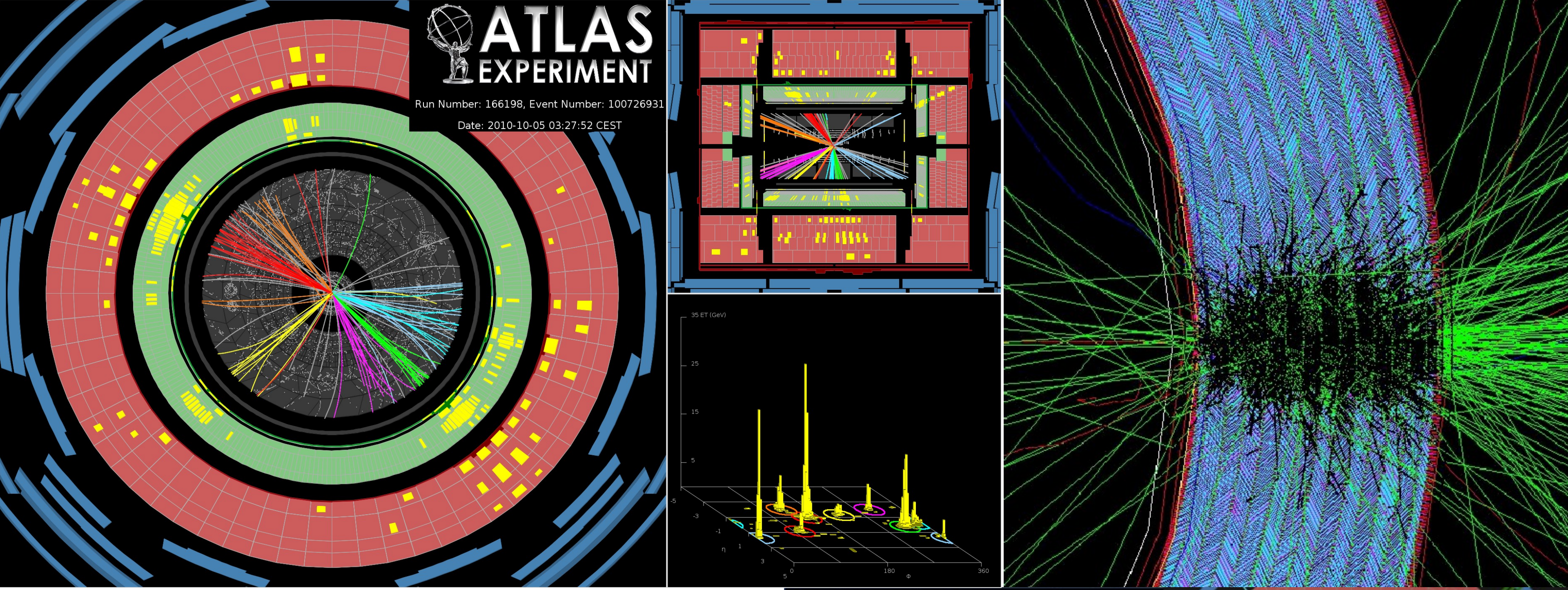
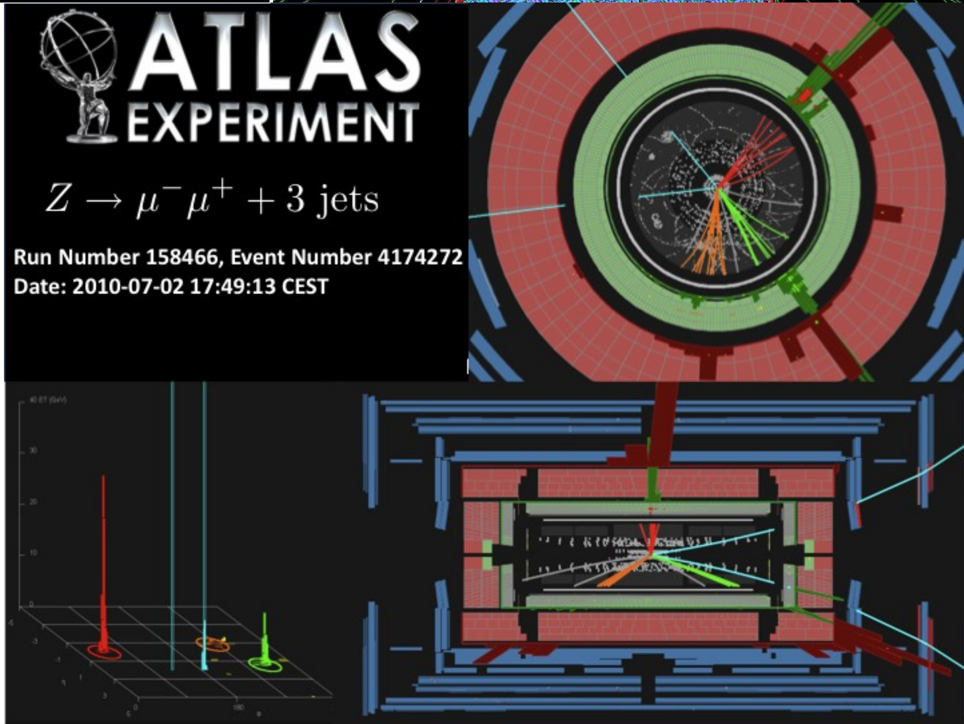


ATLAS Data/MC Comparisons in Hadronic Physics

Zach Marshall (CERN)
For the ATLAS Collaboration

LPCC Simulation Workshop

6 October 2011



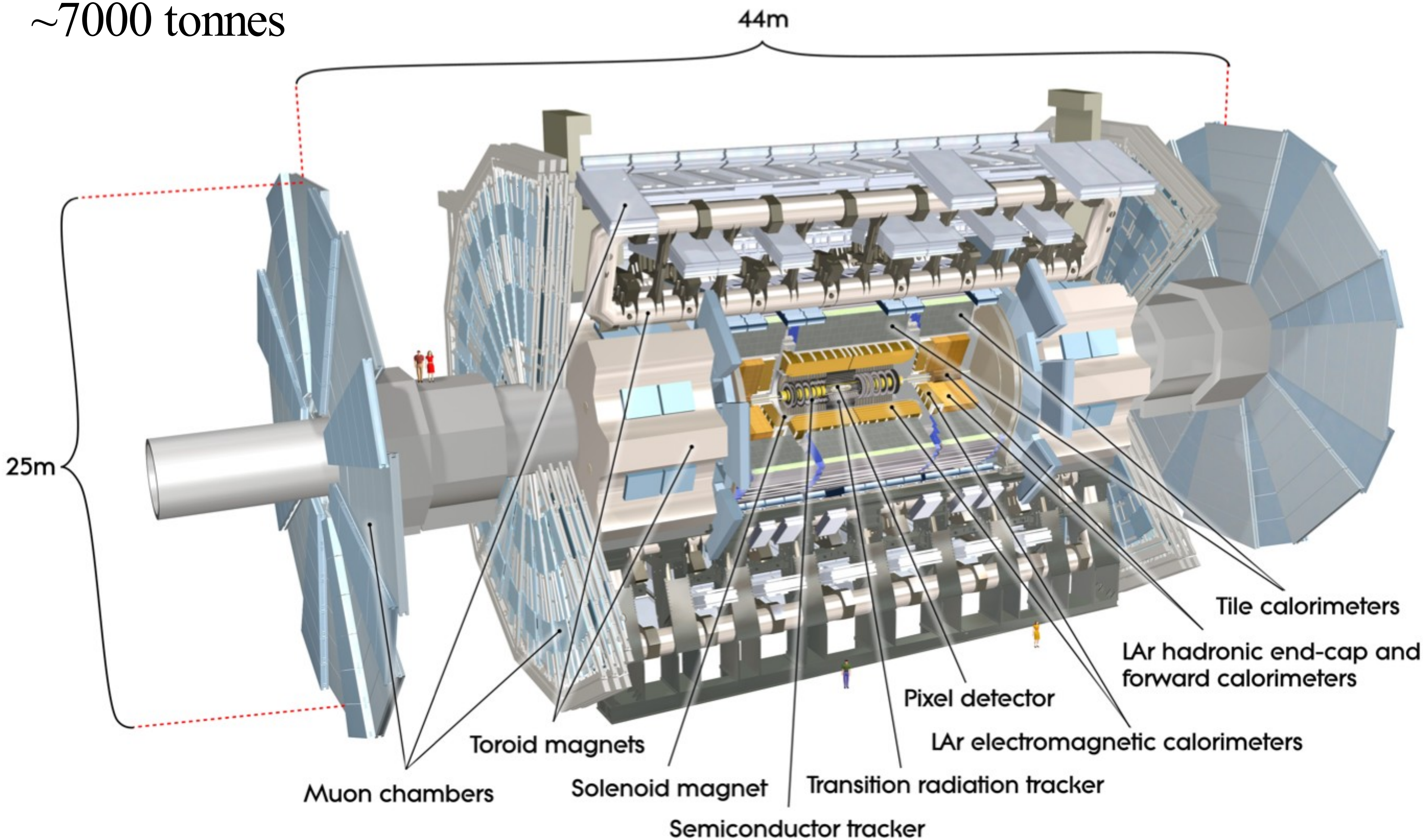
Overview

- Some reminders about the ATLAS calorimetry
- Single particle measurements
  - Response in test-beam and in situ
  - Resolution in test-beam and in situ
  - Shower shapes in test-beam
- Extrapolation from simple (pions) to complex (jets)
  - Particle composition
  - Loosening isolation
  - Pile-up contributions

- ATLAS MC simulation is described in this note
- ATLAS generator tuning is described in this note
- Other results will be linked in the talk

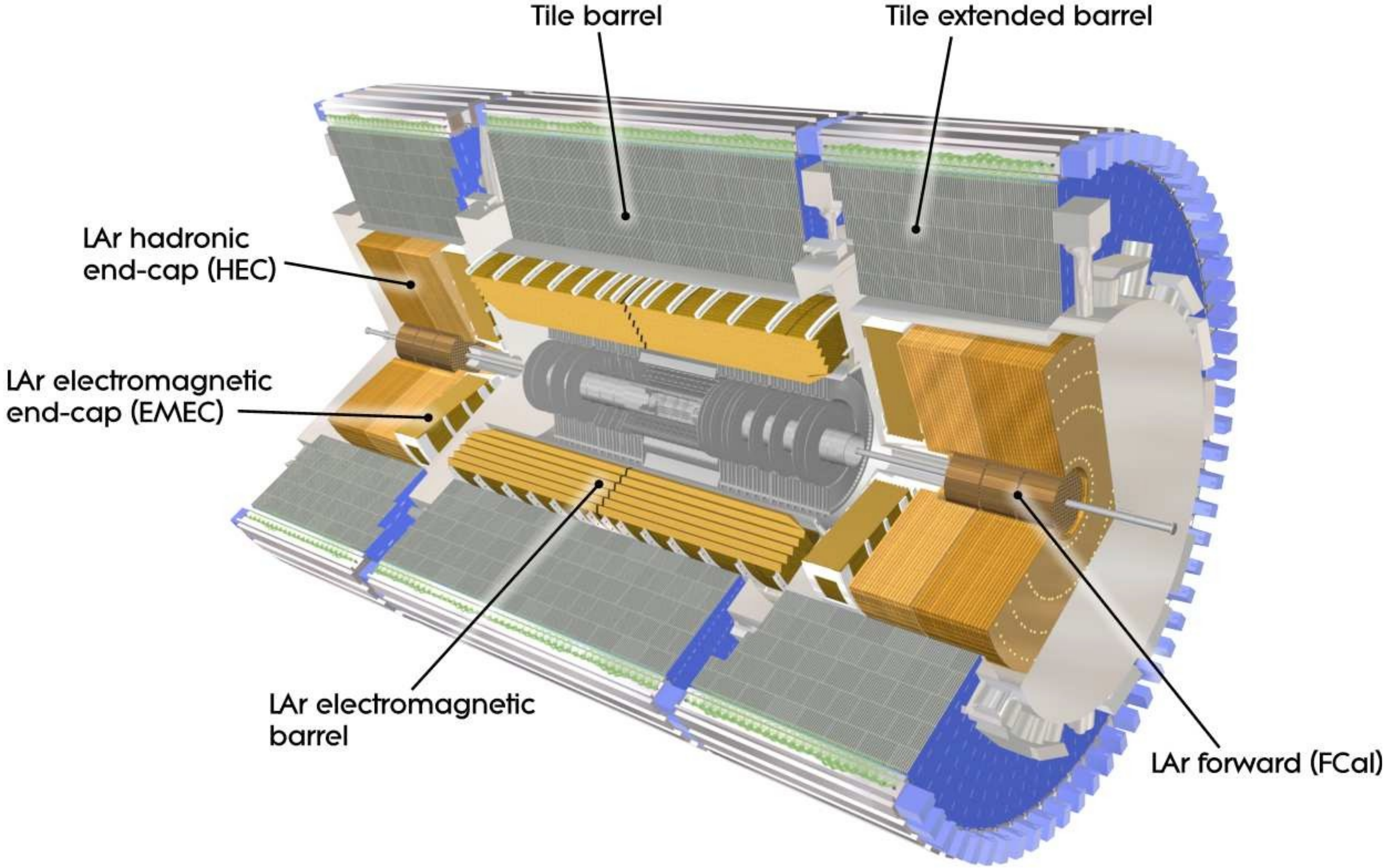
ATLAS

~7000 tonnes



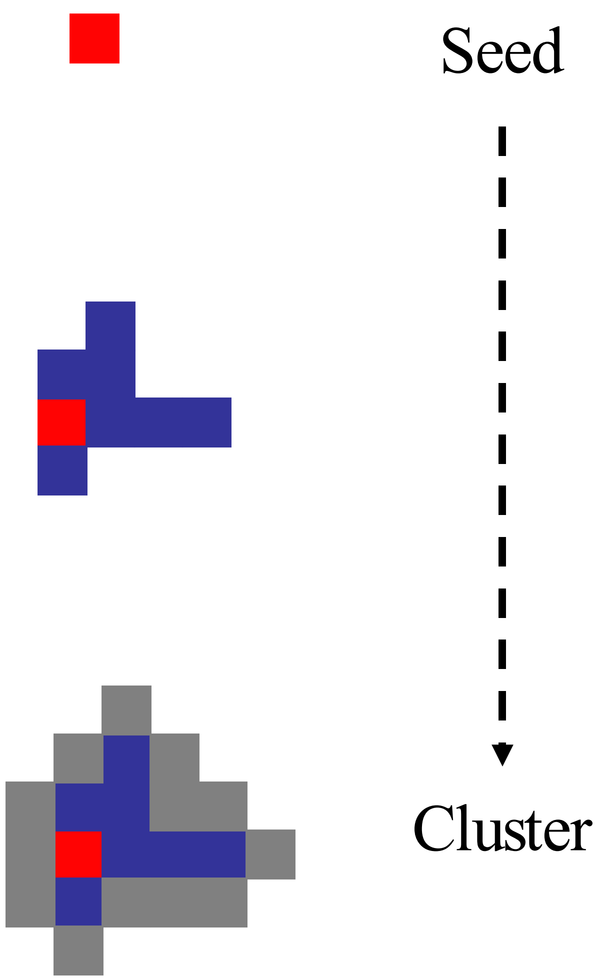
ATLAS Calorimetry

- Designed for hermeticity (to $|\eta|$=4.9), 1% jet energy scale uncertainty, $50\%/\sqrt{E} \oplus 3\%$ central jet resolution, and $50\% \times \sqrt{(\Sigma E_T)}$ MET resolution



