Pileup suppression in ATLAS:
jet-vertex tagging & track-based grooming

Pascal Nef

“Mitigation of pileup effects at the LHC” workshop at CERN

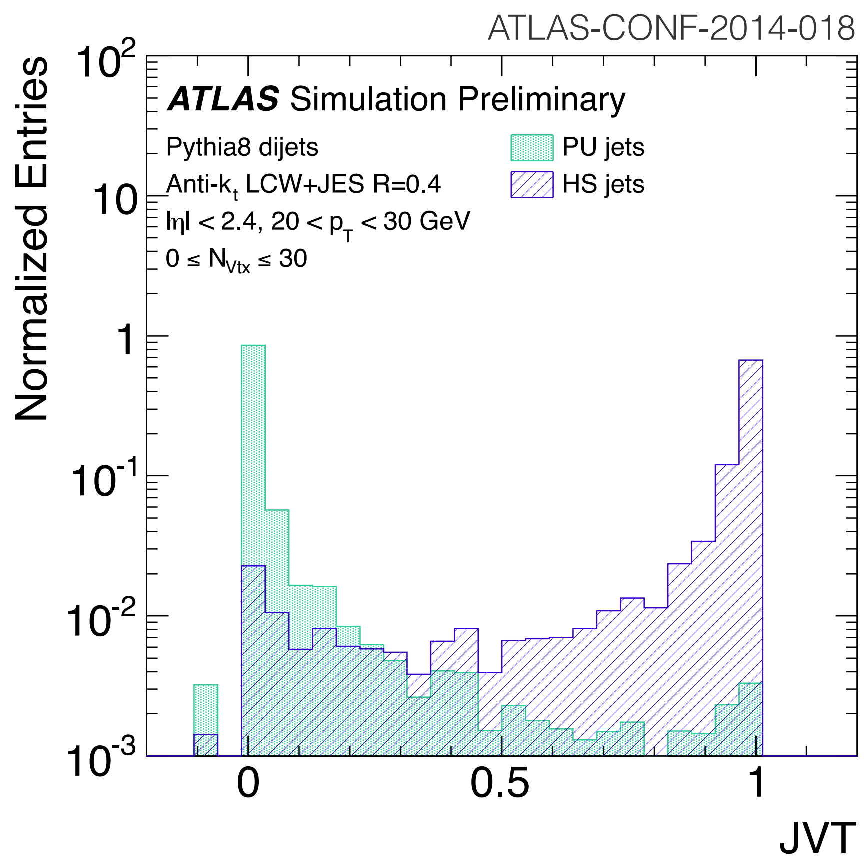
May 16th, 2014

Introduction

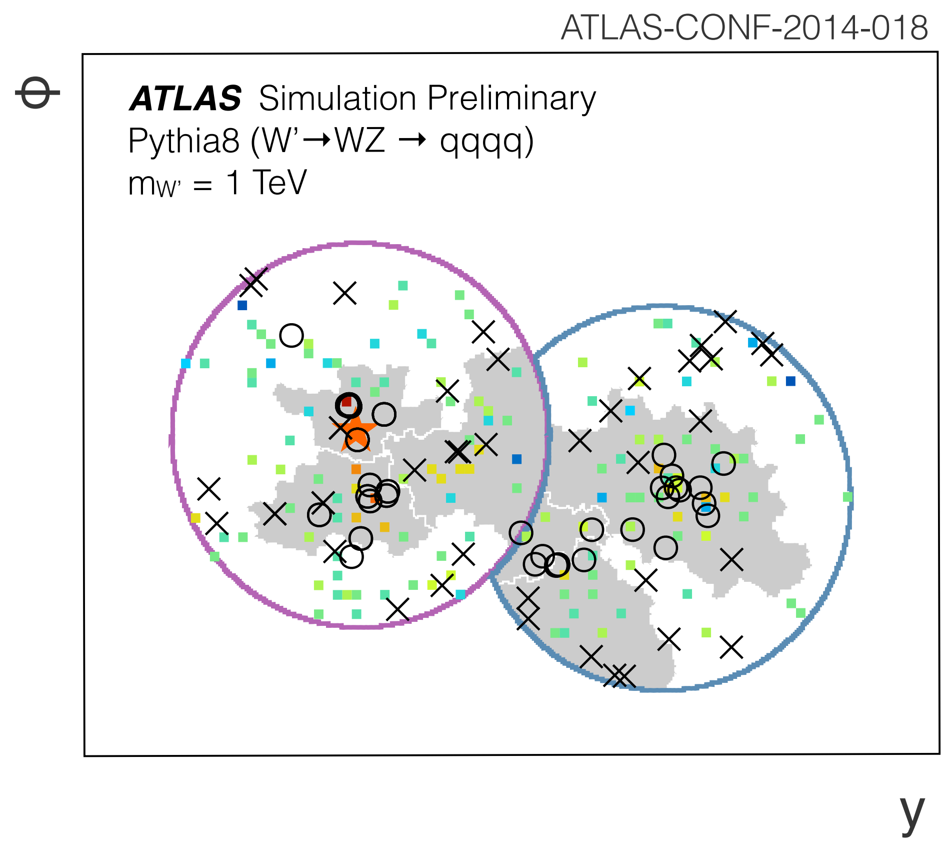
- Pileup is one of the main challenges at the LHC
  - Especially for jets: jet resolution, pileup jets, jet shapes and jet mass, …
  - Major area of research for experimentalist and theorists!

- Talk is organized in two themes: **pileup jet suppression** & **track-based grooming**
  - more details in recently approved CONF note: ATLAS-CONF-2014-018

Track-based pileup jet suppression



Track-based grooming



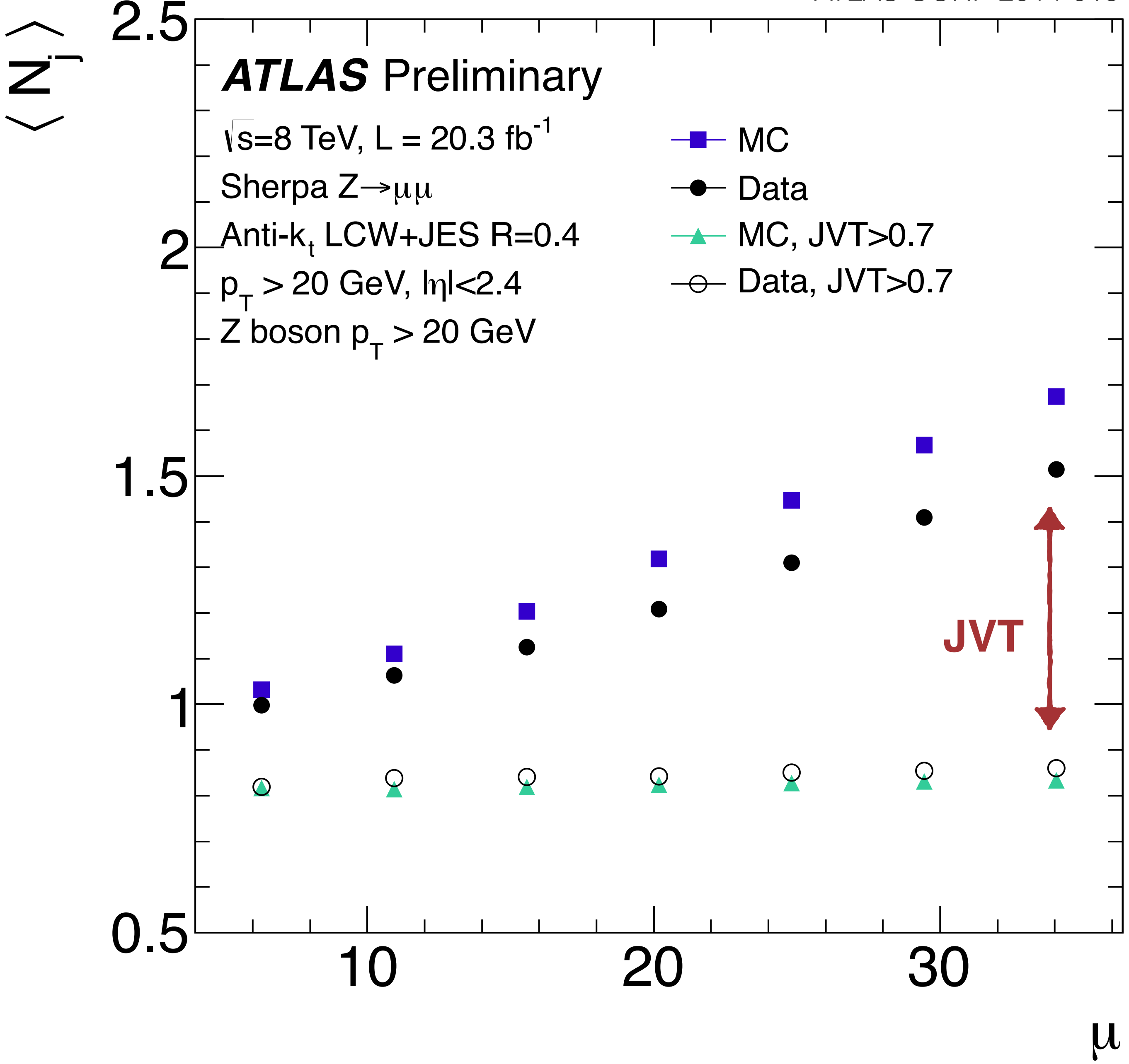
SLAC NATIONAL ACCELERATOR LABORATORY

Pileup jet suppression

The need for pileup jet suppression

- Pileup effect on jets are mitigated by applying the jet-area pileup correction
  - Local fluctuations in the event-by-event pileup activity can give rise to pileup jets

ATLAS-CONF-2014-018



- Mean number of jets per event after jet-area based correction
  - increases with μ
  - ATLAS simulation over-predicts
- After suppressing pileup jets using tracking information:
  - **$\langle N_j \rangle$ is flat**
  - good agreement between data and MC

