X ⟶ *bb* and Top-Tagging in ATLAS

Mike Nelson, University of Oxford

michael.nelson@physics.ox.ac.uk

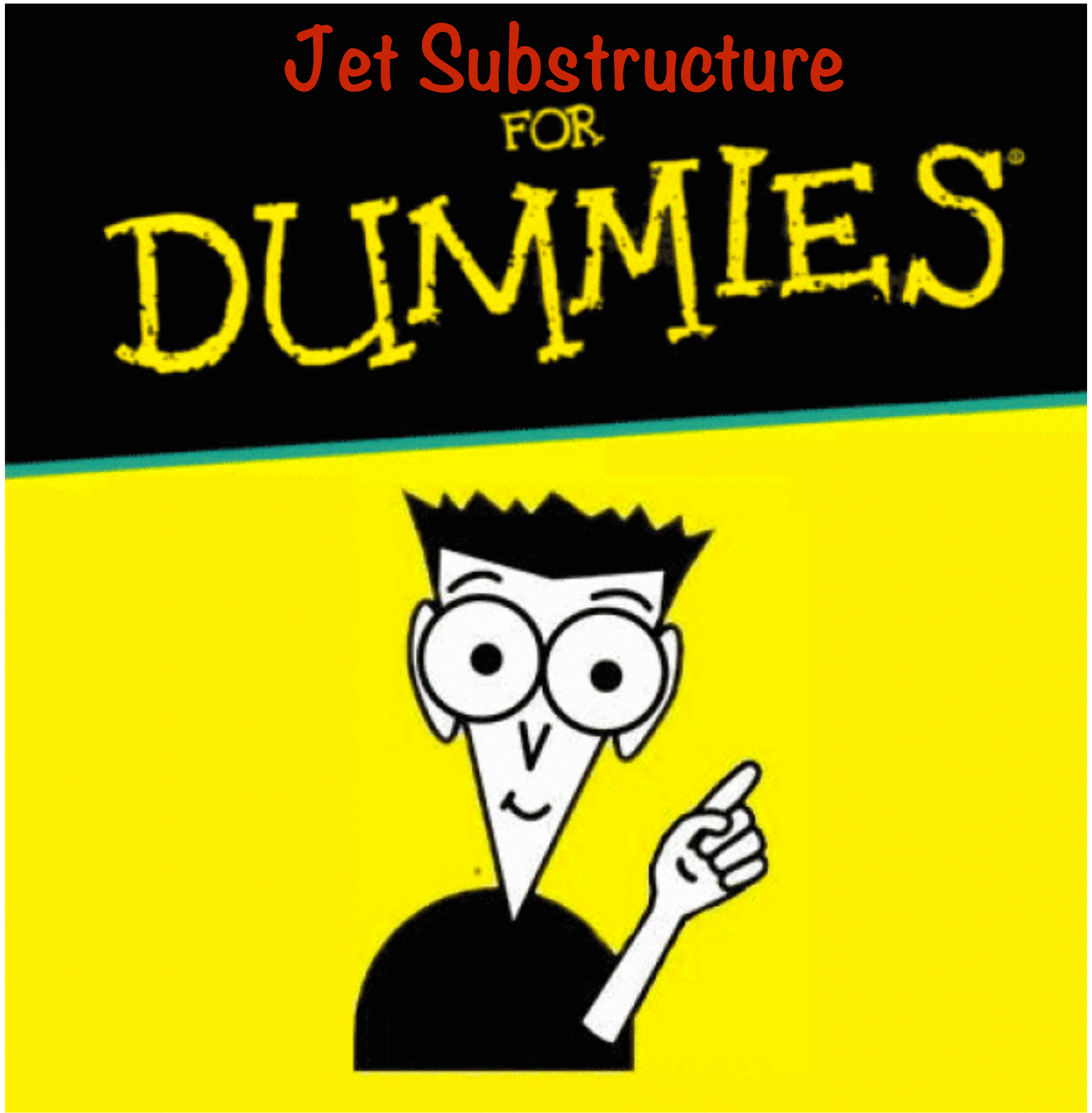
HF@LHC, 2017

Focus of the discussion

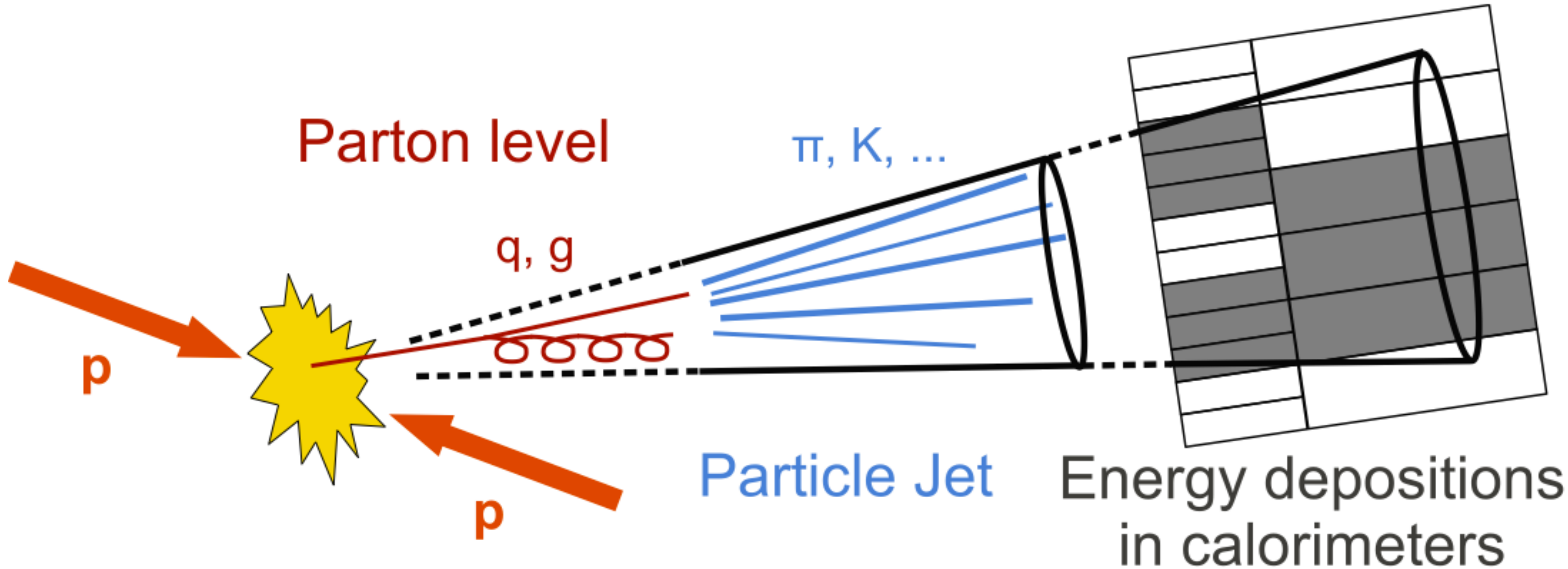
- I want to try and achieve **two** things:
  - Introduce the **basic tools** employed in ATLAS jet taggers … the **jet substructure variables**.
  - Present the **latest** jet substructure and machine-learning-based taggers available as of **BOOST2017** —> new cut-based top-taggers, DNN-based top-taggers, and $X \longrightarrow bb$ taggers using track-jets.

- Why substructure ?
  - Angle between decay products in a jet goes as $\boldsymbol{\Delta R = 2m^{\mathrm{jet}}/p_{\mathrm{T}}^{\mathrm{jet}}}$
  - Leads to **high-p_{T} boosted objects**, which can be captured within a single large-radius jet.

Our Toolbox for Tagging



Jets in ATLAS



- Jet = collimated **spray of hadrons** resulting from the **fragmentation** and **hadronisation** of quarks and gluons produced in *pp* collisions.

- Jets are constructed by applying the **anti-k_t clustering algorithm** to energy deposits (**topoclusters**) reconstructed in the calorimeter. Anti-k_t clusters hardest p_T topoclusters first, working "outwards" to build a 3-dimensional object with a hard p_T core, and radius $R = (\Delta\eta^2 + \Delta\phi^2)^{1/2}$.

- **Small-*R*** jets: combine (electromagnetic scale) topoclusters to form jets of radius $R = 0.4$.

- **Large-*R*** jets: combine (LC scale) topoclusters to form jets of radius $R = 1.0$, and apply trimming ($R_{sub} = 0.2$, $f_{cut} = 0.05$) to mitigate contaminations from pile-up and the underlying event.

Jet Mass ATLAS-CONF-2016-035

- Jet four-momentum = sum of four-momenta of constituent topoclusters. **Jet mass** is the **invariant mass** of the sum.

- "Standard" ATLAS jet mass - **calorimeter mass, m^{calo}** from calo-jet topoclusters.

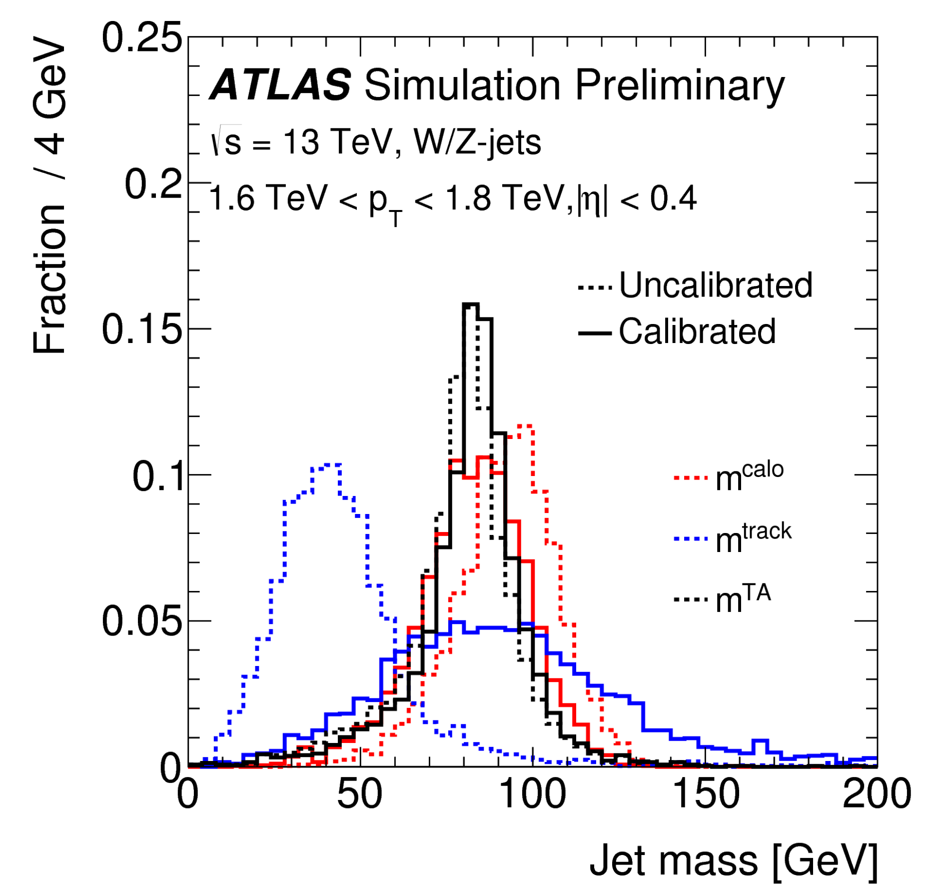
$$\boxed{m^{\text{calo}} = \sqrt{\left(\sum_{i\in J} E_i\right)^2 - \left(\sum_{i\in J} \vec{p}_i\right)^2}}$$