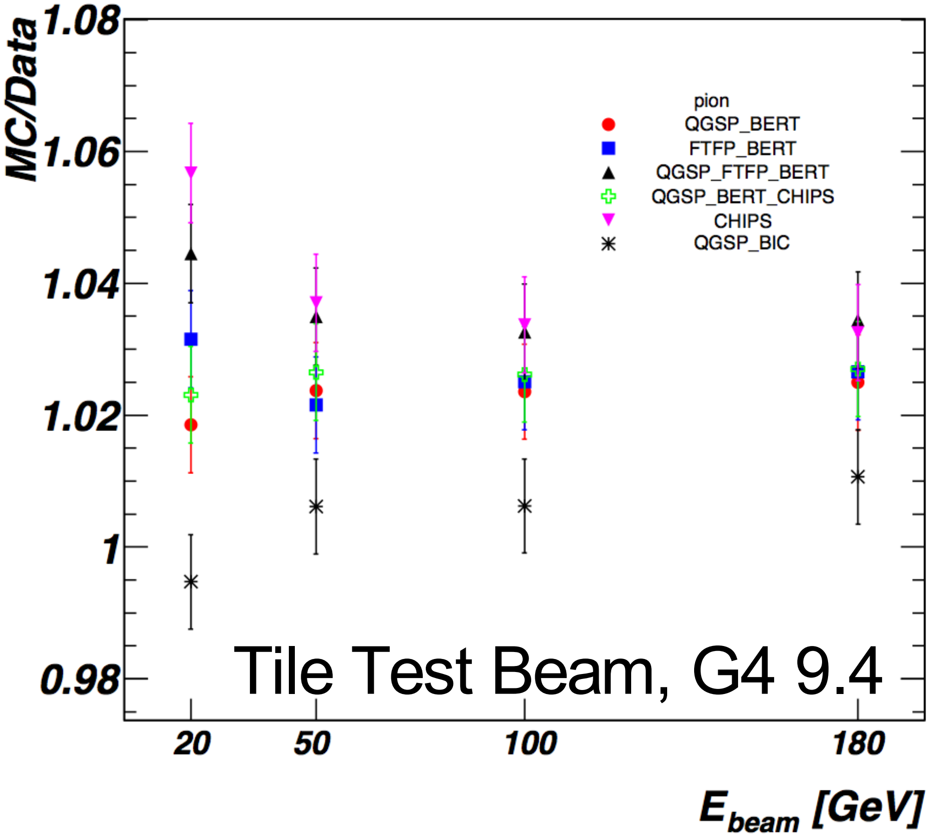
Calorimetry Reminders

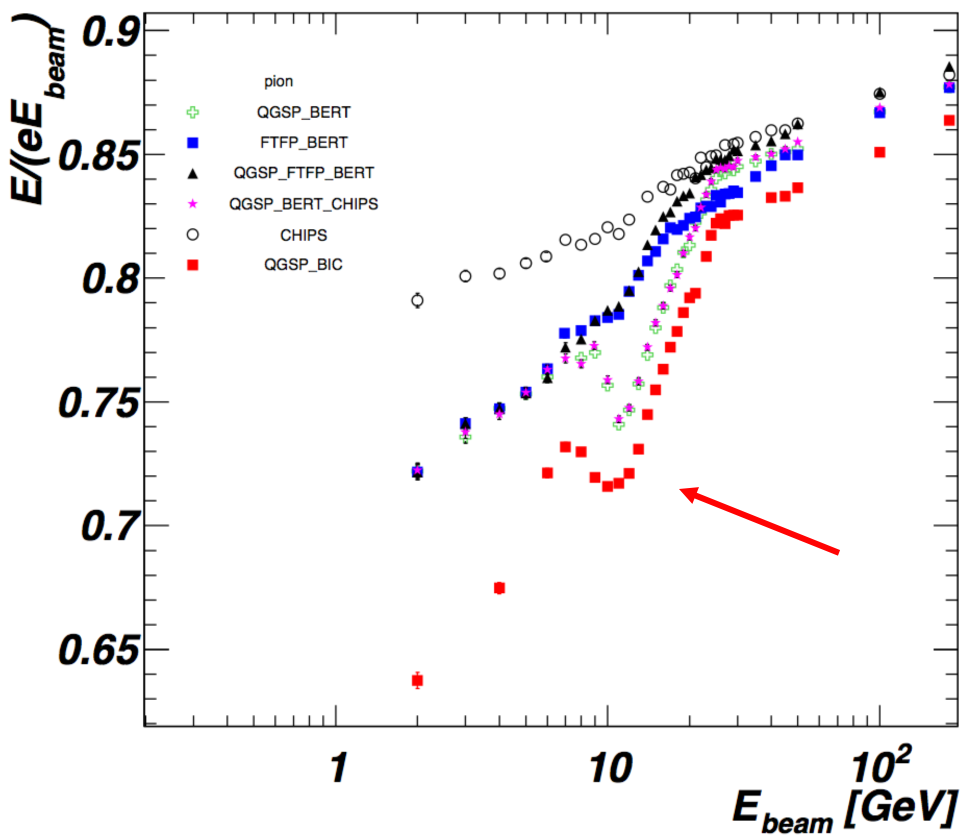
- ATLAS uses a three-step clustering algorithm
  - Our way to deal with noise
  - Seed with $|E_{cell}|>4\sigma$
  - Add cells with $|E_{cell}|>2\sigma$
  - Final layer to improve resolution
  - Rough claim: 1 cluster = 1 particle
    - Breaks down in some cases
    - Classify clusters as "EM" or "Hadronic" – more on this later
- We *almost never* measure the energy that G4 deposited
  - More or less because of noise, clustering effects, and backgrounds
  - Shower shapes are *cluster* shapes!
  - Go to test beam!!



Energy Response

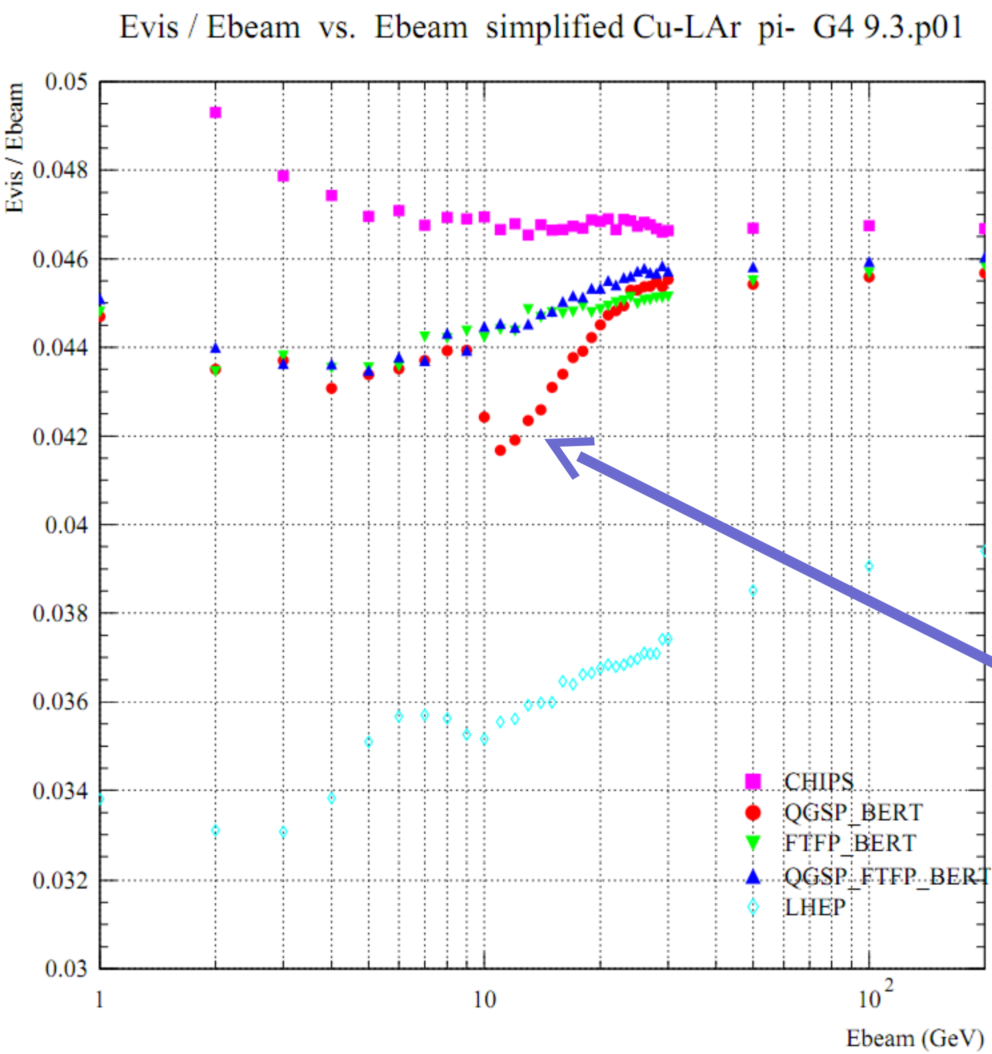
- Obvious first choice: single pion response vs physics list
  - Energy measured in the calorimeter / track p (or beam energy) = E/p
    - Only measured within the tracker coverage, $|\eta|<2.5$
  - Agreement is within a few percent for most lists
- Fine binning reveals some nasty features – very hard to spot in-situ!

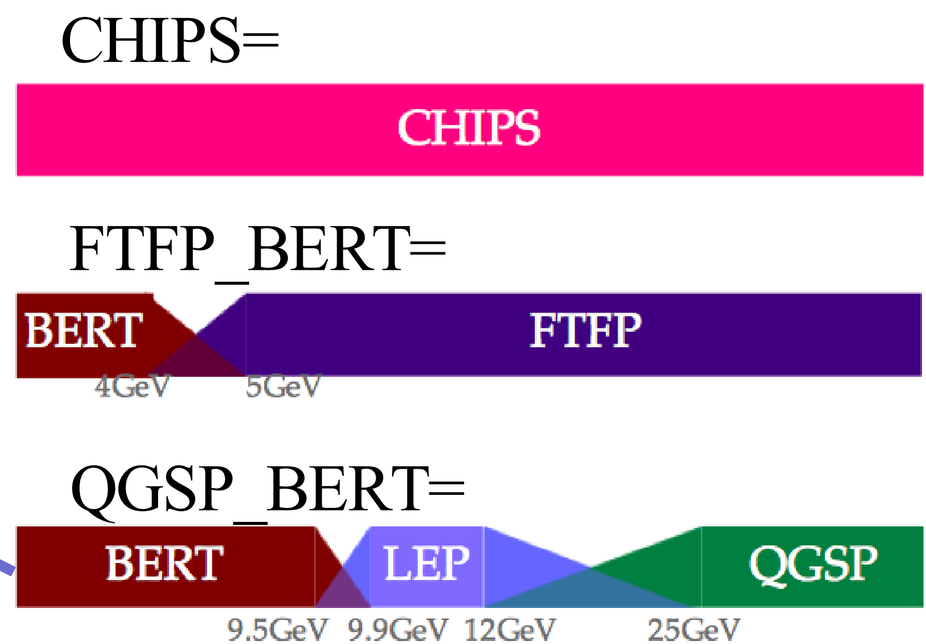




Response Discontinuity

- We knew about that one: entirely to do with transitions between different models!

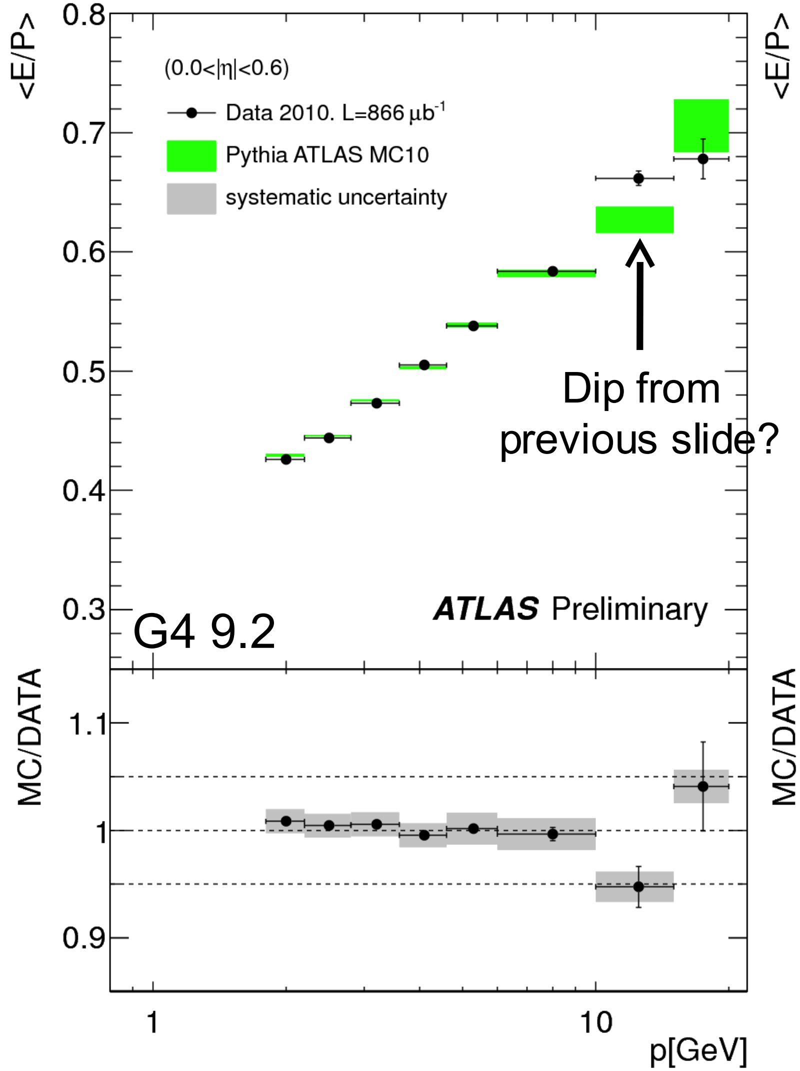


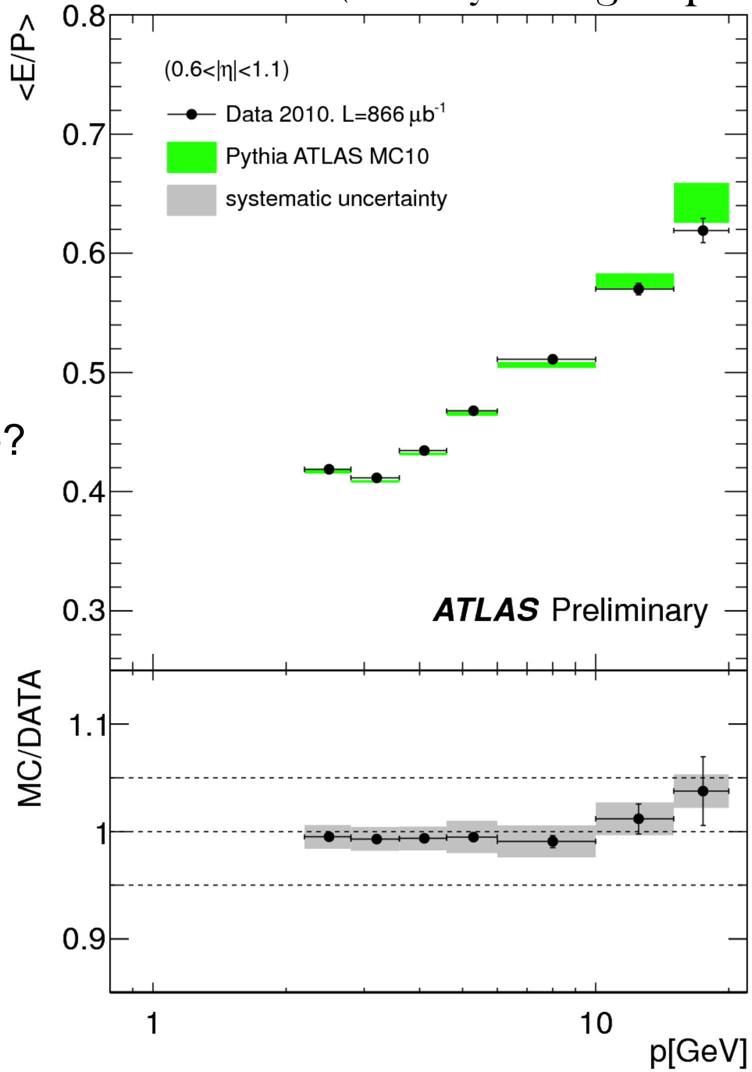


Described in detail in an LCG Simulation Note

Energy Response (II)

- Next, move to *in situ* measurements
  - Here for *inclusive isolated hadrons* (mostly charged pions)





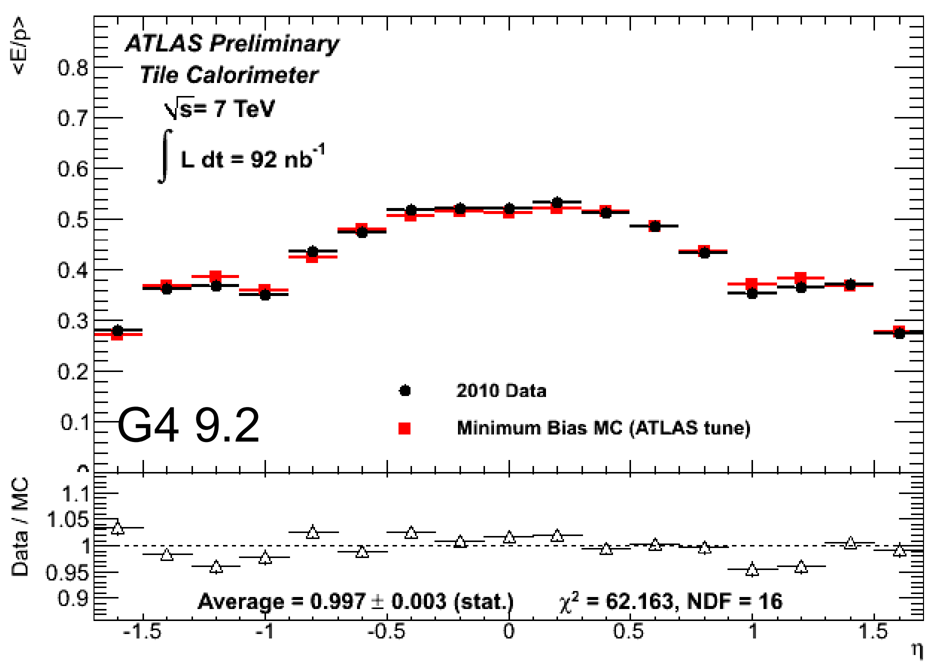
High p:
not isolated, so
low statistics

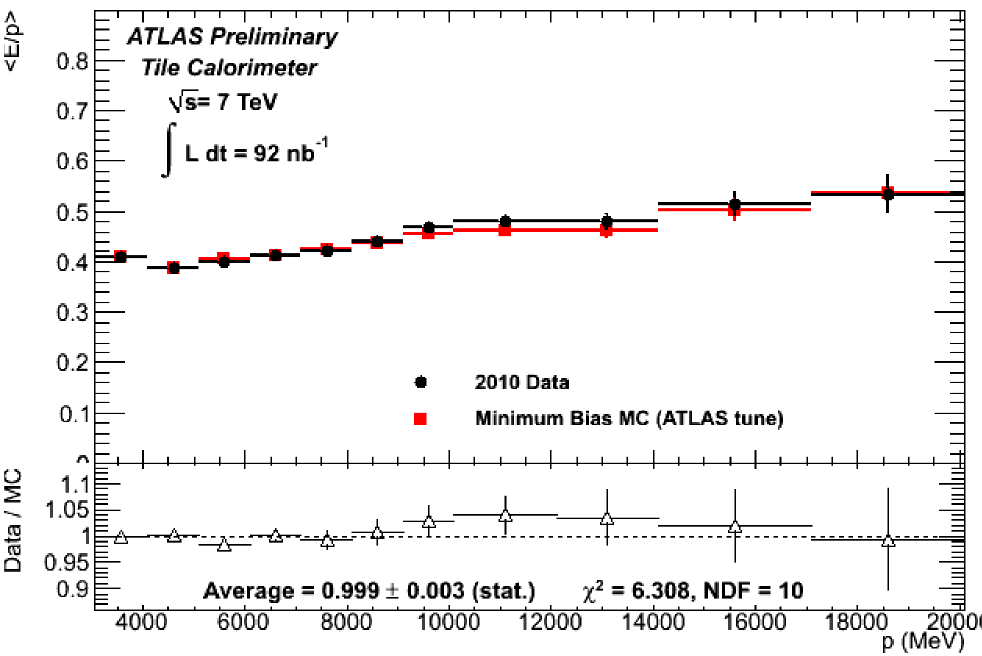
However:
Low p
dominate jet
kinematics!

