Boosted hadronic object identification using jet substructure in ATLAS Run-2

Emma Winkels

on behalf of the ATLAS collaboration

HEPMAD18

UNIVERSITY OF SUSSEX

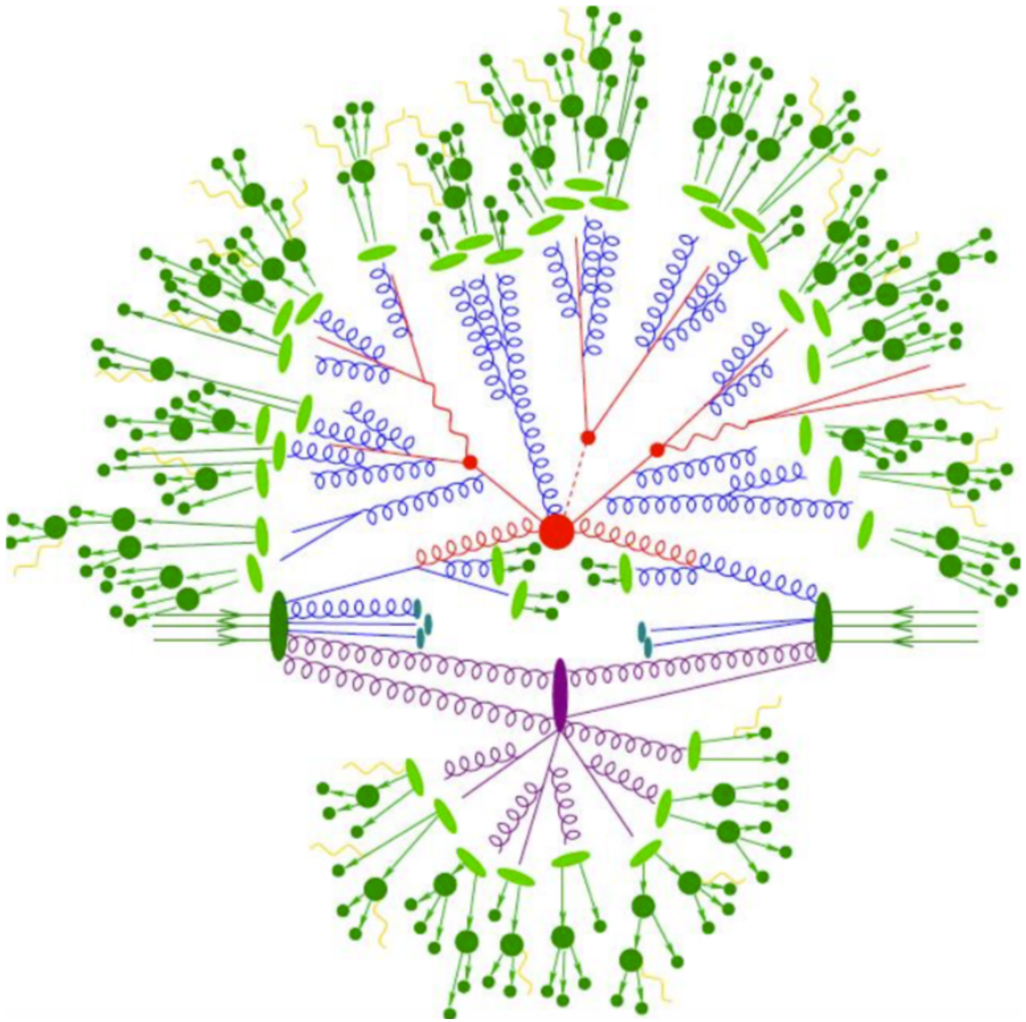
ATLAS EXPERIMENT

Outline

- Jets and jet substructure
- Top and W tagging
- $H \rightarrow bb$ tagging
- Mass-decorrelated taggers
- Summary

Jets

What is a jet?



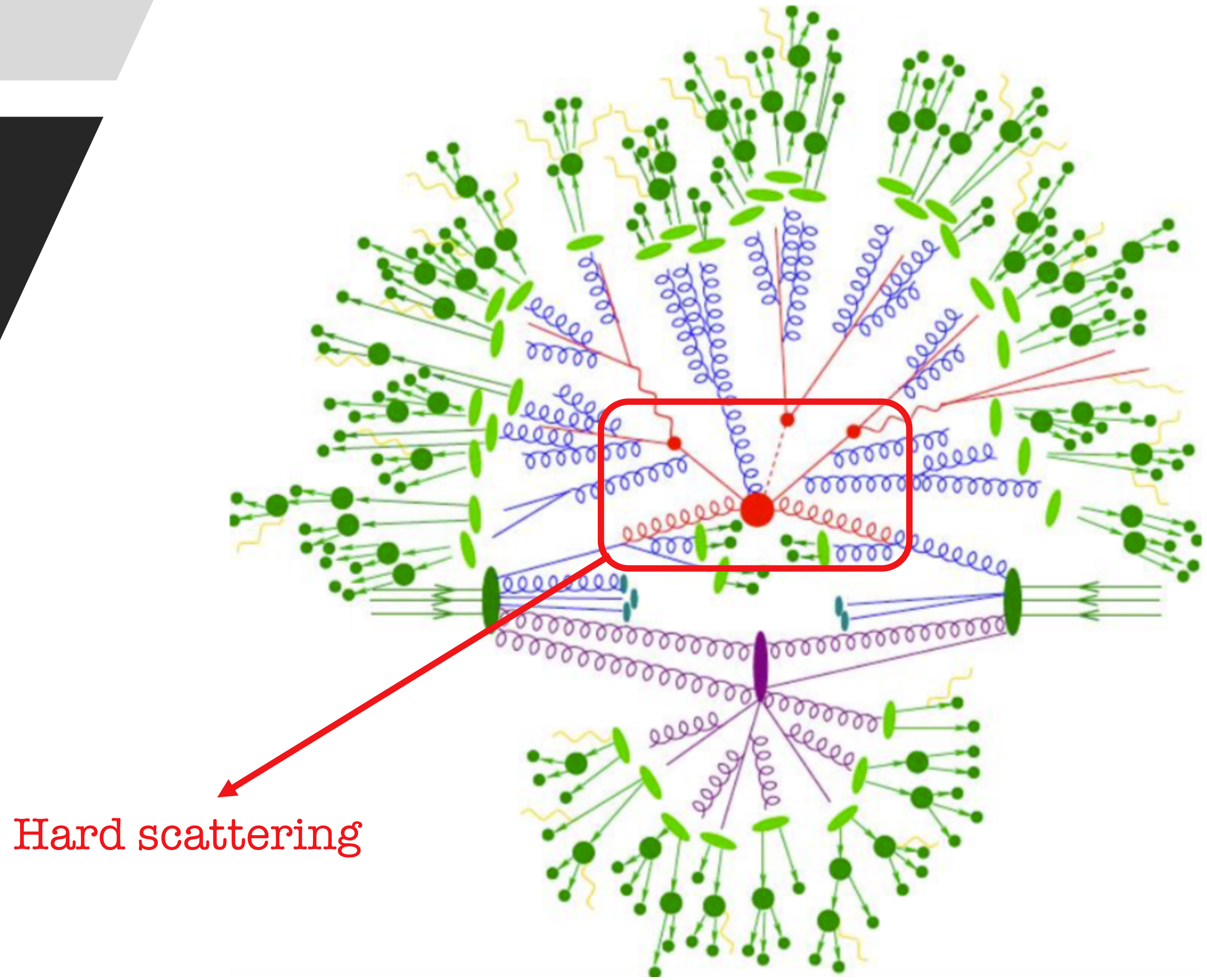
Jets are objects constructed from the energy deposits left by collimated sprays of particles.

Attempt to group inputs from common sources together.

What is a jet?

Jets are objects constructed from the energy deposits left by collimated sprays of particles.

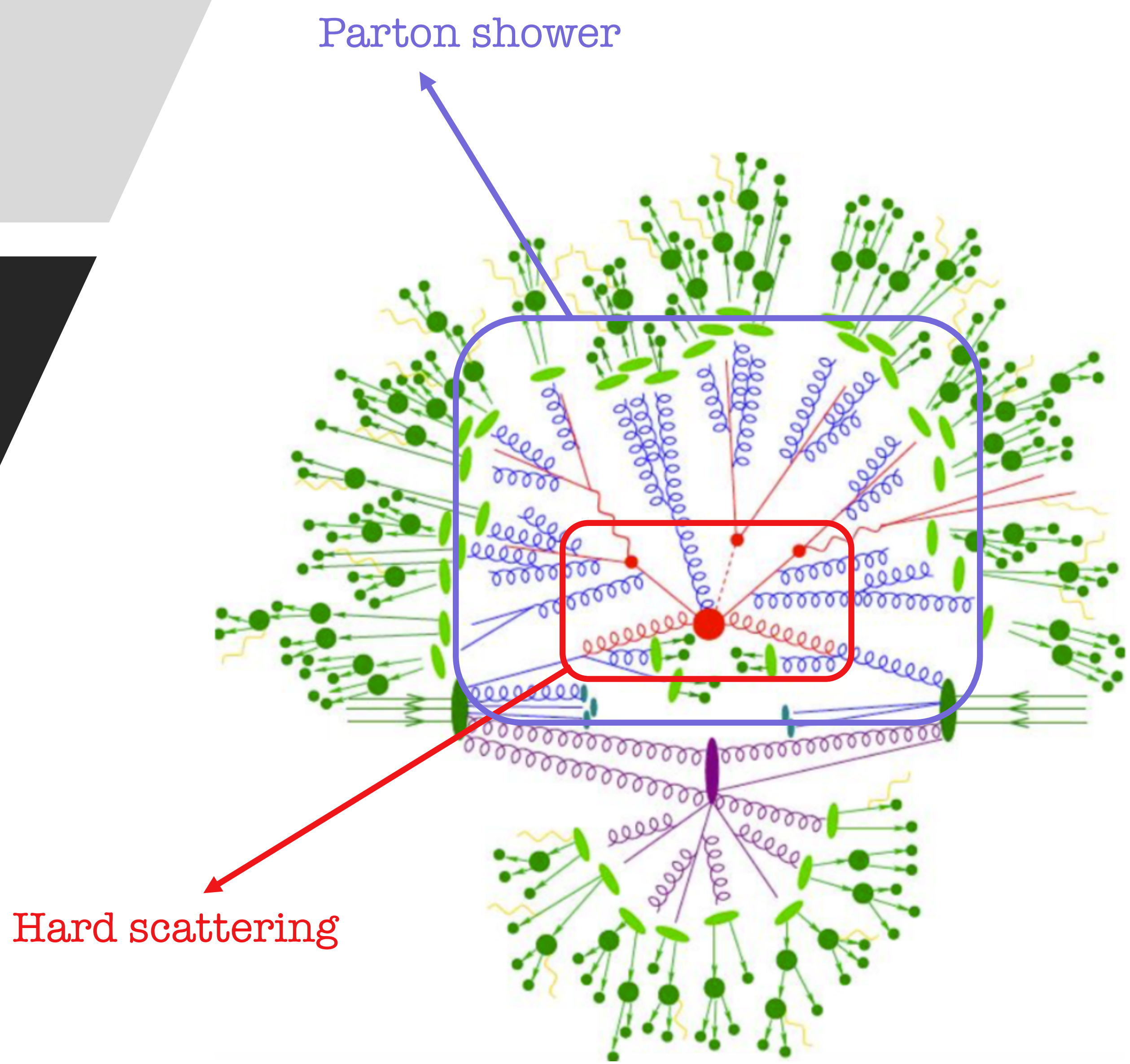
Attempt to group inputs from common sources together.



What is a jet?

Jets are objects constructed from the energy deposits left by collimated sprays of particles.

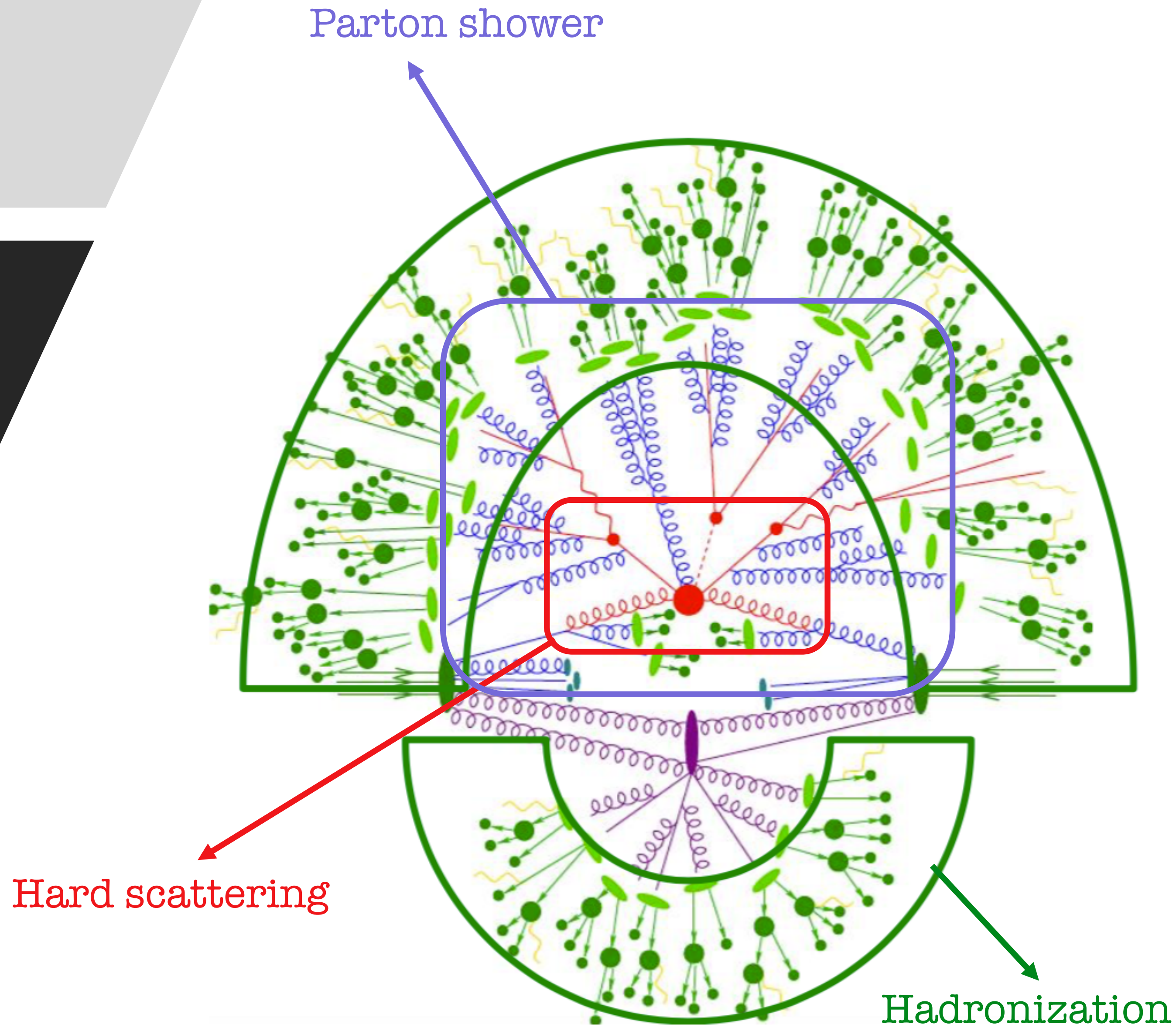
Attempt to group inputs from common sources together.

