Jet tagging at ATLAS

Ayana Arce

for the ATLAS collaboration

BOOST 2013 | Flagstaff, AZ

Motivation for jet tagging tools

linking collision "debris" to parton types allows

- new measurements to improve models of fragmentation
- better tests of the latest new boson's interactions

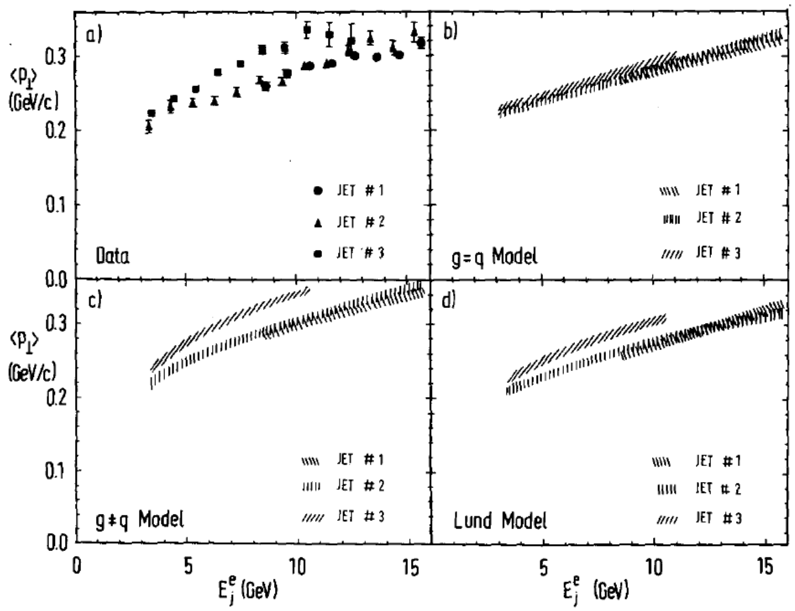


Fig. 2. The average transverse momentum of the charged and neutral particles within a jet relative to the jet axis for the three jets as a function of E_j^e for (a) the experimental data with $E_{cm} > 29$ GeV, and (b), (c), (d) for the prediction of the q = g model, the q ≠ g model ($\sigma_{q,q}$ = 330 MeV, $\sigma_{q,g}$ = 500 MeV), and the Lund model, respectively.

Bartel et. al.; Phys. Lett. 123B (1983)

*tools shown here use observables from era when this was the gluon

Outline

1. evaluating jet taggers at ATLAS
2. tagging $g \rightarrow b\bar{b}$
3. quark/gluon tagging
4. jet charge
5. summary and outlook

common considerations

What are we tagging?

- flavor-labeled jets : we use a functional, leading order definition:
  - apply a jet threshold
  - jet's flavor is the flavor of highest p_T parton near the jet

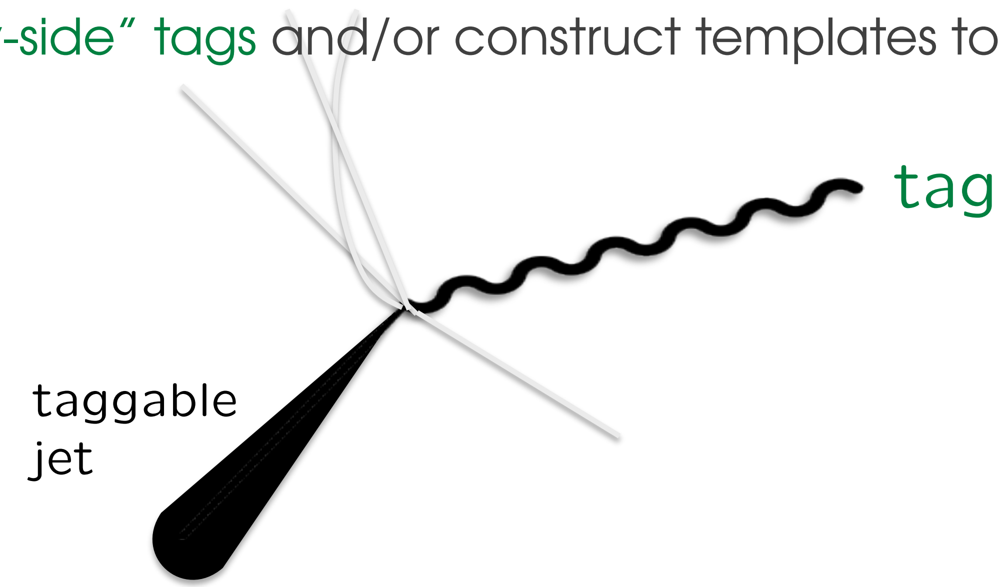
What are we tagging with?

- we want to use taggers now <u>and</u> after 2014, so we must limit pileup sensitivity:
  - prefer track-based observables
  - performance turns out to be close to ideal calo-based tools (and sometimes infinitely better)

common considerations

in-situ validation is…

- especially important for track-based observables
  - these aren't all consistently described by MC models
- a different problem than for boosted object taggers
  - 'signals' and 'backgrounds' are both light jet species so control samples tend to be abundant (but mixed!)
    - use "away-side" tags and/or construct templates to unmix



DOUBLE *B*-HADRON JETS

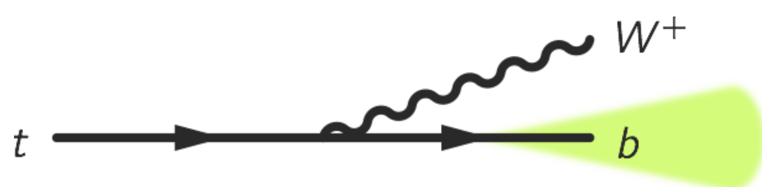
(1) a $g \to b\bar{b}$ tagger

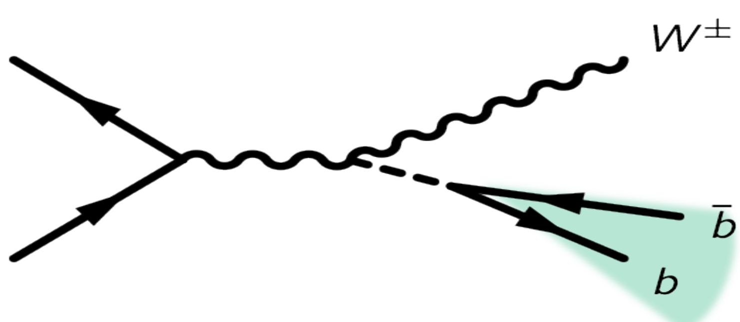
unmerged b-jets

heavy flavor → interesting hard process

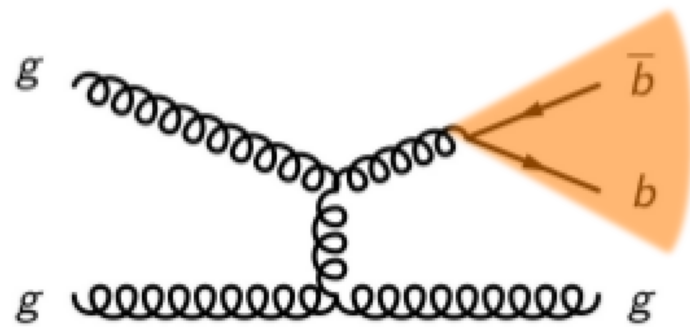
- top quark and H decays
- searches in heavy/3rd generation channels

taggable by displaced vertex (lifetime) tag





ATLAS-CONF-2012-100



merged b-jets

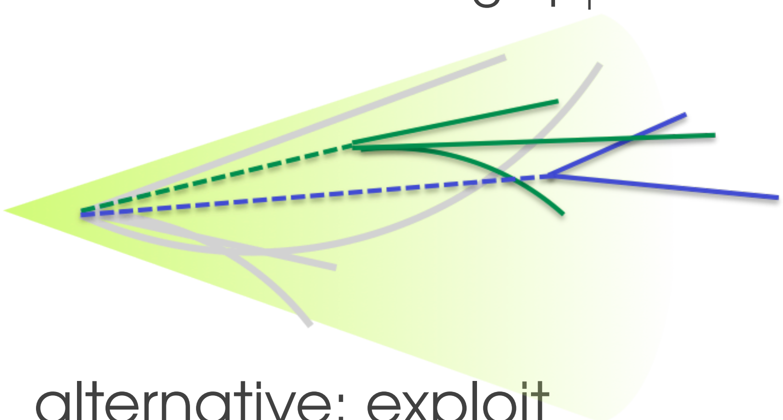
no net heavy flavor

- can't say much about ***t*, *H***
- source of large uncertainty in predicting ***b*** production
- mimic boosted ***X→ bb***

