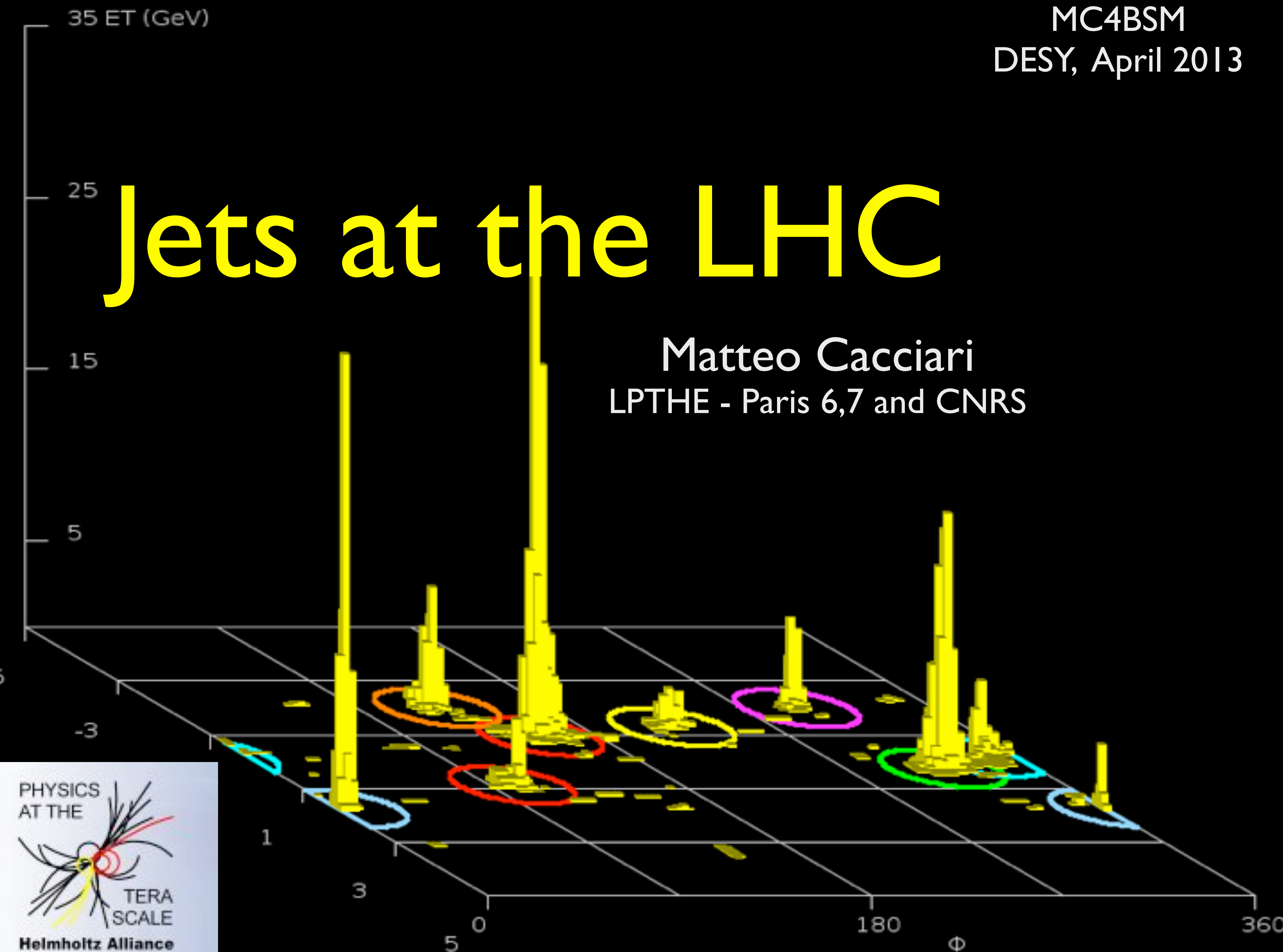
MC4BSM
DESY, April 2013

Jets at the LHC

Matteo Cacciari
LPTHE - Paris 6,7 and CNRS



PHYSICS AT THE TERA SCALE
Helmholtz Alliance

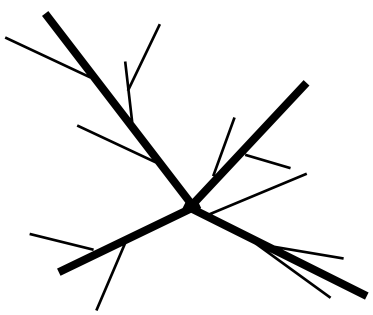
The pervasiveness of jets

- ATLAS and CMS have each published **300+** papers since 2010
  - More than **a third** of these papers make use of **jets**
  - **60%** of the **searches** papers makes use of **jets**

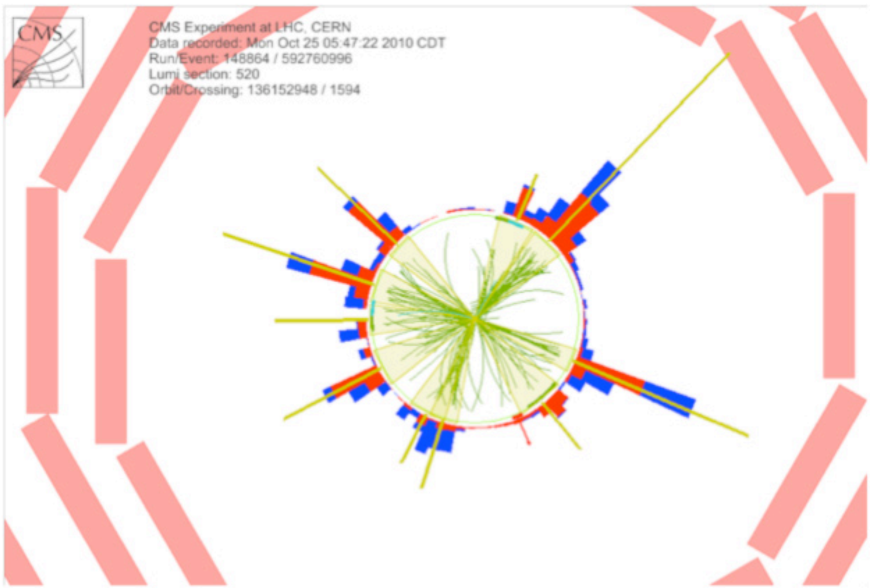
(Source: INSPIRE)

Taming reality

Multileg + PS



??



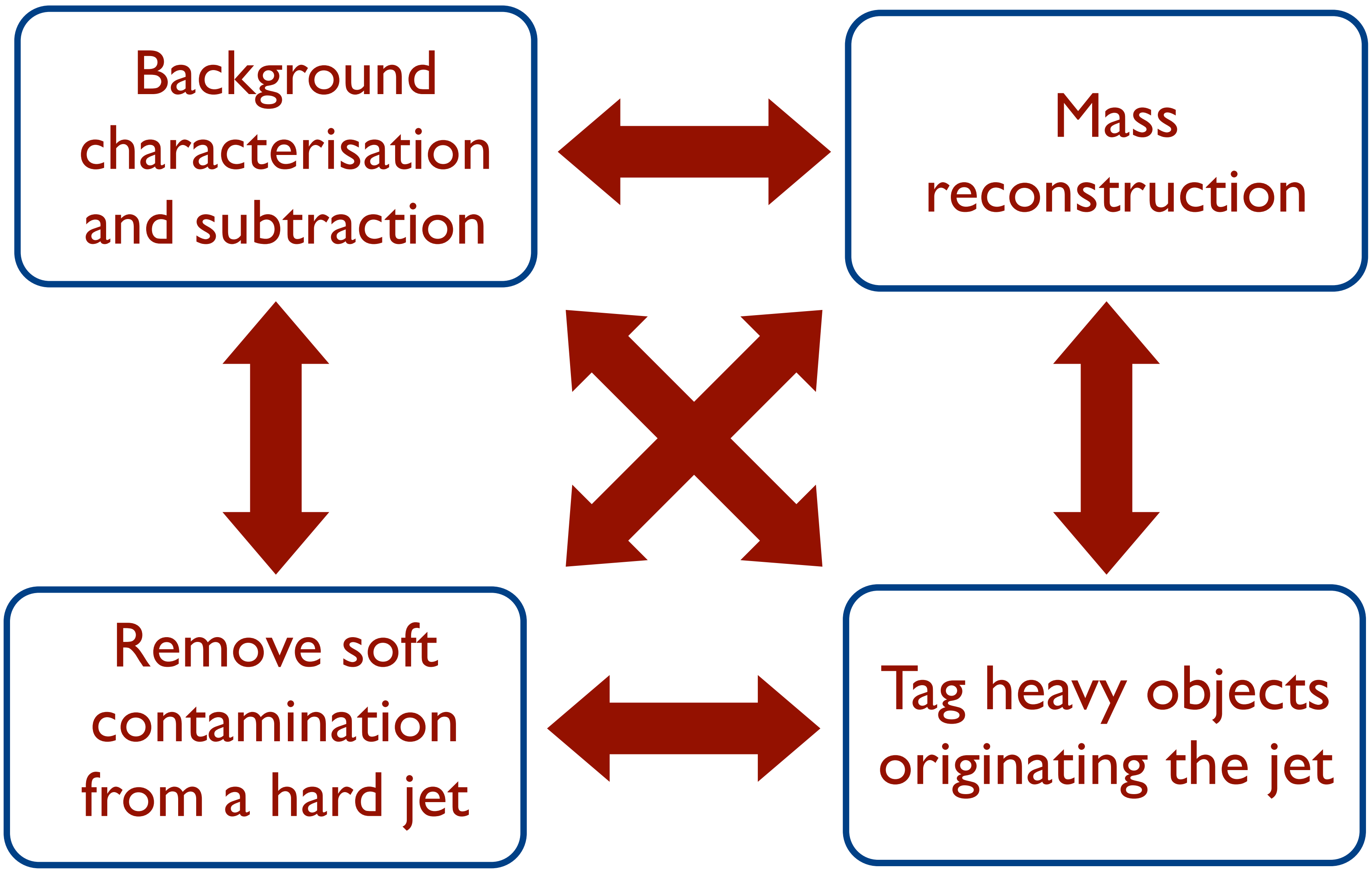
QCD predictions

Real data

Jets

One purpose of a 'jet clustering' algorithm is to **reduce the complexity** of the final state, simplifying many hadrons to **simpler objects** that one can hope to **calculate**

Jets as tools



Evolution of jet algorithms

- Two main classes of jet algorithms: **cones** and **sequential recombination**

- In the past (e.g. Tevatron) cones were mainly used. Many variants, usually not infrared and collinear safe

- The LHC has seen a push towards standardisation, with all experiments using mainly an IRC-safe algorithm, **anti-k_t** (though ATLAS and CMS managed not to pick the same parameter, R=0.4 and 0.6 for ATLAS, 0.5 and 0.7 for CMS) and a common software library

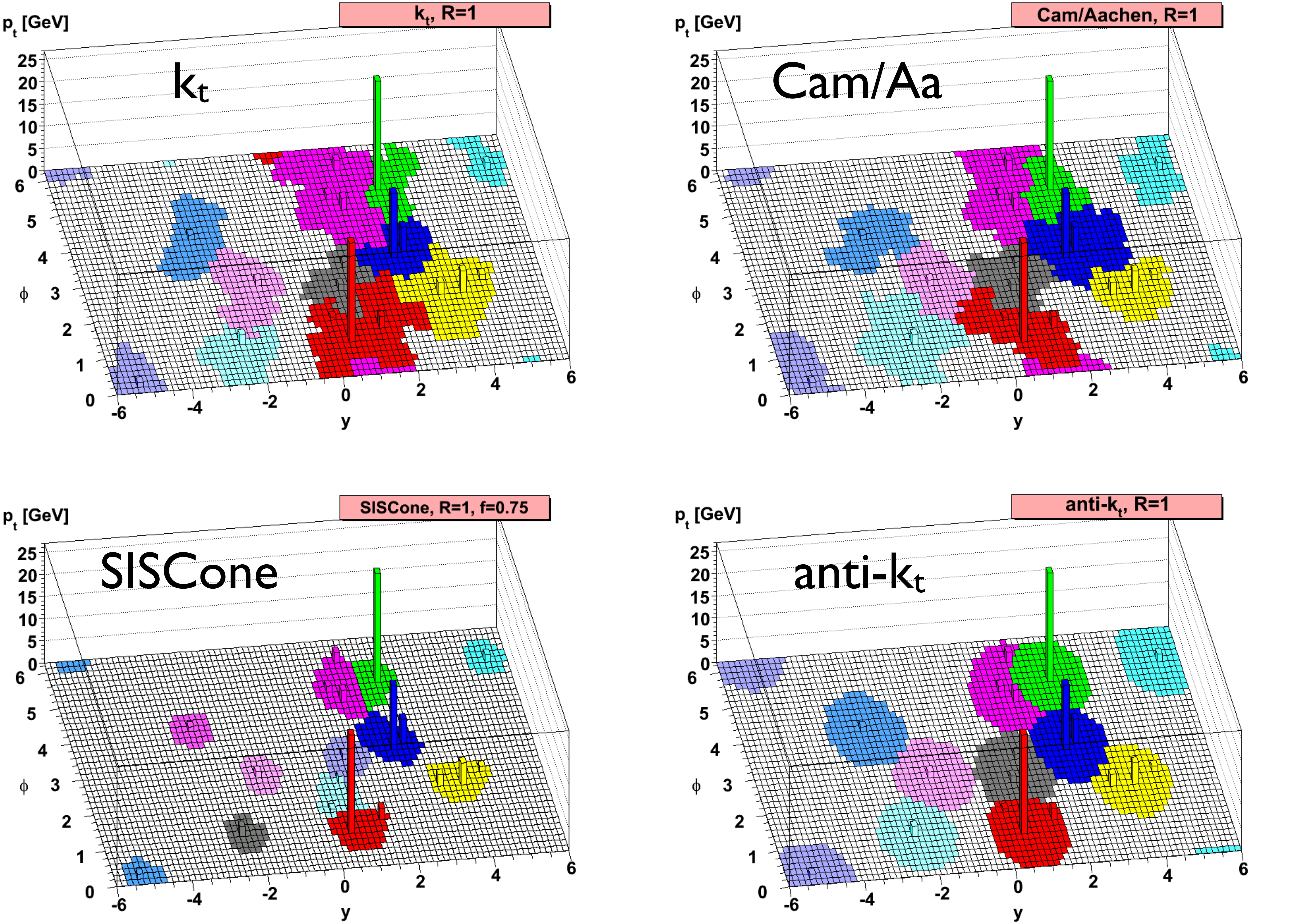
IRC safe algorithms

k_t	SR $d_{ij} = \min(k_{ti}^2, k_{tj}^2)\Delta R_{ij}^2/R^2$ hierarchical in rel p_t	Catani et al '91 Ellis, Soper '93	NlnN
Cambridge/ Aachen	SR $d_{ij} = \Delta R_{ij}^2/R^2$ hierarchical in angle	Dokshitzer et al '97 Wengler, Wobish '98	NlnN
anti-k_t	SR $d_{ij} = \min(k_{ti}^{-2}, k_{tj}^{-2})\Delta R_{ij}^2/R^2$ gives perfectly conical hard jets	MC, Salam, Soyez '08 (Delsart, Loch)	$N^{3/2}$
SISCone	Seedless iterative cone with split-merge gives 'economical' jets	Salam, Soyez '07	$N^2 \ln N$

'second-generation' algorithms

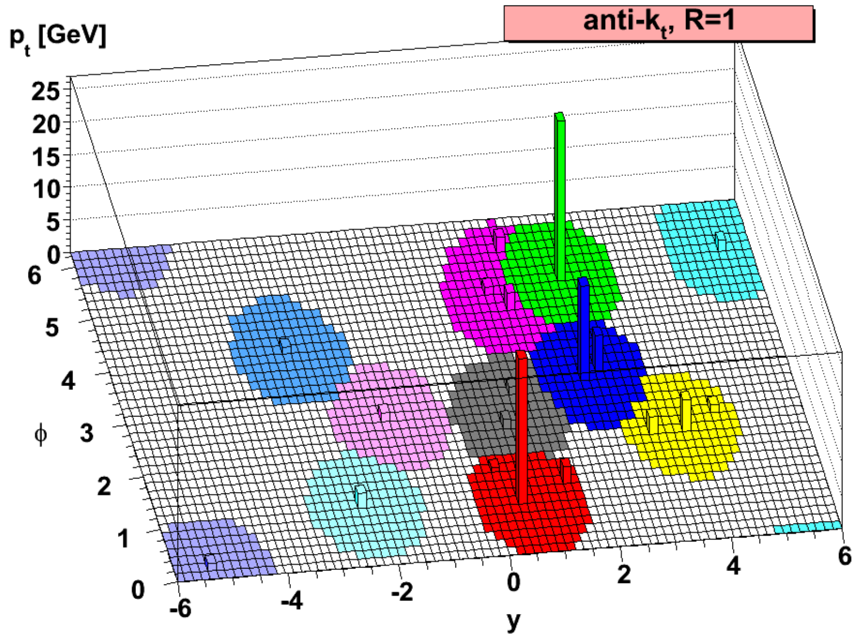
All are available in FastJet, http://fastjet.fr

(As well as many IRC unsafe ones)



k_t, R=1
k_t
Cam/Aachen, R=1
Cam/Aa
SISCone, R=1, f=0.75
SISCone
anti-k_t, R=1
anti-k_t
p_t [GeV]
φ
y

Jet areas



Jet ***areas***, graphically represented by the coloured regions, represent the ***susceptibility*** of each jet to contamination from ***diffuse, soft radiation***

Given an IRC-safe jet algorithm, jet areas can be calculated numerically for each jet, opening the way for a jet-by-jet, rather than average, correction for background contamination

Hard jets and background

Modifications of the hard jet

$$\Delta p_t = \rho A \pm (\sigma\sqrt{A} + \sigma_\rho A + \rho\sqrt{\langle A^2\rangle - \langle A\rangle^2}) + \Delta p_t^{BR}$$

Background momentum density (per unit area)

background — 'susceptibility'

back-reaction — 'resiliency'

Background subtraction

If the background momentum density ρ is known, it can be used to correct the transverse momentum of the hard jets:

$$p_T^{\text{hard jet, corrected}} = p_T^{\text{hard jet, raw}} - \rho \times \text{Area}_{\text{hard jet}}$$

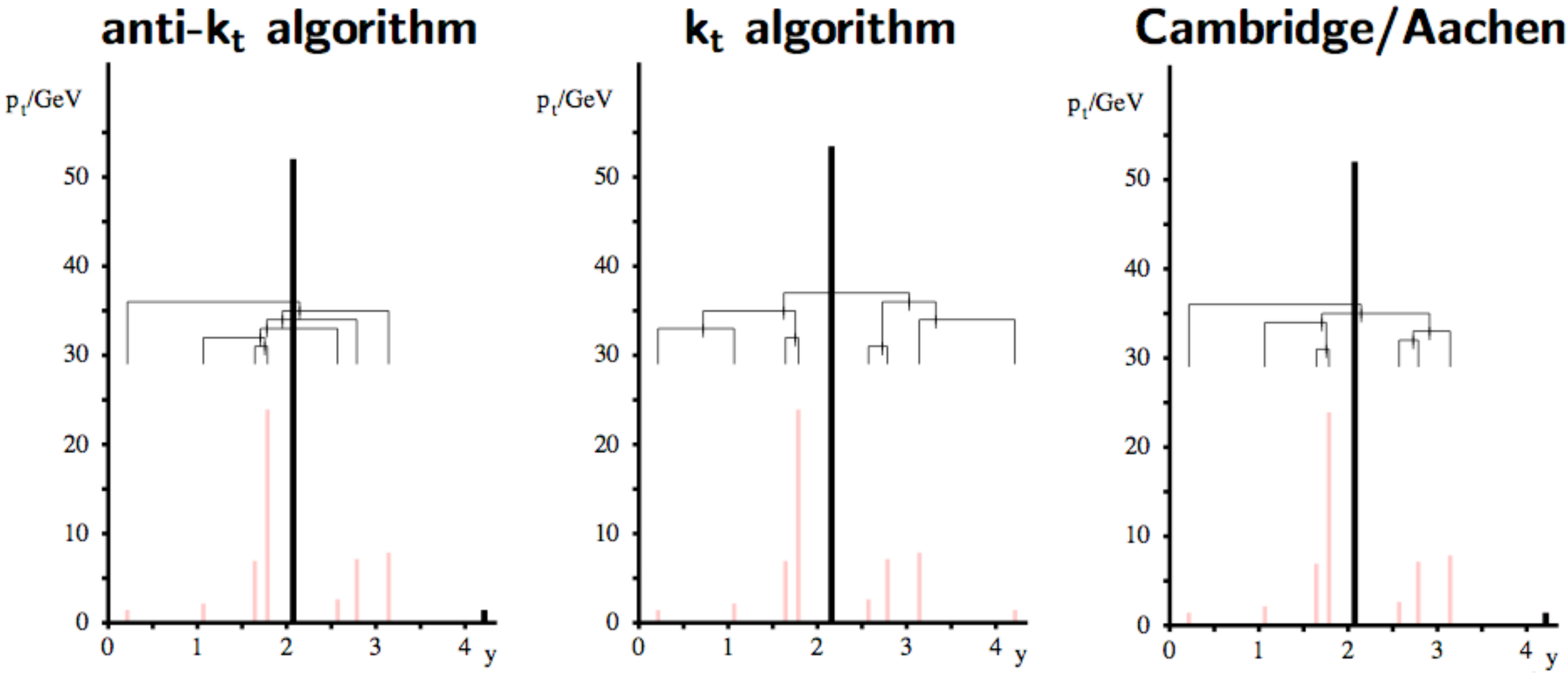
MC, Salam, 0707.1378

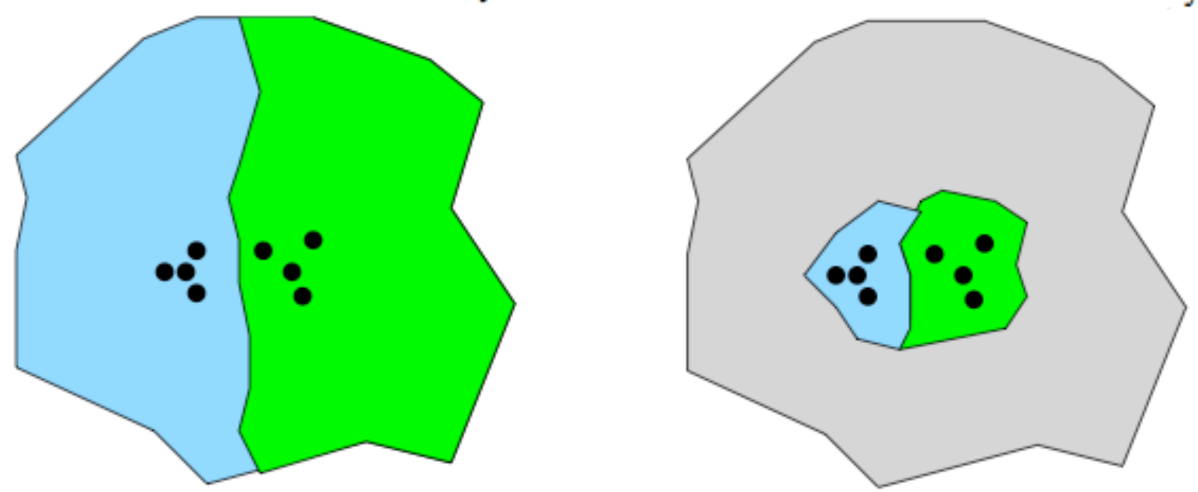
If ρ is measured on an event-by-event basis, and each jet subtracted individually, this procedure will remove many fluctuations and generally improve the resolution of, say, a mass peak

$$\Delta p_t = \rho A \pm (\sigma\sqrt{A} + \sigma_\rho A + \rho\sqrt{\langle A^2\rangle - \langle A\rangle^2}) + \Delta p_t^{BR}$$

