H→bb Tagging in ATLAS

Qi Zeng

On behalf of the ATLAS Collaboration
BOOST, July 18-22, 2016

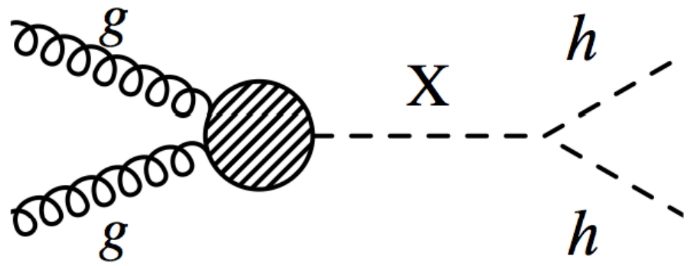
U.S. DEPARTMENT OF ENERGY
Office of Science

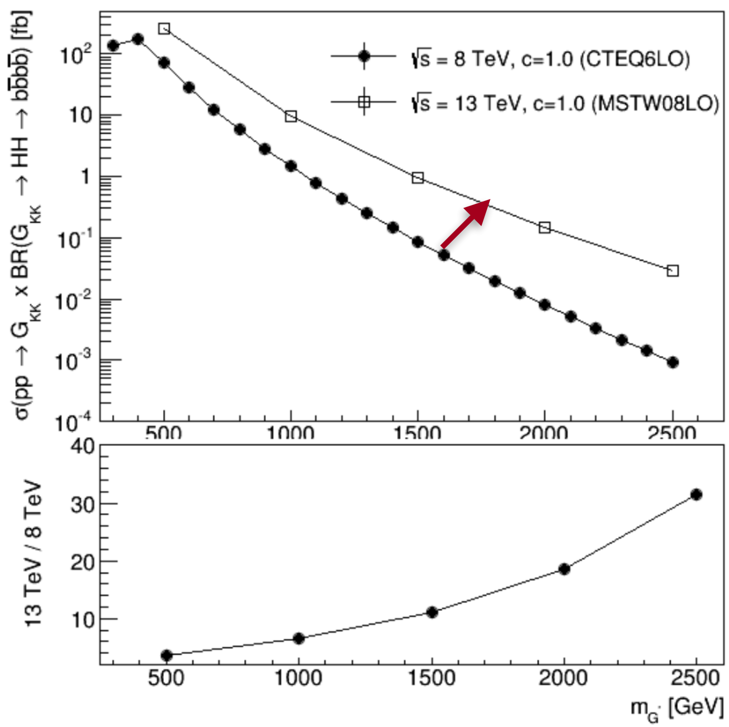
SLAC NATIONAL ACCELERATOR LABORATORY

Introduction

Motivation

- SM decay H→bb: largest branching ratio
- Boosted topology provides various handles to suppress dominant multi-jet background
- Boosted H→bb tagger becomes an essential tool for new physics search in ATLAS





Contents

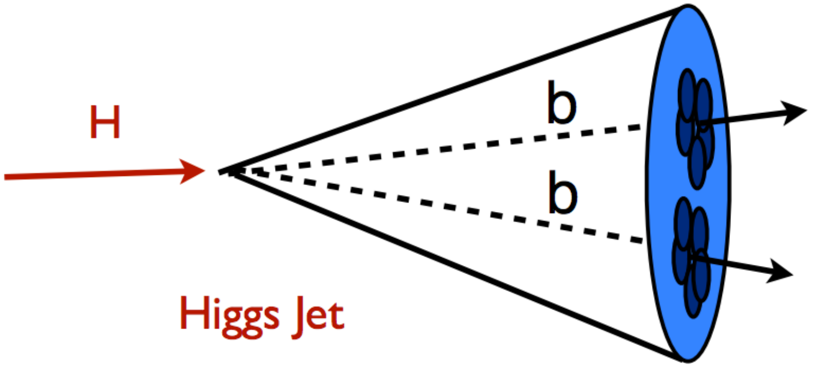
- Boosted H→bb tagger in Run-II
- Modeling in high-pT g→bb enriched sample using 8TeV data

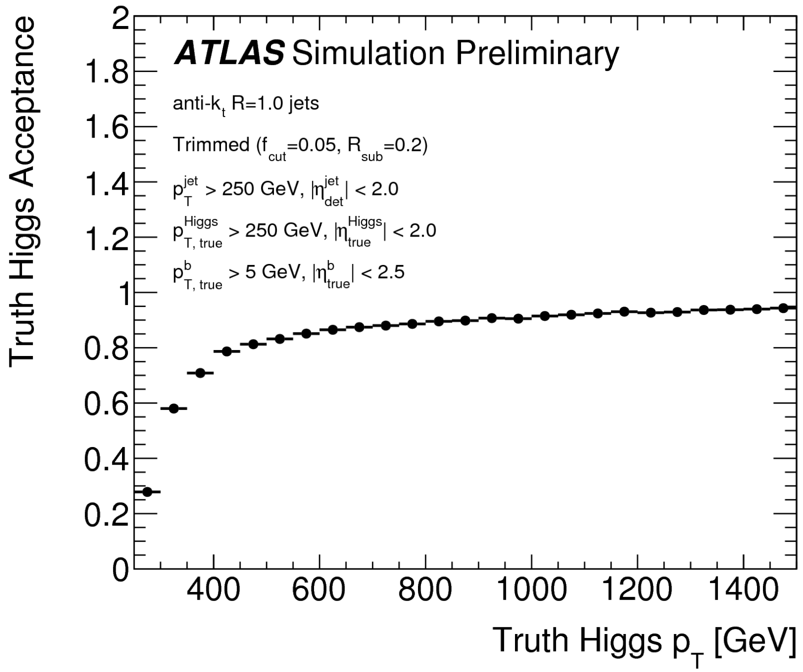
Boosted H→bb Tagging in Run-II

ATL-PHYS-PUB-2015-035

Tagging Boosted H→bb

- Reconstruct H→bb topology with trimmed anti-k_t R=1.0 jet (f_{cut}=0.05, R_{sub}=0.2)
- Major backgrounds considered:
  - Multi-jet events
  - Boosted hadronically decaying top quarks
- Three handles for background rejection:
  - *b*-Tagging
  - Large-R jet mass
  - Large-R jet substructure

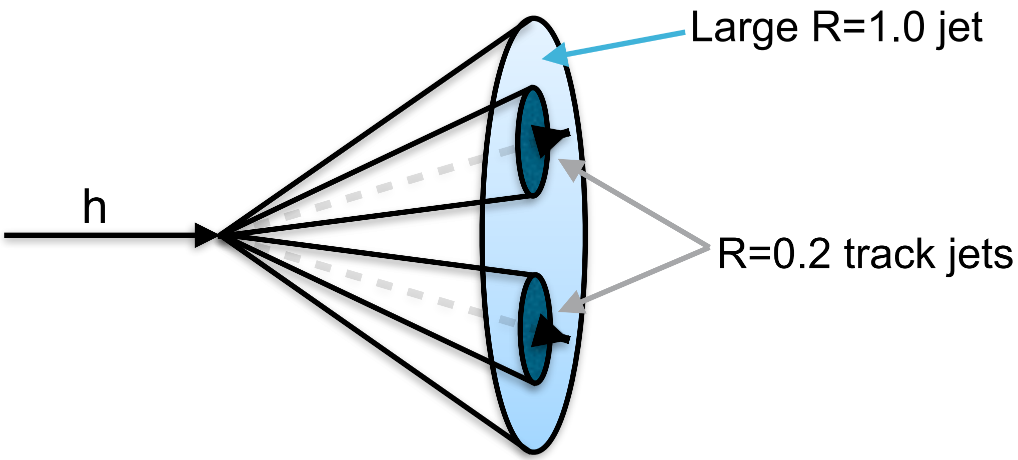


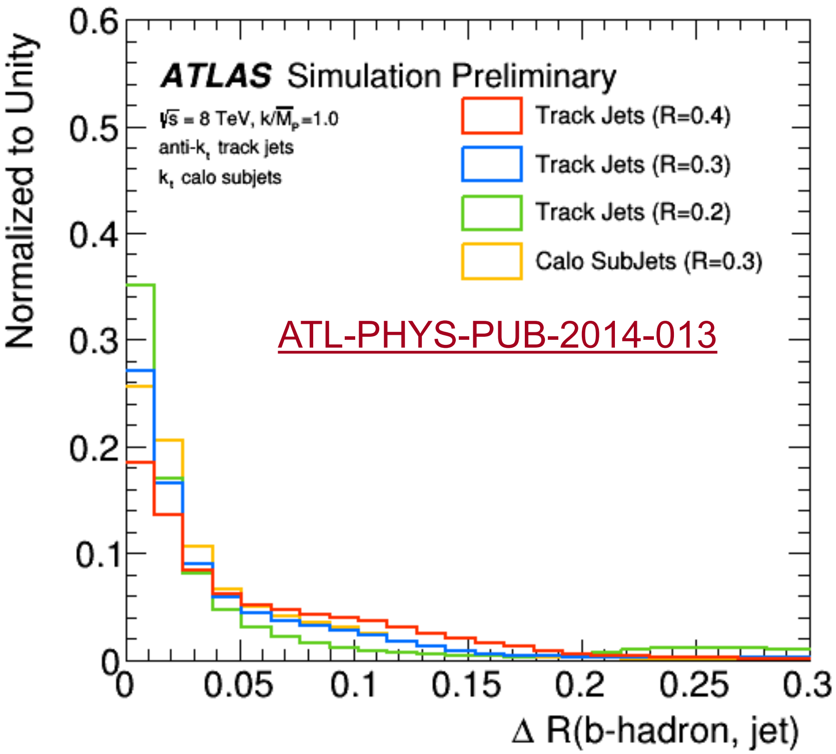


Small-R Track Jet *b*-Tagging

- Use small radius (R=0.2) track jets to resolve close-by b-hadrons
- Ghost association of track jets to ungroomed large-R jets to provide *b*-tagging

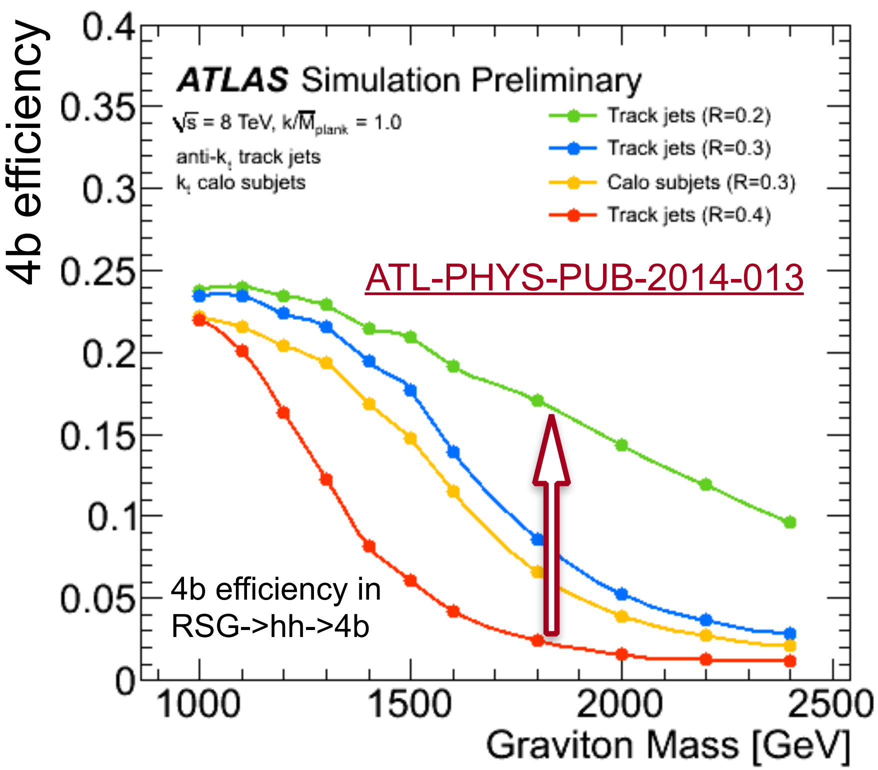
Advantage of track jets
- Better estimate *b*-hadron flight direction
- Pile-up resistant
- *b*-tagging independent of calorimeter jets

