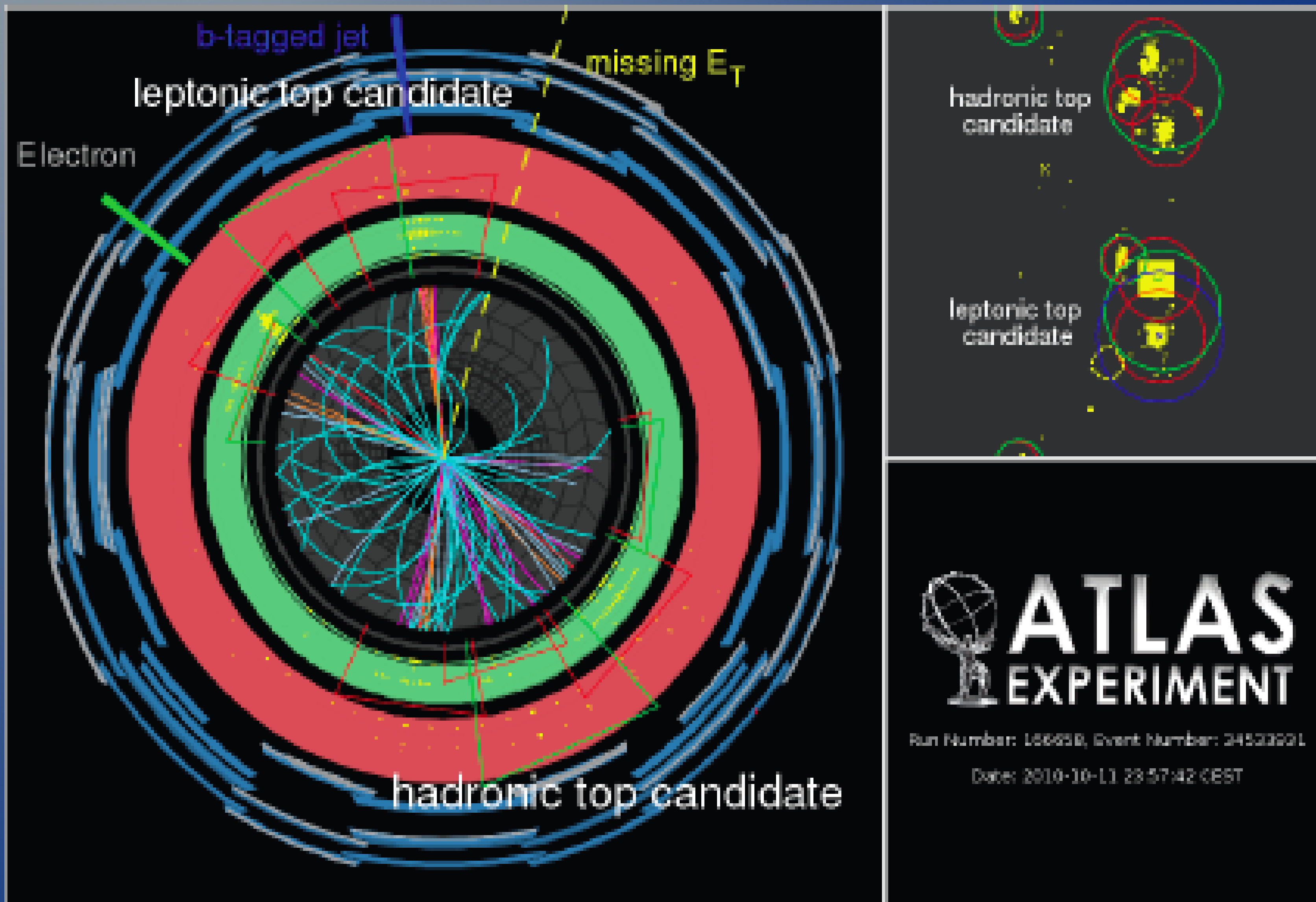
Jets and boosted objects at the LHC

Mario Campanelli/ UCL



Why?

At the end of the first data-taking period of the LHC, after > 20/fb of data, and the discovery of a Higgs-like particle, why do I come here to talk about jets?
Isn't it old, boring physics?

QCD may be old (well, not more than the Higgs mechanism!), but its experimental study can be extremely creative, and has been able to reinvent itself over the years (yes it would have been boring if we were using the same techniques as Tevatron Run I)

In the following I will highlight some aspects of jet physics through the lens of new ideas (some of which I had, most of which I would have like to have)

More specifically, I will talk about

- Algorithms, triggers, calibration
- Inclusive and dijet cross section
- Hadronic decays of boosted objects as a window to new physics
- Jet grooming techniques
- Boosted top
- Quark/gluon jet separation

Jet algorithms

Most of jet measurements at the Tevatron have been made with not fully adequate jet clustering.

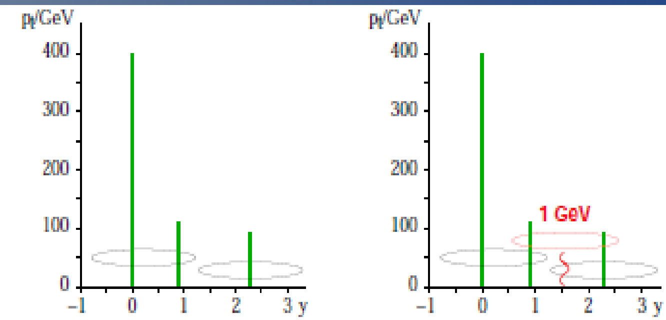
Trigger was including in a jet all calorimeter clusters within a given radius of a major energy deposition

Offline clustering followed the same technique, with some refinement in the case of overlapping jets.

Recombination algorithms too slow for either trigger or offline.

But cone algorithms were shown to be all infrared unsafe: the result critically depends on the threshold at which clusters can initiate jets

Issues with cones

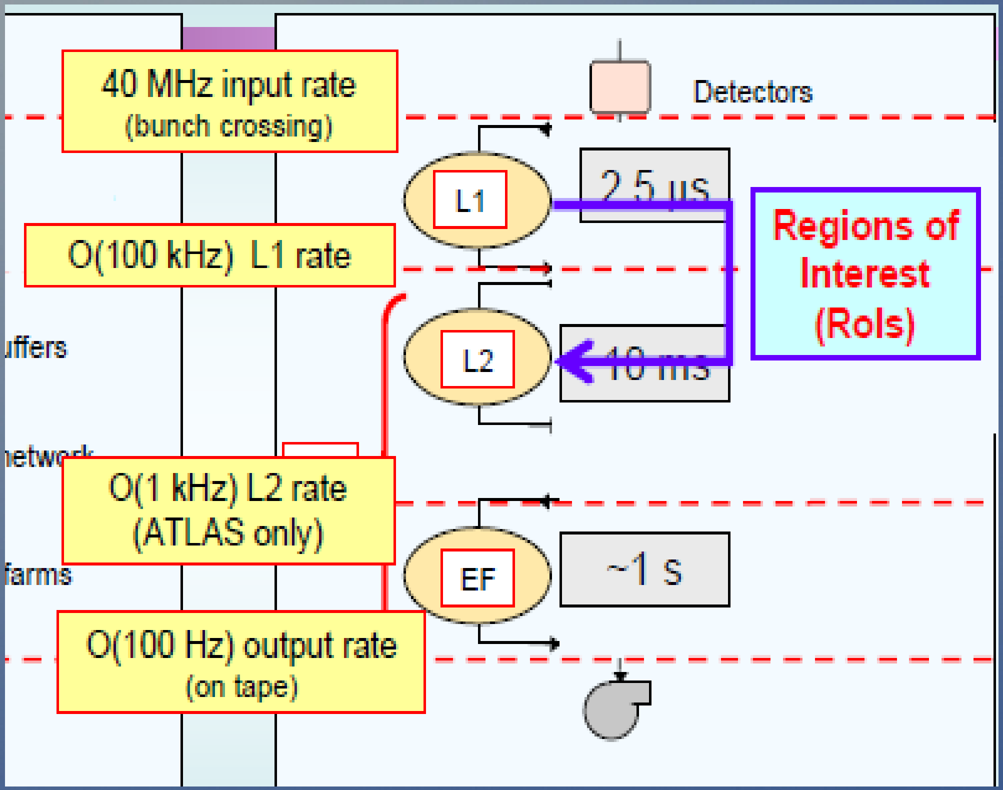


Stable cones with midpoint:	{1,2} & {3}	{1,2} & {2,3} & {3}
Jets with midpoint ($f = 0.5$)	{1,2} & {3}	{1,2,3}

	Last meaningful order			
	JetClu, ATLAS cone [IC-SM]	MidPoint [ICmp-SM]	CMS it. cone [IC-PR]	Known at
Inclusive jets	LO	NLO	NLO	NLO
$W/Z + 1$ jet	LO	NLO	NLO	NLO
3 jets	none	LO	LO	NLO [nlojet++]
$W/Z + 2$ jets	none	LO	LO	NLO [MCFM]
$m_{\rm jet}$ in $2j + X$	none	none	none	LO $\rightarrow$ NLO

An infrared-unsafe algorithm does not guarantee cancellation of infinities performed by high-order theory calculations, therefore no meaningful comparison with higher orders of theory is possible.
All LHC publication now use the infrared-safe Anti-Kt algorithm, (Salam, Soyez), it looks trivial now but it was a very big step forward

Infrared-safe triggering: breaking the ATLAS trigger strategy



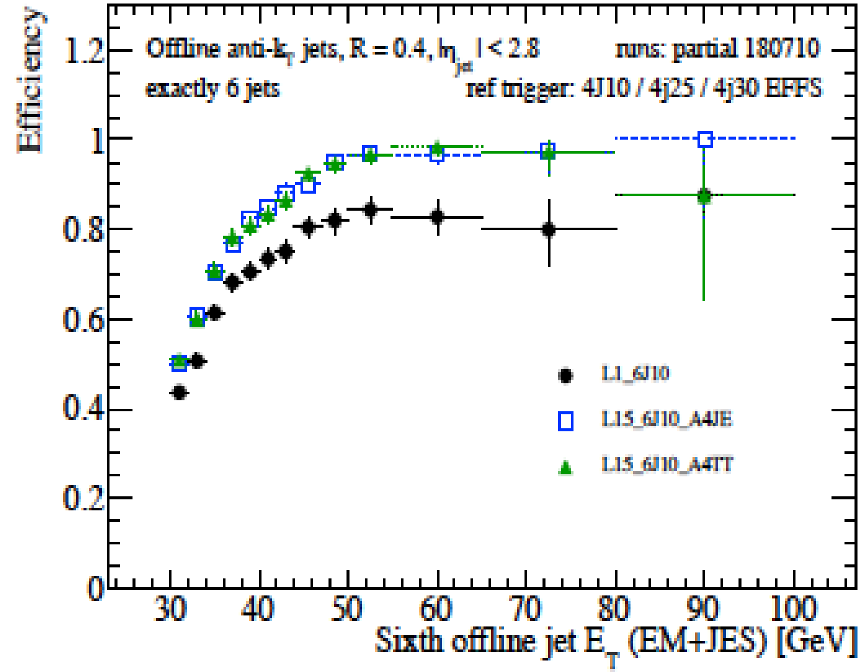
Based on the concept of region of interest (ROI) L1 (calo or muon) tells L2 which parts of the detector had interesting signals, and only those will be read. Same between L2 and EF.

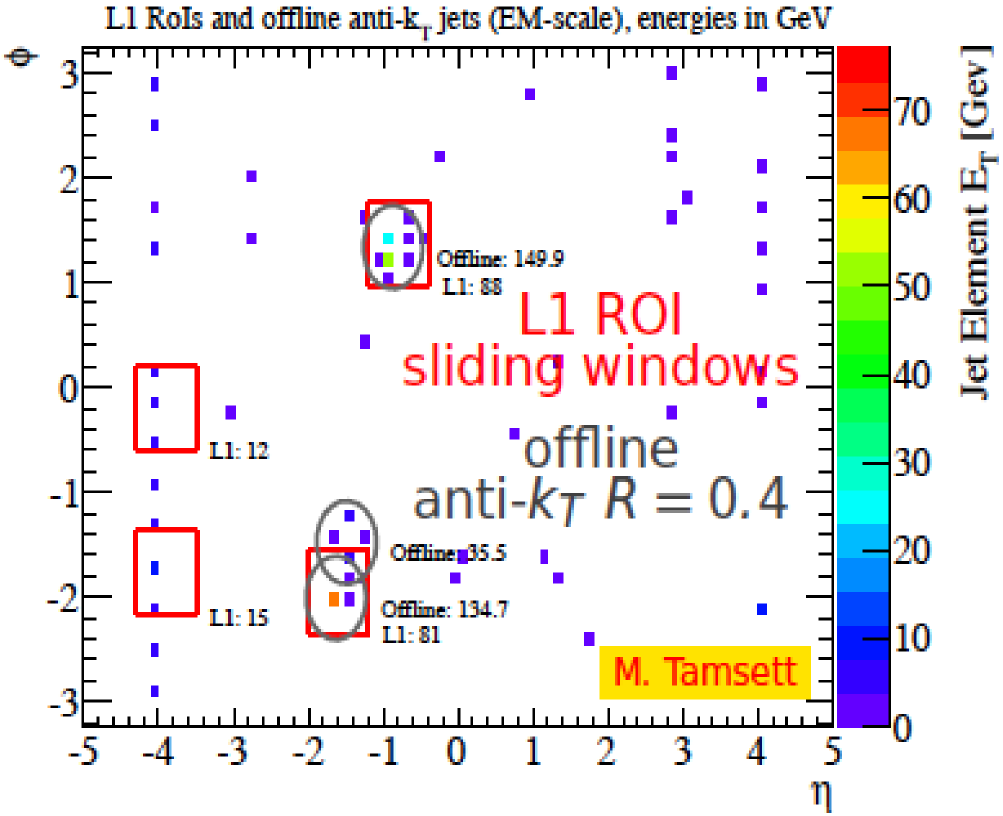
This approach is infrared-unsafe, and it was shown that a single cluster passing or not noise thresholds leads to different ROI configurations, therefore different triggered jets

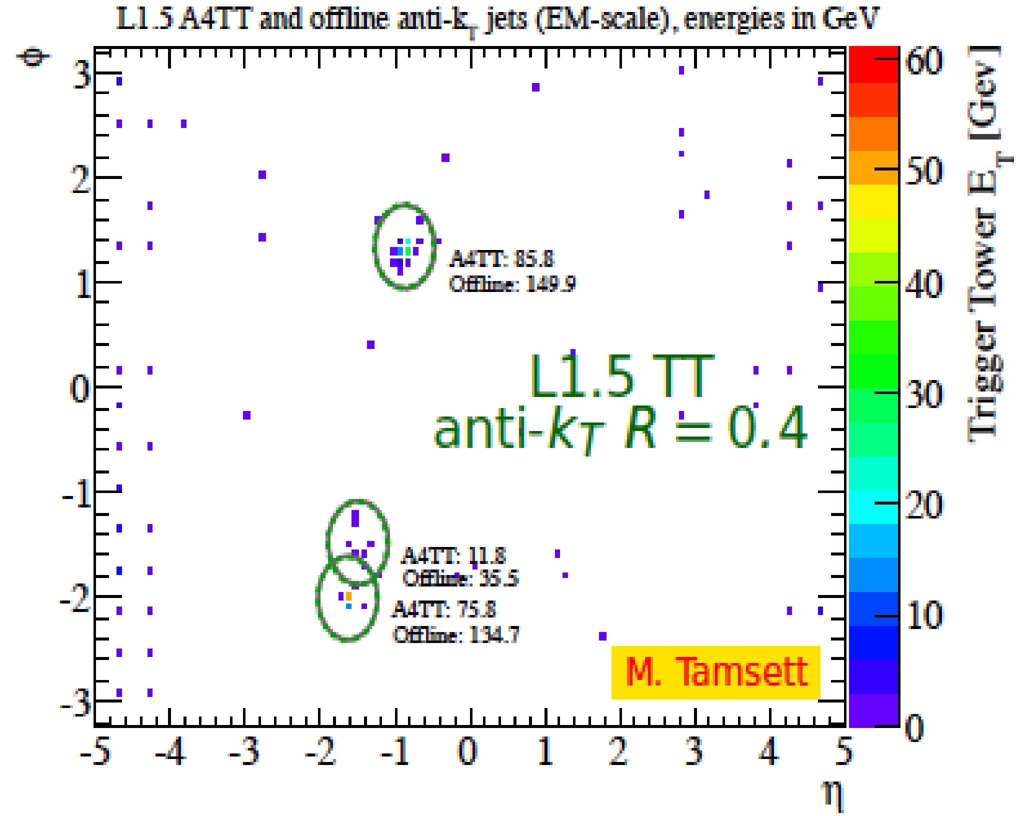
Not much could be done on the hardware-only L1 system, but we introduced full-scan clustering at EF (all events) and L2 (selected chains)

- **Full-scan anti-k_T jets from L1Calo towers**
  - read L1Calo instead of Calo ROS yielding increased L2-input rate
  - *included since run 193834*
- **Adds flexibility & acceptance (multijets, fat jets, jet+X)**
  - reduces inefficiency associated with L1 sliding windows & L1/L2 ROI

M. Tamsett



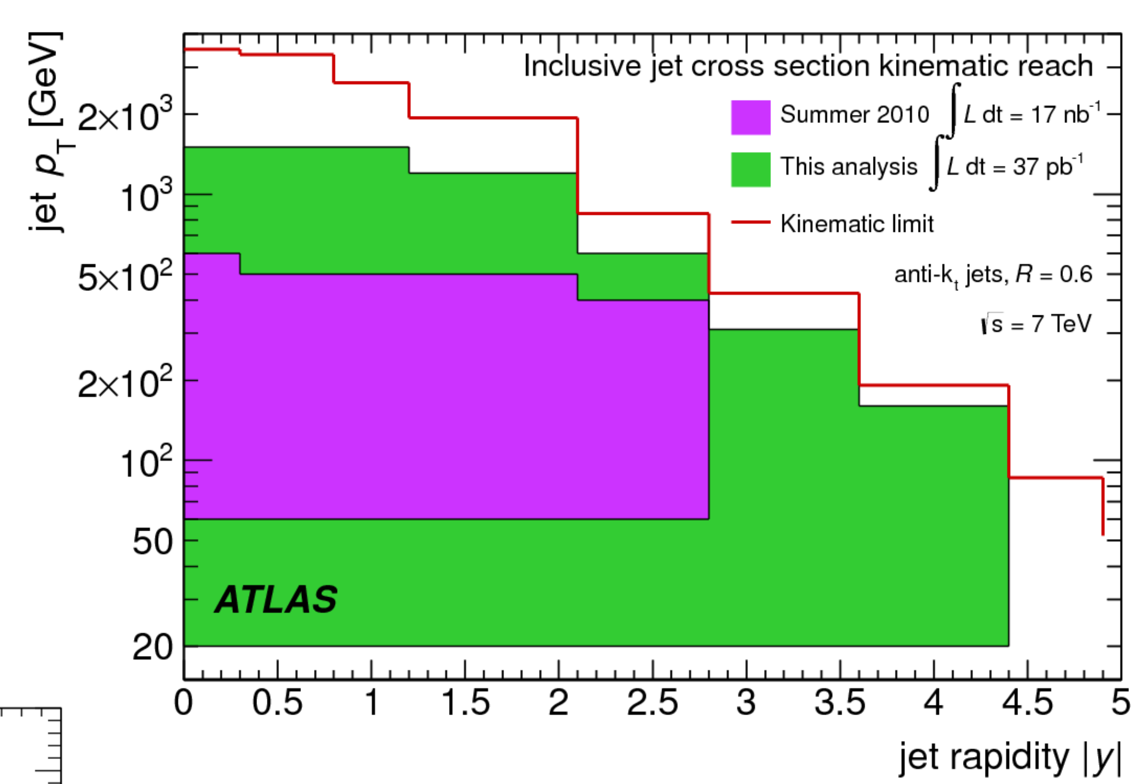


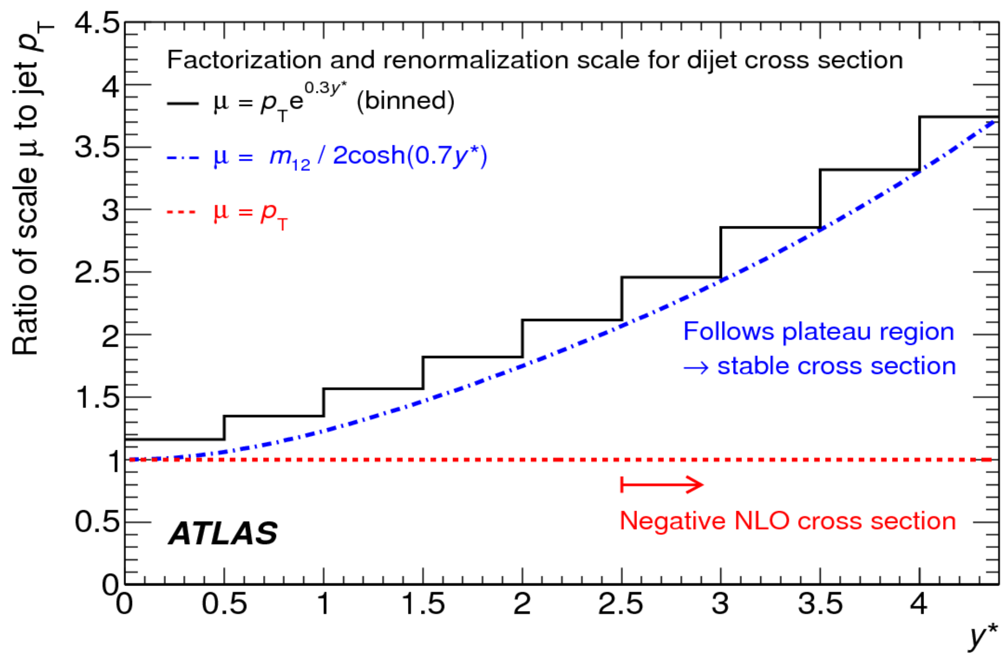


Phys.Rev. D86 (2012) 014022: the ultimate inclusive jet and dijet cross-section paper?

