Quarks and gluons in the Lund plane(s)

Gregory Soyez, with Frederic Dreyer, Andrew Lifson, Gavin Salam and Adam Takacs
based on arXiv:1807.04758, arXiv:2007.06578 and arXiv:2112.09140

IPhT, CNRS, CEA Saclay

CERN, June 3 2022

Plan

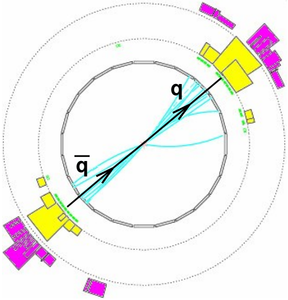
- Context
  - Jets as fundamental objects
  - The onset of jet substructure
- The Lund Plane(s): Picture, logic and construction
- The Lund Plane(s): Applications
  - radiation visualisation
  - analytic viewpoint
  - experimental viewpoint
  - Monte Carlo generators
  - Boosted object tagging
  - Machine Learning
  - heavy ions
  - quark v. gluon

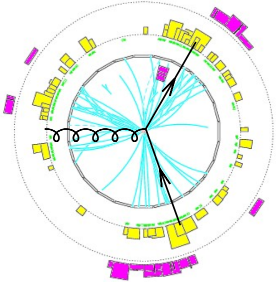
Context:
jets and jet substructure

Jets mimic hard partons

Hard partons (quarks&gluons)
produced in high-energy collisions
branch into more partons mostly at
small angles
→ collimated bunches of hadrons

Jets "collect" these bunches
⇒ jet ≡ proxies to hard partons



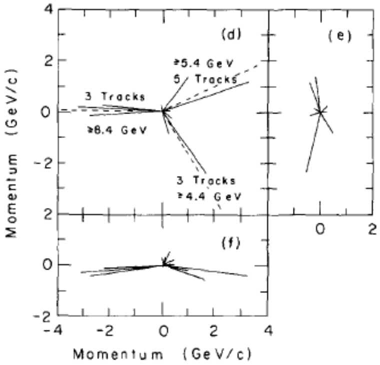


Jets mimic hard partons

Hard partons (quarks&gluons) produced in high-energy collisions branch into more partons mostly at small angles
$\rightarrow$ collimated bunches of hadrons

Jets "collect" these bunches
$\Rightarrow$ jet $\equiv$ proxies to hard partons

- From the discovery of the gluon...
  (as $e^+e^- \rightarrow 3$ jets at TASSO)

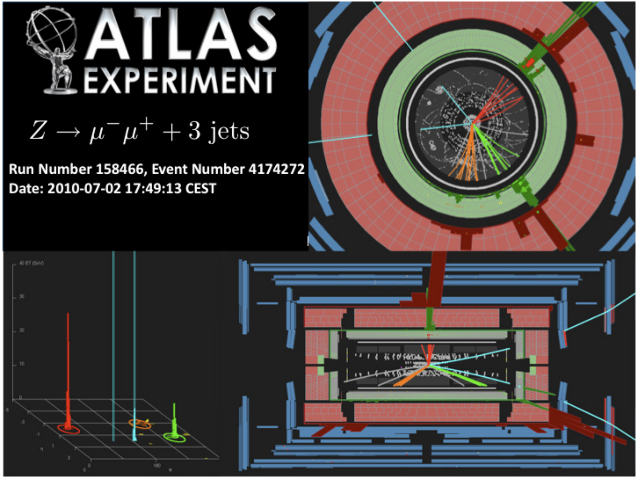


Jets mimic hard partons

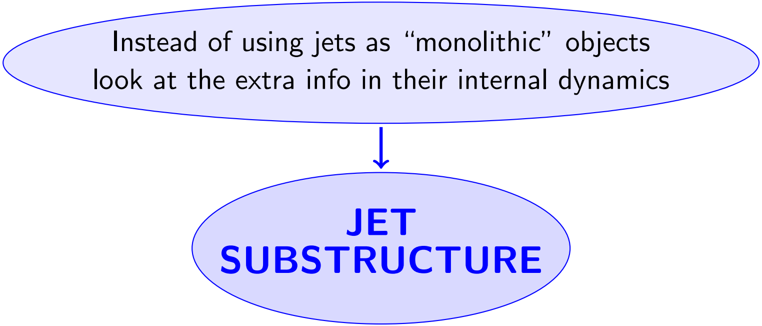
Hard partons (quarks&gluons) produced in high-energy collisions branch into more partons mostly at small angles
$\to$ collimated bunches of hadrons

Jets "collect" these bunches
$\Rightarrow$ jet $\equiv$ proxies to hard partons

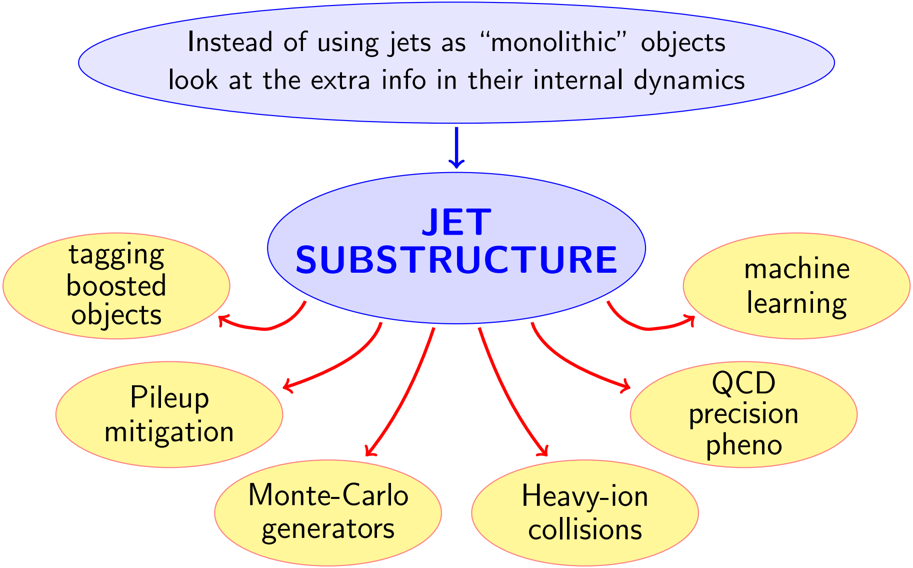
- From the discovery of the gluon... (as $e^+e^- \to 3$ jets at TASSO)
- ... to routine usage at the LHC ($\gtrsim 2/3$ analyses)

