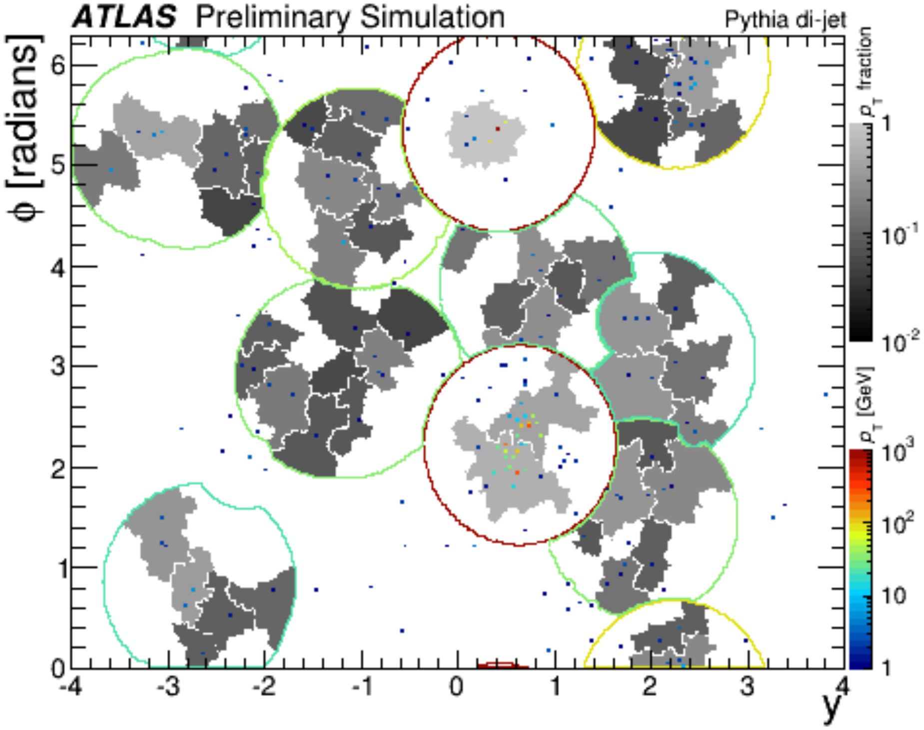
Summary of Jet and E_T^{miss} Reconstruction in ATLAS Analyses



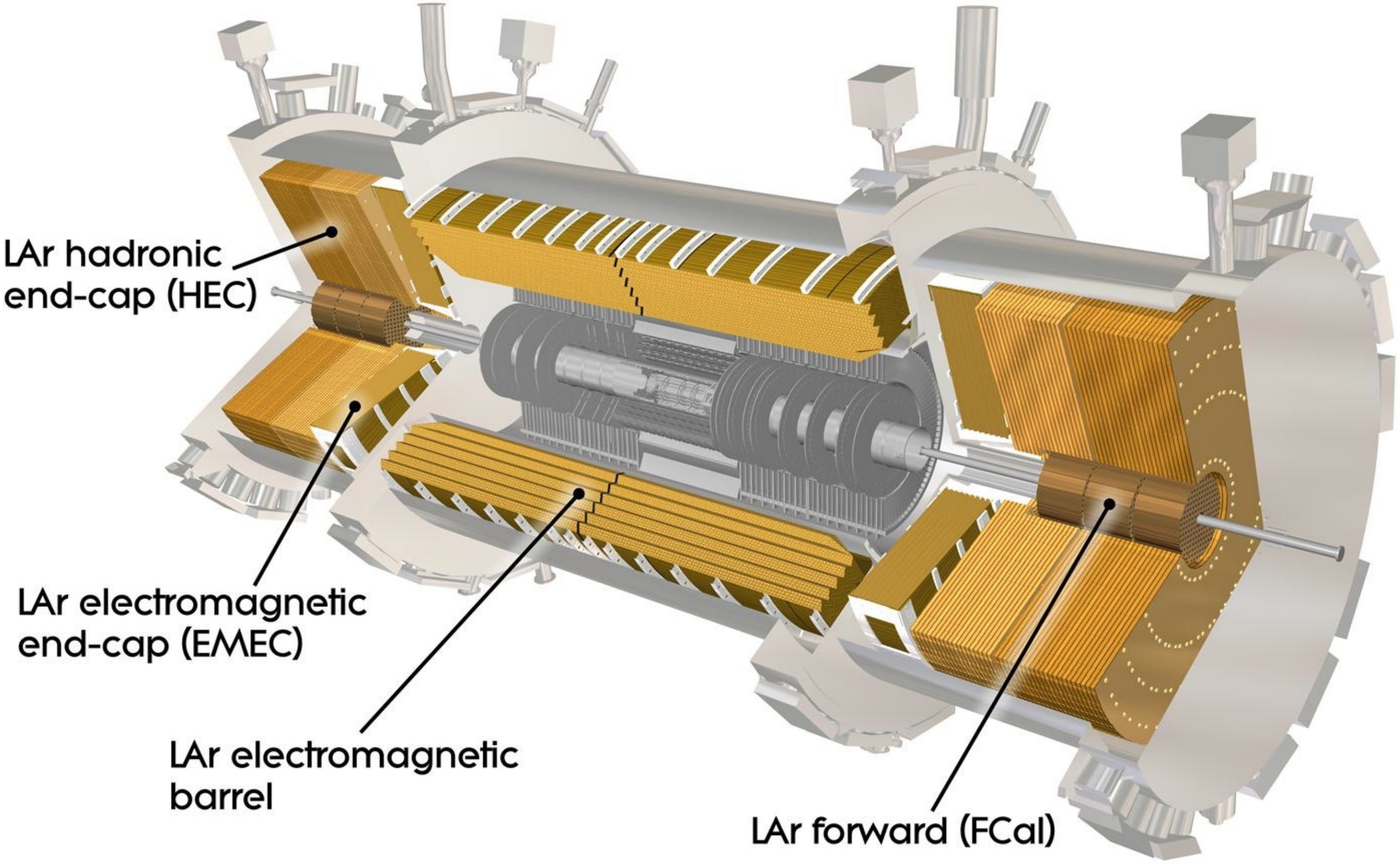
David López Mateos (Harvard University), August 25th, 2014
2014 CMS Jet/MET Workshop

Outline

- Calorimetry and Inputs
- Jet Energy MC-based Calibration
- In-situ techniques and Systematic Uncertainties on the JES
- Beyond the JES: Jet energy resolution and Jet mass calibration
- Jet substructure and hadronic jet tagging
- Missing transverse energy
- Summary and Conclusions

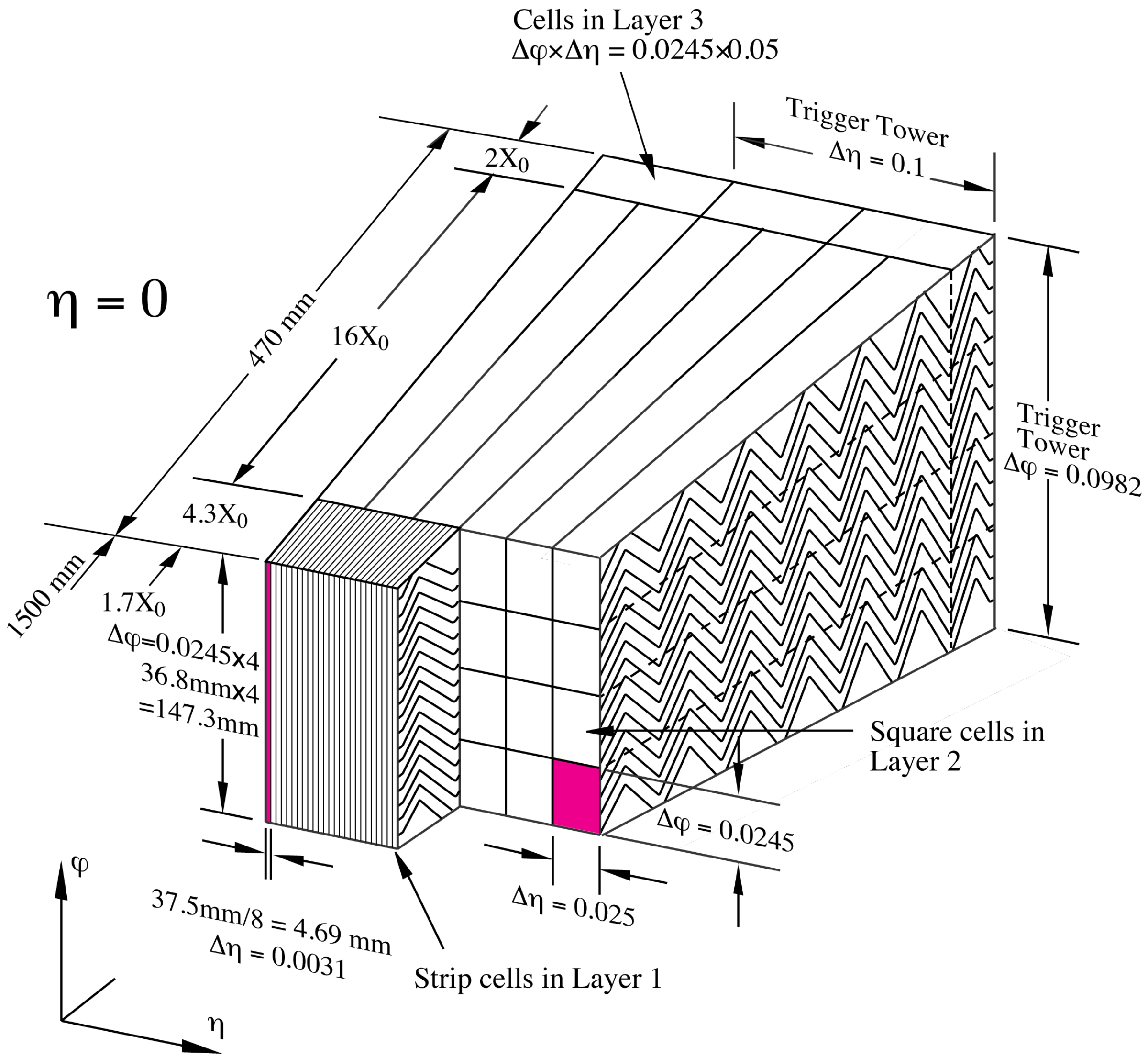
⇒ All the results shown here and more in our public twiki

The ATLAS LAr Calorimetry



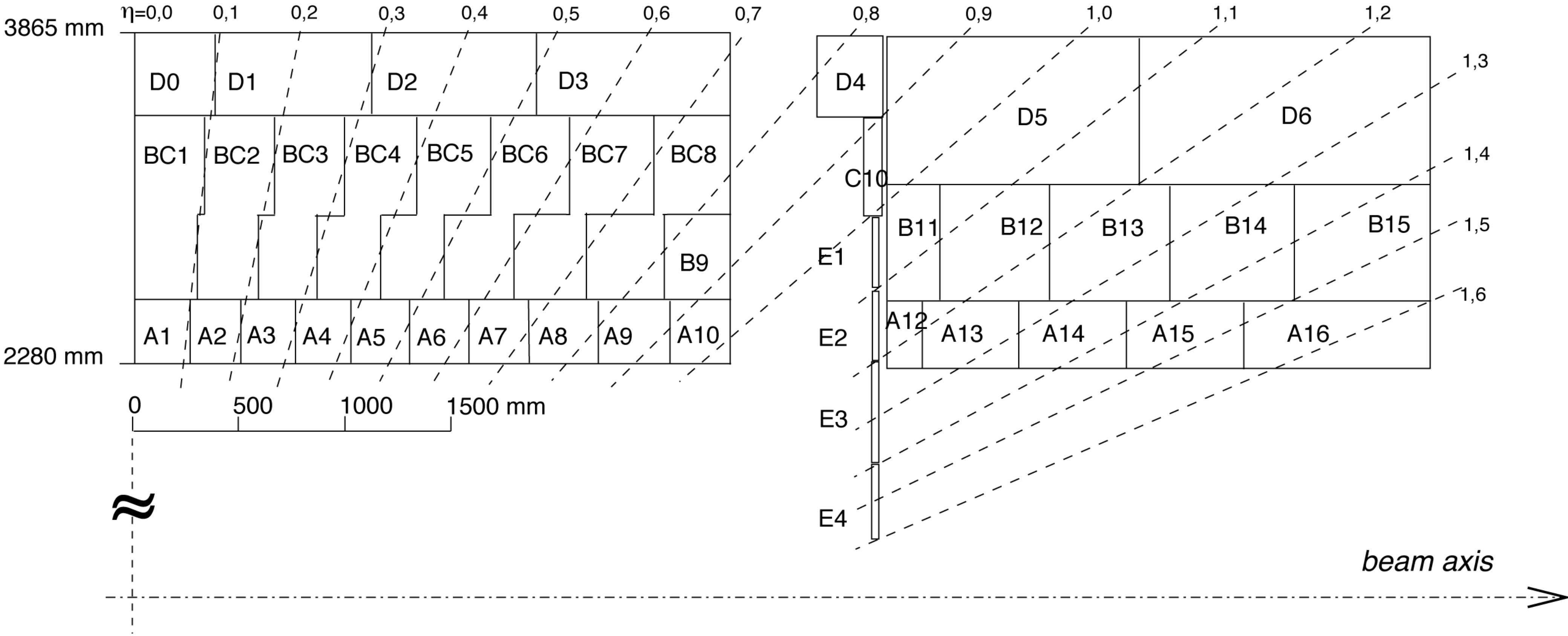
- EM and hadronic end-cap calorimetry use LAr as active medium

The ATLAS EM Calorimetry



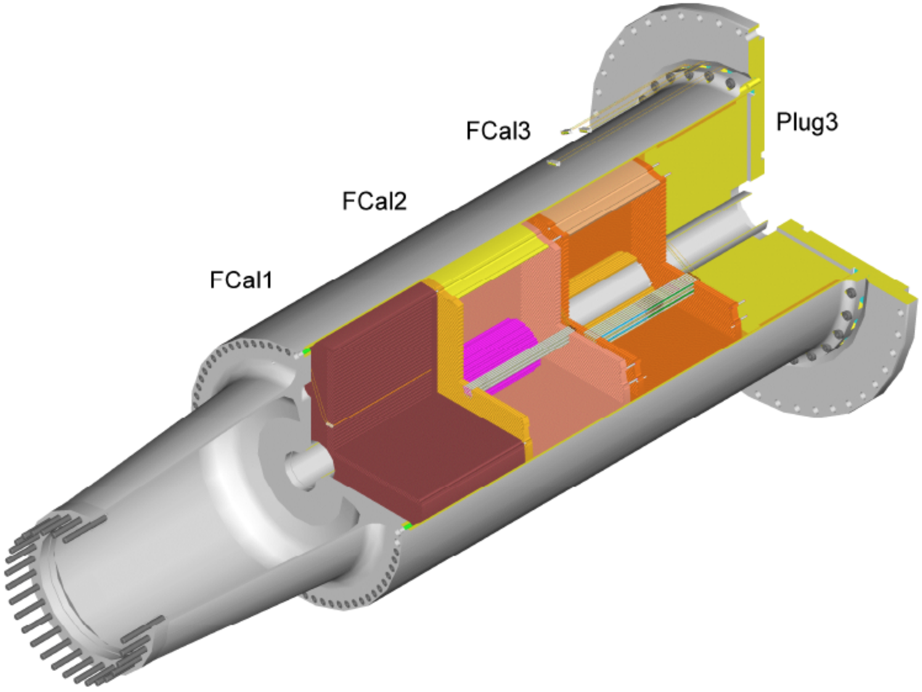
- Very fine read-out segmentation in η in first layer (motivated by H→$\gamma\gamma$)
- Longitudinal segmentation helps following development of the shower
- Additional pre-sampler layer acts as “active medium” for material in front of calo

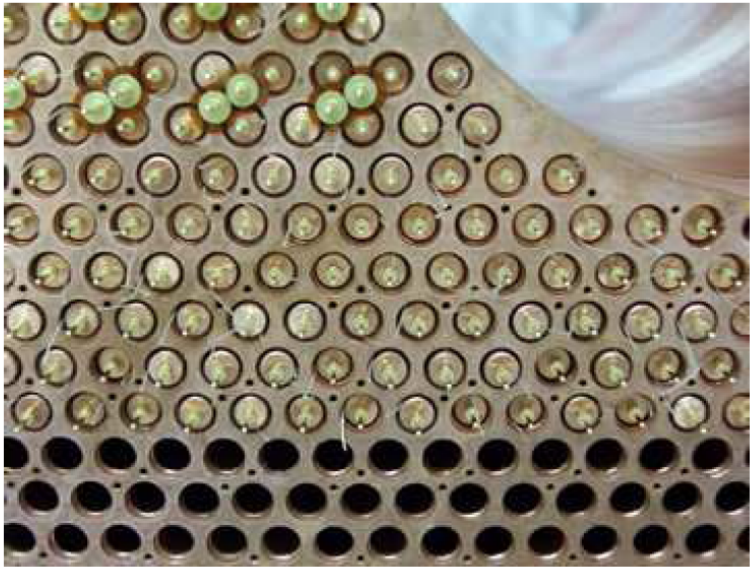
The ATLAS Tile Calorimeter

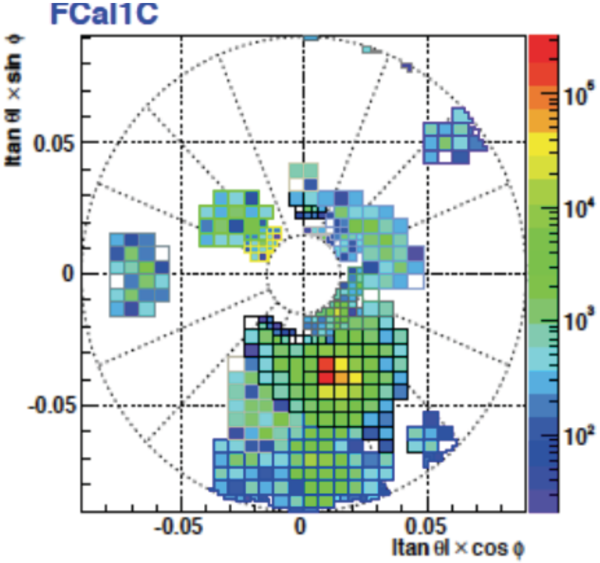


- For precision hadronic barrel calorimetry use scintillating tiles and steel
- Longitudinal segmentation helps follow shower development
- Large fraction of barrel services leave detector at η~0.9

The Forward Calorimeter

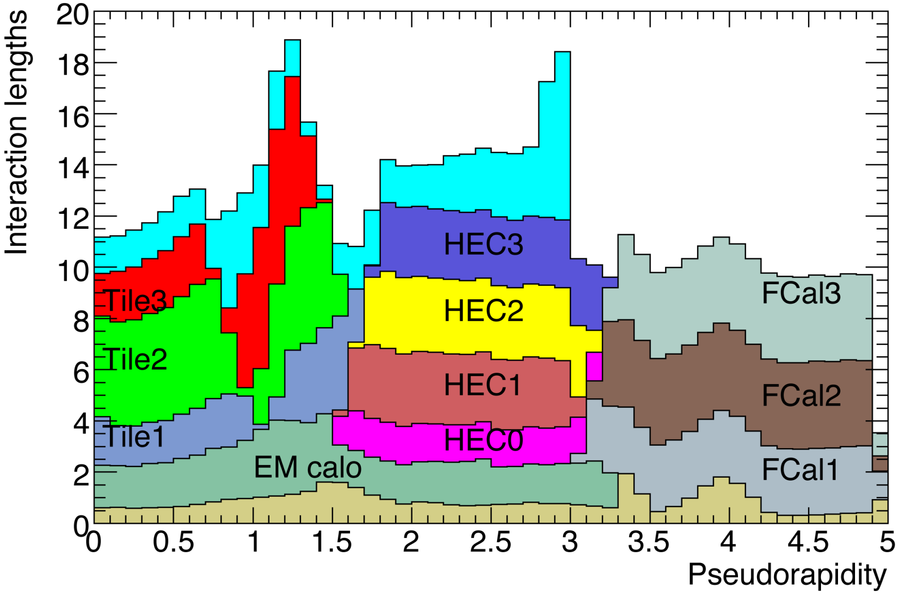


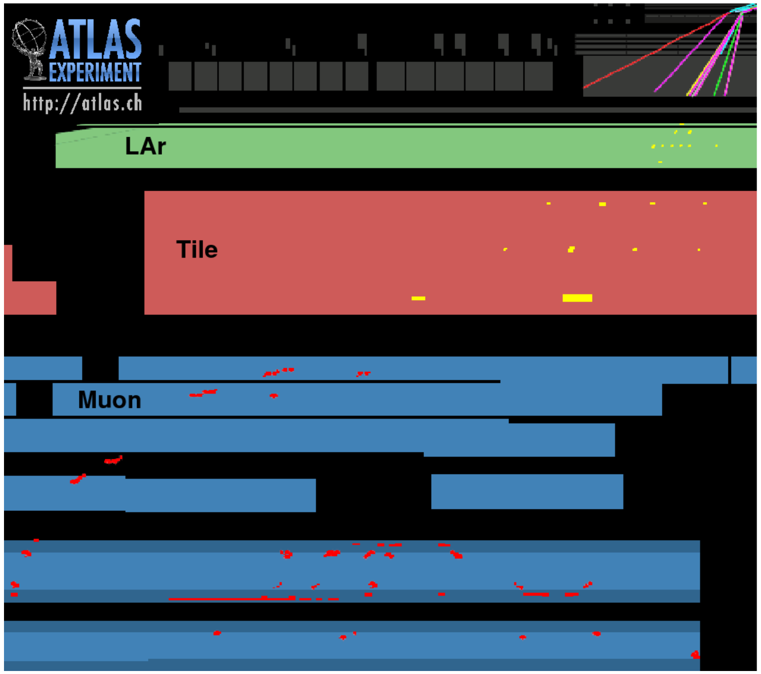




- Projectivity is somewhat lost
- Relatively small calorimeter: use Tungsten to contain showers
- Cluster size is relatively large: harder to use for substructure

