Jet substructure for parton showers and resummations

Gregory Soyez,
with Frederic Dreyer, Andrew Lifson, Gavin Salam, Adam Takacs, and PanScales

IPhT, CNRS, CEA Saclay, CERN

Parton Showers and Resummations (PSR), 6-8 June 2023

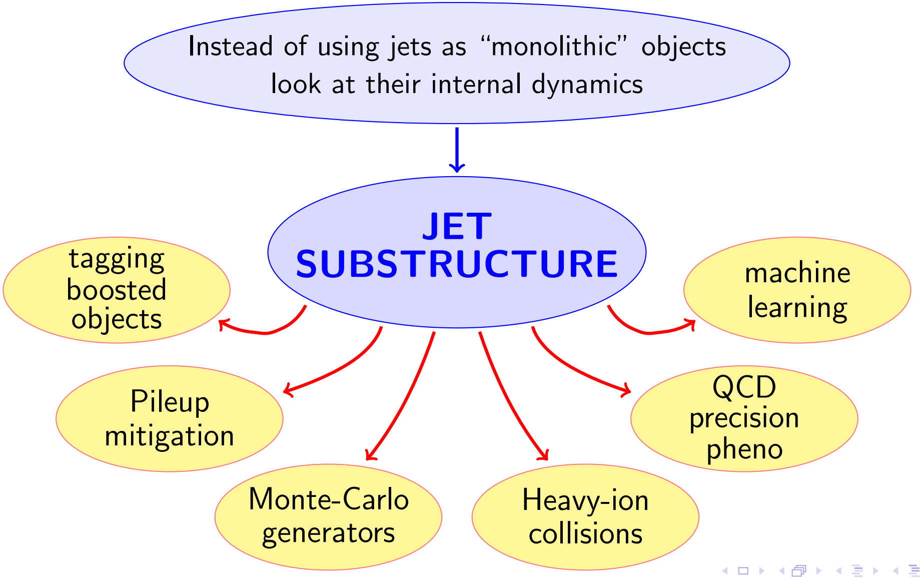
Jet substructure

Instead of using jets as "monolithic" objects
look at their internal dynamics

↓

JET SUBSTRUCTURE

Jet substructure



Almost the organiser's request

This talk

Instead of giving yet-another overview of jet substructure glorifying its wonderful achievements and merits in many areas of QCD, I will instead...

... focus on only 2 examples directly connecting jet substructure to PSR

Almost the organiser's request

This talk

Instead of giving yet-another overview of jet substructure glorifying its wonderful achievements and merits in many areas of QCD, I will instead...

... focus on only 2 examples directly connecting jet substructure to PSR and glorifying its wonderful achievements and merits in 2 areas of QCD!

Main message (in case you fall asleep after this, this should allow you to have sweet dreams)

The substructure of boosted jets is wonderful for PSR

- Boosted jets have (by definition)

$$p_t \gg \Lambda_{\text{QCD}}$$

  i.e. a large phase-space for perturbative emissions
  i.e. fall directly in the area of parton showers and resummations

- a whole library of techniques/observables is readily available

- many possibilities to design new tools focusing on specific tasks

Plan

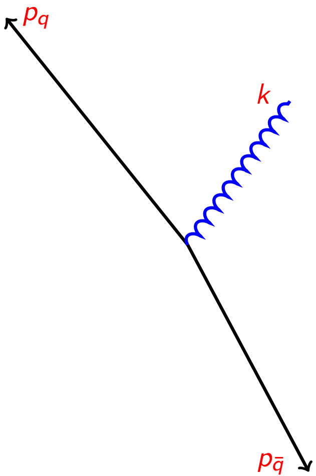
1. Introduction to make sure we are on the same page
   - Lund diagrams: a (historical) conceptual tool for parton showers and resummations
   - promoting to a practical tool for jet physics
2. Example #1: azimuthal correlations in the Lund plane
3. Example #2: quark/gluon tagging

More directly-related examples will be given in Alba Soto Ontoso's talk
See also talks by Basem El-Menoufi, Matt Schwartz, Silvia Ferrario Ravasio and Alexander Karlberg

Warmup: Lund diagrams

A useful representation of radiation in a jet

Take a gluon emission from a $(q\bar{q})$ dipole



Emission:

$$k^\mu \equiv z_q p_q^\mu + z_{\bar{q}} p_{\bar{q}}^\mu + k_\perp^\mu$$

3 degrees of freedom:

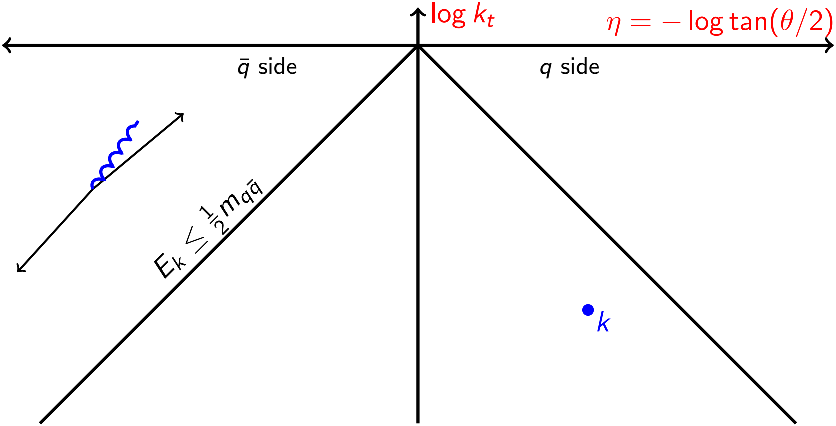
- Rapidity: $\eta = \frac{1}{2} \log \frac{z_q}{z_{\bar{q}}}$
- Transverse momentum: $k_\perp$
- Azimuth: ϕ

In the soft-collinear approximation

$$d\mathcal{P} = \frac{\alpha_s(k_\perp) C_F}{\pi^2} \, d\eta \, \frac{dk_\perp}{k_\perp} \, d\phi$$

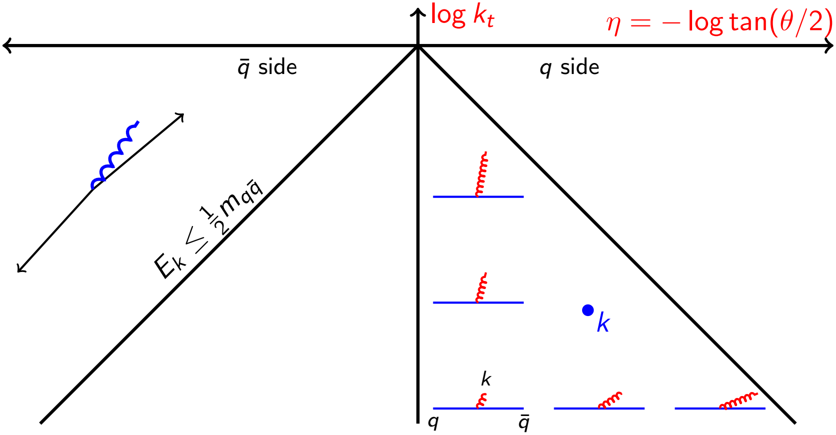
Basic features of QCD radiations: the Lund plane

Lund plane: natural representation uses the 2 “log” variables η and $\log k_\perp$

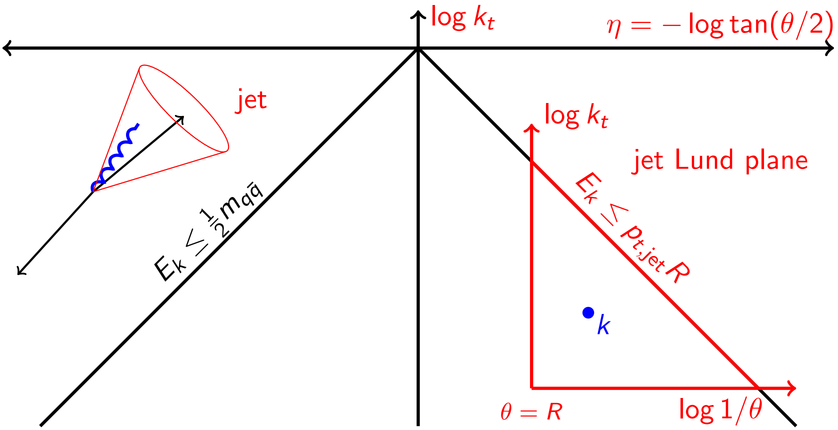


Basic features of QCD radiations: the Lund plane

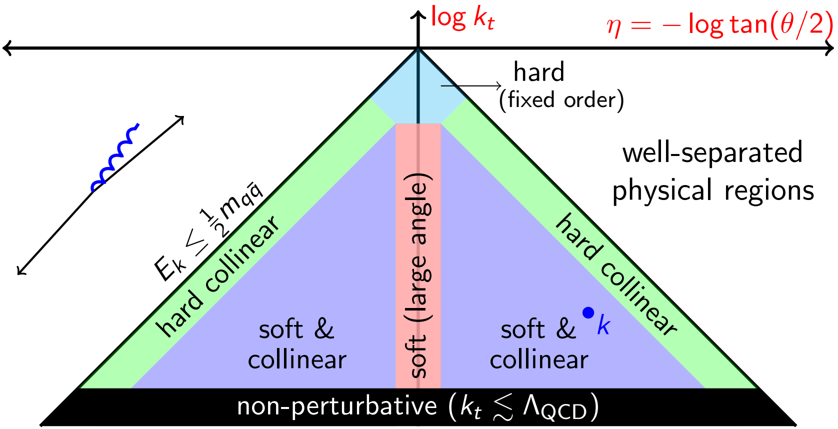
Lund plane: natural representation uses the 2 "log" variables η and $\log k_\perp$

