

Boosted objects tagging in ATLAS

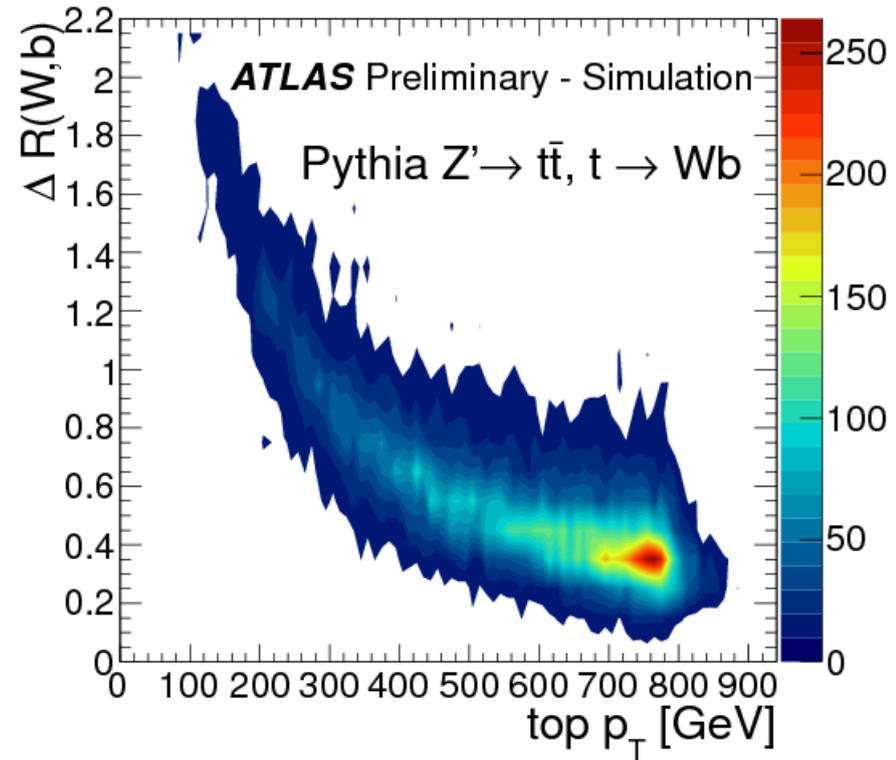
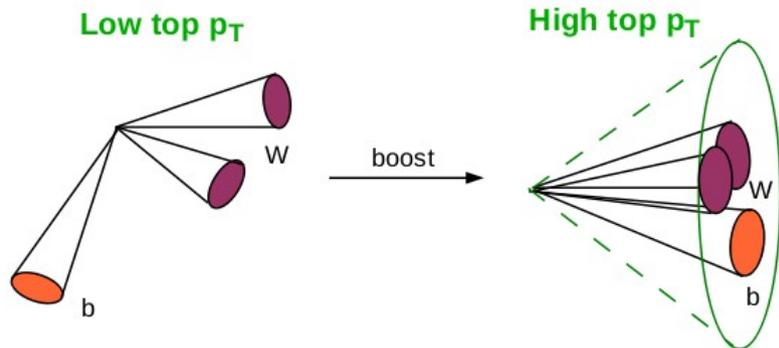
Pierre-Antoine Delsart

Laboratoire de Physique Subatomique et Corpusculaire
CNRS/IN2P3

Motivations

- Now exploring very high p_T regions
- boosted objects (top, W, H...) → **colimated hadronic decays**
- Classic reco techniques fail

➤ Hadronic jets merging



Full colimation within R when

$$P_T > 2 m / R$$

(400 GeV for $R=0.4$ and $m=m_W$)

Boosted reconstruction

Must use large hadronic jets because of collimation

- Pick-up more pile-up and underlying events
 - Fake large mass, decrease resolution, ...
 - Evaluate new procedures to filter/clean jet constituents

Boosted top reconstruction

Must use large hadronic jets because of collimation

- Pick-up more pile-up and underlying events

- Fake large mass, decrease resolution
- Evaluate new procedures to filter/clean jet constituents

=> Jet Grooming

- A single large R jet : more QCD-induced background

- Decays occur through weak force
- Jets have different structures than QCD jets

Boosted top reconstruction

Must use large hadronic jets because of collimation

- Pick-up more pile-up and underlying events

- Fake large mass, decrease resolution **=> Jet Grooming**
- Evaluate new procedures to filter/clean jet constituents

- A single large R jet : more QCD-induced background

- Decays occur through weak force **=> Jet Substructures**
- Jets have different structures than

- Combine substructure variables and kinematics of decay products

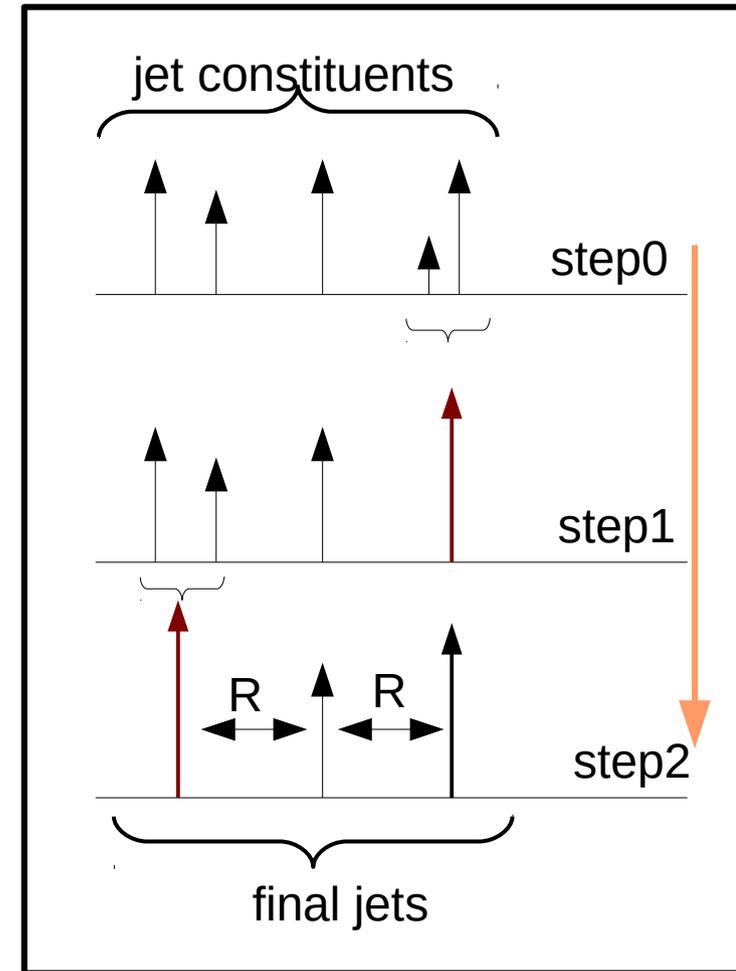
=> Object tagging

Reminder on jet algs

- Depends on a distance parameter R
- Start from input 4-vectors
 - Find *closest* pair p_1, p_2

$$d_{ij} = \min(p_{Ti}^{2p}, p_{Tj}^{2p}) \frac{\Delta R_{ij}^2}{R^2}$$

- Replace pair by p_1+p_2
- Iterate until ~all separated by R
- Parameter $p=1$ (kt), $p=0$ (Cambridge), $p=-1$ (anti-kt)
 - (Low p_T first)
 - (angular ordering)
 - (high p_T first)