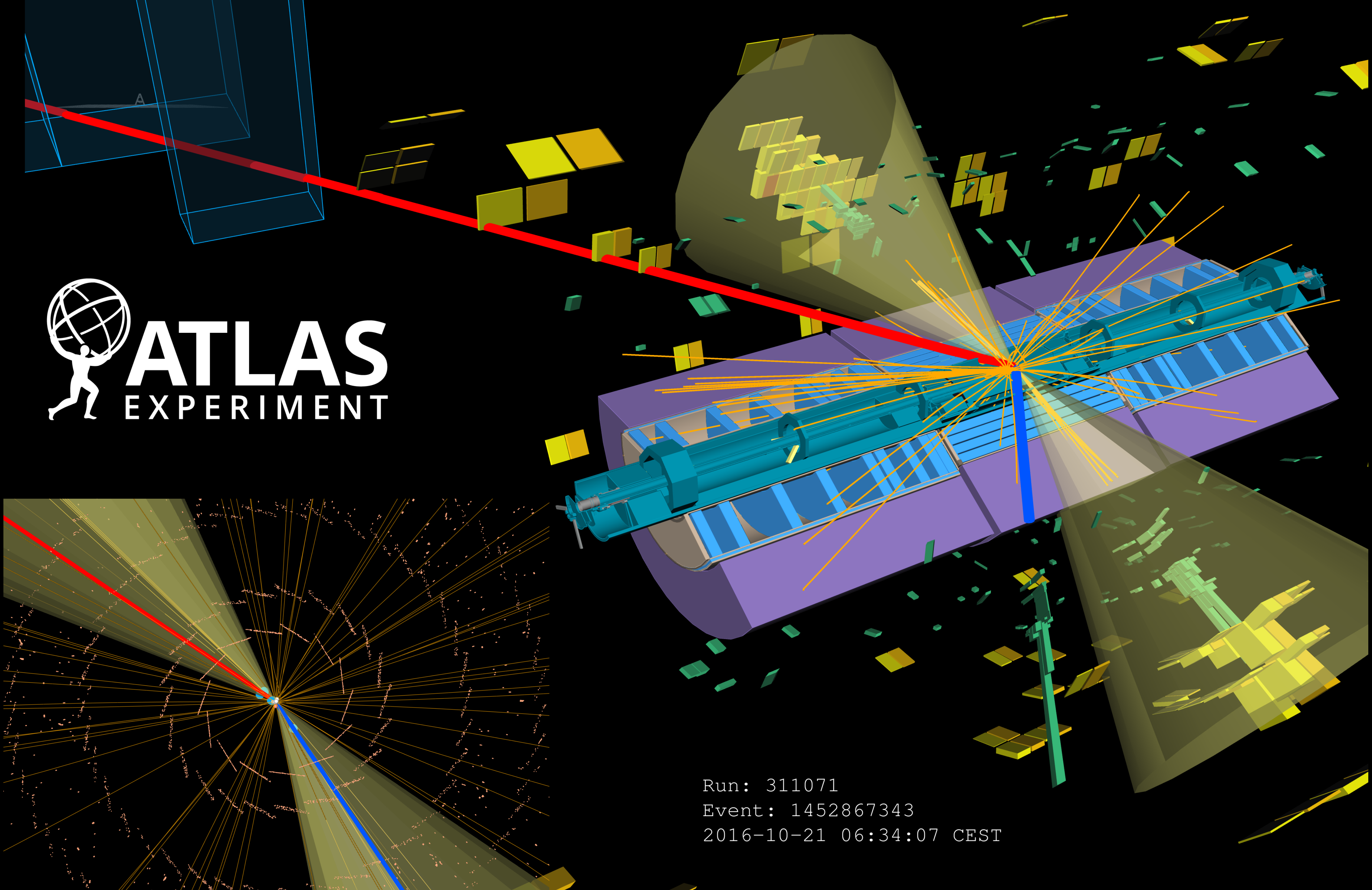


What is a jet?

Jets are objects constructed from the energy deposits left by collimated sprays of particles.

Attempt to group inputs from common sources together.

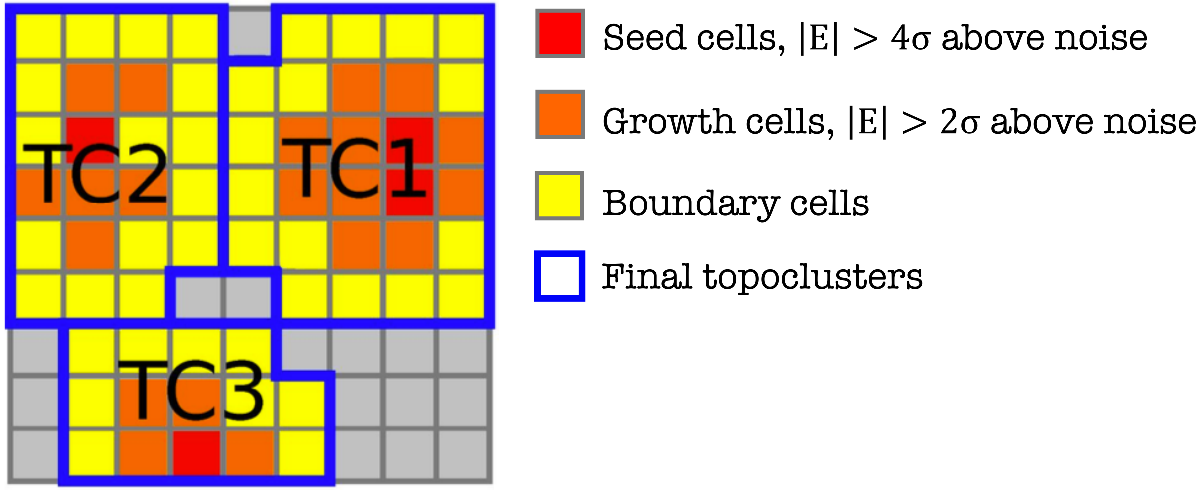


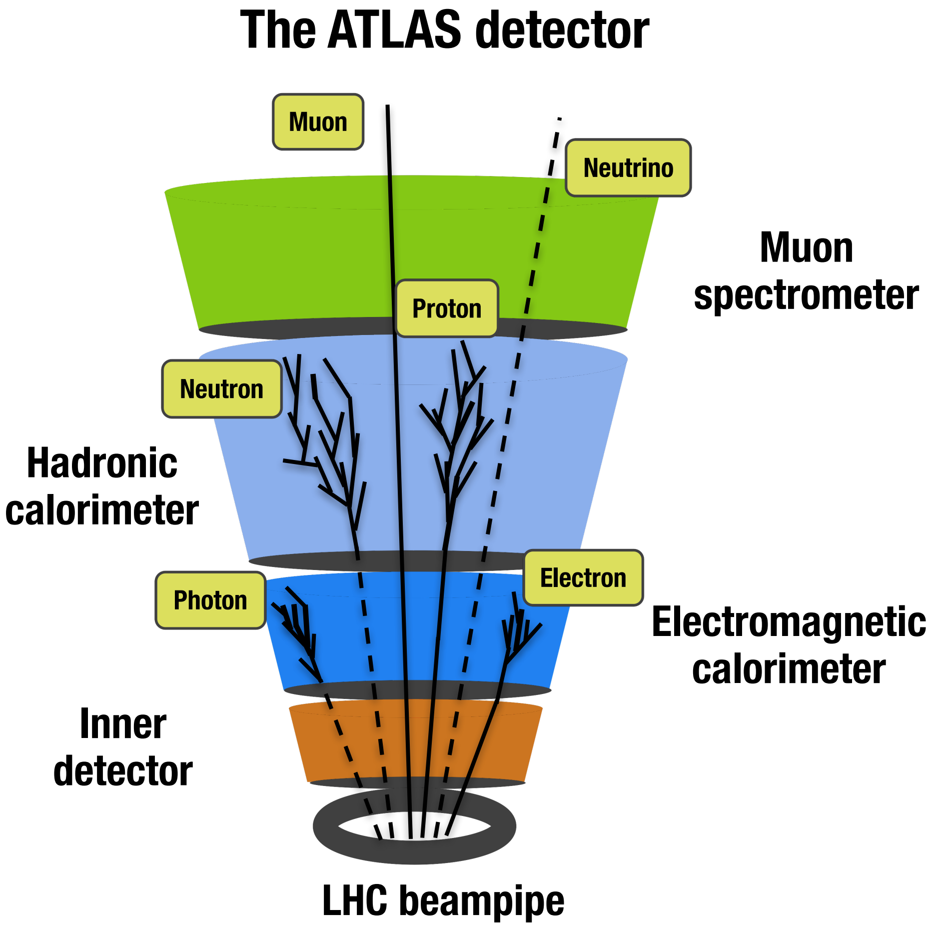

ATLAS
EXPERIMENT
Run: 311071
Event: 1452867343
2016-10-21 06:34:07 CEST

Inputs to jets

ATLAS uses calorimeter objects and tracks as jet inputs. Calorimeter measures energy of particles.

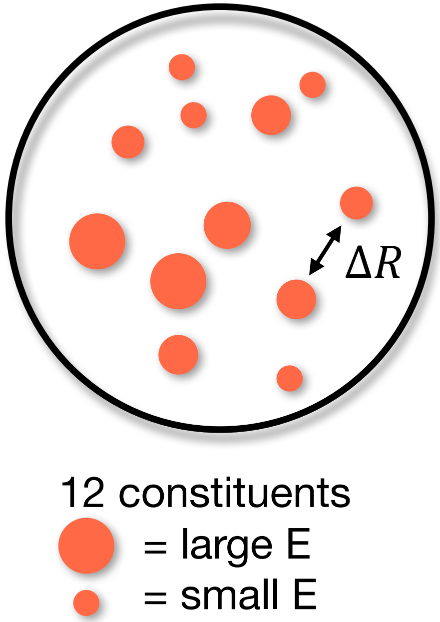
Starts with topological clustering of calorimeter cells.





How to define a jet

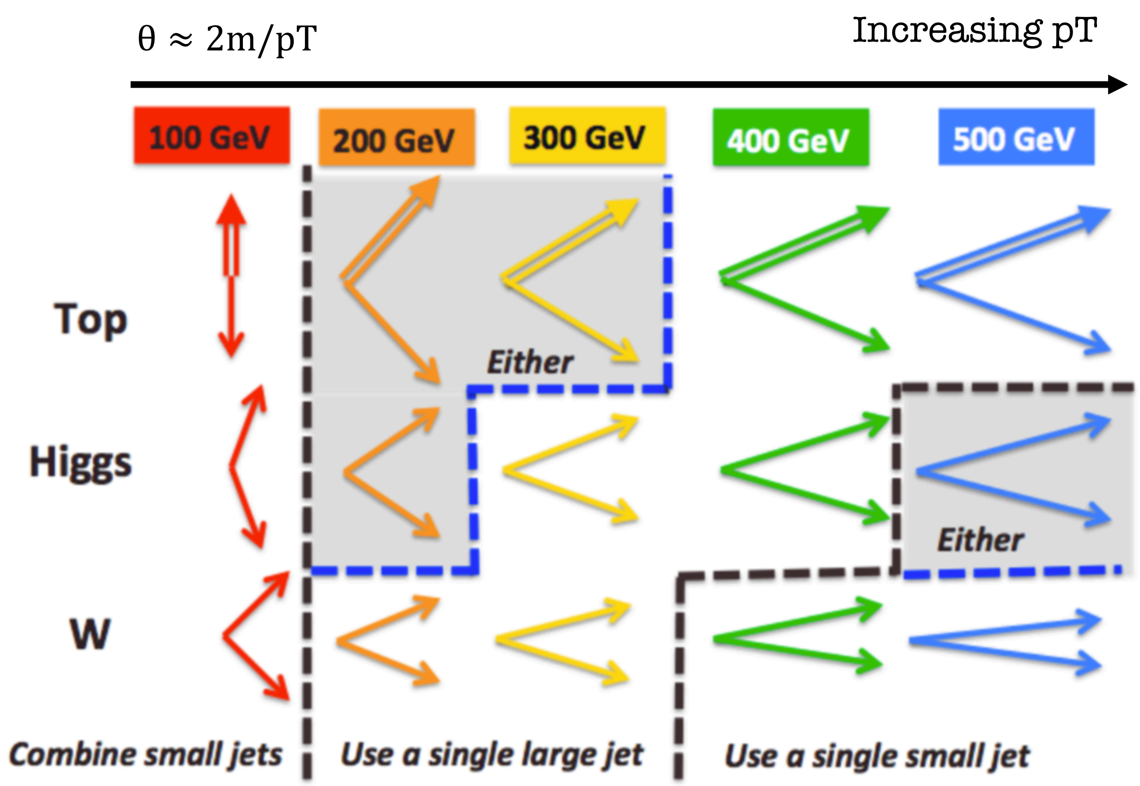
- There is no unique way to define a jet
- Different jet algorithms to cluster energy constituents into a jet:
  - Anti-k_T: cluster hard (high-p_T) and close (small ΔR) energy deposits first
  - k_T: cluster soft (low-p_T) and close
  - Cambridge/Aachen: cluster close



Boosted jets

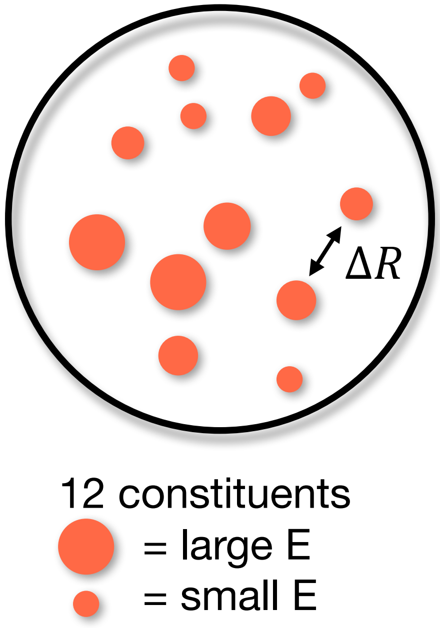
Different jet radii (R) for different purposes:

- Small jets for quarks and gluons
- Large jets for hadronic decays of W, Z, H, top..



Jet substructure

- Access the inner structure of large jets*
- Jet substructure variables are some function of
  - Number of constituents
  - Energy of the constituents
  - Angular separation of the constituents (ΔR)
- Jet substructure helps us in jet tagging



* We use jet grooming to cut away the soft parts of jets, see 1510.05821

Jet tagging

- Identify the particle that produced the jet.
- Used in broad range of physics analyses.
    - Analyses looking at boosted top/Higgs/W: distinguish these large jets from quark/gluon jets
    - Analyses with *b*-hadrons or *c*-hadrons in the final state: heavy flavour tagging (*b*-, *c*-quark jets)

Top/W tagging

ATLAS-CONF-2017-064

Commissioning of a tagger

ATLAS process from idea to tagger used for physics analyses:

Use MC to choose tagger

Two-variable tagging

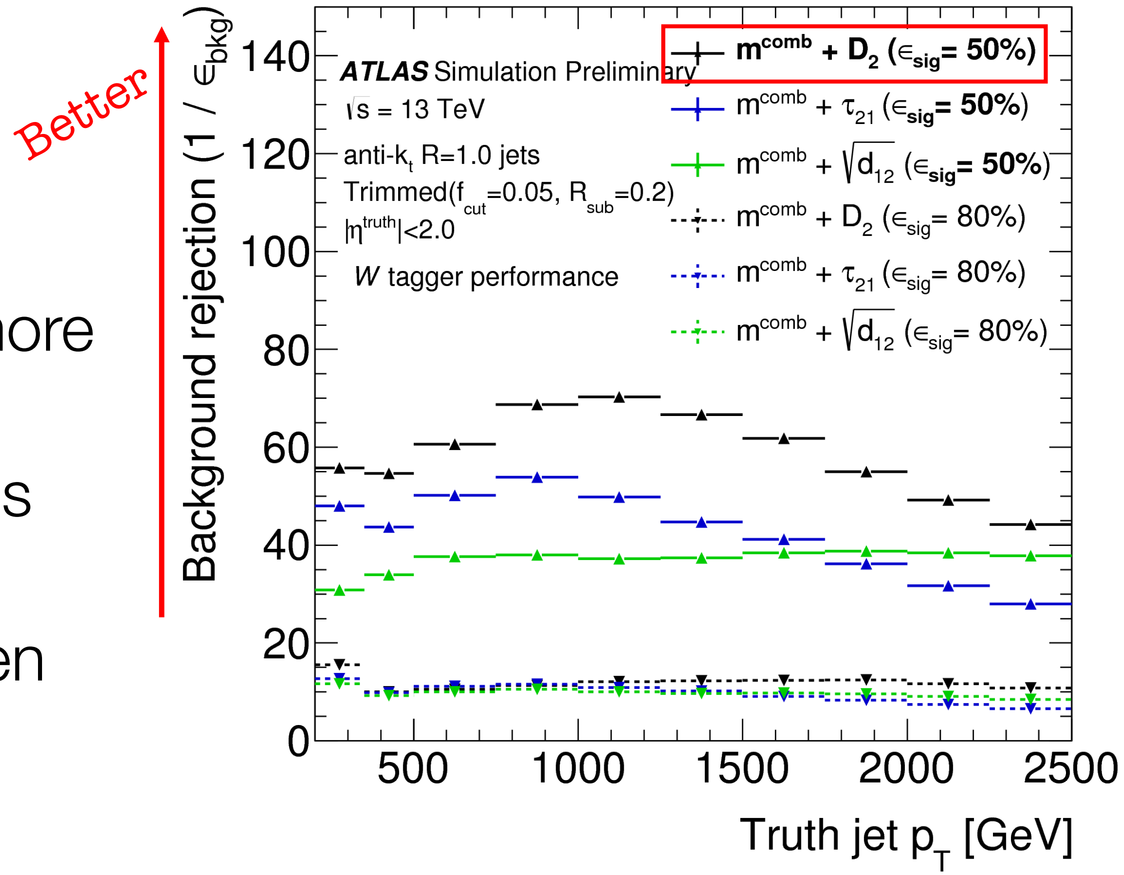
- Simple cut-based tagging works well:
  - **W**: $\mathrm{m}^{\mathrm{comb}} + \mathrm{D}_2$
  - **Top**: $\mathrm{m}^{\mathrm{comb}} + \tau_{32}$
- $\mathrm{m}^{\mathrm{comb}}$: Combined mass, combines calorimeter clusters with tracks to give more stable mass performance at high-p_T.
- D_2 : Energy correlation ratio, distinguishes between one-prong and two-prong jets.
- τ_{32}: N-subjettiness, distinguishes between two-prong and three-prong jets.

