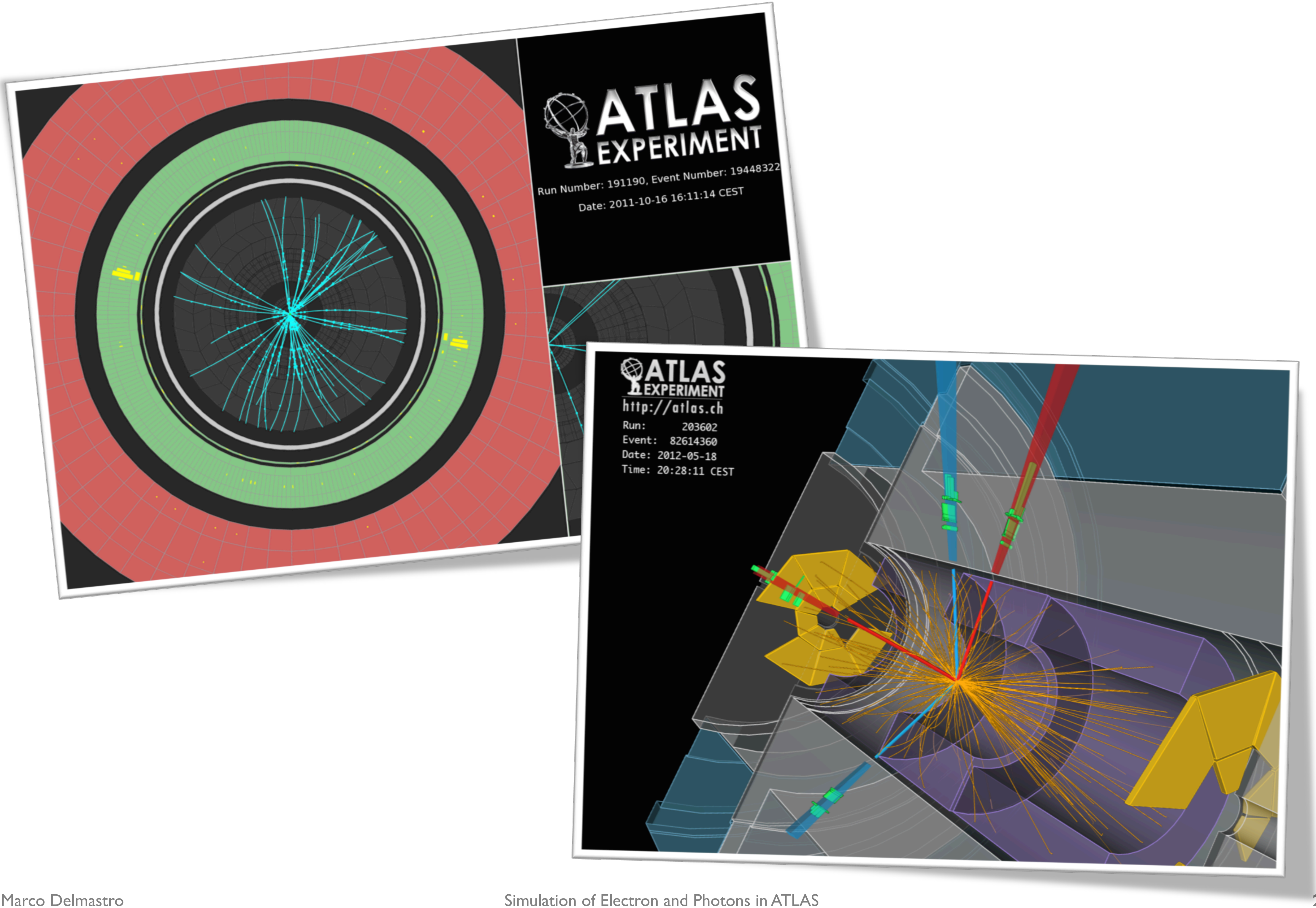
2nd LPCC Workshop on
LHC detector simulations

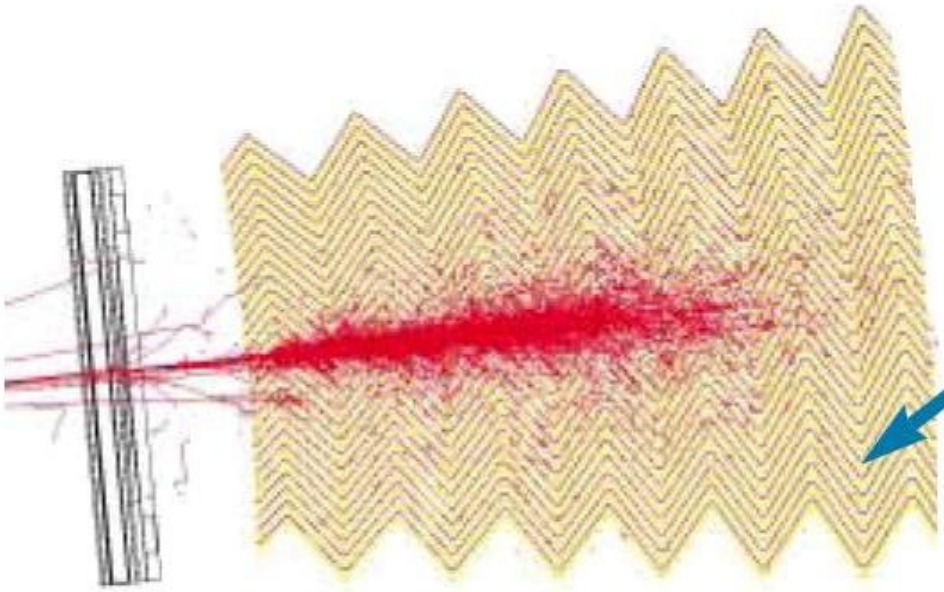
Simulation of Electrons & Photons in ATLAS

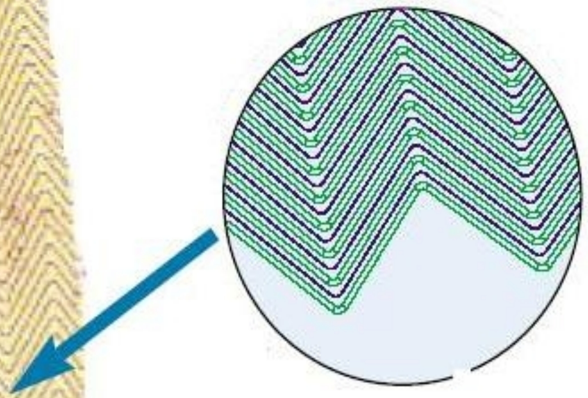
Marco Delmastro
on behalf of the ATLAS Collaboration

March 18th, 2014

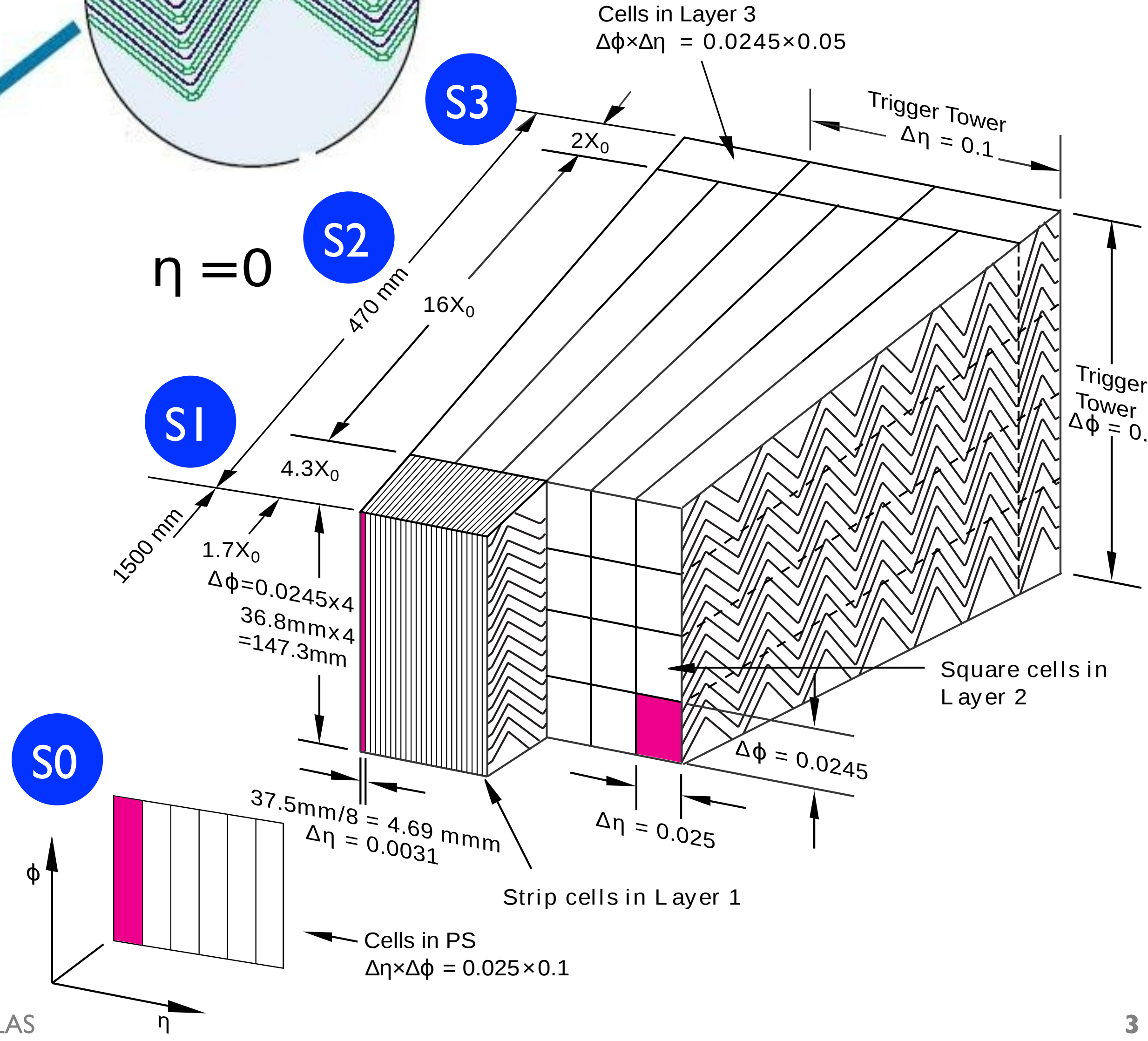

ATLAS EXPERIMENT
Run Number: 191190, Event Number: 19448322
Date: 2011-10-16 16:11:14 CEST
ATLAS EXPERIMENT
http://atlas.ch
Run: 203602
Event: 82614360
Date: 2012-05-18
Time: 20:28:11 CEST

ATLAS liquid argon electromagnetic calorimeters





sampling calorimeter
Pb-LAr

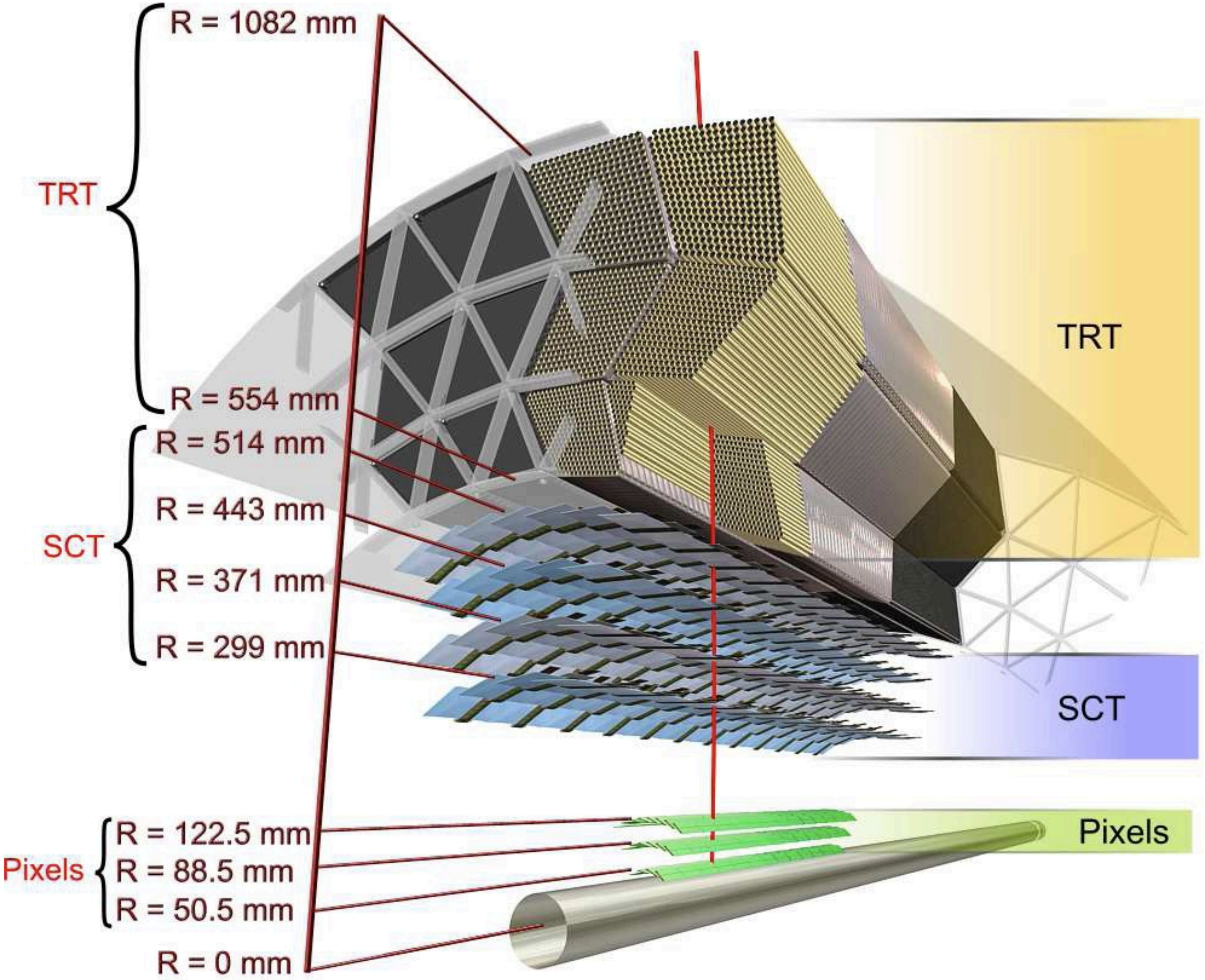


10 % 250 MeV 0.7%

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$$

• S0 (Presampler)	Energy loss correction	
• S1 (Strips)	γ/π^0 separation	4.3 X_0
• S2 (Middle)	Main energy deposit	16 X_0
• S3 (Back)	High energy showers	2 X_0

ATLAS Inner Detector

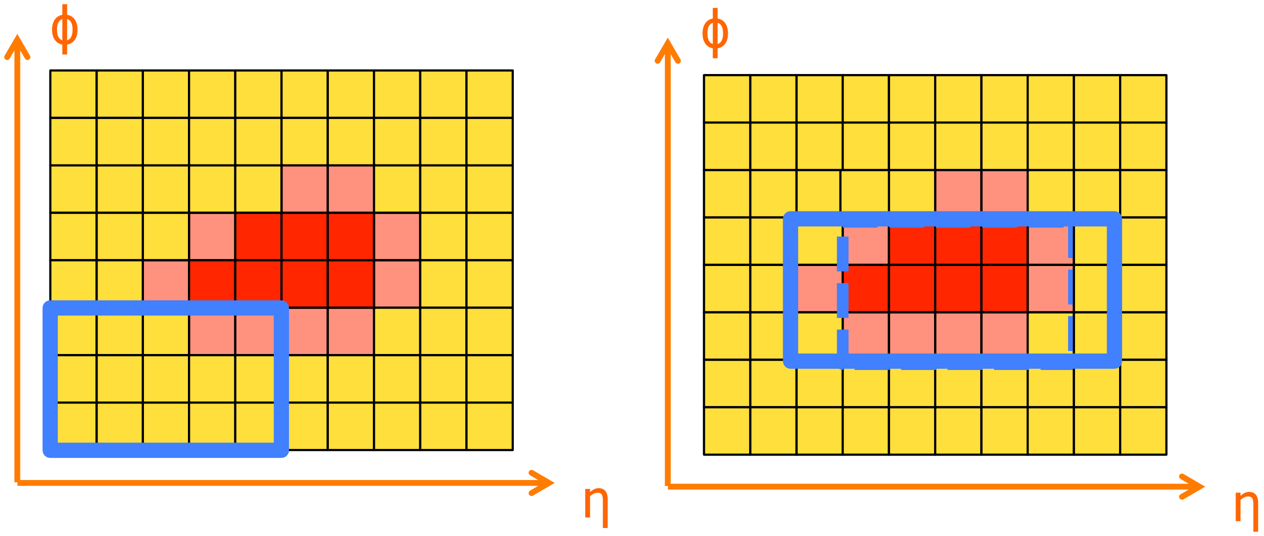


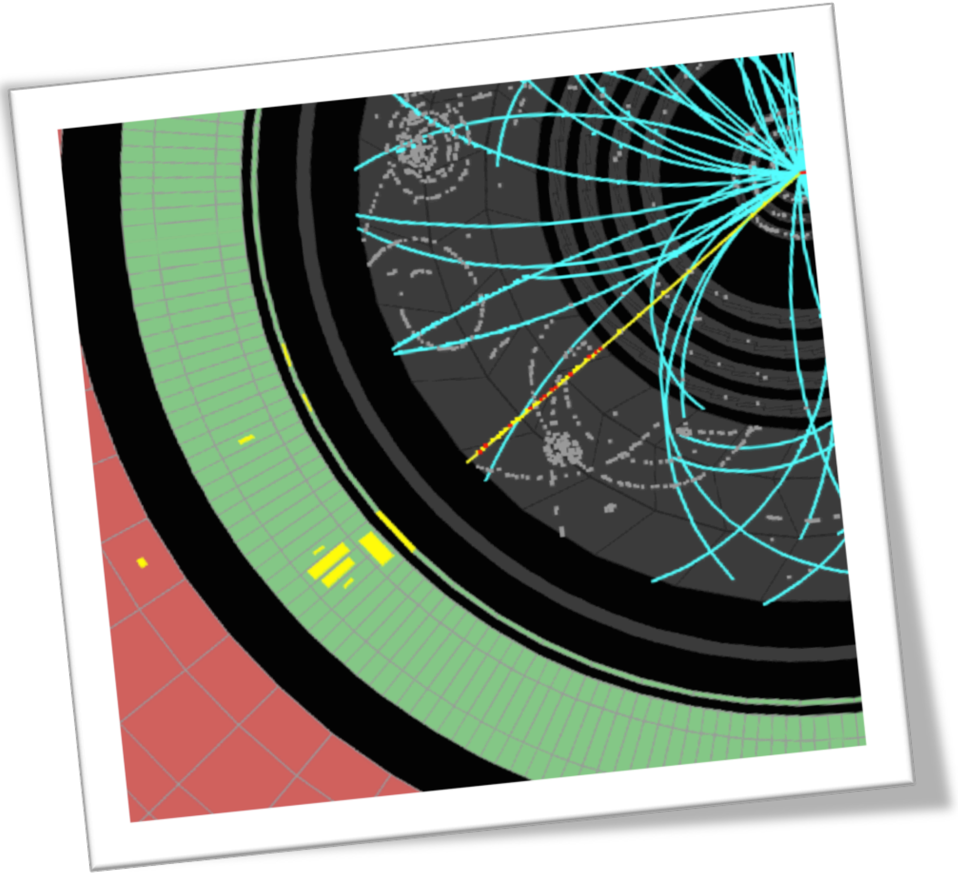
- **Properties of track from charged particles**
  - ✓ momentum
  - ✓ charge
- **Transition Radiation**
  - ✓ e/pi discrimination
- e/γ discrimination
- γ conversion reconstruction

