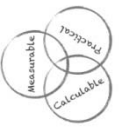


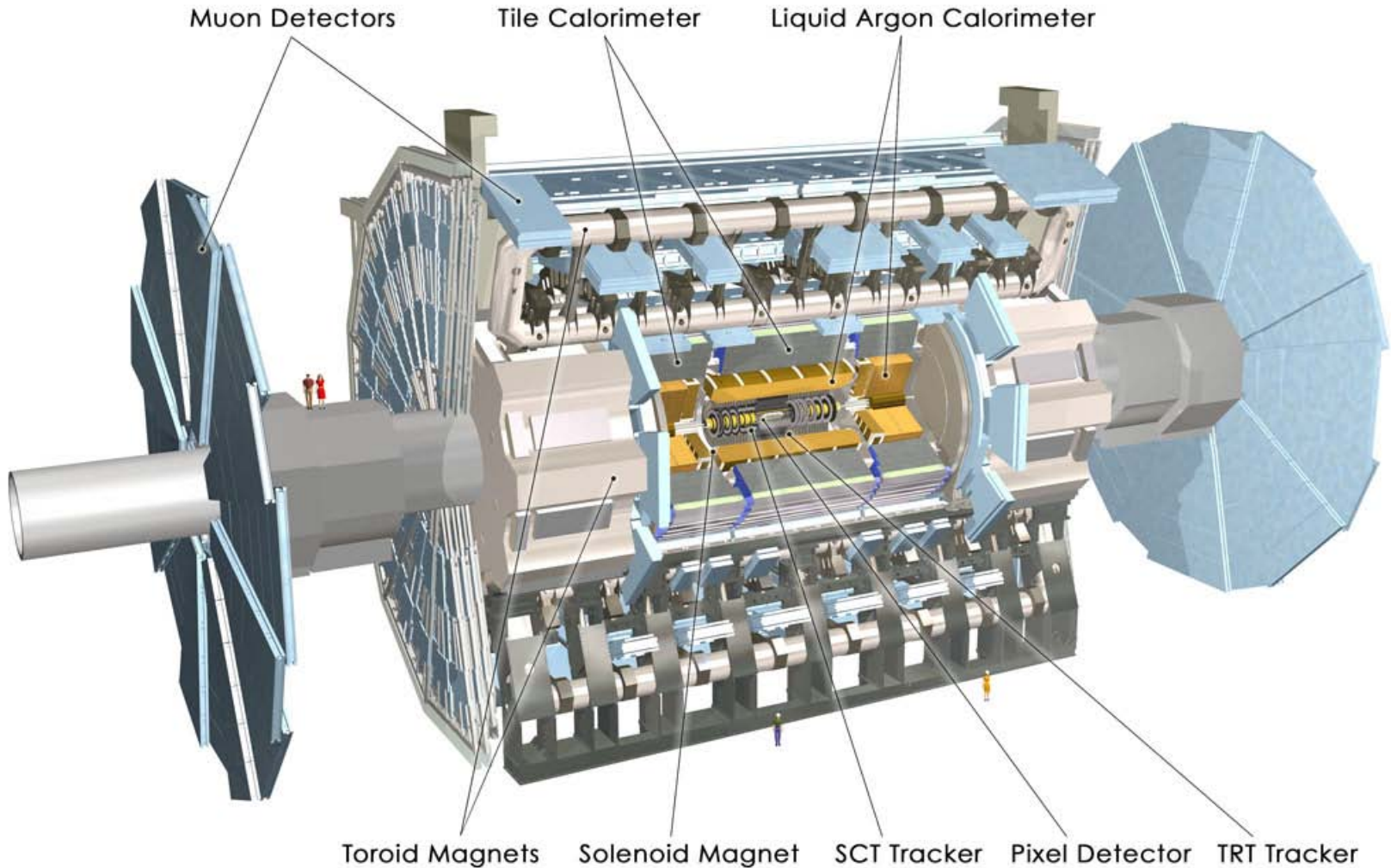
Boston Jet Physics Workshop
Harvard University, January 12-14, 2011



Experimental Aspects of Jet Reconstruction in ATLAS

Peter Loch
University of Arizona
Tucson, Arizona
USA





Total weight : 7000 t
Overall length: 46 m
Overall diameter: 23 m
Magnetic field: 2T solenoid
+ toroid



ATLAS Calorimeters

- EM: $|\eta| < 3.2$,
 - Pb/LAr calorimeter,
 - $22-26 X_0, 1.2 \lambda$,
 - 3 longitudinal sections,
 - $\Delta\eta \times \Delta\Phi = 0.025 \times 0.025 - 0.1 \times 0.1$
 - $\sigma/E \simeq 10\%/\sqrt{E}$.

- Central Hadronic: $|\eta| < 1.7$,
 - Fe/Scintillator sampling calorimeter
 - 7.4λ ,
 - 3 longitudinal sections,
 - $\Delta\eta \times \Delta\Phi = 0.1 \times 0.1 - 0.2 \times 0.1$,
 - $\sigma/E \simeq 50\%/\sqrt{E} \oplus 0.03$.

- EndCap Hadronic: $1.7 < |\eta| < 3.2$,
 - Cu/LAr sampling calorimeter,
 - 4 longitudinal sections,
 - $\Delta\eta \times \Delta\Phi = 0.1 \times 0.1 - 0.2 \times 0.2$

- FCAL: $3 < |\eta| < 4.9$,
 - EM: Cu/LAr, HAD: W/LAr calorimeter,
 - 10λ ,
 - 1 EM + 2 HAD longitudinal sections,
 - $\Delta\eta \times \Delta\Phi = 0.75 \times 0.65 - 5.4 \times 4.7$

~200,000 channels in total
 6-7 longitudinal segments
 for $|\eta| < 3.2$
 3 longitudinal segments for
 $3.2 < |\eta| < 4.9$



Non-compensating calorimeters

Electrons generate larger signal than pions
depositing the same energy

Typically $e/\pi \approx 1.3$

High particle stopping power over whole detector acceptance $|\eta| < 4.9$

$\sim 23\text{-}35 X_0$ electromagnetic calorimetry

$\sim 10 \lambda$ total for hadrons

Hermetic coverage

No significant cracks in azimuth

Non-pointing transition between barrel, endcap and forward

Small performance penalty for hadrons/jets

High granularity

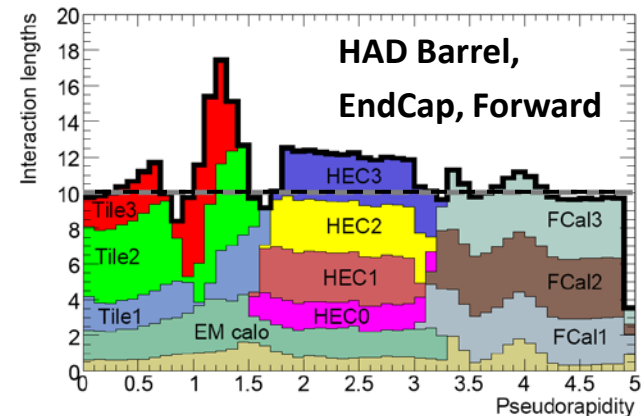
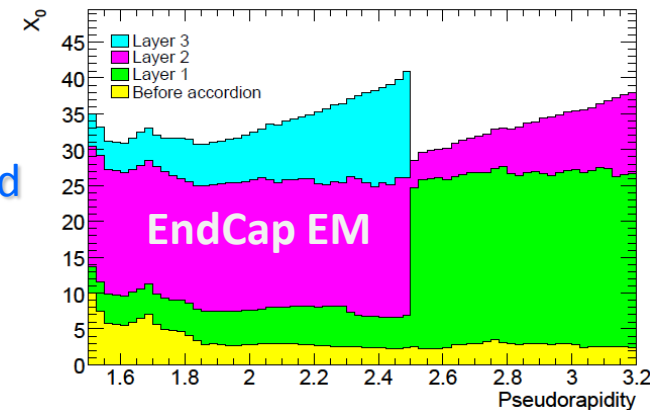
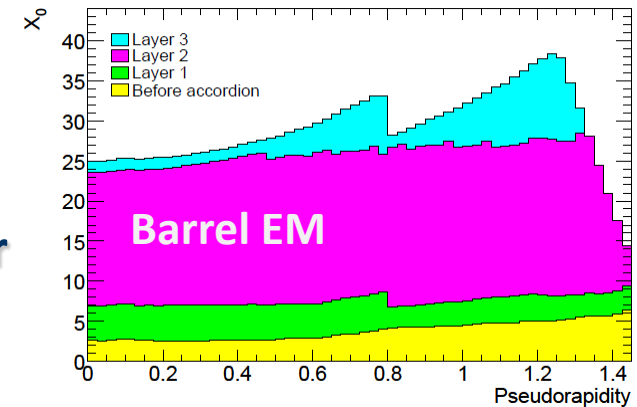
About 190,000 readout channels

Highly efficient particle identification

Jet substructure resolution capabilities

Local hadronic calibration using spatial signal distributions

Signal processing optimized for pile-up (see appendix)



Calorimeter signal scales

Electromagnetic (EM) scale for calorimeter cells

Derived from electron test beams & simulations

Hadronic (HAD) scale from cell signal weighting

2 different schemes – global in jet context, and local in cluster context

Signal summation

Towers (EM scale only)

sum EM energy of all or selected cells on regular $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ grid

Clusters (EM & HAD scale)

Collect topologically connected cells with significant signals into 3-dim energy blobs

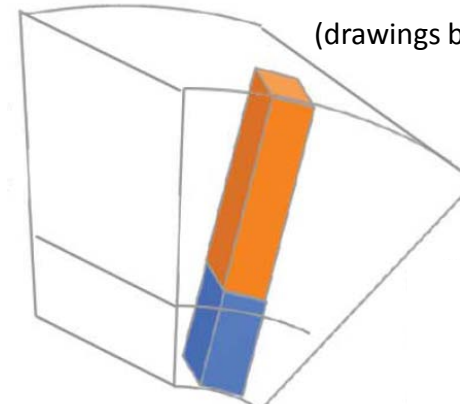
Implies noise suppression at cell level – can be used to select cells before tower formation

Can be used to apply local HAD scale

Towers and clusters are massless pseudo-particles

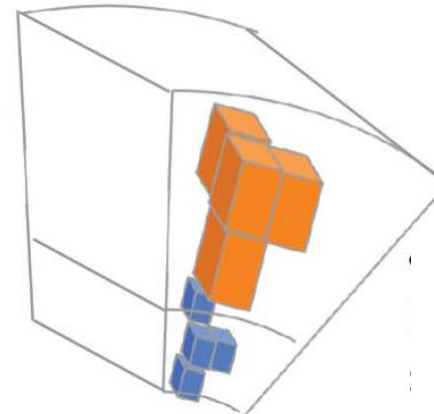
Convention in ATLAS

Can really only accept $E > 0$ input in jet finding

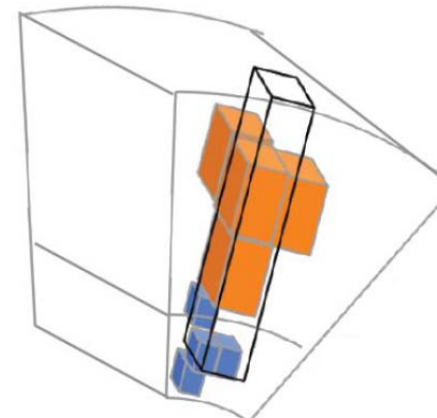


(drawings by K. Perez, Columbia University)

Unbiased calorimeter tower is a "slab" of energy in a regular pseudorapidity-azimuth grid (each tower covers the same area in this frame of reference)



Topological cell cluster is a "blob" of energy dynamically located inside the calorimeter (even crossing sub-detector boundaries)



Noise suppressed towers are sparsely populated slabs of energy in a regular pseudorapidity-azimuth grid (each tower covers the same area in these coordinates)

Calorimeter signal scales

Electromagnetic (EM) scale for calorimeter cells

Derived from electron test beams & simulations

Hadronic (HAD) scale from cell signal weighting

2 different schemes – global in jet context, and local in cluster context

Signal summation

Towers (EM scale only)

sum EM energy of all or selected cells on regular $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ grid

Clusters (EM & HAD scale)

Collect topologically connected cells with significant signals into 3-dim energy blobs

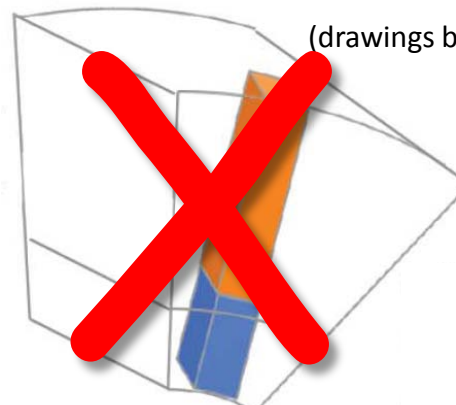
Implies noise suppression at cell level – can be used to select cells before tower formation

Can be used to apply local HAD scale

Towers and clusters are mass-less pseudo-particles

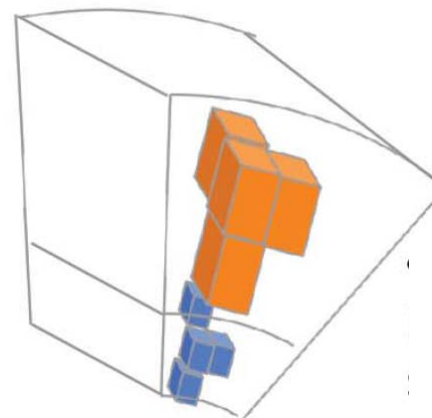
Convention in ATLAS

Can really only accept $E > 0$ input!

