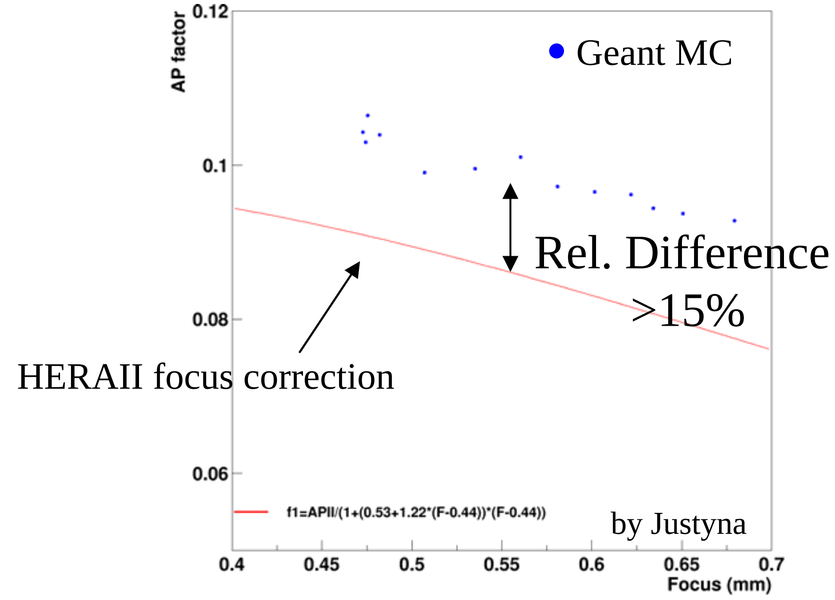
# Status of the TPOL Simulation

**TPOL Geant Simulation**
**in Comparison to Data**
**and**
**Parametrized Calorimeter Response**

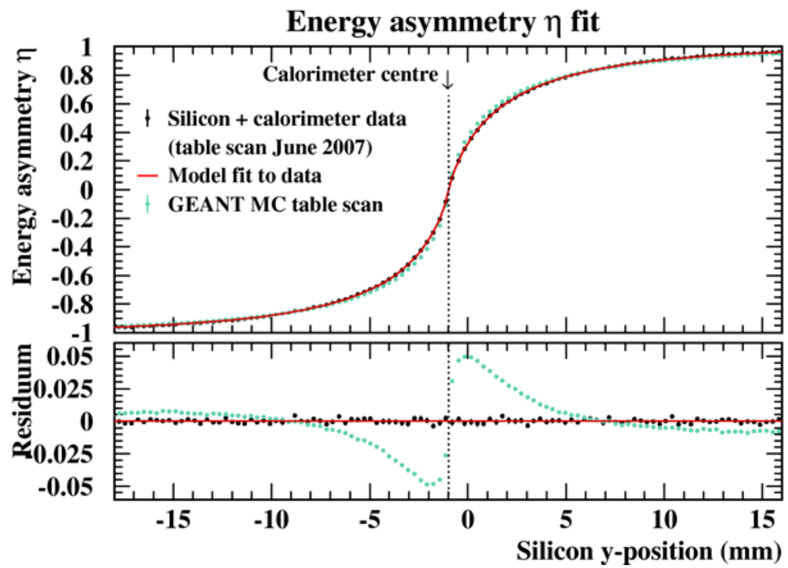
Blanka Sobloher
POL2000 meeting, 10th February 2010

# TPOL Geant Monte Carlo - Status as of first large MC prod.

- First large scale Monte Carlo production last year, which is the status as of November (PRC)
  - → Relied on the general tuning status as traditioned by several ‚tuners' over ~3years
  - → Apparently Analysing power of the shift of means method is way too far off to be correct
  - → Additionaly the energy resolution of this setup is too bad
  - → Need to tune the MC better



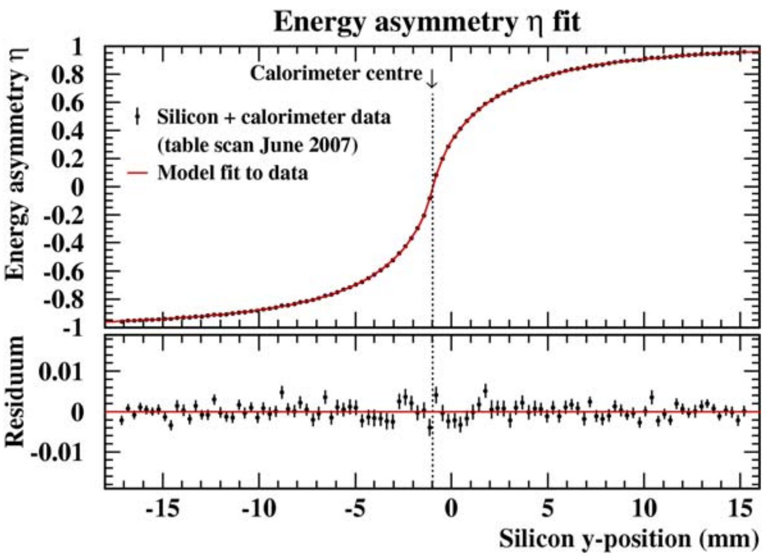

- The reason for this large difference in AP
  - → Energy asymmetry function $\eta_{UD}(y)$ (UP/DOWN)
  - → Energy resolution
  - → Beam size (emittance of beam)
  - → Calibration and centering
  - → …?