Big p bins:
hard to resolve
features, but
excellent
agreement!

Energy Response (III)

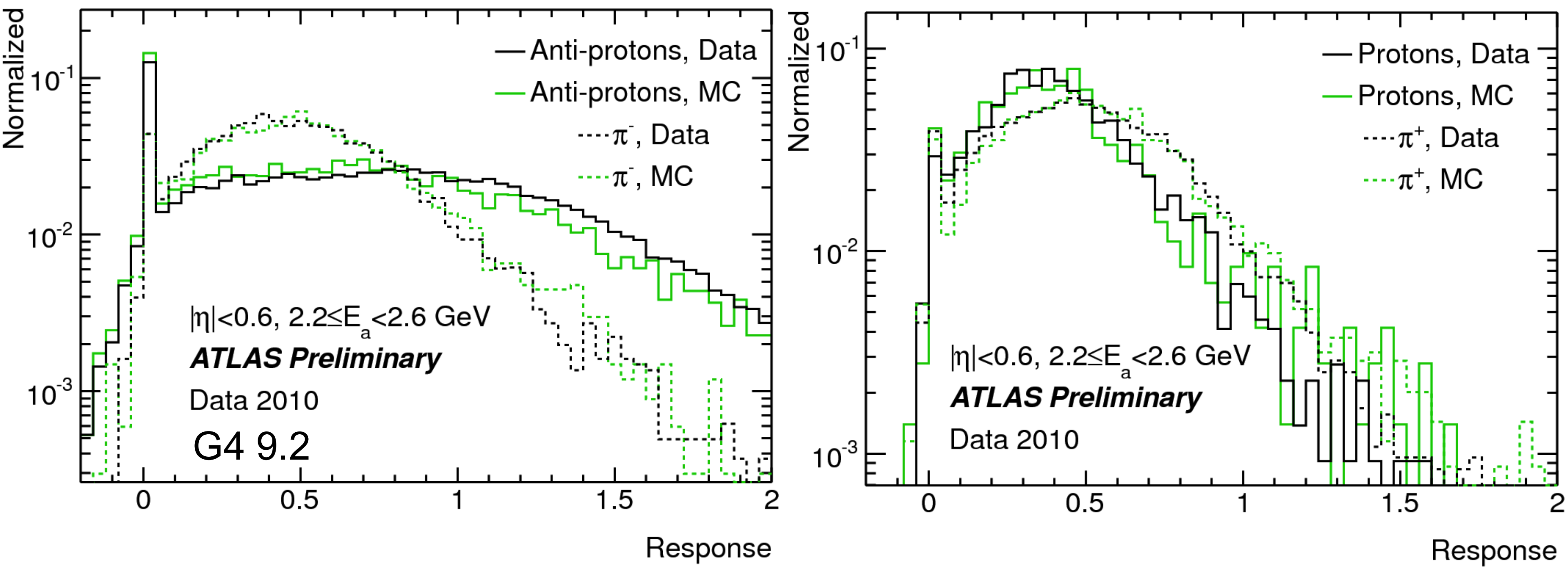
- Also measured in the tile calorimeter in situ using particles that deposit only minimal energy in the EM calorimeter
  - Background is largely caught by the EM calorimeter, so this should really be measuring isolated hadron response in the tile calorimeter
- Excellent agreement with the MC simulation
  - Note: no statement about the fraction of hadrons not interacting in





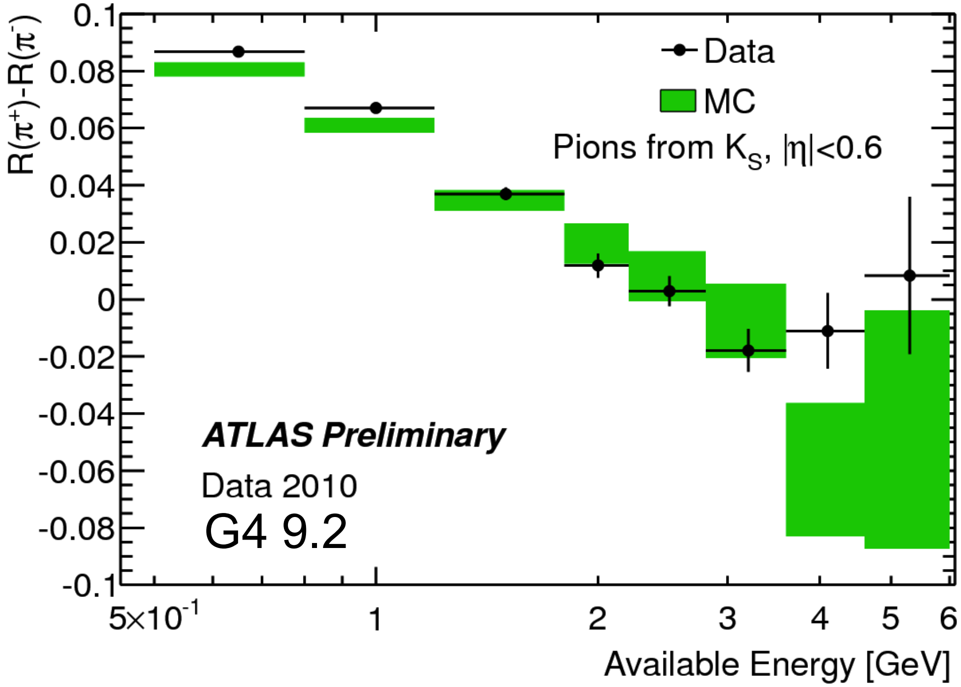
Energy Response (IV)

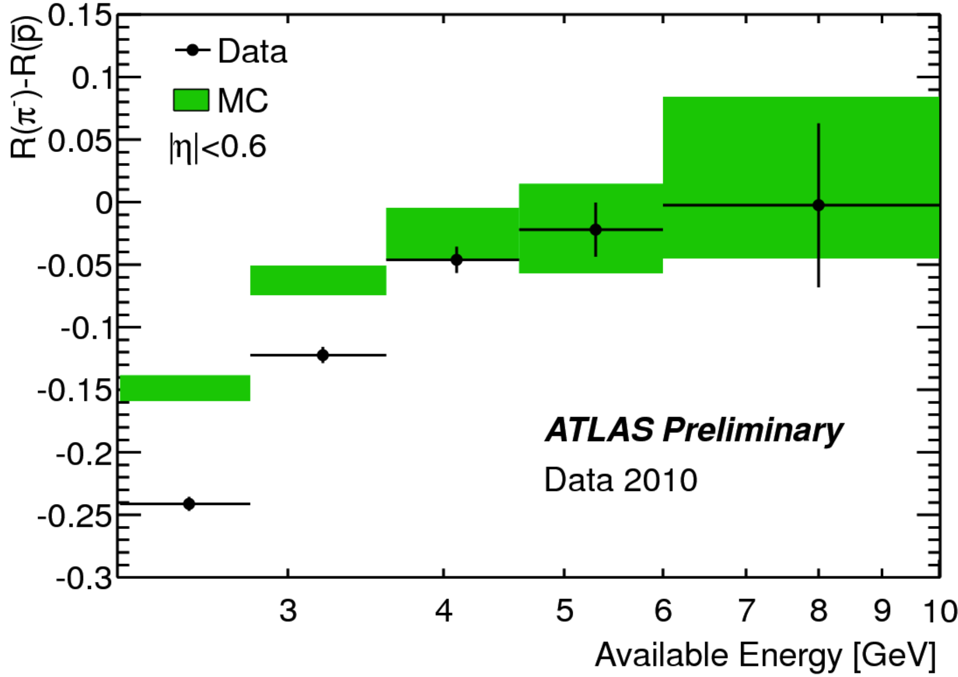
- Now use *resonances reconstructed in situ* for response (E/p)
  - Protons look like pions, but anti-protons look *very* different
  - Binning in E_a, "Available Energy": KE (ps), KE+m (πs), or KE+2m (anti-ps)
- Notice the spike at E=0 from particles that don't deposit much energy in the calorimeter (geometry dependent!)
  - Good agreement gives us confidence in our detector description
- Tail above E=p from neutral bgs and annihilation of the anti-proton



Energy Response (V)

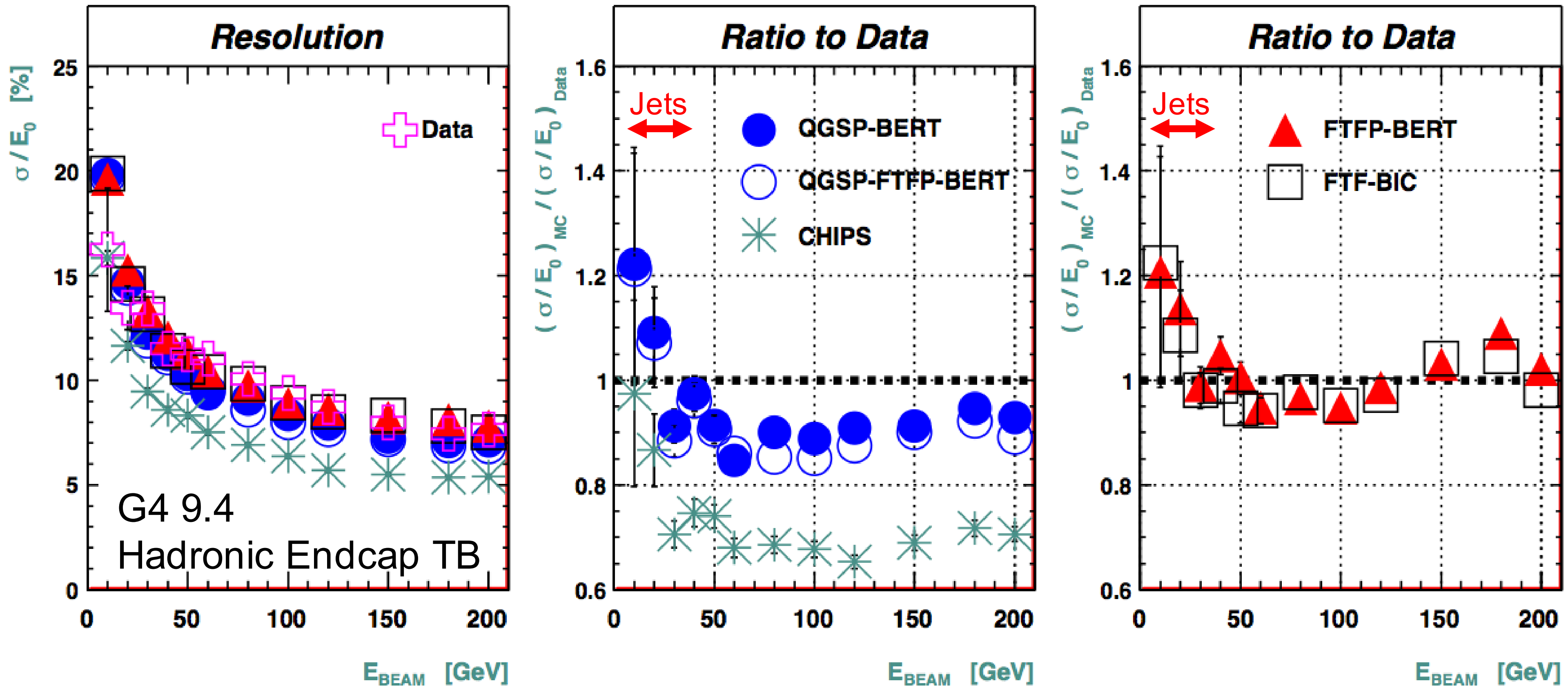
- Use *differences* to avoid background issues
- Now we can see π+ and π- *are* different
- We can also find where G4 does poorly: anti-protons with QGSP_BERT, as expected!
  - This situation is improved somewhat in other physics lists that use CHIPS for anti-protons (e.g. FTFP_BERT as shown by Markus Jügnst)





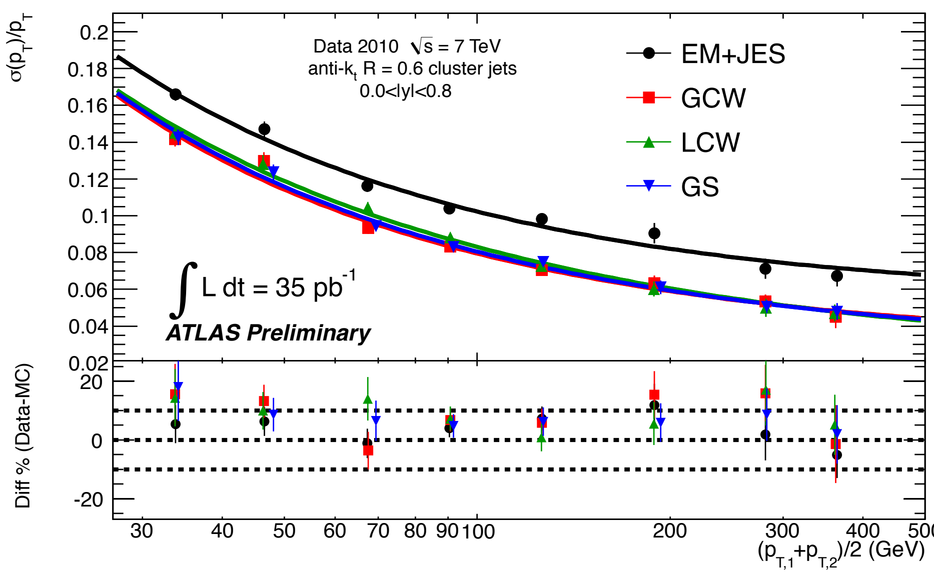
Resolution

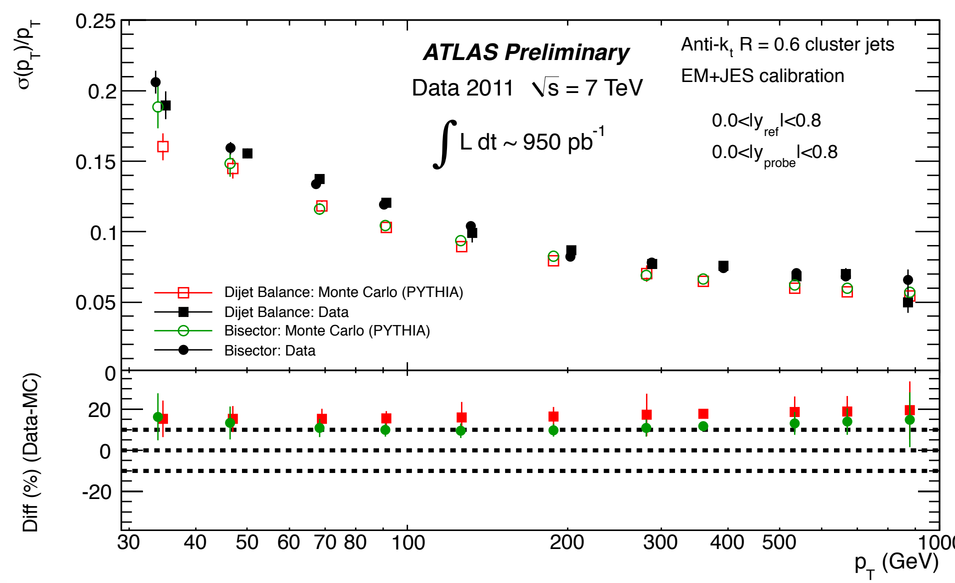
- Considerably more variation in resolution
  - In situ, we have to also deal with background variations, which make resolution measurements *very* tricky - go to test beam!
  - Also varies against G4 release considerably more than response



Resolution (II)

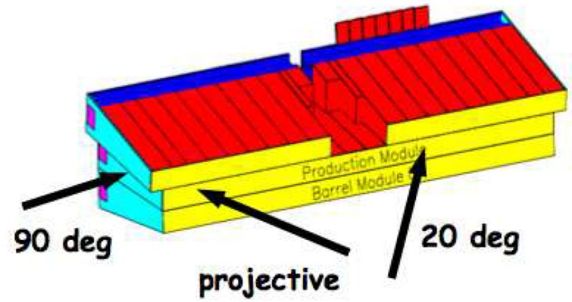
- Jet resolution looked quite good in 2010 (left)
- Some discrepancies appearing in 2011 (right)
- Not clear yet what we can 'blame' that on
  - Lots of basic things have changed, not the least of which is the amount of pile-up in the data! The calorimeter noise-suppression threshold has changed as well (to deal with the higher pile-up)
  - We know single particle resolution isn't perfect, but we're in the 'good' regime
  - Could be a problem was masked in 2010 by some other problem we had…
  - Too early to draw conclusions in this complex system