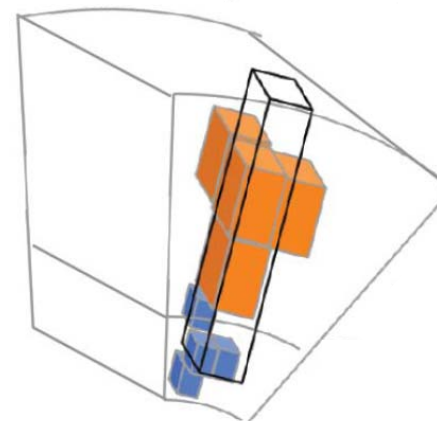


(drawings by K. Perez, Columbia University)

Unbiased calorimeter tower is a "slab" of energy in a regular pseudorapidity-azimuth grid (each tower covers the same area in these coordinates)



Topological cell cluster is a "blob" of energy dynamically located inside the calorimeter (even crossing sub-detector boundaries)



Noise suppressed towers are sparsely populated slabs of energy in a regular pseudorapidity-azimuth grid (each tower covers the same area in these coordinates)

Electromagnetic (EM) scale

Suffers from large fluctuations due to calorimeter signal features

Static calibration without jet-by-jet corrections for hadronic signal characteristics

Least algorithm biases

Good control for systematic uncertainties for first collision data

Still basis of present default jet calibration in ATLAS

Hadronic (HAD) scales

Attempt to reduce signal fluctuations dynamically

Implicit or explicit use of calorimeter signal features, e.g. cell energy density, spatial cell signal distribution

Global Cell Weighting (GCW) in jet context

Apply cell signal weights derived in resolution minimalization fits of matching particle level and calorimeter jets in MC

Correlates dead material and other jet particle energy losses (magnetic field) with weights from non-compensation

Can be applied to cells in clusters or towers

Local Cell Weighting (LCW) in cluster context

Cluster classification tags clusters with photon-like shower shapes

Applying cell signal weights in hadronic cluster reconstructs total deposited energy at cluster location + near-by dead material losses

Factorizes corrections for $e/h \neq 1$, out-of-cluster, and local dead material losses

No correction for acceptance losses (dead material & magnetic field) in jet context



EM scale jet response:

$$\begin{pmatrix} E_{0,\text{jet}} \\ \vec{p}_{0,\text{jet}} \end{pmatrix} = \begin{cases} \sum_{j=1}^{\text{towers} \in \text{jet}} \left(E_{0,\text{tower}} = |\vec{p}_{0,\text{tower}}|, \vec{p}_{0,\text{tower}} \right)_j \\ \sum_{k=1}^{\text{clusters} \in \text{jet}} \left(E_{0,\text{cluster}} = |\vec{p}_{0,\text{cluster}}|, \vec{p}_{0,\text{cluster}} \right)_k \end{cases} \approx \underbrace{[0.6 - 0.9]}_{\sim 20 \text{ GeV} \rightarrow \sim 1 \text{ TeV}} \times \begin{pmatrix} E_{\text{jet,true}} \\ \vec{p}_{\text{jet,true}} \end{pmatrix}$$

Required calibration/corrections for EM scale jet signals

Missing e/h, dead material, acceptance corrections

Energy dependent signal deficits require corresponding corrections

Corrected by application of jet level correction functions

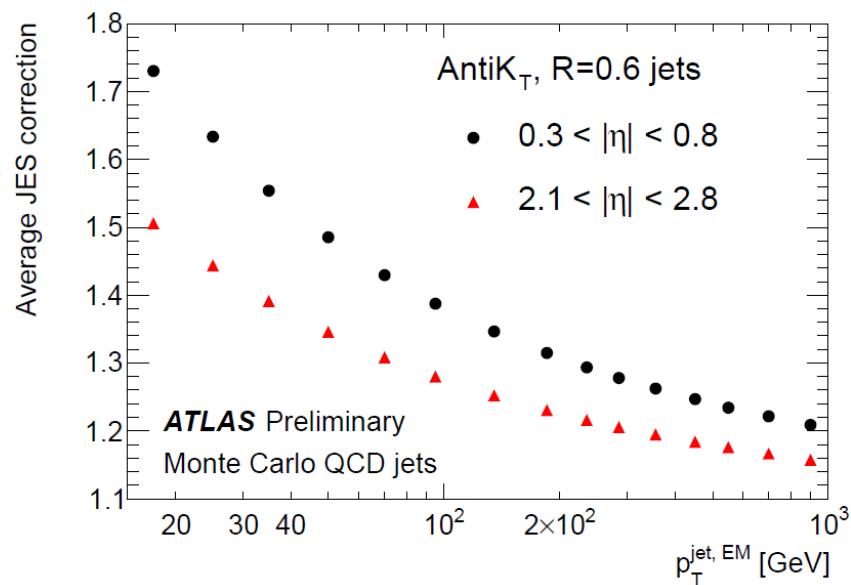
Functions determined with MC (PYTHIA QCD di-jets)

Parameterized in bins of true jet pT and pseudorapidity

For each jet algorithm and configuration (e.g., jet size)

Numerical inversion techniques

Maintain response function shape in given phase space bin when transforming input variables from true to measured scales



Response function for true variables:

$$R_{\text{jet}}(p_{T,\text{true,jet}}, \eta_{\text{jet}}) = \left\langle p_{T,\text{EM,jet}} / p_{T,\text{true,jet}} \right\rangle (p_{T,\text{true,jet}}, \eta_{\text{jet}})$$

Response function for measured variables:

$$R_{\text{jet}}(p_{T,\text{EM,jet}}, \eta_{\text{jet}}) = R_{\text{jet}}(R_{\text{jet}}(p_{T,\text{true,jet}}, \eta_{\text{jet}}) \cdot p_{T,\text{true,jet}}, \eta_{\text{jet}})$$

Determination of contributions from various sources

Determine default calibrations and corrections with given MC

- Best detector geometry knowledge
- Best fitting signal simulation models
- Workhorse QCD di-jet physics generator Pythia

Vary used models and setups within reasonable boundaries

- Models excluded by data not considered
- Only realistic detector misalignments and material budgets
- Use of measured range of experimental conditions

Reconstruct jets in alternative setups

- Use calibrations from default configuration

Physics modeling contributions

Underlying event

- Use alternative tunes for Pythia

Jet shapes

- Herwig/Jimmy instead of Pythia (cluster instead of string fragmentation)

Detector signal modeling

Variations of detector description

- Dead material budget and alignment

Different hadronic shower models in Geant4

- FTFP_BERT instead of default QGSP_BERT

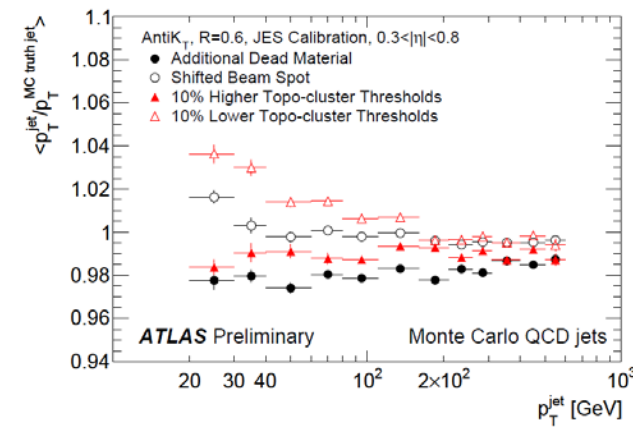
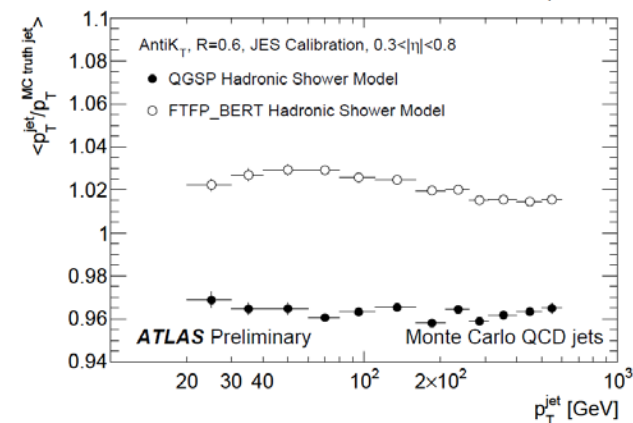
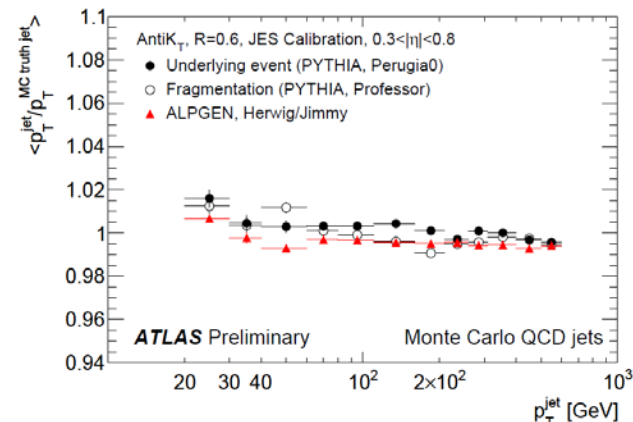
Variation of noise suppression thresholds in clustering

- +/- 10% change in threshold

Experimental conditions

Effect of shifted beam spot

- Corresponding shift in response variations, p_T /energy



Putting it all together

Conservative estimates for overall systematic uncertainty for first data

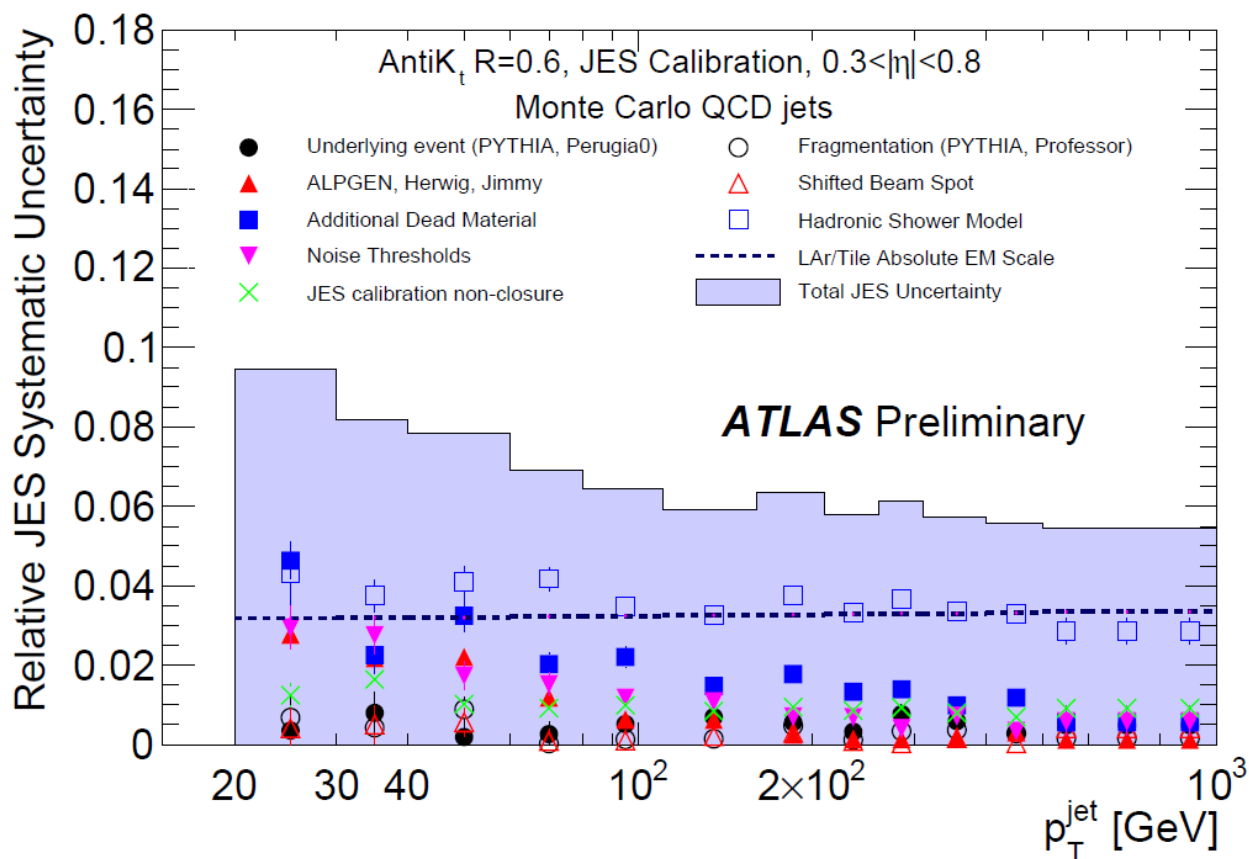
Dominant contribution for all p_T from hadronic shower models

Dead material important for low p_T regime

Similar uncertainties from E/p and testbeam

Jets composed from single particle response in collision and testbeam data and MC

Systematic uncertainty
as of summer 2010!
Significant
improvements can be
expected for winter
2011 conferences!



