Pile up Mitigation in ATLAS and CMS

Richard Polifka
University of Toronto

LPCC HL-LHC Workshop
11.5.2015 CERN

Outline

- What is pile up
- ATLAS and CMS
- jet reconstruction
  - ATLAS
  - CMS
- How to beat pile up
  - basic
  - complex
- HL-LHC environment

- warning - talk is targeting pile up suppression in light jets

Sources for this talk

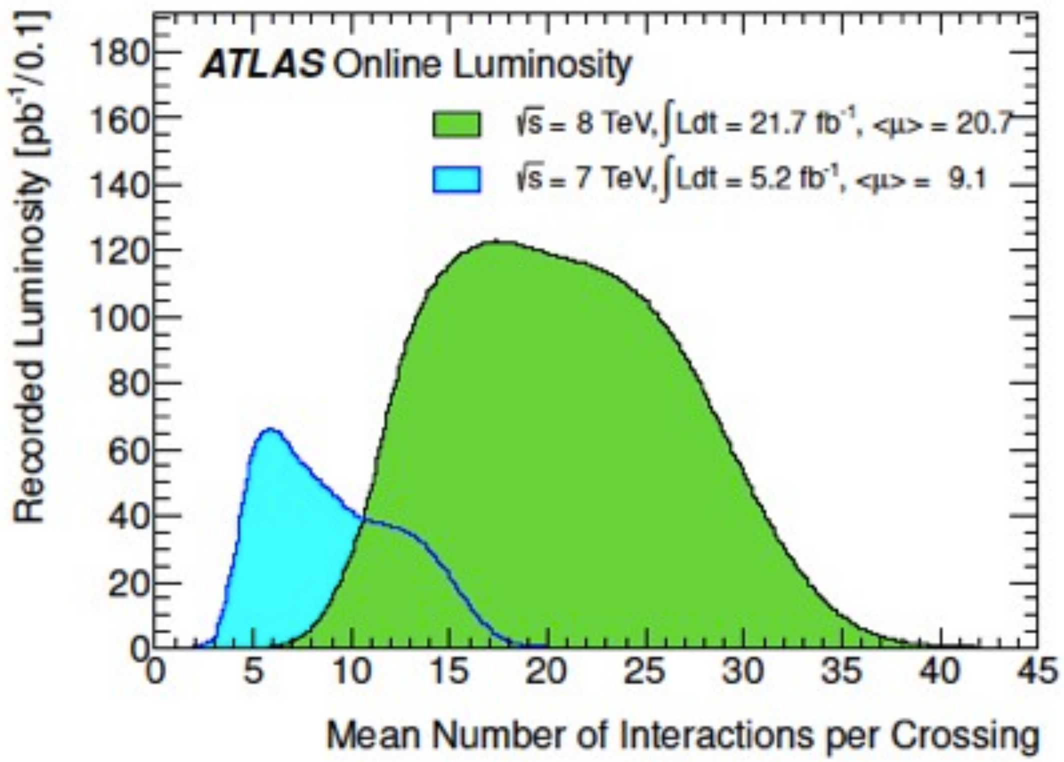
- **ATLAS:**
- https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2013HighMuPileup
- https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2013HighMuEtmiss
- http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2013-004/
- https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LargeEtaECFA2014
- http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2014-018/
- http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2013-085/
- http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CONFNOTES/ATLAS-CONF-2013-083/
- **CMS:**
- https://cds.cern.ch/record/1751454?ln=en
- https://twiki.cern.ch/twiki/pub/CMSPublic/PhysicsResultsFP/ECFA-CMSPublicResults.pdf
- https://cds.cern.ch/record/1247373/files/PFT-10-001-pas.pdf

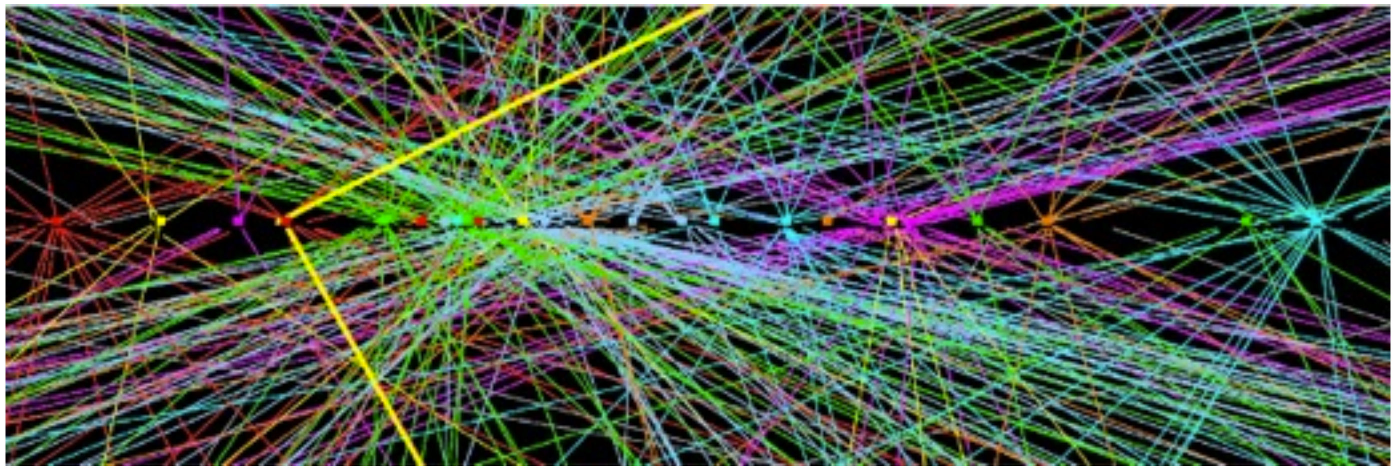
Pile up definition

- in the quest for more luminosity, LHC is delivering collisions with more than one interaction per bunch crossing (BC)
- in each event the vertex with max Σp_T^2 usually defines the primary vertex
- origins of **pile-up** jets (PU, general name for any additional activity coming from non-PV vertices):
  - in-time ... IT ... real (typically) QCD di-jet events emerging from non-primary vertex
  - out-of-time ... OOT ... energy deposit in calorimeters which is coming from other BC due to long readout time
  - stochastic ... random energy fluctuations combined by cluster algorithm

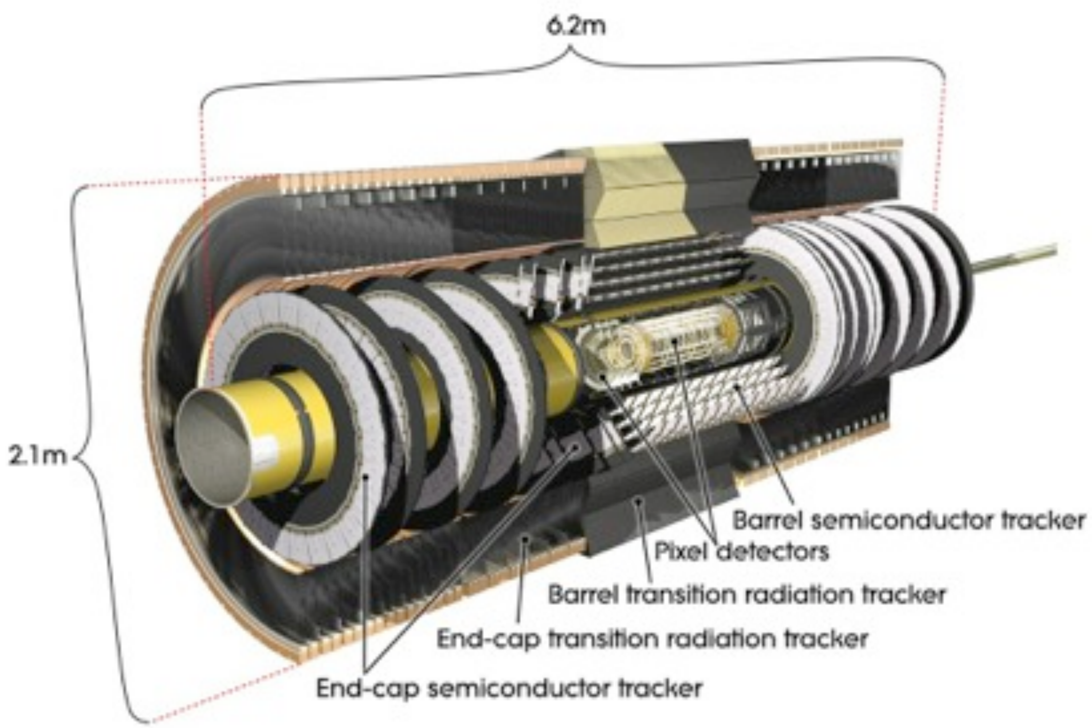
Pile up situation

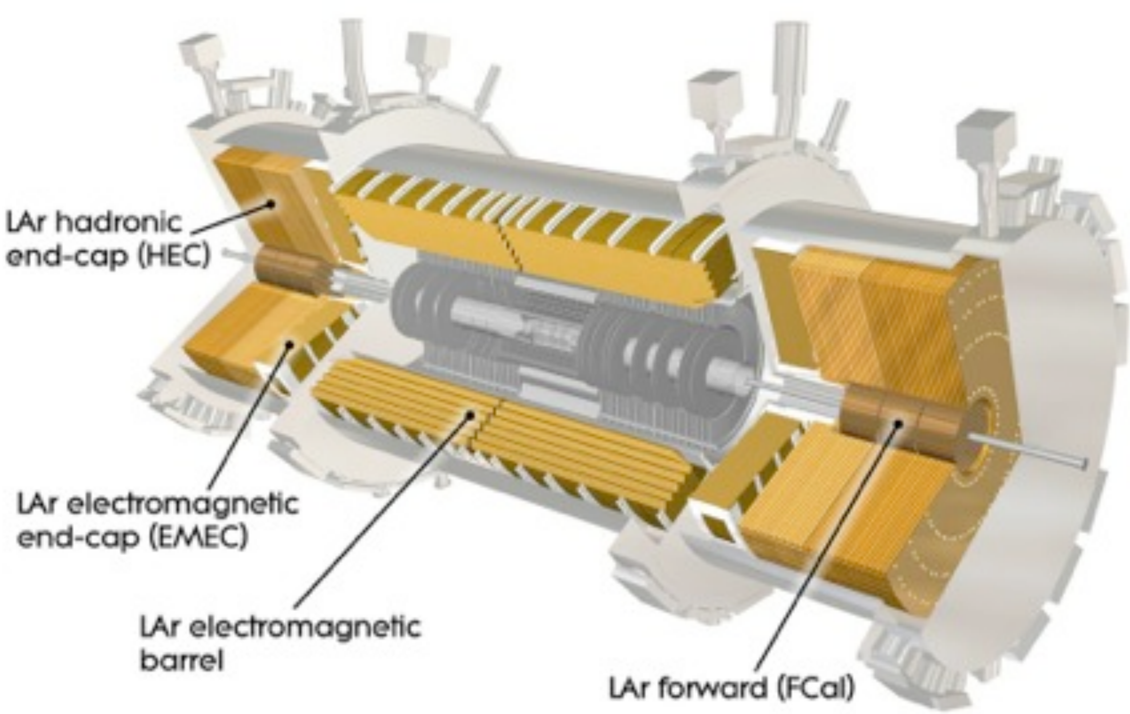
- μ … Average Number of Interactions per BC
  - sensitive to IT and OOT PU
- **NPV** … number of primary vertex "candidates"
  - only in one event -> sensitivity only to IT PU
- pile up through the years
  - **2011** - $\langle\mu\rangle$ = **9.1**
  - **2012** - $\langle\mu\rangle$ = **20.7**
  - **Run3** - $\langle\mu\rangle$ = **80** (used in projections)
  - **HL-LHC** - $\langle\mu\rangle$ = **140 - 200** (used in projections)





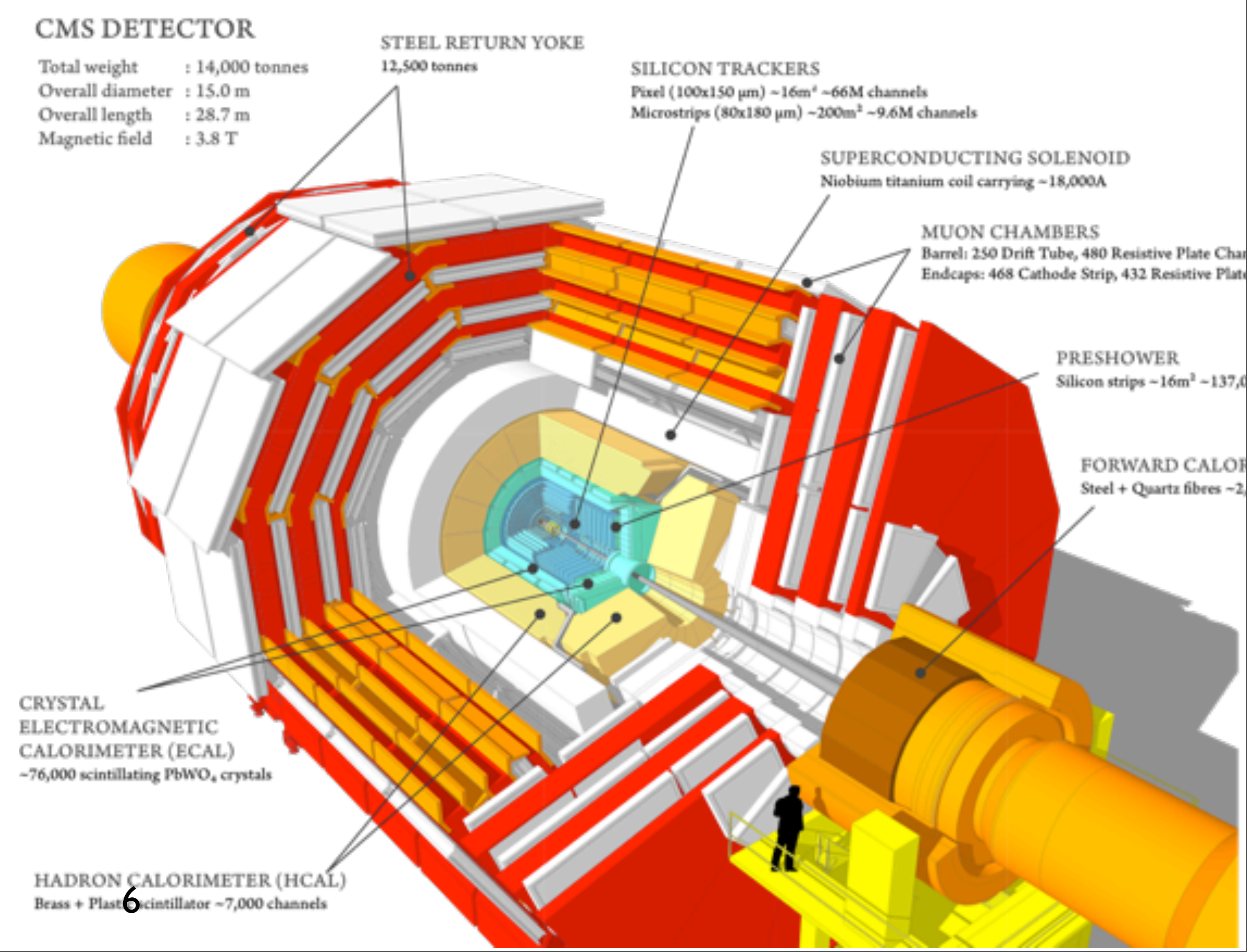
ATLAS and CMS





- ATLAS
- ID: Pixel, Silicon and TRT
- Calo: - EM (all) - LAr+Pb
  - Had: Tile - Fe+scint (plastic)
  - EC+FCAL - LAr+Cu/ Tungsten

- CMS
- ID:Silicon
- Calorimeter:
  - ECAL - crystals+scint ($PbWO4$)
  - HCAL - brass+scint (plastic)



Pile up Sensitivity

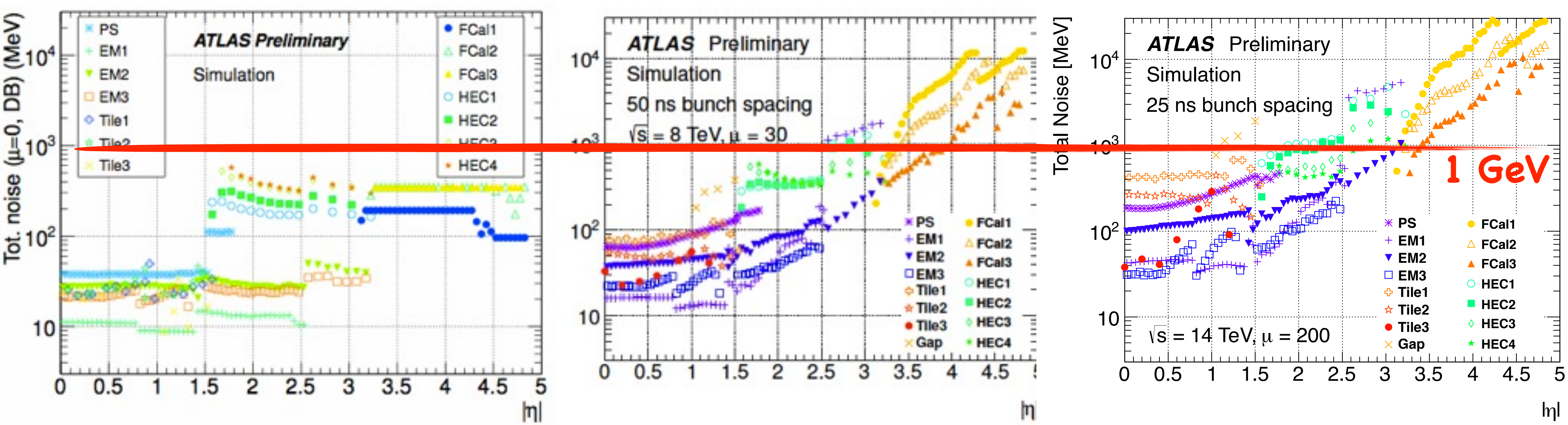
- ATLAS:
  - large pile up sensitivity for objects which are reconstructed in the calorimeter only (jets, photons, hadronic taus)
    - tracker information used subsequently for pile up mitigation
- CMS:
  - idea of particle flow combines measured objects (clusters and tracks) to the level of stable particles
    - cluster-level pu mitigation as part of the particle flow algorithm, subsequent usage of vertex association

from constituents to jets

- clusters -> jets -> jet energy calibration
- particle flow

clusters

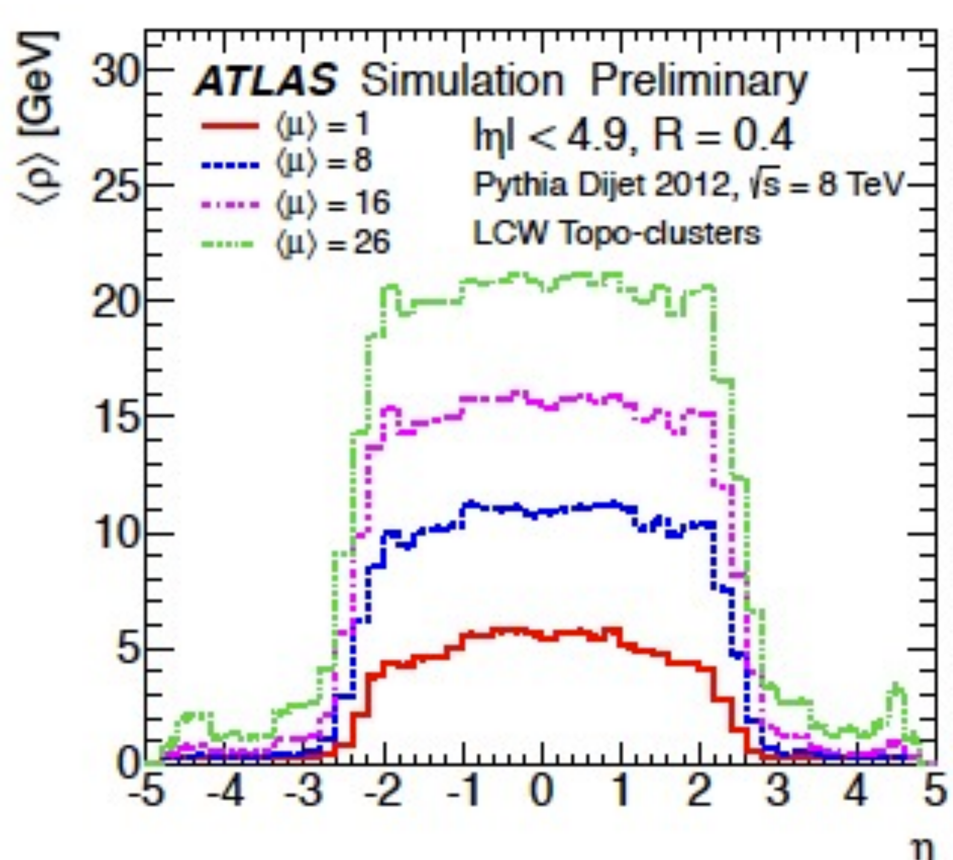
- topological clustering
  - based on nearest neighbor algorithm that clusters calo cells with energy above threshold with scheme $|E_{cell}|/\sigma_{noise}$ > 4 (seed) -> 2 (neighbors) -> 0 (additional layer)
  - $\sigma^2_{noise} = \sigma^2_{electronic} + \sigma^2_{pile\text{-}up}$
  - cell by cell, granularity and μ-dependent