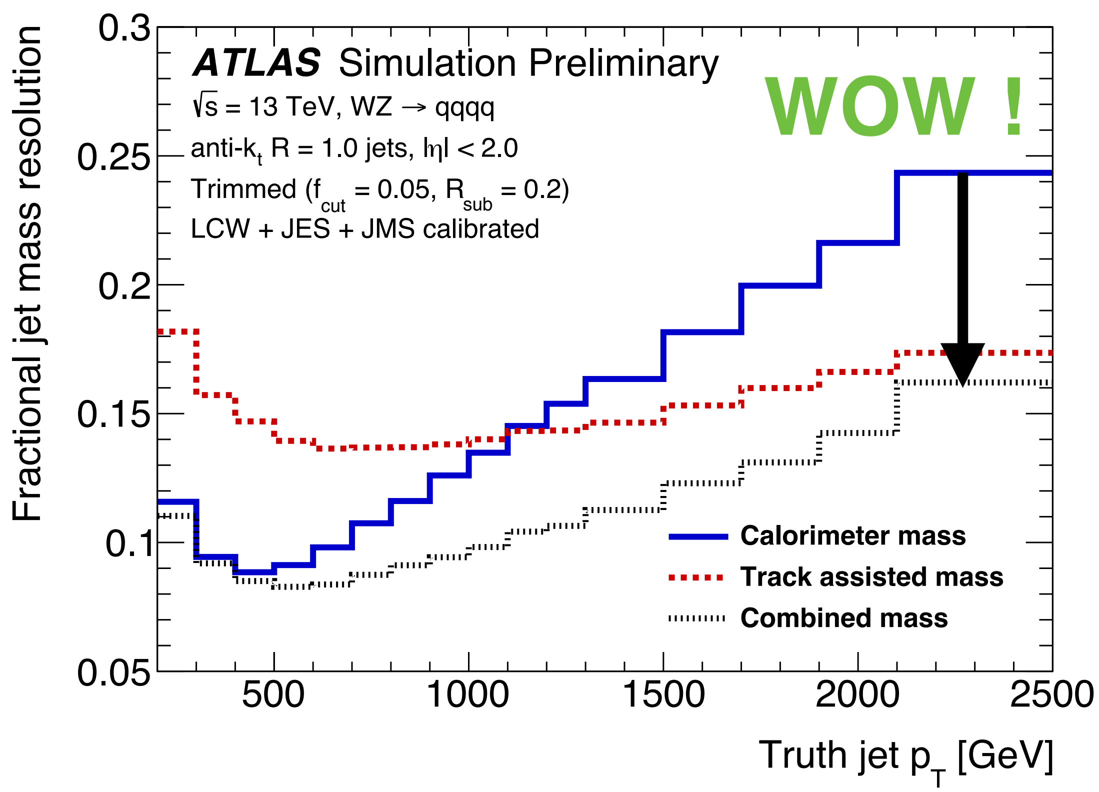
- **Track-assisted mass, m^{TA}** - associate tracks in the inner detector to a calorimeter jet, where the total mass of the associated tracks is m^{track}, which is then scaled to correct for neutral components.

$$\boxed{m^{\text{TA}} = \frac{p_{\text{T}}^{\text{calo}}}{p_{\text{T}}^{\text{track}}} \times m^{\text{track}}}$$

- **Combined mass, m^{comb}** — linear combination of m^{calo} and m^{TA}, weighted to minimise the jet mass resolution. New for Moriond, 2017.

$$\boxed{m^{\text{comb}} = a \times m^{\text{calo}} + b \times m^{\text{TA}}}$$



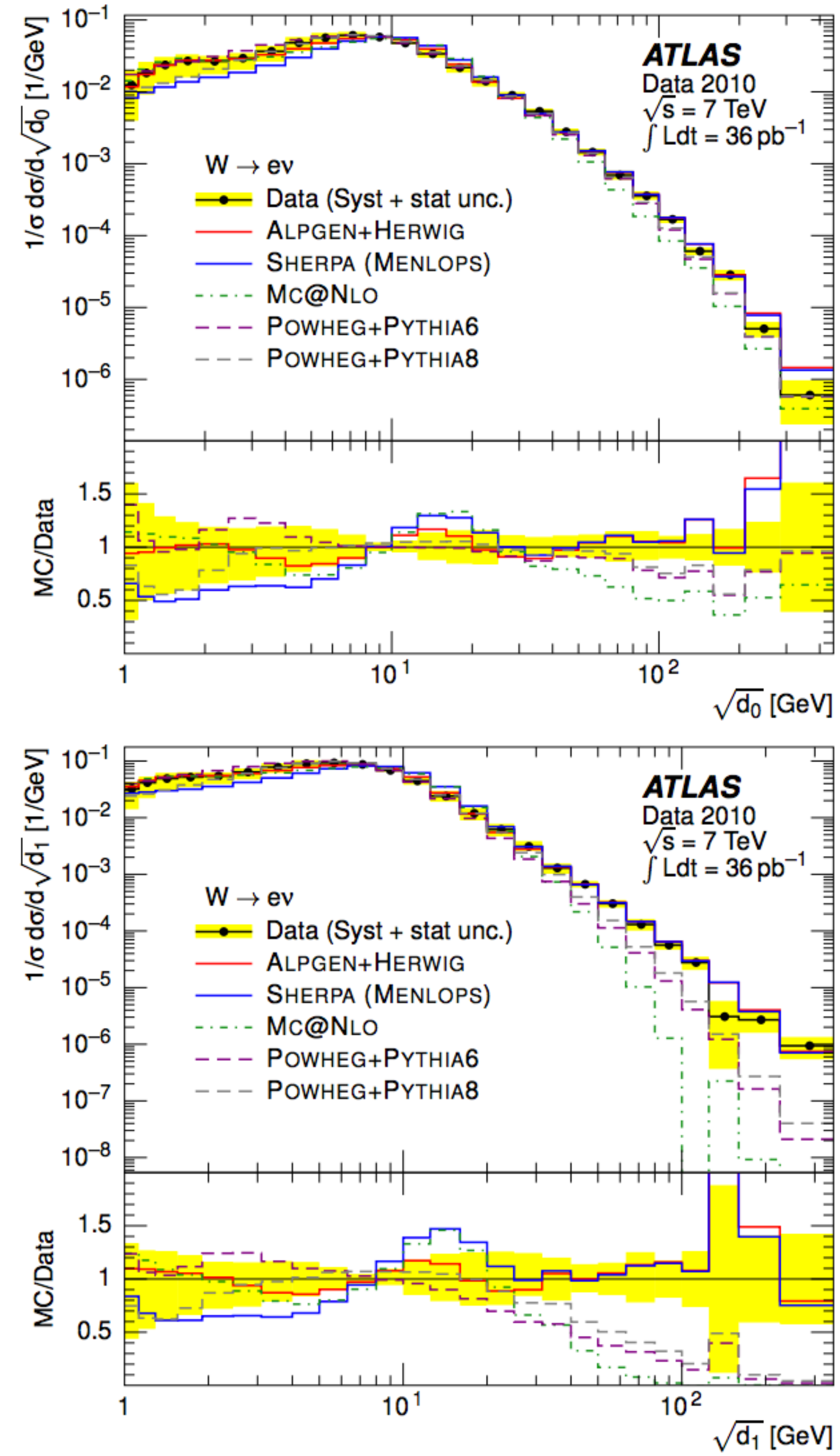


Jet Mass Splitting Scales arXiv:1302.1415

UNIVERSITY OF OXFORD

- Can reclusters the constituents of a jet applying the ***k*$_t$ algorithm**.
    - Final recombination step: jet is split into **two subjets**, with a mass-splitting characterised by
      $d_{12} = \min(p_{T,1}^2, p_{T,2}^2)\Delta R_{12}^2/R^2$
    - Penultimate recombination step: jet is split into **three subjets**, with a mass-splitting characterised by
      $d_{23} = \min(p_{T,2}^2, p_{T,3}^2)\Delta R_{23}^2/R^2$
- For bosonic jets, expect $d_{12}^{1/2} \sim m^{jet}/2$ due to the **2-prong** structure of the W/Z decay.
- For top jets, expect $d_{23}^{1/2} \sim m^{jet}/3$ due to the **3-prong** structure of the top decay.

Right: Run-1 measurement on splitting scale in a $W(e\nu)$ signal.



N-subjettiness arXiv1011.2260

- Variable τ_N quantifies the radiation pattern in a large-R jet which contains (as a hypothesis) N subjets.

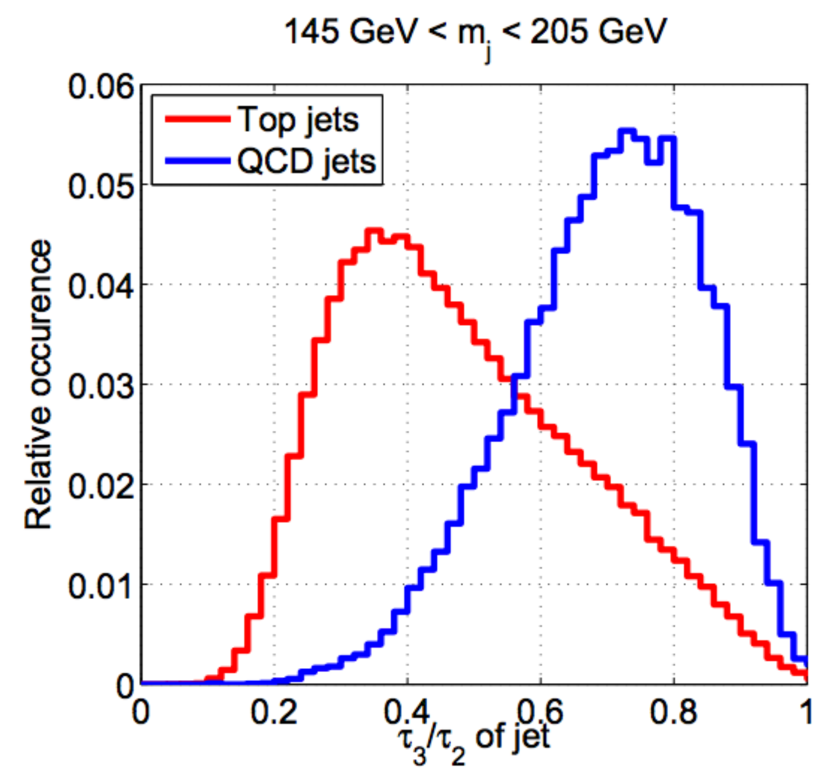
- Begin with an N-subjet hypothesis for the large-R jet and sum over k clusters in the jet.

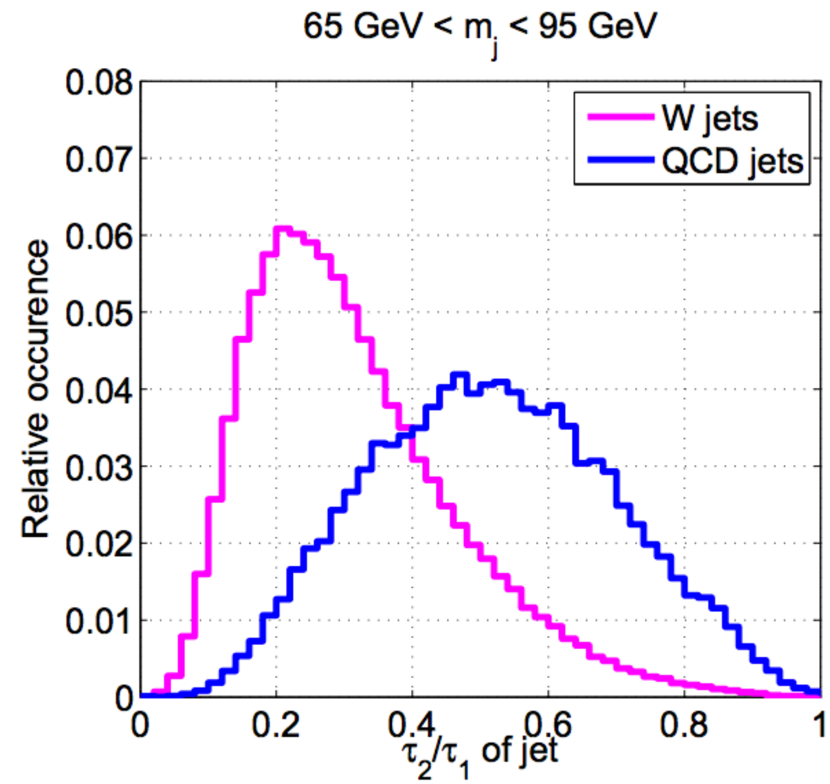
$$\boxed{\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \cdots, \Delta R_{N,k}\}}$$

- Small τ_N —> radiation strongly aligned with the axes of the N-subjets —> **N-prong radiation pattern**.

