MANCHESTER 1824

The University of Manchester

Recent progress in jet substructure studies

Mrinal Dasgupta
University of Manchester

LHCP 2016 Lund, 14th June 2016

Based on work with Salam, Marzani, Fregoso, Powling, Siodmok, Soyez and Schunk

Boosted objects and jet substructure

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. Jet substructure is powerful handle on signal vs QCD background.

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

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

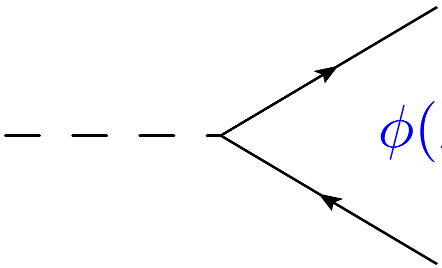
where z, $1-z$ are the momentum fractions of the two quarks. In particular one should capture the $b, \bar{b}$ and any

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

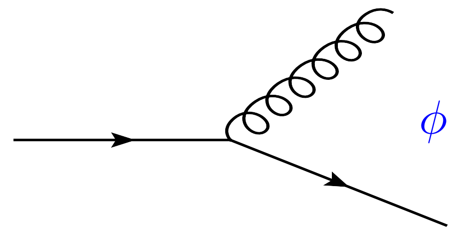
Butterworth, Davison Rubin, Salam 2008. Published in PRL. Builds on work by Seymour 1993.

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

Signal vs background



$\phi(z) \propto 1$



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

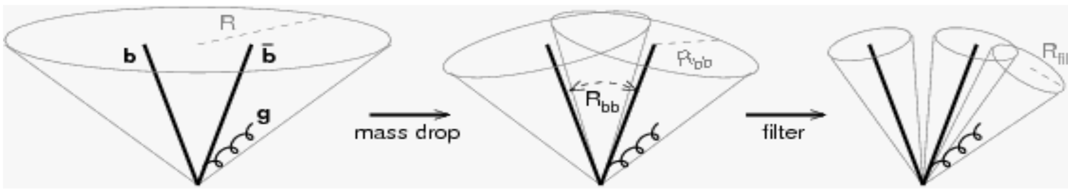
BDRS studied the process $pp \to VH,\ H \to b\bar{b}$

- 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.

The University of Manchester

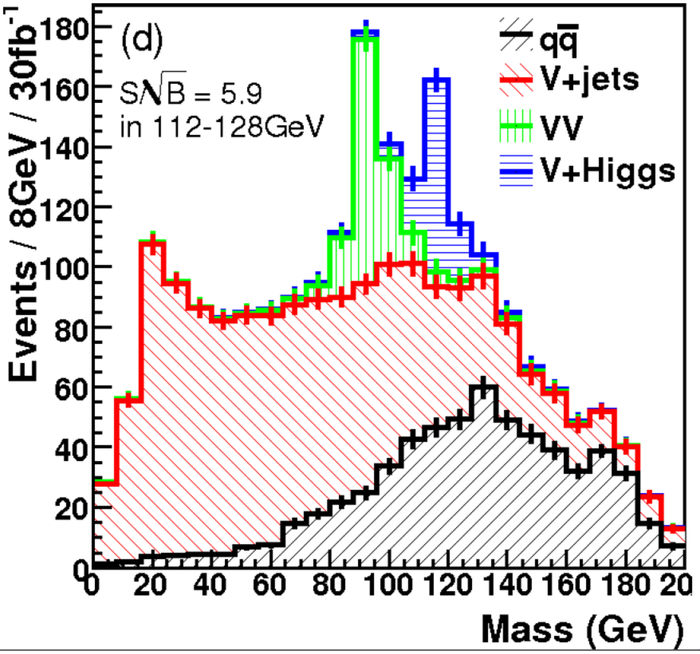
MANCHESTER 1824

BDRS mass drop+filtering



- Break the jet into two subjets j_1 and j_2 such that $mj_1 > mj_2$.
- If there is a mass drop $m_{j1} < \mu m_j$ and the splitting is not too asymmetric $y = \min\left(p_{tj1}^2, p_{tj2}^2\right) \Delta R_{j1j2}^2 / m_j^2 > y_{\rm cut}$ then deem the jet tagged or if not discard j_2 and continue.
- Also called the "mass drop" tagger (MDT). more about this later……