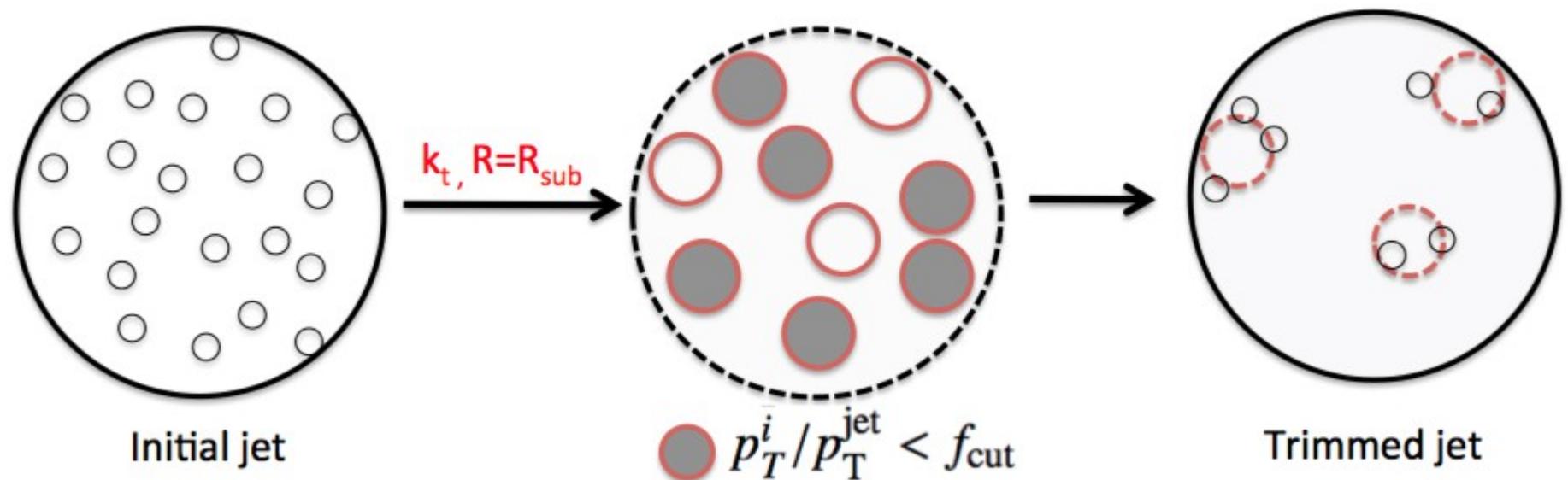
Jet Grooming techniques

- Several techniques studied in Atlas
- **Trimming**
- **Pruning**
- **Mass-Drop**

Jet Grooming techniques

■ Trimming

- Recluster jets with small R k_t alg
- Remove subjects with low pt fraction. Remaining constituents form the trimmed jet.

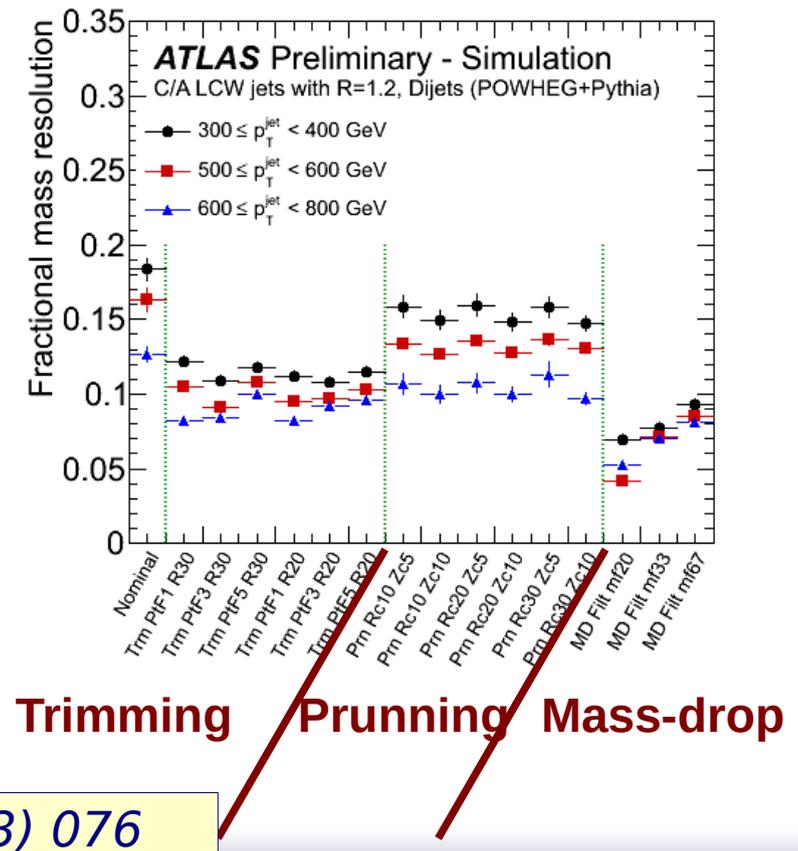
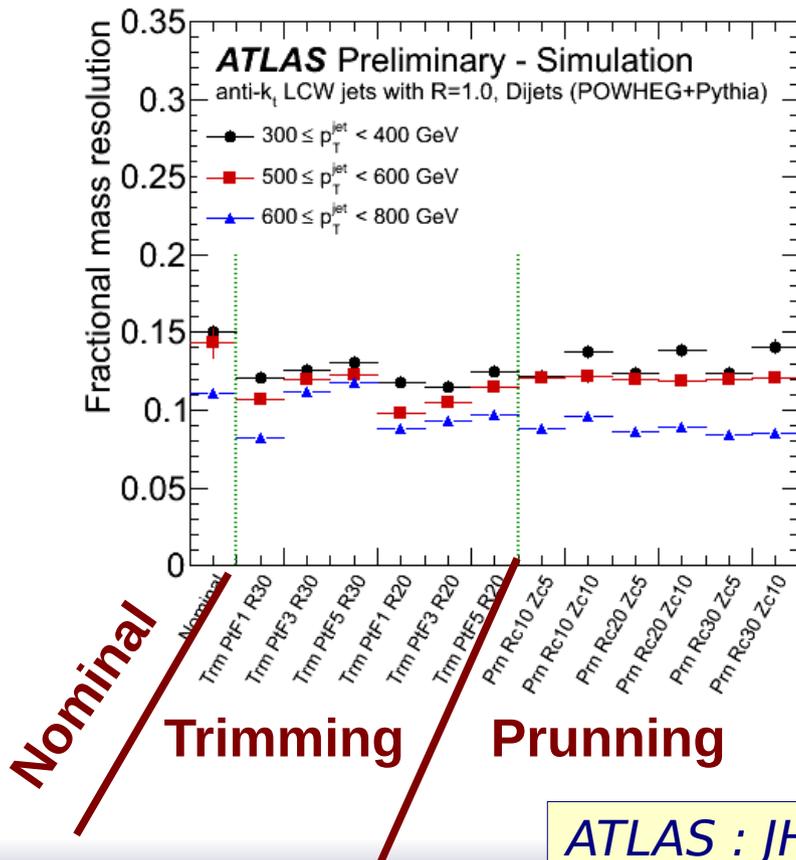