Use MC to choose tagger

Two-variable tagging

- Simple cut-based tagging works well:
  - **W**: m^{comb} + D_2
  - **Top**: m^{comb} + τ_{32}
- m^{comb}: Combined mass, combines calorimeter clusters with tracks to give more stable mass performance at high-p_T.
- D_2 : Energy correlation ratio, distinguishes between one-prong and two-prong jets.
- τ_{32}: N-subjettiness, distinguishes between two-prong and three-prong jets.

W tagging performance



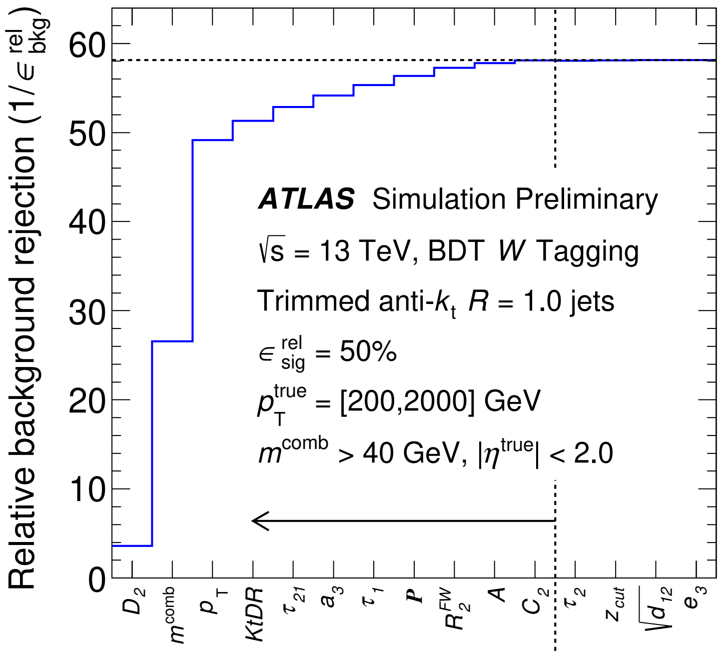
ATLAS-CONF-2017-064

Use MC to choose tagger

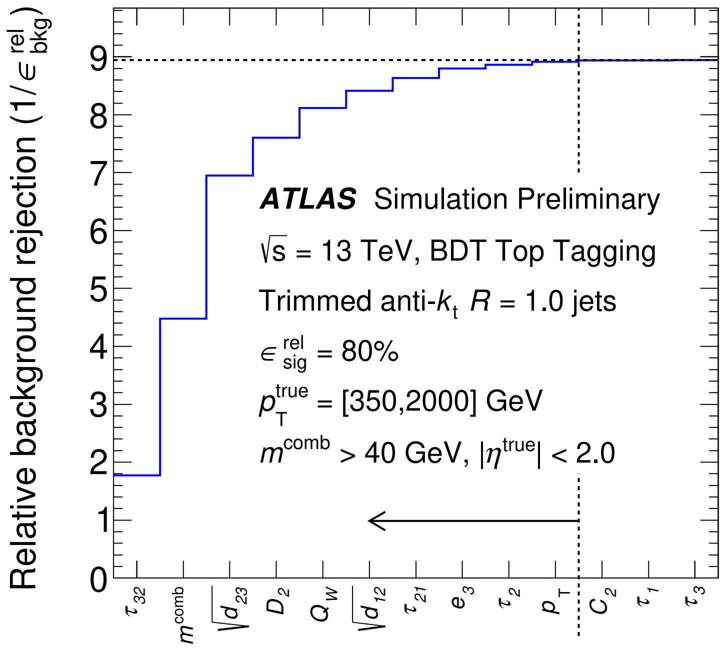
Machine learning taggers - BDT

- Machine learning techniques allow for the use of multiple variables
- Boosted decision tree (BDT): sequentially adding variables improves classification over two-variable tagging

W tagging



Top tagging



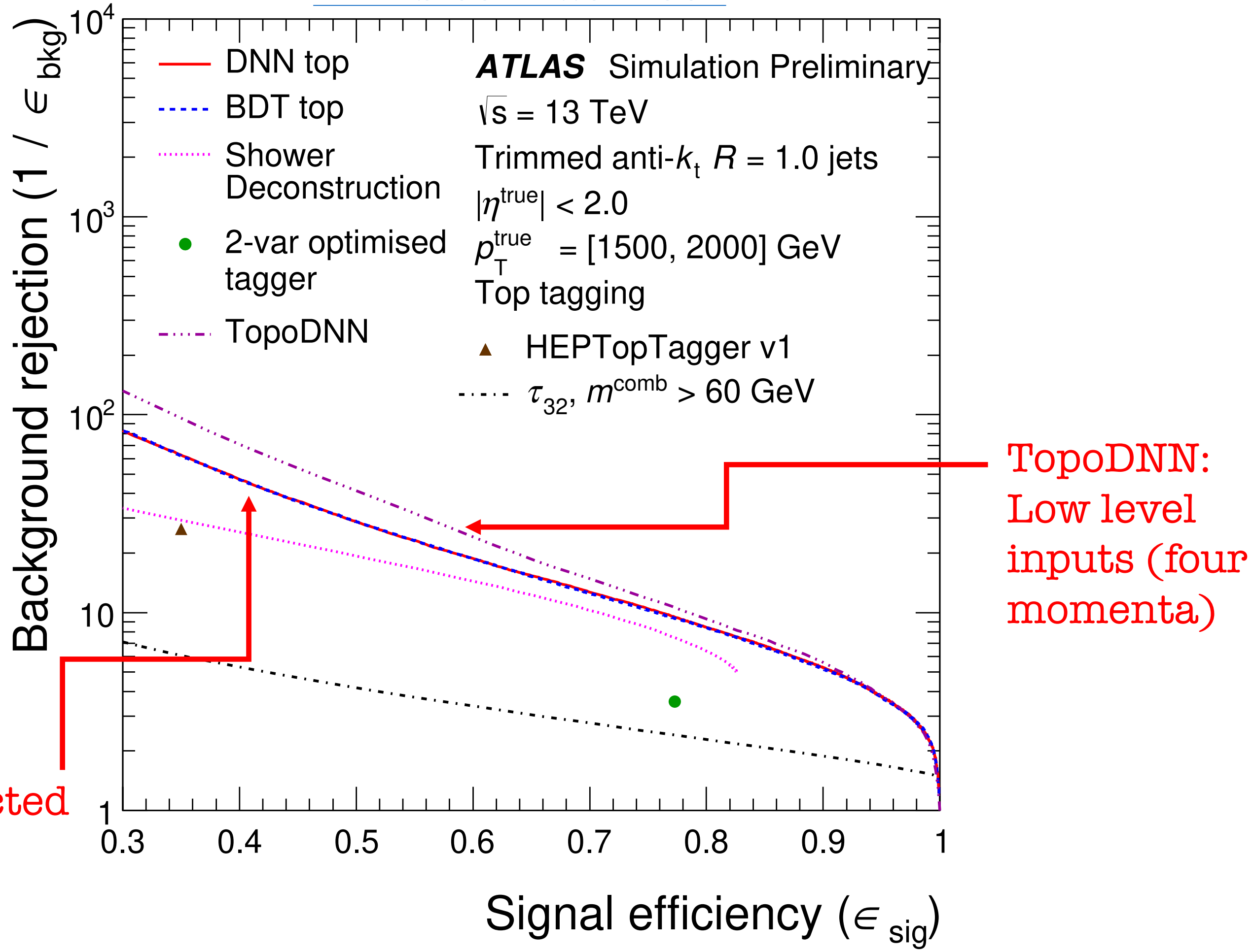
ATLAS-CONF-2017-064

Use MC to choose tagger

Machine learning taggers - DNN

- ATLAS evaluated TopoDNN* top tagger.
- Uses deep neural network (DNN) with topocluster jet constituents as inputs.
- Performance is better with low-level inputs than standard machine learning taggers.

ATLAS-CONF-2017-064



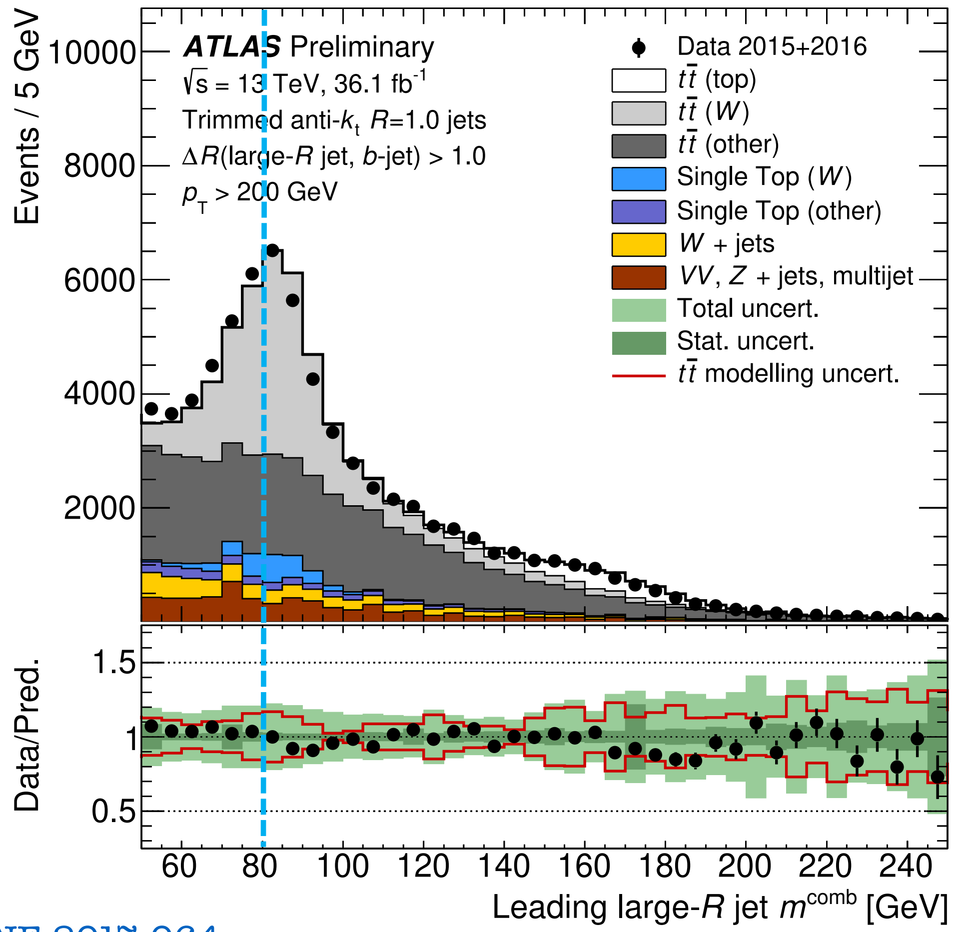
* More details on TopoDNN: 1704.02124.

Use MC to choose tagger

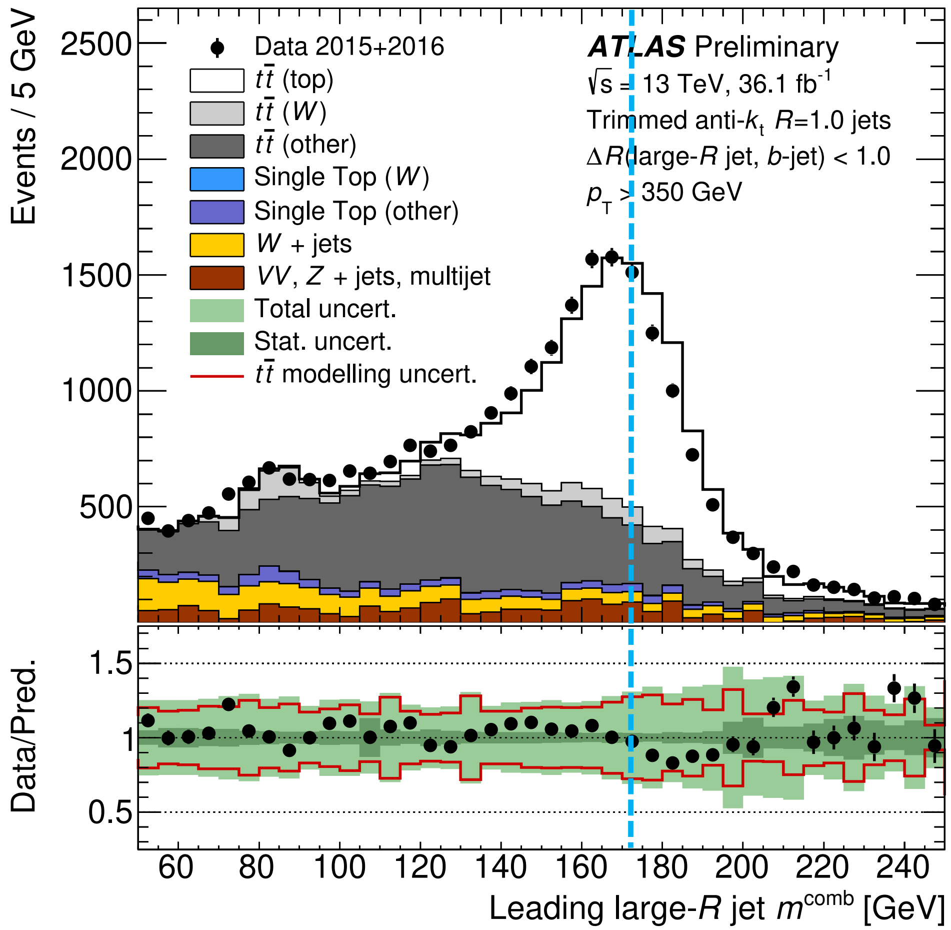
Check data/MC

Measurements in data

W enriched sample



Top enriched sample



ATLAS-CONF-2017-064

W/top tagging efficiency

Need to measure efficiency in data and get uncertainty on this efficiency*.

- Full ATLAS 2015-2016 dataset of $36.1 - 36.7\ \mathrm{fb}^{-1}$
- Measure top/W tagging efficiency in $t\bar{t}$ lepton+jets samples.
- Measure multijet rejection in dijet and γ +jets samples.

* Also done in V+jets: [ATLAS-CONF-2018-016]

W/top tagging efficiency vs. pile-up

Pile-up is the resulting signal in the detector from other interactions besides the hard scatter we want to look at. Expressed as the mean number of interactions per bunch crossing.

Use MC to choose tagger → Check data/MC → Measure efficiencies in data & MC → Calculate uncertainties on efficiencies

W/top tagging efficiency vs. pile-up

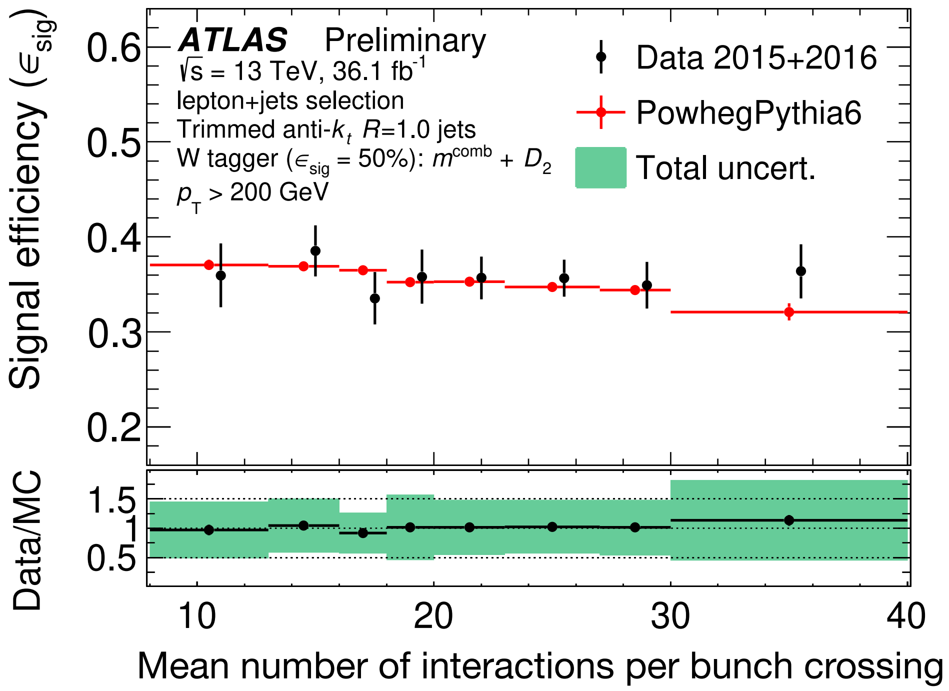
pre-fit

$$\epsilon_{\mathrm{MC}} = \frac{N_{\mathrm{signal}}^{\mathrm{tagged}}}{N_{\mathrm{signal}}^{\mathrm{tagged}} + N_{\mathrm{signal}}^{\mathrm{not\ tagged}}}$$