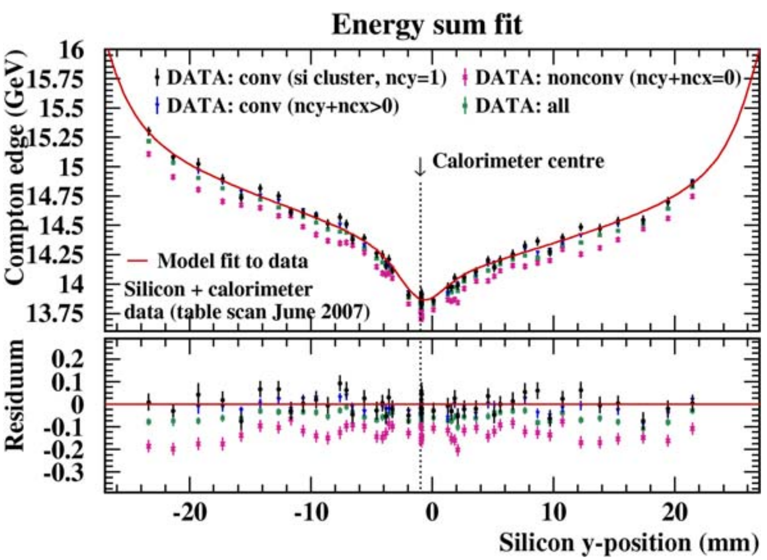



- Right: $\eta_{UD}(y)$ of converted photons of Geant MC in comparison to the measurement using the combined silicon calorimeter data including a model fit

# Silicon Calorimeter Data - Working Horse and Guinea Pig

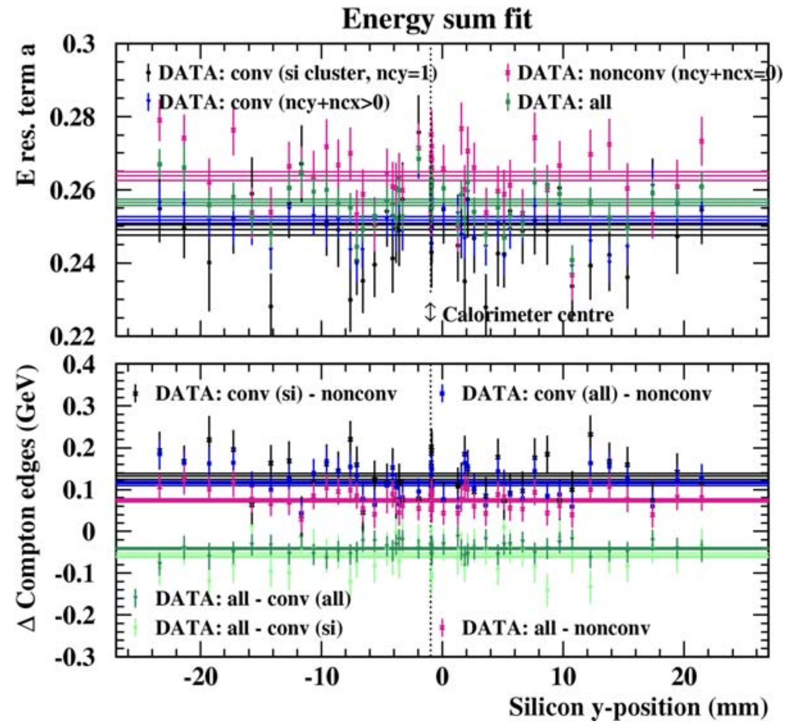
- Combined silicon calorimeter data
  - → Mainly table scans taken at the end of june 2007, but comparing to scans and other data taken throughout the HERA II running period
  - → Cutting for different photon classes using silicon clusters
    - ➢ Converted (silicon): 1 cluster in y-plane, charge>35.
    - ➢ Converted (any): any number of clusters in either x- or y-plane
    - ➢ Nonconverted: no clusters at all (discard first 1k events)
    - ➢ All: no cuts
  - → Cuts certainly not perfect (purity<100%), but already quite good, judging from comparisons with Geant
  - → Different markers in data, which the MC should follow (minimum requirement): to reproduce
    - ➢ energy resolution
    - ➢ differences of Compton edges (leakage),
    - ➢ shape of $\eta_{UD}(y)$!
    - ➢ shape of total energy $E_{UD}(y)$ (mostly detector effects)
    - ➢ possibly changing/specific behaviour of resolution and edges with y (e.g. inside/outside gap, etc)

