MANCHESTER
1824
The University of Manchester

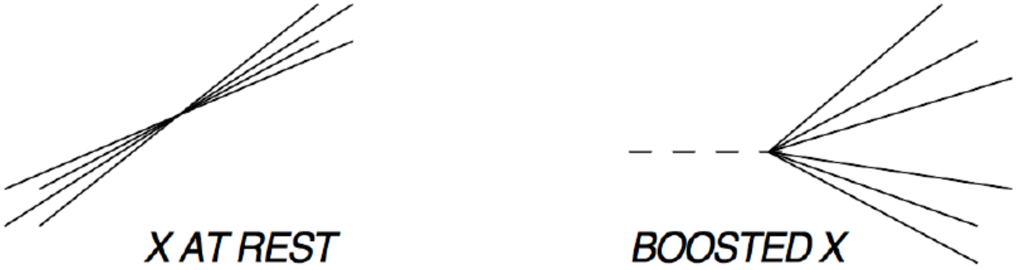
Jet Substructure : a theory perspective

QMUL, 3rd December 2020

Outline

- Introduction and basics
- Theory issues from QCD
- Recent progress
- Challenges and prospects

Basic ideas



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

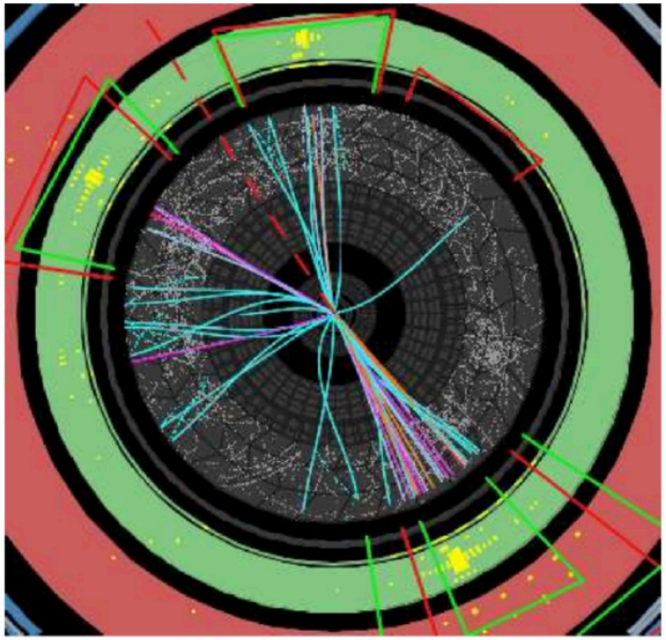
Either from going to high p_T or from decay of heavy new particle

Key idea : for tagging a particle with mass M exploit boosted regime i.e. $P_T >> M$.

Hadronic decays reconstructed in single “fat” jet. Use our knowledge of QCD jets to distinguish this from background.

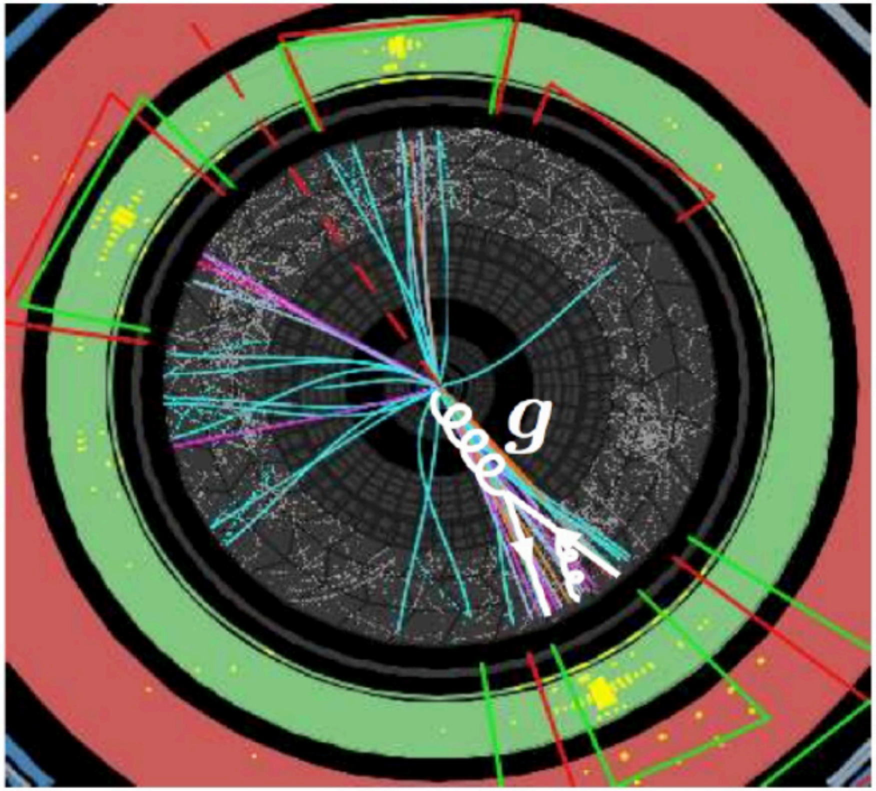
Jets from QCD vs boosted heavy particles

What jet do we have here?



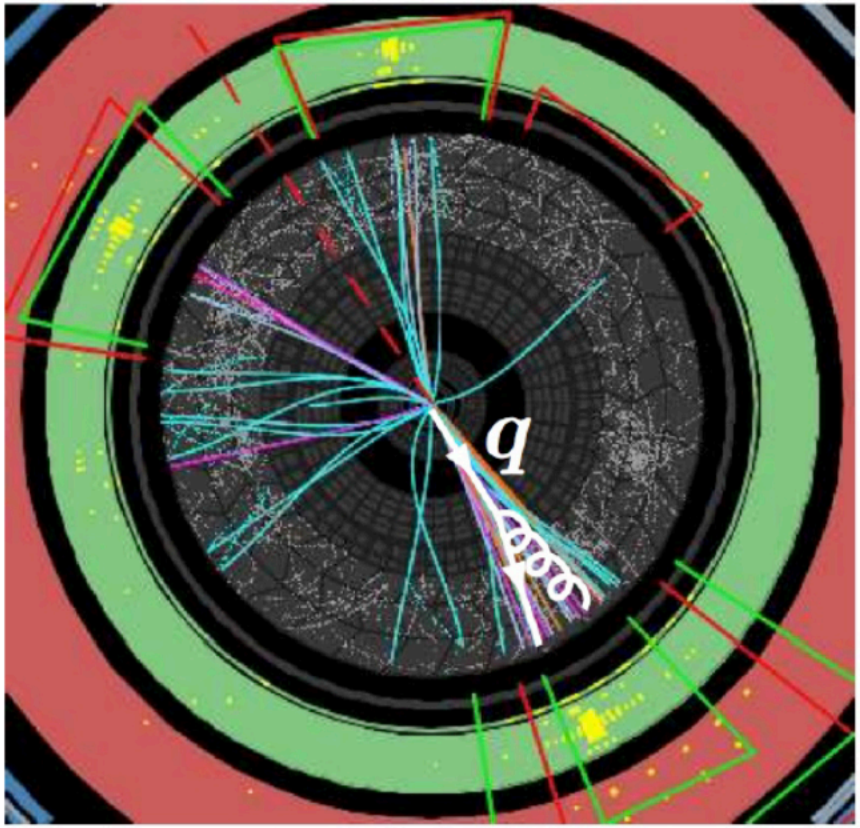
Jets from QCD vs boosted heavy particles

A gluon jet?



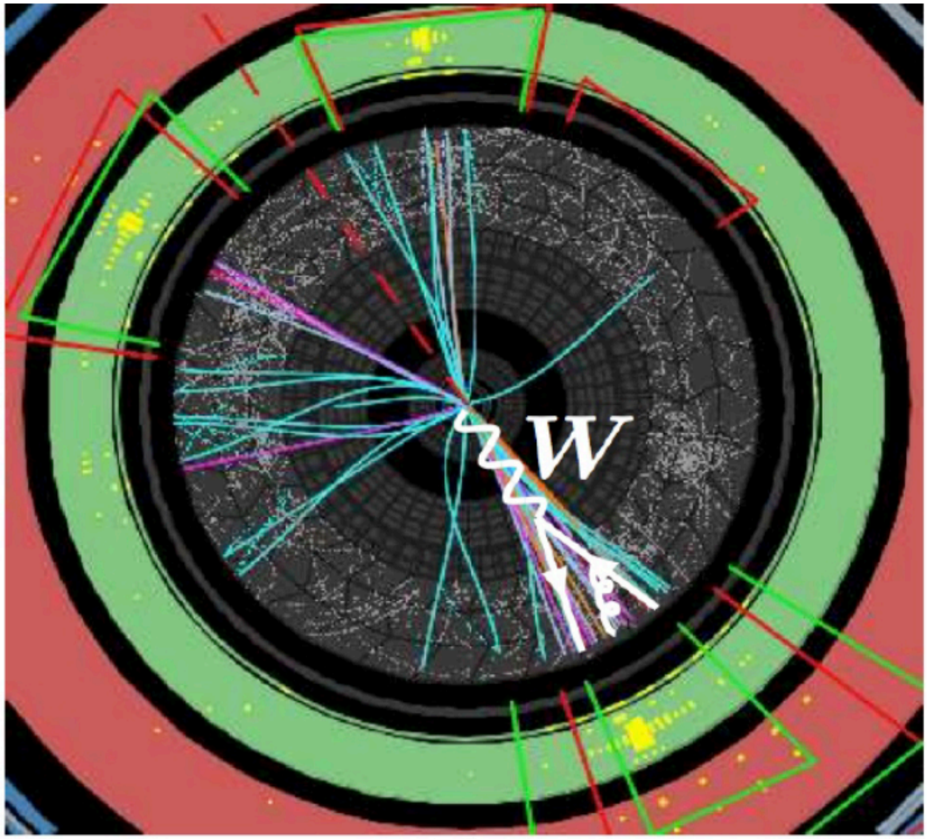
Jets from QCD v boosted heavy particles

A quark jet?



