Boosted Measurements in ATLAS+CMS

Dan Marley
On behalf of the ATLAS & CMS collaborations

17-22 September 2017

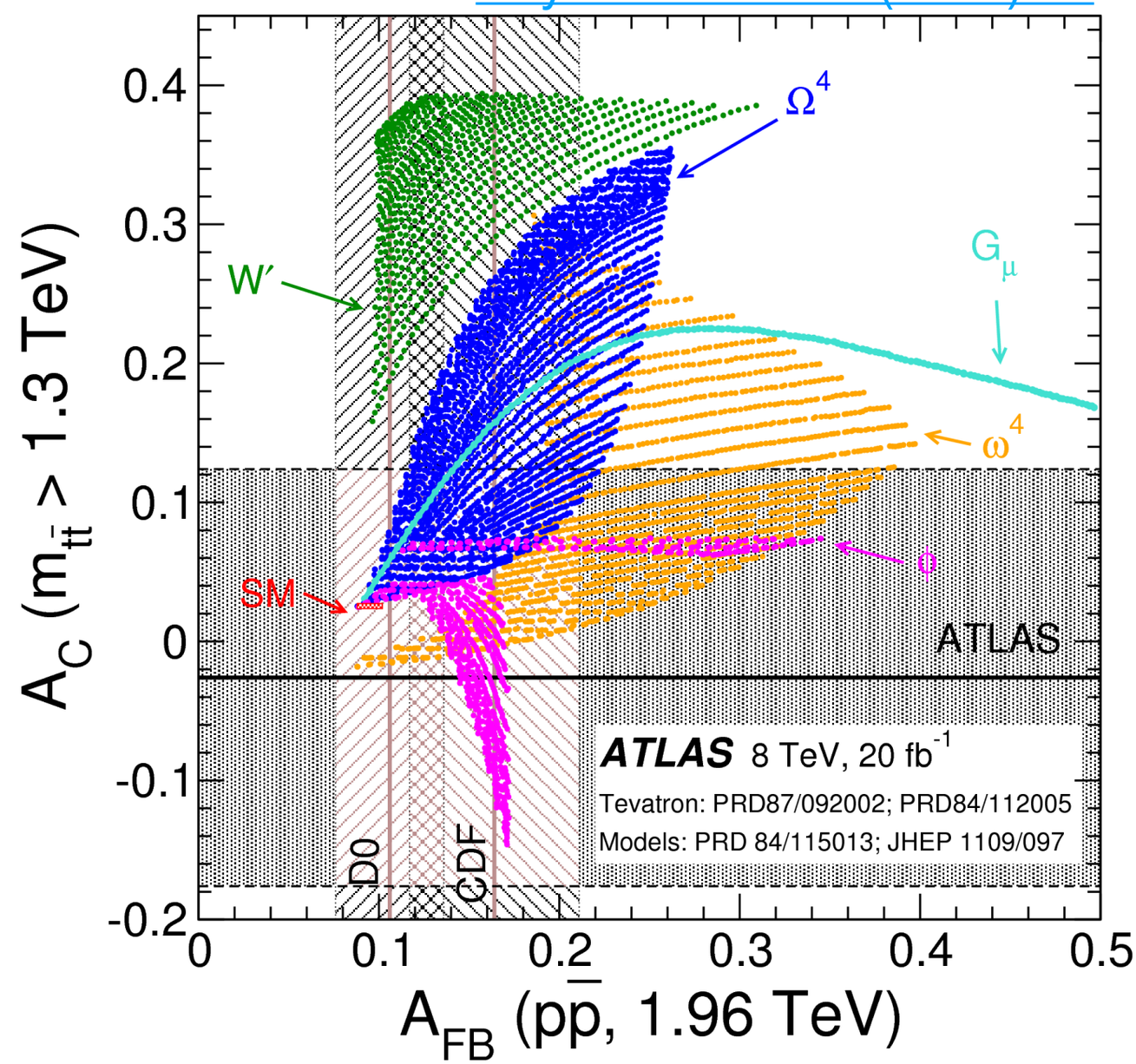
TOP2017 Braga, Portugal
10th International Workshop on Top Quark Physics

TEXAS A&M UNIVERSITY

Motivation

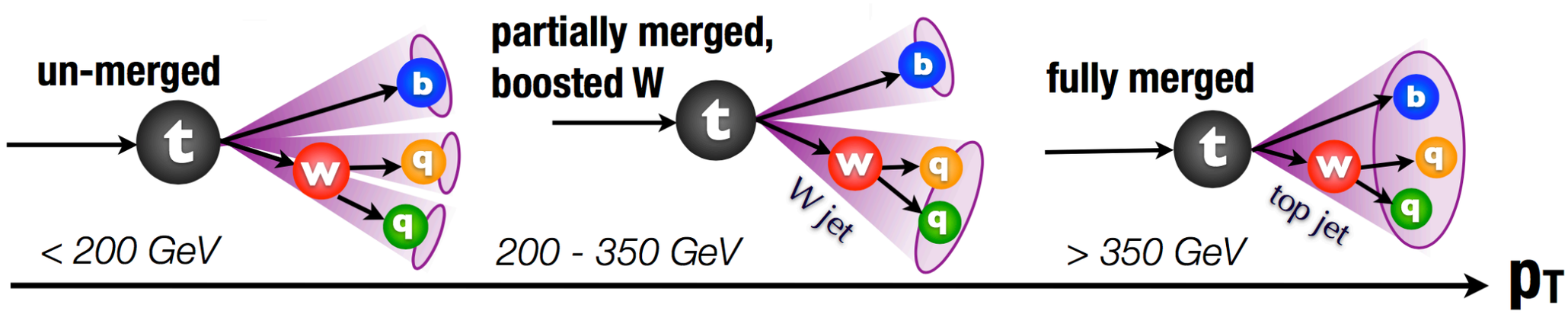
- Test the Standard Model — specifically ttbar — at very high energies
  - Compare measured values with multiple SM predictions
- Complementary way to search for new physics
  - Exclude BSM physics incompatible with measurements

Phys. Lett. B756 (2016) 52



Challenges

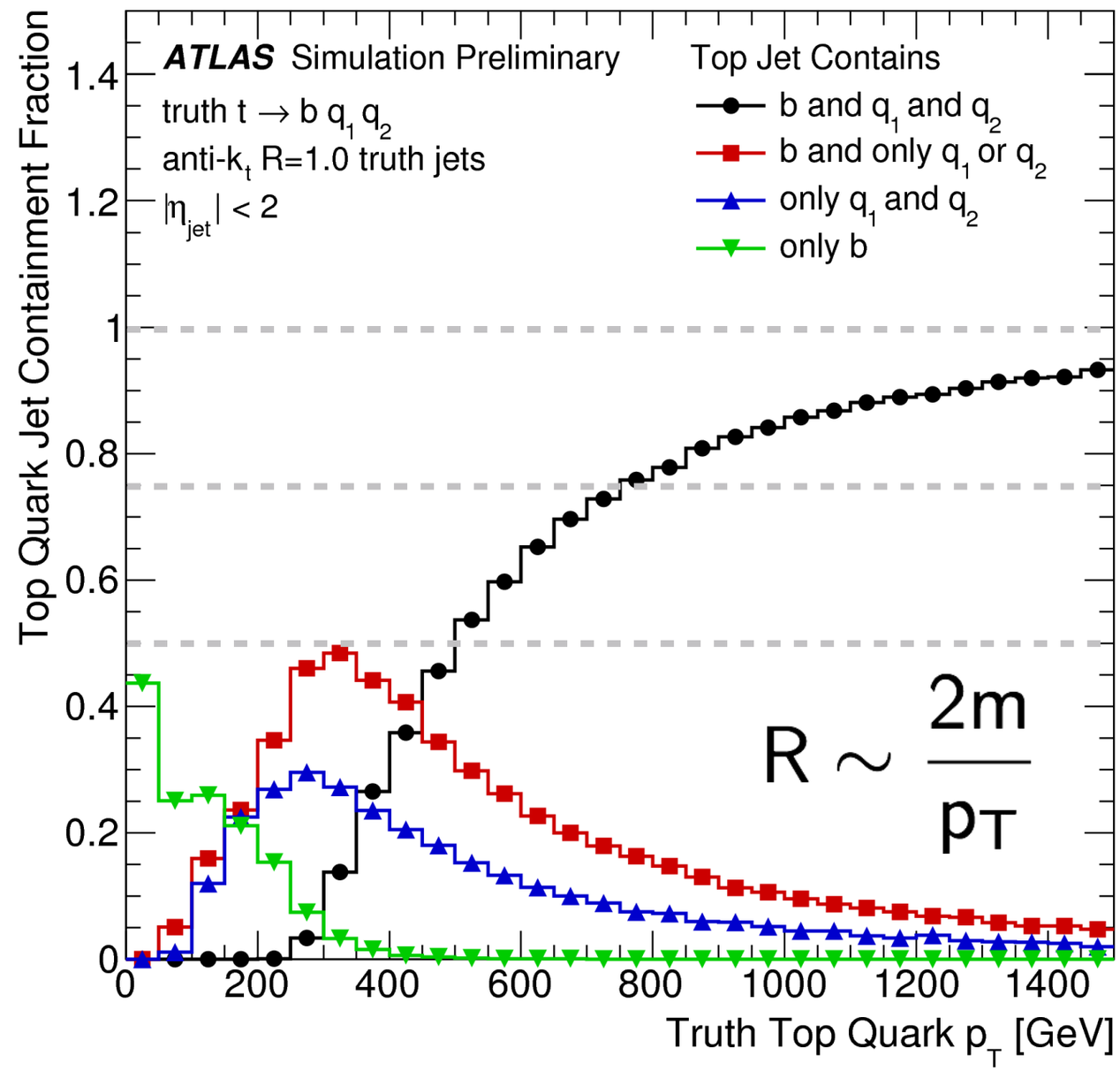
- Reconstruction of high-p_T hadronically- and leptonically-decaying top quarks
  - Merged decay products require unique algorithms
- Reducing systematic uncertainties associated with boosted top quark identification
- Modeling of top quarks in high-p_T regime
- Limited statistics: fewer events available to study relative to “resolved” analyses



Reconstruction

ATLAS-CONF-2017-064

- Large-R jets (R≥0.8) capture boosted top quark decay products
- Use jet substructure information to select pure sample of top quarks
  - Ensure "good" reconstruction of ttbar system (~diagonal response matrix)
- Mitigate pileup by using grooming procedures



Boosted top tagging is part of a much larger community! See BOOST2017

See talks by S. Egan, G. Kasieczka, & S. Lee

Top Measurements

ATLAS	8 TeV	13 TeV
	Differential Cross Section ℓ+jets (PRD 93 (2016) 032009) **Charge Asymmetry** ℓ+jets (PLB 756 (2016) 52)	**Differential Cross Section** ℓ+jets (arXiv:1708.00727) ☆ All-hadronic (*In preparation*) ☆

CMS	8 TeV	13 TeV
	Differential Cross Section ℓ+jets (p_T PRD 94 (2016) 072002; jet mass: EPJC 77 (2017) 467) ☆ All-hadronic (CMS-PAS-TOP-16-018) ☆	**Differential Cross Section** All-hadronic (CMS-PAS-TOP-16-013)

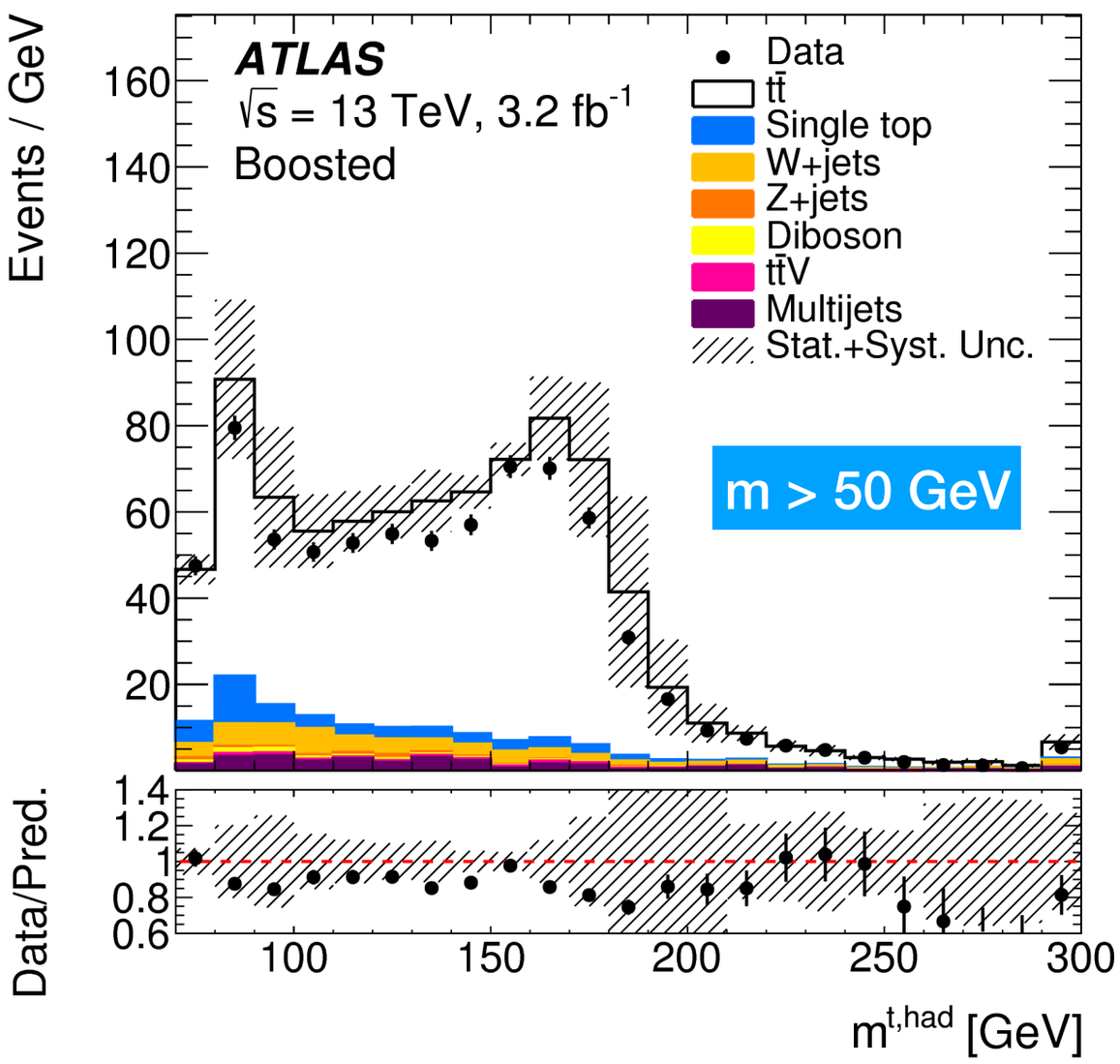
☆ Latest updates presented here

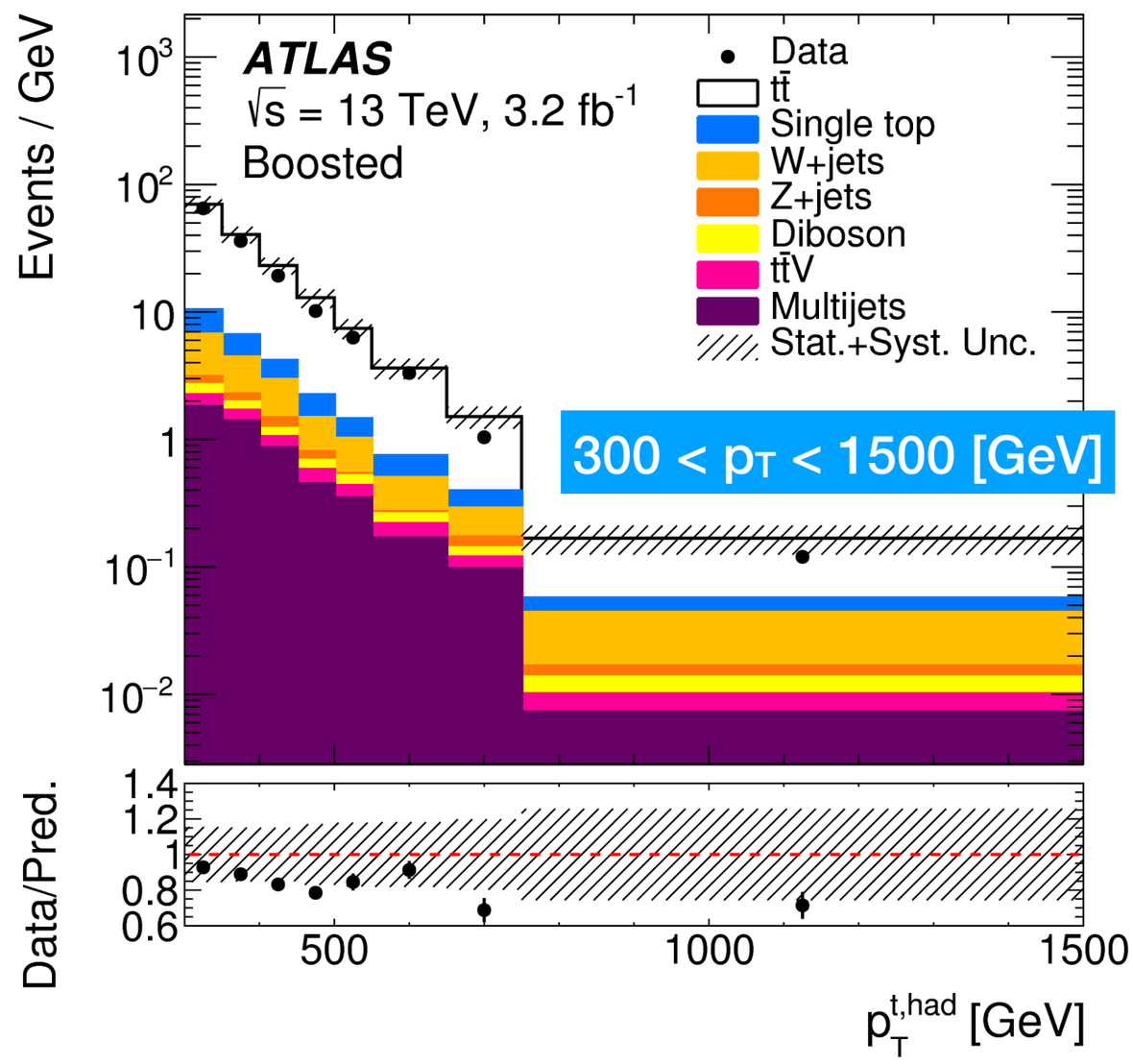
arXiv:1708.00727

ℓ+jets (13 TeV)

See talk by M. Giannelli

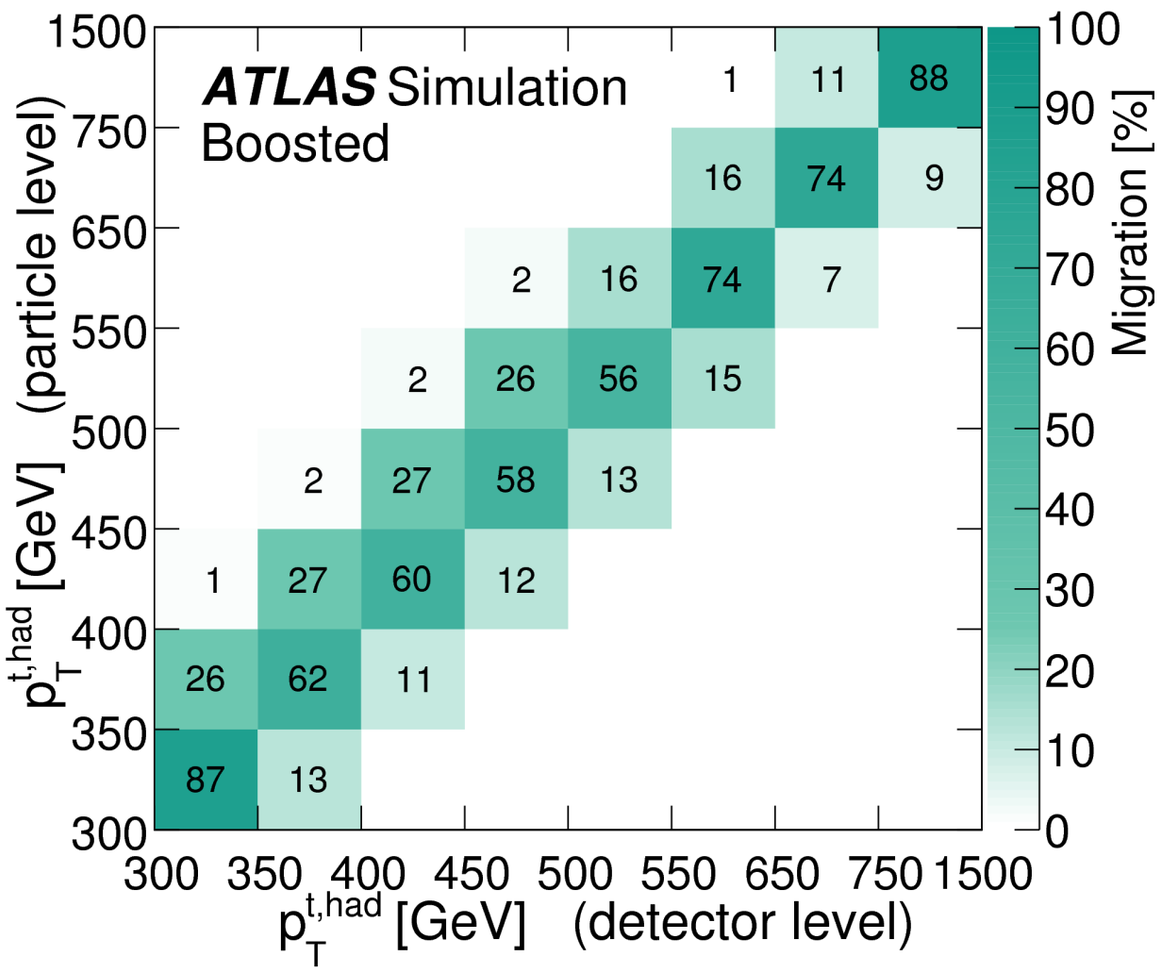
- **ℓ+jets** final state with large-R jet representing the hadronic top:
  - Trimmed (f_{cut}=0.05, R_{sub}=0.2), R=1.0 anti-k_T large-R jet
  - Top-tagged using p_T-dependent cuts on τ_{32} and large-R jet mass (80% ε)
  - *Particle-level*: mass > 100 GeV; τ_{32}<0.75

