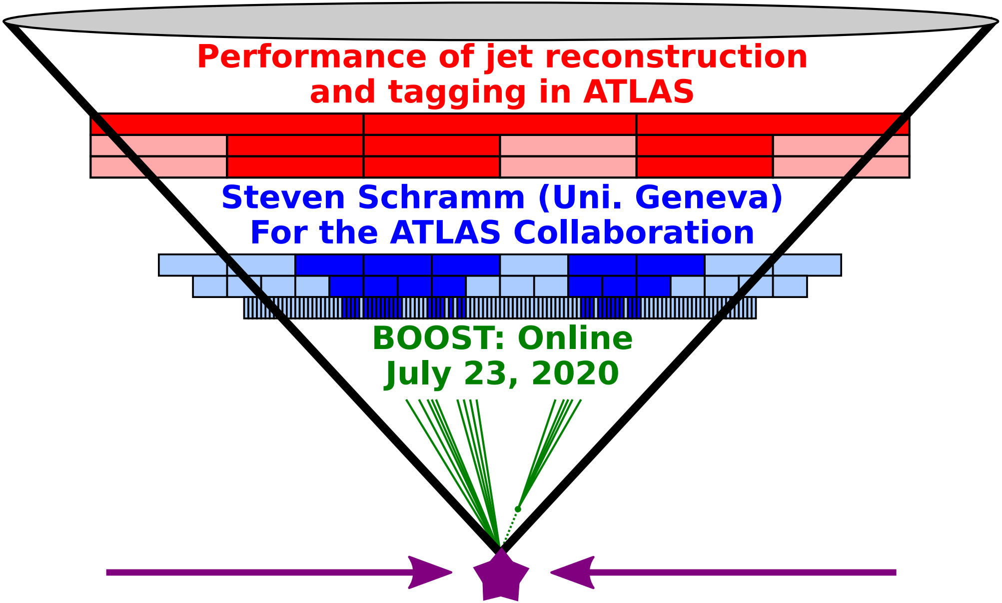


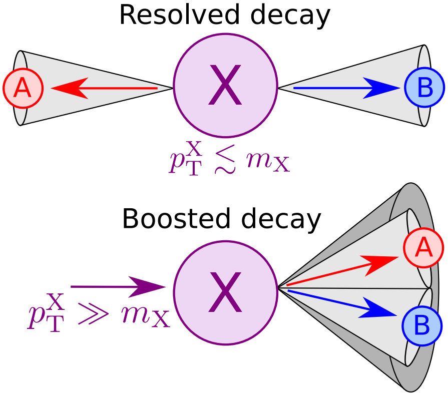
Performance of jet reconstruction and tagging in ATLAS

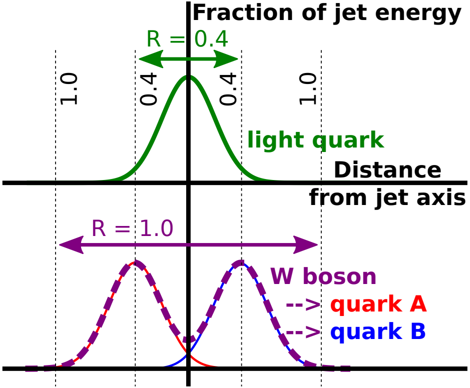
Steven Schramm (Uni. Geneva)
For the ATLAS Collaboration

BOOST: Online
July 23, 2020

Introduction and jet types

- Jet substructure and tagging is increasingly used throughout the ATLAS physics program
  - Includes *b*-tagging (typically $R = 0.4$ jets) and boosted decay tagging (typically $R = 1.0$ jets)
- This talk will cover both, with an emphasis on boosted decay substructure and tagging

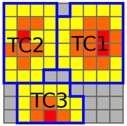




Inputs to jet reconstruction

- The vast majority of jets in ATLAS are built using the anti-k_t algorithm, but inputs differ
- Traditionally, jets have been built from topo-clusters: calorimeter 4-vector
  - Typically $R = 0.4$ jets use EM topo-clusters, while $R = 1.0$ jets use LC topo-clusters
- Increasing move for $R = 0.4$ jets toward particle flow objects: calorimeter+track 4-vector
  - Similar or improved performance ∼everywhere, especially useful at low p_{T}
- Sometimes $R = 1.0$ jets use Track-CaloClusters: calorimeter energy, track angle
  - Improves W/Z tagging performance at high p_{T}, for some trade-offs

Topo-clusters (EM or LC)
[Traditional ATLAS usage]



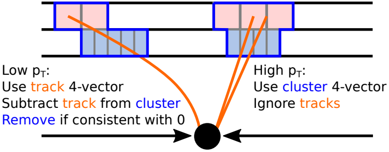
🟥 seed cells $|E| > 4\sigma$

🟧 growth cells $|E| > 2\sigma$

🟨 boundary cells

☐ final topoclusters $\eta \times \phi = \text{dynamic}$

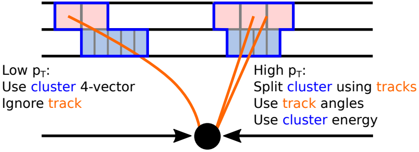
Particle Flow Objects (PFOs)
[More recent ATLAS usage]



Low p_{T}:
Use track 4-vector
Subtract track from cluster
Remove if consistent with 0

High p_{T}:
Use cluster 4-vector
Ignore tracks

Track-CaloClusters (TCCs)
[Used for highest p_{T} results]



Low p_{T}:
Use cluster 4-vector
Ignore track

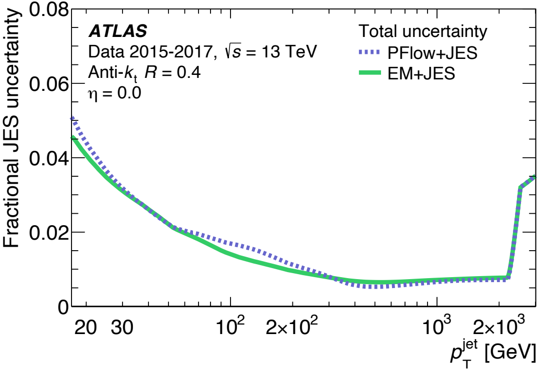
High p_{T}:
Split cluster using tracks
Use track angles
Use cluster energy

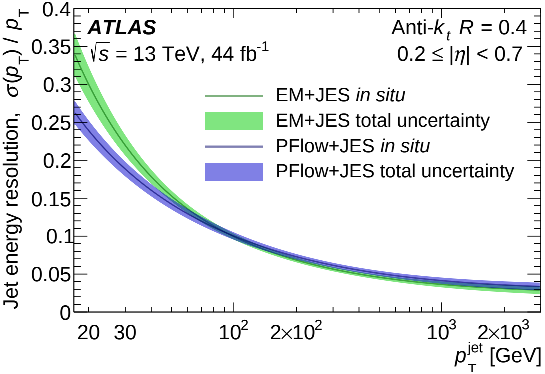
$R = 0.4$ jet energy scale and resolution

New @ BOOST2020

Both: JETM-2018-05

- *In situ* JES combination reaches the level of 1% precision over large range of p_T
  - PFlow and topo-cluster jets have very comparable JES uncertainties
  - Uncertainties stable at low p_T compared to 8 TeV, despite much higher pileup - see backup
- PFlow significantly improves the low p_T JER and its uncertainties vs topo-clusters
  - Resolution uncertainties are now typically below 1% absolute (5–15% relative - see backup)



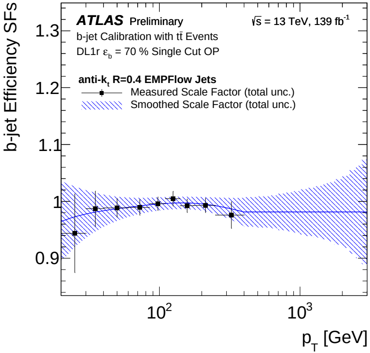


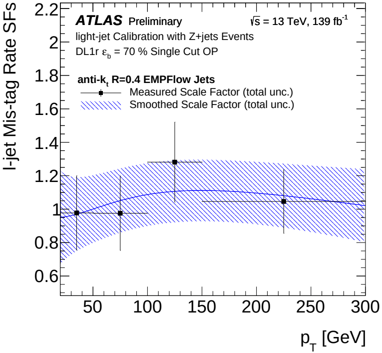
b-tagging of PFlow jets

Recent result (April)

Both: FTAG-2020-001

- We have just seen that PFlow jets can improve upon topo-cluster jets
  - It is thus important to also consider potential *b*-tagging implications
- To first order, *b*-tagging is independent of the exact jet definition
  - Inputs depend primarily on the tracks near the jet, not the jet itself
- Still important to derive new tagging scale factors to account for data/MC agreement



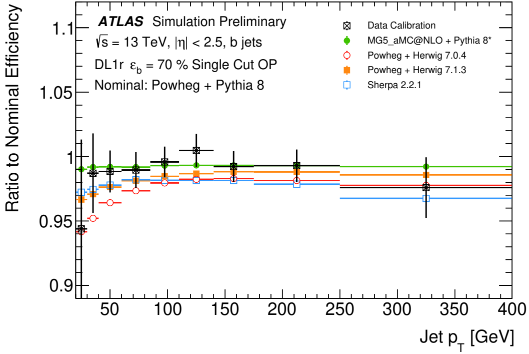


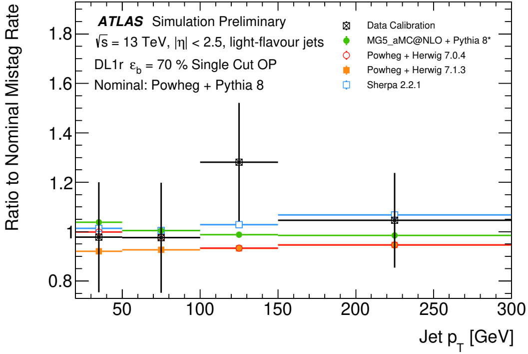
MC to MC *b*-tagging scale factors

Recent result (May)

Both: PUB-2020-009

- Correct for differences between generators with MC-to-MC scale factors
  - Differences can be quite large, theory inputs have a significant impact
- Precise calculations/predictions are crucial to continue improving tagger discrimination
  - Especially true for parton shower, fragmentation+hadronization, and decays



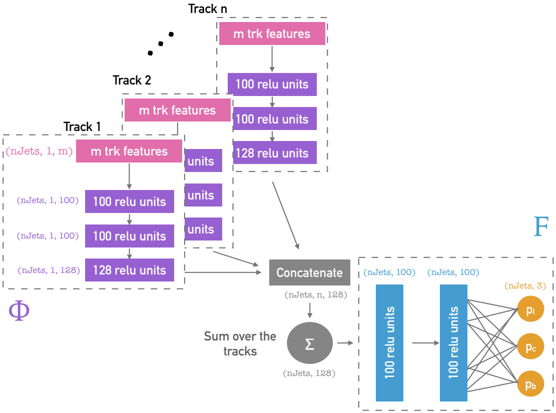


Deep Impact Parameter Sets (DIPS)

Recent result (May)

Plot: PUB-2020-014

- Run 1 and Run 2: ATLAS used a 3D impact parameter (IP) likelihood
    - IP3D was a key input to *b*-tagging
- In 2017, ATLAS studied an RNN for impact parameter (IP) tagging
    - RNN requires defining an ordering
    - Sequential operation, not parallelizable
- Now studied a Deep IP Set (DIPS)
    - Permutation invariant, parallelizable
    - Faster to train and evaluate!
- However, what about performance?

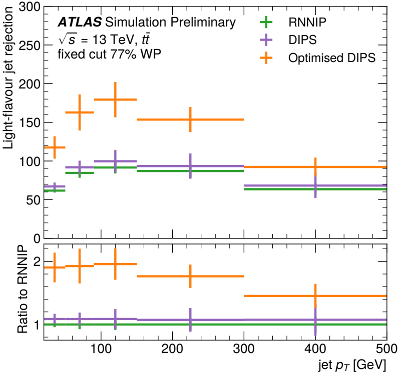


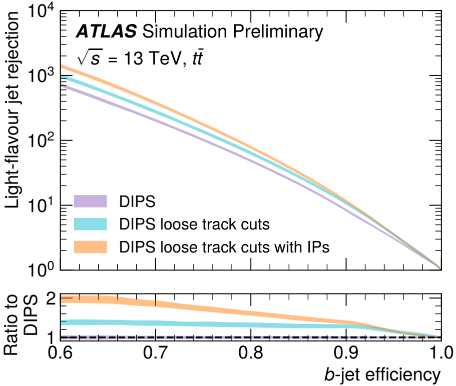
DIPS *b*-tagging performance

Recent result (May)

Both: PUB-2020-014

- With the same inputs, DIPS already matched or exceeded the performance of RNNIP
- DIPS input optimization led to significant further gains
  - Up to 50% gain from loose track selection, up to 100% with additional IP inputs

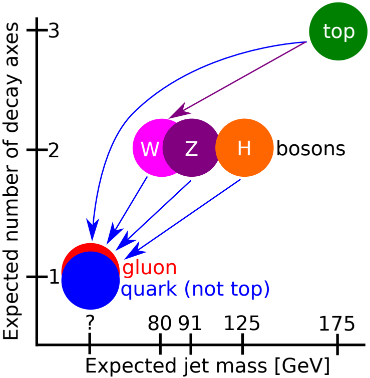


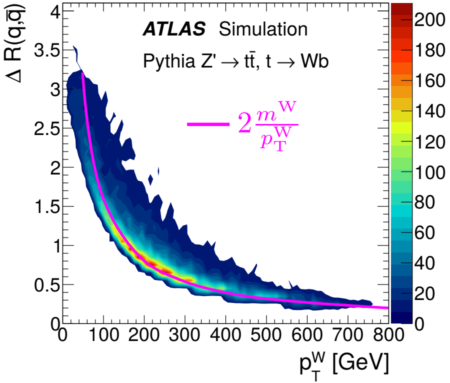


Boosted hadronic decays

Right: PERF-2012-02

- Hadronic decays of massive particles can come from many sources
  - Main examples: $W \to qq'$, $Z \to q\bar{q}$, $H \to b\bar{b}$, top $\to bW \to bqq'$
- For high p_T parent particles, the decay products are collimated and fall into a single jet
  - Basic idea of tagging: look at the jet mass and number of decay product axes

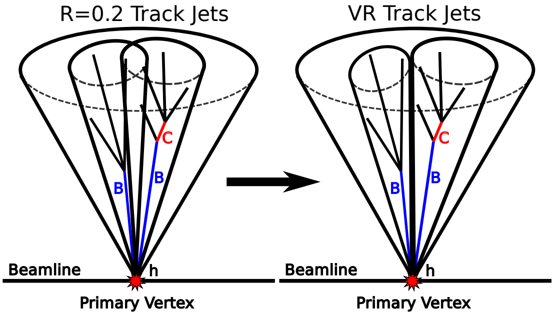


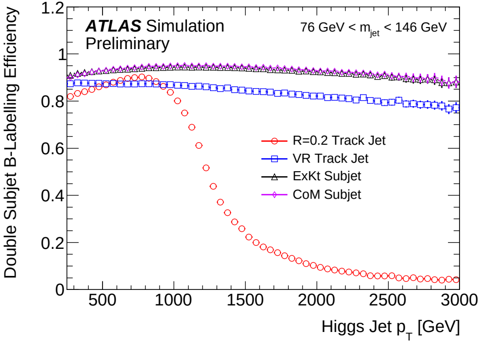


Hbb tagging

Both: PUB-2017-010

- Let's start with a $H \rightarrow bb$ tagger: most important aspect is double-b-tagging
  - Identifies that there are two decay axes, and that they both are compatible with b-quarks
- Using $R = 0.4$ jets won't work: doesn't contain full H decay, instead use $R = 1.0$
  - Then geometrically match smaller jets to the $R = 1.0$ jet and b-tag those small jets
  - Even $R = 0.2$ can be too large, instead use Variable-Radius (VR) track-jets: R shrinks vs p_T

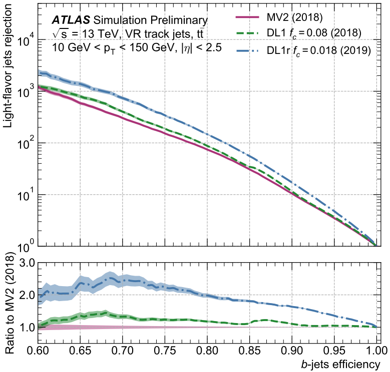


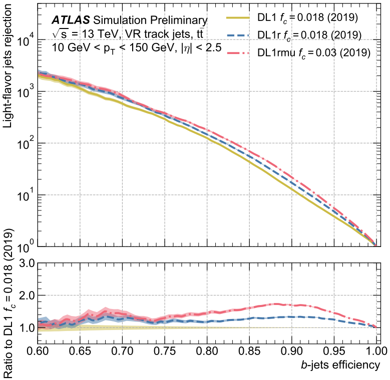


Variable-radius track-jet *b*-tagging

Both: FTAG-2019-006

- VR track-jet *b*-tagging is a key component to the boosted hadronic physics program
  - Dedicated VR track-jet optimization improved *b*-tag performance by roughly a factor of 2
  - Additional input variables can extend this even further for high *b*-tagging efficiency