- Radios of τ_N useful discriminating different jet substructures:

    - Low $\tau_{32} = \tau_3/\tau_2$ ($\tau_{21} = \tau_2/\tau_1$) characteristic of **3-prong** (**2-prong**) energy distributions, typically expected from the decay products of **boosted top** (***W/Z/H***) jets.



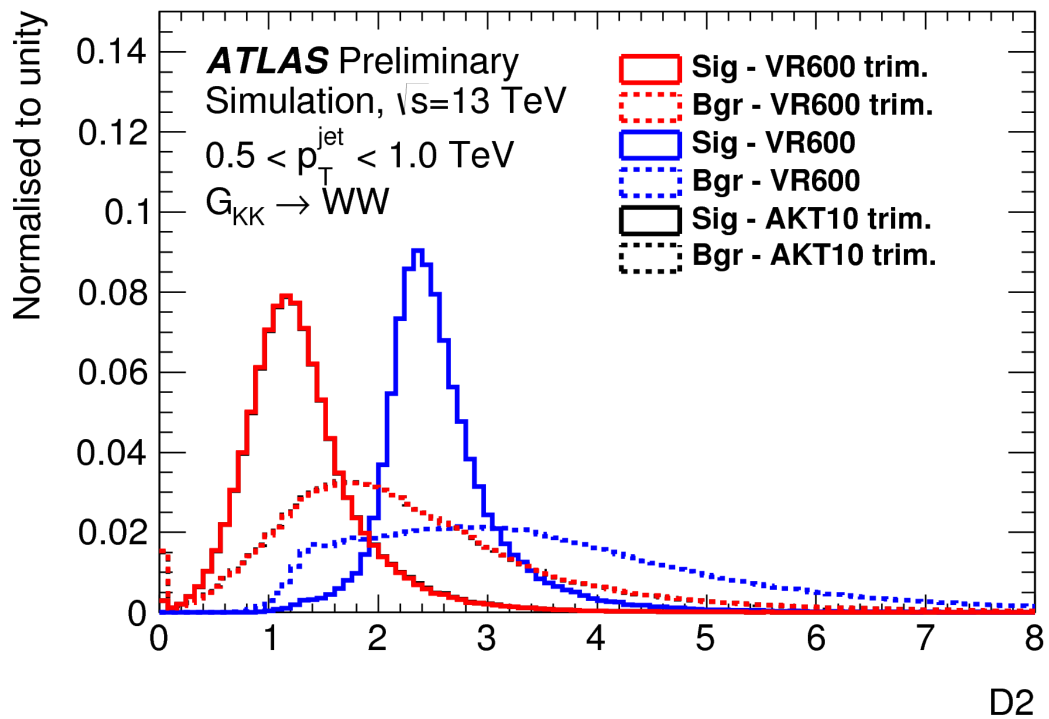


Energy Correlation Functions arXiv:1305.0007

- Instead of finding subjets, energy correlation functions rely on the **energies and the angles** between the jet constituents.

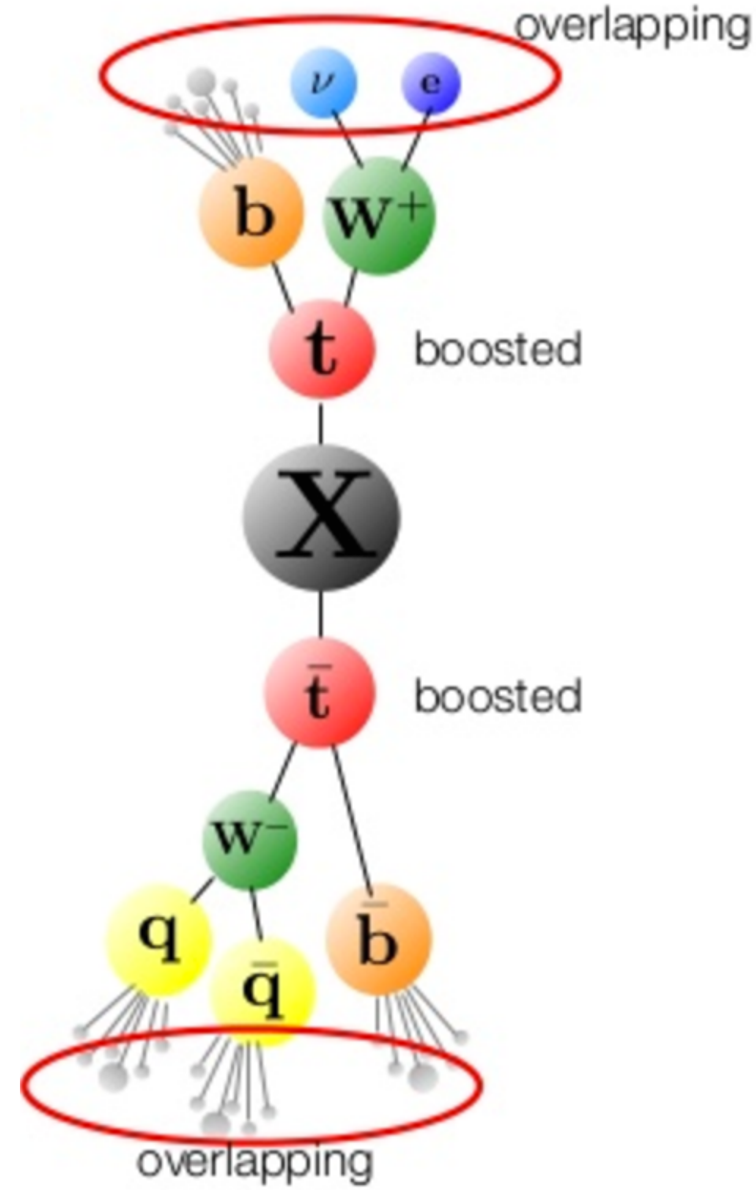
$$e_1^\beta = \frac{1}{p_{T,J}} \sum_{i \le n_J} p_{T_i}$$
$$e_2^\beta = \frac{1}{p_{T,J}^2} \sum_{1 \le i < j \le n_J} p_{T_i} p_{T_j} R_{ij}^\beta$$
$$e_3^\beta = \frac{1}{p_{T,J}^3} \sum_{1 \le i < j < k \le n_J} p_{T_i} p_{T_j} p_{T_k} R_{ij}^\beta R_{ik}^\beta R_{jk}^\beta$$

- $e_N = 0$ if there are (N-1) subjets in a jet, and, if there are N subjets, e_{N+1} should be much smaller than e_N.
- As with N-subjettiness, takes ratios of e_Ns in order to better discriminate **prong-y jets from backgrounds**.
  - Example: $D_2 = e_3/e_2^3$ is a powerful discriminator for 2-pronged jets ($W/Z/H$ jets)



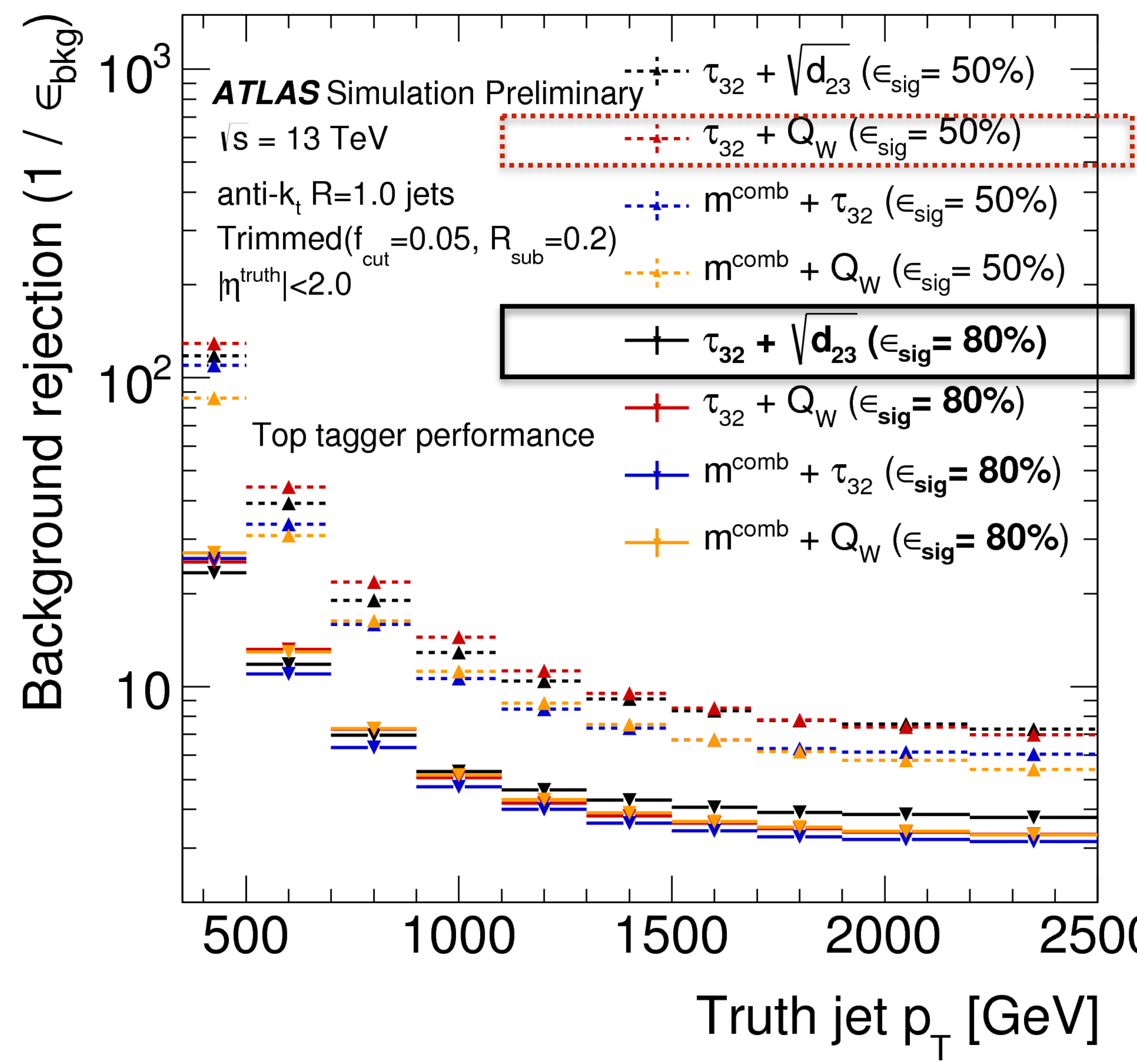
Above: D_2 distributions for a boosted W signal (solid lines) and background (dashed lines) in a variable-R jet study — ATL-PHYS-PUB-2016-013.