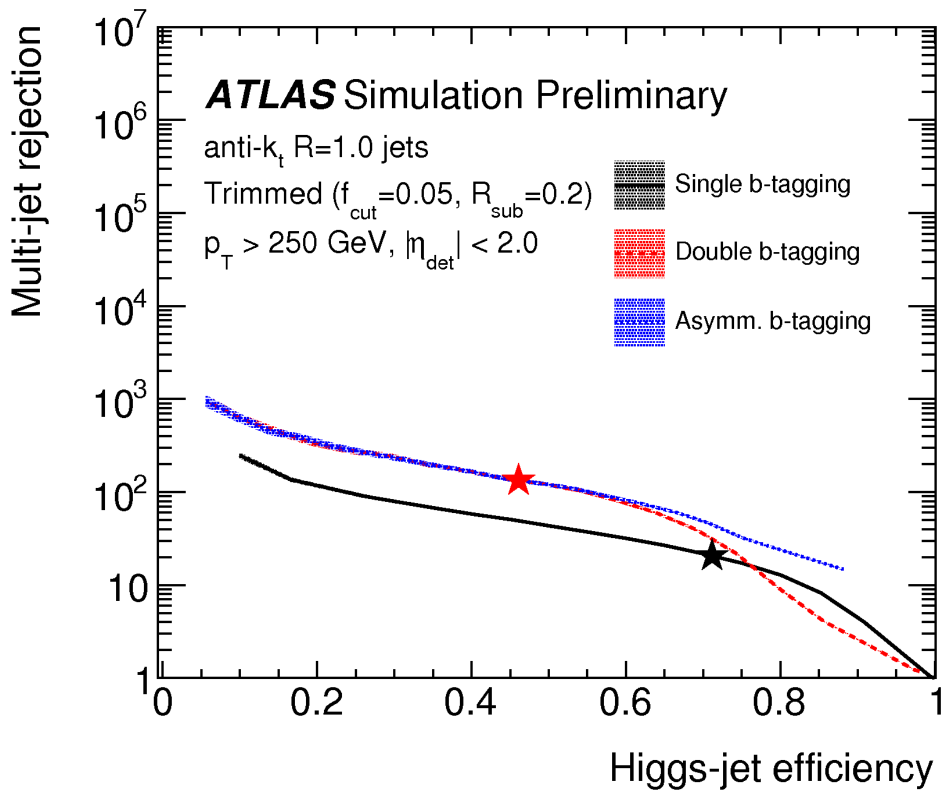




Small-R Track Jet *b*-Tagging

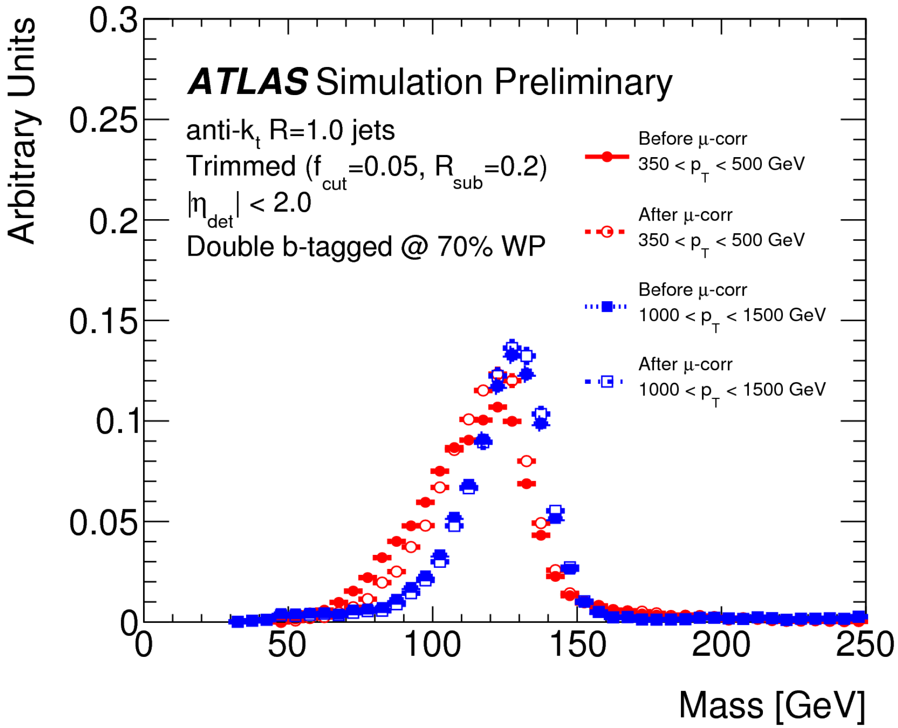
- Large improvement in efficiency to find boosted Higgs jet from small radius
- Flexible in track jet *b*-tagging
  - Independent of calorimeter jets
  - Single / double / one-tight-one-loose schemes

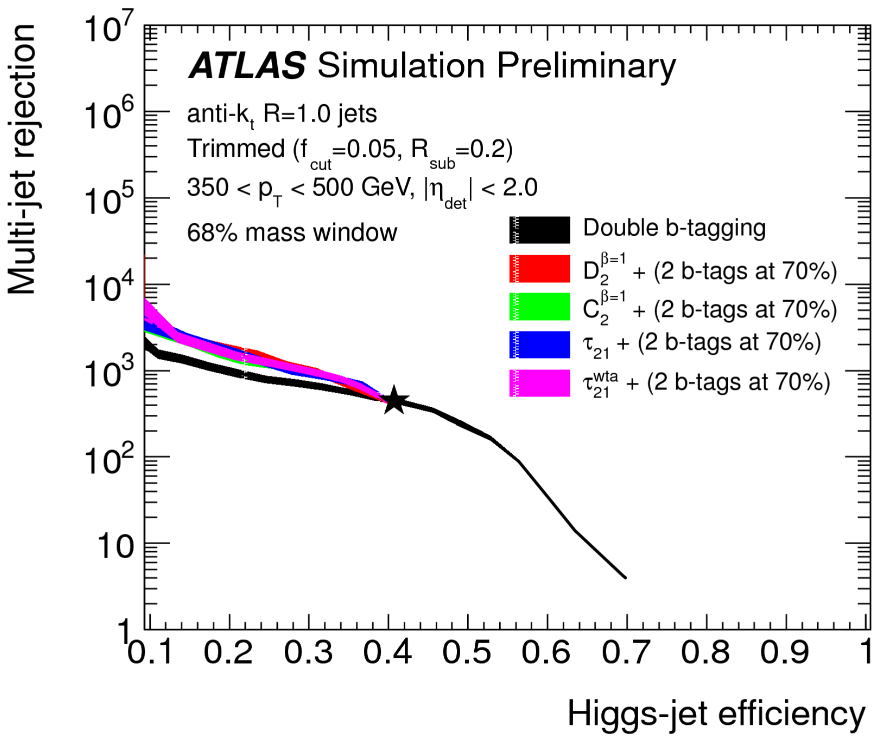




Large-R Jet Mass and Substructure

- Improve large-R jet mass resolution by:
  - Trimming with $f_{cut} = 0.05, R_{sub} = 0.2$
  - Muon-in-b-jet correction correcting for semi-leptonic b hadron decays
  - Variable-R jets(ATL-PHYS-PUB-2016-013): More details in talks from A. Dattagupta and C. Anders
- Substructure information considered in addition to mass cut and b-tagging:
  - Similar performance across $D_2^{\beta=1}, C_2^{\beta=1}$ and τ_{21}^{wta}
  - $D_2^{\beta=1}$ is chosen due to better modeling in data



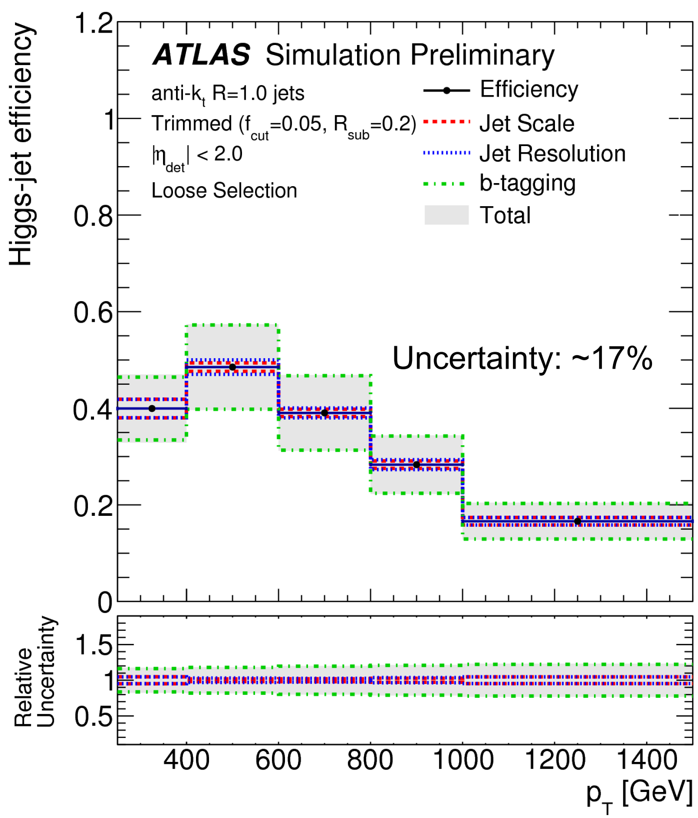


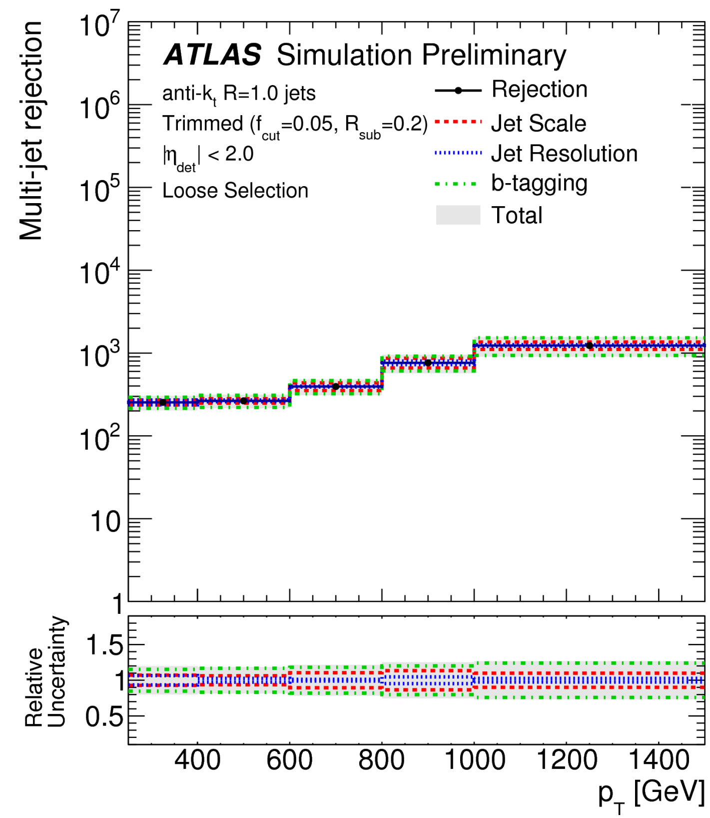
H→bb Tagger Performance

- Three working points (WP) defined
- Systematic uncertainties:
  - *b*-tagging largest for loose selection
  - Jet energy/mass scale & resolution larger for tight selection

Selection	double *b*-tagging	large-*R* jet Mass	$D_2^{(\beta=1)}$
Loose	70% WP	90% window, $m \in [76, 146]$ GeV	-
Medium	70% WP	68% window, $m \in [93, 134]$ GeV	-
Tight	70% WP	68% window, $m \in [93, 134]$ GeV	p_T-dependent cut

Table 1: Criteria used for the different Higgs-jet tagging selections.