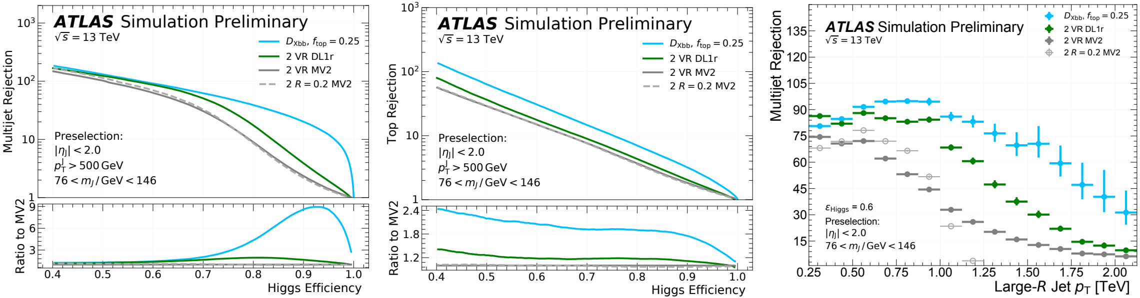


Hbb tagging performance

New @ BOOST2020

Both: PUB-2020-019

- The dedicated VR track-jet optimization is directly beneficial for $H \to bb$ tagging
  - Substantial improvements for double-tagging Higgs jets using optimized vs old b-taggers
- Even more significant improvements are possible with a dedicated $X \to bb$ tagger
  - Fully-connected DNN taking the DL1r scores $(\mathcal{P}_l, \mathcal{P}_c, \mathcal{P}_b)$ for three associated VR track-jets
  - Light-jet rejection improves at high p_T (shown), top-jet rejection mostly p_T-stable [backup]

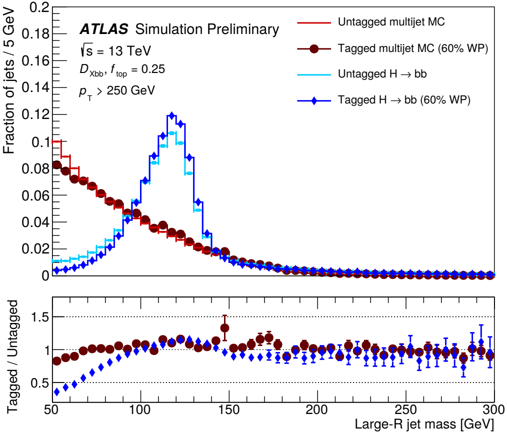


Hbb tagging and mass sculpting

New @ BOOST2020

Plot: PUB-2020-019

- Ideally want the $H \to bb$ tagger (D_{Xbb}) not to shape Higgs or background mass
    - D_{Xbb} is not given any substructure variables (other than the three subjets)
    - This limits the mass sculpting
- Important for $H \to bb$ measurements in the boosted regime
    - Good identification performance with manageable mass sculpting!
    - Sculpting consistent with asking for matched VR jets (typically done)



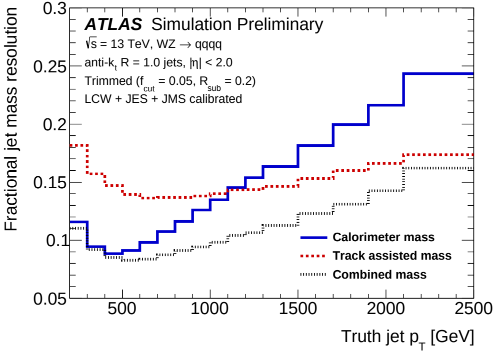
Large-radius jet mass

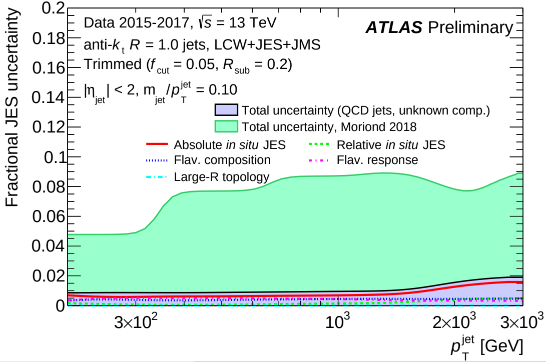
Left: JETM-2019-05

Right: JETM-2017-002

- $H \to bb$ tagging is a rare case where the jet mass is less important for tagging
  - Still important to identify as a Higgs (reduce $Z \to bb$ or $g \to bb$), but less than other cases
- Use "combined mass" to improve high p_{T} performance: $m_{\mathrm{jet}}^{\mathrm{comb}} = A \cdot m_{\mathrm{jet}}^{\mathrm{calo}} + B \cdot m_{\mathrm{jet}}^{\mathrm{TA}}$

  $m_{\mathrm{jet}}^{\mathrm{TA}} = m_{\mathrm{jet}}^{\mathrm{track}} \cdot \frac{p_{\mathrm{T}}^{\mathrm{calo}}}{p_{\mathrm{T}}^{\mathrm{track}}}$, A and B are resolution weights derived in QCD dijet MC samples
- ATLAS calibrates the $R = 1.0$ jet mass scale (JMS) after applying the JES



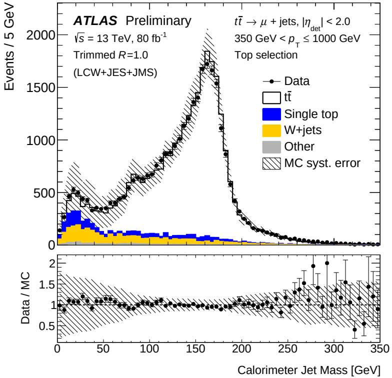


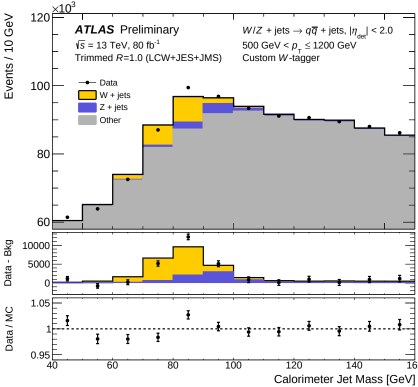
Forward folding: jet mass scale and resolution

New @ BOOST2020

Plot: CONF-2020-022

- Use visible mass peaks to precisely calibrate mass differences between data and MC
  - Traditionally top quarks and W bosons from $t\bar{t}$ events, now also V from $V+$jets events
- Use MC templates in Forward Folding: $m^{\text{fold}} = s \cdot m^{\text{reco}} + (m^{\text{reco}} - m^{\text{truth}} \cdot \mathcal{R}_m)(r - s)$
  - s (r) is the JMS (JMR) difference between data and MC, $\mathcal{R}_m$ is the MC mass response



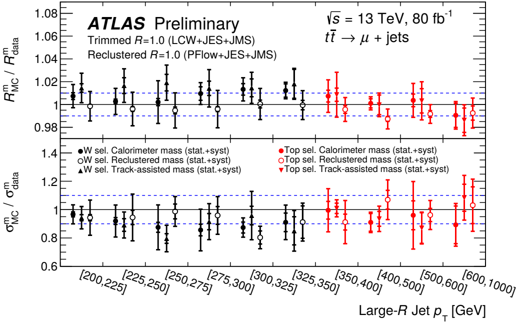


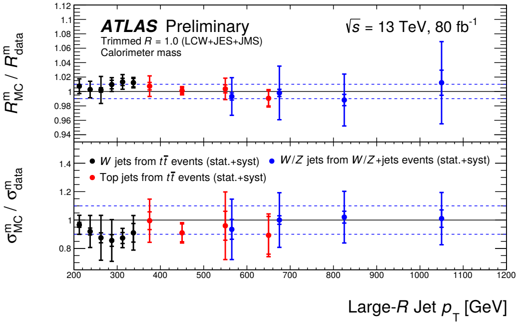
JMS and JMR results

New @ BOOST2020

Plot: CONF-2020-022

- Extracted the *in situ* JMS and JMR (after applying the *in situ* JES) using forward folding
  - Left: W and top jets from $t\bar{t}$ events, for calorimeter + track-assisted + reclustered masses
    - Calorimeter mass: from jet constituents (topo-clusters); track-assisted mass: $m_{\text{jet}}^{\text{track}} \cdot \frac{p_{\text{T}}^{\text{calo}}}{p_{\text{T}}^{\text{track}}}$
    - Reclustered mass: build $R = 1.0$ jets using $R = 0.4$ jets as input, take the mass of that jet
  - Right: Calorimeter mass for W jets ($t\bar{t}$), top jets ($t\bar{t}$), and W/Z jets (V + jets)
- JMS is consistent with 1 (within ∼1% uncertainty), JMR sometimes underestimated



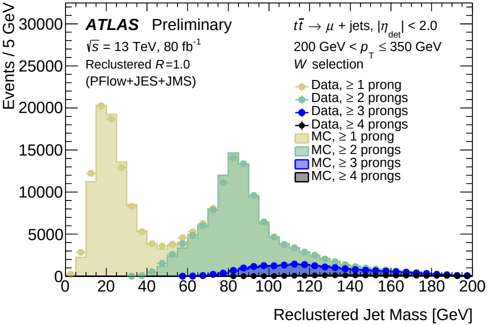


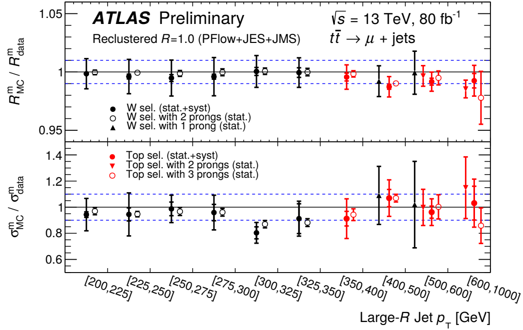
JMS and JMR dependence on jet structure

New @ BOOST2020

Plot: CONF-2020-022

- Also used Forward Folding to study the dependence of the JMS and JMR on substructure
- Reclustered jets: $R = 1.0$ jets built using inputs of fully-calibrated $R = 0.4$ jets
  - Count the number of $R = 0.4$ jets in the reclustered jet
  - Compare different $R = 0.4$ jet multiplicities for a given particle interpretation
- No significant differences observed $\Longrightarrow$ JMS and JMR appear stable vs substructure

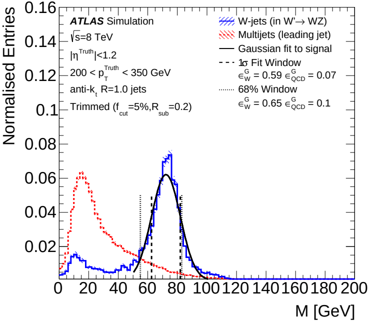


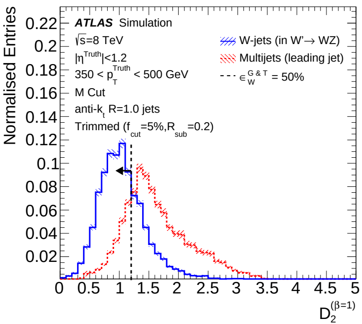


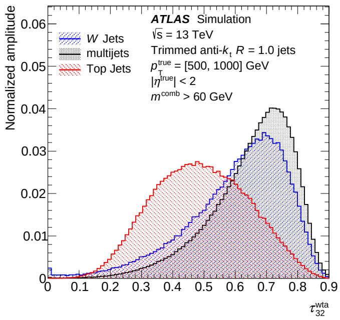
Going beyond the jet mass for W/Z/top decays

Left: PERF-2015-03
Mid: PERF-2015-03
Right: JETM-2018-03

- As mentioned earlier, to first order tagging is a simple procedure
  - Cut on the jet mass and other substructure variable(s) correlated to the number of decay axes
  - Common variables: $D_2^{(\beta=1)}$ for two-body decays (W/Z), τ_{32}^{wta} for three-body decays (top)
- Such simple two-variable cut-based taggers provide a strong reference point
  - Typical for early run 2, but we have moved to more powerful combinations





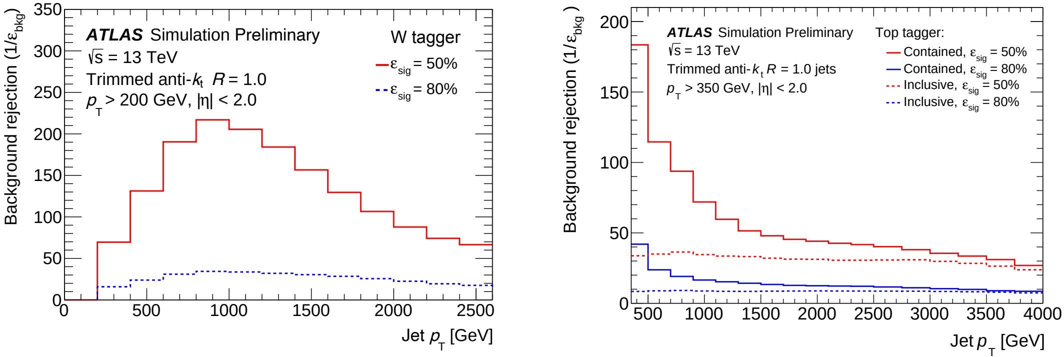


W and top tagger performance in MC

New @ BOOST2020

Both: PUB-2020-017

- W/Z-taggers are optimized simultaneously for three-variables: mass, $D_2^{(\beta=1)}$, and N_{trk}
  - N_{trk} can provide a large gain, at the cost of some modelling → fix with scale factors
- Top taggers use 13 jet-level variables in a fully-connected Deep Neural Network (DNN)
  - Significant benefits when using a DNN as compared to a cut-based tagger
  - Note: "inclusive" top also shown, only requires truth top quark match (more in backup)

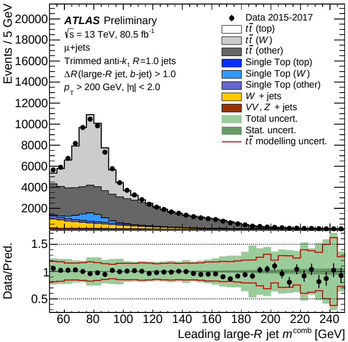


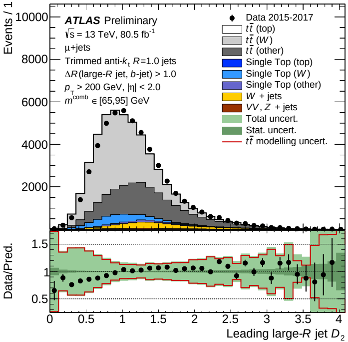
W tagger selections in data

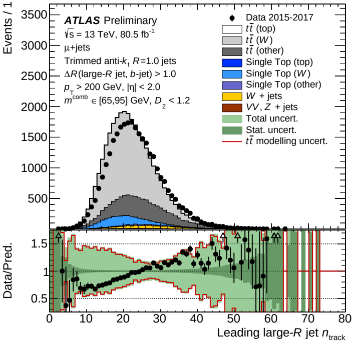
New @ BOOST2020

All: PUB-2020-017

- Studied data/MC agreement for W bosons in $t\bar{t}$ events, at each stage of the tagger
  - Cut values evolve with p_T, shown here for cuts corresponding to the lowest p_T range
- Some significant data/MC differences, but all within the modelling uncertainties
  - N_trk actually very well modelled inclusively, disagreement arises after cuts [backup]
- Account for these data/MC differences using tagging efficiency scale factors





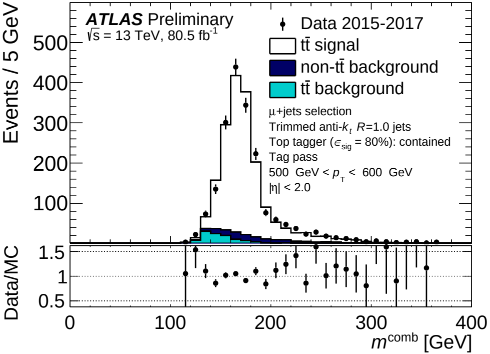


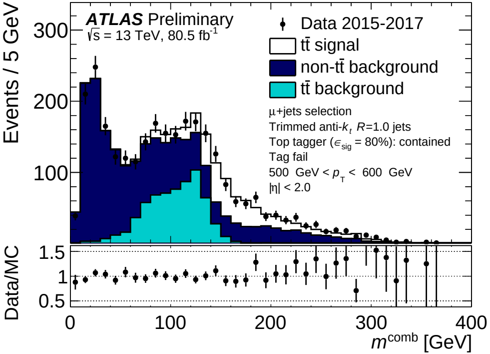
Extracting signal efficiency scale factors