Electrons and photons in ATLAS

Electron and photon candidates in ATLAS are built from EM “sliding window” fixed-size rectangular clusters, that can be associated to tracks

- ✓ $\Delta\eta \times \Delta\phi$ (Middle) = 3 × 5, 3 × 7, 5 × 5
- ✓ Size is trade-off between energy leakage and noise pickup



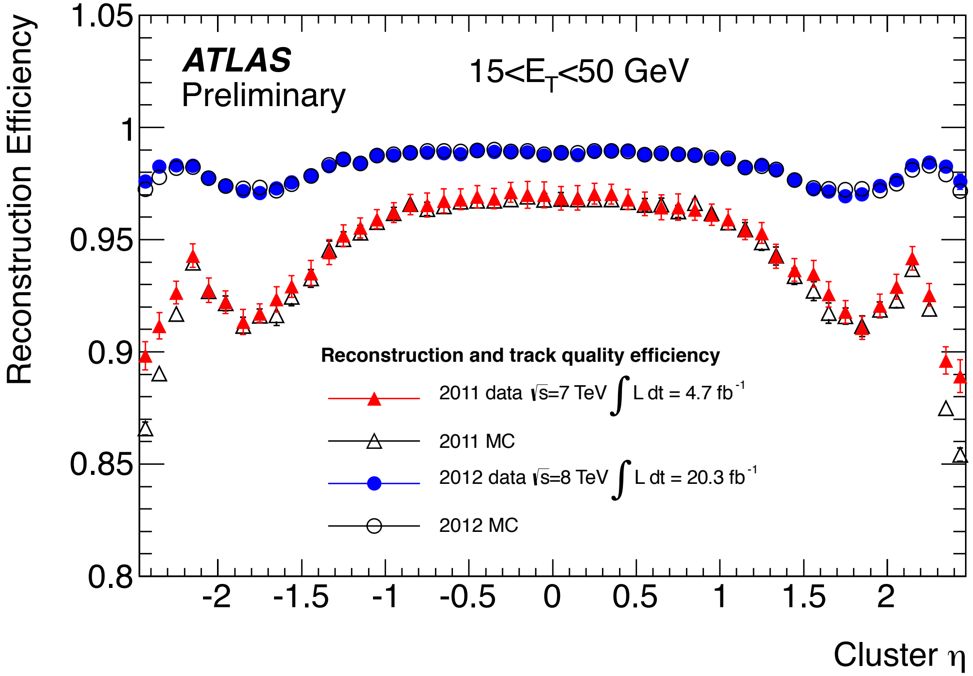


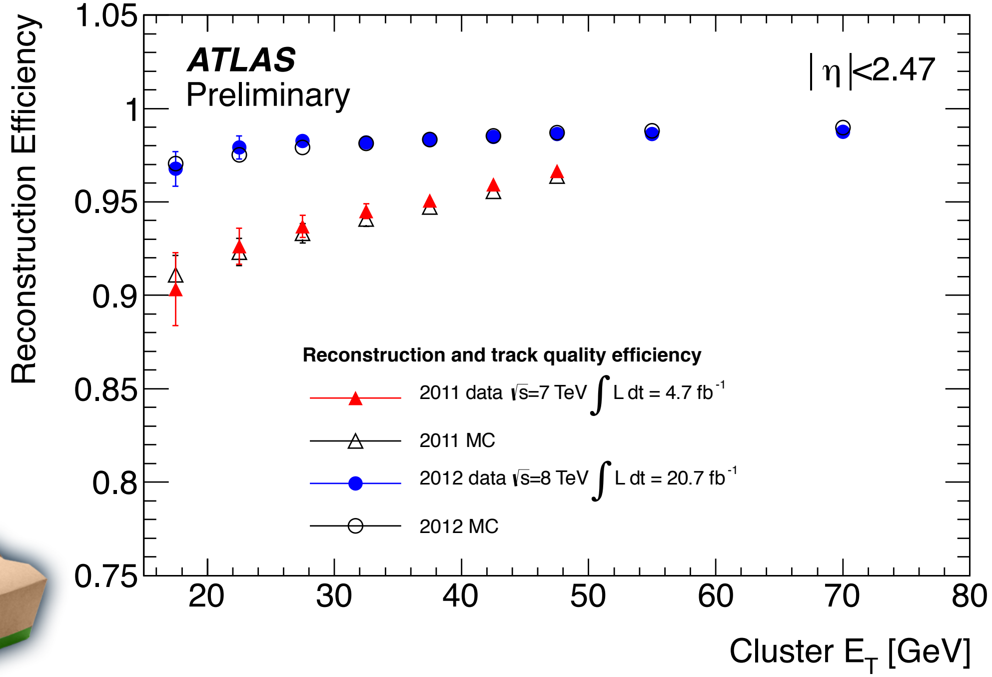
- **Cells in EM clusters are calibrated to “EM scale”**
  - ✓ Including expected average sampling fraction, from simulated electron at 100 GeV
- **Several corrections needed for e and γ, mostly based on simulation**
  - ✓ e.g. energy losses outside cluster, losses in upstream material, variable sampling fraction…

Efficiency of electron reconstruction

- **Cluster reconstruction efficiency better then 99 %**
  - ✓ … for MC electrons with E_T > 15 GeV
- **Track reconstruction + track matching efficiency measured in data**
  - ✓ background estimated in high m_{ee} side-band
  - ✓ estimate electrons without a matched track from fit
- **Improved reconstruction in 2012**
  - ✓ Bremsstrahlung recovery with Gaussian Sum Filter
- **Data/MC scale factors are close to unity,** thanks to excellent description of detector setup and EMC-ID alignment

- Total uncertainty on measured reconstruction efficiencies is better the 0.5% for E_T > 35 GeV
  - ✓ Small Data/MC differences cannot be ignored, Scale Factors are needed on all phase space!



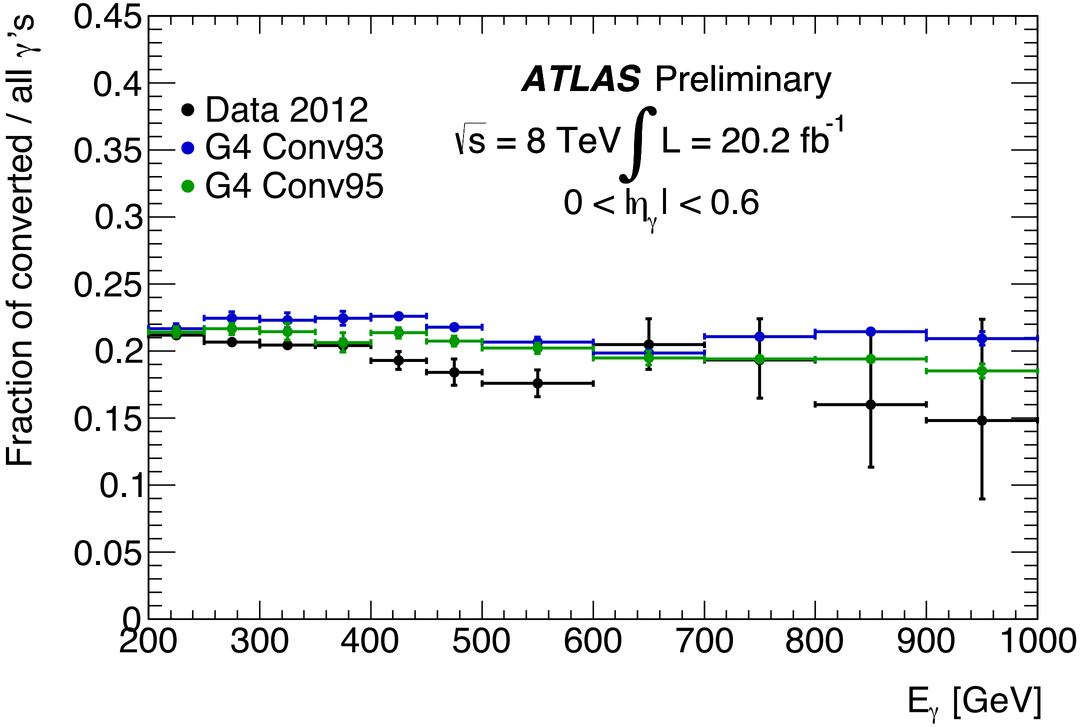


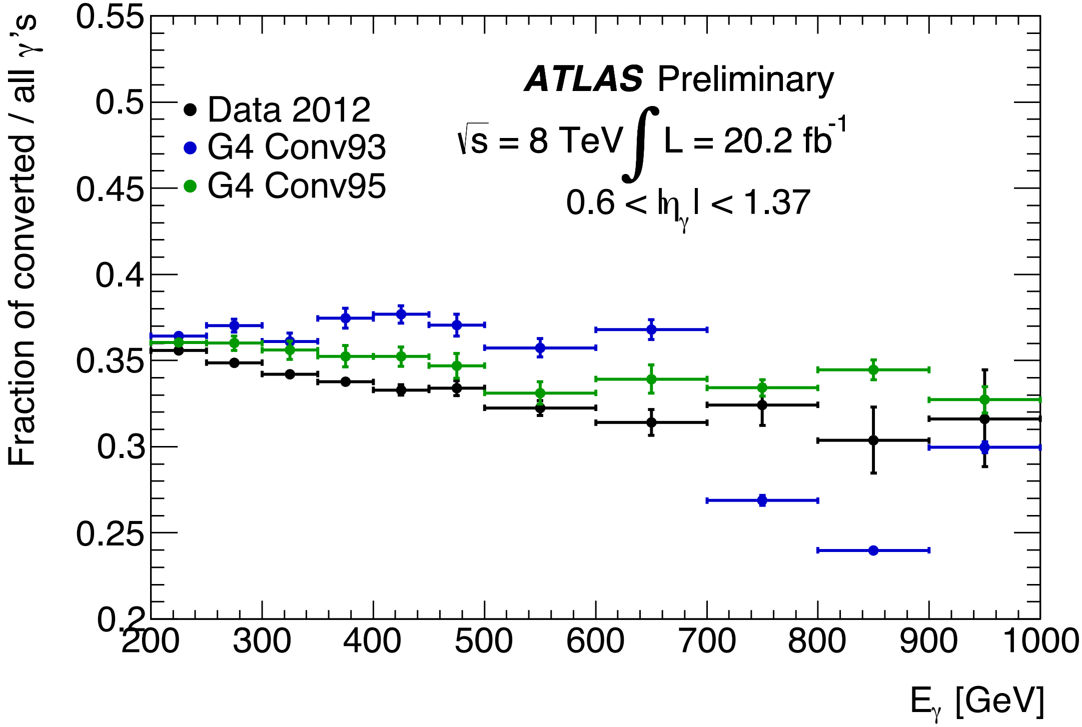
Photon conversion reconstruction

- **Candidate photon conversion vertices reconstructed from tracks pre-selected as loosely matching EMC clusters**
  - ✓ 1 or 2 tracks
  - ✓ 1-track conversions from tracks that missing the hit in innermost ID layer
- Photon reconstruction expected efficiency
  - ✓ ~ 98% for photons $E_T > 25$ GeV
  - ✓ > 99% for unconverted photons
  - ✓ ~ 95% for converted (R < 80 cm)
- **Expected fraction of converted photons**
  - ✓ ~ 20% at $|\eta| \sim 0$ - ~ 45% $|\eta| \sim 1.6$
- Relative fraction of reconstructed photon conversion depends on:
  - ✓ Material upstream EMC
  - ✓ In MC, on conversion model…

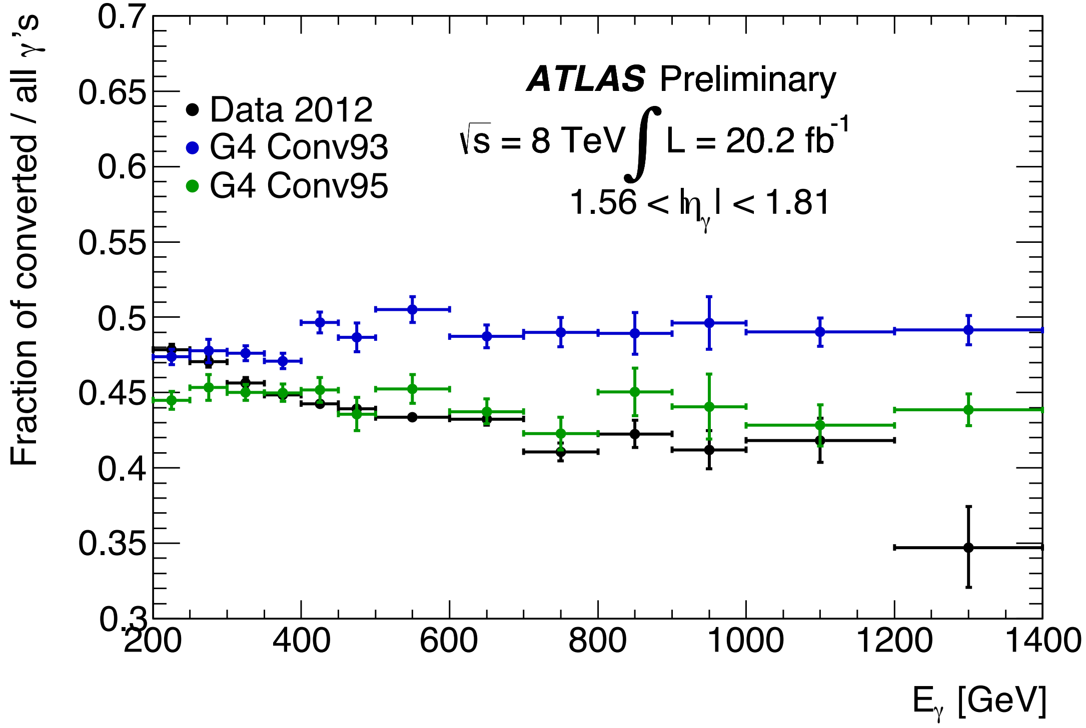
Photon conversion fraction vs. G4 Conv model

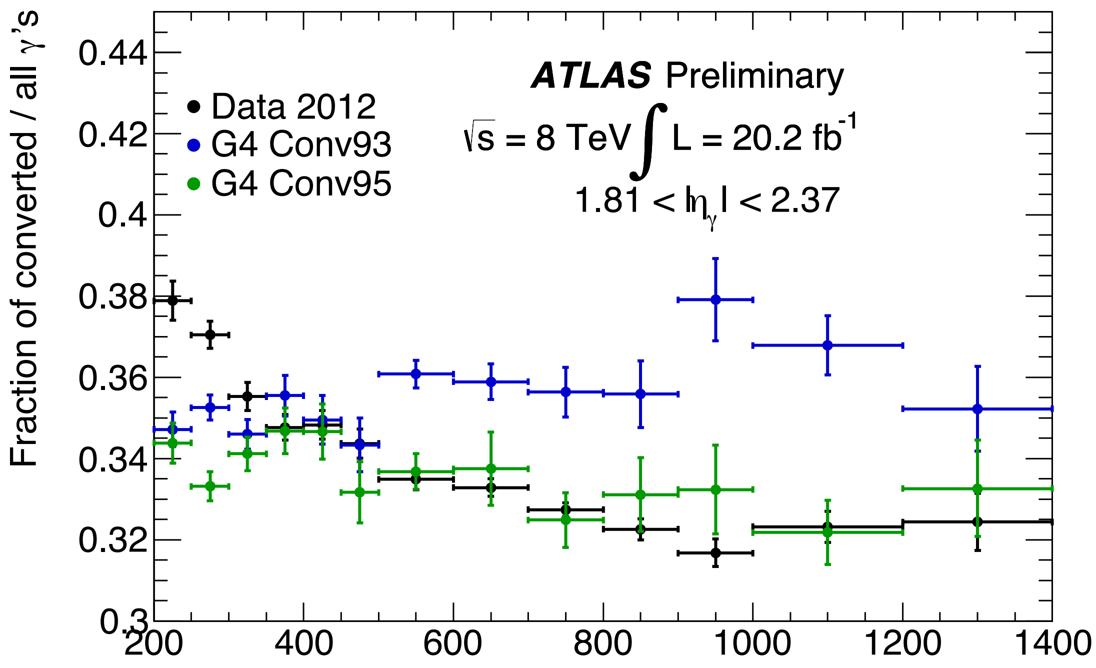
- Photon candidates passing the tight quality selection, and having E_{Tiso} < 3 GeV
  - ✓ Purity > 80% at E_T ~ 25 GeV, better then 99% for E_T > 200 GeV
- **Direct photons (from hard scattering and fragmentation) simulated in detector with G4 using different conversion models**
  - ✓ Conv93 *(known to overestimate probability of conversion in silicon tracker)*
  - ✓ Conv95 *(accurate cross section above 100 GeV + ultra-relativistic conversion model accounting for LPM effect)*
- **Fraction of reconstructed photon conversion vs. candidate energy**





Photon conversion fraction vs. G4 Conv model

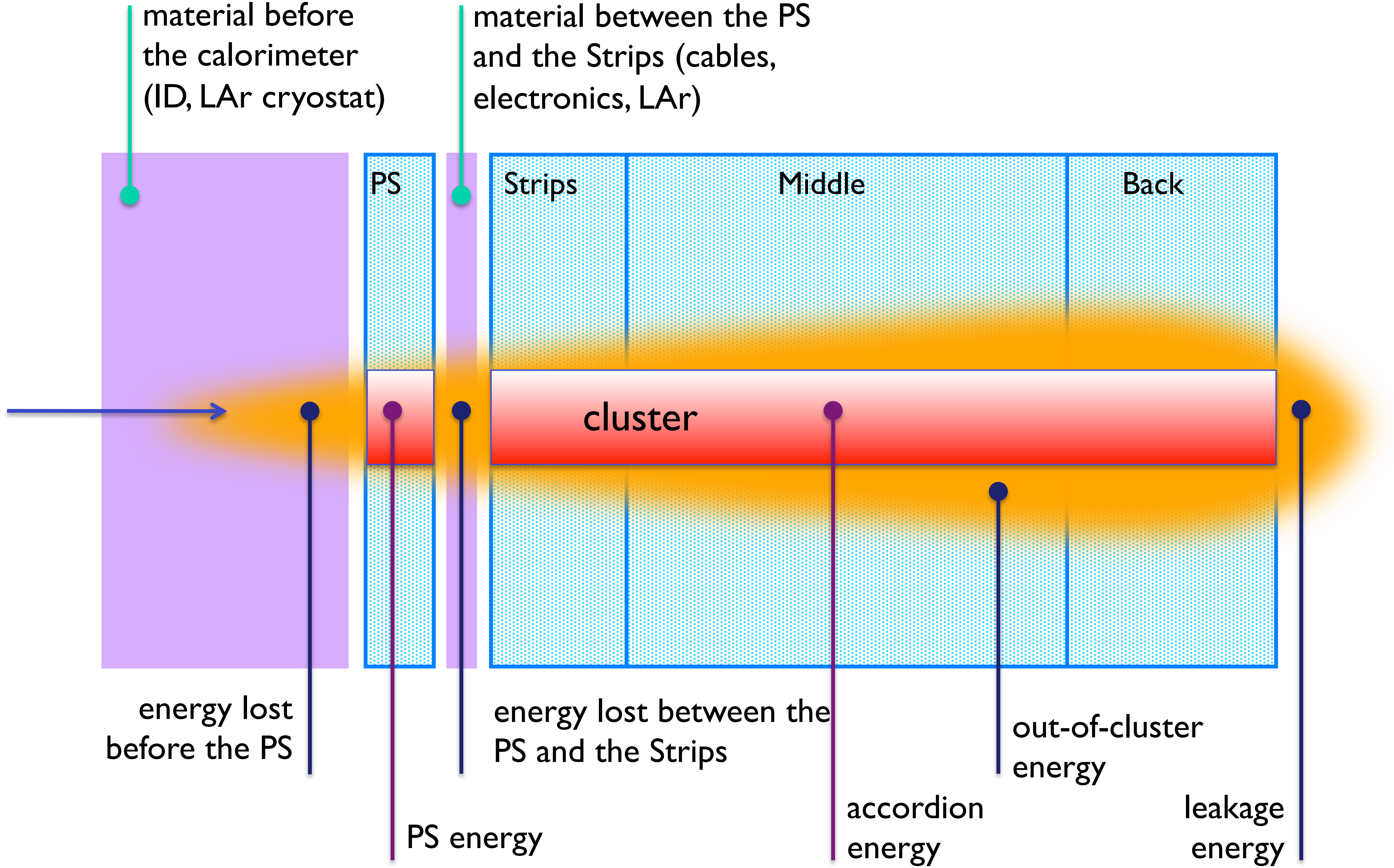




- **Uncertainties are statistical only!**
  - ✓ They do not account for any background contamination in data, nor for possible systematic uncertainties associated with the reconstruction of photon conversions.