Grooming performances

Studied in Atlas on inclusive jet samples (Data and MC)

■ Jet mass resolution

$$\sigma \left(\frac{Mass_{truth} - Mass_{reco}}{Mass_{truth}} \right)$$



ATLAS : JHEP09 (2013) 076

Jet Grooming usage in ATLAS

- Atlas prefers Trimming and Mass-drop
 - better mass resolution
 - less sensitive to PU
- Atlas uses trimmed jets in physics searches
 - Large $R(=1.0)$ jet in $t\bar{t}$ resonances

Substructure variables

Many substructure variables invented.

Atlas studied several of them

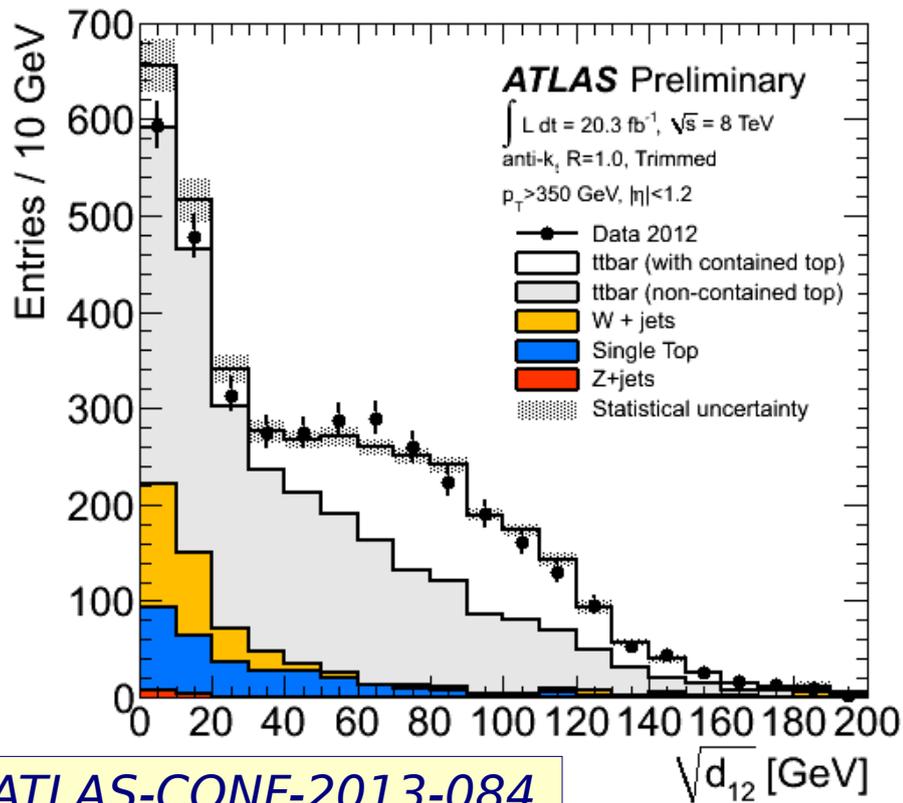
- Jet Mass
 - Kt split scale
 - N-subjettiness
 - Q_w
 - Shape variables
 - Shower deconstruction
 - Q-jets
 -
- Show only a selection based on recent public results

Substructure : d_{12} (split scale)

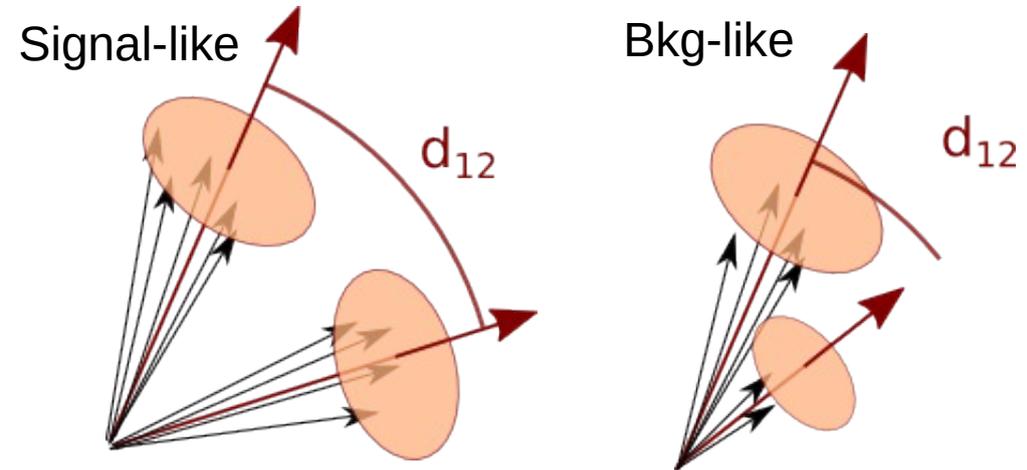
- Last recombination distances used in **Kt** alg :

$$\sqrt{d_{ij}} = \min(p_{Ti}, p_{Tj}) \times \Delta R_{ij}$$

Example : in a boosted $t\bar{t}$ selection :