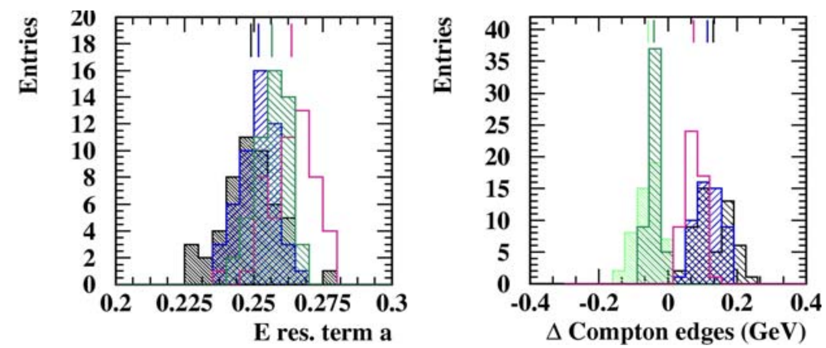


# Silicon Calorimeter Data - Working Horse and Guinea Pig

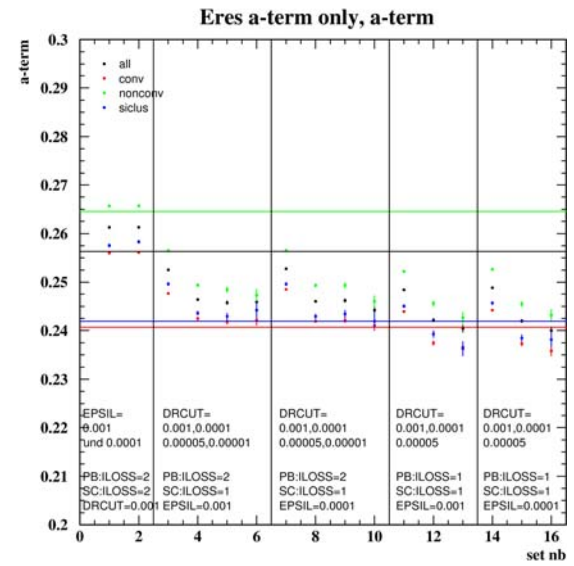
- Combined silicon calorimeter data
  - → Mainly table scans taken at the end of june 2007, but comparing to scans and other data taken throughout the HERA II running period
  - → Cutting for different photon classes using silicon clusters
    - ➢ Converted (silicon): 1 cluster in y-plane, charge>35.
    - ➢ Converted (any): any number of clusters in either x- or y-plane
    - ➢ Nonconverted: no clusters at all (discard first 1k events)
    - ➢ All: no cuts
  - → Cuts certainly not perfect (purity<100%), but already quite good, judging from comparisons with Geant
  - → Different markers in data, which the MC should follow (minimum requirement): to reproduce
    - ➢ energy resolution
    - ➢ differences of Compton edges (leakage),
    - ➢ shape of $\eta_{UD}(y)$!
    - ➢ shape of total energy $E_{UD}(y)$ (mostly detector effects)
    - ➢ possibly changing/specific behaviour of resolution and edges with y (e.g. inside/outside gap, etc)

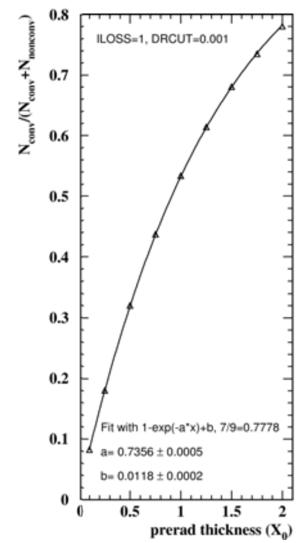


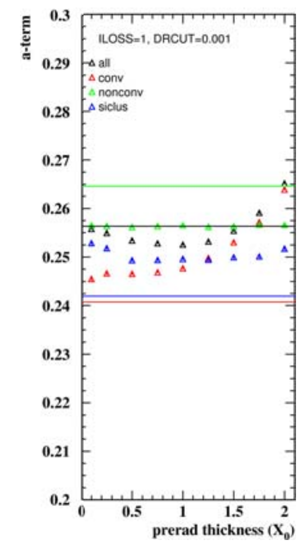



# Tuning of Geant MC - A never-ending Story

- Variations and checks to improve the response
  - → Geometry: materials, stacking (Densimet, lead frames, scintillators, air), sizes, thicknesses, position and thickness of preradiator and silicon planes
  - → Obvious things like beam size (by emittance), optics, etc…
  - → Took out any additional smearing to simulate photon statistics
  - → Performed variations with particle gun (same beam spot, fixed energies 1-30GeV):
    - ➢ Preradiator thickness
    - ➢ Gap width
    - ➢ Parameters of the simulation: ILOSS, EPSIL, DRCUT in various combinations
    - ➢ Scintillator thickness/density
    - ➢ Absorber/Lead thickness, Absorber density
    - ➢ Blinding deliberately the first scintillator layer (rad. damage)
  - → Check, if derived resolution terms differ from those obtained by fitting Compton edges in Compton setups…. Yes, mostly
- Most variations allow to change obvious things
  - → e.g. preradiator thickness ↔ conversion factor
  - → But most worsen the resolution (like too thick prerad, too large gap,etc)
  - → But most of them do not influence the relative behaviour of leakage of photon classes, i.e. the difference between the Compton edges
  - → Most of them do not help in improving the resolution or the distance of the resolution terms of both classes…

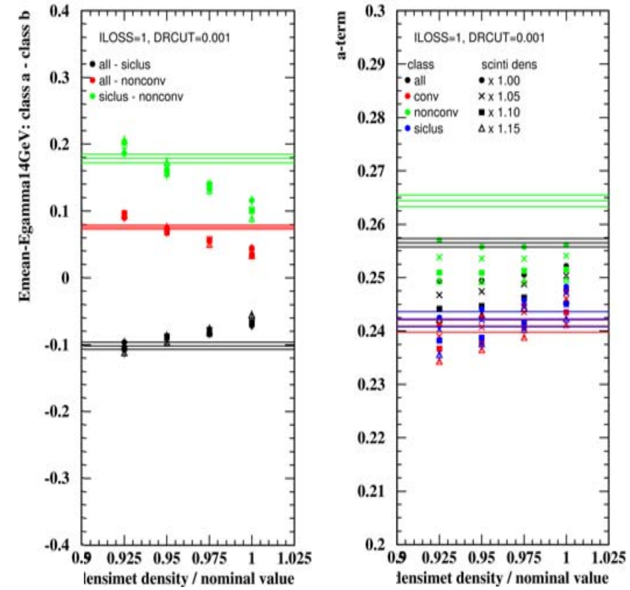


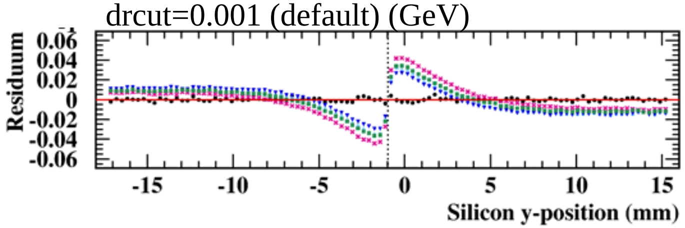
# Tuning of Geant MC - Variations and Checks

- Variations and checks to improve the response
  - → Found only two promising candidates
    - ➢ densities of the materials
    - ➢ DRCUT
  - → Decreasing the density of the absorber (Densimet only)
    - ➢ distance between the Compton edges rises as well as the distance in resolution terms
  - → Increasing the scintillator density
    - ➢ almost nothing except the resolution changes, and it improves with larger density
  - → OK, this does ot change the $\eta_{UD}(y)$
    - ➢ shower form stays mainly the same
    - ➢ Is not really justified judging from the calorimeter inspection
  - → Changing the energy cut driving the threshold for gamma and electron/positron transport
    - ➢ changes the long shower component (halo), but not the inner one (core)
    - ➢ core is driven by multiple Coulomb scattering, no parameters free for this…
    - ➢ improves resolution

