A brief look inside jet substructure

Grégory Soyez

IPhT, CEA Saclay, CNRS

Getting to grips with QCD, April 4-6 2018

Jets are routine QCD objects

- ubiquitous in collider physics
- around since 40 years
- used in at least 60% of LHC analyses

Jets are routine QCD objects

- ubiquitous in collider physics
- around since 40 years
- used in at least 60% of LHC analyses

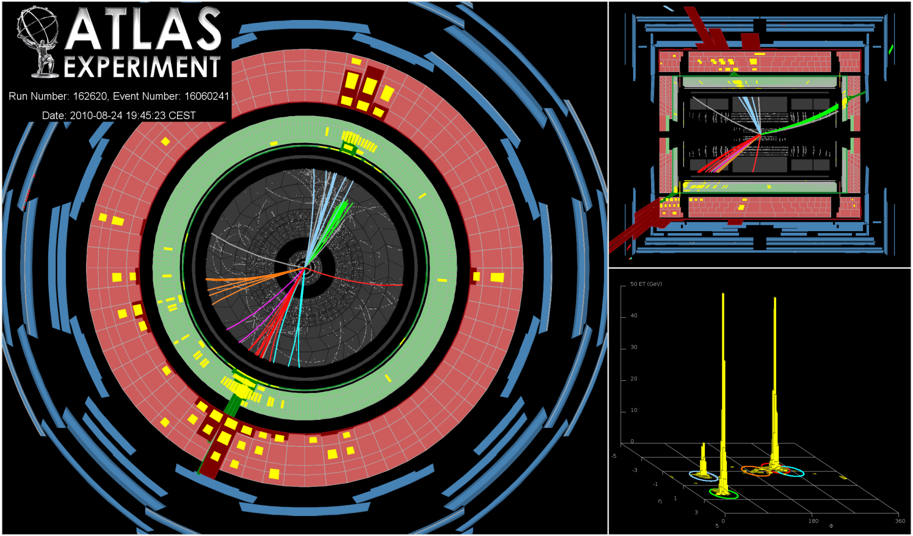
You could live a happy life by just knowing a few things

Concepts:

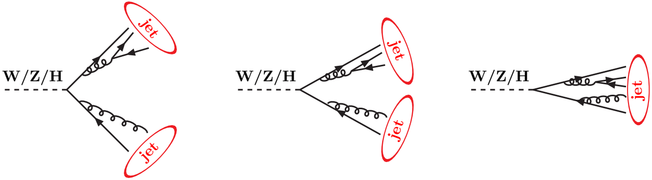
- Jets are proxies to hard partons produced in collisions
- infrared-and-collinear safe
- capture collimated parton cascades from hard scale Q to $\mathcal{O}(1\text{ GeV})$

Practically:

- obtained by running a clustering algorithm
- the LHC uses the anti-k_t algorithm
- FastJet covers covers all your numerical needs for clustering


ATLAS
EXPERIMENT
Run Number: 162620, Event Number: 16060241
Date: 2010-08-24 19:45:23 CEST
50 ET (GeV)
40
30
20
10
0
-5
-3
-1
η
1
3
5
0
180
360
Φ

$p_t \lesssim m$ $\qquad\qquad$ $p_t \gg m$



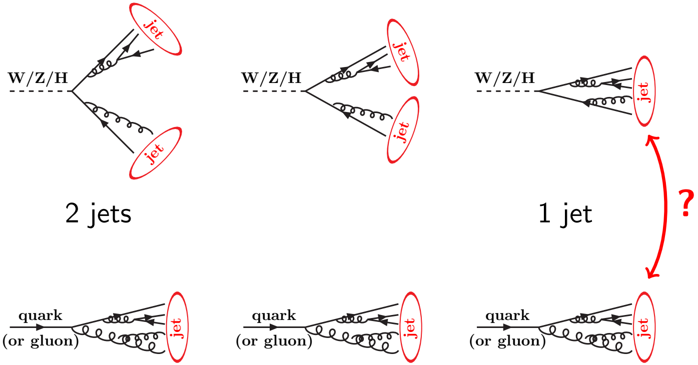
2 jets $\qquad\qquad$ 1 jet

(massive) objects produced boosted (energy $\gg$ mass) are seen as 1 jet:

$$\theta_{q\bar{q}} \sim \frac{m}{p_t}$$

Boosted objects

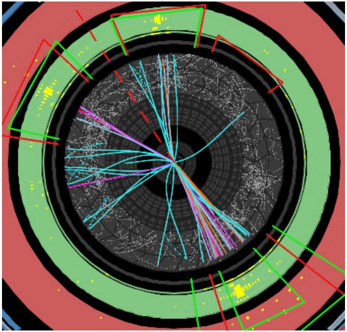
$p_t \lesssim m$ $\qquad\qquad$ $p_t \gg m$



use substructure to separate from QCD jets

Other examples

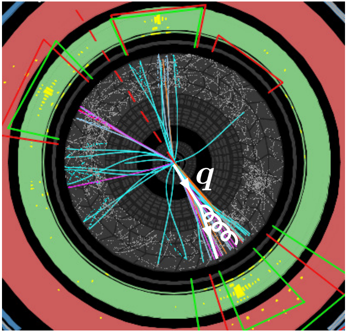
What jet do we have here?



Other examples

What jet do we have here?

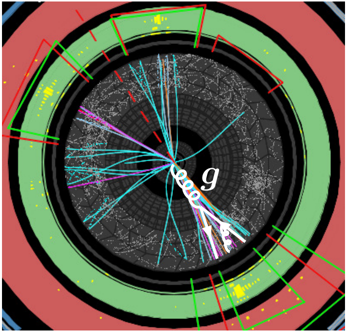
- a quark?



Other examples

What jet do we have here?

