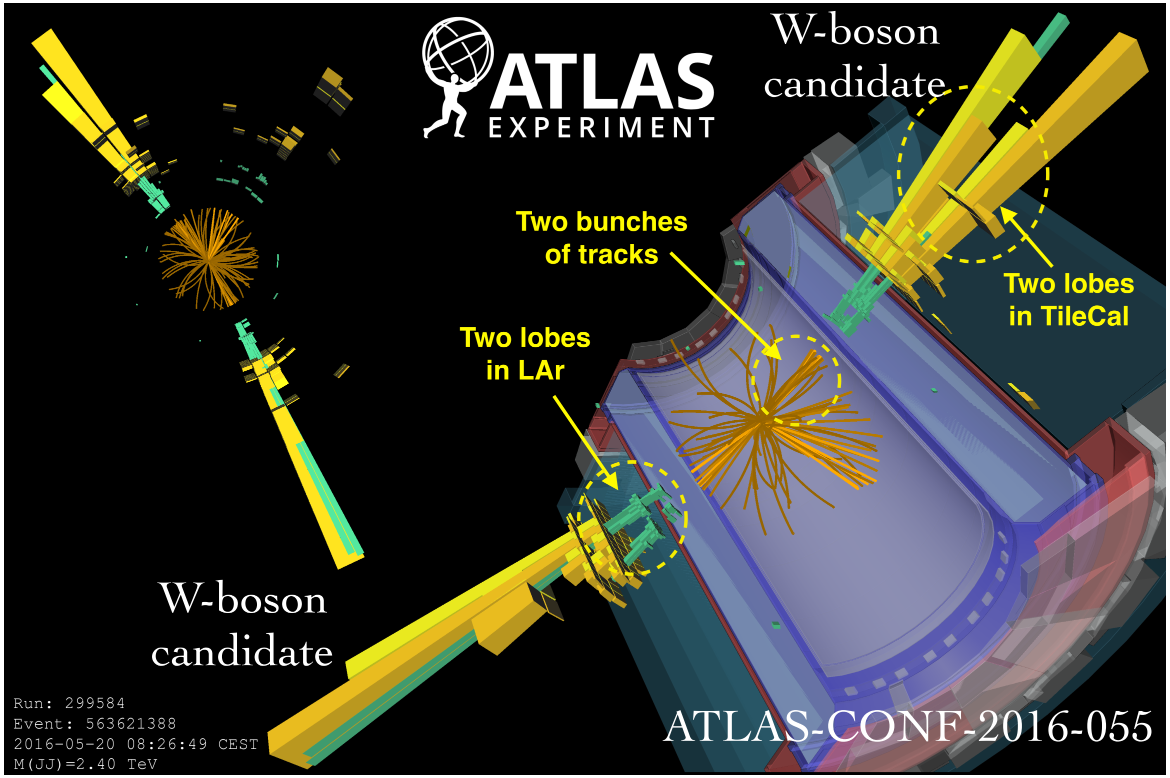
Measurement of jet substructure observables using the ATLAS detector.

Iain Bertram
Lancaster University
Monday 15 July
QCD@LHC2019

Jet Substructure

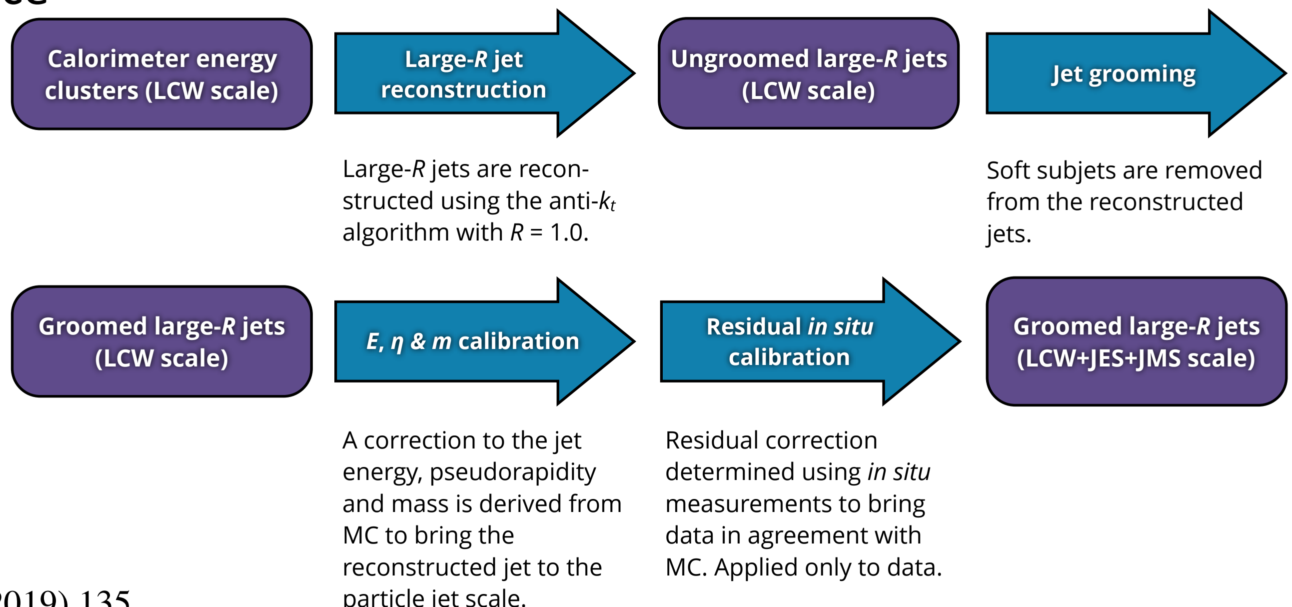
- The internal structure of jets can be used to
  - study QCD
  - distinguish the origin of jets between
    light quarks, gluons and hadronic decays of heavy particles
    (W, Z, H or top quark)



- Many types of jets used in ATLAS
  - calorimeter-based, particle-flow, track-based
  - radius R = 1.0, 0.8, 0.4, 0.2, variable-R
- Calorimeter-based R = 0.8 and 1.0 jets reconstructed with anti-k_T algorithm are used for the substructure results presented here.

Large-R Calorimeter Jets

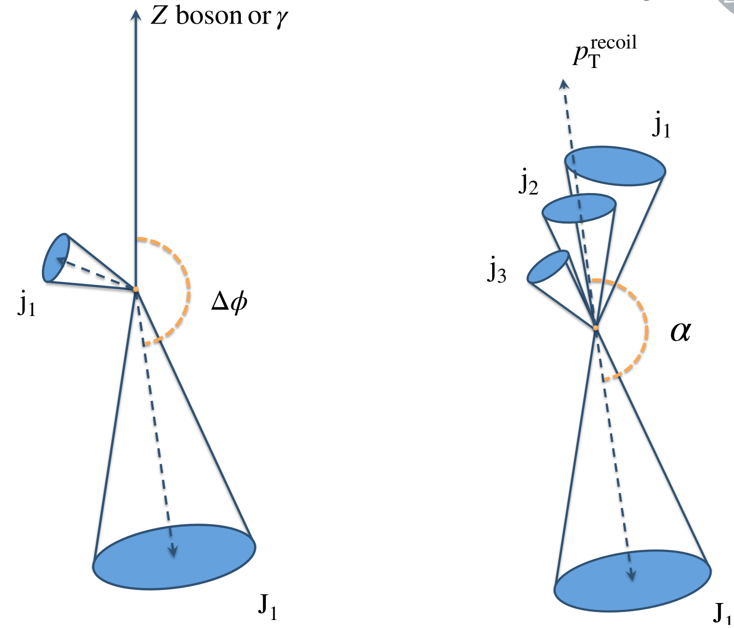
- Calorimeter jets are built from topological clusters (representing particle deposition)
  - start with a cell 4σ above noise and add neighbouring cells 2σ above noise and the surrounding layer
  - splitting algorithm separates nearby clusters
  - clusters are calibrated based on properties related to shower development (Local Cluster Weighting)
  - clusters are combined into jets using the anti-k_T algorithm
- **Trimming**: removes R = 0.2 subjets with p_T < 5% of jet p_T, to reduce pile-up dependence

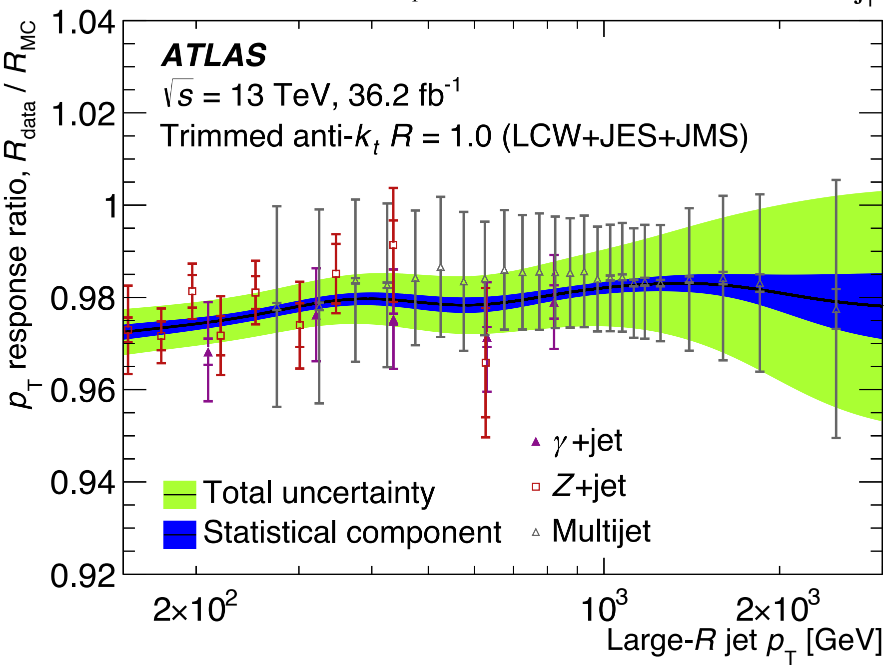


Eur. Phys. J. C 79 (2019) 135

Residual In Situ Calibration

- In situ calibration uses "balancing" with well-measured objects: photons, $Z \rightarrow ll$, small-R jets
- p_T-dependent scale factors are derived to rescale the jet 4-momentum
- provides systematic uncertainties on the jet energy scale - "**top-bottom**" approach
- Average correction
  – no access to differential quantities