Irreducible fluctuations: uncertainty of the subtraction

Hierarchical substructure





Slide by
Gavin Salam

The IRC safe algorithms

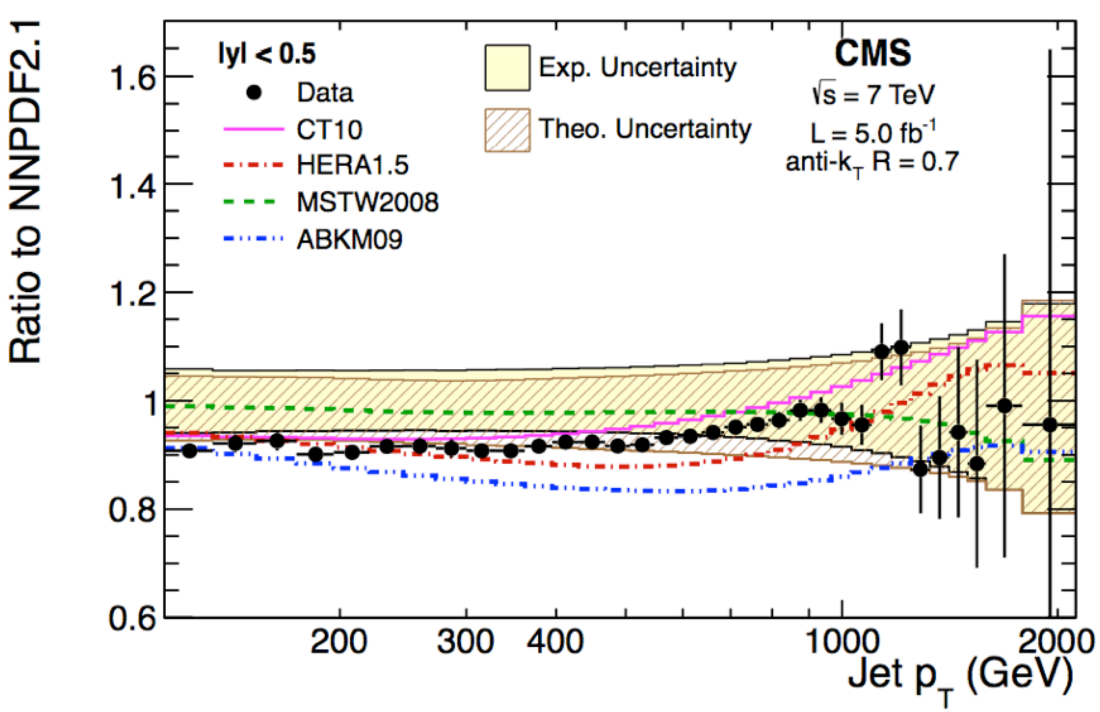
	Speed	Regularity	UE contamination	Backreaction	Hierarchical substructure
k_t	☺ ☺ ☺	☂	☂☂	☁☁	☺ ☺
Cambridge /Aachen	☺ ☺ ☺	☂	☂	☁☁	☺ ☺ ☺
anti-k_t	☺ ☺ ☺	☺ ☺	☁/ ☺	☺ ☺	✘
SISCone	☺	☁	☺ ☺	☁	✘

Recent activity on jets

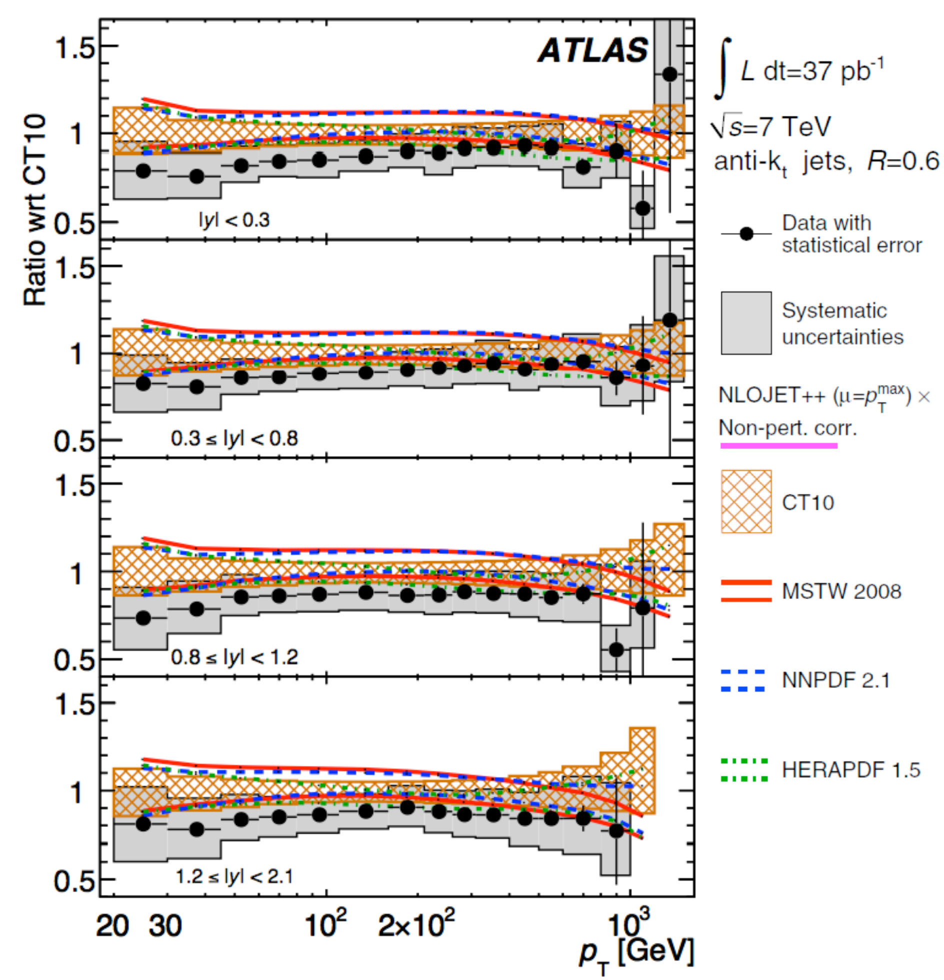
- LHC measurements
- Higher order calculations/matching techniques
- Jet substructure tools and calculations

Inclusive jet cross section

arXiv:1212.6660



arXiv:1112.6297

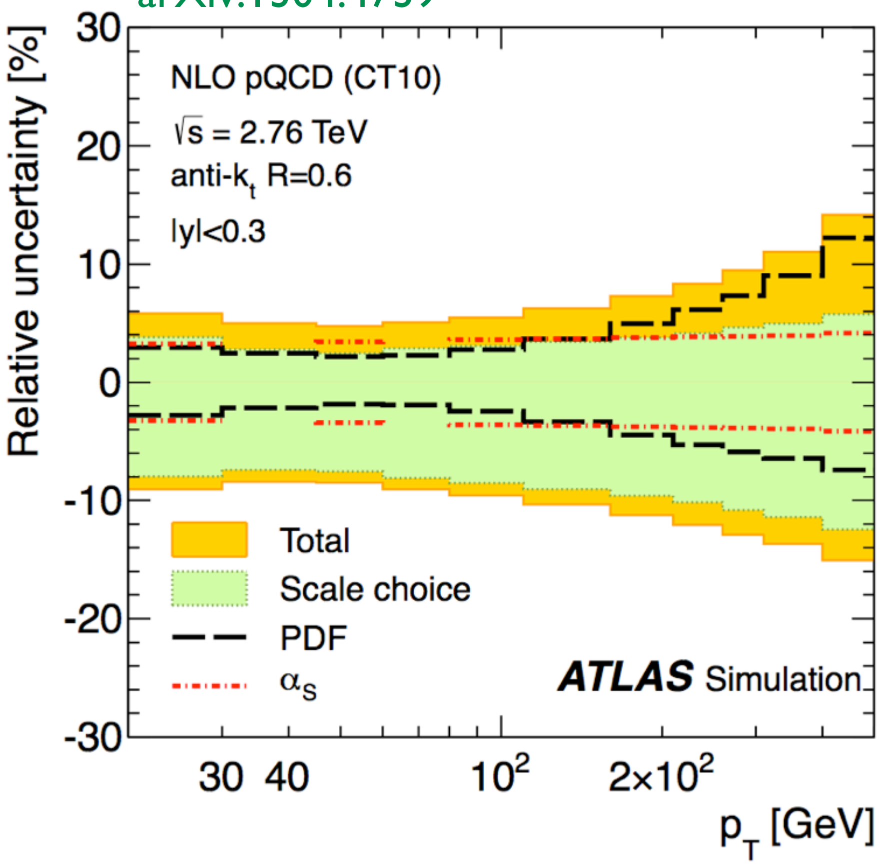


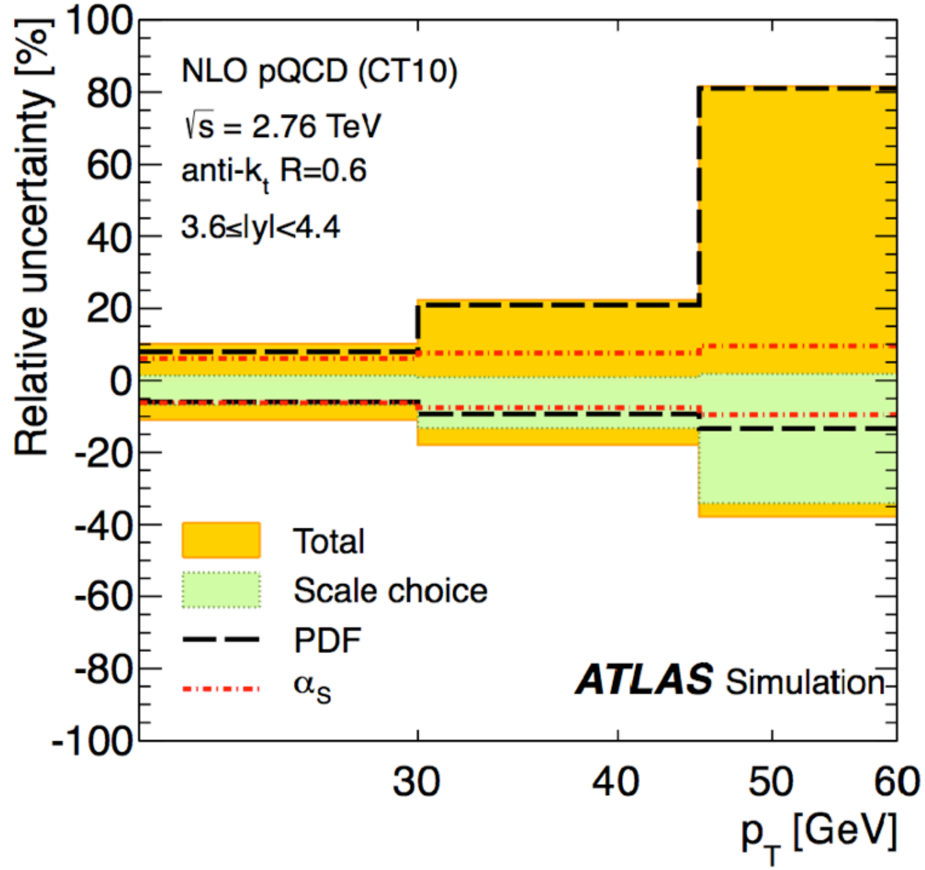
Broad agreement within uncertainties

(<10% for theory, similar and decreasing fast for experiments)

(Theoretical) uncertainties

arXiv:1304.4739



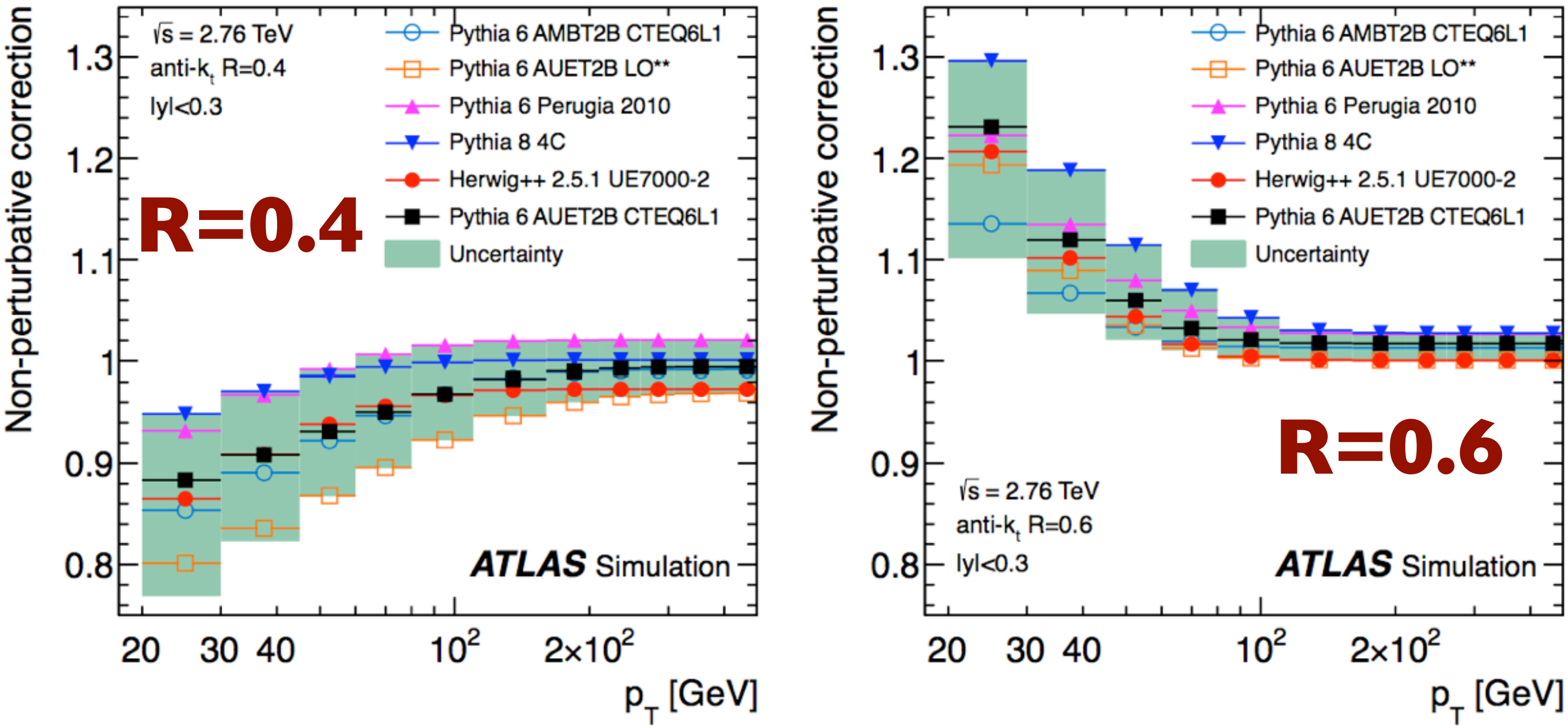


- Uncertainties dominated by unknown higher orders in the central region, and by PDFs uncertainties in the forward and high-p_T regions

Non-perturbative corrections

Non-perturbative contributions from ***Underlying Event*** and ***Hadronisation*** can be estimated by running MCs with and without them

arXiv:1304.4739



Note the ***very different situation*** at R=0.4 and 0.6, due to ***cancellations*** (or lack thereof) between UE and hadronisation corrections

R-dependent effects

Perturbative radiation: $$\Delta p_t \simeq \frac{\alpha_s(C_F, C_A)}{\pi} p_t \ln R$$

Hadronisation: $$\Delta p_t \simeq -\frac{(C_F, C_A)}{R} \times 0.4 \text{ GeV}$$

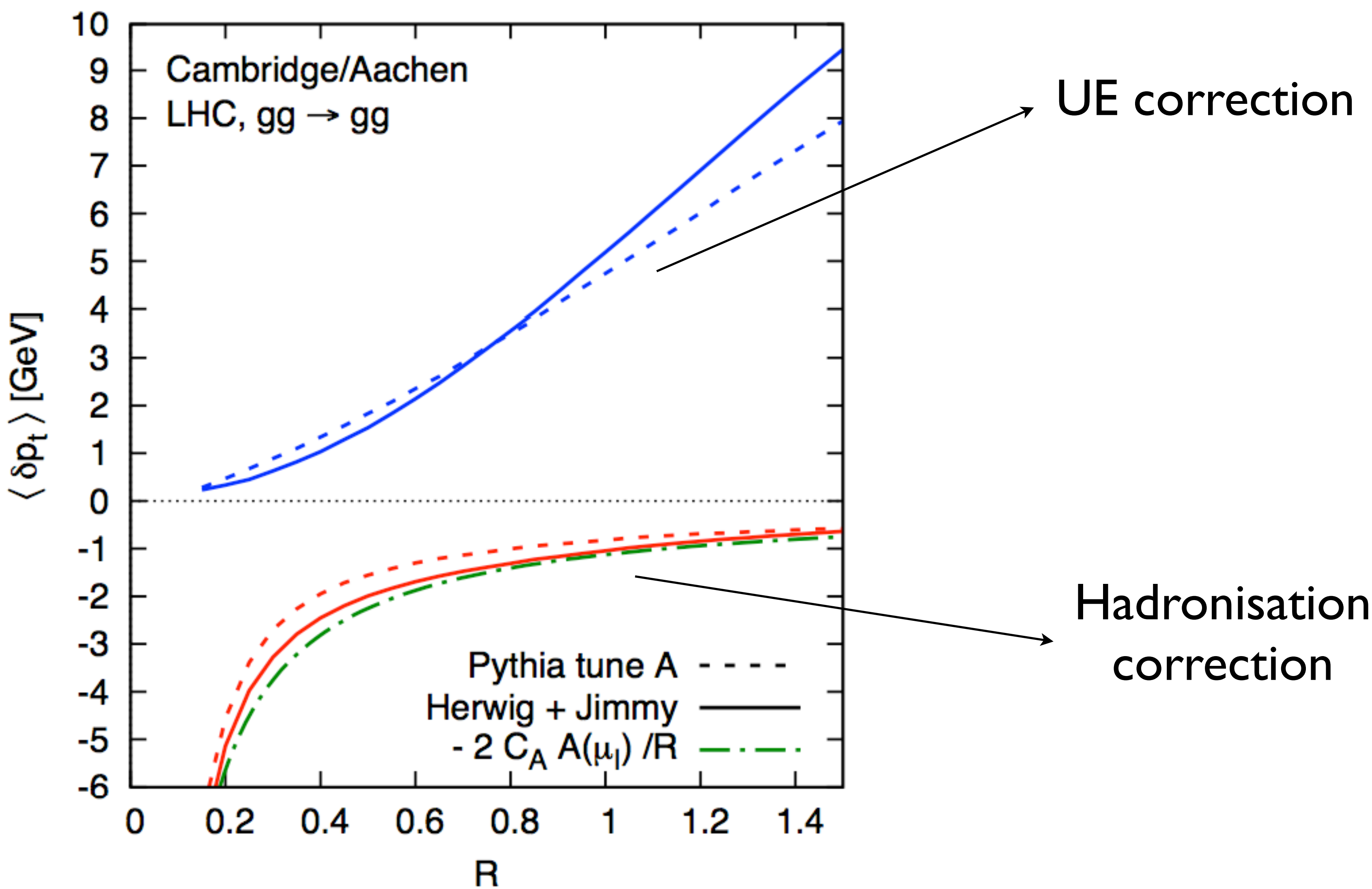
Underlying Event: $$\Delta p_t \simeq \frac{R^2}{2} \times (\underset{\text{Tevatron}}{2.5} \text{ — } \underset{\text{LHC}}{15} \text{ GeV})$$

