Off-line compensation of a SDHCAL, a Monte Carlo study

CALICE collaboration meeting
Hamburg/DESY, 19-22 March 2013

M. Chefdeville, I. Koletsou, Y. Karyotakis, CNRS/IN2P3/LAPP, Annecy

Overview

- Simulation and testbeam data
  - Simulation: geometry, digitisation, data set, ntuple content
  - Testbeam: data set, cuts
- Pure digital response, electromagnetic fraction and saturation
- Offline compensation techniques with additional thresholds
  - Chi2 minimisation: optimal weights with 2 and 3 thresholds
  - Maximum likelihood with 3 thresholds
- Application to testbeam data, a status
- Conclusion

Monte Carlo simulation

Geant4 version 4.9.5, physics list QGSP_BERT

Geometry = deep SDHCAL

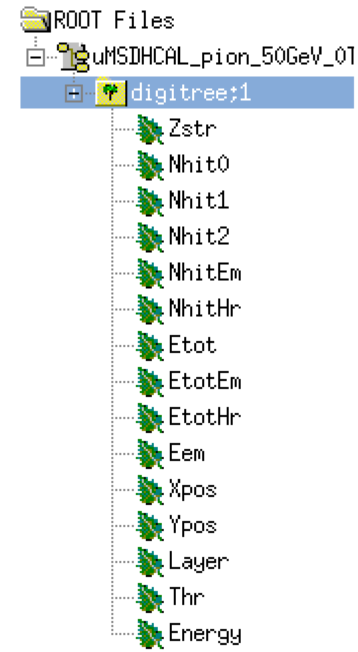
100 layers of 1x1 m2 (~ 3 mm of gas + 2 mm of steel) with 1x1 cm2 cells

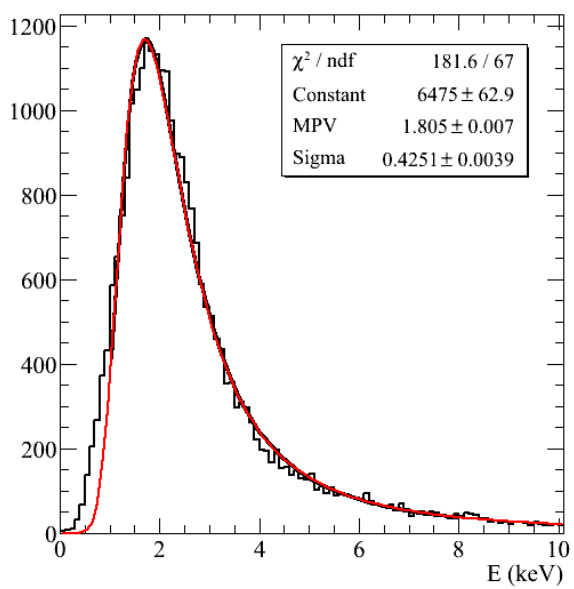
Steel absorbers, 15 mm thick

Data sets

2 times 10k pions at 5 GeV and from 10-70 GeV, every 10 GeV

First set used for the optimisation of parameters, second for estimation of performance





Shower start and electromagnetic (EM) fraction

Identify layer of first inelastic nuclear reaction

Identify neutral mesons and energy deposited through daughters

→ Store EM visible energy

Digitisation

Low threshold of 1 primary electron (~15 eV), as low as possible

Medium and high thresholds set like during testbeam : 5 and 15 MIPs

MIP determined from muon sample ~ 1.8 keV

GRPC-SDHCAL testbeam data

Period: August-September 2012, SPS/H6 line

Data set

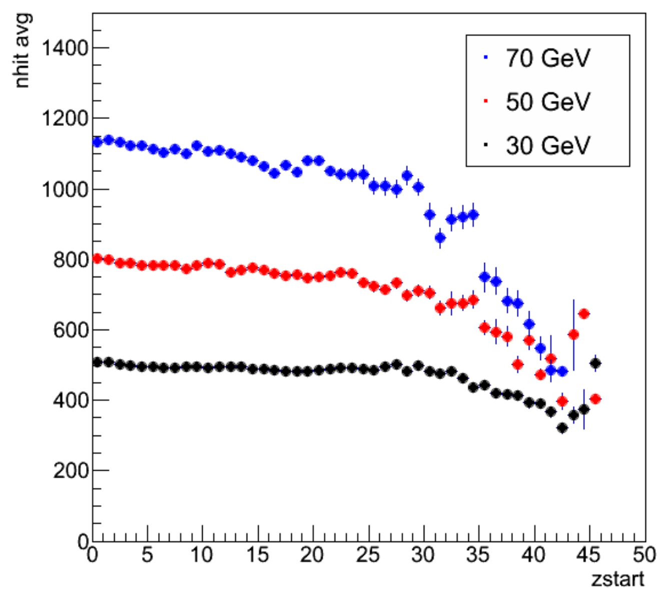
> 10k pions at 20, 25, 30, 40, 50, 60, 70, 80, 90, 100 GeV