- **A qualitatively better agreement between data and MC is observed when the Conv95 conversion model is used**

From cluster energy to electron and γ energy



material before the calorimeter (ID, LAr cryostat)

material between the PS and the Strips (cables, electronics, LAr)

PS

Strips

Middle

Back

cluster

energy lost before the PS

PS energy

energy lost between the PS and the Strips

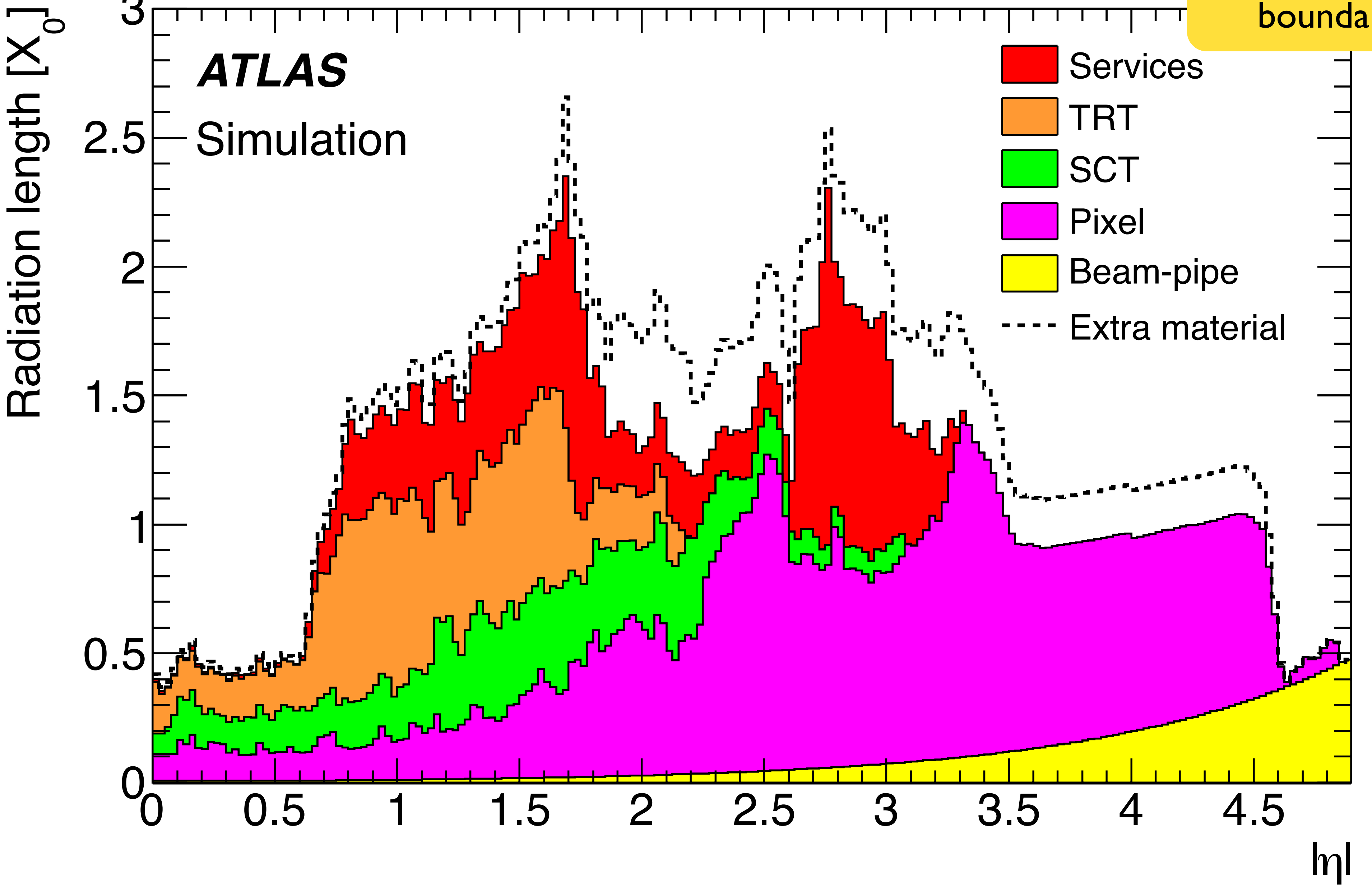
accordion energy

out-of-cluster energy

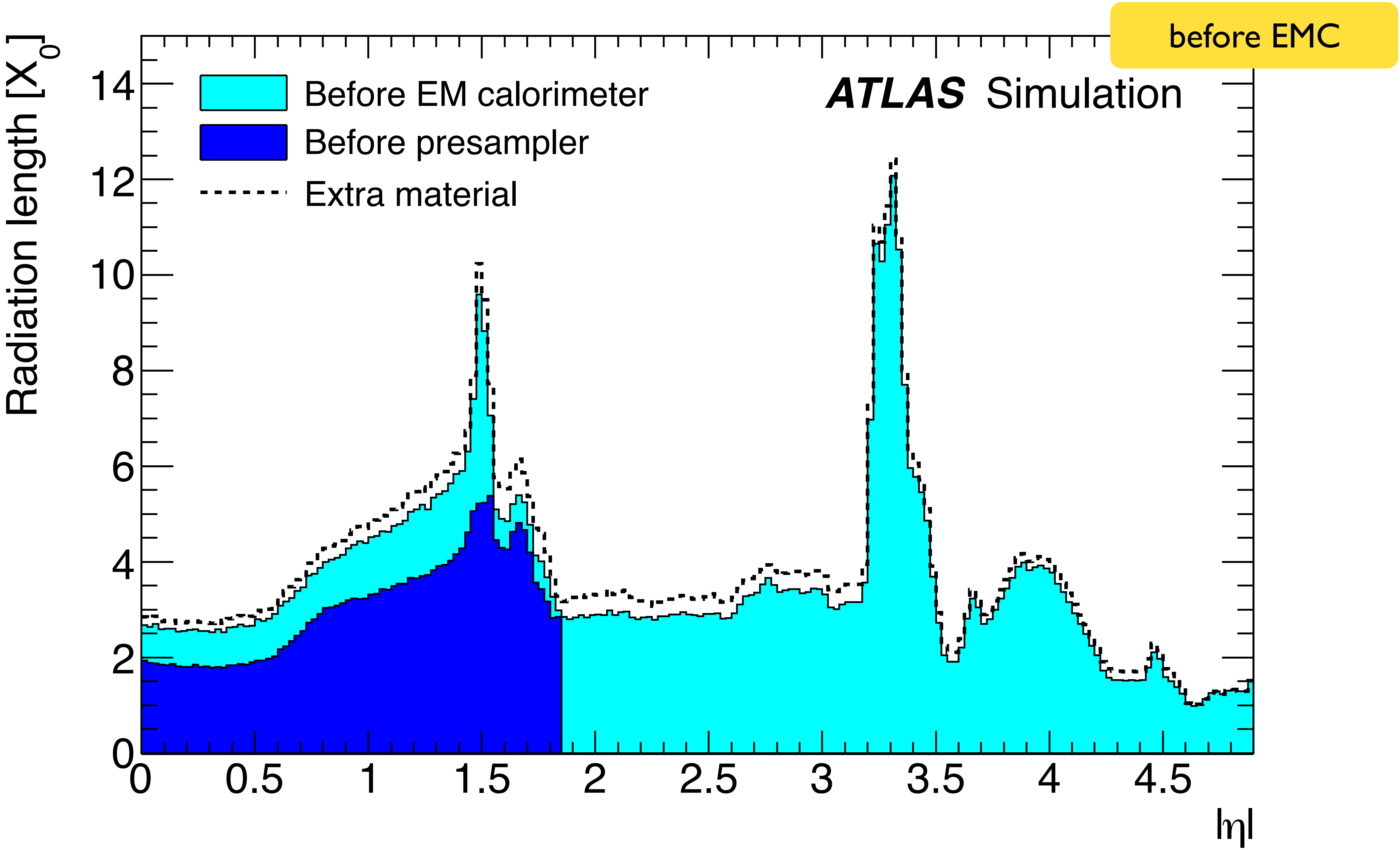
leakage energy

How much material do we have in front of EMC?

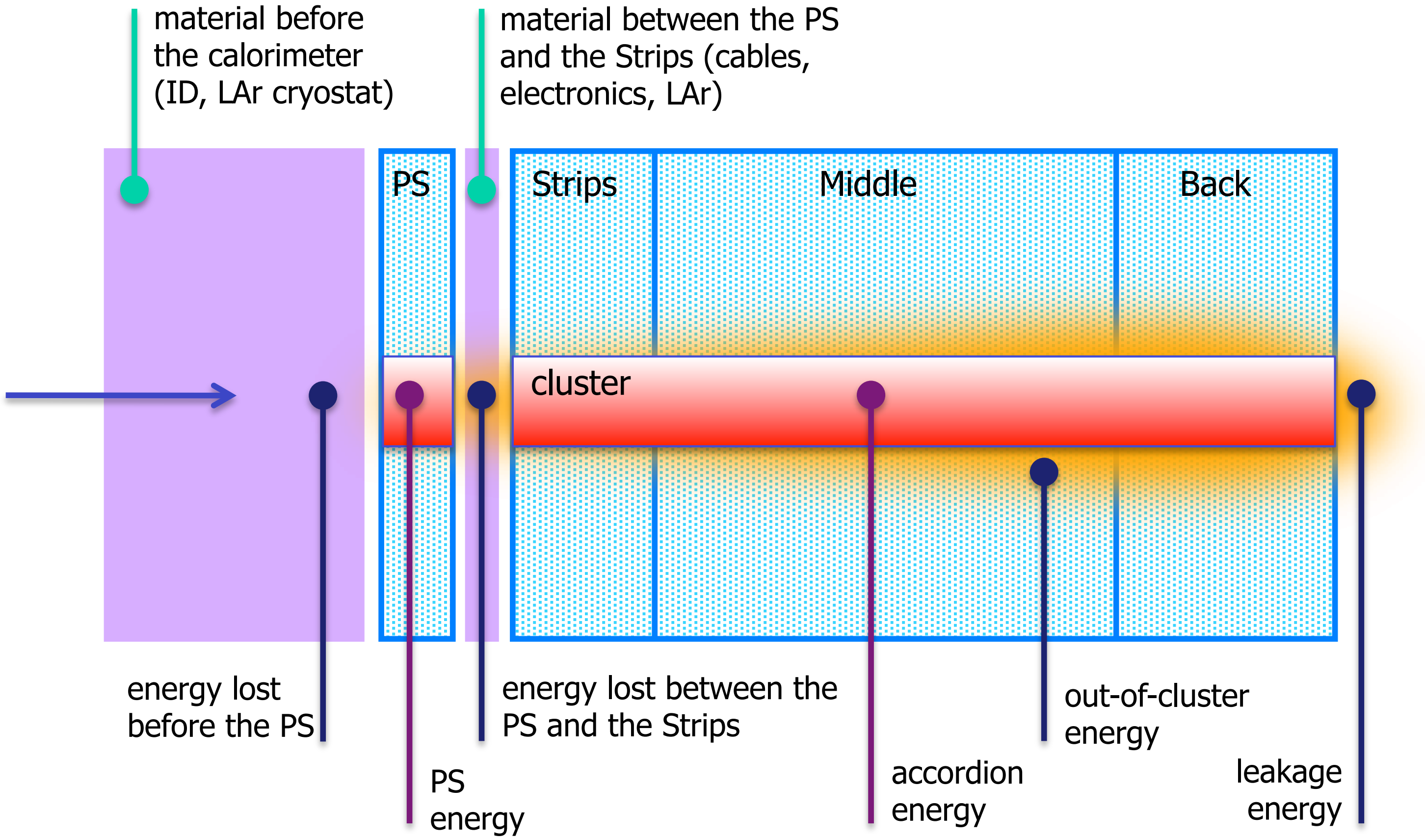
up to ID boundaries



How much material do we have in front of EMC?



How do we calibrate?



material before the calorimeter (ID, LAr cryostat)

material between the PS and the Strips (cables, electronics, LAr)

energy lost before the PS

PS energy

energy lost between the PS and the Strips

accordion energy

out-of-cluster energy

leakage energy

$$X = \frac{\sum_{i=1}^{4} E_i^{cl\,LAr} X_i}{\sum_{i=1}^{4} E_i^{cl\,LAr}}$$

X_0 ($\eta=0$) ~ 1.1 X_0
X_1 ($\eta=0$) ~ 3.7 X_0
X_2 ($\eta=0$) ~ 13.8 X_0
X_3 ($\eta=0$) ~ 23.3 X_0

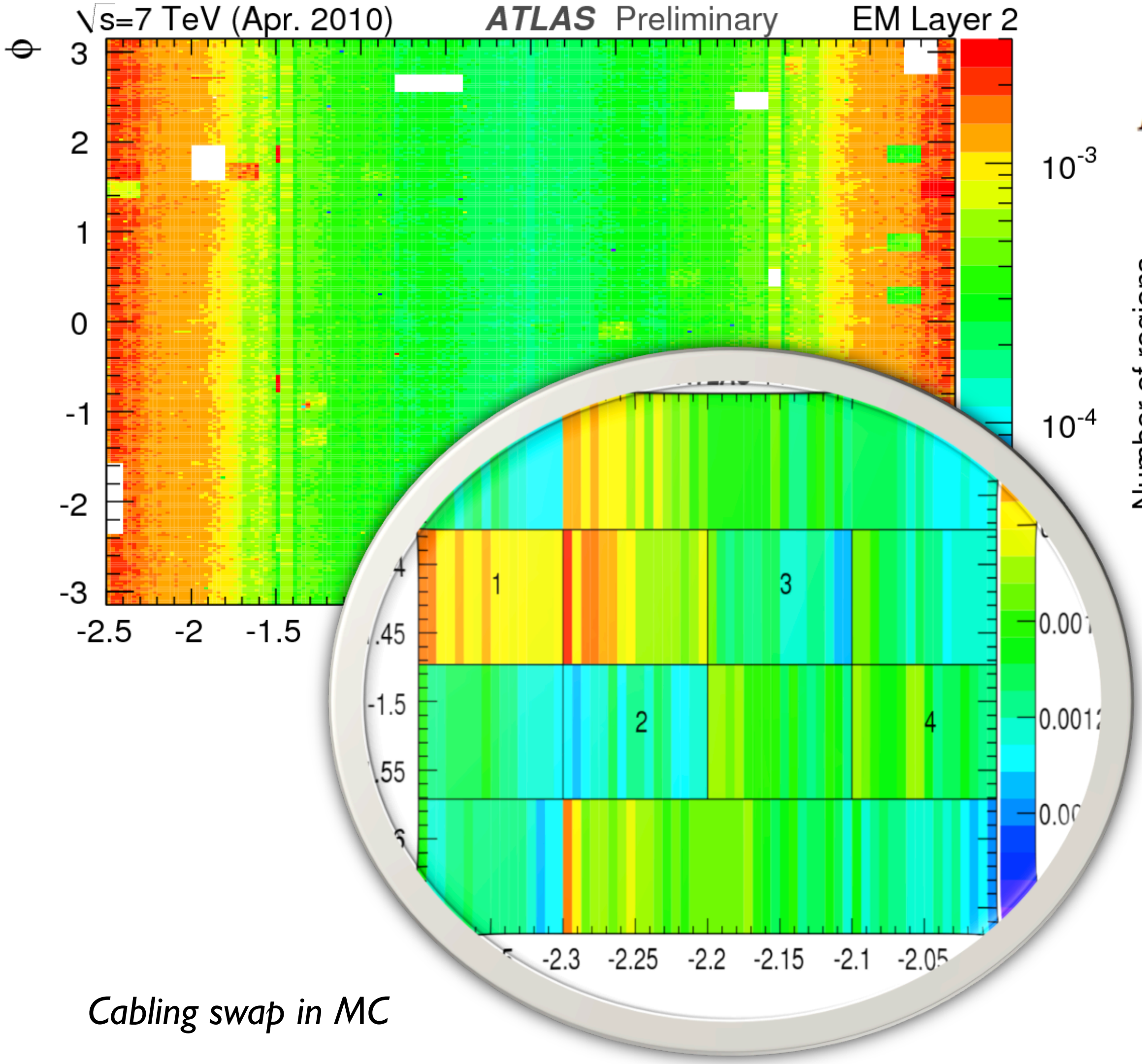
$$E^{cal}_{reco} = a(E^{acc}_{reco},|\eta|) + b(E^{acc}_{reco},|\eta|)E^{cl\ LAr}_{ps} + c(E^{acc}_{reco},|\eta|)(E^{cl\ LAr}_{ps})^2$$
$$+ f_{acc}(X,|\eta|) \times (1+f_{out}(X,|\eta|)) \times \left(\sum_{i=1}^{3} E_i^{cl\ LAr}\right)\left(1 + f_{leak}(X,|\eta|)\right) \times F(\eta,\varphi)$$

- **"Calibration hits"**
  - ✓ Parameterize energy deposits as accounted by G4 vs. visible energy information
- **"MVA"**
  - ✓ Train BDT with visible energy information to return best energy response (not yet public)
- Separately for electrons, unconverted and converted photons!

- Regardless of calibration approach, **knowledge of the material distribution in data** is crucial!