New @ BOOST2020

Both: PUB-2020-017

- Define three templates from the shapes in MC, but allow their normalization to float
  - Fit normalizations to pass and fail distributions to extract number of signal events in data
- $\epsilon_{\text{data}}^{\text{signal}} = \frac{N_{\text{fit,signal}}^{\text{tagged}}}{N_{\text{fit,signal}}^{\text{tagged}} + N_{\text{fit,signal}}^{\text{not tagged}}}$, $\epsilon_{\text{MC}}^{\text{signal}} = \frac{N_{\text{signal}}^{\text{tagged}}}{N_{\text{signal}}^{\text{tagged}} + N_{\text{signal}}^{\text{not tagged}}}$, $\text{scale factor} = \epsilon_{\text{data}}^{\text{signal}} / \epsilon_{\text{MC}}^{\text{signal}}$



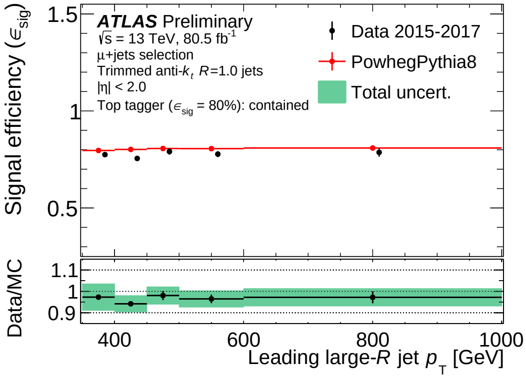


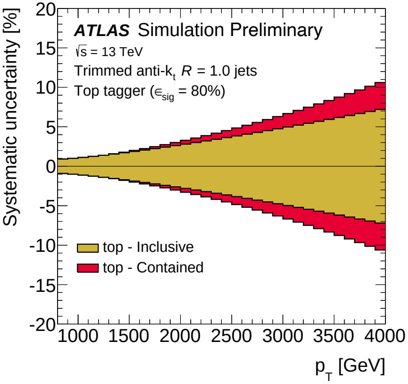
Signal efficiency scale factors and extrapolation

New @ BOOST2020

Both: PUB-2020-017

- Data and simulation generally agree within uncertainties, even for complex taggers
  - Contained DNN top tagger $\epsilon_{\text{sig}}^{\text{data}}/\epsilon_{\text{sig}}^{\text{MC}} \approx 1$; 3-variable W-tagger $\epsilon_{\text{sig}}^{\text{data}}/\epsilon_{\text{sig}}^{\text{MC}} \approx 0.8$ [backup]
- Then extrapolate the scale factors to higher p_{T} using MC variations
  - Variations: Geant4 physics lists, ATLAS detector geometry model, alternative generators



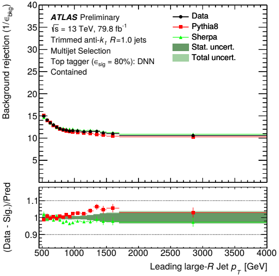


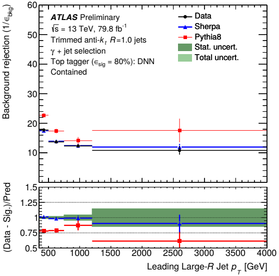
Background efficiency scale factors

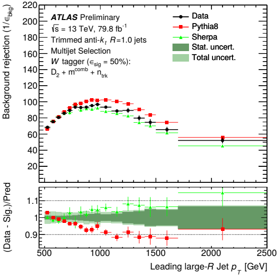
New @ BOOST2020

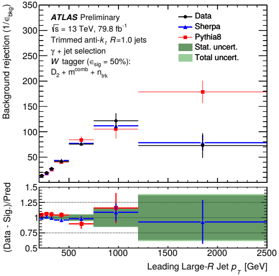
All: PUB-2020-017

- Much simpler than for the signal efficiency, as the sample is very pure in background
  - $\epsilon_{\text{data}}^{\text{background}} = \frac{N_{\text{background}}^{\text{tagged}}}{N_{\text{background}}^{\text{total}}}$, $\epsilon_{\text{MC}}^{\text{background}} = \frac{N_{\text{background}}^{\text{tagged}}}{N_{\text{background}}^{\text{total}}}$, scale factor $= \epsilon_{\text{data}}^{\text{background}} / \epsilon_{\text{MC}}^{\text{background}}$
- Study performance of taggers in background samples: QCD multijets and γ+jet
  - Study modelling differences between gluon-enriched (QCD) and light-quark-enriched (γ+jet)
- Left two plots are for a top-tagger, right two plots are for a W-tagger





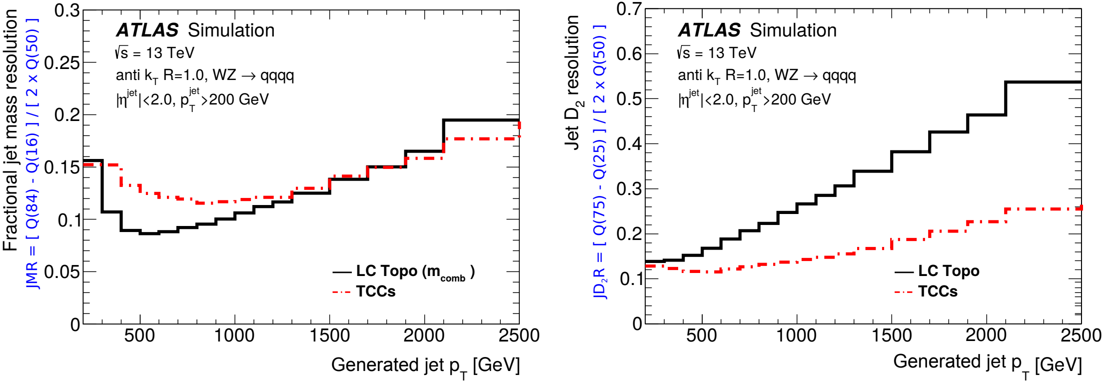




TrackCaloCluster jets

Both: HDBS-2018-31

- Track-CaloClusters (TCCs) split and assign cluster energy to matching track(s)
  - Degrades low p_T mass performance, but significantly improves $D_2^{(\beta=1)}$
- TCC jets support powerful taggers at high p_T; topo-clusters are better at low p_T
  - Some high-p_T analyses have thus used TCC jet taggers to increase sensitivity

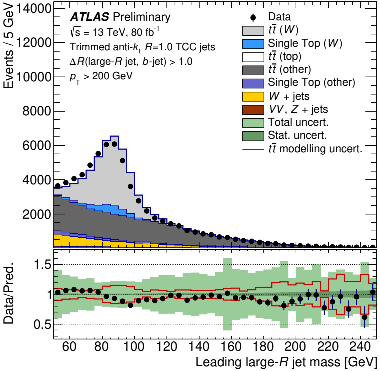


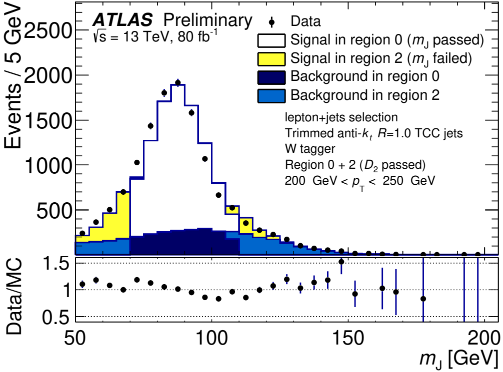
Scale factors for TCC jets

Recent result (May)

Both: PUB-2020-008

- No *in situ* JES calibration for TCC jets $\Longrightarrow$ mass peak offset between data and MC
  - Differences are corrected for with scale factors and are within uncertainties
  - Simultaneously extract multi-region scale factors (4×): [pass, fail] × [mass, $D_2^{(\beta=1)}$]



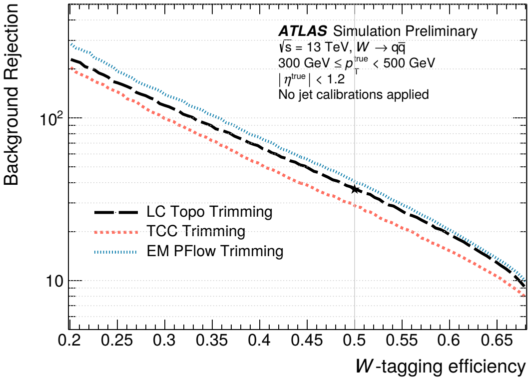


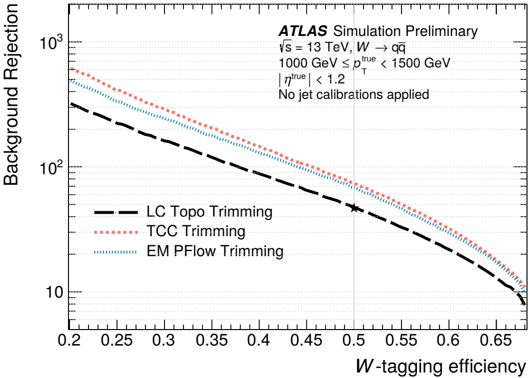
The importance of jet inputs

New @ BOOST2020

Both: CONF-2020-021

- As mentioned, $R = 0.4$ jets are increasingly using Particle Flow (PFlow)
  - These work very well for the lower p_T regime, also improves low-p_T tagging of large-R jets
- However, TCC jets dominate at high p_T (relevance continues to grow at higher p_T)
  - Two different algorithms for two different domains: let's combine them!

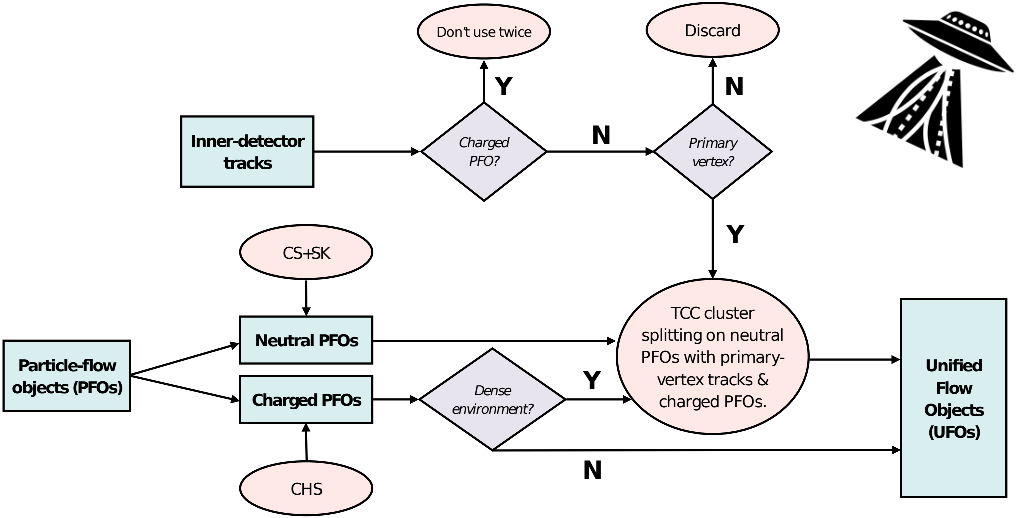




Unified Flow Objects (UFOs) ≈ PFOs + TCCs

New @ BOOST2020

Plot: CONF-2020-021



Inner-detector tracks

Charged PFO?

Y

Don't use twice

N

Primary vertex?

N

Discard

Y

CS+SK

Neutral PFOs

Particle-flow objects (PFOs)

Charged PFOs

CHS

Dense environment?

Y

N

TCC cluster splitting on neutral PFOs with primary-vertex tracks & charged PFOs.

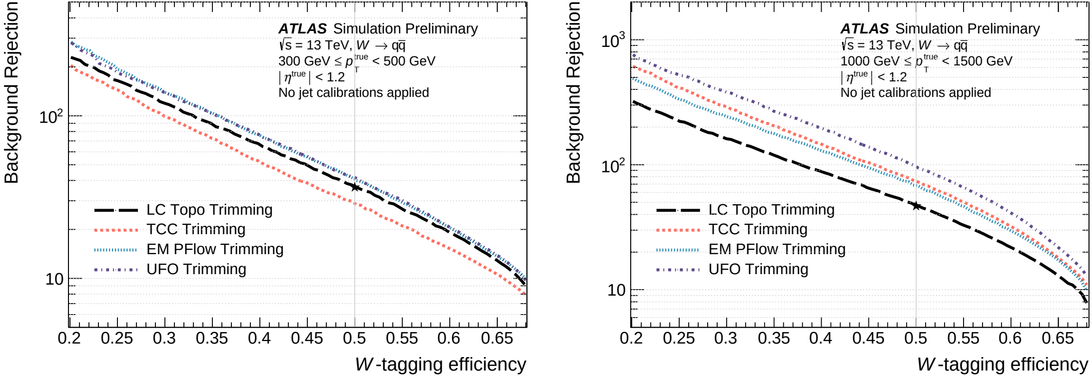
Unified Flow Objects (UFOs)

Revisiting the W-tagging performance

New @ BOOST2020

Both: CONF-2020-021

- This is the same plot as the previous slide, but with the UFO lines added
  - Equivalent to PFlow at low p_T, superior to TCC at high p_T (more TCC-like at very high p_T)
- Significant improvements from using UFOs, without even re-optimizing the jet definition!
- This and next slides are simple taggers: 68% mass window + $[D_2^{(\beta=1)}$ or $\tau_{32}^{\text{wta}}]$ cut

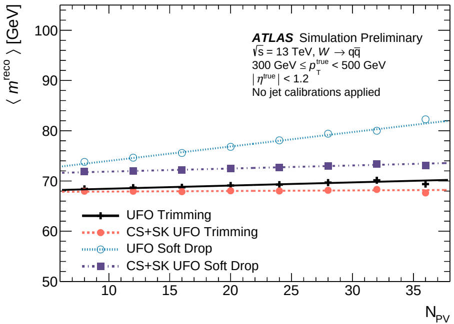


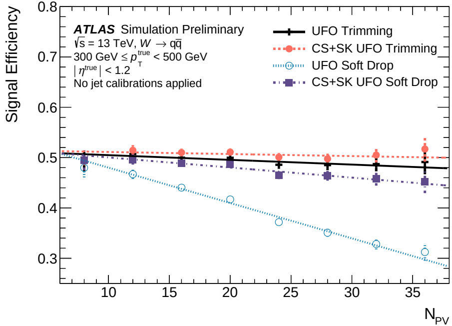
Surveying jet definitions and pileup stability

New @ BOOST2020

Both: CONF-2020-021

- UFOs already work very well, but we can still improve by surveying jet definitions
- Jet definitions have come a long way since ATLAS settled on the trimming algorithm
  - New groomers such as Soft Drop, Recursive Soft Drop, and Bottom-Up Soft Drop
  - New pileup mitigation techniques such as Constituent Subtraction and SoftKiller
- Consider pileup stability: clear balance between groomers and pileup mitigation techniques
  - Left: stability of the W mass peak; right: stability of the W-tagging efficiency (50% tagger)



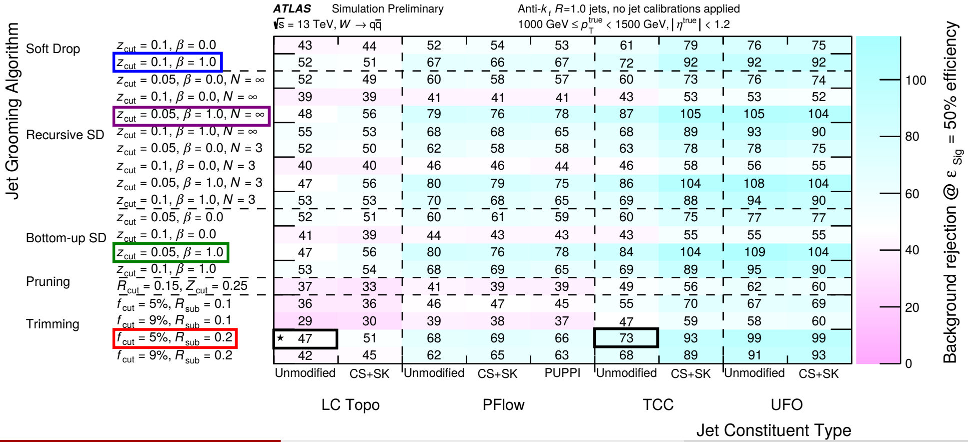


W-tagging performance, $1 < p_T/\text{TeV} < 1.5$

New @ BOOST2020

Plot: CONF-2020-021

- Consider everything at once: jet input type, pileup mitigation, and grooming algorithm
  - The below plot shows the QCD multijet rejection for a 50% W-tagger
  - Several other mega-plots for different metrics can be found in the backup