Use of hadronic scales

Resolution improvements (GCW, LCW)

More dynamic calibrations sensitive to individual shower character in jet (GCW) or clusters (LCW)

Jet constituent calibrations

So far jets reconstructed from EM scale calorimeter signals in towers or clusters

Using input signals on hadronic scale improves performance and stability of kT and Anti-kT due to better relative signal calibration between photon and hadron component of jet

Consistent hadronic final state

LCW allows use of signals on the same scale for jets and missing transverse energy

Hadronic scales still need final JES corrections

LCW misses jet energy losses from particles completely escaping the calorimeters (dead material, magnetic field) – GCW not fitted in calorimeter crack regions etc.

Complex calibrations derived from MC

Use of cell signal density

Cluster (LCW) and jet context (GCW) cell weighting

LCW uses cluster structures

Location, shape and sizes directly and indirectly used

Concern about model dependence

Observables feeding calibrations need to be well simulated

Some dependence on shower model details

Validation of calibration input in collision data

First task to gain confidence in models and parameters

Comparison of signal spectra in minimum bias and jet events

Factorized approach in LCW

Allows understanding each calibration and correction step individually

Jet shapes and energy sharing

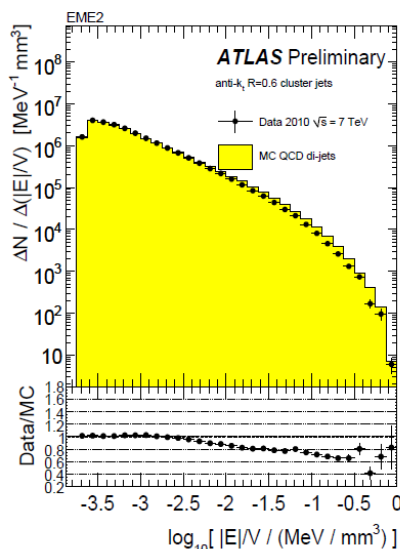
Direct and indirect input on hadronic scale

Level of understanding also important for jet physics and hadronic event shapes!

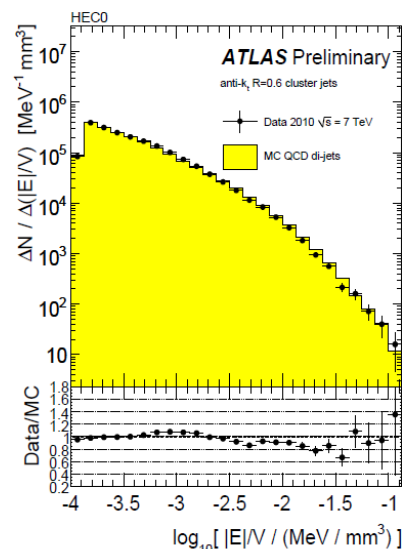


Calorimeter cell
energy density
distributions (e.g.,
for GCW, LCW)

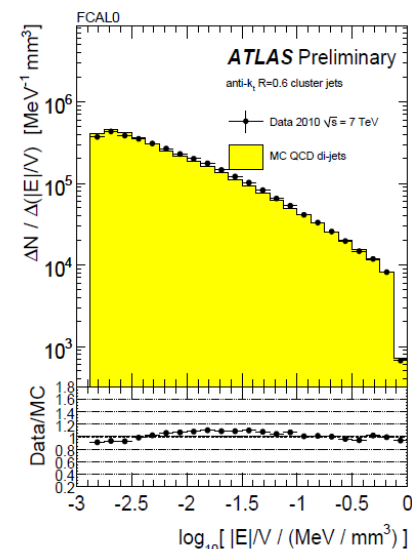
Cells and clusters
inside jets!



(a) Second layer of EM endcap

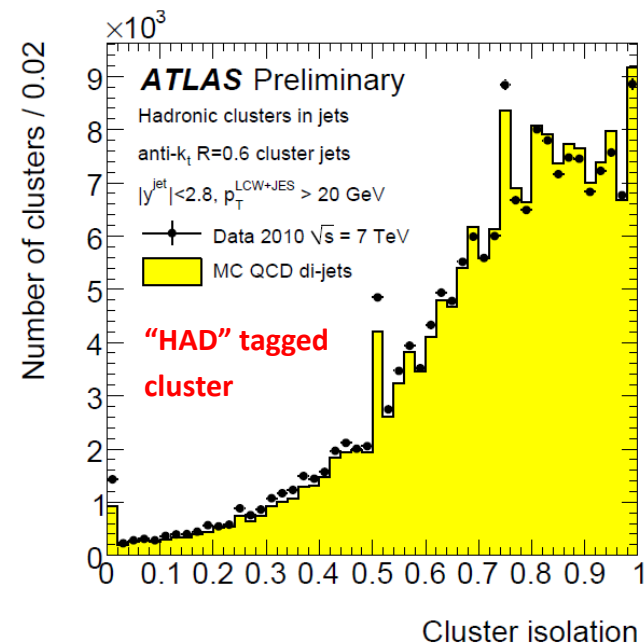
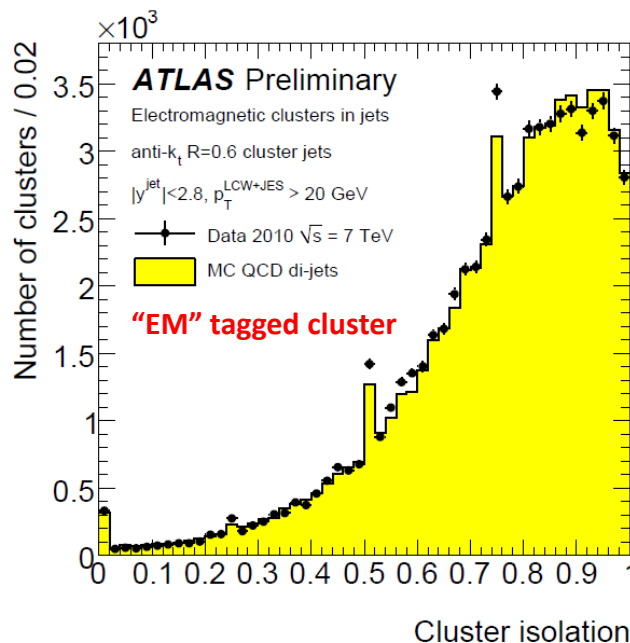


(b) First layer of HEC



(c) First layer of FCAL

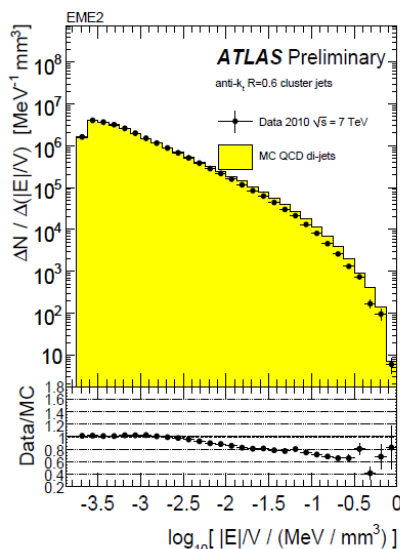
Calorimeter cluster
isolation measure
(LCW)



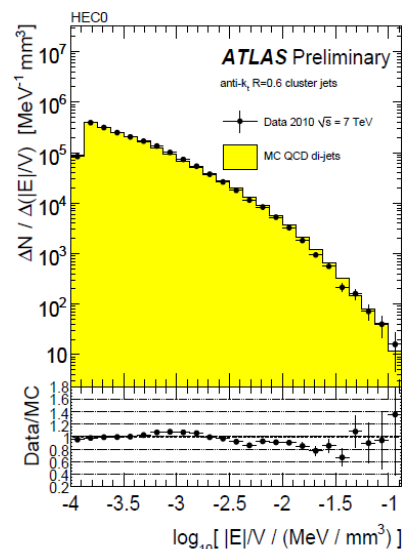
(ICHEP2010)

Calorimeter cell
energy density
distributions (e.g.,
for GCW, LCW)

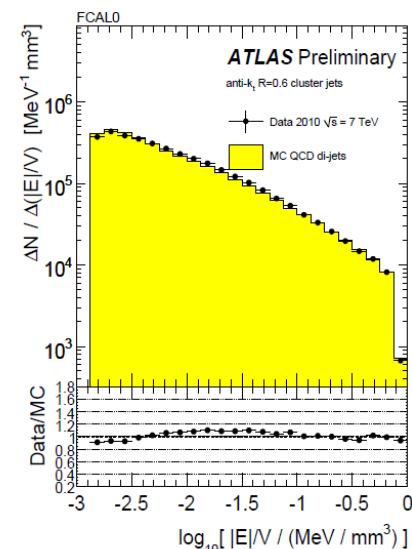
Cells and clusters
inside jets!



(a) Second layer of EM endcap

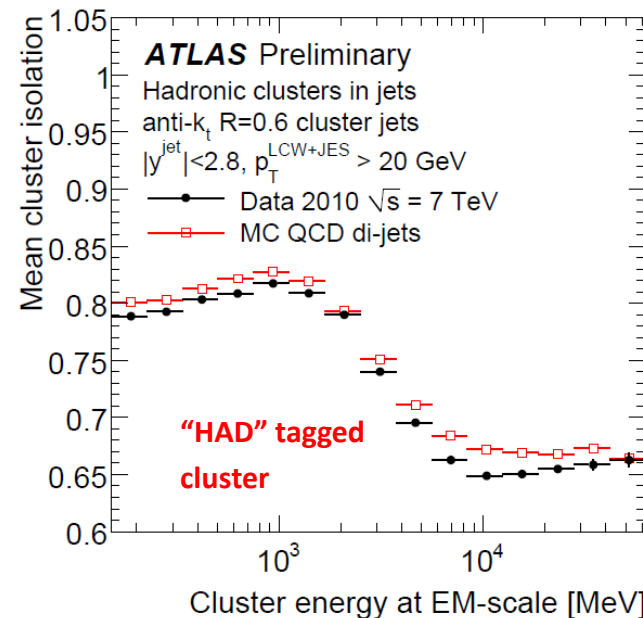
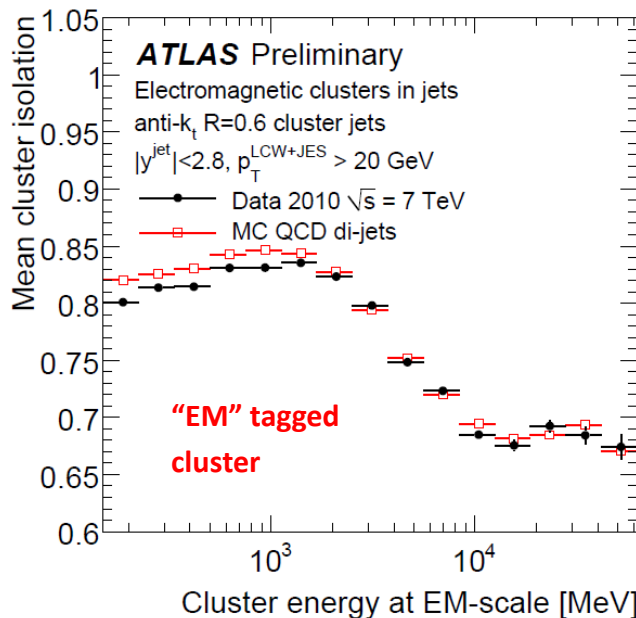