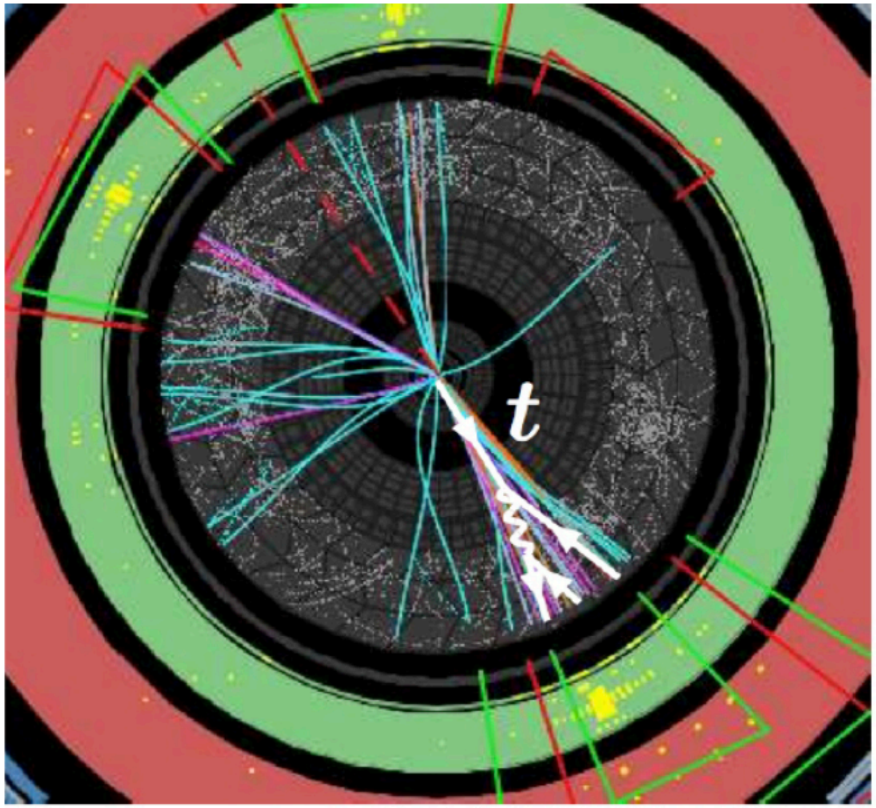
Jets from QCD vs boosted heavy particles

A W/Z/H?



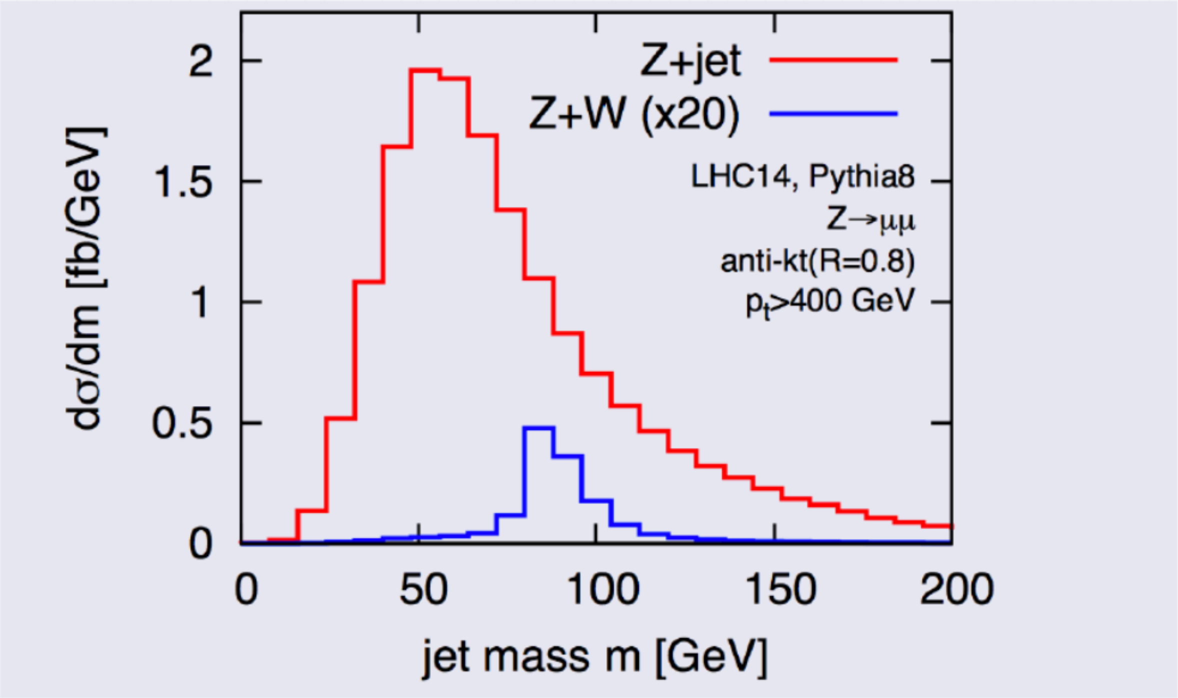
Jets from QCD vs boosted heavy particles

A top quark?



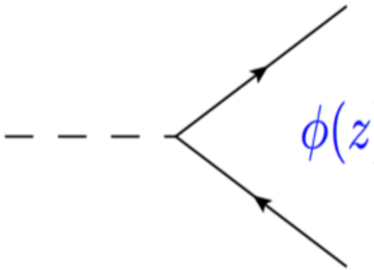
Source : ATLAS boosted top candidate

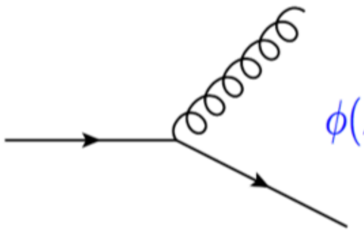
Isn’t the jet mass a clue?



Looking at jet mass is not enough!

Jet substructure and tagging





- Exploit the asymmetric nature of QCD splittings. Produce jets with single hard core or prong versus 2 pronged W/Z/H and 3 pronged t.

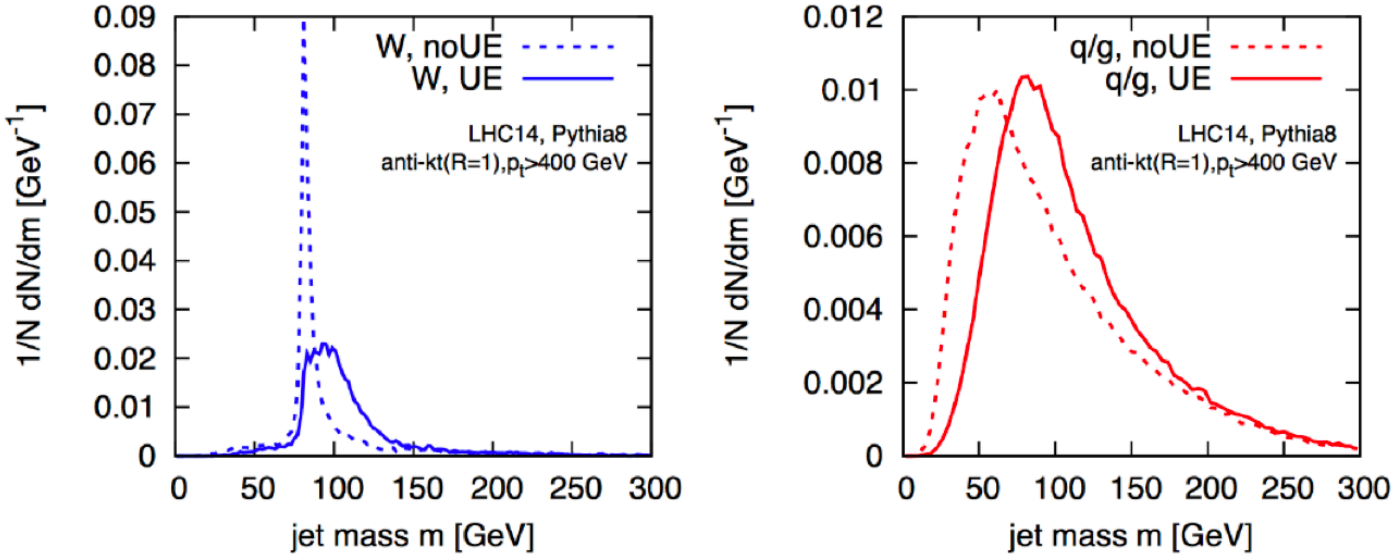
- Colour singlet nature of W/Z/H suppressing soft large angle radiation.

