MANCHESTER 1824

The University of Manchester

Jet substructure and boosted object studies

Mrinal Dasgupta

CTEQ//DESY/MCNET School, DESY Hamburg

14 July 2016

Outline

- Introduction to boosted particle searches and jet substructure
- Introduction to some jet substructure methods
- Jet substructure from QCD first principles?
- Calculations for substructure methods.
- Improving substructure tools and enhancing performance.

Boosted object hadronic decays

Boosted regime implies studying particles with $p_T >> M_X$. Important at the LHC with access to TeV scales in p_T.

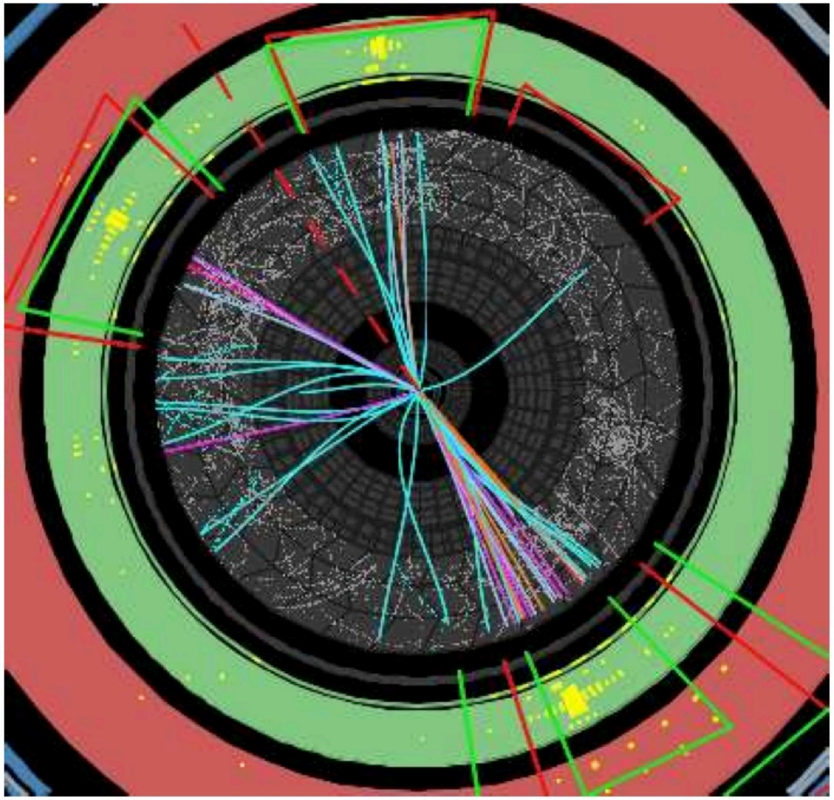
Decay products are collimated.

$$\theta^2 = \frac{M^2}{p_T^2 z(1-z)}$$

Hadronic two-body decays often reconstructed in single jet.

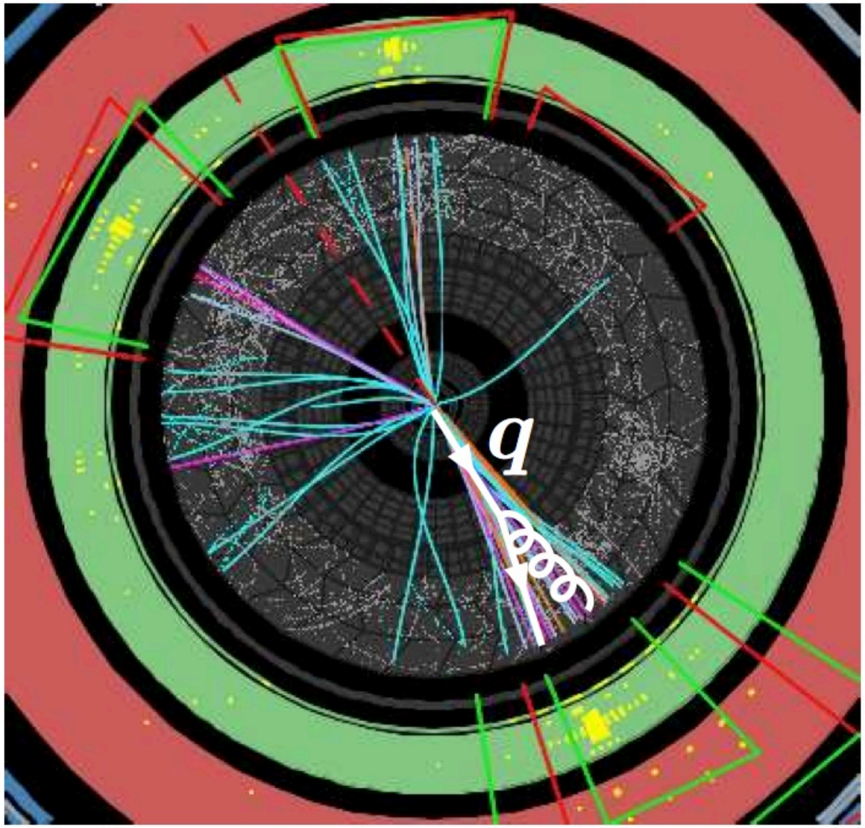
Jets from QCD vs boosted heavy particles

What jet do we have here?



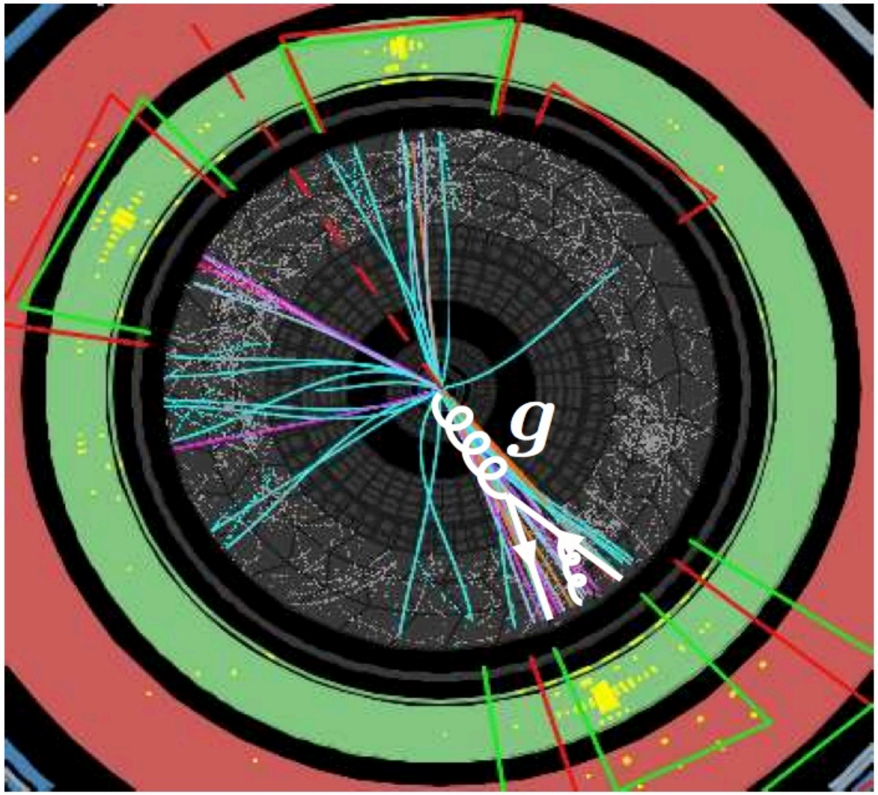
Jets from QCD vs boosted heavy particles

A quark jet ?



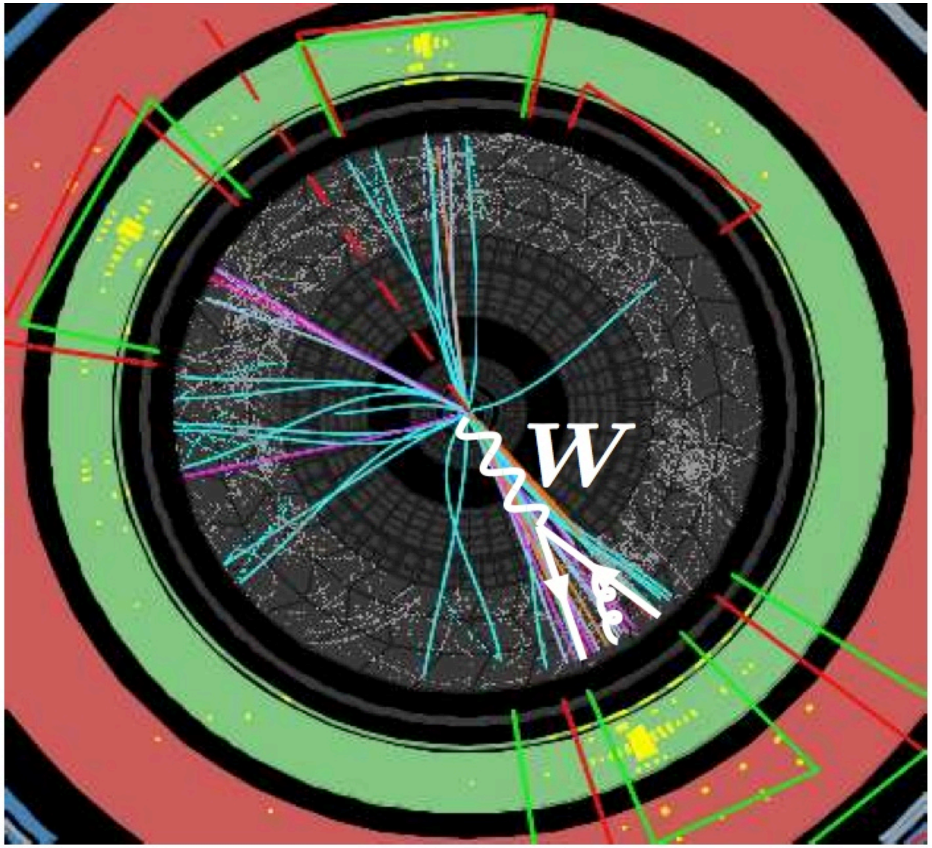
Jets from QCD vs boosted heavy particles

A gluon jet ?