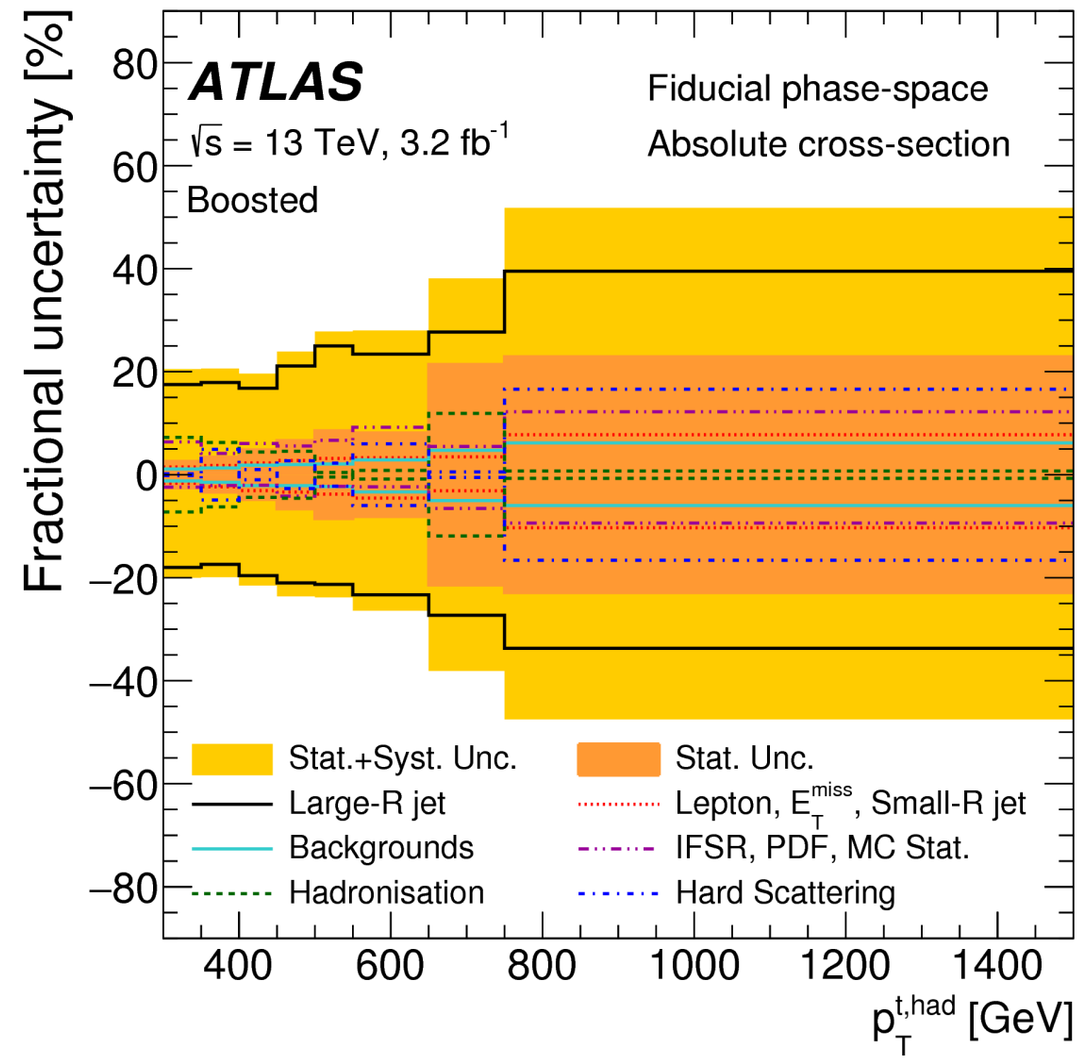




arXiv:1708.00727

ℓ+jets (13 TeV)

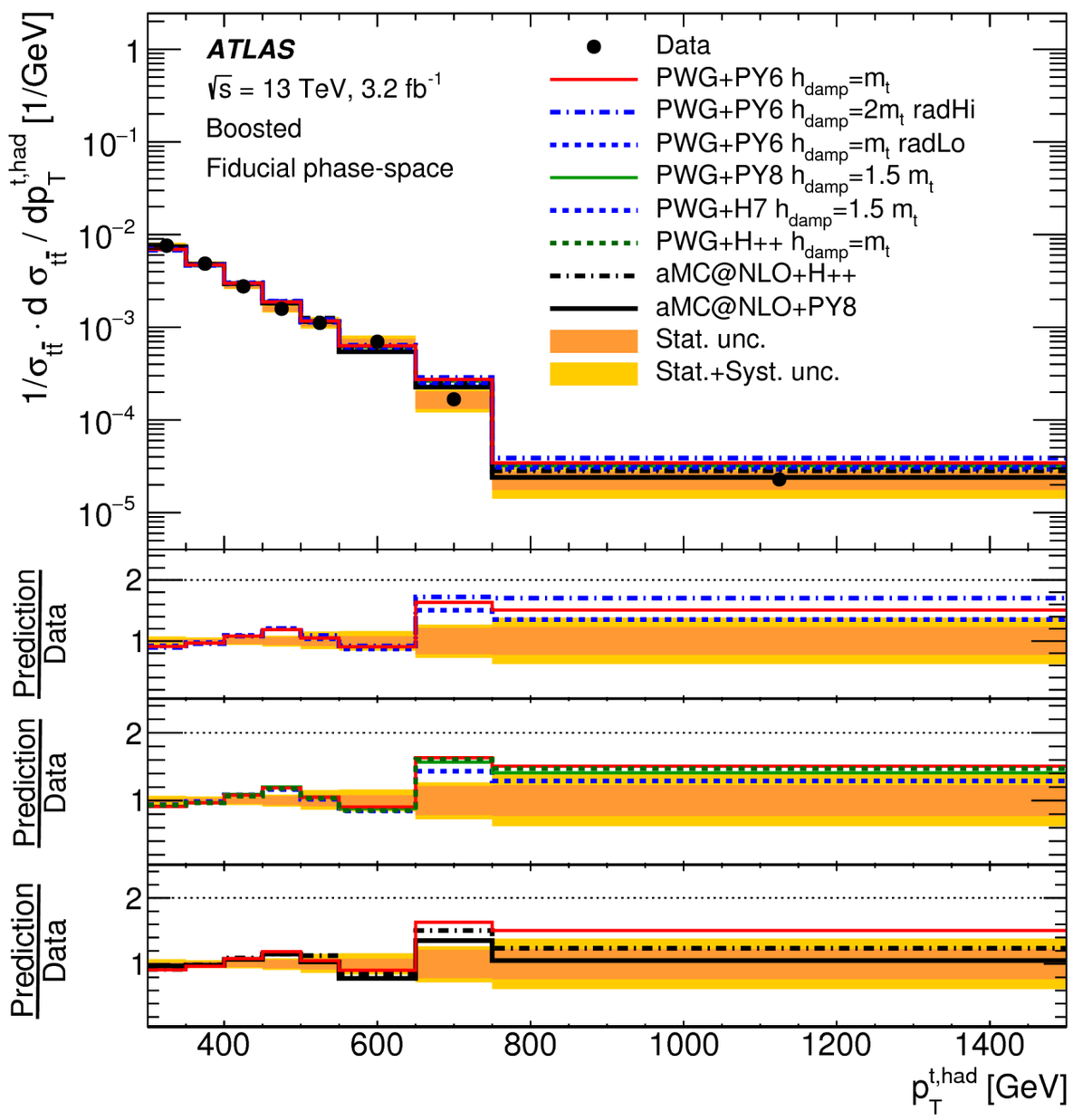




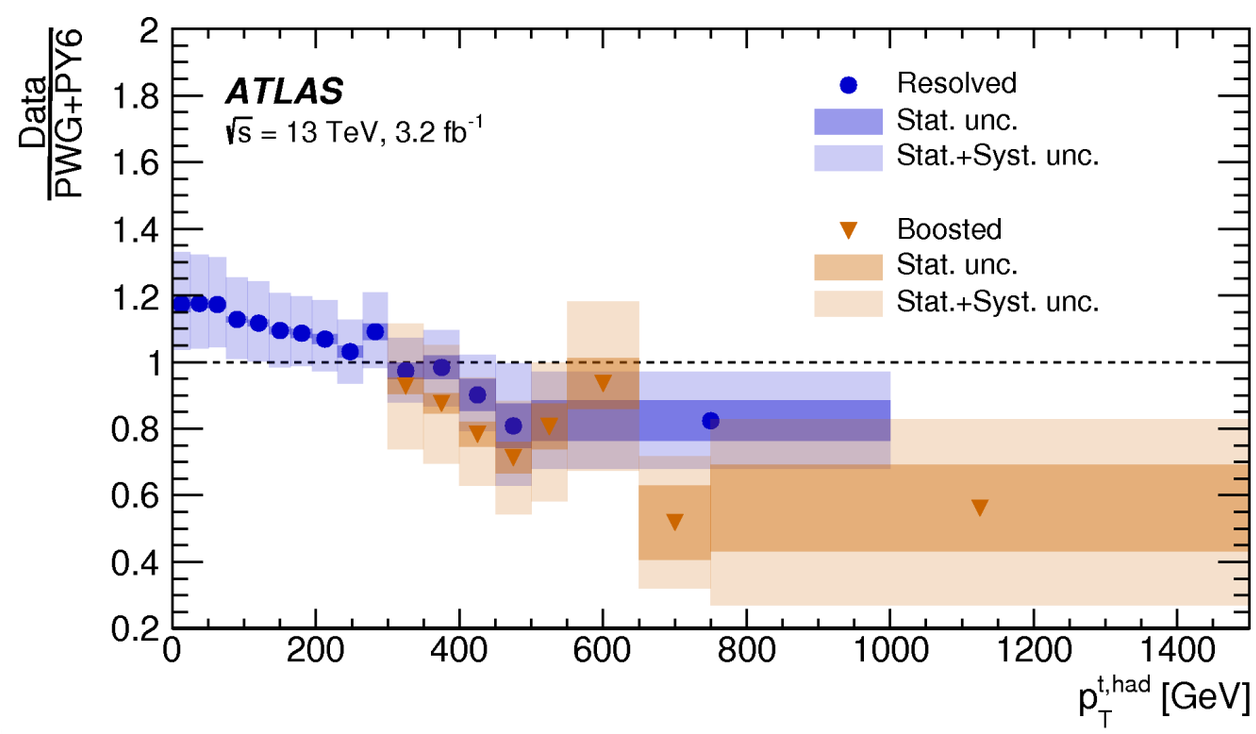
- Good reconstruction and nearly diagonal migration matrix (>85% of events along diagonal in rapidity migration matrix)
- Systematic uncertainties dominated by large-R jet and ttbar modeling components
  - Total is nearly twice the magnitude uncertainties in resolved channel!

arXiv:1708.00727

ℓ+jets (13 TeV)



- Iterative Bayesian Unfolding used to obtain final result
- Resolved and boosted channels consistently show overestimation at large top p_T



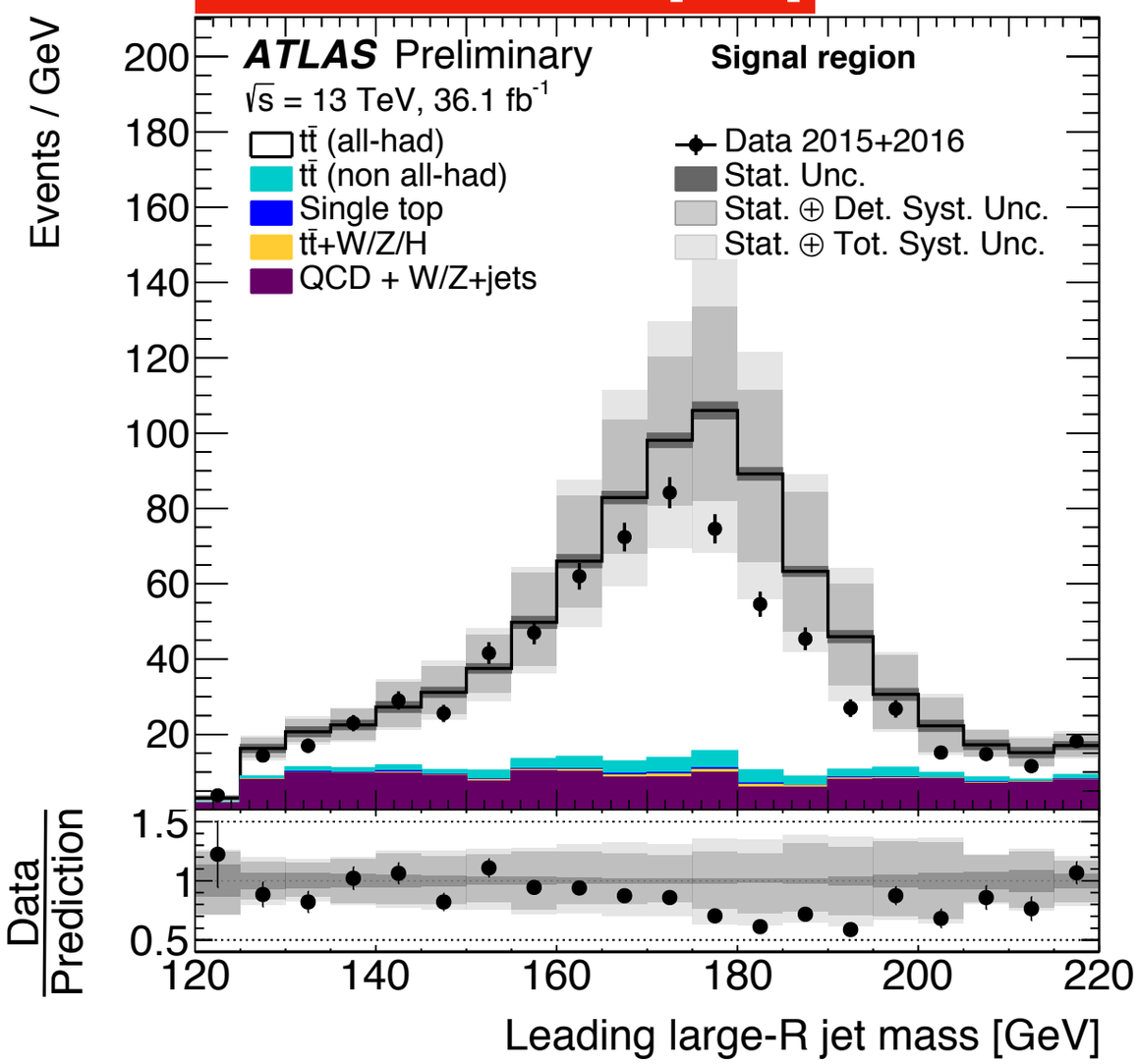
In preparation

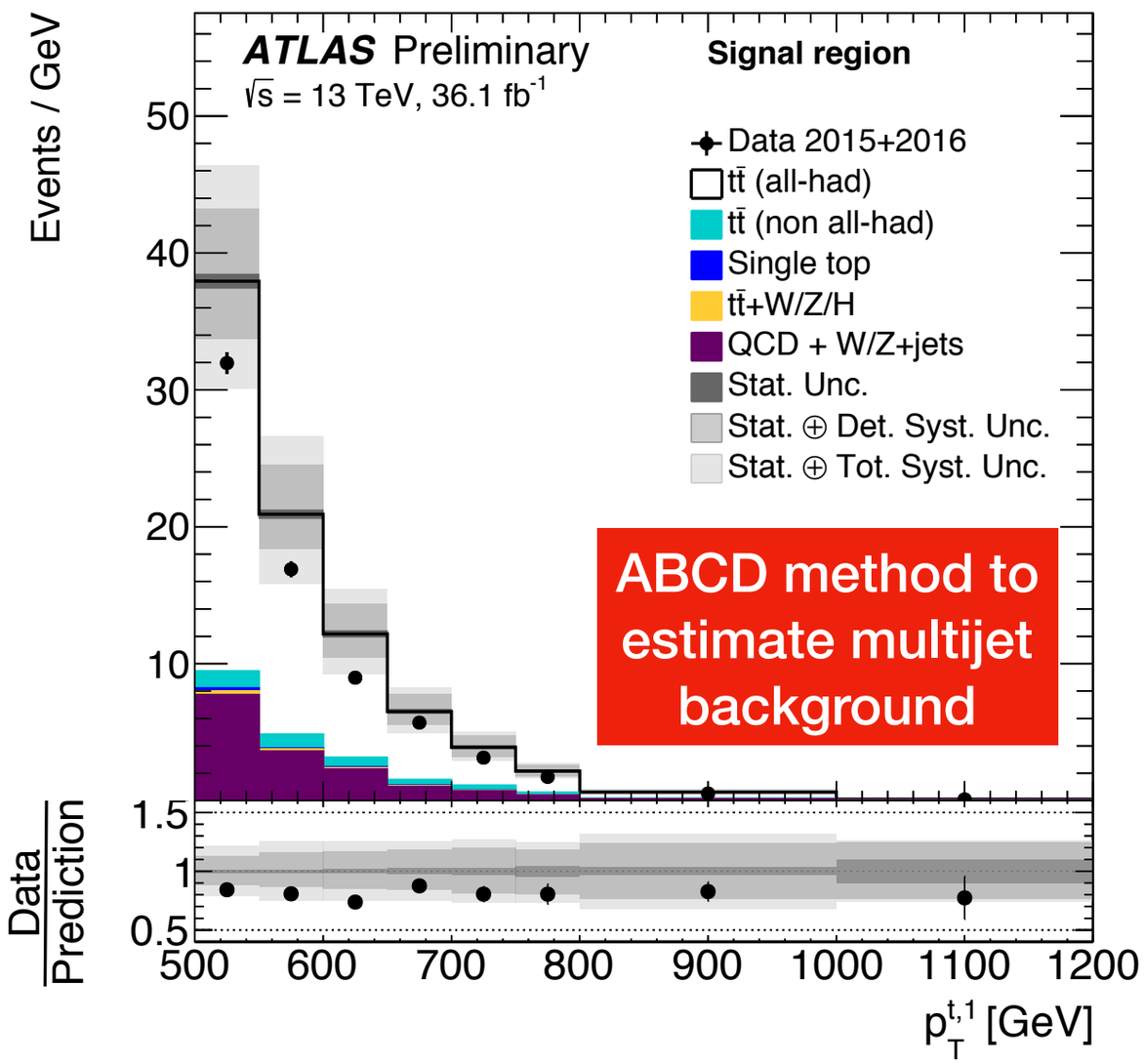
All-hadronic (13 TeV)

See talks by K. Cormier & M. Giannelli
See poster by P. Jacka

- **All-hadronic** event selection: two large-R jets used to identify top quarks
  - Trimmed (f_{cut}=0.05, R_{sub}=0.2), R=1.0 anti-k_T large-R jet p_T>500 (350) GeV
  - Top-tagged using p_T-dependent cuts on τ_{32} and large-R jet mass (50% ε)

122.5 < m < 222.5 [GeV]

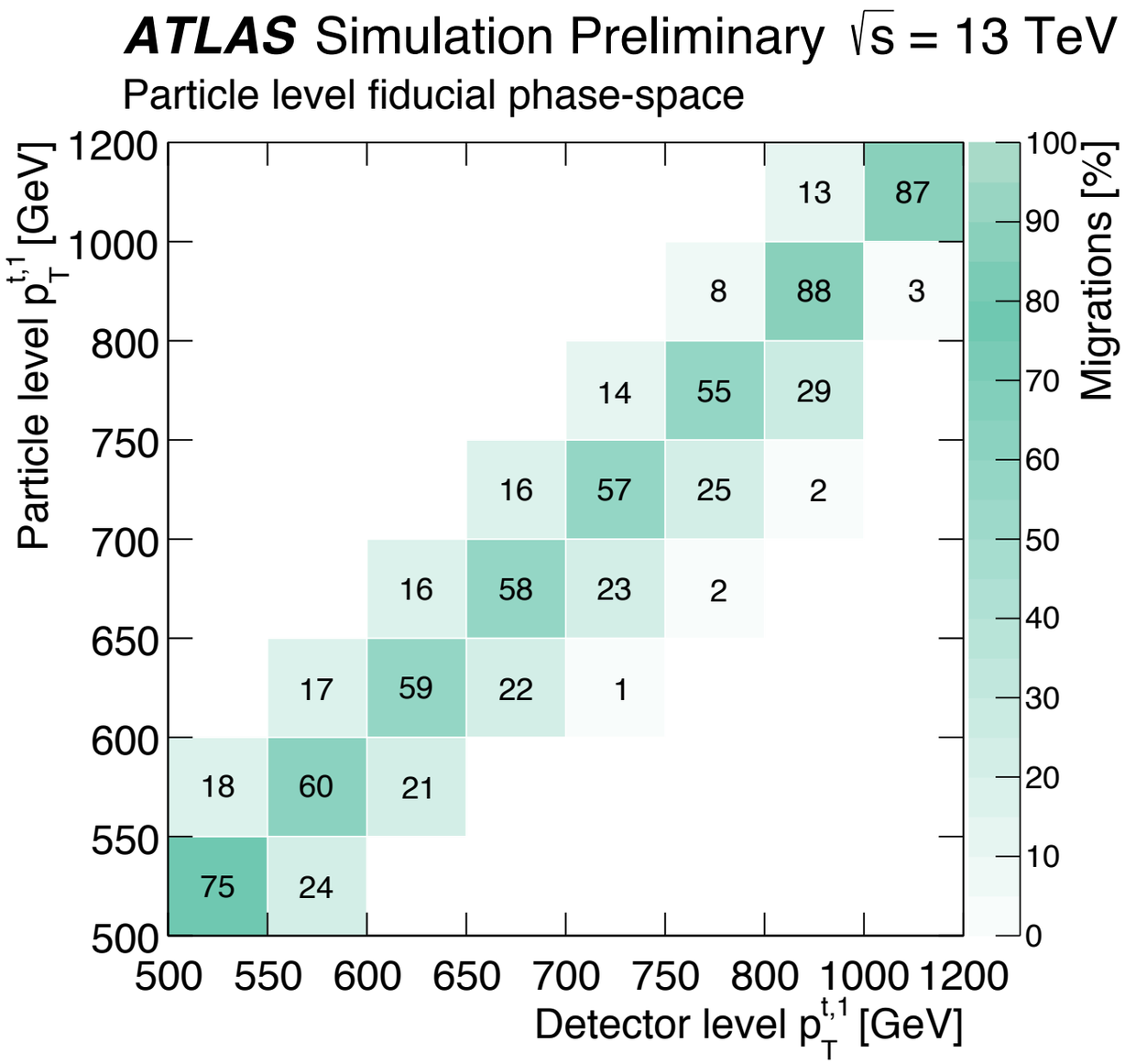




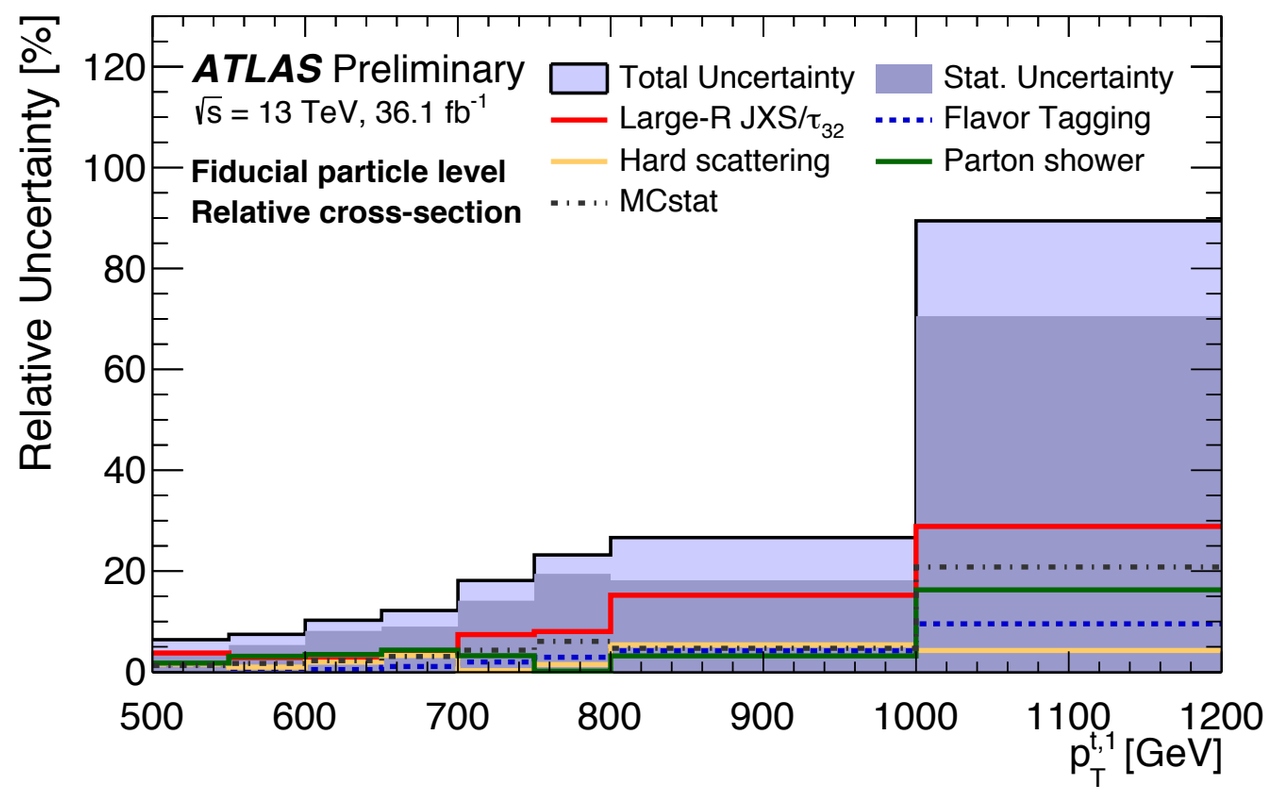
ABCD method to estimate multijet background