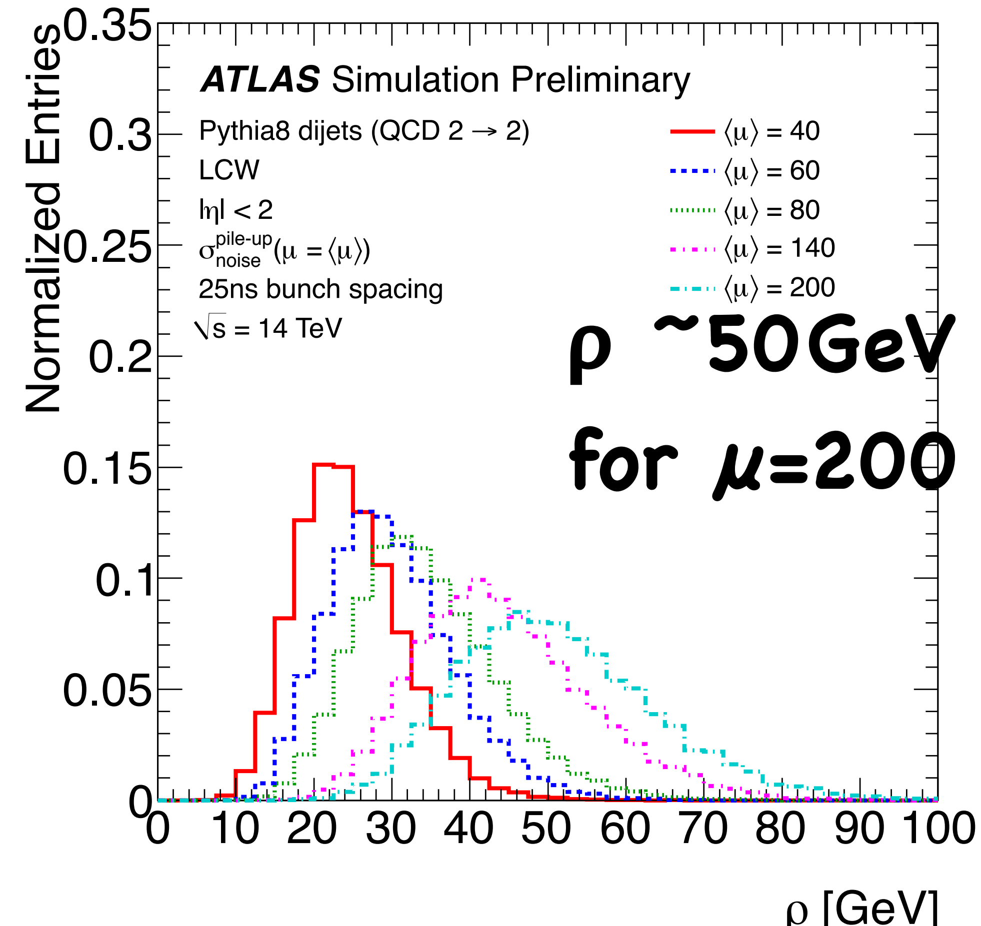


μ=0 (2009) μ=30 (Run1) μ=200 (HL-LHC)

jet energy calibration

- jets - AntiKt4, 6
- idea of jet real energy sitting on a "pedestal" caused by pile-up
    - is ~uniform in central region and can be subtracted
    - $p_T^{corr} = p_T^{jet} - \rho^* A^{jet}$
    - calculated event by event as median of distribution of density of many jets constructed with no p_T threshold



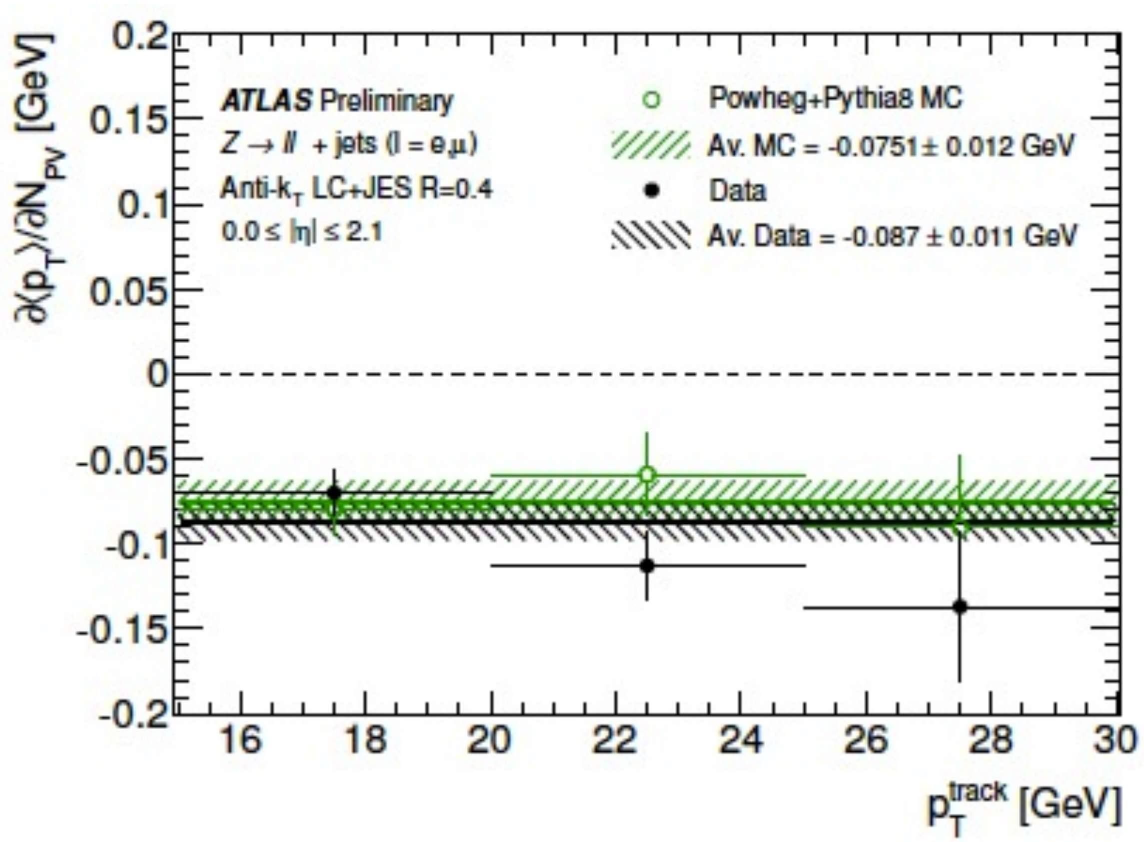


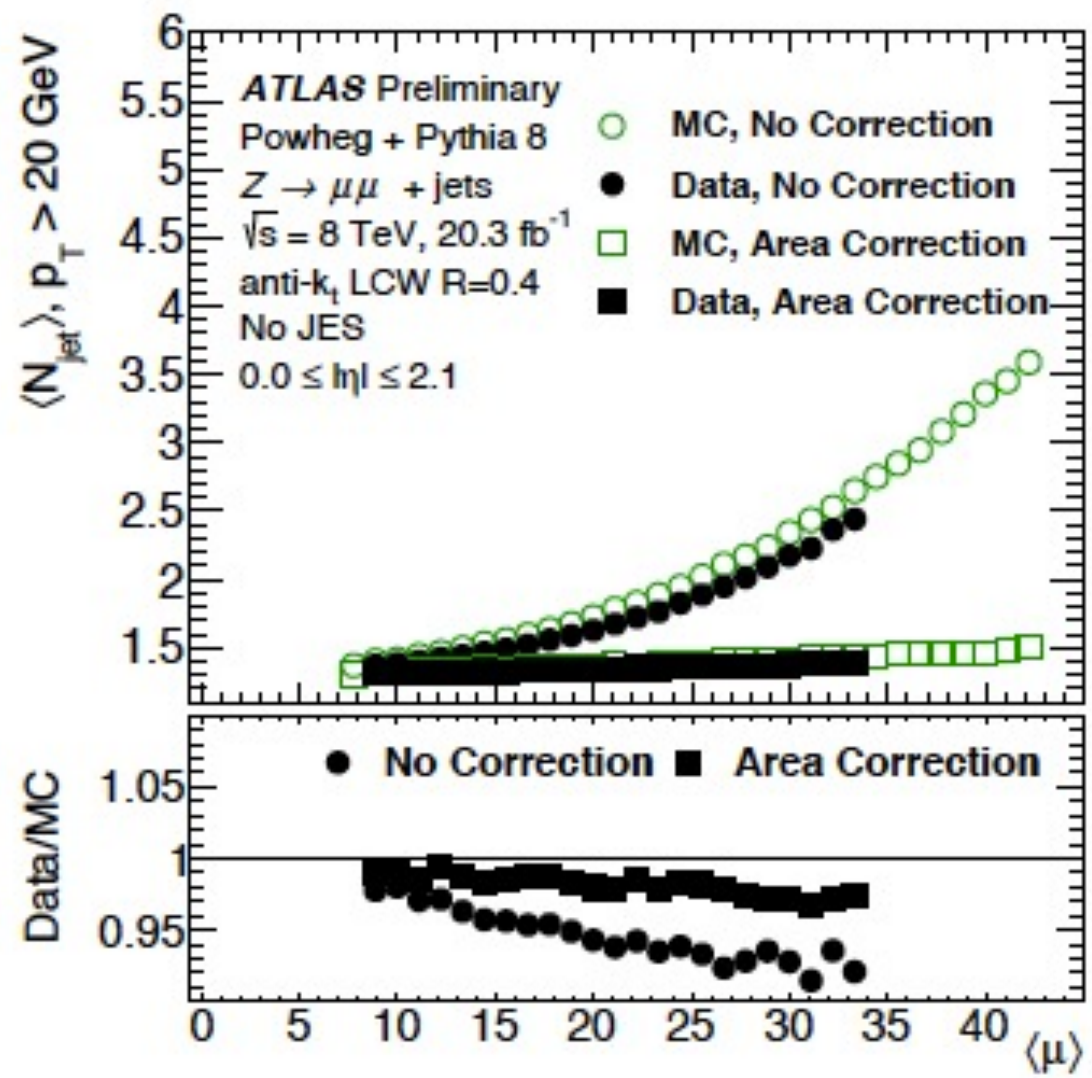
ρ ~50GeV for μ=200

at HL-LHC, p_T of 25 GeV (on average) will be subtracted

jet energy calibration

- after MC JES (E_{true}/E_{reco}), residual correction based on μ and NPV removes all PU sensitivity of JES
- validated with data - track jet (jets formed from tracks only) analysis on Z->ll events
  - final ~5% dependence goes to JES systematics

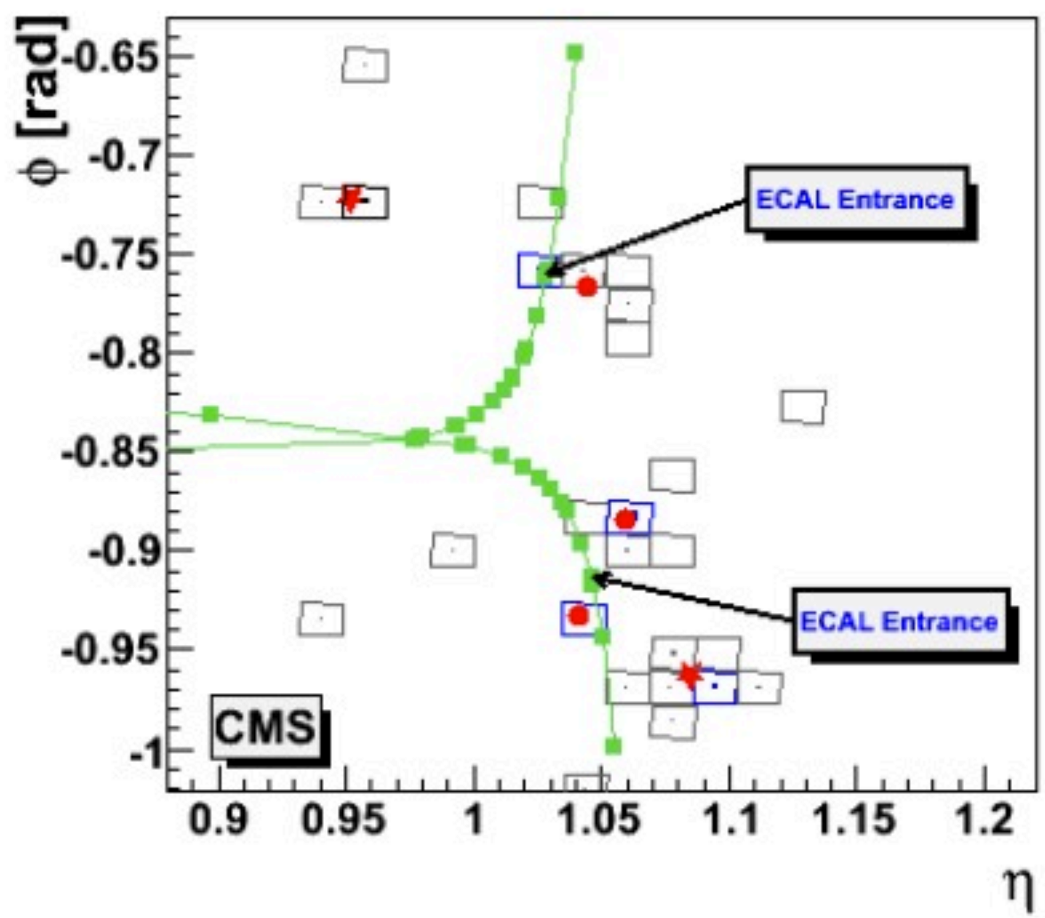


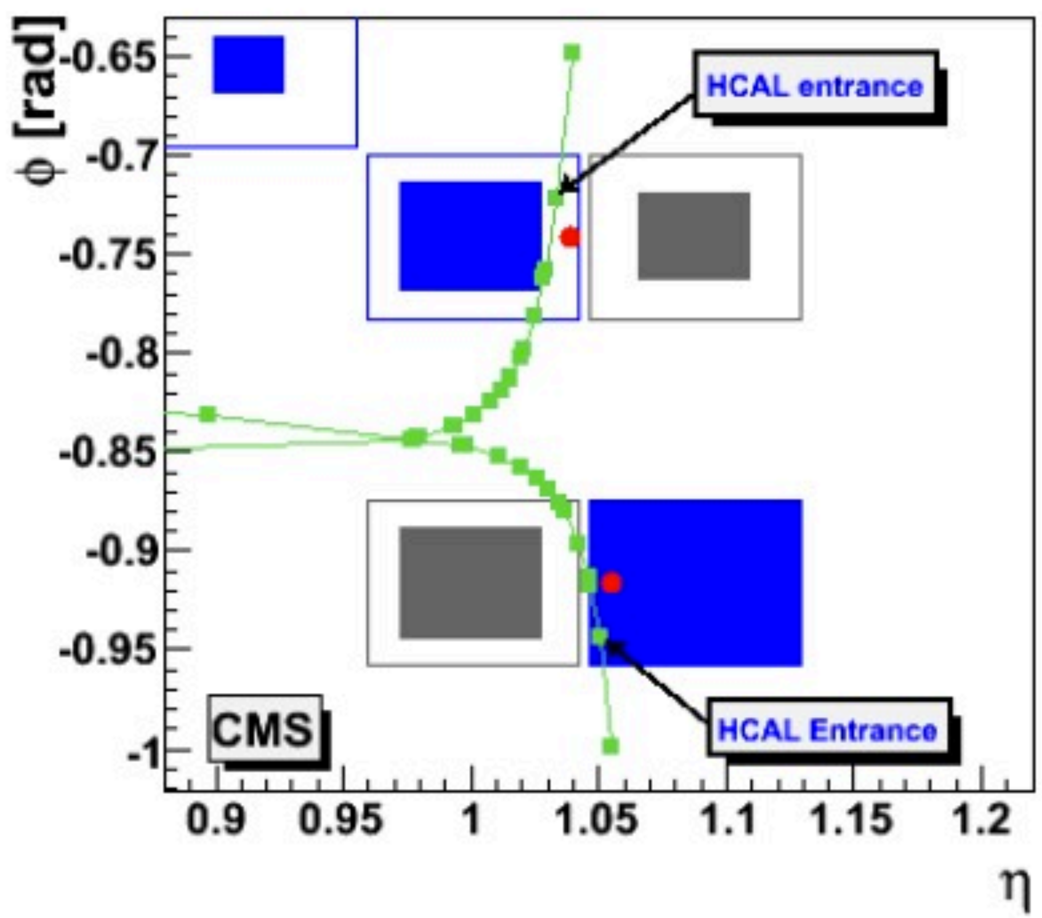


stable vs μ and NPV

particle flow

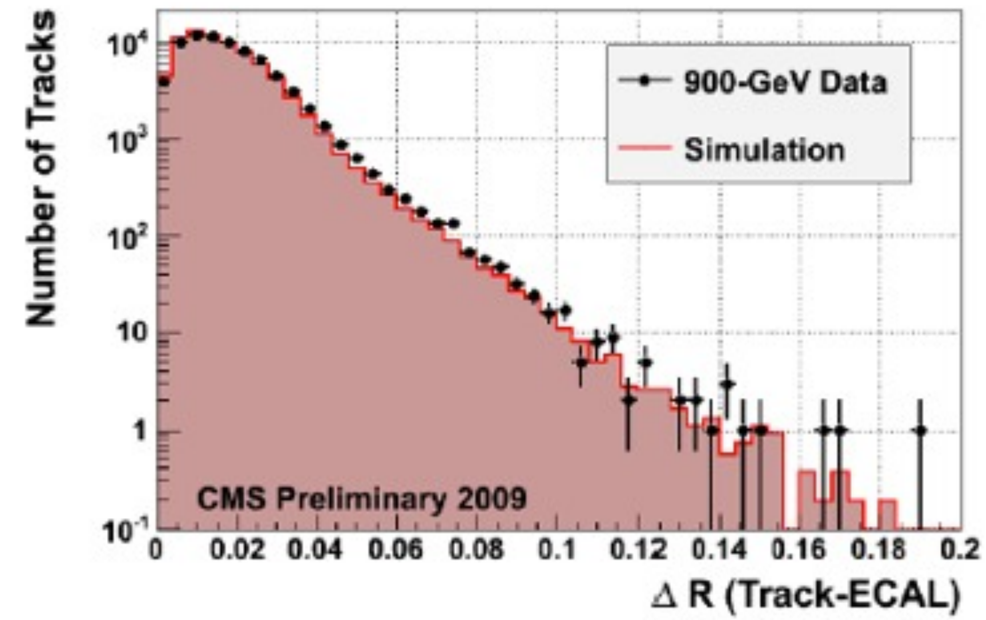
- track-cluster linking
- particle 4-vector reconstruction
- calorimeters cleaned for noise

