Jets, E_T^{miss}, and boosted jet identification in high-pileup conditions with ATLAS

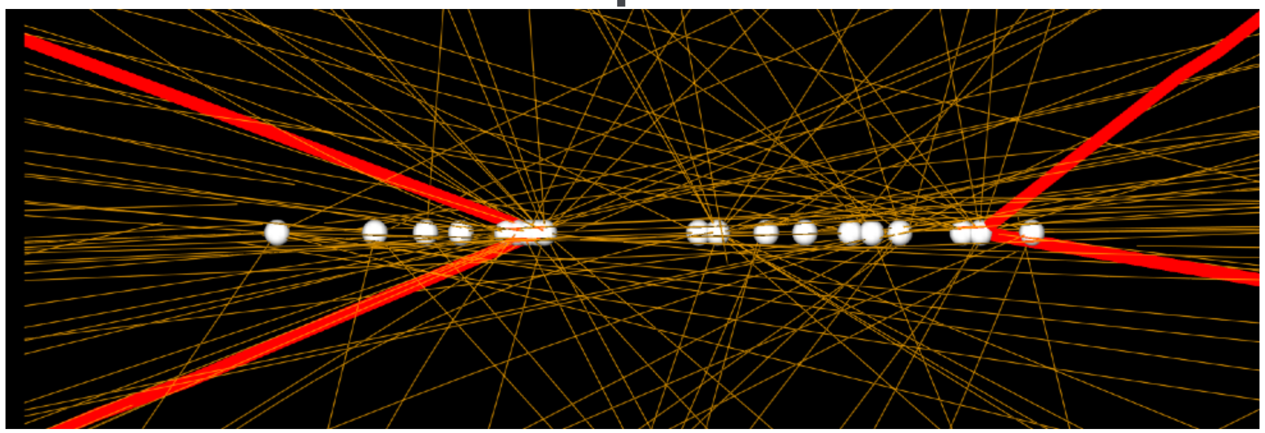
Jeff Dandoy
University of Pennsylvania
on behalf of the ATLAS Collaboration

34th International Conference on High Energy Physics
Seoul, South Korea
7 July 2018

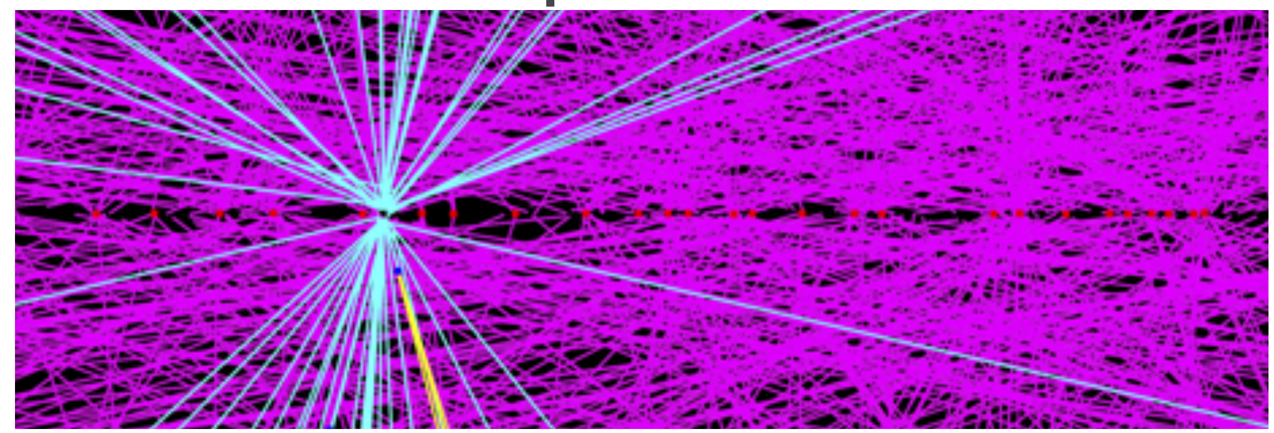
The Problem with Pileup

- Pileup from additional collisions is a big problem, and will be even worse at the High Luminosity LHC

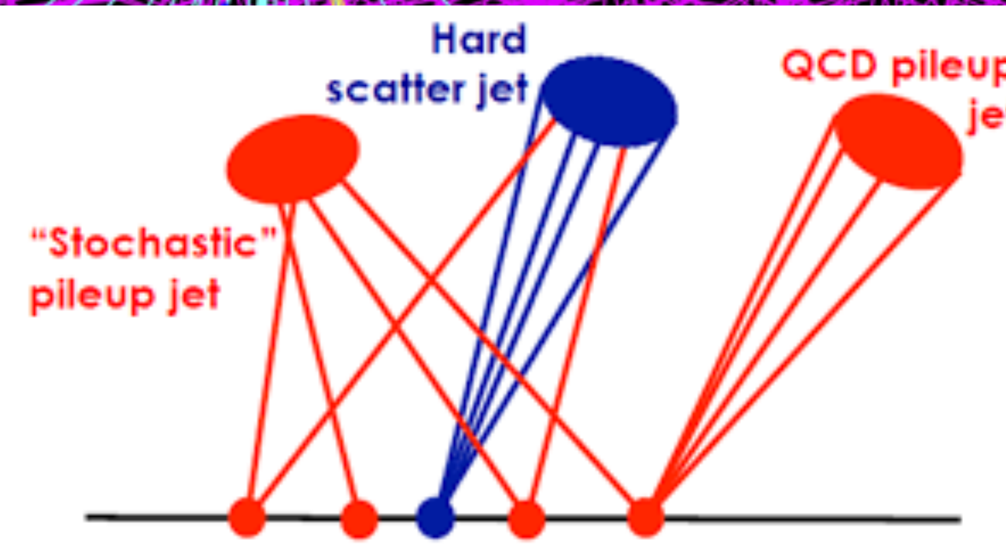
Pileup now

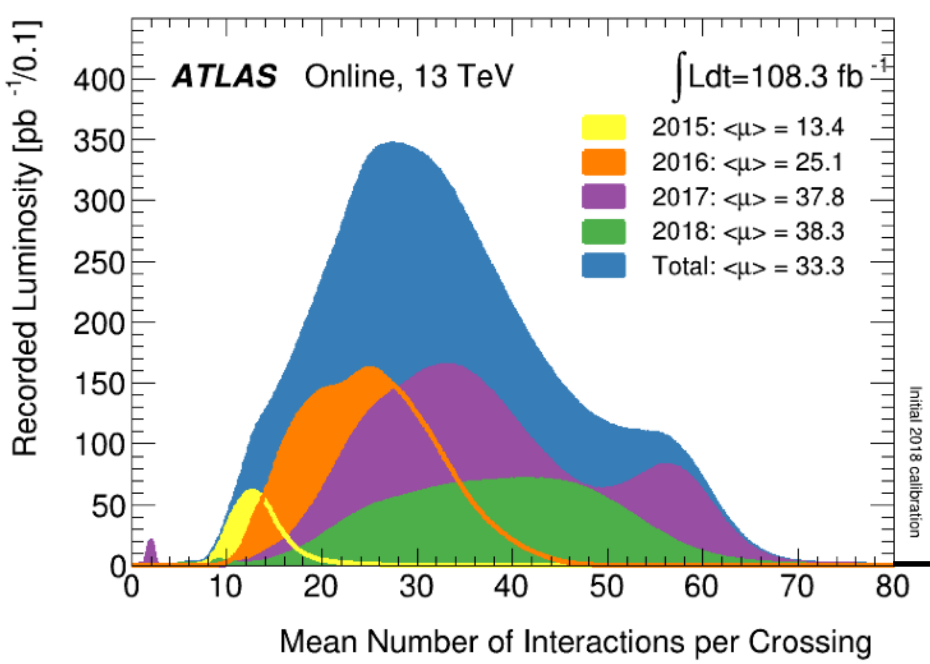


Pileup at HL-LHC



- Particles from pileup collisions can:
  - Add **additional jets** not from the hard-scatter
  - **Overlap with hard scatter jets**, altering their energy & structure

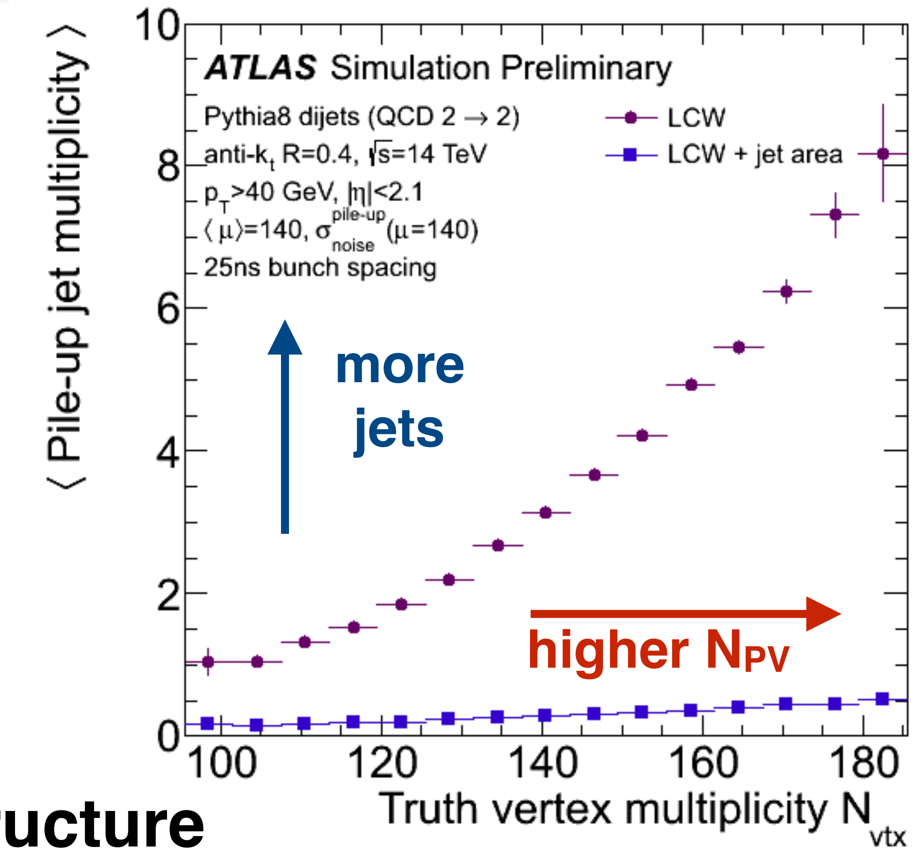




Every year the **# of interactions per event (μ)** increases, and will reach ~200 at the HL-LHC!

200

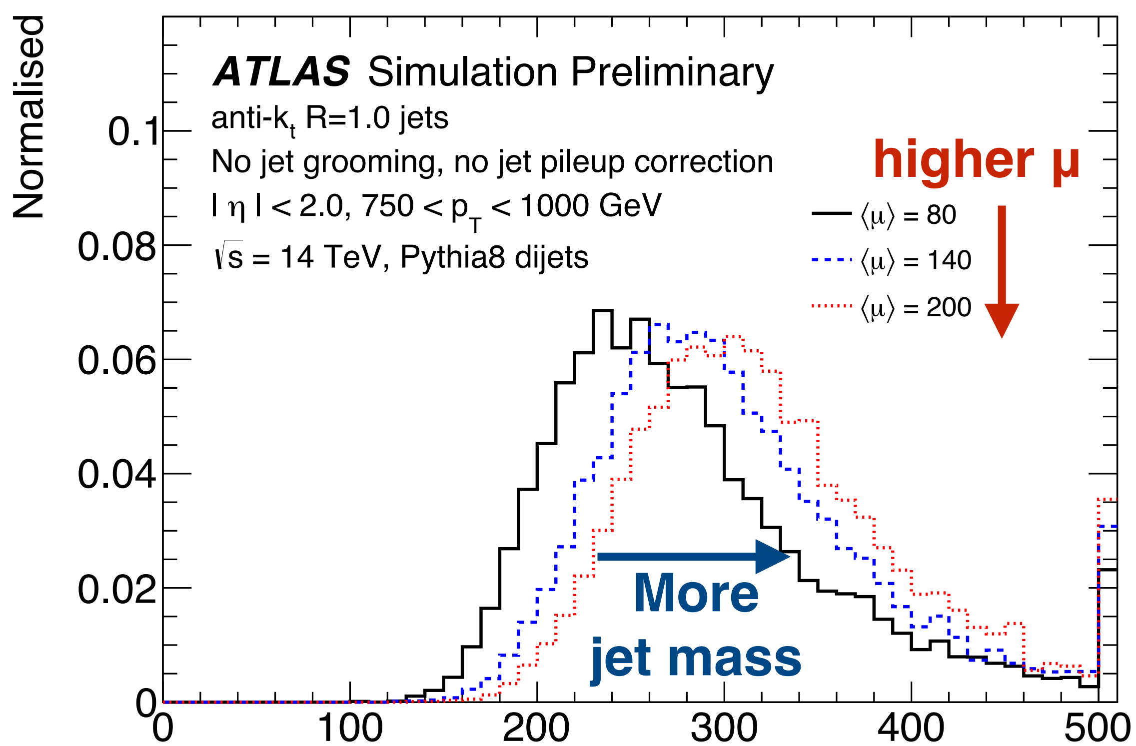
The Effect of Pileup



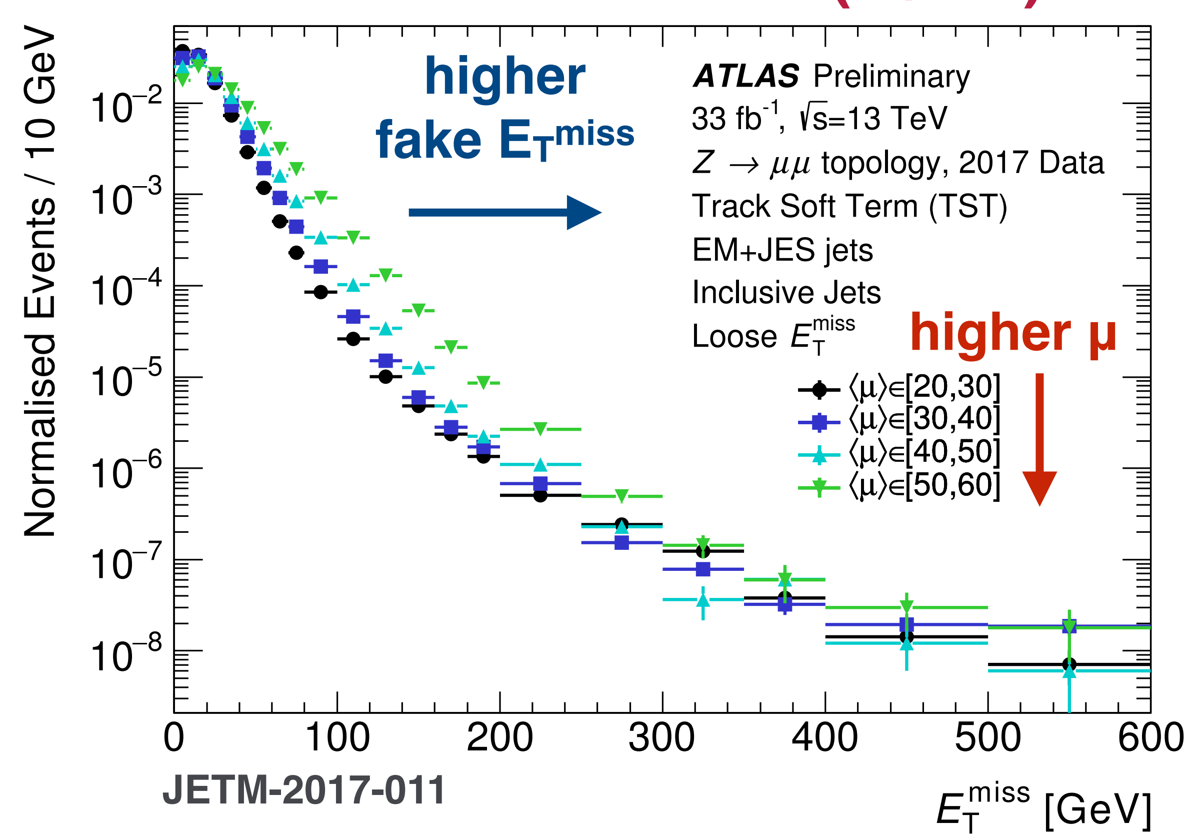
More jets vs # vertices (N_{PV})