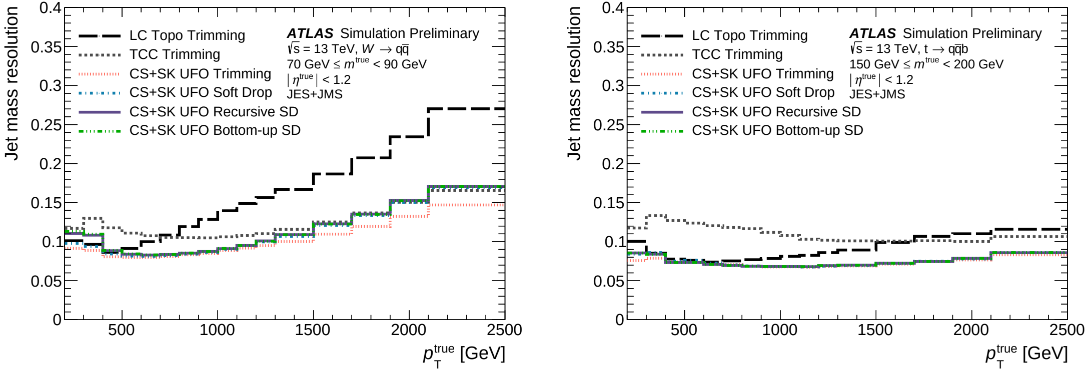
Jet Grooming Algorithm		LC Topo		PFlow			TCC		UFO	
		Unmodified	CS+SK	Unmodified	CS+SK	PUPPI	Unmodified	CS+SK	Unmodified	CS+SK
Soft Drop	z_{cut} = 0.1, β = 0.0	43	44	52	54	53	61	79	76	75
	z_{cut} = 0.1, β = 1.0	52	51	67	66	67	72	92	92	92
Recursive SD	z_{cut} = 0.05, β = 0.0, $N = \infty$	52	49	60	58	57	60	73	76	74
	z_{cut} = 0.1, β = 0.0, $N = \infty$	39	39	41	41	41	43	53	53	52
	z_{cut} = 0.05, β = 1.0, $N = \infty$	48	56	79	76	78	87	105	105	104
	z_{cut} = 0.1, β = 1.0, $N = \infty$	55	53	68	68	65	68	89	93	90
	z_{cut} = 0.05, β = 0.0, N = 3	52	50	62	58	58	63	78	78	75
	z_{cut} = 0.1, β = 0.0, N = 3	40	40	46	46	44	46	58	56	55
	z_{cut} = 0.05, β = 1.0, N = 3	47	56	80	79	75	86	104	108	104
	z_{cut} = 0.1, β = 1.0, N = 3	53	53	70	68	65	69	88	94	90
Bottom-up SD	z_{cut} = 0.05, β = 0.0	52	51	60	61	59	60	75	77	77
	z_{cut} = 0.1, β = 0.0	41	39	44	43	43	43	55	55	55
	z_{cut} = 0.05, β = 1.0	47	56	80	76	78	84	104	109	104
	z_{cut} = 0.1, β = 1.0	53	54	68	69	65	69	89	95	90
Pruning	R_{cut} = 0.15, Z_{cut} = 0.25	37	33	41	39	39	49	56	62	60
Trimming	f_{cut} = 5%, R_{sub} = 0.1	36	36	46	47	45	55	70	67	69
	f_{cut} = 9%, R_{sub} = 0.1	29	30	39	38	37	47	59	58	60
	f_{cut} = 5%, R_{sub} = 0.2	★ 47	51	68	69	66	73	93	99	99
	f_{cut} = 9%, R_{sub} = 0.2	42	45	62	65	63	68	89	91	93

Finalists: jet mass resolution

New @ BOOST2020

Both: CONF-2020-021

- Combining information from the mega-plots like the last slide, we can identify "finalists"
  - A subset of four new definitions that are candidates for an optimal new ATLAS jet definition
  - All use UFO inputs and CS+SK pileup mitigation, only the grooming algorithm varies
- Compare finalists to the currently used options: trimming with topo-clusters or TCCs
  - Able to match or improve the current jet mass resolution **everywhere**

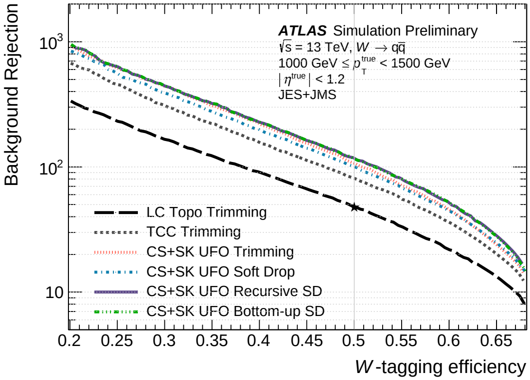


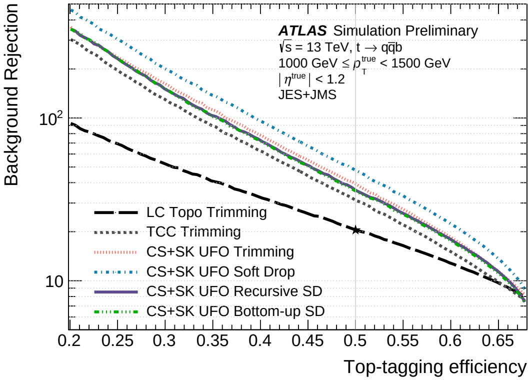
Finalists: tagging performance

New @ BOOST2020

Both: CONF-2020-021

- Next, consider the W- and top-tagging performance of the new jet definitions
  - Once again, large improvements are possible **everywhere**
  - Soft Drop seems to work very well for top, only small losses vs other options for W-tagging



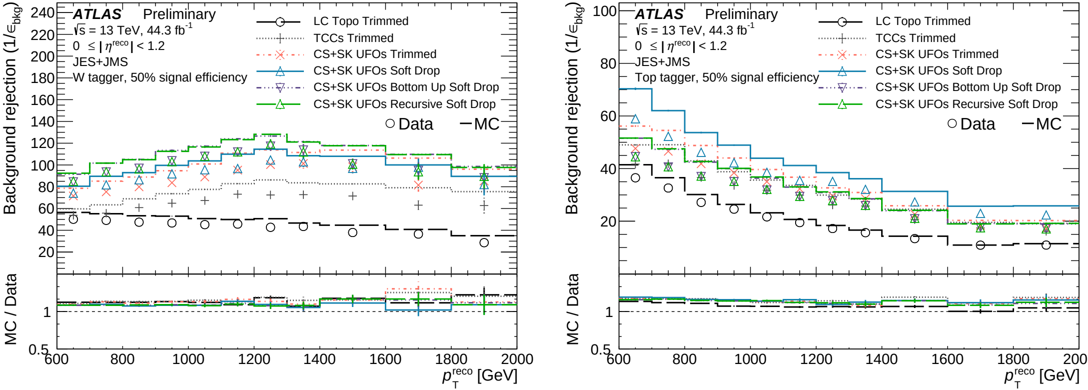


Finalists: tagging performance, in data

New @ BOOST2020

Both: CONF-2020-021

- It is important to understand if the tagging improvements are real
  - Applied to an inclusive jet data sample and compared to Pythia8 QCD dijet MC
- Results generally match the ordering of the MC tagging performance studies
  - Additionally, reasonable data/MC agreement ∼everywhere (only statistical uncertainties)

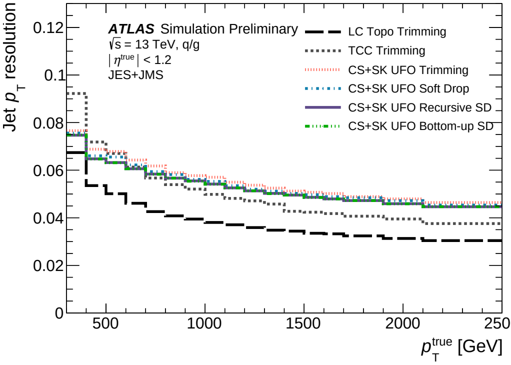


Finalists: jet energy resolution

New @ BOOST2020

Plot: CONF-2020-021

- UFO, CS+SK, Soft Drop seems to be working excellent in all of the metrics
  - However, there is one downside to UFO
- UFO uses “EM-scale” topo-clusters
  - Degraded JER compared to LC clusters
- All of the improvements make the new definition very useful for many analyses
  - However, still can be improved...
- An LC UFO variant could improve JER
  - Could also be mitigated via dedicated calibration (similar to $R = 0.4$ GSC)

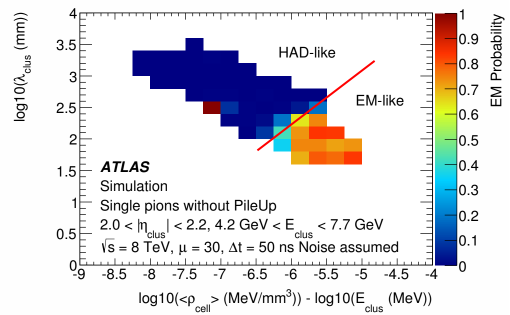


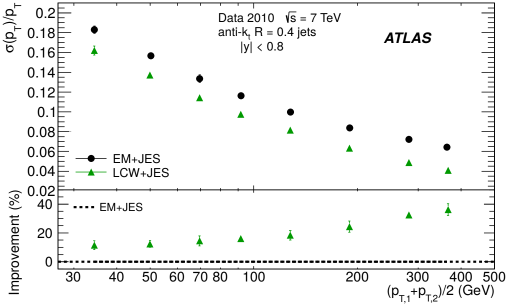
Topo-cluster calibrations

Left: PERF-2014-07

Right: PERF-2011-04

- ATLAS $R = 1.0$ jets have used topo-clusters at the Local Cell Weighting (LCW) scale
  1. Derive a probability of being a cluster from an electromagnetic or hadronic shower
  2. Apply a calibration weighted by that probability: $\text{LCW} \approx \mathcal{C}_{\text{EM}} \cdot \mathcal{P}_{\text{EM}} + \mathcal{C}_{\text{had}} \cdot (1 - \mathcal{P}_{\text{EM}})$
- This historically improves the JER by better representing showers within the jet
  - Let's try improving this correction using machine learning
  - See also Dewen's poster/talk later today for more details



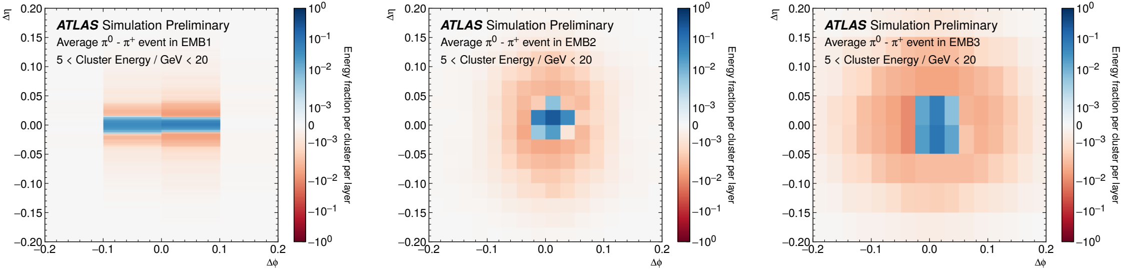


Using machine learning for cluster calibration

New @ BOOST2020

Both: PUB-2020-018

- LCW performs two tasks: classification (EM or hadronic) and calibration
- Do the same with machine learning and the cluster energy in each layer of the calorimeter
  - Tried DNNs, CNNs, and DenseNets for both tasks [network architectures in backup]
- Below is single-pion MC of the ATLAS EM barrel calorimeter layers
  - Shows the difference between the average shower energy distributions: π^0 - π^+
  - Narrow EM showers of $\pi^0 \to \gamma\gamma$ are visible compared to the wider showers of hadronic π^+

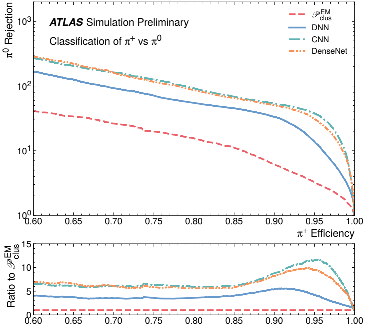


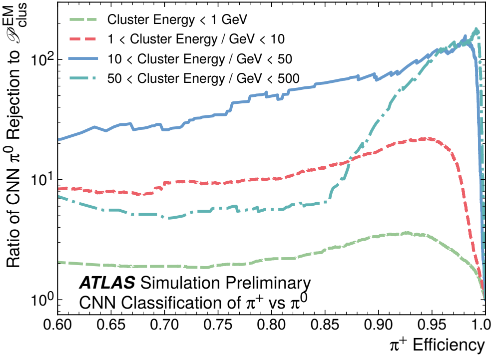
Cluster identification performance

New @ BOOST2020

Both: PUB-2020-018

- The ML techniques all do an excellent job of distinguishing π^0 from π^+ showers
  - Dramatic improvements are possible compared to the LCW classification performance
  - Overall, CNN seems to give the best performance for classifying EM vs hadronic showers
- Does well across p_T even though the networks were never actually given the cluster energy



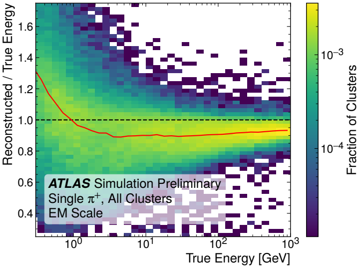


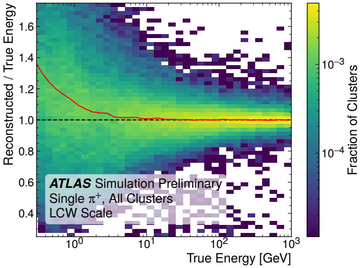
Regression performance, charged pions

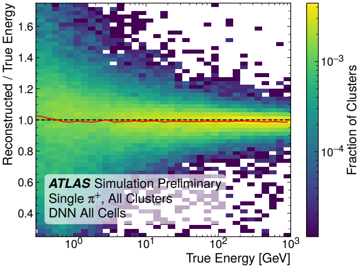
New @ BOOST2020

Both: PUB-2020-018

- Below is the cluster energy response $\mathcal{R} = E^{\text{reco}}_{\text{cluster}} / E^{\text{truth}}_{\text{cluster}}$ for charged pions
  - See the backup for plots showing performance for π^0 and a pion mixture $(\frac{1}{3}\pi^+, \frac{1}{3}\pi^-, \frac{1}{3}\pi^0)$
- Raw “EM” scale under-estimates $\mathcal{R}$; LCW over-estimates $\mathcal{R}$ at low-energy
  - In contrast, a DNN regression does an excellent job nearly everywhere





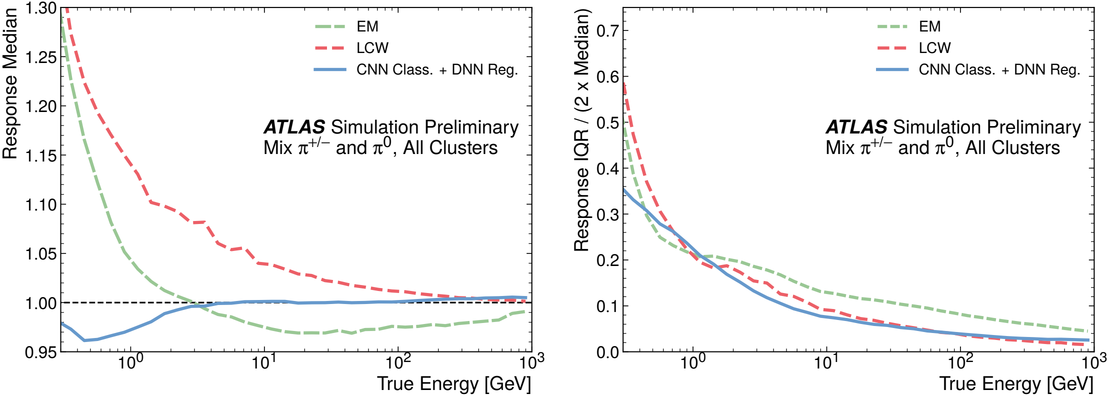


Cluster response and resolution, pion mixture

New @ BOOST2020

Both: PUB-2020-018

- Compare the EM, LCW, and CNN(classifier)+DNN(regression) performance
  - The CNN+DNN combination has the best scale and similar/better resolution ∼everywhere
- Very promising method to improve the topo-cluster calibration in the future



Summary

- There are lots of interesting new jet reconstruction and tagging developments in ATLAS!
  - Improved JES and JER determination for $R = 0.4$ jets, with a full PFlow calibration [link]
  - Newly optimized b-taggers for PFlow jets and with MC-to-MC scale factors [1, 2]
  - Improved b-tagging identification using deep impact parameter sets [link]
  - Dedicated $X \rightarrow bb$ taggers with enhanced Higgs-jet tagging capabilities [link]
  - $R = 1.0$ JMS and JMR measurements, including $V+$jets events and reclustered jets [link]
  - Improved W/Z/top taggers and scale factors, for both topo-cluster and TCC jets [1, 2]
  - A survey of $R = 1.0$ jet definitions, including a new UFO input type [link]
  - Improved topo-cluster calibrations using machine learning techniques [link]

- This is only a brief look through all of these new results
  - Please look at all of the associated notes/papers using the above links!
  - A compiled list of new ATLAS results for the BOOST audience can be found **here**