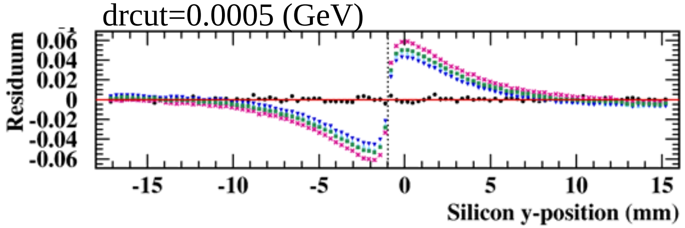



$\eta_{UD}(y)$ Residua to fit

drcut=0.001 (default) (GeV)

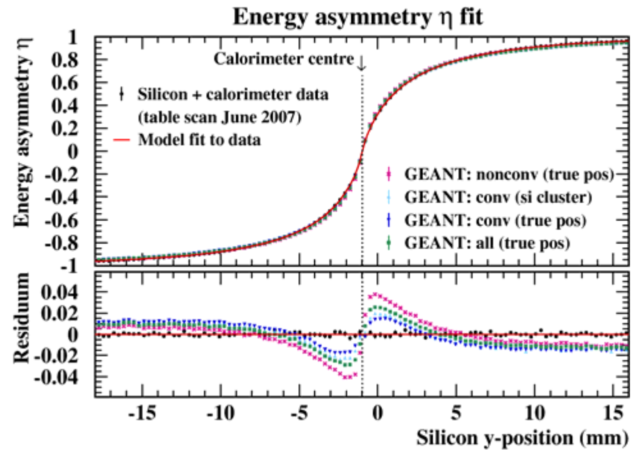


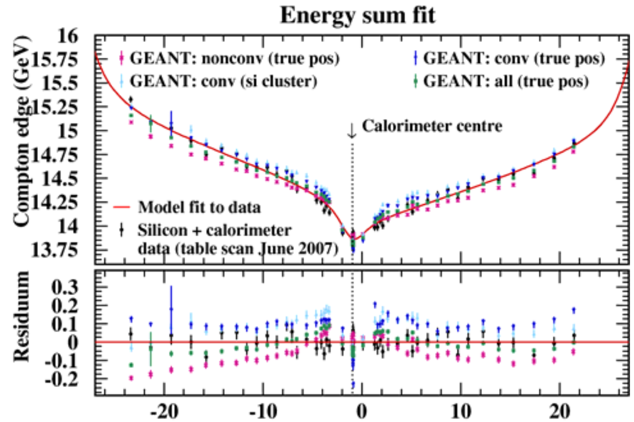

drcut=0.0005 (GeV)

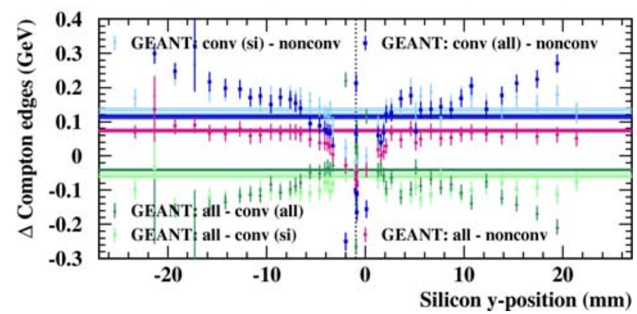


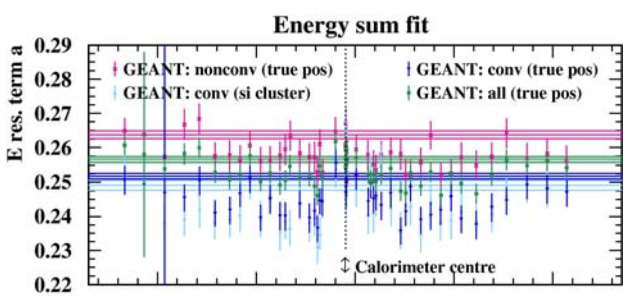
# Tuning of Geant MC - Variations and Checks

- Variations and checks to improve the response
  - → What about blinding of the first scintillator layer?
- The first layer sees a very small/dense spray of particles with high energies
  - → In that region the scintillators are more likely to be damaged than in other regions or layers deeper inside of the calorimeter
- Radiation damage would make the scintillator yellow
  - → decrease of transmission of scintillator light
  - → decrease of generation of scintillator light
- Simple model: add an inefficiency, gaussian shape with approximate beam sizes in the center of the first layer
  - → simulates the loss of generated and transmitted light output of that region
  - → does not affect the transmission of light through this region but generated at different places
  - → noticeable improvement in $\eta_{UD}(y)$, but behaviour of edges and resolution in blinded region contradicts data!
  - →Not the right track!