(b) First layer of HEC



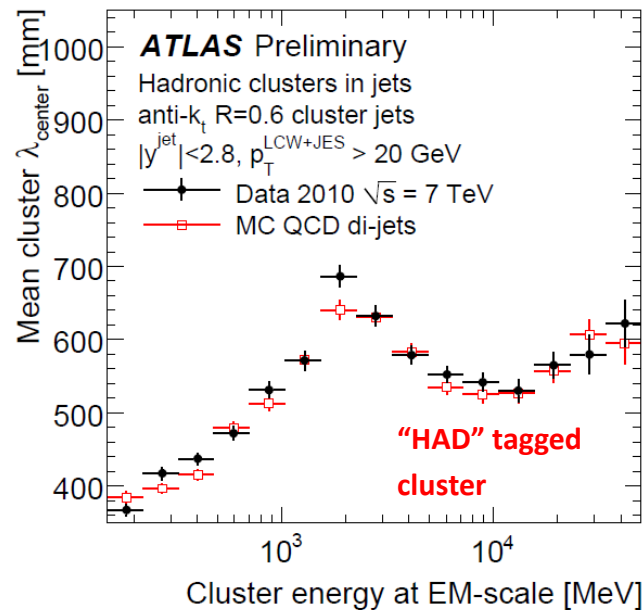
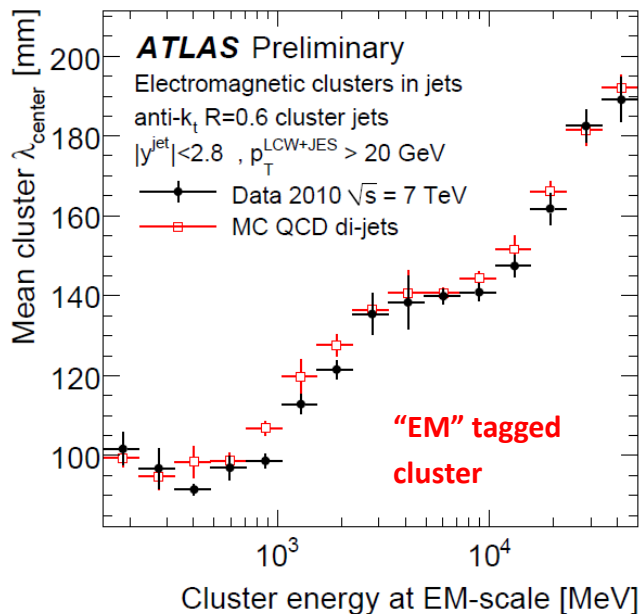
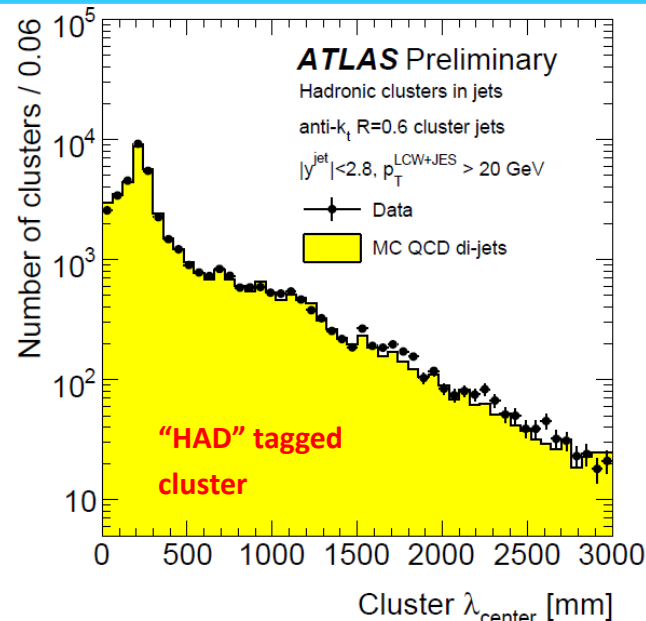
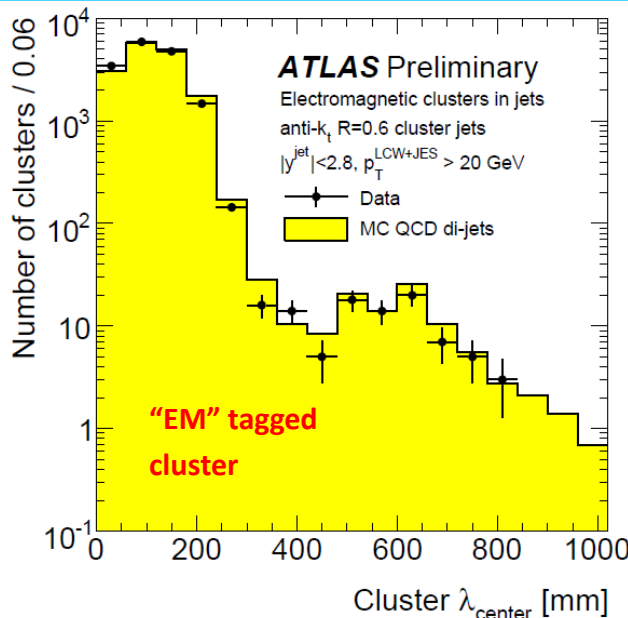
(c) First layer of FCAL0

Calorimeter cluster
isolation measure
(LCW)



Calorimeter cluster depth in calorimeter (LCW)

Clusters inside jets!



(ICHEP2010)

Test factorized calibration model

All functions and parameters derived from single particle simulations

Non-trivial assumption that they work in collision physics environment!

Factorization allows validation of each calibration/correction

HAD clusters: cell weighting, out-of-cluster, dead material corrections

EM clusters: out-of-cluster, dead material (not the same as for HAD!)

Checked for clusters in and outside of jets

Shown here for clusters inside jets

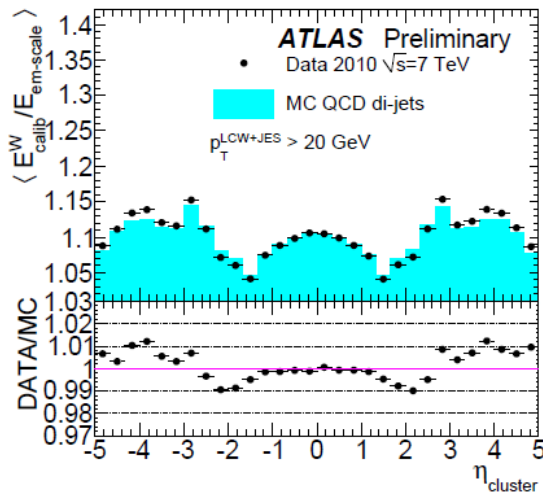
Promising for sub-jet calibration

Local hadronic scale is jet constituent scale

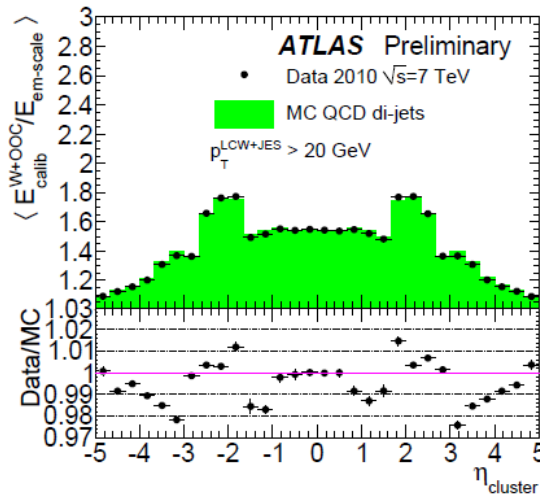
Can be understood at the 1-3% level!

Needs to be verified with high p_T jets

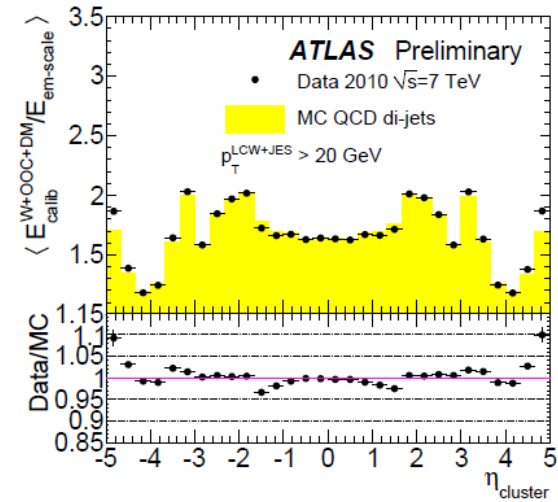
Boosted hadronic W and top decays become slowly available in data



(a) Hadronic response weights



(b) Out-of-cluster weights



(c) Dead material weights

Kinematics and composition

Understanding dependence of reconstruction quality on calorimeter signal choice

Direction and p_T for cluster and tower jets

Number of constituents

Tower and cluster jets compared to particle level

Other observables related to composition

Longitudinal energy sharing – access to photon content of jets by measuring energy fractions in EM calorimeters (?)

Jet shapes

Width

Calorimeter width measurement affected by particle flow and shower spread

Track jets provide (reference) width with independent bias

Annular transverse energy flow

Jet fragmentation and source (light quark/gluon) sensitivity

Use in jet-by-jet calibration to be explored

Note

Nearly all comparisons in the following plots are to Pythia with the ATLAS MC09 tune!

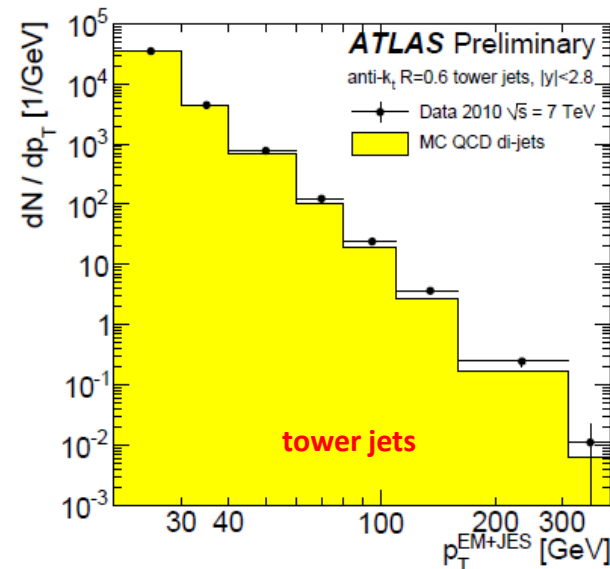
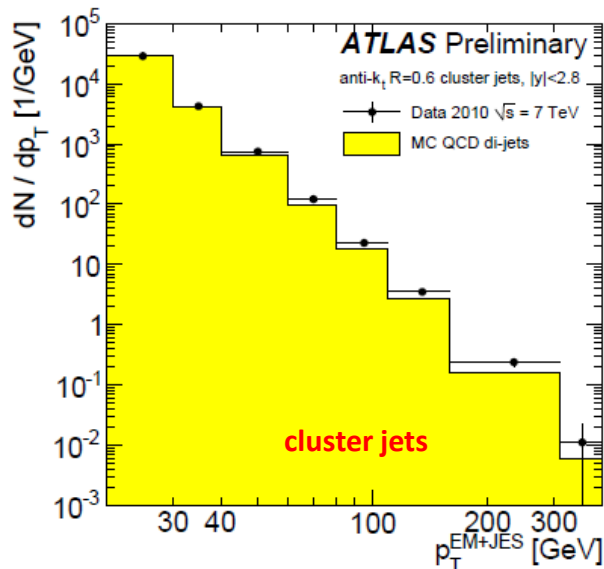
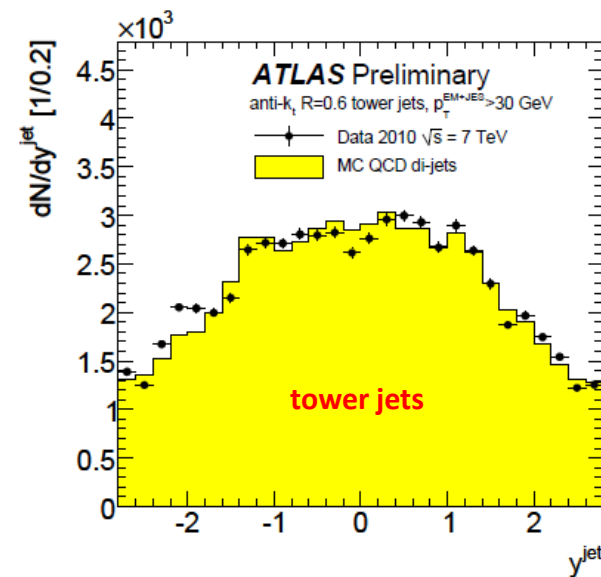
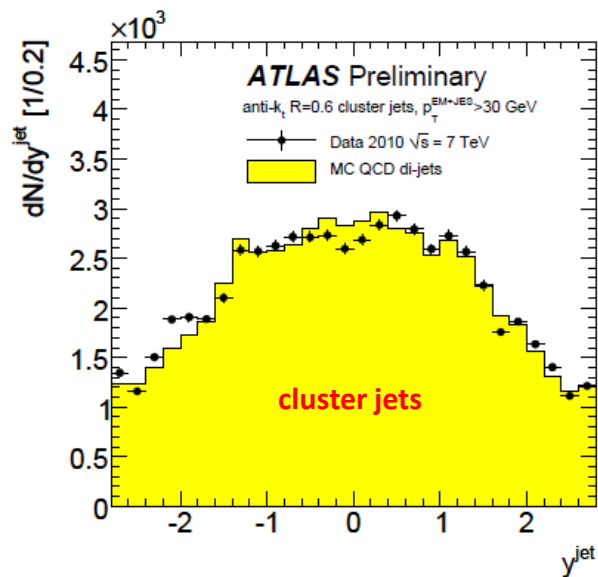
... if not specifically mentioned otherwise!