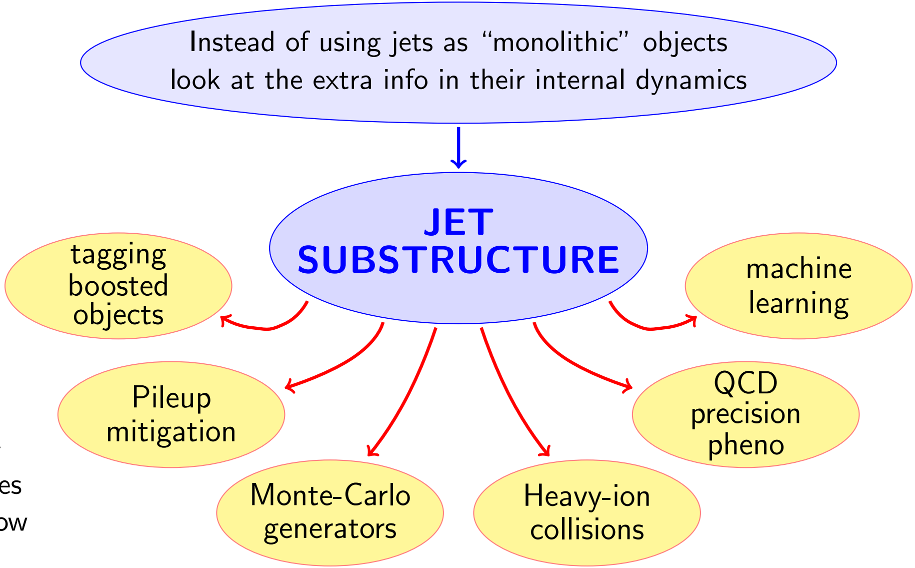


Jet substructure



Jet substructure





Many examples will follow

A decade of substructure tools

(modified)
MassDrop
Tagger

(generalised)
angularities

(recursive)
SoftDrop

N-subjettiness

Trimming

Energy
Correlation
Functions

Pruning

Shower
deconstruct^{n}

Energy flow
Polynomials

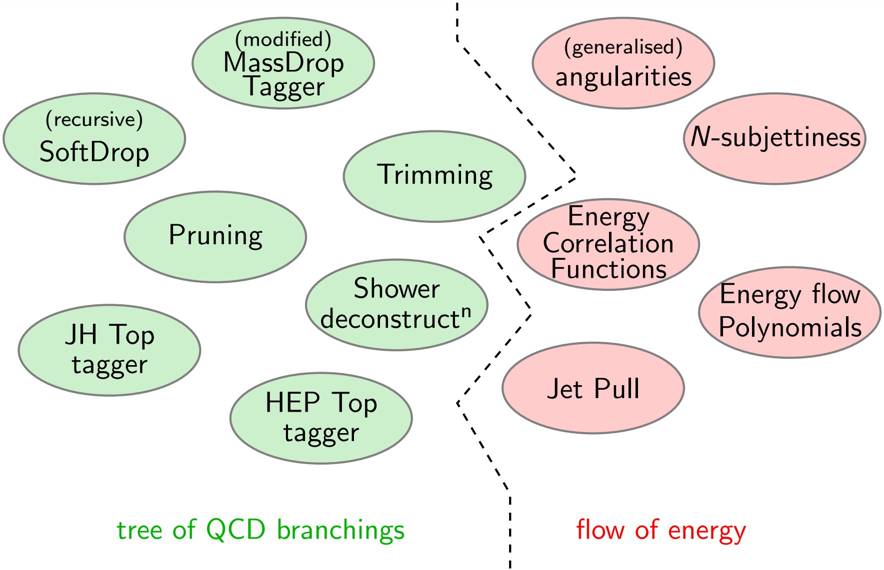
JH Top
tagger

Jet Pull

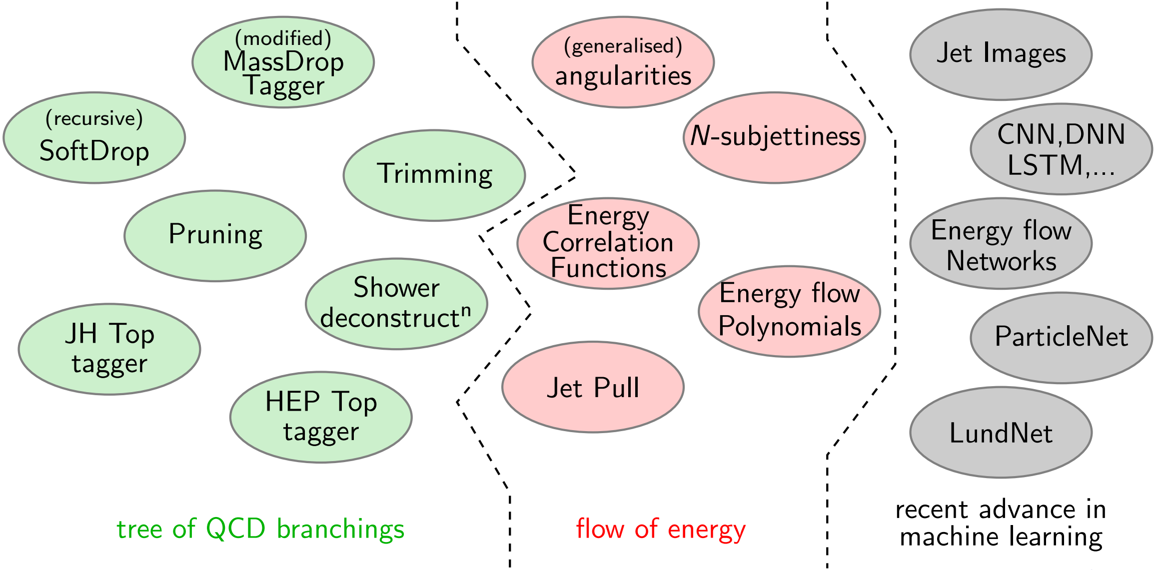
HEP Top
tagger

* Non-exhaustive/biased/... list

A decade of substructure tools



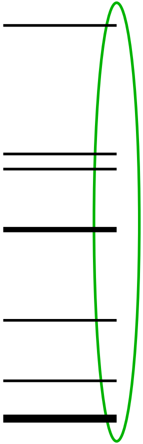
A decade of substructure tools



The Lund Jet Plane(s)
definition/logic

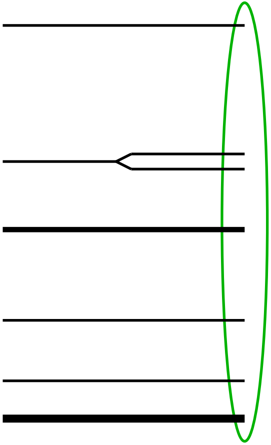
The Lund plane(s) representation (1/3)

use Cambridge/Aachen to iteratively recombine the closest pair



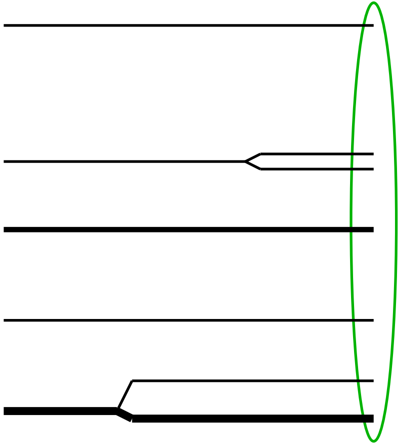
The Lund plane(s) representation (1/3)

use Cambridge/Aachen to iteratively recombine the closest pair

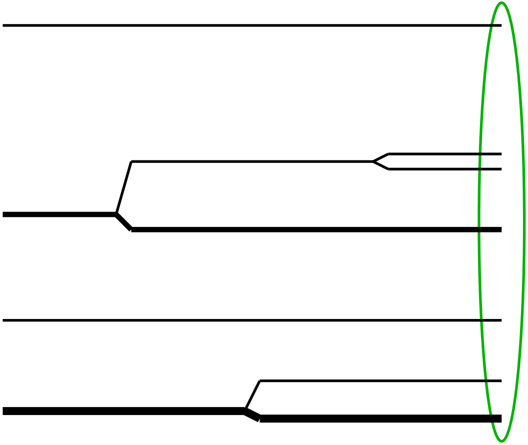


The Lund plane(s) representation (1/3)

use Cambridge/Aachen to iteratively recombine the closest pair

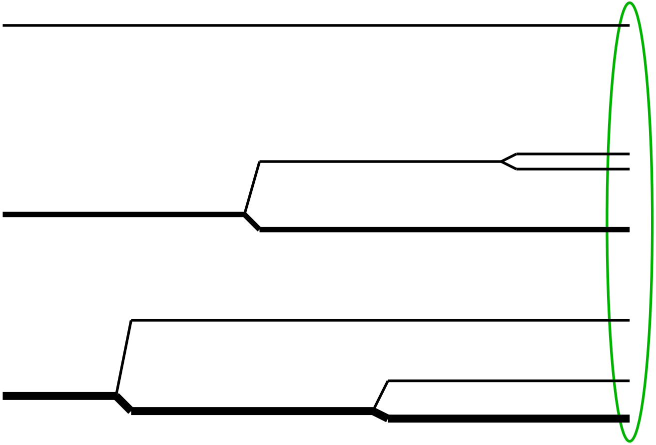


use Cambridge/Aachen to iteratively recombine the closest pair



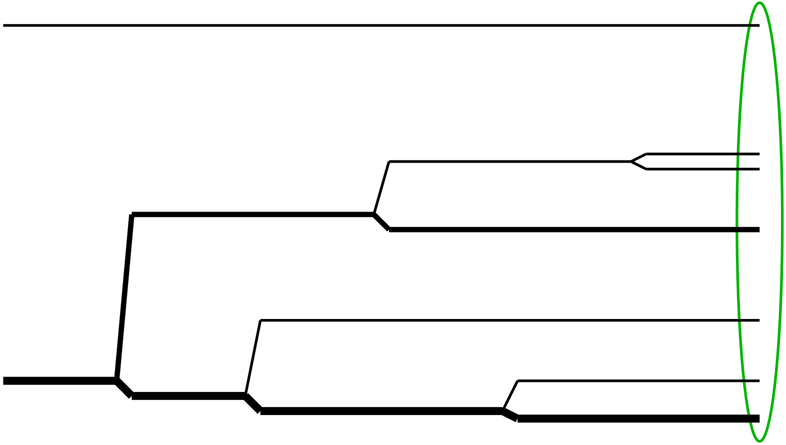
The Lund plane(s) representation (1/3)

use Cambridge/Aachen to iteratively recombine the closest pair



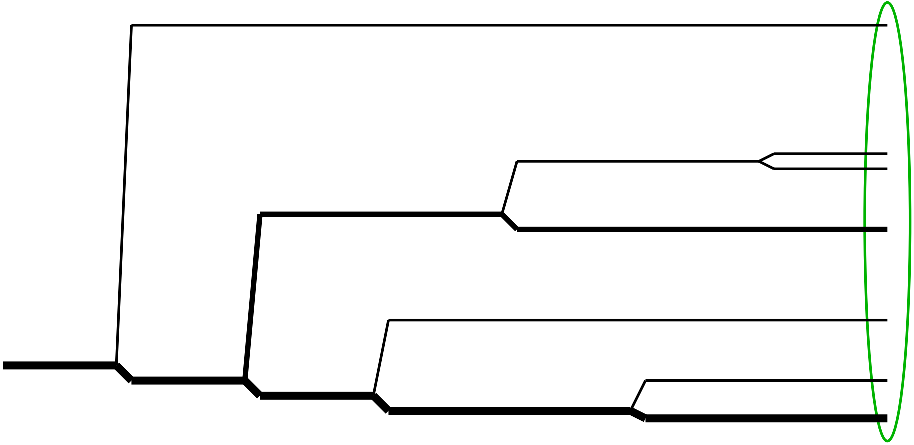
The Lund plane(s) representation (1/3)