ATLAS Taggers: Latest and Greatest



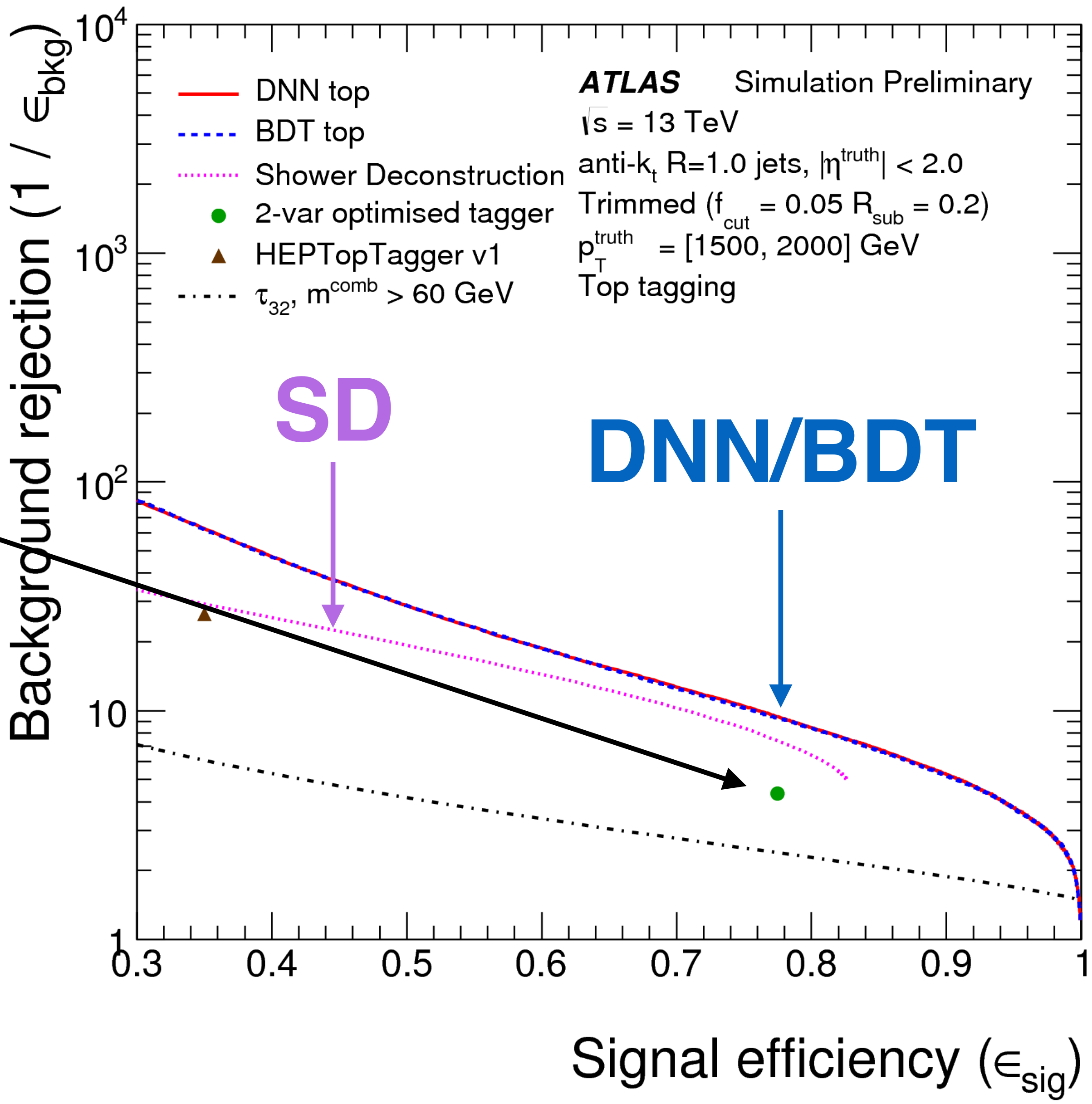
Smooth Top-Tagger ATLAS-CONF-2017-064

- Uses anti-k_t $R = 1.0$ trimmed jets, and re-optimised for BOOST2017.
- Performed a scan over combinations of two variables, determining the two variables which provide the **largest background rejection for fixed signal efficiency** working points.
- Two signal efficiency working points: **50.0 %** and **80.0 %** (used by many analyses).
- Optimised to give largest background rejection at very high p_T.
  - 50.0 % : τ_{32} and Q_w ($\sim m_W$)
  - 80.0 % : τ_{32} and $d_{23}^{1/2}$



Beyond Cut-based Taggers ATLAS-CONF-2017-064

- More sophisticated tagging techniques can be employed to make taggers which give a larger background rejection for a fixed signal efficiency, compared to the **smooth top-tagger**.
- Particularly promising performance from **DNN/BDT-based taggers** and the **shower deconstruction tagger**.
- These are brand new to ATLAS in 2017 !



Shower Deconstruction ATLAS-CONF-2017-064

arXiv:1211.3140

- Split the jet into subjets of four-momenta

$$\boxed{\{p\}_N = \{p_1, p_2, \ldots, p_N\}}$$

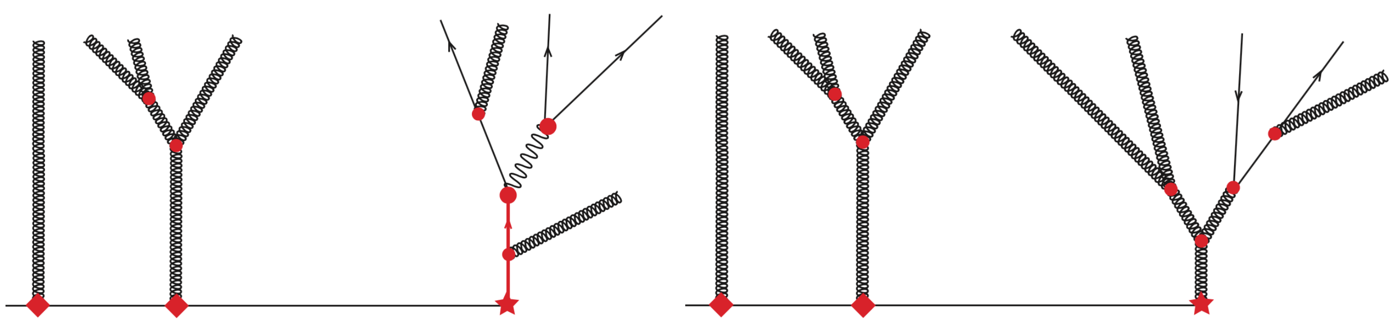
- Calculate the probabilities that a **simplified approximation to a shower Monte Carlo would generate $\{p\}_N$** according to separate signal and background hypotheses.

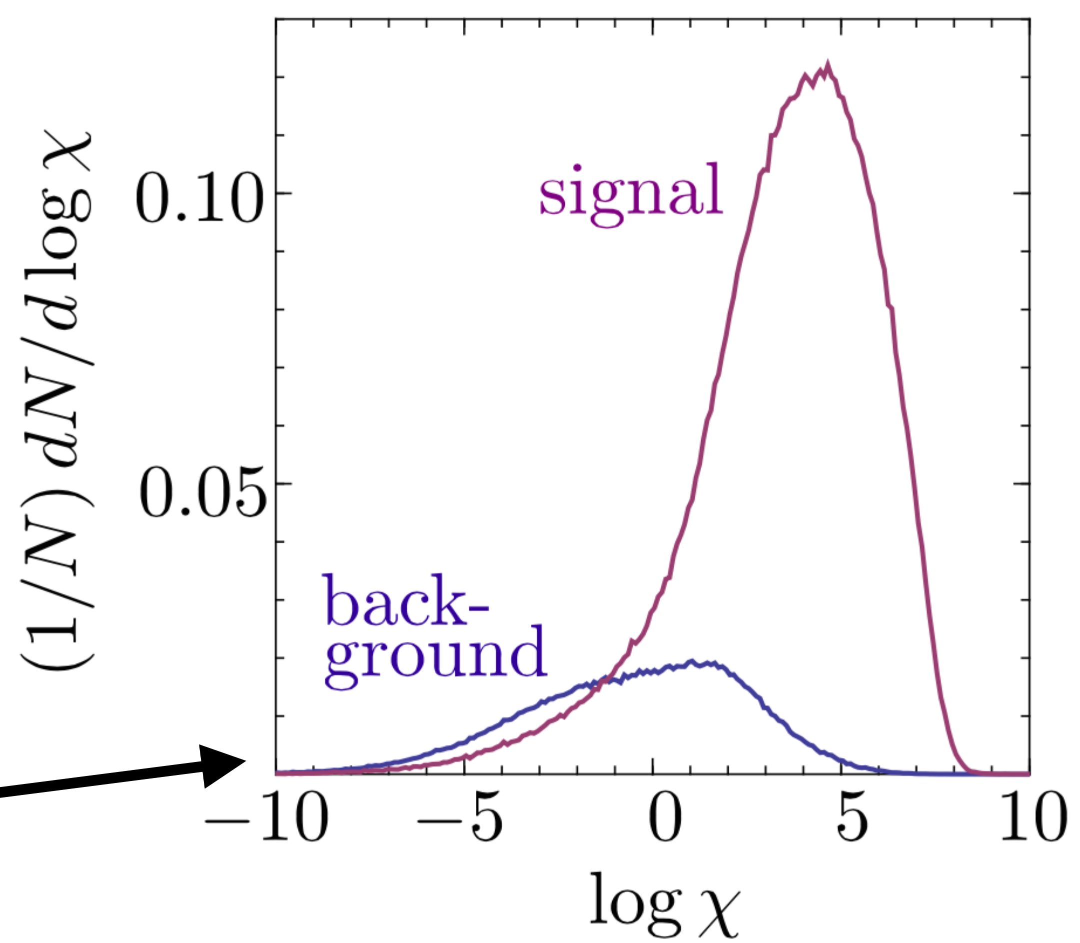
- Construct **likelihood ratio** that is large when the likelihood that the jet is a top is high. Sum of the **parton shower histories** of signal and background hypotheses.

$$\boxed{\chi(\{p\}_N) = \frac{\sum_{\text{histories}} P(\{p\}_N|S)}{\sum_{\text{histories}} P(\{p\}_N|B)}}$$

