ATLAS Boosted Object Tagging: Vector Boson

Julien Caudron
on behalf of the ATLAS Collaboration

Johannes-Gutenberg Universität Mainz

August 11th 2015, BOOST2015, University of Chicago

PRiSMA Cluster of Excellence

BMBF-Forschungsschwerpunkt ATLAS Experiment
Physics on the TeV-scale at the Large Hadron Collider
FSP 101 ATLAS

Introduction

Motivation:

Including hadronically decaying boson increases sensitivity,
especially because the hadronic decay has a large branching ratio
→ discriminate *boson-jets* from *QCD-jets*
Boosted region provides a better S/B ratio

W-tagging techniques can be extended to tag other bosons.
Additionally, the next natural step consists in distinguishing between W and Z
(WZ/WW separation, BSM diboson, leptophobic FCNC)

Content:

- **W-tagging in Run1** (Grooming and Substructure variables) — **PERF-2015-03**
- **W- and Z-tagging in Run 2** — **ATL-PHYS-PUB-2015-033**
- **W/Z separation in Run 1** — **PERF-2015-02**

W-tagging in Run1

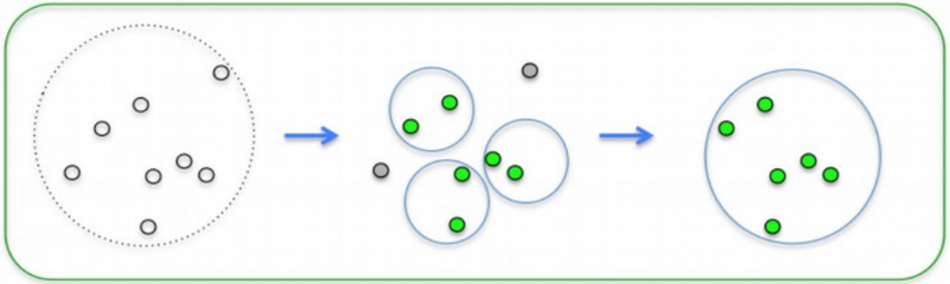
Material from **PERF-2015-03**

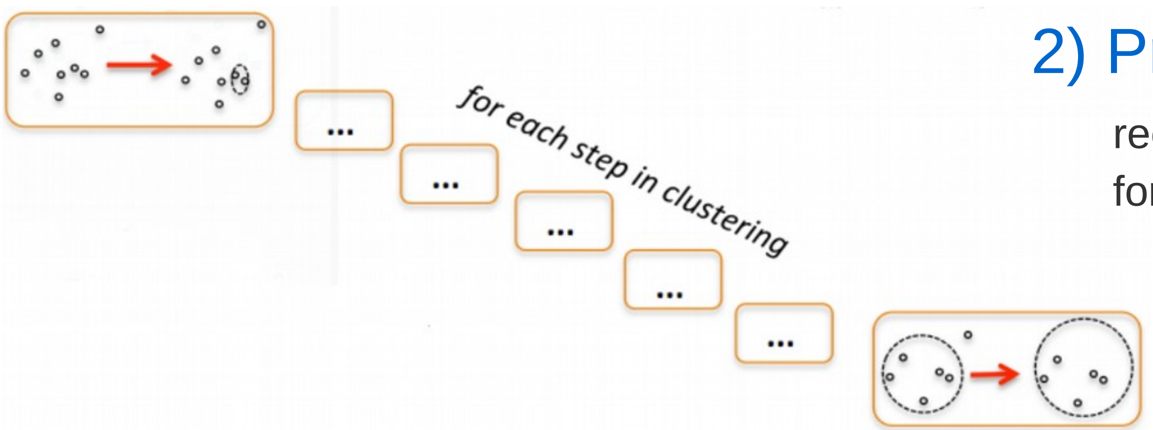
Grooming 1/4

1) Trimming

reclustering: kt algorithm, with $R=R_{sub}$

If p_T(subjet) < f_{cut} x p_T(jet) → discard subjet





2) Pruning

reclustering: C/A algorithm

for pair ij with $p_T(i) < p_T(j)$, if $\Delta R_{ij} > R_{cut} \times 2m/p_T$

or $p_T(i)/p_T(j) < z_{cut}$

→ discard object i

3) Split-Filtering (mass-drop, BDRS)

1) Splitting:

de-clustering C/A algorithm

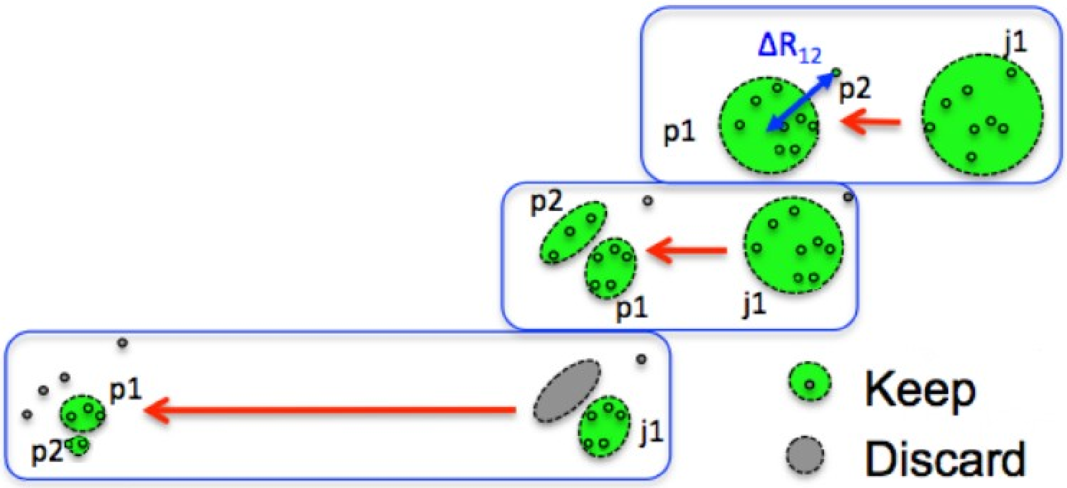
stop when $\max(m_1,m_2)/m_{12} < \mu_{frac}$ and $y_{12} > y_{cut}$

otherwise, discard lowest mass object

2) Filtering:

reclustering with $R=R_{sub}$,

only the 3 hardest subjets are kept



Grooming 2/4

Three grooming techniques have been studied → >500 jet collections

1) Trimming

Trimming configurations

Input jet algorithms	R	R_{sub}	f_{cut} (%)
C/A, anti-k_t	0.6, 0.8, 1.0, 1.2	0.1, 0.2, 0.3	1, 2, 3, 4 , 5, 7, 9, 11, 13, 15

→ 2 x 4 x 3 x 10 = 240

2) Pruning

Pruning configurations

Input jet algorithm	R	Reclust. alg.	Z_{cut} (%)	R_{cut}
C/A, anti-k_t	0.8, 1.0, 1.2	C/A	10, 15, 20, 25, 30	$\frac{1}{100}, \frac{1}{10}, \frac{1}{8}, \frac{1}{4}, \frac{1}{2}$, 1.0

→ 2 x 3 x 1 x 5 x 6 = 180

3) Split-Filtering

Split-filtering configurations

Input jet algorithm	R	R_{sub}	μ_{frac}	y_{cut}
C/A	0.8, 1.0, 1.2	0.3, $min(0.3, \Delta R/2)$	67, 78, 89, 100	0.06, 0.07, ..., 0.20

→ 1 x 3 x 2 x 4 x 11 = 264

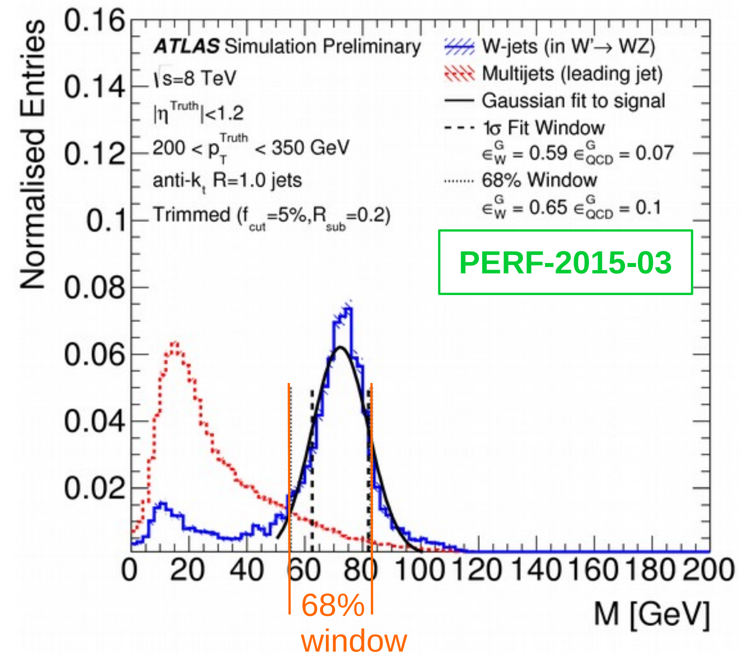
Study of the grooming techniques

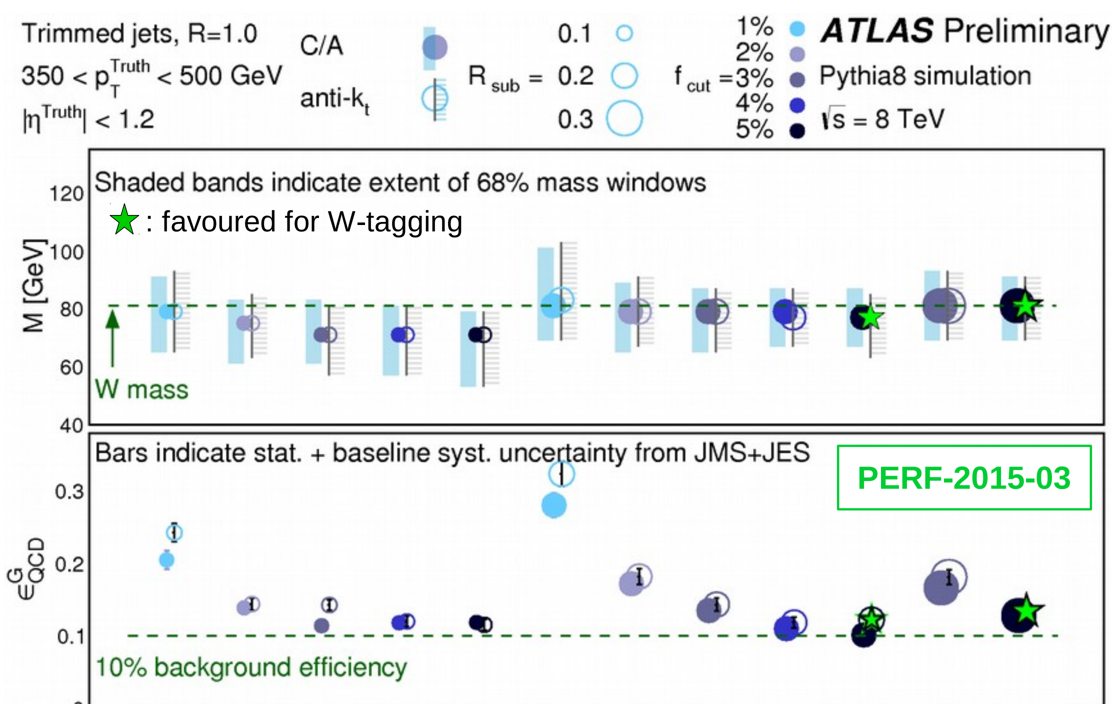
Smallest mass windows containing 68% of signal are defined

For 3 region of truth-jet p_T (200 - 350 - 500 - 1000 GeV)

Reject pathological cases:

- the 68% mass window contains m_W
- mass peak is too asymmetric
- background in the window has irregular shape
- not sufficiently stable w.r.t. Pile-Up





Sort by best QCD multijets background rejection in the mass window

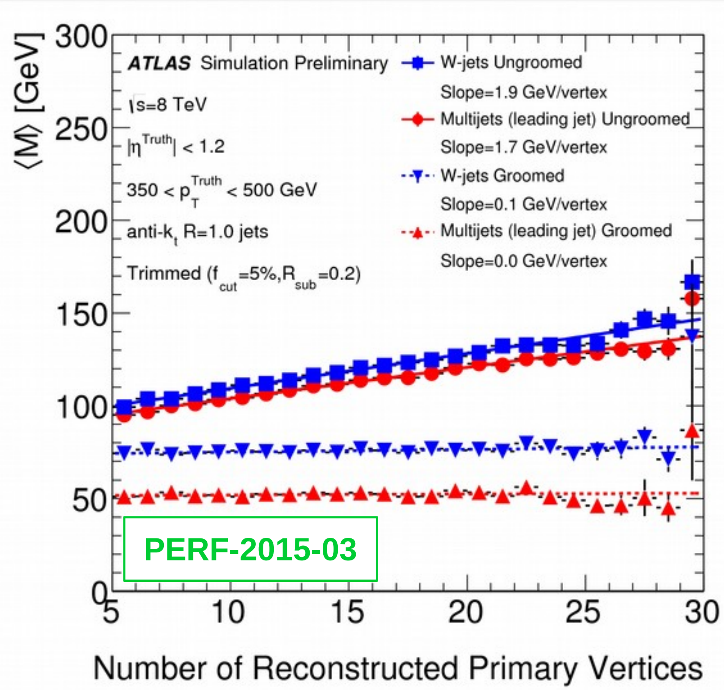
← example (trimming)

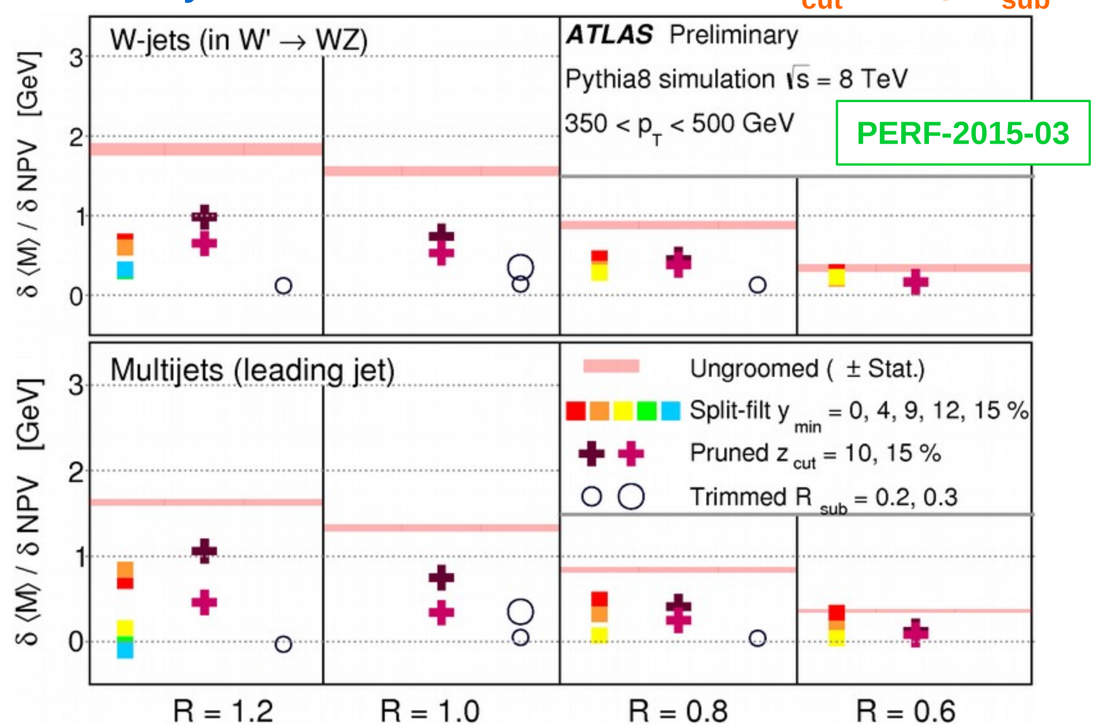
→ 27 best jet collections are kept

Study observations:

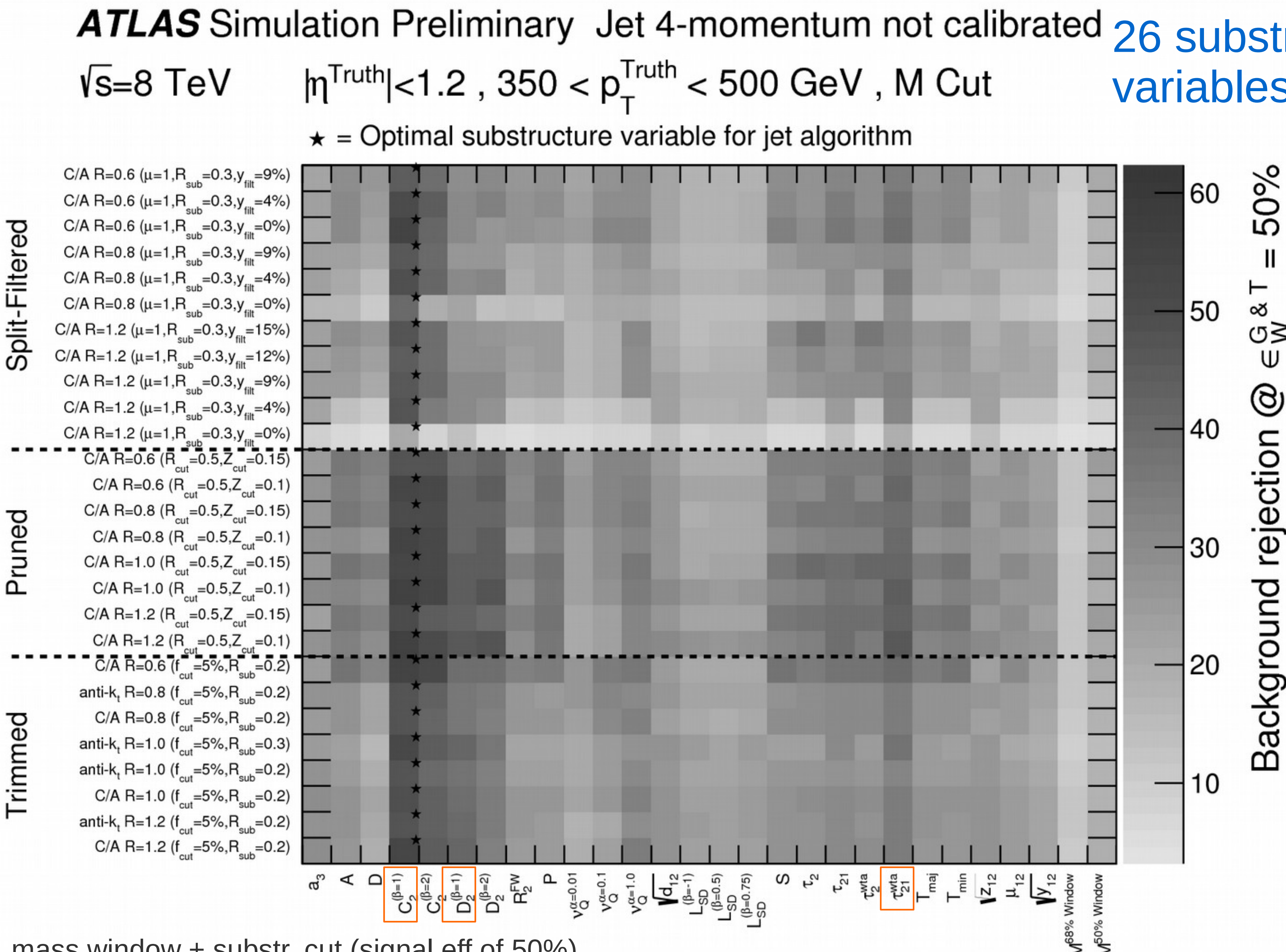
- R=0.6 or R=0.8 are too small (loss of constituents for p_T < 500 or 350GeV respectively)
- For high p_T, performances of the different grooming techniques are similar
- For **trimming**, C/A and anti-kt are similar, and best for large values of f_{cut}
- For **pruning**, C/A always better than kt, and for low p_T, other grooming methods preferred
- For **split-filtering**, larger y_{cut} is better, optimal param. depends of p_T, but less dependence with p_T and other param for y_{cut} > 0.09
- Grooming reduces Pile-Up dependency (so, no additional P-U removal procedure needed)

→best grooming: low bkgd eff. + good P-U stability → **anti-kt R=1.0 trimmed f_{cut}=5%, R_{sub}=0.2**





Substructure variables 1/3



26 substructure variables studied

mass window + substr. cut (signal eff of 50%)

Substructure variables 2/3

Best variables:

For the constituents *i* inside the large-R jet J

Mass: $$M^2 = (\sum_i E_i)^2 - (\sum_i p_i)^2$$

τ_{21}^{wta}:

$$\tau_0(\beta) = \sum_{i\in J} p_{T_i} \Delta R^\beta,$$

$$\tau_1(\beta) = \frac{1}{\tau_0(\beta)} \sum_{i\in J} p_{T_i} \Delta R^\beta_{a_1,i},$$

$$\tau_2(\beta) = \frac{1}{\tau_0(\beta)} \sum_{i\in J} p_{T_i} \min(\Delta R^\beta_{a_1,i}, \Delta R^\beta_{a_2,i}),$$

τ_N: exactly N subjets kt reconstructed

a_i : axis of the I[th] subjet (if Winner-Take-All (wta), axis of the hardest constituent)

here: $\beta = 1$

$$\tau_{21} = \frac{\tau_2}{\tau_1}, \quad \tau_{21}^{wta} = \frac{\tau_2^{wta}}{\tau_1^{wta}}$$

$C_2^{(\beta=1)}$ and $D_2^{(\beta=1)}$:

$$E_{CF}0(\beta) = 1,$$

$$E_{CF}1(\beta) = \sum_{i\in J} p_{T_i},$$

$$E_{CF}2(\beta) = \sum_{i<j\in J} p_{T_i} p_{T_j} \left(\Delta R_{ij}\right)^\beta,$$

$$E_{CF}3(\beta) = \sum_{i<j<k\in J} p_{T_i} p_{T_j} p_{T_k} \left(\Delta R_{ij} \Delta R_{ik} \Delta R_{jk}\right)^\beta,$$

→

$$e_2^{(\beta)} = \frac{E_{CF}2}{E_{CF}1^2},$$

$$e_3^{(\beta)} = \frac{E_{CF}3}{E_{CF}1^3}$$

→

$$C_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^2},$$

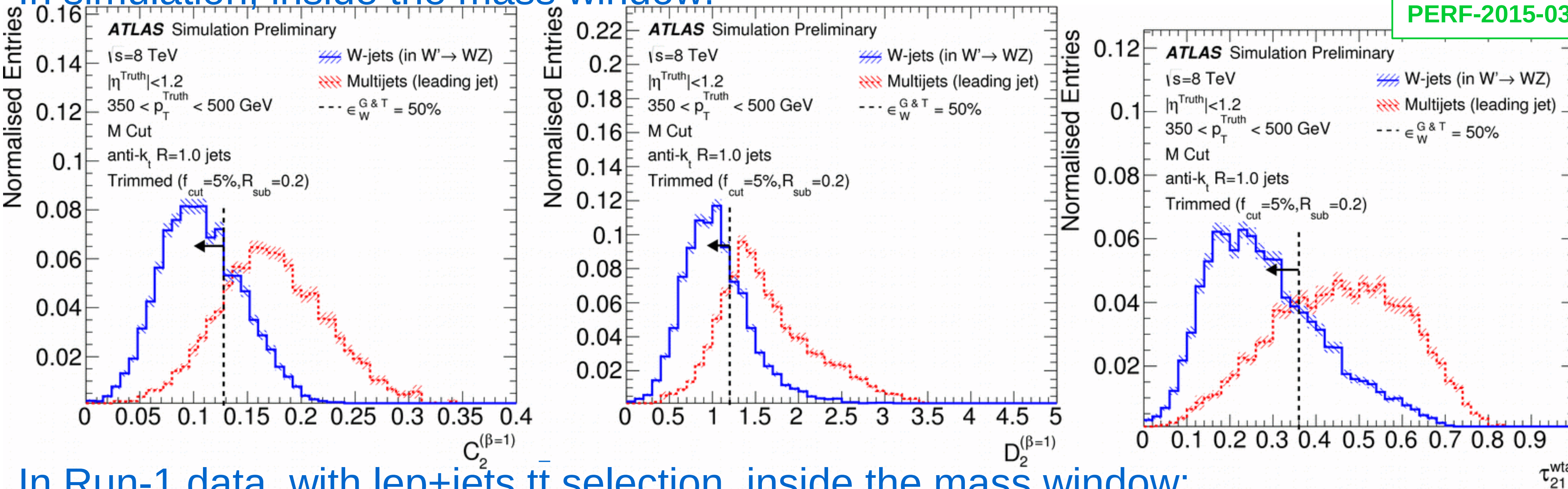
$$D_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3}$$

→ W-Tagger defined as grooming + mass window + cut on one substructure variable

Substructure variables 3/3

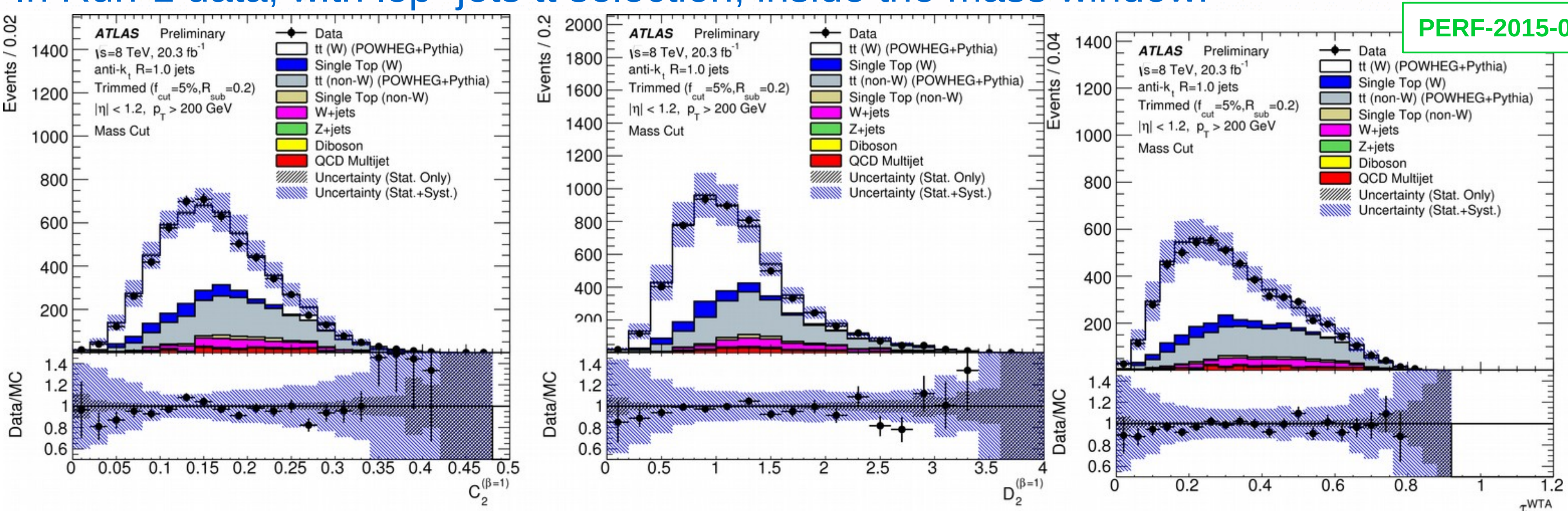
In simulation, inside the mass window:

PERF-2015-03



In Run-1 data, with lep+jets $t\bar{t}$ selection, inside the mass window:

PERF-2015-03



W-tagging efficiencies 1/2

Two W-tagger working points: medium (50% signal eff.), tight (25% signal eff.)

Efficiency definition:

$$\varepsilon_{MC} = \frac{N^{\text{tagged}}_{\text{Boosted}-W}}{N^{\text{pre-tagged}}_{\text{Boosted}-W}} \qquad \varepsilon_{\text{Data}} = \frac{N^{\text{tagged}} - N^{\text{tagged}}_{\text{Bkg}}}{N^{\text{pre-tagged}} - N^{\text{pre-tagged}}_{\text{Bkg}}}$$

Boosted-W = jet matched (ΔR<1.0) with both W decay partons
Bkg = multijet QCD and W+Jets from data-driven methods, others from simulation

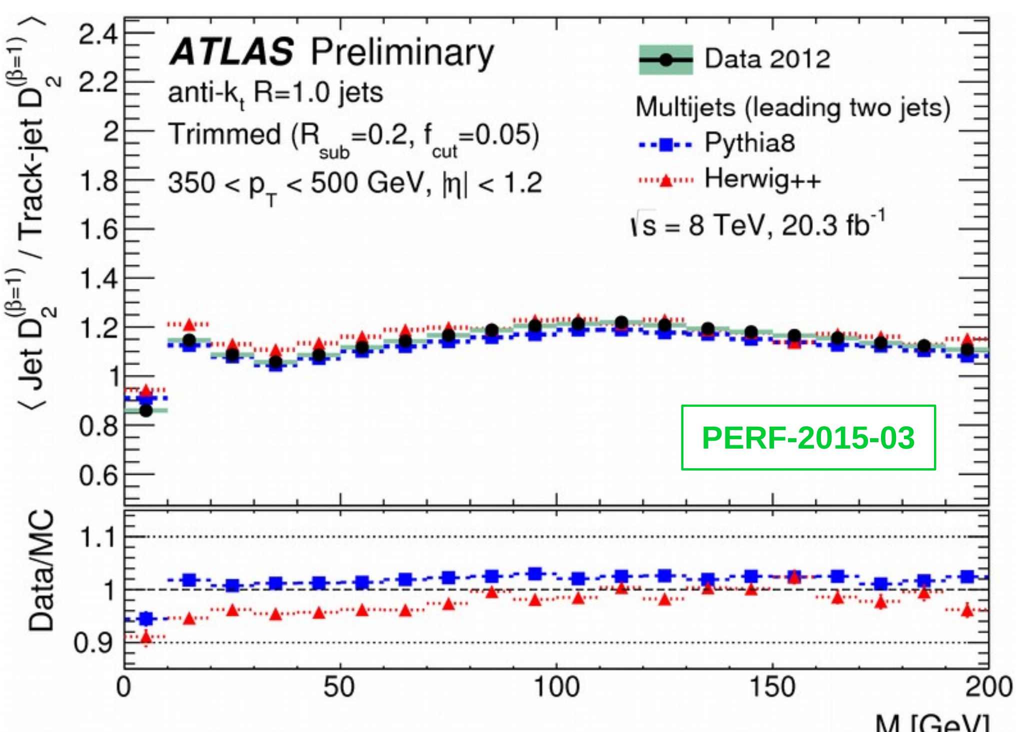
Systematic uncertainties:

- Large-R jets mass and energy scale/resolution
- Substructure variables (~5%)
- MC generator, ISR/FSR
- Backgrounds normalisation

→ 6-13%

Example for $D_2^{(\beta=1)}$ tagging with akt 1.0 trimming f_{cut}=5%, R_{sub}=0.2

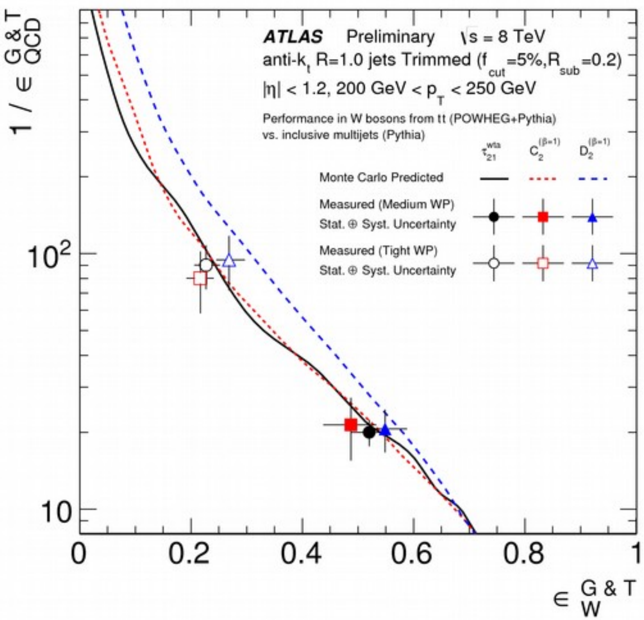
$D_2^{(\beta=1)}$ unc. from data-driven track-jet double-ratio method



