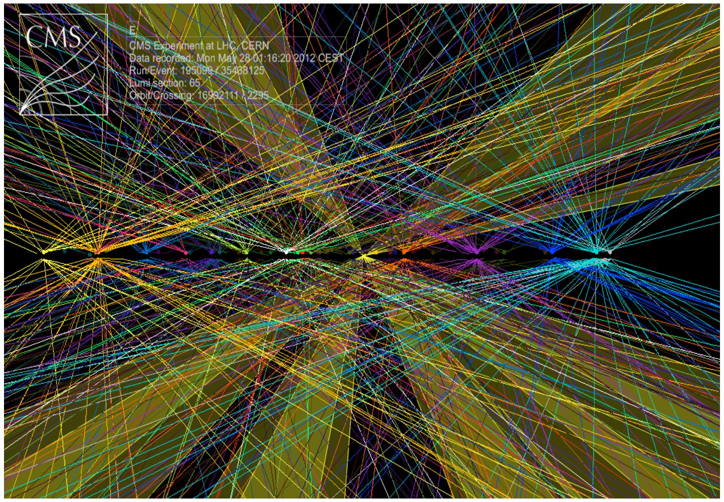
Performance of Jet Substructure with Pileup

Kalanand Mishra

Fermilab, CMS Collaboration



BOOST 2012 at Valencia, July 24, 2012

Outline

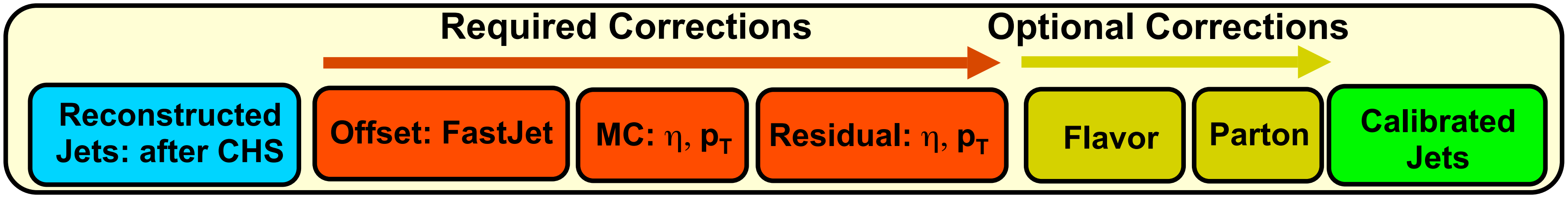
- Jet calibration in CMS
  - Effect of pileup and pileup subtraction
  - Jet correction uncertainties and contribution from pileup

- Jet composition
  - energy fractions contributed by charged and neutral hadrons, photons, electrons, and muons

- p_T response and resolution for groomed jets
- Jet mass for ungroomed and groomed jets
  - Effect of pileup on jet mass
  - Performance vs pileup by jet size
  - Performance vs pileup by grooming technique

- Summary

Jet energy calibration: overview



Required Corrections

Optional Corrections

Reconstructed Jets: after CHS

Offset: FastJet

MC: η, p_T

Residual: η, p_T

Flavor

Parton

Calibrated Jets

✦ Factorization facilitates the use of data-driven corrections

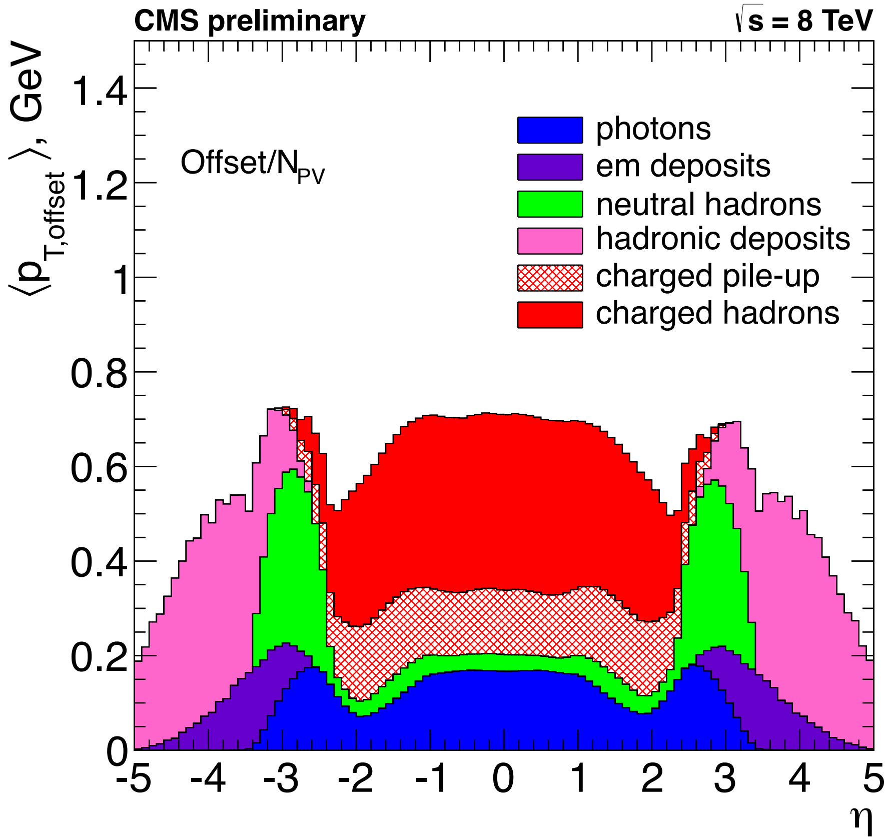
- Breaking the correction into pieces that are naturally measured in collider data:
  - **Offset**: pile-up and noise measured in zero-bias events.
  - **MC**: jet response vs. η, P_T using MC truth.
  - **Residual**: jet response vs. η, P_T using dijet balance and γ/Z +jet in data.

In CMS the most widely used jet is anti-k_T 0.5 (0.7 for QCD measurements). For pileup studies, consider anti-k_T 0.5, 0.7, 0.8 with various grooming techniques: filtering, trimming, pruning.

Pileup contribution to jet energy

✦Pileup (PU) measured with Zero Bias data and MC

- Random cone allows to separate contribution per detector
- Most charged hadrons can be associated to pileup vertices and removed



•Part that can be removed is labeled "charged hadrons"

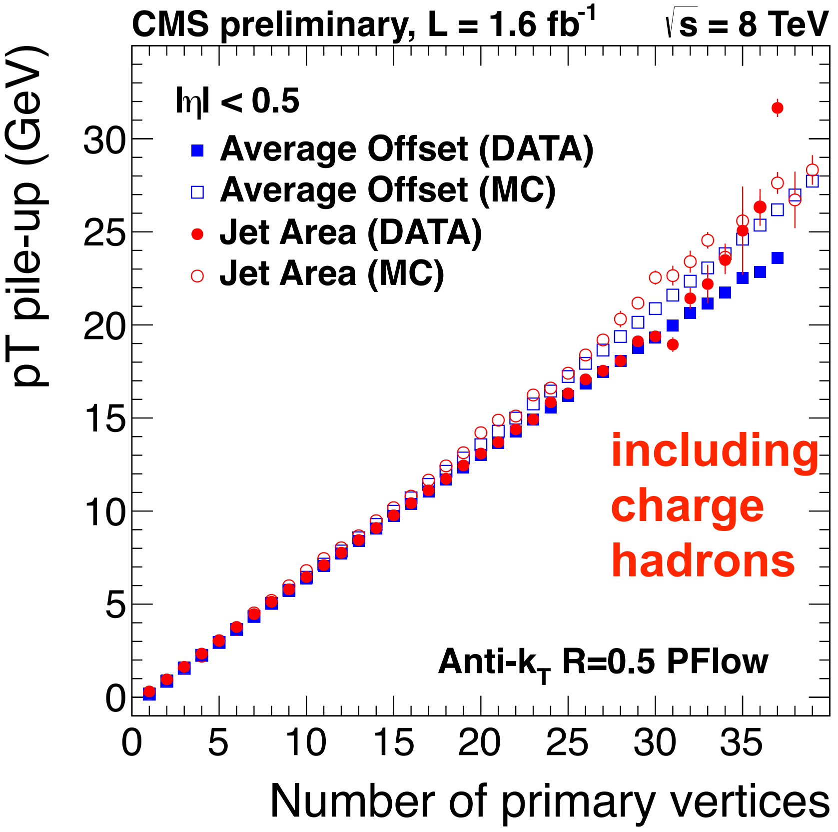
•Part that remains as PU after this needs to be subtracted

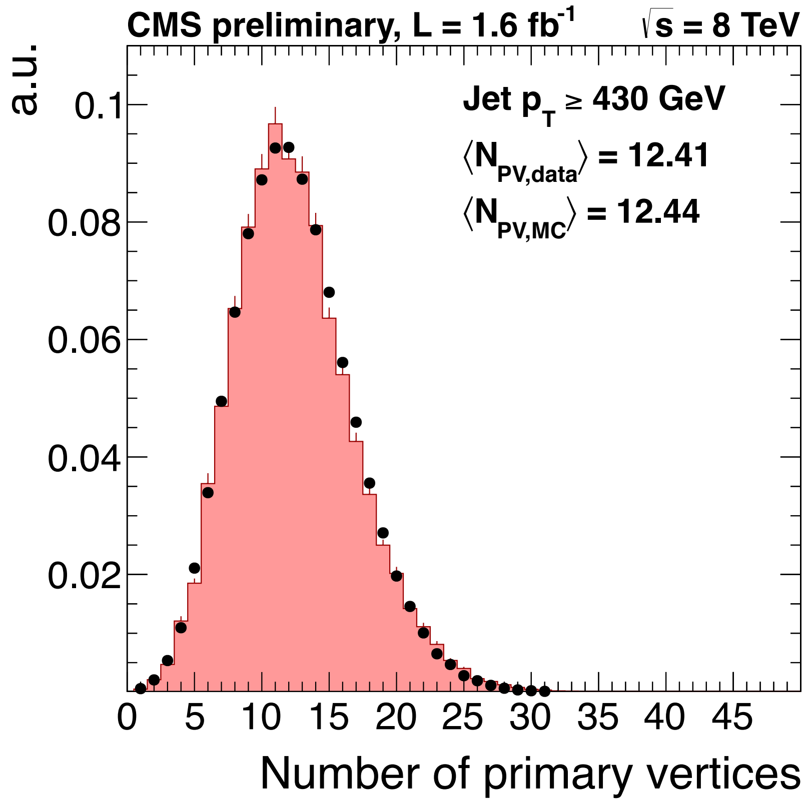
PU density x Effective area
(FastJet-ρ)

•PU density depends on the # of primary vertex in the event

Pileup correction: using NPV and [$\rho_{FastJet}$ x Area]

