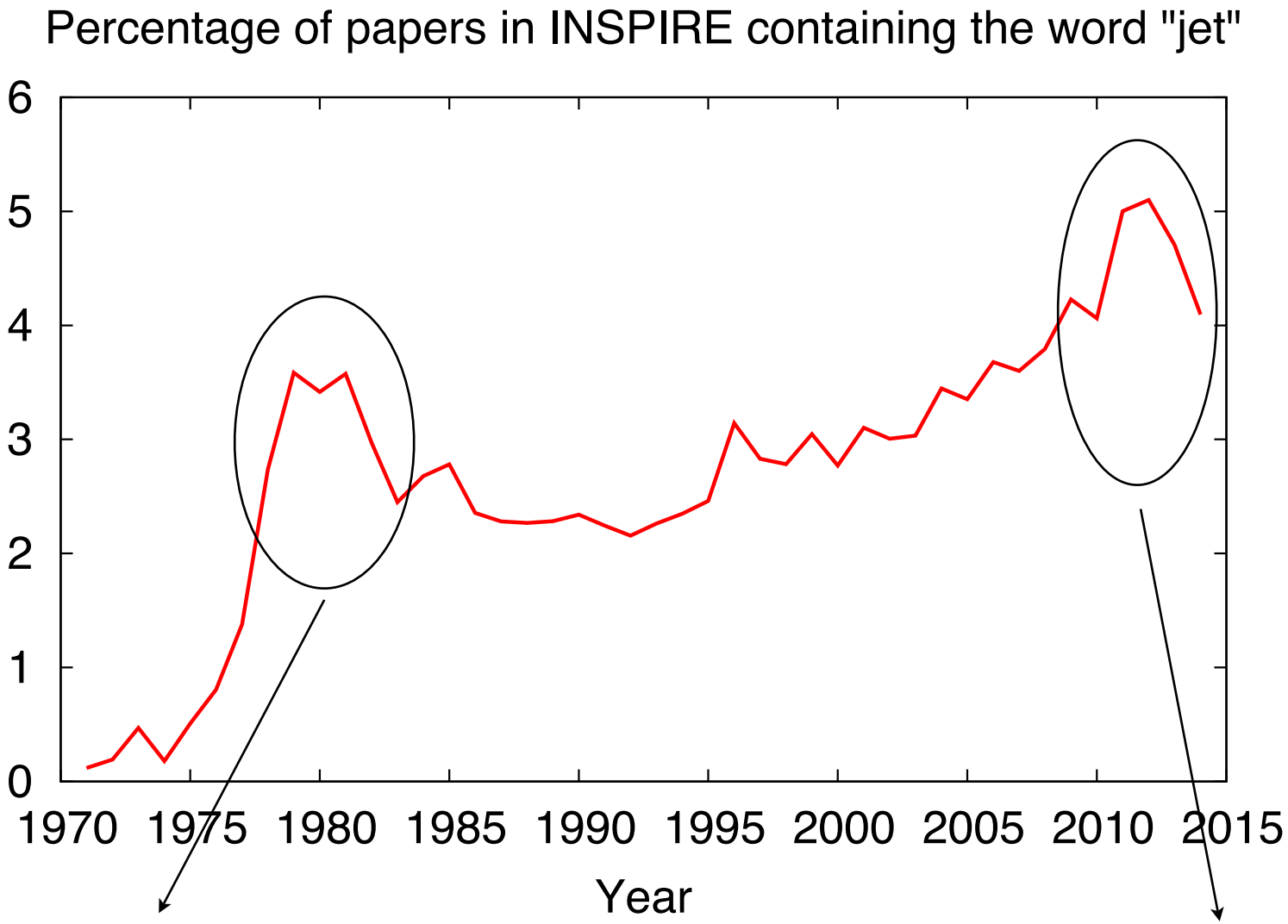
Jets Physics
Algorithms and Substructure Techniques

Matteo Cacciari
LPTHE Paris

ICHEP VALENCIA 2014
37th INTERNATIONAL CONFERENCE ON HIGH ENERGY PHYSICS
2-9-JULY-2014-VALENCIA

Jet revolutions



First 'jet revolution'

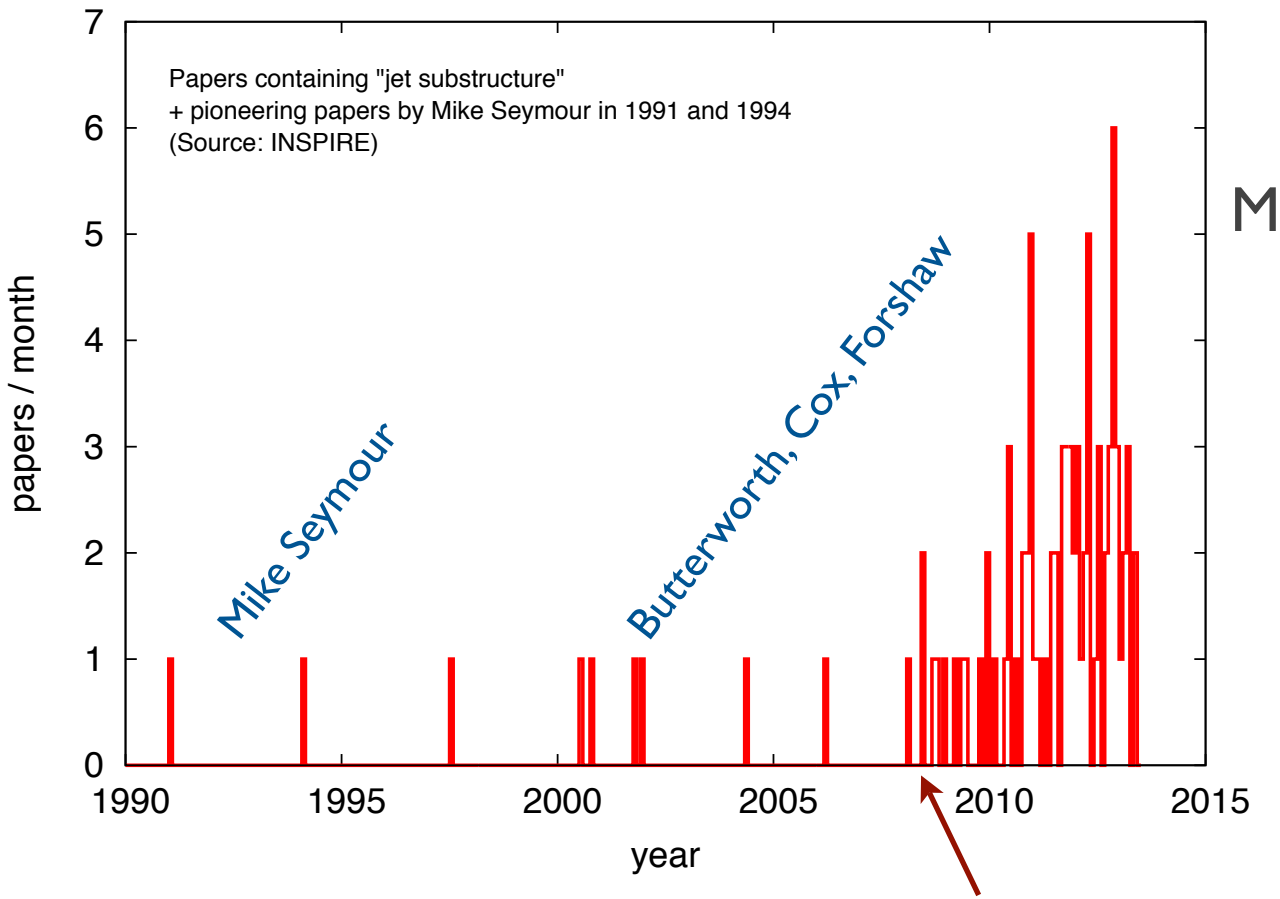
Second 'jet revolution'?