Material Budget and Jet Containment

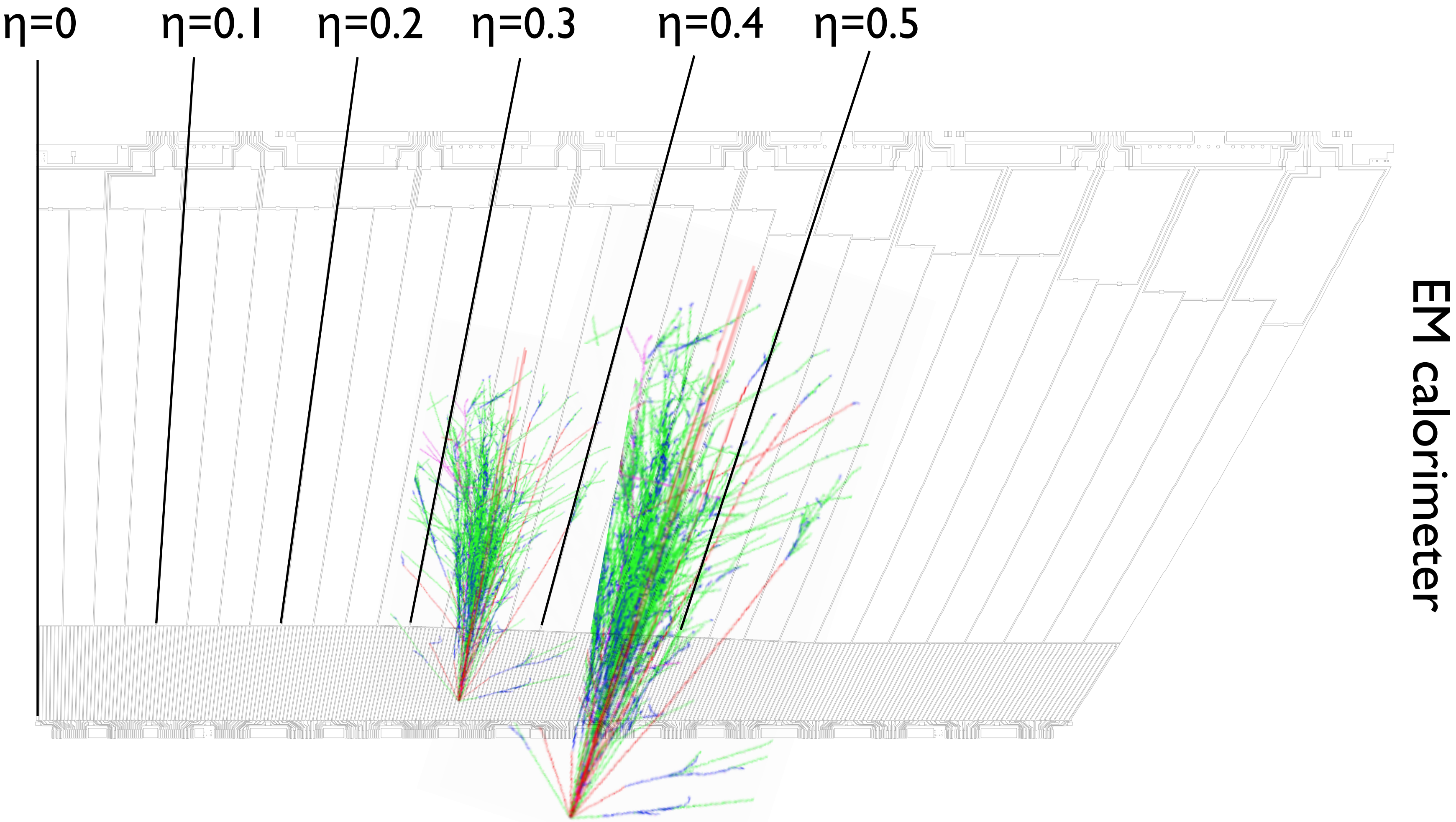




- Material budget in hadronic calorimeter is quite high, to contain guarantee containment of the shower

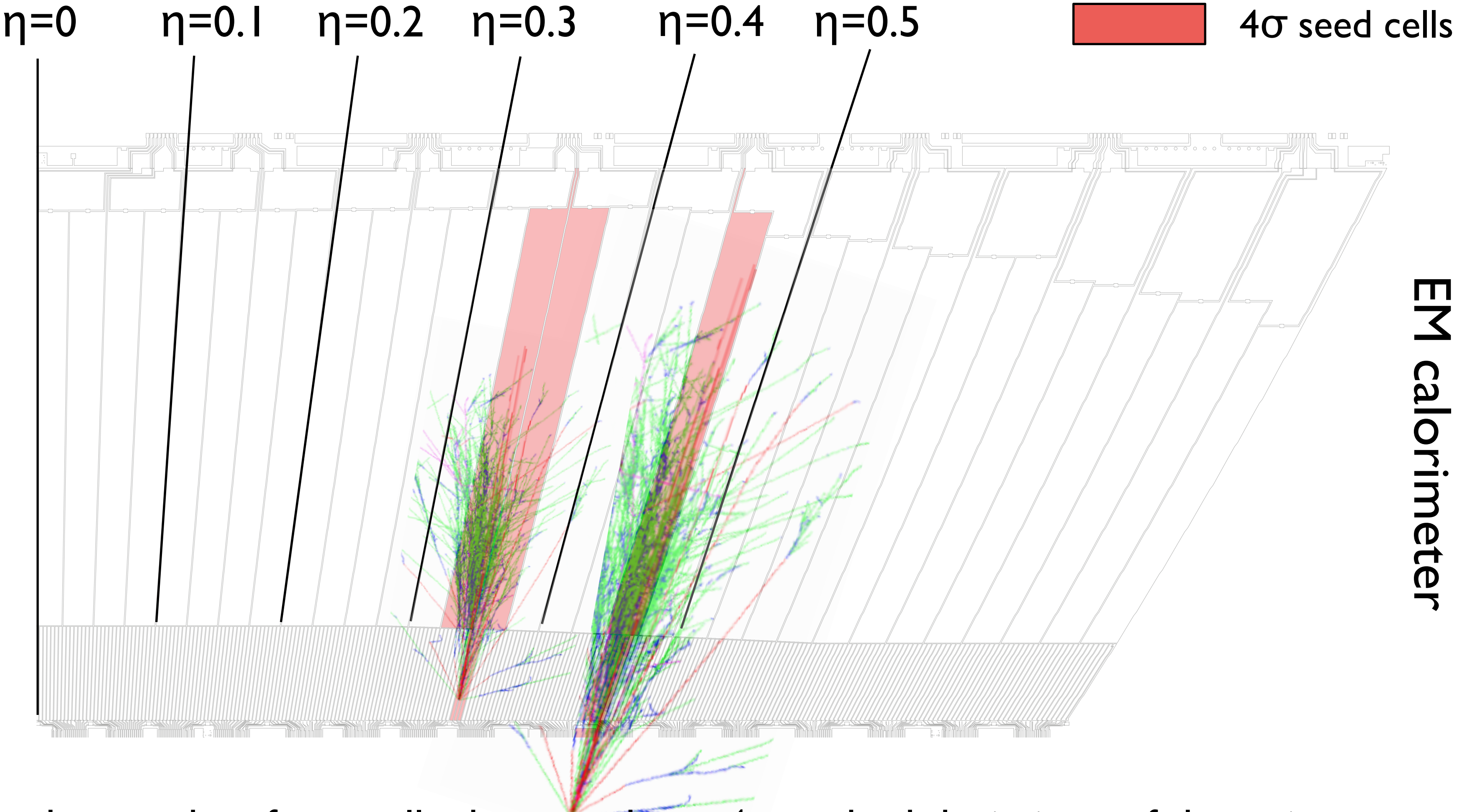
- Tails can be relevant at p_T~1 TeV, can be corrected by looking at activity in the muon spectrometer

Cluster Formation



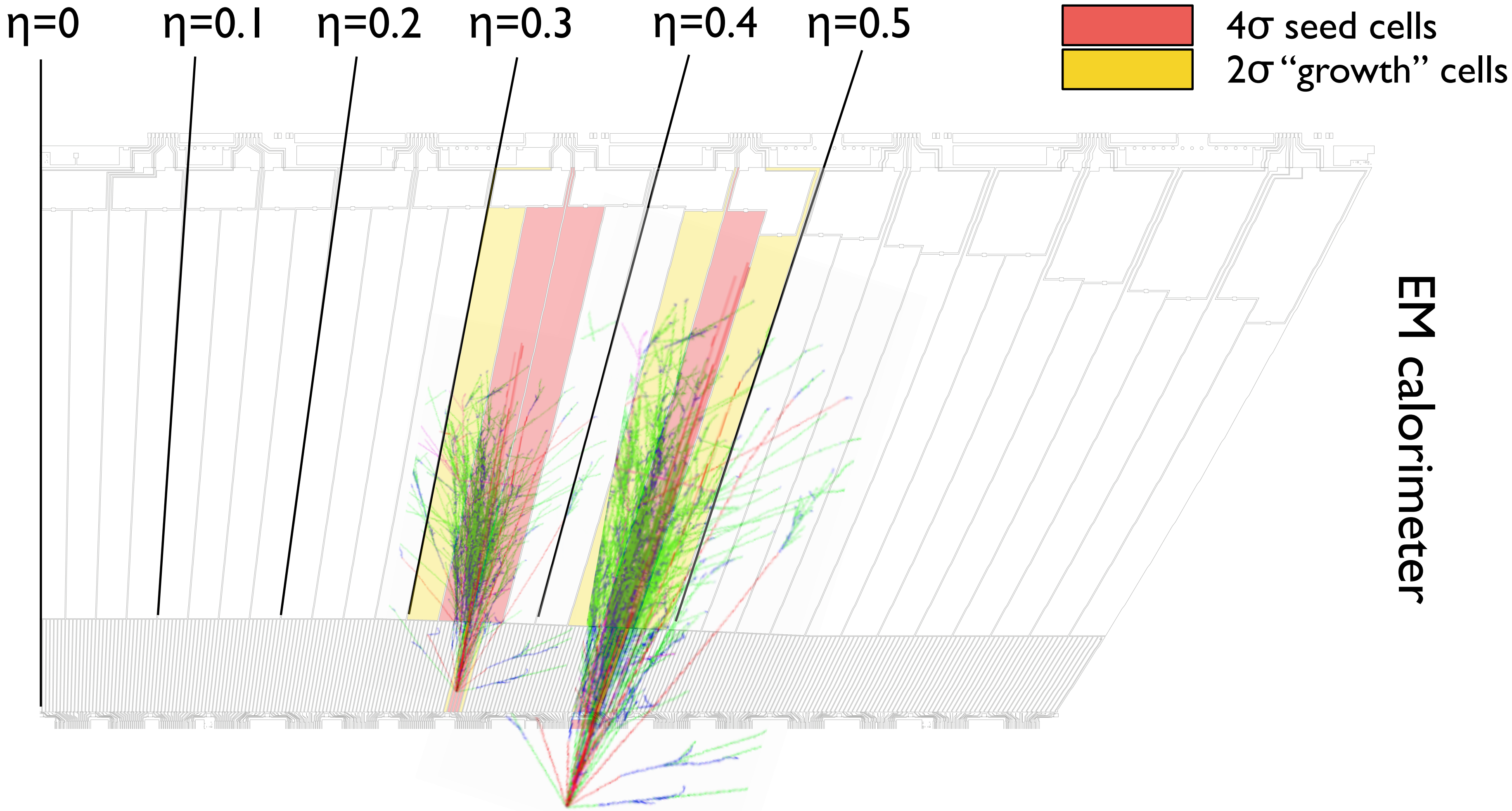
- Clusters are built starting from the fine readout granularity of the ATLAS calorimeter (above the EM calorimeter in the central region)

Cluster Formation



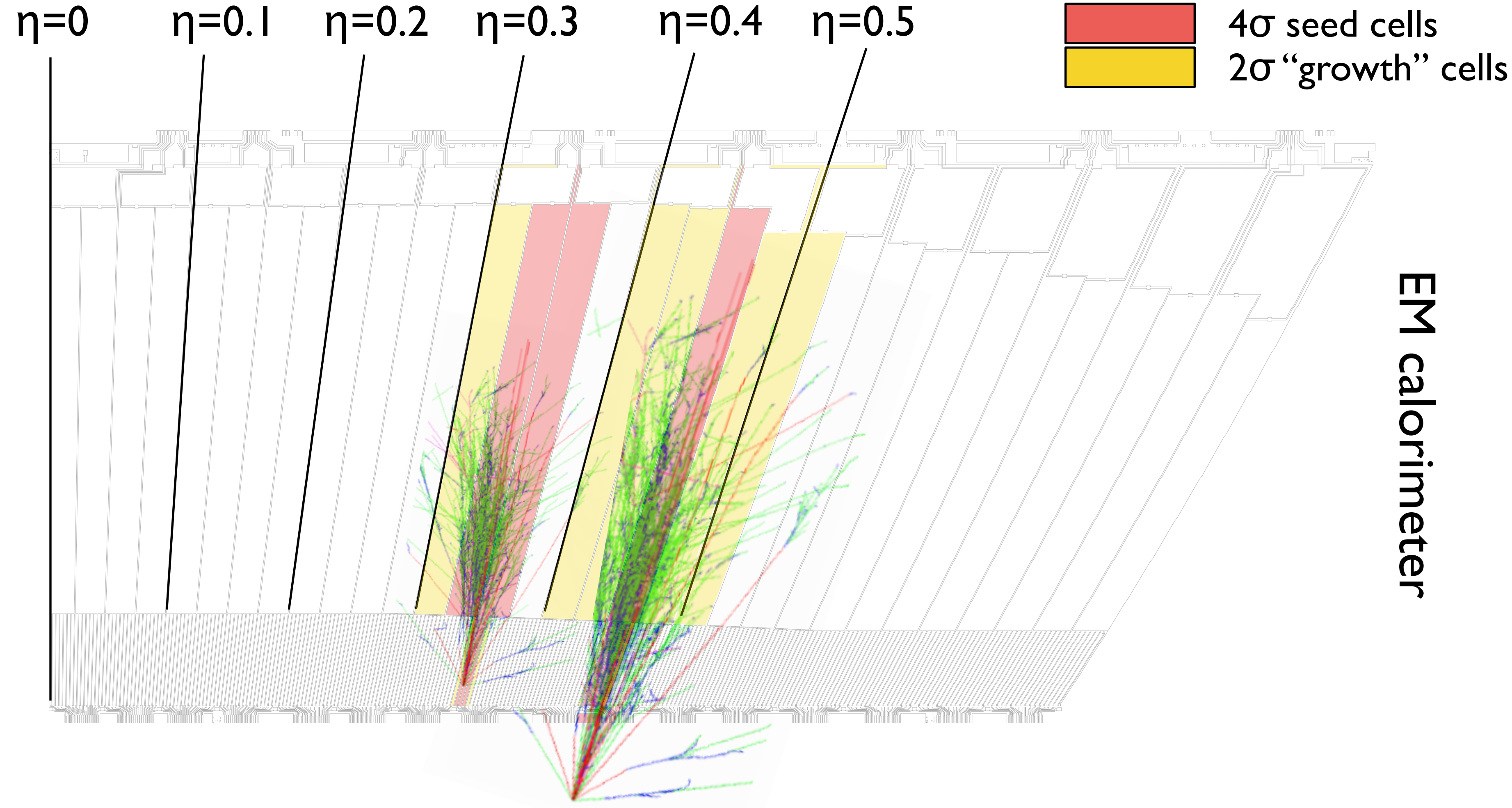
- Seeds are taken from cells that are above 4 standard deviations of the noise
- Noise includes electronic noise and average energy readings from pile-up
- Each cell has its value of noise stored in a database and that value is validated in data

Cluster Formation



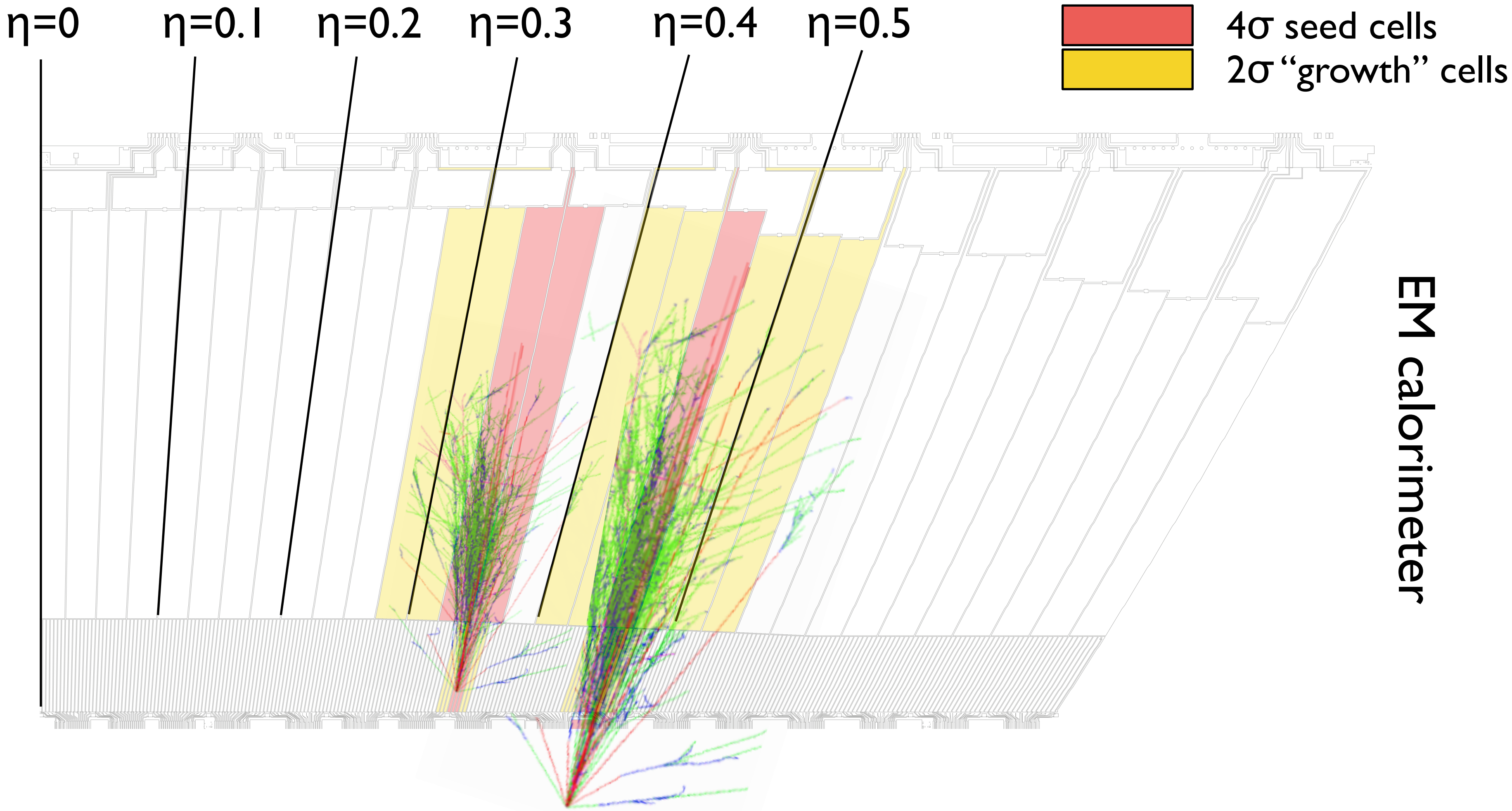
- Cluster grows (in 3 dimensions) into adjacent cells where a deposition >2σ is found

Cluster Formation



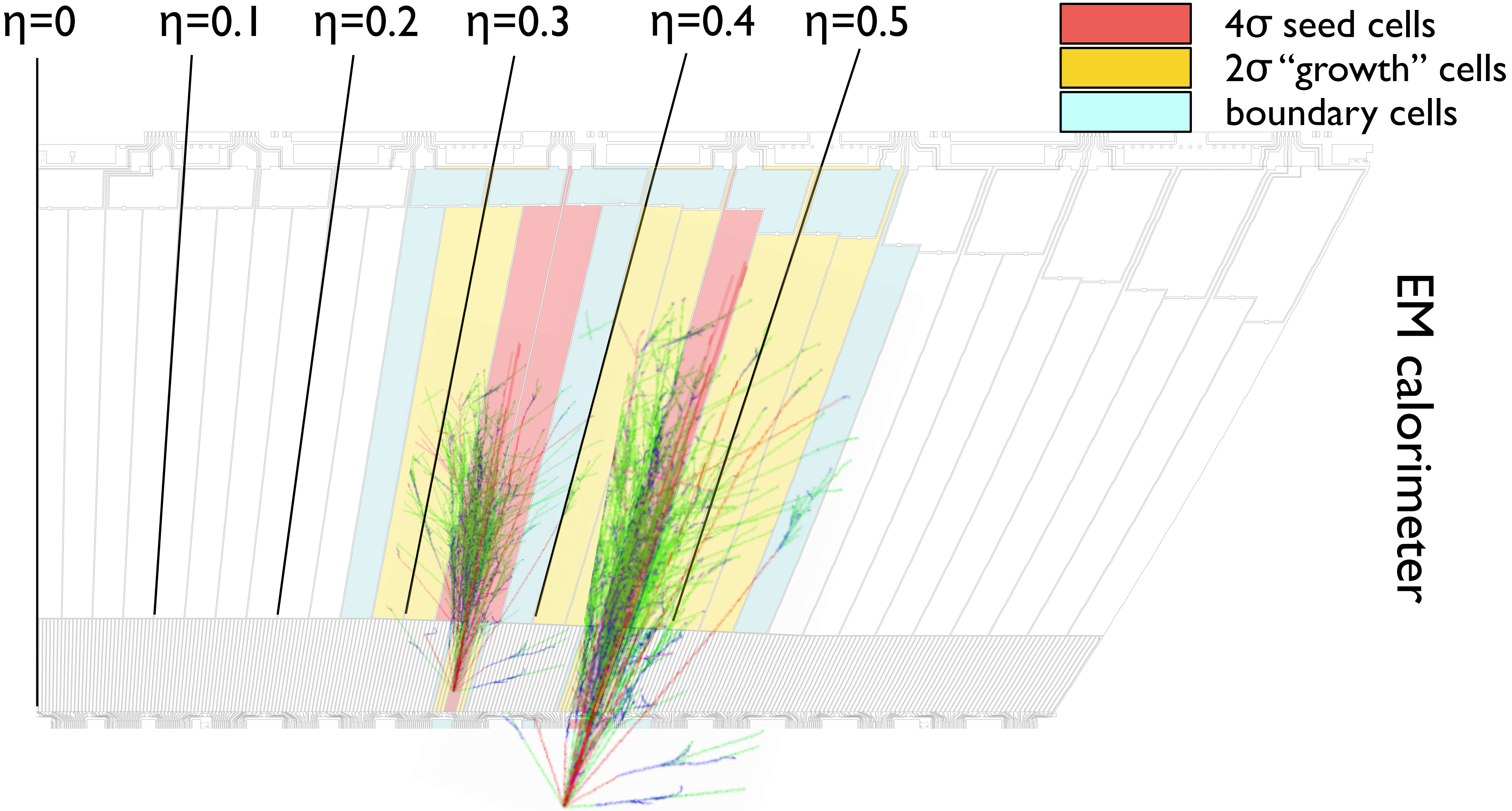
- Cluster grows (in 3 dimensions) into adjacent cells where a deposition >2σ is found
- Growth continues while adjacent cells with >2σ are found