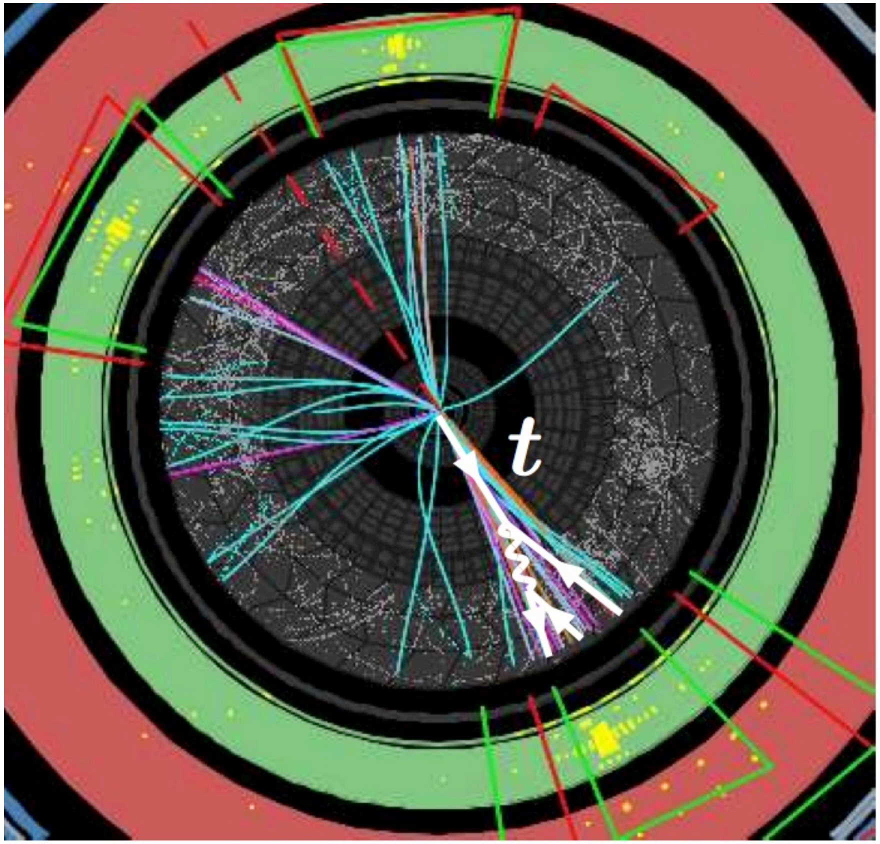
Jets from QCD vs boosted heavy particles

A W/Z/H ?



Jets from QCD vs boosted heavy particles

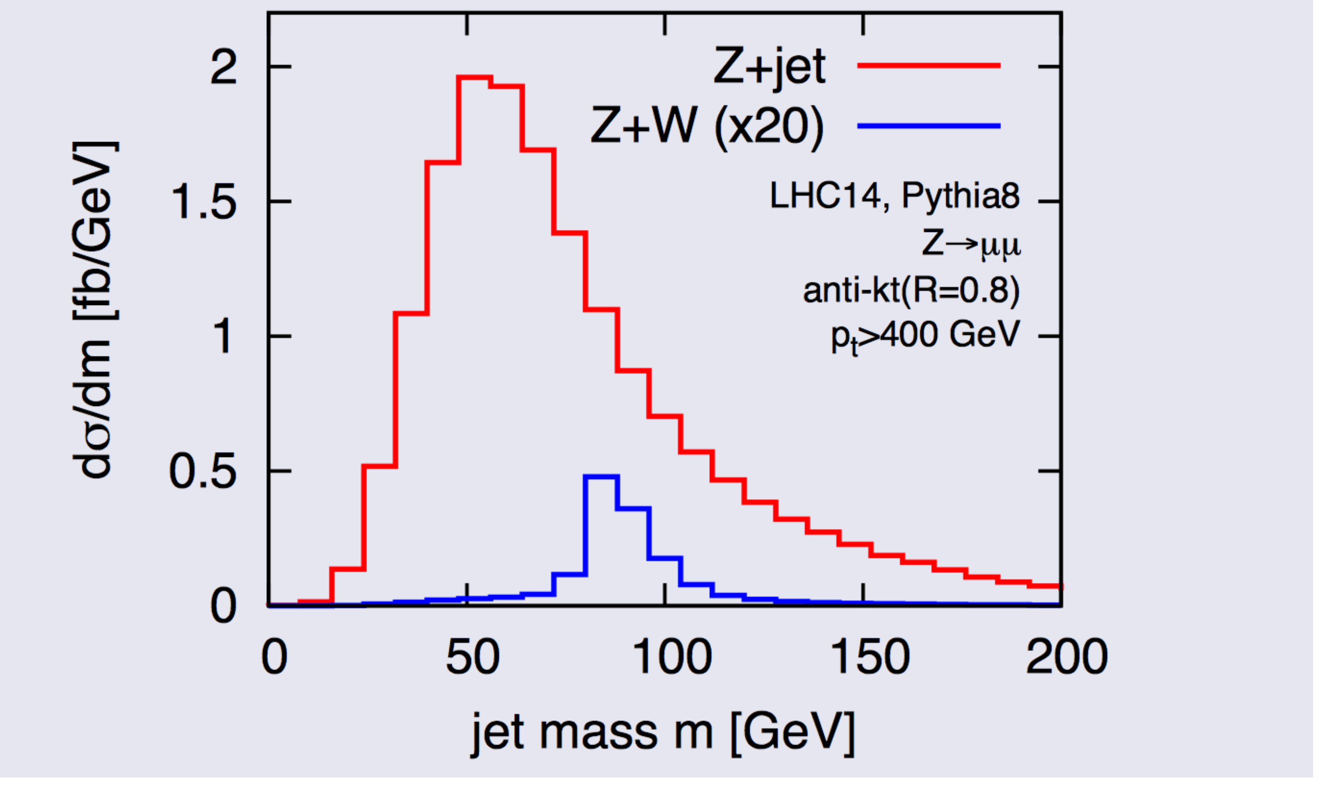
A top quark?



Source: An ATLAS boosted top candidate

The boosted regime implies a change in paradigm in that jets can be more than quarks and gluons.

Isn't the jet mass a clue?



Looking at jet mass is not enough!

Jet substructure for LHC searches

Jet substructure as a new Higgs search channel at the LHC

Jonathan M. Butterworth, Adam R. Davison
Department of Physics & Astronomy, University College London.

Mathieu Rubin, Gavin P. Salam
LPTHE; UPMC Univ. Paris 6; Univ. Denis Diderot; CNRS UMR 7589; Paris, France.

It is widely considered that, for Higgs boson searches at the Large Hadron Collider, WH and ZH production where the Higgs boson decays to $b\bar{b}$ are poor search channels due to large backgrounds. We show that at high transverse momenta, employing state-of-the-art jet reconstruction and decomposition techniques, these processes can be recovered as promising search channels for the standard model Higgs boson around 120 GeV in mass.

arXiv:0802.2470v2 [hep-ph] 19 Jun 2008

A key aim of the Large Hadron Collider (LHC) at CERN is to discover the Higgs boson, the particle at the heart of the standard-model (SM) electroweak symmetry breaking mechanism. Current electroweak fits, together with the LEP exclusion limit, favour a light Higgs boson, i.e. one around 120 GeV in mass [1]. This mass region is particularly challenging for the LHC experiments, and any SM Higgs-boson discovery is expected to rely on a combination of several search channels, including gluon fusion $\to H \to \gamma\gamma$, vector boson fusion, and associated production with $t\bar{t}$ pairs [2, 3].

Two significant channels that have generally been considered less promising are those of Higgs-boson production in association with a vector boson, $pp \to WH, ZH$, followed by the dominant light Higgs boson decay, to two b-tagged jets. If there were a way to recover the WH and ZH channels it could have a significant impact on Higgs boson searches at the LHC. Furthermore these two channels also provide unique information on the couplings of a light Higgs boson separately to W and Z bosons.

