Systematic uncertainties in recent AHCAL analyses

Marina Chadeeva (ITEP, MEPhI), Eva Sicking (CERN) and Felix Sefkow (DESY)

Overview

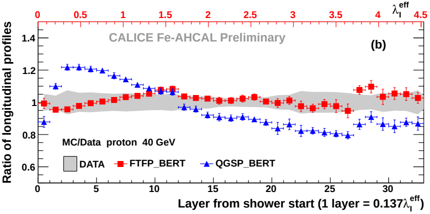
Two example analyses

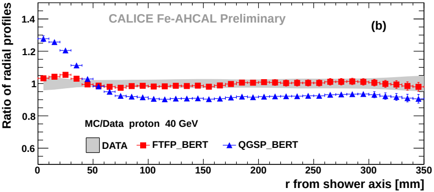
- Fe AHCAL:
  Pion and proton showers in the CALICE scintillator-steel analogue hadron calorimeter, *arXiv:1412.2653*, accepted by JINST
  Parametrisation of hadron shower profiles in the CALICE Sc-Fe AHCAL, *CAN-048*
- W AHCAL:
  Shower development of particles with momenta from 15 GeV to 150 GeV in the CALICE Scintillator-Tungsten Hadronic Calorimeter, *paper draft in EB*, **not even preliminary**

General approach

- Select pure particle type, possibly also restrict topology (shower start)
- Calibrate data (equalise cell response), correct for saturation and temperature effects
- Simulate optical cross-talk and noise, beam profile and dead cells
- Compare global and differential shower observables with Geant 4 simulations
- Systematics limited, quantify significance of deviations of simulations from data

Fe AHCAL proton analyses





Test beam data

CERN 2007 runs, π^+ @ 30-80 GeV (Si-W ECAL + Fe-AHCAL + TCMT)
FNAL 2009 runs, protons @ 10 and 15 GeV (Fe-AHCAL + TCMT)
Reconstruction with calice_soft v04-08

Simulations

GEANT4.9.6p01, Mokka v08_02
Physics lists: QGSP_BERT and FTFP_BERT
calice_soft v04-08, 846 keV/MIP, 0.15 light crosstalk for the AHCAL

Sources of uncertainties and biases

Mean observables

- MIP to GeV conversion factor (e.m. scale)
- Calibration constants: gain and saturation correction scale
- Leakage from the AHCAL
- Identification of shower axis
- Contamination of samples

Shower profiles

- Leakage from the AHCAL
- Layer-to-layer variations
- Identification of shower start layer
- Identification of shower axis
- Contamination of samples

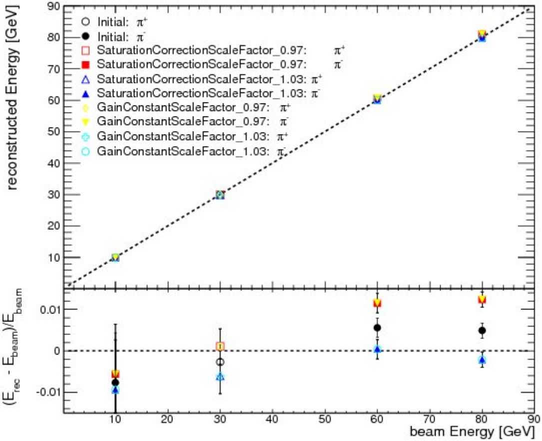
Uncertainties of calibration constants

MIP to GeV conversion factor

- 0.9% (energy independent)
- extracted from e.m. calibration

Gain and saturation correction

- increase with energy
- very conservative estimate: simultaneous shift of constants for all cells ⇒ effect $<1\%$



Leakage

- impact on both mean observables and profile parameters
- minimized by selecting start at the beginning of AHCAL
- does not exceed a few percent at 80 GeV
- negligible impact on data-MC comparisons

Layer-to-layer variations

Observation:
jagged profile, variations increase with energy

Origin:

- dead and noisy cells (few %, some layers more affected, reproduced in MC)
- **saturation correction issues** (cells with unknown scaling factor and gain, not reproduced in MC)

Estimation:

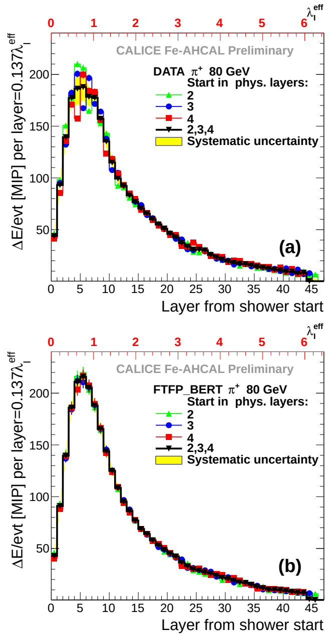
Find average of several normalised profiles from different fixed shower starts