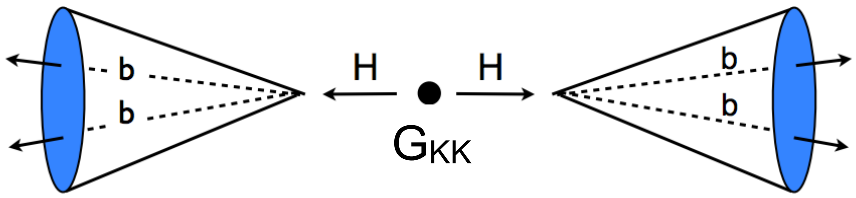


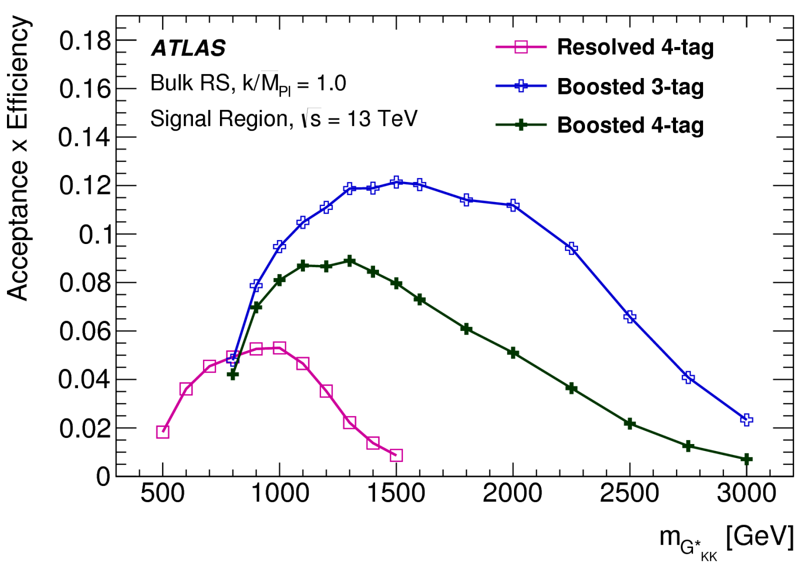
Application in 13 TeV New Physics Search

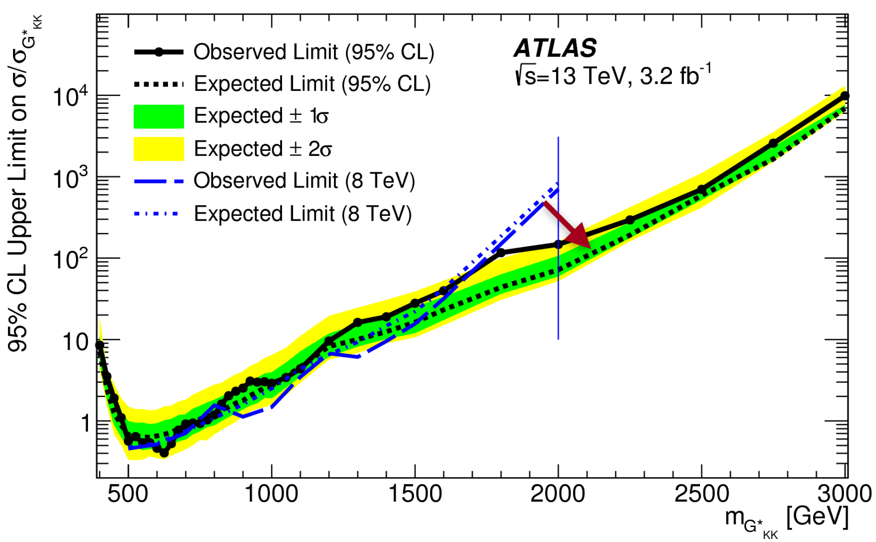
Higgs tagger has been widely used in ATLAS 13TeV analysis, in particular the new physics search:

- A selected list of examples here
- More details in C. Pollard's talk on Thursday

Di-Higgs 4b [EXOT-2015-11]:
- Require 3 or 4 *b*-tagged track jets
- 77% WP is used
- Customized jet mass requirement

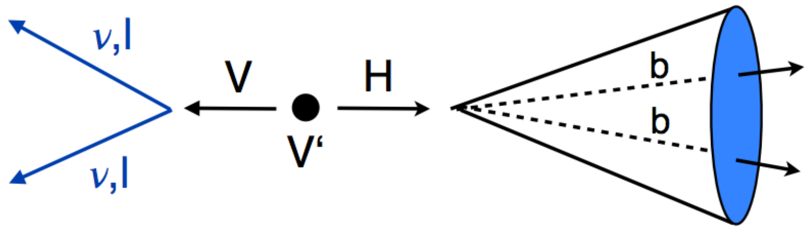




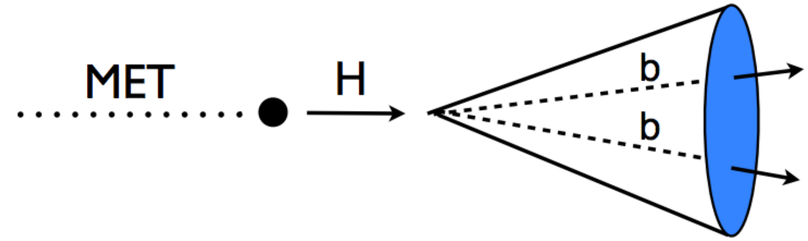


Application in 13 TeV New Physics Search (Continued)

VH resonance search [ATLAS-CONF-2015-074]:
- Loose WP
- Additional 1 *b*-tag WP



mono-Higgs [ATLAS-CONF-2016-019]:
- 1 or 2 *b*-tagged track jets
- Full mass distribution used in likelihood fit

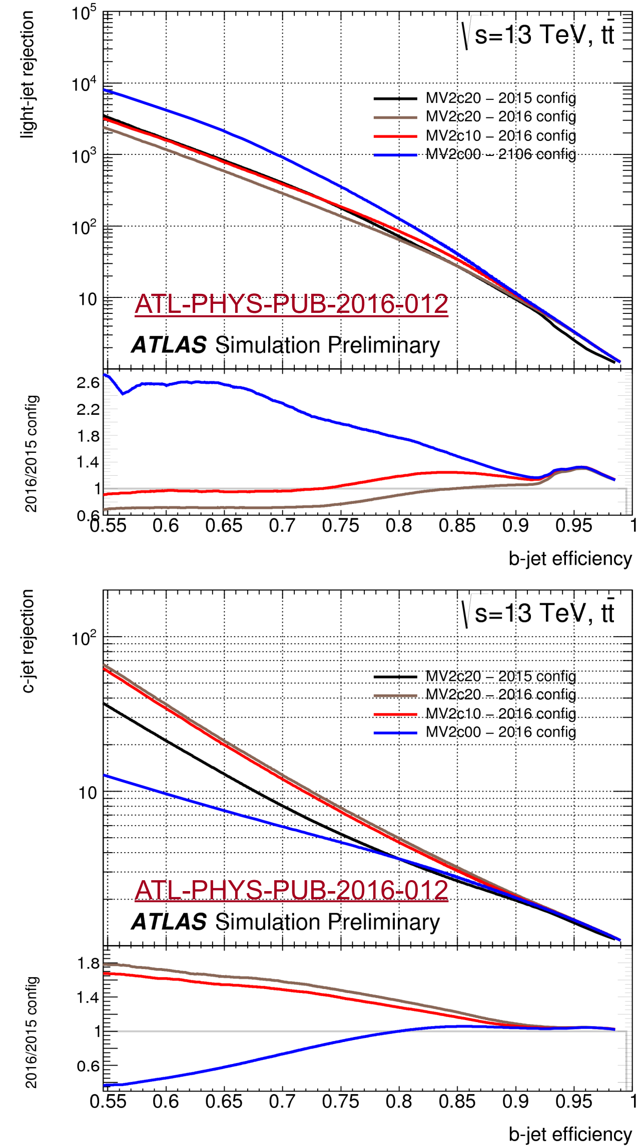


Lessons learnt from analysis:

- *b*-Tagging contributes the most to sensitivity
- Analysis benefits more from efficiency than rejection in boosted topology
  - Alternative: Single *b*-tagging (+ substructure)
  - Alternative: Looser *b*-tagging WP (especially high pT)

Updates of H→bb Tagger in 2016: *b*-Tagging

- In 2016, we are following the similar tagging strategy as previous H→bb tagger
- Improvement in *b*-tagging:
  - Algorithm optimization in impact parameter based tagging (IP3D) and secondary vertex finder (SV)
  - Optimization of final multivariable *b*-tagging discriminant (MV2)
- Various discriminants available for *b*-tagging
  - Suit for various background composition
  - Improved rejection against light/c-jet in anti-k_T R=0.4 calorimeter jets
  - Similar improvement expected in track jet b-tagging



Modeling in g→bb Enriched Data

ATLAS-CONF-2016-002

Challenges

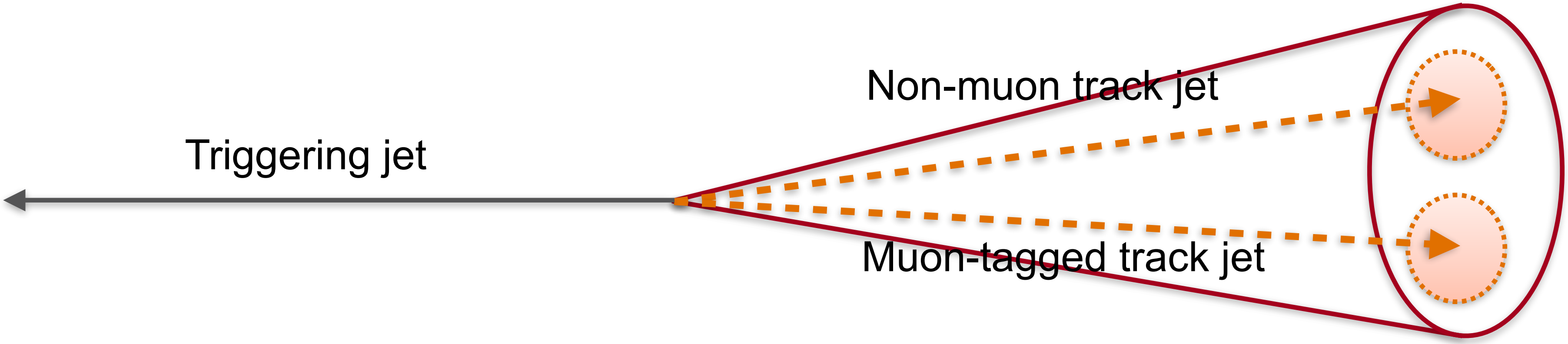
- Not enough H→bb in data — not even found yet :)

- Solution: *g→bb* provides a copious sample of close-by b-jets
  - Double b-tagging systematics for track jets
  - Check modeling of large-R jet substructure variables
  - Cross-check large-R jet energy scale (JES) / jet mass scale (JMS) / $D_2^{\beta=1}$ uncertainty