Jet number density in pseudorapidity

$$p_{T,\text{jet}} > 30 \text{ GeV}$$

Jet p_T spectra



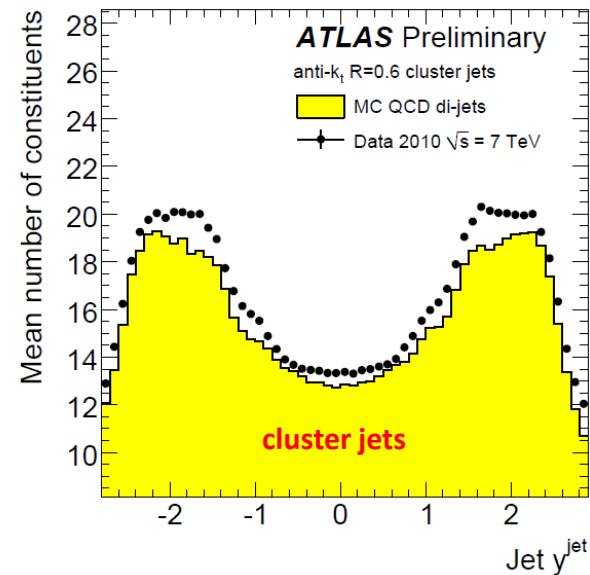
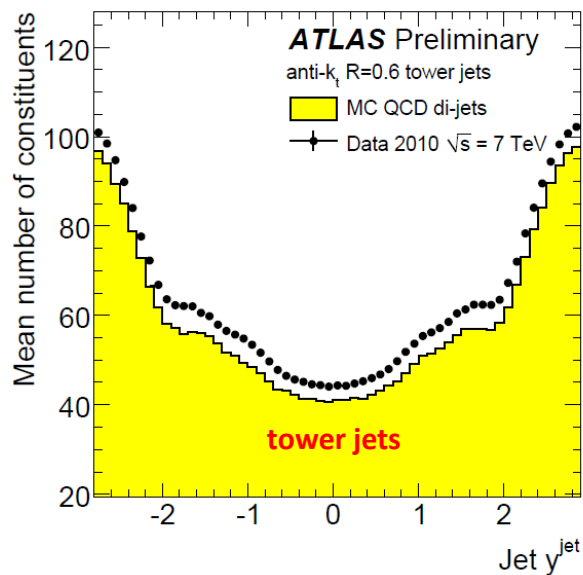
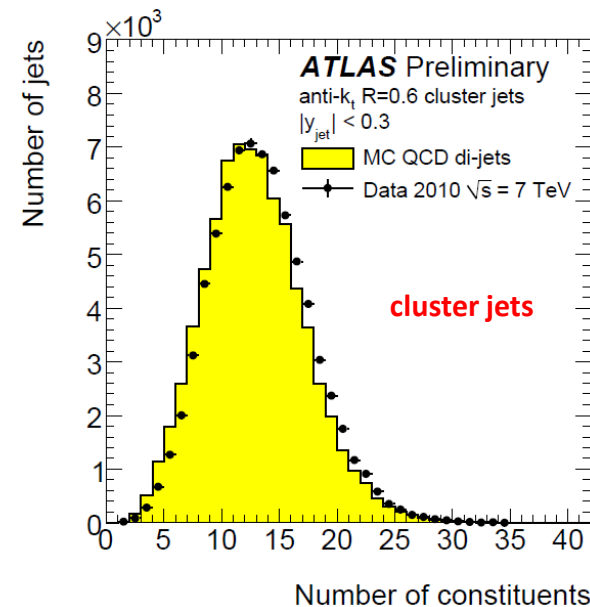
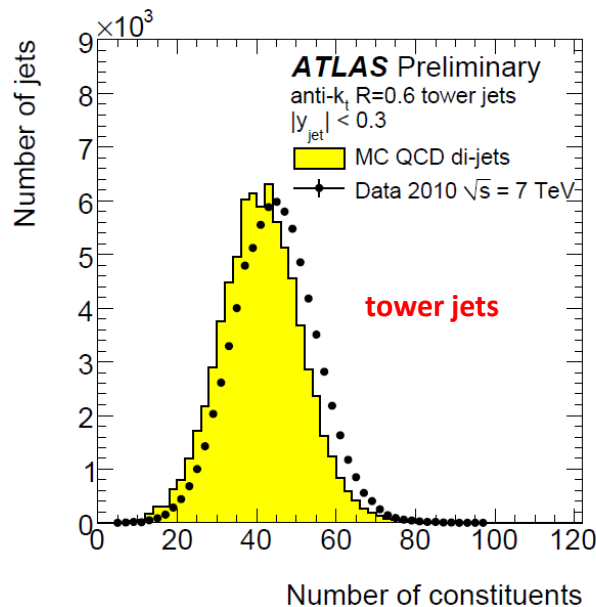
(a) Topological cluster jets

(b) Noise-suppressed tower jets

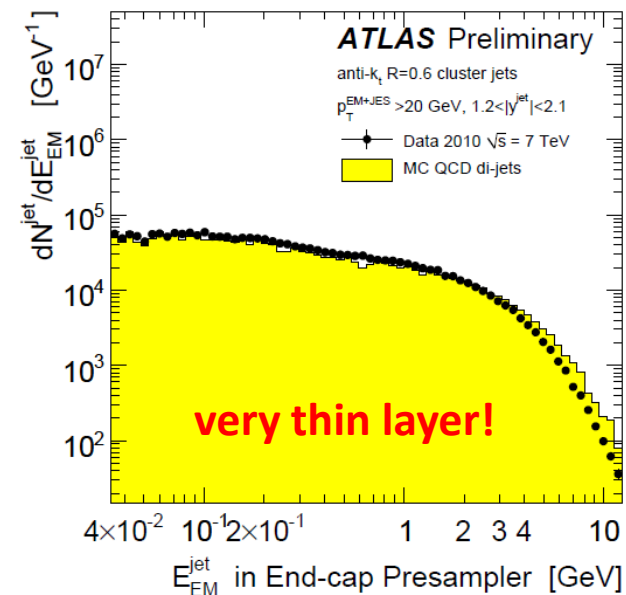
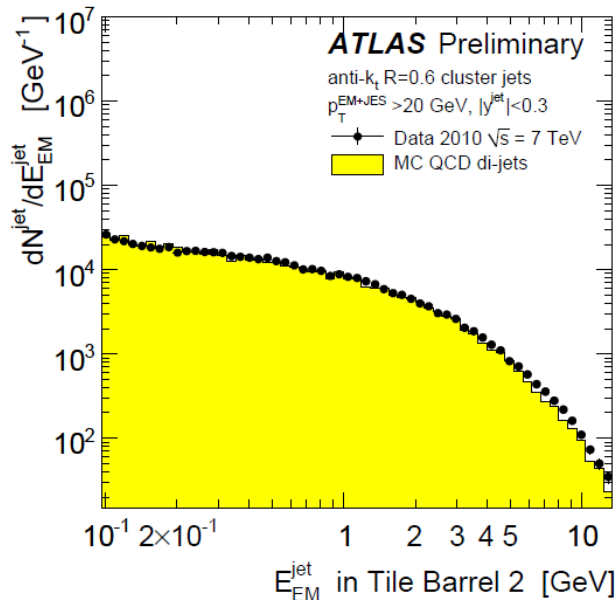
Number of constituents in jets

$$p_{T,\text{jet}} > 30 \text{ GeV}$$

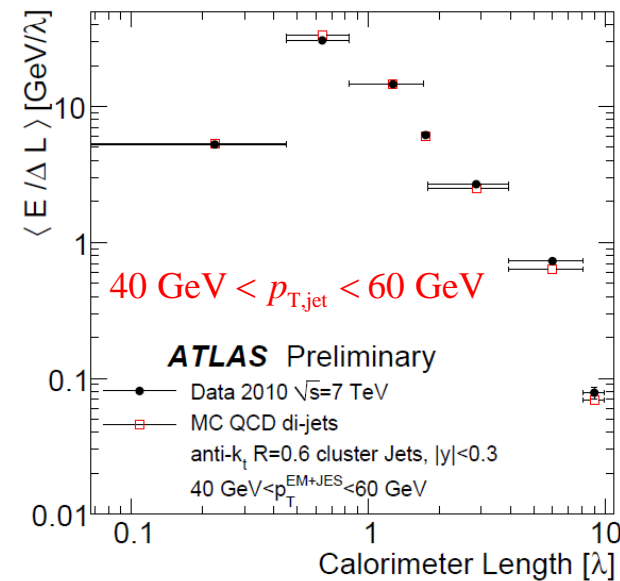
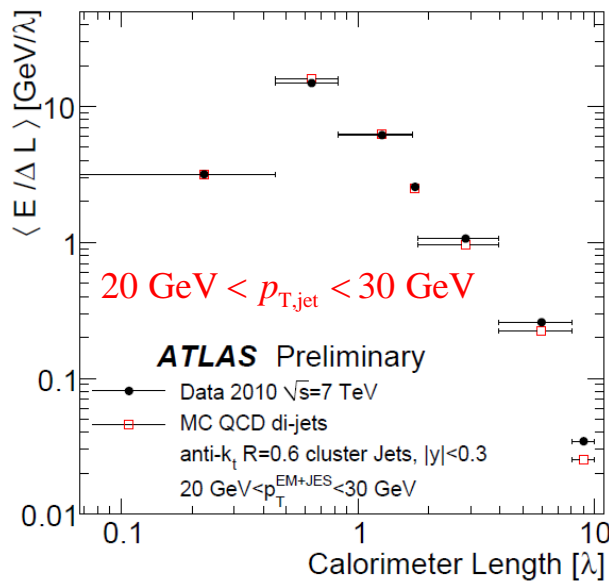
Jets in data have more constituents on average



Energy distribution in calorimeter samplings from jets – spectra

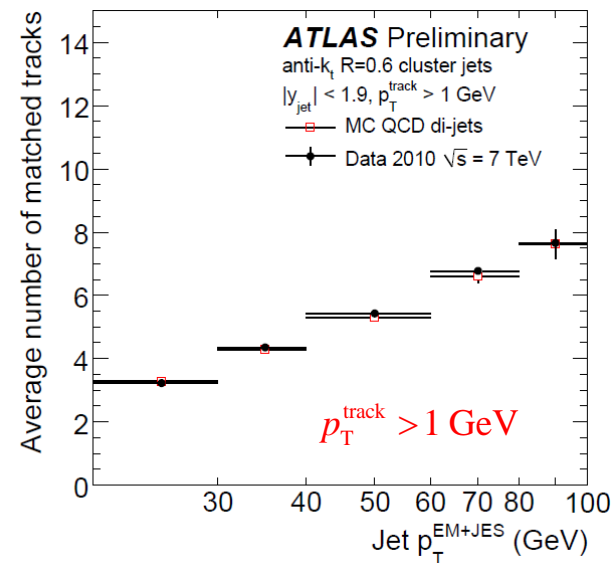
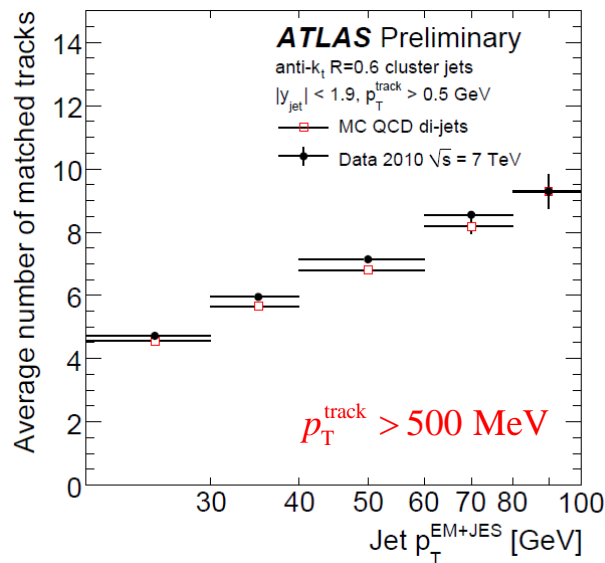
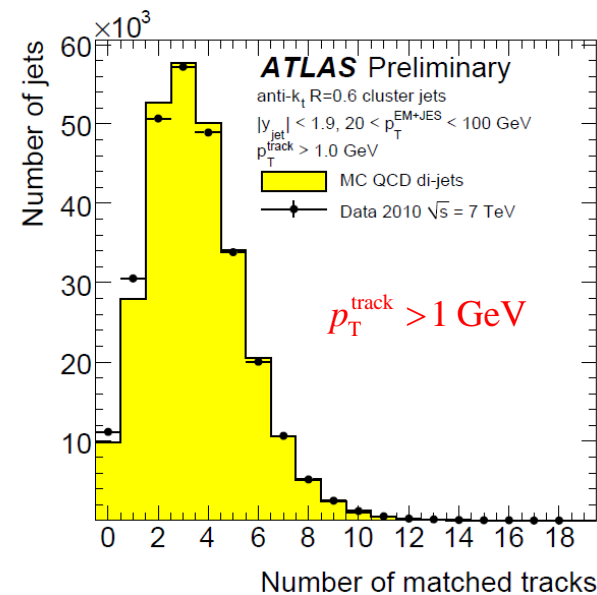
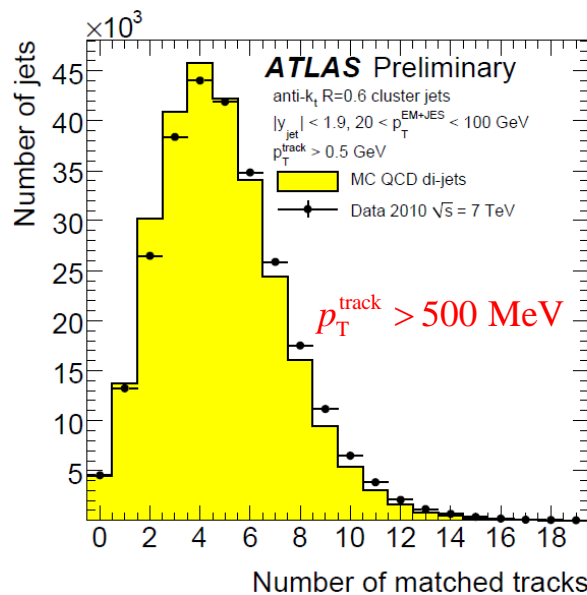


Energy sharing in calorimeter samplings – longitudinal shower profiles in jets



Number of tracks pointing to calorimeter jets – sensitive to fragmentation (model)