(1% ≈ 500 papers)

NB: Data not noise-subtracted

‘Jet substructure’ papers in INSPIRE

Number of papers containing the words ‘jet substructure’



More than 100 papers since 2008
(+ some background noise)

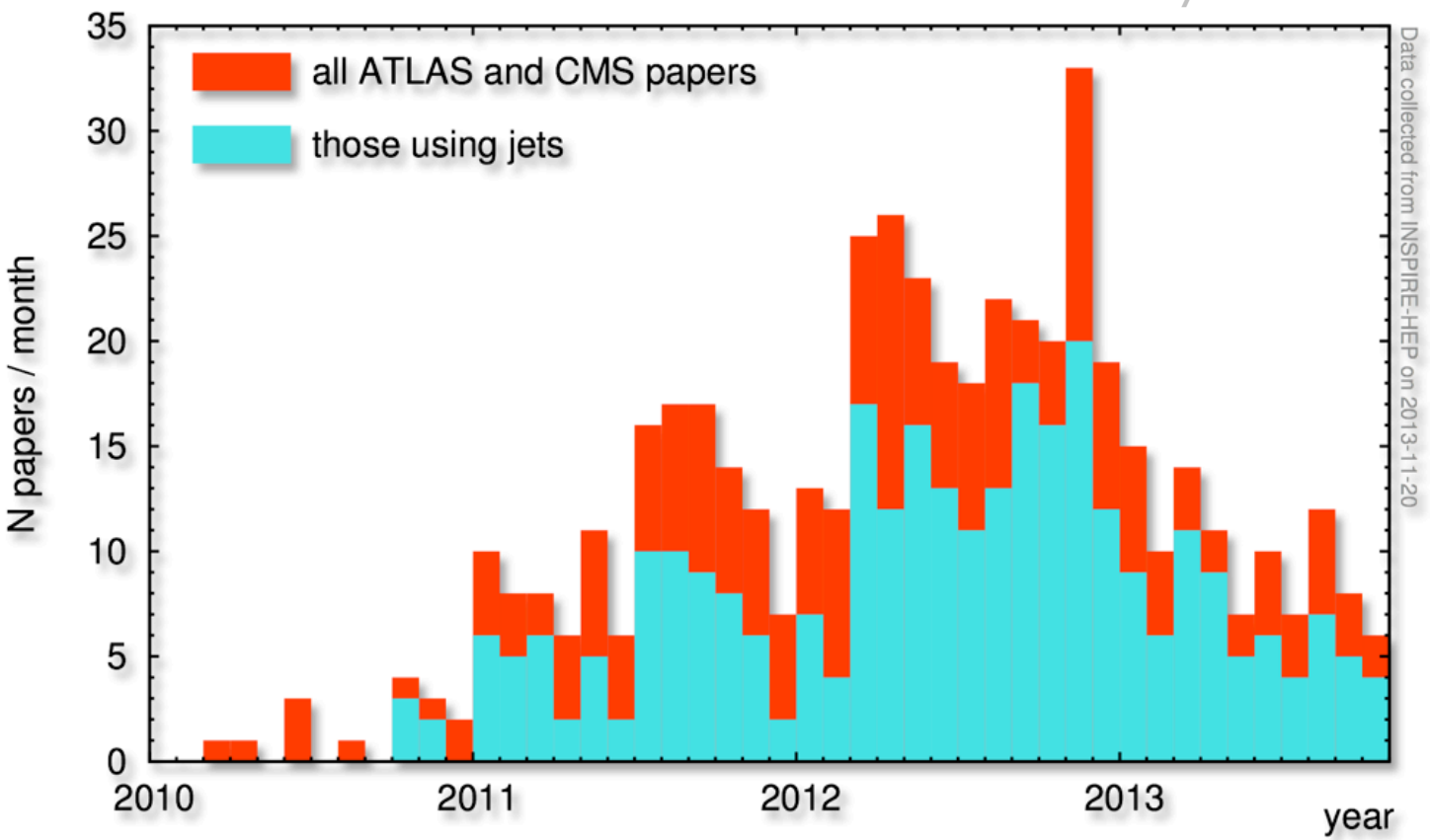
Pioneered by M. Seymour in the early ‘90s, rebooted by BDRS paper

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]**

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**

Plot by G. Salam

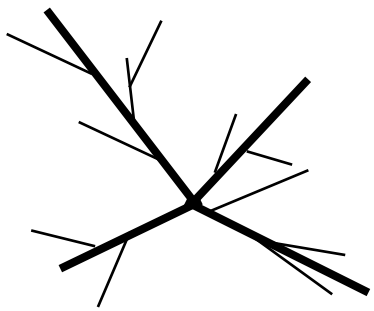


(Source: INSPIRE. Results may vary when employing different search keywords)

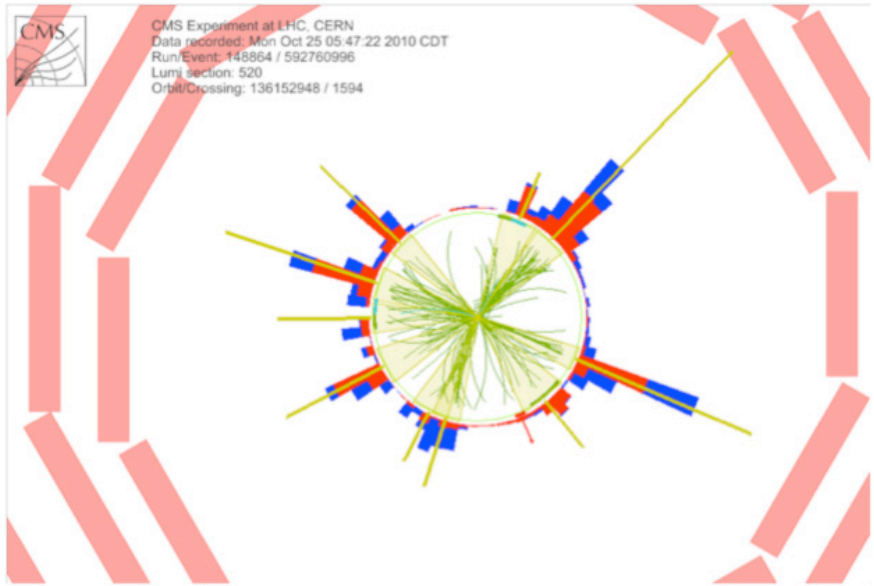
Why are jets so important?

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**

Janus Jets

From Wikipedia:

In ancient Roman religion and myth, **Janus** (Latin: *Ianus*, pronounced ['iaː.nus]) is the god of beginnings and transitions, and thereby of gates, doors, passages, endings and time. He is usually depicted as having two faces, since he looks to the future and to the past. The Romans named the month of January *(Ianuarius)* in his honor.

Like Janus, jets can serve two purposes:

- ***Observables***

  to be defined, calculated, measured

- ***Tools***

  to be used to extract specific properties of the final state

Requirements for jet algorithms

- In order to be most useful as an *observable*, a jet algorithm should be usable in perturbation theory and its jets not too affected by non-perturbative physics
  - Infrared and collinear safety
  - Controlled sensitivity to hadronisation and underlying event
  - Possibility to subtract effectively pile-up contamination

- Two main classes of jet algorithms:
  - cones
  - sequential recombination algorithms
    - Sequential recombination algorithms (k_t, Cambridge/Aachen, anti-k_t) provide a clustering history
      - When physically meaningful, the clustering history can be exploited to explore the substructure of a jet → jet as a tool

