Boson Tagging in ATLAS

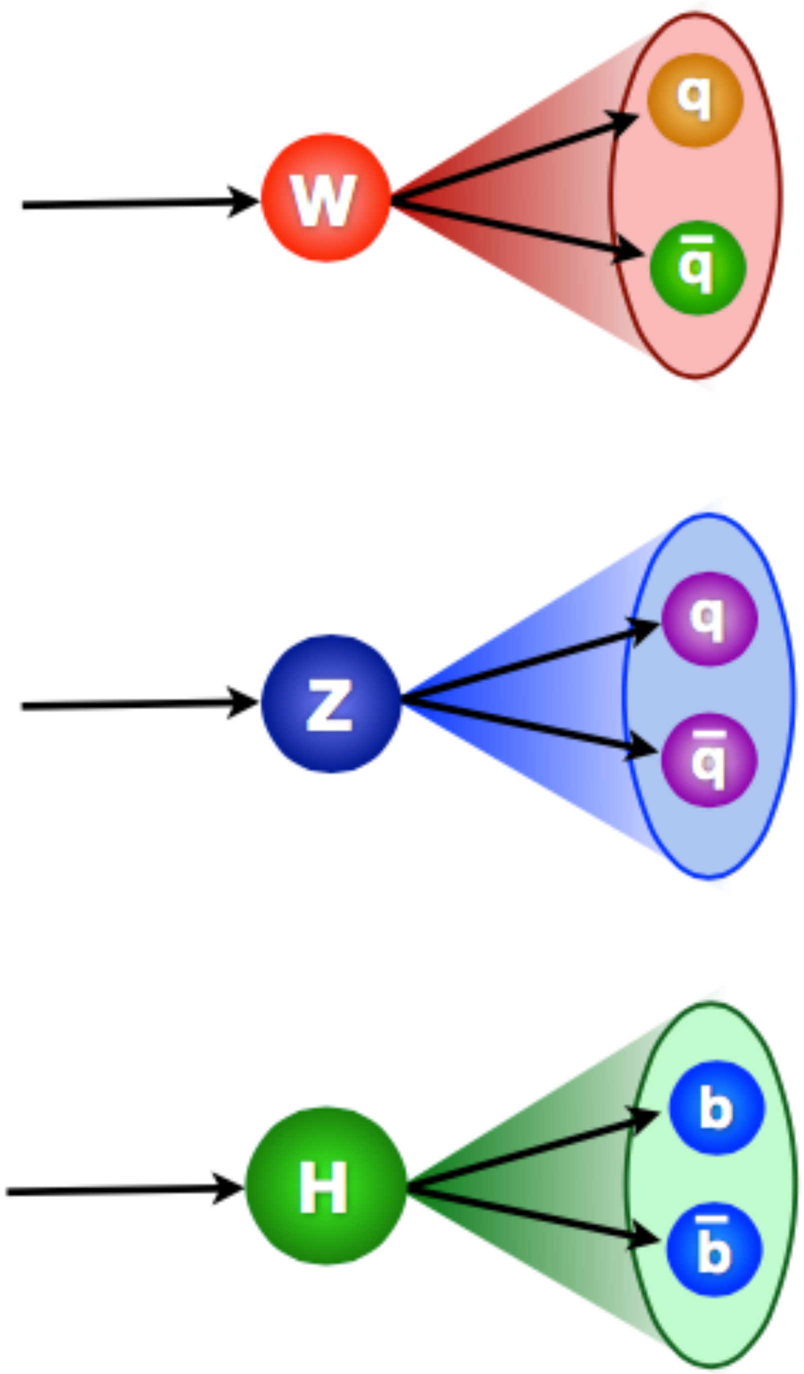
Shih-Chieh Hsu

Oct 22 2015

HEP Seminar in Stockholm University

Introduction

- Boosted bosons (W/Z/H) identifications are techniques to discriminate highly boosted hadronically decaying bosons from QCD jets
- The focus of this talk is the development in Run1 and prospects for Run2



- Reference
  - W-tagging arXiv:1510.05821
  - H-tagging ATL-PHYS-PUB-201 New 3

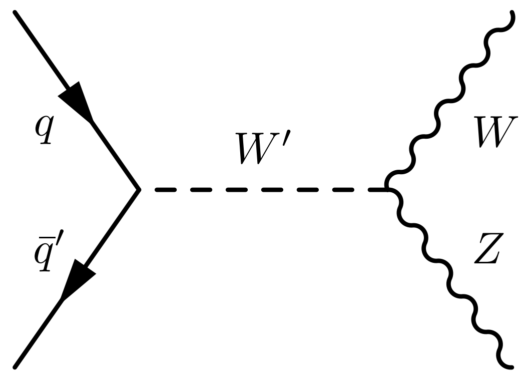
Motivation

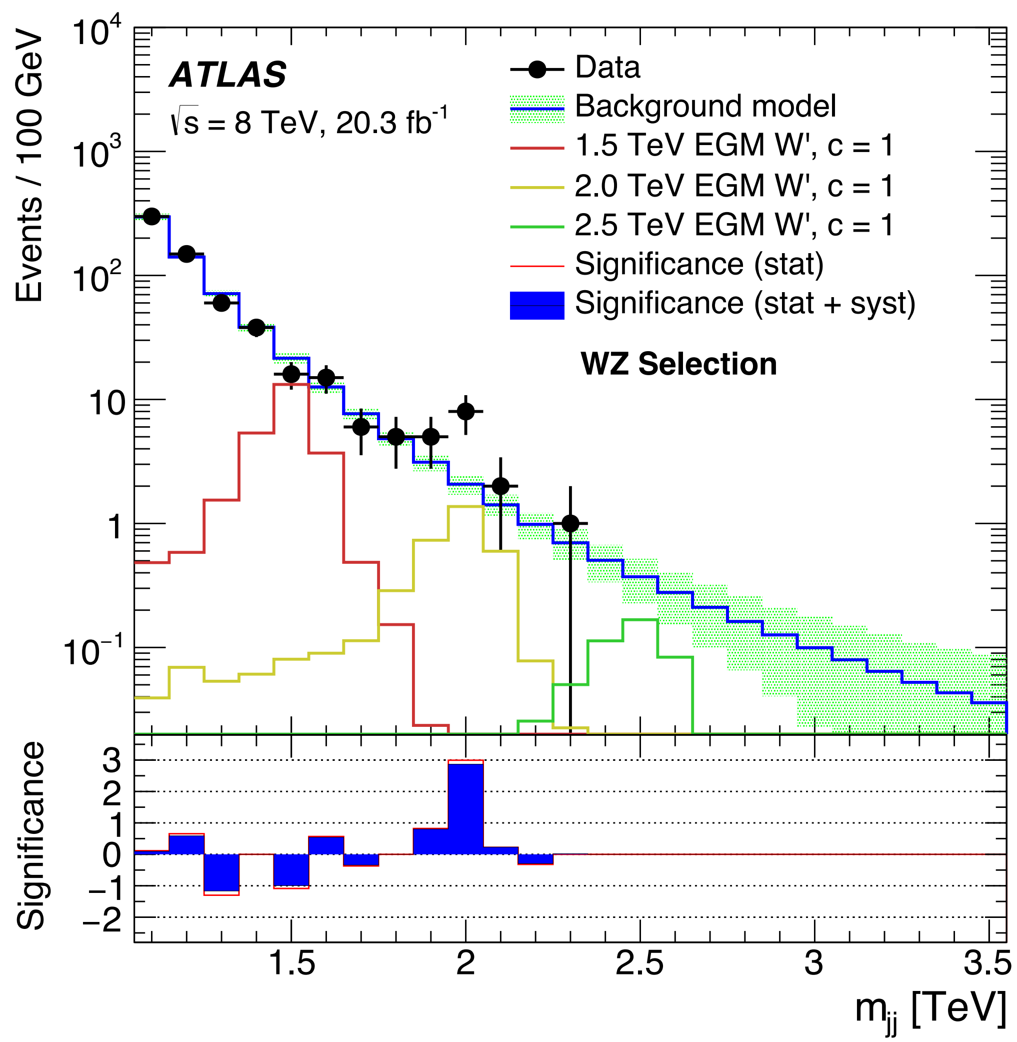
Boosted boson techniques have been used in several LHC run1 analyses in ATLAS

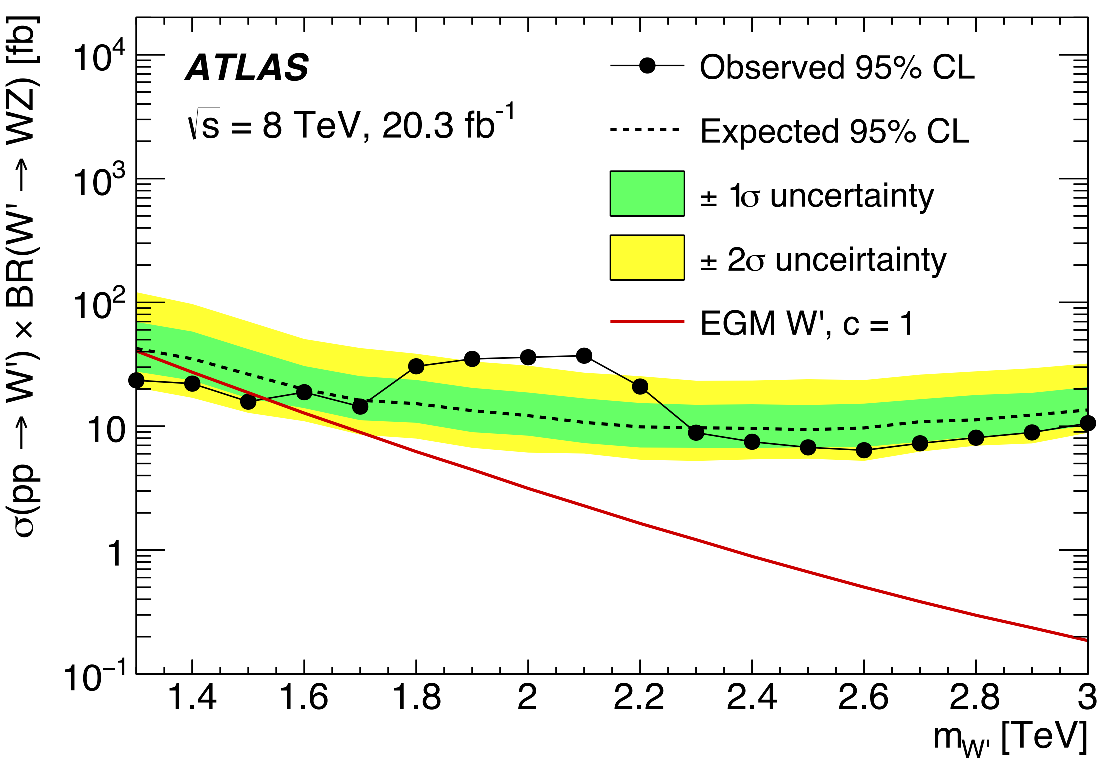
VV Excess in Run I

- Many Beyond the Standard Model theories predicts existence of new particles decaying into vector-boson pairs
- KK-Graviton, Extended gauge model, Technicolor
- Interesting 3.4 σ excess in full hadronic WZ search in Run1

VV→→lnuqq: EJP C75 (2015) 209
VV→→llqq: EPJ C75 (2015) 69
VV→→JJ: arXiv: 1506.00962



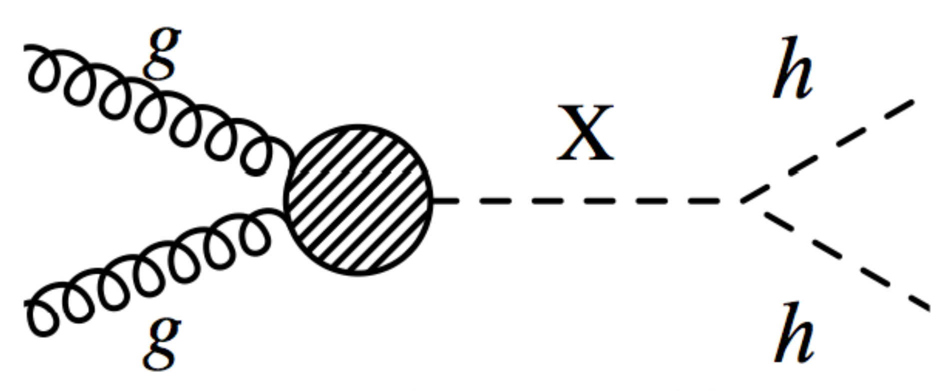


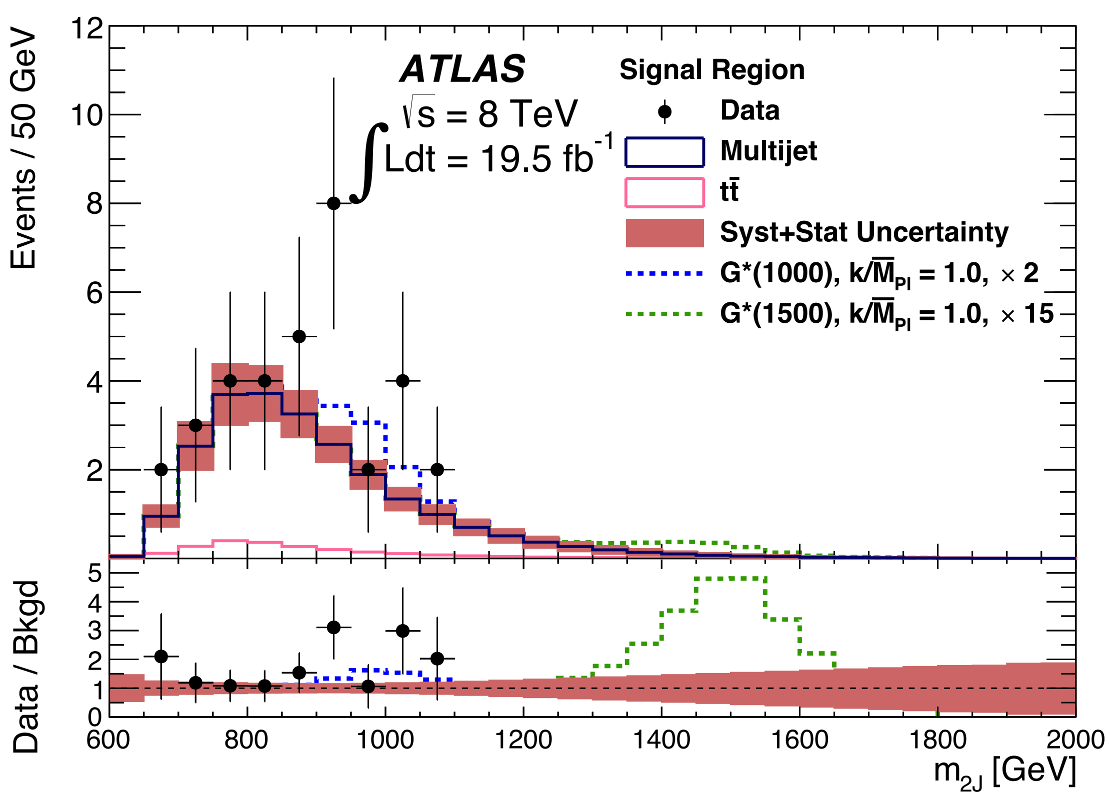


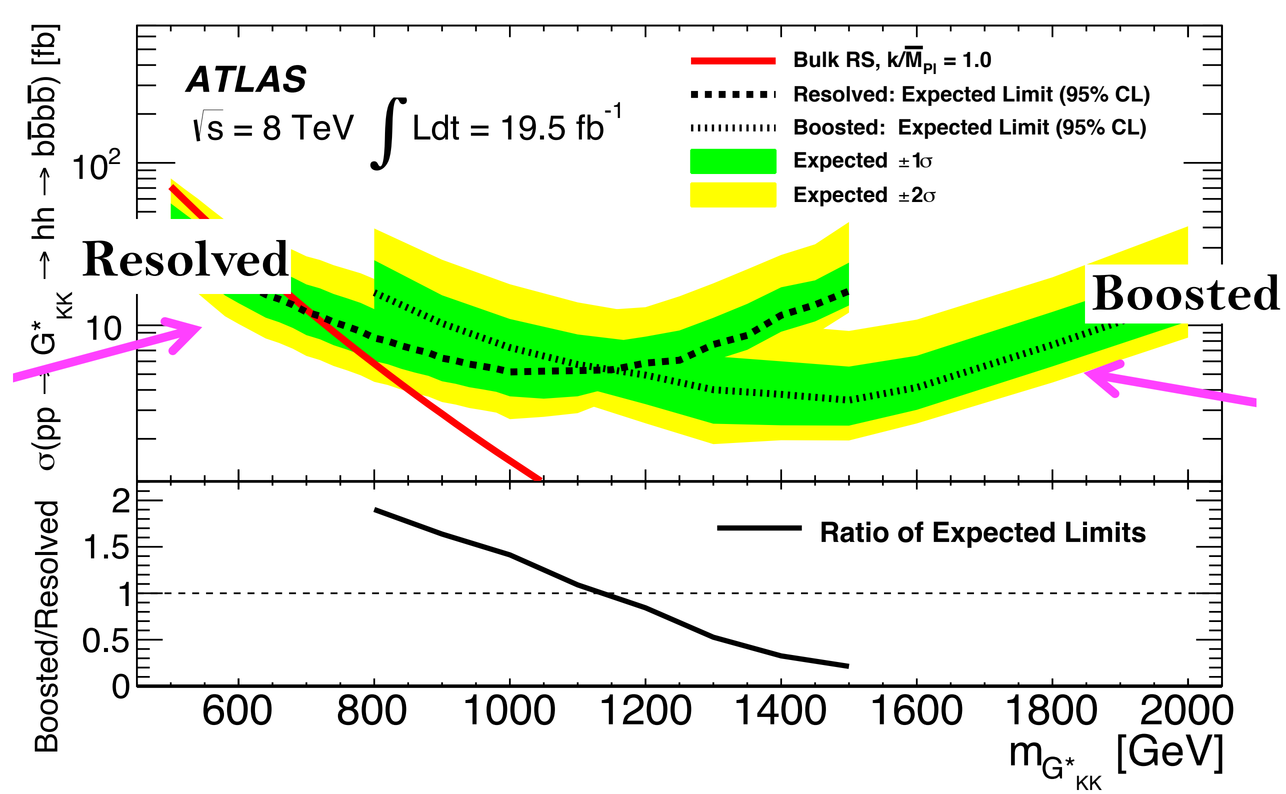
hh Resonance in Run I

- SM hh production
  - Direct test of Higgs potential
  - Small cross section: O(40 fb) at 14 TeV
- SBM Higgs could potentially enhance hh production
  - KK-Graviton, 2HDM, new scalar in Higgs portal…
- Boosted Higgs dominant the sensitivity of heavy resonance search

EPJ C75 (2015) 9, 412



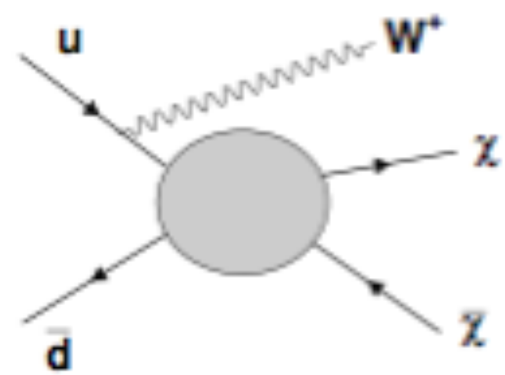


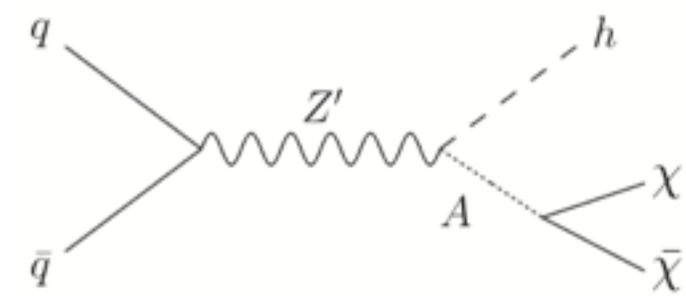


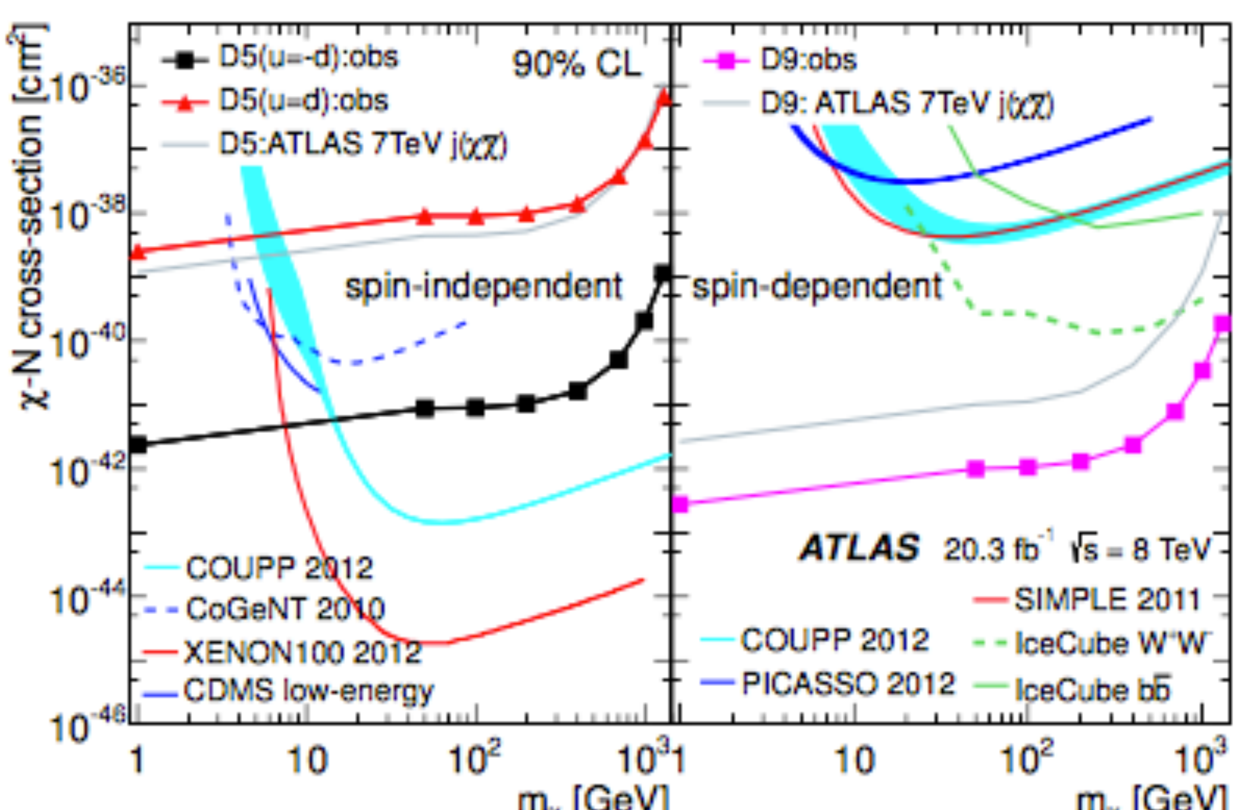
mono-W/Z/H in Run I

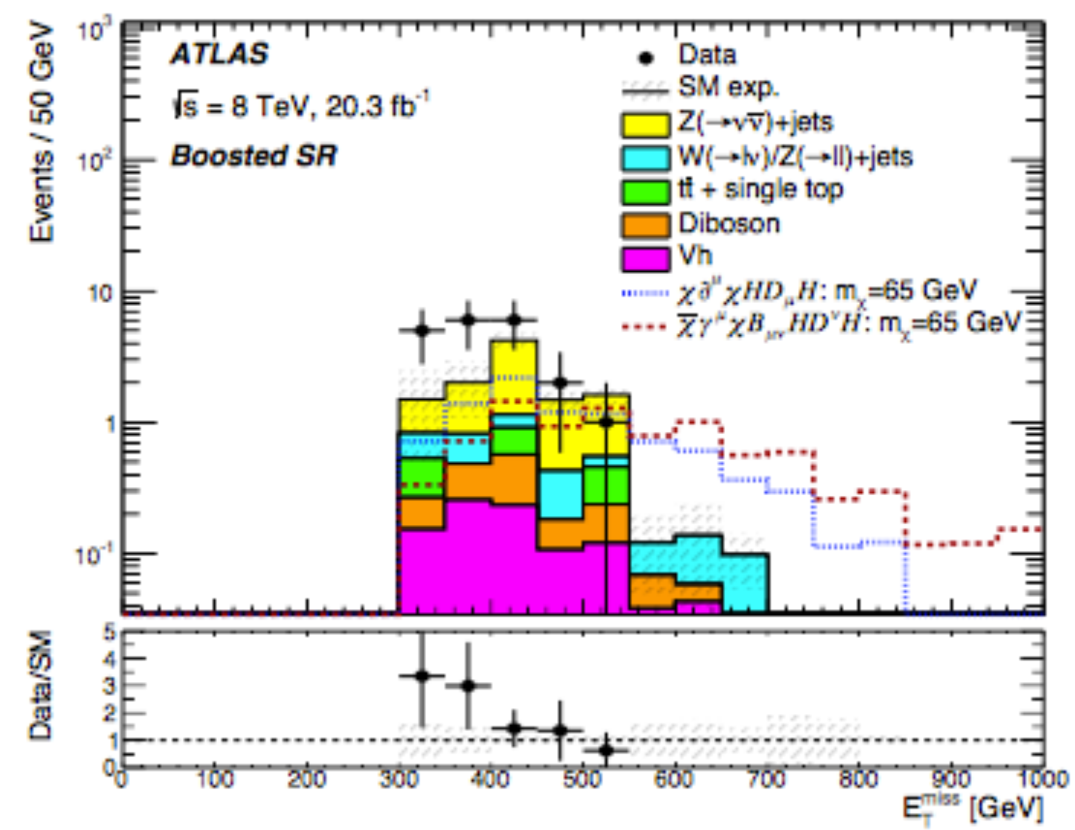
- mono Boson = Bosons + Met signature is predicted in many BSM models.
  - dark matter pair + Boson, VZ(vv) resonance, Z'-2HDM, …
- Unique and complementary to non-collider dark matter searches

mono-W/Z: [PRL 112, 041802 (2014)]
mono-H: [arXiv:1510.06218] New





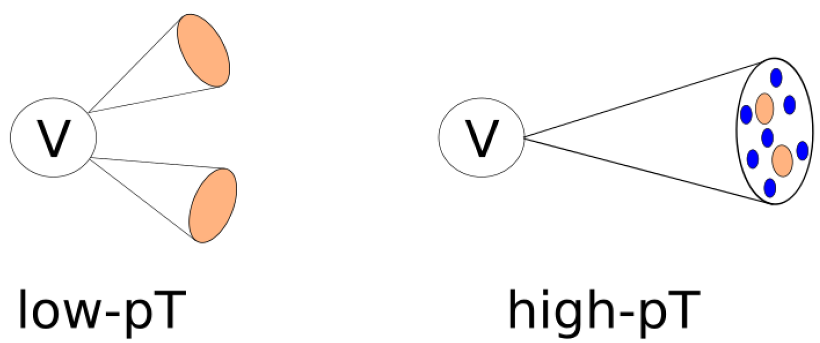




W tagging

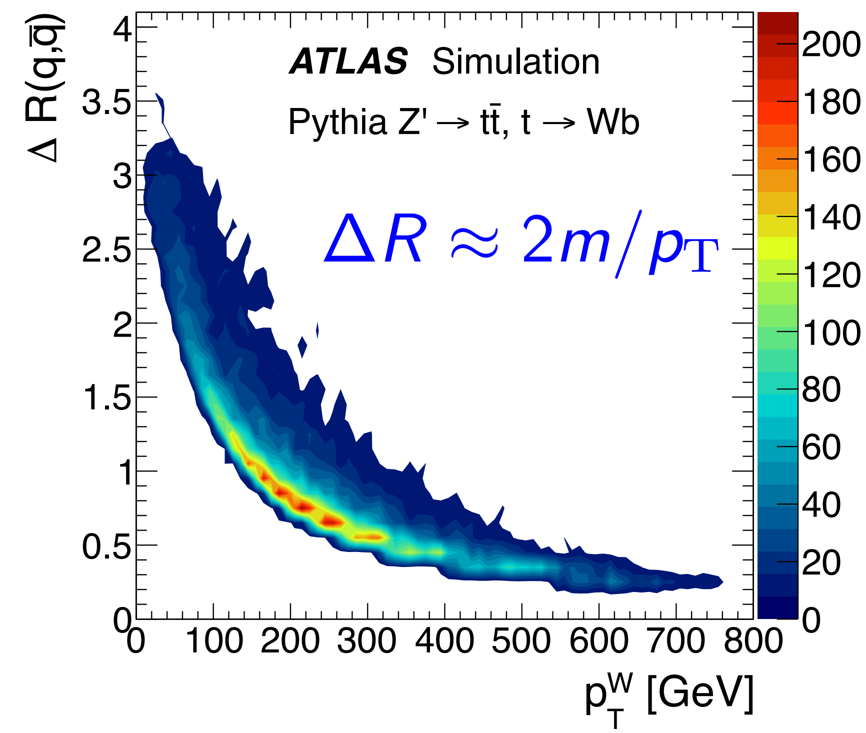
Basic Idea

- Reconstruction as one single large-radius jet

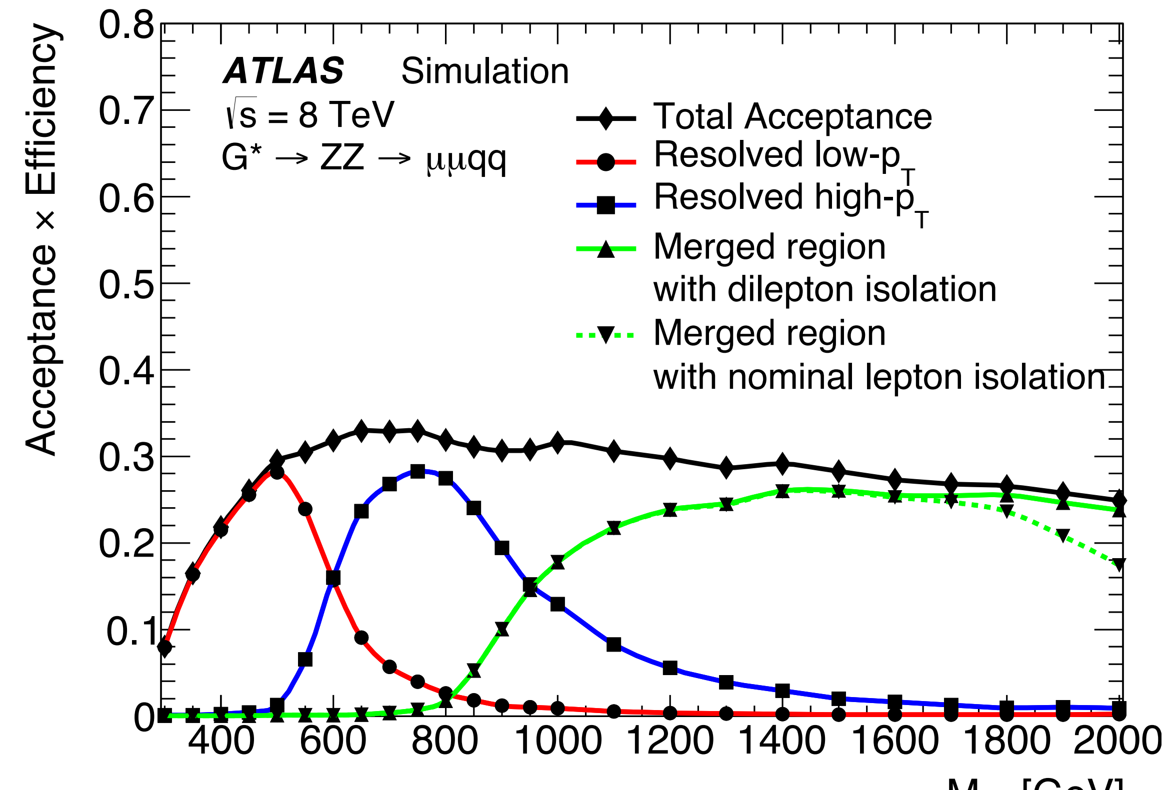


- Grooming techniques
  - Clean the large-R jet from soft gluon radiation and pile-up effects that diminish jet mass resolution
  - Techniques: BDRS* (mass-drop/filtering), trimming, pruning