Eur. Phys. J. C 79 (2019) 135

Substructure Uncertainties

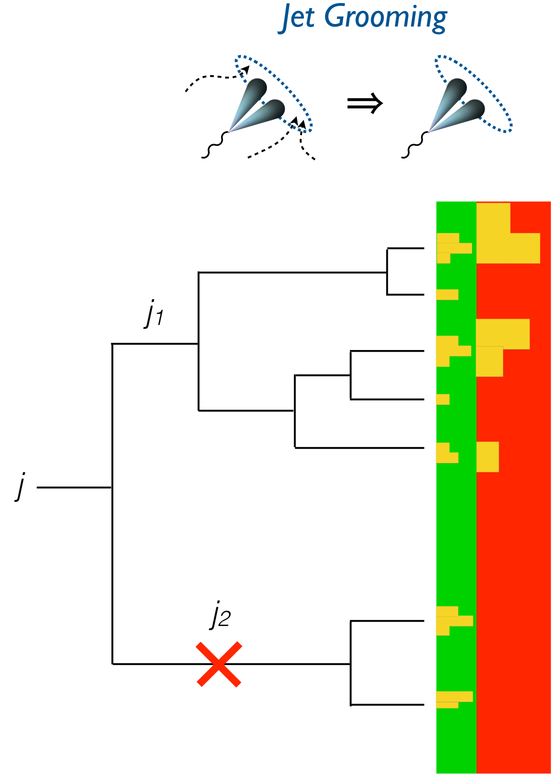
- uncertainties are computed from the topological clusters and propagated to the jet substructure observables ("**bottom-up**" approach):
  - isolated calorimeter cell clusters are matched to tracks: the mean and the standard deviation of E/p is used for the cluster energy scale and resolution uncertainties; the standard deviation of the relative position is used for the angular resolution uncertainty
  - the reconstruction efficiency uncertainty is evaluated from the fraction of tracks without a matched cluster.
  - Uncertainties are also applied to account for:
    - energy correlations between the clusters
    - fraction of hadrons that are not pions
    - effect of cluster splitting and merging
- Uncertainties from calorimeter/track ratios are of the same size (can only be measured for average quantities)

Soft-Drop Grooming

- The Trimming procedure is not analytically calculable
  - trimmed jet measurements can only be compared to the NLL predictions
- The soft-drop procedure allows comparison to NLO+NLL and LO+NNLL predictions for the jet mass
  - clusters jet constituents with Cambridge-Aachen algorithm and retraces the clustering history from the last branching
  - consider a jet of radius R with only two constituents. The soft-drop procedure removes the softer component unless

$$\frac{\min(p_{T,j_1}, p_{T,j_2})}{p_{T,j_1} + p_{T,j_2}} > z_{cut} \left(\frac{\Delta R_{12}}{R}\right)^{\beta},$$

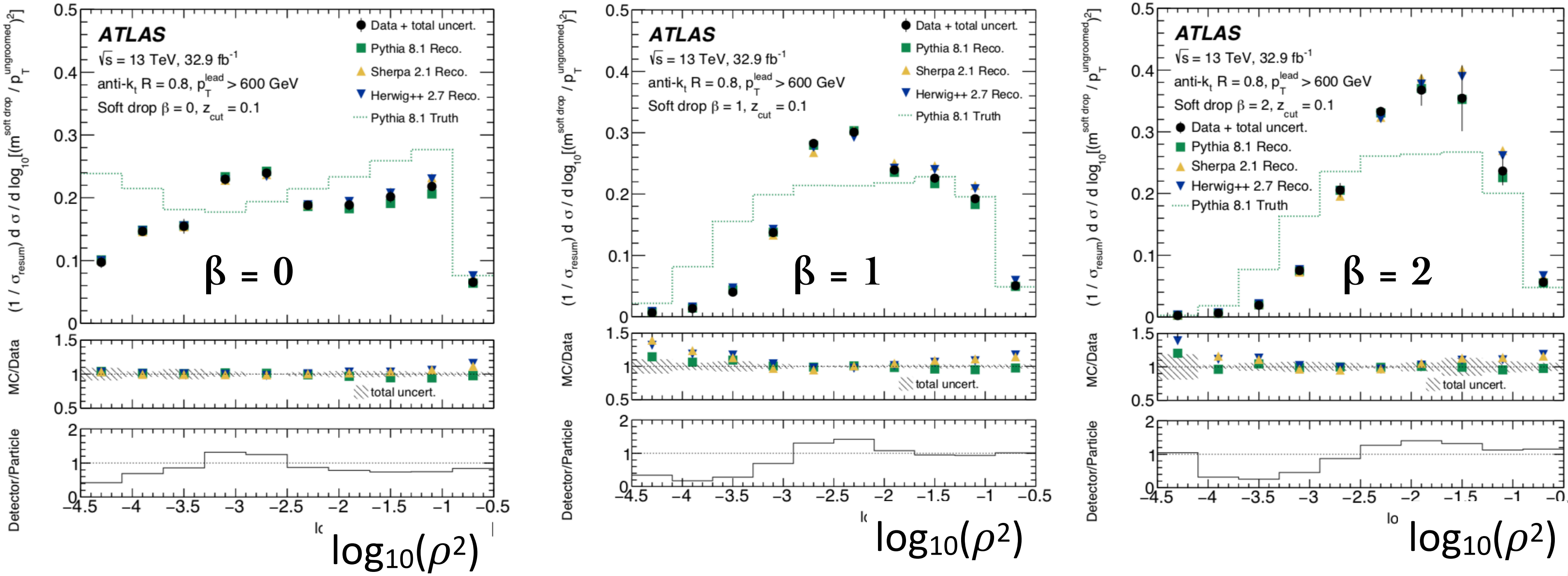
  - z_{cut} is the soft-drop threshold
  - β is angular exponent (β -> infinity: ungroomed jet)



Phys. Rev. Lett. 121 (2018) 092001

Soft-drop Jet Mass

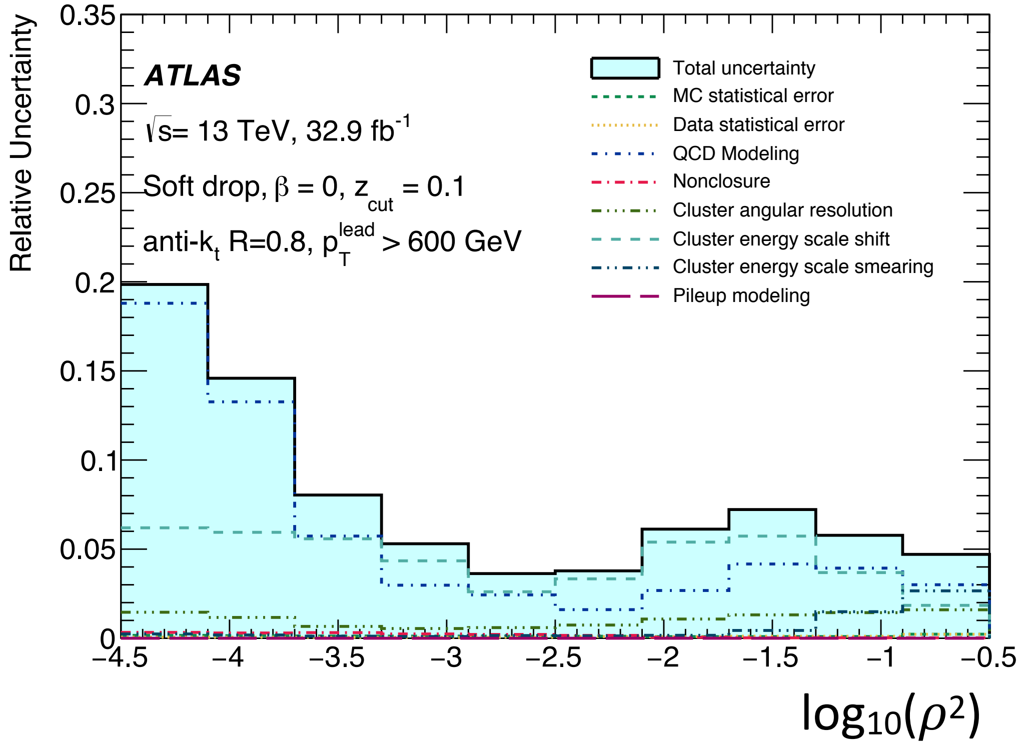
- Ungroomed dijet events with radius R=0.8.
  - Leading jet $p_T > 600$ GeV
  - two leading jets have similar p_T: $p_{T,1}/p_{T,2} < 1.5$
- To reduce p_T dependence we examine: $\rho = m^{\text{soft drop}} / p_T^{\text{ungroomed}}$



- Data compared with reconstructed Pythia, Sherpa and Herwig simulations and truth-level Pythia.