⇓

Calculate mean and variance in each bin, due to different physical layer

⇓

Error of mean = systematic uncertainty



Identification of shower start layer

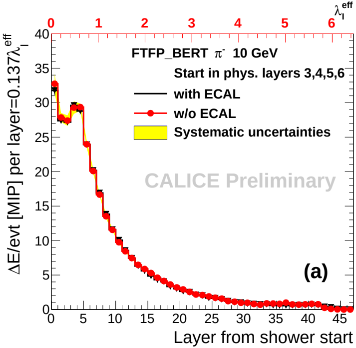
Uncertainties of the algorithm:

- distribution of (found start - true start): r.m.s.≈1 layer
- increases with decreasing energy
- higher for first few layers
  ⇒ higher for runs w/o ECAL

Impact:

- negligible for data and MC comparison: found start layer used in both cases
- profile shape might be affected for low energy samples taken w/o ECAL

Estimation:
Crosscheck with simulations: difference within layer-to-layer variations



Identification of shower axis

Algorithm with ECAL:

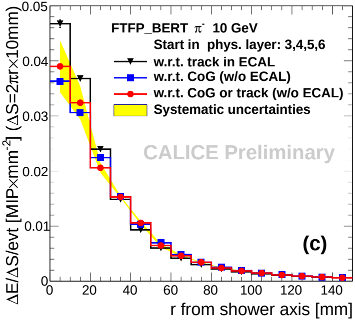
- track in ECAL (granularity 1×1 cm^2, 25-30 track hits)

Algorithm w/o ECAL:

- track in AHCAL if start behind fourth physical layer (granularity 3×3 cm^2, >4 track hits)
- shower centre of gravity (>100 shower hits)

Problem w/o ECAL:

- both approaches not accurate
- CoG underestimates contribution near shower axis



Rough estimate:
from variance extracted from difference between different approaches

N.B.: found start for all ⇒ negligible impact on data and MC comparison

Contamination biases to observables

Electron/positron contamination

- crucial for runs w/o ECAL (e.g. fraction of e^+ can exceed 35% at 10 GeV)
- achieved purity $>$97% due to dedicated selections
- impact on pion samples, negligible for protons (due to event selection by Cherenkov)
- overestimation of reconstructed energy
- underestimation of mean shower radius and longitudinal CoG

Pion contamination of proton samples

- sample purity varies significantly
- overestimation of reconstructed energy
- underestimation of mean shower radius and longitudinal CoG

Beam momentum GeV/c	Purity of proton sample η
10	0.74 ± 0.17
15	0.80 ± 0.11
30	0.95 ± 0.01
40	0.84 ± 0.07
50	0.78 ± 0.08
60	0.86 ± 0.06
80	0.83 ± 0.04

Correction for bias

Mean observables

$$A_{\mathrm{corr}} = \frac{1}{\eta} A_{\mathrm{meas}} + \left(1 - \frac{1}{\eta}\right) A_{\mathrm{cont}},$$

η - estimated sample purity from contaminating admixture
A_{meas} - mean of the measured (mixed) sample
A_{cont} - mean for contaminating admixture

Shower profiles

Correction applied layer-by-layer
The corrected content for *i*-th layer (bin):

$$E_i^{\mathrm{corr}} = \frac{1}{\eta} E_i^{\mathrm{meas}} + \left(1 - \frac{1}{\eta}\right) E_i^{\mathrm{cont}}$$

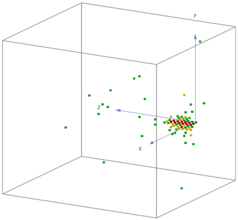
E_i^{meas} - from *i*-th bin in the measured (mixed) sample
E_i^{cont} - from *i*-th bin in the sample of contaminating admixture

Uncertainties

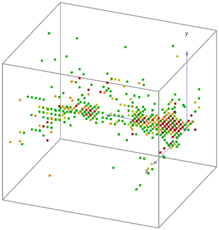
The uncertainties of corrected values are calculated using error propagation technique and taking in account the uncertainties of involved means and purities.