Pioneering work in 1991 by Seymour and 2008 by Butterworth, Davison, Rubin and Salam

Also a 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)

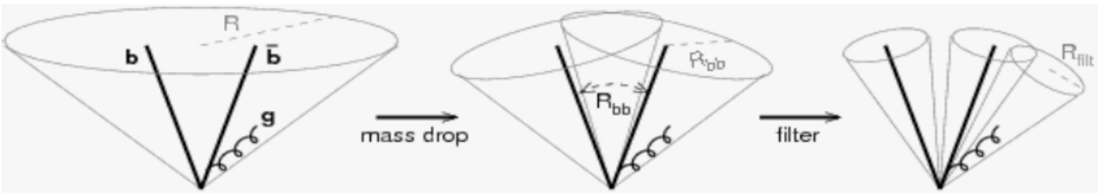


Non-perturbative effects which degrade signal
and shift can background into signal region

MANCHESTER 1824

The University of Manchester

BDRS mass drop tagger (MDT)



Butterworth Davison Rubin and Salam 2008

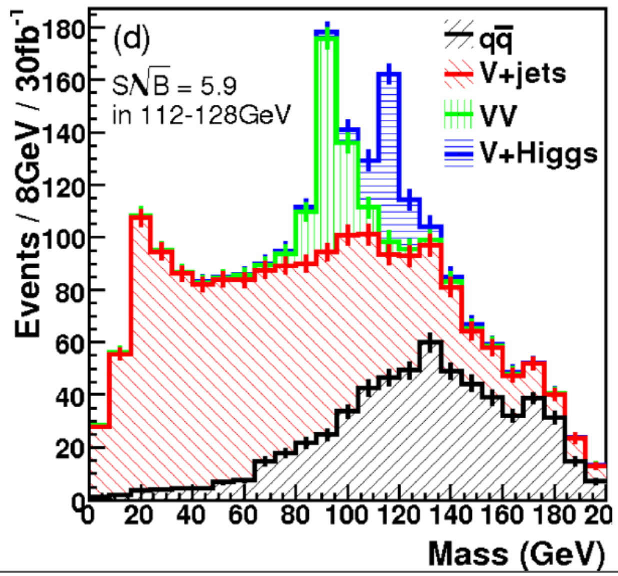
- Break the jet into two subjets j_1 and j_2 such that $mj_1 > mj_2$.
- Require mass drop $m_{j1} < \mu m_j$ and $\frac{\min(p_{t1}, p_{t2})}{\max(p_{t1}, p_{t2})} > y_{\mathrm{cut}}$

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.

MDT tags and grooms. Last step is pure grooming element needed at moderate p_T

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.

Led to a rapid proliferation of tools !

MANCHESTER 1824

The University of Manchester

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 intrinsic scales close to a light Higgs mass. For example, $t\bar{t}$ events can produce a leptonically decaying W, and in each top-quark rest frame, the b-quark has an energy of ~ 65 GeV, a value uncomfortably close to the $m_H/2$ that comes from a decaying light Higgs boson. If the second W-boson decays along the beam direction, then such a $t\bar{t}$ event can be hard to distinguish from a WH signal event.

In this letter we investigate VH production in a boosted regime, in which both bosons have large transverse momenta and are back-to-back. This region corresponds 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 gluons they emit, while discarding as much contamination as possible from the underlying event (UE), in order to maximise resolution on the jet mass. One should also correlate the momentum structure with the directions of the two b-quarks, and provide a way of placing effective cuts on the z fractions, both of these aspects serving to eliminate backgrounds.

To flexibly resolve different angular scales we use the inclusive, longitudinally invariant Cambridge/Aachen (C/A) algorithm [7, 8]: one calculates the angular distance $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ between all pairs of

Butterworth, Davison, Rubin and Salam 2008

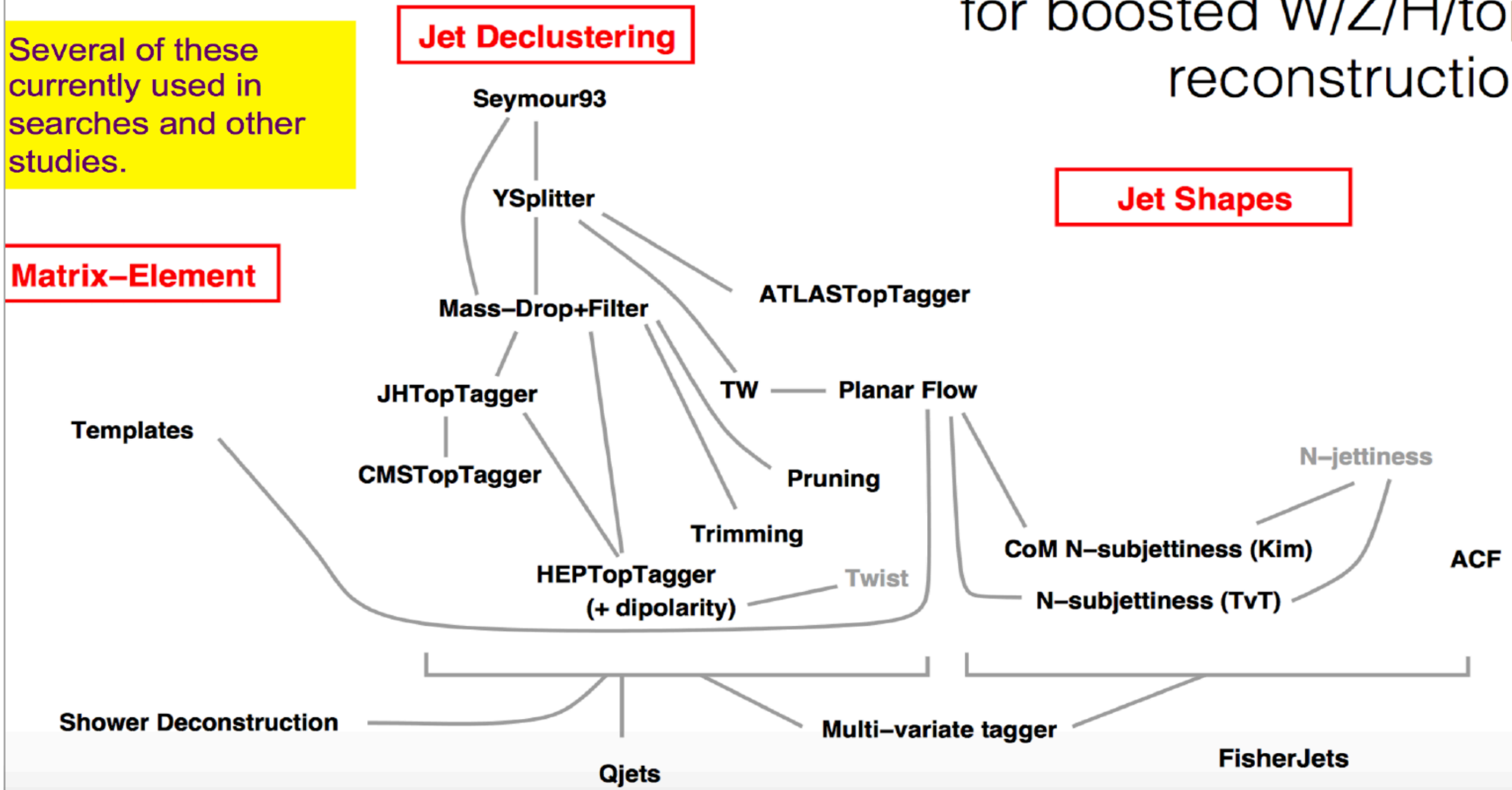
Since 2008 a vibrant research field emerged based on developing and exploiting jet substructure.

BDRS paper has > 1000 citations

Lots of tools

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

Several of these currently used in searches and other studies.



Open questions

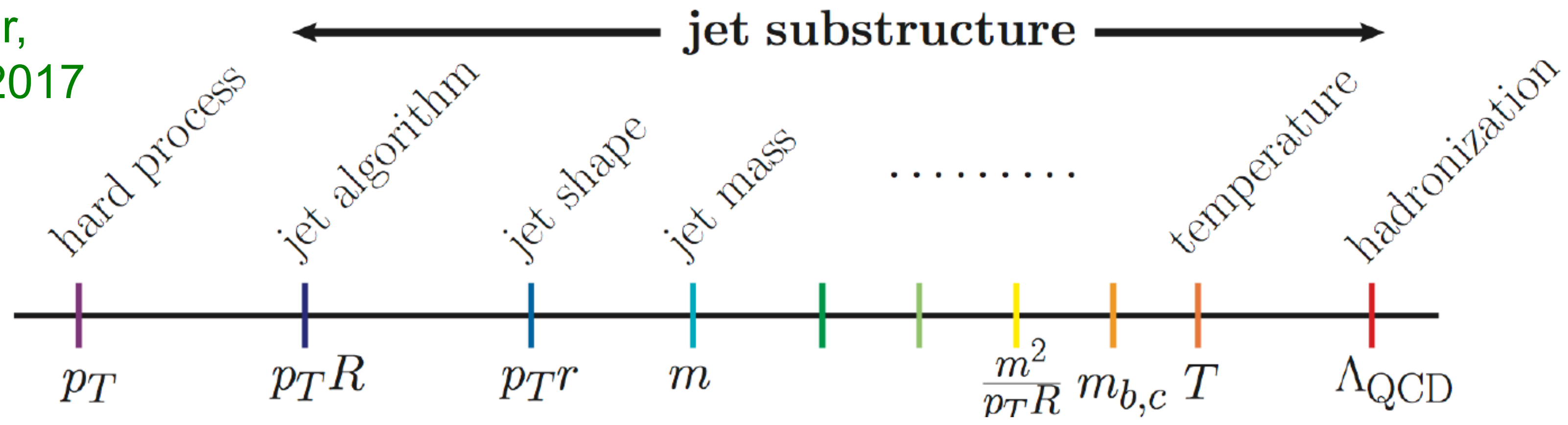
- Why so many tools to exploit a few physics principles?
- Do we understand the physics behind tool performance?
- How robust are various tools? E.g. does performance change with kinematics and parameters?

How to decide which tools to use in searches and data/theory comparisons? Look to guidance from theory?