Phys. Rev. Lett. 121 (2018) 092001

Soft-drop Mass Uncertainties

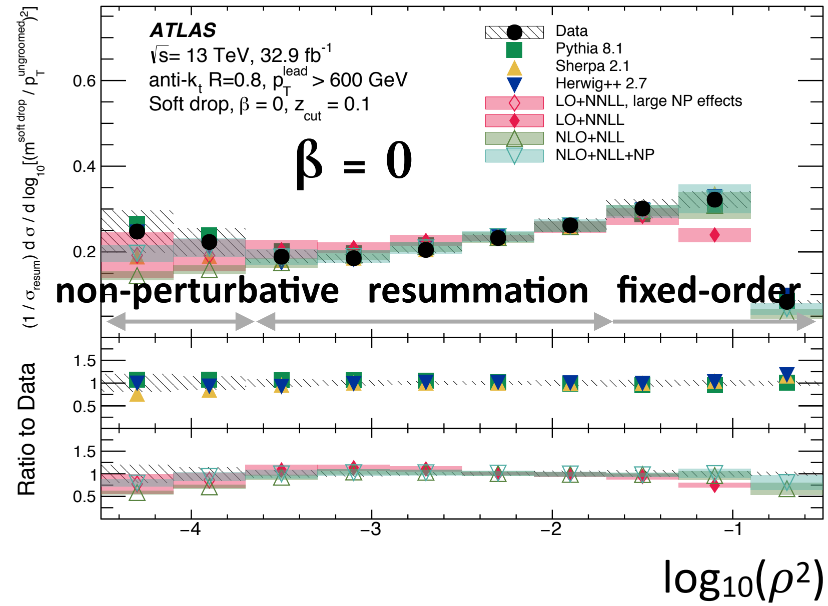


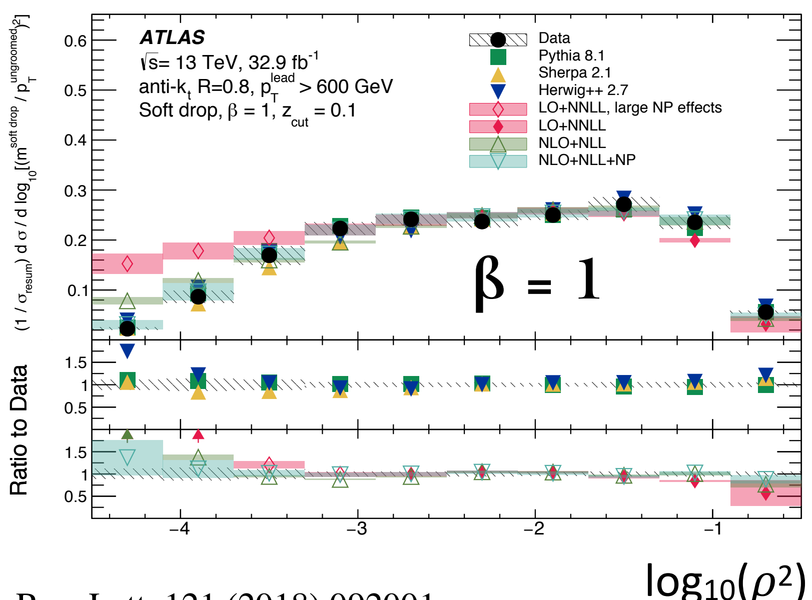
- Dominant Uncertainties
  - QCD Modelling: Pythia 8.186 vs Sherpa 2.1.1
    (Herwig++ 2.7.1 produces uncertainties of the same order)
  - calorimeter cluster energy scale shift (smearing) at low (high) ρ

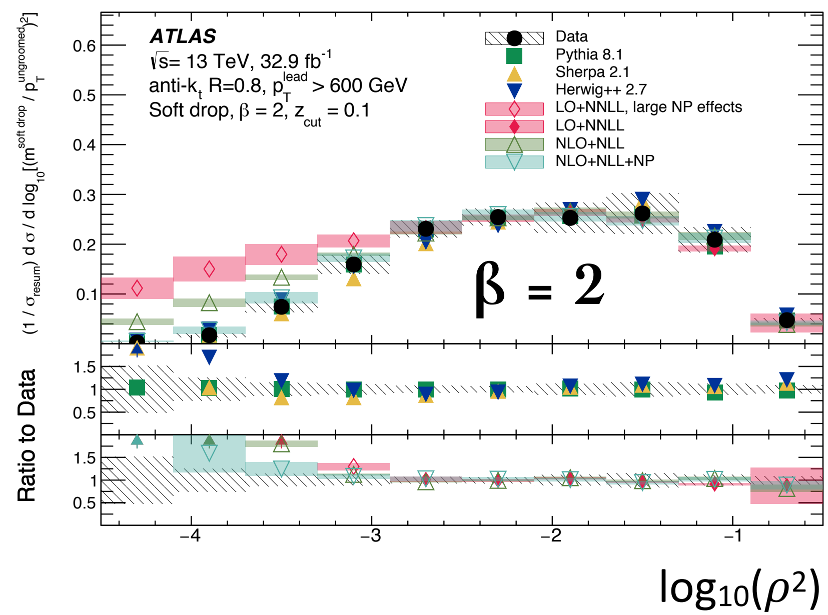
Phys. Rev. Lett. 121 (2018) 092001

Soft-drop Mass Results

- Mostly good agreement with generator predictions
    - The exception is at low values of ρ
    - To get good agreement requires LO+NNLL with large non-perturbative corrections.







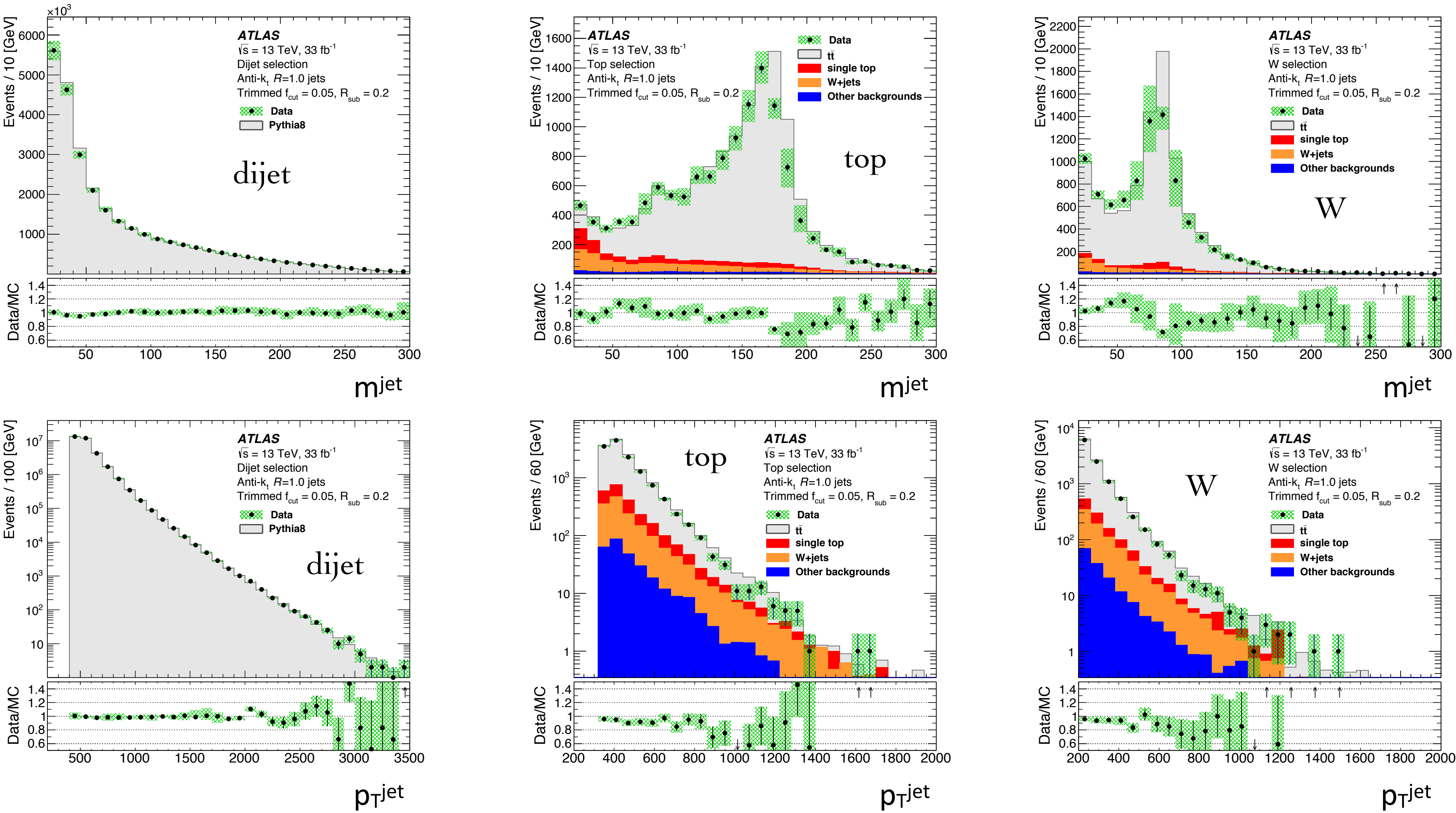
Phys. Rev. Lett. 121 (2018) 092001

Jet Substructure Measurements in Di-jet and semileptonic tt Events

- Two Data sets:
  - dijet: two large R=1 calorimeter jets
  - semileptonic tt: leptonic top and top-quark large-R jet or W-boson large-R jet
- Selection uses trimmed jets.
  - measurements made for both trimmed and soft-drop jets.
- Measure many different observables
  - relevant for W/top tagging.
  - Les Houches angularity
  - number of sub-jets (k_T-algorithm, $R = 0.2$ and $p_T > 10$ GeV)
  - energy correlation functions.
  - Ratios of N-subjettiness (see later for definition)