Most of the time, more than one run per energy

Containment cut

→ Select shower starting in 12 first layers

A cut on Nhit in last layers is not allowed
as it would bias the pion sample



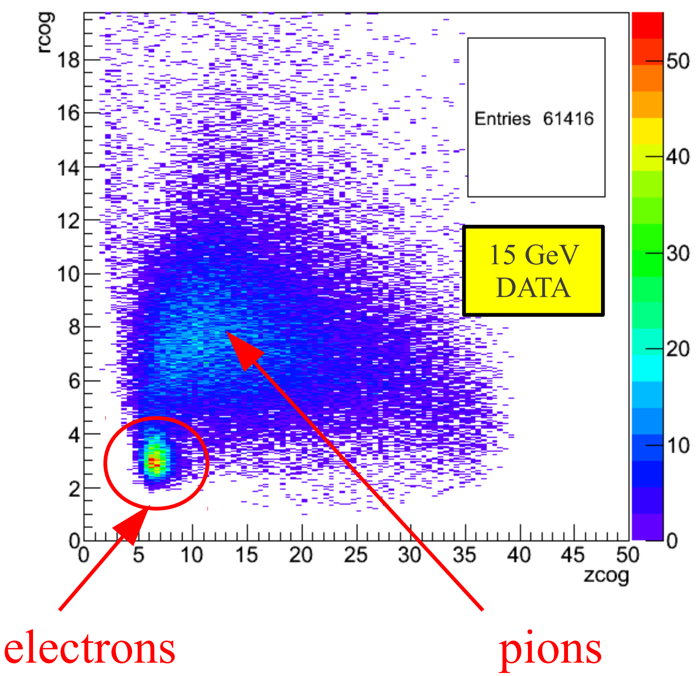
PID cut

SDHCAL is ~ compensated at low energy

→ PID e/h based on Nhit useless

→ Use transverse and longitudinal information

→ C.o.g. radial and along Z



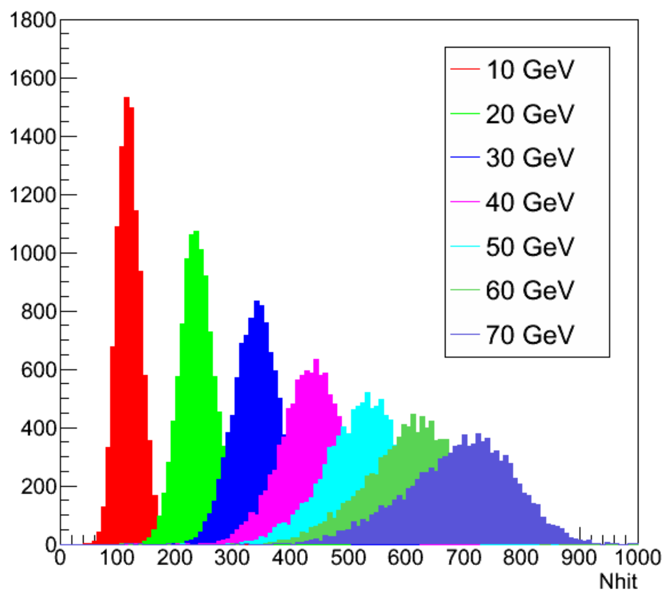
The pure digital pion response

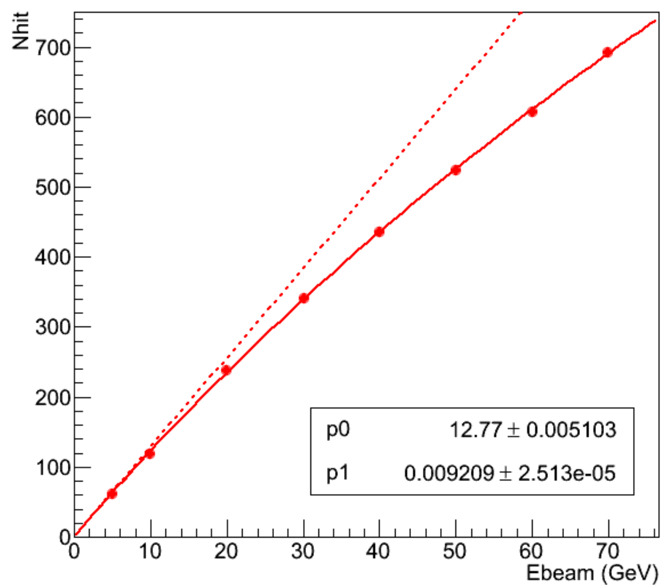
Energy reconstruction

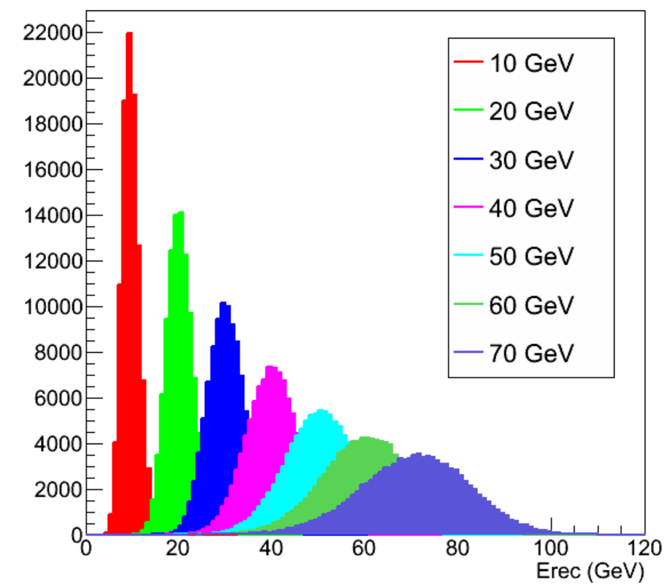
Use of the <u>inverse of the Nhit response</u> which we describe by a <u>logarithmic function of Ebeam</u>

Other parametrisation exists, this one works up to 150 GeV on testbeam data (see Micromegas talk in DHCAL session)

Nhit distributions | *Pion response in Nhit* | *Smoothed Erec distributions*







<u>**Mean and sigma of the distributions**</u> extracted in the following way (<u>no correction for tails!</u>)

Fit Novosibirsk function $f \rightarrow$ Fill empty histo h with 100000 entries from $f \rightarrow$ Get Mean and RMS of h

Electromagnetic fraction and saturation

Performance after energy reconstruction

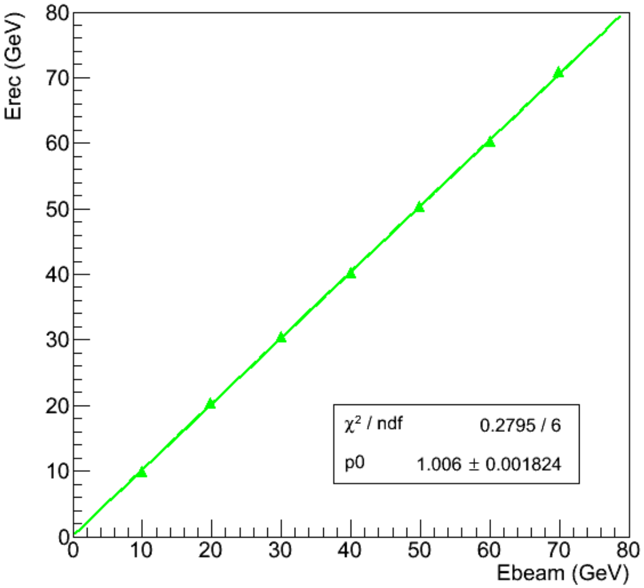
Linearity almost perfect, not surprising, we used the inverse of the response!

The energy resolution degrades above at 30 GeV as Nhit distribution develops a left-hand tail

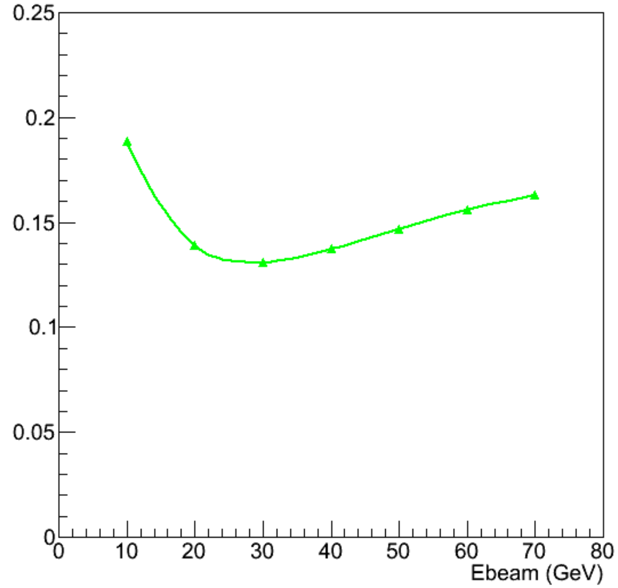
Events with significantly lower than average Nhit have a high electromagnetic energy

As expected, the EM fraction is responsible for the saturation of the DHCAL response