W AHCAL analysis

Positron at 15 GeV



Pion at 80 GeV



Data and simulation

- 2011 CERN SPS data 15 – 150 GeV
- positrons, pions, protons, kaons
- global and differential observables
- Geant 4 v9.6p2, QGSP_BERT_HP, FTFP_BERT_HP

Sources of uncertainties

Data

- Saturation correction scale factors
- MIP calibration constants
- Detector stability
- Shower start

Simulation

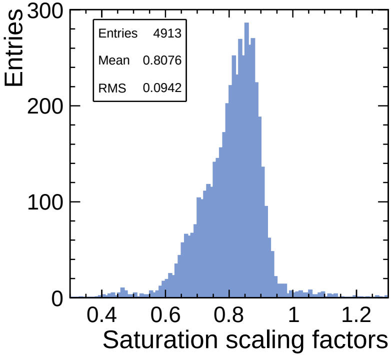
- Optical cross-talk between tiles

Negligible

- Particle type purities
- temperature corrections
- Noise
- timing cuts

Saturation scaling

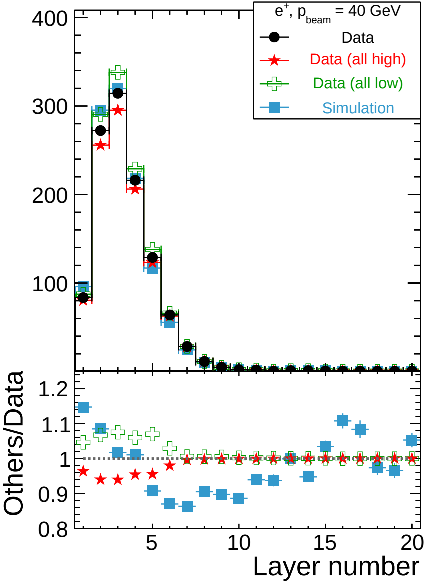
- Measured saturation scaling factors
  - Measured for 60% of all cells
  - Mean = 0.8
  - RMS = 0.09 due to device-to-device spread and measurement uncertainties
- Estimation of systematic uncertainty
  - Vary individual values by $\pm 0.5 \cdot \mathrm{RMS} = \pm 0.045$
  - Vary other values by $\pm \mathrm{RMS} = \pm 0.09$



→ Systematic uncertainty on E_{sum}

$$\rightarrow \frac{\sigma}{E} = \frac{\sum_{\mathrm{layer}=1}^{\mathrm{nLayer}} (E_{\mathrm{layer,\ default}} - E_{\mathrm{layer,\ modified}})^2}{E_{\mathrm{sum}}}$$

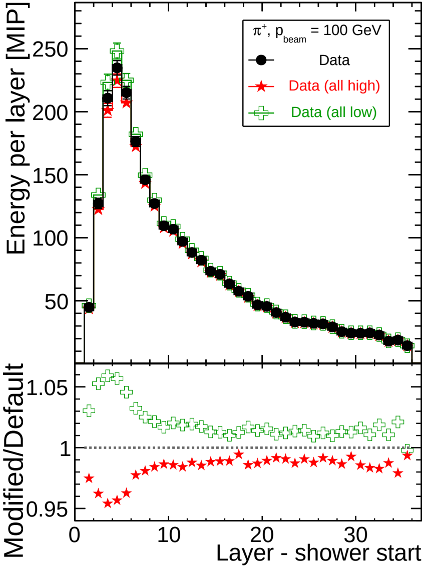
Positrons at 40 GeV



- **Data**
  - Individual value if available
  - 0.8 for all other cells
- Data (all high)
  - Individual value $+0.5 \cdot \text{RMS}$ if available
  - $0.8 + \text{RMS}$ for all other cells
- Data (all low)
  - Individual value $-0.5 \cdot \text{RMS}$ if available
  - $0.8 - \text{RMS}$ for all other cells

- Systematic uncertainty on energy sum
  $\sigma_{\text{high}} / E_{\text{sum}} = 2.46\%$
  $\sigma_{\text{low}} / E_{\text{sum}} = 3.01\%$
- Envelope of green and red, divided by $\sqrt{12}$ as systematic uncertainty for longitudinal profile; highest difference per layer: -6.07%, +7.52%

Pions at 100 GeV



- **Data**
  - Individual value if available
  - 0.8 for all other cells
- Data (all high)
  - Individual value $+0.5 \cdot \text{RMS}$ if available
  - $0.8 + \text{RMS}$ for all other cells
- Data (all low)
  - Individual value $-0.5 \cdot \text{RMS}$ if available
  - $0.8 - \text{RMS}$ for all other cells

- Systematic uncertainty on energy sum
  $\sigma_{\text{high}}/E_{\text{sum}} = 0.70\%$
  $\sigma_{\text{low}}/E_{\text{sum}} = 0.91\%$
- Envelope of green and red, divided by $\sqrt{12}$ as systematic uncertainty for longitudinal profile; highest difference per layer: -4.59%, +5.91%