ATLAS-CONF-2013-084



Variant :

$$\sqrt{y_f} = \frac{\sqrt{d_{12}}}{m}$$

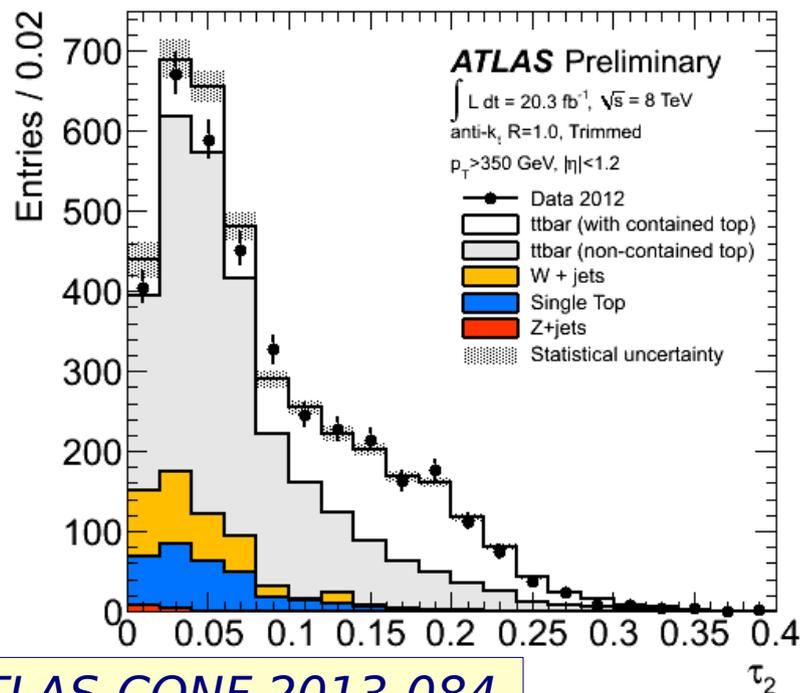
Substructure : n-subjettiness

- Recluster a jet in N kt subjects, then compute :

$$\tau_N = \frac{1}{d_0} \sum_k p_{Tk} \times \min(\delta R_{1k}, \delta R_{2k}, \dots, \delta R_{Nk}), \text{ with } d_0 \equiv \sum_k p_{Tk} \times R$$

- Consider ratios of these quantities (ex : $\tau_{21} = \tau_2/\tau_1$)

Example : in a boosted tbar selection :



ATLAS-CONF-2013-084

τ_N small when jet is N-prong

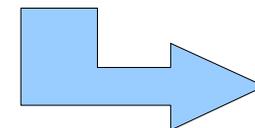
Ex : ratio τ_{32} quantifies how a jet is 3-prong rather than 2-prong

Substructure : energy correlation

- 2 or 3-points energy correlation sum in a jet
 - Sum over jet constituents
 - $e_N = 0$ when exactly N constituents

$$e_2 = \frac{1}{p_T^2} \sum_{i < j} p_{Ti} p_{Tj} R_{ij}$$

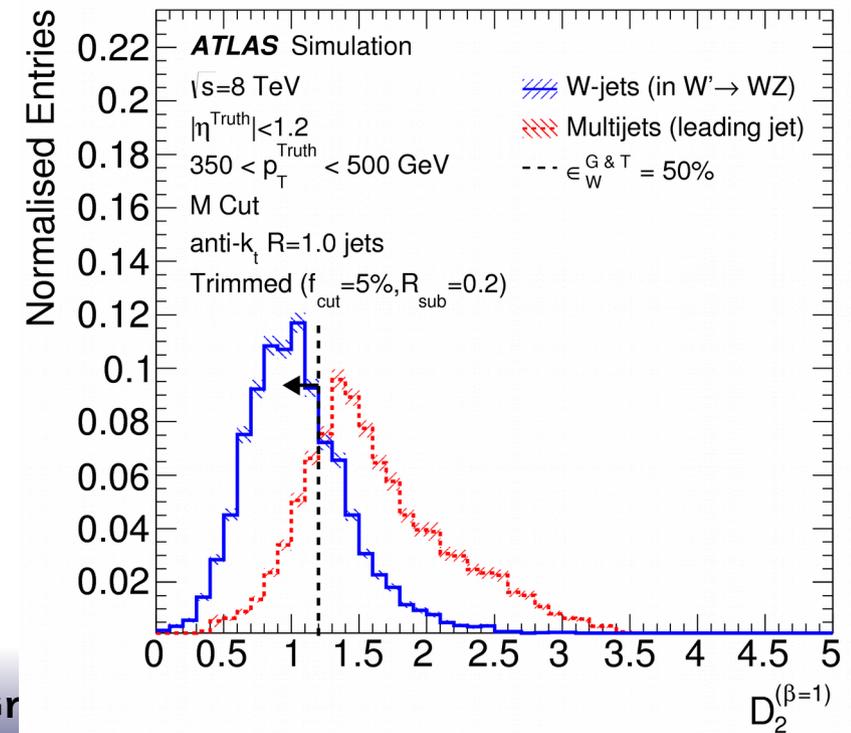
$$e_3 = \frac{1}{p_T^3} \sum_{i < j < k} p_{Ti} p_{Tj} p_{Tk} R_{ij} R_{ik} R_{jk}$$



$$D_2 = \frac{e_3}{(e_2)^3}$$

- Preferred variable : D2
 - optimized to discriminate 1-prongs vs 2-prongs

ATLAS-PERF-2015-03

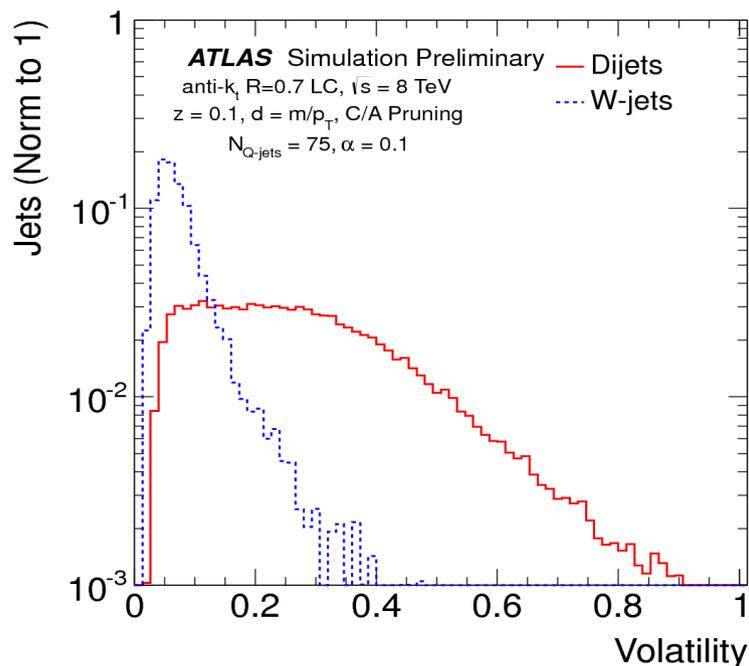


Substructure : Q-jets

- Run **pruning** clustering several times introducing **randomness** in merging criteria

- Select pair for merging randomly as in $\exp\left\{-\alpha\frac{d_{ij} - d^{\min}}{d^{\min}}\right\}$
- This simulates different possible showering history

- Obtain a mass distribution for a given jet : take the RMS



Ellis et al.