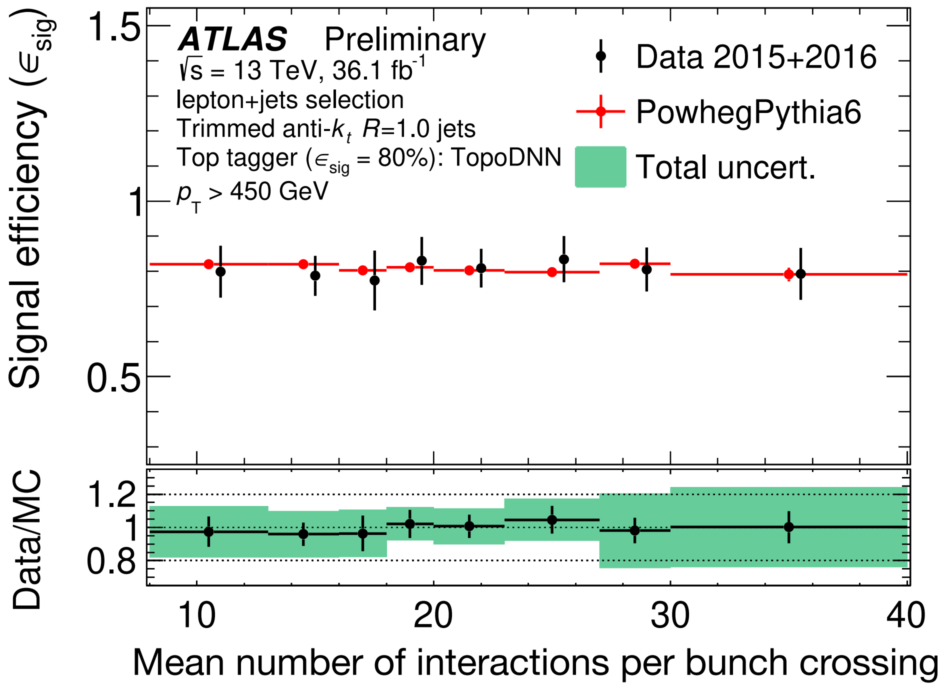
$$\epsilon_{\mathrm{data}} = \frac{N_{\mathrm{fitted\ signal}}^{\mathrm{tagged}}}{N_{\mathrm{fitted\ signal}}^{\mathrm{tagged}} + N_{\mathrm{fitted\ signal}}^{\mathrm{not\ tagged}}}$$

post-fit

W tagger



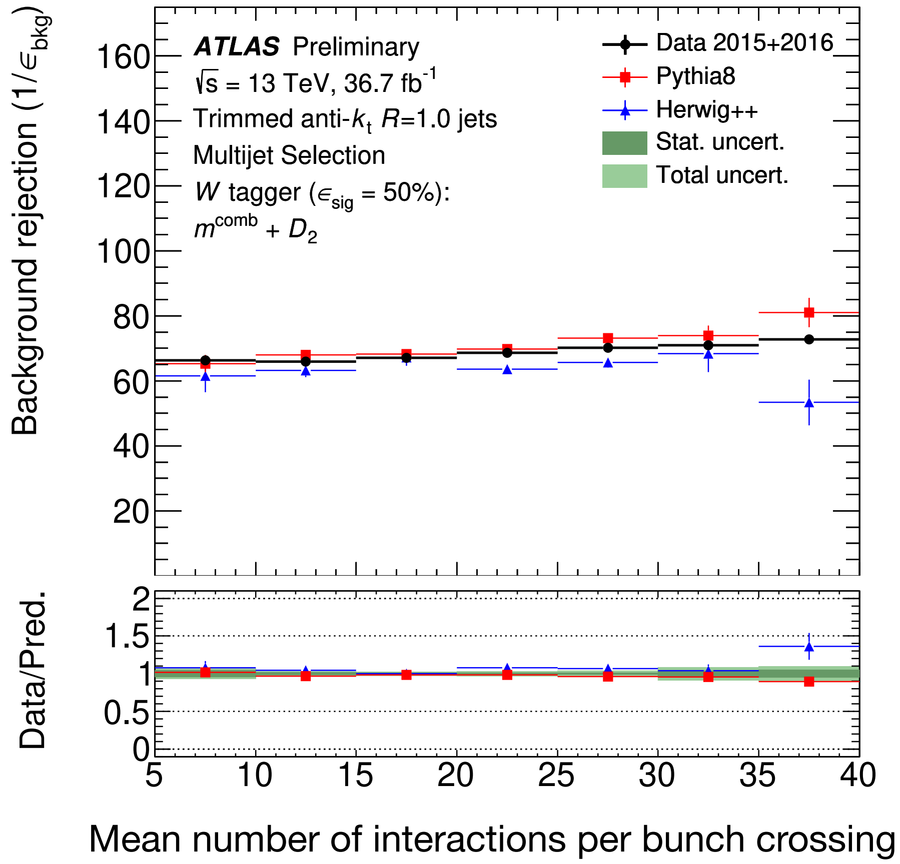
TopoDNN tagger



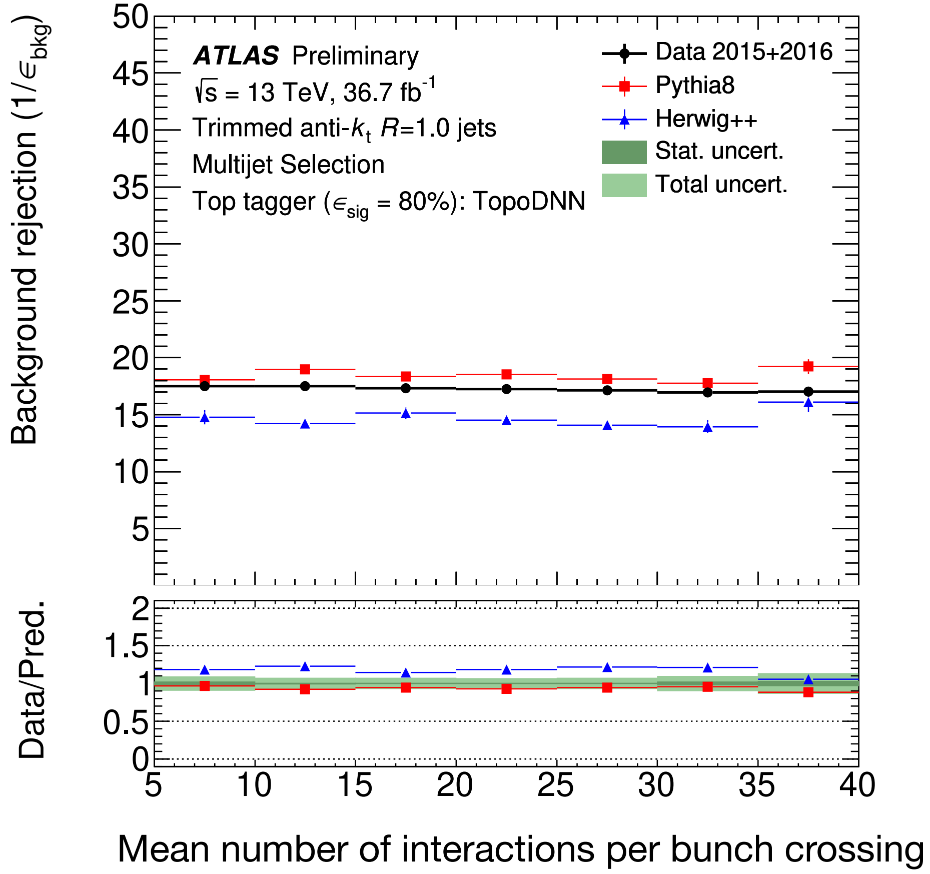
ATLAS-CONF-2017-064

Multijet rejection vs. pile-up

W tagger



TopoDNN tagger

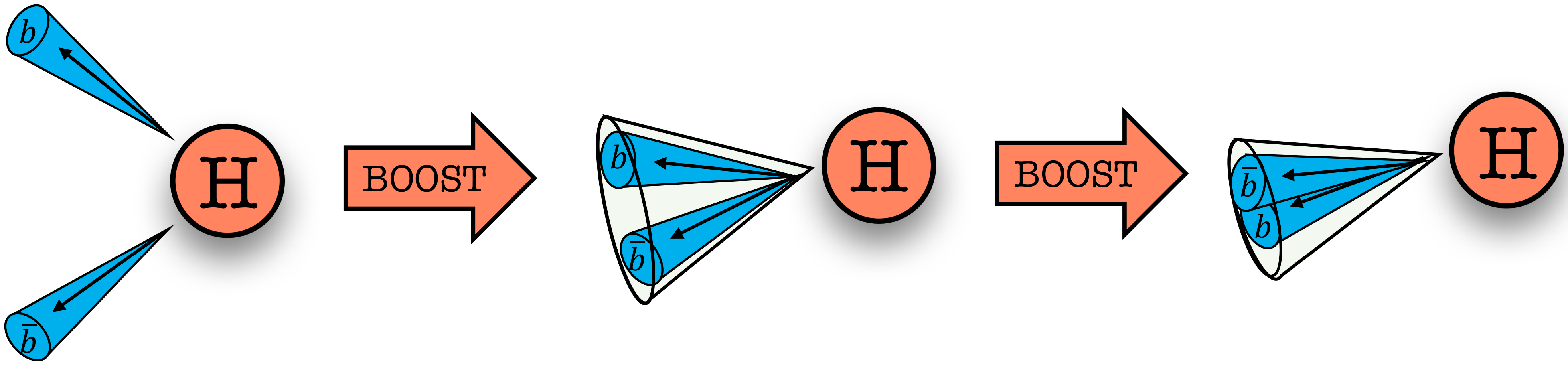


ATLAS-CONF-2017-064

$H \rightarrow bb$ tagging

ATL-PHYS-PUB-2017-010

Overview



- Current nominal tagger identifies b-jets with multivariate algorithm on anti-k_T R=0.2 track-jets. Loses efficiency at high-p_T due to b-jets merging.
- 3 new subjet reconstruction techniques to mitigate this loss.

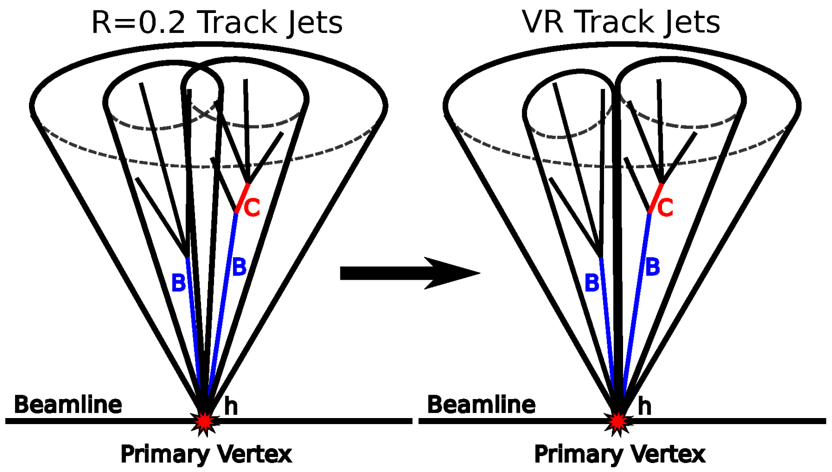
1 – Variable radius track jets

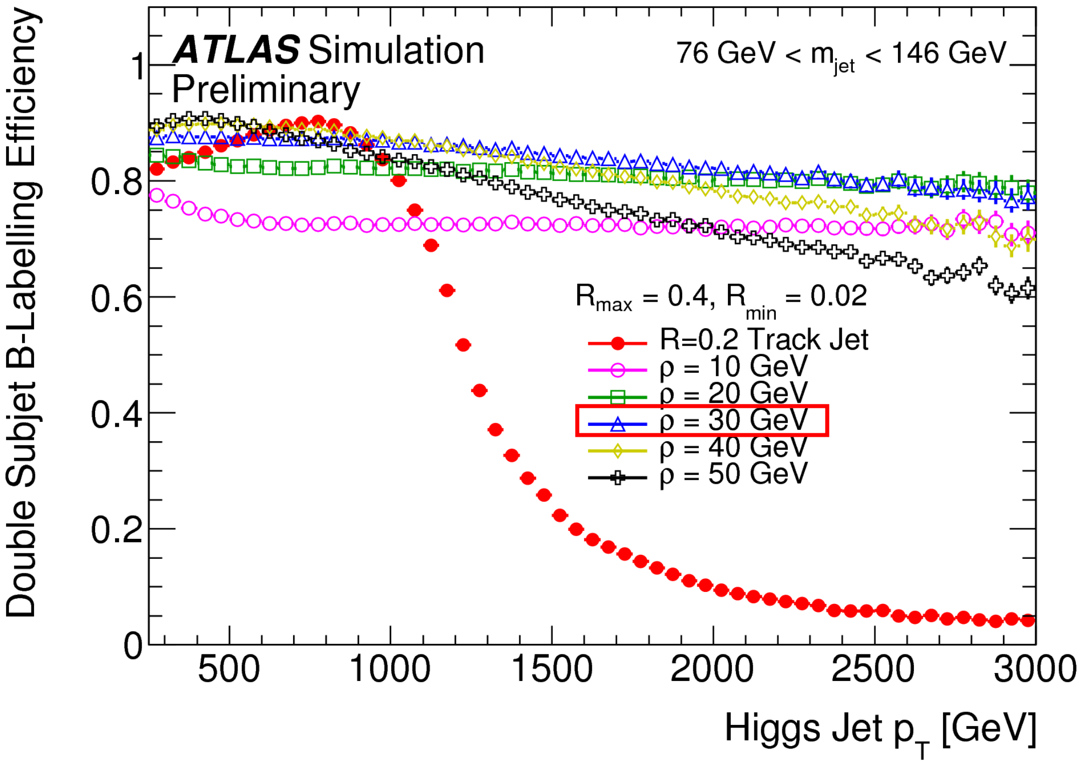
ATL-PHYS-PUB-2017-010

Subjets with dynamic radius parameter:

$$R_{eff}(p_T) = \frac{\rho}{p_T}$$

with low (R_{min}) and high (R_{max}) cutoff.

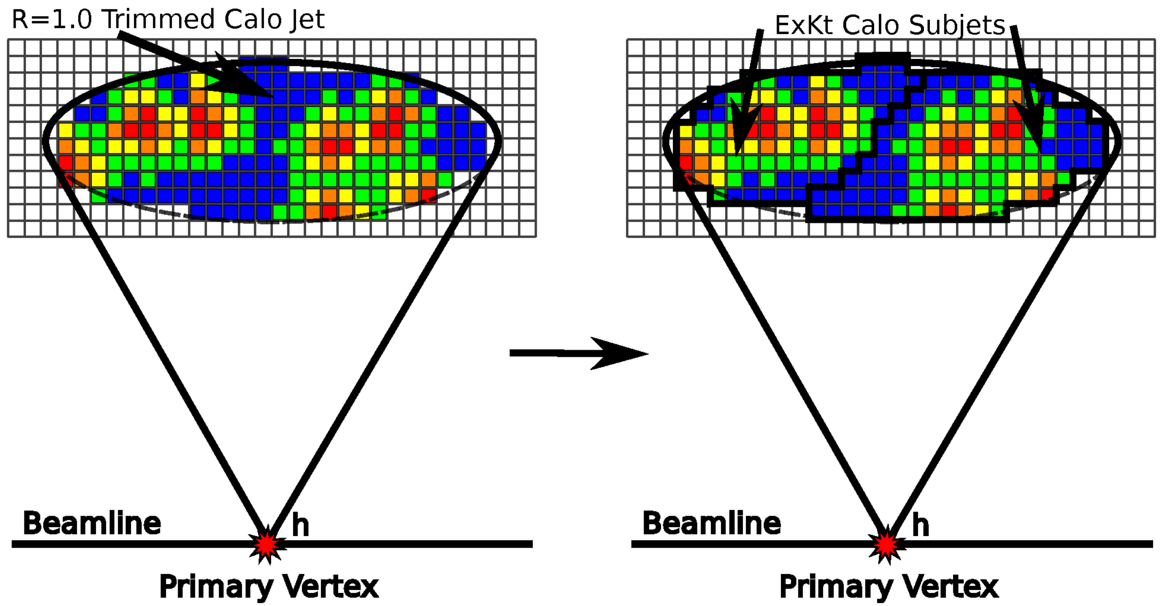




Scans performed over ρ, R_{min}, R_{max} to find optimal values

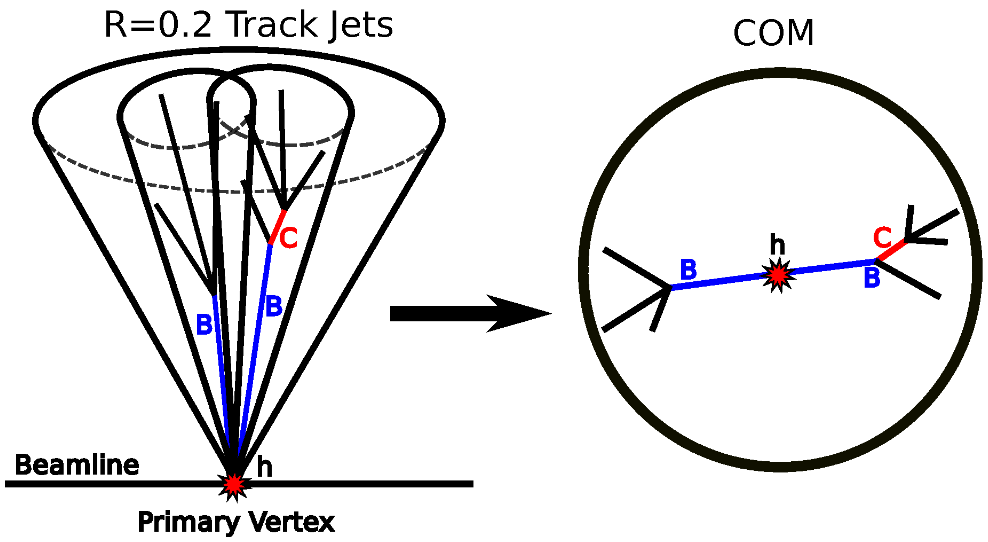
2 – Exclusive-k_T

- Recluster trimmed large-R jet calorimeter constituents with k_T algorithm and stop when 2 jets are obtained.
- Splits the large-R jet into two parts, each of which is expected to contain one *b*-hadron.