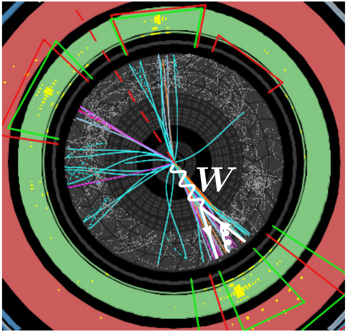
- a quark?
- a gluon?



Other examples

What jet do we have here?

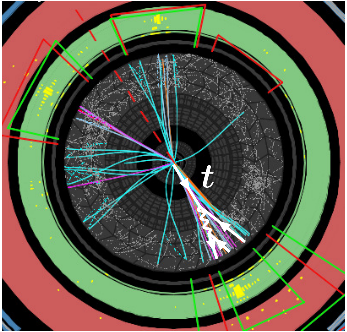
- a quark?
- a gluon?
- a W/Z (or a Higgs)?



Other examples

What jet do we have here?

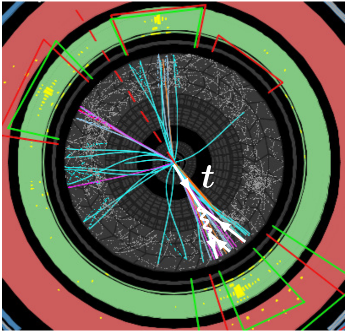
- a quark?
- a gluon?
- a W/Z (or a Higgs)?
- a top quark?



What jet do we have here?

- a quark?
- a gluon?
- a W/Z (or a Higgs)?
- a top quark?

Source: ATLAS boosted top candidate

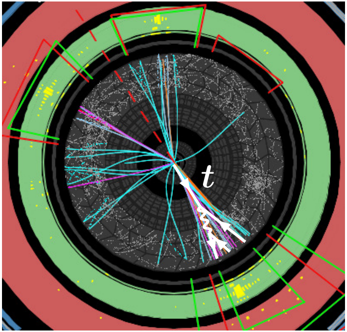


Other examples

What jet do we have here?

- a quark?
- a gluon?
- a W/Z (or a Higgs)?
- a top quark?

Source: ATLAS boosted top candidate



Substructure means looking at the internal dynamics of jets
(as opposed to consider jets as monolithic objects)

Main concepts:

- Compared to standard jets, substructure uses a **large toolkit**

Main concepts:

- Compared to standard jets, substructure uses a **large toolkit**
- Basic tools are organised around $2_{(3)}$ major concepts:
  - "peak/prong finders":
    - $W/Z/H/t$ decay into hard partons $\Rightarrow$ jets with multiple hard cores
    - QCD (q/g) jets dominated by soft radiation $\Rightarrow$ single cores

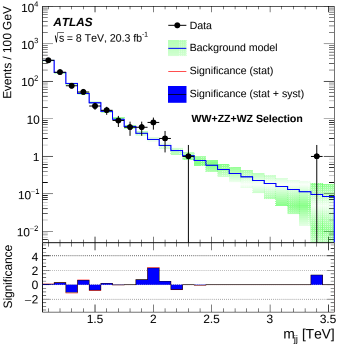
    Tools search for multiple hard cores
  - Radiation patterns:
    colourless $W/Z/H$ has less radiation than q/g jets
    Tools (jet shapes) to quantify radiation
  - (also grooming to remove soft contamination from fat jets)

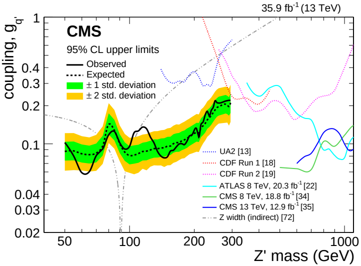
Main concepts:

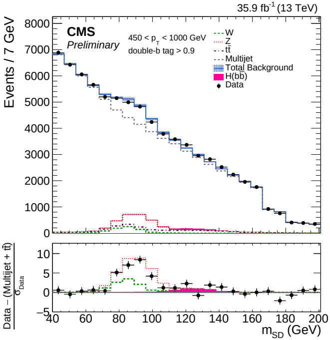
- Compared to standard jets, substructure uses a **large toolkit**
- Basic tools are organised around $2_{(3)}$ major concepts:
  - "peak/prong finders":
    - $W/Z/H/t$ decay into hard partons $\Rightarrow$ jets with multiple hard cores
    - QCD (q/g) jets dominated by soft radiation $\Rightarrow$ single cores

    Tools search for multiple hard cores
  - Radiation patterns:
    colourless $W/Z/H$ has less radiation than q/g jets
    Tools (jet shapes) to quantify radiation
  - (also grooming to remove soft contamination from fat jets)
- Increasing range of applications
  - Searches and measurements at the LHC
  - Precision QCD calculations
  - Measurements in heavy-ion collisions
  - Machine Learning

Searches and measurements







↑ (now-gone) di-boson excess (end of Run-I)

↗ Search for $X \to q\bar{q}$
Region inaccessible otherwise

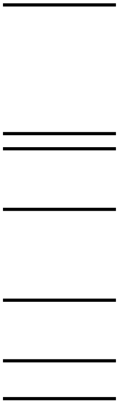
→ Clear Z peak, hint of a H peak

Plan for the talk

Give an overview of these aspects

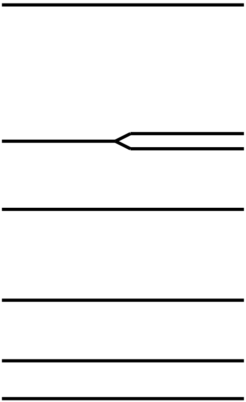
Articulated around a work in progress
with Frederic Dreyer and Gavin Salam