We did our best to produce the most luxurious bread-and-butter you can think of

Measure jets clustered with akt04 and akt06 almost to kinematic limit



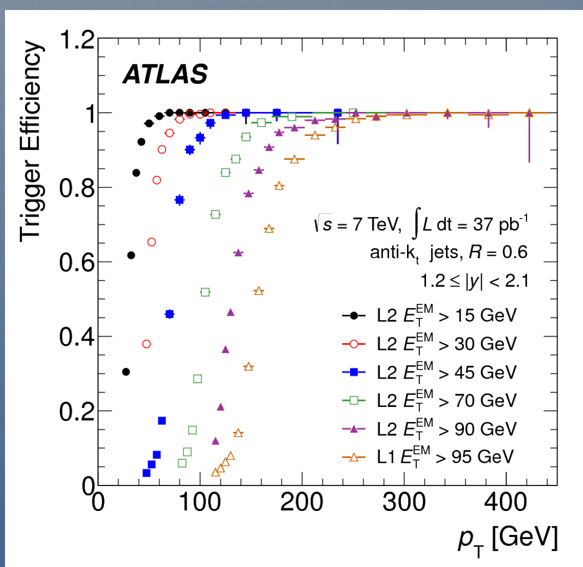


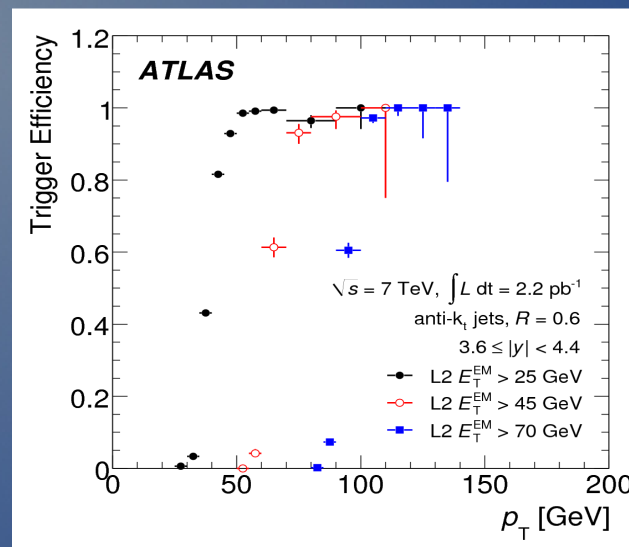
For very forward jets NLO calculation becomes negative, and we had to use a special scale definition to make sense of it

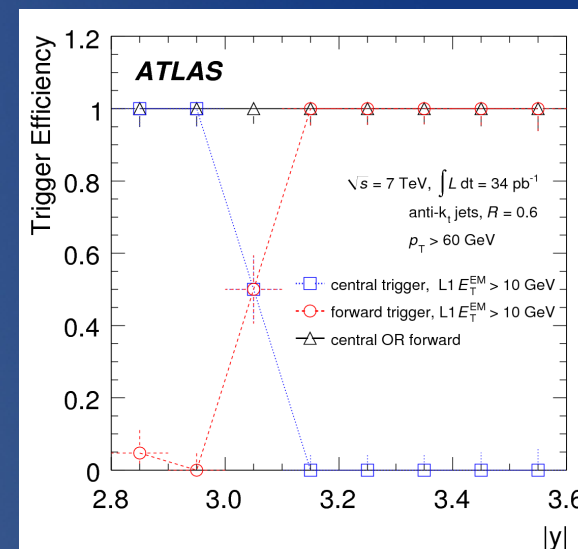
Two-jet trigger strategy

Central and forward jet triggers are independent and with different turn-on and prescales

No single-jet trigger strategy was optimal to select central-forward nor forward-forward events





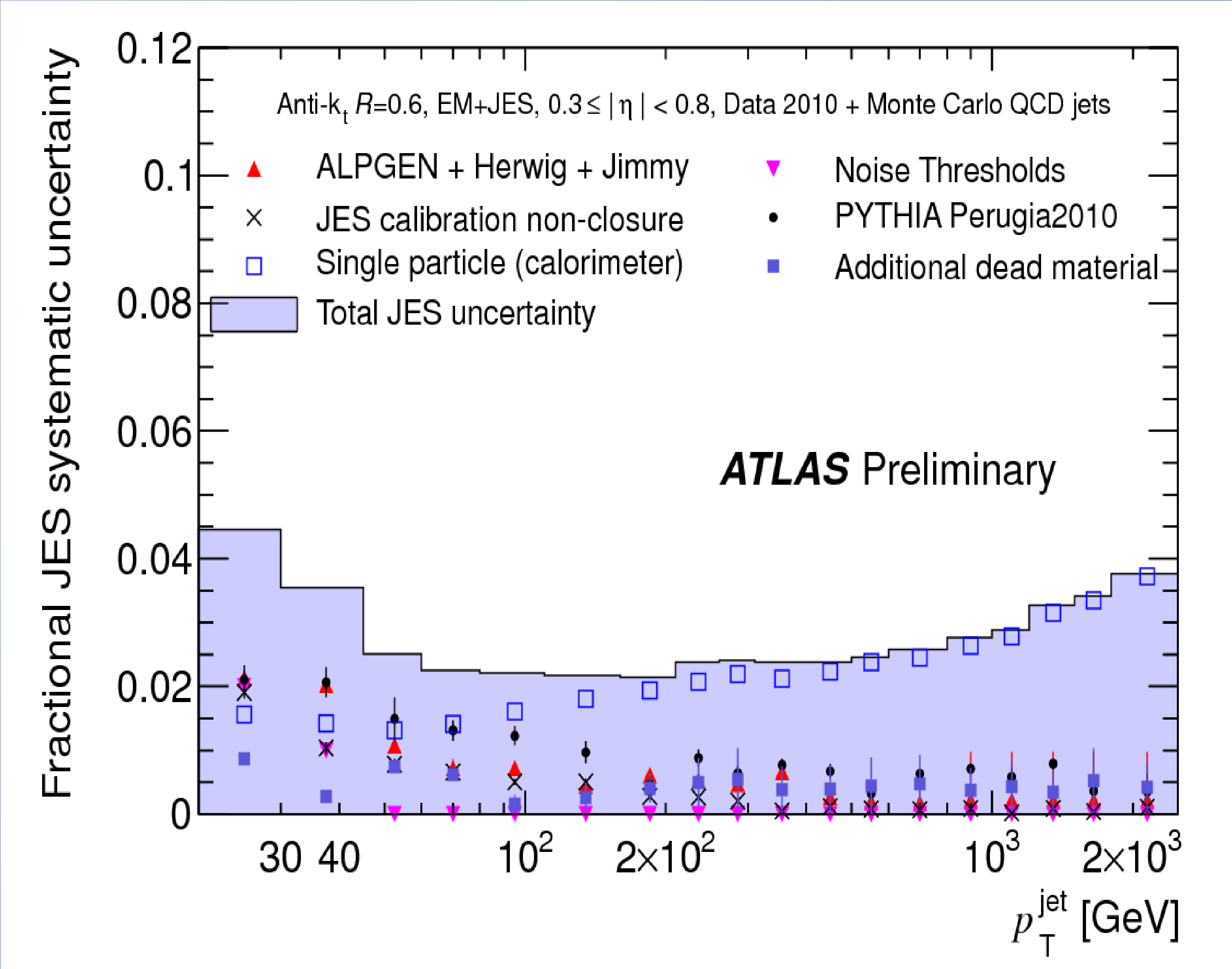


We divided dijet phase-space into hundreds of regions according to jet pt and eta. At least one jet was at trigger plateau.

Equivalent luminosities for each region were computed using trigger prescales, and everything was combined back again accounting for overlaps

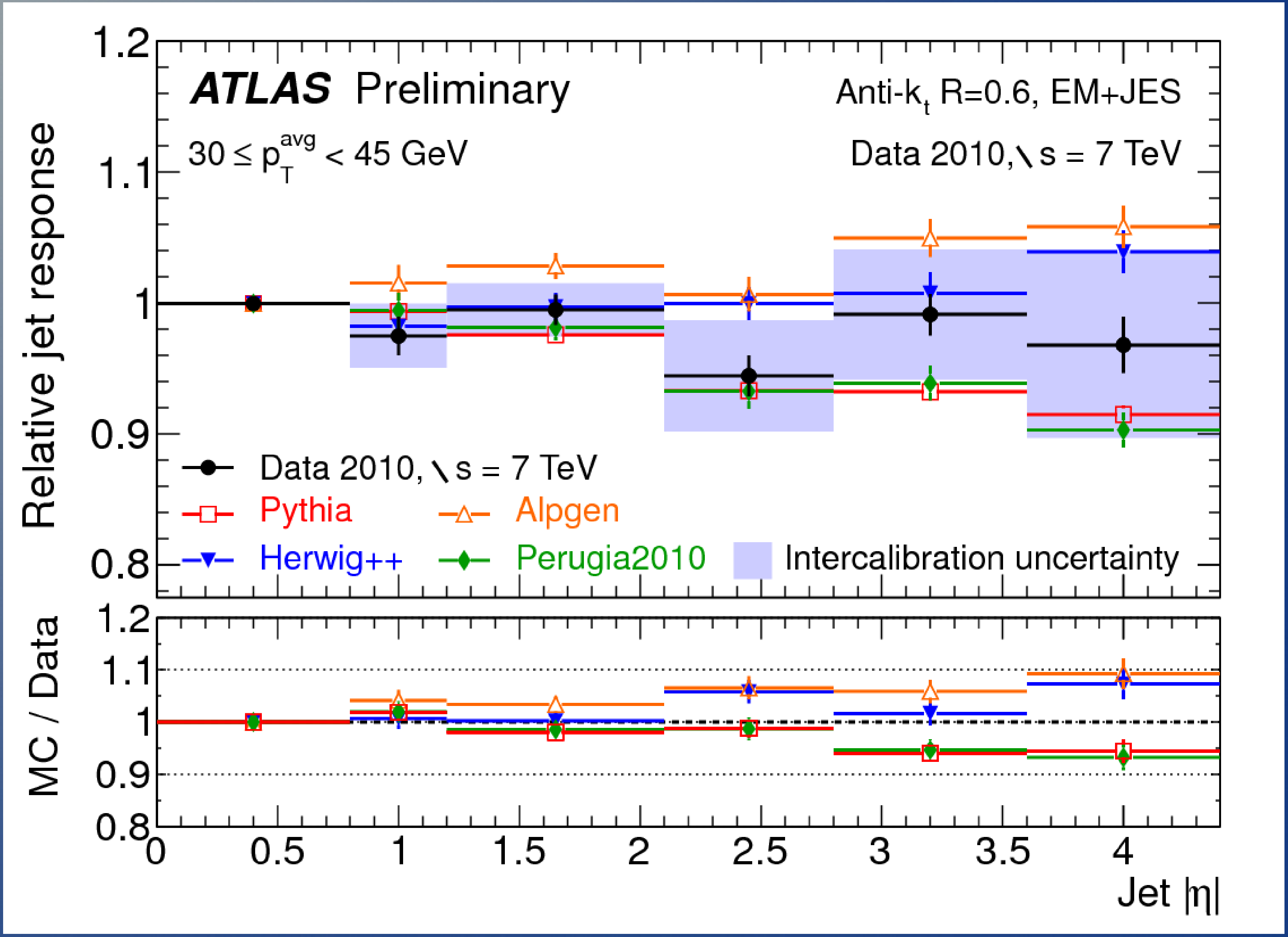
This technique maximised the dijet output for every kinematical configuration

JES uncertainty in 2010



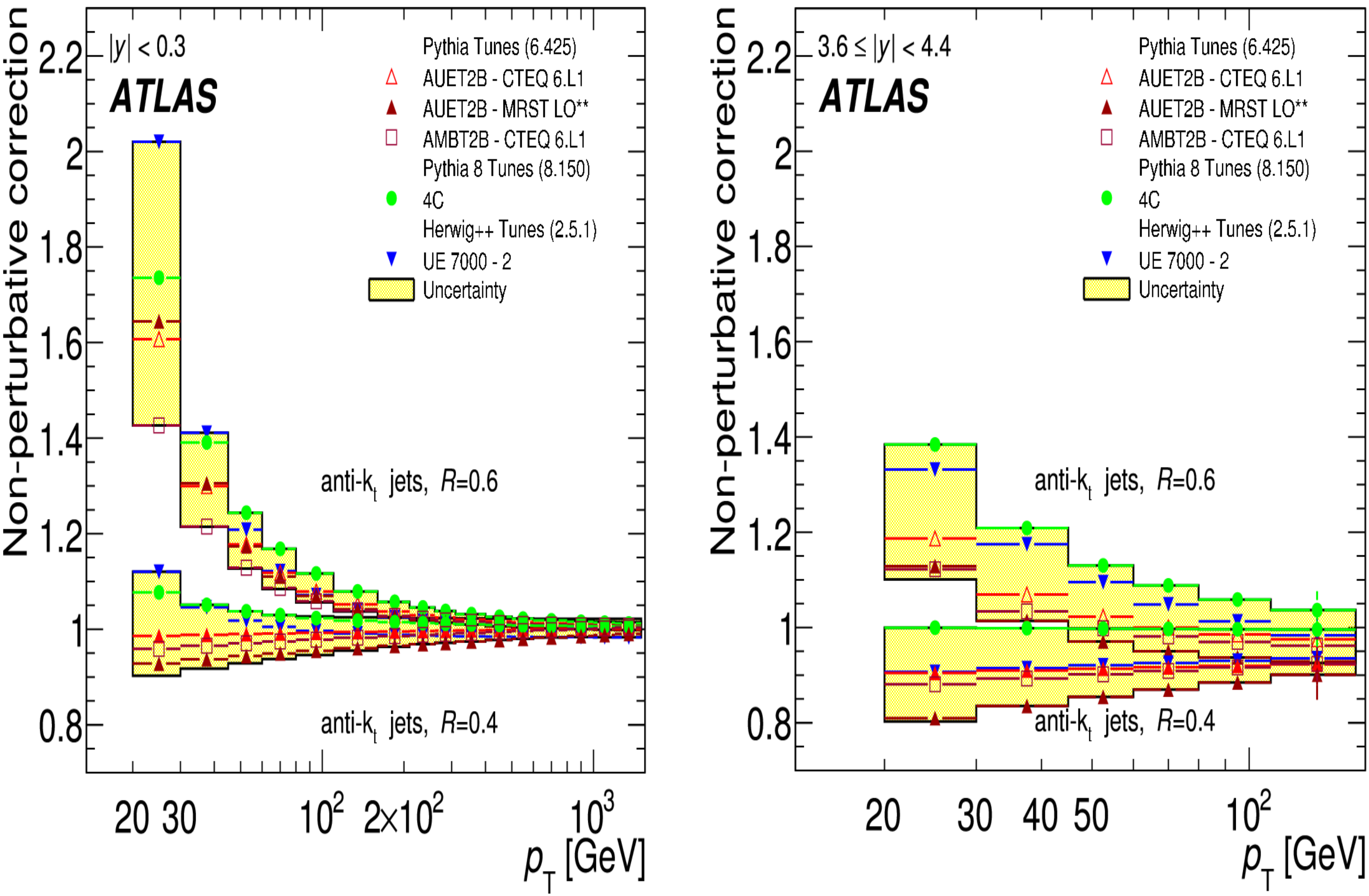
Largest source of systematics for any steeply-falling distribution, and a lot of effort to reduce it.
Dominated by statistics of single particle response
Now we use a much more refined calibration

Forward JES calibration and (more) uncertainty

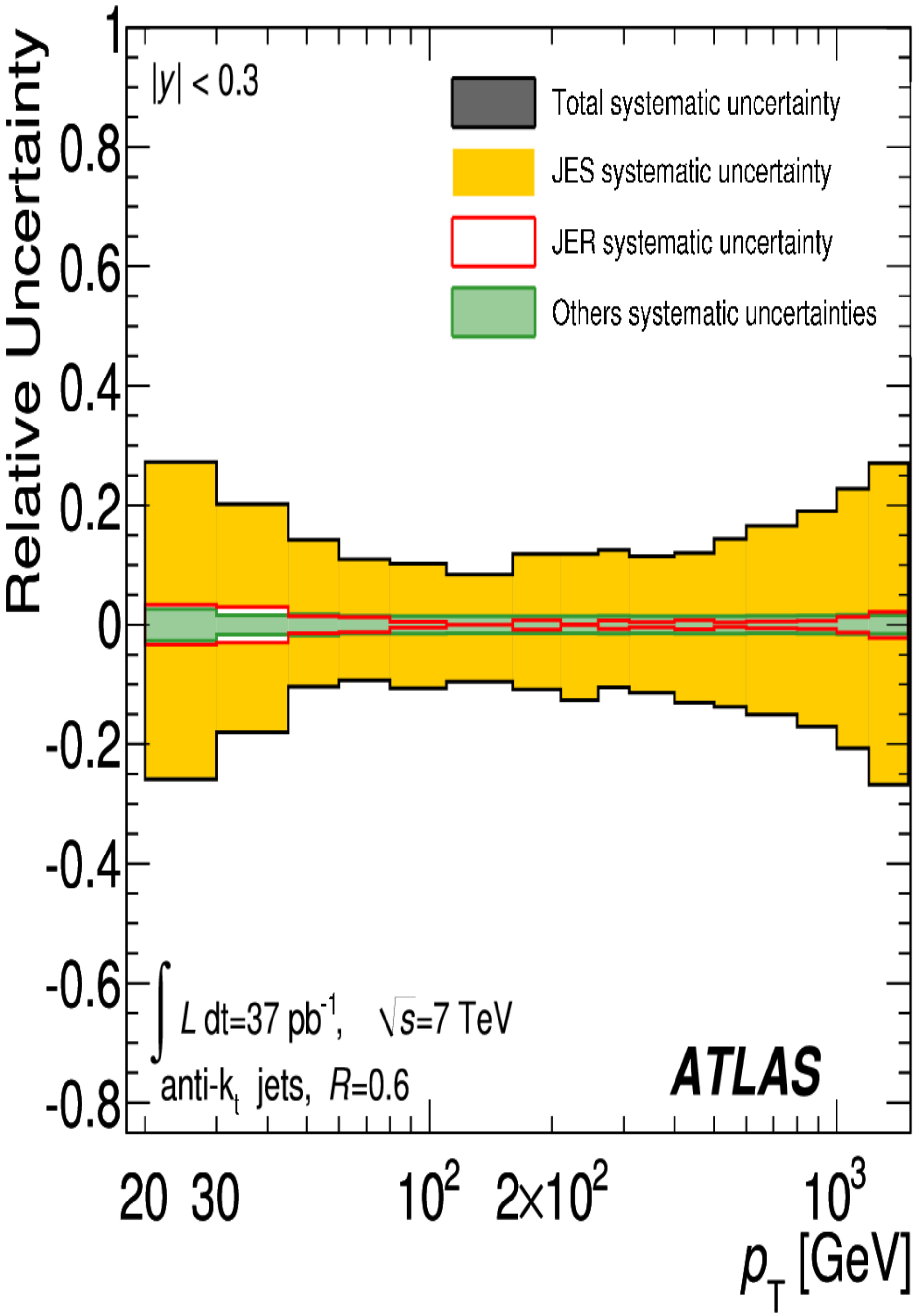


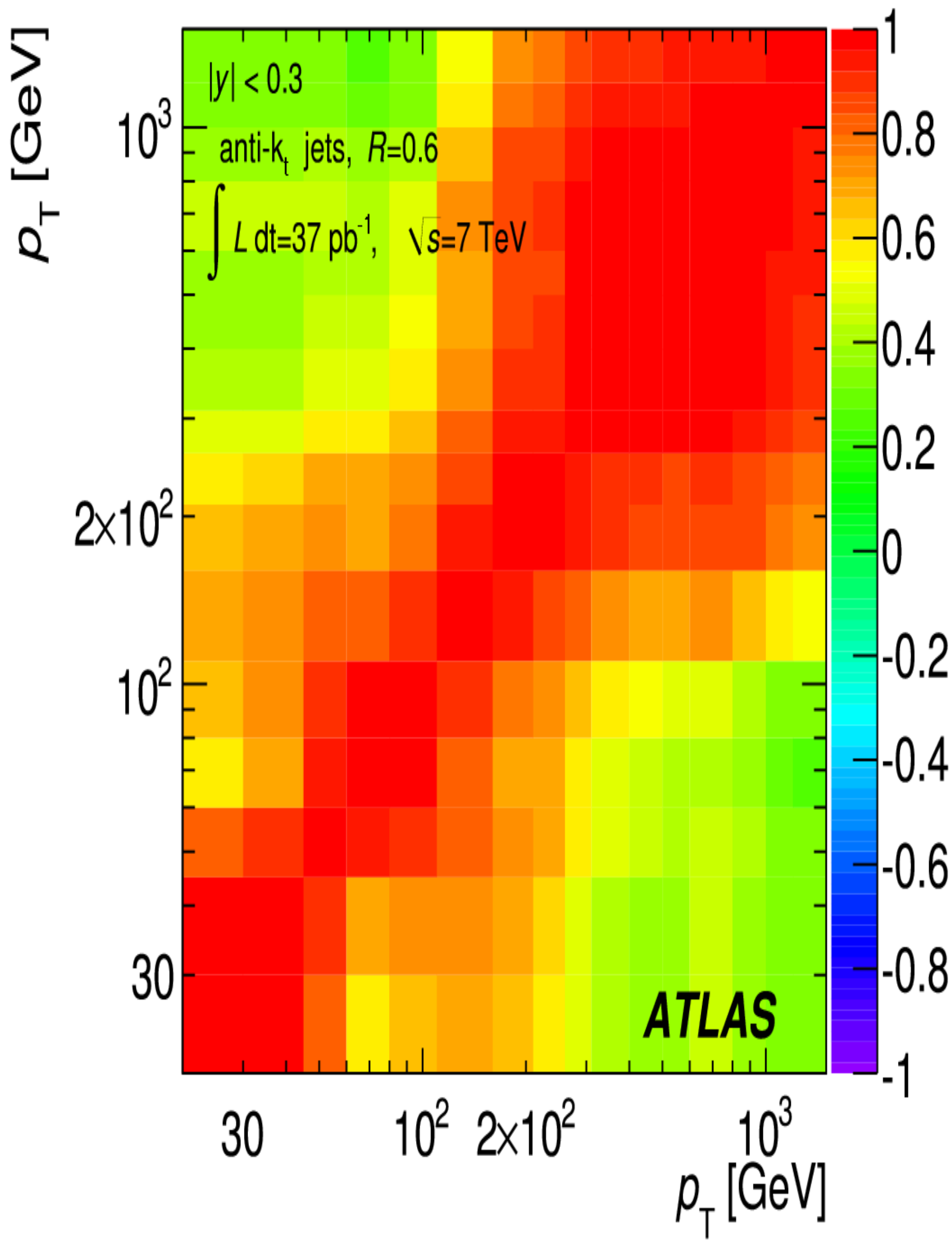
Without tracking, the only way to cross-check the JES in the forward region is central-forward jet balance, in the limit of vanishing third jet. Dicrepancies O(10%) have been found, going in opposite directions between Pythia and Herwig (Alpgen + Herwig) showering. Additional systemetic uncertainties applied.