Cluster Formation



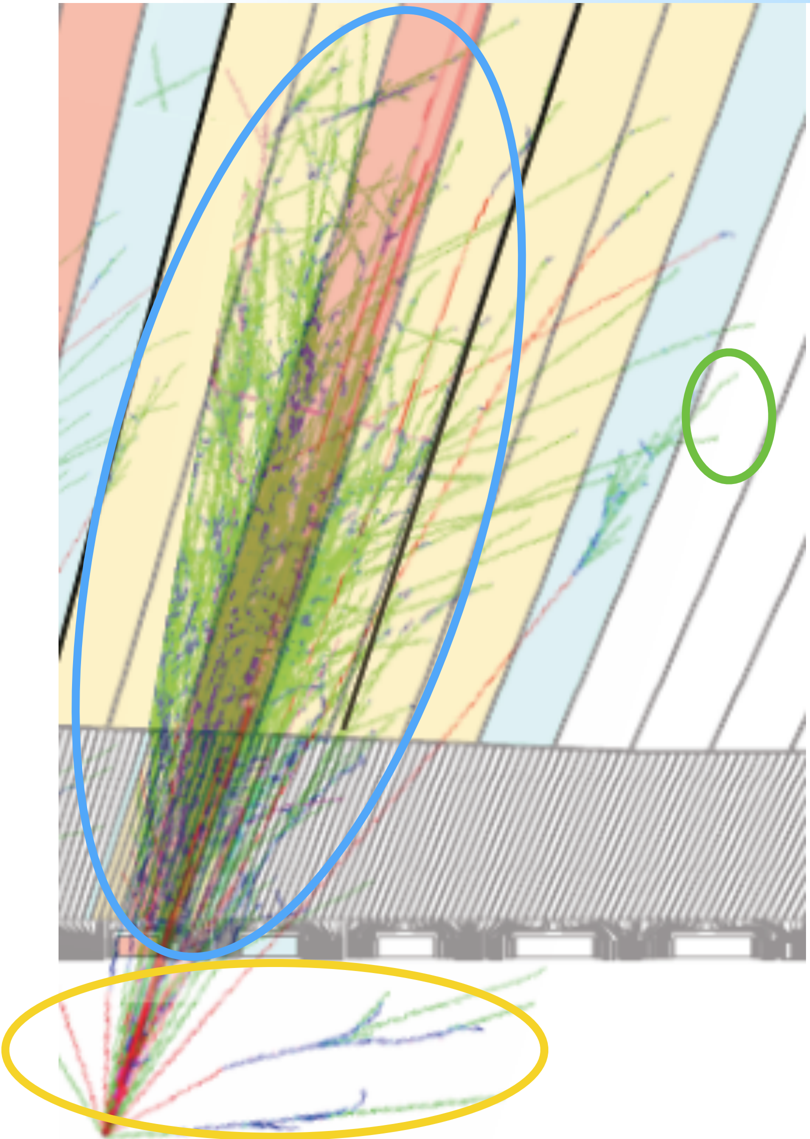
- Cluster grows (in 3 dimensions) into adjacent cells where a deposition >2σ is found
- Growth continues while adjacent cells with >2σ are found

Cluster Formation

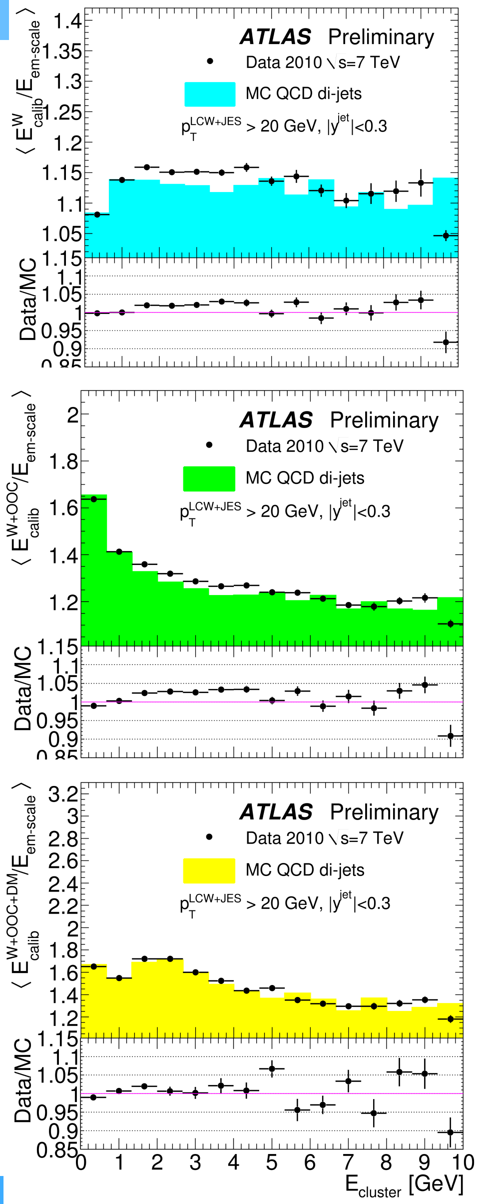


- Once growth is no longer possible, an additional set of boundary cells is added (irrespective of their energy)

Cluster Calibration



- Weights for non-compensation
  - Cluster energy
  - Cluster depth
  - Cell energy density
- Weights for energy out of the cluster
  - Cluster depth
  - Cluster isolation
- Weights for energy in dead material
  - Cluster energy
  - Energy deposited in each layer
  - Cluster depth



Calibration Sequence

Pile-up Suppression

Origin Correction

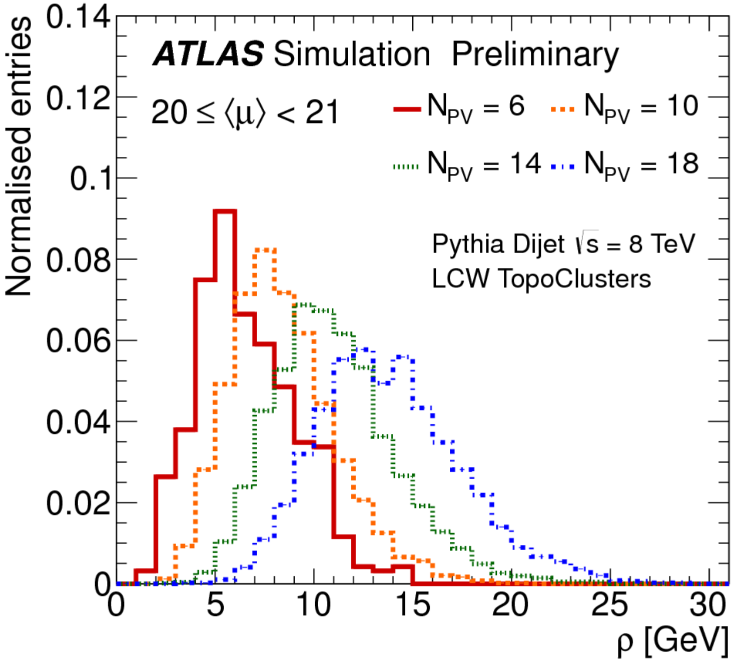
MC JES Calibration

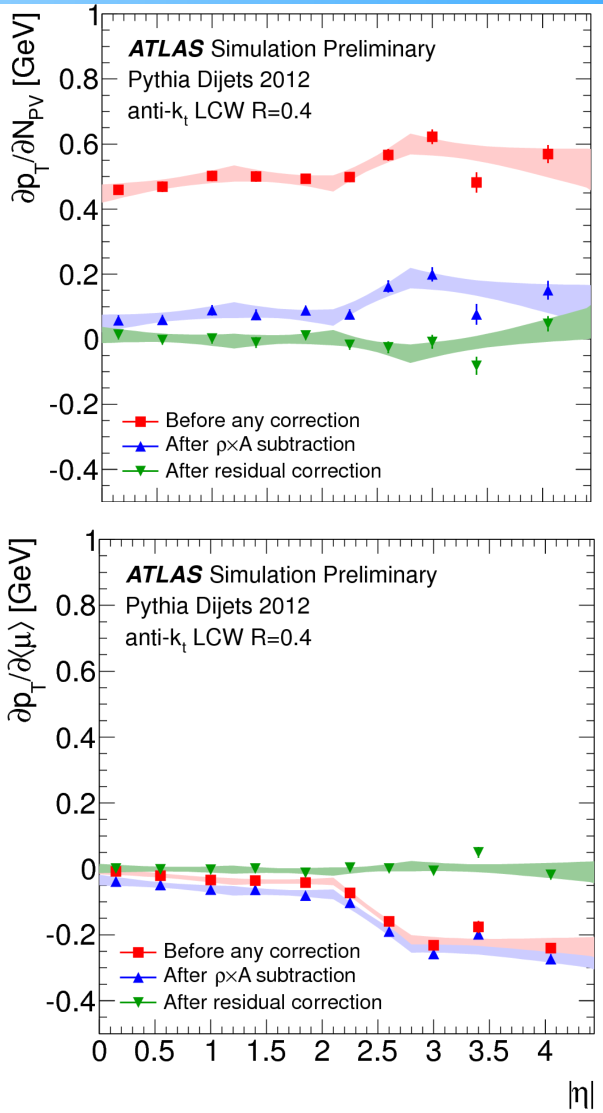
Global Sequential Calibration

In-situ Correction

- Not very different from CMS
- MC JES calibration also includes an eta calibration
- Global sequential calibration is based on properties: important to reduce flavor dependence of JES

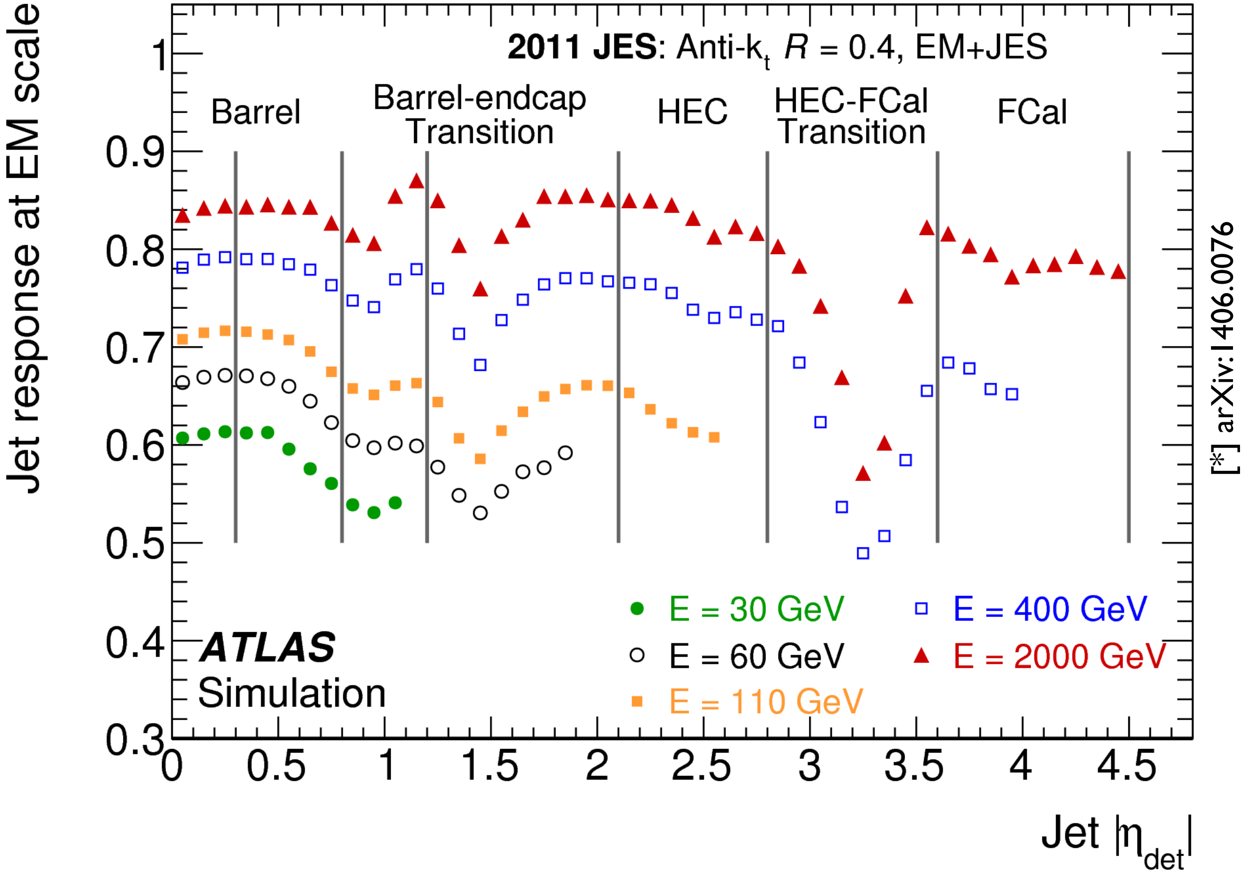
Area Subtraction and Residual Pile-up Corrections





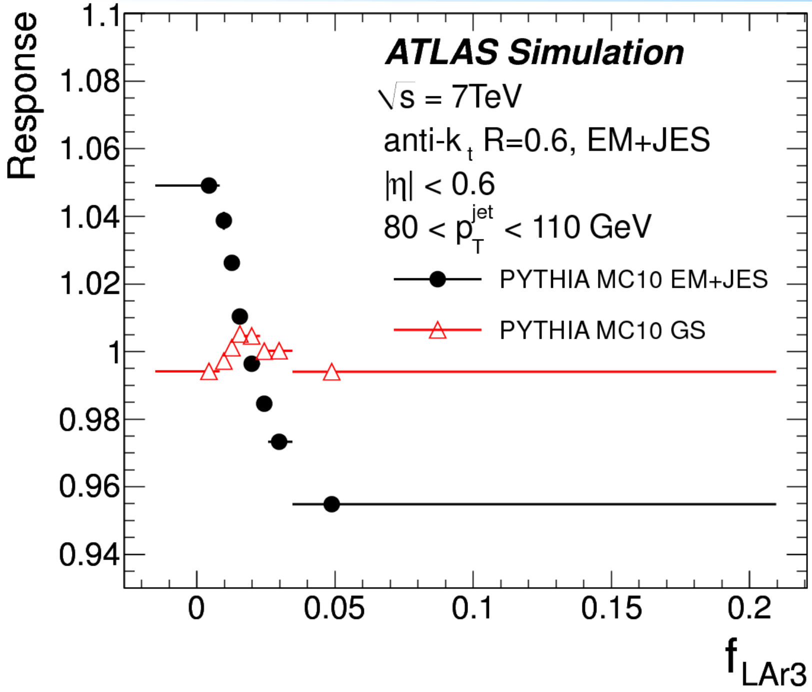
- ρ calculated in central region
- Residual correction required for out-of-time pileup

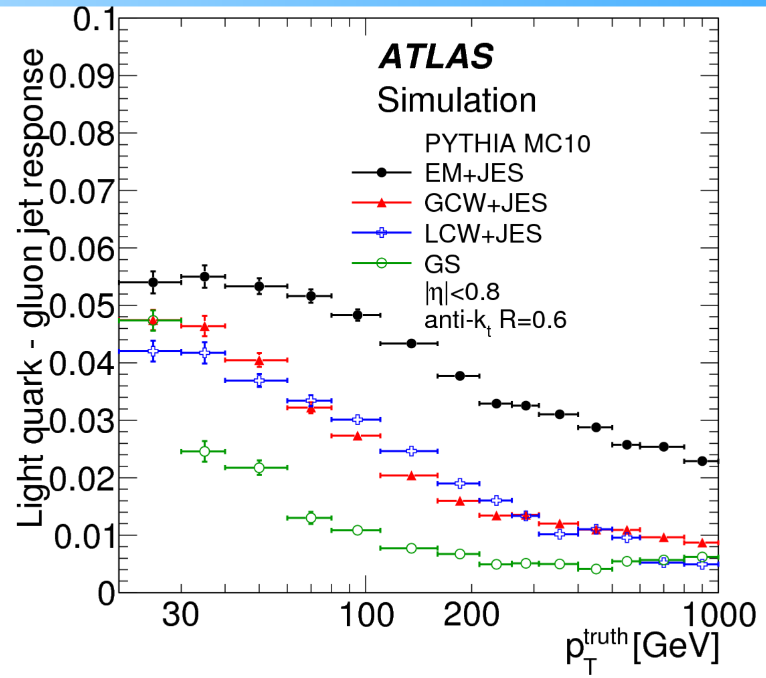
Energy and η Calibration



- Just invert the response as a function of energy (and η)

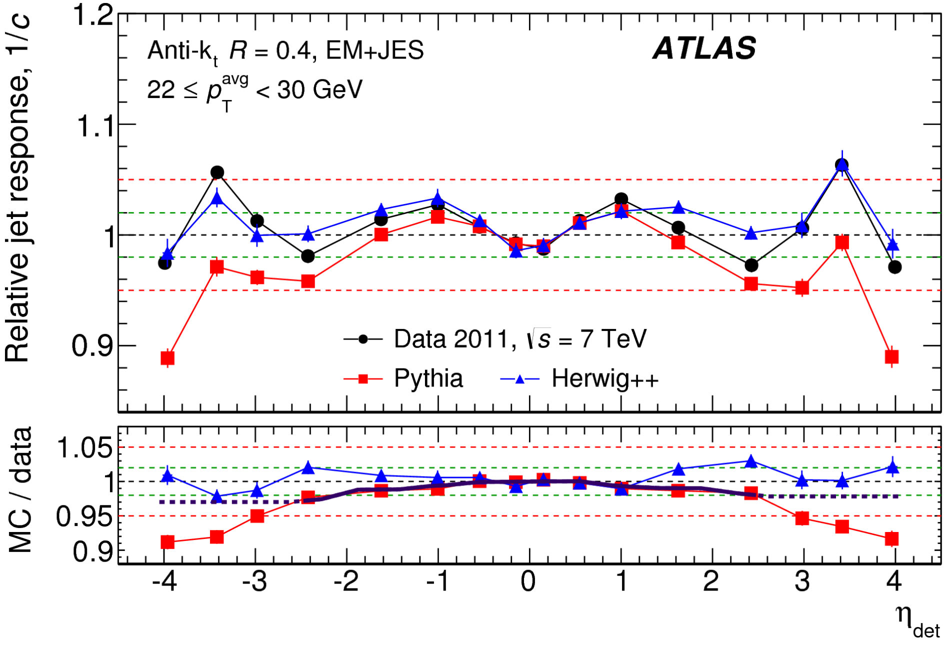
Global Sequential Corrections

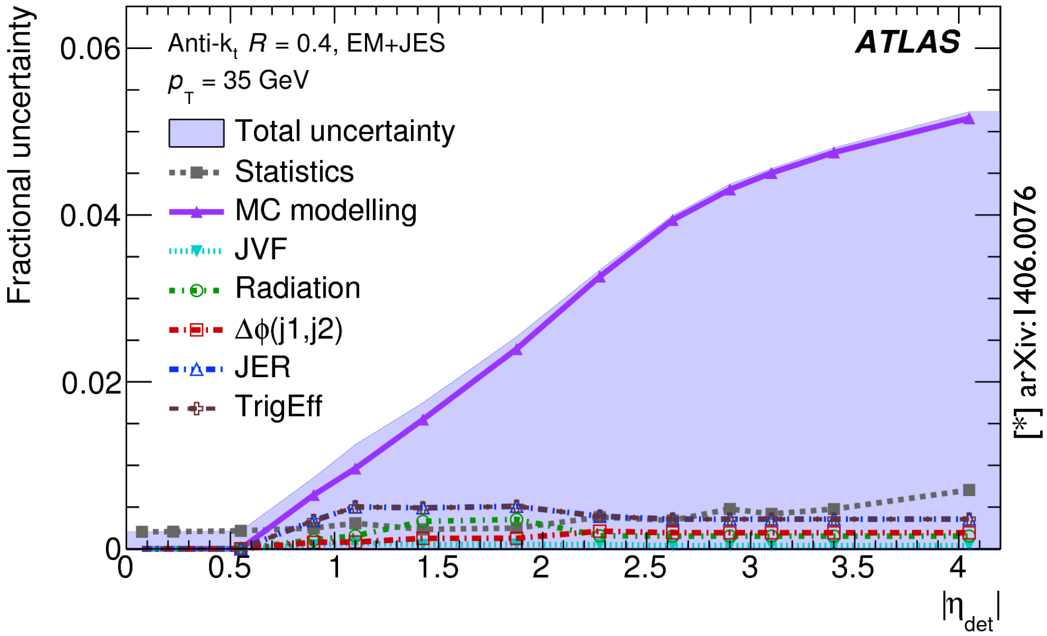




- Just invert the response as a function of energy and something else (and η)
- Currently, that something else is: n_{trk}, track width, f_{EM3}, f_{Tile1} $N_{segments}$ behind the jet
- Missing correlations don't win us much