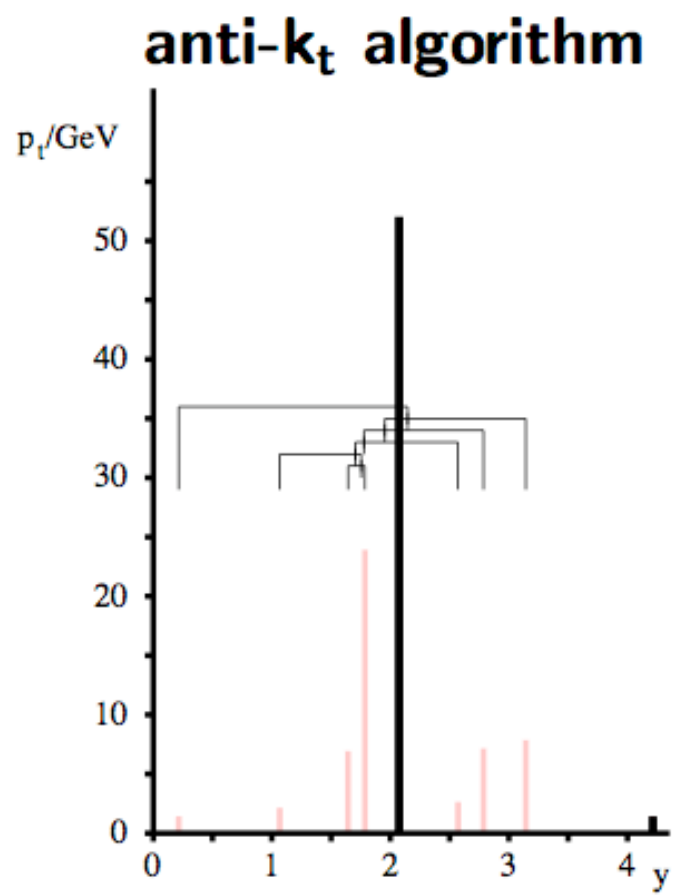
IRC safe algorithms

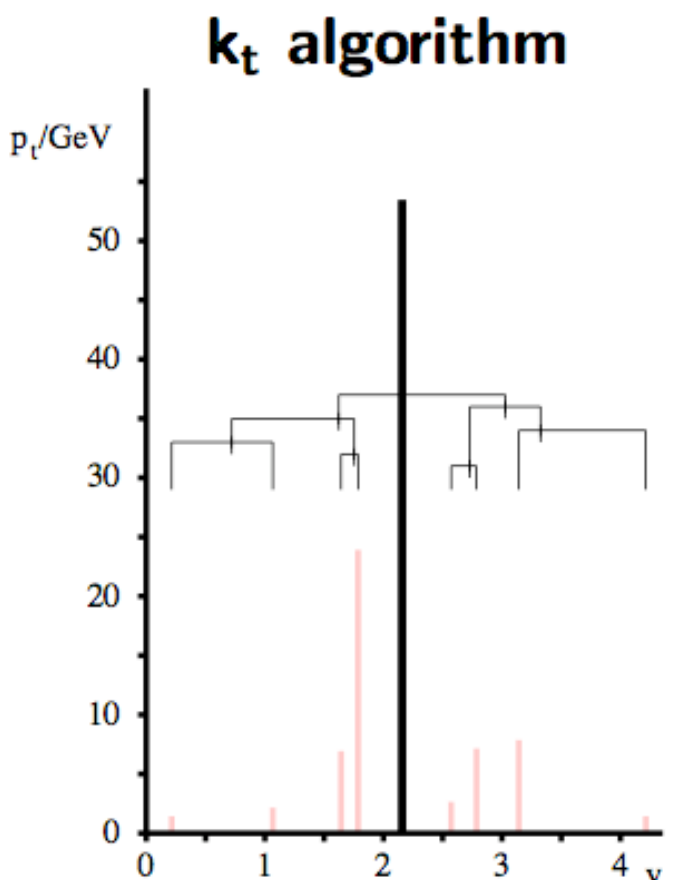
k_t	SR $d_{ij} = \min(k_{ti}^2, k_{tj}^2)\Delta R_{ij}^2/R^2$ hierarchical in rel p_t	Catani et al '91 Ellis, Soper '93	NlnN
Cambridge/ Aachen	SR $d_{ij} = \Delta R_{ij}^2/R^2$ hierarchical in angle	Dokshitzer et al '97 Wengler, Wobish '98	NlnN
anti-k_t	SR $d_{ij} = \min(k_{ti}^{-2}, k_{tj}^{-2})\Delta R_{ij}^2/R^2$ 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^2lnN

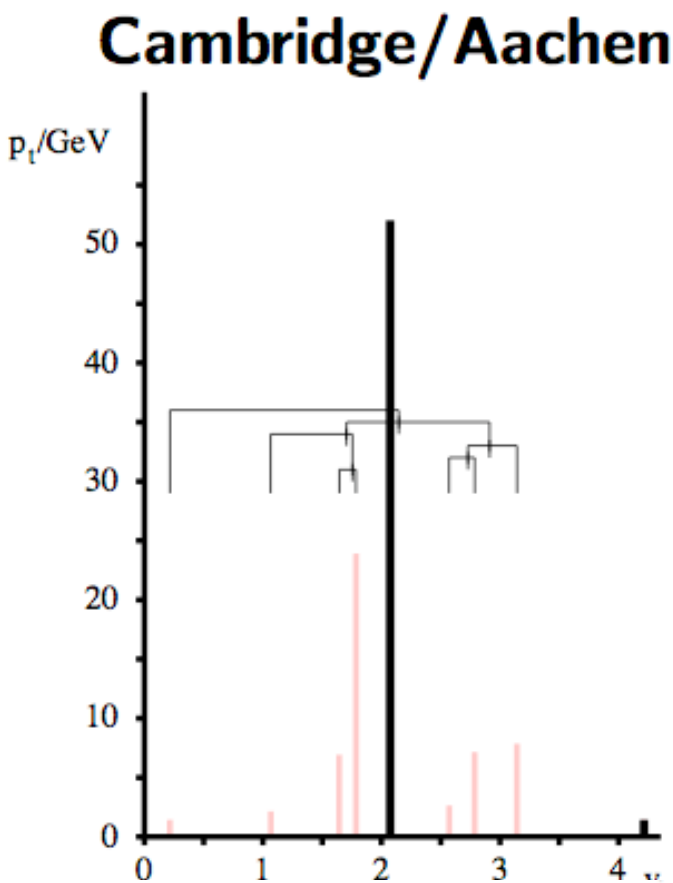
'second-generation' algorithms

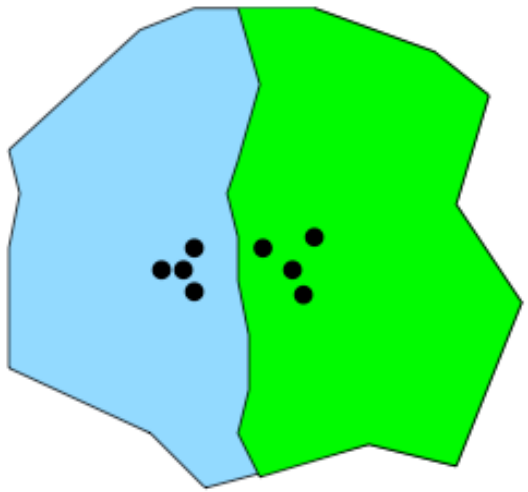
All (as well as many IRC unsafe ones) are available in FastJet, http://fastjet.fr

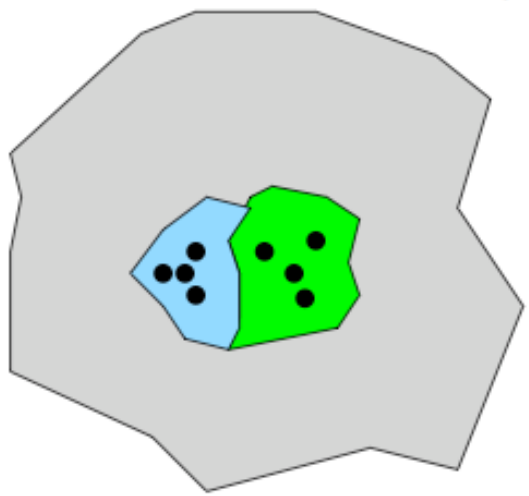
Hierarchical substructure





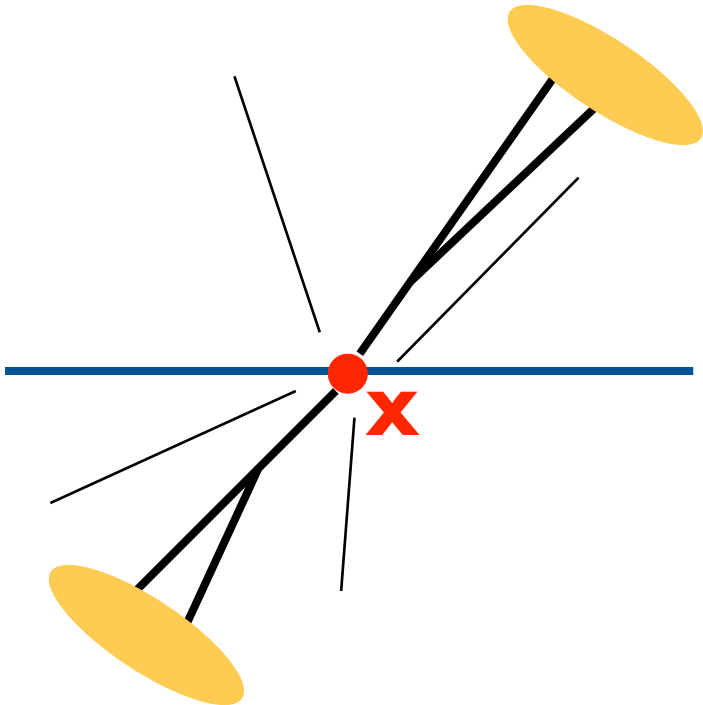




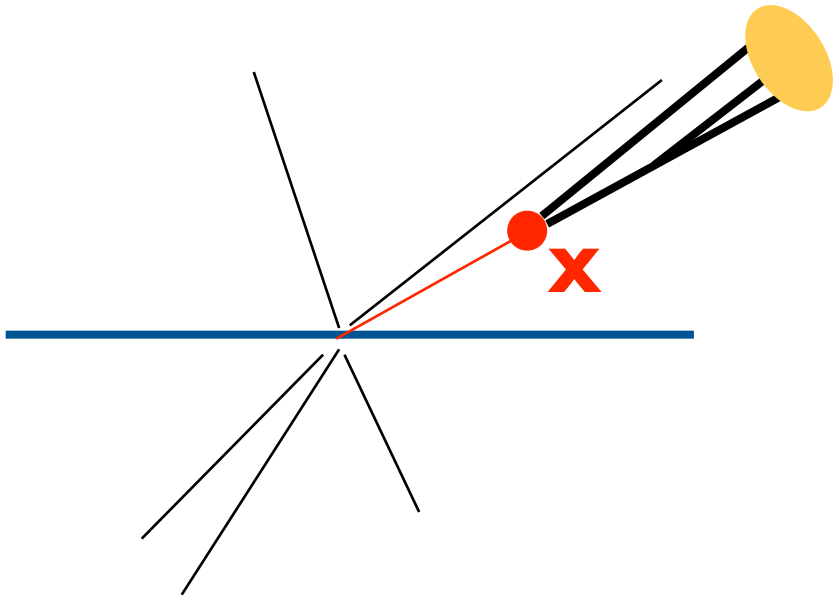


Slide by
Gavin Salam

Why boosted objects



Heavy particle X at **rest**



Boosted heavy particle X

Easy to resolve jets and calculate invariant mass, but signal very likely swamped by background (eg H→bb v. tt →WbWb)