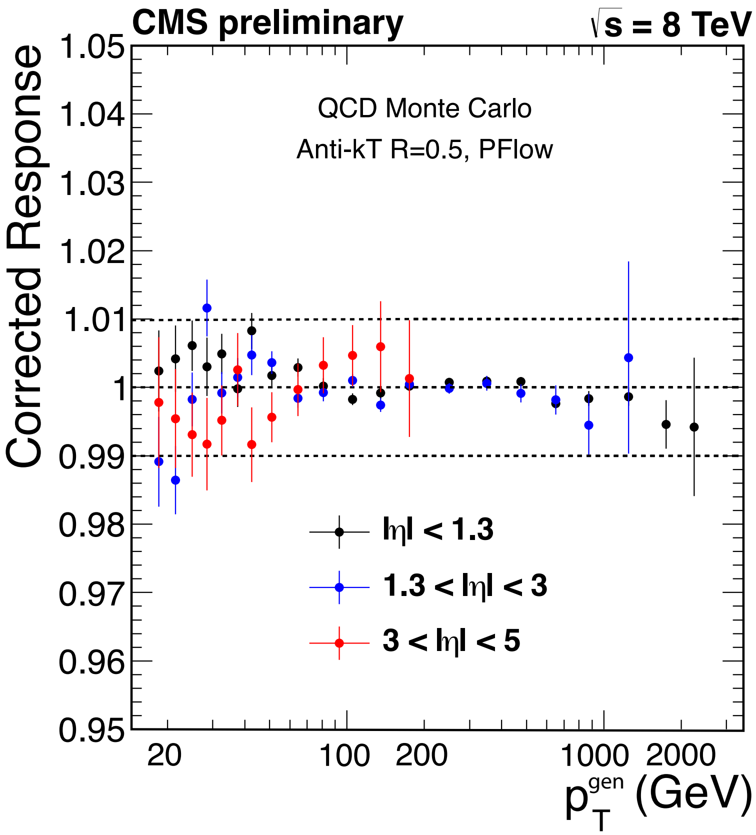
- Both NPV-based and FastJet-ρ-based corrections are in agreement
- Remaining Data/MC difference accounted for with separate PU corrections
  - Reweight pileup Poisson mean in MC to data. Poisson mean determined from measured luminosity and Minimum bias cross section.

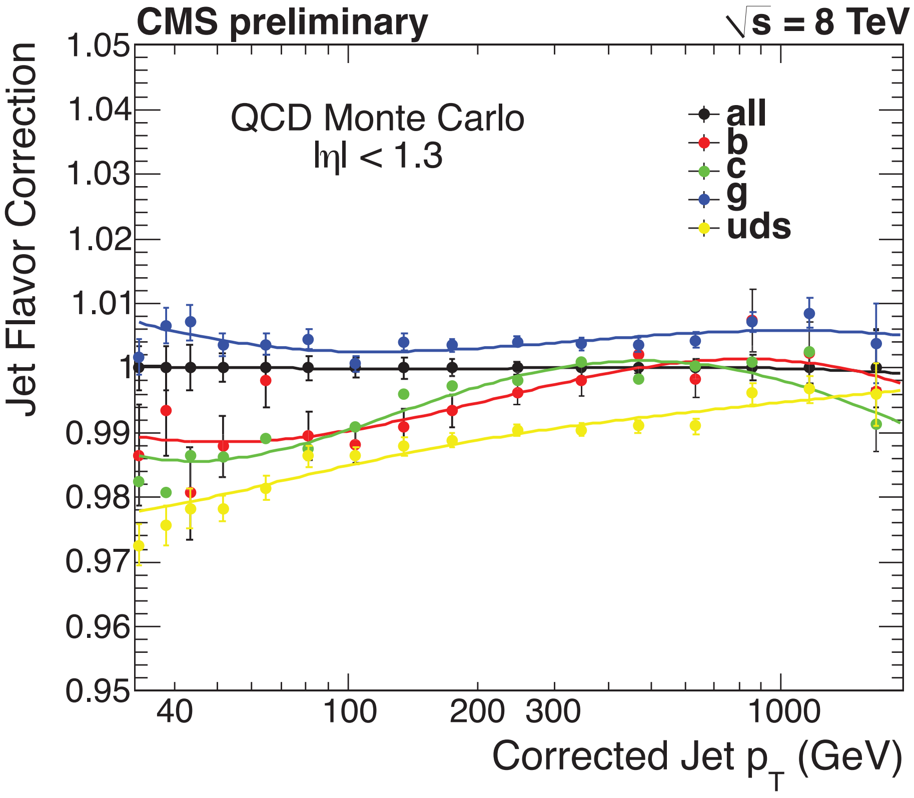




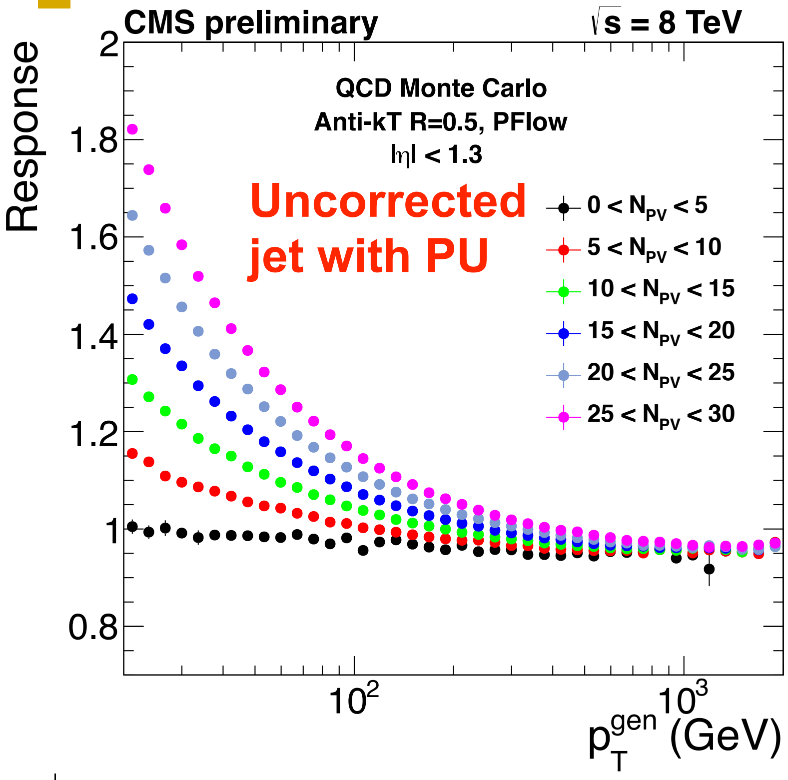
After PU offset: start with MC-based correction

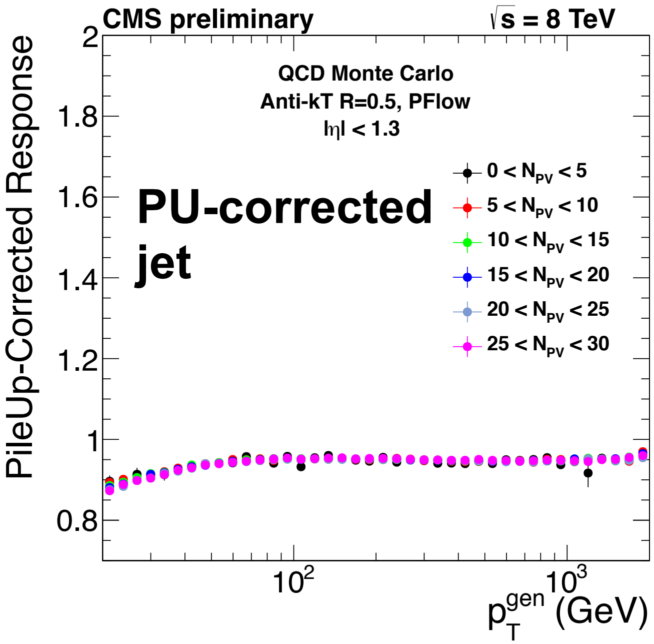
- ✦Eta and p_T corrections derived from QCD MC sample
  - Corrected response closes well in MC.
- ✦Particle flow minimizes flavor response differences
  - Maximum flavor difference within 3% in barrel for p_T > 30 GeV.

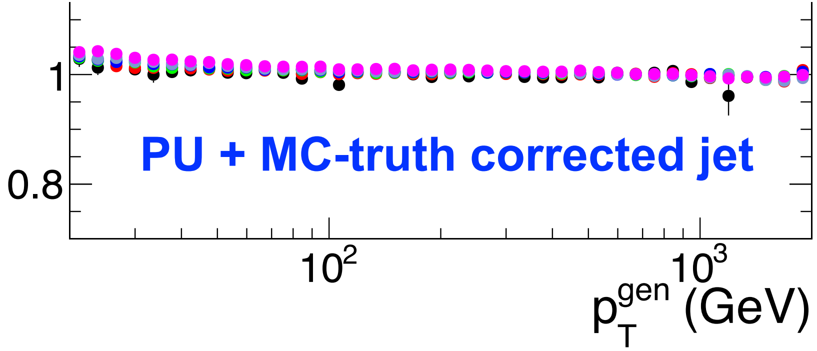




Combined pileup and MC truth effects





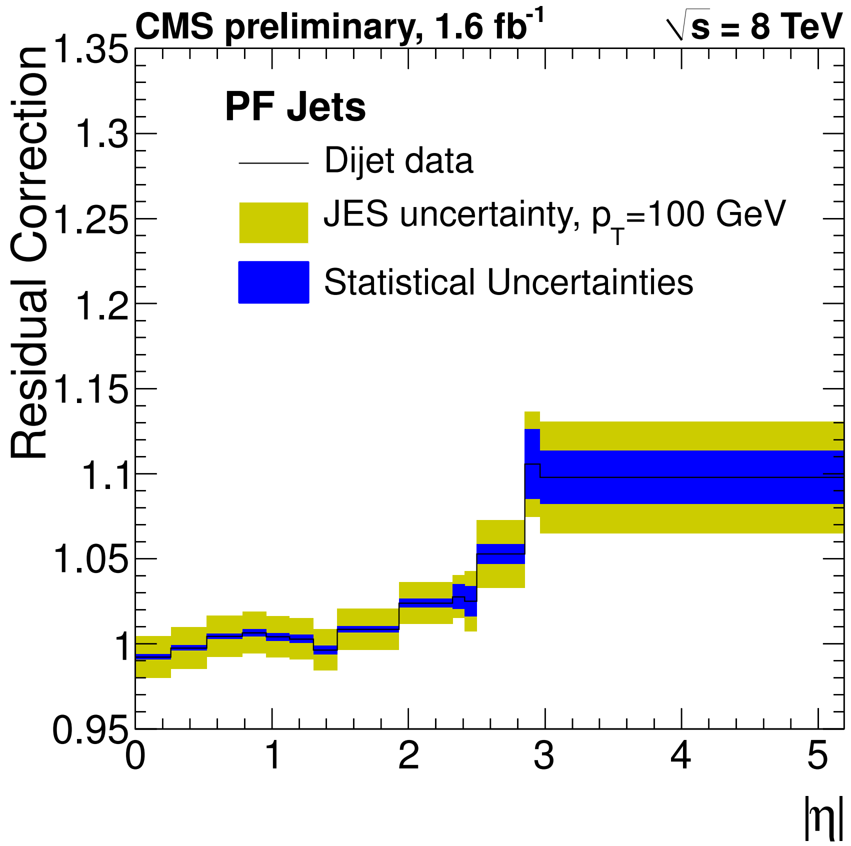


- ✦PU corrections remove response dependance vs N_{PV}.
- ✦MC truth correction brings the closure back to one.