In preparation

All-hadronic (13 TeV)



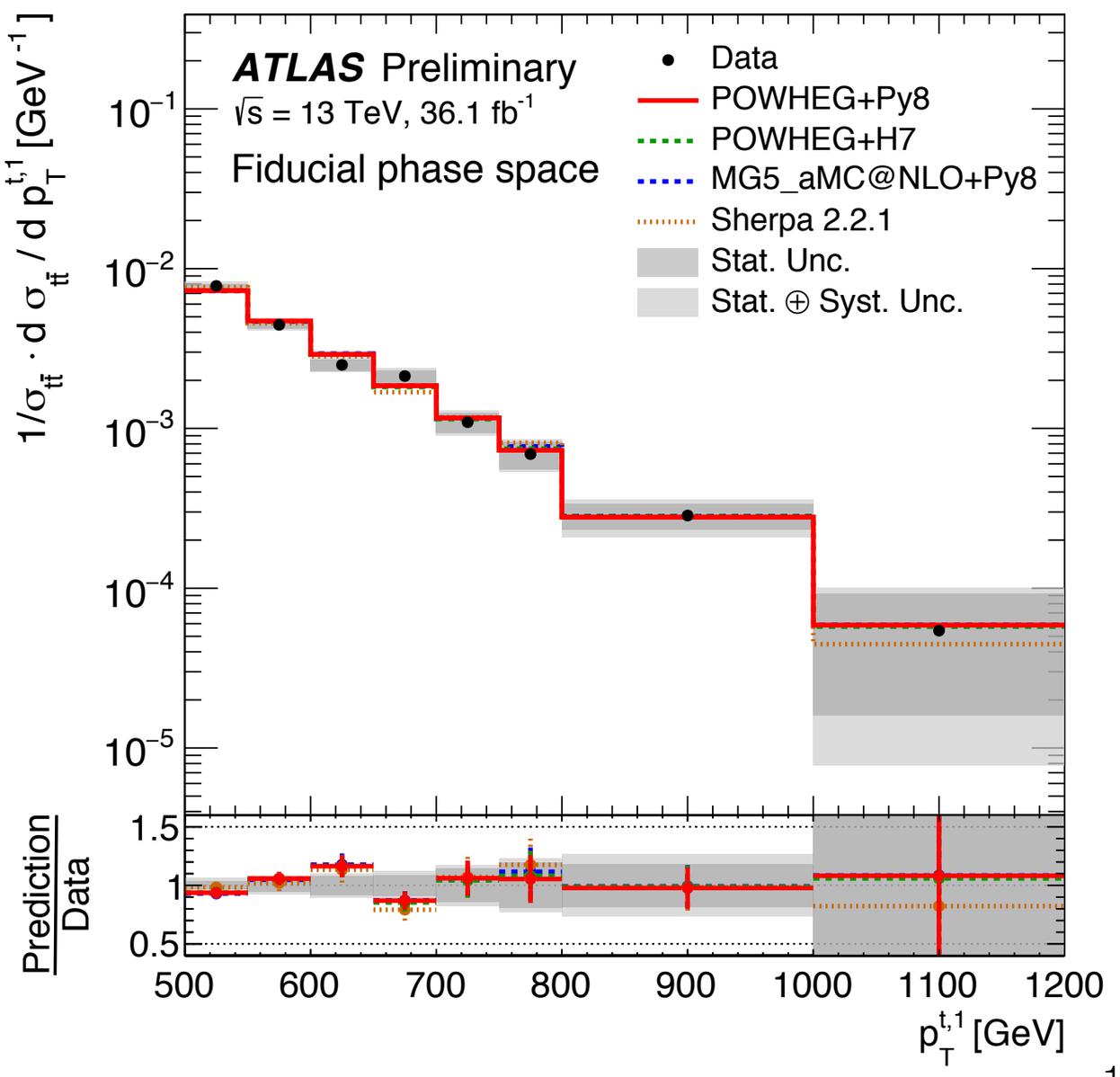
- Nearly diagonal response matrix in leading jet p_T
- Systematic uncertainties dominated by large-R jets and ttbar modeling

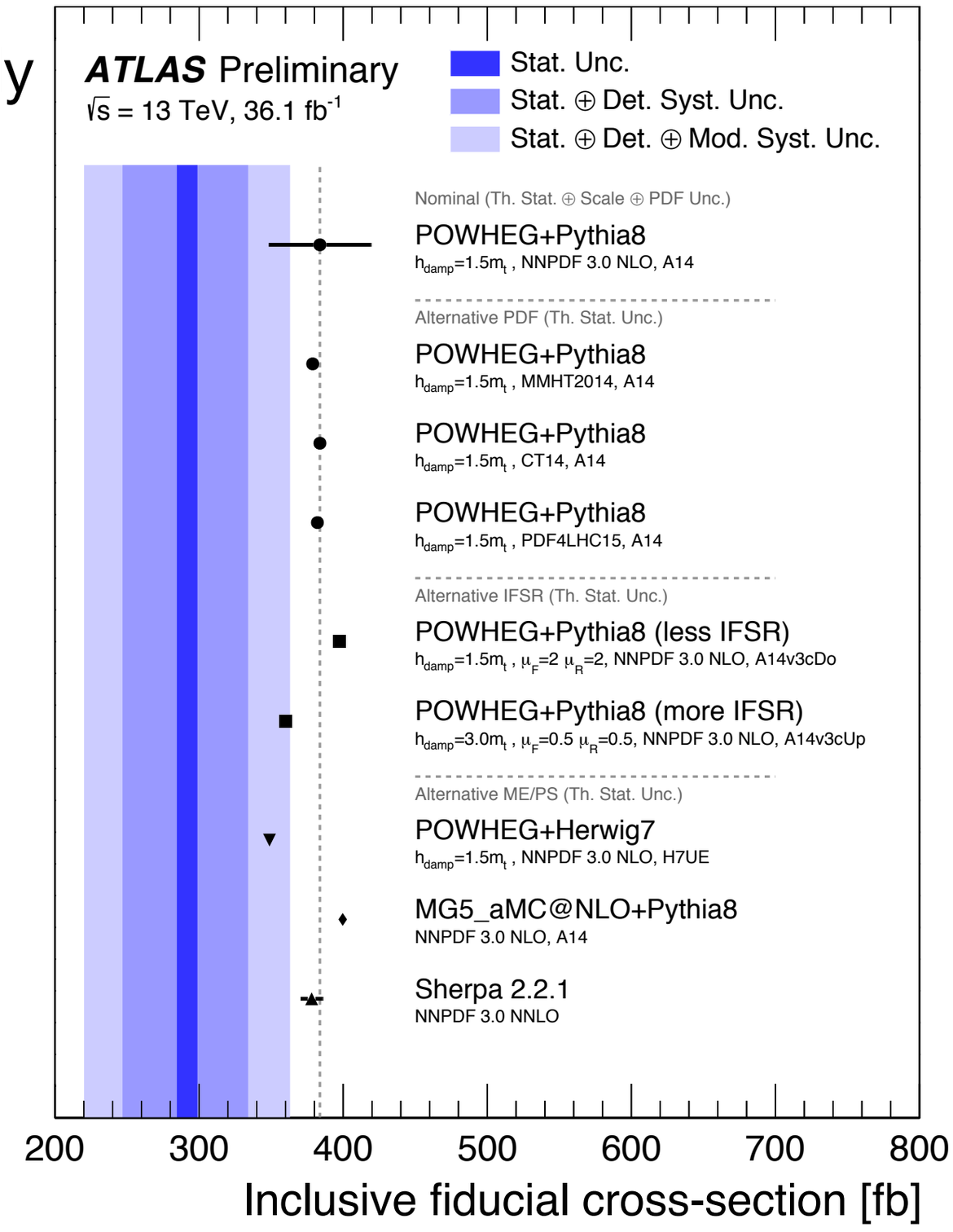


In preparation

All-hadronic (13 TeV)

- Iterative Bayesian unfolding inclusively and differentially to particle-level



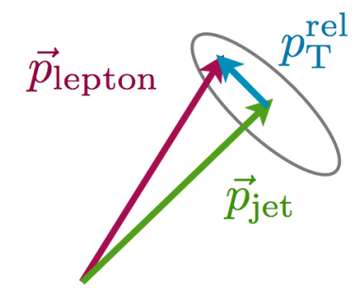


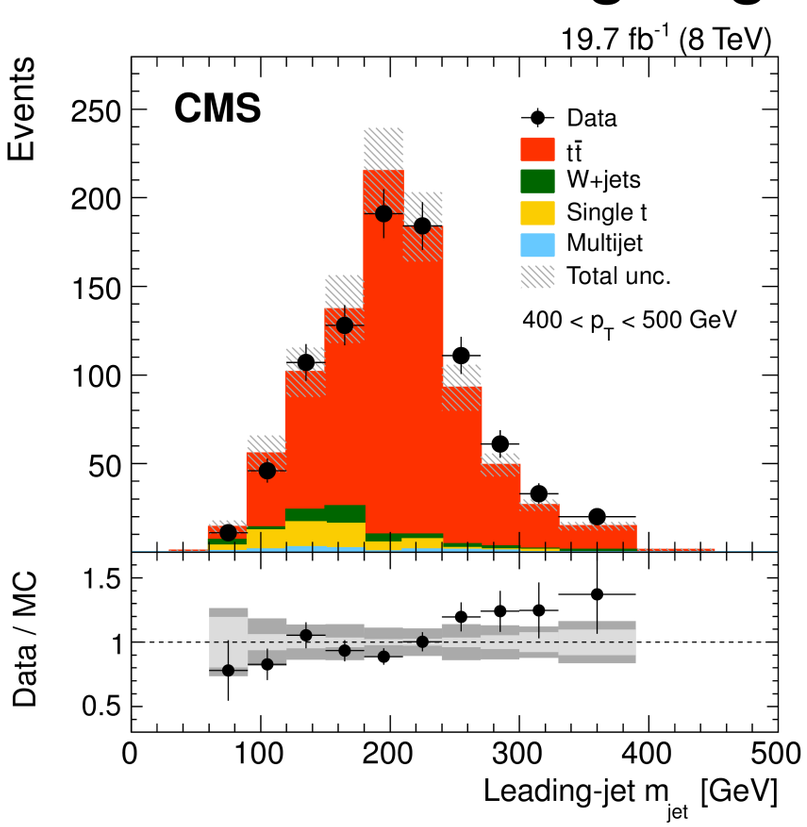
Eur. Phys. J. C 77 (2017) 467

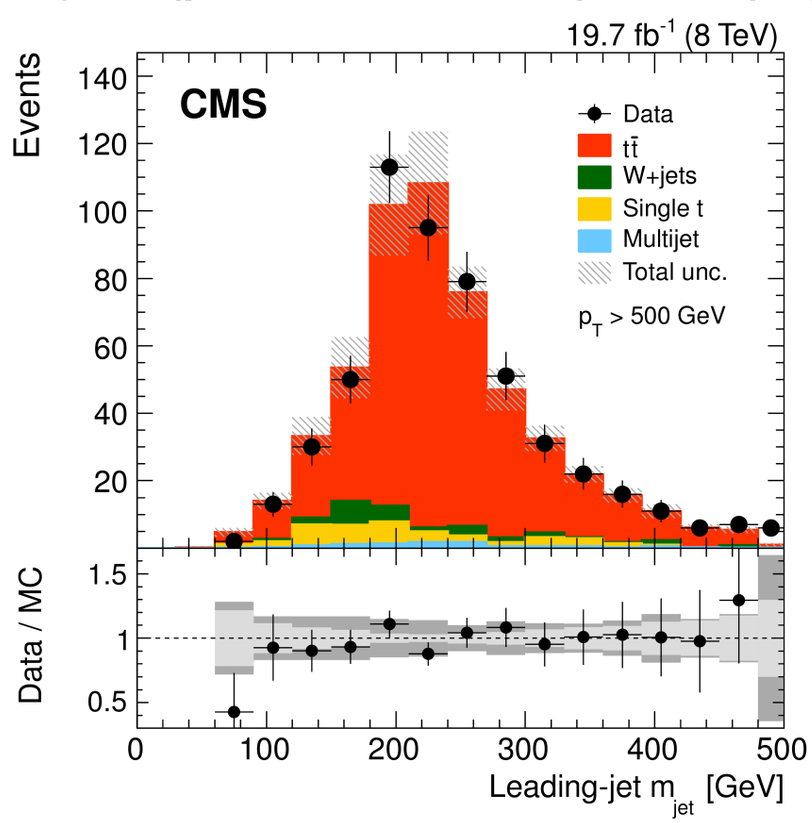
ℓ+jets (8 TeV)

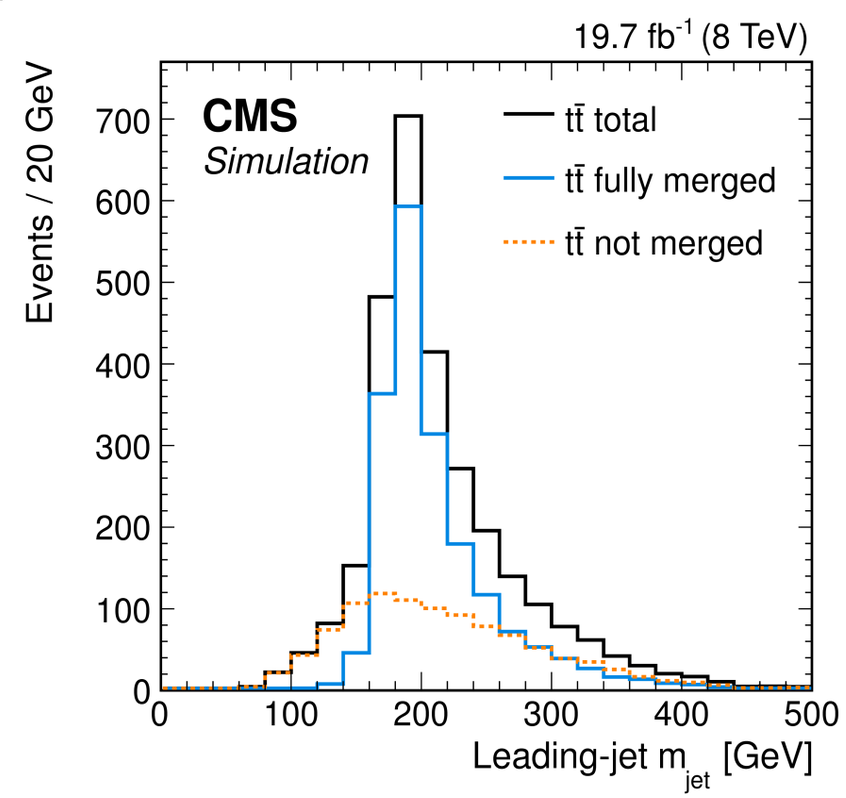
See talk by A. Castro

- **ℓ+jets** final state: differential cross section in the jet mass
- Hadronic Top: R=1.2 C/A large-R jet with p_T>400 GeV
  - **No selection on jet substructure!**
    Mass of leading jet required to be larger than sub-leading jet
- Leptonic Top:
  - high-p_T lepton (p_T>45 GeV; ΔR(ℓ,AK5) > 0.5 or p_T^{rel}>25 GeV) & sub-leading large-R jet (p_T>150 GeV): ΔR(ℓ,j) < 1.2









C/A = Cambridge/Aachen

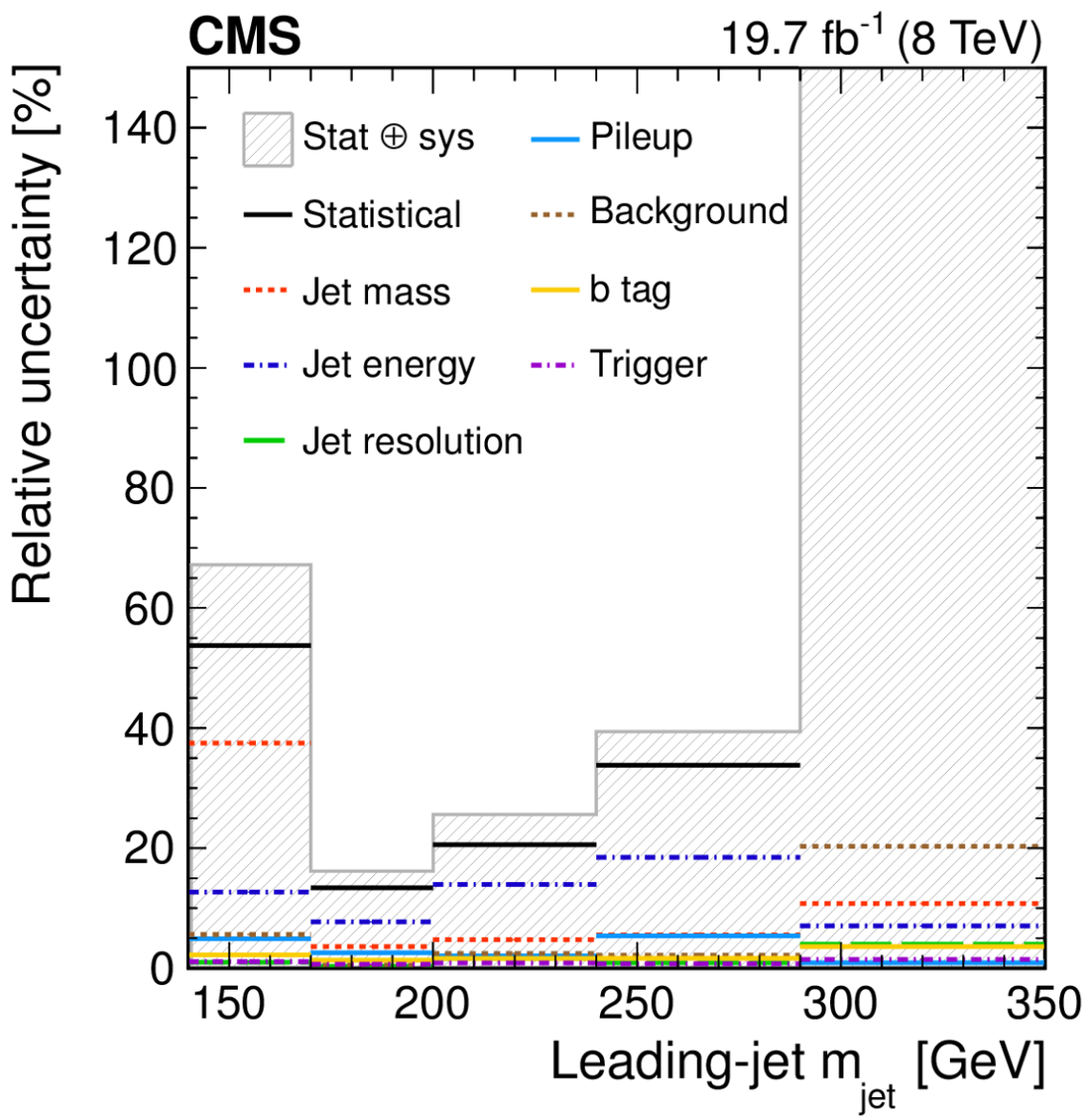
Particle-level p_T>400 GeV

Eur. Phys. J. C 77 (2017) 467

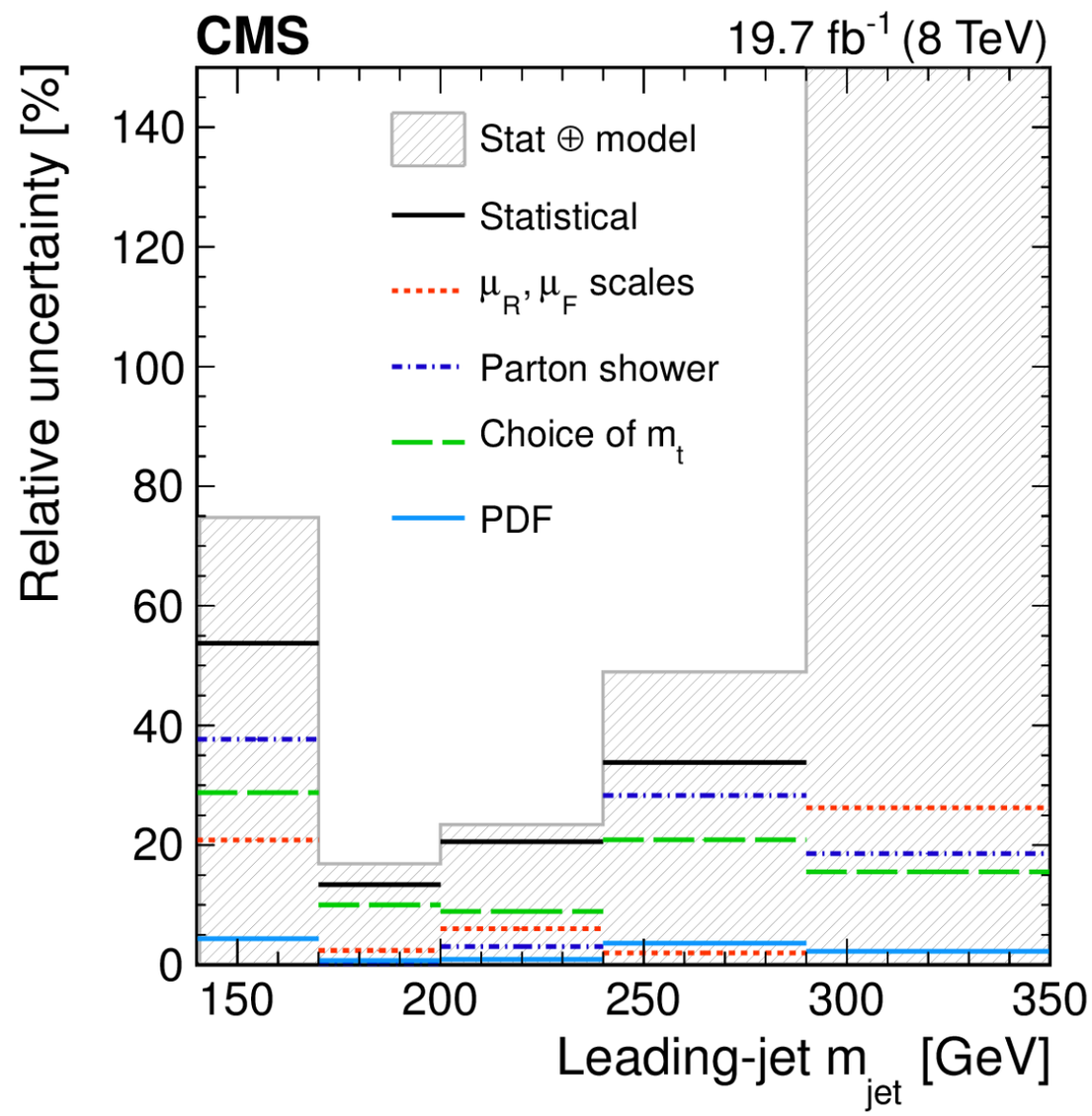
ℓ+jets (8 TeV)

- Systematic uncertainties dominated by large-R jets and ttbar modeling. Significant statistical uncertainties at large values for the jet mass