In Atlas : used in boosted W studies

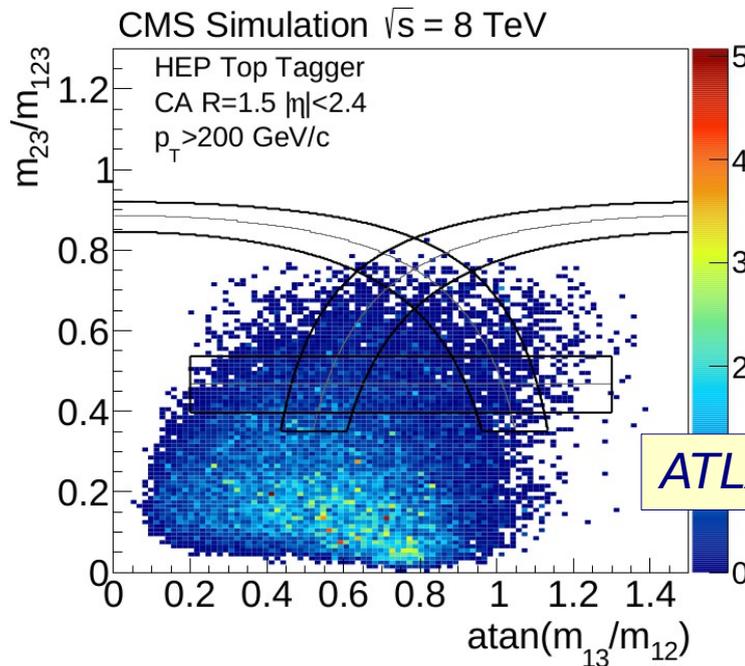
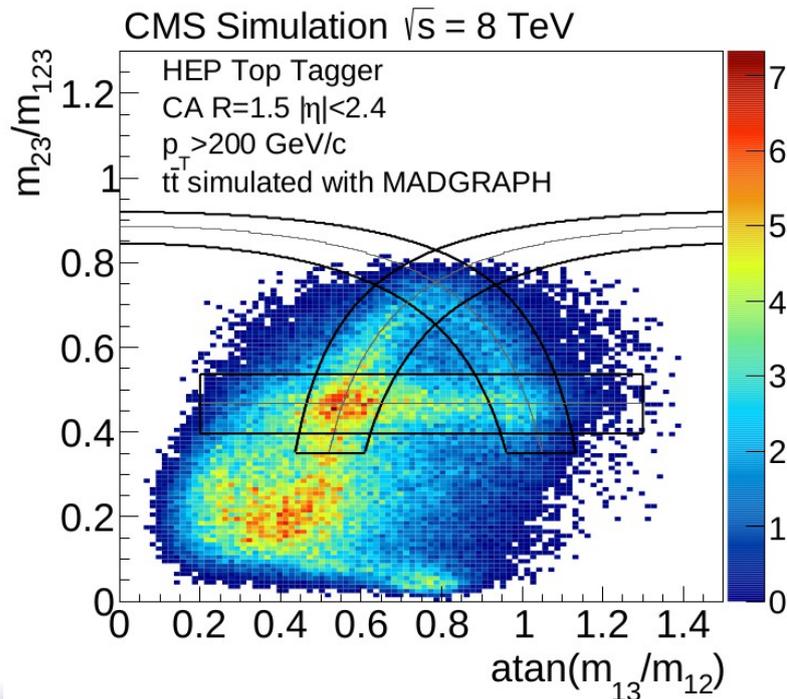
ATLAS-CONF-2013-087

Object tagging

- Cuts on substructure variables
 - ex : $m > X \text{ GeV} \ \&\& \ \tau_{32} < Y$
 - Often use pT-dependent cuts to take boost into account
- Combine variable in multivariate-analysis
 - Boosted Decision Trees
 - deep neural network...
- Dedicated procedure have been proposed : Top taggers
 - HepTopTagger
 - Template TopTagger

HEP TopTagger

- Multi steps algorithm on large R(1.5) Cam jet
 - Mass-drop technique to identify hard subjets
 - Filtering of soft components, re-cluster to 3 subjets
 - Kinematic constraints on final subjets (W, top mass...)
- If succeed, results in 3 identified subjets

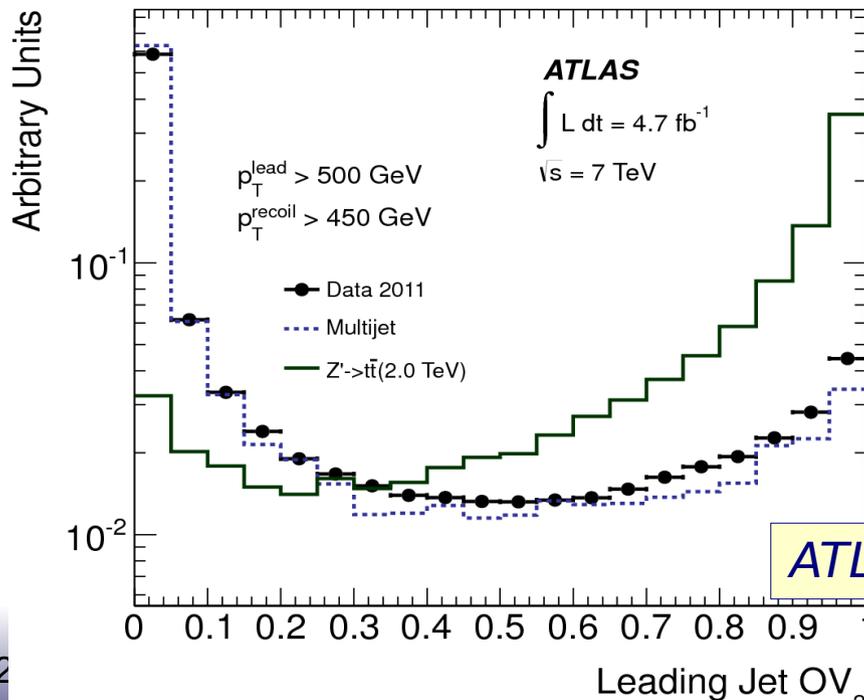


Plehn et al.