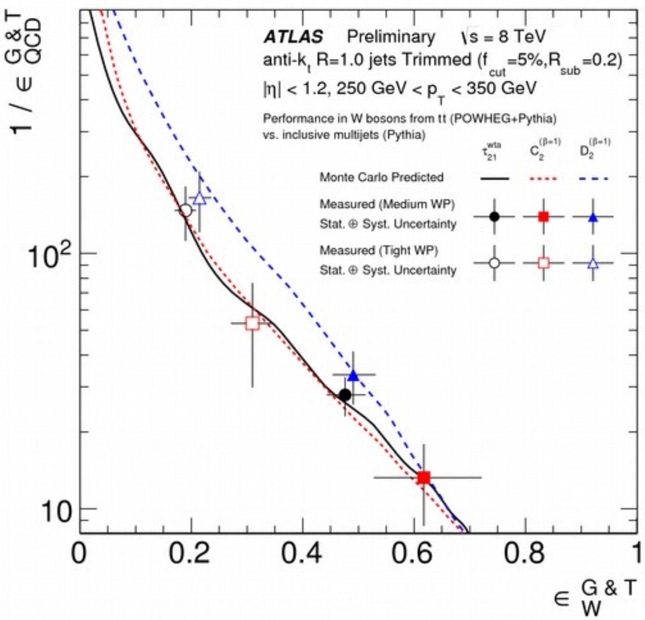
W-tagging efficiencies 2/2

Two W-tagger working points:

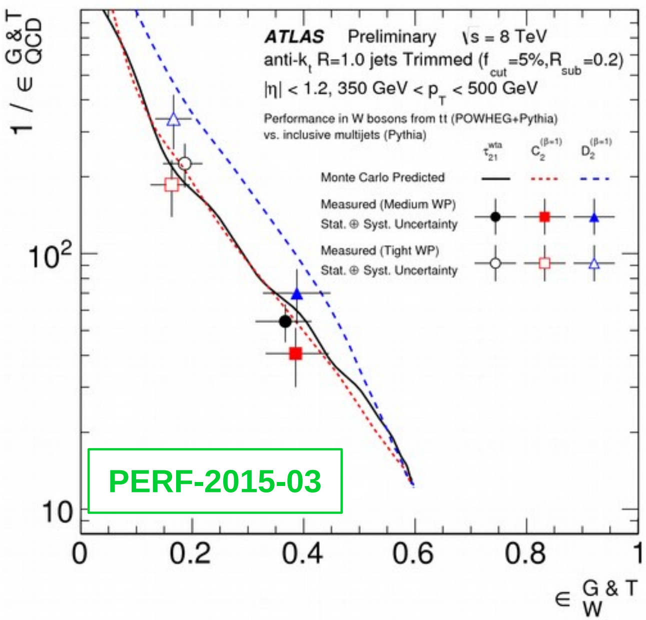
for anti-kt R=1.0 trimmed f_{cut}=5%, R_{sub}=0.2, cut on substr. variable



200 < p_T < 250 GeV



250 < p_T < 350 GeV



350 < p_T < 500 GeV

Data points are from fits using templates from Powheg+Pythia
Compared with simulation scan

Rem.:
- data from $t\bar{t}$ topology, while the working-points have been optimized in W' topology
- different MC generators (Powheg+Pythia vs. MC@NLO + Herwig) lead to different shapes

W- and Z-tagging in Run2

Material from **ATL-PHYS-PUB-2015-033**

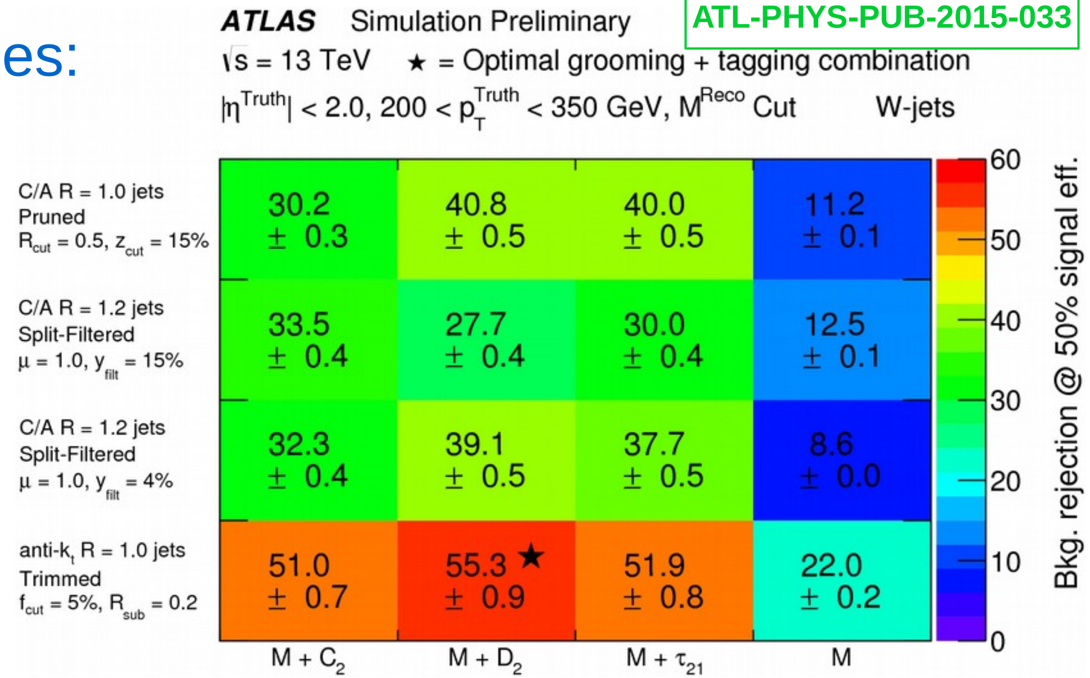
Strategy

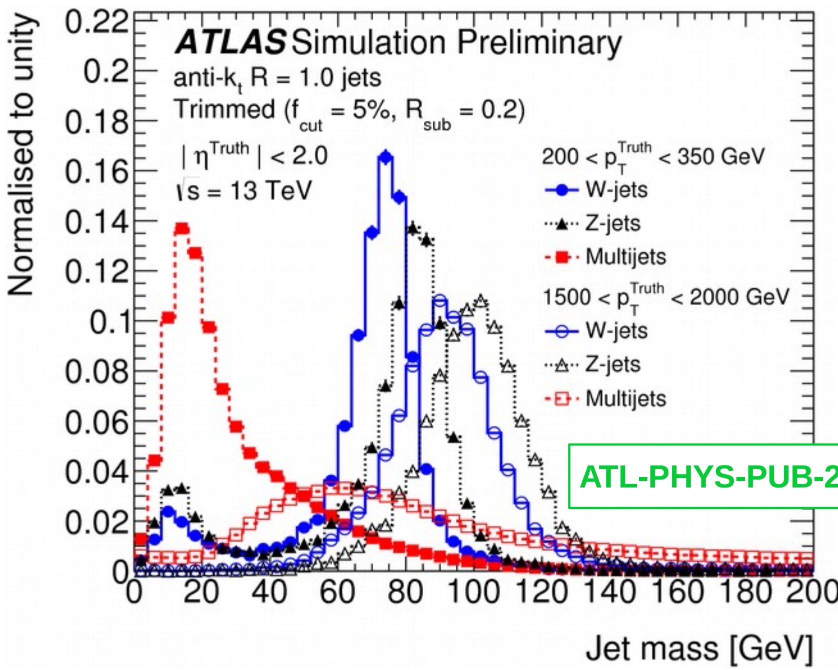
Early W-tagger for early Run2 analyses:

similar approach than in Run1:
grooming + substructure cut

based on lesson from Run1:

- selected number of grooming techniques (4 instead of 27)
- selected number of substructure variables (3: τ_{21}^{wta}, $C_2^{(\beta=1)}$, $D_2^{(\beta=1)}$)





For W and Z-tagger

Syst. unc. from Run1 + Run2 MC

Final tagger based on:

- anti-kt R=1.0 trimmed f_{cut}=5%, R_{sub}=0.2
- $D_2^{(\beta=1)}$ variable
- p_T parametrisation of the cut on $D_2^{(\beta=1)}$

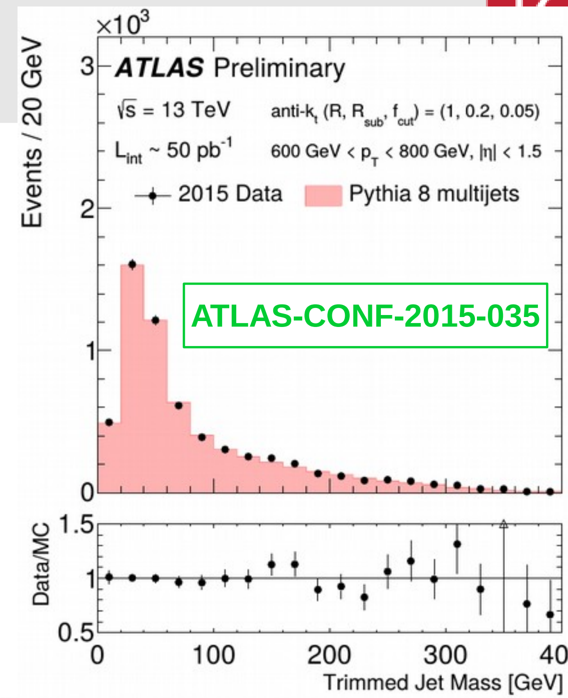
Variables in Run2

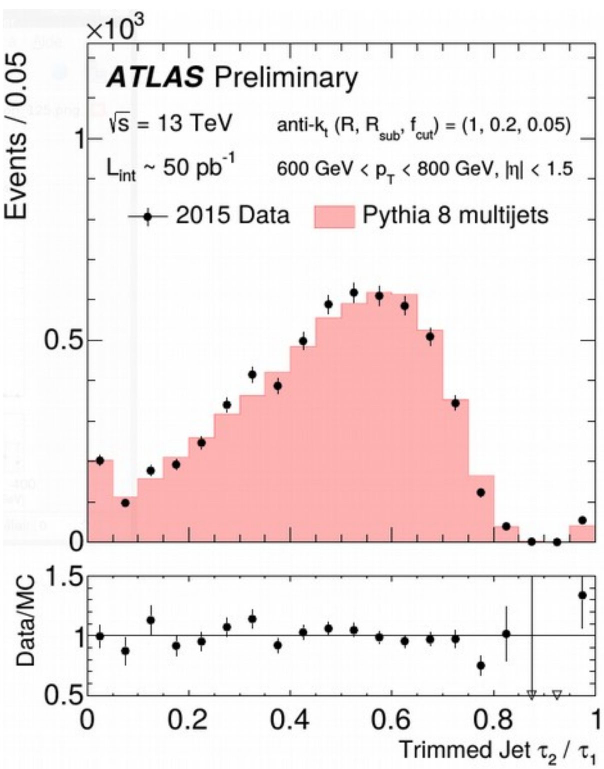
First data in Run2:

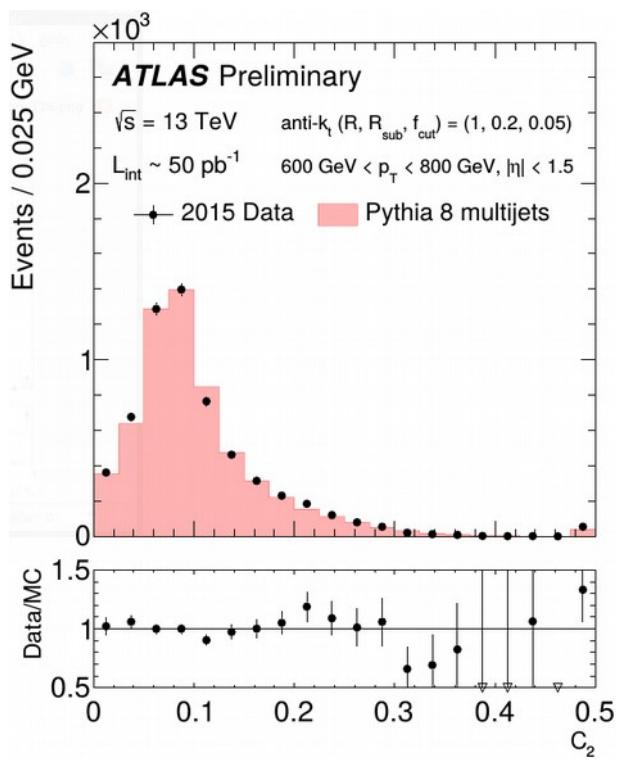
50 pb^{-1}, $\sqrt{s}$ = 13 TeV

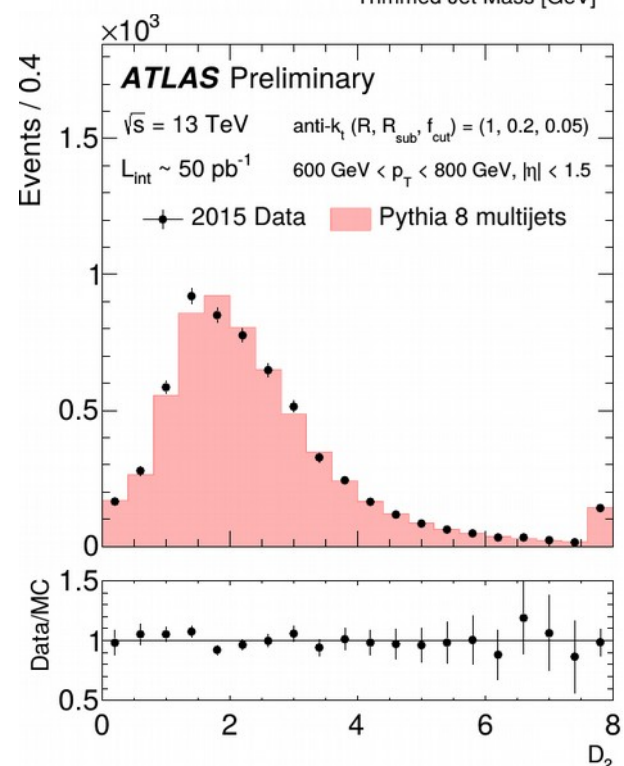
Anti-kt R=1.0 trimmed f_{cut}=5% R_{sub}=0.2

mass, τ_{21}^{wta}, $C_2^{(\beta=1)}$, $D_2^{(\beta=1)}$









Systematic uncertainties

Uncertainties on p_T, m and $D_2^{(\beta=1)}$: p_T < 1TeV: 3 Run2 variations + Run1 uncert. (quad. sum)

p_T > 1TeV: 7 Run2 variations + Stat. uncert. (quad. sum)

Run2 variations: (uncert. = deviation from nominal in truth-reco jet response)

Noise thr. : noise threshold variation in the calorimeter cells clustering procedure