Cross section very much reduced, but acceptance better and some backgrounds smaller/ reducible

Why jet substructure

- The substructure of a jet can be exploited to
  - ***tag*** *a particular structure inside the jet, i.e. a massive particle*
    - *First examples: Higgs (2-prongs decay), top (3-prongs decay)*
  - *remove background contamination from the jet or its components, while keeping the bulk of the perturbative radiation*
    - *First examples: filtering, trimming, pruning*
      - *This operation is often generically denoted as* ***grooming***

Generic tagging/grooming

Fat-jet finding
Often anti-kt, R ≈ 1

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

Tagging step

signal jet candidate, still
partly background-contaminated

grooming step

final candidate, potentially with little background contamination

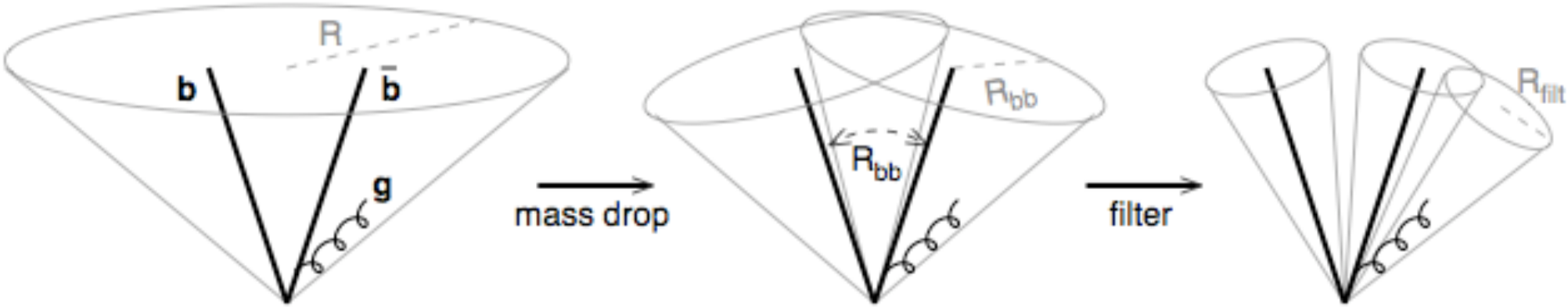
Note that in some taggers (e.g. pruning) the tagging and grooming steps are not explicitly factorised

Also, some tools may actually not follow rigidly this scheme

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

The BDRS tagger/groomer

Butterworth, Davison, Rubin, Salam, 2008

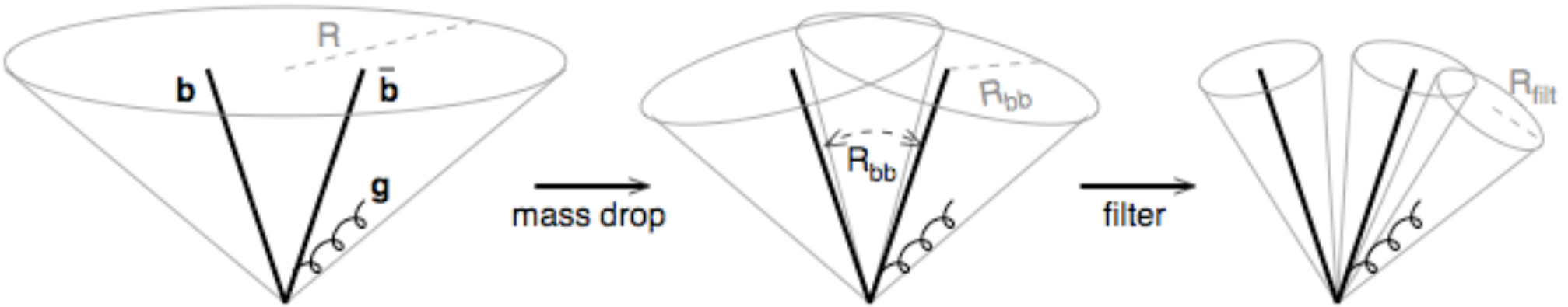


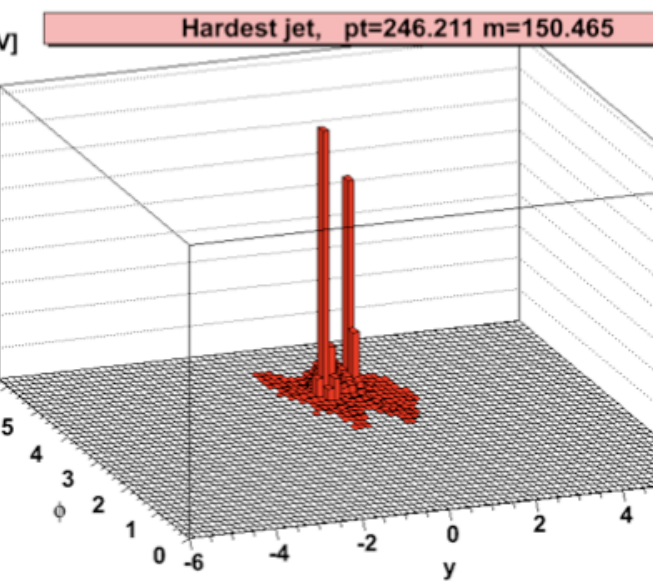
- A two-prongs tagger/groomer for boosted Higgs, which
  - Uses the **Cambridge/Aachen** algorithm
  - Employs a **Mass-Drop** condition, as well as an **asymmetry cut** to find the relevant splitting (i.e. '**tag**' the heavy particle)
  - Includes a post-processing step, using 'filtering' (introduced in the same paper) to clean as much as possible the resulting jets of UE contamination ('**grooming**')

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

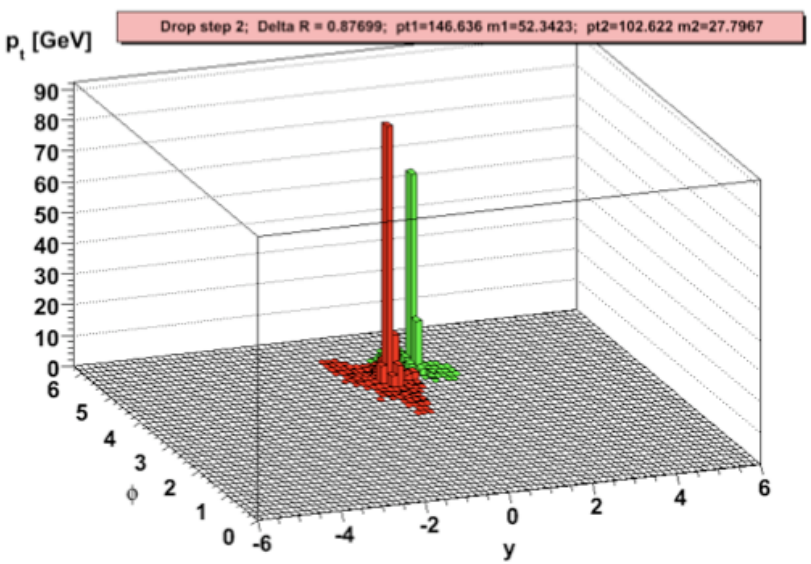
Visualisation of BDRS

Butterworth, Davison, Rubin, Salam, 2008

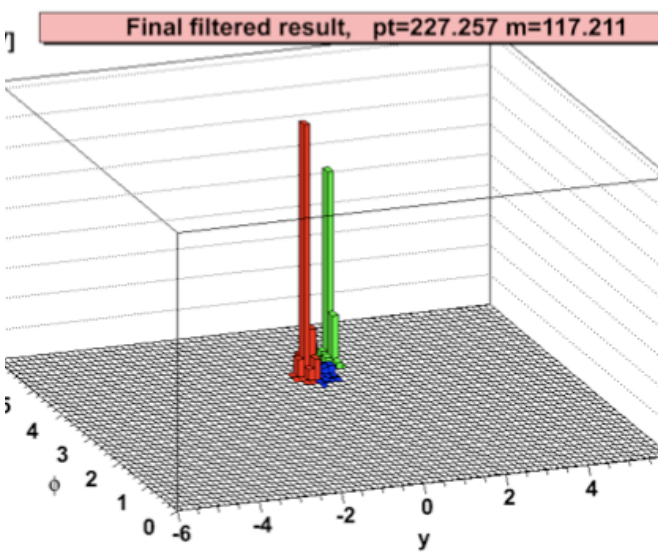




Cluster with a large R



Undo the clustering into subjets, until a large mass drop is observed: tagging step

