Jet structure at high energy colliders

Nhan Tran

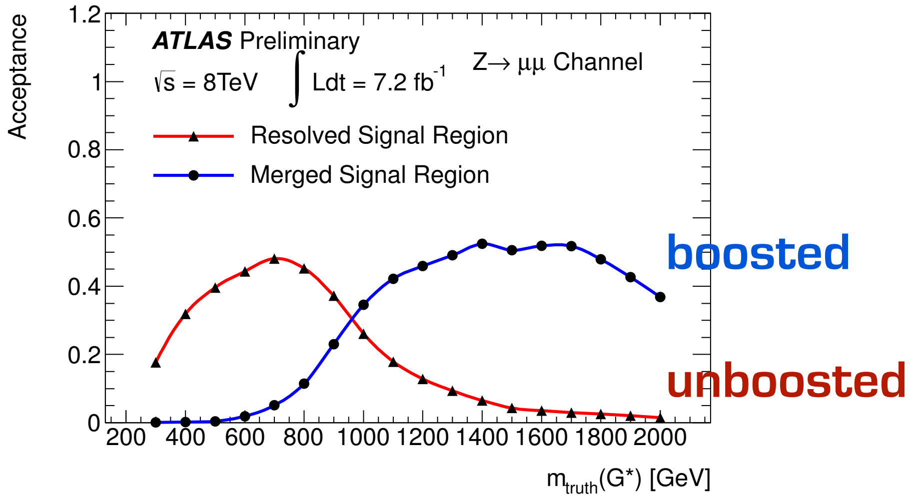
Fermi National Accelerator Laboratory

on behalf of the ATLAS, CMS and H1 collaborations

June 25th, 2013

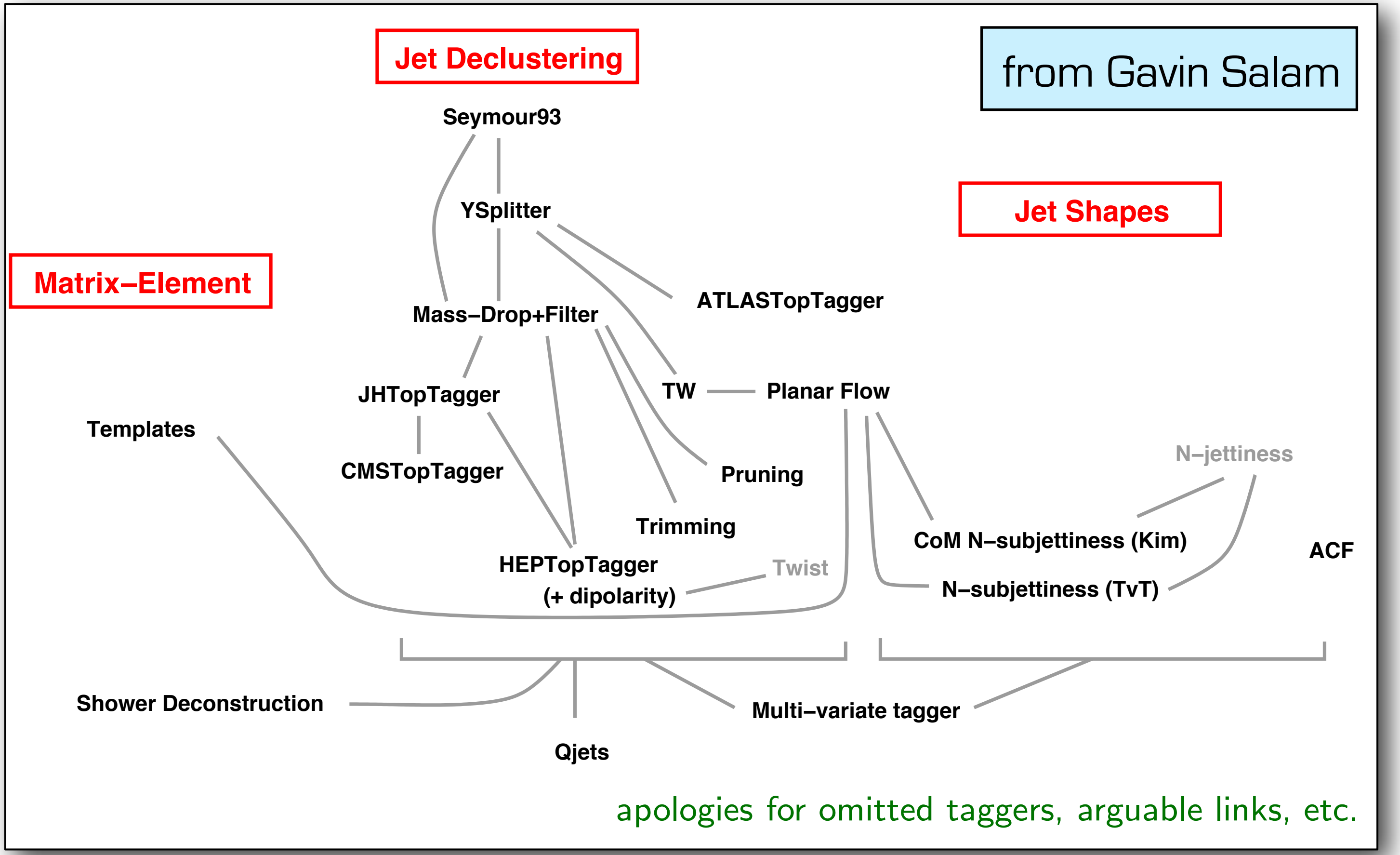
Lepton Photon 2013

- At the LHC, **searches for new physics require jet substructure techniques** for a large range of models and final states
  - Heavy exotic resonances
  - H→bb, heavy Higgs
  - Vector boson scattering
  - anomalous gauge couplings
  - High jet multiplicity SUSY
  - etc...



  - As objects become more boosted, jet structure techniques become a necessity to identify certain new physics signatures
- In addition, jet structure tools can improve performance of jets – i.e. pileup mitigation, distinguishing quark and gluon jets
- Requires a deep understanding of perturbative QCD

In recent years, several techniques and applications proposed by theorists.



Experiments are now adopting these techniques and applying them in many ways.

- Properties of “background” quark and gluon jets in pQCD
  - Parton shower modeling at H1, CMS, ATLAS
  - Jet structure observables at CMS and ATLAS in inclusive dijet and V+jet QCD processes
- Searches employing jet structure observables from CMS and ATLAS
  - Merged W and Z bosons
  - Merged top quarks
  - Boosted non-SM particles

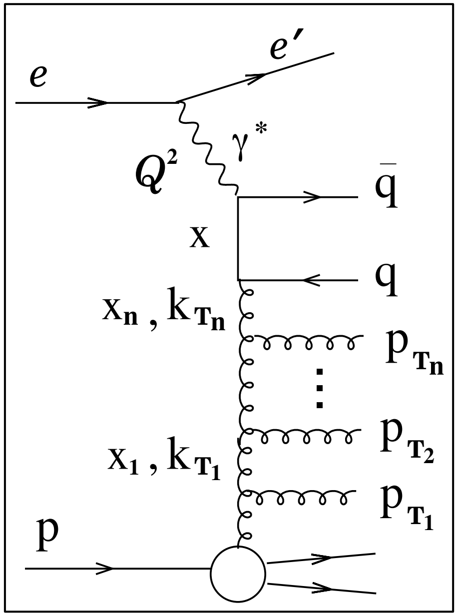
inclusive jet structure studies

parton showering and pQCD
jet mass and grooming
more substructure observables
identifying pileup jets
quark and gluon jet comparisons
g to bbar identification

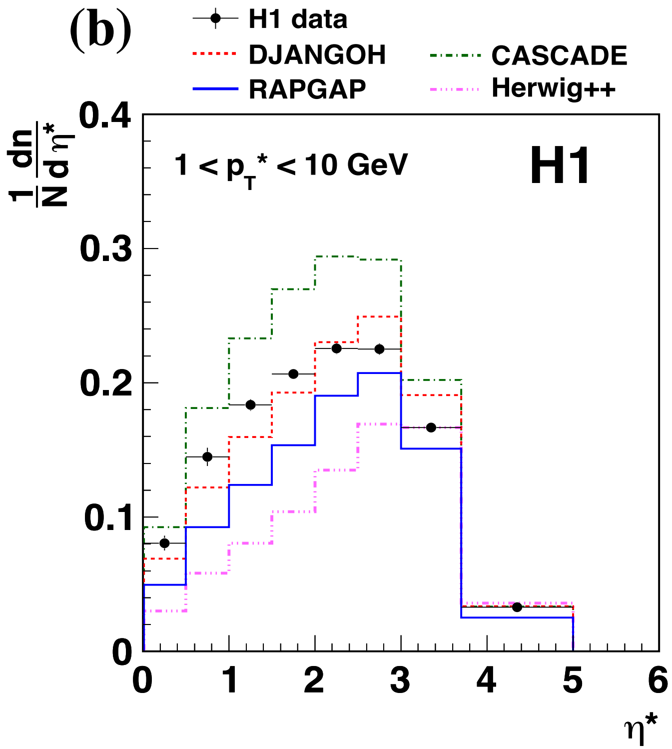
Eur. Phys. J. C73 (2013) 2406

In deep inelastic ep scattering, charged particle pT spectra at low x and high Q^2 a good probe of parton dynamics

Study performed from
$5\ GeV^2 < Q^2 < 100\ GeV^2$
$10^{-4} < x < 10^{-2}$



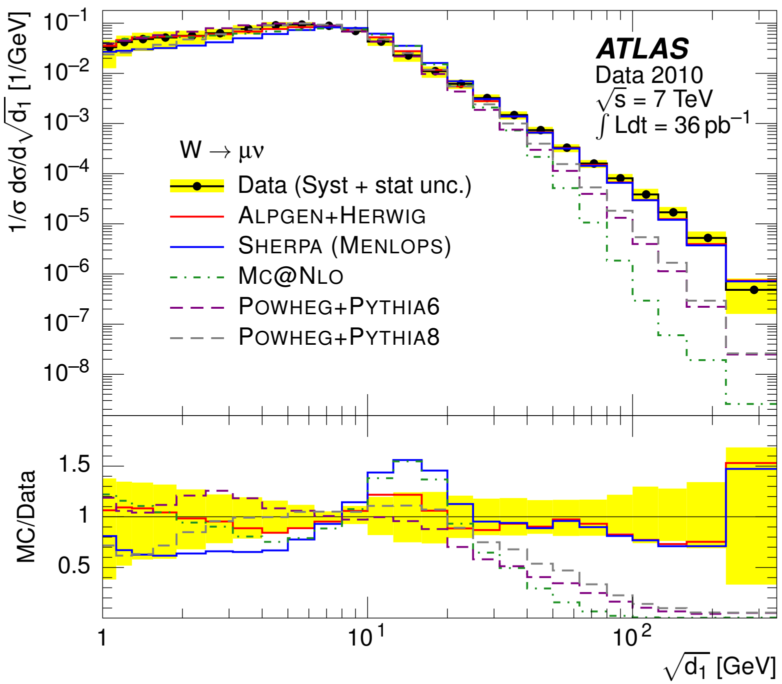
- Tests of parton evolution models: DGLAP, BKFL, CCFM
- Many generators and several tunes studied
- Standard DGLAP, LHC-like (Rapgap, Herwig++) does not do well at high p_T
- Color dipole model in Djangoh gives best agreement over all η and p_T



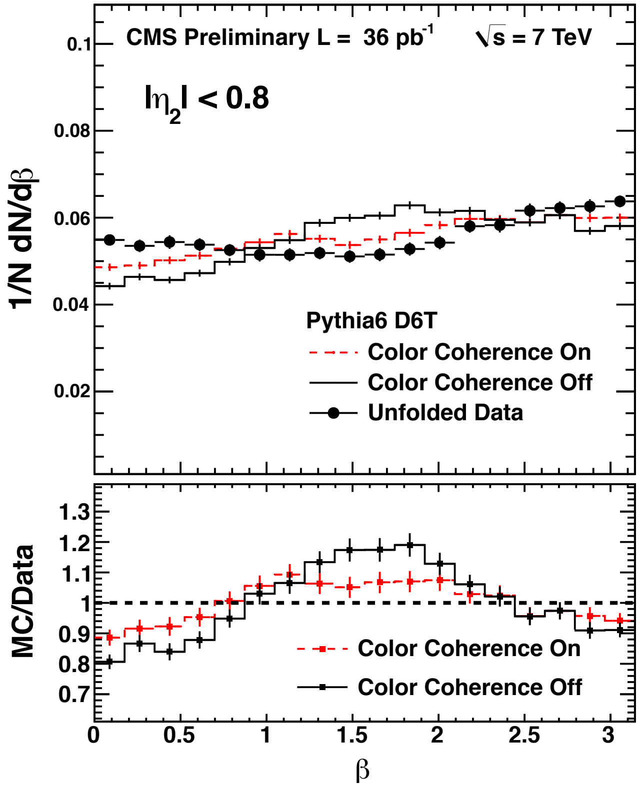
CMS-PAS-SMP-12-010, Eur. Phys. J. C, 73 5 (2013) 2432

Highlighting two recent LHC studies on pQCD

kT clustering algorithm splitting scales in W+jet events compared against many MC models



Tests of color coherence in multijet events, correlations between 2nd and 3rd jet



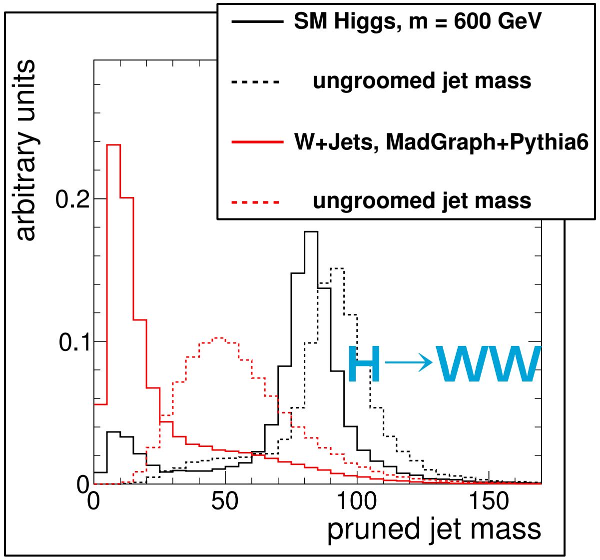
H1, CMS, ATLAS pQCD results fed back into parton shower tunes.