still plenty of displaced vertices

observables for $g \rightarrow b\bar{b}$ tagging

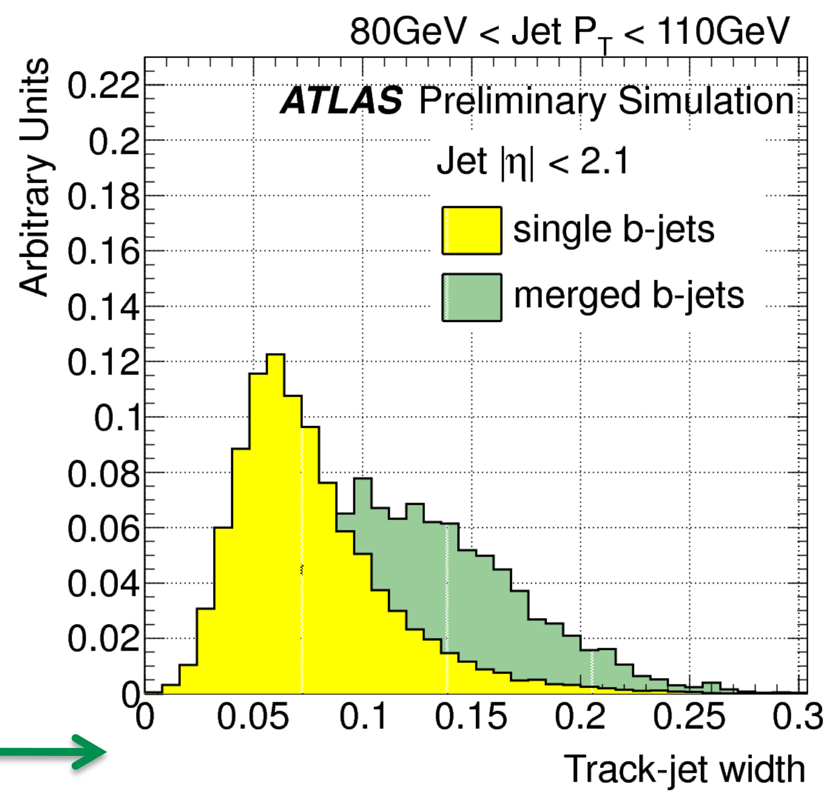
- multiple displaced vertices
  - could be used, but inefficient at high p_T



- alternative: exploit differences in substructure

e.g. track-jet width:

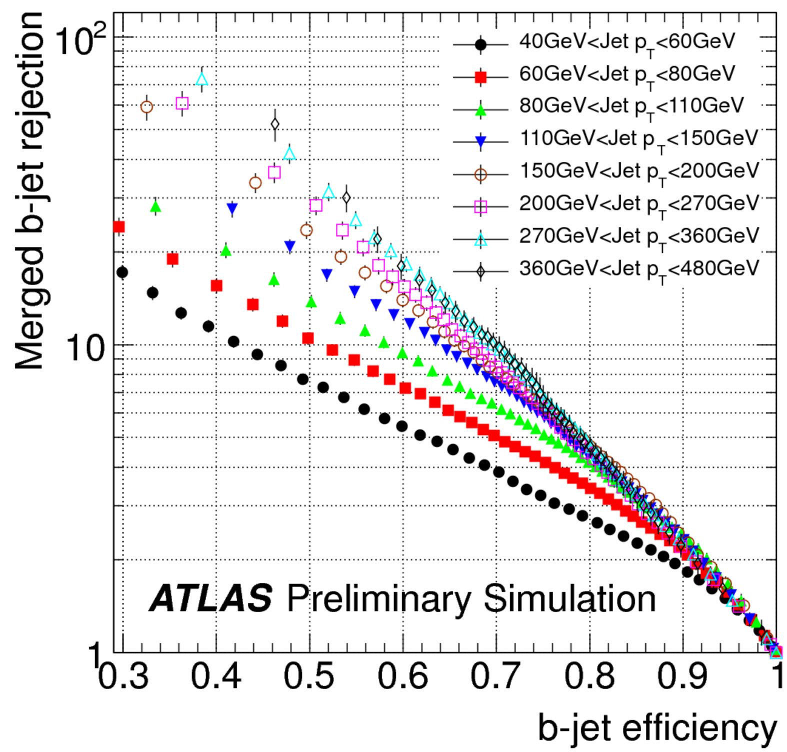
$$\frac{\sum_{i=1}^{N(\mathrm{trk})} \left(p_{T\,i}\, \Delta R(t_i, \mathrm{jet})\right)}{\sum_{i=1}^{N(\mathrm{trk})} p_{T\,i}}$$



ATLAS-CONF-2012-100

tagging $g \rightarrow b\bar{b}$

- likelihood tagger trained on
  - track-jet width
  - track multiplicity
  - separation of first k_T subjets
- separation is good and *improves with p_T*
  - different trend than expected from a vertex approach



ATLAS-CONF-2012-100

QUARK/GLUON JETS

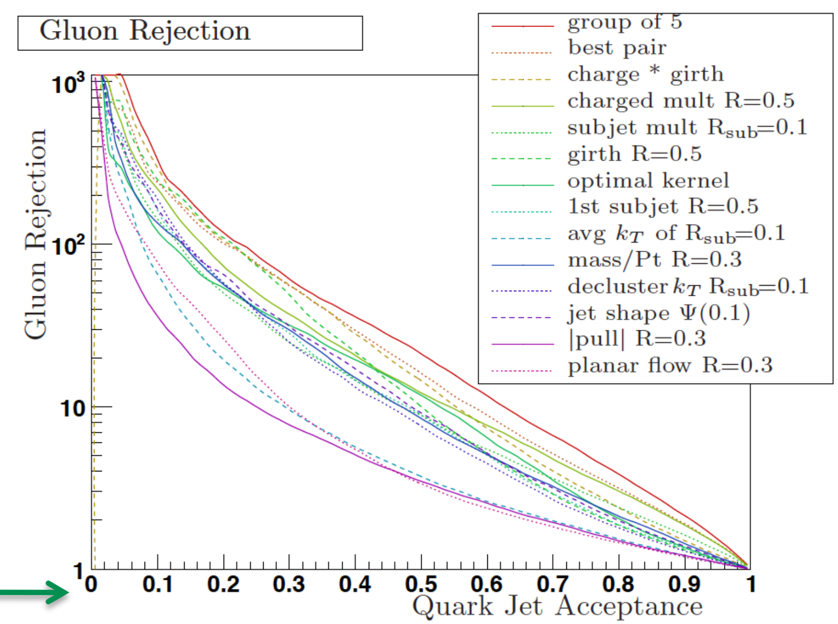
(2) a quark-gluon tagger

use cases:

- many! new physics signals are often quark-enriched
- including q/g likelihood also helps decipher hadronic final states:
  - combinatoric background from ISR gluons

useful observables:

- many! see MC studies here
- fortunately, a few suffice

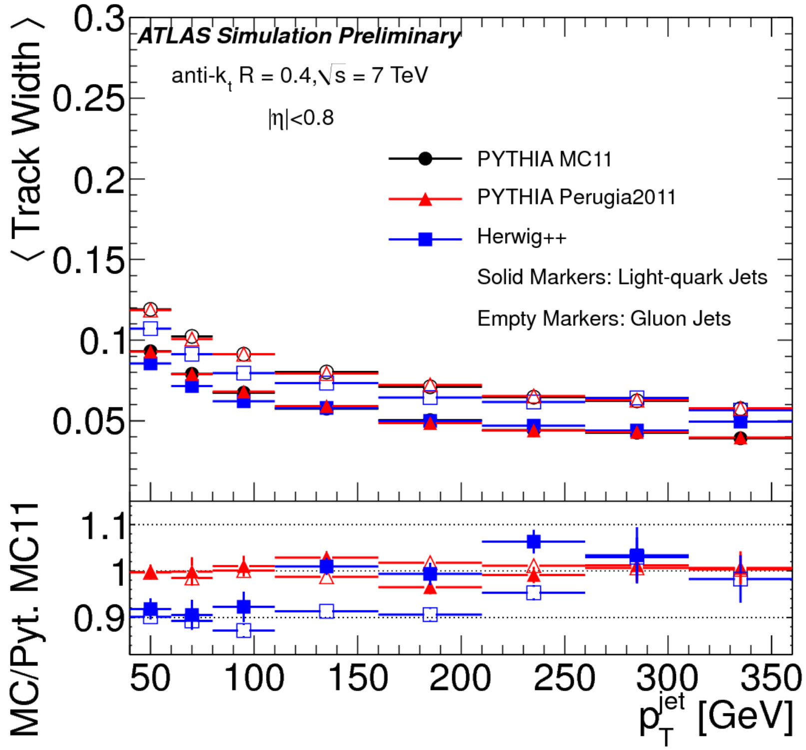


Gallicchio, Schwartz, PRL 107 (2011)