Theory Issues

A multiscale problem

J.Thaler,
Boost 2017



MANCHESTER
1824

The University of Manchester

Multiple scales and large logs

Higher order pQCD calculations are standard way to obtain better precision but are not used here.

$$\frac{1}{\sigma}\frac{d\sigma}{dm_j^2} \sim \frac{1}{m_j^2}\frac{C_i\alpha_s}{\pi}\ln\left(\frac{R^2p_t^2}{m_j^2}\right)$$

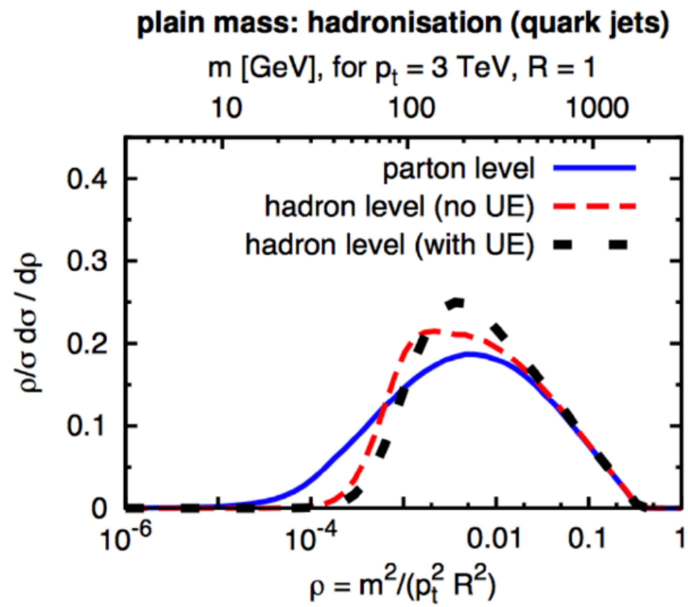
$p_t >> m_j$

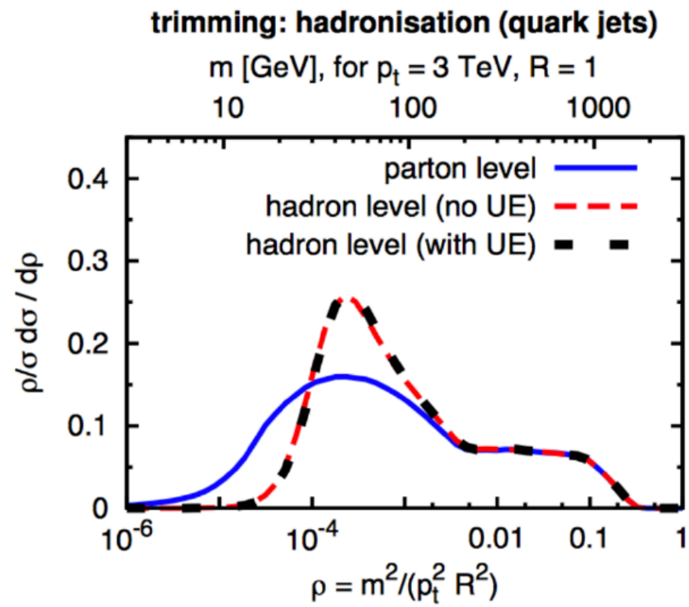
Affect convergence of perturbation theory

- Perturbation theory at fixed-order fails
- Can analytically ‘resum’ the logarithms for selected observables.
- Parton showers in GPMC codes also resum logs to a limited and ill-understood accuracy

Progress here!

Non-perturbative effects





Are these important in the TeV region? Consider that a 1 GeV gluon inside an R=1 3 TeV jet can produce a jet mass of 55 GeV.

$$m_j^2 \sim \Lambda p_T R^2$$

NP bumps visible but where NP = Non-Perturbative!

Models for hadronisation and UE but no first principles theory.

MANCHESTER 1824

Main approaches

- Develop substructure taggers using rough intuition and study performance with **Monte Carlo methods.**
- Look for some guidance from perturbative QCD theory - **analytic resummation**.
- Exploit recent advances **in machine learning** to develop more performant tools.

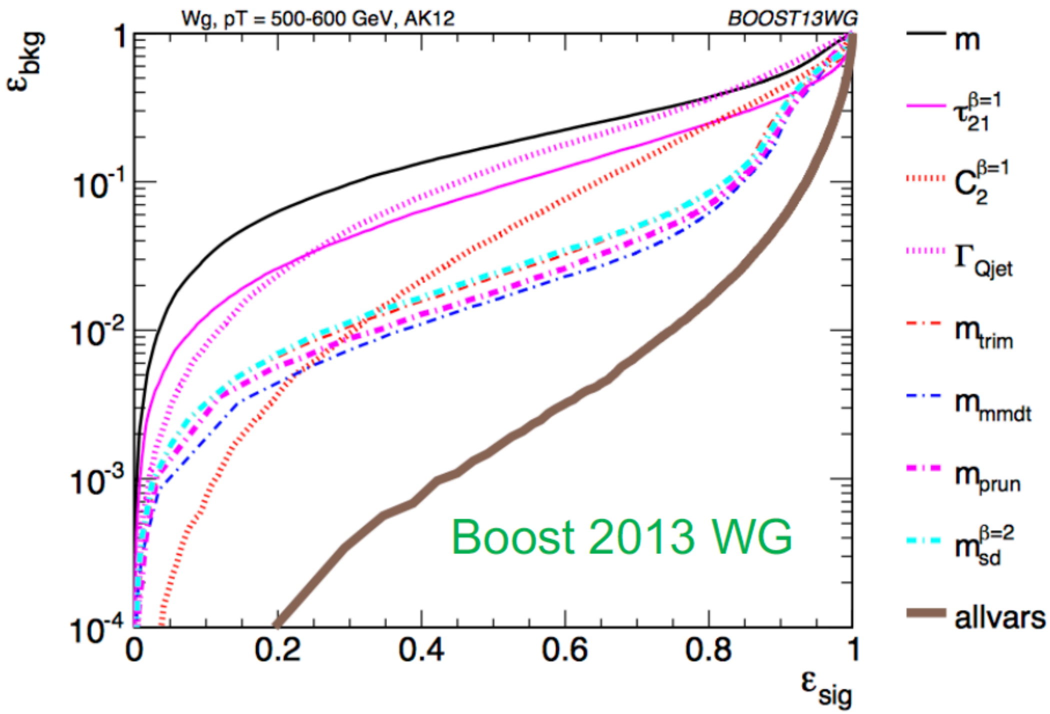
All offer advantages but its crucial to recognise the limitations in each case.

MANCHESTER 1824

The University of Manchester

Performance in MC studies

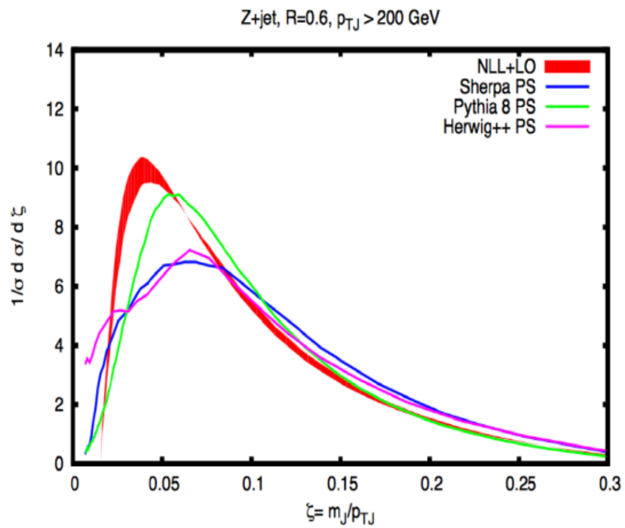
Imagine these same results with large uncertainty bands



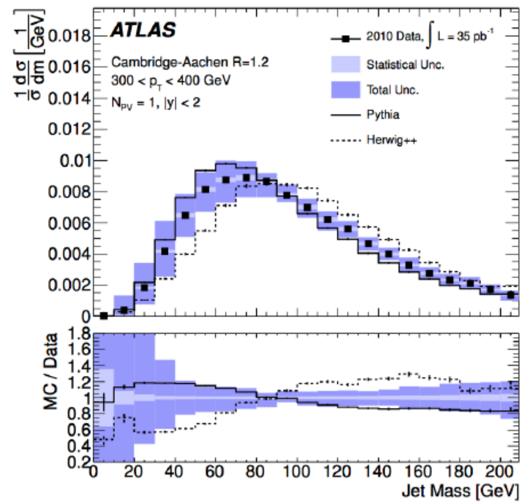
Estimating uncertainties depends on assumptions about shower accuracy. At LL level 50-100% is possible.

Simplest most common approach is to use MC results.
But gives no idea about uncertainties or insight into origin of gains. Complete reliance on MC models.