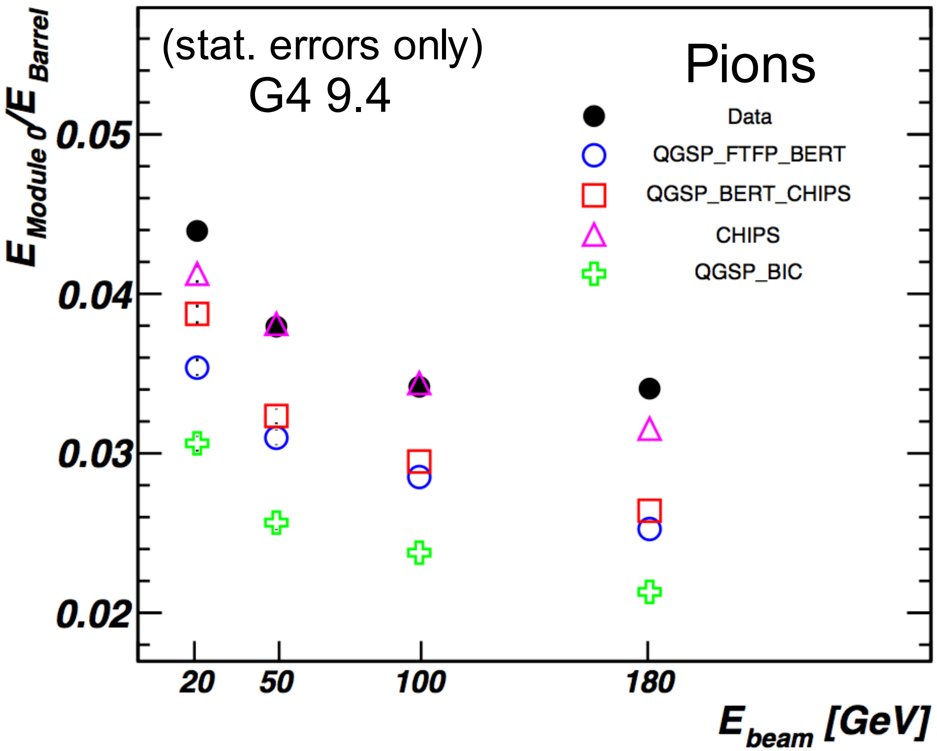


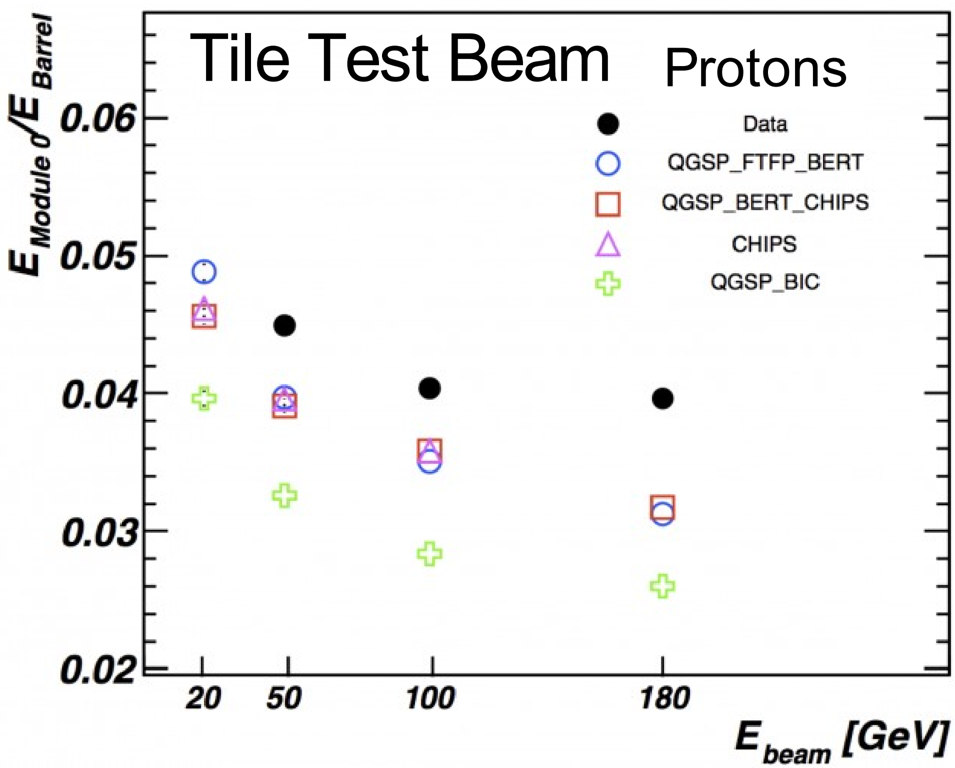


Lateral Shower Shapes

- In most models, showers are laterally too narrow
  - By about 30%, depending on the metric
- Long-standing problem for G4
  - Low-energy brem or low-energy n physics problem?
  - See Olivier Arnaez's talk for some part of the story in the EM calorimeter

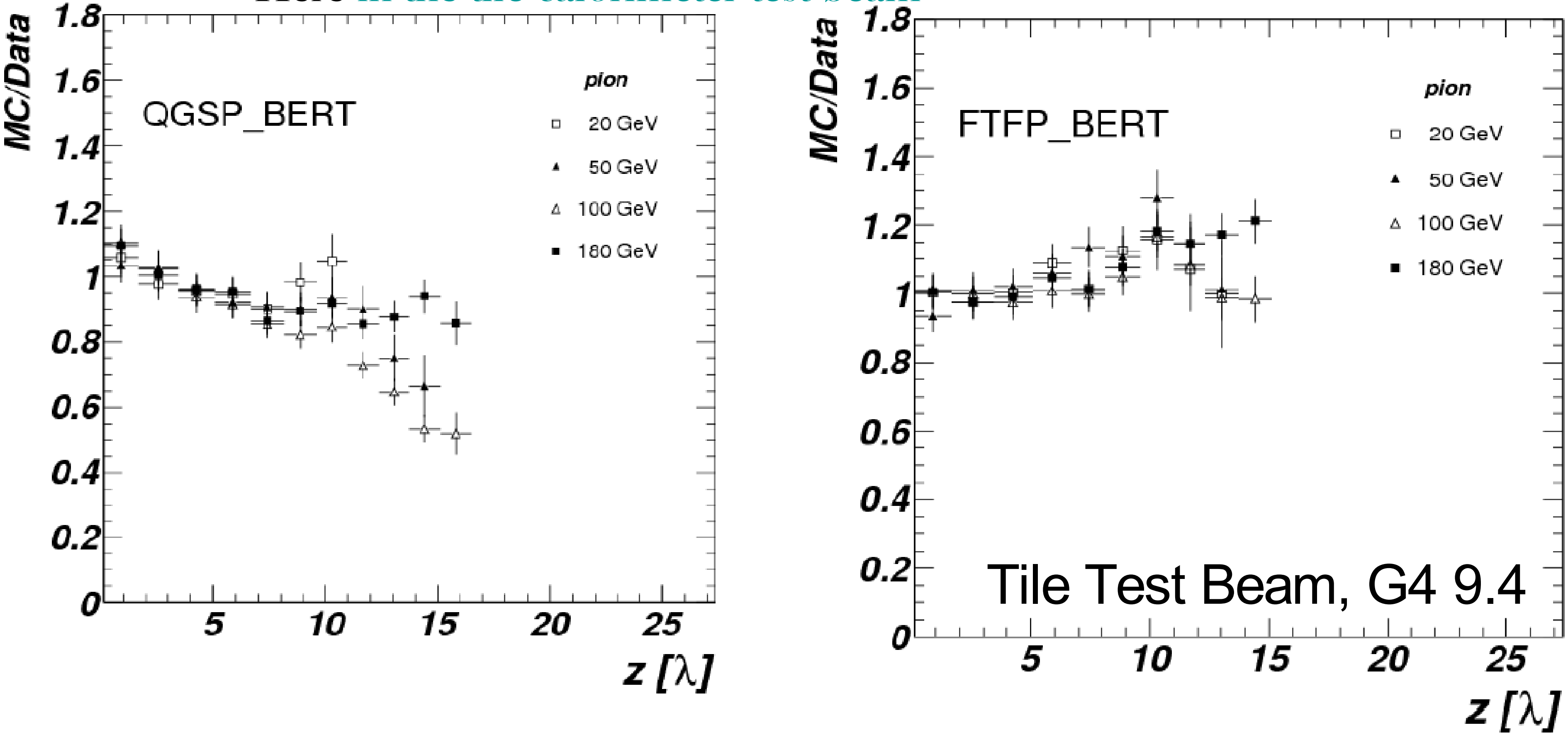






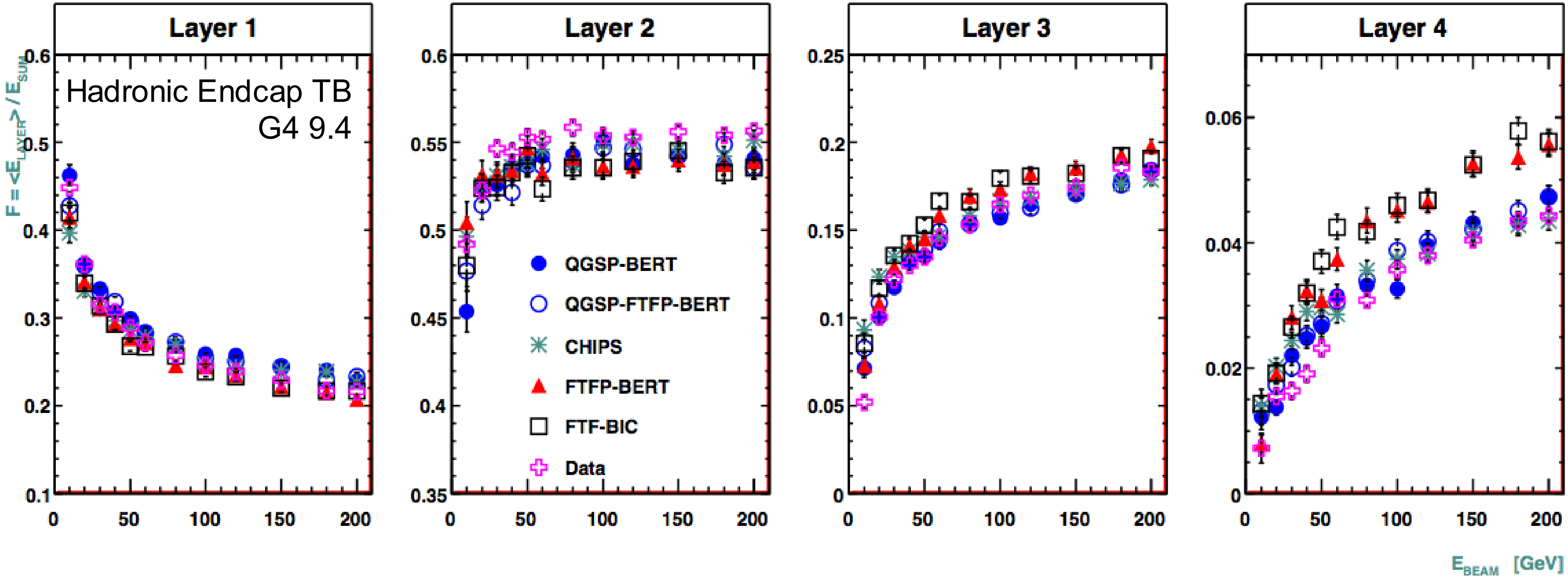
Longitudinal Shower Shapes

- Shower depth was measured in test beam(s)
  - Without background this is an easier measurement to make!
  - QGSP_BERT too short, FTFP_BERT too long
  - Here in the *tile calorimeter test beam*



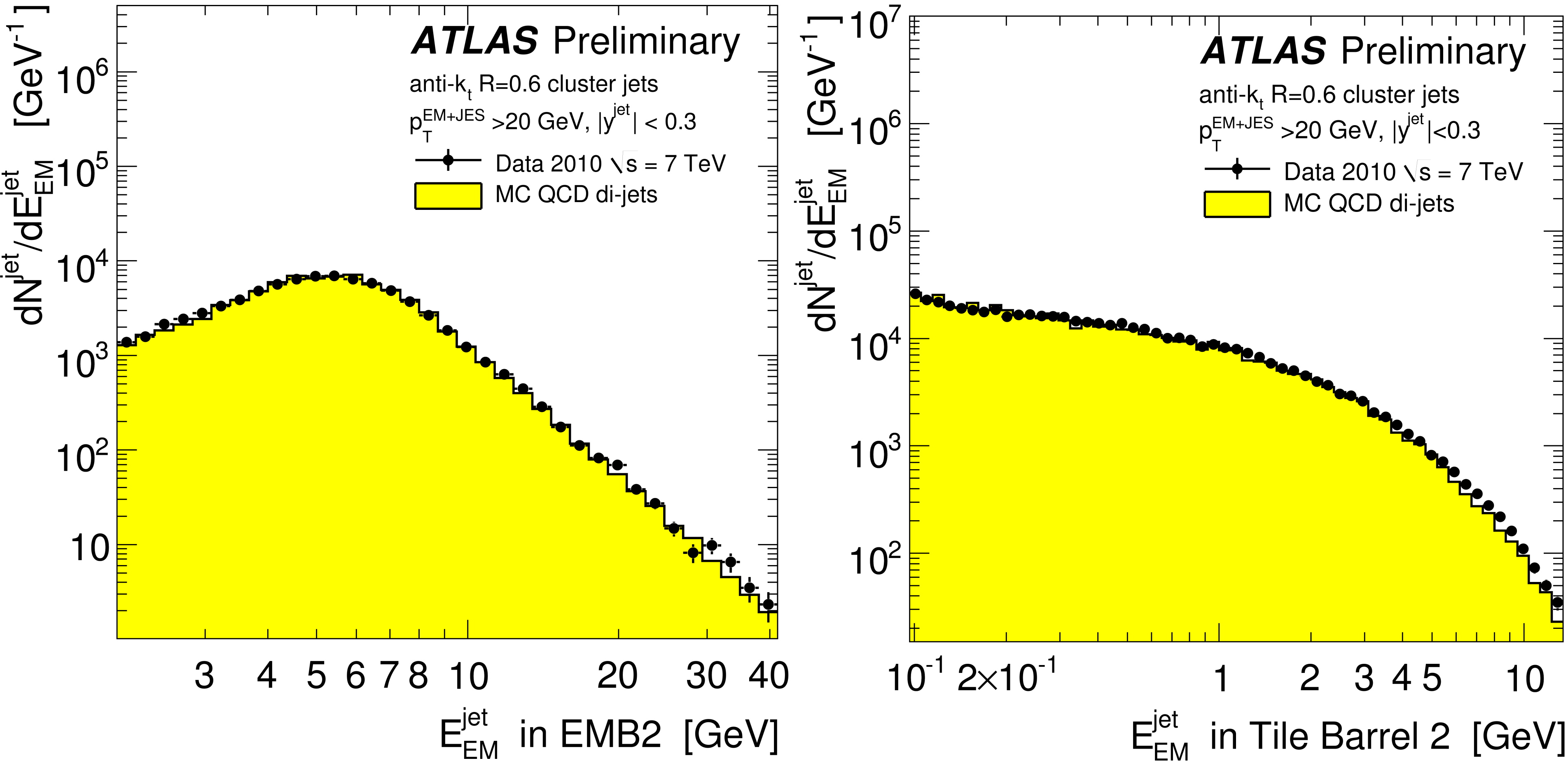
Longitudinal Shower Shapes (II)

- Shower depth was measured in test beam(s)
  - The tail shows significant differences from list to list
- Data agree best with QGSP_BERT in the Hadronic Endcap testbeam
  - Another indication that material matters - the tile test beam (previous slide) was iron-scintillator, this is liquid argon-copper
  - May have implications when picking an optimal physics list for the entire detector – we may not always have the best agreement everywhere