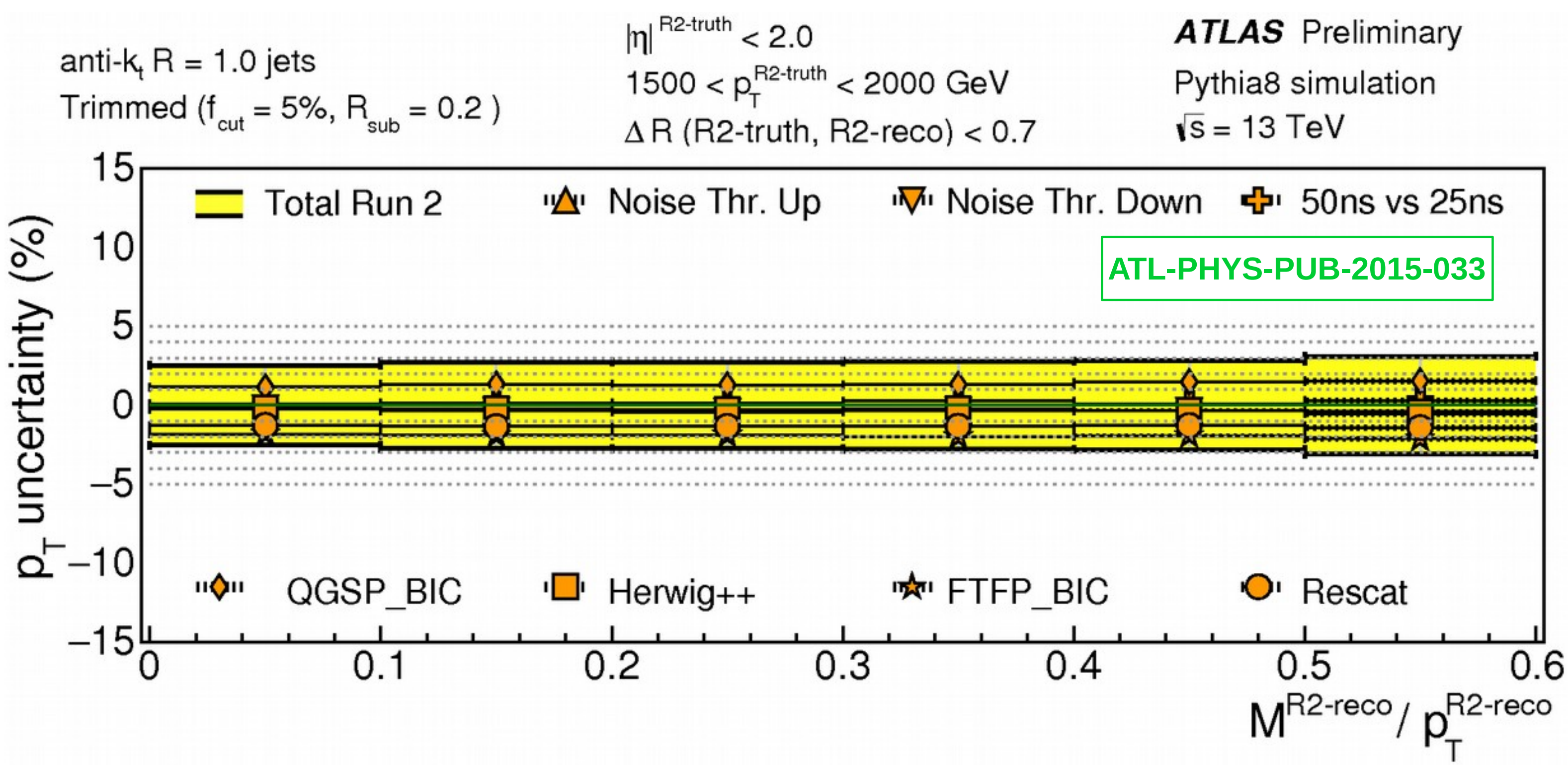
QGSP_BIC (FTFP_BIC): Quark Gluon String model (Fritiof String model) + Binary Intranuclear Cascade

Rescat.: High-energy re-scattering simulated with Binary Intranuclear Cascade

50ns vs 25ns:
bunch spacing condition var.

Herwig++:
Herwig++ generator

For p_T < 1TeV:
Run1 from track-jet double-ratio data-driven method



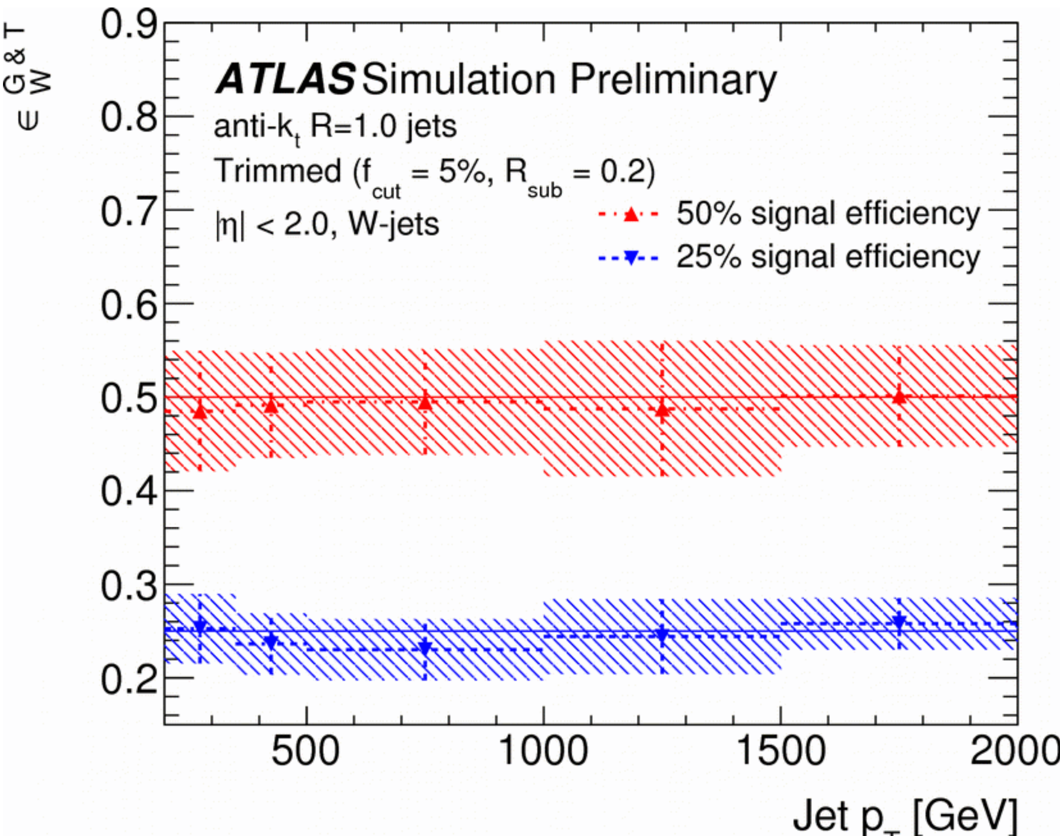
Efficiencies 1/2

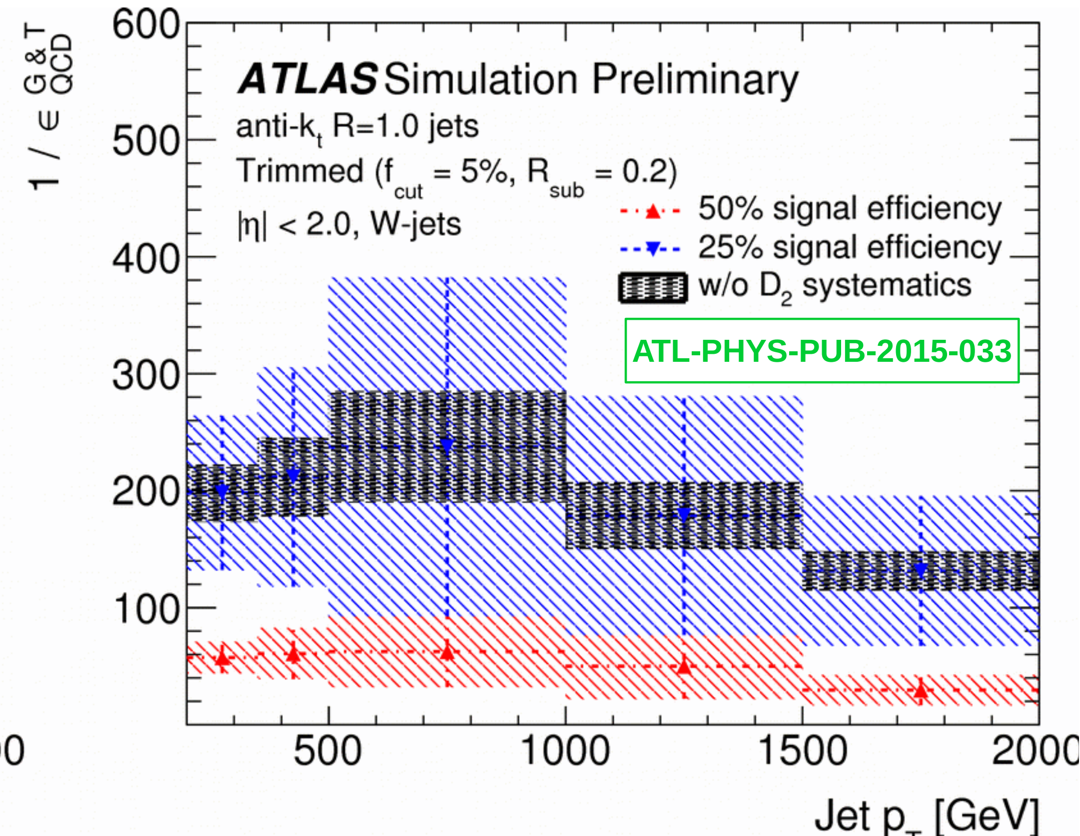
Two W-tagger working points: medium (50% signal eff.), tight (25% signal eff.)

Similar results available for Z-tagging (next slide)

Signal efficiencies and bkgd rejection, with respect to reco jet p_T

Larger uncertainties at larger p_T, usually from D_2 uncertainties



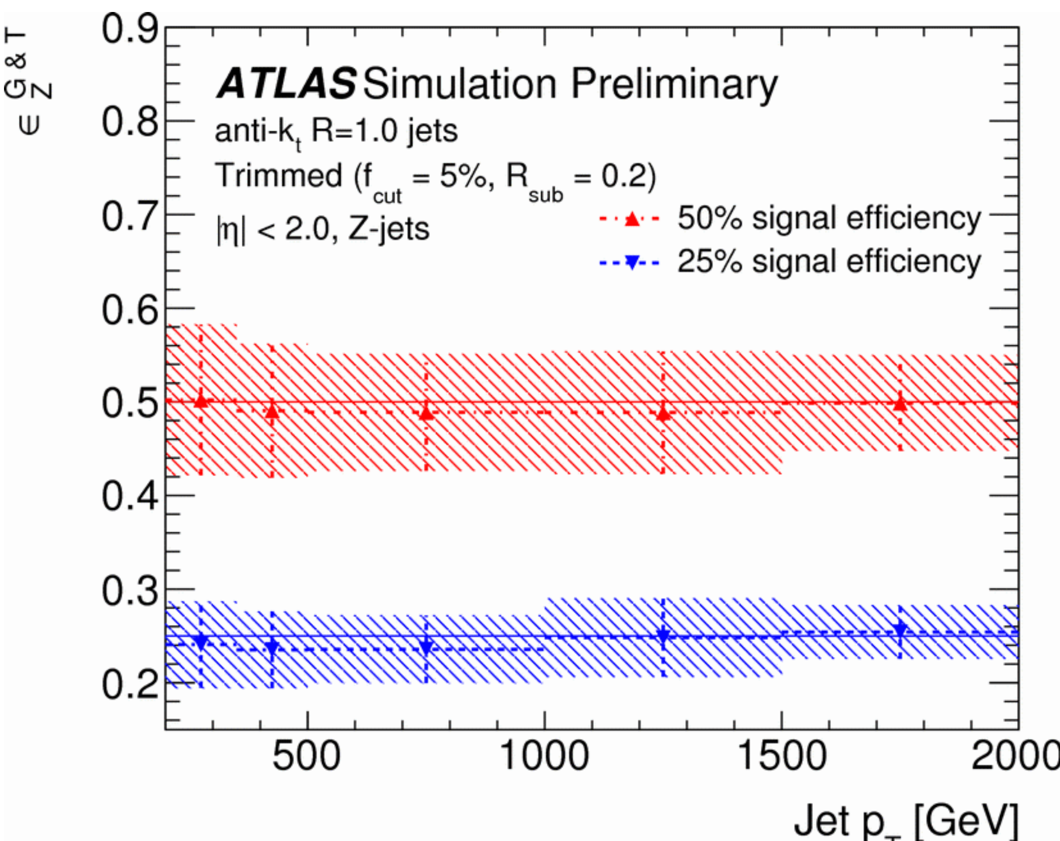


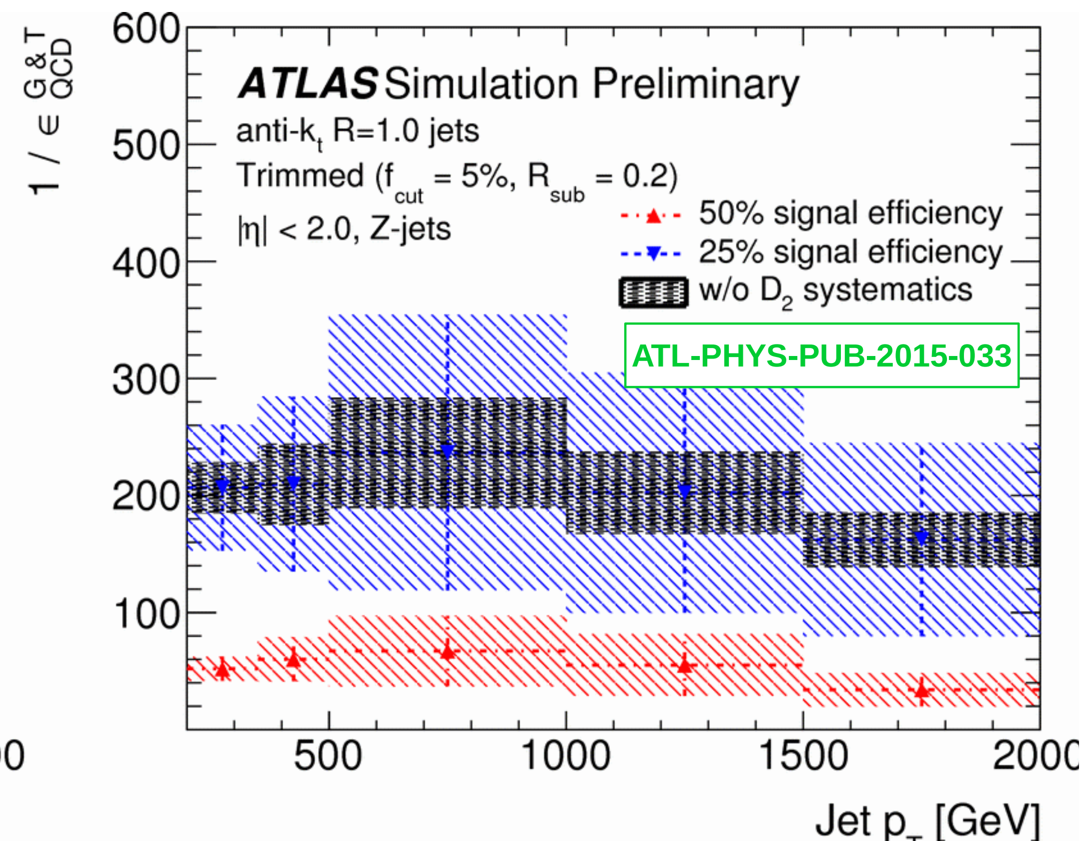
Two Z-tagger working points: medium (50% signal eff.), tight (25% signal eff.)