BDRS method results

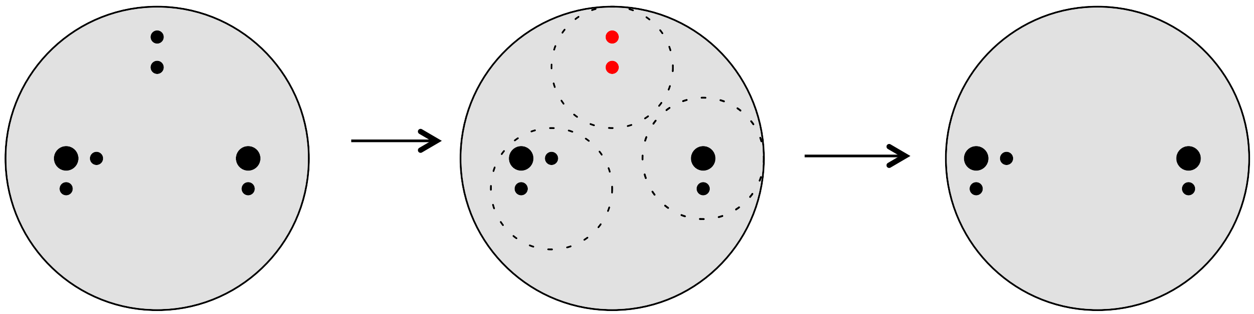


Signal significance of 4.5σ was demonstrated in MC studies for a Higgs boson of 115 GeV. Turned this unpromising channel into one of the best discovery channels for light Higgs.

MANCHESTER 1824

The University of Manchester

Several other methods exist



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

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.

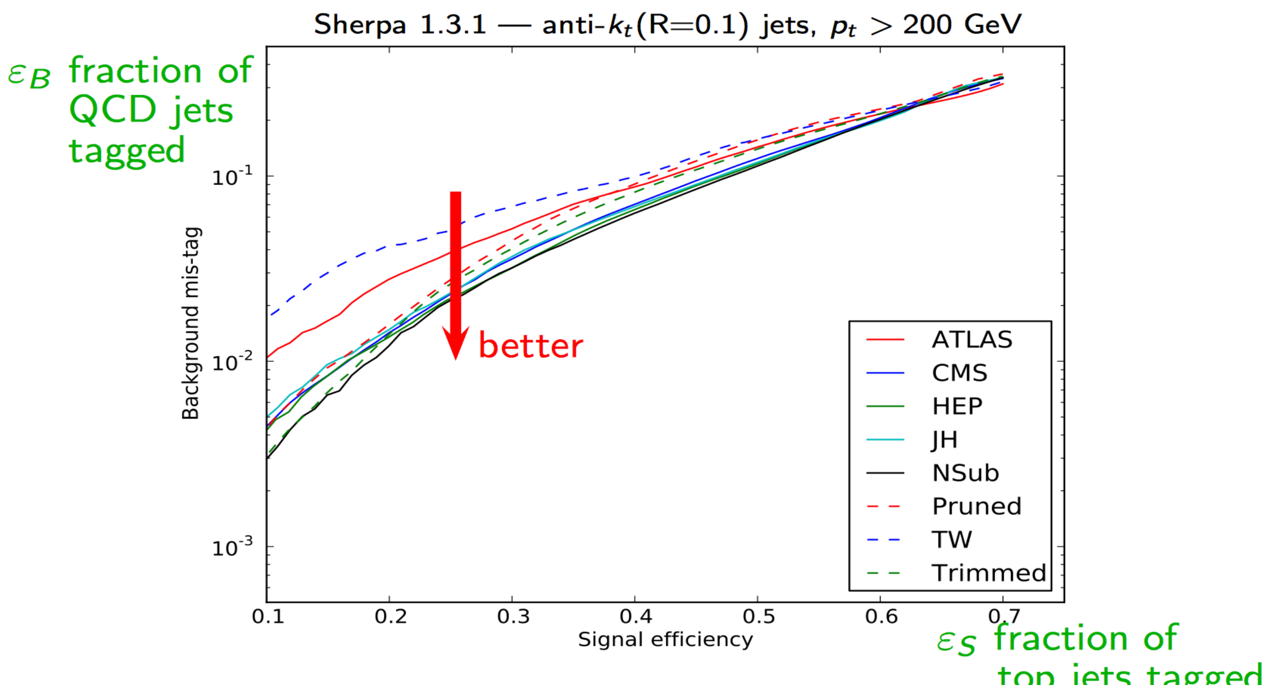
Some open questions

- 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]