Reconstructing W or Z associated $H \to b\bar{b}$ production would typically involve identifying a leptonically decaying vector boson, plus two jets tagged as containing b-mesons. Two major difficulties arise in a normal search scenario. The first is related to detector acceptance: leptons and b-jets can be effectively tagged only if they are reasonably central and of sufficiently high transverse momentum. The relatively low mass of the VH (i.e. WH or ZH) system means that in practice it can be produced at rapidities somewhat beyond the acceptance, and it is also not unusual for one or more of the decay products to have too small a transverse momentum. The second issue is the presence of large backgrounds with intrin-responds to only a small fraction of the total VH cross section (about 5% for $p_T > 200$ GeV), but it has several compensating advantages: (i) in terms of acceptance, the larger mass of the VH system causes it to be central, and the transversely boosted kinematics of the V and H ensures that their decay products will have sufficiently large transverse momenta to be tagged; (ii) in terms of backgrounds, it is impossible for example for an event with on-shell top-quarks to produce a high-p_T $b\bar{b}$ system *and* a compensating leptonically decaying W, without there also being significant additional jet activity; (iii) the HZ with $Z \to \nu\bar{\nu}$ channel becomes visible because of the large missing transverse energy.

One of the keys to successfully exploiting the boosted VH channels will lie in the use of jet-finding geared to identifying the characteristic structure of a fast-moving Higgs boson that decays to b and $\bar{b}$ in a common neighbourhood in angle. We will therefore start by describing the method we adopt for this, which builds on previous work on heavy Higgs decays to boosted W's [4], WW scattering at high energies [5] and the analysis of SUSY decay chains [6]. We shall then proceed to discuss event generation, our precise cuts and finally show our results.

When a fast-moving Higgs boson decays, it produces a single fat jet containing two b quarks. A successful identification strategy should flexibly adapt to the fact that the $b\bar{b}$ angular separation will vary significantly with the Higgs p_T and decay orientation, roughly

$$R_{b\bar{b}} \simeq \frac{1}{\sqrt{z(1-z)}}\frac{m_H}{p_T}\,, \qquad (p_T \gg m_H)\,, \tag{1}$$

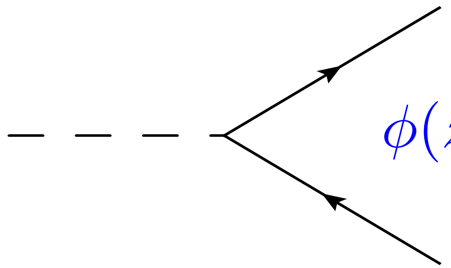
where z, $1-z$ are the momentum fractions of the two quarks. In particular one should capture the $b, \bar{b}$ and any

Since 2008 a vibrant research field emerged based on developing and exploiting jet substructure.

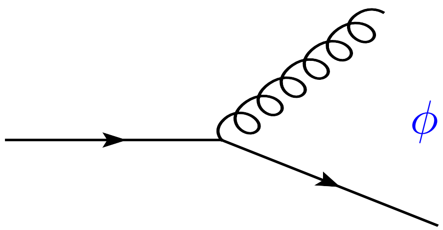
Butterworth, Davison Rubin, Salam 2008. Published in PRL. Builds on work by Seymour 1993.

BDRS paper has over 600 citations. “Jet substructure” title search on arXiv gives > 100 papers post BDRS.

Signal vs background



$$\phi(z) \propto 1$$

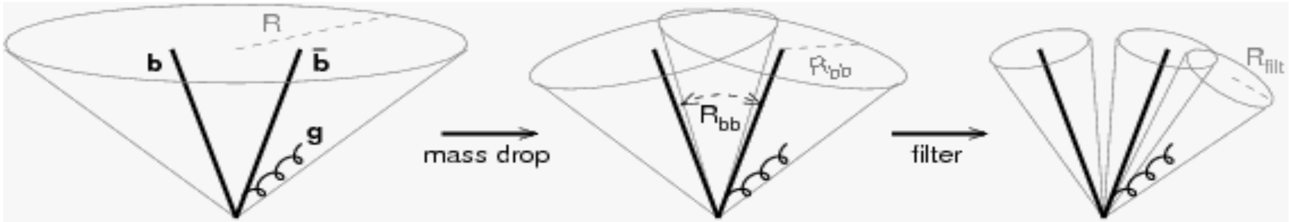


$$\phi(z) \propto \frac{1+z^2}{1-z}$$

BDRS studied the process $pp \to VH,\ H \to b\bar{b}$

- This was considered an unpromising channel for Higgs discovery due to large QCD backgrounds.
- In boosted limit Higgs decay products are reconstructed in a single fat jet and need to distinguish a signal jet from a plain QCD jet.
- One key is that QCD branchings have soft enhancements. Asymmetric sharing of energy compared to Higgs case.

BDRS mass drop tagger

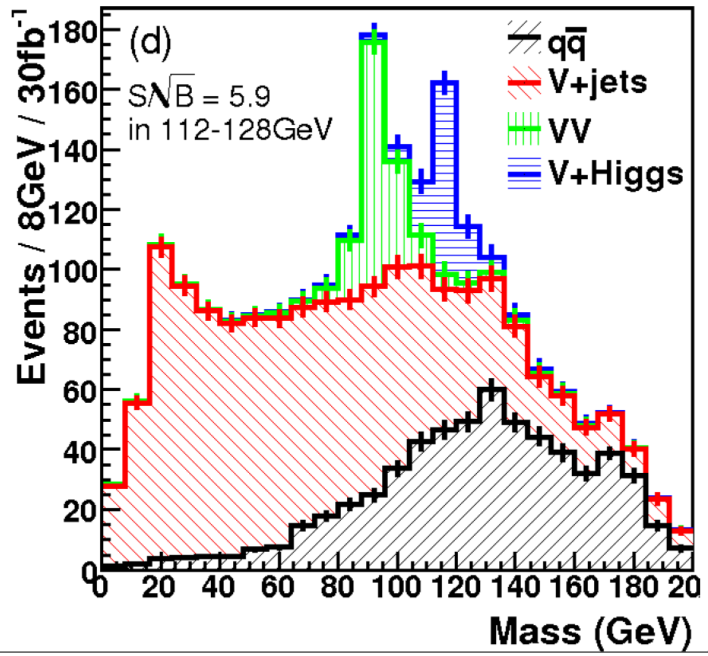


- Break the jet into two subjets j_1 and j_2 such that $mj_1 > mj_{2}$.
- Require mass drop $m_{j1} < \mu m_j$ and $\frac{\min(p_{t1}, p_{t2})}{\max(p_{t1}, p_{t2})} > y_{\mathrm{cut}}$

Then deem the jet tagged or if not discard j_2 and continue.

- Additional filtering step involves reclustering with smaller radius and retaining only n_{filt} hardest subjets.

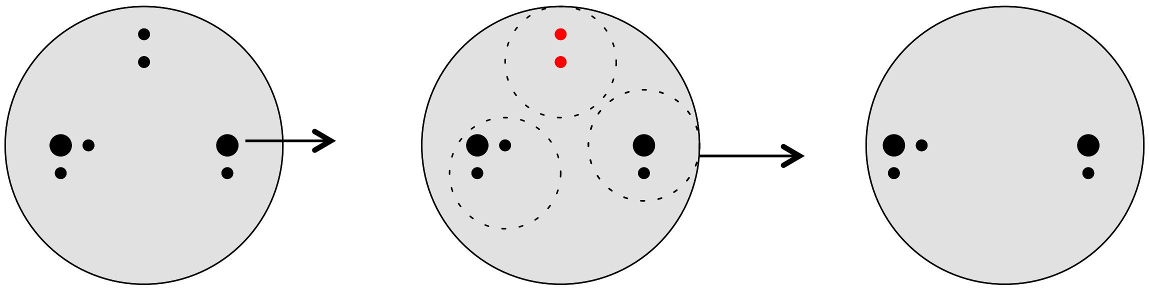
BDRS method results



MDT + filtering

Signal significance of 4.5σ was demonstrated in MC studies for a Higgs boson of 115 GeV. Turned this unpromising channel into one of the best discovery channels for light Higgs.

Several other methods exist



Trimming re-clusters jet with smaller radius R_{trim}.
Discards subjets with $p_{t,subjet} < z_{cut}\ p_{tjet}$.

Krohn, Thaler, Wang 2010

Pruning is similar but uses a dynamical radius $R_{prune} \sim m_j/p_t$.

Ellis, Walsh, Vermillion 2009

Many other methods: Y-splitter, Atlas top tagger, HEP top tagger, CMS top tagger, JH top tagger, Template Overlap, Planar Flow, Shower Deconstruction, Qjets, N-subjettiness, ECF's etc.

