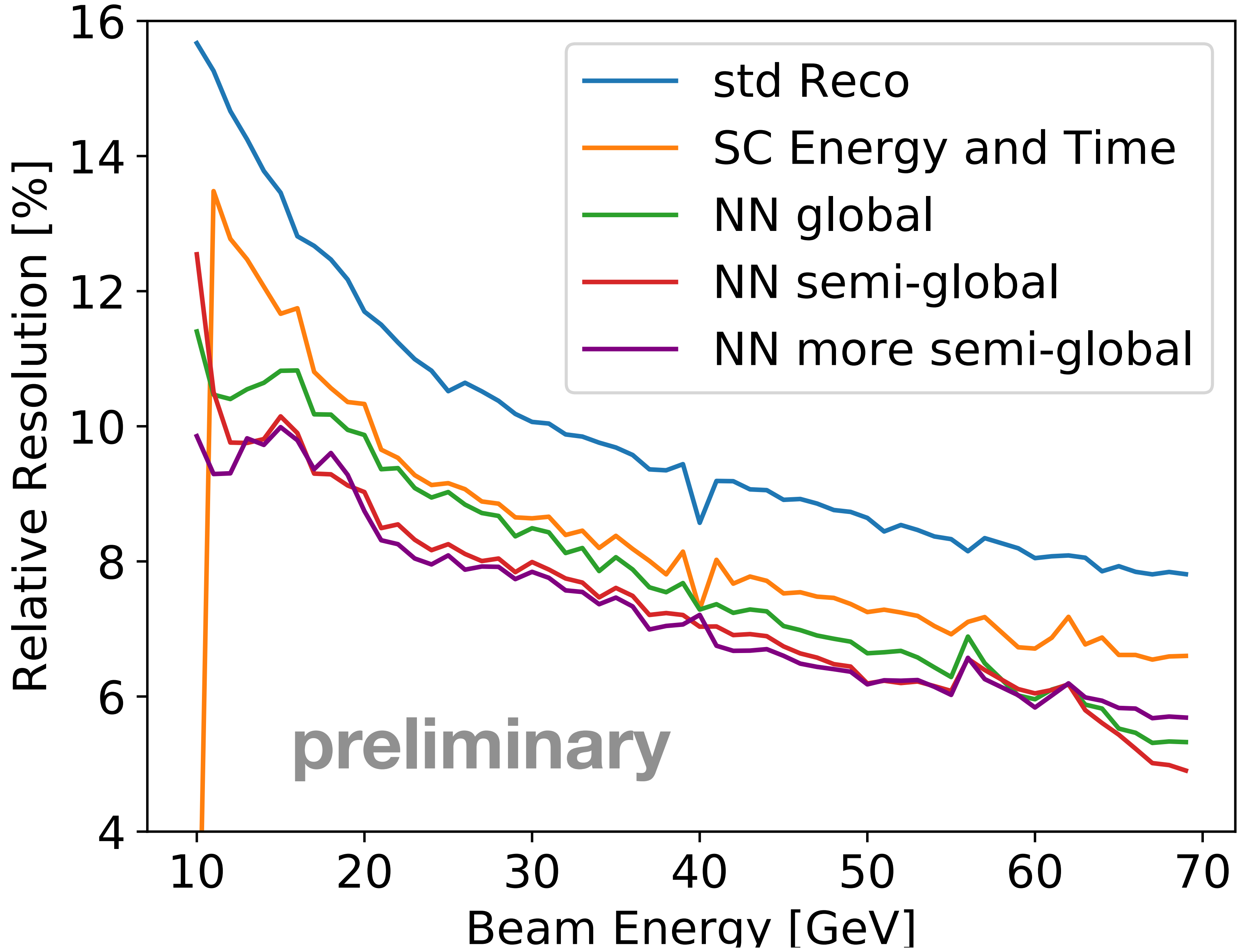
40 GeV




fraction of late hits in shower as additional observable

Additional parameters, such as layer-wise energy in neural network




*Study still in an early stage*

- Timing adds some improvement to global SC (~ 10% in quadrature)
- Energy, in particular also local, more powerful

# Summary & Outlook

- Hadronic energy reconstruction in calorimeters is a challenge - and a limiting factor for overall detector performance
- Highly granular calorimeters provide detailed information on the shower substructure on an event-by-event level that can be use to improve the energy reconstruction & resolution: Used in ***Software Compensation***

- Different techniques developed and studied in CALICE with test beam data:
  Global and local software compensation, semi-digital reconstruction,
  global software compensation with neural networks
  - Successfully applied to single detectors and combined ECAL and HCAL systems - with typical resolution improvement of 20% - 30% for pions with energies above ~ 15 GeV
  - Implemented in PandoraPFA for the AHCAL

- Substantial potential for further improvement: Addition of new variables (time), more sophisticated machine learning techniques, extension to electromagnetic showers, …