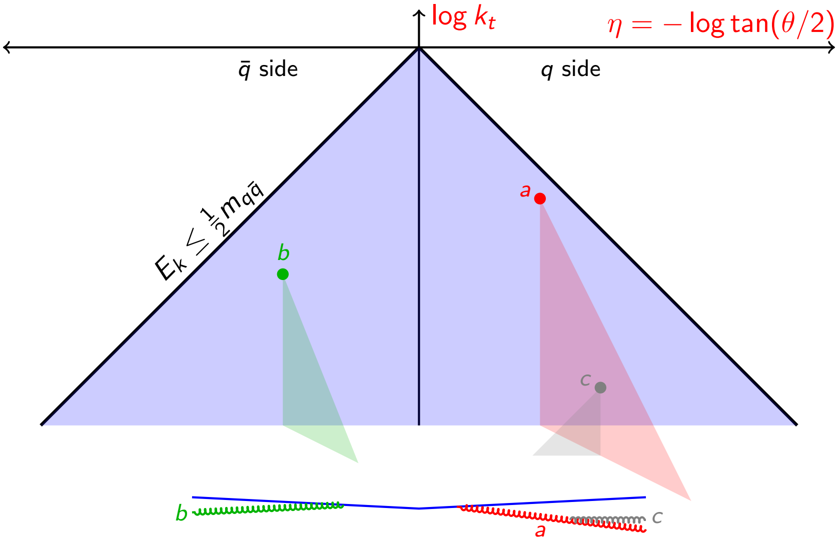


Lund plane: natural representation uses the 2 "log" variables η and $\log k_\perp$



Lund plane: natural representation uses the 2 "log" variables η and $\log k_\perp$

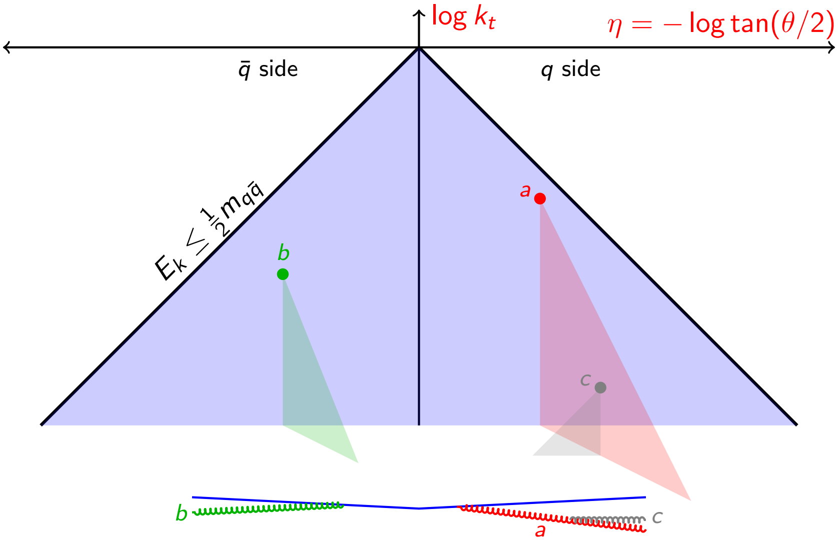




Each emission spawns its own plane/leaf

- a, b primary
- c secondary
- …

Multiple emissions in the Lund plane



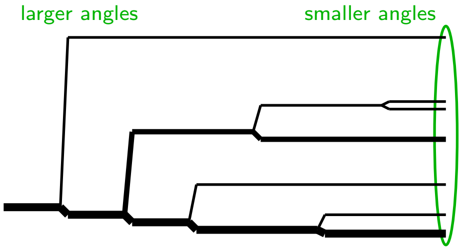
Each emission spawns its own plane/leaf

- a, b primary
- c secondary
- ...

Respects angular ordering ($\theta_c < \theta_a$)

Lund planes:
promoting Lund diagrams to a practical tool

The Lund plane(s) representation



For a given jet

- recluster (the constituents) with the Cambridge/Aachen algorithm

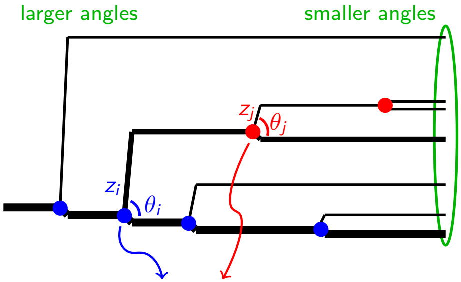
$$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

i.e. cluster from small to large angular distance

- gives a tree structure on the jet

[F.Dreyer,G.Salam,GS,arXiv:1807.04758]

The Lund plane(s) representation



$$\mathcal{T}_i \equiv \{\theta_i, k_{t,i}, z_i, \psi_i, m_i, \dots\}$$

Lund coordinates at each vertex

For a given jet

- recluster (the constituents) with the Cambridge/Aachen algorithm
  $$\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$
  i.e. cluster from small to large angular distance
- gives a tree structure on the jet

For jets in pp: (similar for ee)