Issues with showers



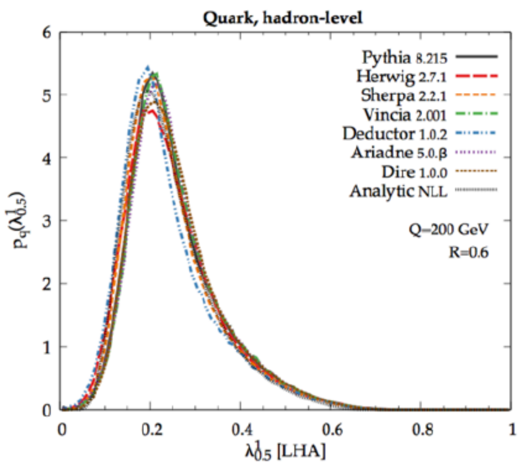
Dasgupta, Khelifa-Kerfa, Marzani, Spannowsky 201

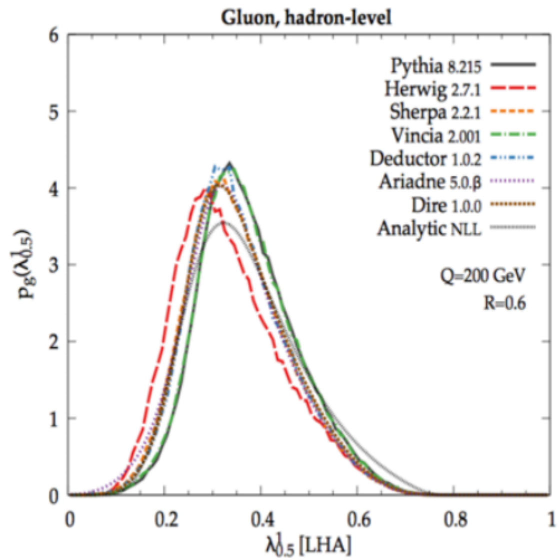


ATLAS collaboration 2012

- Shower predictions often show **a substantial spread**.
- At high p_t > 1 TeV the above differences would be large for 100 GeV or more in jet mass.

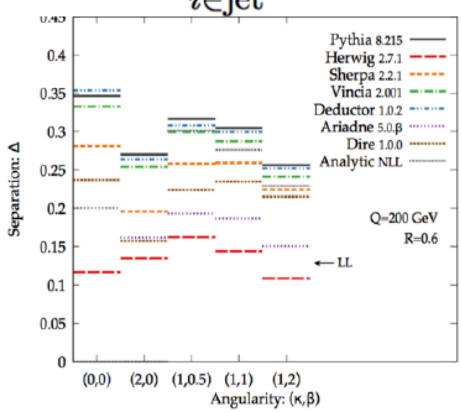
Quark/gluon discrimination





Gras et. al 2017

$$\lambda^{\kappa}_{\beta} = \sum_{i \in \text{jet}} z_i^{\kappa} \theta_i^{\beta},$$



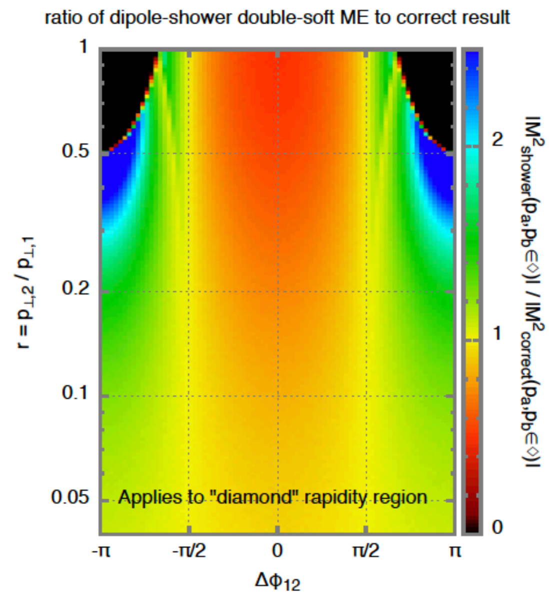
Focus on LHA with $\kappa = 1, \beta = 0.5$

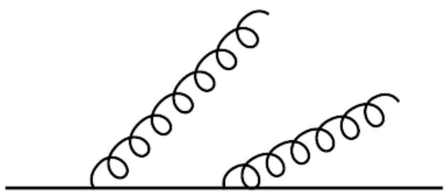
Wide spread between generators for gluon jets. Call for LHC measurements of angularities to tune event generators and constrain gluon jets.

$$\Delta = \frac{1}{2} \int \mathrm{d}\lambda \, \frac{\left(p_q(\lambda) - p_g(\lambda)\right)^2}{p_q(\lambda) + p_g(\lambda)} = 1 - 2 \int \mathrm{d}\lambda \, \frac{p_q(\lambda)\, p_g(\lambda)}{p_q(\lambda) + p_g(\lambda)}$$

Shower versus QCD matrix elements

MD, Dreyer, Hamilton, Monni, Salam 2018





Pythia and Dire Shower two emission matrix element fails to reproduce known QCD results in logarithmically enhanced regions.

A concern for methods that exploit pattern of correlation between emissions e.g. machine learning based approaches?

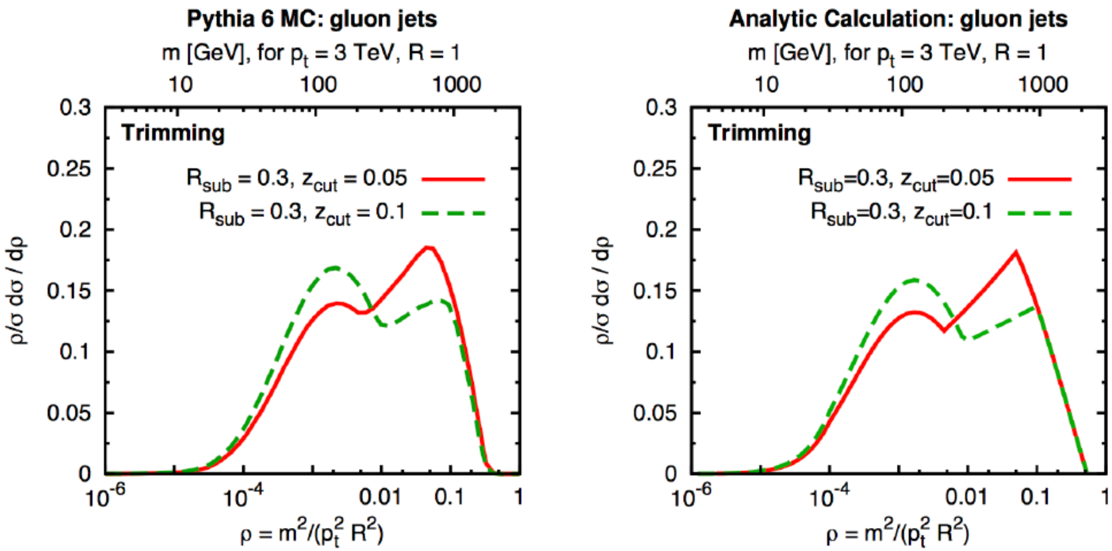
Recent Progress

MANCHESTER 1824

The University of Manchester

Jet substructure from analytics

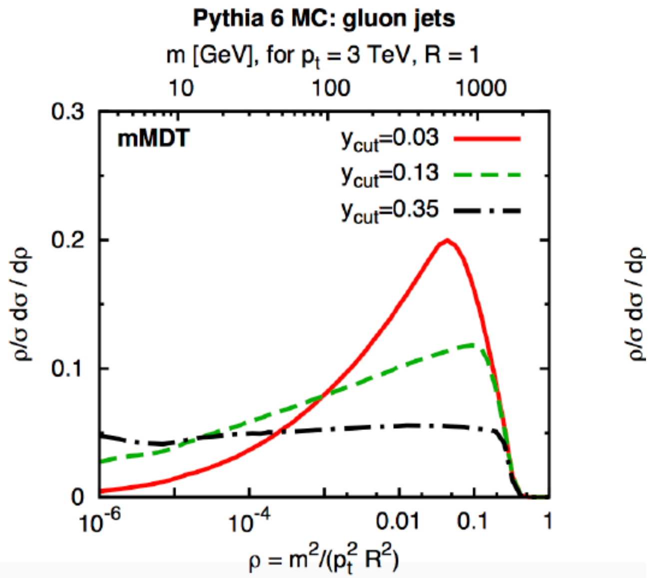
Dasgupta, Fregoso, Marzani and Salam 2013

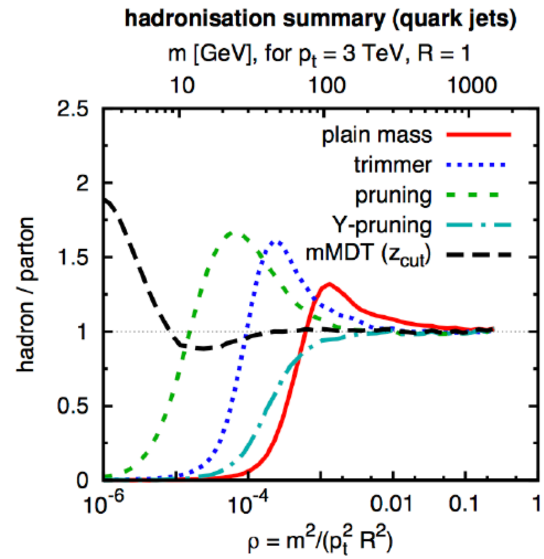


- Since 2013 analytical calculations for substructure observables developed.
- Based on perturbative QCD resummed calculations.
- Give considerable insight into taggers and their features.