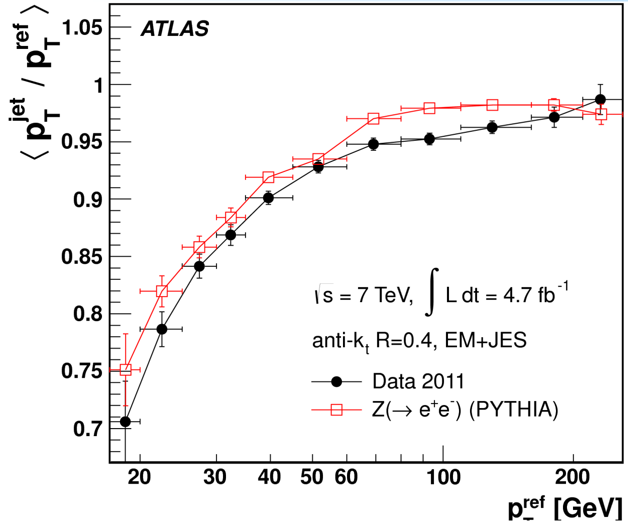
In-situ Analyses: η Intercalibration

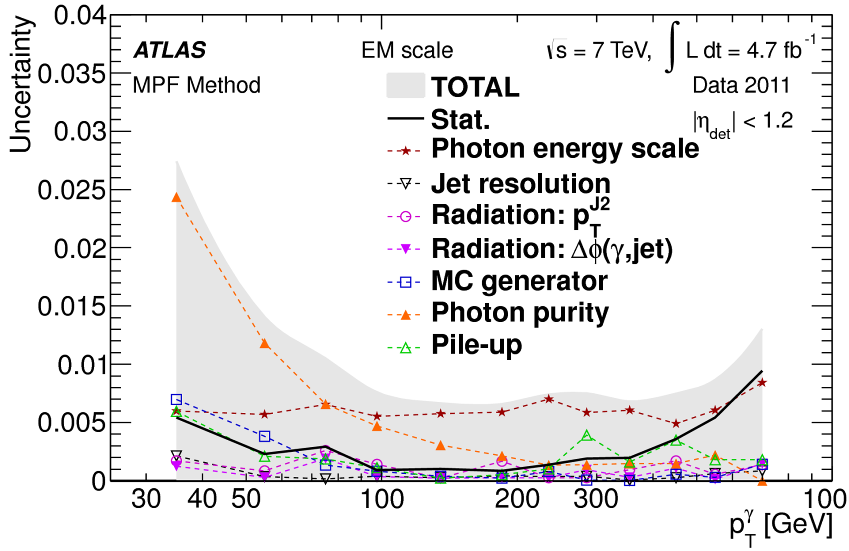




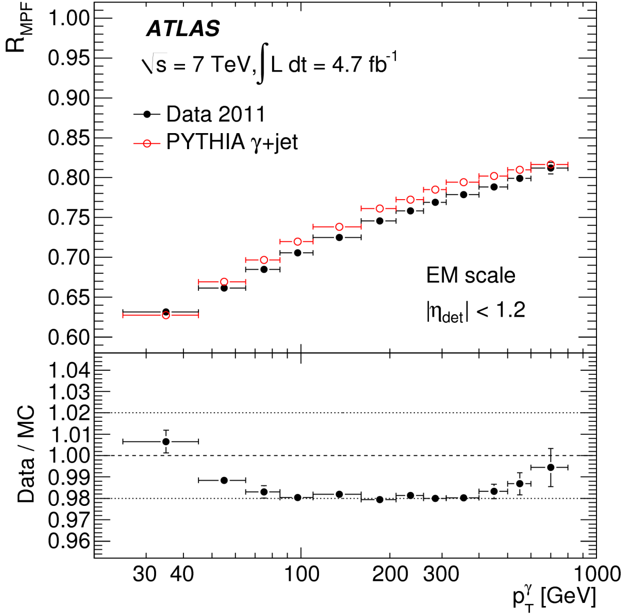
- In-situ calibration freezes out at around η~3
- Systematics mostly come from modeling of 3rd jet radiation in different MCs
- Expect less dependence (smaller systematics) with NLO MCs

In-situ Analyses: V+jet



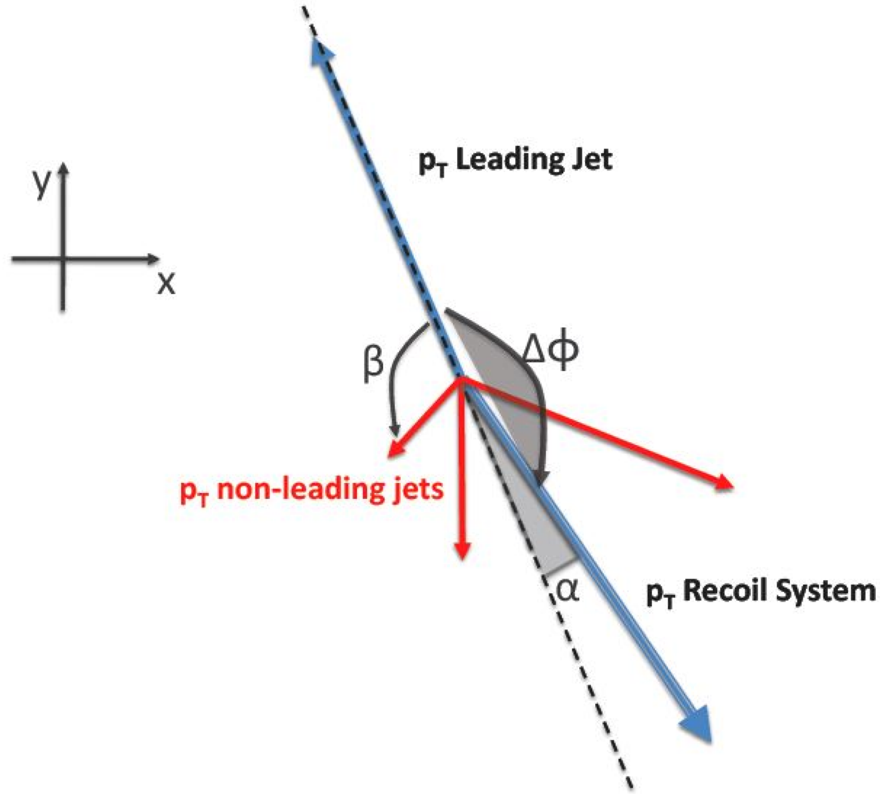


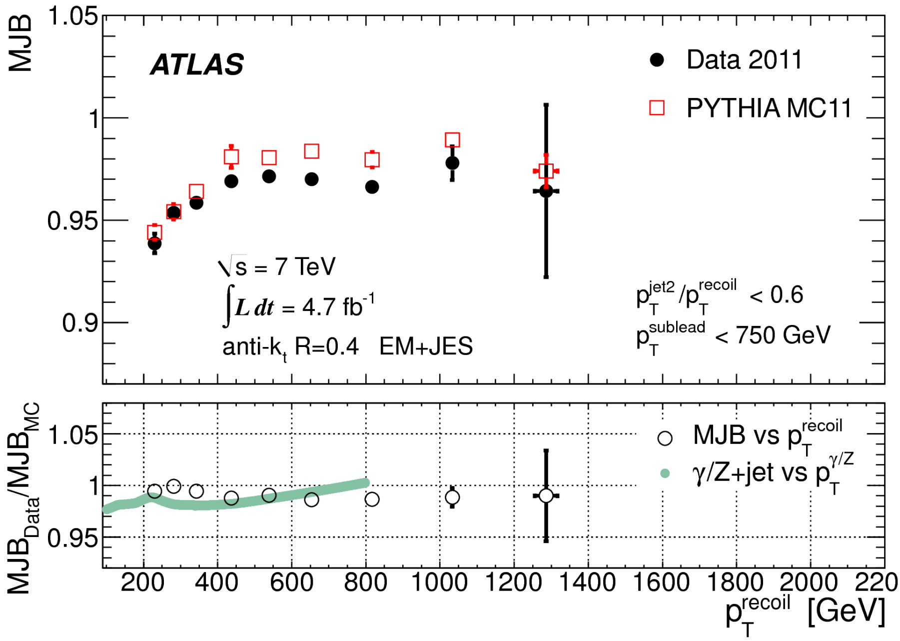
[*] arXiv:1406.0076



- Both MPF and direct balance techniques give compatible results
- Uncertainties dominated by EM scale
- Reach to about 1 TeV using 2012 data

In-situ Analyses: Multijet

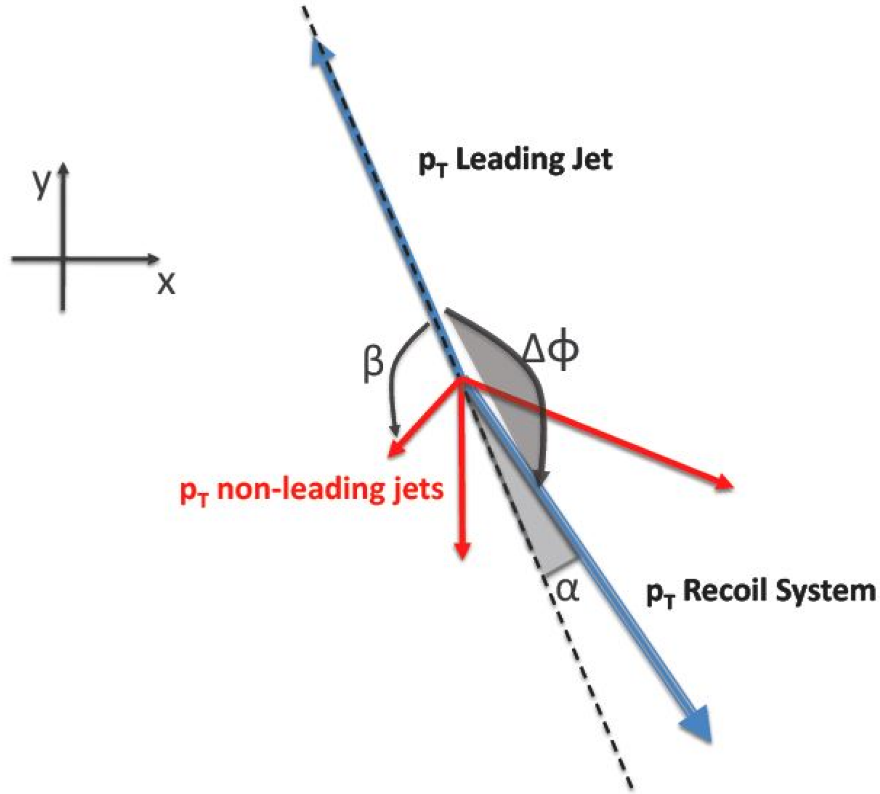


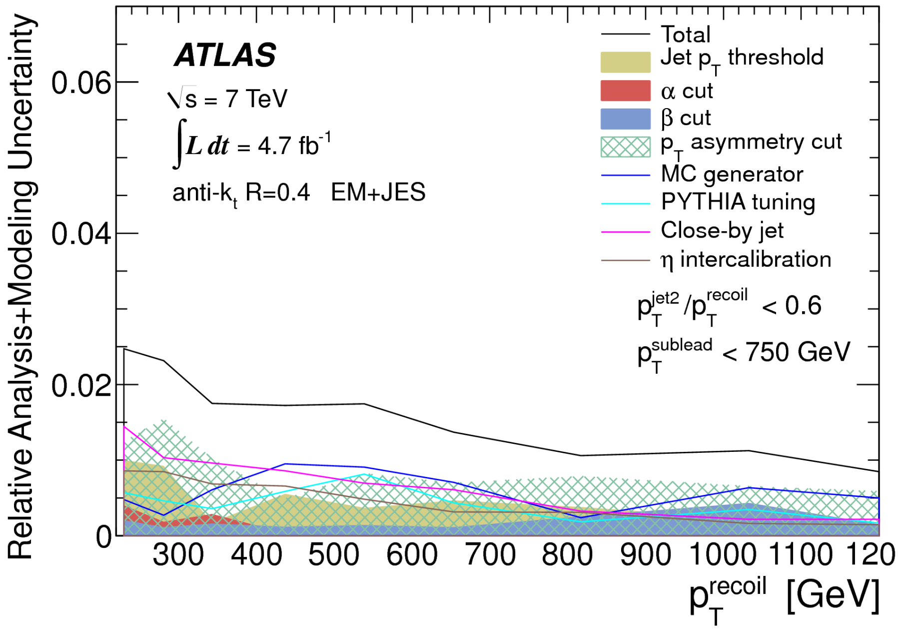


[*] arXiv:1406.0076

- Get to higher p_T (up to 1.8 TeV with full 2012 dataset) by using single jet recoiling against multiple (calibrated jets)

In-situ Analyses: Multijet

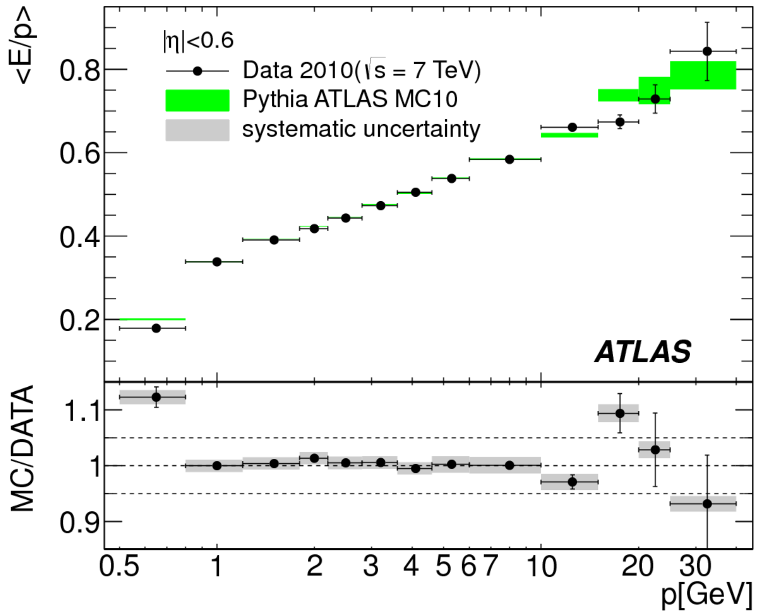




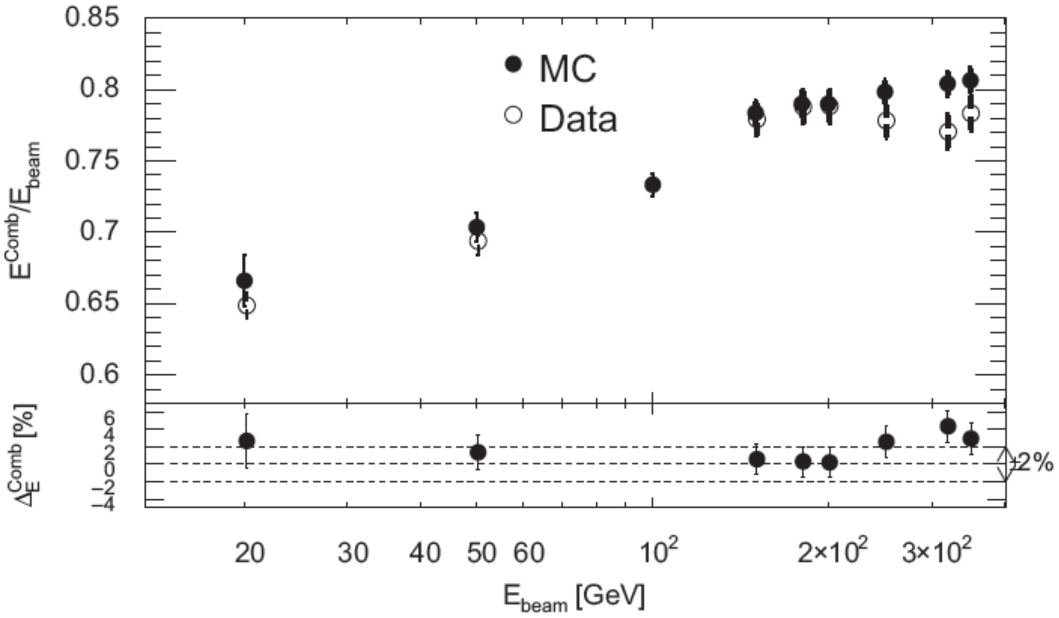
[*] arXiv:1406.0076

- Uncertainties quite comparable (~1%) to V+jet uncertainties
- Uncertainties on the topology (and the JES) of the recoil are most important

In-situ Analyses: Single-particle



[*] Eur. Phys. Jour. C, 73 3 (2013) 2305

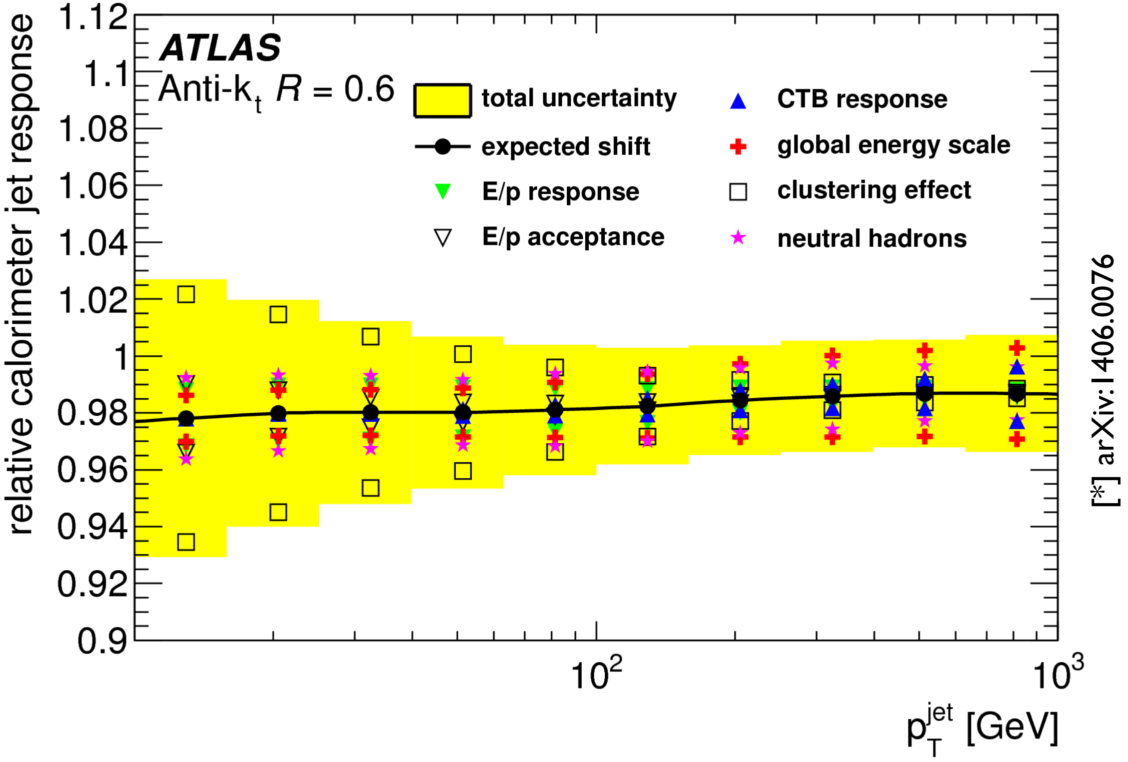


[*] Nucl. Instrum. Meth. A621 (2010) 134-150

- Large effort to extrapolate to higher p_Ts using data

- Most of it coming from the test-beam, but also from isolated hadron (pions and protons/antiprotons) data

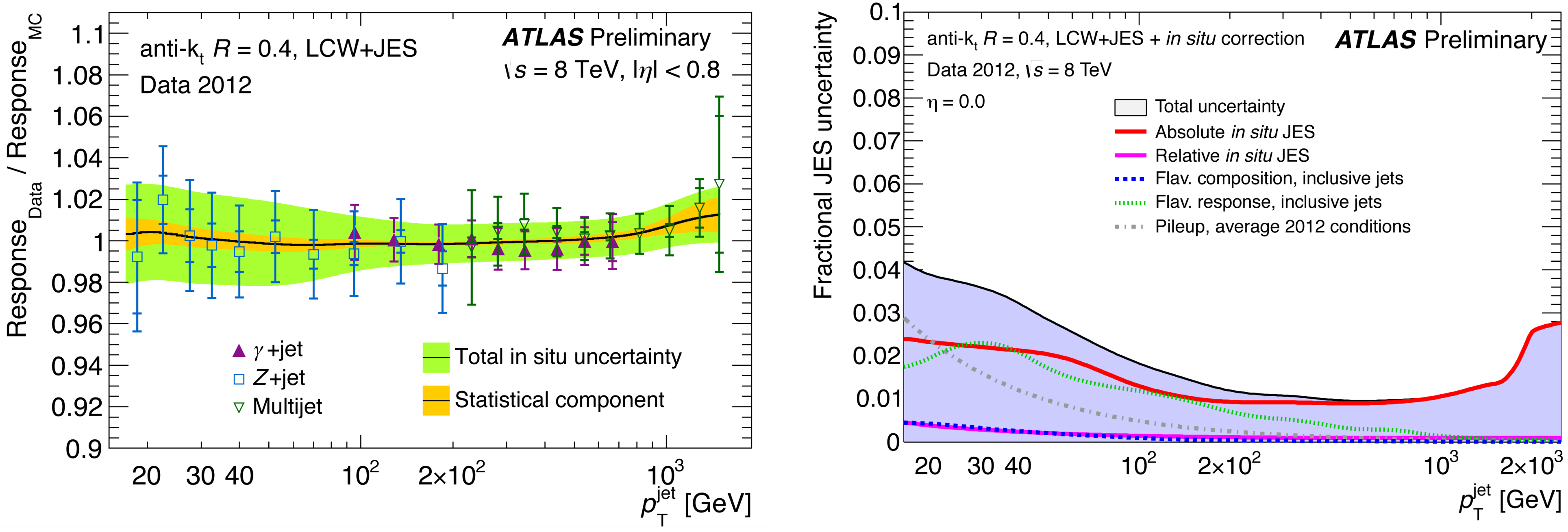
In-situ Analyses: Single-particle



- Single-particle results allow propagating shifts and uncertainties to JES

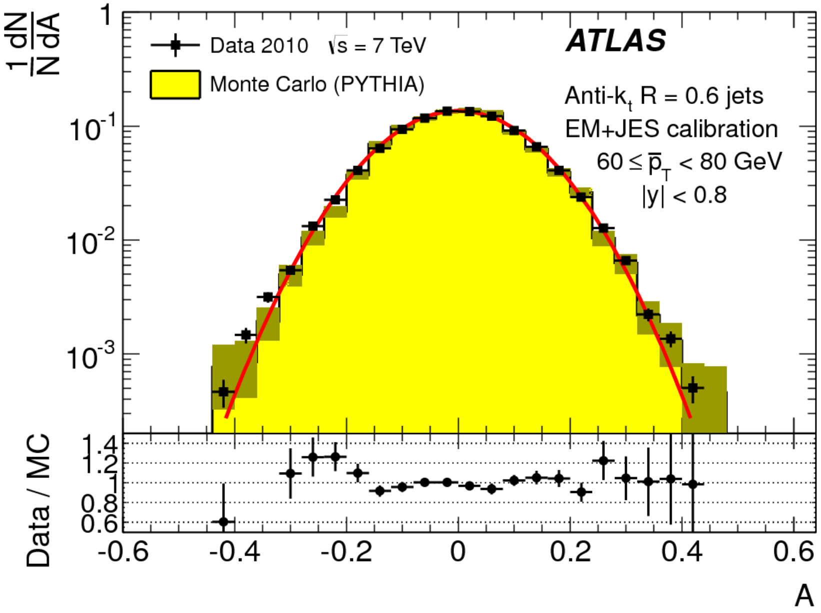
- Shift predicted by single particle analysis compatible with shift observed in data within <1%!