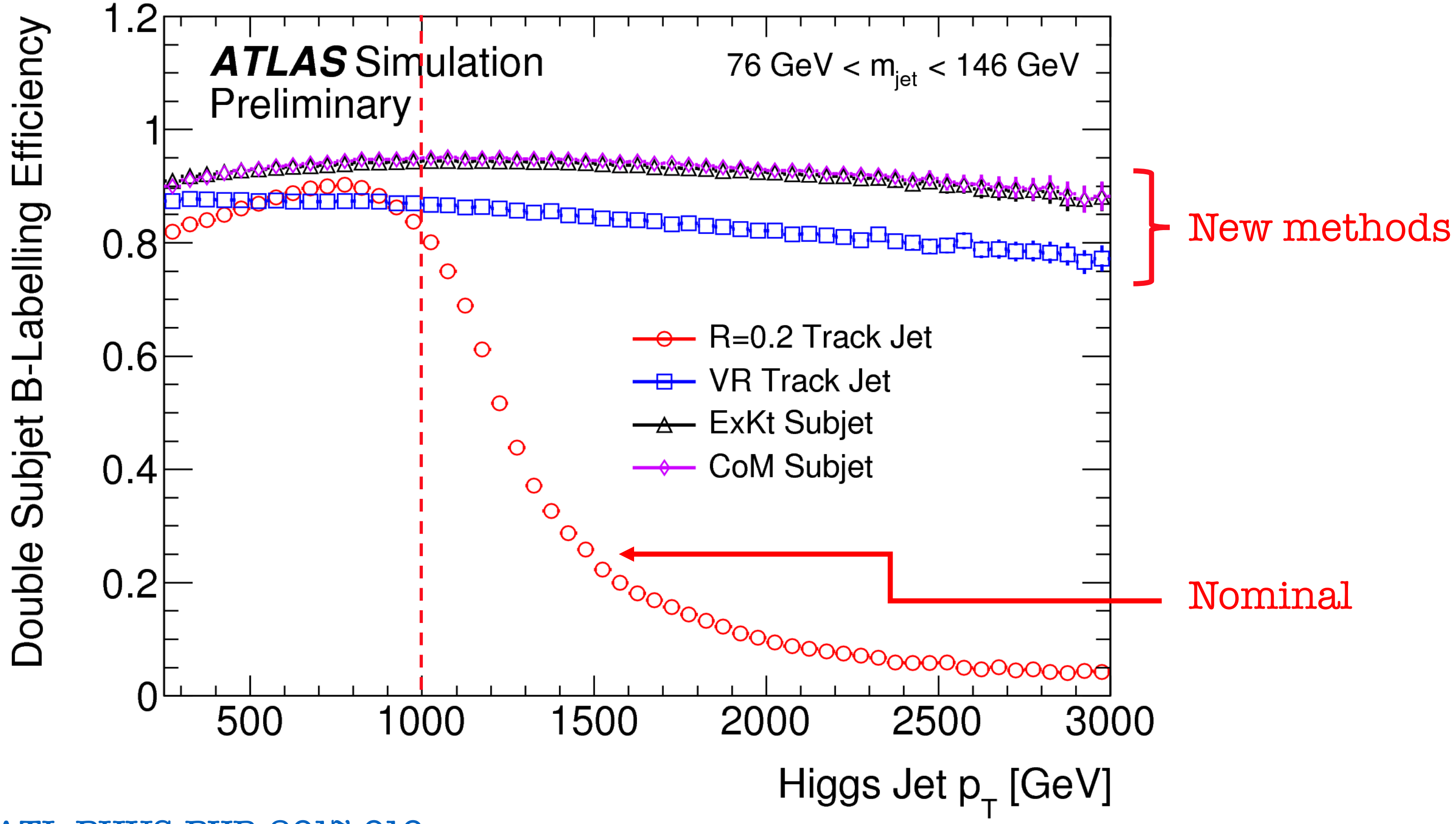
3 – Centre-of-mass

- Boost jet calorimeter clusters to the centre-of-mass frame of the large-R jet (jet rest frame) and reconstruct subjets.
- Tracks for *b*-tagging are also boosted to the centre-of-mass frame.



Results

New methods show large improvement over nominal tagger for $p_T > 1000$ GeV.



ATL-PHYS-PUB-2017-010

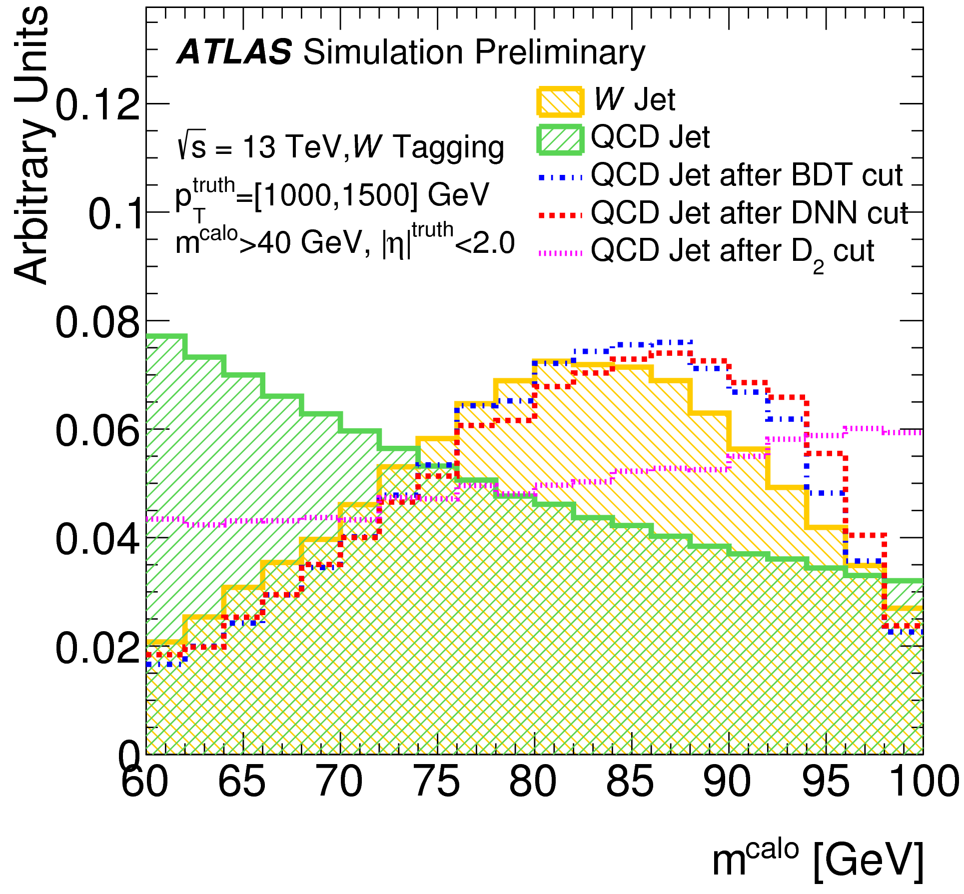
Mass decorrelation

ATL-PHYS-PUB-2018-014

Mass decorrelated taggers

- Jet substructure variables are correlated with jet mass. When you put many of them into a multivariate analysis the correlation gets very strong.
    - **Sculpting of the multijet background** -> resembles the resonance jet mass distribution
    - Depopulates side-band regions
- Aim to decorrelate jet substructure classifiers from jet mass

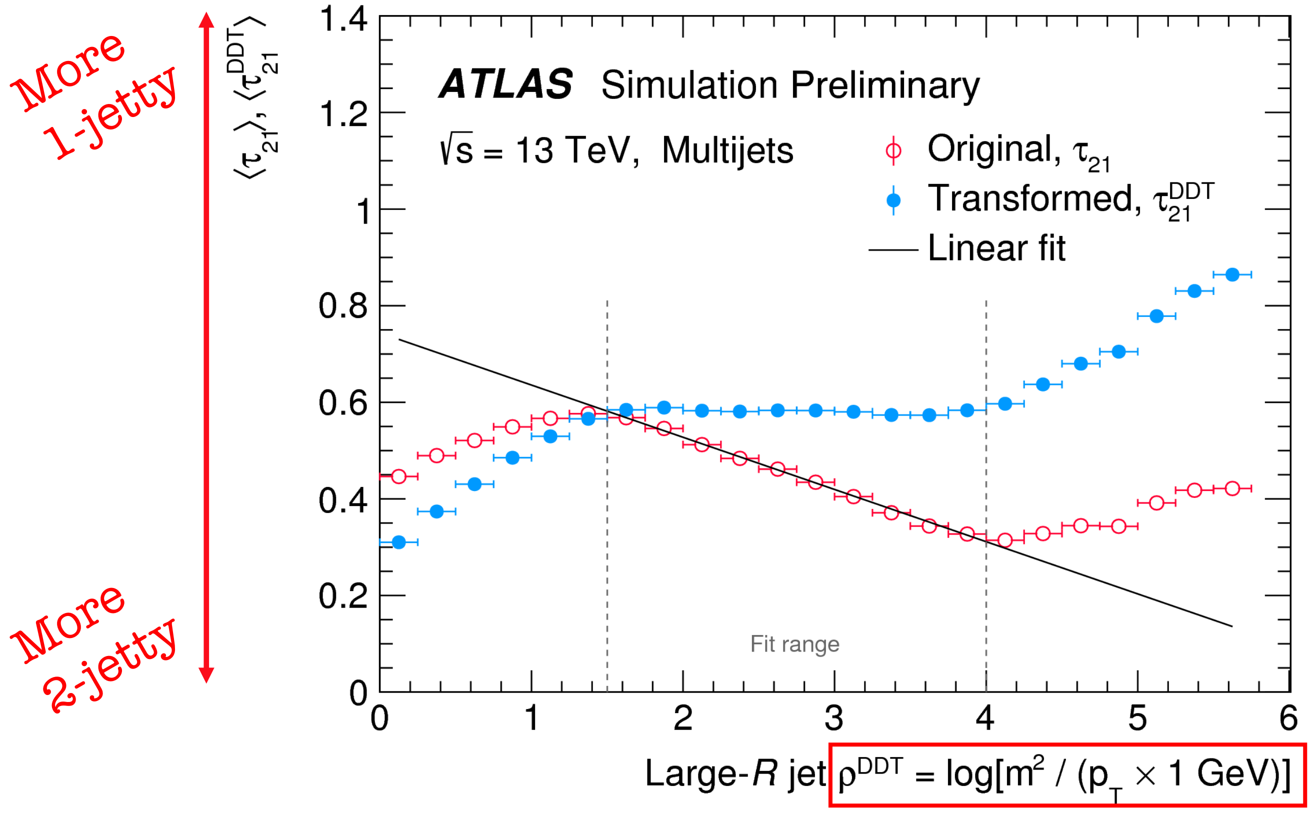
ATL-PHYS-PUB-2017-004



Designed decorrelated taggers (DDT)

- τ_{21} variable distinguishes 1-prong from 2-prong jets
- Has a linear relationship to jet scaling variable ρ^{DDT} for masses > 80 GeV

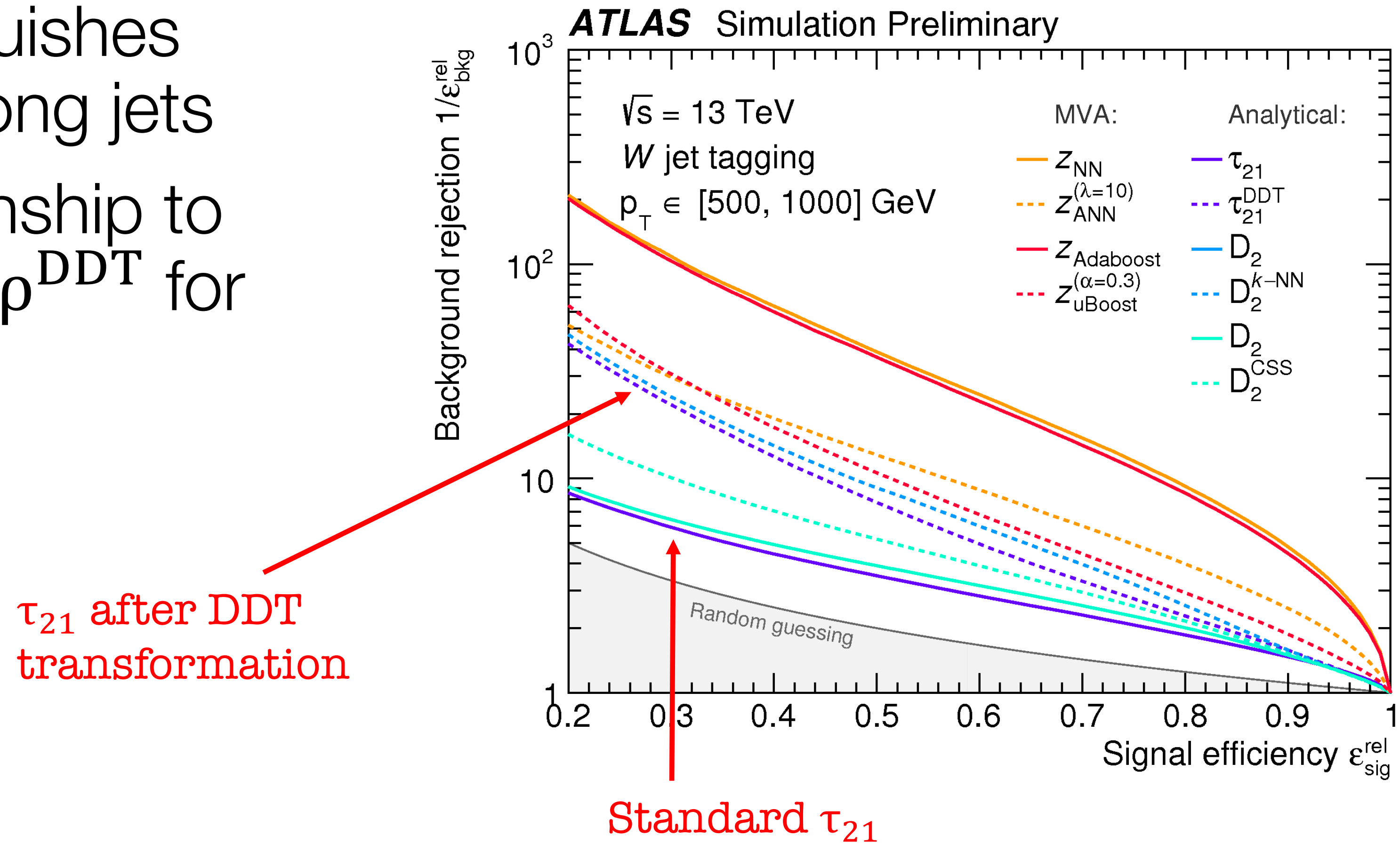
ATL-PHYS-PUB-2018-014



Designed decorrelated taggers (DDT)

ATL-PHYS-PUB-2018-014

- τ_{21} variable distinguishes 1-prong from 2-prong jets
- Has a linear relationship to jet scaling variable ρ^{DDT} for masses > 80 GeV