# Extras

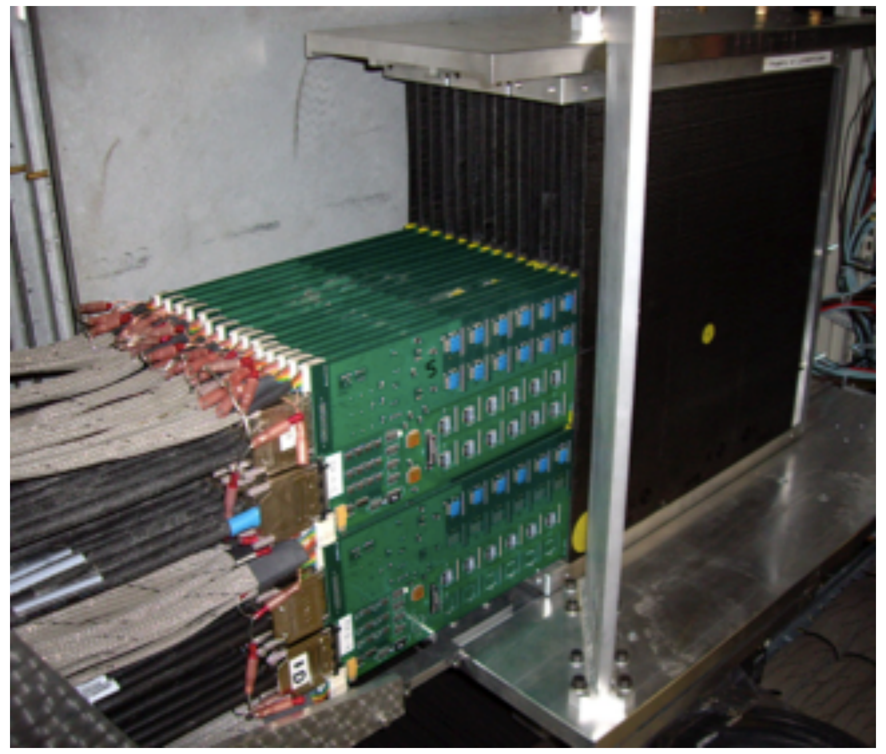
# CALICE Technologies

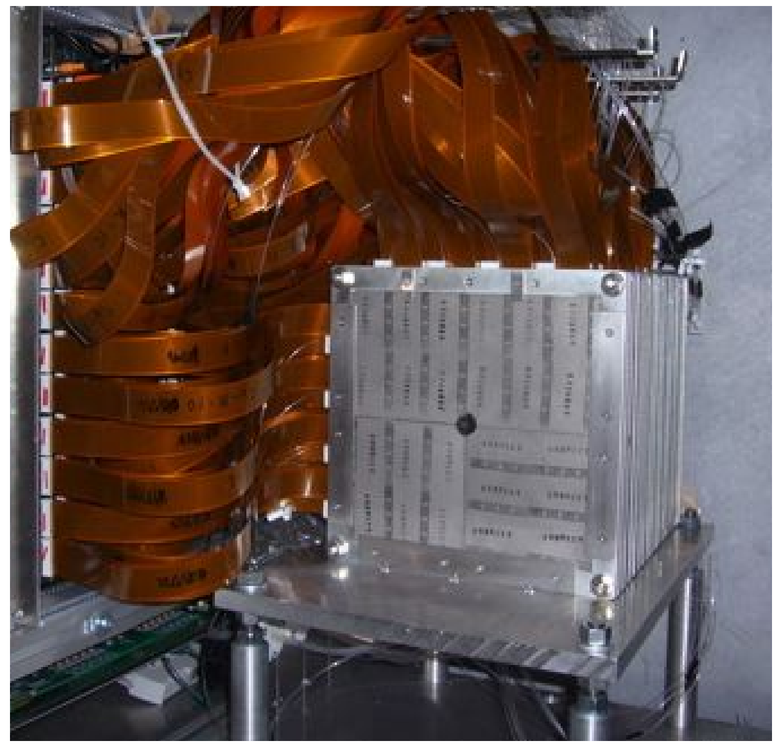
*A wide range of prototypes*

- A rich test beam program, with a variety of different prototypes

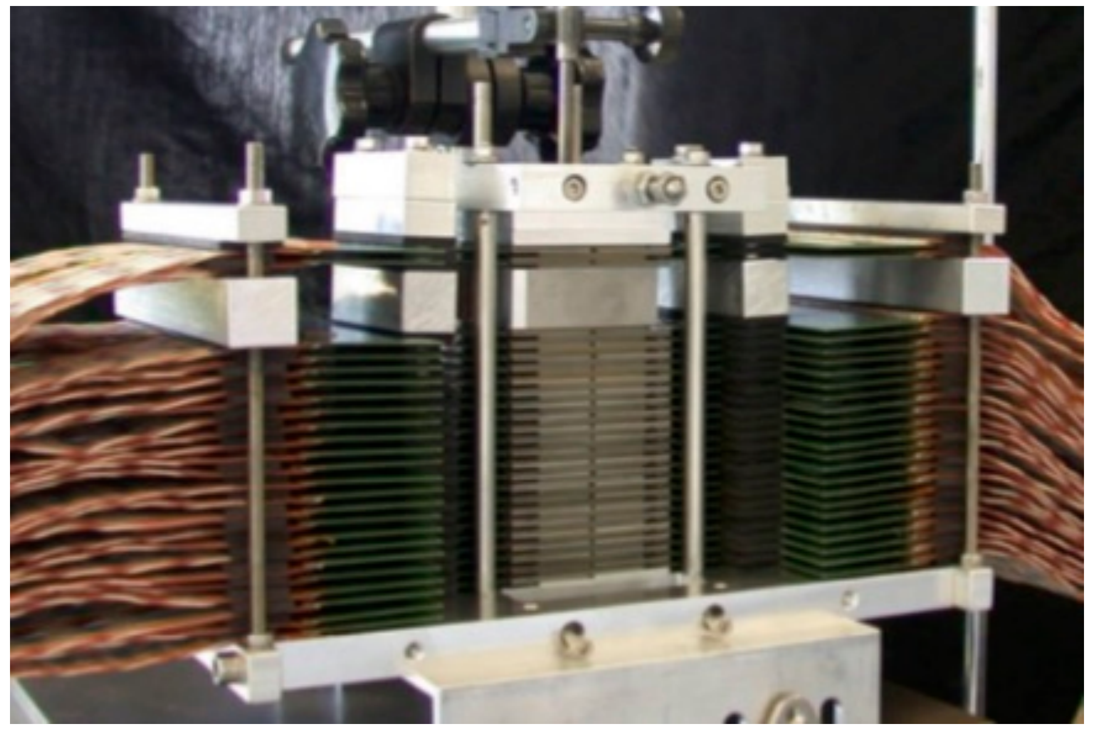
Electromagnetic - Tungsten absorbers

analog: Silicon and Scintillator/SiPM





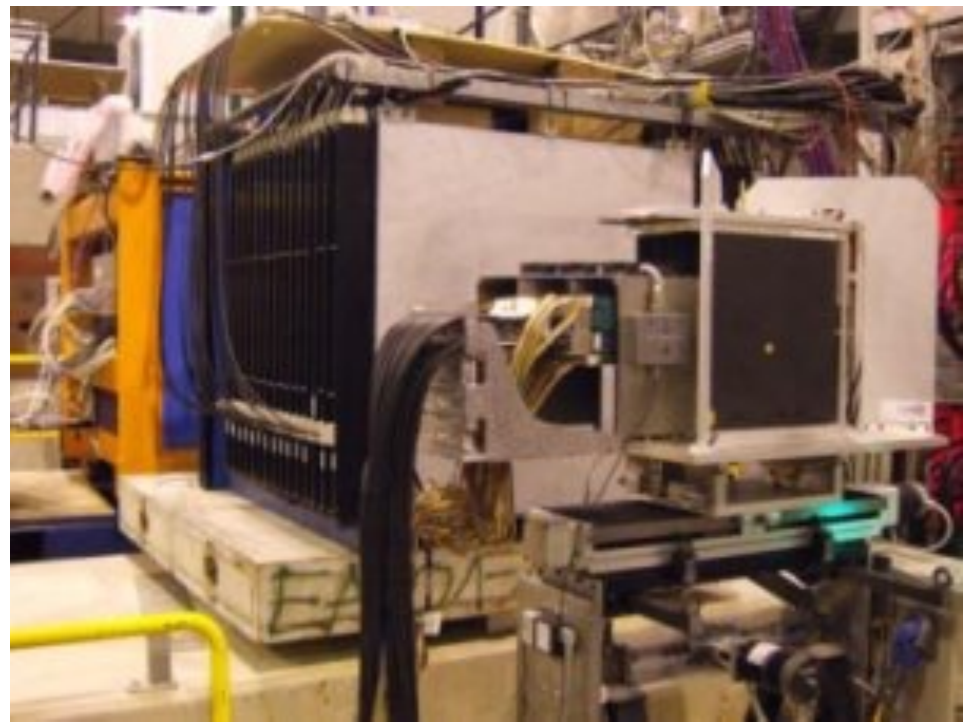
digital: Silicon (MAPS)