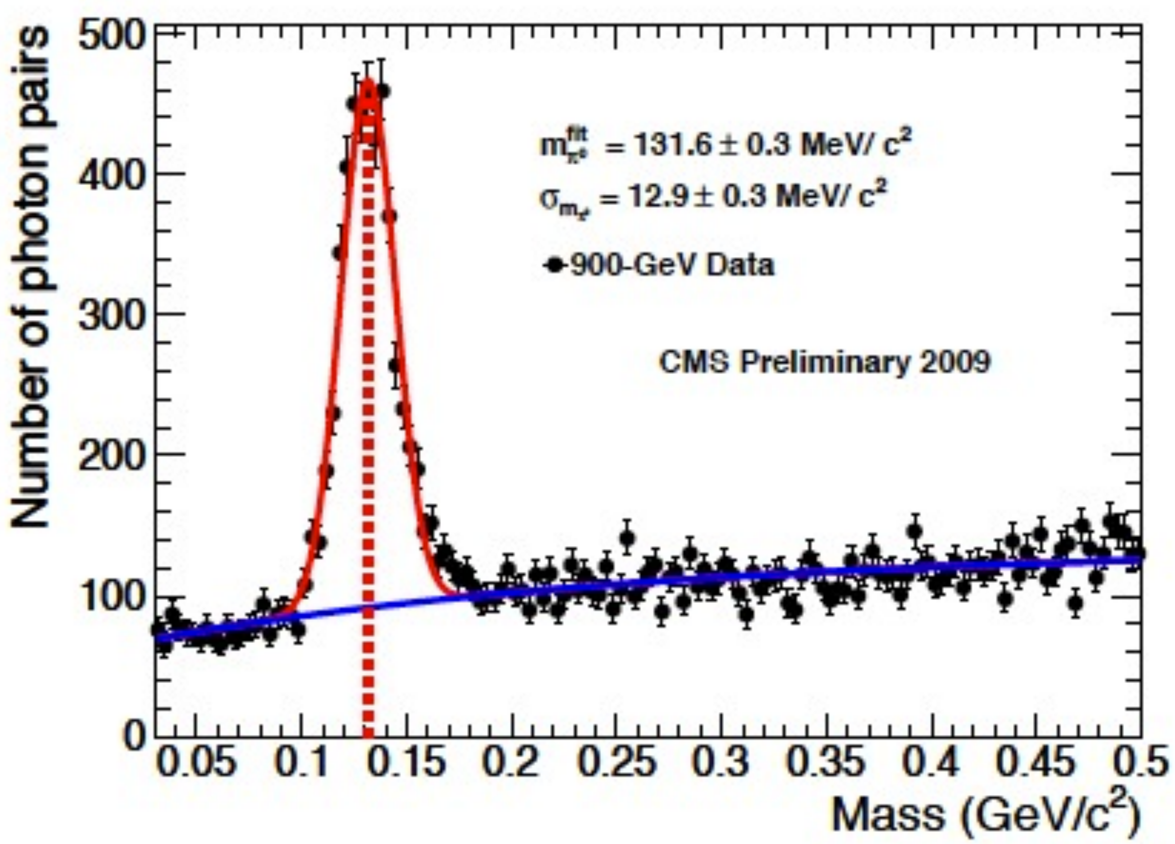


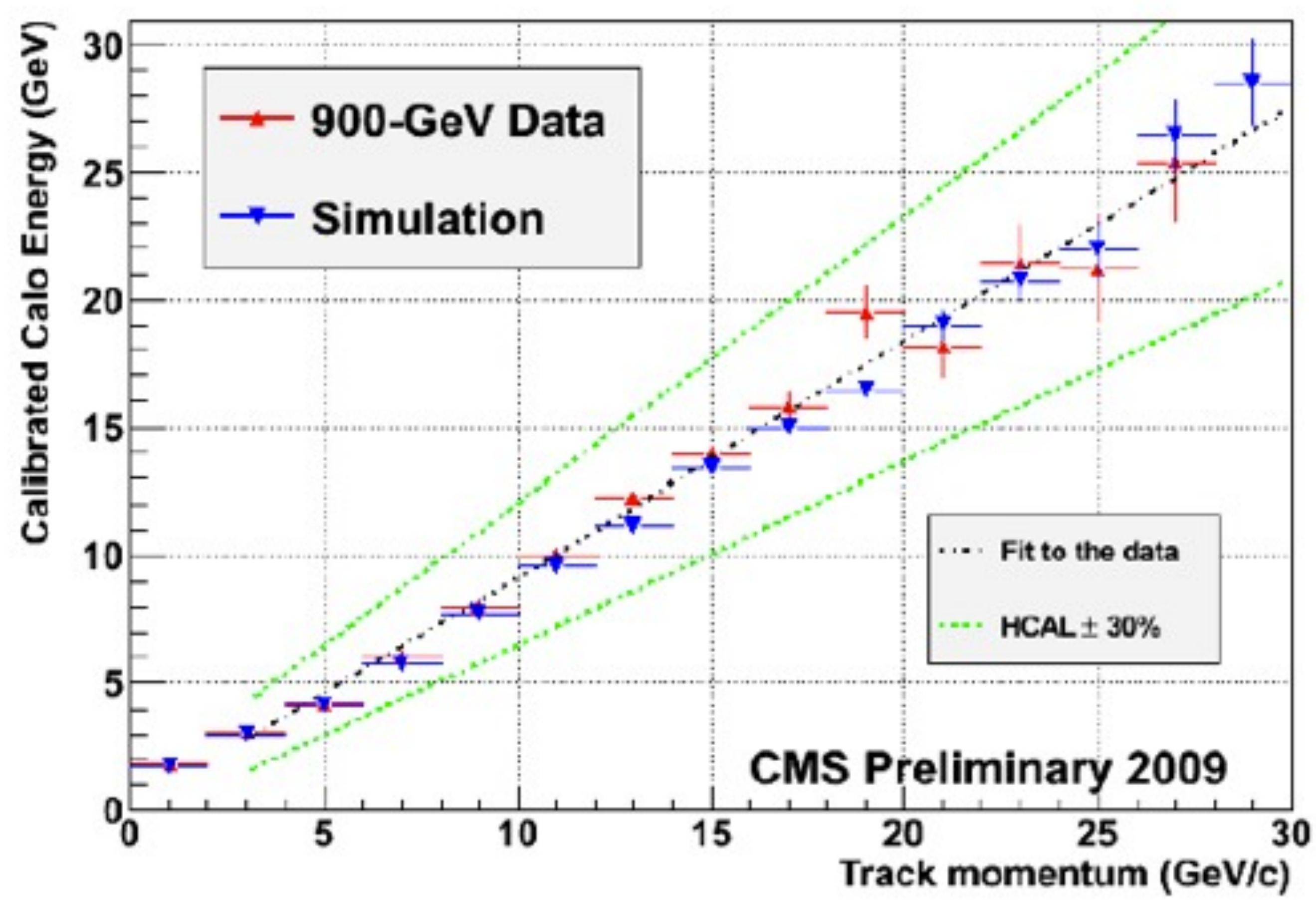


particle flow

- very good data-MC description of link variables
- even π^0 peak reconstructed in ECAL
- seeding of clusters:
  - EndCaps - p_T dependent
  - rest - p_T in-dependent







response within 1.5%

jets

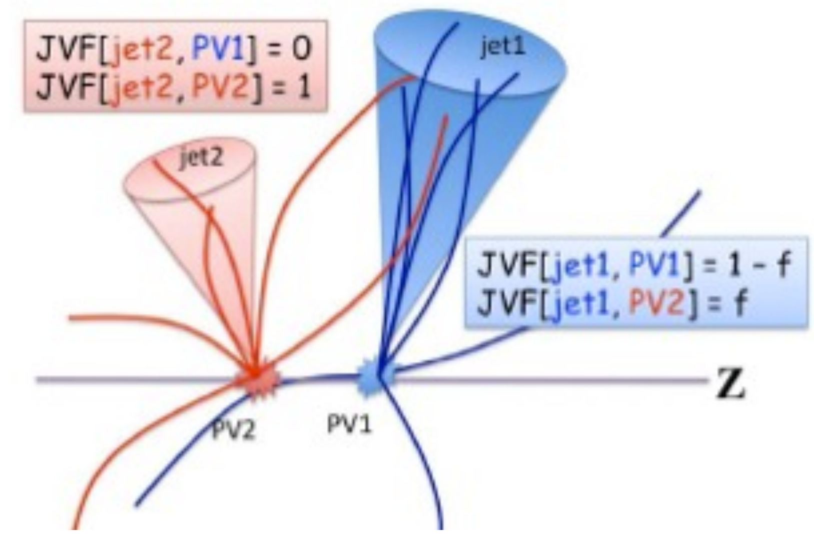
- charged hadrons - final calibration based on matching between tracks and clusters
- neutral hadrons - calibration taken from simulation once charged hadron calibration is validated
- jets: AntiKt R=0.3, 0.5 on particles from particle flow
  - charged hadrons corrected for zero-suppression and non-linearity of calorimeters (higher in forward region)
  - all available particles added to jet in the first step

vertex association

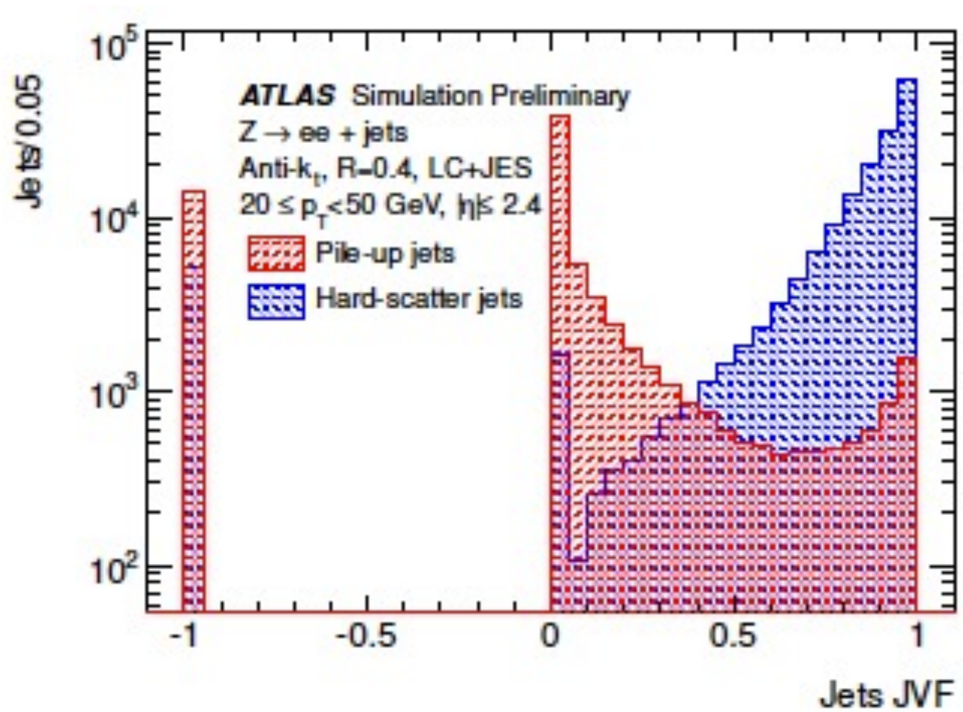
- **JVF -> JVT**
- **CHS**

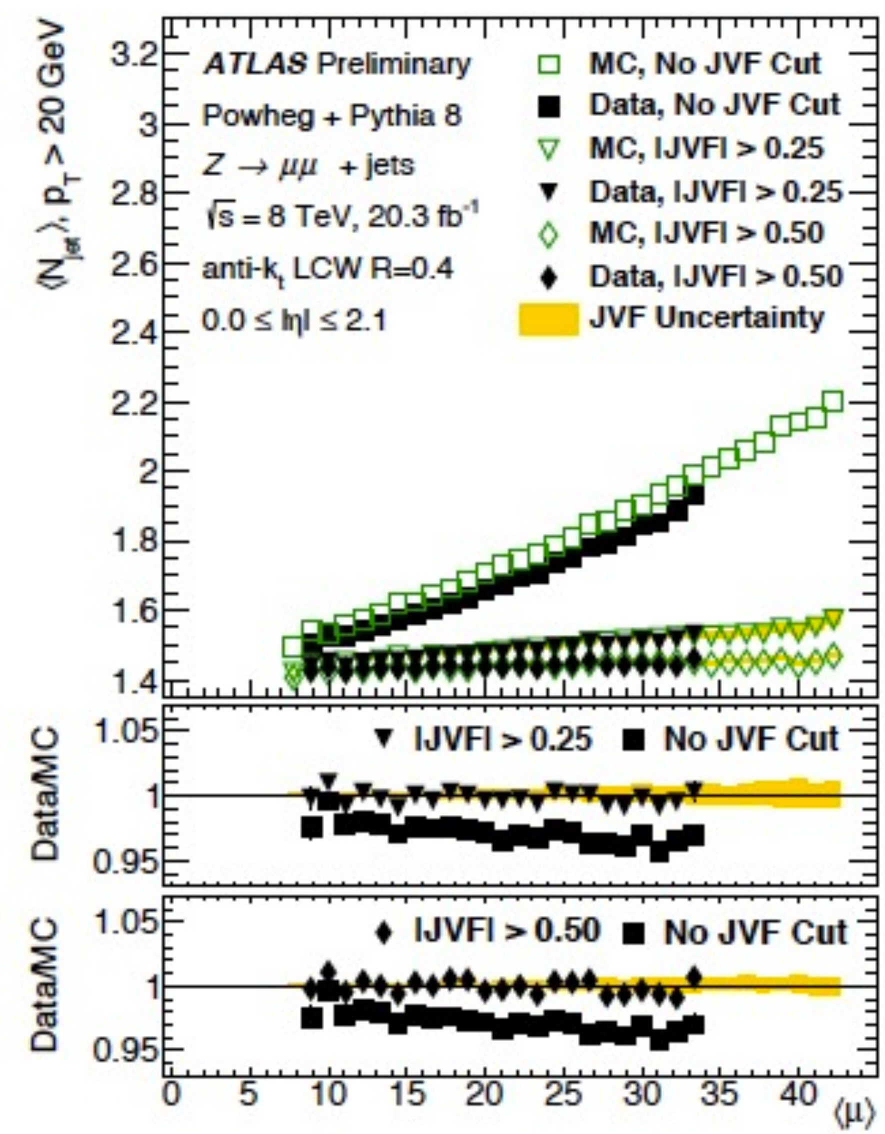
PU Mitigation - tracking confirmation

- some PU jets remain even after subtraction techniques, mainly genuine non-PV QCD dijets
  - -> usage of tracking and vertex association
  - Jet Vertex Fraction (JVF) in Run1



$$\mathrm{JVF}(\mathrm{jet}_i, \mathrm{PV}_j) = \frac{\sum_k p_{\mathrm{T}}(\mathrm{track}_k^{\mathrm{jet}_i}, \mathrm{PV}_j)}{\sum_n \sum_l p_{\mathrm{T}}(\mathrm{track}_l^{\mathrm{jet}_i}, \mathrm{PV}_n)}$$





PU dependent by construction - denominator increases with NPV

PU Mitigation - corrJVF & RpT

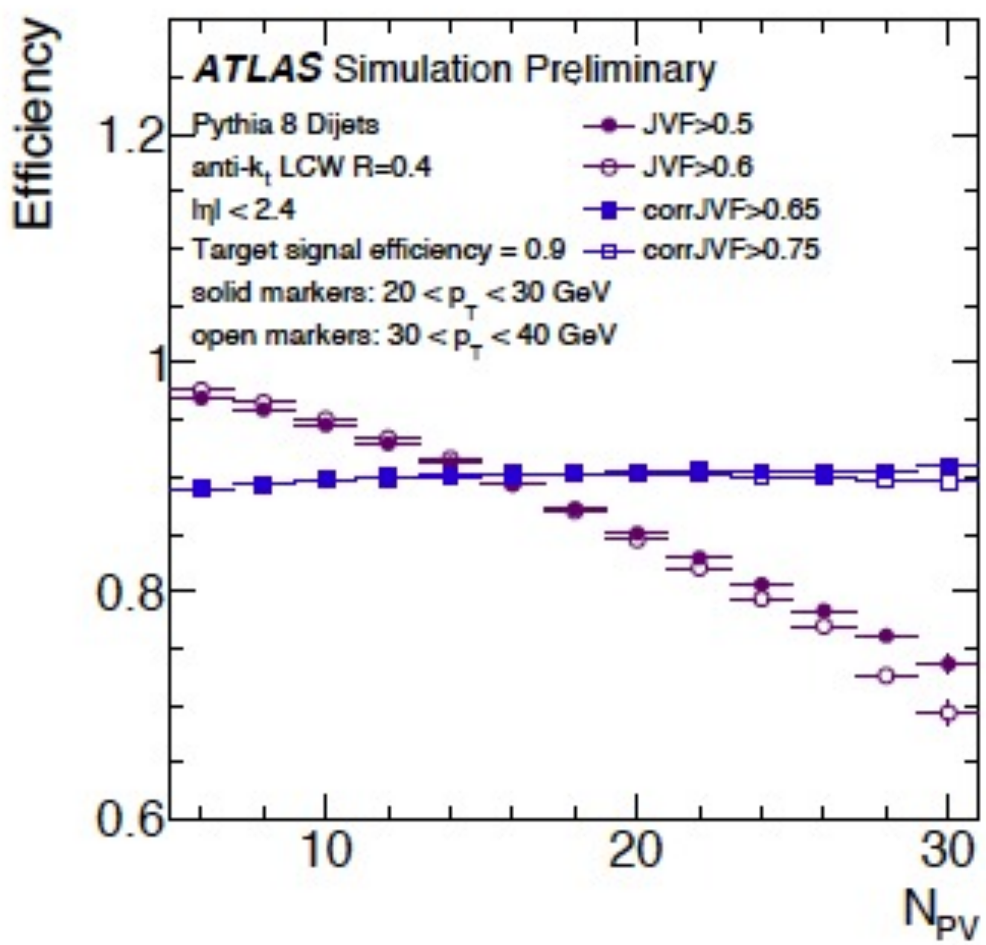
- NPV independent variables studied:
  - corrJVF and charged fraction (RpT)