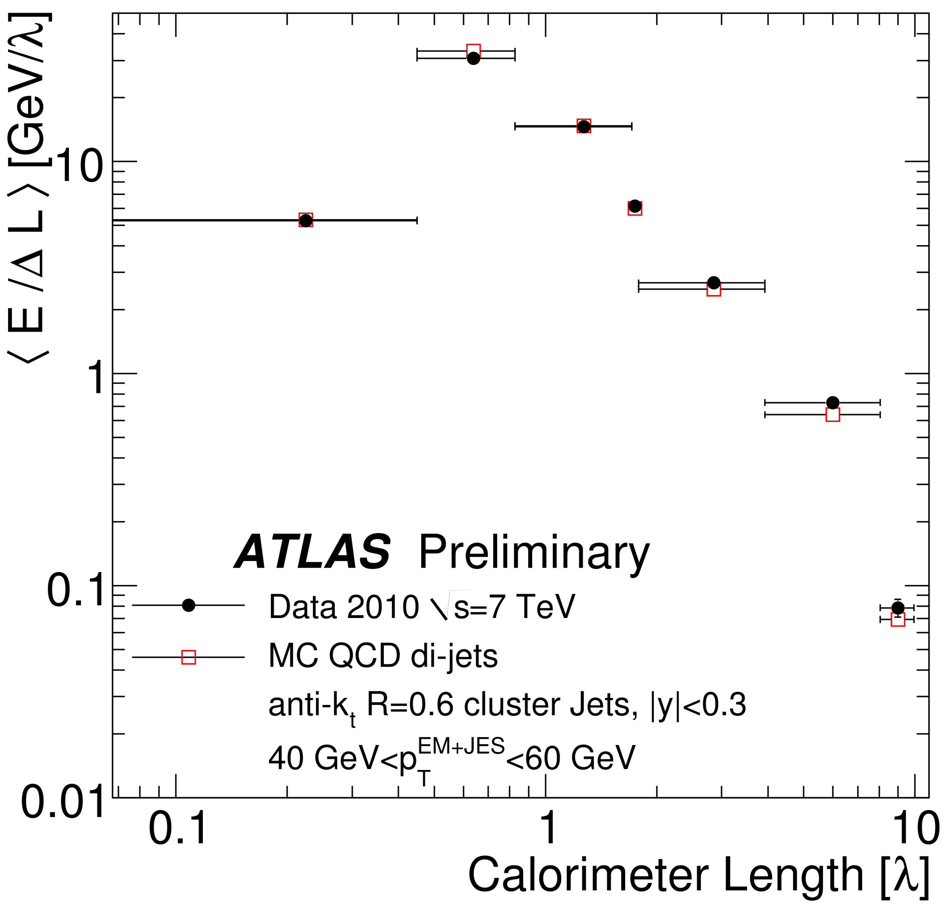
Longitudinal Shower Shapes (III)

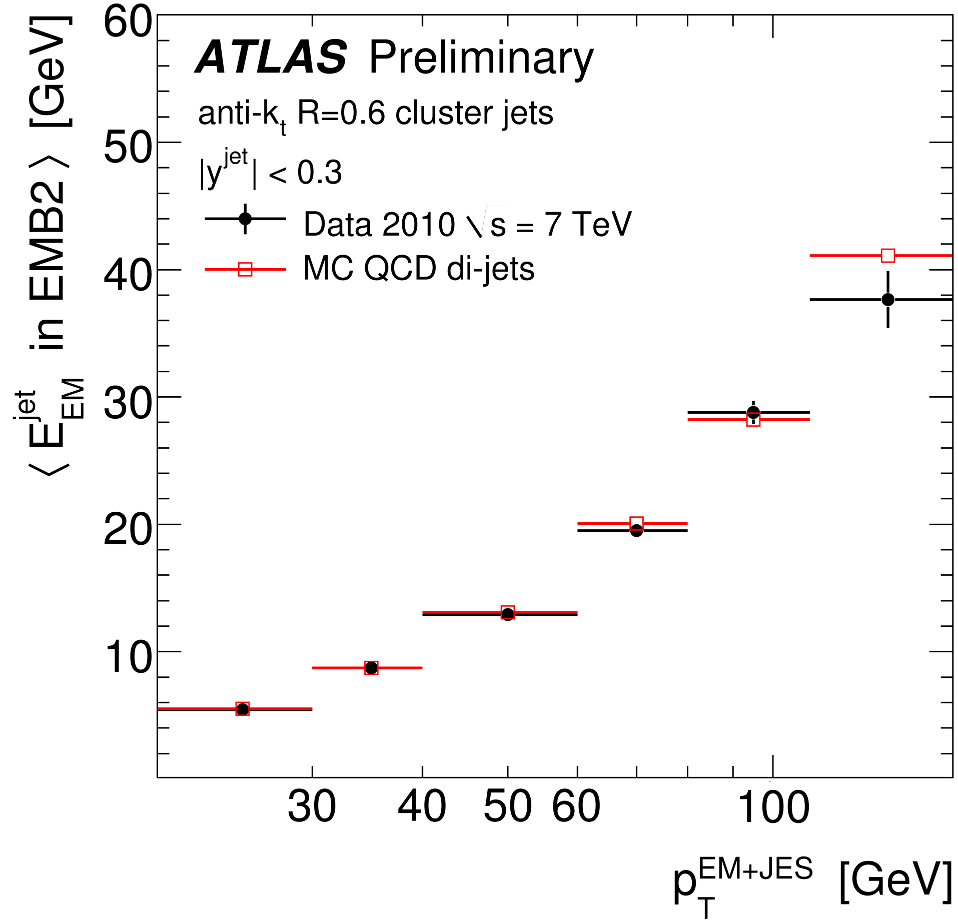
- Jet layer fraction is the in situ longitudinal shower shape
  - Tough to measure – depends on jet kinematics, generator spectra...
  - Still, quite good data/MC agreement, even very early on!



Longitudinal Shower Shapes (IV)

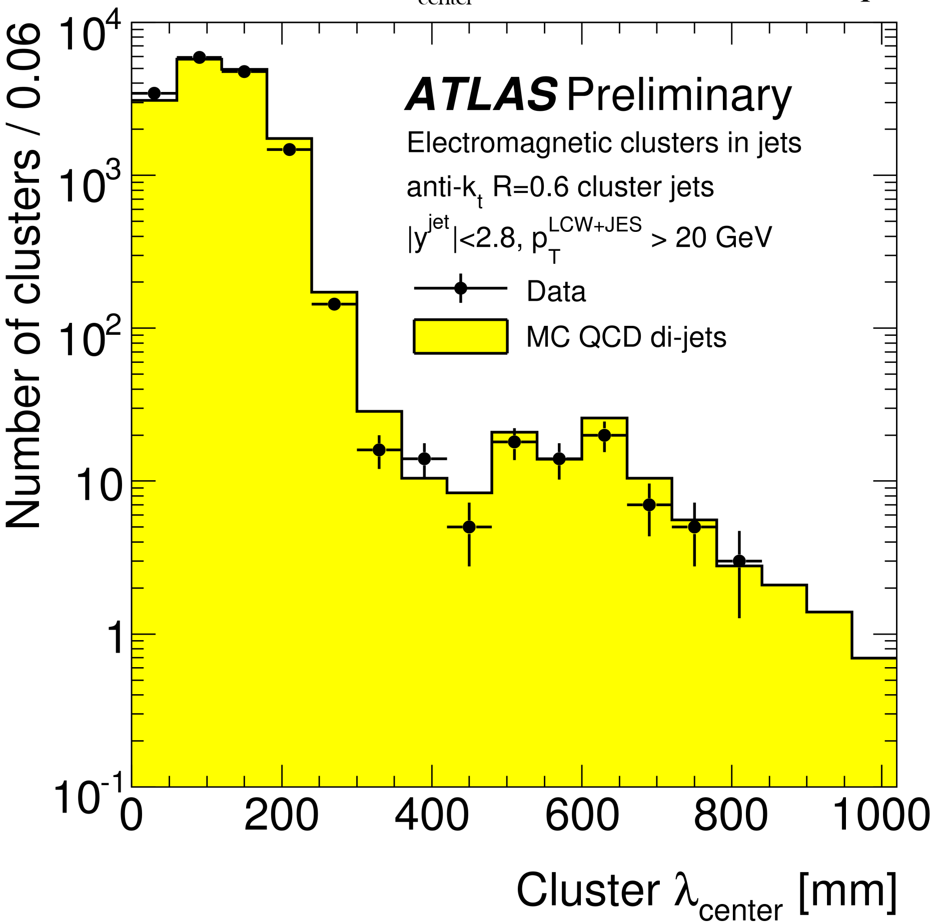
- Quite good agreement; data is *slightly* deeper than MC
  - Very difficult to deconvolve the effects of a poor G4 description and a poor generator description
  - Likely this is from both (Pythia was known to give too narrow jets)

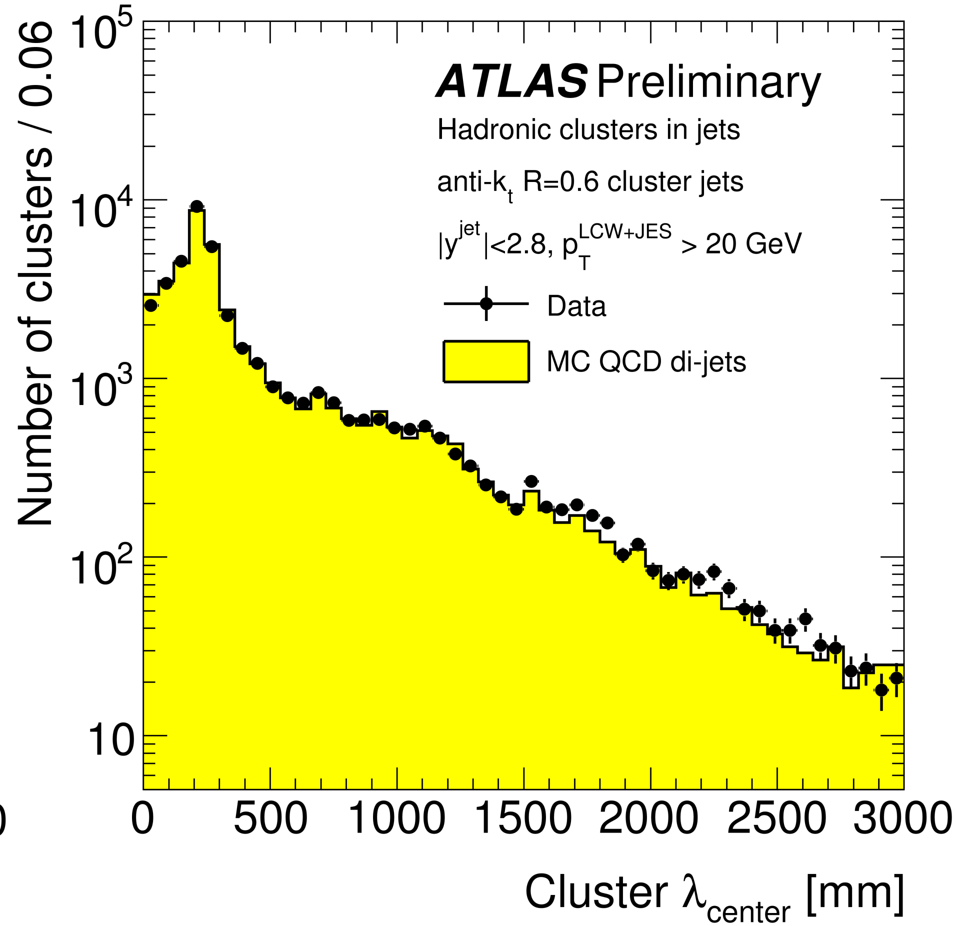




Longitudinal Cluster Distribution

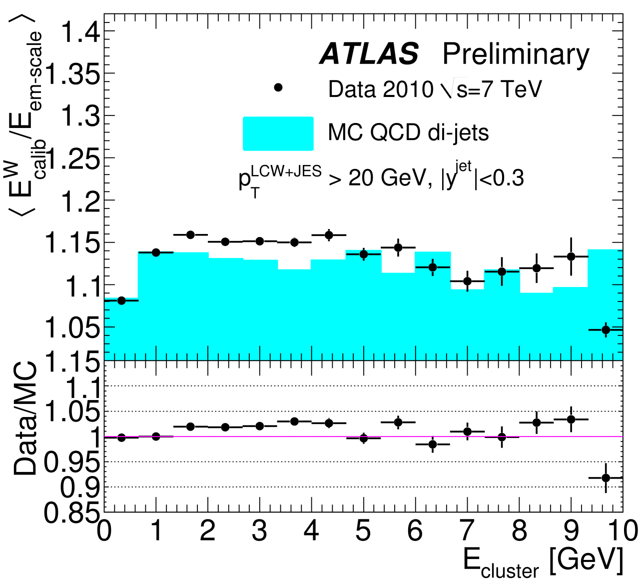
- Clusters classified as EM or Hadronic based on properties
- Distribution of clusters in the calorimeter shows excellent agreement – more a 'jet shape' than a 'pion shape'
  - Here Λ_{center} a measure of cluster depth in the calorimeter

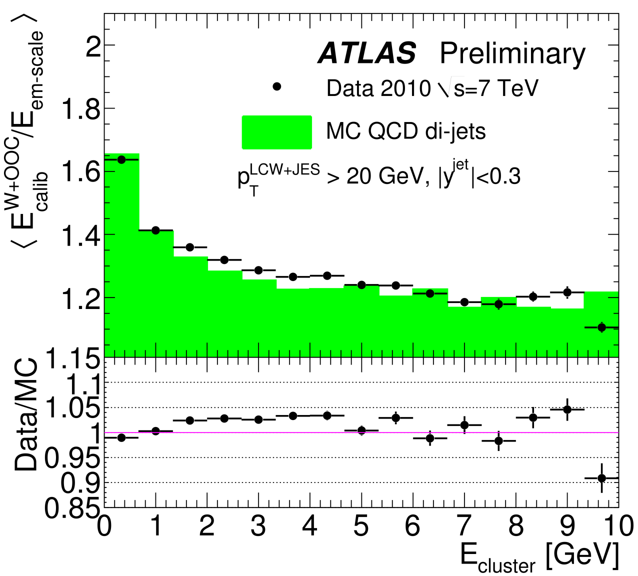


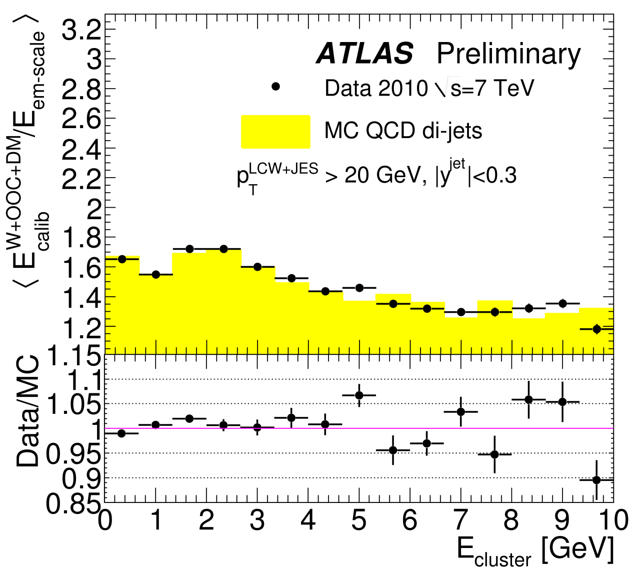


Cluster Properties

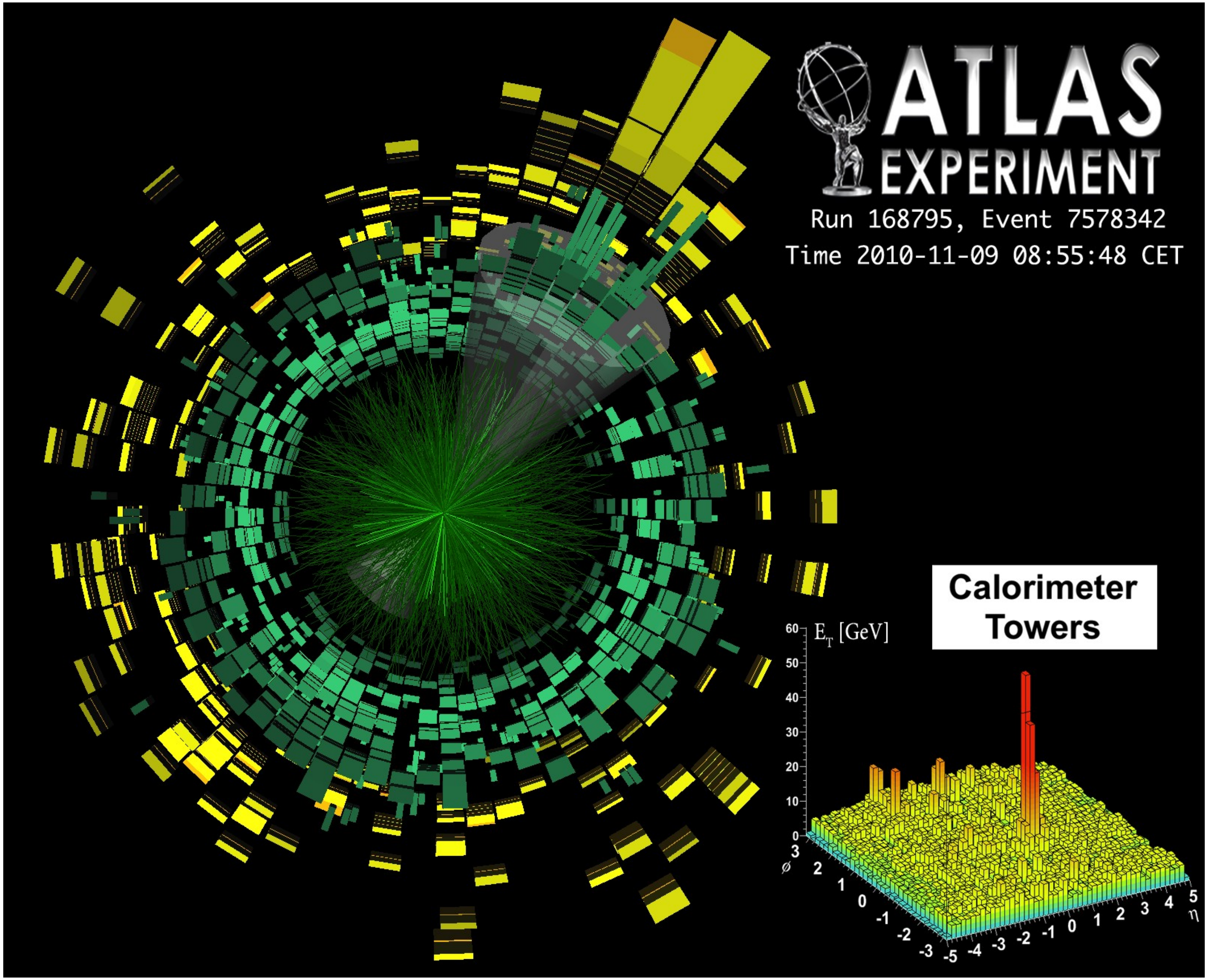
- Clusters are calibrated to take into account three effects
  - Response of hadrons in the calorimeter (i.e. non-compensation)
  - Out of cluster showering (expected loss from clustering effects)
  - Dead material effect (portion of shower not measured)
- All corrections show good agreement between data and MC
  - Good understanding of our jet constituents in this complicated environment
  - Ability to calibrate clusters (jet constituents) removes some dependence of the calibration on the jet algorithm