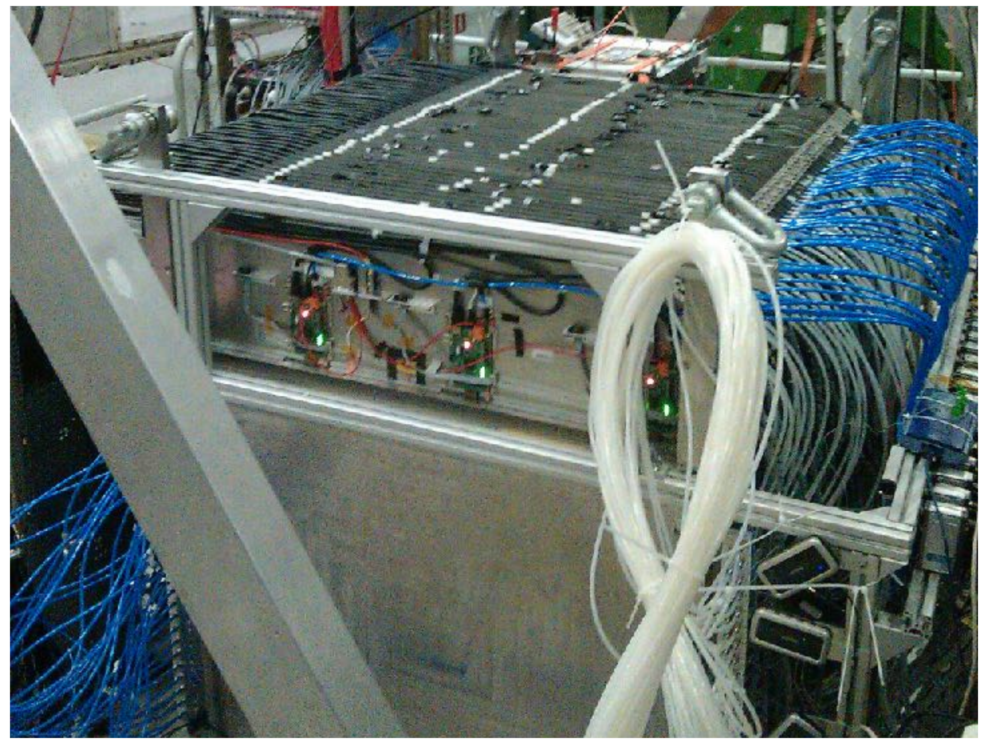
39 Mpixels in 160 $cm^2$

Hadronic - Steel and Tungsten absorbers

analog:
Scintillator/SiPM
(Fe and W)



(Semi)digital: RPCs (Fe, W digital only)