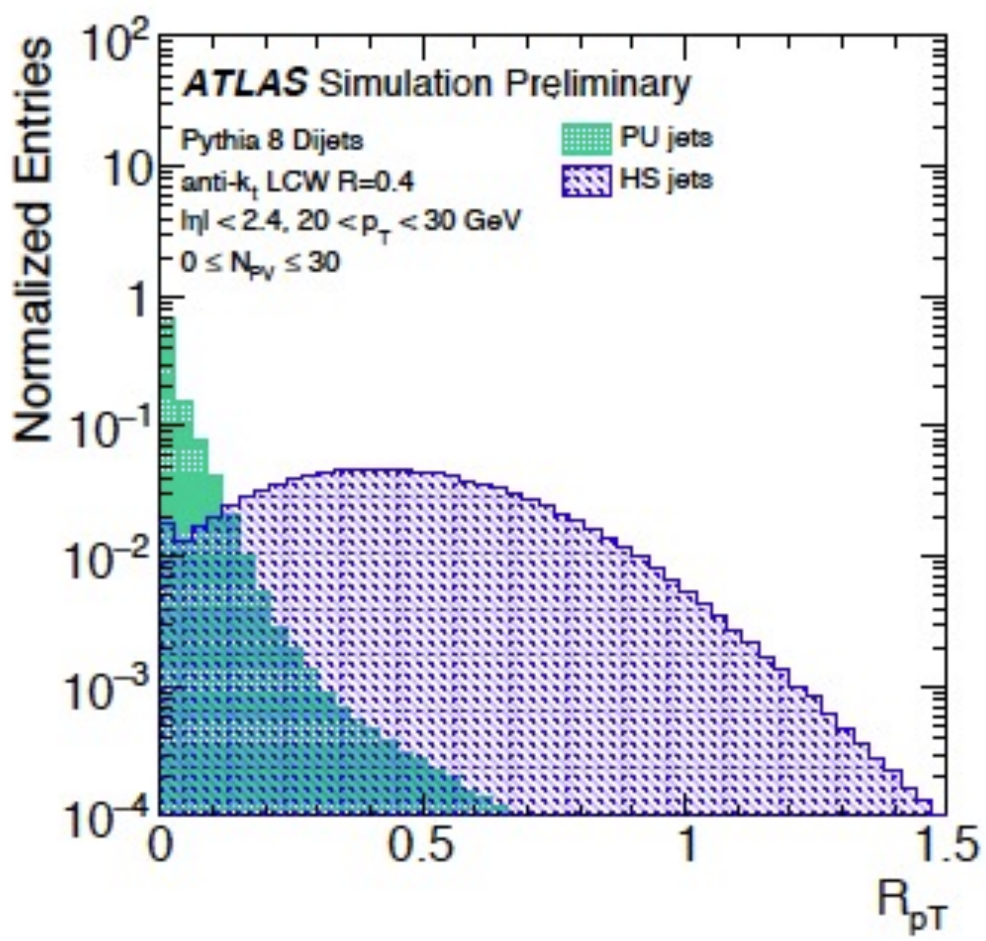
$$\mathrm{corrJVF} = \frac{\sum_k p_T^{\mathrm{trk}_k}(\mathrm{PV}_0)}{\sum_l p_T^{\mathrm{trk}_l}(\mathrm{PV}_0) + \frac{\sum_{n\geq 1}\sum_l p_T^{\mathrm{trk}_l}(\mathrm{PV}_n)}{(k\cdot n_{\mathrm{trk}}^{\mathrm{PU}})}}$$

non-PV dependence cancels, results k-independent

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

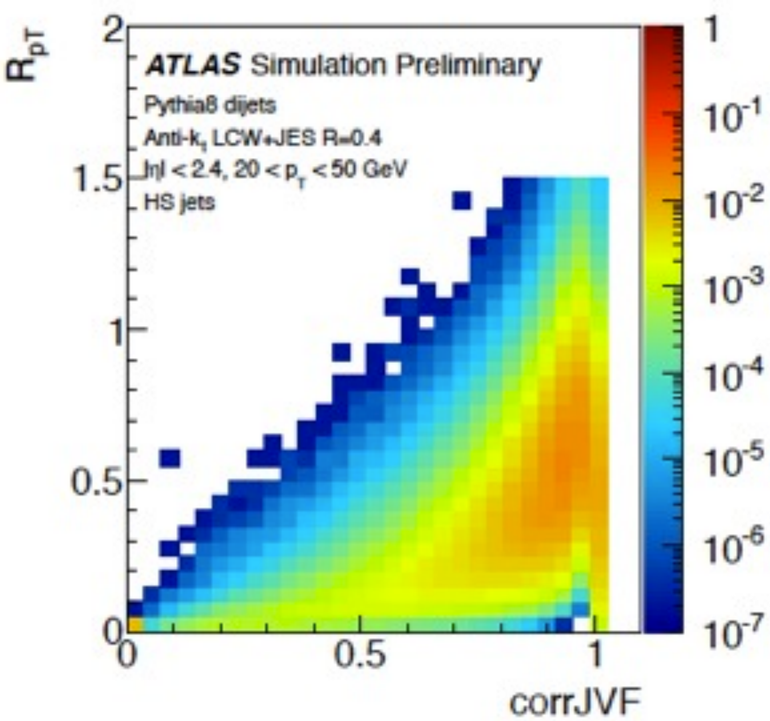
only with respect to primary vertex

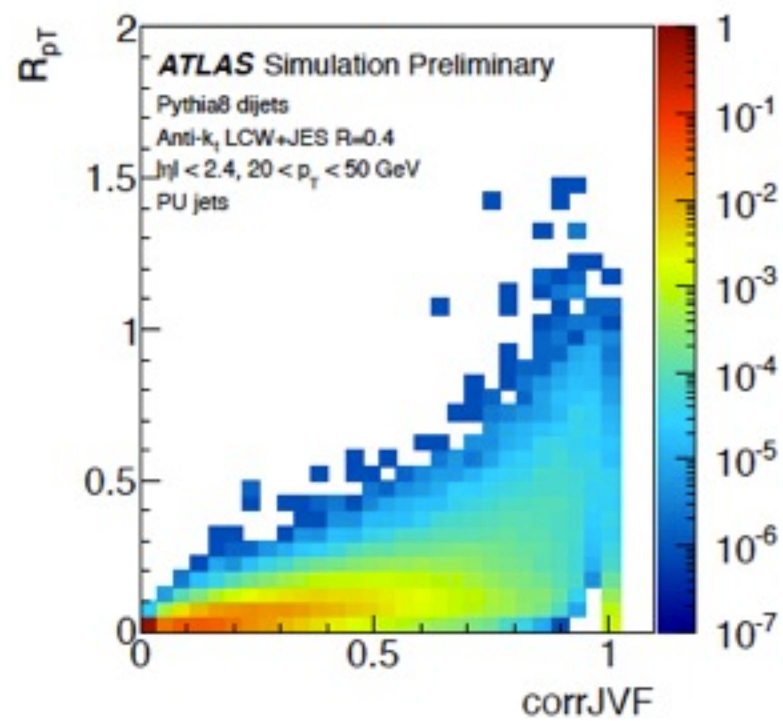


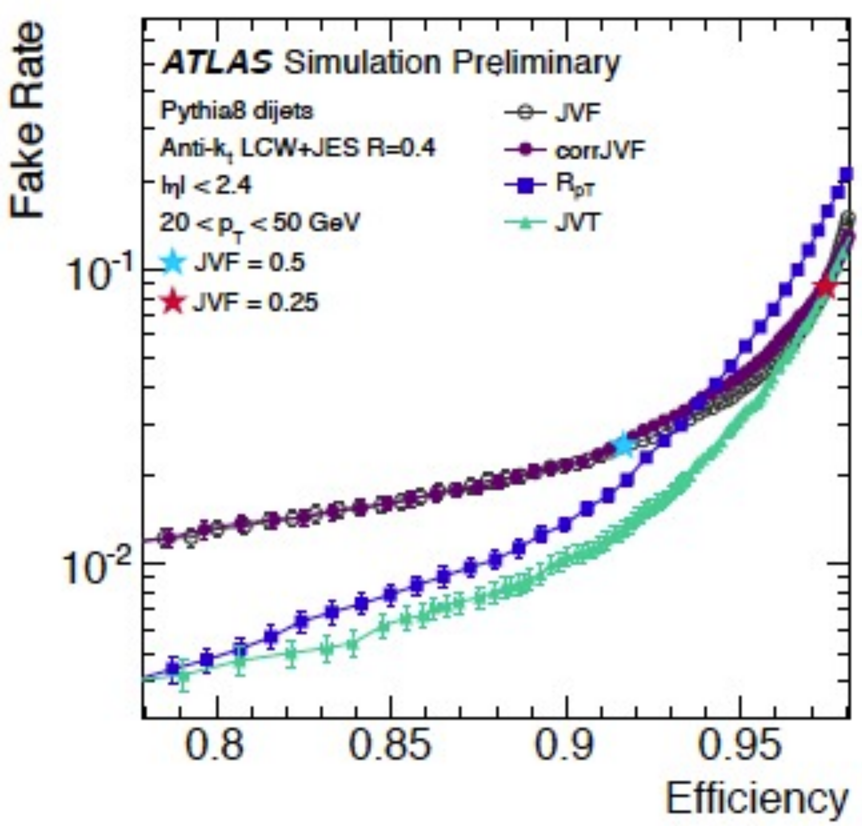


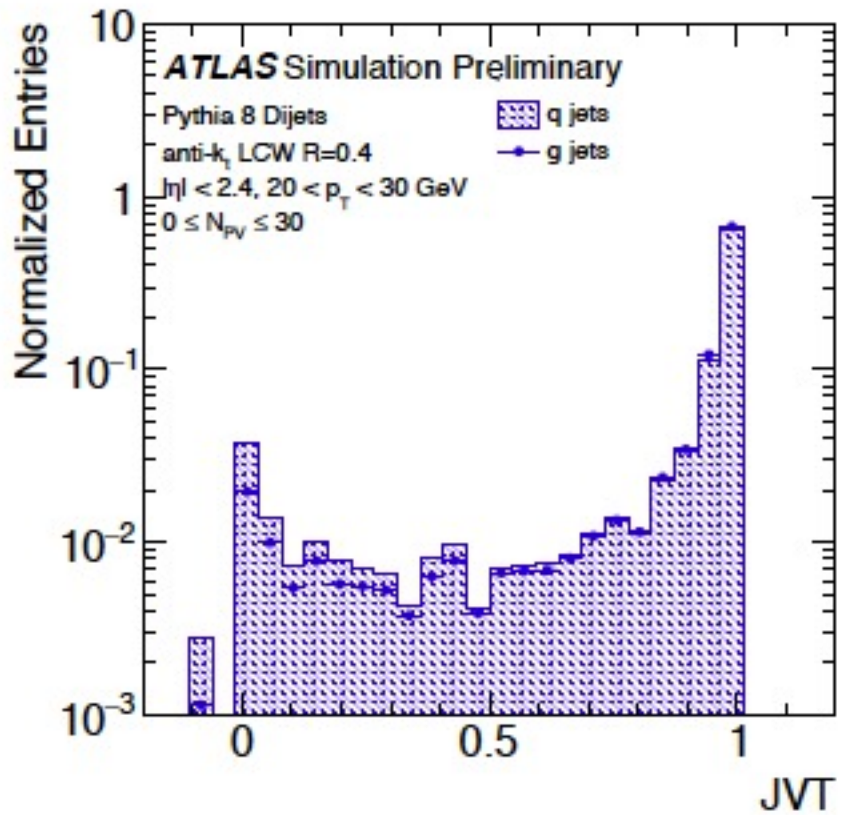
PU Mitigation - JVT

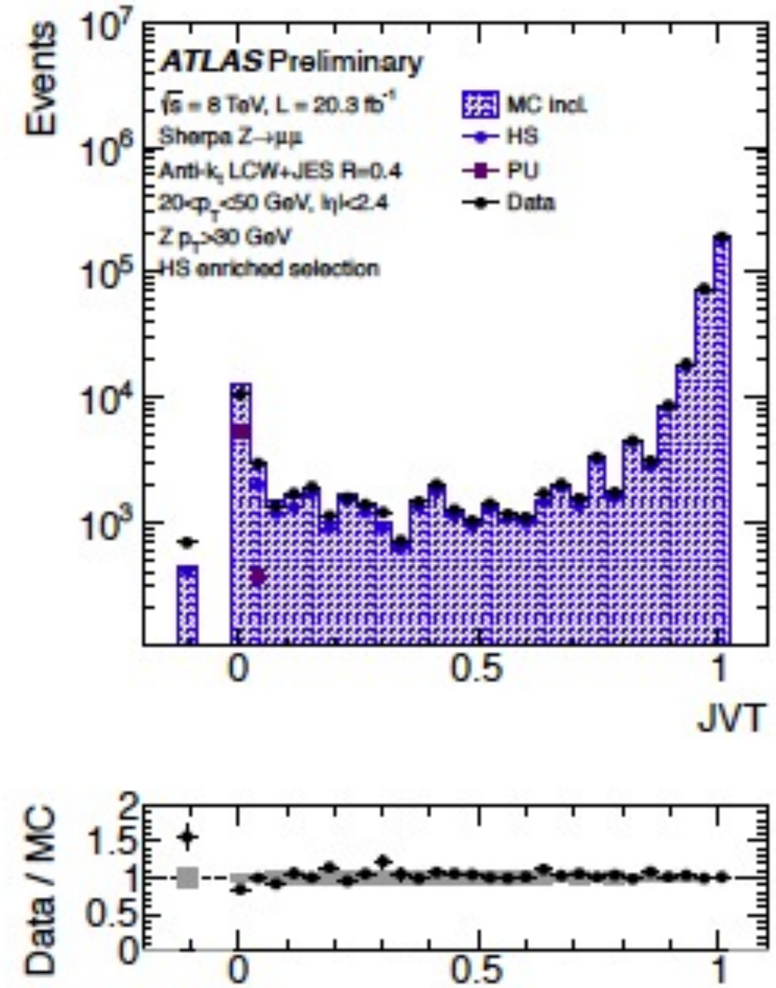
corrJVF and RpT form a 2D likelihood based on a kNN algorithm





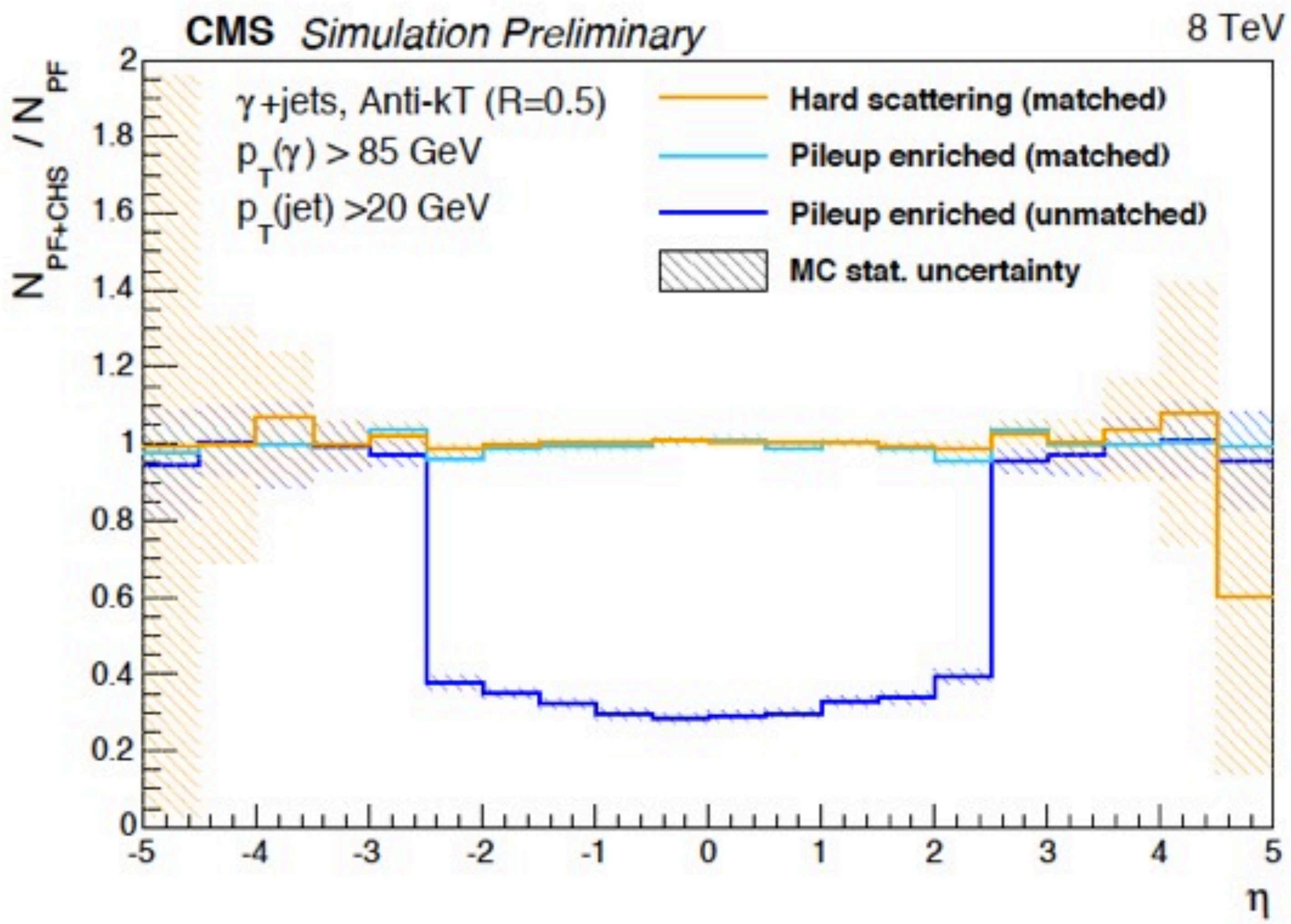






Charged Hadron Subtraction (CHS)

- removal of charged hadrons from the particle collection which is input for the jet algorithm
  - based on track-vertex association within $|\eta| < 2.5$
  - performance of PF only and CHS+PF compared
  - tracks not associated to any vertex are kept



hard scatter-matching (j-γ)

-> unmatched = "pile up"

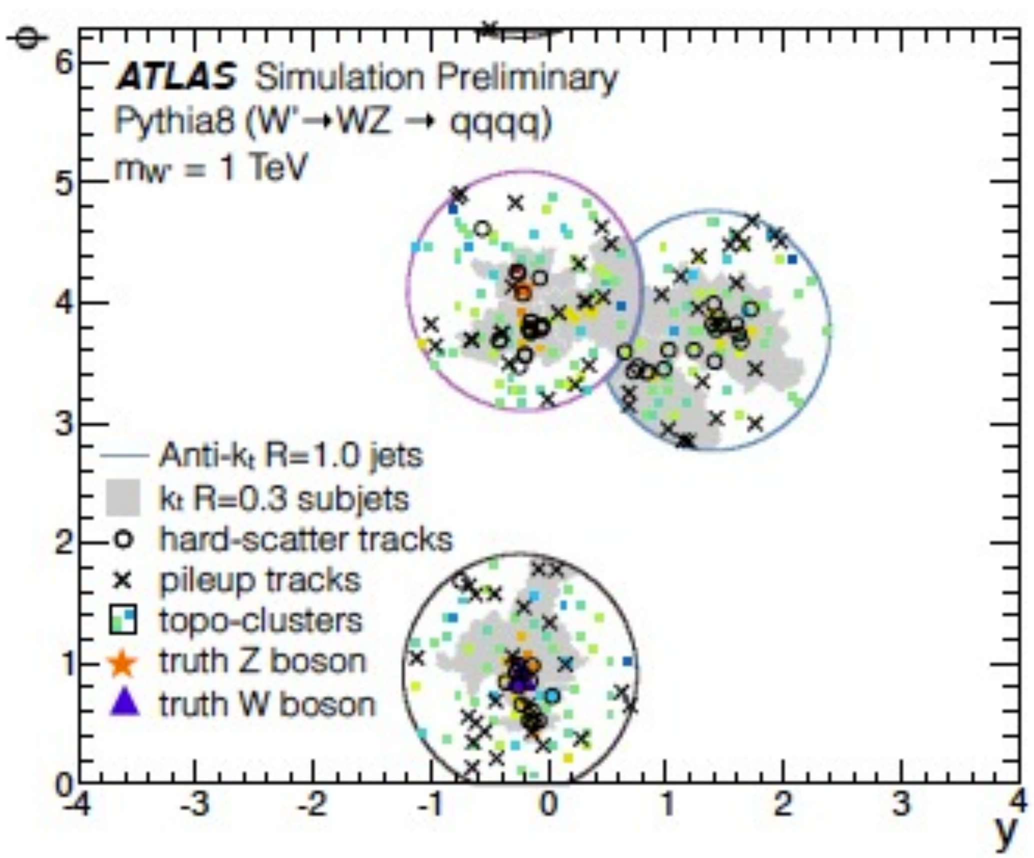
PU rate down by factor of ~3

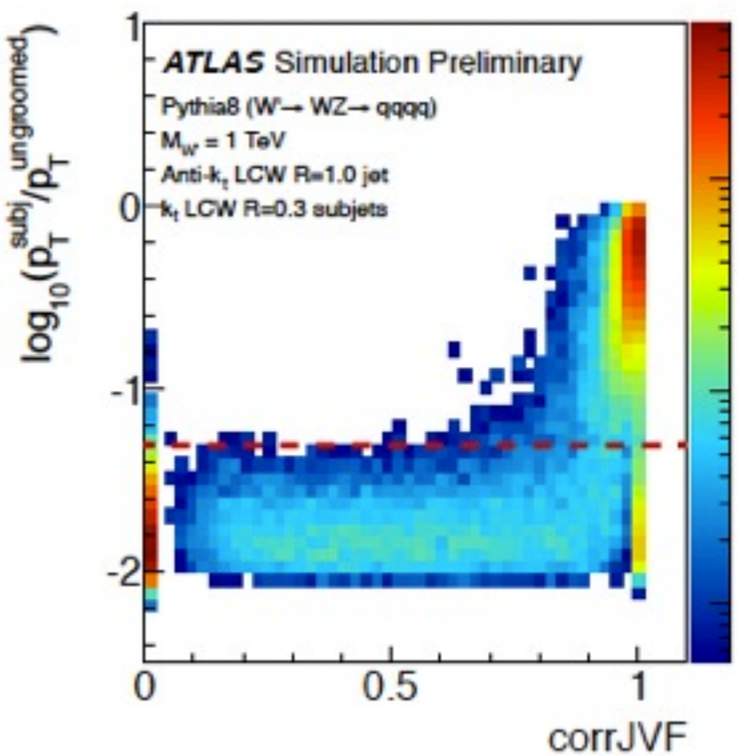
complex algorithms

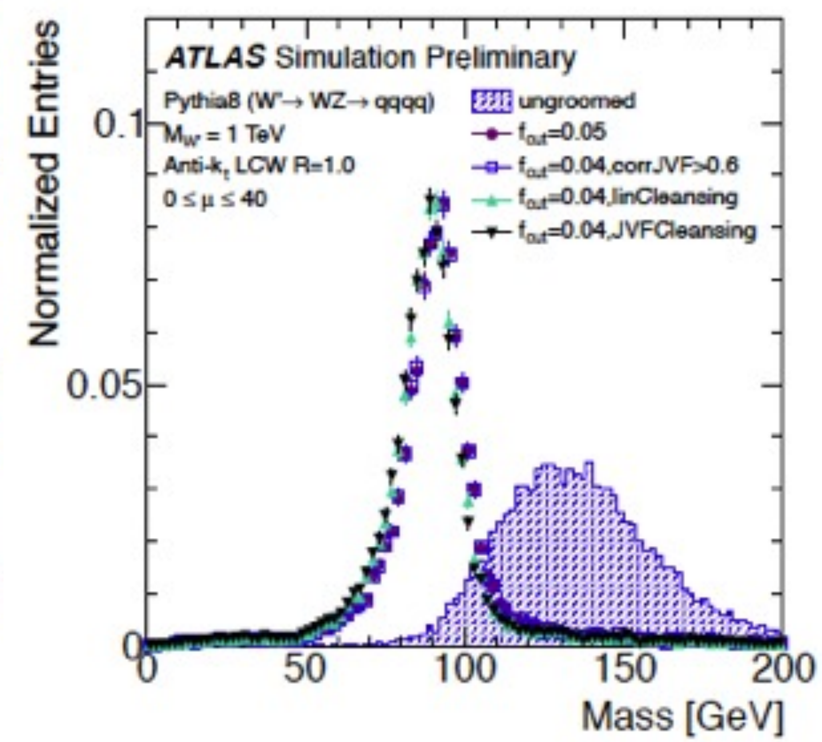
- cleansing
- pruning, trimming, soft drop -> constituent subtraction, linear cleansing and PUPPI