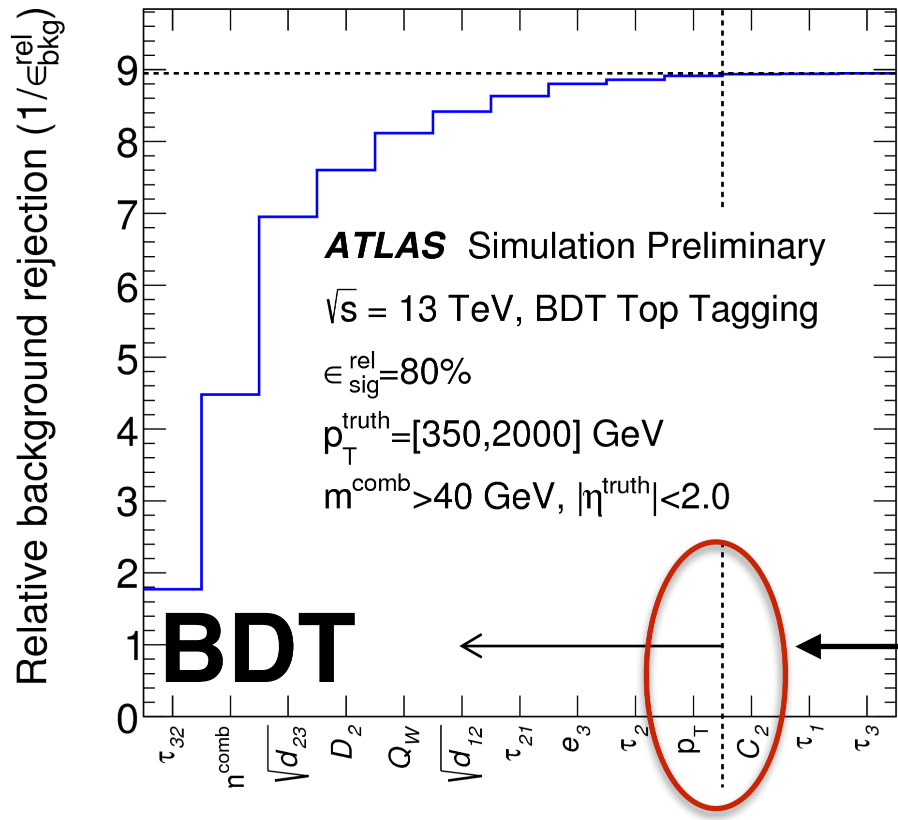
Top shower history

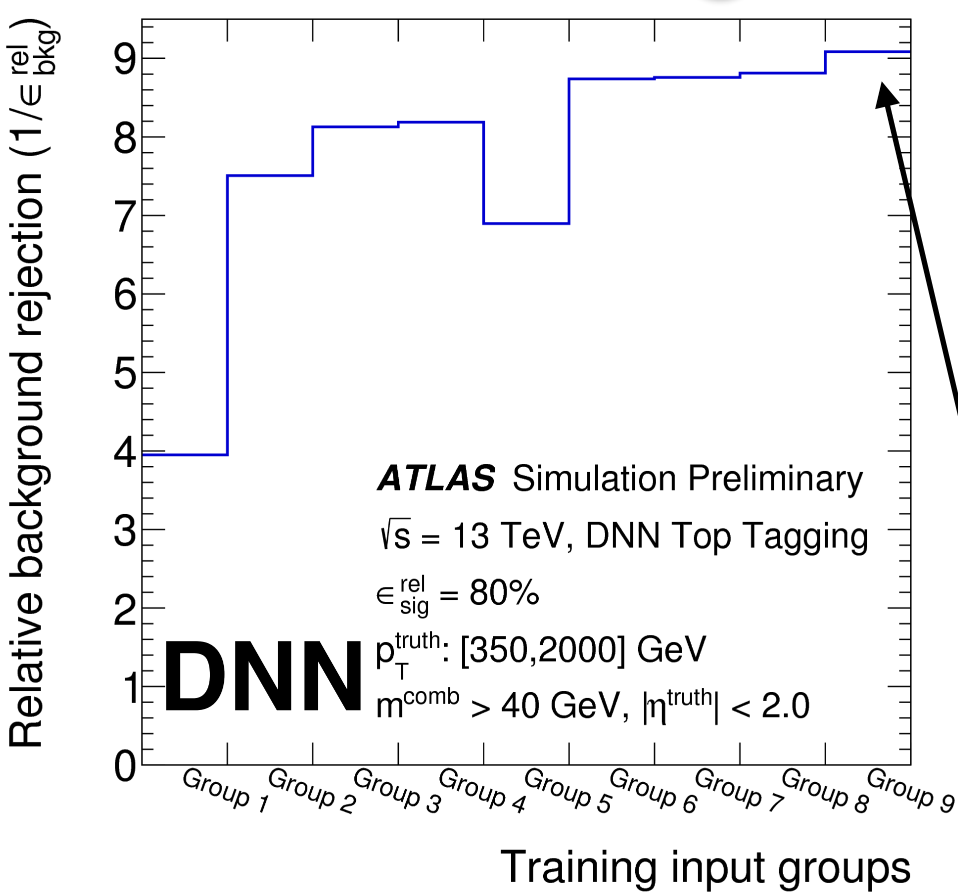
QCD shower history





ML Top-Taggers





ATLAS-CONF-2017-064

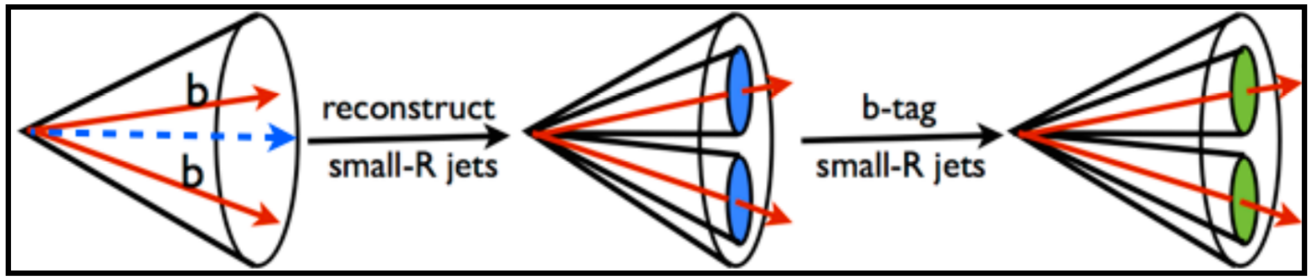
- Basic BDT strategy: single input variables which give the largest increase in performance are sequentially added to the network.
- **BDT**: At each step, the variable which gives the **greatest increase in relative background rejection**, for a fixed relative signal efficiency of 80.0 %, is retained until there is a **minimum number of variables** required to achieve the highest possible relative background rejection.
- **DNN**: Test with different input groups of variables. Performance of the DNN depends on both the **number of variables** and the **information content** in the group.

Group 1	$C_2, D_2, \tau_{21}, \tau_{32},$
Group 2	$C_2, D_2, \tau_{21}, \tau_{32}, m^{comb}$
Group 3	$C_2, D_2, \tau_{21}, \tau_{32}, m^{comb}, p_T$
Group 4	$\tau_1, \tau_2, \tau_3, e_3, m^{comb}, p_T$
Group 5	$C_2, D_2, \tau_{21}, \tau_{32}, \sqrt{d_{12}}, \sqrt{d_{23}}, Q_W$
Group 6	$C_2, D_2, \tau_{21}, \tau_{32}, \sqrt{d_{12}}, \sqrt{d_{23}}, Q_W, m^{comb}$
Group 7	$\tau_1, \tau_2, \tau_3, e_3, m^{comb}, p_T, \sqrt{d_{12}}, \sqrt{d_{23}}, Q_W$
Group 8	$C_2, D_2, \tau_{21}, \tau_{32}, \sqrt{d_{12}}, \sqrt{d_{23}}, Q_W, m^{comb}, p_T$
Group 9	$\tau_1, \tau_2, \tau_3, \tau_{21}, \tau_{32}, \sqrt{d_{12}}, \sqrt{d_{23}}, Q_W, C_2, D_2, e_3, m^{comb}, p_T$

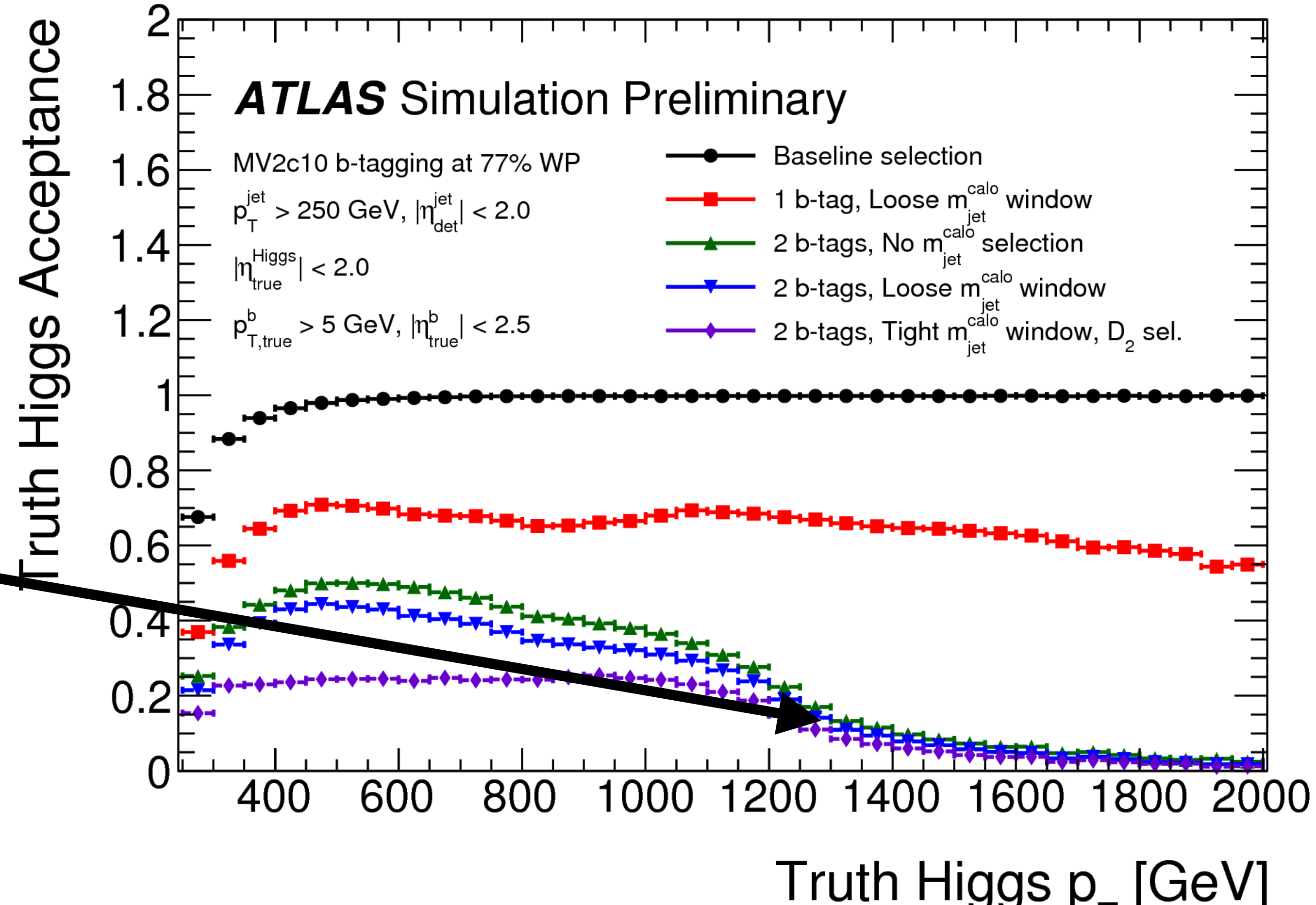
Baseline $H \rightarrow bb$ Tagger ATLAS-CONF-2016-039

UNIVERSITY OF OXFORD

- Reconstruct boosted Higgs decays using R = 1.0 trimmed jets.