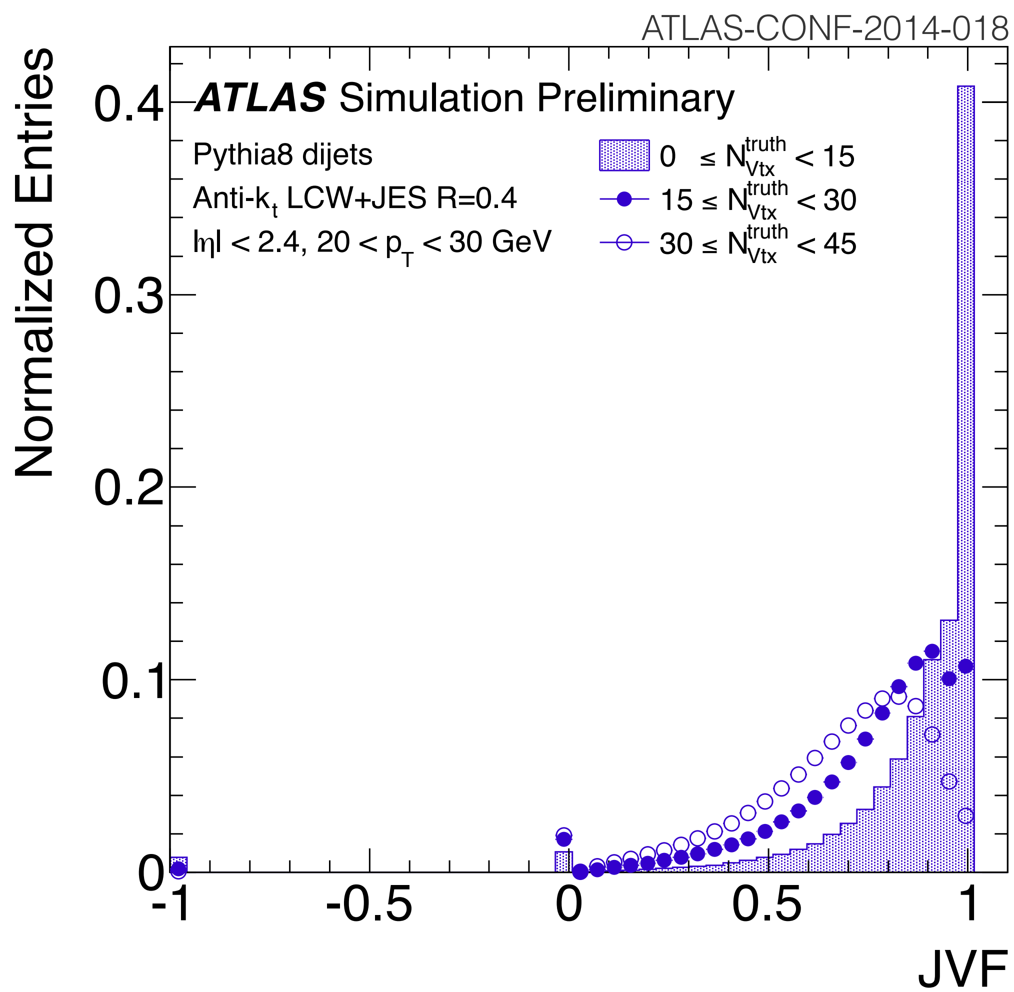
Pileup jet suppression with JVF

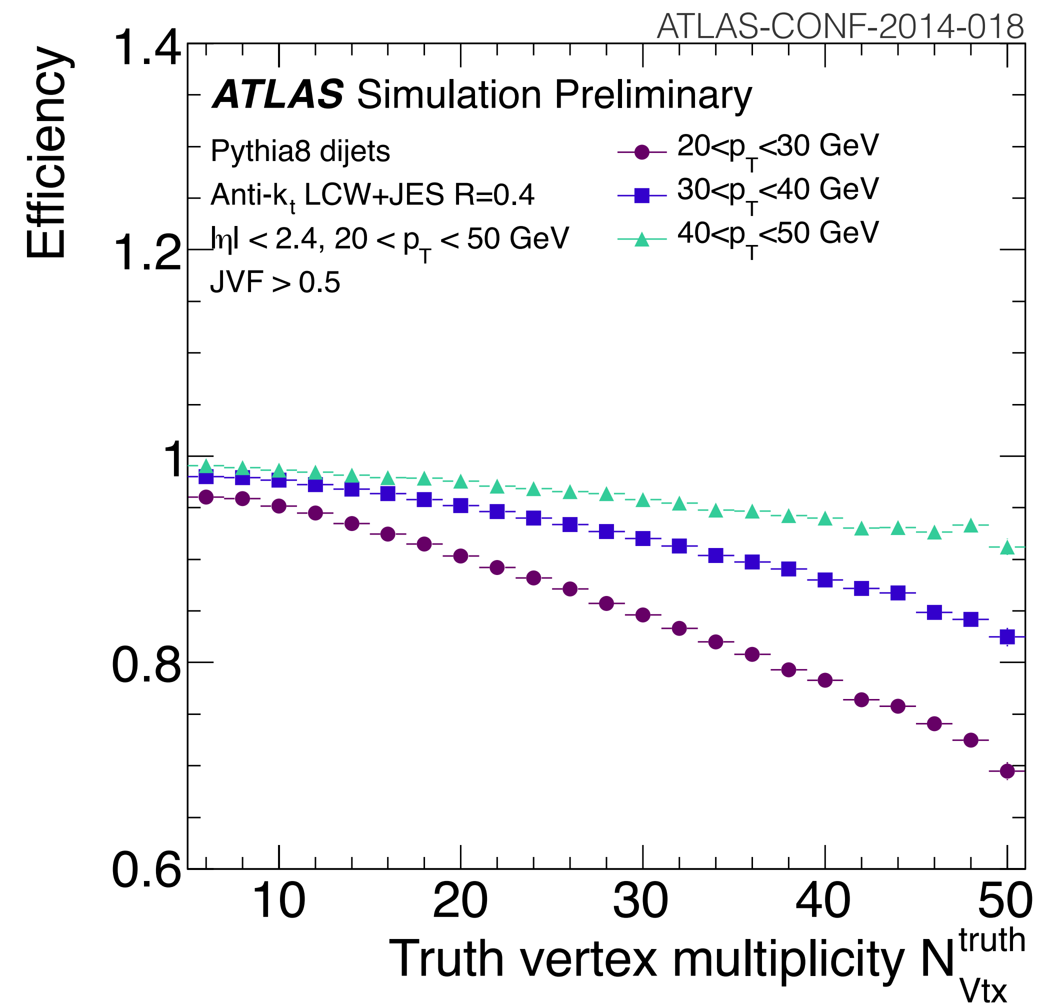
- Track-based pileup jet suppression:
  - associate tracks with vertices
  - associate tracks with calorimeter jets

$$\mathrm{JVF} = \frac{\sum_k p_{\mathrm{T}}^{\mathrm{trk}_k}(\mathrm{PV}_0)}{\sum_l p_{\mathrm{T}}^{\mathrm{trk}_l}(\mathrm{PV}_0) + \sum_{n\geq 1} \sum_l p_{\mathrm{T}}^{\mathrm{trk}_l}(\mathrm{PV}_n)}$$

- Pileup jet suppression in ATLAS used to be entirely based on the jet-vertex fraction JVF
  - **JVF itself is pileup dependent**: p_T sum of associated pileup tracks is pileup dependent

- find alternative variables variables that …
  - optimize performance
  - give N_{Vtx} insensitive hard-scatter jet efficiencies





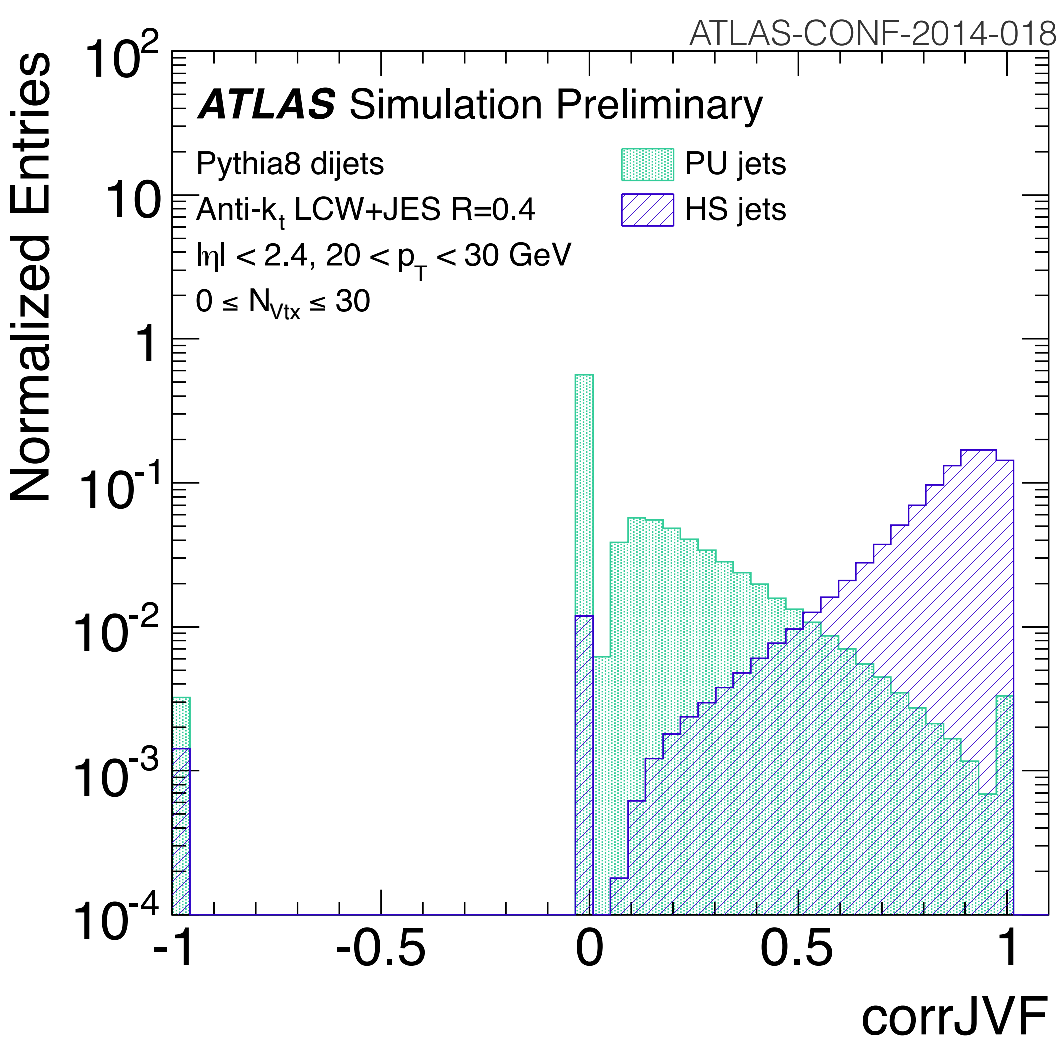
SLAC NATIONAL ACCELERATOR LABORATORY

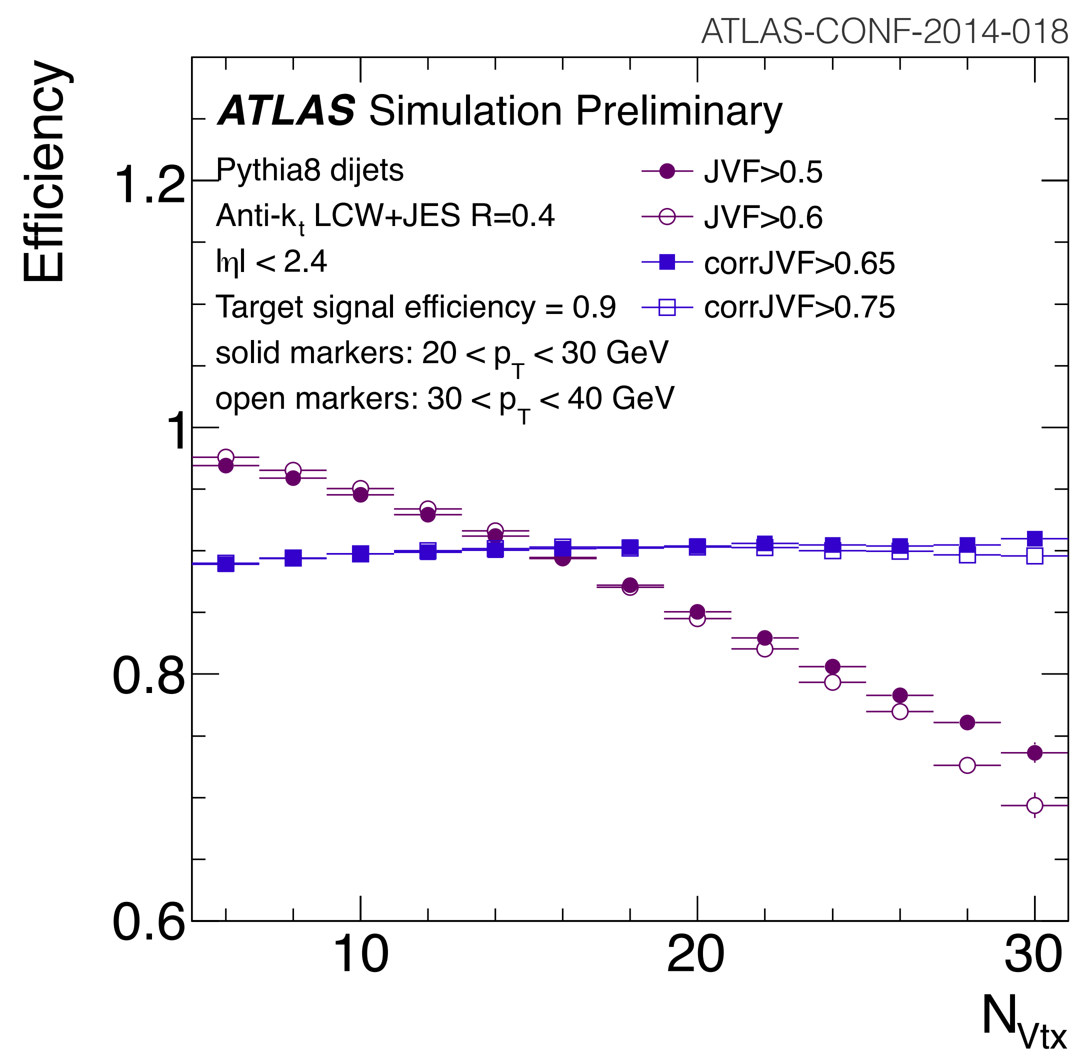
from JVF to corrJVF

- Correcting JVF (in average) for its pileup dependence:

$$\text{corrJVF} = \frac{\sum_k p_T^{\text{trk}_k}(\text{PV}_0)}{\sum_l p_T^{\text{trk}_l}(\text{PV}_0) + \frac{\sum_{n\geq 1}\sum_l p_T^{\text{trk}_l}(\text{PV}_n)}{(k\cdot n_{\text{trk}}^{\text{PU}})}}$$

- mean p_T from pileup tracks associated with jets increases linearly with the **total number of pileup tracks in the event** (n_{trk}^{PU})
  - n_{trk}^{PU} is a proxy for the event pileup activity:
  - tried different variables: μ, N_{Vtx}, p_T density ρ





The charged fraction R_{pT}

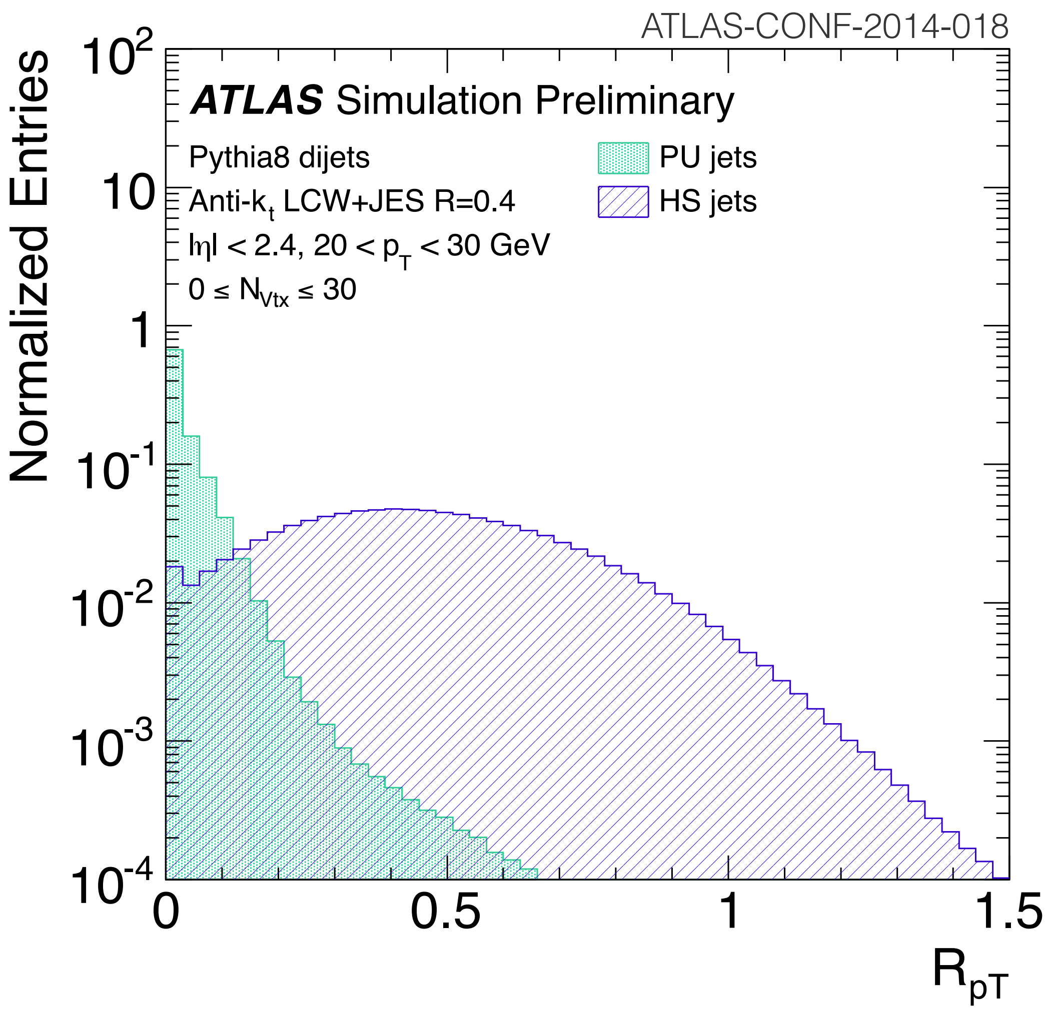
- R_{pT} is the charged p_T of a jet and defined as:

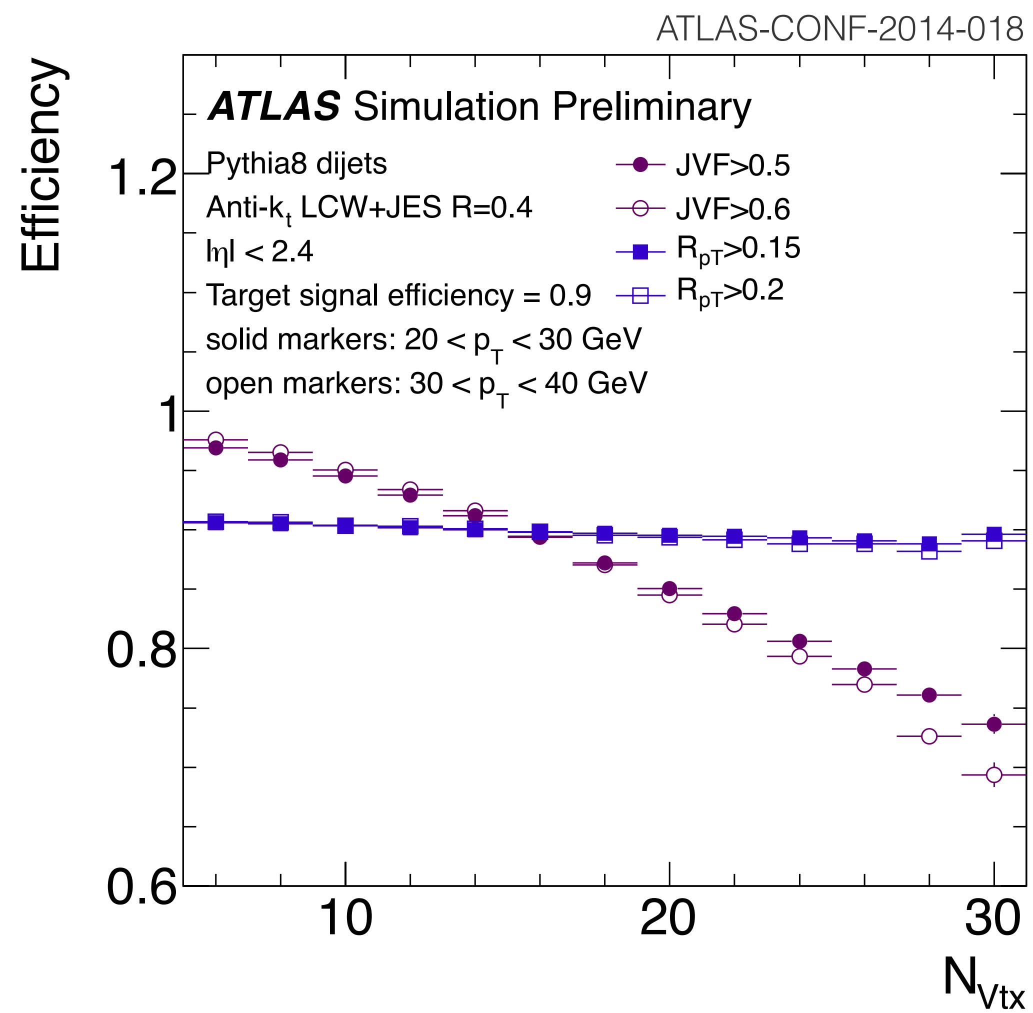
$$R_{\mathrm{pT}} = \frac{\sum_k p_{\mathrm{T}}^{\mathrm{trk}_k}(\mathrm{PV}_0)}{p_{\mathrm{T}}^{jet}}$$

→ hard-scatter tracks only

→ fully calibrated (pileup corrected) calorimeter p_T

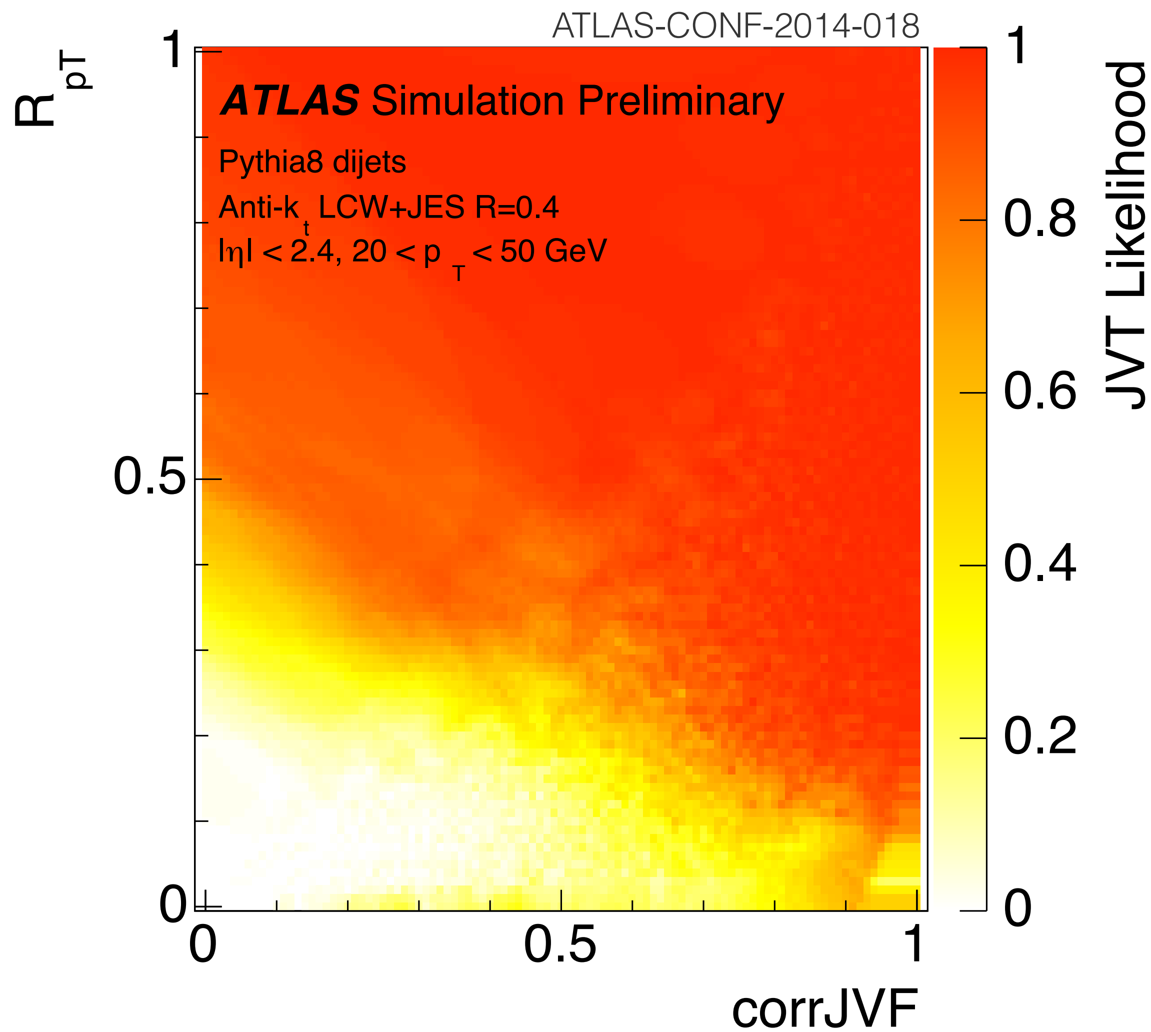
  - constructed from pileup corrected / insensitive variables
  - excellent discrimination between hard-scatter and pileup jets

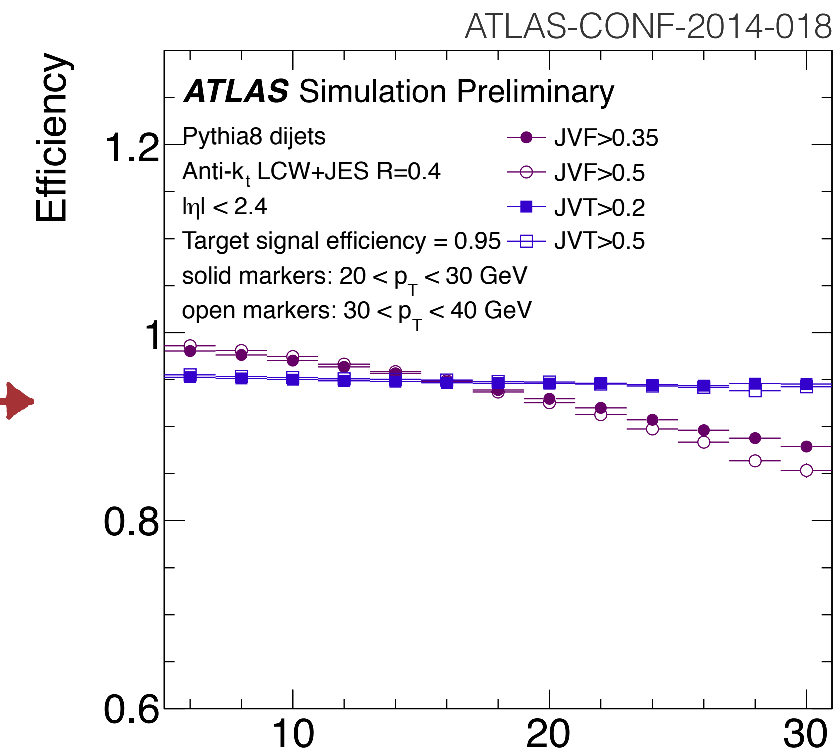


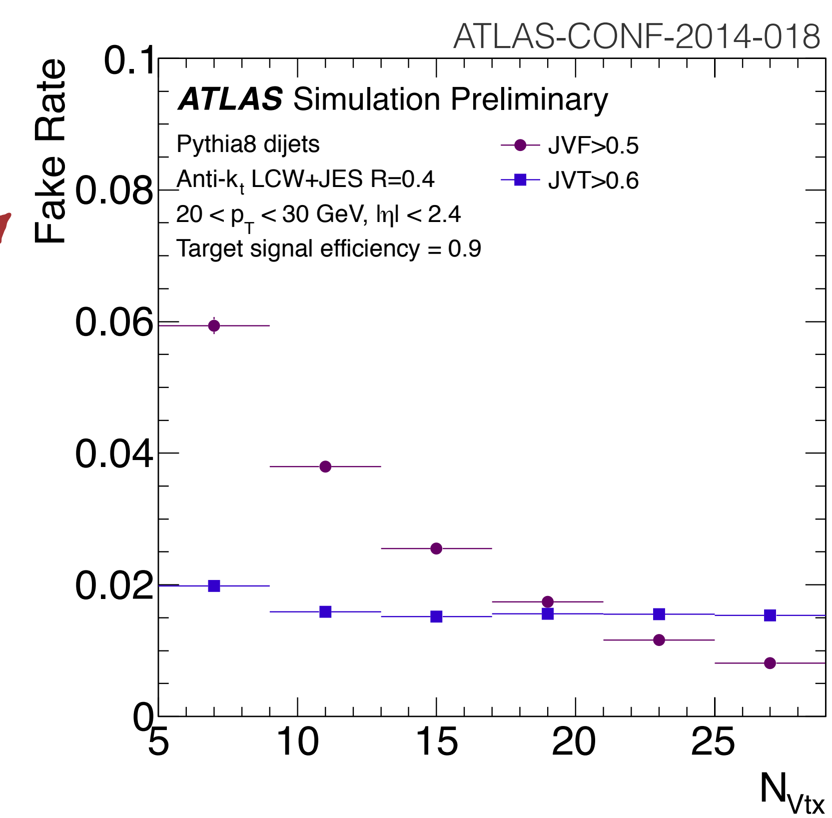


the JVT likelihood

- constructing the jet-vertex tagger (JVT) as a 2D likelihood from corrJVF and R_{pT}
  - using a nearest neighbor algorithm to avoid statistical fluctuations