Average number of reconstructed tracks pointing to calorimeter jet – more soft tracks in data



$20 \text{ GeV} < p_{\text{T,jet}} < 100 \text{ GeV}$

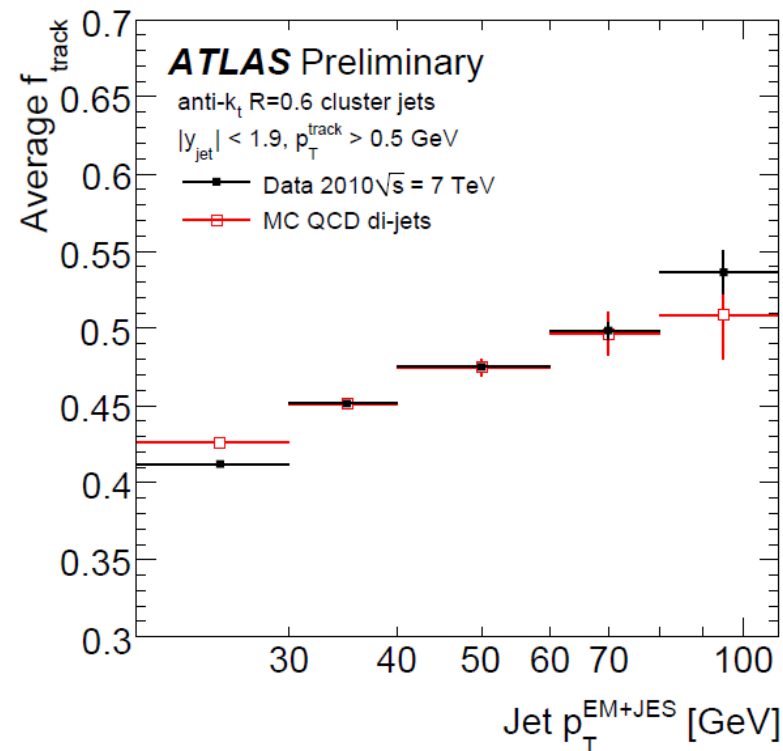
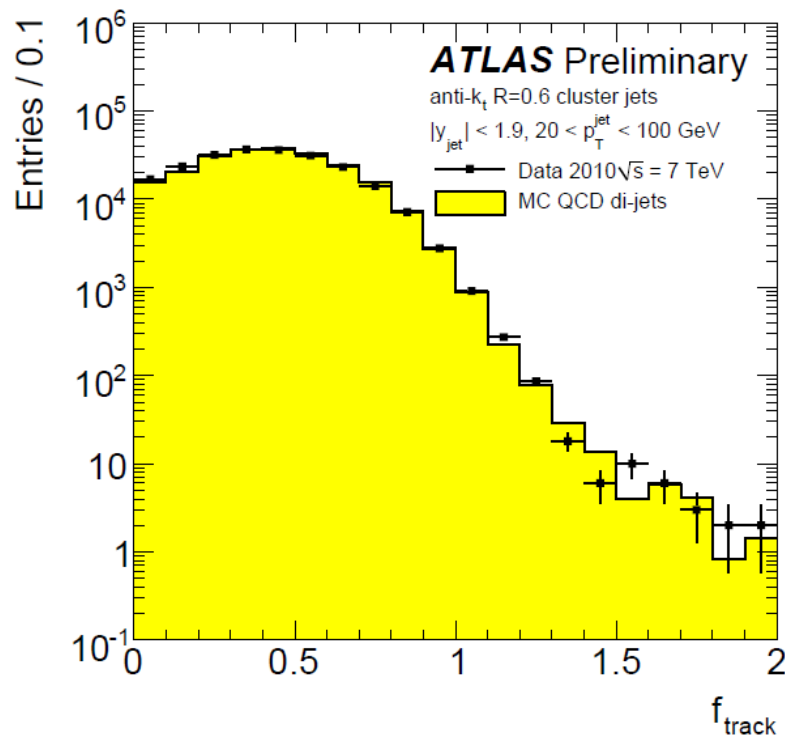
Track p_T fraction

Tool for fragmentation validation

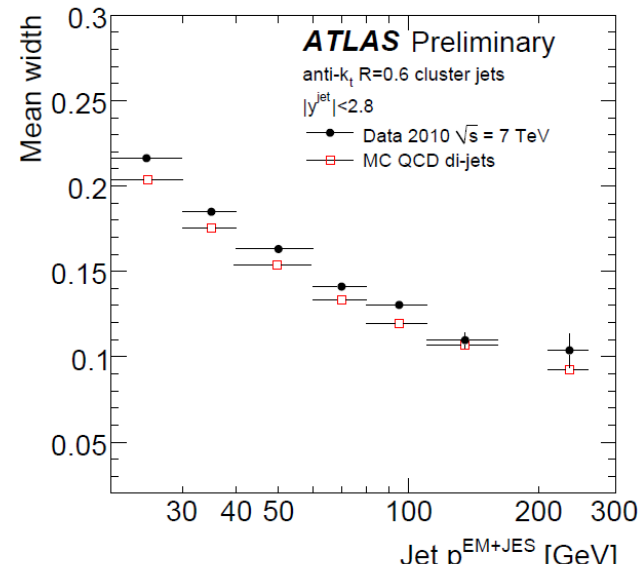
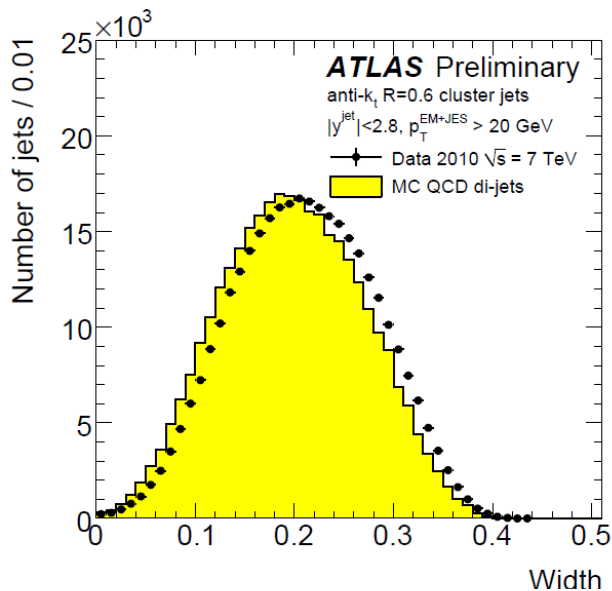
Needs to be understood in the context of pile-up in jets (least biased estimator?)

Handle for jet-by-jet calibration

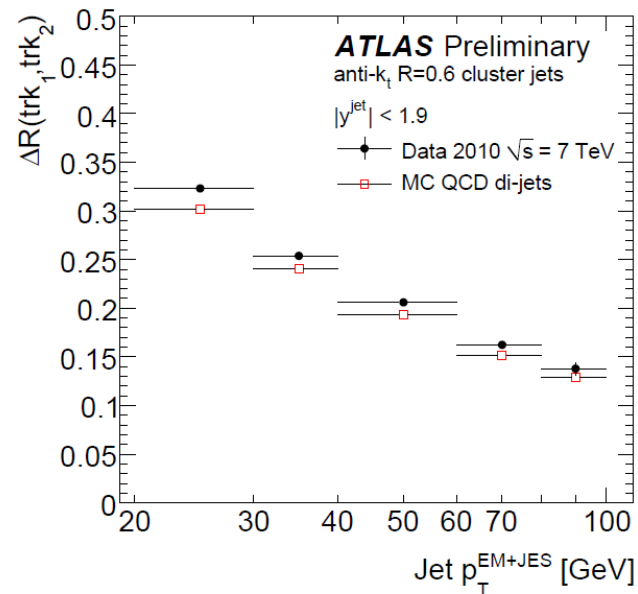
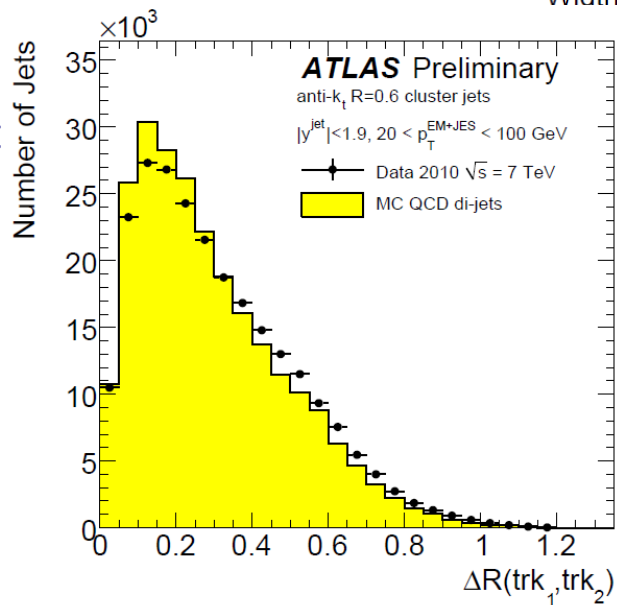
Large track momentum fraction indicates hadron rich jet – can be exploited for jet calibration

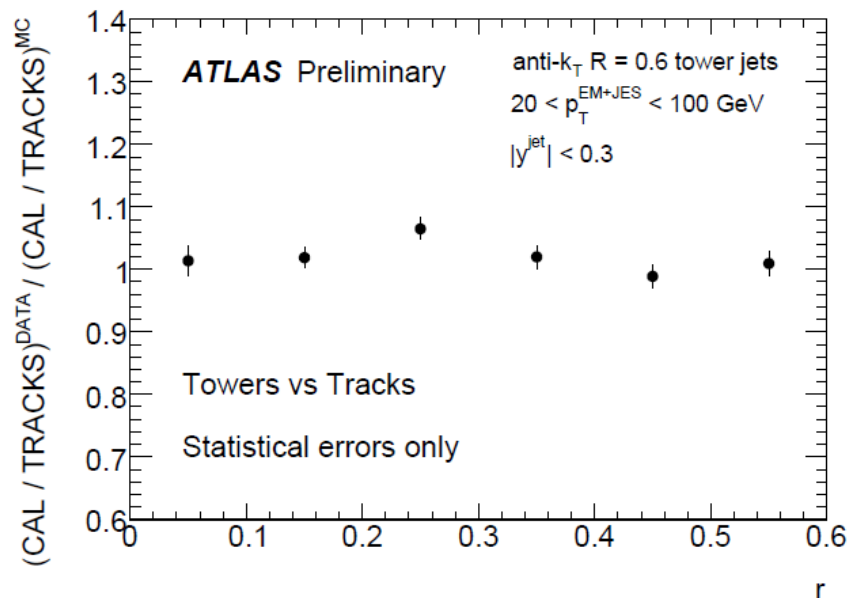
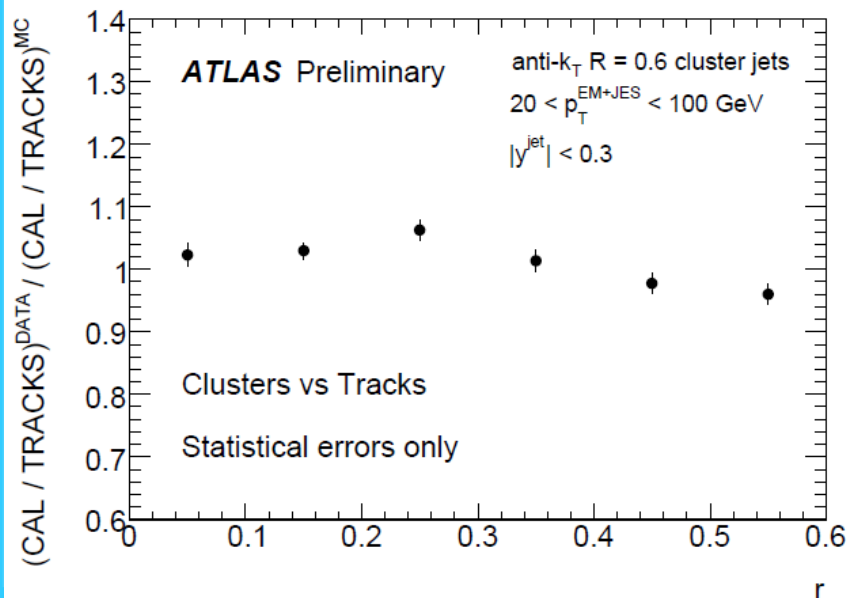
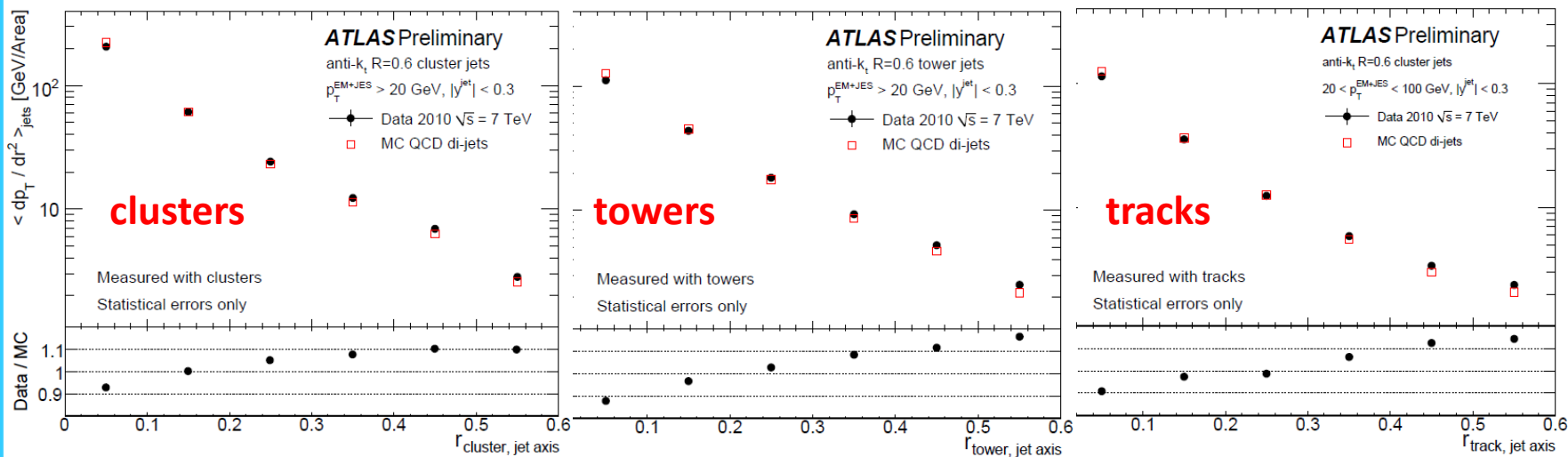


Calorimeter jet width in data and MC



Track jet width estimate (independent measurement)

(a) $\Delta R(\text{trk}_1, \text{trk}_2)$ of the two hardest tracks in jets(b) Dependence of $\Delta R(\text{trk}_1, \text{trk}_2)$ with p_T^{jet}



(ICHEP2010)

Potential physics message: fragmentation/UE different in data and MC – to be followed up!