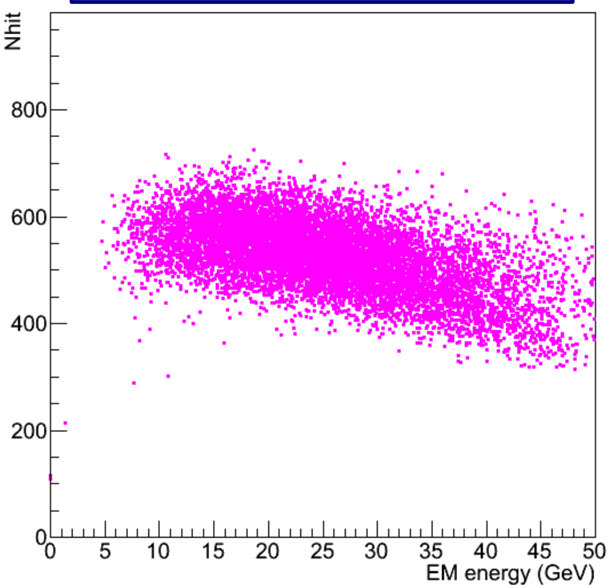
Pion response in energy



Energy resolution



Nhit VS EM energy @ 50 GeV



Offline compensation with 2 thresholds

Energy reconstruction with 2 thresholds

Define N0 as Nhit above first threshold and N1 as hit above second threshold

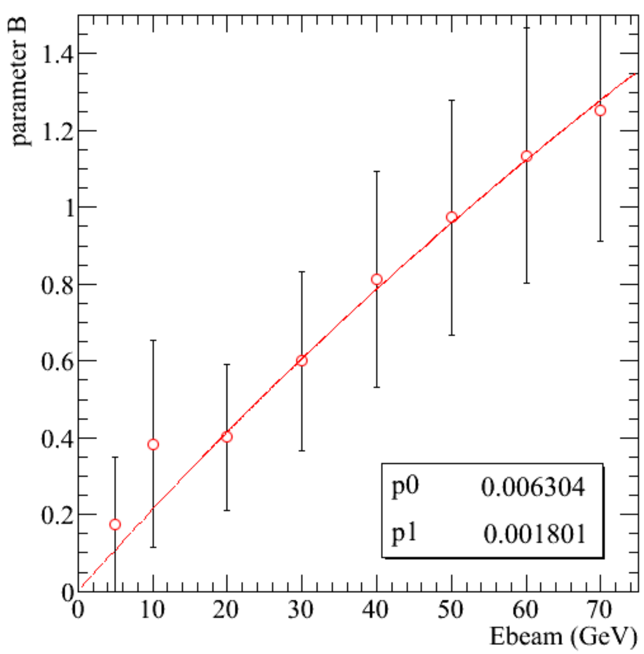
Erec = A (N0 + B.N1) with A constant in GeV/hit and B a parameter

Set A from calorimeter response → A = 1/12.77 = 0.078 GeV/hit and find B such that Erec = Ebeam

Parameter B given by (Ebeam/A – N0) /N1

Calculate B event by event and take average and RMS

Important spread (error bars = RMS) at a given energy but smooth behaviour above 10 GeV (well described by a log)



Parameter B VS Ebeam
Error bar = RMS

With thr0 ~ 0 and thr1 = 5 MIP

(N1 hits are already included in N0)

→ N1 hits "over-contribution" is 40% at 20 GeV and 125% at 70 GeV

Results with 2 thresholds (0-5) MIP

In real life, Ebeam is not known so we use a parametrisation of B versus Nhit1

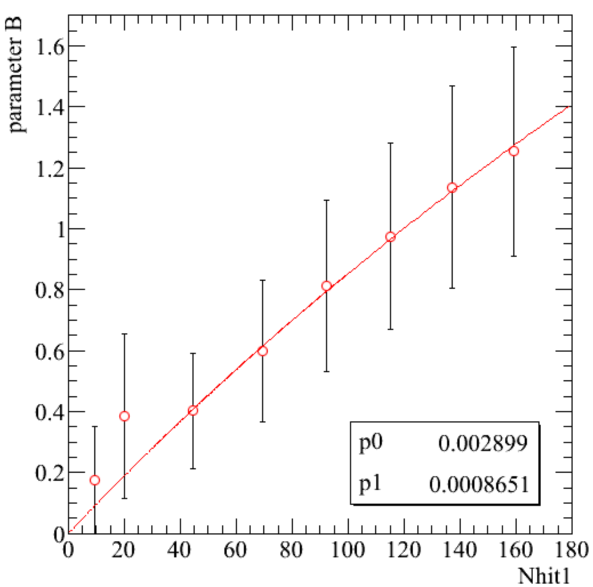
And use it to reconstruct the energy of events from the 2nd data set

Performance with thr0 ~ 0 and thr1 = 5 MIP

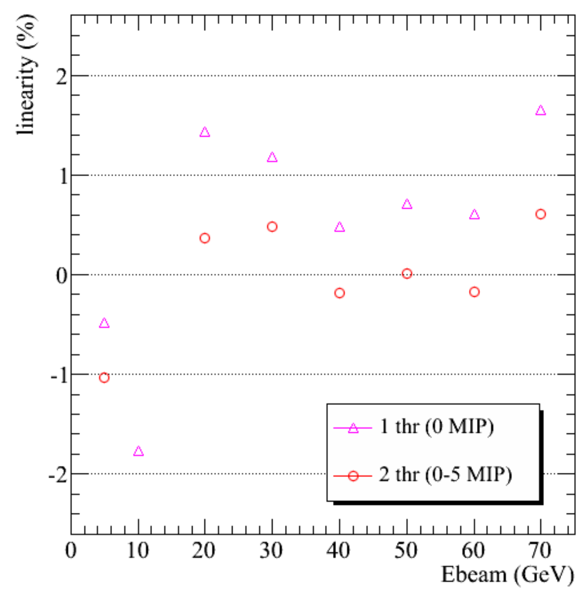
Linearity measured as (Erec – Ebeam) / Ebeam: within ± 2% (except at 10 GeV: ± 3%)

Quite an <u>improvement on resolution w.r.t. pure digital case</u> (3 terms for the fit: stochastic, constant, saturation)

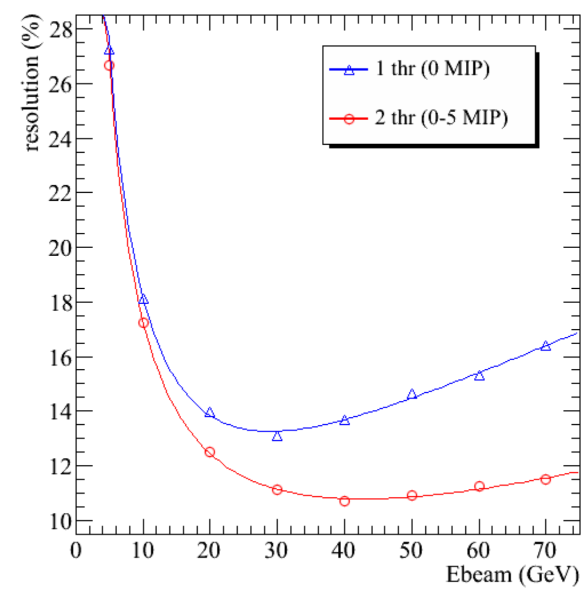
Parameter B VS Nhit1



Linearity



Resolution



After compensation, the saturation appears between 40-50 GeV → what with a higher second threshold?

Results with 2 thresholds (0-15) MIP

What has changed?