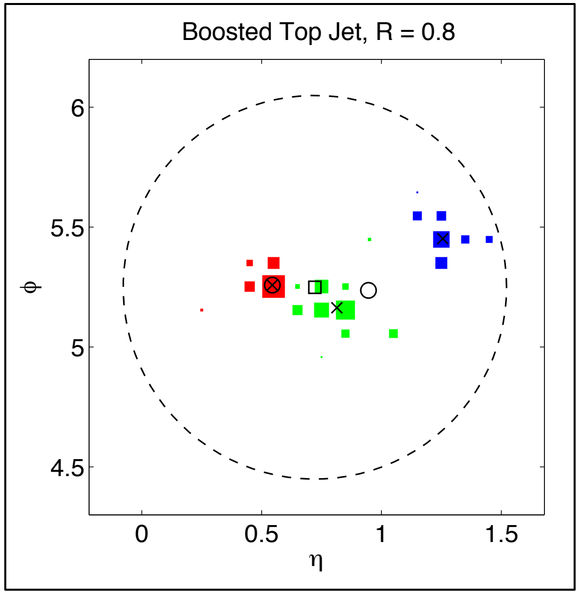
CMS/ATLAS jet structure

- Inclusive jet structure measurements in dijets and V+jets improve understanding of pQCD and backgrounds for searches
  - Gives insight into parton shower modeling
  - Studies performed for large-R jets (R > 0.6) – improved acceptance for new particle searches
- Typically "search" observables examined:
  - Primary observable: jet mass
  - Several additional observables and algorithms separate signal from background

CMS-PAS-HIG-13-008

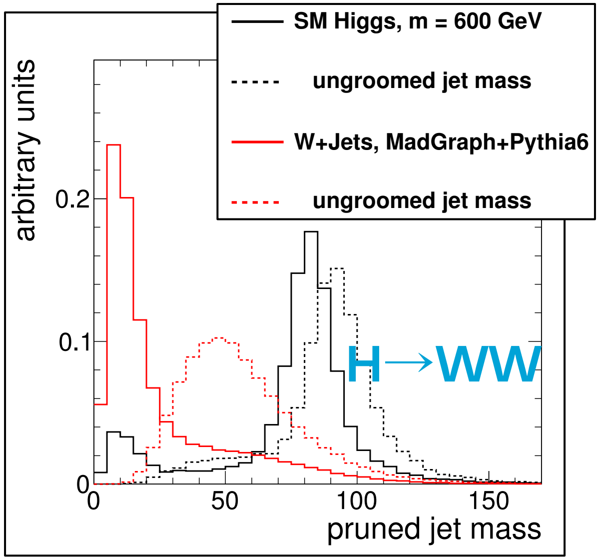
- **Jet mass is most discriminant observable** for heavy objects
  - **Grooming**: a procedure to remove soft radiation and pileup contributions to jet; **used to improve background rejection**
- Additional observables and algorithms provide further background rejection
  - Correlations with jet mass important
  - A few presented today, many explored by experimentalists
    - i.e. N-subjettiness

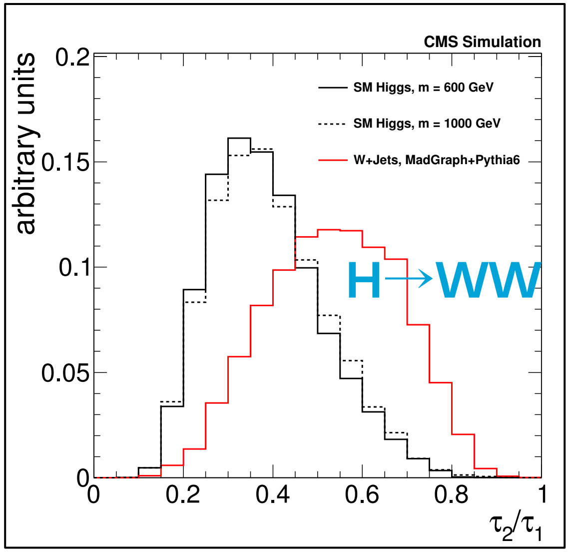




CMS-PAS-HIG-13-008

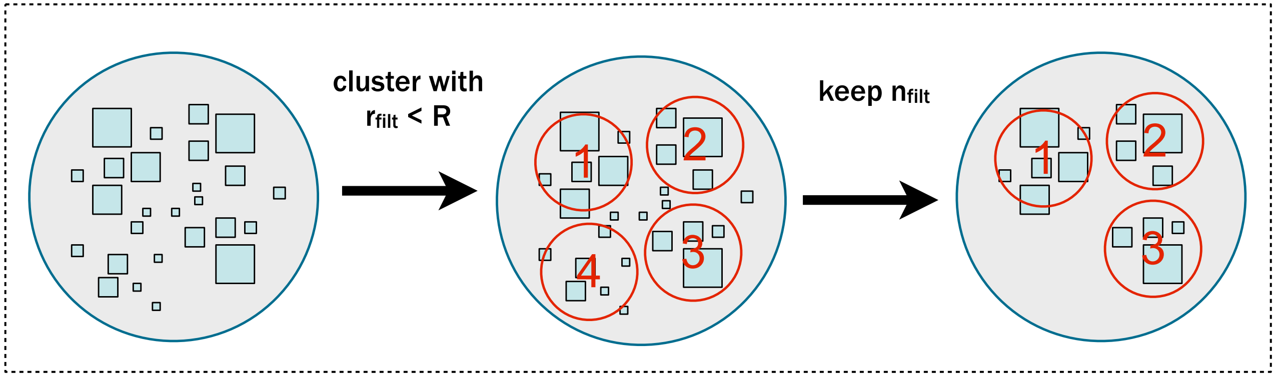
- **Jet mass is most discriminant observable** for heavy objects
  - **Grooming**: a procedure to remove soft radiation and pileup contributions to jet; **used to improve background rejection**
- Additional observables and algorithms provide further background rejection
  - Correlations with jet mass important
  - A few presented today, many explored by experimentalists
    - i.e. N-subjettiness





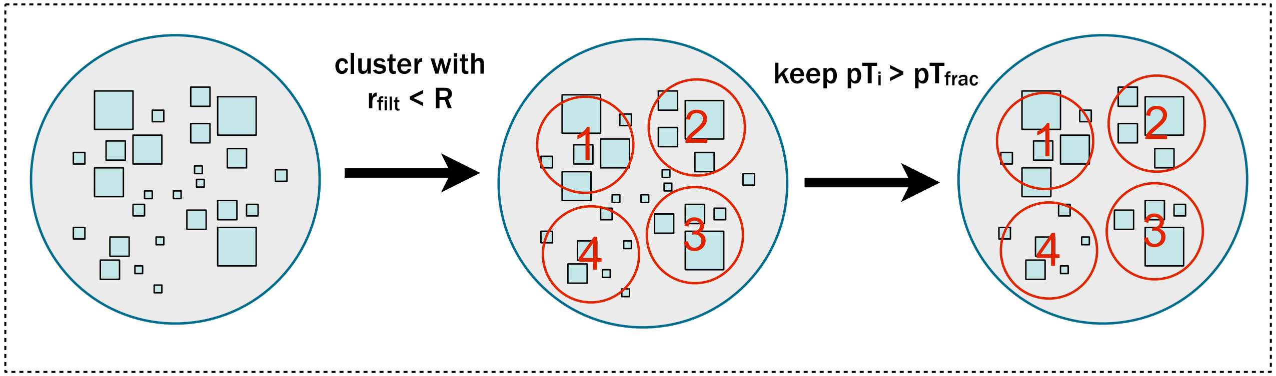
grooming

filtering
[0802.2470]



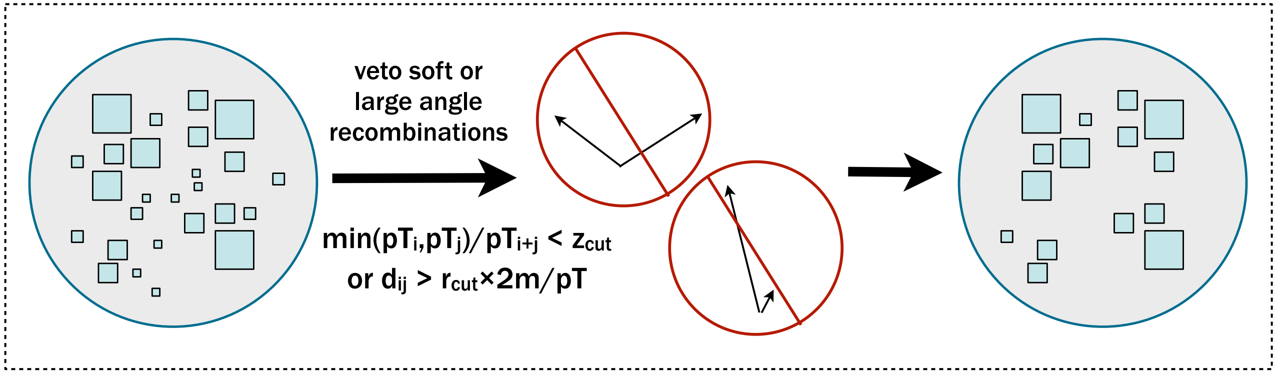
reference values:
$r_{filt} = 0.3$
$n_{filt} = 3$

trimming
[0912.1342]



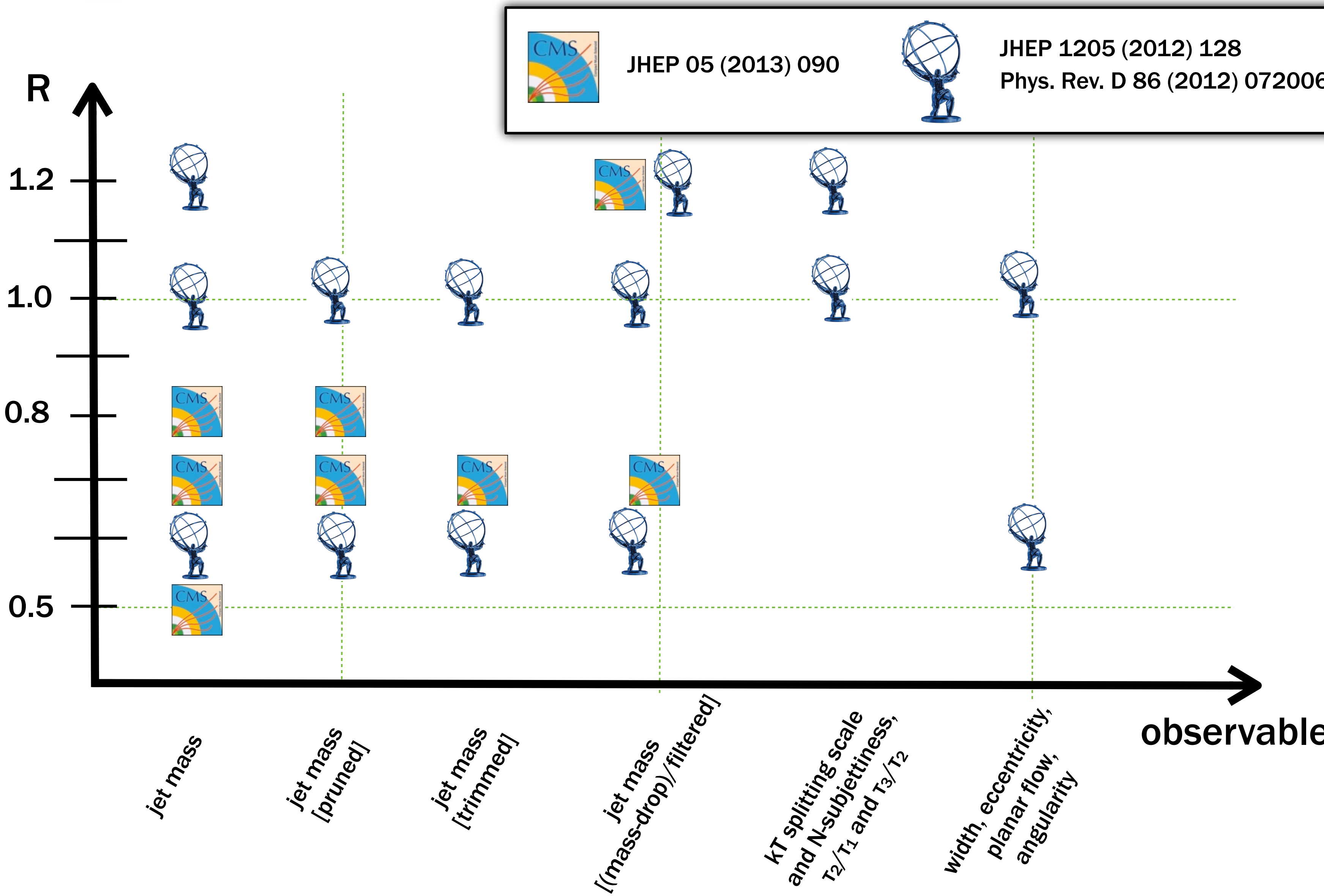
reference values:
$r_{filt} = 0.2$
$pT_{frac} = 0.03$

pruning
[0903.5081]