Residual correction in data: η & p_T dependence

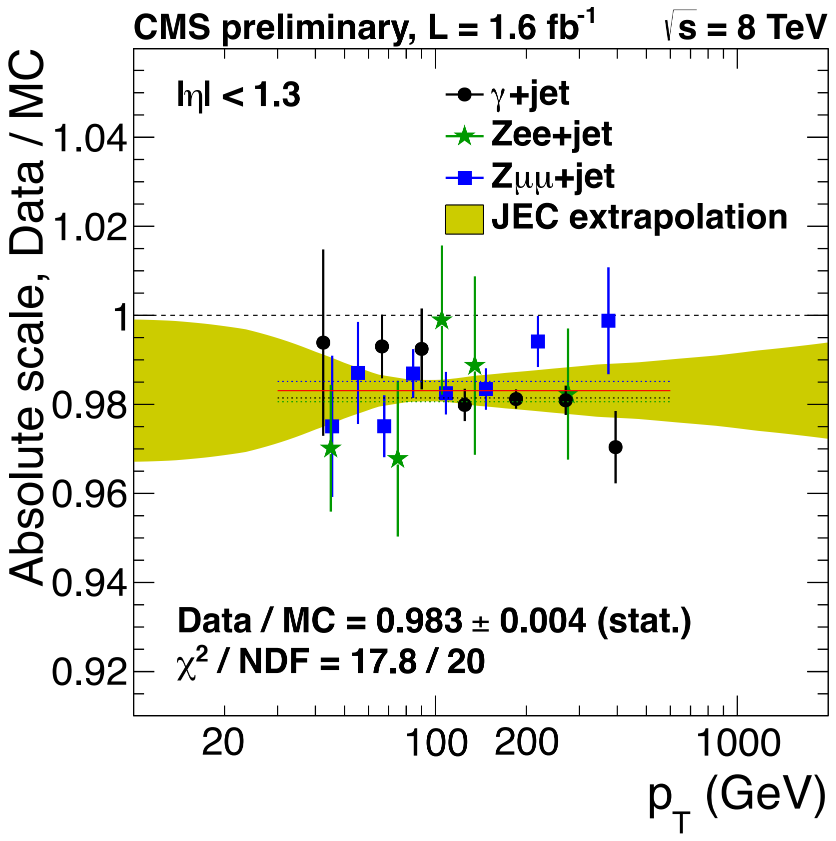
η dependence

- Using dijet events
- < 2.5% for jets in |η|<2.4
- HF modeling requires 5−10% correction



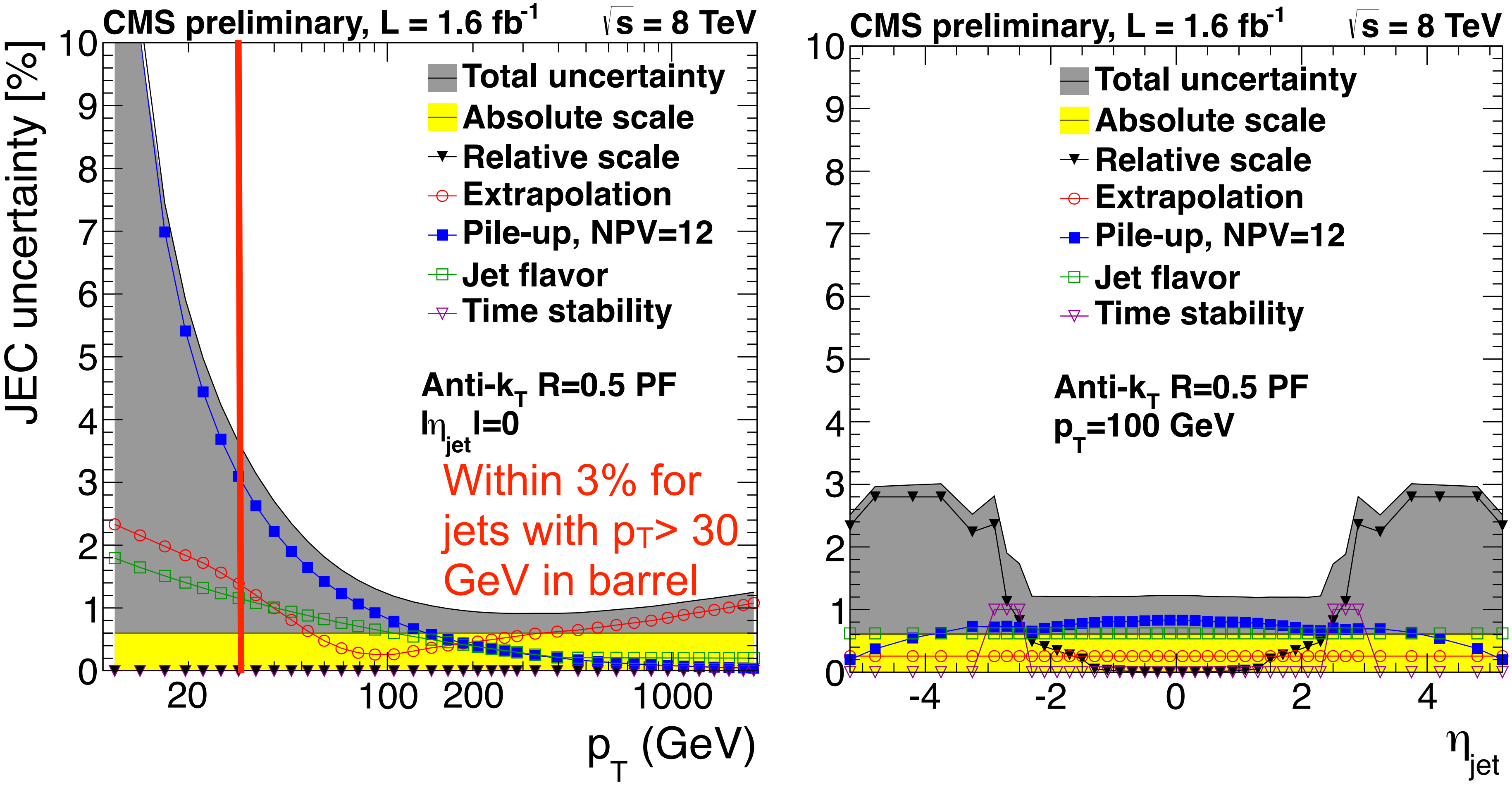
p_T dependence

- Using Zμμ+jet, Zee+jet, γ+jet events
- No significant p_T dependance observed, so a single flat scale factor is used.



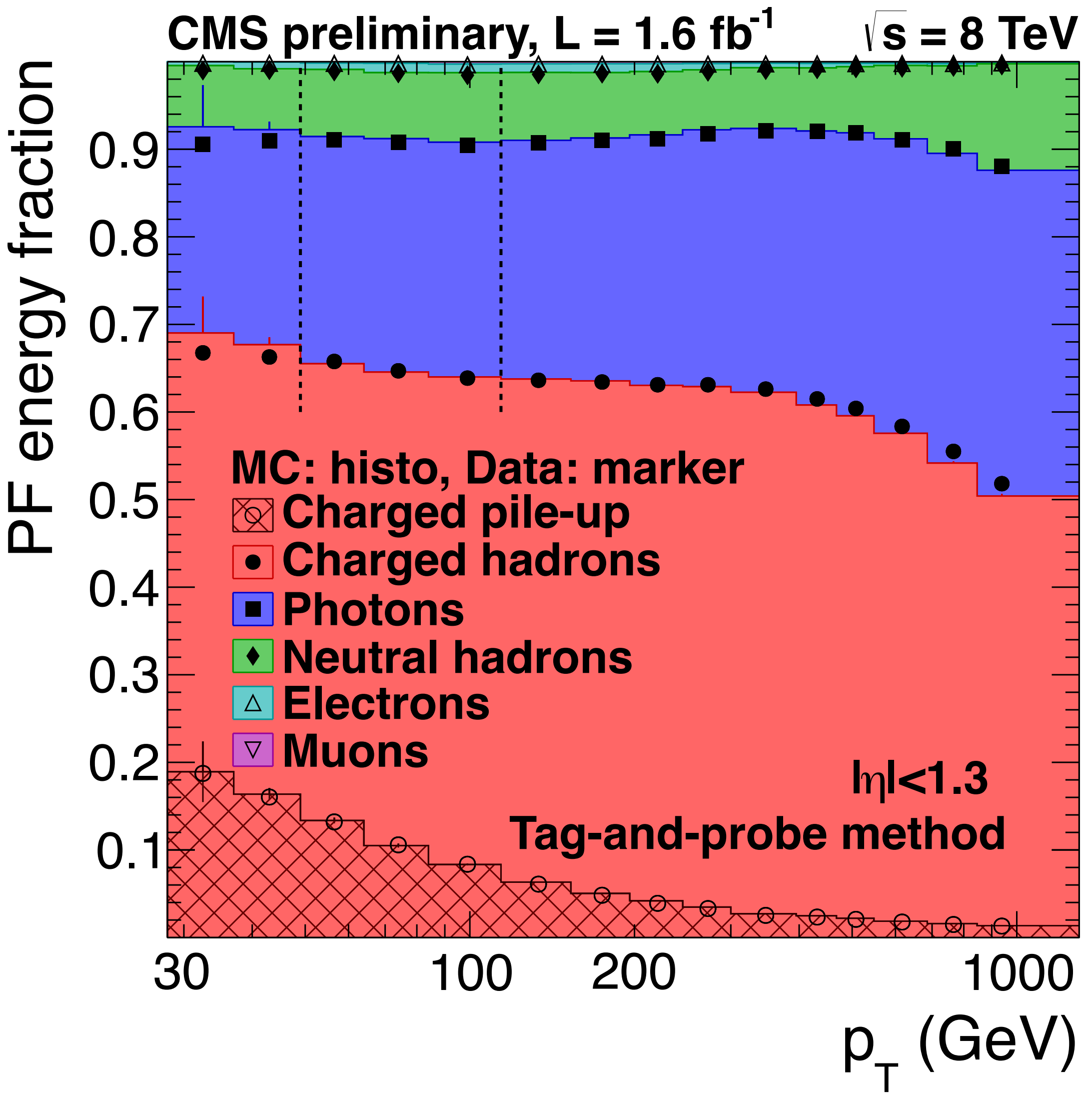
Correction uncertainties

- ✦Uncertainties in 2012 data comparable to 2010, 2011.
  - Pileup uncertainties increasing due to higher average pileup.