Jets seem wider in data than in MC

Direct width measurement

Track width confirms calorimeter observation

Different composition

More towers and clusters in data

Indication of physics sources

Nearby jet activity

More low p_T jets close to hardest jet in data than in MC

Measurement

Distance from hard jet to nearest (soft) jet

Hard jet $p_{Tmin} < p_T < p_{Tmax}$

Neighbouring jet $p_T > 7$ GeV

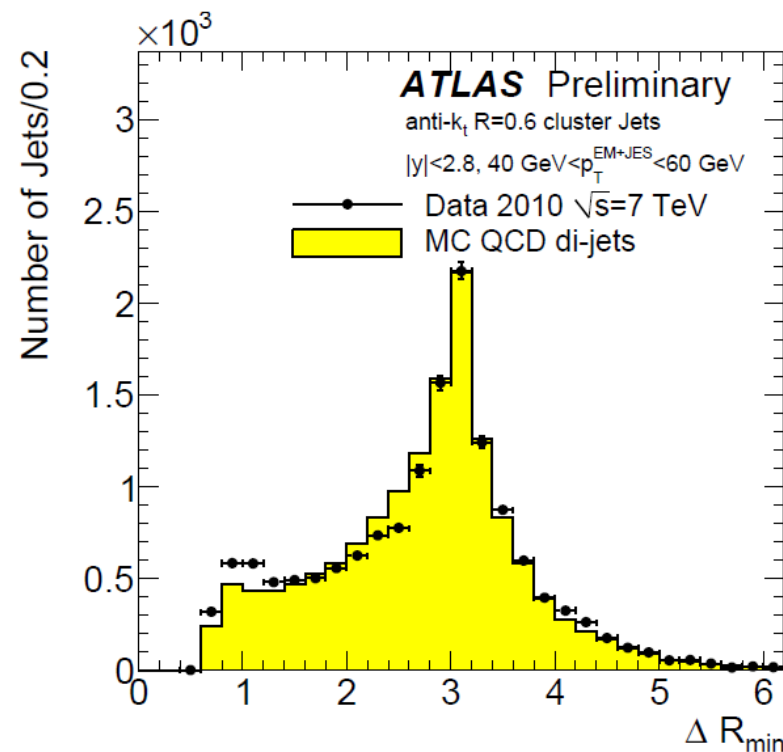
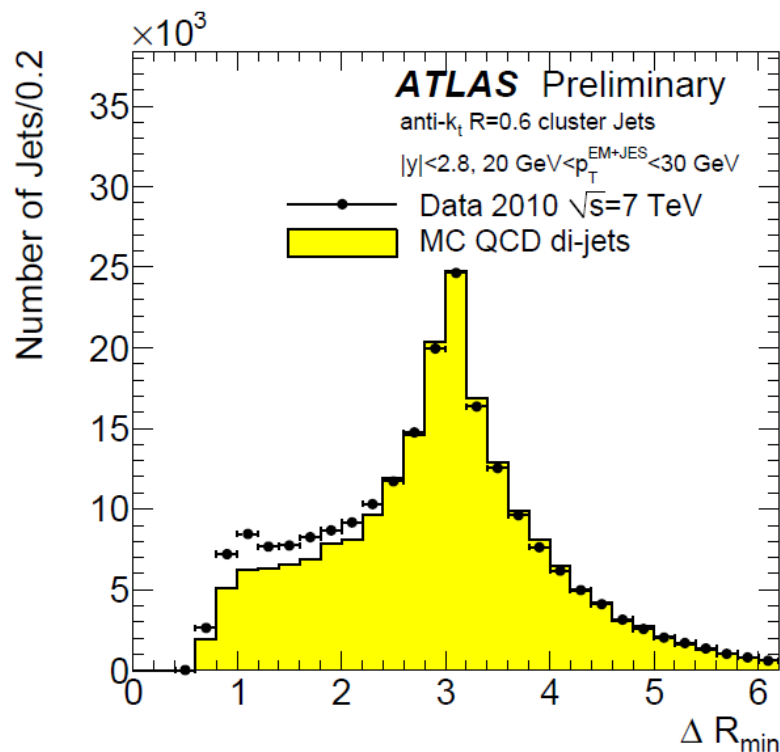
Di-jet structure clearly emerges

Peak at π as expected from back-to-back scattering

Far distance distribution well modeled

More small angle radiation in data?

Unclassified jet fragments?



Testing various MC models and tunes

Differential radial jet shape

Reconstructed from clusters,
corrected to particle level

Unfolding procedure

Bin-by-bin (p_T and η) using fully
simulated Pythia-Perugia2010 (best
match to data)

Correction factors 0.95-1.1, typically

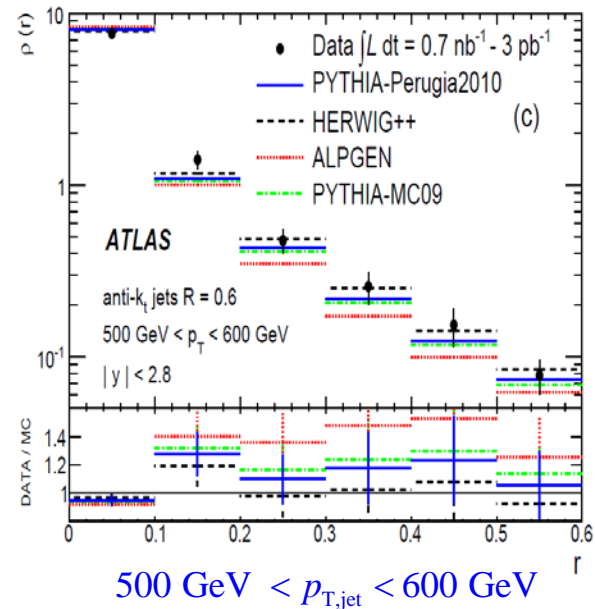
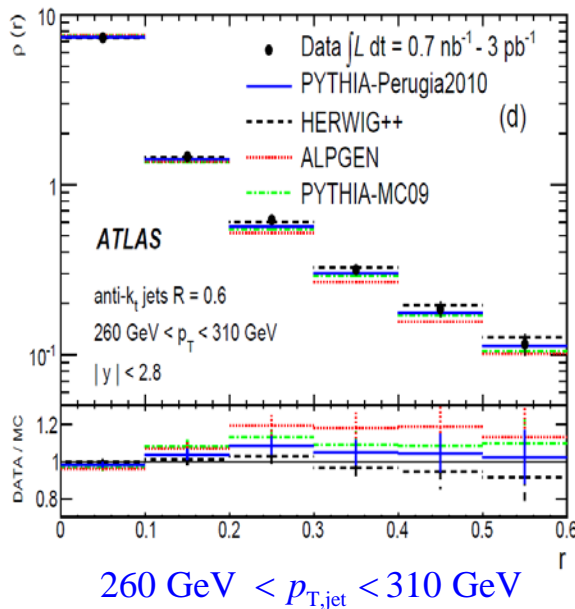
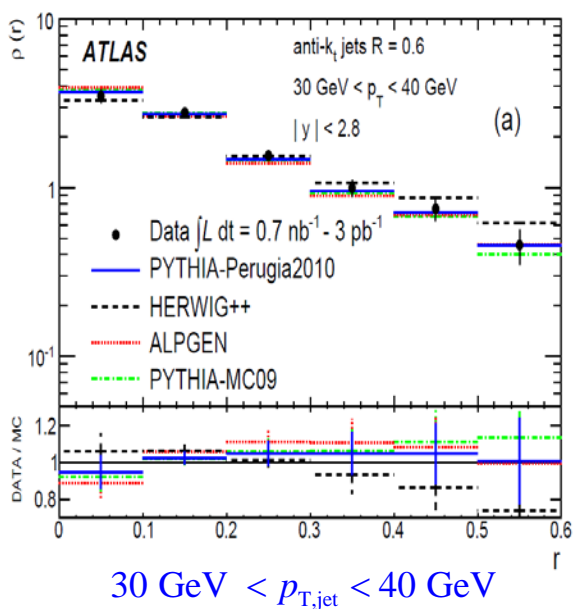
Differential jet shape :

$$\rho(r) = \frac{1}{\Delta r} \frac{1}{N_{\text{jet}}} \sum_{\text{jets}} \frac{p_T(r - \Delta r/2, r + \Delta r/2)}{p_T(0, R)}$$

with $\Delta r/2 \leq r \leq R - \Delta r/2$

Integrated jet shape :

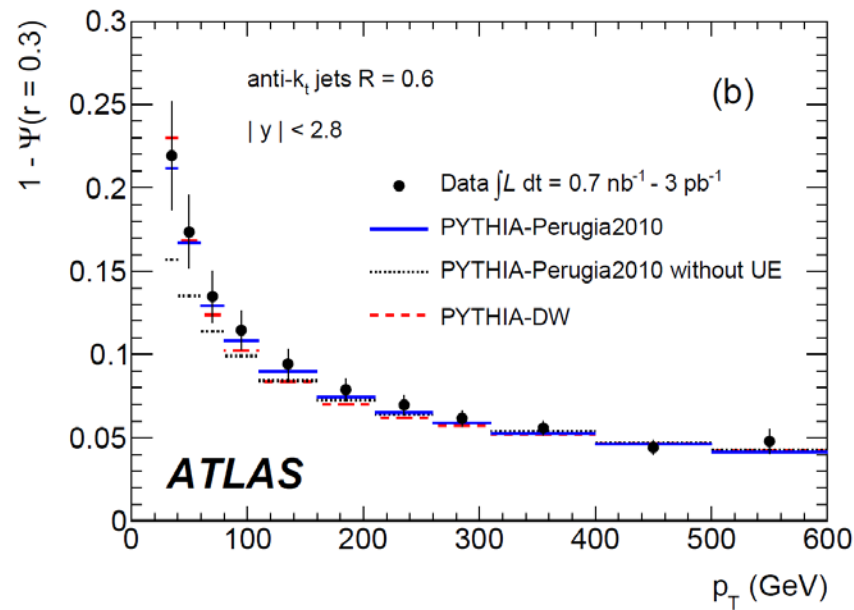
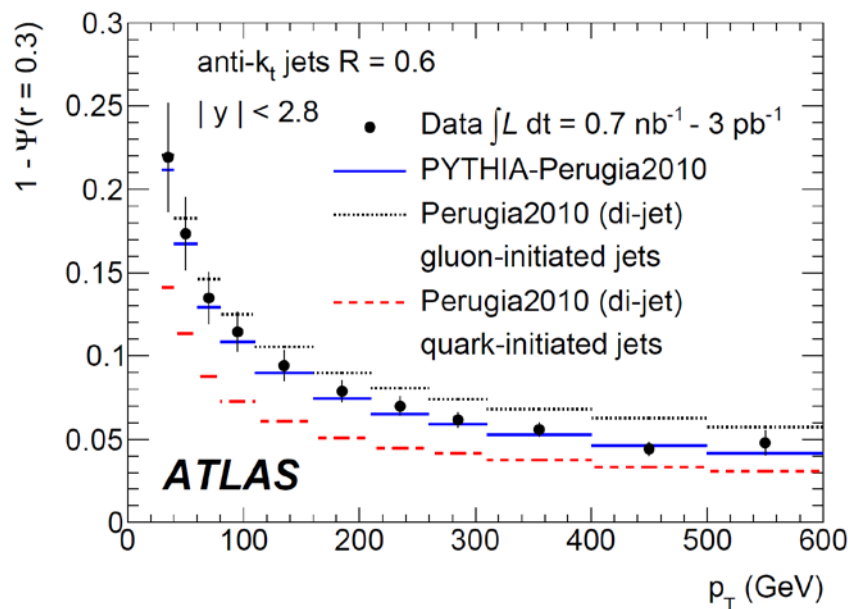
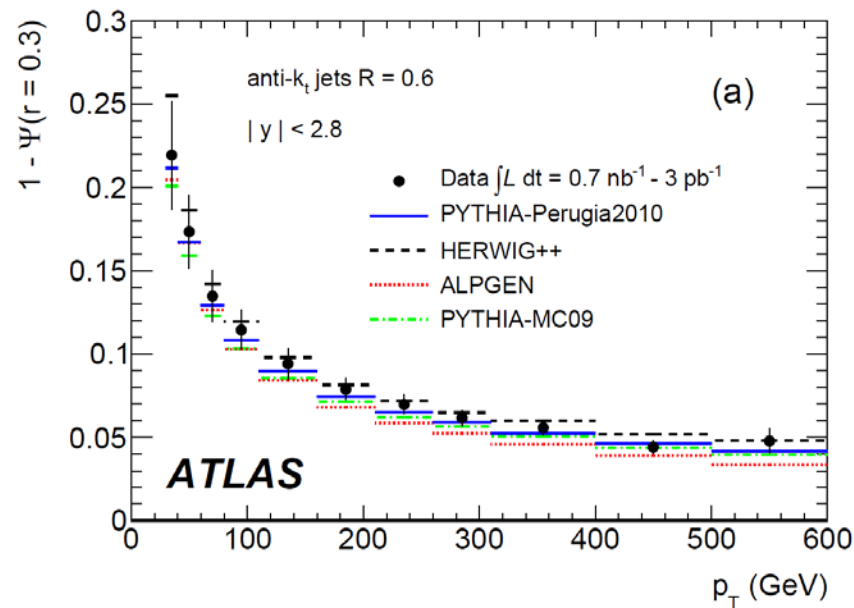
$$\Psi(r) = \frac{1}{N_{\text{jet}}} \sum_{\text{jets}} \frac{p_T(0, r)}{p_T(0, R)} \quad \text{with } 0 \leq r \leq R$$