Analytical calculations exposed crucial flaws in many existing methods

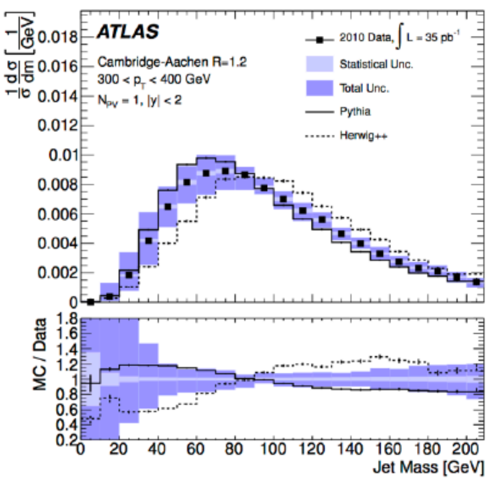
More robust tools

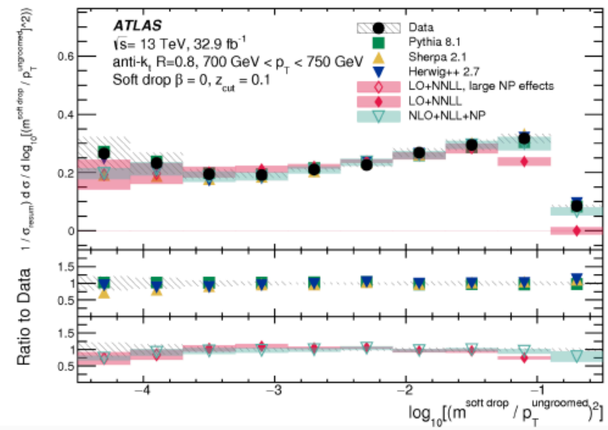




- Modifed mass drop tagger and its descendent SoftDrop
- Unique features implying very high accuracy perturbative calculations possible
- Widely used as jet grooming tools at the LHC.

Precise Calculations and phenomenology

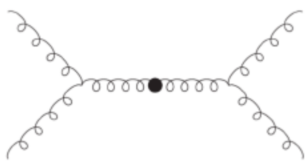


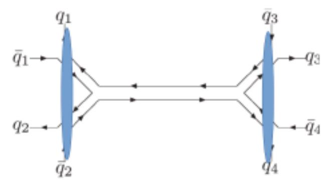


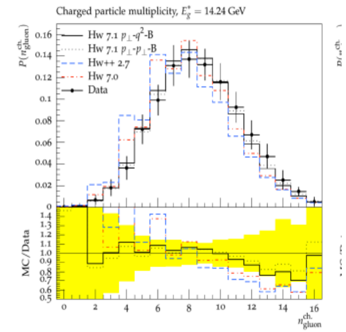
Frye, Larkoski, Schwartz, Yan, 2016

- Direct comparisons between data and first principles QCD theory
- mMDT/SoftDrop are widely used so confidence in tools is key.
- Significant development since 2012 ATLAS comparisons

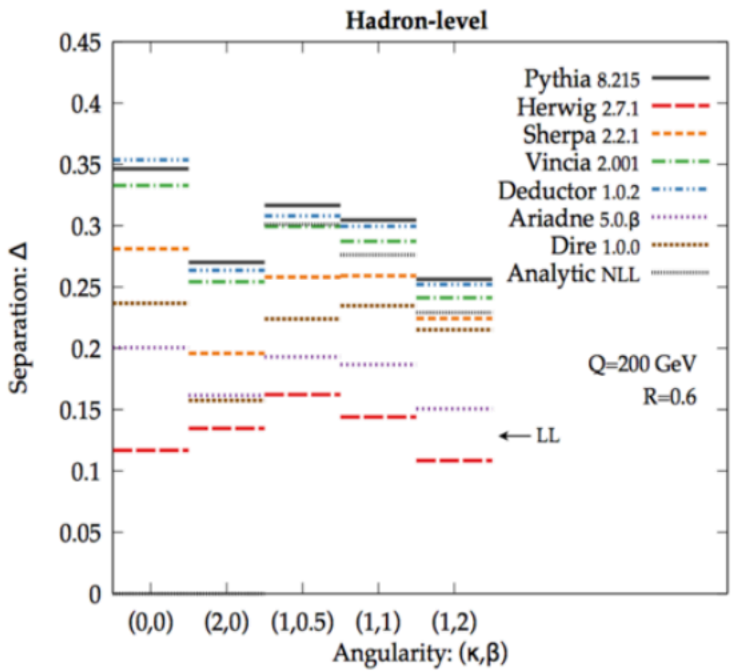
Improvements to showers

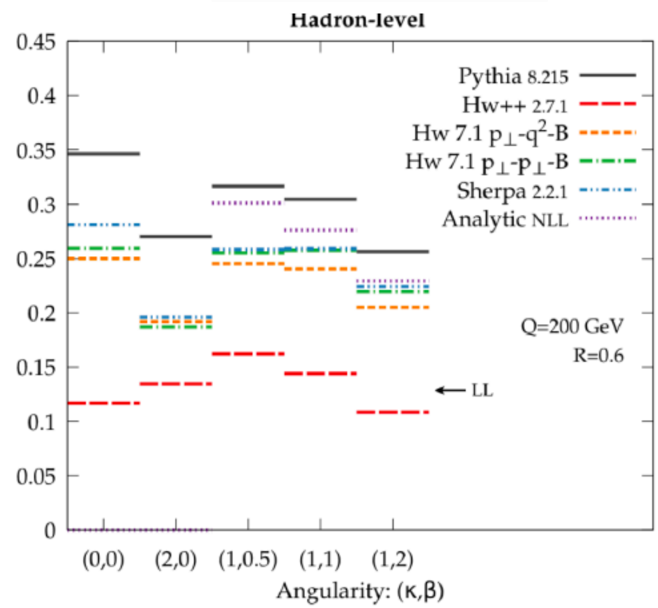






OPAL data for gluon jet charged multiplicity. Not used for tuning before!





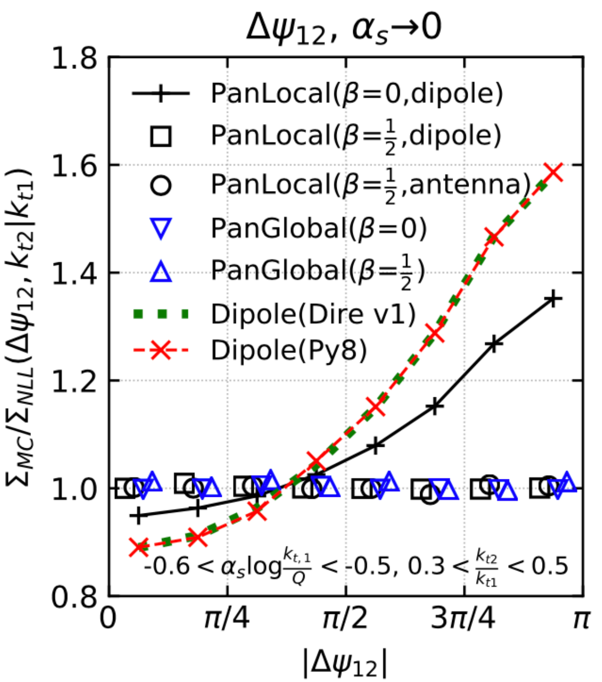
Richardson, Reichelt, Siodmok 2017

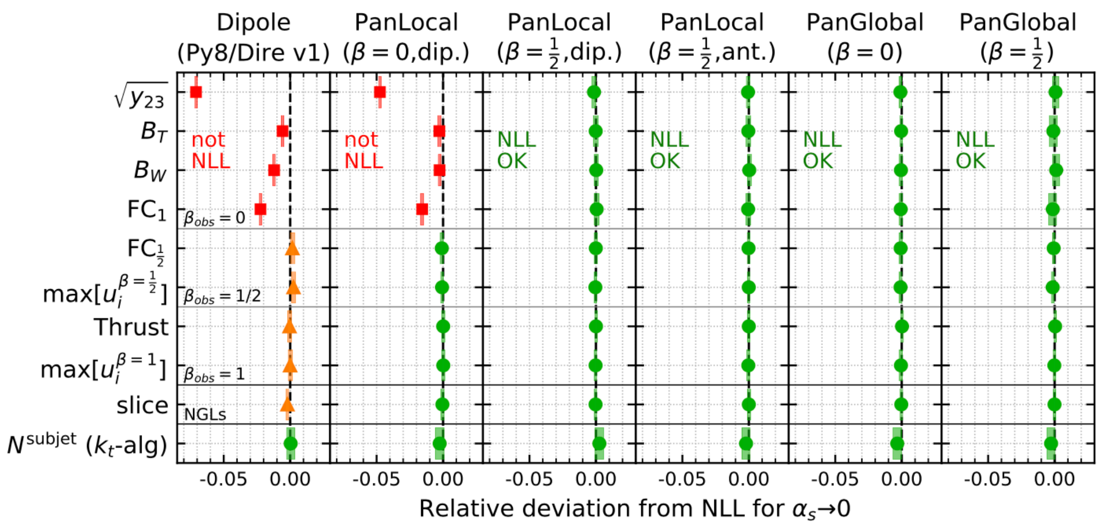
- Herwig is now more optimistic when it comes to distinguishing q/g jets.
- Spread of predictions is reduced.