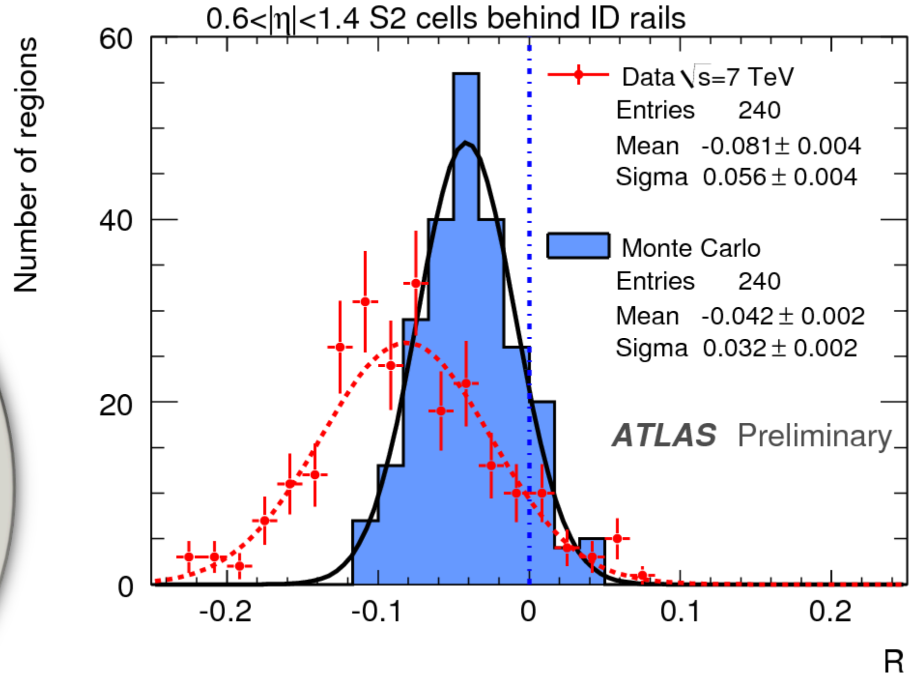
Probing the material before the EMC in ATLAS

- Effort started a long time ago, using minimum bias events…
  - ✓ *"Probing the response of the ATLAS electromagnetic calorimeter and material upstream with energy flow from $\sqrt{s}$ = 7 TeV minimum bias events" (ATLAS-CONF-2010-037)*



Cabling swap in MC

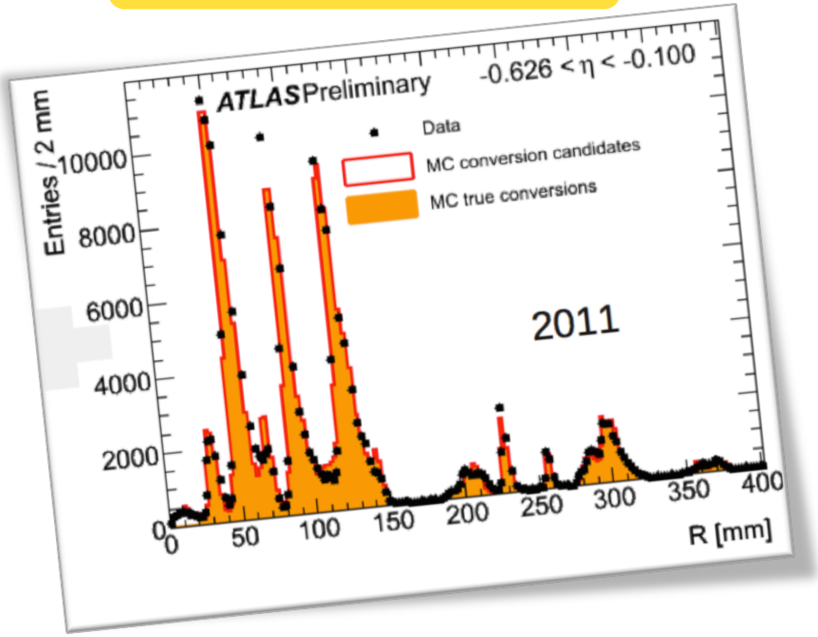
$$R = \frac{Occ_{\phi,\eta-slice} - < Occ_{\eta-slice} >}{< Occ_{\eta-slice} >}$$



Region with amount of material in MC (0.7 X_0) underestimated by factor ~ 2 (interpreted as extra services running parallel to ID rails beyond | η | = 0.6)

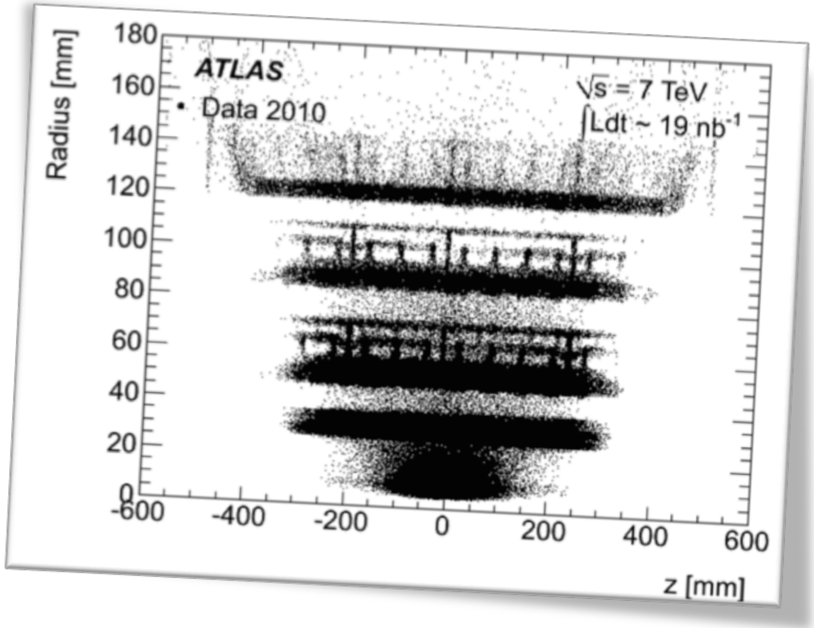
Probing the material before the EMC in ATLAS

photon conversion



Measure radiation length
(i.e. amount of material at a given distance)

hadronic interactions



Measure interaction length
Compared with diffractive MC (PYTHIA6 AMBT1)

- Both methods allowed to constrain **local variations of "early" material at ~5% X_0**
  - ✓ Crucial to define new ATLAS geometry for refined calibration
  - ✓ **Cannot constrain "late" material (after ID, before EMC)**

- More recently, **EMC longitudinal segmentation** used with muons, electron and unconverted photons to **constrain "late" material up to ~2% X_0**!
  - ✓ *Apologies, not public yet (but used in some of the plots shown later)…*

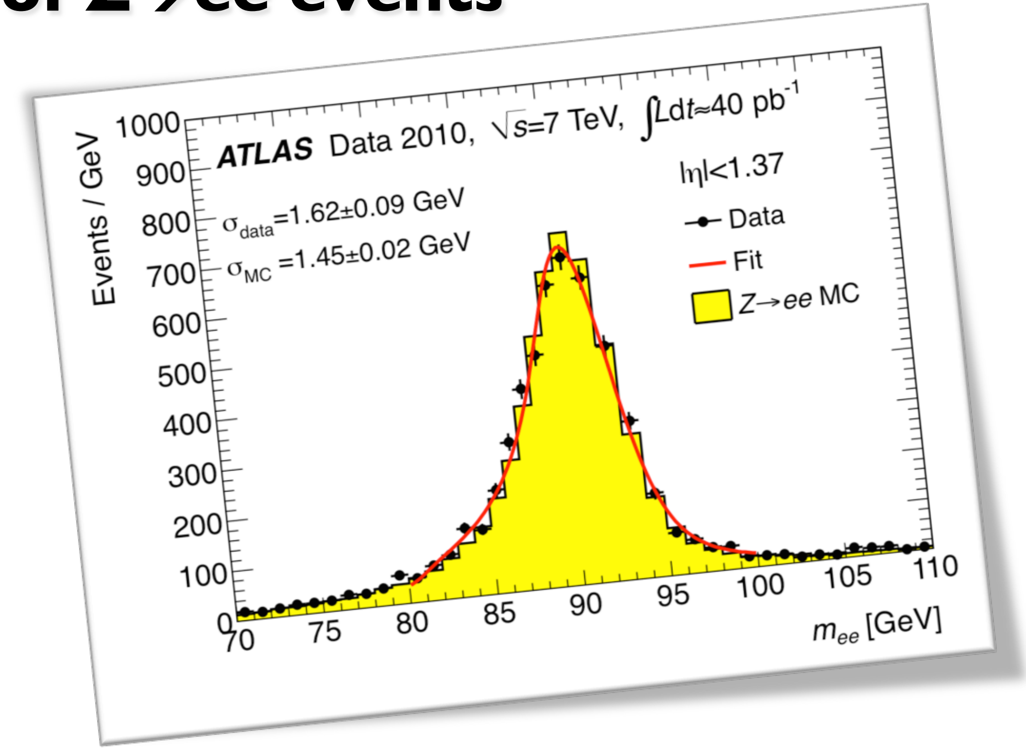
In-situ calibration with Z→ee events

- Whatever the calibration applied to a cluster associated to an electron (e.g. Calibration Hits, MVA, longitudinal weights), **EM scale is constrained in data by comparing the response of Z→ee events**

$$m = \sqrt{2E_1E_2\left(1-\cos(\theta_{12}\right)}$$

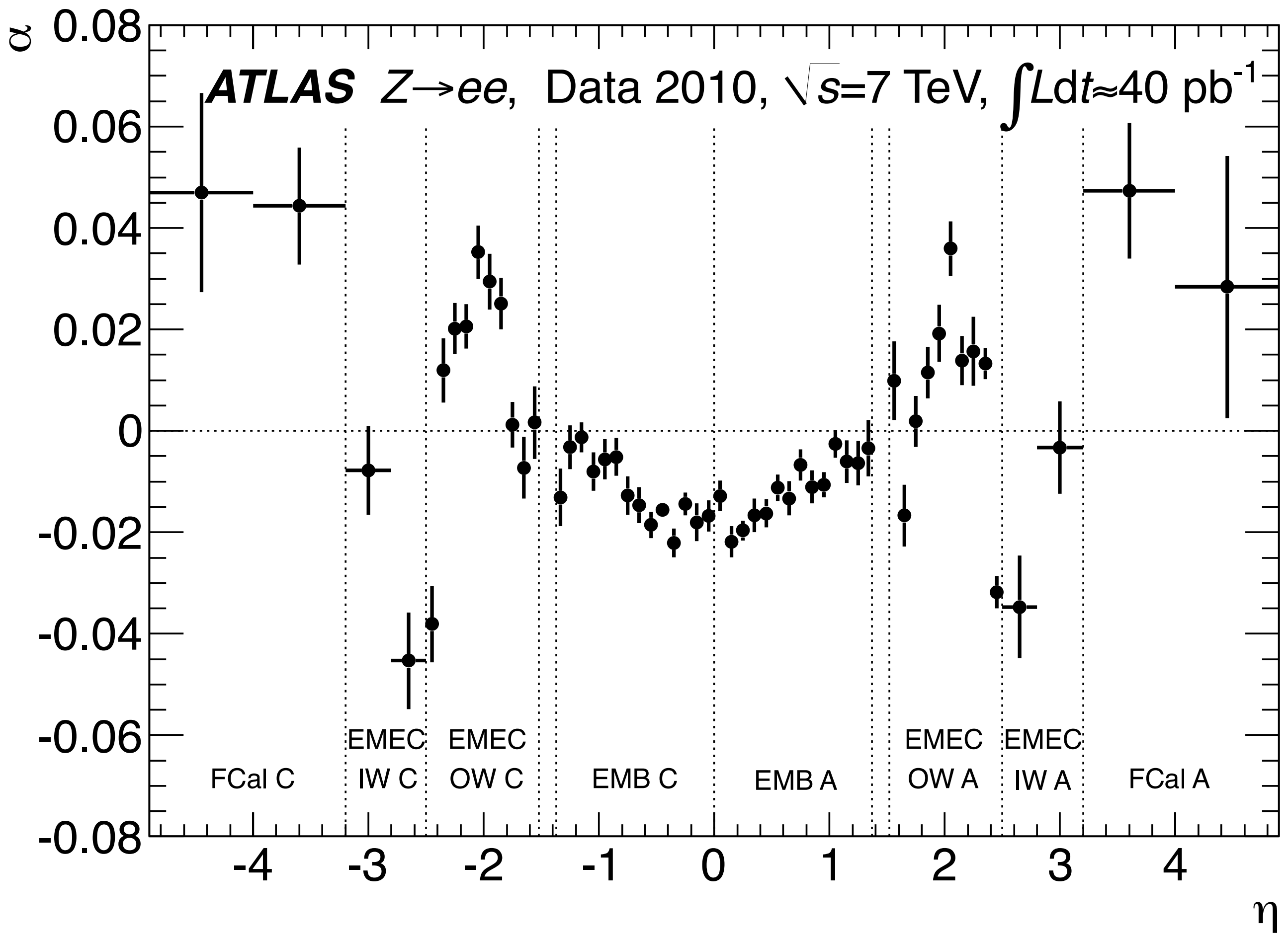
$$E^{\mathrm{corr}} = E\left(1+\alpha_i\right)$$

$$m_{ij}^{\mathrm{corr}} \simeq m_{ij}\left(1+\frac{\alpha_i+\alpha_j}{2}\right)$$



- α_i in-situ scales constrained by comparing the m_{ee} spectra in data and MC
  - ✓ either simultaneously with an unbinned log-likelihood fit, or through a bi-dimensional templated regression of the M_{ee} response
- Smearing applied to MC to account for larger resolution constant term in data

In-situ calibration with Z→ee events



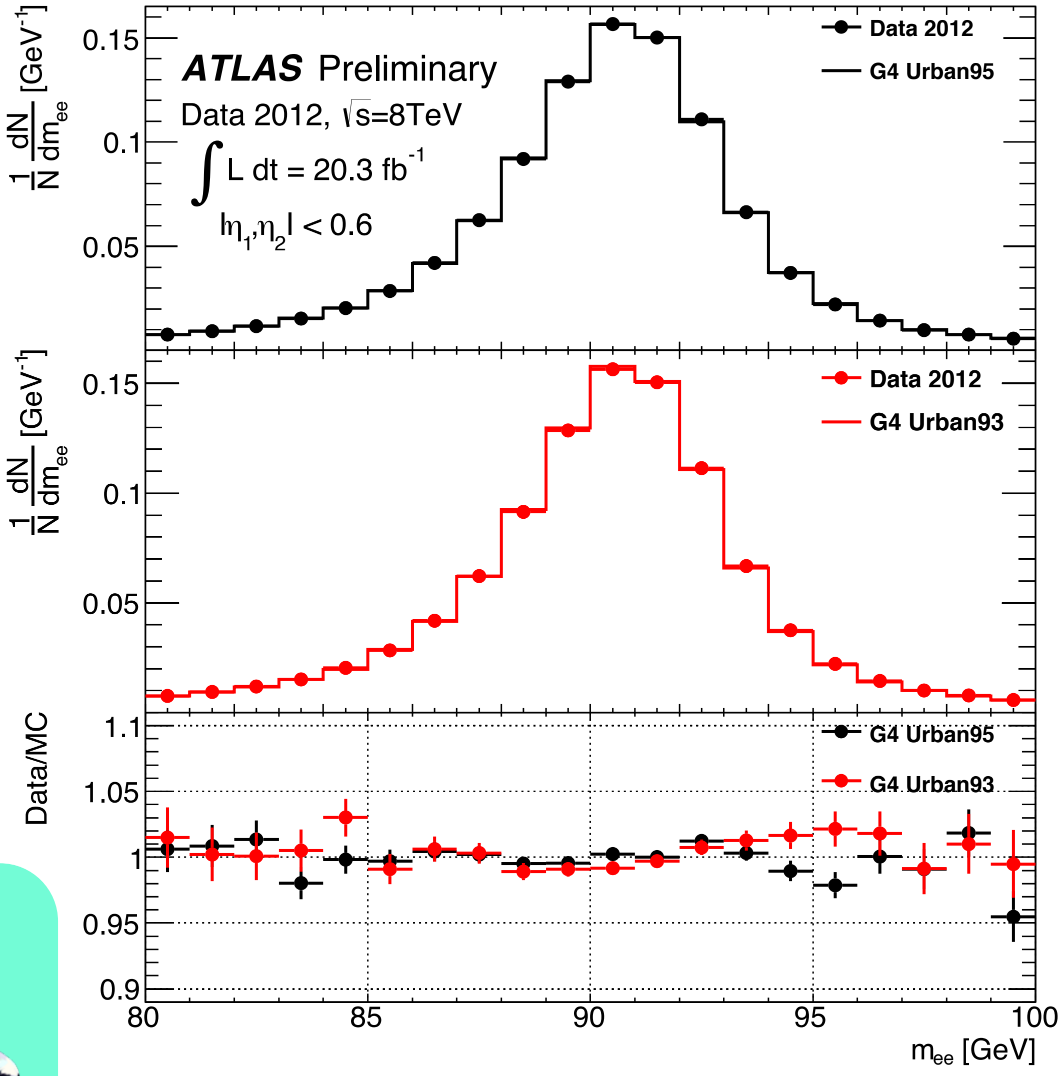
Current EM calibration accuracy (8 TeV data)

- Electrons from *Z* decays: **2×10^{-4}** in most of acceptance, rising to **0.2%** in regions with large amounts of passive material
- Electrons at E_T = 10 GeV: **0.2%-1%**
- Photons: **0.2- 0.3%**

How well is Z→ee line-shape described by MC?

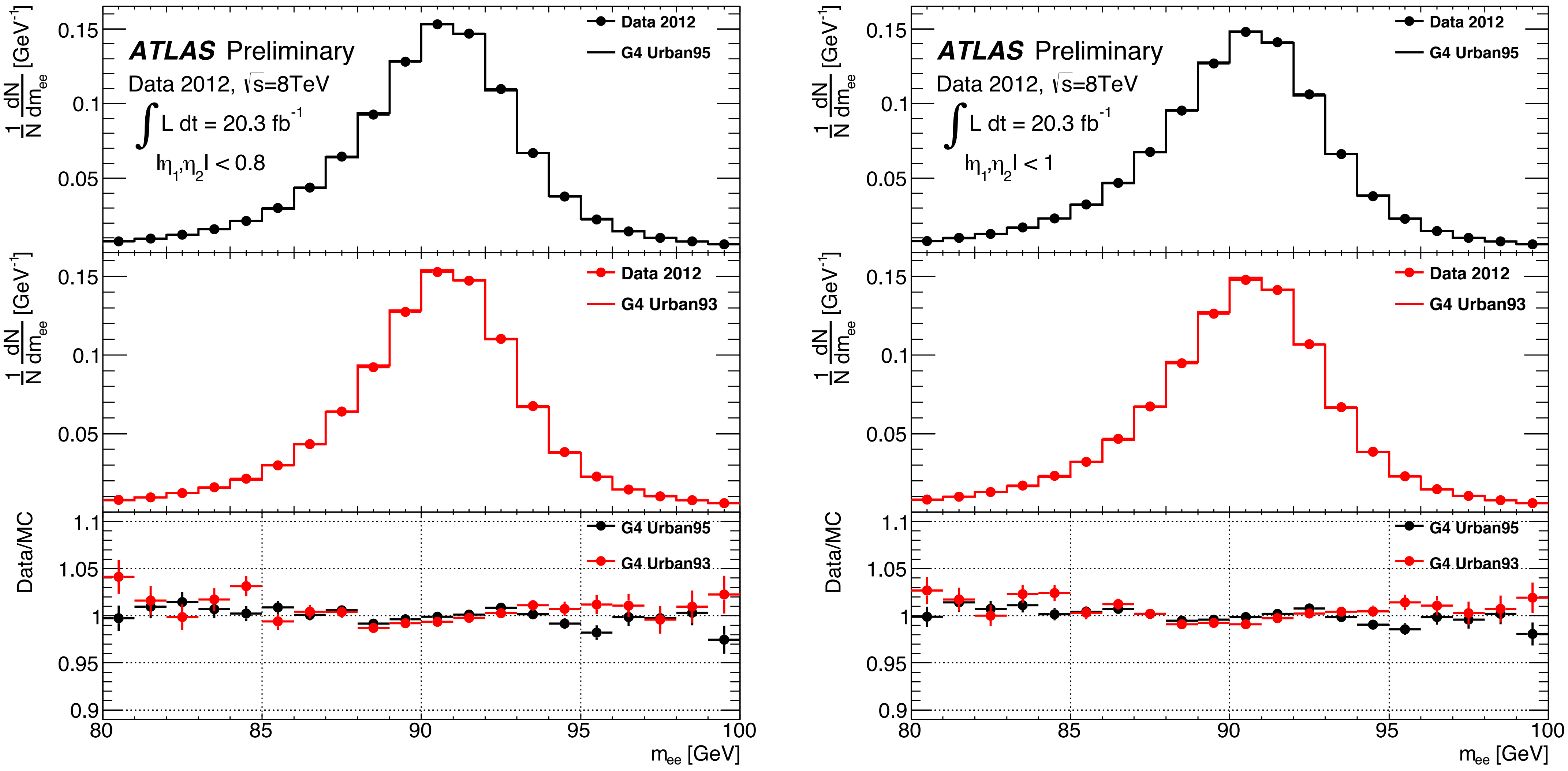
- Z→ee events simulated with POWHEG interfaced to PYTHIA8 with AU2CT10 tuning
- Same selection in data and MC
  - ✓ 2 electrons with $|\eta| < 2.47$, both medium quality and ET > 27 GeV
  - ✓ 2 candidates must have opposite charges $80 < m_{ee} < 100$ GeV

- **Z→ee MC simulated with G4 9.4 using different Multiple Scattering models**
  - ✓ Urban93 (with fix for "bad scattering events", e.g. high E electrons deflected in small steps)
  - ✓ Urban95



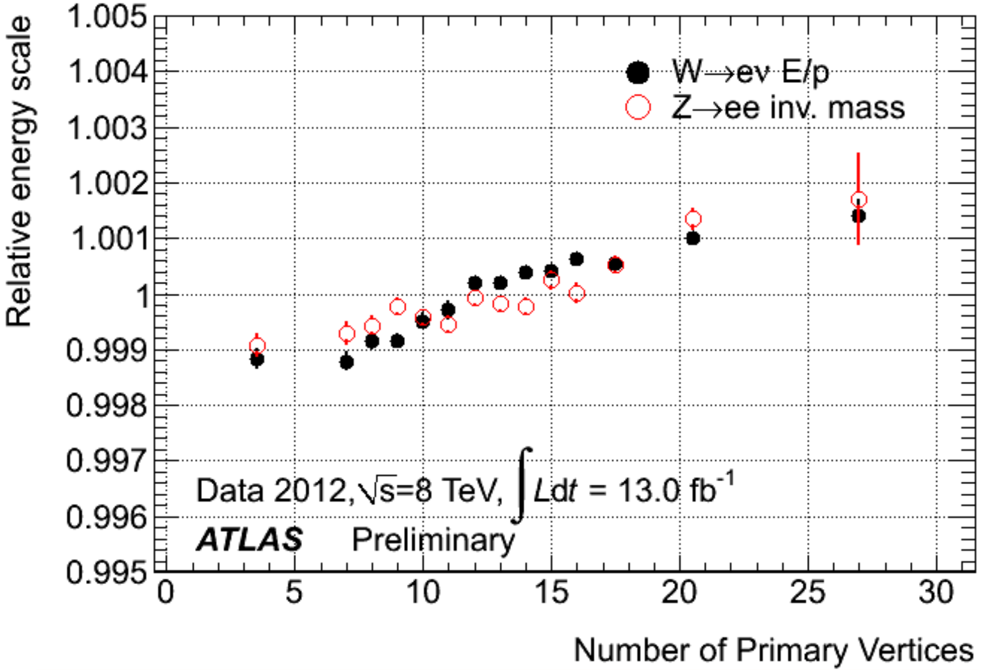
- Uncorrected Z→ee MC line-shapes agree < 1% in m_{ee} = 85-95 GeV, while **show systematic differences ~2-3% in** ***m_{ee}*** **tails**.