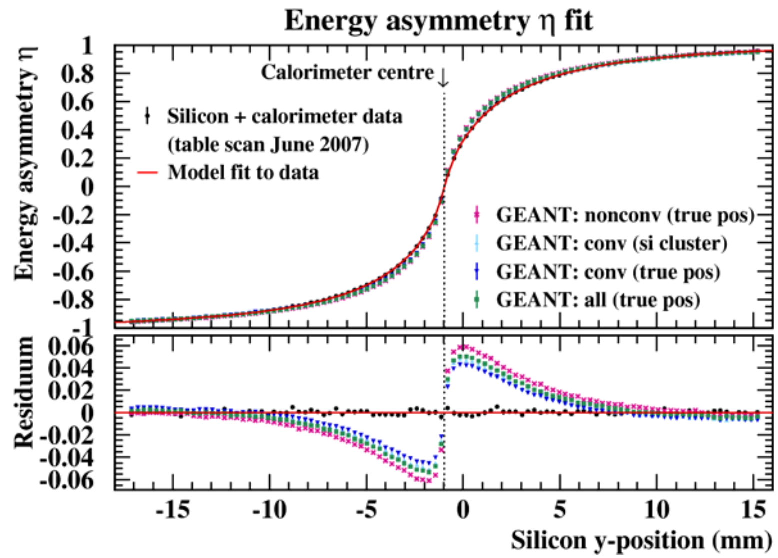



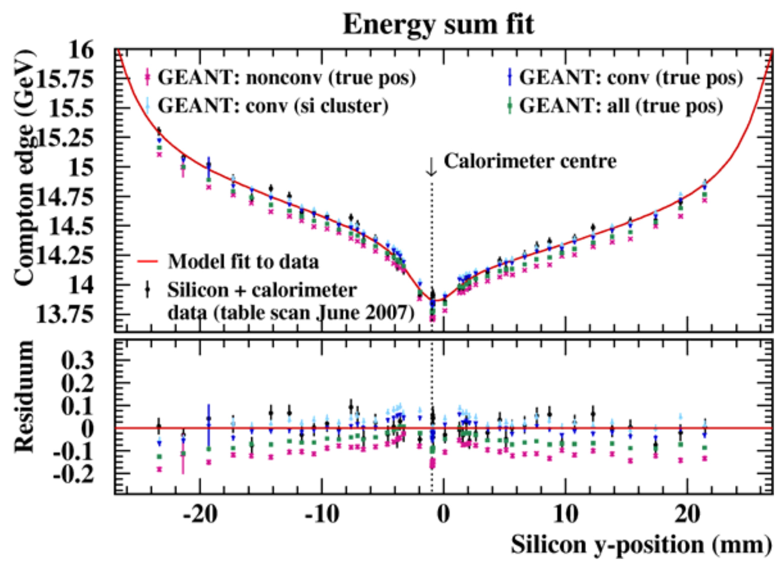



# Tuning of Geant MC - Summary

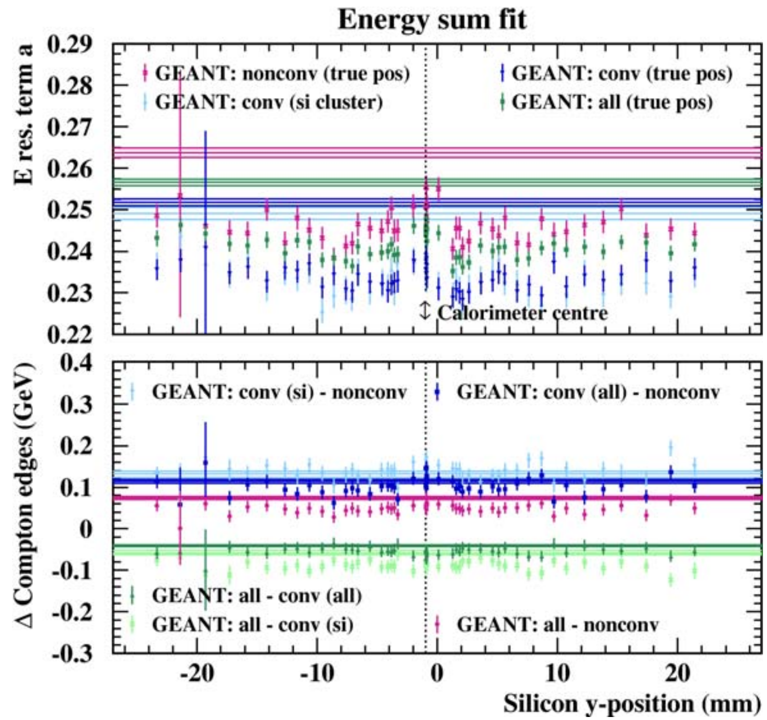
- What has been learned from these variations?
  - → Tuning Geant is VERY consuming: (wo)man time(!), CPU time, disk space, my and my colleagues patience… - to be compared to improvements achieved
- I.e. 8 weeks later, ~1600 ntuples, ~15000 CPUh and ~1TB disk space
  - → FLC has used currently 96% resource share on the BIRD farm (should be only about 6%...) – we are sharing with Atlas, Opera, Astro, IT, FLA, Theory,…
    - ➢ Current Malus sets a natural end to these studies
- Quasi final Geant setup
  - → The best of all variations combined
  - → Energy resolution better, leaves room to add some photoastatistics
  - → Edge differences better than before
  - → Didn‘t change much according to $\eta_{UD}(y)$, still wiggly
  - → Currently two versions of parameter sets
    - ➢ either having more wiggles with smaller residua and worse energy shape
    - ➢ or shower has at least the desired halo length, then energy shape better, absolute differences are smaller, but $\eta_{UD}(y)$ residuals in the center are larger
  - → Didn‘t change much, so analysing power from that MC should still be a problem (didn‘t check that explicitly due to time/computing power/etc)