Analytical estimates:
Dasgupta, Magnea, Salam, arXiv:0712.3014
G. Soyez, arXiv:1006.3634

Non-perturbative corrections

Jet radius dependence

Dasgupta, Magnea, Salam, 0712.3014

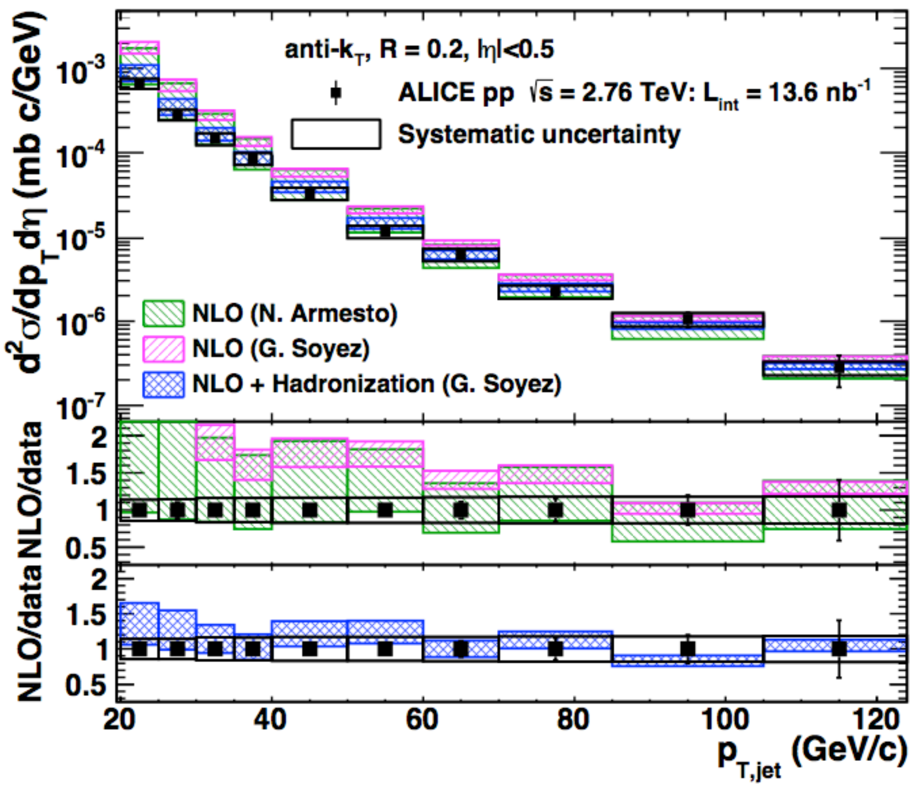


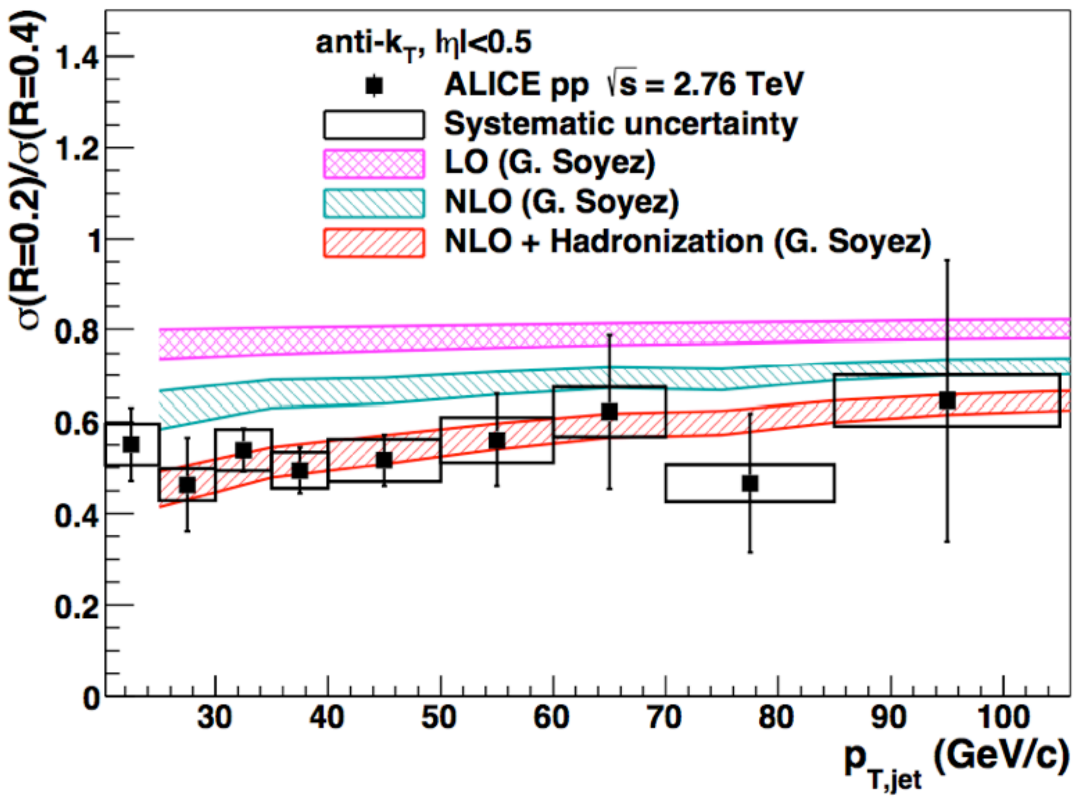
Inclusive jet cross secion

ALICE @ 2.76 TeV

(reference for PbPb data)

arXiv:1301.3475



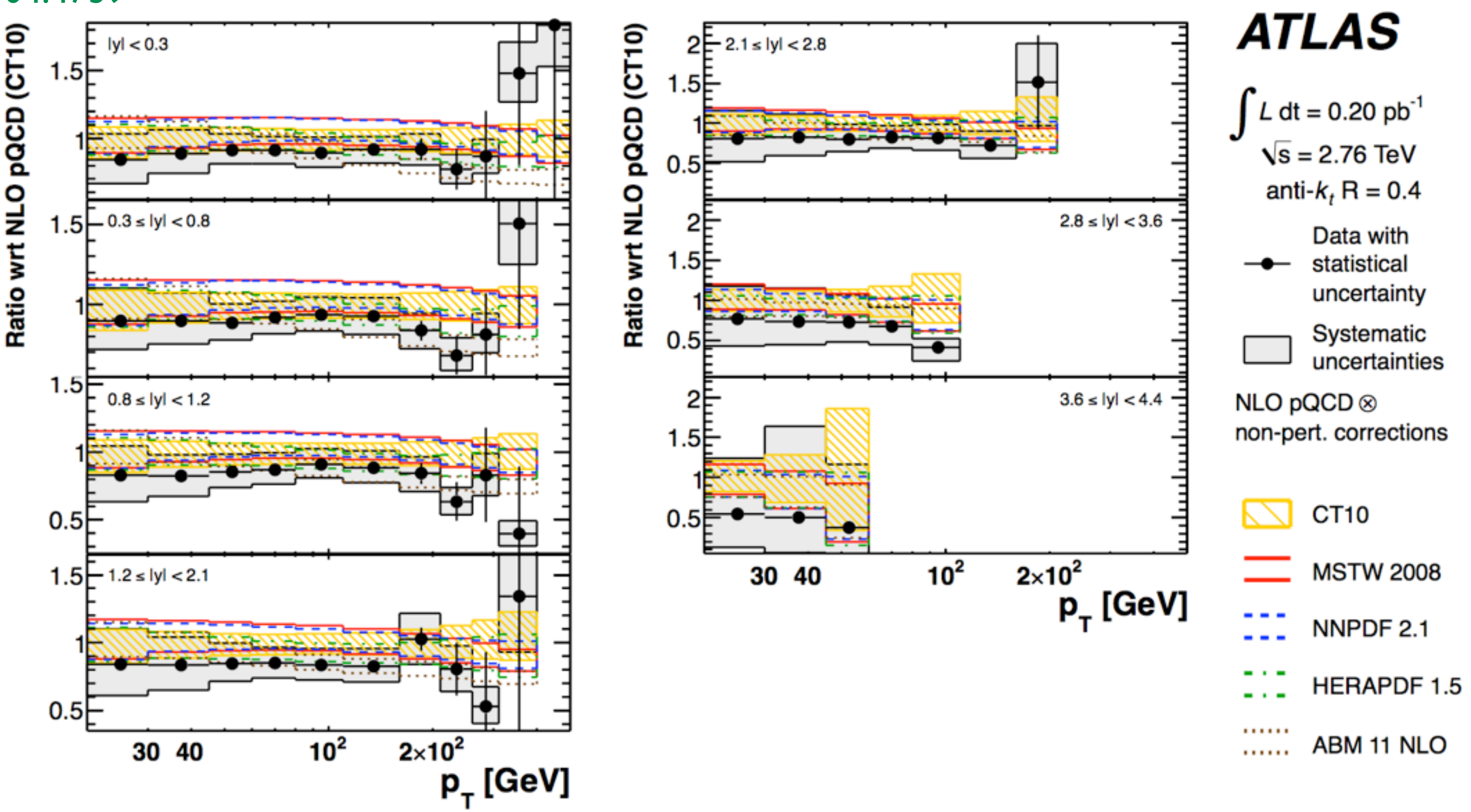


Ratios of two significantly different radii emphasizes hadronisation corrections

Inclusive jet cross sections

2.76TeV/7 TeV ratios from ATLAS

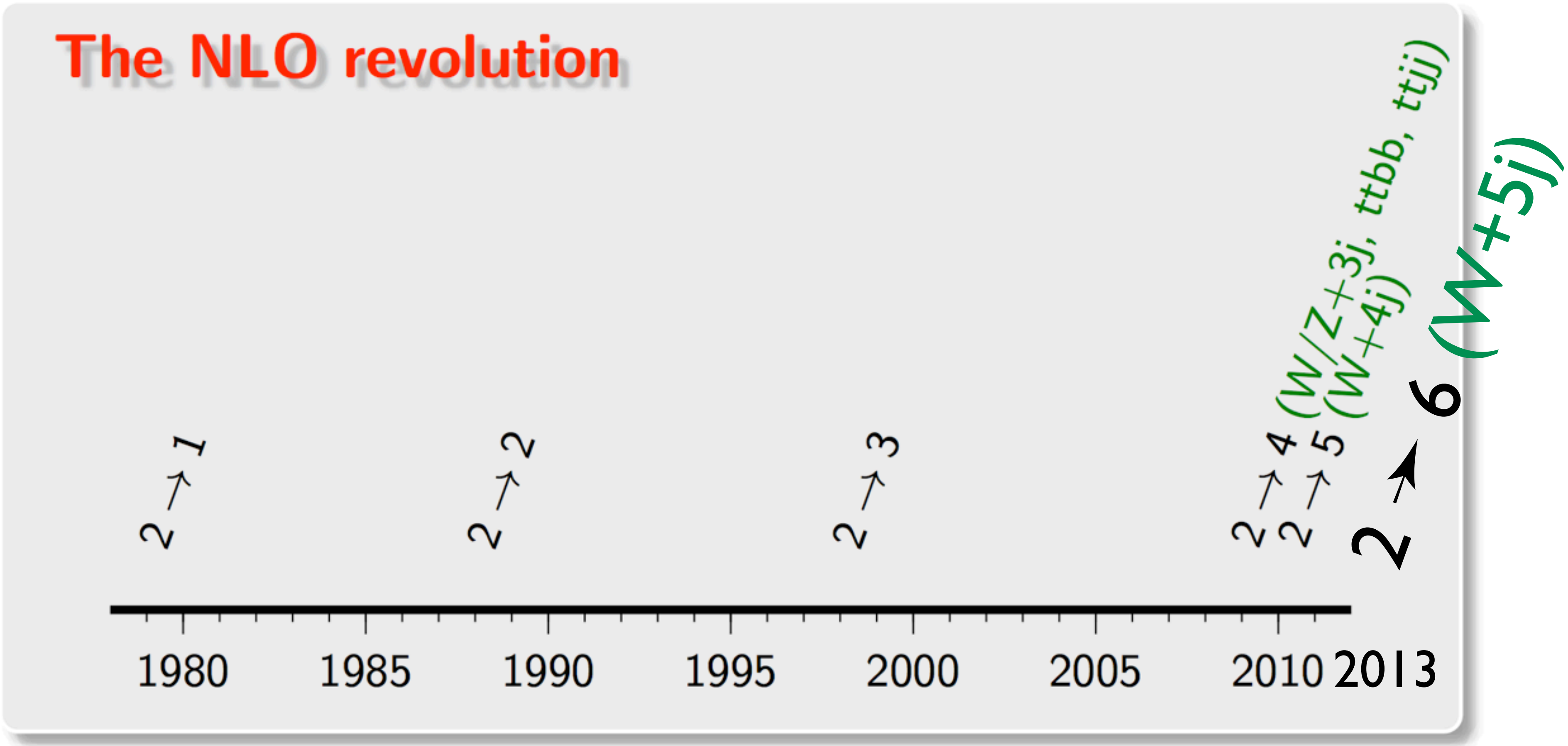
arXiv:1304.4739



- Smaller uncertainties due to cancellations
- Enhanced sensitivity to PDF set
- Data have been used to complement a PDF fit from HERA data

The revolution lives on

G. Salam, ICHEP 2010

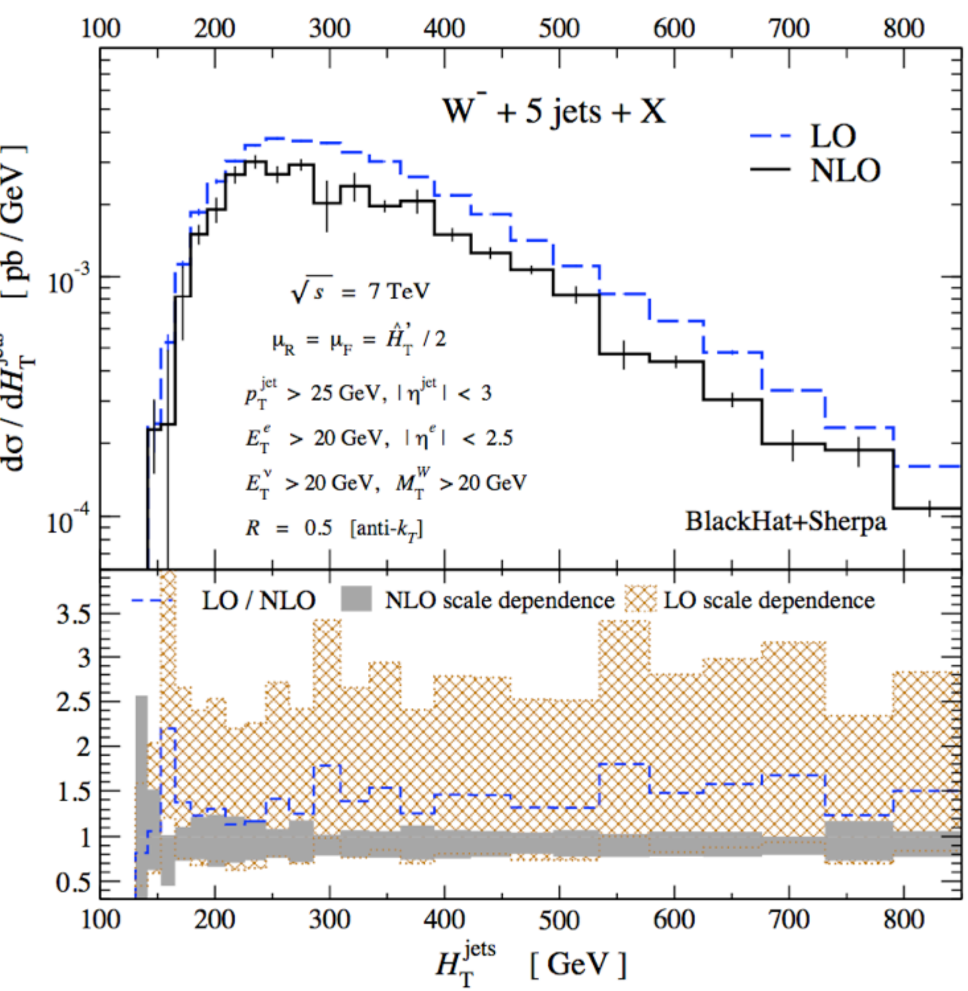


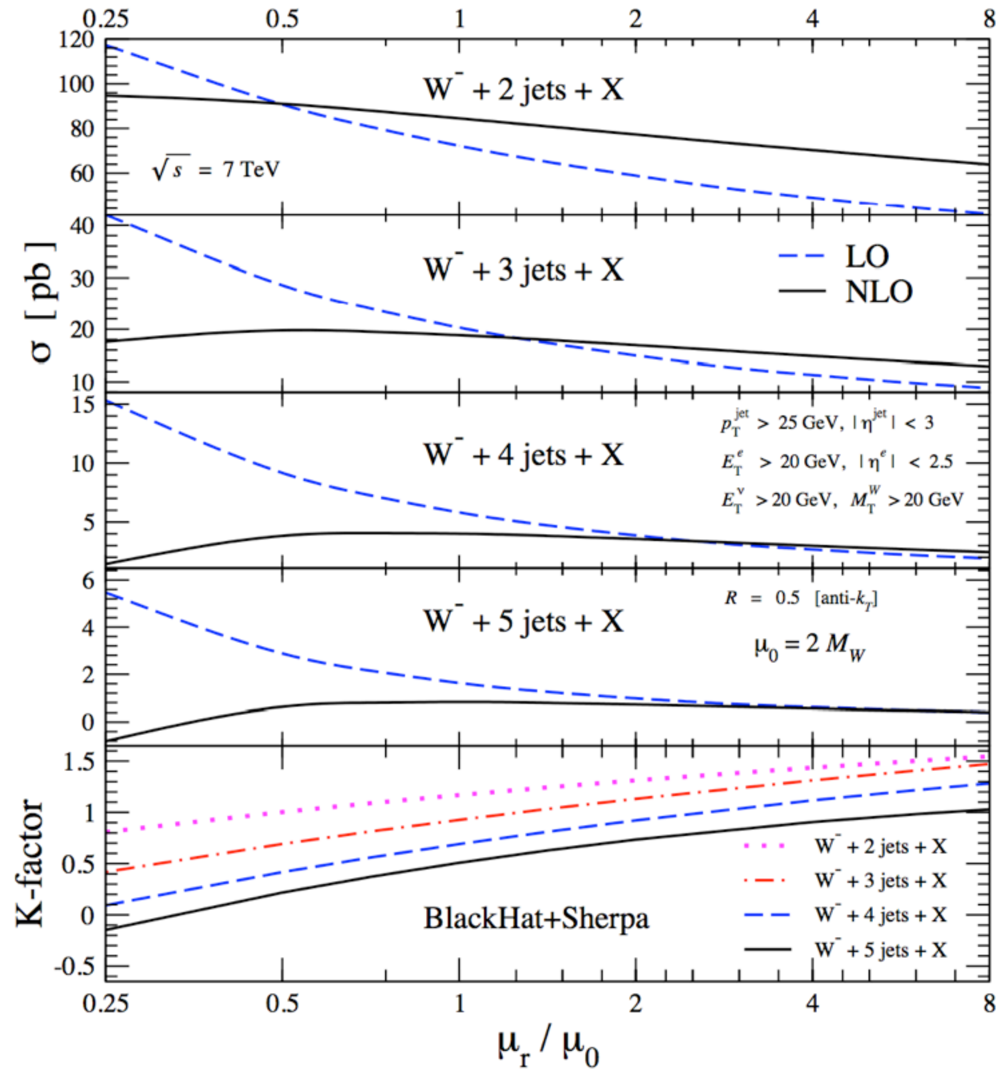
Bern, Dixon, Fabres Cordero, Hoeche, Ita, Kosower, Maitre, Ozeren, 1304.1253
(BLACKHAT+SHERPA)

To this revolution, one should of course add the automated-NLO/MCs one

W + 5 (and even 6) jets

arXiv:1304.1253





- Reduced scale dependence
- NB: NLO corrections are not always positive.....