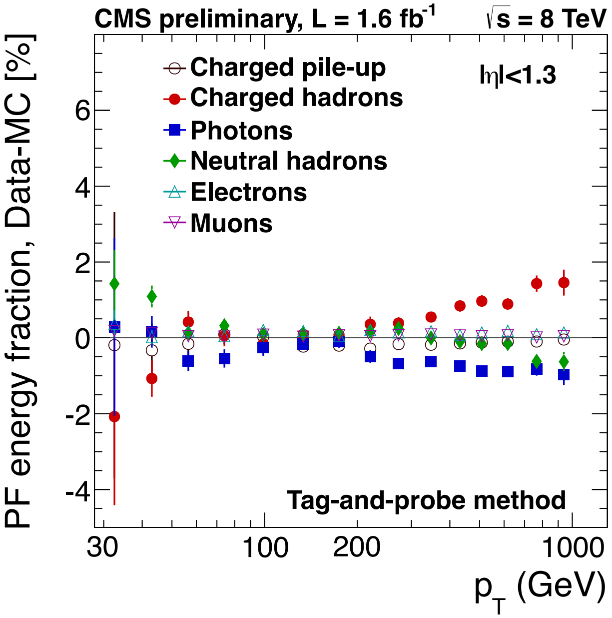


Within 3% for jets with p_T> 30 GeV in barrel

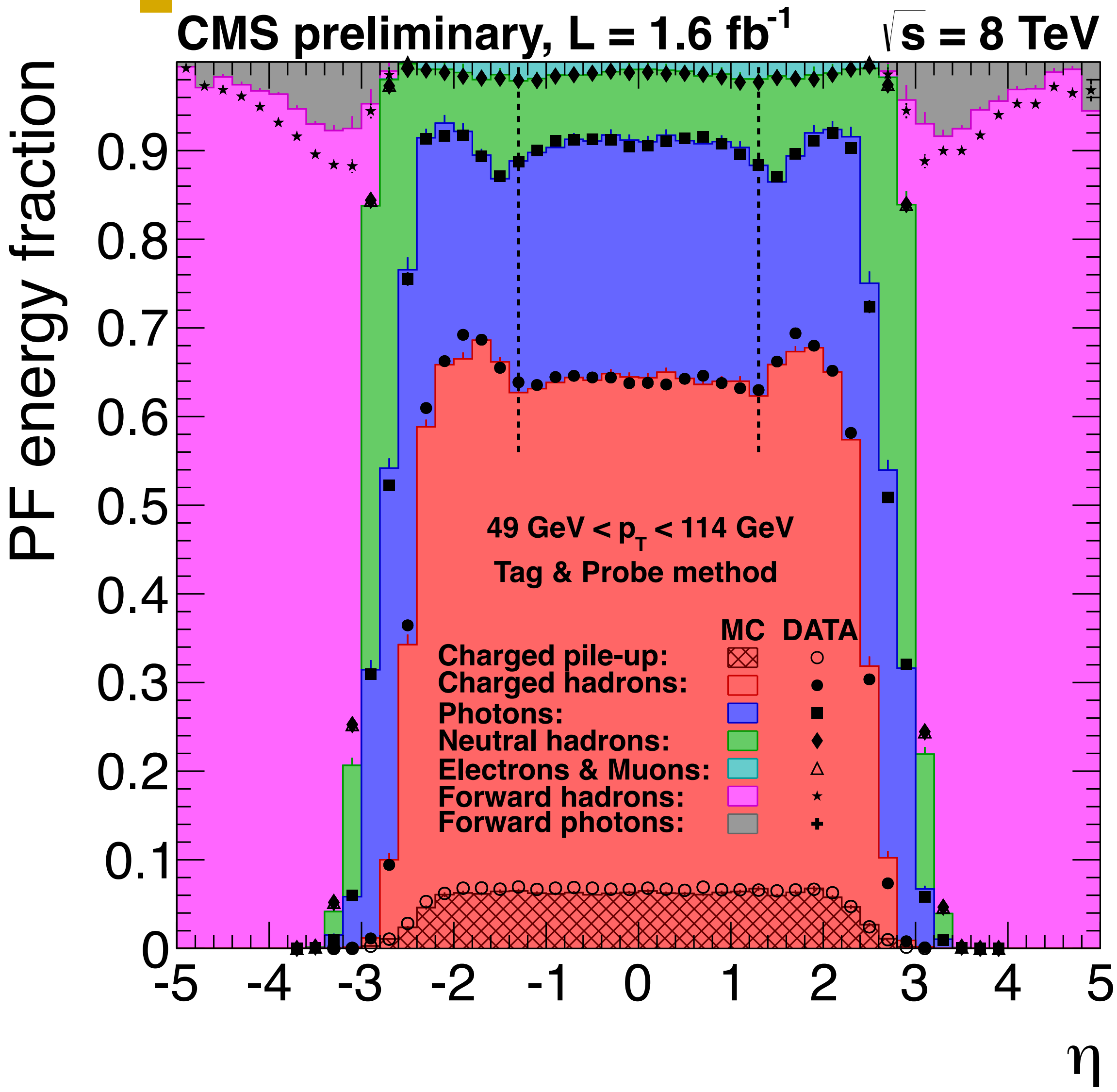
Jet composition vs p_T in barrel



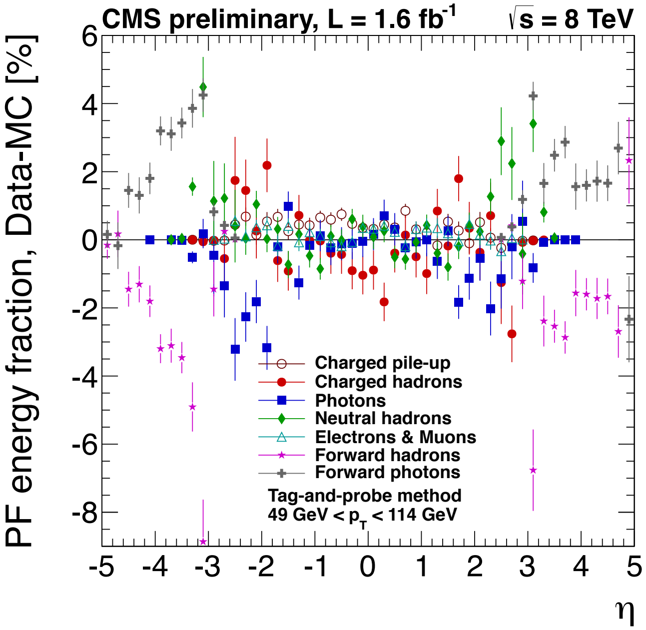
- ✦Jet composition agrees well between Data and MC
  - •consistent with small residual JEC at the 1-2% level.



Jet composition vs η



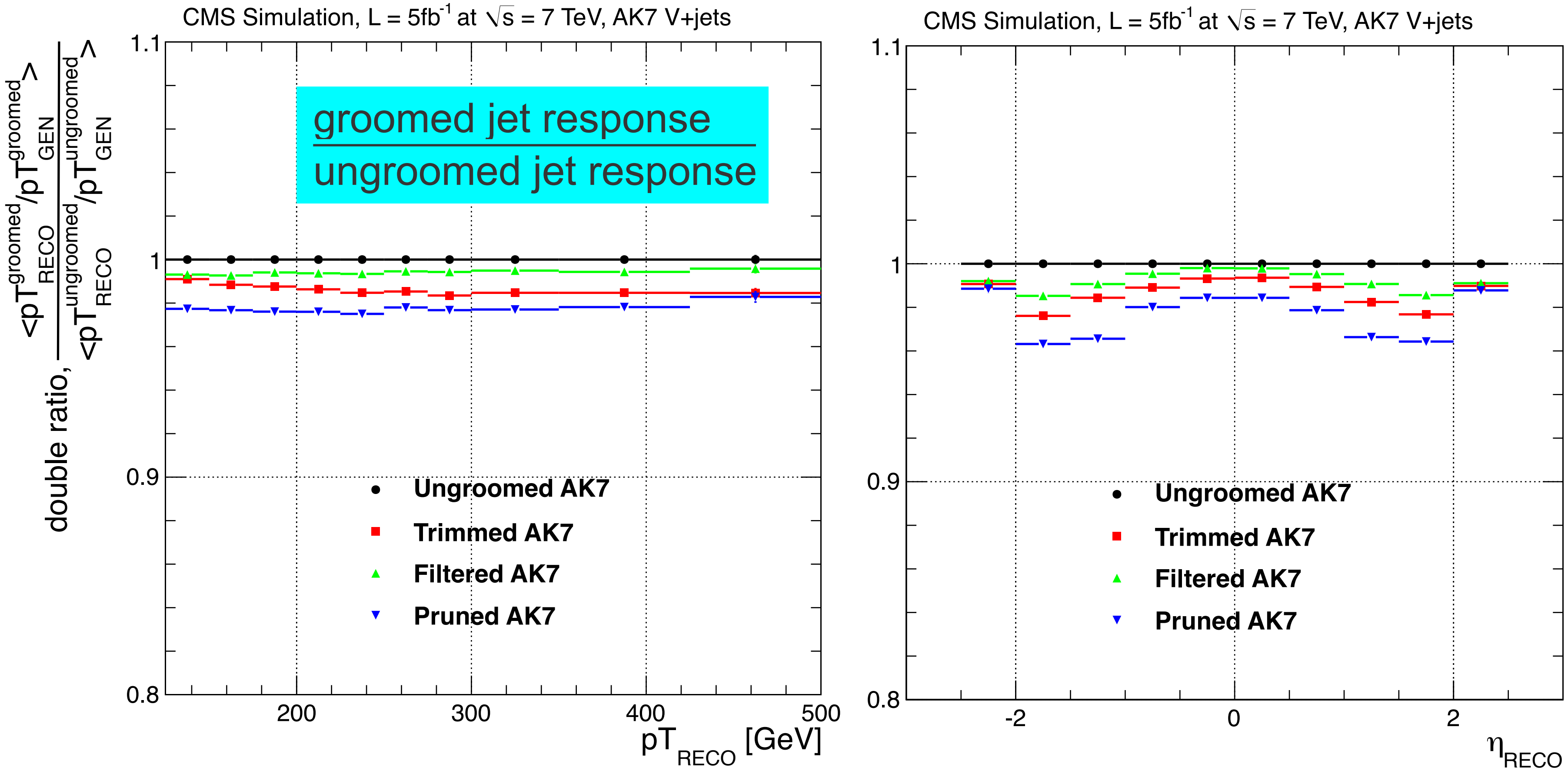
- ✦Jet composition shows increasing differences in the forward region
  - •consistent with JEC at 2-13% level.