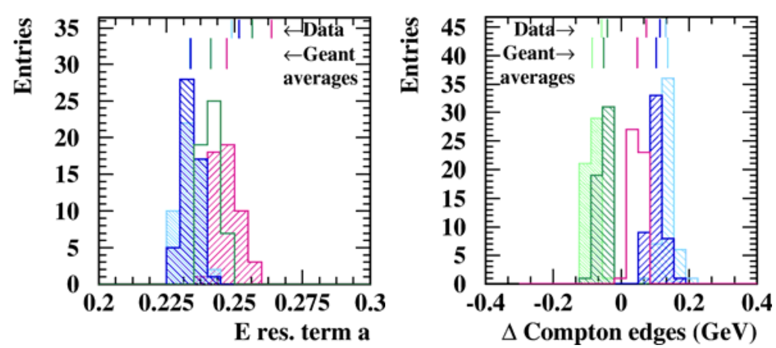


# Tuning of Geant MC - Summary

- What has been learned from these variations?
  - → Tuning Geant is VERY consuming: (wo)man time(!), CPU time, disk space, my and my colleagues patience… - to be compared to improvements achieved
- I.e. 8 weeks later, ~1600 ntuples, ~15000 CPUh and ~1TB disk space
  - → FLC has used currently 96% resource share on the BIRD farm (should be only about 6%...) – we are sharing with Atlas, Opera, Astro, IT, FLA, Theory,…
    - ➢ Current Malus sets a natural end to these studies
- Quasi final Geant setup
  - → The best of all variations combined
  - → Energy resolution better, leaves room to add some photoastatistics
  - → Edge differences better than before
  - → Didn‘t change much according to $\eta_{UD}(y)$, still wiggly
  - → Currently two versions of parameter sets
    - ➢ either having more wiggles with smaller residua and worse energy shape
    - ➢ or shower has at least the desired halo length, then energy shape better, absolute differences are smaller, but $\eta_{UD}(y)$ residuals in the center are larger
  - → Didn‘t change much, so analysing power from that MC should still be a problem (didn‘t check that explicitly due to time/computing power/etc)

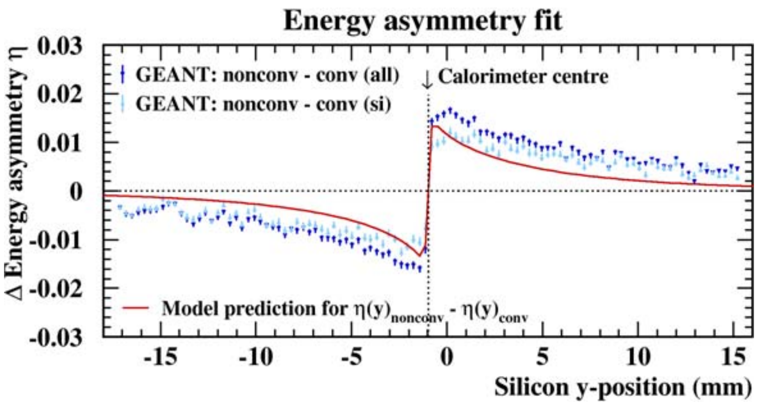


# Parametrized Calorimeter Response - An Offer

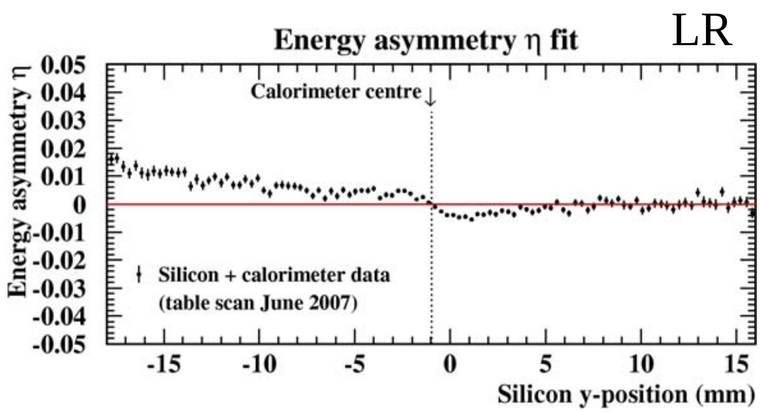
- A parametrized modeling of the calorimeter response
  - → Implemented into the MC (in cvs), block ‚calopara‘
  - → Incorporates the extensive model used to fit the silicon calorimeter data
  - → Can be run in addition to the Geant MC (takes conversion decision from Geant), and allows to switch off the Geant calorimeter part (makes it some 2-4 orders of magnitude faster, and allows thus for much higher statistics)
    - ➢ Event generation rate on standard computers as of today: 1.3k evts/min with higher drcut, some 600 evts/min for smaller drcut
    - ➢ Compare this to a 3% statistical error if we generate ~3M events per point…
    - ➢ Last large MC production: 870 points with 1M each, untuned (even faster) MC, took about 2weeks computing time on bird (on ~170CPU‘s on average)
- Parametrized response is a faster solution
  - → $\eta_{UD}(y)$, $E_{UD}(y)$ are most important for the AP
  - → Model is fitted to converted photons
  - → As a physical model it prescribes how to extrapolate to nonconverted photons → handles those too
- But there is more: we have also the LR channels and what about the horizontal direction?
  - → Also $\eta_{LR}(y)$, $E_{LR}(y)$ and if one allows for a horizontal dependence (i.e. not entirely flat), then we have four more functions(x)! Assuming that the problem factorizes in x and y for all…

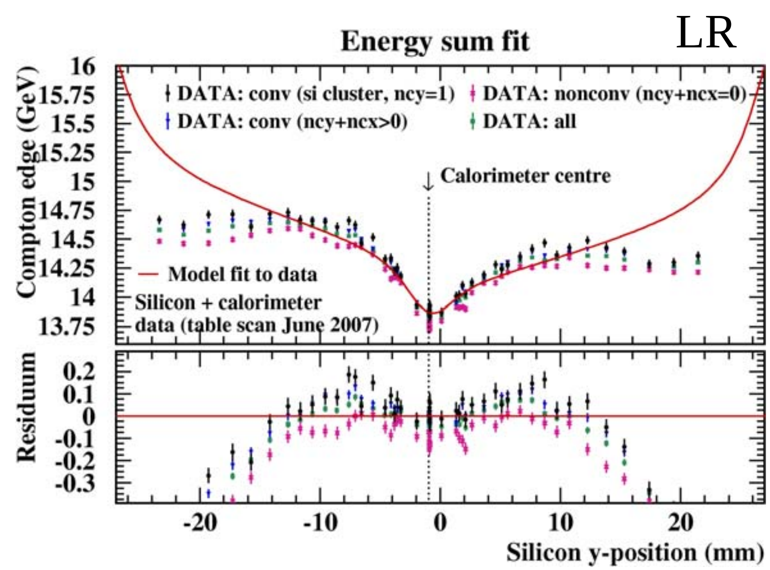
- Nonconverted photon
  - → No conversion width folded into the single particle shower distribution
    - ➢ change corresponding length=0
  - → Different energy loss due to leakage at the backplane
    - ➢ mainly the second shower component, the socalled halo, is leaking
    - ➢ adapt relative leakage factor for halo to fit to data: eloss = 0.991 (total relative loss = $E_{nonconverted}/E_{converted}$)