- Extrapolation to W+6j:

$$W^- + 6 \text{ jets}: \quad 0.15 \pm 0.01 \text{ pb},$$
$$W^+ + 6 \text{ jets}: \quad 0.30 \pm 0.03 \text{ pb},$$

Back to the inclusive jet and dijet cross sections

The state of the art until a few months ago was:

- NLO corrections [Ellis, Kunzst, Soper '92, Giele, Glover, Kosower '94, Nagy '02]
- NLO + parton shower [Alioli, Hamilton, Nason, Oleari, Re '11]

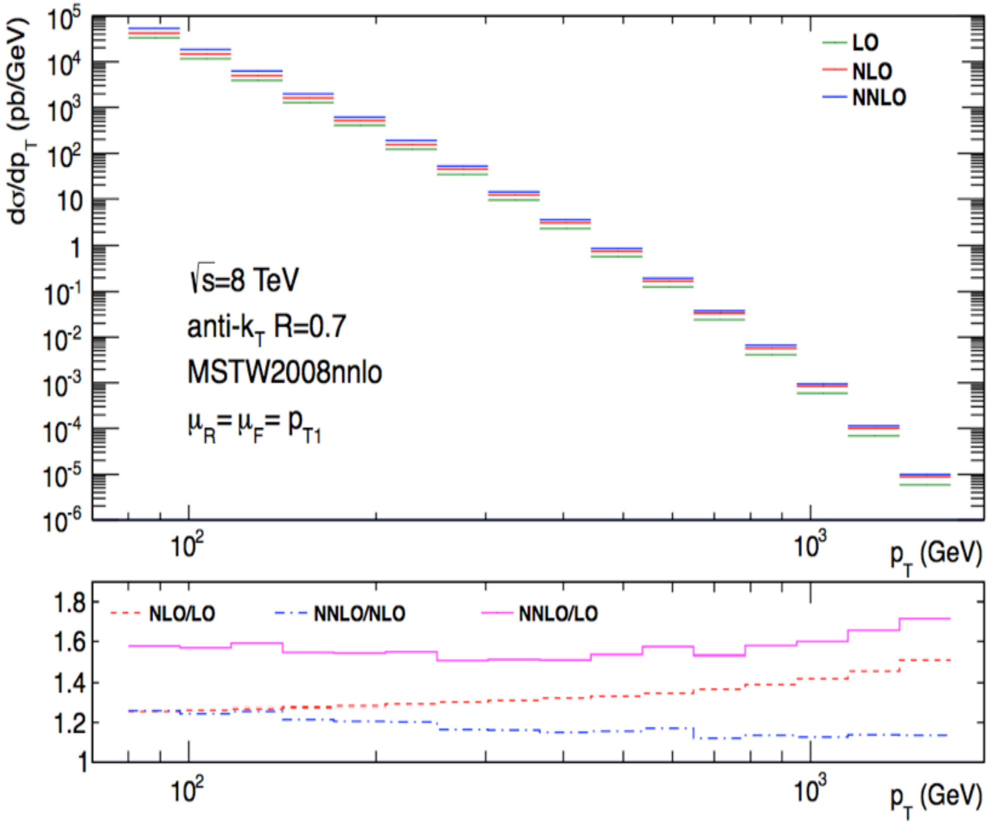
The experimental uncertainty already at the level of the theoretical one (and decreasing fast) calls for an NNLO calculation

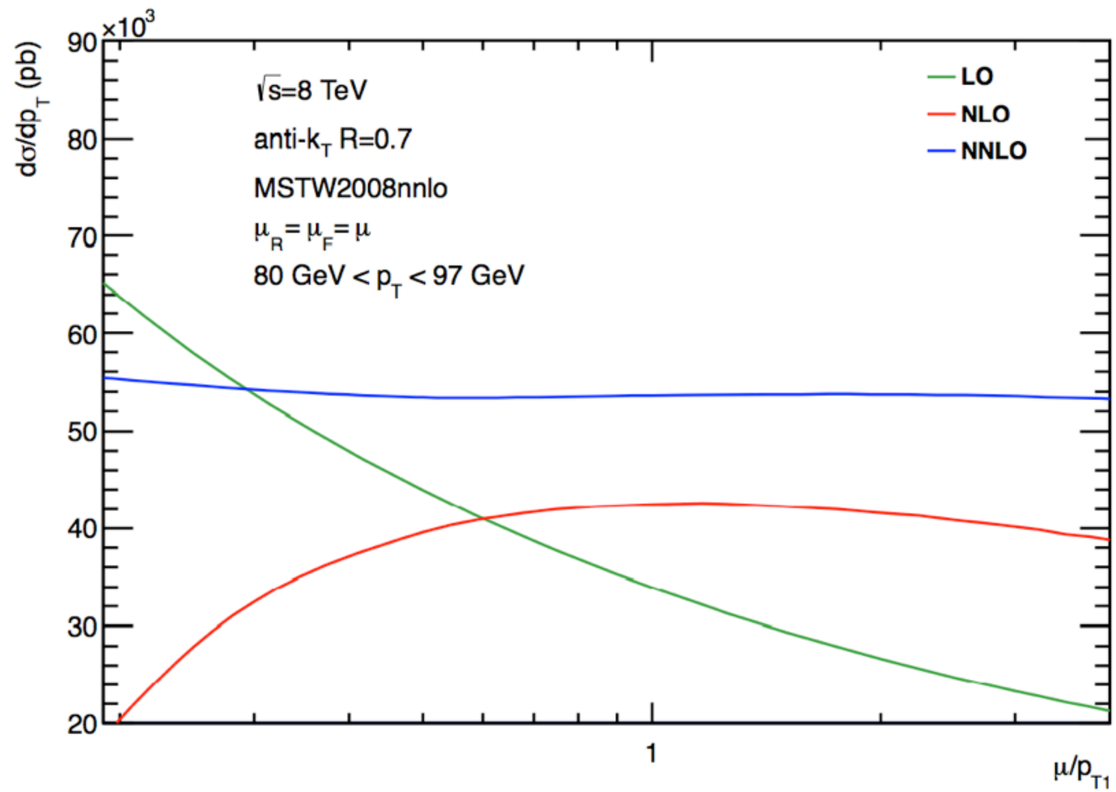
Very recent result: $pp \to 2j$ to NNLO)

(gg channel only, and in the leading colour approximation)

[Gehrmann-De Ridder, Gehrmann, Glover, Pires, 1301.7310]

2 jets to NNLO

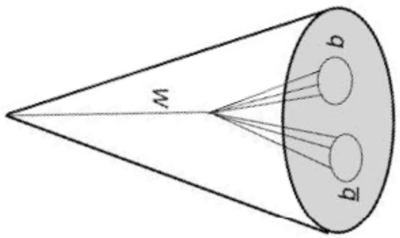


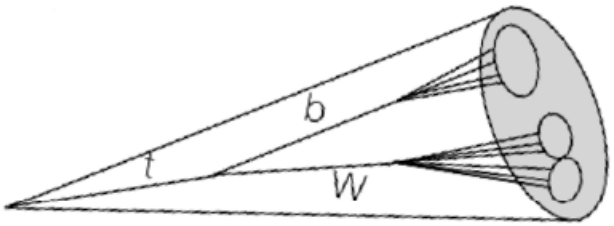


- NNLO/NLO K-factor ~ 1.2, quite independent of p_T value

- Very small residual scale dependence (<1%)
- No obvious sign of convergence of the series at small (~ 100 GeV) p_T

Original motivation: tag a boosted massive particle whose decay products end up in a single 'fattish' jet

W 

top 

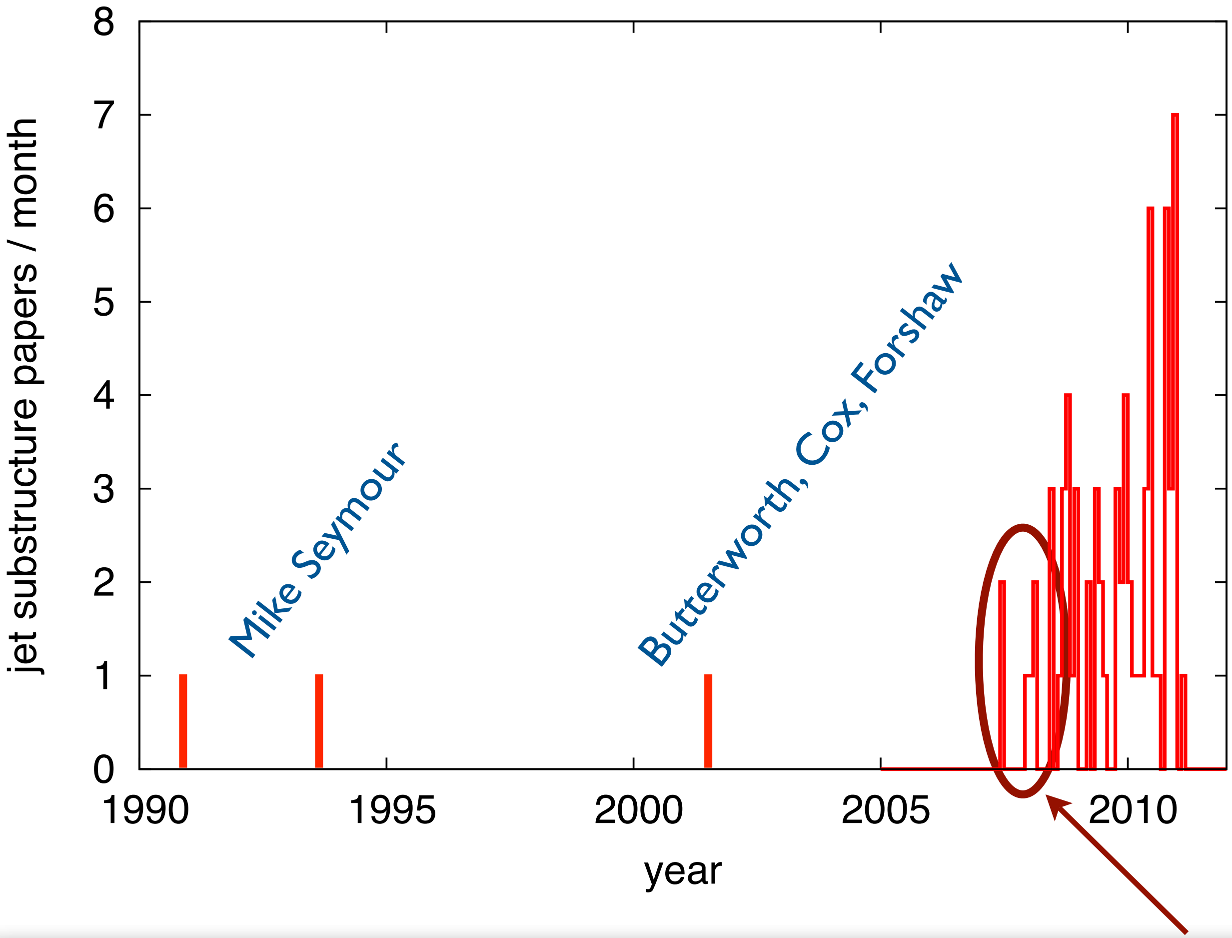
Cone aperture: $R \sim 2m/p_T$

- Electroweak-scale particles (m ~ 100 GeV) boosted at a few hundreds GeV (e.g. coming from the decay of a TeV-scale BSM particle) mean **R ~ 1**
  - Too large for a single 'standard' jet with R=0.4-0.7 to catch all decay products
  - Too small to get separate jets for the decay products

Need a completely new strategy

‘Jet substructure’ papers in SPIRES

Number of papers containing the words ‘jet substructure’ and ‘LHC’



More than 70 papers after 2008.
(the histogram dates from early 2011)

Pioneered by M. Seymour in the early ‘90s, rebooted by BDRS paper

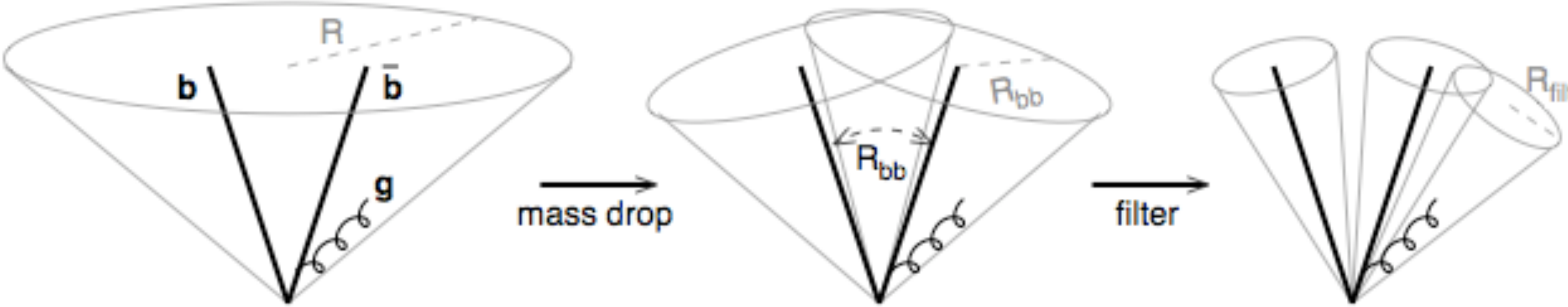
15. **Jet substructure as a new Higgs search channel at the LHC.**
Jonathan M. Butterworth, Adam R. Davison (University Coll. London), Mathieu Rubin, Gavin P. Salam (Paris, LPTHE).
Published in **Phys.Rev.Lett. 100 (2008) 242001**
e-Print: **arXiv:0802.2470 [hep-ph]**

Jet substructure

- The substructure of a jet (i.e. the ability to further resolve smaller components) can be exploited to
  - *tag a particular structure inside the jet, i.e. a massive particle*
    - *Examples: Higgs (2-prongs decay), top (3-prongs decay)*
  - *remove background contamination from the jet or its components*
    - *Examples: filtering, trimming, pruning*

In the following I'll be mainly illustrating the BDRS tagger/filter as a pedagogical example, and also list other approaches

The BDRS tagger



It's a two-prongs tagger for boosted Higgs, which

- Uses the **Cambridge/Aachen** algorithm (see why in the next slide)
- Employs a **Mass-Drop** condition (as well as an asymmetry cut) to find the relevant splitting (i.e. 'tag' the heavy particle)
- Includes a post-processing step, using '**filtering**' (introduced in the same paper) to clean as much as possible the resulting jets of UE contamination

Jet substructure as filter

The ***jet substructure*** can be exploited to help **removing contamination** from a soft background

- Jet 'filtering' — Butterworth, Davison, Rubin, Salam, 2008

  Break jet into subjets at distance scale R_{filt}, retain n_{filt} hardest subjets

- Jet 'trimming' — Krohn, Thaler, Wang, 2009

  Break jet into subjets at distance scale R_{trim}, retain subjets with $p_{t,subjet} > \varepsilon_{trim}\, p_{t,jet}$

- Jet 'pruning' — S. Ellis, Vermilion, Walsh, 2009

  While building up the jet, discard softer subjets when $\Delta R > R_{prune}$ and $\min(p_{t1},p_{t2}) < \varepsilon_{prune}\,(p_{t1}+p_{t2})$

Aim: limit sensitivity to background while retaining bulk of perturbative radiation

Filtering, trimming and pruning can appear quite similar.
These and similar tools are collectively called **groomers**

A generic substructure approach to tagging will

- Cluster initially with a large R, so as to collect all the decay products of a boosted heavy particle into a single jet
- Decluster this jet into subjets, using some conditions to decide when to stop the declustering (i.e. find the ‘relevant splitting’), possibly including kinematical cuts to reduce the QCD background.
  - *The stopping criterion automatically finds the ‘right size’ for the distance between the two prongs of the heavy particle decay*
    - Alternatively to declustering, one can employ one of the jet-shapes based tagging methods, i.e. N-subjettiness ratios
- Optionally add a final ‘cleaning’ procedure to remove as much as possible spurious soft/background radiation

Generic tagging/grooming

Fat-jet finding

Often anti-kt, $R \approx 1$

↓ large p_t, large mass fat-jet, signal or background

Tagging step

↓ signal jet candidate, still background-contaminated

grooming step

↓

final candidate, potentially with little background contamination

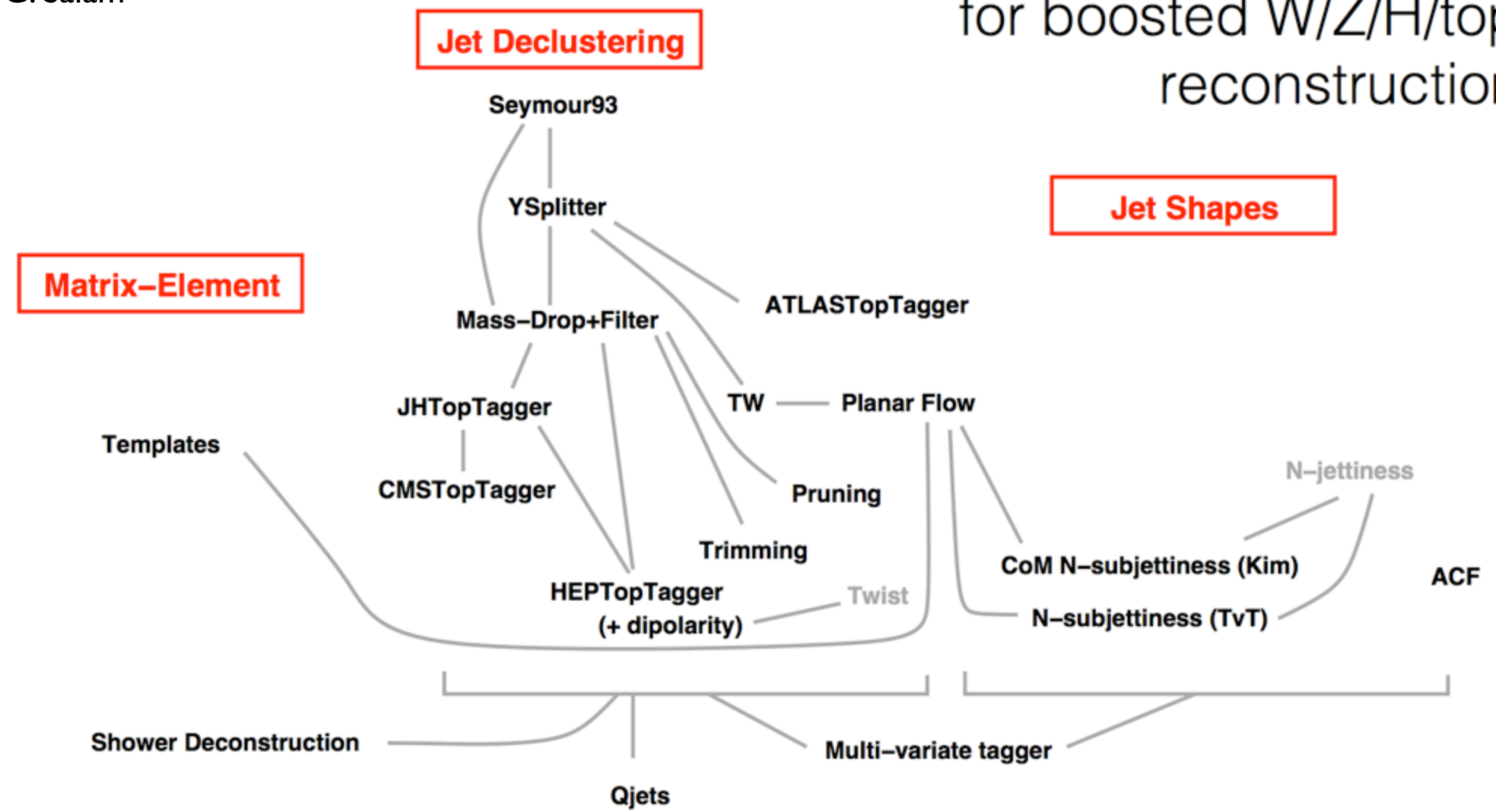
Note that in some taggers (i.e. pruning) the tagging and grooming steps are not explicitly factorised

Also, some tools may actually not follow rigidly this scheme

The jet substructure maze

Slide by
G. Salam

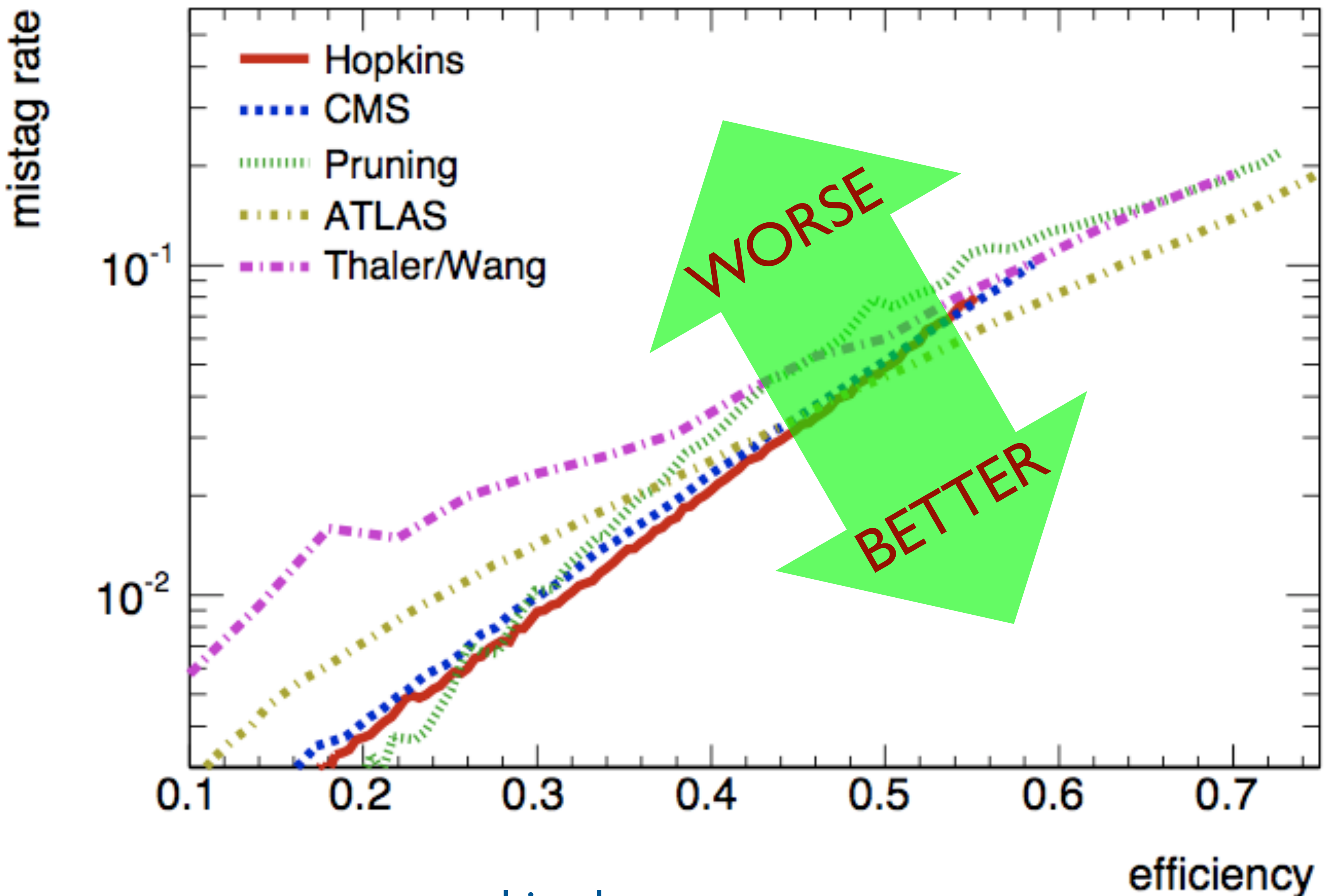
Some of the tools developed for boosted W/Z/H/top reconstruction