Bias mass, resolution, & substructure

Bias event-level observables (E_T^{miss})



JETM-2016-012



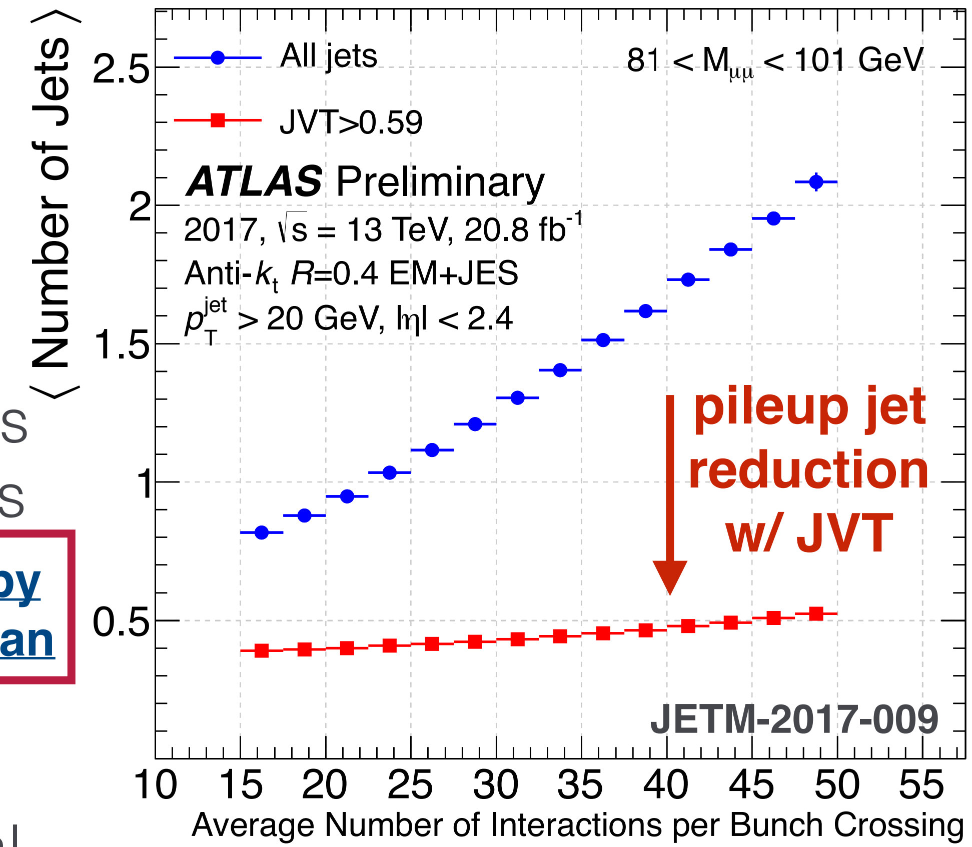
JETM-2017-011

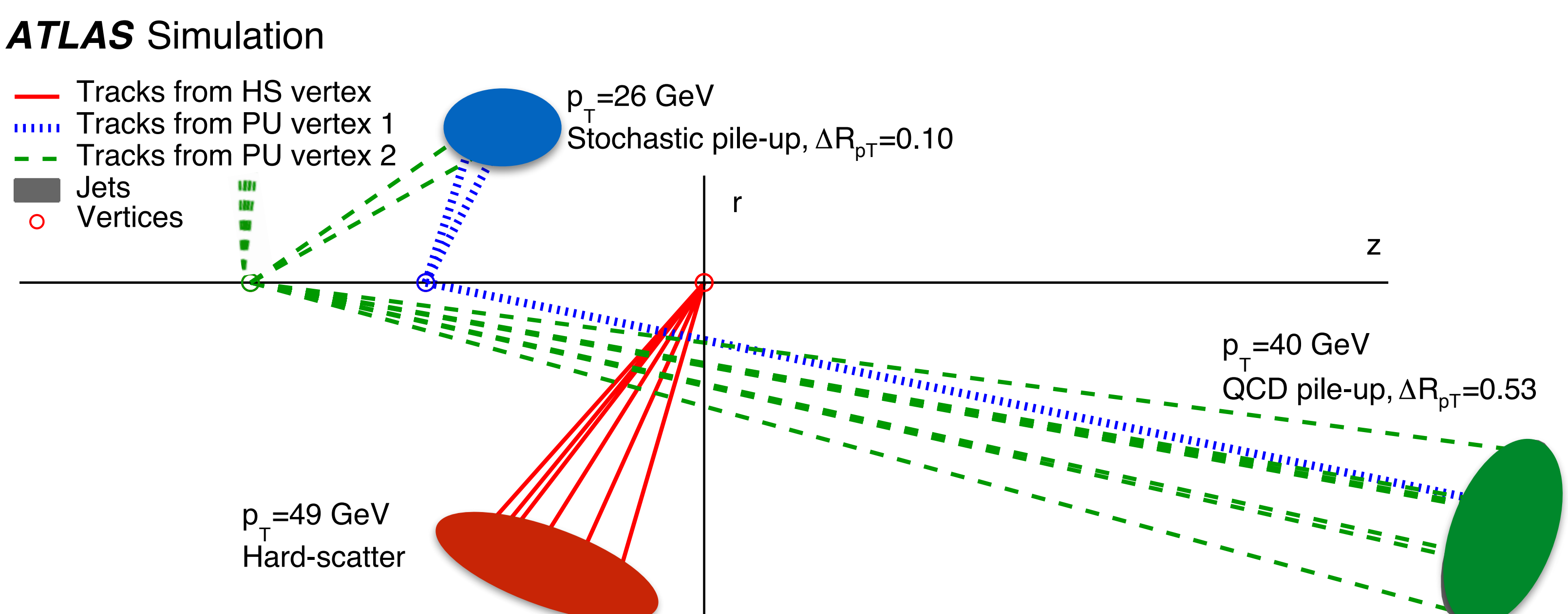
Jet Vertex Tagger

- **Multivariate JVT** connects jets to pileup vertices using **tracking** information
  - Large reduction in pileup jets within tracker (|η| < 2.4)
- HL-LHC: improve & **extend tracker to |η| = 4.0**
  - Proposed **High Granularity Timing Detector** (HGTD) aims for **30 ps timing resolution** to match jets to pileup vertices

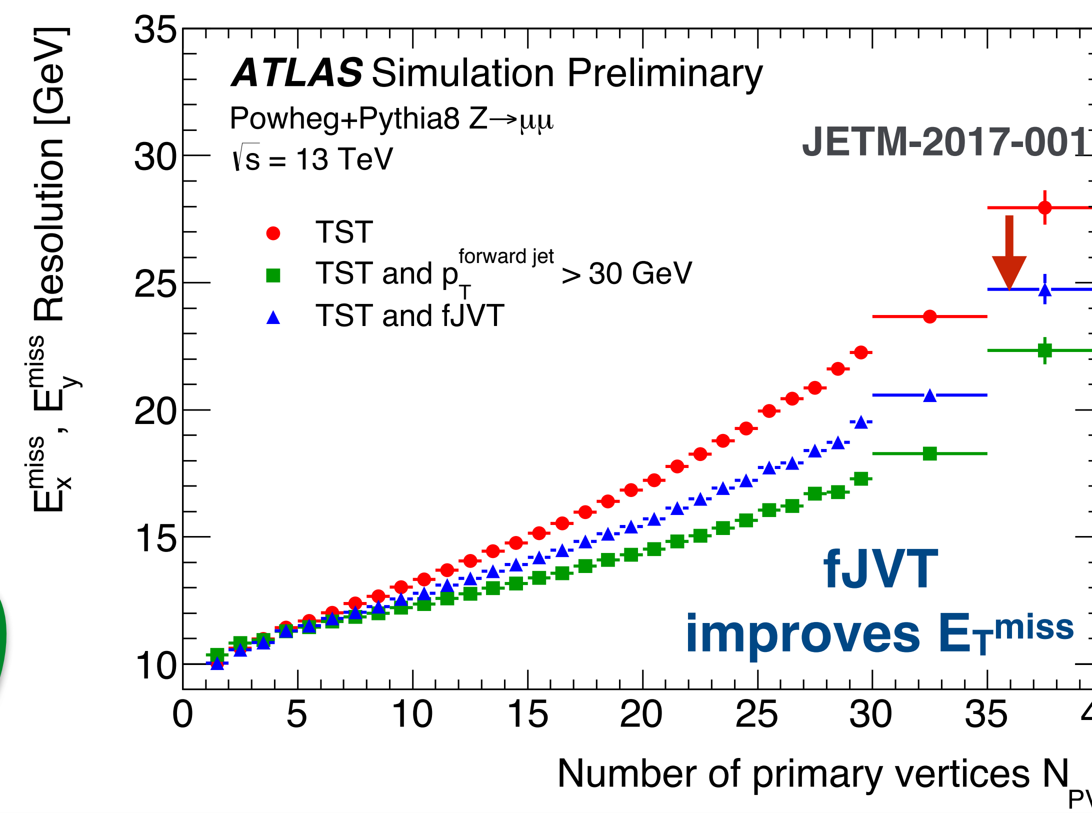
See HGTD talk by Ariel Schwartzman

- Beyond tracker use **Forward JVT**:
  - Remove **forward QCD jets** that balance a pileup vertex
  - Remove **stochastic jets** (from many vertices) using spatial and timing correlations of clusters



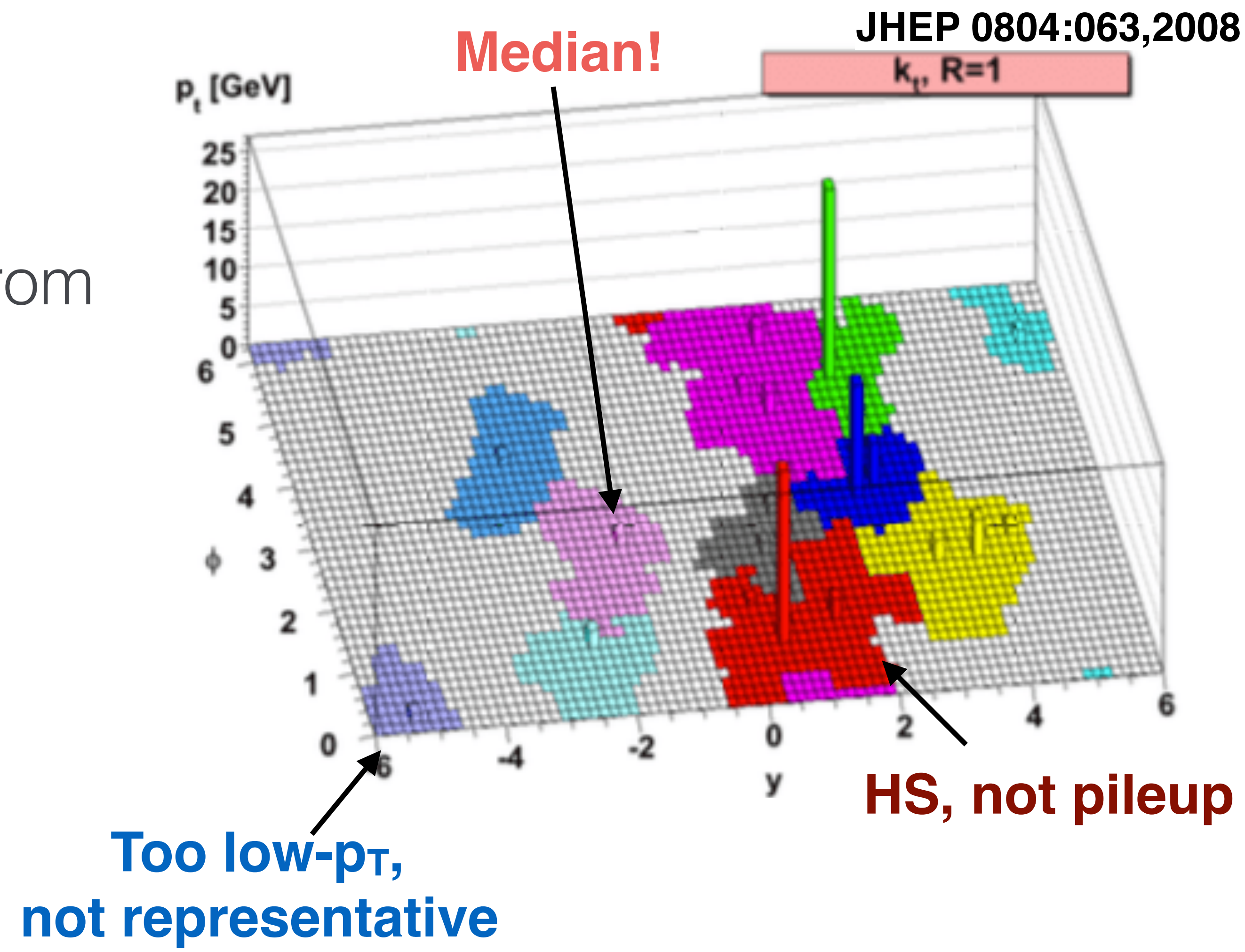


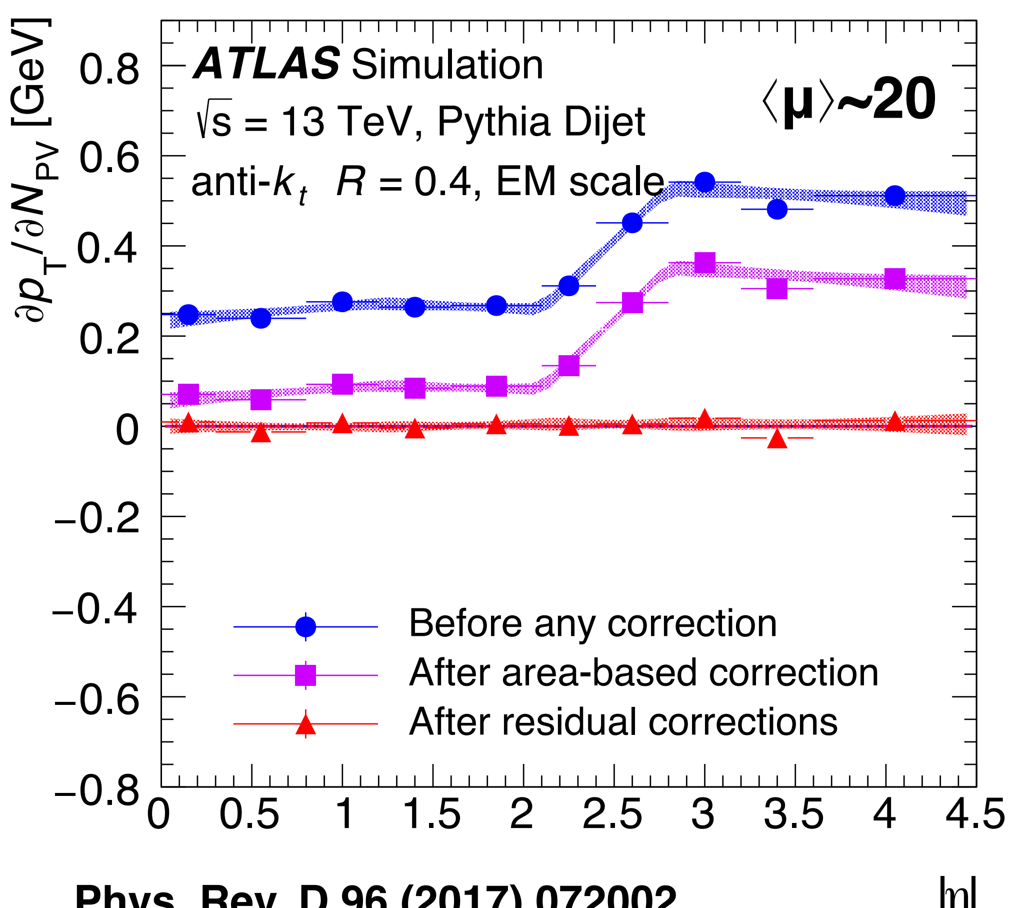
Eur. Phys. J. C 77 (2017) 580



Area-based subtraction

- Hard-scatter jets can include overlapping **pileup energy** & **pileup clusters**
- Event-wide ambient p_T density (**ρ**) taken from median p_T of (pileup sensitive) k_t jets
- Subtract **ρ** from each anti-k_t jet according to its area