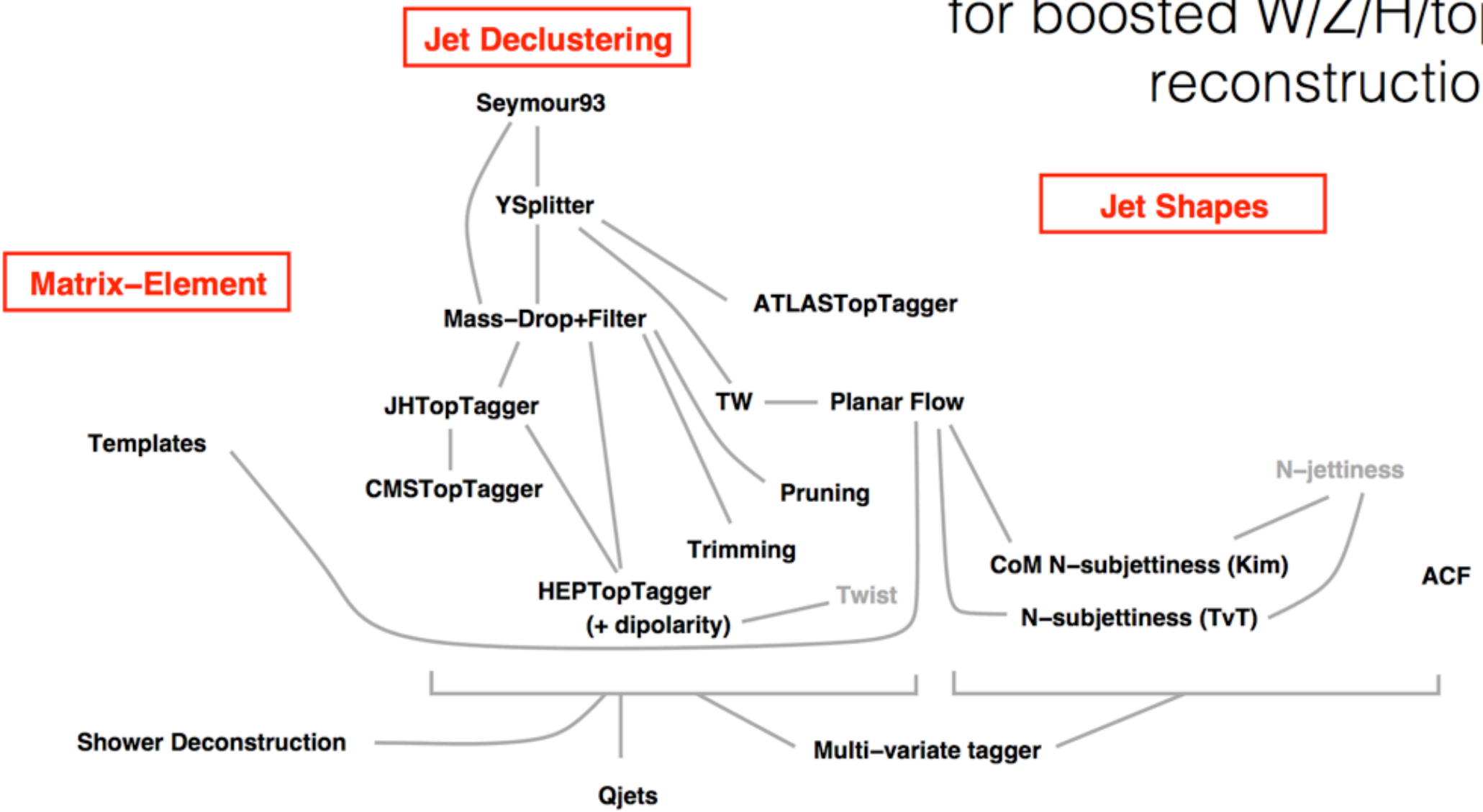


Re-cluster with smaller R, and keep only 3 hardest jets: grooming step

The jet substructure maze

Slide by
G. Salam

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



First taggers/groomers

- **Mass Drop + Filtering** — Butterworth, Davison, Rubin, Salam, 2008
  Decluster with mass drop and asymmetry conditions
  Recluster constituents into subjets at distance scale R_{filt}, retain n_{filt} hardest subjets

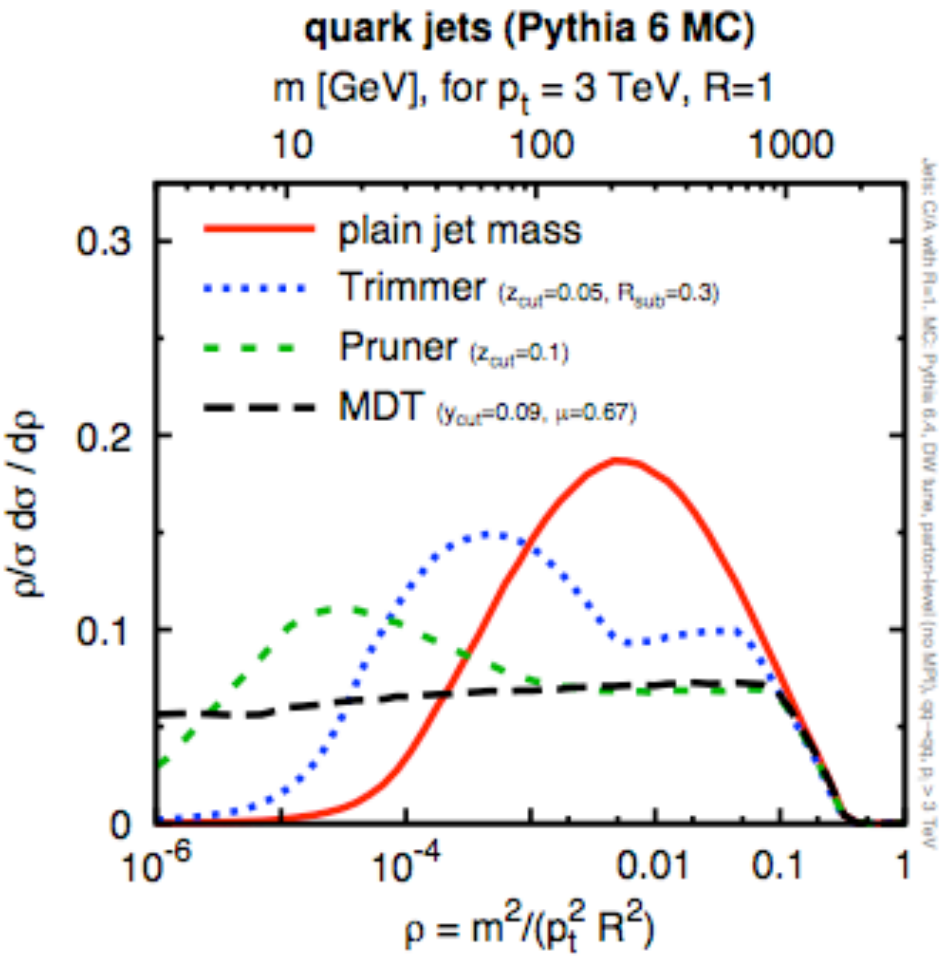
- **Jet 'trimming'** — Krohn, Thaler, Wang, 2009
  Recluster constituents into subjets at distance scale R_{trim}, retain subjets with $p_{t,subjet} > \varepsilon_{trim}\ p_{t,jet}$

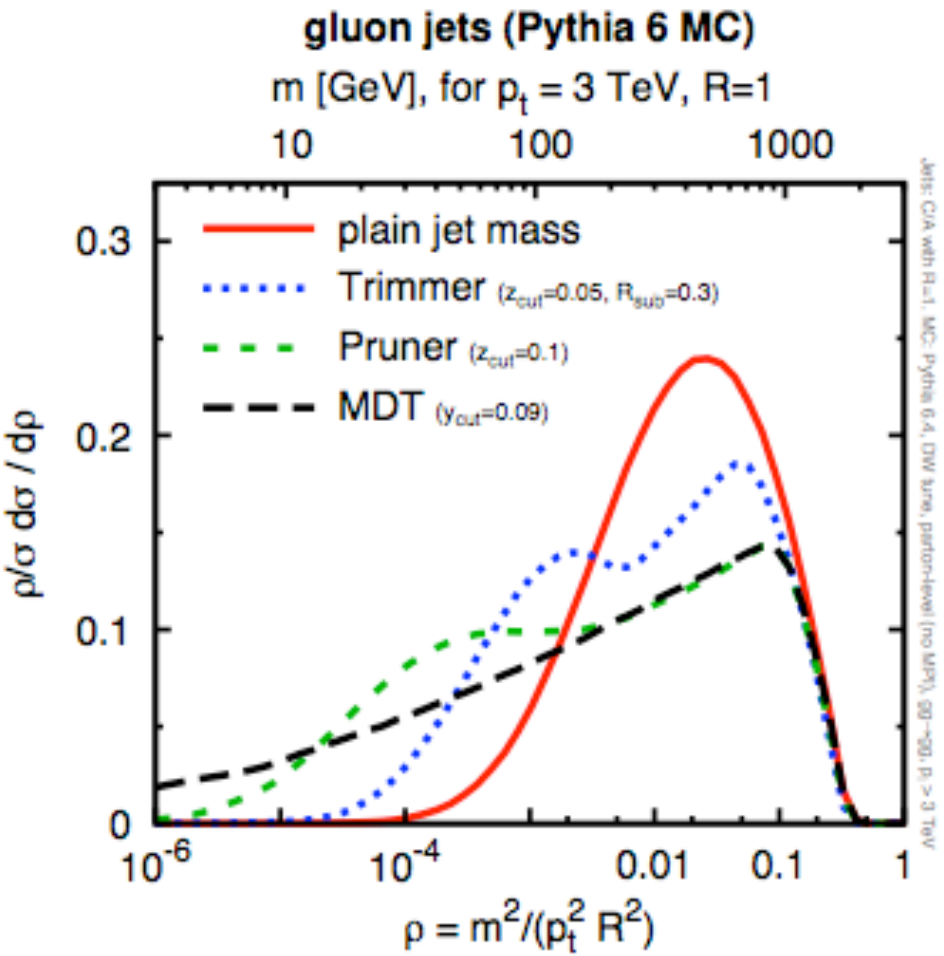
- **Jet 'pruning'** — S. Ellis, Vermilion, Walsh, 2009
  While building up the jet, discard softer subjets when $\Delta R > R_{prune}$ and $\min(p_{t1}, p_{t2}) < \varepsilon_{prune}\ (p_{t1}+p_{t2})$

