MadGraph School
Jiao Tong University - Shanghai
November 2015

Jets

Lecture 1: jet algorithms

Lecture 2: jet substructure

Matteo Cacciari
LPTHE Paris

Glossary

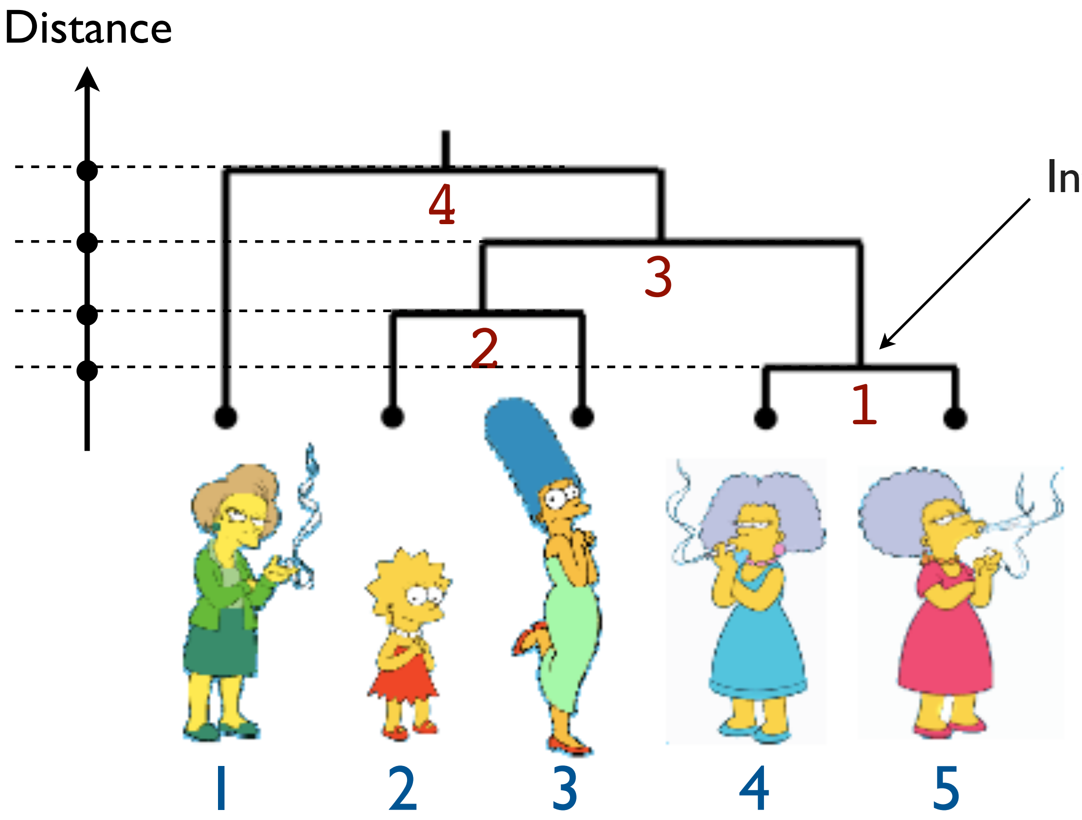
What	i.e.	When	Ref.
AKT	Anti-kt algorithm	2008	0802.1189
CA	Cambridge/Aachen algorithm	1999	9907280
BDRS	mass-drop tagger, includes filtering	2008	0802.2470
trimmed	Trimming, tagger/groomer	2009	0912.1342
pruned	Pruning, tagger/groomer	2009	0903.5081
HTT	HepTopTagger	2009	0910.5472
N-subjettiness	jet shape function, used in tagging	2010	1011.2268
WTA	Winner-Take-All (recombination scheme)	2013	1310.7584
one-pass	choice of axis for N-subjettiness	2010	
JVT	Jet Vertex Tagger (used in pileup subtr.)	2014	
ρ	background density (used in pileup subtr.)	2007	0707.1378
D2	jet shape function, used in tagging	2014	1409.6298
PUPPI	particle-by-particle pileup subtr.	2014	1407.6013
Soft Drop	tagger/groomer	2014	1402.2657

Outline

- Algorithms, speed
- Infrared and collinear safety
- Background (pileup)
- Substructure

Dendrogram

Used to represent graphically the sequence of clustering steps in a sequential recombination algorithm

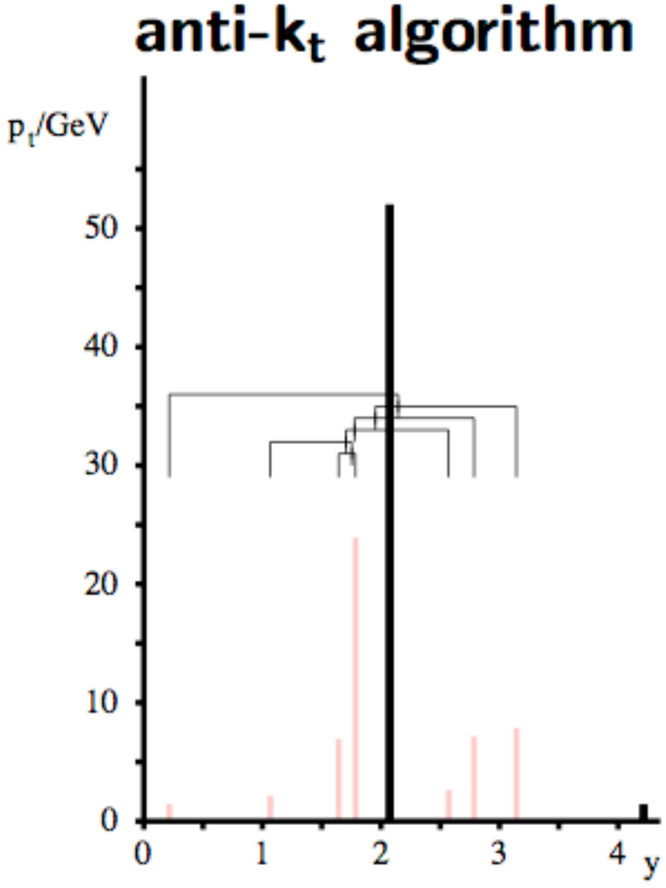


Distance between two objects is given by the **height** of the lowest internal node that they share.

Order of clustering here is 1,2,3,4

The **clustering sequence** is 4-5 (1), 2-3 (2), 23-45 (3), 1-2345 (4)

Hierarchical substructure

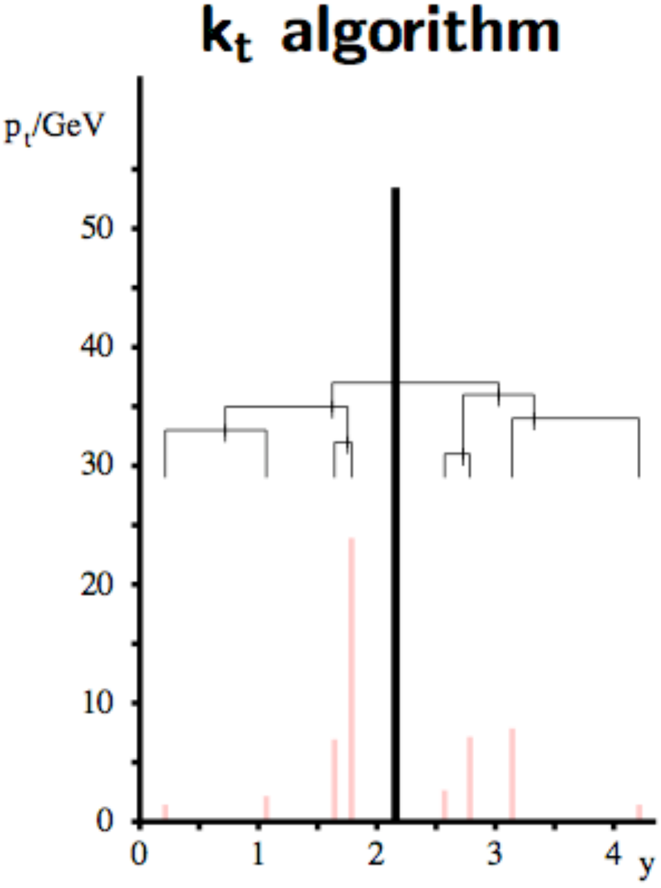


Anti-k_t distance measure

$$d_{ij} = \min\left(\frac{1}{k_{ti}^2}, \frac{1}{k_{tj}^2}\right) \frac{\Delta y^2 + \Delta\phi^2}{R^2}$$

Cluster by merging
to the **hardest/closest** particle

Hierarchical substructure

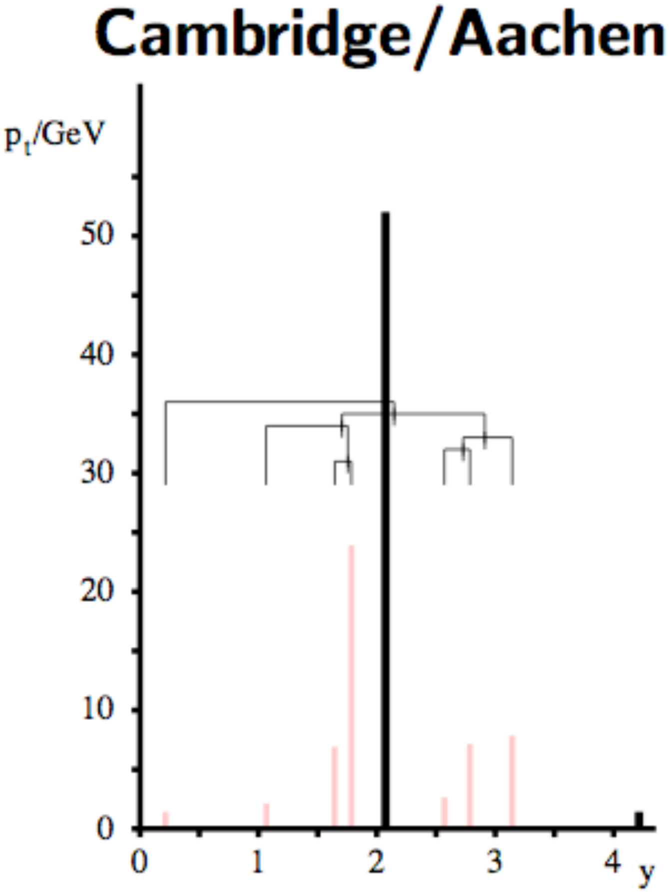


$\mathrm{k_t}$ distance measure

$$d_{ij} = \min(k_{ti}^2, k_{tj}^2)\frac{\Delta y^2 + \Delta\phi^2}{R^2}$$

Cluster by merging
the **softest/closest** particles

Hierarchical substructure



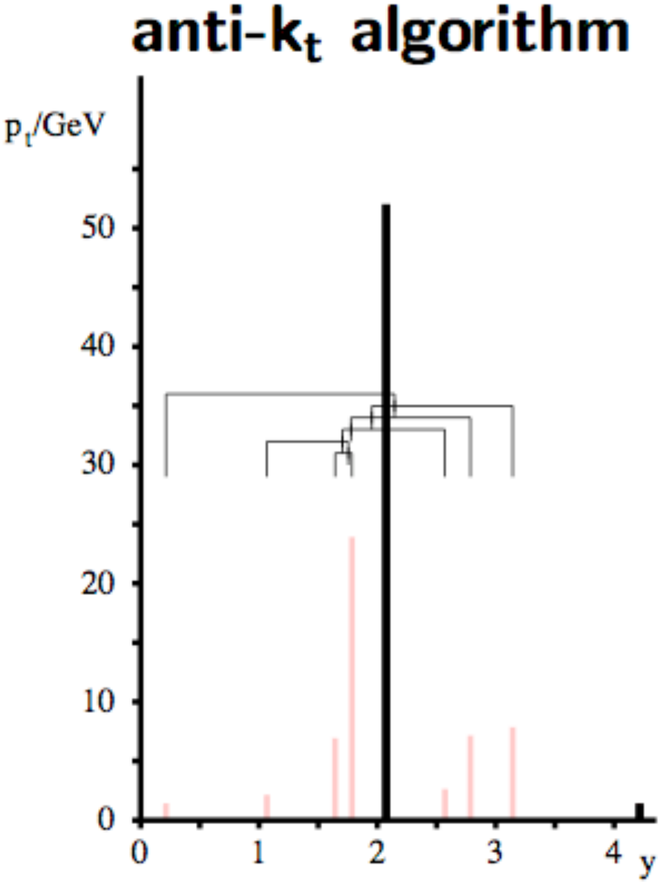
C/A distance measure

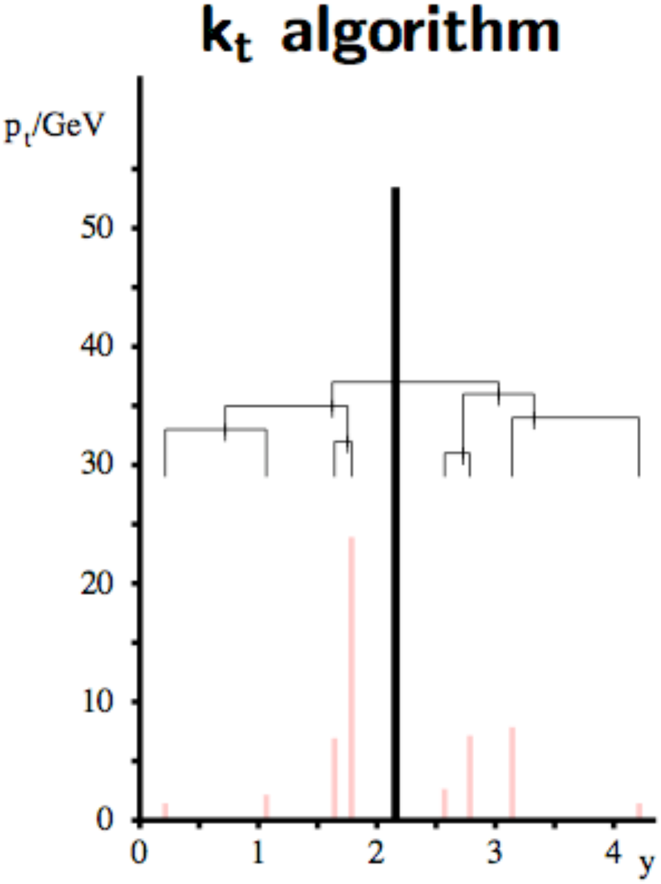
$$d_{ij} = \frac{\Delta y^2 + \Delta\phi^2}{R^2}$$