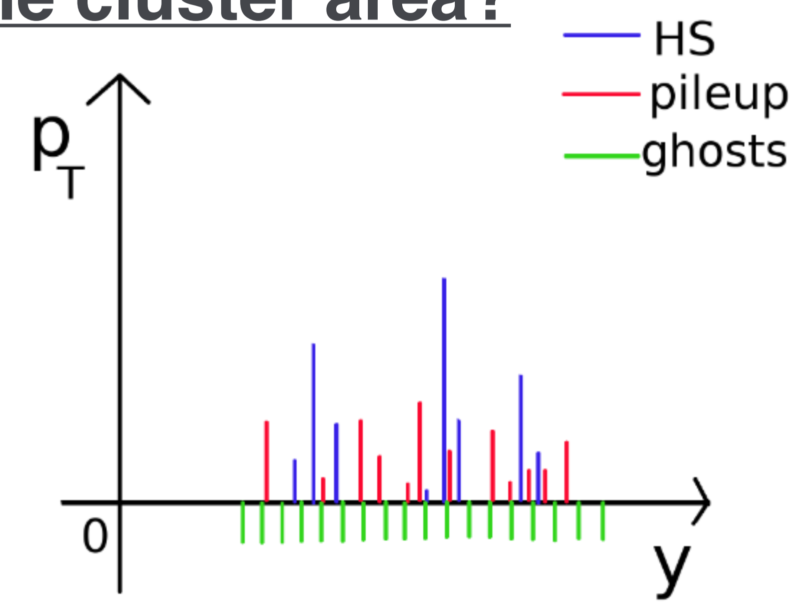
Phys. Rev. D 96 (2017) 072002

- **Residual correction vs NPV and μ** due to changing calorimeter geometry
- Works well currently, but difficult at HL-LHC

Cluster-level Subtraction

ATLAS-CONF-2017-065

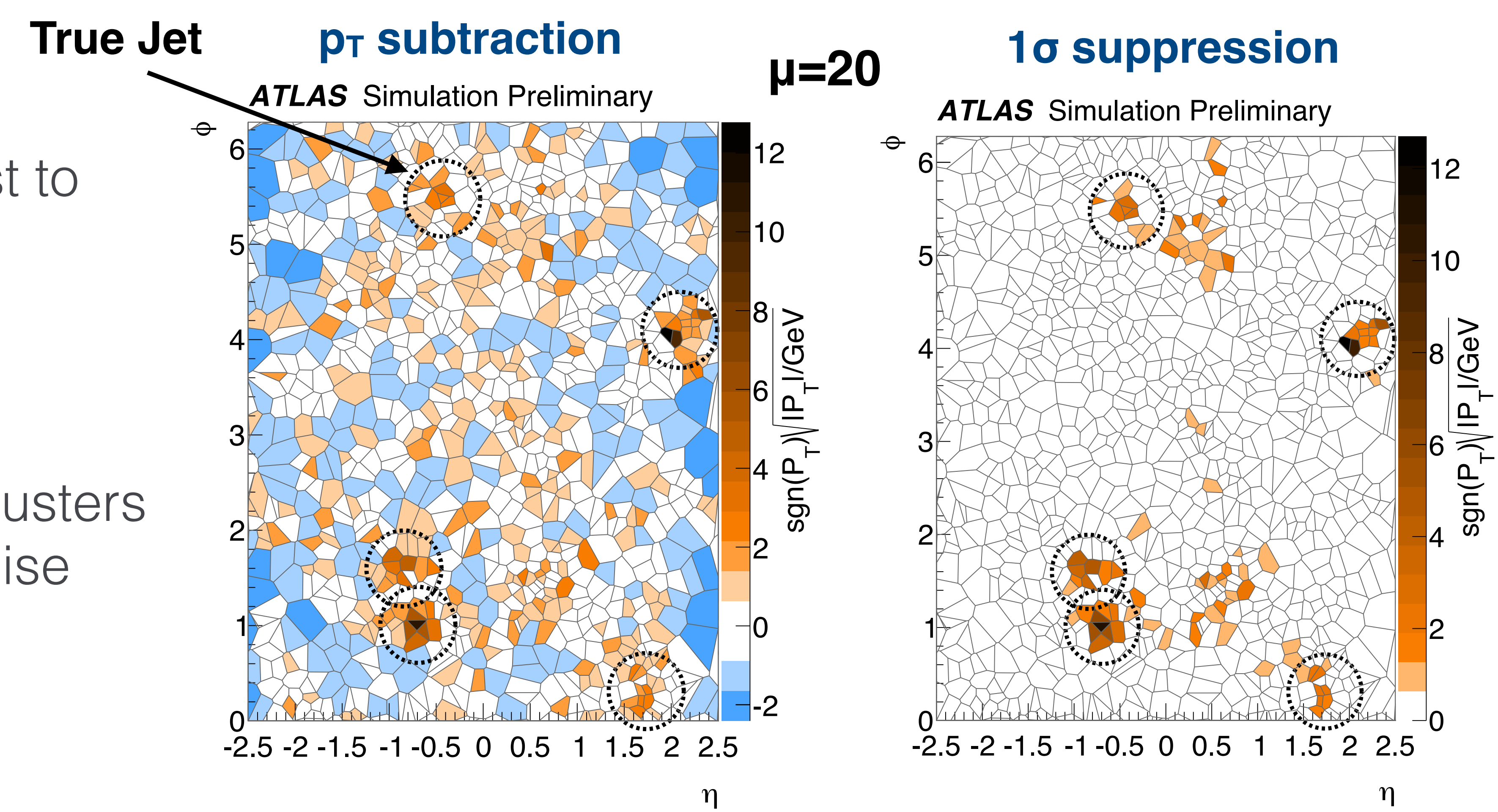
Translate area-based subtraction to clusters. **How to define cluster area?**



- **Constituent Subtraction:** Add fake "ghost particles" uniformly to event, and **cluster alongside cells**
  - Number of clustered ghosts ~ area
- Correct topoclusters according to N_{ghost} & event ρ (i.e. **give ghosts negative p_T**)

Clusters & Voronoi Area

- **Voronoi Area:** η-ϕ area closest to each cluster
- Subtract **ρ from each cluster** according to Voronoi area
- **1σ suppression:** remove all clusters with low significance above noise

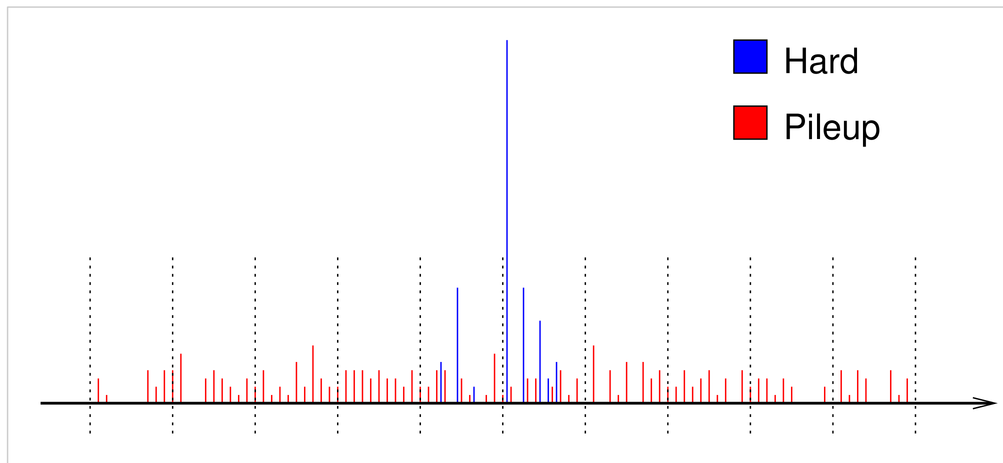


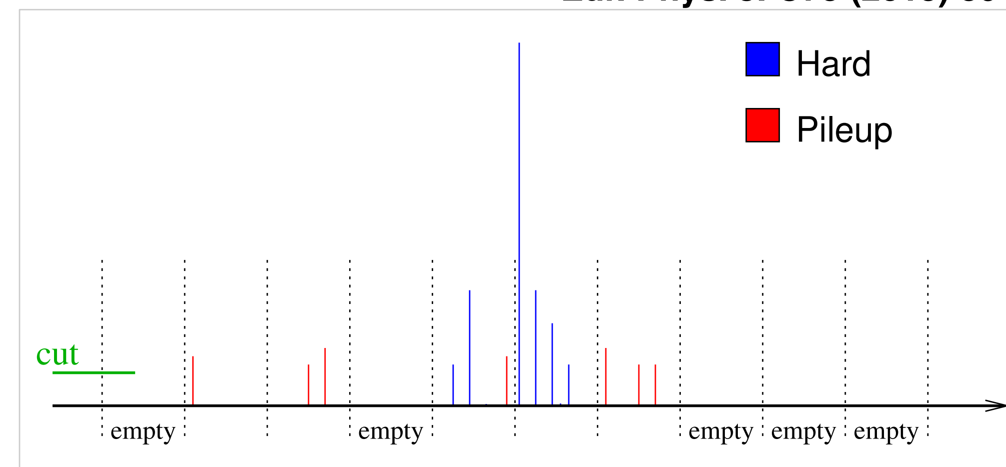
SoftKiller

ATLAS-CONF-2017-065

- SoftKiller **targets individual pileup clusters** surviving Constituent or Voronoi subtraction
- Reject all clusters below an event-specific **p_T cut**
- **p_T cut** chosen so **detector is half empty** for the event

Eur. Phys. J. C75 (2015) 59

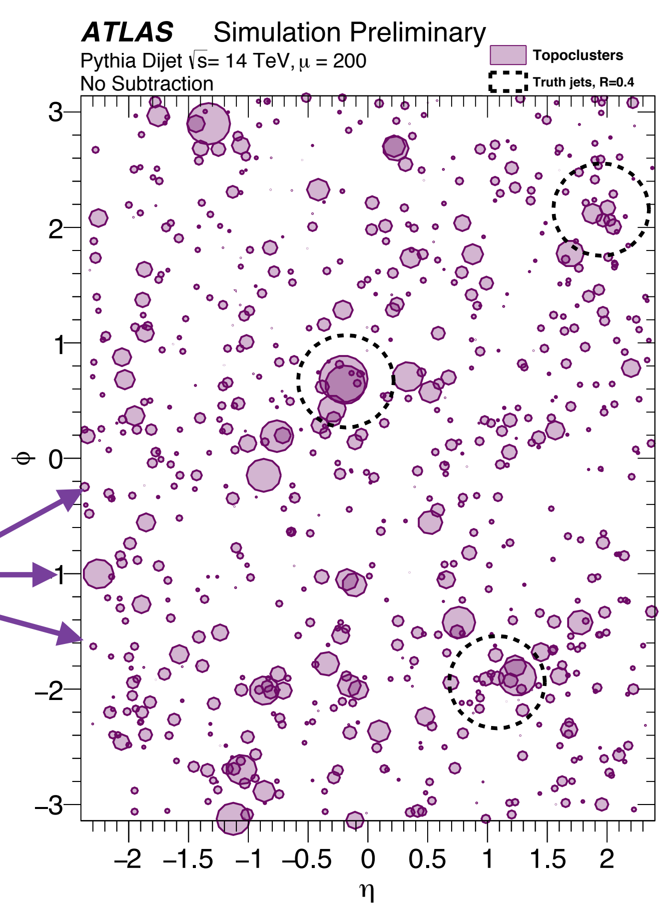




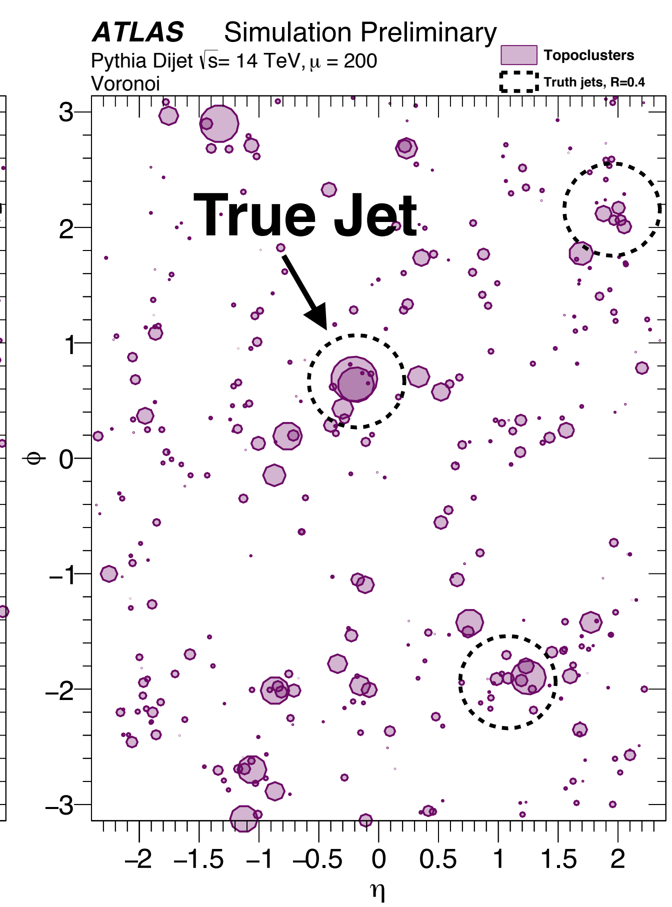
VS+SoftKiller in action:

Topoclusters

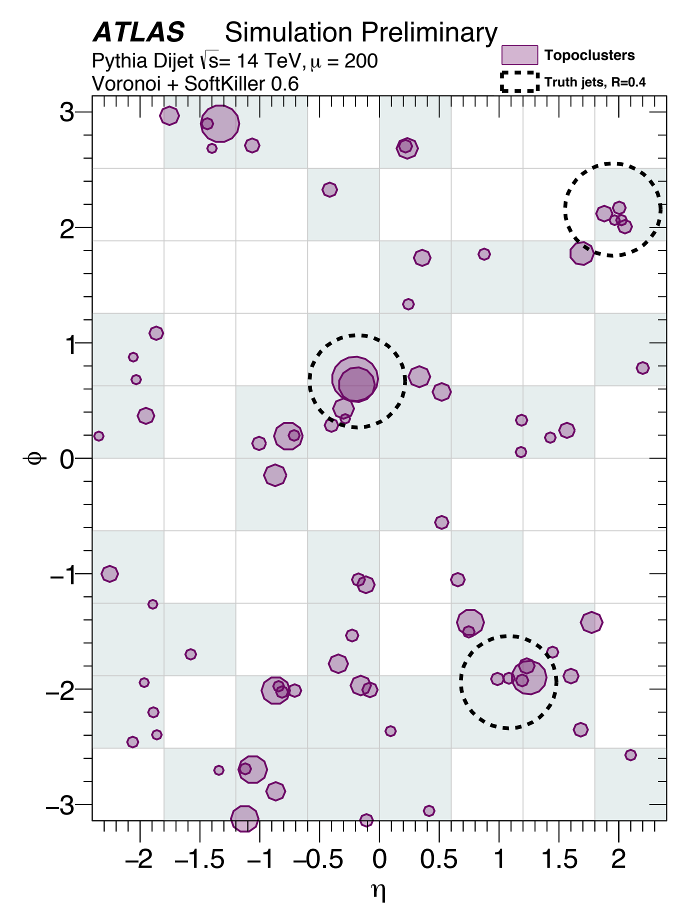
Event w/ pileup



Voronoi Subtraction



VS + SoftKiller



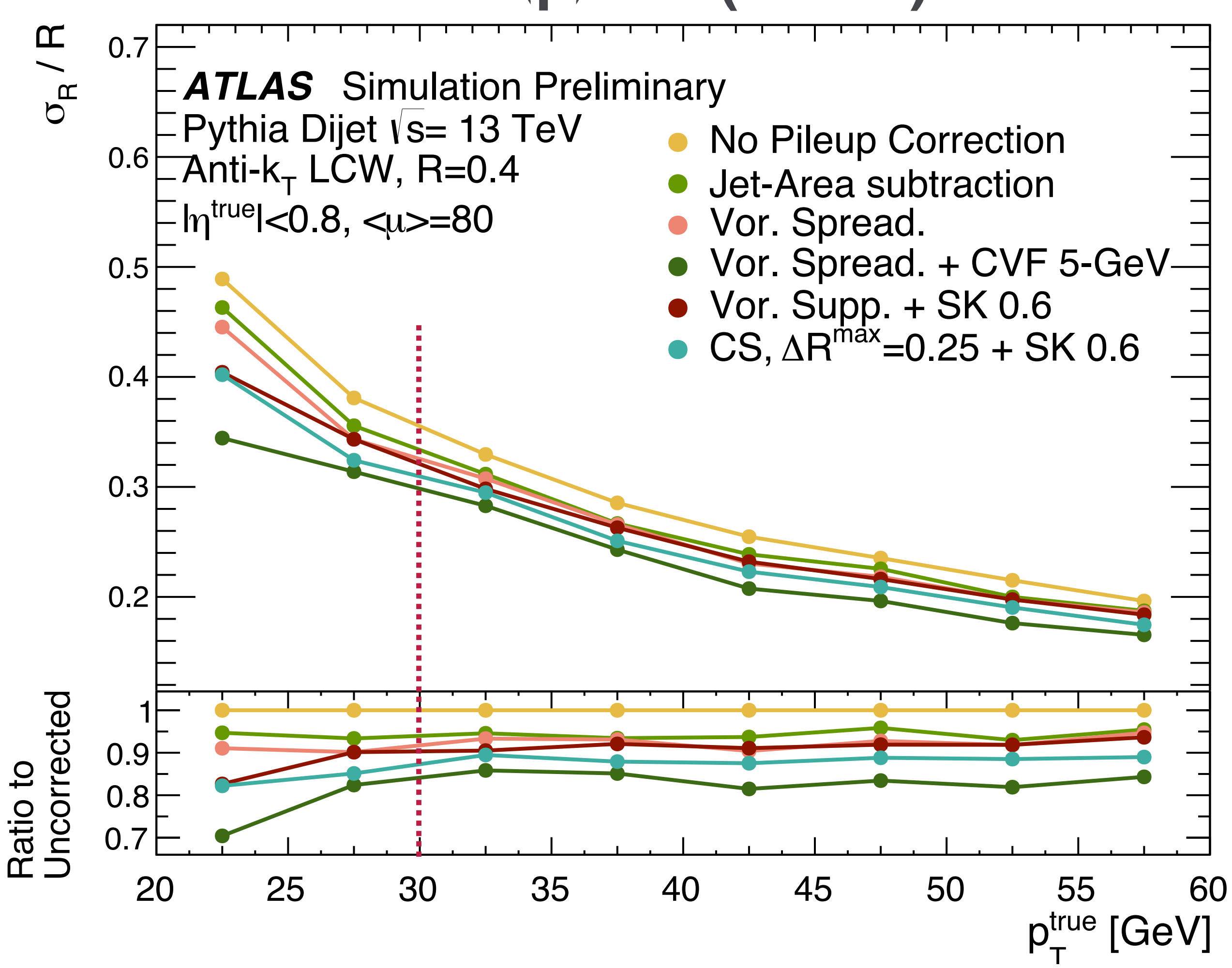
μ=200

Cluster-level Performance

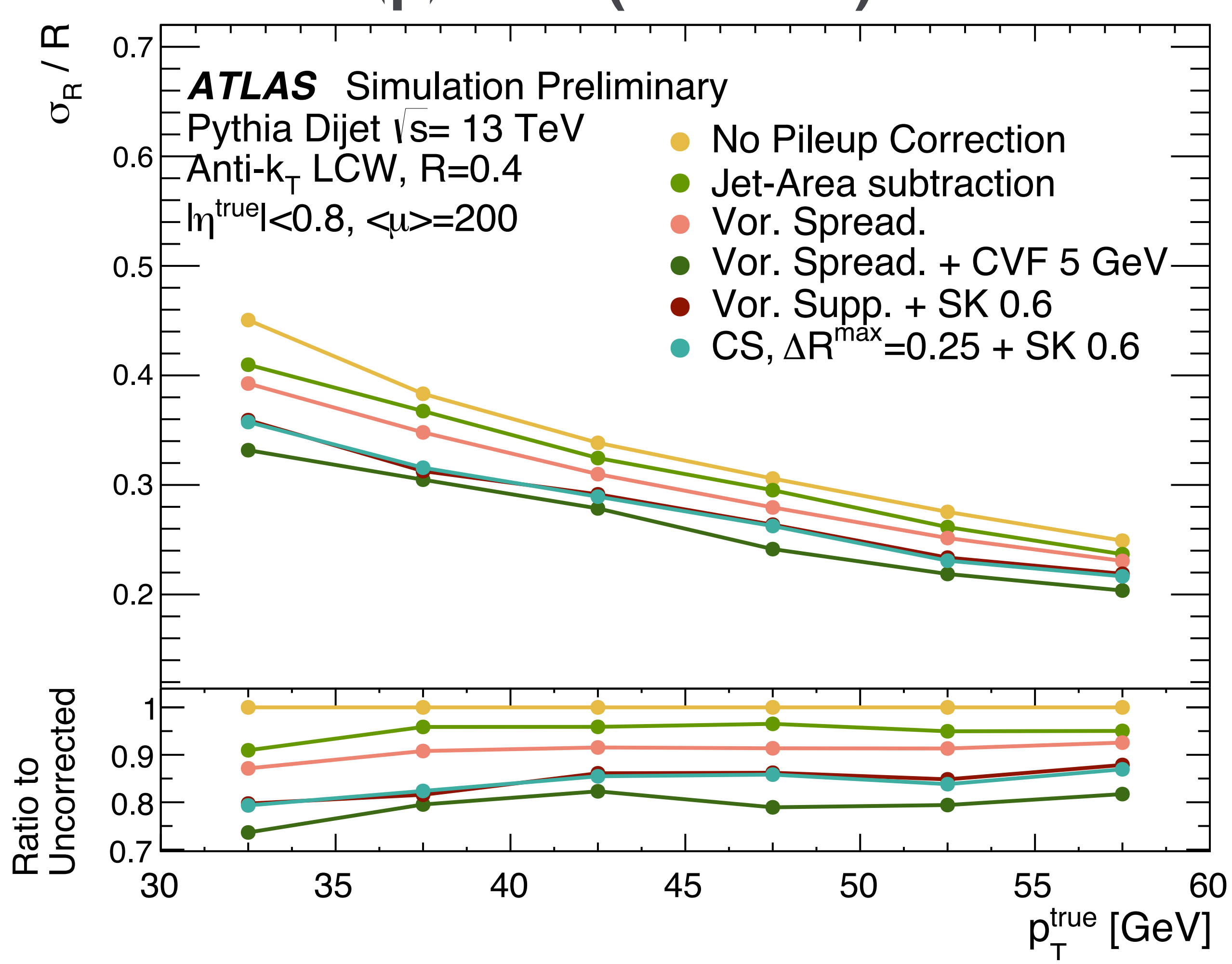
ATLAS-CONF-2017-065

- Large **reduction in jet energy resolution** compared to jet-area subtraction in high pileup conditions
- **Gains even at lower μ** → ongoing studies for use Run 2 & 3 data

⟨μ⟩=80 (Run 2)



⟨μ⟩=200 (HL-LHC)



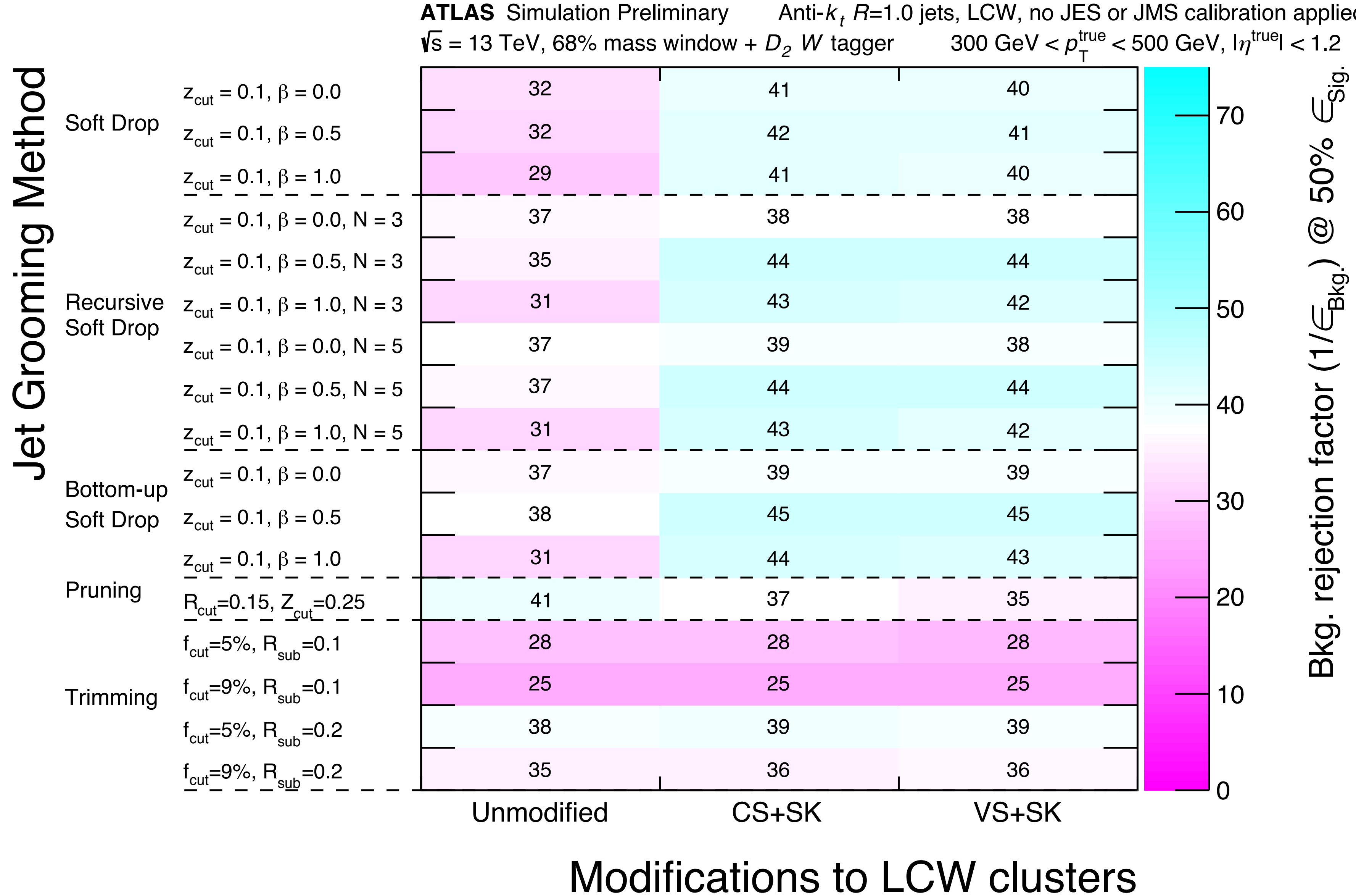
Jet Tagging Performance

ATL-PHYS-PUB-2017-020

- Correcting clusters also **improves jet substructure**
  - Essential for **tagging** boosted top, W, Z, & Higgs
- Scan over **clusters types** & **grooming methods** for best combination
  - Tagging efficiency, fake rejection, and mass stability w.r.t. pileup

⟨μ⟩~20

Background rejection of W tagger at 50% signal efficiency for various jet types

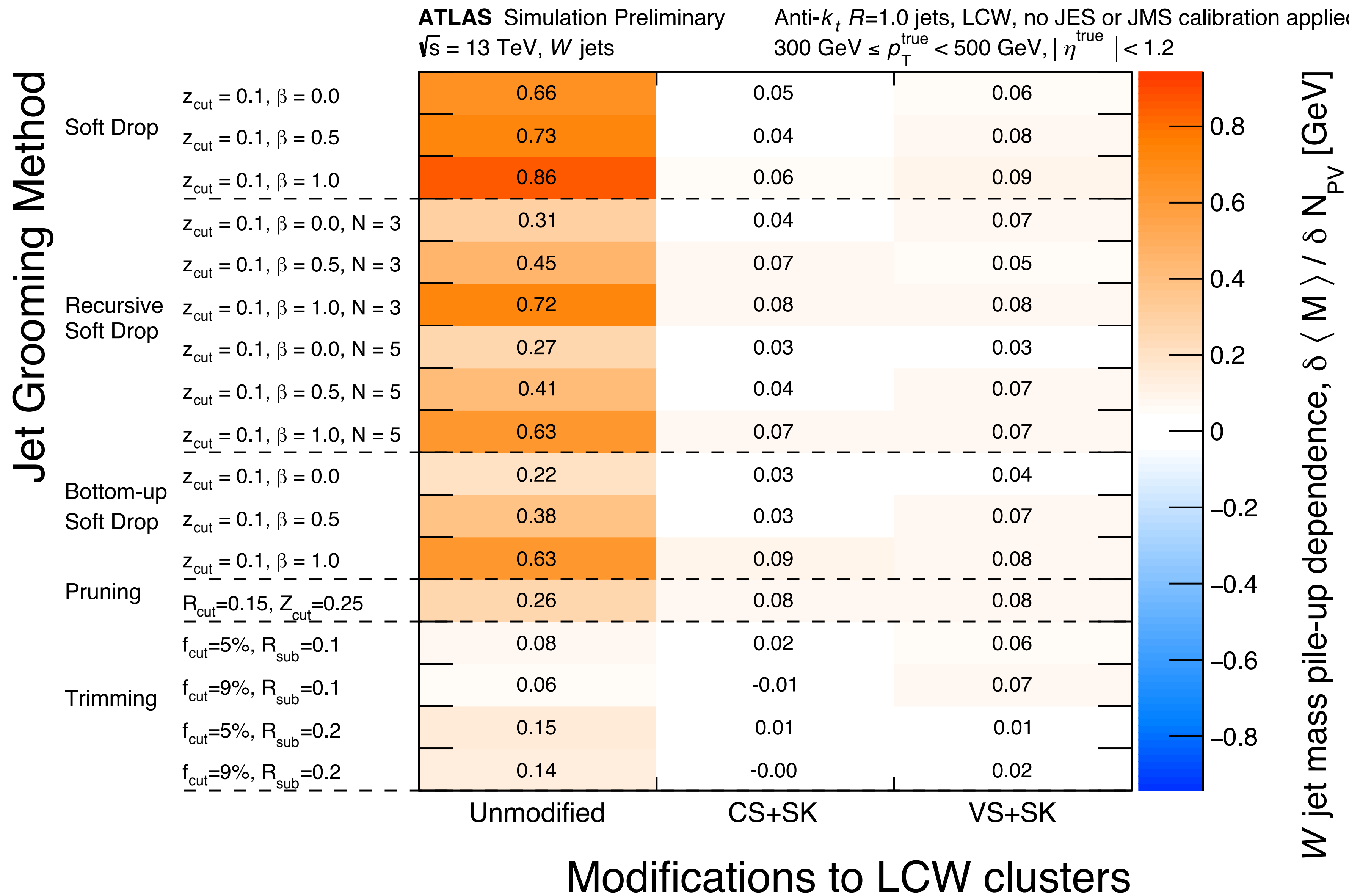


JETM-2018-003

- Good performance from **trimmed LCtopo** & **soft drop CS+SK / VS+SK jets**
  - Final algorithms will be calibrated with full *in situ* chain used for small-R jets