Tagging and grooming

There are two main ideas:

Idea 1:
Find $N = 2, 3, \ldots$ hard cores

Works because different splitting

QCD jets: $P(z) \propto 1/z$

$\Rightarrow$ dominated by soft emissions

$\Rightarrow$ "single" hard core

Idea 2:
Constrain radiation patterns

Works because different colours

Radiation pattern is different for

- colourless $W \to q\bar{q}$
- coloured $g \to q\bar{q}$

Taggers try and exploit the above differences.
But we also need jet grooming.

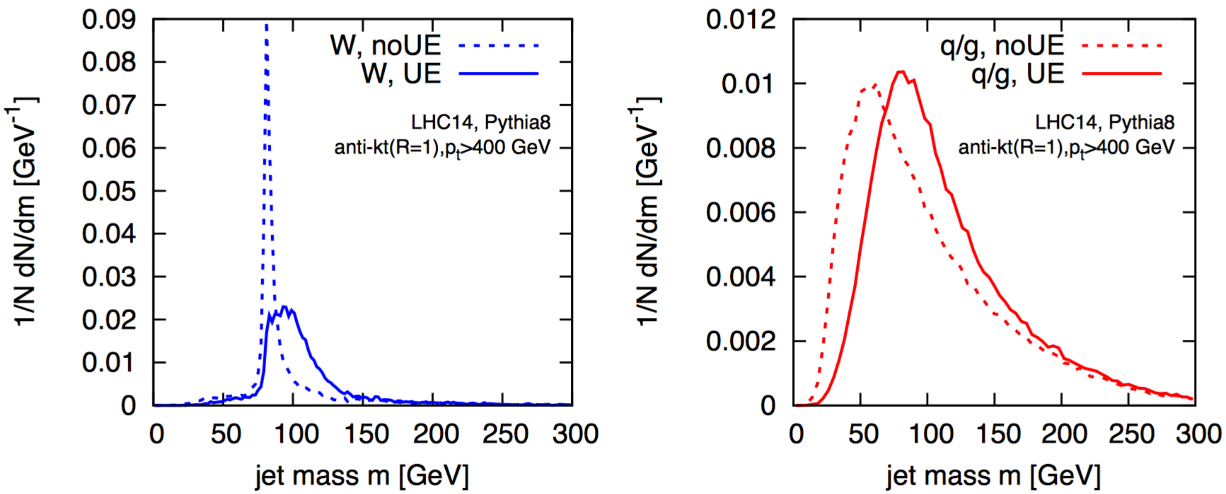
MANCHESTER 1824

The University of Manchester

Need for grooming

Fat Jets

One usually work with large-R jets ($R \sim 0.8 - 1.5$)
$\Rightarrow$ large sensitivity to UE (and pileup)



Example of groomer is filtering used in BDRS method. Most tools including mDT pruning and filtering both tag and groom. We can collectively use the name taggers.

Some open questions

Given the limited number of main ideas involved in tagging and grooming we can ask:

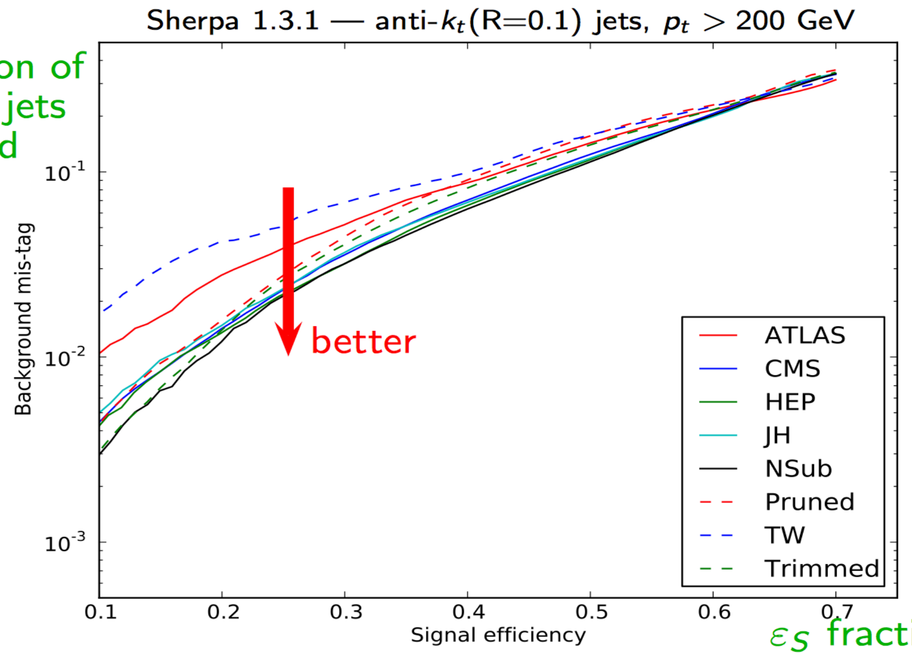
- Why so many methods?
- Are they really different?
- How to compare methods: number of parameters, vast kinematic range?
- Are tools robust? What is the connection to QCD predictions?

Monte Carlo studies alone are insufficient to provide detailed answers to these and other questions.

Monte Carlo studies

[Boost 2011 proceedings]

ε_B fraction of QCD jets tagged



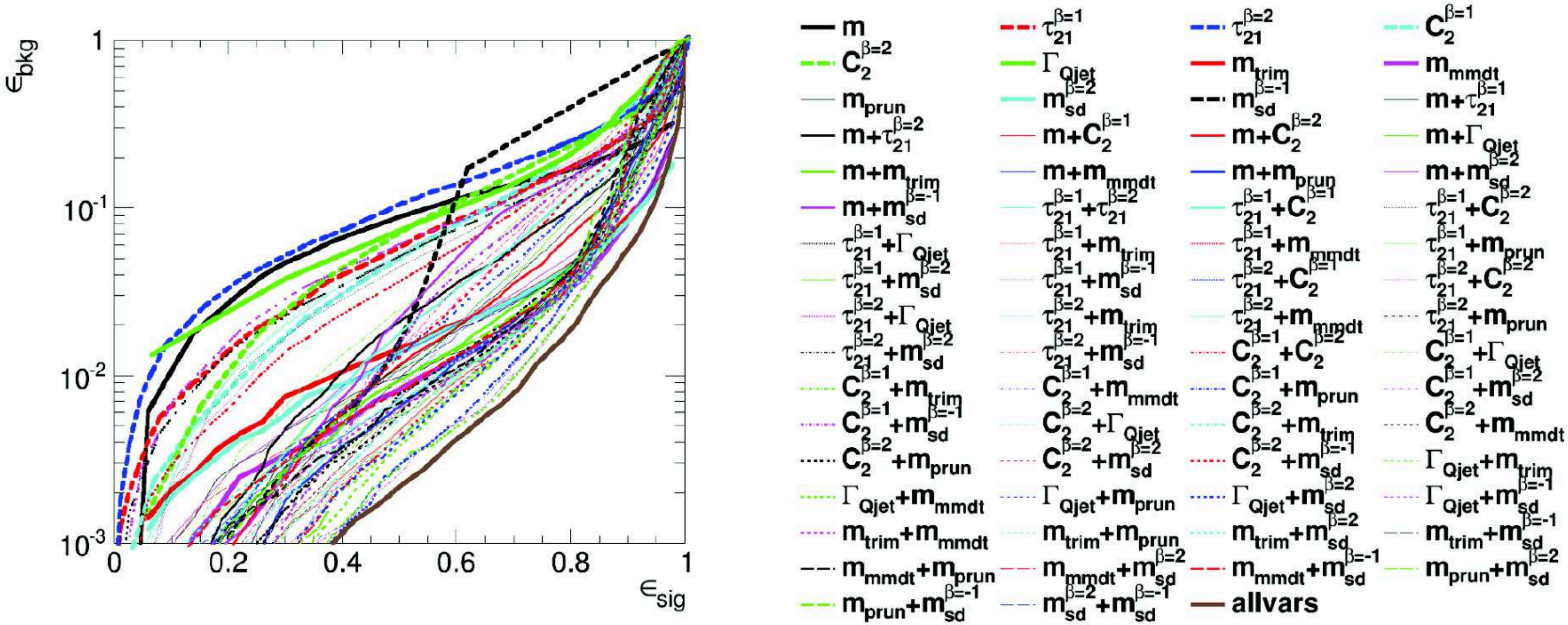
ε_S fraction of top jets tagged

Studies are for fixed parameter settings. No idea about why something works better or if picture changes with parameters.

More games with Monte Carlo

[Boost 2013 WG]

W v. q jets: combination of "2-core finder" + "radiation constraint"



Combinations help but details far from obvious.

A theoretical framework?

- Can we go back to basics? Understand the results from first principles of QCD?

Schwartz, Boost 2012

Precision QCD

- Or is that impossible?

Building a framework

A key observable is the jet mass distribution since a jet mass cut is often the first step in tagging.

Let us start by computing just the plain jet mass.