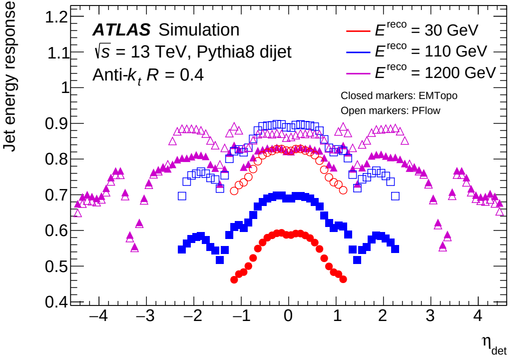
Backup Material

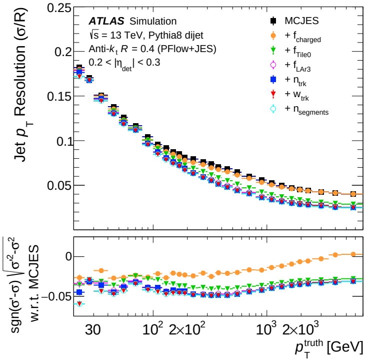
Simulation-based jet calibration

New @ BOOST2020

Both: JETM-2018-05

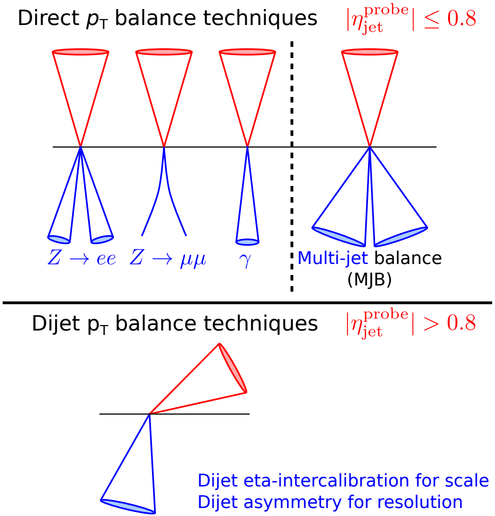
- The jet energy is only partially observed in the detector
  - Neutral particles: only measured by calorimeter, and only the energy in active layers
  - Charged particles: measured by calorimeter (in active layers) and tracker (full energy)
- Use the MC response $= \langle x_{\text{jet}}^{\text{reco}} / x_{\text{jet}}^{\text{truth}} \rangle$ of key variables to correct for unmeasured energy
  - PFlow measures the full scale of charged particles $\implies$ reduces required calibration factors





Accounting for differences between data and MC

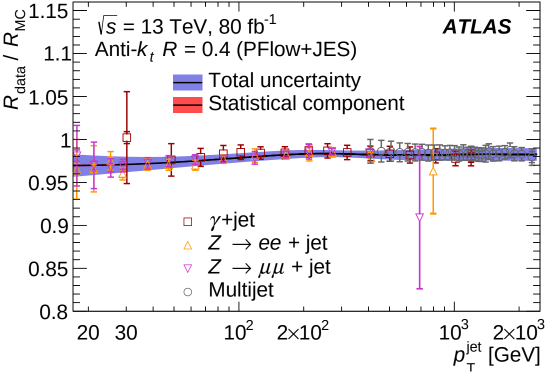
- The calibration was derived using MC
  - However, we apply it also to data
  - We need to account for potential data/MC differences after calibration
- Study data/MC agreement *in situ*
  - Direct p_T balance: probe jet balancing a well-understood reference object
  - Dijet p_T balance: forward probe jet balancing a different η reference jet
  - Dijet p_T asymmetry: unbalanced probe and reference jets quantify resolution
- Use to evaluate data/MC agreement
  - Correct the **data** to match the **MC**

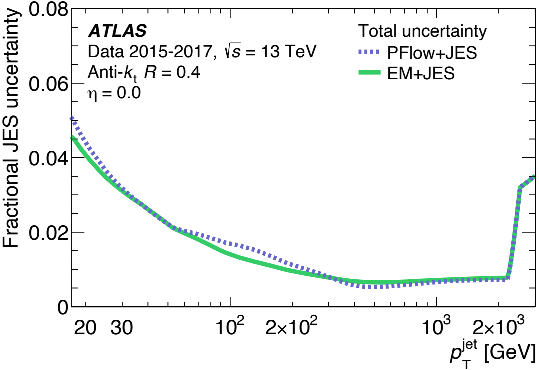


JES and its uncertainties

New @ BOOST2020
Both: JETM-2018-05

- *In situ* JES combination is reaching the level of 1% precision over large range of p_T
  - Additional uncertainties (flavour, etc) increase this as shown in right plot
  - PFlow and topo-cluster jets have very comparable JES uncertainties
- Uncertainties stable at low p_T compared to 8 TeV, despite much higher pileup [backup]

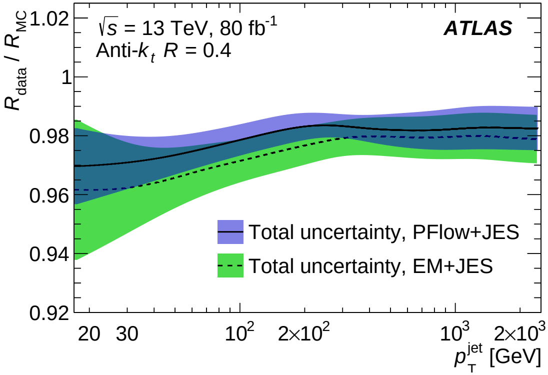


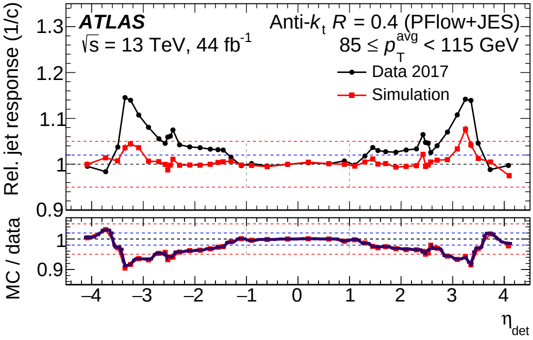


JES combination results

New @ BOOST2020

Both: JETM-2018-05

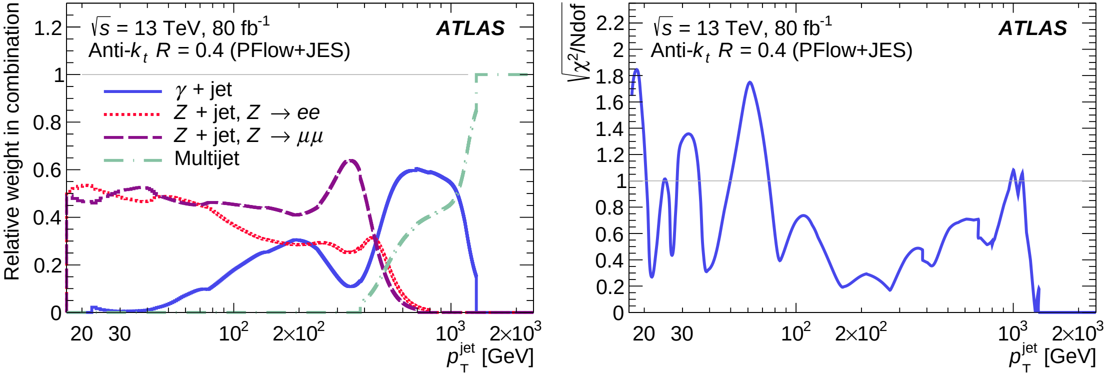




JES combination weights and χ^2

New @ BOOST2020

Both: JETM-2018-05



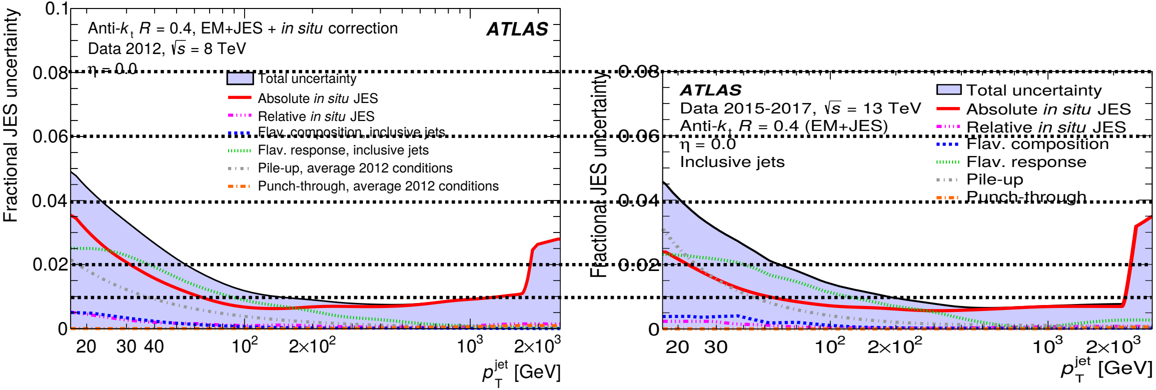
Comparing JES uncertainties at 8 and 13 TeV

Left: PERF-2014-02

New @ BOOST2020

Right: JETM-2018-05

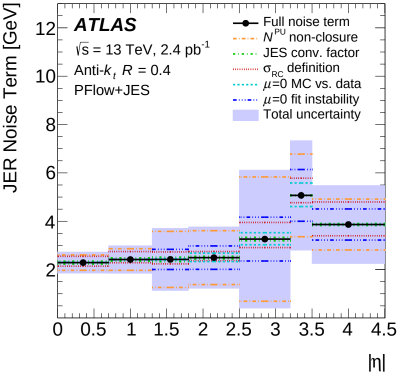
- Note different y-axis scale: uncertainties are very similar despite challenging conditions!

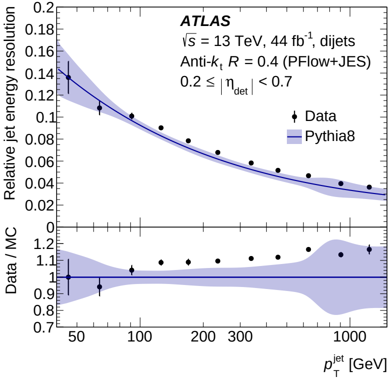


JER extraction

New @ BOOST2020

Both: JETM-2018-05



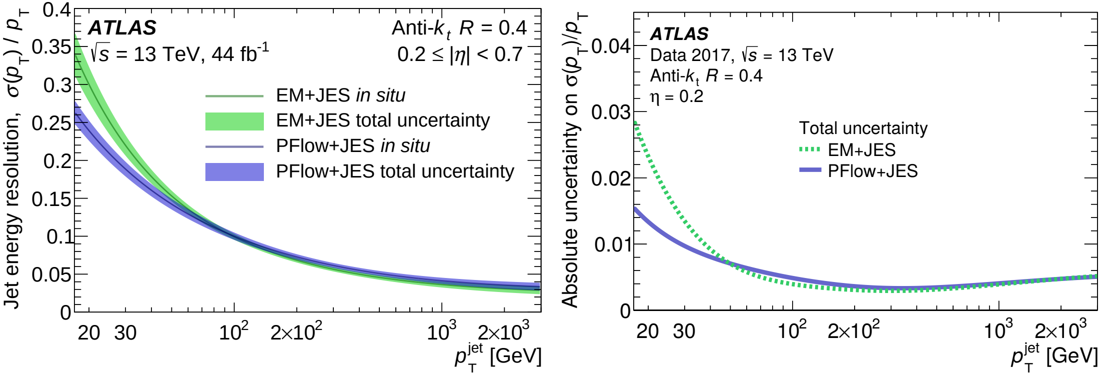


JER and its uncertainties

New @ BOOST2020

Both: JETM-2018-05

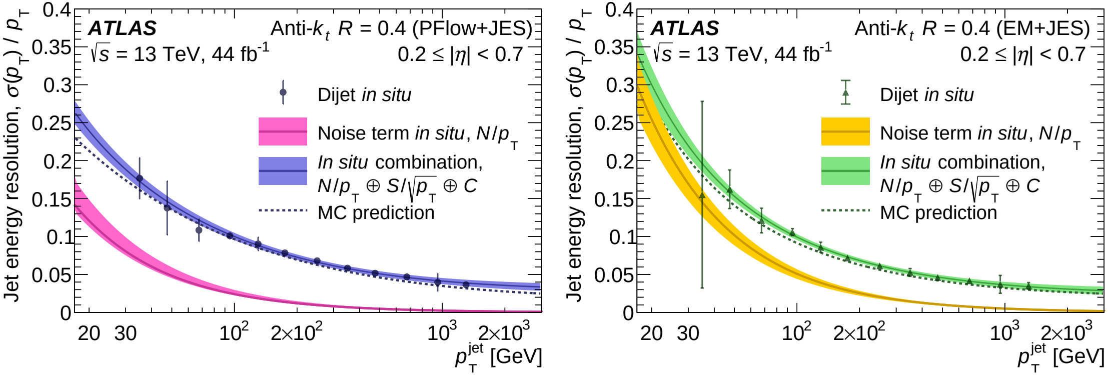
- JER is evaluated combining dijet asymmetry measurements with a noise term constraint
  - Constraint comes from random cones in zero-bias data to get ambient noise in a given area
- PFlow significantly improves the low p_T JER and its uncertainties vs topo-clusters
  - Resolution uncertainties are now typically below 1% absolute (5–15% relative - see backup)



JER combination

New @ BOOST2020

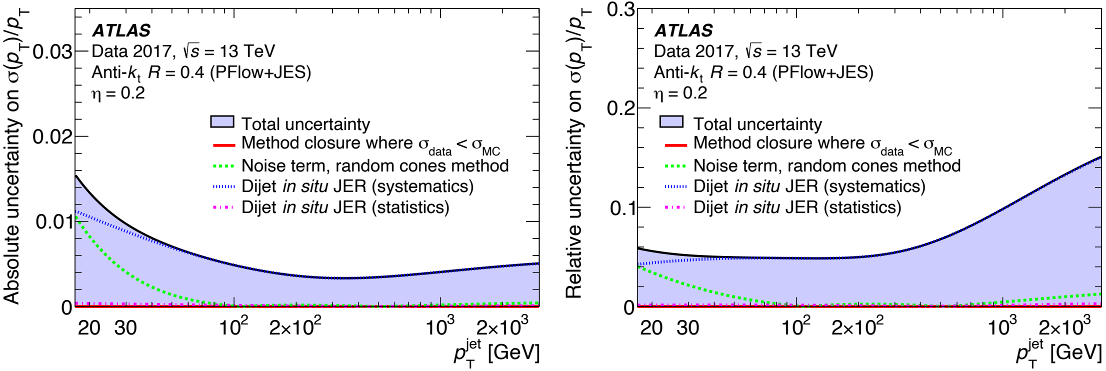
Both: JETM-2018-05



JER uncertainties, absolute and relative

New @ BOOST2020

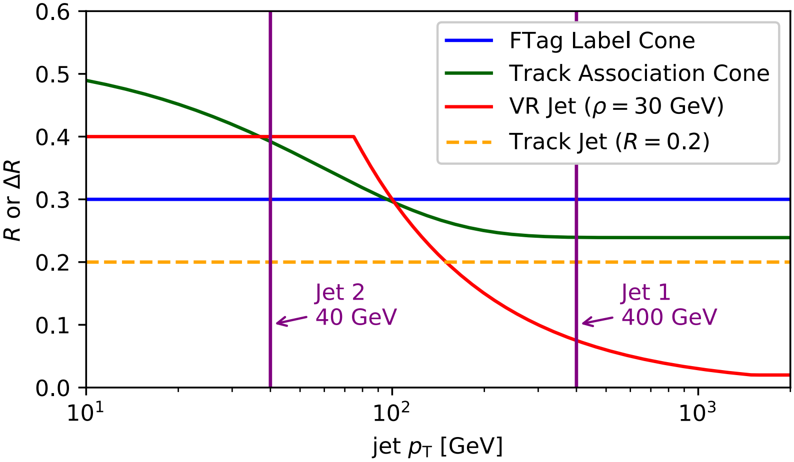
Both: JETM-2018-05



Labels, associations, and sizes for *b*-tagging

Plot: D. Guest

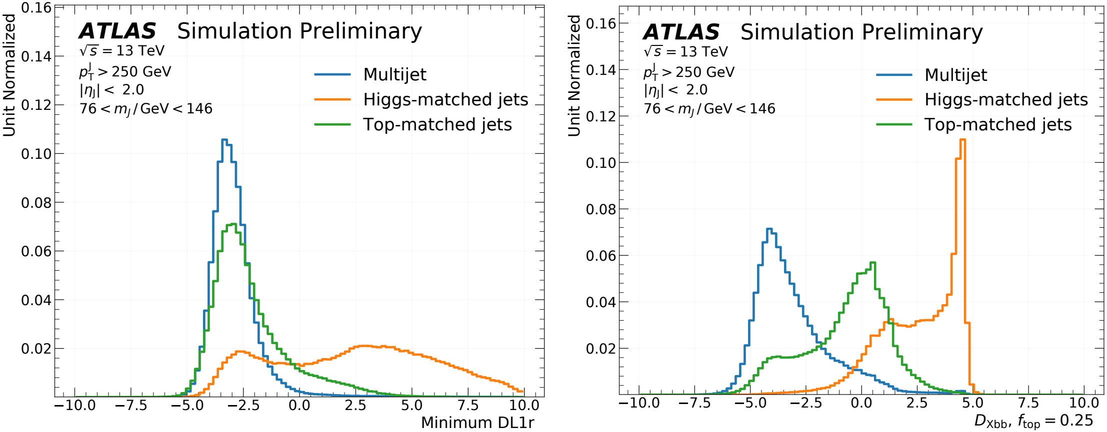
- Track association is exclusive: matched to closest jet in ΔR if multiple candidates