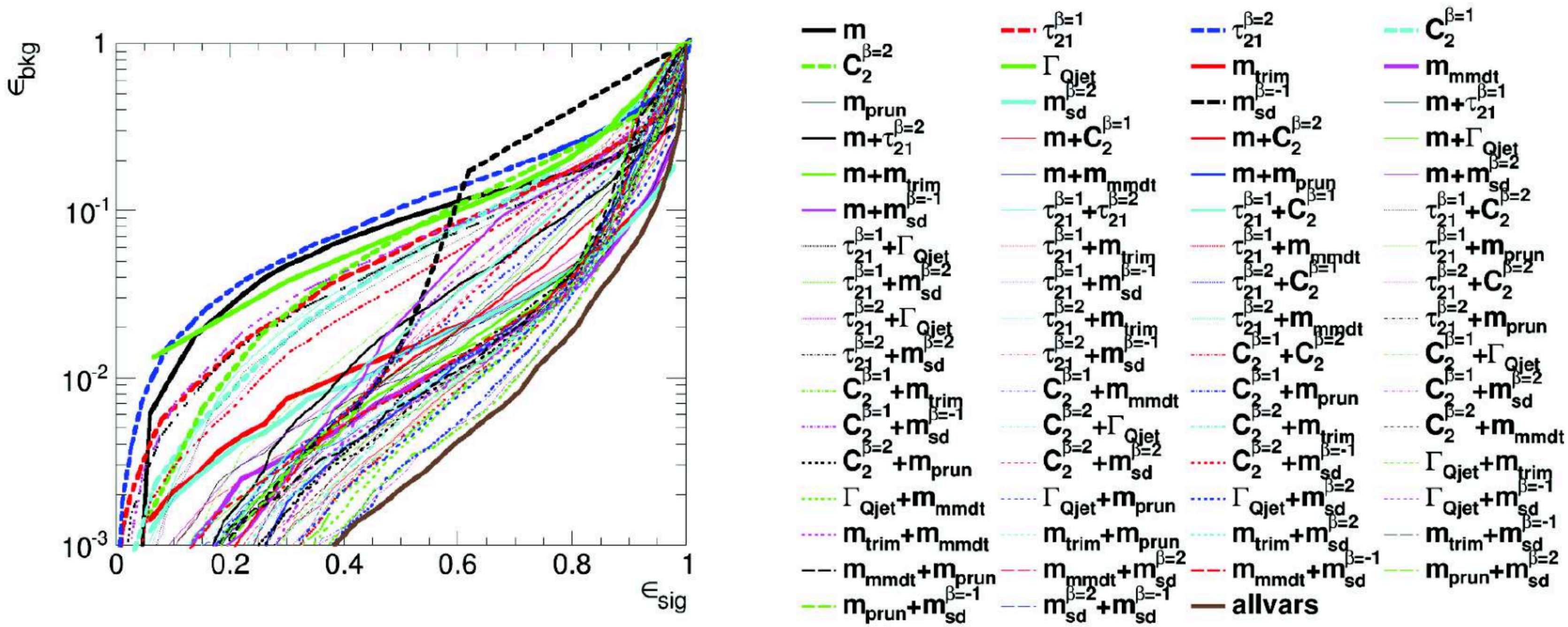


Studies are for fixed parameter settings. No idea about why something works better or if picture changes with parameters.

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?

Analytical understanding

Need a concrete link between results and underlying principles.

Putting together

- A soft/collinear approximation of QCD matrix elements for n gluon emission.