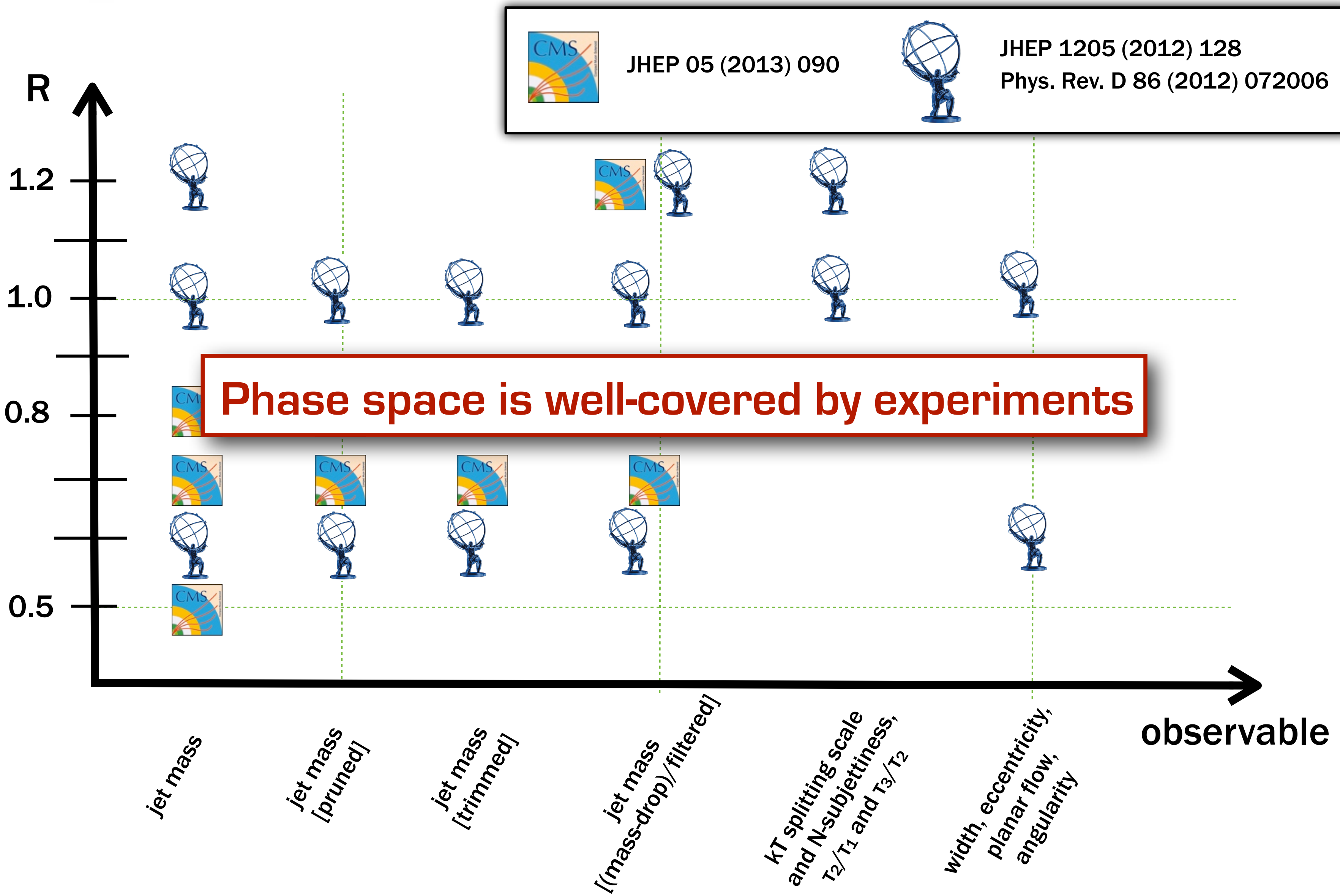


reference values:
$z_{cut} = 0.1$
$r_{cut} = 0.5$

CMS/ATLAS inclusive jet structure



CMS/ATLAS inclusive jet structure

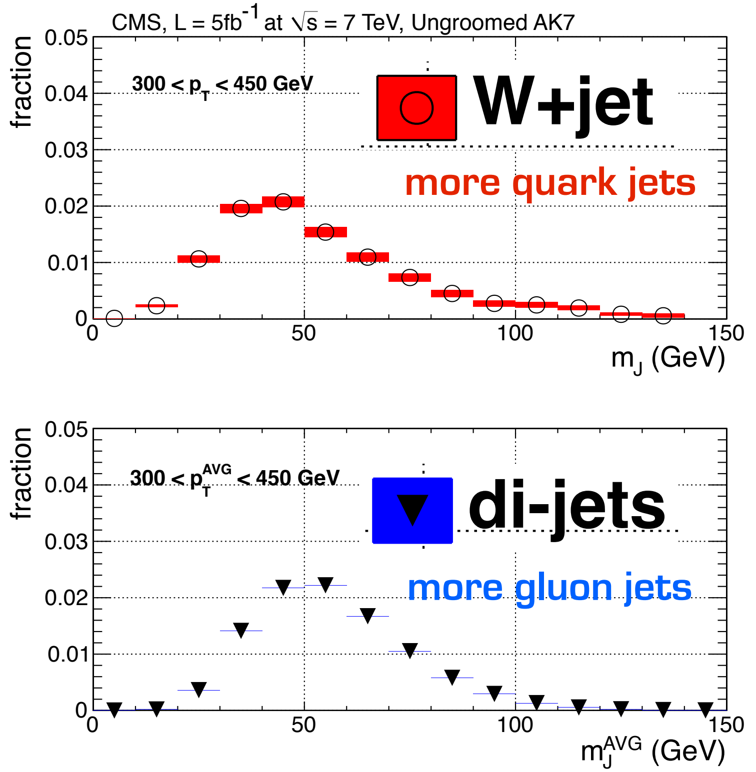


inclusive measurements

JHEP 1205 (2012) 128, Phys. Rev. D 86 (2012) 072006, JHEP 05 (2013) 090

- Analysis of 7 TeV ATLAS/CMS data for dijet and V+jet events
    - Quark and gluon jets have different properties (more later)
- Probe several different jet radii and jet finding algorithms
- Study jet mass and grooming algorithms as well as other jet structure observables
- Provide detector-unfolded jet mass distributions for comparison against simulation

ungroomed jet mass



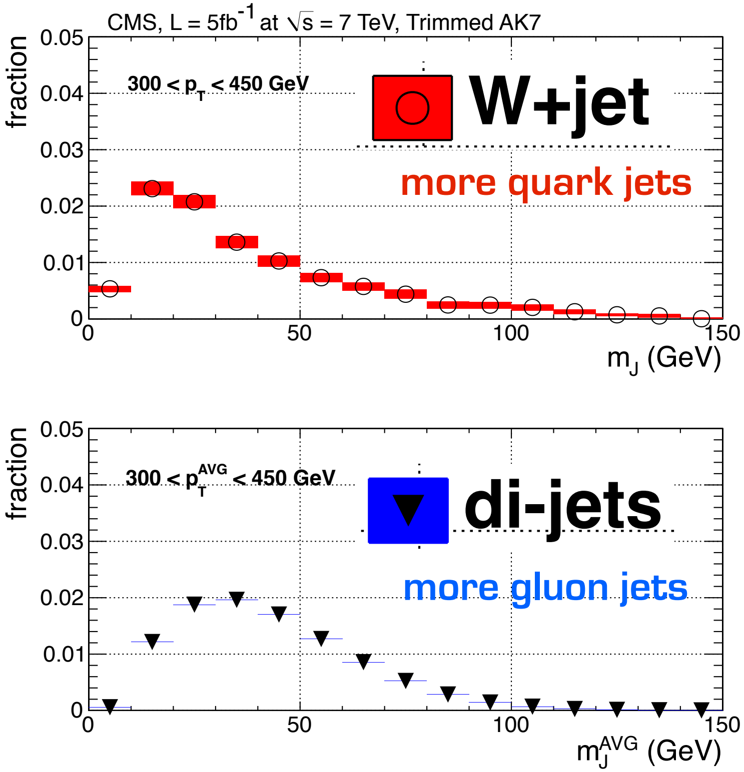
For visualization:
N.B. stat. err. only

inclusive measurements

JHEP 1205 (2012) 128, Phys. Rev. D 86 (2012) 072006, JHEP 05 (2013) 090

- Analysis of 7 TeV ATLAS/CMS data for dijet and V+jet events
  - Quark and gluon jets have different properties (more later)
- Probe several different jet radii and jet finding algorithms
- Study jet mass and grooming algorithms as well as other jet structure observables
- Provide detector-unfolded jet mass distributions for comparison against simulation

trimmed jet mass



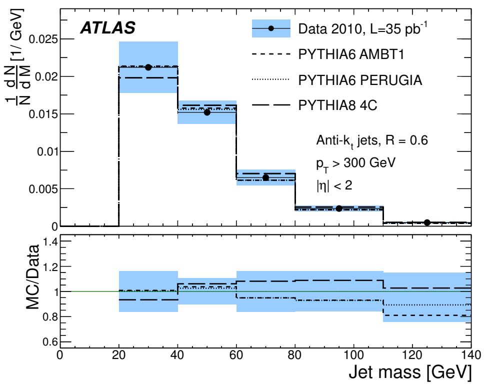
For visualization:
N.B. stat. err. only

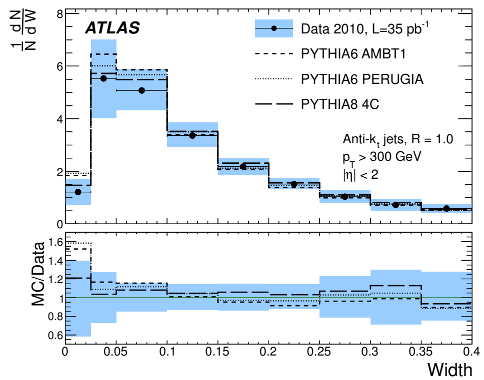
comparison to simulation

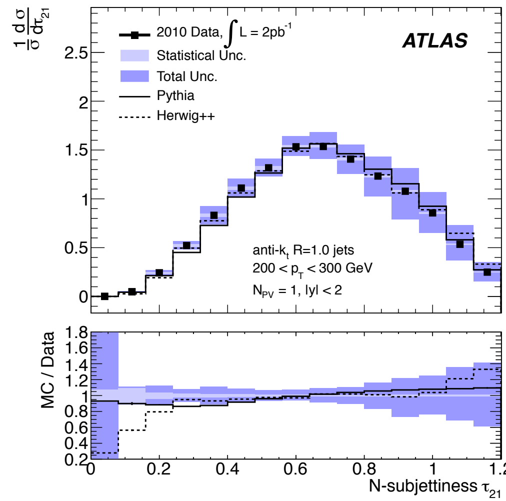
JHEP 1205 (2012) 128, Phys. Rev. D 86 (2012) 072006, JHEP 05 (2013) 090

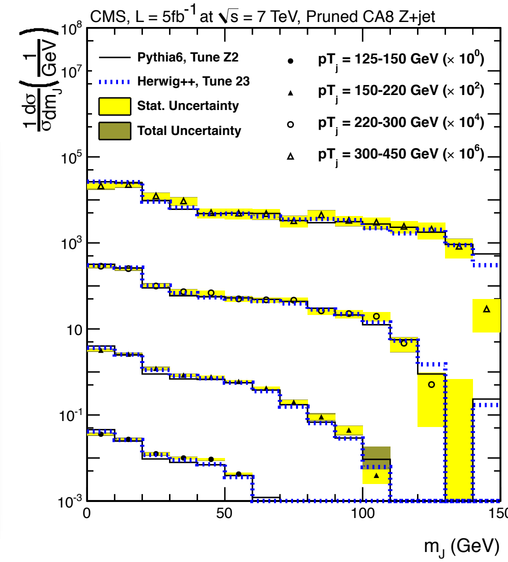
Comparison of unfolded distributions against MC

jet mass, kT splitting scale, N-subjettiness, width, eccentricity, angularity, planar flow examined