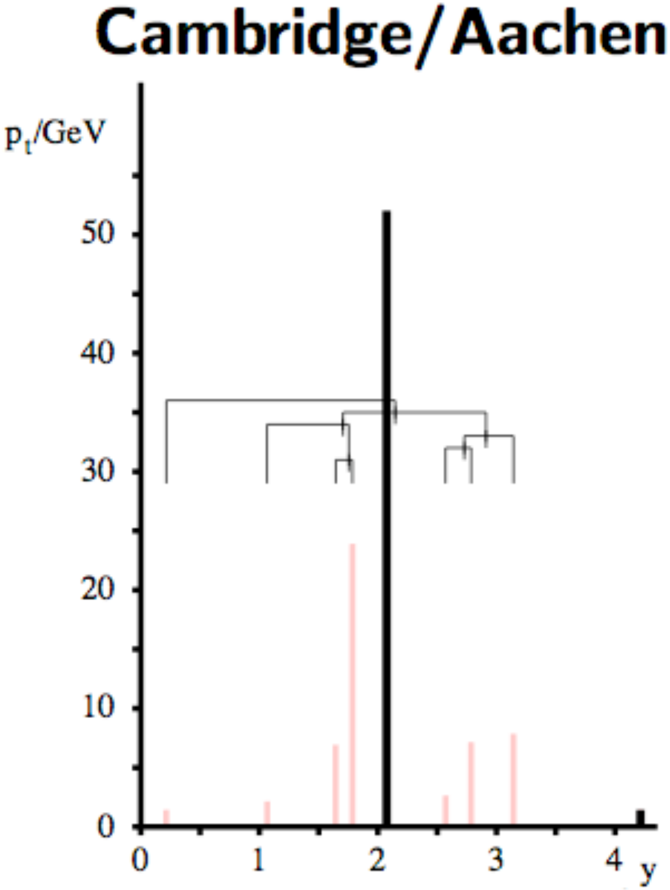
Cluster by merging the **closest** particles

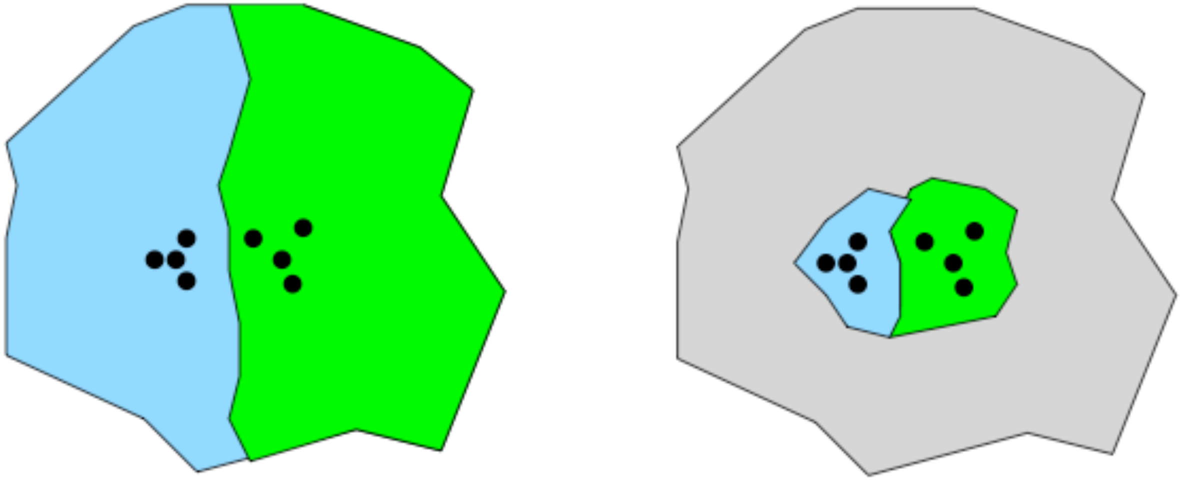
Slide by
Gavin Salam

Hierarchical substructure









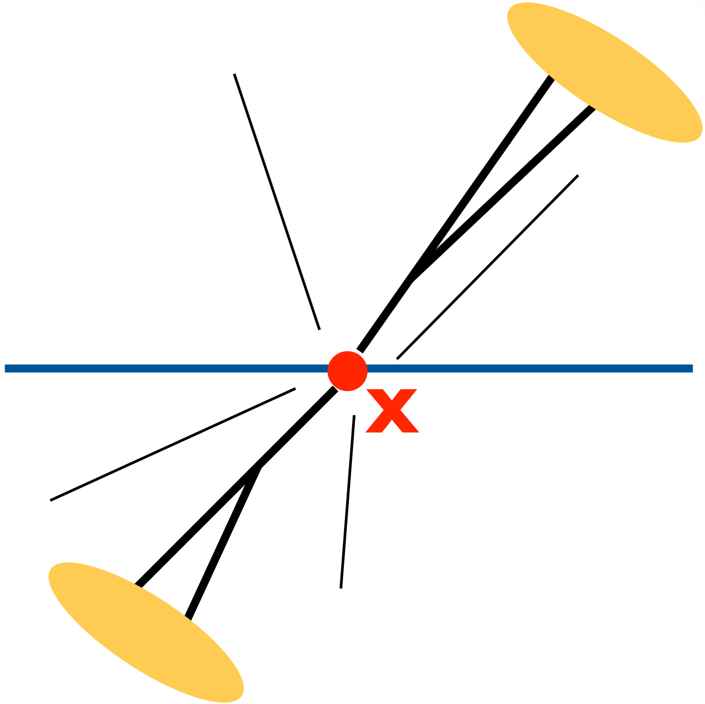
Undo the last clustering step(s) →

The IRC safe algorithms

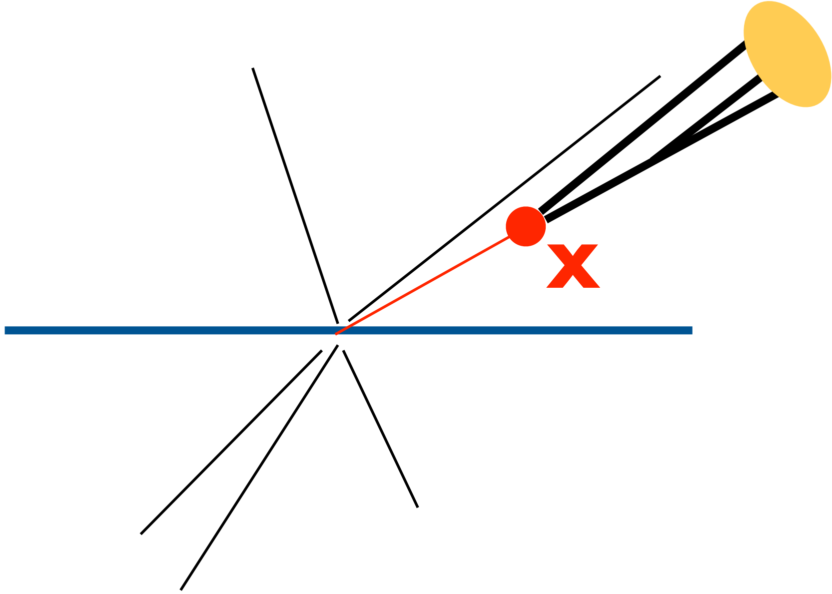
	Speed	Regularity	UE contamination	Backreaction	Hierarchical substructure
k_t	☺ ☺ ☺	☂	☂☂	☁☁	☺ ☺
Cambridge /Aachen	☺ ☺ ☺	☂	☂	☁☁	☺ ☺ ☺
anti-k_t	☺ ☺ ☺	☺ ☺	☁/ ☺	☺ ☺	✘
SISCone	☺	☁	☺ ☺	☁	✘

Array of tools with different characteristics.
Pick the right one for the job

Why boosted objects



Heavy particle X at **rest**



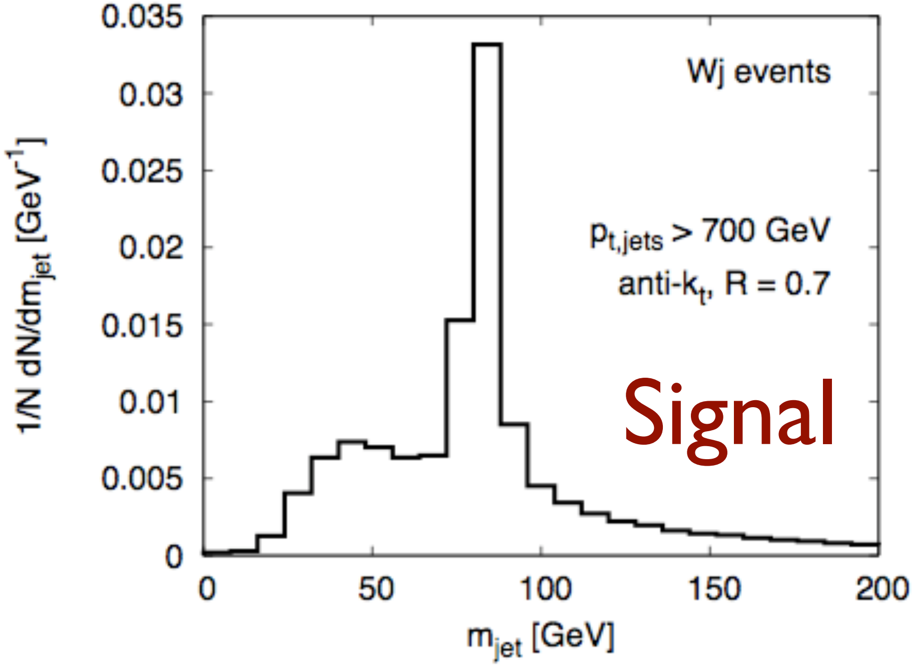
Boosted heavy particle X

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

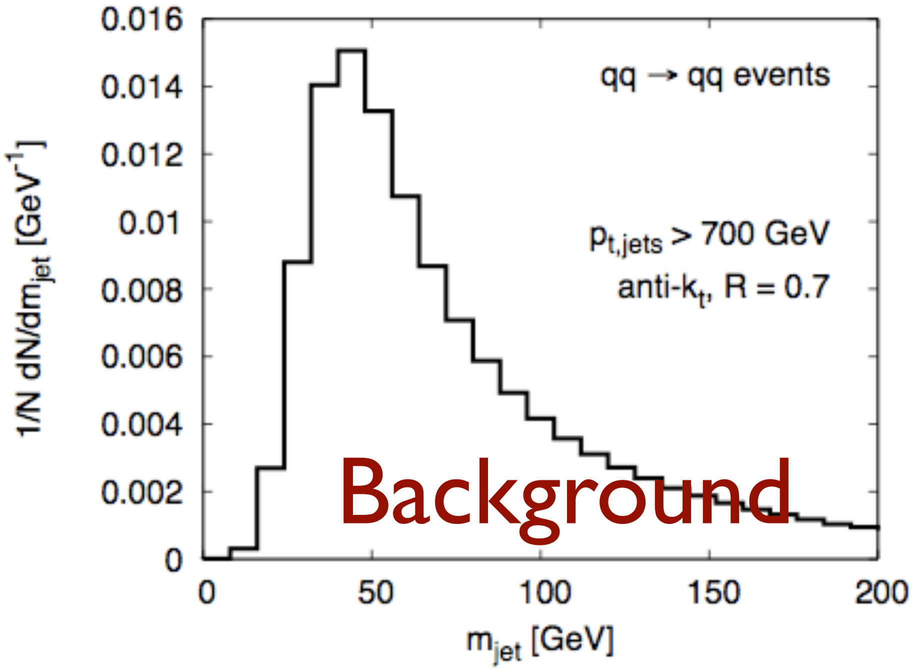
Cross section very much reduced, but acceptance better and some backgrounds smaller/reducible

Mass of a single jet

G. Salam



A heavy object decaying into a single jet naturally gives it a mass...

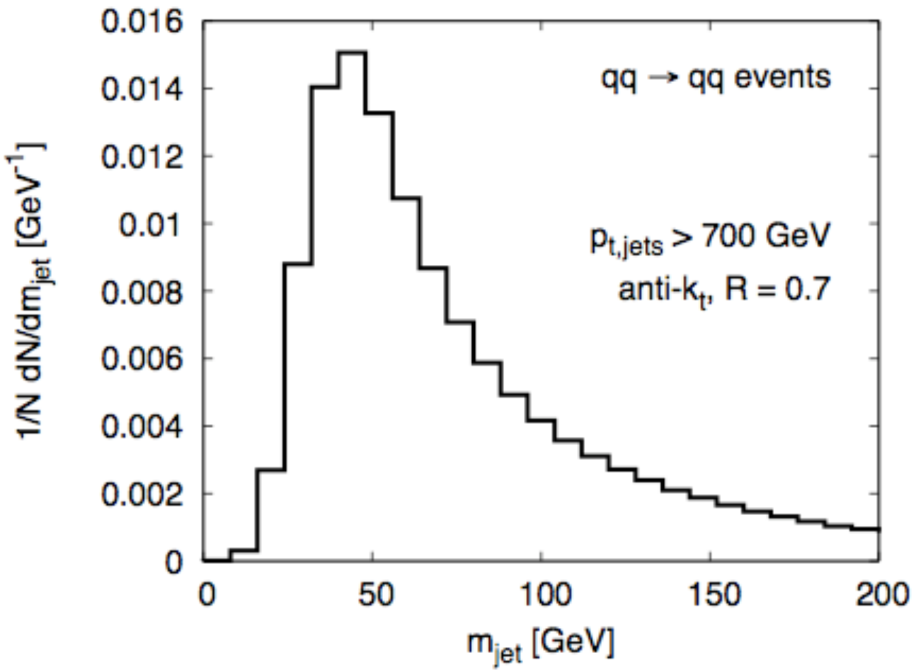


... but pure QCD jets can be massive too:

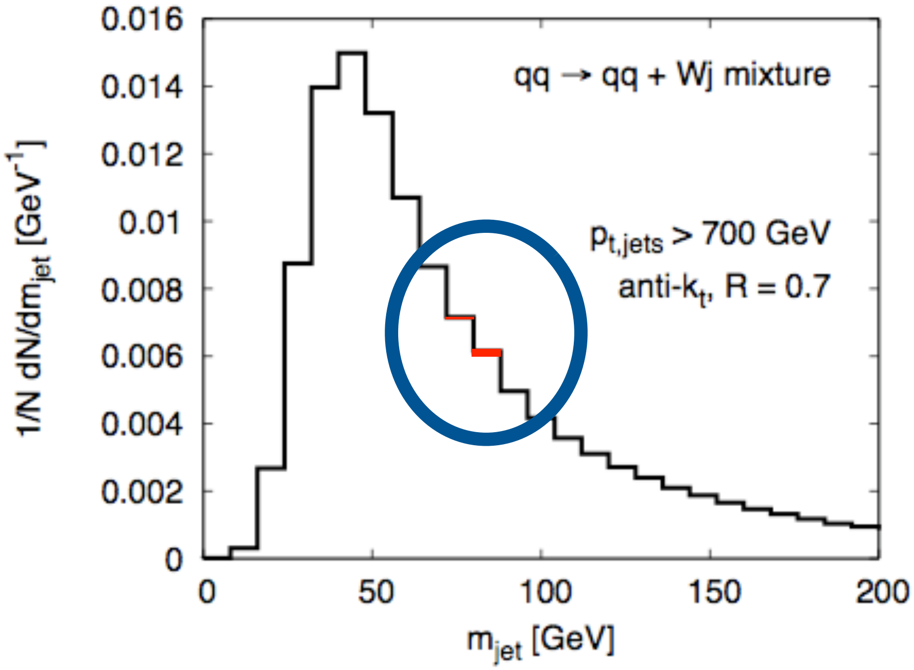
$$\frac{dN}{d\ln m} \sim \alpha_s \ln\frac{p_t R}{m} \times \text{Sudakov}$$

Mass of a single jet

Summing ‘signal’ and ‘background’ (with appropriate cross sections) shows how much the background dominates



Background only



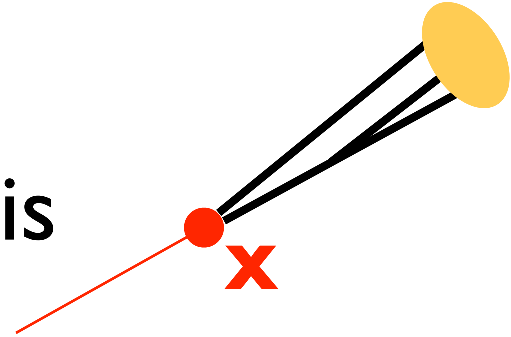
Signal + background

Practically identical

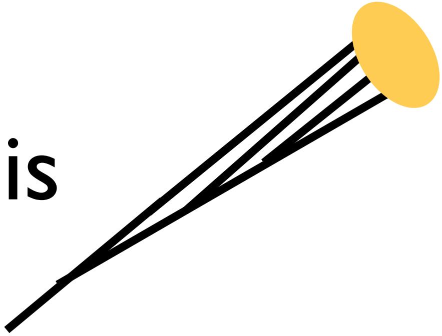
This means that one can’t rely on the invariant mass only.
An appropriate strategy must be found to reduce the background and enhance the signal

Tagging

How to tell this from this ?