Detector-level



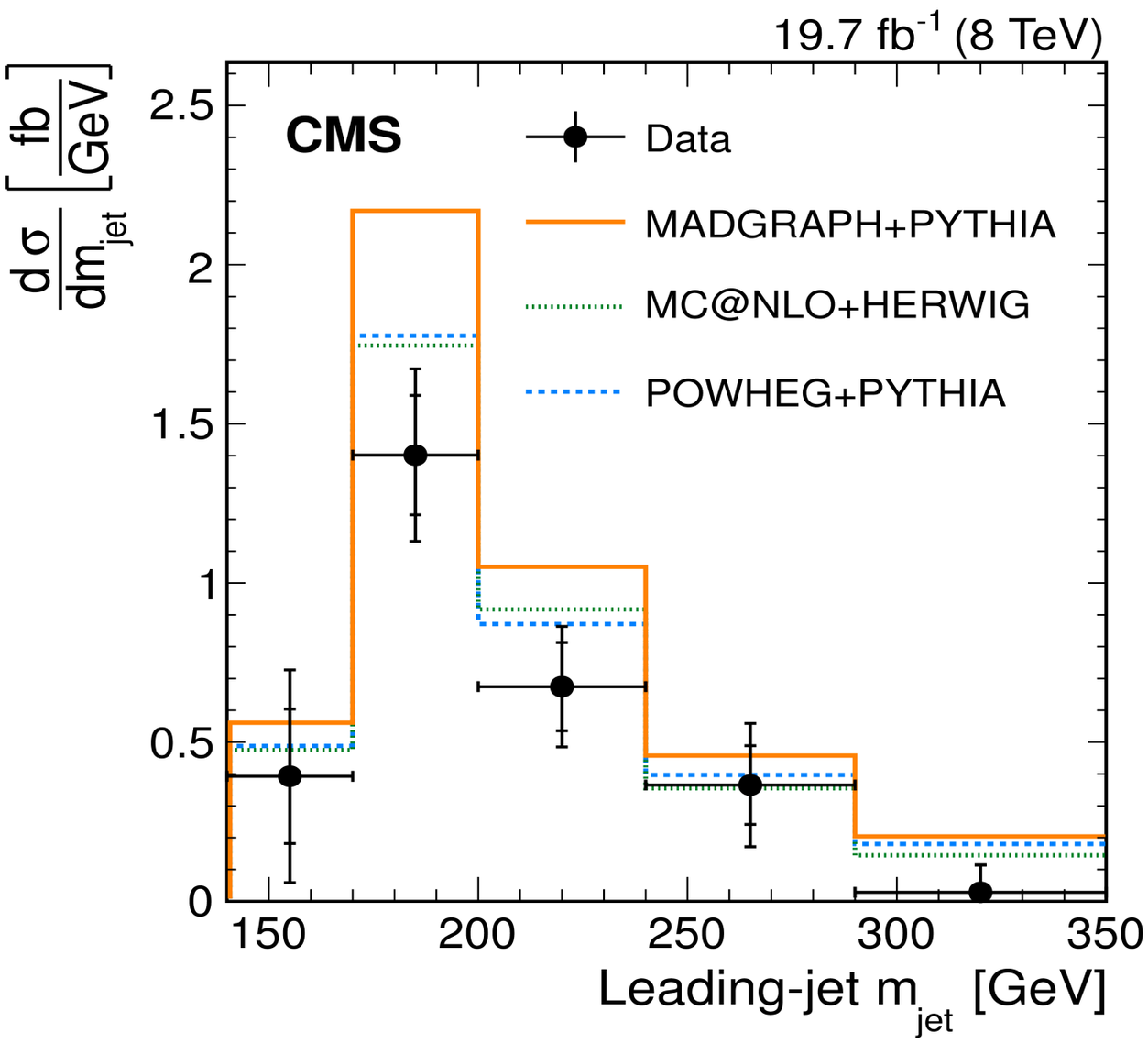
Modeling

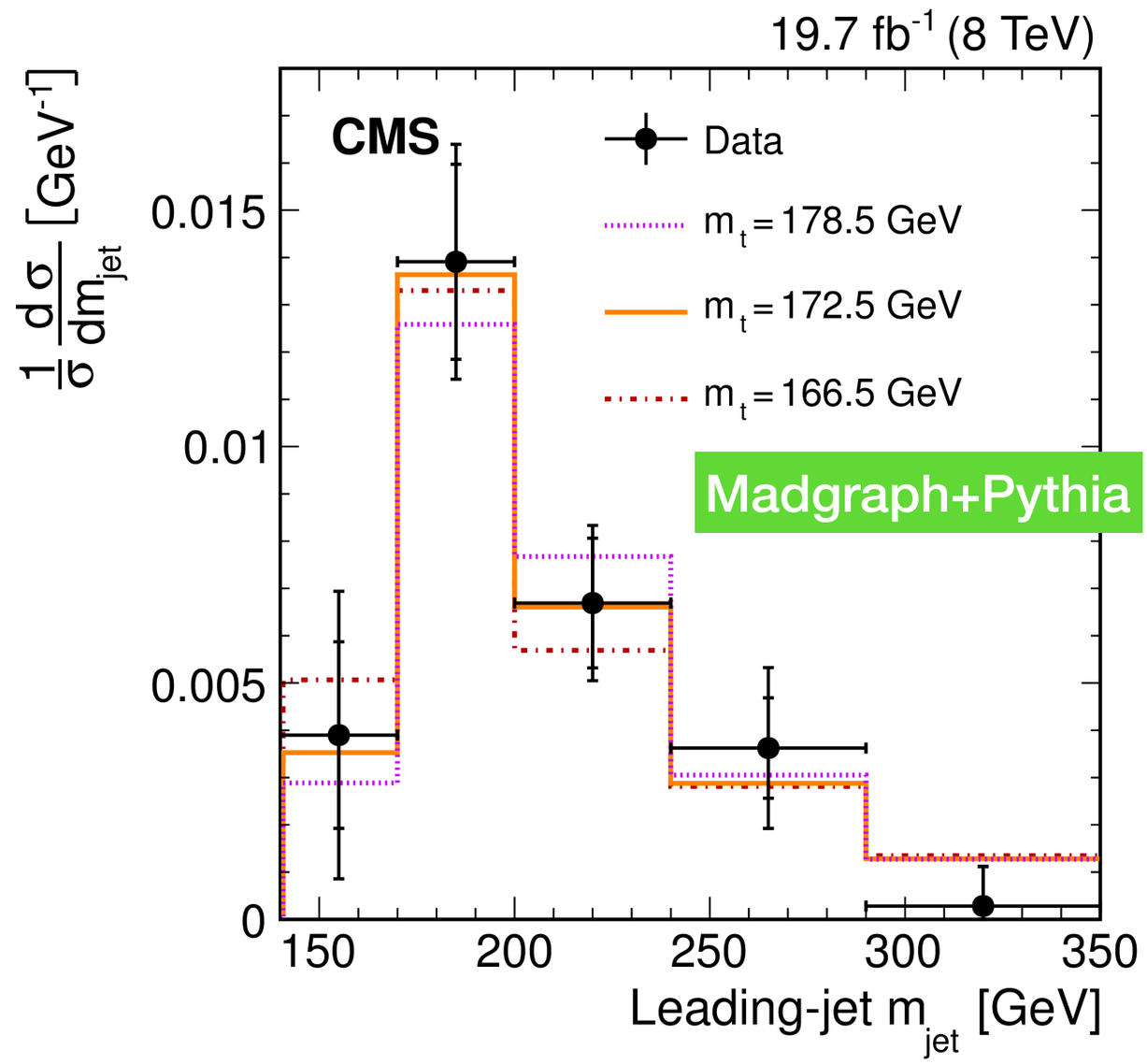


Eur. Phys. J. C 77 (2017) 467

ℓ+jets (8 TeV)

- Result unfolded to particle-level using TUnfold
- *Proof-of-concept to extract the top mass*!



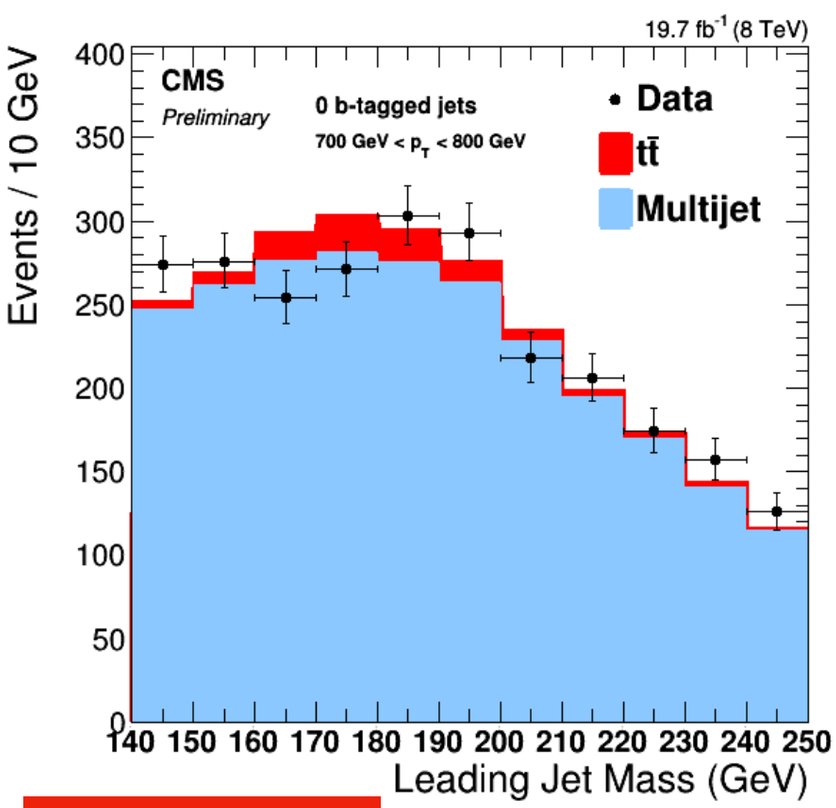


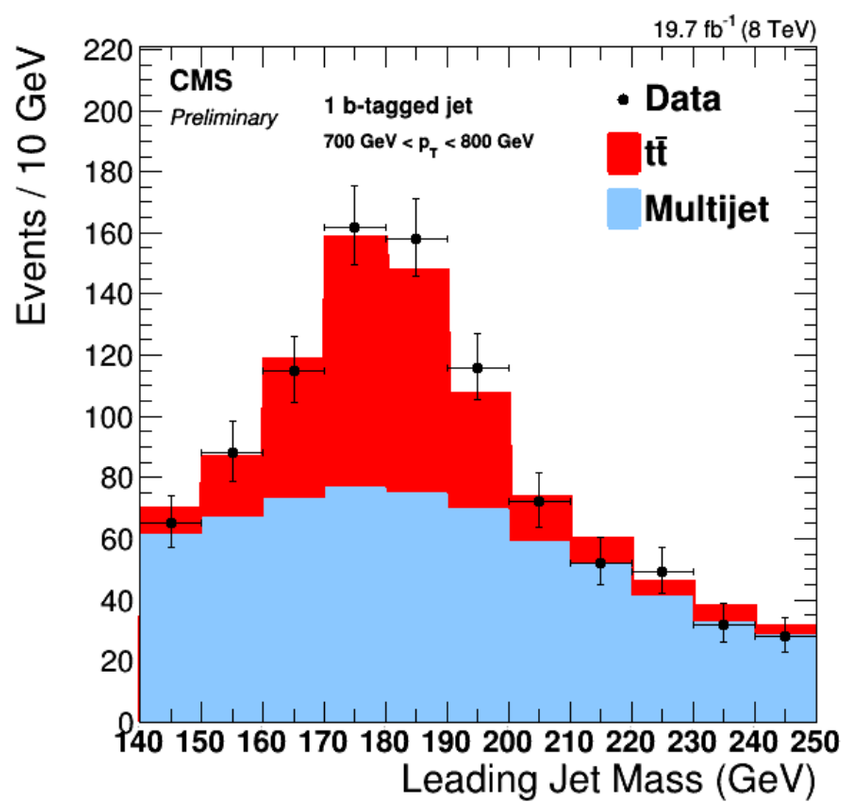
CMS-PAS-TOP-16-018

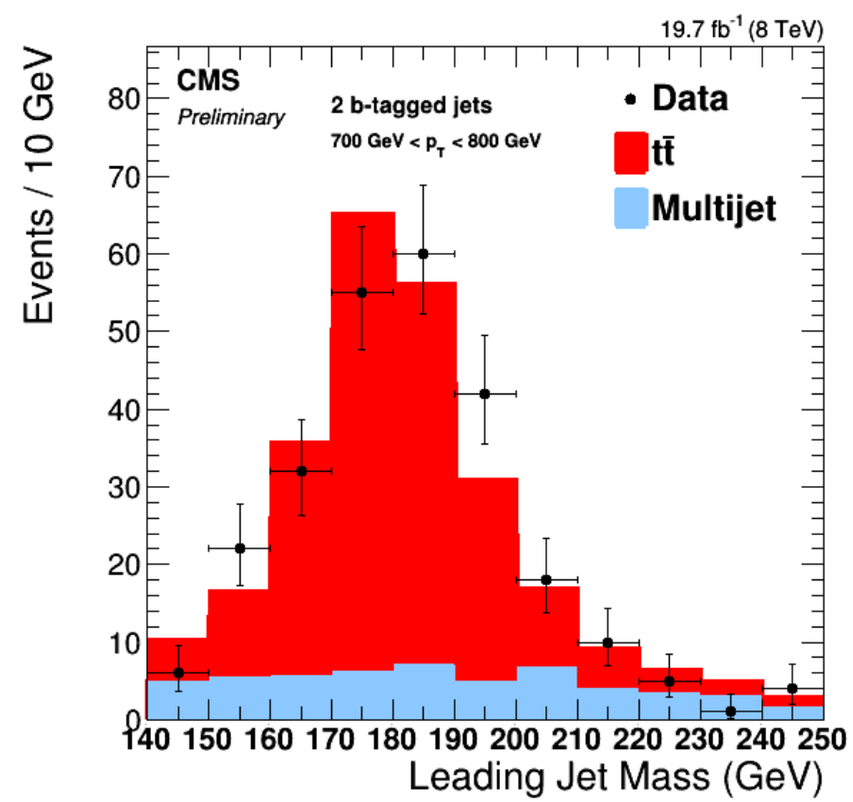
All-hadronic (8 TeV)

See talk by O. Hindrichs

- **All-hadronic** event selection: two C/A (R=0.8) large-R jets used to identify top quarks
- High threshold on H_T (>750 GeV) and large-R jet p_T (>400 GeV)
  - Top tagging with mass window 140-250 GeV and $\tau_{32}<0.55$ & $\tau_{21}>0.1$ (for leading jet)





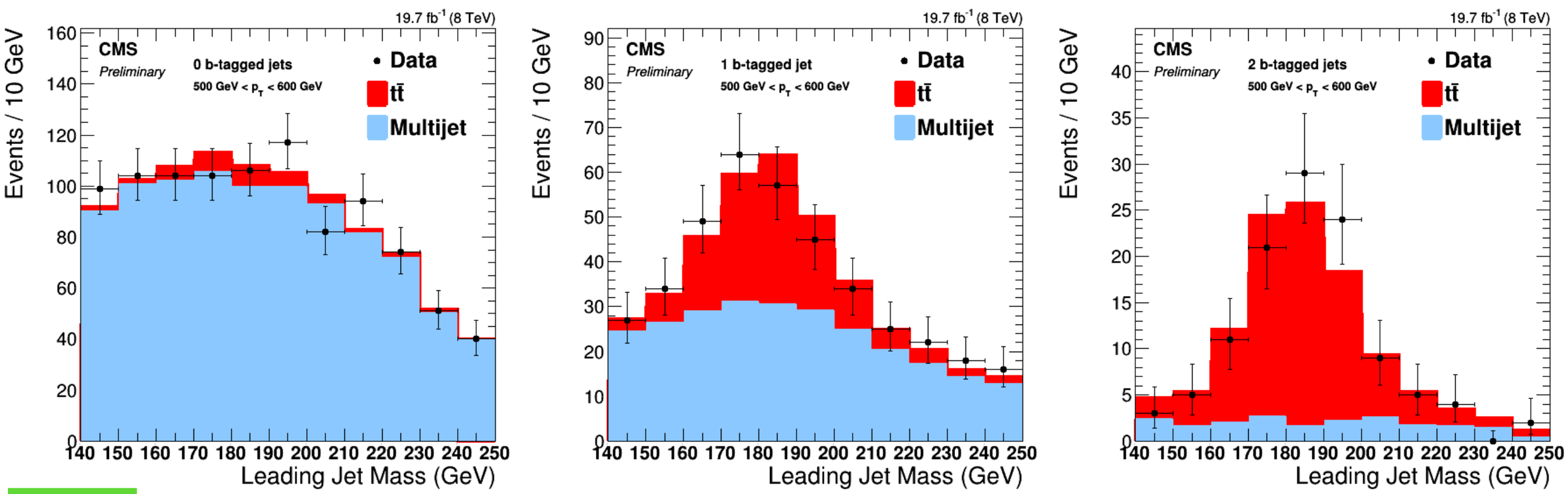


p_T > 500 GeV

CMS-PAS-TOP-16-018

All-hadronic (8 TeV)

- Data-driven technique using jet mass sideband and pass/fail τ_{32} regions (Alphabet method) to estimate the multijet background
- Ttbar yield is extracted using a maximum likelihood fit to the leading jet mass in signal enriched and depleted regions ($\tau_{32} \gtreqless 0.55$)