Cambridge/Aachen: iteratively recombine the closest pair

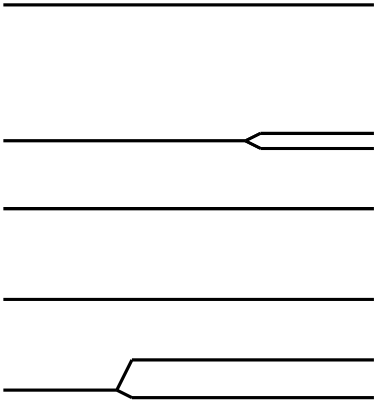


Pre-requisite number 1: Cambridge/Aachen (de-)clustering

Cambridge/Aachen: iteratively recombine the closest pair

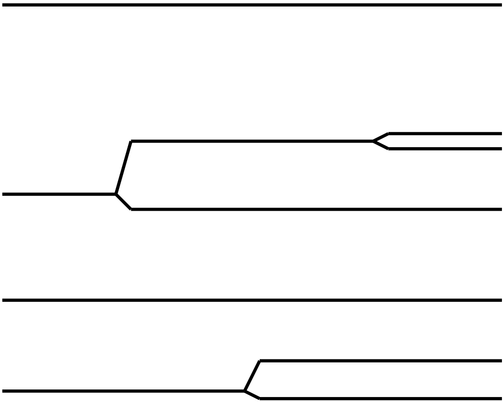


Cambridge/Aachen: iteratively recombine the closest pair

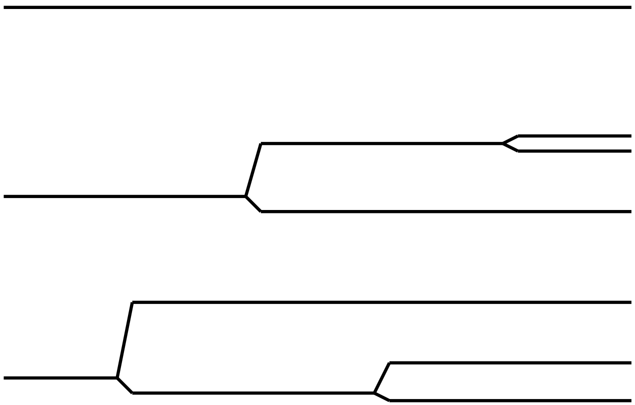


Pre-requisite number 1: Cambridge/Aachen (de-)clustering

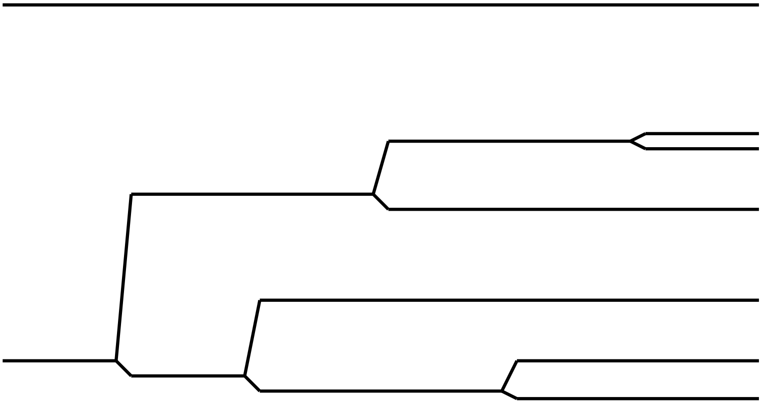
Cambridge/Aachen: iteratively recombine the closest pair



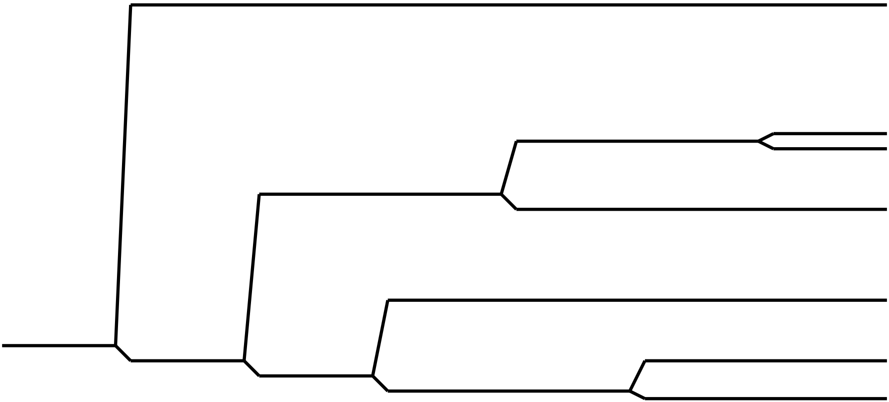
Cambridge/Aachen: iteratively recombine the closest pair



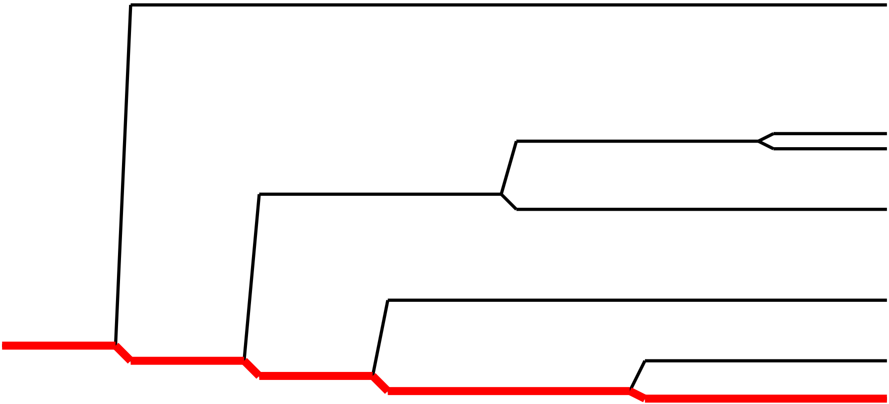
Cambridge/Aachen: iteratively recombine the closest pair