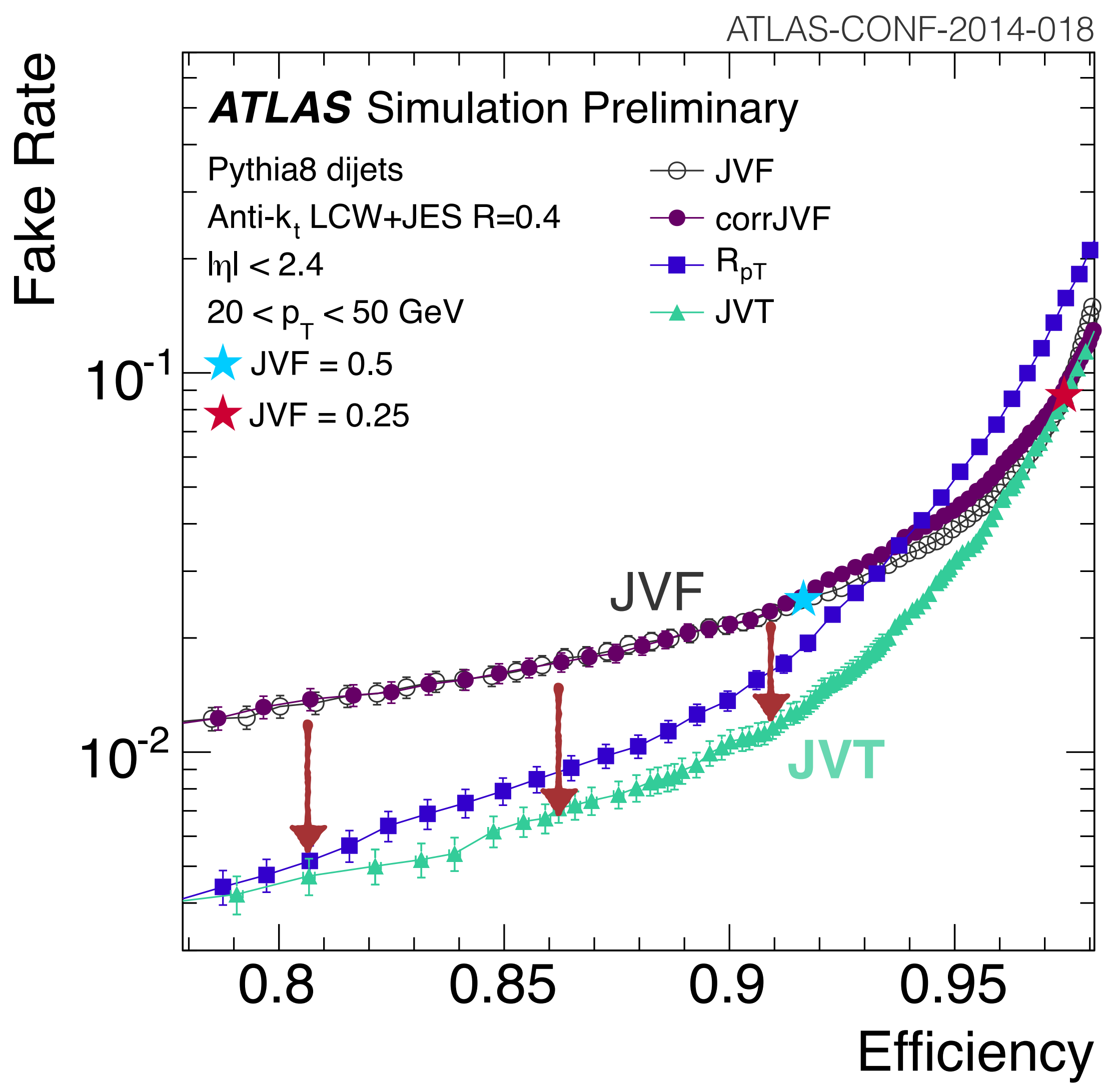


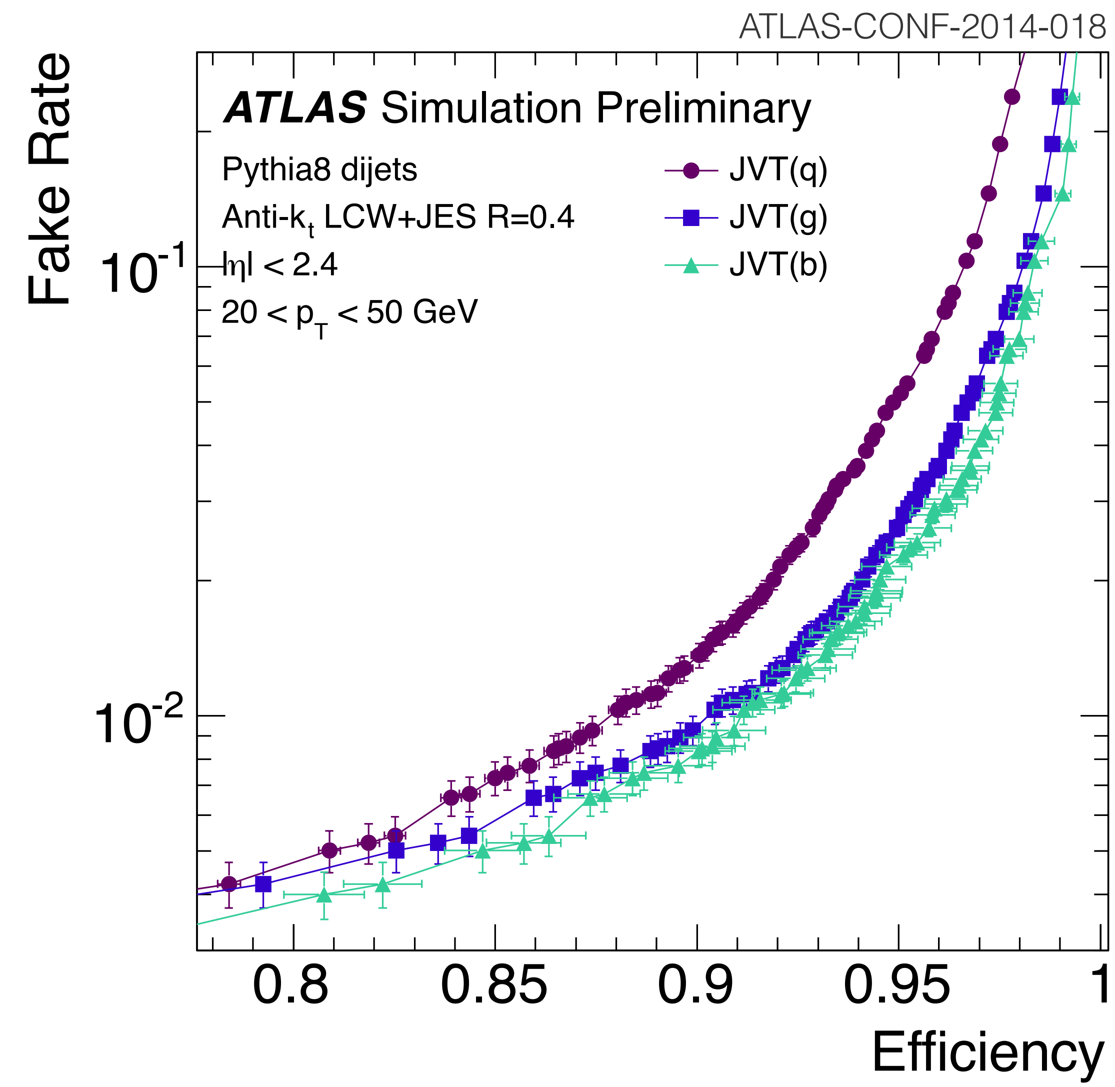
- hard-scatter jet efficiency & pileup fake rate are stable with pileup
  - **no need to re-optimize** cut values as beam conditions change!

SLAC NATIONAL ACCELERATOR LABORATORY

Efficiency vs. fake-rate curves

- efficiency vs. fake-rate curves (ROC) for JVF and JVT
  - large improvement in pileup jet rejection at fixed efficiency





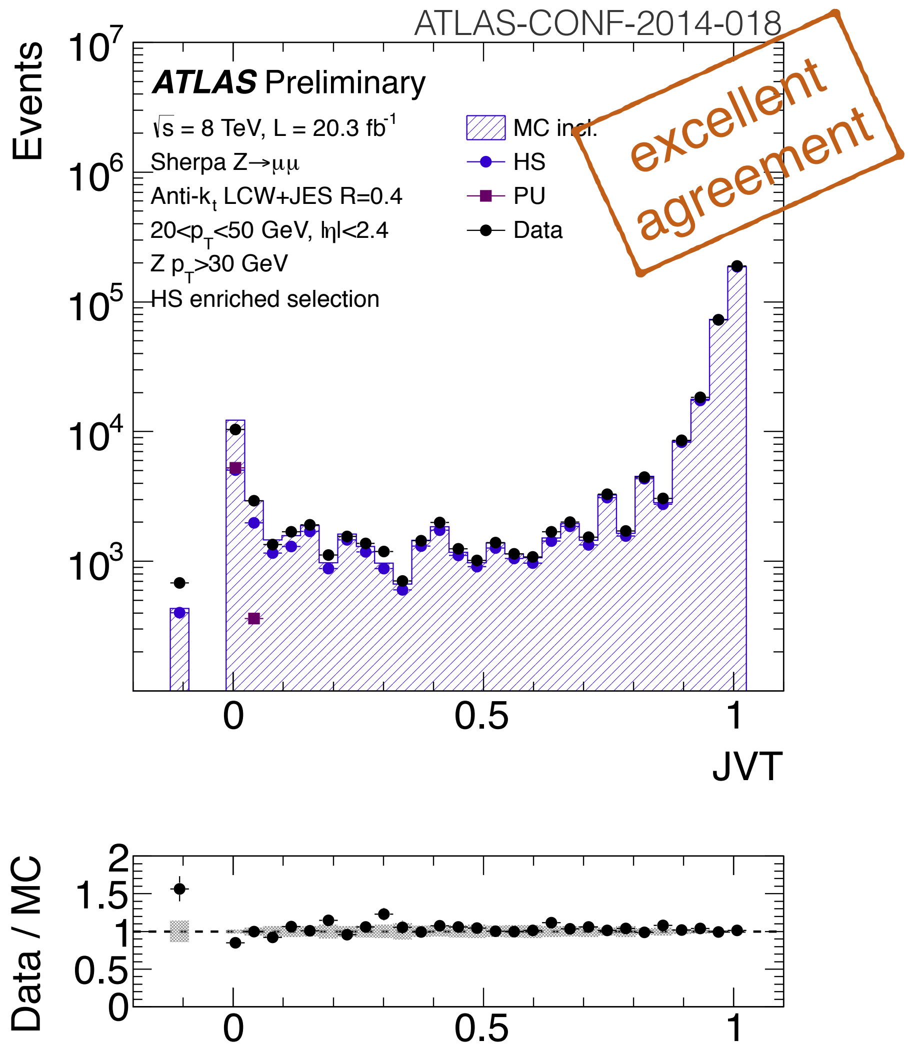
- performance is worse for light (uds)-quark initiated jets than for b-quark and gluon jets
  - light quark jets have higher response and radiate less

SLAC NATIONAL ACCELERATOR LABORATORY

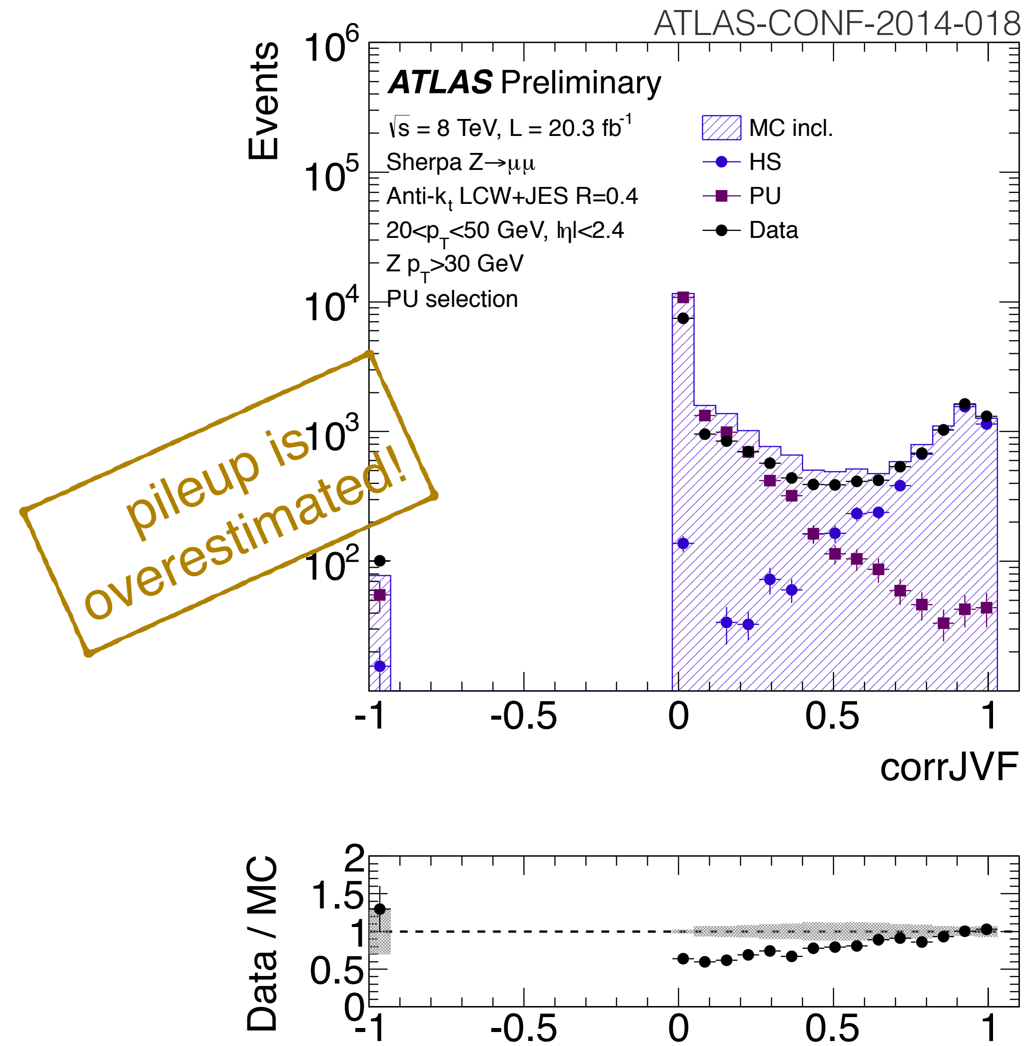
Data to MC comparison in Z(μμ)+jets

- Validate modeling of JVT (and corrJVF and R_{pT}) using Z(μμ)+jets and semileptonic ttbar events
  - separately in hard-scatter and pileup dominated regions