PU Mitigation - grooming

- algorithmic removal of substructures within a jet based on kinematic criteria
- mainly used for boosted analyses using “fat jets” (CA algo with R > 1.0)
- calo-based trimming and filtering successful so far - combine with tracking -> cleansing







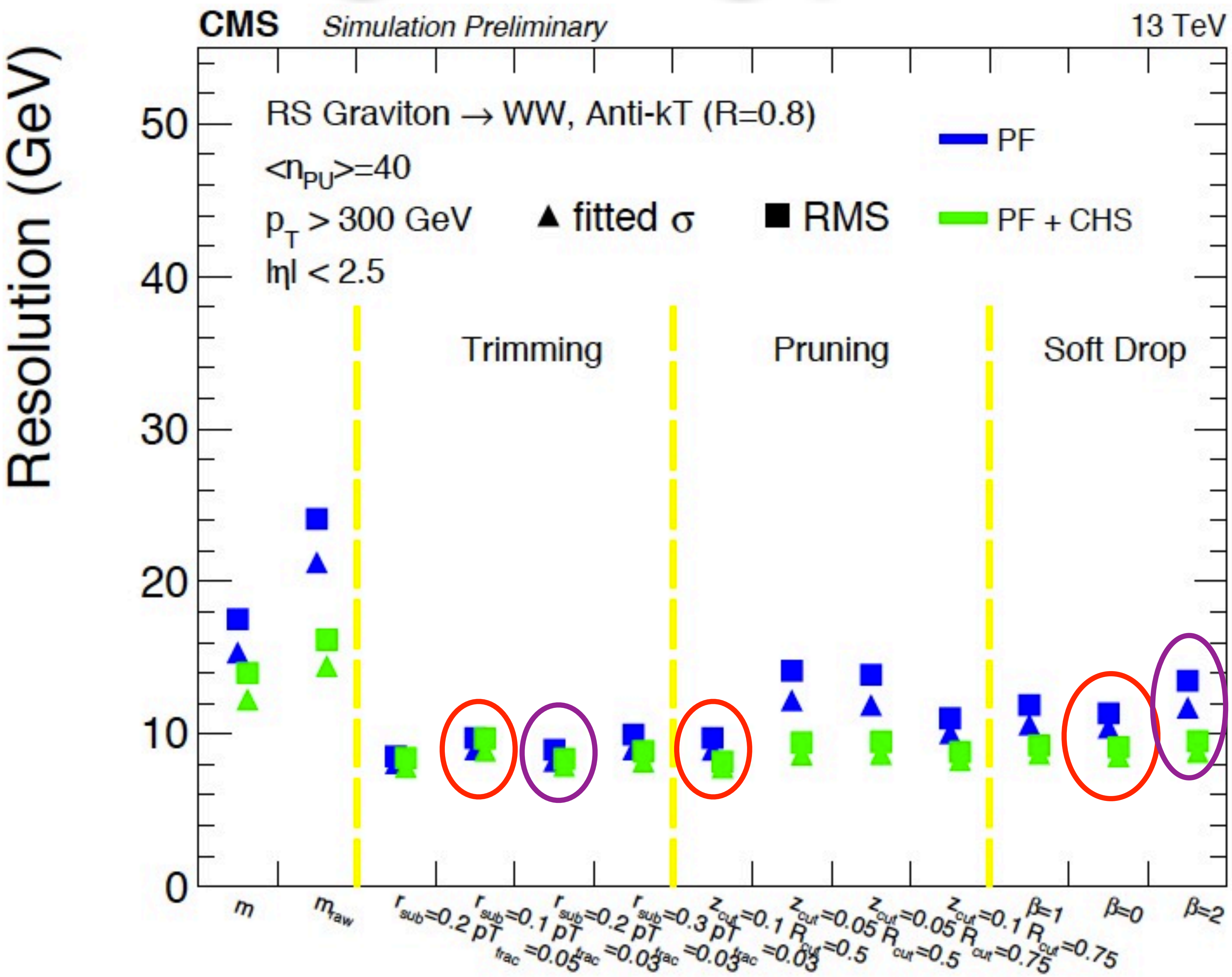
no subjet p_T cut

shall be more efficient in high PU environment

PU Mitigation - grooming

- removal of a subset of jet constituents
- to reduce jet mass dependence on PU for large jets (R > 0.8)
- alters the jet shape
- pruning: at each step, the softer from any particle pair is rejected if its momentum fraction (and distance) from the other is too small (large)
- trimming: removes particles below a dynamical p_T threshold
- soft drop/modified mass drop tagger: declustering of a jet, dropping sub-jets based on fractional p_T and size
- groomed jets are corrected through "safe subtraction" - subtract pile up p_T density

grooming performance



m … only "safe subtraction"

m_{raw} … no correction

aggressive groomers (in terms of mass reduction)

less aggressive

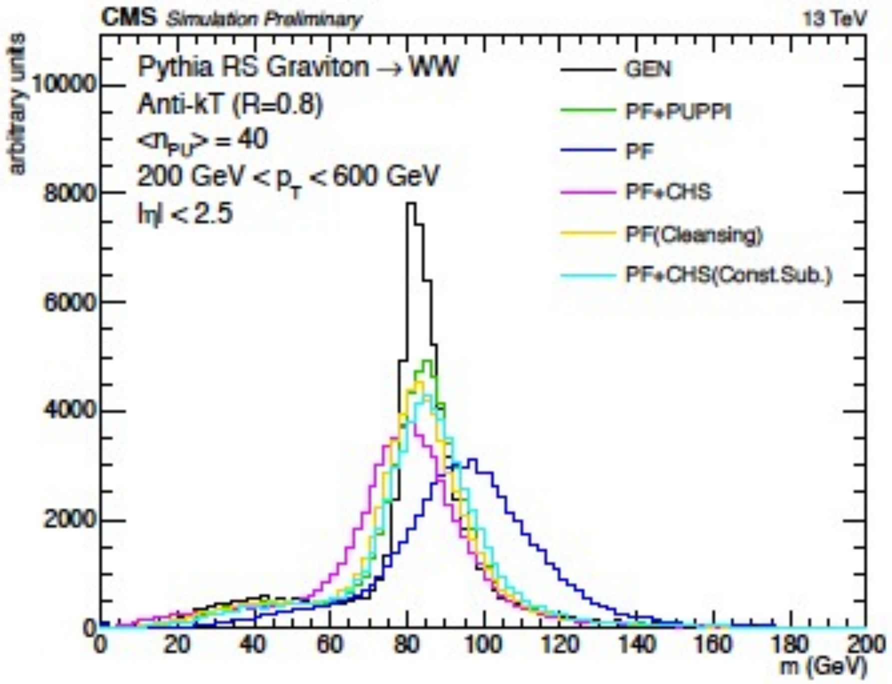
depends on physics case (in searches)

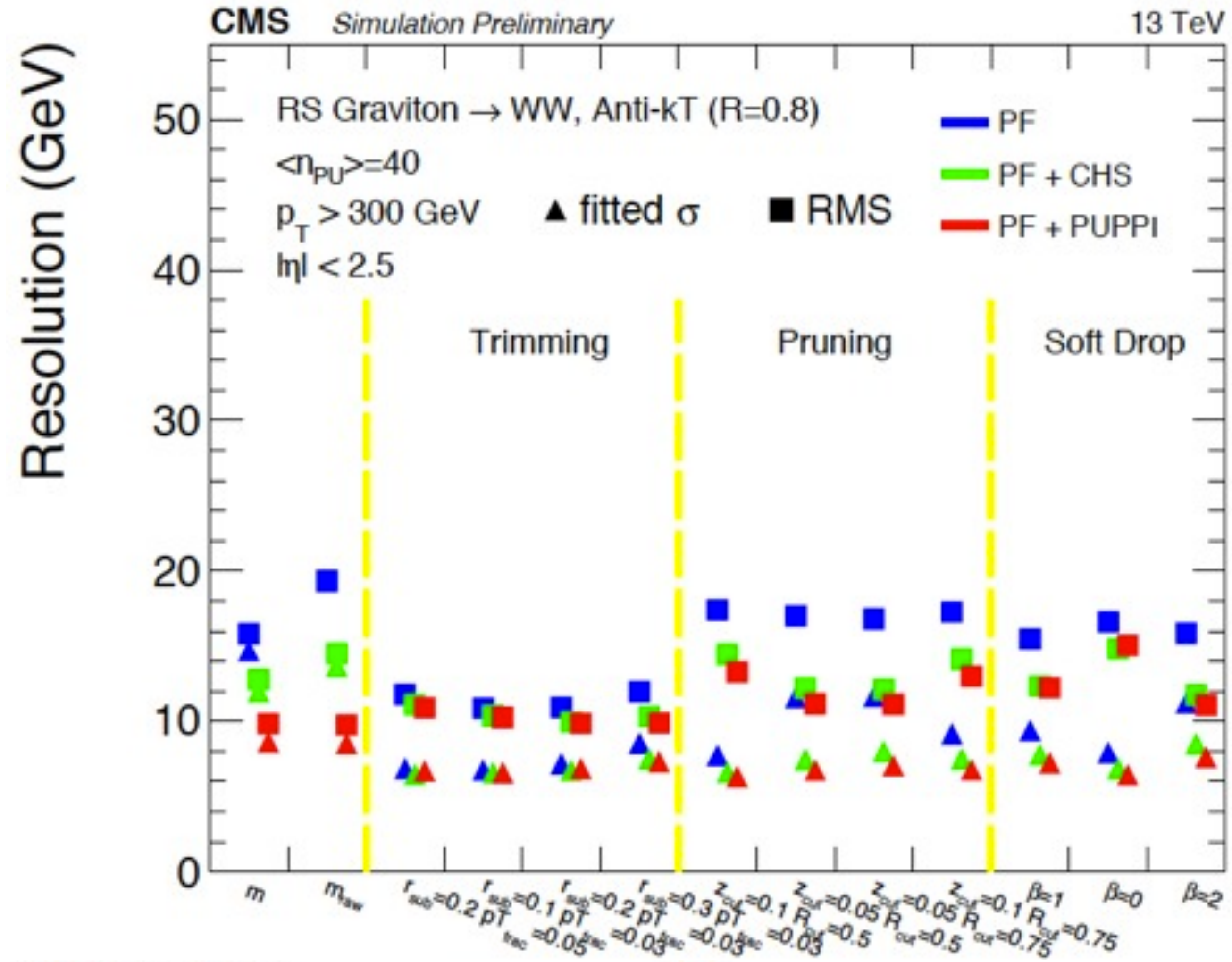
- algorithms using CHS subtraction have ~same resolution and better than without and are more stable wrt PU

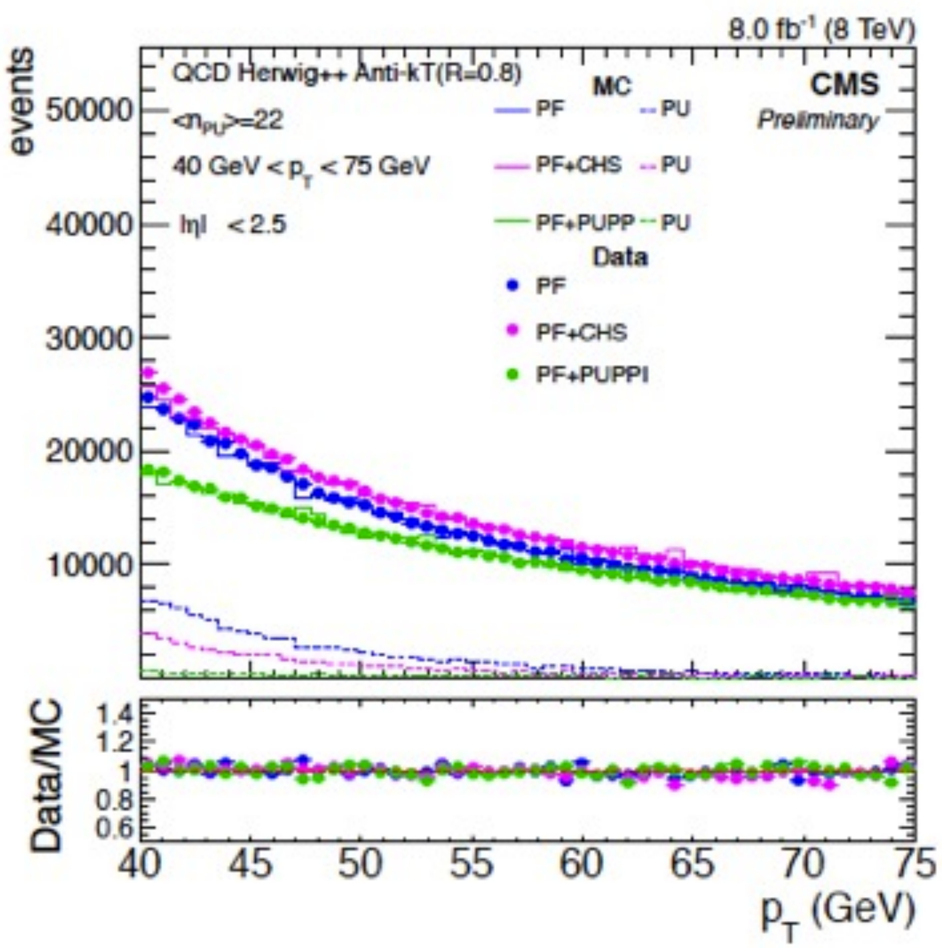
PU Mitigation - complex techniques

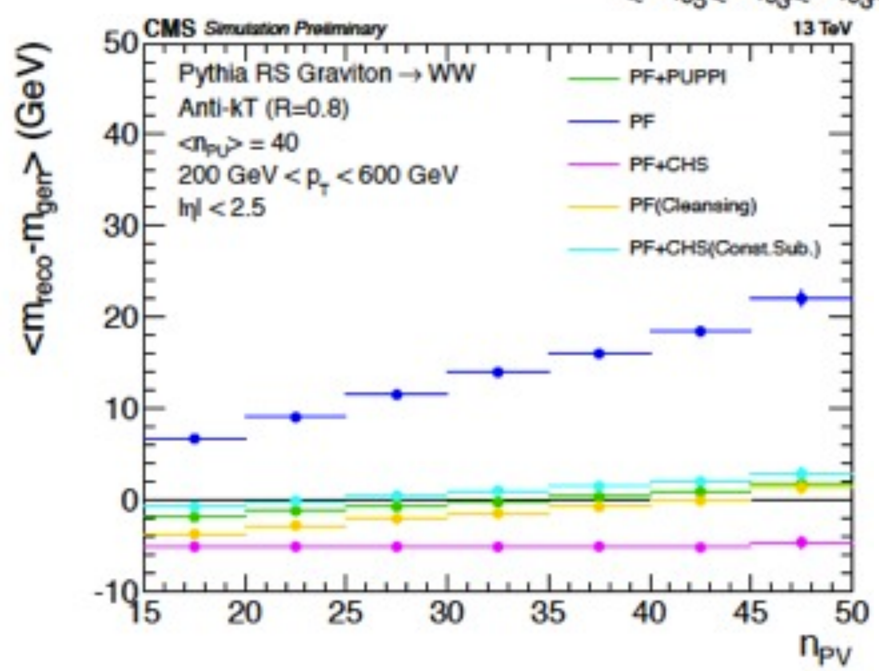
- constituent subtraction: pile up subtraction from individual particles, inputs are jets and ρ
- linear cleansing: combined tracks and shapes
  - at subjet level, fraction of PV and non-PV tracks decide about removing the subjet
- pileup per particle identification (PUPPI): works on particles before jet clustering, particle-weight is calculated from p_T, shapes to label each as "PU-like" / "HS like"
  - extended to forward region (without tracking)

complex techniques - performance









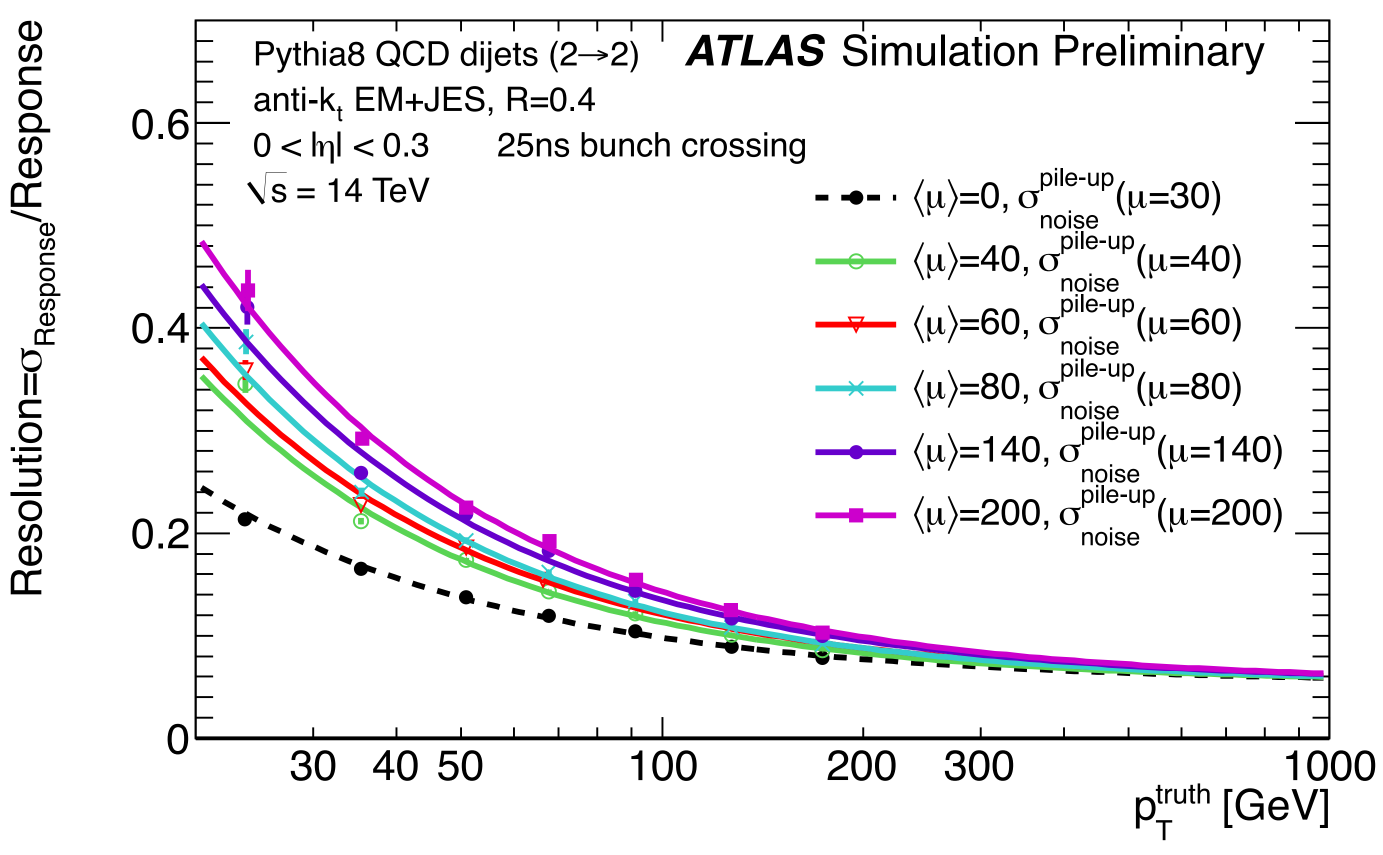
- nice resolution improvement through PUPPI
- PU stability
- good data/MC