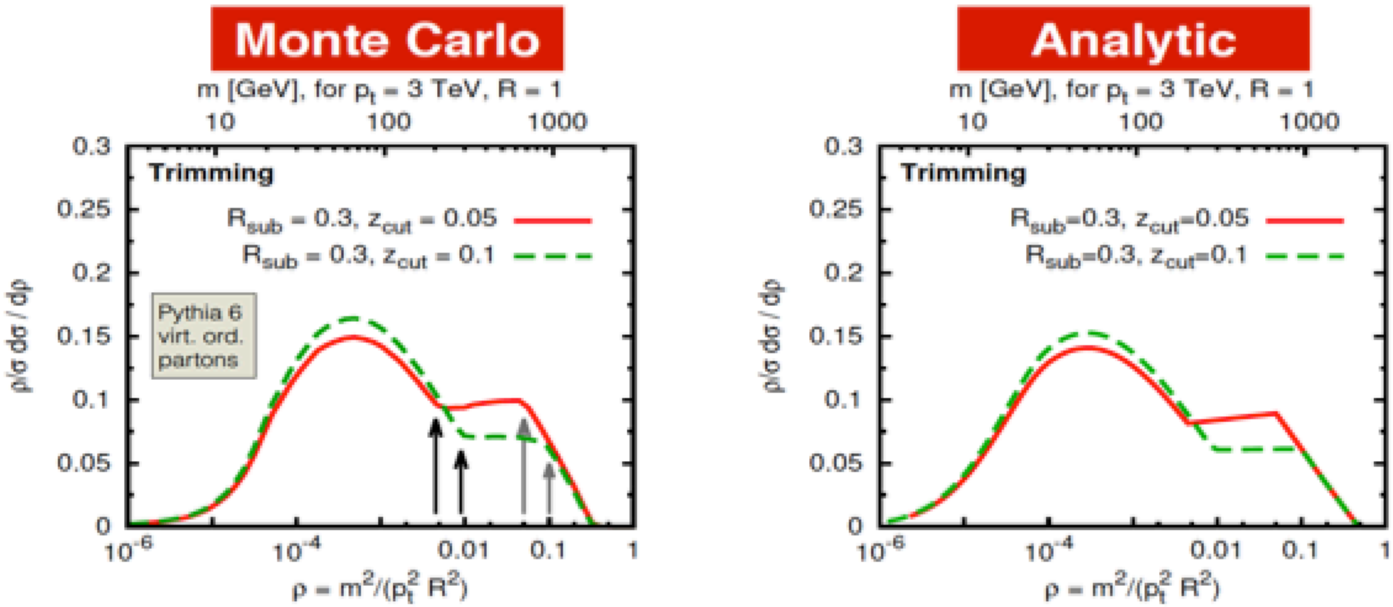
$$\frac{1}{n!}\prod_i \frac{C_F \alpha_s\left(z_i \theta_i p_t^{\text{jet}}\right)}{\pi}\frac{dz_i}{z_i}\frac{d\theta_i^2}{\theta_i^2}$$

- Virtual corrections from unitarity.

- Understanding of taggers for multiple emissions.

Dasgupta, Fregoso,Marzani, Salam 2013. Dasgupta, Fregoso, Marzani, Powling 2013.

Analytics v Monte Carlo

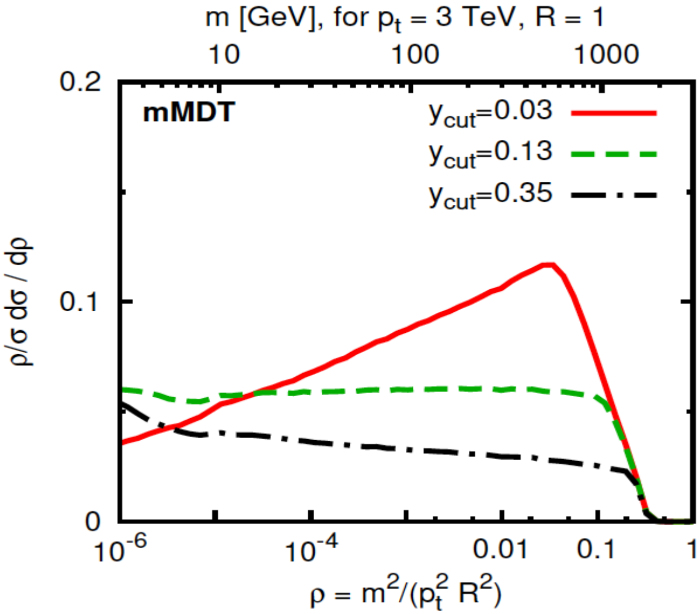


- Excellent overall agreement which captures the dependence on parameters and transition points.
- Indicates flaws in existing methods. Reveals distinct regimes for tagger behaviour.
- Undesirable behaviour in region relevant for pheno.