arXiv:1903.02942

Detector Level Comparisons

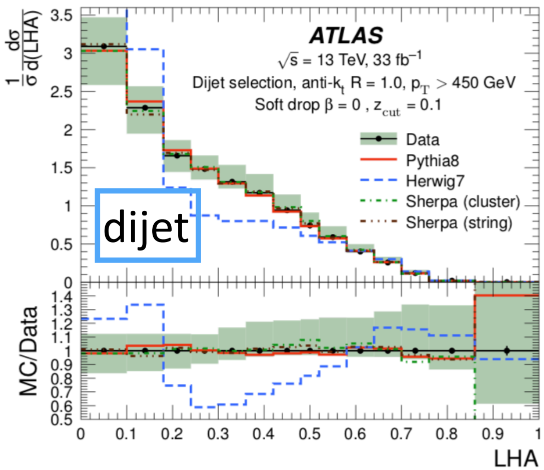


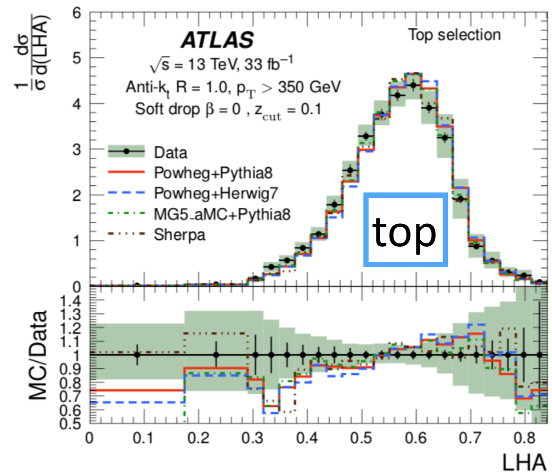
Good agreement for jet p_T. The discrepancy in the jet mass distributions is a known effect due to missing in-situ calibration and mass scale uncertainties.

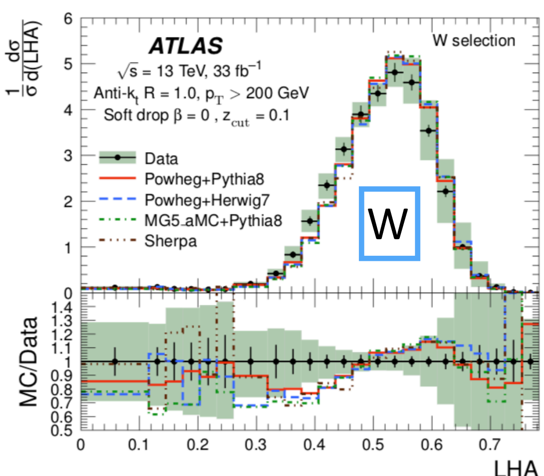
arXiv:1903.02942

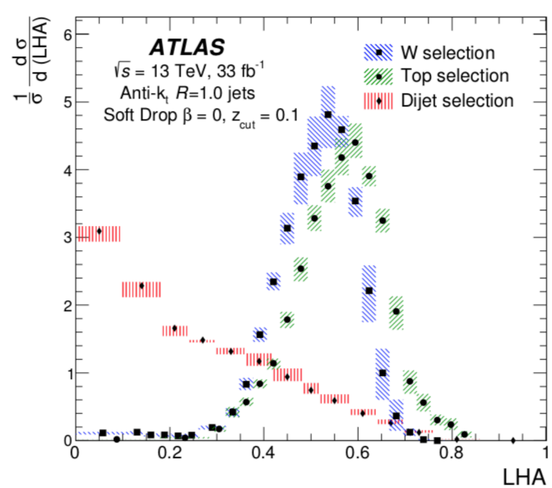
Les Houches Angularity

$$\lambda^{\kappa}_{\beta^{\mathrm{LHA}}} = \sum_{i \in J} z_i^{\kappa} \theta_i^{\beta^{\mathrm{LHA}}},$$







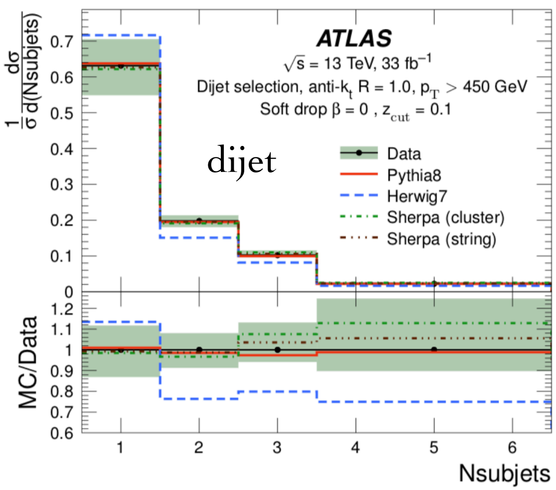


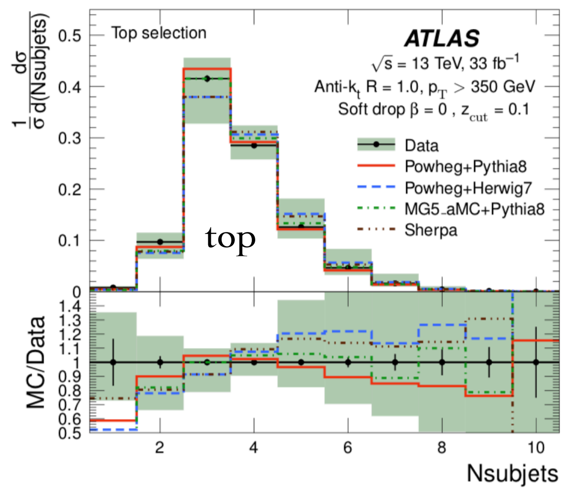
- where z_i is the transverse momentum fraction p_T^i/p_T^{jet}
  θ_i is the angle of the constituent relative to the jet axis
  and $\kappa = 1$ and $\beta^{LHA} = 0.5$
- It is an infrared-safe version of the jet-shape angularity, and provides a measure of the broadness of a jet.
- discrepancies in generator predictions for top and W selections
- Dijets: good agreement except for Herwig 7.
- Good discrimination between q/g and W/top jets.

arXiv:1903.02942

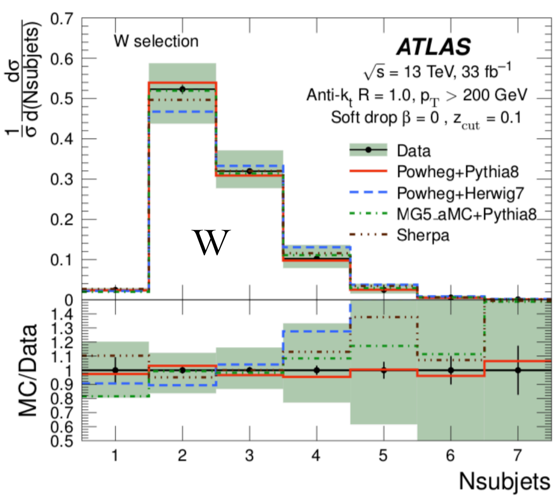
Subjet Multiplicity

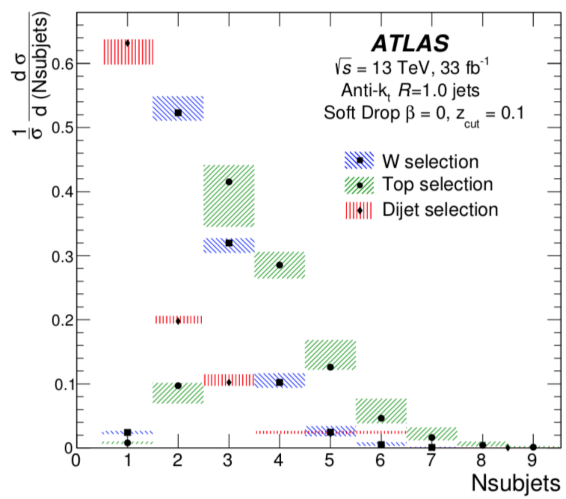
- number of sub-jets (k_T-algorithm, R = 0.2 and p_T > 10 GeV)





- Pythia8, Herwig7, Sherpa and MadGraph.
- Dijets: good agreement except for Herwig 7.
- W and top samples show good agreement for all generators.





- Largest contribution to uncertainties from mass and jet p_T calibration as well as calorimeter cluster corrections.
- Good discrimination between samples.

arXiv:1903.02942

Energy Correlation Functions

$$\mathrm{ECF1} = \sum_{i \in J} p_{\mathrm{T}_i}, \quad \text{(jet } p_{\mathrm{T}}\text{)}$$

$$\mathrm{ECF2}(\beta^{\mathrm{ECF}}) = \sum_{i<j \in J} p_{\mathrm{T}_i} p_{\mathrm{T}_j} \left(\Delta R_{ij}\right)^{\beta^{\mathrm{ECF}}}, \quad \text{(jet mass)}$$

$$\mathrm{ECF3}(\beta^{\mathrm{ECF}}) = \sum_{i<j<k \in J} p_{\mathrm{T}_i} p_{\mathrm{T}_j} p_{\mathrm{T}_k} \left(\Delta R_{ij} \Delta R_{ik} \Delta R_{jk}\right)^{\beta^{\mathrm{ECF}}},$$

- Use $\beta^{\mathrm{ECF}}=1$ for this analysis and use the correlation functions in ratios:

$$e_2 = \frac{\mathrm{ECF2}}{(\mathrm{ECF1})^2}, \qquad e_3 = \frac{\mathrm{ECF3}}{(\mathrm{ECF1})^3}.$$