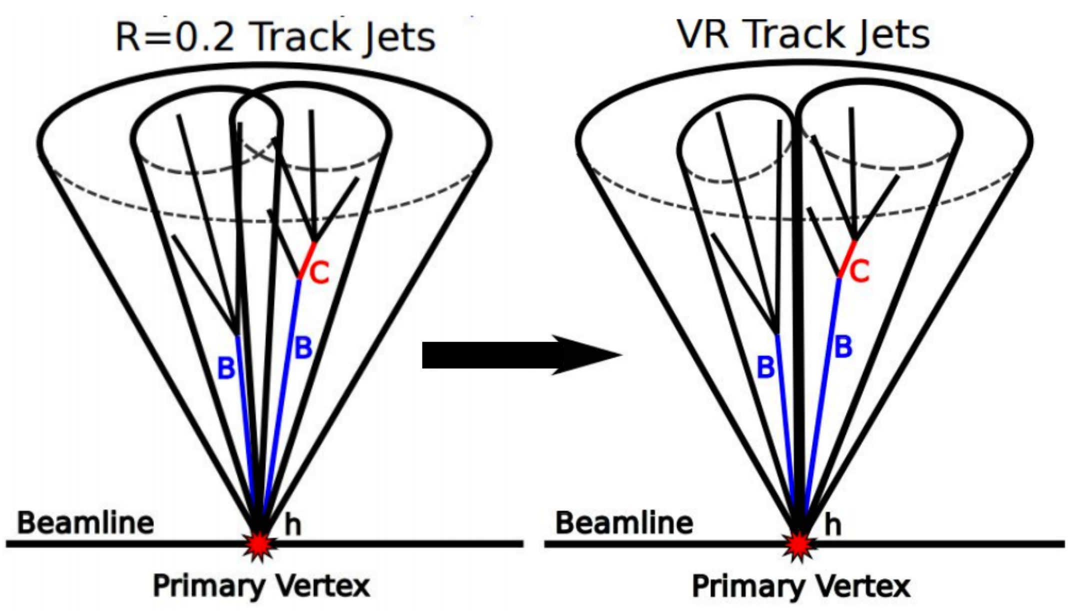


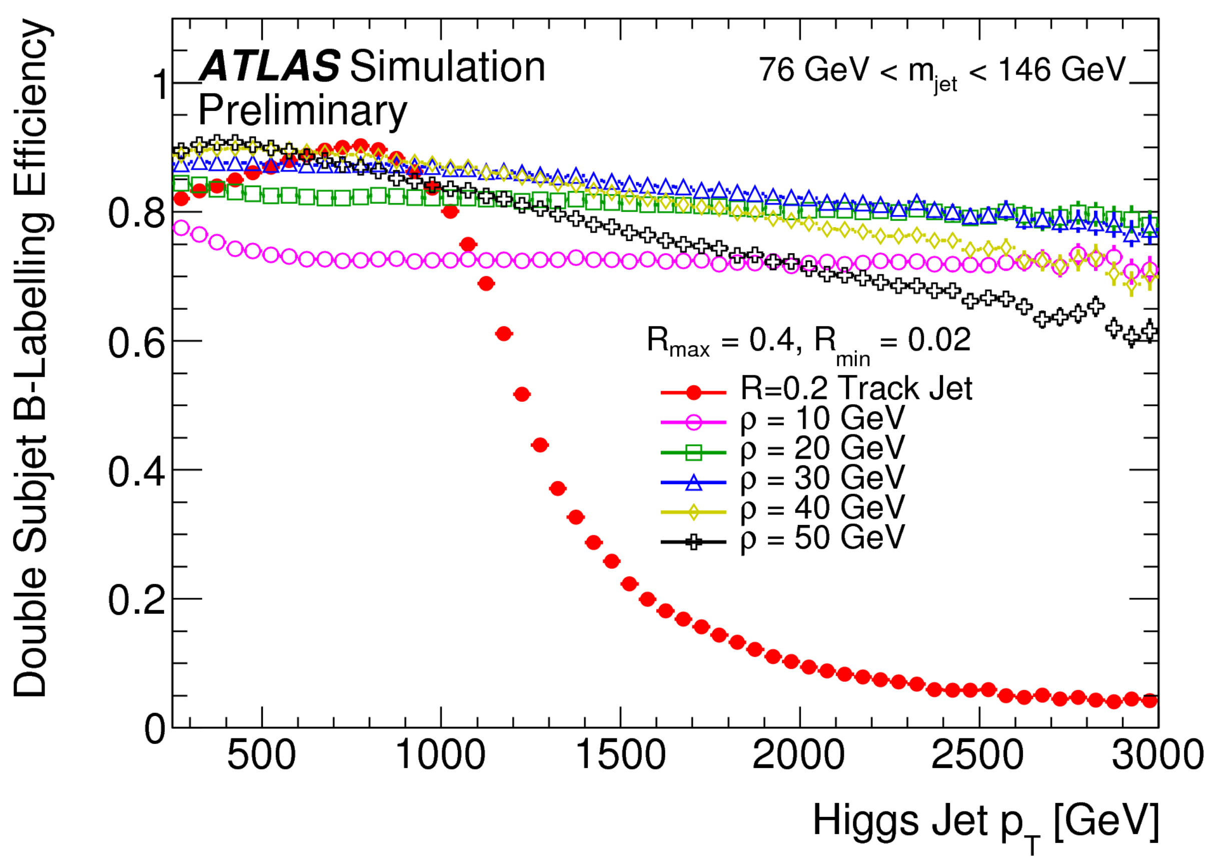
- Identify b-jets by matching R = 0.2 track-jets to the R = 1.0 calorimeter jet and using the **MV2c10** standard tagger ($w_{b\text{-tag}}$ of track-jet > w_X, typically using 70.0 % or 77.0 % efficiency working points).
- Different numbers of b-tags, with m^{calo} mass windows, and m^{calo} mass windows with a D_2 (2-prong) cut investigated.
- Requiring 2 b-tags **kills the acceptance** at much higher p_T. Why? …
- **Track-jet merging ! New approaches required …**



$X \longrightarrow bb$ Tagger: Variable-R Track Jets

"I've lost my track jets!"





ATLAS-PUB-2017-010

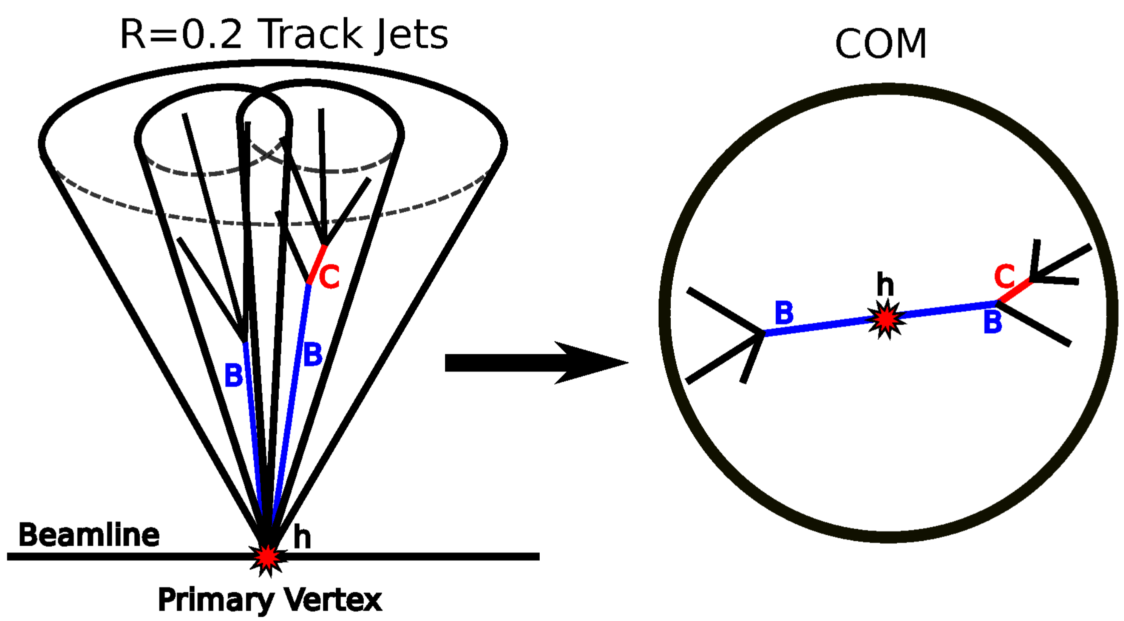
- Variable-R jet approach: build jets where the radius scales directly with **$1/p_T$** arXiv:0903.0392
- Build the subjets with a variable radius, R_{eff}, parametrised in the following way:

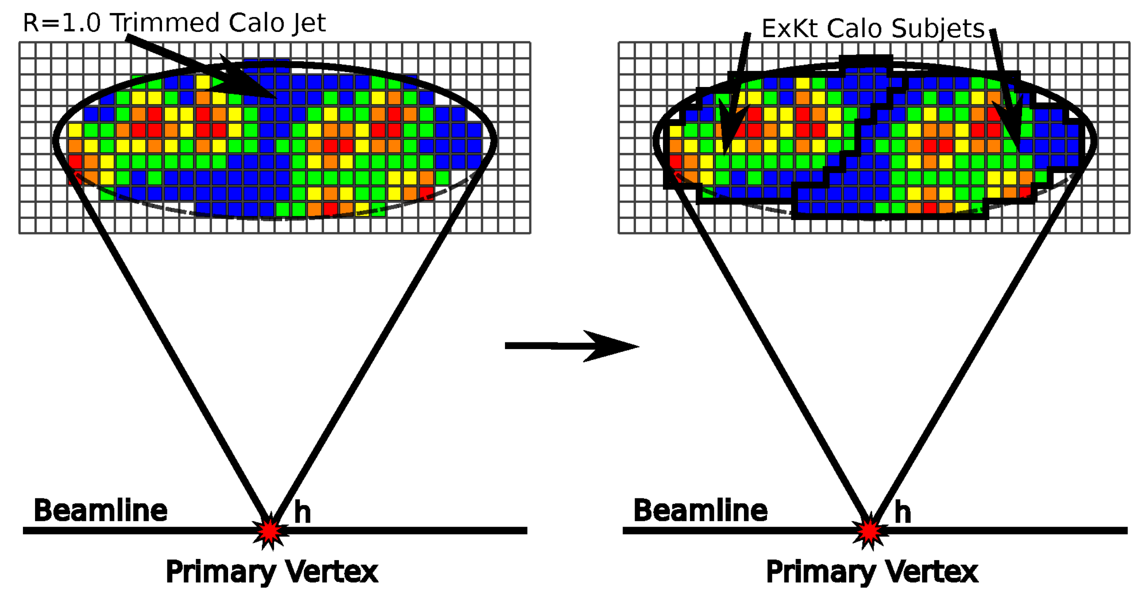
$$\boxed{R \longrightarrow R_{\mathrm{eff}}(p_{\mathrm{T}}) = \frac{\rho}{p_{\mathrm{T}}}}$$

ATLAS $H \longrightarrow bb$ optimisation:

- R_{min} = 0.2 (original track-jet radius)
- R_{max} = 0.4 (standard small-R jet radius)
- ρ = 30 GeV (dimensionful parameter)