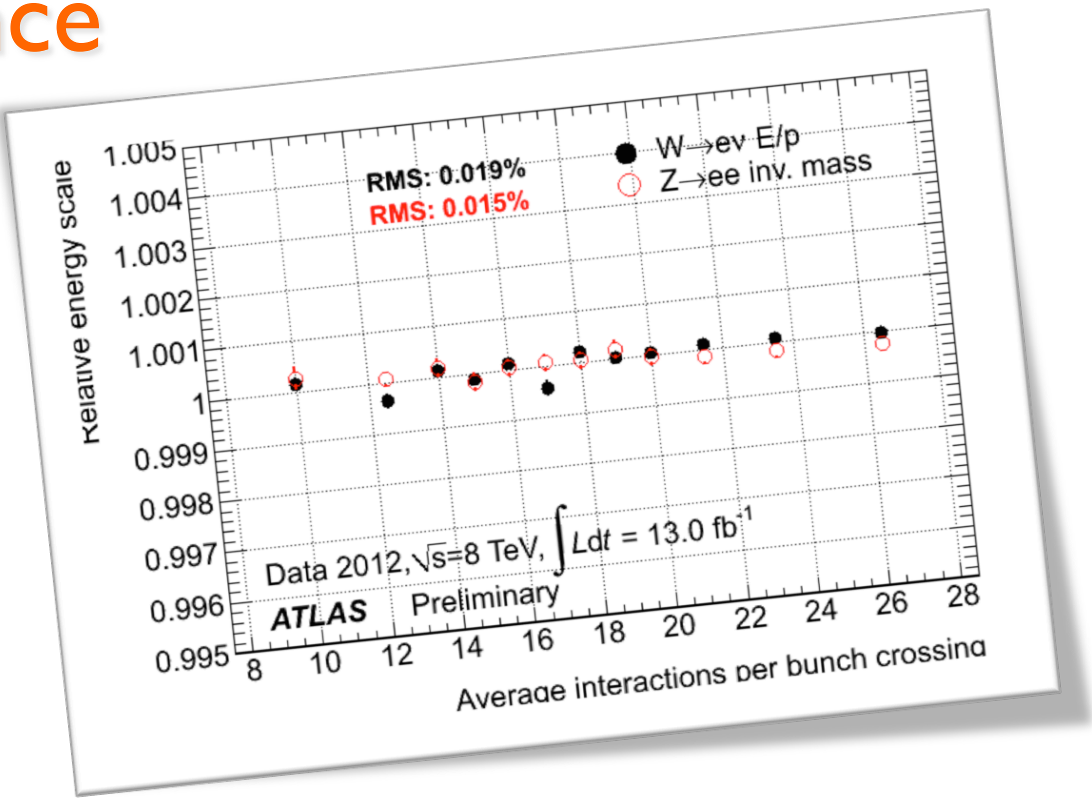
How well is Z→ee lineshape described by MC?

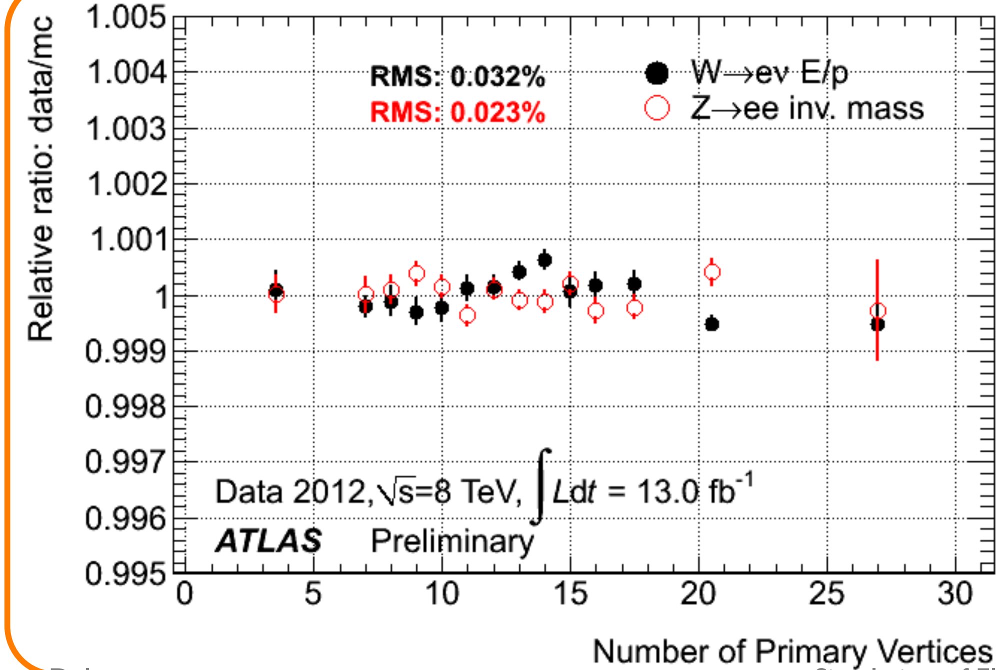


- When compared to data after all corrections, both **G4 versions seem to reproduce the data Z → e+e− lineshape with similar accuracy**
  - ✓ None of the difference between MC simulations account for specific data features

EM scale pileup dependence



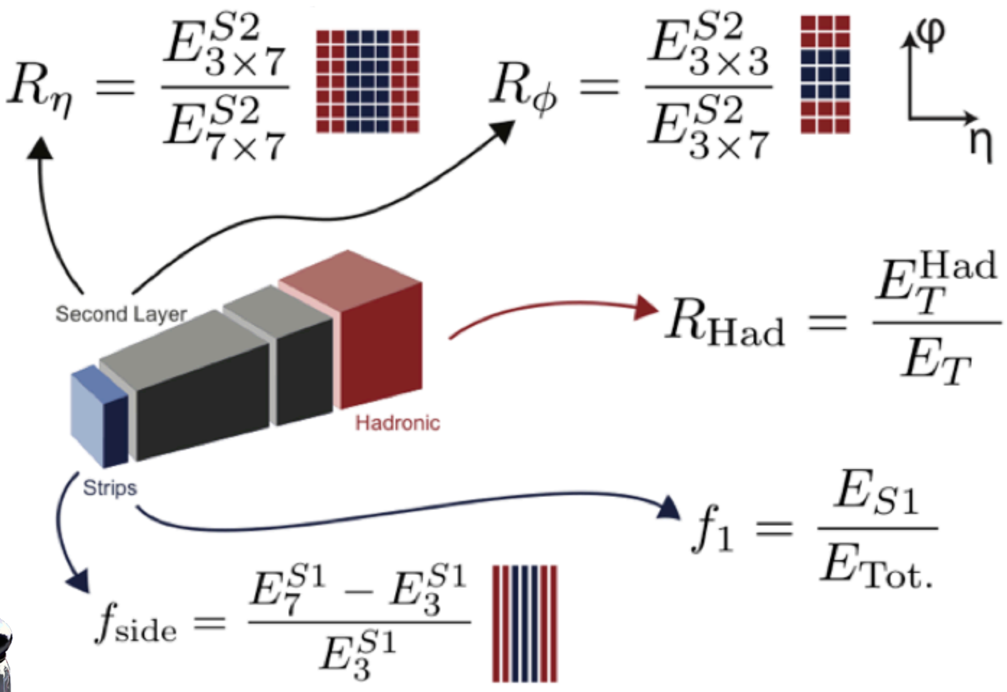




- EM response is extremely stable as a function of pileup (< 0.1%)
- ATLAS EMC response is on average insensitive to pileup
- Local fluctuations are extremely well described by MC!

e/γ selection over background: shower shapes

- Shower shape in EMC allows for rejection of a large fraction of background
  - ✓ R~1000-5000
- This EMC granularity and segmentation of the calorimeter allows to compute shower properties in all directions
  - ✓ η, ϕ, longitudinal, structure..



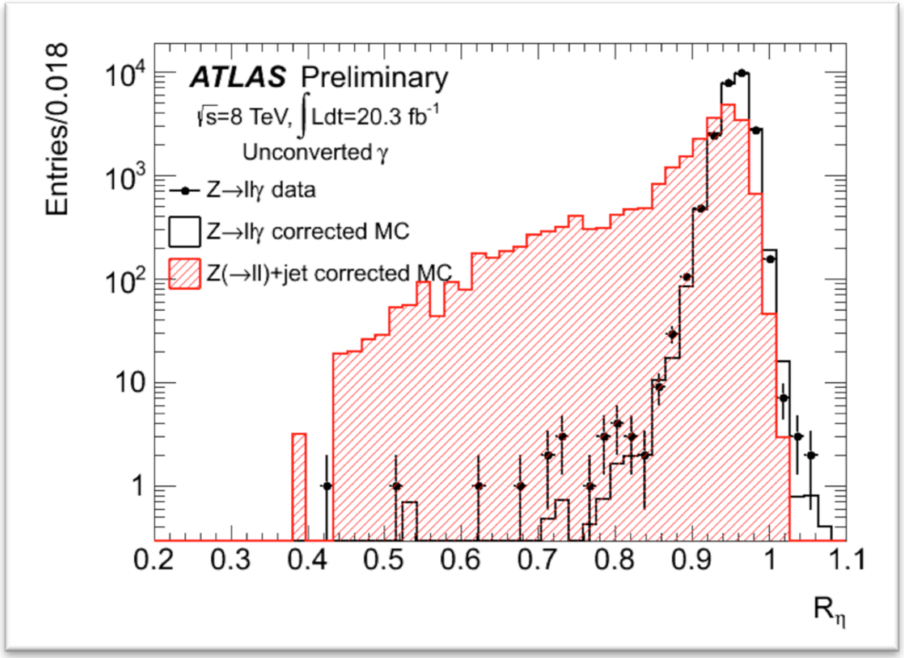
$$R_\eta = \frac{E^{S2}_{3\times7}}{E^{S2}_{7\times7}}$$

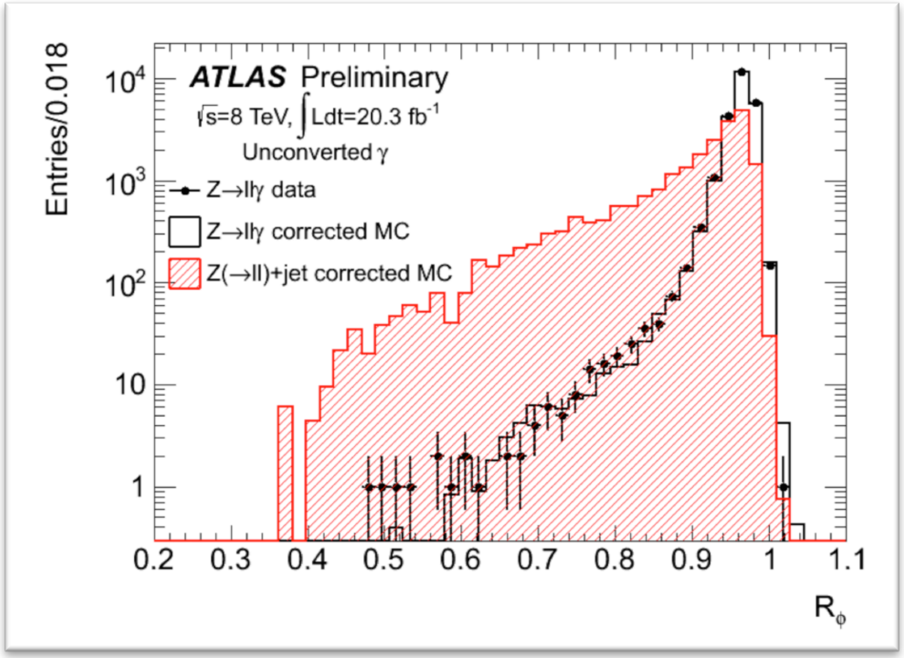
$$R_\phi = \frac{E^{S2}_{3\times3}}{E^{S2}_{3\times7}}$$

$$R_{\mathrm{Had}} = \frac{E_T^{\mathrm{Had}}}{E_T}$$

$$f_1 = \frac{E_{S1}}{E_{\mathrm{Tot.}}}$$

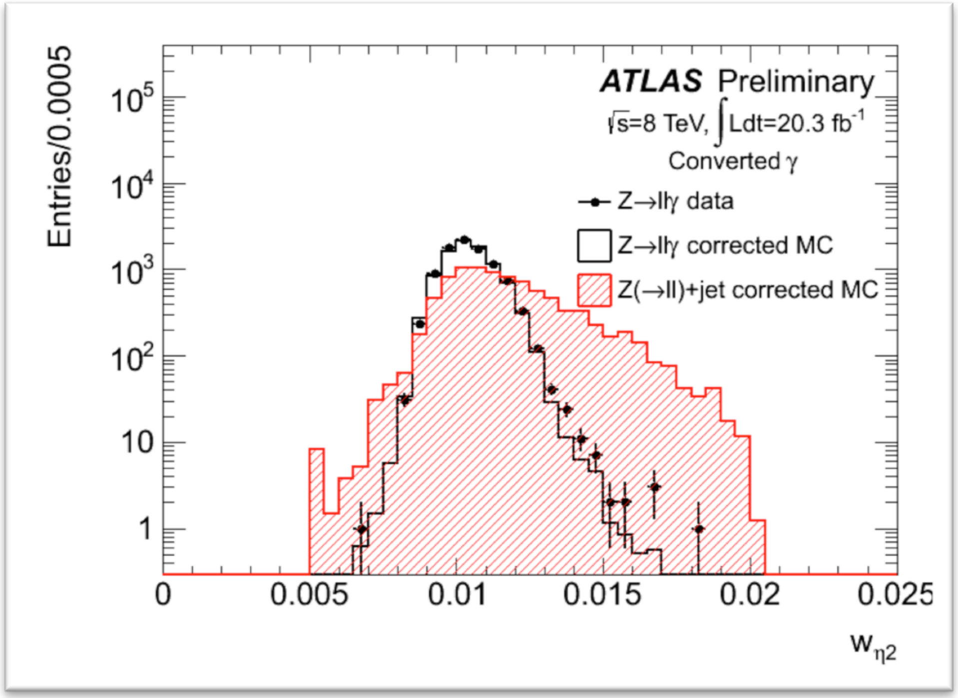
$$f_{\mathrm{side}} = \frac{E_7^{S1} - E_3^{S1}}{E_3^{S1}}$$





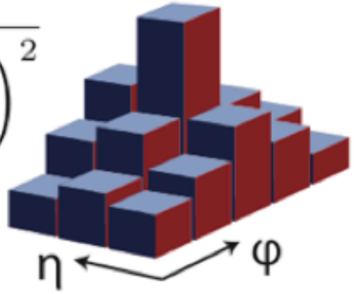
MC in these plots is “corrected” (data-driven shift) to match data

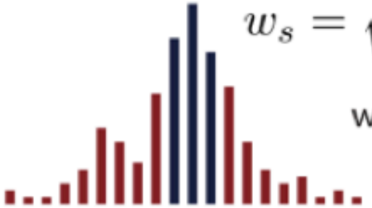
e/γ selection over background: shower shapes



$$w_{\eta,2} = \sqrt{\frac{\sum E_i \eta_i^2}{\sum E_i} - \left(\frac{\sum E_i \eta_i}{\sum E_i}\right)^2}$$

Width in a 3×5 (Δη×Δφ) region
of cells in the second layer.