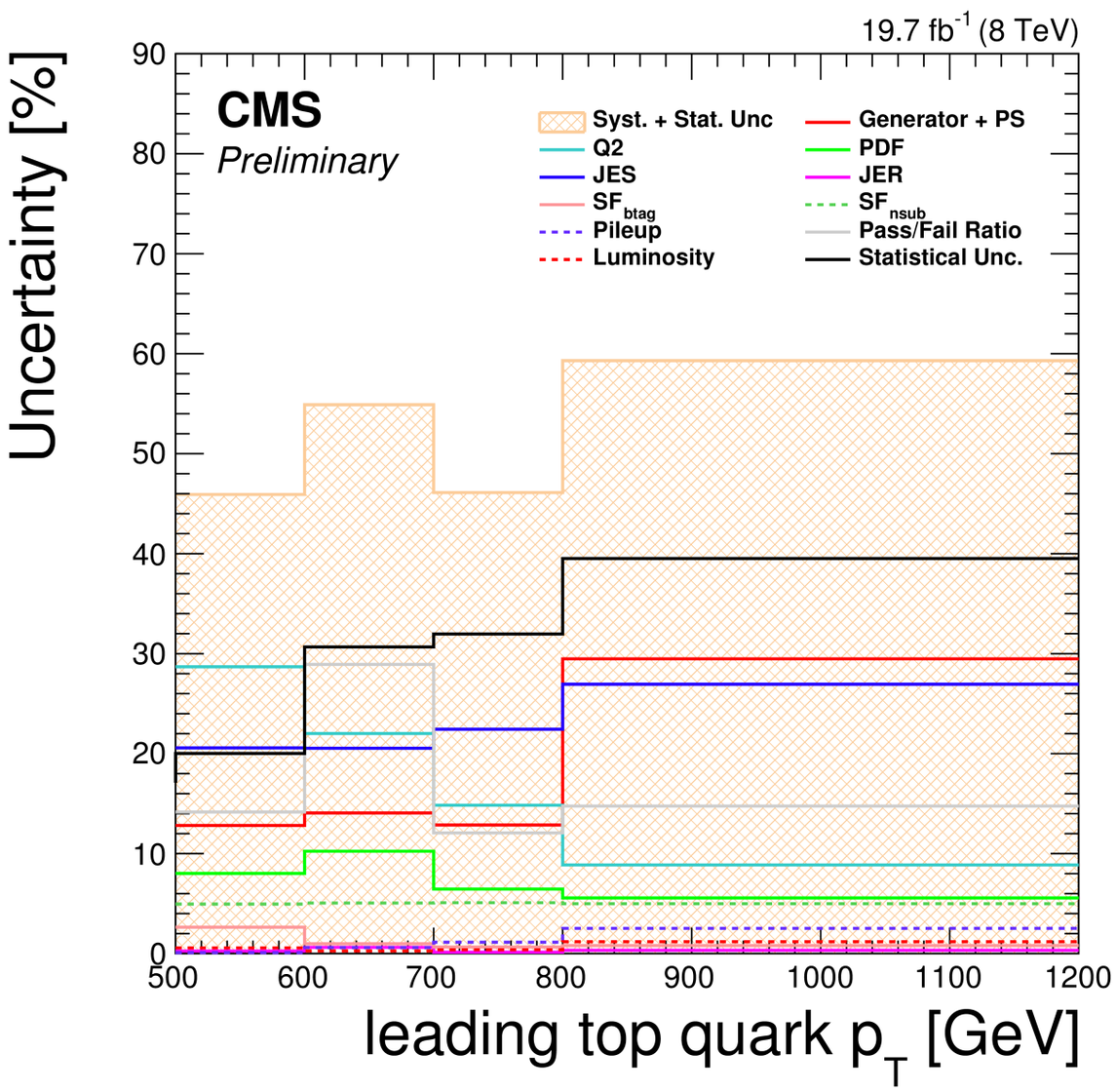


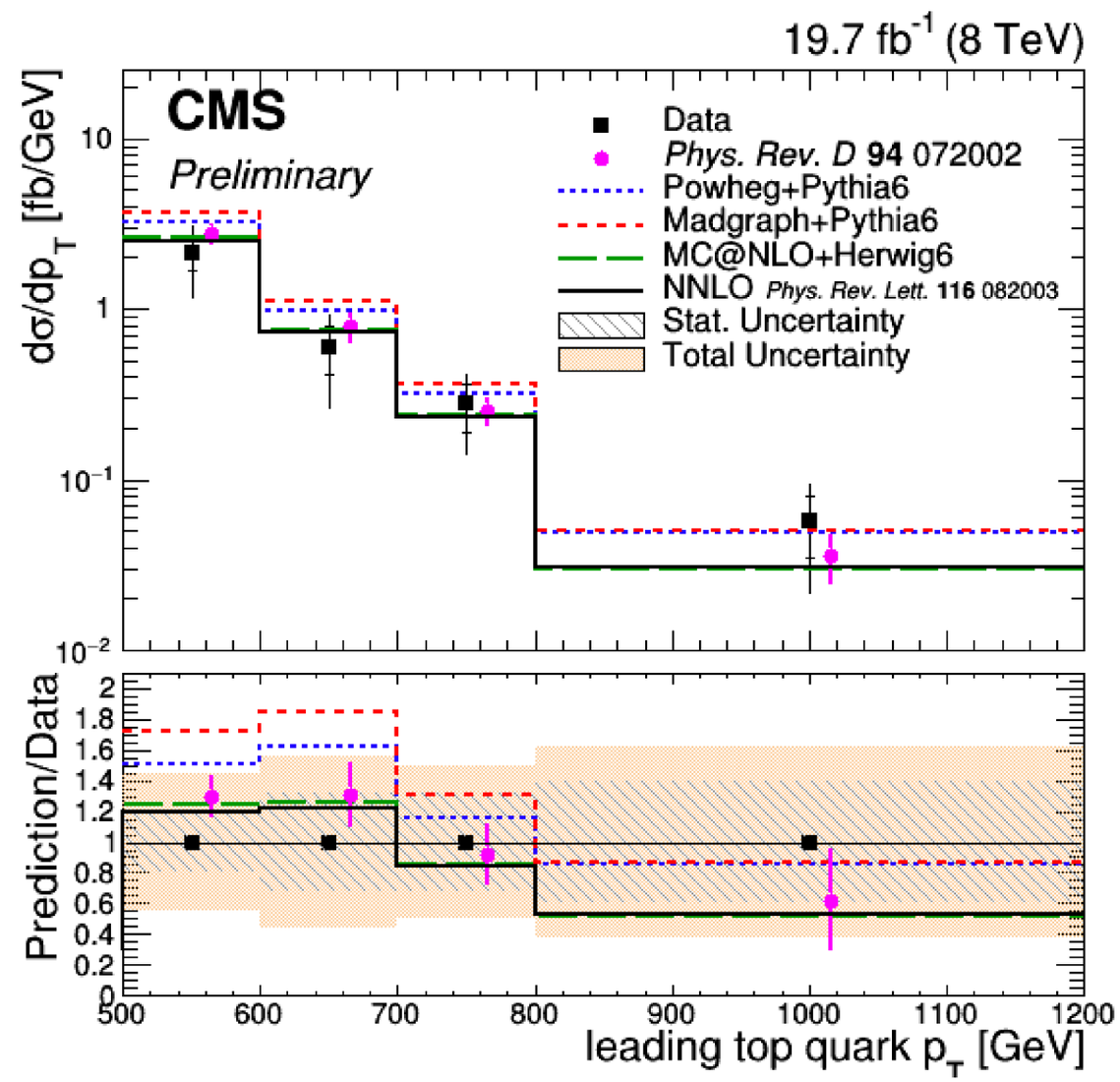
Post-fit

CMS-PAS-TOP-16-018

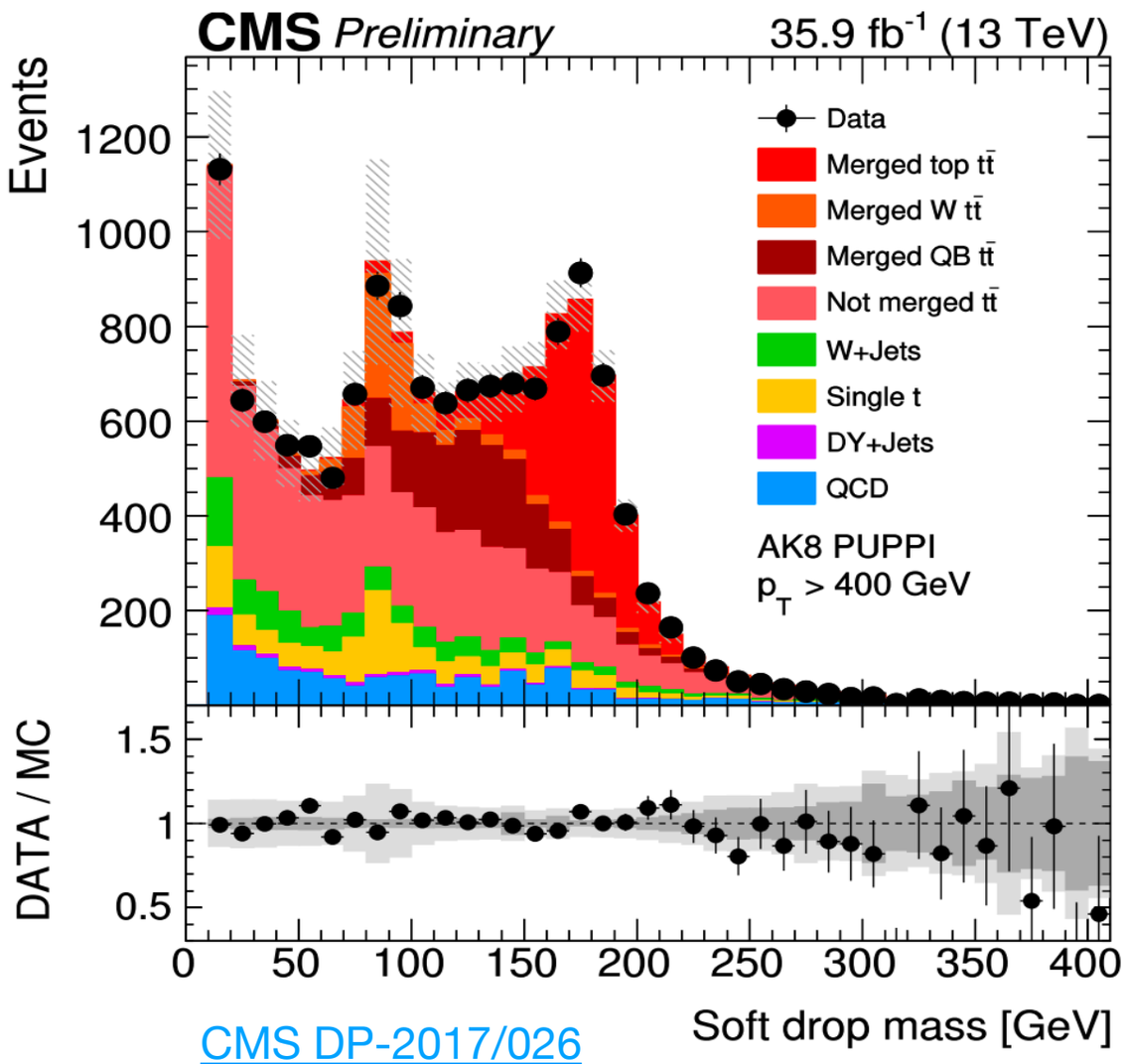
All-hadronic (8 TeV)

- Before unfolding, ttbar yield is extracted using likelihood fit
- SVD is used to unfold the result to parton-level
- Jet and ttbar modeling systematics dominate uncertainty

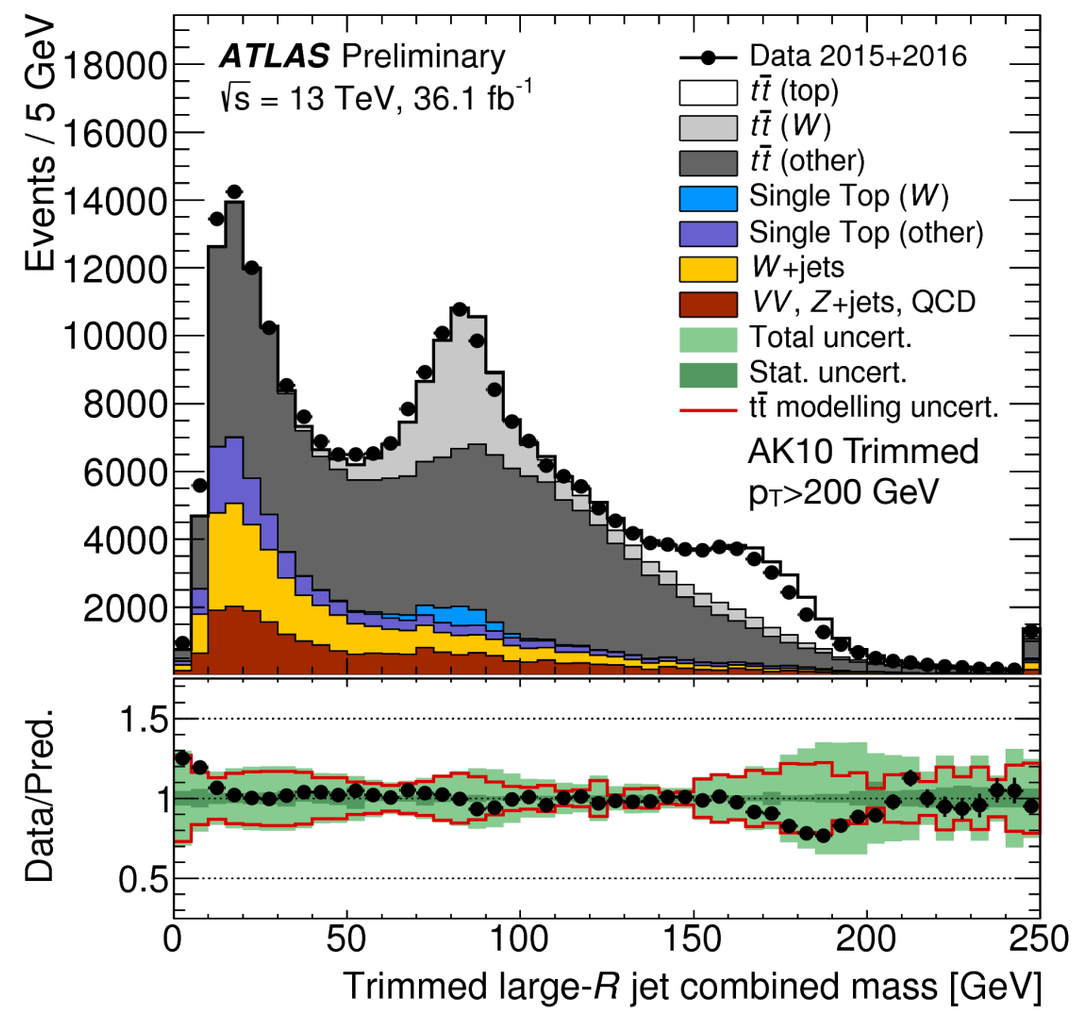




Outlook



CMS DP-2017/026



ATLAS-CONF-2017-064

Lots of on-going development in boosted top tagging!
Top tagging techniques largely developed for top vs qcd discrimination in searches (ttbar resonances, SUSY, VLQ, etc.).
For future measurements, we should expect top tagging algorithms (and boosted ttbar event selections) more tailored to ***measurements****.*

Conclusions

- Many exciting results from the LHC using boosted top quarks!

- Early 13 TeV results still suffer from significant uncertainties on the large-R jets & ttbar modeling

  - With more data being collected in 2017/2018, these results will become more precise and competitive with resolved measurements

- Upcoming boosted measurements of top properties can be used as a stronger test of the SM and complement many of the existing resolved measurements

Backup

More Info

For additional applications of boosted tops:

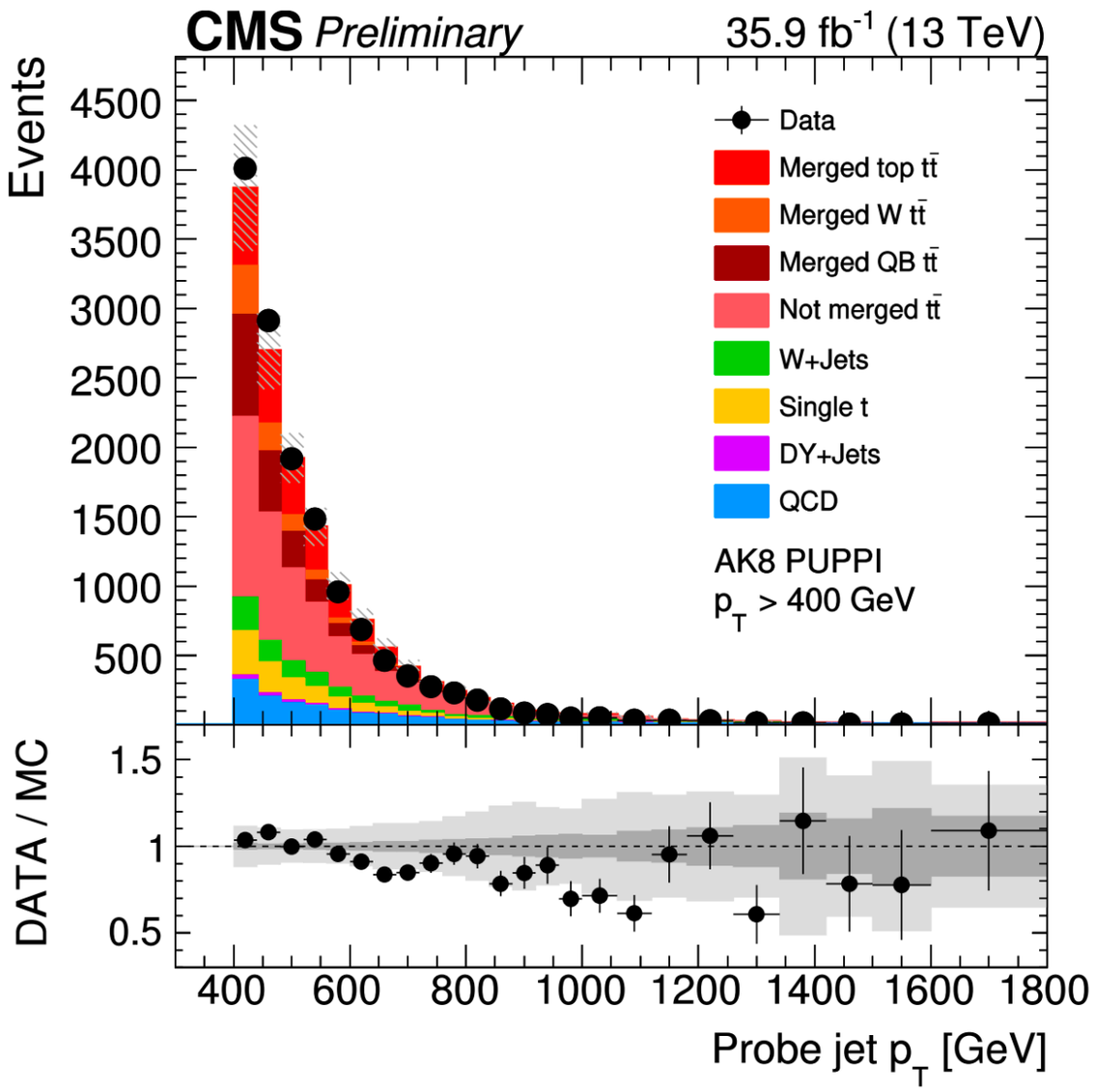
1. BOOST2017
    1. ATLAS
    2. CMS

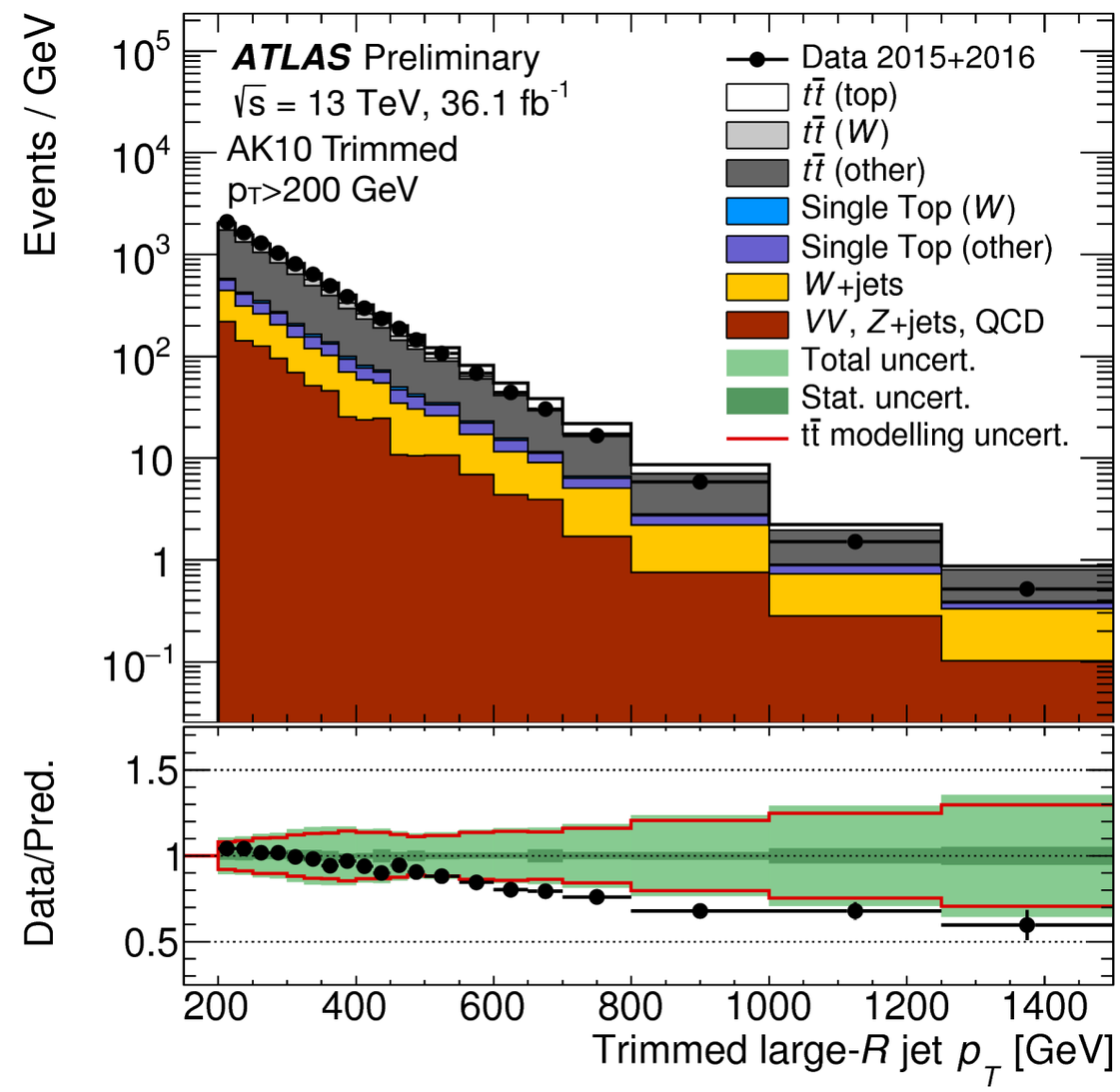
TOP2017

1. ATLAS Differential X-Sections (M. Giannelli)
2. CMS Differential X-Sections (O. Hindrichs)
3. SUSY/DM (D. Guest)
4. VLQ (G. Van Onsem)
5. YSF (S. Egan)

Previous presentations:

Top2016: (K. Kousouris)
Top2015 (L. Skinnari)


CMS Preliminary
35.9 fb-1 (13 TeV)
Events
Data
Merged top tt̄
Merged W tt̄
Merged QB tt̄
Not merged tt̄
W+Jets
Single t
DY+Jets
QCD
AK8 PUPPI
pT > 400 GeV
DATA / MC
Probe jet pT [GeV]


ATLAS Preliminary
√s = 13 TeV, 36.1 fb-1
AK10 Trimmed
pT>200 GeV
Data 2015+2016
tt̄ (top)
tt̄ (W)
tt̄ (other)
Single Top (W)
Single Top (other)
W+jets
VV, Z+jets, QCD
Total uncert.
Stat. uncert.
tt̄ modelling uncert.
Events / GeV
Data/Pred.
Trimmed large-R jet pT [GeV]

A_C (8 TeV)

The clouds of points in Fig. 3 correspond to a number of models in Refs. [80] ; [81]: a heavy W' boson exchanged in the t -channel, a heavy axi-gluon G_μ exchanged in the s-channel, and doublet (ϕ), triplet (ω^4) or sextet (Ω^4) scalars. Each point corresponds to a choice of the new particle's mass, in the range between 100 GeV and 10 TeV, and of the couplings to SM particles, where all values allowed give a total cross-section for top-quark pair production at the Tevatron compatible with observations and a high-mass $t\bar{t}$ production cross-section ($m_{t\bar{t}} > 1$ TeV) at the LHC that is at most three times the SM prediction. The contribution from new physics to the Tevatron A_{FB} is moreover required to be positive. The predictions of the Tevatron forward–backward asymmetry and the LHC high-mass charge asymmetry are calculated using PROTOS [83], which includes the tree-level SM amplitude plus the one(s) from the new particle(s), taking into account the interference between the two contributions. This measurement extends the reach of

[80] J. Aguilar-Saavedra, M. Perez-Victoria
Asymmetries in $t\bar{t}$ production: LHC versus Tevatron
Phys. Rev. D, 84 (2011), Article 115013 arXiv:1105.4606 [hep-ph]

CrossRef

[81] J. Aguilar-Saavedra, M. Perez-Victoria
Simple models for the top asymmetry: constraints and predictions
J. High Energy Phys., 09 (2011), Article 097 arXiv:1107.0841 [hep-ph]

ATLAS	8 TeV	13 TeV
	Differential Cross Section ℓ+jets (PRD 93 (2016) 032009) **Charge Asymmetry** ℓ+jets (PLB 756 (2016) 52)	**Differential Cross Section** ℓ+jets (arXiv:1708.00727) All-hadronic (In preparation)

CMS	8 TeV	13 TeV
	Differential Cross Section ℓ+jets (p_T PRD 94 (2016) 072002; jet mass: EPJC 77 (2017) 467) All-hadronic (CMS-PAS-TOP-16-018)	**Differential Cross Section** All-hadronic (CMS-PAS-TOP-16-013)

Top Tagger

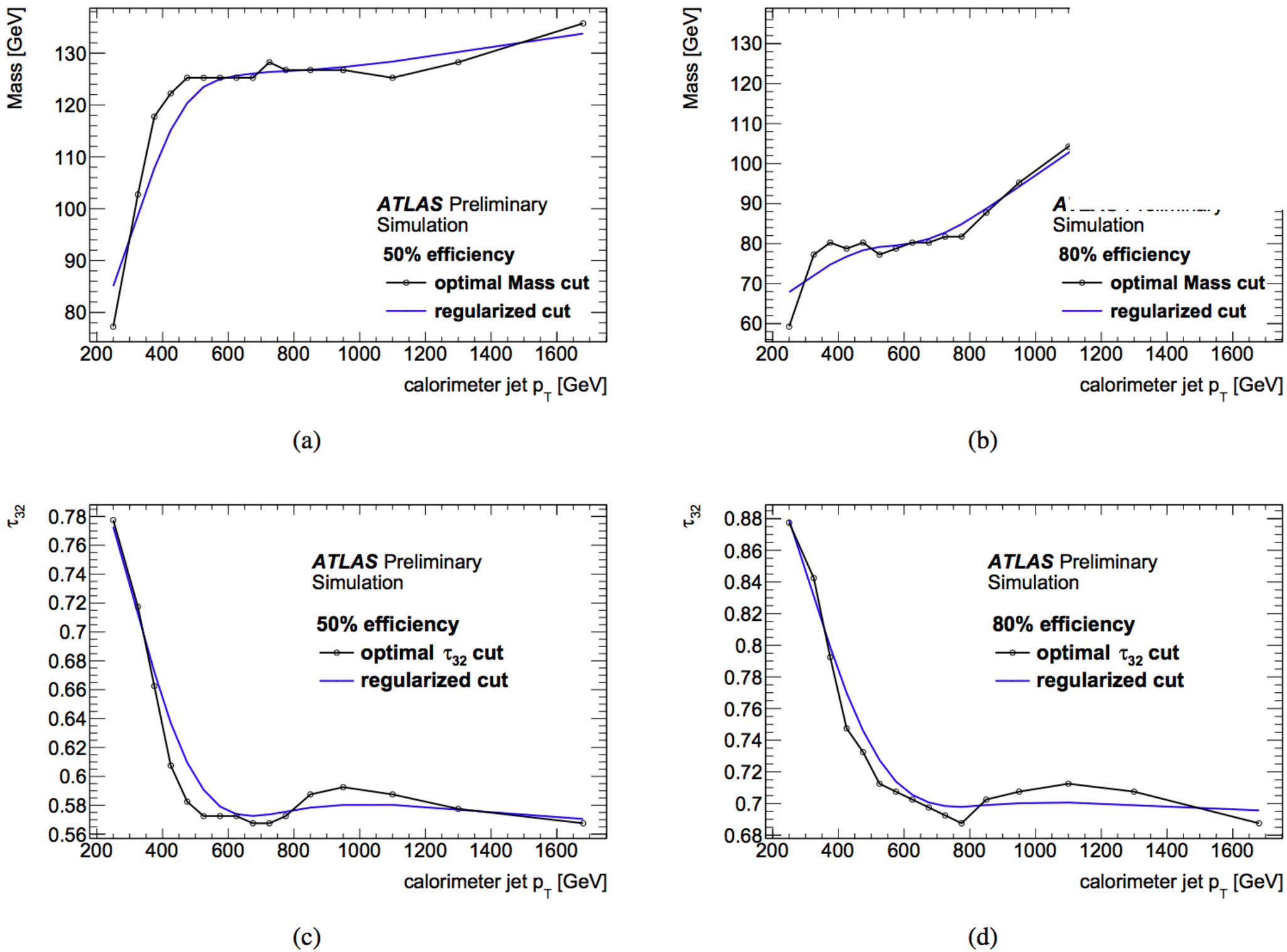


Figure 2: Values of the selection requirements in the top-tagging algorithm: (a) the lower threshold on $m_{\text{jet}}^{\text{calib}}$ for the 50% working point; (b) the lower threshold on $m_{\text{jet}}^{\text{calib}}$ for the 80% working point; (c) the upper threshold on τ_{32} for the 50% working point and (d) the upper threshold on τ_{32} for the 80% working point. The black points show the explicitly evaluated requirements, the red line shows the regularised interpolation.

arXiv:1708.00727

ℓ+jets (13 TeV)