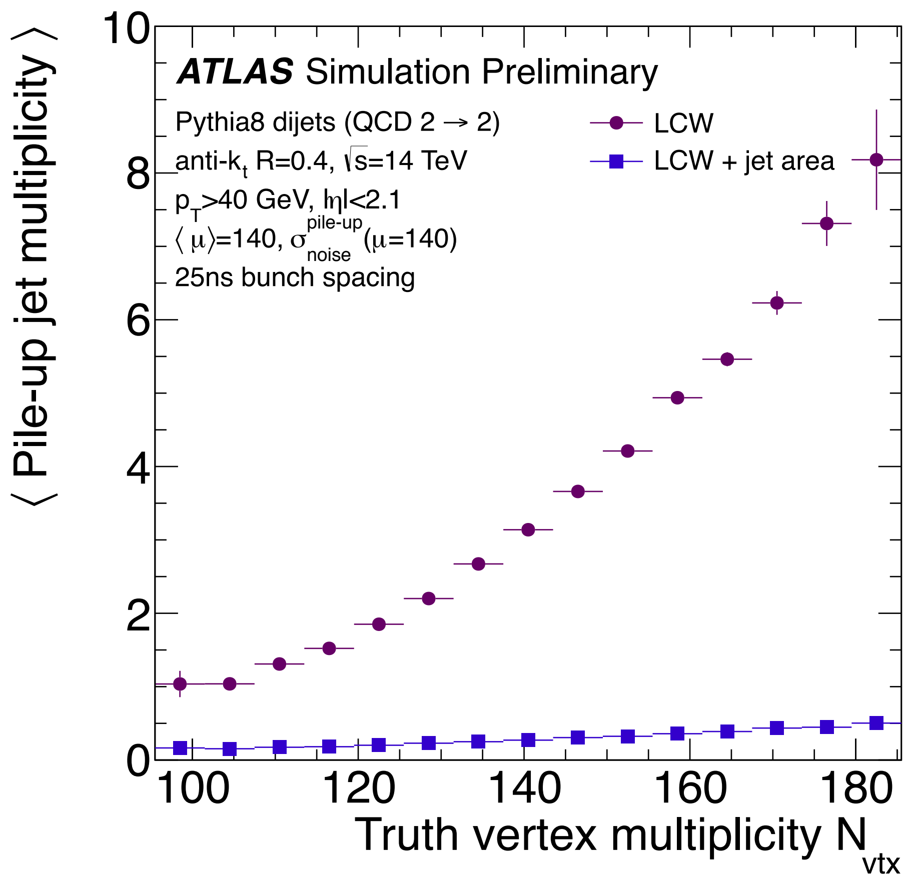
HL-LHC Projections

- jet response, vtx association (central +forward), grooming
- jet response, PF, CHS, PUPPI and timing

HL-LHC - ATLAS

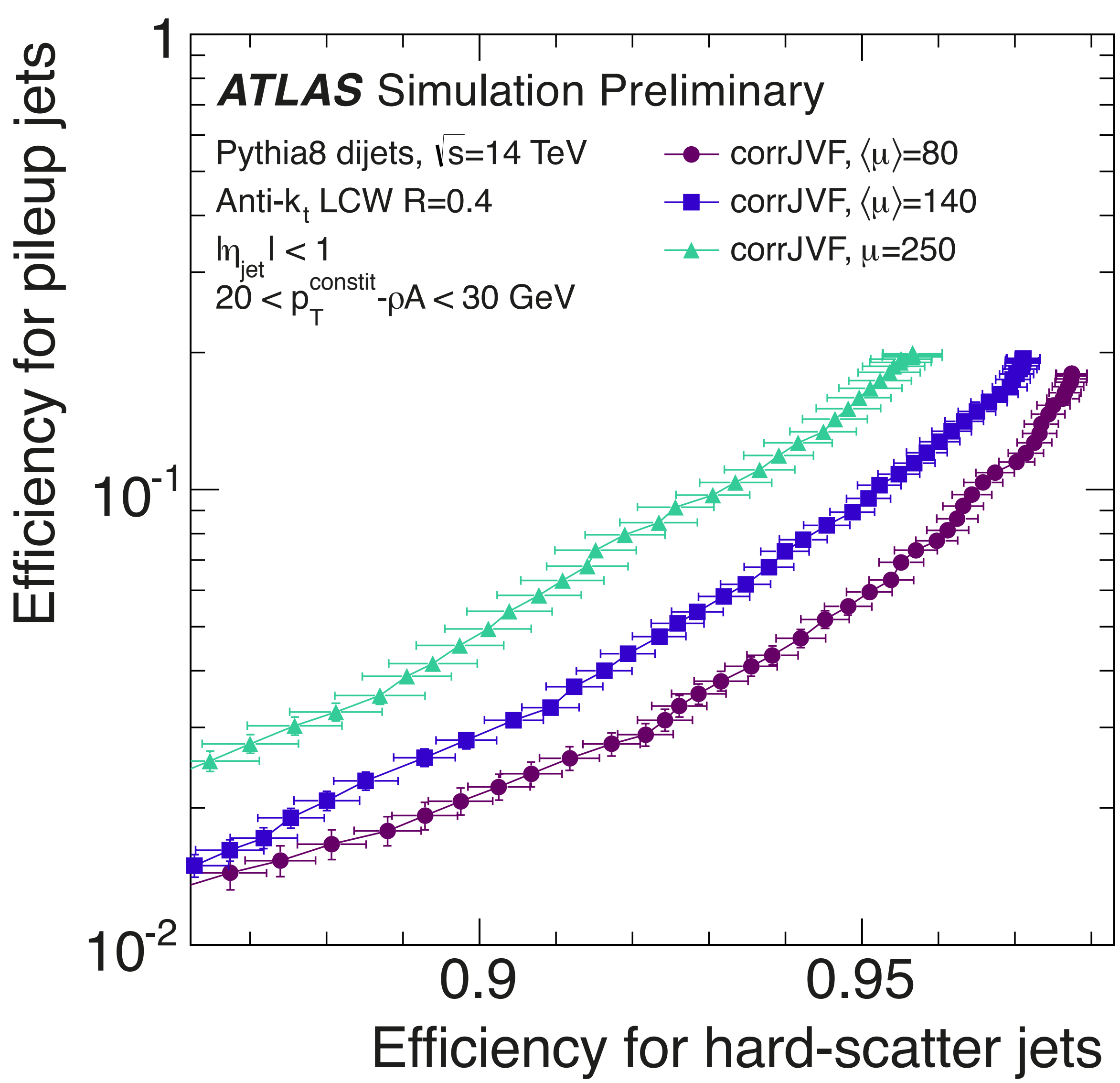
- currently **post Run3** - 300fb^{-1}, <μ>=80 and **HL-LHC** - 3ab^{-1}, <μ>=140-200
- Three Scoping scenarios for upgrade - work in progress
- dramatic worsening of jet p_T resolution with PU
- hope for straightforward application of current methods -> ρ-Area subtraction seems to restore stability
- for PU mitigation, tracking extension is considered up to 4.0

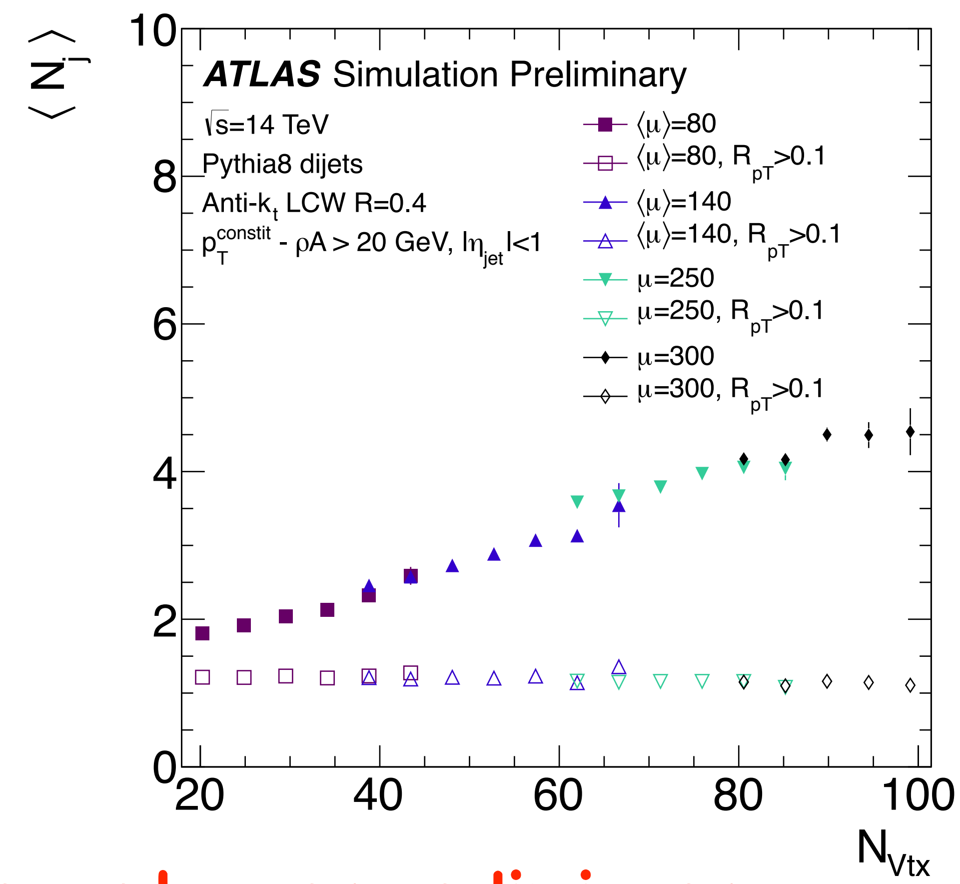




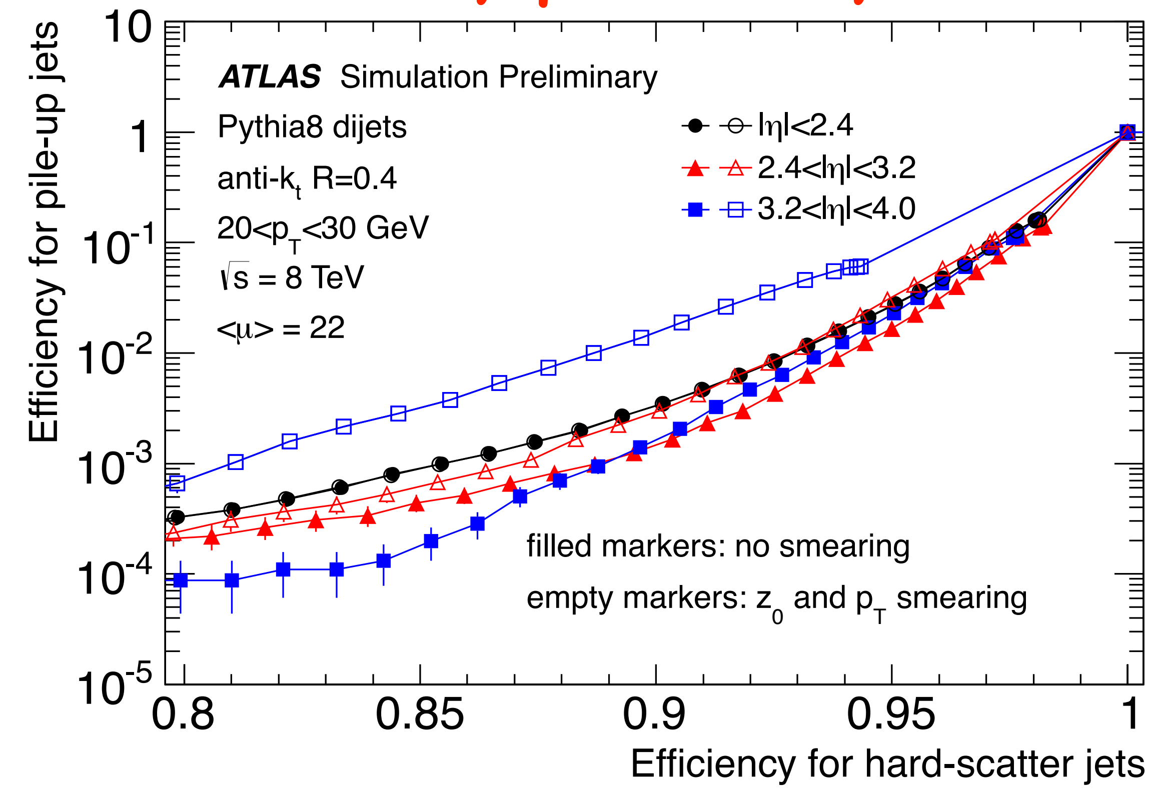
Tracking Confirmation

- corrJVF and RpT





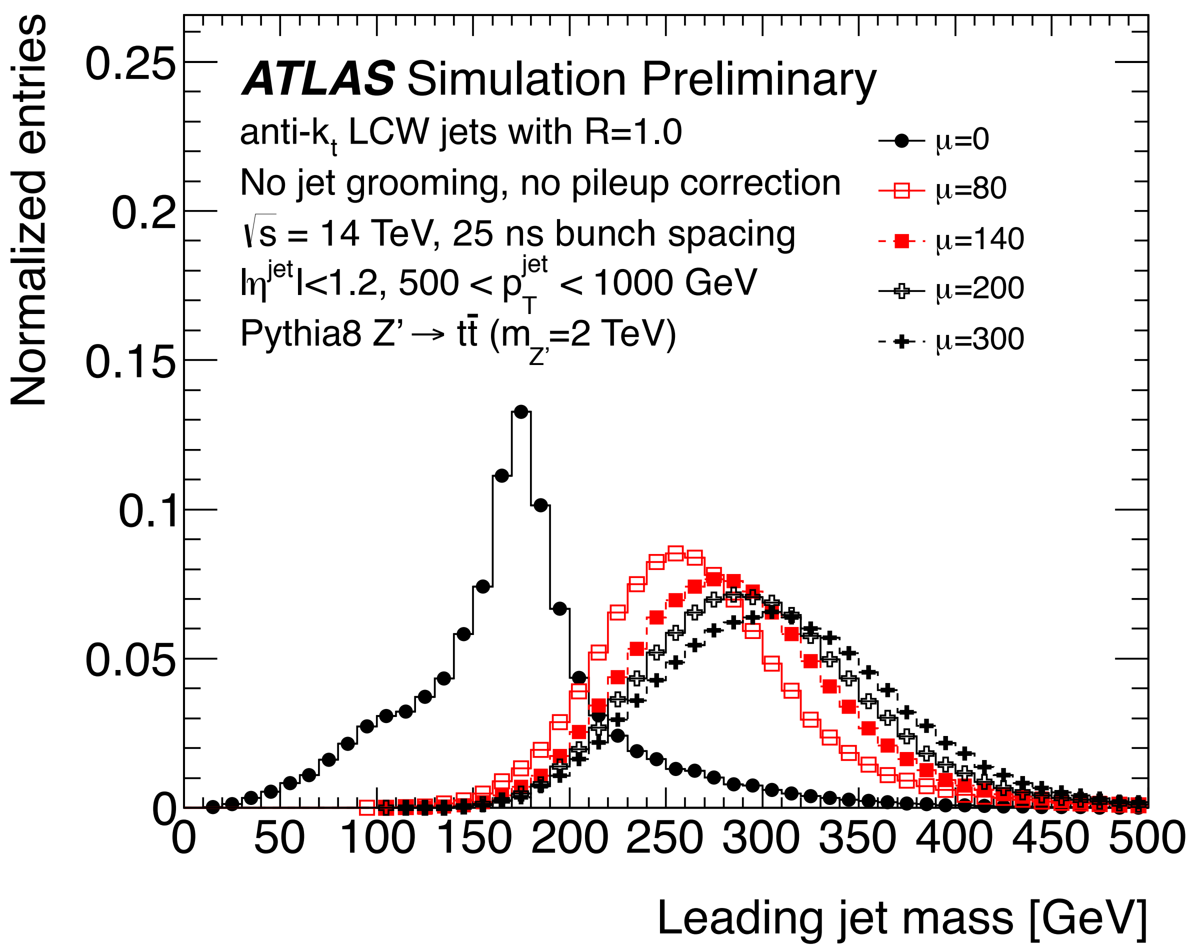
forward, very preliminary

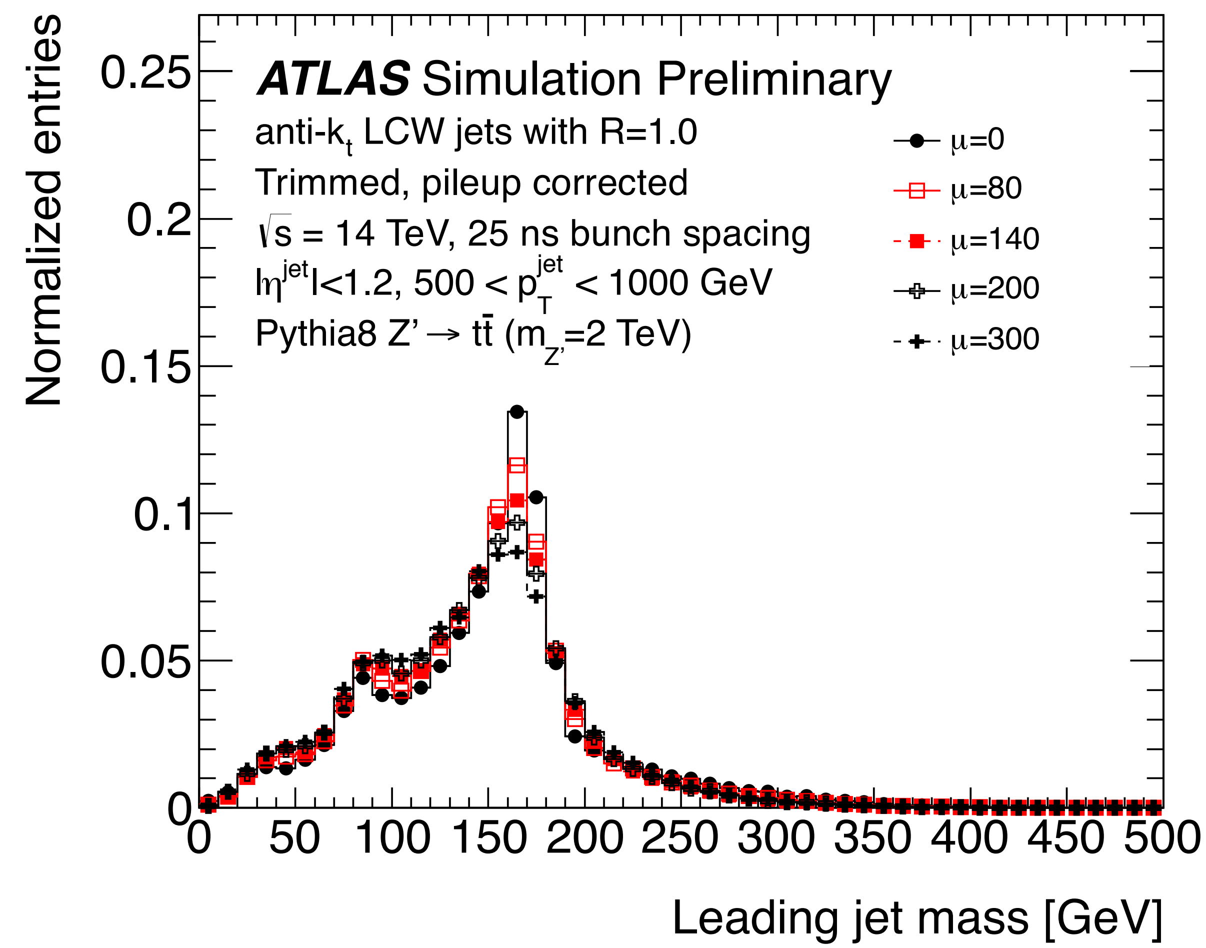


R_{pT} restores NPV independence, but PU rejection efficiency degrades with $\langle\mu\rangle$