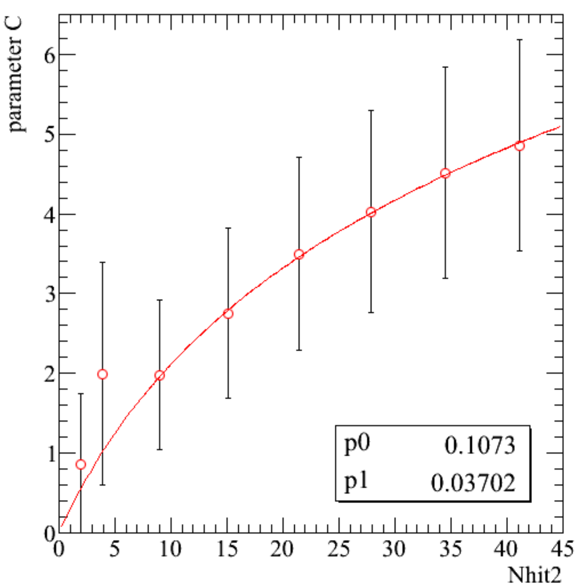
Energy reconstruction is **Erec = A (N0 + C.N2)**, A fixed and C given by (Ebeam/A – N0) /N2

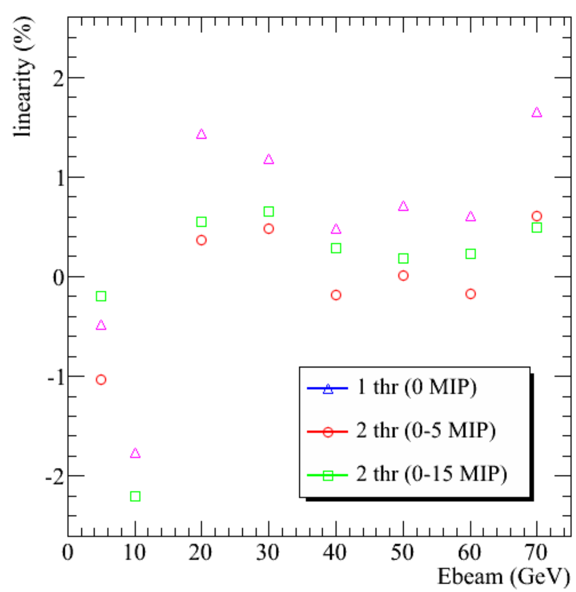
Weight of Nhit2 is now ranging from 200% at 20 GeV to 500% at 70 GeV, also well described by a log function.

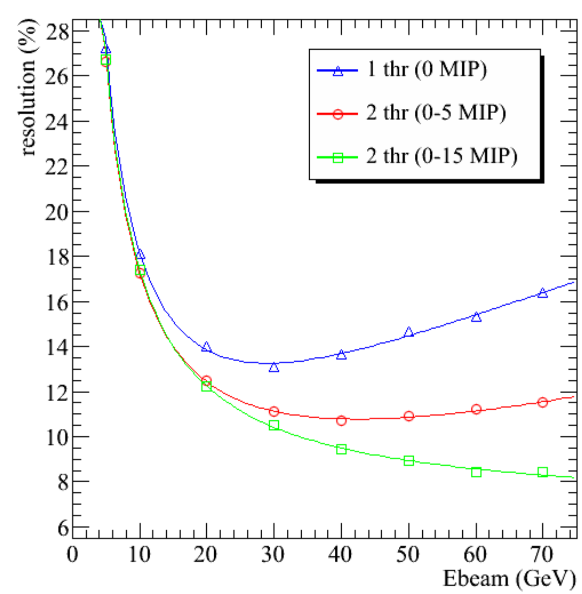
Performance with thr0 ~ 0 and thr1 = 15 MIP

Linearity measured as **(Erec – Ebeam) / Ebeam**: slightly better than in the (0-5) MIP threshold configuration

Quite some improvement of resolution: ~ 9% at 70 GeV compared to ~16% in the pure digital case







Can we gain even more with 3 thresholds?

Offline compensation with 3 thresholds

Energy reconstruction with 3 thresholds

Define N0 as Nhit above first threshold and N1 and N2 as hit above second and high thresholds

Erec = A (N0 + B.N1 + C.N2) with A constant in GeV/hit and B and C parameters

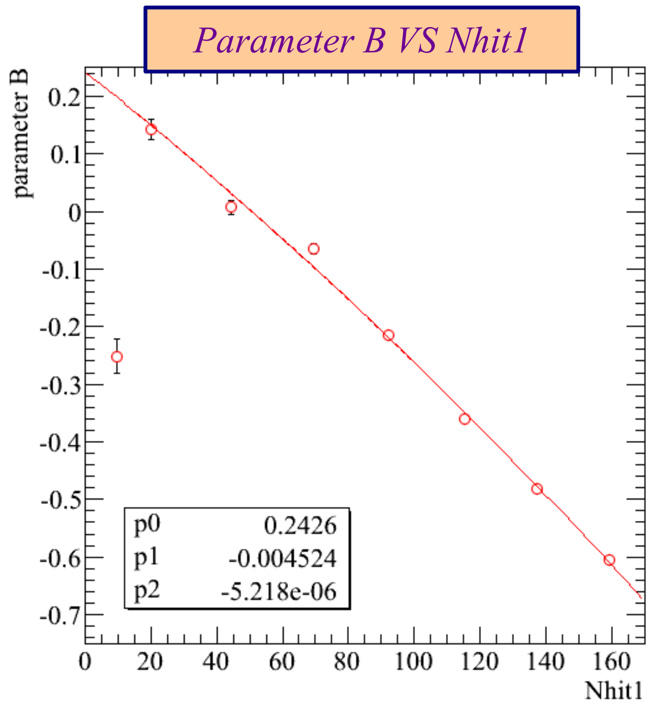
Set A from calorimeter response → A = 1/12.77 = 0.078 GeV/hit

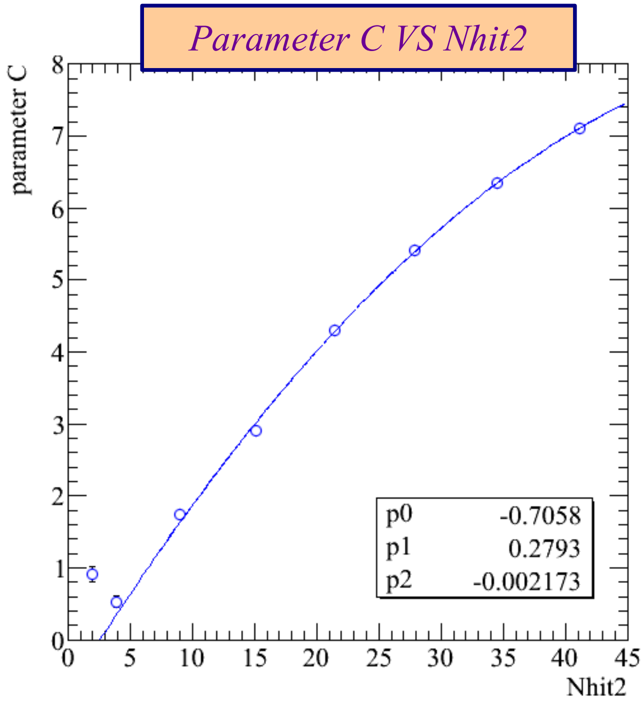
Find B and C by minimising the Chi2 = (Erec - Ebeam)2

Trends of weights

Corrections to N0 are dominated by N2 (N1 contribution actually decreases with energy)