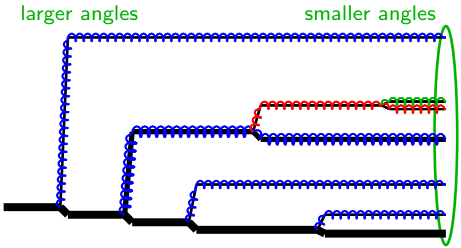
$$\eta = -\ln \Delta R$$

$$k_t = p_{t,\text{soft}} \Delta R, \text{ or } z = \frac{p_{t,\text{soft}}}{p_{t,\text{parent}}}$$

$$\psi \equiv \text{azimuthal angle}$$

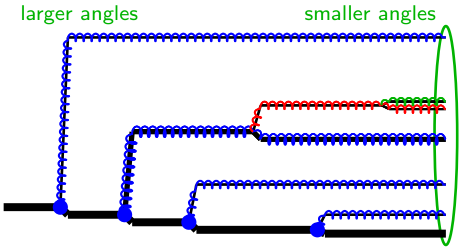
[F.Dreyer,G.Salam,GS,arXiv:1807.04758]

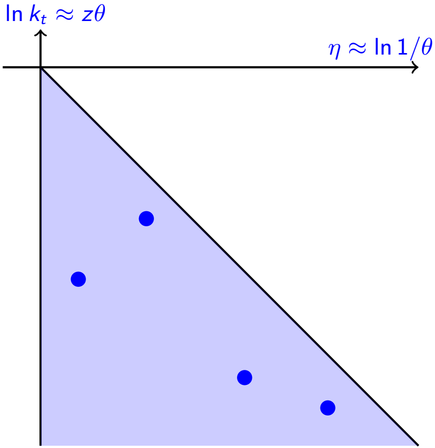
The Lund plane(s) representation



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

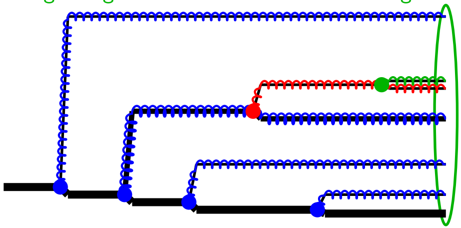
The Lund plane(s) representation

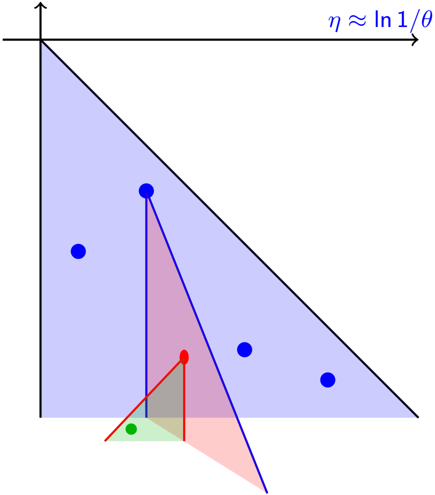




- closely follows angular ordering
  i.e. mimics partonic cascade
- can be organised in Lund planes
  - primary

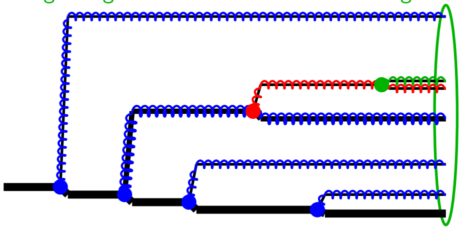
The Lund plane(s) representation



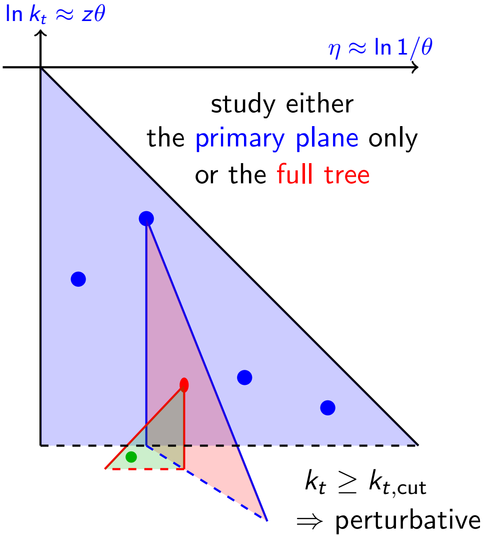


- closely follows angular ordering
  i.e. mimics partonic cascade
- can be organised in Lund planes
  - primary
  - secondary
  - ...

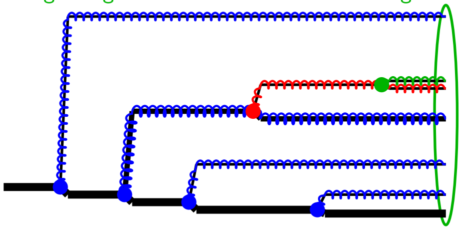
The Lund plane(s) representation



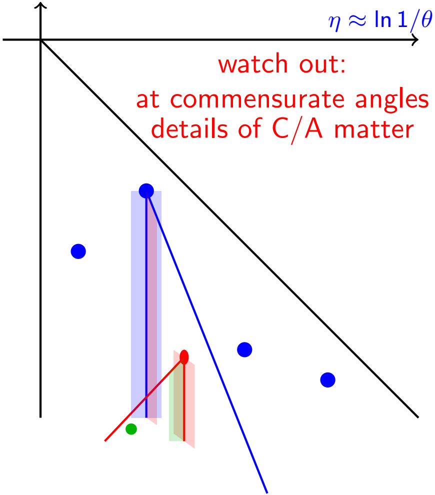
- closely follows angular ordering
  i.e. mimics partonic cascade
- can be organised in Lund planes
  - primary
  - secondary
  - ...