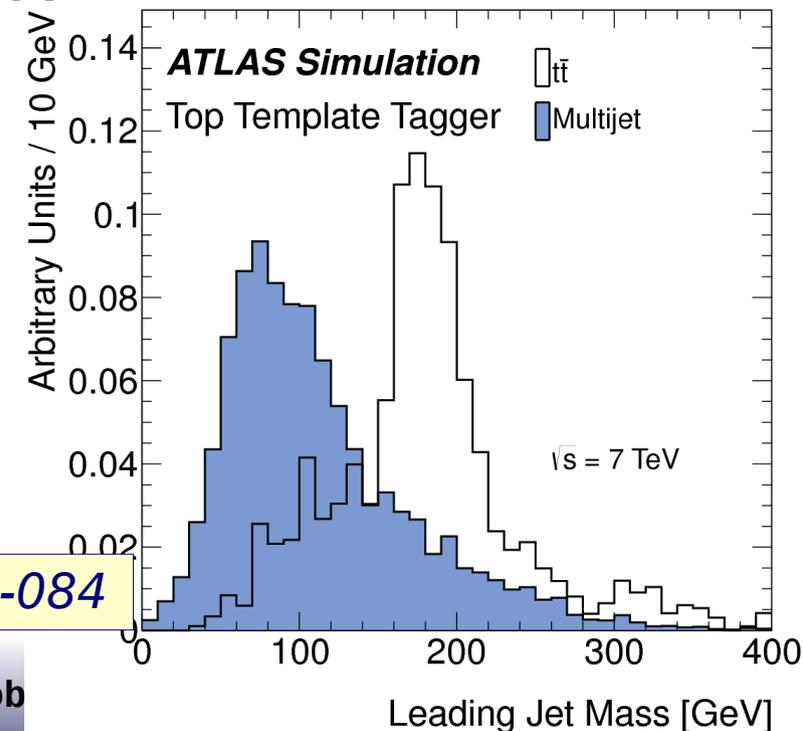
ATLAS-CONF-2013-084

Template top tagger

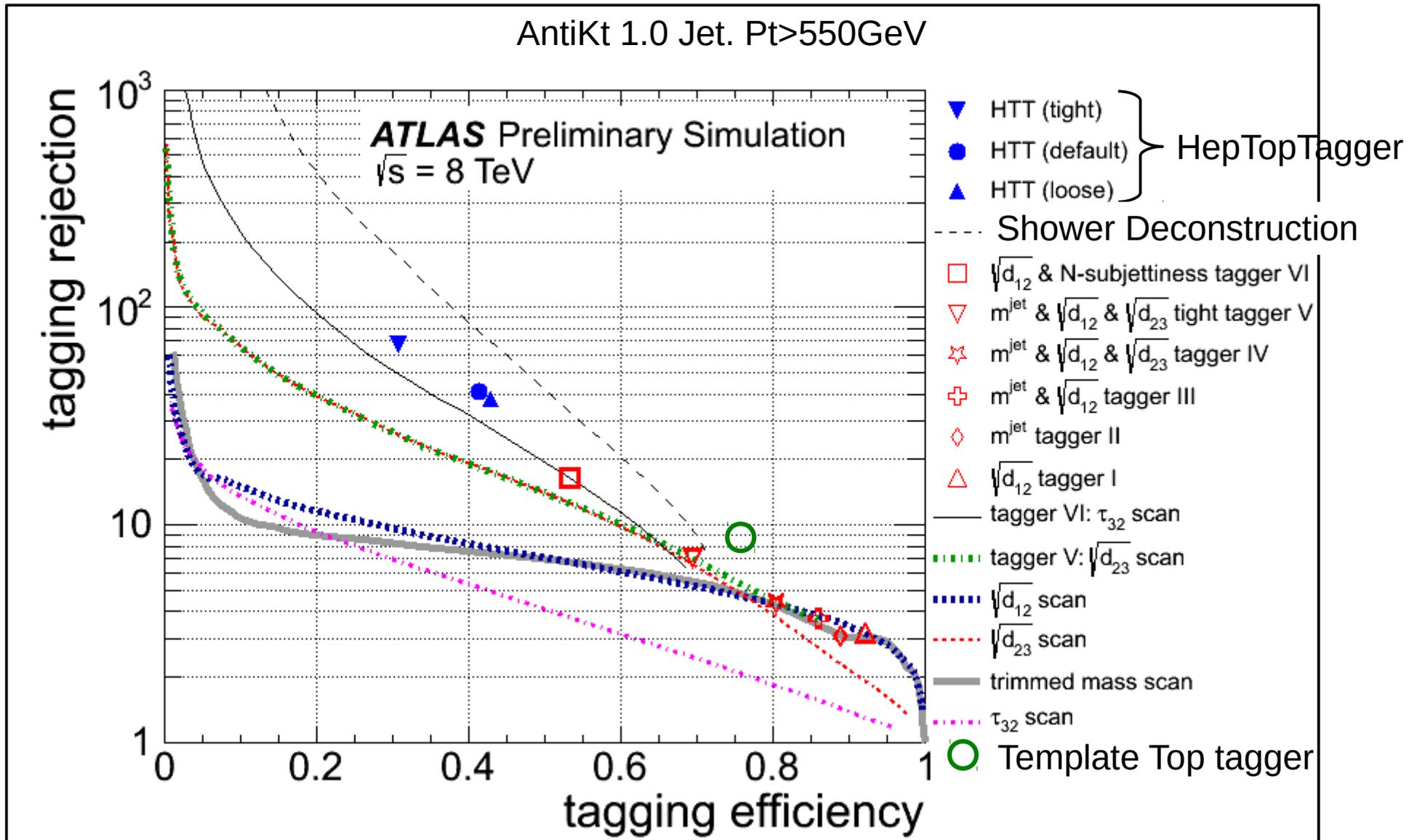
- For a top at given p_T , generate all possible 3 partons decays (== all possible templates)
- For a jet candidate, measure overlap between constituents (clusters) and each template
 - Overlap returns a number in $[0,1]$
 - Computed from E differences between close constituents and partons
 - Take the **max overlap** over all templates



ATLAS-CONF-2013-084



Top Tagging performances (Atlas)



Object tagging in Run1 analysis

■ Top resonances

- semi-leptonic : mass & d_{12} cuts

JHEP08 (2015) 14822

- full hadronic : HEPTopTagger, TemplateTopTagger

Model	Obs. Limit (TeV)	Exp. Limit (TeV)
HEPTopTagger		
Z'	$0.70 < m_{Z'} < 1.00$ $1.28 < m_{Z'} < 1.32$	$0.68 < m_{Z'} < 1.16$
KK gluon	$0.70 < m_{g_{KK}} < 1.48$	$0.70 < m_{g_{KK}} < 1.52$
Top Template Tagger		
KK gluon	$1.02 < m_{g_{KK}} < 1.62$	$1.08 < m_{g_{KK}} < 1.62$