Cambridge/Aachen: iteratively recombine the closest pair



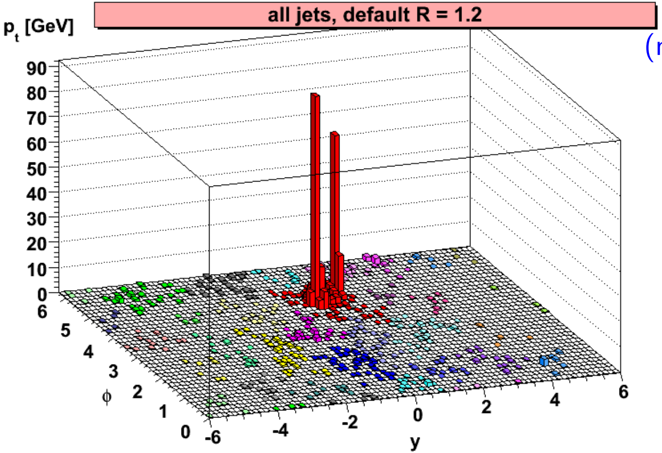
Cambridge/Aachen: iteratively recombine the closest pair



Usage: iteratively undo the clustering to study internal jet dynamics
Typically: follow the hardest branch (largest p_t or z)

Example: mMDT/SD prong finder

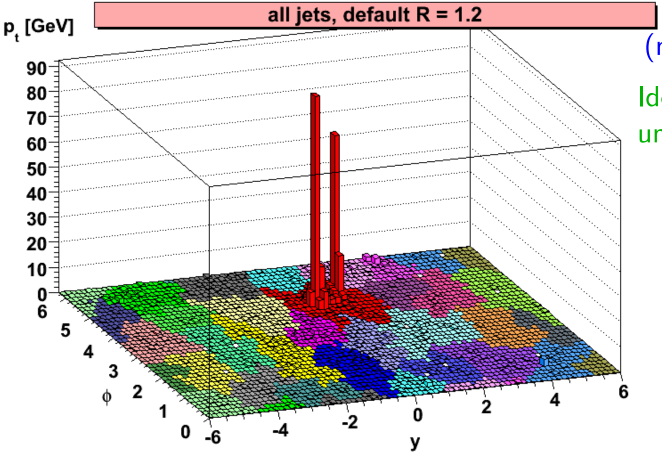
[J.Butterworth,A.Davison,M.Rubin,G.Salam,08]



(modified)MassDrop

Example: mMDT/SD prong finder

[J.Butterworth,A.Davison,M.Rubin,G.Salam,08]

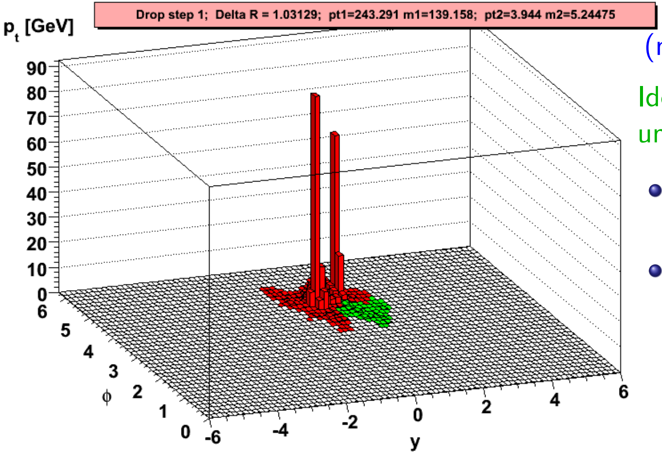


(modified)MassDrop

Idea: undo clustering until $z > z_{\text{cut}} \approx 0.1$

Example: mMDT/SD prong finder

[J.Butterworth,A.Davison,M.Rubin,G.Salam,08]



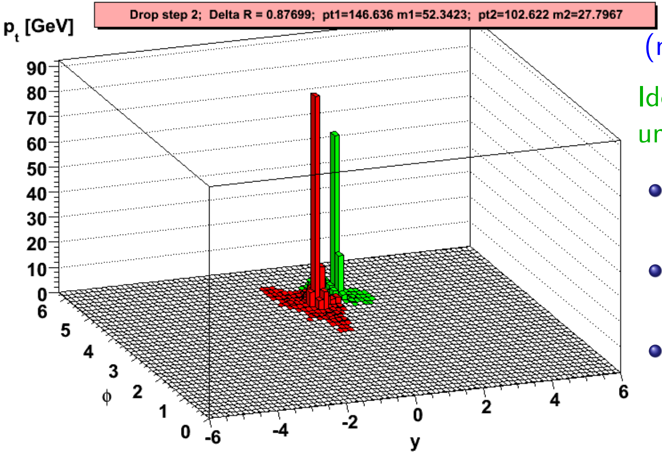
(modified)MassDrop

Idea: undo clustering until $z > z_{\rm cut} \approx 0.1$

- undo the last clustering step
- $z = 0.016 < 0.1$ carry on

Example: mMDT/SD prong finder

[J.Butterworth,A.Davison,M.Rubin,G.Salam,08]

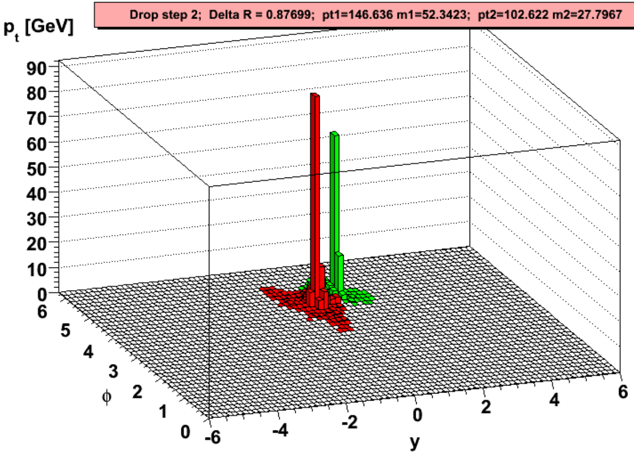


(modified)MassDrop

Idea: undo clustering until $z > z_{\text{cut}} \approx 0.1$

- undo the last clustering step
- $z = 0.016 < 0.1$ carry on
- $z = 0.41 > 0.1$ stop