- We check these using 2012 8TeV data

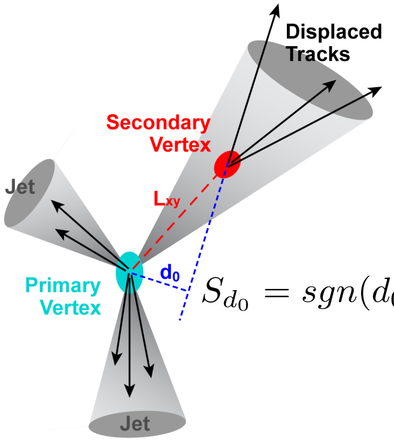
Strategy

- Goal: Increase g→bb purity
- Key: At least one of small radius track jets should be matched to muon
  - Select semi-leptonic b-hadron decays
  - Enrich events with jets containing b-hadrons
- Further double b-tagging on small-R track jets to obtain high purity g→bb samples

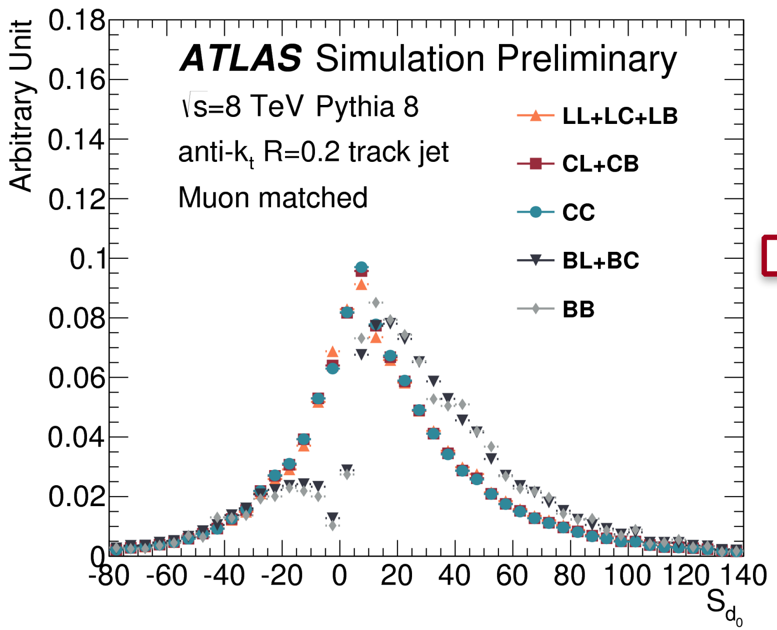


Flavor Fraction Correction

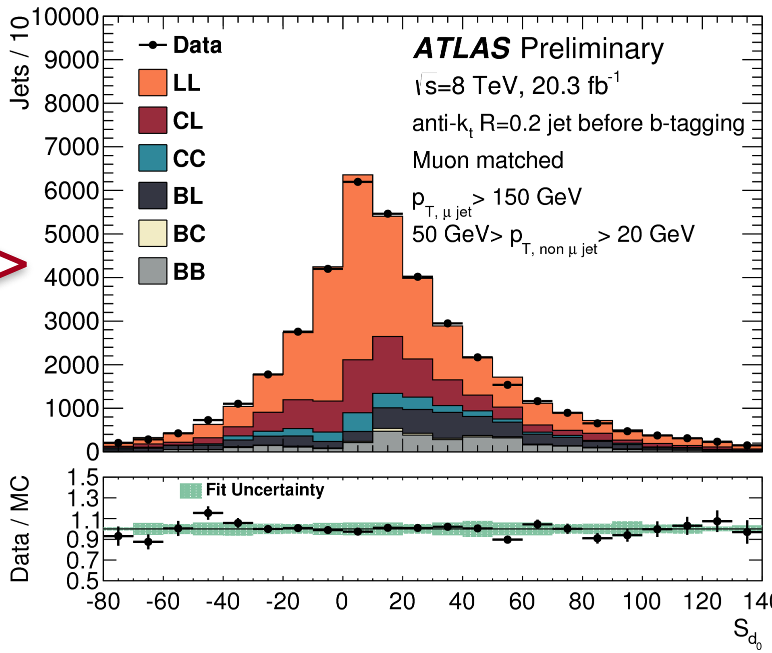
- Issue: MC does not model the heavy flavor content in data
- Method: Fit variable sensitive to flavor composition to data
- Discriminant: Largest track impact parameter significance inside track jet
- Fit procedure:
  - Build S_{d_0} template for each flavor component
  - Simultaneous binned likelihood fit on 1-D S_{d_0} distribution for muon and non-muon track jet



$$S_{d_0} = sgn(d_0) \times \frac{d_0}{\sigma(d_0)}$$



Fit

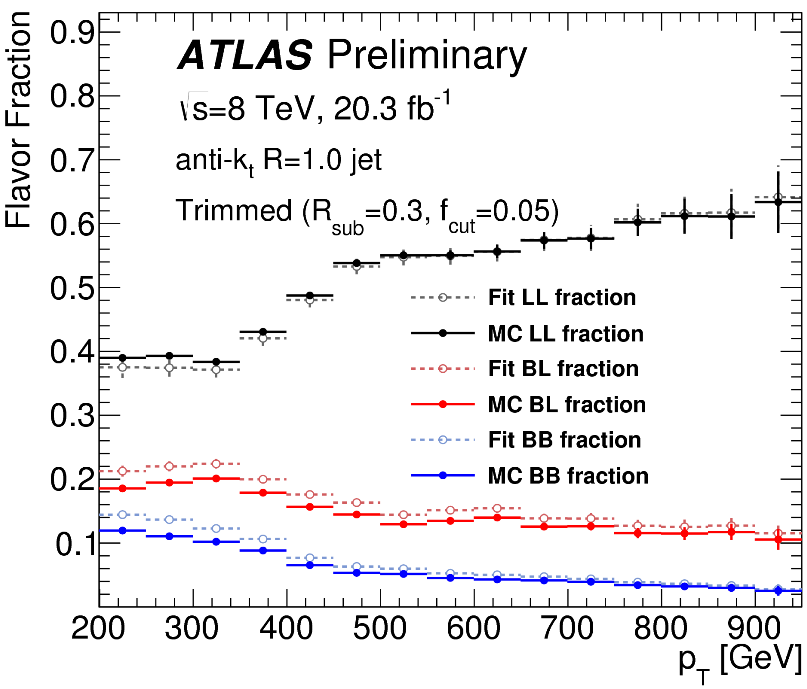


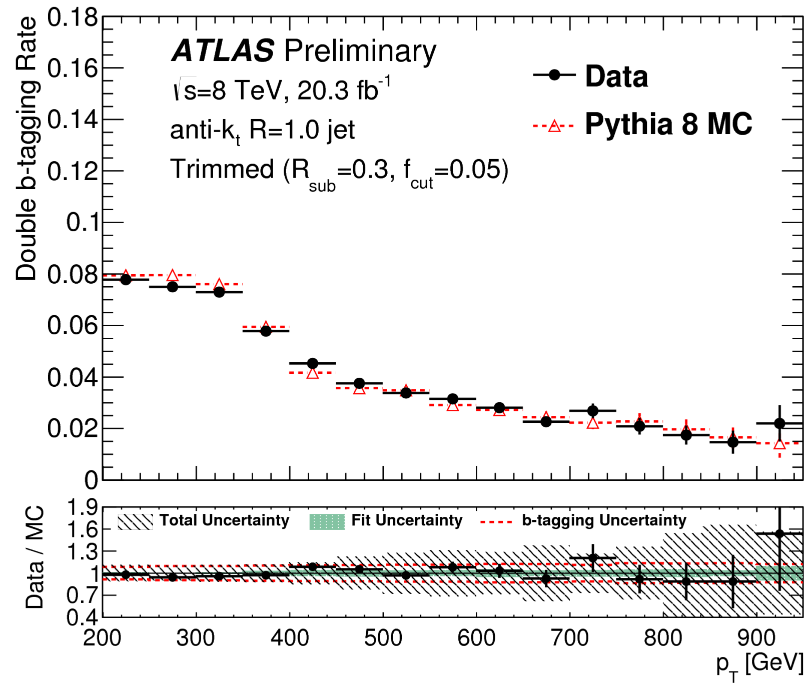
Flavor Fraction Correction

- A clear discrepancy between data/MC on flavor fraction can be seen for large-R jet, especially when pT < 500 GeV

  → Flavor correction becomes essential

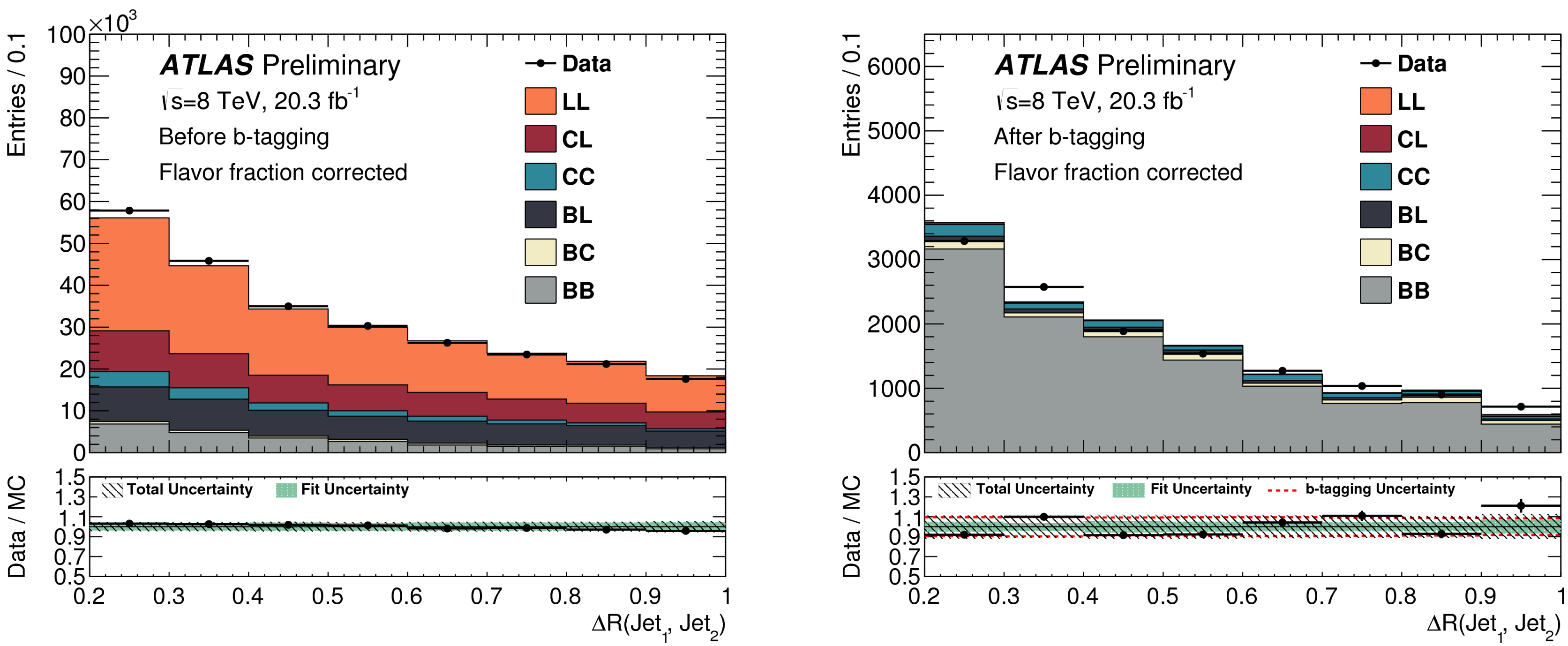
- Double b-tagging rate very well modeled after flavor correction