ATLAS-CONF-2012-138

separating quarks and gluons

by jet characteristics



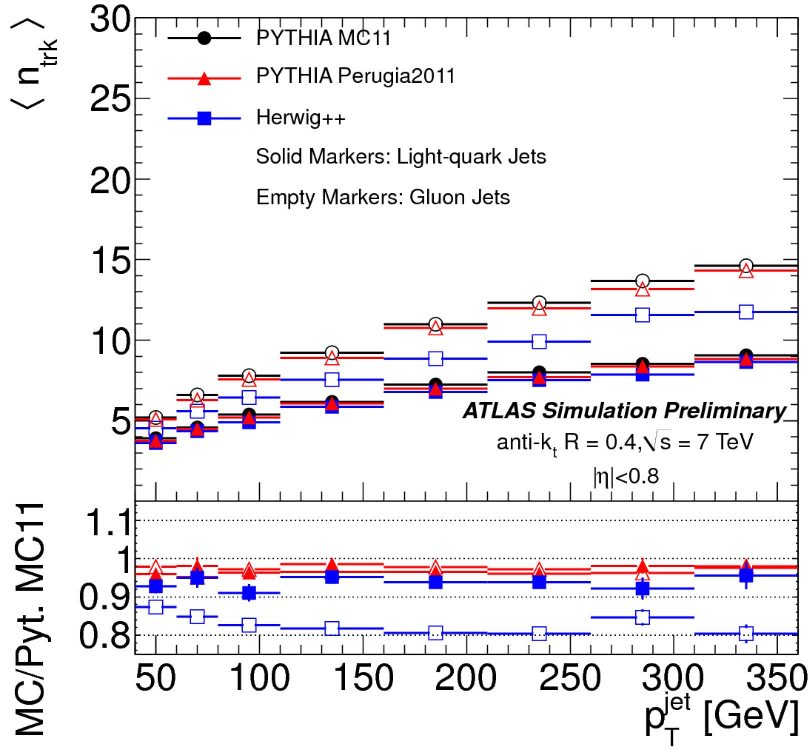
ATLAS-CONF-2012-138

- Good separation found in both 2011 paper and ATLAS study combining
  - track width
  - track multiplicity
- But note strong dependence of $\langle n_{trk} \rangle$ on MC model!

separating quarks and gluons

by jet characteristics

ATLAS-CONF-2012-138

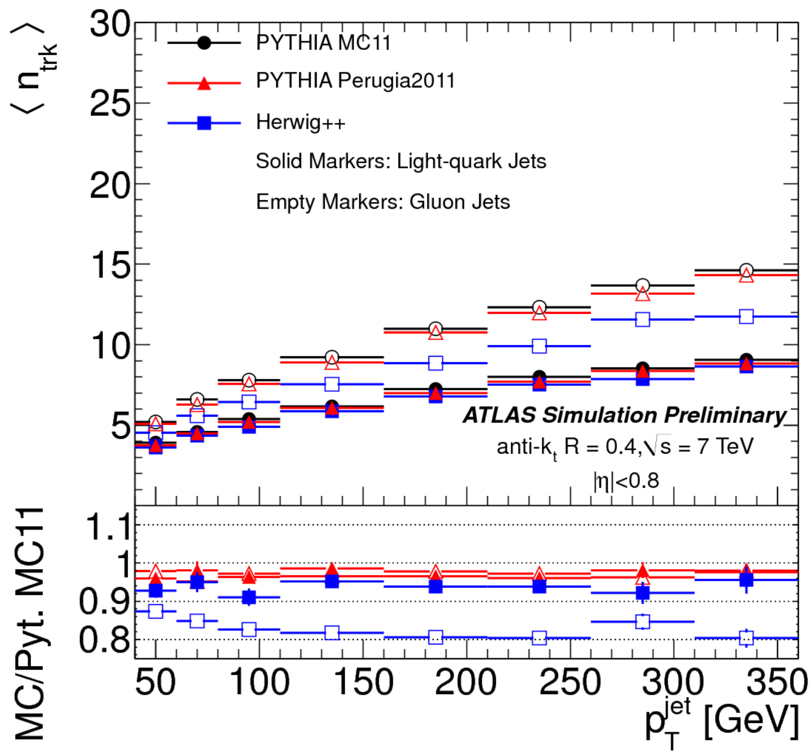


Good separation found in both 2011 paper and ATLAS study combining
  - track width
  - track multiplicity
- But note strong dependence of $\langle n_{trk} \rangle$ on MC model!

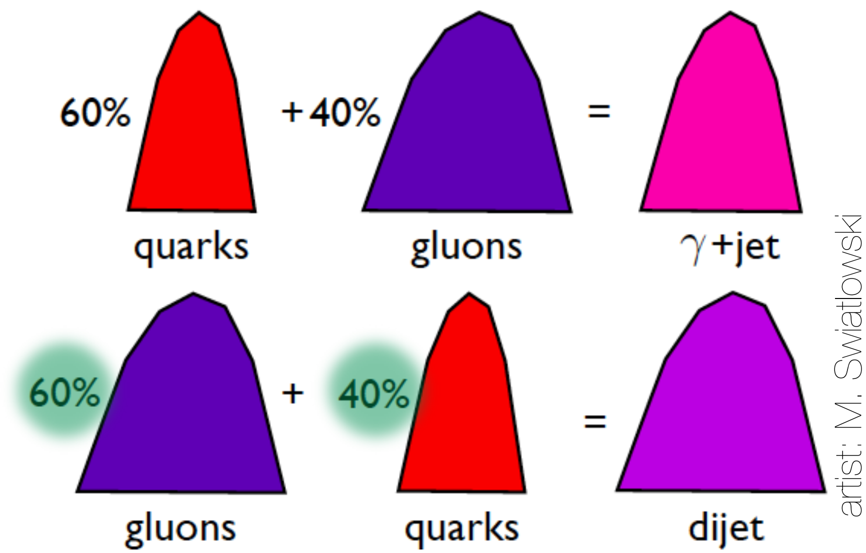
→ need to study performance in data

separating quarks and gluons

by jet characteristics



by event characteristics

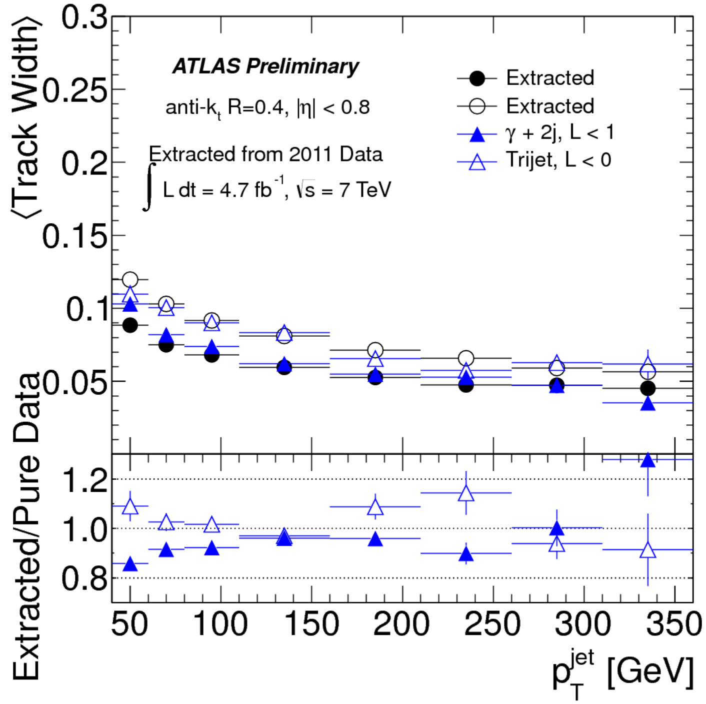


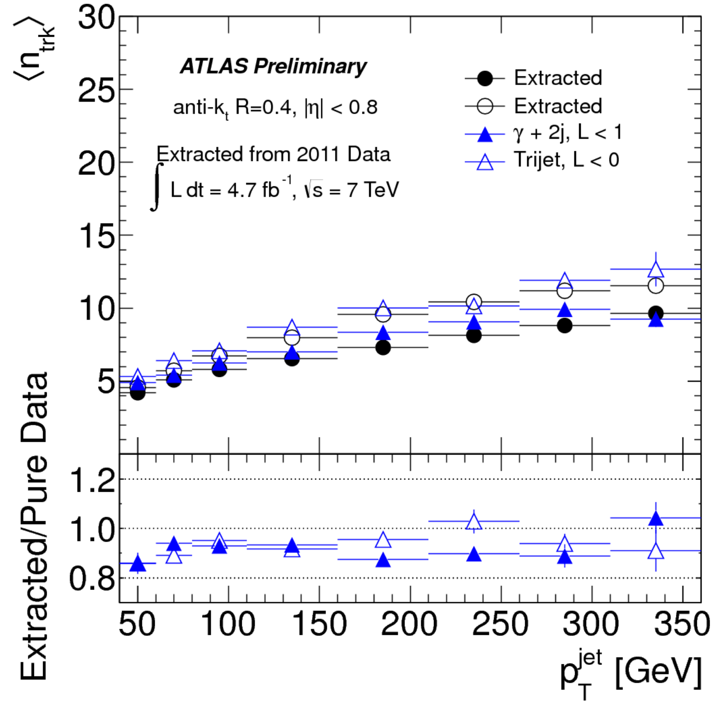
construct width and n_{trk} distributions expected for pure samples

- bin in jet p_T, η ; fix **flavor ratios** to MC predictions
  - also fix heavy flavor templates (shape and normalization)

Solve to extract pure templates

quark/gluon observables in data: width and $\langle n_{trk} \rangle$





ATLAS-CONF-2012-138

- Quark and gluon $\langle n_{trk} \rangle$ are more similar in data
- Cross-check: ~80-90% purified samples
  - close to photon ($\eta_{j1}\eta_{j2} + \Delta R(\gamma, j) < 2$): quark
  - third jet in 3-jet event: gluon if $|\eta_{j3}| - |\eta_{j1} - \eta_{j2}| < 0$

JET CHARGE

(3) a jet-charge tagger

p_T-weighted jet charge was recently studied for LHC
(Krohn et. al PRL 110 (2013))

In an ATLAS toolkit, jet charge would be useful at different scales:

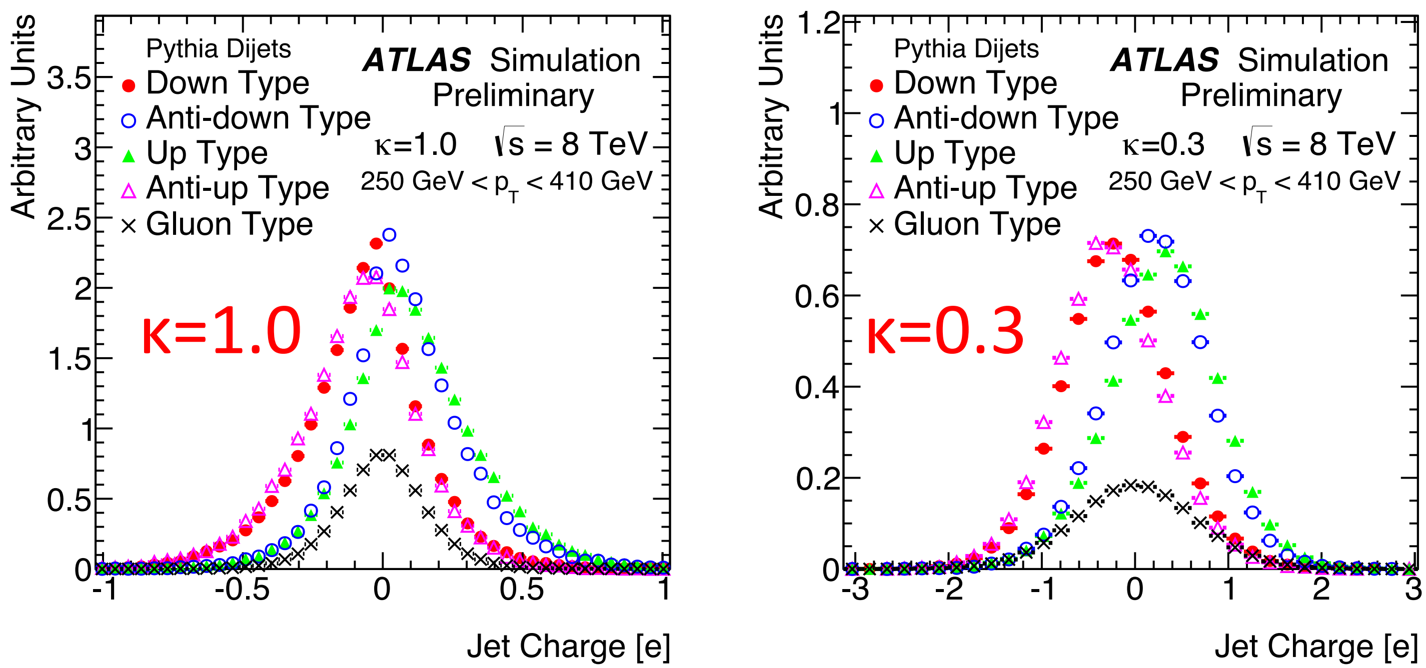
- already proven in top physics (conclusively determined q_{top} = 2/3)
- could differentiate dijet resonances in high m_{jj} searches
- can constrain QCD models and allow studies of tracks in jets from a new angle

$$Q_{\text{jet}} \equiv \frac{\sum_{i=1}^{N(\text{trk})} q_i \left(p_{T\,i}\right)^{\kappa}}{\left(p_T^{\text{calorimeter jet}}\right)^{\kappa}}$$

ATLAS-CONF-2013-086: *new!*

jet charge flavor-separation

- κ parameter tunes sensitivity to soft tracks
  - doesn't significantly affect power to separate charges

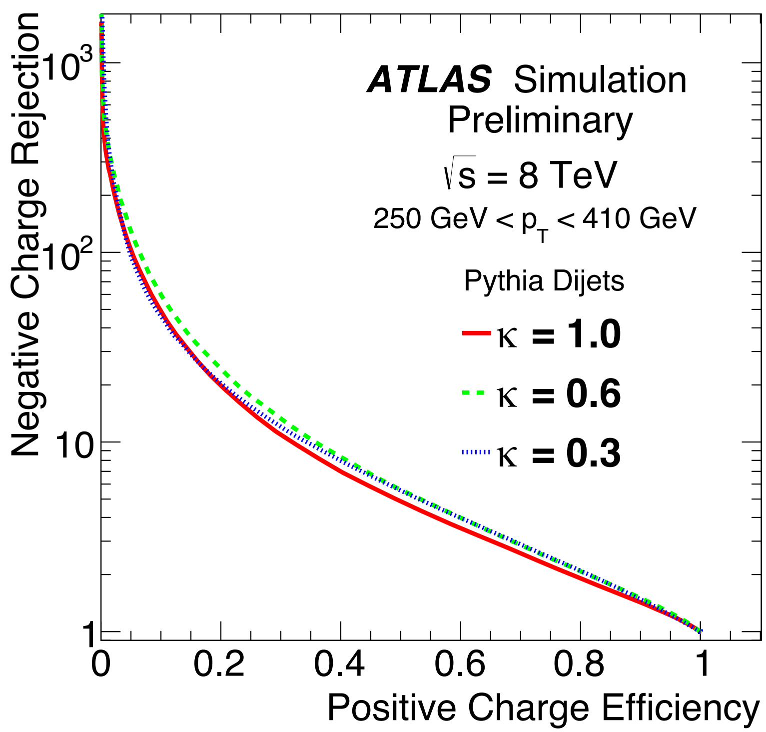


ATLAS-CONF-2013-086

(gluon distributions scaled down)

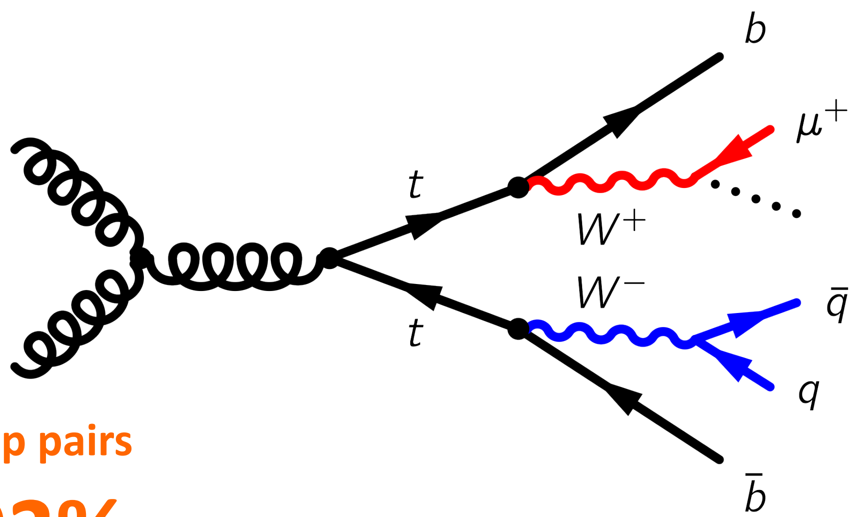
jet charge flavor-separation

- κ parameter tunes sensitivity to soft tracks
  - doesn't significantly affect power to separate charges



ATLAS-CONF-2013-086

testing jet charge in data



Process	N_{events} with μ^+	N_{events} with μ^-
$t\bar{t}$	3575 ± 29	3522 ± 20
Single Top	126 ± 3	97 ± 3
W+jets	170 ± 29	91 ± 15
Z+jets	23± 5	18 ± 3
Dibosons	3 ± 0.4	3 ± 0.3
Total MC	3895 ± 36	3729 ± 25
2012 Data	4095	3893

top pairs

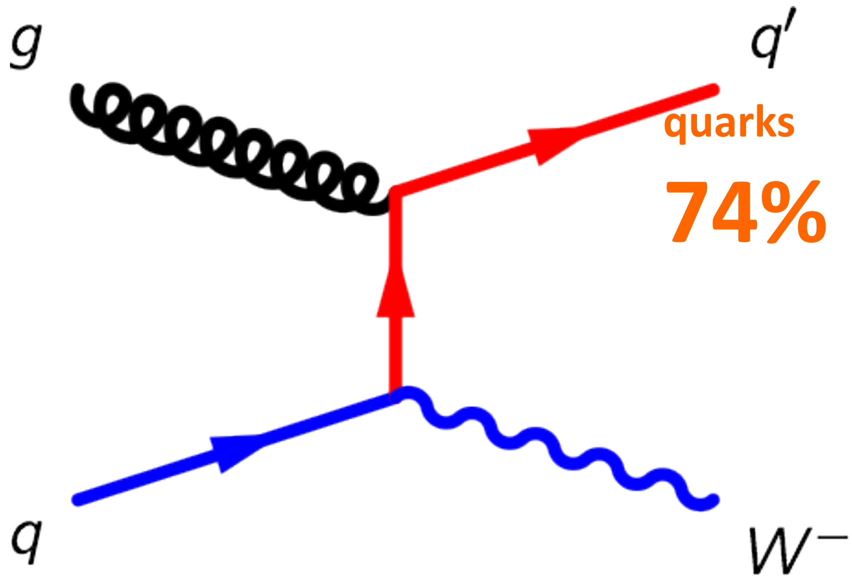
92%

ATLAS-CONF-2013-086

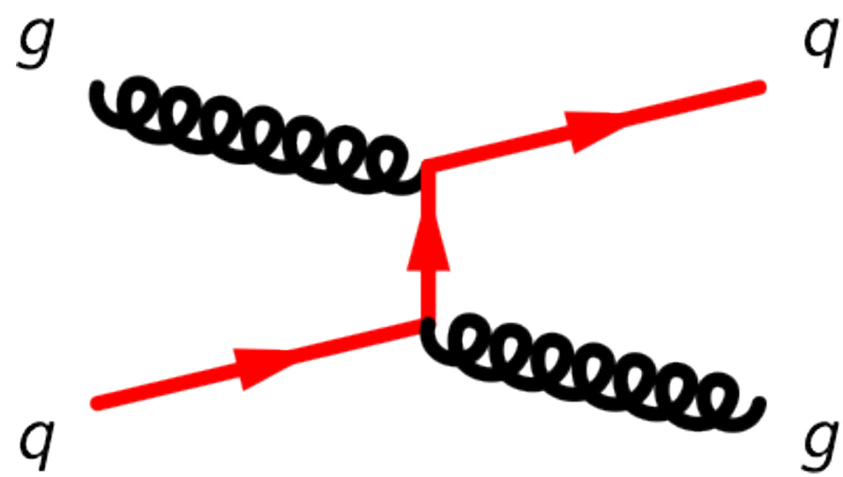
Very pure sample of jet pairs from a particle of known charge:

- use semi-leptonic top pair sample, 5.8 fb^{-1}
  - μ +E_T(miss), ≥4 jets, 2 b-tagged
- kinematically select the two jets from the hadronic W and sum their charges

testing jet charge in data

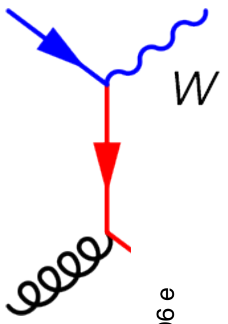


- for single jets, use leptonic W+jets sample, 15.2 fb^{-1}
  - μ + E_T(miss), ≥1 jet: ~80% pure W
  - quark charge is opposite to W charge

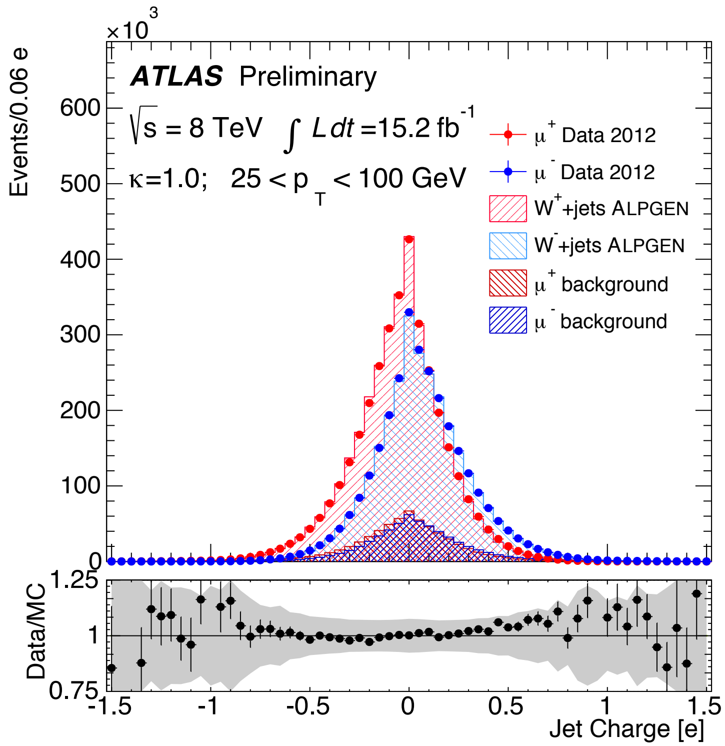


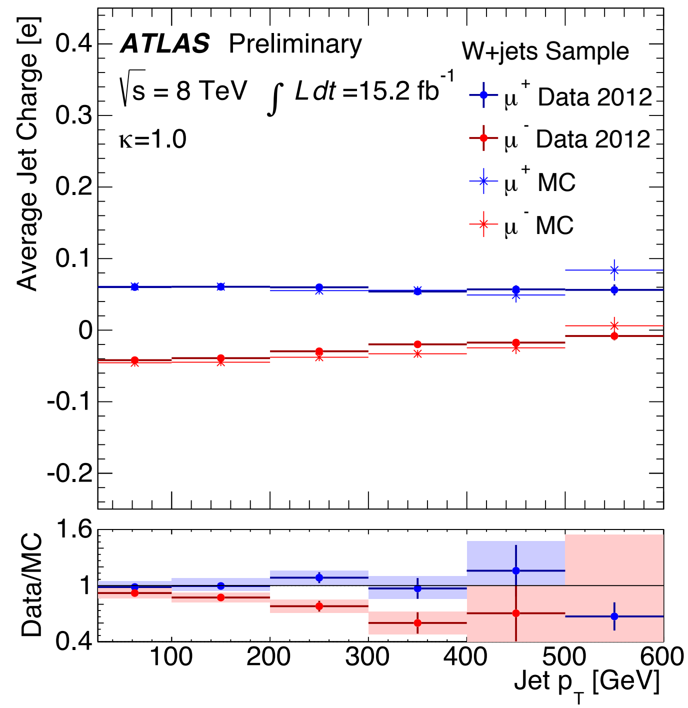
- better statistics but very mixed samples from dijets
  - total dijet charge provides sensitivity to mean jet charge, flavor composition

ATLAS-CONF-2013-086



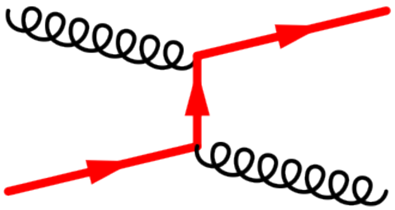
quark-jet charge separation





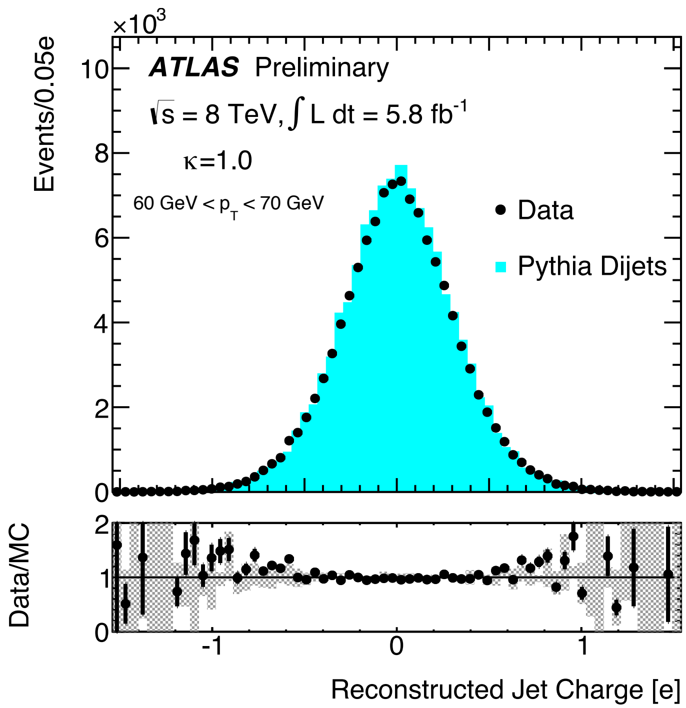
ATLAS-CONF-2013-086

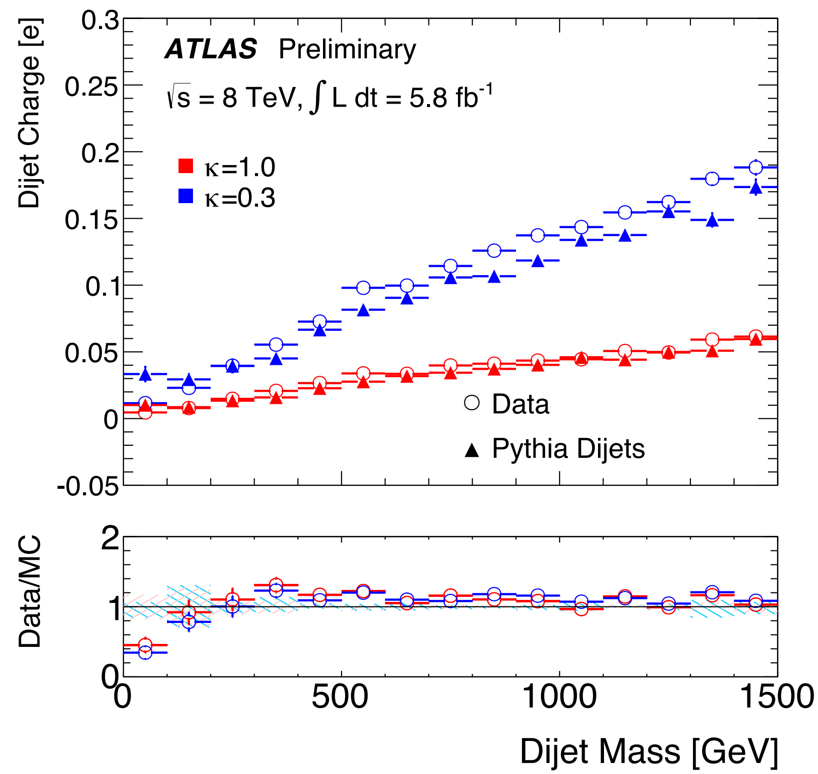
- model (Alpgen+Pythia W+jets) includes W+g
  - mostly charge-symmetric backgrounds (shaded) : multi-jets, top, Z+jets
- a measurement of the p_T dependence can eventually be compared to calculations (Waalewjin, PRD 86 (2012)