- Hard-scatter jet dominated region



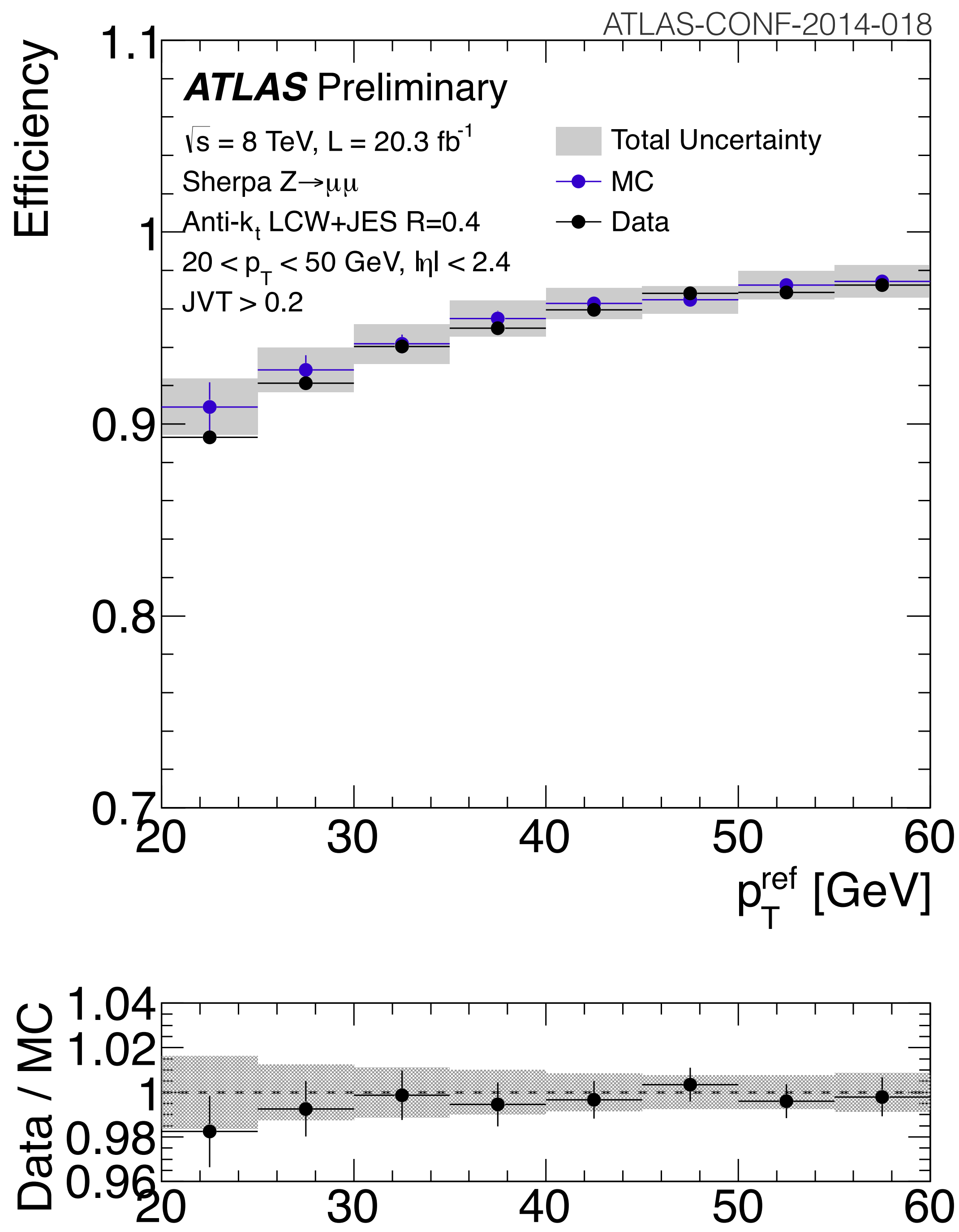
- Pileup jet dominated region



- Ultimately care about modeling of corrJVF, R_{pT} and JVT for hard-scatter, which is good.
- Rate of pileup jets are too high in simulation, corrJVF, R_{pT} and JVT shapes are reasonably well modeled

JVT efficiency measurement

- JVT efficiency measurement from tag & probe in Z+jets events:
- JVT efficiency measured in the signal region with $|\Delta\phi(Z, jet)| > 2.8$
- **pileup background** in the signal region **estimated from data** using a pileup control region $|\Delta\phi(Z, jet)| < 1.2$
- dominant systematics:
  - difference in efficiency between Sherpa & Powheg+Pythia (different fragmentation model)

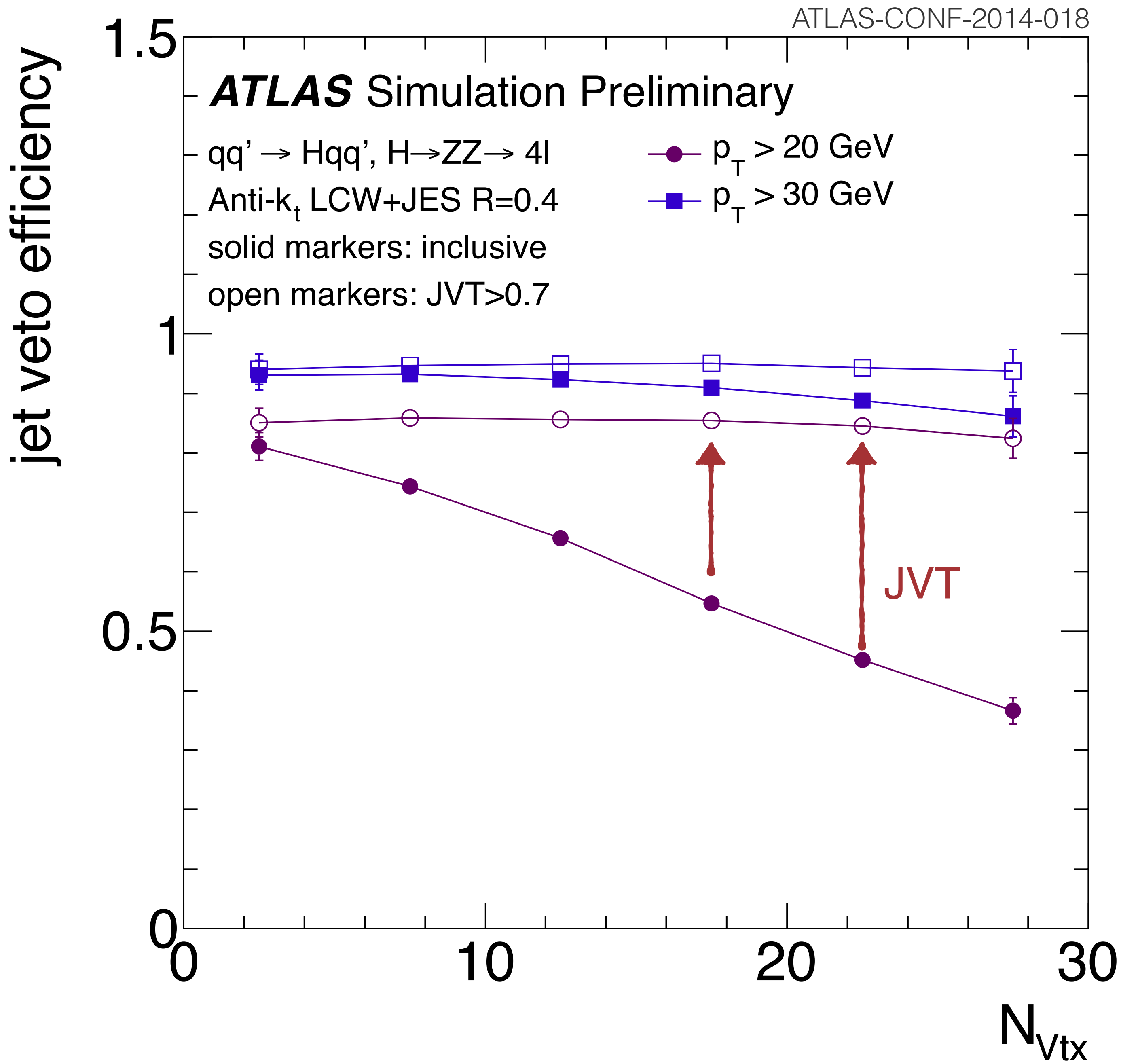


- provide efficiency measurements for:
  - three different operating points
  - vs p_t and vs. η
- scale factors are consistent with unity
  - 1% to 2% uncertainty

Physics application: VBF Higgs

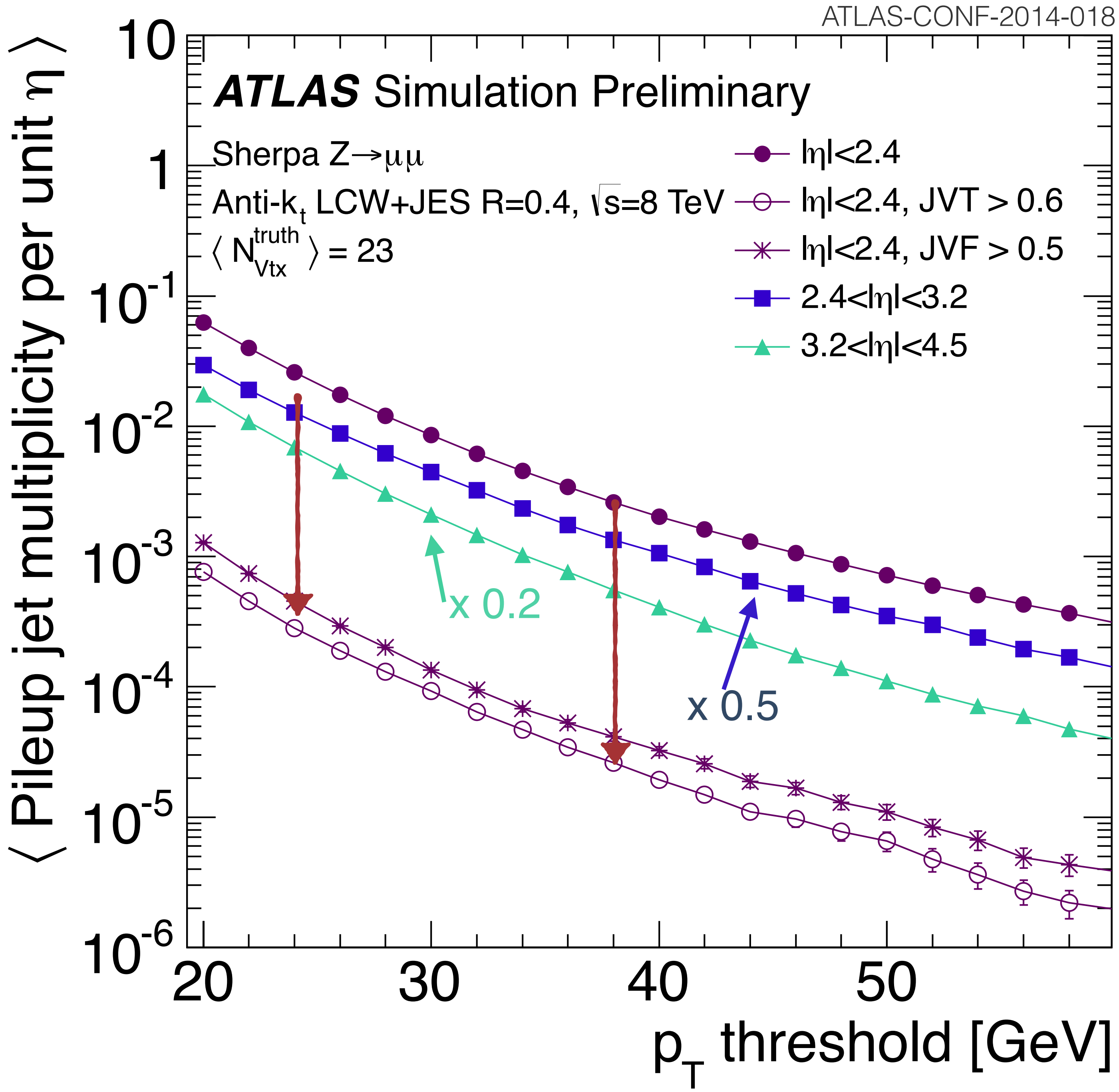
- Application of JVT to a VBF H->4l analysis:
- Typical event selection
  - require two p_T > 30 GeV jets to be separated in η by more than 3 units
- Define the jet-veto efficiency as:
  - fraction of events with no additional (third) jet within the tracker coverage

- Veto efficiency for p_T > 20 GeV jets spoilt by pileup jets
- **Flat efficiency** is recovered if **JVT** is used to suppress pileup jets



Pileup jet rates: where do we stand?