MANCHESTER 1824

The University of Manchester

New showers with unprecedented accuracy

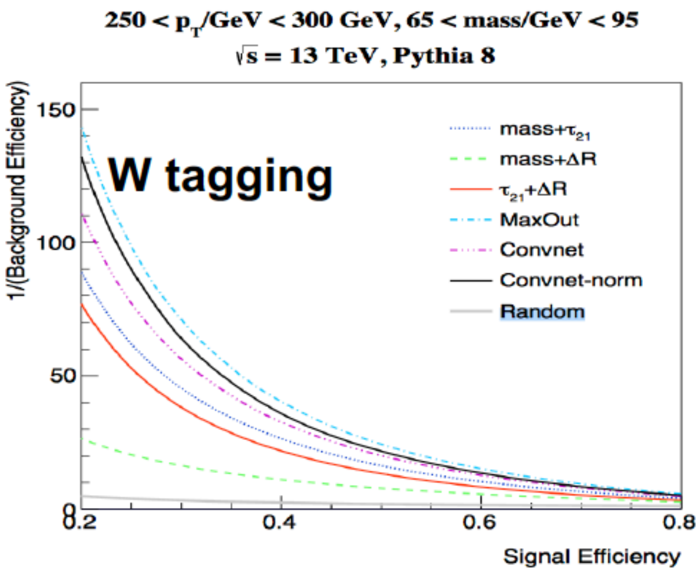


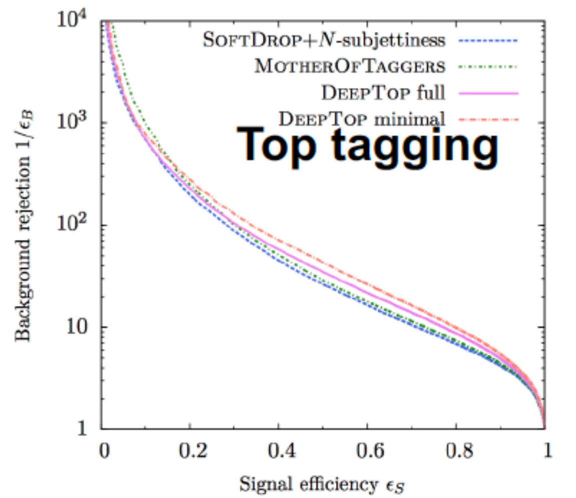


Dasgupta, Dreyer, Hamilton, Monni, Salam, Soyez 2020

- For the first time showers **constructed** and **proven** to have NLL accuracy.
- Implies that formal accuracy would change from as high as 50% to about the 10% percent level.

Powerful machine learning tools





De Oliveira, Kagan, Mackey, Nachman, Schwartzmann 2016

Kasieczka, Plehn, Russell. Schell 2016

- **Very active area. Perfect playground for ML approaches**. Wide range of methods used
- Often **substantially better than manually constructed observables** for performance
- Do we pay a price? **What features are learnt? Are they well modelled by showers, detector simulations?**

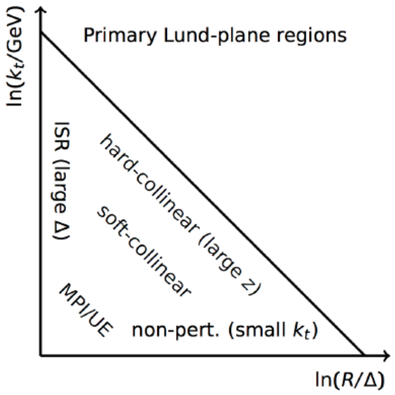
The University of Manchester

MANCHESTER 1824

Learning from the Lund plane

Lund diagrams in the $(\ln z\theta, \ln \theta)$ plane are a very useful way of representing emissions.

Dreyer, Salam, Soyez 2018



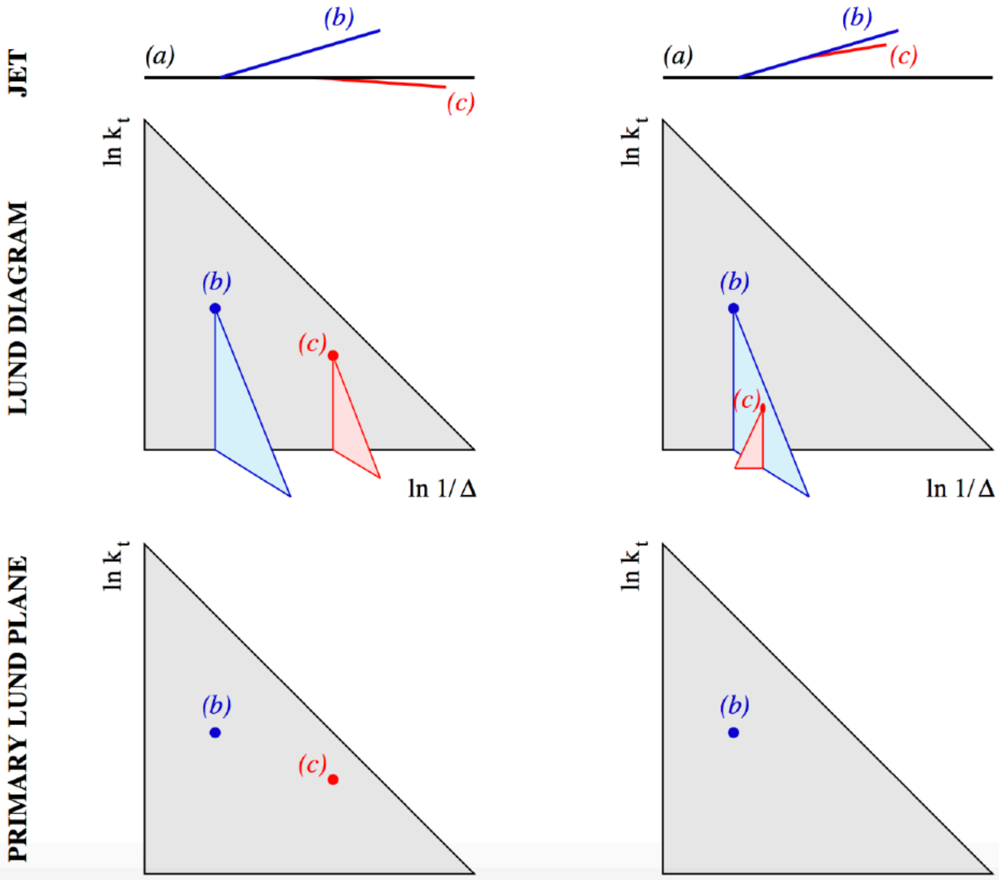
Separation of different physics effects inluding non-pert.

Density of emissions in primary Lund plane well understood theoretically.

At the heart of analytic approaches and parton showers. Can be used as an input to ML.

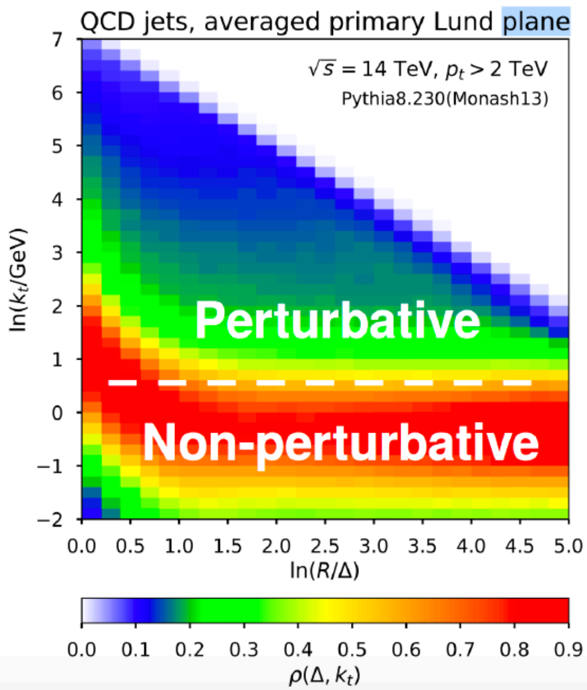
Bridges the gap between Deep learning and “Deep thinking” approaches?

Learning from the Lund plane



Dreyer, Salam, Soyez, 2018

QCD jets in the Lund plane



Dreyer, Salam, Soyez, 2018

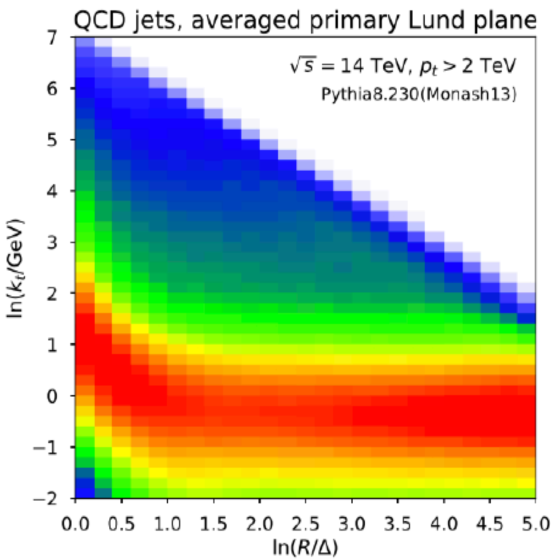
ln(k_t/GeV)

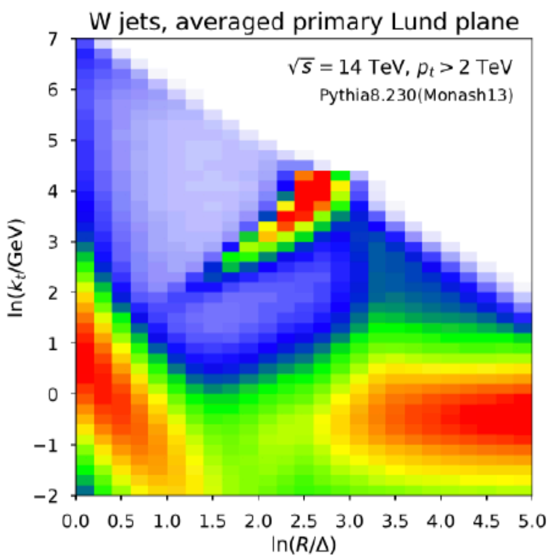
In the soft collinear region

$$\rho \sim 2C_F \frac{\alpha_s(k_t)}{\pi}$$

Well understood theoretically. Applications include constraining event generator models, input to machine learning, manually designing optimal observables and direct measurements

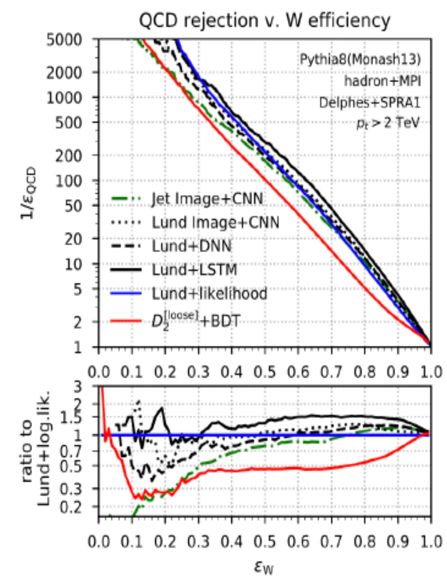
W tagging





Jet Tagging Applications

(Direct Likelihood or Deep Network)



Used both log-likelihood and machine learning approach.