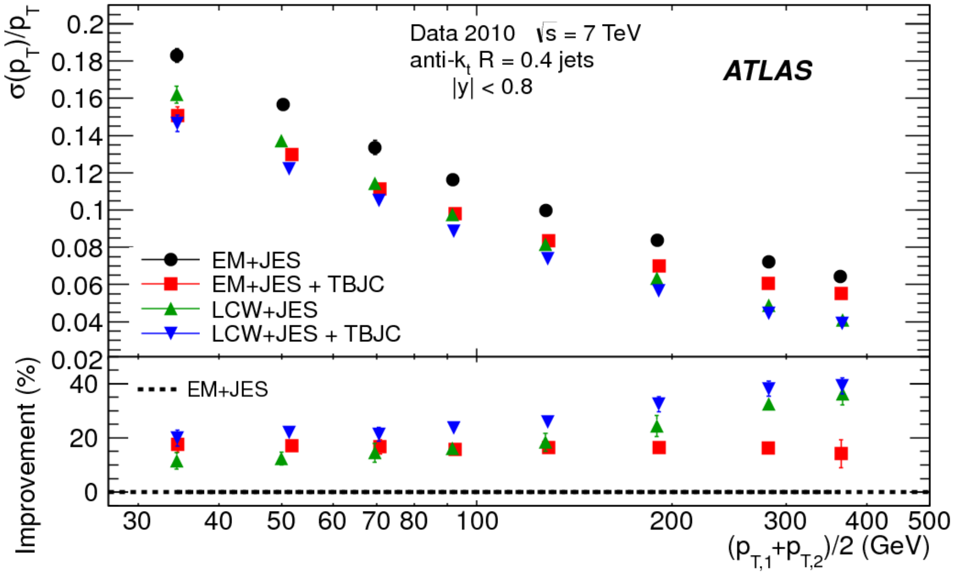
In-situ Combination



- In-situ calibration and uncertainties come from statistical combination between all methods
- Dedicated pile-up (N$_{PV}$ and μ dependent) and flavor uncertainties

Beyond the JES: Jet Energy Resolution

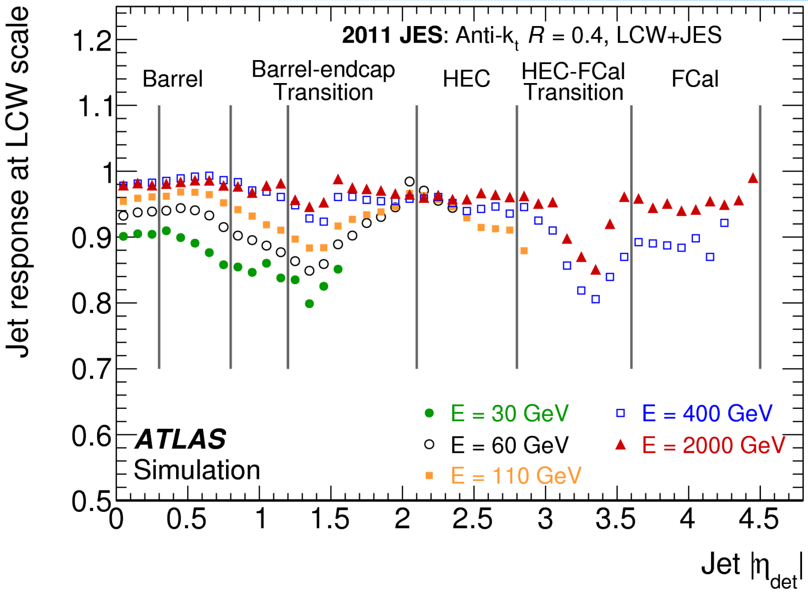




[*] EPJC 73 3 (2013) 2306

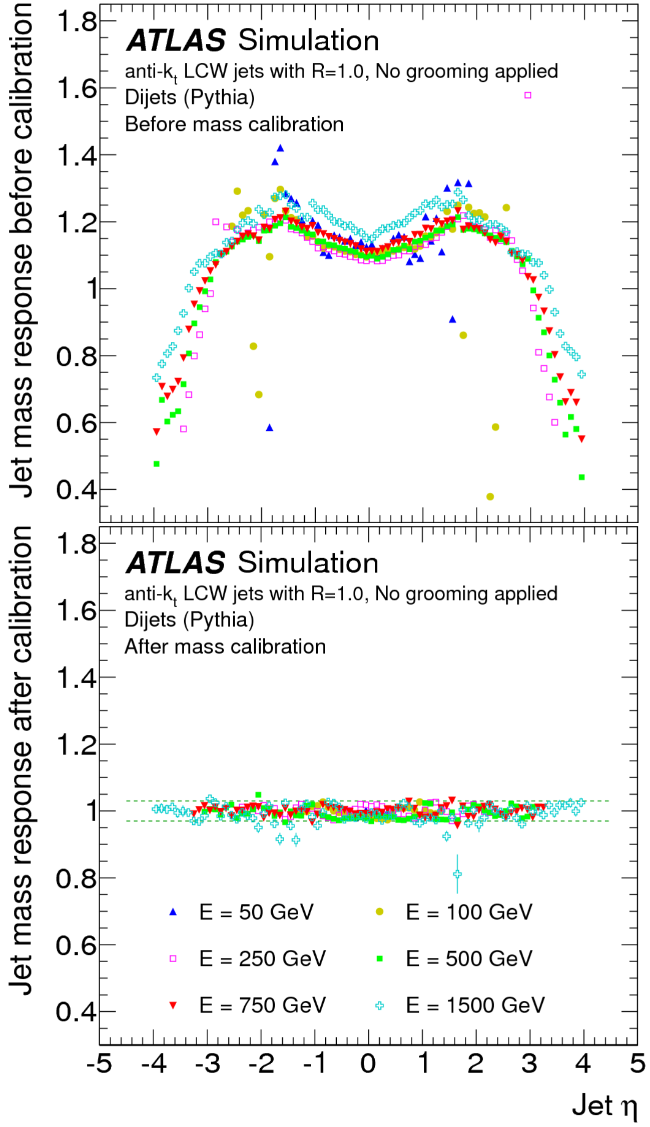
- Concerned mostly with the core (2σ) of the distribution
- Two methods used to subtract radiation (and in good agreement)
- Methods confirm improvements obtained with global sequential calibrations

Beyond the JES: Calibrating the Jet Mass



[*] arXiv:1406.0076

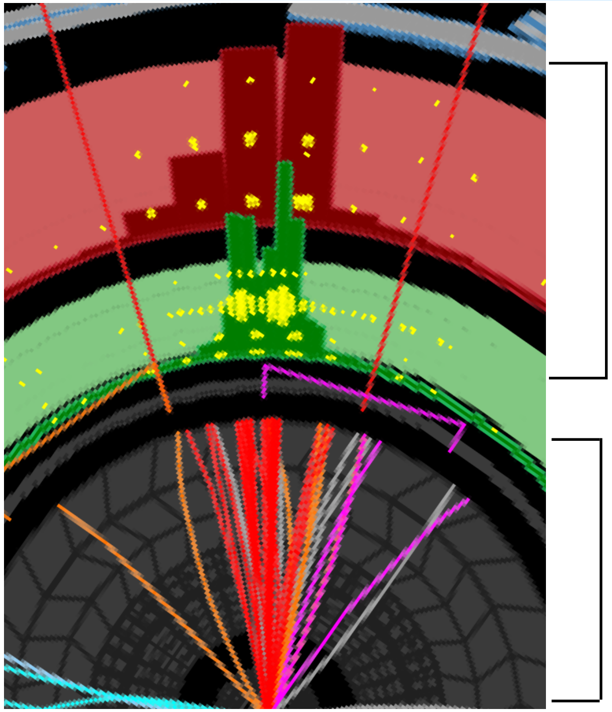
Same technique



[*] JHEP09 (2013) 076

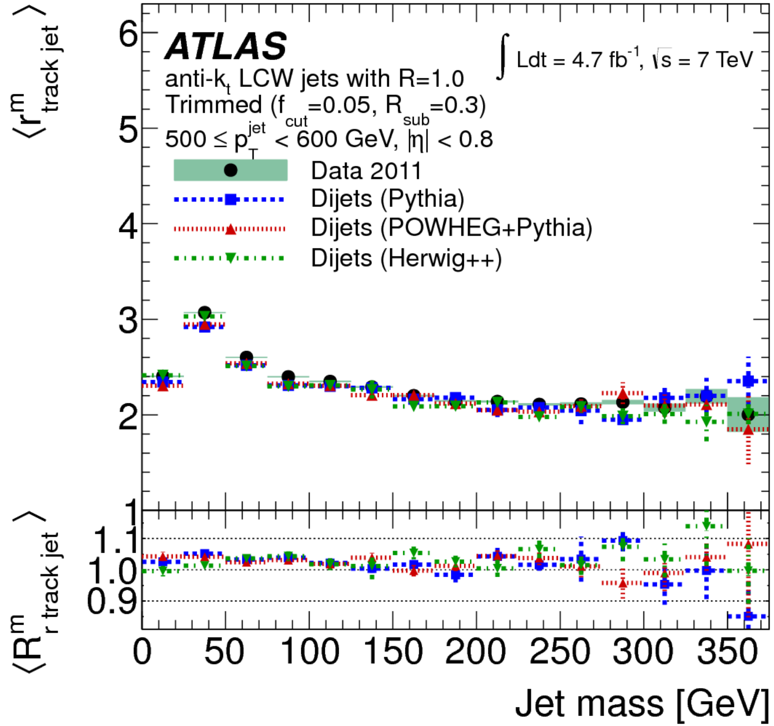
- Calibrated energy <u>doesn't mean</u> calibrated mass (same goes for systematics)
- Calibration improves resolution and teaches us many things about detector response
- Generic mass calibration trickier at low masses, easier for EW jets

Jet Mass Uncertainties: Tracking Measurements



calorimeter measurement

reference



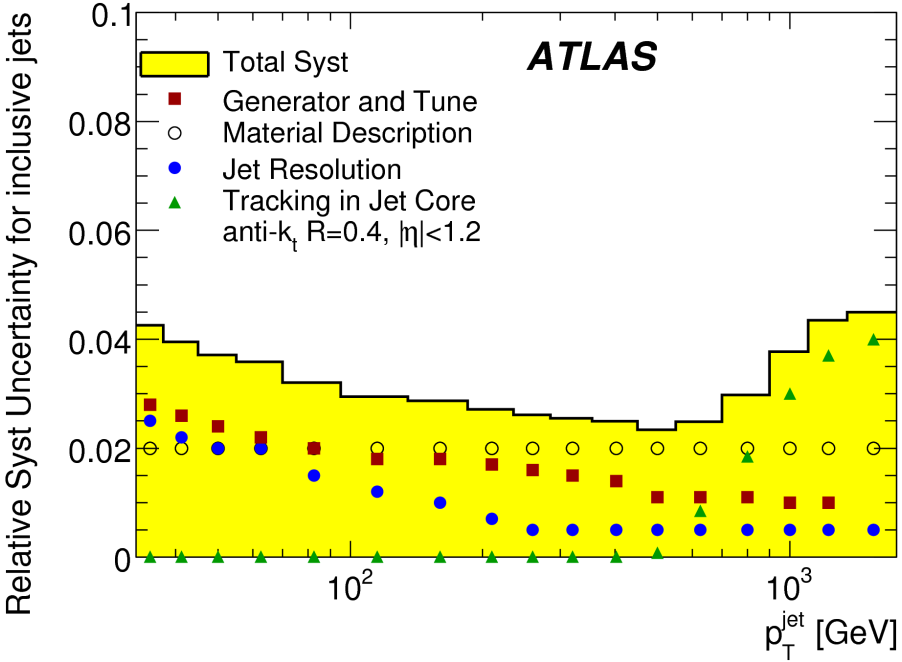
[*] JHEP09 (2013) 076

- Reference measurement is very precise, but of a quite different quantity than that of interest (large fragmentation systematics)

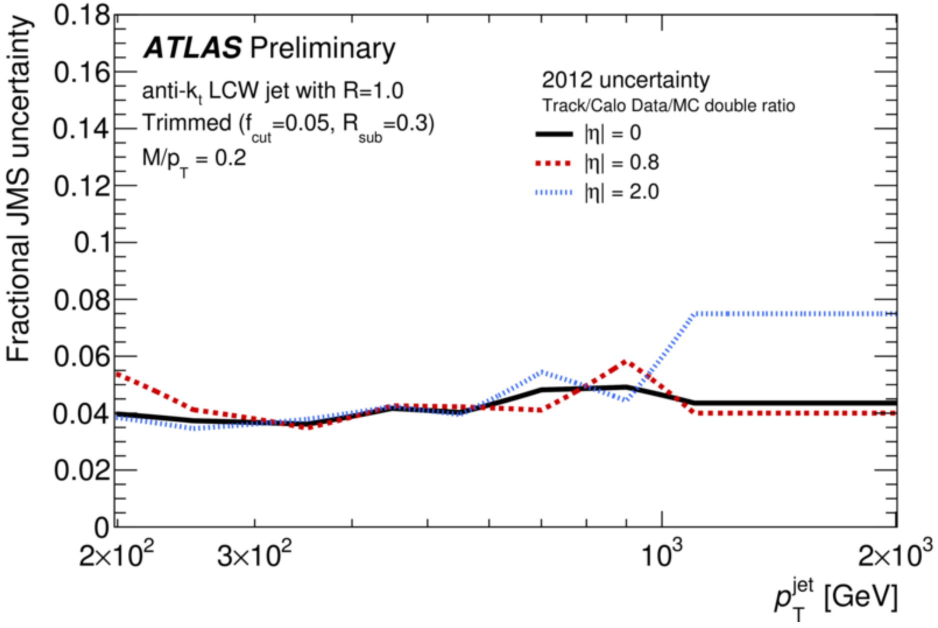
- Much more generic (do not exploit balance, can be applied to different topologies/ variables)

- Used in ATLAS for mass scale, splitting scales and N-subjetiness uncertainties

Jet Mass Uncertainties: Tracking Measurements



[*] arXiv:1406.0076

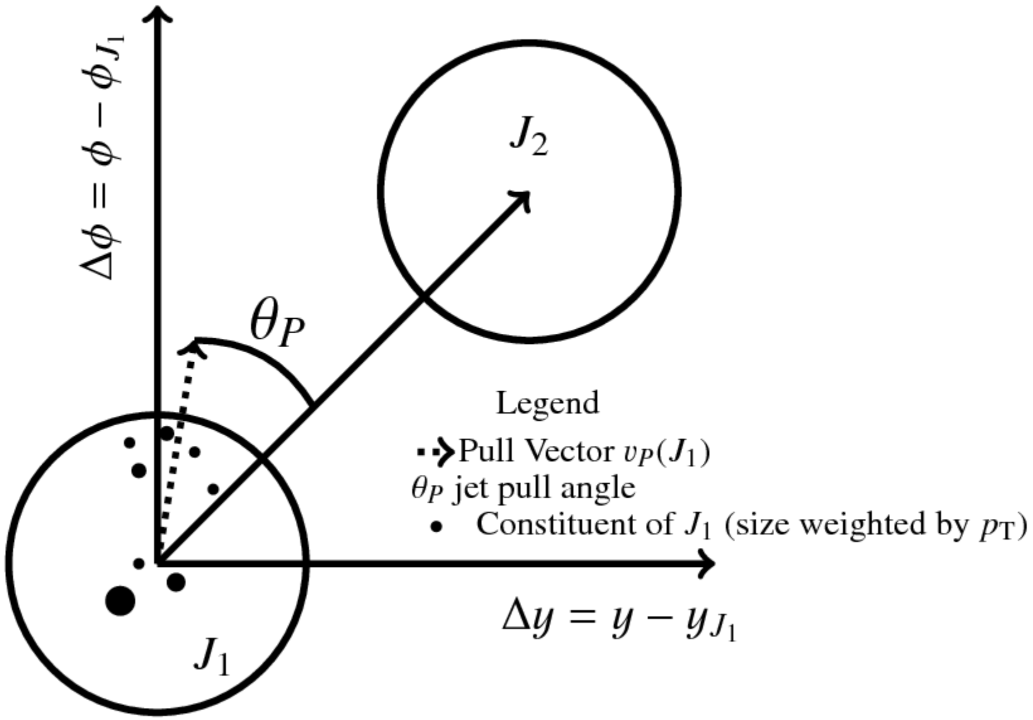


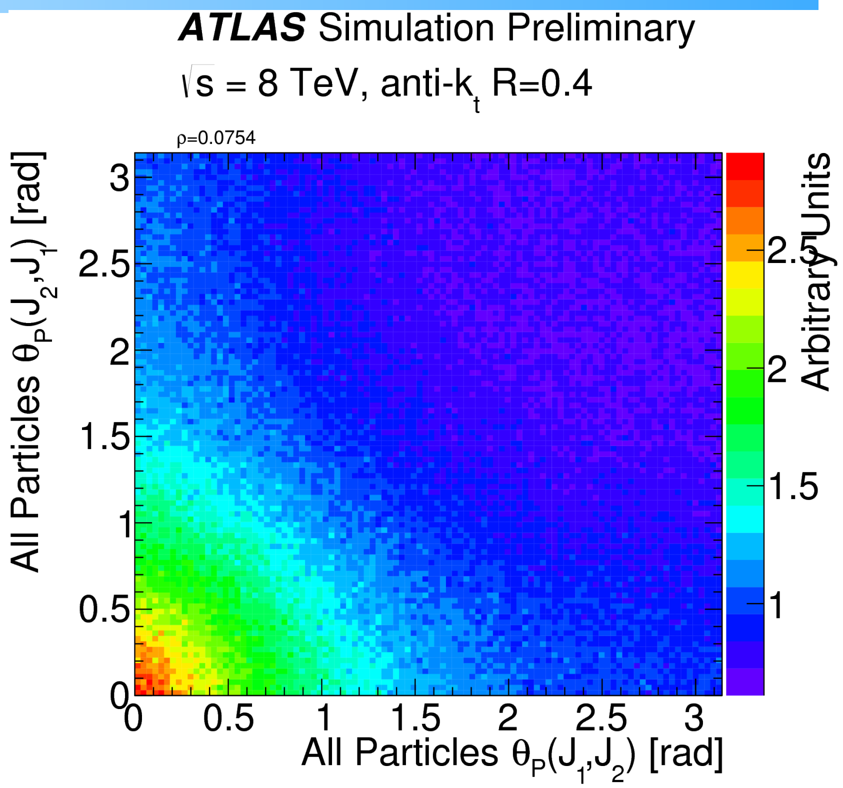
- Measurement of p_T^{trk}/p_T^{calo} pioneered early in the run to estimate the JES in a dijet sample (used also for b-jets, for instance)

- Versatile because reference doesn't require specific topology

- Used for mass/d_{12}/τ_n uncertainties

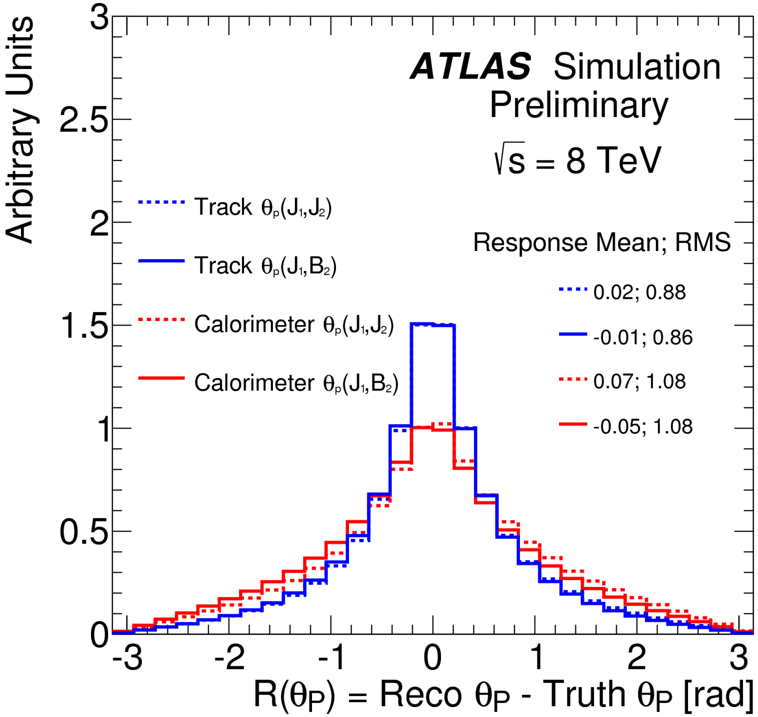
Sub+Superstructure: Jet Pull

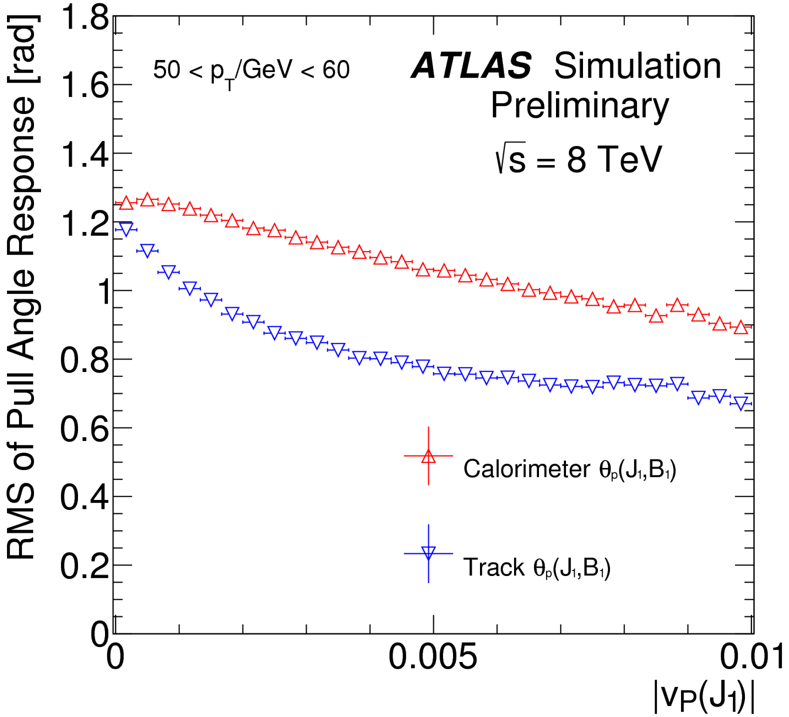




- Performance studies to understand whether we can measure and understand color flow
- Definition of pull angle same as CMS, both jets in a W "pull" towards each other