Apologies for missing or misplaced items or links

Comparison of top taggers

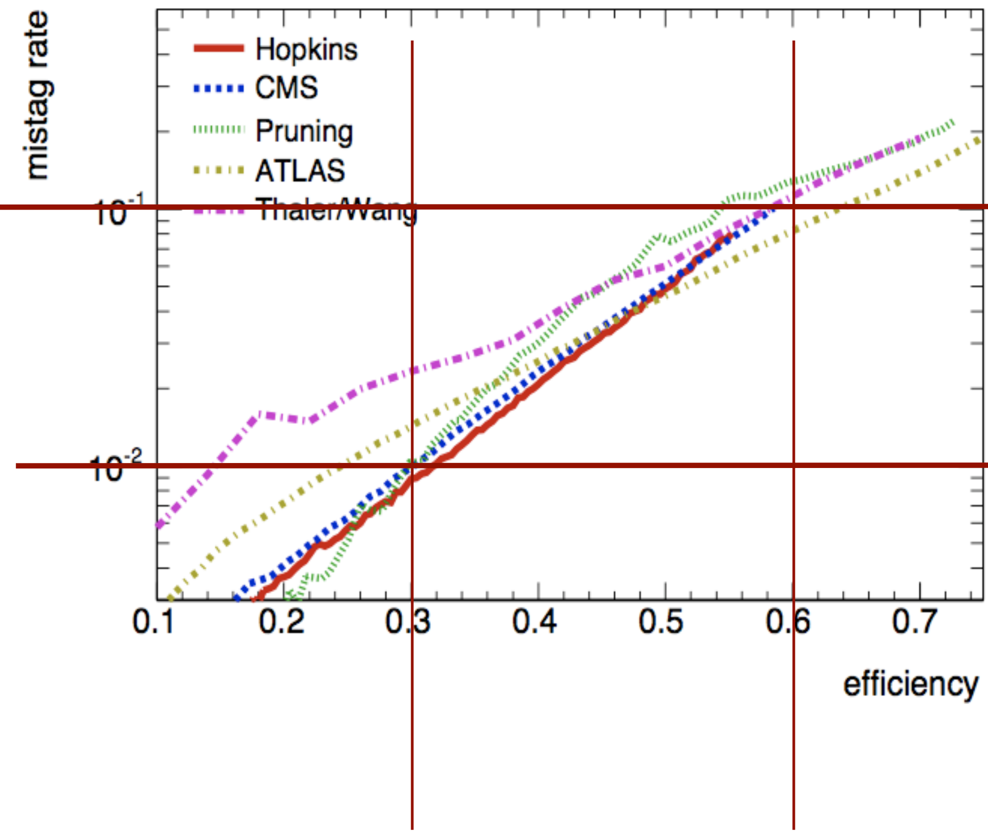
Boost 2010 proceedings, arXiv:1012.5412



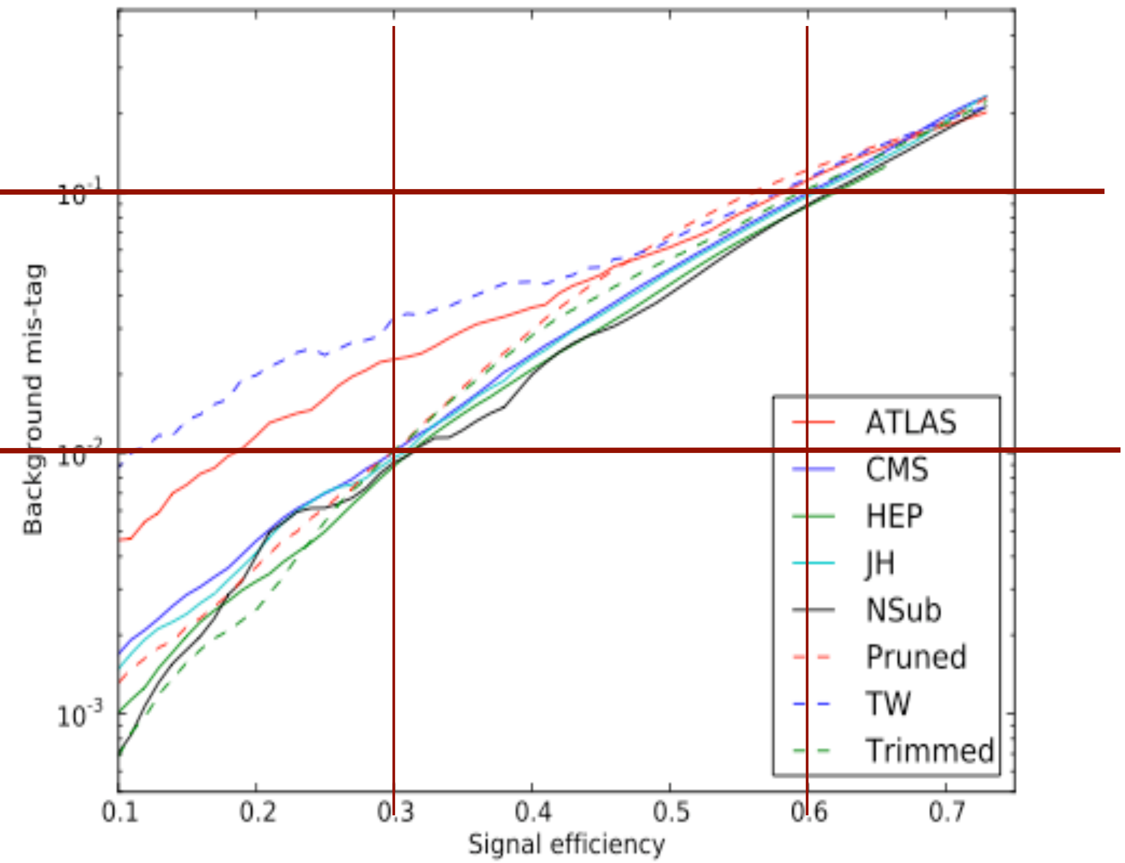
Even more curves now on this plot

Comparison of top taggers

Boost **2010** proceedings, arXiv:1012.5412



Boost **2011** proceedings, arXiv:1201.0008



Law of diminishing returns: improvement has become very hard

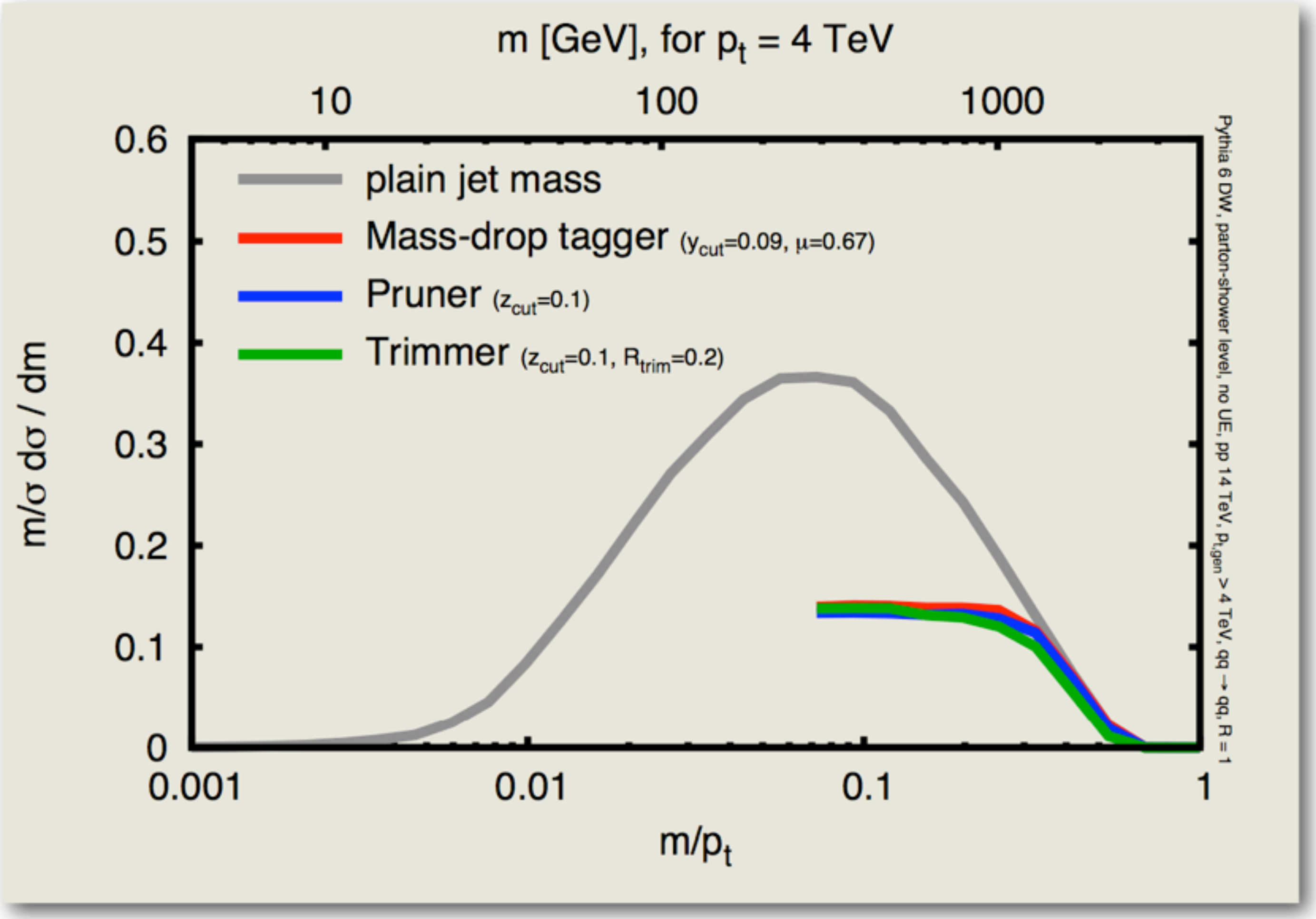
Jet substructure tools

- Darwinian evolution will eventually (hopefully!) select a few best tools, through:
  - checks that MCs ***reproduce data*** for critical variables/tools
  - checks that one can effectively eliminate contamination from ***pileup***
  - ***Effectiveness***
  - checks that the tools are ***robust***, and possibly can be ***understood analytically***

On the importance of checks

Taggers are relatively easy to write, once the basic ideas were spelled out in 2008. Hence their O(20+) proliferation.

Testing them properly, even only at the MC level, can be more tricky. An example from Dasgupta, Marzani, Salam

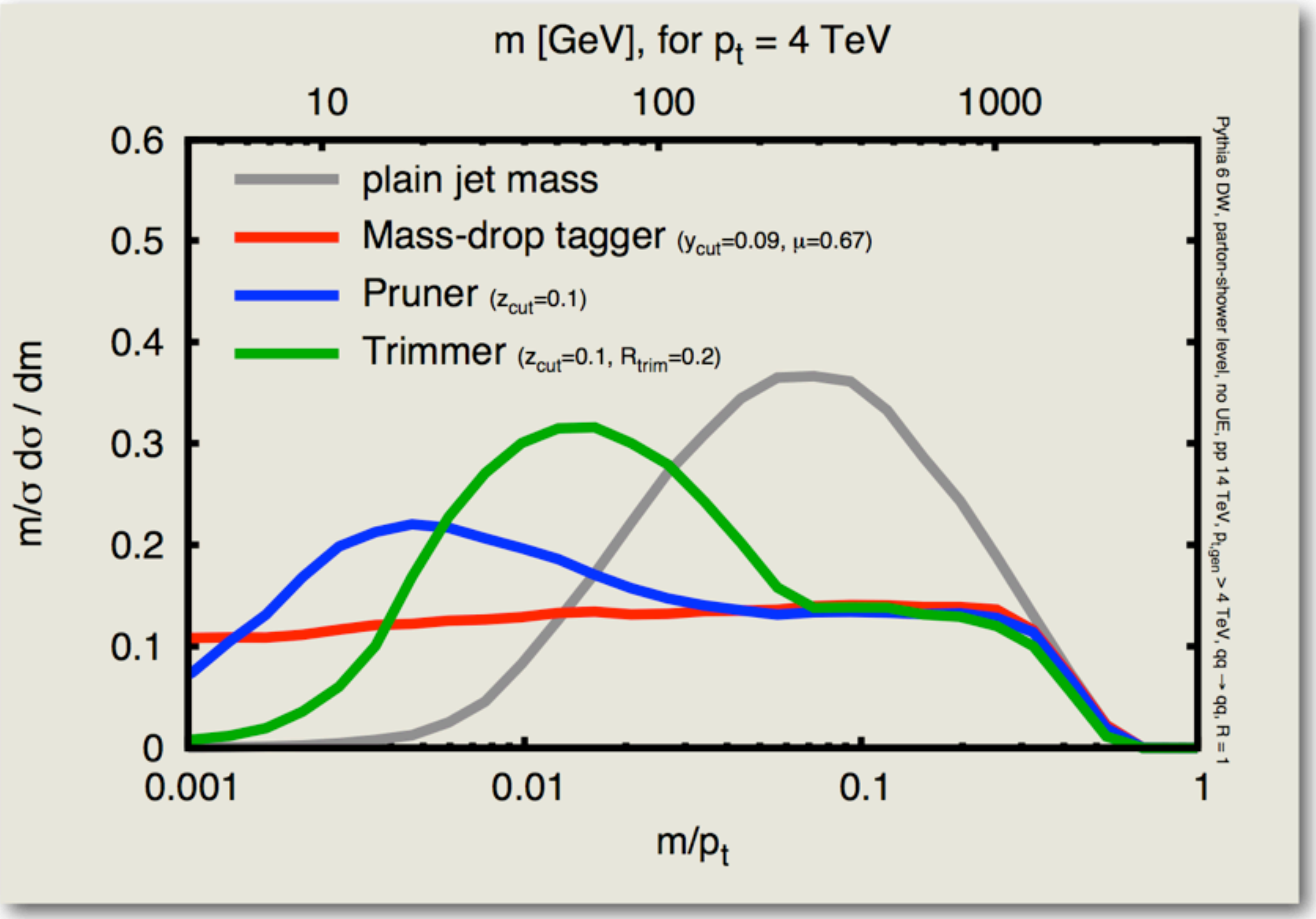


- Different taggers can appear to behave quite similarly in a limited range of masses.
- Mass-drop, trimming and pruning have often been considered as almost equivalent

On the importance of checks

Taggers are relatively easy to write, once the basic ideas were spelled out in 2008. Hence their O(20+) proliferation.

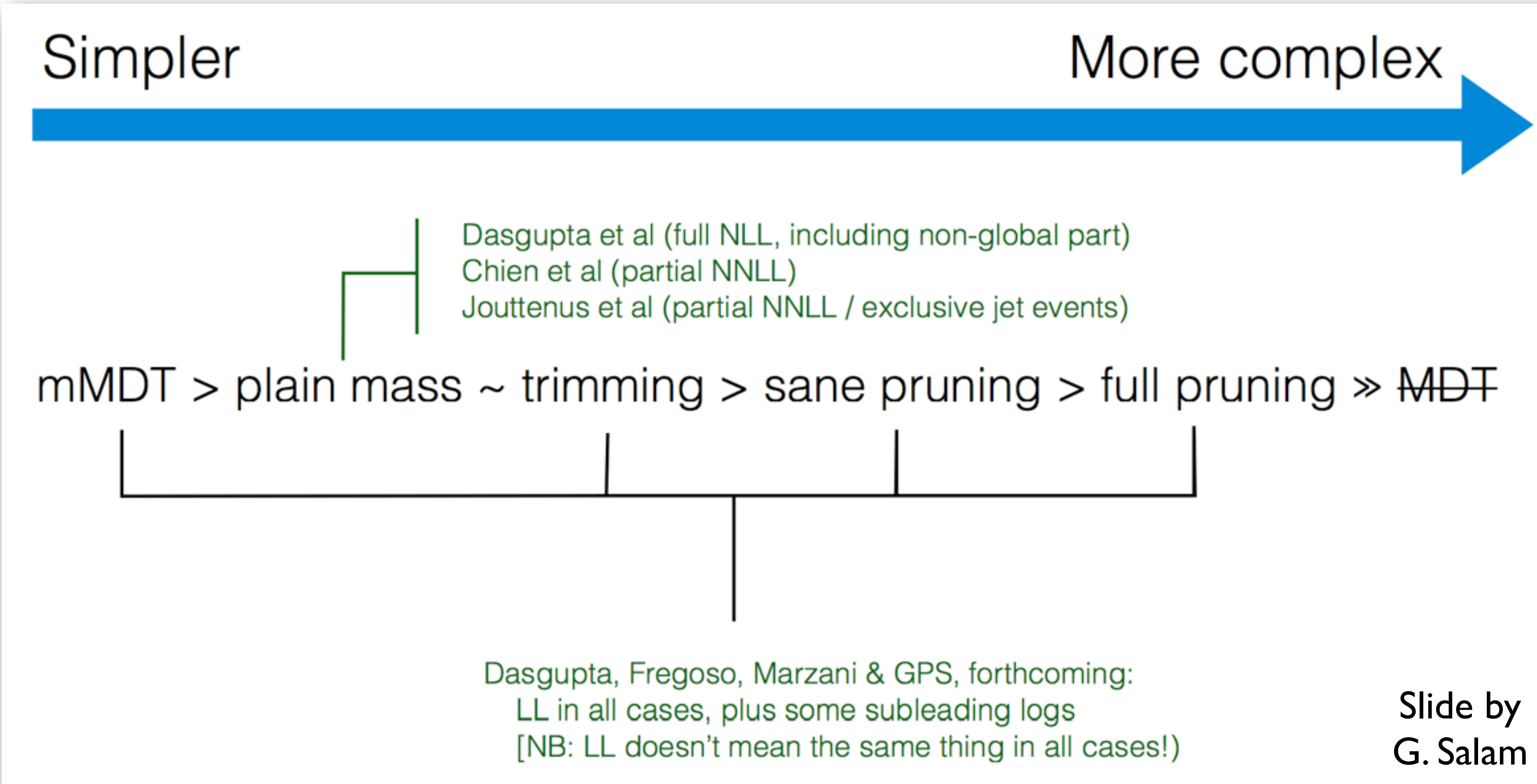
Testing them properly, even only at the MC level, can be more tricky.
An example from Dasgupta, Marzani, Salam



- Different taggers can appear to behave quite similarly in a limited range of masses.
- Mass-drop, trimming and pruning have often been considered as almost equivalent
- Extending the check one can see them to differ significantly

Analytic analysis of taggers

Analytic understanding allows one to study the behaviour of taggers without relying on a MC simulation (that may itself be less than exact)

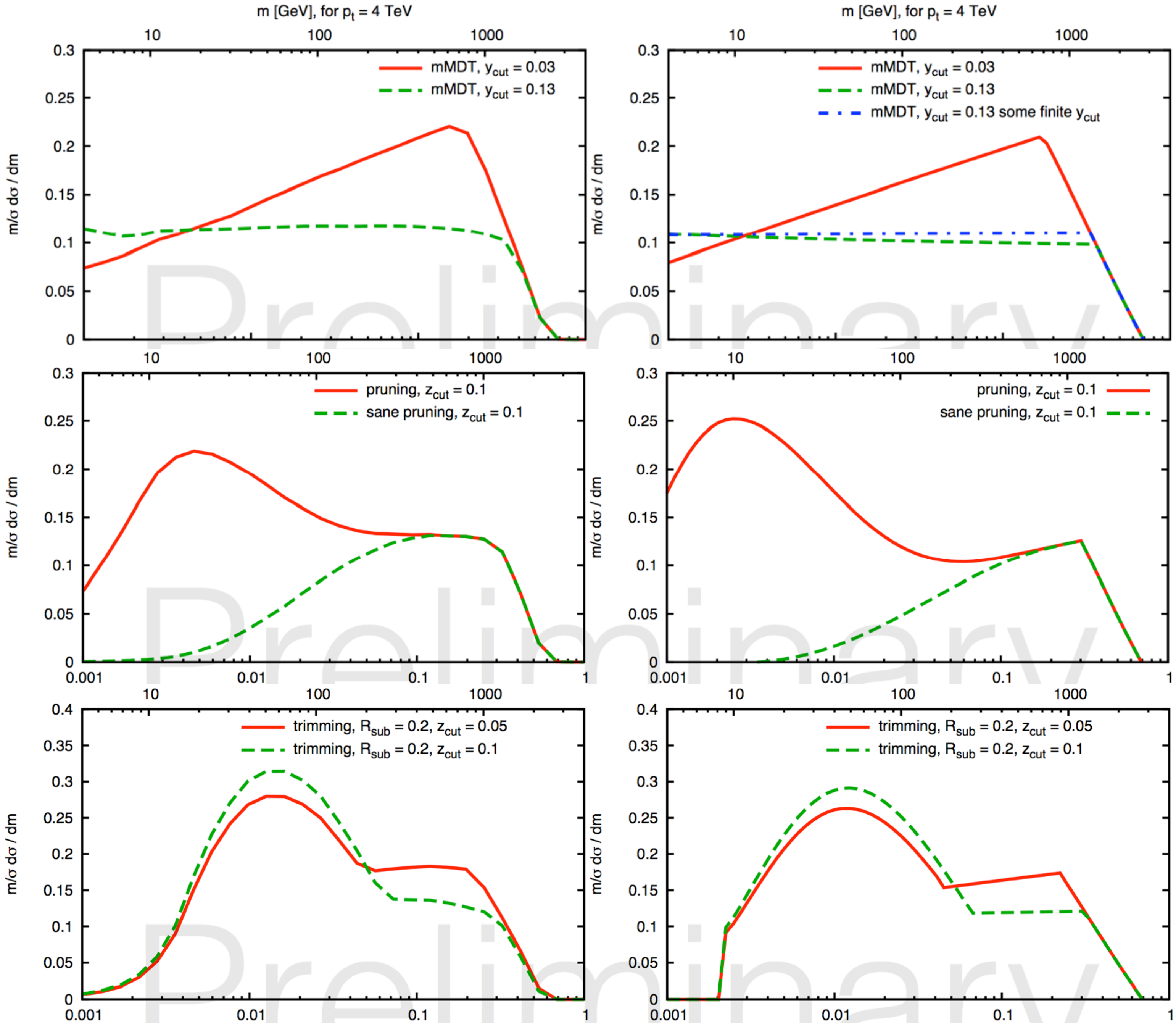


Slide by G. Salam

Analytic analysis of taggers

[Dasgupta, Fregoso, Marzani, Salam, preliminary]