Aim: limit contamination from QCD background while retaining bulk of perturbative radiation

Robustness of substructure tools

Dasgupta, Fregoso, Marzani, Salam, 2008



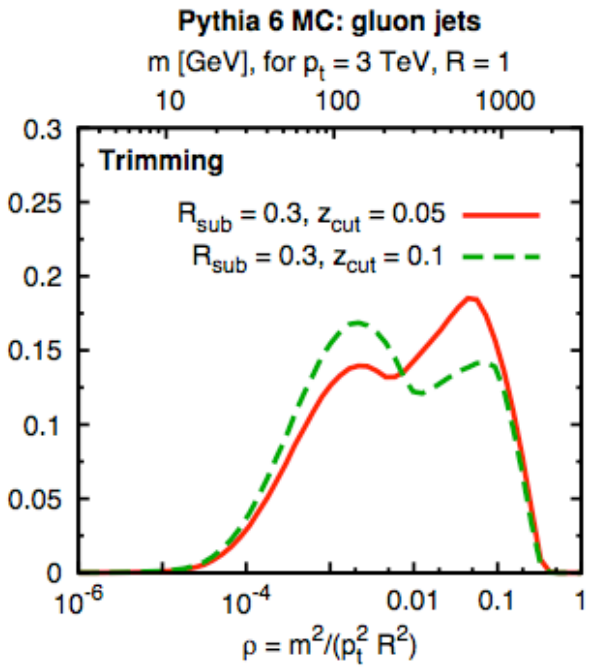


Tools that are considered (or can be seen in Monte Carlo tests) to behave 'similarly' could cease to do so in different parameter regions

Analytic calculations of jet substructure

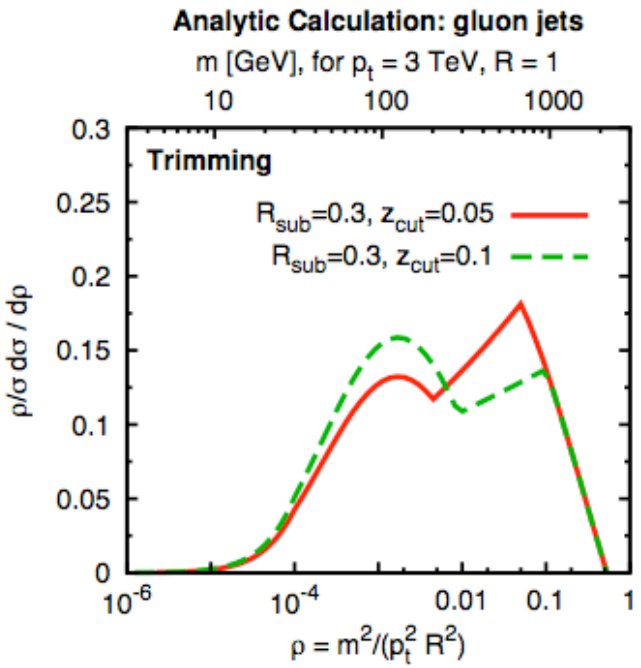
Dasgupta, Fregoso, Marzani, Salam, 2013

Monte Carlo



Analytic

(resummed pQCD)



- Analytical understanding of ‘kinks’ in distributions
- Check of Monte Carlo predictions
- Other analytical investigations: Rubin 2010 (filtering), Walsh, Zuberi 2011 (jet substructure with SCET), Feige Schwartz, Stewart, Thaler 2012 (N-subjettiness), Dasgupta, Marzani, Powling 2013 (groomed jet mass), ...

$$\frac{1}{\sigma}\frac{d\sigma}{dm^2}^{(\text{trim, LO})} = \frac{\alpha_s C_F}{\pi}\int_0^1 dz\, p_{gq}(z)\int\frac{d\theta^2}{\theta^2}\,\delta\left(m^2 - z(1-z)p_t^2\theta^2\right)\times$$
$$\times\left[\Theta\left(z-z_{\text{cut}}\right)\Theta\left(1-z-z_{\text{cut}}\right)\Theta(\theta^2-R_{\text{sub}}^2)+\Theta(R_{\text{sub}}^2-\theta^2)\right]\Theta\left(R^2-\theta^2\right)$$

Alternatives to (standard) substructure

- If what we are interested in is the structure of the constituents of a jet, the "jet" itself is not the most important feature.
- A different algorithm, or simply the study of the constituents in a certain patch will also do. Selected alternatives are:
  - Use of jet-shapes to characterise certain features
    - e.g. *N-subjettiness*: how many subjets a jets appears to have

      Thaler, van Tilburg, 2011
  - Alternative ways of clustering
    - e.g. *Qjets*: the clustering history not deterministic, but controlled by random probabilities of merging. Can be combined with, e.g. pruning

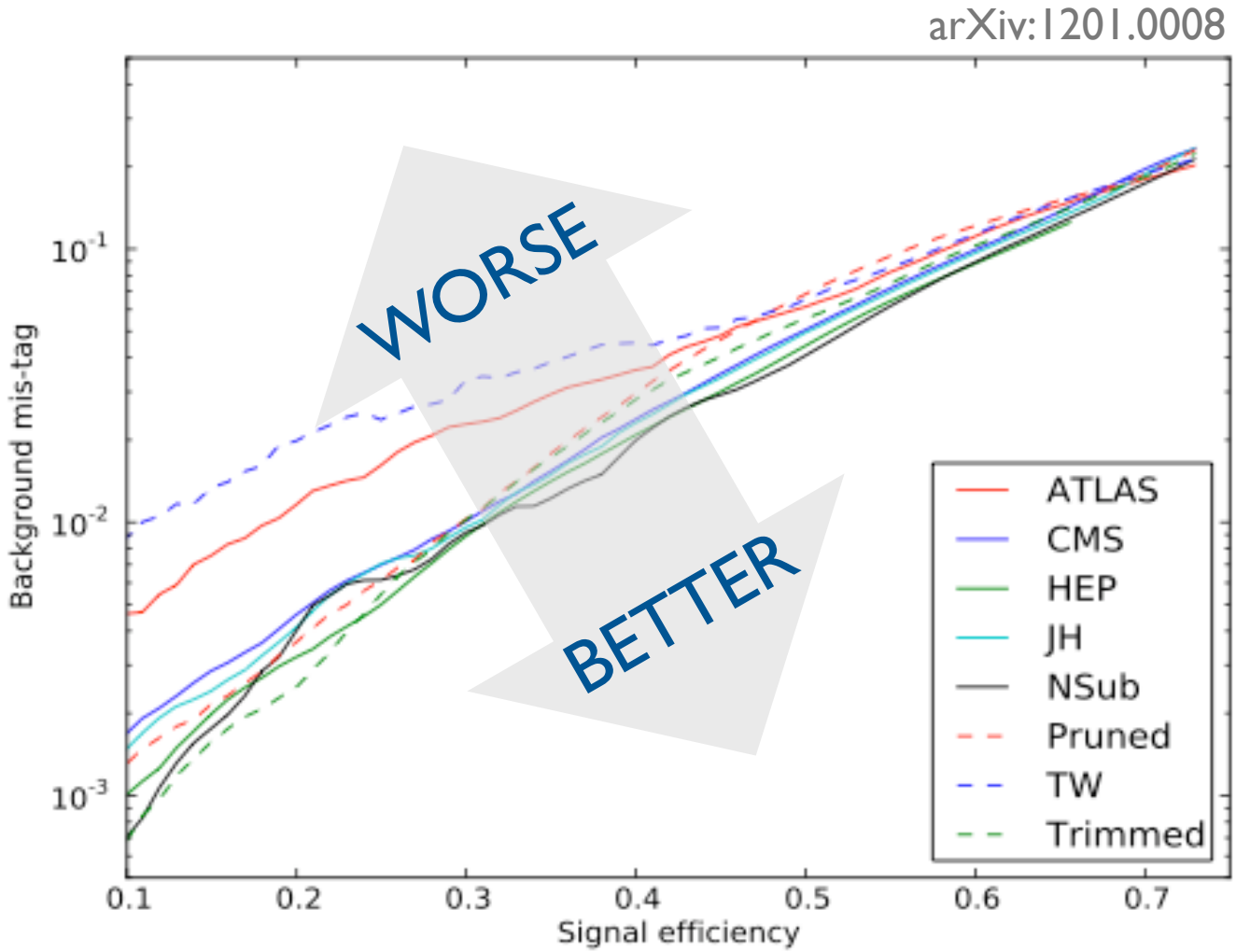
      Ellis, Hornig, Roy, Krohn, Schwartz, 2012
  - Use information from matrix element
    - e.g. *shower deconstruction*: use analytic shower calculations to estimate probability that a certain configuration comes from signal or from background