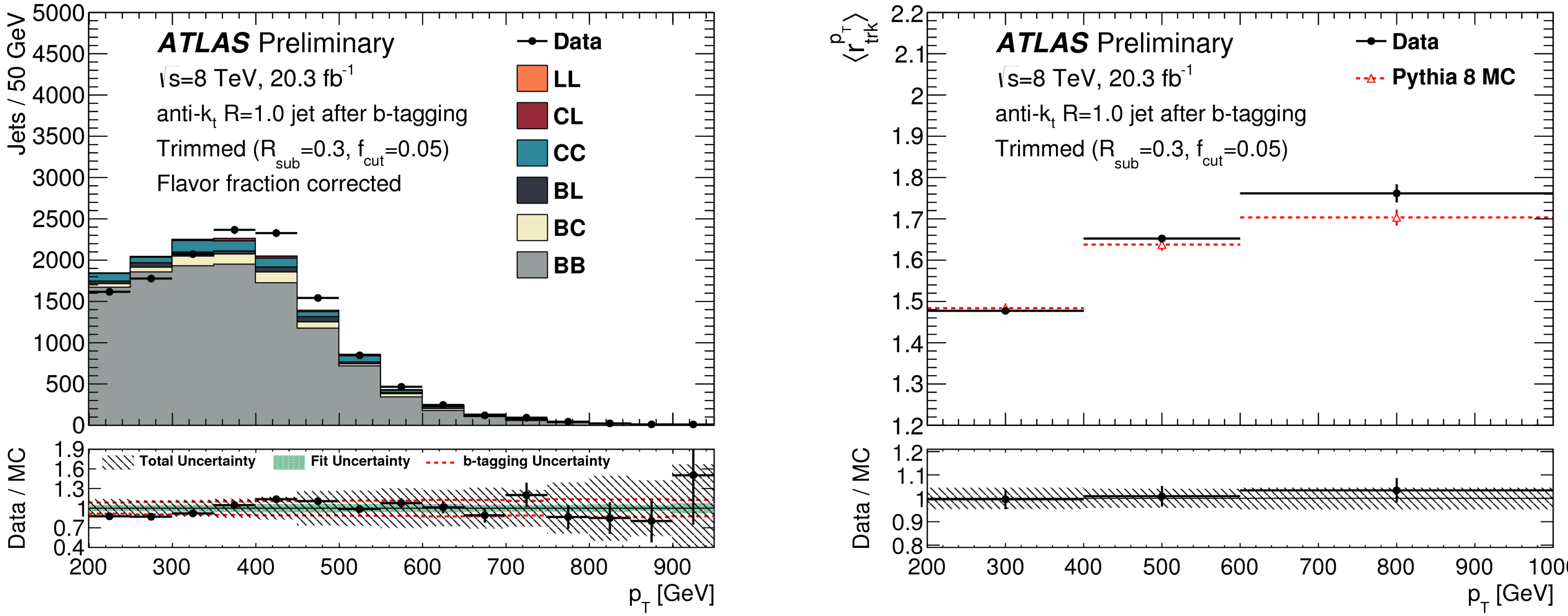


Results: Double b-Tagging

- Data/MC agrees well within the uncertainties
- It is noticeable that same level of agreement is seen at low ΔR in which jets are not isolated from each other



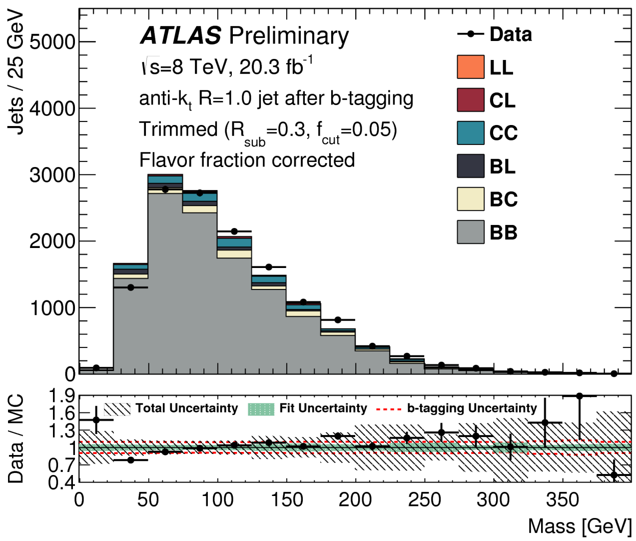
Results: Large-R Jet p_T Modeling

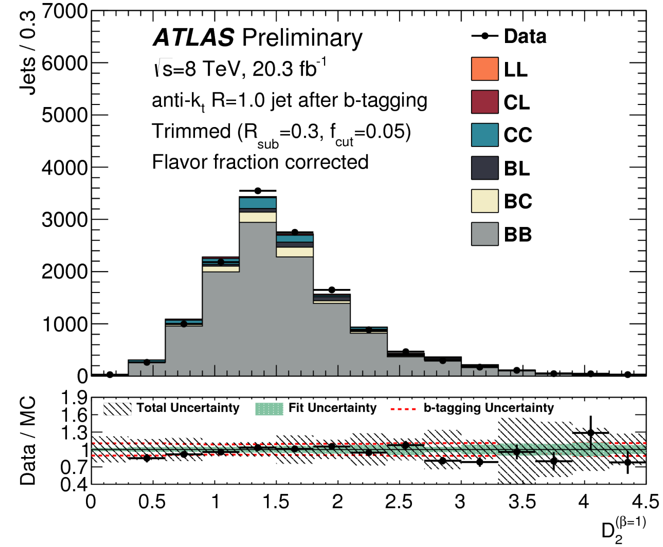


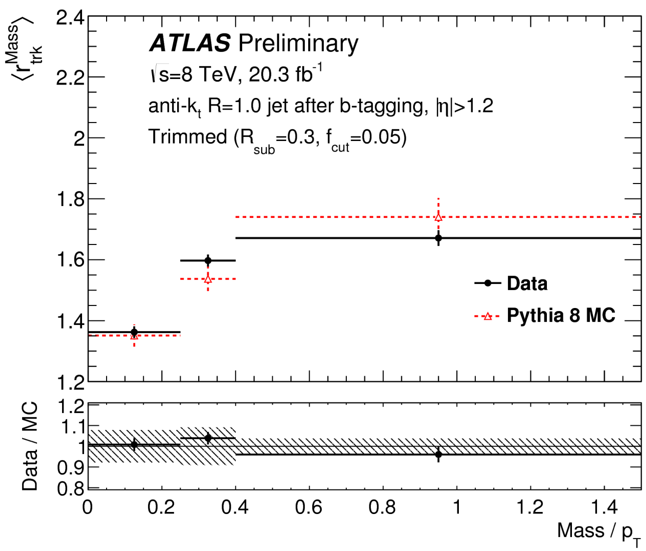
- Large-R jet pT well modeled within uncertainties
- Large-R jet JES uncertainty cross-check
  - We use data and MC ratio of variable $r_{trk}^{p_T} = p_T^{calo}/p_T^{trk}$ for in-situ calibration
  - Existing JES uncertainty derived from inclusive sample is large enough to cover topology dependence
- Same cross-check on JMS / $D_2^{\beta=1}$ uncertainty

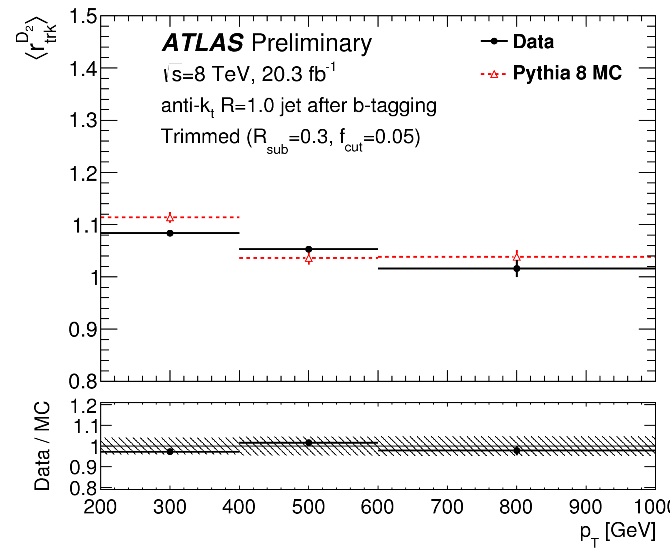
Results: Large-R Jet Mass / Substructure Modeling

SLAC









Concluding Remarks

Boosted H→bb tagger is highly successful in ATLAS for new physics search with 13TeV data

Pushing the physics search limits in ATLAS through

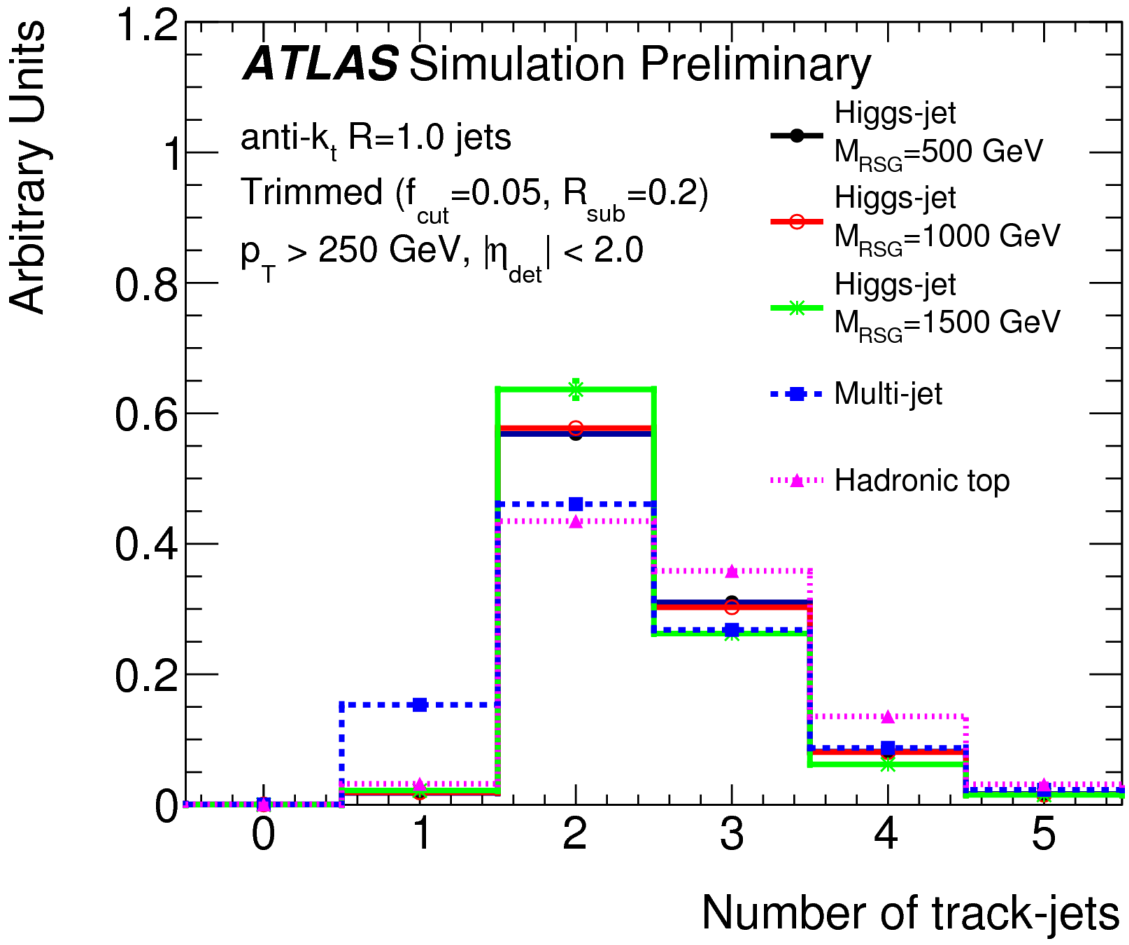
- Jet substructure techniques
- Small-R track jet *b*-tagging