Non-perturbative corrections



Systematic uncertainties: magnitude and correlations



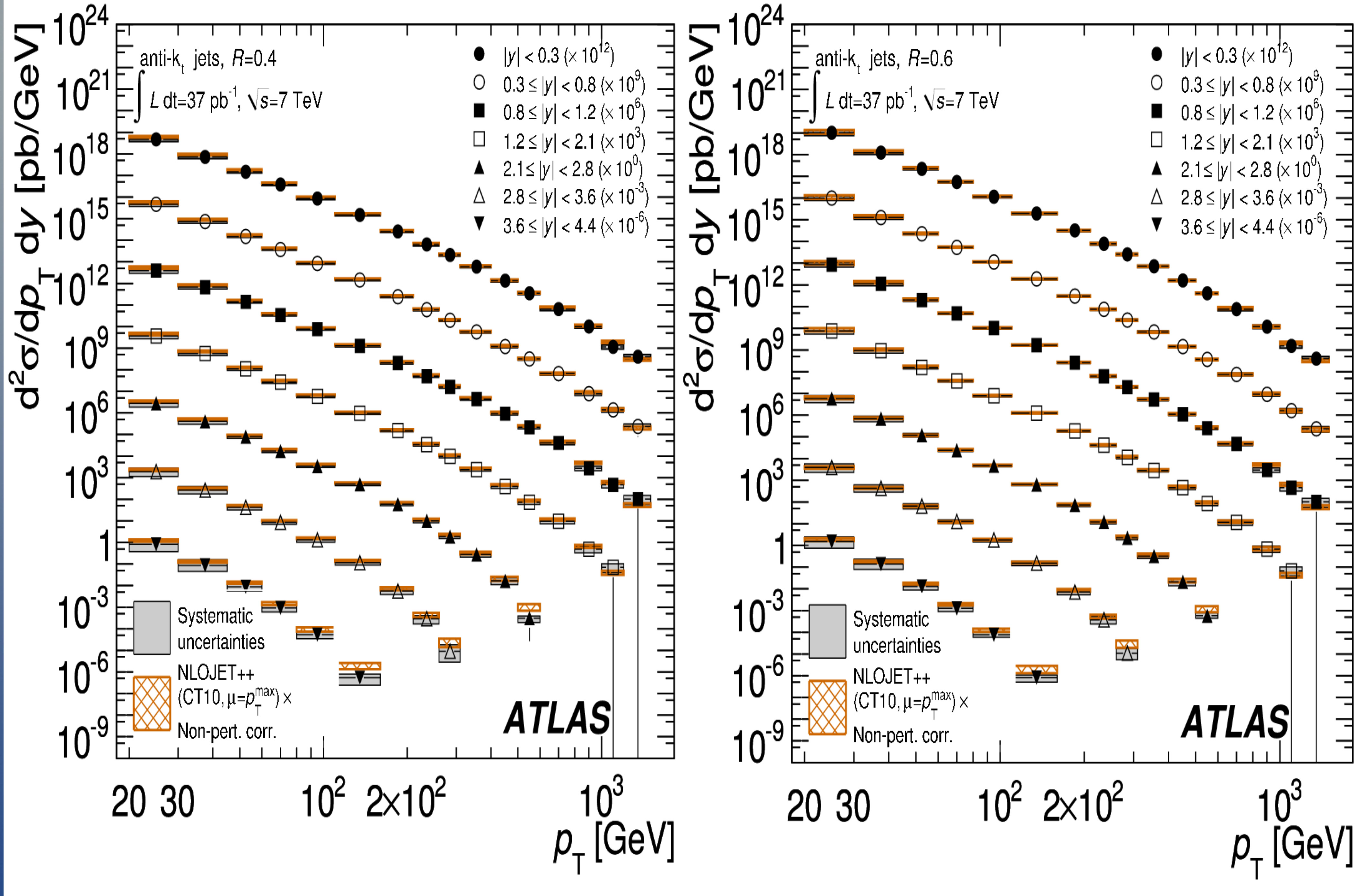


All bin correlations have been provided in the form of nuisance parameters to maximise information to be used in Pdf fits

C.Doglioni

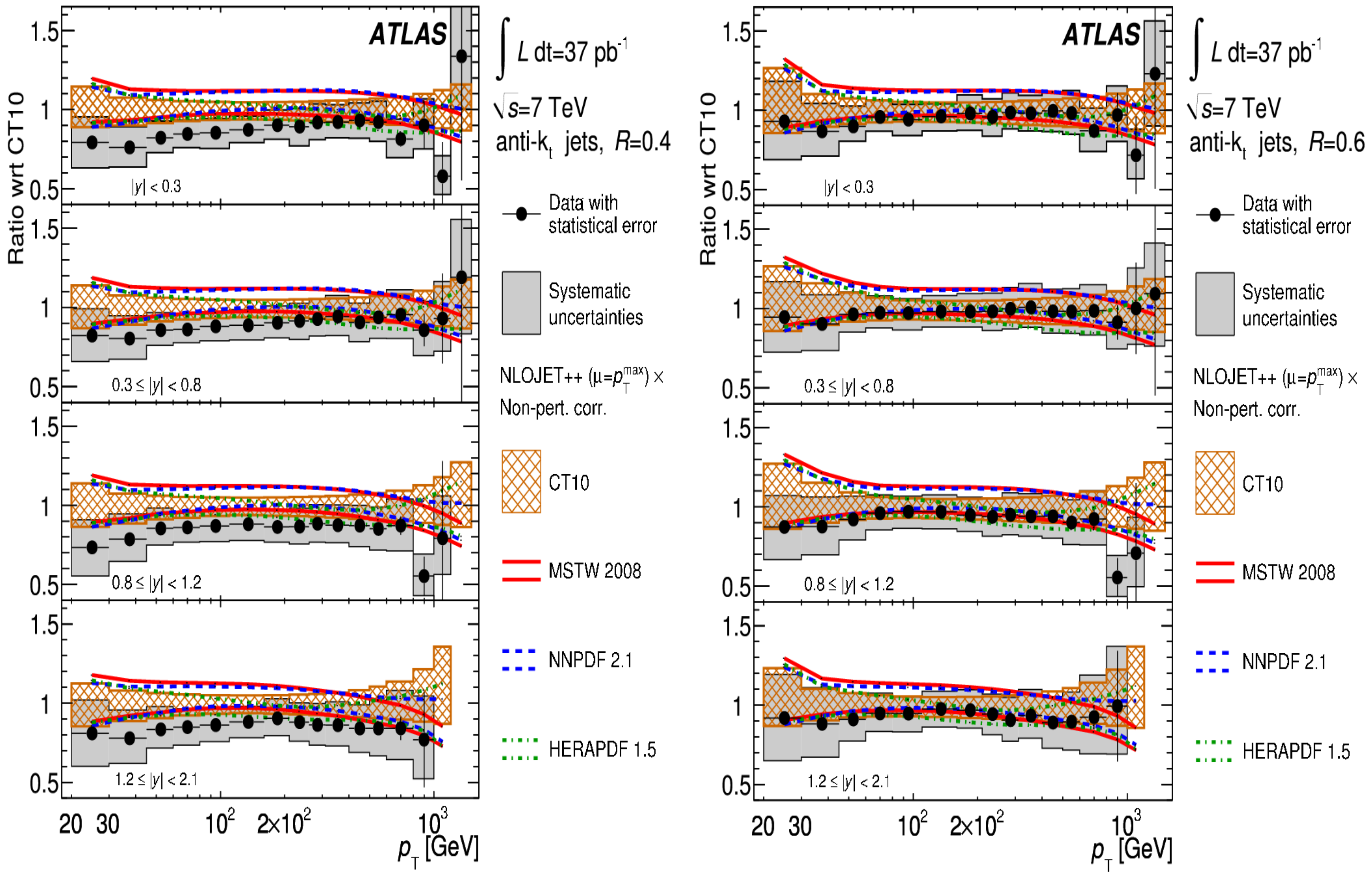
	$|y|$-bins						
Uncertainty Source	0-0.3	0.3-0.8	0.8-1.2	1.2-2.1	2.1-2.8	2.8-3.6	3.6-4.4
JES 1: Noise threshold	1	1	2	3	4	5	6
JES 2: Theory UE	7	7	8	9	10	11	12
JES 3: Theory Showering	13	13	14	15	16	17	18
JES 4: Non-closure	19	19	20	21	22	23	24
JES 5: Dead material	25	25	26	27	28	29	30
JES 6: Forward JES	31	31	31	31	31	31	31
JES 7: E/p response	32	32	33	34	35	36	37
JES 8: E/p selection	38	38	39	40	41	42	43
JES 9: EM+neutrals	44	44	45	46	47	48	49
JES 10: HAD E-scale	50	50	51	52	53	54	55
JES 11: High p_T	56	56	57	58	59	60	61
JES 12: E/p bias	62	62	63	64	65	66	67
JES 13: Test-beam bias	68	68	69	70	71	72	73
Unfolding	74	74	74	74	74	74	74
Jet matching	75	75	75	75	75	75	75
Jet energy resolution	76	76	77	78	79	80	81
y-resolution	82	82	82	82	82	82	82
Jet reconstruction eff.	83	83	83	83	84	85	86
Luminosity	87	87	87	87	87	87	87
JES 14: Pile-up (u_1)	u	u	u	u	u	u	u
Trigger (u_2)	u	u	u	u	u	u	u
Jet identification (u_3)	u	u	u	u	u	u	u

TABLE III. Description of bin-to-bin uncertainty correlation for the inclusive jet measurement. Each number corresponds to a nuisance parameter for which the corresponding uncertainty is fully correlated versus p_T. Bins with the same nuisance parameter are treated as fully correlated, while bins with different nuisance parameters are uncorrelated. The sources indicated by the letter "u" are uncorrelated both between p_T- and $|y|$-bins. The one-standard-deviation amplitude of the systematic effect associated with each nuisance parameter is detailed in Tables V–XVIII in Appendix B. The JES uncertainties for jets with $|y| \geq 0.8$ are determined relative to the JES of jets with $|y| < 0.8$. As a consequence, several of the uncertainties that are determined using jets with $|y| < 0.8$ are also propagated to the more forward rapidities (such as the E/p uncertainties). Descriptions of the JES uncertainty sources can be found in Refs. [59] and [69].

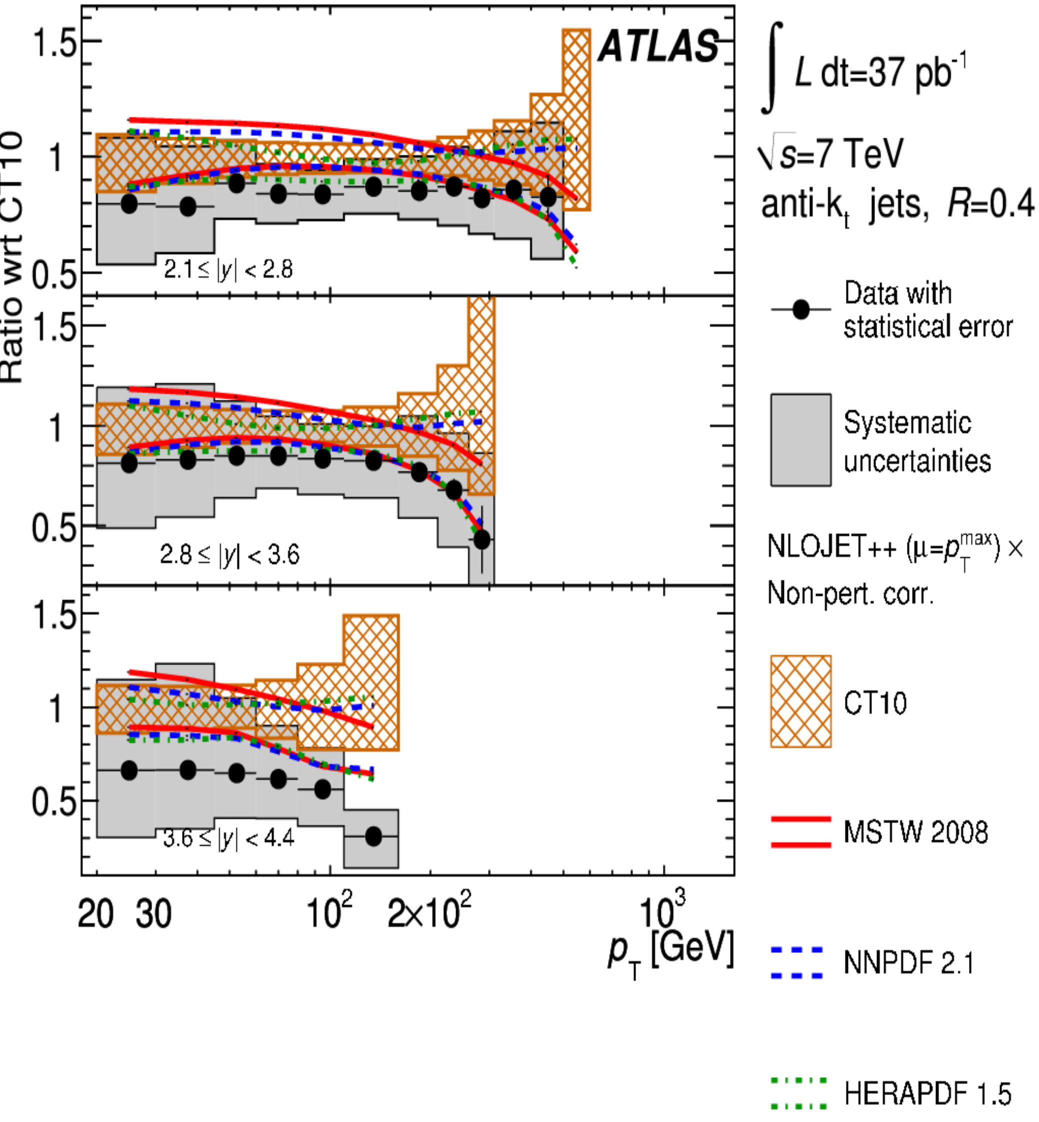


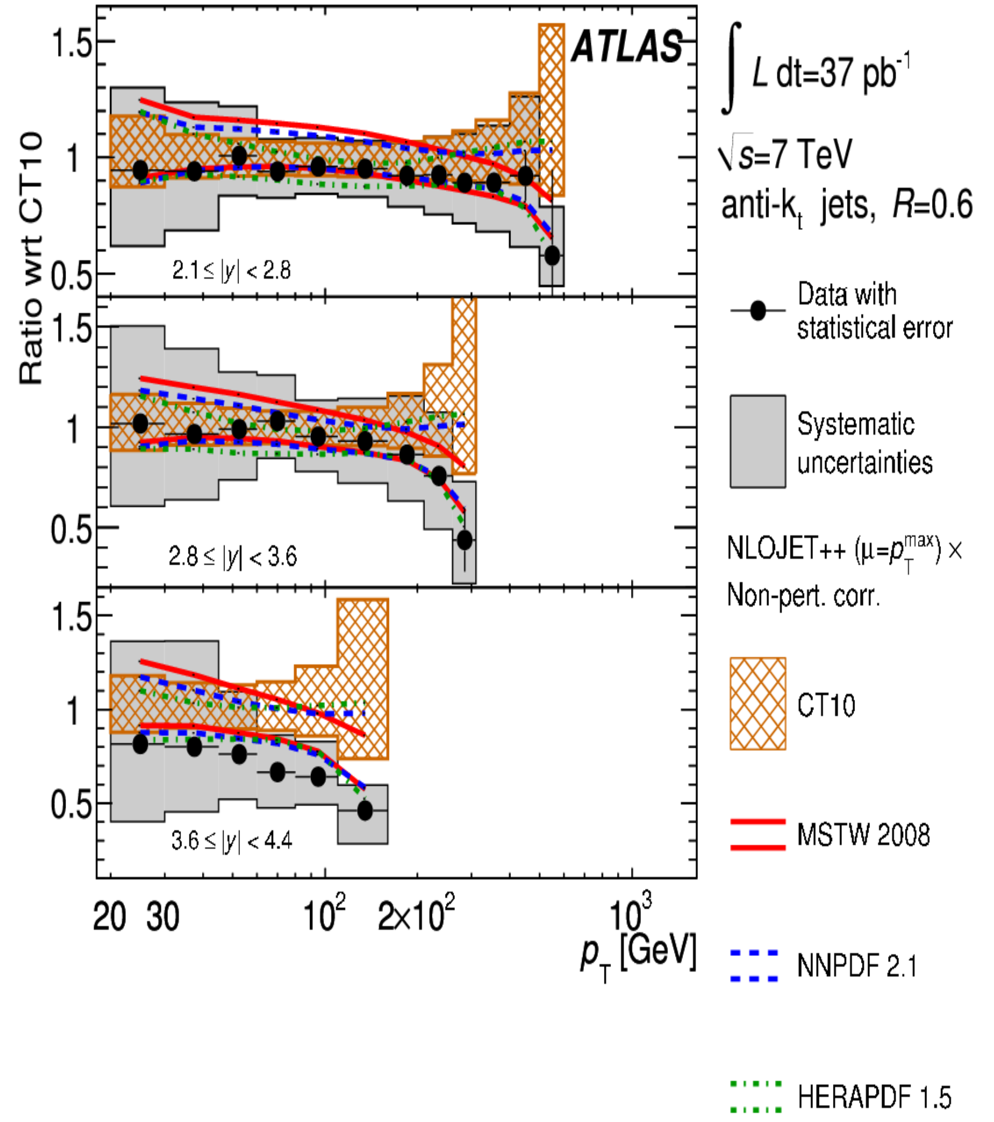
Inclusive jet cross section for antikt 06 jets after detector unfolding.
Pt range from 60 to 600 GeV, rapidity < 4.4

Ratio with NLO + soft corrections comparison between jet sizes (central)