Similar results available for W-tagging

Signal efficiencies and bkgd rejection, with respect to reco jet p_T

Larger uncertainties at larger p_T, usually from D_2 uncertainties





W/Z separation in Run1

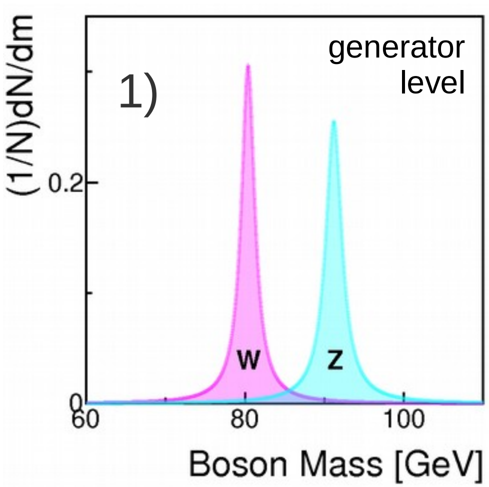
Material from **PERF-2015-02**

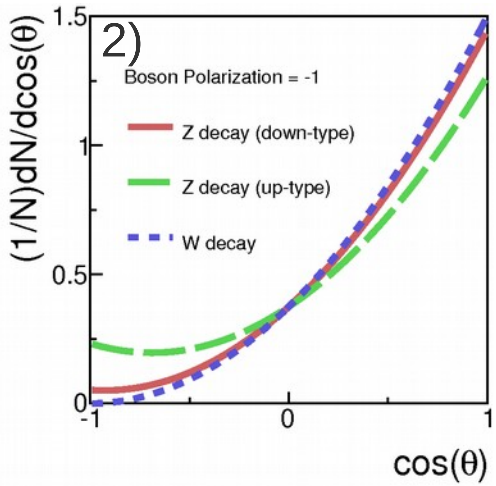
W/Z discriminating variables 1/2

Use of substructure variables to distinguish hadronically decaying W and Z

Differences between W and Z:

1) Mass
2) ~~Angular distributions~~ ← small at parton level, likely to be washed out by combinatorics, non-perturbative physics and reconstruction
3) Flavor decay ratio
4) Boson charge



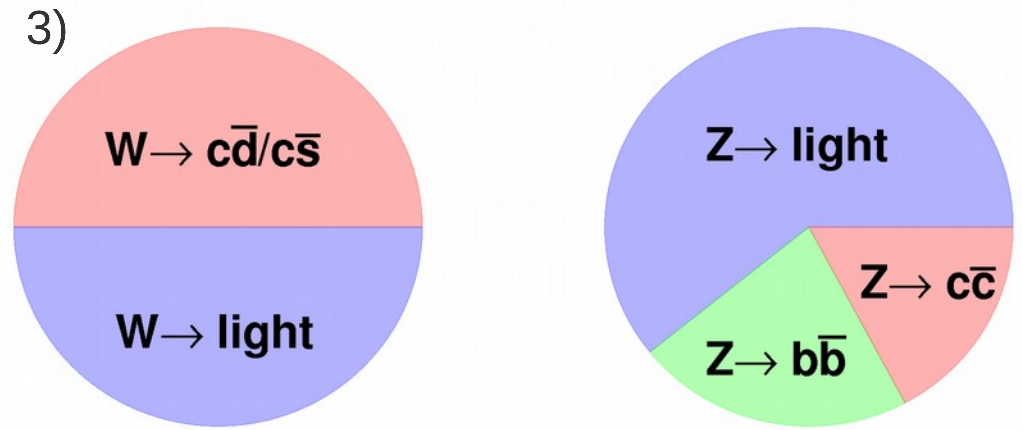


The boson is contained in a large-R jet:

anti-kt R=1.0, kt trimmed f_{cut}=5%, R_{sub}=0.3 calibrated

This large-R jet can be matched with small-R jets:

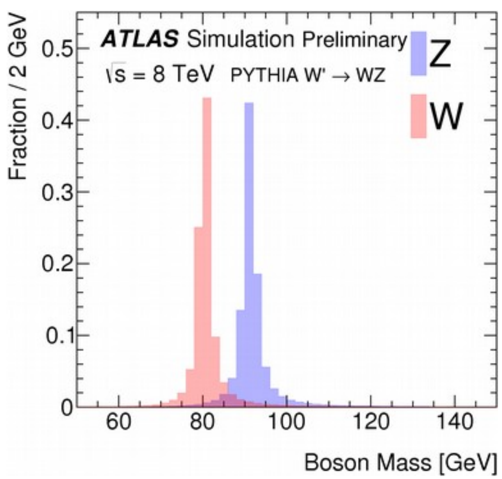
anti-kt R=0.4, matched with ΔR(j,J)<1.0 calibrated

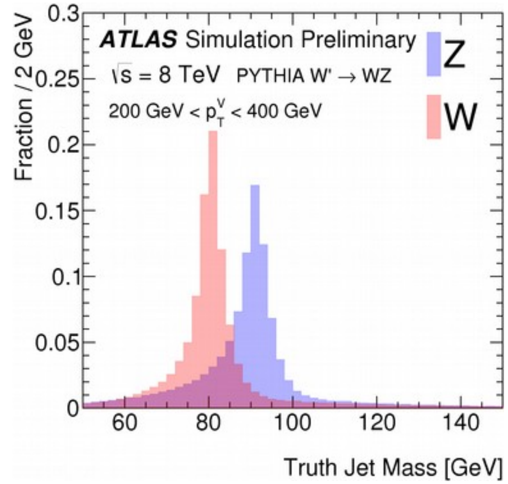


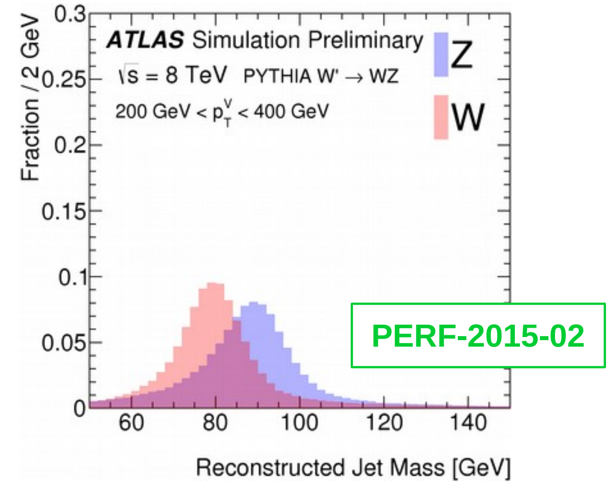
PERF-2015-02

W/Z discriminating variables 2/2

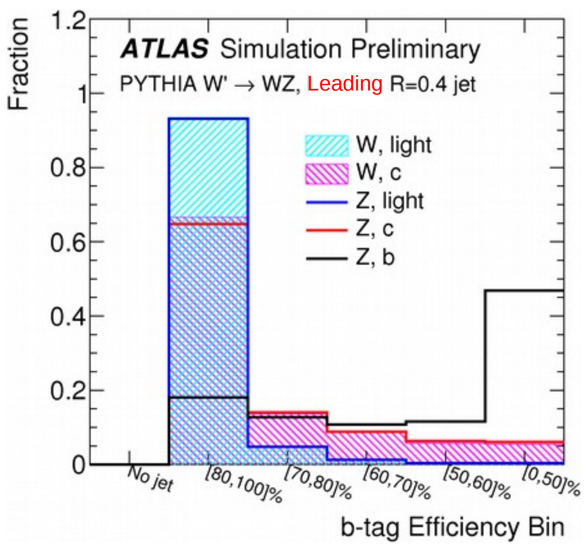
Mass resolution is degraded by parton fragmentation and reconstruction

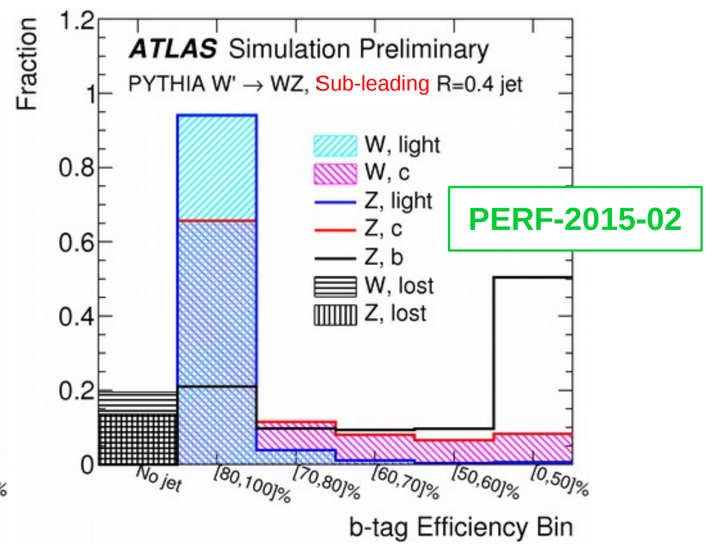




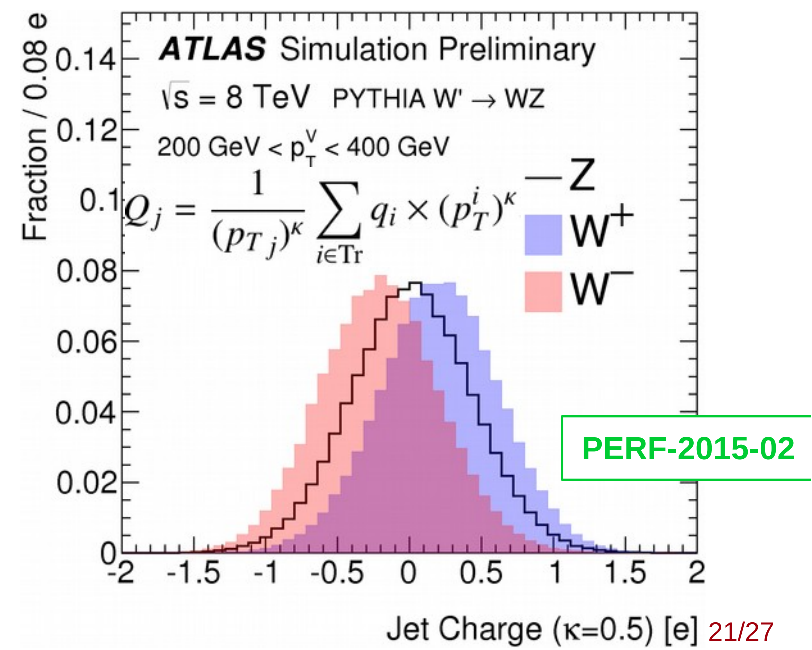


2 first leading small jets matched with large jet, b-tagging (MV1)





Charge (cf. B. Nachman's talk)

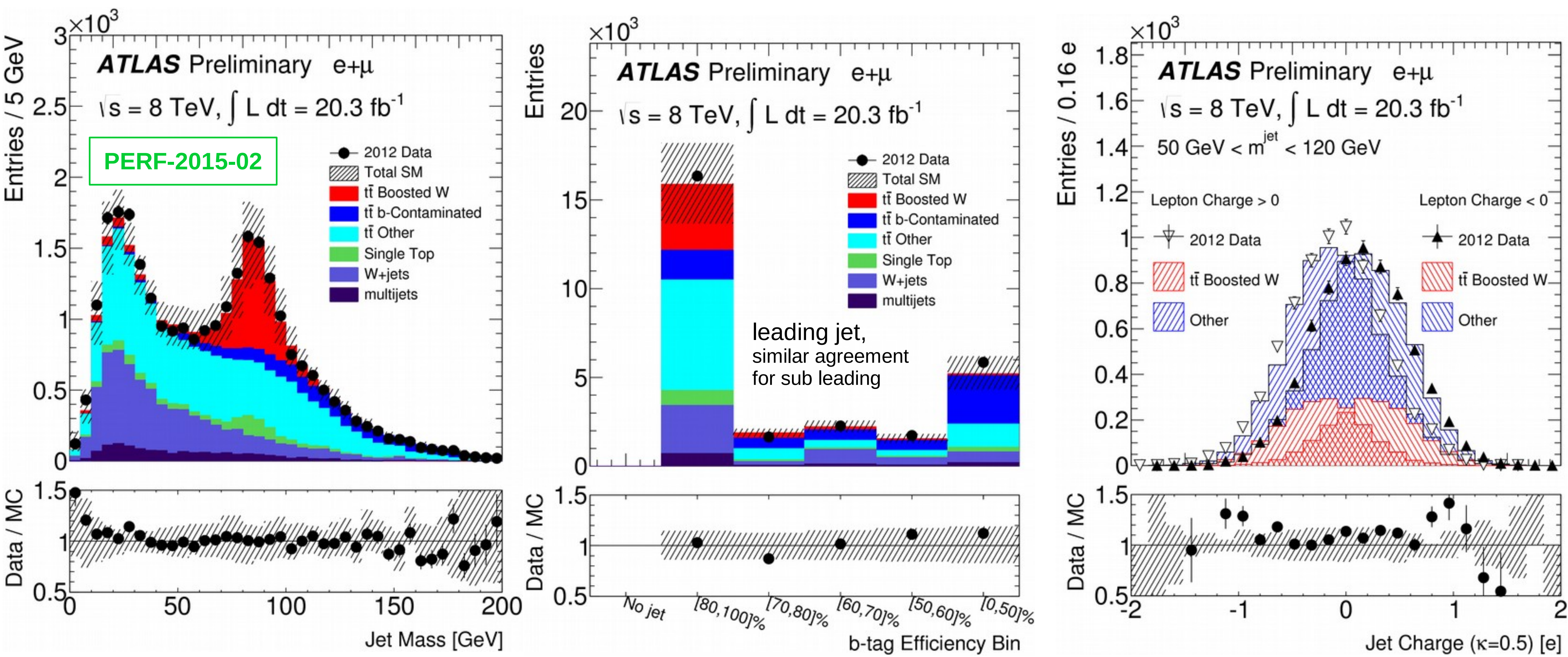


Comparison with Data

Semi-leptonic $t\bar{t}$ selection of hadronically decaying W

- MC $t\bar{t}$ divided into: $t\bar{t}$ Boosted W, $t\bar{t}$ b-contaminated, $t\bar{t}$ others (using MC matching)
- W+jets and multi-jets QCD from data-driven methods
- Others from simulation

Tests of the 3 different input variables in data → well modelled



Tagger definition

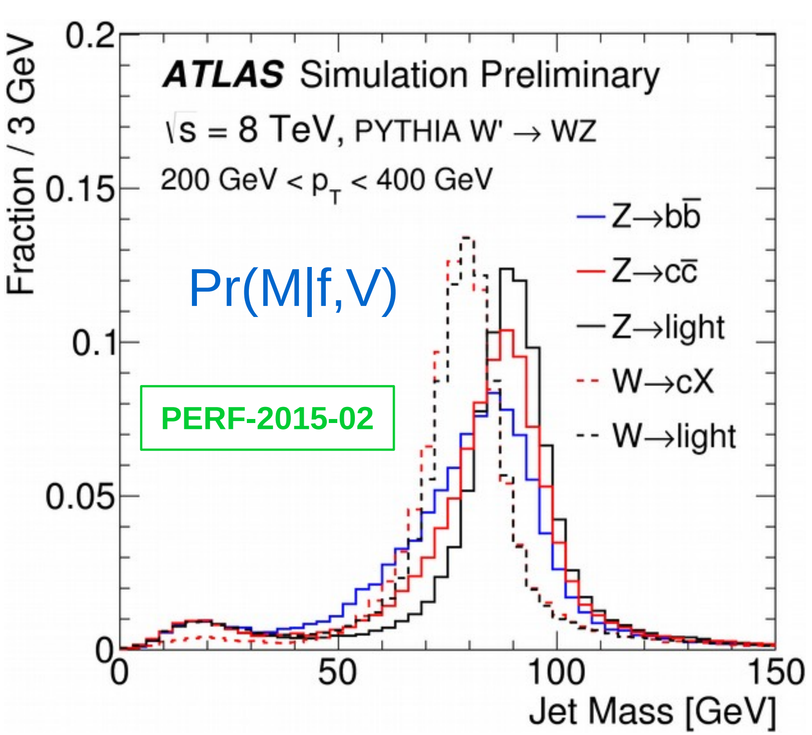
Likelihood tagger: templates build for each possible boson decay channel

$$\Pr(M,Q,B|V) = \sum_f \Pr(f|V)\Pr(M|f,V)\Pr(Q|f,V)\Pr(B|f,V)$$

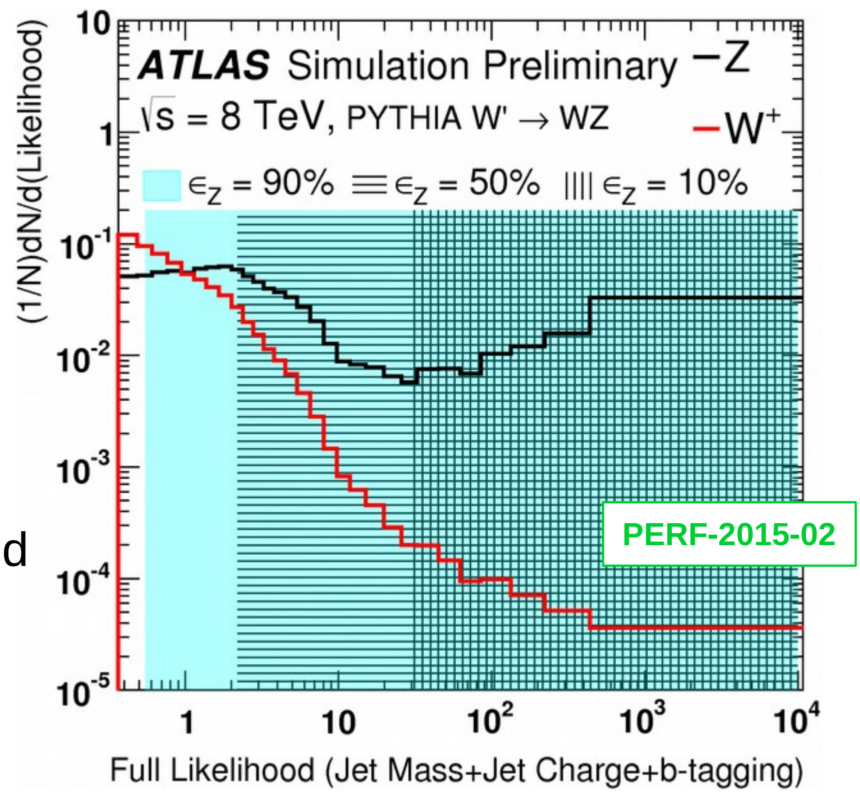
f: {bb,cc,cs,cd,light}
V: {W,Z}

with $p(B|f,V) = p(B_{\text{lead}}|f,V)p(B_{\text{sub-lead}}|f,V)$

This factorisation based on the boson decay channel is very good.
The tagger is then build from likelihood ratio.