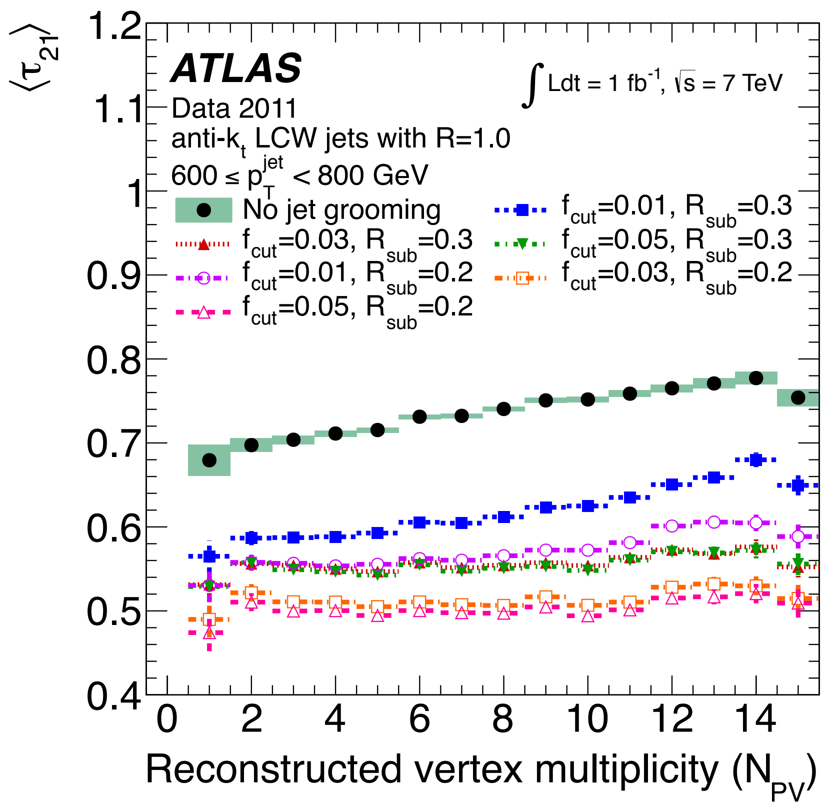


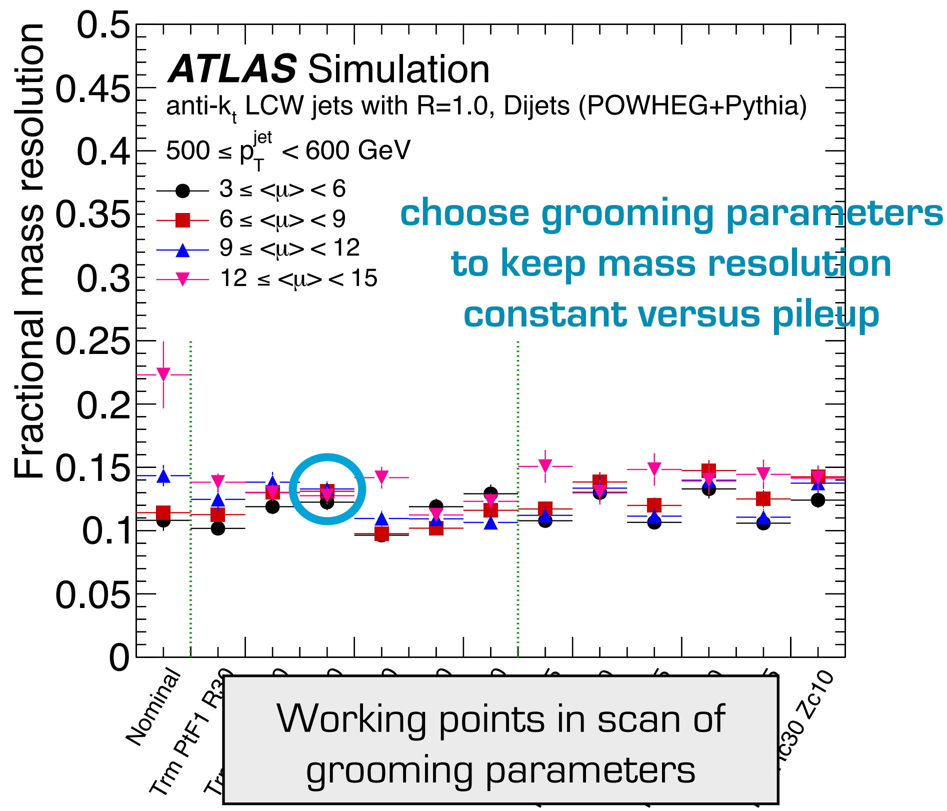
- Both experiments find reasonable agreement for both Herwig and Pythia – not a given!
- Agreement depends on observable and tunes; Herwig++ tends to do a bit better than Pythia for jet mass

arXiv:hep-ex/1306.4945

Groom to reduce pileup effects, especially for large-R jets

Scans in grooming parameter space probe jet structure and pileup characteristics





choose grooming parameters to keep mass resolution constant versus pileup

Working points in scan of grooming parameters

identifying pileup jets

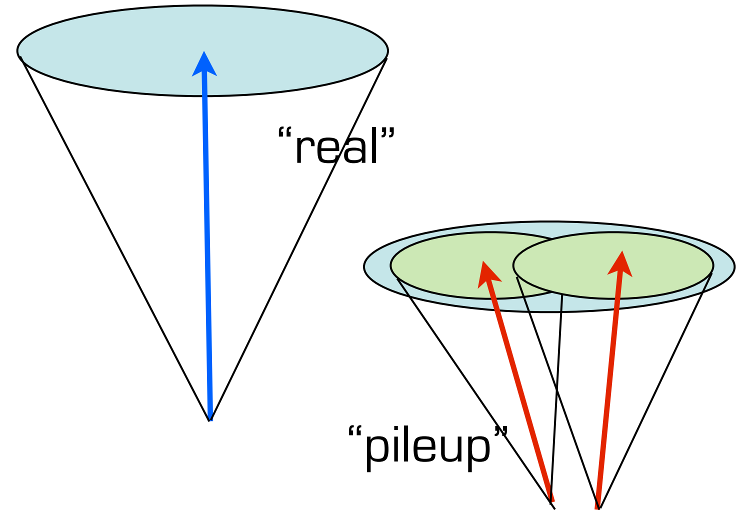
CMS-PAS-HIG-12-043

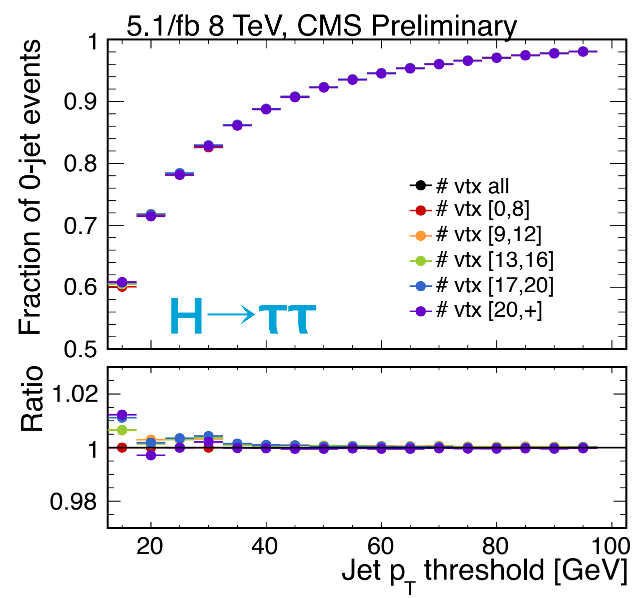
- Pileup jets are several low energy jets from pileup vertices accumulating on each other
- Certain jet substructure variables are found to separate PU jets from real jets
  - Fraction of jet pT from charged tracks coming from primary vertex
    {CMS = β, ATLAS = Jet vertex fraction}
  - **Jet width, shapes**
  - **Charged track multiplicity**

$$\beta = \frac{\sum_{\Delta z(track, v0) < 0.2\mathrm{cm}} p_T^{cand}}{\sum p_T^{cand}}$$

$$RMS = \frac{\sum p_{Ti}^2 \Delta R_i^2}{\sum p_{Ti}^2}$$

- Successfully deployed in Hττ analysis for stabilizing jet vetoes versus pileup

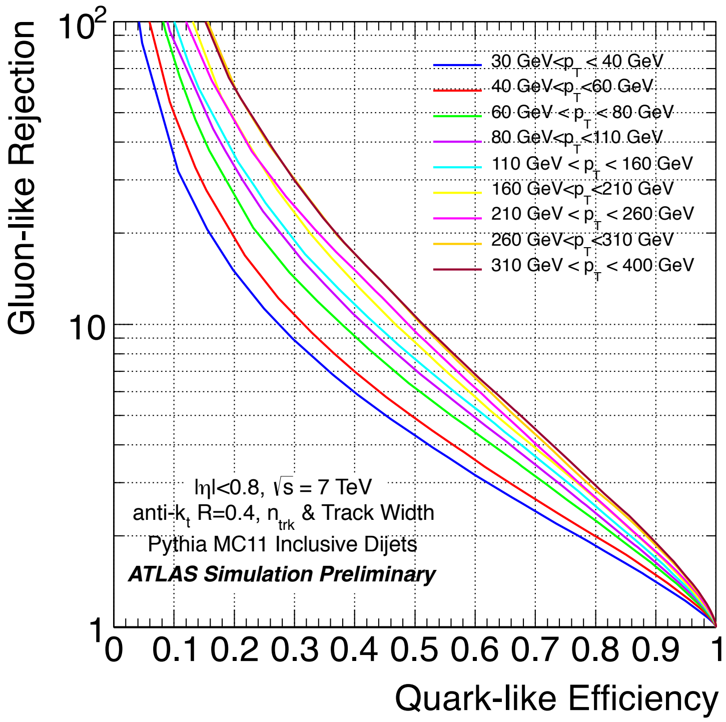




quark and gluon comparisons

ATLAS-CONF-2012-138

- Quark- and gluon-initiated jets have different properties
- Many search applications for distinguishing quarks and gluon jets
  - Hadronically decaying vector bosons
  - monojet, dijet searches
  - SUSY searches with high quark jet multiplicity
- **Jet width and number of charged tracks** provide good discrimination



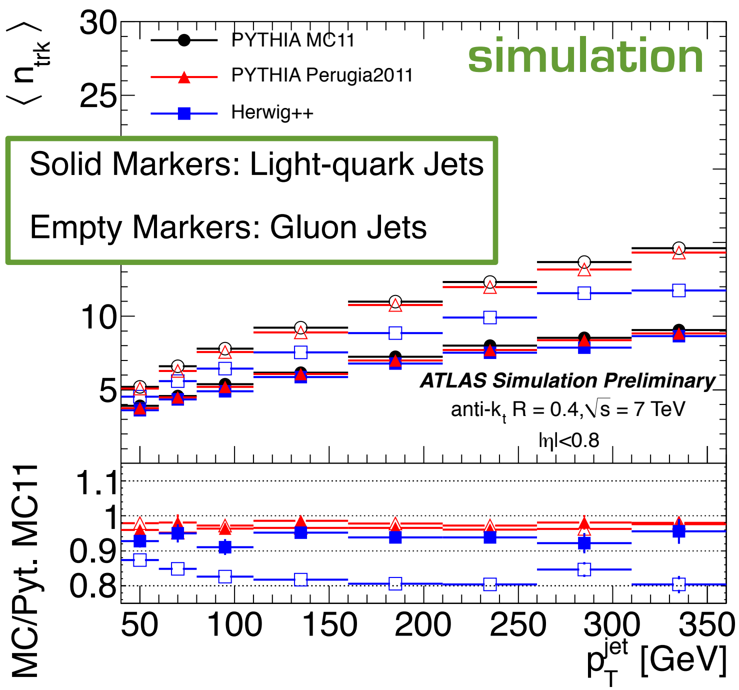
Example: for 50% quark jet efficiency, we can reject 90% gluon jets

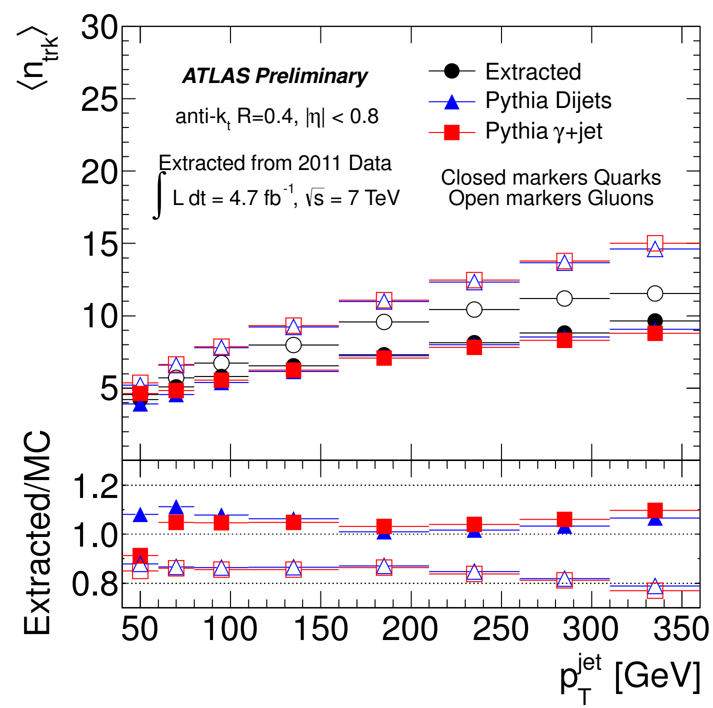
More discriminant at higher pTs

quark and gluon comparisons

ATLAS-CONF-2012-138

- Extract in-situ data distributions of quark and gluons
- Pythia and Herwig++ describe quark jets similarly, but larger difference for gluon jets
- Herwig++ seems to describes gluons better, particularly for n_{trk}