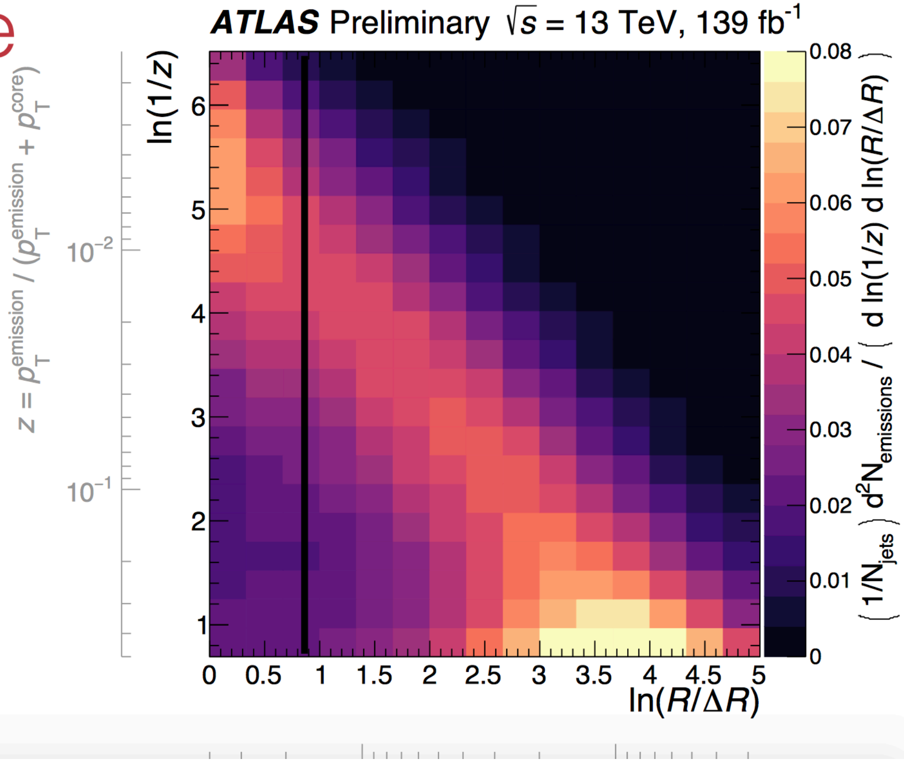
Dreyer, Salam, Soyez, 2018

Lund plane measurement

The Lund Jet Plane

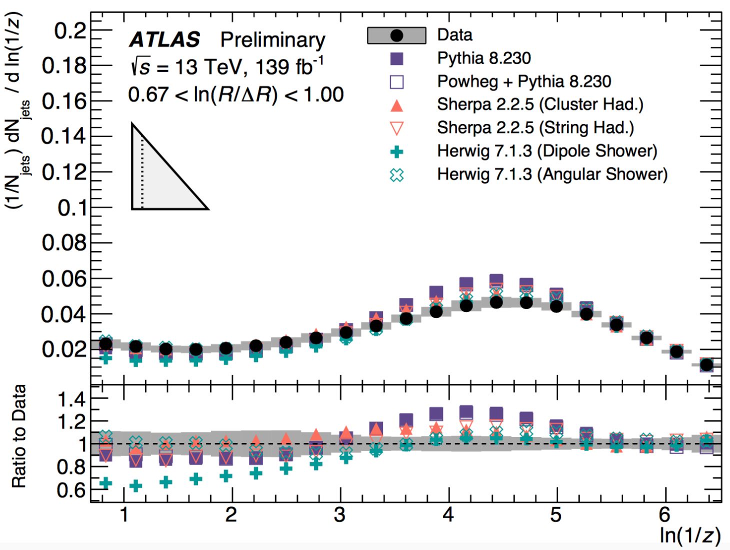
- Unfolded the primary Lund plane in dijet events
  - Using R = 0.4 jets
- Using tracks associated to the jets in order to have precise measurements for small splittings
  - Unfolding to charged particle level

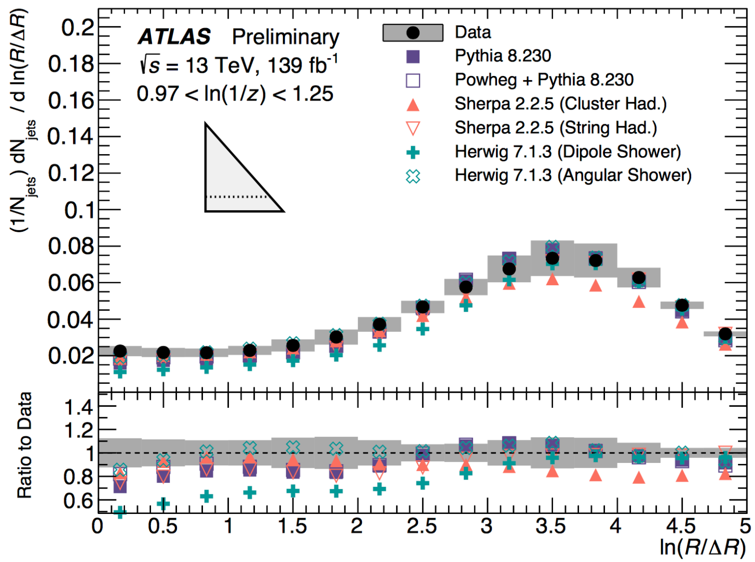
ATLAS-CONF-2019-035



Roloff Boost 2019

Lund plane measurement

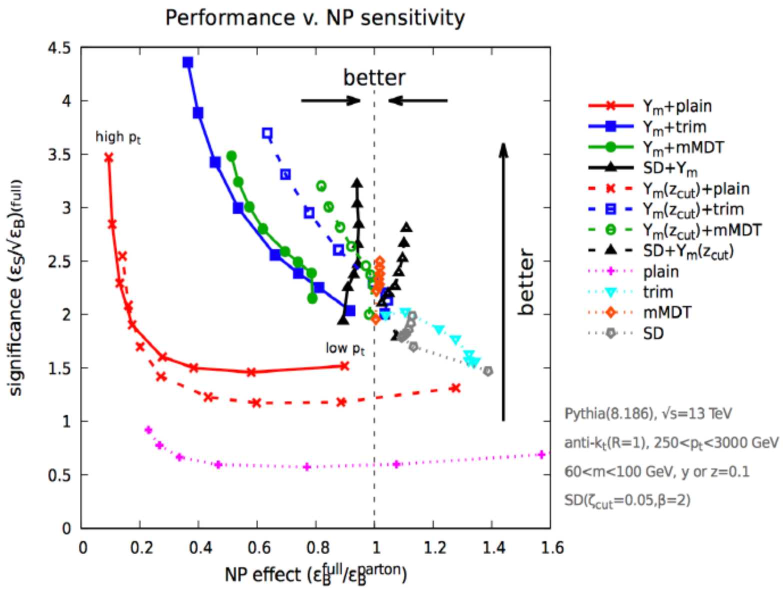




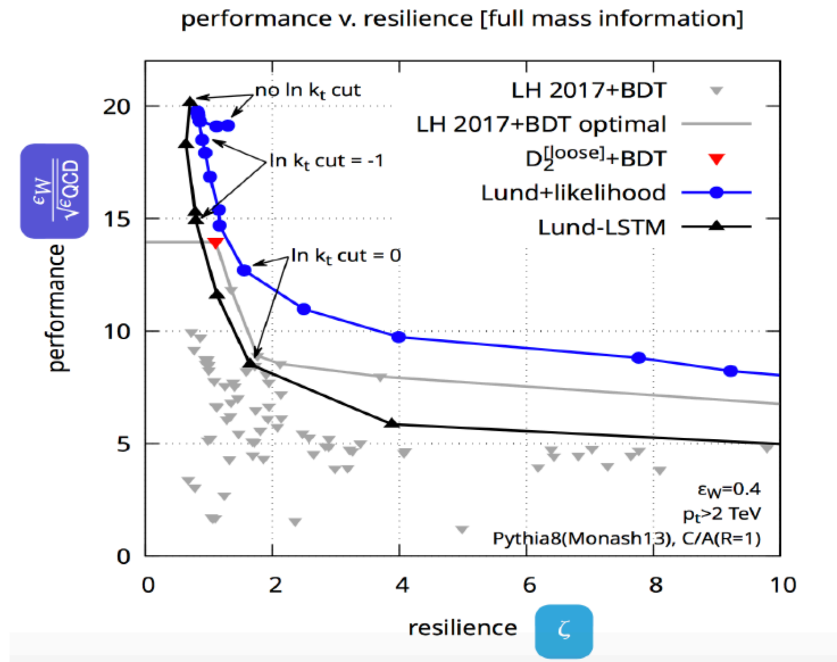
Non-trivial differences between generators and data.

Roloff Boost 2019

Final thoughts : performance v resilience



Dasgupta, Powling, Schunk, Soyez 2016



Dreyer, Salam, Soyez, 2018

To what extent do we want to rely on our knowledge of QCD at 1 GeV for TeV scale physics?