- Pileup jet rates as a function of jet p_T threshold and η in ATLAS simulation

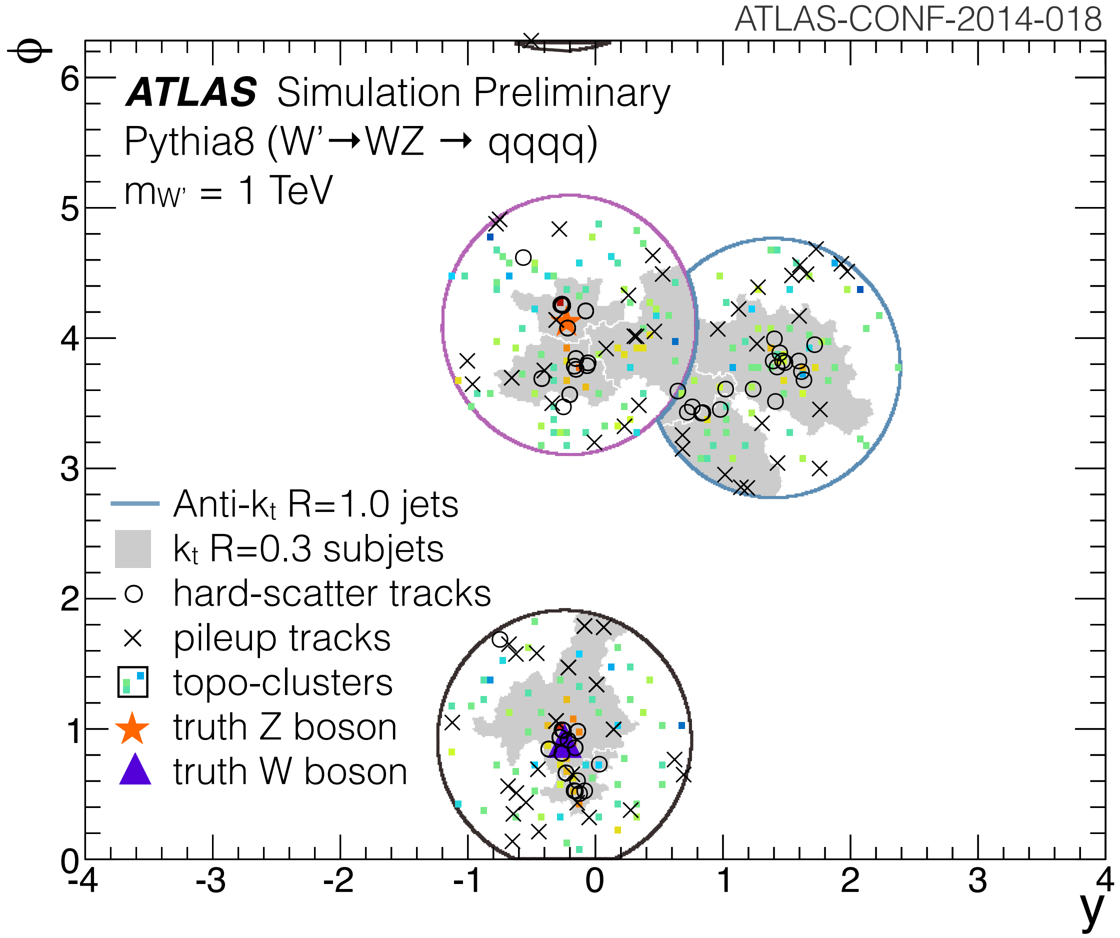


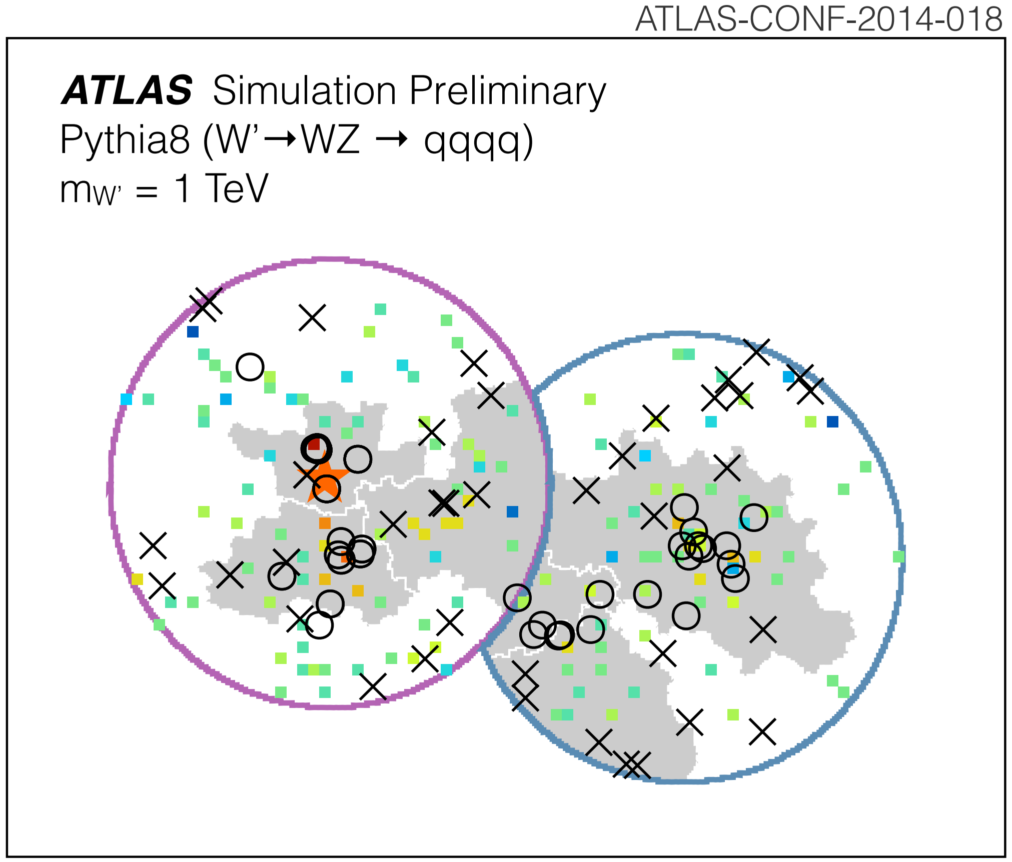
- This is for an average number of truth interactions of 23 (ranging between 0 and 50)
- Pileup jet rates in the forward region are strongly suppressed w.r.t. the central region:
  - Mainly due to granularity and noise thresholds of the ATLAS calorimeter
- Using JVT the pileup jet rate is suppressed by factor ~ 100 in the central region
  - Forward region now become important
  - As of today: No calorimeter (jet shape) based pileup jet suppression in the forward region in ATLAS

Track-based grooming

Track-based grooming

- Event display of a W' at m=1 TeV, decaying via WZ -> qqqq
  - reconstruct anti-k_t R=1.0 jets and their k_t R=0.3 subjets
  - ghost associate hard-scatter and pileup tracks to the subjets: calculate corrJVF and related variables





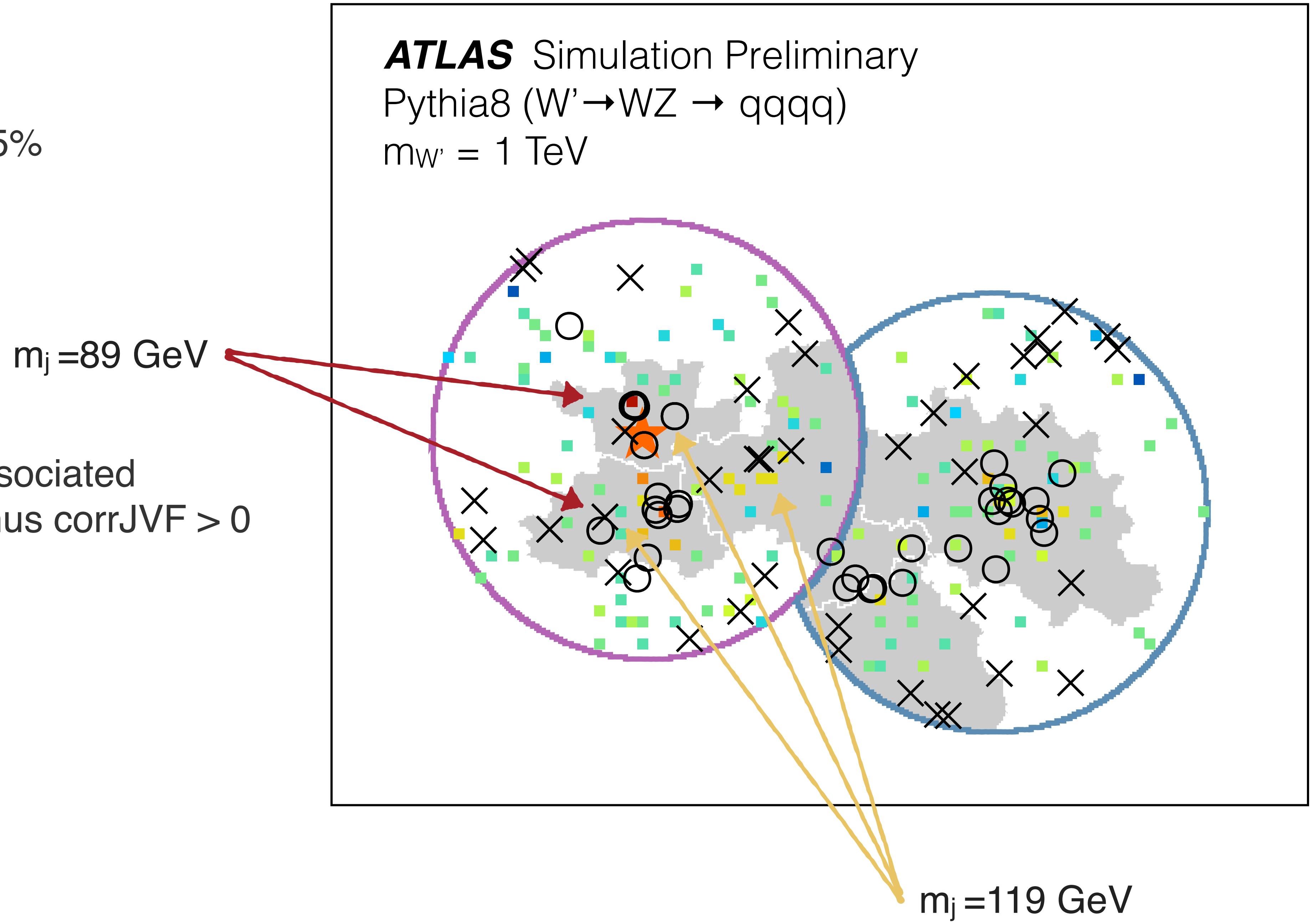
Track-based grooming

- Event display of a W' at m=1 TeV, decaying via WZ -> qqqq
  - reconstruct anti-k_t R=1.0 jets and their k_t R=0.3 subjets
  - ghost associate hard-scatter and pileup tracks to the subjets: calculate corrJVF and related variables

- The gray subjets pass a $f_{Cut} = p_T^{subj} / p_T^{ungroomed} > 5\%$

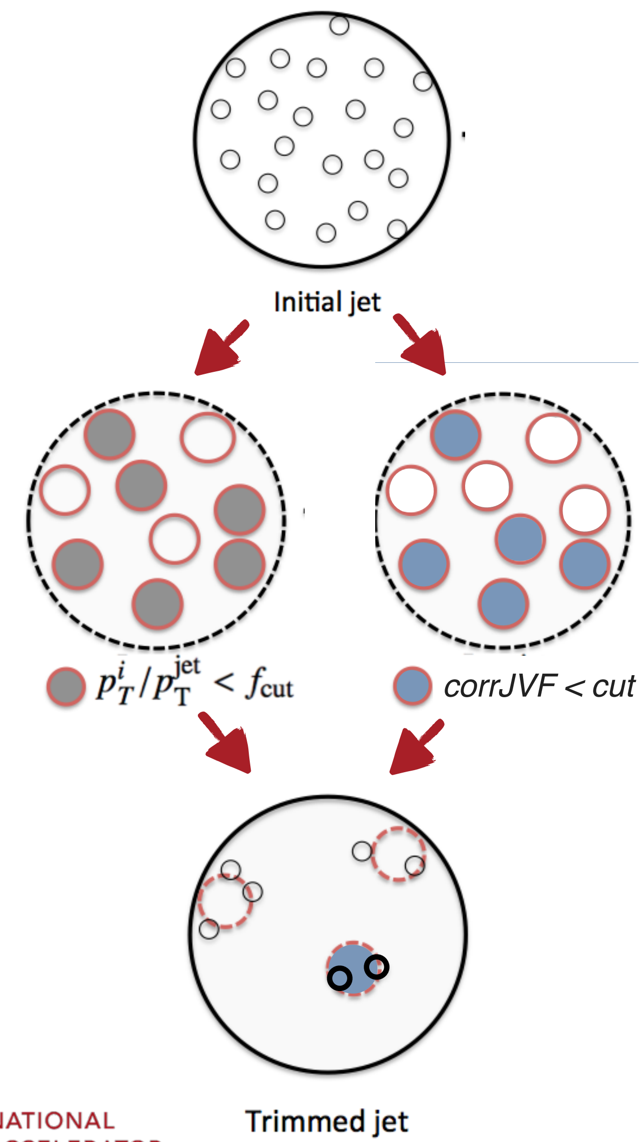
- Only two subjets have associated hard-scatter tracks and thus corrJVF > 0