Jet p_T response for groomed jets

https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSMP12019

Most aggressive grooming algorithm, pruned, still exhibits response within a few percent of ungroomed case.

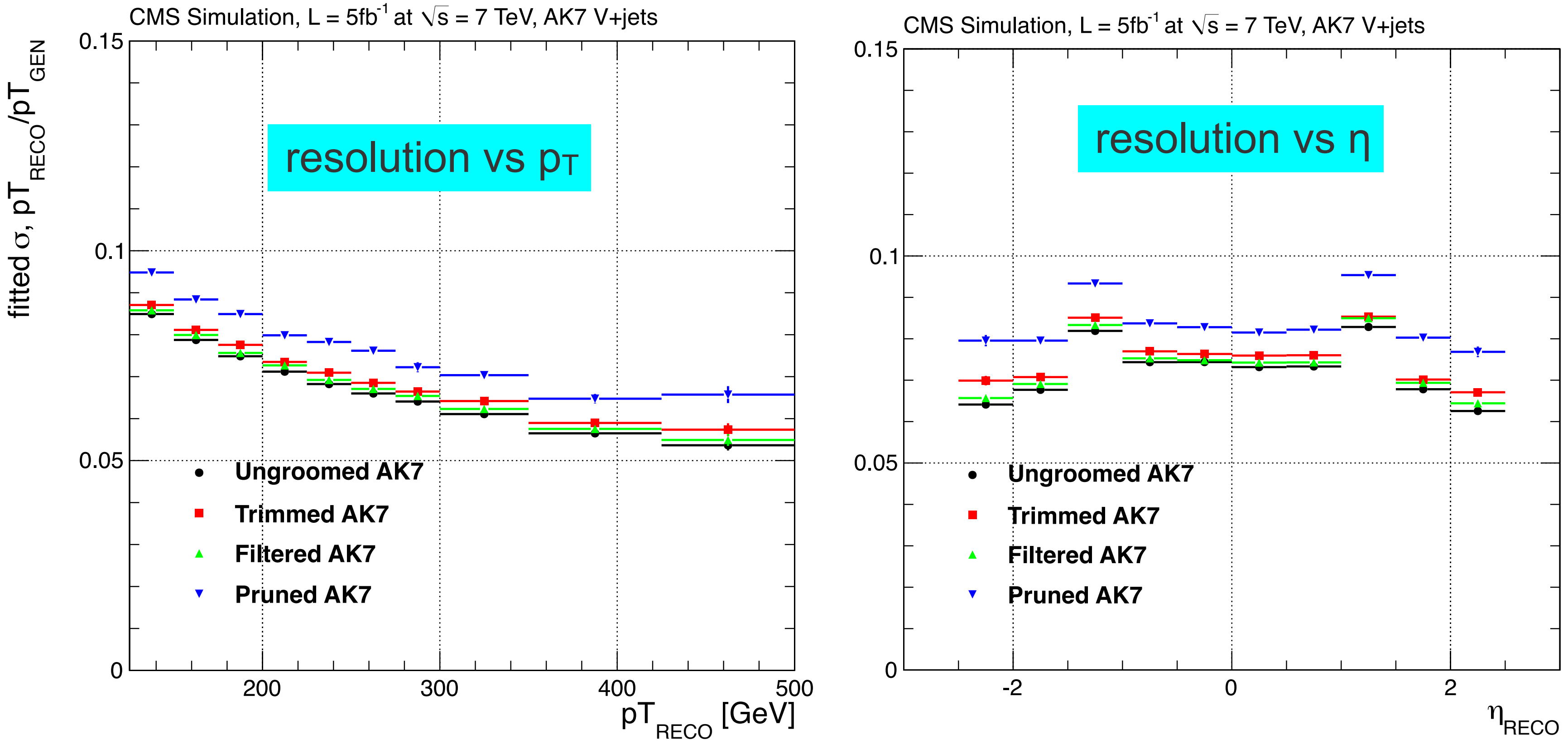


$$\frac{\text{groomed jet response}}{\text{ungroomed jet response}}$$

Jet p_T resolution for groomed jets

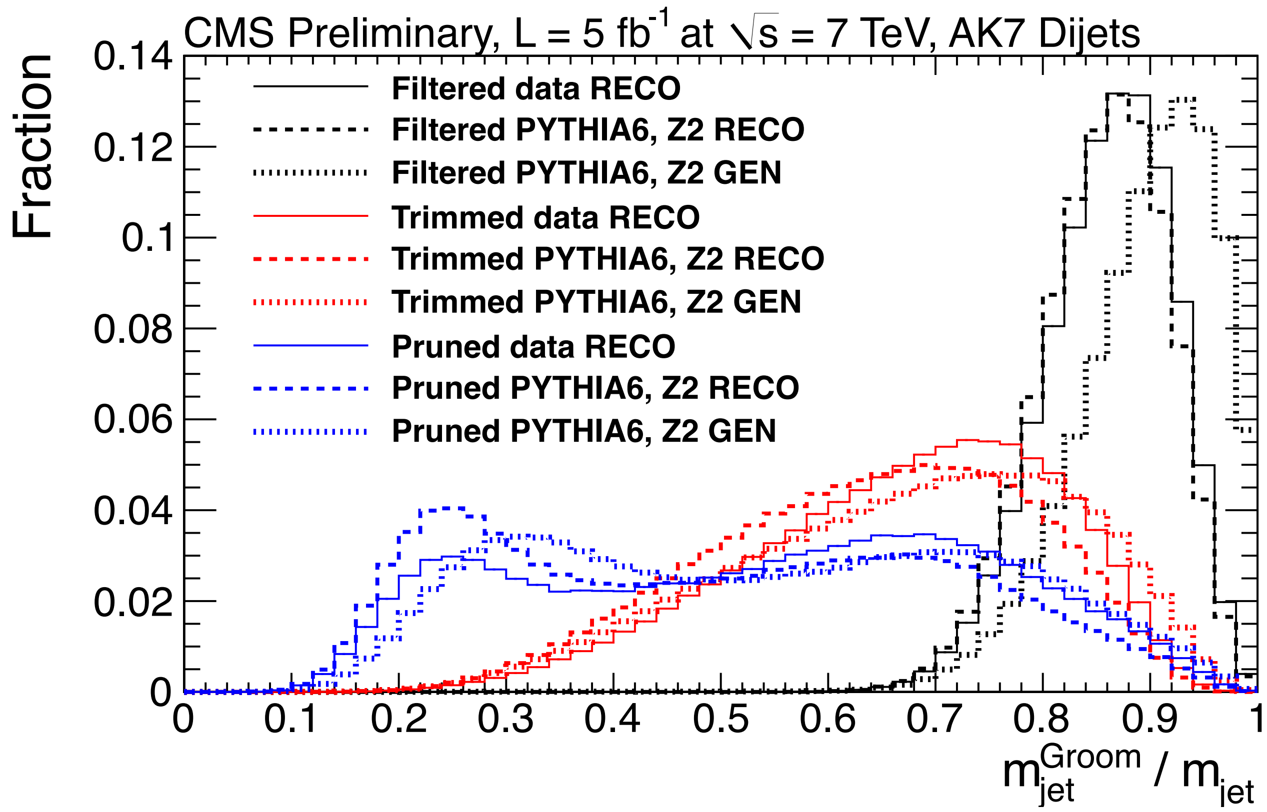
Jet p_T resolution for various grooming algorithms also shows good agreement to within a few percent.

- **Pruned jet p_T resolution degraded slightly.**



Jet mass for groomed jets

Ratio of the groomed to ungroomed jet mass, encapsulates behavior of each grooming algorithm



Pruning is the most aggressive, filtering is the least aggressive

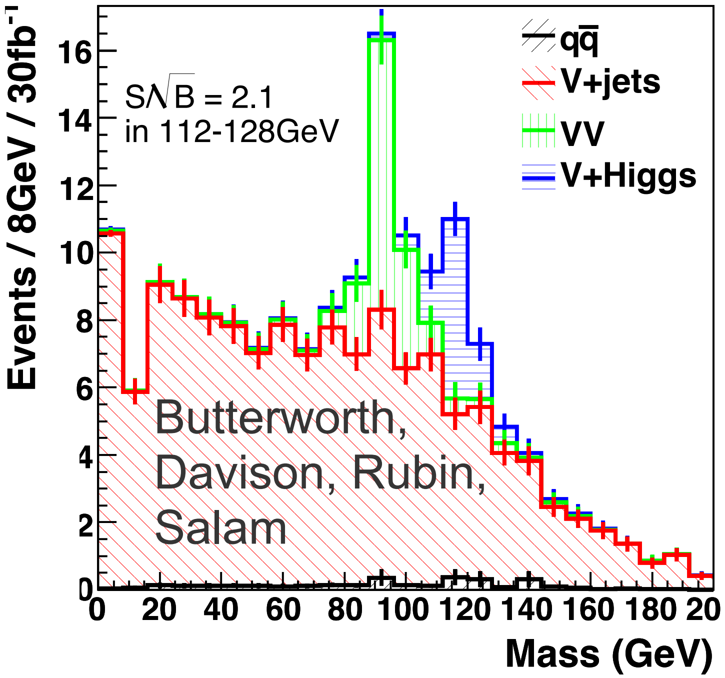
See Nhan Tran's talk (Monday) for details.