- Substructure information (tagging)
  Use hard substructure of jet (not present in e.g. gluon jet) to improve signal efficiency and background rejection

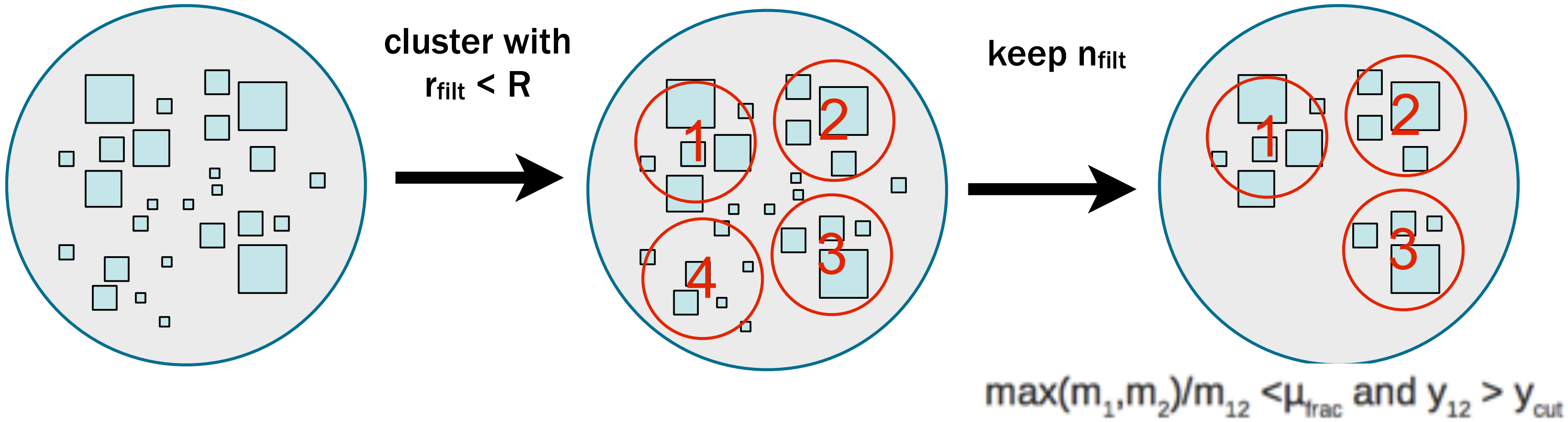


JHEP09 (2013) 076

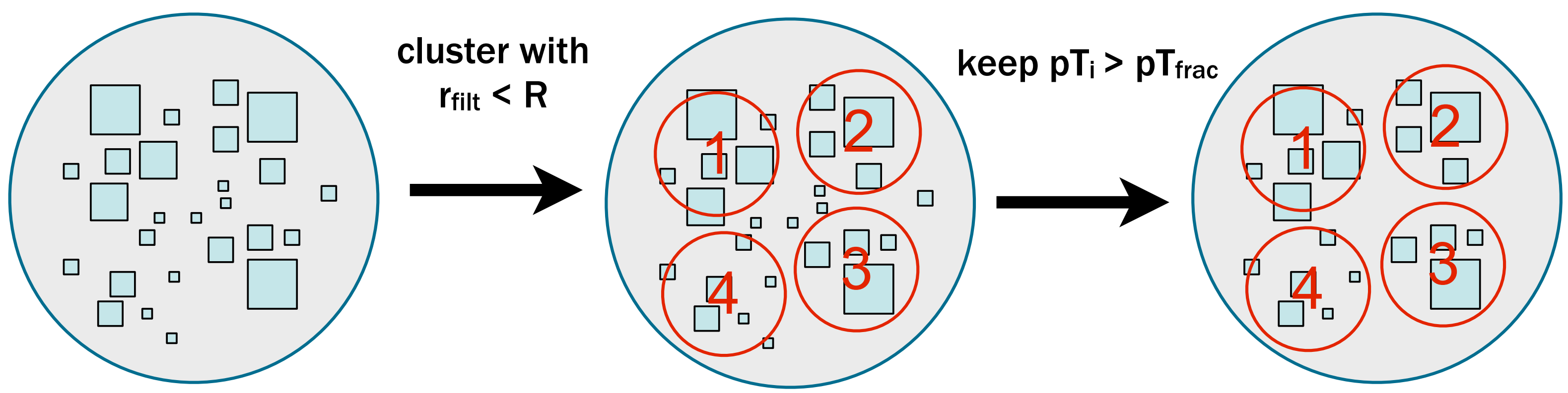


Jet Grooming

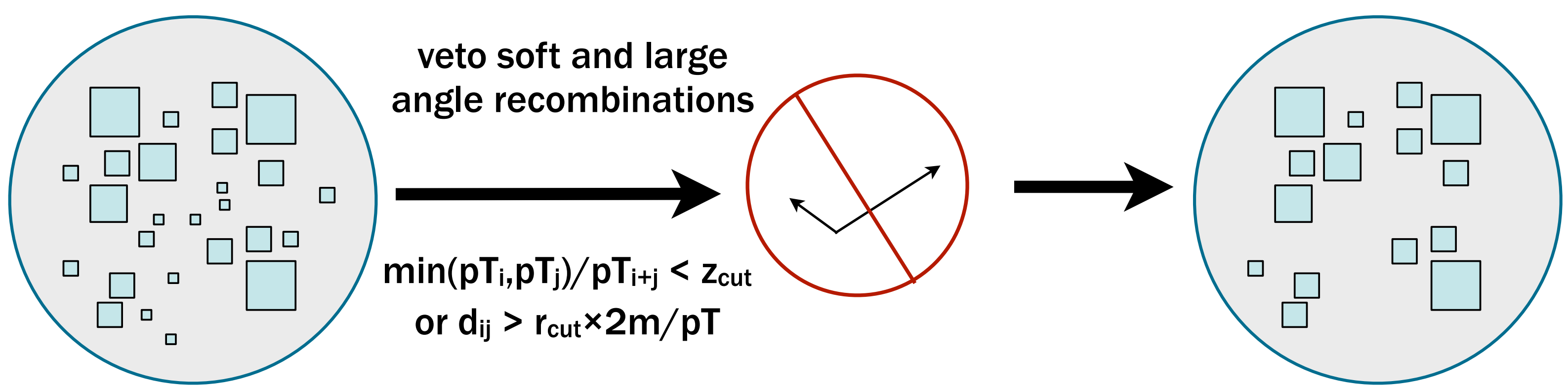
filtering



trimming



pruning



Jet Grooming optimization

Test three grooming techniques with >500 jet collections

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

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

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

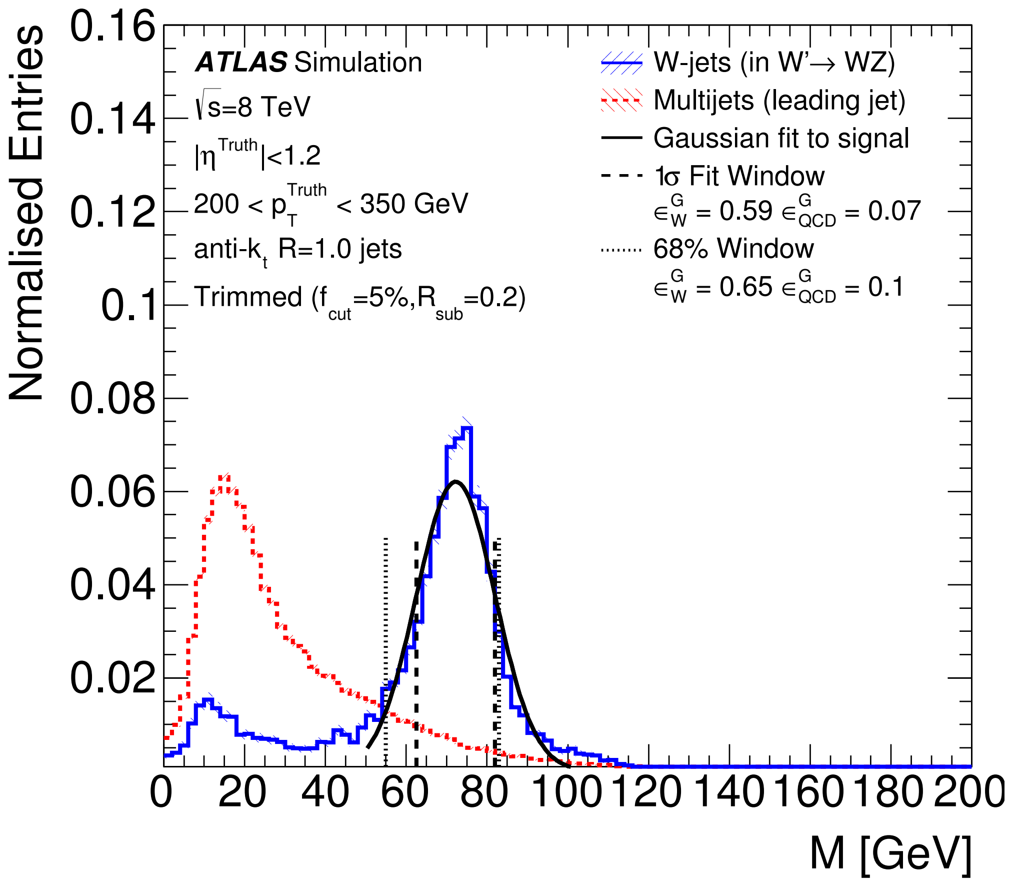
→ 2 x 3 x 1 x 5 x 6 = 180

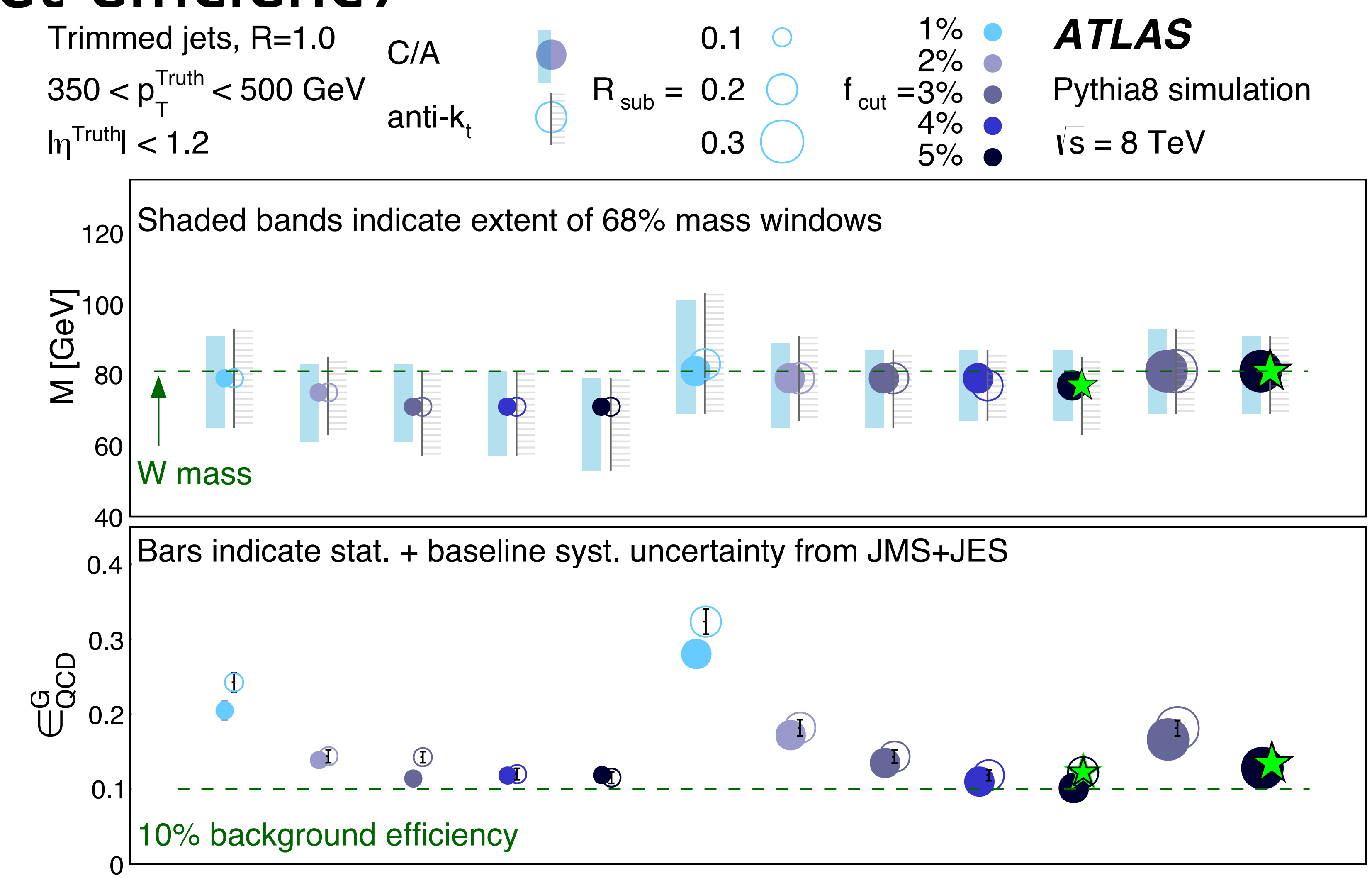
Mass optimization

Identify 68% mass window for 3 truth-jet pT (200,350,500~1000GeV)

Optimization figure of merit:

- mass peak is relative symmetric
- minimal QCD jet efficiency



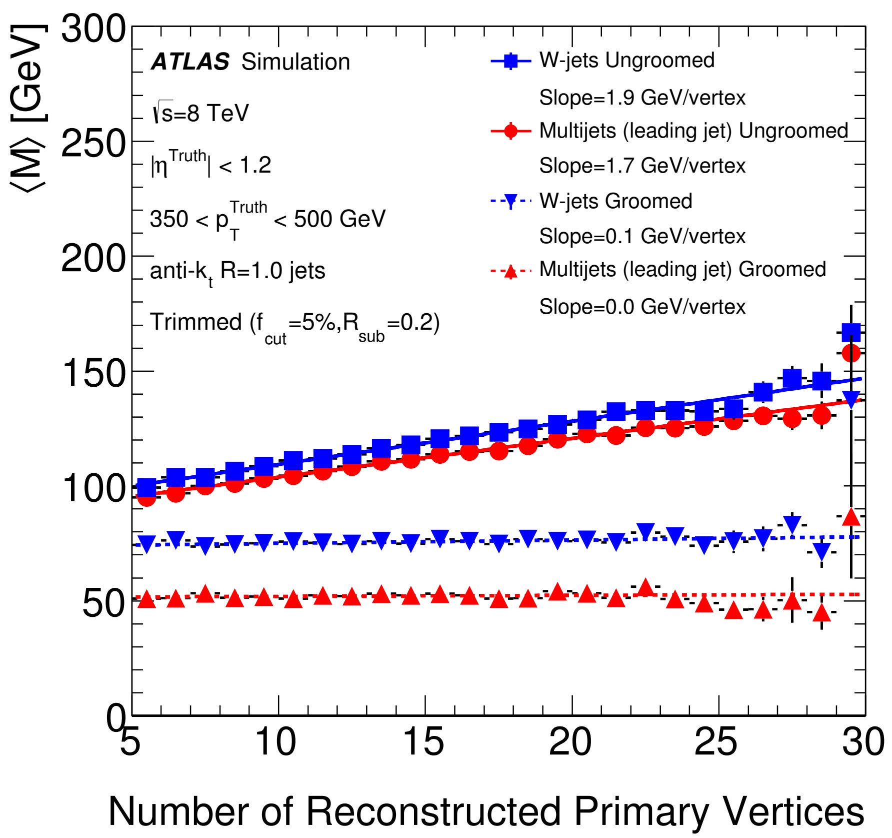


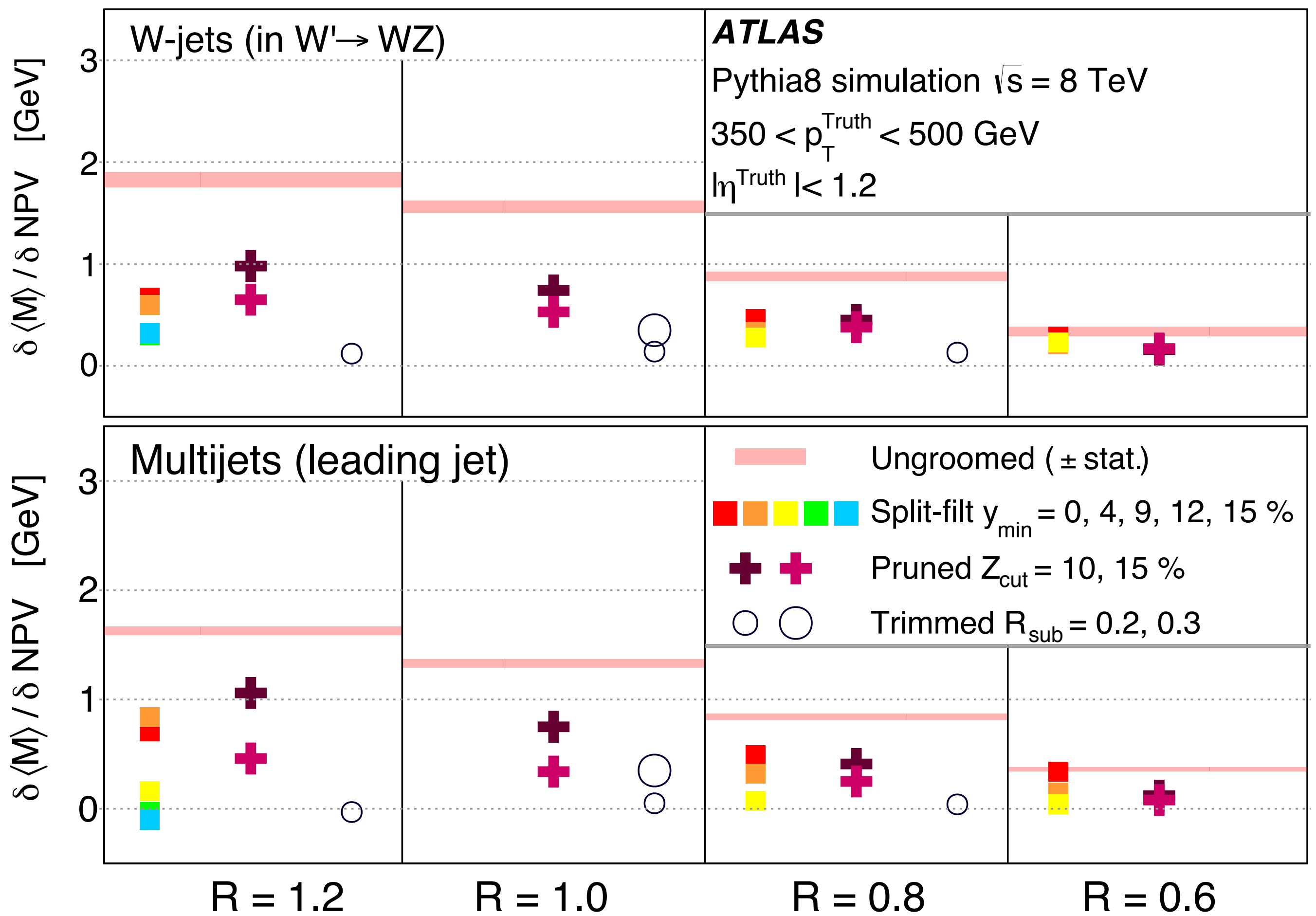
Identify best groomer in each

Optimal groomer in each algorithm can achieve equivalent bkg rejection.

best grooming: low bkgd eff. + good P-U stability

→ **anti-kt R=1.0 trimmed f_{cut}=5%, R_{sub}=0.2**





Jet substructure variables

Energy Correlation Function

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

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

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

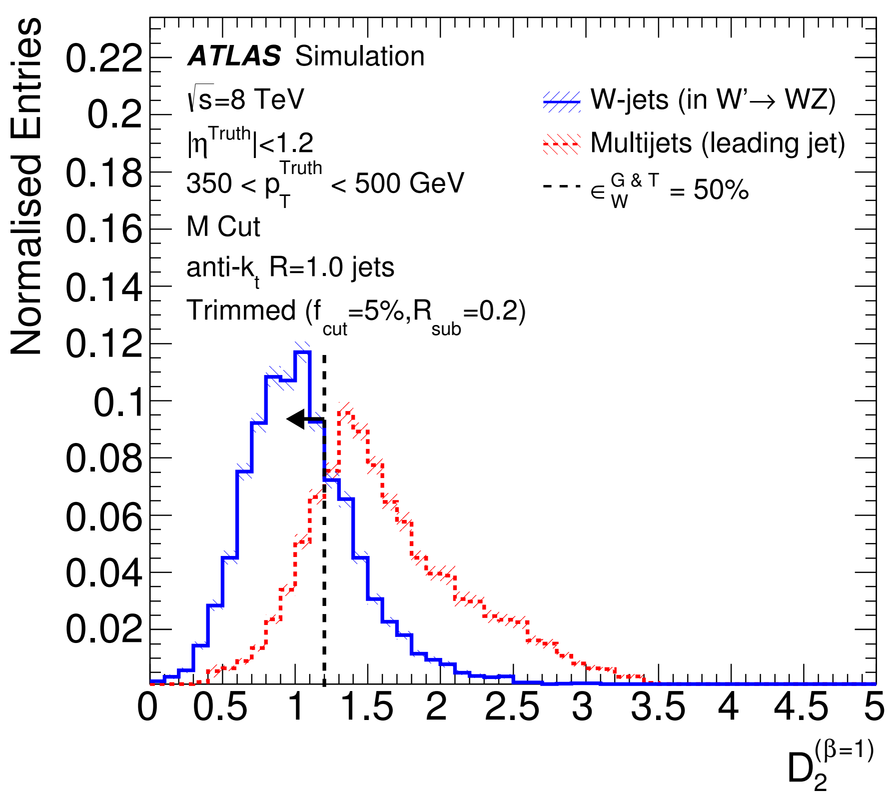
$$E_{CF}3(\beta) = \sum_{i<j<k \in J} p_{\mathrm{T}_i} p_{\mathrm{T}_j} p_{\mathrm{T}_k} \left(\Delta R_{ij} \Delta R_{ik} \Delta R_{jk}\right)^{\beta},$$

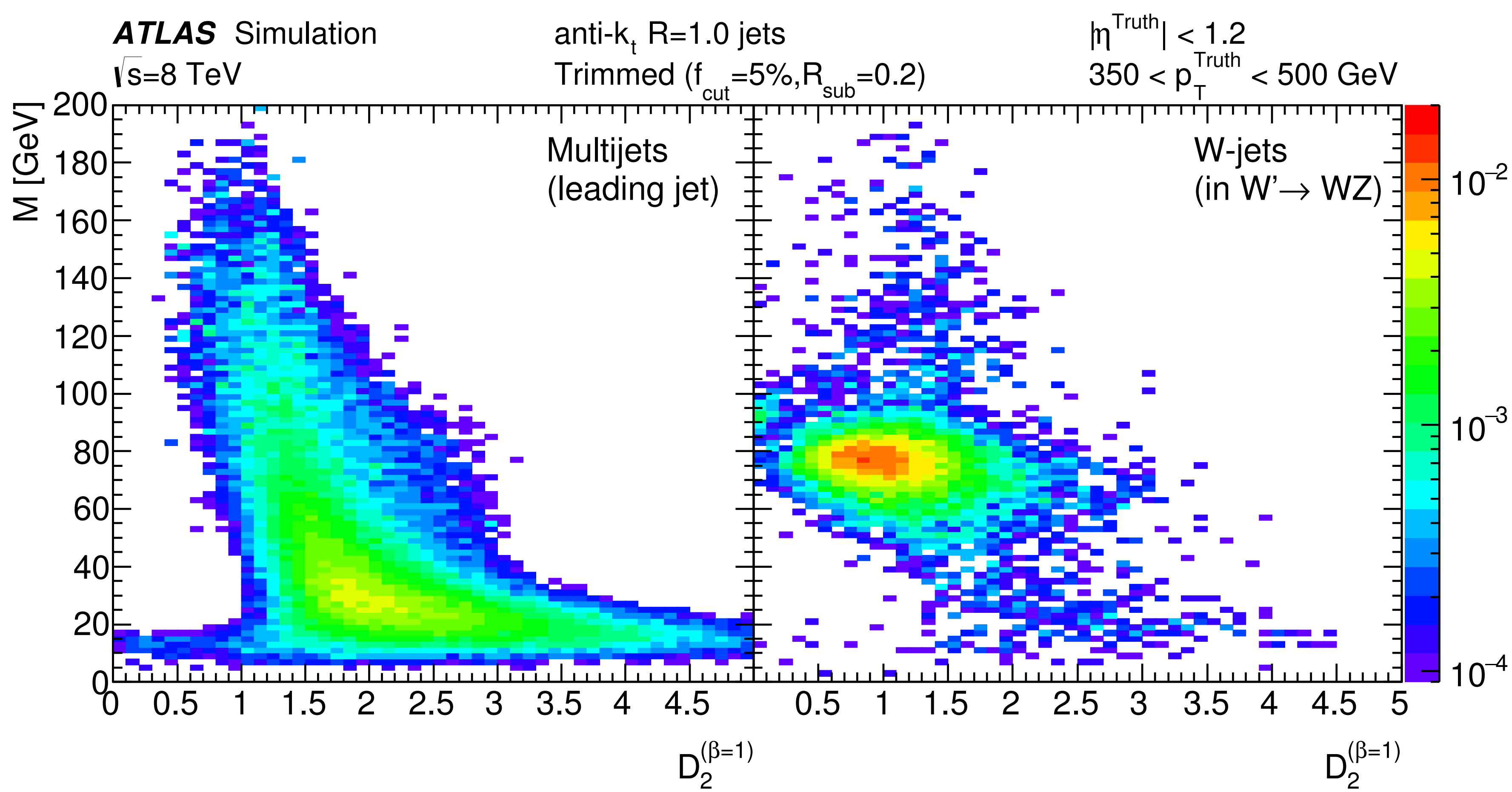
$$e_2^{(\beta)} = \frac{E_{CF}2}{E_{CF}1^2},$$