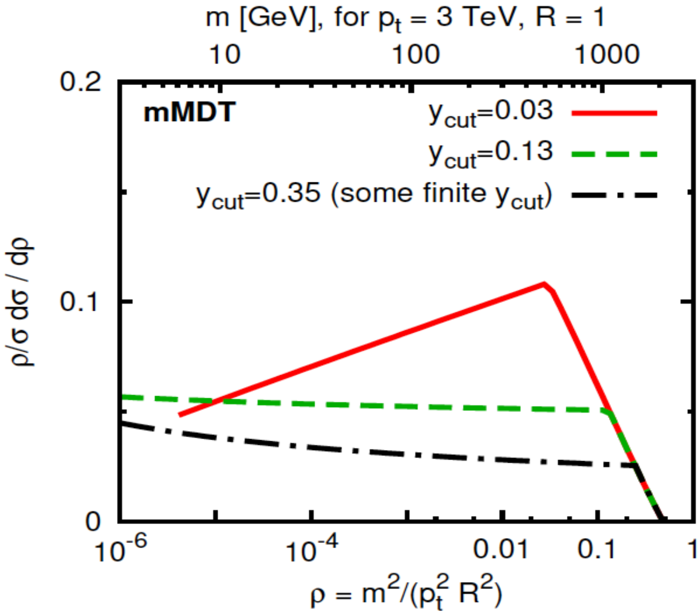
Dasgupta, Fregoso, Marzani and Salam 2013.

Analytics v Monte Carlo

Monte Carlo



Analytic



- Results for modified mass-drop method.
- Analytics show that mass-drop condition is not essential.
- Similar understanding also for pruning.

Dasgupta, Fregoso, Marzani, Salam 2013

How can this be exploited?

- More informed rather than blind MC comparisons of tools
- Paves way to understand more methods and put whole field on firmer ground.
- Developing new superior tools and improve existing methods.
- Devise variables that lend themselves to precision QCD studies.

MANCHESTER
1824

The University of Manchester

Understanding jet shapes

- Analytic understanding also achieved for other variables.

- Important category of variables are radiation constraining jet shapes.
- Examples are N-subjettiness and Energy Correlation Functions (ECFs). Thaler and Van Tilburg 2011 Larkoski, Salam, Thaler 2013

Shape variables : N subjettiness

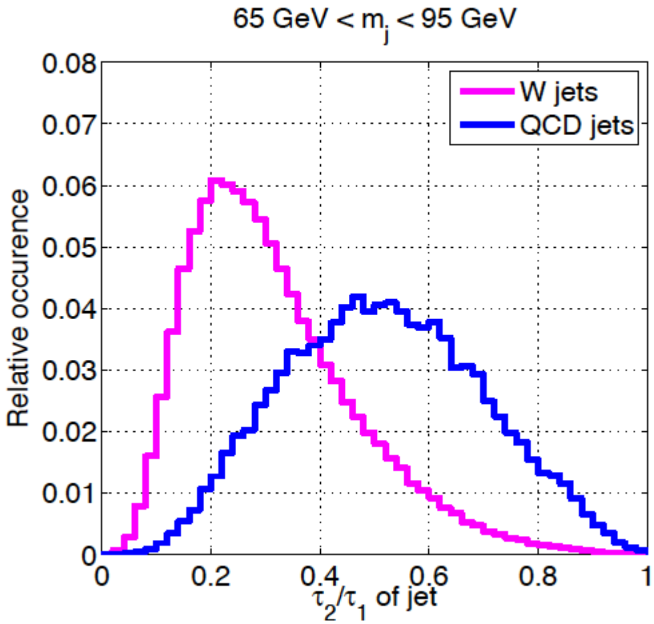
Find N prongs (axes) in a jet using e.g. kt algorithm.