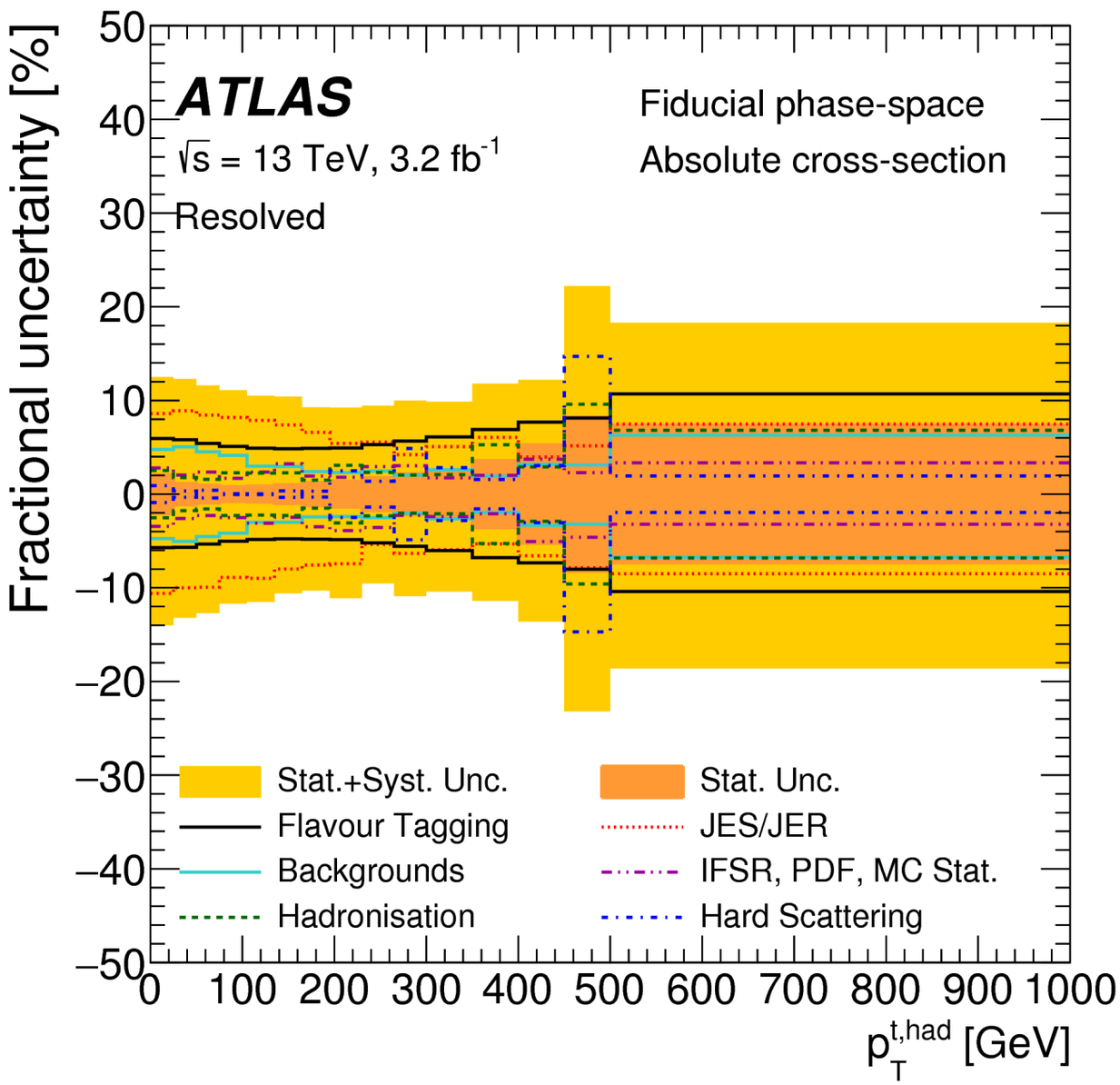
Level	Detector		Particle
Topology	Resolved	Boosted	
Leptons	$\lvert d_0\rvert/\sigma(d_0) < 5$ and $\lvert z_0 \sin\theta\rvert < 0.5$ mm Track and calorimeter isolation $\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.47$ (e), $\lvert\eta\rvert < 2.5$ (μ) E_{T} (e), p_{T} (μ) > 25 GeV		$\lvert\eta\rvert < 2.5$ $p_{\mathrm{T}} > 25$ GeV
Small-R jets	$\lvert\eta\rvert < 2.5$ $p_{\mathrm{T}} > 25$ GeV JVT cut (if $p_{\mathrm{T}} < 60$ GeV and $\lvert\eta\rvert < 2.4$)		$\lvert\eta\rvert < 2.5$ $p_{\mathrm{T}} > 25$ GeV
Num. of small-R jets	≥ 4 jets	≥ 1 jet	Same as detector level
$E_{\mathrm{T}}^{\mathrm{miss}}$, m_{T}^{W}		$E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV, $E_{\mathrm{T}}^{\mathrm{miss}} + m_{\mathrm{T}}^{W} > 60$ GeV	Same as detector level
Leptonic top	Kinematic top-quark reconstruction for detector and particle level	At least one small-R jet with $\Delta R(\ell$, small-R jet$) < 2.0$	
Hadronic top	Kinematic top-quark reconstruction for detector and particle level	The leading-p_{T} trimmed large-R jet has: $\lvert\eta\rvert < 2.0$, 300 GeV $< p_{\mathrm{T}} <$ 1500 GeV, $m > 50$ GeV, Top-tagging at 80% efficiency ΔR(large-R jet, small-R jet associated with lepton) > 1.5, $\Delta\phi(\ell$, large-R jet$) > 1.0$	**Boosted:** $\lvert\eta\rvert < 2.0$ $300 < p_{\mathrm{T}} < 1500$ GeV Top-tagging: $m > 100$ GeV, $\tau_{32} < 0.75$
b-tagging	At least 2 b-tagged jets	At least one of: 1) the leading-p_{T} small-R jet with $\Delta R(\ell$, small-R jet$) < 2.0$ is b-tagged 2) at least one small-R jet with ΔR(large-R jet, small-R jet$) < 1.0$ is b-tagged	Ghost-matched b-hadron

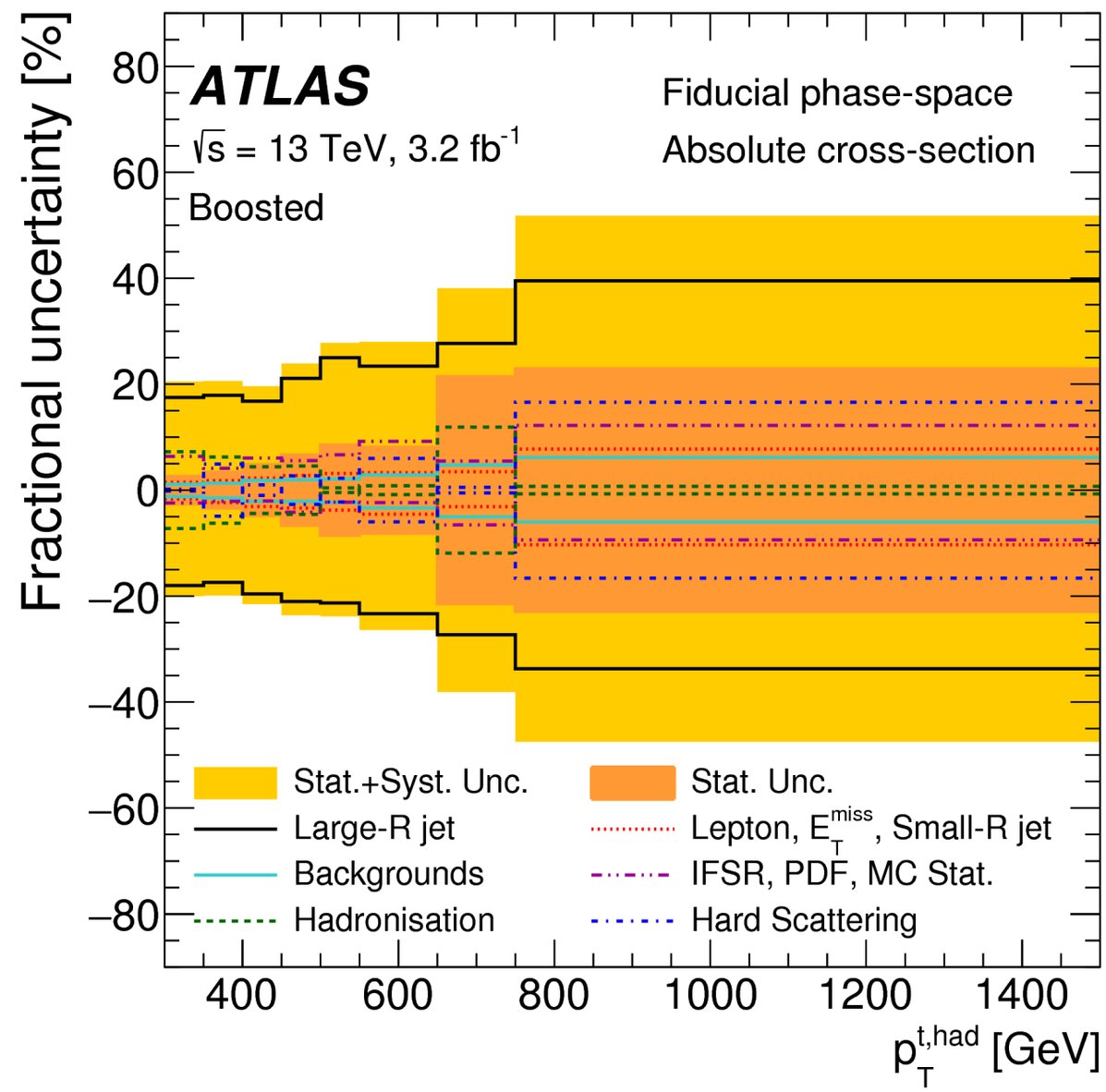
Process	Expected events	
	Resolved	Boosted
$t\bar{t}$	123800 ± 10600	7000 ± 1100
Single top	6300 ± 800	500 ± 80
Multijets	5700 ± 3000	300 ± 80
W+jets	$3600\ ^{+2000}_{-2400}$	500 ± 200
Z+jets	1300 ± 700	60 ± 40
$t\bar{t}V$	400 ± 100	70 ± 10
Diboson	300 ± 200	60 ± 10
Total prediction	$142000\ ^{+11000}_{-12000}$	8300 ± 1300
Data	155593	7368

 # Boosted = 5% # Resolved

arXiv:1708.00727

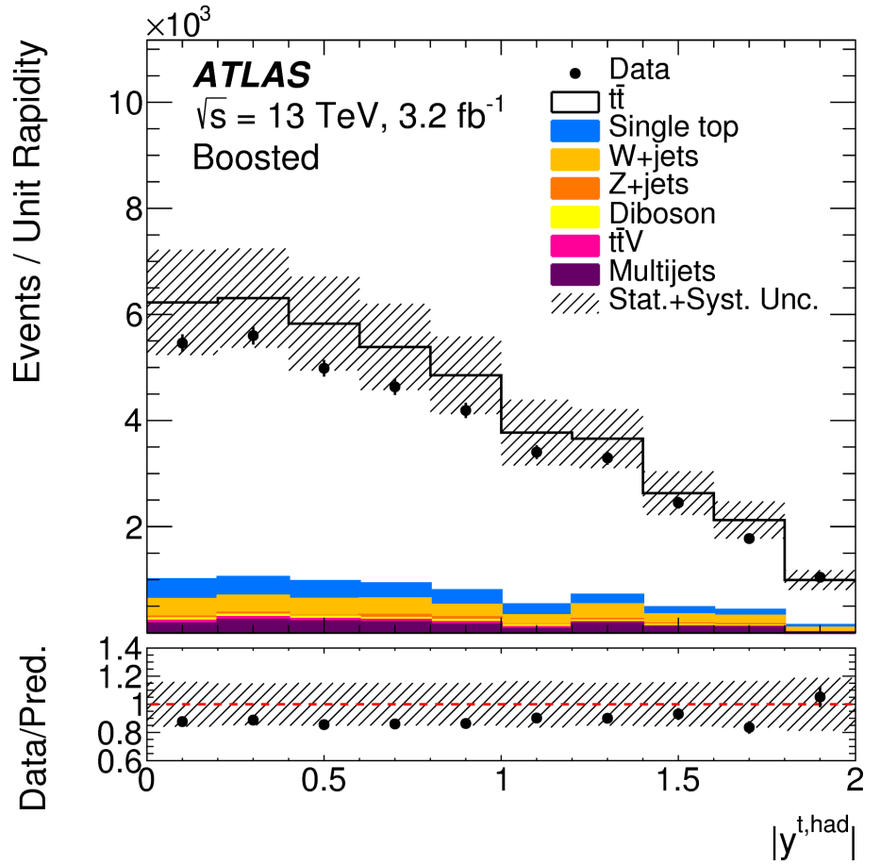
ℓ+jets (13 TeV)

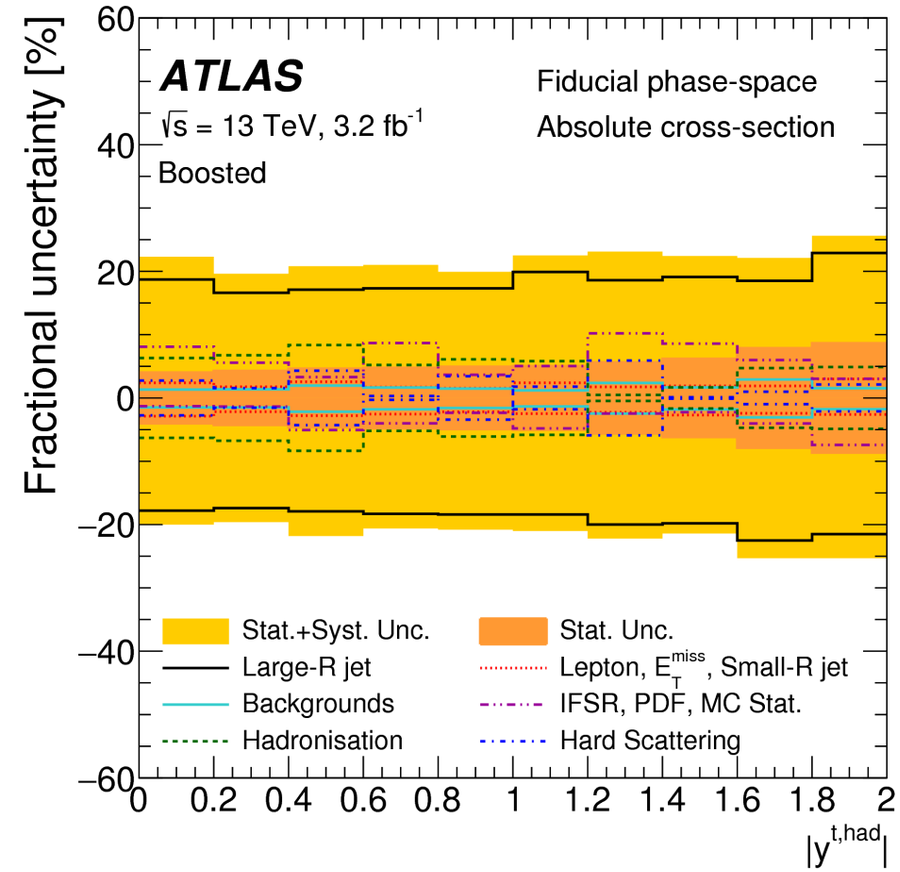


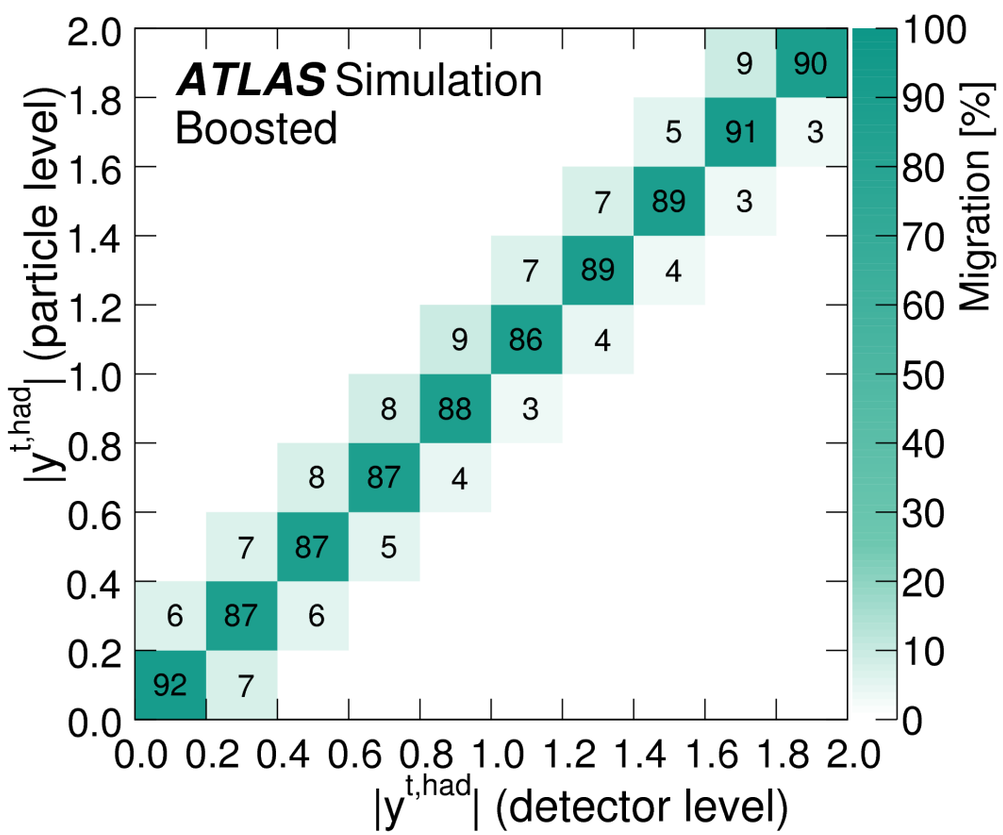


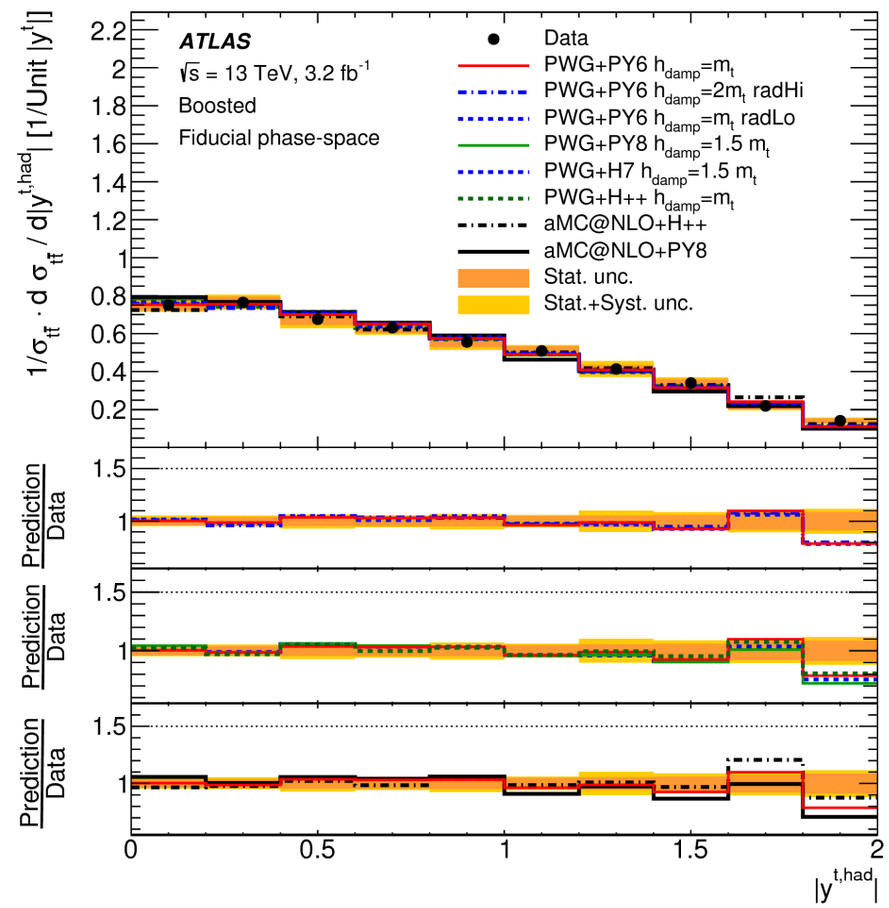
ℓ+jets (13 TeV)