Decay of a heavy (boosted) object



Light parton fragmentation

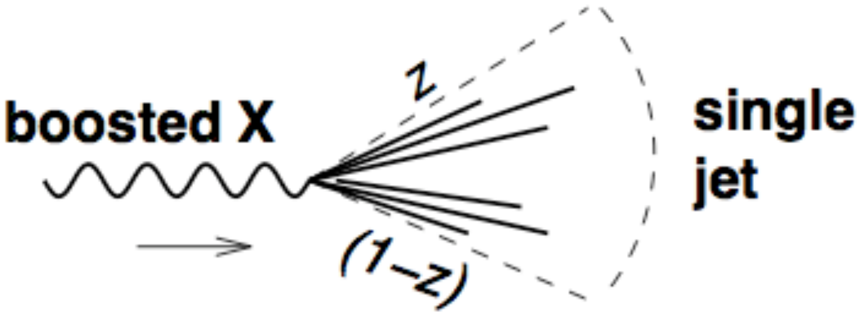
Tagging and Grooming

- 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-prong decay), top (3-prong decay)*
  - *remove background contamination from the jet or its components, while keeping the bulk of the perturbative radiation (often generically denoted as* ***grooming****)*
    - *First examples: filtering, trimming, pruning*

Why substructure

Scales: m ~ 100 GeV, p_t ~ 500 GeV

(e.g. electroweak particle from decay of ~ 1TeV BSM particle)



$$R \gtrsim \frac{m}{p_t} \frac{1}{\sqrt{z(1-z)}}$$

- need small R (< $2m/p_t$ ~ 0.4) to resolve two prongs
- need large R (>~ $3m/p_t$ ~ 0.6) to cluster into a single jet

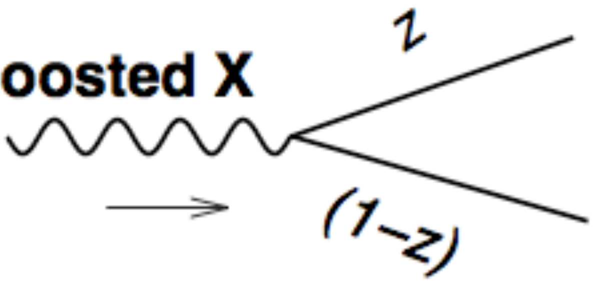
Possible strategies

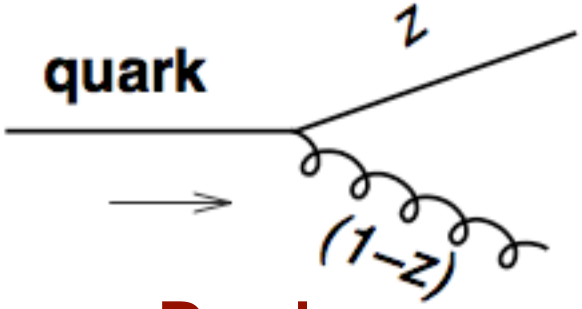
- Use large R, get a single jet : background large
- Use small R, resolve the jets : what is the right scale?
  - Also: small jets lead to huge combinatorial issues

Let an algorithm find the 'right' substructure

QCD v. heavy decay

A possible approach for reducing the QCD background is to identify the two prongs of the heavy particle decay, and put a cut on their momentum fraction





Signal:

$$P(z) \sim 1$$

Will split mainly
symmetrically

Background:

$$P(z) \sim \frac{1+z^2}{1-z} \qquad P(z) \sim \frac{1+(1-z)^2}{z}$$

Will split mainly
asymmetrically

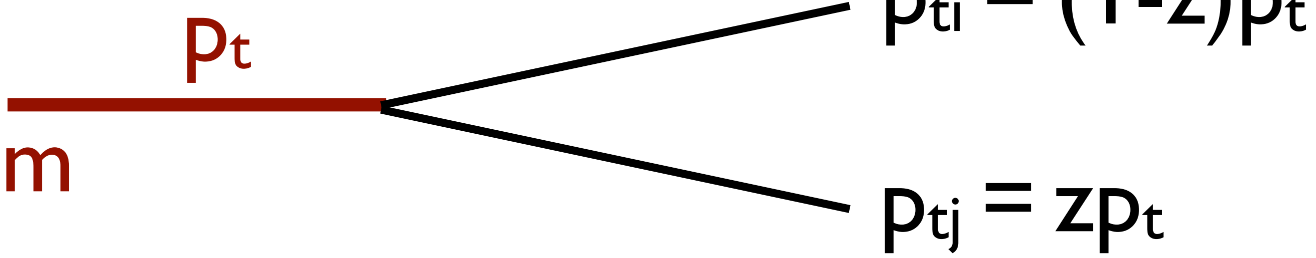
Potential tagger: asymmetric splitting

Possibly
implemented
via a cut on

$$y = min(p_{ti}^2, p_{tj}^2)\frac{\Delta R_{ij}^2}{m^2} \simeq \frac{min(p_{ti}, p_{tj})}{max(p_{ti}, p_{tj})}$$

Splittings and distances

Quasi-collinear splitting ($p_{tj} < p_{ti}$)



Invariant mass:

$$m^2 \simeq p_{ti}p_{tj}\Delta R_{ij}^2 = (1-z)zp_t^2\Delta R_{ij}^2$$

k_t distance:

$$d_{ij} \overset{(p_{tj}<p_{ti})}{=} z^2 p_t^2 \Delta R_{ij}^2 \simeq \frac{z}{1-z} m^2$$

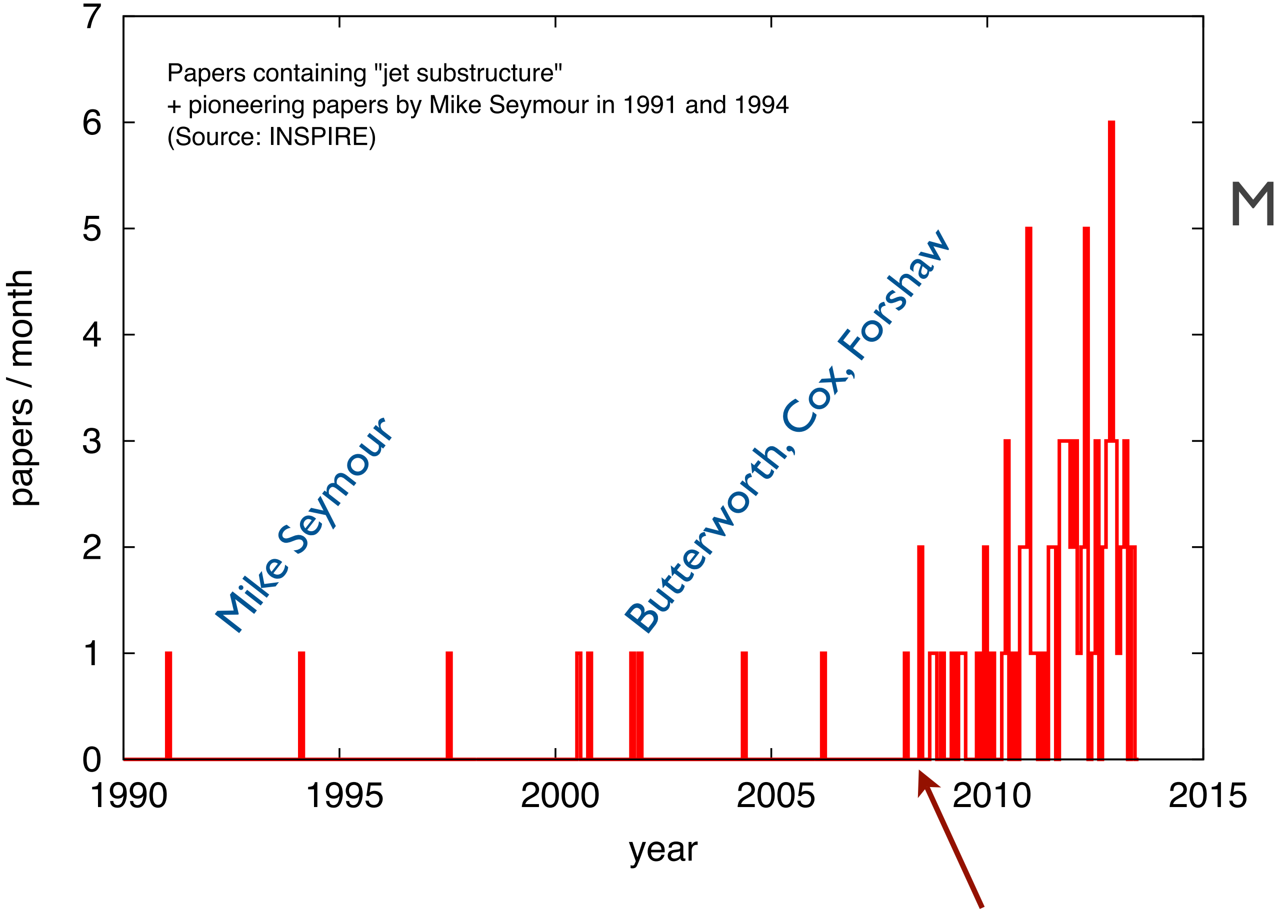
For a given mass, the **background** will have smaller distance d_{ij} than the signal, i.e. it will tend to **cluster earlier** in the k_t algorithm

Potential tagger: last clustering in k_t algorithm

This is where the hierarchy of the k_t algorithm becomes relevant.
QCD radiation is clustered first, and only at the end the symmetric, large-angle splittings due to decays are reclustered

‘Jet substructure’ papers in INSPIRE

Number of papers containing the words ‘jet substructure’



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

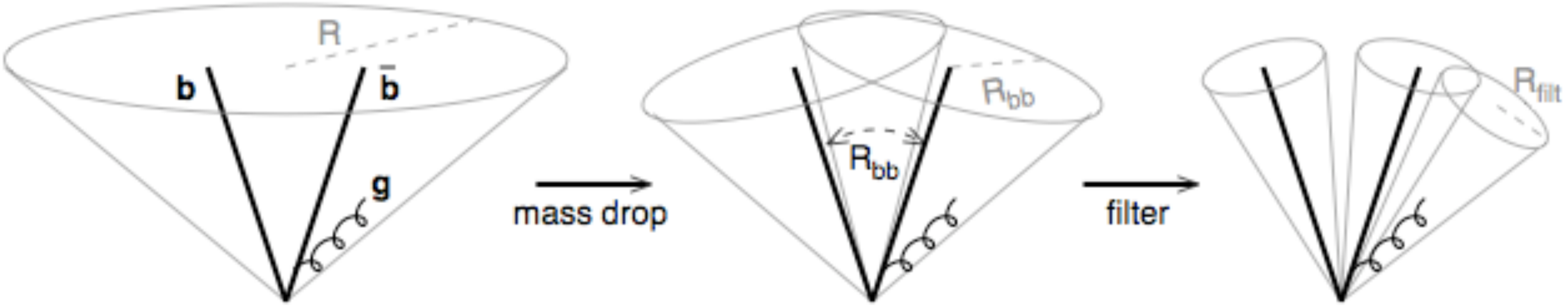
Pioneered by M. Seymour in the early ‘90s, rebooted by BDRS paper

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]**

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

The BDRS tagger/groomer

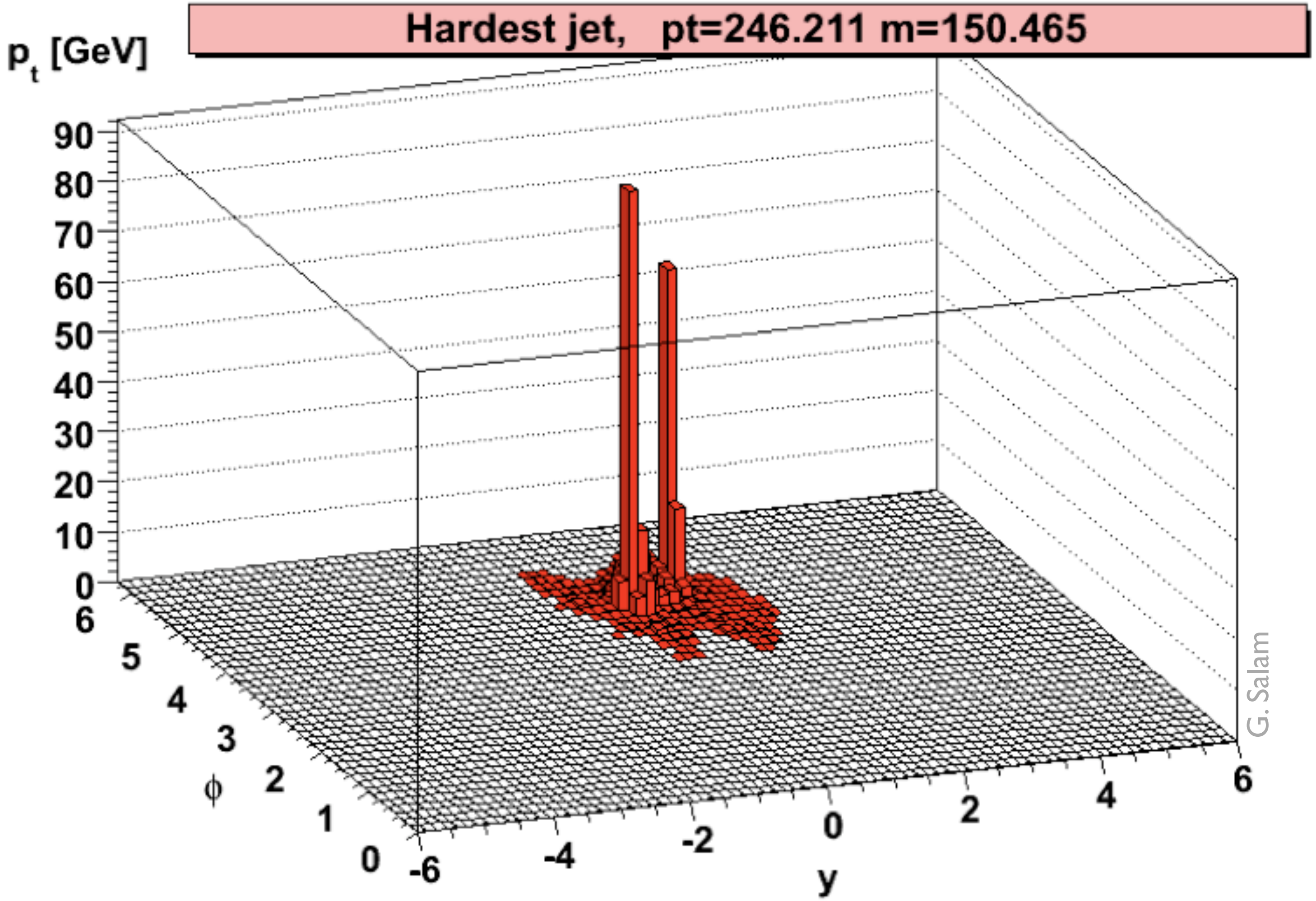
Butterworth, Davison, Rubin, Salam, 2008