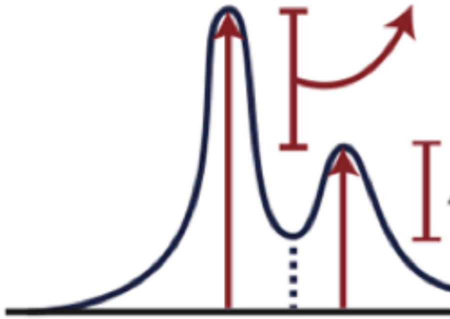




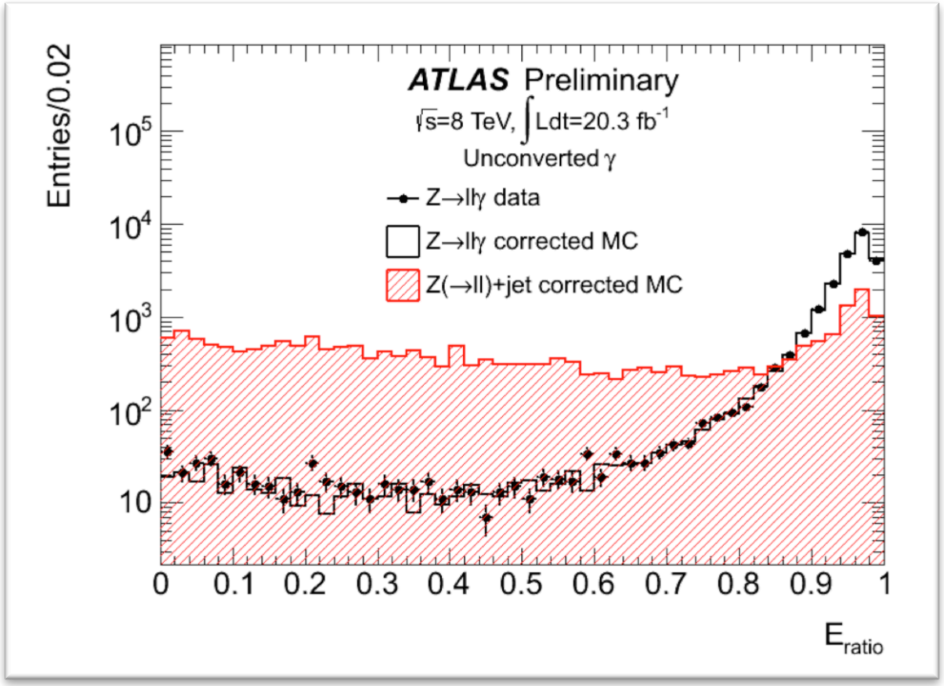
$$w_s = \sqrt{\frac{\sum E_i (i - i_{\max})^2}{\sum E_i}}$$

ws3 = w_1 uses 3 strips in η;
wstot is defined similarly,
but uses 20 strips.

$$E_{\text{ratio}} = \frac{E^{S1}_{\max,1} - E^{S1}_{\max,2}}{E^{S1}_{\max,1} + E^{S1}_{\max,2}}$$



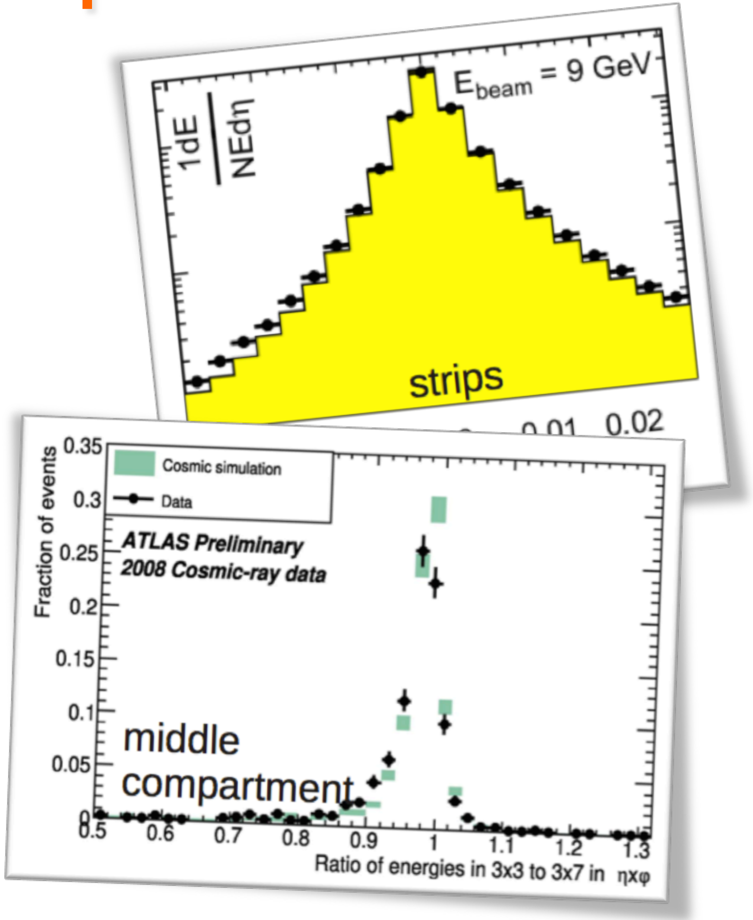
$$\Delta E = E^{S1}_{\max,2} - E^{S1}_{\min}$$

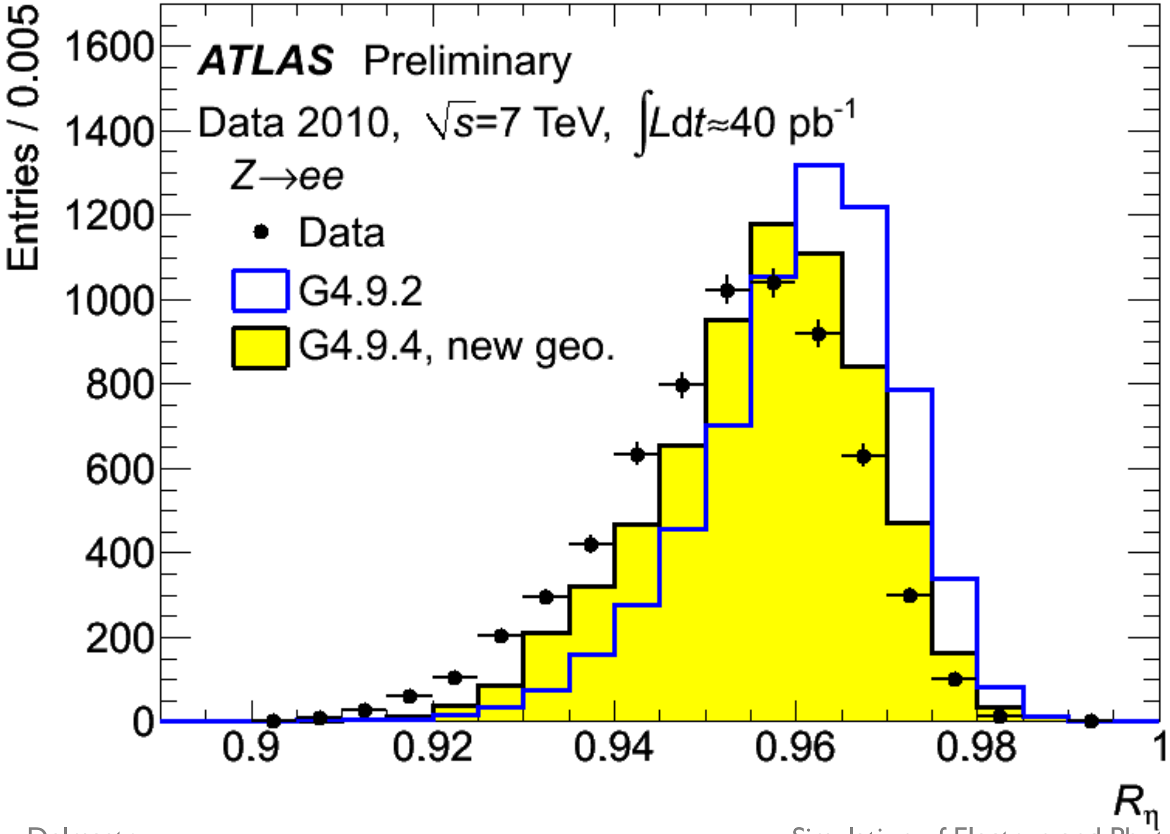


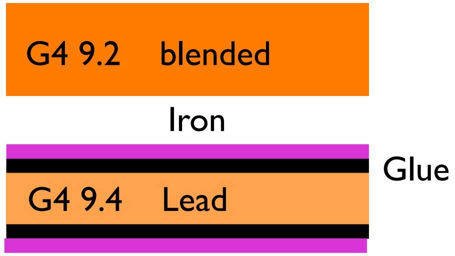
MC in these plots is "corrected" to match data

MC description of EM shower shapes

- ATLAS simulation does not perfectly predict the key distributions for the lateral shower development in collision data, as already observed in test-beam and cosmic data
  - ✓ **EM shower in data are broader then in MC!**
- Part of the disagreement was tracked to a G4 9.2 bug treating blended material, and mitigated by transition to the detailed simulation of LAr accordion






G4 9.2 blended

Iron

G4 9.4 Lead

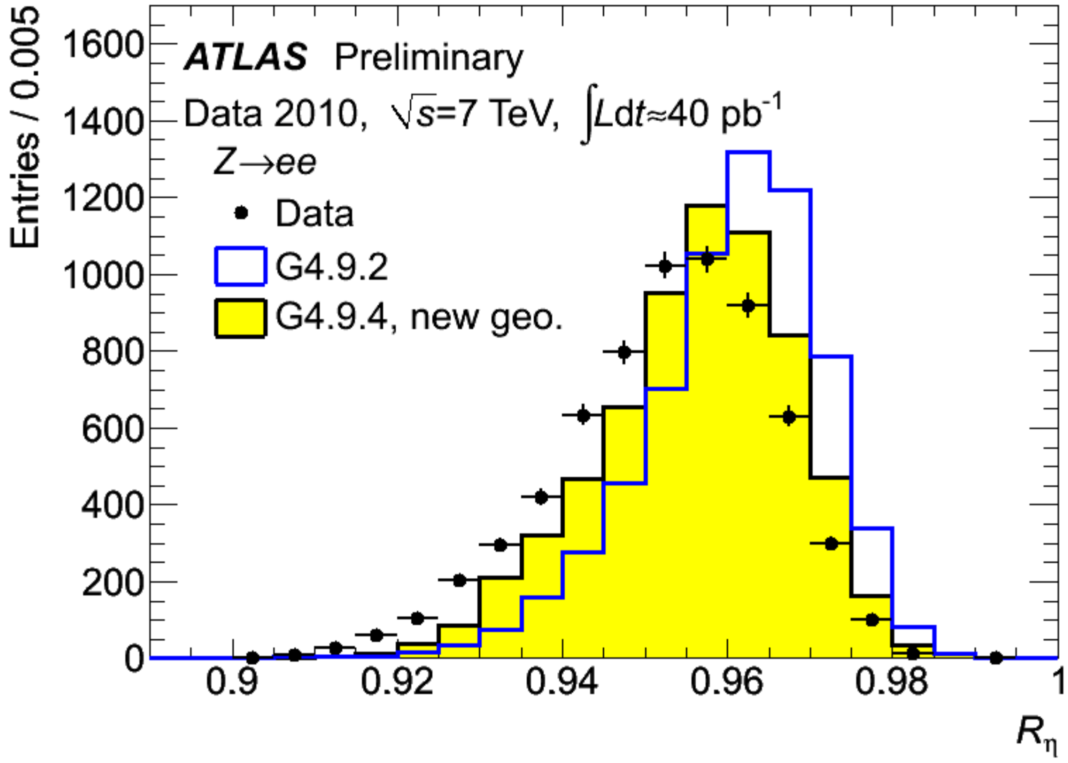
Glue

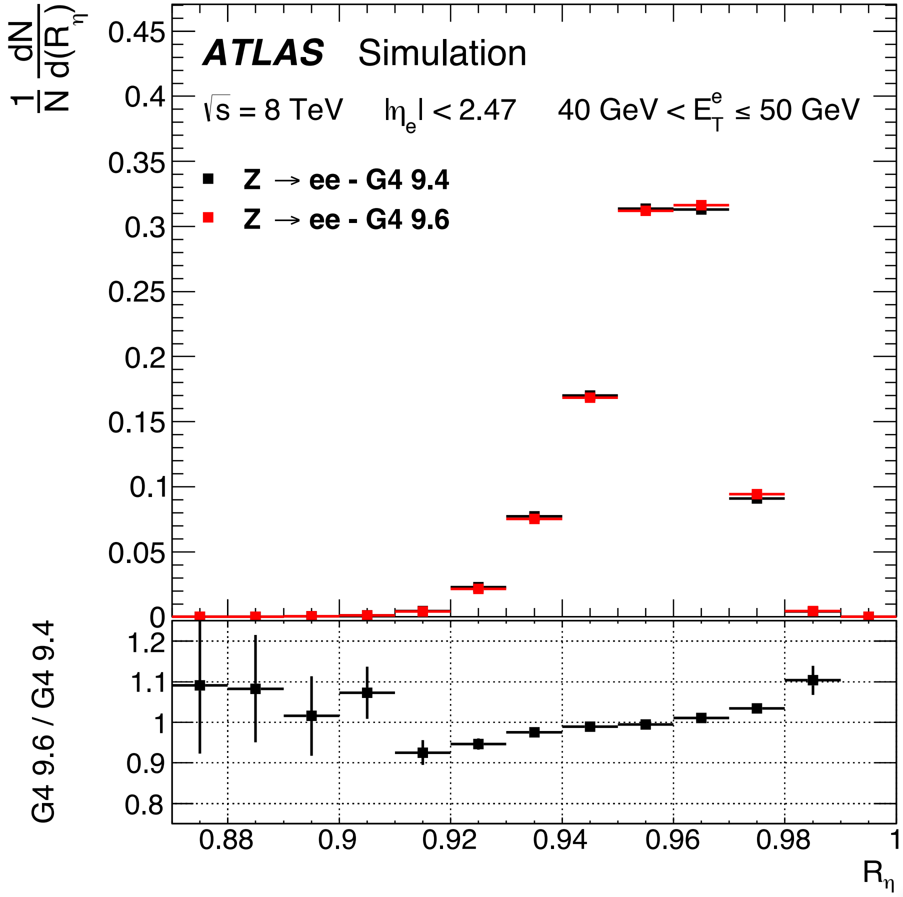
30-60% more CPU time for EM showers

Detailed simulation accounts only for ~50% of the effect…

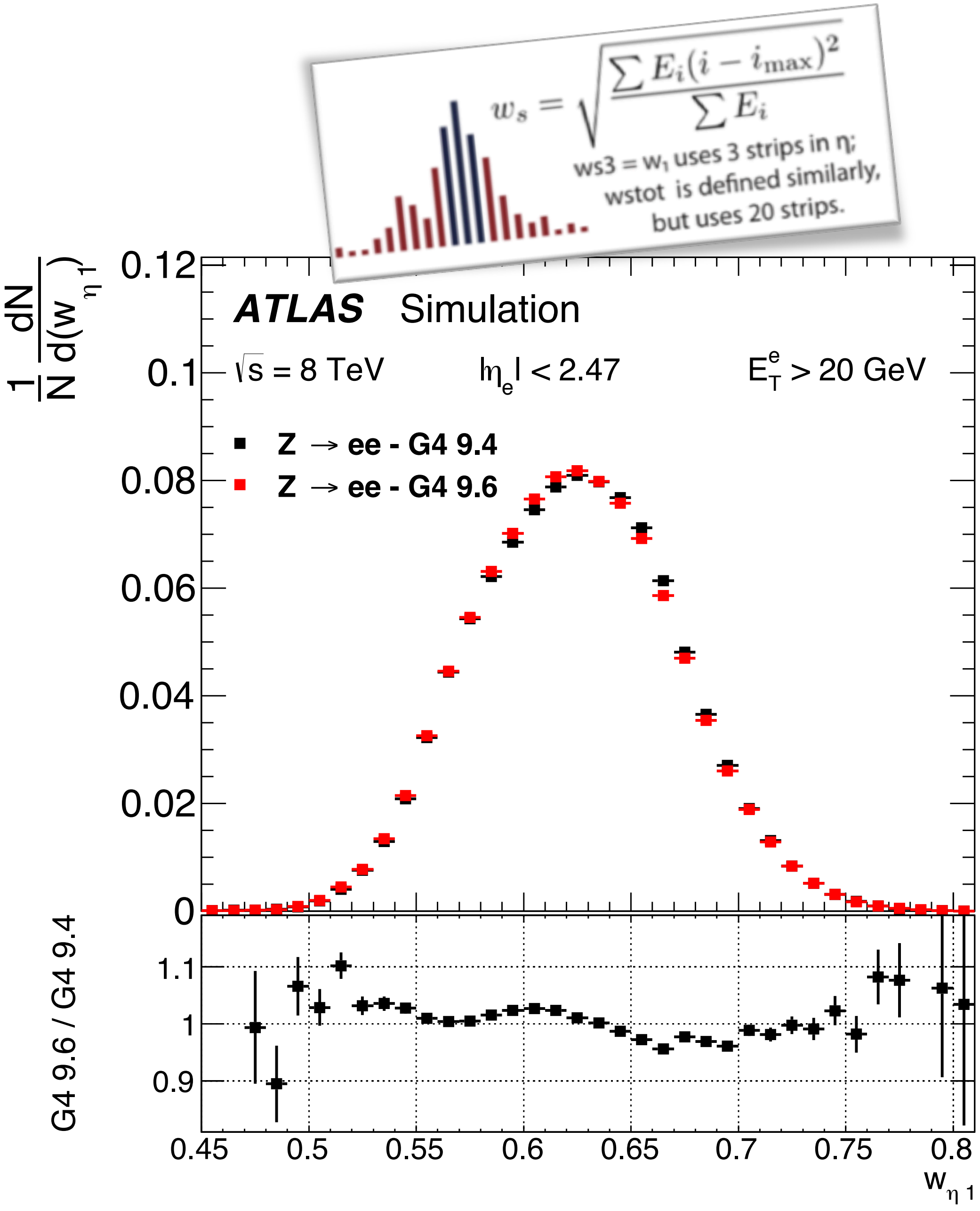
Can disagreement be explained with G4 physics?

- In the past years we have checked impacts of numerous possible sources for remaining discrepancies, not finding any good explanation
  - ✓ Cross-talk, material, geometry (accordion, sagging,...), misalignment, ...
- **One of the candidates among many was the G4 physics...**
  - ✓ G4 9.4 → Urban93 MSc model
  - ✓ G4 9.6 → Urban95 MSc model



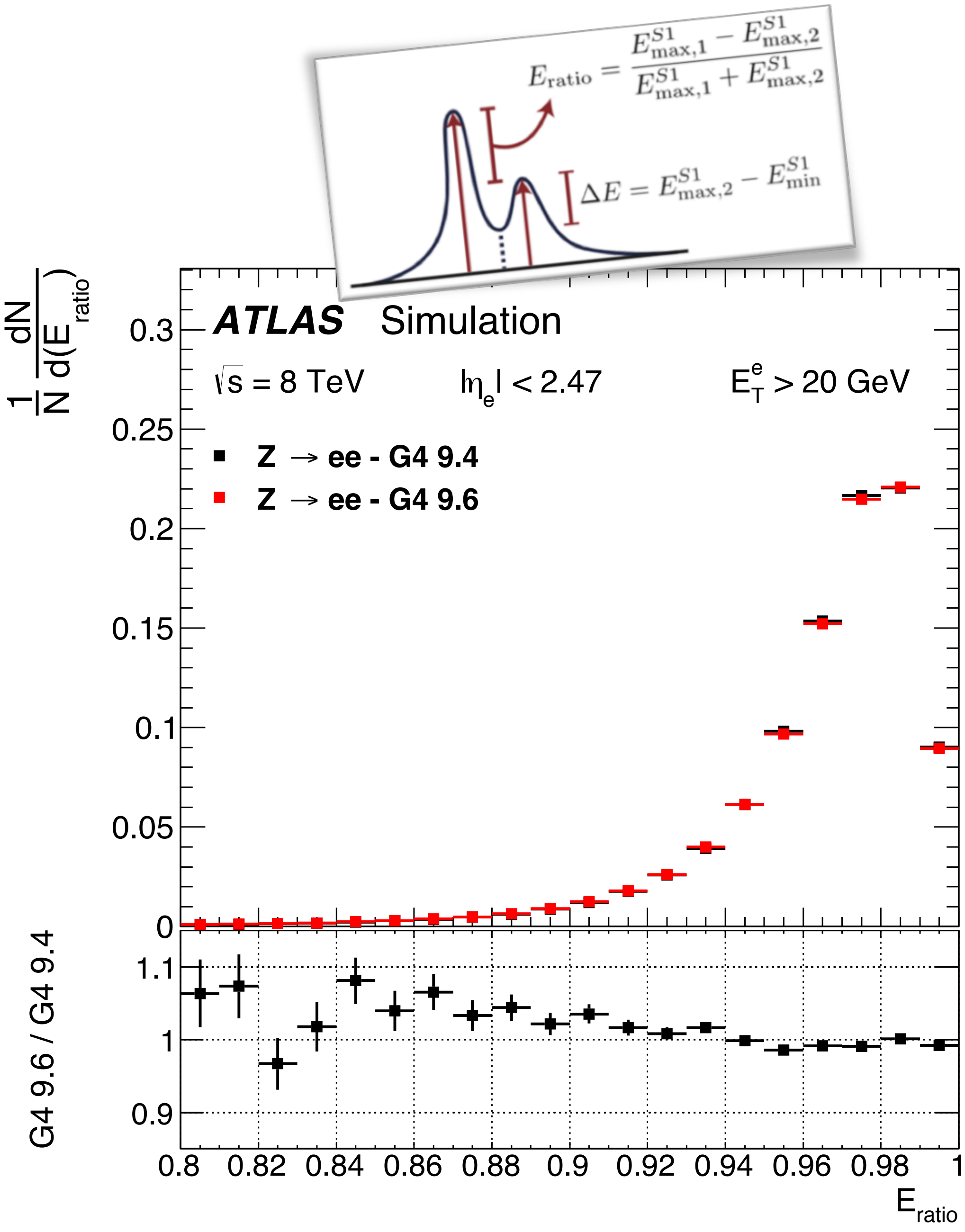


Can disagreement be explained with G4 physics?