Monte Carlo | **Analytic**



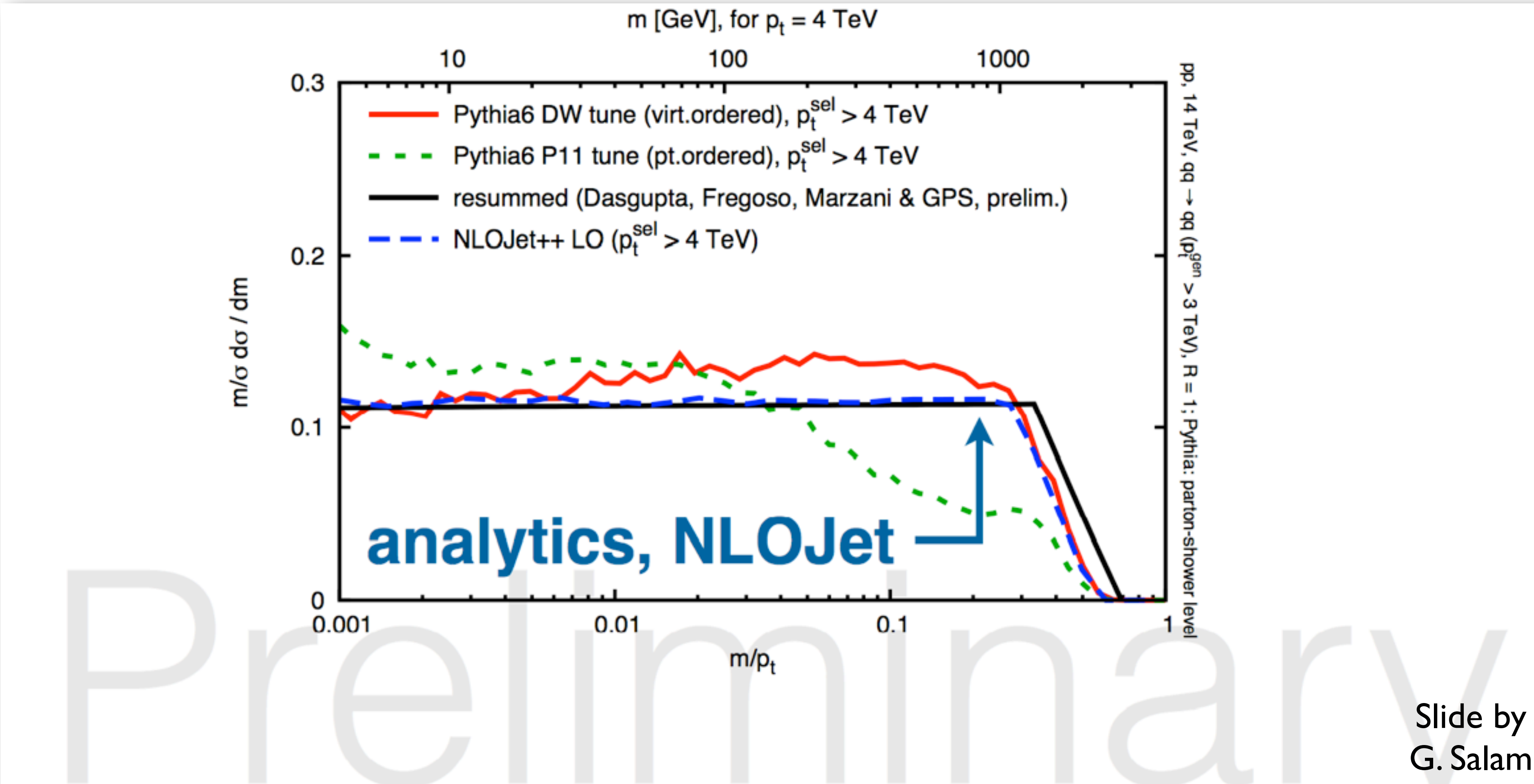
modified Mass-Drop

pruning

trimming

More on importance of analytics

Besides the ease of use and the wider reach, analytic control can be useful because MC simulations do not always agree



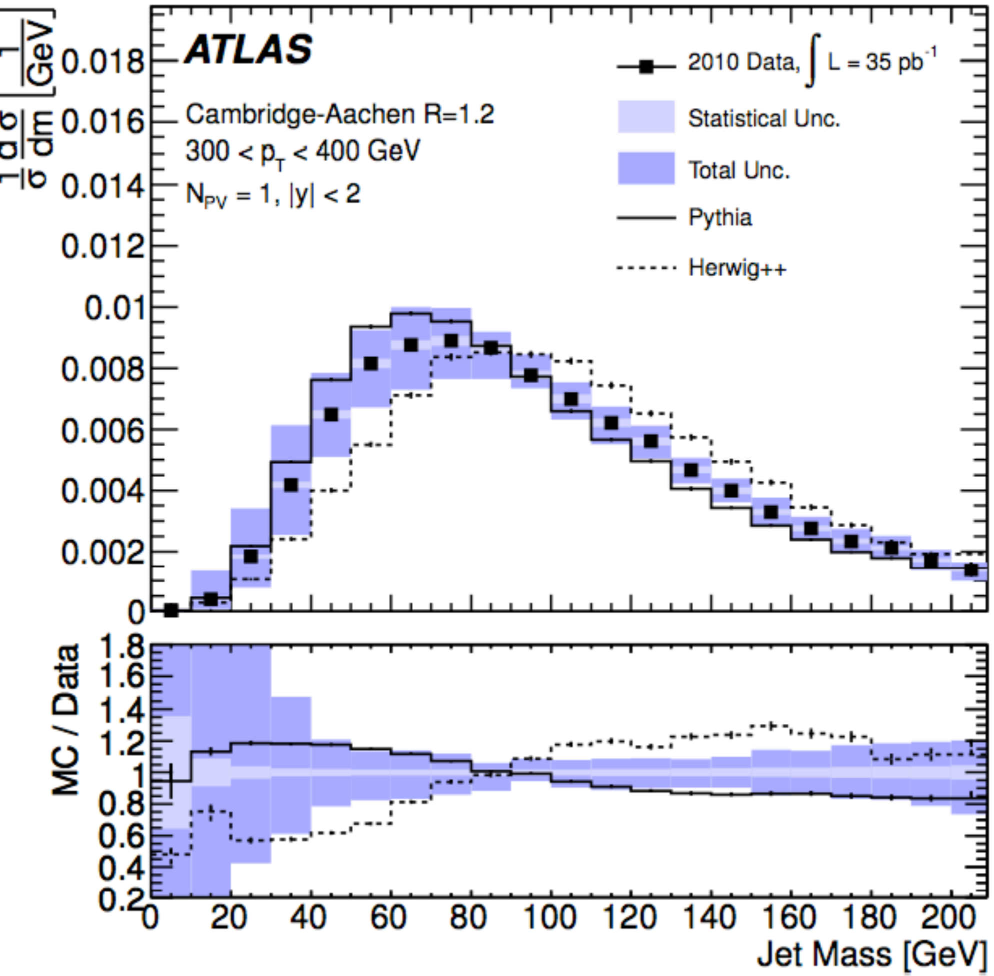
Slide by G. Salam

ATLAS 1203.4606:
First measurement of substructure variables at LHC

- ATLAS measured and compared to MCs' predictions *(splitted/filtered) jet mass*, k_t *splitting scales* and *N-subjettiness ratios*
- Effects of *pile-up*, and role of *grooming* techniques to reduce them, were also considered

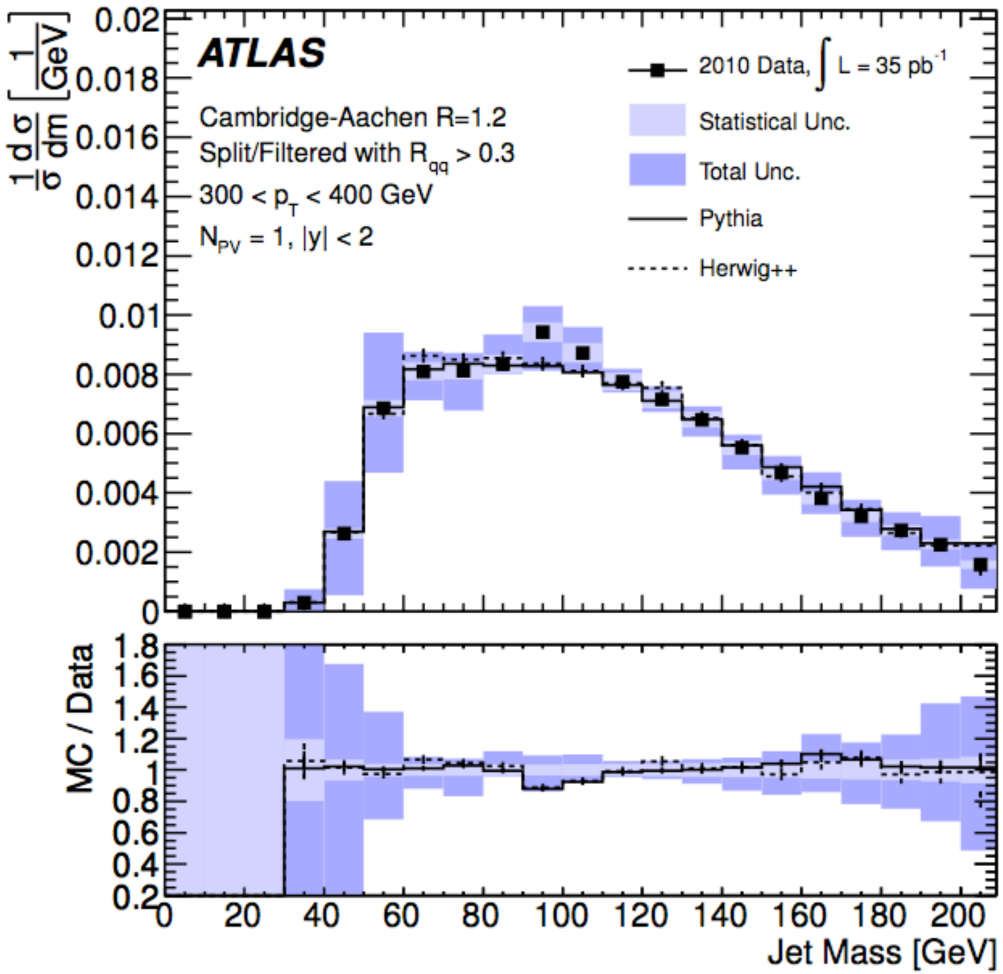
Jet substructure in ATLAS

jet mass



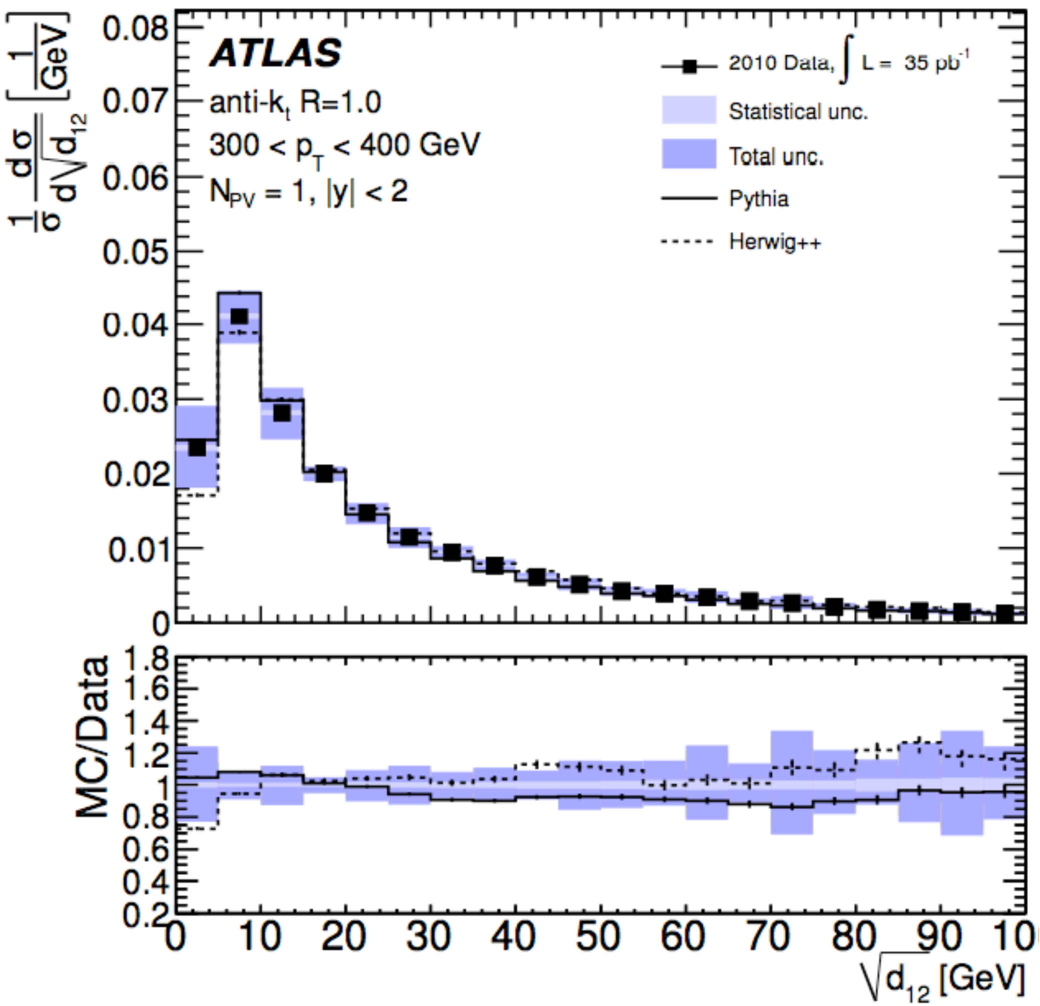
- “Plain” jet mass

split/filtered mass



- Jet mass after Mass-Drop tagger (with minimum ΔR_{12} requirement) and filtering

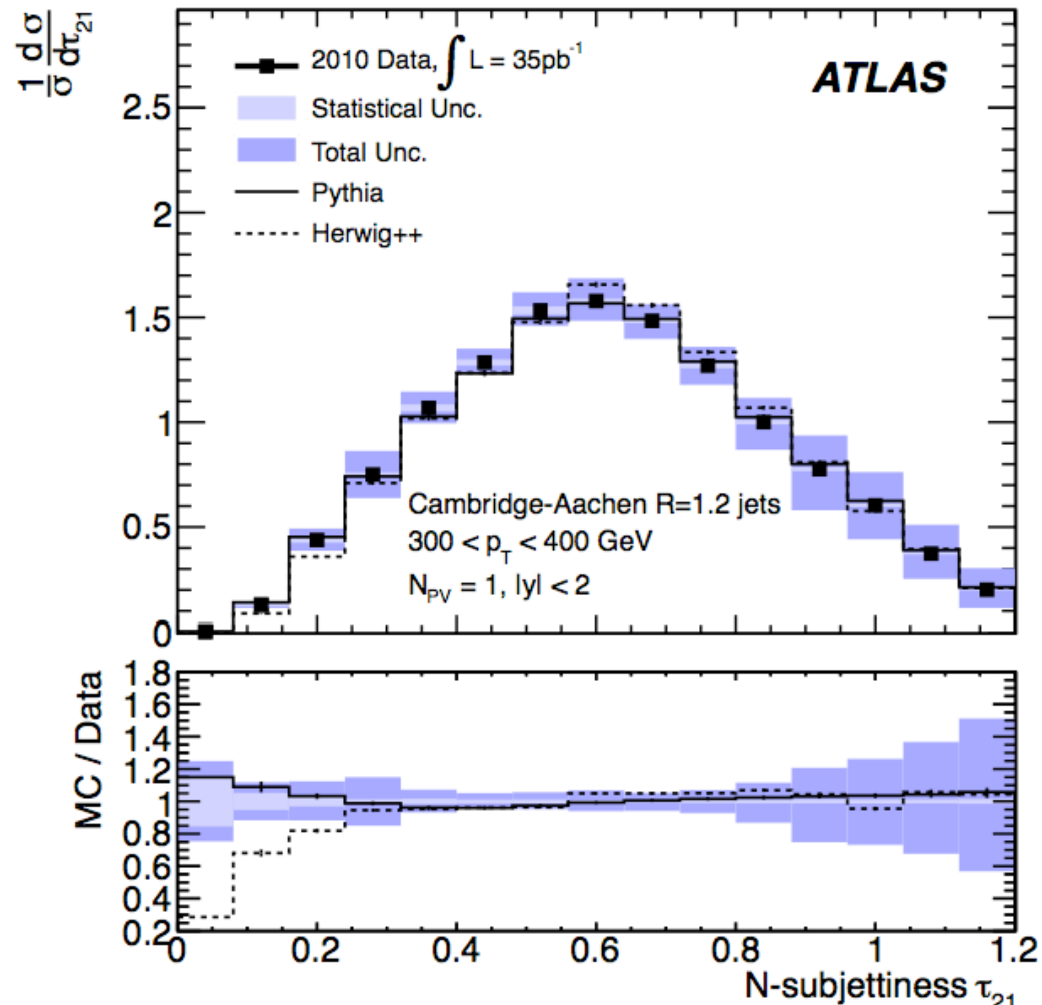
k_t distance



- Recluster the jet's' constituents with k_t
- Consider the scale of the last clustering step,

$$\sqrt{d_{12}} = \min(p_{t,1}, p_{t,2})\Delta R_{1,2}$$

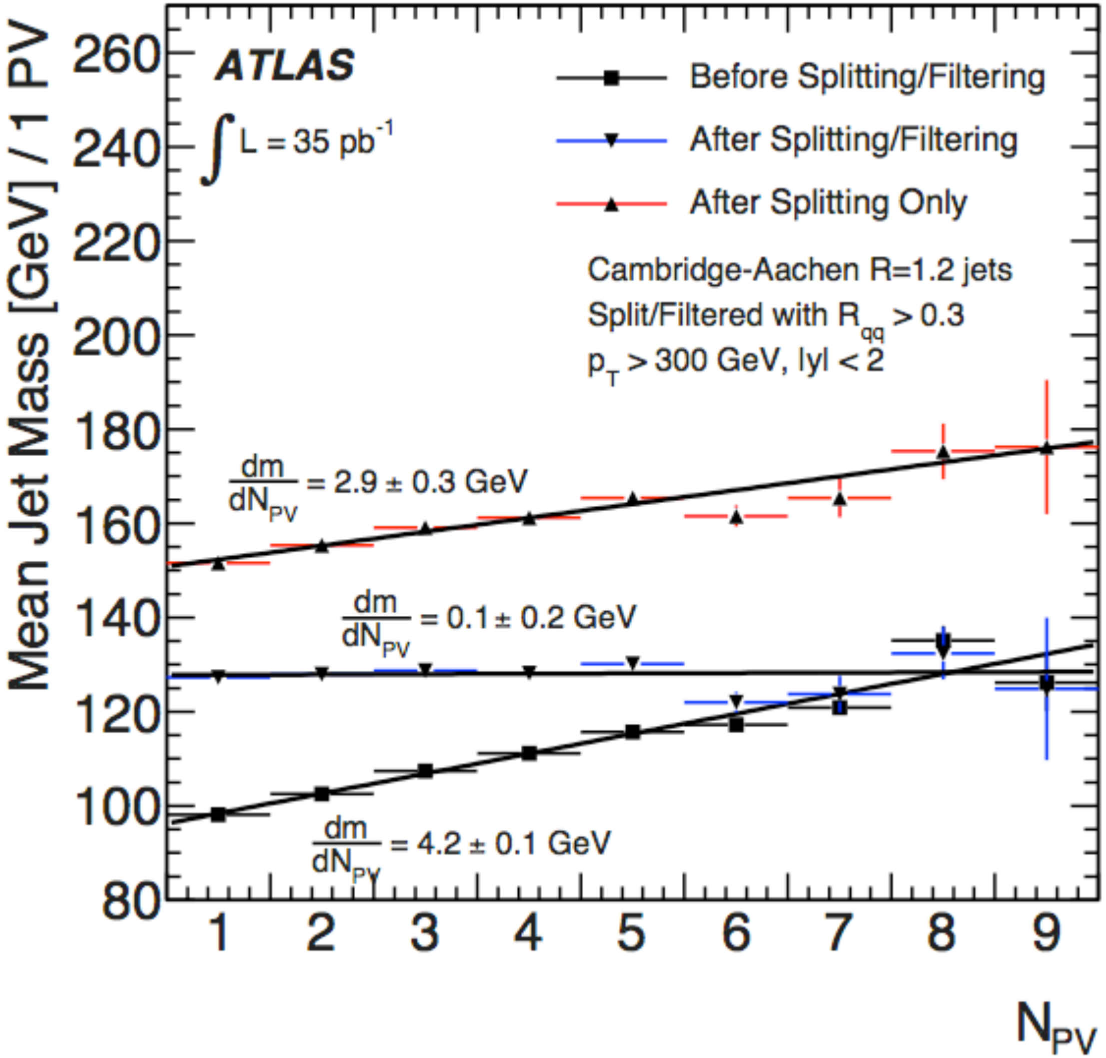
N-subjettiness



- τ_N measures how well a jet can be seen as being composed of N subjets
- Ratios are defined as