use Cambridge/Aachen to iteratively recombine the closest pair



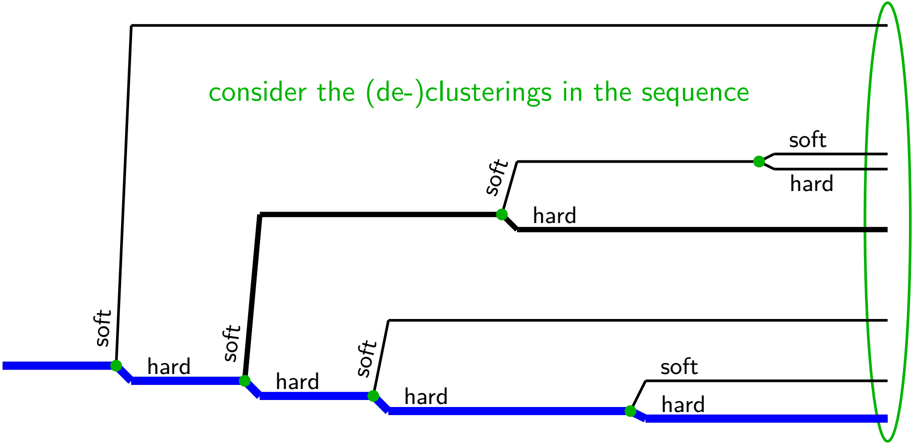
The Lund plane(s) representation (1/3)

use Cambridge/Aachen to iteratively recombine the closest pair



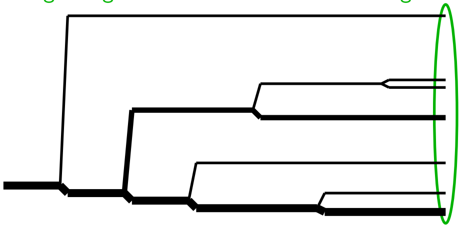
The Lund plane(s) representation (1/3)

use Cambridge/Aachen to iteratively recombine the closest pair



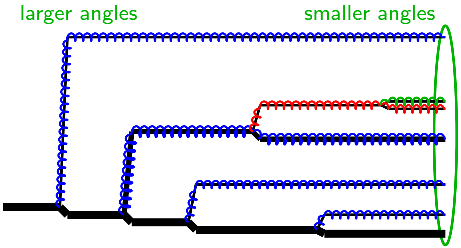
Note: conceptually the largest-energy (p_t or z) branch ≡ emissions from the "leading parton"

The Lund plane(s) representation (2/3)



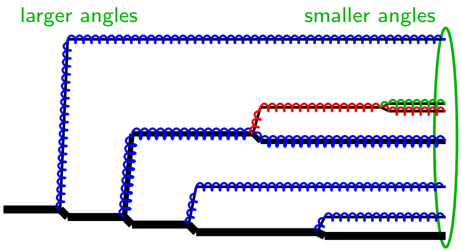
- closely follows our beloved
  angular ordering

The Lund plane(s) representation (2/3)

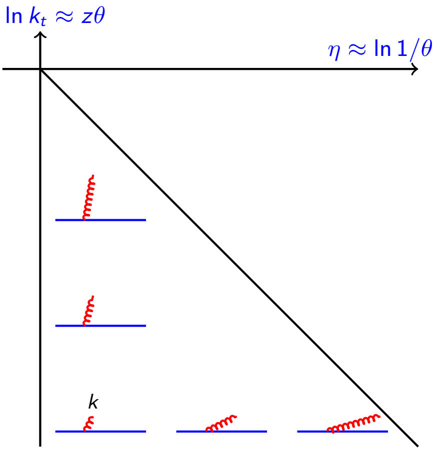


- closely follows our beloved angular ordering
- i.e. mimics partonic cascade

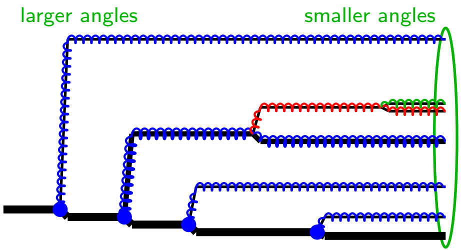
The Lund plane(s) representation (2/3)

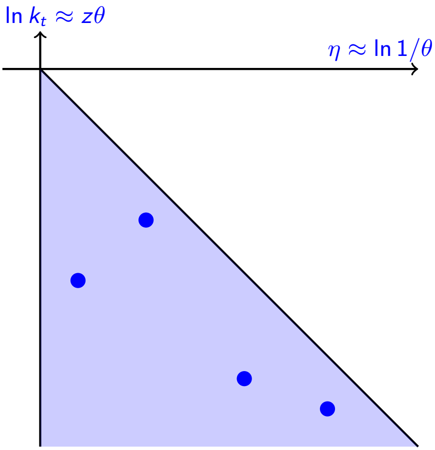


- closely follows our beloved angular ordering
- i.e. mimics partonic cascade
- can be organised in Lund planes



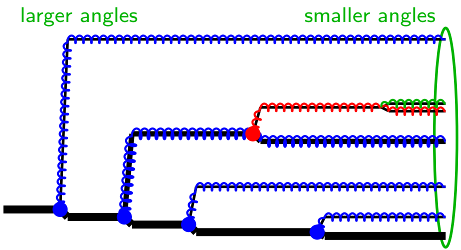
The Lund plane(s) representation (2/3)



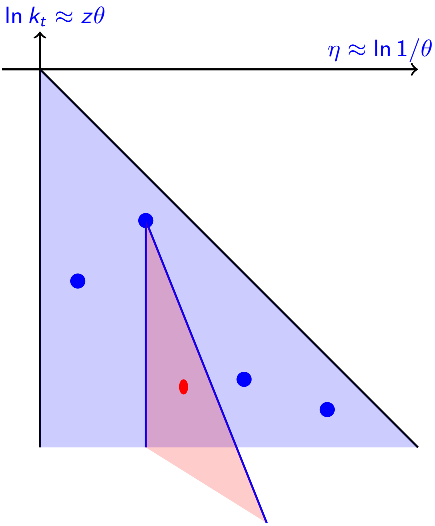


- closely follows our beloved angular ordering
- i.e. mimics partonic cascade
- can be organised in Lund planes
  - primary

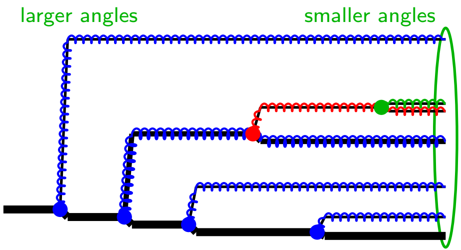
The Lund plane(s) representation (2/3)



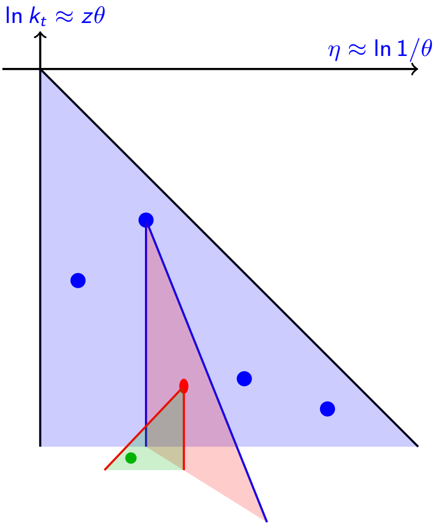
- closely follows our beloved angular ordering
- i.e. mimics partonic cascade
- can be organised in Lund planes
  - primary
  - secondary



The Lund plane(s) representation (2/3)



- closely follows our beloved angular ordering
- i.e. mimics partonic cascade
- can be organised in Lund planes
  - primary
  - secondary
  - ...
- Other interesting variables: ψ, z, m, ...