[J.Butterworth,A.Davison,M.Rubin,G.Salam,08]

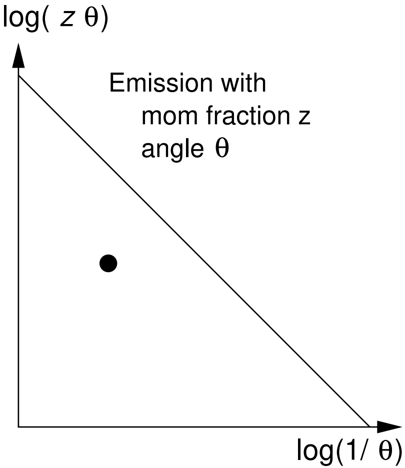


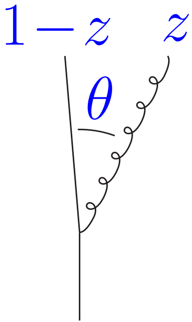
(modified)MassDrop

Idea: undo clustering until $z > z_{\text{cut}} \approx 0.1$

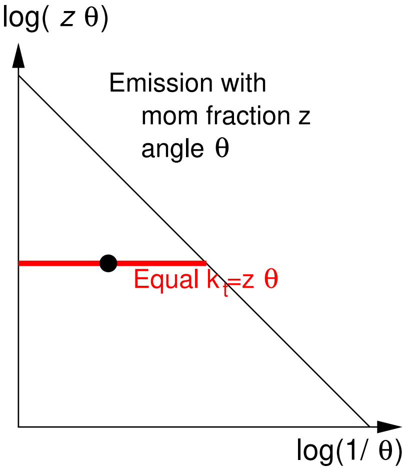
- undo the last clustering step
- $z = 0.016 < 0.1$ carry on
- $z = 0.41 > 0.1$ stop

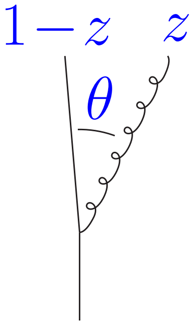
Variant: **SoftDrop**: impose $z > z_{\text{cut}}\theta^{\beta}$ [A.Larkoski,S.Marzani,GS,J.Thaler,14]


log(z θ)
Emission with
mom fraction z
angle θ
log(1/ θ)

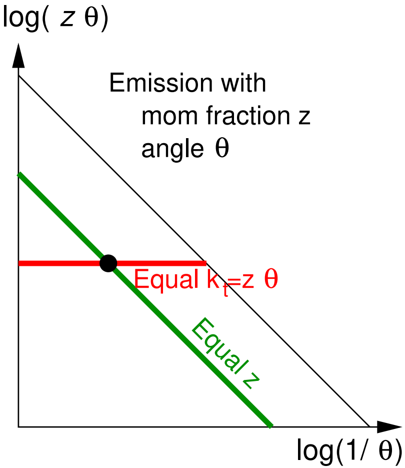

1−z
z
θ

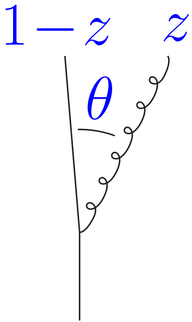
Pre-requisite number 2: Lund planes



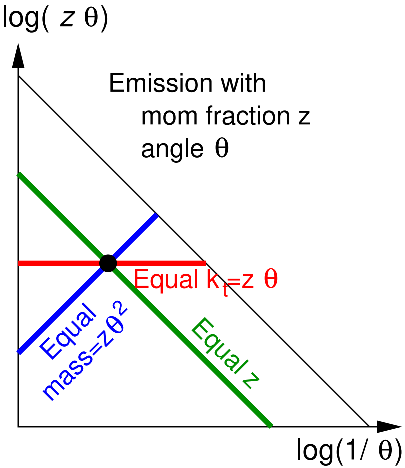


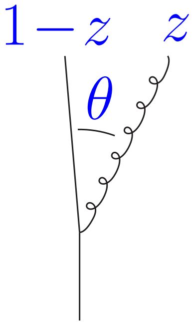
Pre-requisite number 2: Lund planes

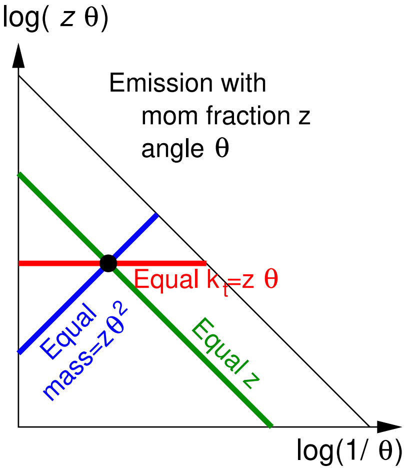




Pre-requisite number 2: Lund planes





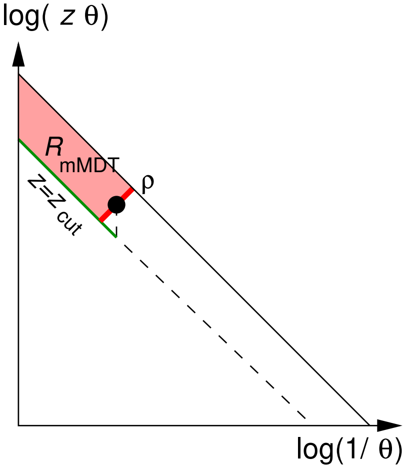


Observables in the boosted limit

Jet "mass": $(z_1\theta_1^2 \gg z_2\theta_2^2 \gg \ldots)$

$$\rho \equiv \frac{m^2}{p_t^2 R^2} = \sum_{i\in\text{jet}} z_i\theta_i^2 \approx z_1\theta_1^2$$