Hbb tagging discriminants

New @ BOOST2020

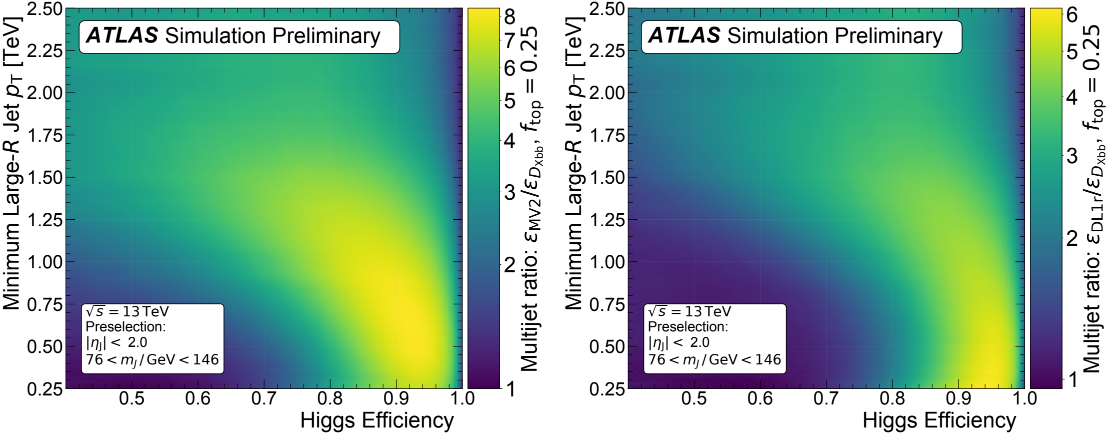
Both: PUB-2020-019



Hbb tagging vs 2x b-tag, light jets

New @ BOOST2020

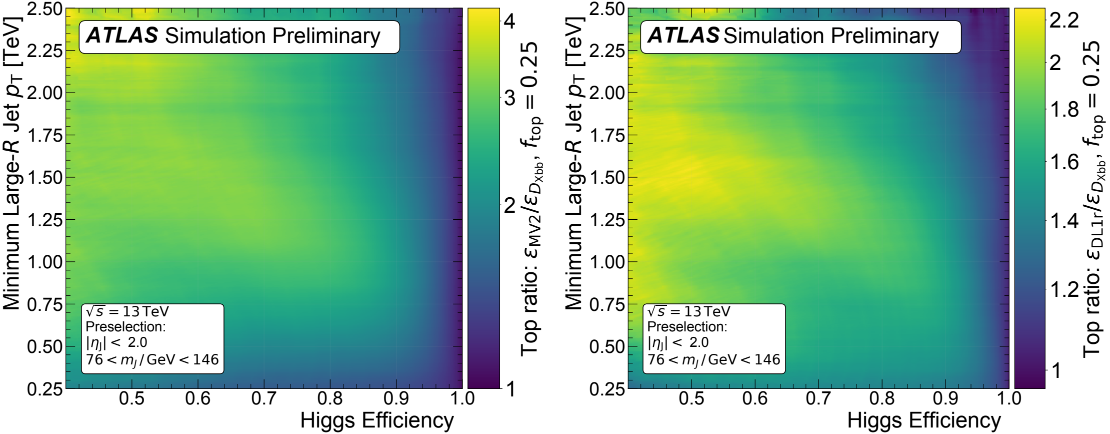
Both: PUB-2020-019



Hbb tagging vs 2x b-tag, top jets

New @ BOOST2020

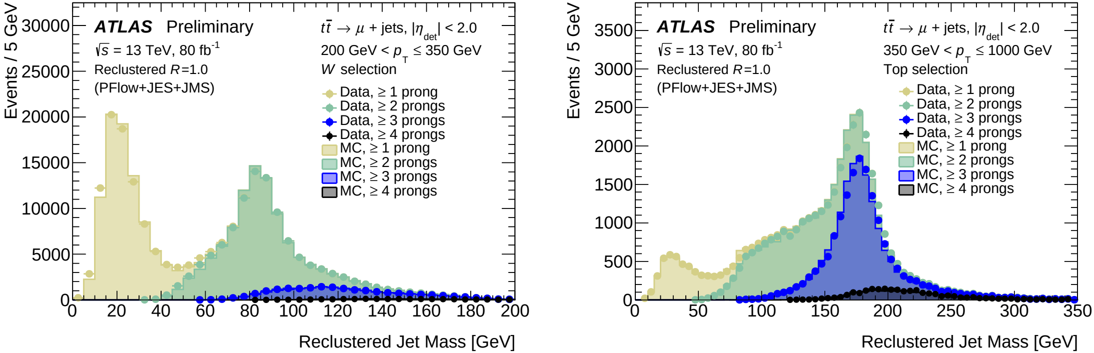
Both: PUB-2020-019



JMS and JMR dependence on jet structure

New @ BOOST2020

Plot: CONF-2020-022

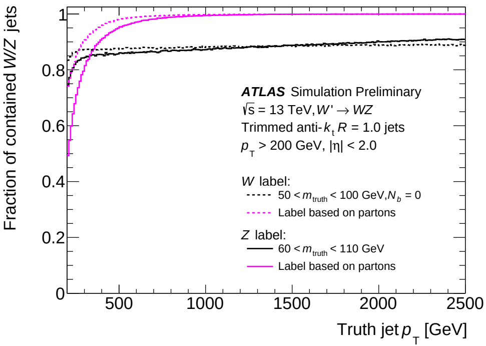


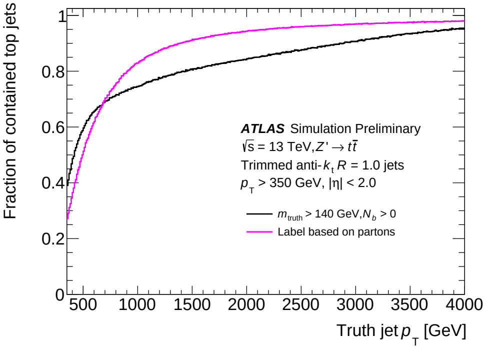
Signal jet definitions: W, Z, and top

New @ BOOST2020

Both: PUB-2020-017

- We need to define our target when designing a tagger: the objective we want to identify
  - W/Z: jet is matched to a truth W/Z boson, $m_{\text{truth}}^{\text{trimmed}}$ window, no b-hadrons (for W)
  - Top: jet is matched to a truth top quark, $m_{\text{truth}}^{\text{trimmed}} > 140\,\text{GeV}$, at least one b-hadron
- Referred to as "containment": the truth jet includes most of the truth particles
  - Changed from previous parton-based definition to reduce modelling dependence

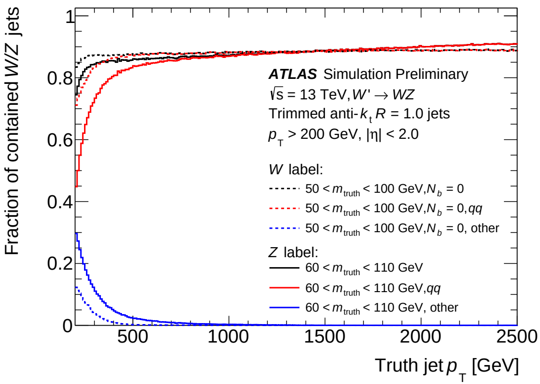


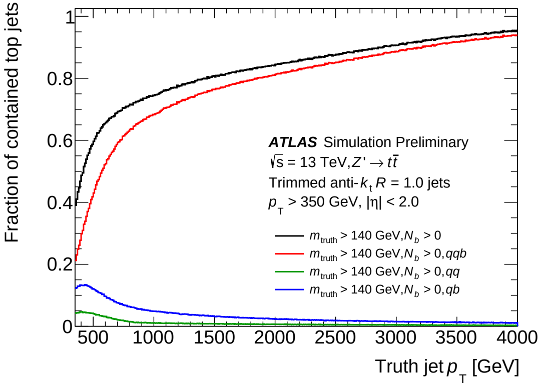


Signal jet definitions “purity”

New @ BOOST2020

Both: PUB-2020-017



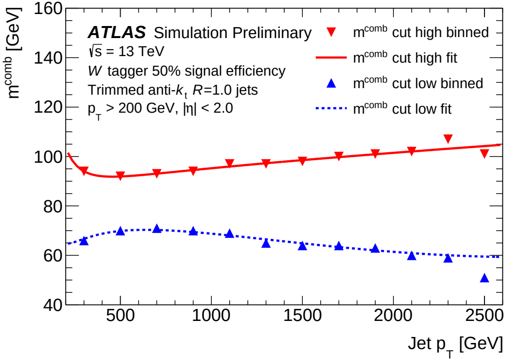


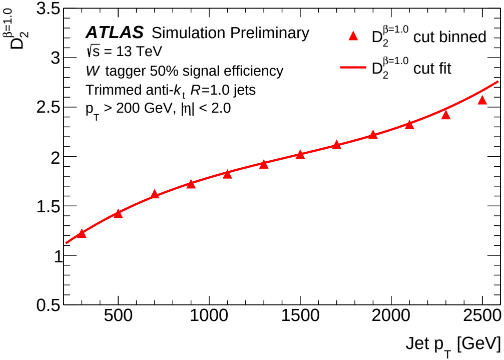
W-tagger optimization

New @ BOOST2020

Both: PUB-2020-017

- N_{trk} cut is fixed to 26 (does not evolve with p_{T}, found to be optimal everywhere)