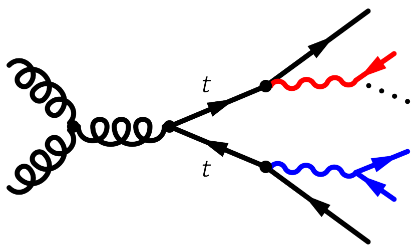
dijet total charge vs. mass

- *sum* of leading two jet charges in inclusive dijet sample is well modeled by Pythia at lower jet p_T



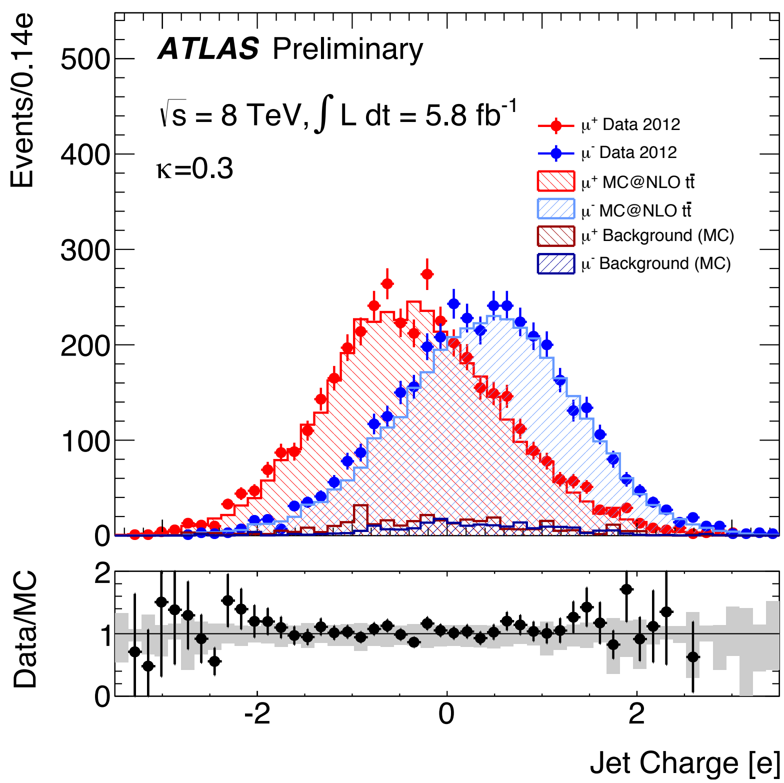


ATLAS-CONF-2013-086

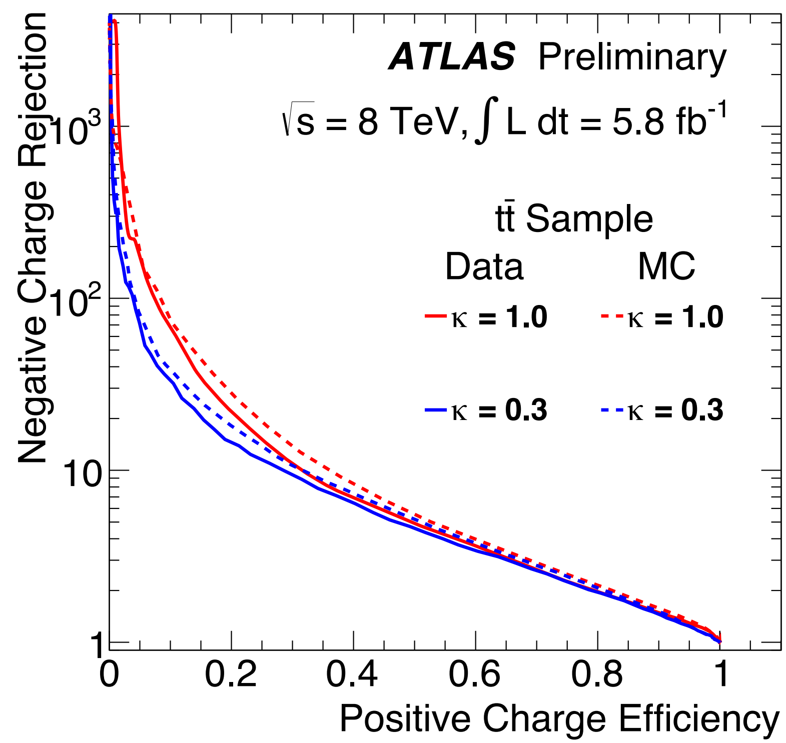


W boson charge tagging (resolved)

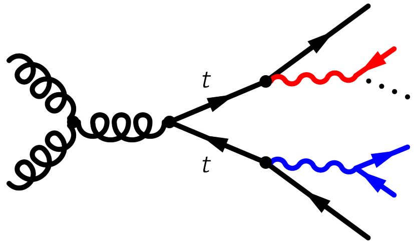
W → qq candidate charge



Performance of a W+ tagger

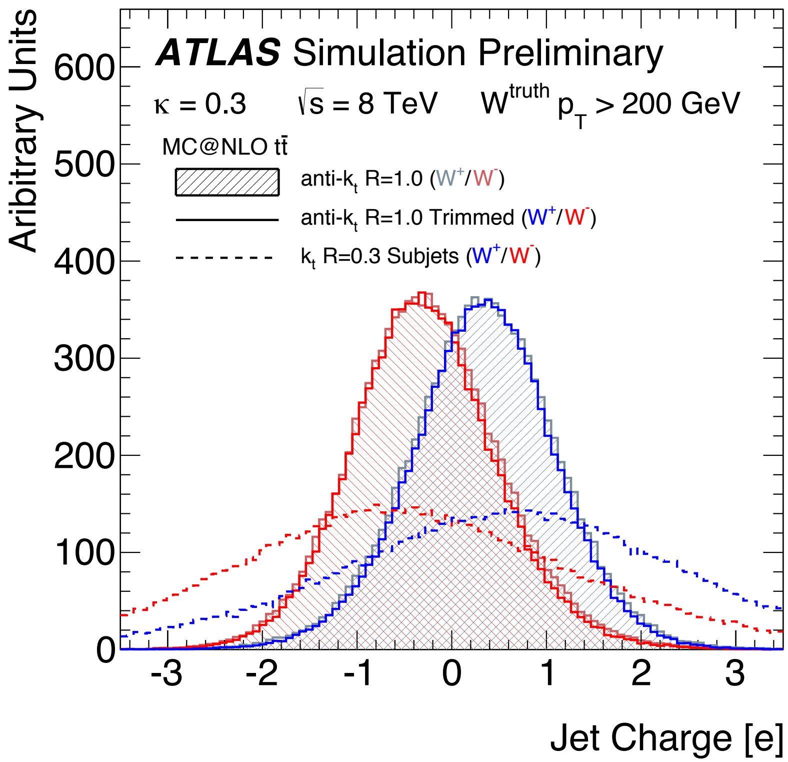


ATLAS-CONF-2013-086

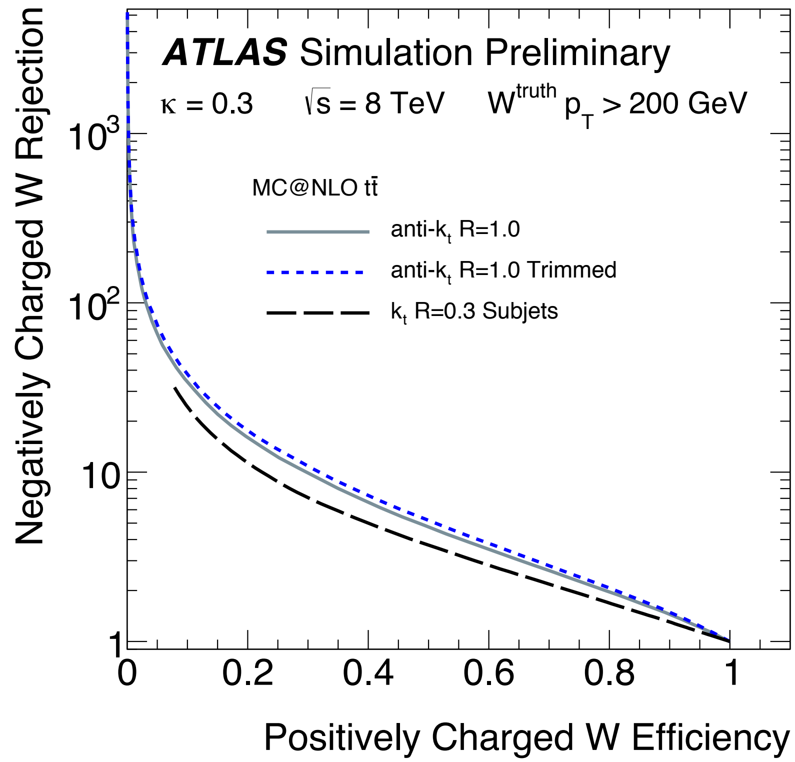


W boson charge tagging (boosted)