The Lund plane(s) representation



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



Take-home messages

Message #1

Lund diagrams represent (multiple) radiation across scales

- natural for thinking about resummations and parton showers
- different physical regions (soft, collinear, hard, non-perturbative) well separated
- organised in planes respecting angular ordering

Message #2

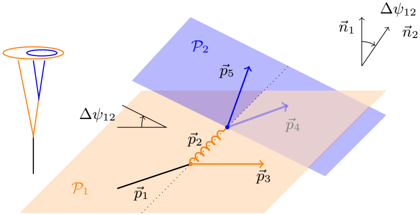
For a jet (or an *ee* event) one can **construct** a Lund-plane(s) structure capturing the properties of Lund diagrams

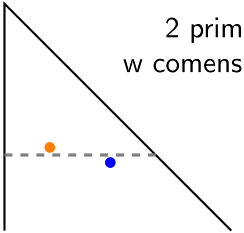
Imposing a k_t cut allows one to stay perturbative

Application series #1: angular correlations

Crafted observables: example $\Delta\Psi_{12}$

Azimuth between 1st and 2nd prim. declust.





2 primaries
w comensurate k_t

Selection

select the 2 emissions with the largest k_t

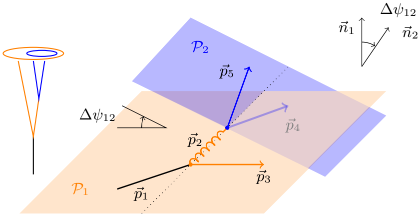
$$-0.6 < \alpha_s \ln \frac{k_{t1}}{Q} < -0.5,\ 0.3 < \frac{k_{t2}}{k_{t1}} < 0.5,\ \alpha_s \to 0$$

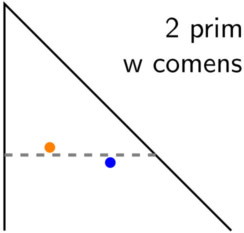
QCD expectation

$$\Sigma_{\mathrm{NLL}}(\Delta\Psi_{12}) = \text{constant}$$

Crafted observables: example $\Delta\Psi_{12}$

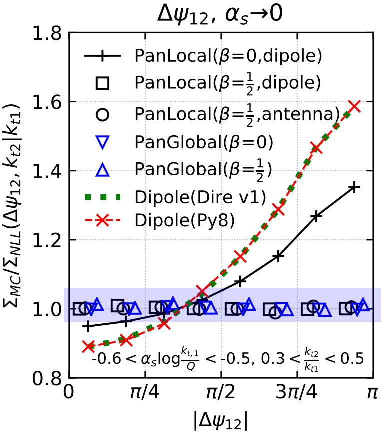
Azimuth between 1st and 2nd prim. declust.





2 primaries
w comensurate k_t

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

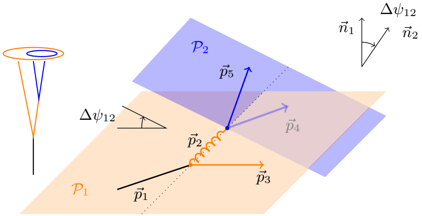


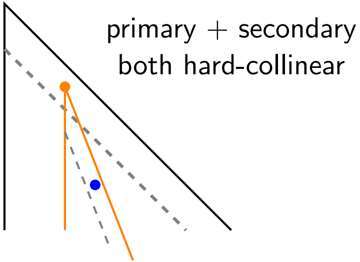
Expected ratio of 1 at NLL

NLL failures for "standard" showers

"New" PanScales shower OK at NLL

Azimuth between 1st and 2nd prim. declust.





Selection

first (primary) emission (k_1) with $z > z_{cut}$
+ first 2^{ndary} emission from k_1 with $z > z_{cut}$

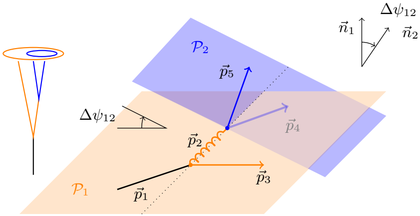
fixed $z_{cut} = 0.1$; $\alpha_s \to 0$

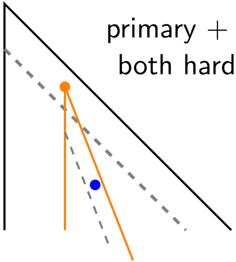
QCD expectation

- some $\Delta\Psi_{12}$ dependence due to (collinear) spin correlations
- analytic expressions available for EEEC (2011.02492)

Crafted observables: example $\Delta\Psi_{12}$

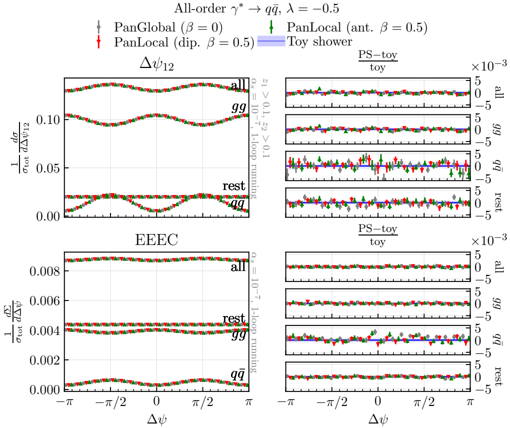
Azimuth between 1[st] and 2[nd] prim. declust.