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

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

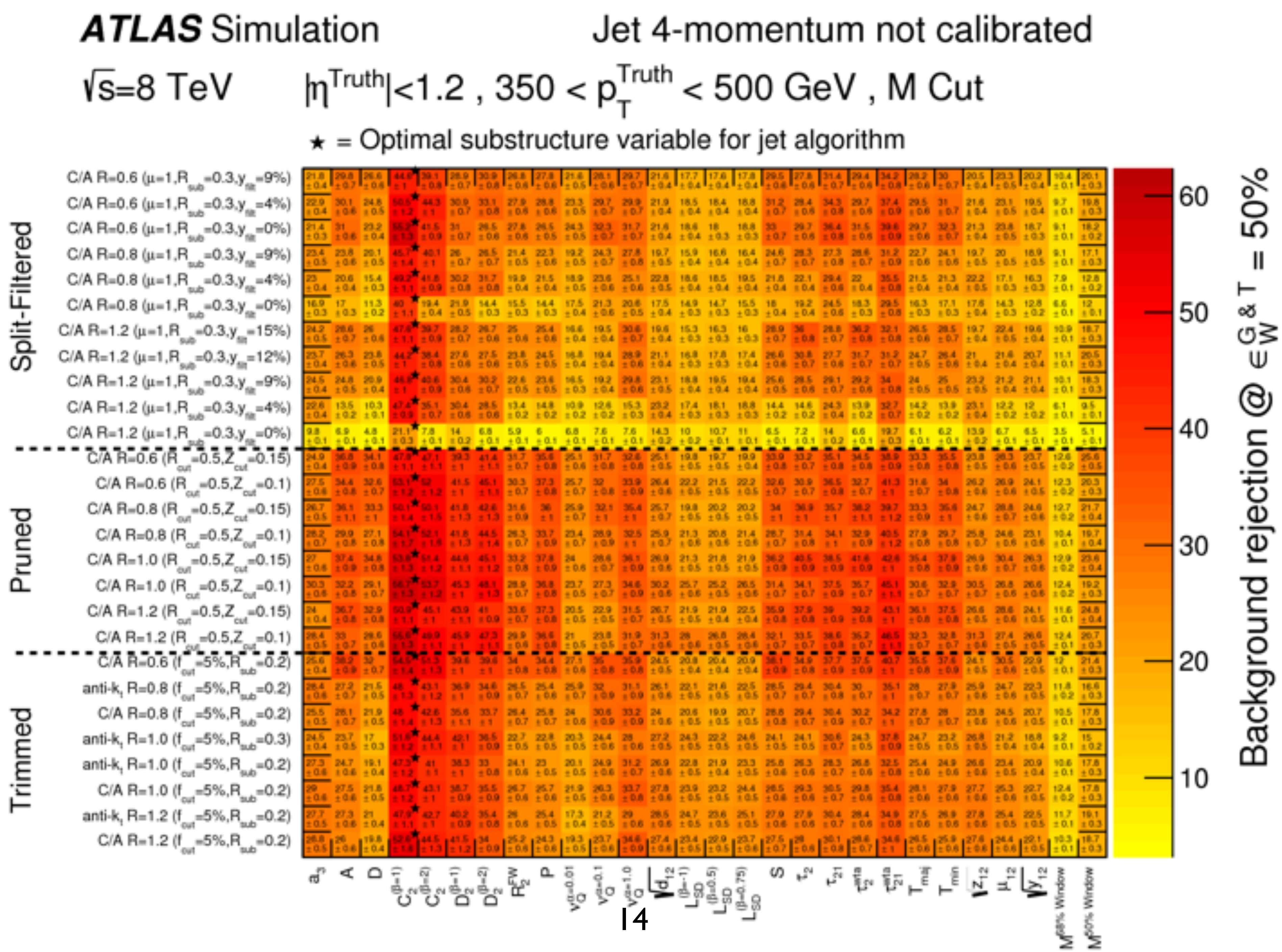
$$\boxed{D_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3}}$$





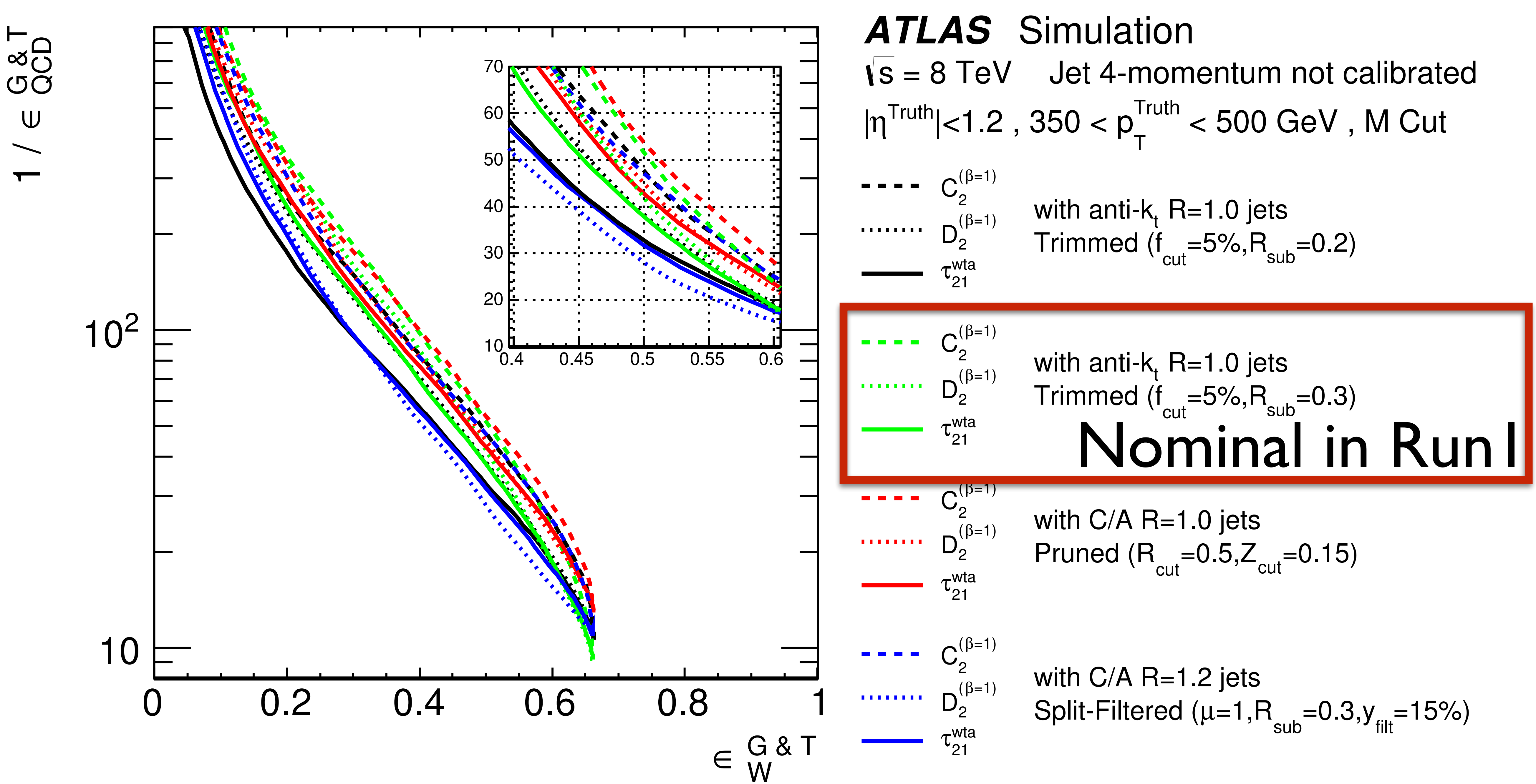
Correlation Studies

26 subsutructure variables studied



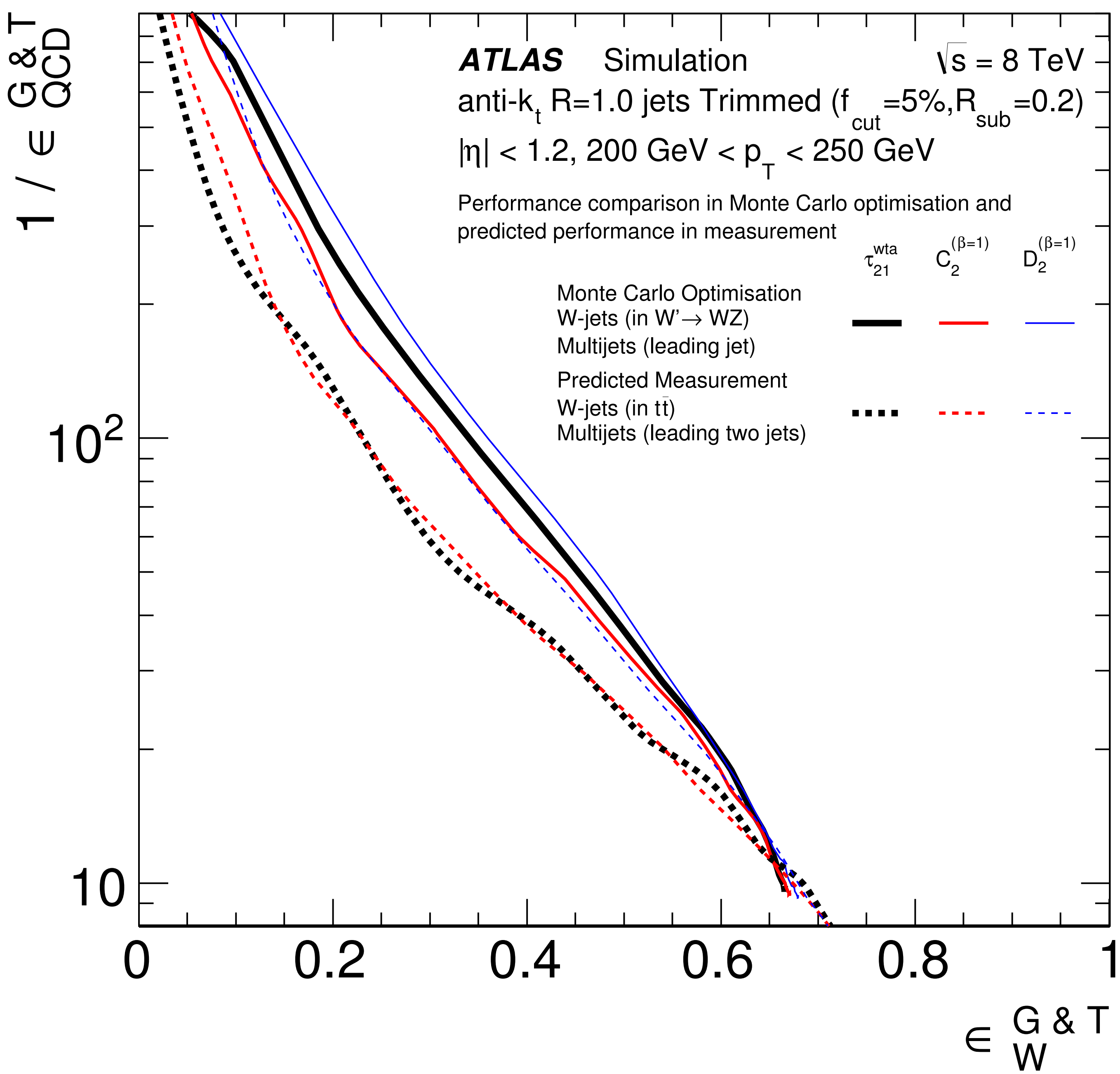
ROC curves comparison

Comparison among grommers + taggers



Best Tagger in Trimming

Reoptimize based on reconstructed jet PT

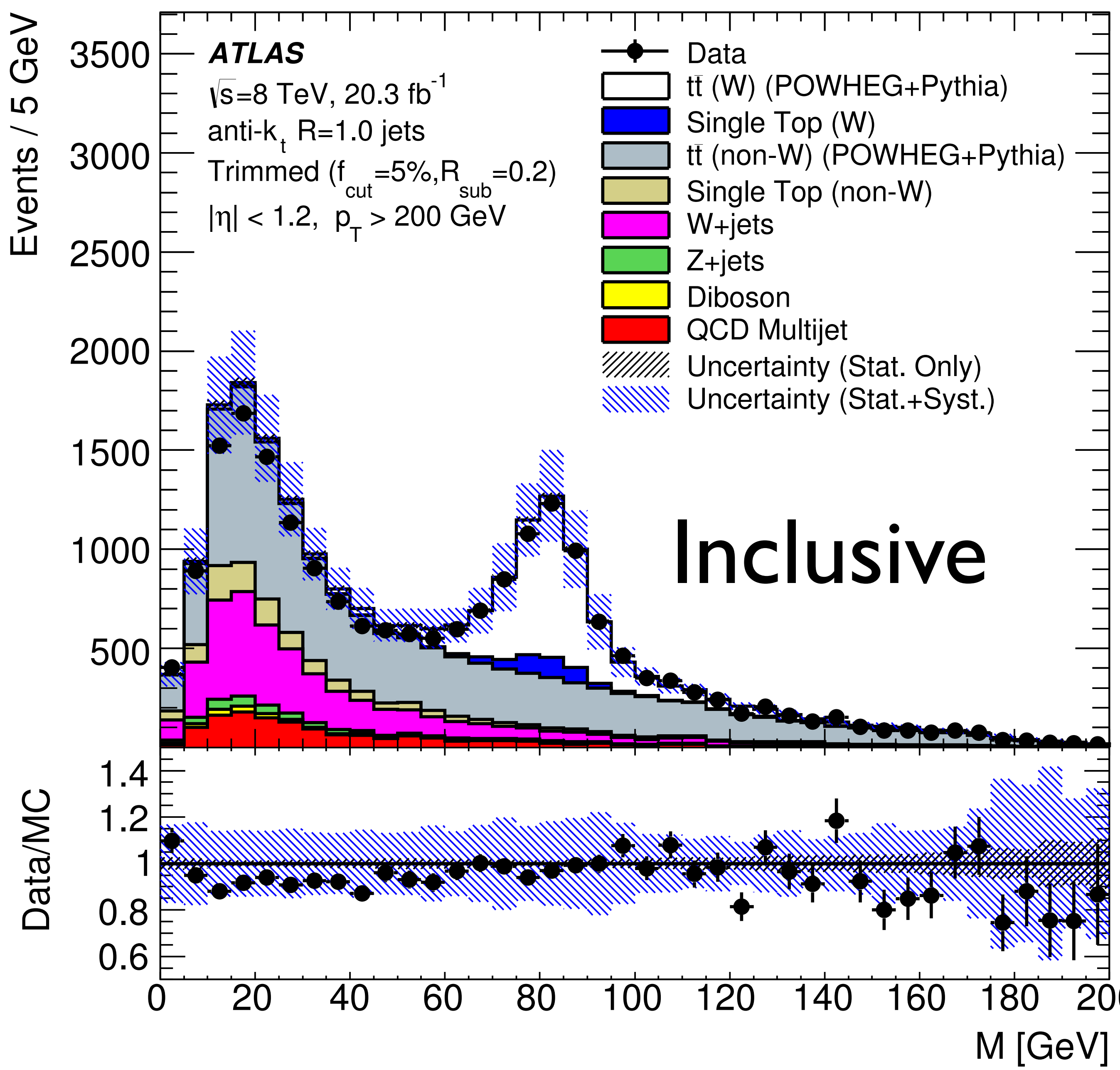


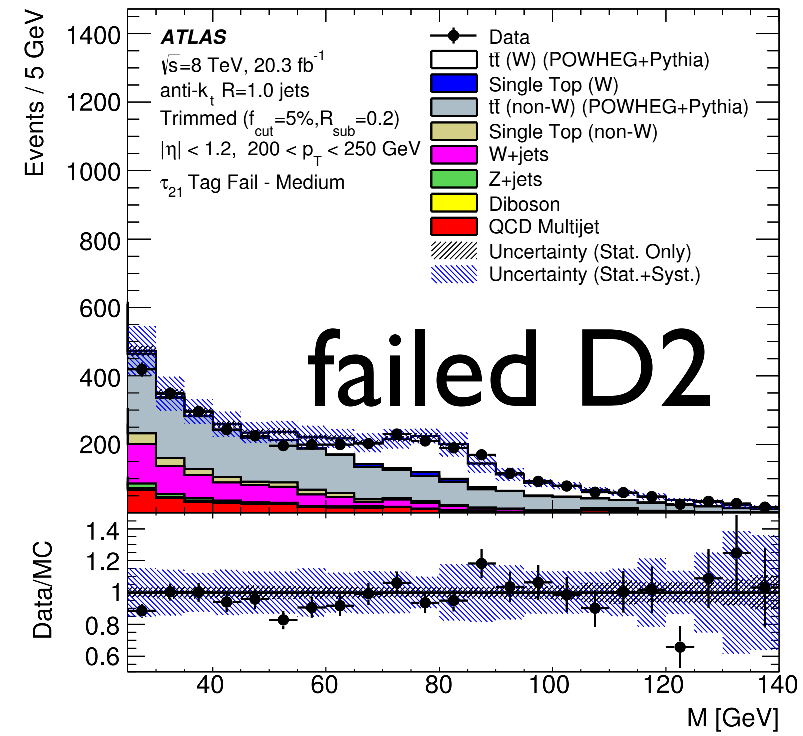
50% Working Point

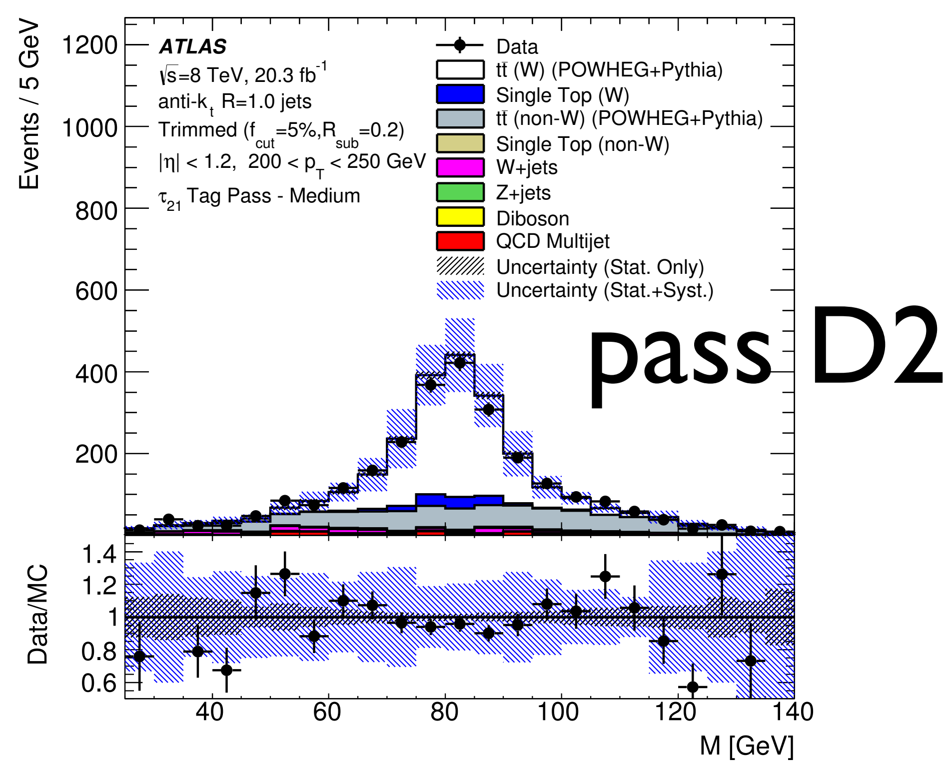
Variable		Tagging criteria in p_T range		
		200–350 GeV	350–500 GeV	500–1000 GeV
ϵ_W^G = 68% mass range		61–93 GeV	71–91 GeV	73–91 GeV
$\epsilon_W^{G\&T}$ = 50%	$C_2^{(\beta=1)}$	< 0.18	< 0.13	< 0.10
	$D_2^{(\beta=1)}$	< 1.14	< 1.23	< 1.35
	τ_{21}^{wta}	< 0.32	< 0.36	< 0.40