grooming

- jet mass is increasingly washed out with pile up

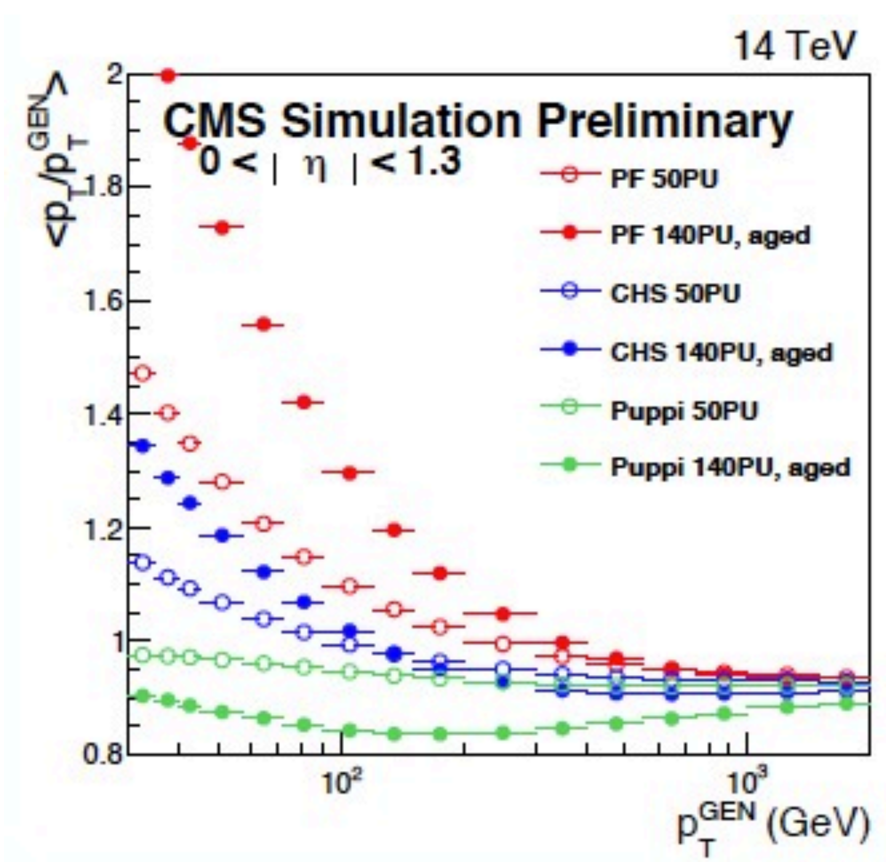


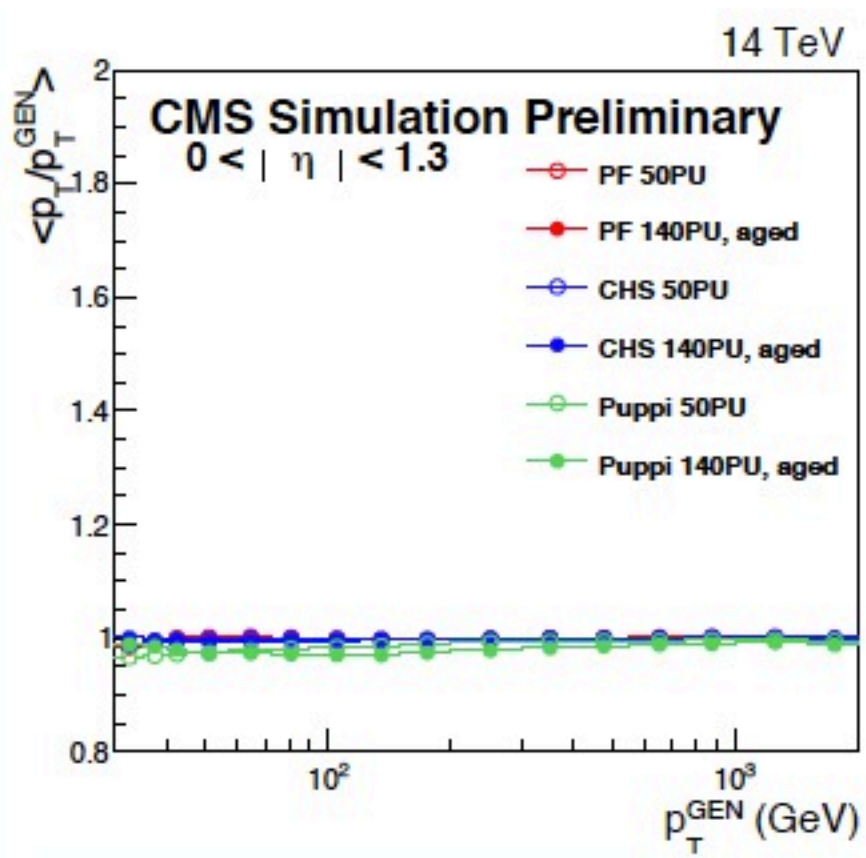


grooming and pile up corrections restore the mass distribution

HL-LHC - CMS

- Phase I, <μ> = 50, no aging
  - post Run3
- Phase I, <μ> = 140, aging except for pixel, 1ab^{-1}
  - demonstrates need for updates
- Phase II, <μ> = 140, aging except for pixel, 1ab^{-1}
  - benchmark for Phase II performance
- main upgrade is in tracking extension up to |η| < 4.0

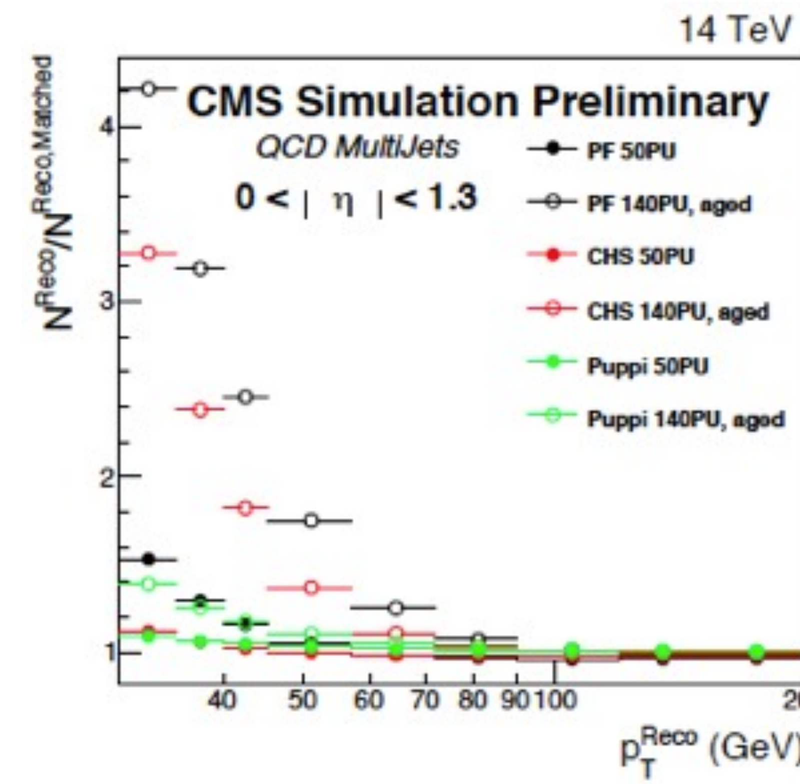


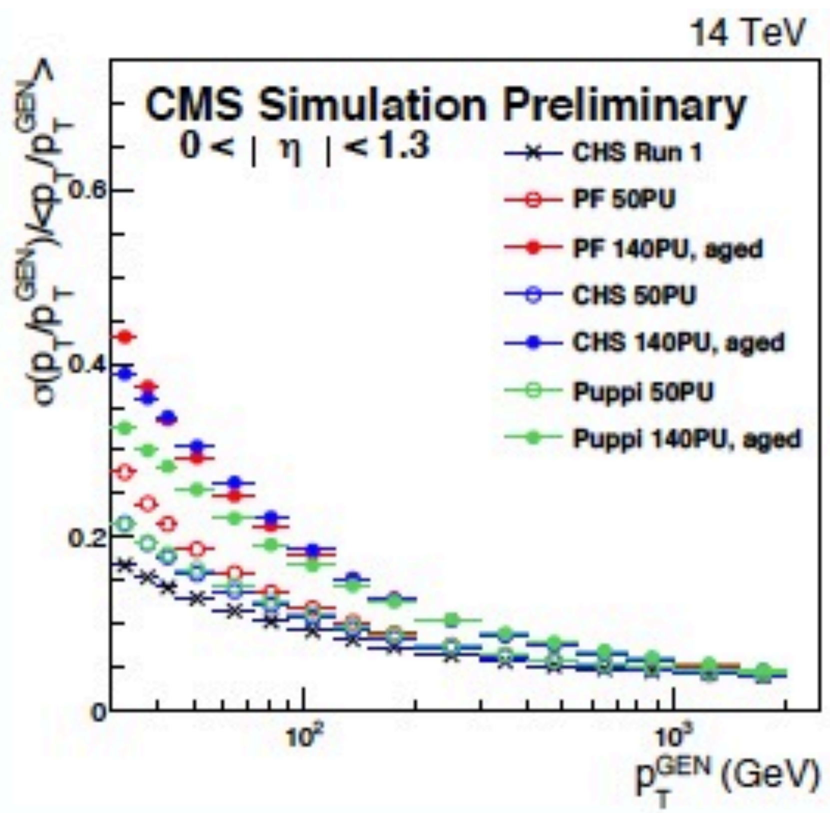


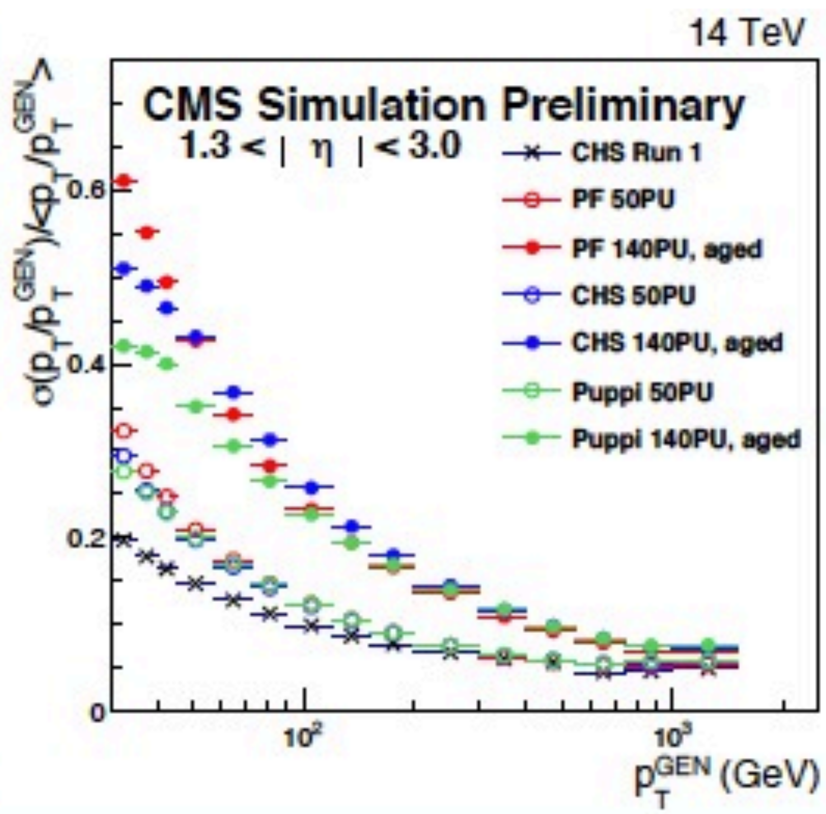
gain in stability through jet correction (ρ subtraction, detector response,...)

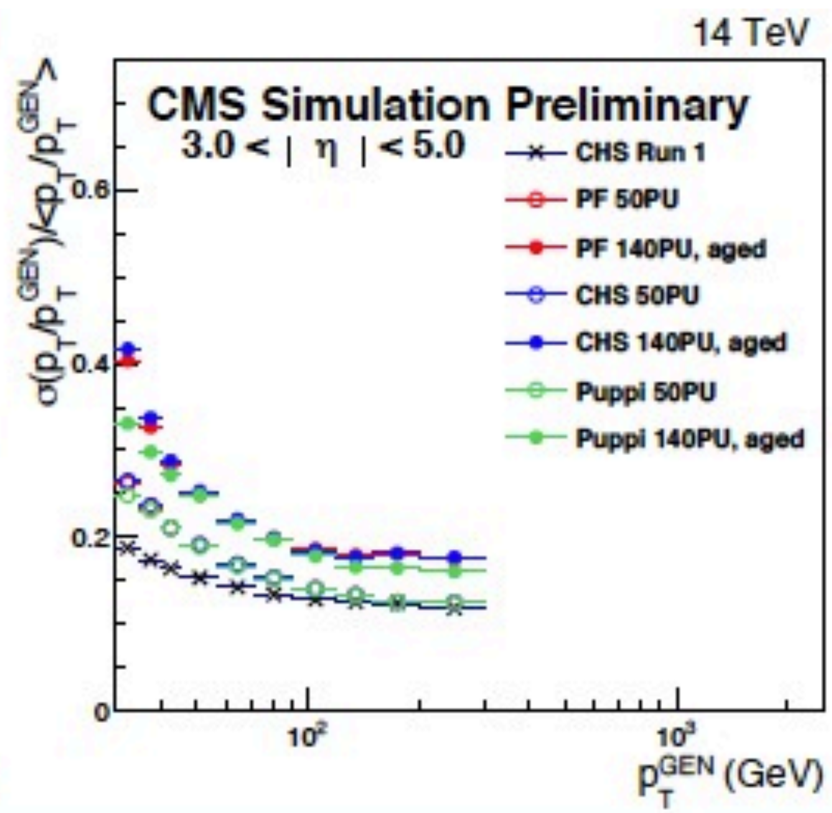
HL-LHC - jet response

- jet response for post Run3 is at best ~25% worse than Run1
- aged Phase2 detector shows 50-100% worsening
- in all cases, PUPPI performs the best