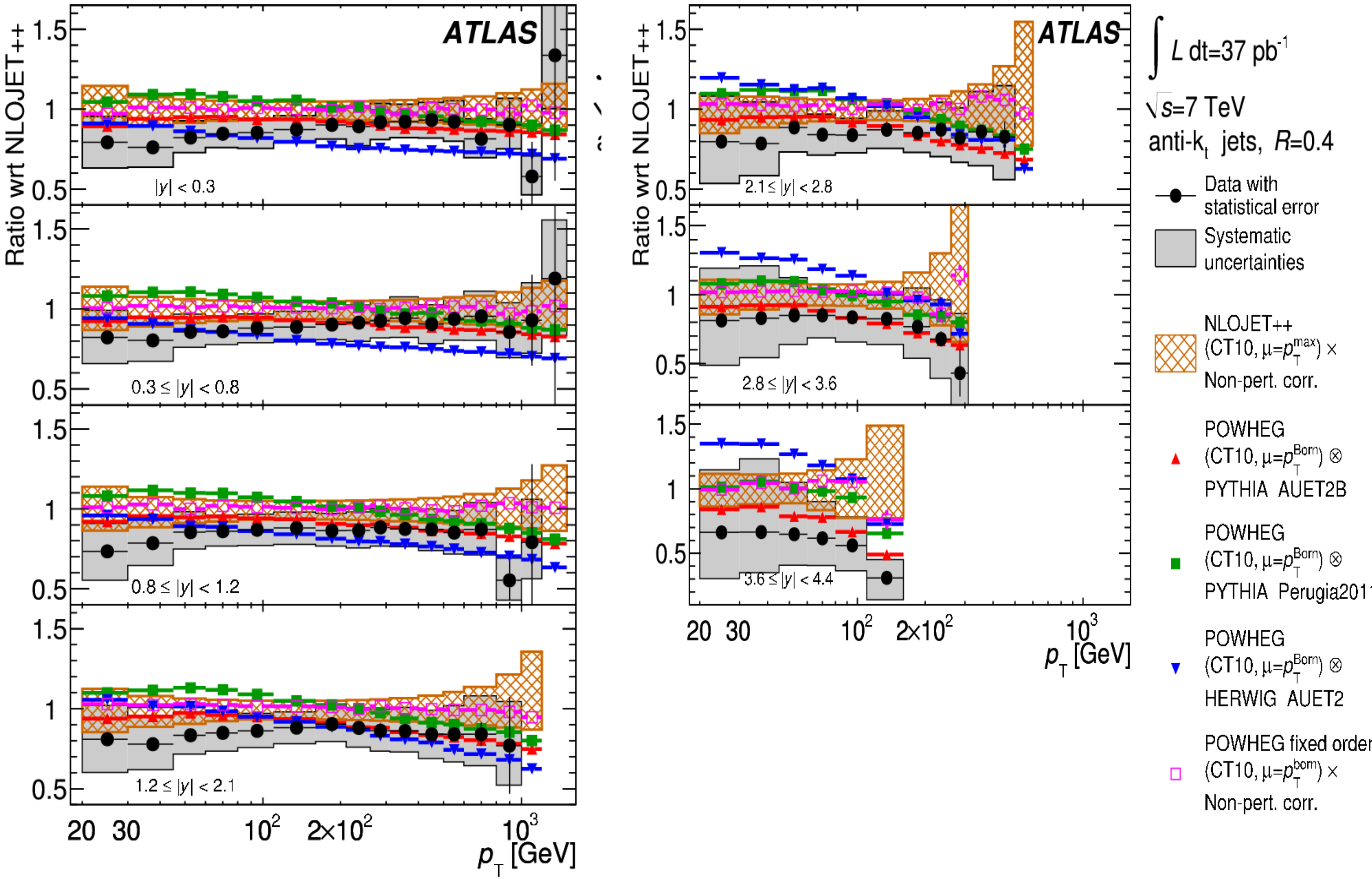
Ratio with NLO + soft corrections comparison between jet sizes (forward)



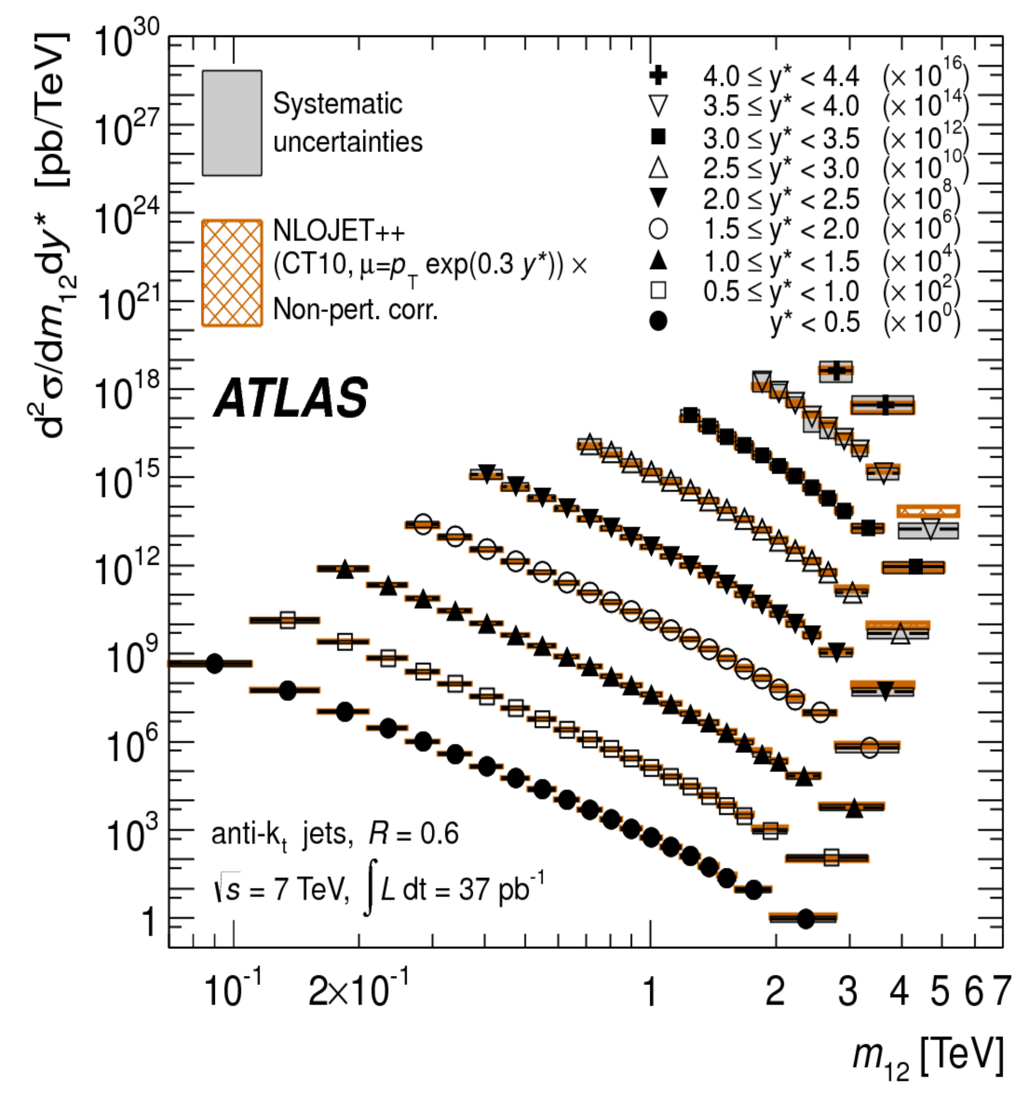


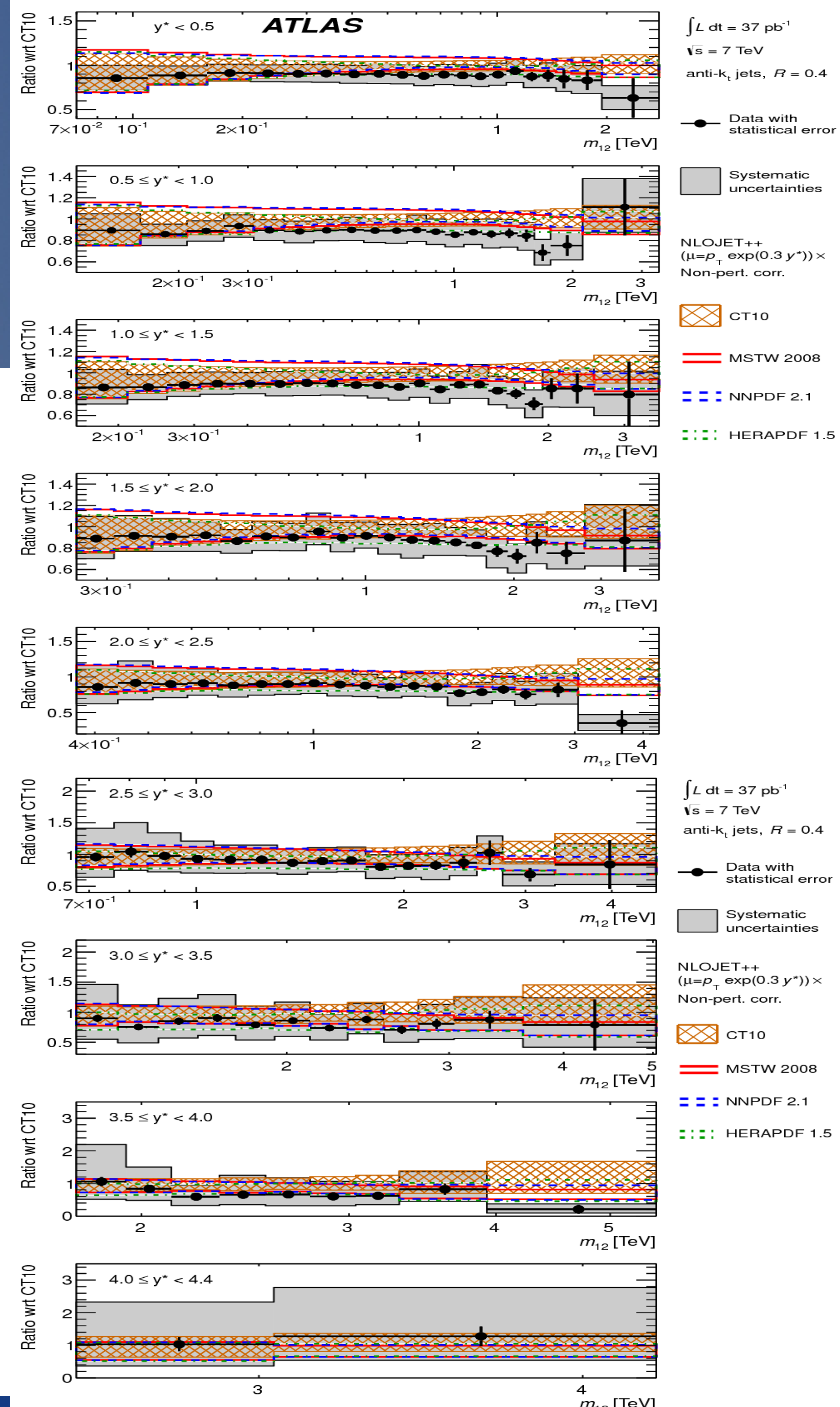
No big differences between cone sizes once soft corrections are applied

Comparison with Powheg (0.4)

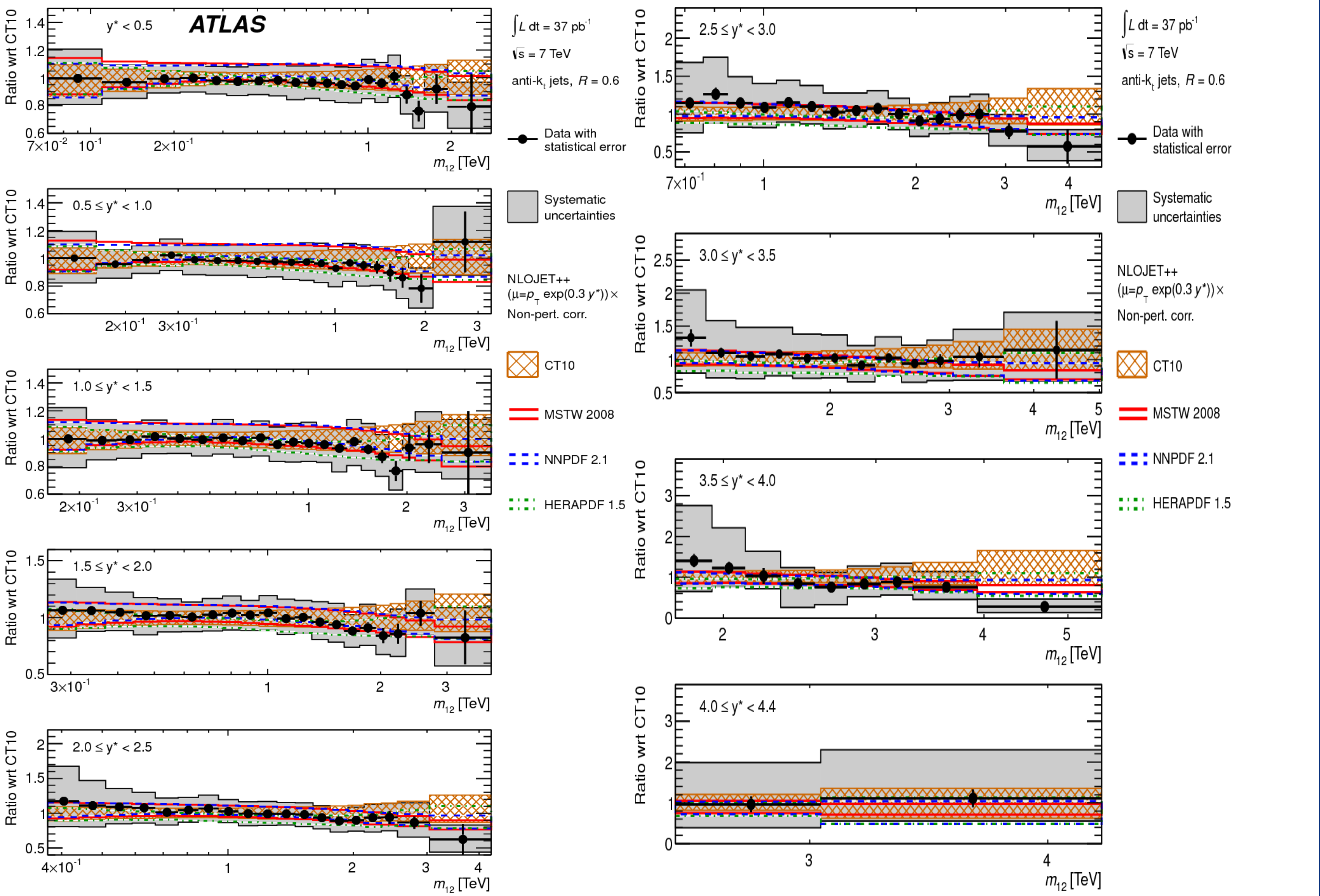


Dijet cross-section and ratio (0.6)

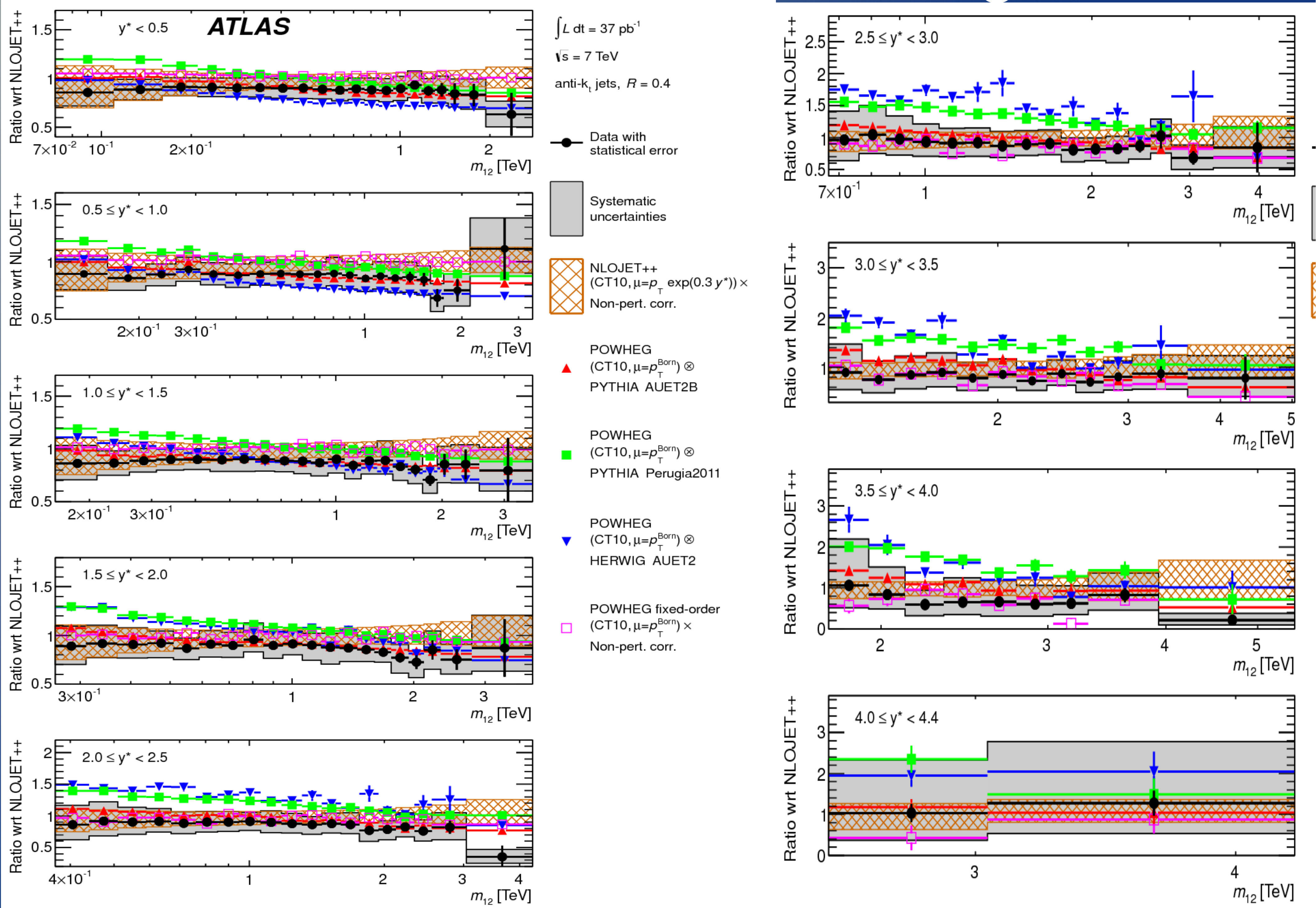




Pdf comparisons (0.4)

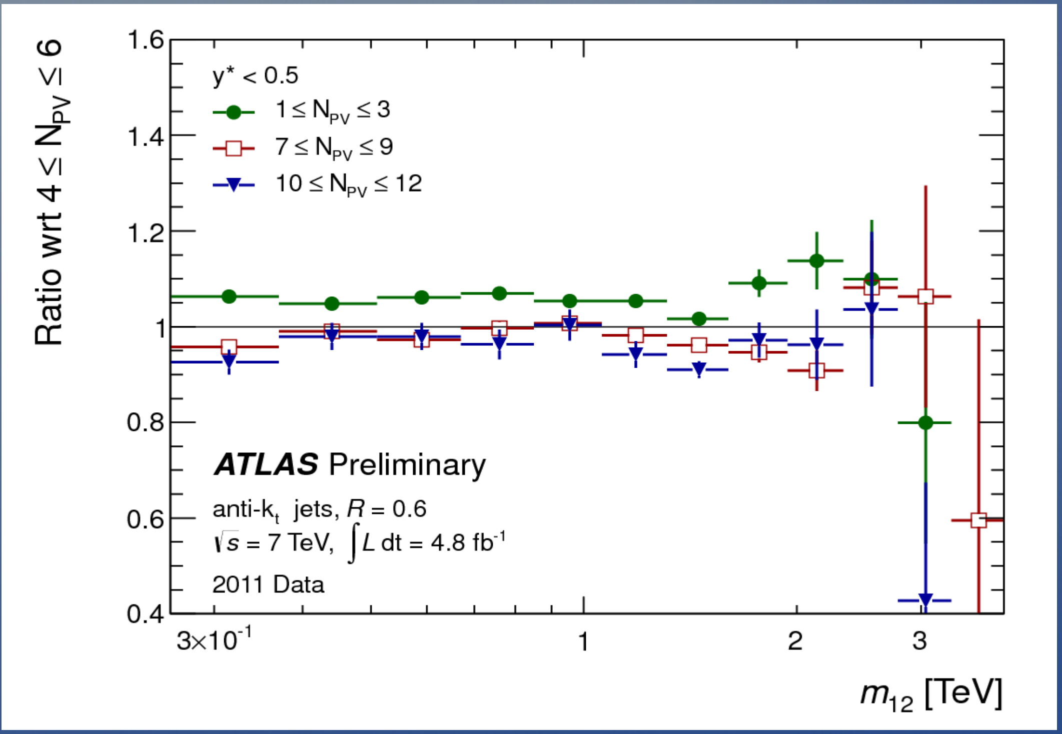


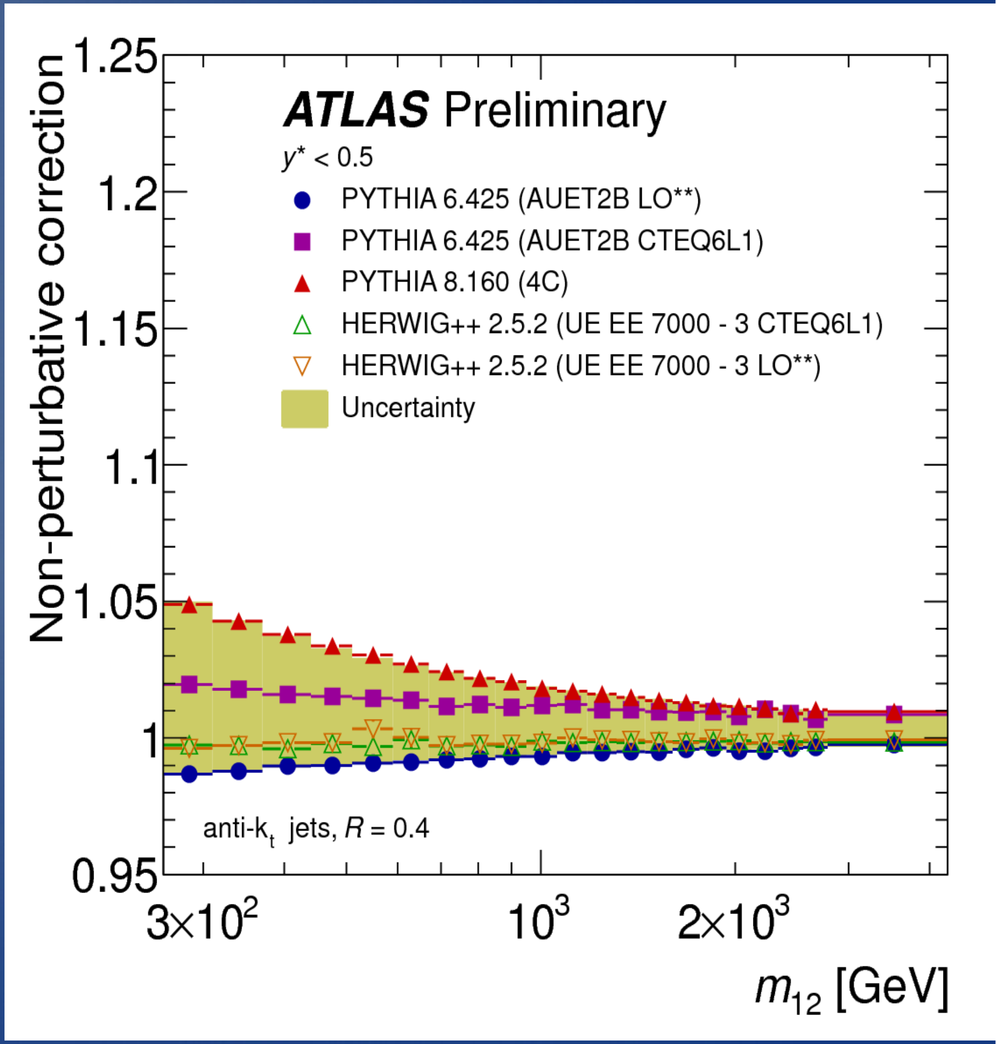
Comparison with Powheg



The 2011 dijet measurement

The big issue is controlling pileup, first measurement performed in the central region (|eta|<2.8) to have tracking, and with pt>60 GeV