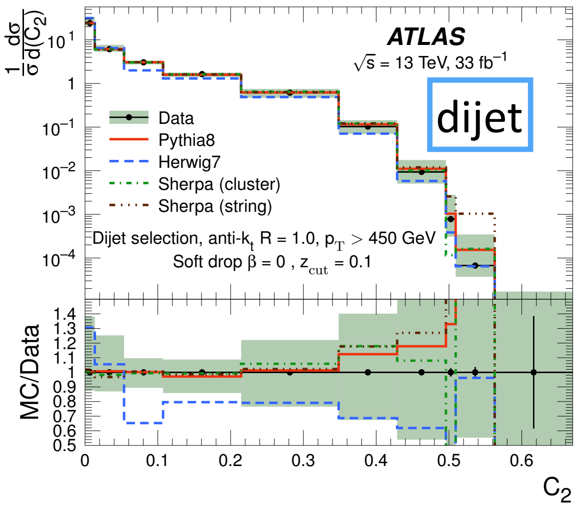
to form

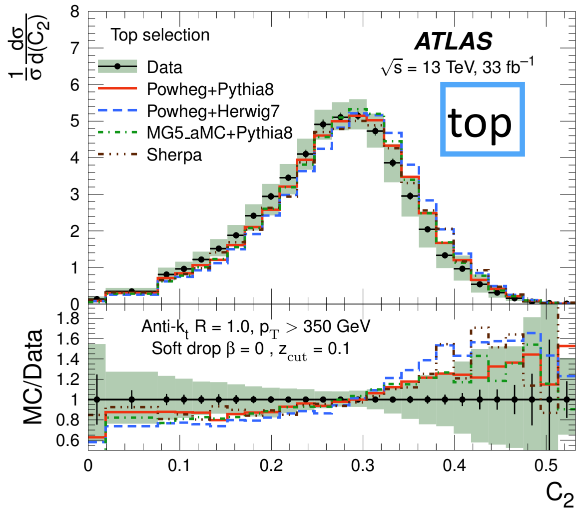
$$C_2 = \frac{e_3}{(e_2)^2}, \qquad D_2 = \frac{e_3}{(e_2)^3}.$$

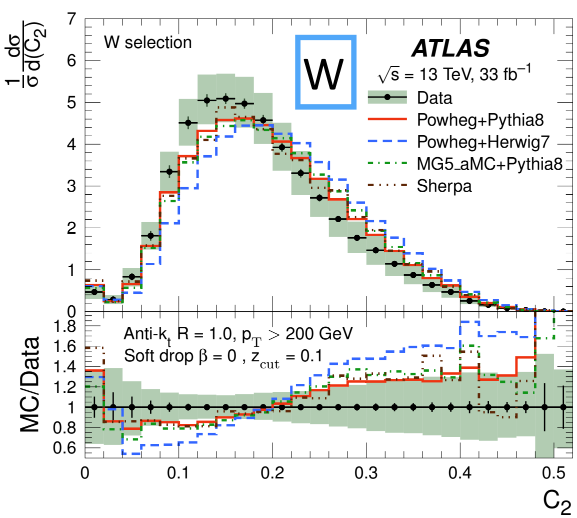
- which are useful for identifying two-body structures within jets..

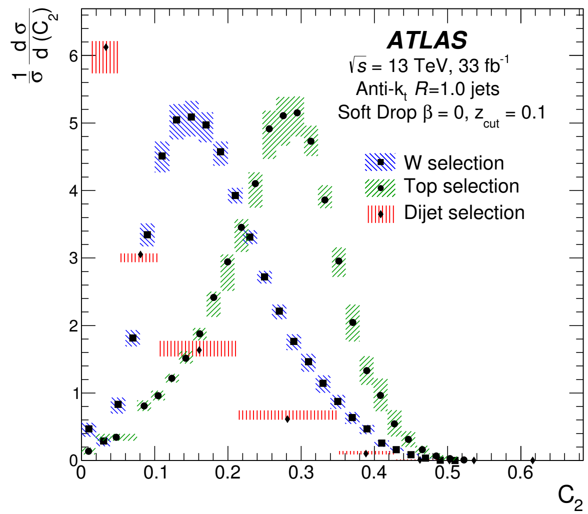
arXiv:1903.02942

Particle Level: C_2









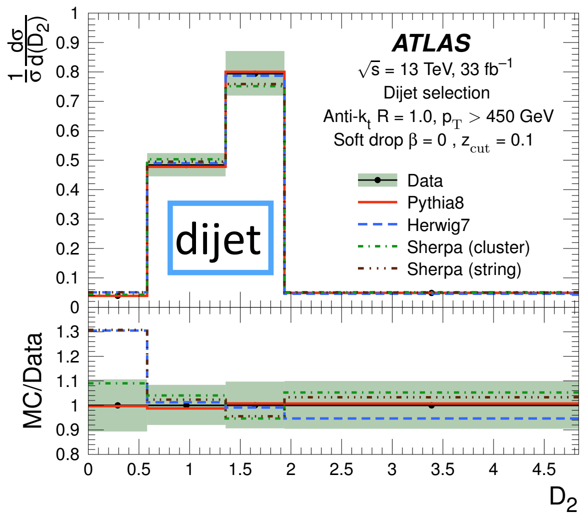
- Useful in W/top tagging
- Largest discrepancy for W-dataset.
- Excellent separation between q/g, W and top events.

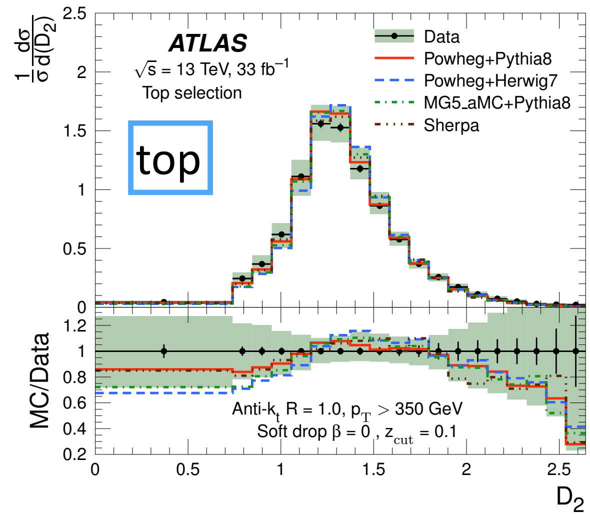
arXiv:1903.02942

ATLAS EXPERIMENT

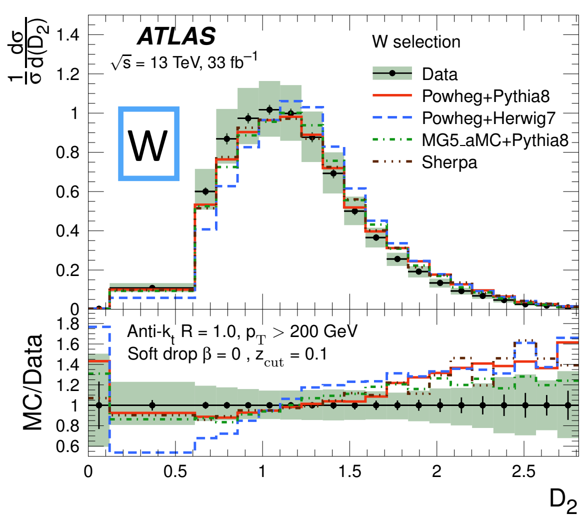
Lancaster University

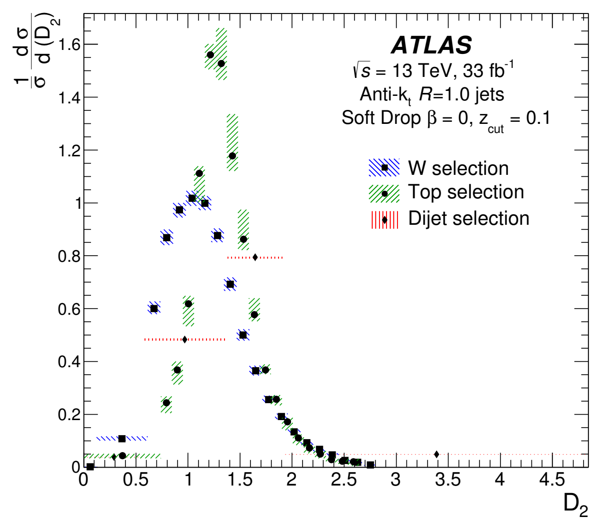
Particle Level: D_2





- Useful in W/top tagging
- Largest discrepancy for W-dataset.
- Good separation between q/g, W and top events.





arXiv:1903.02942

Subjettiness

- N-subjettiness describes to what degree the substructure of a given jet is compatible with being composed of N or fewer objects

$$\tau_0(\beta^{\mathrm{NS}}) = \sum_{i \in J} p_{\mathrm{T}_i} R^{\beta^{\mathrm{NS}}},$$

$$\tau_1(\beta^{\mathrm{NS}}) = \frac{1}{\tau_0(\beta^{\mathrm{NS}})} \sum_{i \in J} p_{\mathrm{T}_i} \Delta R_{a_1,i}^{\beta^{\mathrm{NS}}},$$