Parametrisation by a polynomial of 2nd order (works well except for points below saturation, i.e. 5 and 10 GeV)





Results with 3 thresholds (0-5-15) MIP

Comparison (0-15) and (0-5-15) MIP

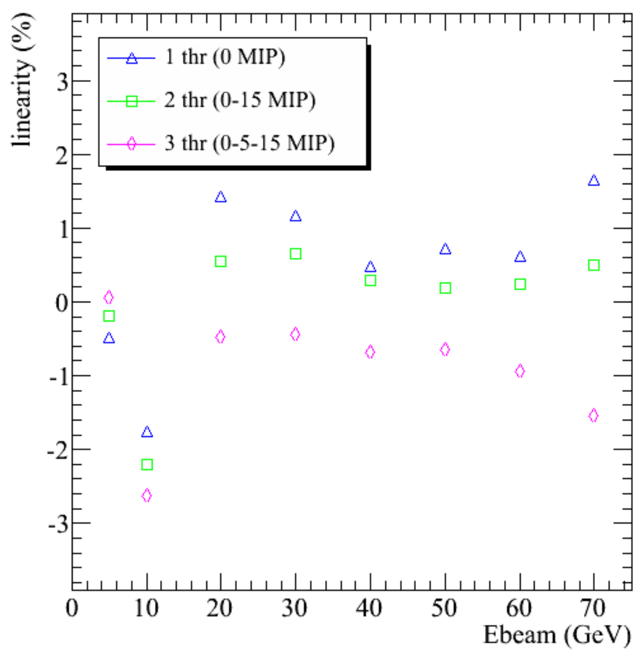
Slight improvement of the energy resolution above 30 GeV

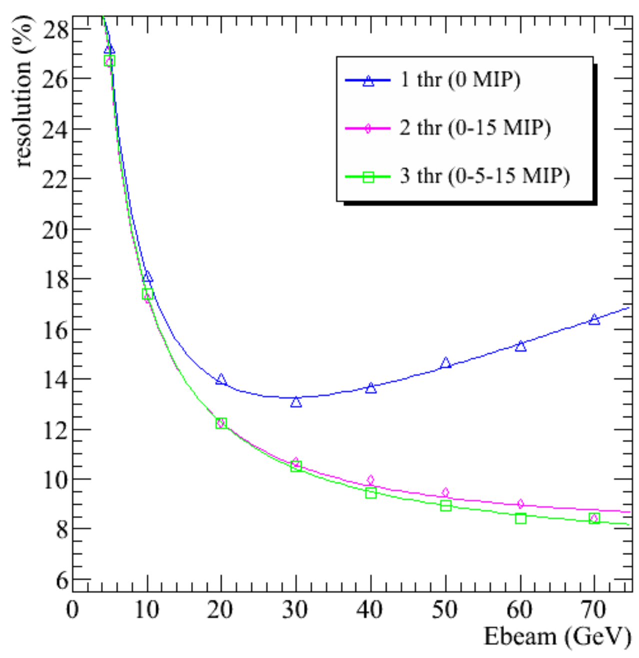
At the price of a slightly worse linearity (maybe due to the polynomial parametrisation of C(Nhit2) which is a parabola!)

With these sets of thresholds and over the studied energy range, there is not much gain in performance from 2 to 3 thresholds

→ Can we do better with a different set of medium and high thresholds? E.g. (0,15,30) MIP.

To be continued...





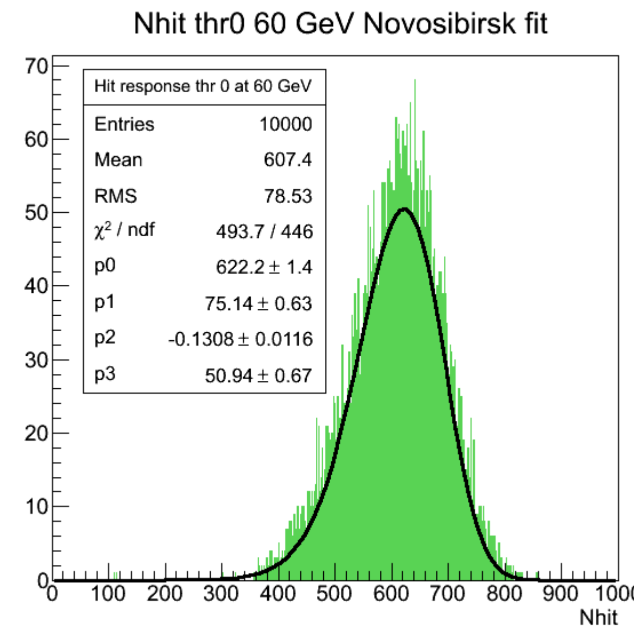
Offline compensation with a likelihood method

Maximum likelihood method

Calculate the probability to observe (N0,N1,N2) hits versus energy

The best energy estimate is the one for which the probability is maximum

Advantage: other discriminating variables can be added to the p.d.f. (hit position!)

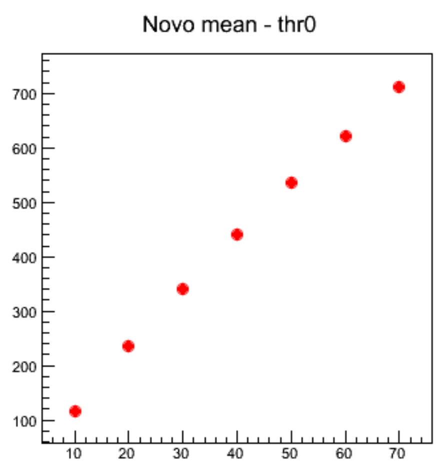


Calculation of probability, for N0 only

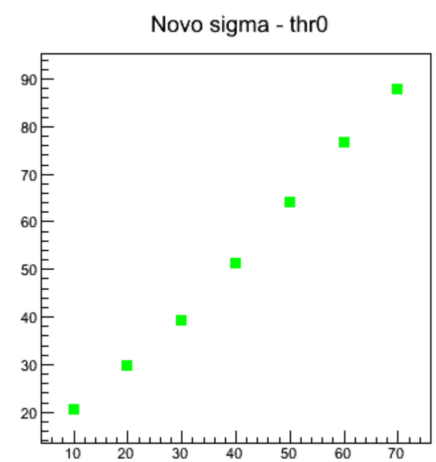
Fit the parameters of a Novosibirsk function to the Nhit distribution

Parametrise the energy dependence of fit parameters (mu,sig,tail,norm)

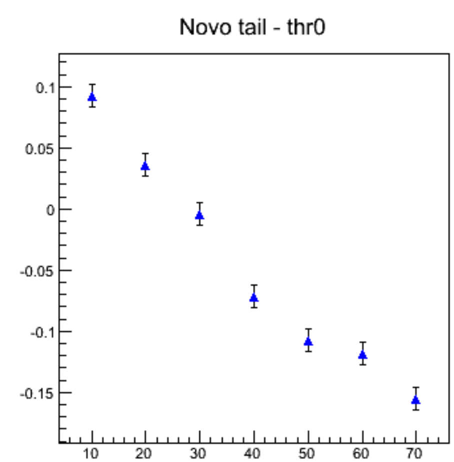
Normalised distributions → <u>p(N0,E) at any energy in the parametrisation range</u>