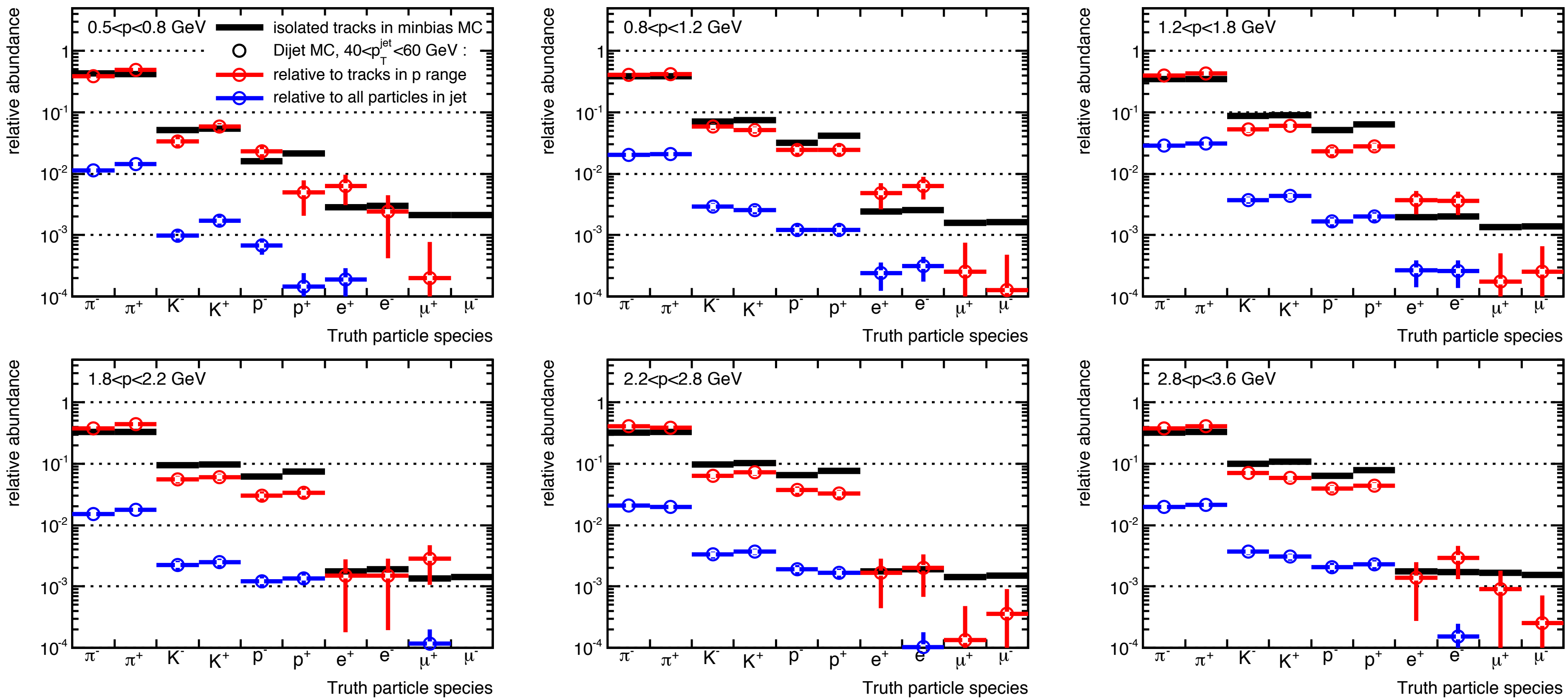


Extrapolation



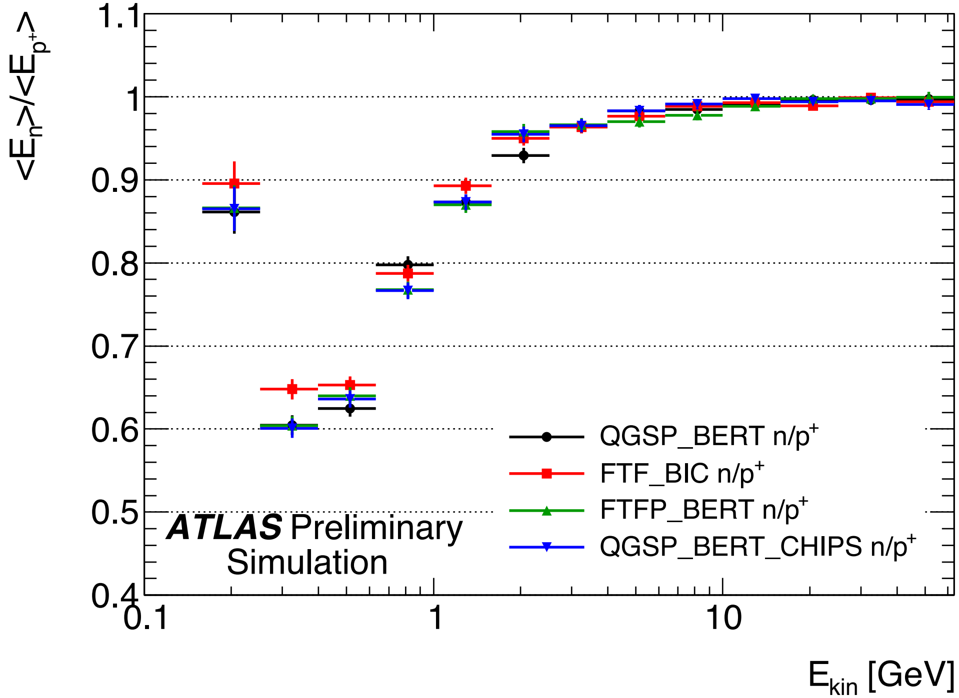
Extrapolating to Jets

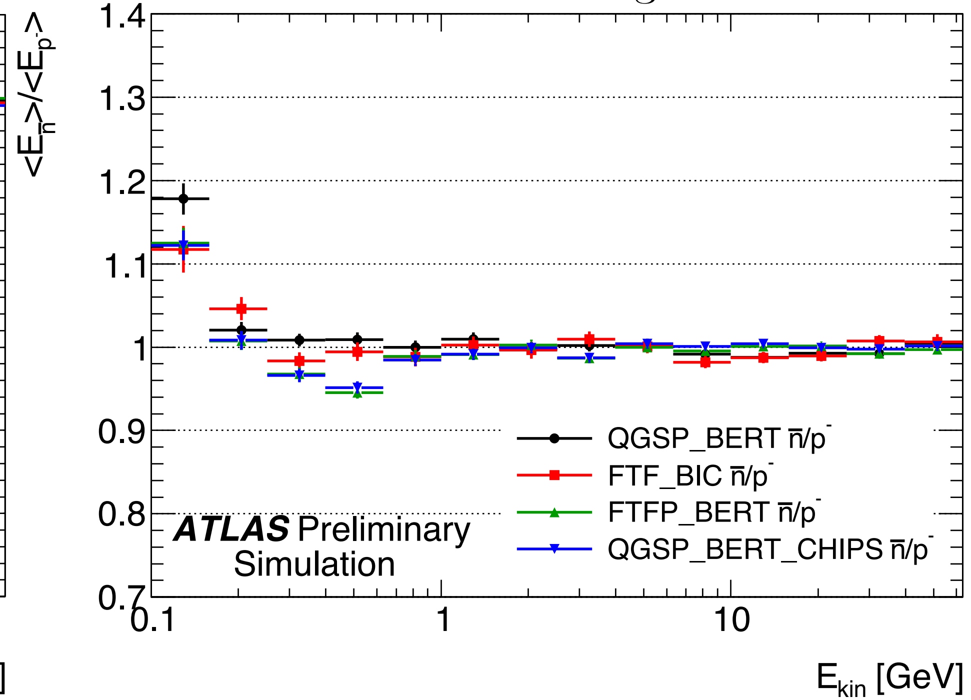
- What changes when we move to jets?
  - If we can measure all the components and how those components go together, we will have great confidence in jet physics
- The particle composition might! Is the spectrum of particles we measure in E/p a fair representation of the particles we find in a jet?
  - Yes! Those in a jet may be high energy, so we extrapolate or use test beam



Extrapolating to Jets (II)

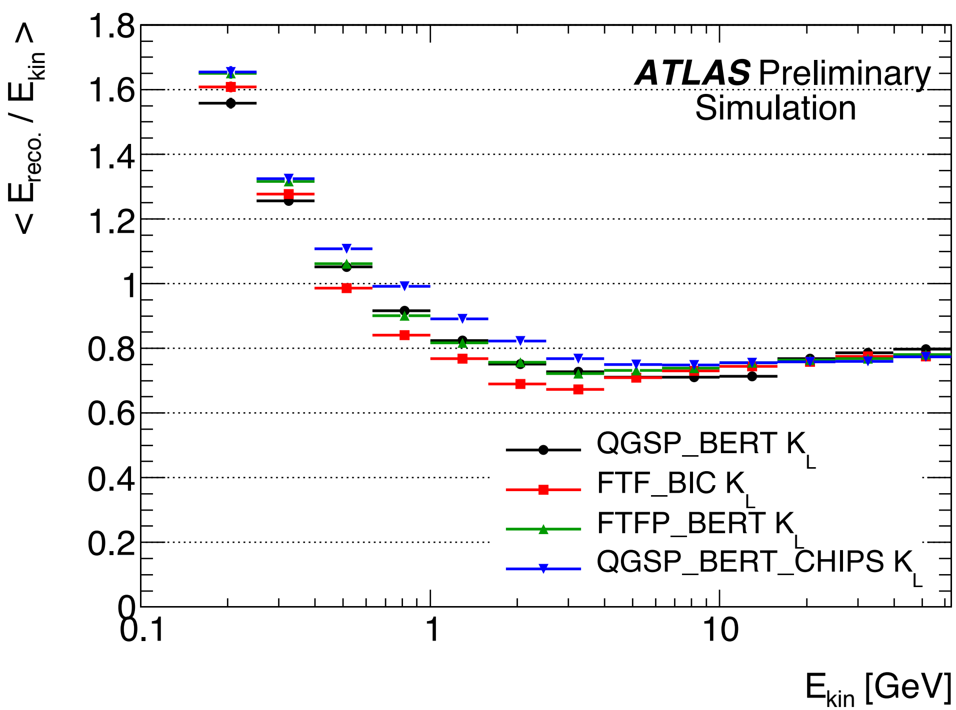
- Response for neutral particles is *very* hard for us to measure
  - But there is some data available to G4 from other experiments
  - So how well do we trust their modeling of the data?
- Compare response ratios of charged particles to neutrals
  - Ratios are fairly constant across various physics lists
  - Perhaps we can say from this that "if we measure protons then we have some information about neutrons" - but we're stretching...

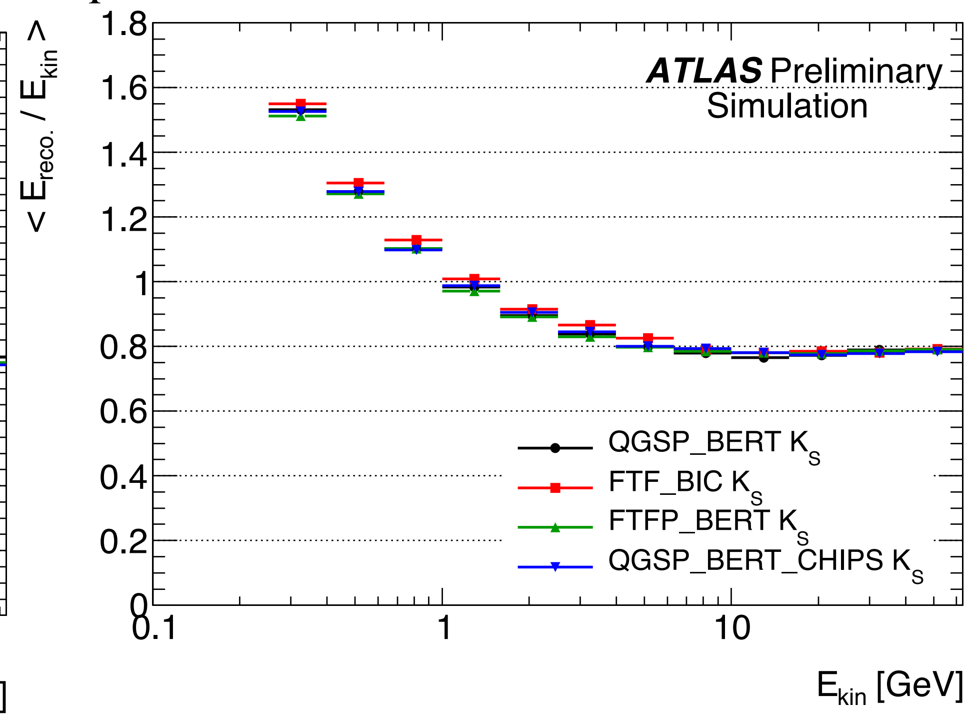




Extrapolating to Jets (III)

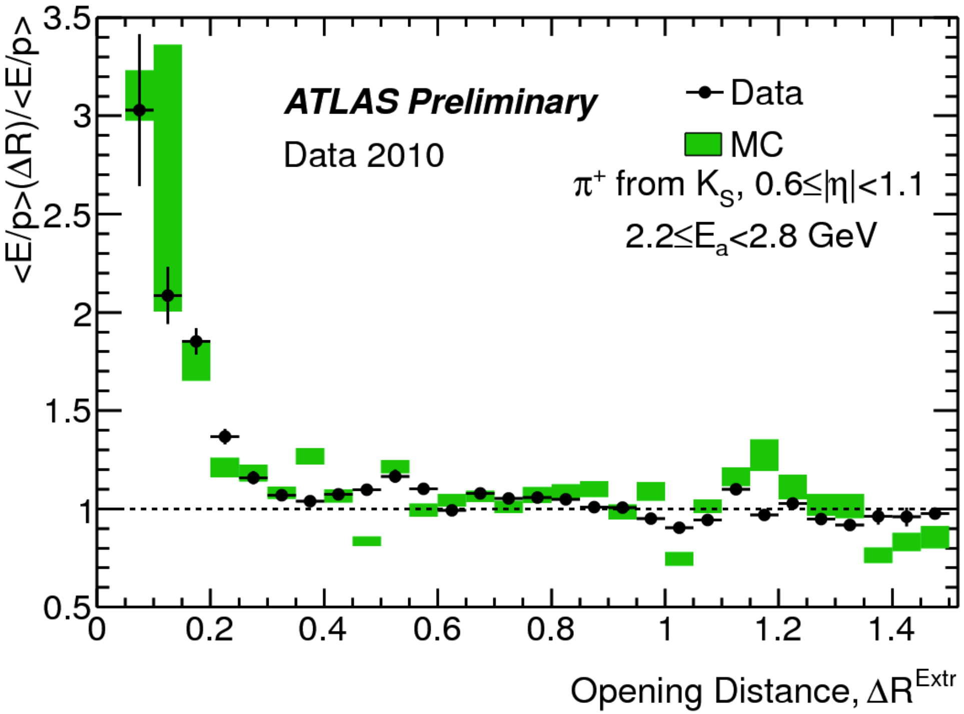
- Similar story for $K_{L/S}$, but much more variation in response
- Some (limited) hope to measure $K^{+/-}$ response using D^*
  - Limited statistics in 2010 and high pile-up in 2011 make this a very difficult measurement to make (may have to rely on the same 'trick')
  - K's are a non-negligible fraction of QCD jets – and analysis selections (e.g. tagging) may enhance the fraction
  - Perhaps ALICE or LHCb can help us out here!