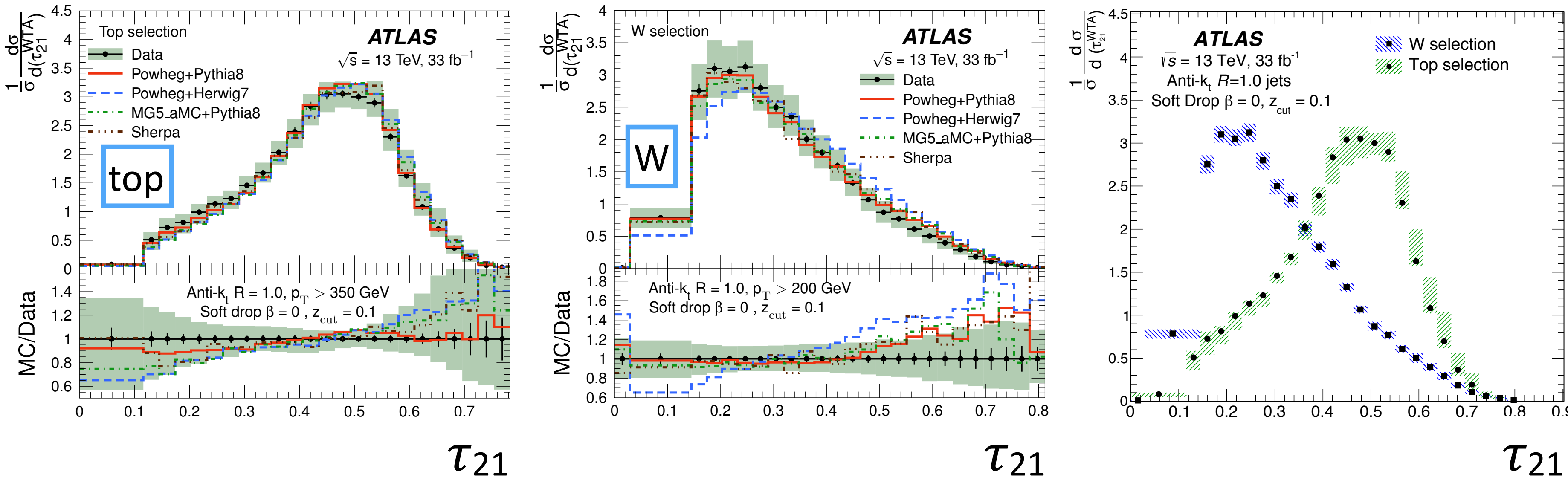
$$\tau_2(\beta^{\mathrm{NS}}) = \frac{1}{\tau_0(\beta^{\mathrm{NS}})} \sum_{i \in J} p_{\mathrm{T}_i} \min(\Delta R_{a_1,i}^{\beta^{\mathrm{NS}}}, \Delta R_{a_2,i}^{\beta^{\mathrm{NS}}}),$$

$$\tau_3(\beta^{\mathrm{NS}}) = \frac{1}{\tau_0(\beta^{\mathrm{NS}})} \sum_{i \in J} p_{\mathrm{T}_i} \min(\Delta R_{a_1,i}^{\beta^{\mathrm{NS}}}, \Delta R_{a_2,i}^{\beta^{\mathrm{NS}}} \Delta R_{a_3,i}^{\beta^{\mathrm{NS}}}),$$

- R is the radius parameter of the jet,
  β^{NS}=1 is the angular separation weight
  $\Delta R_{a,n}$ is the angular distance between constituent i and the axis of the n^{th} subjet
- The ratios: $\tau_{21} = \tau_2/\tau_1$ and $\tau_{32} = \tau_3/\tau_2$ are useful in identifying two- and three-body structures in jets.
- Calculated using the direction of the hardest constituent in the subjet (WTA)

arXiv:1903.02942

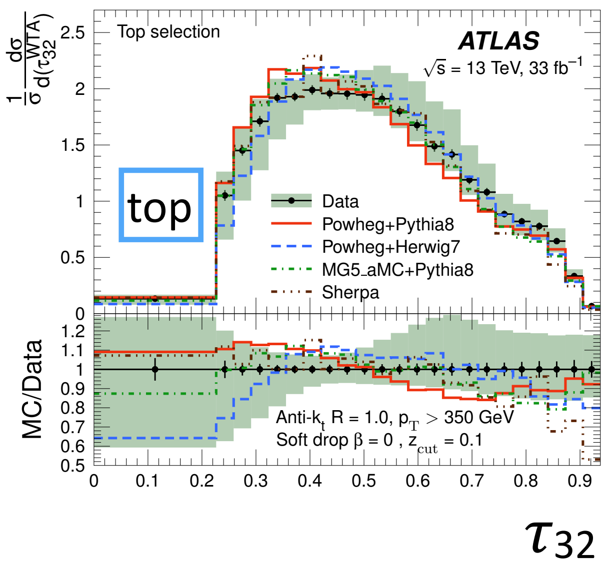
Subjettiness: τ_{21}

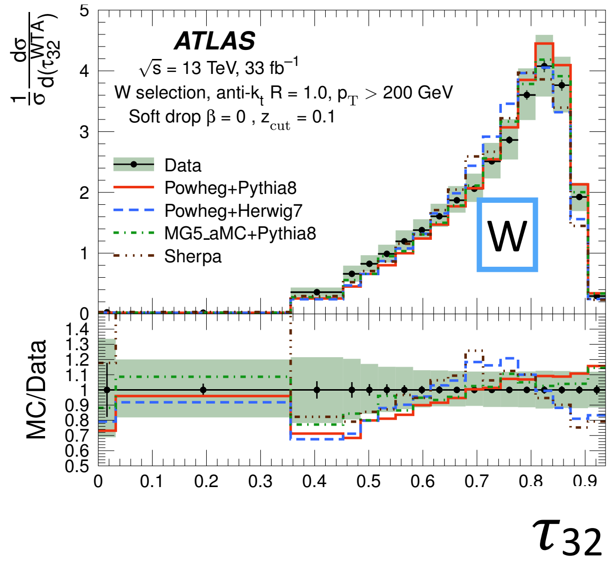


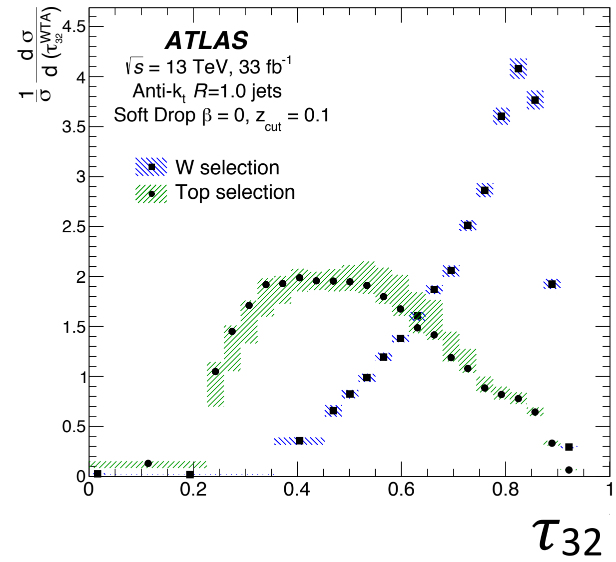
- Used in the top/W taggers
  - larger discrepancies for the W selection
  - Excellent discrimination between W and top events
- Not used on the dijet sample as these events have less hard splitting and are highly sensitive to the splitting and merging of clusters.

arXiv:1903.02942

Subjettiness: τ_{32}







- Used in the top/W taggers
  - larger discrepancies for the W selection
  - Excellent discrimination between W and top events
- Not used on the dijet sample as these events have less hard splitting and are highly sensitive to the splitting and merging of clusters.

arXiv:1903.02942

Conclusions

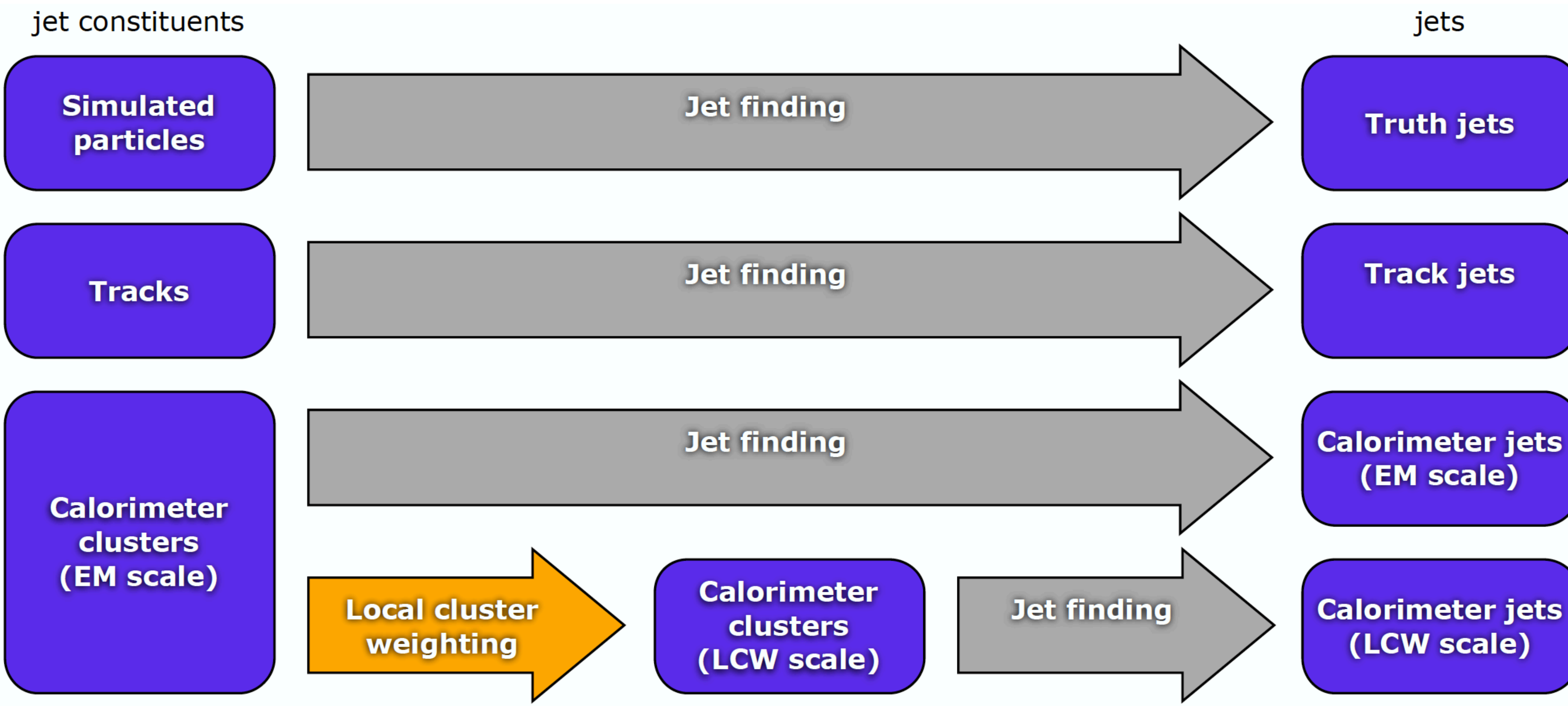
- jet substructure observables are widely used for W/top/H tagging in BSM searches
  - used in cut-based and multivariate taggers
- jet substructure uncertainties are of the order of 10 to 20% and allows
  - small enough to differentiate between the various MC models
- soft-drop groomer allows a comparison to NLO+NNL and LO+NNNL predictions
  - important to probe a new regime of QCD at LHC
  - improve the understanding of jet substructure properties
- dedicated measurements validate MC models, also show larger discrepancies for some observables
  - results can be used to tune MC generators!