Measuring W-tagging efficiency

Selected high purity boosted W from semileptonic Top events

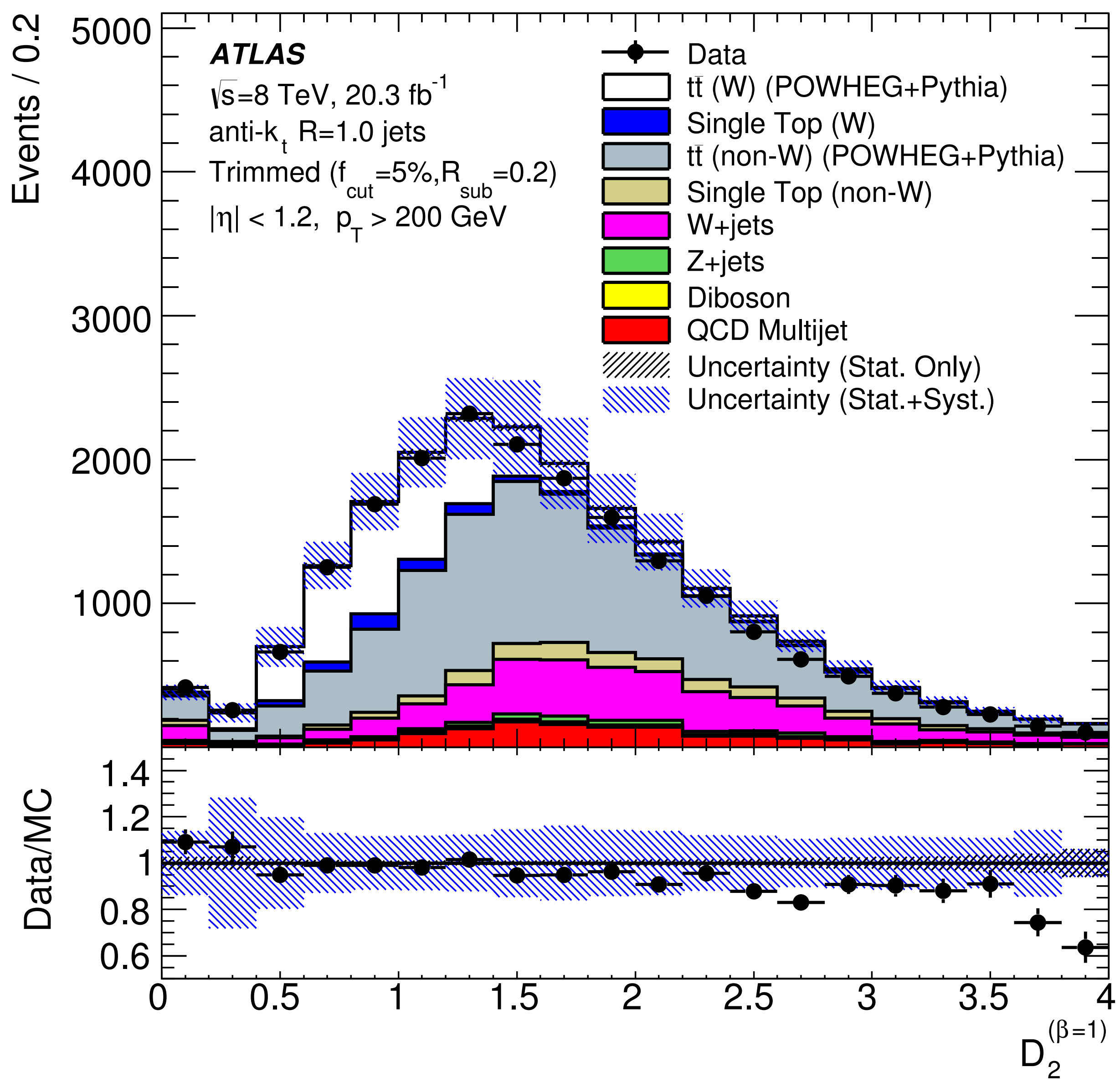




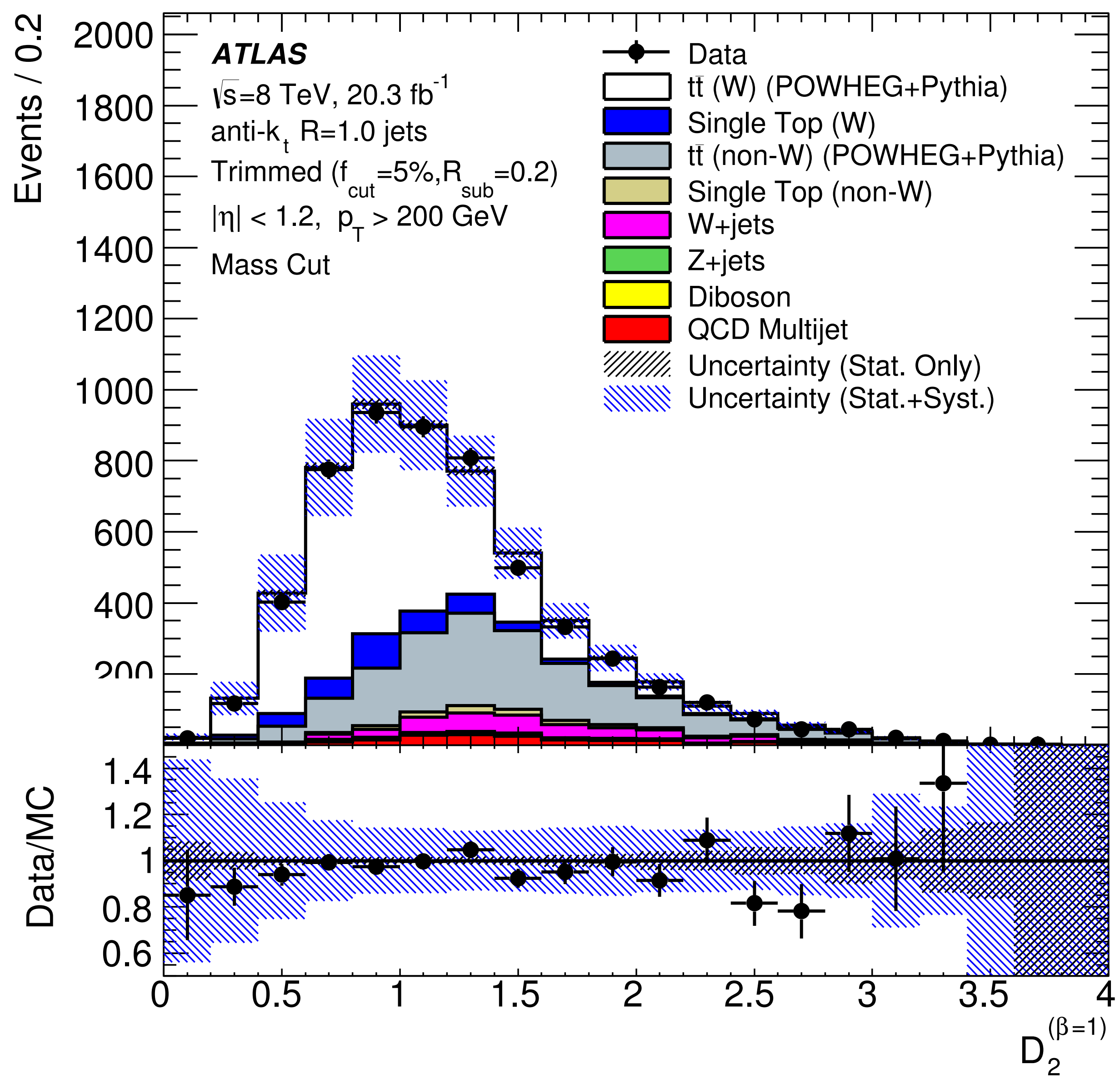


Jet substructure in data

Before 68% mass cut



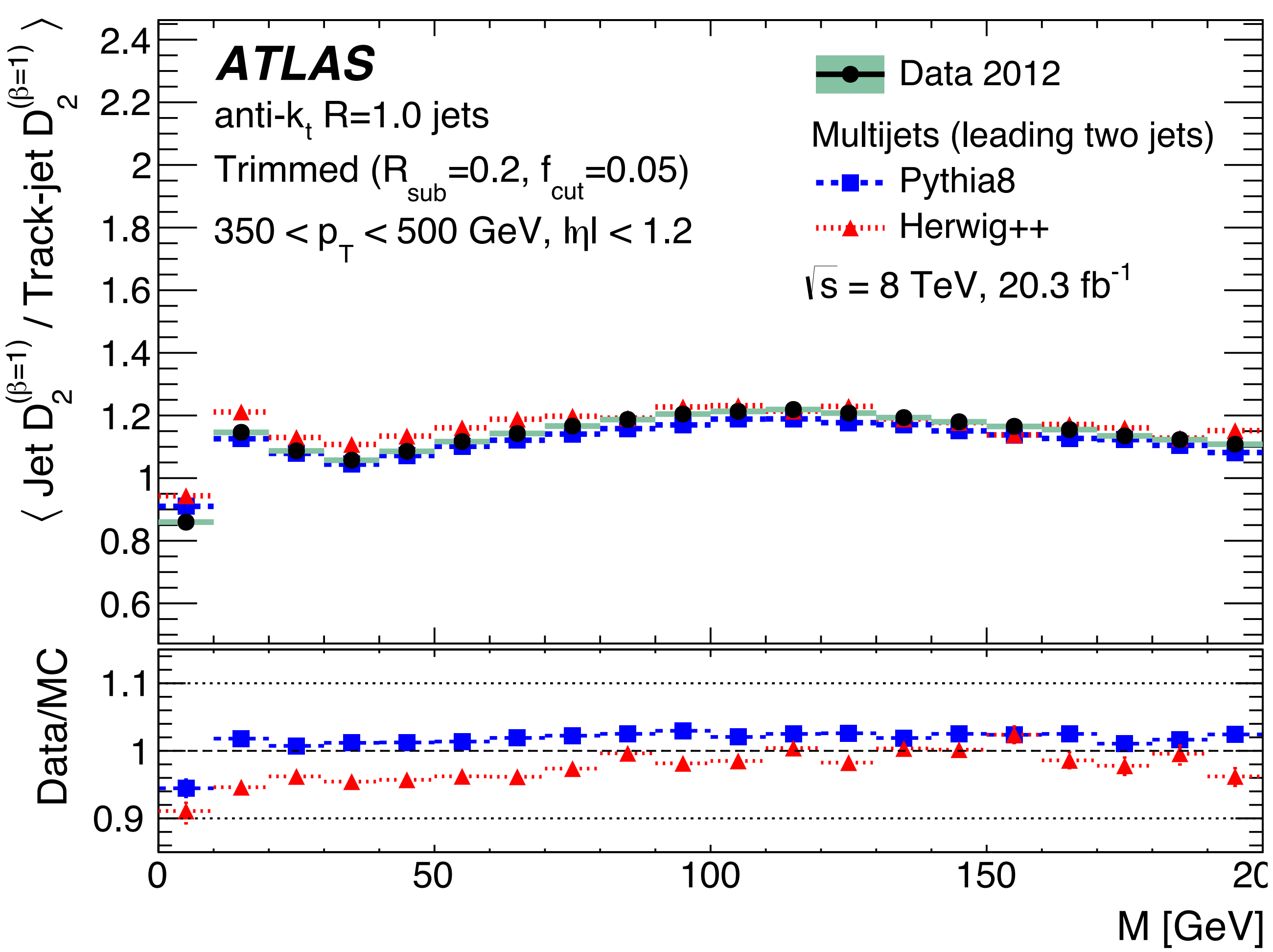
After 68% mass cut

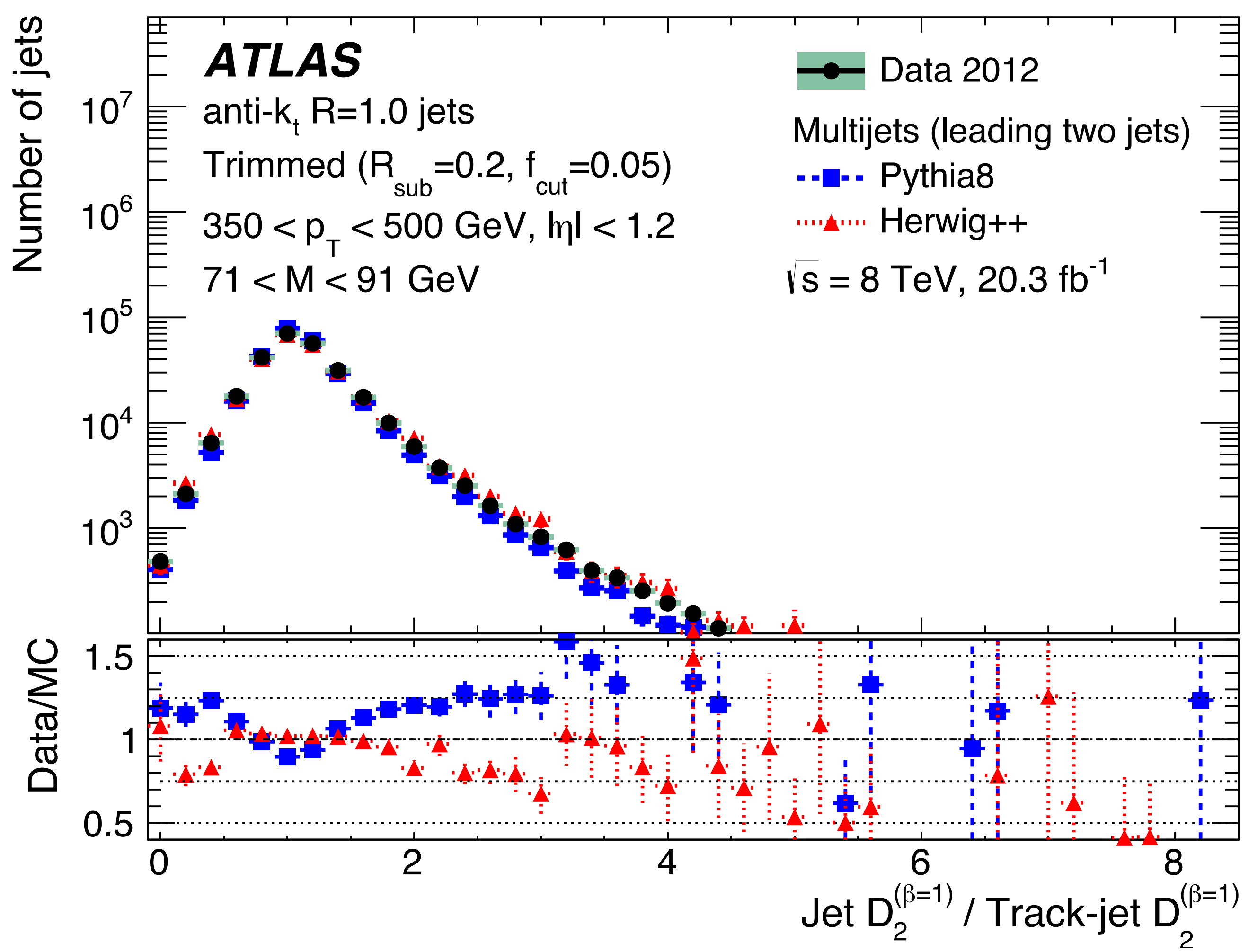


Jet-substructure Uncertainty

Ratio of Jet-substructure from calo/track and calculate Data/MC

It gives us estimate of instrumentation errors as systematic uncertainties





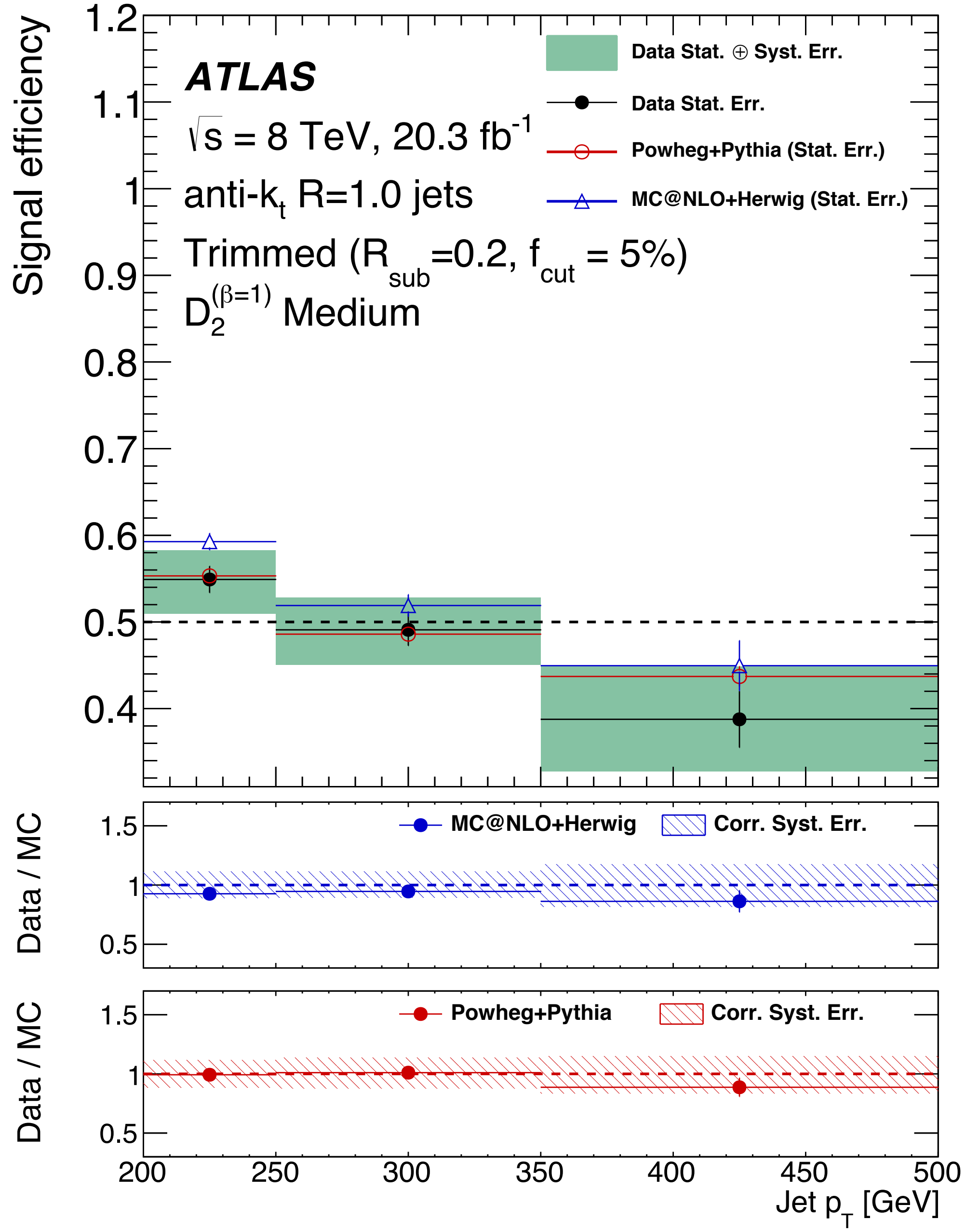
Systematic Uncertainties

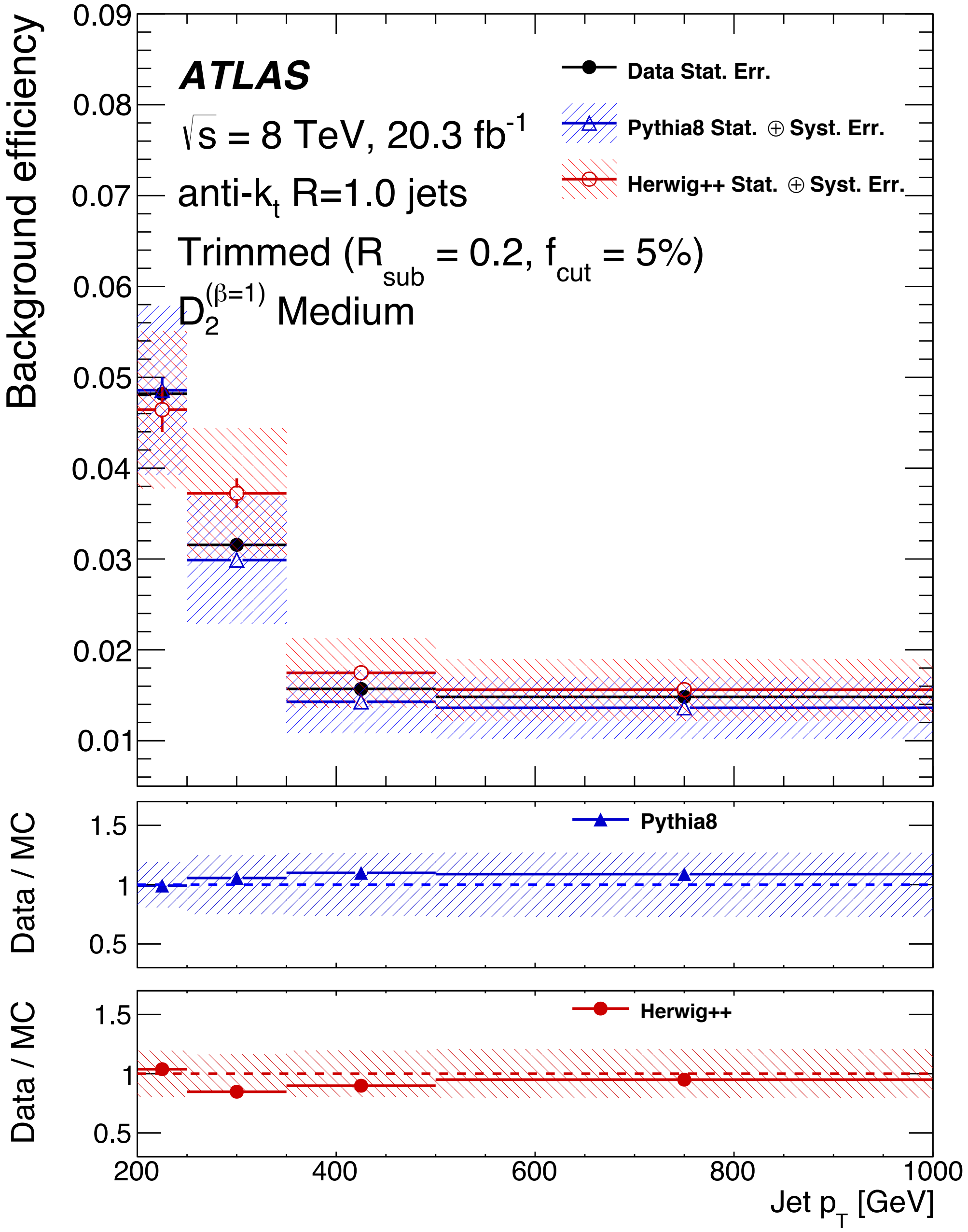
Dominant uncertainty from jet substructure variables

Source	p_{T} range [GeV]		
	200–250	250–350	350–500
JMS	+1.1	+1.1	+9.6
JES	−3.5 / +3.6	−1.7 / +2.5	+1.6 / −2.3
JER	−0.1	+1.0	+1.0
JMR	+2.7	+3.7	+4.3
JSS ($D_2^{(\beta=1)}$)	+4.3 / −2.9	+4.2 / −4.5	+5.1 / −4.8
MC generator	−0.9	+1.9	−3.2
ISR/FSR	+1.6 / −2.2	+2.7 / −4.0	+4.4 / −5.6
Multijet normalisation	−0.4 / +0.4	−0.3 / +0.3	+0.1 / −0.1
Single-top normalisation	−0.1 / +0.1	−0.1 / +0.1	−0.1 / +0.1
$t\bar{t}$ normalisation	0.6 / −0.5	+0.6 / −0.6	+0.5 / −0.5
W+jets normalisation	−0.3 / +0.3	−0.4 / +0.4	−0.5 / +0.4
MC statistics	−1.0	−1.5	−3.5
Total	+6.6 / −5.4	+7.3 / −6.6	+13.1 / −13.2

Efficiency Measurement

- Overall good agreement between simulation and data.
- Herwig has relative larger deviation but still within one sigma uncertainty

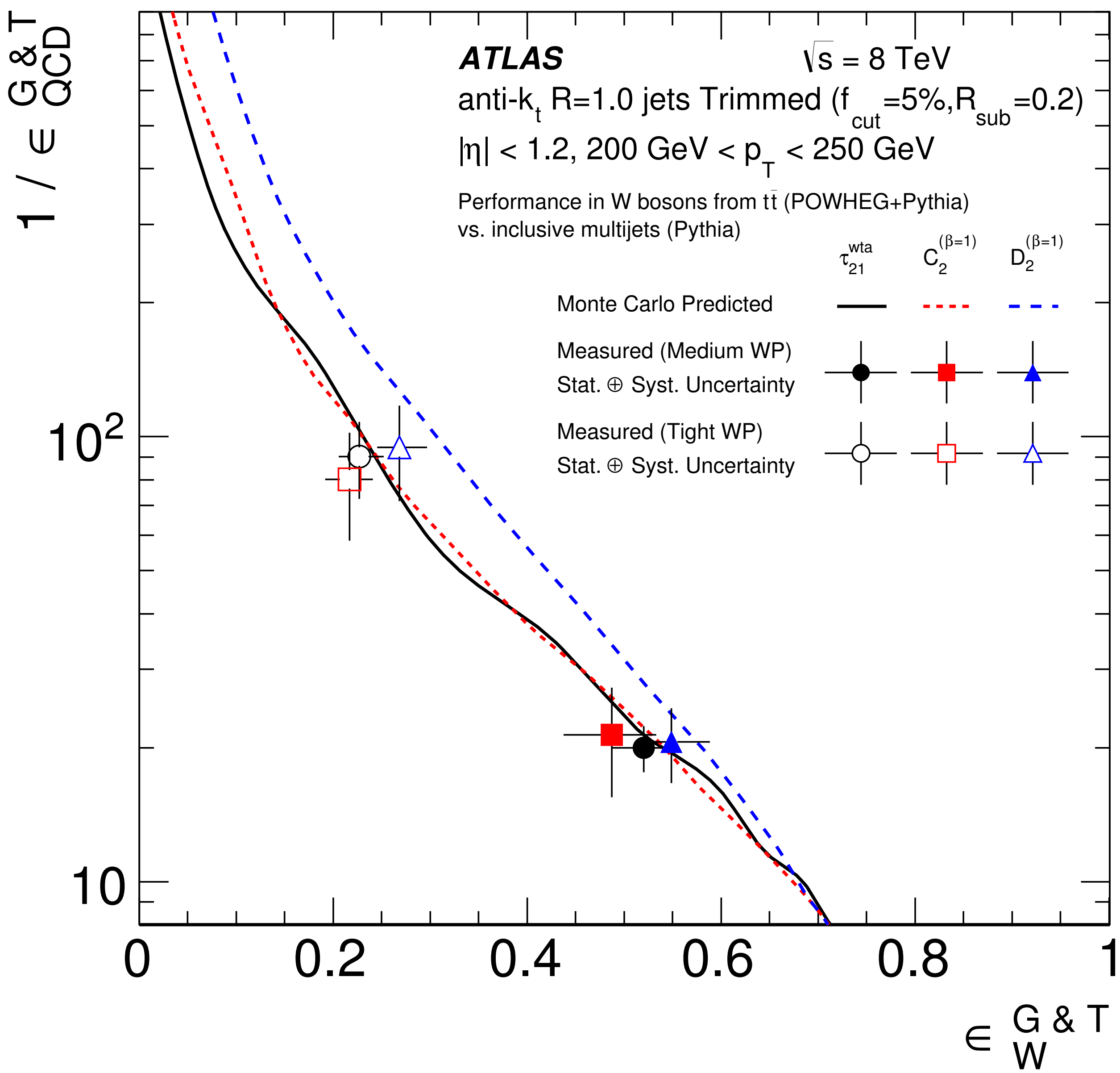


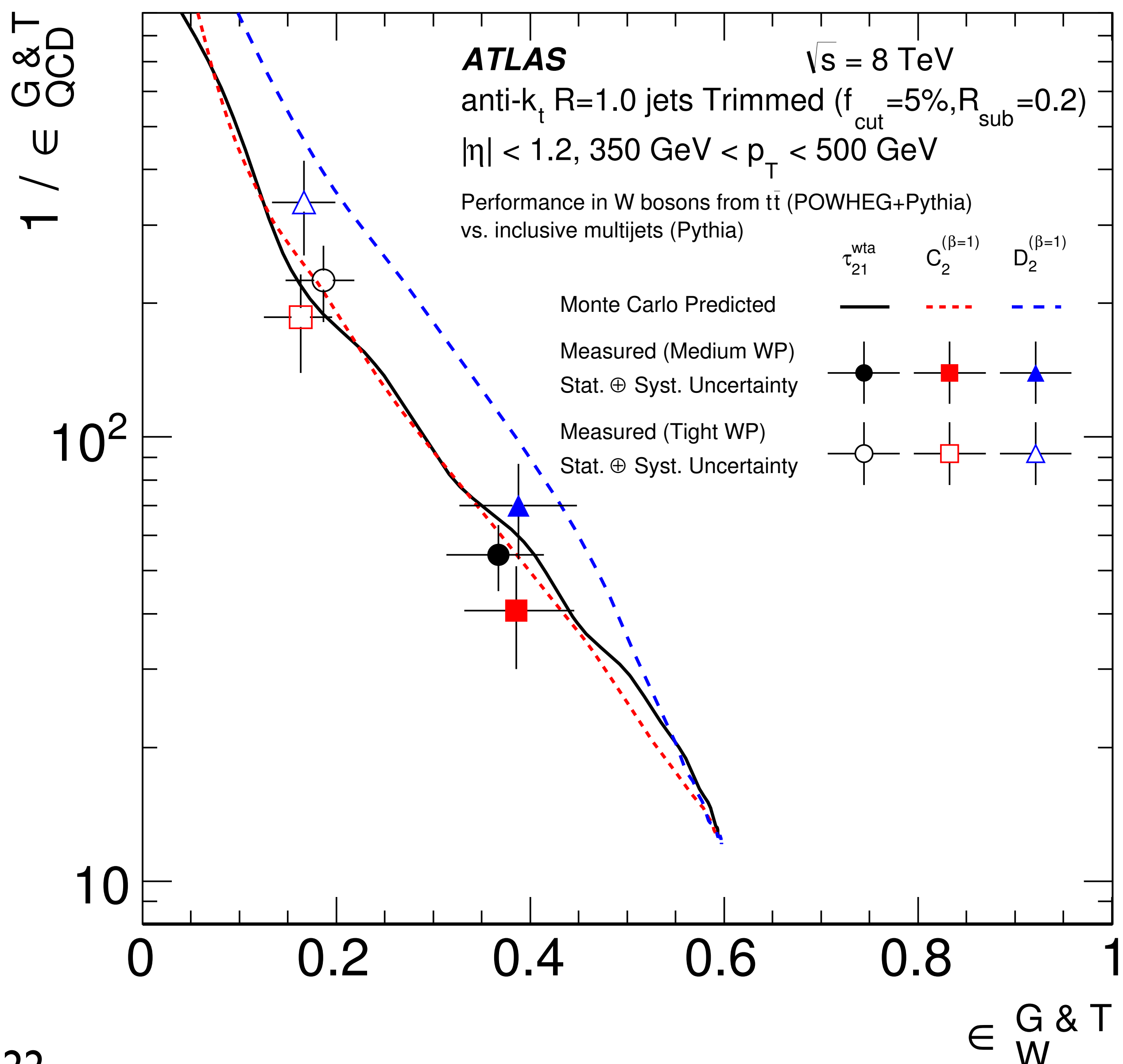


Money Plots

Possibility to derive in-situ W-tagging efficiency measurements and derive scale factor as correction to simulation. Open questions:

- Can it be applied to different physics analysis?
  What is the topological dependence of tagger performance?