Comparison of grooming algorithms at particle level (GEN), reconstructed simulation (RECO) and data

Grooming helps big in reconstructing resonances

Can be used to improve sensitivity of hadronic decays of boosted heavy particles such as Higgs, W/Z, and top

This is what we aim to do



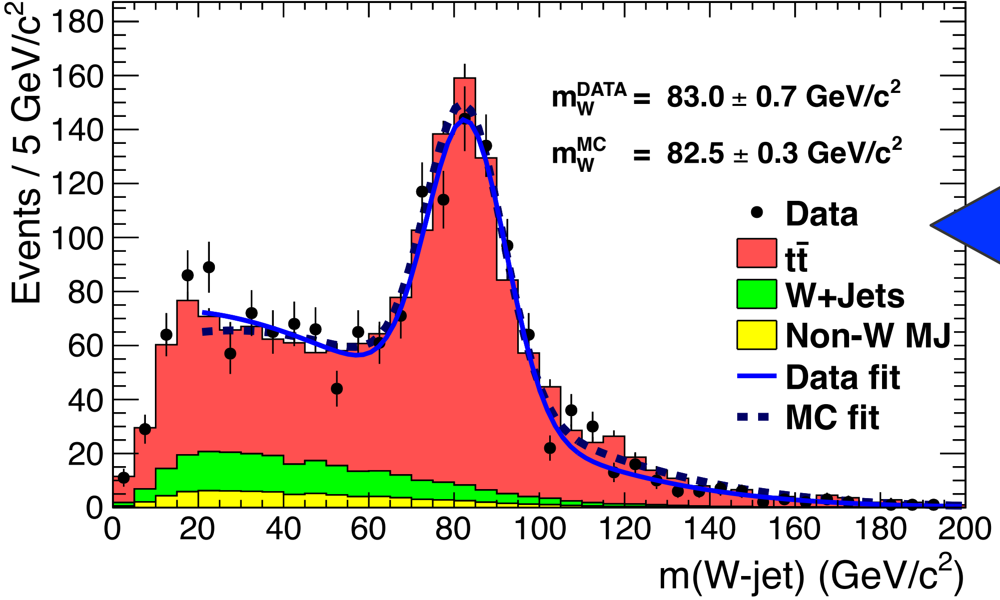
Butterworth, Davison, Rubin, Salam

arXiv: 0802.2470

Started with hadronic W in boosted top events

http://cdsweb.cern.ch/record/1370237

CMS, L = 5 fb^{-1} at $\sqrt{s}$ = 7 TeV



See talk by Derek Strom (Wednesday) for more details.

Performance versus pileup

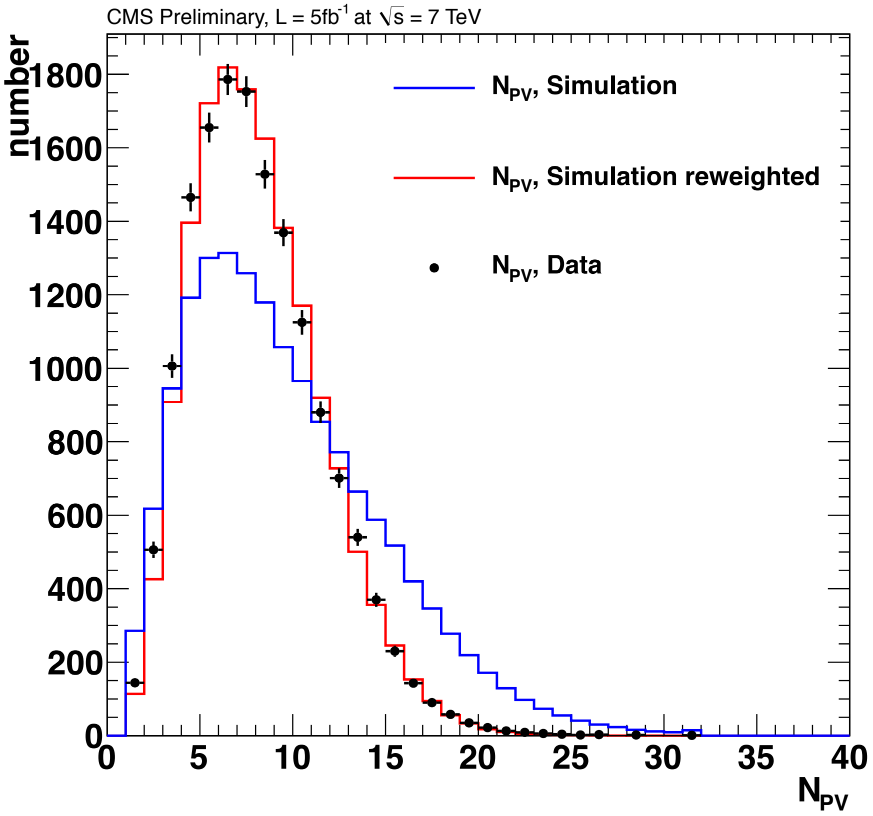
https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSMP12019

- It is of particular interest to understand the sensitivity of large size jets in presence of pileup

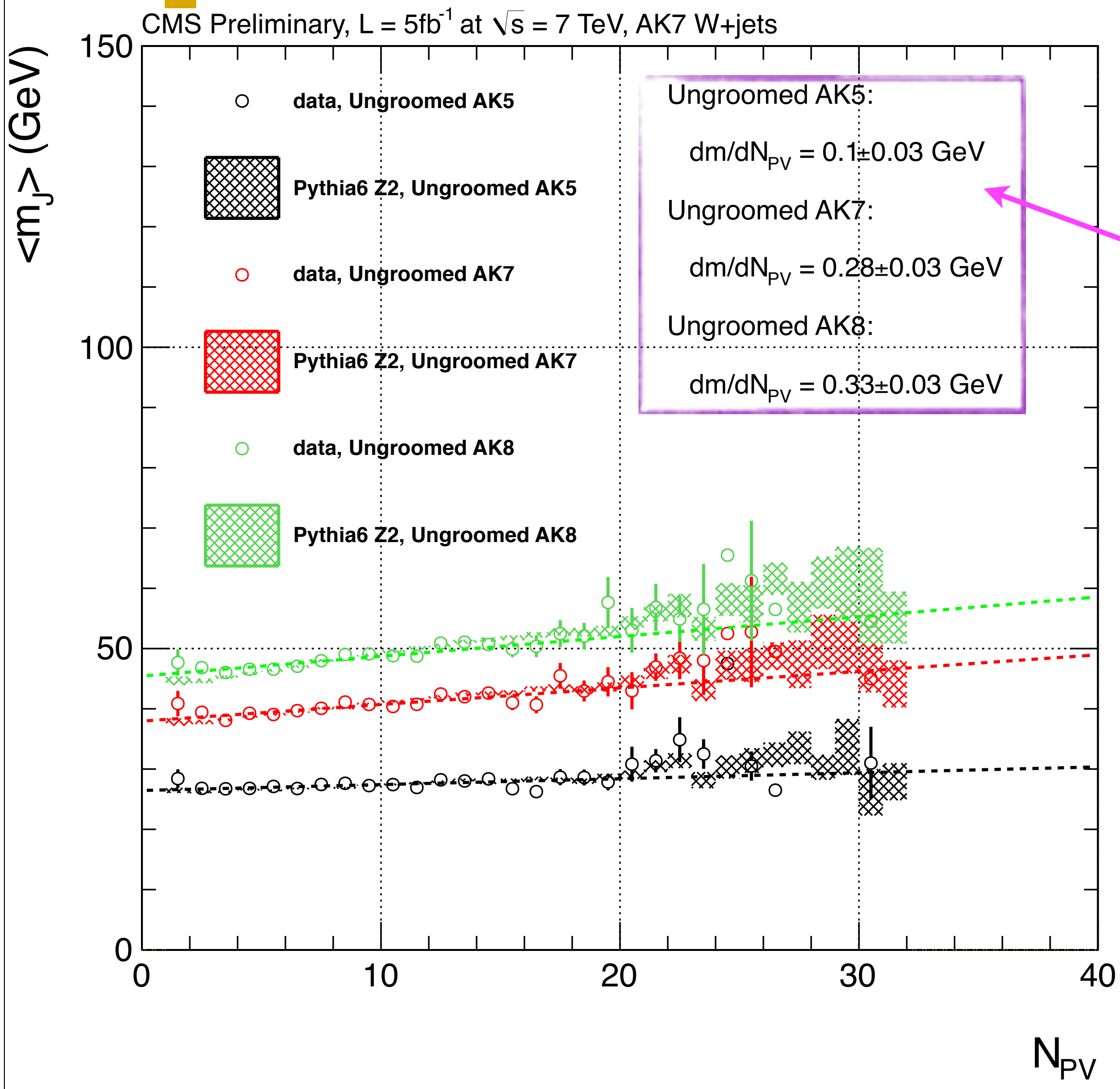
- Grooming techniques may serve to mitigate pileup sensitivity by effectively reducing the jet area

- Understand performance of mean jet mass as a function of number of primary vertices

Pileup profile in 7 TeV data

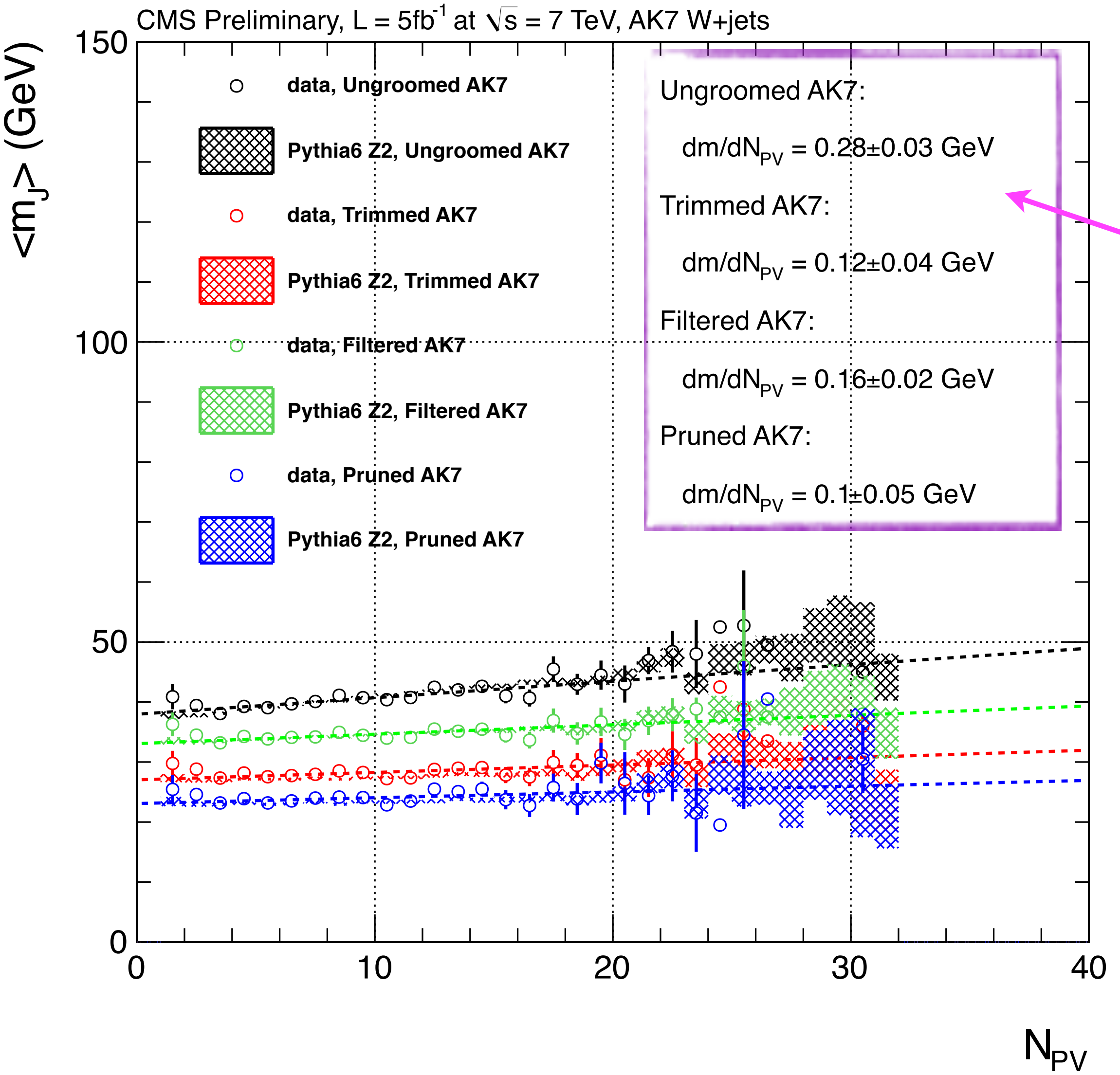


Performance versus pileup by jet size



- ✦Ungroomed jet mass is very sensitive to PU
  - •<mJ> increases linearly as a function of the number of primary vertices
- ✦Effect becomes more pronounced as the jet size increases
  - •AK8 shows much worse effect than AK5