Pre-requisite number 2: Lund planes



(mMDT) jet mass spectrum at LL

$$\frac{\rho}{\sigma}\frac{d\sigma}{d\rho} = R'_{\text{mMDT}} \exp(-R_{\text{mMDT}})$$

1. emission of given mass

$$R'_{\text{mMDT}} \sim \frac{\alpha_s C_R}{\pi} \log(1/z_{\text{cut}})$$

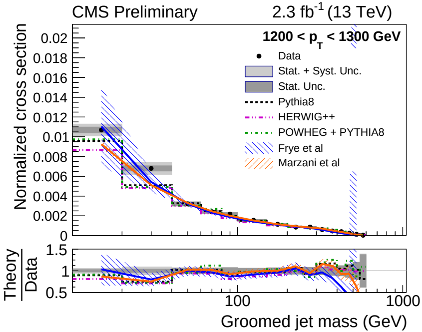
2. veto on larger-mass (Sudakov)

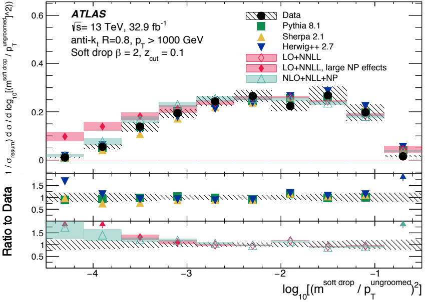
$$R_{\text{mMDT}} \sim \frac{\alpha_s C_R}{\pi} \log(1/\rho) \log(1/z_{\text{cut}})$$

[M.Dasgupta,A.Fregoso,S.Marzani,G.Salam]

Precision physics at the LHC

- LHC Measurements (CMS-PAS-SMP-16-010, ATLAS CERN-EP-2017-231)
- NNLL+LO in SCET (Frye,Larkoski,Schwartz,Yan; assumes small z_{cut})
- NLL+NLO in "standard QCD" (Marzani,Schunk,GS; includes (LL) finite z_{cut})

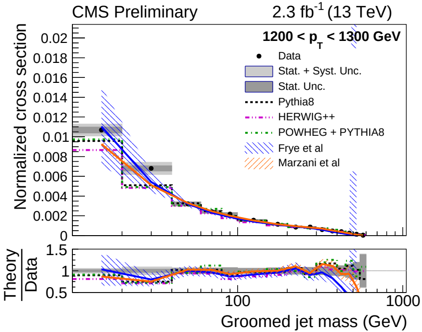


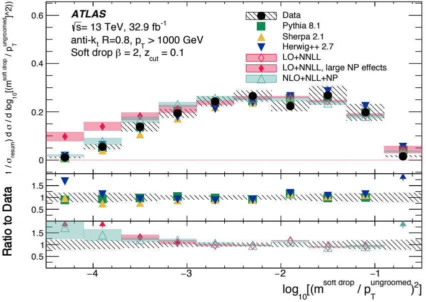


good overall agreement with the data

Precision physics at the LHC

- LHC Measurements (CMS-PAS-SMP-16-010, ATLAS CERN-EP-2017-231)
- NNLL+LO in SCET (Frye,Larkoski,Schwartz,Yan; assumes small z_{cut})
- NLL+NLO in "standard QCD" (Marzani,Schunk,GS; includes (LL) finite z_{cut})





good overall agreement with the data

Precise observable with limited sensitivity to NP effects

$\Rightarrow$ **possibility to extract** α_s (on-going study)

Opening many options

Based on

1. Cambridge/Aachen de-clusterings
2. Lund-plane variables

Opening many options

Based on

1. Cambridge/Aachen de-clusterings
2. Lund-plane variables

We can

- learn about jet dynamics
- constrain Monte-Carlo generators (separately for pQCD and non-pQCD)
- perform analytic studies
- learn about the quark-gluon plasma
- Tag boosted objects
- Use Machine Learning techniques

Opening many options

Based on

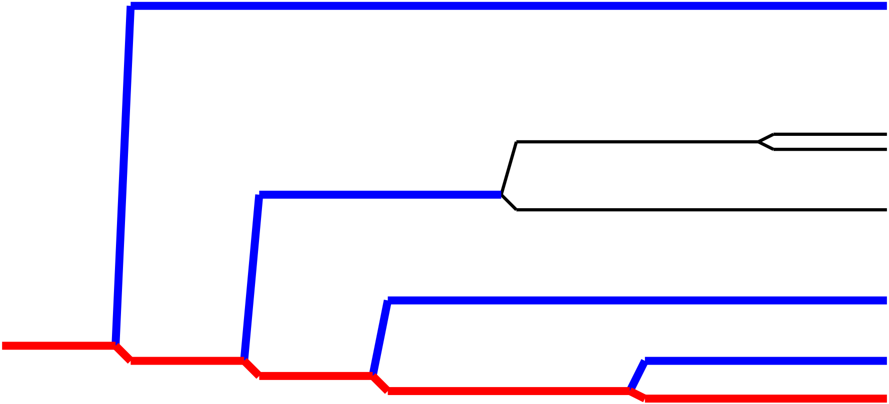
1. Cambridge/Aachen de-clusterings
2. Lund-plane variables

We can

- learn about jet dynamics
- constrain Monte-Carlo generators
  (separately for pQCD and non-pQCD)
- perform analytic studies
- learn about the quark-gluon plasma
- Tag boosted objects
- Use Machine Learning techniques

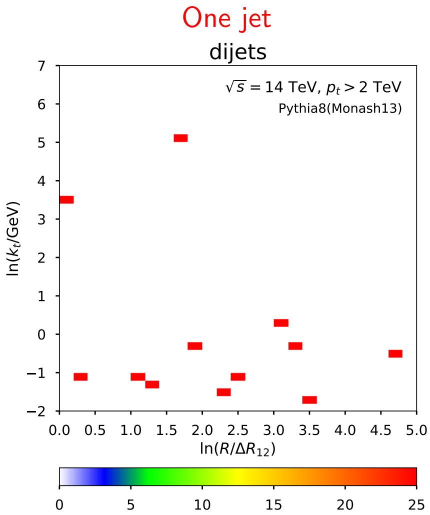
I will illustrate this in the next slides