$$\tau_N^{(\beta)} = \frac{1}{p_T R^\beta} \sum_{i \in jet} p_{T,i} \min\left(\theta_{i,a_1} \cdots \theta_{i,a_N}\right)$$

$$\tau_{N,N-1} = \tau_N / \tau_{N-1}$$

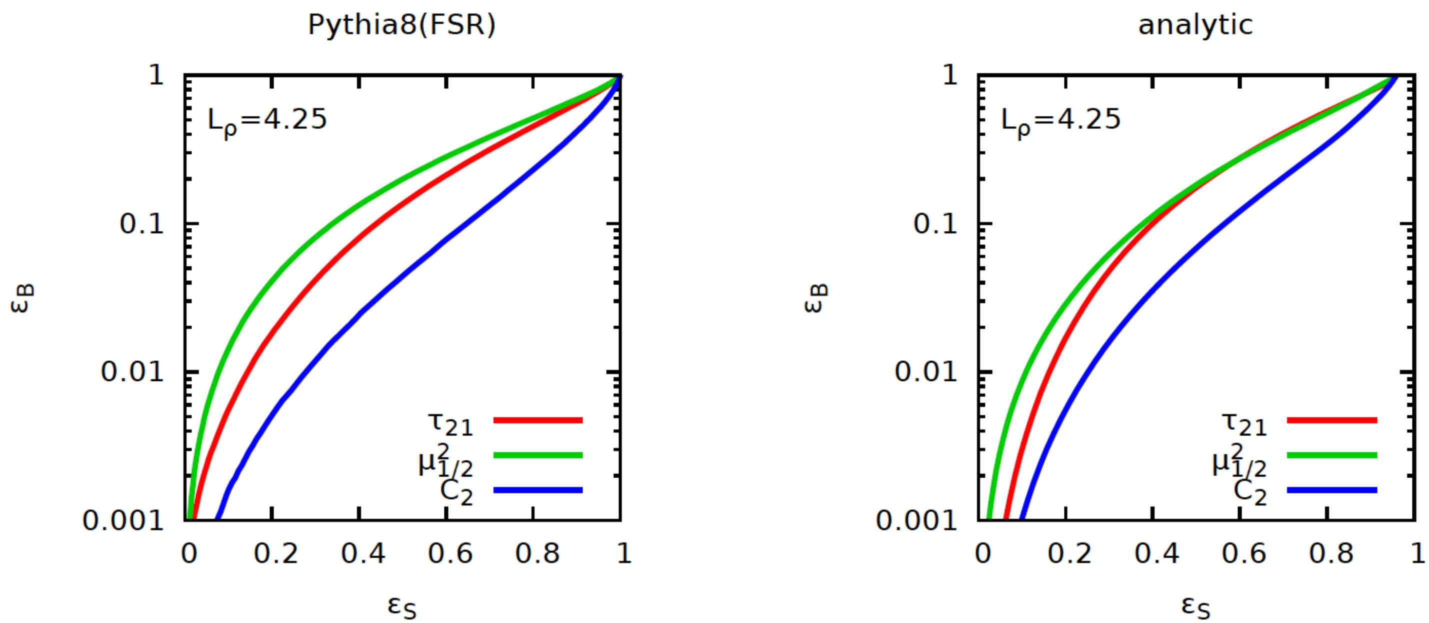
Discriminates N pronged signal from QCD

$\tau_{2,1}$ Smaller for W/Z/H than for QCD



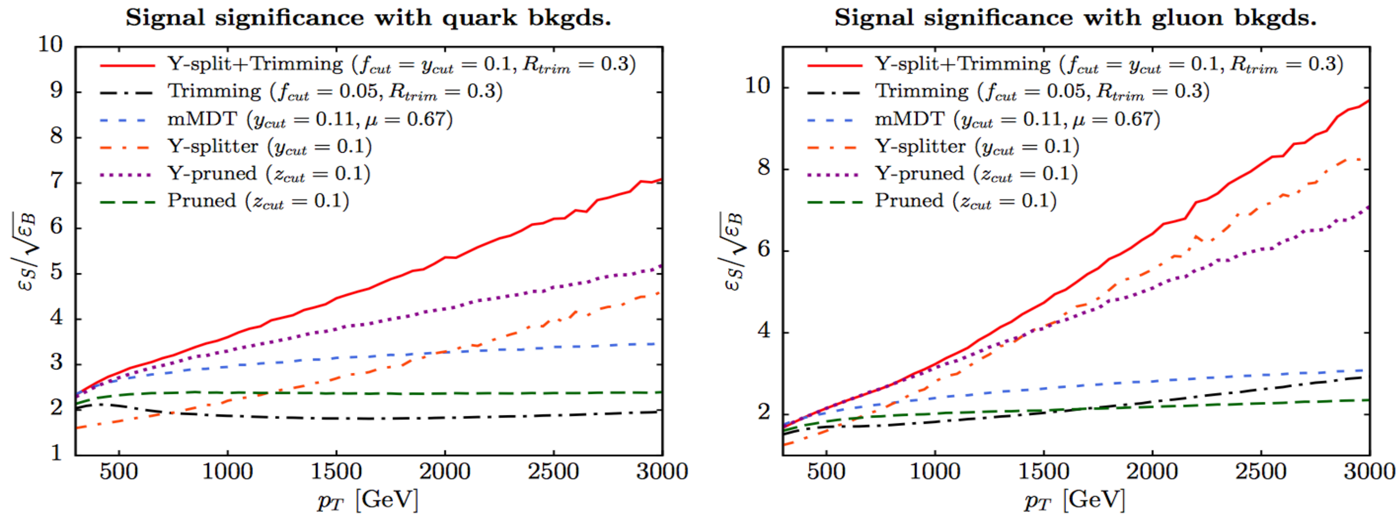
Monte Carlo v Analytic

First understanding for jet shapes:



Dasgupta, Soyez and Schunk 2015. See also Larkoski, Moult and Neill 2015 for dedicated calculation of C_2.

New tools and combinations



- Y-splitter is existing method. Analytics reveal desirable Sudakov suppression of background but poor signal efficiency.
- Combining with grooming improves signal vastly and leaves background suppression intact. Also understood analytically.