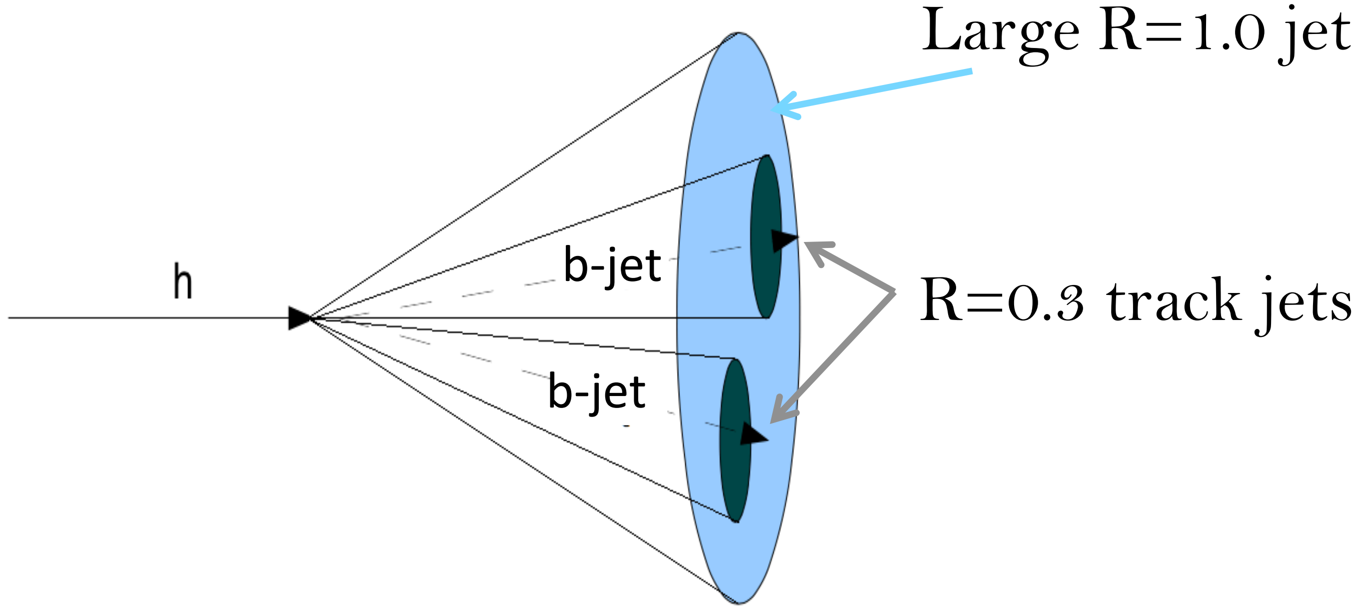


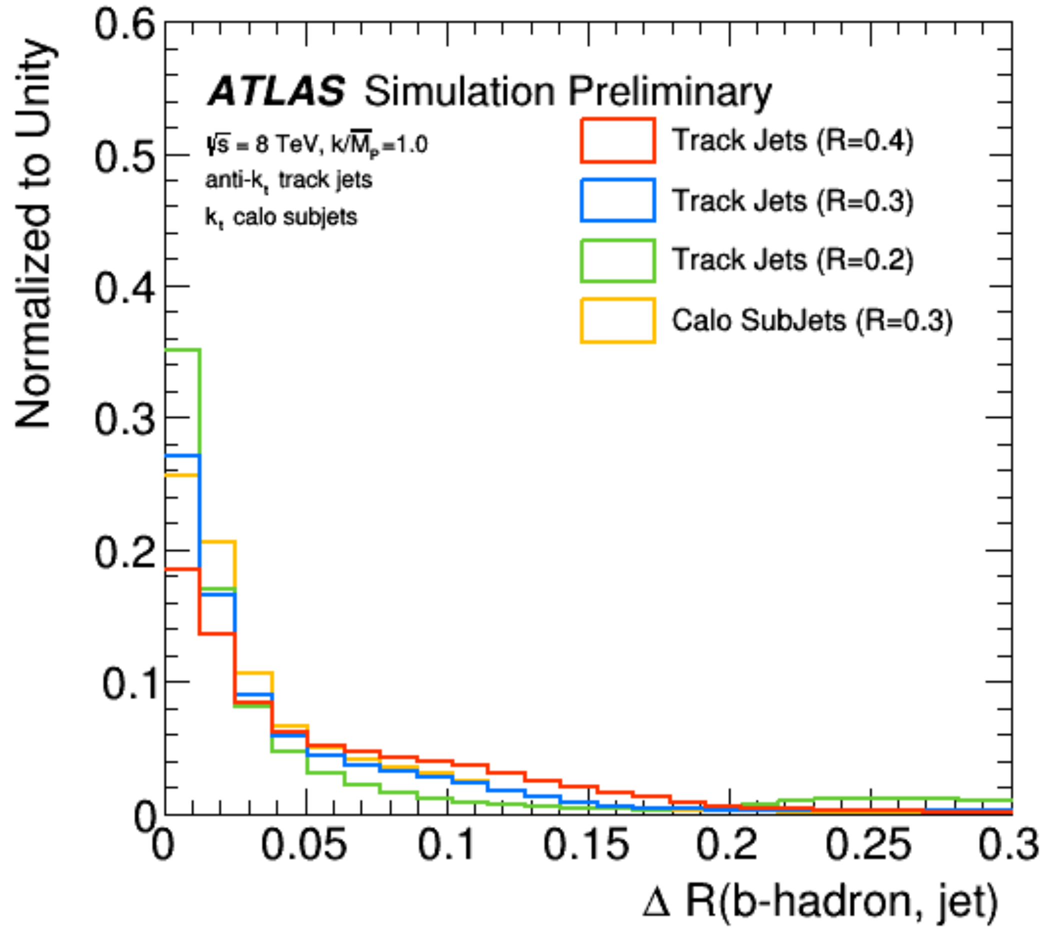
Hbb tagging

Higgs tagging

ATL-PHYS-PUB-2014-013

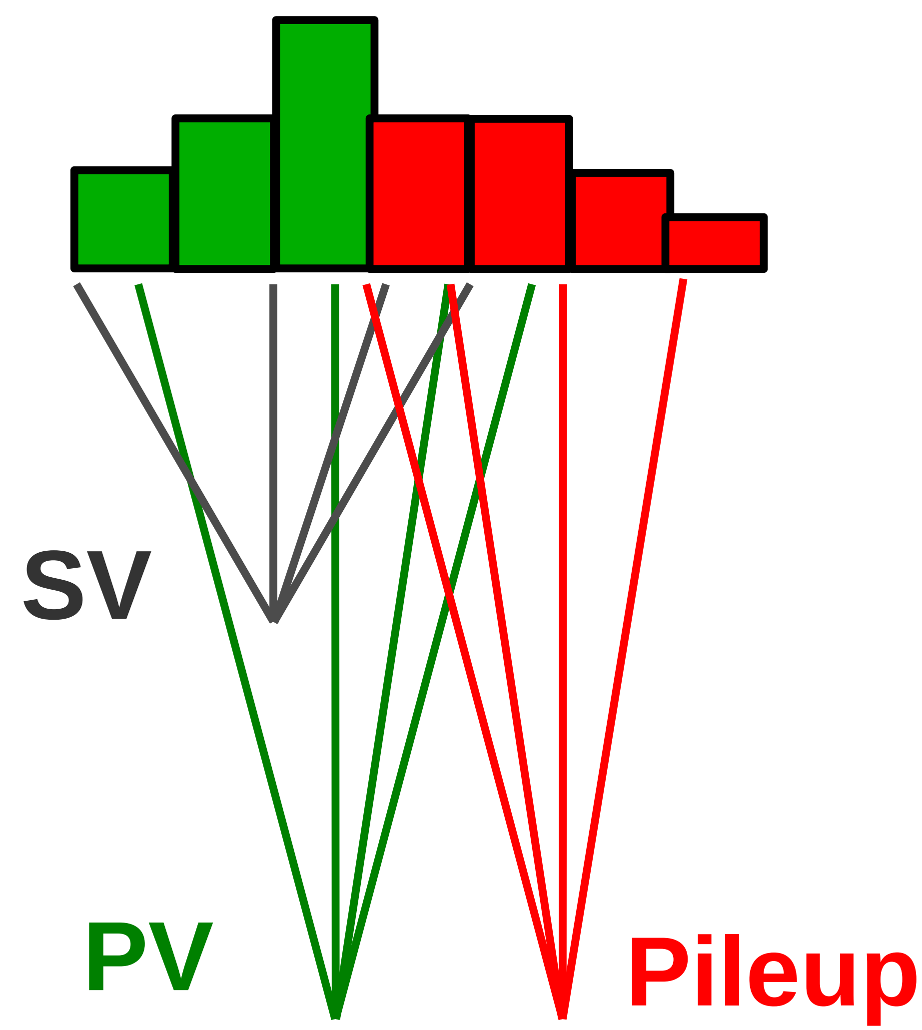
- R=1.0 calorimeter jets trimmed with kT R=0.3 subjects and fcut=0.05 to measure kinematics and substructure
- b-tagging with small R=0.3 track jets to resolve close-by b-hadrons and allow calibration of each track jet individually





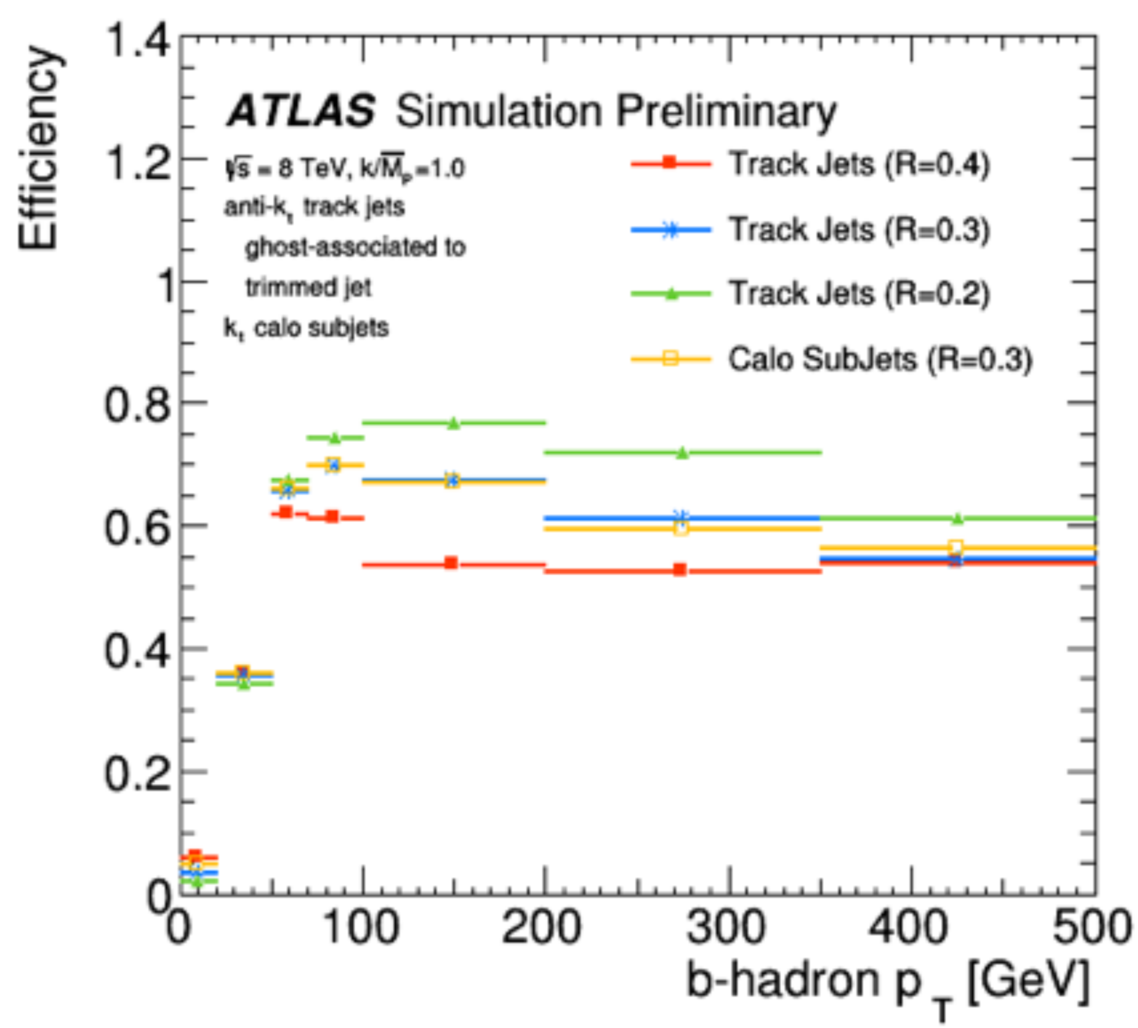
Track Jet b-tagging

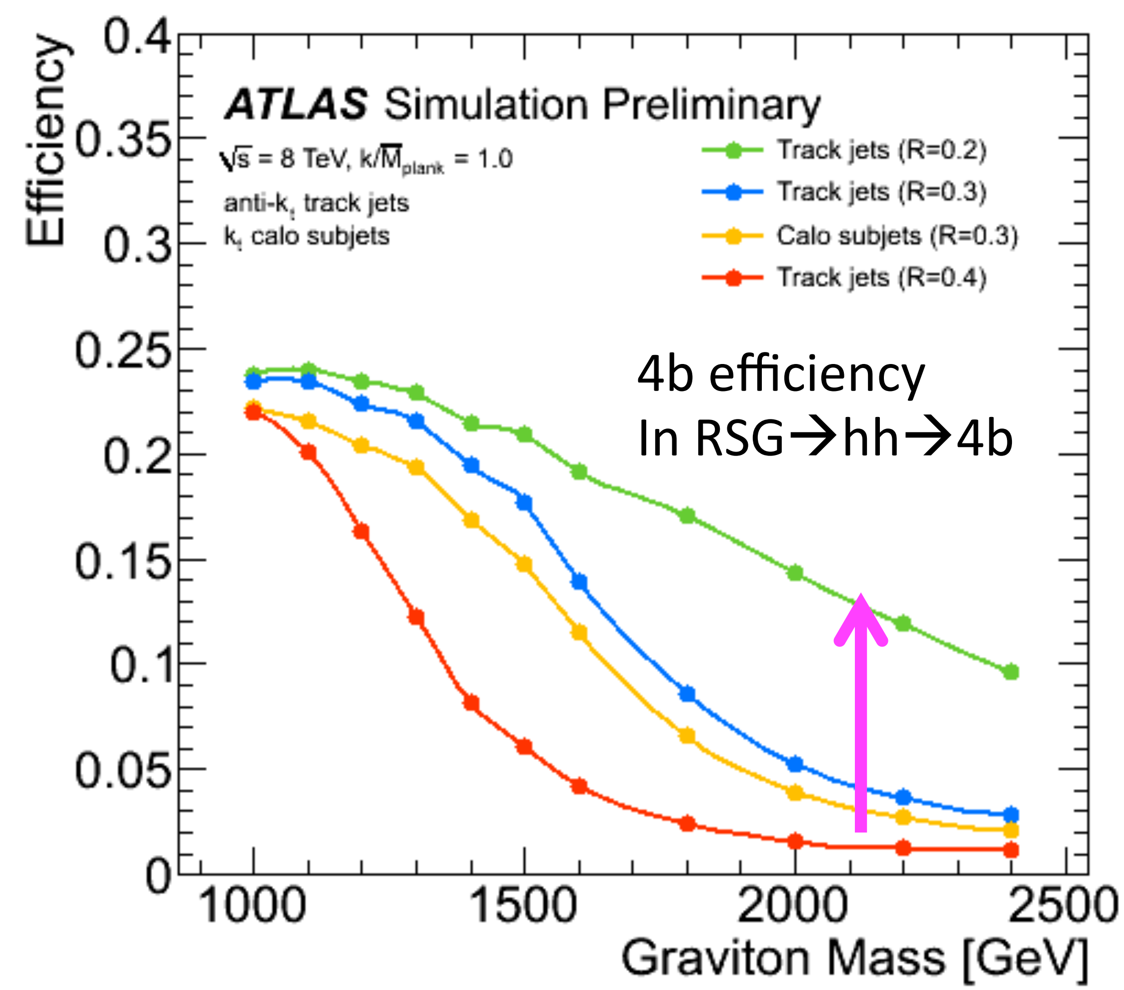
- Use charged particles to build track jets
- B-tagging using track jets
  - insensitive to pileup
  - small radius to identify close-by objects
  - Independent of calorimeter
  - good angular resolution w.r.t. b-hdaron
- Ghost association
  - associate track jets by setting pT of each track jet to a number close to 0 (ghost particles)
  - cluster ghost particle together with calorimeter jets



Higgs tagging

- Ghost association of track jets to large-R to provide b-tagging (matching to ungroomed parent jet)
- Large improvements in efficiency to find Higgs jets!

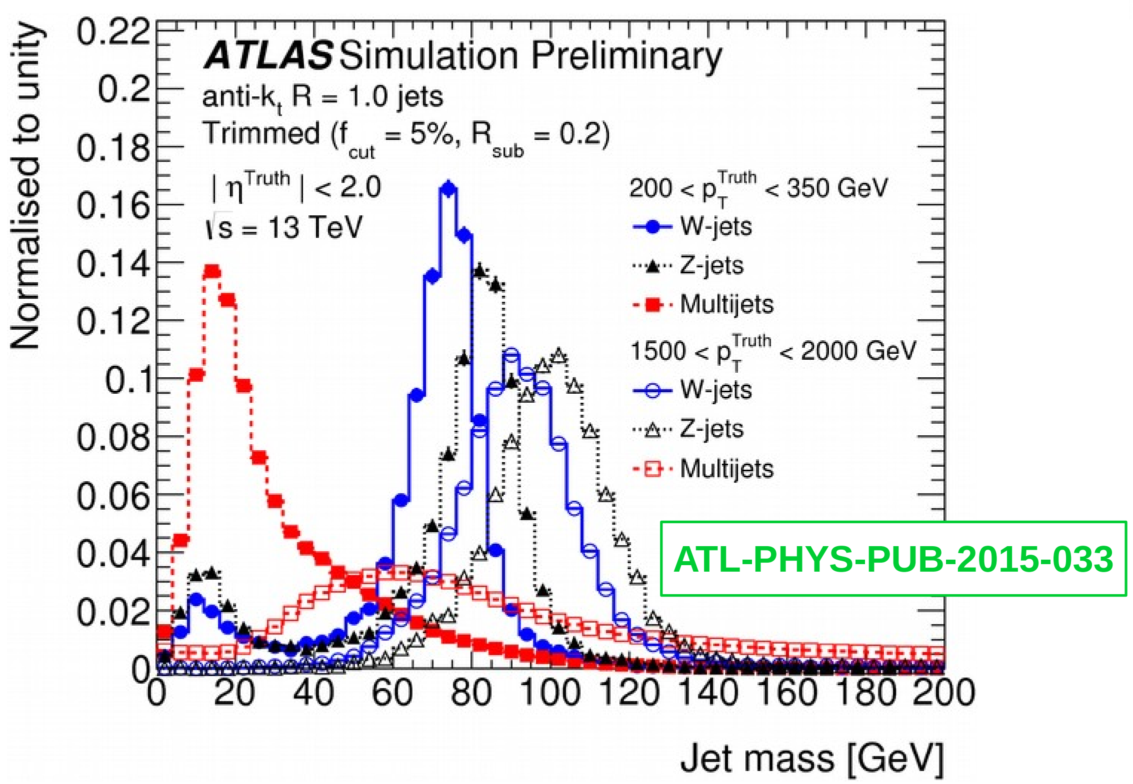


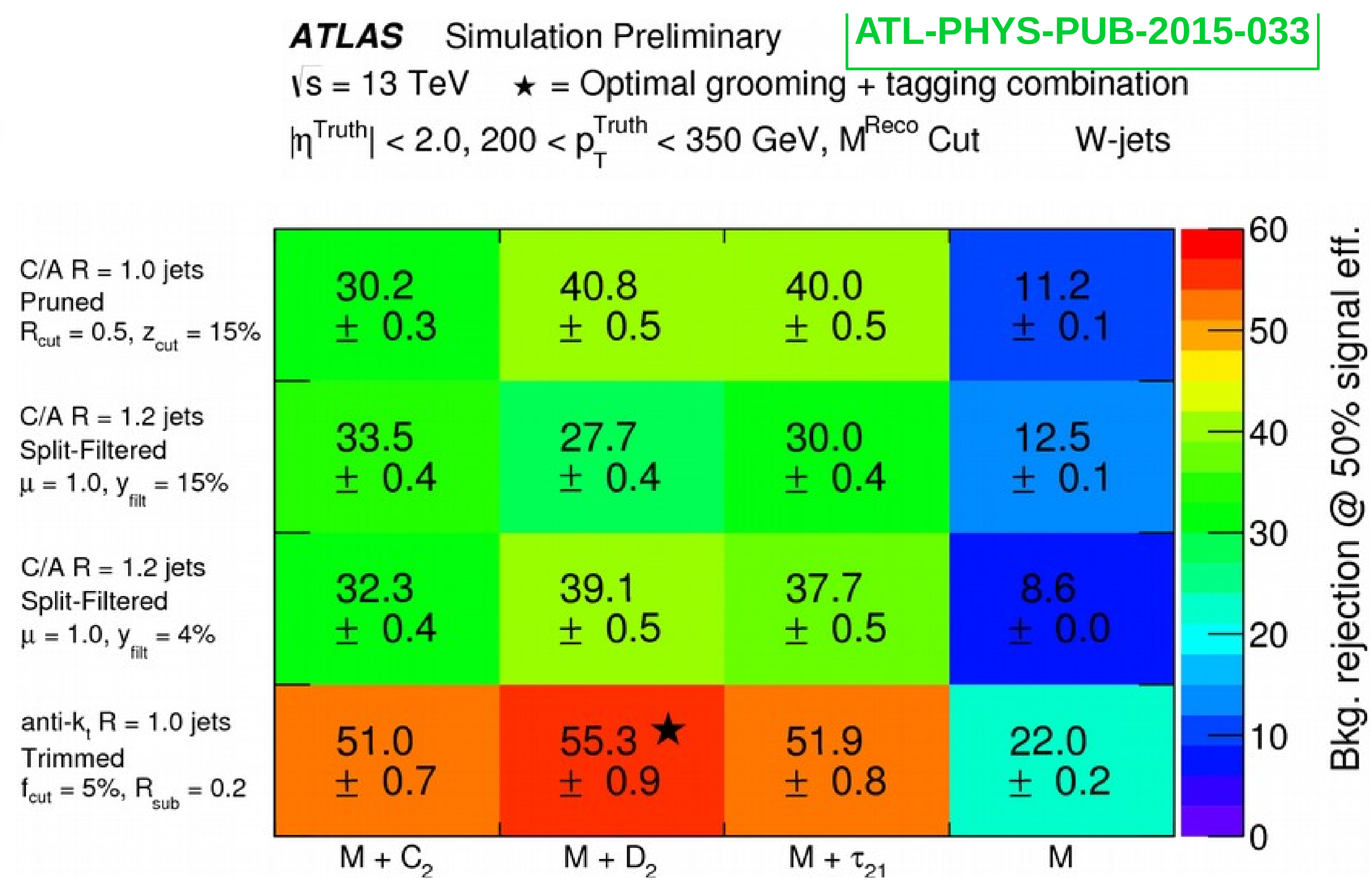


Run2 Prospects

W/Z-tagging

- Similar approach as Run1 except smaller subject radius, i.e. Rsub=0.2
- Continuous pT parametrization of D2 cut
- Performance still to be checked with Run2 data?

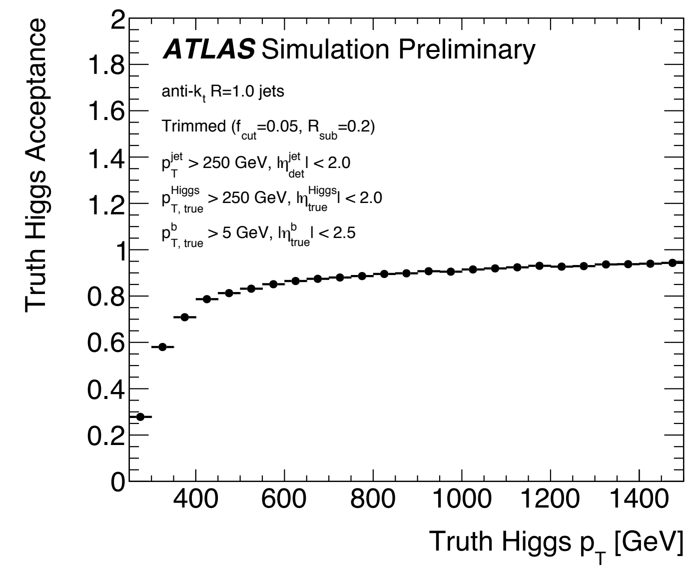


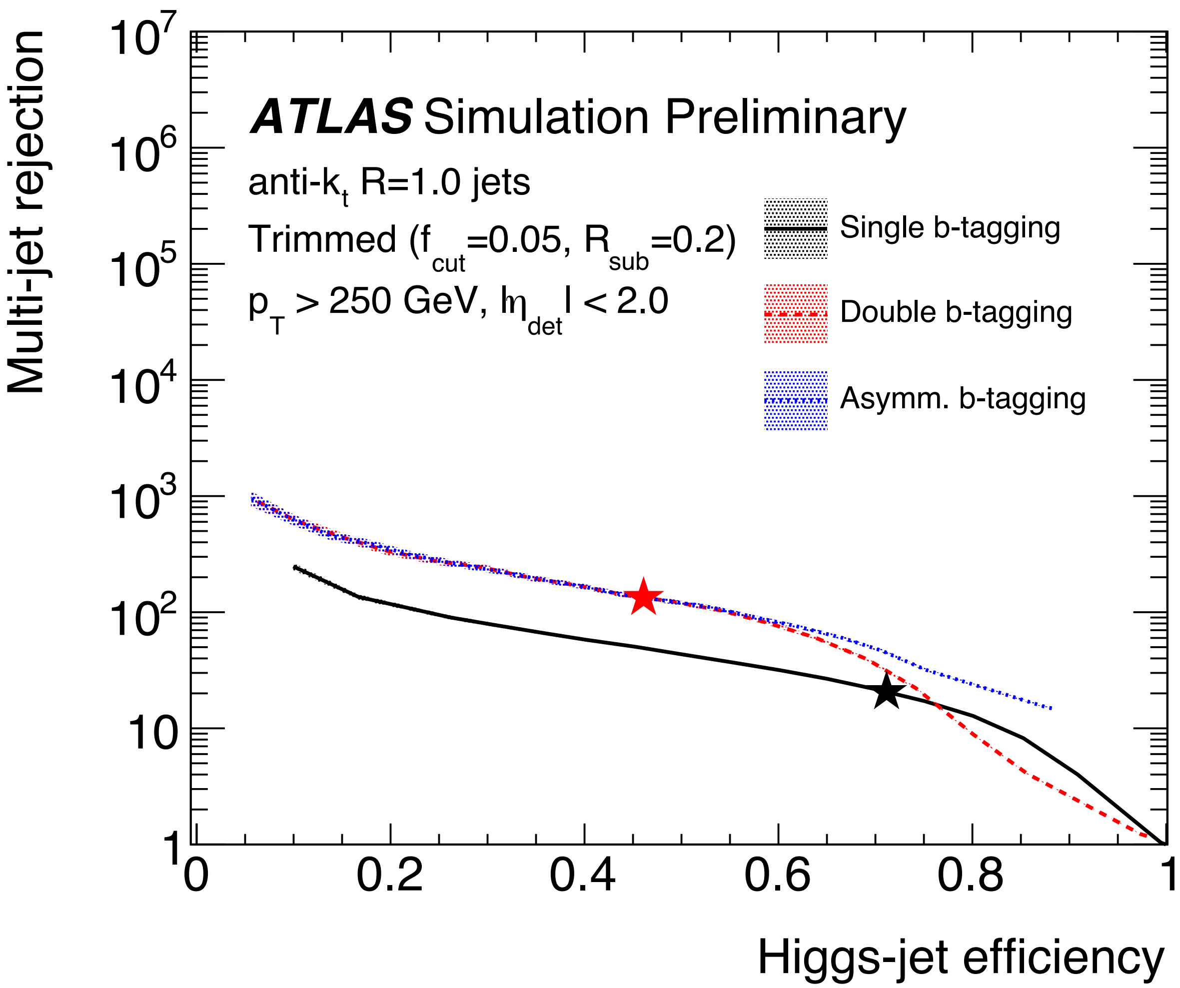


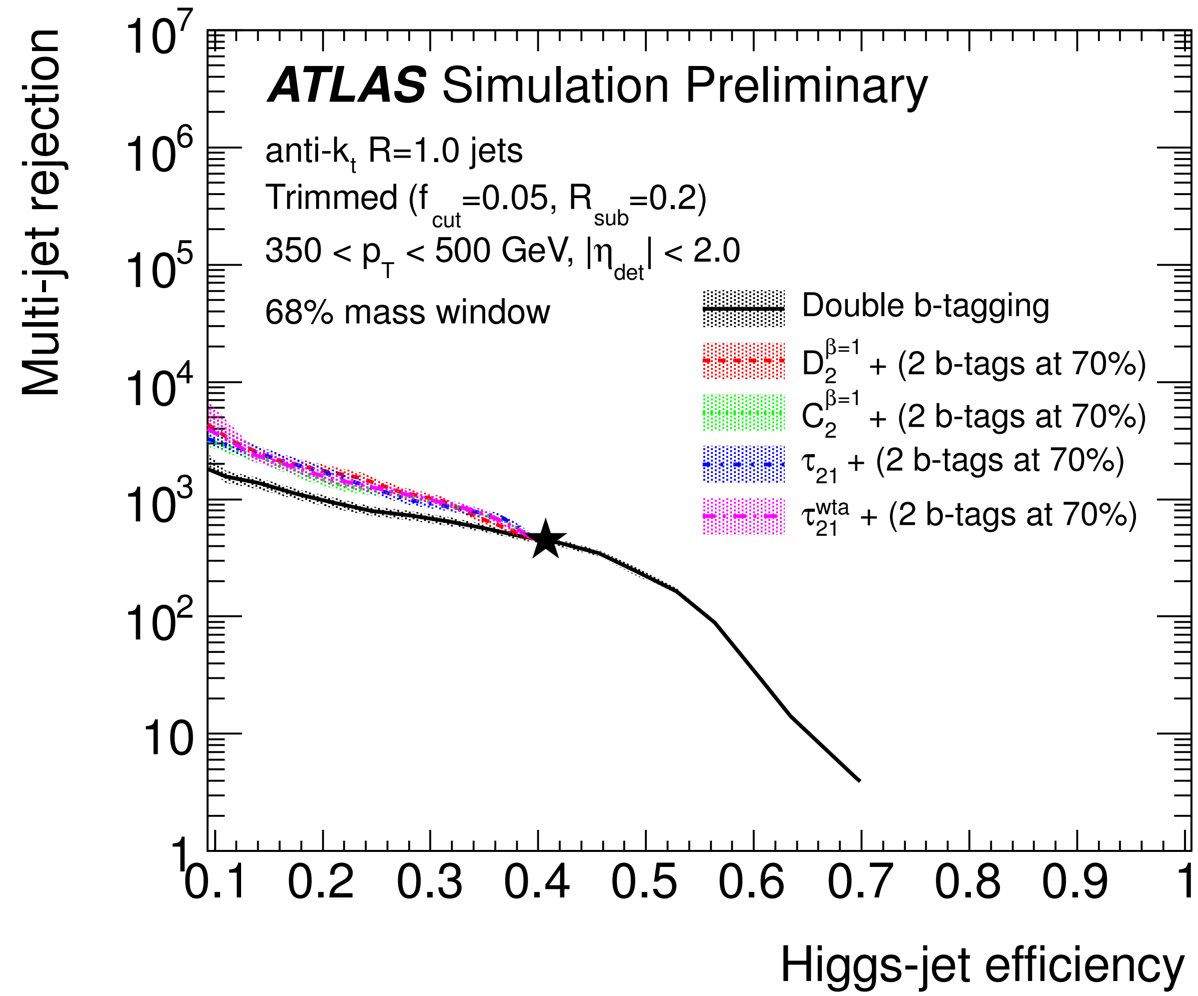
Hbb-tagging

ATL-PHYS-PUB-2015-035

- Similar approach as Run1 except smaller
  subject radius, i.e. Rsub=0.2
- Combination with D2 cut