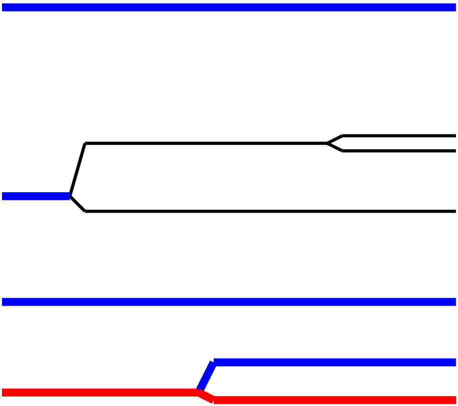
Lund plane density



Consider all the emissions from the hardest branch

Lund plane density

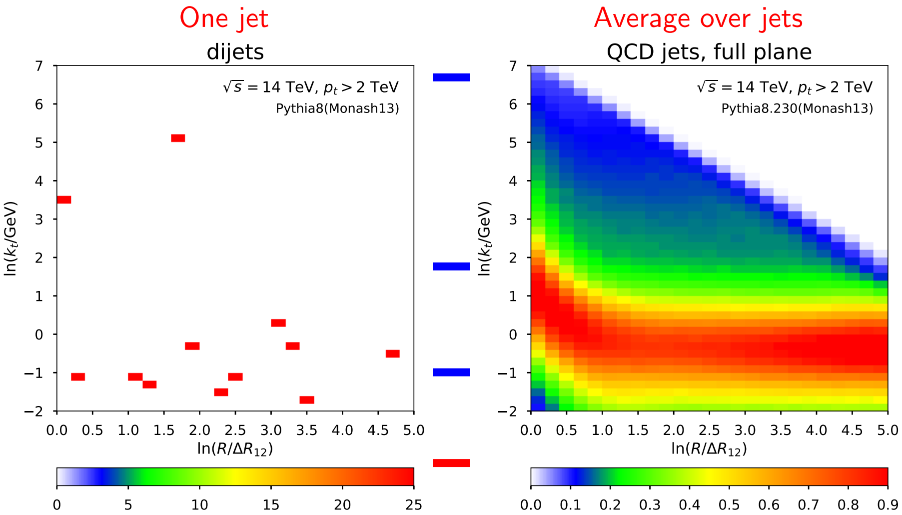




Consider all the emissions from the hardest branch

Put them in the Lund plane

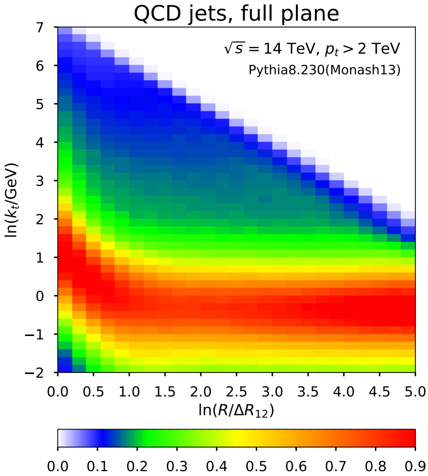
Lund plane density



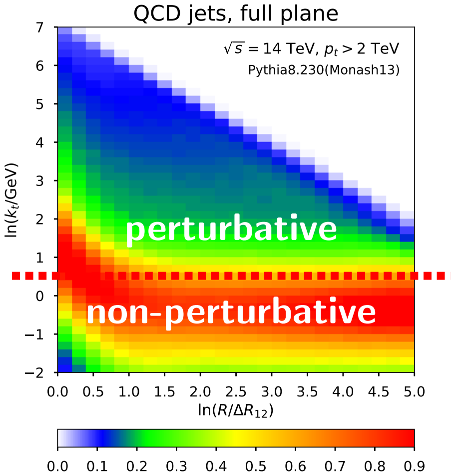
Consider all the emissions from the hardest branch

Put them in the Lund plane

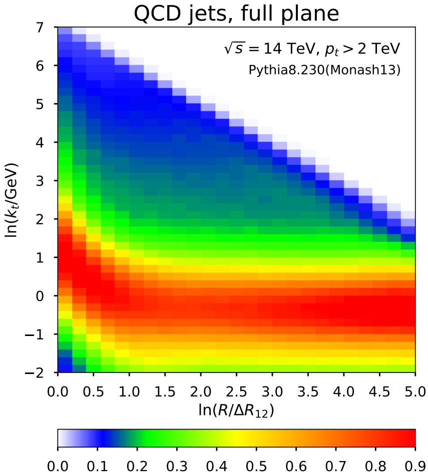
Lund plane and Monte-Carlo generators



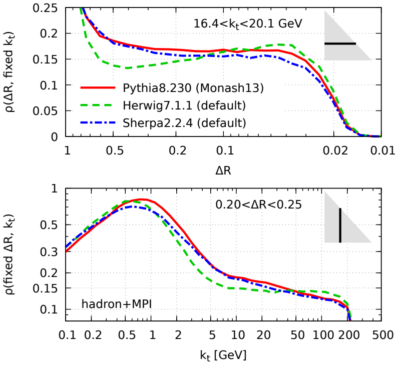
Lund plane and Monte-Carlo generators



Lund plane and Monte-Carlo generators



Slices through the 2D plane

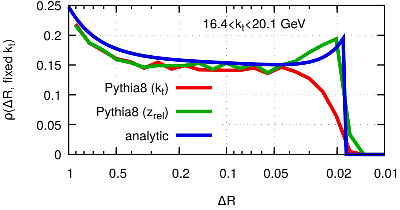


- Flat at small angles (until ISR)
- $\alpha_s(k_t)$ then non-pert bump
- Differences between MCs
- calls for measurement

Lund plane and QCD analytic calculations

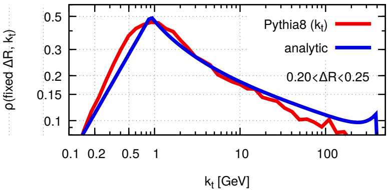
Roughly:

$$\rho(\Delta R, k_t) = \frac{\alpha_s(k_t) C_R}{\pi} z p(z)$$



Additional effects:

- ISR (included here)
- clustering effects
- flavour changes
- energy conservation
- ...