+ few-layer SD prototype with Micromegas

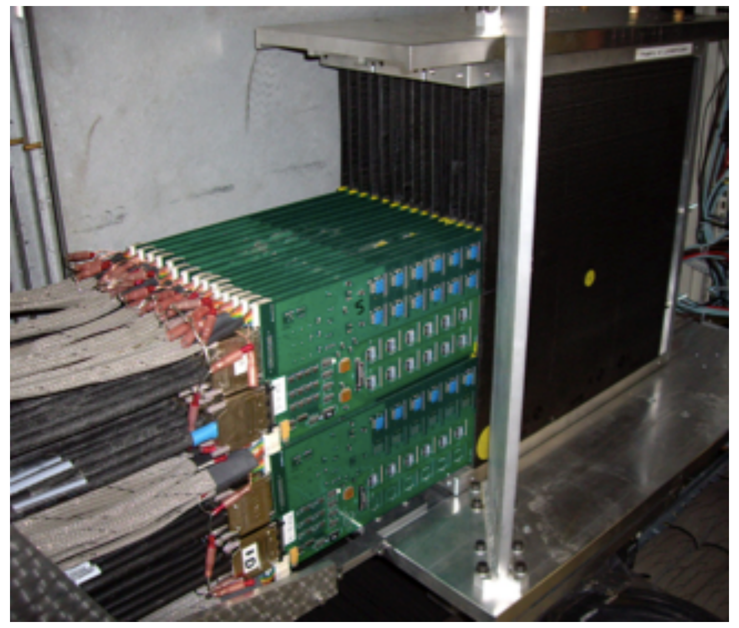
# CALICE Prototypes

*Evolution with Time*

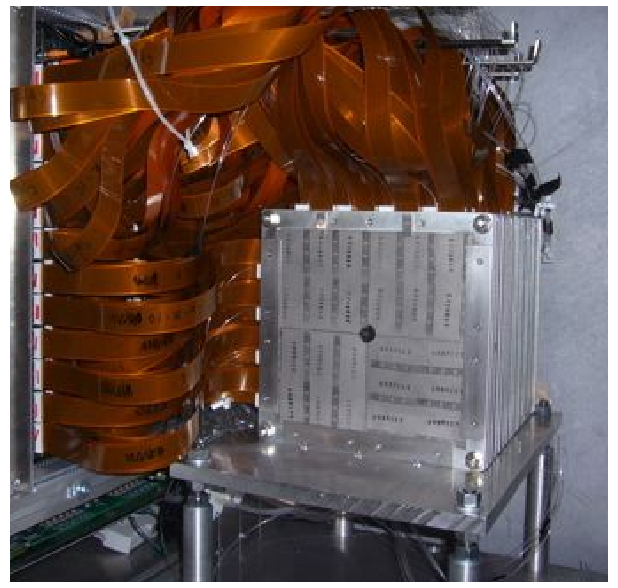
Physics Prototypes

Technological Prototypes

SiW ECAL

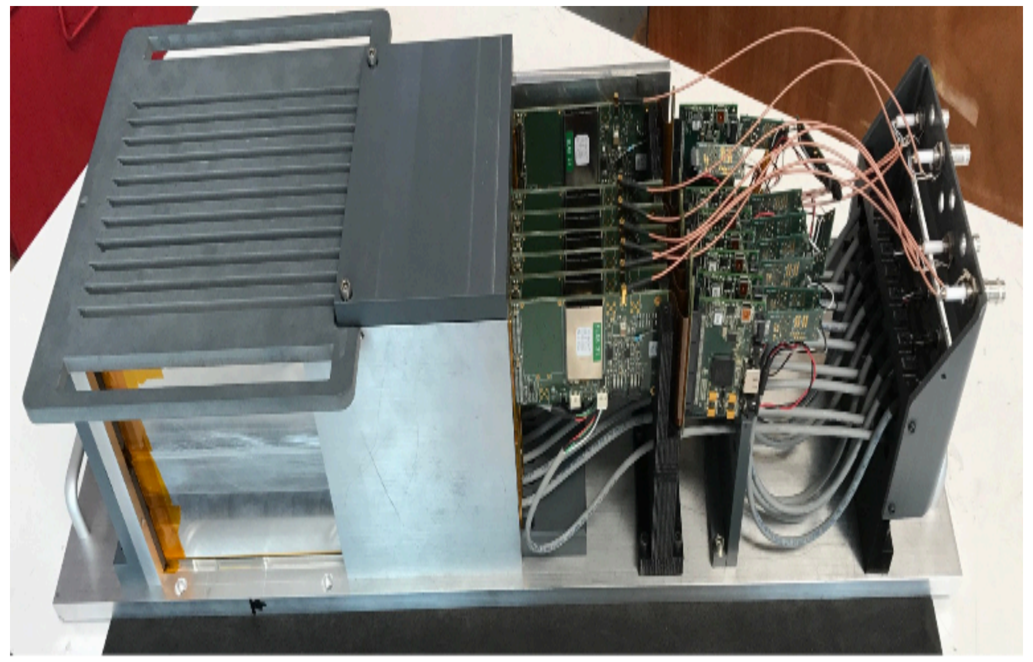


ScintW ECAL



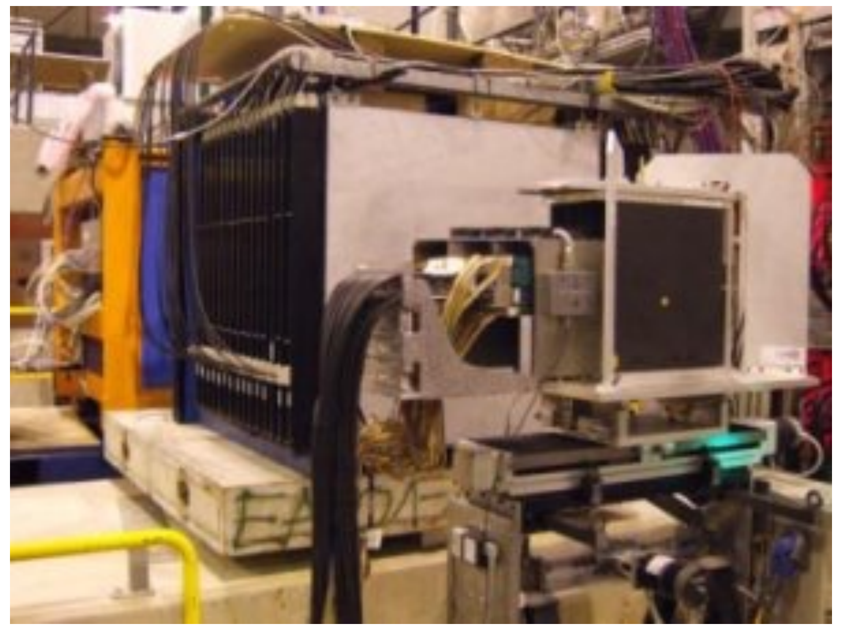
SDHCAL

SiW ECAL