$X \rightarrow bb$ Tagger: Exclusive k_t and COM Approaches ATLAS-PUB-2017-010





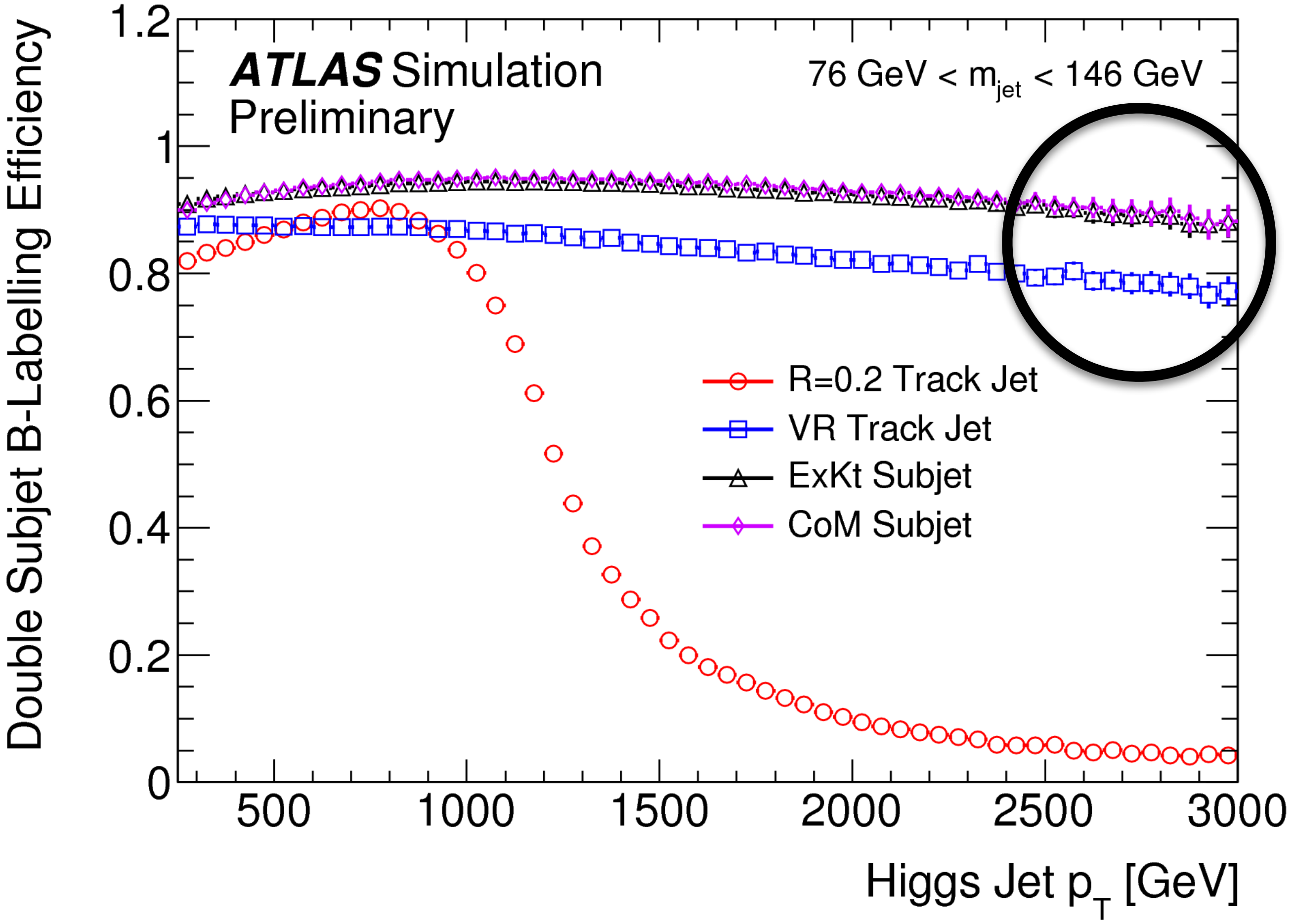
COM approach: boost the track-jets matched to the large-R jet **in the COM frame**, so that they are back-to-back. Measure the angular distances between tracks and subjets, and associate tracks to subjets. Finally, **boost back to the lab frame** and b-tag. Really intuitive and nice.

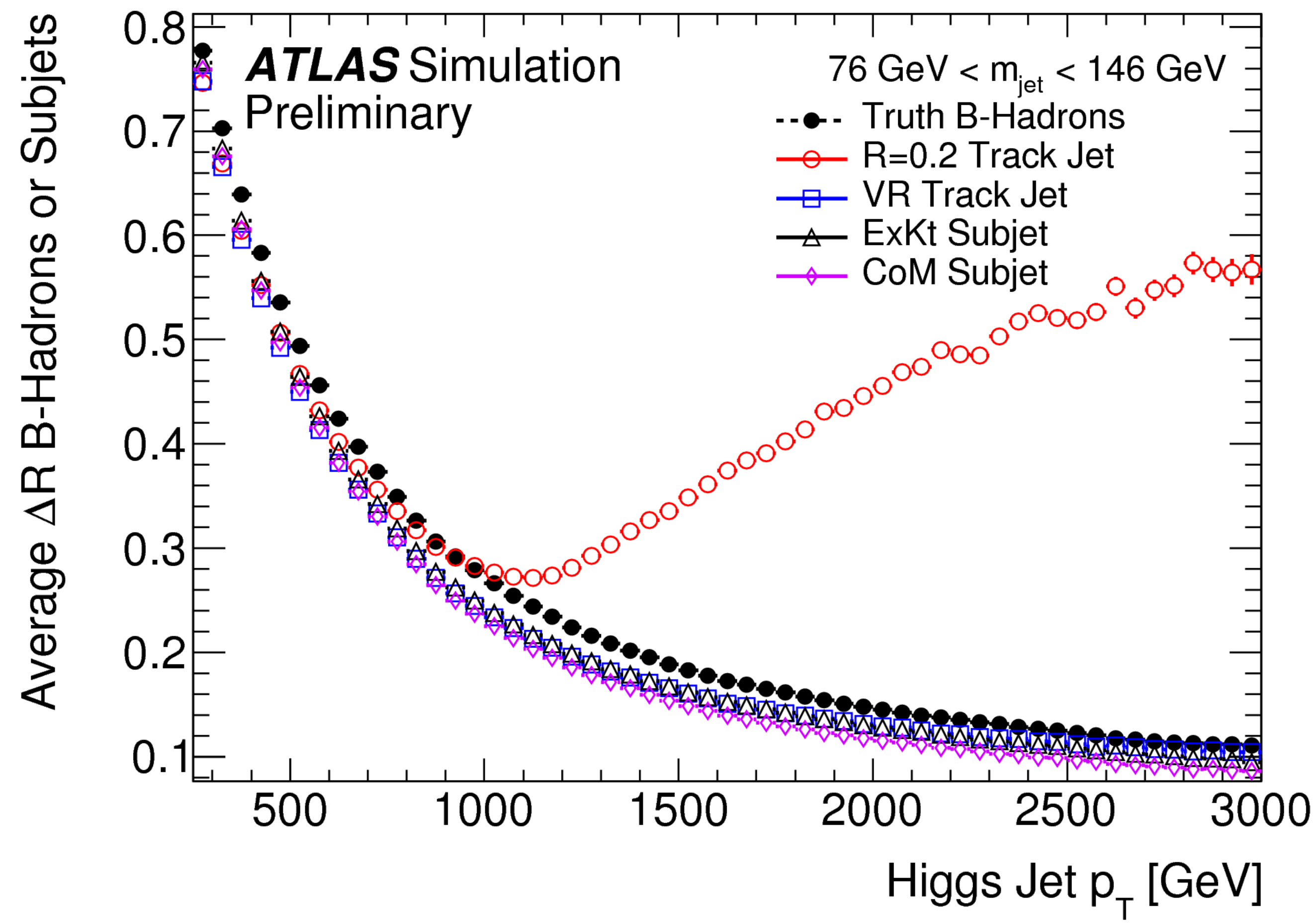
Exclusive k_t approach: undo the anti-k_t algorithm by clustering the $R = 1.0$ (trimmed, ungroomed, track-jet associated) jet calorimeter cluster constituents into **two subjets using the k_t algorithm**.

Problem with COM and exclusive k_t approaches: **dependence on jet topology**, making calibration (traditionally done using QCD dijets) potentially very difficult. Analysis feedback will be important here.

$X \longrightarrow bb$ Tagger: Putting Everything Together

ATLAS-PUB-2017-010





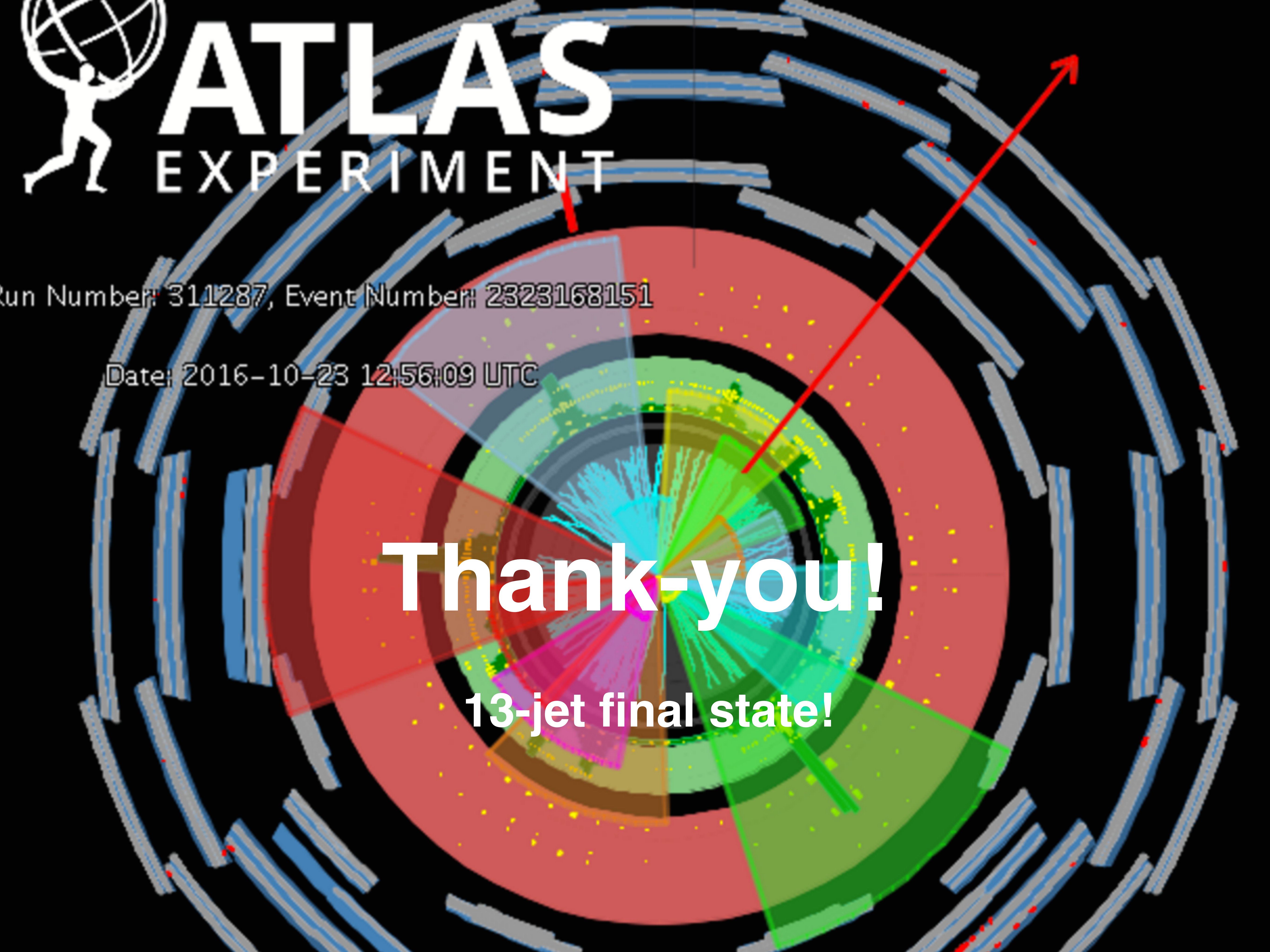
- Substantial improvement in the double *B*-labelling efficiency using the new *X*(*bb*) methods.
- Largest improvement from the **COM** and **exclusive k_t** approaches. **VR** also highly efficient.
- New methods also scale with $1/p_T$, as expected.

Where do we go from here ?

Where do we go from here?