Backup

Bibliography

- *Measurement of jet-substructure observables in top quark, W boson and light jet production in proton-proton collisions at √s=13 TeV with the ATLAS detector*, arXiv:1903.02942, submitted to JHEP, (Link to ATLAS Repository).
- *A measurement of the soft-drop jet mass in pp collisions at √s=13 TeV with the ATLAS detector,* Phys. Rev. Lett. 121 (2018) 092001, (Link to ATLAS Repository).
- *In situ calibration of large-R jet energy and mass in 13 TeV proton-proton collisions with the ATLAS detector*, Eur. Phys. J. C 79 (2019) 135, (Link to ATLAS Repository).
- *A measurement of the soft-drop jet mass in pp collisions at √s=13 TeV with the ATLAS detector*, Phys. Rev. Lett. 121 (2018) 092001, (Link to ATLAS Repository).

Jet Reconstruction



Calibrates clusters based on cluster properties related to shower development

Data Selection for Substructure

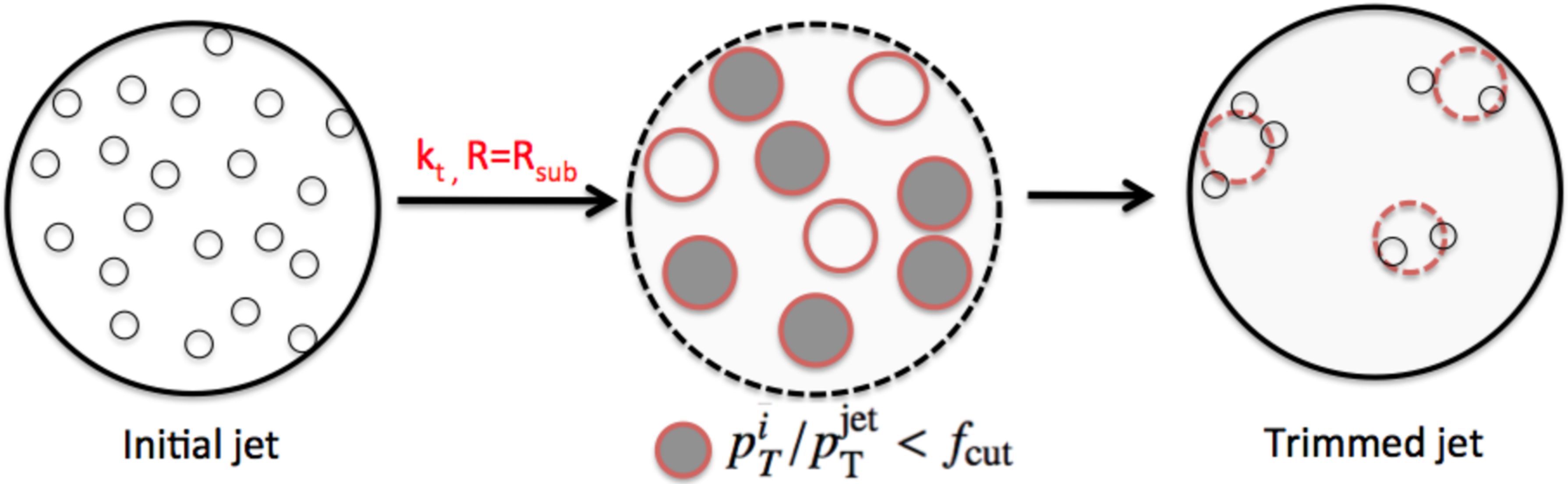
		Detector level	Particle level
	Dijet selection:		
Dijets	Two trimmed anti-k_t $R = 1.0$ jets	$p_{\mathrm{T}} > 200$ GeV $\lvert\eta\rvert < 2.5$	$p_{\mathrm{T}} > 200$ GeV $\lvert\eta\rvert < 2.5$
Dijets	Leading-p_{T}trimmed anti-k_t $R = 1.0$ jet	$p_{\mathrm{T}} > 450$ GeV	
	Top and W selections:		
Semiletonic tt	Exactly one muon	$p_{\mathrm{T}} > 30$ GeV $\lvert\eta\rvert < 2.5$ $\lvert z_0 \sin(\theta)\rvert < 0.5$ mm and $\lvert d_0/\sigma(d_0)\rvert < 3$	$p_{\mathrm{T}} > 30$ GeV $\lvert\eta\rvert < 2.5$
Semiletonic tt	Anti-k_t $R = 0.4$ jets	$p_{\mathrm{T}} > 25$ GeV $\lvert\eta\rvert < 4.4$ JVT output > 0.5 (if $p_{\mathrm{T}} < 60$ GeV)	$p_{\mathrm{T}} > 25$ GeV $\lvert\eta\rvert < 4.4$
Semiletonic tt	Muon isolation criteria	If $\Delta R(\mu, \mathrm{jet}) < 0.04 + 10$ GeV $/p_{\mathrm{T},\mu}$: muon is removed, so the event is discarded	None
Semiletonic tt	$E_{\mathrm{T}}^{\mathrm{miss}}$, $m_{\mathrm{T}}^{\mathrm{W}}$	$E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV , $E_{\mathrm{T}}^{\mathrm{miss}} + m_{\mathrm{T}}^{\mathrm{W}} > 60$ GeV	
Semiletonic tt	Leptonic top	At least one small-radius jet with $0.4 < \Delta R(\mu, \mathrm{jet}) < 1.5$	
	Top selection:		
hadronic top	Leading-p_{T}trimmed anti-k_t $R = 1.0$ jet	$\lvert\eta\rvert < 1.5$, $p_{\mathrm{T}} > 350$ GeV , mass > 140 GeV $\Delta R(\text{large−radius jet}, b\text{-tagged jet}) < 1$ $\Delta\phi(\mu, \text{large−radius jet}) > 2.3$	
	W selection:		
hadronic W	Leading-p_{T}trimmed anti-k_t $R = 1.0$ jet	$\lvert\eta\rvert < 1.5$, $p_{\mathrm{T}} > 200$ GeV , mass > 60 GeV and mass < 100 GeV $1 < \Delta R(\text{large−radius jet}, b\text{-tagged jet}) < 1.8$ $\Delta\phi(\mu, \text{large−radius jet}) > 2.3$	

Monte Carlo Generators

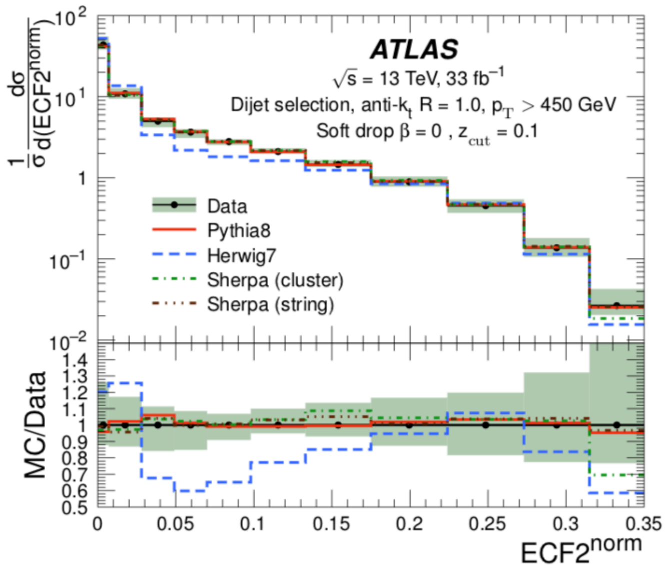
Process	Generator	Version	PDF	Tune	Use
Dijet	Pythia8	8.186	NNPDF23LO	A14	Nominal for unfolding
	Sherpa	2.2.1	CT10	Default	Validation of unfolding (with two different hadronisation models)
	Herwig7	7.0.4	MMHT2014	H7UE	Comparison
$t\bar{t}$	Powheg	v2	NNPDF30NLO		Nominal for unfolding
	+ Pythia8	8.186	NNPDF23LO	A14	
	Powheg	v2	CT10		Validation of unfolding
	+Herwig++	2.7	CTEQ6L1	UE-EE-5 tune	
	Powheg	v2	CT10		Comparison
	+Herwig7	7.0.4	MMHT2014	H7UE	
	MG5_aMC@NLO	2.6.0	NNPDF30NLO		Comparison
	+ Pythia8	8.186	NNPDF23LO	A14	
	Sherpa	2.2.1	CT10	Default	Comparison
Single top	Powheg	v1	CT10		Nominal for unfolding
	+ Pythia6	6.428	CTEQ6L1	Perugia2012	
Z+jets	Sherpa	2.2.1	CT10	Default	Background estimation
W+jets	Sherpa	2.2.1	CT10	Default	Background estimation (nominal)
W+jets	MG5_aMC@NLO	2.2.5	CT10		Background estimation (cross-check)
	+ Pythia8	8.186	NNPDF23LO	A14	
Diboson	Sherpa	2.2.1	CT10	Default	Background estimation