ATLAS : JHEP 1301 (2013) 116

■ VV resonances

- Mass-Drop grooming technique
- mass & y_f & n_{trk} cuts

*ATLAS-EXOT-2013-08
(seen yesterday!)*

Object tagging in RunII

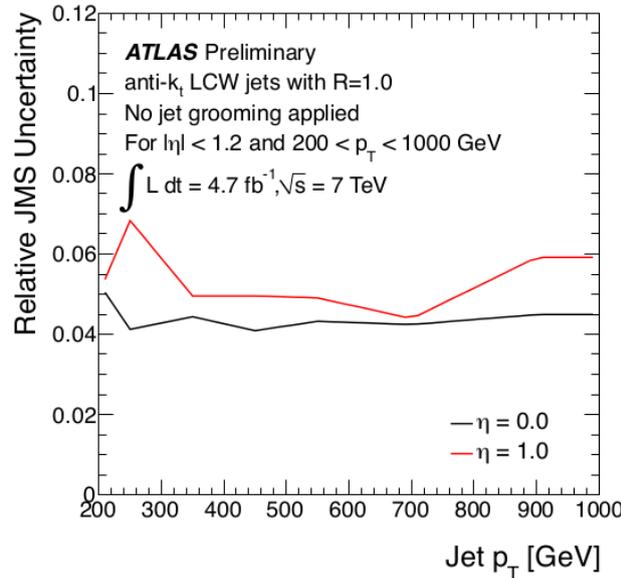
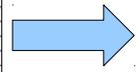
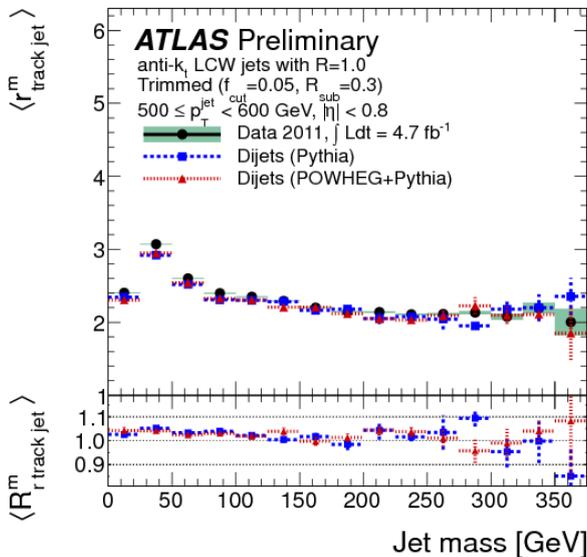
- Atlas provides recommendations for early analysis
 - simple, robust substructure variable cuts
 - uncertainties on variable prepared
 - Top : **mass** & τ_{32}
 - W,Z : **mass** & D_2
 - Higgs : **mass** & b-tagging & D_2
- Several performance analysis on-going for later analysis
 - deep neural network
 - elaborate technique (HEPTopTagger,...)

Systematics & substructures

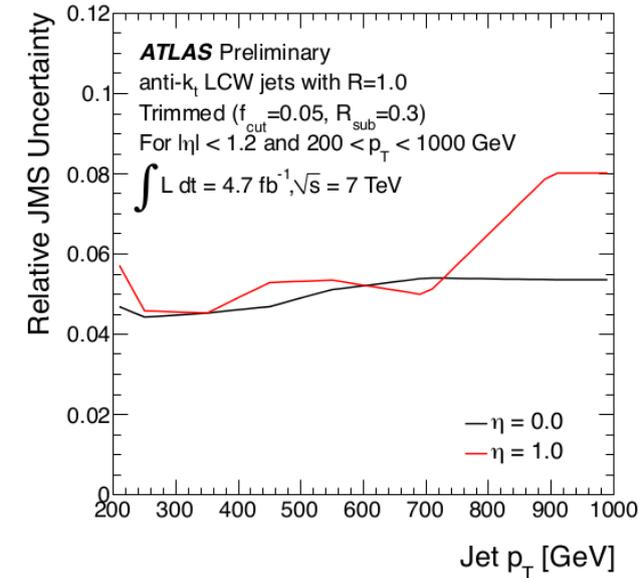
- Substructure uncertainties evaluated with "double ratios"

$$\frac{(V_{calo}/V_{track})_{data}}{(V_{calo}/V_{track})_{MC}}$$

Example : Mass



(a) anti- k_t , $R = 1.0$ (no jet grooming)



(b) anti- k_t , $R = 1.0$ (trimming)

- Elaborate taggers :

- Uncertainties on constituents & propagation ? (calculation time ?)
- Uncertainties on subjets ? (interaction with 3rd party software)
- Is final sensitivity really improved ?

Summary & Conclusion

- Boosted topologies now important parts of analysis
- Lots of ideas have been tested at the LHC
 - Many perf studies done and on-going
 - Some analysis used complex taggers
- State of the art taggers usage still lagging
 - Optimal use is much more complex/time consuming
 - planned for next rounds of analysis (or next-to-next...)
- More challenges ahead : approaching calo granularity limit ?
 - At $p_T \sim 1.5\text{TeV}$ (W,Z) or 3TeV (tops)
 - Make more use of tracker ?

Other refs

■ Atlas references :

- W-tagging: <https://cds.cern.ch/record/1690048>
- top-tagging: <https://cds.cern.ch/record/1571040>
- shower deconstruction: <https://cds.cern.ch/record/1648661>
- Q-jets: <https://cds.cern.ch/record/1572981>

■ CMS references

- Boosted Top Jet Tagging
- Jet Substructure Algorithms



Improved n-subjetiness

Optimized n-subjettines

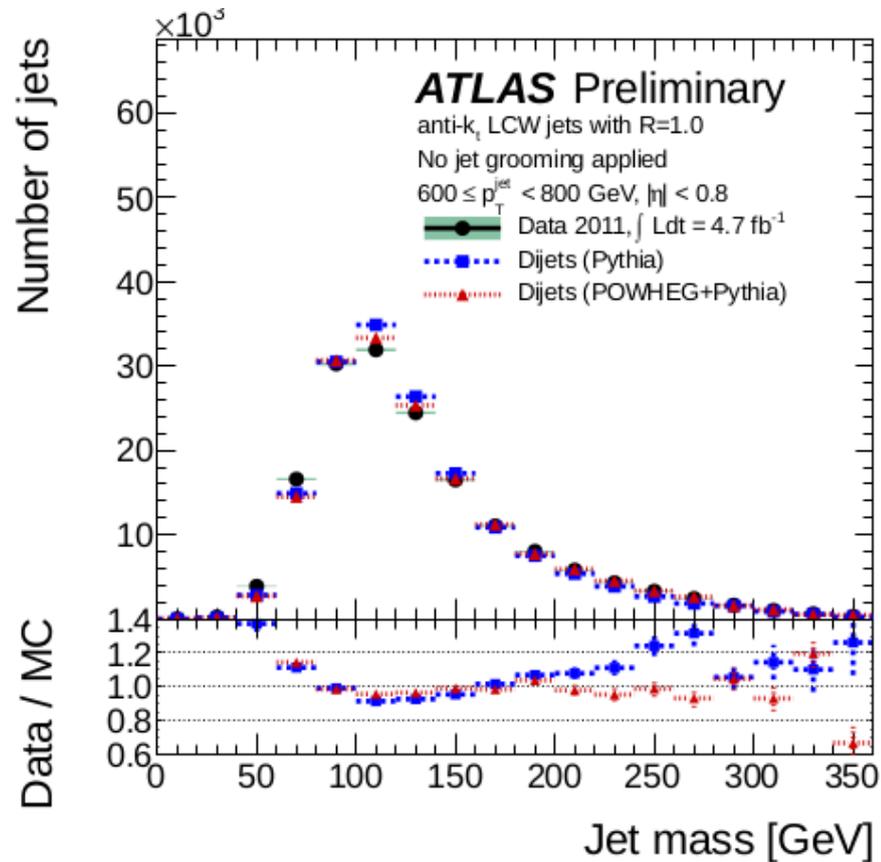
- Do not consider only subjects
- instead vary axis
- take the minimum τ_N