⟨μ⟩~20

W mass stability vs N_{PV} for various jet types



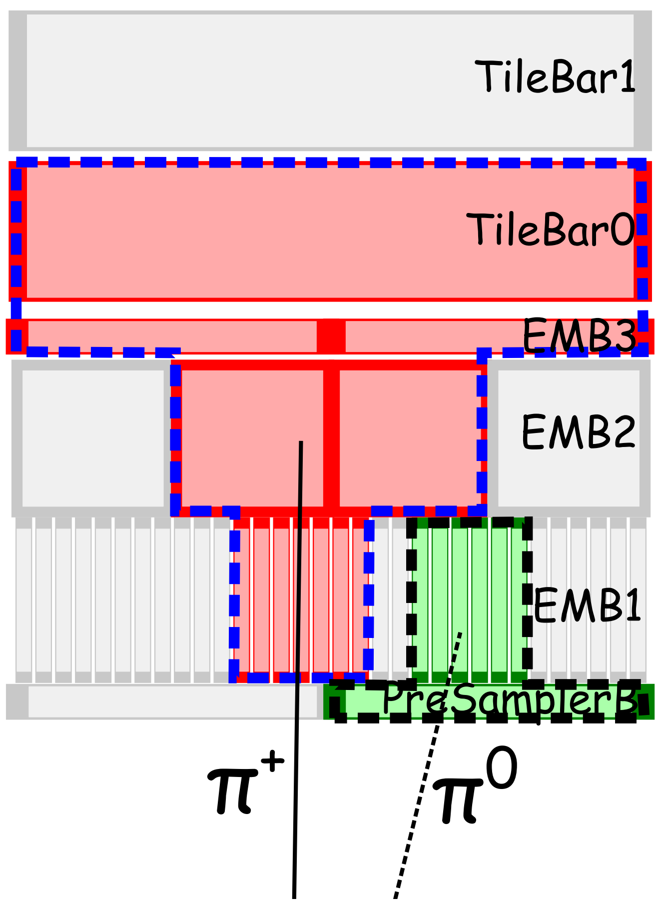
JETM-2018-003

- Correcting clusters also **improves jet substructure**
  - Essential for **tagging** boosted top, W, Z, & Higgs
- Scan over **clusters types** & **grooming methods** for best combination
  - Tagging efficiency, fake rejection, and mass stability w.r.t. pileup
- Good performance from **trimmed LCtopo** & **soft drop CS+SK / VS+SK jets**
  - Final algorithms will be calibrated with full *in situ* chain used for small-R jets

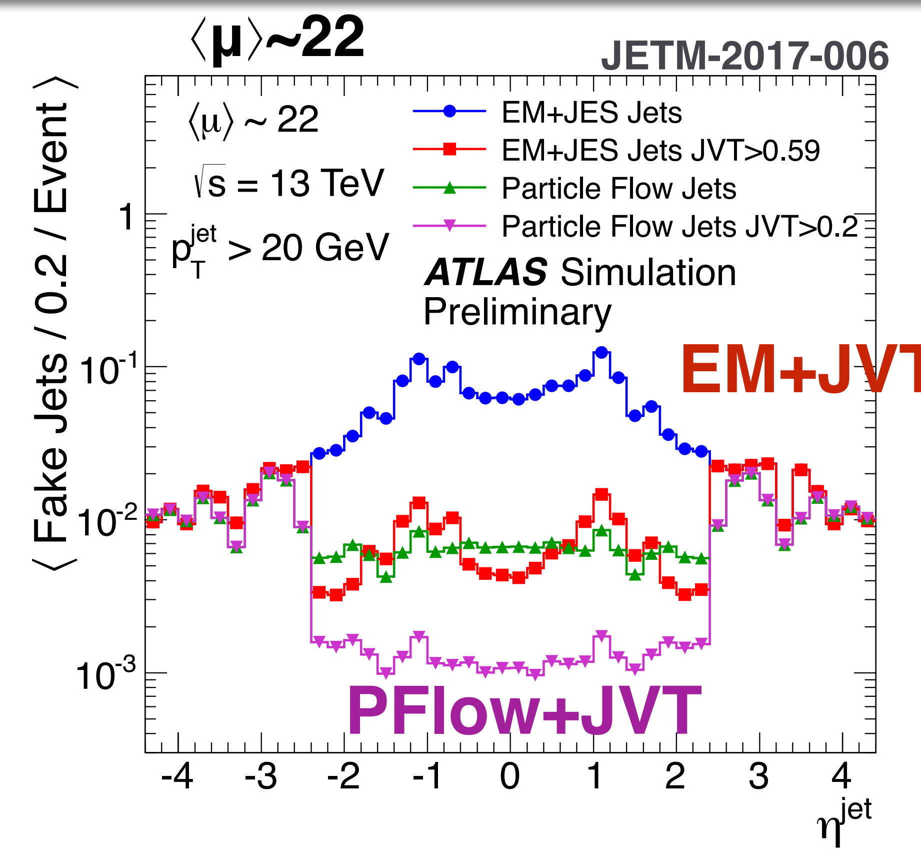
Particle flow jets

Eur. Phys. J. C 77 (2017) 466

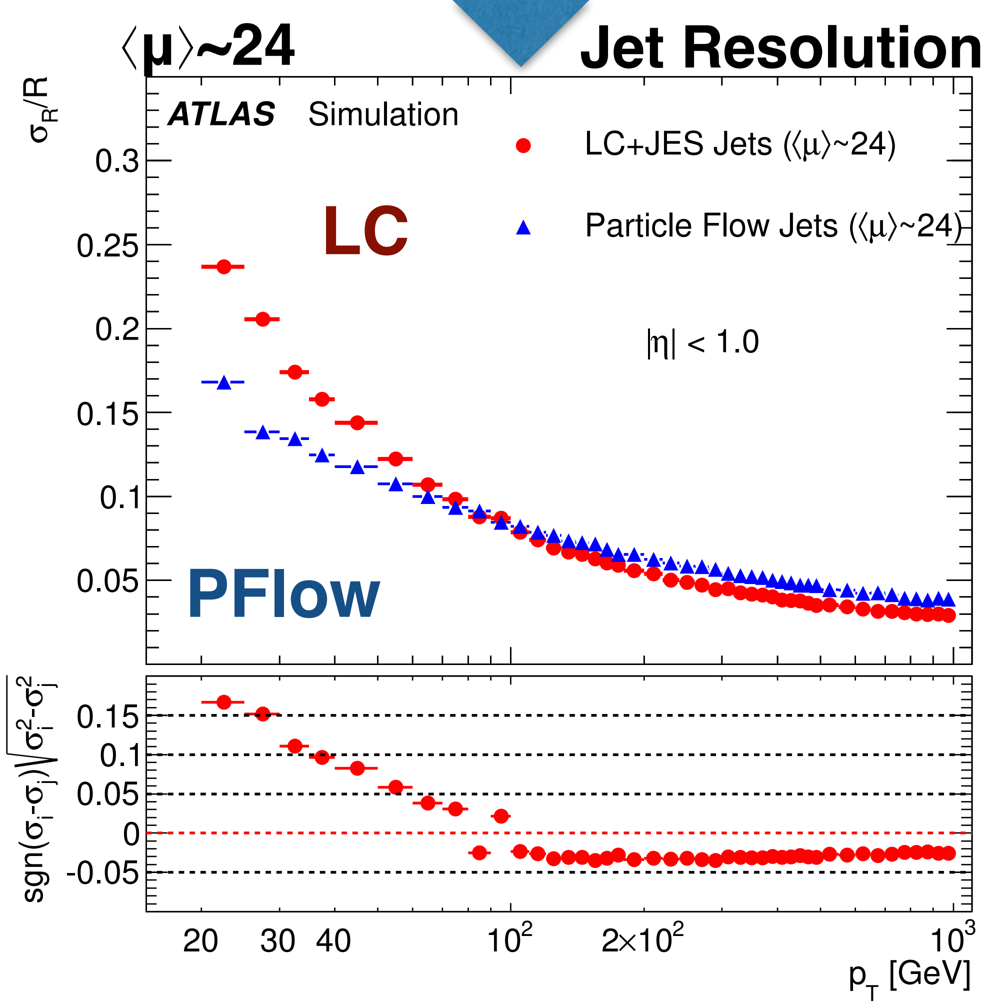
Match tracks to topoclusters, removing charged energy while keeping neutral component



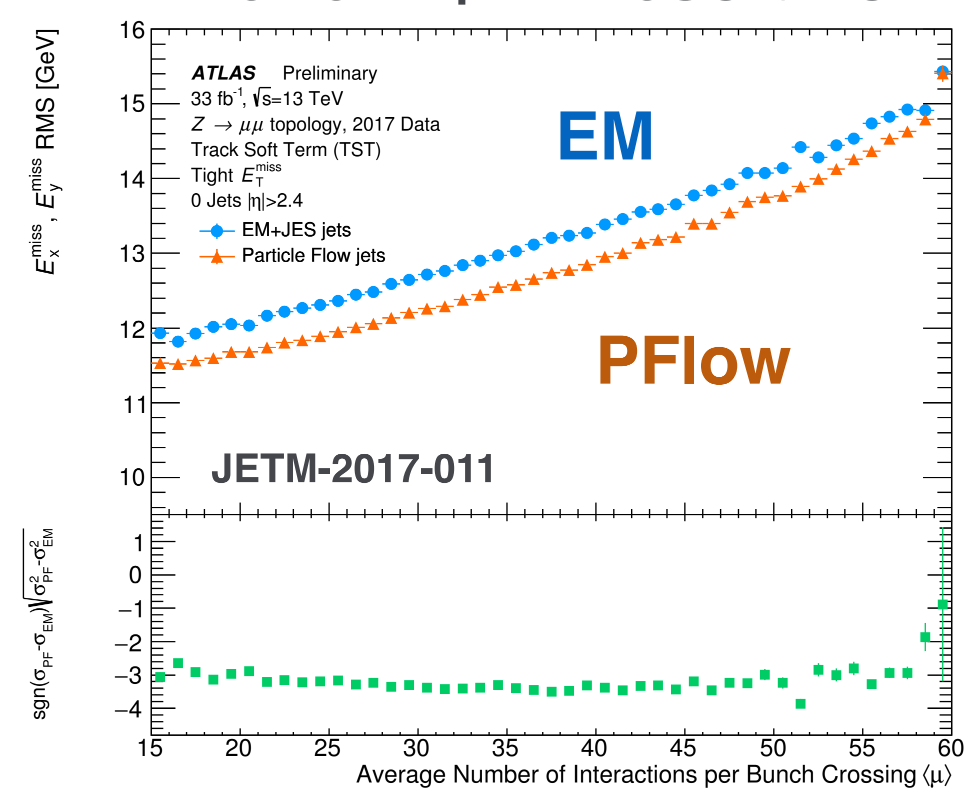
Topoclusters consistent with pileup tracks are rejected, **reducing pileup**



Improved energy resolution at low p_T, driven by accurate track measurements



Better E_T^{miss} resolution

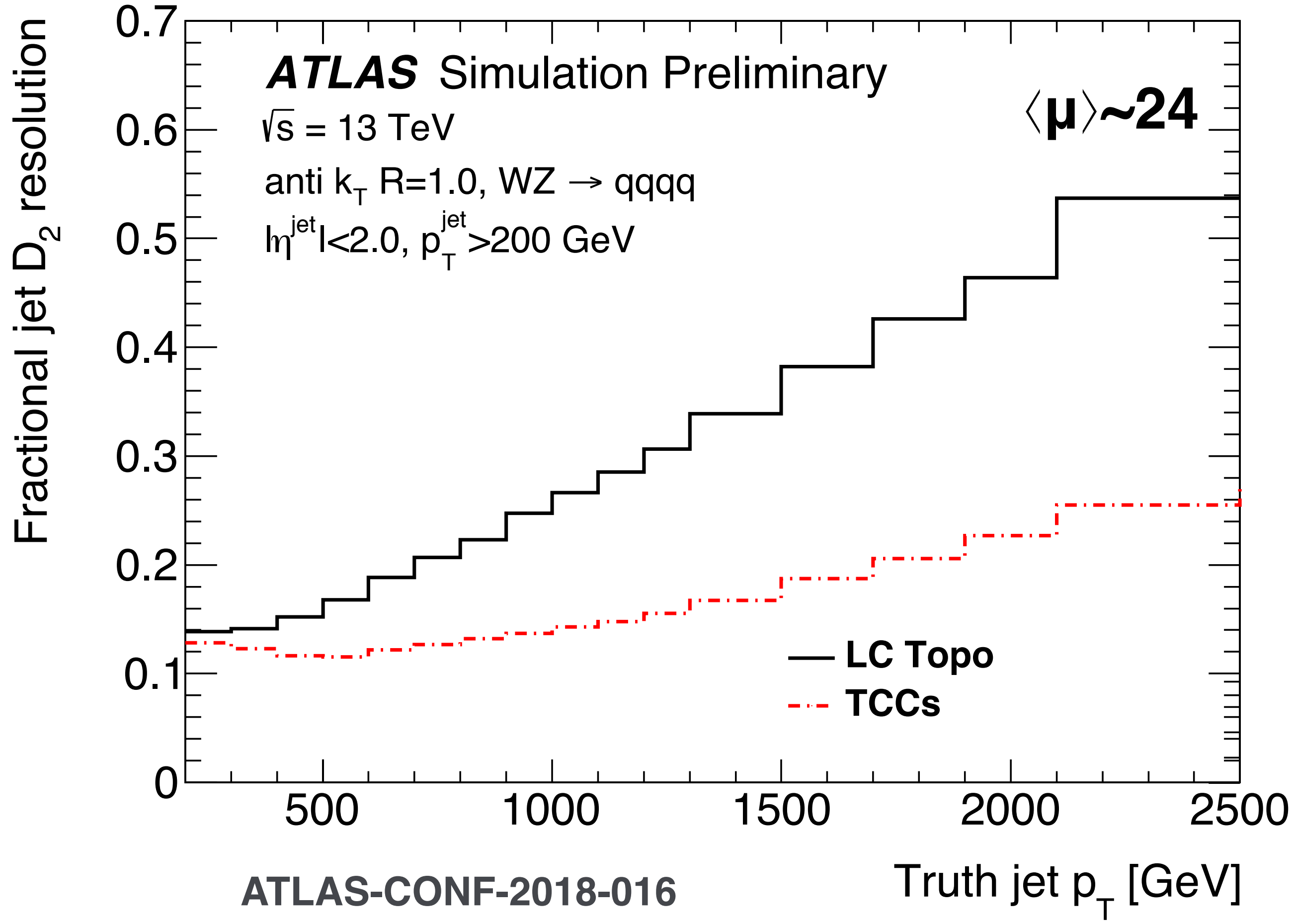


Track-Calo Clusters (TCC)