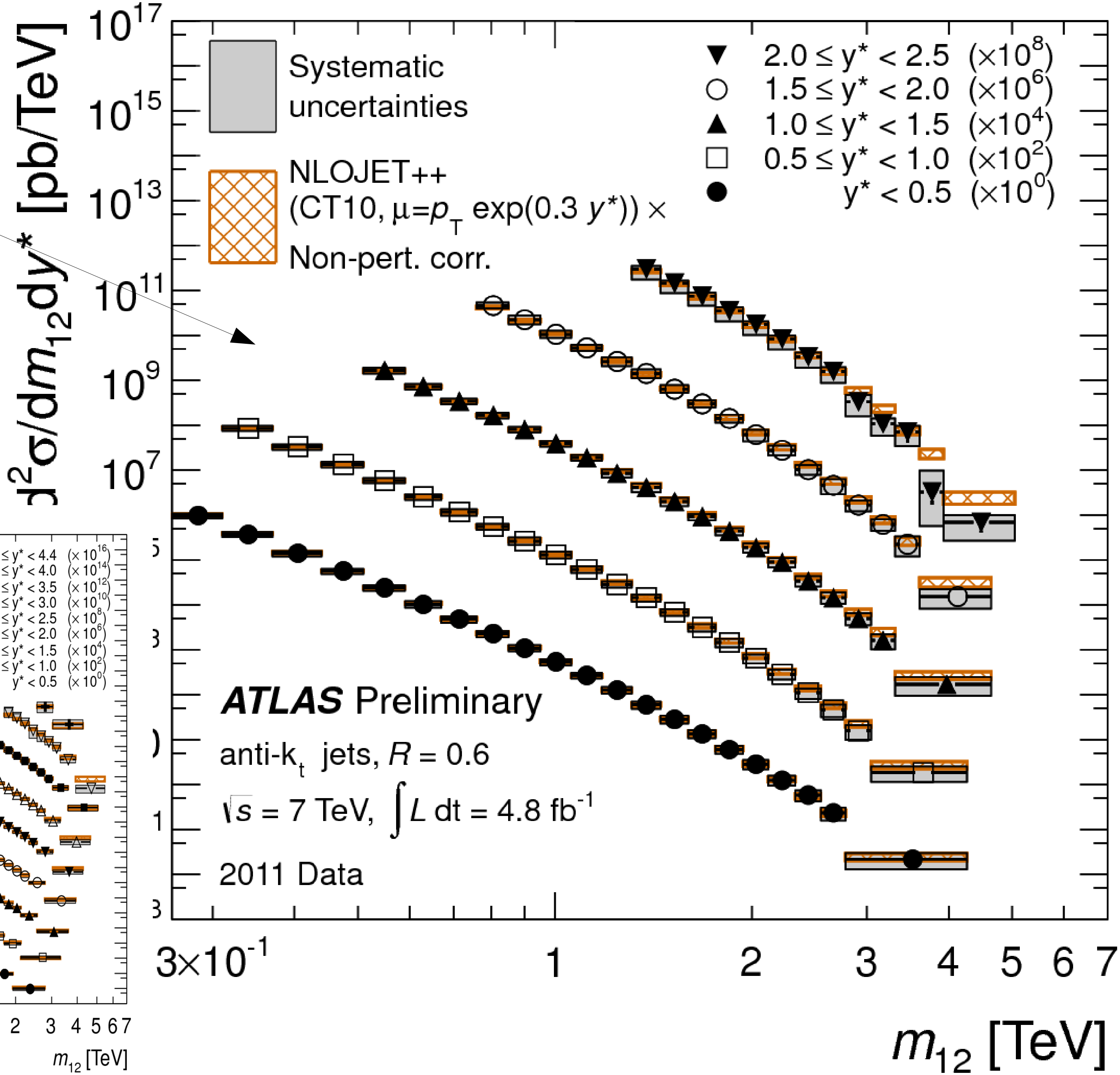




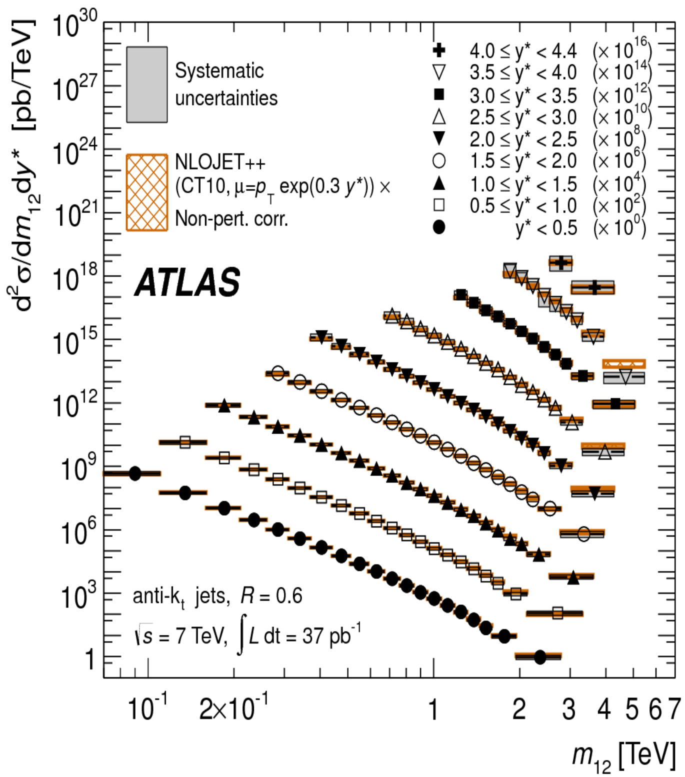
More MC statistics and different tunes

New

Old



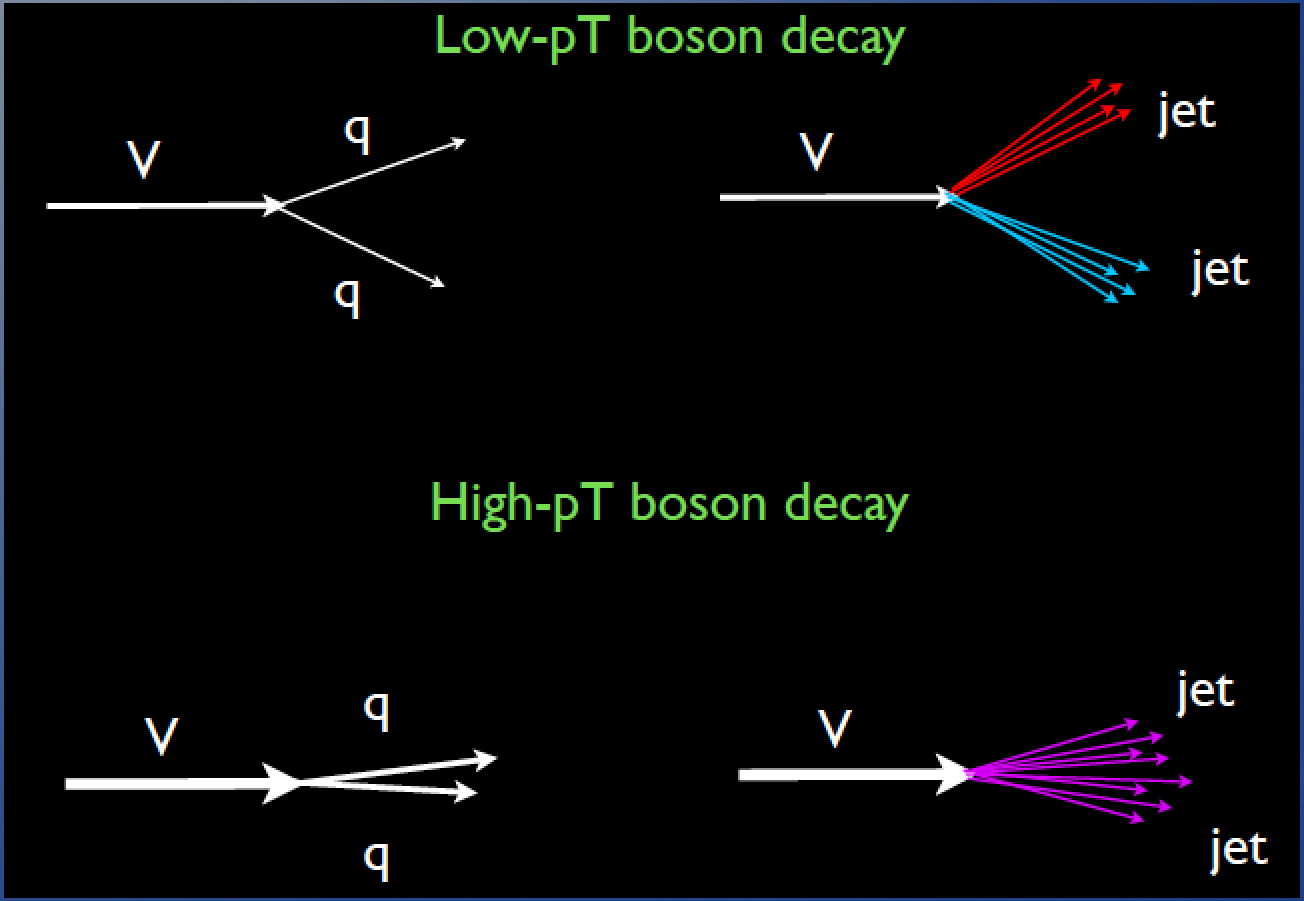
Systematic uncertainties
NLOJET++ (CT10, μ=p_T exp(0.3 y^*)) × Non-pert. corr.
$2.0 \le y^* < 2.5$ ($\times 10^8$)
$1.5 \le y^* < 2.0$ ($\times 10^6$)
$1.0 \le y^* < 1.5$ ($\times 10^4$)
$0.5 \le y^* < 1.0$ ($\times 10^2$)
$y^* < 0.5$ ($\times 10^0$)
ATLAS Preliminary
anti-k_t jets, R = 0.6
$\sqrt{s}$ = 7 TeV, $\int L\,dt$ = 4.8 fb^{-1}
2011 Data
$d^2\sigma/dm_{12}dy^*$ [pb/TeV]
m_{12} [TeV]



Systematic uncertainties
NLOJET++ (CT10, μ=p_T exp(0.3 y^*)) × Non-pert. corr.
$4.0 \le y^* < 4.4$ ($\times 10^{16}$)
$3.5 \le y^* < 4.0$ ($\times 10^{14}$)
$3.0 \le y^* < 3.5$ ($\times 10^{12}$)
$2.5 \le y^* < 3.0$ ($\times 10^{10}$)
$2.0 \le y^* < 2.5$ ($\times 10^8$)
$1.5 \le y^* < 2.0$ ($\times 10^6$)
$1.0 \le y^* < 1.5$ ($\times 10^4$)
$0.5 \le y^* < 1.0$ ($\times 10^2$)
$y^* < 0.5$ ($\times 10^0$)
ATLAS
anti-k_t jets, R = 0.6
$\sqrt{s}$ = 7 TeV, $\int L\,dt$ = 37 pb^{-1}
$d^2\sigma/dm_{12}dy^*$ [pb/TeV]
m_{12} [TeV]

Boosted decays of heavy objects

At the LHC for the first time EW-scale particles can have such large momenta that their hadronic decays can be reconstructed as a single “fat” jet



Idea is not new: M.Seymour, Z.Phys 62 (1994!) 127

Higgs->bb

The real excitement came when Butterworth Davison Rubin and Salam (PRL 100, 242001 2008) have shown that the neglected W/Z + H ->bb could be seen in the boosted configuration

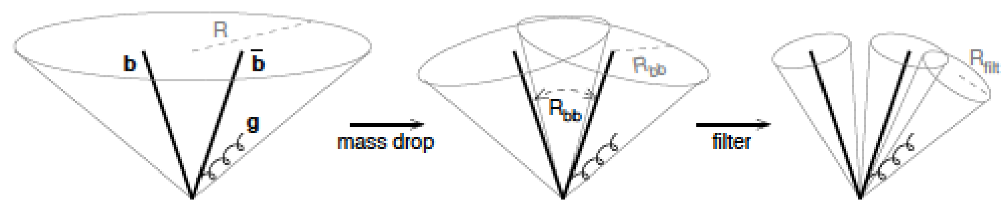


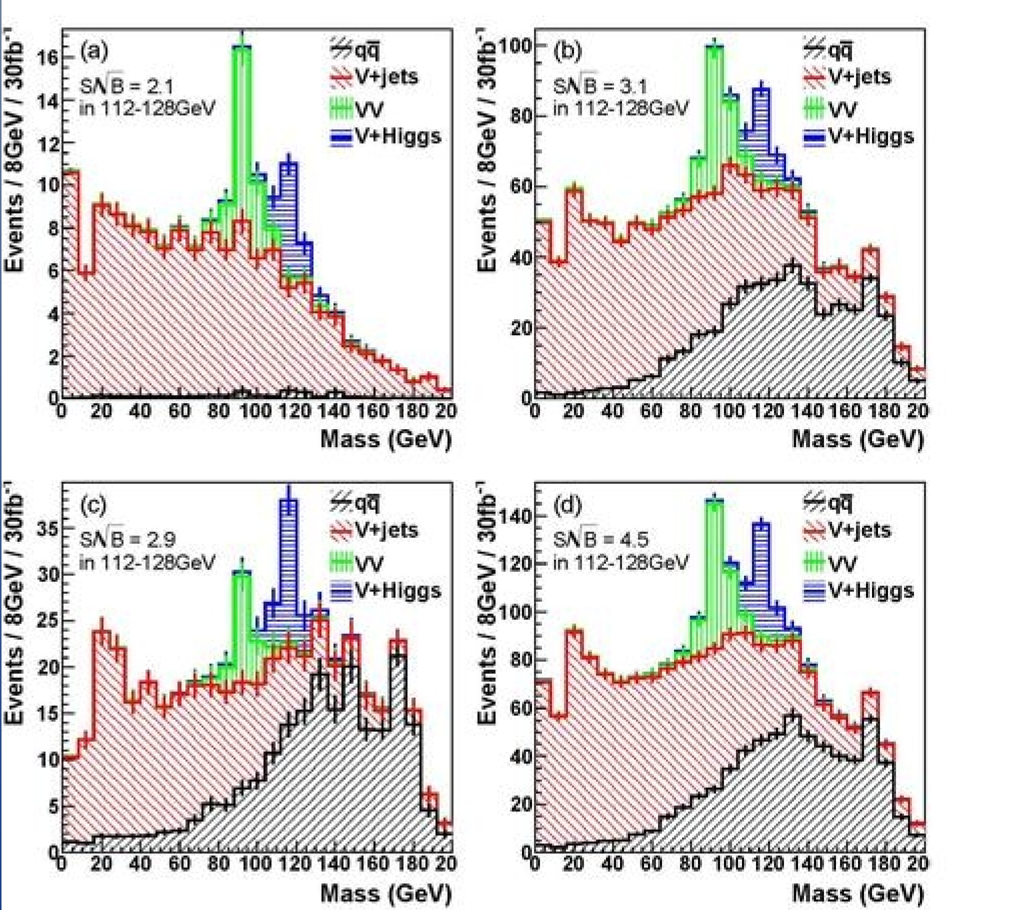
FIG. 1: The three stages of our jet analysis: starting from a hard massive jet on angular scale R, one identifies the Higgs neighbourhood within it by undoing the clustering (effectively shrinking the jet radius) until the jet splits into two subjets each with a significantly lower mass; within this region one then further reduces the radius to $R_{\rm filt}$ and takes the three hardest subjets, so as to filter away UE contamination while retaining hard perturbative radiation from the Higgs decay products.