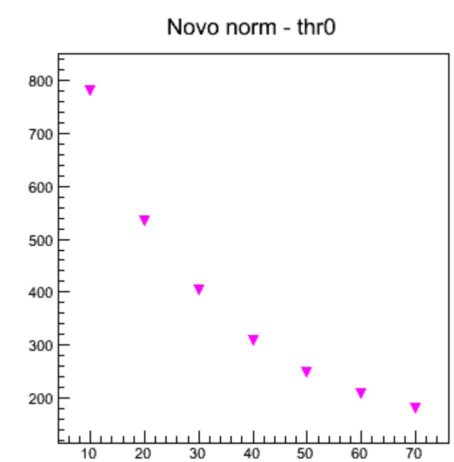
Novosibirsk mean



Novo. sigma



Novo. tail



Normalisation

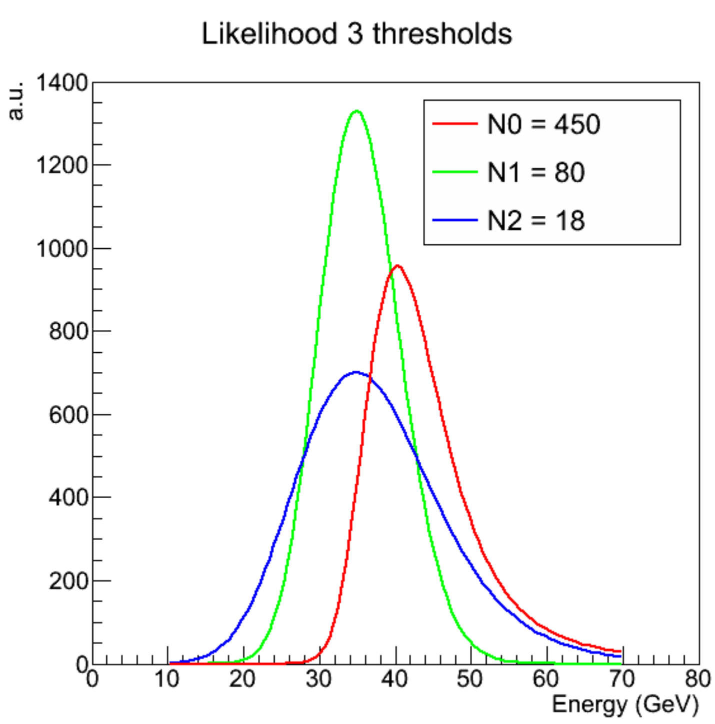
Probability distribution functions

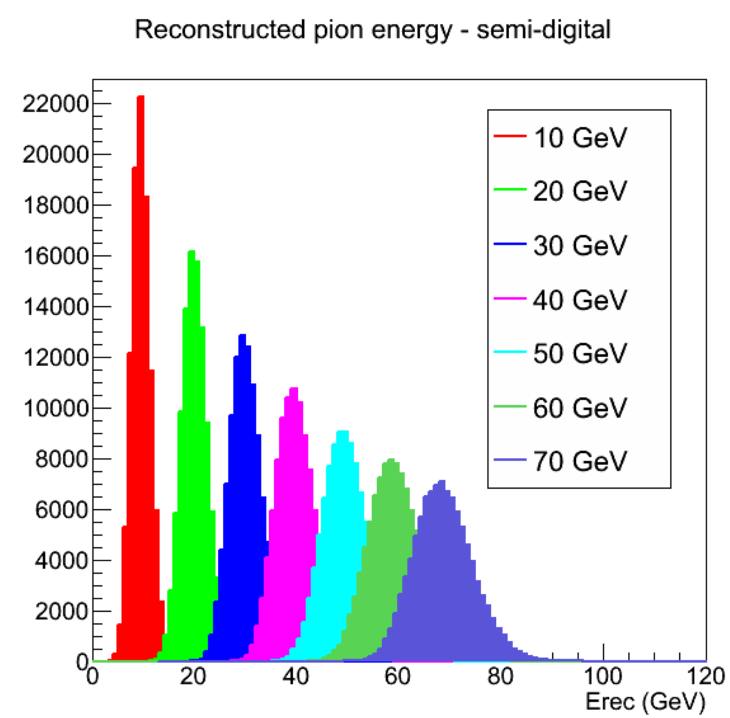
First data set (10k pions): calculation of probability of (N0,N1,N2)

A probability distribution function for (N0,N1,N2) can be built based on the individual p.d.f. for (N0), (N1) and (N2)

Second data set (10k pions): use likelihood functions

For each event, call the probability distribution function of (N0,N1,N2) and find energy at maximum





Results with a likelihood method

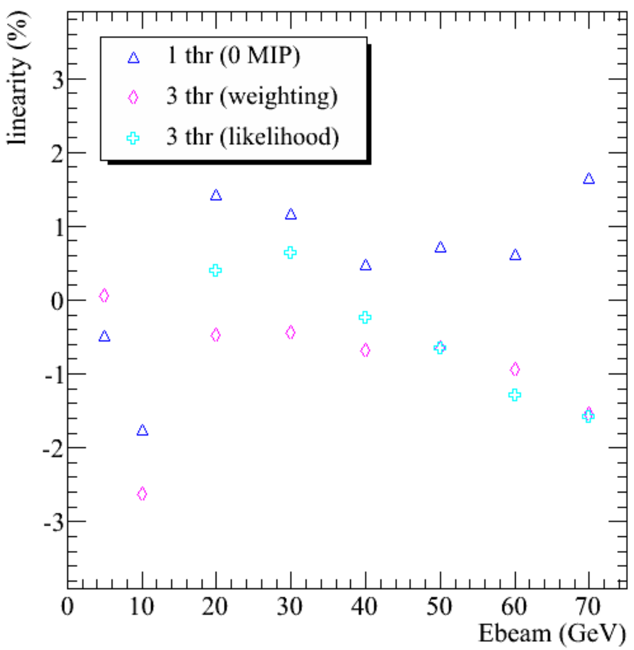
No big difference compared to a weighting method

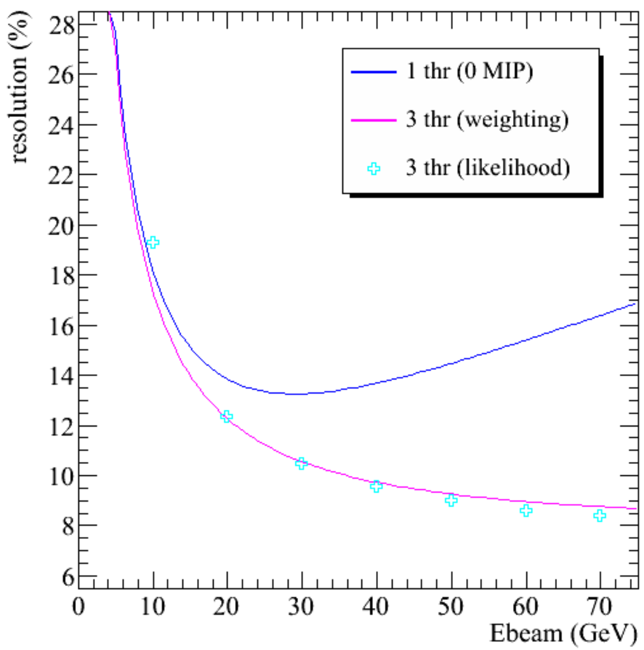
Comparable linearity than with the weighting method