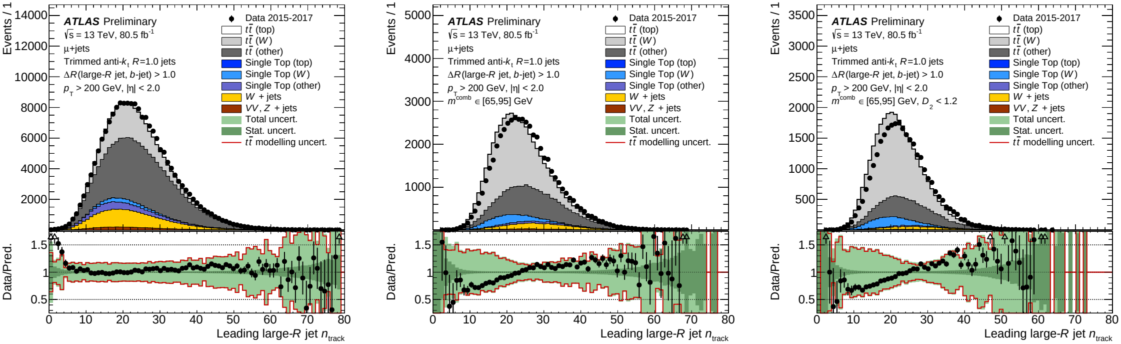




W tagger selections and N_{trk}

New @ BOOST2020

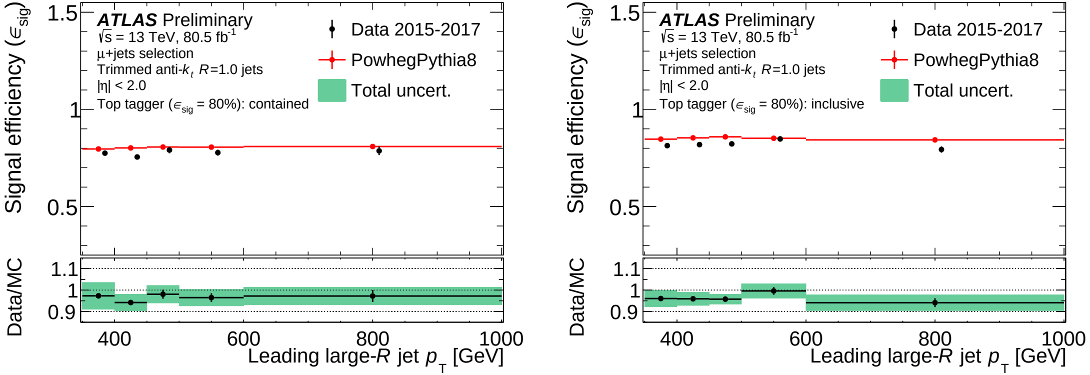
All: PUB-2020-017



Top signal efficiency scale factors

New @ BOOST2020

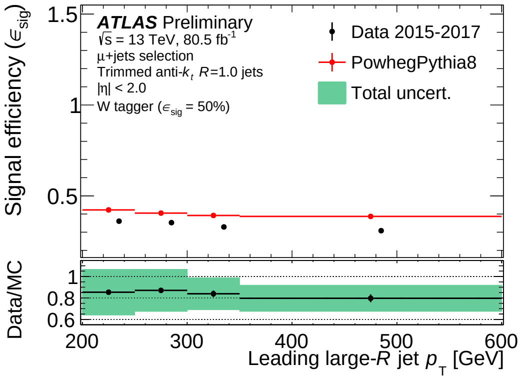
Both: PUB-2020-017

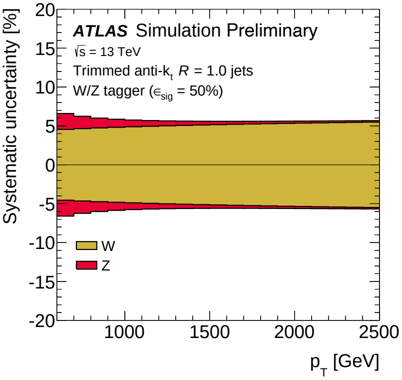


W signal efficiency scale factors

New @ BOOST2020

Both: PUB-2020-017

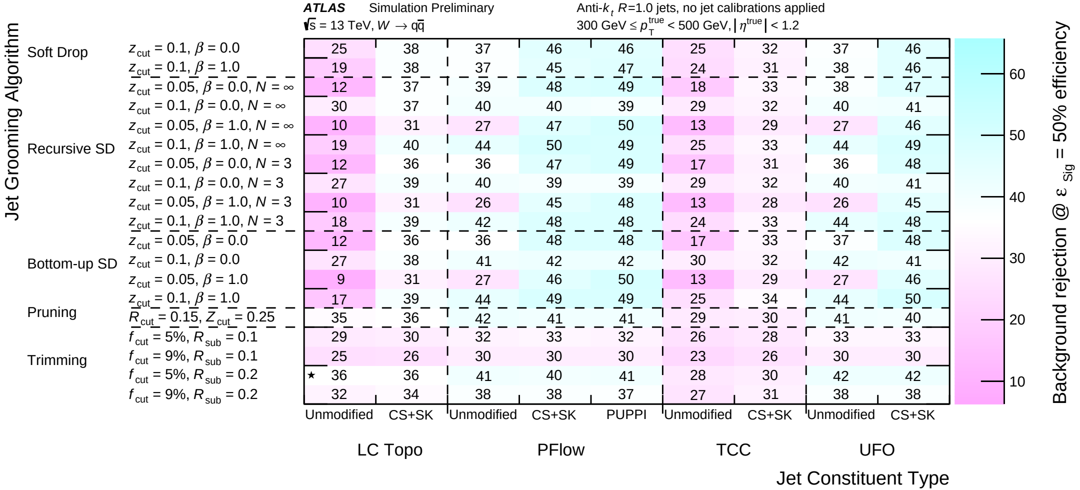




W-tagging performance, $300 < p_T/\text{GeV} < 500$

New @ BOOST2020

Plot: CONF-2020-021

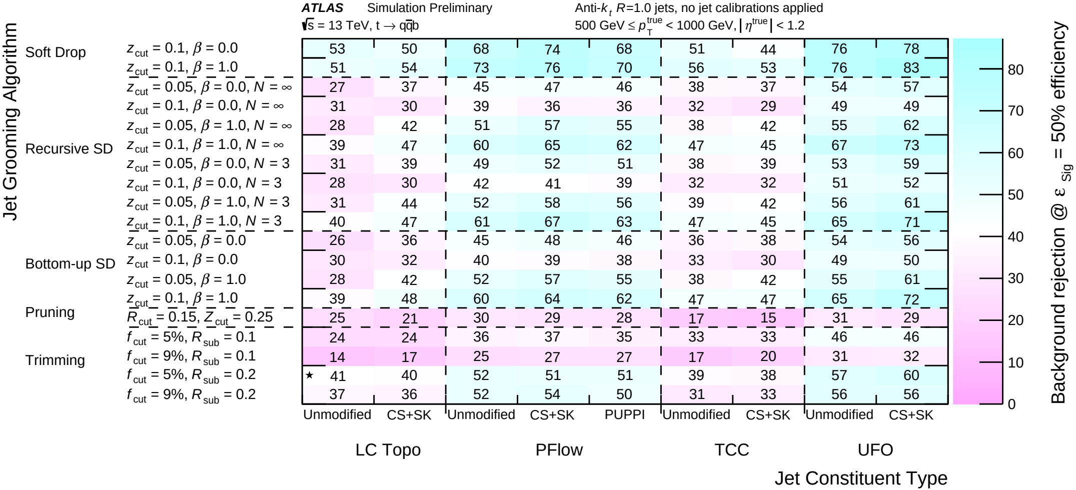


Jet Grooming Algorithm		LC Topo Unmodified	LC Topo CS+SK	PFlow Unmodified	PFlow CS+SK	PFlow PUPPI	TCC Unmodified	TCC CS+SK	UFO Unmodified	UFO CS+SK
Soft Drop	z_{cut} = 0.1, β = 0.0	25	38	37	46	46	25	32	37	46
	z_{cut} = 0.1, β = 1.0	19	38	37	45	47	24	31	38	46
Recursive SD	z_{cut} = 0.05, β = 0.0, $N = \infty$	12	37	39	48	49	18	33	38	47
	z_{cut} = 0.1, β = 0.0, $N = \infty$	30	37	40	40	39	29	32	40	41
	z_{cut} = 0.05, β = 1.0, $N = \infty$	10	31	27	47	50	13	29	27	46
	z_{cut} = 0.1, β = 1.0, $N = \infty$	19	40	44	50	49	25	33	44	49
	z_{cut} = 0.05, β = 0.0, N = 3	12	36	36	47	49	17	31	36	48
	z_{cut} = 0.1, β = 0.0, N = 3	27	39	40	39	39	29	32	40	41
	z_{cut} = 0.05, β = 1.0, N = 3	10	31	26	45	48	13	28	26	45
	z_{cut} = 0.1, β = 1.0, N = 3	18	39	42	48	48	24	33	44	48
Bottom-up SD	z_{cut} = 0.05, β = 0.0	12	36	36	48	48	17	33	37	48
	z_{cut} = 0.1, β = 0.0	27	38	41	42	42	30	32	42	41
	z_{cut} = 0.05, β = 1.0	9	31	27	46	50	13	29	27	46
	z_{cut} = 0.1, β = 1.0	17	39	44	49	49	25	34	44	50
Pruning	R_{cut} = 0.15, Z_{cut} = 0.25	35	36	42	41	41	29	30	41	40
Trimming	f_{cut} = 5%, R_{sub} = 0.1	29	30	32	33	32	26	28	33	33
	f_{cut} = 9%, R_{sub} = 0.1	25	26	30	30	30	23	26	30	30
	f_{cut} = 5%, R_{sub} = 0.2	★ 36	36	41	40	41	28	30	42	42
	f_{cut} = 9%, R_{sub} = 0.2	32	34	38	38	37	27	31	38	38

Top-tagging performance, $0.5 < p_T/\mathrm{TeV} < 1$

New @ BOOST2020

Plot: CONF-2020-021

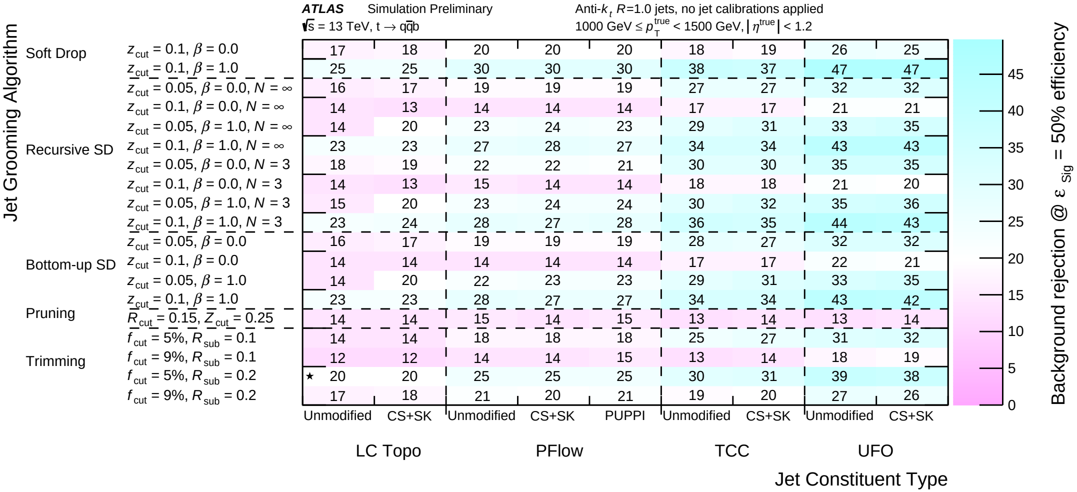


Jet Grooming Algorithm		LC Topo Unmodified	LC Topo CS+SK	PFlow Unmodified	PFlow CS+SK	PFlow PUPPI	TCC Unmodified	TCC CS+SK	UFO Unmodified	UFO CS+SK
Soft Drop	z_{cut} = 0.1, β = 0.0	53	50	68	74	68	51	44	76	78
	z_{cut} = 0.1, β = 1.0	51	54	73	76	70	56	53	76	83
Recursive SD	z_{cut} = 0.05, β = 0.0, $N = \infty$	27	37	45	47	46	38	37	54	57
	z_{cut} = 0.1, β = 0.0, $N = \infty$	31	30	39	36	36	32	29	49	49
	z_{cut} = 0.05, β = 1.0, $N = \infty$	28	42	51	57	55	38	42	55	62
	z_{cut} = 0.1, β = 1.0, $N = \infty$	39	47	60	65	62	47	45	67	73
	z_{cut} = 0.05, β = 0.0, N = 3	31	39	49	52	51	38	39	53	59
	z_{cut} = 0.1, β = 0.0, N = 3	28	30	42	41	39	32	32	51	52
	z_{cut} = 0.05, β = 1.0, N = 3	31	44	52	58	56	39	42	56	61
	z_{cut} = 0.1, β = 1.0, N = 3	40	47	61	67	63	47	45	65	71
Bottom-up SD	z_{cut} = 0.05, β = 0.0	26	36	45	48	46	36	38	54	56
	z_{cut} = 0.1, β = 0.0	30	32	40	39	38	33	30	49	50
	z_{cut} = 0.05, β = 1.0	28	42	52	57	55	38	42	55	61
	z_{cut} = 0.1, β = 1.0	39	48	60	64	62	47	47	65	72
Pruning	R_{cut} = 0.15, Z_{cut} = 0.25	25	21	30	29	28	17	15	31	29
Trimming	f_{cut} = 5%, R_{sub} = 0.1	24	24	36	37	35	33	33	46	46
	f_{cut} = 9%, R_{sub} = 0.1	14	17	25	27	27	17	20	31	32
	f_{cut} = 5%, R_{sub} = 0.2	★ 41	40	52	51	51	39	38	57	60
	f_{cut} = 9%, R_{sub} = 0.2	37	36	52	54	50	31	33	56	56

Top-tagging performance, $1 < p_T/\text{TeV} < 1.5$

New @ BOOST2020

Plot: CONF-2020-021

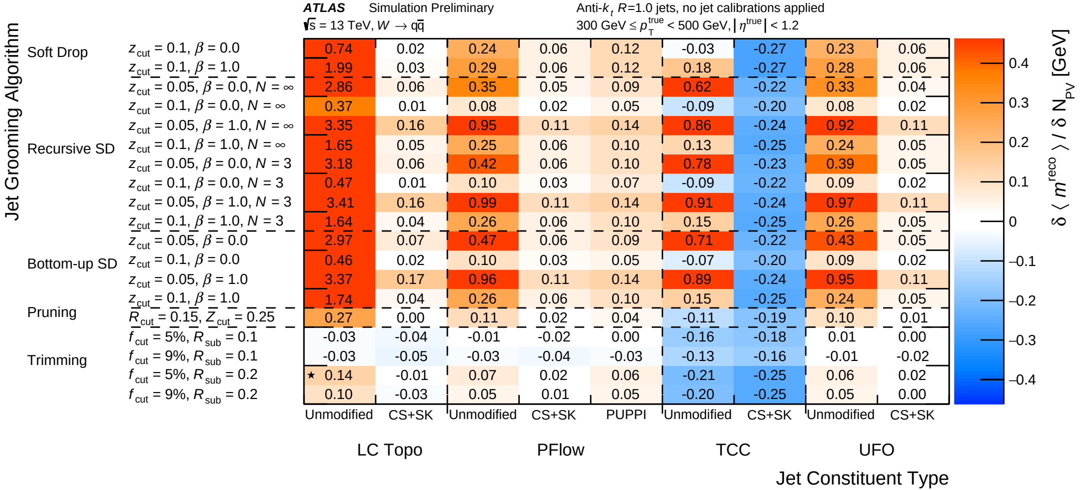


		LC Topo		PFlow			TCC		UFO	
		Unmodified	CS+SK	Unmodified	CS+SK	PUPPI	Unmodified	CS+SK	Unmodified	CS+SK
Soft Drop	$z_{cut} = 0.1, \beta = 0.0$	17	18	20	20	20	18	19	26	25
	$z_{cut} = 0.1, \beta = 1.0$	25	25	30	30	30	38	37	47	47
Recursive SD	$z_{cut} = 0.05, \beta = 0.0, N = \infty$	16	17	19	19	19	27	27	32	32
	$z_{cut} = 0.1, \beta = 0.0, N = \infty$	14	13	14	14	14	17	17	21	21
	$z_{cut} = 0.05, \beta = 1.0, N = \infty$	14	20	23	24	23	29	31	33	35
	$z_{cut} = 0.1, \beta = 1.0, N = \infty$	23	23	27	28	27	34	34	43	43
	$z_{cut} = 0.05, \beta = 0.0, N = 3$	18	19	22	22	21	30	30	35	35
	$z_{cut} = 0.1, \beta = 0.0, N = 3$	14	13	15	14	14	18	18	21	20
	$z_{cut} = 0.05, \beta = 1.0, N = 3$	15	20	23	24	24	30	32	35	36
	$z_{cut} = 0.1, \beta = 1.0, N = 3$	23	24	28	27	28	36	35	44	43
Bottom-up SD	$z_{cut} = 0.05, \beta = 0.0$	16	17	19	19	19	28	27	32	32
	$z_{cut} = 0.1, \beta = 0.0$	14	14	14	14	14	17	17	22	21
	$z_{cut} = 0.05, \beta = 1.0$	14	20	22	23	23	29	31	33	35
	$z_{cut} = 0.1, \beta = 1.0$	23	23	28	27	27	34	34	43	42
Pruning	$R_{cut} = 0.15, Z_{cut} = 0.25$	14	14	15	14	15	13	14	13	14
Trimming	$f_{cut} = 5\%, R_{sub} = 0.1$	14	14	18	18	18	25	27	31	32
	$f_{cut} = 9\%, R_{sub} = 0.1$	12	12	14	14	15	13	14	18	19
	$f_{cut} = 5\%, R_{sub} = 0.2$	★ 20	20	25	25	25	30	31	39	38
	$f_{cut} = 9\%, R_{sub} = 0.2$	17	18	21	20	21	19	20	27	26

W mass peak pileup dependence

New @ BOOST2020

Plot: CONF-2020-021

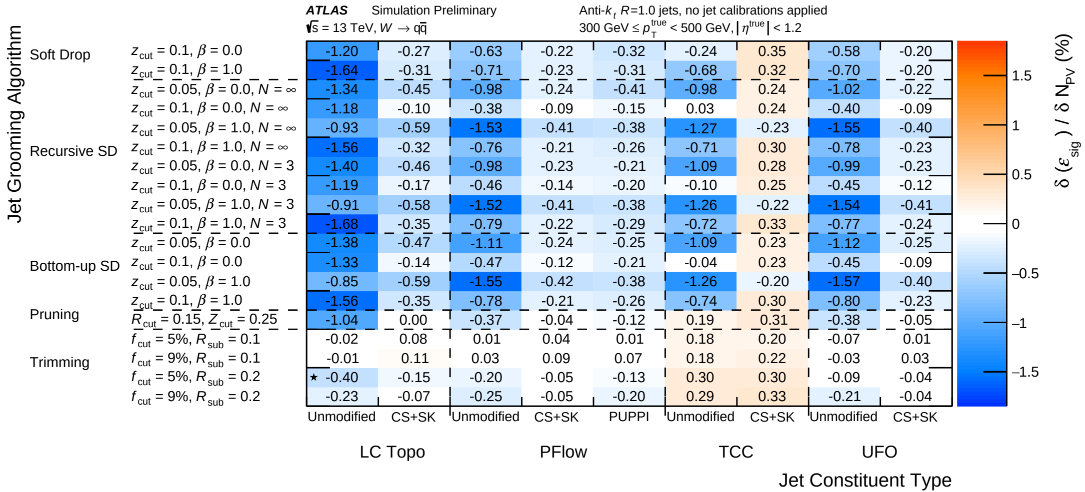


Jet Grooming Algorithm		LC Topo		PFlow			TCC		UFO	
		Unmodified	CS+SK	Unmodified	CS+SK	PUPPI	Unmodified	CS+SK	Unmodified	CS+SK
Soft Drop	$z_{cut} = 0.1, \beta = 0.0$	0.74	0.02	0.24	0.06	0.12	-0.03	-0.27	0.23	0.06
	$z_{cut} = 0.1, \beta = 1.0$	1.99	0.03	0.29	0.06	0.12	0.18	-0.27	0.28	0.06
Recursive SD	$z_{cut} = 0.05, \beta = 0.0, N = \infty$	2.86	0.06	0.35	0.05	0.09	0.62	-0.22	0.33	0.04
	$z_{cut} = 0.1, \beta = 0.0, N = \infty$	0.37	0.01	0.08	0.02	0.05	-0.09	-0.20	0.08	0.02
	$z_{cut} = 0.05, \beta = 1.0, N = \infty$	3.35	0.16	0.95	0.11	0.14	0.86	-0.24	0.92	0.11
	$z_{cut} = 0.1, \beta = 1.0, N = \infty$	1.65	0.05	0.25	0.06	0.10	0.13	-0.25	0.24	0.05
	$z_{cut} = 0.05, \beta = 0.0, N = 3$	3.18	0.06	0.42	0.06	0.10	0.78	-0.23	0.39	0.05
	$z_{cut} = 0.1, \beta = 0.0, N = 3$	0.47	0.01	0.10	0.03	0.07	-0.09	-0.22	0.09	0.02
	$z_{cut} = 0.05, \beta = 1.0, N = 3$	3.41	0.16	0.99	0.11	0.14	0.91	-0.24	0.97	0.11
	$z_{cut} = 0.1, \beta = 1.0, N = 3$	1.64	0.04	0.26	0.06	0.10	0.15	-0.25	0.26	0.05
Bottom-up SD	$z_{cut} = 0.05, \beta = 0.0$	2.97	0.07	0.47	0.06	0.09	0.71	-0.22	0.43	0.05
	$z_{cut} = 0.1, \beta = 0.0$	0.46	0.02	0.10	0.03	0.05	-0.07	-0.20	0.09	0.02
	$z_{cut} = 0.05, \beta = 1.0$	3.37	0.17	0.96	0.11	0.14	0.89	-0.24	0.95	0.11
	$z_{cut} = 0.1, \beta = 1.0$	1.74	0.04	0.26	0.06	0.10	0.15	-0.25	0.24	0.05
Pruning	$R_{cut} = 0.15, Z_{cut} = 0.25$	0.27	0.00	0.11	0.02	0.04	-0.11	-0.19	0.10	0.01
Trimming	$f_{cut} = 5\%, R_{sub} = 0.1$	-0.03	-0.04	-0.01	-0.02	0.00	-0.16	-0.18	0.01	0.00
	$f_{cut} = 9\%, R_{sub} = 0.1$	-0.03	-0.05	-0.03	-0.04	-0.03	-0.13	-0.16	-0.01	-0.02
	$f_{cut} = 5\%, R_{sub} = 0.2$	★ 0.14	-0.01	0.07	0.02	0.06	-0.21	-0.25	0.06	0.02
	$f_{cut} = 9\%, R_{sub} = 0.2$	0.10	-0.03	0.05	0.01	0.05	-0.20	-0.25	0.05	0.00