2005 2006 2007 2008 2010 2012 2018 year of (first) TB

AHCAL

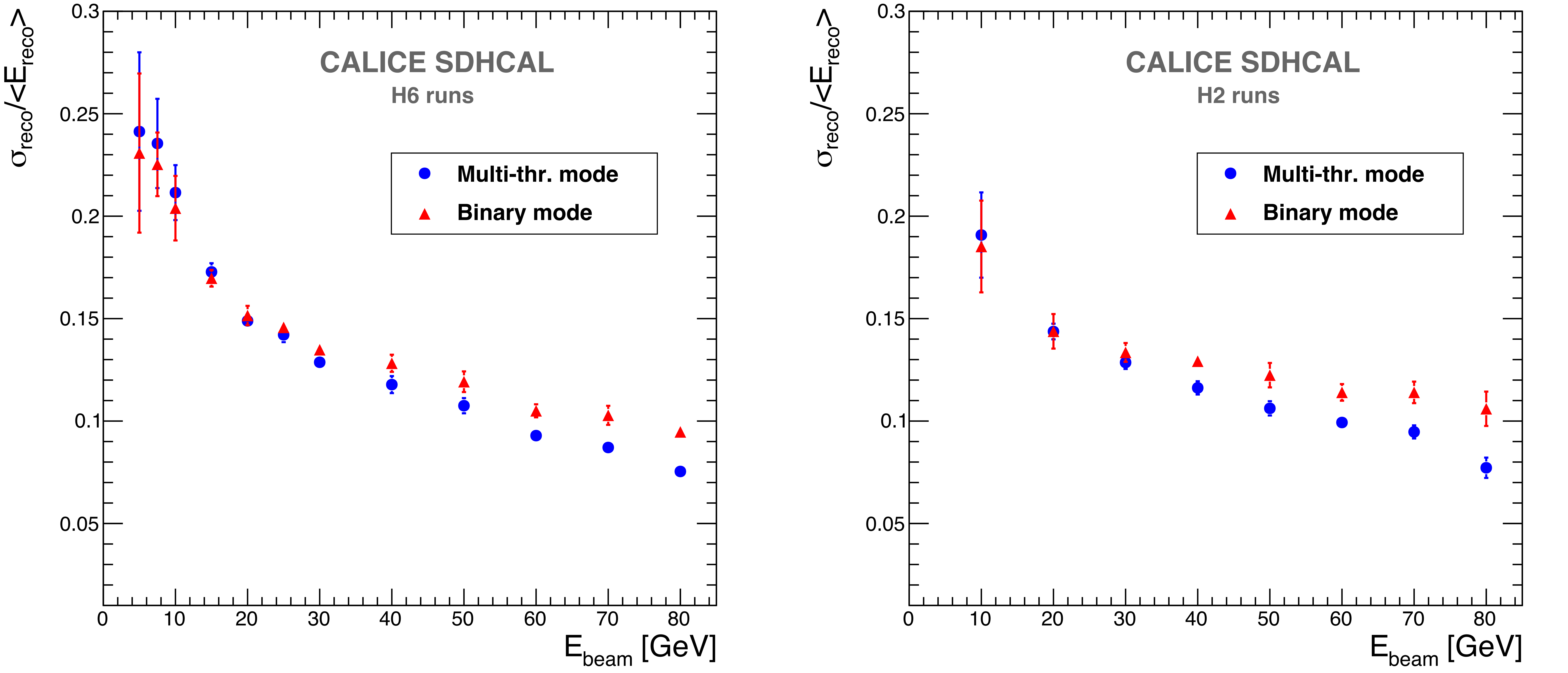


DHCAL

AHCAL

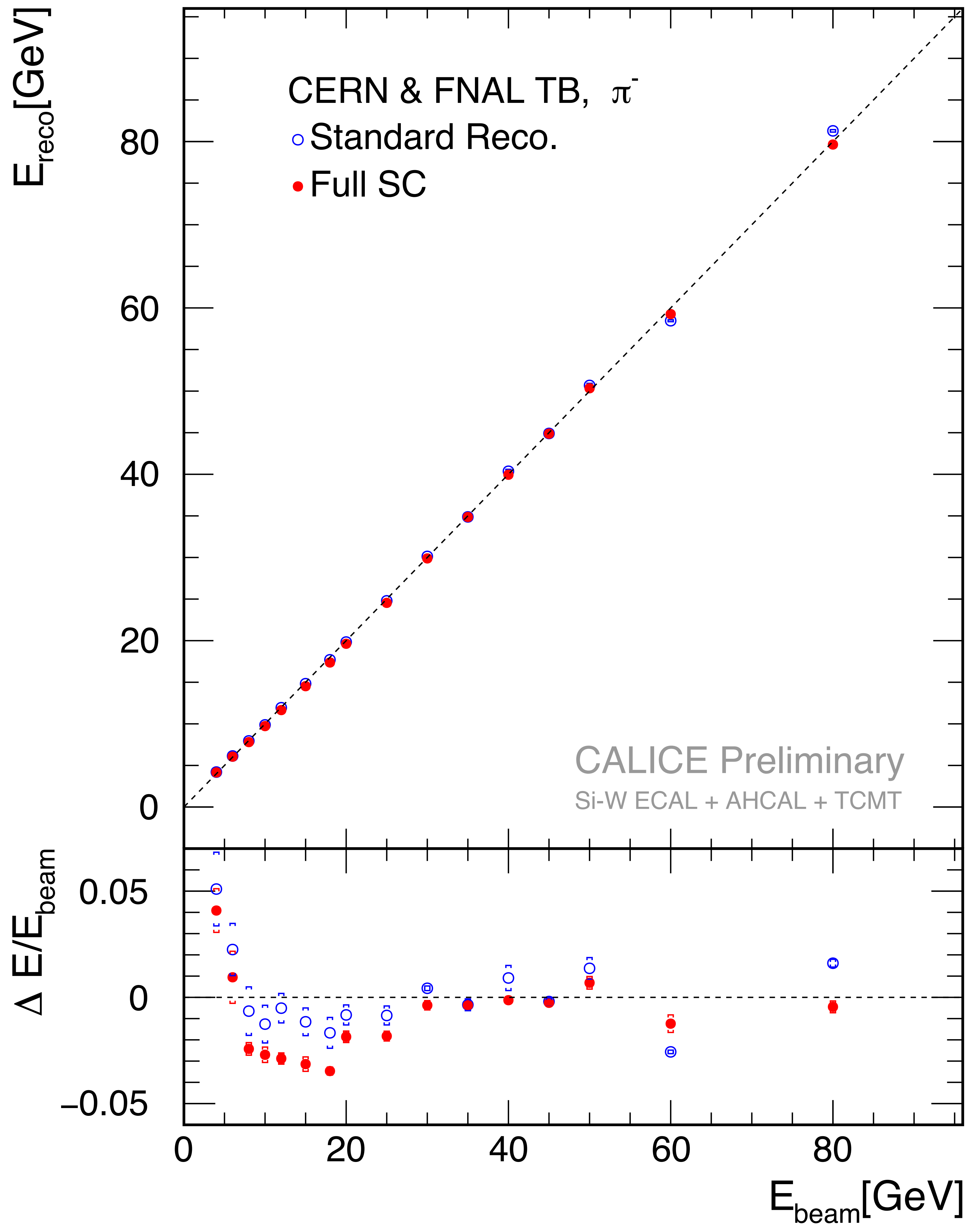
also: W-AHCAL also: W-DHCAL

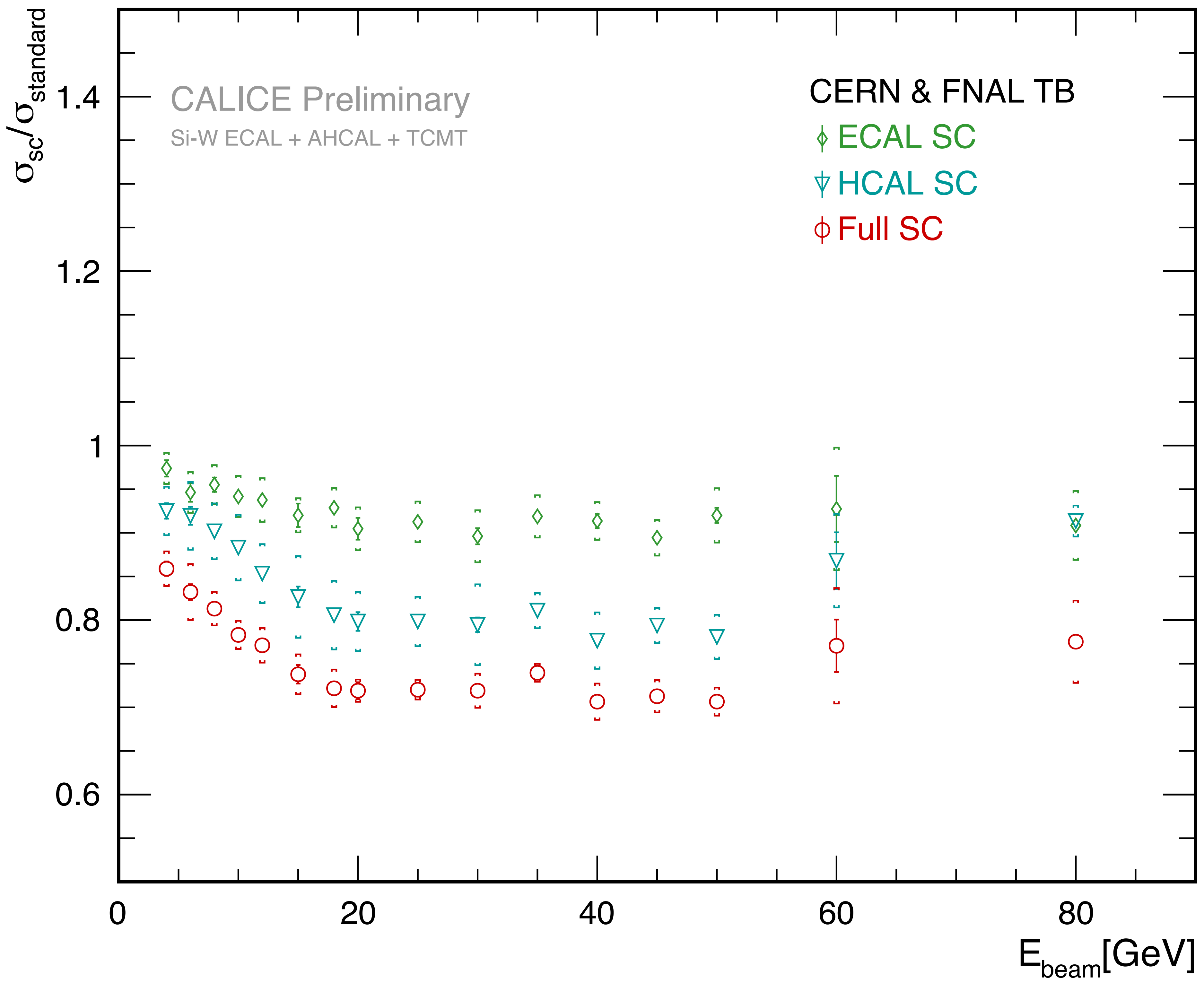
- At energies above 30 GeV semi-digital reconstruction provides a substantial performance advantage wrt digital reconstruction: Saturation effects become relevant