Slight improvement of resolution above 50 GeV, degradation below 10 GeV

Possible improvements

Include centre of gravity of hits along shower axis in probability distribution





Application of methods to testbeam data

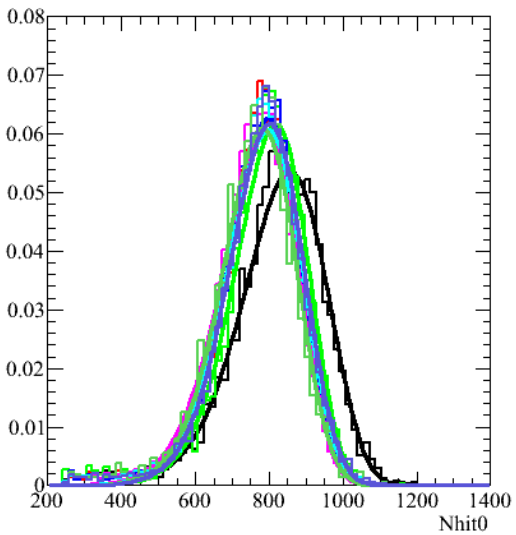
The presented compensation methods are based on the parametrisation of weights (A,B,C) and Nhit distributions (Novosibirsk parameters mu, sig, tail, norm) VS beam energy

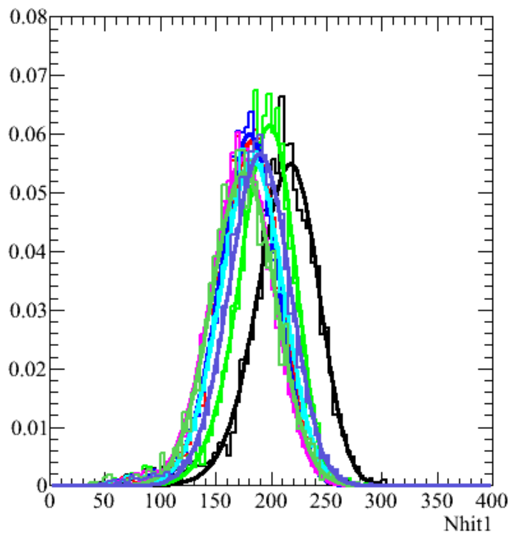
→ It is crucial that (N0,N1,N2) are smooth functions of beam energy

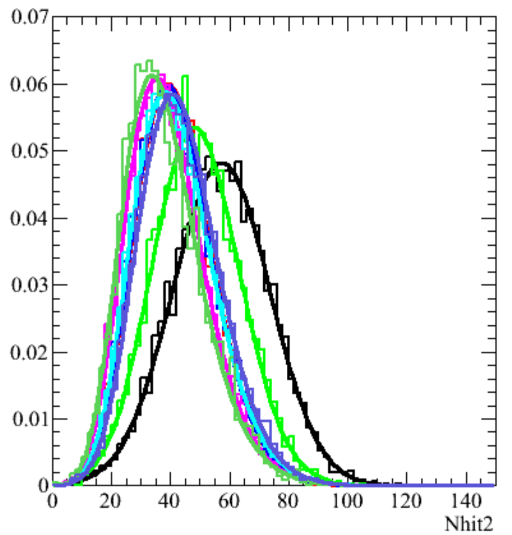
This condition is not realised for the GRPC-SDHCAL testbeam of August

Nhit distributions for muons are fairly the same from one run to the other

Nhit distributions for pions of given energy show significant spread, e.g. at 50 GeV below







Nhit0 @ 50 GeV *Nhit1 @ 50 GeV* *Nhit2 @ 50 GeV*

Application of methods to testbeam data

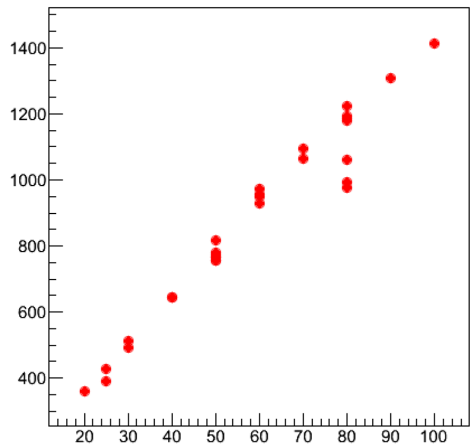
Compensation methods are based on the parametrisation of weights (A,B,C)
and Nhit distributions (Novosibirsk parameters mu, sig, tail, norm) VS beam energy

→ It is crucial that (N0,N1,N2) are smooth functions of beam energy

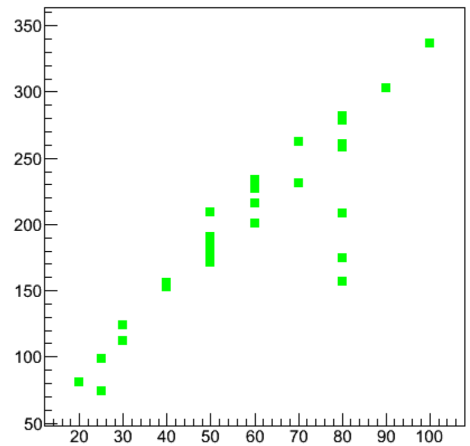
This condition is not realised for the GRPC-SDHCAL testbeam of August

Nhit distributions for muons are fairly the same from one run to the other

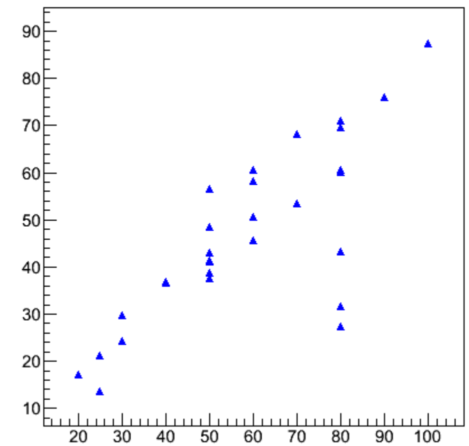
Nhit distributions for pions of given energy show significant spread, e.g. at 50 GeV below