Sub+Superstructure: Jet Pull

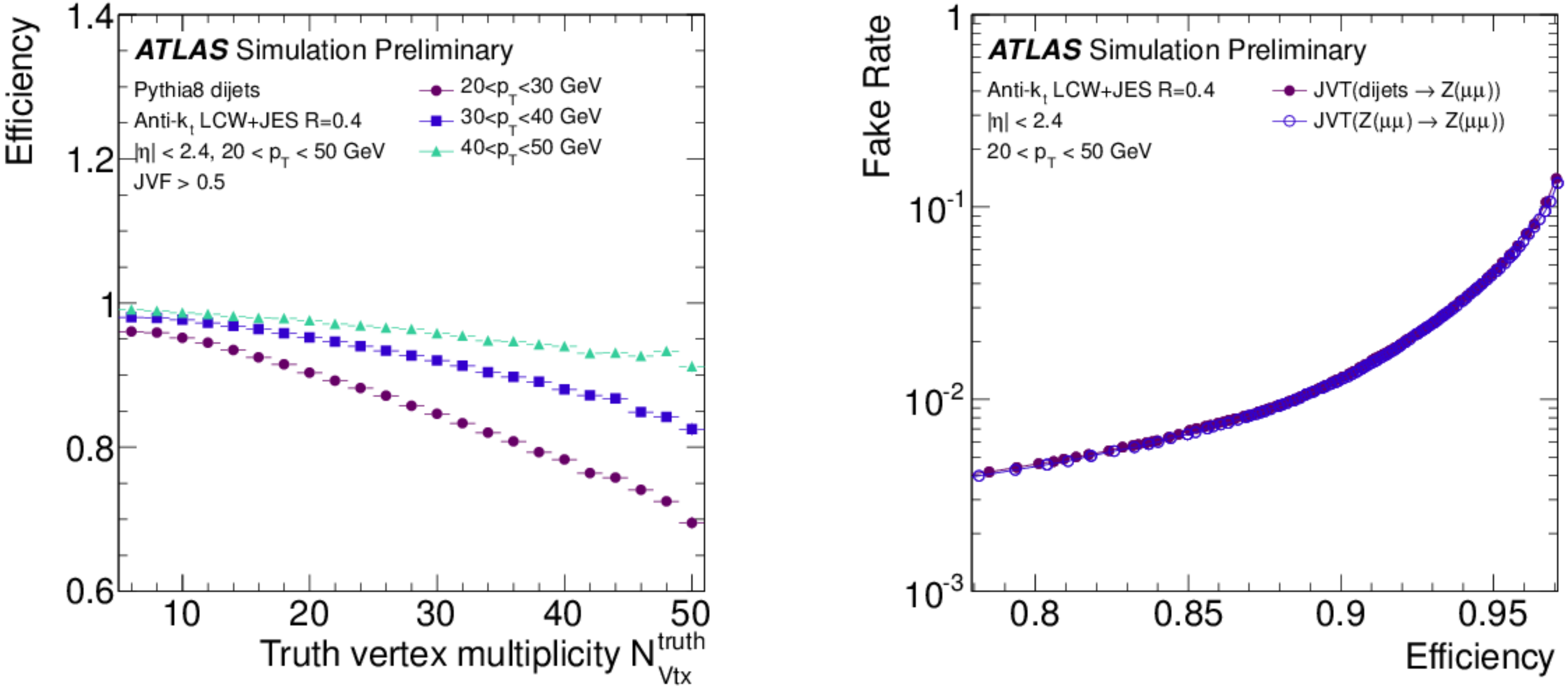




- Color flow is a subtle effect, detector resolution is not particularly good, so it becomes even more subtle

- Tracks and cuts on magnitude can be used to help obtaining better performance

Pile-up Jet Tagging



- Progress in pile-up rejection: solve the problem of pile-up dependent efficiency
- Demonstration of small q/g dependence (and also for b-jets)

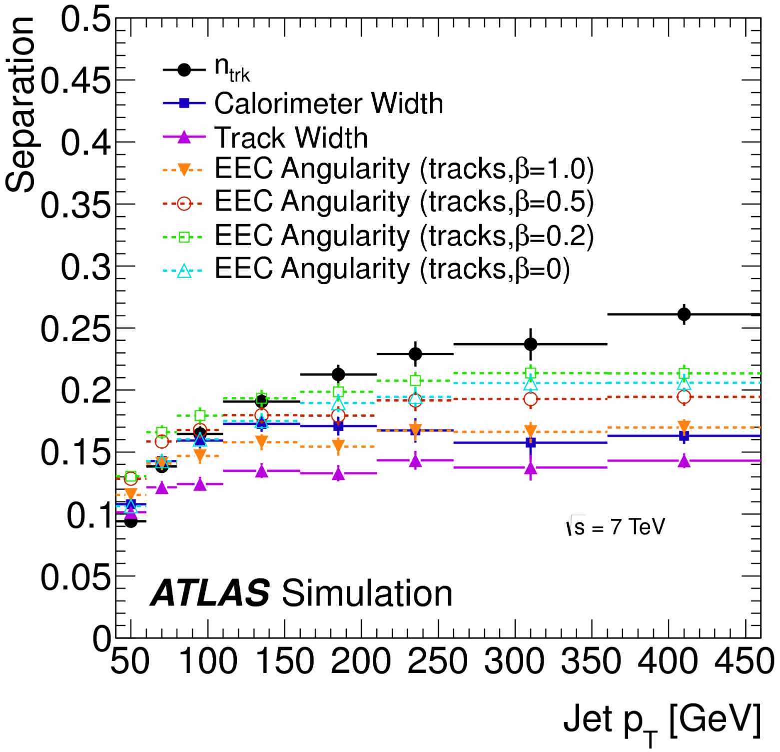
Quark/Gluon Tagging

- Important and challenging

- Large phase space is relevant (low p_T and large η)

- MCs show differences among them (and also differences with data)

- ATLAS has published a detailed study based on 2011 data



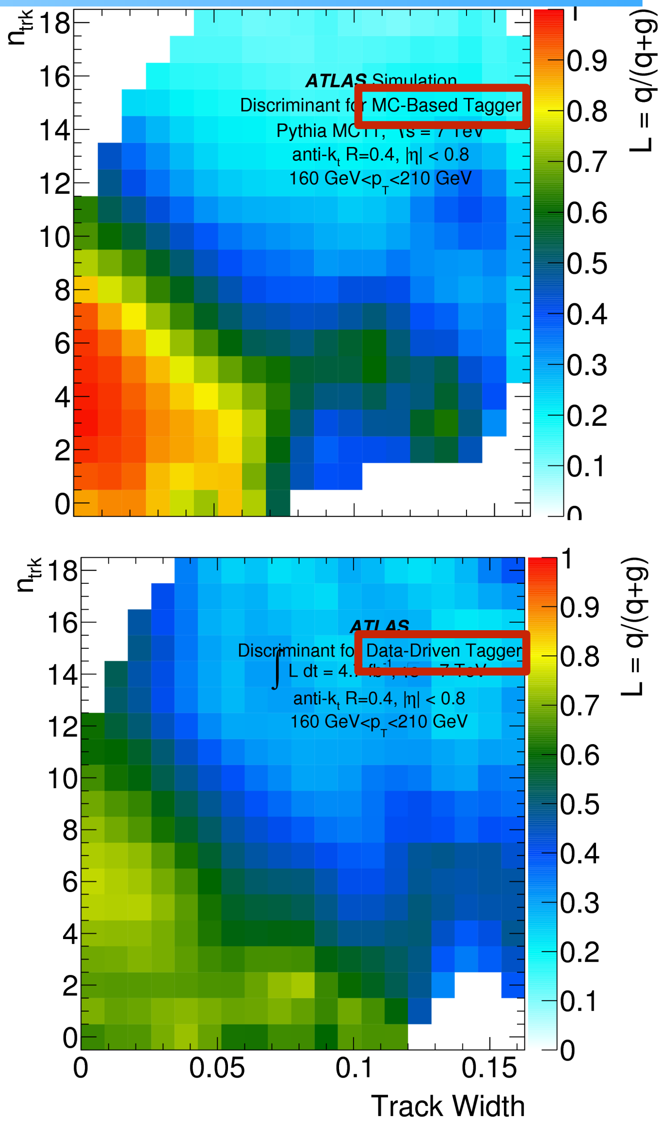
[*] arXiv:1405.6583

Quark/Gluon Tagging: Data Extraction

$$\begin{aligned} P_i(\eta, p_{\mathrm{T}}) = & \boxed{f_q(\eta, p_{\mathrm{T}})} \times \boxed{P_{q,i}(\eta, p_{\mathrm{T}})} \\ & + \boxed{f_g(\eta, p_{\mathrm{T}})} \times \boxed{P_{g,i}(\eta, p_{\mathrm{T}})} \\ & + \boxed{f_c(\eta, p_{\mathrm{T}}) \times P_{c,i}(\eta, p_{\mathrm{T}})} \\ & + \boxed{f_b(\eta, p_{\mathrm{T}}) \times P_{b,i}(\eta, p_{\mathrm{T}})} \\ & + \boxed{f_{\mathrm{fake},i}(\eta, p_{\mathrm{T}}) \times P_{\mathrm{fake},i}(\eta, p_{\mathrm{T}})} \end{aligned}$$

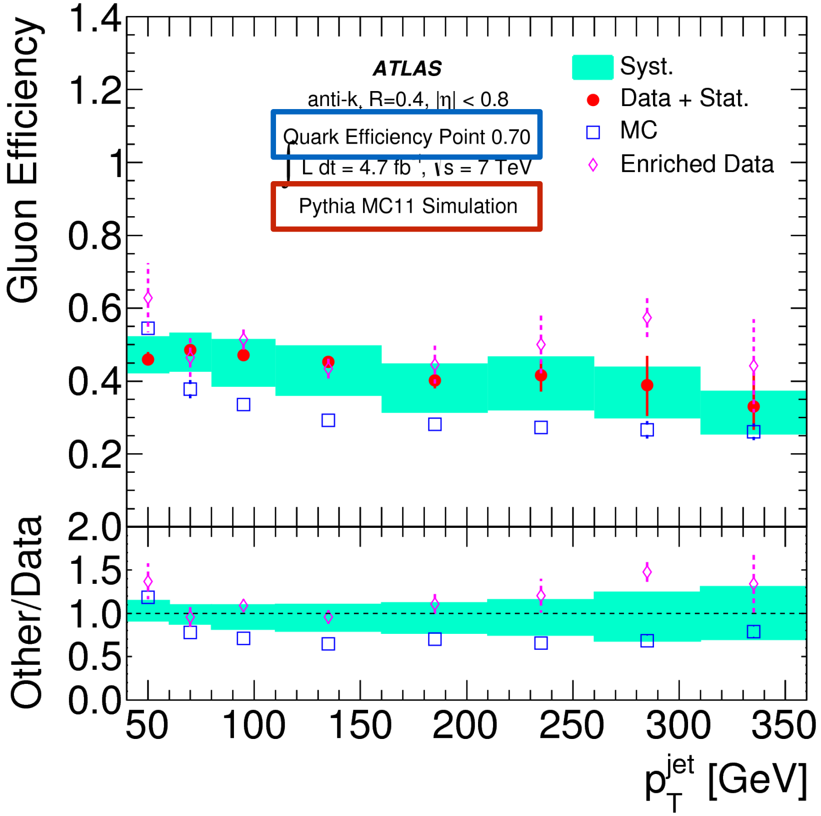
for γ+jet only →

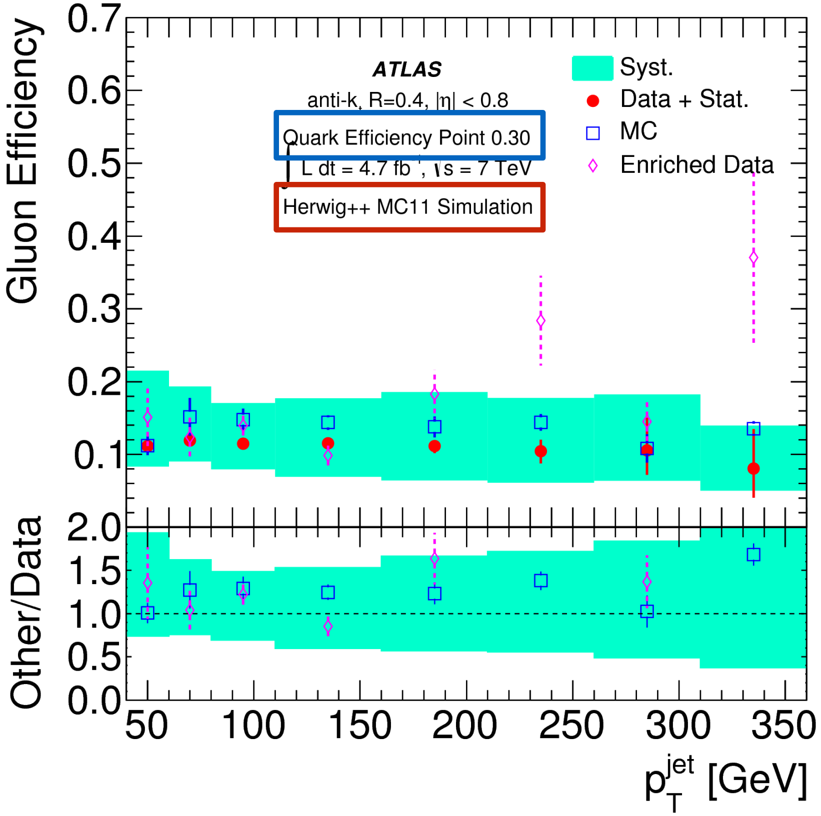
- Build a data-driven tagger to:
  - Extract properties independently for quarks and gluons
  - Depend on the MC only at second order
- Likelihood distribution is compressed: less discrimination in data



[*] arXiv:1405.6583

Quark-Gluon Tagging: Performance

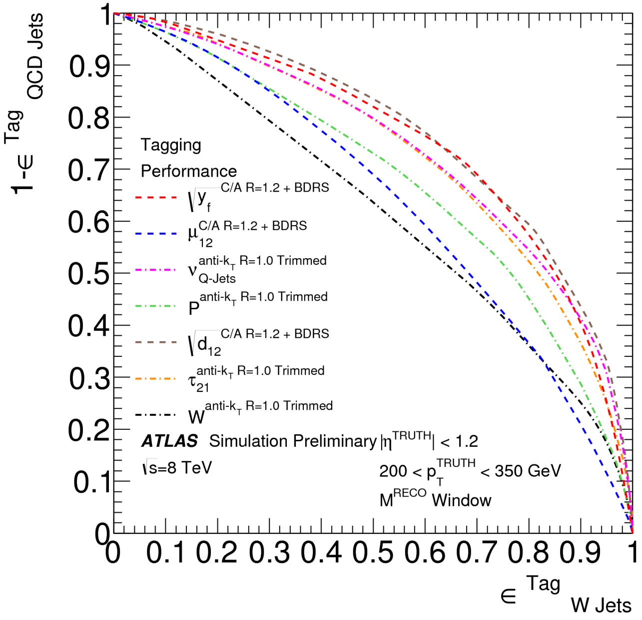


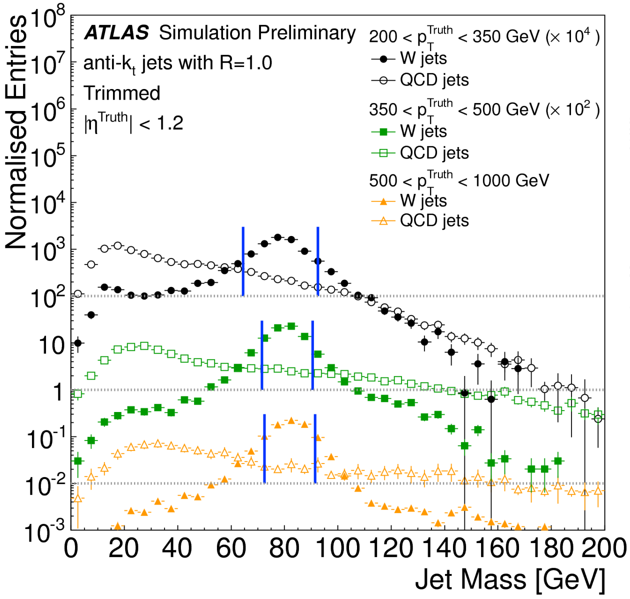


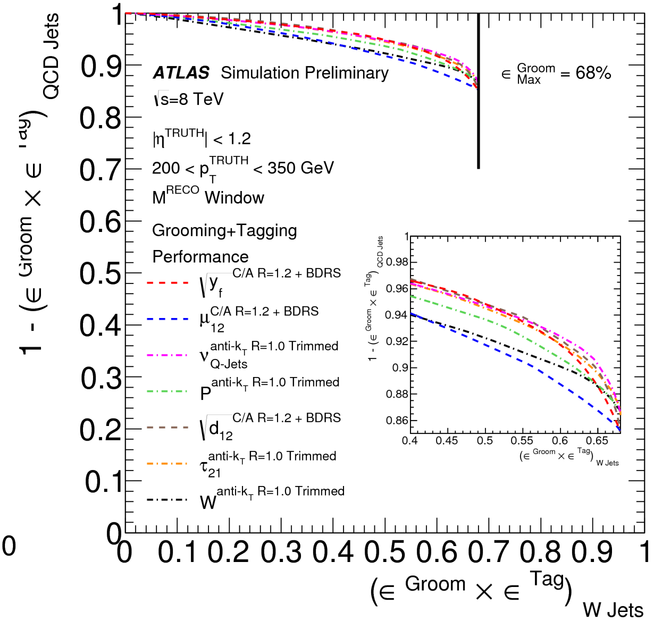
[*] arXiv:1405.6583

- Data tagger is more performant than Herwig++, less than Pythia
- Systematic uncertainties do not cover the difference for all operating points

Boosted Boson Tagging: Optimizations





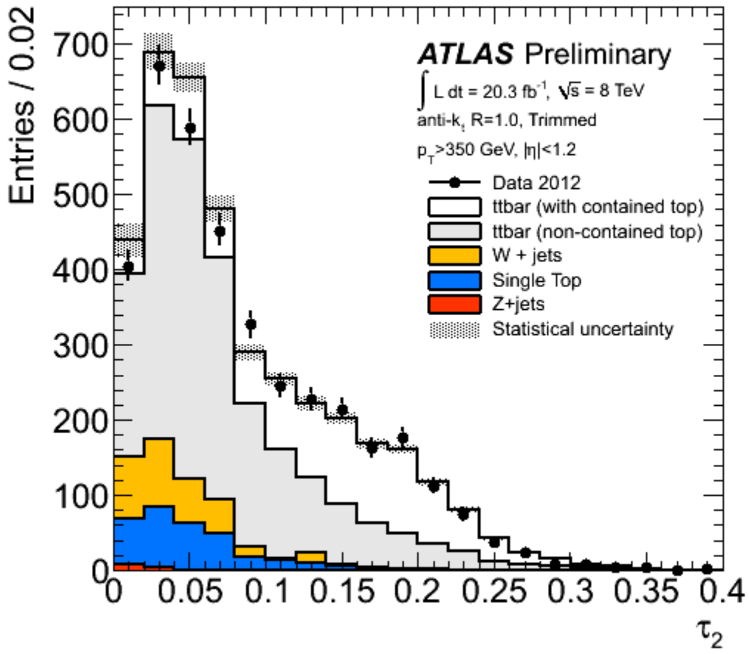


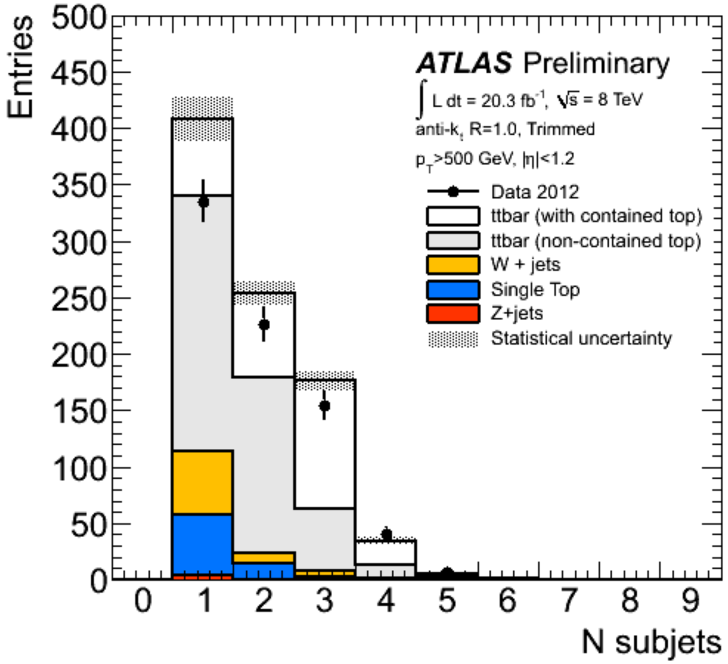
- First results made available earlier this year (even though some version already used for boosted W/Z cross-section measurement)

- Mass variables clearly extremely powerful, but can get better

⇒Emerging from BOOST: how precisely can we determine the y axis?