Performance versus pileup for groomed jets (I)



Ungroomed AK7:

dm/dN_{PV} = 0.28±0.03 GeV

Trimmed AK7:

dm/dN_{PV} = 0.12±0.04 GeV

Filtered AK7:

dm/dN_{PV} = 0.16±0.02 GeV

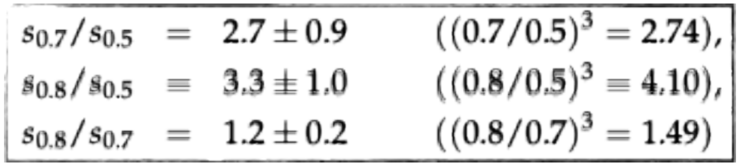
Pruned AK7:

dm/dN_{PV} = 0.1±0.05 GeV

✦Grooming techniques are less sensitive to PU

- $\langle m_J \rangle$ vs NPV slope becomes flatter
- ~ perfectly flat for pruned jets

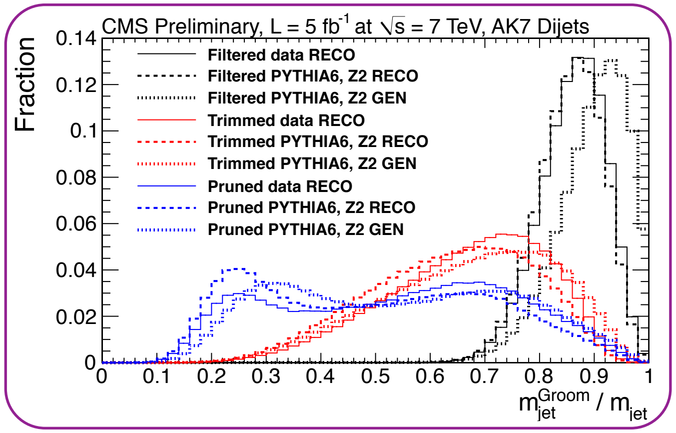
✦Observe the expected behavior that $\langle m_J \rangle$ typically scales as R^3



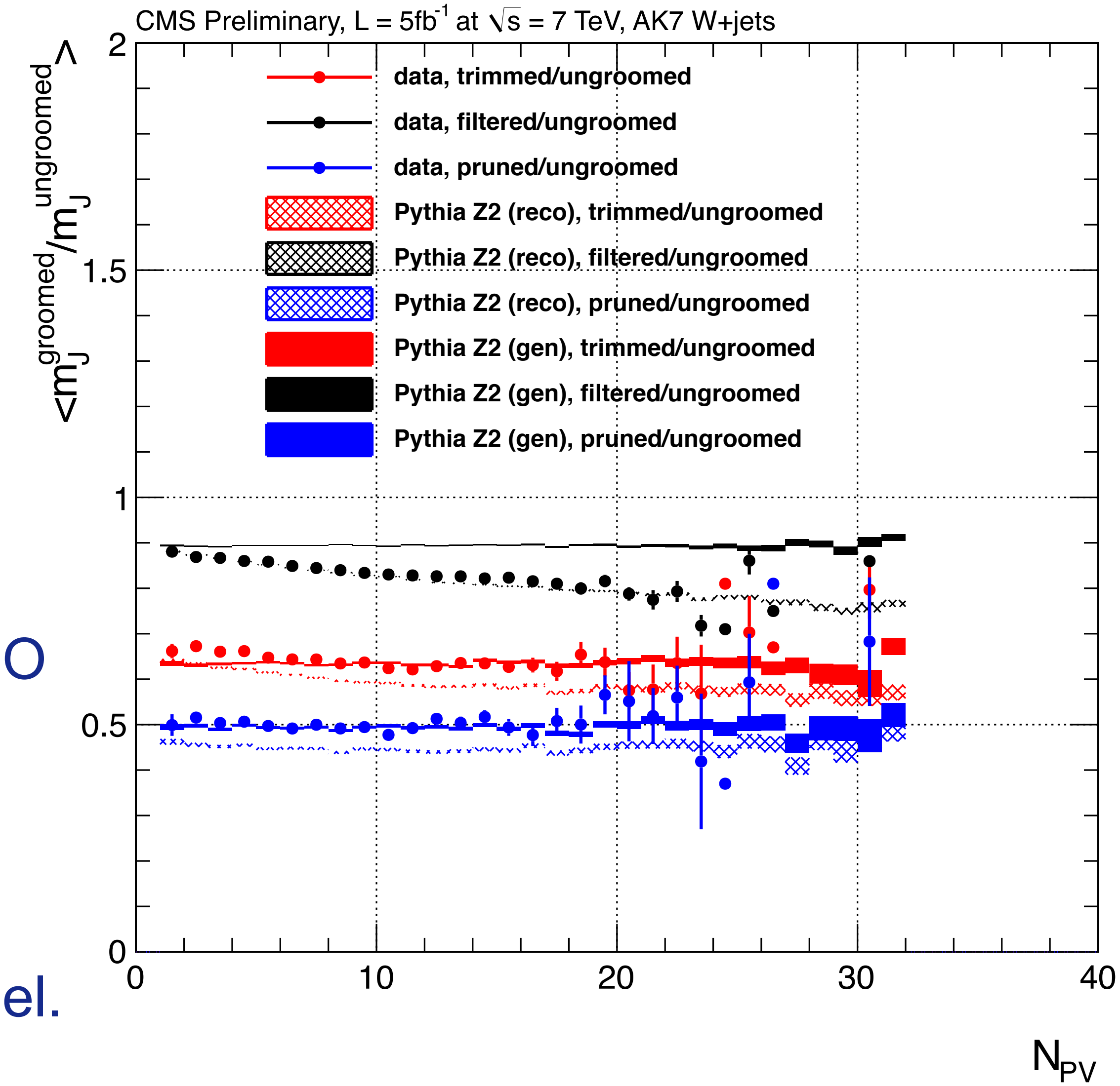
$$s_{0.7}/s_{0.5} = 2.7 \pm 0.9 \qquad ((0.7/0.5)^3 = 2.74),$$
$$s_{0.8}/s_{0.5} = 3.3 \pm 1.0 \qquad ((0.8/0.5)^3 = 4.10),$$
$$s_{0.8}/s_{0.7} = 1.2 \pm 0.2 \qquad ((0.8/0.7)^3 = 1.49)$$

Performance versus pileup for groomed jets (II)

slide 14



- Main differences between GEN Filtered jets and RECO Filtered jets from pileup.
- Trimmed and Pruned jets differences are convoluted with parton showering model.



Kalanand Mishra, Fermilab

Summary

☑ Excellent understanding of jet energy calibrations

☑ Jet composition well understood for jets in tracker coverage

- at the level of a few % (not much room for improvement)
- larger uncertainty for forward jets: 5−10%; need to improve; important for vector boson fusion

☑ Performance of jet p_T and mass versus pileup shows that the grooming techniques lessen sensitivity to pileup

- Important for future high luminosity LHC runs

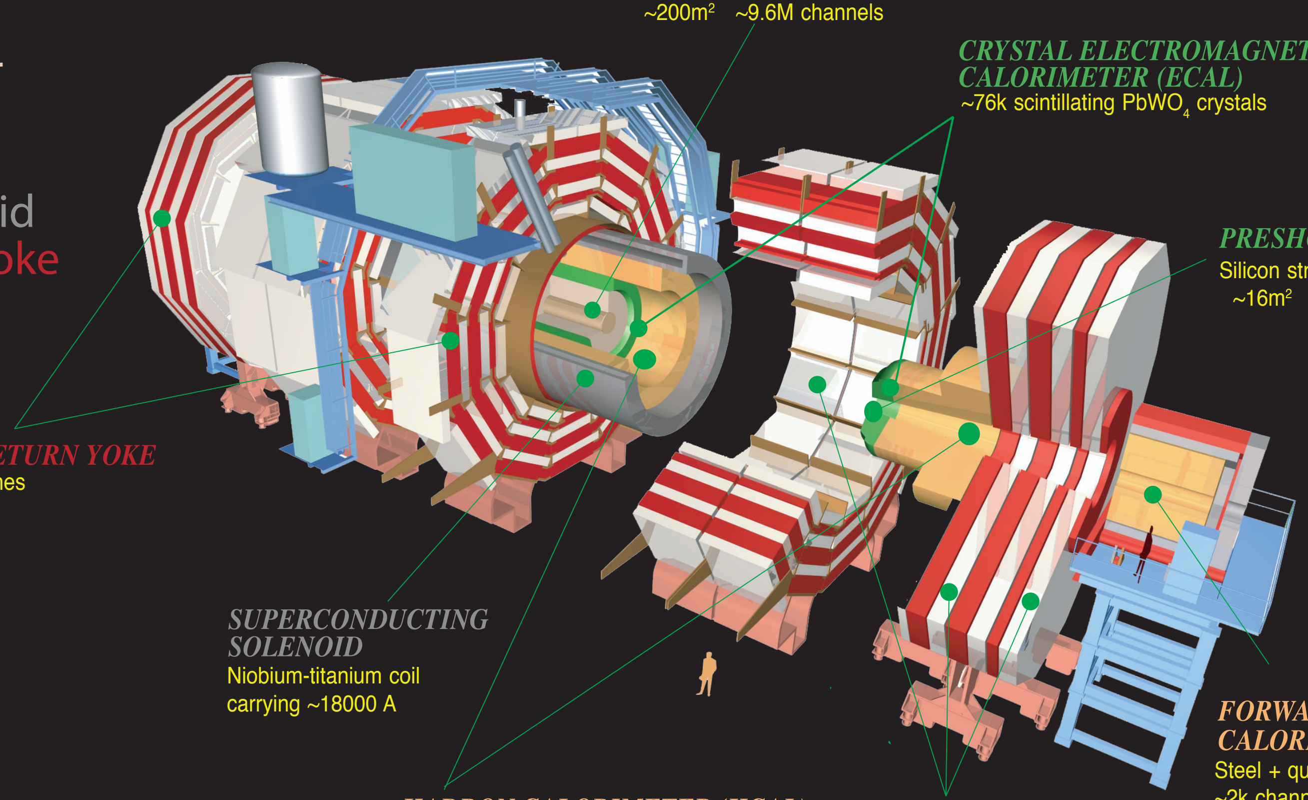
☑ Understanding interplay of various algorithms vs pileup is an important benchmark for searches for new physics