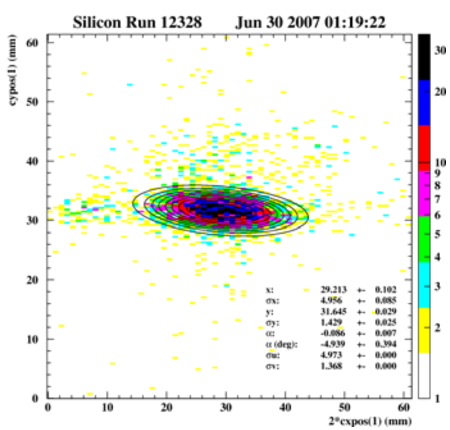


# Silicon Calorimeter Data - LR Channels

- Combined silicon calorimeter data for the LR channels
  - → Same table scan of june 2007 as for the UD curves, but this time showing $\eta_{LR}(y)$ and $E_{LR}(y)$
- Expect to first order a flat behaviour from $\eta_{LR}(y)$
  - → Wiggle in center can be explained with tilt of beam ellipse
  - → Left hand side not understood
- $E_{LR}(y)$ also affected
  - → by e.g. a gap in the center, but less from light attenuation (L and R see the same scintillator areas, attenuation thus cancels to first order)
  - → Expect a larger effect of a geometric light collection factor attenuating the signal in the center and to the sides
  - → Gap appears to be larger, effect of Lead frames (sampling fraction and Moliere Radius change are diminished
  - → Energy response can be partly understood and by heuristic change of parameters even fairly modelled (approx.), though no explicit modelling applied for this channels

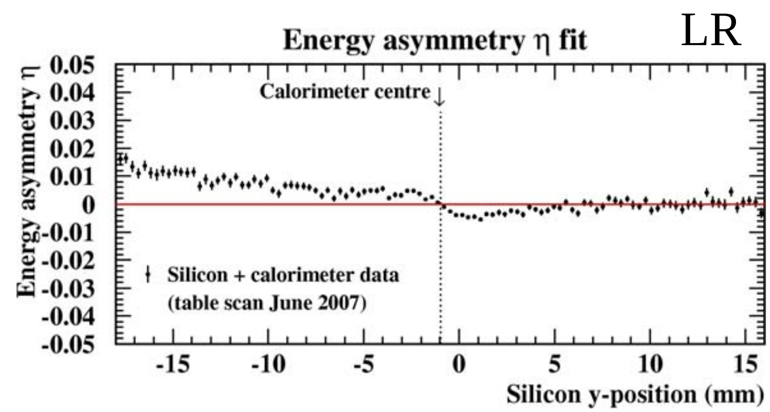


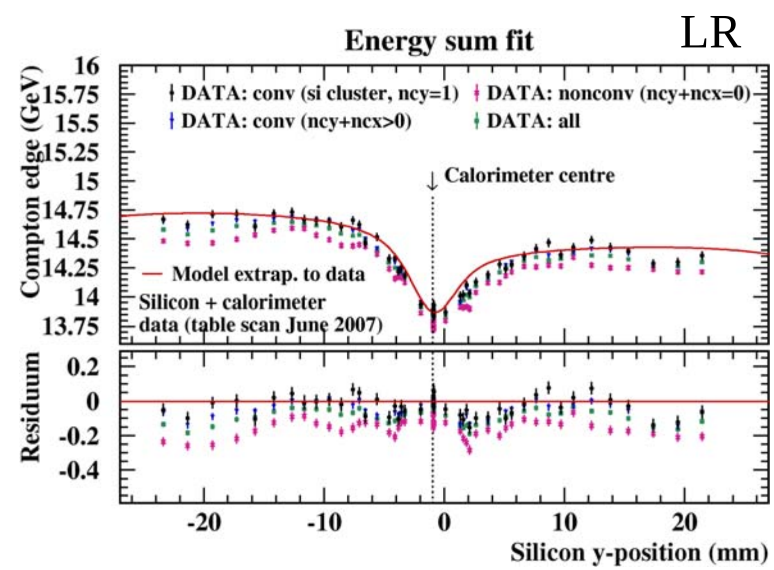



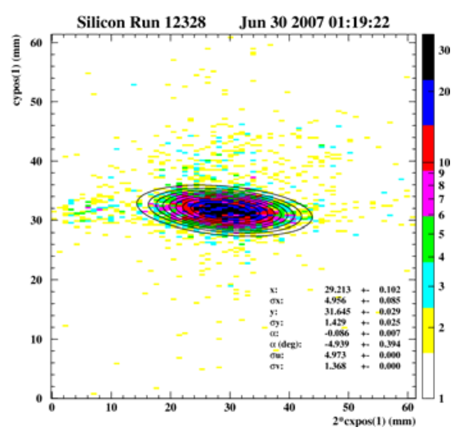



# Silicon Calorimeter Data - LR Channels

- Combined silicon calorimeter data for the LR channels
  - → Same table scan of june 2007 as for the UD curves, but this time showing $\eta_{LR}(y)$ and $E_{LR}(y)$
- Expect to first order a flat behaviour from $\eta_{LR}(y)$
  - → Wiggle in center can be explained with tilt of beam ellipse
  - → Left hand side not understood
- $E_{LR}(y)$ also affected
  - → by e.g. a gap in the center, but less from light attenuation (L and R see the same scintillator areas, attenuation thus cancels to first order)
  - → Expect a larger effect of a geometric light collection factor attenuating the signal in the center and to the sides
  - → Gap appears to be larger, effect of Lead frames (sampling fraction and Moliere Radius change are diminished
  - → Energy response can be partly understood and by heuristic change of parameters even fairly modelled (approx.), though no explicit modelling applied for this channels