Summary

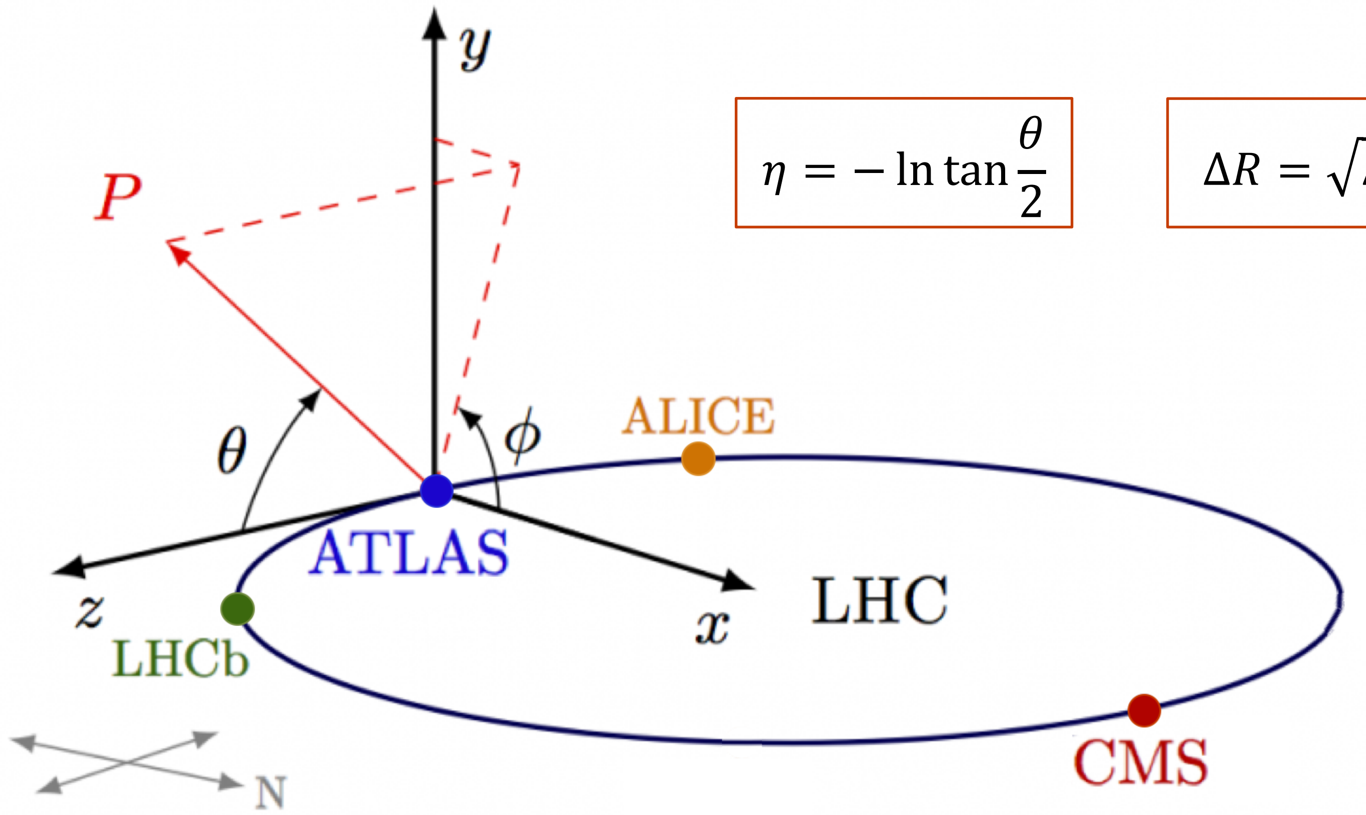
Summary

- Top/W tagging:
  - Machine learning taggers perform better than 2-variable taggers
  - Machine learning tagger with low-level inputs (TopoDNN) performs the best for top tagging
  - Signal efficiency & background rejection in data are well modelled by MC
- H → bb tagging:
  - 3 new subjet reconstruction techniques to overcome loss of efficiency at high p_T due to b-jet merging
  - Large improvement over nominal tagger for $p_T > 1000$ GeV
- Mass decorrelated taggers:
  - Various approaches studied with promising results

Fin.

Back-up

The ATLAS coordinate system

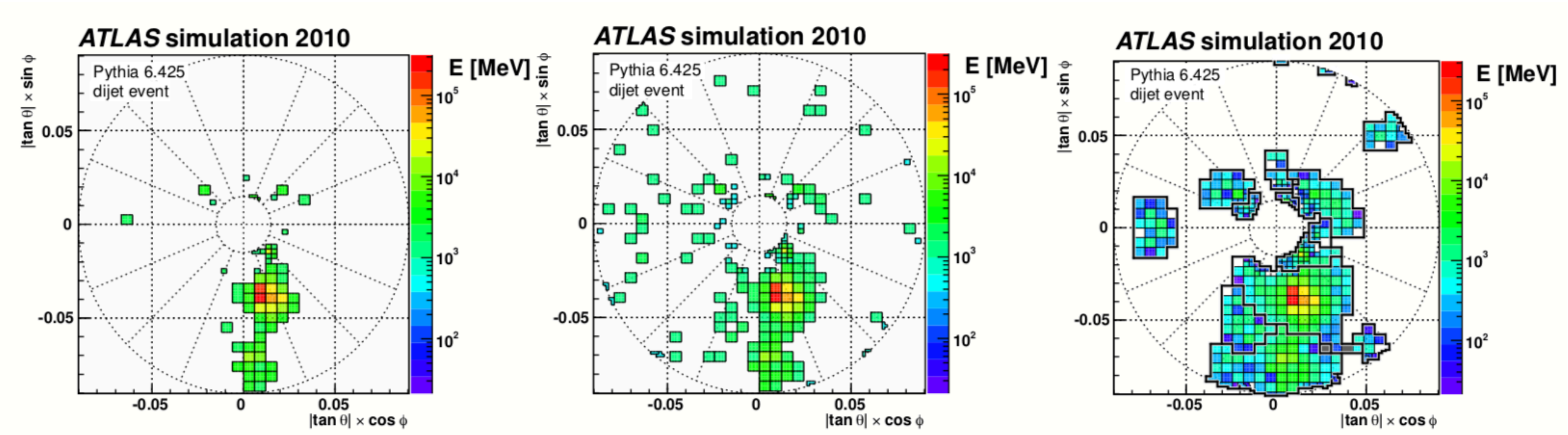


$$\eta = -\ln \tan \frac{\theta}{2}$$

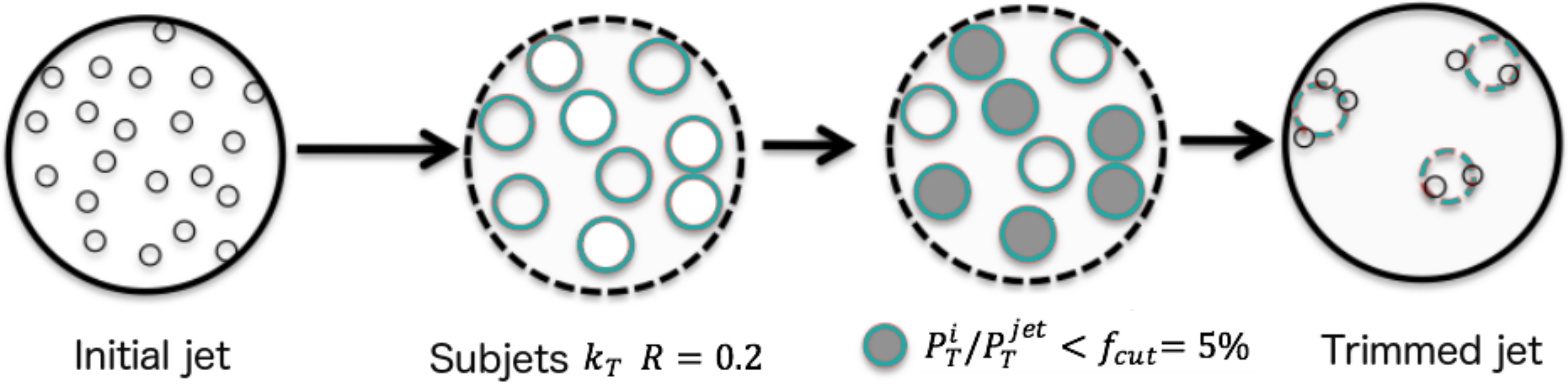
$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$$

Granularity of ATLAS calorimeter

- The hadronic calorimeter is coarser than the EM calorimeter. If we want to use all calorimeter layers we are limited by coarsest layer (R_{min}~0.2). But we can also ignore coarse layers and use only very fine layers.
  - EM granularity: $\Delta\eta\times\Delta\phi \approx 0.025\times\pi/128$
  - Hadronic granularity: $\Delta\eta\times\Delta\phi \approx 0.1\times\pi/32$



Jet trimming



Combined mass

Track-assisted mass: $\mathrm{m}^{\mathrm{TA}} = \mathrm{m}^{\mathrm{track}} \times \frac{\mathrm{p}_{\mathrm{T}}^{\mathrm{calo}}}{\mathrm{p}_{\mathrm{T}}^{\mathrm{track}}}$

in which $\mathrm{m}^{\mathrm{track}}$ and $\mathrm{p}_{\mathrm{T}}^{\mathrm{track}}$ are the invariant mass and pT calculated from tracks associated with the large-R trimmed calorimeter jet and $\mathrm{p}_{\mathrm{T}}^{\mathrm{calo}}$ is the pT of the original trimmed large-R jet.

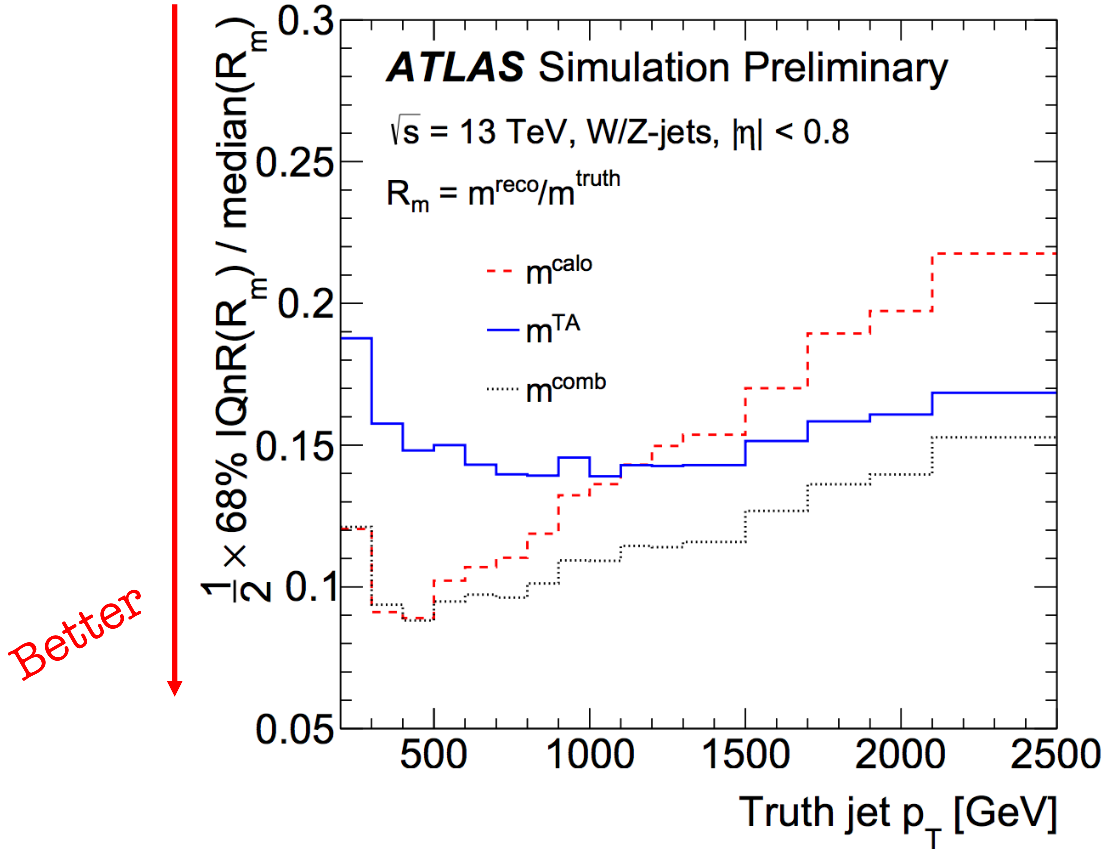
Calorimeter mass: $m^{calo} = \sqrt{\left(\sum_i E_i\right)^2 - \left(\sum_i \vec{p}_i\right)^2}$

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

with $\mathrm{a} = \frac{\sigma_{\mathrm{calo}}^{-2}}{\sigma_{\mathrm{calo}}^{-2} + \sigma_{\mathrm{TA}}^{-2}}$ and $\mathrm{b} = \frac{\sigma_{\mathrm{TA}}^{-2}}{\sigma_{\mathrm{calo}}^{-2} + \sigma_{\mathrm{TA}}^{-2}}$

where σ_{calo} and σ_{TA} are the calorimeter-based jet mass resolution function and the track-assisted mass resolution function respectively.

ATLAS-CONF-2016-035



JSS variables

- Energy correlation ratio: $\mathrm{D}_2 = \frac{\mathrm{e}_3}{(\mathrm{e}_2)^3}$ where e_2 and e_3 are the 2- and 3-prong energy correlation functions which are sensitive to the 2- and 3-prong structure in a jet.
- N-subjettiness: $\tau_{32} = \tau_3/\tau_2$ where

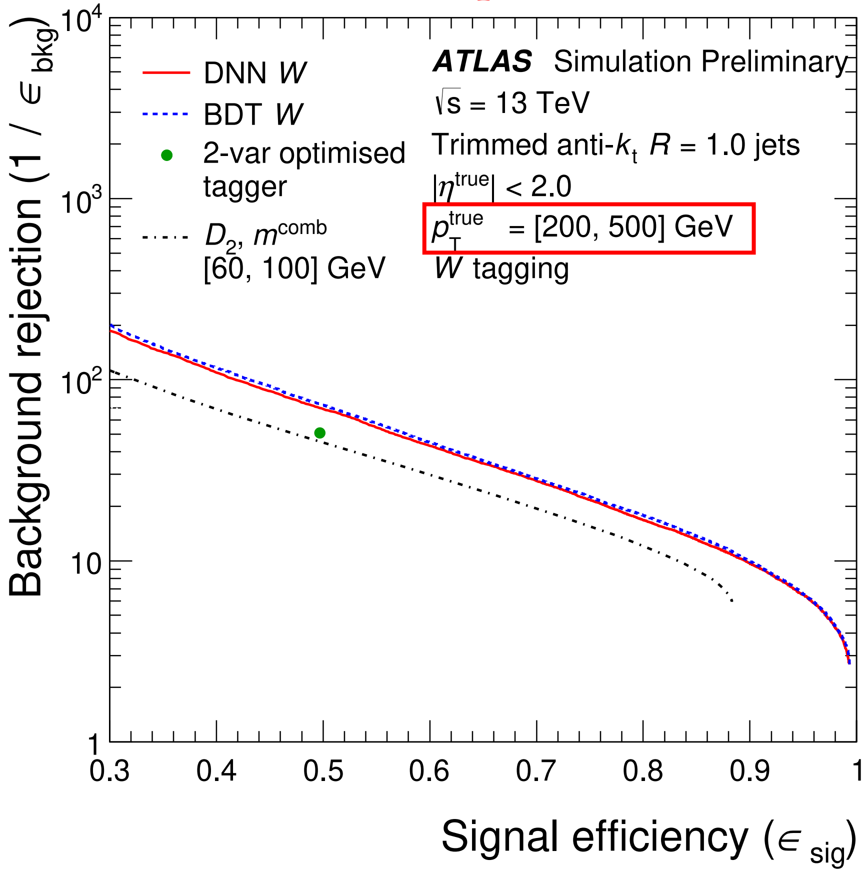
$$\tau_\mathrm{N} = \frac{1}{\mathrm{d}_0}\sum_\mathrm{k} \mathrm{p}_{\mathrm{T,k}} \times \min(\Delta\mathrm{R}_{1\mathrm{k}}, \Delta\mathrm{R}_{2\mathrm{k}}, \dots, \Delta\mathrm{R}_{\mathrm{Nk}}), \qquad \mathrm{d}_0 = \sum_\mathrm{k} \mathrm{p}_{\mathrm{Tk}} \times \mathrm{R}$$

R is radius parameter of the jet, p_{Tk} is transverse momentum of constituent k, and $\delta\mathrm{R}_{\mathrm{ik}}$ is the distance between the subjet i and the constituent k. The N-subjettiness variable τ_N expresses how well a jet can be described as containing N subjets.