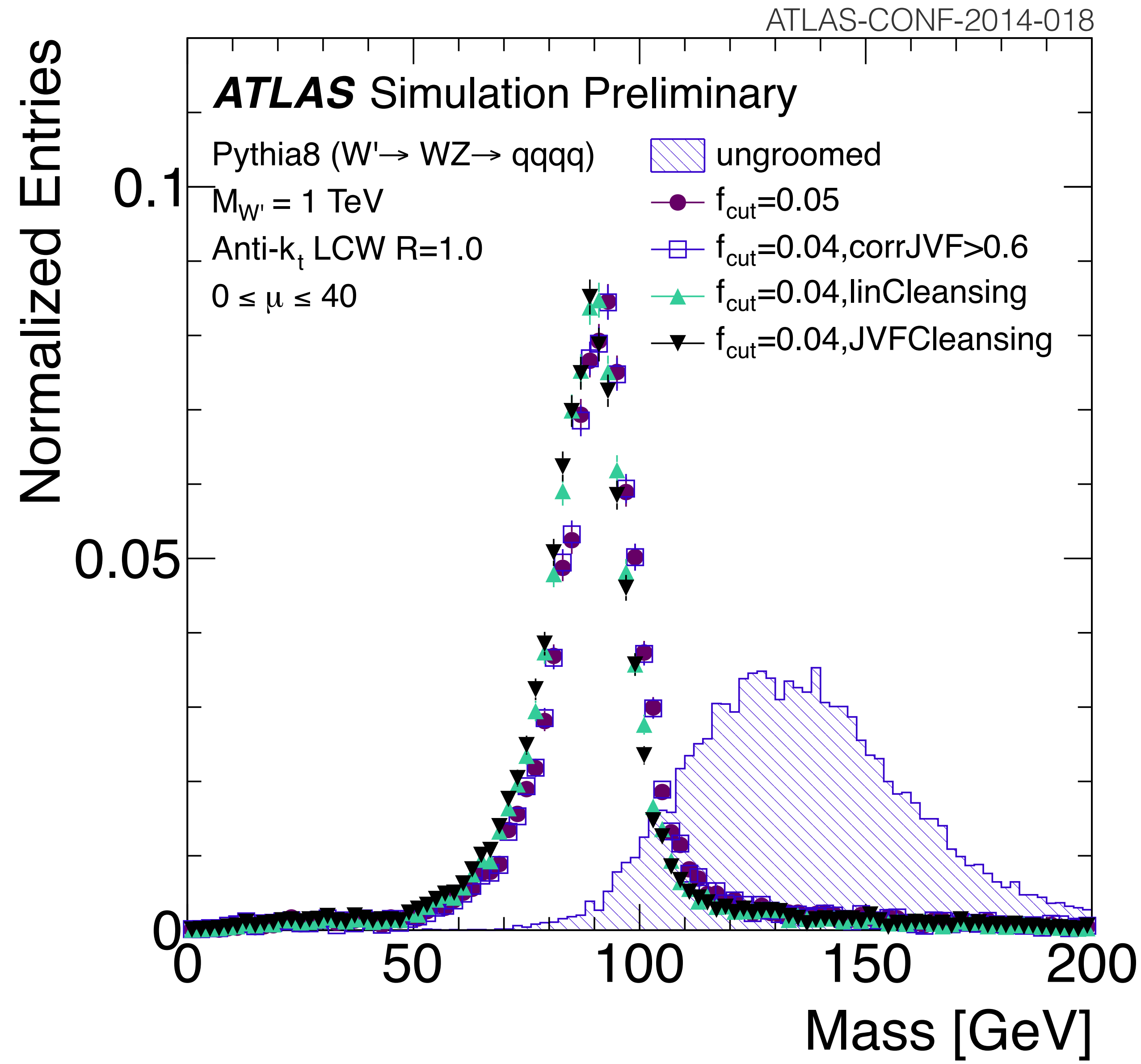
ATLAS-CONF-2014-018



corrJVF-based grooming

- corrJVF-based grooming
  - use subjet corrJVF in combination with trimming to groom large-R jets
- this is different
  not a **‘neutral-proportional-to-charge’** approach!

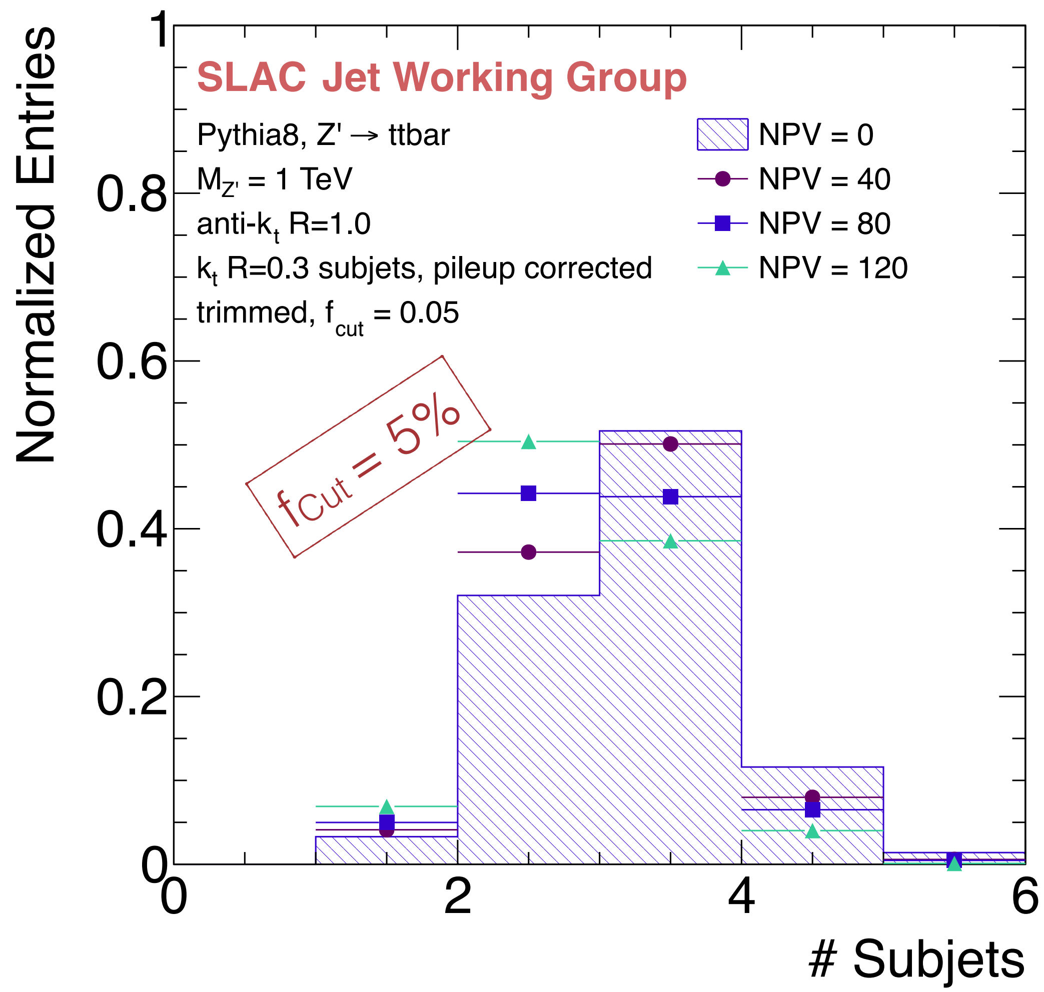


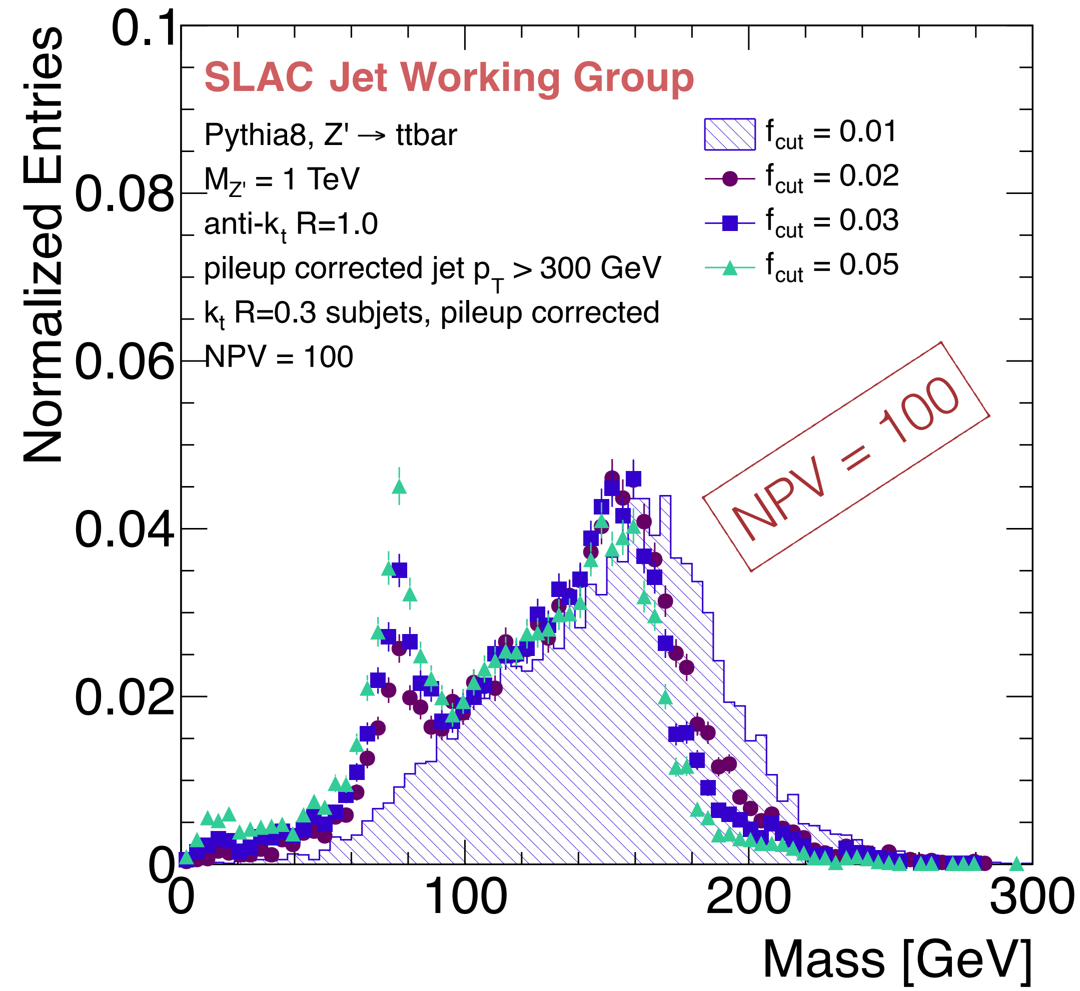


- for 2012 pileup conditions, no improvement observed w.r.t. calorimeter-only based trimming
  - similar conclusion for jet cleansing [1309.4777] (updated due to bug)

truth-level study for corrJVF grooming

- **truth-level study** at high pileup: this is ***not*** ATLAS material
  - Z'-> ttbar (all hadronic), $M_{Z'}$ = 1 TeV
  - looking at leading anti-k_t R=1.0 jets, with pileup corrected p_T > 300 GeV