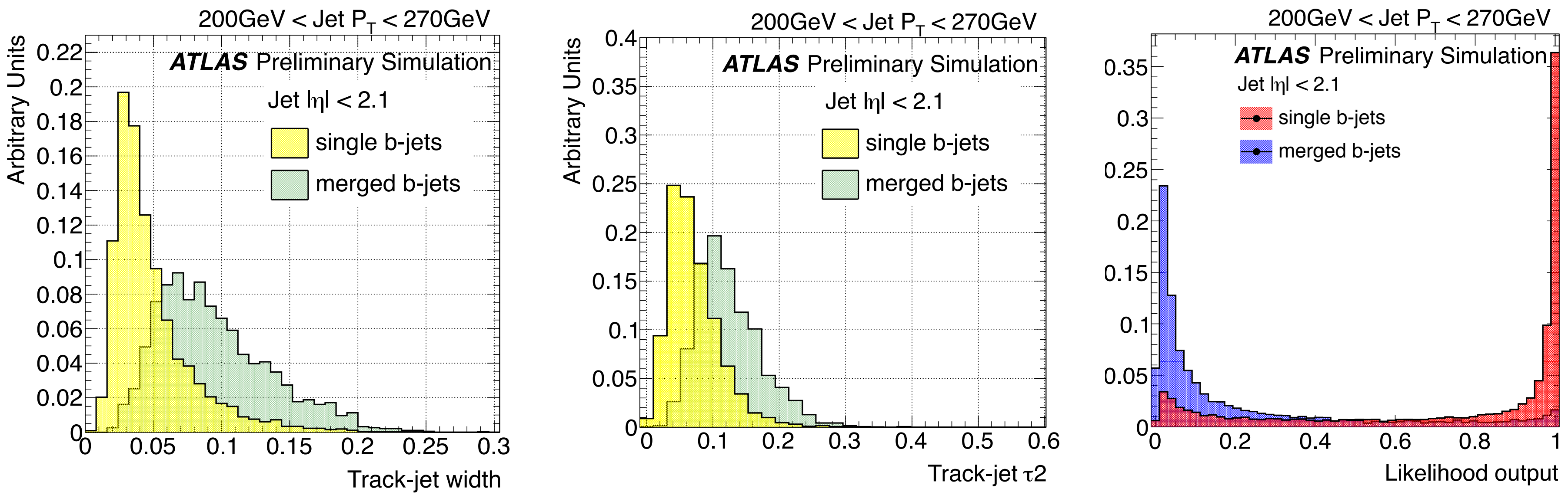


data

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

ATLAS-CONF-2012-100

- **Identify b-jets originating from gluon splitting at small angles**
  - Reduces backgrounds in b-jet searches and estimating efficiency for signatures with double b-tagged jets (i.e. Higgs)
- Identify a b-tagged jet, then use jet structure observables to distinguish between single b-jets and merged b-jets
  - Jet width, N_{trk}, $\Delta R_{\text{(max, tracks)}}$, τ_2
  - Very good discrimination power, better at high pTs



searches with jet structure

merged vector bosons
merged top quarks
non-SM boosted objects

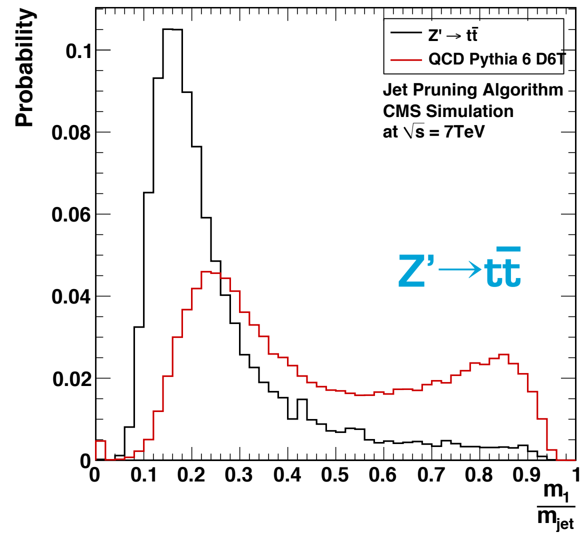
An emphasis will be placed on techniques instead of limits.

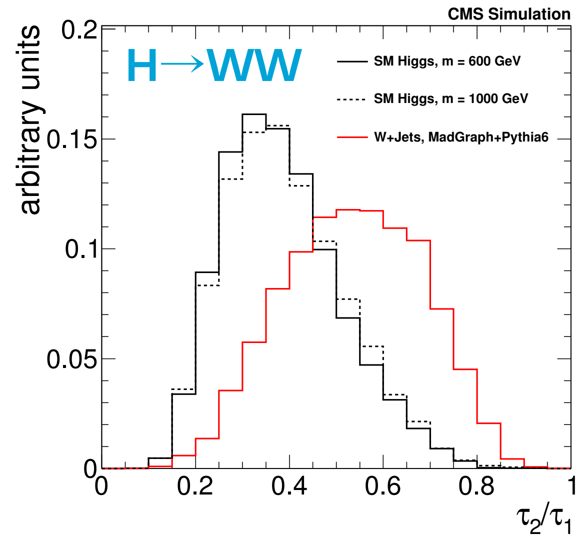
identifying boosted vector bosons

CMS-PAS-JME-09-001, CMS-PAS-HIG-13-008

Techniques are developing rapidly, different analyses using varying selections

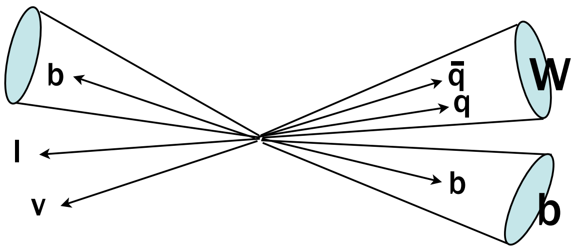
- In all cases, main discriminating variable is jet mass, sometimes cut on additional variables
  - Groomed jet mass improves background rejection in searches
- Jet-finding with a standard or large radius algorithm depending on search phase space
- Cut on the mass drop, $\mu = m_1 / m$
  - m_1 is mass of highest mass subjet
  - subjets defined by un-clustering last step
- Cut on N-subjettiness, τ_2 / τ_1

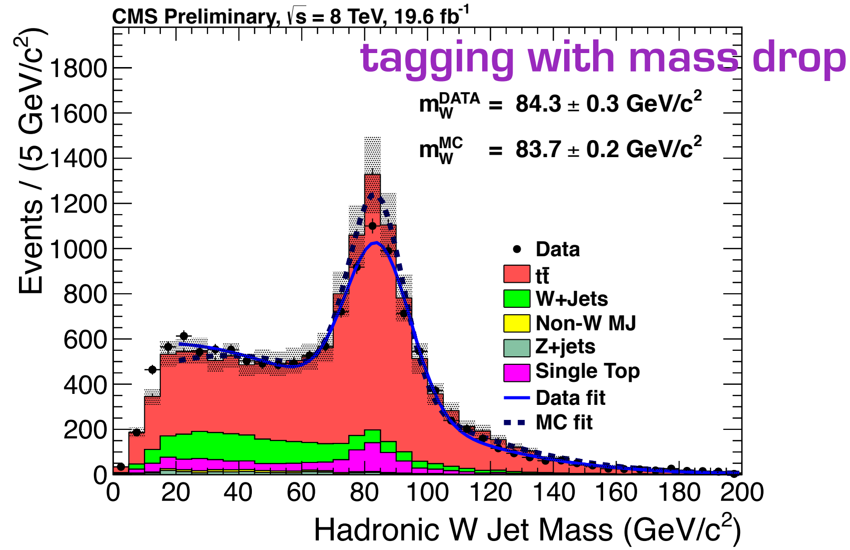




validating merged W bosons

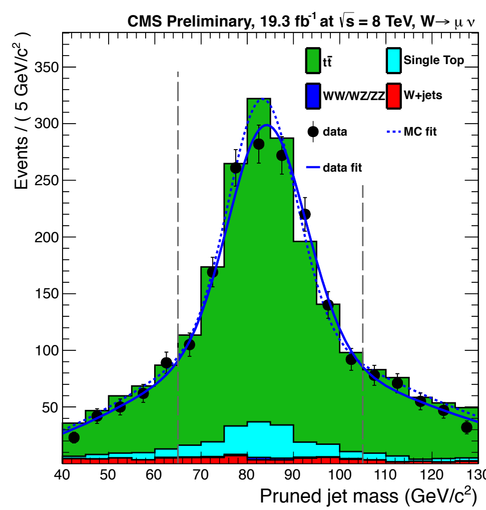
CMS-PAS-B2G-13-005, CMS-PAS-HIG-13-008

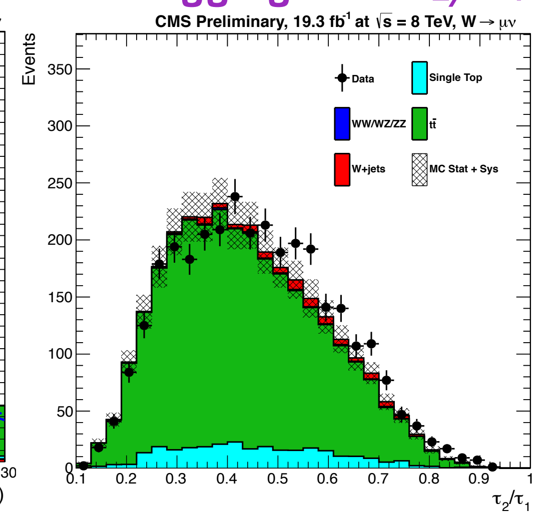




- Validation using merged W bosons in semi-leptonic tt sample
- Clear observation of merged W's
- Used as a valuable sample for understanding jet mass scale and resolution
- Statistics are becoming sufficient for estimating data-driven efficiencies