Dasgupta, Powling and Siodmok 2015 and work in progress.

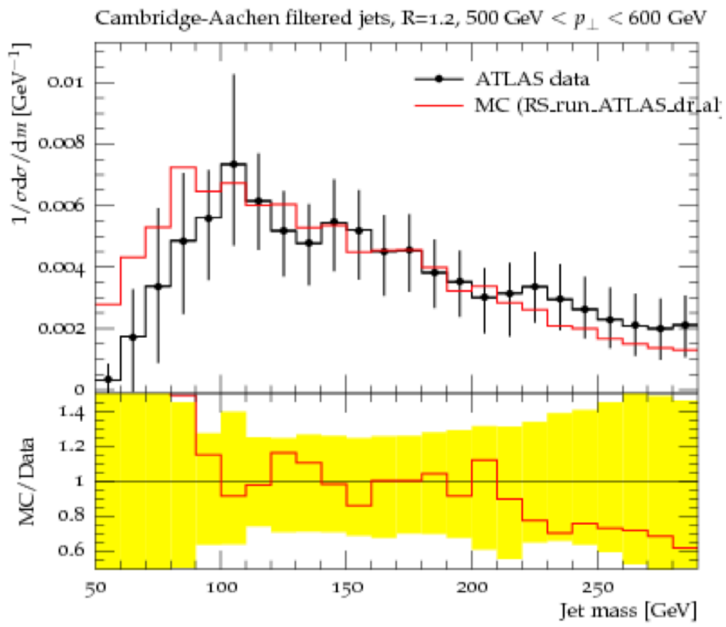
MANCHESTER 1824

The University of Manchester

Jet observables for precision QCD

Plot by

A. Siodmok



- mMDT shown to have remarkable properties. Freedom from non-global logs and minimal NP effects. Even fixed-order (LO) predictions already give good agreement with data.
- This motivated the generalisation to SoftDrop set of variables. Recently calculated to (N)NLL. Larkoski, Marzani, Soyez, Thaler 2014. Frye, Larkoski, Schwartz, Yan 2016.

Conclusion

- Recent analytical insight into jet substructure is a potential game changer.
- Full potential yet to be exploited but many interesting studies in progress.
- Hopefully these will bring substantial benefit to Run 2 of LHC and beyond.