W → jet candidate charge



Performance of a W+ tagger



ATLAS-CONF-2013-086

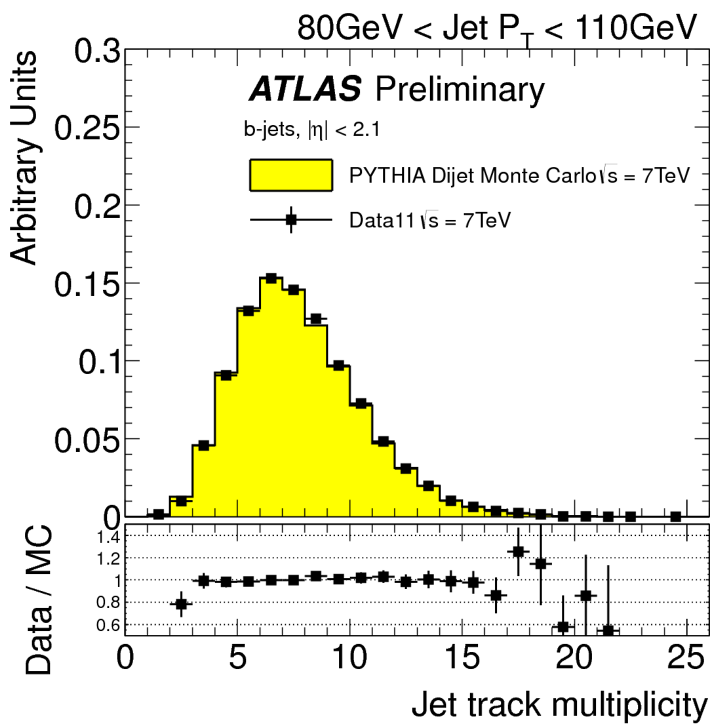
- Restricting to tracks in subjets loses useful information
  - trimming seems to support the pileup rejection inherent in the track selection

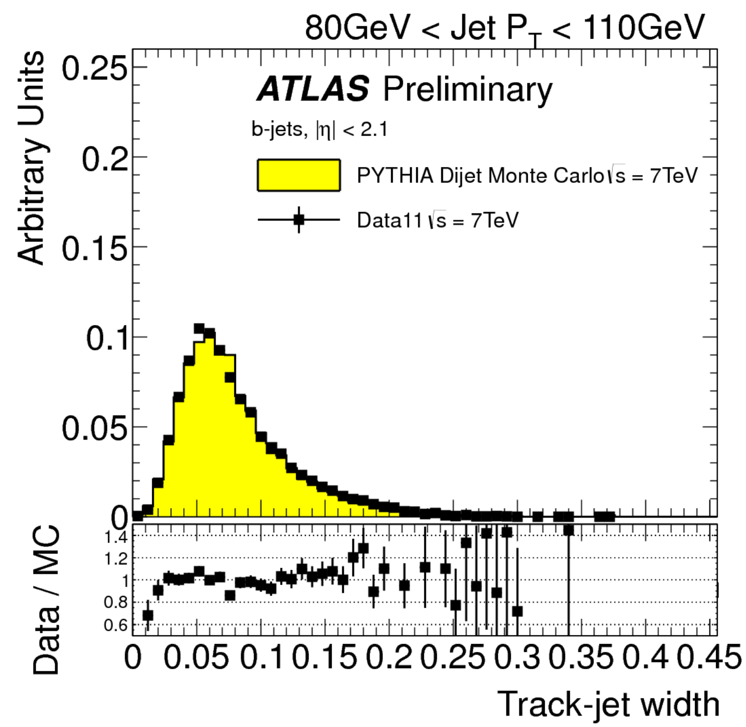
Summary and outlook

- Data comparisons show that these tagger ideas will (still) work: let's use them!
  - prospects for quark-gluon tagging are more pessimistic than Pythia
    - but jet charge study may provide a new handle
- tools shown all emphasize track-based observables
  - this highlights both a need and an opportunity to better understand track distributions in jets
- There's lots to expect even in 2012 data
  - unfolded measurements (jet charge) to compare directly to calculations
  - update (paper) for quark/gluon tagging
  - improved measurements, searches in quarky/heavy flavor/large-charge final states…and possibly, deciphering the nature of a new signal!

ADDITIONAL MATERIAL

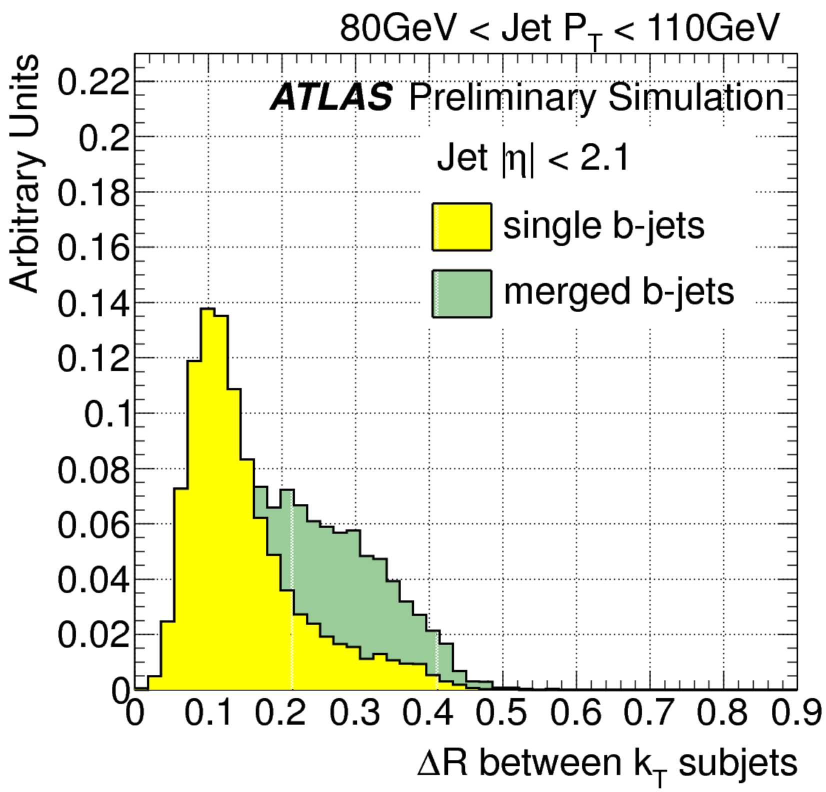
$g \to b\bar{b}$ data validation

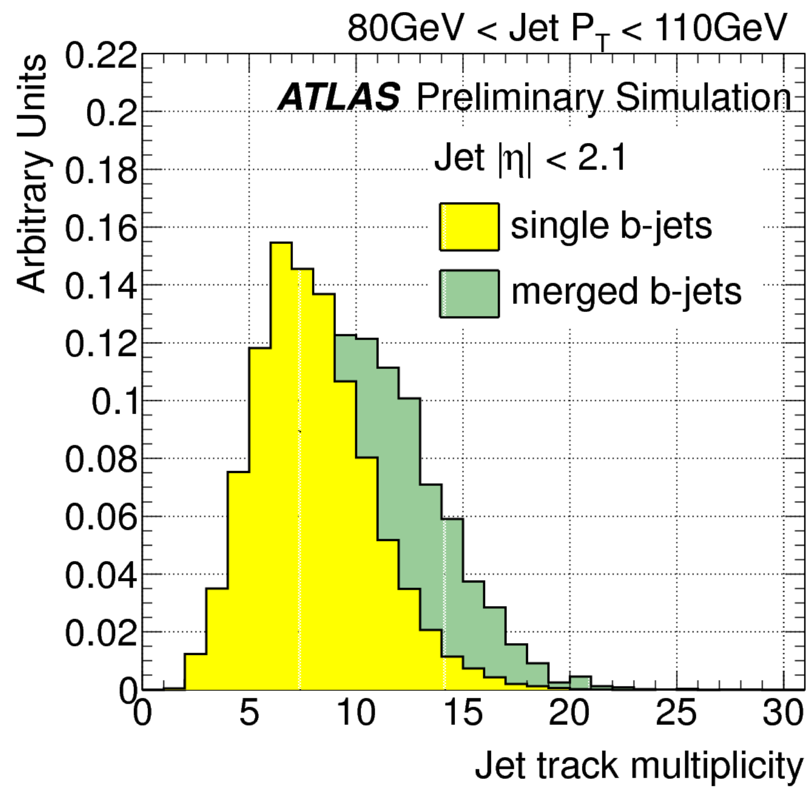




- first check: track-based observables agree for <u>inclusive</u> (single + merged b-jet) distributions