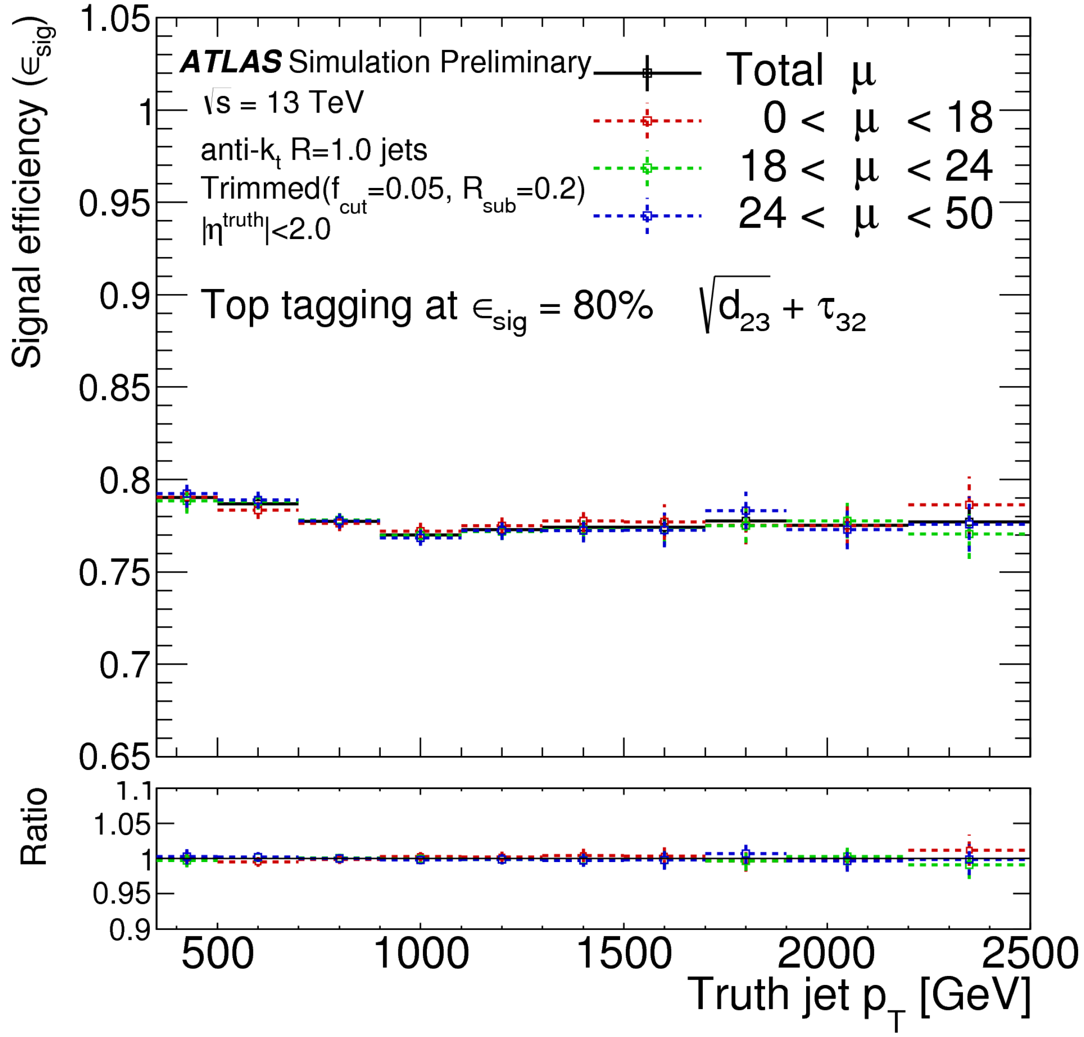
UNIVERSITY OF OXFORD

- **Jet substructure techniques** provide a natural framework for developing powerful boosted object taggers — focused today on boosted top and Higgs.
- More sophisticated **machine learning techniques** are entering mainstream **heavy-flavour** tagging in ATLAS.
- Machine learning approaches offer potential **performance improvements** over traditional cut-based taggers.
- New approaches to **tagging track-jets** (VR, COM, exclusive k_t) are greatly enhancing the performance of tagging Higgs and heavy bosons at **much higher p_T** than before. Can now do better than standard double *B*-taggers with a mass window.
- Need more analysis feedback towards the end of the Run-2 for these taggers.
- Success and new sensitivity to the **heavy-flavour, boosted physics** regime beckons ! ?


ATLAS
EXPERIMENT
Run Number: 311287, Event Number: 2323168151
Date: 2016-10-23 12:56:09 UTC
Thank-you!
13-jet final state!

Back-up

ML Top-Tagging



- **[BoostedJetTaggers code](#)**, developed by JSS, has been updated for Moriond, 2017.

- Two-variable optimisation procedure: for each two-variable combination, find a set of p_T-dependent cuts that maximises background rejection at a fixed working point. Fitting these cuts then defines the tagger, which is smooth in p_T.

- Scan over many substructure variables (*N*-subjettiness, calorimeter/combined/track-assisted jet mass, energy correlation functions, …) and find the combination of two giving the best background rejection, for a fixed tagging efficient (50 % and 80 % WP) -> **recommended taggers**.

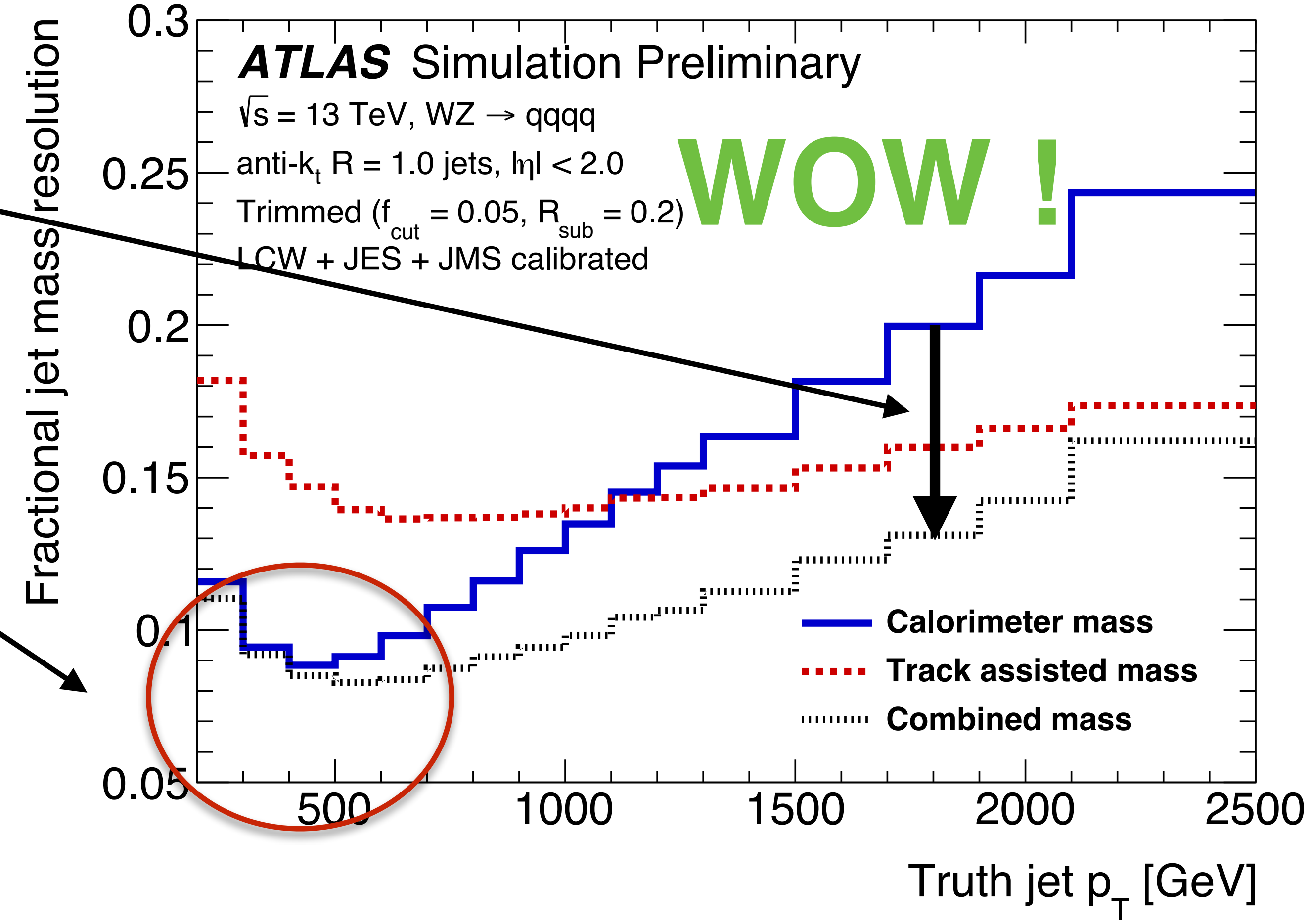
(more later)

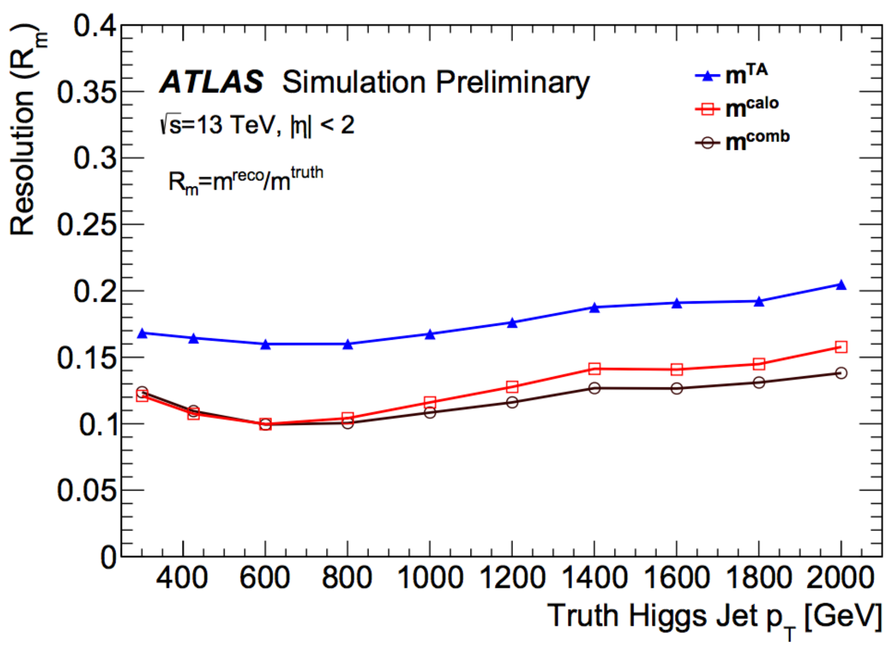
Jet Mass Scale and Jet Mass Resolution

- Combined jet mass: take a linear combination of the calorimeter and track-assisted jet mass, weighted such that the jet mass resolution is minimal across jet p_T.

- **Major improvement for exotics**: at large boost, **track-component of resolution dominates** (finite granularity of the calorimeter) => significant improvement from combined mass.

- **SUSY**: high-multiplicity, moderately boosted final states means that **standard calorimeter mass** gives competitive resolution performance with the combined mass.





- Higgs example: combined mass resolution boosted $H \rightarrow bb$ decays captured with trimmed $R = 1.0$ jets. Nice improvements, particularly in the boosted regime!

Quark/gluon Tagging in Exotics

- Good progress has been made with quark/gluon tagging since Run-1. New quark/gluon tagger based on the charged-particle multiplicity of the jets.

- Basic idea: gluon radiation off a gluon adds a C_A factor to the A-P splitting function, and C_F for gluon radiation off a quark. C_A/C_F = 9/4 ~ 2 => gluon jets have more constituents than quark jets, with a broader radiation pattern. **Discrimination** using $\langle n_{charged} \rangle$ natural and intuitive. Run-1 implementation below.