Testing various MC models and tunes

Integrated jet shape

Measured as energy leaking out of a cone with radius 0.3 around the jet axis



ATLAS jet reconstruction performed very well

Detailed understanding of jet reconstruction in first data

Topological cell clusters perform very well – formation and local calibration well understood

Also confirmed by missing transverse energy reconstruction

Well motivated systematic jet energy scale error delivered quickly

Significant improvements expected in the next few weeks

Some jet structures and shapes already well measured

Cluster jets promising for substructure analysis

Need more studies for high p_T jets to understand possible merging problem

Indications that jets in data are wider than in MC (Pythia)

Calorimeter width measurement confirmed by jet width in tracker

Jet substructure reconstruction and calibration under study

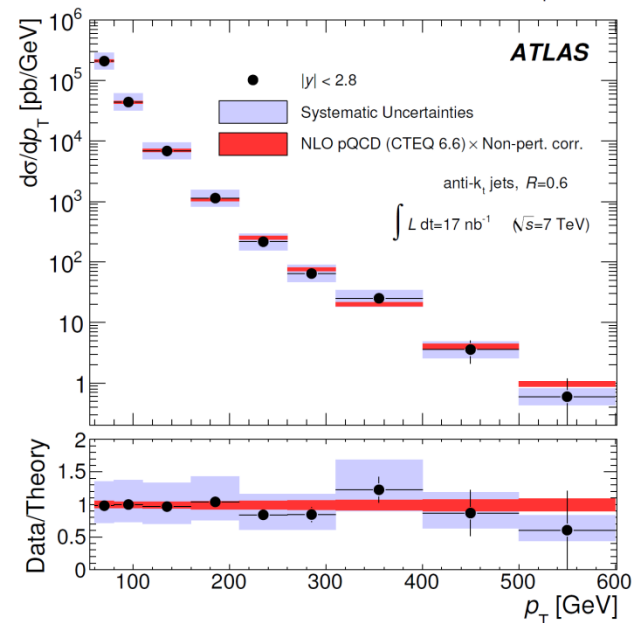
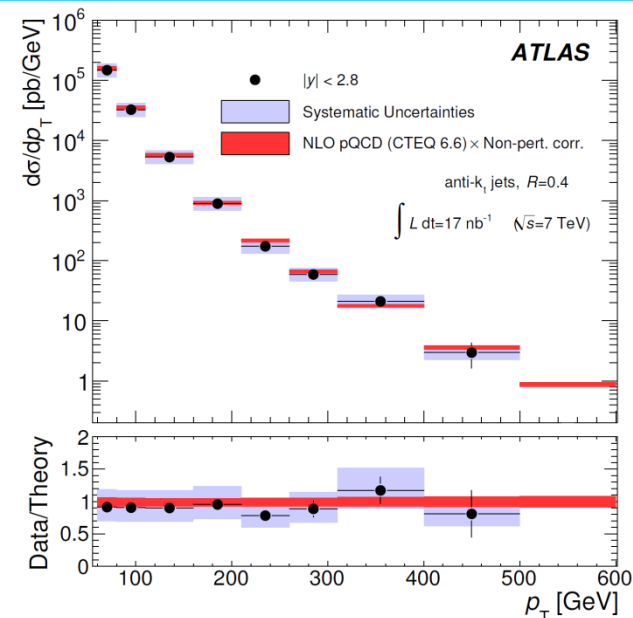
Local hadronic calibration seems applicable within small systematics – promising starting point for sub-jet calibration
 Hope to increase data sample with jets including 2-prong (W) and 3-prong (top) hadronic decays

Single jet mass reconstruction to be validated

Performance not yet completely understood – geometry effect in towers adds mass while hadronic shower merging in clusters reduces reconstructed mass

Lots of interesting aspects still to look at!

And sorry for not being in Boston in person!





Signal extraction tool

Attempt reconstruction of individual particle showers

Reconstruct 3-dim clusters of cells with correlated signals

Use shape of these clusters to locally calibrate them

Explore differences between electromagnetic and hadronic shower development and select best suited calibration

Suppress noise with least bias on physics signals

Often less than 50% of all cells in an event with “real” signal

Some implications of jet environment

Shower overlap cannot always be resolved

Clusters represent merged particle showers in dense jets

Clusters have varying sizes

No simple jet area as in case of towers

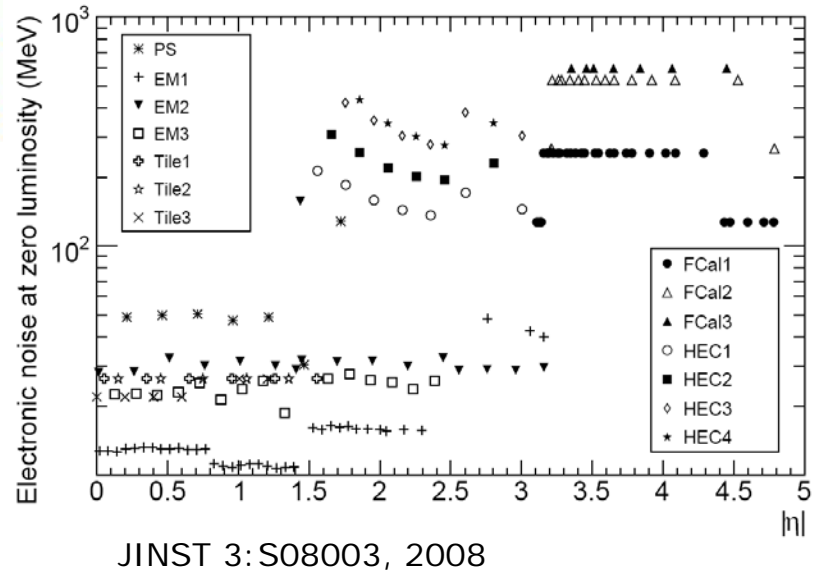
Clusters are mass-less 4-vectors (as towers)

No “artificial” mass contribution due to showering

Issues with IR safety at very small scale insignificant

Pile-Up environment triggers split as well as merge

Note that calorimeters themselves are not completely IR safe



Cluster seeding

Cluster seed is cell with significant signal above a primary threshold

Cluster growth: direct neighbours

Neighbouring cells (in 3-d) with cell signal significance above some basic threshold are collected

Cluster growth: control of expansion

Collect neighbours of neighbours for cells above secondary signal significance threshold

Secondary threshold lower than primary (seed) threshold

Cluster splitting

Analyze clusters for local signal maxima and split if more than one found

Signal hill & valley analysis in 3-d

Final “energy blob” can contain low signal cells

Cells survive due to significant neighbouring signal

Cells inside blob can have negative signals

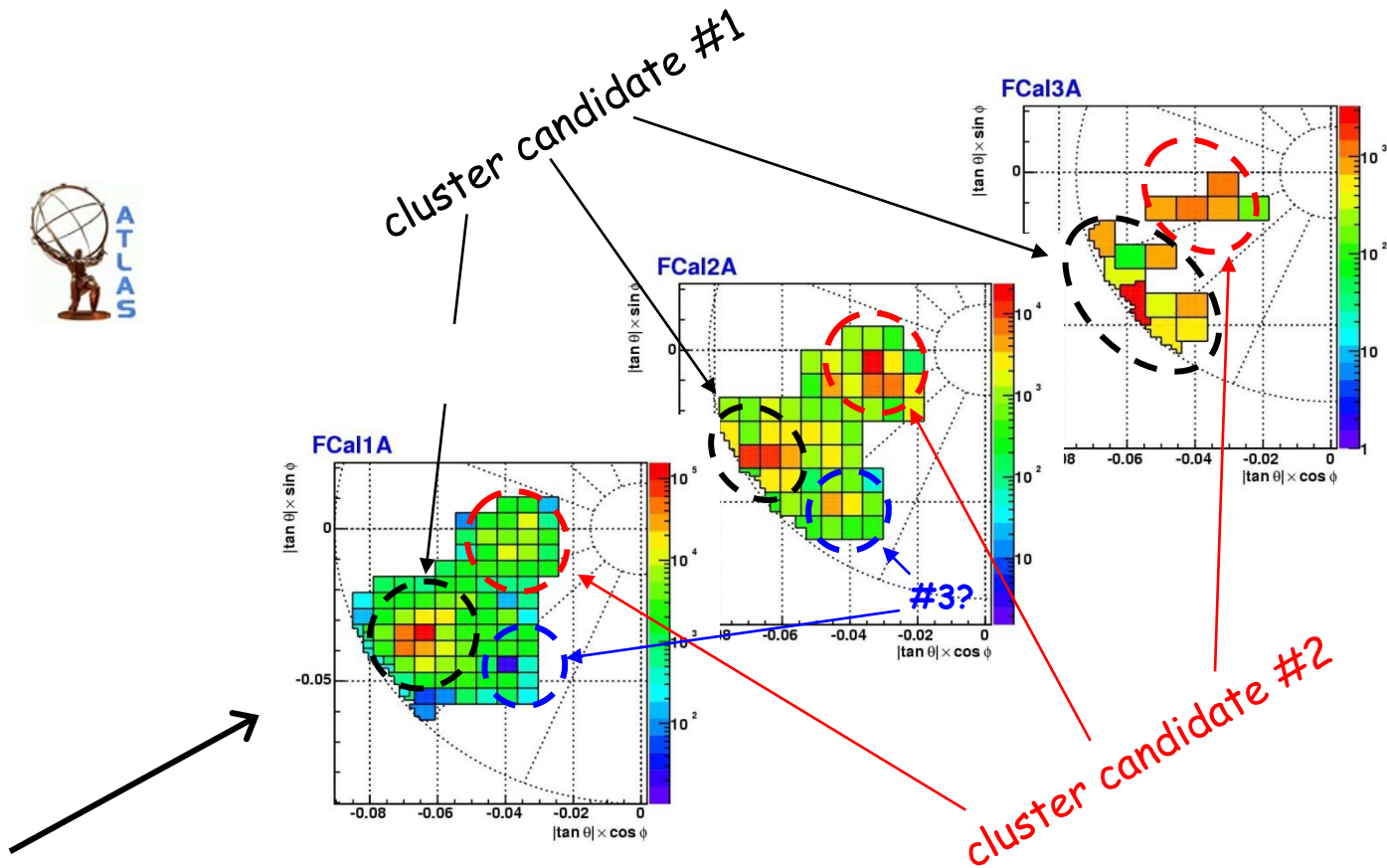
ATLAS also studies “TopoTowers”

Use topological clustering as noise suppression tool only

Distribute only energy of clustered cells onto tower grid

Motivated by DZero approach





Local hadronic energy scale restoration depends on origin of calorimeter signal

Attempt to classify energy deposit as electromagnetic or hadronic from the cluster signal and shape

Allows to apply specific corrections and calibrations

Local calibration approach

Use topological cell clusters as signal base for a hadronic energy scale

Recall cell signals need context for hadronic calibration

Basic concept is to reconstruct the locally deposited energy from the cluster signal first

This is not the particle energy

Additional corrections for energy losses with some correlation to the cluster signals and shapes extend the local scope

True signal loss due to the noise suppression in the cluster algorithm (still local)

Dead material losses in front of, or between sensitive calorimeter volumes (larger scope than local deposit)

After all corrections, the reconstructed energy is on average the isolated particle energy

E.g., in a testbeam

But not the jet energy



Quality of True Deposited Energy Estimate
 