      Soper, Spannowsky, 2011
  - Use event shapes mimicking jet properties
    - e.g. *JetsWithoutJets*, mimicking trimming

      Bertolini, Chen, Thaler, 2013

Performance comparisons

Many of the substructure-based tools have been systematically compared, for instance in the context of the proceedings of the BOOST conferences

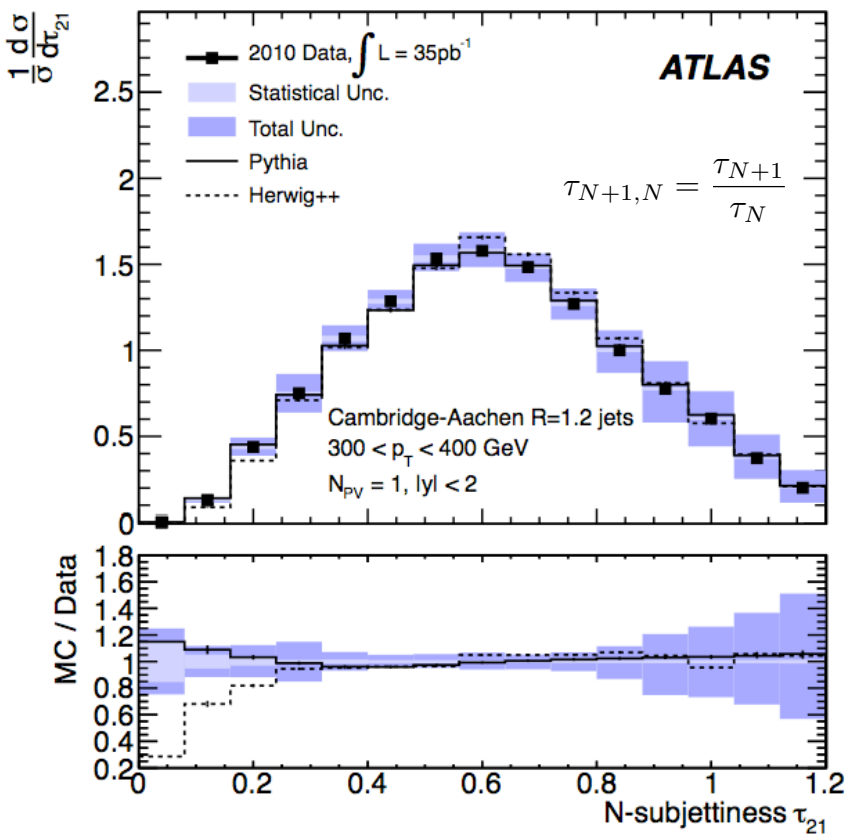


Efficiency-mistag curves for many top taggers.
Improvements upon earlier tools have become hard to obtain

Experimental validation

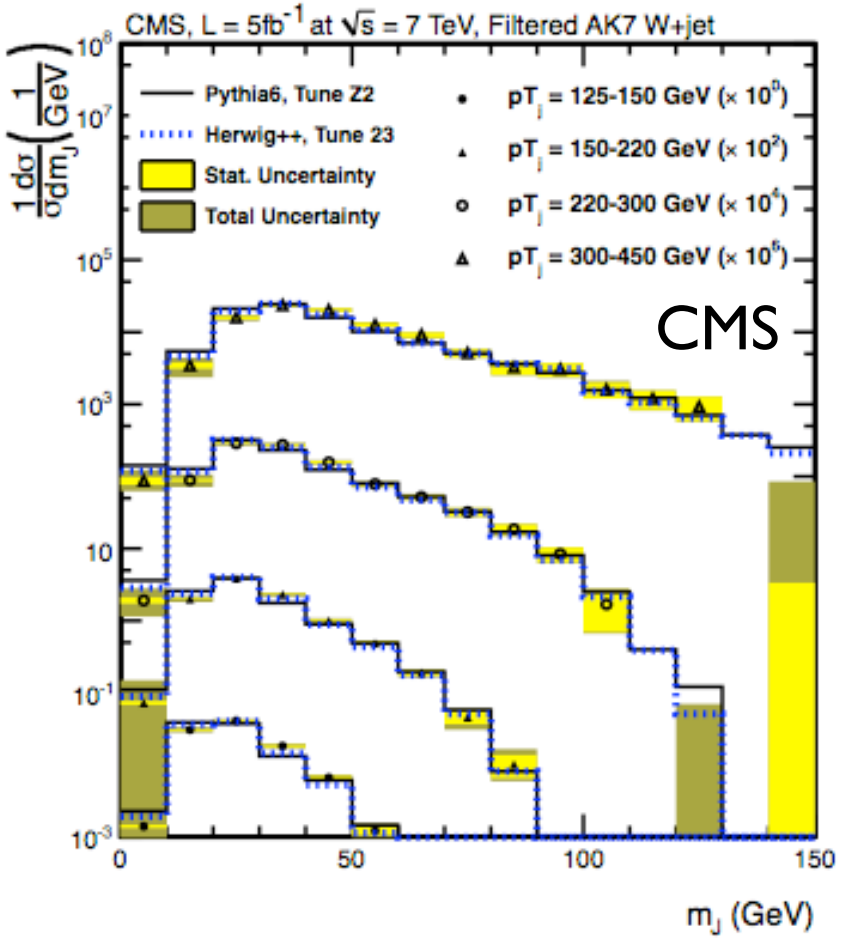
ATLAS and CMS have extensively validated the new grooming techniques, and also produced the first results using them

arXiv:1203.4606



N-subjettiness

arXiv:1303.4811



Filtered jet mass

Jet substructure in experiments

Slide by G. Salam

Extensive experimental work

Last 12 ATLAS & CMS preprints citing jet substructure work

Jet Cross-Section Measurements In CMS
CMS Collaboration
Inspire. arXiv:1306.6604 (ps, pdf). Int.J.Mod.Phys. A28 (2013) 1330030.

Performance of jet substructure techniques for large-R jets in proton-proton collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector
ATLAS Collaboration
Inspire. arXiv:1306.4945 (ps, pdf). JHEP 1309 (2013) 076. 16 cites [co]

Measurement of jet shapes in top pair events at $\sqrt{s}$ = 7 TeV using the ATLAS detector
ATLAS Collaboration
Inspire. arXiv:1307.5749 (ps, pdf).

Searches for New Physics in Multijet Final States
for the CMS Collaboration
Inspire. arXiv:1307.2518 (ps, pdf).

Search for Single and Pair-Production of Dijet Resonances with the CMS Detector
CMS Collaboration
Inspire. arXiv:1307.1400 (ps, pdf). J.Phys.Conf.Ser. 455 (2013) 012034. 1 cites [co]

Search for dark matter in events with a hadronically decaying W or Z boson and missing transverse momentum in pp collisions at $\sqrt{s}$=8 TeV with the ATLAS detector
ATLAS Collaboration
Inspire. arXiv:1309.4017 (ps, pdf). 5 cites [co]

Searches for anomalous ttbar production in pp collisions at $\sqrt{s}$=8 TeV
CMS Collaboration
Inspire. arXiv:1309.2030 (ps, pdf). 6 cites [co]

Search for heavy resonances decaying to top quarks
for the CMS Collaboration
Inspire. arXiv:1310.8183 (ps, pdf).

Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks
CMS Collaboration
Inspire. arXiv:1310.3687 (ps, pdf). 3 cites [co]

Search for the SM Higgs Boson Produced in Association with a Vector Boson and Decaying to Bottom Quarks
for the CMS Collaboration
Inspire. arXiv:1310.3551 (ps, pdf).

Inclusive search for a vector-like T quark with charge 2/3 in pp collisions at $\sqrt{s}$=8 TeV
CMS Collaboration
Inspire. arXiv:1311.7667 (ps, pdf).

Search for top-quark partners with charge 5/3 in the same-sign dilepton final state
CMS Collaboration
Inspire. arXiv:1312.2391 (ps, pdf).