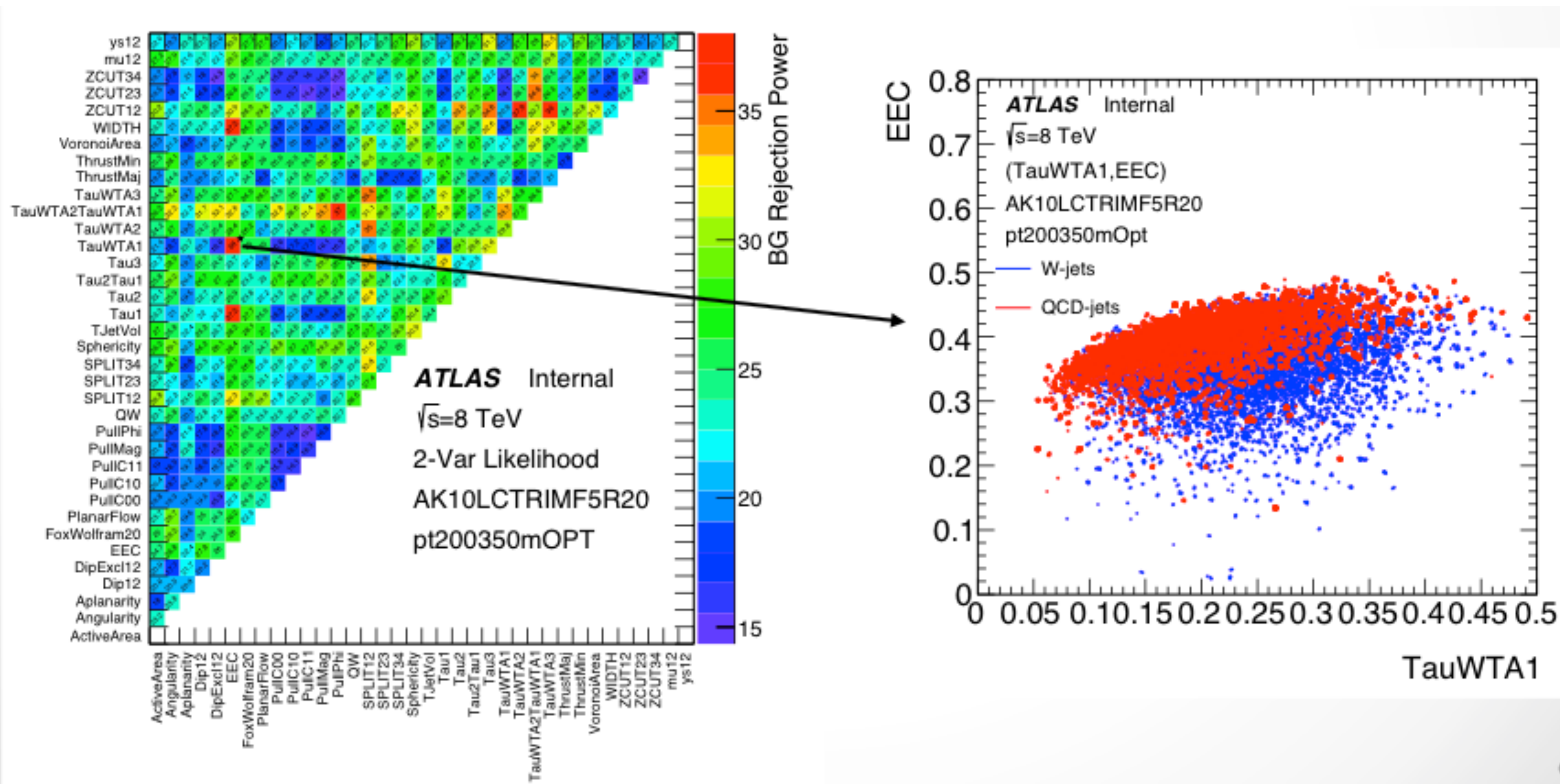




New ideas

MVA Tagger combination

Starting from all two-tagger combinations to determine ordering of combing various taggers



Q-Jet: A new approach to parton shower

arXiv:1201.1914
ATLAS-CONF-2013-087

A typical jet clustering is to inverse the parton shower

Q-Jet: A new approach to parton shower

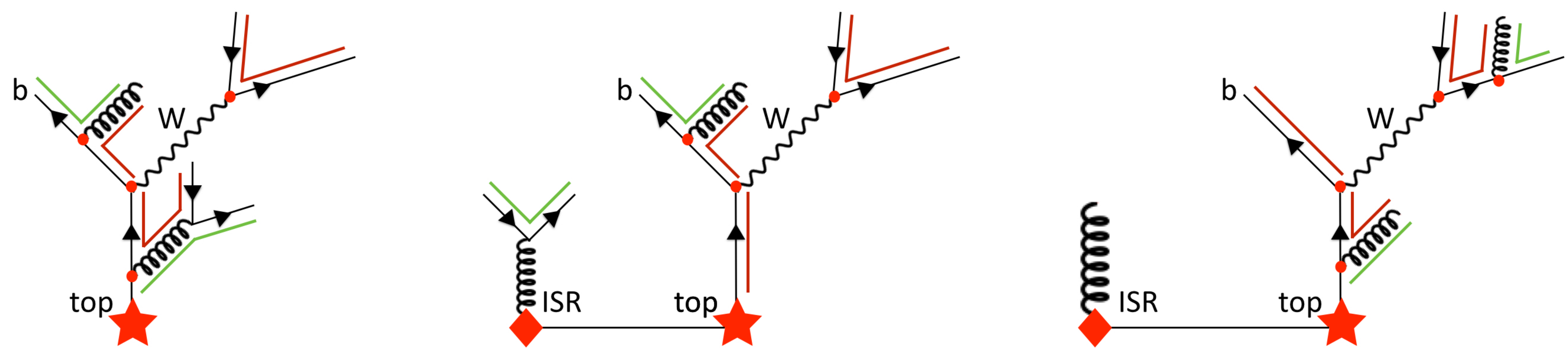
arXiv:1201.1914
ATLAS-CONF-2013-087

A typical jet clustering is to inverse the parton shower

But the parton shower is not really invertible
Many parton shower can produce the same jet
Q-jet asks: study all possible inverse parton shower path.

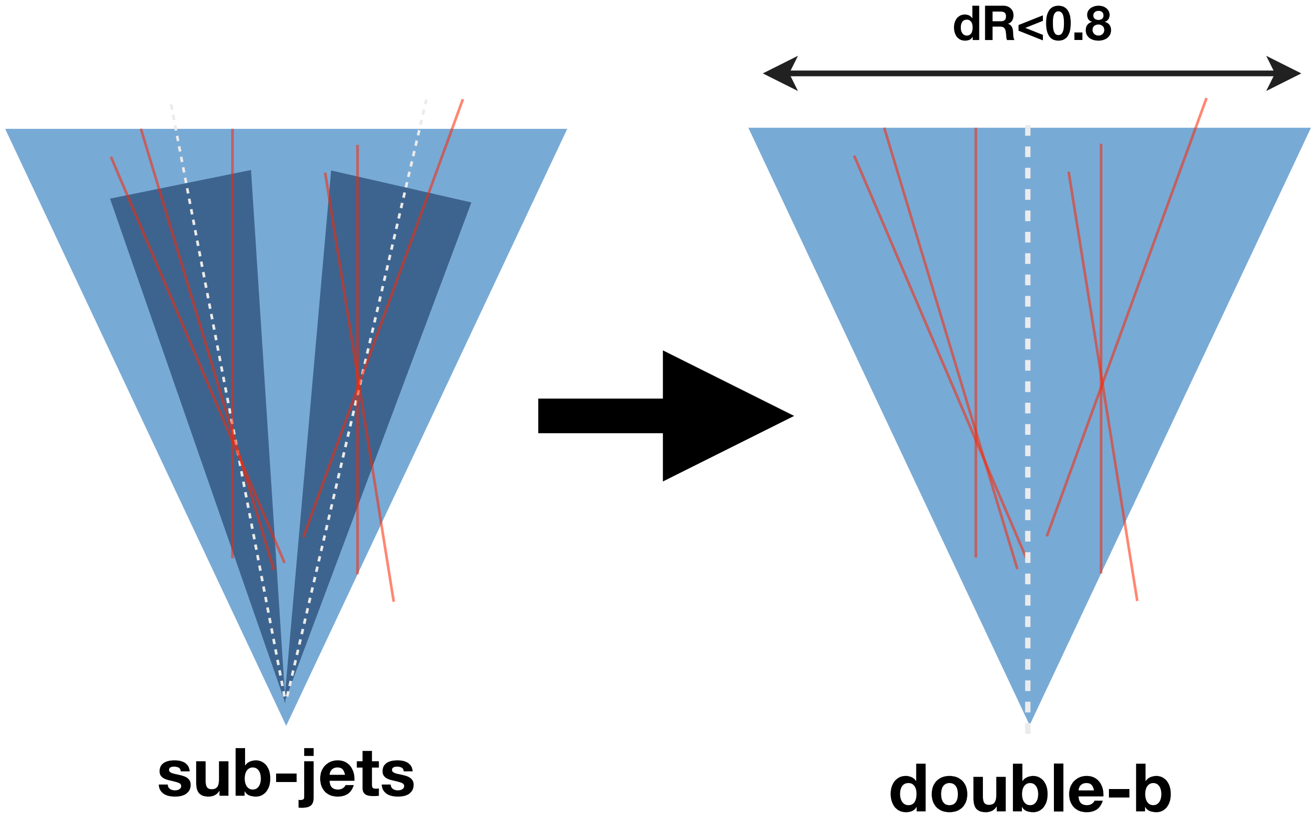
Shower Deconstruction

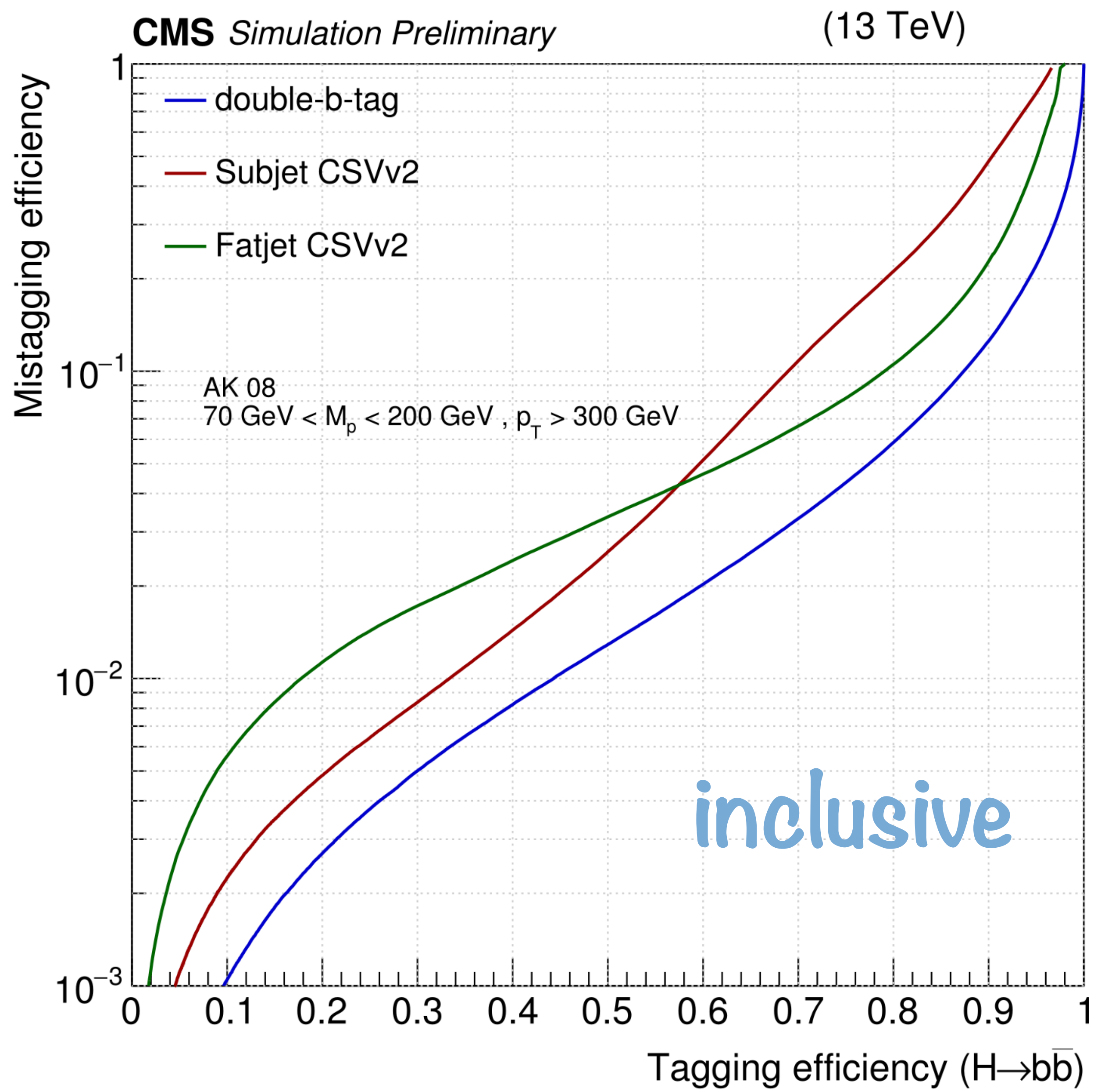
- Calculate, for each subjet of the input, the probability that the subjet is associated with a certain source of radiation (ISR, light quark, etc.).
- Discriminant χ is ratio of sum of signal probabilities to sum of background probabilities.



Multi-Secondary Vertex

Using inclusive multi-secondary vertex for double-b tagging





proof of ideas work in CMS

Summary

- Boson tagging as a new tool have been successfully
  used in Run1.
- Many innovative ideas under developments for Run2
  - How to use all the information using Machine Learning tool?
  - How to extract information from QCD splitting processes?
  - How to improve input object reconstructions by improving reconstruction algorithm?
- Gor for unltra-boost - challenges on the ultimate detector resolution!

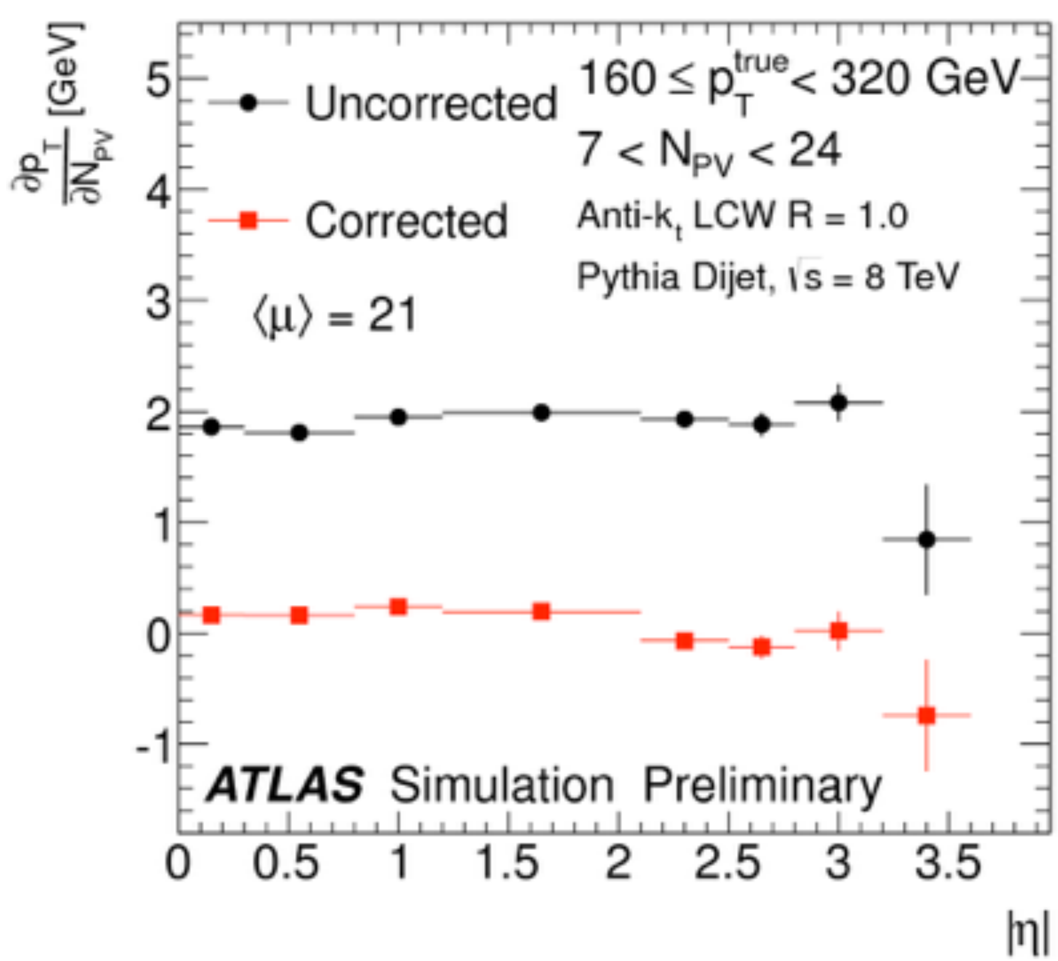
Backup

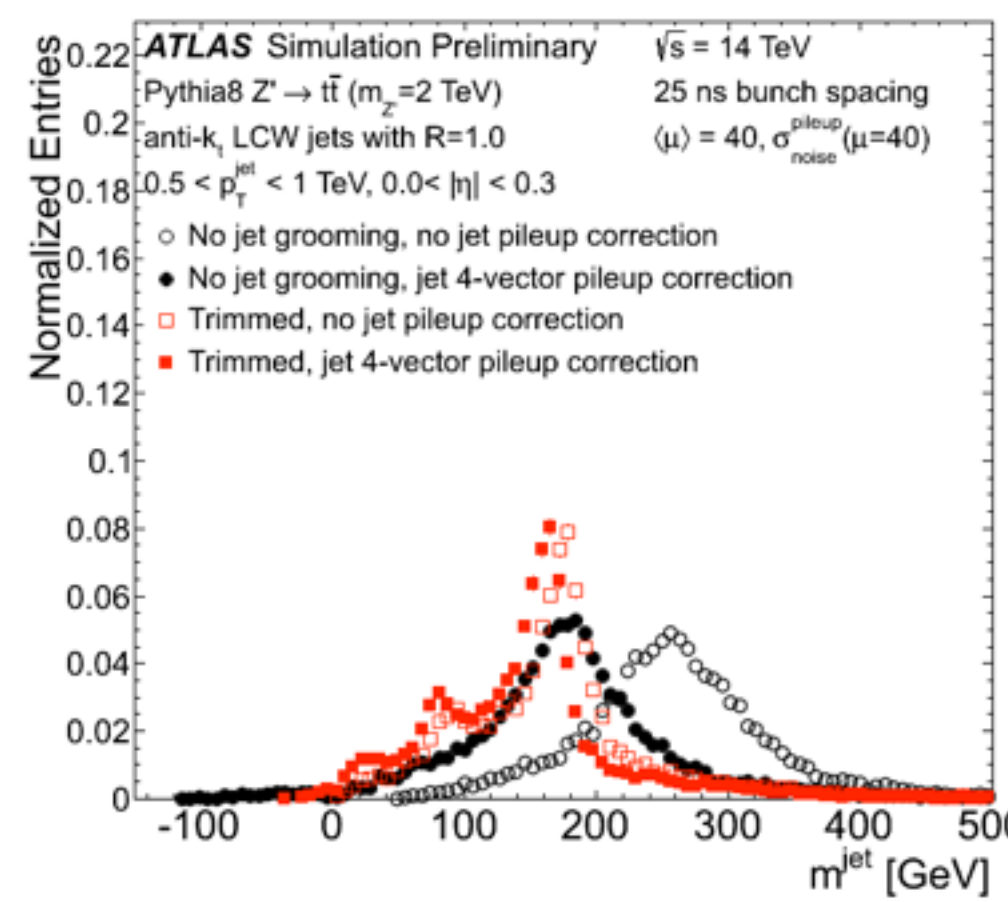
Pile-up correction I

Jet area based 4-vector correction (ATLAS-CONF-2013-083)
It only applies corrections to jet pT and jet mass.

$$p_{\mathrm{T}}^{\mathrm{jet,corr}} = p_{\mathrm{T}}^{\mathrm{jet}} - \rho \cdot A \qquad \rho = \mathrm{median}\left\{\frac{p_{T,i}^{\mathrm{jet}}}{A_i^{\mathrm{jet}}}\right\} \qquad A_j = \frac{1}{\nu_g\langle g_t\rangle}\sum_{g_i \in j} g_i$$

where the index i enumerates the jets found when clustering the event with the k_t algorithm
where $\nu_g\langle g_t\rangle$ is the transverse momentum density of the ghosts.

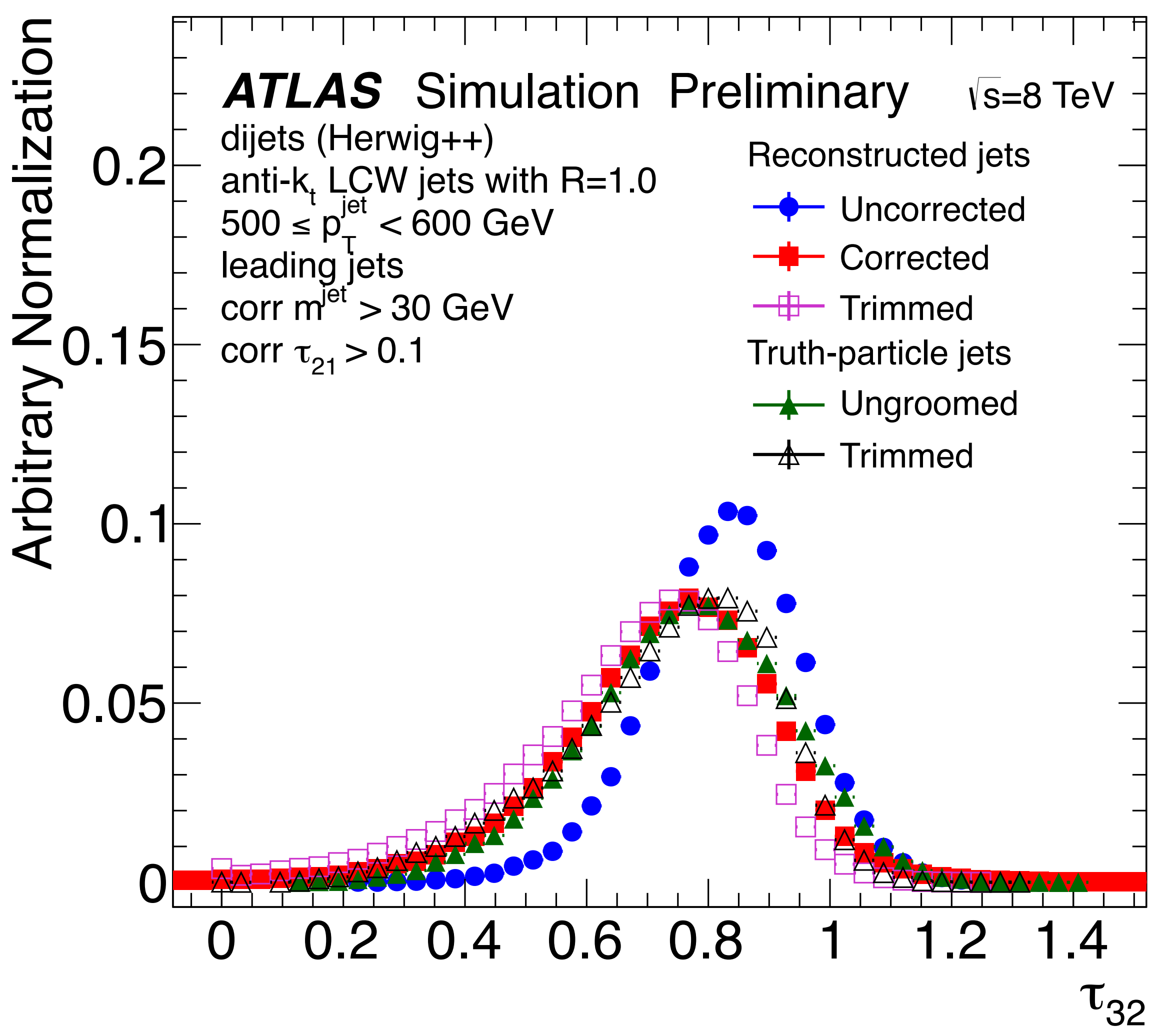


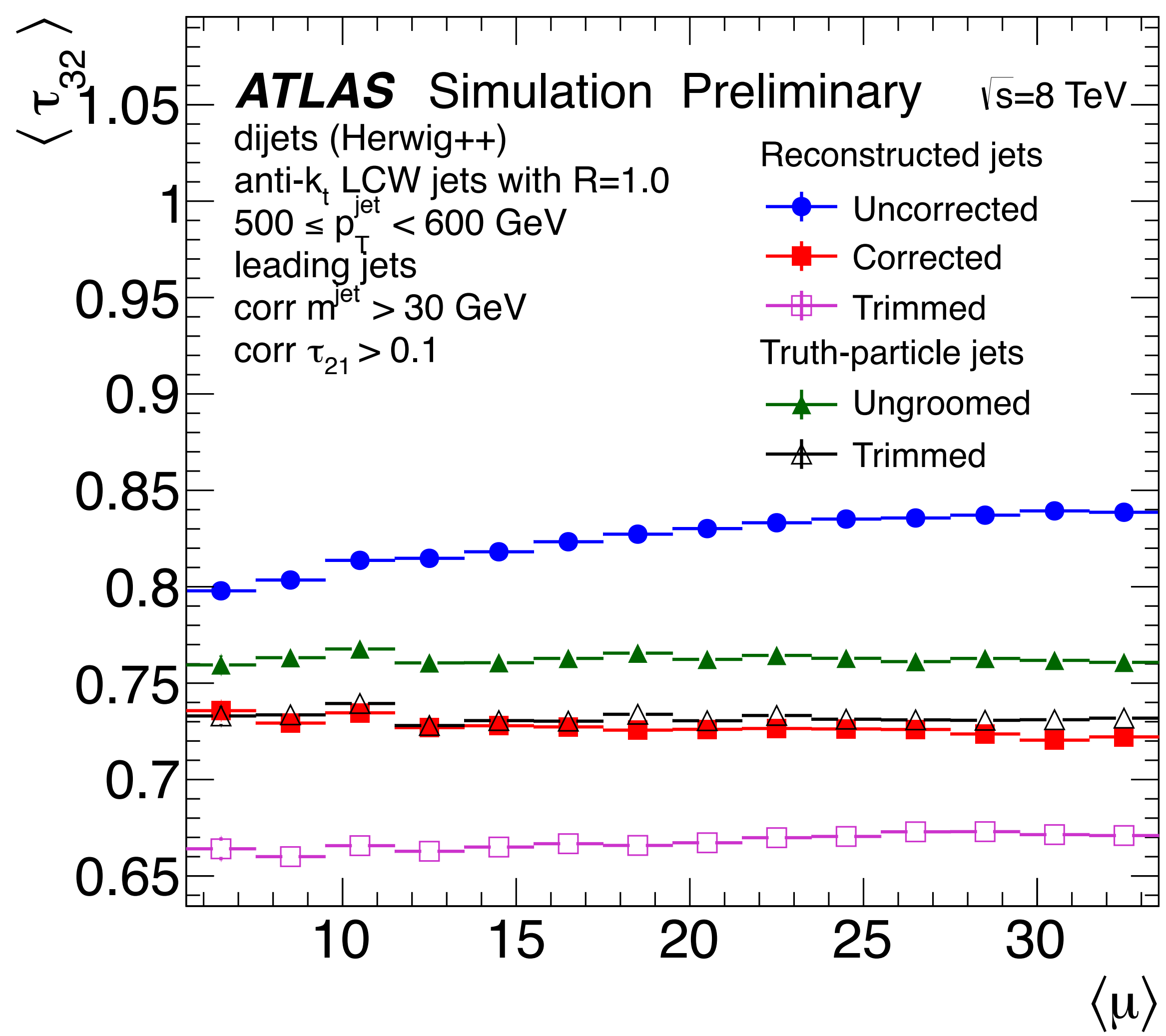


JetEtmissApproved2013HighMuSubstructure

Pile-up correction II

Do we study all jet shape variables with pile-up correction (ATLAS-CONF-2013-085)?
How to add systematic uncertainty?