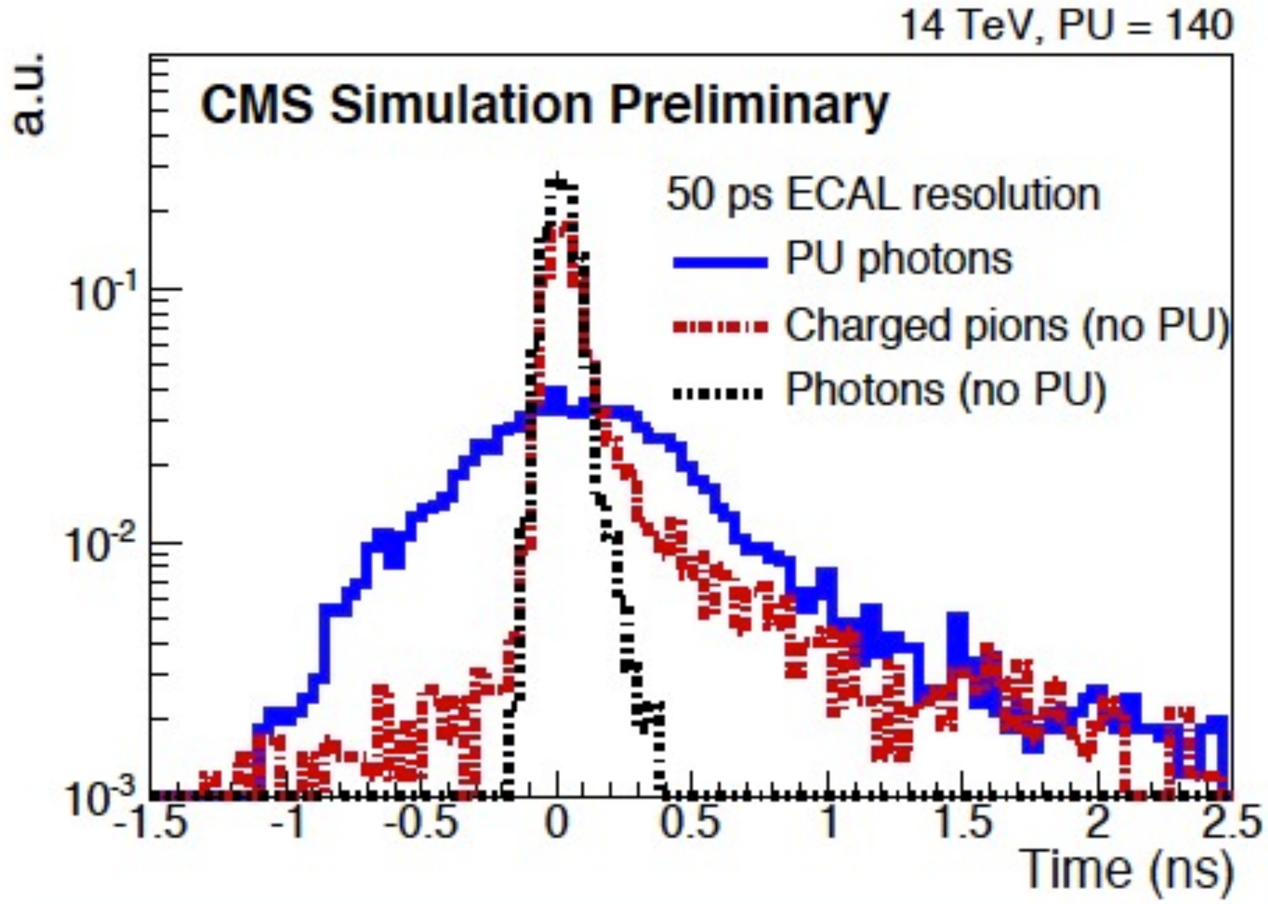


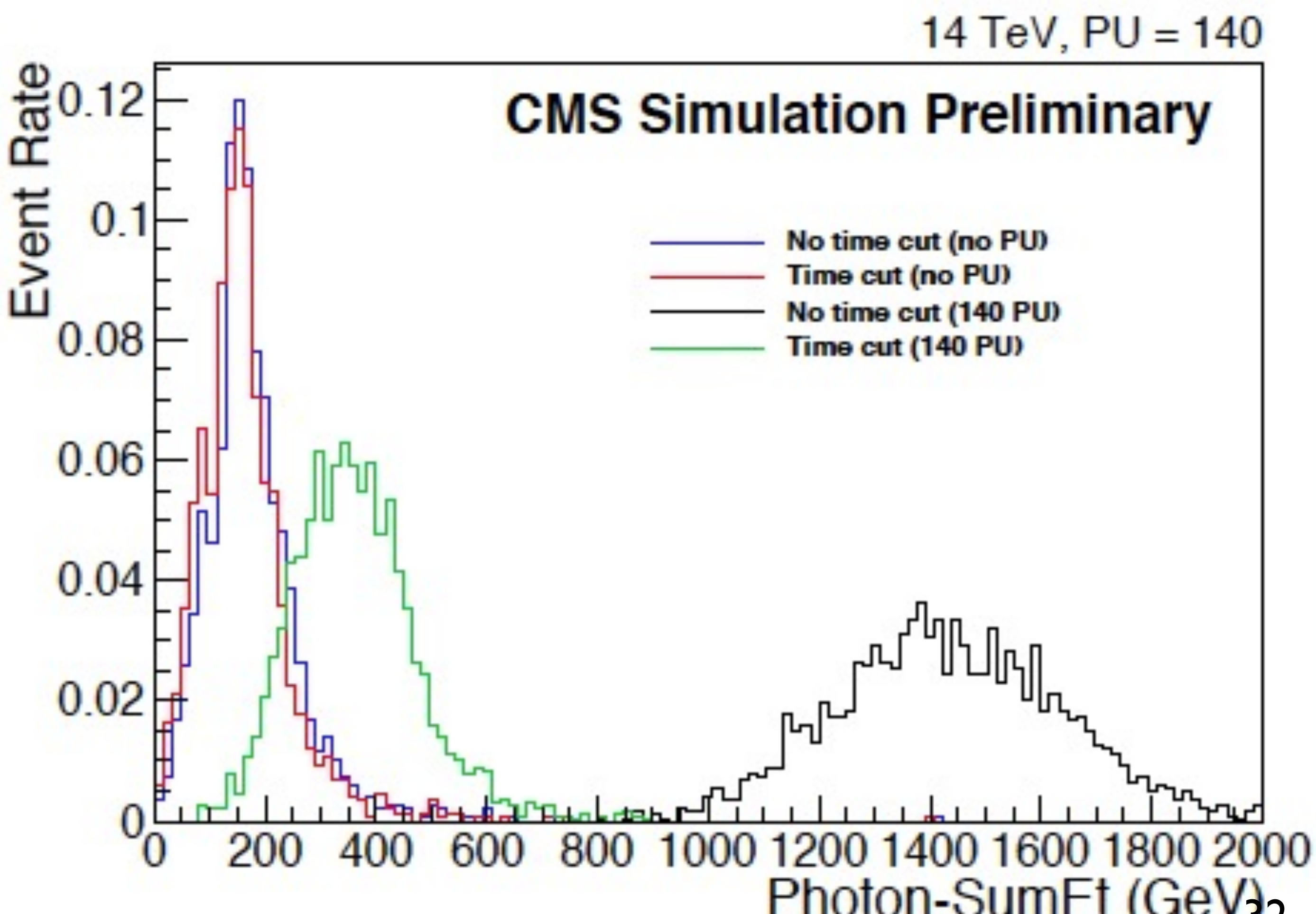






HL-LHC - precision timing





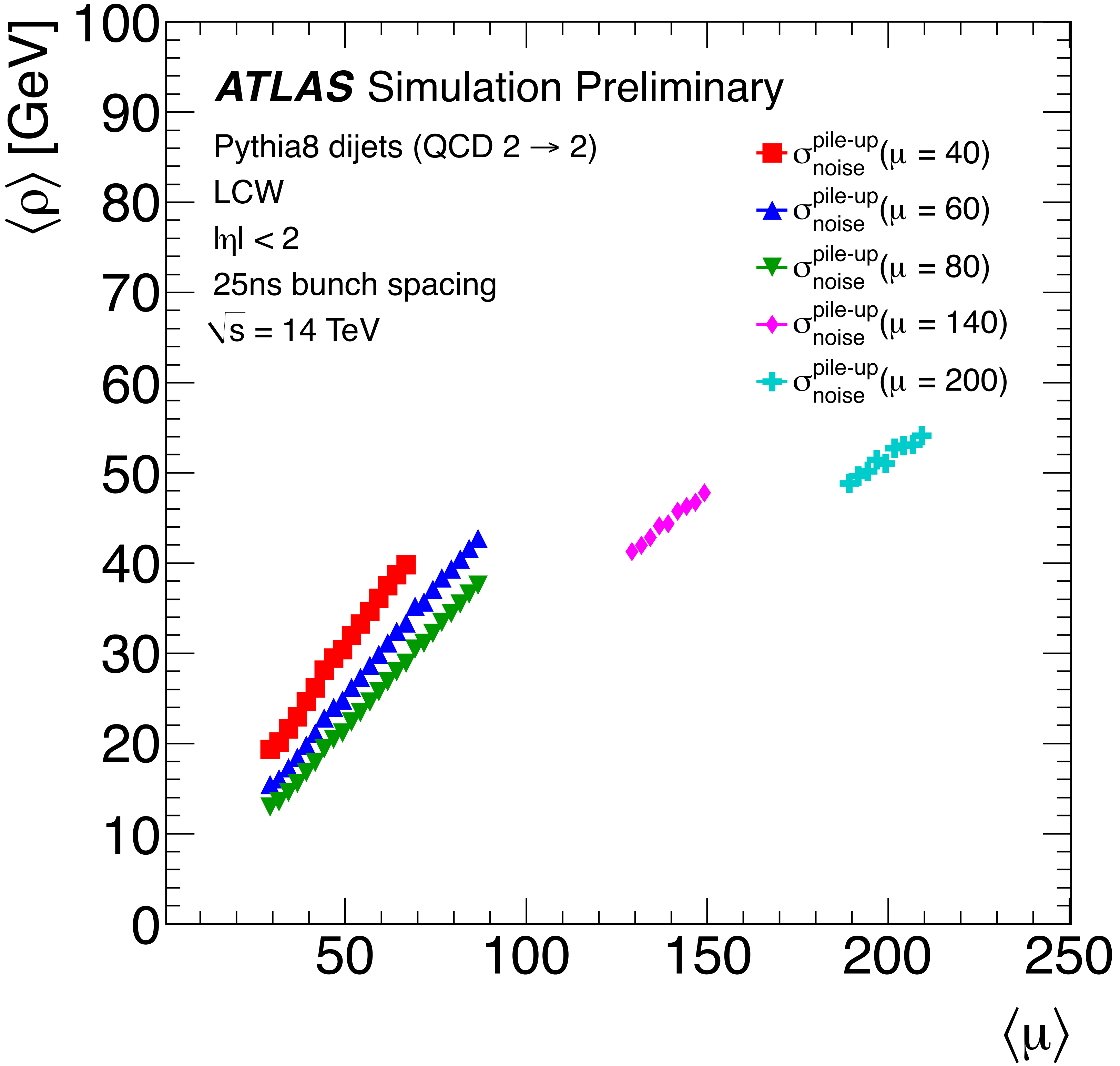
- 50ps timing would hugely help to reject pile up
- even with a time cut, mass peak for the H->γγ VBF signal does not correspond to m_H, but much still huge improvement visible

Summary

- pile up is the price to pay for high luminosity ... and it will get much worse
- ATLAS and CMS have slightly different approach to jet reconstruction, rich variety of pile up mitigation techniques is nevertheless based on similar ideas
  - subtracting pile up density from jets
  - usage of tracking - CHS, JVT, grooming, ...
  - advanced combination - cleansing and PUPPI
- **already upcoming Run2 with higher CME and pile up conditions will allow for more careful validation and further development towards HL-LHC conditions**

backup

<ρ> vs <μ>



ATLAS Simulation Preliminary
Pythia8 dijets (QCD 2 → 2)
LCW
|η| < 2
25ns bunch spacing
√s = 14 TeV
σ noise pile-up (μ = 40)
σ noise pile-up (μ = 60)
σ noise pile-up (μ = 80)
σ noise pile-up (μ = 140)
σ noise pile-up (μ = 200)
⟨ρ⟩ [GeV]
⟨μ⟩
0
10
20
30
40
50
60
70
80
90
100
50
100
150
200
250

groomers vs NPV

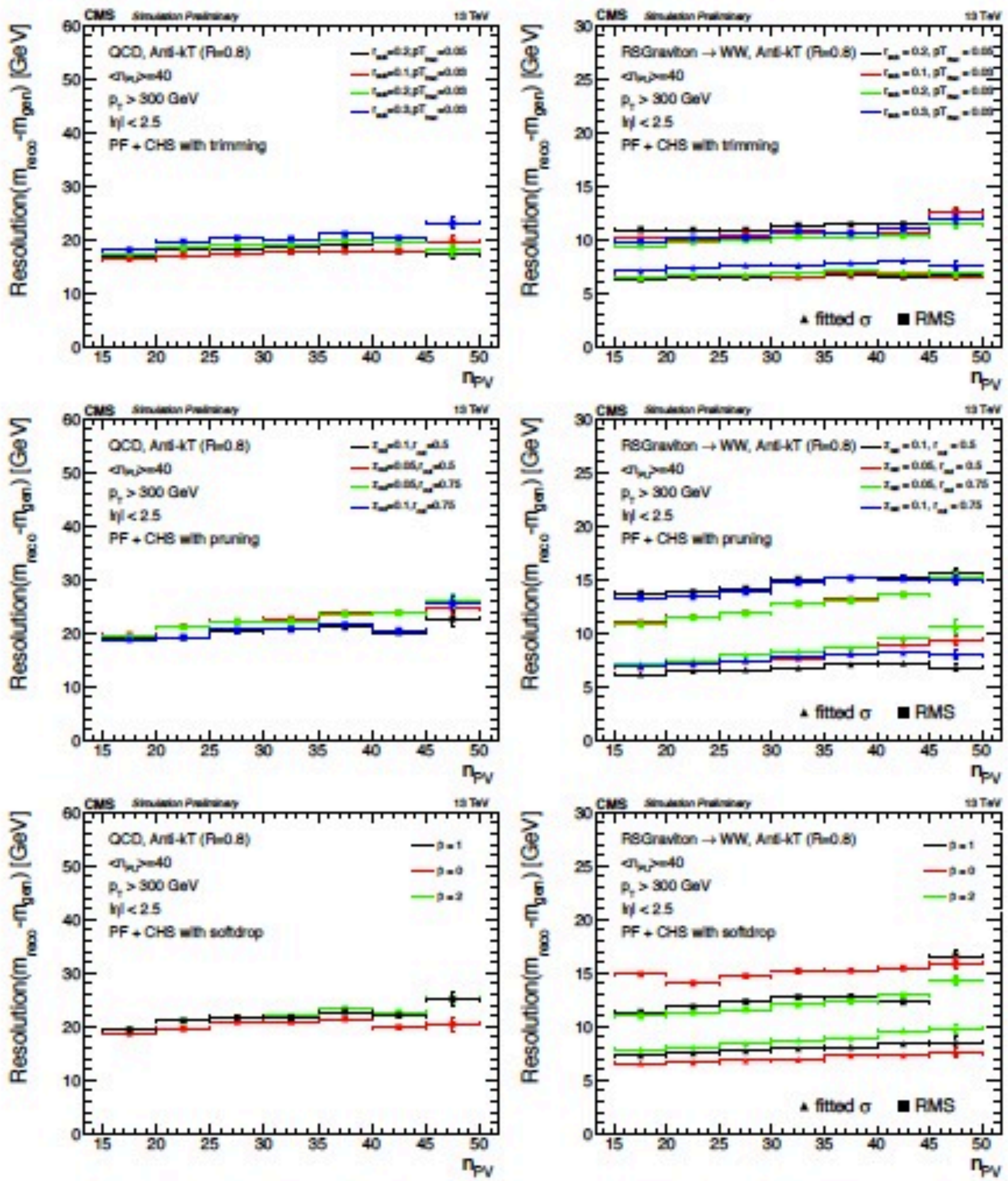


Figure 9: Pileup dependence versus n_{PV} of the mass resolution for QCD jets (left) and jets matched to generated W bosons (right) for various parameters of the grooming algorithms. For jets matched to W bosons, both the RMS and the σ from a Gaussian fit to the $(m_{\rm reco} - m_{\rm gen})$ distribution are reported.

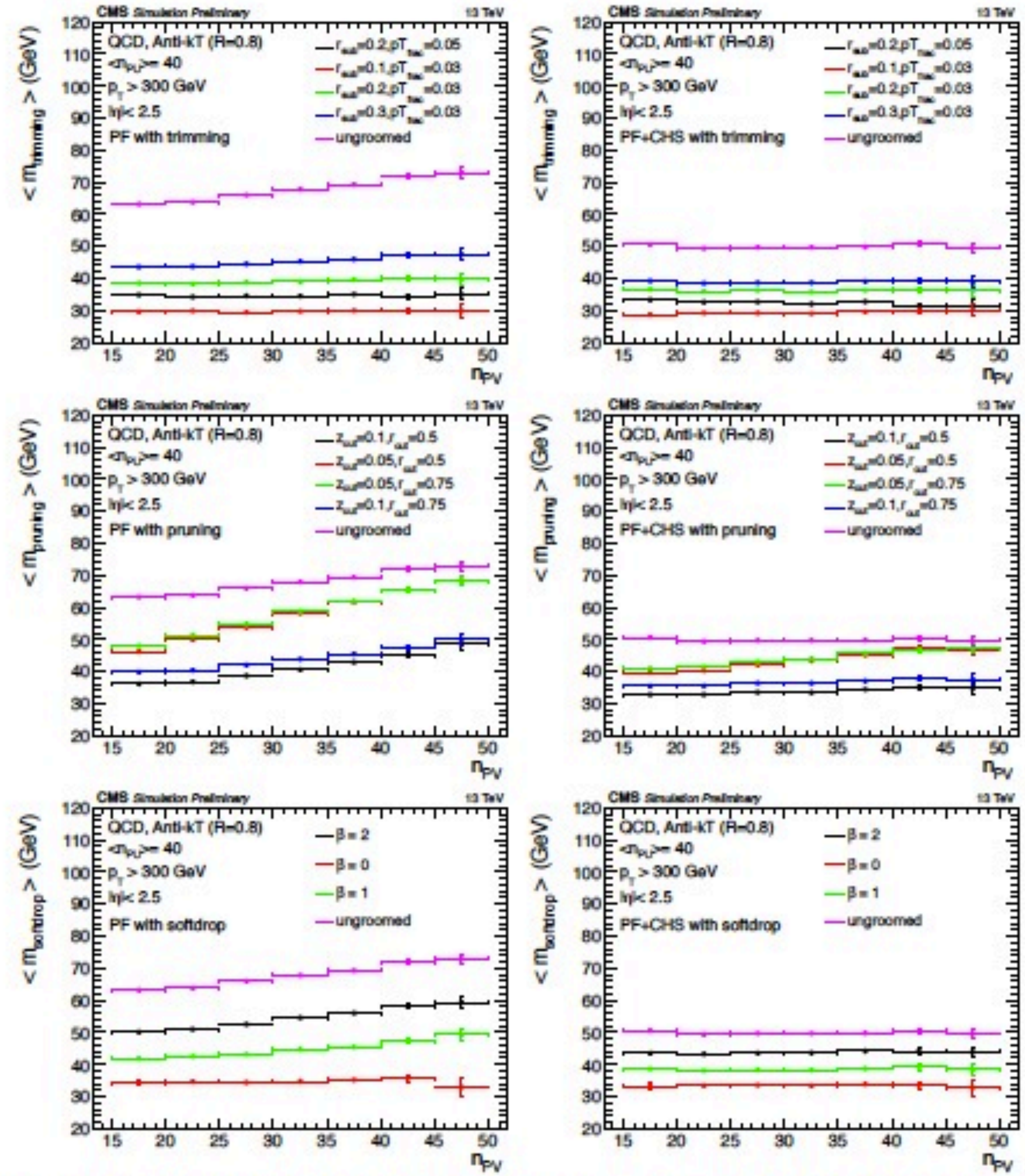
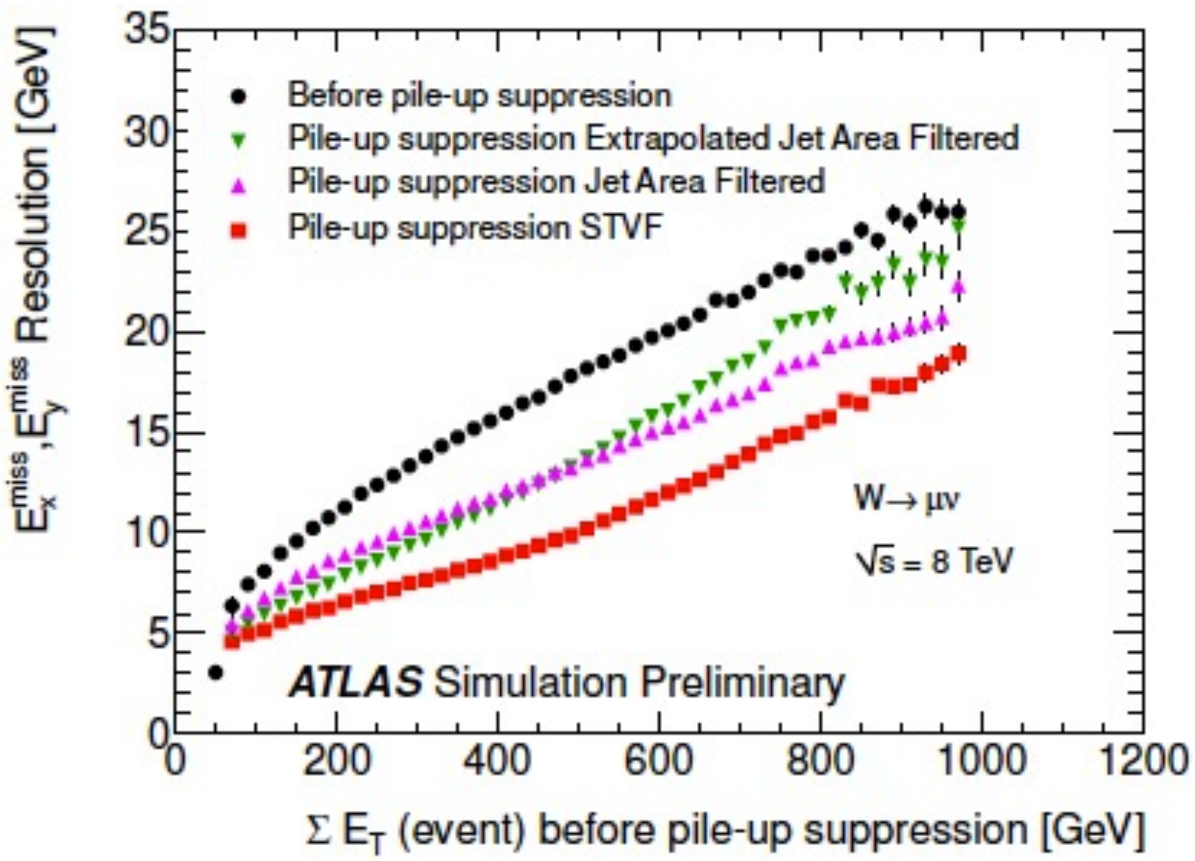
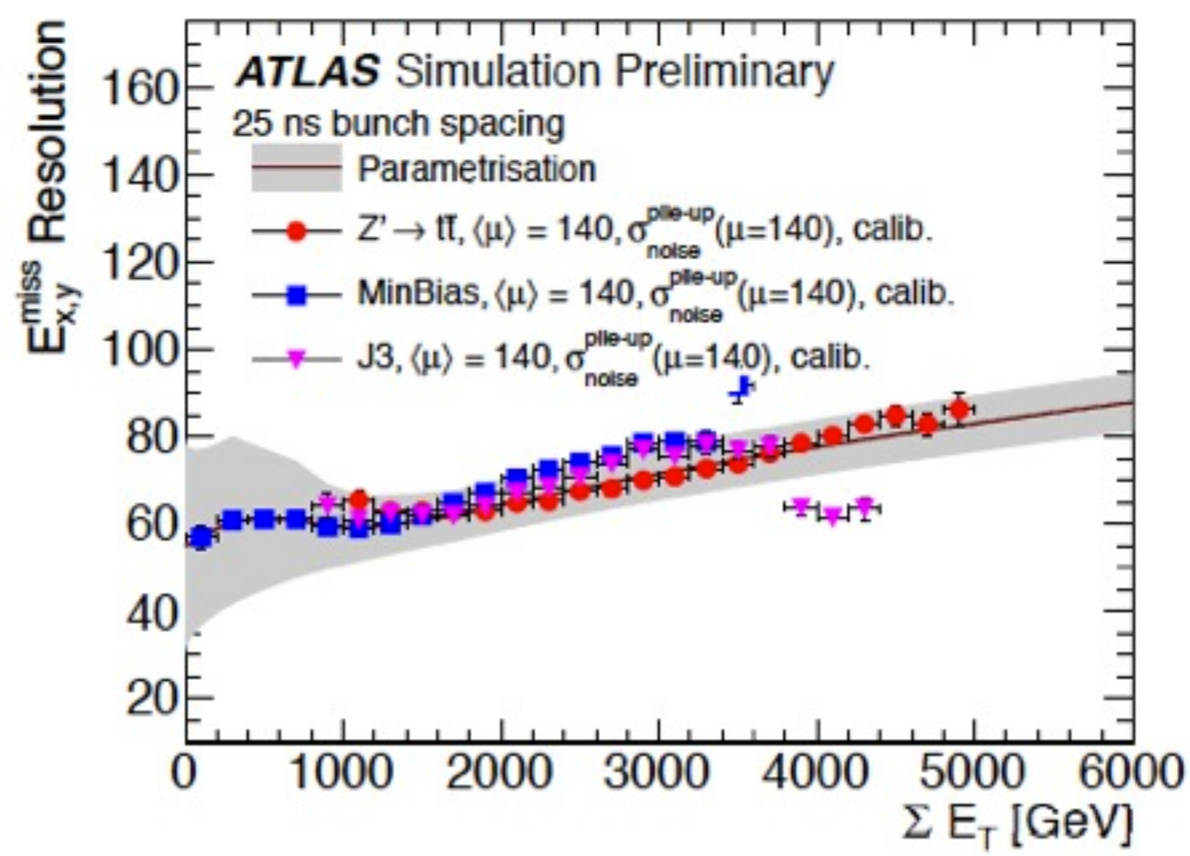


Figure 8: Pileup dependence versus n_{PV} of the average jet mass for PF jets (left) and PF+CHS jets (right) for several grooming algorithms and parameters. The top row is trimmed jets, middle row is pruned jets and bottom row is jets with soft drop applied.

E_T^{miss}

- is a vectorial sum of hard scatter objects (jets with PU suppression) - biggest work on the residual "Soft Term" (for example using track association - STVF)
- for HL-LHC, soft term parametrizations used
- resolution worsens significantly





pile up suppression is under study