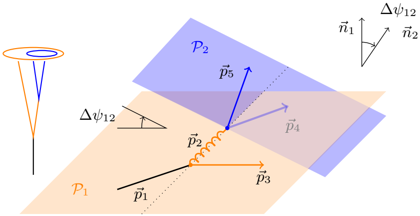
primary + secondary
both hard-collinear

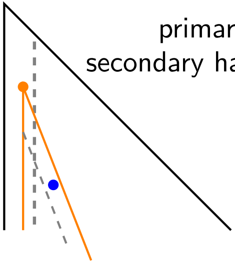
[A.Karlberg,G.Salam,L.Scyboz,R.Verheyen,2103.16526]



clear sensitivity to (collinear) spin
"New" PanScales shower have spin at NLL
EEEC also OK albeit less sensitive

Azimuth between 1st and 2nd prim. declust.





Selection

first (primary) emission (k_1) with $|\eta| < \eta_{\text{cut}}$

+ first 2ndary emission from k_1 with $z > z_{\text{cut}}$

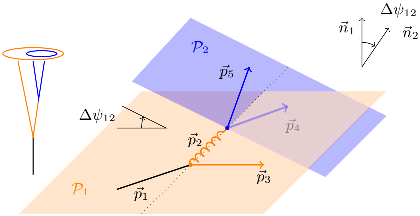
fixed $\eta_{\text{cut}} = 1$; $z_{\text{cut}} = 0.1$; $\alpha_s \to 0$

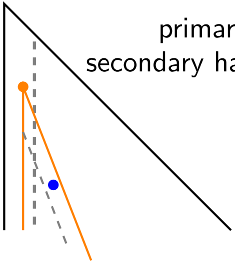
QCD expectation

- some $\Delta\Psi_{12}$ dependence due to (soft) spin correlations
- no (all-order) analytic expressions known

Crafted observables: example $\Delta\Psi_{12}$

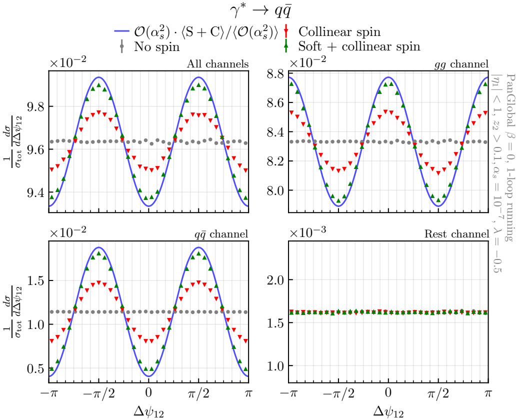
Azimuth between 1st and 2nd prim. declust.





primary soft
secondary hard-collinear

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



Sensitive to (soft) spin

"New" PanScales shower have spin at NLL

shower gives first NLL all-order result

Quark/gluon discrimination

Quark/gluon discrimination

Note: not totally trivial to define what is a "quark jet" or a "gluon jet".
Let us say that we work at small jet radius R so that we can at least focus on "universal" effects
i.e. aspects depending on the overall process are suppressed as R^2.
One can then test e.g. quark/gluon jets in Z+jet v. dijets. (see also arXiv:1704.03878)

Quark v. gluon jets, part I: approaches

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, part I: approaches

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, part I: approaches

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 $\mathcal{L}_{\text{tree}}$

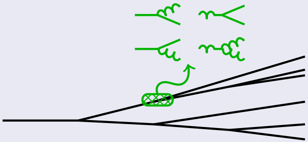
Approach #2

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

- Only splittings with $k_t \geq k_{t,\text{cut}}$ to stay perturbative
- Resum logs to all orders in α_s, up to single logs
  - single logs from "DGLAP" collinear splittings

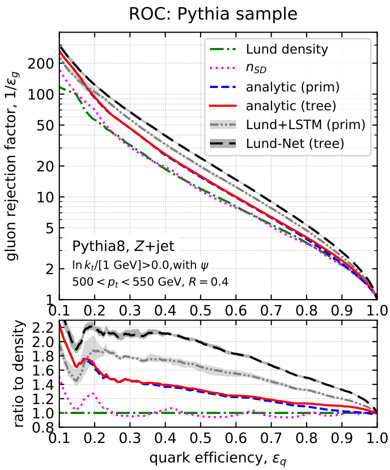
$$P_{i=q,g}(\mathcal{L}_{\text{parent}}) = S_i(\Delta_{\text{prev}}, \Delta) \sum_{j,k=q,g} \tilde{P}_{i\to jk}(z) p_j(\mathcal{L}_{\text{hard}}) p_k(\mathcal{L}_{\text{soft}})$$