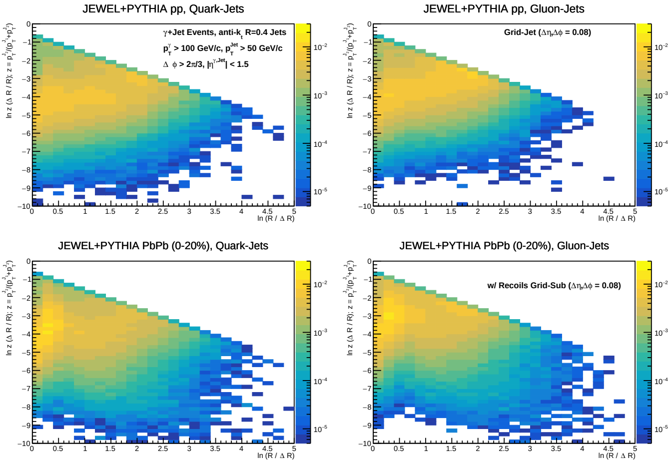


Can be systematically calculated and improved

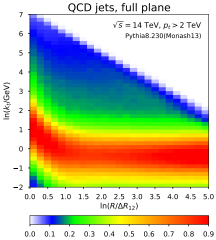
Lund plane and Heavy-ion collisions

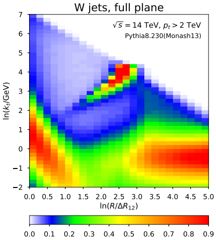
[Y-T. Chien, R. K. Elayavalli]

Quark v. gluon jets in Heavy-ion collisions



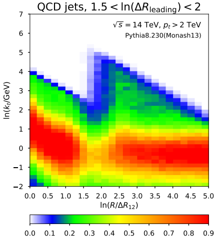
Lund plane and W tagging

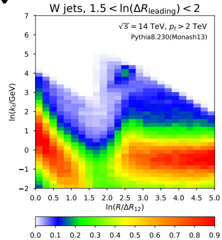




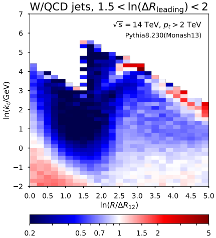
- In practice: split into a “leading” emission and Lund plane for all the others
- Use ratio to build a log-likelihood discriminant

Lund plane
after leading removal

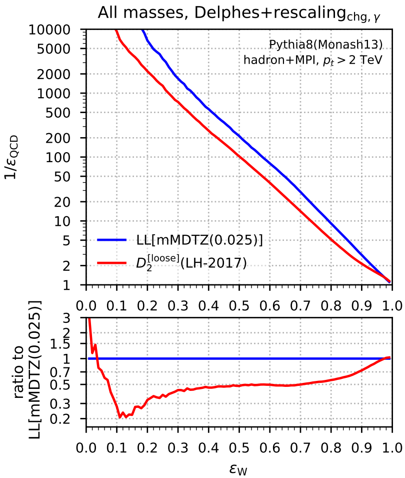




ratio



Lund plane and W tagging



better than "standard" approach
(2-prong tagger + cut on shape)

[D_2^{loose} from Les-Houches PhysTeV 2017]

Lund plane and Machine learning

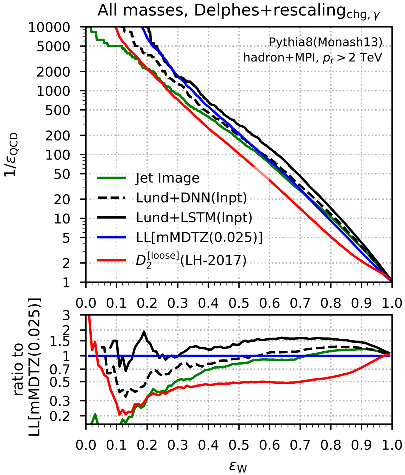
Previous approach can be used as input to Machine Learning:

- "Lund image" fed to a Dense Neural Netword
- "Lund coordinates" fed to a Long Short Term Memeory network

Lund plane and Machine learning

Previous approach can be used as input to Machine Learning:

- "Lund image" fed to a Dense Neural Netword
- "Lund coordinates" fed to a Long Short Term Memeory network



- better than previous ML approaches
- LSTM does even better than Log-Likelihood
- TODO: Log-Likelihood from analytics

[Jet Image: J.Cogan, M.Kagan, E.Strauss, A.Schwartzman, 14; L.de Oliveira, M.Kagan, L.Mackey, B.Nachman and A.Schwartzman, 15]

Conclusions

Jet substructure has gained a lot of importance in the past decade

- Important tool for LHC physics (searches and measurements)
- exciting pQCD phenomenology
  - understanding and development of tools
  - precision pheno at the LHC
  - interesting QCD structure emerging (not covered here)
- Expansion towards new horizons:
  - heavy-ion hard probes
  - machine learning
- Recent work: Lund planes are useful in many aspects

BOOST Annual meeting
around 100 theorists and experimentalists discussing
discussing latest progress in substructure

BOOST Annual meeting
around 100 theorists and experimentalists discussing
discussing latest progress in substructure

July 16-20: BOOST 2018 in Paris
https://indico.cern.ch/e/boost2018