They suggested a jet substructure technique, now known as filtering:

- Start from a fat jet reconstructed with Cambridge-Aachen (only angular information used), then deconstruct the last steps of clustering

- Require the 2-jet configuration to have a large mass drop with respect to the fat jet

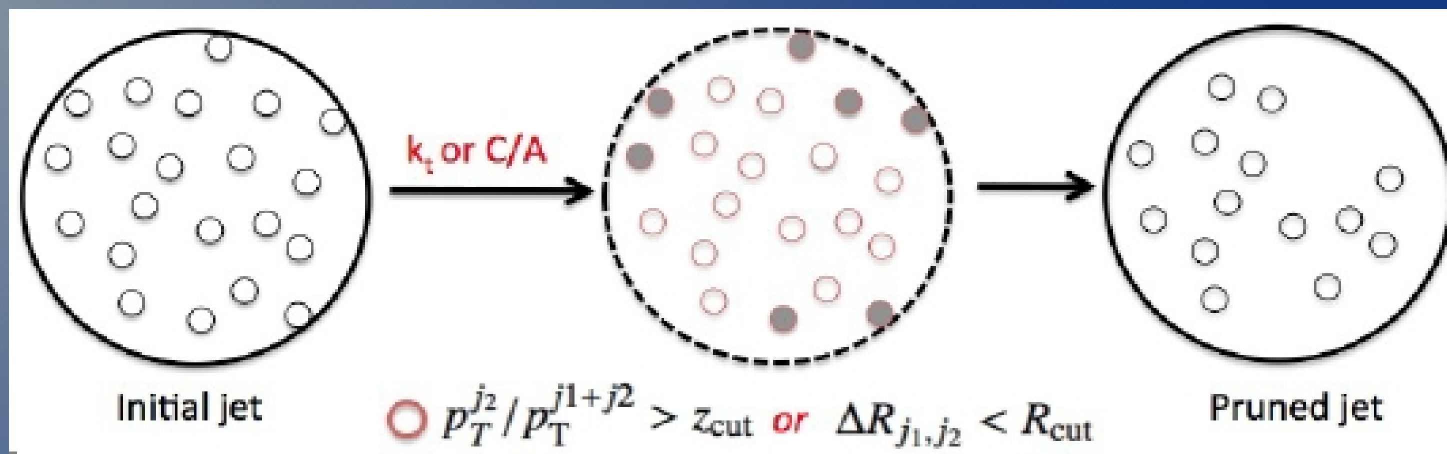
- Go back one more step, and keep the three hardest subjets (to remove soft components)



Other jet grooming techniques

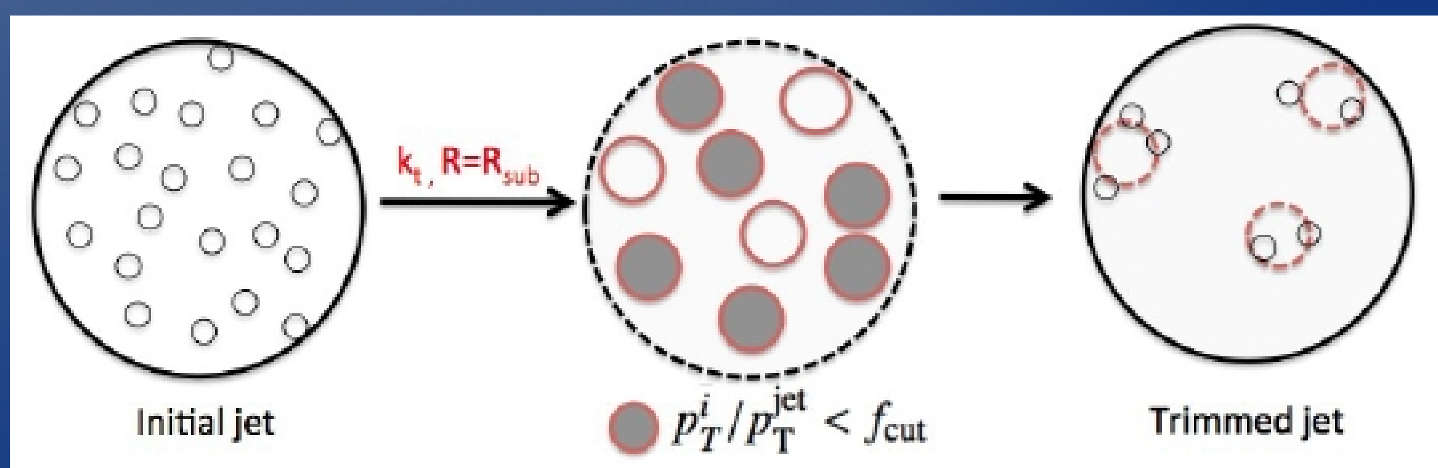
- Pruning (Ellis Vermillon Walsh 0912.0033)

Take a fat jet, and recombine constituents after removing wide angle and soft ones

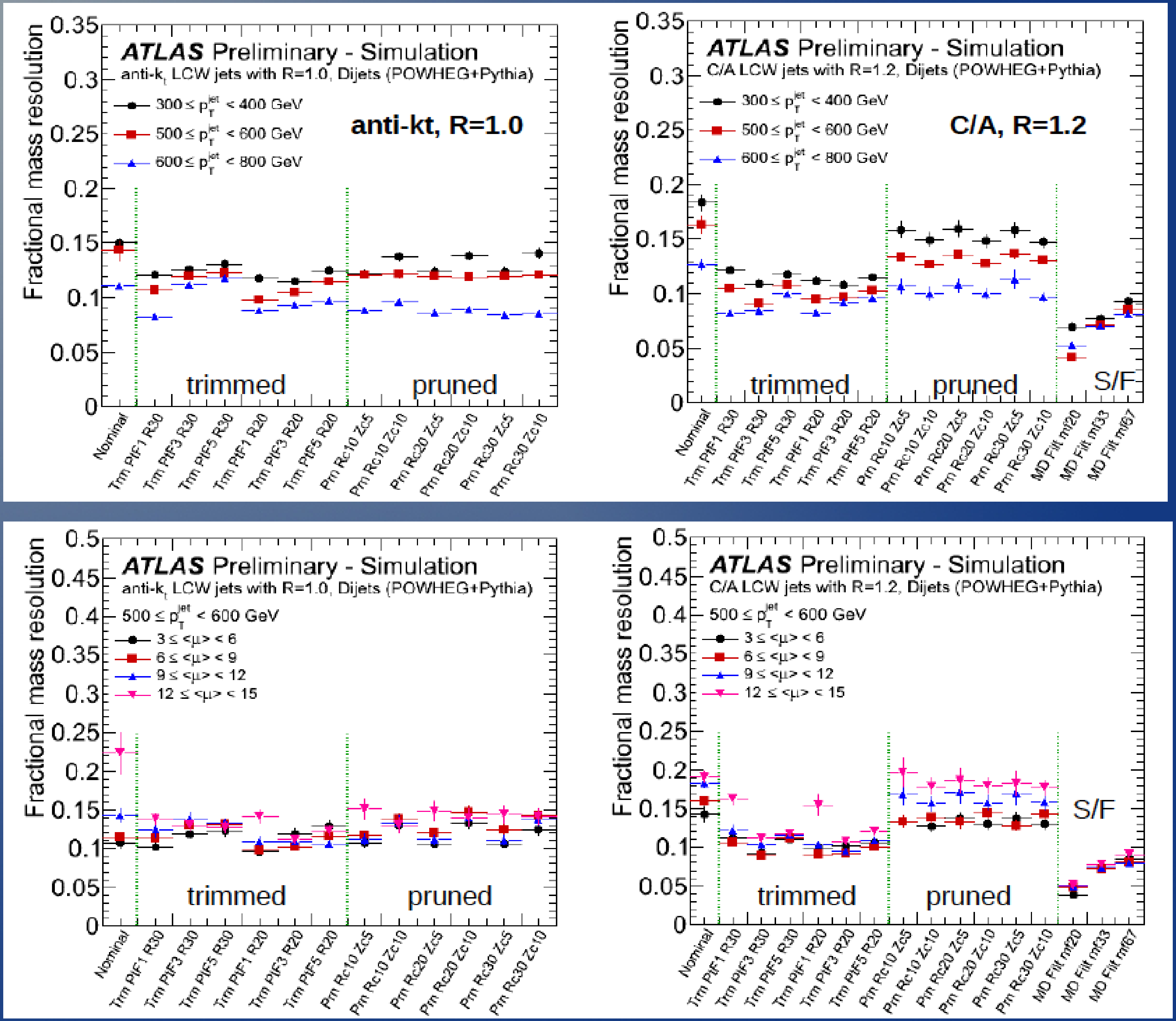


- Trimming (Krohn, Thaler, Wang 0912.1342)

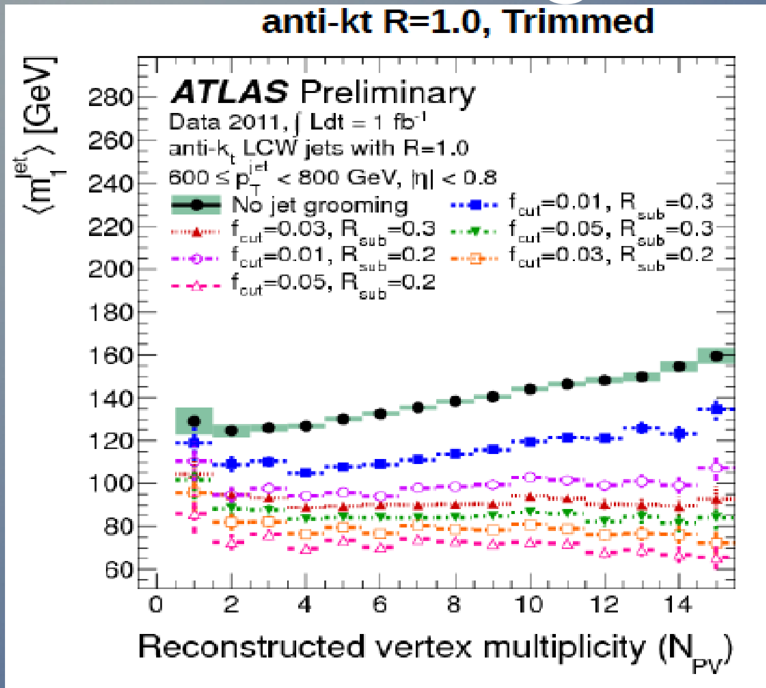
build fixed-radius subjets using kt, then remove those with pt/pt(jet)< fcut

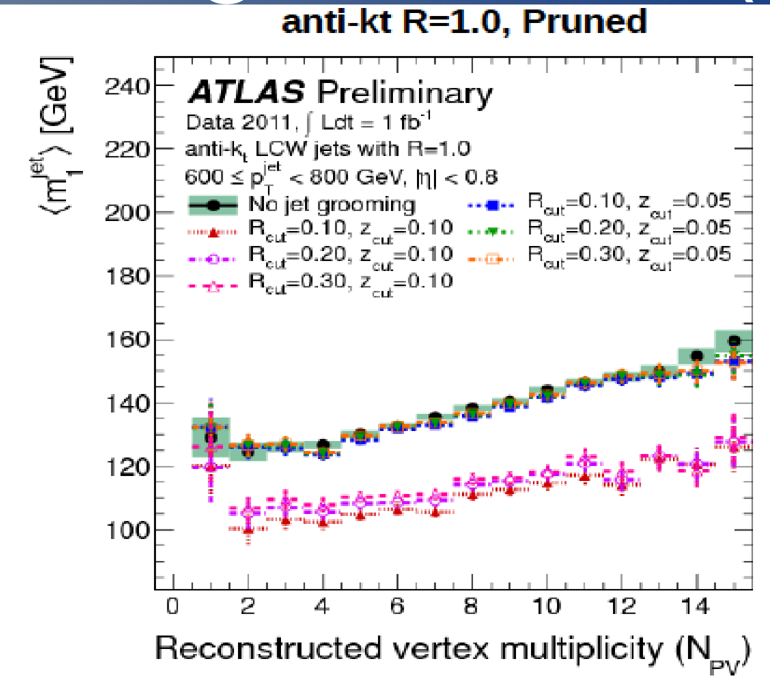


Simulated grooming performances

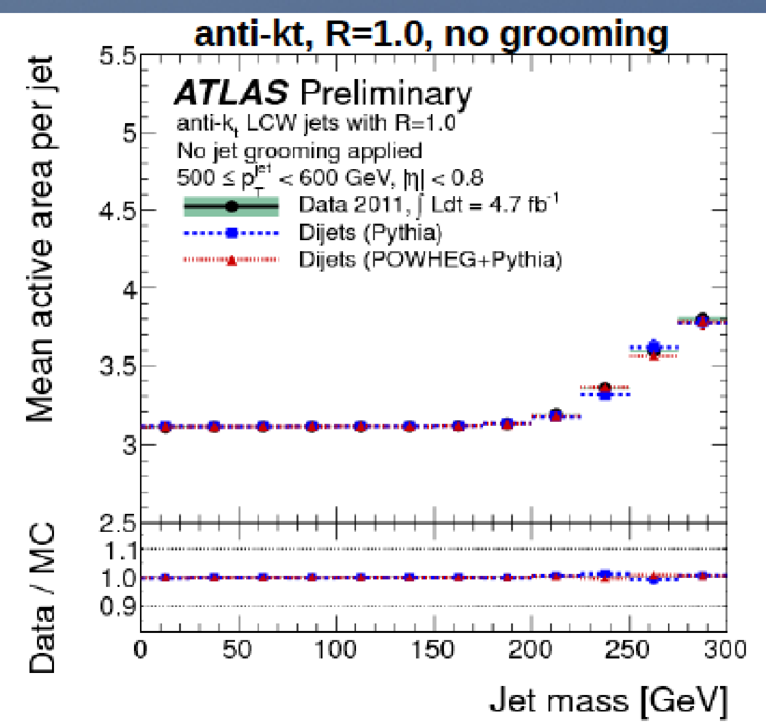


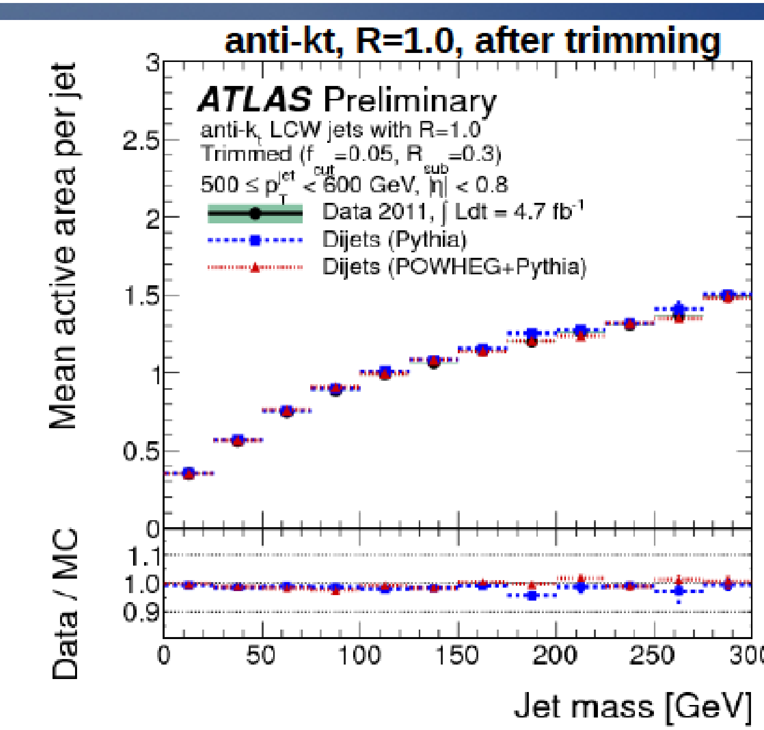
Effect of grooming on data (QCD)





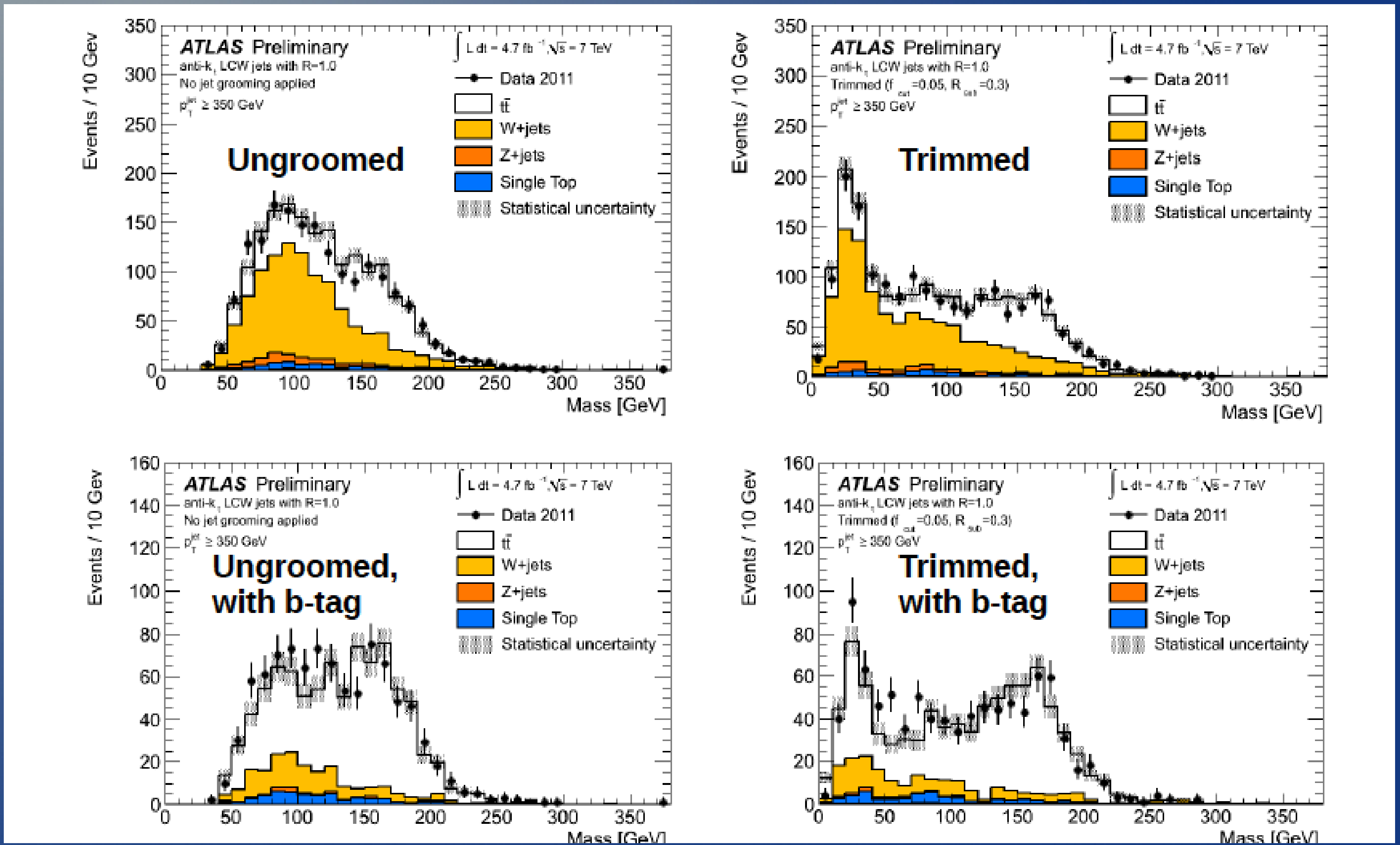
Mass dependence on pileup



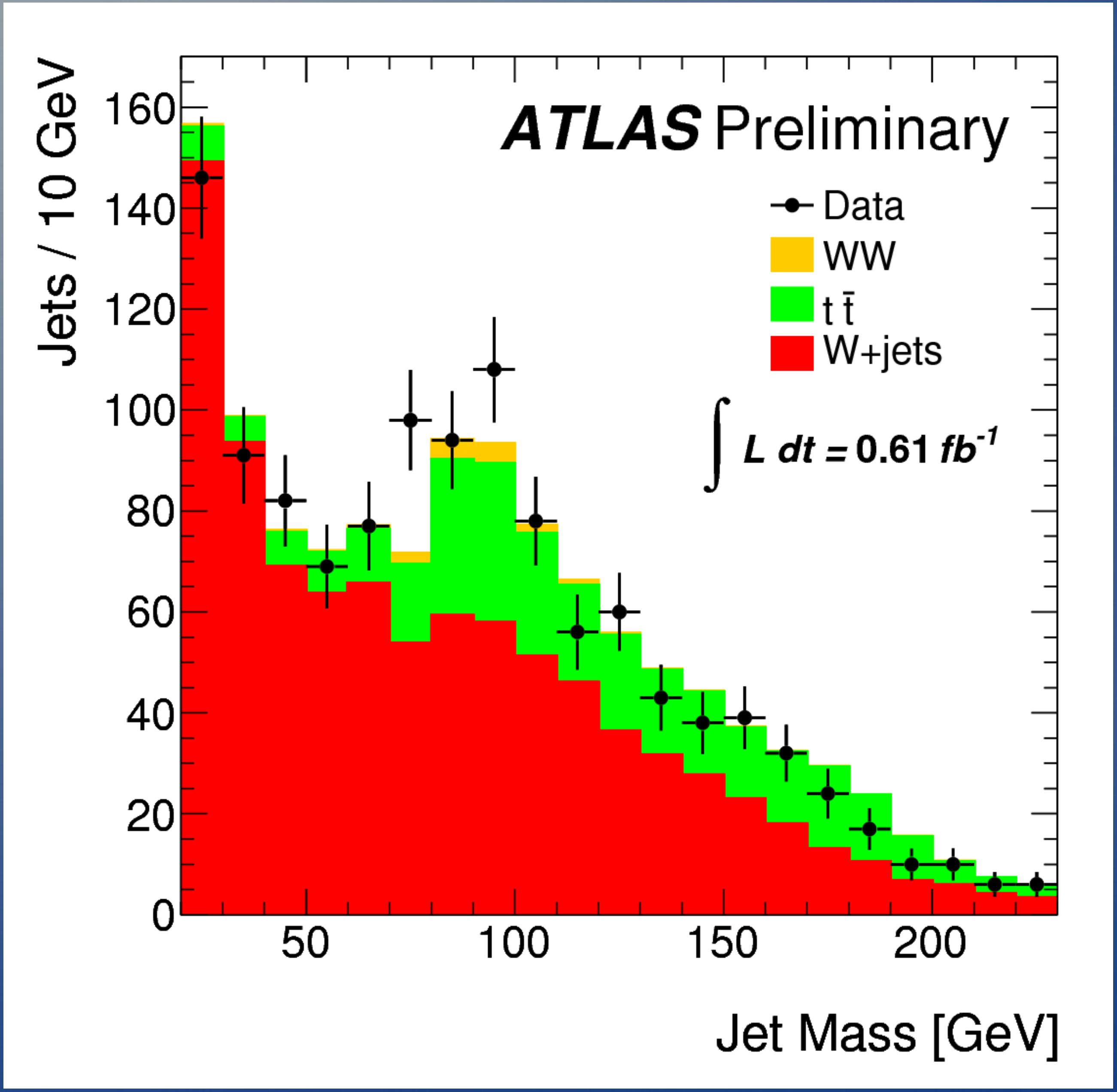


Active jet area

Effect of grooming on data (top)



W mass peak from semileptonic top



This is an old plot, a newer one is not available yet.