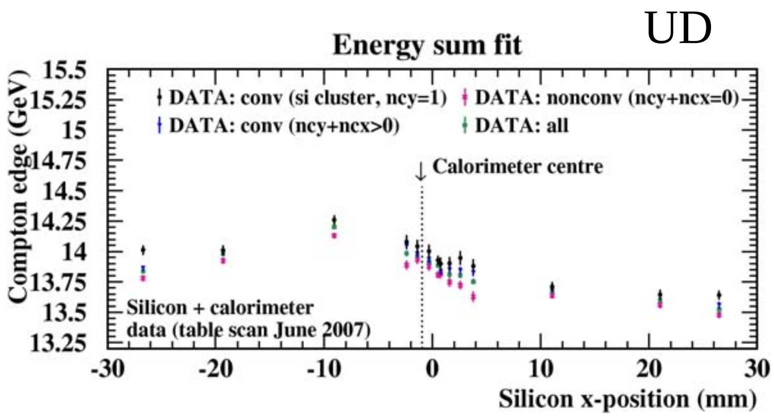


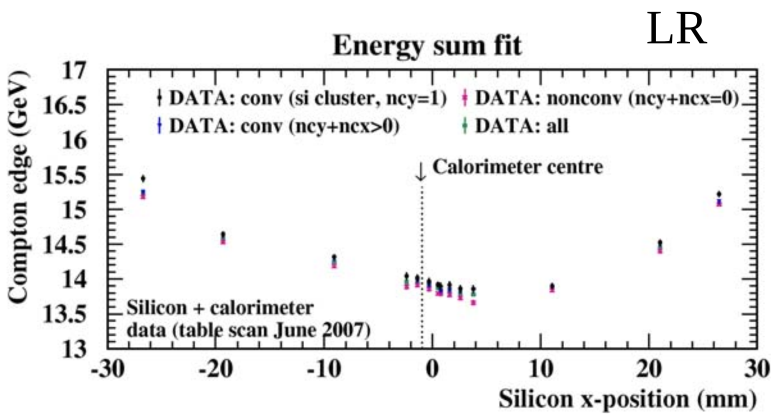



# Silicon Calorimeter Data - Horizontal Response

- Combined silicon calorimeter data
  - → Horizontal table scans of june 2007
  - → Expecting to first order flat $E_{UD}(x)$
    - ➢ Wiggle is not understood
    - ➢ What to do with it? Ignore it, or take some heurisitic wiggly function? Is it important?
  - → Expecting symmetric $E_{LR}(x)$, symmetric around center
    - ➢ Centered around ~10mm on right side and even then asymmetric?
  - → $\eta_{UD}(x)$ and $\eta_{LR}(x)$ still missing
    - ➢ Didn‘t manage to look into this up to now (not trivial, needs reinterpretation because of change from silicon y to x clusters)
- Behaviour of Geant MC concerning LR or x dependencies not checked
  - → Nothing expected there, as only exponential light attenuation is implemented, but no x dependencies or asymmetric responses, no geometrical factors (naive calculations are too large, when compared to data, don‘t work)
  - → LR modelling is fair, as there a geometrical factor should be much more important than for UD and any x-dependency would essentially point to a geometrical factor…
    - ➢ So, the Geant response still has some more things it doesn‘t reproduce…
- Easier to tune the parametrized response to data though…



# Summary - Conclusions so far

- Quite extensive Geant tuning not entirely successfull
  - → Many parameters tuned to data, many paramters studied
  - → Still the main issues of UD response not solved
  - → Prediction power of this MC concerning an AP is fair
    - ➢ derived correction functions are not independent of the absolute scale
  - → This tuning has to come to an end
- Parametrized calorimeter response seems to be more promising
  - → Model is successfully fitted to silicon calorimeter data
  - → Incorporates shower related and many detector related effects
  - → Can be used for a parametrized modelling of the calorimeter response
    - ➢ Would reproduce data better
    - ➢ Allows for much higher statistics and more iterations in systematic studies
    - ➢ BUT: How to incorporate digitization / cross talk of cables, etc?
  - → Basic implementation is ready, tuning of final parameters under way

# Outlook - A very personal View

- A possible way to proceed
  - → **Tune** the parameters to have energy resolution, $E_{UD}(y)$ and $\eta_{DU}(y)$ as close to data as possible + some approximate functions and parameters to get the other dependencies (LR and horizontal)
  - → **Rerun MC** production with caloparam, e.g with optics for 2005
  - → Large statistics possible: $O(10^8)$ events per point feasible
  - → Fill no ntuples, fill immediately into histograms instead, calculate moments online
    - ➢ Need to take care of calibration and centering prior to production
  - → Add linear light $S_1$ production for the same set of ‚MC points‘
  - → **Incorporate $S_1$** for given values $S_1$, $S_2$ in correction function for RMSs and the shift of means function (i.e. the AP)
  - → Then **try a glance at data** – with $S_1$, $S_2$ and moments on ntuple basis
    - ➢ New binning and linear light numbers should be available by then
    - ➢ First possibility to have a look at the correction functions, the new analysis procedure and (hopefully) get a glimpse at the polarization scale
  - → **Study systematic errors,** set the focus on largest (main) errors
    - ➢ IP distance and focus as given by procedure
    - ➢ Table centering
    - ➢ Photomultiplier gain difference, role of pedestal shifts
    - ➢ $S_1$ systematic error
  - → Need to study digitization effects, cross talk of cables, etc
    - ➢ How to incorporate such, if the response is already tuned to data? Digitization module?
  - → How to incorporate peculiarities of data, e.g. arising from the pilot, the way we were calibrating, centering, etc?
  - → How to get this stuff through ‚the data software chain‘?
    - ➢ E.g. finish the correction functions without them and then generate MC with them and study the systematic influence? Only feasible, if influences are small…
  - → What else?...