Systematic uncertainty of $\langle E_{\text{visible}} \rangle$

Positrons:

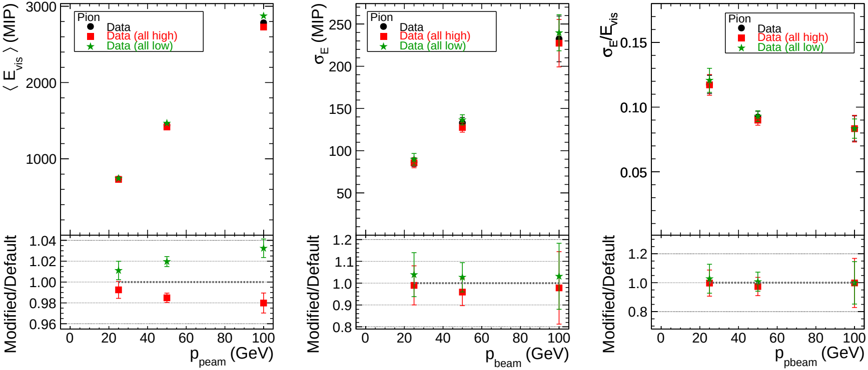
p_{beam} (GeV)	$s_{\text{indi}} = s_{\text{indi}}$ $s_{\text{others}} = 0.8$	$s_{\text{indi}} = s_{\text{indi}} - 0.045$ $s_{\text{others}} = 0.71$	$s_{\text{indi}} = s_{\text{indi}} + 0.045$ $s_{\text{others}} = 0.89$
15	434 MIPs	$(+)1.44\%$	$(-)1.12\%$
20	578 MIPs	$(+)1.92\%$	$(-)1.46\%$
30	867 MIPs	$(+)2.65\%$	$(-)2.07\%$
40	1147 MIPs	$(+)3.01\%$	$(-)2.46\%$

Pions:

p_{beam} (GeV)	$s_{\text{indi}} = s_{\text{indi}}$ $s_{\text{others}} = 0.8$	$s_{\text{indi}} = s_{\text{indi}} - 0.045$ $s_{\text{others}} = 0.71$	$s_{\text{indi}} = s_{\text{indi}} + 0.045$ $s_{\text{others}} = 0.89$
25	726 MIPs	$(+)0.35\%$	$(-)0.28\%$
50	1423 MIPs	$(+)0.60\%$	$(-)0.47\%$
100	2755 MIPs	$(+)0.91\%$	$(-)0.70\%$

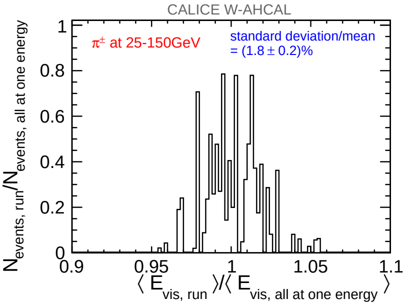
Pion energy resolution

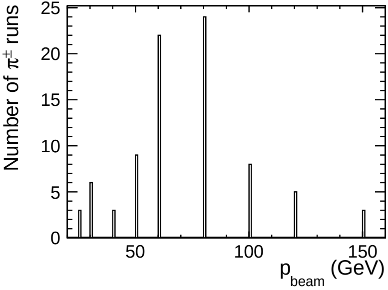
- Test full impact of modification of individual scaling on energy resolution



- While E_{vis} and σ are effected by the modification, the effect partially cancels out in the relative energy resolution for pions
- Values of energy resolution agree within uncertainties
  - 100 GeV: $(-0.10 \pm 14.68)\%$, $(-0.17 \pm 16.98)\%$

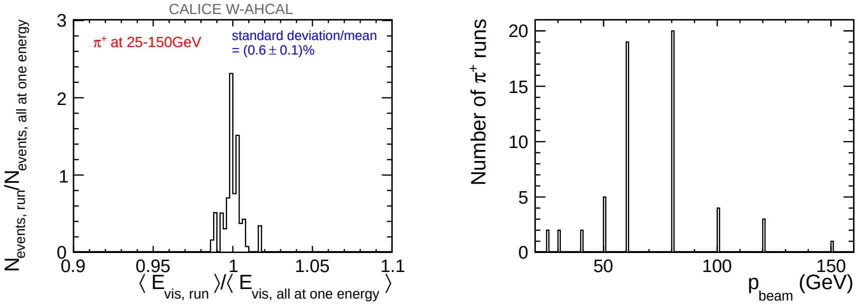
Detector instability all beam momenta: $\pi^{\pm}$