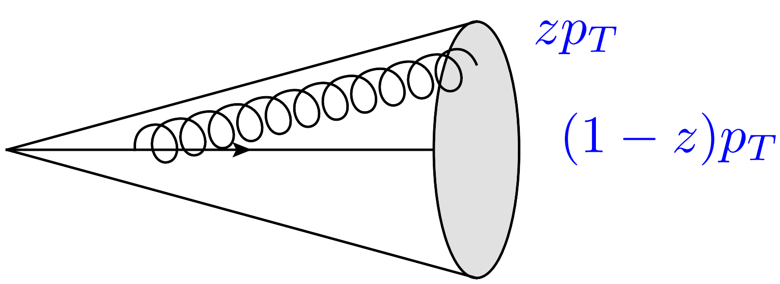
In boosted regime p_t >> m so there are large logarithms of p_t/m.

Introduce variable $\rho = \frac{m^2}{R^2 p_t^2}$ invariant under boosts along jet axis.

Want to compute $\frac{1}{\sigma}\frac{d\sigma}{d\rho}$ in soft-collinear limit valid for $\rho \ll 1$

Plain mass at LO

In leading order QCD in soft-collinear limit we have (for a quark jet)



$$m^2 = 2z^2 p_T^2 \theta^2 \left(1-\cos\theta\right) \approx z p_T^2 \theta^2$$

$$\frac{1}{\sigma}\frac{d\sigma}{d\rho} = \frac{C_F \alpha_s}{\pi}\int \frac{dz}{z}\frac{d\theta^2}{\theta^2}\delta\left(\rho - z\theta^2/R^2\right)\Theta\left(R^2-\theta^2\right)$$

Plain mass at LO

To correct for hard collinear radiation use full splitting function

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

The result is (check yourself)

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho} \simeq \frac{\alpha_s C_F}{\pi}\left(\ln\frac{1}{\rho} - \frac{3}{4}\right)$$

Leading "double logarithm".

Next to leading "single logarithm"

MANCHESTER 1824

The University of Manchester

Resummation for jet mass

- Factorisation of multi-gluon emission in soft-collinear limit

$$\frac{1}{n!}\prod_i \frac{C_F \alpha_s\left(z_i \theta_i p_t^{\text{jet}}\right)}{\pi}\frac{dz_i}{z_i}\frac{d\theta_i^2}{\theta_i^2}$$

Classical nature of soft brehmsstrahlung

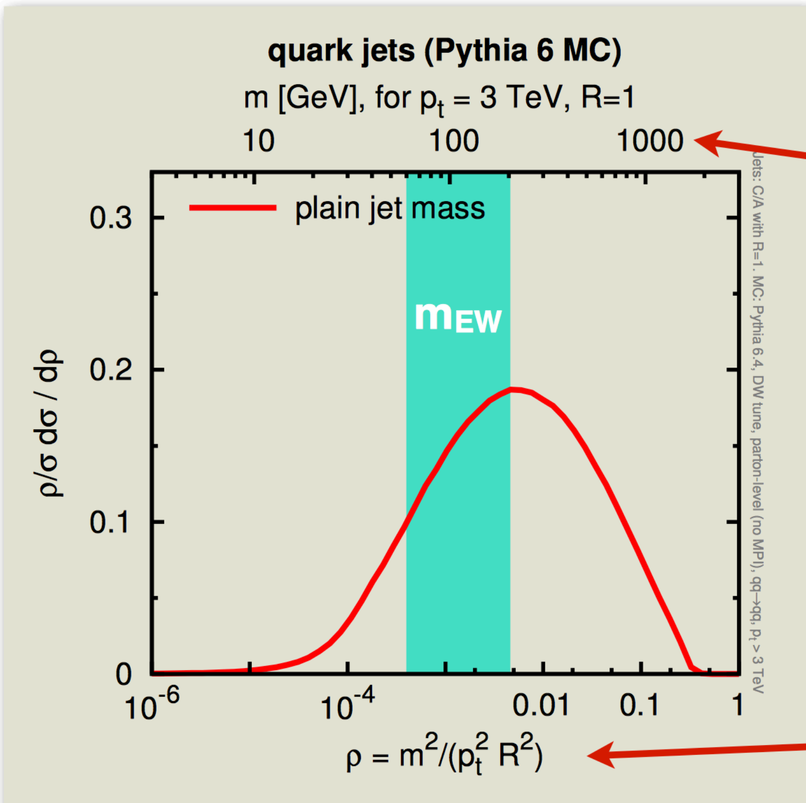
- Unitarity for including virtual corrections

Fixed-coupling resummed result is :

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho} \simeq \frac{\alpha_s C_F}{\pi}\left(\ln\frac{1}{\rho}-\frac{3}{4}\right) e^{-\frac{\alpha_s C_F}{2\pi}\left(\ln^2\frac{1}{\rho}-\frac{3}{2}\ln\frac{1}{\rho}+\mathcal{O}(1)\right)}$$

The University of Manchester

Resummation for jet mass

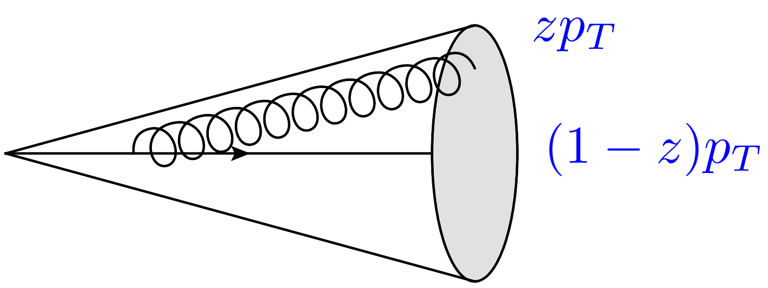


Physical mass for 3 TeV, R=1 jets

ρ ~ Rescaled mass^2
(i.e. the QCD variable)

Resummation gives a good general picture of the main features of the jet mass distribution including Sudakov peak. Can we do the same for taggers?

The mass drop tagger (BDRS)



Let us try and do a LO calculation for the mDT for QCD background jets.

At LO the mass drop condition is automatically satisfied.

The asymmetry condition gives $z, 1-z > y_{\text{cut}}$

MDT at leading order

The leading order result for MDT is

Asymmetry cuts

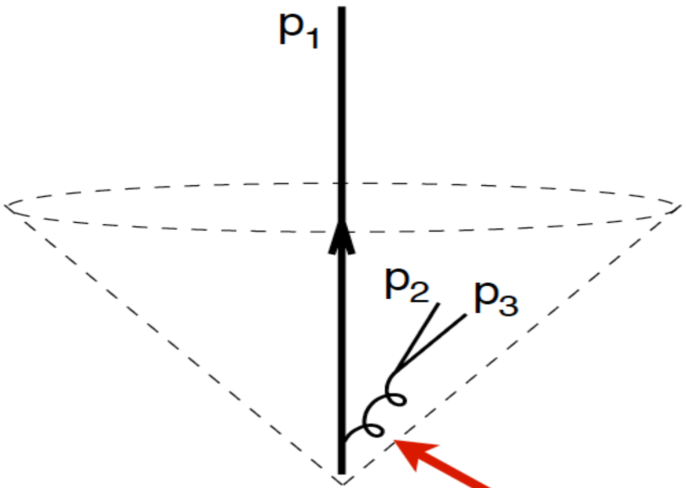
$$\frac{1}{\sigma}\frac{d\sigma}{d\rho} = C_F \frac{\alpha_s}{\pi} \int \frac{d\theta^2}{\theta^2}\frac{dz}{z} \Theta\left(z - y_{\mathrm{cut}}\right) \Theta\left(1 - y_{\mathrm{cut}} - z\right) \delta\left(\rho - z\theta^2/R^2\right)$$

One can also include hard collinear effects by using the full splitting function as for jet mass.

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho}^{(\mathrm{MDT,\ LO})} = \frac{\alpha_s C_F}{\pi}\left[\Theta(\rho - y_{\mathrm{cut}})\ln\frac{1}{\rho} + \Theta(y_{\mathrm{cut}} - \rho)\ln\frac{1}{y_{\mathrm{cut}}} - \frac{3}{4}\right].$$

The mDT has a single logarithmic behaviour at small ρ.
mDT reduces background by replacing $\ln\rho$ by modest $\ln y_{\mathrm{cut}}$

Beyond LO and a flaw in MDT



What MDT does wrong beyond LO:

Follows a soft branch ($p_2+p_3 < y_{cut}\, p_{jet}$) with "accidental" small mass, when the "right" answer was that the (massless) hard branch had no substructure