$$w_s = \sqrt{\frac{\sum E_i (i - i_{\max})^2}{\sum E_i}}$$

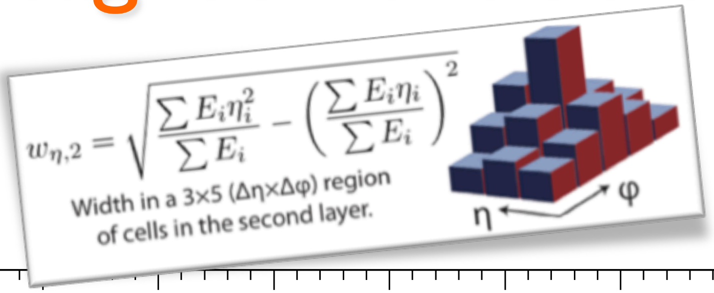
ws3 = w_1 uses 3 strips in η; wstot is defined similarly, but uses 20 strips.



$$E_{\text{ratio}} = \frac{E^{S1}_{\max,1} - E^{S1}_{\max,2}}{E^{S1}_{\max,1} + E^{S1}_{\max,2}}$$

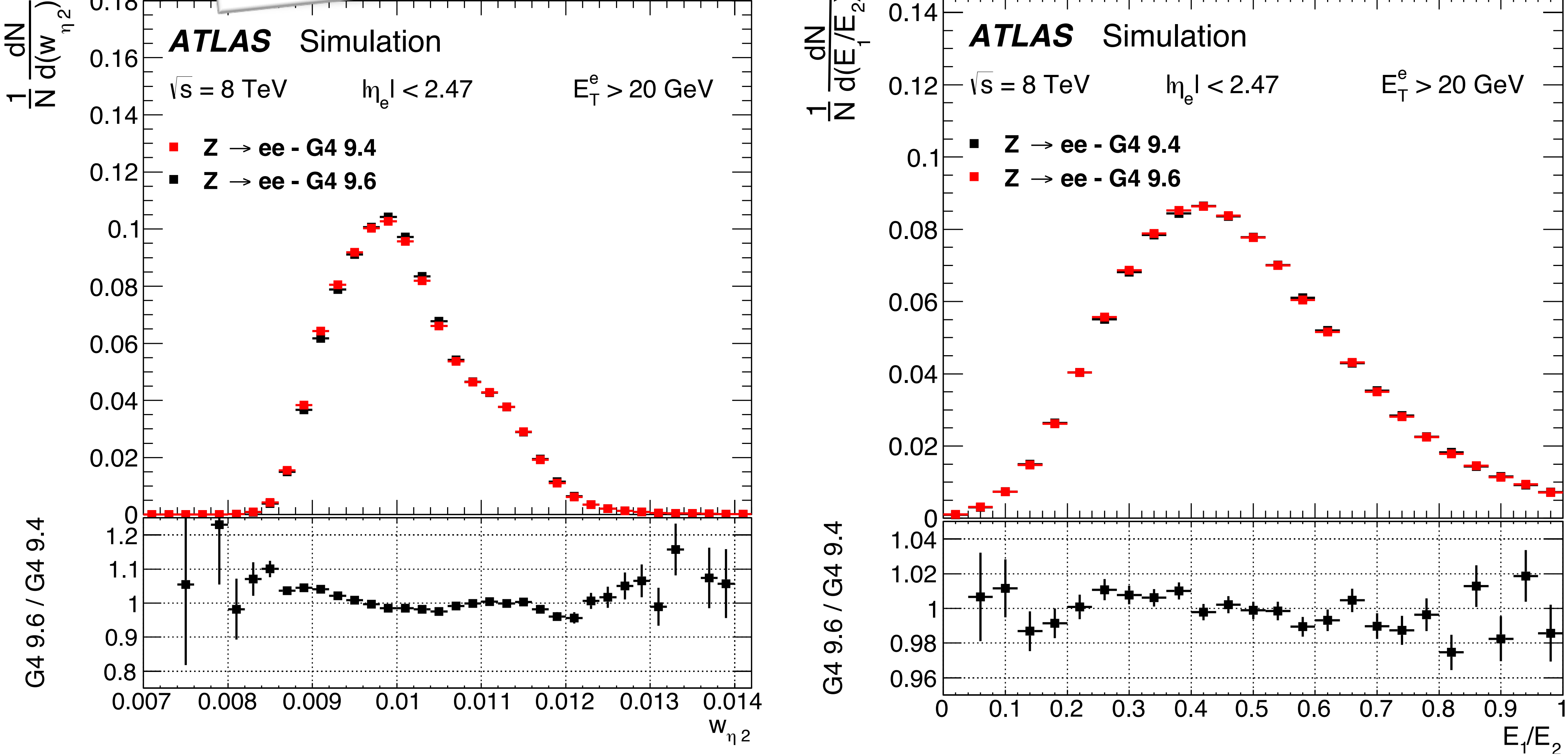
$$\Delta E = E^{S1}_{\max,2} - E^{S1}_{\min}$$

Can disagreement be explained with G4 physics?



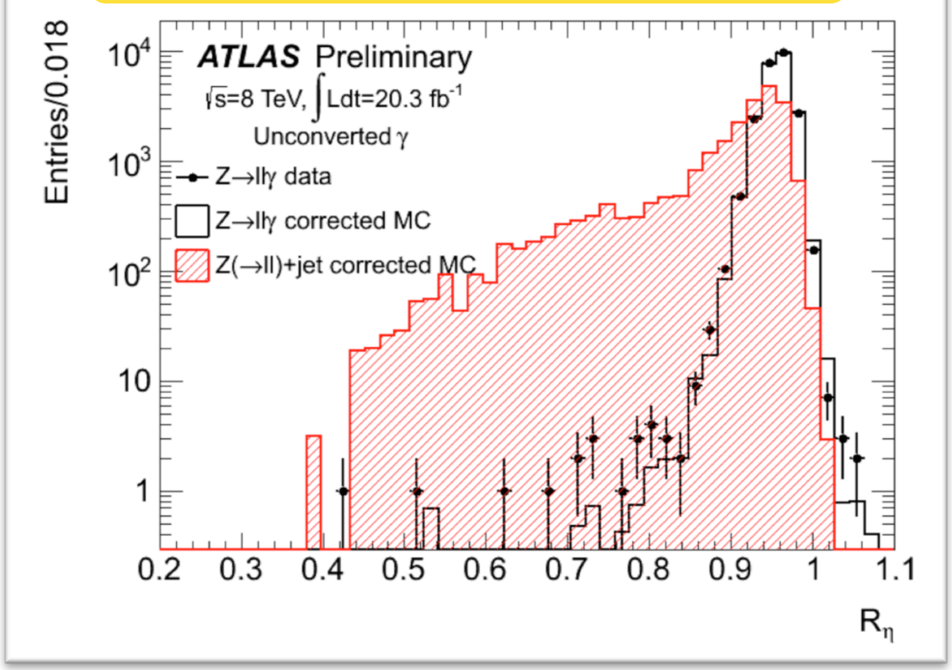
$$w_{\eta,2} = \sqrt{\frac{\sum E_i\eta_i^2}{\sum E_i} - \left(\frac{\sum E_i\eta_i}{\sum E_i}\right)^2}$$

Width in a 3×5 (Δη×Δφ) region of cells in the second layer.



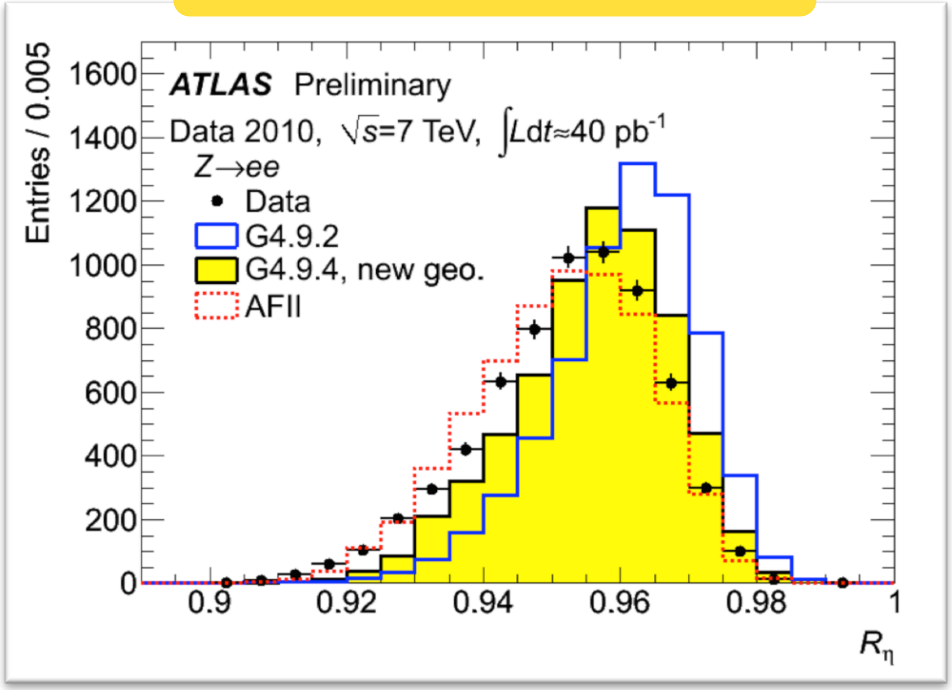
Ways to mitigate the data/MC discrepancies

Data-MC "alignment" using data



- **Shower shapes compared in data and MC**
  - ✓ γ: after background injection in MC following purity measurement, or using Z→eeγ events
  - ✓ e: using Z→ee or J/ψ→ee s-Plots
- **MC "shift" is computed by minimizing χ² between PDF**
  - ✓ Shower shape "alignment" is applied to signal MC, dramatically reducing efficiency data/MC differences
  - ✓ Data-driven measurements provide SF and uncertainties

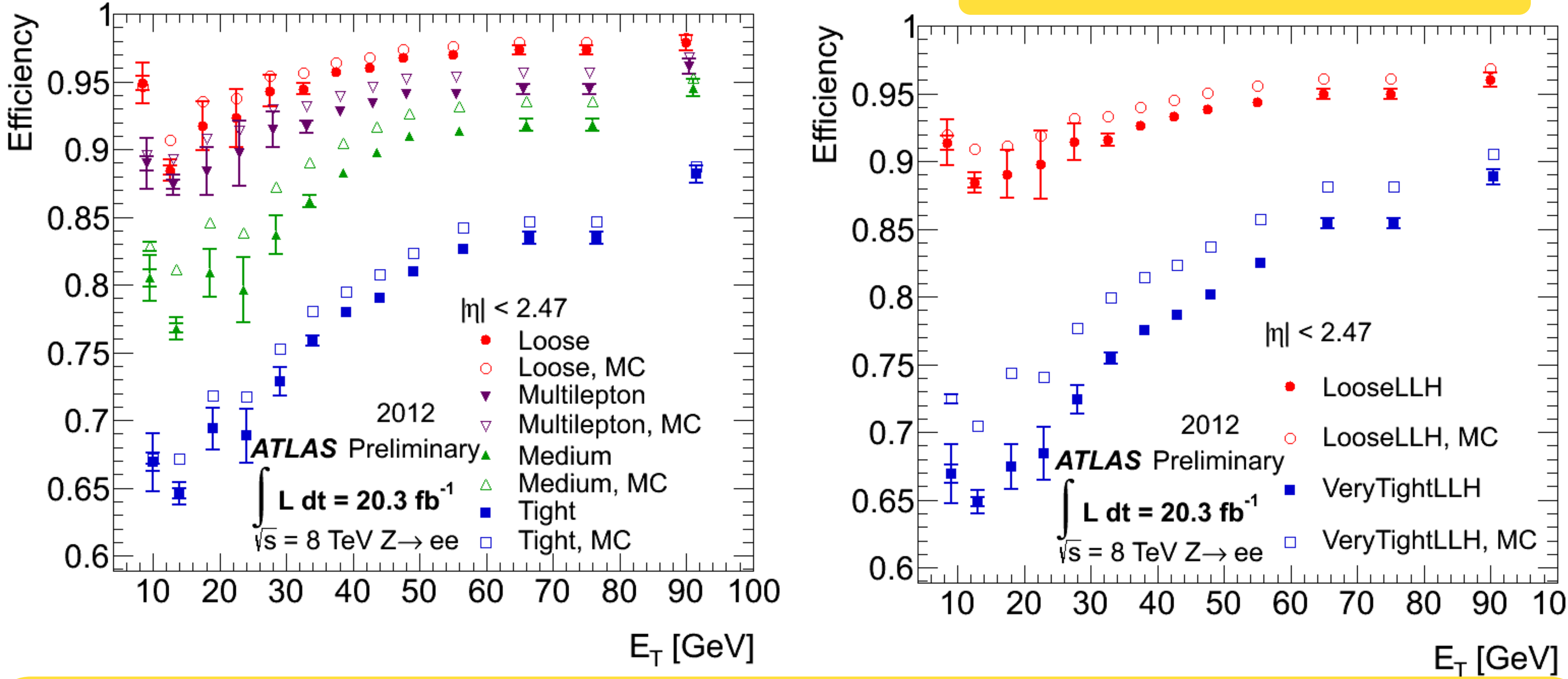
ATLAS Fast Simulation



- **G4 simulation of ID, FastCaloSim for calorimeters**
  - ✓ All calorimeter layers, with full granularity
  - ✓ Based on pre-data-taking single particles (e,γ, π±
  - ✓ Parameterization of energy response and resolution
  - ✓ Parameterization of average (i.e no fluctations) ateral shower shape
- **Fast and reliable for most studies!**
  - ✓ Data-driven measurements provide SF and uncertainties

Impact of data/MC shower shape differences

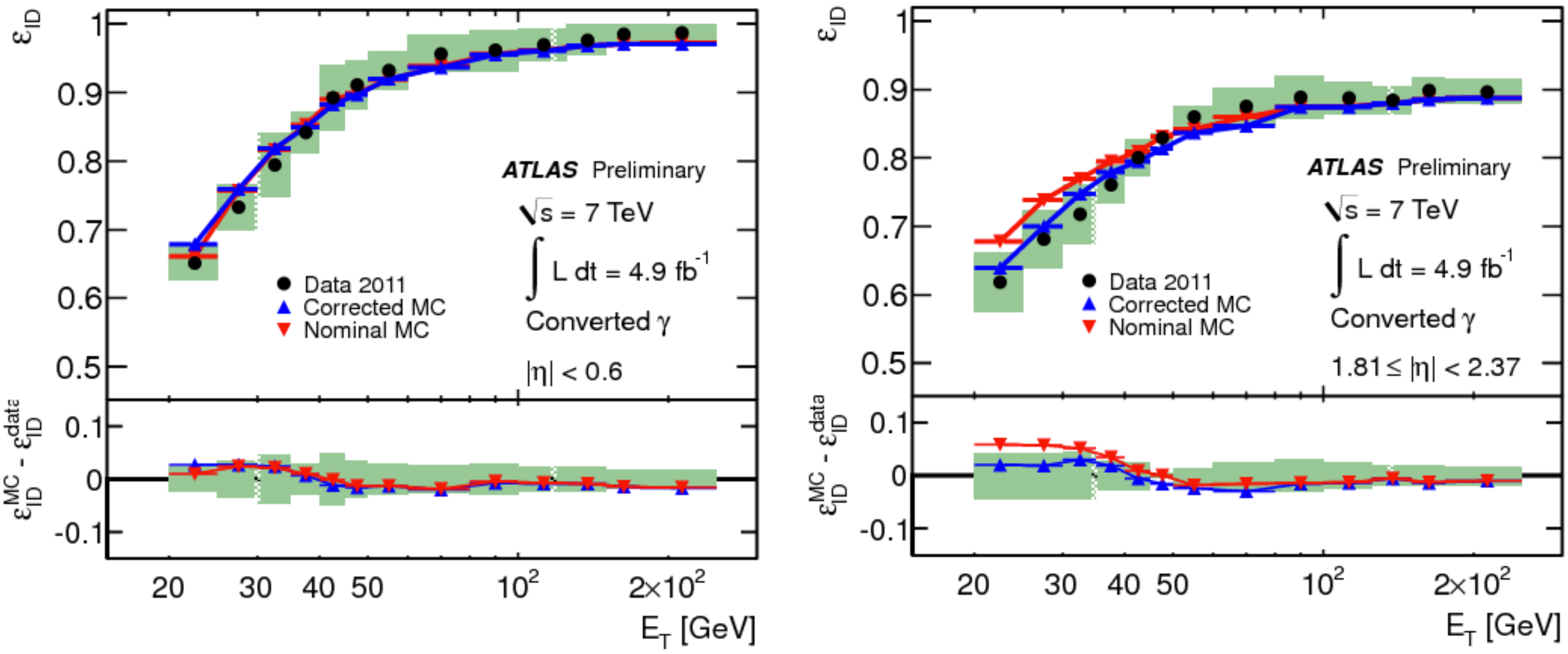
Electron identification efficiencies



- **When no correction to shower shapes is applied, efficiencies can show up to a few % differences between data and MC**, depending on tightness of selection
  - ✓ Difference is measured in data and accounted for with MC scale factors
- **More complex selection using full shower shape distribution** (e.g. LLH, MVA) **show larger discrepancies** (as expected)

Impact of data/MC shower shape differences

Photon identification efficiencies



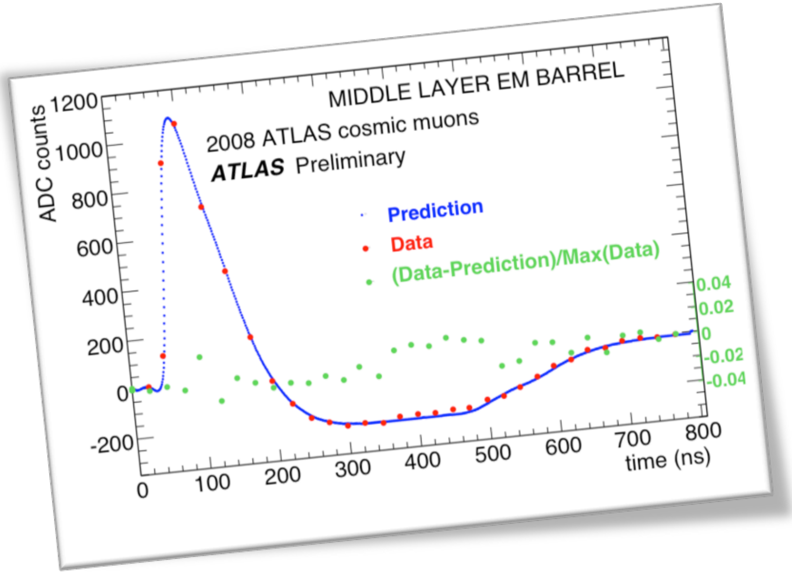
- **When “data-MC alignment” is used, efficiency difference is greatly mitigated**
  - ✓ Even in difficult regions!
  - ✓ Residual difference is measured in data and accounted for with MC scale factors