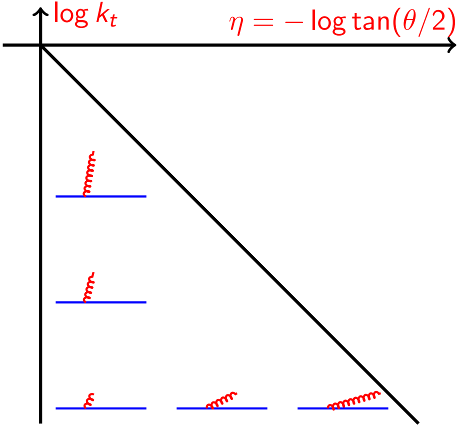


The Lund Jet Plane(s)

(many) applications

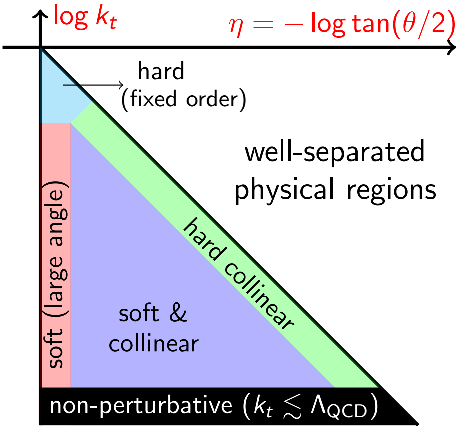
Application: different regions of sensitivity

Concentrate on the **primary plane**

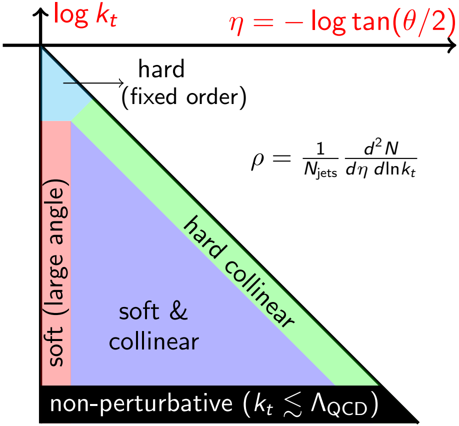


Application: different regions of sensitivity

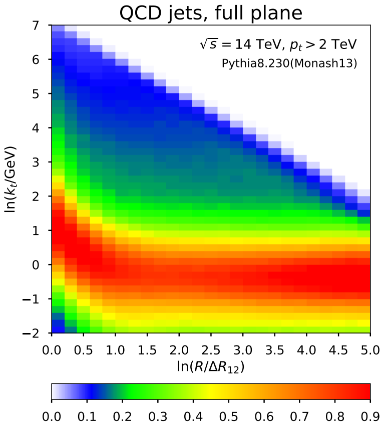
Concentrate on the **primary plane**



Concentrate on the **primary plane**

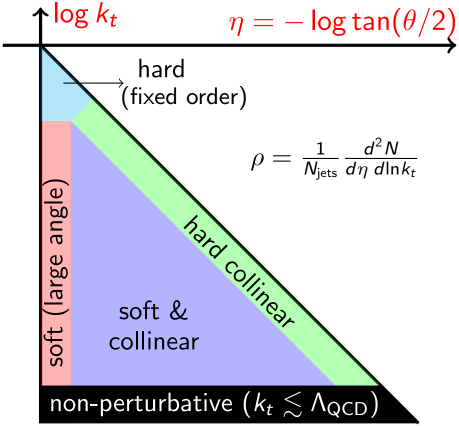


- meaningfull radiation pattern in each region
  $\alpha_s(k_t)$ running, NP at $\lesssim$ 5 GeV, ISR+MPI effects at large angles, ...

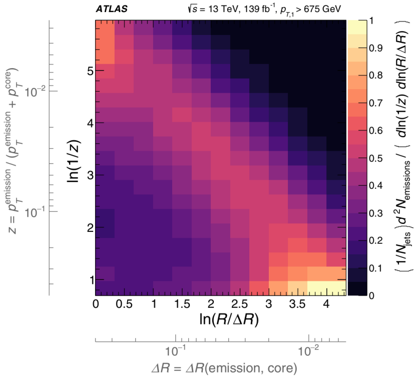


Application: different regions of sensitivity

Concentrate on the **primary plane**

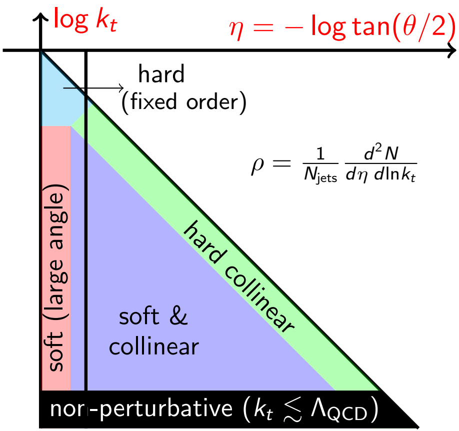


- meaningfull radiation pattern in each region
- measured by ATLAS [ATLAS, 2004.03540]
  watch out: different projection: $\ln k_t \to \ln z$

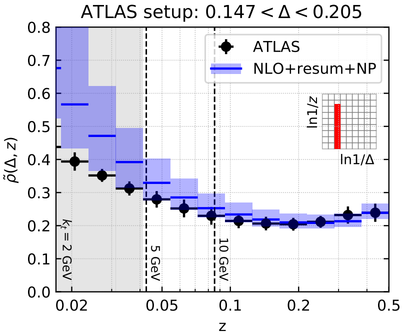


Application: different regions of sensitivity

Concentrate on the **primary plane**

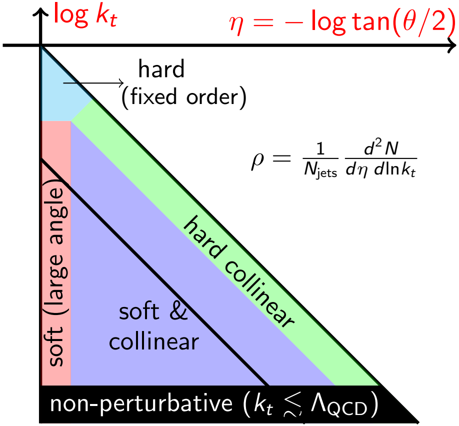


- meaningfull radiation pattern in each region
- measured by ATLAS [ATLAS, 2004.03540]
- helpful comparison to analytics
  NLO(exact $\mathcal{O}(\alpha_s^2)$)+NLL(all-orders separated emissions)+NP(from MC)

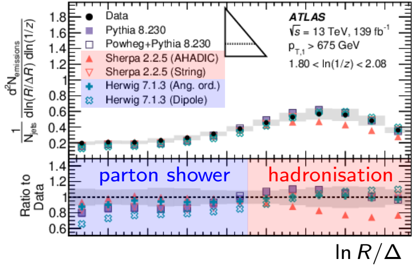


see also [R.Medves,A.Soto,GS,2205.02861] for a multiplicity observable

Concentrate on the **primary plane**

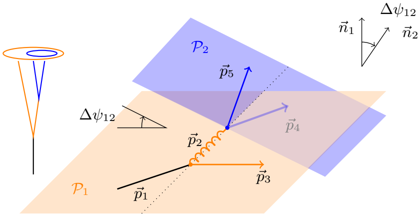


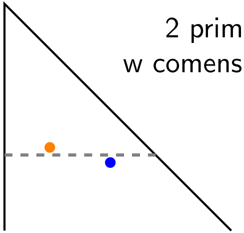
- meaningfull radiation pattern in each region
- measured by ATLAS [ATLAS, 2004.03540]
- helpful comparison to analytics
- helpful comparison to MC generators



Crafted observables for MC studies: example $\Delta\Psi_{12}$

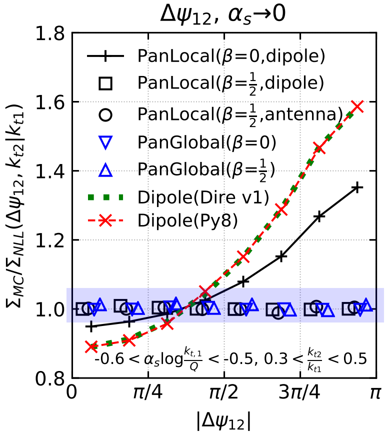
Azimuth between 1[st] and 2[nd] prim. declust.