  - some single logs for emissions at commensurate angles
    (correct only for (any number of) pairs of commensurate-angle emissions via a matrix-element correction)
- At double-log: $\frac{p_g}{p_q} = \left(\frac{C_A}{C_F}\right)^{n_{\text{prim}}} \Rightarrow$ reproduces the Iterated SoftDrop multiplicity n_{prim}



Quark v. gluon jets, part II: performance

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

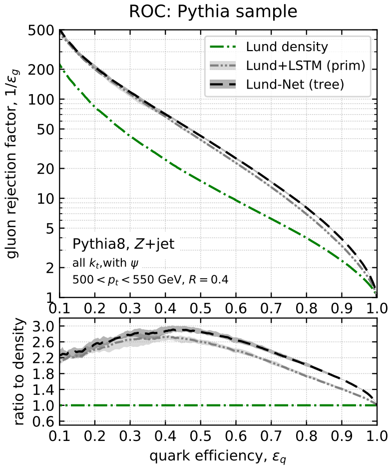


- $k_t > 1$ GeV, clear performance ordering:
  1. Lund+ML > Lund analytic > ISD
  2. tree > prim

several potential effects "learned" by network: subleading, large R, fixed order, > 2 commensurate angles, non-pert, MPI, ...

Quark v. gluon jets, part II: performance

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



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

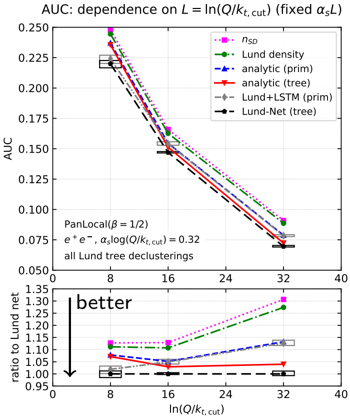
  several potential effects "learned" by network: subleading, large R, fixed order, > 2 commensurate angles, non-pert, MPI, ...

- larger gains with no k_t cut

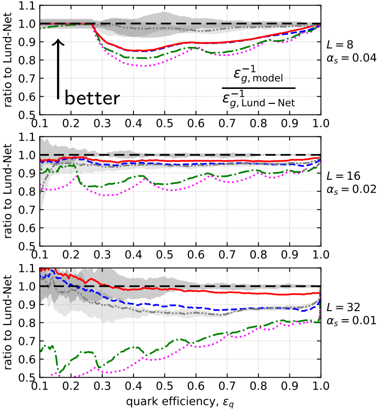
  Suggests that there is quite a lot of differences between quarks and gluon in the NP region ("learned" by the network)

Quark v. gluon jets, part III: towards asymptotics

Ares Under Curve:
lower is better



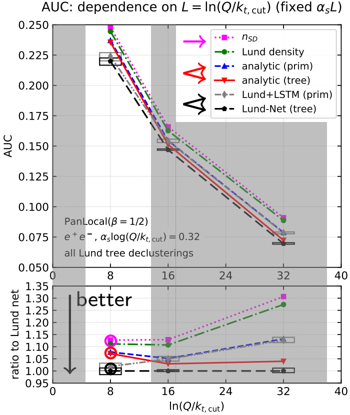
gluon rejection:
higher is better



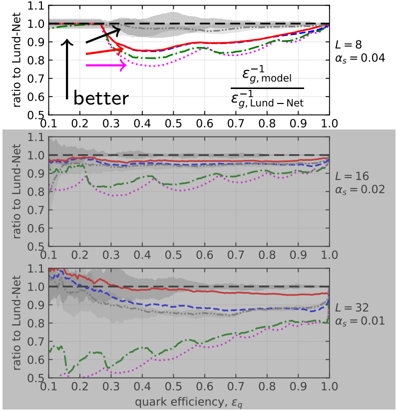
Idea

Asymptotics towards NLL
$\alpha_s L = \text{cst},\ \alpha_s \to 0\ (L \to \infty)$

Ares Under Curve:
lower is better



gluon rejection:
higher is better



Idea

Asymptotics towards NLL

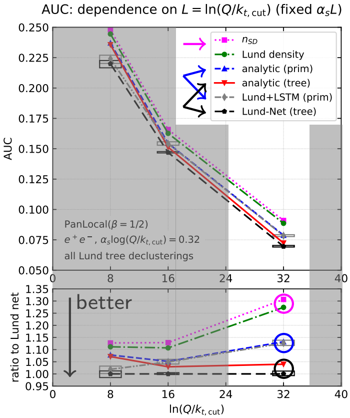
$\alpha_s L = \mathrm{cst}$, $\alpha_s \to 0$ ($L \to \infty$)

Larger α_s (lower L)