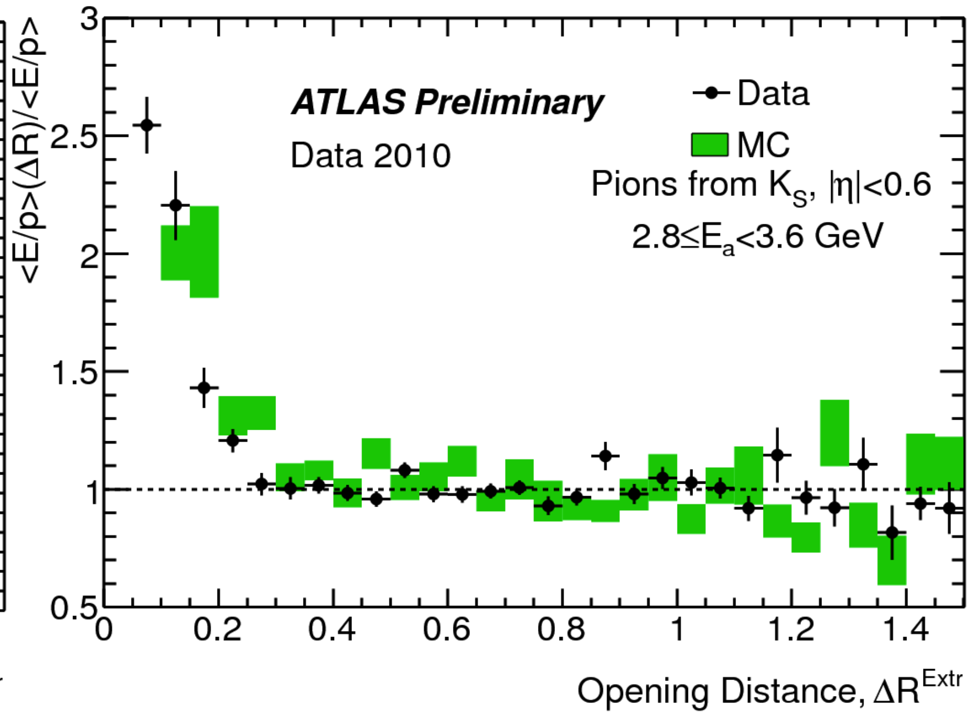




Extrapolating to Jets (IV)

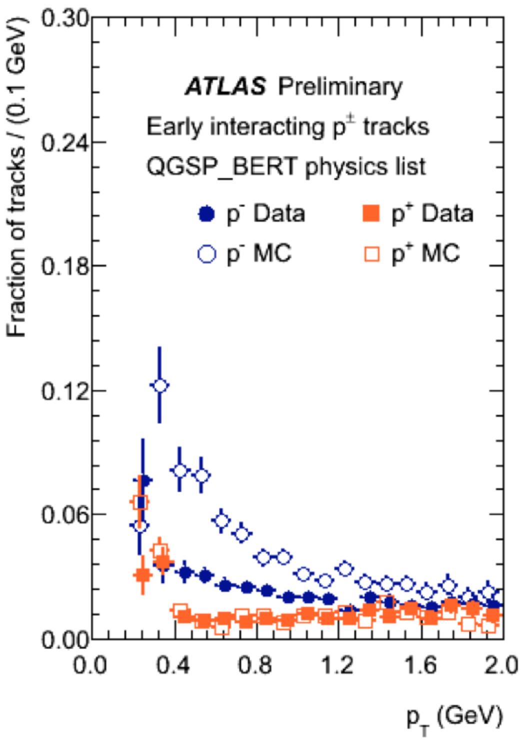
- In E/p we measure *isolated* particles, and jets are *groups* of particles - does the lack of isolation hurt?
  - Good news: No! In ATLAS, at least, the transition from "isolated" to "not-isolated" seems to be well described by G4
  - This includes effects from both lateral shower shapes and calorimeter clustering thresholds (hard to deconvolve!!)

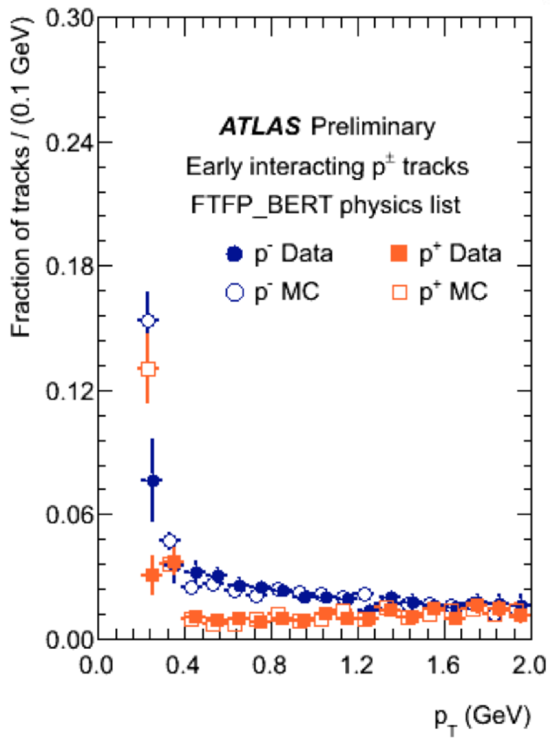


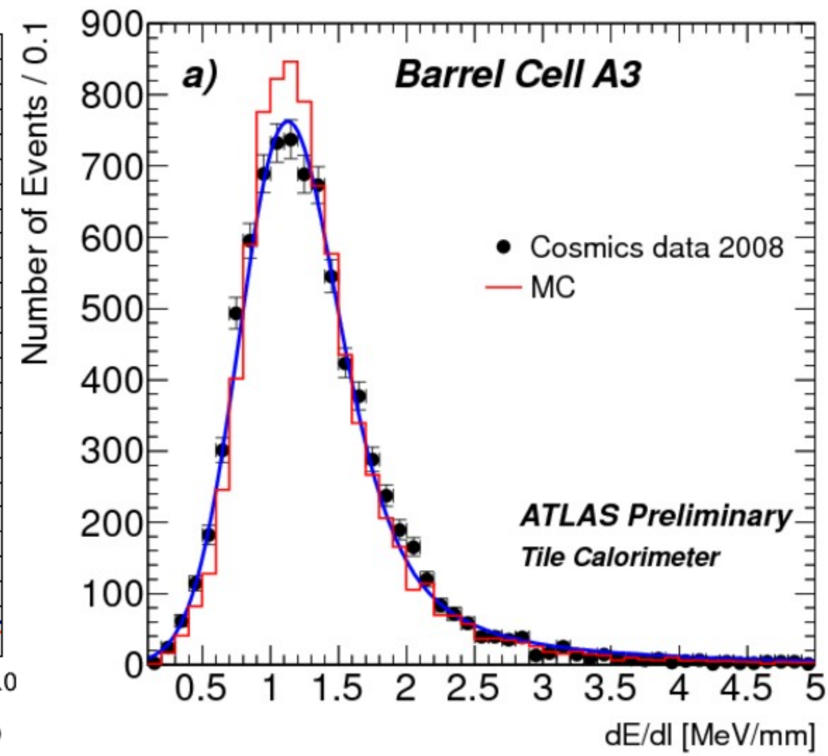


Deconvolving the Problems

- When things don't match up, we often *must* go to other measurements to test specific simulation issues
- That may mean testing hadronic interactions in the inner detector…
  - Test G4 hadronic physics – might measure the # of charged secondaries!
- Or measuring dE/dx for muons in the calorimeter
  - Nice test of detector description, since the physics is well-understood

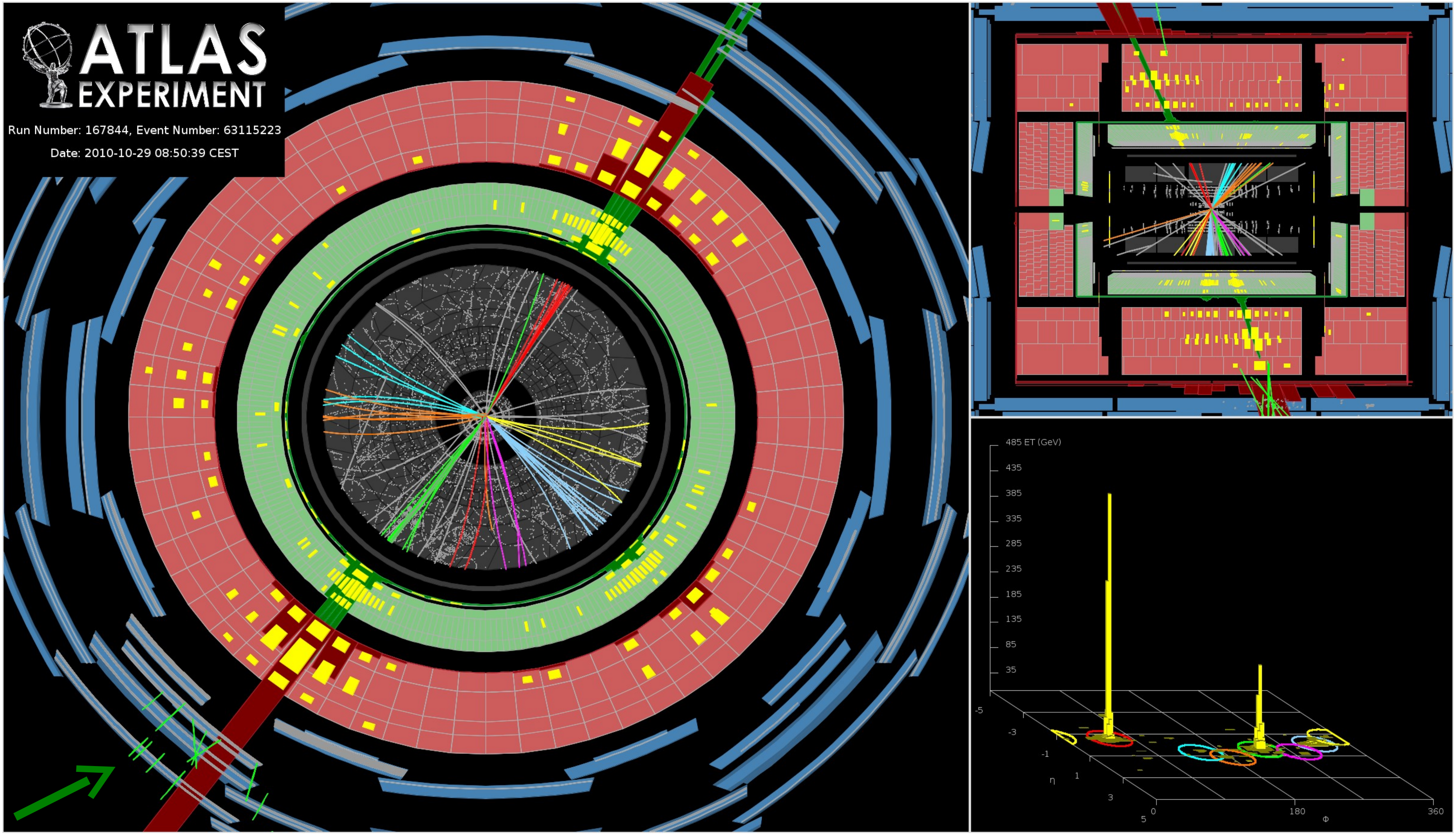






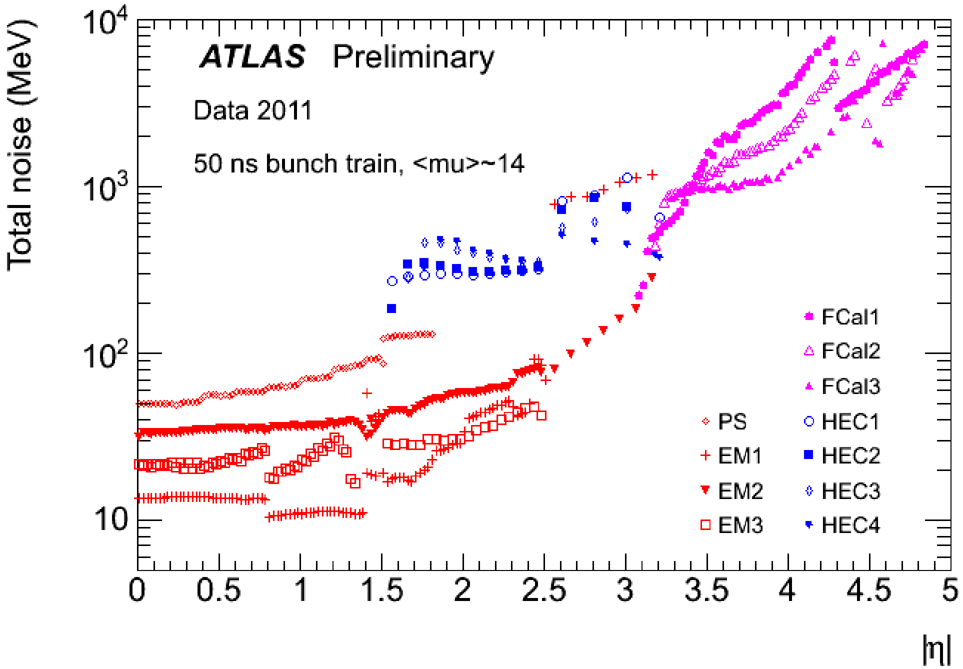
Punch-Through

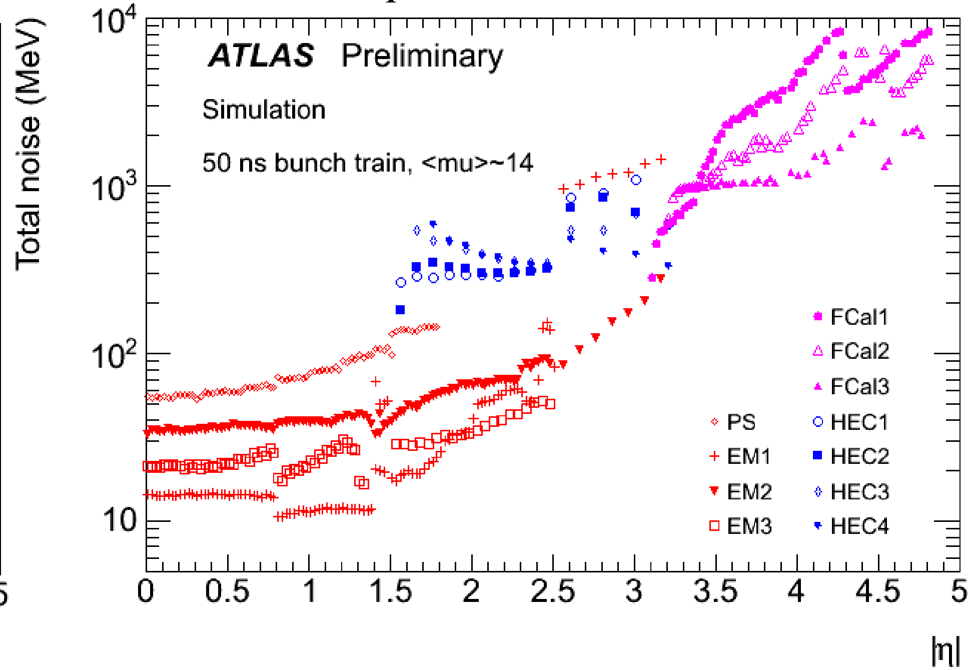
- Geant4 seems to do an excellent job of describing even difficult events, despite reasonable concerns that it might not



Detector Noise

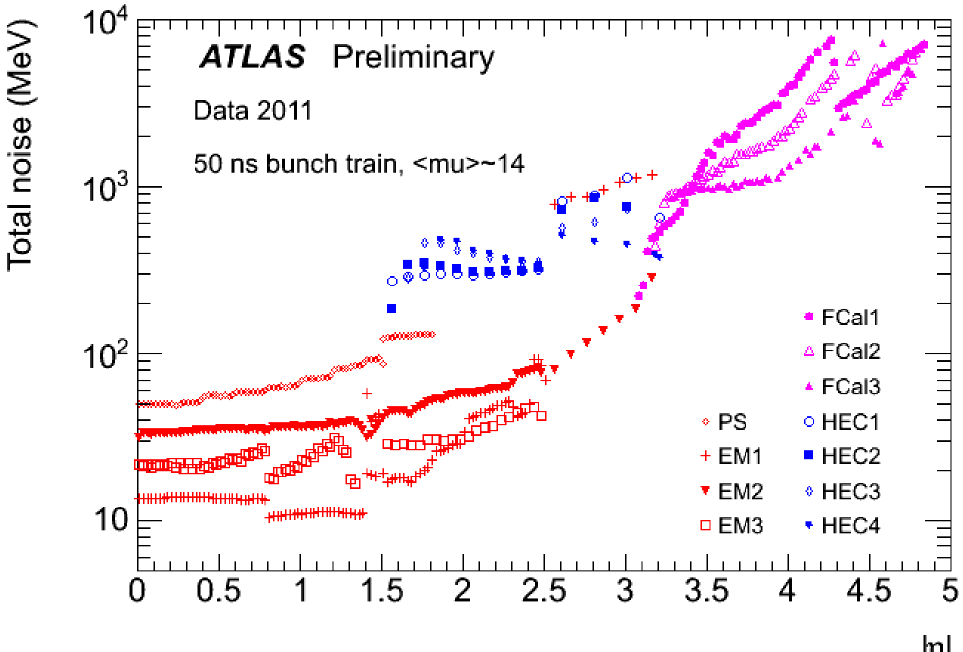
- The calorimeter noise description is quite good, even with moderate to high pile-up (below an average of 14 MB interactions per crossing)
  - The simulation follows the structure of the data quite well
- Some departure in the forward region, particularly deep in the FCAL
  - Origin of the difference is unclear – this could "just" be physics!
  - Considerable improvement in the description of the forward region in the last year, though the detector description may still be an area of some concern – no tracking here to provide constraints / help