ATL-PHYS-PUB-2017-015

- At high-p_T, track p_T resolution degrades, but extrapolated **angular resolution improves**
- TCC uses tracks to **correct spatial resolution** of coarser calorimeter clusters, not their energy
  - Retains benefits of **pileup vertex rejection**
  - Large improvement to **substructure variables** (like D_2), **benefiting taggers**
  - Robust against pileup

TCC jets improve D_2 resolution



ATLAS-CONF-2018-016

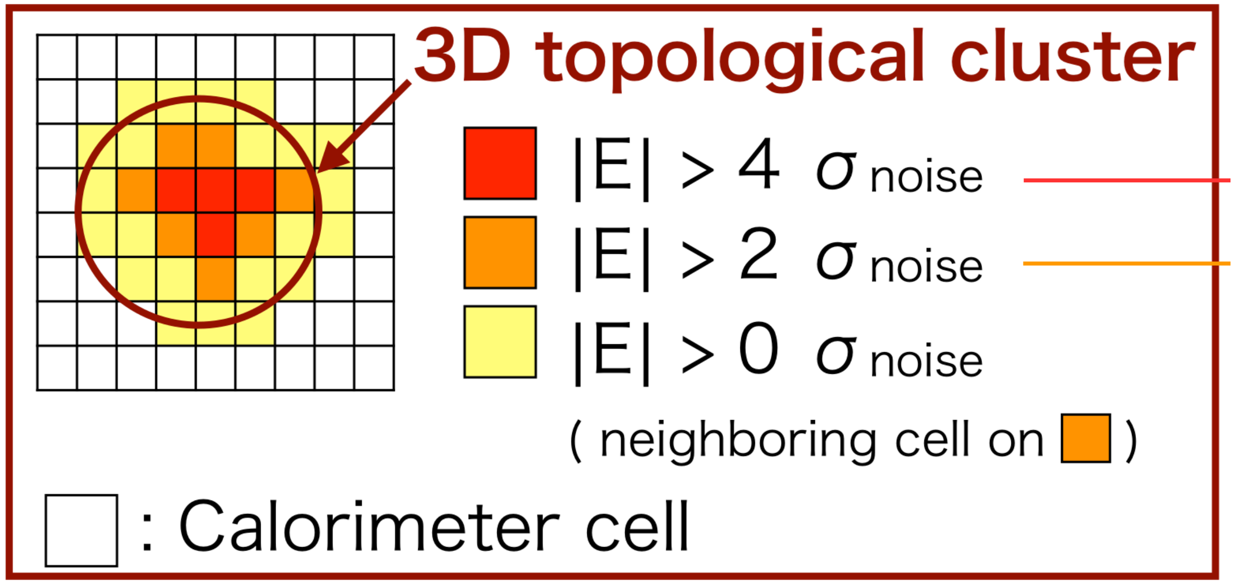
Conclusion

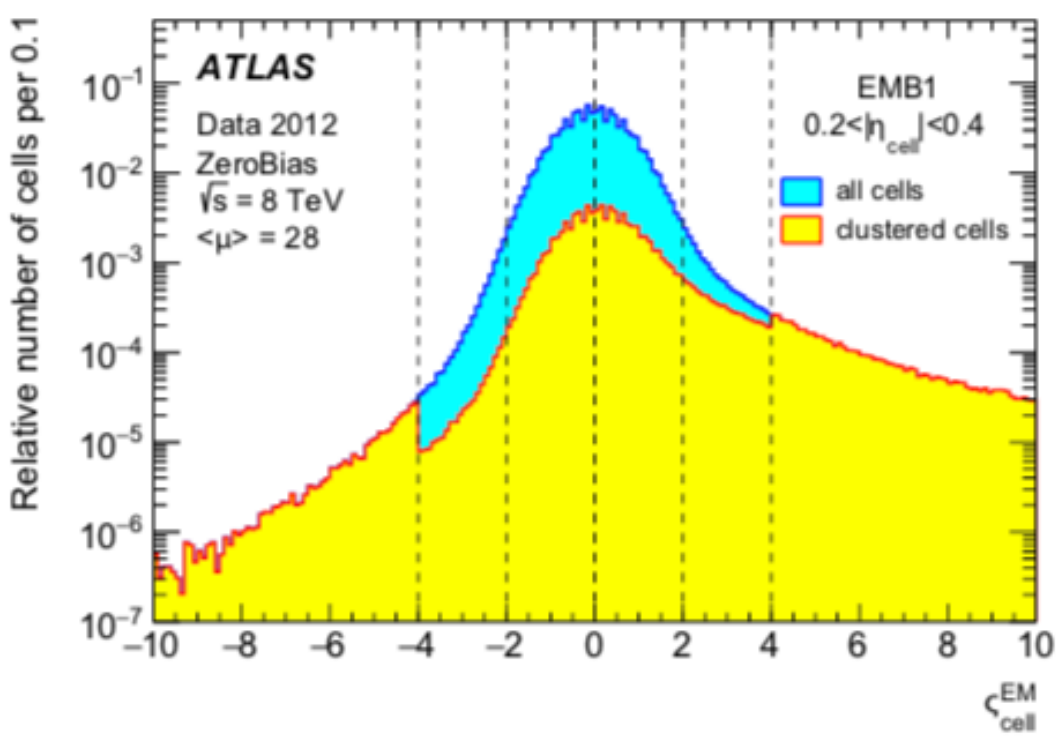
- **High pileup environment** now & at the HL-LHC offers challenges for jet calibration & tagging
- **Various new techniques** for mitigating pileup impact on jet measurements, E_T^{miss}, and identification of high-p_T jets (boson & top tagging)
  - Great deal of experimentation - significant effort by software team to ease implementation
  - Plan to converge in Run 3 on optimal combination of techniques
- **Significant HL-LHC upgrades** will improve track-based pileup tagging

Backup

Jet Reconstruction

- Inputs to jets are clusters (collections of neighboring calorimeter cells)
- **Inherent noise suppression** from 4-2-0 clustering algorithm:
  - Low energy pileup rejected
- **Anti-k$_T$** jet-finding algorithm focuses on hardest energy deposits, w/ reduced shaping by pileup
- However:
  - Higher-pt pileup jets still get through
  - Selected clusters are still affected by pileup

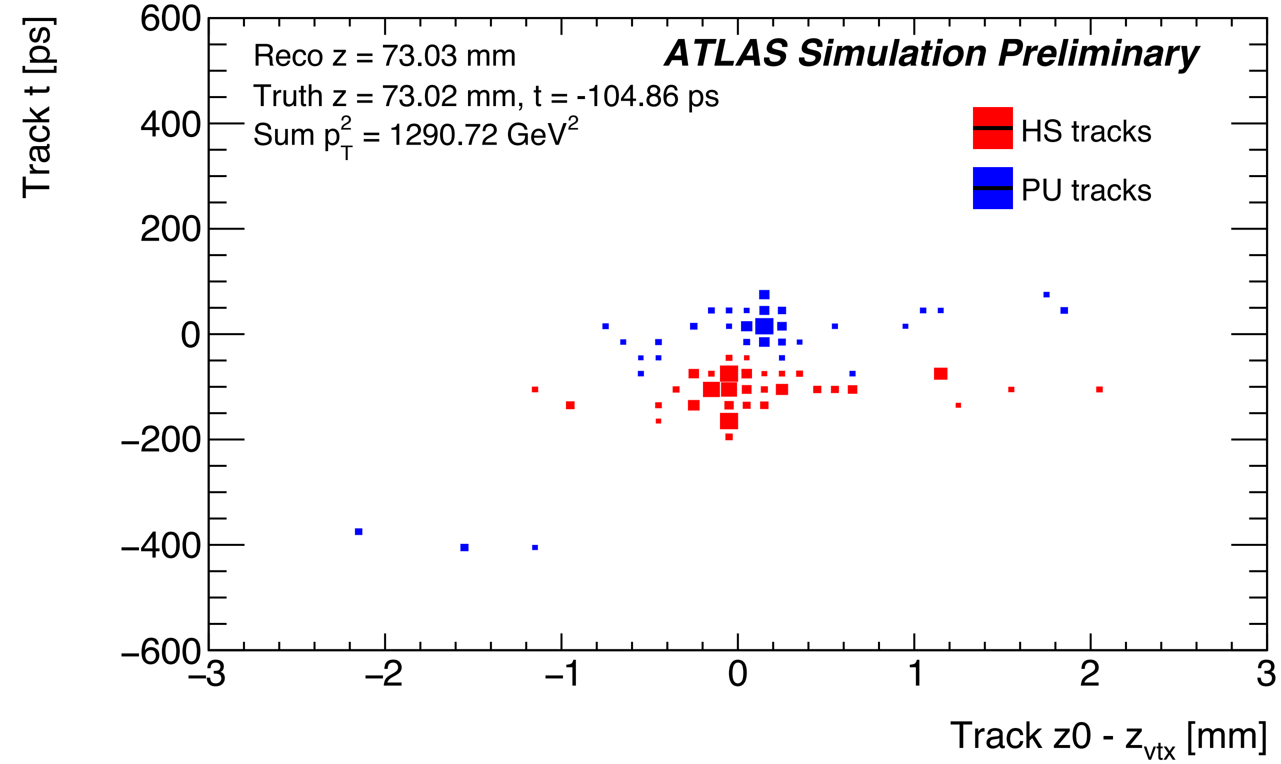


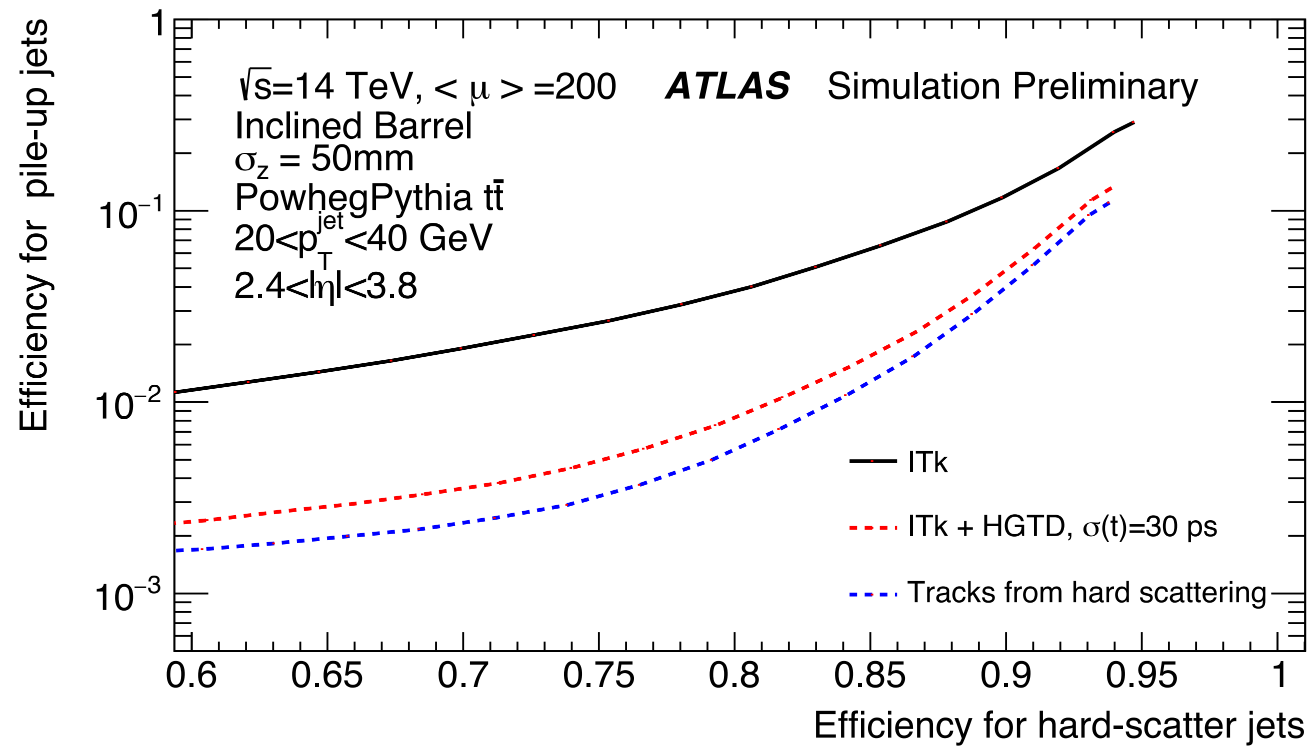


Most **low-significance cells** removed by clustering

High granularity timing detector

- HL-LHC will see <mu>=200, with ~1.8 vertex per mm
  - Impossible to distinguish pileup vs hard-scatter tracks via geometry only
  - Within a bunch crossing, collisions occur with σ_t = 180 ps
- HGTD can resolve track time within 30 ps
- Large reduction in tracks from pileup vertices close to hard scatter





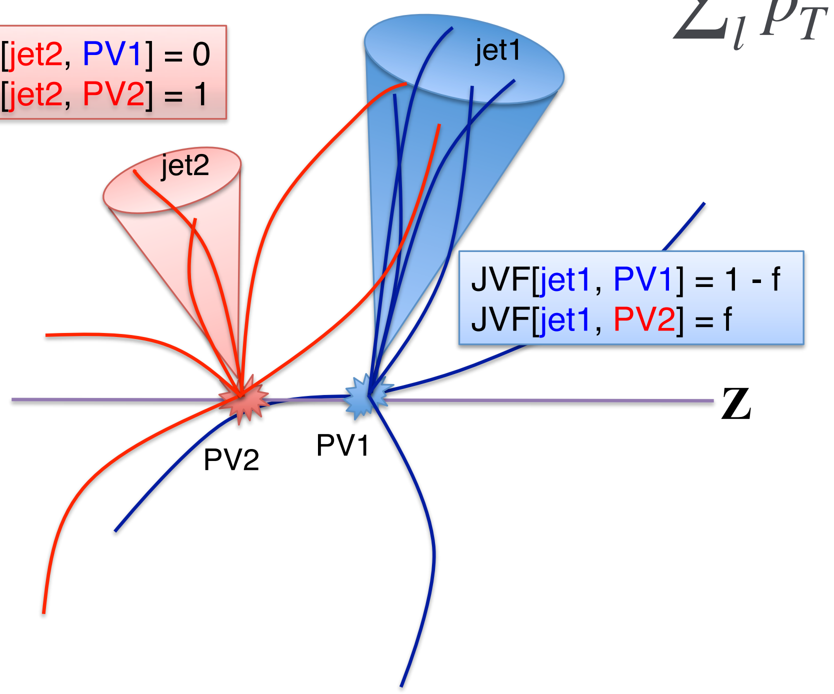
Jet Vertex Tagger (JVT)