1D histo build from the ratios signal/bkgd of each Pr(M,Q,B|V)



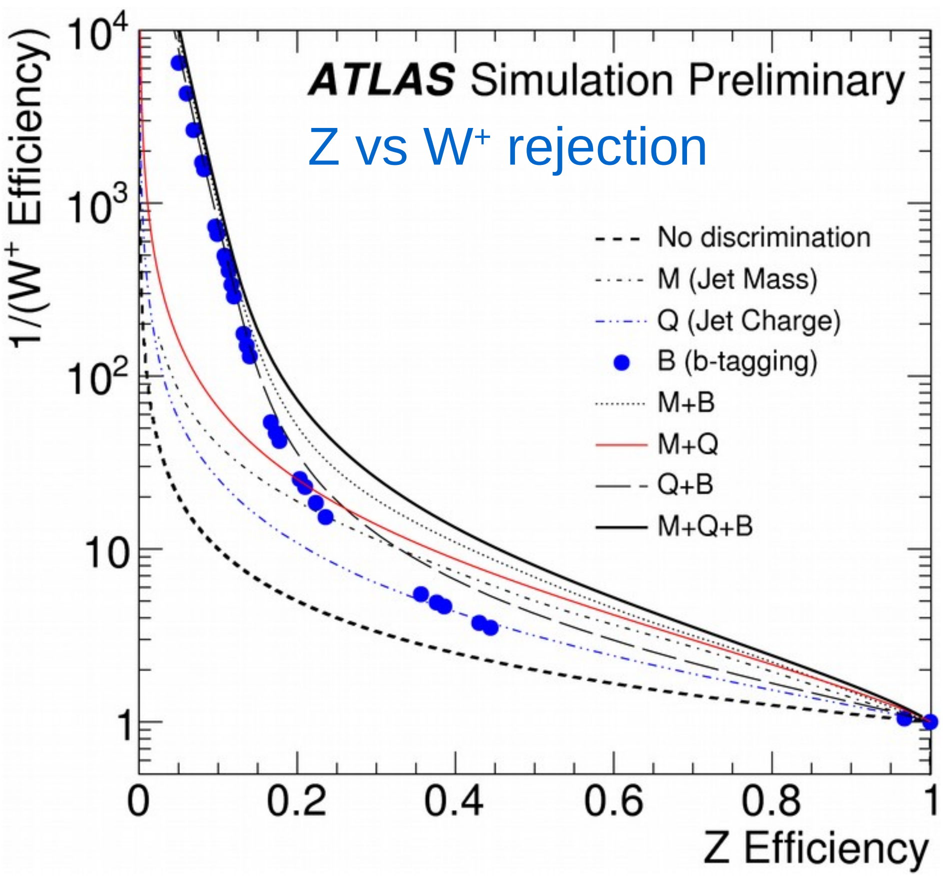
Tagger efficiencies

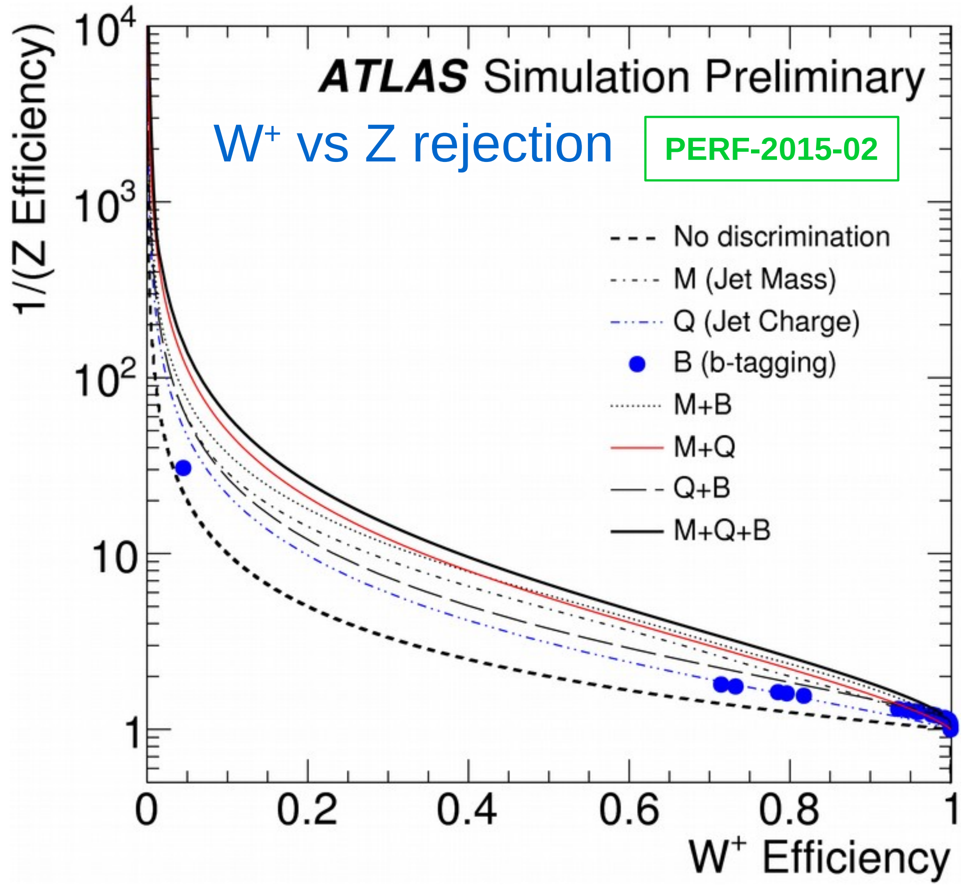
Efficiencies for   Z vs W⁺ rejection   and for W⁺ vs Z rejection

Example of application: (50% x σ(WZ)) / (12% x σ(WW)) = 50/12 x 20% = 83%

For Z-tagging, at low eff.: big gain from b-tagging because of Z → $b\bar{b}$

No syst. uncertainties (next slide)



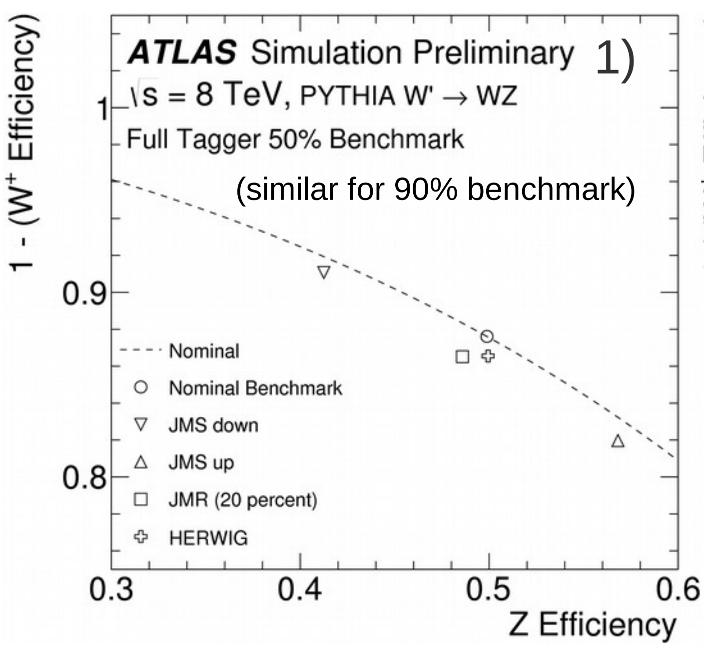


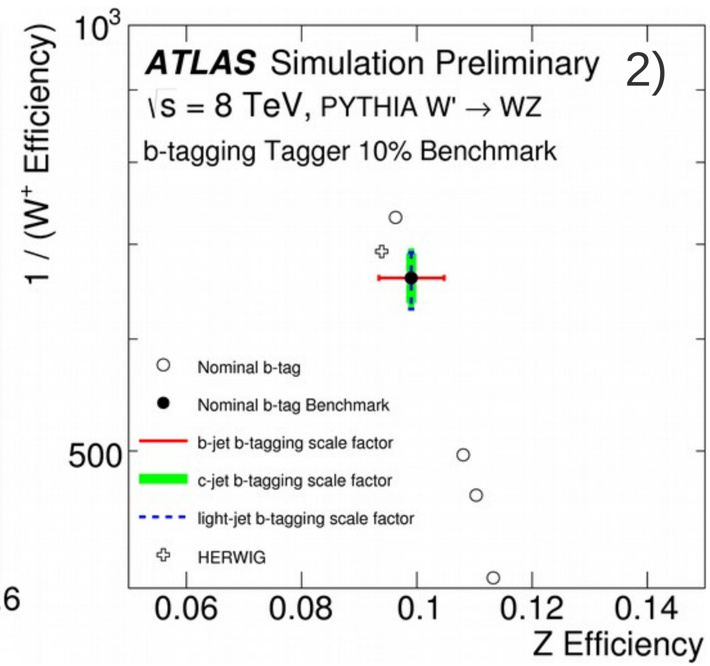
Systematic uncertainties

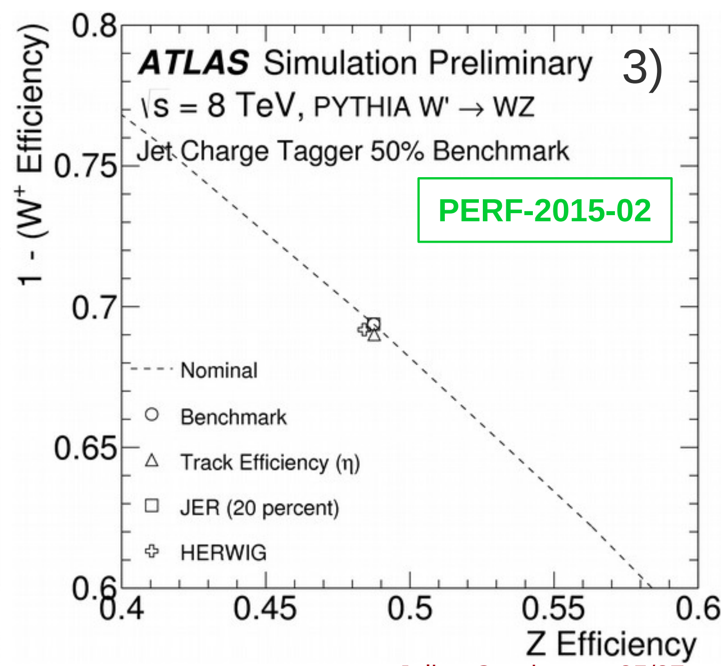
Sources of systematic uncertainties:

1) Calibration of large-R and small-R jets: JMS moves on the ROC curve
   JMR reduces performances
2) b-tagging uncertainties: from MC/Data discrepancies
   relevant at low Z-tagger eff.,
   b-tag affecting Z, c- and light-tag affecting W
3) modelling of track reconstruction: small impact in this p_T range

Dominant uncertainties: jet mass scale and resolution







Conclusion

Conclusion

Many grooming techniques and variables have been tested in Run1:

- grooming: Background rejection in mass window
  Pile-Up stability
  → **anti-kt R=1.0 trimmed f_{cut}=5%, R_{sub}=0.2**
- variables: Best performers: **τ_{21}^{wta}, $C_2^{(\beta=1)}$, $D_2^{(\beta=1)}$**

W-tagger in Run1, 2 working points (medium and tight)

Early W-tagger and Z-tagger for Run2:

Based on Run1 study, subset tested

→ **anti-kt R=1.0 trimmed f_{cut}=5%, R_{sub}=0.2 + $D_2^{(\beta=1)}$**

Uncertainties from Run1 + Run2 simulation variations

W-tagger and Z-tagger in Run2, 2 working points (medium and tight)

W and Z separation in Run1:

Based on 3 discriminating variables: mass, charge and b-tagging

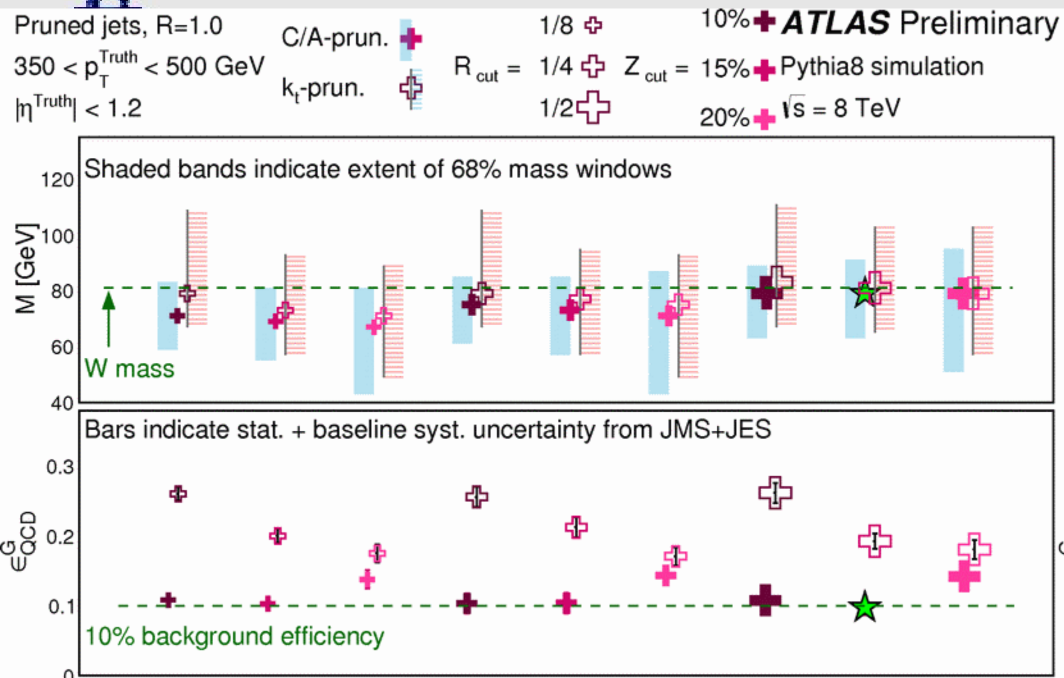
Z selection tagger, W selection tagger provided