# Comined ECAL/HCAL Software Compensation

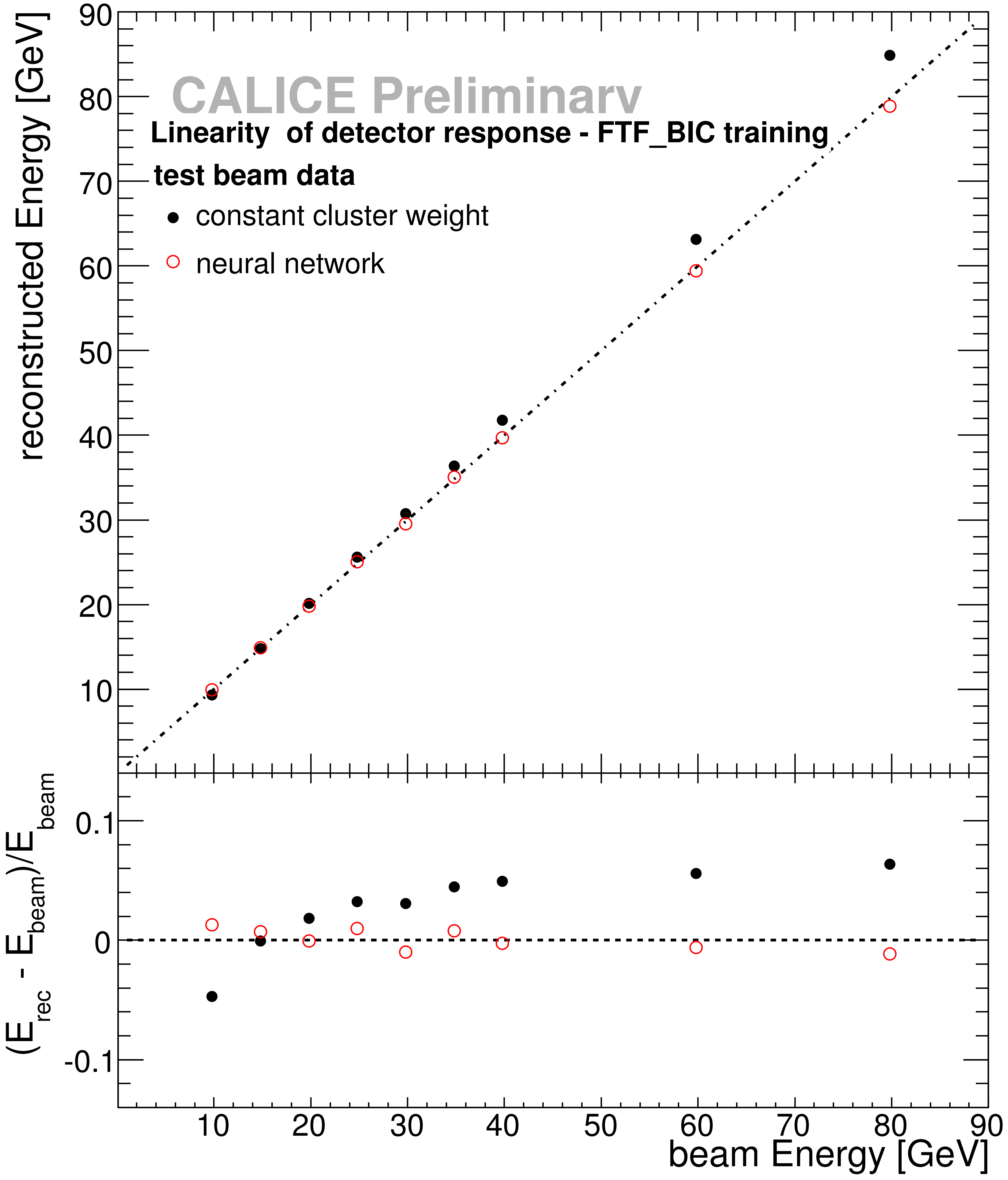
*Linearity & Resolution Improvement*



# Software Compensation with Neural Networks

*CALICE AHCAL: Linearity*

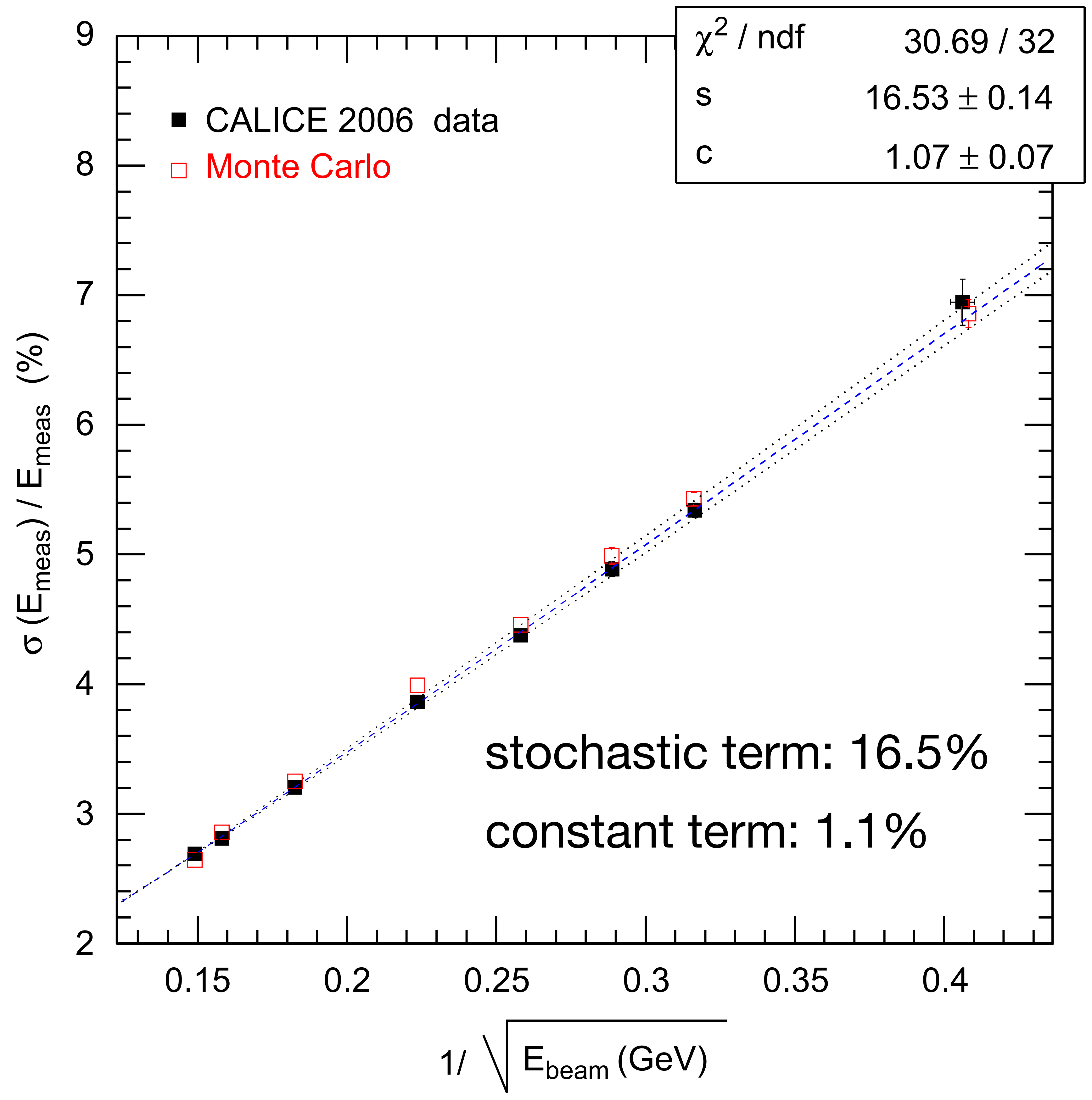