more observables for $g \to b\bar{b}$ tagging

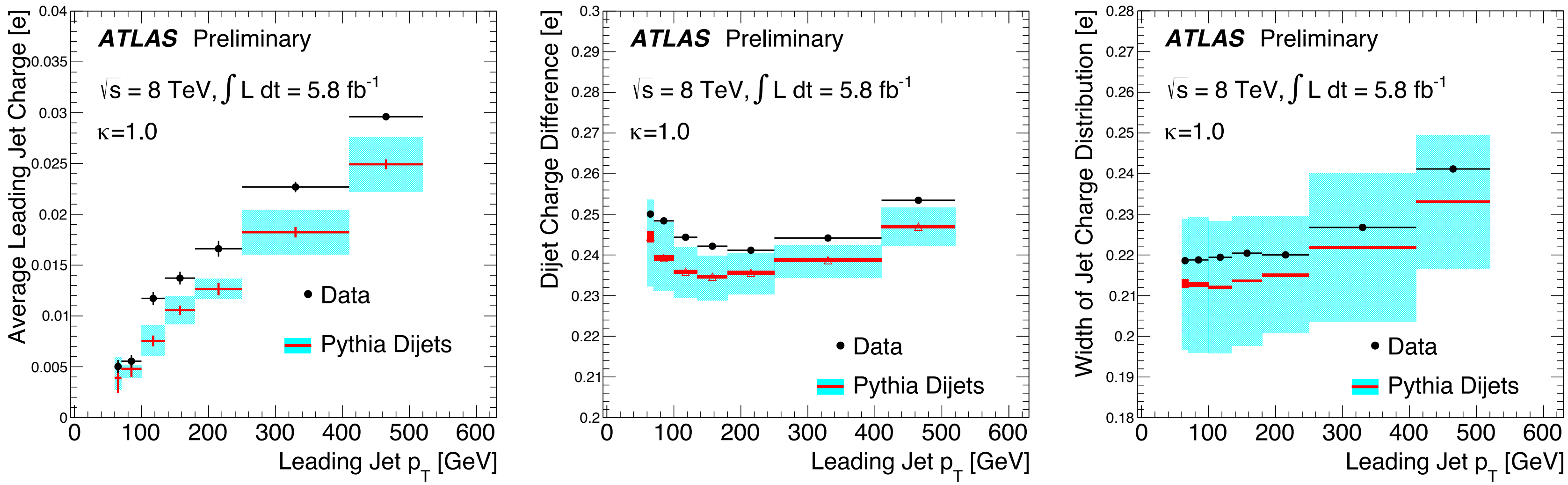




ATLAS-CONF-2012-100

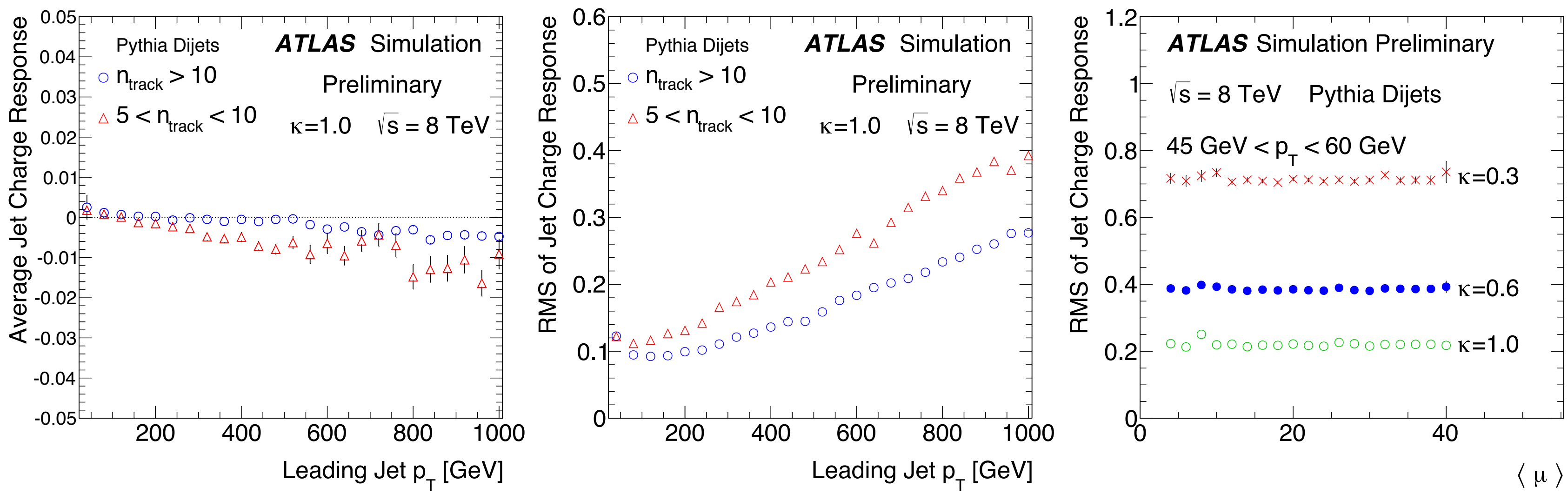
dijet charge observables

- Charge of the leading jet in dijet events is underestimated
- Absolute value of charge difference is also low



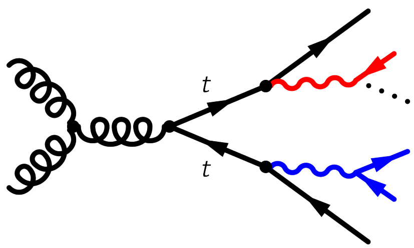
jet charge response

- response defined with respect to stable-particle truth-jets: $Q_{\text{jet}}^{\text{reconstructed}} - Q_{\text{MC jet}}^{\text{stable MC particles}}$

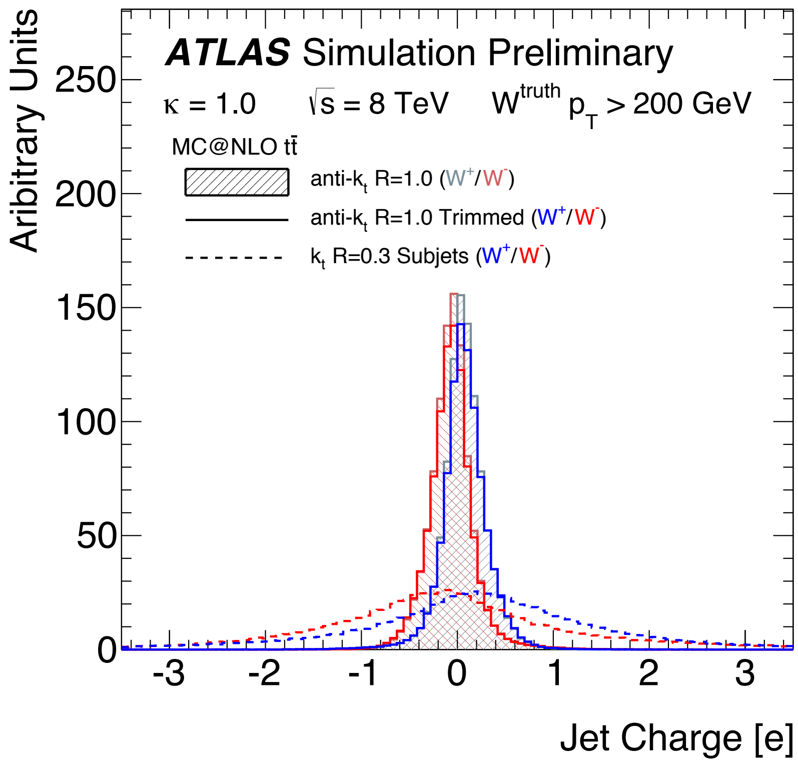


- broadened in hard jets by degraded high p_T track resolution

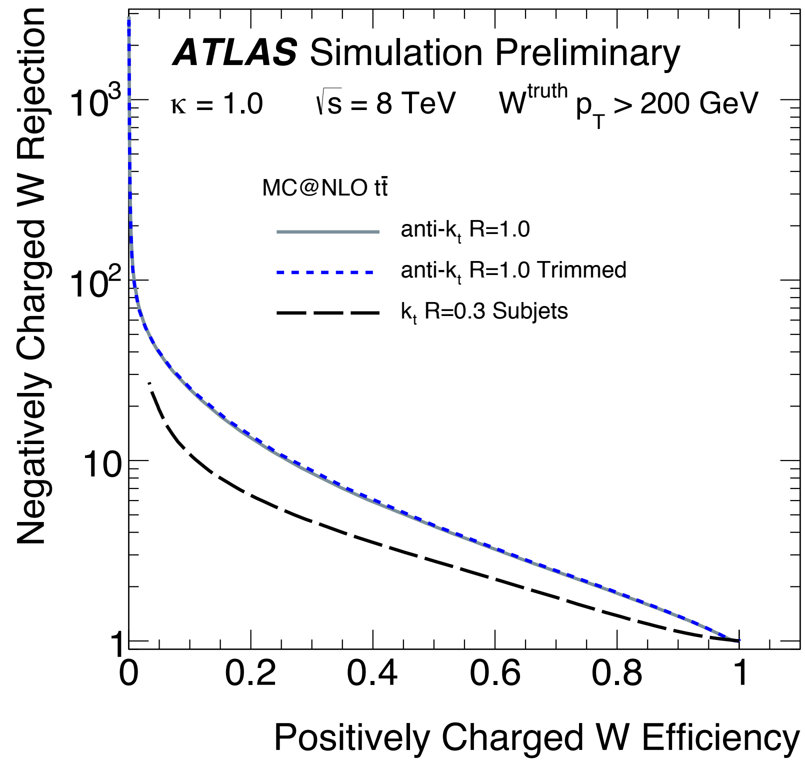
W boson charge tagging (boosted)



W → jet candidate charge



Performance of a W⁺ tagger



- subjet charge sum does even worse with κ=1

ATLAS-CONF-2012-086: *new!*