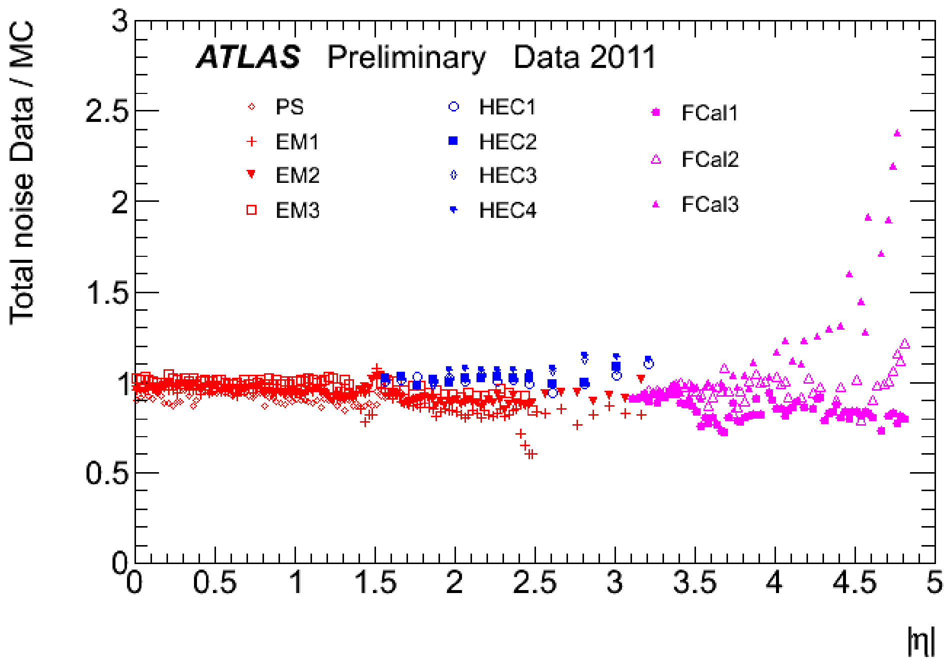




Detector Noise

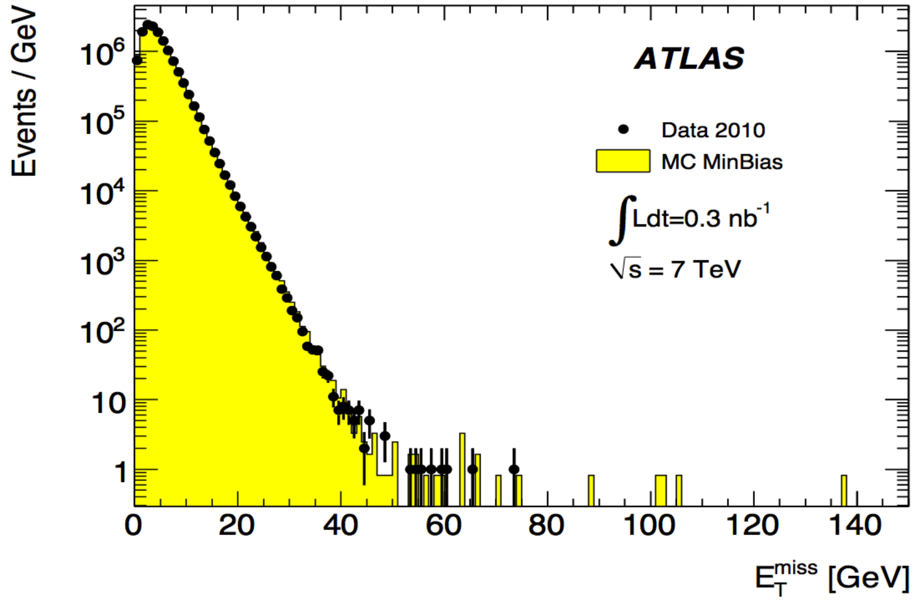
- The calorimeter noise description is quite good, even with moderate to high pile-up (below an average of 14 MB interactions per crossing)
  - The simulation follows the structure of the data quite well
- Some departure in the forward region, particularly deep in the FCAL
  - Origin of the difference is unclear – this could "just" be physics!
  - Considerable advancement in the description of the forward region in the last year (particularly in ID services), though the detector description may still be an area of some concern

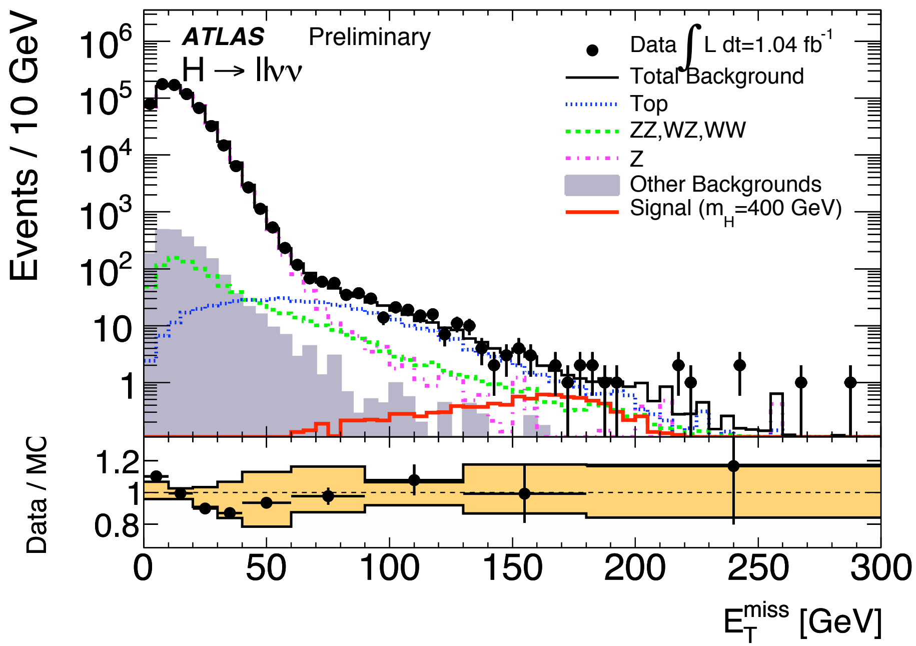


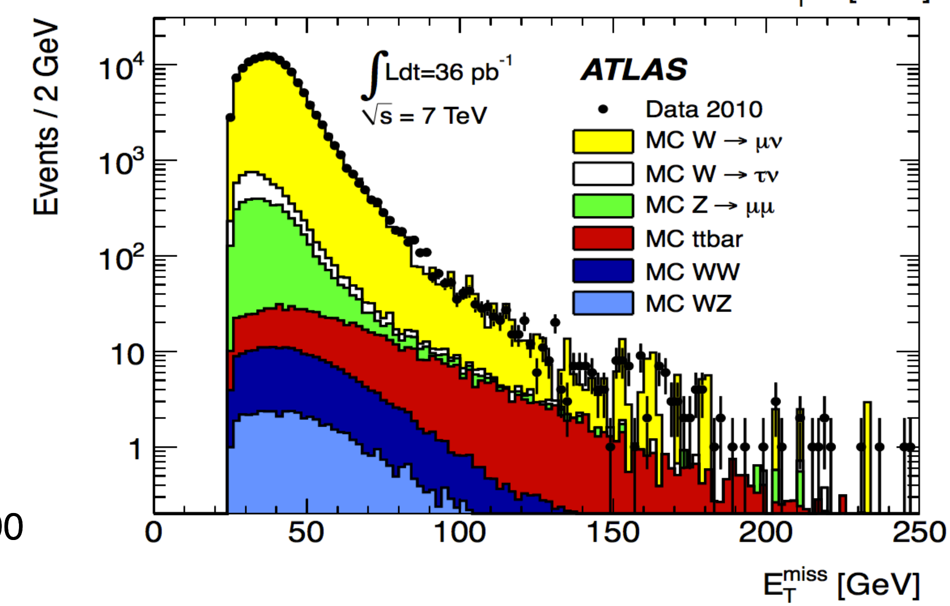


Missing Energy

- Missing energy and ΣE_T are, of course, *very* complicated variables, with many effects entering
- Still, excellent agreement with data, in both the core and tails, in events with and without true MET
- Builds more confidence in our modeling of all aspects of the data







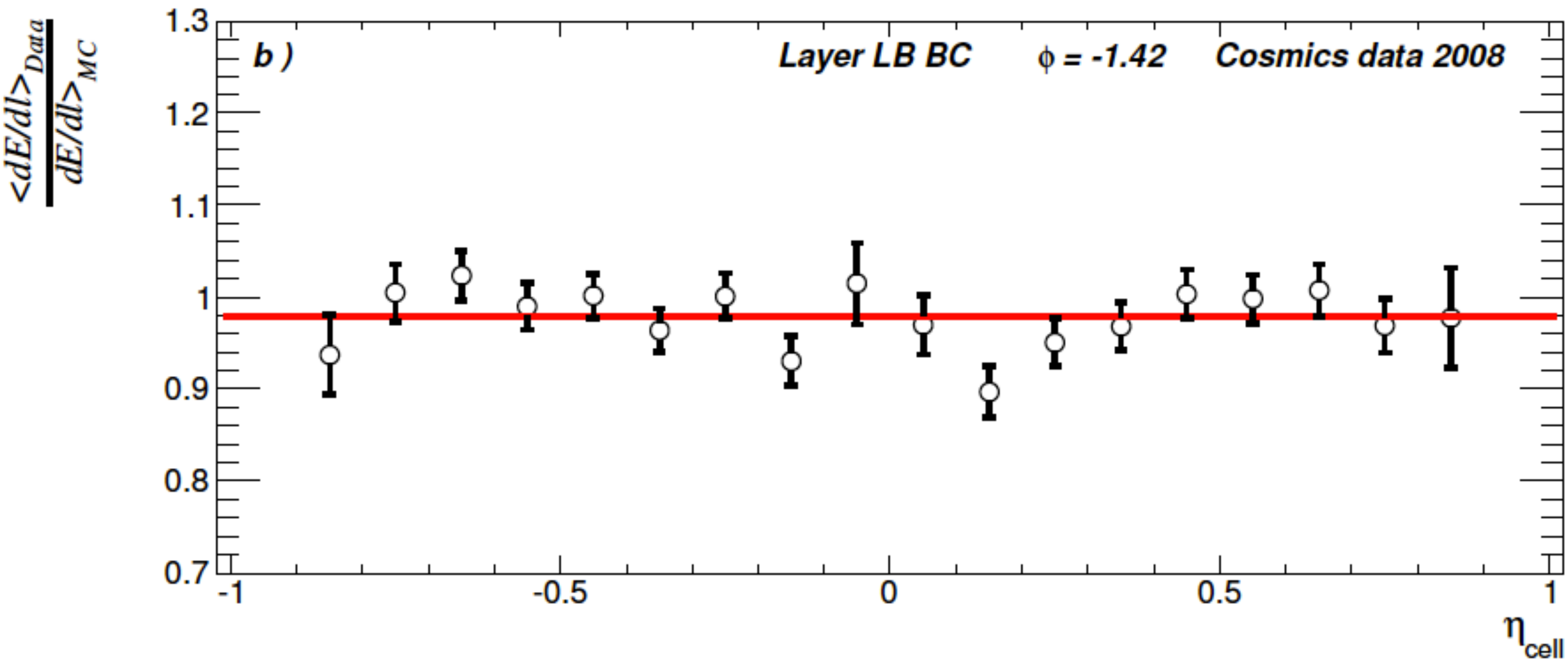
Summary and Conclusions

- Generally, hadronic physics simulation is in good shape
  - In many plots it does well; still some issues (lateral shower shapes!!)
  - Some give a confusing picture: average jet response looks okay, but QGSP_BERT shows an unphysical dip… FTFP_BERT too high?
- Wary of extrapolation from the simple to the complex
  - Getting a handle on single particles is critical
  - Jet kinematics and E_T^{miss} are *very* complicated observables – rarely obvious when things don't look good, what is going wrong
    - More on pile-up tomorrow during John Chapman's talk
- Using all the available handles can give a *lot* of information
  - Need to keep going back to test beam, and need to think about which effects could be in play when data and simulation don't agree
  - We can rely on other aspects (tracking, isolated leptons) for hints about what causes some disagreements
  - Working hard to deconvolve the effects of G4 from our detector description and electronics modeling

BACKUPS

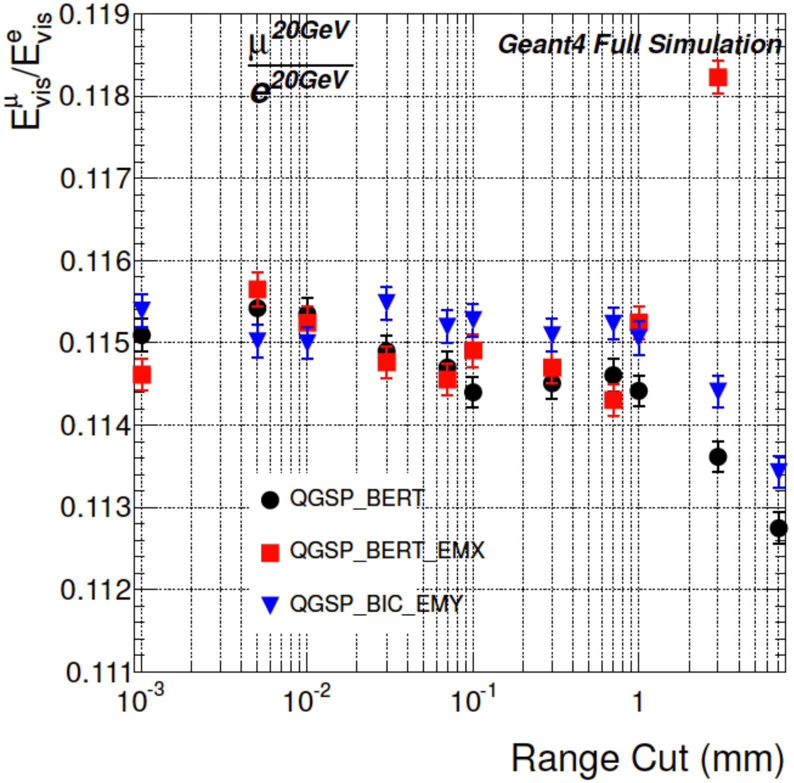
Sampling Fractions

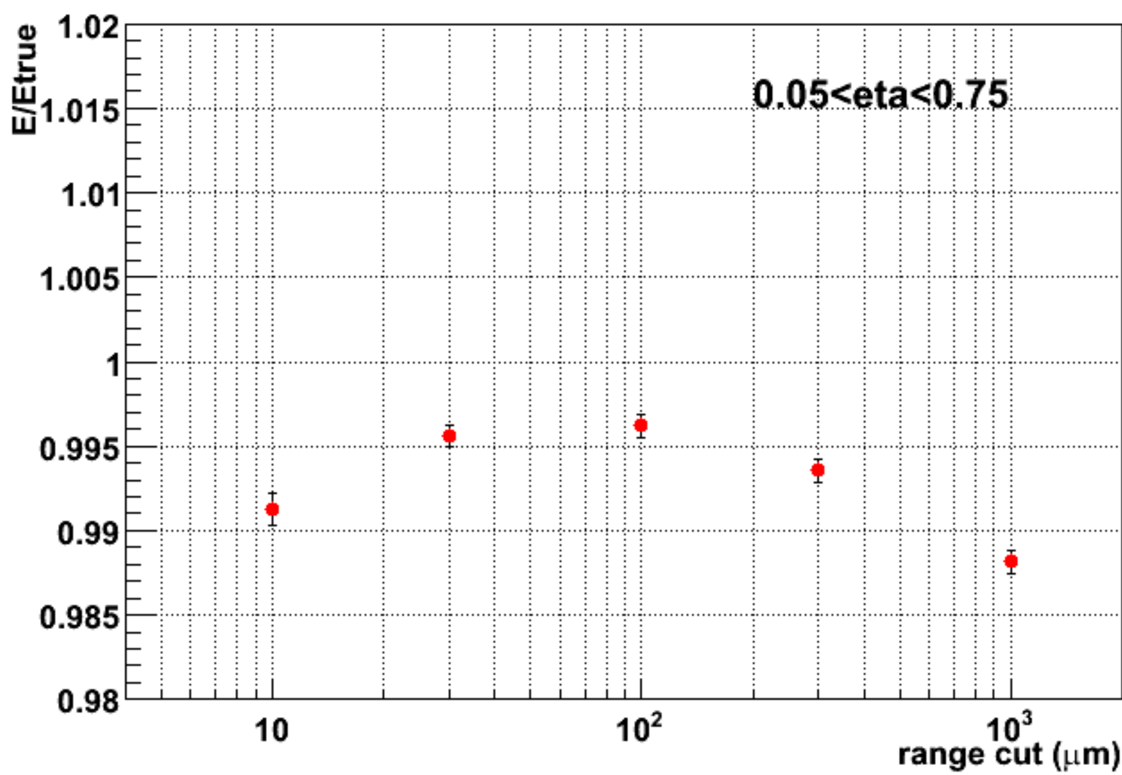
- Measure LAr and Tile sampling fractions several ways
  - Cosmics, Z->ee, any other handle we can find…
  - In LAr, sensitive to temperature - monitor temp carefully!
    - Can't compare well to test beam (no temperature measurement)
- Calibrate constantly in situ with a cesium source



Sampling Fractions

- Varies with range cuts (left: tile; right: LAr)
- Need to be in the stable band (where we claim independence)
  - For us, this means range cuts significantly at least a factor of two smaller than the smallest detector feature (e.g. LAr gap width)
  - Only stable with "new" (range-limiting / non-EMV) multiple scattering
  - Noticeable CPU effects, so watching this is a good idea





Pile-Up

- Pulse shaping in the liquid argon calorimeter is included in the simulation
- Jet offsets from pile-up is modeled to <50%
  - Remaining differences from bunch-to-bunch current variation not modeled in the MC
  - Could be included in the future – see talk from John Chapman in tomorrow's session