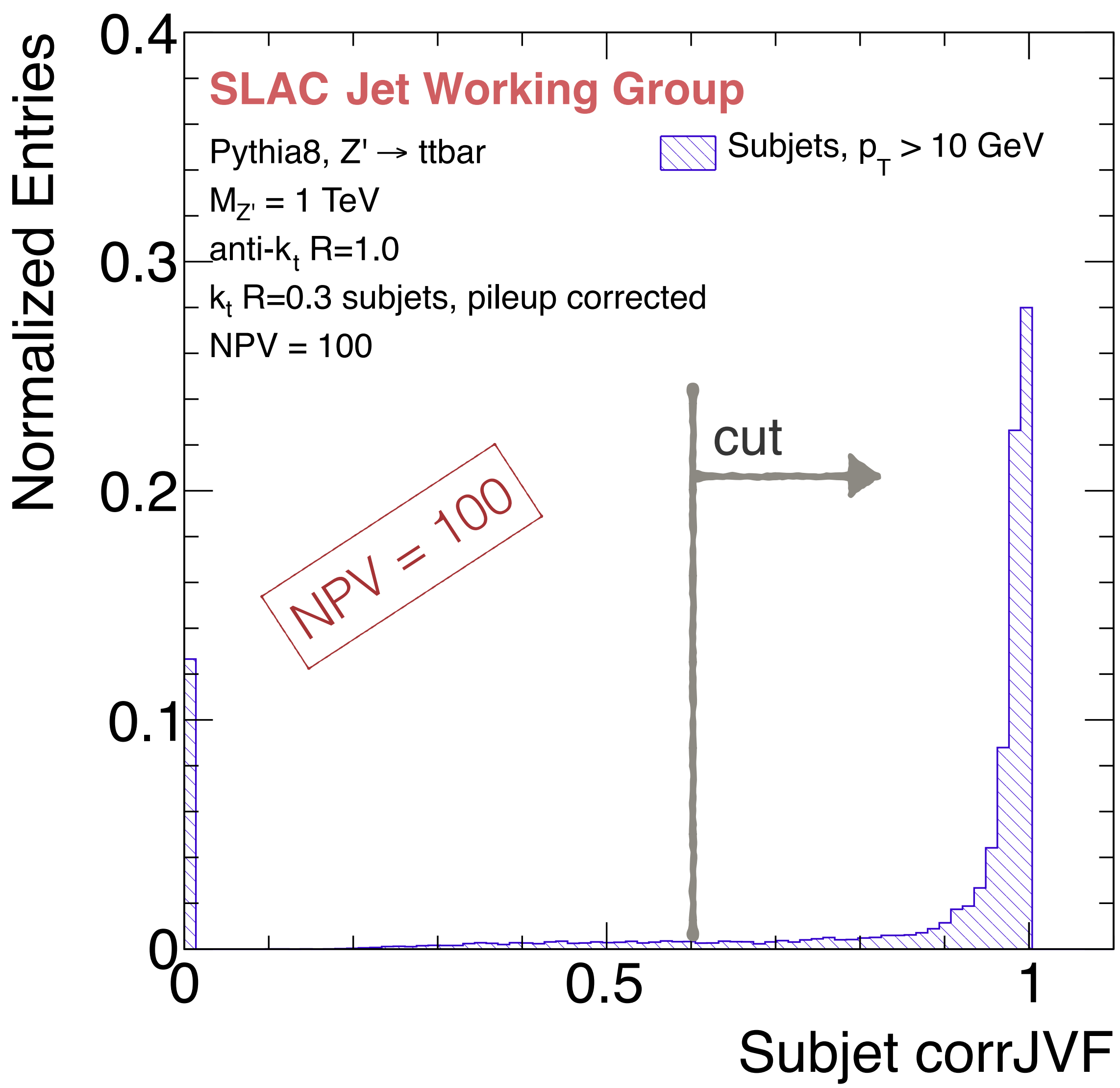


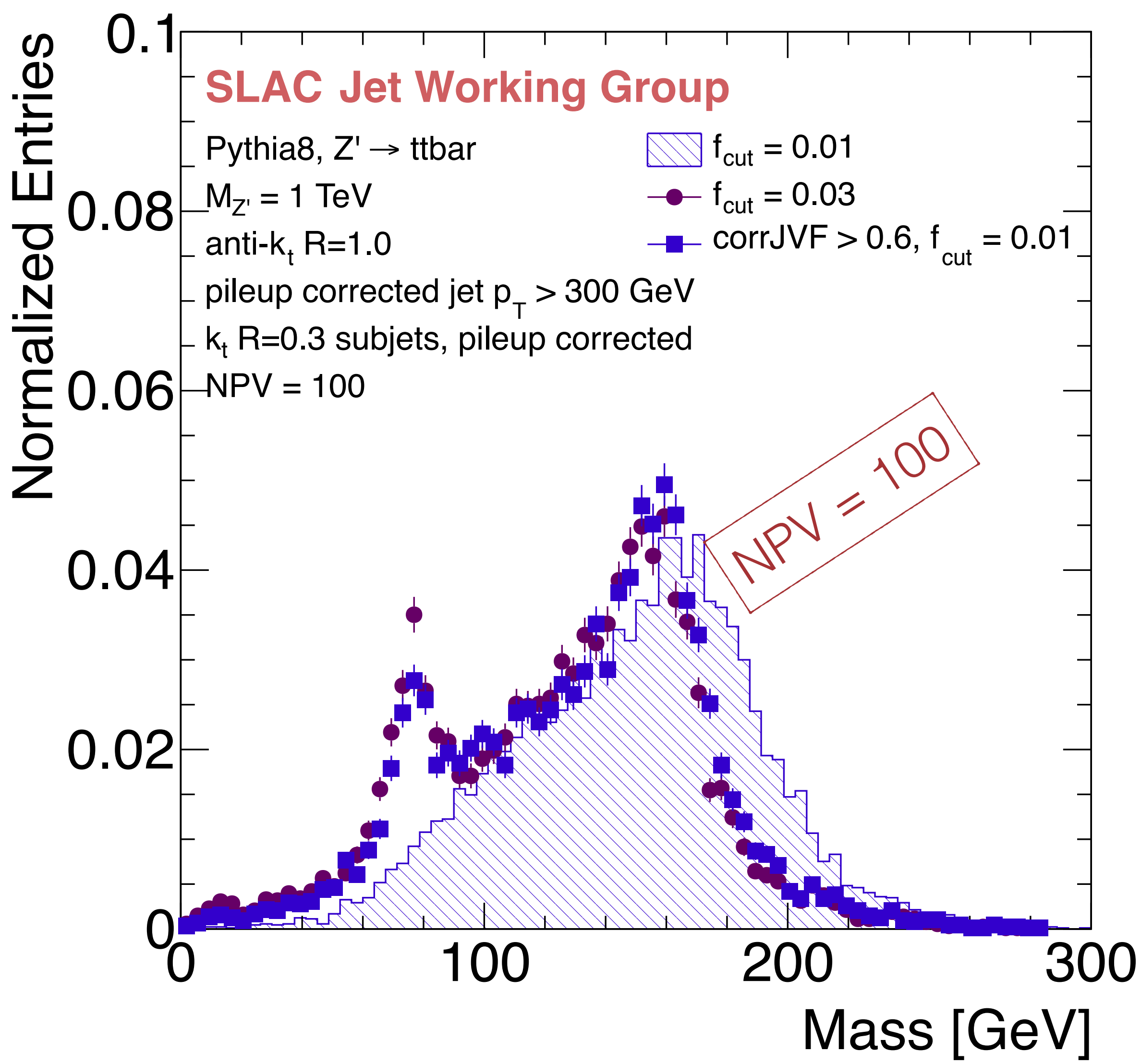


- in trimming, f_{cut} is defined as p_T^{subjet} (pileup corrected) / $p_T^{ungroomed}$
- f_{cut} = 5% is too tight for high N_{Vtx}
  - losing subjets from the hard-scatter

truth-level study for corrJVF grooming

- Only small fraction of subjets have corrJVF < 0.6
  - it seems it's not very likely to get a pileup subjet within the anti-k_t R = 1.0 jet …

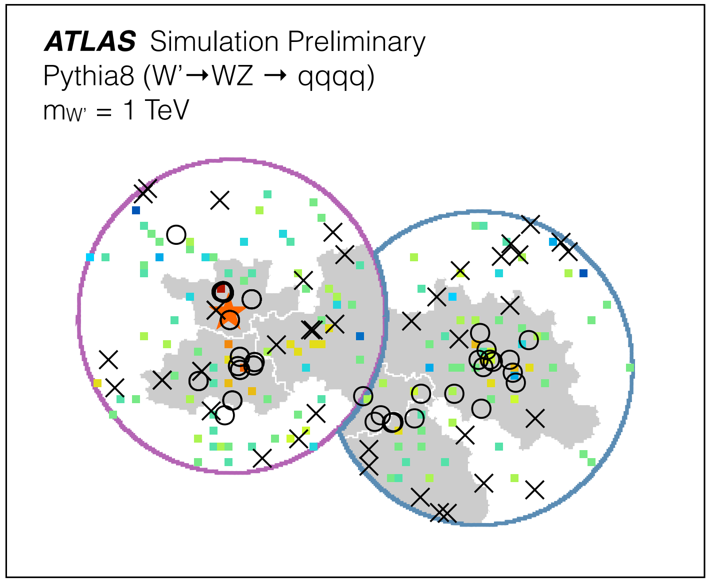


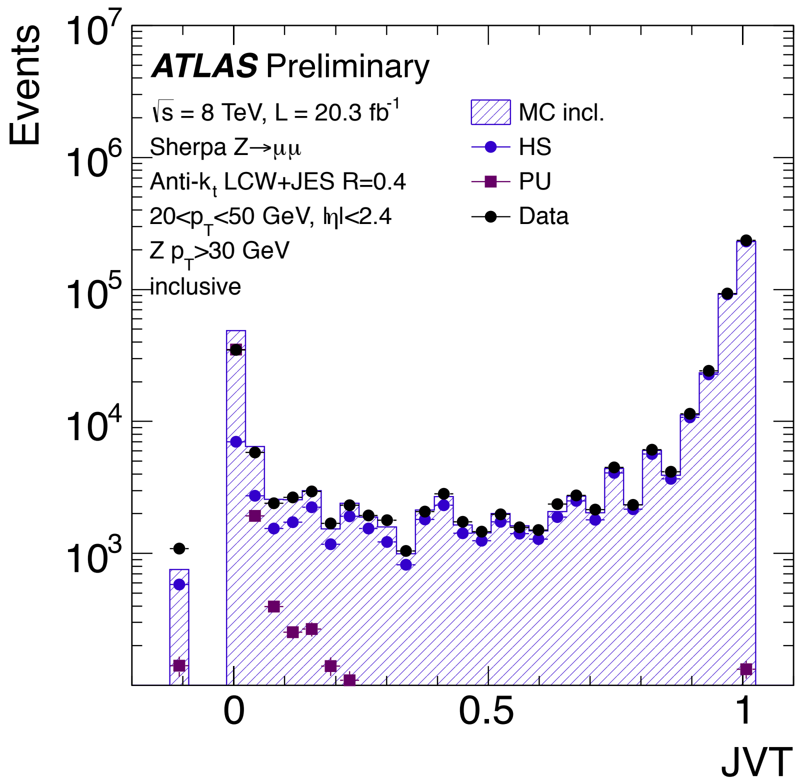


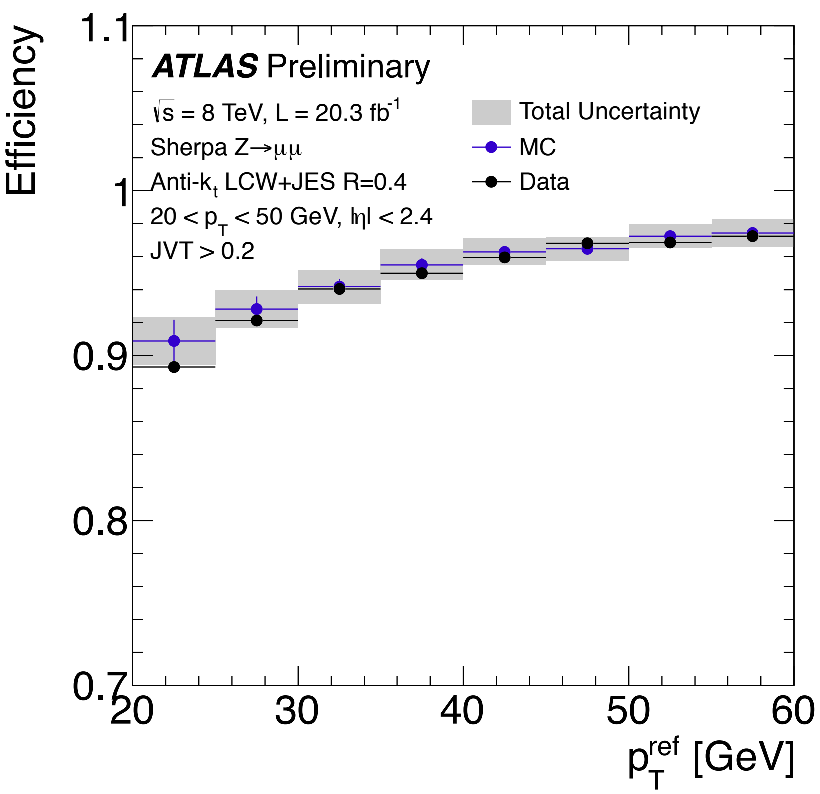
- cutting on corrJVF allows to lower the f_{cut}
  - improvement in mass resolution is marginal at N_{Vtx} = 100

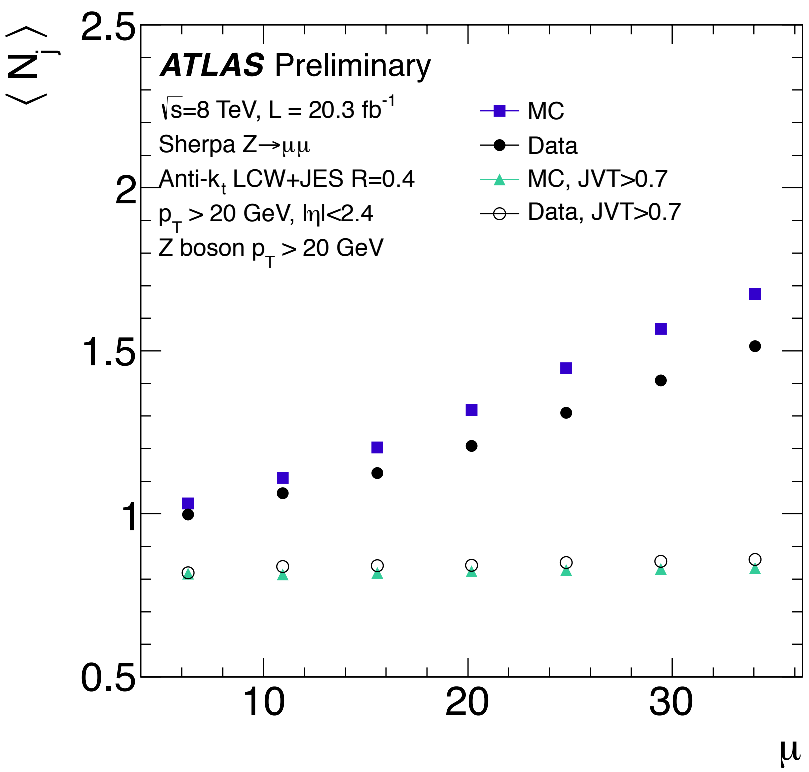
Summary

- New ATLAS CONF note on pileup jet suppression:
  - ATLAS-CONF-2014-018
- Key advantages of JVT-based PU jet suppression
  - resulting hard-scatter efficiency is N_{Vtx} independent
  - excellent performance (fake-rate vs. efficiency)
- corrJVF-based grooming:
  - rejecting pileup subjets allows to lower the trimming f_{cut}
  - performance improvement w.r.t. trimming is marginal
  - test-version of **fastjet contribution** is on svn (implementation & interface not yet stable…)









SLAC NATIONAL ACCELERATOR LABORATORY