tagging with τ_2/τ_1

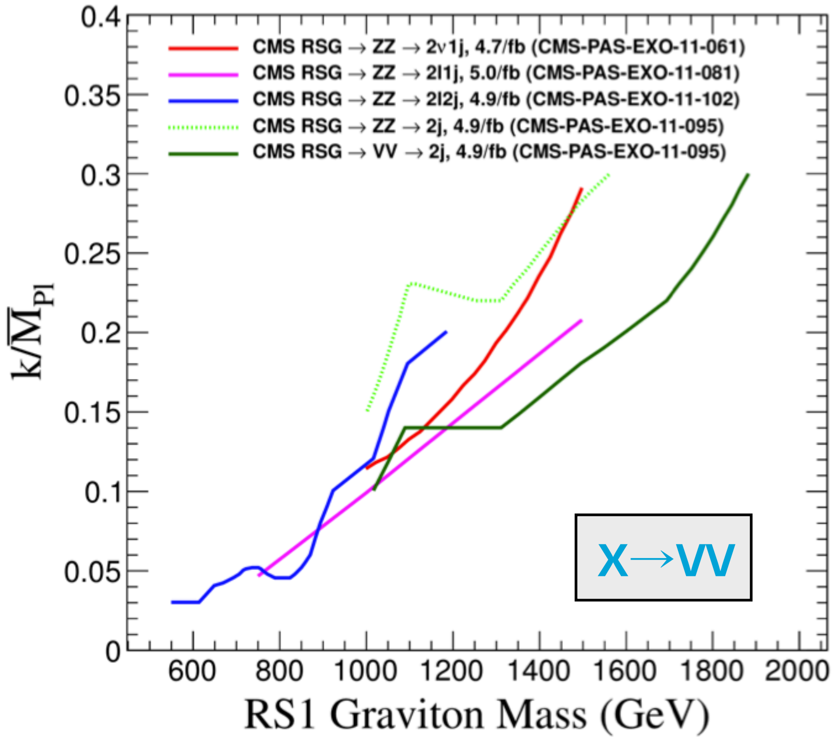




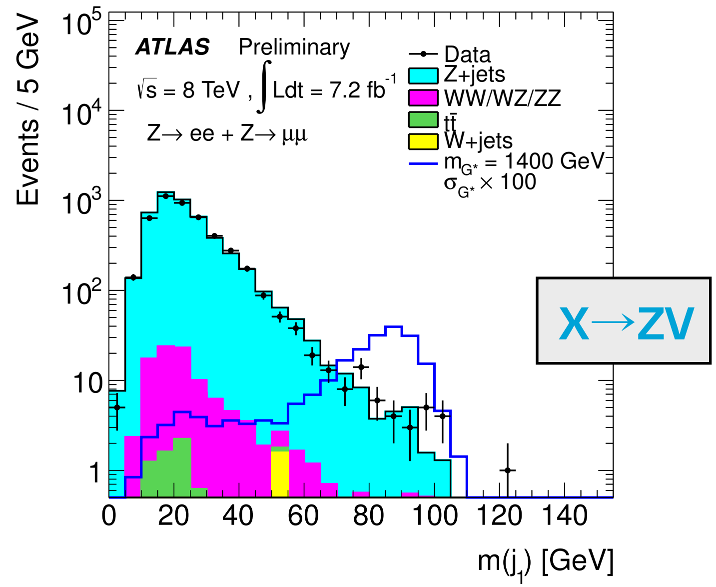
searches with boosted vector bosons

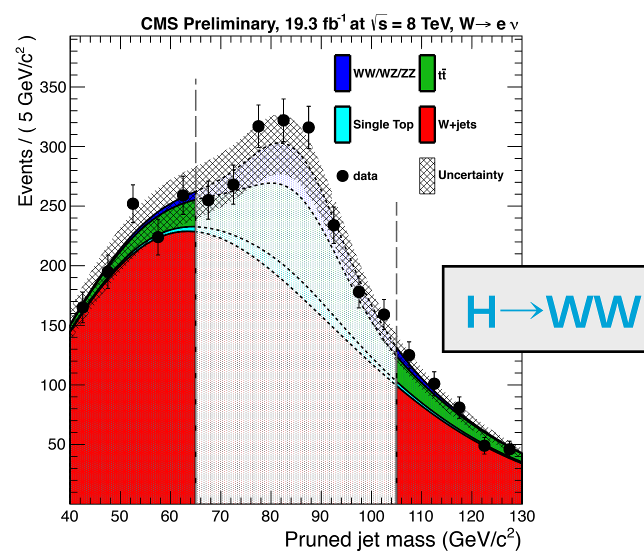
ATLAS-CONF-2012-150, CMS-PAS-HIG-13-008

Combination of CMS searches
for RS1 Graviton in di-boson final states

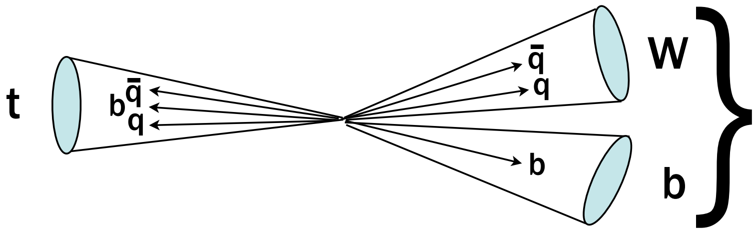


4 of 5 analyses using jet
substructure observables

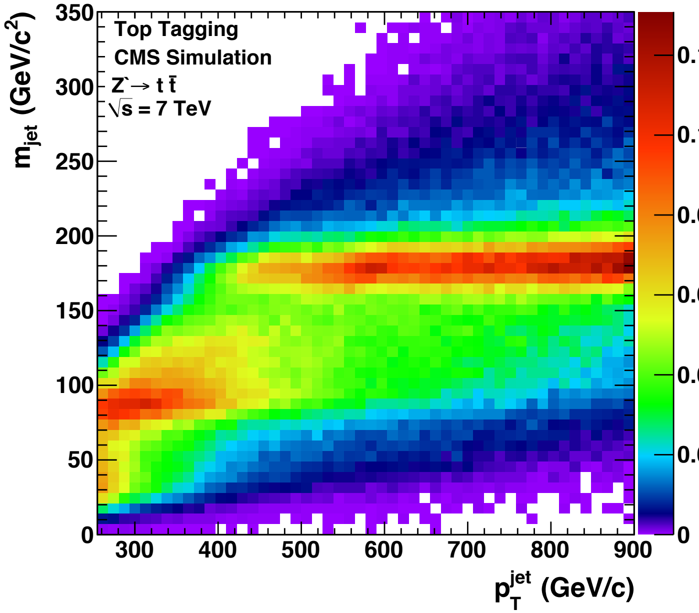




CMS-PAS-JME-09-001, JHEP 1010:078 (2010)



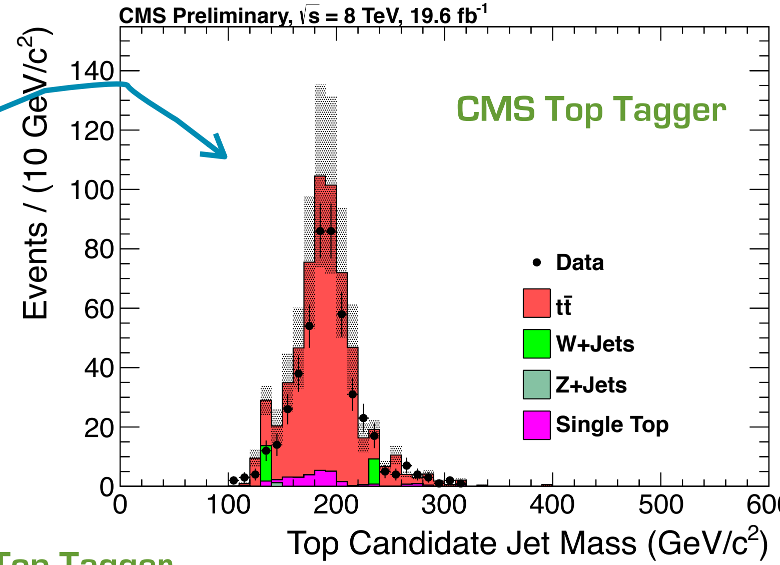
- Many algorithms for identifying top quarks
- Currently implemented in public analyses
  - **Tagging using kT-splitting scale**: require the last kT clustering step to be hard
  - **HEP Top Tagger**: decluster jet into subjets, apply kinematic constraints on all mass pairings: $\{m_{12}, m_{23}, m_{13}\}$
  - **Template Top Tagger**: test compatibility of jet with O(300k) top decay templates
  - **CMS Top Tagger**: decluster jet into subjets, apply kinematic constraints on subjets
- Not the end of the story, existing algorithms to be further optimized and more to be tested

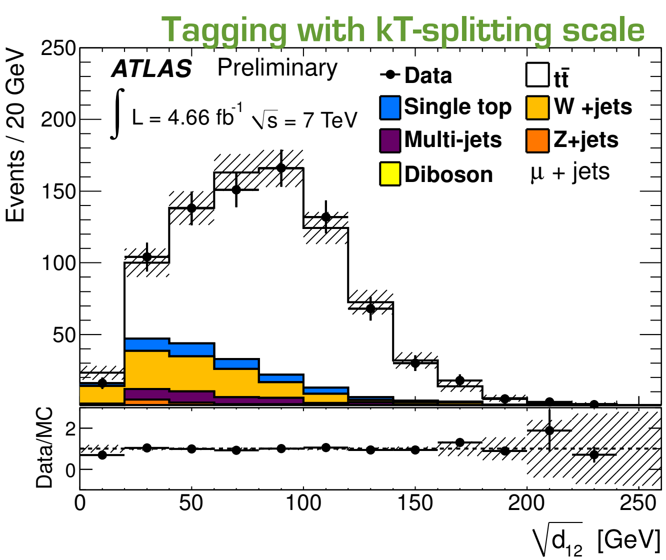


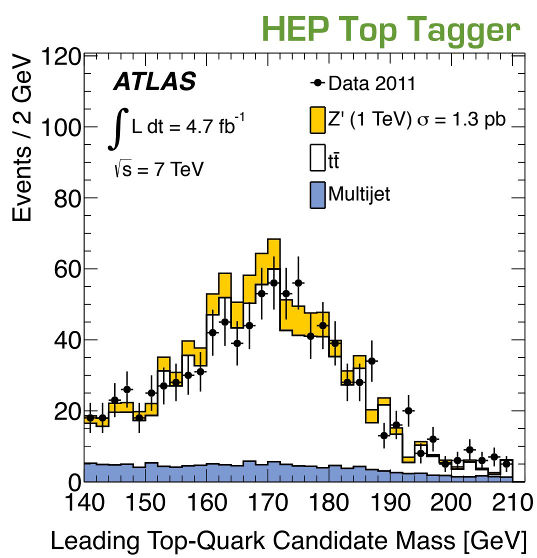
validation of boosted tops

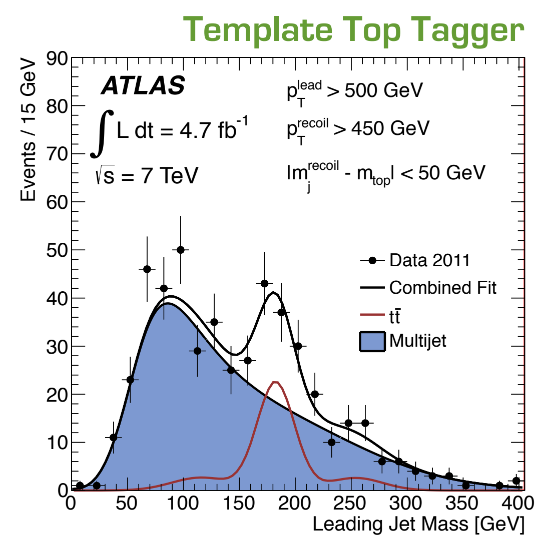
Phys. Lett. B 718 (2013) 1284-1302, ATLAS-CONF-2012-136, CMS-PAS-B2G-13-005

- **With 8 TeV data, enough semi-leptonic tt events to validate boosted tops at high pT**
- Search signal regions validation show good agreement for data vs. simulation
  - B-tagging in boosted environments helps to further reduce multijet backgrounds







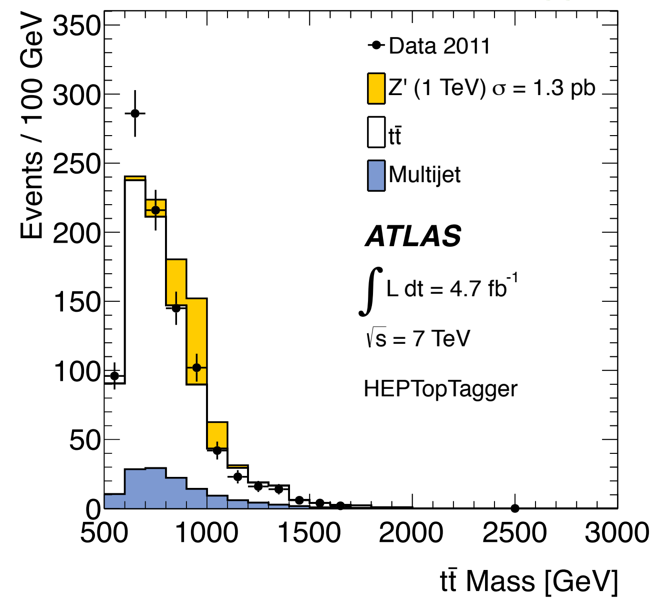


searches with boosted tops

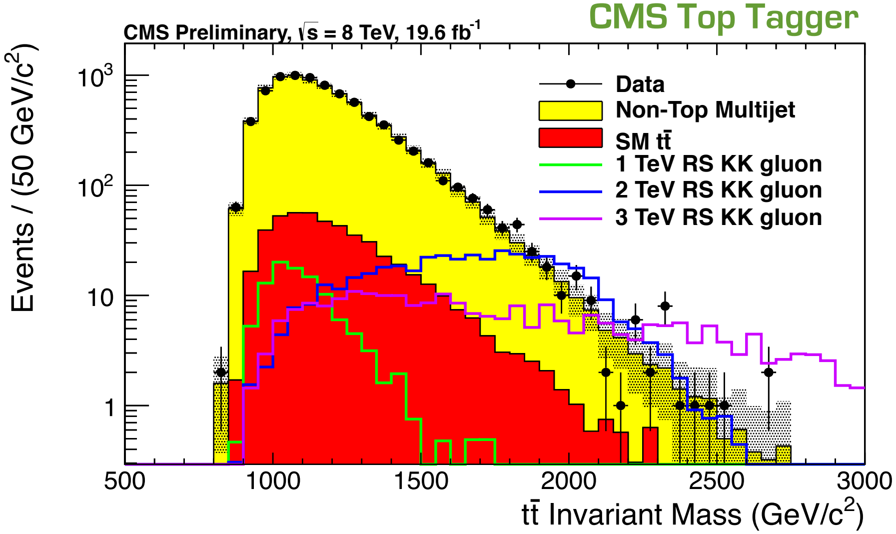
Phys. Lett. B 718 (2013)1284-1302, ATLAS-CONF-2012-136, CMS-PAS-B2G-13-005

- Boosted tops currently in searches for tt resonances both in all-hadronic and semi-leptonic channel
- Many additional applications:
  3rd generation final states (W’, b’, etc.)
  Moderately boosted tops in SUSY stop searches