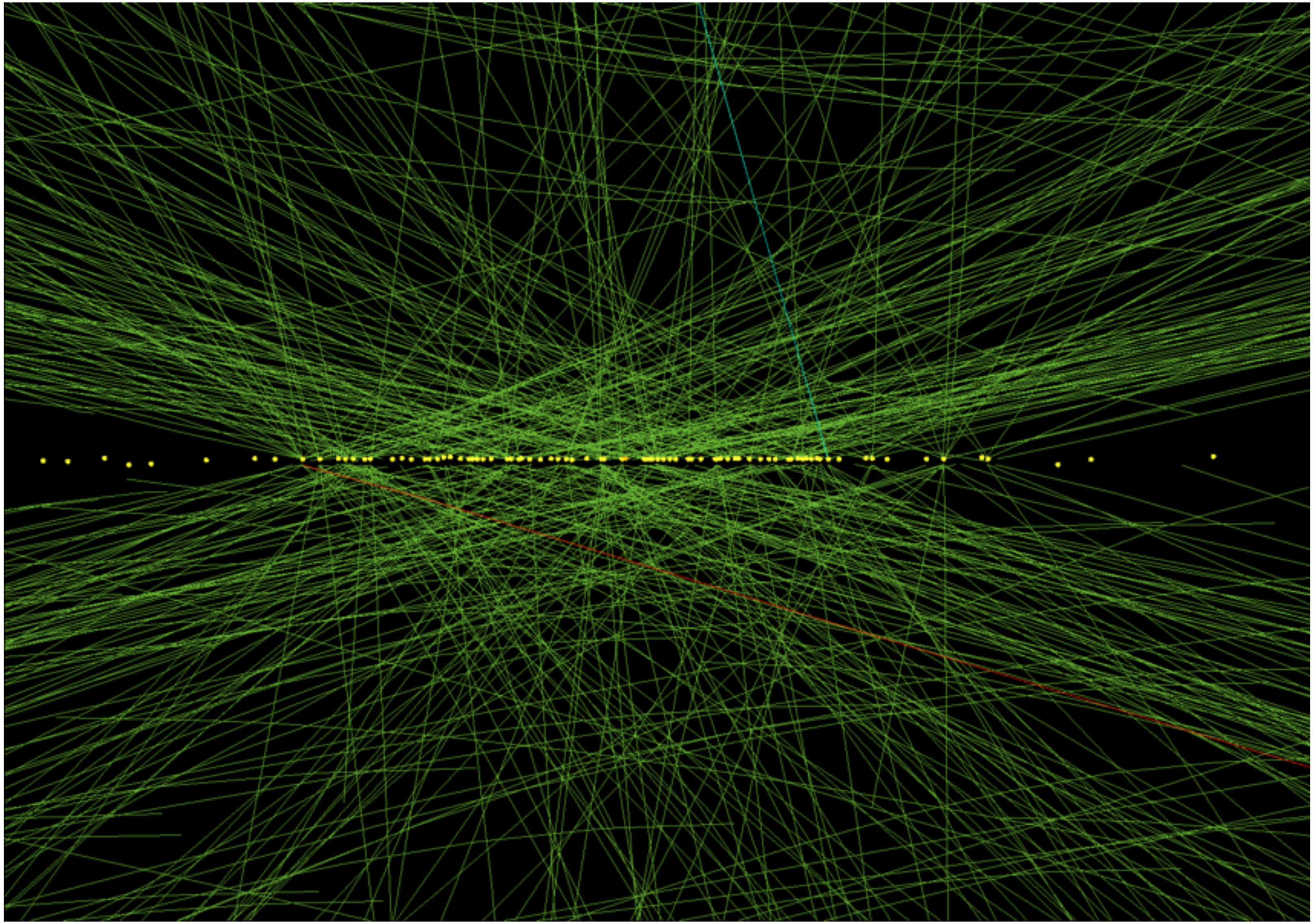
$$\tau_{N+1,N} = \frac{\tau_{N+1}}{\tau_N}$$

Jet substructure in ATLAS



- Effect of pileup: multiple interactions distort the observable.
- Independence from pileup is recovered for "tagged" mass after filtering
- NB: pilup conditions are/will be much higher than 9 simultaneous interactions

Pileup

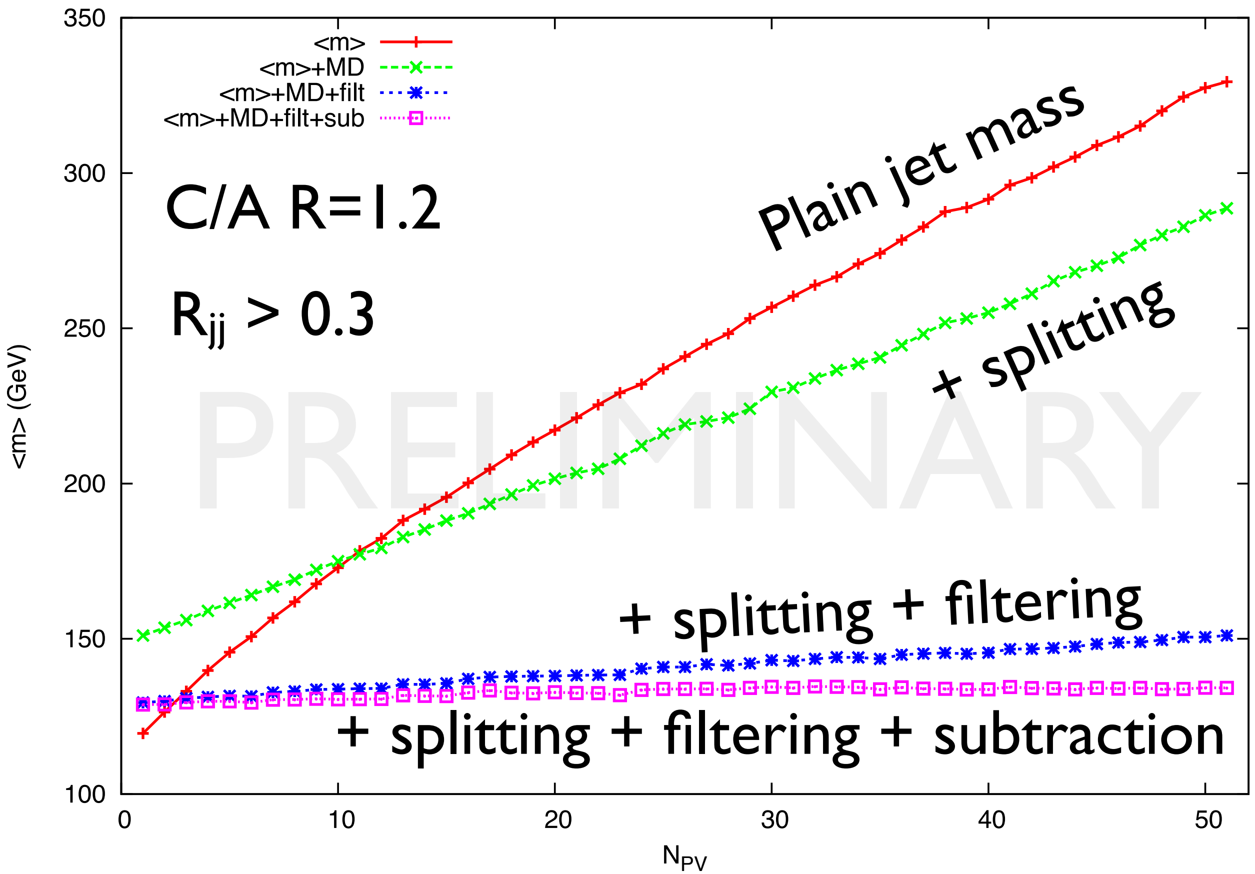


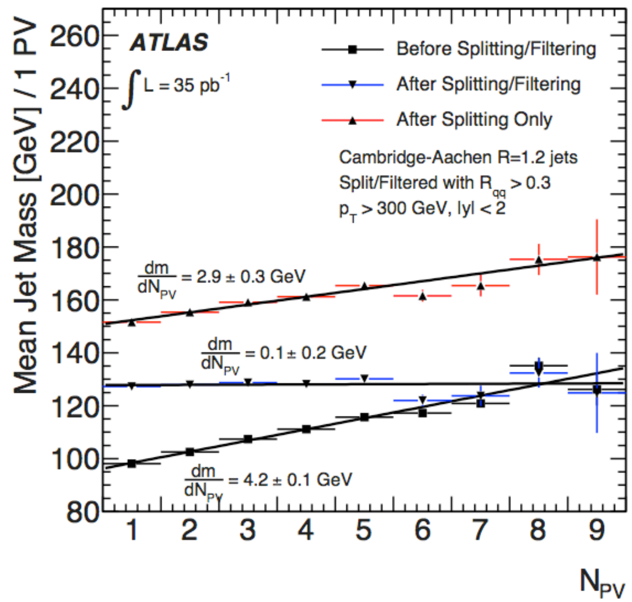
78-vertices event from CMS

https://cds.cern.ch/record/1479324

Effect of pileup

Extend ATLAS analysis to higher N_{PV}





- Filtering not sufficient for $N_{PV} \gtrsim 10$, but additional ***subtraction*** largely recovers the independence from the pileup level
- Combination of techniques can help

More on pileup subtraction

The $\mathbf{p_T^{raw}}$**-ρA** technique (also called **area/median**) only corrects a jet's transverse momentum

Each jet shape has its own specific sensitivity to background contamination.

How to correct them?

- One option is to study analytically each shape [Alon et al. 1101.3002]. Can be time consuming and cumbersome

- Alternatively, determine **numerically** the *susceptibility* of any IRC-safe jet shape to contamination [Soyez et al. 1211.2811] (this generalises the jet area)

Numerical jet shape correction

Numerical derivative w.r.t. ghosts momenta

Ghosts area

Jet shape as a function of the jets's constituents momenta

$$V_{\rm jet}^{[n]} \equiv A_g^n \frac{d^n}{dr_{t,g}^n} V(\{p_i\}_{\rm jet})$$

Pileup momentum density

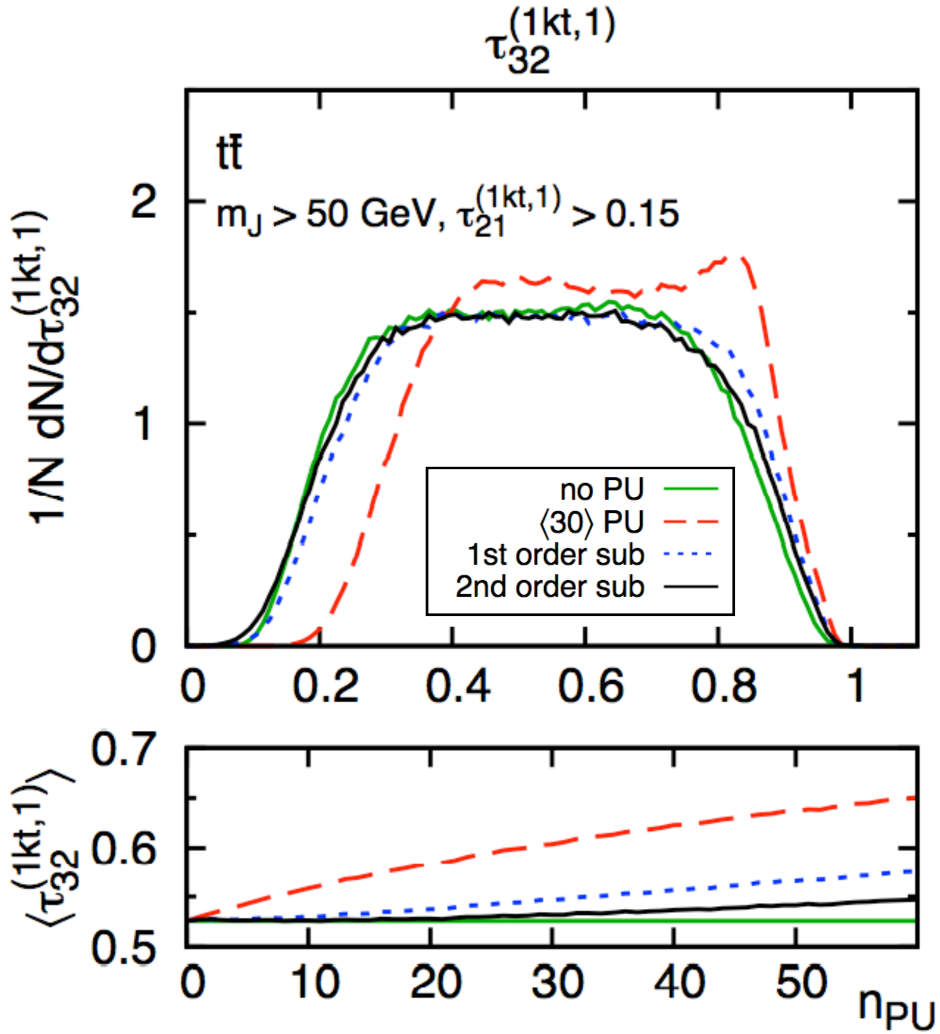
$$V_{\rm jet,sub} = V_{\rm jet} - \rho V_{\rm jet}^{[1]} + \frac{1}{2}\rho^2 V_{\rm jet}^{[2]} + \cdots$$

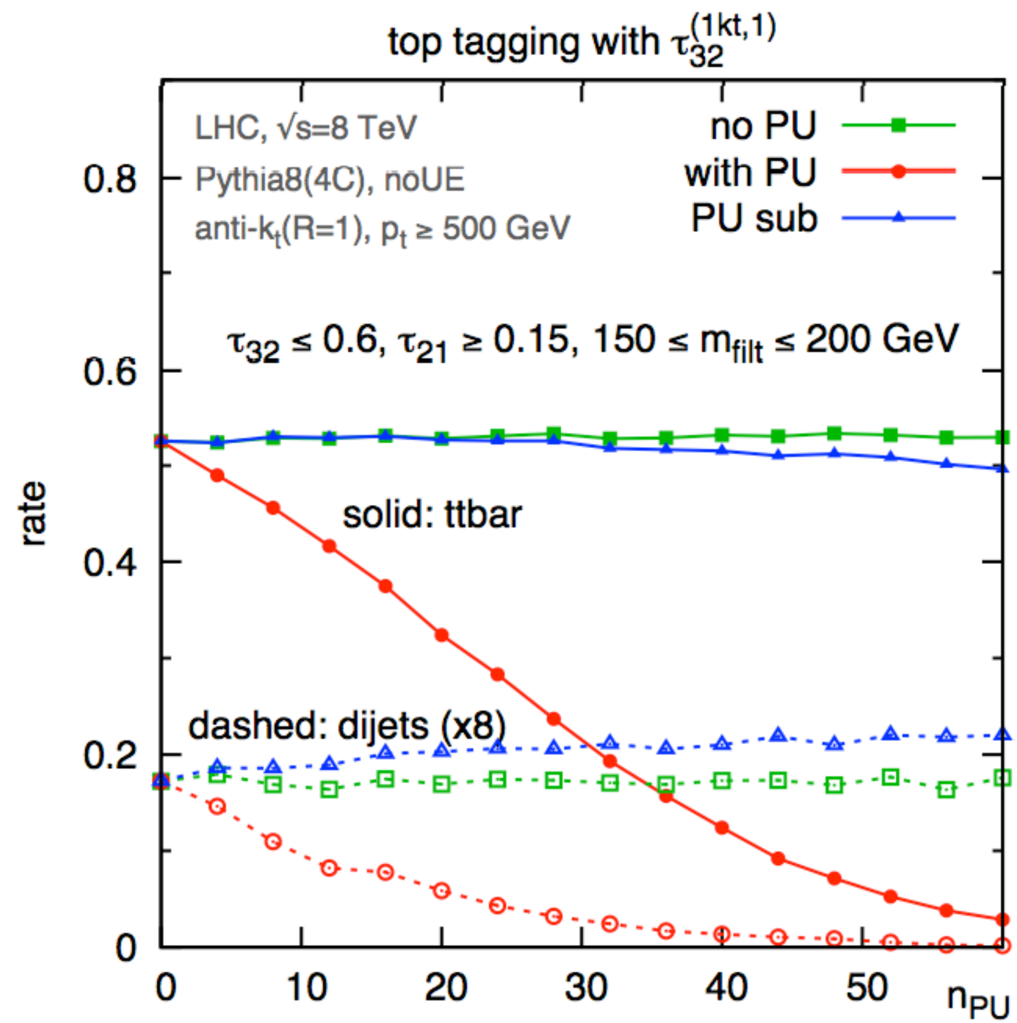
Numerical derivative w.r.t. ghosts momenta

This procedure generalises the transverse momentum correction to any jet shape

Numerical jet shape correction

Example: τ_{32} correction and top tagging





- Original distribution reproduced after pileup subtraction
- **Tagging rates independent of amount of pileup after correction of the jet shapes involved in the tagging**

FastJet Contrib

A public repository for 3rd party extensions of FastJet

FastJet - Hepforge

fastjet.hepforge.org/contrib/

FastJet is hosted by Hepforge, IPPP Durham

- fastjet.fr
- fastjet-contrib
- contrib svn

FastJet Contrib

The fastjet-contrib space is intended to provide a common location for access to 3rd party extensions of FastJet.

Download the current version: fjcontrib-1.002 (released 12 April 2013), which contains these contributions. Changes relative to earlier versions are briefly described in the NEWS file.

After downloading and unpacking, enter the fjcontrib-1.002/ directory and then run

```
./configure [--fastjet-config=FILE] [--prefix=...] [...]
make
make check          # optional
make install
```

By default the package installs to the same directories as the FastJet installation.

A contribution named "SomeContrib" is usually accessed by including "fastjet/contrib/SomeContrib.hh" in your C++ file, and linking with -lSomeContrib.

Developers who wish to develop their own contribution or submit new ones should use svn to checkout the contrib framework,

```
svn checkout http://fastjet.hepforge.org/svn/contrib/trunk fjcontrib
```

then run a local script to get the current set of individual contribs

```
cd fjcontrib/
scripts/update-contribs.sh
```

and follow the instructions in the README and DEVEL-GUIDELINES files.