- A two-prong tagger/groomer for boosted Higgs, which
  - Uses the **Cambridge/Aachen** algorithm (because it's 'physical')
  - 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**')

BDRS: tagging

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



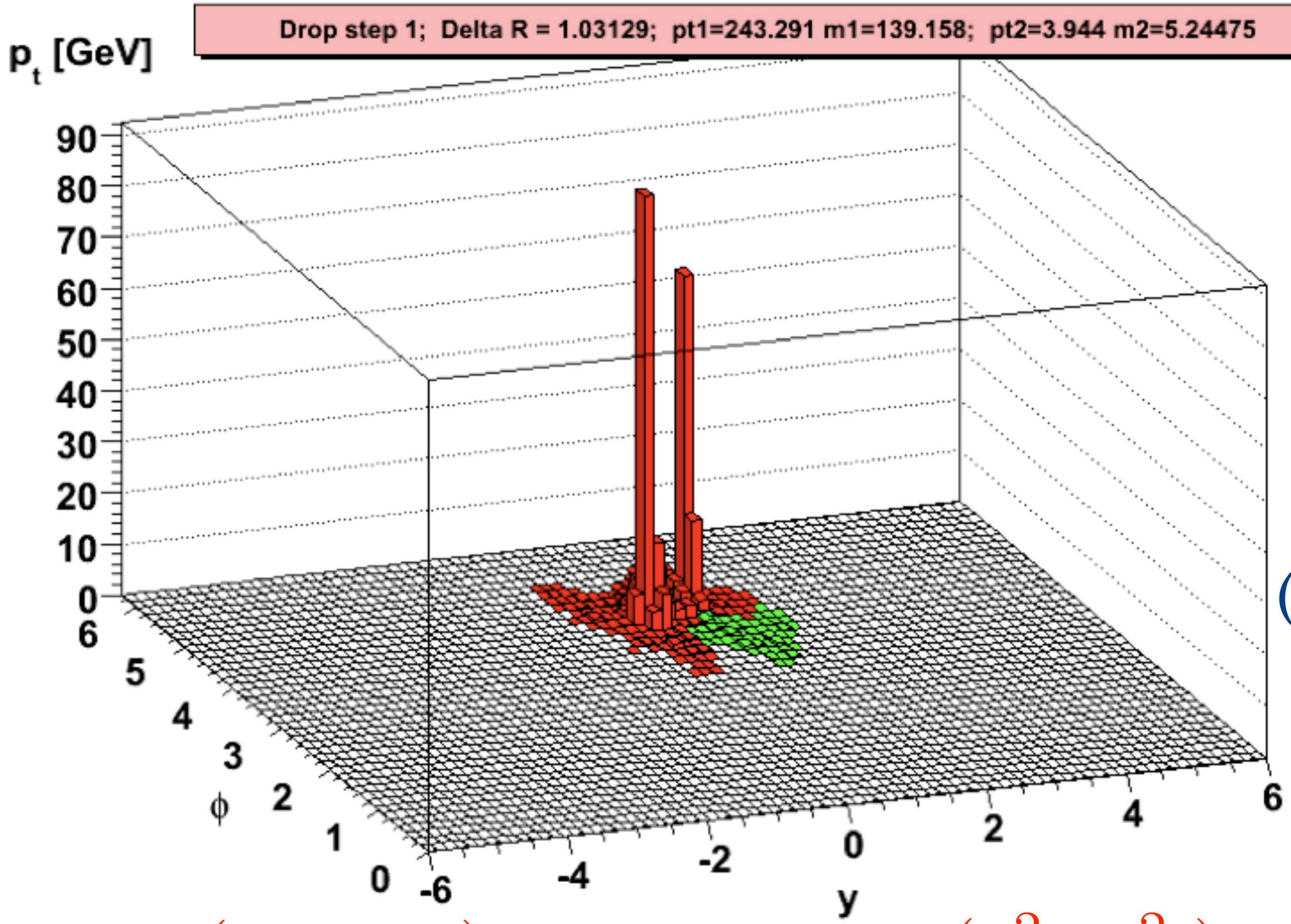
Start with the hardest jet

Use C/A with large R=1.2

m_j = 150 GeV

BDRS: tagging

pp → ZH → ννbb



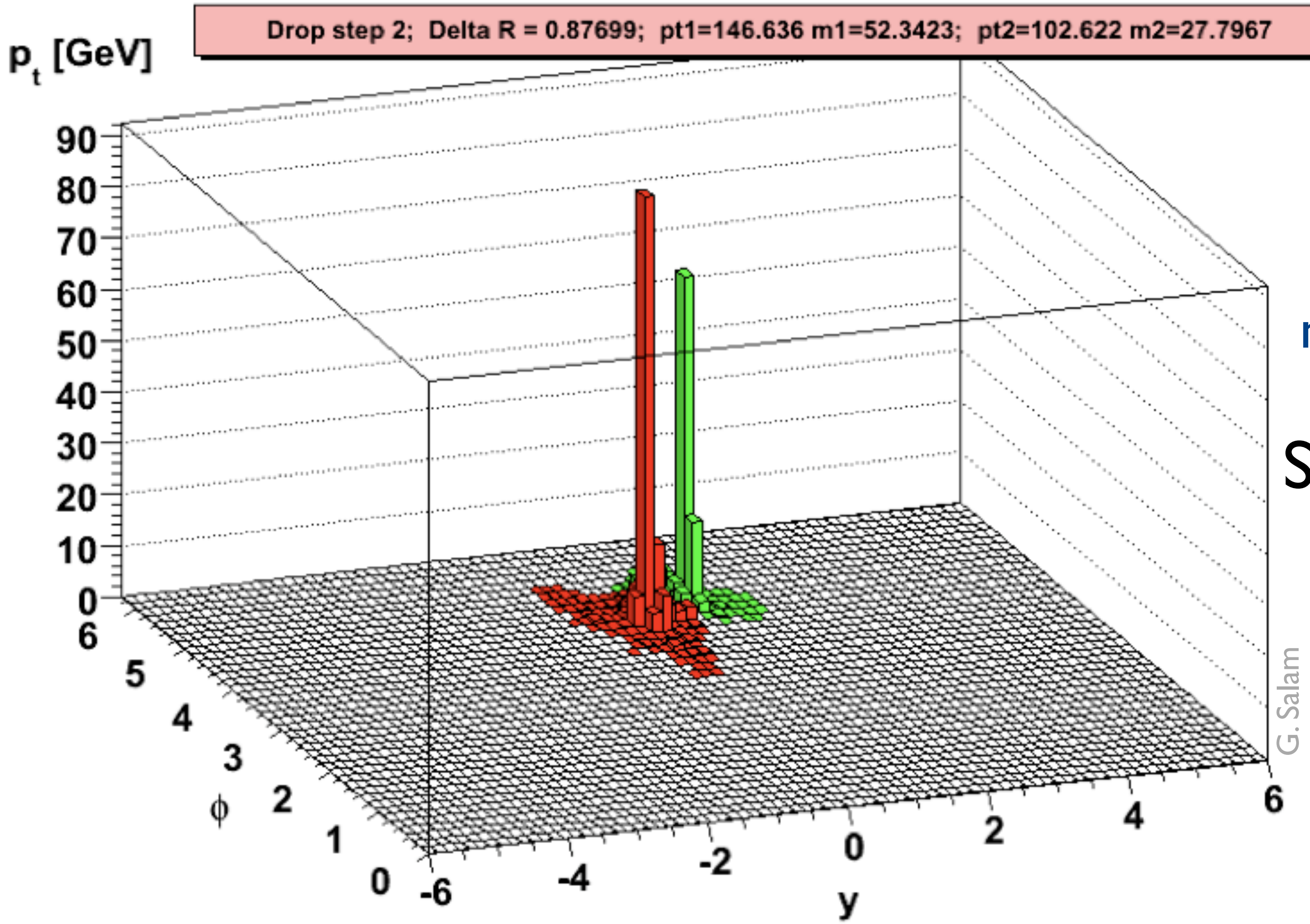
Undo last step of clustering

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

If $\frac{max(m_1, m_2)}{m_j} > \mu$ or $\frac{min(p_{t1}^2, p_{t2}^2)}{m_j^2} \Delta R_{12}^2 < y_{cut}$ repeat

BDRS: tagging

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



m_1 = 52 GeV, m_2 = 28 GeV

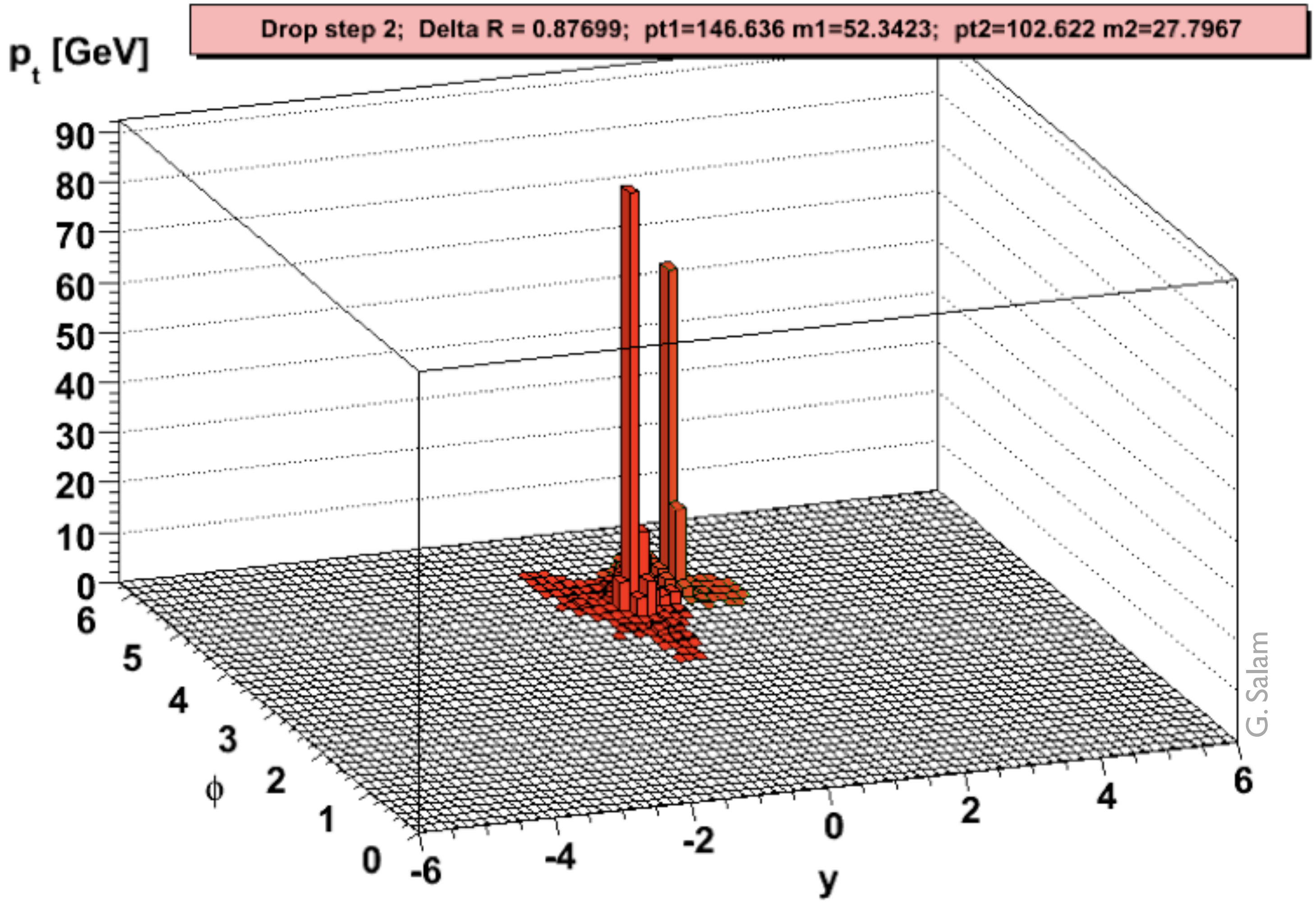
Stop when a large mass drop is observed

(and **recombine** these two jets)

[NB. Parameters used $\mu = 0.67$ and $y_{cut} = 0.09$]