# Performance of Highly Granular Calorimeters

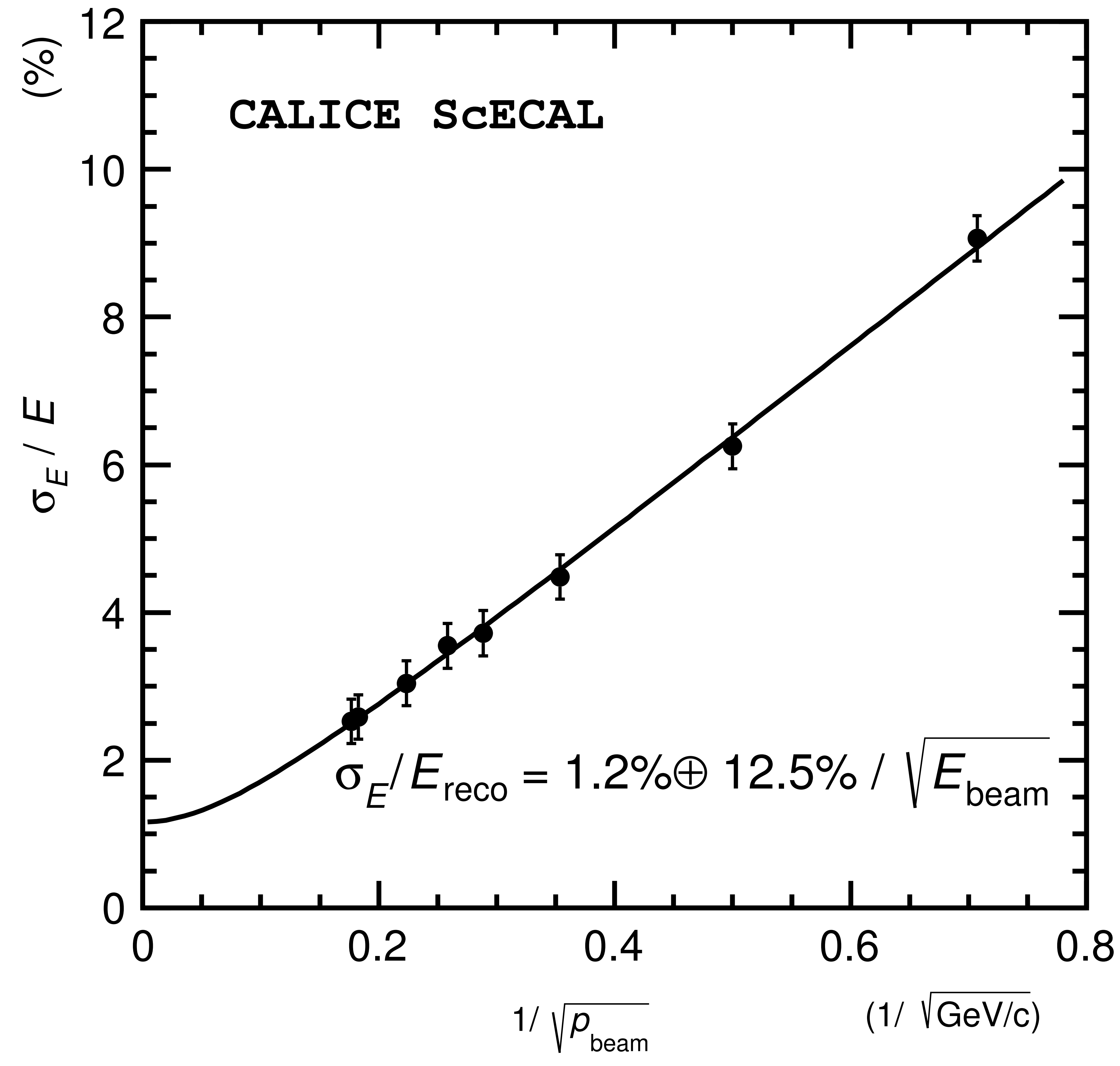
*Energy resolution - Electromagnetic*

[N.B. Detector optimized for particle separation, not single particle resolution]

Silicon-Tungsten ECAL:




Scintillator-Tungsten ECAL:

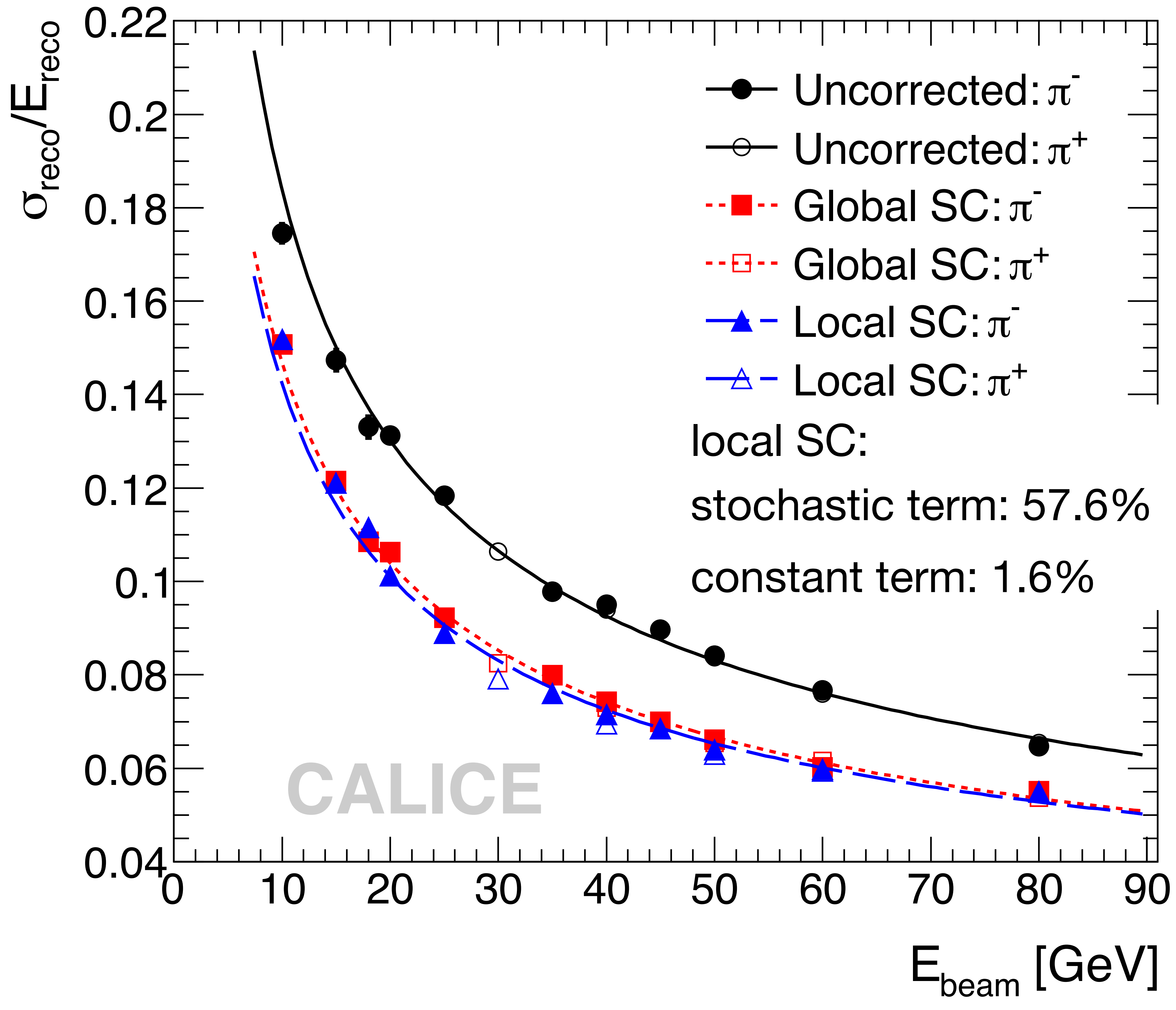



Scintillator provides better energy resolution due to larger sampling fraction, with a reduced compactness

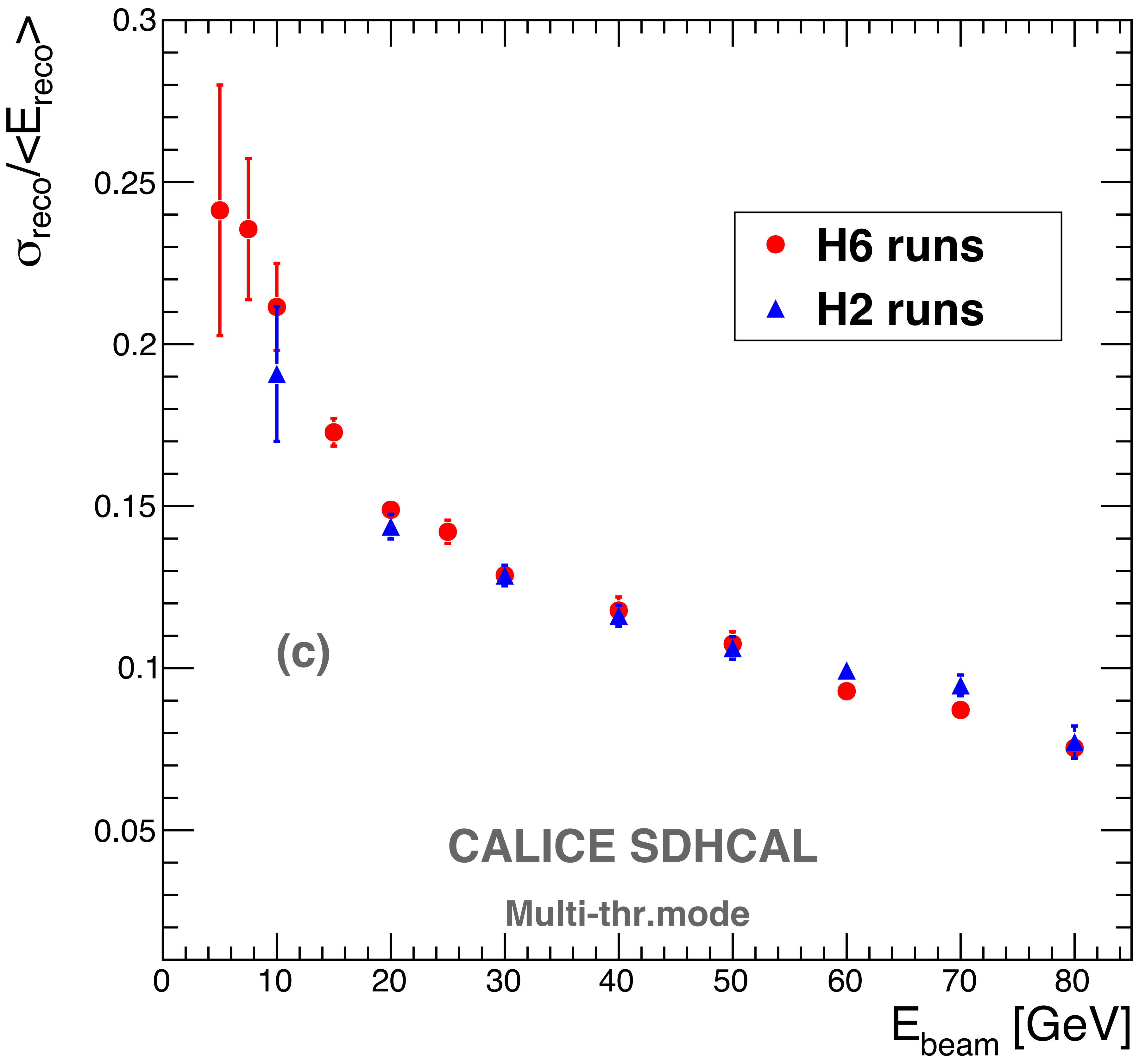
# Performance of Highly Granular Calorimeters

*Energy resolution -Hadronic*

Analog (Scintillator + SiPM)




semi-digital (RPCs)




Software compensation (SC) and semi-digital reconstruction use weighting factors to optimise energy resolution