arXiv:1708.00727

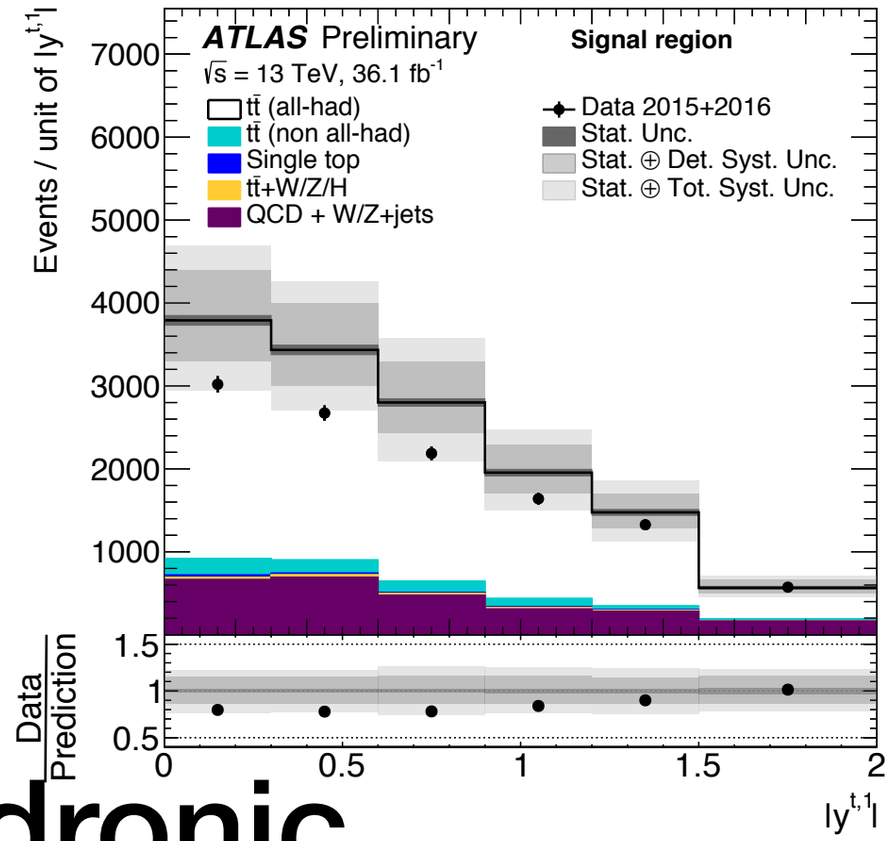


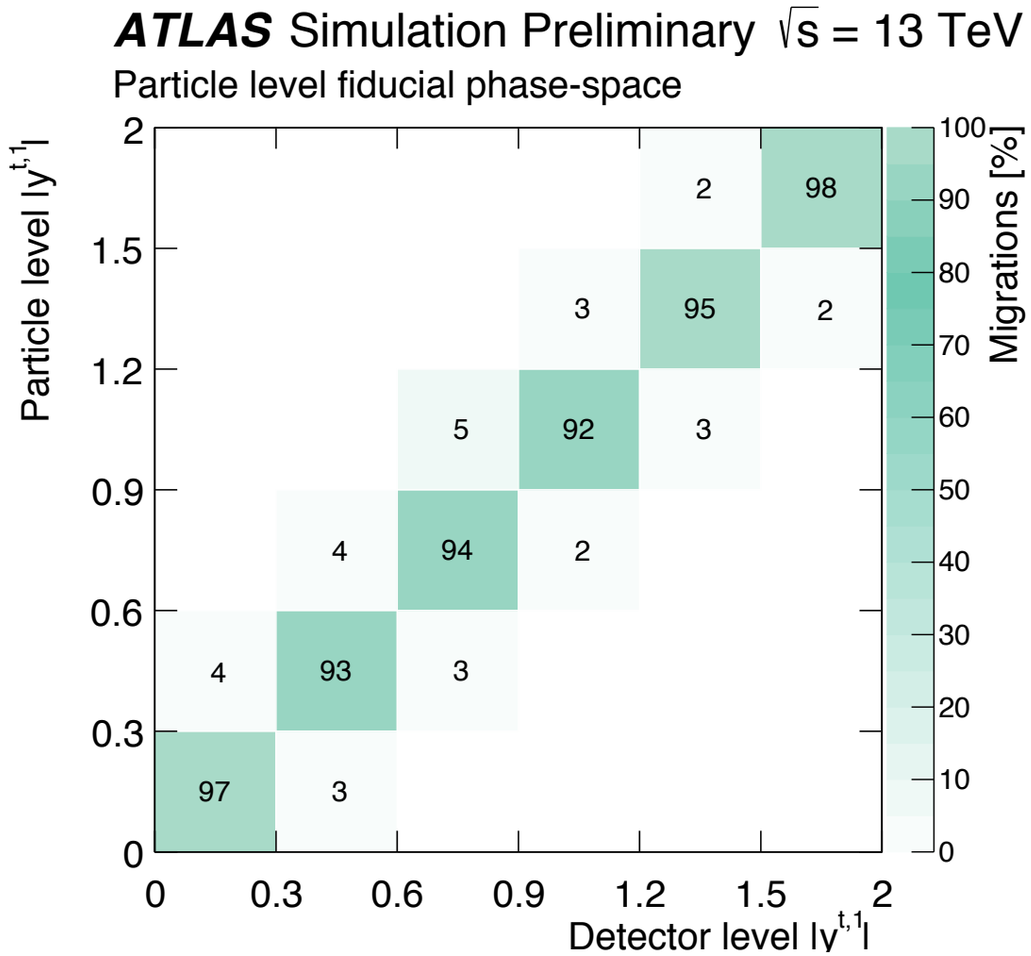






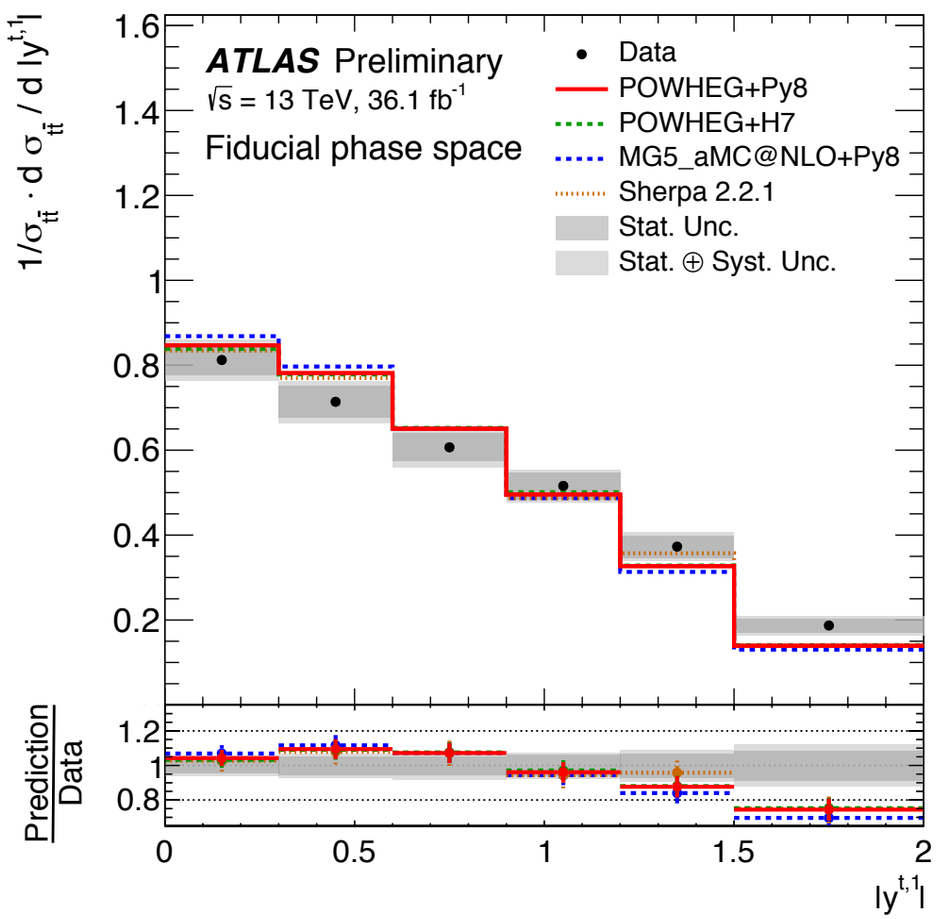
All-hadronic (13 TeV)



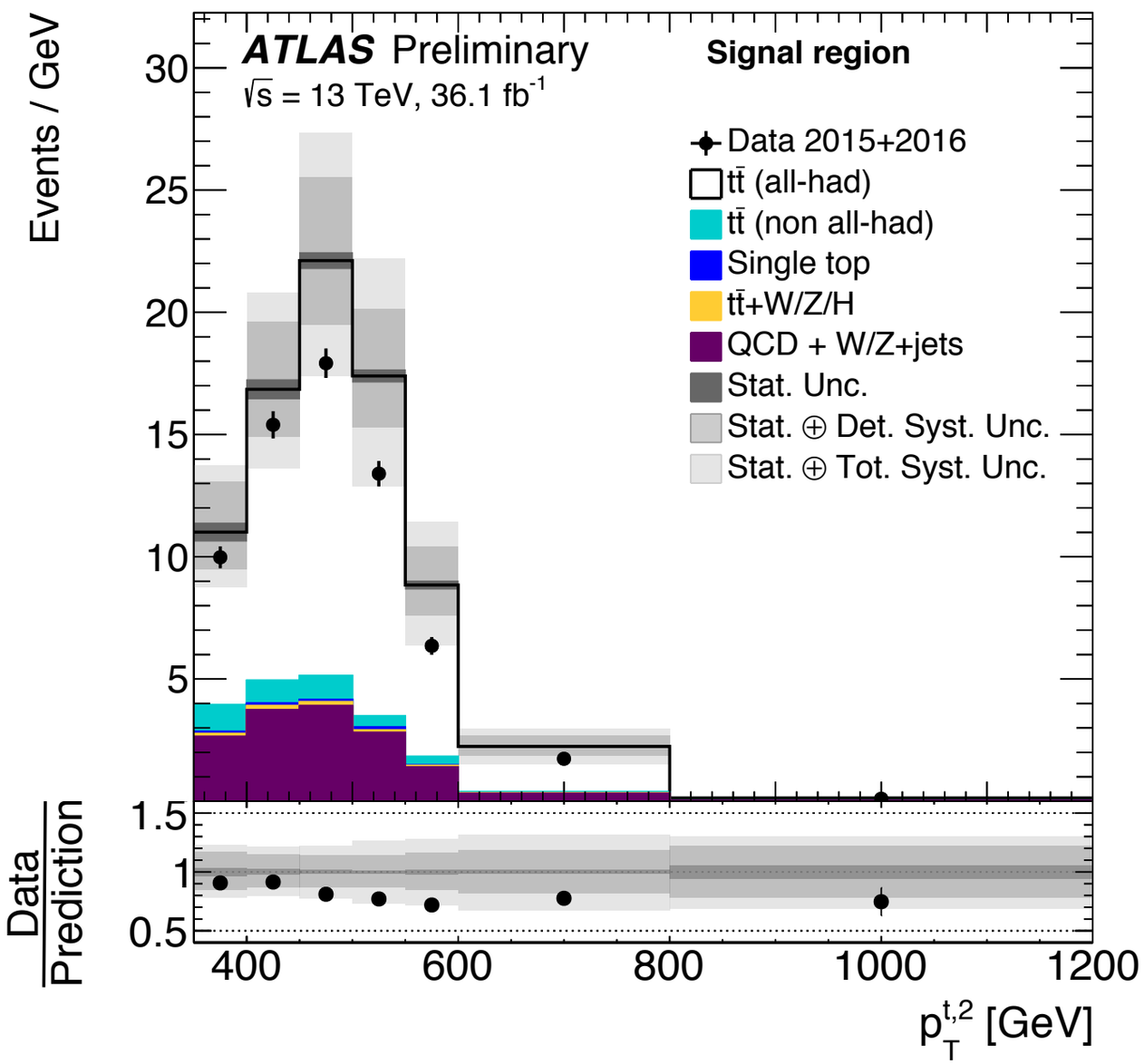


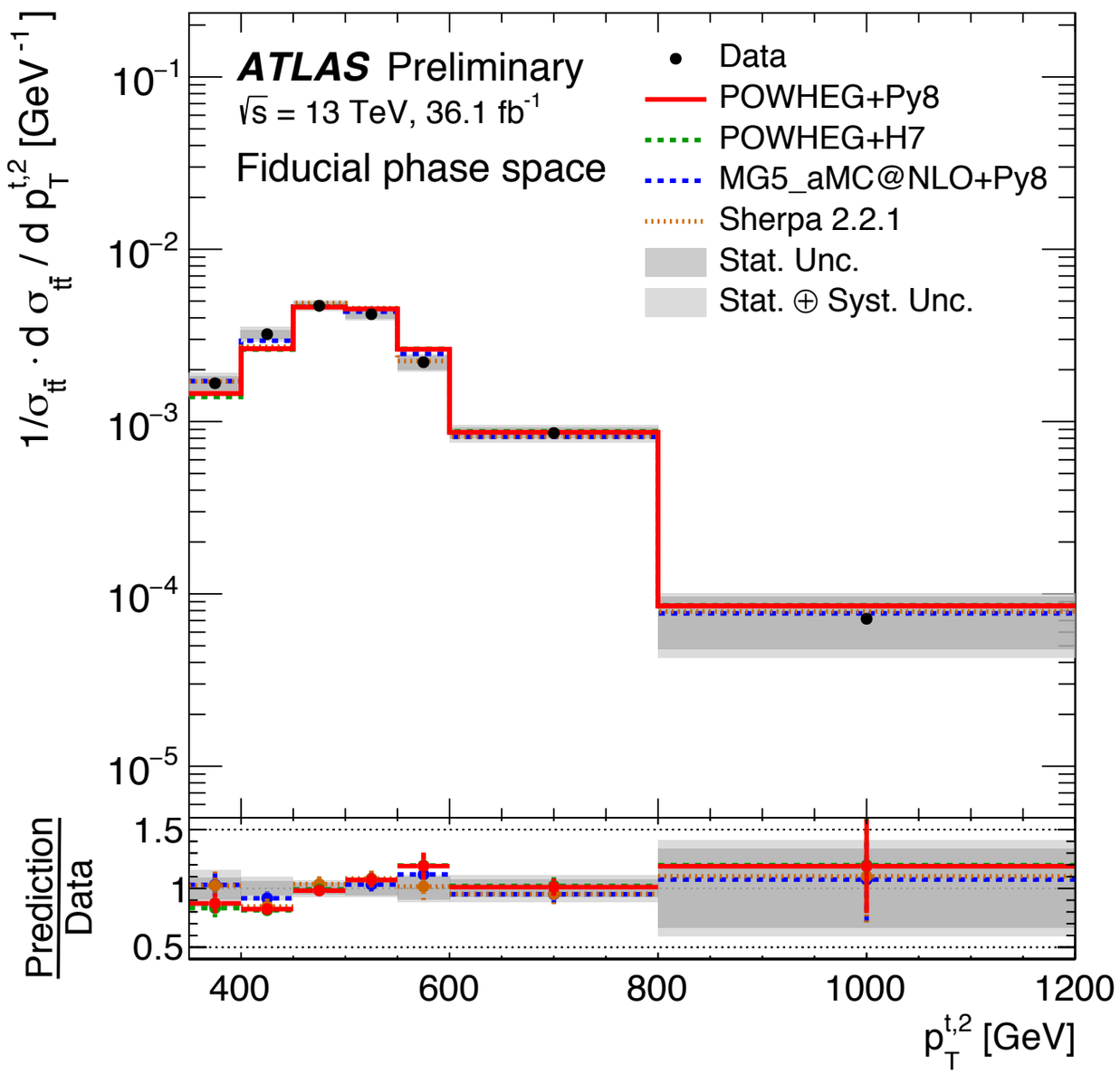
$$S = \frac{J\times O}{A}\cdot\frac{D\times A}{B\times C}\cdot\frac{G\times A}{E\times I}\cdot\frac{F\times A}{E\times C}\cdot\frac{H\times A}{B\times I} = \frac{J\times O\times H\times F\times D\times G\times A^3}{(B\times E\times C\times I)^2},$$

Sub-leading Large-R Jet \ Leading Large-R Jet	0t0b	1t0b	0t1b	1t1b
1t1b	J	K	L	S
0t1b	B	D	H	N
1t0b	E	F	G	M
0t0b	A	C	I	O



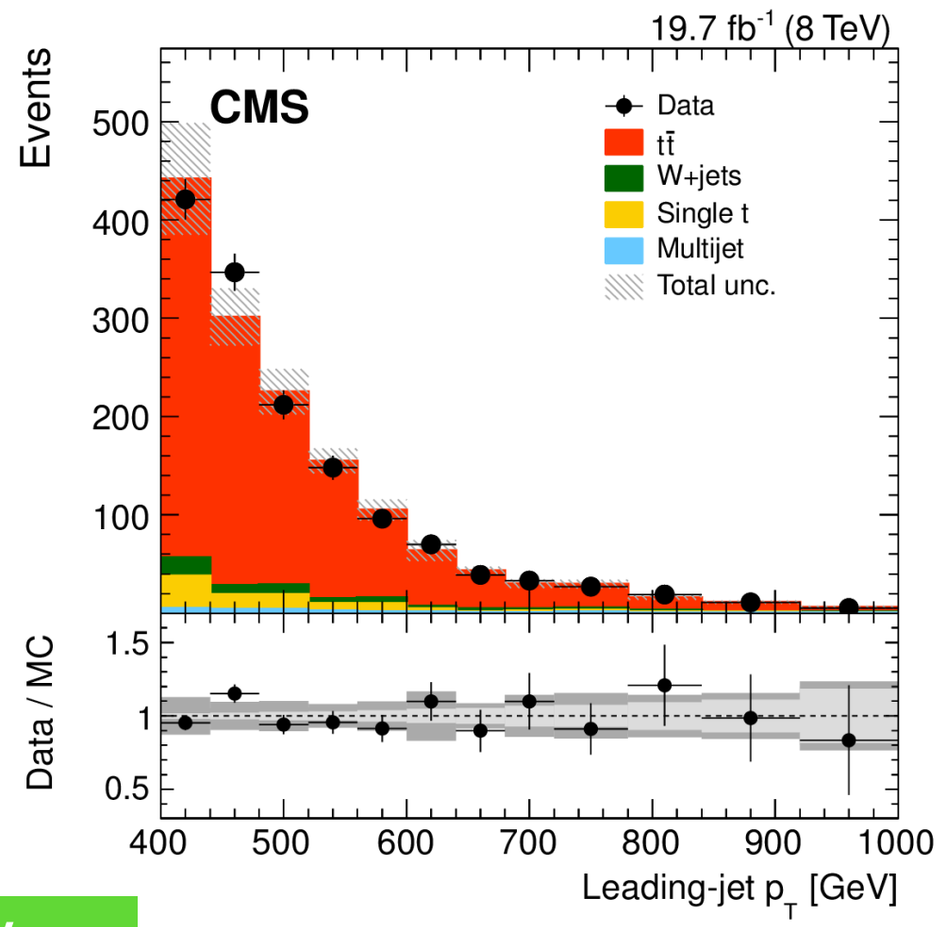
All-hadronic (13 TeV)

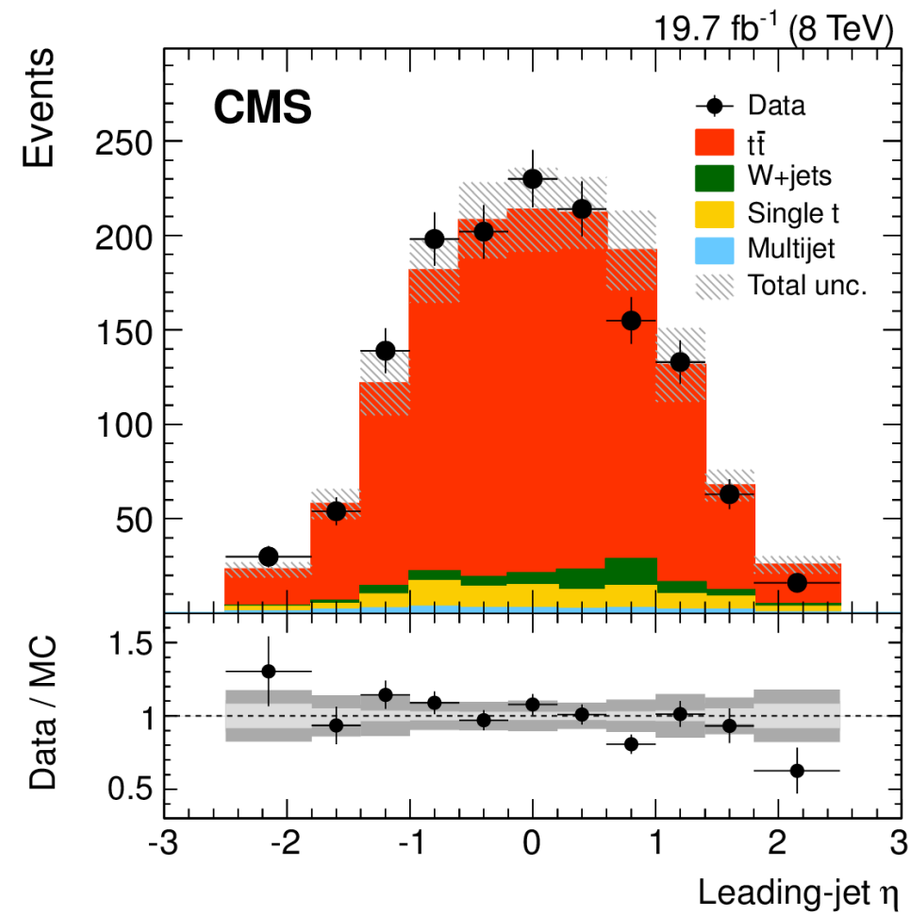




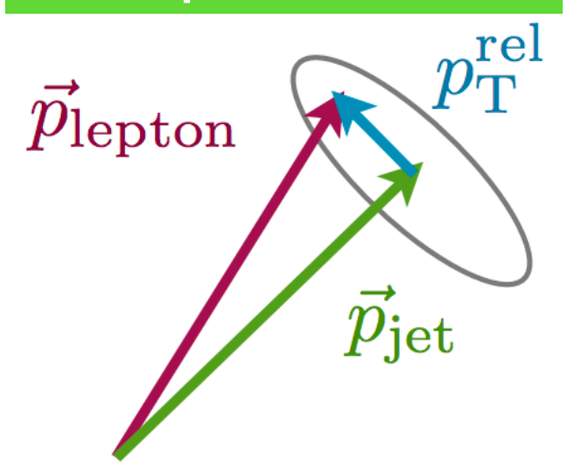
Eur. Phys. J. C 77 (2017) 467

ℓ+jets (8 TeV)





p_T^{rel}>25 GeV
(only using AK5 jets
w/ pT>25 GeV

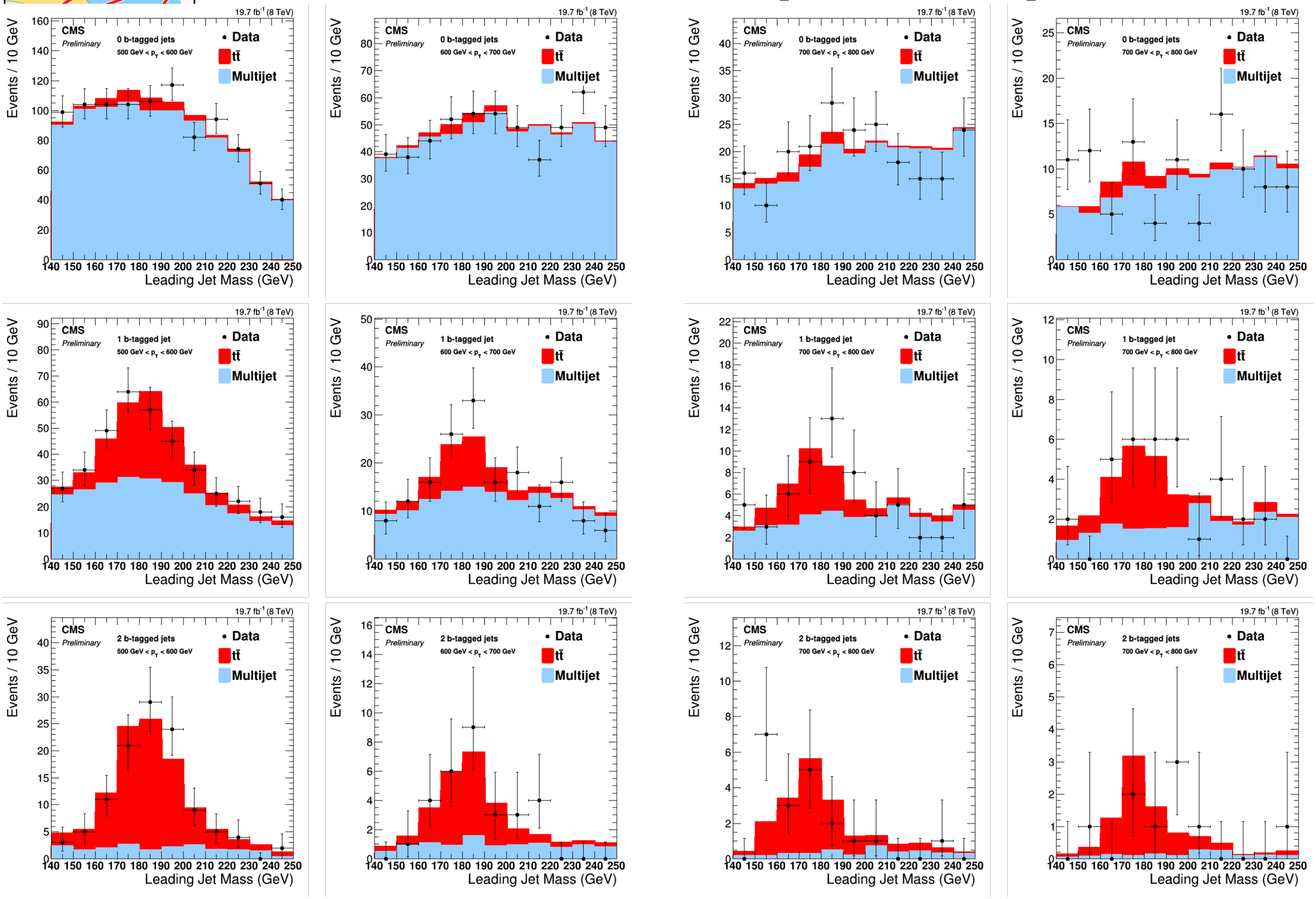


Leptons	$p_T^{\ell} > 45\,\text{GeV}$	$\lvert\eta^{\ell}\rvert < 2.1$
Jets	$p_{T,1} > 400\,\text{GeV}$ $p_{T,2} > 150\,\text{GeV}$ $p_{T,veto} > 150\,\text{GeV}$	$\lvert\eta\rvert < 2.5$
Event	$\Delta R(\ell, j_2) < 1.2$ $m_{jet,1} > m_{jet,2+\ell}$	

Range in m_{jet} [GeV]	140–170	170–200	200–240	240–290	290–350
Integrated cross section [fb]	12	42	27	18	1.7
Statistical uncertainty [%]	54	13	21	34	300
Systematic uncertainty [%]	40	9	16	20	25
Model uncertainty [%]	52	10	11	35	36
Total uncertainty [%]	85	19	28	53	300

All-hadronic (8 TeV)