BDRS: filtering

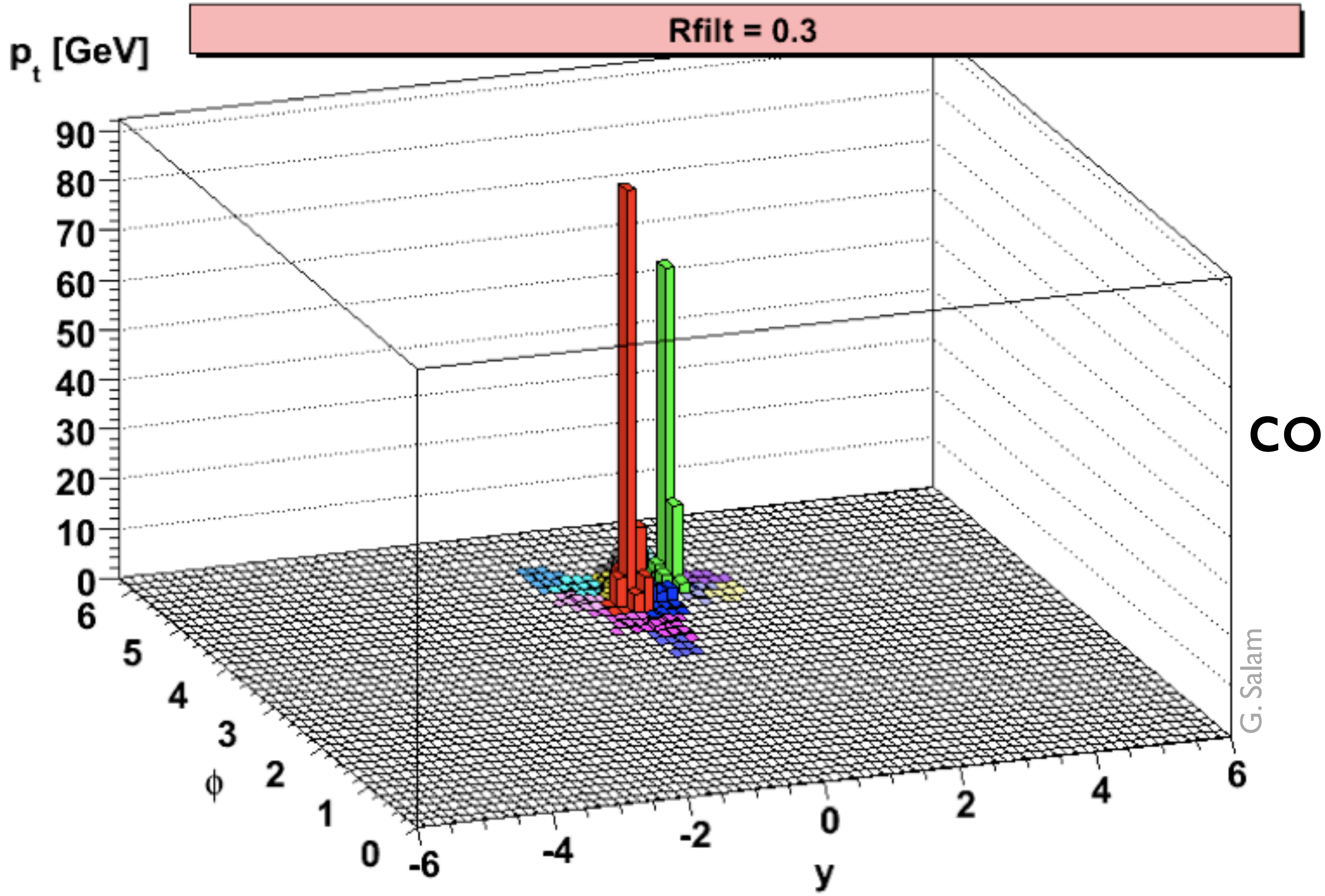
$pp \to ZH \to \nu\nu bb$



Start with the recombined jet

BDRS: filtering

pp →ZH → ννbb

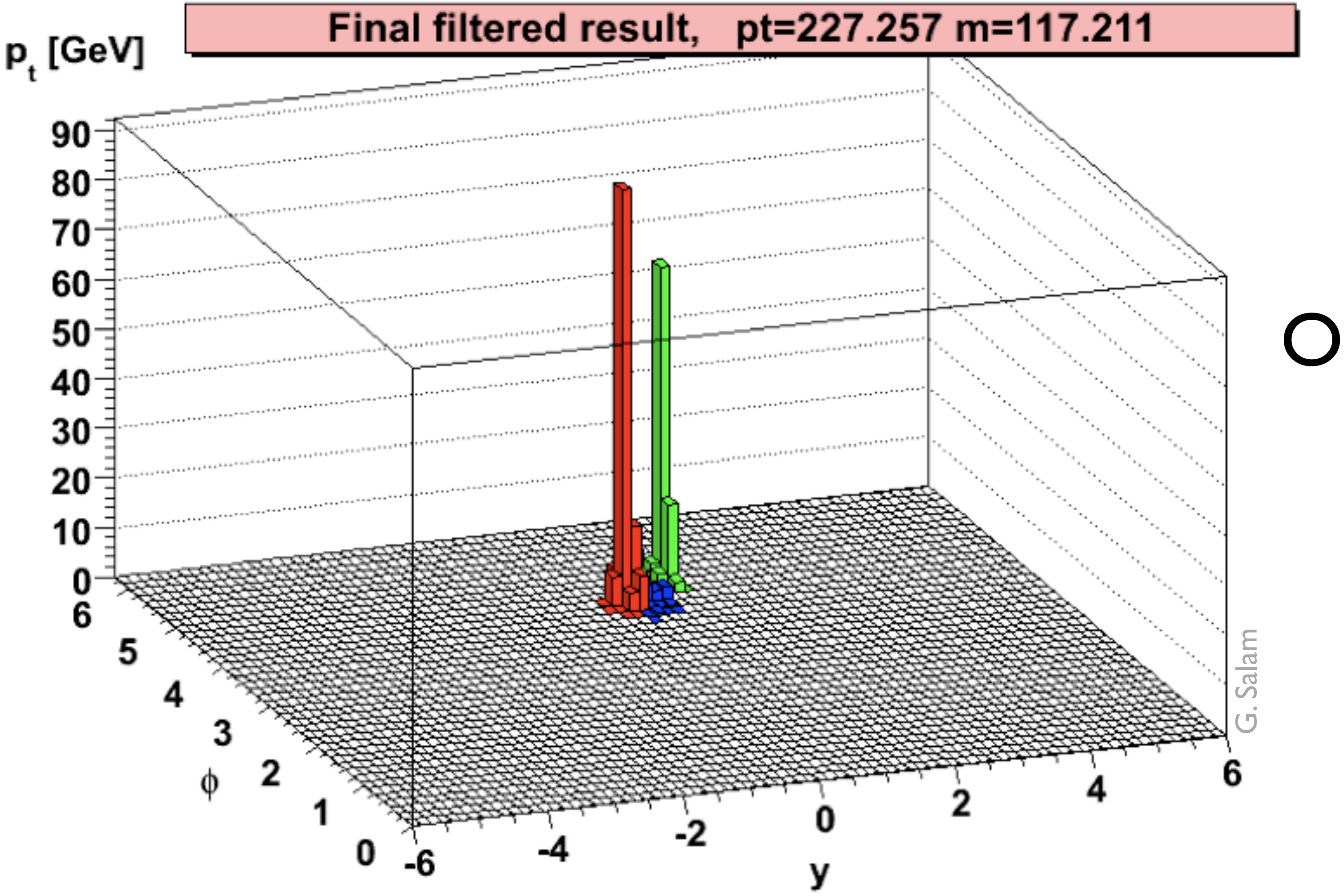


G. Salam

Recluster the contituents with R_{filt}

BDRS: filtering

pp →ZH → ννbb



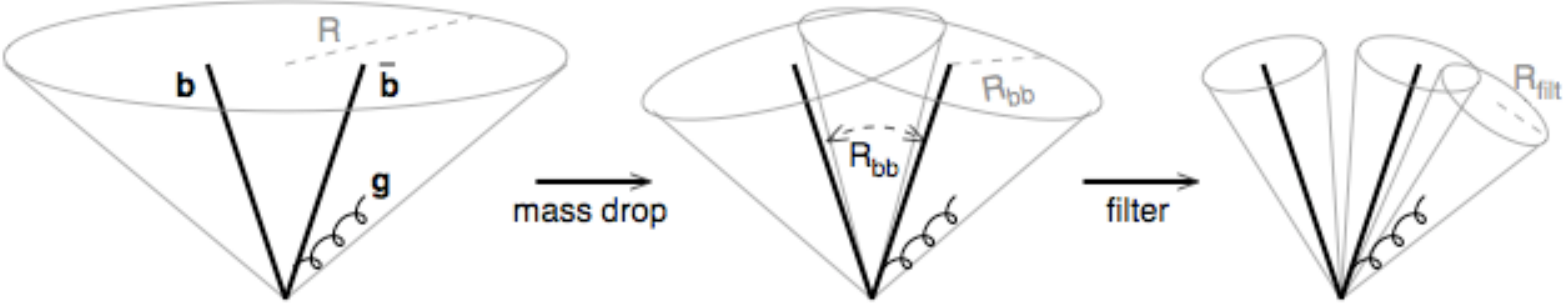
Only keep the n_{filt} hardest jets

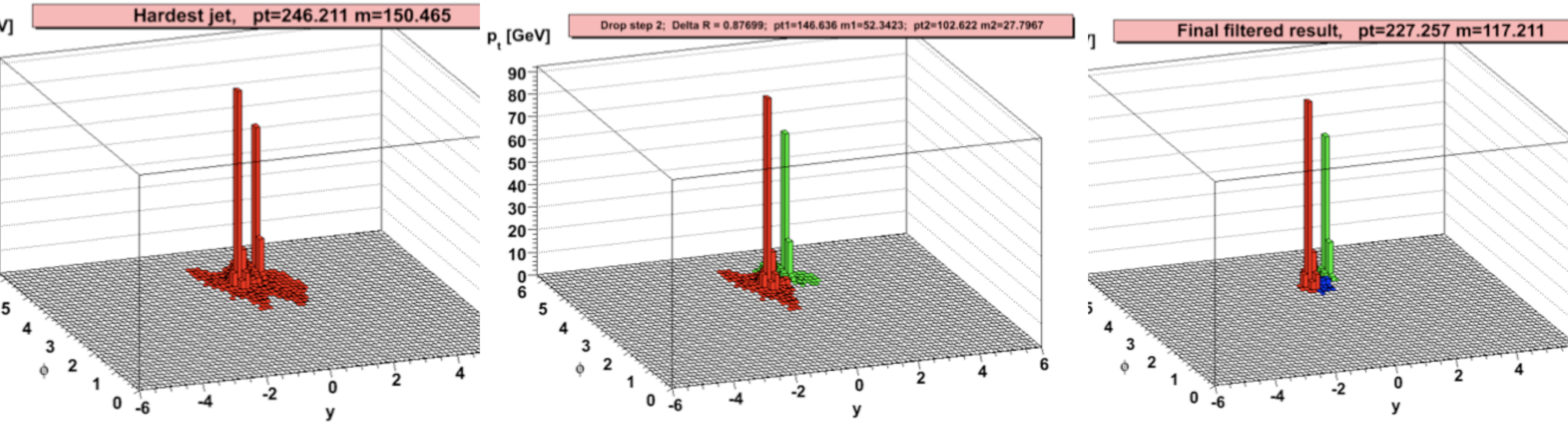
The low-momentum stuff surrounding the hard particles has been removed

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

Visualisation of BDRS

Butterworth, Davison, Rubin, Salam, 2008





Cluster with a large R

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

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

BDRS in FastJet

In FastJet

```
#include “fastjet/tools/MassDropTagger.hh”
#include “fastjet/tools/Filter.hh”

JetDefinition jet_def(cambridge_algorithm, 1.2);
ClusterSequence cs(input_particles, jet_def);

// define the tagger and use it
MassDropTagger md_tagger(0.667, 0.09);
PseudoJet tagged = md_tagger(jets[0]);

// define the filter and use it
Filter filter(0.3,SelectorNHardest(3));
Pseudojet higgs = filter(tagged);     // this is the Higgs!!
```

The real analysis is slightly more refined (b-tagging, dynamical filter radius, etc) but the main features are already present here

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

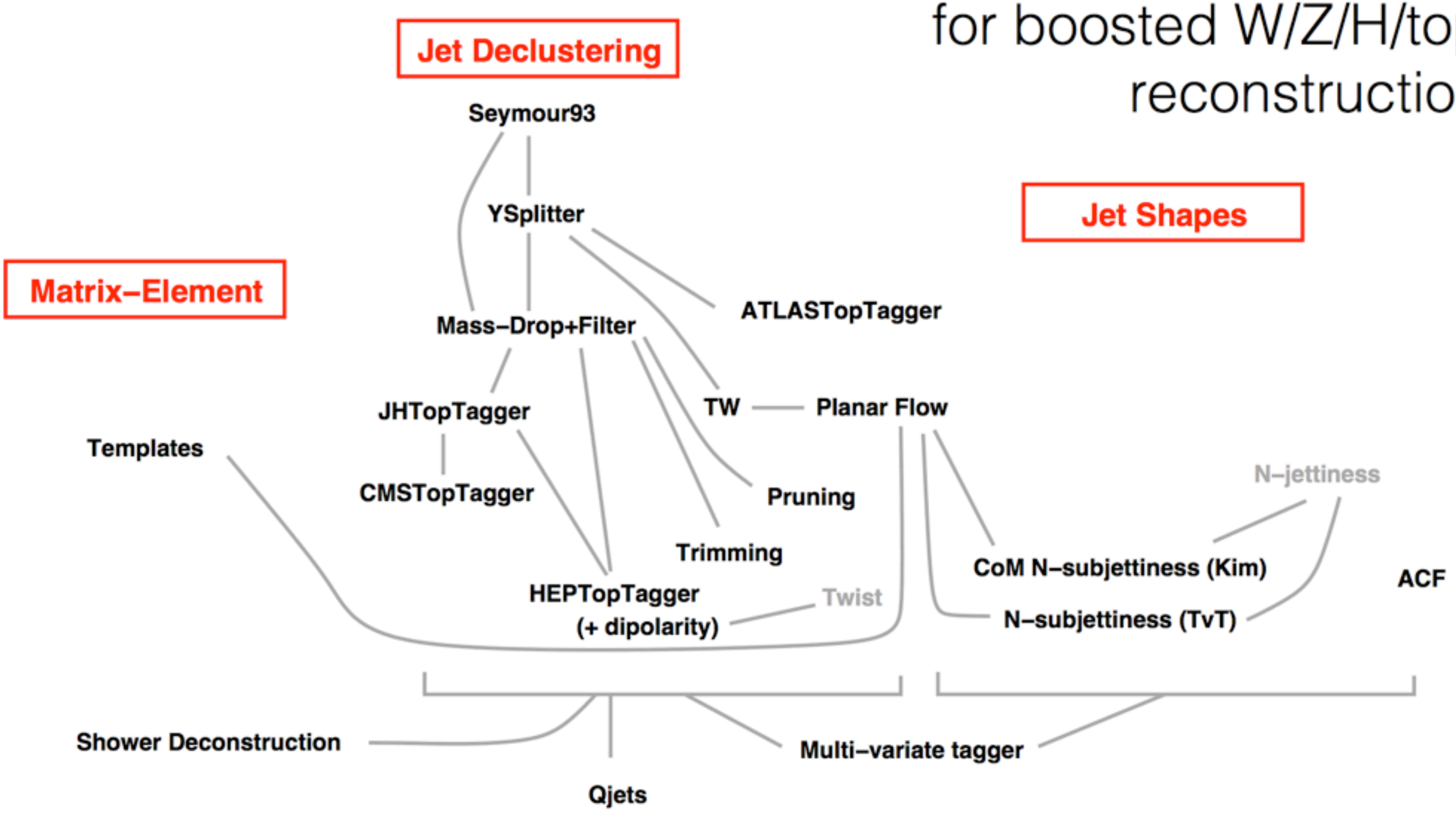
  While building up the jet, discard softer subjets when $\Delta R > R_{prune}$
  and $\min(p_{t1},p_{t2}) < \varepsilon_{prune}\ (p_{t1}+p_{t2})$