2 primaries
w comensurate k_t

[M.Dasgupta,F.Dreyer,K.Hamilton,P.Monni,G.Salam,GS,2002.11114]

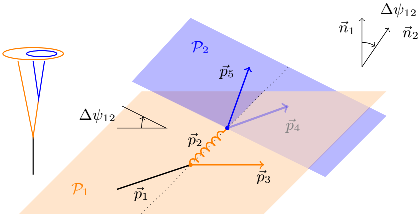


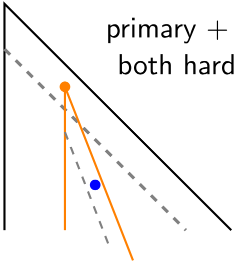
Expected ratio of 1 at NLL

NLL failures for "standard" showers

"New" PanScales shower OK at NLL

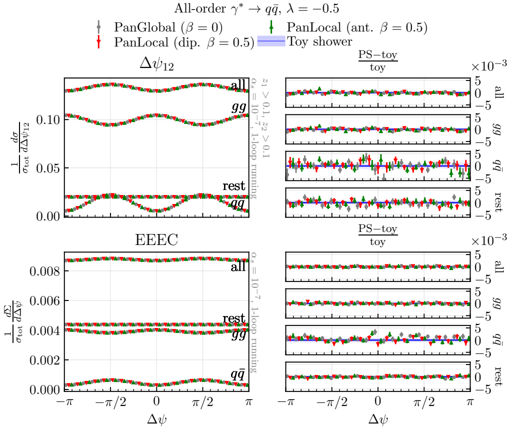
Azimuth between 1[st] and 2[nd] prim. declust.





primary + secondary
both hard-collinear

[A.Karlberg,G.Salam,L.Scyboz,R.Verheyen,2103.16526]

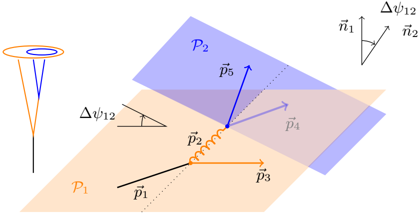


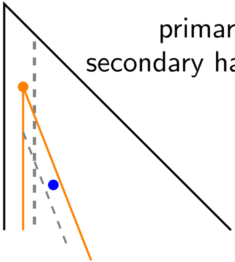
Sensitive to (collinear) spin
"New" PanScales shower have spin at NLL
agrees w EEEC from 2011.02492 (EEEC less sensitive)

Crafted observables for MC studies: example $\Delta\Psi_{12}$

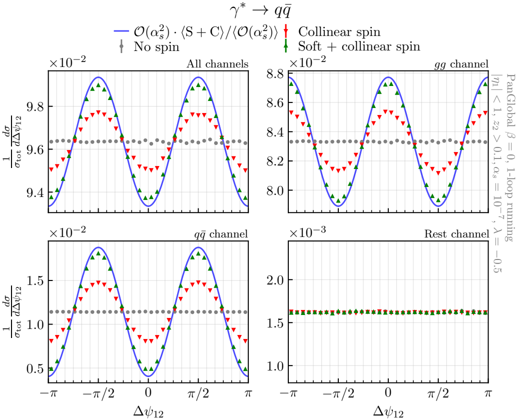
[K.Hamilton,A.Karlberg,G.Salam,L.Scyboz,R.Verheyen,2111.01161]

Azimuth between 1st and 2nd prim. declust.





primary soft
secondary hard-collinear



Sensitive to (soft) spin

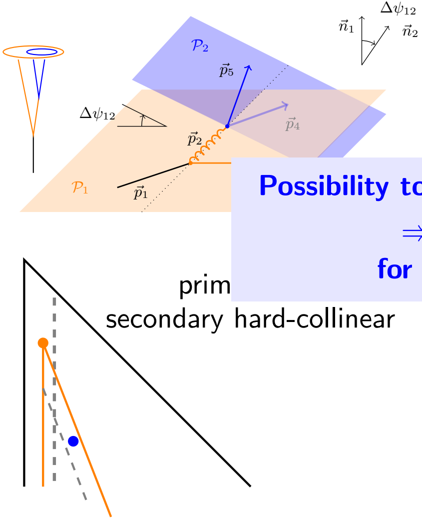
"New" PanScales shower have spin at NLL

first all-order result

Crafted observables for MC studies: example $\Delta\Psi_{12}$

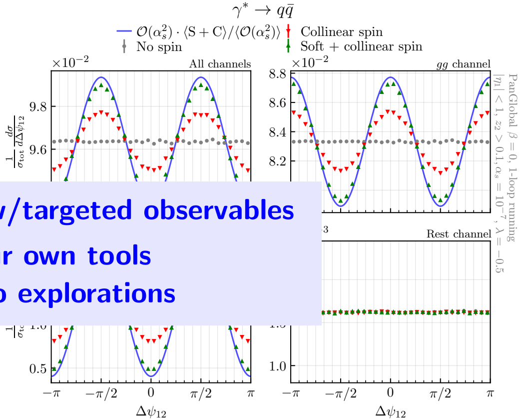
[K.Hamilton,A.Karlberg,G.Salam,L.Scyboz,R.Verheyen,2111.01161]

Azimuth between 1^{st} and 2^{nd} prim. declust.



prim

secondary hard-collinear



Possibility to build new/targeted observables

⇒ build your own tools

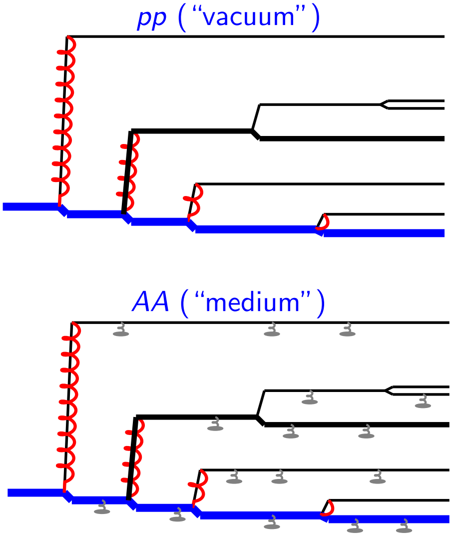
for new pheno explorations

Sensitive to (soft) spin

"New" PanScales shower have spin at NLL

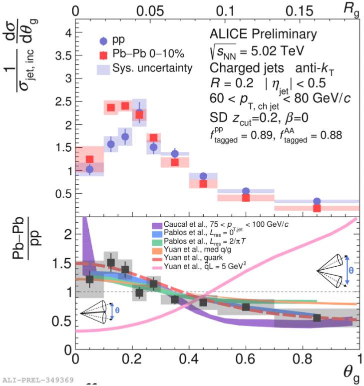
first all-order result

Application: heavy-ion collisions



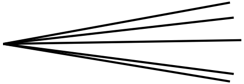
Check how radiation changes when interacting with the QGP

Example: largest-θ emission with $z > z_{\text{cut}}$



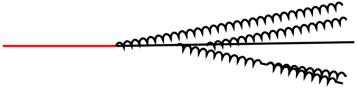
Application to boosted object tagging

THE typical substructure application: given a high-p_t jet

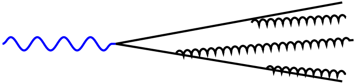


$^{(*)}$ or Z, H, top, ...

Is it a "standard" QCD jet...

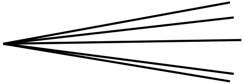


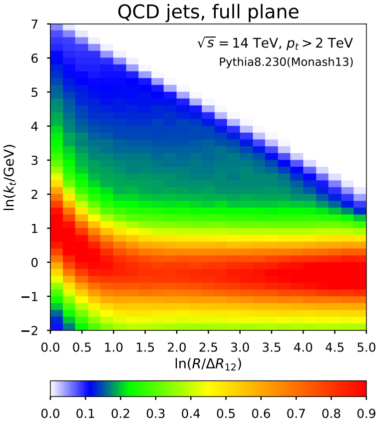
...or a **boosted** W-boson$^{(*)}$ decay?



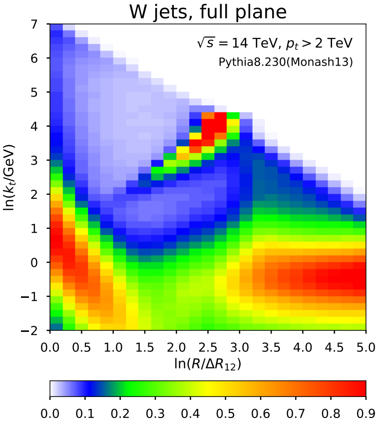
Decay angle: $\theta \propto \frac{m}{p_t}$

THE typical substructure application: given a high-p_t jet

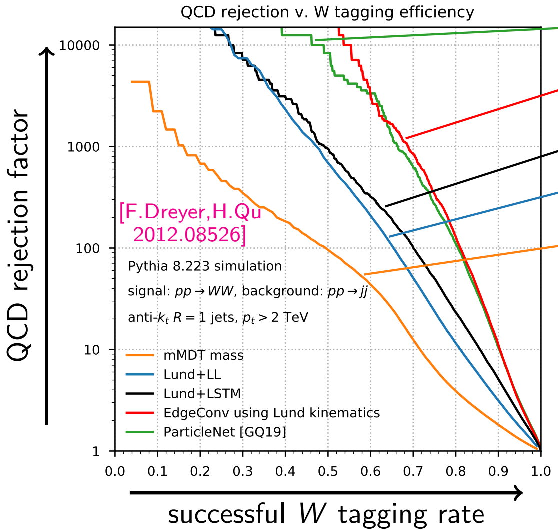




clearly different radiation patterns



Example performance



[graph network using 4-vector(more complex)]