BACKUP SLIDES

Understanding CMS detector

CMS Detector

Pixels
Tracker
ECAL
HCAL
Solenoid
Steel Yoke
Muons



SILICON TRACKER
Pixels (100 x 150 μm^2)
~1m^2 ~66M channels
Microstrips (80-180μm)
~200m^2 ~9.6M channels

CRYSTAL ELECTROMAGNETIC CALORIMETER (ECAL)
~76k scintillating $PbWO_4$ crystals

PRESHOWER
Silicon strips
~16m^2 ~137k channels

STEEL RETURN YOKE
~13000 tonnes

SUPERCONDUCTING SOLENOID
Niobium-titanium coil
carrying ~18000 A

FORWARD CALORIMETER
Steel + quartz fibres
~2k channels

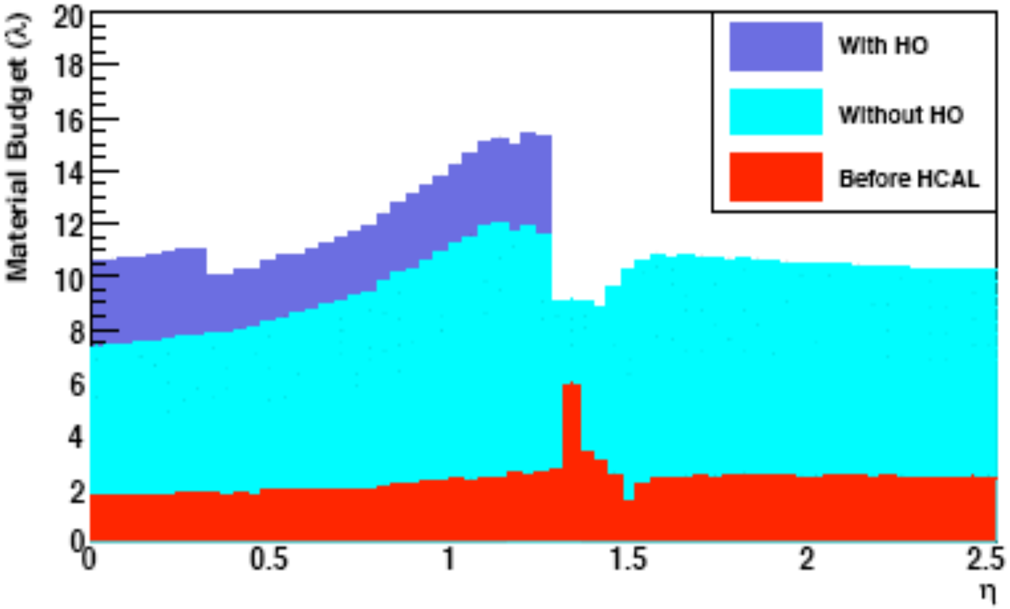
HADRON CALORIMETER (HCAL)
Brass + plastic scintillator
~7k channels

MUON CHAMBERS
Barrel: 250 Drift Tube & 480 Resistive Plate Chambers
Endcaps: 473 Cathode Strip & 432 Resistive Plate Chambers

Total weight	**: 14000 tonnes**
Overall diameter	**: 15.0 m**
Overall length	**: 28.7 m**
Magnetic field	**: 3.8 T**

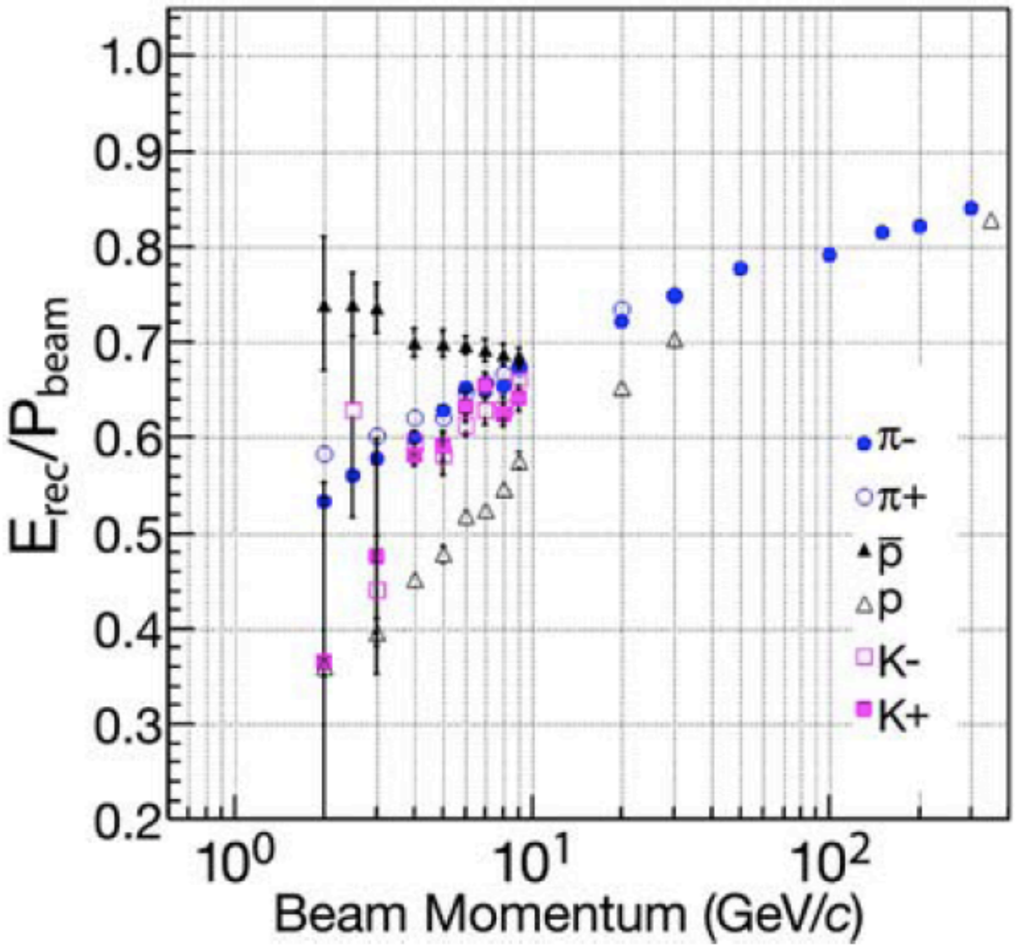
Material budget of the calorimeter

Thickness of HCAL in terms of interaction lengths



7-8 Interaction Lengths at η=0 with HCAL alone and is insufficient to fully contain the shower generated by pions above 100 GeV

Calorimeter response in test beam data

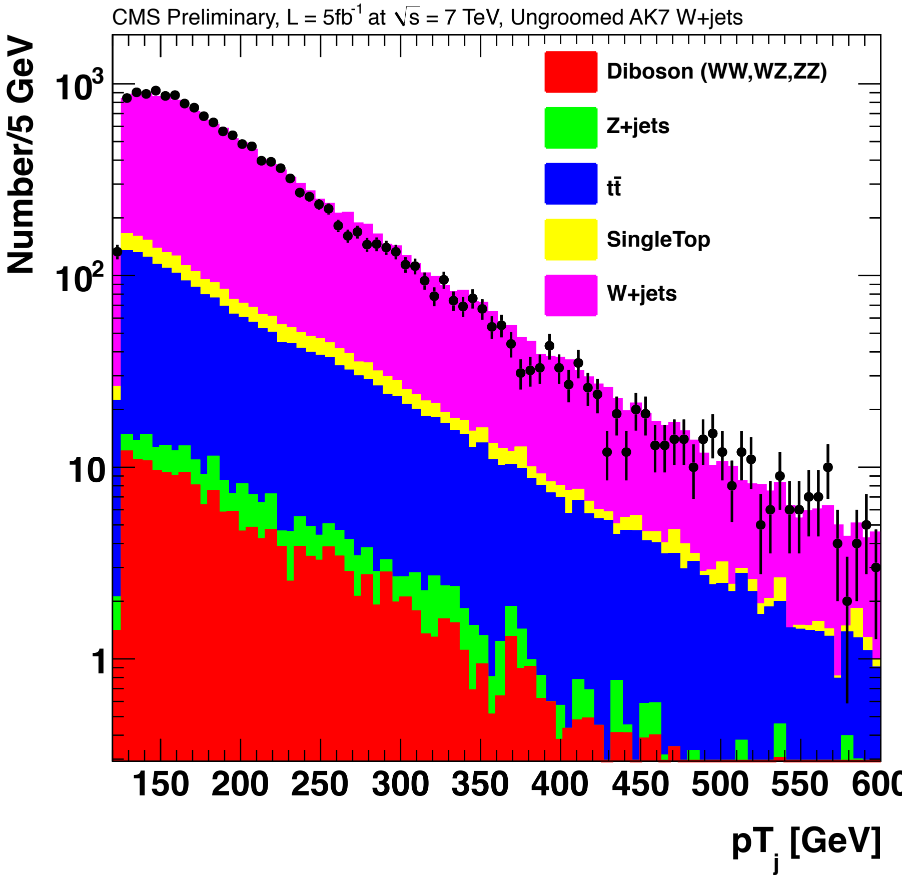


- The figure shows the combined response of EB+HB to different particles as a function of beam momentum.
  - -The response is normalized to 1 for electron.
  - -At 100 GeV, the pion response is 80% of electron.
  - -The proton response is always lower than pion.

- In collision data the response is lower than in test beam because of additional material in front of the calorimeter.
- The calorimeter response is clearly non-linear.

Sample composition

W+jets



Z+jets