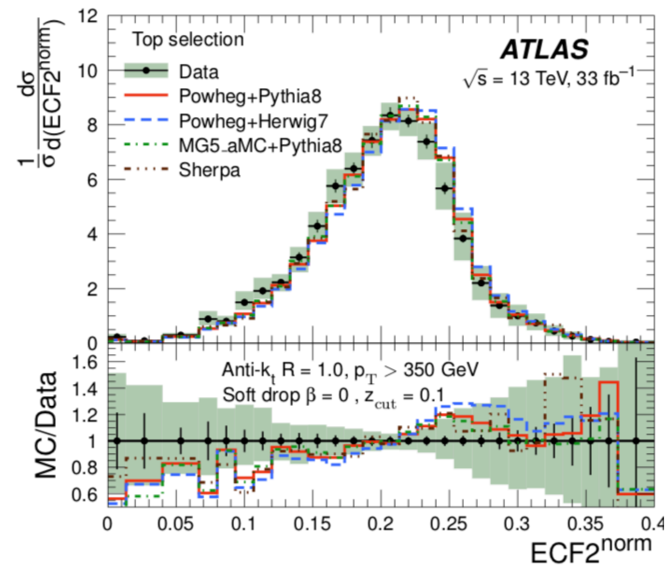
Jet Trimming

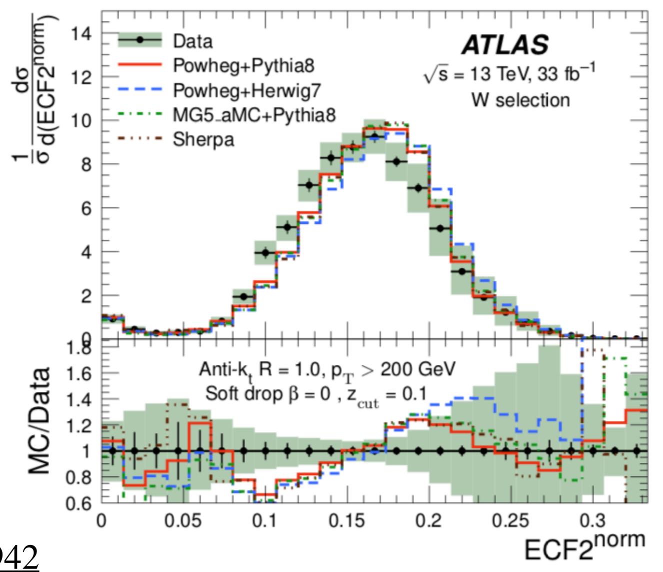
- Default grooming procedure: trimming, removes R = 0.2 subjets with p_T < 5% of jet p_T

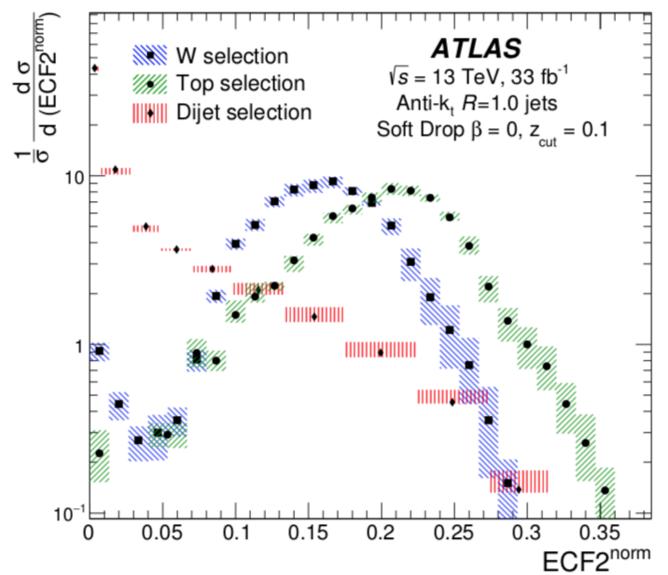


Particle Level: ECF2



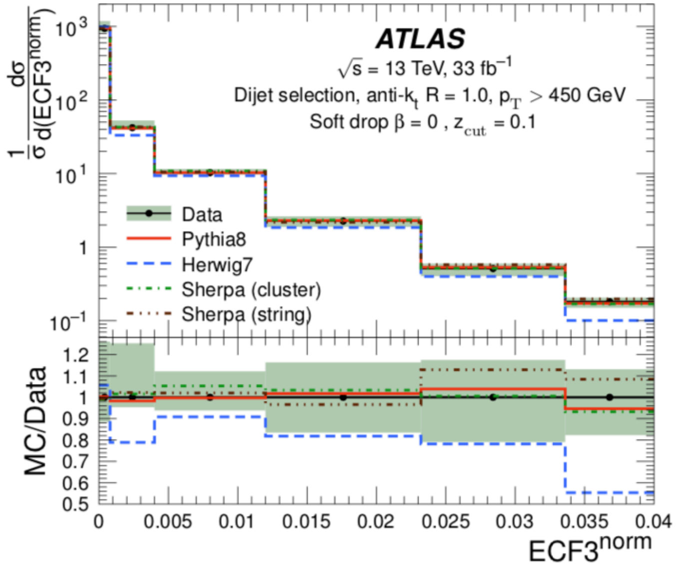


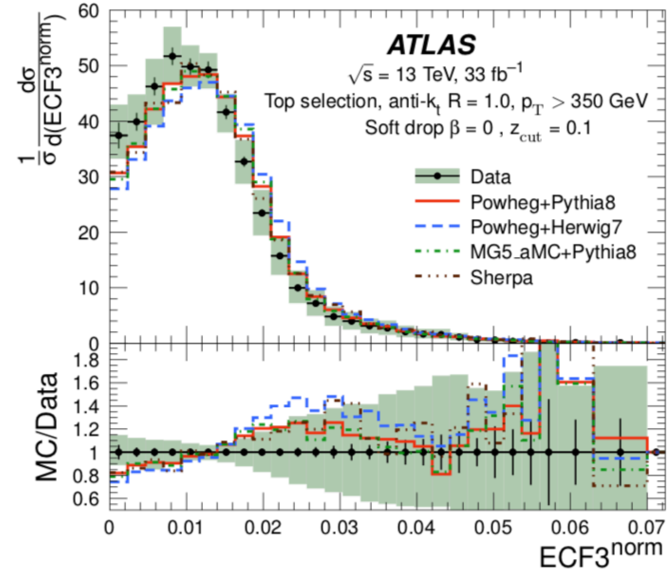


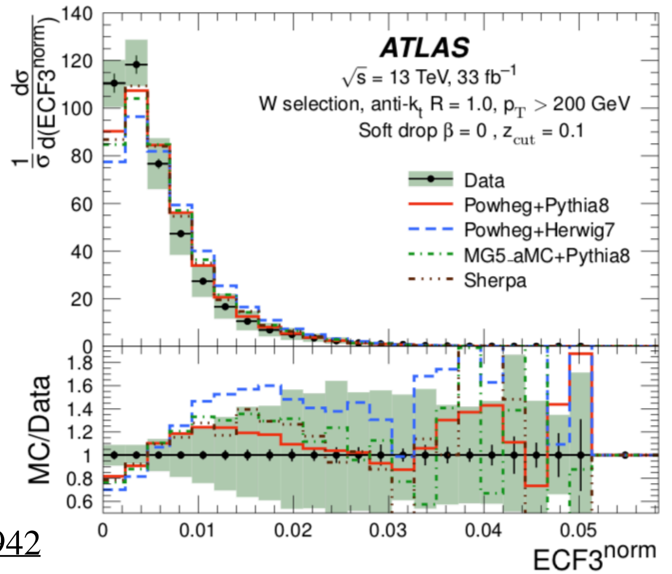


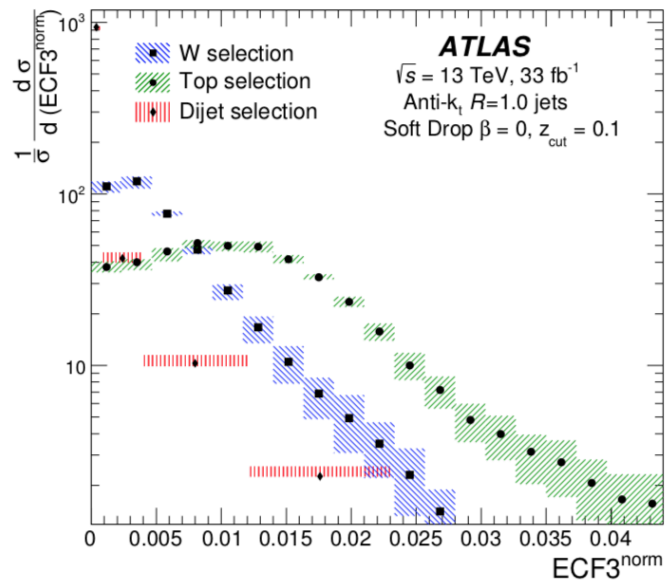
arXiv:1903.02942

Particle Level: ECF3









arXiv:1903.02942