Modeling of b-tagging performance and large-R jet properties studied in g→bb topology for the first time in ATLAS using 8TeV data

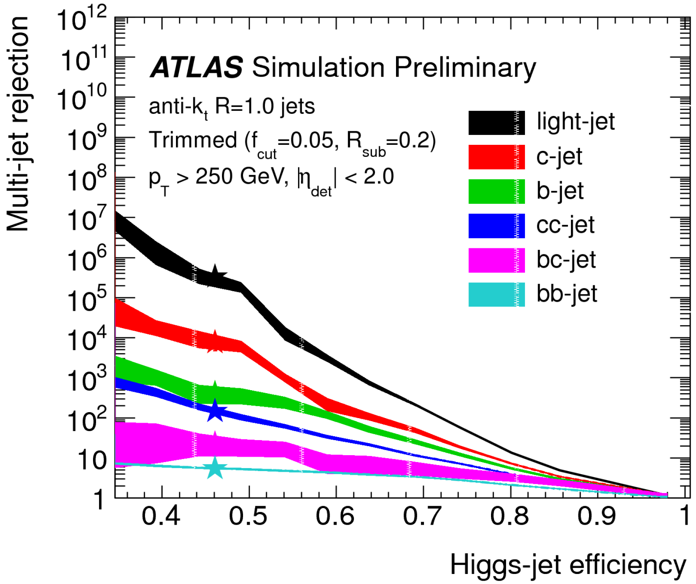
- Useful to cross-check flavor dependency of existing calibrations
- Modeling is in good agreement within uncertainties