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

Trimming and pruner are a priori groomers, but can become taggers
when combined with an invariant mass window test
(if you can groom everything then there's no heavy particle in the jet)

The jet substructure maze

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



Jet Declustering

Jet Shapes

Matrix–Element

Seymour93

YSplitter

ATLASTopTagger

Mass–Drop+Filter

JHTopTagger

TW

Planar Flow

Templates

CMSTopTagger

Pruning

N–jettiness

Trimming

CoM N–subjettiness (Kim)

ACF

HEPTopTagger (+ dipolarity)

Twist

N–subjettiness (TvT)

Shower Deconstruction

Multi–variate tagger

Qjets

Slide by G. Salam, now a few years old

Soft Drop declustering

Larkoski, Marzani, Soyez, Thaler, 2014

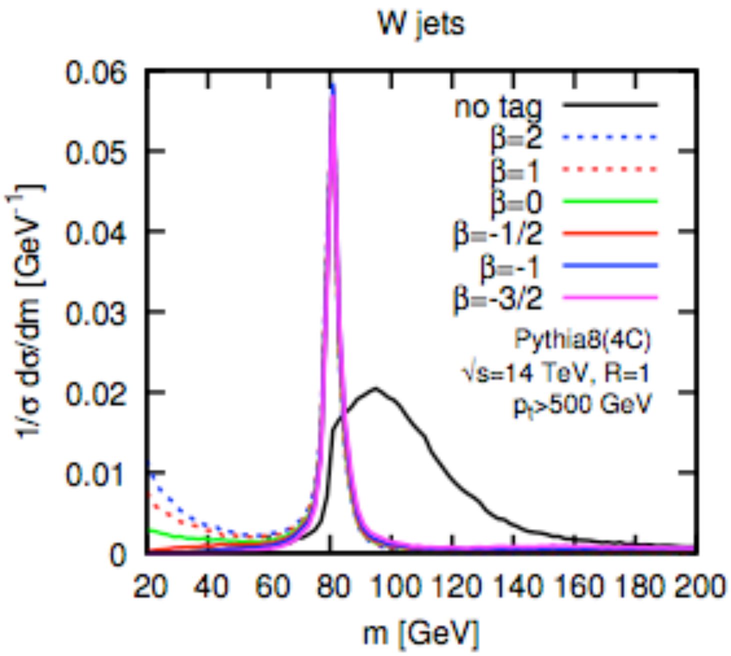
Decluster and drop softer constituent unless

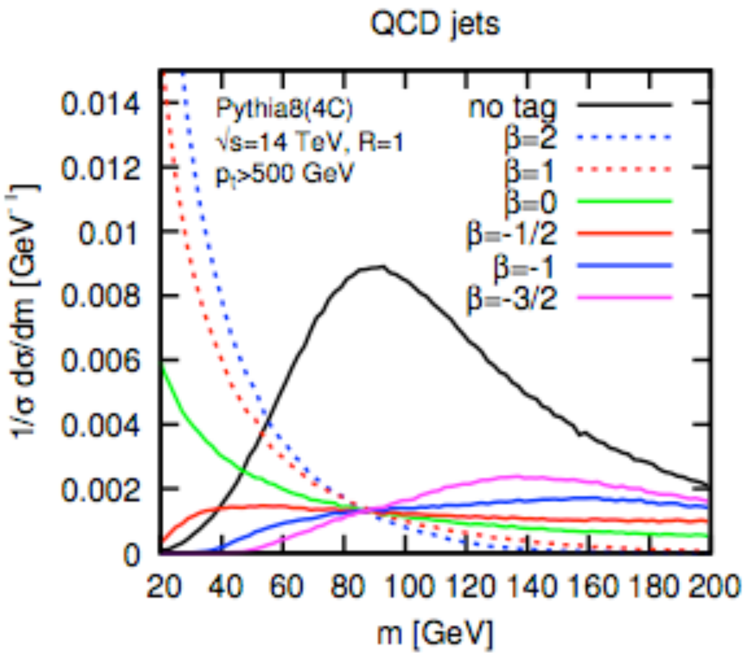
$$\text{Soft Drop Condition:} \quad \frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{\text{cut}} \left(\frac{\Delta R_{12}}{R_0} \right)^{\beta}$$

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

Alternatives to hierarchical substruct.

- 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

N-subjettiness

Thaler, van Tilburg, 2010

$$\tau_N^{(\beta)} = \sum_i p_{Ti} \min\left\{R_{1,i}^{\beta}, R_{2,i}^{\beta}, \ldots, R_{N,i}^{\beta}\right\}$$

Sum over constituents of a jet

Distances to axes of N subjets

τ_N measures departure from N-parton energy flow:

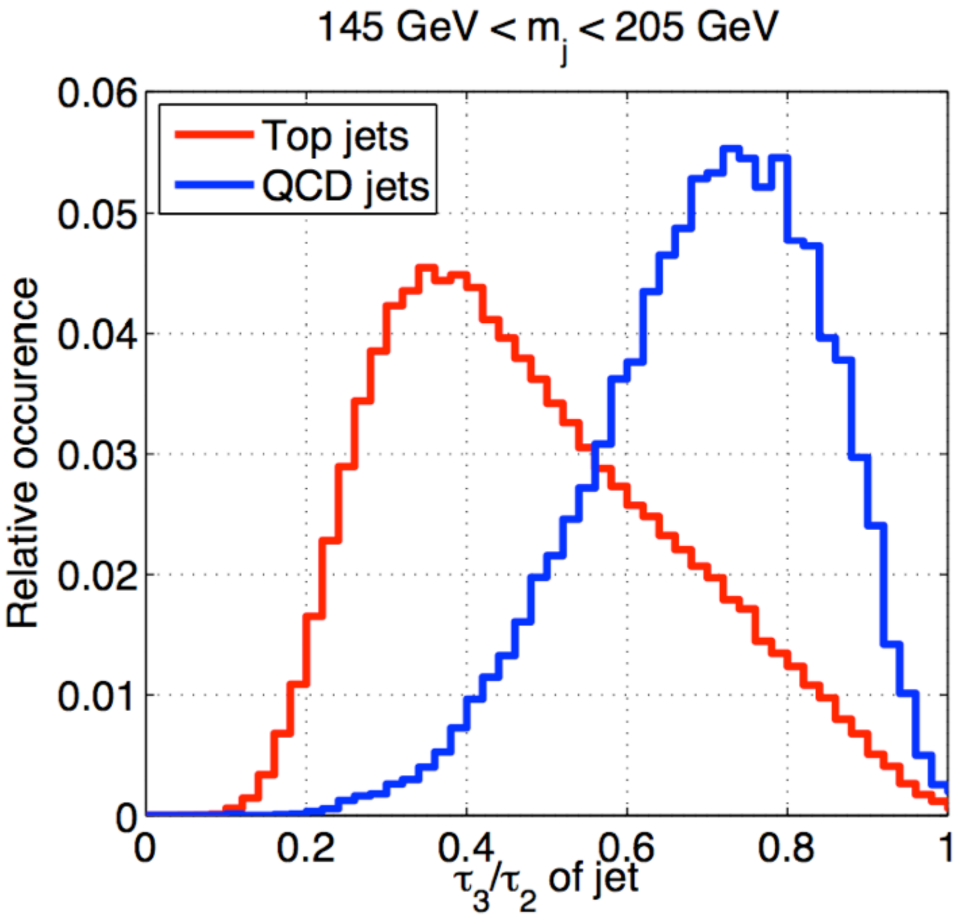
if a jet has N subjets, τ_{N-1} should be much larger than τ_N

N-subjettiness

Thaler, van Tilburg, 2010

$$\tau_{N,N-1}^{(\beta)} \equiv \frac{\tau_N^{(\beta)}}{\tau_{N-1}^{(\beta)}}$$

A jet with a **small** $\tau_{N,N-1}$ is more likely to have N than N-1 subjets



(from 1011.2268, with β=1)

Larkoski, Salam, Thaler 2013

Energy correlation functions

Probes of N-prong structures without requiring identification of subjets

$$\mathrm{ECF}(N,\beta) = \sum_{i_1<i_2<\ldots<i_N\in J}\left(\prod_{a=1}^{N} p_{T i_a}\right)\left(\prod_{b=1}^{N-1}\prod_{c=b+1}^{N} R_{i_b i_c}\right)^{\beta}$$

Angular (y-φ) distances between constituents

ECF(N+1) is zero if there are only N particles

More generally, *if there are N subjets one expects ECF(N+1) to be much smaller than ECF(N)*

[because radiation will be mainly soft/collinear to subjets]

Larkoski, Salam, Thaler 2013

Discriminators

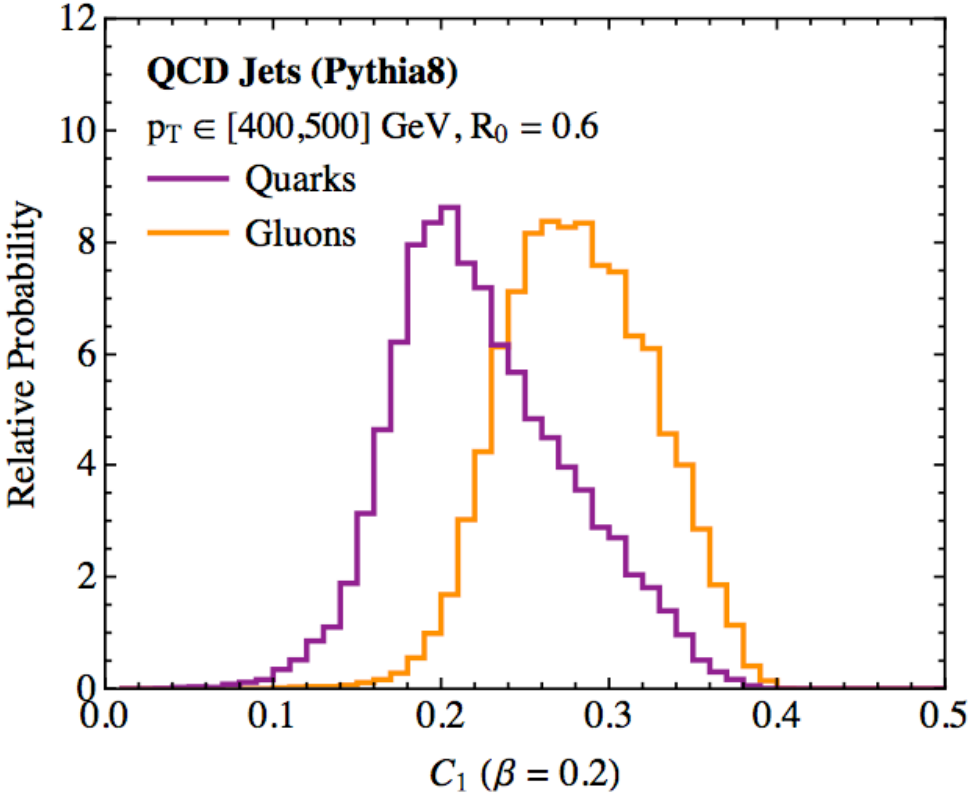
$$r_N^{(\beta)} \equiv \frac{\mathrm{ECF}(N+1,\beta)}{\mathrm{ECF}(N,\beta)}$$

small for N prongs:
if N hard partons, small if radiation
only soft-collinear

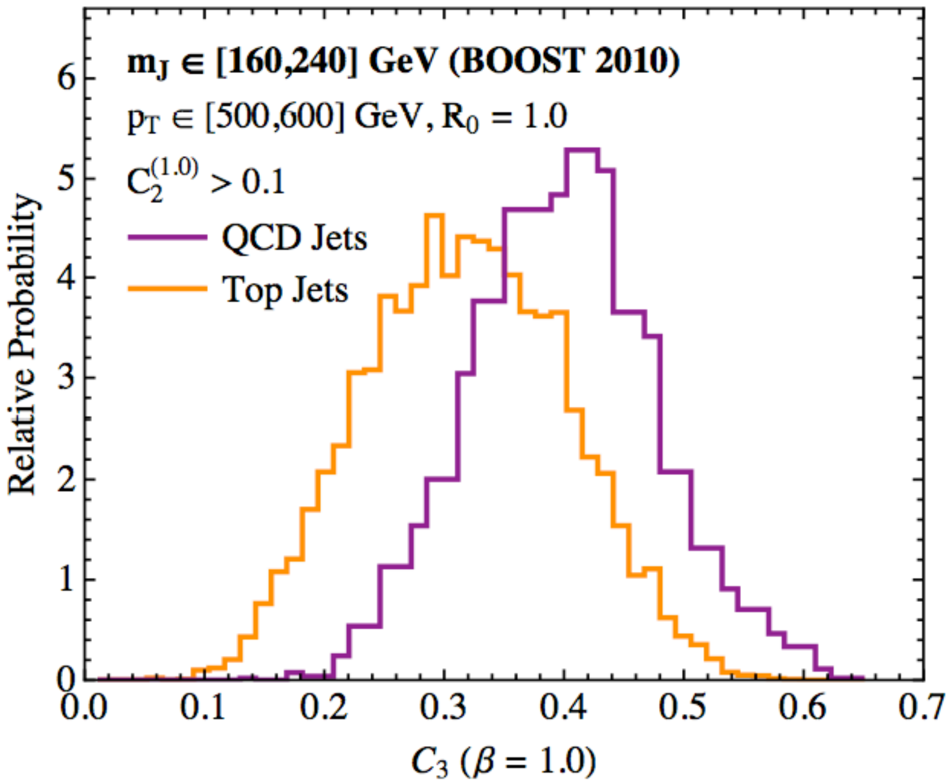
$$C_N^{(\beta)} \equiv \frac{r_N^{(\beta)}}{r_{N-1}^{(\beta)}} = \frac{\mathrm{ECF}(N+1,\beta)\,\mathrm{ECF}(N-1,\beta)}{\mathrm{ECF}(N,\beta)^2}$$

A jet with a **small** C_N is more likely
to have N prongs and at most soft/coll radiation

C_1
quark-gluon discriminator



C_3
top tagging



Note different values of β
(chosen to maximise discriminating power)