Subjet is soft, but has more substructure than hard subjet

MDT's leading logs (LL, in Σ) are:

$$\alpha_s L,\ \alpha_s^2 L^3,\ \ldots .\ \text{I.e. } \boldsymbol{\alpha_s^n L^{2n-1}}$$

quite complicated to evaluate

All orders results for mMDT

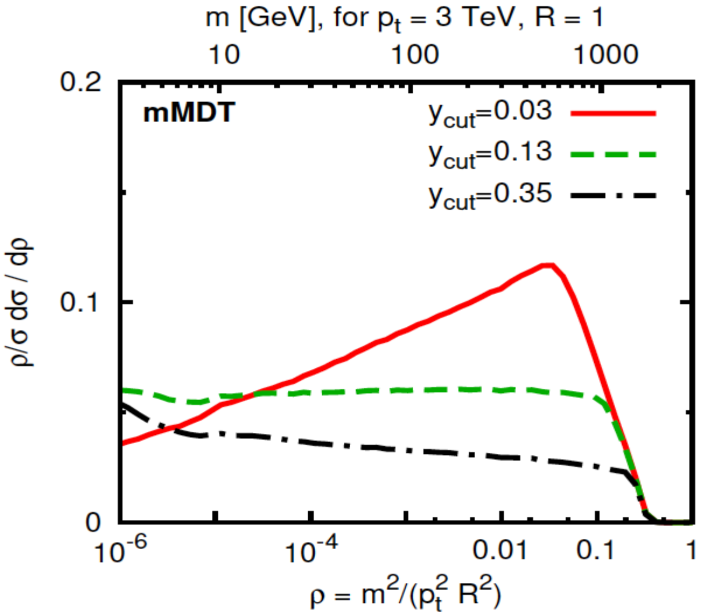
Possible after we modify the mass-drop to follow harder rather than more massive branch

$$\left(\rho\frac{d\sigma}{d\rho}\right)^{\mathrm{fixed-coupling}} = \rho\frac{\partial}{\partial\rho}\exp\left[-C_F\frac{\alpha_s}{\pi}\left(\ln\frac{1}{y_{\mathrm{cut}}}-\frac{3}{4}\right)\ln\frac{1}{\rho}\right] \quad \rho < y_{\mathrm{cut}}$$

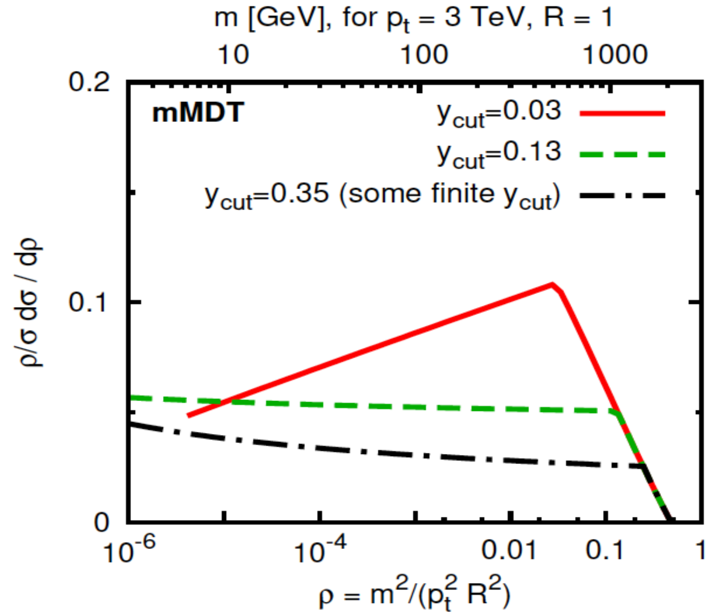
- Transition to plain jet mass at large masses.
- Only collinear logs
- First time a jet observable of this type was ever seen.
- No dependence on mass-drop cut but only on asymmetry parameter.

Comparison to MC

Monte Carlo

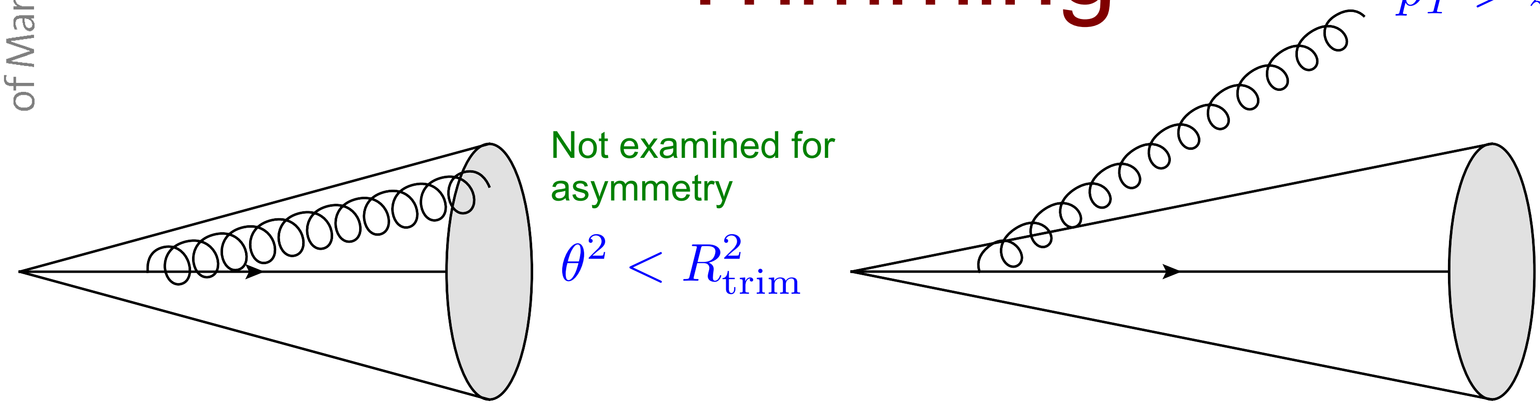


Analytic



Excellent agreement of analytic and MC results indicate we have captured the relevant physics with our simple formulae.

Trimming



3 distinct regions seen

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho} = \frac{C_F\alpha_s}{\pi}\;\ln\frac{1}{\rho} - \frac{3}{4} \qquad \rho > z_{\rm cut}$$

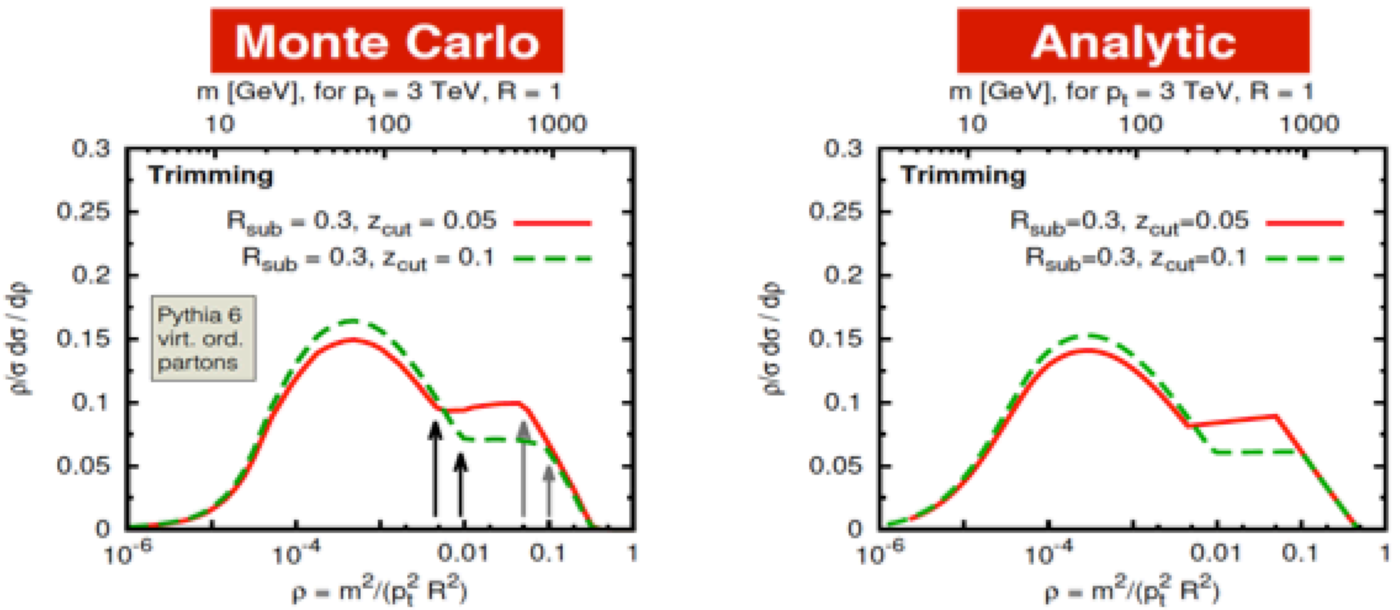
$$= \frac{C_F\alpha_s}{\pi}\;\ln\frac{1}{z_{\rm cut}} - \frac{3}{4} \qquad \rho < z_{\rm cut}$$

$$= \frac{C_F\alpha_s}{\pi}\;\ln\frac{z_{\rm cut} r^2}{\rho} \qquad \rho < z_{\rm cut} r^2 \qquad r = \frac{R_{\rm trim}}{R}$$

This leading order result exponentiates at all orders.