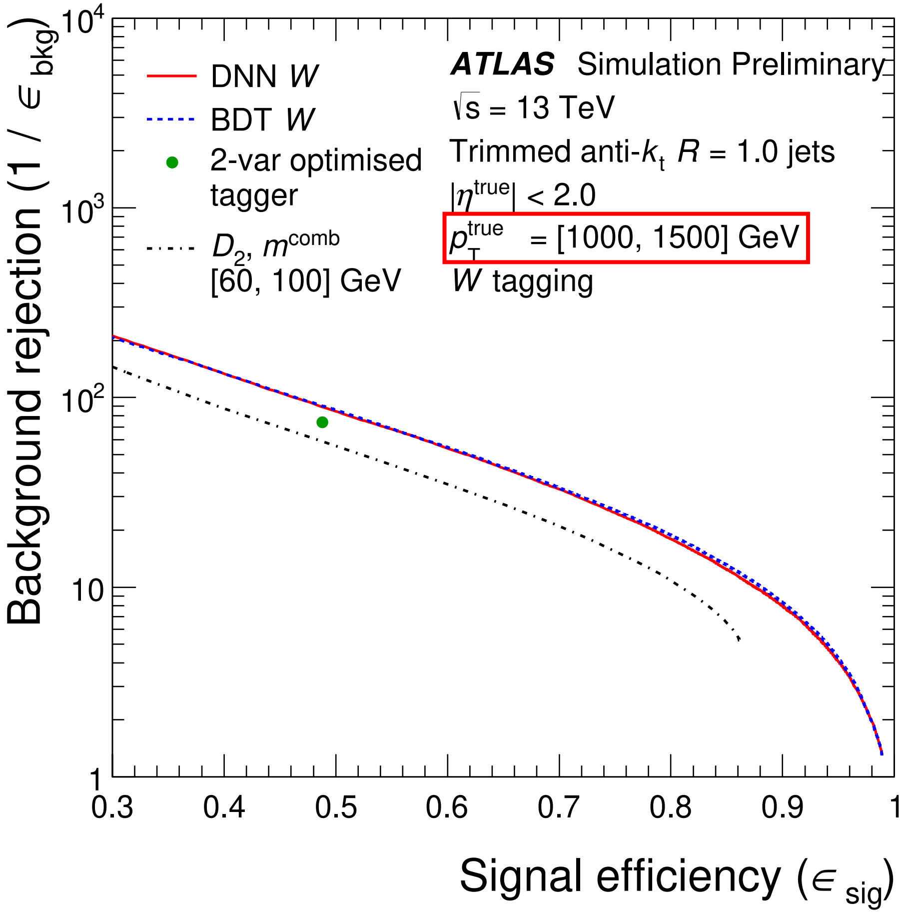
- Jet mass calculated from constituents: $m^2 = (\sum_i E_i)^2 - (\sum_i \vec{p}_i)^2$

W tagging with ML methods

Low pT

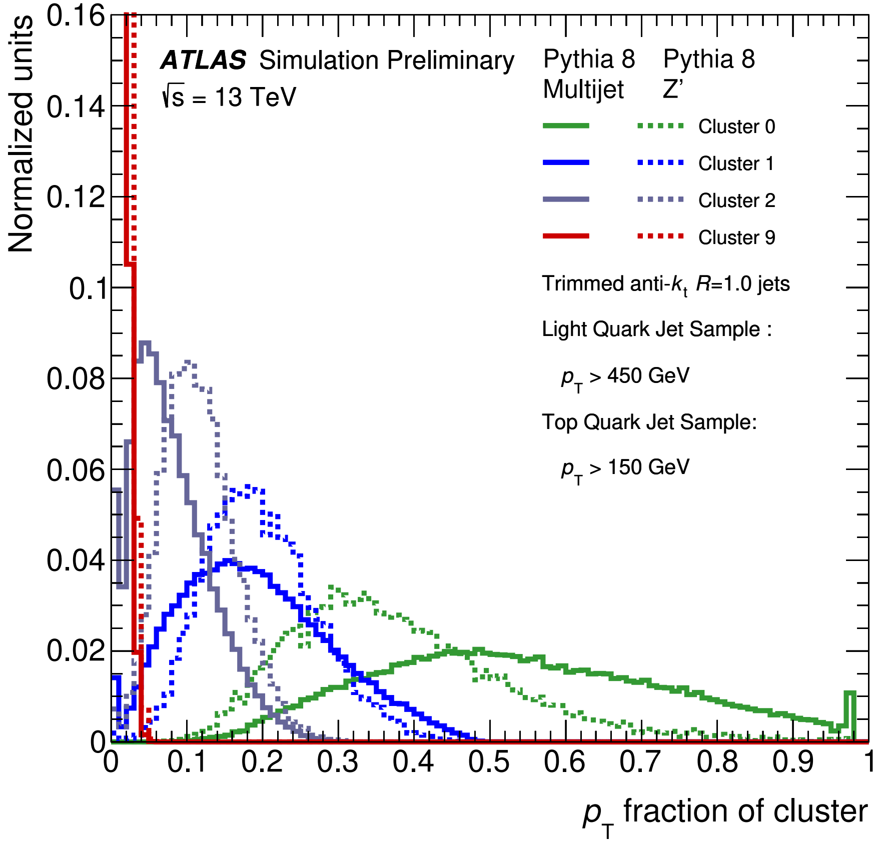


High pT



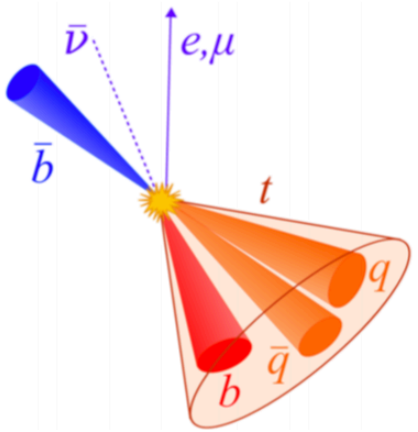
TopoDNN

Uses p_T, η, φ of 10 leading topoclusters in trimmed large-R jet.



Top/W tagging in data

- Full ATLAS 2015 – 2016 dataset, L = 36.1 – 36.7 fb^{-1}
- ***W*/top tagging efficiency**
  - $t\bar{t}$ decay to single lepton + jets topology
  - ***W*:** ΔR(*b*-jet, large-*R* jet) > 1.0
    $p_T(J)$ > 200 GeV
  - **Top:** ΔR(*b*-jet, large-*R*/HTT jet) < 1.0/1.5
    $p_T(J)$ > 350 GeV
- **Multijet rejection**
  - Dijets: $p_T(J)$ > 450 GeV
  - γ + jets: $p_T(J)$ > 200 GeV



Single-lepton $t\bar{t}$ decay

Dijets

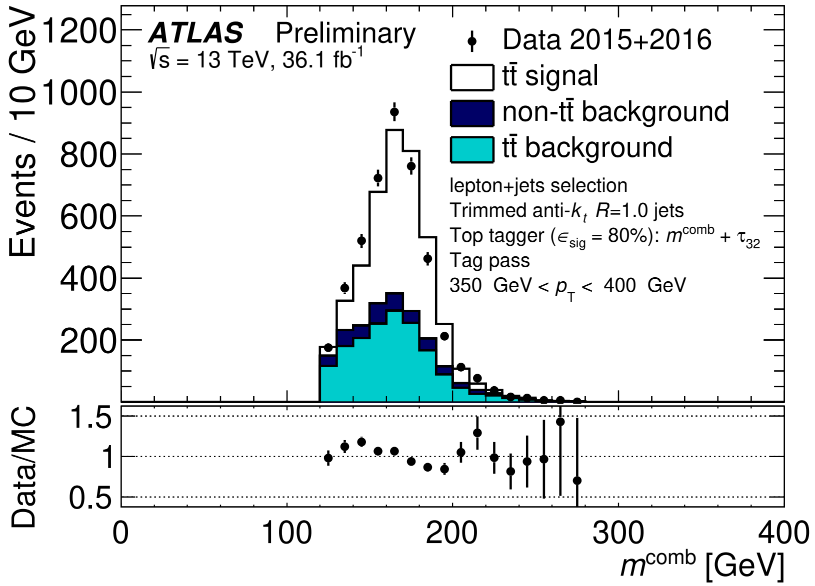
γ+jet

W/top tagging efficiency

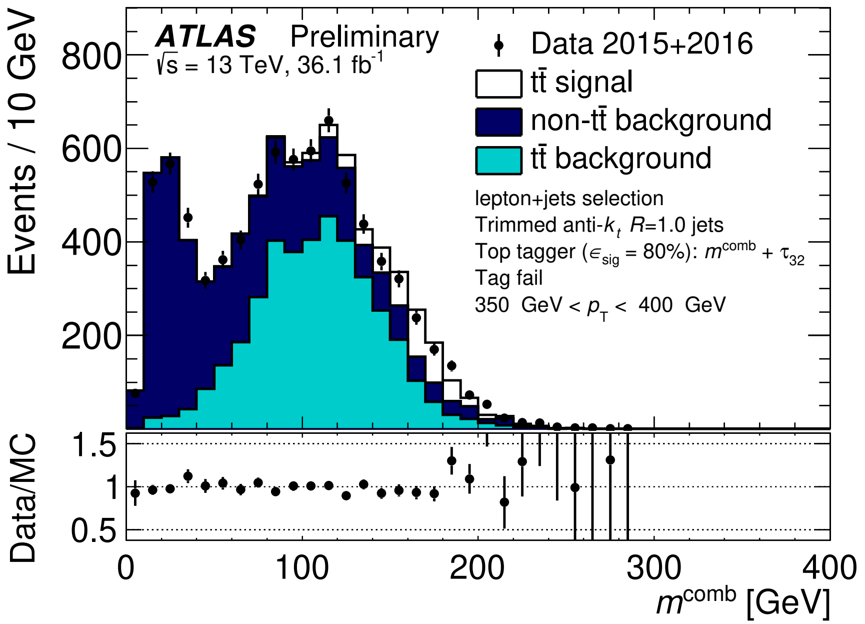
Need to measure efficiency in data and get uncertainty on this efficiency.

Measure signal-like events in data to fit large-R jet mass for events passing/failing a tagger

Passing tag



Failing tag

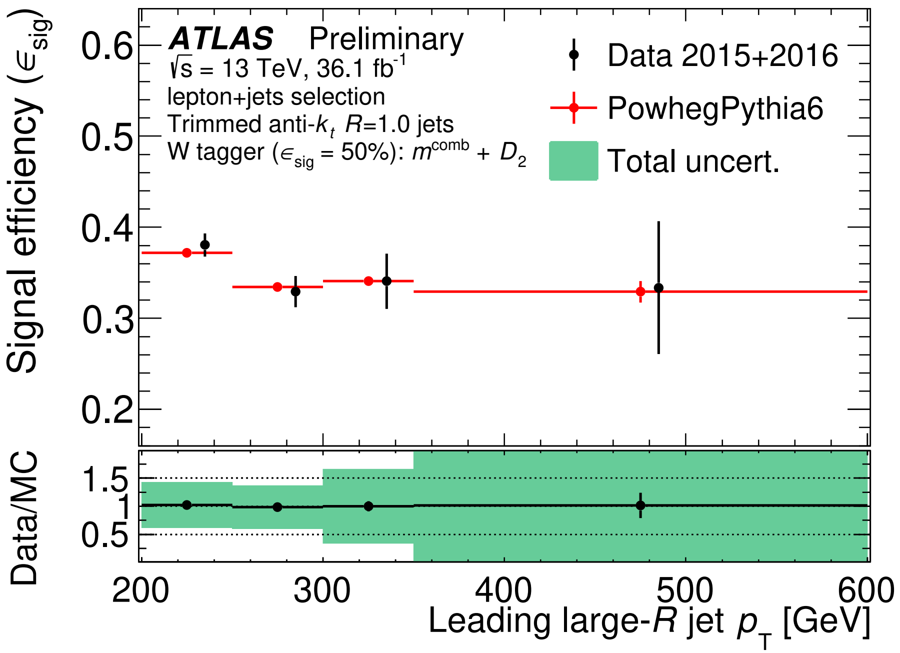


W/top tagging efficiency vs. jet pT

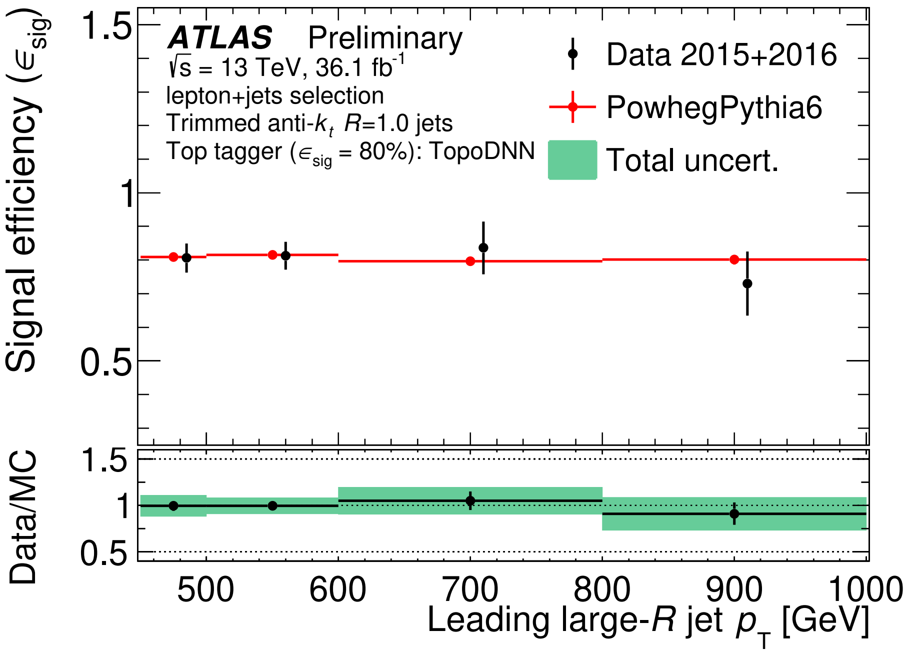
pre-fit $\epsilon_{MC} = \frac{N^{\text{tagged}}_{\text{signal}}}{N^{\text{tagged}}_{\text{signal}} + N^{\text{not tagged}}_{\text{signal}}}$

$\epsilon_{\text{data}} = \frac{N^{\text{tagged}}_{\text{fitted signal}}}{N^{\text{tagged}}_{\text{fitted signal}} + N^{\text{not tagged}}_{\text{fitted signal}}}$ post-fit