Larkoski, Moult, Neill, 2014

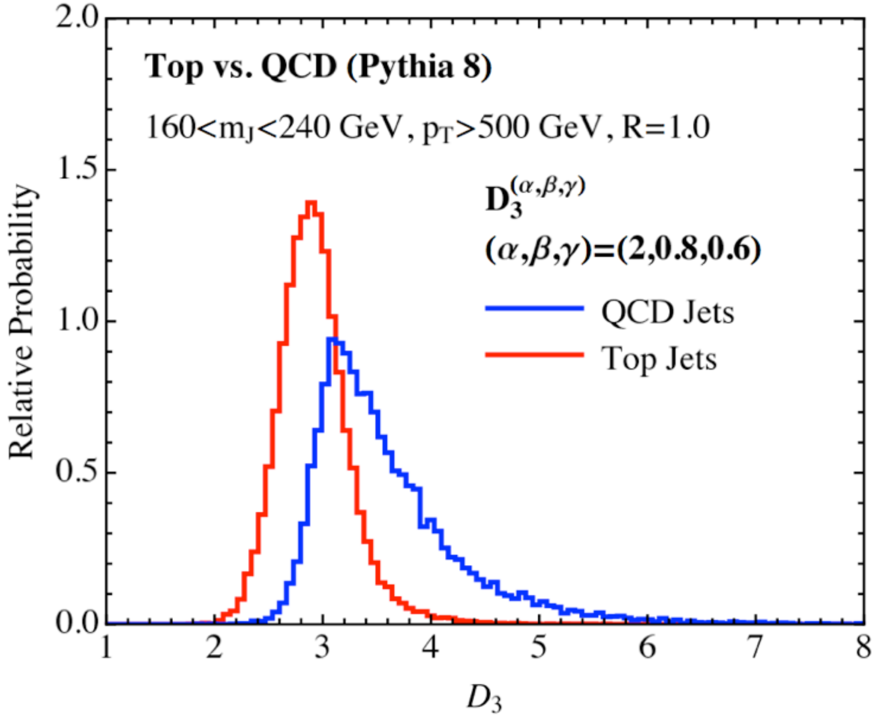
The D functions are variations of the C ones

Instead of $C_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^2}$ $C_3^{(\beta)} = \frac{e_4^{(\beta)} e_2^{(\beta)}}{(e_3^{(\beta)})^2}$

define $D_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3}$ $D_3^{(\alpha,\beta,\gamma)} = \frac{e_4^{(\gamma)} \left(e_2^{(\alpha)}\right)^{\frac{3\gamma}{\alpha}}}{\left(e_3^{(\beta)}\right)^{\frac{3\gamma}{\beta}}} + x \frac{e_4^{(\gamma)} \left(e_2^{(\alpha)}\right)^{\frac{2\gamma}{\beta}-1}}{\left(e_3^{(\beta)}\right)^{\frac{2\gamma}{\beta}}} + y \frac{e_4^{(\gamma)} \left(e_2^{(\alpha)}\right)^{\frac{2\beta}{\alpha}-\frac{\gamma}{\alpha}}}{\left(e_3^{(\beta)}\right)^2}$

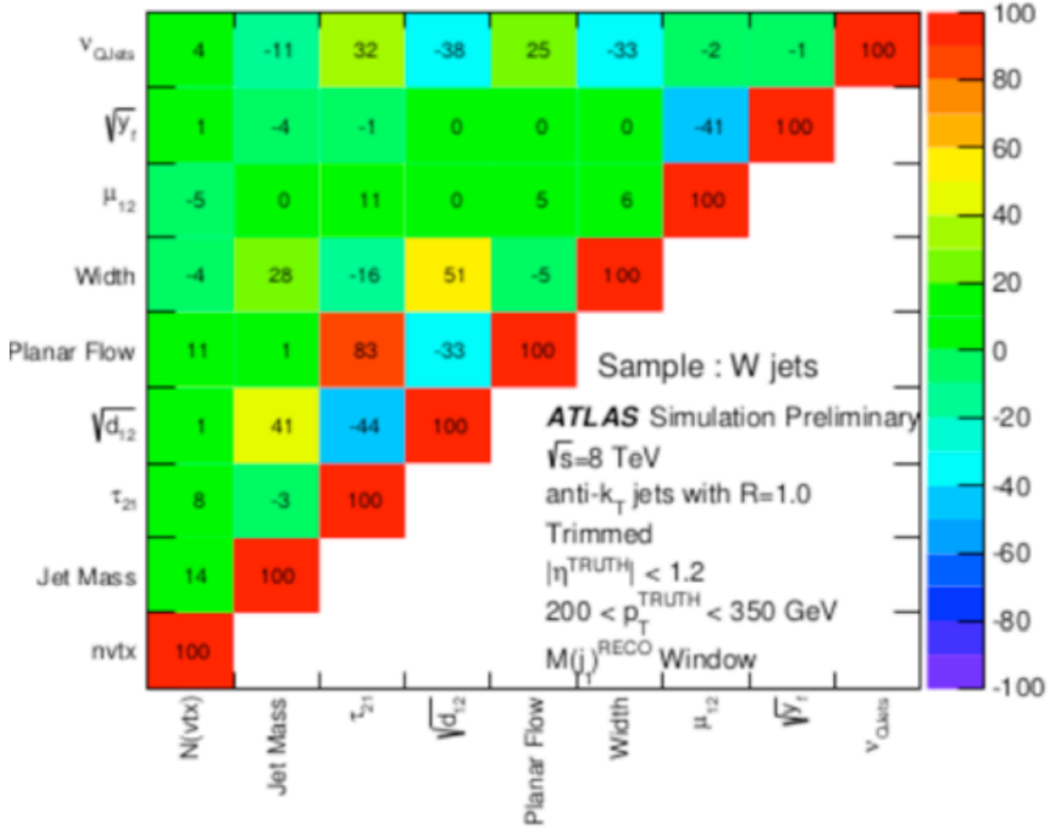
Attempt to improve the discriminating power, and to account for different regions of phase space of radiation

[also, gives an idea of increasing 'sophistication', or complexification]

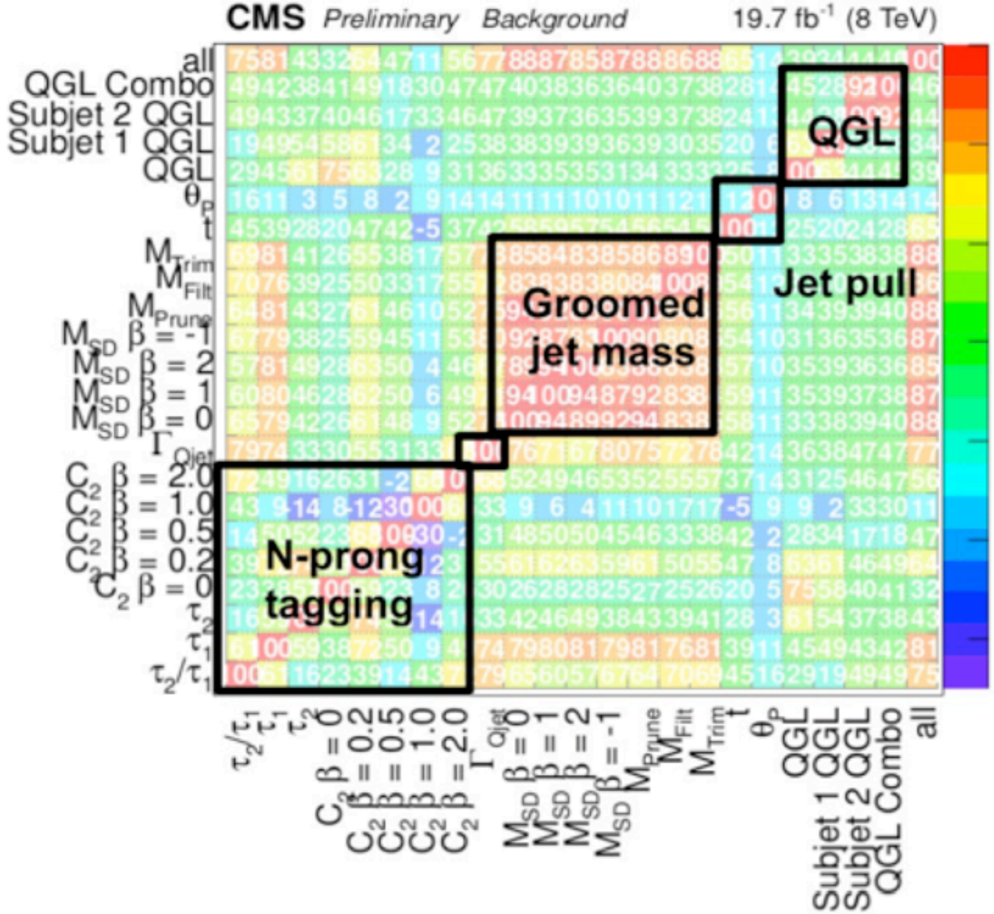


Zoology of substructure tools

Experimental studies of W/Z taggers



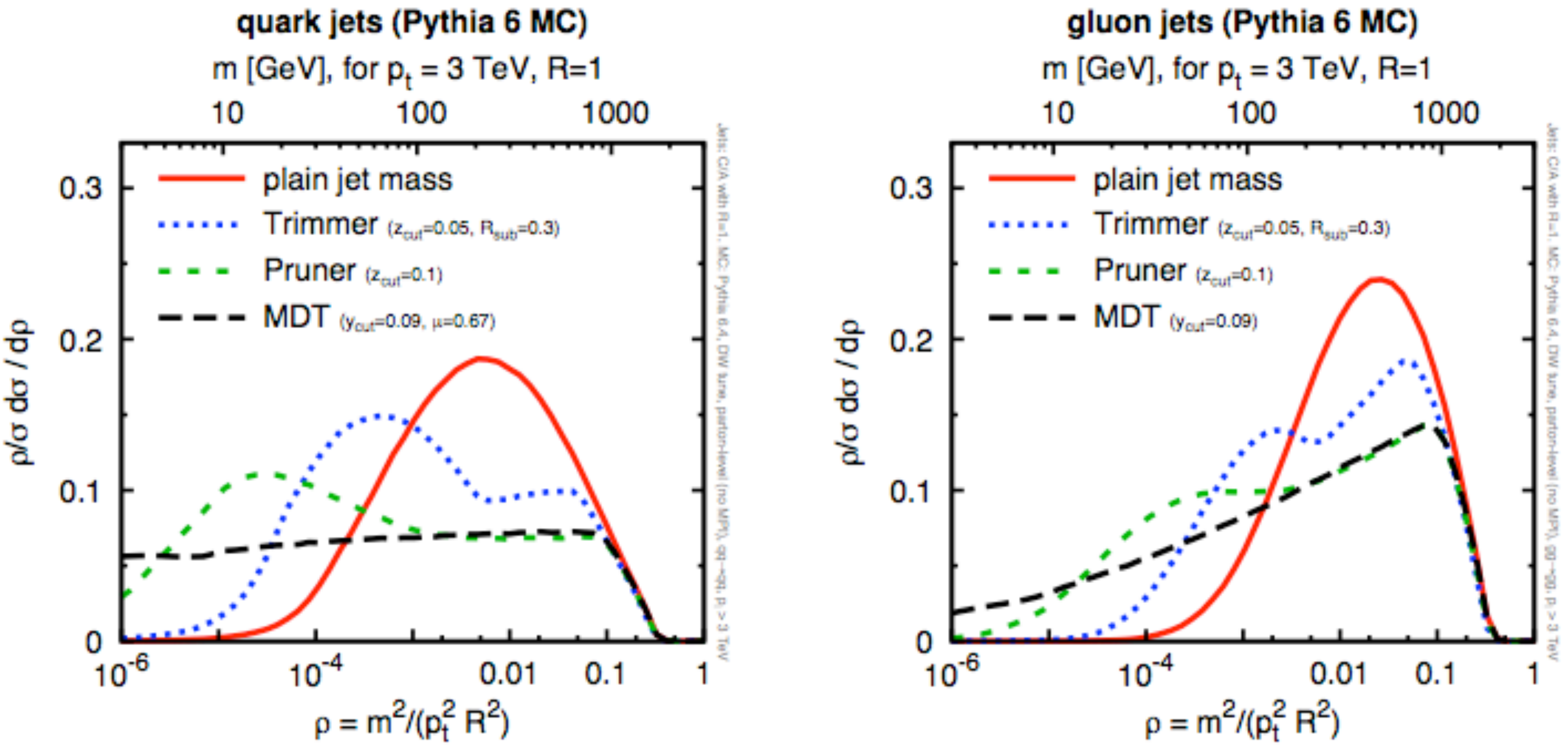
	N(vtx)	Jet Mass	τ_{21}	$\sqrt{d_{12}}$	Planar Flow	Width	μ_{12}	$\sqrt{y_f}$	v_{QJets}
v_{QJets}	4	-11	32	-38	25	-33	-2	-1	100
$\sqrt{y_f}$	1	-4	-1	0	0	0	-41	100	
μ_{12}	-5	0	11	0	5	6	100		
Width	-4	28	-16	51	-5	100			
Planar Flow	11	1	83	-33	100				
$\sqrt{d_{12}}$	1	41	-44	100					
τ_{21}	8	-3	100						
Jet Mass	14	100							
nvtx	100								



- Many variables/techniques are correlated
- Which ones are more robust?

Robustness of substructure tools

Dasgupta, Fregoso, Marzani, Salam, 2013

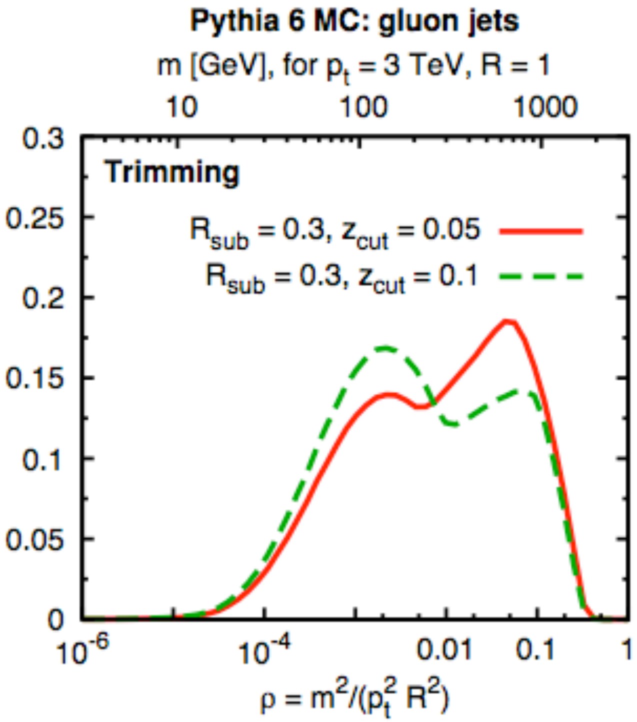


Tools that are considered (or can be seen in Monte Carlo tests) to behave ‘similarly’ could cease to do so in different parameter regions

Analytic calculations of jet substructure

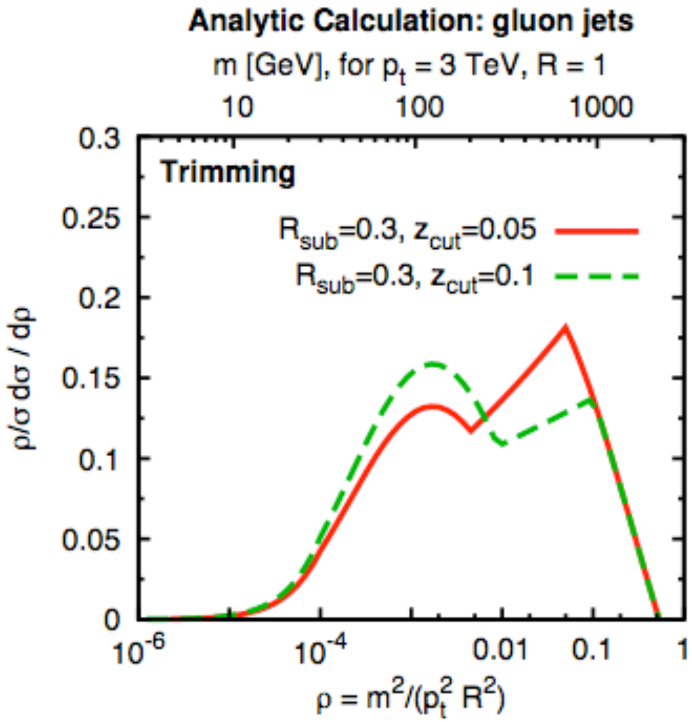
Dasgupta, Fregoso, Marzani, Salam, 2013

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), ...