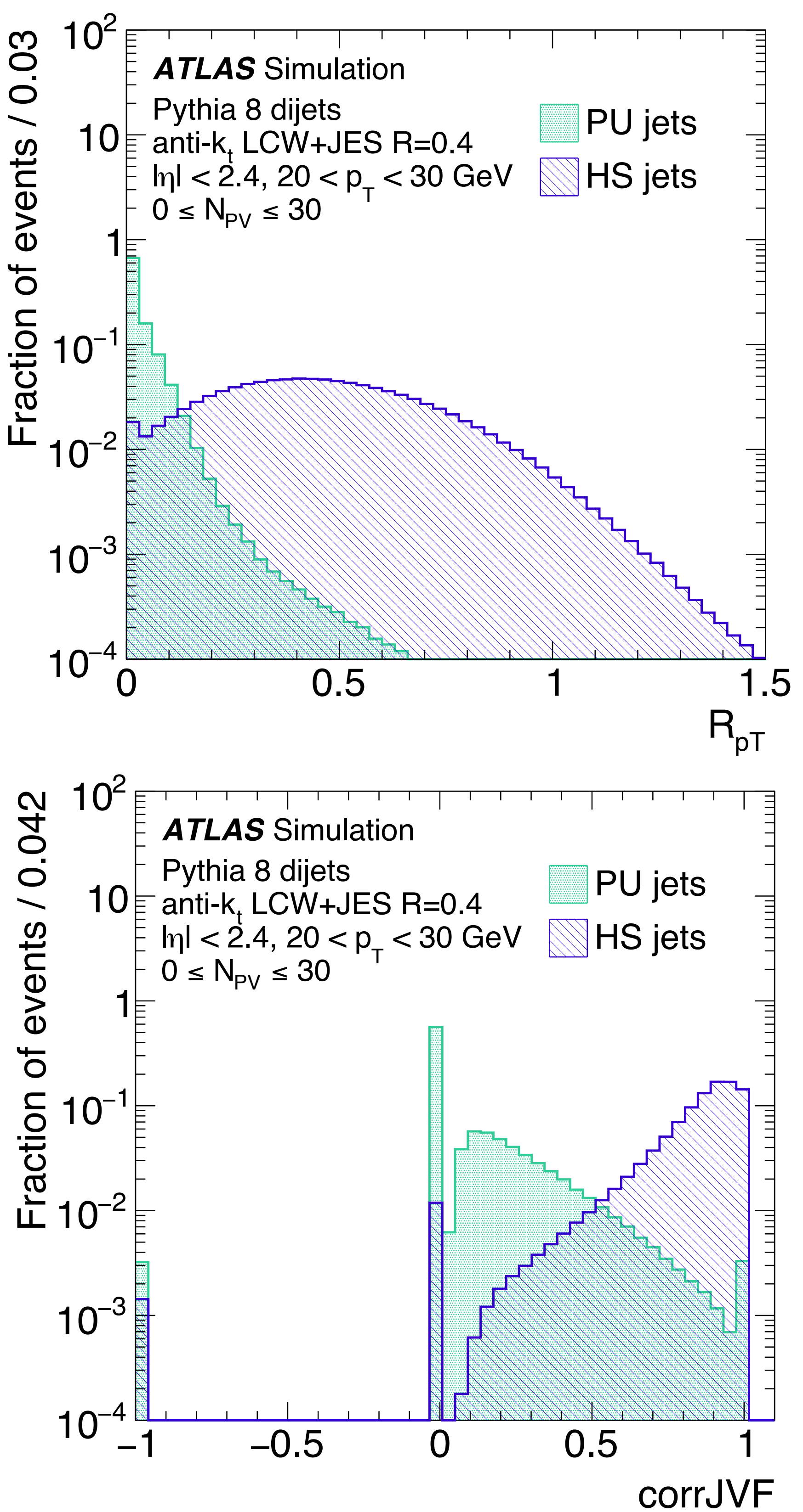
Eur. Phys. J. C (2016) 76:581

$$R_{pT} = \frac{\sum_k p_T^{trk_k}(PV_0)}{p_T^{jet}}$$

- Multivariate using R_{pT} and corrJVF
- R_{pT} is ratio of jet's p_T matched to hard scatter tracks
- corrJVF compare fraction of hard scatter tracks against pileup tracks
  - Corrected by # pileup tracks to remove N_{PV} dependence

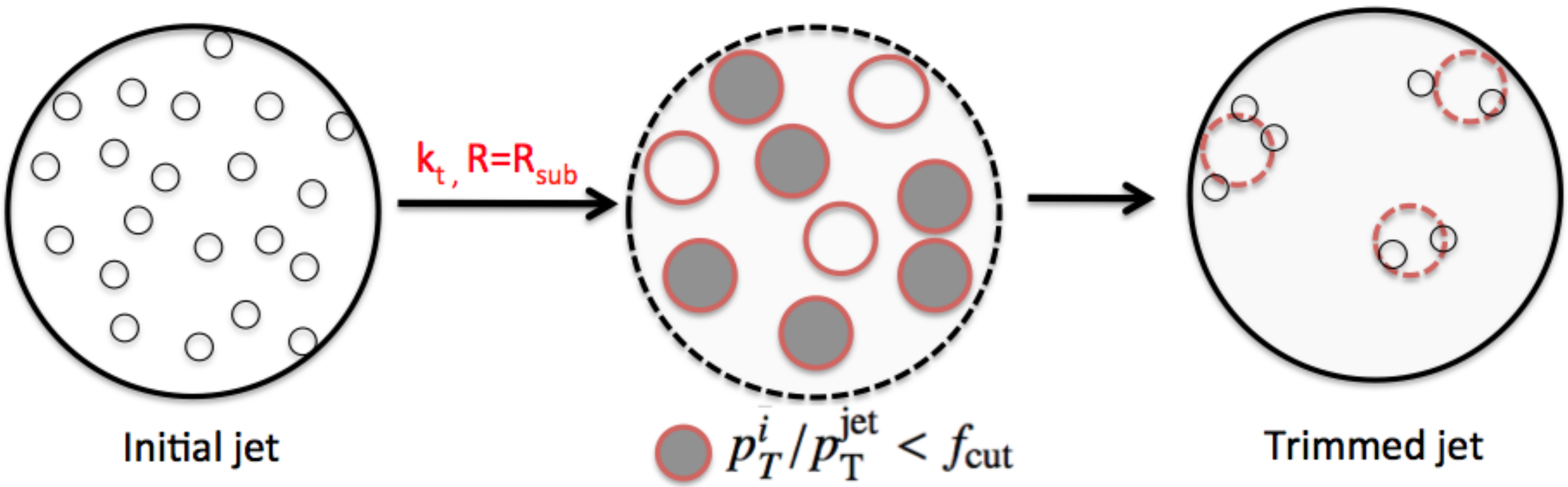
$$corrJVF = \frac{\sum_k p_T^{trk_k}(PV_0)}{\sum_l p_T^{trk_l}(PV_0) + \frac{\sum_{n\geq 1}\sum_l p_T^{trk_l}(PV_n)}{(k \cdot n_{trk}^{PU})}}$$





Jet Grooming: Trimming

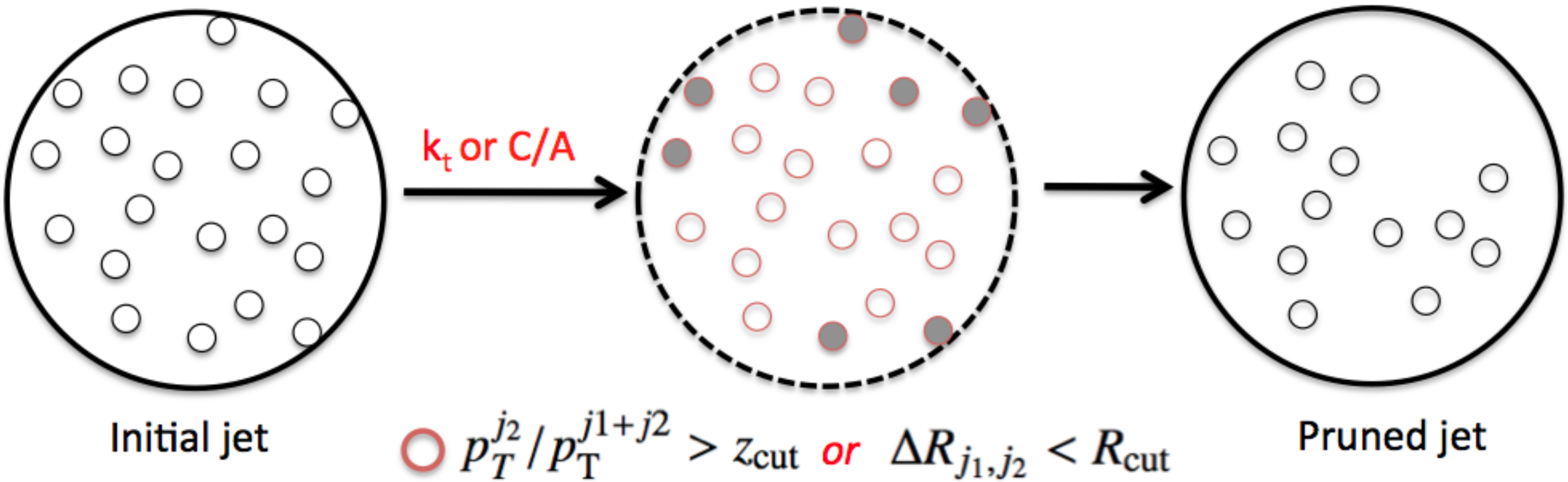
- ATLAS standard grooming procedure
- Target softer radiation from pileup, MPI, & ISR
- Recluster constituents into small-R sub-jets $R_{sub} \sim 0.2$
- Remove sub-jets with fractional $p_T < f_{cut} \sim 3\%$



JHEP 09 (2013) 029

Jet Grooming: Pruning

- Remove soft low-p_T clusters, but keep large-angle radiation
- Redo jet clustering (C/A or k_t), and at each stage cluster if either:
  - Not soft: p_T fraction of second constituent is > z_{cut}
  - Close-by: $\Delta R_{1,2} < R_{cut}$
- Otherwise, reject 2nd constituent



JHEP 09 (2013) 029

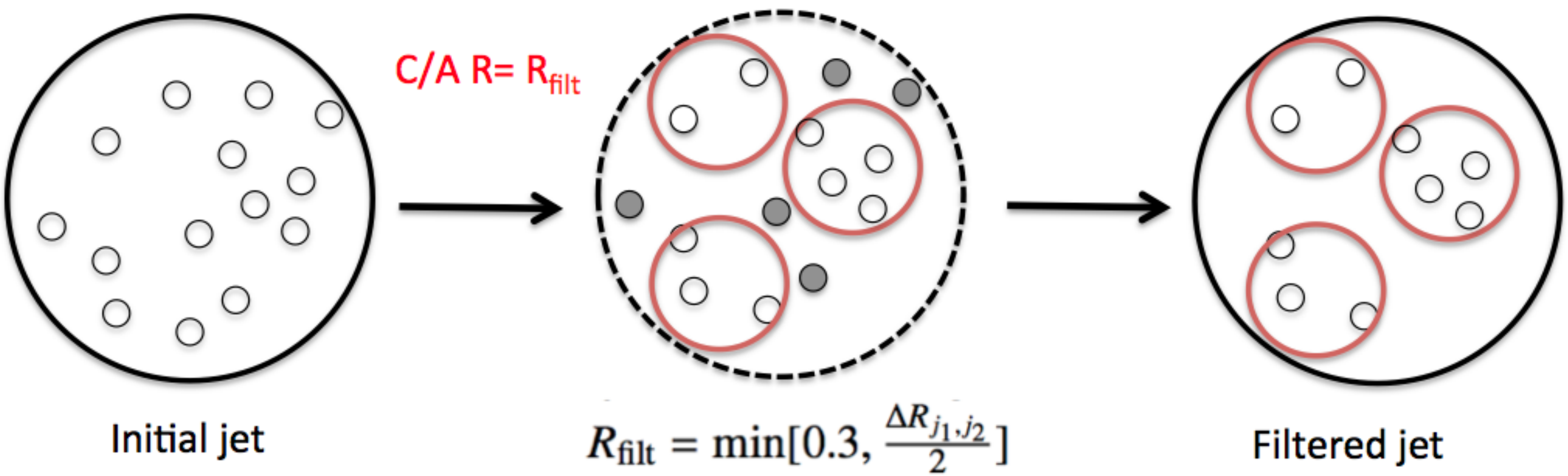
Jet Grooming: Modified Mass Drop + Filtering

- Iterative declustering of a C-A jet targeting **soft and wide-angle radiation**
  - Remove branches with p_T imbalance, provided no large drop in mass
- Requirements on 2 subjets from last clustering stage:

$$\text{(i - Mass Drop)} : m(j_1) < \mu \times m(j_2)$$

$$\text{(ii - Balanced Splitting)} : \frac{min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} > z_{cut}$$

  - If either fails, remove softest jet j_2, and continue procedure
  - If both pass, end procedure and keep jet
- **Filtering**: Recluster constituents into 3 C-A jets of radius $\mathbf{R_{filt}}$ (discard extra clusters)



$$R_{\text{filt}} = \min[0.3, \frac{\Delta R_{j_1,j_2}}{2}]$$

JHEP 09 (2013) 029

Jet Grooming: Soft drop

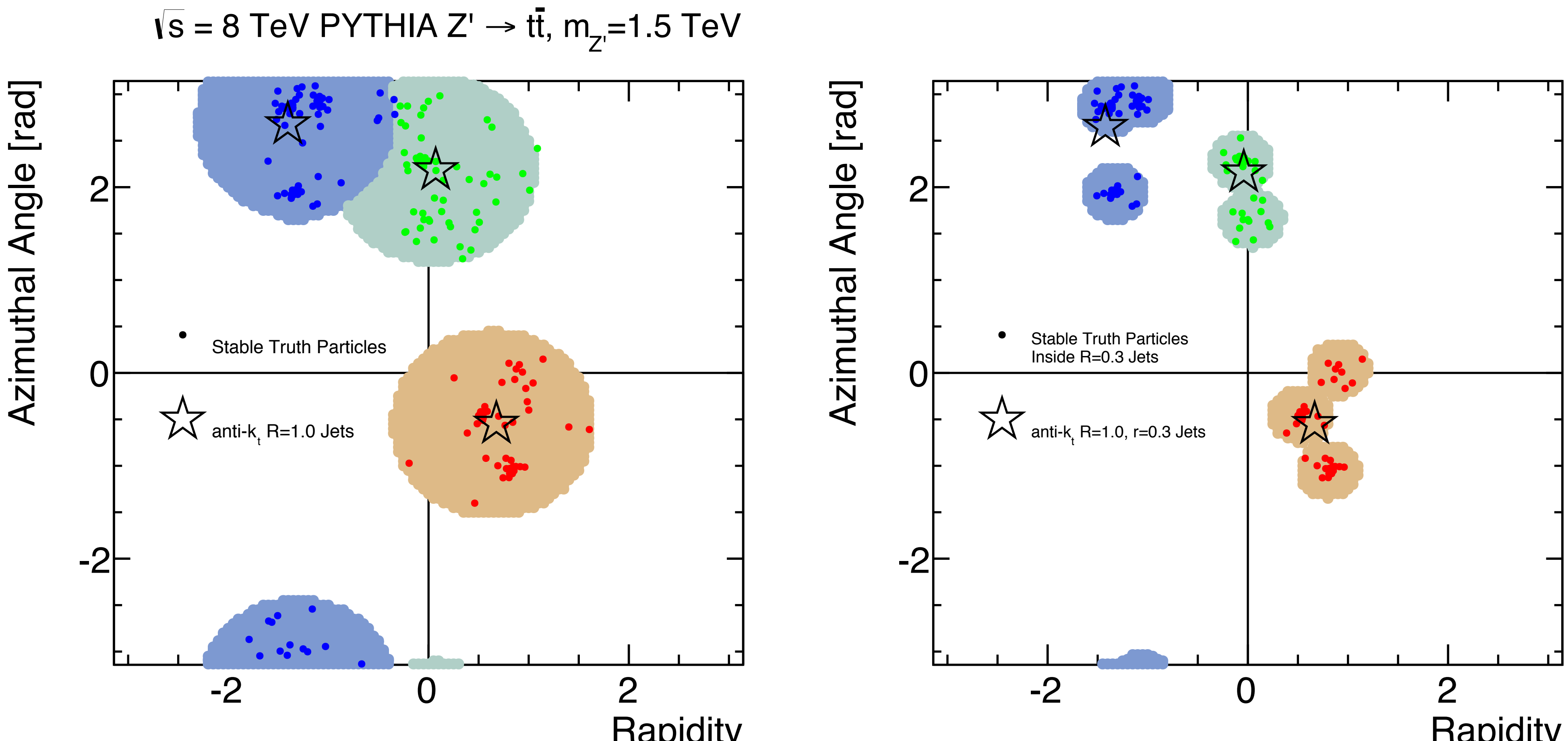
- Extends mMDT to reject wide-angled radiation
- Run backwards through clustering of C/A jet, removing constituent if:

$$\frac{min(p_{T,1}, p_{T,2})}{p_{T,1} + p_{T,2}} < z_{cut}\left(\frac{\Delta R_{12}}{R_0}\right)^{\beta}$$