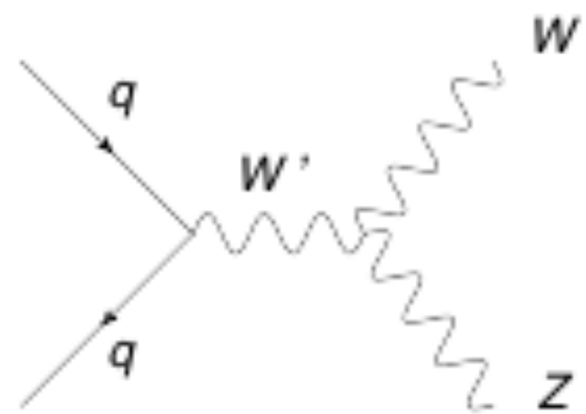


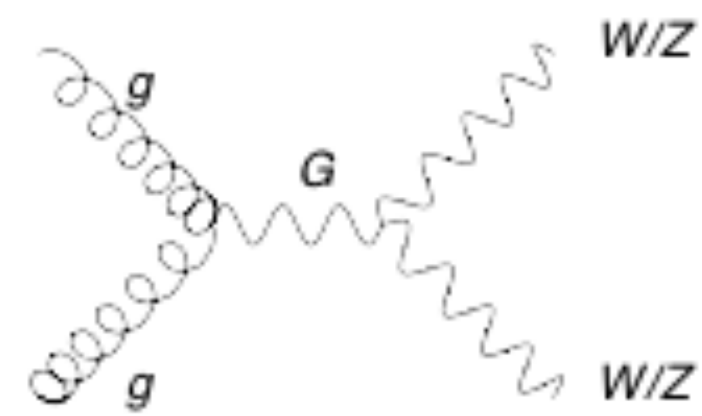
Motivation2

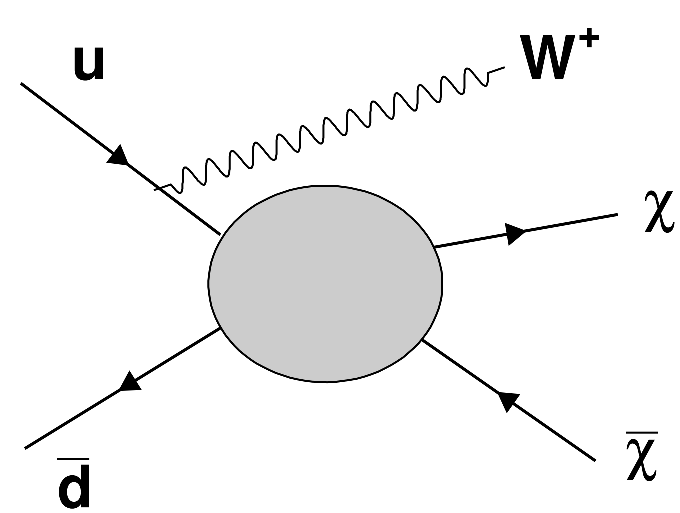
- mono Boson = Bosons + Met signature is predicted in many BSM models.

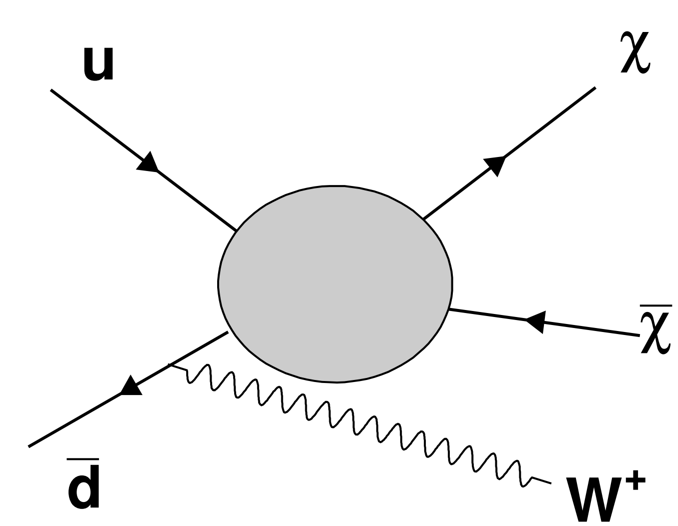
  - Warped Extra Dim. G→ZZ(ν ν)
  - heavy H→ZZ(ν ν)
  - SSM W'→WZ(ν ν)
  - VH(→ invisible)
  - DM V$\chi\chi$

  ATLAS, PRL 112, 041802 (2014)



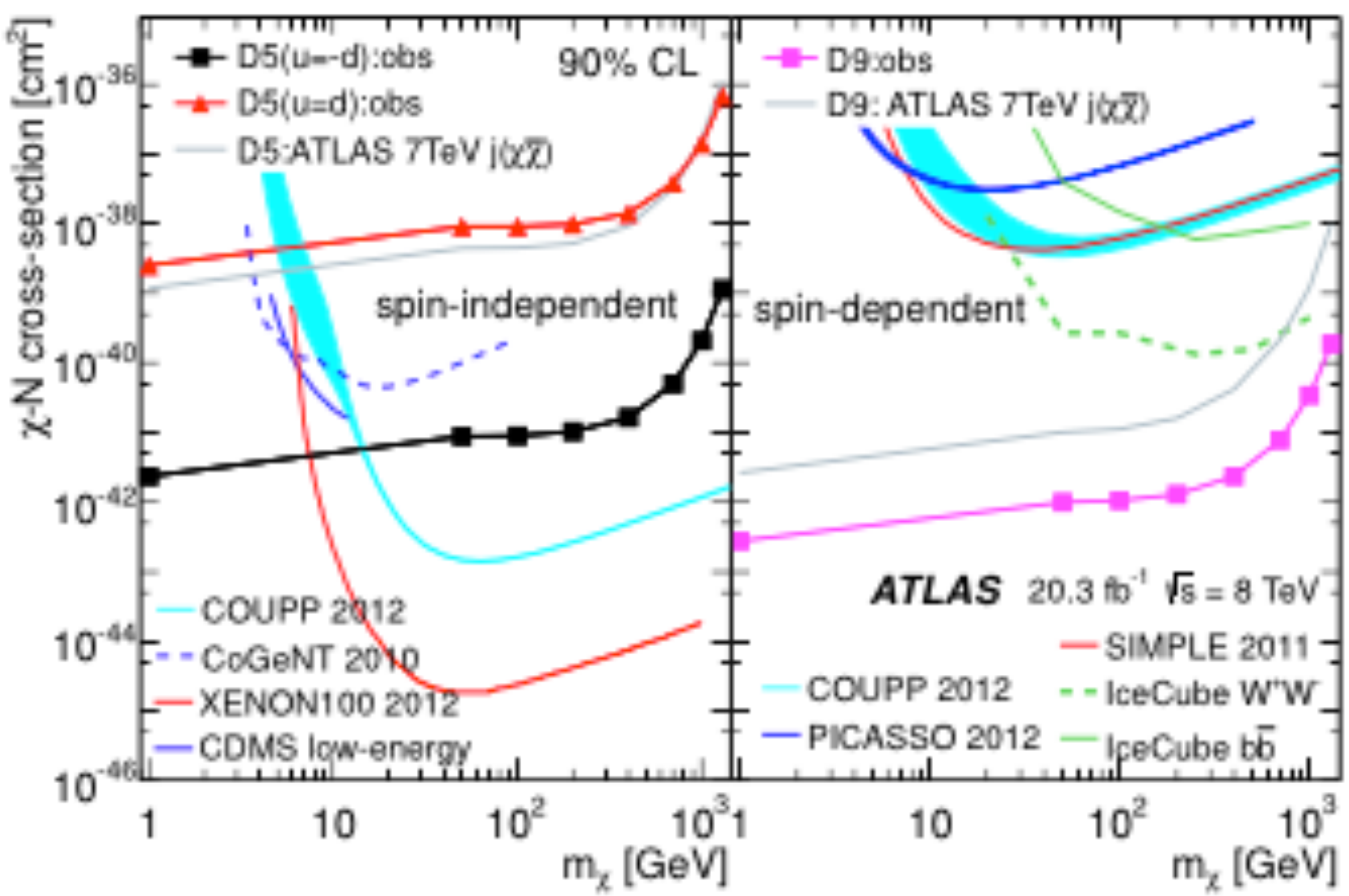


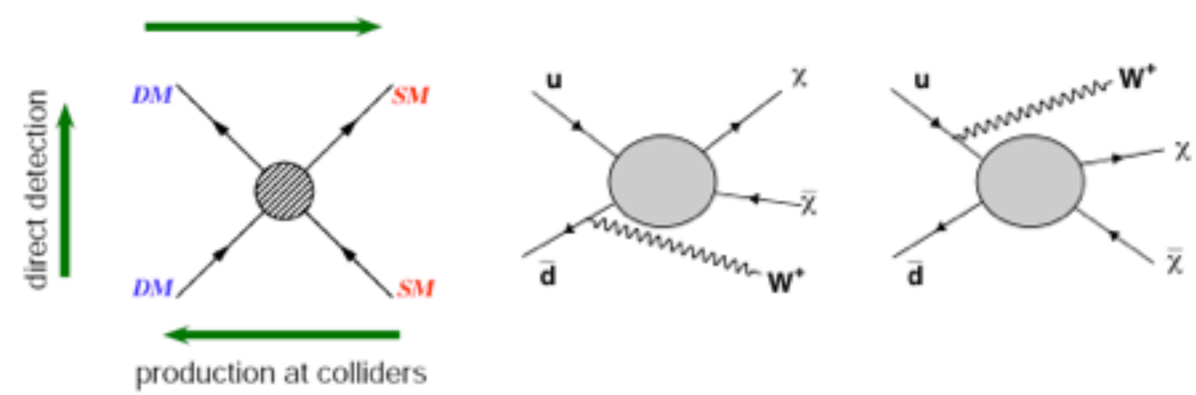


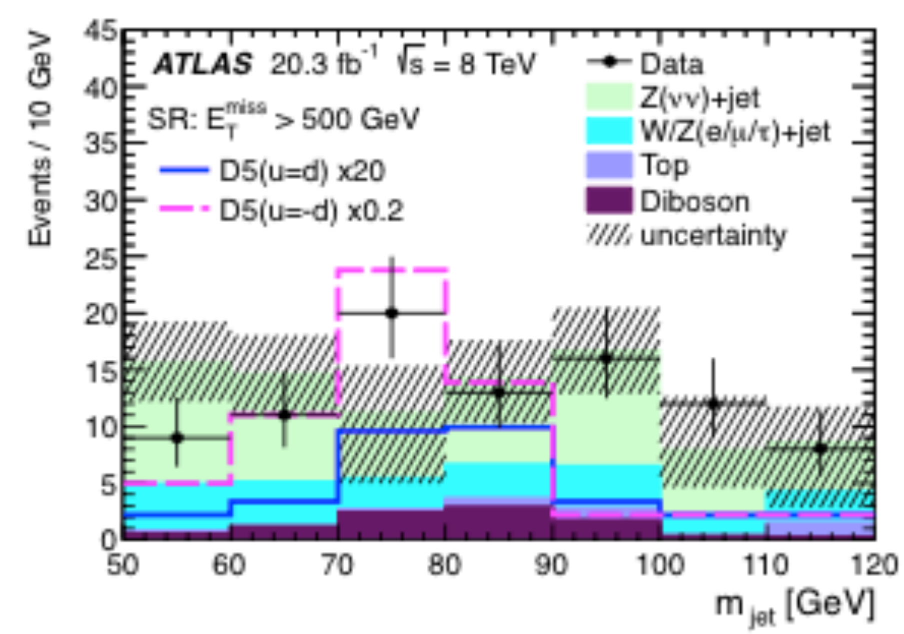


monoV in Run I

- Reconstruct W/Z with central large-R jet, pT > 250GeV, mJet[50,120 GeV], ETmiss > 350, 500 GeV
- Sensitive to the sign of DM couplings to up/down quarks
  Unique for mono-W
- WIMP-Nucleon reinterpretation complements (in)direct search

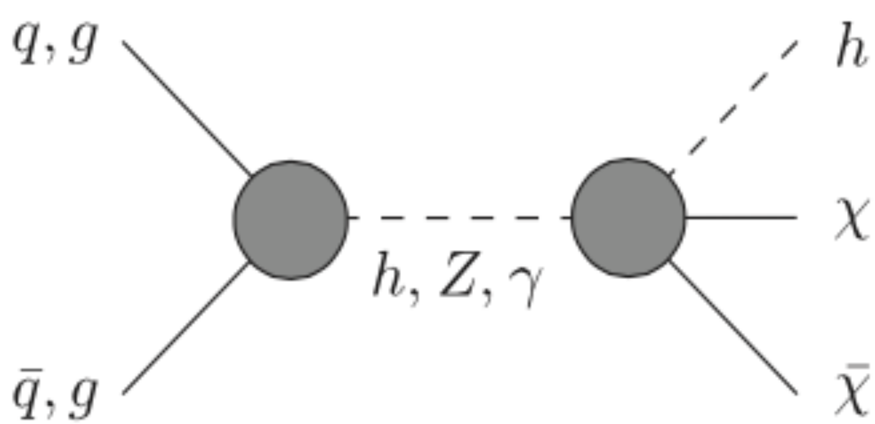






$$\sigma^{D5,C(u)=C(d)} = 1.38\times10^{-37}\mathrm{cm}^2\left(\frac{\mu_\chi}{1\ \mathrm{GeV}}\right)^2\left(\frac{300\ \mathrm{GeV}}{M_*}\right)^4$$

monoH EFT

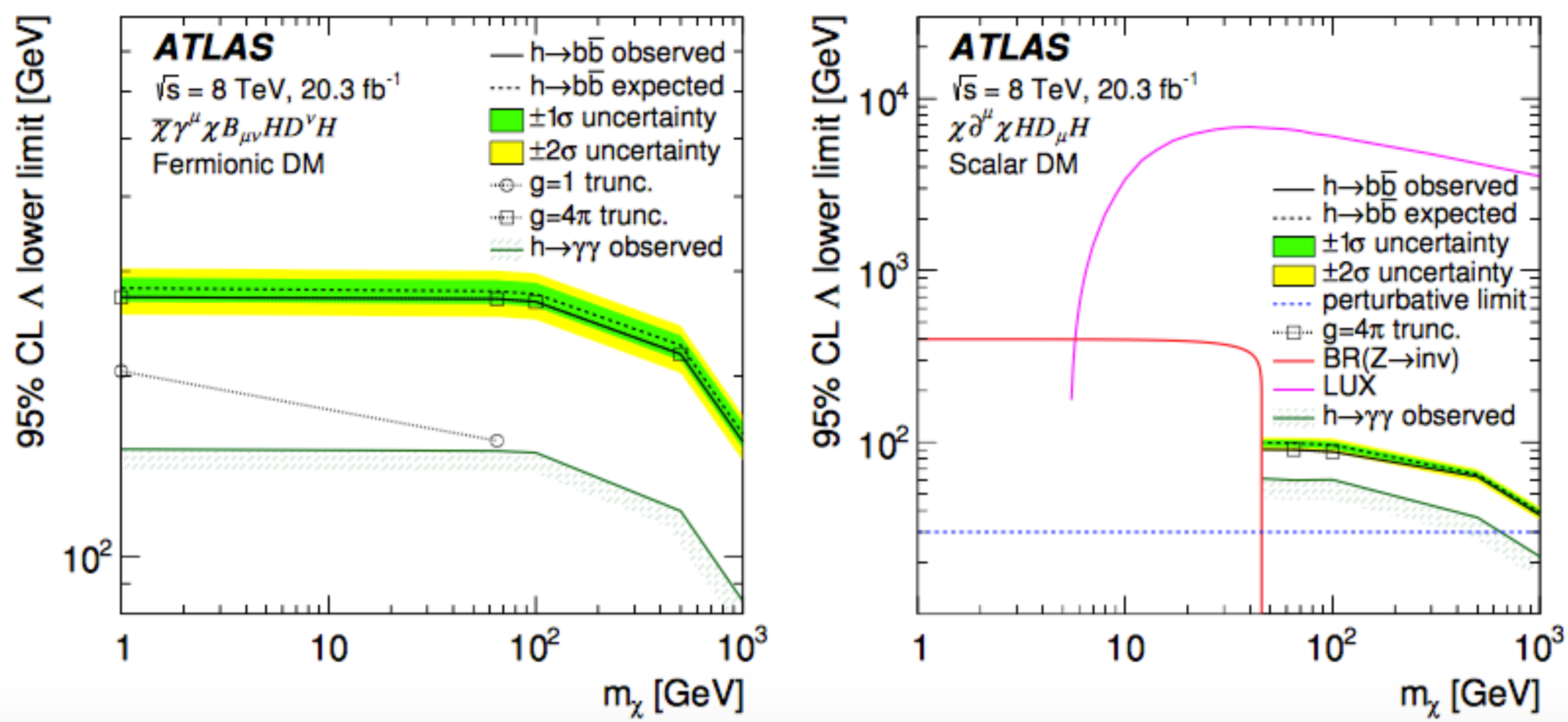


$$\lambda|\chi|^2|H|^2 \qquad \text{(Scalar DM, dimension-4)}$$

$$\frac{1}{\Lambda}\bar{\chi}i\gamma_5\chi|H|^2 \qquad \text{(Fermionic DM, dimension-5)}$$

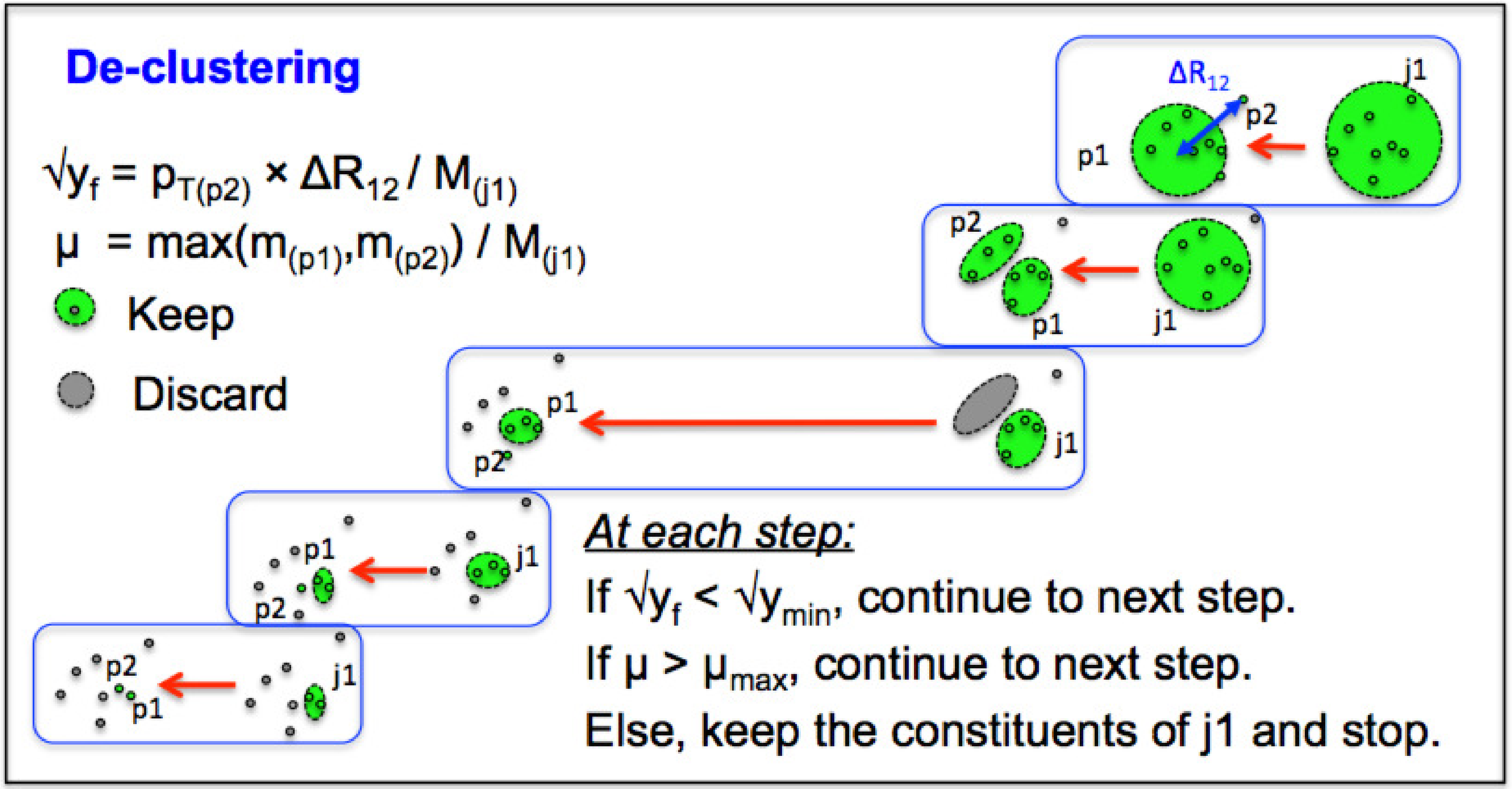
$$\frac{1}{\Lambda^2}\chi^\dagger\partial^\mu\chi H^\dagger D_\mu H \qquad \text{(Scalar DM, dimension-6)}$$

$$\frac{1}{\Lambda^4}\bar{\chi}\gamma^\mu\chi B_{\mu\nu}H^\dagger D^\nu H \qquad \text{(Fermionic DM, dimension-8)}$$



W-tagging in VV

- Jet reconstruction:
  All diboson searches presented here use C/A R = 1.2 jets groomed with the BDRS (split-filtering) algorithm
- Splitting:
  - Require symmetric splitting sqrt(yf) = min(pT1,pT2)/m12 × ΔR12 with sqrt(ymin) = 0.2
  - No mass drop-criterion is used in ATLAS (μ = 100%)
  - Slightly modified version of BDRS using a fixed reclustering distance parameter
- Filtering: remove soft radiation