Graph Net trained on full Lund tree

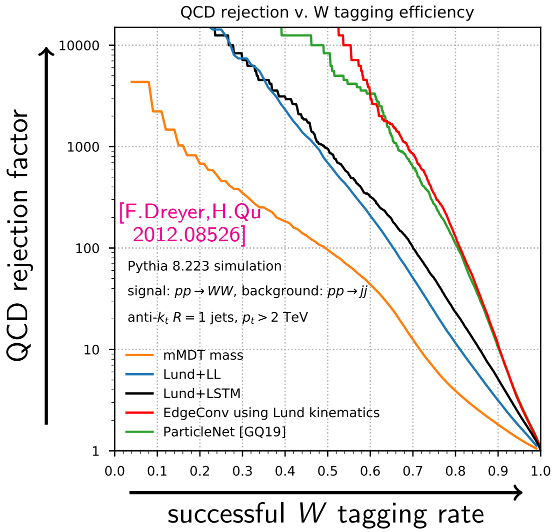
Deep-learning (LSTM) using Lund primaries

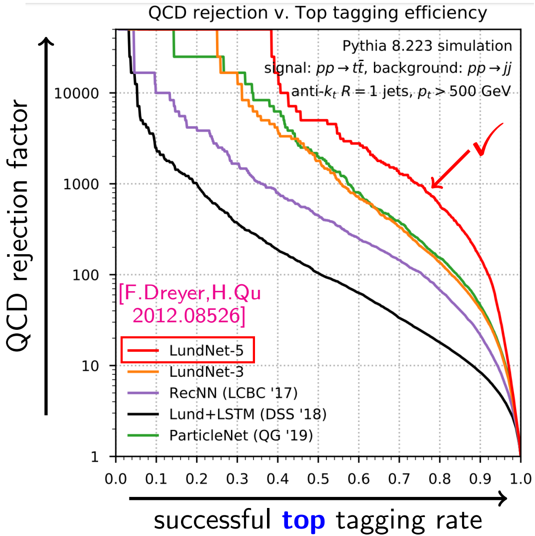
Likelihood ratio based on prim. Lund images

Historical mMDT/SoftDrop

Main messages

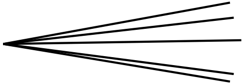
- Combination with Deep-Learning methods
- Large gain from info in the primary plane
- Yet another gain from the full Lund tree


QCD rejection v. W tagging efficiency
QCD rejection factor
successful W tagging rate
[F.Dreyer,H.Qu 2012.08526]
Pythia 8.223 simulation
signal: $pp \to WW$, background: $pp \to jj$
anti-k_t $R = 1$ jets, $p_t > 2$ TeV
mMDT mass
Lund+LL
Lund+LSTM
EdgeConv using Lund kinematics
ParticleNet [GQ19]

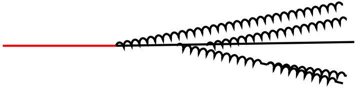

QCD rejection v. Top tagging efficiency
QCD rejection factor
successful top tagging rate
Pythia 8.223 simulation
signal: $pp \to t\bar{t}$, background: $pp \to jj$
anti-k_t $R = 1$ jets, $p_t > 500$ GeV
[F.Dreyer,H.Qu 2012.08526]
LundNet-5
LundNet-3
RecNN (LCBC '17)
Lund+LSTM (DSS '18)
ParticleNet (QG '19)

Application: quark v. gluons

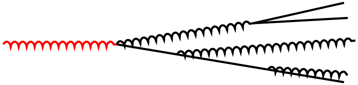
Last application for today: given a high-p_t jet



Is it a quark-initiated jet...



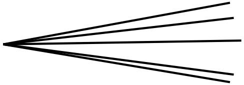
...or a gluon-initiated jet?



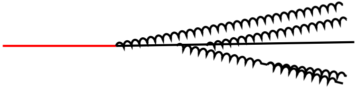
WATCH OUT: technically "quark v. gluon" is not a well-defined concept in QCD (see arXiv:1704.03878)

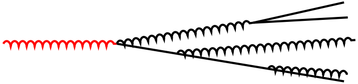
Application: quark v. gluons

Last application for today: given a high-p_t jet



Is it a quark-initiated jet... ...or a gluon-initiated jet?





Question: can we answer given the Lund dweclusterings in a jet?

Quark v. gluon jets: I. approach

Optimal discriminant (Neyman–Pearson lemma)

$$\mathbb{L}_{\text{prim,tree}} = \frac{p_g(\mathcal{L}_{\text{prim,tree}})}{p_q(\mathcal{L}_{\text{prim,tree}})}$$

Quark v. gluon jets: I. approach

Optimal discriminant (Neyman–Pearson lemma)

$$\mathbb{L}_{\text{prim,tree}} = \frac{p_g(\mathcal{L}_{\text{prim,tree}})}{p_q(\mathcal{L}_{\text{prim,tree}})}$$

Approach #1

Deep-learn $\mathbb{L}_{\text{prim,tree}}$

LSTM with $\mathcal{L}_{\text{prim}}$ or Lund-Net with $\mathcal{L}_{\text{tree}}$

Quark v. gluon jets: I. approach

Optimal discriminant (Neyman–Pearson lemma)

$$\mathbb{L}_{\text{prim,tree}} = \frac{p_g(\mathcal{L}_{\text{prim,tree}})}{p_q(\mathcal{L}_{\text{prim,tree}})}$$

Approach #1

Deep-learn $\mathbb{L}_{\text{prim,tree}}$

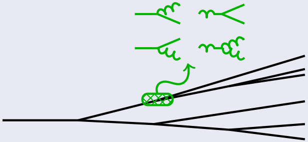
LSTM with $\mathcal{L}_{\text{prim}}$ or Lund-Net with $\mathcal{L}_{\text{tree}}$

Approach #2

Use pQCD to *calculate* $p_{q,g}(\mathcal{L}_{\textbf{prim,tree}})$

- Consider $k_t \geq k_{t,\text{cut}}$ to stay perturbative
- Leading order: $\mathbb{L}_{\text{prim,tree}} \leftrightarrow$ number of primary emissions!
  - Primary emissions get factor $\frac{2\alpha_s(k_t)C_i}{\pi}$ $(C_q = C_F,\ C_g = C_A)$
  - Subsidiary emissions get a factor $\frac{2\alpha_s(k_t)C_A}{\pi}$
- Next order: include collinear effects (incl. flavour changing)

  $+$ running coupling effects $+$ Sudakov for virtuals $+$ clustering effects at commensurate angles



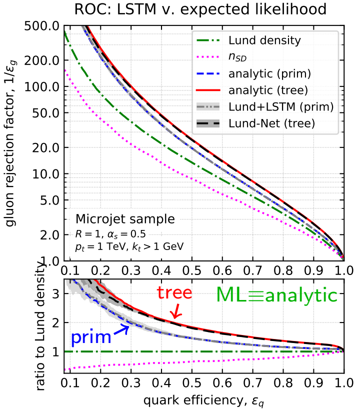
Quark v. gluon jets: II. ML validation

our analytic discriminant is exact/optimal in the dominant collinear limit $\theta_1 \gg \theta_2 \gg \cdots \gg \theta_n$

$\Rightarrow$ ML expected to give the same performance

Quark v. gluon jets: II. ML validation

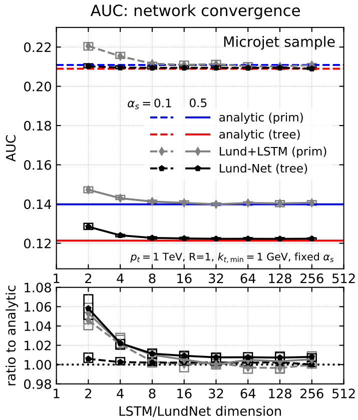
our analytic discriminant is exact/optimal in the dominant collinear limit $\theta_1 \gg \theta_2 \gg \cdots \gg \theta_n$
$\Rightarrow$ ML expected to give the same performance



ROC curves agree

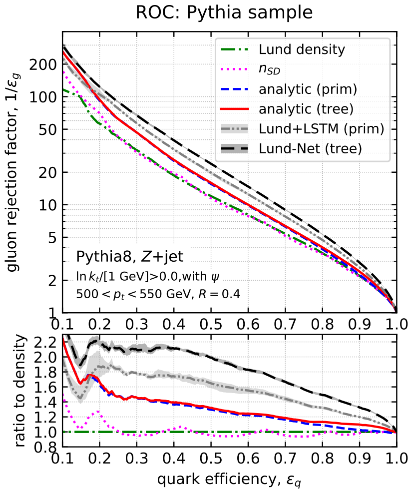
Microjet
≡
exact
pure-collinear

[M.Dasgupta, F.Dreyer
G.P.Salam, G.Soyez,
1411.5182]



Converges for large-enough networks

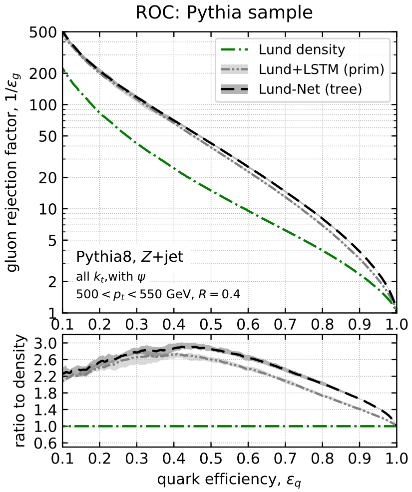
$pp \to Zq$ v. $pp \to Zg$ $\quad (p_t \sim 500$ GeV, $R = 0.4)$



- clear performance ordering:
  1. Lund+ML > Lund analytic > ISD
  2. tree > prim

Quark v. gluon jets: III. performance

$pp \to Zq$ v. $pp \to Zg$ $\quad (p_t \sim 500$ GeV, $R = 0.4)$



- clear performance ordering:
  1. Lund+ML > Lund analytic > ISD
  2. tree > prim
- larger gains with no k_t cut

  (several potential reasons)
- Q: analytics to other systems ($W/Z/H$, top)?