Top-tagging: Inputs

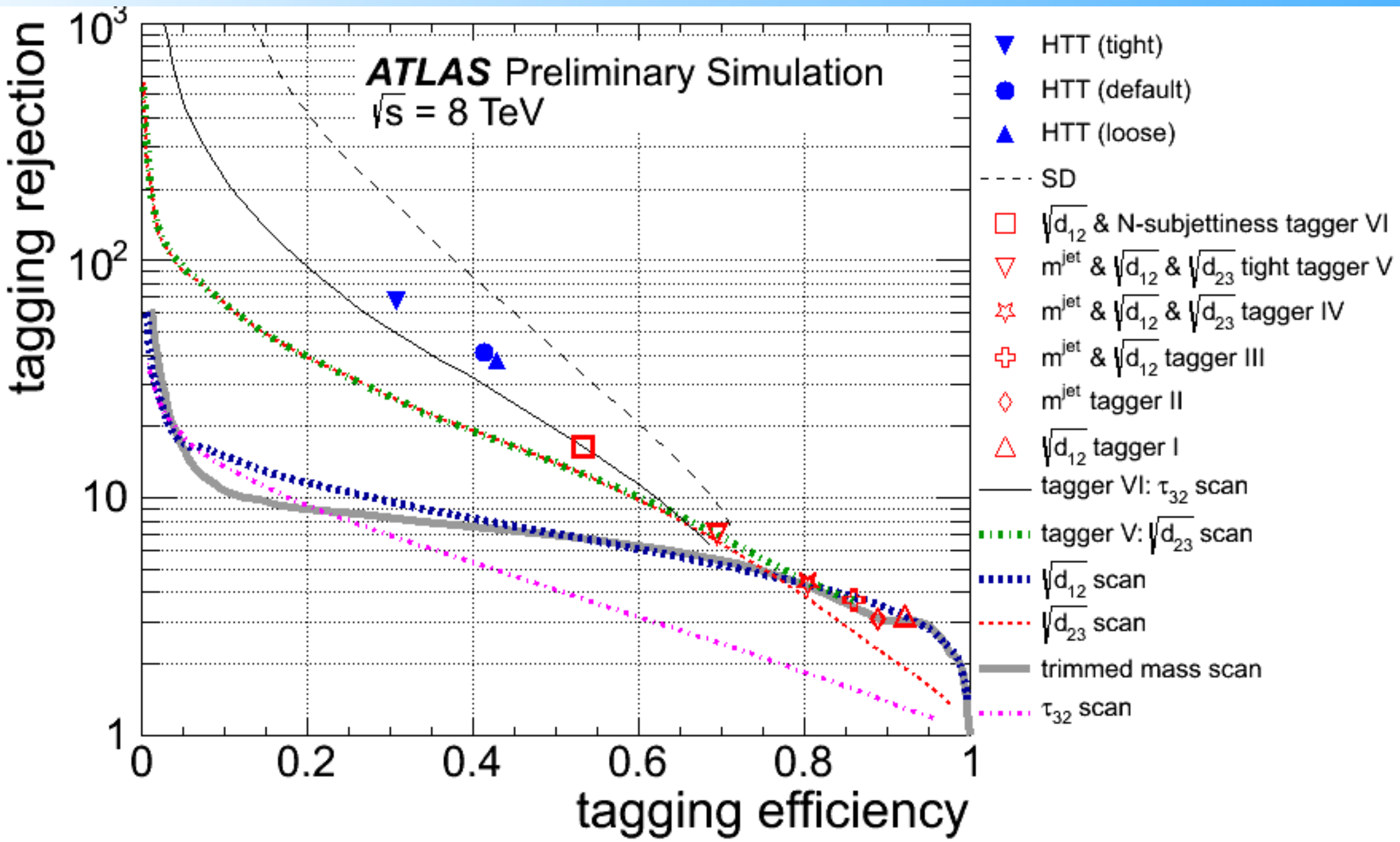




- Top-tagging already a bit more sophisticated

- Not perfect agreement in all variables used for tagging, but pretty good agreement for the most part

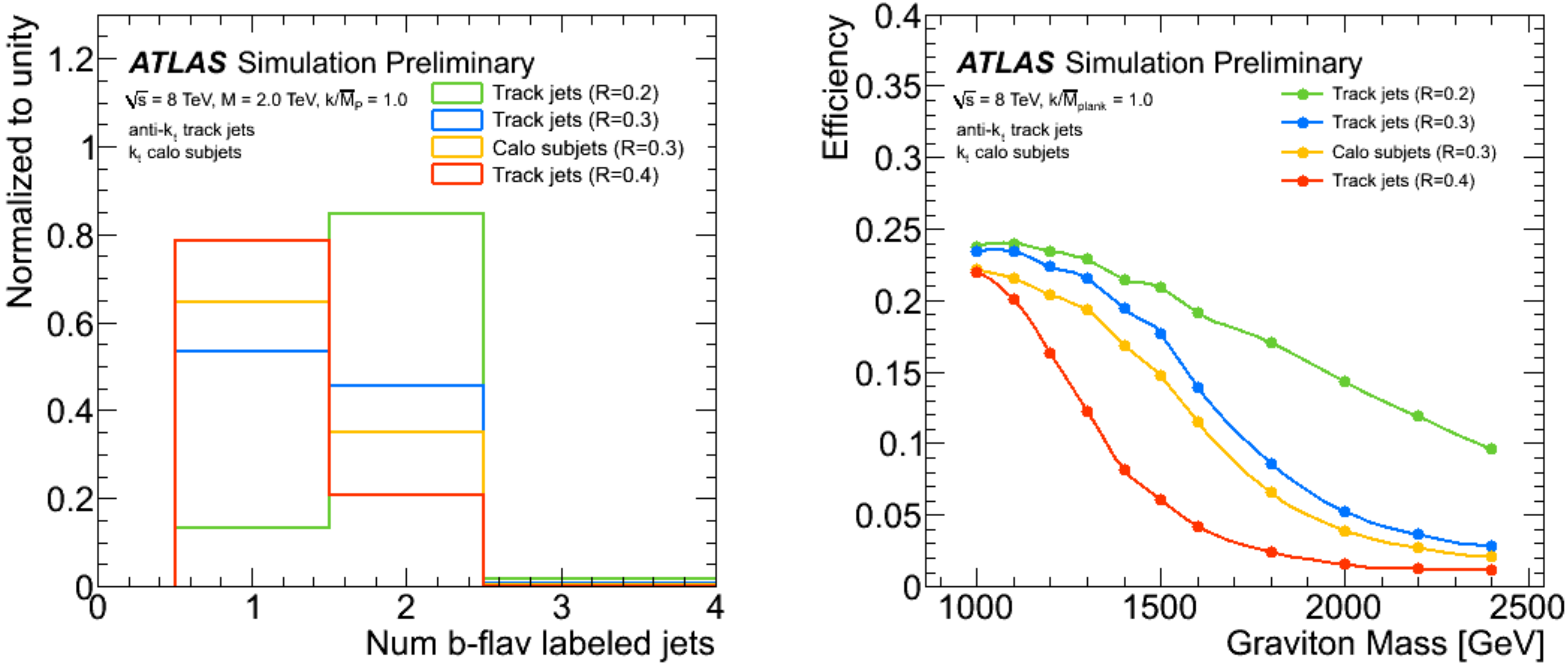
Top-tagging: Performance



- Summary of the latest performance in ATLAS (updates to the HTT to come soon)

⇒ Some more focus recently on the systematics for these curves

Fat-Bottomed Jets: b-tagging in Boosted Objects



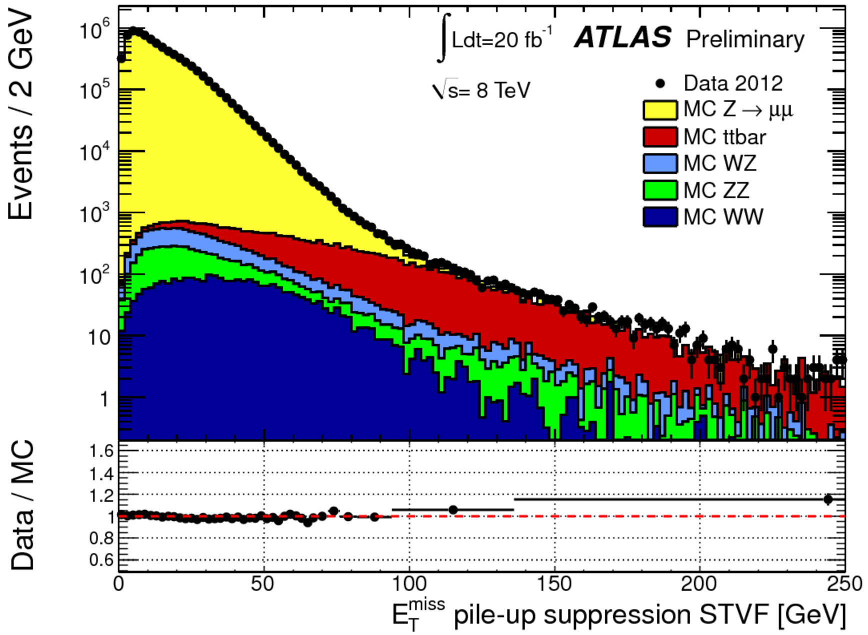
- Effort ongoing to decouple b-tagging from jet reconstruction

⇒ Flexibility for boosted object reconstruction and tagging optimization

Missing E_T Reconstruction

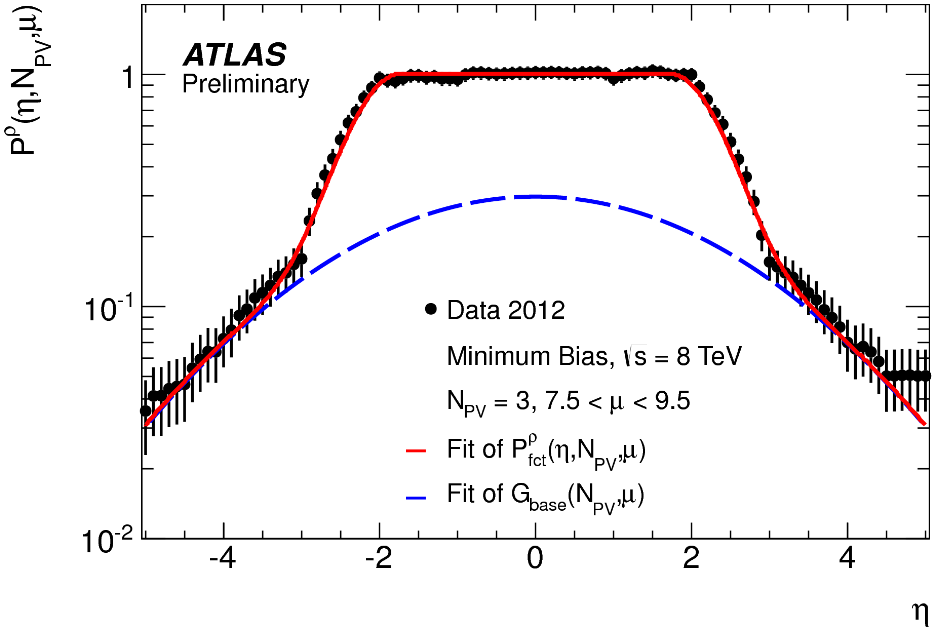
E_T(jets) + $E_T(\gamma/e/\mu/\tau)$ + Soft term

- For most analyses, selection of objects is the same
- Sensitive analyses performed dedicated selection
- Mostly care about the understanding the soft term (but hard terms are important to understand performance)



Pile-up Suppression in the Soft Terms

- Energy density in forward region heavily suppressed



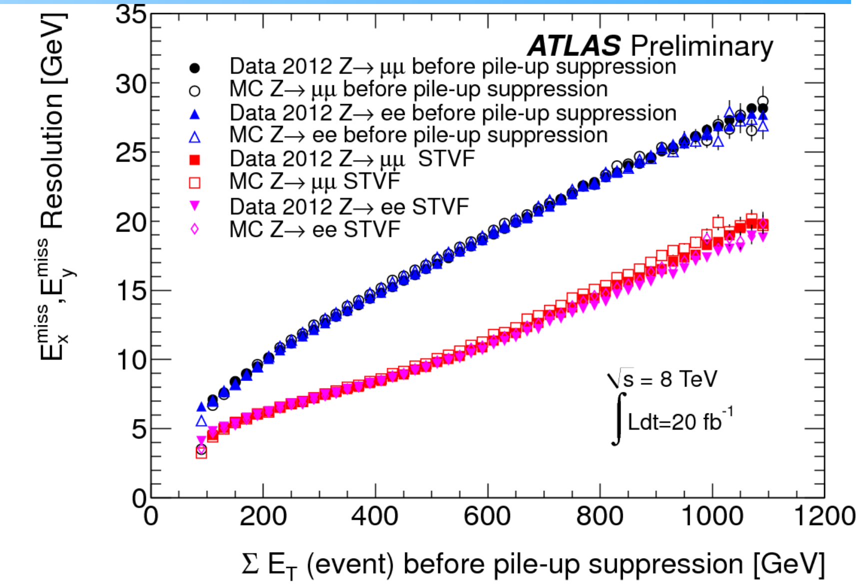
- Indication that pile-up deposits merge with signals in the same cluster

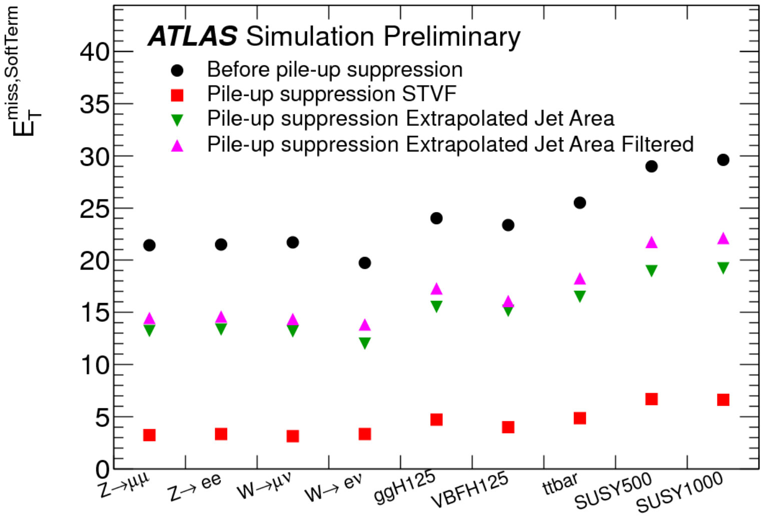
- Use of tracks in analogy to JVF

$$STVF = \left(\sum p_T^{track,PV} \Big/ \sum p_T^{track}\right)_{\substack{unmatched\\objects}}$$

Pile-up Suppression in the Soft Terms

- Resolution seems best with STVF
- Soft term scale is heavily suppressed
- Appropriate for certain final states, but clearly not optimal

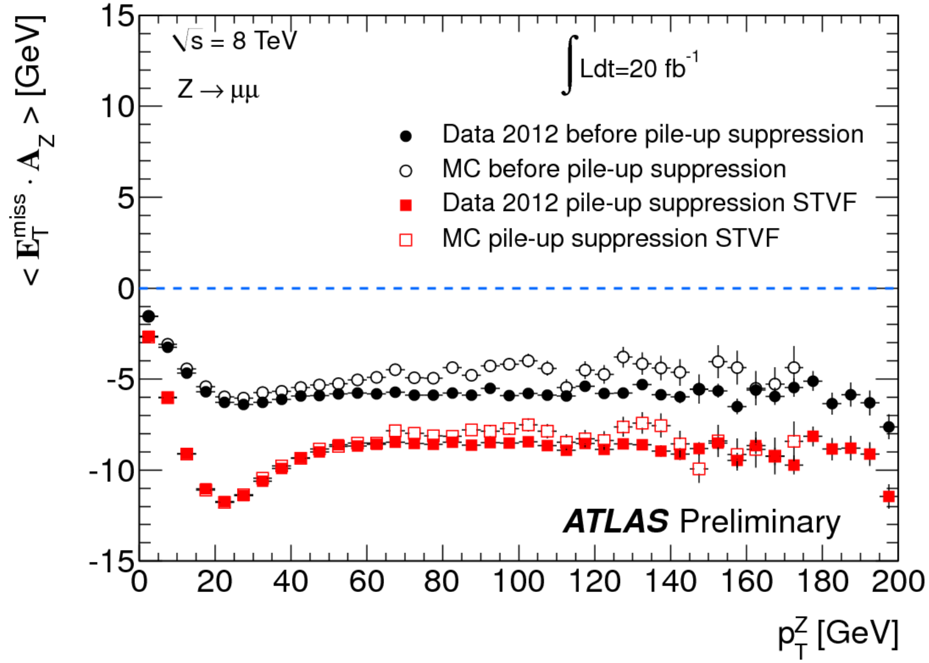


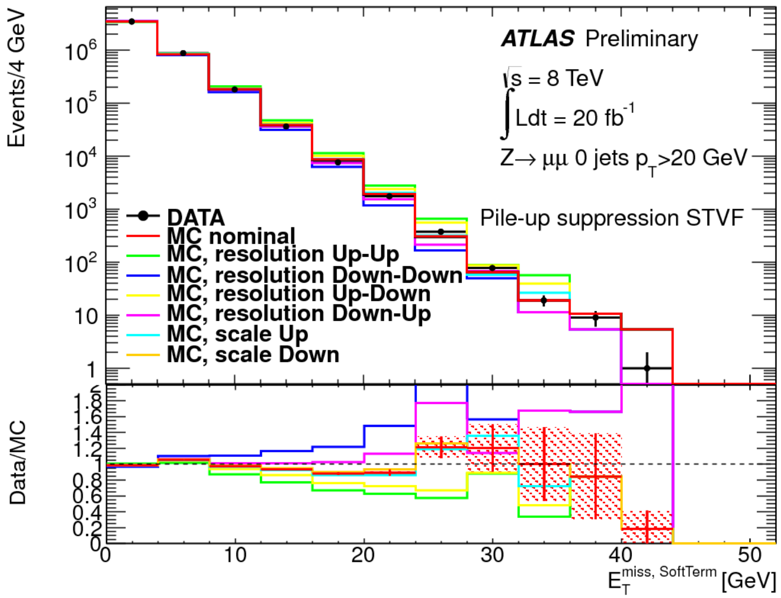


Soft Terms Validation in Data

- Clear issue with scale at low p_T

- Area-based methods are somewhere in the middle





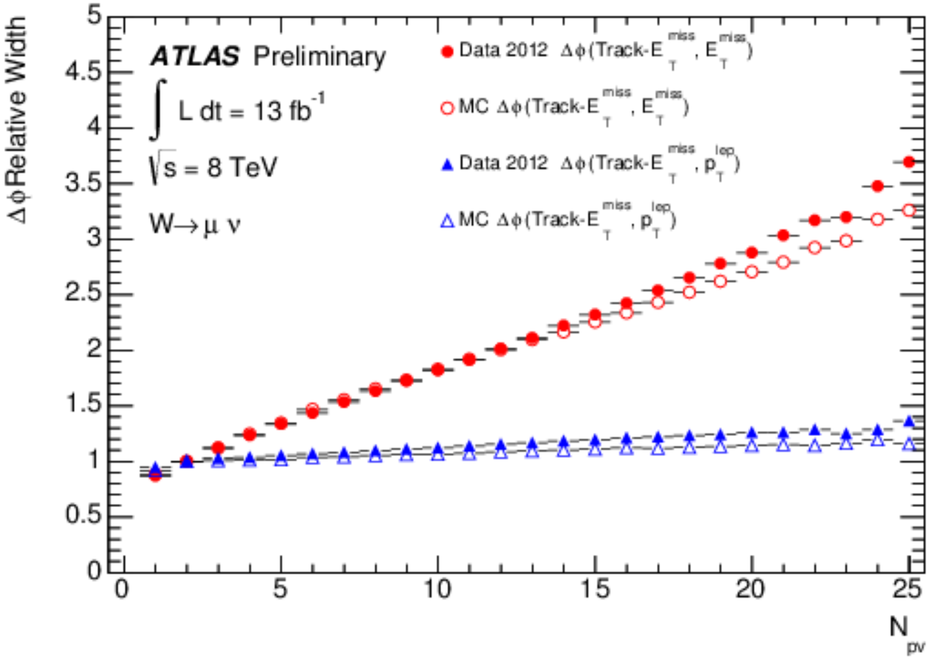
- Differences between data and MC covered by
  - Soft terms scale
  - Transverse resolution
  - Longitudinal resolution

Track-Based Soft Terms

- Soft term performance degradation is due to pile-up

- Track-based E_T^{miss} used quite often for background rejection

- Pile-up dependence much reduced, data/MC agreement equally good



Summary and Conclusions

- ATLAS has a very mature program for understanding JES and JER

- New developments coming through, in particular in the aspects of pile-up suppression

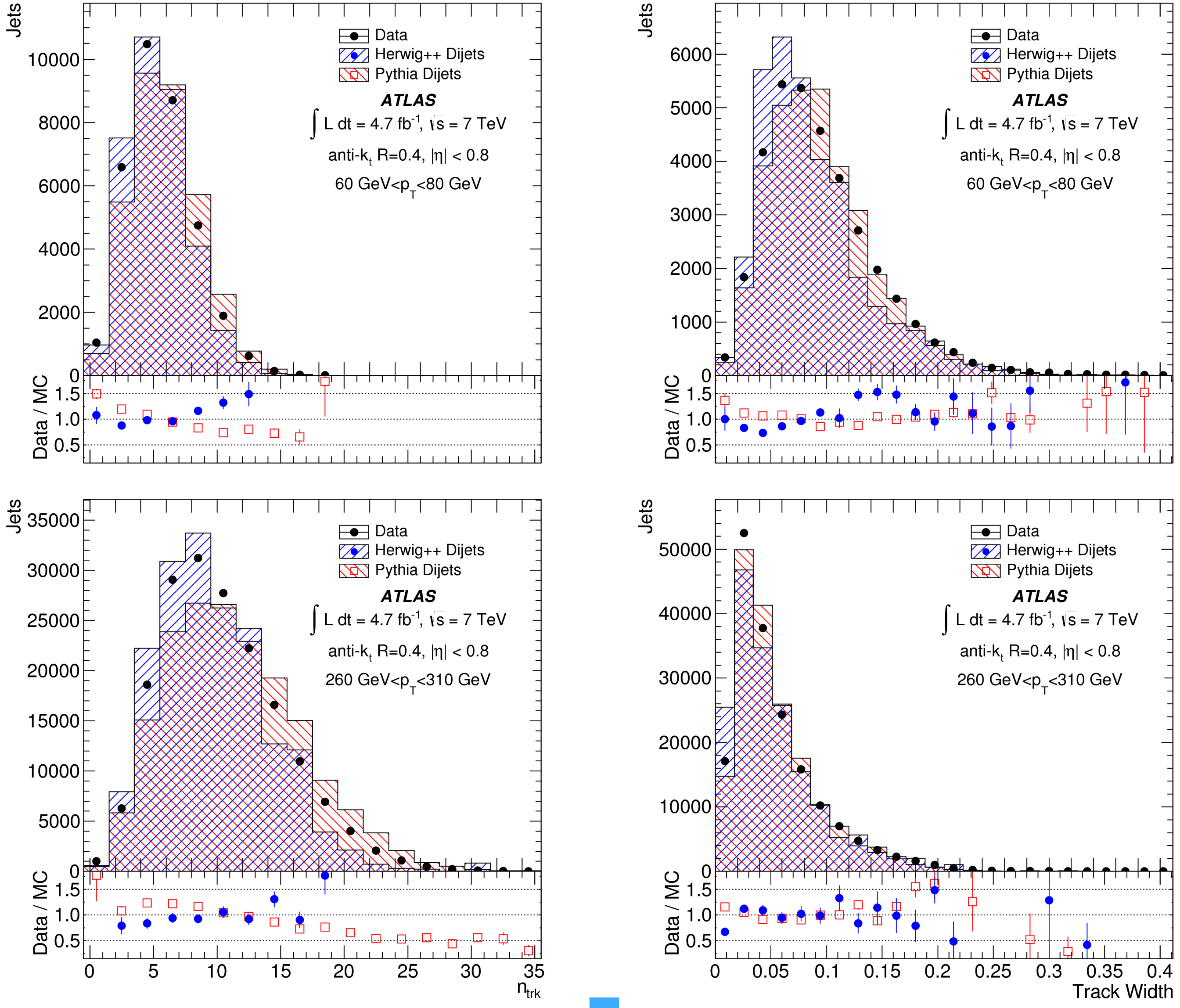
- Techniques in the JES program have found their way to jet substructure studies