W tagging efficiency pileup dependence

New @ BOOST2020

Plot: CONF-2020-021

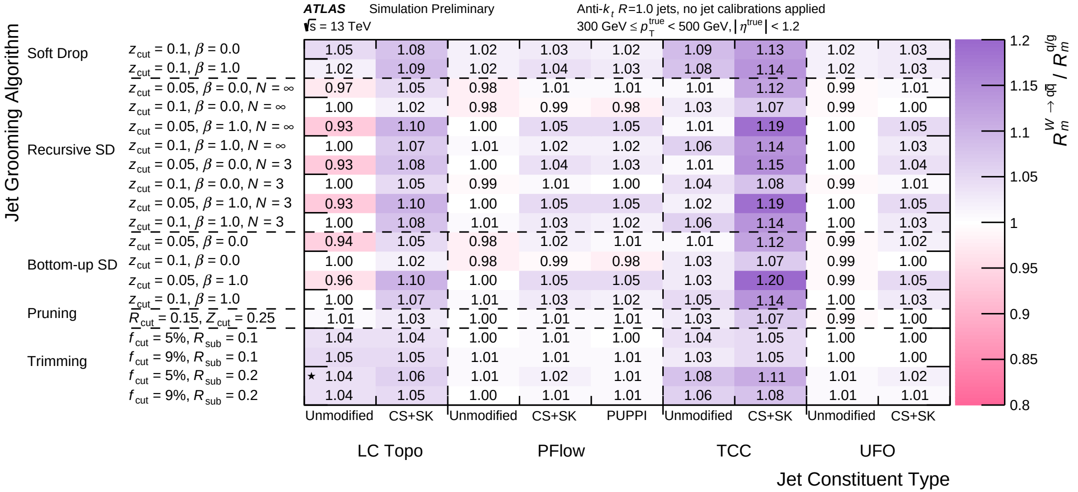


		LC Topo		PFlow			TCC		UFO	
Jet Grooming Algorithm		Unmodified	CS+SK	Unmodified	CS+SK	PUPPI	Unmodified	CS+SK	Unmodified	CS+SK
Soft Drop	z_{cut} = 0.1, β = 0.0	-1.20	-0.27	-0.63	-0.22	-0.32	-0.24	0.35	-0.58	-0.20
	z_{cut} = 0.1, β = 1.0	-1.64	-0.31	-0.71	-0.23	-0.31	-0.68	0.32	-0.70	-0.20
Recursive SD	z_{cut} = 0.05, β = 0.0, $N = \infty$	-1.34	-0.45	-0.98	-0.24	-0.41	-0.98	0.24	-1.02	-0.22
	z_{cut} = 0.1, β = 0.0, $N = \infty$	-1.18	-0.10	-0.38	-0.09	-0.15	0.03	0.24	-0.40	-0.09
	z_{cut} = 0.05, β = 1.0, $N = \infty$	-0.93	-0.59	-1.53	-0.41	-0.38	-1.27	-0.23	-1.55	-0.40
	z_{cut} = 0.1, β = 1.0, $N = \infty$	-1.56	-0.32	-0.76	-0.21	-0.26	-0.71	0.30	-0.78	-0.23
	z_{cut} = 0.05, β = 0.0, N = 3	-1.40	-0.46	-0.98	-0.23	-0.21	-1.09	0.28	-0.99	-0.23
	z_{cut} = 0.1, β = 0.0, N = 3	-1.19	-0.17	-0.46	-0.14	-0.20	-0.10	0.25	-0.45	-0.12
	z_{cut} = 0.05, β = 1.0, N = 3	-0.91	-0.58	-1.52	-0.41	-0.38	-1.26	-0.22	-1.54	-0.41
	z_{cut} = 0.1, β = 1.0, N = 3	-1.68	-0.35	-0.79	-0.22	-0.29	-0.72	0.33	-0.77	-0.24
Bottom-up SD	z_{cut} = 0.05, β = 0.0	-1.38	-0.47	-1.11	-0.24	-0.25	-1.09	0.23	-1.12	-0.25
	z_{cut} = 0.1, β = 0.0	-1.33	-0.14	-0.47	-0.12	-0.21	-0.04	0.23	-0.45	-0.09
	z_{cut} = 0.05, β = 1.0	-0.85	-0.59	-1.55	-0.42	-0.38	-1.26	-0.20	-1.57	-0.40
	z_{cut} = 0.1, β = 1.0	-1.56	-0.35	-0.78	-0.21	-0.26	-0.74	0.30	-0.80	-0.23
Pruning	R_{cut} = 0.15, Z_{cut} = 0.25	-1.04	0.00	-0.37	-0.04	-0.12	0.19	0.31	-0.38	-0.05
Trimming	f_{cut} = 5%, R_{sub} = 0.1	-0.02	0.08	0.01	0.04	0.01	0.18	0.20	-0.07	0.01
	f_{cut} = 9%, R_{sub} = 0.1	-0.01	0.11	0.03	0.09	0.07	0.18	0.22	-0.03	0.03
	f_{cut} = 5%, R_{sub} = 0.2	★ -0.40	-0.15	-0.20	-0.05	-0.13	0.30	0.30	-0.09	-0.04
	f_{cut} = 9%, R_{sub} = 0.2	-0.23	-0.07	-0.25	-0.05	-0.20	0.29	0.33	-0.21	-0.04

W vs QCD response (topology dependence)

New @ BOOST2020

Plot: CONF-2020-021

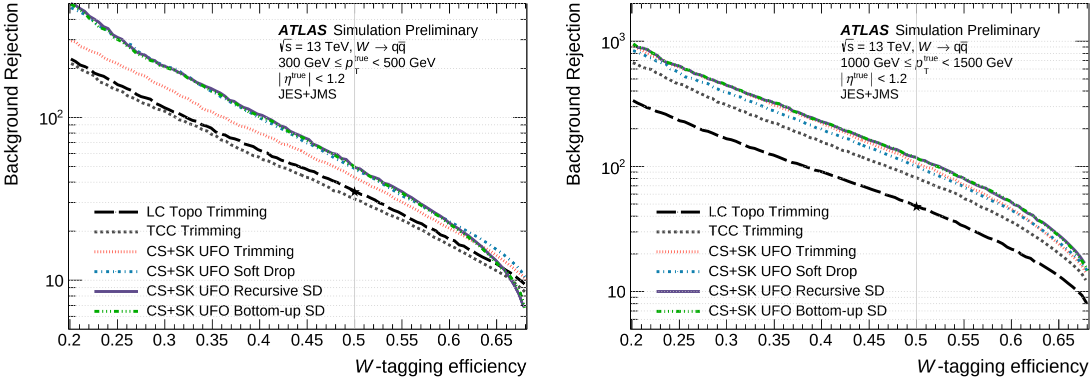


Jet Grooming Algorithm		LC Topo		PFlow			TCC		UFO	
		Unmodified	CS+SK	Unmodified	CS+SK	PUPPI	Unmodified	CS+SK	Unmodified	CS+SK
Soft Drop	z_{cut} = 0.1, β = 0.0	1.05	1.08	1.02	1.03	1.02	1.09	1.13	1.02	1.03
	z_{cut} = 0.1, β = 1.0	1.02	1.09	1.02	1.04	1.03	1.08	1.14	1.02	1.03
Recursive SD	z_{cut} = 0.05, β = 0.0, $N = \infty$	0.97	1.05	0.98	1.01	1.01	1.01	1.12	0.99	1.01
	z_{cut} = 0.1, β = 0.0, $N = \infty$	1.00	1.02	0.98	0.99	0.98	1.03	1.07	0.99	1.00
	z_{cut} = 0.05, β = 1.0, $N = \infty$	0.93	1.10	1.00	1.05	1.05	1.01	1.19	1.00	1.05
	z_{cut} = 0.1, β = 1.0, $N = \infty$	1.00	1.07	1.01	1.02	1.02	1.06	1.14	1.00	1.03
	z_{cut} = 0.05, β = 0.0, N = 3	0.93	1.08	1.00	1.04	1.03	1.01	1.15	1.00	1.04
	z_{cut} = 0.1, β = 0.0, N = 3	1.00	1.05	0.99	1.01	1.00	1.04	1.08	0.99	1.01
	z_{cut} = 0.05, β = 1.0, N = 3	0.93	1.10	1.00	1.05	1.05	1.02	1.19	1.00	1.05
	z_{cut} = 0.1, β = 1.0, N = 3	1.00	1.08	1.01	1.03	1.02	1.06	1.14	1.00	1.03
Bottom-up SD	z_{cut} = 0.05, β = 0.0	0.94	1.05	0.98	1.02	1.01	1.01	1.12	0.99	1.02
	z_{cut} = 0.1, β = 0.0	1.00	1.02	0.98	0.99	0.98	1.03	1.07	0.99	1.00
	z_{cut} = 0.05, β = 1.0	0.96	1.10	1.00	1.05	1.05	1.03	1.20	0.99	1.05
	z_{cut} = 0.1, β = 1.0	1.00	1.07	1.01	1.03	1.02	1.05	1.14	1.00	1.03
Pruning	R_{cut} = 0.15, Z_{cut} = 0.25	1.01	1.03	1.00	1.01	1.01	1.03	1.07	0.99	1.00
Trimming	f_{cut} = 5%, R_{sub} = 0.1	1.04	1.04	1.00	1.01	1.00	1.04	1.05	1.00	1.00
	f_{cut} = 9%, R_{sub} = 0.1	1.05	1.05	1.01	1.01	1.01	1.03	1.05	1.00	1.00
	f_{cut} = 5%, R_{sub} = 0.2	★ 1.04	1.06	1.01	1.02	1.01	1.08	1.11	1.01	1.02
	f_{cut} = 9%, R_{sub} = 0.2	1.04	1.05	1.00	1.01	1.01	1.06	1.08	1.01	1.01

W-tagging performance of finalist definitions

New @ BOOST2020

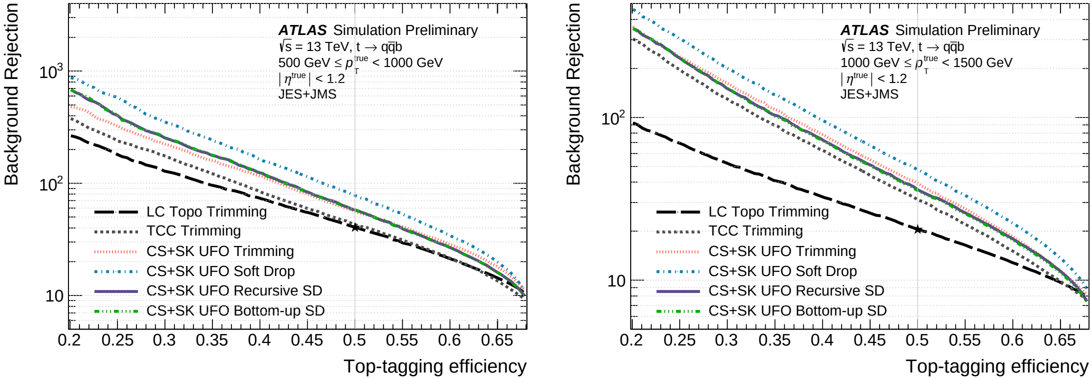
Both: CONF-2020-021



Top-tagging performance of finalist definitions

New @ BOOST2020

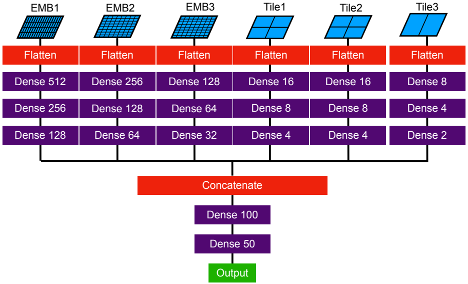
Both: CONF-2020-021



Classification architectures

New @ BOOST2020

Both: PUB-2020-018



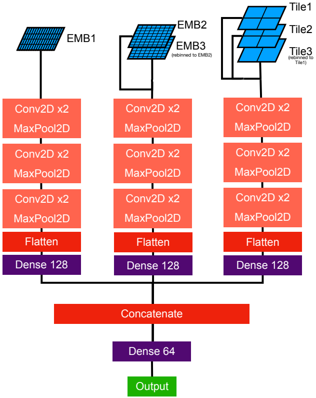
EMB1	EMB2	EMB3	Tile1	Tile2	Tile3
Flatten	Flatten	Flatten	Flatten	Flatten	Flatten
Dense 512	Dense 256	Dense 128	Dense 16	Dense 16	Dense 8
Dense 256	Dense 128	Dense 64	Dense 8	Dense 8	Dense 4
Dense 128	Dense 64	Dense 32	Dense 4	Dense 4	Dense 2

Concatenate

Dense 100

Dense 50

Output



EMB1	EMB2 EMB3 (rebinned to EMB2)	Tile1 Tile2 Tile3 (rebinned to Tile1)
Conv2D x2 MaxPool2D	Conv2D x2 MaxPool2D	Conv2D x2 MaxPool2D
Conv2D x2 MaxPool2D	Conv2D x2 MaxPool2D	Conv2D x2 MaxPool2D
Conv2D x2 MaxPool2D	Conv2D x2 MaxPool2D	Conv2D x2 MaxPool2D
Flatten	Flatten	Flatten
Dense 128	Dense 128	Dense 128

Concatenate

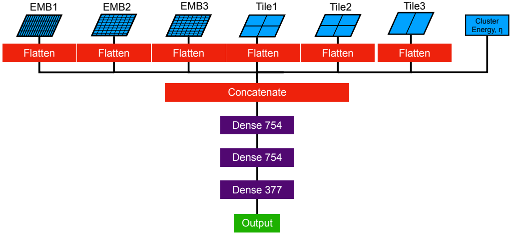
Dense 64

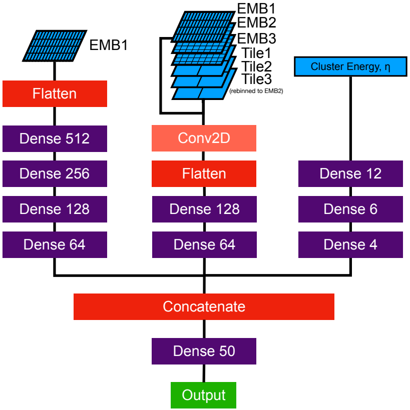
Output

Regression architectures

New @ BOOST2020

Both: PUB-2020-018

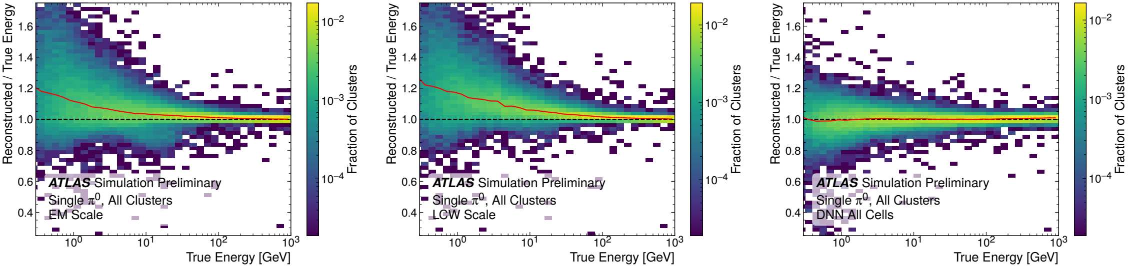




Regression performance, neutral pions

New @ BOOST2020

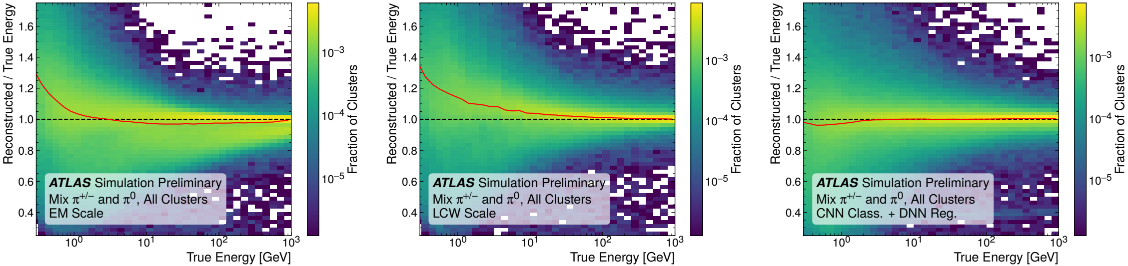
Both: PUB-2020-018



Regression performance, mixed pions

New @ BOOST2020

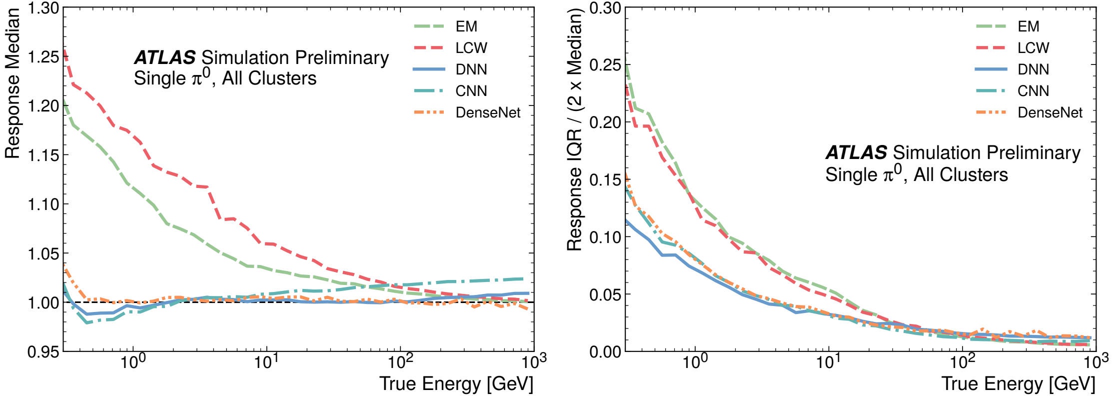
Both: PUB-2020-018



Regression performance, neutral pions

New @ BOOST2020

Both: PUB-2020-018



Regression performance, charged pions

New @ BOOST2020

Both: PUB-2020-018