CMS-PAS-TOP-16-018



CMS-PAS-TOP-16-018

All-hadronic (8 TeV)

$$L = \prod_{i}^{btags} \prod_{m}^{mass} \frac{N[i,m]^{n_i} e^{-N[i,m]}}{n_i!} \prod_{j}^{syst} e^{P_j(i,m)} \prod_{k}^{MCstat} e^{P_k(i,m)}$$

$$N[i,m] = N_{t\bar{t}}^{tag} \times P_{t\bar{t}}^{tag}(i,m) + N_{NTMJ}^{tag} \times P_{NTMJ}^{tag}(i,m) + N_{t\bar{t}}^{at} \times P_{t\bar{t}}^{at}(i,m) + N_{NTMJ}^{at} \times P_{NTMJ}^{at}(i,m)$$

p_T [GeV]	Data	Stat.	Stat. [%]	Exp.	Exp. [%]	Theory	Theory [%]	Total	Total [%]
500-600	2.51	0.251	10.0	0.510	20.3	0.487	19.4	0.75	29.8
600-700	0.789	0.108	13.7	0.335	42.5	0.055	7.0	0.357	45.2
700-800	0.266	0.051	19.0	0.062	23.2	0.025	9.3	0.084	31.4
800-1200	0.042	0.001	23.5	0.013	31.3	0.008	19.7	0.018	43.8

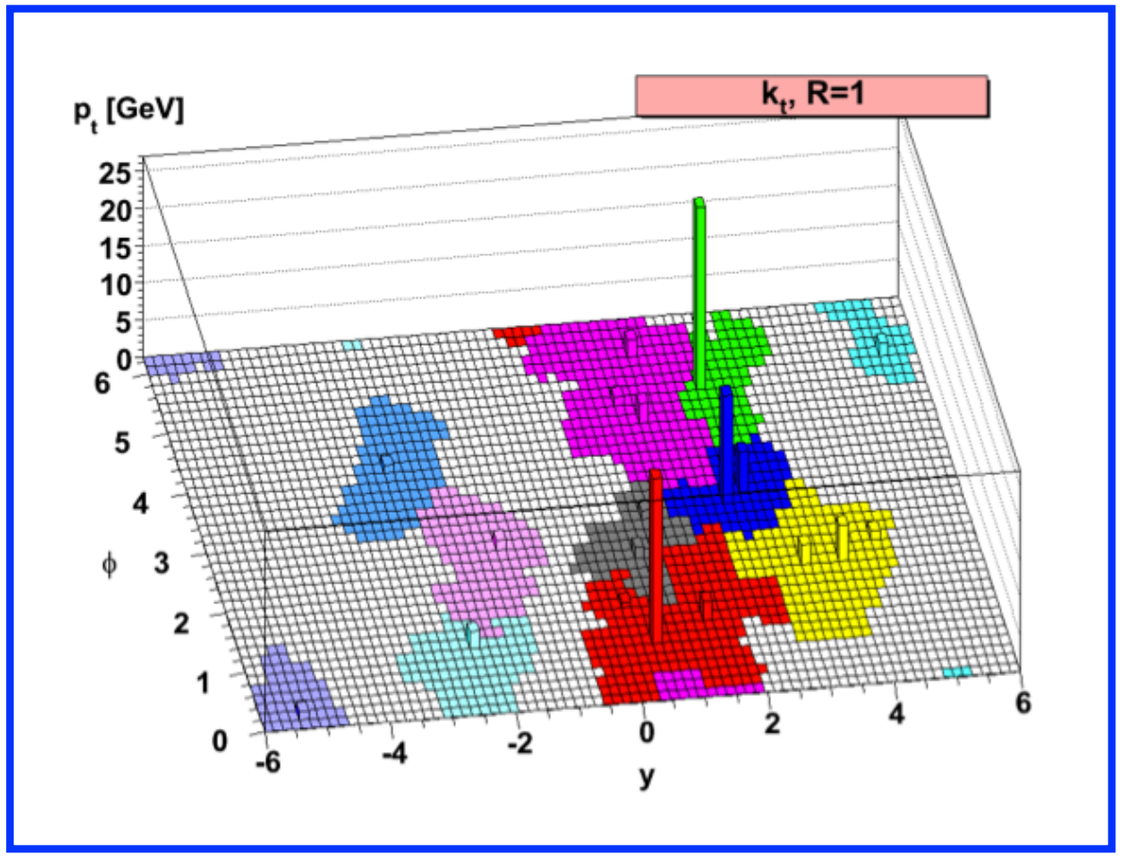
p_T [GeV]	Data $\pm$ Tot. Unc.	POWHEG	MC@NLO	MADGRAPH	SemiLep	CT14
500-600	2.51 ± 0.75	3.20	2.63	3.64	2.74	2.52
600-700	0.789 ± 0.357	0.972	0.754	1.11	0.786	0.733
700-800	0.266 ± 0.083	0.322	0.238	0.363	0.254	0.233
800-1200	0.042 ± 0.018	0.049	0.030	0.050	0.036	0.031

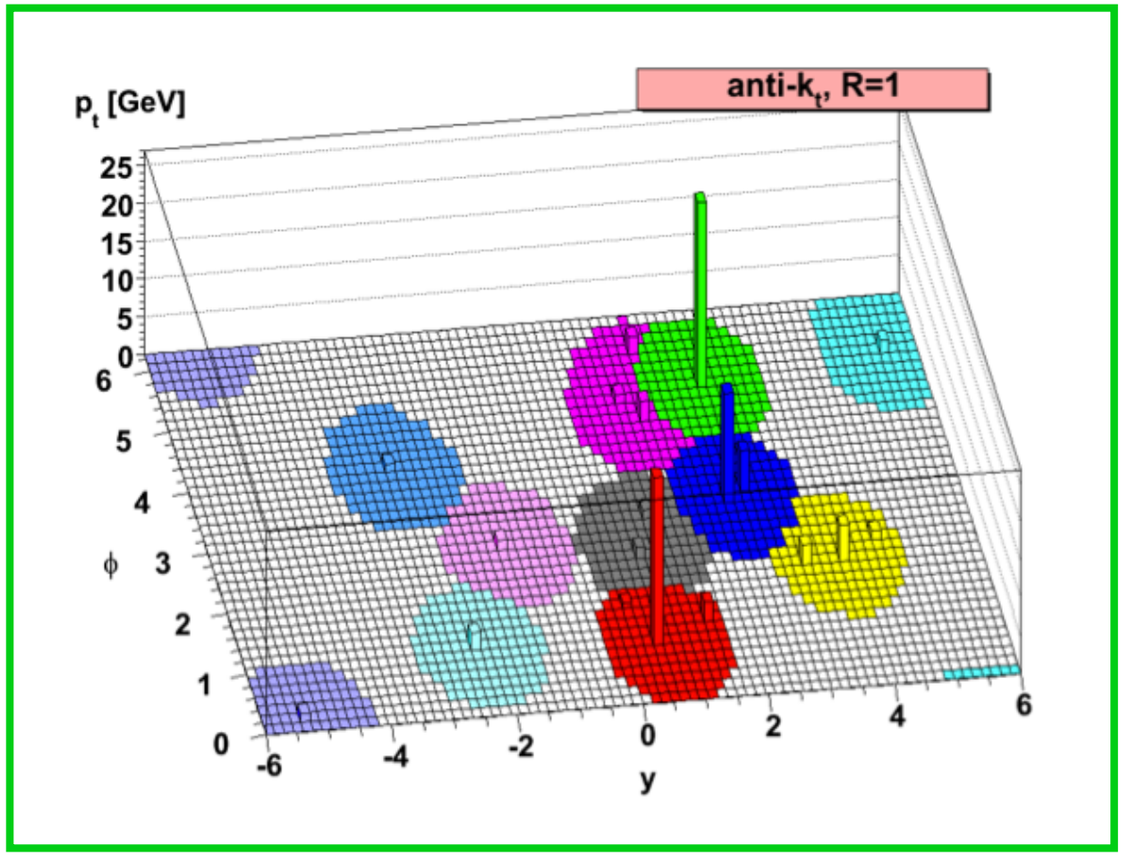
Jet Clustering

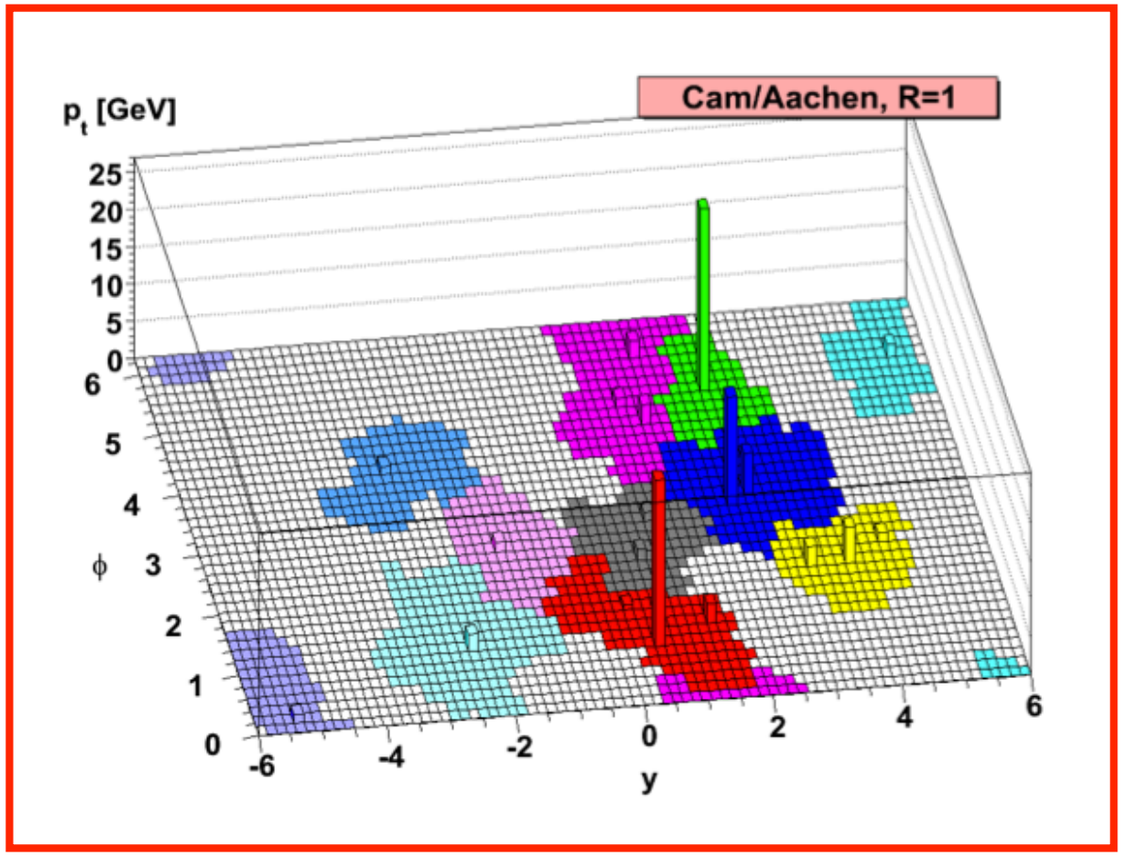
$$d_i = p_{T_i}^{2a}$$

$$d_{ij} = \frac{\Delta R_{ij}^2}{R^2} \min(p_{T_i}^{2a}, p_{T_j}^{2a})$$

a = 1 : k_T
a = 0 : C/A
a = -1 : anti-k_T







Jet Clustering

1. Define the splitting scales for each input and pairs of inputs (Topoclusters or PF objects).

2. Find min{d_q}, where d_q includes all of the distance scales d_i and d_{ij}

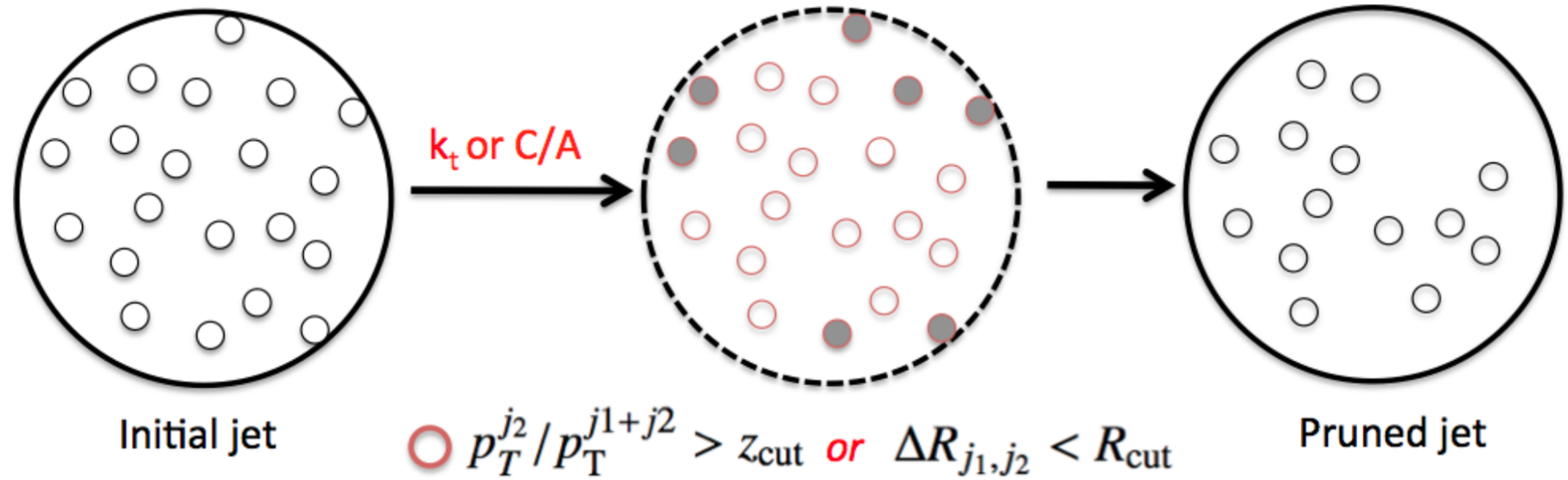
3.(a) If min is a d_{ij}: redefine as d_k.

$$
\begin{aligned}
p_{T_k} &= p_{T_i} + p_{T_j} \\
\eta_k &= (p_{T_i}\eta_i + p_{T_j}\eta_j)/p_{T_k} \\
\phi_k &= (p_{T_i}\phi_i + p_{T_j}\phi_j)/p_{T_k}
\end{aligned}
$$

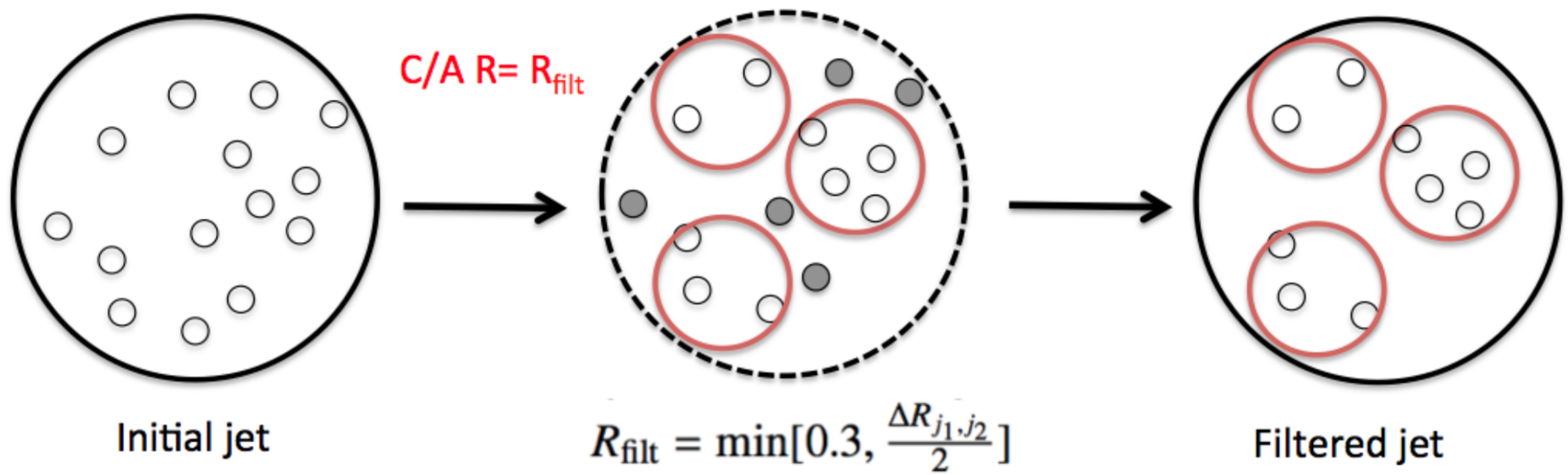
3.(b) If min is a d_i: Remove from list and move to JET list

4. Repeat until all JETs formed

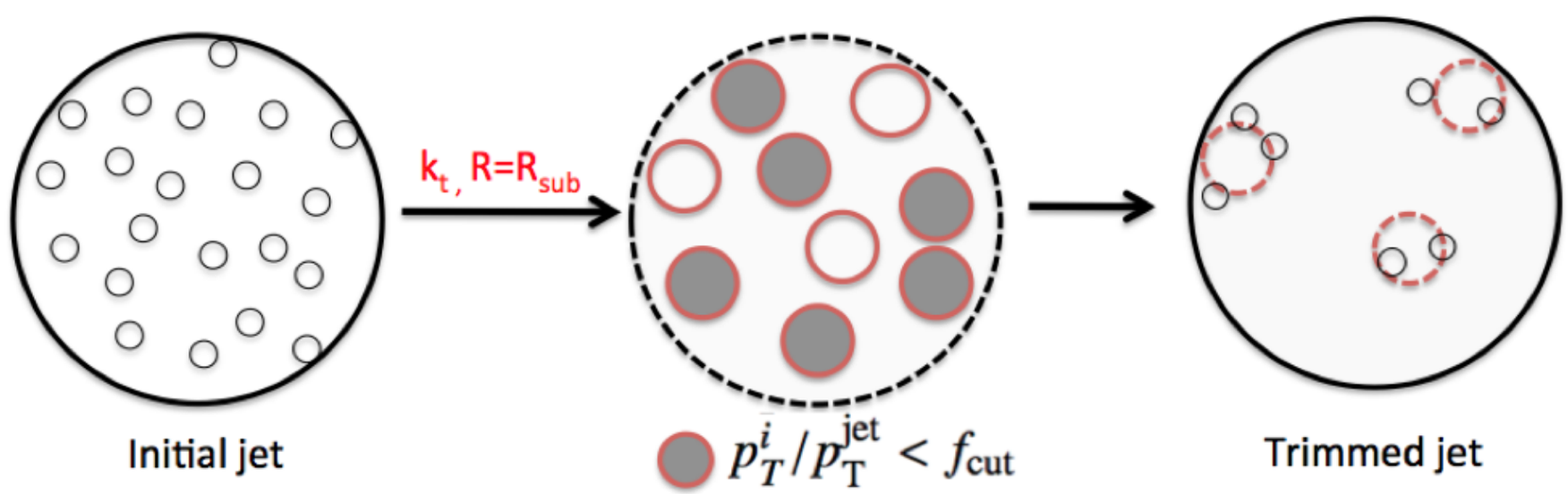
Jet Grooming



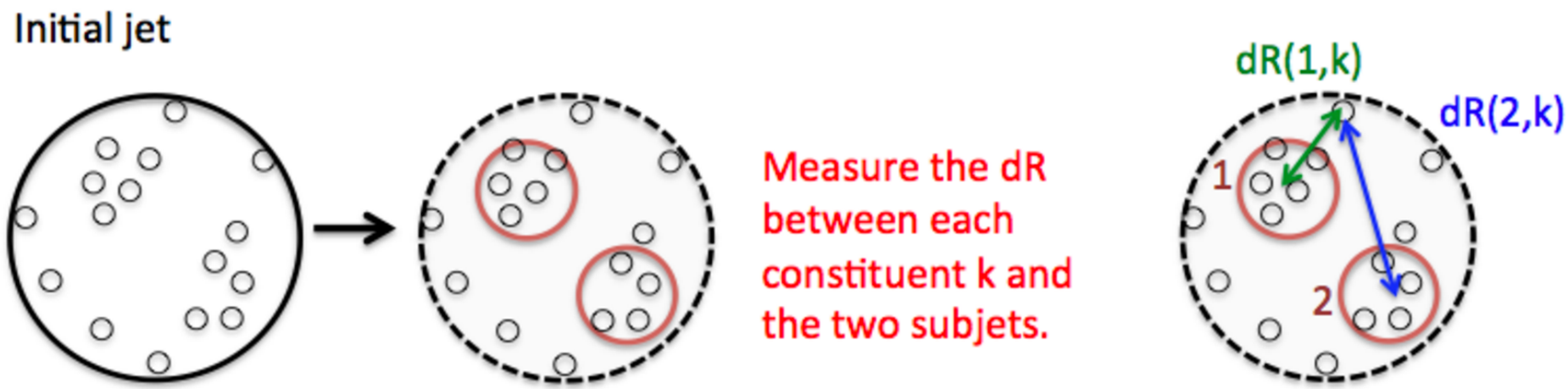
Pruning



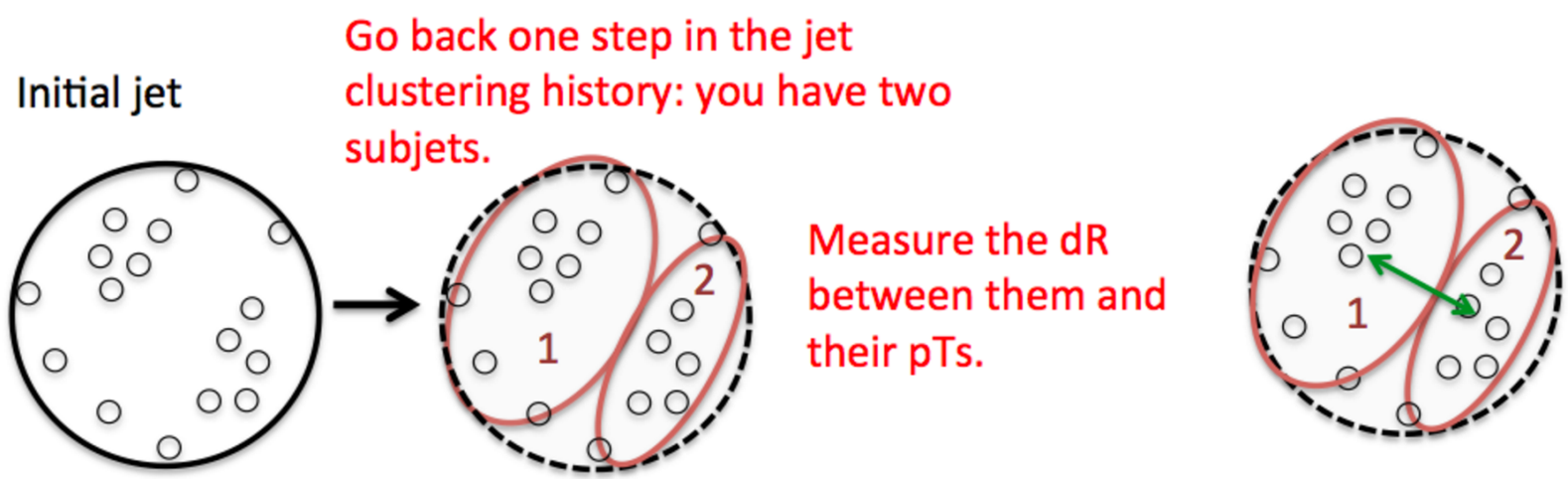
Filtering



Trimming



$d_k = pT(k) \times \min(dR(1,k), dR(2,k))$



$\sqrt{d_{12}} = \min(pT(1), pT(2)) \times dR(1,2)$

$$\tau_N = \frac{1}{d_0} \sum_i p_{T_i} \times min(\Delta R_{1,i}, \Delta R_{2,i}, ..., \Delta R_{N,i}). \quad (1)$$

Here $\Delta R_{j,i}$ is the angular distance between the subjet axis j and the candidate jet i. The normalization constant d_0 takes p_T into account, with $d_0 = \sum p_{T_i} R_0$, and R_0 is the distance parameter used in the jet clustering algorithm (here $R_0 = 0.8$). If a jet is N-pronged, then τ_N is small, whereas τ_{N-1} and lower moments are all larger. This analysis primarily makes use of τ_{32} defined as τ_3/τ_2.