Take-home messages

Jets are ubiquitous at colliders

Jet substructure

- Jets have a substructure (internal dynamics) which is worth exploiting
- Now routinely used at the LHC
- Broad applications: tagging, pQCD, measurements, Monte Carlo, heavy-ions, machine-learning, ...

Physics with Lund-plane(s)

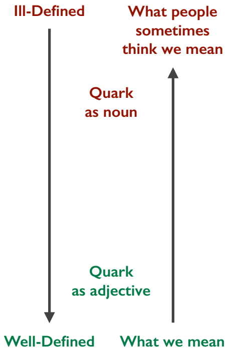
- Construction with clear physics properties
  - Organised in trees respecting angular ordering
  - Different physics effects contribute to different regions
  - Opens possibilities to craft your own observables
- Broad applications: tagging, pQCD, measurements, Monte Carlo, heavy-ions, machine-learning, ...

Backup

Quark v. gluon jets: 0. basic considerations

What is a Quark Jet?

From lunch/dinner discussions



A quark parton

A Born-level quark parton

The initiating quark parton in a final state shower

An eikonal line with baryon number 1/3 and carrying triplet color charge

A quark operator appearing in a hard matrix element in the context of a factorization theorem

A parton-level jet object that has been quark-tagged using a soft-safe flavored jet algorithm (automatically collinear safe if you sum constituent flavors)

A phase space region (as defined by an unambiguous hadronic fiducial cross section measurement) that yields an enriched sample of quarks (as interpreted by some suitable, though fundamentally ambiguous, criterion)

pedestrian summary

- there is no such thing as a "quark" or a "gluon" jet
- well-defined: tagging process A ("quark-enriched"$^{(*)}$) against process B ("gluon-enriched"$^{(*)}$)

$^{(*)}$ ambiguous

Our approach(es)

- discuss process-independent aspects (at least analytically)
- probe changes for different processes

Resilience (1/2)

Question: is your tagger resilient to uncontrolled effects?

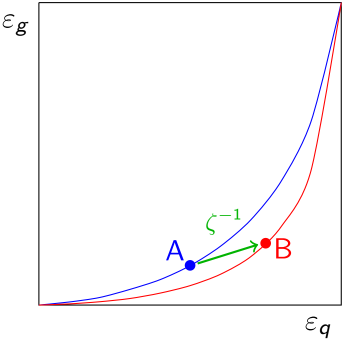
One has:

- a reference sample A
  (e.g. network trained+tested w Pythia)
- an alternate sample B
  (e.g. network tested w Herwig)

We want (for a given working point)

$$\zeta = \left[\left(\frac{\Delta \varepsilon_q}{\langle \varepsilon_q \rangle} \right)^2 + \left(\frac{\Delta \varepsilon_g}{\langle \varepsilon_g \rangle} \right)^2 \right]^{-1}$$

as small as possible.



(would probably deserve a study on its own)

Question: is your tagger resilient to uncontrolled effects?

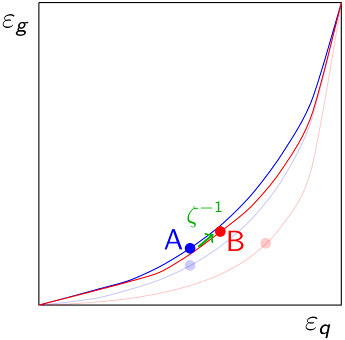
One has:

- a reference sample A
  (e.g. network trained+tested w Pythia)
- an alternate sample B
  (e.g. network tested w Herwig)

We want (for a given working point)

$$\zeta = \left[\left(\frac{\Delta \varepsilon_q}{\langle \varepsilon_q \rangle} \right)^2 + \left(\frac{\Delta \varepsilon_g}{\langle \varepsilon_g \rangle} \right)^2 \right]^{-1}$$

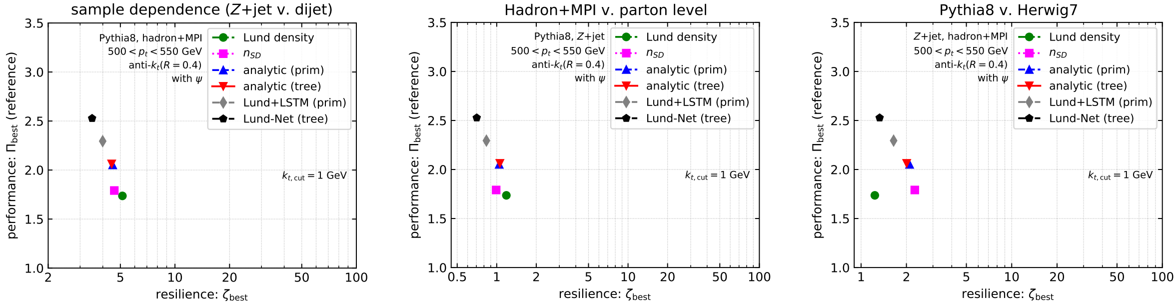
as small as possible.



Less performant
More resilient

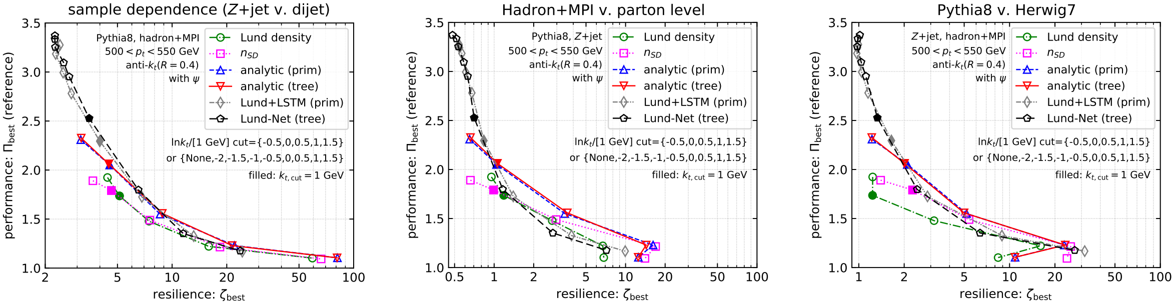
(would probably deserve a study on its own)

Resilience (2/2)



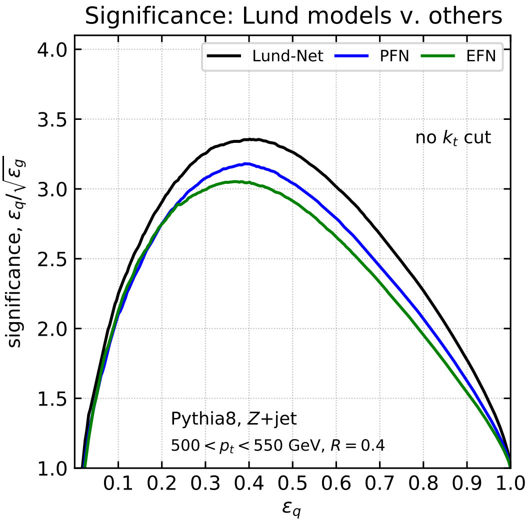
- performance $= \varepsilon_q / \sqrt{\varepsilon_g}$
- working point: $k_{t,\mathrm{cut}} = 1$ GeV, optimal performance (reference: Pythia, hadron+MPI, Z+jet)
- 3 studies: sample (Z+jet v. dijets), NP effects (hadron v. parton), generator (Pythia v. Herwig)
- performance: same ordering as before
- resilience: network-based < Lund analytics $\lesssim n_{SD}$

Resilience (2/2)



- same, varying $k_{t,\text{cut}}$
- for each curve: "standard" trade-off between performance and resilience
- Overall: better behaviour for the new Lund-based approaches:
  - At "large" resilience: better envelope for the Lund analytic approaches
  - At "small" resilience: ML performance gain pays off

Comparison to other approaches: ML-based

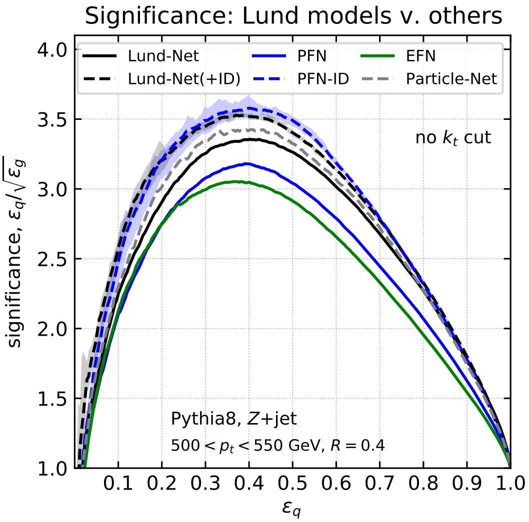


Approaches:

- Lund-Net (full tree)
- Particle-flow network
- Energy-flow network

- small performance gain for Lund
- differences might come from details

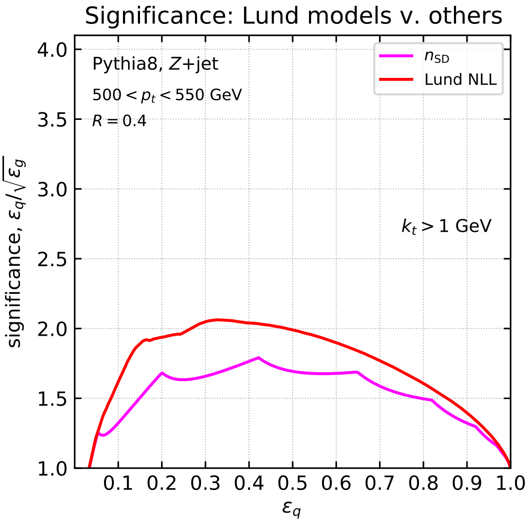
Comparison to other approaches: ML-based



Approaches:

- Lund-Net (full tree)
- Particle-flow network
- Energy-flow network
- Dashed: with PDG-ID
- Particle-Net

- small performance gain for Lund
- differences might come from details
- with PDG-ID: PFN$\sim$Lund$\gtrsim$PNet

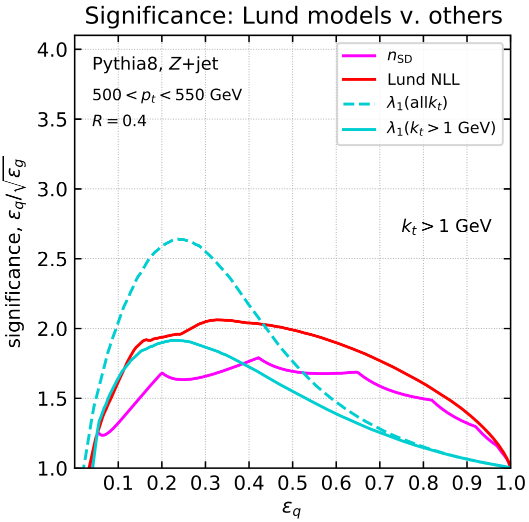


Approaches:

- ISD mult (n_{SD})
- Lund (full tree, analytic)

▶ clear gain from our analytic approach

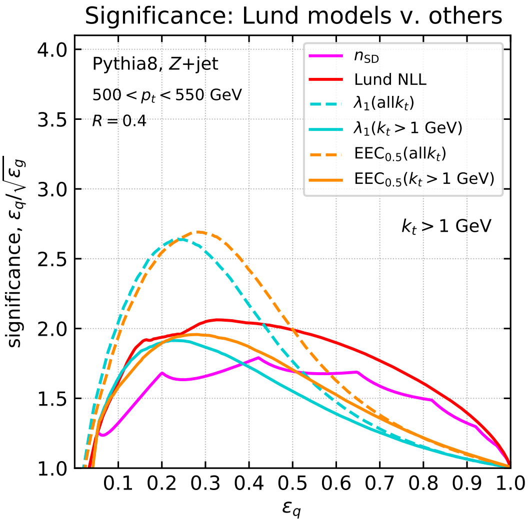
Comparison to other approaches: analytics/shapes



Approaches:

- ISD mult (n_{SD})
- Lund (full tree, analytic)
- width ($\sum_i p_{ti}\Delta R_i$)
- Dashed: use subjets with $k_t > 1$ GeV

- clear gain from our analytic approach
- Different behaviour for shapes
- Lund (expectably) better for same info



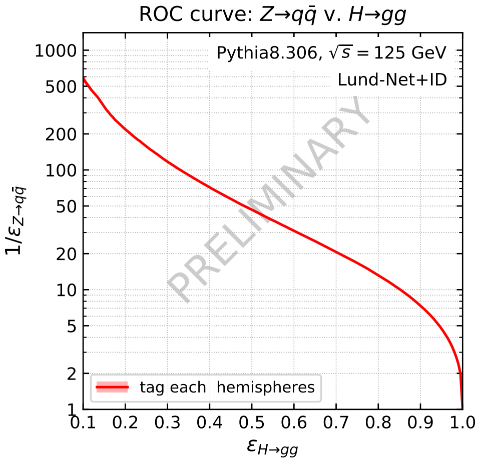
Approaches:

- ISD mult (n_{SD})
- Lund (full tree, analytic)
- width ($\sum_i p_{ti} \Delta R_i$)
- EE correlation ($\sum_{i,j} p_{ti} p_{tj} \Delta R_{ij}^{\beta}$)
- Dashed: use subjets with $k_t > 1$ GeV

- clear gain from our analytic approach
- Different behaviour for shapes
- Lund (expectably) better for same info

Towards full-event tagging

$$e^+e^- \to Z \to q\bar{q} \text{ v. } e^+e^- \to H \to gg \qquad (\sqrt{s} = 125 \text{ GeV, no ISR})$$



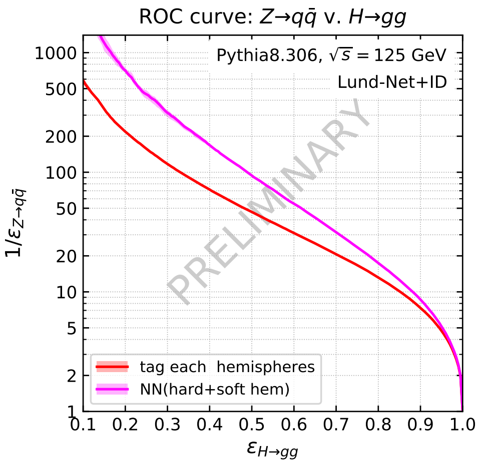
observed performance:

- tagging both hemispheres
  i.e. both jets should be tagged

  full event clearly worse that (jet)^2

Towards full-event tagging

$e^+e^- \to Z \to q\bar{q}$ v. $e^+e^- \to H \to gg$ $(\sqrt{s} = 125$ GeV, no ISR)



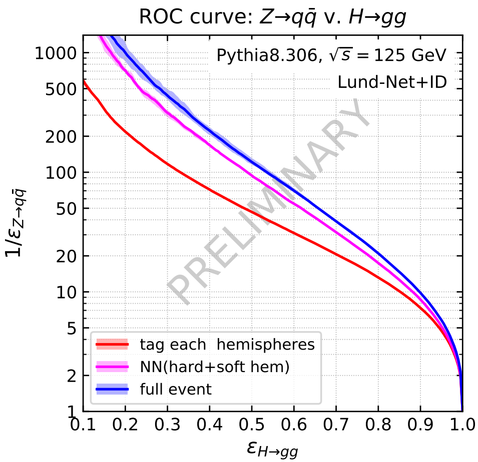
observed performance:

- tagging both hemispheres
- double Lund-Net tag
  train separately on hard & soft hemispheres
  use another NN (or MVA) to combine the two

  clear performance gain

Towards full-event tagging

$$e^+e^- \to Z \to q\bar{q} \text{ v. } e^+e^- \to H \to gg \qquad (\sqrt{s} = 125 \text{ GeV, no ISR})$$

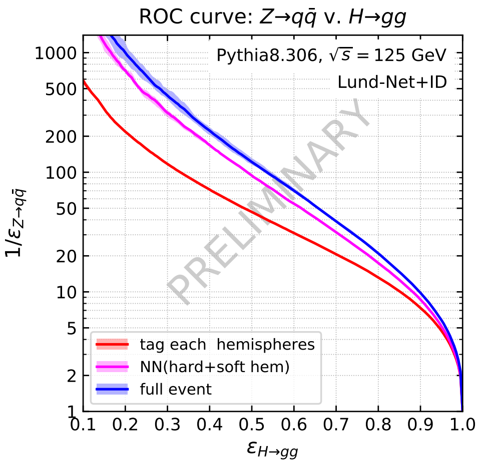


observed performance:

- tagging both hemispheres
- double Lund-Net tag
- Lund-Net for the full event
  Another performance gain

Towards full-event tagging

$$e^+e^- \to Z \to q\bar{q} \text{ v. } e^+e^- \to H \to gg \qquad (\sqrt{s} = 125 \text{ GeV, no ISR})$$



observed performance:

- tagging both hemispheres
- double Lund-Net tag
- Lund-Net for the full event
  Another performance gain

Open questions/work in progress

- How does the analytic do?
  e.g. what gain from full-event tagging?
- Applications to other cases (e.g. at the LHC)?