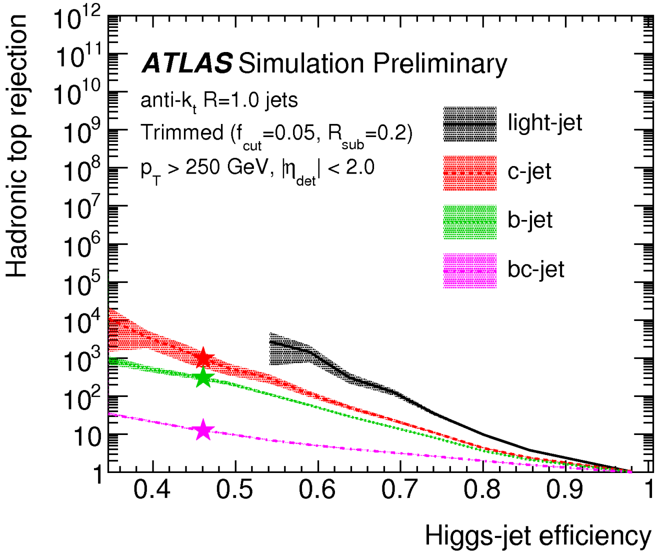
Back Up

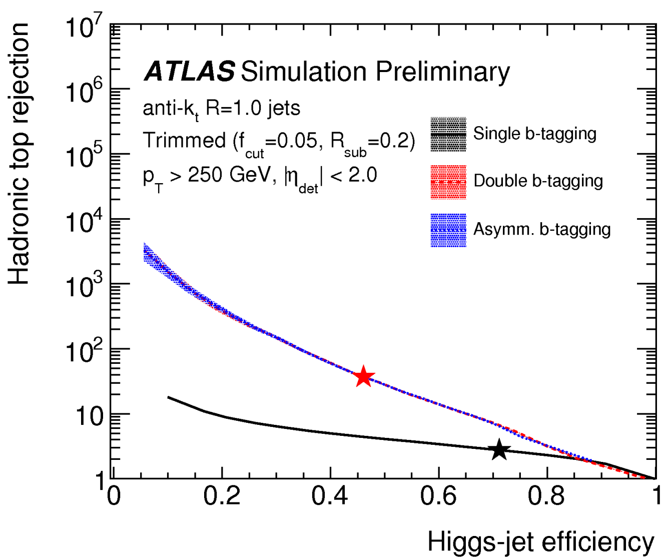
H→bb: Number of Track Jets



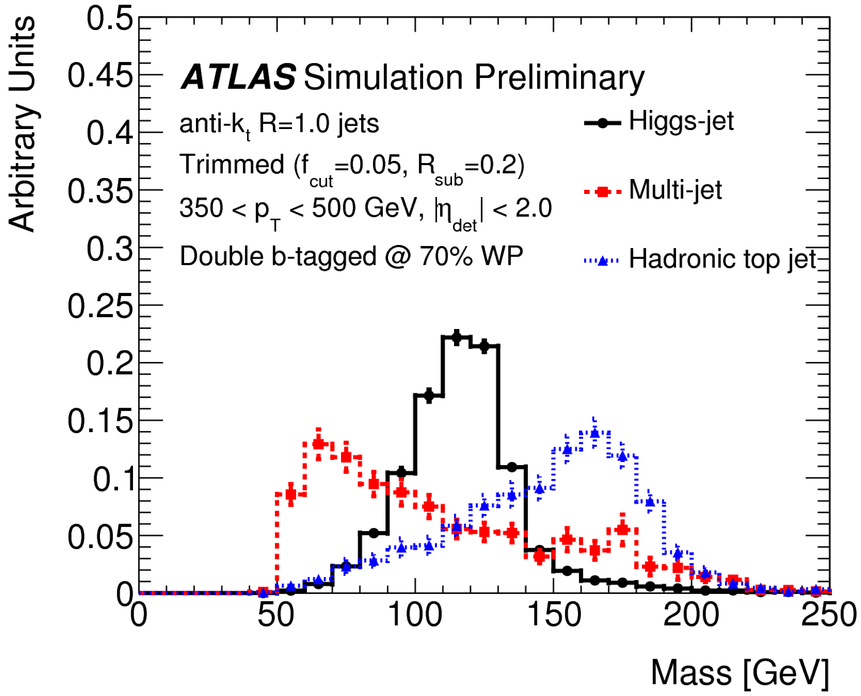
H→bb: *b*-Tagging

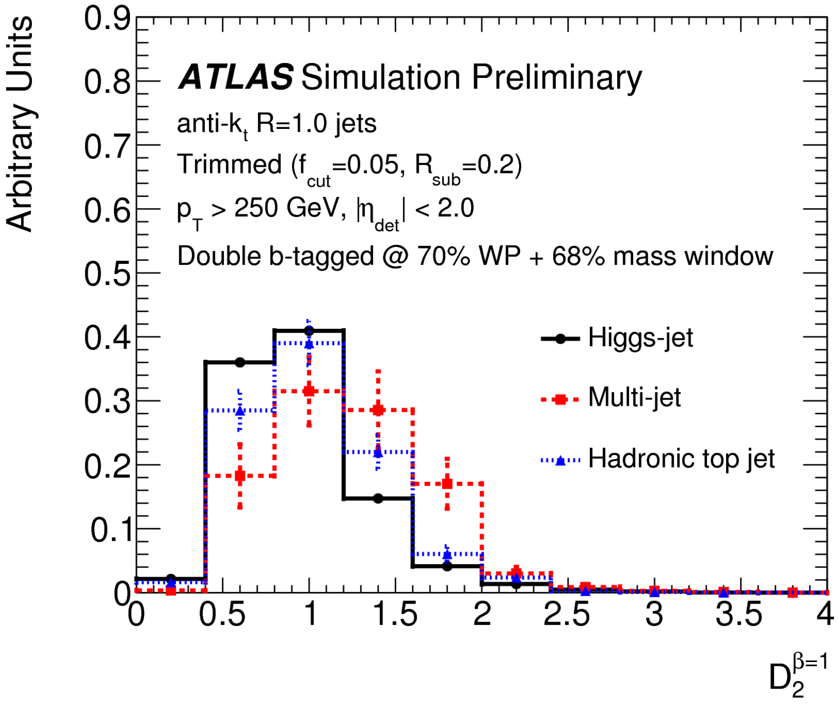




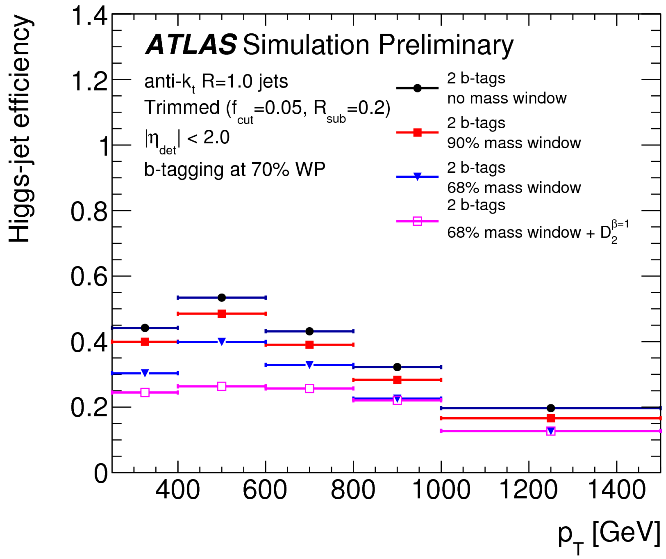


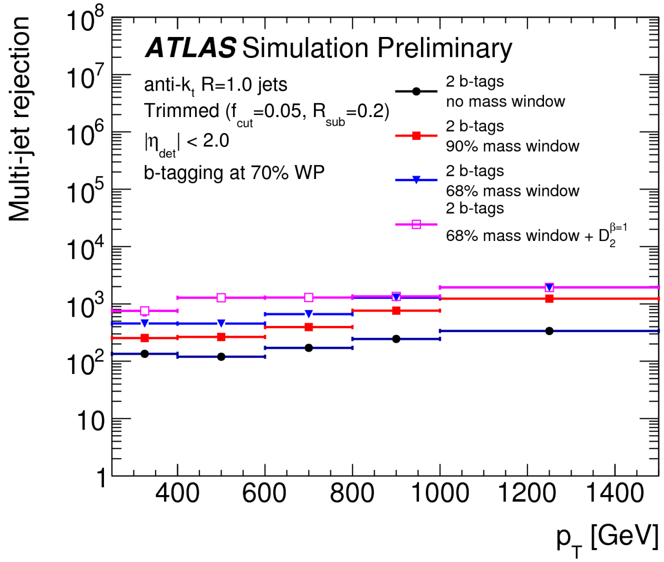
H→bb: Large-R Jet Mass and Substructure

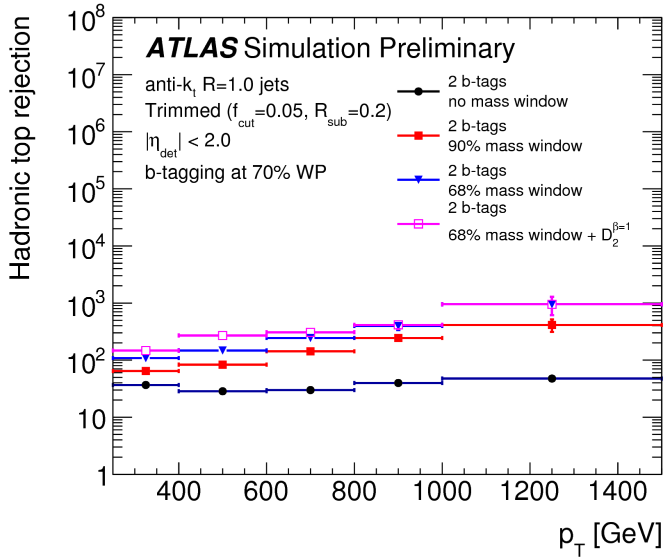




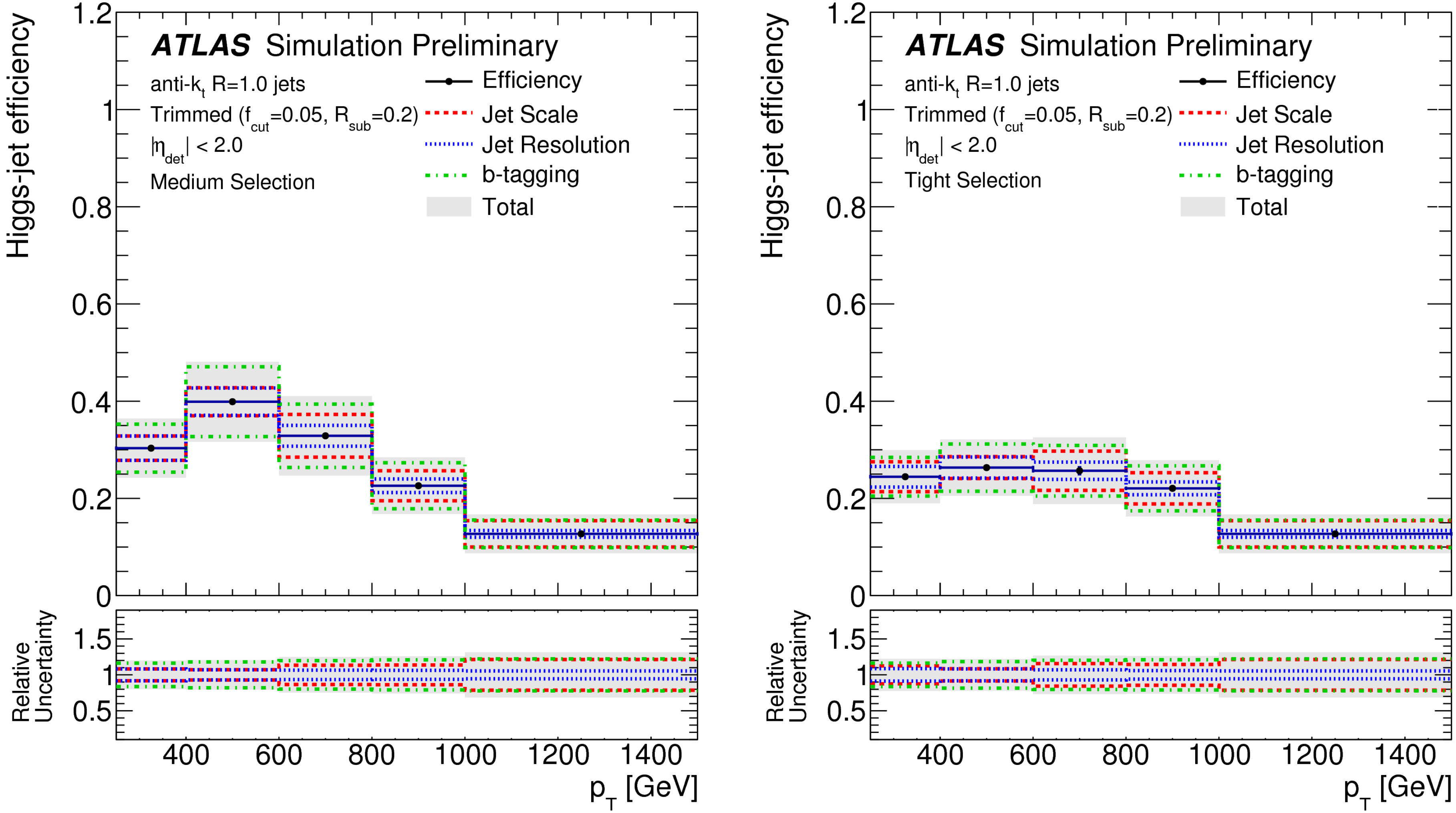
H→bb Taggers


ATLAS Simulation Preliminary
anti-k$_t$ R=1.0 jets
Trimmed (f_{cut}=0.05, R_{sub}=0.2)
$|\eta_{det}|$ < 2.0
b-tagging at 70% WP
2 b-tags no mass window
2 b-tags 90% mass window
2 b-tags 68% mass window
2 b-tags 68% mass window + $D_2^{\beta=1}$
Higgs-jet efficiency
p_T [GeV]

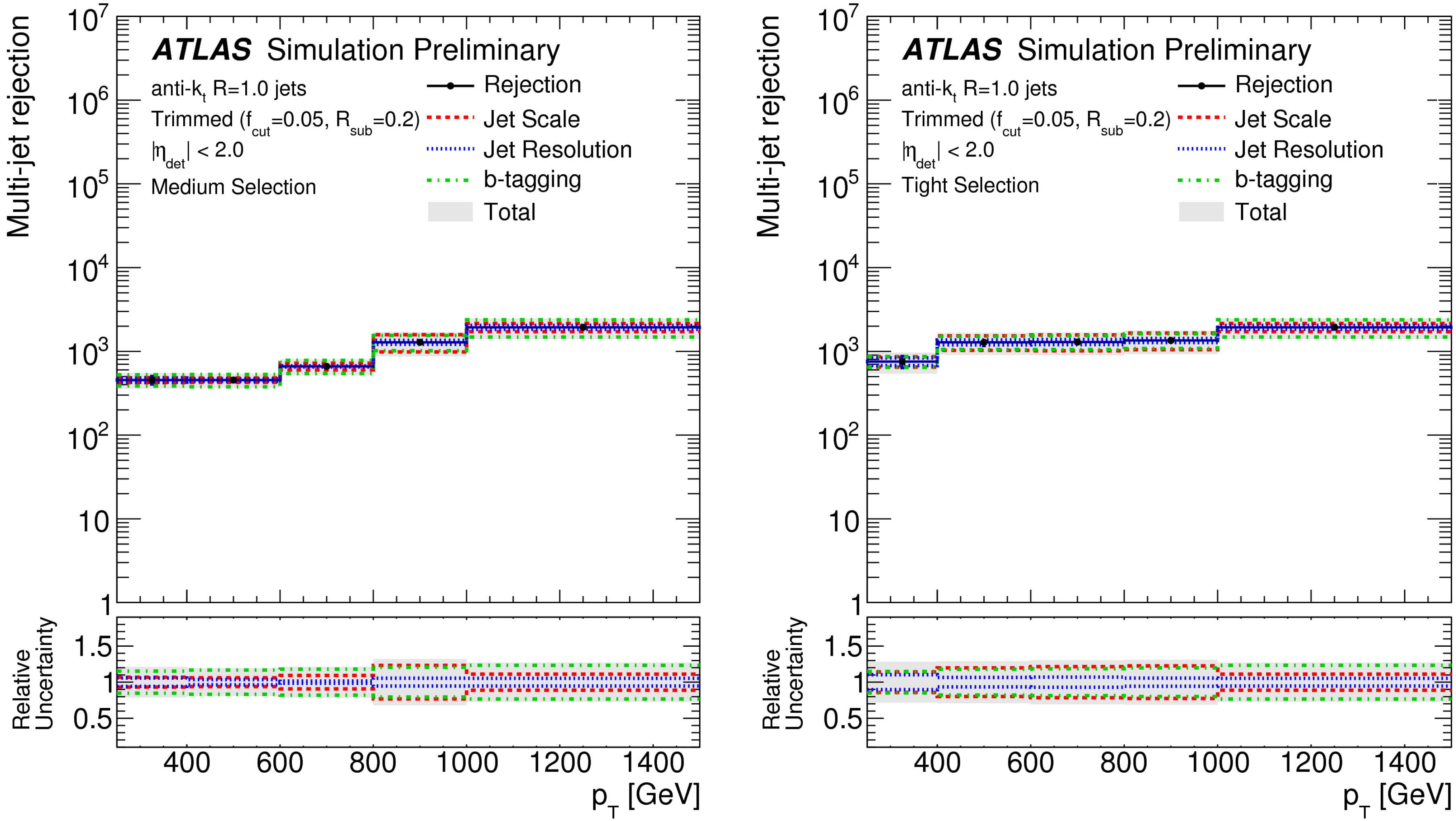

ATLAS Simulation Preliminary
anti-k$_t$ R=1.0 jets
Trimmed (f_{cut}=0.05, R_{sub}=0.2)
$|\eta_{det}|$ < 2.0
b-tagging at 70% WP
2 b-tags no mass window
2 b-tags 90% mass window
2 b-tags 68% mass window
2 b-tags 68% mass window + $D_2^{\beta=1}$
Multi-jet rejection
p_T [GeV]


ATLAS Simulation Preliminary
anti-k$_t$ R=1.0 jets
Trimmed (f_{cut}=0.05, R_{sub}=0.2)
$|\eta_{det}|$ < 2.0
b-tagging at 70% WP
2 b-tags no mass window
2 b-tags 90% mass window
2 b-tags 68% mass window
2 b-tags 68% mass window + $D_2^{\beta=1}$
Hadronic top rejection
p_T [GeV]

H→bb Taggers

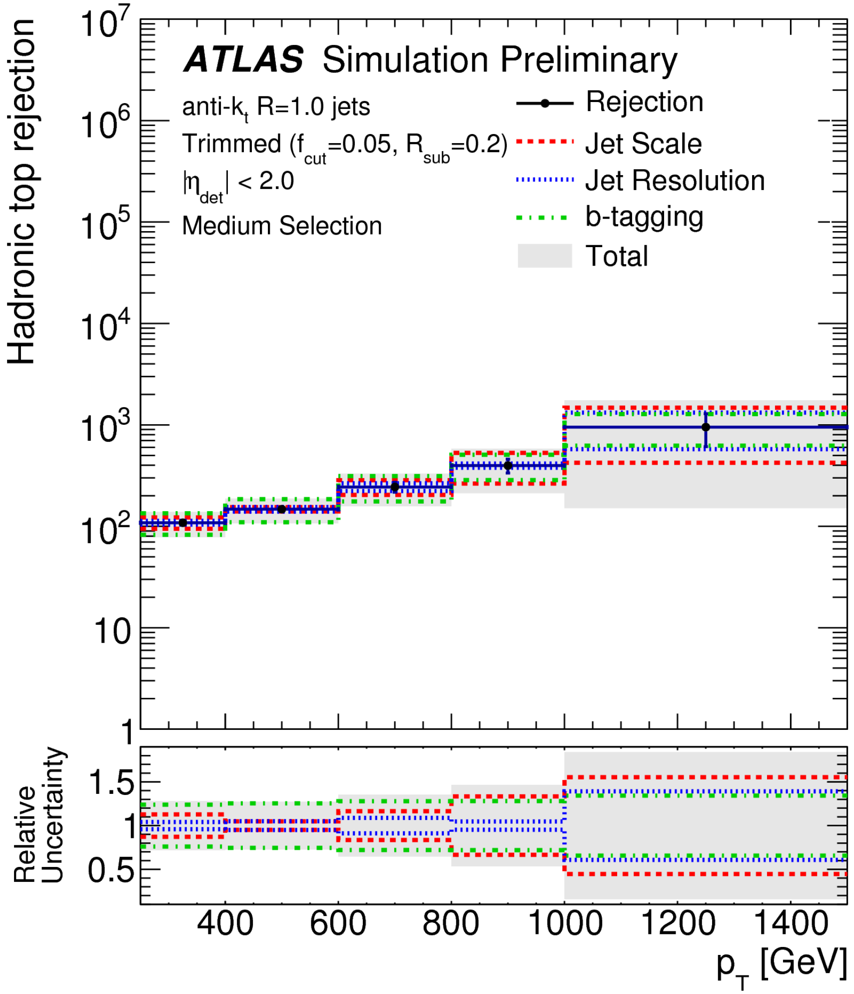

ATLAS Simulation Preliminary
anti-k_t R=1.0 jets
Trimmed (f_{cut}=0.05, R_{sub}=0.2)
$|\eta_{det}|$ < 2.0
Medium Selection
Efficiency
Jet Scale
Jet Resolution
b-tagging
Total
Higgs-jet efficiency
Relative Uncertainty
p_T [GeV]
ATLAS Simulation Preliminary
anti-k_t R=1.0 jets
Trimmed (f_{cut}=0.05, R_{sub}=0.2)
$|\eta_{det}|$ < 2.0
Tight Selection
Efficiency
Jet Scale
Jet Resolution
b-tagging
Total
Higgs-jet efficiency
Relative Uncertainty
p_T [GeV]