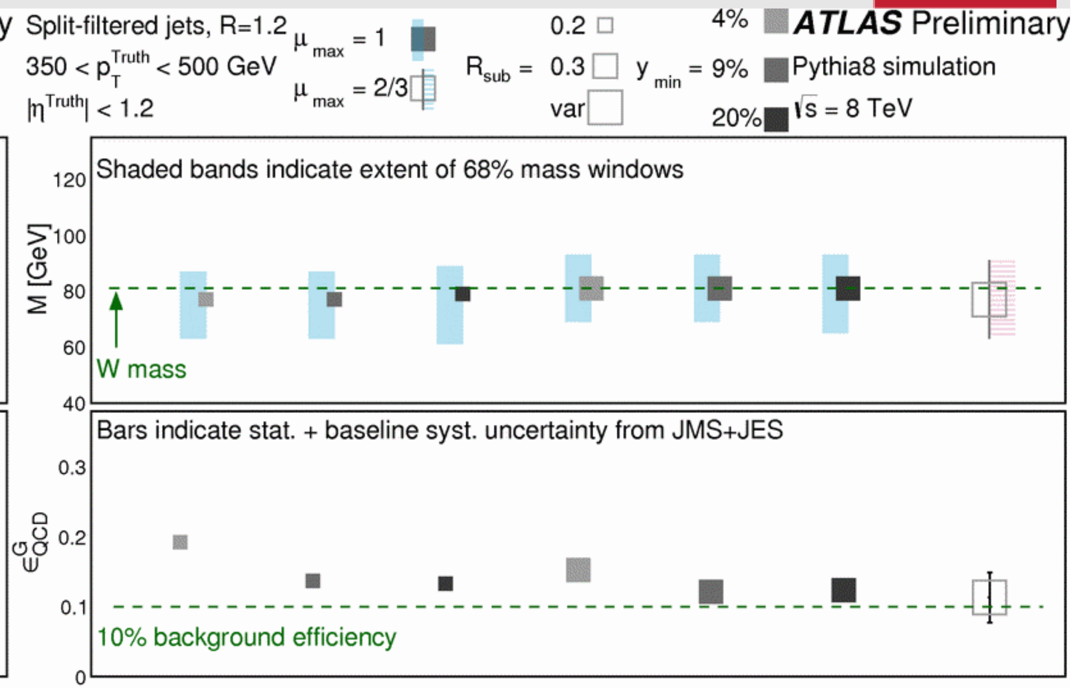
Calibration of the tagger in the data in Run2

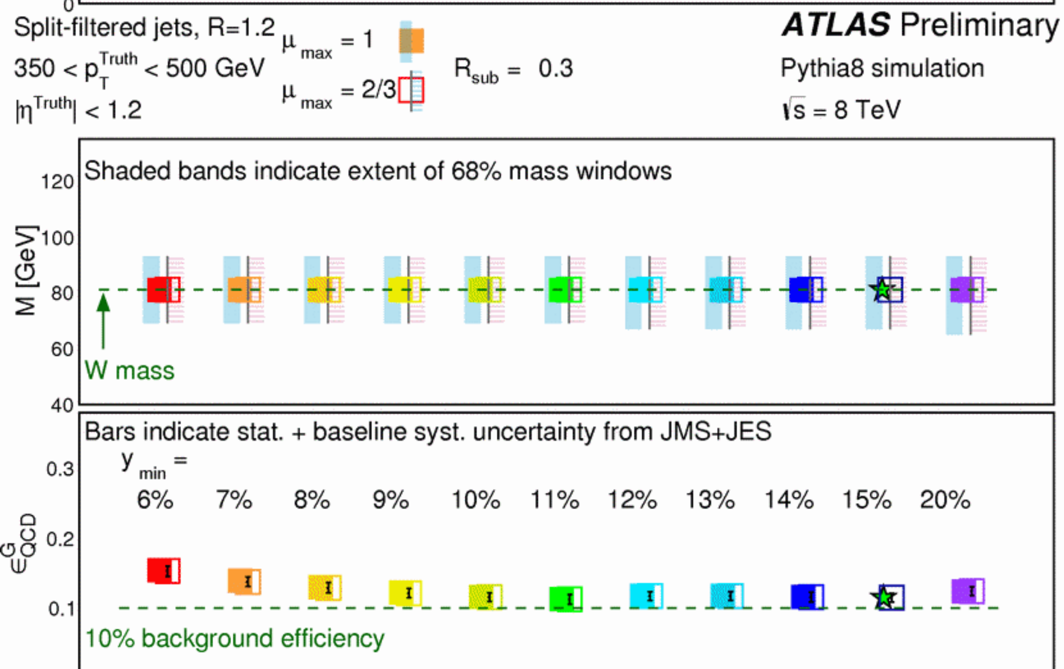
Backup slides

W-tagging in Run1

PERF-2015-03



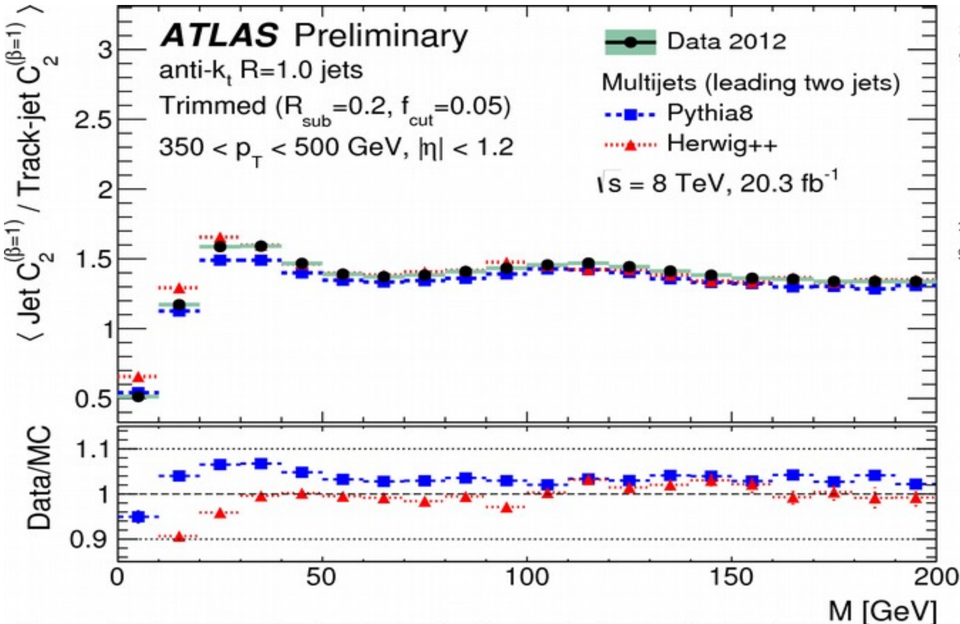


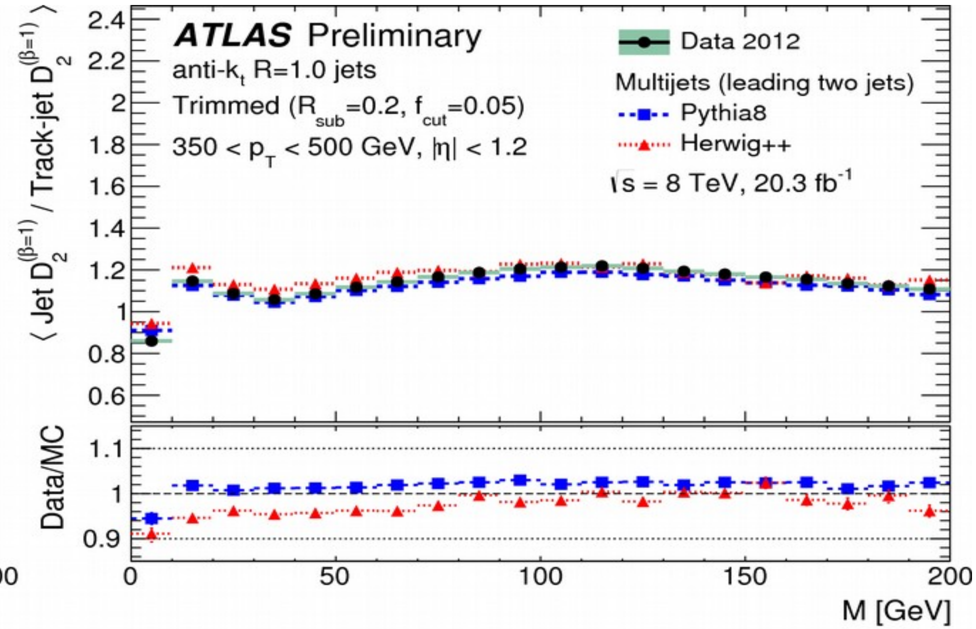


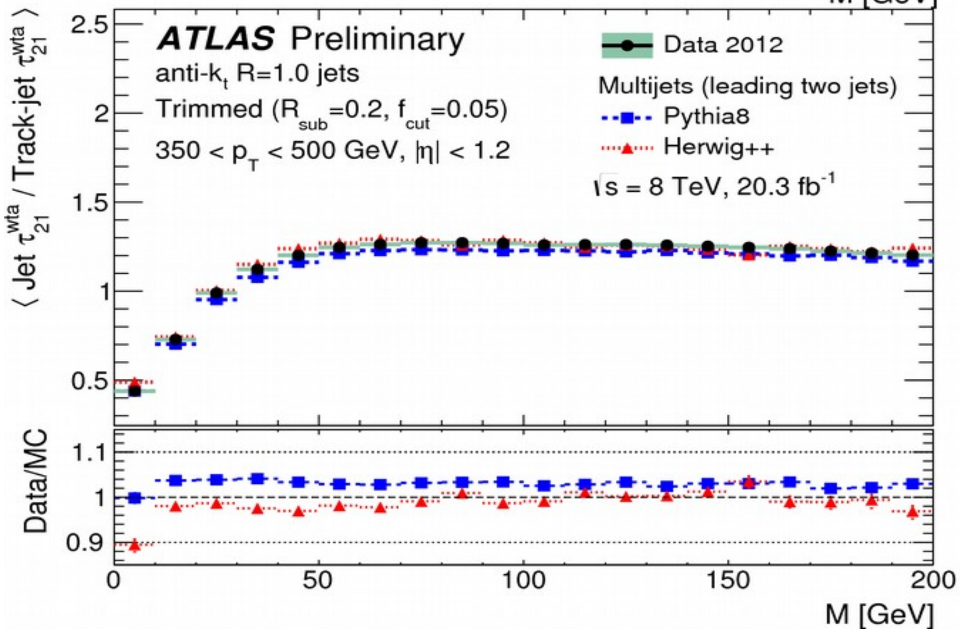
Performance of different grooming techniques

W-tagging in Run1

PERF-2015-03

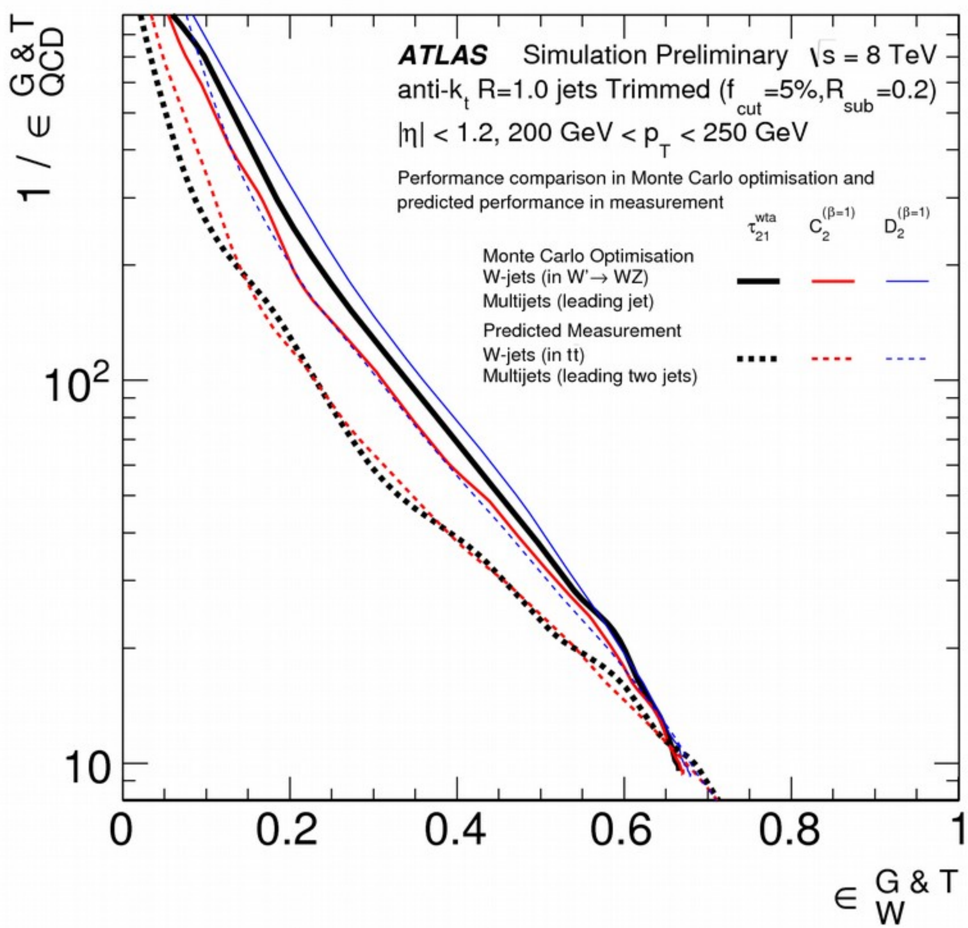


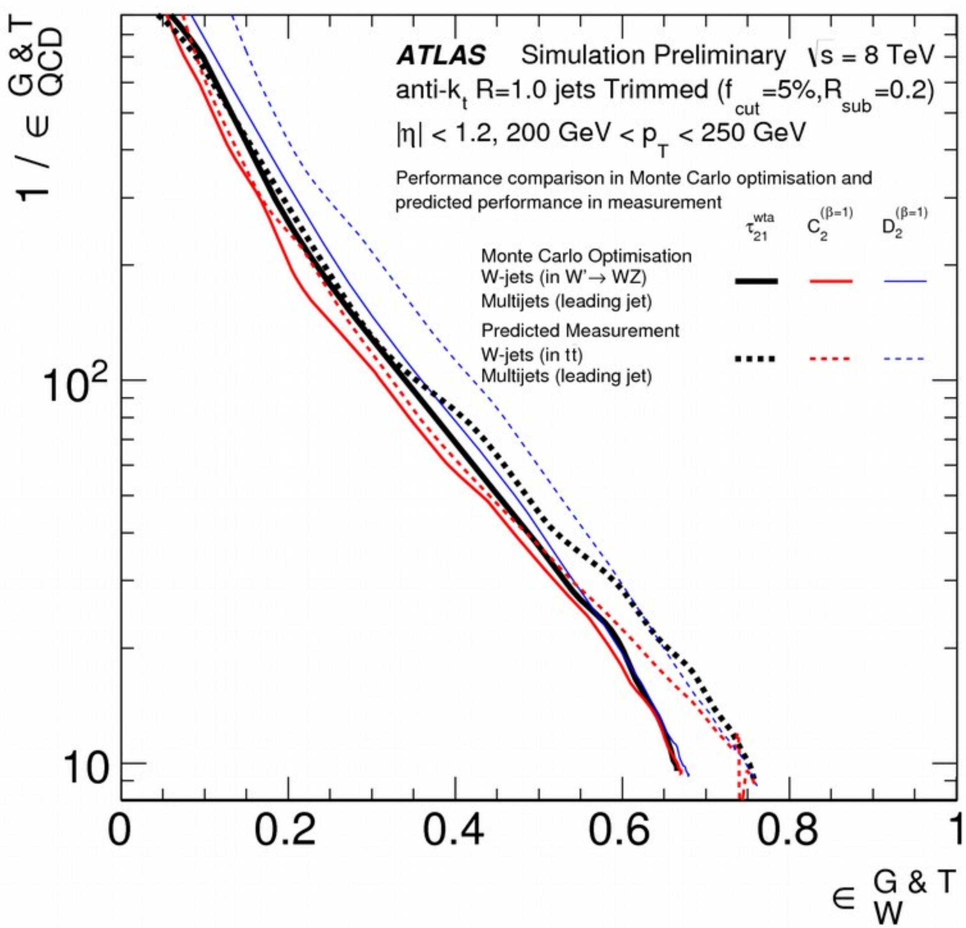




Track-jet double-ratio for C_2, D_2, tau_{21}^{wta}

Differences between W-jets in W' simulations and W-jets in $t\bar{t}$ predicted measurement