Thaler & Van Tilburg
hep-ph 1108.2701

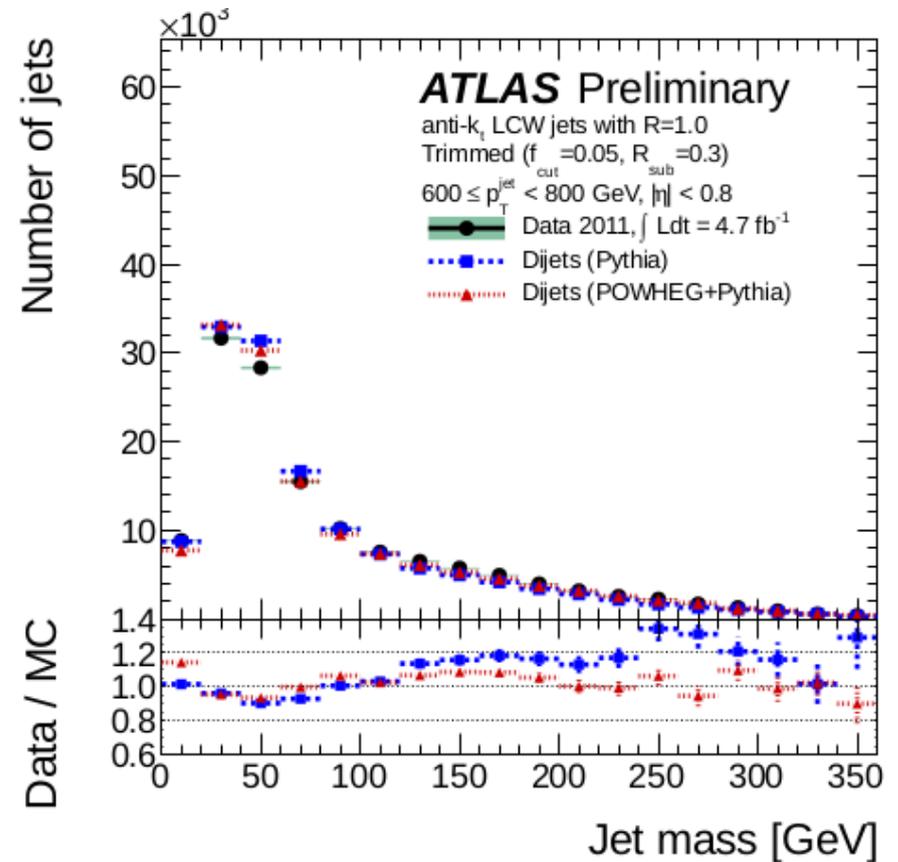
Better discrimination power

Grooming performances

Data/MC, effect on mass distribution



No grooming

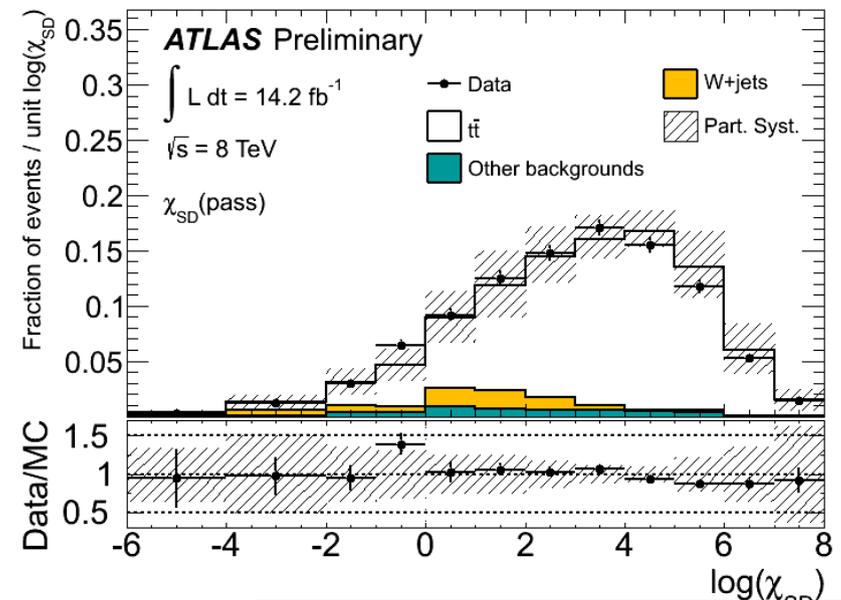
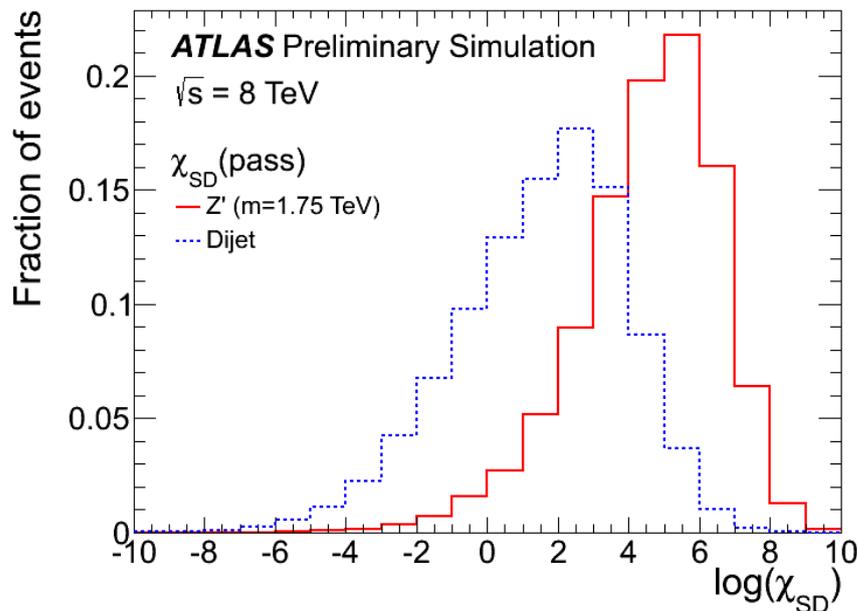


Trimmed

Substructure : shower deconstruction

- Cluster small subjets from a large jet => configuration of N subjets
- Calculate probabilities for this subjet configuration assuming signal (**top**) or assuming background (**QCD**)
 - Probabilities calculated as in parton shower generator, using Sudakov factors
- Build a likelihood ratio from these proba χ_{SD}

*D. E. Soper and
M. Spannowsky*



ATLAS-CONF-2014-003