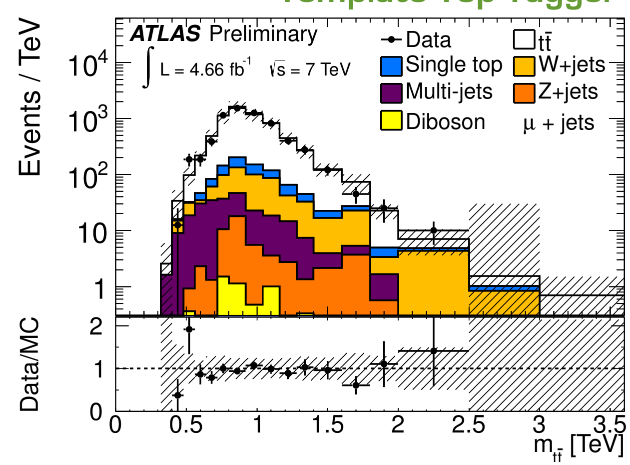
HEP Top Tagger



CMS Top Tagger



Template Top Tagger

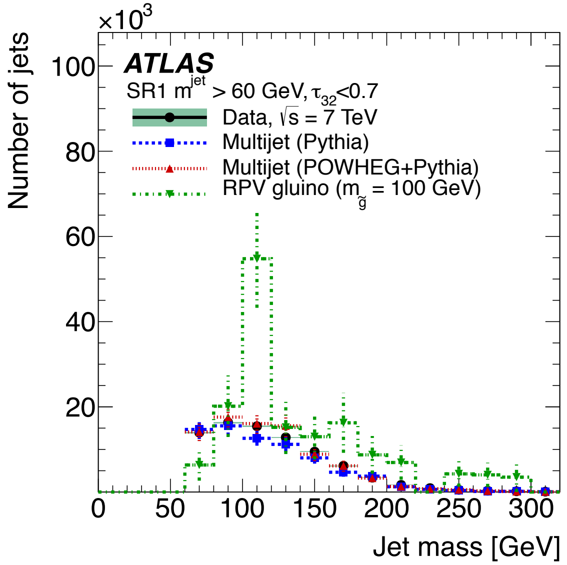


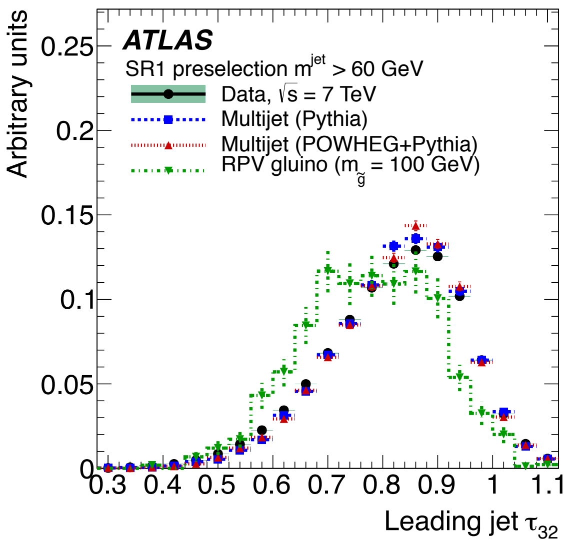
boosted non-SM objets: RPV gluinos

CERN-PH-EP-2012-281

- First example: exotic resonance search: $\tilde{g} \rightarrow \tilde{q}q \rightarrow qqq$, pair produced RPV gluinos
- For light gluinos, decaying quarks can be highly collimated
- Use large radius jets with high pT
- N-subjettness, τ_3/τ_2, variable used to identify jets with 3 subjets

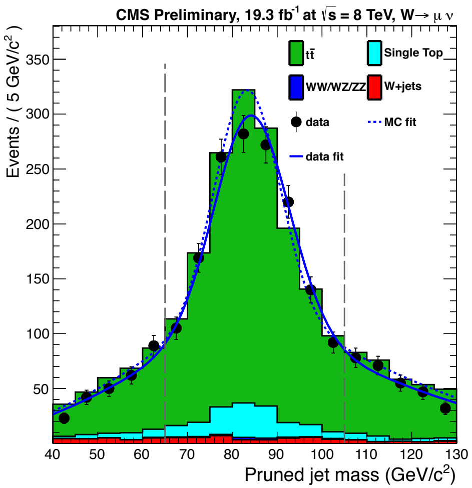
- New phenomenological ideas for low MET SUSY with high jet multiplicities using jet structure





summary

- **Large amount of experimental results studying jet structure**
  - probes of pQCD
  - inclusive jet structure measurements
  - searches with boosted objects in wide range of physics models
- Many new experimental results expected at the BOOST 2013 conference in August



- The 7 TeV and 8 TeV LHC data already proven to be an excellent dataset for jet structure studies;
  Relevance will only increase for the upcoming 13 TeV run!

summary

- **Large amount of experimental results studying jet structure**
  - probes of pQCD
  - inclusive jet structure measurements
  - searches with boosted objects in wide range of physics models
- Many new experimental results expected at the BOOST 2013 conference in August

- The 7 TeV and 8 TeV LHC data already proven to be an excellent dataset for jet structure studies;
  Relevance will only increase for the upcoming 13 TeV run!

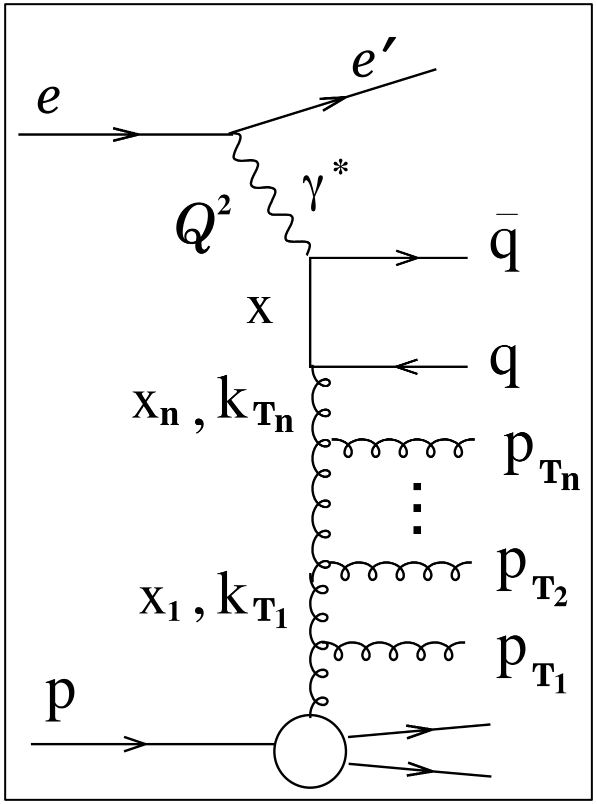
additional material

charged particle production at H1

Eur. Phys. J. C73 (2013) 2406

Study performed from
5 GeV^2 < Q^2 < 100 GeV^2
$10^{-4} < x < 10^{-2}$

- In deep inelastic ep scattering, charged particle pT spectra at low x and high Q^2 a good probe of parton dynamics
- Tests of parton evolution models: DGLAP (large Q, moderate x), BKFL (low x), and CCFM (unifying over full range)
- Compare against different MC programs
  - Rapgap*, DGLAP LL approximations
  - Djangoh*/Ariadne, BKFL-like
  - Cascade*, CCFM
  - Herwig++, DGLAP-like with angular ordering and cluster fragmentation model



* Hadronization with Lund fragmentation model (Pythia)

Additional jet observables explored by ATLAS

kT splitting scales: $\sqrt{d_{12}}$, kT distance of last clustering step in kT clustering

N-subjettiness: τ_N a measure of how many subjets a jet has

Width: small width, pT distributed closer to jet core; close to 1, pT distributed near edges

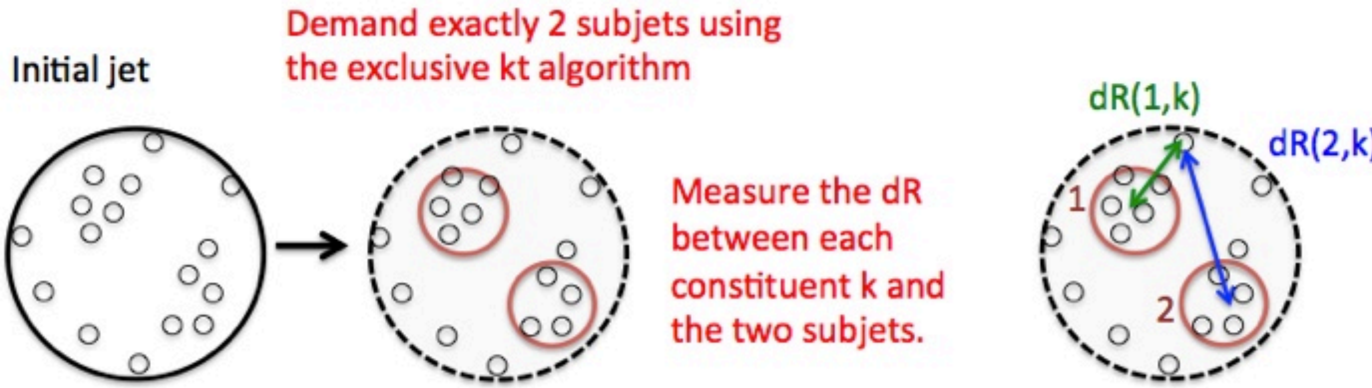
Eccentricity: measures deviation of jet profile from a perfect circle

Angularity: measures the degree of symmetry in jet energy flow

Planar Flow: measures if a jet is spread evenly over a plane or linearly

Cartoons for the different observables in the additional material

N-subjettiness

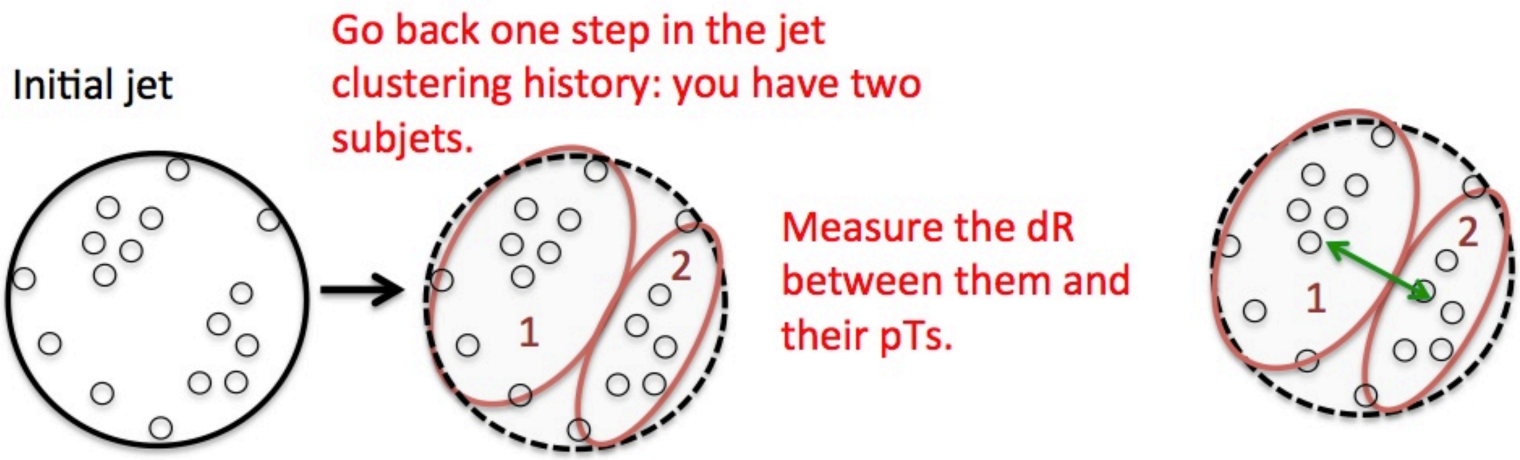


d_k = pT(k) x min(dR(1,k),dR(2,k))

- If constituent k is within, or close to, a subjet, the d_k will be small.
- Sum over all the d_k, and divide by $d_0 = \Sigma(p_T(k) \times R)$, where R is the initial jet radius.
- Now $\Sigma d_k / d_0$ is the two-subjettiness, τ_2. If this is small, the jet is very two-subjetty. If it is close to 1 (or above- see note * below) then it is not.
- To get τ_1, demand a single subjet. To get τ_3, demand exactly three, and take the minimum of the three dR(i,k) values.

* Note: the min(dR(1,k),dR(2,k) *can* be larger than R. If the *average* min is larger than R (unlikely but possible), we get a value for τ_2 that is larger than 1.

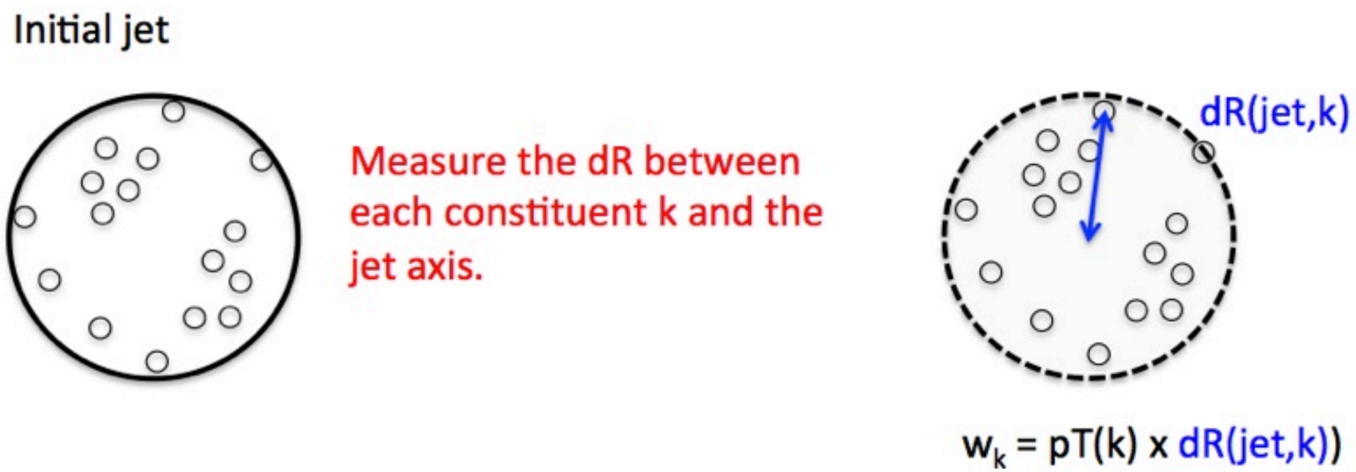
kT splitting scale



$\sqrt{d_{12}} = \min(pT(1), pT(2)) \times dR(1,2)$

- If the distance between the subjets is large, $\sqrt{d_{12}}$ is large.
- If the softer of the two subjets in the last clustering has high pT, then $\sqrt{d_{12}}$ is large.
- Both these things indicate large $\sqrt{d_{12}}$ in symmetric two body decays.