W tagger

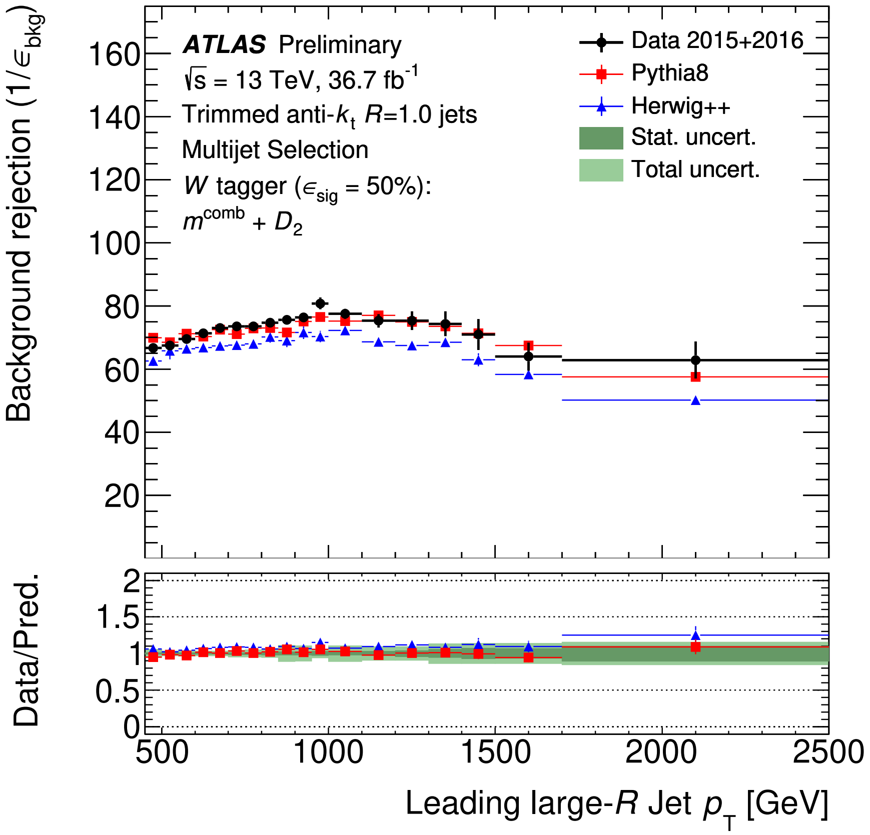


TopoDNN tagger

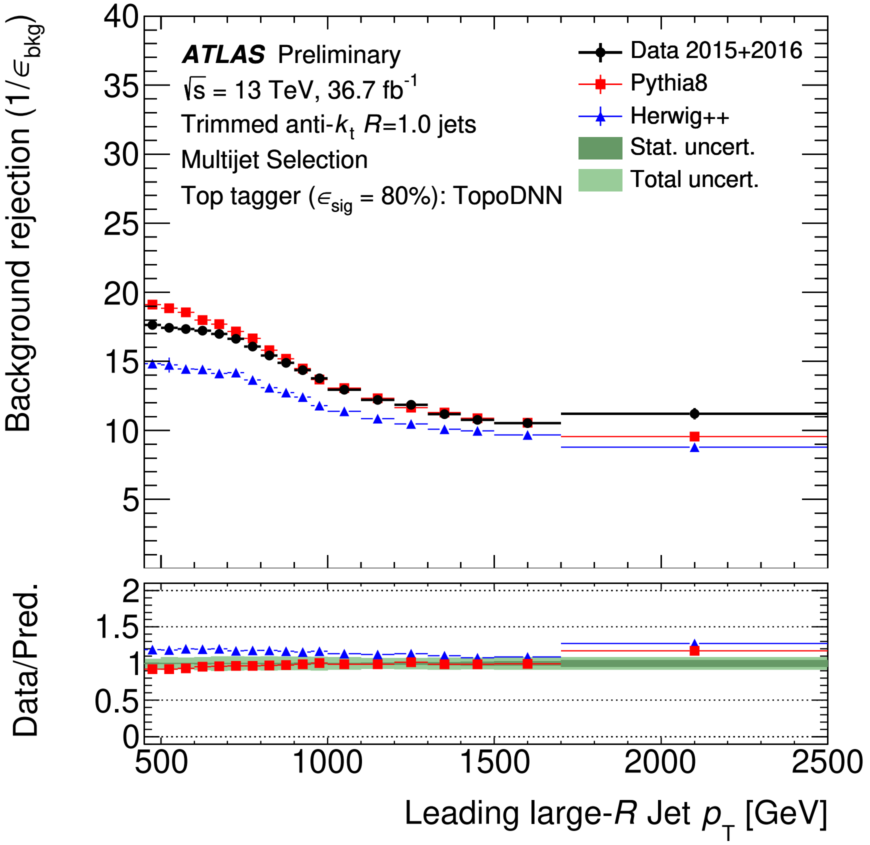


Multijet rejection vs. jet pT

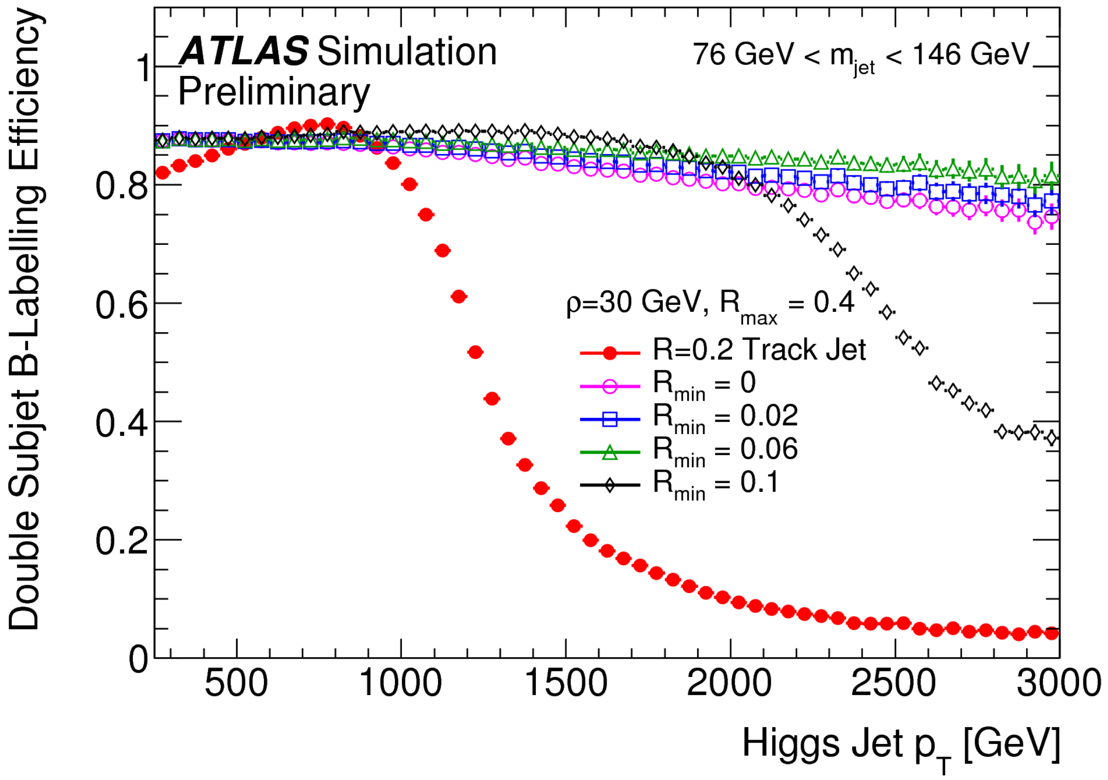
W tagger

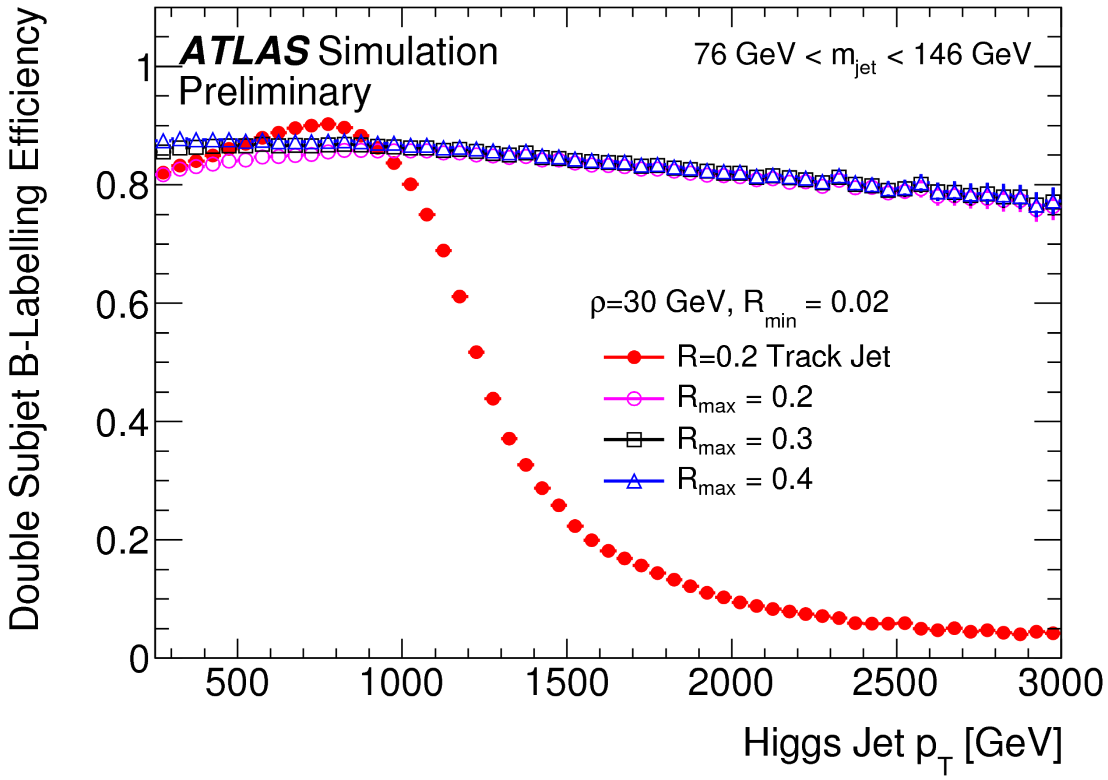


TopoDNN tagger

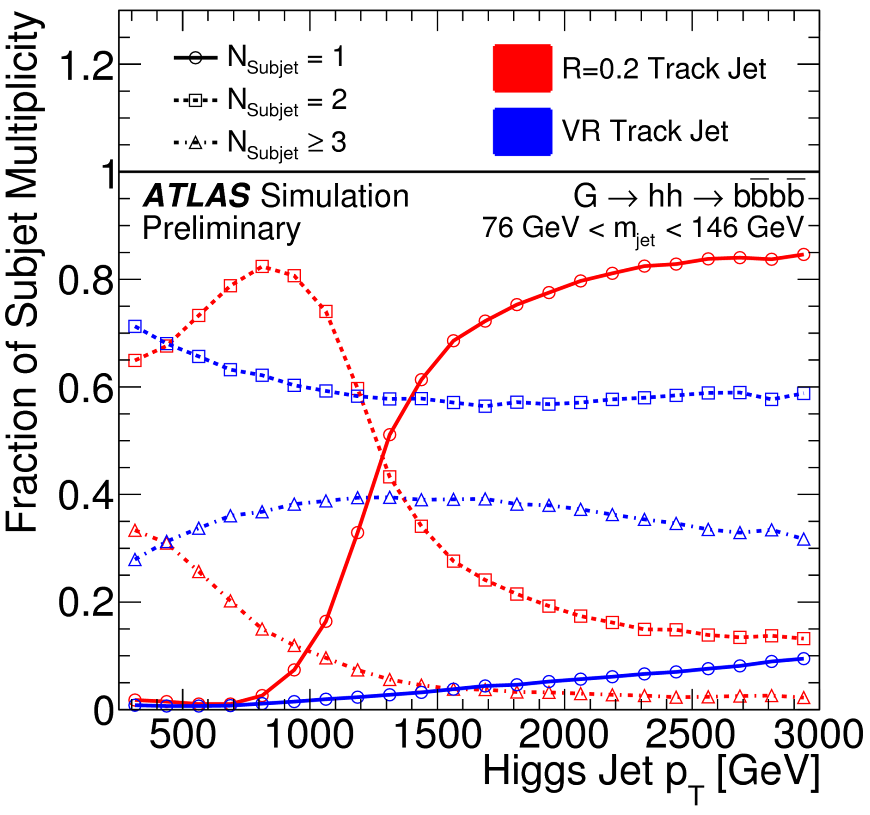


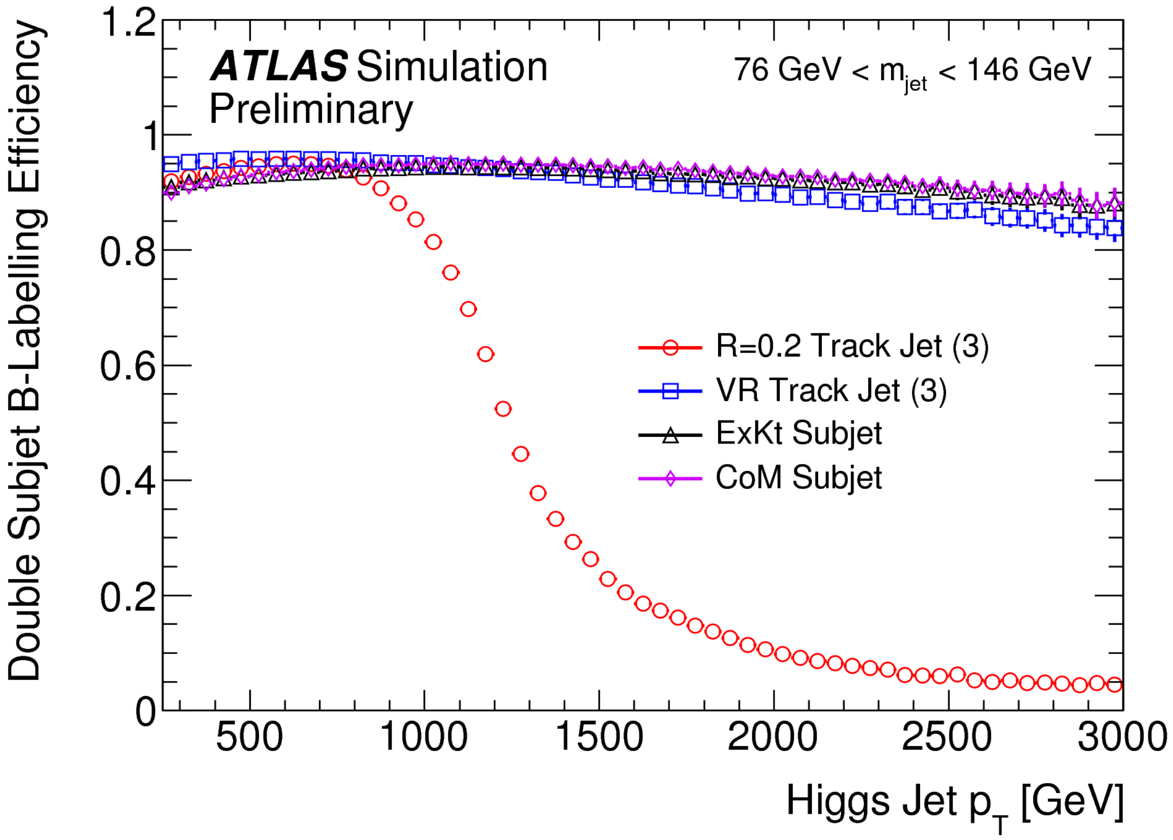
Variable radius track jets



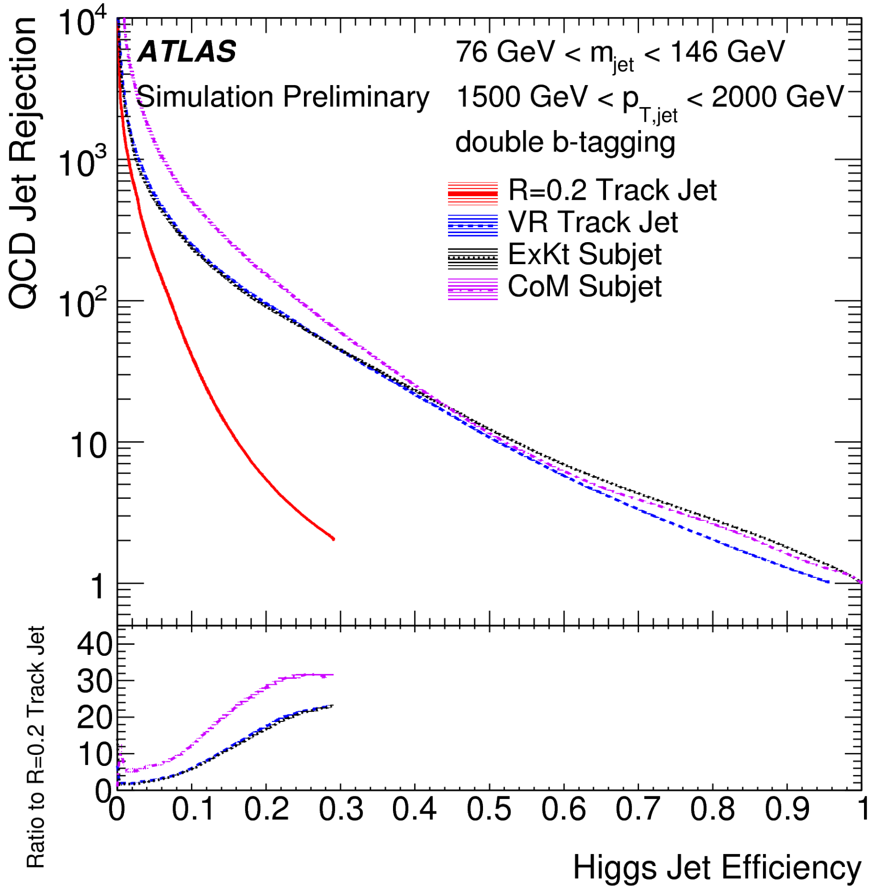


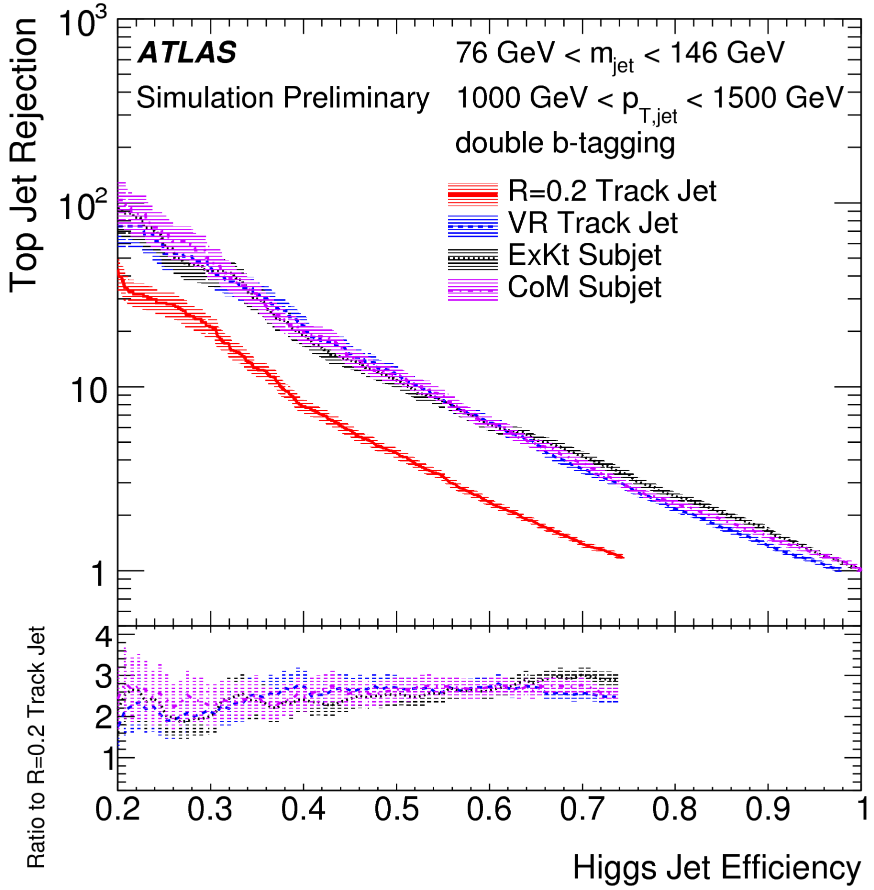
Results H->bb tagging techniques





H->bb tagging results





Designed decorrelated taggers

- Linear fit on the relationship between τ_{21} and rho. Transformation is:

$$\tau_{21}^{DDT} = \tau_{21} - a \times (\rho^{DDT} - 1.5)$$

a is the slope of the fit in the plot.

- DDT transformation removes the linear correlation of τ_{21} with ρ. Since ρ has info on kinematics of the jet (m and pT), the DDT transform yields a JSS discriminant which is decorrelated from the jet mass.

Mass decorrelation