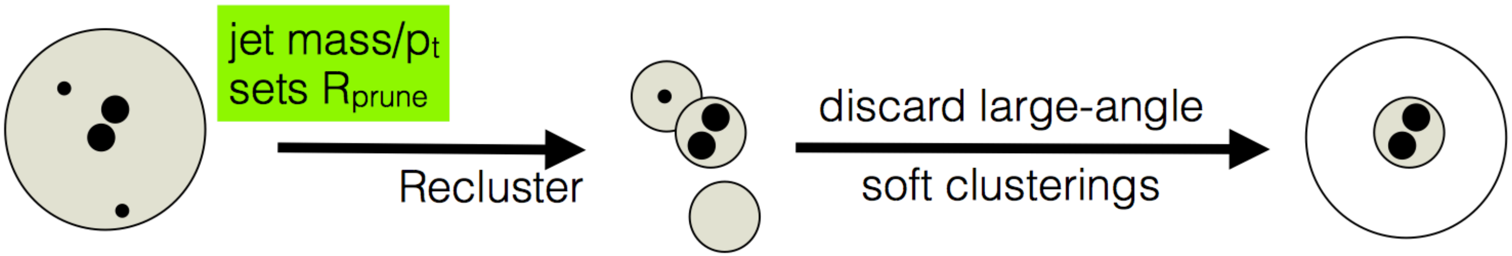
The University of Manchester

All-orders v Monte Carlo



- Excellent overall agreement which captures the dependence on parameters and transition points.
- Indicates flaws in existing methods. Reveals distinct regimes for tagger behaviour. Taggers can be worse than doing nothing!
- Undesirable behaviour lies in region relevant for pheno.

Pruning



Similar to trimming but with dynamical radius choice.

Pruning results

LO result is single logarithmic like (m)MDT.

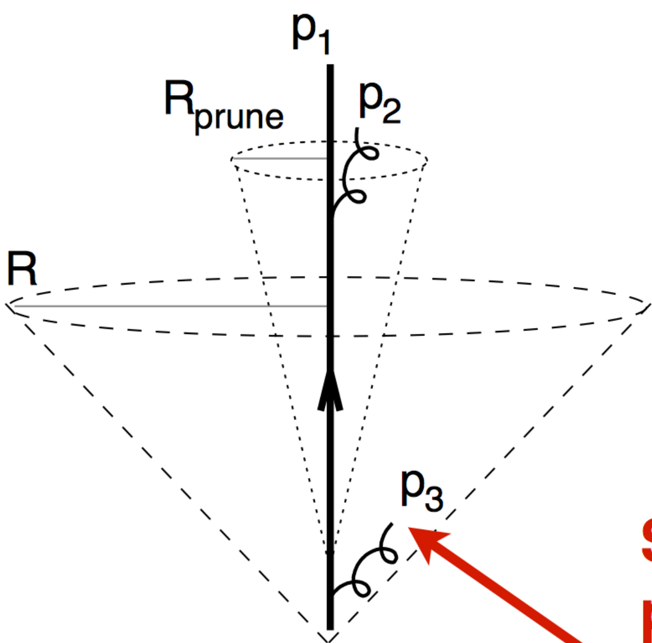
$$\rho \frac{d\sigma}{d\rho} \sim \alpha_s \ln \frac{1}{z_{\mathrm{cut}}}$$

However at NLO one encounters terms as singular as the plain jet mass i.e. double logarithms.

$$\rho \frac{d\sigma}{d\rho} \sim \alpha_s^2 \ln^3 \frac{1}{\rho}$$

Pruning can be thought of as a sum of two distinct components.

Pruning results



What pruning sometimes does

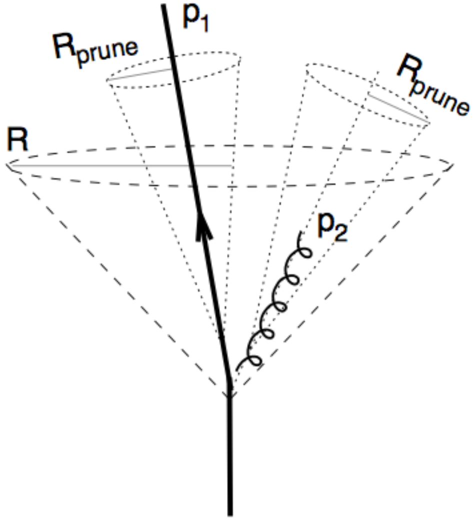
Chooses R_{prune} based on a soft p_3 (dominates total jet mass), and leads to a single narrow subjet whose mass is also dominated by a soft emission (p_2, within R_{prune} of p_1, so not pruned away).

Sets pruning radius, but gets pruned away → "wrong" pruning radius → makes this ~ trimming

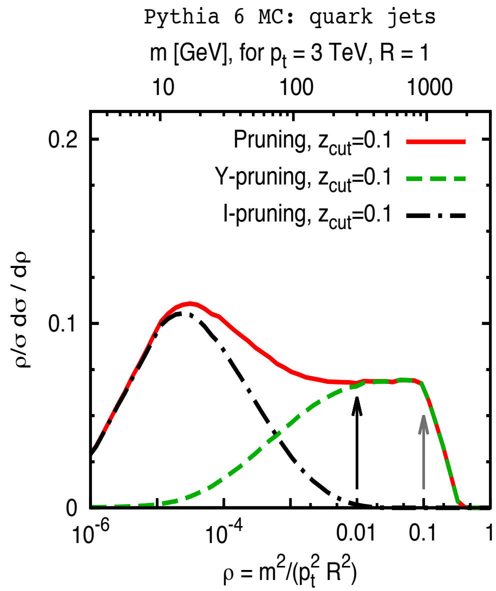
Y- Pruning

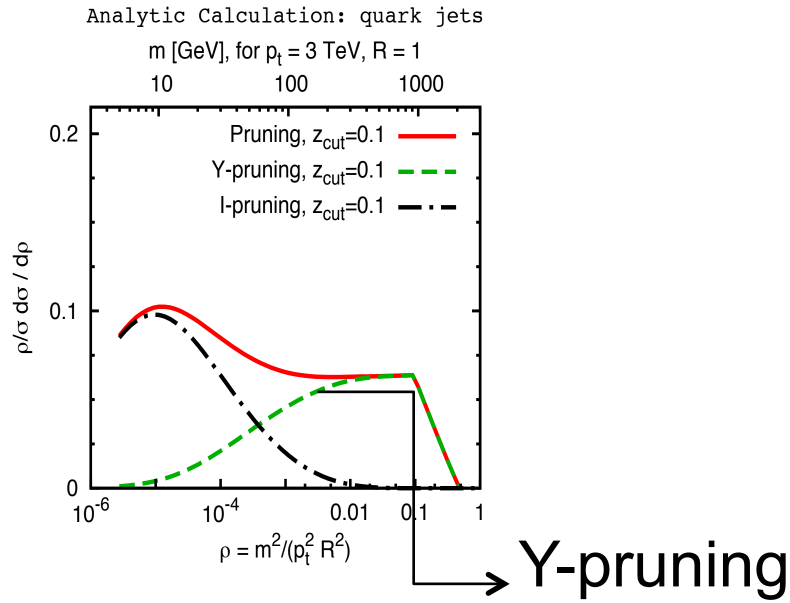
A simple fix: “Y” pruning

Require at least one successful merging with $\Delta R > R_{prune}$ and $z > z_{cut}$ — forces 2-pronged (“Y”) configurations



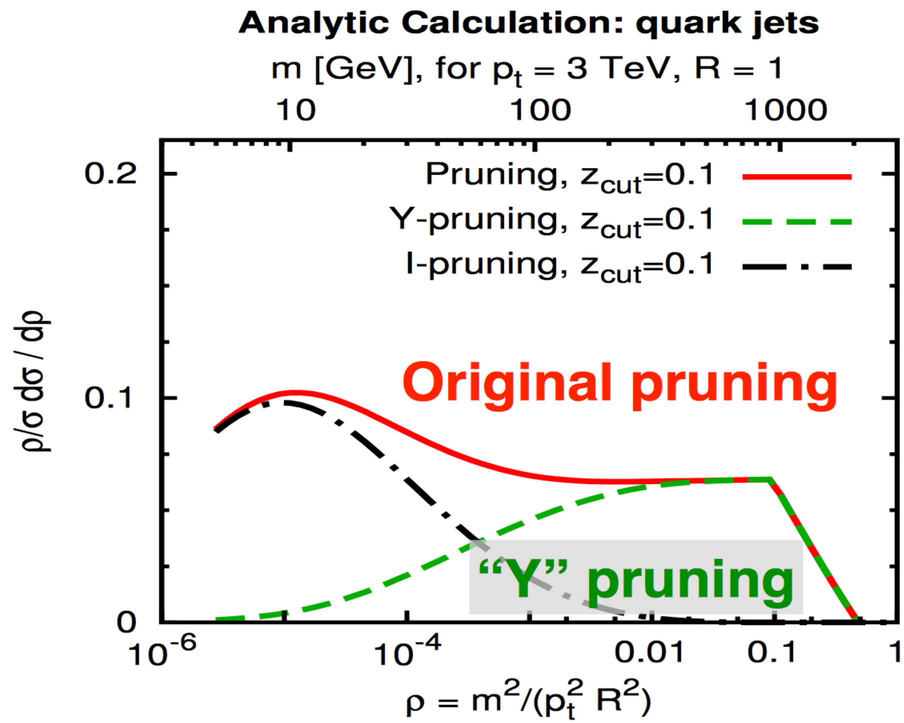
Pruning MC v analytics





The black line denotes the anomalous component (I-pruning). The green line is the sane component (Y-pruning).