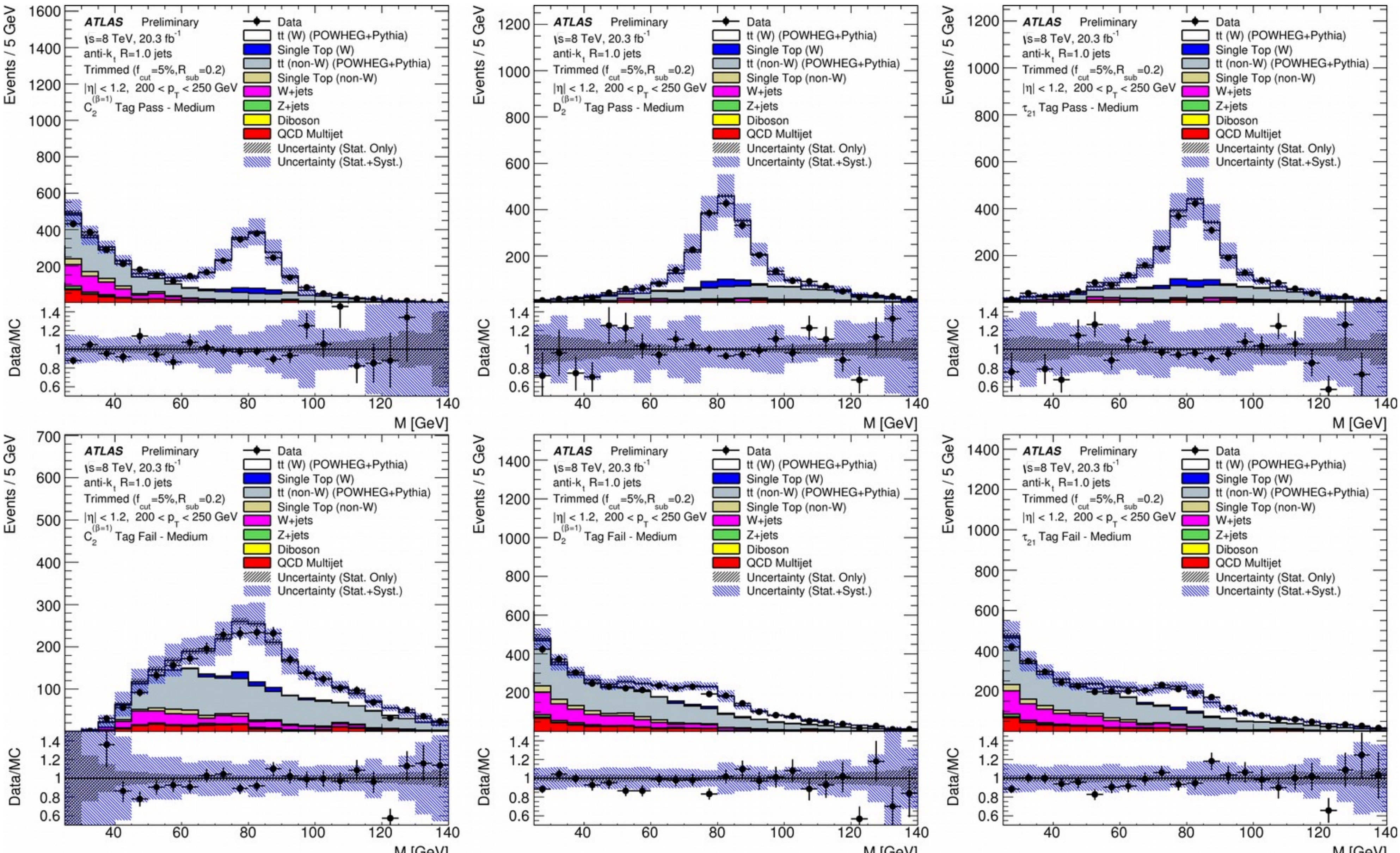




W-tagging in Run1

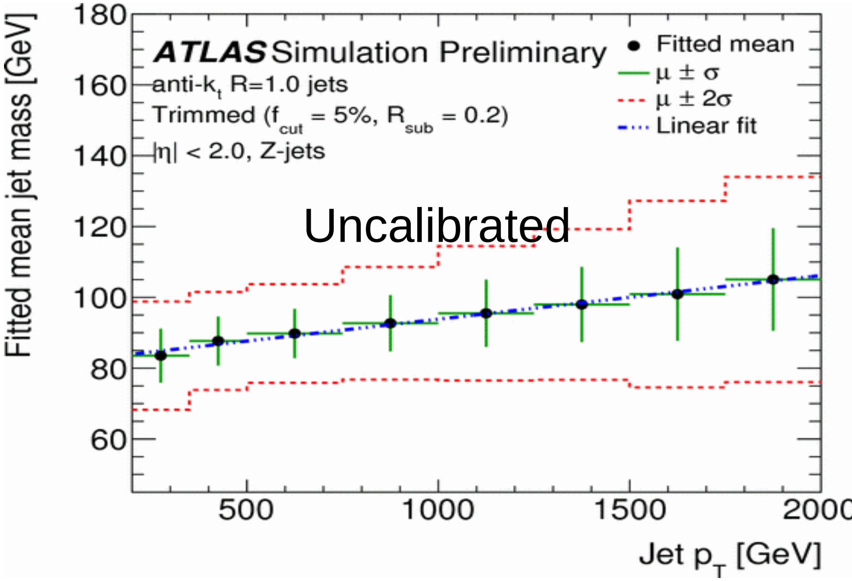
PERF-2015-03

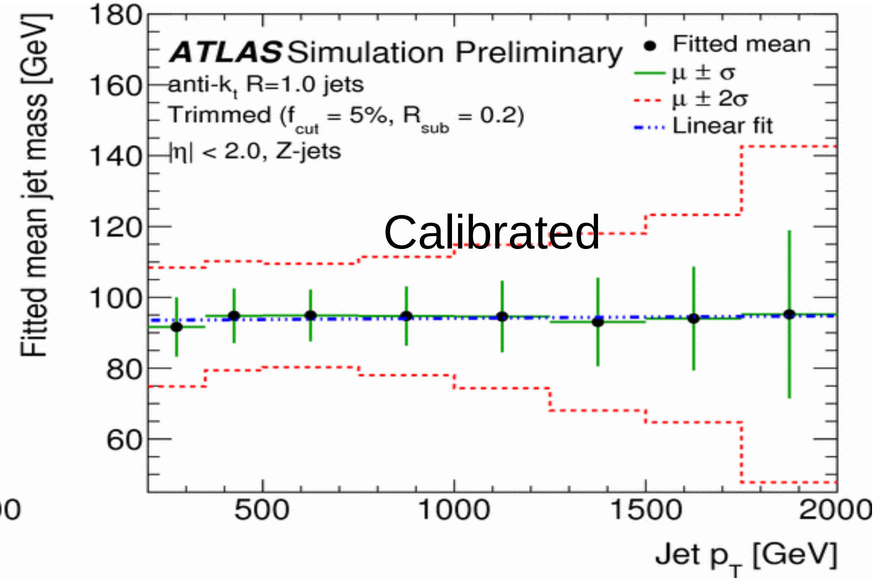
Mass distribution for passing (top) and failing (bottom) medium WP



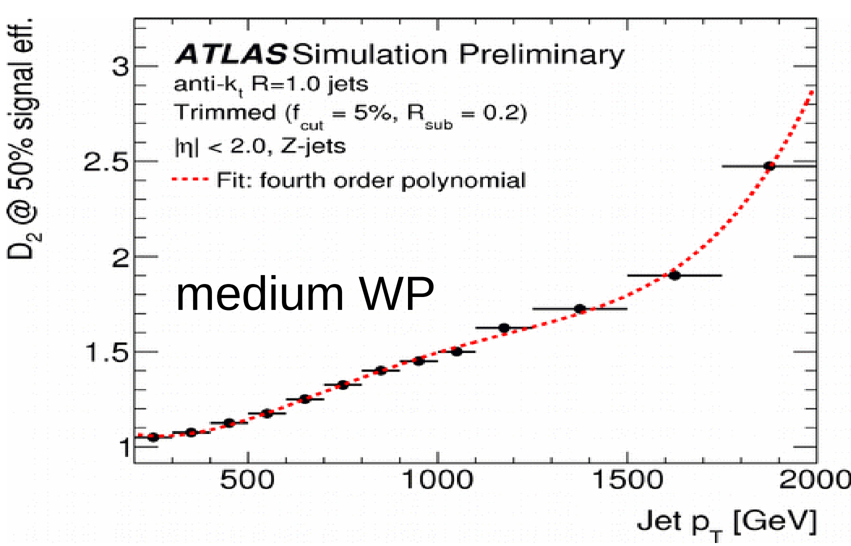
W-tagging in Run2

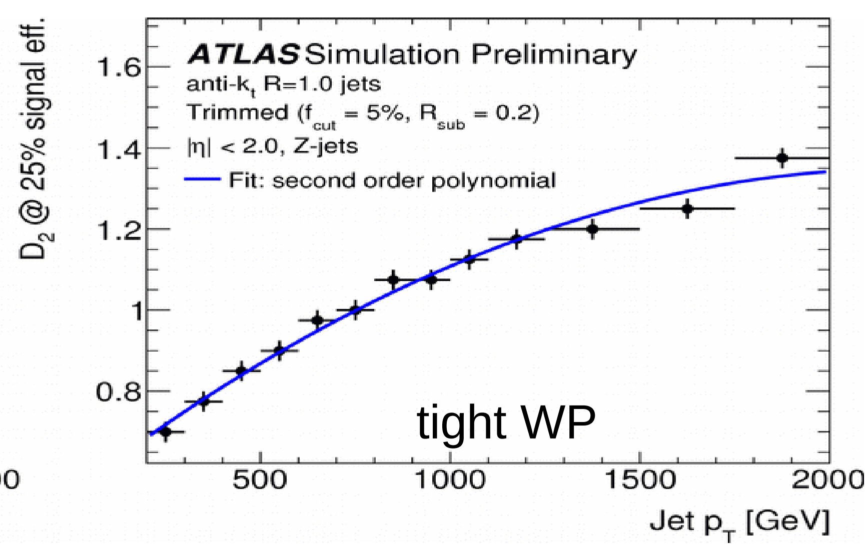
Mass calibration: (here for Z-boson)





Cut parametrization: (here for Z-boson)

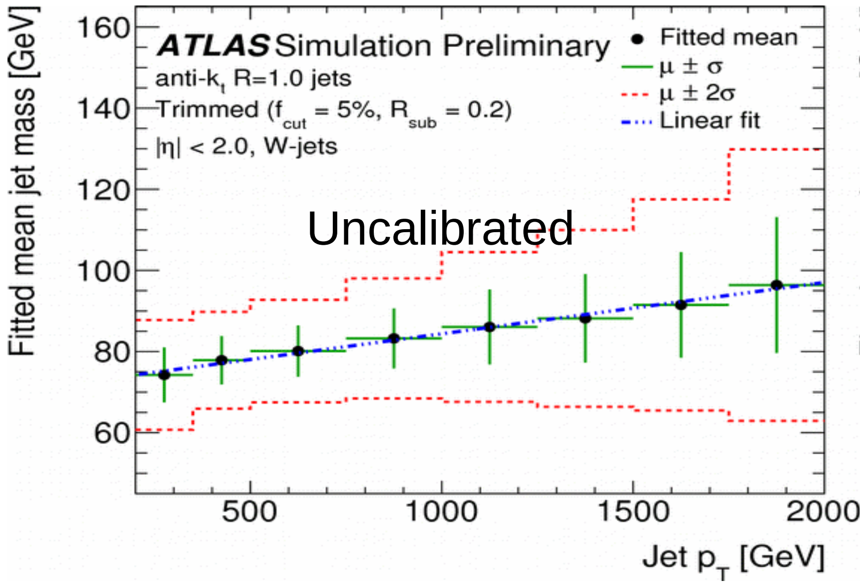


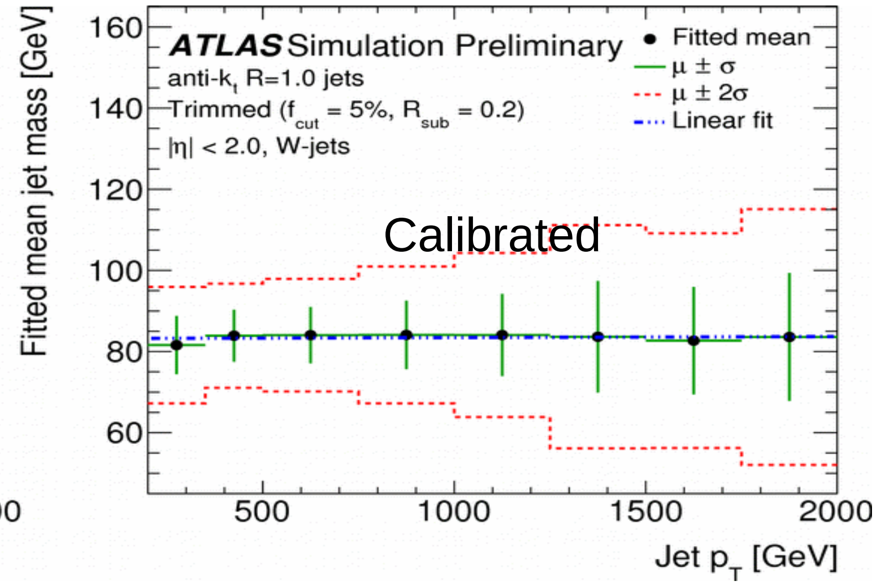


W-tagging in Run2

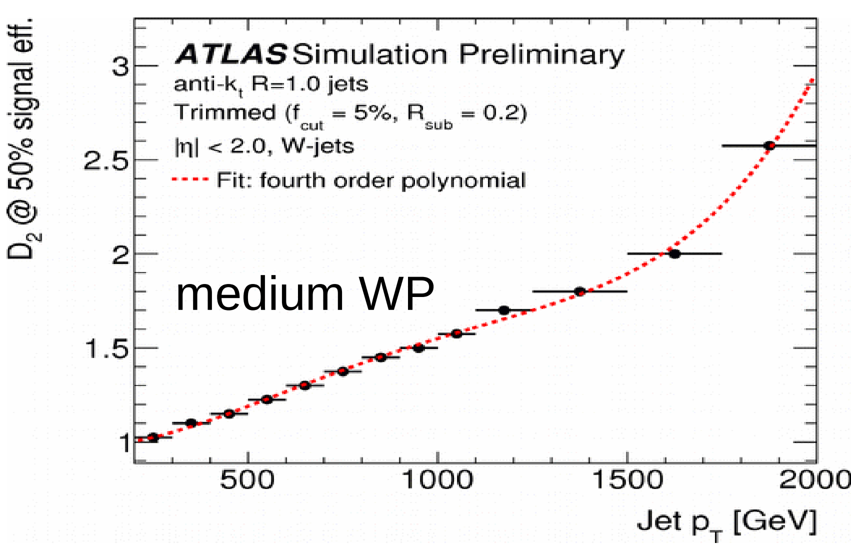
ATL-PHYS-PUB-2015-033

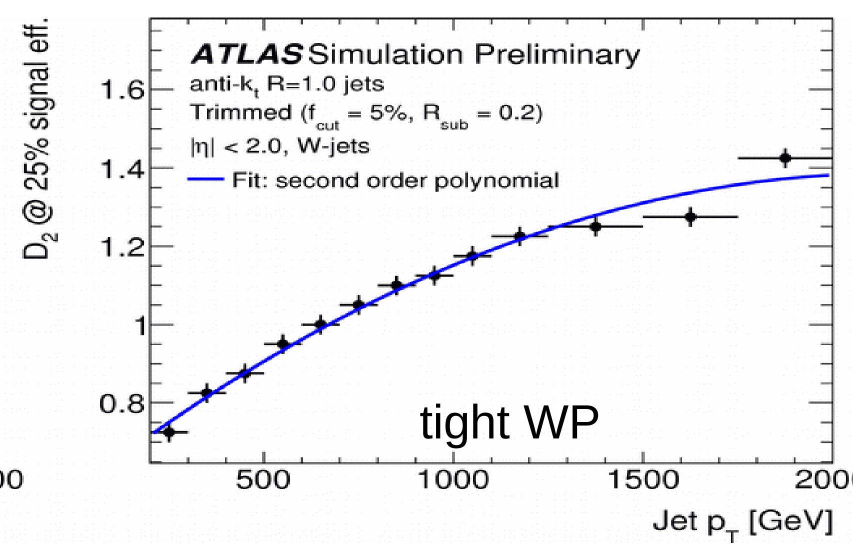
Mass calibration: (here for W-boson)





Cut parametrization: (here for W-boson)





W-tagging in Run2

ATL-PHYS-PUB-2015-033