Nhit0 VS Ebeam



Nhit1 VS Ebeam

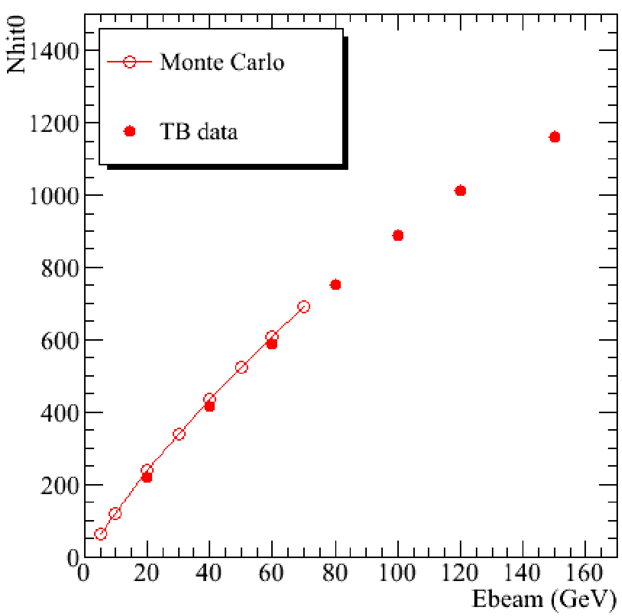


Nhit2 VS Ebeam

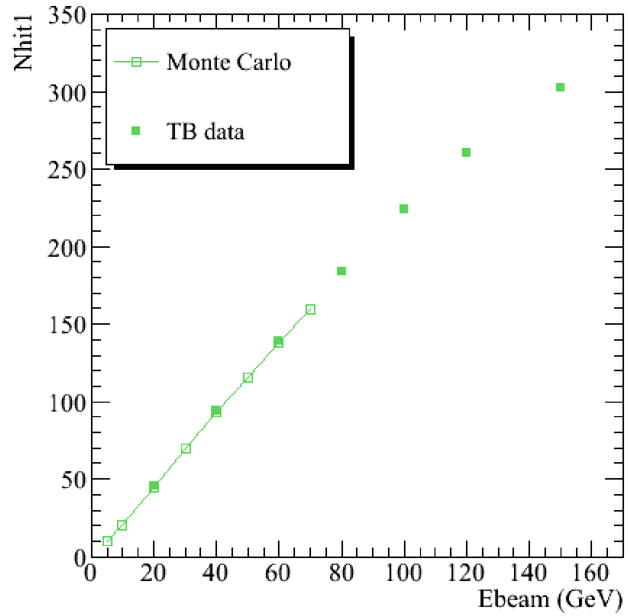
Interlude Micromegas

This is what we measured in testbeam with 4 Micromegas chambers inside SDHCAL (Nov. 12)

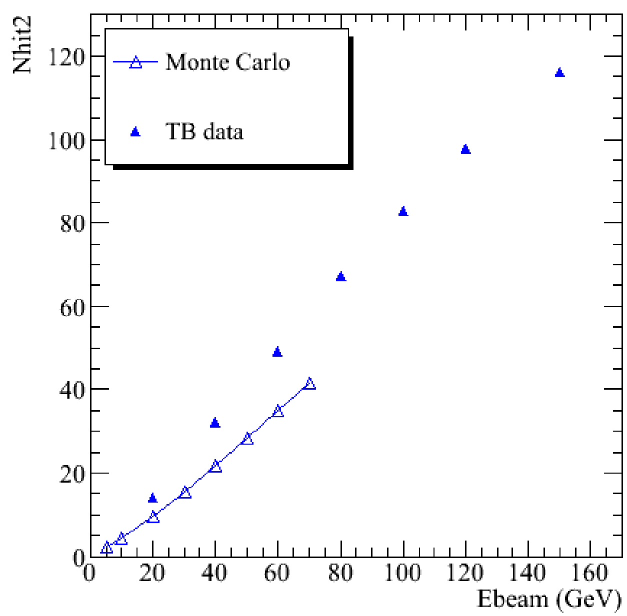
(See Micromegas SDHCAL talk tomorrow)



Nhit0 VS Ebeam



Nhit1 VS Ebeam



Nhit2 VS Ebeam

Application of methods to testbeam data

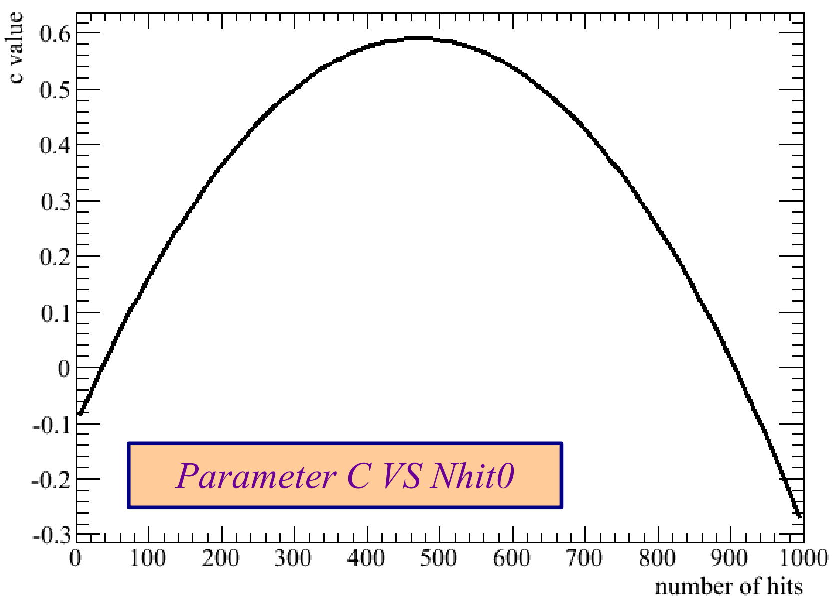
No parametrisation as previously described possible

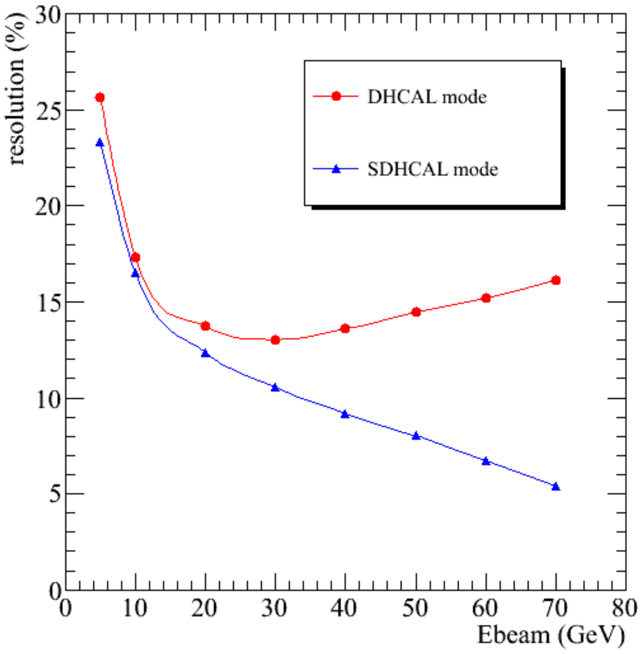
Other methods, however, are possible as the one described in the SDHCAL CALICE note 037

Erec = AN0 + BN1 + CN2 where A, B, C are 2nd order polynomial functions of N0 → 9 parameters!

The Chi2 is minimised over all energies and not energy per energy

→ this accounts for the spread but limits the improvement of the SDHCAL versus DHCAL mode





Application of this reconstruction method to Monte Carlo data gives non-physical meaning:

For instance the parameter C rises and decreases with energy. Resolution improvement questionable.

Conclusions

- A digital HCAL is expected to suffer from saturation
  - With steel absorbers and 1x1 cm2 pads,
    a Monte Carlo study indicates that this should appear at 20-30 GeV
- We are developing offline compensation methods to
  fully exploit the potential of a semi-digital readout of a hadron calorimeter
  - Weight and likelihood methods give promising results
  - On-going efforts to optimise the threshold settings
- Application to testbeam data on-going
  - SDHCAL data affected by some spread
    → more work needed to understand
  - DHCAL data may give some hints on the saturation
    → interaction with B. Bilky in Argonne