Summary and perspectives

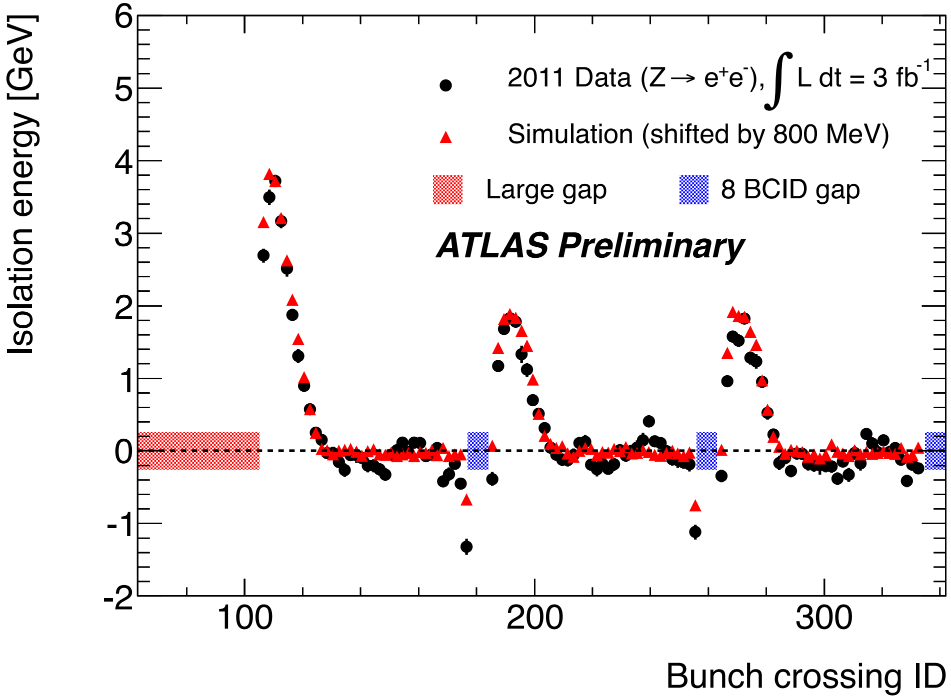
- **Electrons and photons in ATLAS are well describes by MC simulation at all levels**
  - ✓ e.g. reconstruction, identification, …
- **ATLAS heavily rely on simulation to calibrate its response to EM showers**
  - ✓ Material mapping is a crucial ingredients, and has progressed a lot thanks to the analysis of Run 1 data (today constrained at 2-5% X_0 level)
  - ✓ More iteration expected toward definition of geometry for Run 2 simulation
  - ✓ Understanding of Z→ee lineshape fundamental for in-situ determination of EM scale
- **Lot of work done in synergy with G4 team to understand impact of physics models on EM response**
  - ✓ e.g. conversion, multiple scattering
- **Long-standing issue of imperfect description of EM lateral shower development still there**
  - ✓ Partially mitigated by detailed description of LAr absorber
  - ✓ Effective corrections used (or being developed)
  - ✓ Fast simulation tuned on data is one aspect of solution, to be developed further

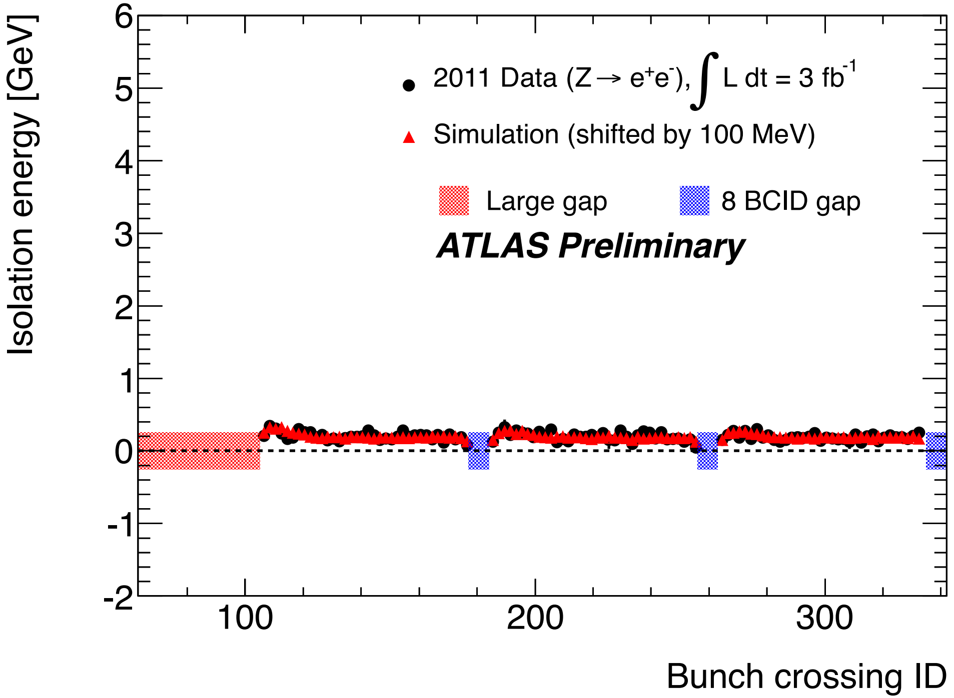

THIS MIGHT LOOK LIKE AN ORDINARY POWERPOINT SLIDE.
REVENUE
$
BUT IT IS ACTUALLY A PORTAL TO ANOTHER DIMENSION IN WHICH FANTASY AND REALITY HAVE TRADED PLACES.
REVENUE
$
STOP PLAYING WITH MY SLIDES.
BEWARE THE HORNED BEAST THAT CROSSES OVER.
$

In-time and out-of-time pileup effect



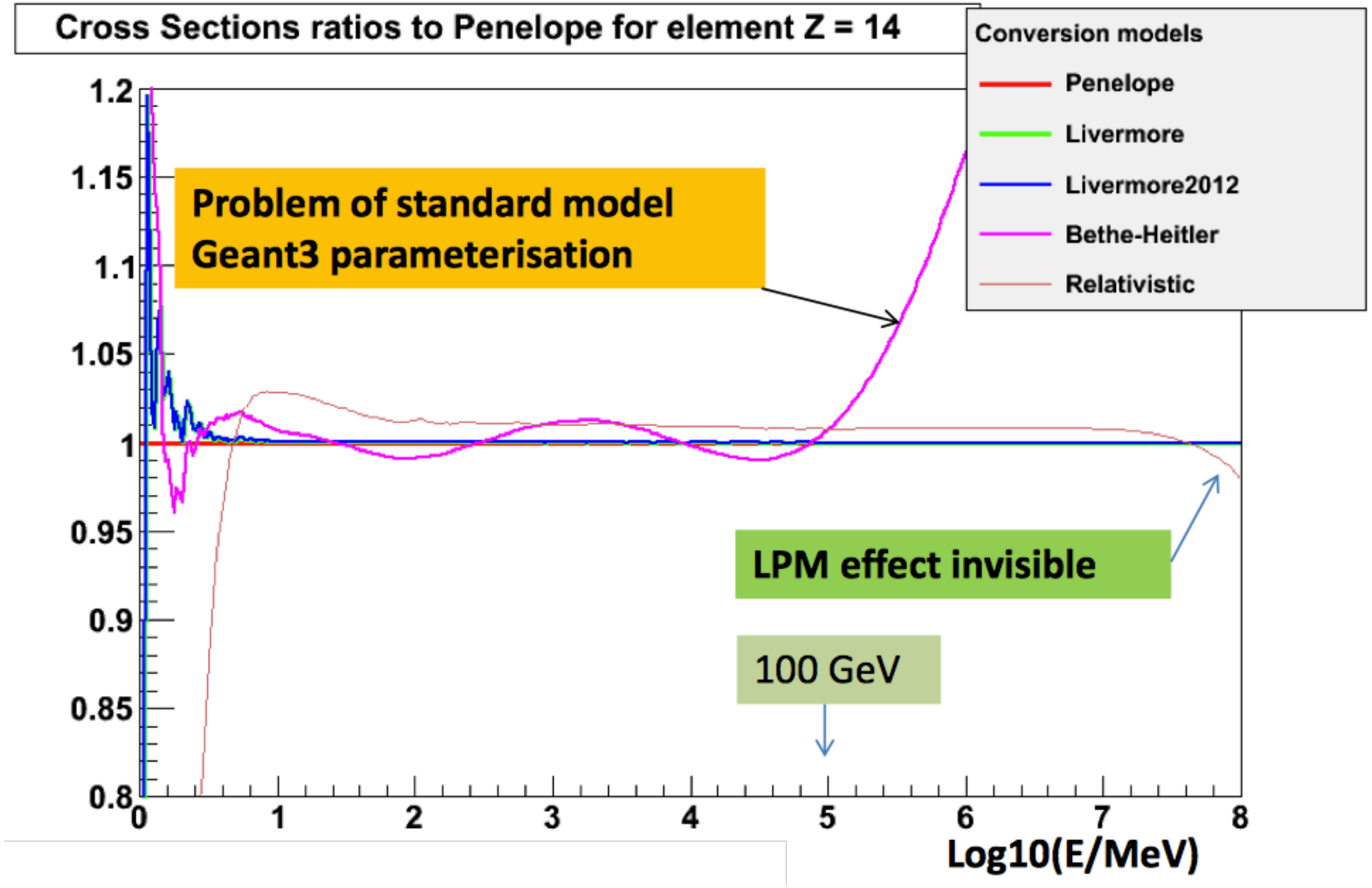
- **Ideally, infinite bunch trains with same luminosity per bunch**
  - ✓ Average energy per cell = 0
  - ✓ Pileup fluctuations translate in noise per cell
    - dominant w.r.t. electronics noise for eta>2.5
- **In real life, bunch train structure and bunch-to-bunch luminosity variations**
  - ✓ No cancellation of in-time pileup shifts for first bunches
  - ✓ Residual shifts for bunches with luminosity different from average
  - ✓ Only the first effect simulated in ATLAS MC
- **Correction computed from pulse shape, OFC, bunch-per-bunch luminosity (data) or bunch structure (MC), and average energy per cell for in-time only energy deposit**





G4 photon conversion models

- http://indico.cern.ch/event/216885/session/0/material/slides/1?contribId=1 (V.N.Ivanchenko)
- Since Geant4 9.4 ultra-relativistic conversion model exists taking LPM effect into account
  - ✓ Not considered for LHC because LPM effect is important at higher energies
  - ✓ Included in Geant4 9.6 in all EM constructors above 80 GeV
- Gamma conversion cross section ratio for Silicon between different Geant4 models



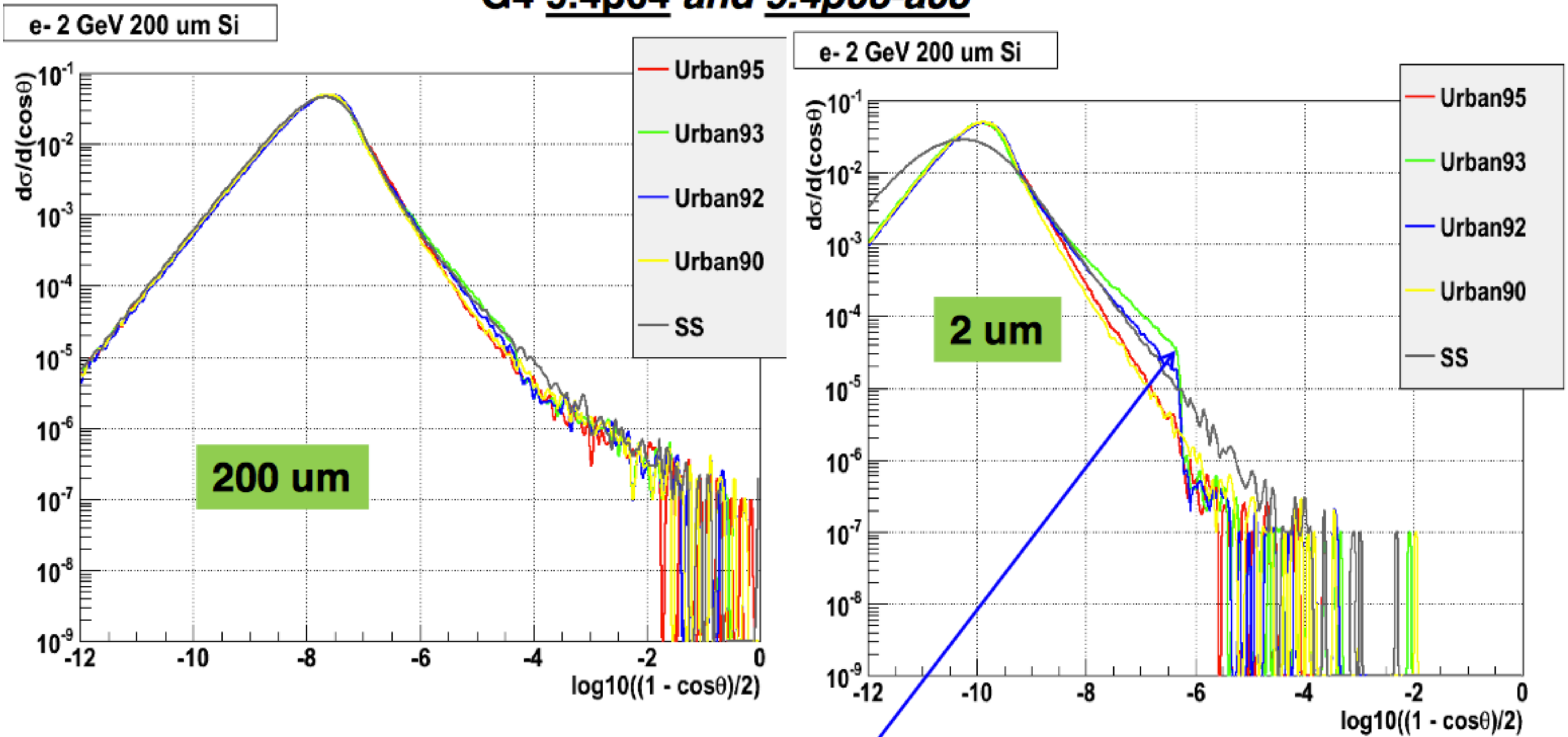
Problem of standard model Geant3 parameterisation

LPM effect invisible

100 GeV

G4 multiple scattering models

G4 9.4p04 *and 9.4p03-a03*



Effect is not visible for larger step size

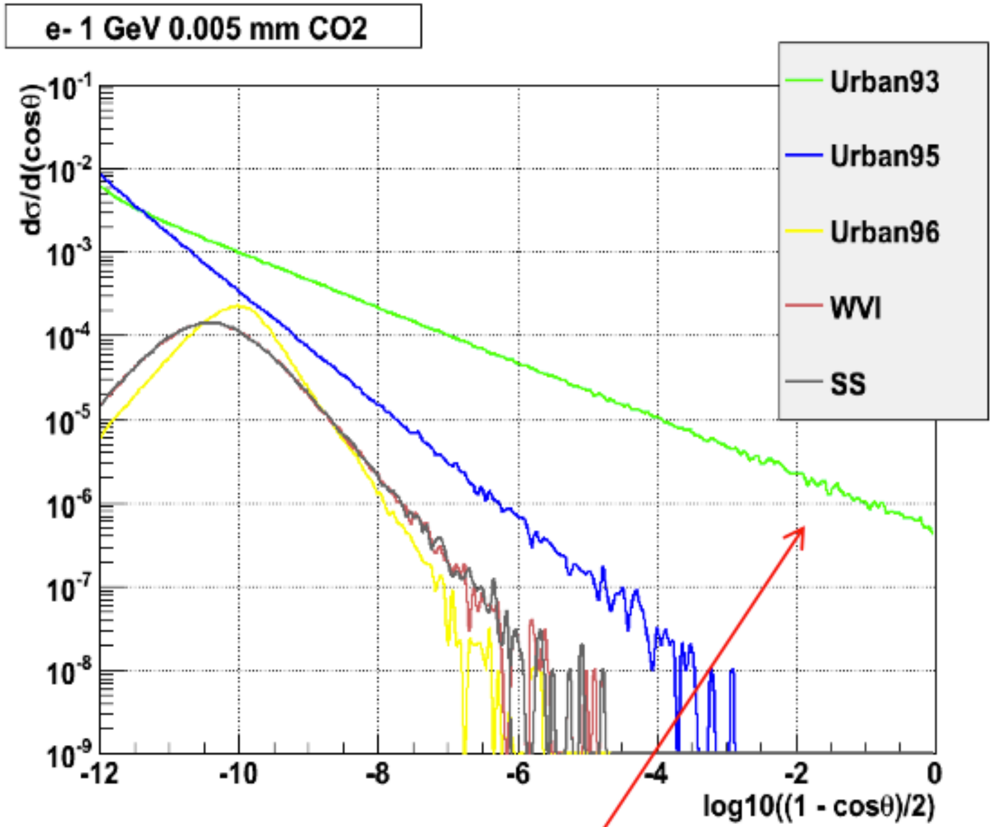
9.5p01

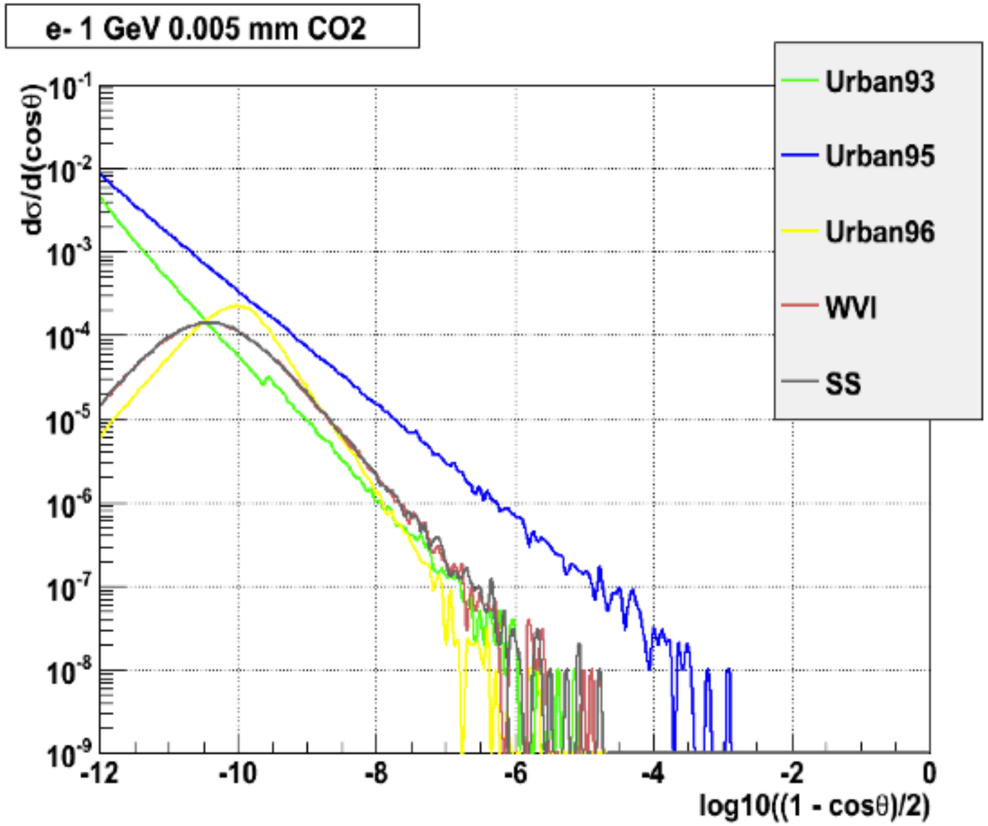
Protection is Cut to reduce the excessive tail of Urban 93

G4 multiple scattering models

Msc93 no fix 5 um step

Msc93 new fix





Msc93 tail is too big
It is the main problem of the model

Msc tail similar to single scattering