  - Larger β allows for more soft & wide-angled radiation in a jet
- **Recursive soft drop**: Continue the procedure the soft-drop requirements are passed N times
  - Continues through good constituents, grooming them as well
- **Bottom-up soft drop**: Apply soft drop criteria during jet reconstruction

Jet Reclustering

- Build large-R jets from **fully-calibrated R=0.4 jets**
  - Benefit from small-R pileup suppression
  - Propagate full suite of small-R uncertainties
  - No additional large-R calibration needed - flexible choice in large-R radius
- Can use other grooming methods with R=0.4 jets

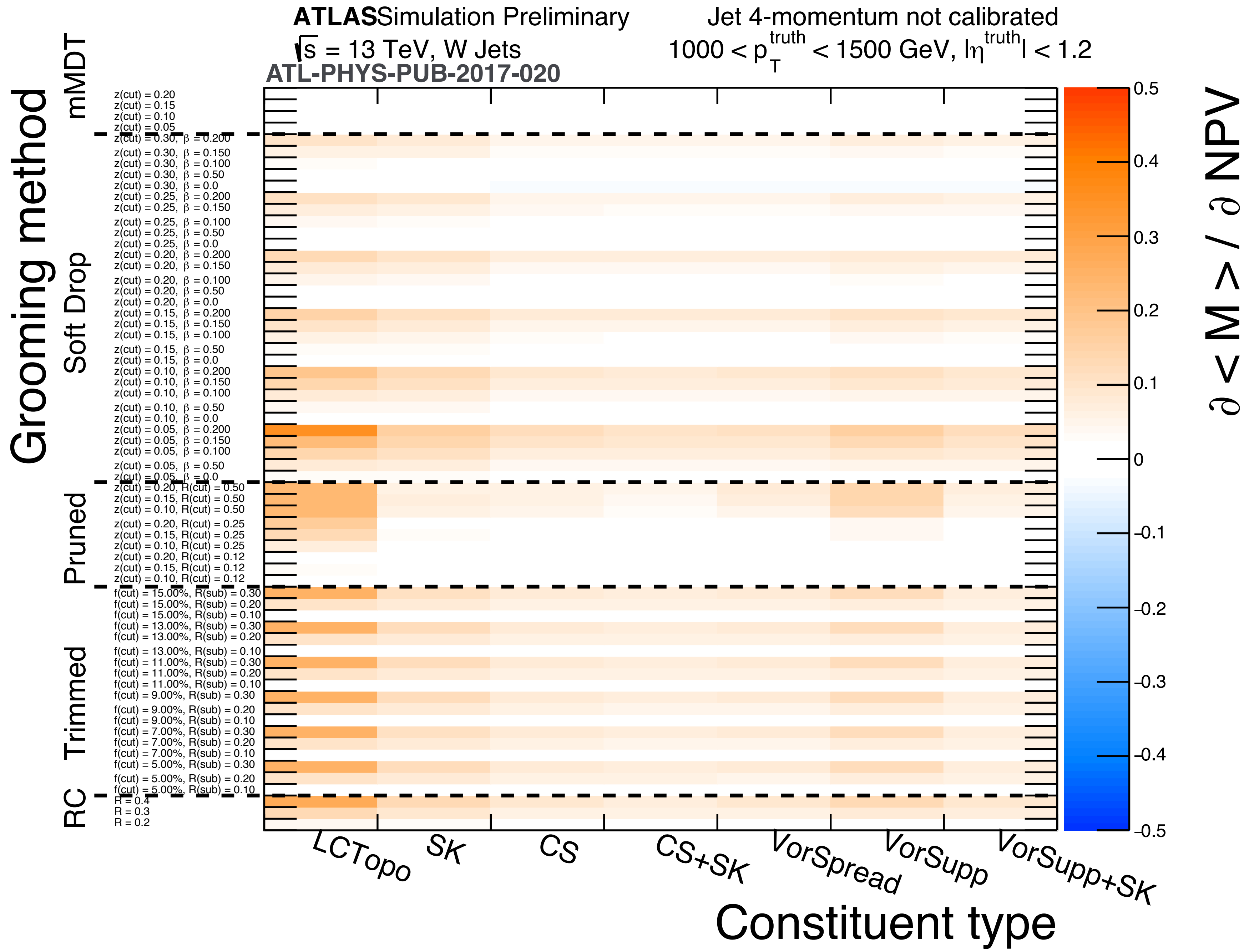


R=1 Jets **Reclustered Jets**

arXiv:1407.2922

Jet Tagging Performance

W mass stability vs N_{PV} for various jet types **⟨μ⟩~20**

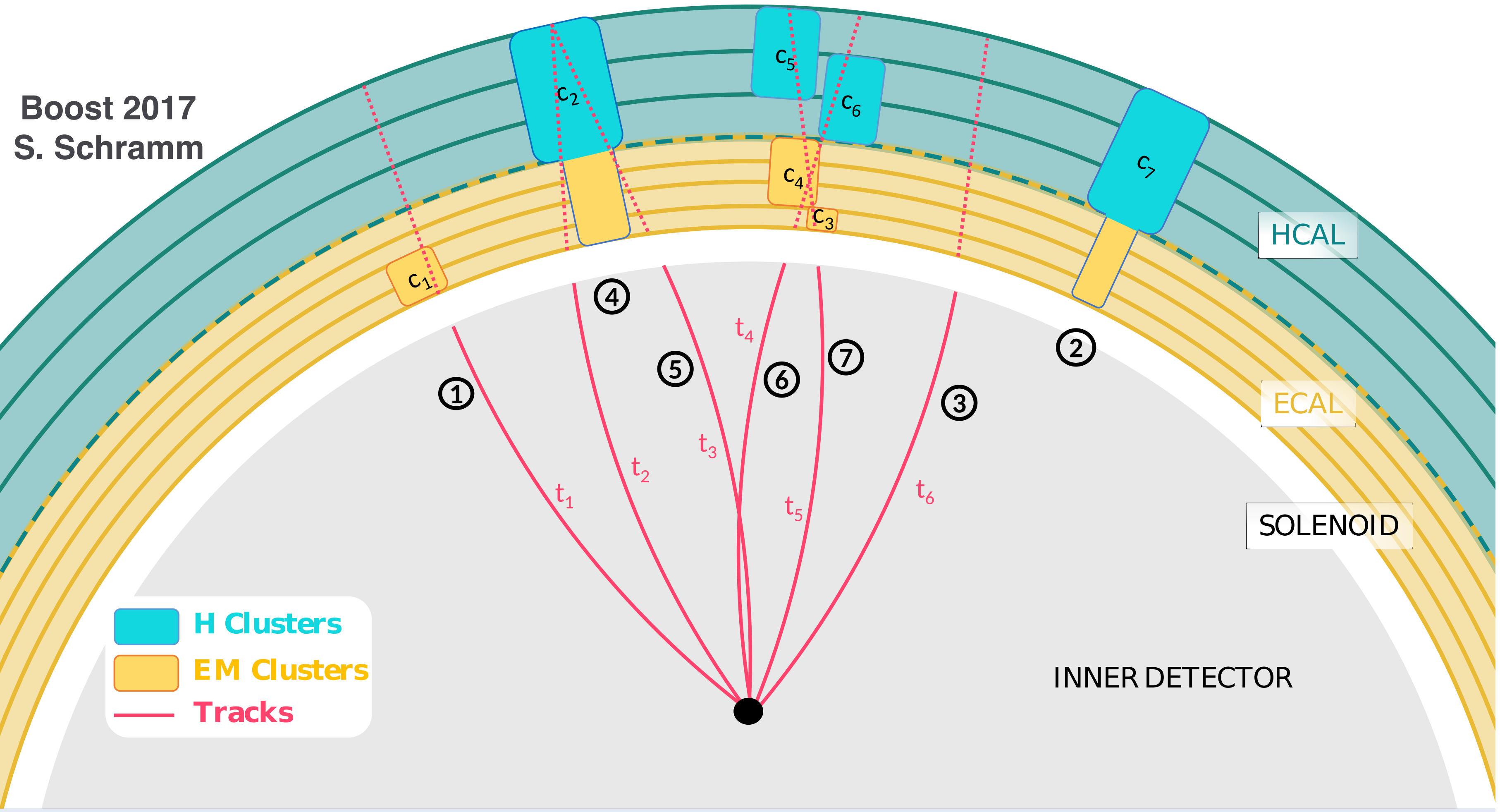


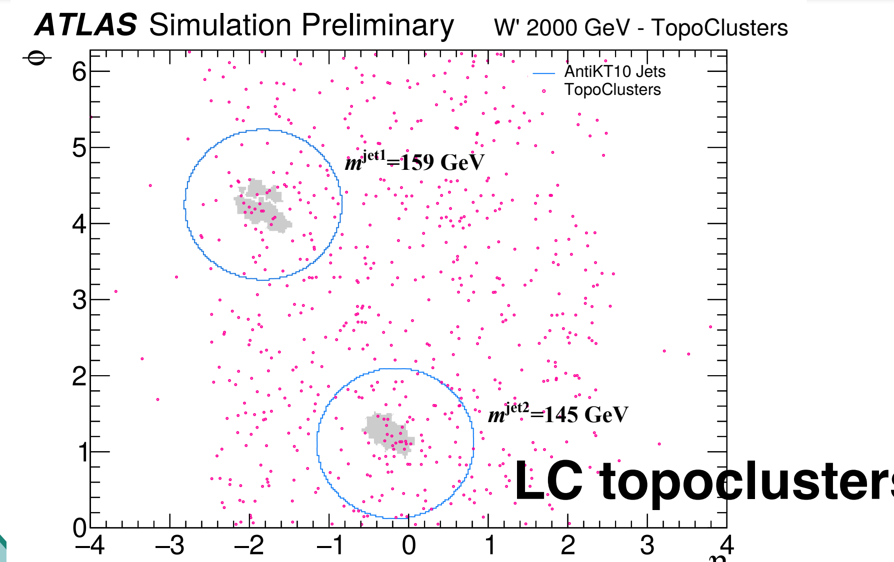
- Scan over many **clusters types** & **grooming methods** for best combination
- Tagging efficiency, fake rejection, and **mass stability w.r.t. pileup**

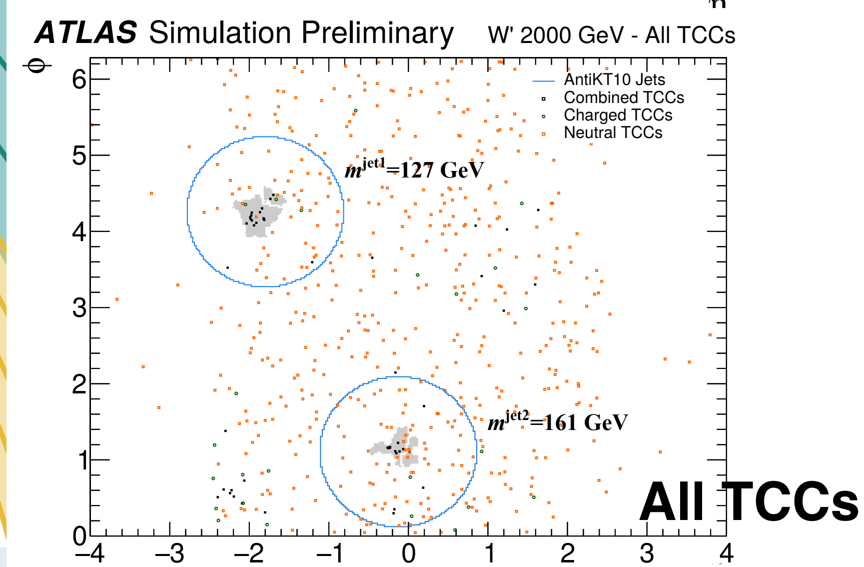
Track-Calo Clusters Method

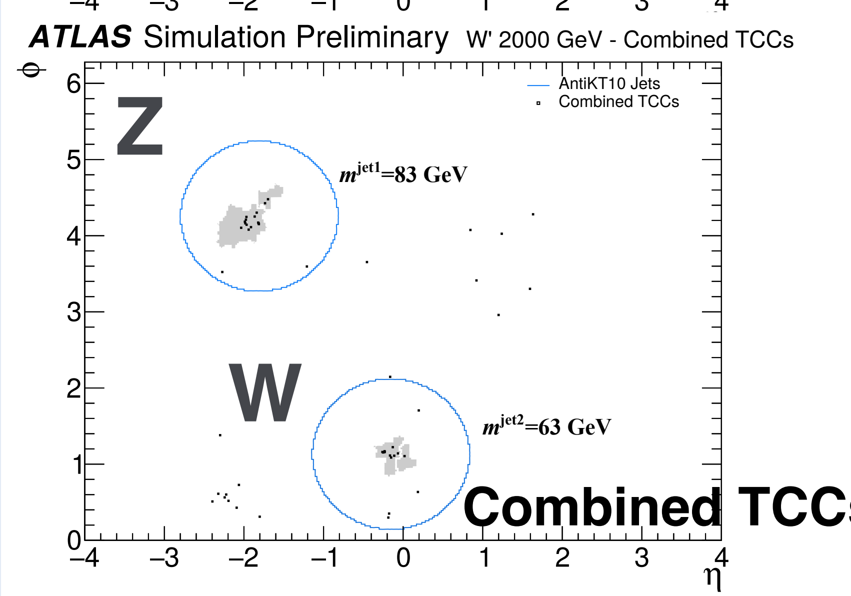
ATLAS-CONF-2018-016

Boost 2017
S. Schramm









TCC reconstruction: use track spatial coordinates and cluster energy components

Unique track-cluster match:

$$\mathrm{TCC}_{①} = (p_{\mathrm{T}}^{c_1}, \eta^{t_1}, \phi^{t_1}, m^{c_1} = 0)$$

Unmatched cluster:

$$\mathrm{TCC}_{②} = (p_{\mathrm{T}}^{c_7}, \eta^{c_7}, \phi^{c_7}, m^{c_7} = 0)$$

Unmatched track:

$$\mathrm{TCC}_{③} = (p_{\mathrm{T}}^{t_6}, \eta^{t_6}, \phi^{t_6}, m^{t_6} = 0)$$

In case of track-cluster-multi-matches, create one TCC object per hard-scatter PV track, and share the energy based on p_T ratios:

$$\mathrm{TCC}_{④} = (\alpha p_{\mathrm{T}}^{c_2}, \eta^{t_2}, \phi^{t_2}, \alpha m^{c_2} = 0)$$

$$\mathrm{TCC}_{⑤} = (\beta p_{\mathrm{T}}^{c_2}, \eta^{t_3}, \phi^{t_3}, \beta m^{c_2} = 0)$$

$$\alpha = \frac{p_{\mathrm{T}}^{t_2}}{p_{\mathrm{T}}\left[\mathbf{p}^{t_2} + \mathbf{p}^{t_3}\right]} \qquad \beta = \frac{p_{\mathrm{T}}^{t_3}}{p_{\mathrm{T}}\left[\mathbf{p}^{t_2} + \mathbf{p}^{t_3}\right]}$$