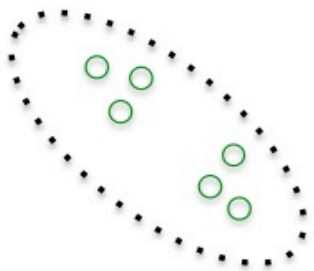
width



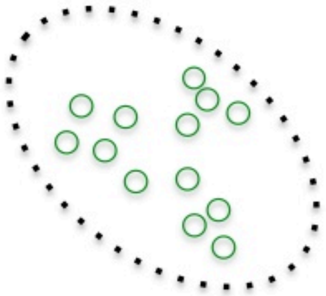
- If constituent k is close to the jet axis, the w_k will be small.
- Sum over all the w_k, and divide by $d_0 = \Sigma(p_T(k) \times R)$, where R is the initial jet radius.
- Now $\Sigma w_k / d_0$ is the jet width. If this is small, the jet has much of its energy concentrated near the core. If it close to 1, then the energy is distributed further out towards the edges.

planar flow

This jet has planar flow ->0 because it has a fairly linear deposition of energy in eta,phi.

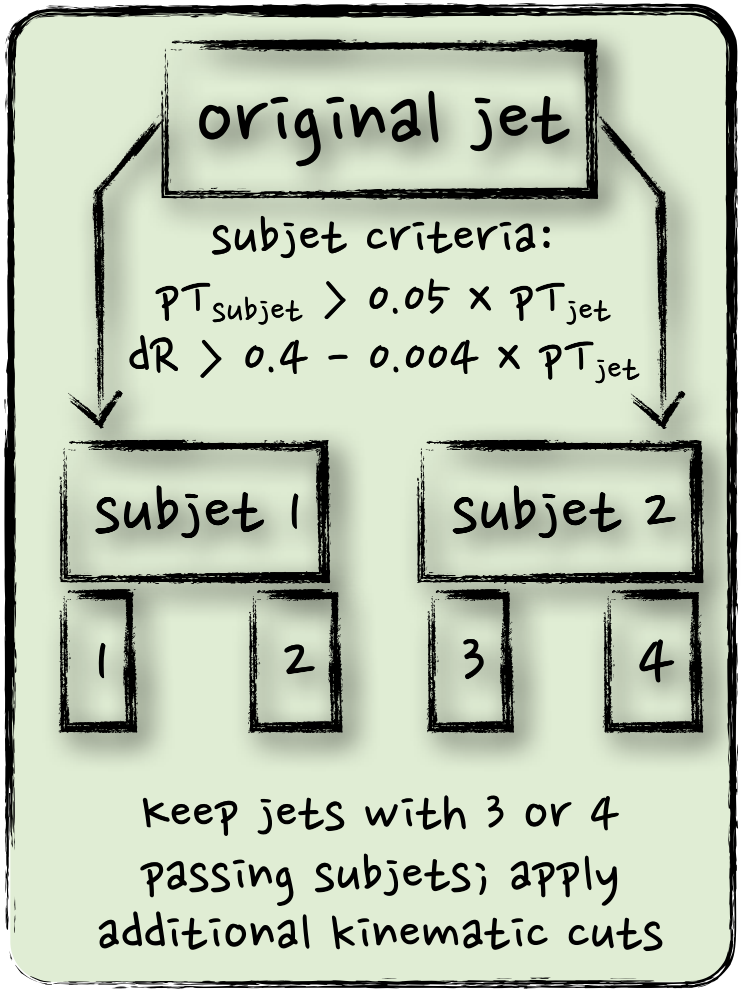


This jet has planar flow ->1 because it has a planar deposition of energy in eta,phi.



CMS top tagger

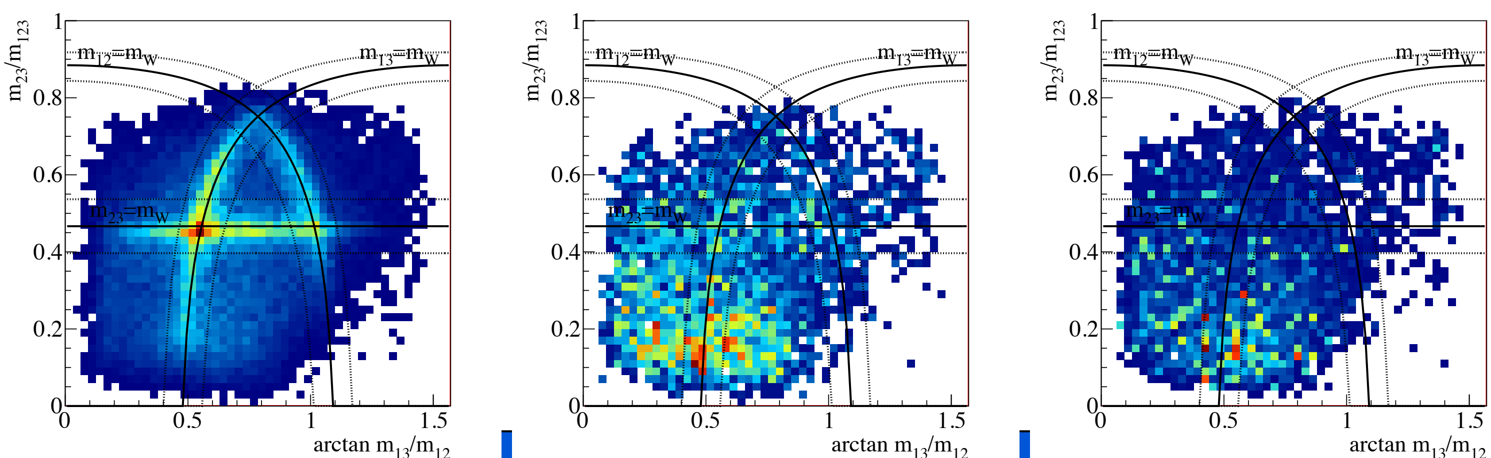
- Based on the JHU top tagger:
  PRL 101/142001 (2008)
  Kaplan et al.
- Cluster jets with CA8 algorithm
- Reverse clustering algorithm to find subjets, keep subjets passing following criteria
  - $pT_{subjet} > 0.05 \times pT_{jet}$
  - $dR > 0.4 - 0.004 \times pT_{jet}$
- Keep original jets with 3 or 4 passing subjets
  - Jet mass is [100-250] GeV
  - Minimum pairwise mass of hardest 3 subjets, $m_{min} > 50$ GeV



arXiv:1006.2833
Plehn et al.

HEP top tagger

- For identifying moderately boosted tops, CA fat jets (R = 1.5, 1.8) with pT > 200 GeV
- Decluster jet keeping subjets that pass the mass drop criterion, $m_{j1} > m_{j2}$ and $m_{j1} < 0.8 \times m_j$ until each subjet each subjet has $m_{j,i} <$ 30 GeV
- Filter all combinations of triplets of subjets to remove UE/PU contributions, keeping 5 hardest filtered consituents to compute the jet mass; keep triplet with jet mass closest to m_t
- Apply kinematic constraints on all mass pairings: $\{m_{12}, m_{23}, m_{13}\}$

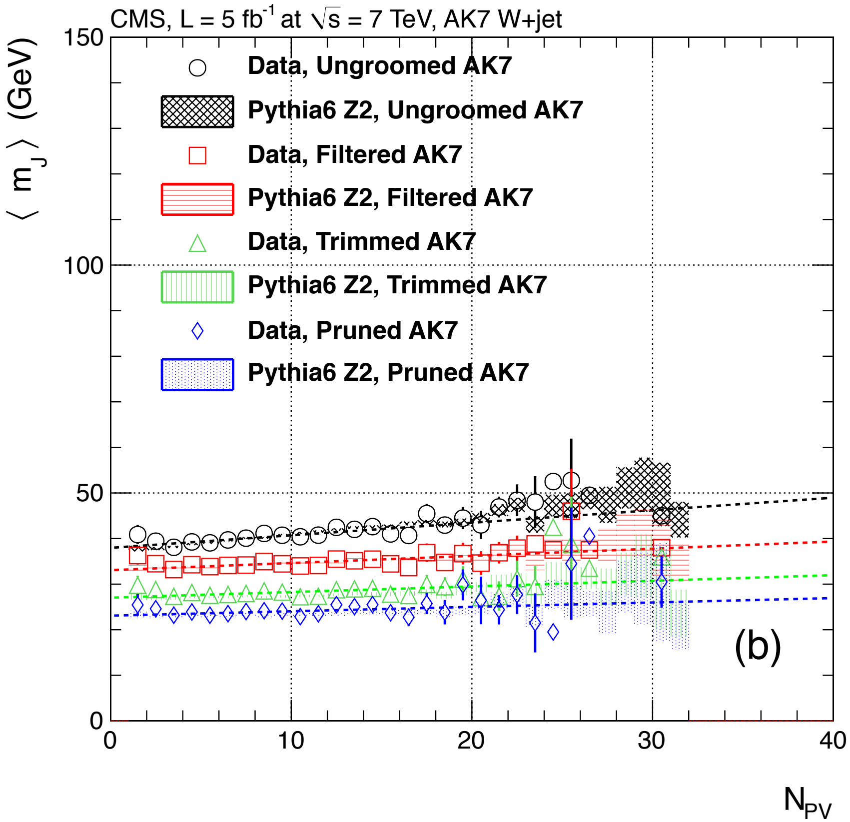


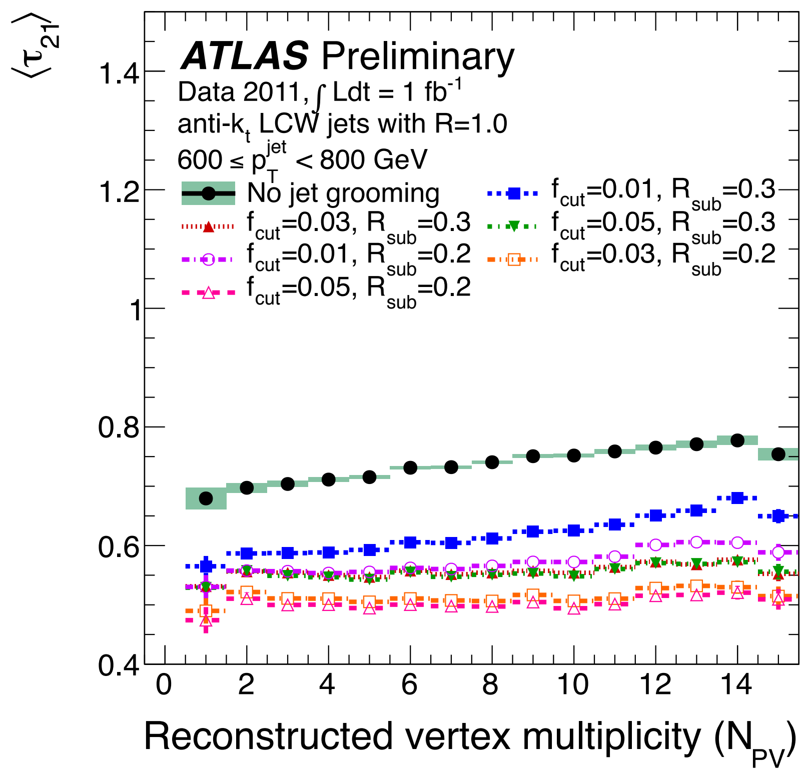
ATLAS-CONF-2012-066, JHEP 05 (2013) 090

Examine jet structure observables as function of primary vertices

Grooming jet mass reduces pileup dependence

Can also reduce pileup dependence for other jet structure observables such as τ_2/τ_1 using only groomed jet constituents





Top Template Tagger

- Selection: AK10 jets with pT_1 (pT_2) > 500 (450) GeV
- Energy flow inside a jet compatibility with top quark decay
- Given a library of ~300k templates, encode the overlap into a single observable OV_3 (from 0-1)
- Libraries in bins of 100 GeV starting from 450 GeV

$$OV_3 = \max_{\{\tau_n\}} \exp\left[-\sum_{i=1}^{3} \frac{1}{2\sigma_i^2}\left(E_i - \sum_{\substack{\Delta R(\text{topo},i)\\ <0.2}} E_{\text{topo}}\right)^2\right]$$

  - τ_n is set of templates
  - i sums over top-quark decay daughters, $\sigma_i = E_i/3$ is weight factor, Etopo is energy of topocluster required to be within $\Delta R < 0.2$
- Selection, make a cut on $OV_3 > 0.7$