The W peak (obtained by BDRS filtering) is used as a standard candle for mass scale and resolution studies

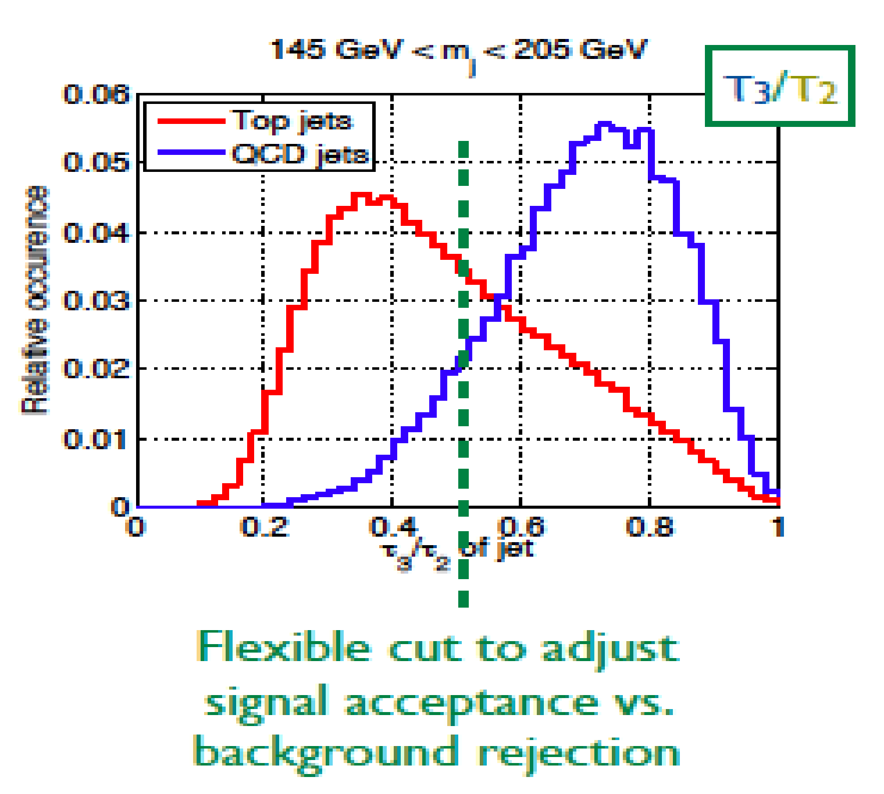
An analysis on inclusive hadronic W's and Z's (not from top) is also very advanced

N-subjettiness (Thaler-ValTilburg 1011.2268)

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\{\Delta R_{k,1}, \Delta R_{k,2}, \ldots, \Delta R_{k,N}\}$$

A measure of how much a jet can be naturally split in N subjets: 0 if <=N, 1 if >N

Typically ratios are used to tag 3-body vs 2-body decays



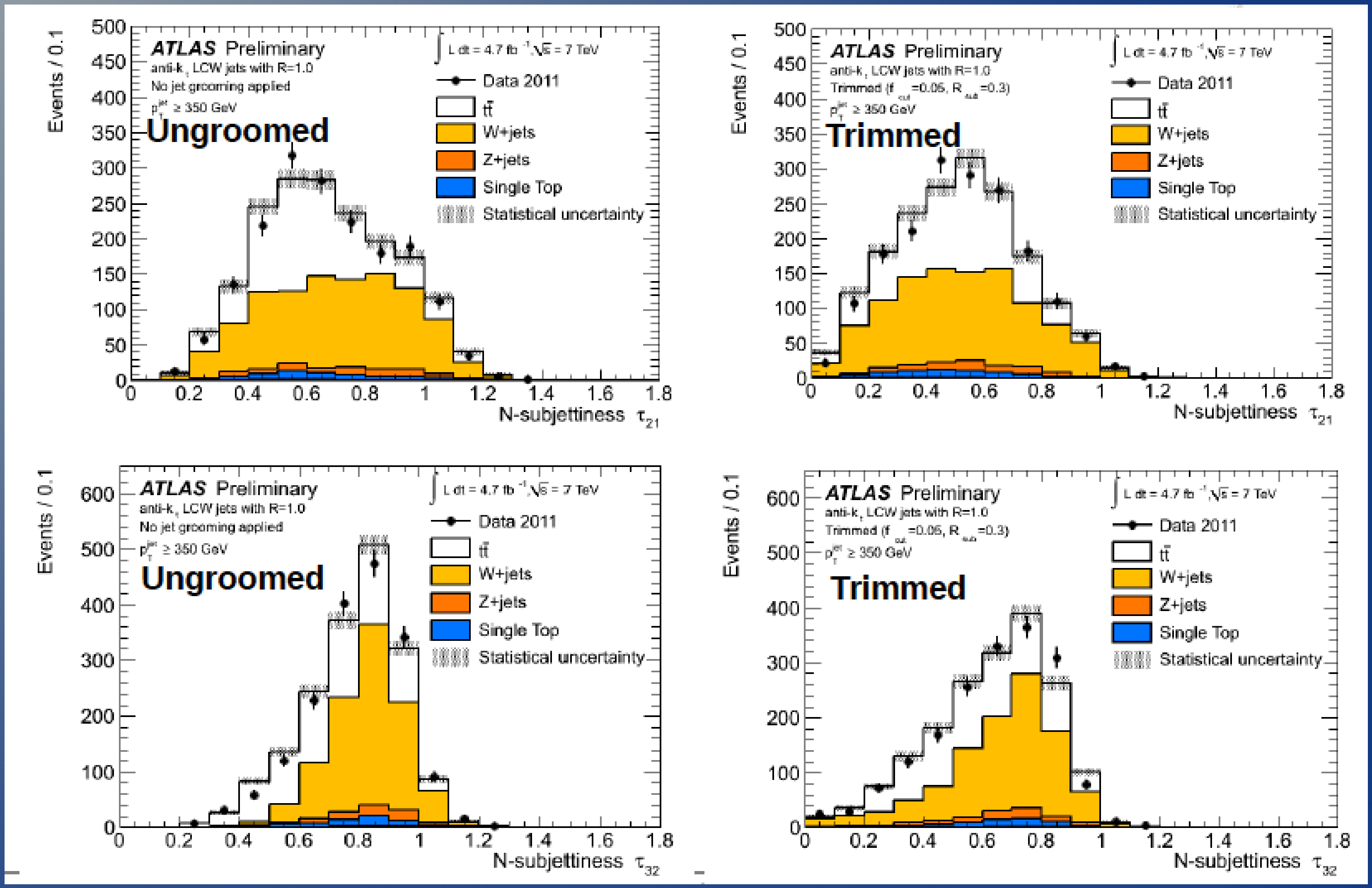
In a similar way, t2/t1 can discriminate W/Z from QCD

Can be calculated theoretically (Fiege et al. 1204.3898)

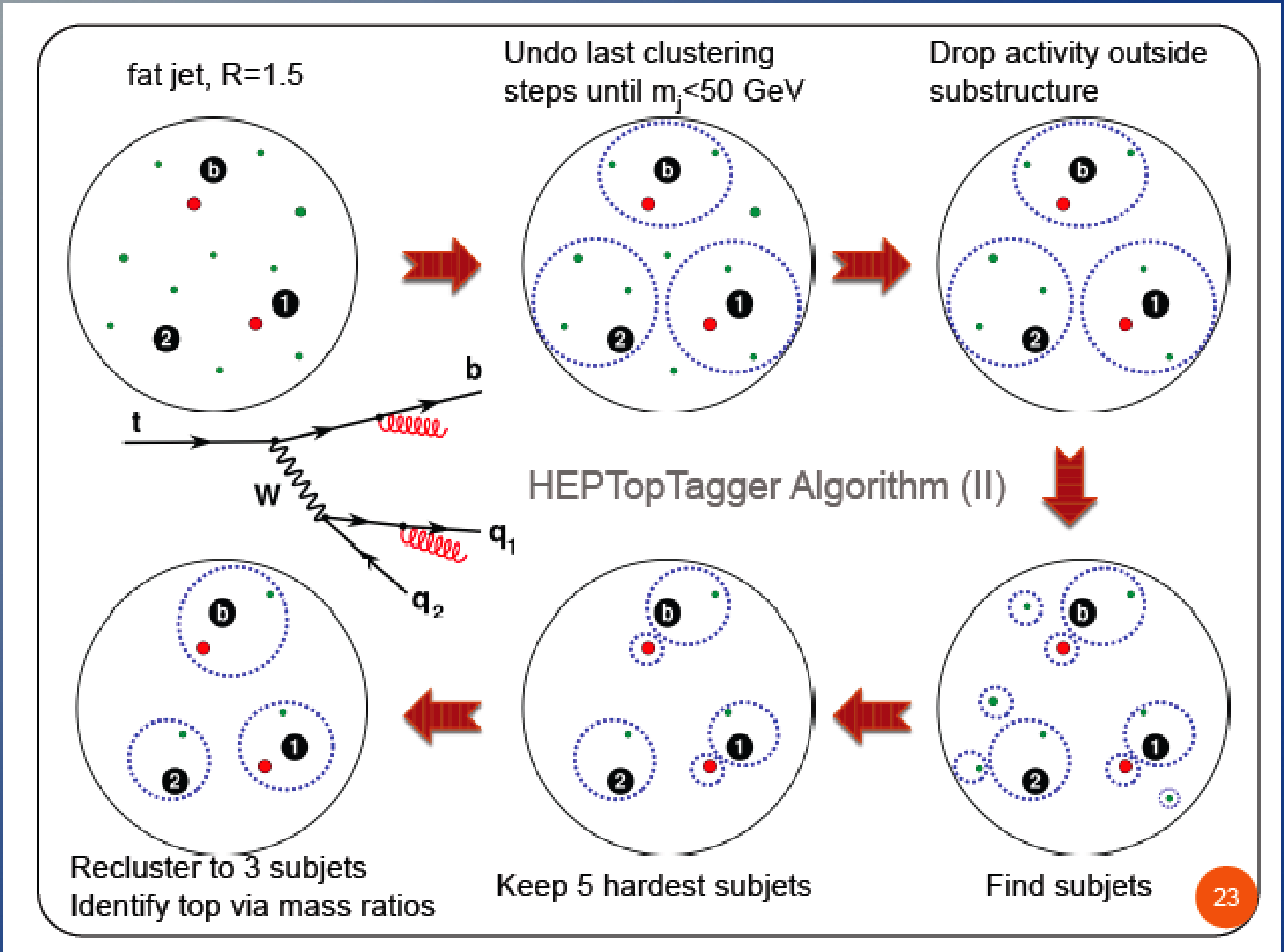
Can be used as a jet algorithm

Can be combined with grooming techniques

N-subjettiness in data (top)

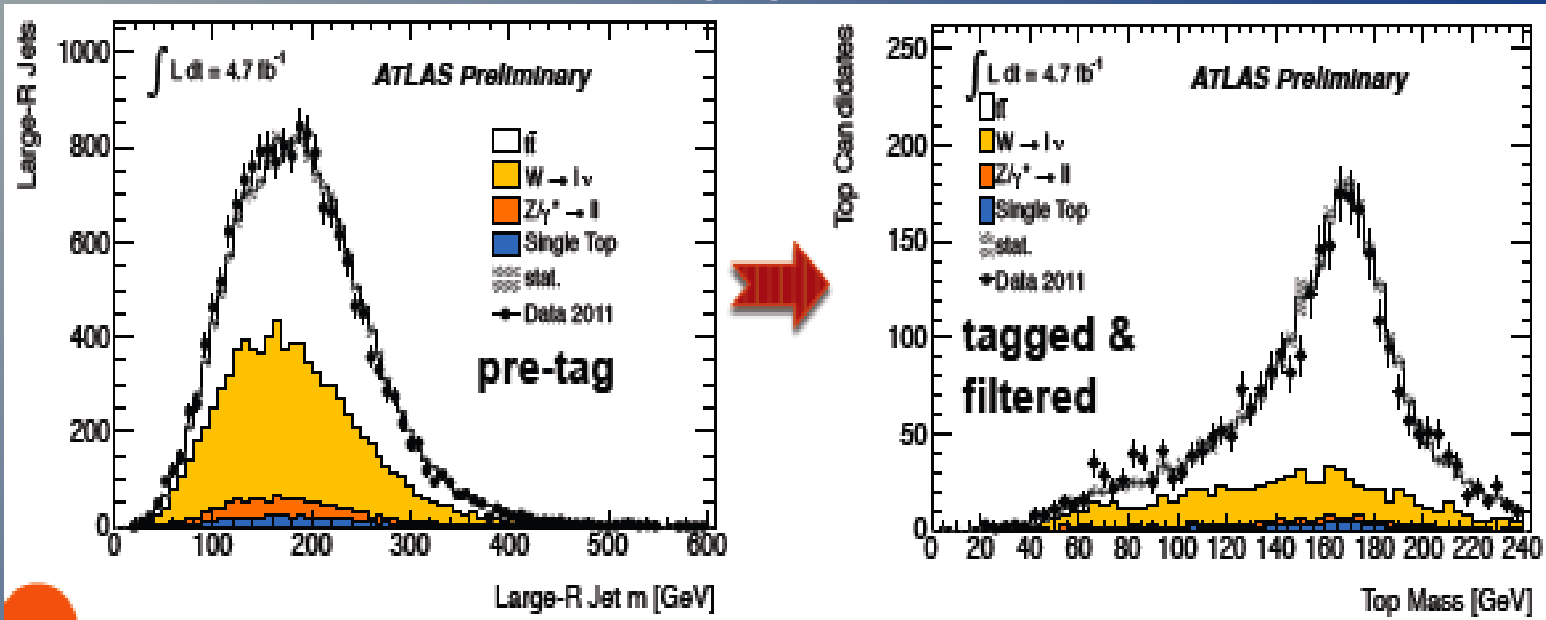


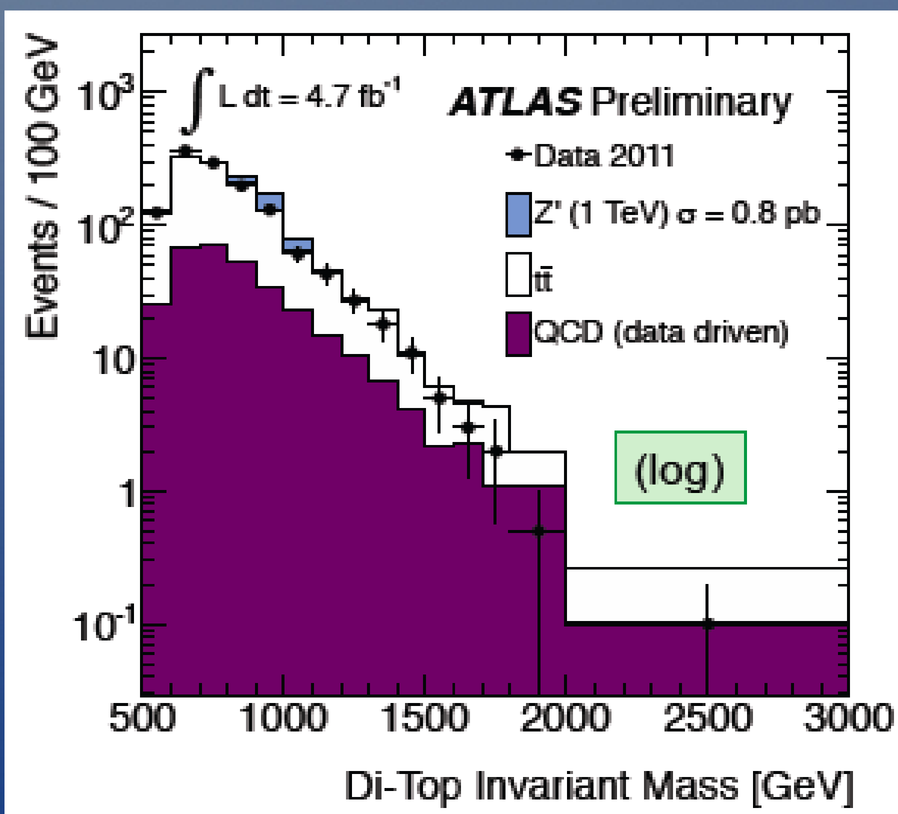
3-body final states: HepTopTagger



Plehn, Spannowsky, Takeuki, Zerwas 1006.2833

HepTopTagger results in ATLAS

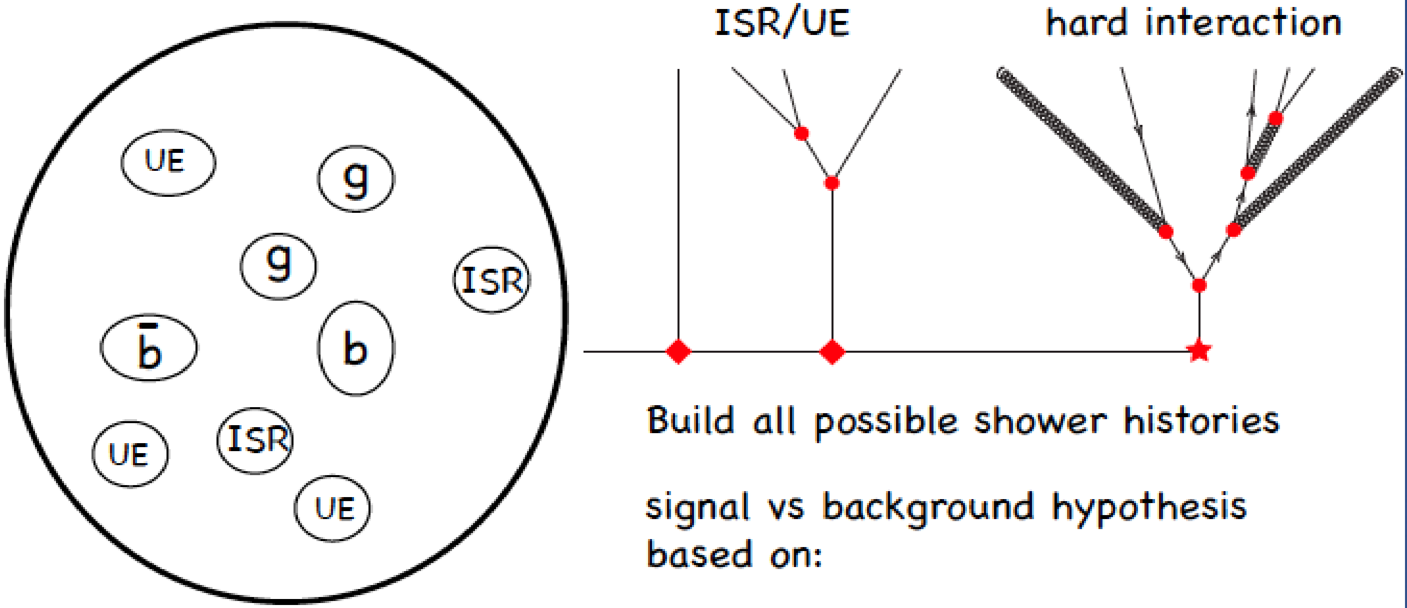




3-body tagging improves S/B for top searches, narrows the top mass peak and allows the search for heavy resonances decaying into boosted top

Shower deconstruction (Soper-Spannowsky 1102.3480)

Fat jet: R=1.2, anti-kT



microjets
R=0.15, kT

Build all possible shower histories

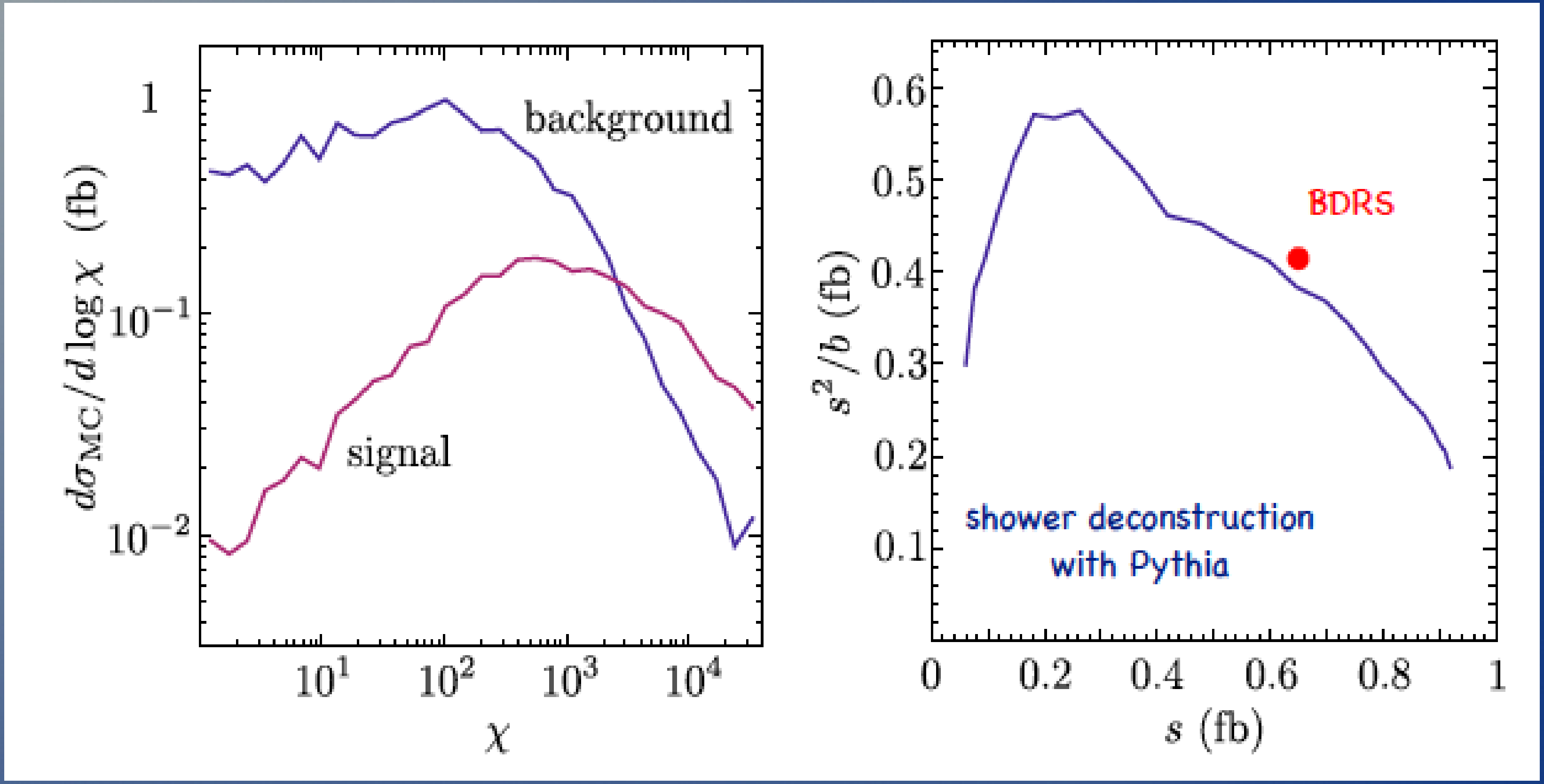
signal vs background hypothesis based on:

- Emission probabilities
- Color connection
- Kinematic requirements
- b-tag information

Maximal information approach to separate various types of events using Feynman rules

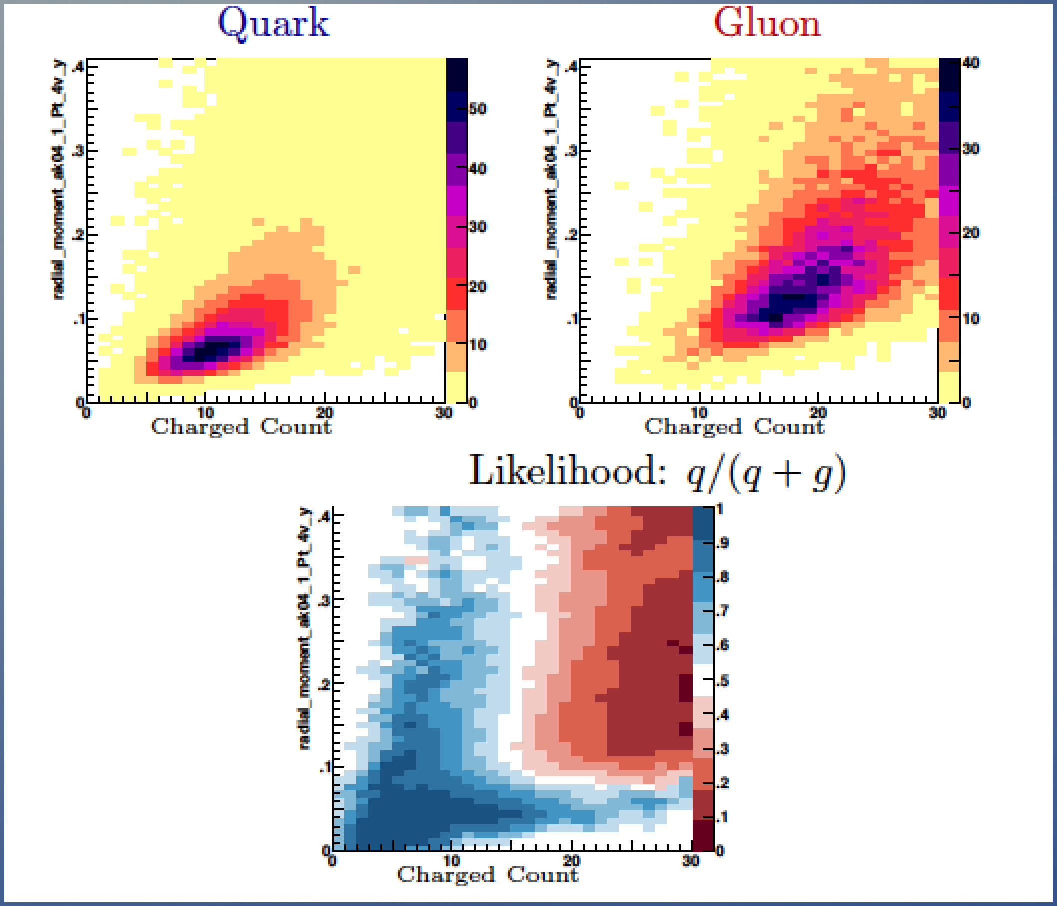
$$\chi(\{p,t\}_N) = \frac{P(\{p,t\}_N|\mathrm{S})}{P(\{p,t\}_N|\mathrm{B})} = \frac{\sum_{\text{histories}} H_{ISR} e^{-S_{I1}} \ldots \sum_{\text{histories}} H_H e^{-S_{s1}} H^s_{bg} e^{-S_{s2}} \ldots}{\sum_{\text{histories}} H_{ISR} e^{-S_{I1}} \ldots \sum_{\text{histories}} H_{gbb} e^{-S_{b1}} H^b_{bg} e^{-S_{b2}} \ldots}$$

Performance of shower deconstruction



In principle, the most powerful method since it uses full matrix element
in practice relies on finding (and calibrating!) micro-jets
An ATLAS implementation currently underway for top and hadronic W's

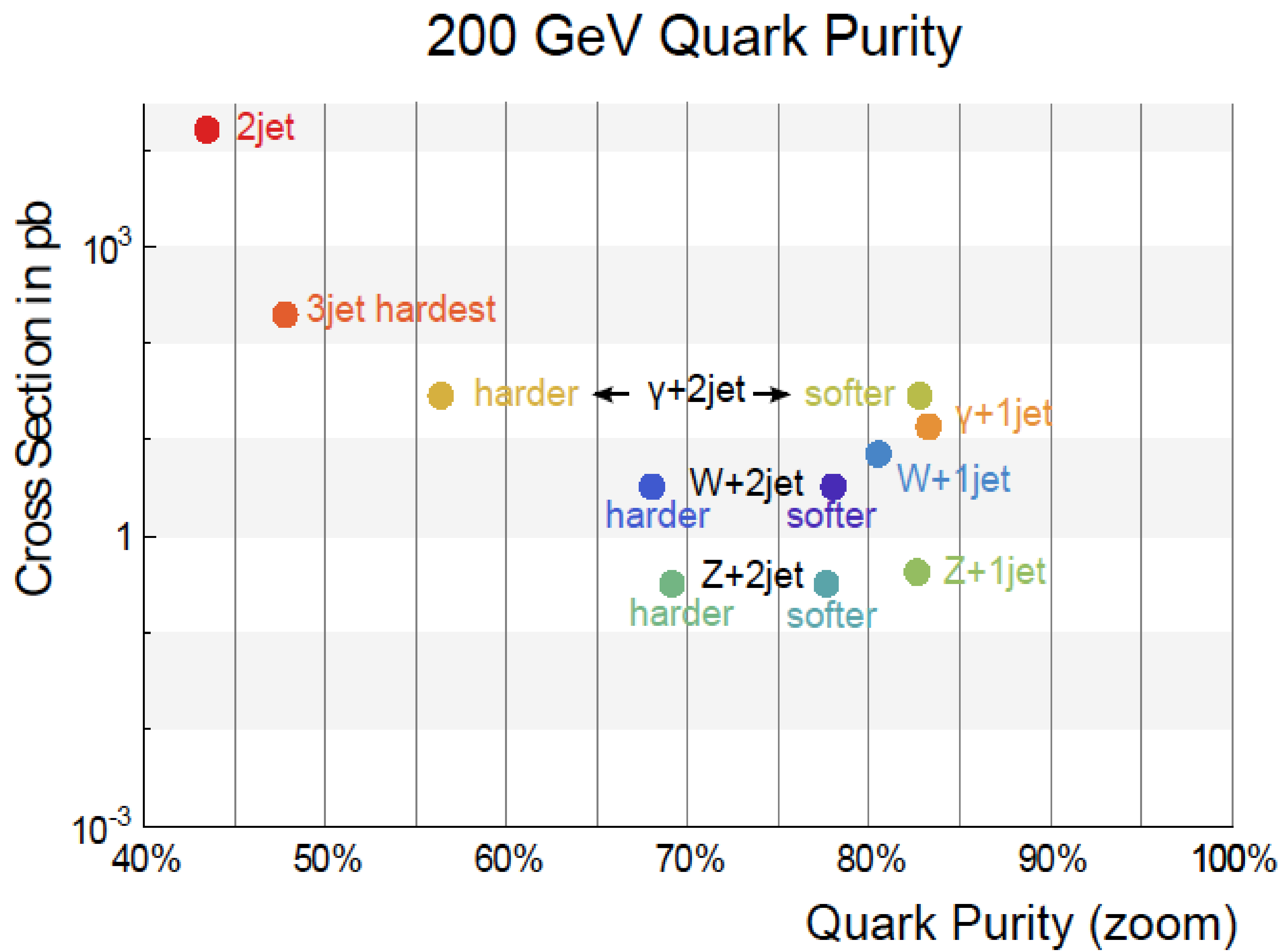
Quark-gluon separation



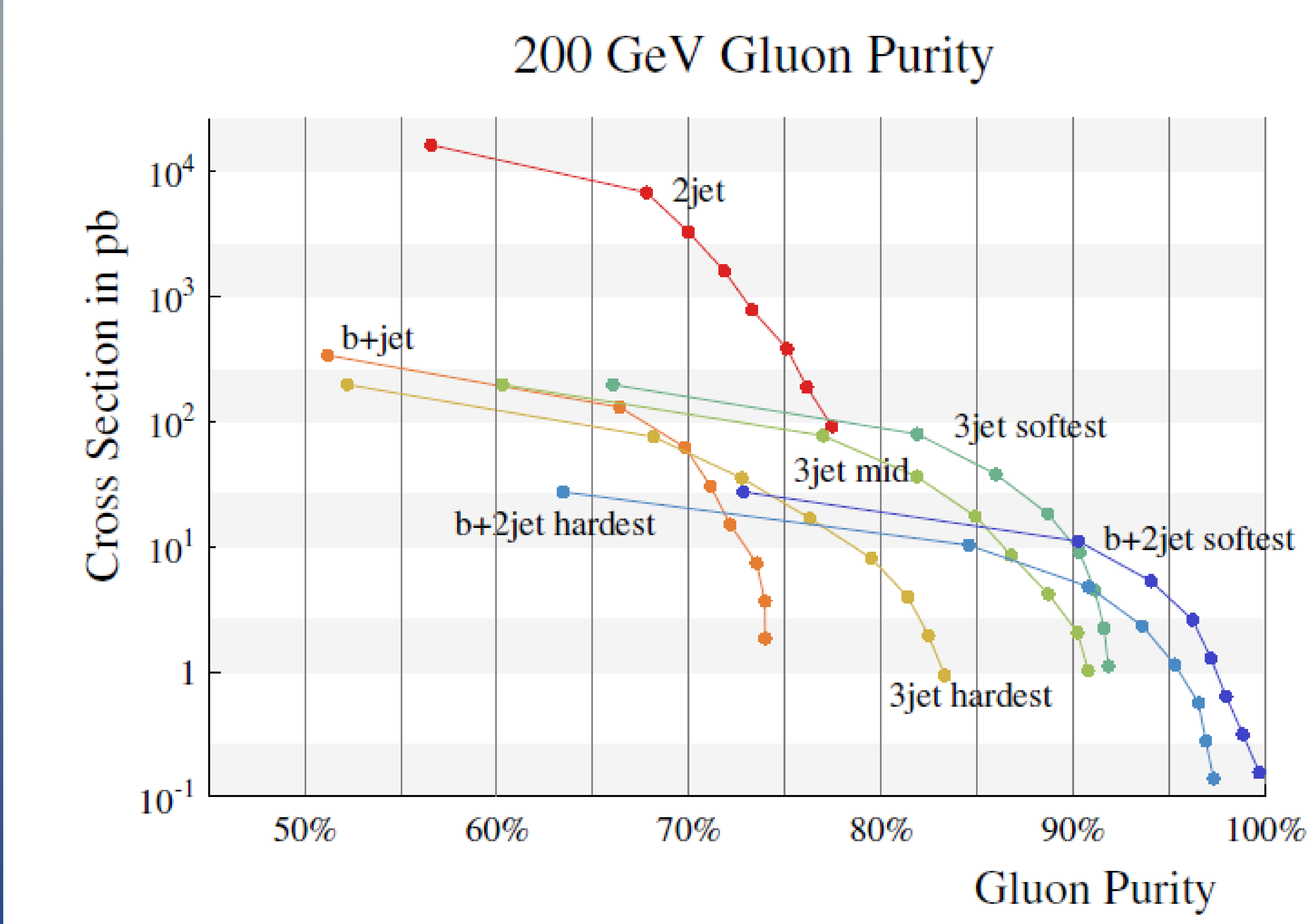
Several times it was proposed to use the fact that gluon jets tend to radiate more than quark jets as a tool to discriminate between them.

At LHC, a simple combination of two variables is not much less powerful than the combination of 10 or more (Gallicchio, Schwartz 1106.3076)
The new thing at the LHC is that it is possible to get pure samples of quark or gluon jets from data, overcoming the main problem that these approaches had in the past

Caveat: there is no obvious rigorous definition of quark or gluon jet! For what follows, people just used the first parton of Pythia event record, waiting for a more robust approach

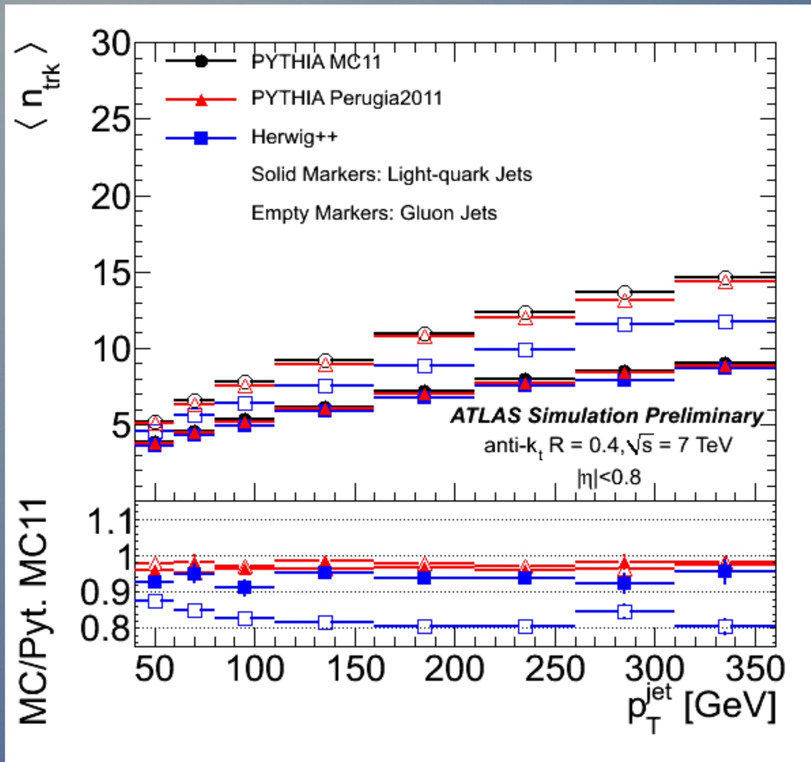


200 GeV Quark Purity
Cross Section in pb
10^3
1
10^{-3}
40%
50%
60%
70%
80%
90%
100%
Quark Purity (zoom)
2jet
3jet hardest
harder
γ+2jet
softer
γ+1jet
W+1jet
W+2jet
harder
softer
Z+2jet
Z+1jet
harder
softer



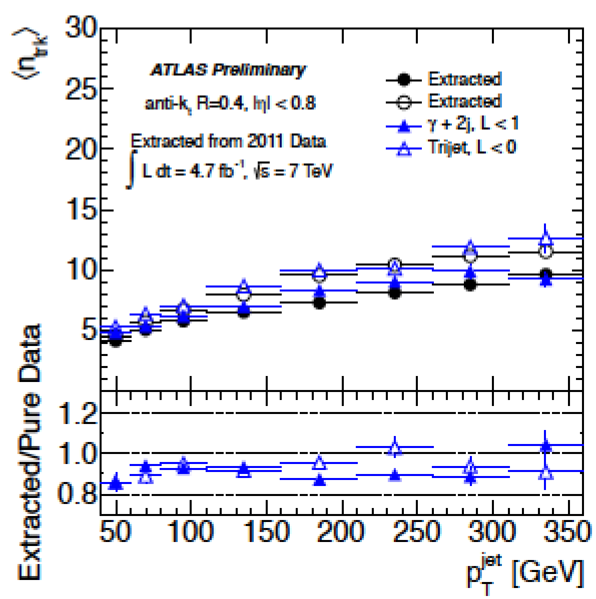
We already have thousands of pure jets to calibrate/cross-check tool

Quark-gluon separation in ATLAS

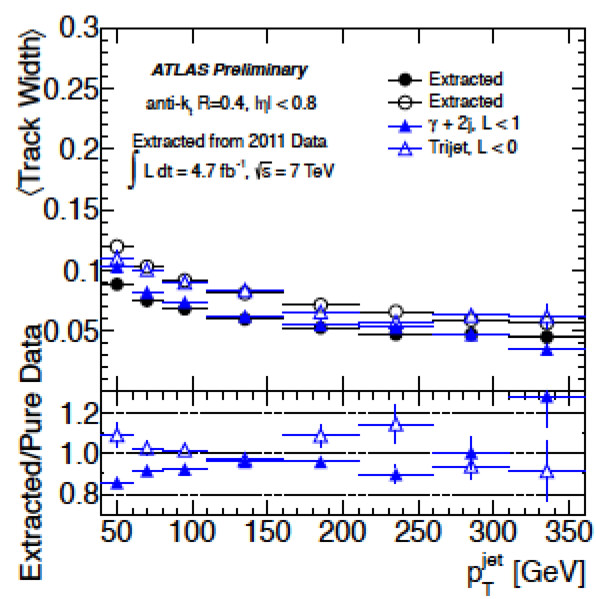


Big model-dependence of basic variables, especially for gluon jets; an approach based on data is mandatory

Properties of quark and gluon jets are extracted from data using QCD and γ+jet samples, weighting them with the expected quark/gluon fractions in Pythia



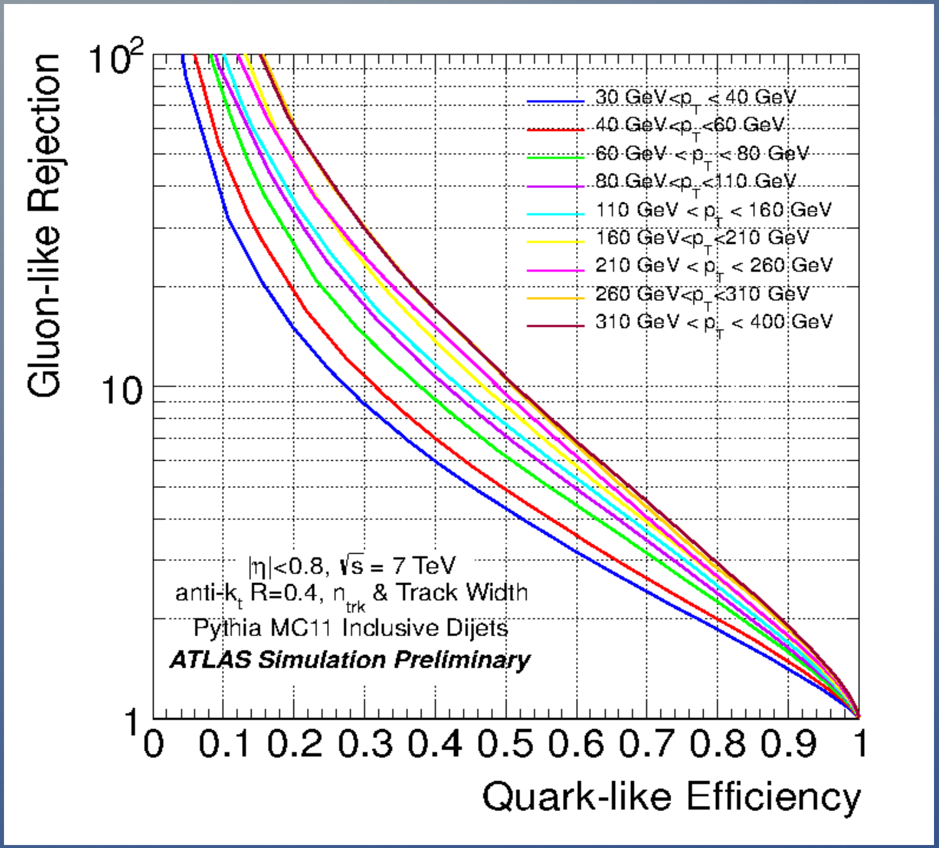
(a) $< n_{trk} >$



(b) $< W >$

After this procedure, a 10% agreement between n. of tracks and jet width is found on the pure samples (there was not enough statistics yet to use the pure samples as templates rather than as a cross-check)

Perspectives and applications of quark-gluon separation



With more collected data, it will be possible to extract templates directly from pure samples

With a proper signal definition, we aim at measuring quark-like and gluon-like inclusive jet and dijet cross-section

Tagger information will be used to improve calibration

Will be used to discriminate Contact Interactions from harder high-x Pdf's

Conclusions

QCD is everywhere at the LHC, and should be studied with the best possible tools. We are taking “traditional” measurements to new quality standards

The field received a big boost a few years ago, when it was realised that jet substructure could be used to highlight hadronic decays of heavy objects.

Since then, several techniques has been developed and proposed, and the field is still booming

First results on data are coming out now, big hopes for a boosted discovery with the 13 TeV data