$$\frac{1}{\sigma}\frac{d\sigma}{dm^2}^{(\text{trim, LO})} = \frac{\alpha_s C_F}{\pi}\int_0^1 dz\, p_{gq}(z)\int\frac{d\theta^2}{\theta^2}\,\delta\left(m^2 - z(1-z)p_t^2\theta^2\right)\times$$

$$\times\left[\Theta\left(z - z_{\text{cut}}\right)\Theta\left(1 - z - z_{\text{cut}}\right)\Theta(\theta^2 - R^2_{\text{sub}}) + \Theta(R^2_{\text{sub}} - \theta^2)\right]\Theta\left(R^2 - \theta^2\right)$$

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***

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)

Recap of lecture 2

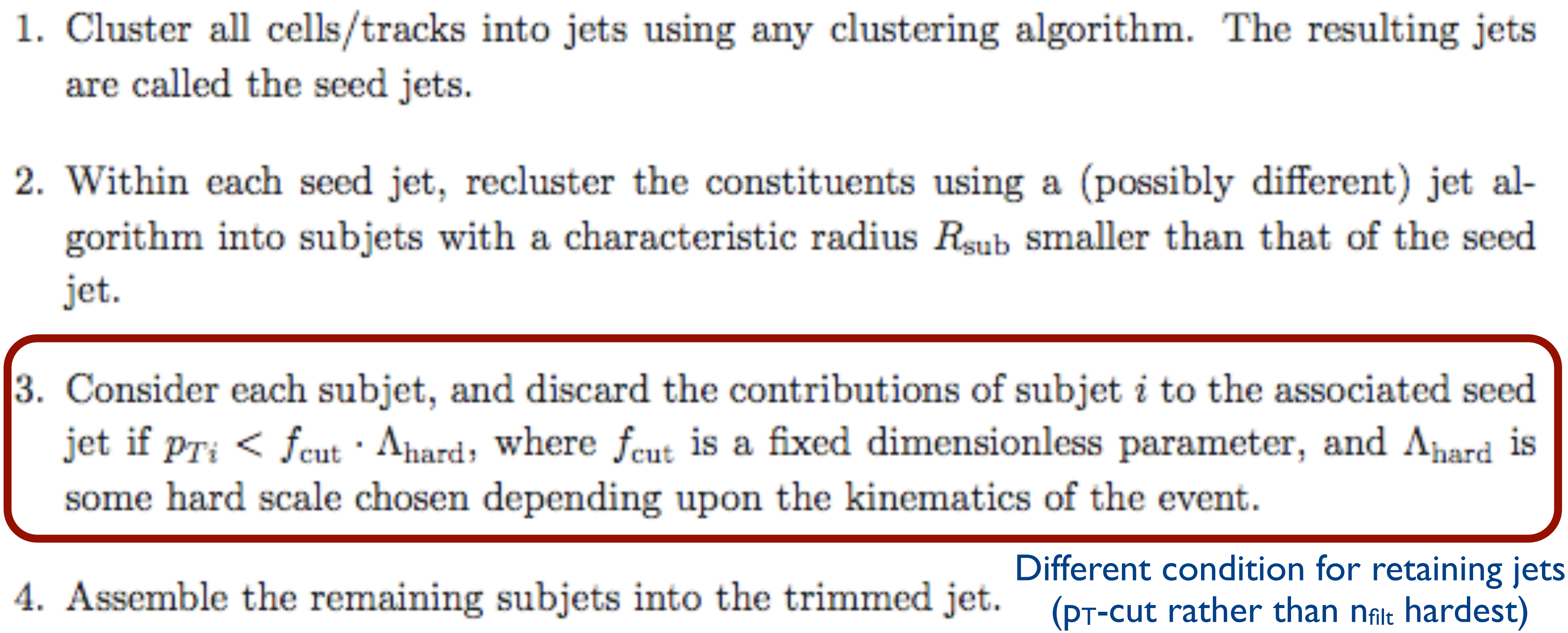
The big news of the past few years has been the emergence of jet-based taggers and groomers

- 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 balkanization and uncontrolled multiplication of not fully tested tools
    - resist the temptation of MVAisation
    - rather, ***try to grow a coherent, theoretically sound, robust, well tested and standardised library of tools***

Backup

Jet trimming

Krohn, Thaler, Wang, 2009

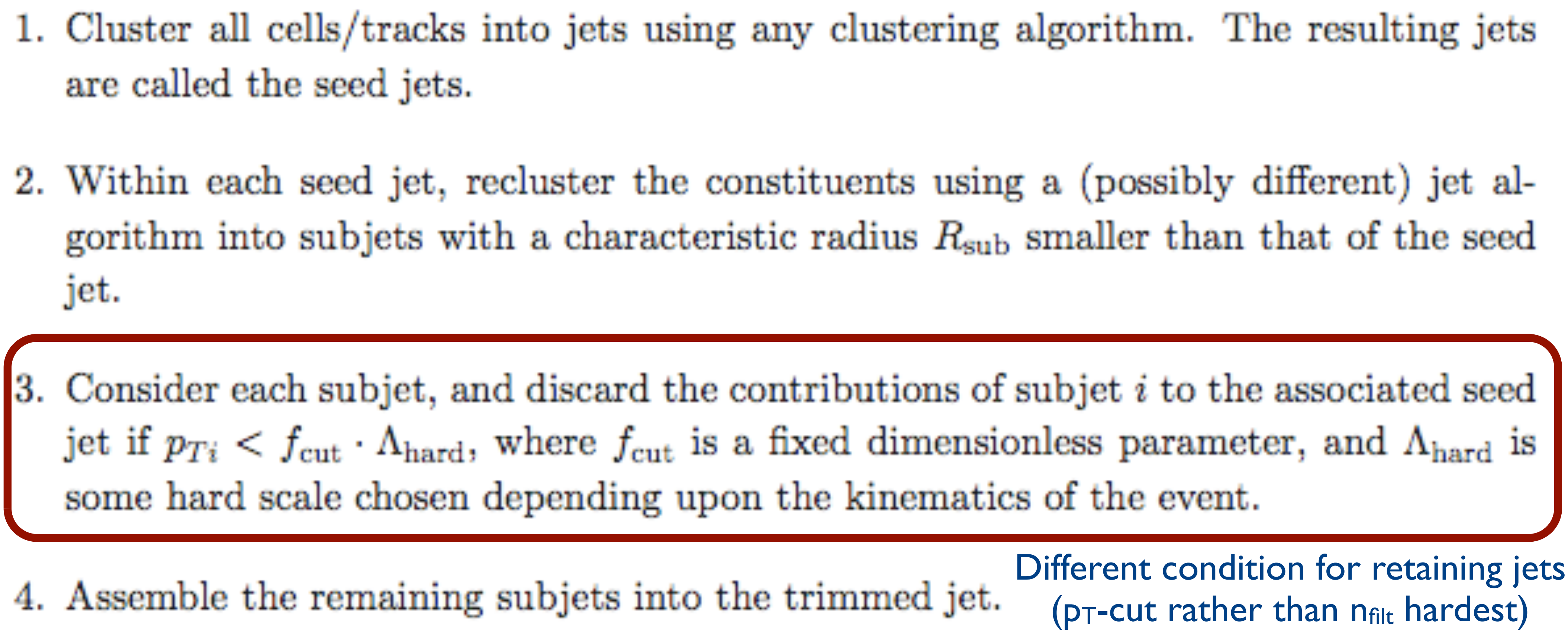


1. Cluster all cells/tracks into jets using any clustering algorithm. The resulting jets are called the seed jets.
2. Within each seed jet, recluster the constituents using a (possibly different) jet algorithm into subjets with a characteristic radius R_{sub} smaller than that of the seed jet.
3. Consider each subjet, and discard the contributions of subjet i to the associated seed jet if $p_{Ti} < f_{\text{cut}} \cdot \Lambda_{\text{hard}}$, where f_{cut} is a fixed dimensionless parameter, and Λ_{hard} is some hard scale chosen depending upon the kinematics of the event.
4. Assemble the remaining subjets into the trimmed jet.

Different condition for retaining jets (p_T-cut rather than n_{filt} hardest) with respect to filtering, but otherwise identical

Jet trimming

Krohn, Thaler, Wang, 2009



1. Cluster all cells/tracks into jets using any clustering algorithm. The resulting jets are called the seed jets.

2. Within each seed jet, recluster the constituents using a (possibly different) jet algorithm into subjets with a characteristic radius $R_{\rm sub}$ smaller than that of the seed jet.

3. Consider each subjet, and discard the contributions of subjet i to the associated seed jet if $p_{Ti} < f_{\rm cut} \cdot \Lambda_{\rm hard}$, where $f_{\rm cut}$ is a fixed dimensionless parameter, and $\Lambda_{\rm hard}$ is some hard scale chosen depending upon the kinematics of the event.

4. Assemble the remaining subjets into the trimmed jet.

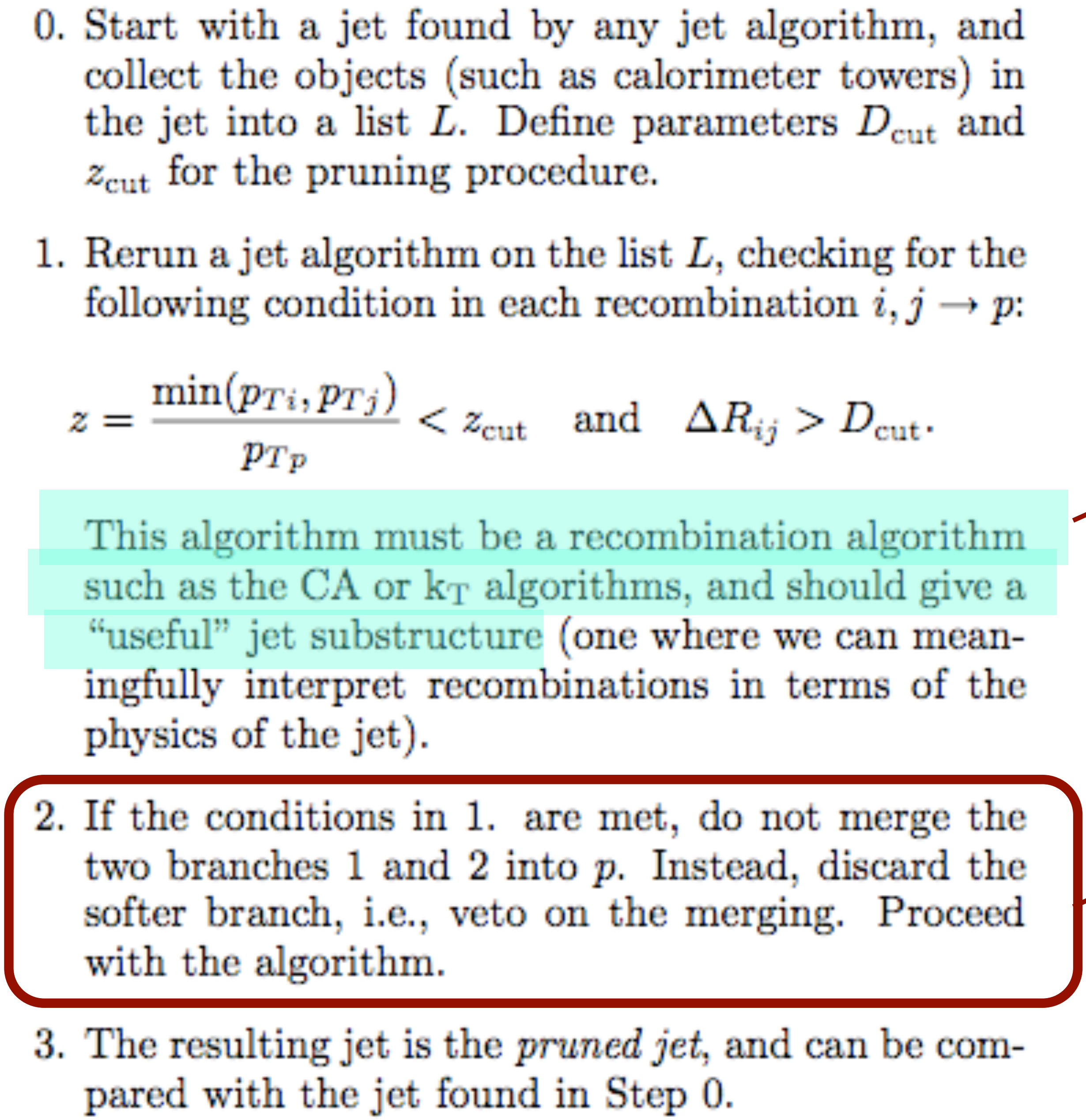
Different condition for retaining jets (p_T-cut rather than $n_{\rm filt}$ hardest) with respect to filtering, but otherwise identical

```
#include “fastjet/tools/Filter.hh”

// define trimmer
Filter trimmer(0.3,SelectorPtFractionMin(0.03));
```

Jet pruning

S. Ellis, Vermilion, Walsh, 2009



0. Start with a jet found by any jet algorithm, and collect the objects (such as calorimeter towers) in the jet into a list L. Define parameters $D_{\rm cut}$ and $z_{\rm cut}$ for the pruning procedure.

1. Rerun a jet algorithm on the list L, checking for the following condition in each recombination $i, j \to p$:

$$z = \frac{\min(p_{Ti}, p_{Tj})}{p_{Tp}} < z_{\rm cut} \quad \text{and} \quad \Delta R_{ij} > D_{\rm cut}.$$

   This algorithm must be a recombination algorithm such as the CA or k_T algorithms, and should give a "useful" jet substructure (one where we can meaningfully interpret recombinations in terms of the physics of the jet).

2. If the conditions in 1. are met, do not merge the two branches 1 and 2 into p. Instead, discard the softer branch, i.e., veto on the merging. Proceed with the algorithm.

3. The resulting jet is the *pruned jet*, and can be compared with the jet found in Step 0.

True in general for substructure studies

Exclude soft stuff and large angle recombinations from clustering

Pruning in FastJet

In FastJet

```
#include “fastjet/tools/Pruner.hh”

JetDefinition jet_def(cambridge_algorithm, 1.2);
ClusterSequence cs(input_particles, jet_def);

// define the pruner and use it
double zcut = 0.1;
double rcut_factor = 0.5;

Pruner pruner(cambridge_algorithm, zcut, rcut_factor);

PseudoJet tagged = pruner(jets[0]);
```