Summary

- ‘Standard’ jets pretty much under control at LHC
  - Theory and measurements in good agreement, experimental precision going below the theoretical one (at NLO)
  - Theory continues to progress: multijet to NLO, dijet to NNLO

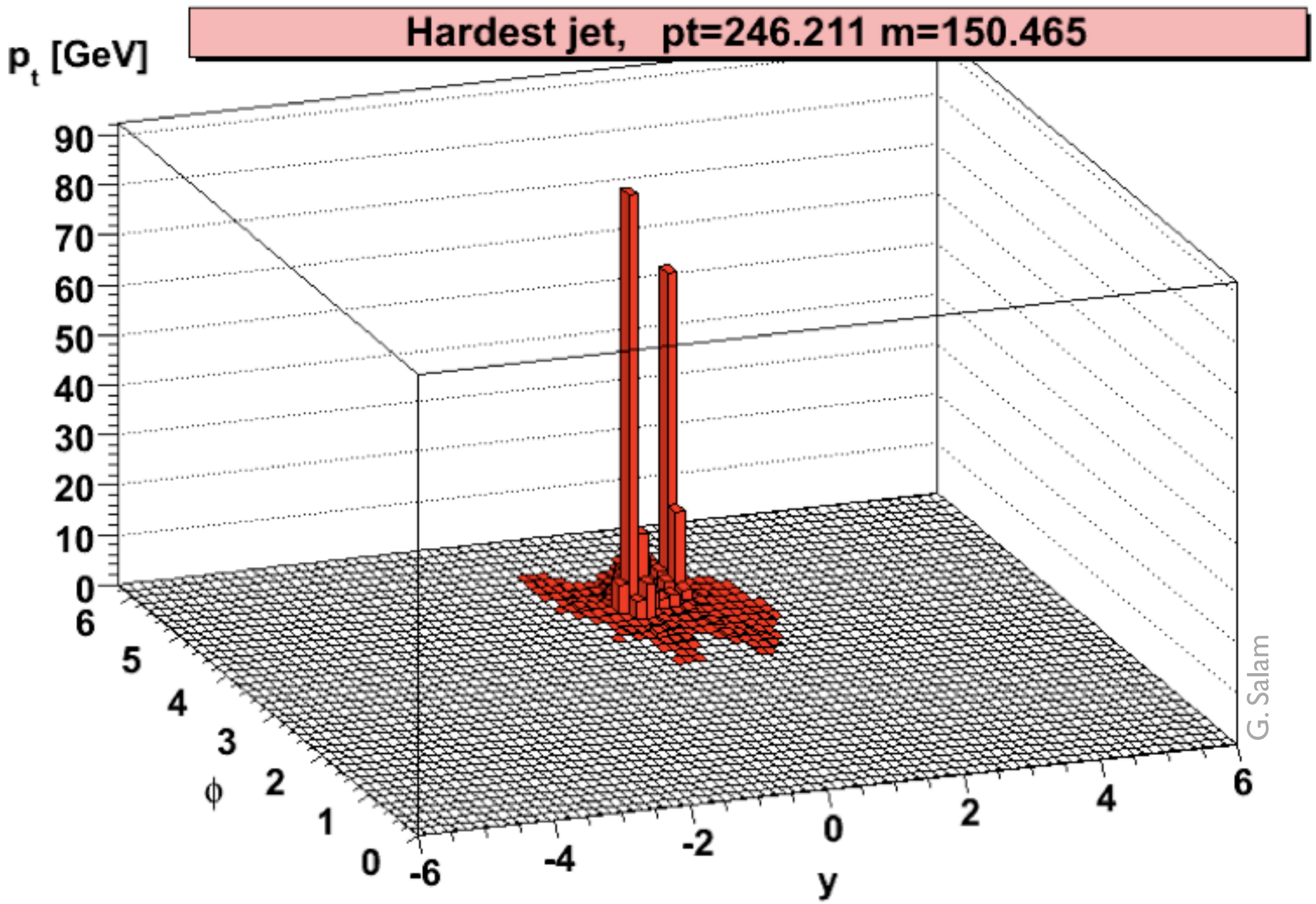
- The big news of the past 4-5 years has been the emergence of jet-based taggers and groomers
  - They have proven their worth in ‘Standard Model’ analyses
  - They are being implemented in BSM searches
  - **A word of caution**: we should try to avoid the balkanization that happened in the past with cone algorithms, and rather *try to grow a coherent, theoretically sound, robust, well tested and standardised library of tools*

Extra material

Boosted Higgs tagger

$pp \to ZH \to \nu\bar{\nu}b\bar{b}$

Butterworth, Davison, Rubin, Salam, 2008



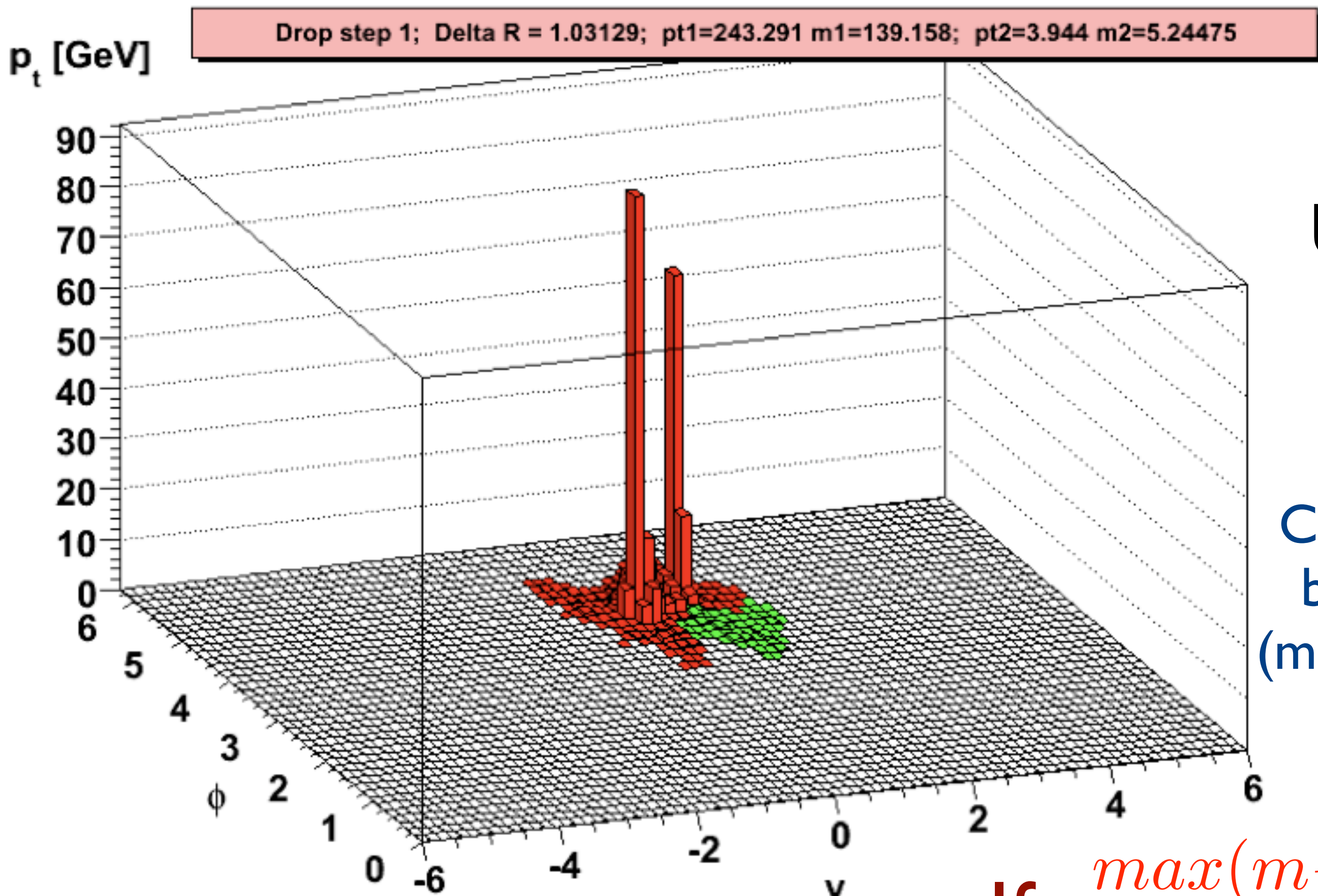
Start with the hardest jet

Use C/A with large R=1.2

m = 150 GeV

Boosted Higgs tagger

$pp \to ZH \to \nu\nu bb$



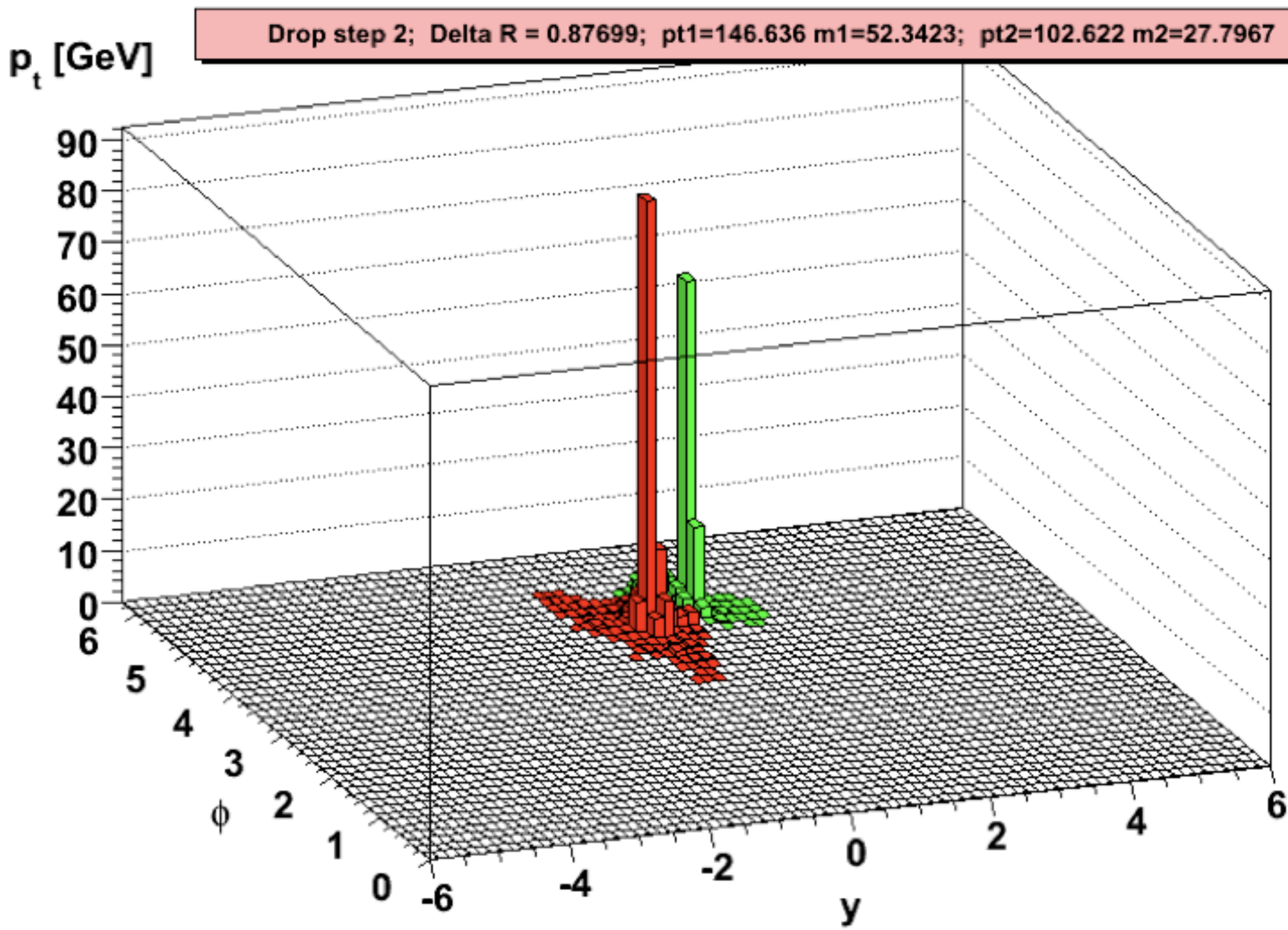
Undo last step of clustering

Check how the mass splits between the two subjets (m_1 = 139 GeV, m_2 = 5 GeV)

If $\frac{max(m_1, m_2)}{m} > \mu$ repeat

Boosted Higgs tagger

pp → ZH → ννbb



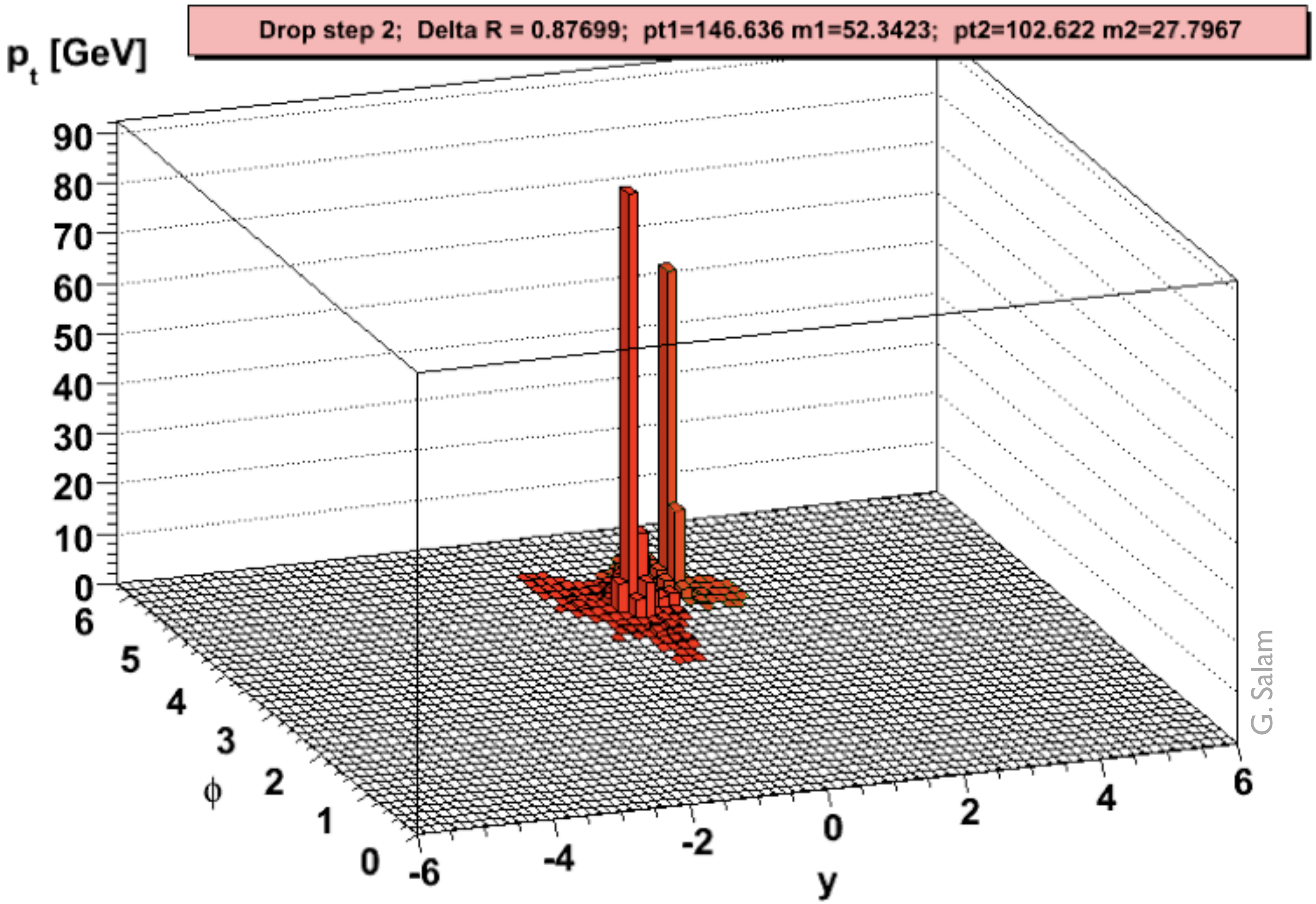
m_1 = 52 GeV, m_2 = 28 GeV

Stop when a large mass drop is observed

(and recombine these two jets)

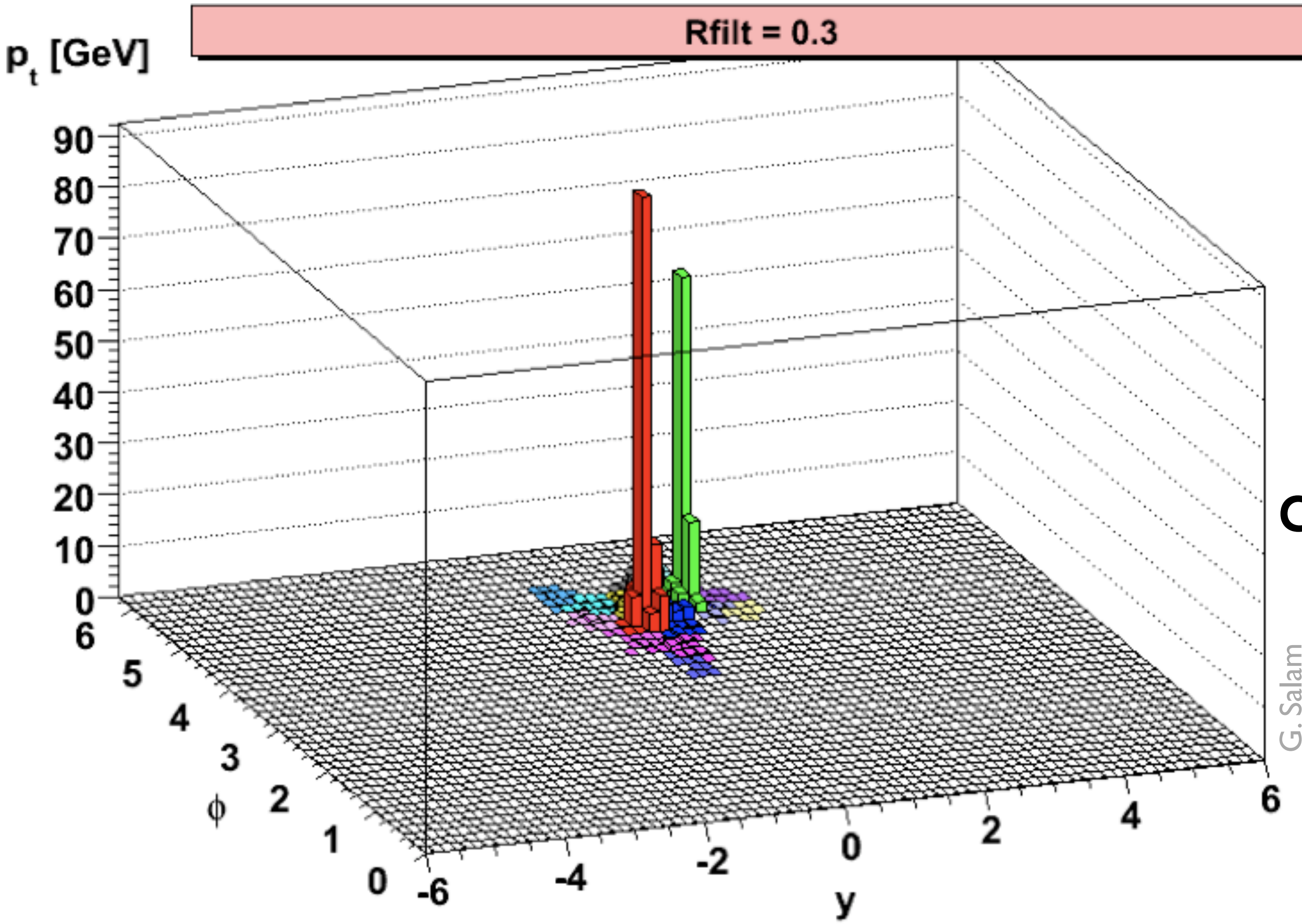
Filtering in action

Butterworth, Davison, Rubin, Salam, arXiv:0802.2470



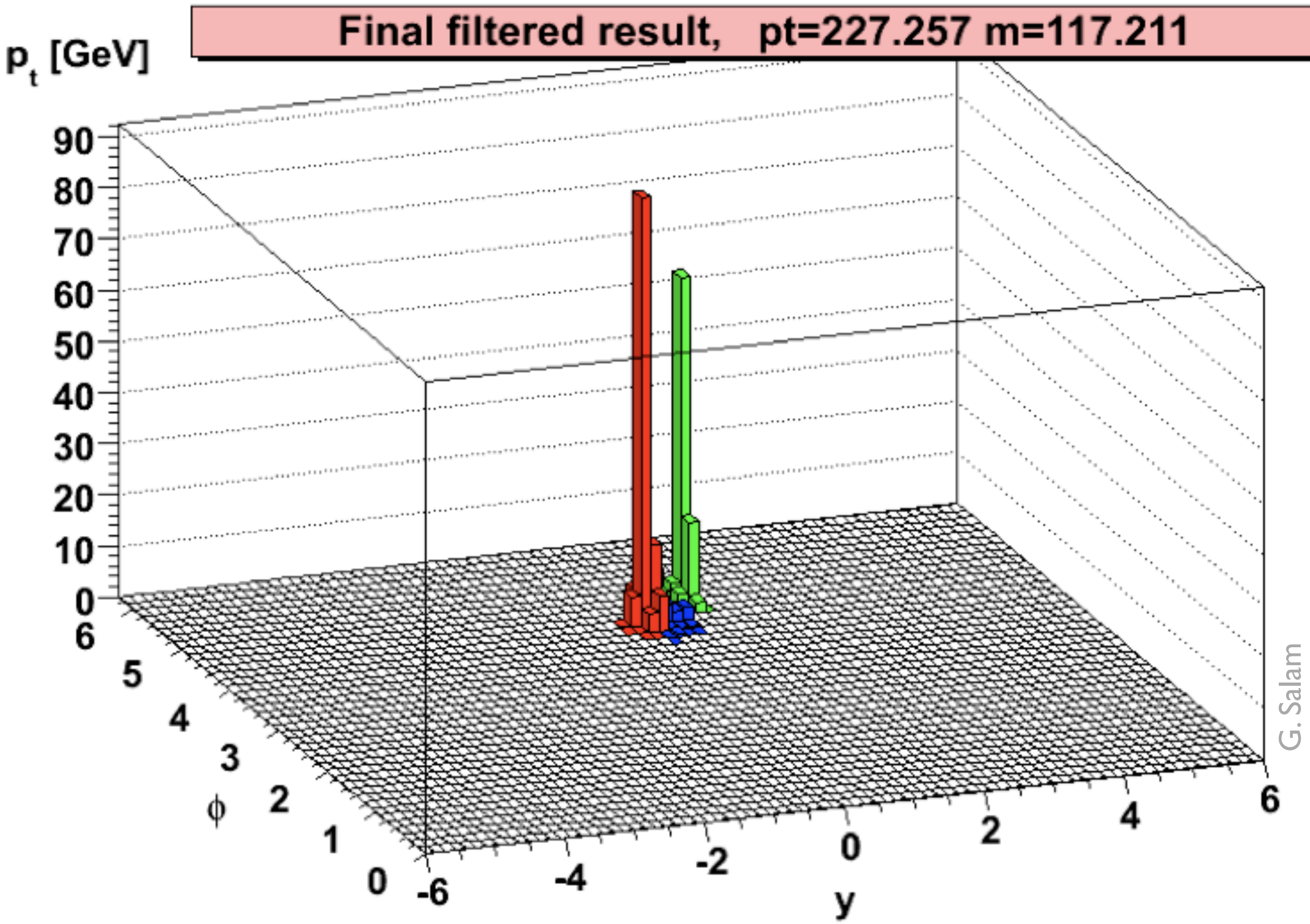
Start with a jet

Filtering in action



Recluster the contituents with R_{filt}

Filtering in action



Only keep the n_{filt} hardest jets

The low-momentum stuff surrounding the hard particles has been removed