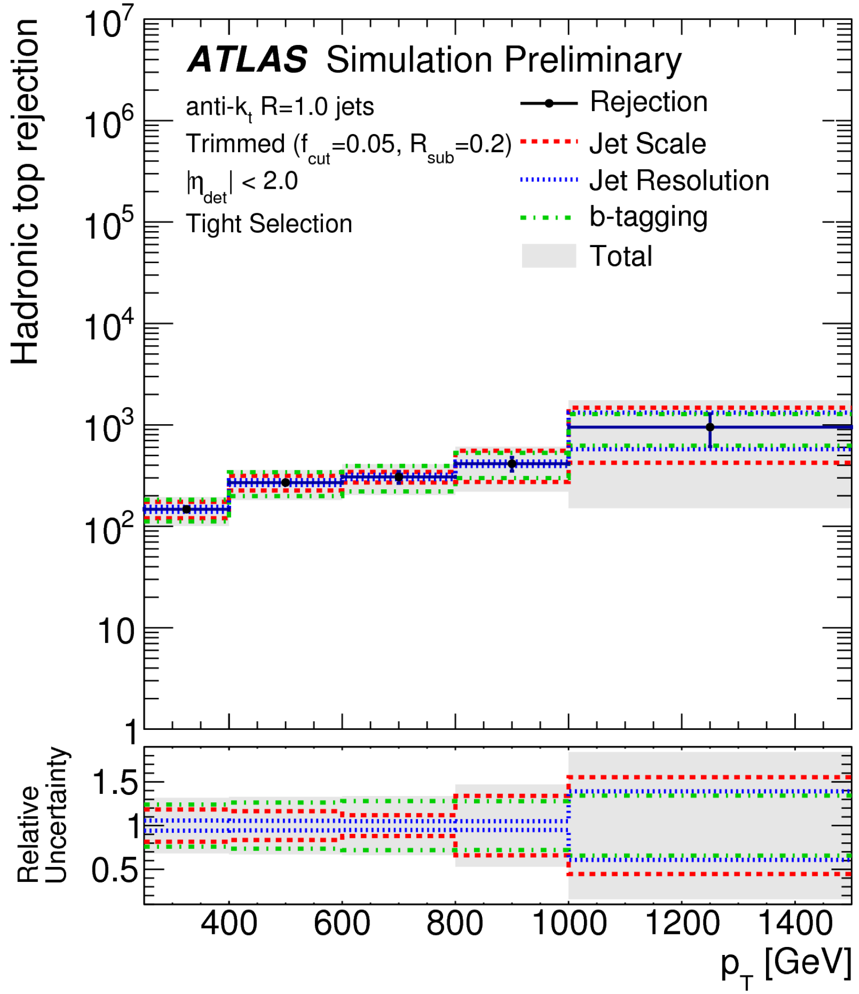
H→bb Taggers



ATLAS Simulation Preliminary
anti-k$_t$ R=1.0 jets
Trimmed (f$_{cut}$=0.05, R$_{sub}$=0.2)
|η$_{det}$| < 2.0
Medium Selection
Rejection
Jet Scale
Jet Resolution
b-tagging
Total
Multi-jet rejection
Relative Uncertainty
p$_T$ [GeV]
ATLAS Simulation Preliminary
anti-k$_t$ R=1.0 jets
Trimmed (f$_{cut}$=0.05, R$_{sub}$=0.2)
|η$_{det}$| < 2.0
Tight Selection
Rejection
Jet Scale
Jet Resolution
b-tagging
Total
Multi-jet rejection
Relative Uncertainty
p$_T$ [GeV]

H→bb Taggers

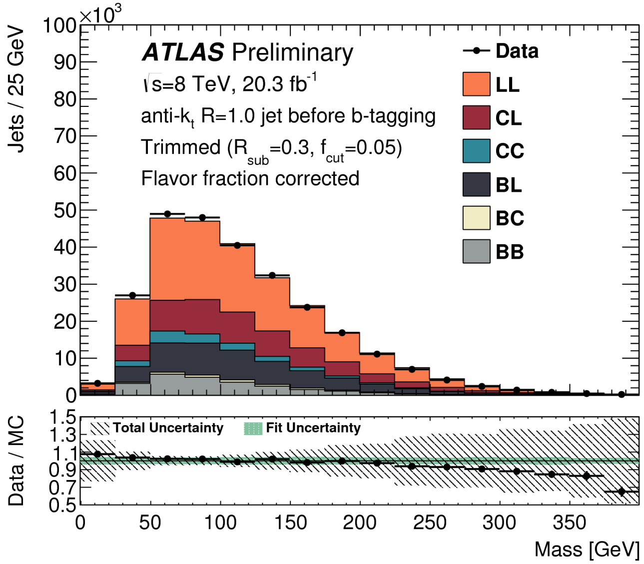

ATLAS Simulation Preliminary
anti-k_t R=1.0 jets
Trimmed (f_{cut}=0.05, R_{sub}=0.2)
$|\eta_{det}|$ < 2.0
Medium Selection
Rejection
Jet Scale
Jet Resolution
b-tagging
Total
Hadronic top rejection
Relative Uncertainty
p_T [GeV]

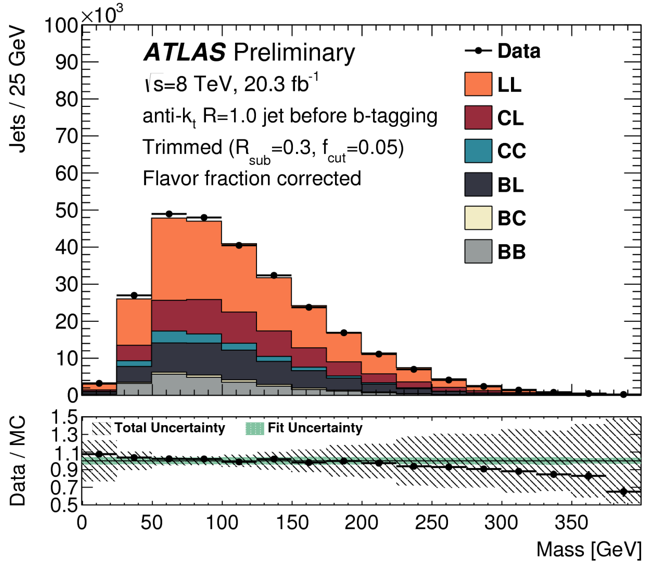

ATLAS Simulation Preliminary
anti-k_t R=1.0 jets
Trimmed (f_{cut}=0.05, R_{sub}=0.2)
$|\eta_{det}|$ < 2.0
Tight Selection
Rejection
Jet Scale
Jet Resolution
b-tagging
Total
Hadronic top rejection
Relative Uncertainty
p_T [GeV]

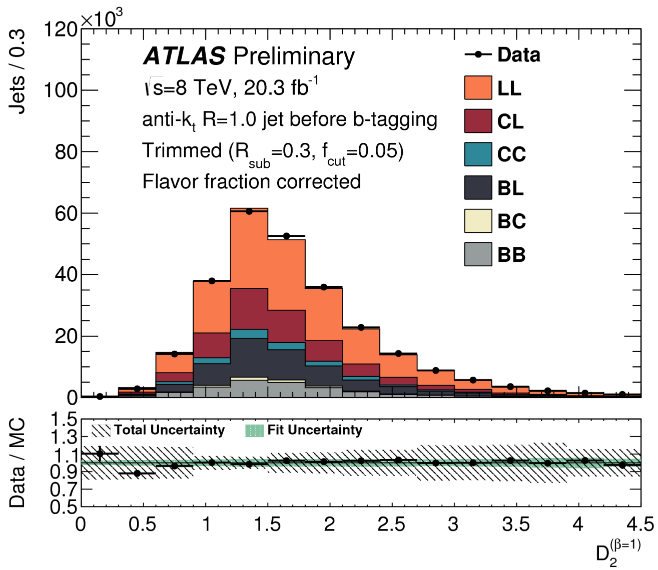
H→bb Taggers

	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

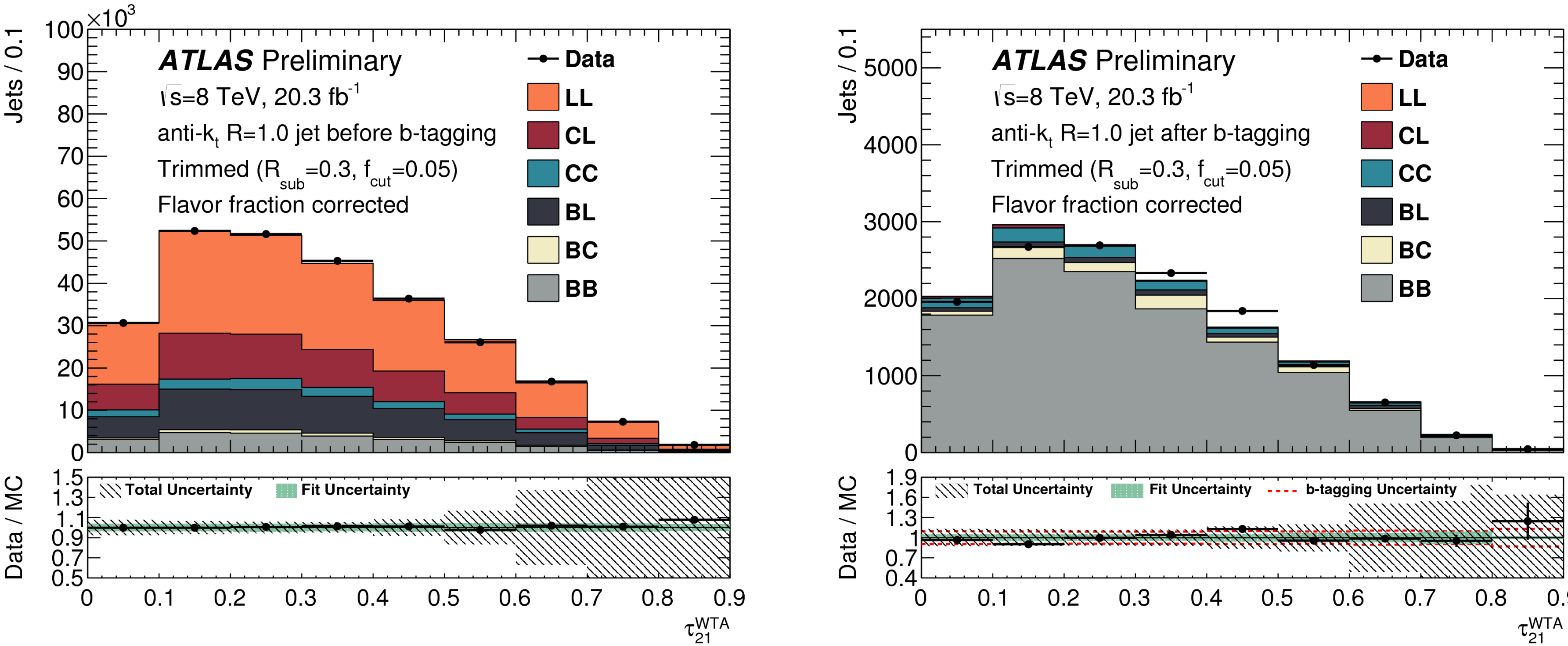
Large-R Jet Modeling before *b*-tagging







Large-R Jet Other Substructure Variables



Variable-R Jets in H→bb