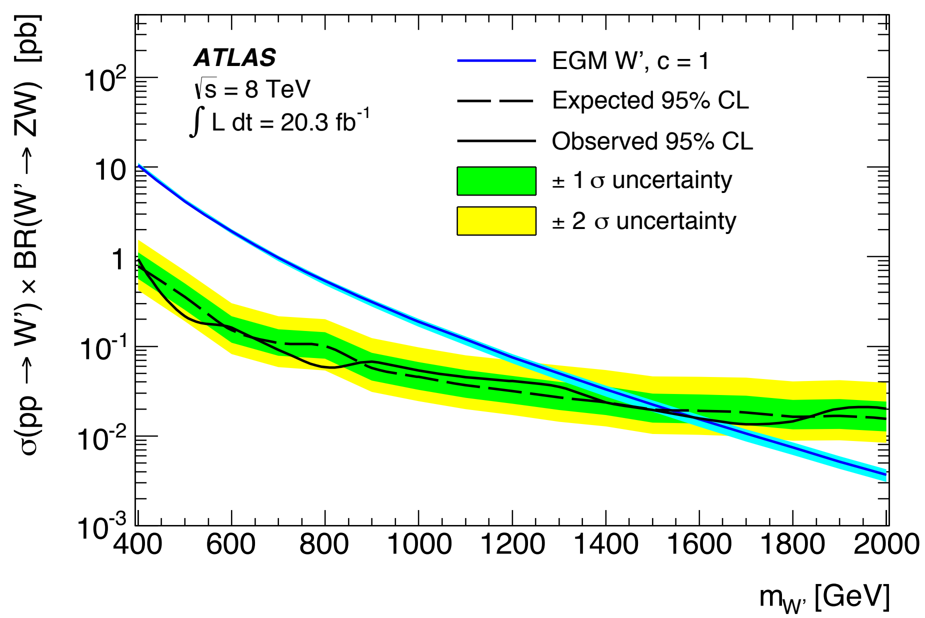


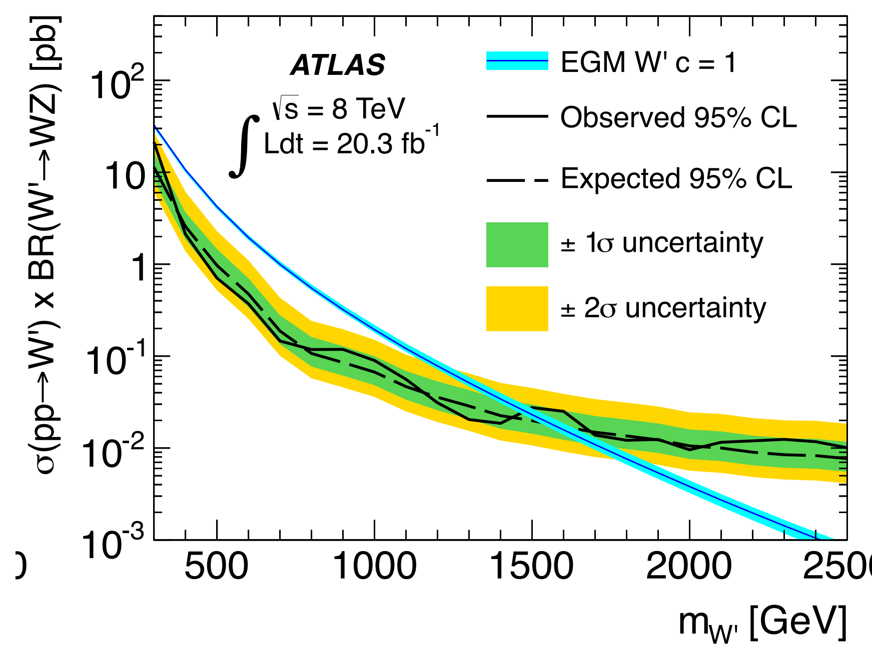
VV in Run I

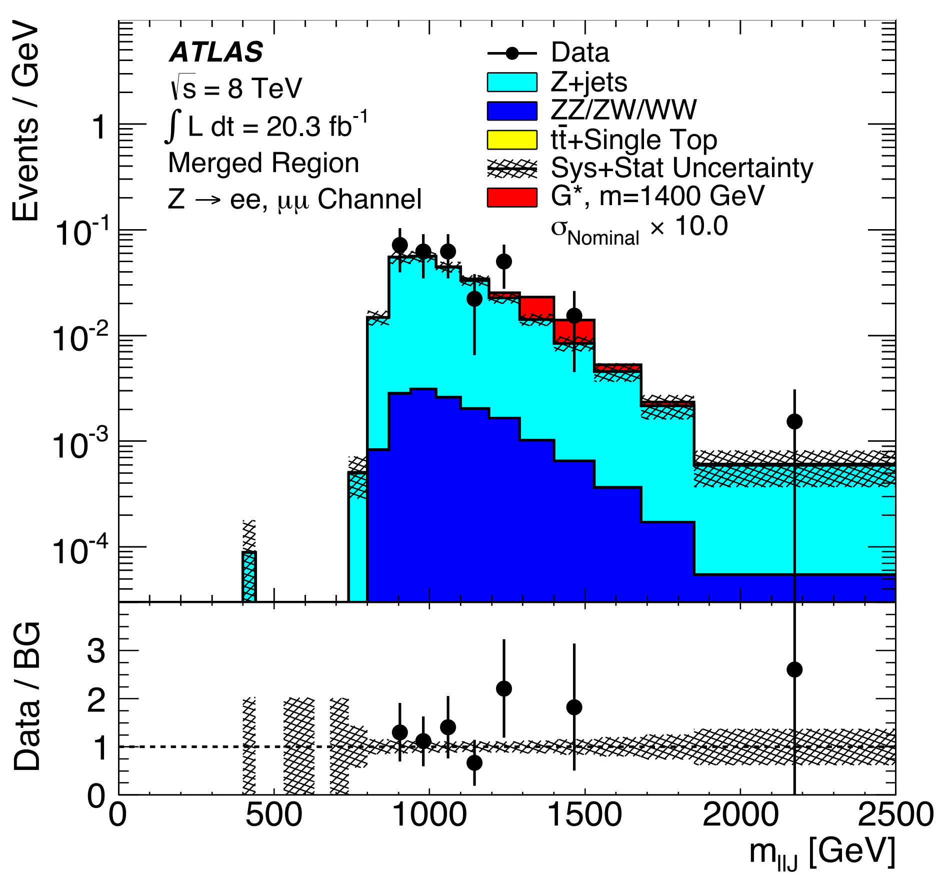
No excess in llqq/lnuqq channel

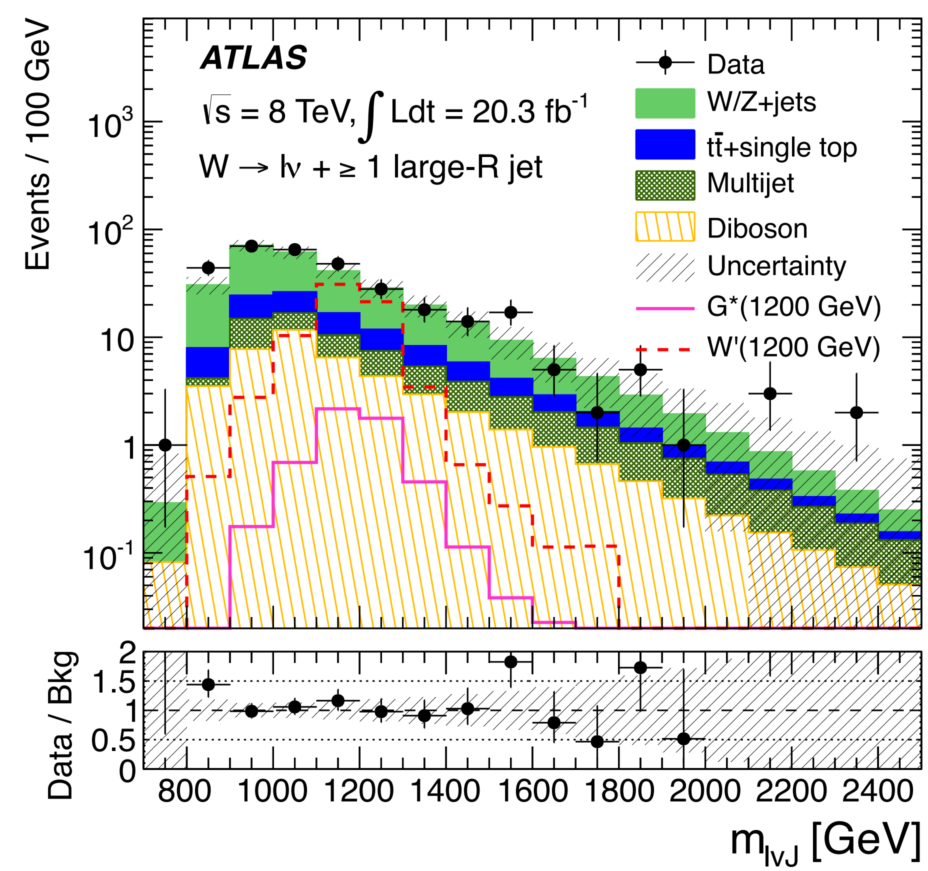
Eur. Phys. J. C (2015) 75:69

Eur. Phys. J. C (2015) 75:209









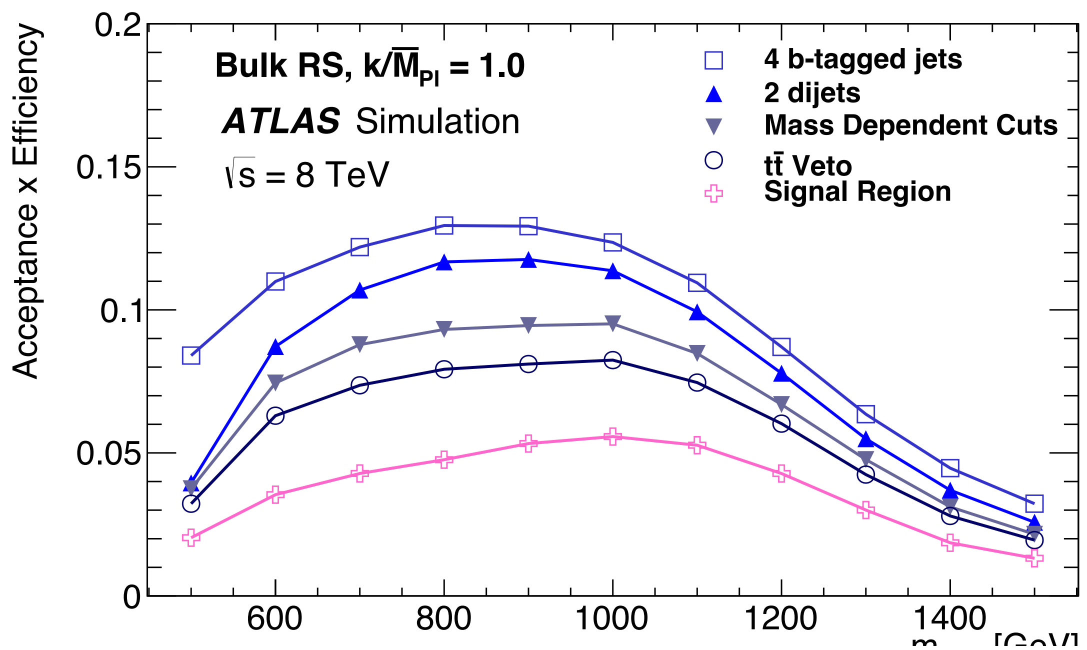
hh overview

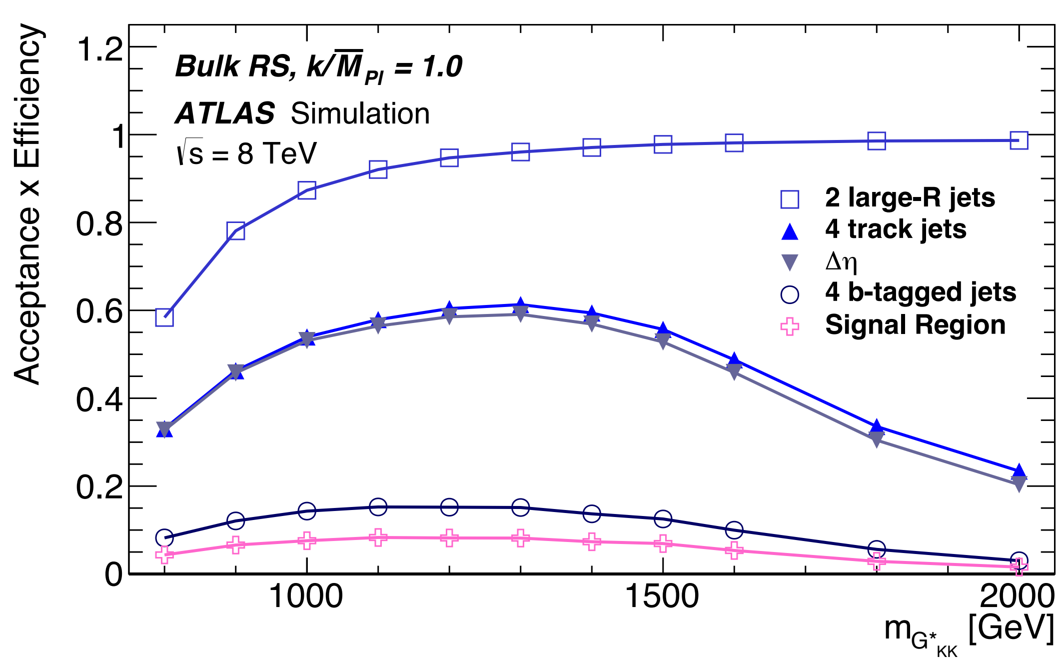
Resolved analysis:

1. Four b-tagged anti-kT R=0.4 jets
2. Arrange into close-by pairs, ΔR<1.5
3. Mass dependent p_T and |Δη| cuts
4. ttbar veto, using 5^{th} jet to test consistency with m_W / m_{top}

Boosted analysis:

1. Two anti-kT R=1.0 jets, trimmed with R_{sub}=0.3 and f_{cut}=0.05
2. Each with 2 b-tagged R=0.3 track jets
3. p_T and |Δη| cuts





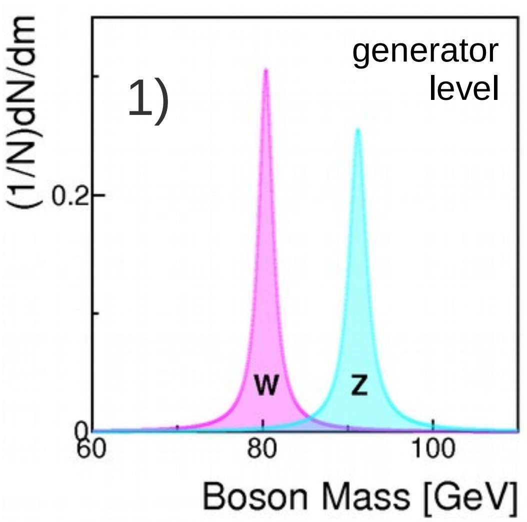
Signal region: $X_{hh} = \sqrt{\left(\frac{m_{\mathrm{lead}} - 124}{0.1 \times m_{\mathrm{lead}}}\right)^2 + \left(\frac{m_{\mathrm{subl}} - 115}{0.1 \times m_{\mathrm{subl}}}\right)^2} < 1.6$

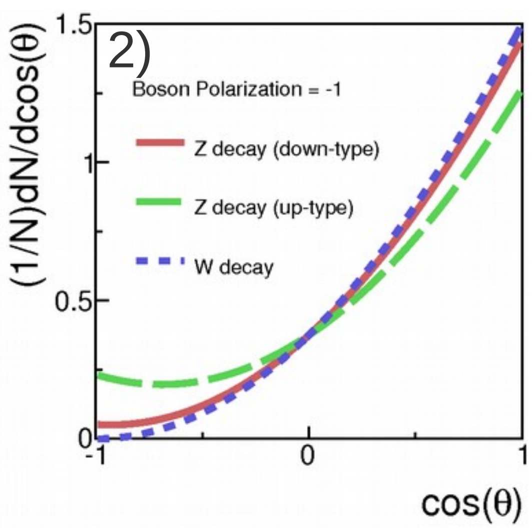
W/Z Sepration

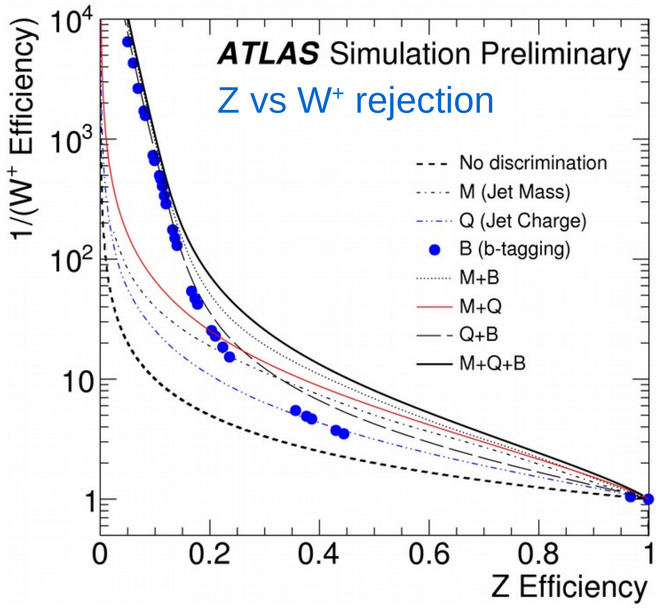
$$\Pr(M,Q,B|V) = \sum_f \Pr(f|V)\,\Pr(M|f,V)\,\Pr(Q|f,V)\,\Pr(B|f,V)$$

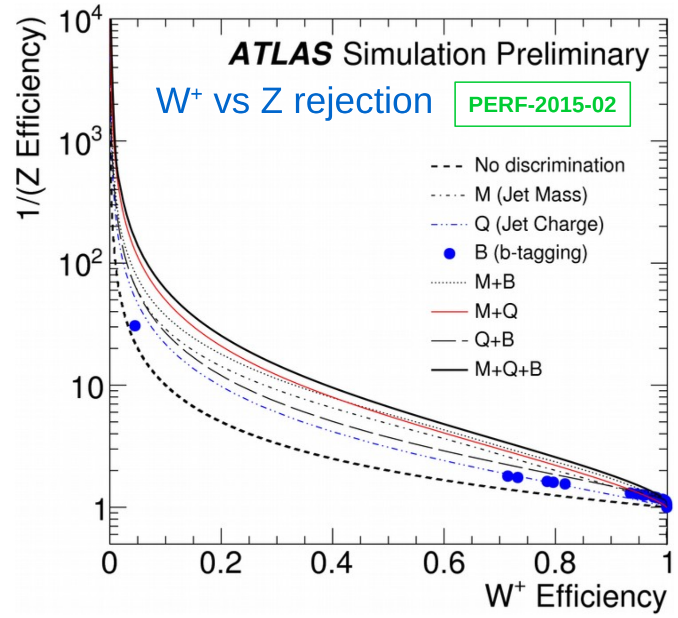
with $p(B|f,V) = p(B_{\text{lead}}|f,V)\,p(B_{\text{sub-lead}}|f,V)$

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

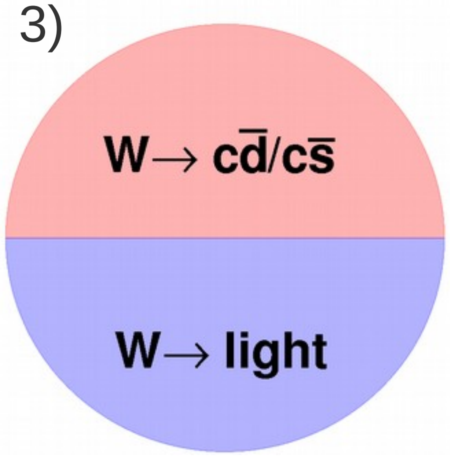


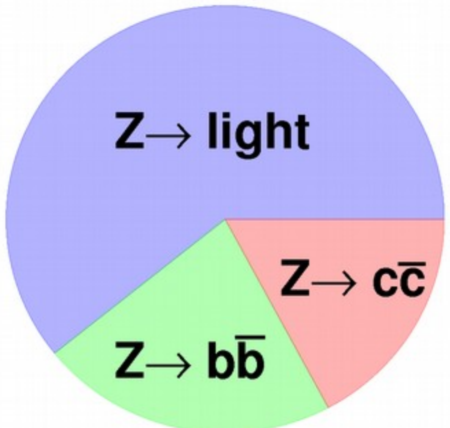




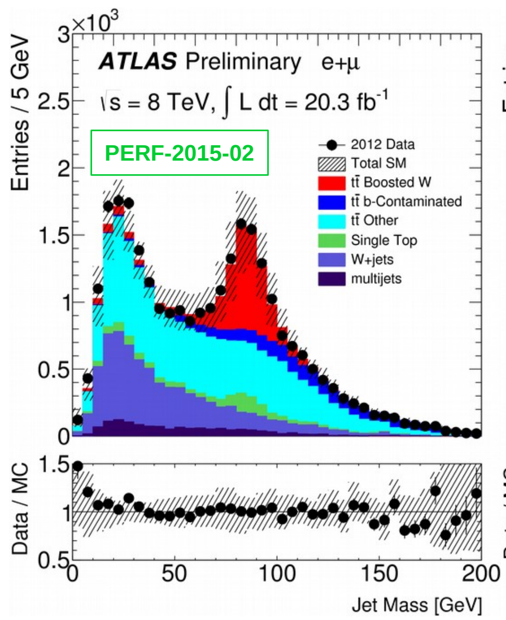


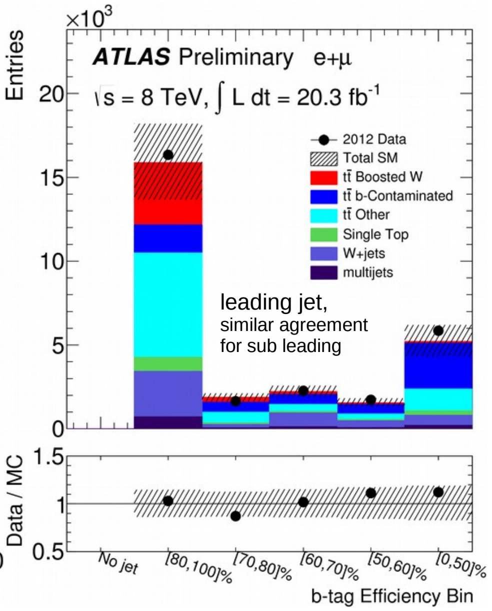
3)

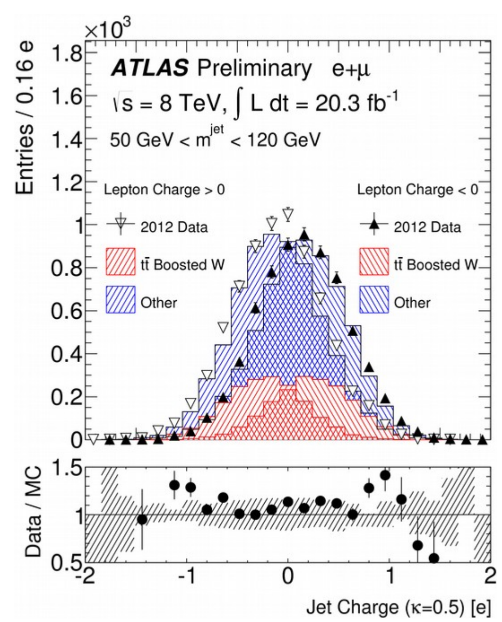




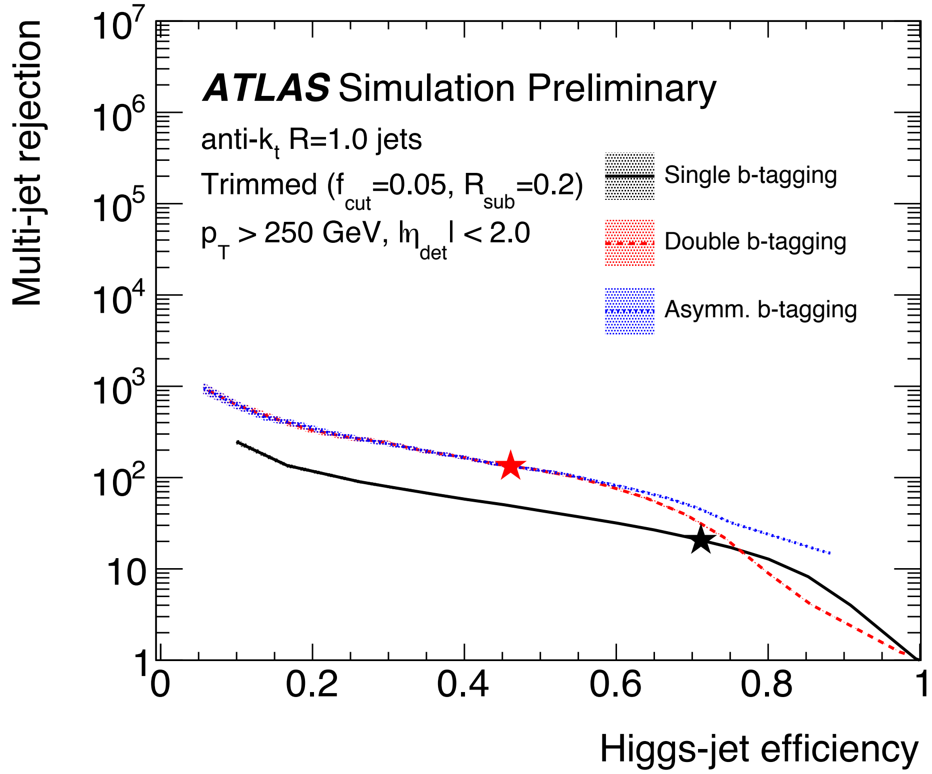
PERF-2015-02

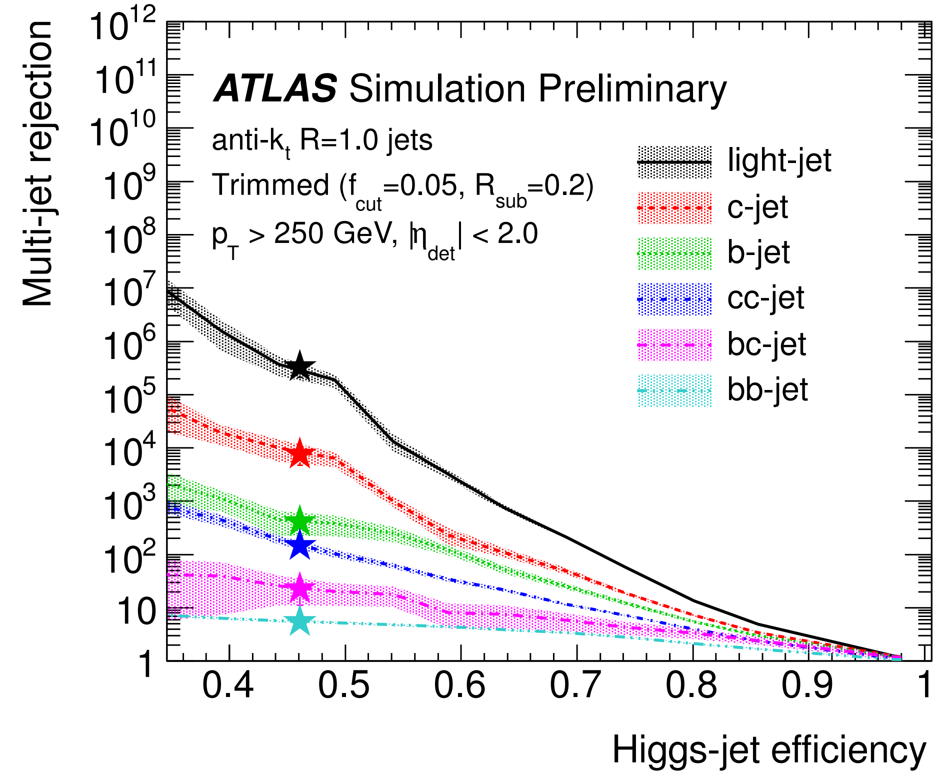


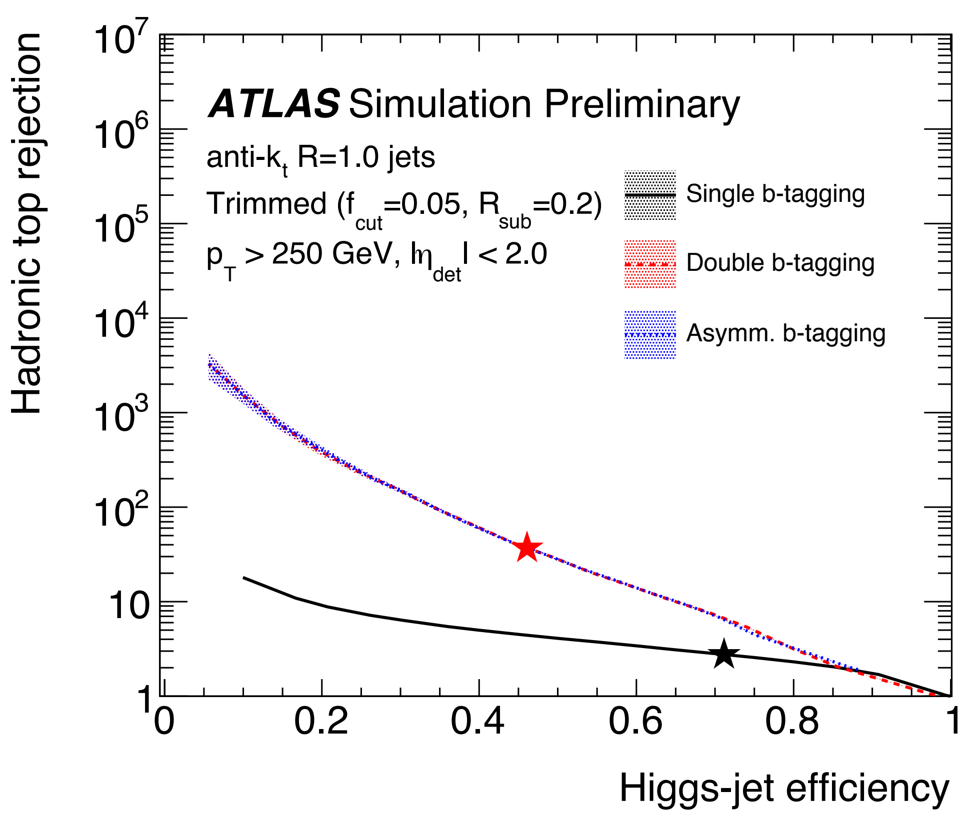


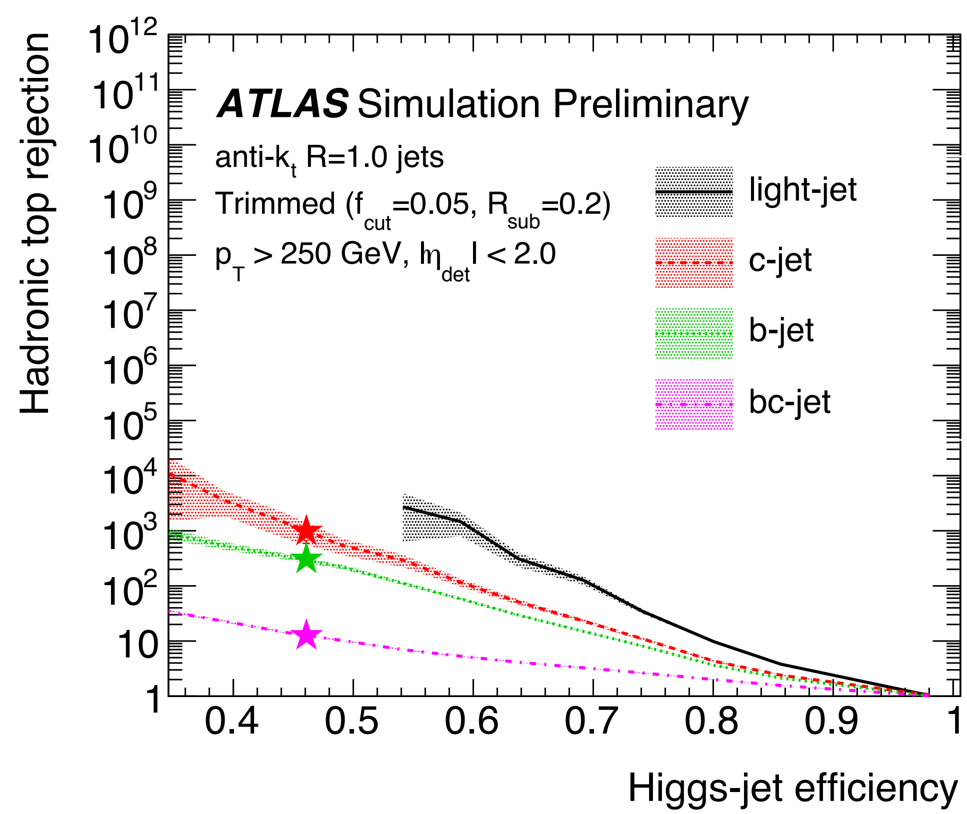


h-tagging


ATLAS Simulation Preliminary
anti-k_t R=1.0 jets
Trimmed (f_{cut}=0.05, R_{sub}=0.2)
p_T > 250 GeV, $|\eta_{det}|$ < 2.0
Single b-tagging
Double b-tagging
Asymm. b-tagging
Multi-jet rejection
Higgs-jet efficiency


ATLAS Simulation Preliminary
anti-k_t R=1.0 jets
Trimmed (f_{cut}=0.05, R_{sub}=0.2)
p_T > 250 GeV, $|\eta_{det}|$ < 2.0
light-jet
c-jet
b-jet
cc-jet
bc-jet
bb-jet
Multi-jet rejection
Higgs-jet efficiency


ATLAS Simulation Preliminary
anti-k_t R=1.0 jets
Trimmed (f_{cut}=0.05, R_{sub}=0.2)
p_T > 250 GeV, $|\eta_{det}|$ < 2.0
Single b-tagging
Double b-tagging
Asymm. b-tagging
Hadronic top rejection
Higgs-jet efficiency


ATLAS Simulation Preliminary
anti-k_t R=1.0 jets
Trimmed (f_{cut}=0.05, R_{sub}=0.2)
p_T > 250 GeV, $|\eta_{det}|$ < 2.0
light-jet
c-jet
b-jet
bc-jet
Hadronic top rejection
Higgs-jet efficiency

h-tagging

Largest uncertainty from flavor tagging

	Loose	Medium	Tight
efficiency	0.41 ±0.07	0.32 ±0.06	0.25 ±0.05
Multi-jet rejection			
Inclusive	260 ±50	460 ±90	800 ±210
Light-flavor	$\mathcal{O}(10^5)$	$\mathcal{O}(10^5)$	$\mathcal{O}(10^6)$
cl	$\mathcal{O}(10^3)$	$\mathcal{O}(10^3)$	$\mathcal{O}(10^4)$
bl	$\mathcal{O}(10^2)$	$\mathcal{O}(10^2)$	$\mathcal{O}(10^3)$
bc	$\mathcal{O}(10)$	$\mathcal{O}(10)$	$\mathcal{O}(10^2)$
cc	250 ±150	480 ±310	1200 ±900
bb	11 ±2	19 ±4	31 ±9
Hadronic top rejection			
Inclusive	67 ±17	110 ±30	160 ±50
bl	360 ±230	660 ±460	810 ±600
bc	24 ±6	39 ±11	53 ±16

muon-in-jet correction