Jet substructure in BSM

Certainly incomplete list of phenomenology papers

Discovering baryon-number violating neutralino decays at the LHC
Jonathan M. Butterworth, John R. Ellis, Are R. Raklev, Gavin P. Salam, **arXiv:0906.0728**

Stop Reconstruction with Tagged Tops
Tilman Plehn, Michael Spannowsky, Michihisa Takeuchi, Dirk Zerwas, **arXiv:1006.2833**

Boosted Semileptonic Tops in Stop Decays
Tilman Plehn, Michael Spannowsky, Michihisa Takeuchi, **arXiv:1102.0557**

Scale-invariant resonance tagging in multijet events and new physics in Higgs pair production
Maxime Gouzevitch, Alexandra Oliveira, Juan Rojo, Rogerio Rosenfeld, Gavin P. Salam, Veronica Sanz, **arXiv:1303.6636**

Pulling Out All the Stops: Searching for RPV SUSY with Stop-Jets
Yang Bai, Andrey Katz, Brock Tweedie, **arXiv:1309.6631**

Measuring boosted tops in semi-leptonic ttbar events for the standard model and beyond
Mihailo Backovic, Ofir Gabizon, Jose Juknevich, Gilad Perez, Yotam Soreq, **arXiv:1311.2962**

Reconstructing singly produced top partners in decays to Wb
Nicolas Gutierrez Ortiz, James Ferrando, Deepak Kar, Michael Spannowsky, **arXiv:1403.7490**

Hadronic b′ search at the LHC with top and W taggers
Shuo Yang, Ji Jiang, Qi-Shu Yan, Xiaoran Zhao, **arXiv:1405.2514**

Summary

- "Plain" jet clustering is mainstream
  - All main algorithms implemented in a standardized, efficient and public way, used by all LHC collaborations and phenomenologists

- Jet substructure techniques also largely mature
  - tens of different approaches:
    - risk of fragmentation and poor readability
    - use of standard, public tools rather than *ad hoc* ones helps consolidation of the field
  - are further improvements in performance becoming hard to achieve?
  - many approaches only tested at MC level: robustness?
    - analytical control should become a must

Substructure TODO

- There may still be room for further improvement in jet substructure techniques

- To avoid fragmenting the field, and make progress efficient, we should
  - Introduce techniques motivated by analytical arguments, not simply MC testing
  - Ensure that they enjoy a **good analytical calculability**
    - very little reason to introduce today a novel substructure technique that does not enjoy a decent calculability, unless HUGE improvement can be shown (and still, it should be justifiable and robust)
  - Provide a **public implementation** (e.g. in the FastJet contrib project, http://fastjet.hepforge.org/contrib, public repository for third-party contributions)

Soft Drop declustering

Larkoski, Marzani, Soyez, Thaler, van Tilburg, 2014

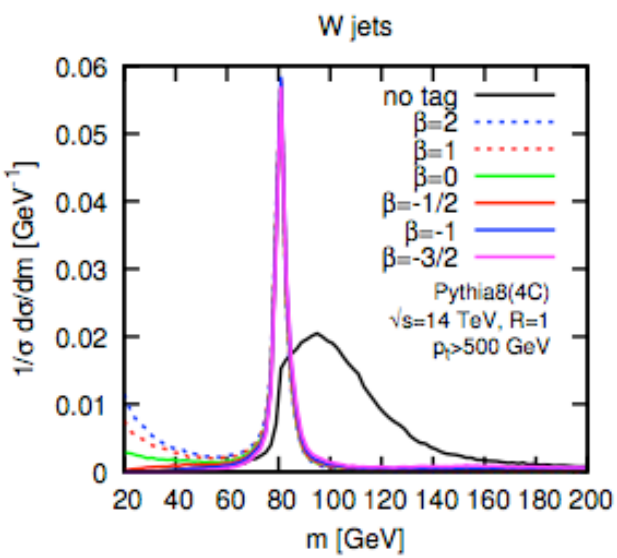
Decluster and drop softer constituent unless

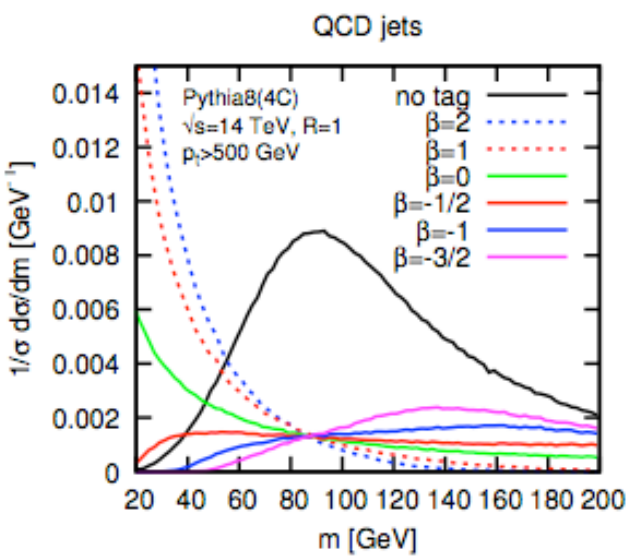
$$\text{Soft Drop Condition:} \quad \frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{\text{cut}} \left(\frac{\Delta R_{12}}{R_0} \right)^{\beta}$$

i.e. remove wide-angle soft radiation from a jet

The paper contains

- ✓ analytical calculations and comparisons to Monte Carlos
- ✓ study of effect of non-perturbative corrections
- ✓ performance studies

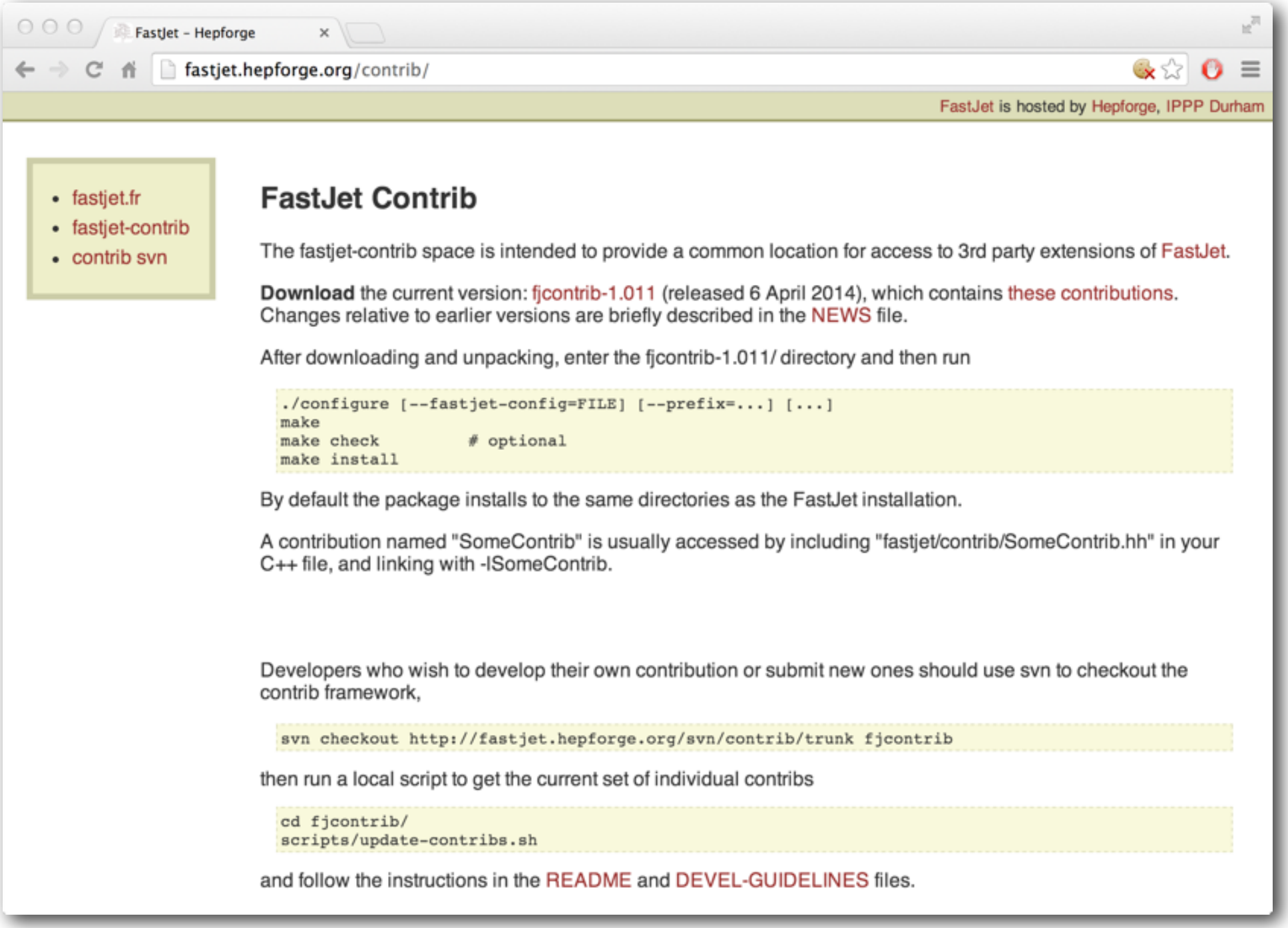




Example of SoftDrop performance when used as a boosted W tagger

Backup

FastJet Contrib



FastJet Contrib