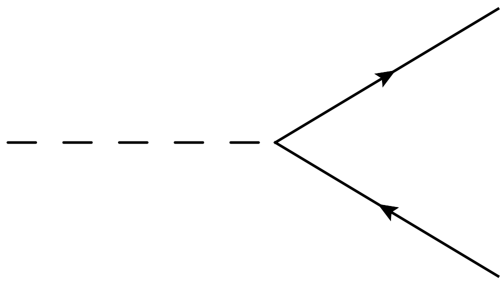
Y-pruning



Plain mass Sudakov exponent

$$\frac{\rho}{\sigma}\frac{d\sigma^{\mathrm{Y-prune}}}{d\rho} \sim \frac{C_F \alpha_s}{\pi}\left(\ln\frac{1}{z_{\mathrm{cut}}}\right)\exp\left(-\frac{C_F\alpha_s}{2\pi}\ln^2\rho\right)$$

Signal jets

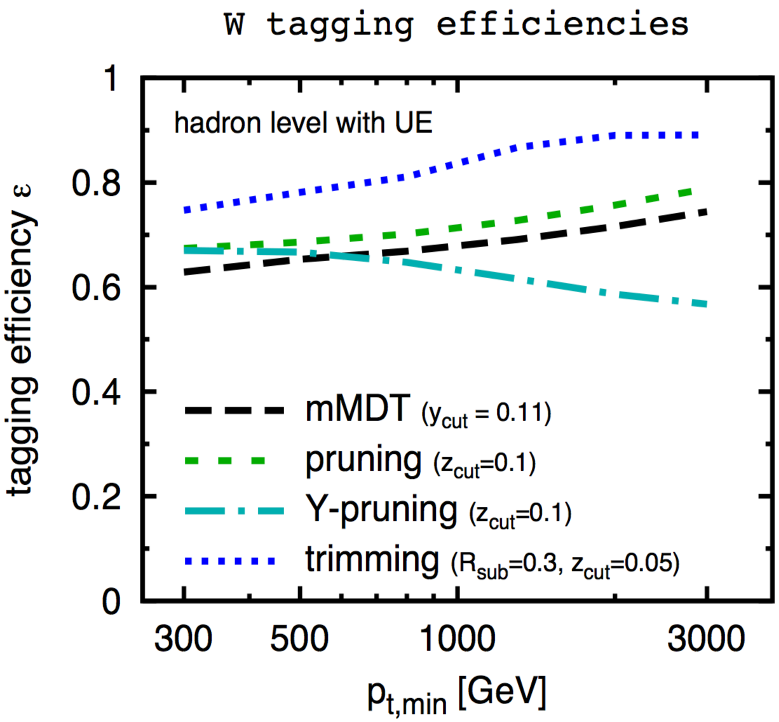


Also possible to analytically understand signal jets.

$$\epsilon_s = \int_{y_{\rm cut}}^{1-y_{\rm cut}} dz = 1 - 2y_{\rm cut}$$

Tree level result for mMDT and pruning receives only modest higher order corrections.

Signal efficiencies with taggers



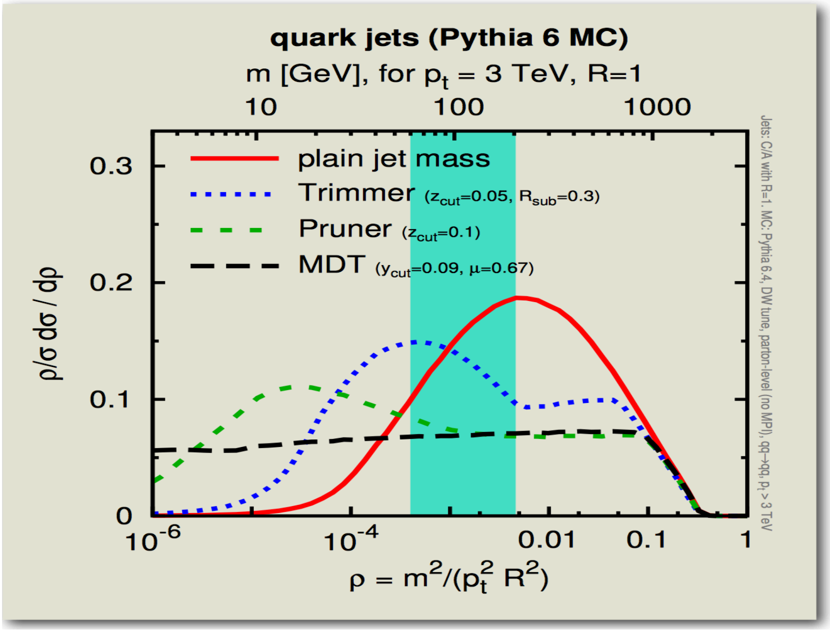
Tree level is a good approximation with small effects from ISR and FSR effects.

Y-pruning suffers a loss of efficiency at high p_t

All this also understood analytically and with MC studies.

Dasgupta, Powling and Siodmok 2015

What does all this buy us?

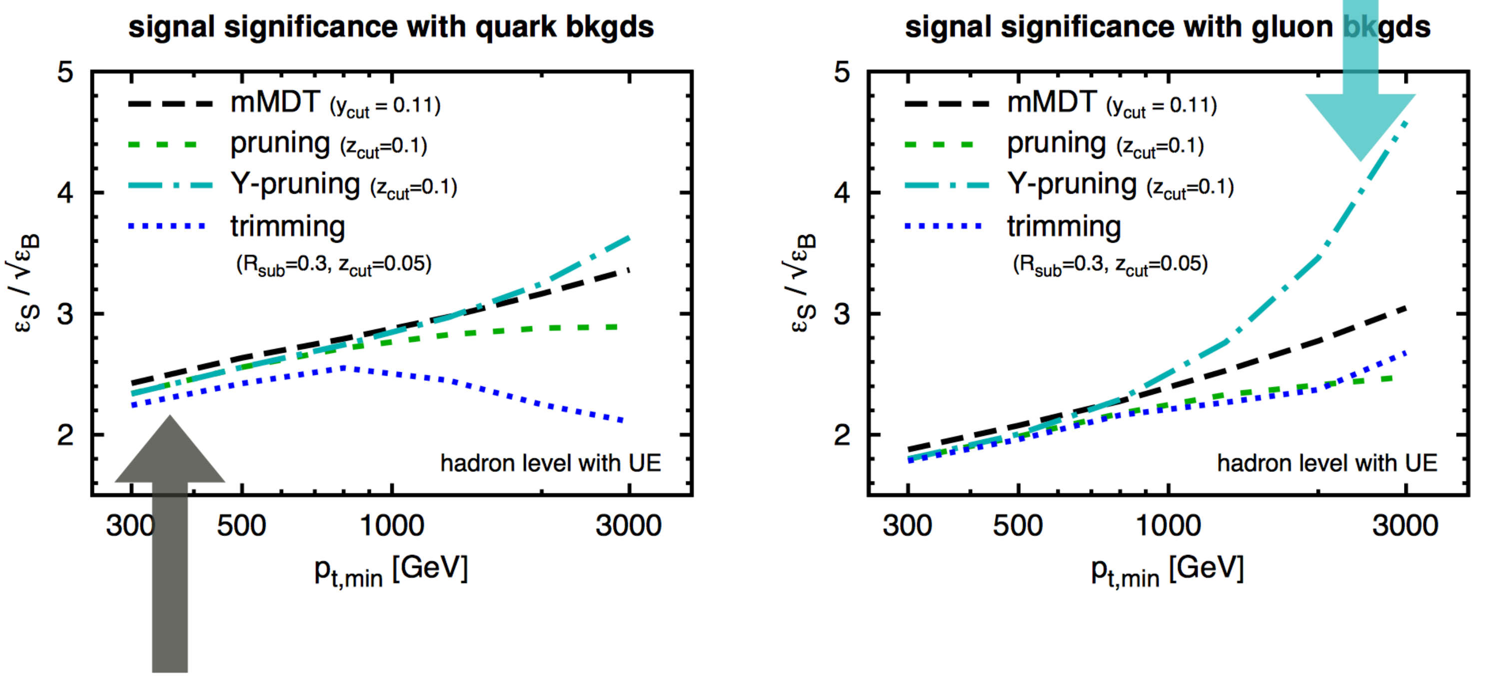


No longer need to run Monte Carlo blindly.

Can do “the right” MC studies to meaningfully compare tools and bring out their main features.

Performance for finding signals

At high p_t, substantial gains from new Y-pruning
(probably just indicative of potential for doing better)



At low p_t (moderate m/p_t), all taggers quite similar

Summary

- Analytical insight into jet substructure proving to be a powerful complement to MC studies.
- This insight is helping to compare tools, assess robustness and to design better tools such as mMDT and Y-pruning.
- The basic message remains the same as that from yesterday's lecture : calculations based on QCD principles can often go a long way in providing information complementary to that from MC and other studies.