- All energies have more than 1 run

Detector instability all beam momenta: π^+



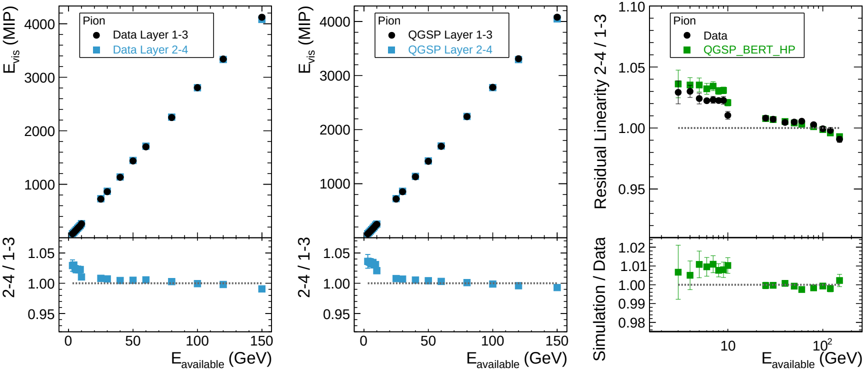
- Ignore energies with only one 1 run in left figure (i.e. 150 GeV)

Shower start

- Estimation of systematic uncertainty of shower start selection
- Compare results when selecting events with
  - Layer 1–3
  - Layer 2–4
- Build residuals of observables for 2–4 / 1–3
- Compare residuals of results between data and simulations

- Average difference between residuals of data and simulation is the systematic uncertainty of the shower start selection used in the comparison of data and simulation
- The systematic uncertainty is applied to data only

W-AHCAL Linearity



- Dependence similar for data and simulations
- Residuals differ in average (weighted by uncertainties) by 0.9% at PS energies and by 0.1% at SPS energies

Systematic uncertainty of shower start selection

Based on difference between effect on data and MC

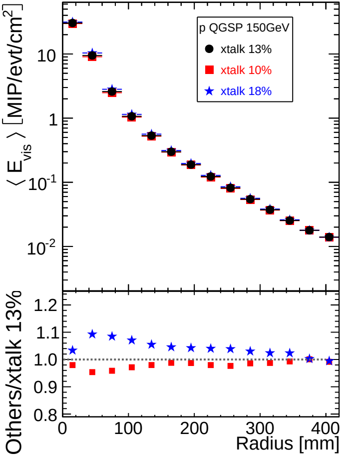
Observable	Systematic uncertainty for $\pi^{\pm}, K^{\pm}, p$	
	PS energies	SPS energies
E_{vis}	0.9%	0.1%
$\sigma_E / E_{\text{vis}}$	1.1%	2.1%
z_{cog}	-	0.1%
R	-	0.3%
E/layer @ 25 GeV	-	0.8%
E/layer @ 100 GeV	-	1.5%
E/ring @ 25 GeV	-	0.6%
E/ring @ 100 GeV	-	0.8%

Optical cross/talk between tiles

- Estimate systematic uncertainty of MC results due to limited knowledge of cross-talk between AHCAL scintillator tiles
- Several measurements are available (for few tiles only)
  - 10%: Original ITEP measurement ▸ arXiv:1012:4343
  - 15–18%: Measurements by Clemens Günter at DESY
  - 13%: best guess based on comparisons between positive data and simulations
- Geant4 simulations
  - Cross-talk settings
    - 10%
    - 13%
    - 18%
  - Energy points
    - Positrons at 40 GeV
    - Hadrons at 150 GeV

- Difference between the cross-talk settings is a worst case scenario
- Consider distribution as rectangular function
- Take RMS of rectangular function as systematic uncertainty
- → value/$\sqrt{12}$

Radial profile



- Difference between 10% and 18% of cross-talk: up to 14.5% per ring

Summary: Systematic uncertainty due to cross-talk

Observable	– Systematic uncertainty due to cross-talk –	
	Positrons@40 GeV	Hadrons@150 GeV
$\langle E_{\mathrm{vis}} \rangle$	2.7%	2.7%
σ_E / E	0.2%	1.0%
ΔE/layer	$\leq 2.8\%$	$\leq 3.9\%$
$\langle z_{\mathrm{cog}} \rangle$	-	0.1%
ΔE/ring	-	$\leq 4.2\%$
$\langle \mathrm{radius} \rangle$	-	0.4%

Summary

- Improving adequateness of Geant 4 simulations amplifies our challenge
- sophisticated procedures developed
- often result in *reduced* uncertainties
- possibilities for more unifrom treatment in different papers should be evaluated