- Those techniques and their application to boosted object tagging systematics are still evolving, expect a lot of activity during Run 2

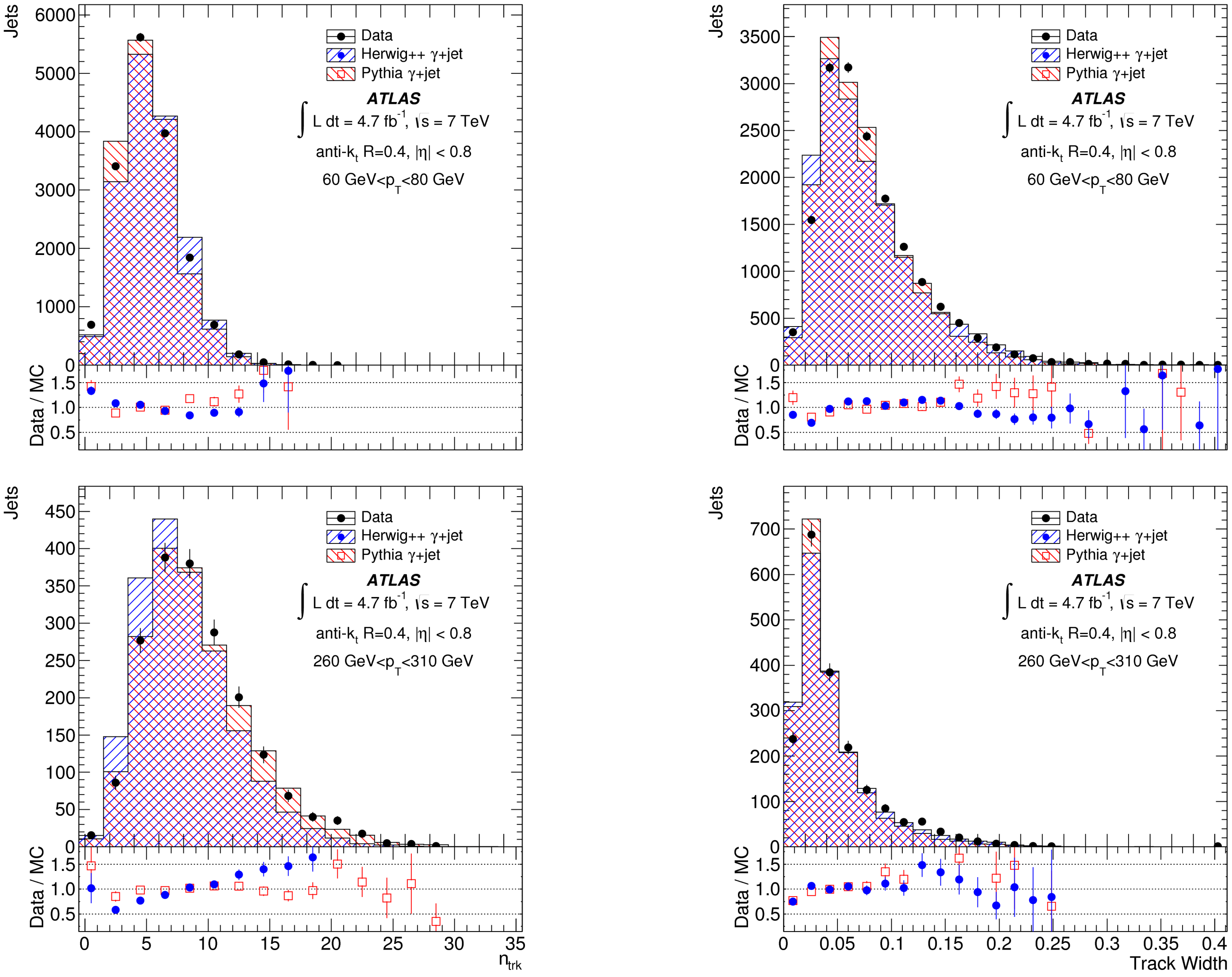
- No silver bullet still found for suppressing pile-up for the missing ET soft terms, but this will remain important through to the HL-LHC

Back-up Slides

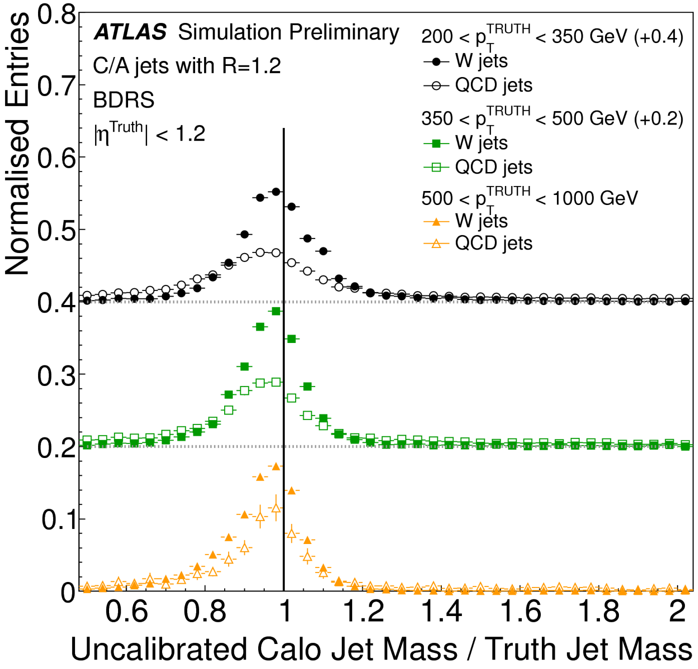
Quark/Gluon Tagging: Dijet Distributions



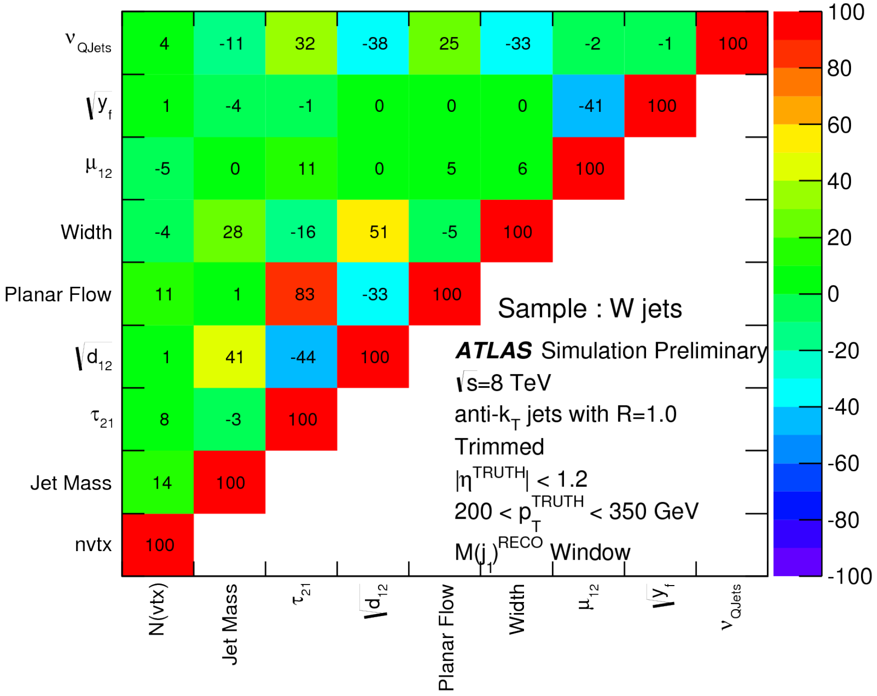
Quark/Gluon Tagging: γ+jet Distributions



Jet Mass Response



Boosted Boson Tagging Correlations



Sample : W jets

ATLAS Simulation Preliminary

$\sqrt{s}$=8 TeV

anti-k_T jets with R=1.0

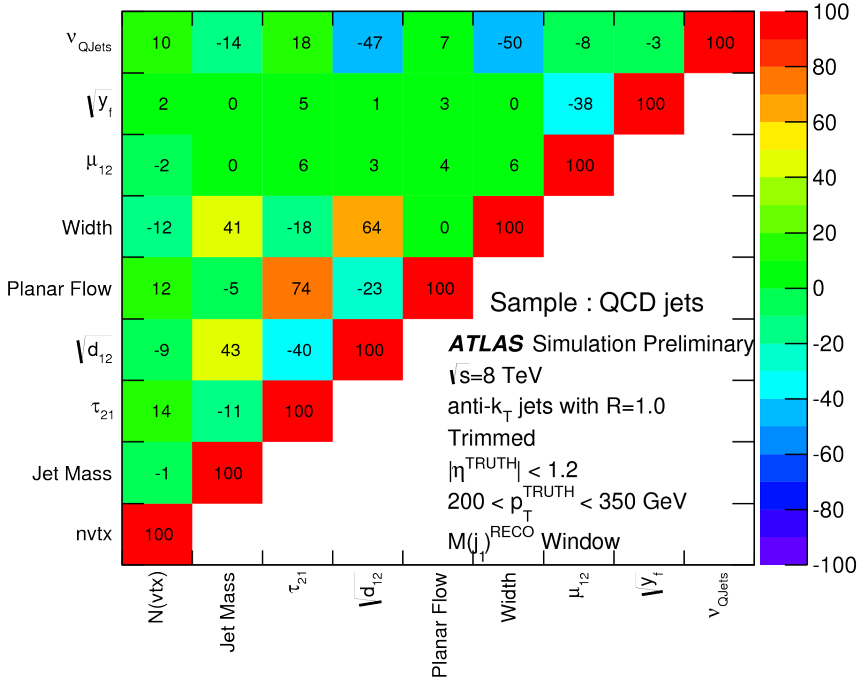
Trimmed

$|\eta^{TRUTH}| < 1.2$

$200 < p_T^{TRUTH} < 350$ GeV

$M(j_1)^{RECO}$ Window

	N(vtx)	Jet Mass	τ_{21}	$\sqrt{d_{12}}$	Planar Flow	Width	μ_{12}	$\sqrt{y_f}$	ν_{QJets}
ν_{QJets}	4	-11	32	-38	25	-33	-2	-1	100
$\sqrt{y_f}$	1	-4	-1	0	0	0	-41	100	
μ_{12}	-5	0	11	0	5	6	100		
Width	-4	28	-16	51	-5	100			
Planar Flow	11	1	83	-33	100				
$\sqrt{d_{12}}$	1	41	-44	100					
τ_{21}	8	-3	100						
Jet Mass	14	100							
nvtx	100								



Sample : QCD jets

ATLAS Simulation Preliminary

$\sqrt{s}$=8 TeV

anti-k_T jets with R=1.0

Trimmed

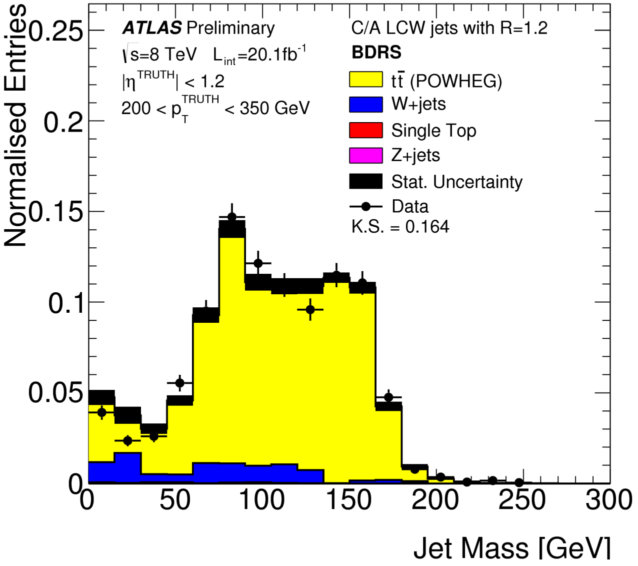
$|\eta^{TRUTH}| < 1.2$

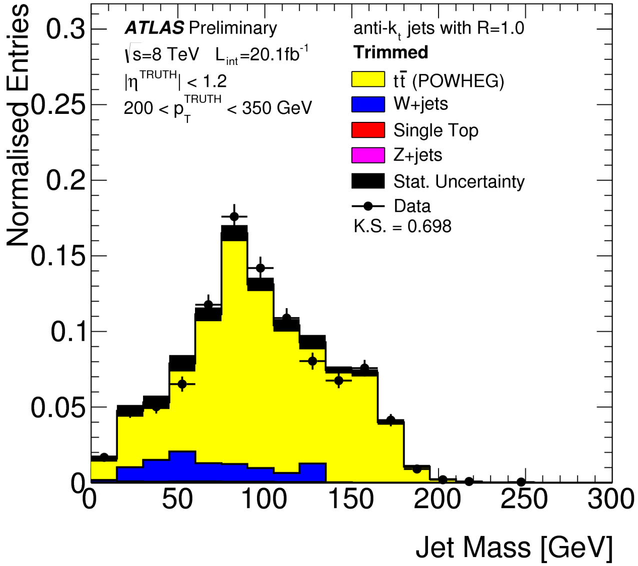
$200 < p_T^{TRUTH} < 350$ GeV

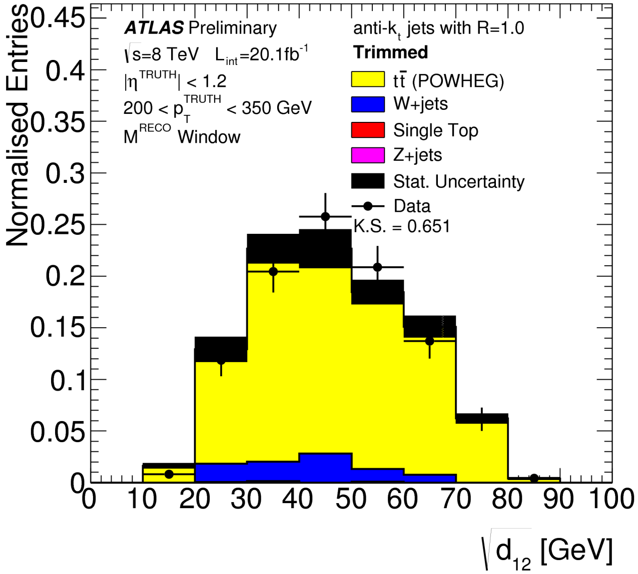
$M(j_1)^{RECO}$ Window

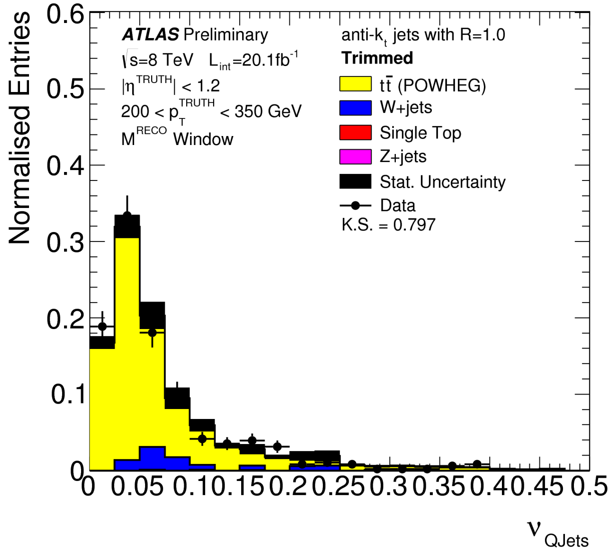
	N(vtx)	Jet Mass	τ_{21}	$\sqrt{d_{12}}$	Planar Flow	Width	μ_{12}	$\sqrt{y_f}$	ν_{QJets}
ν_{QJets}	10	-14	18	-47	7	-50	-8	-3	100
$\sqrt{y_f}$	2	0	5	1	3	0	-38	100	
μ_{12}	-2	0	6	3	4	6	100		
Width	-12	41	-18	64	0	100			
Planar Flow	12	-5	74	-23	100				
$\sqrt{d_{12}}$	-9	43	-40	100					
τ_{21}	14	-11	100						
Jet Mass	-1	100							
nvtx	100								

Boosted Boson Tagging Validation

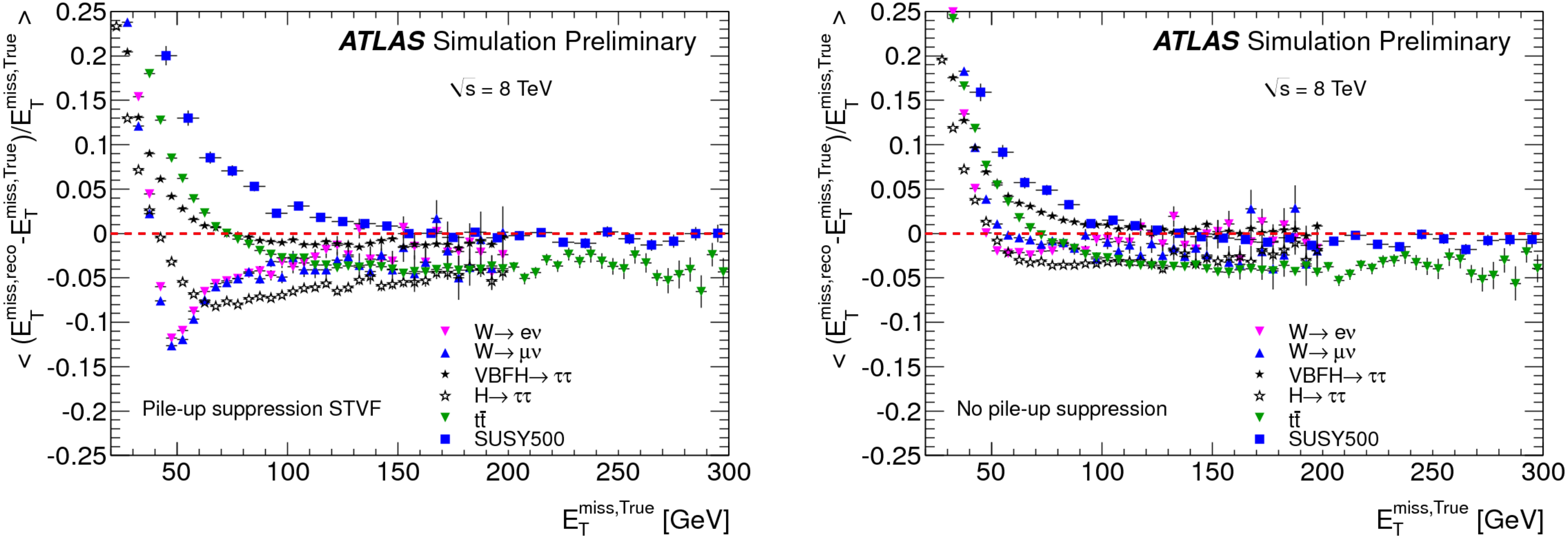




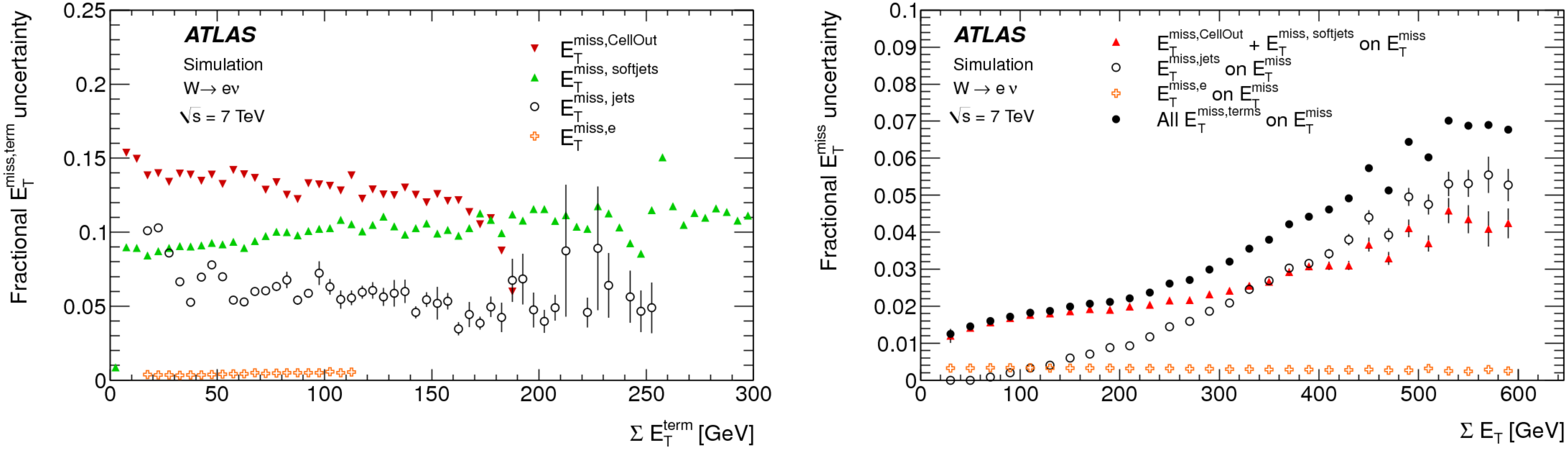




ETmiss Sample Dependence



ETmiss Systematic Uncertainties



ETmiss in Data