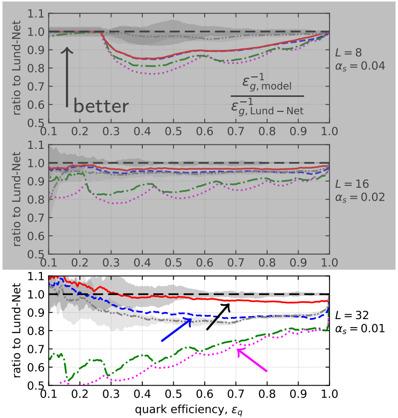
ML > analytics > n_{SD}

little help beyond primary

Ares Under Curve:
lower is better



gluon rejection:
higher is better



Idea

Asymptotics towards NLL
$\alpha_s L = \text{cst}$, $\alpha_s \to 0$ ($L \to \infty$)

Larger α_s (lower L)

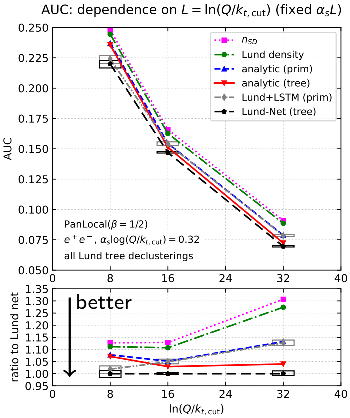
ML > analytics > n_{SD}
little help beyond primary

Larger α_s (lower L)

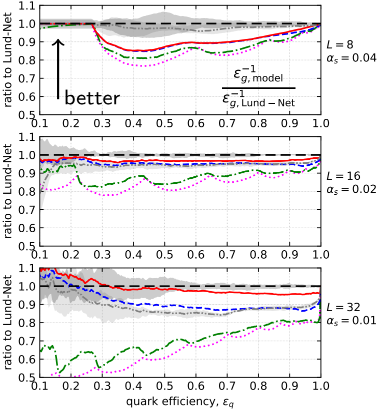
tree > primary > n_{SD}
ML ≈ analytics

Quark v. gluon jets, part III: towards asymptotics

Ares Under Curve:
lower is better



gluon rejection:
higher is better



Idea

Asymptotics towards NLL
$\alpha_s L = $ cst, $\alpha_s \to 0$ ($L \to \infty$)

Larger α_s (lower L)

ML > analytics > n_{SD}
little help beyond primary

Larger α_s (lower L)

tree > primary > n_{SD}
ML $\approx$ analytics

develop accurate
parton-showers for ML

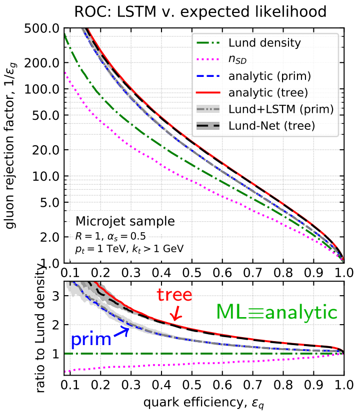
Quark v. gluon jets, part IV: ML validation

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

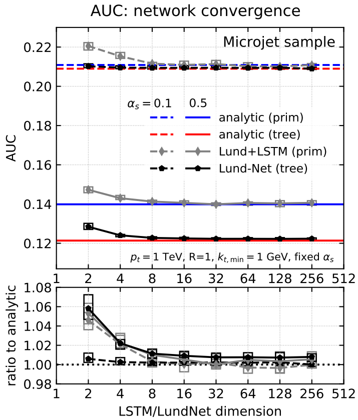
Quark v. gluon jets, part IV: ML validation

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



Converges for large-enough networks

Non-global logarithms

Recent progress

next-to-single-log resummations available!

[Banfi,Dreyer,Monni,2104.06416,2111.02413]

[Becher, Rauh, Xu, 2112.02108] + Thomas' talk

- Improved accuracy on a delicate part of resummations
- Should provide an extra bone to chew on for parton-shower developments

Other things worth noticing:

- Beyond leading-N_c Hatta,Ueda,1304.6930
- Inclusion of heavy quarks [Balsiger, Becher, Ferroglia, 2006.00014]

NGLs and substructure

applying grooming techniques (mMDT/SoftDrop) largely removes NGLs

still left with non-trivial clustering effects at some point

Conclusions

Basic "substructure" facts

1. Substructure tools are now well establishes
2. Many interesting techniques based on reconstructing Lund diagrams/planes
   mimics angular ordering, separate different physical effects (e.g. $k_{t,\mathrm{cut}}$ reduces NP effects)

Conclusions

Basic "substructure" facts

1. Substructure tools are now well establishes
2. Many interesting techniques based on reconstructing Lund diagrams/planes
   mimics angular ordering, separate different physical effects (e.g. $k_{t,\mathrm{cut}}$ reduces NP effects)

Take-home message from this talk: connections between substructure and PSR

PSR → substructure

- parton showers have helped designing many substructure tools
- boosted jets ⇒ resummations

substructure → PSR

can design substructure variables sensitive to specific parton-shower/resummation effects

⇒ connected to several recent parton-shower developments
⇒ connected to several QCD measurements at the LHC

Conclusions

Basic "substructure" facts

1. Substructure tools are now well establishes
2. Many interesting techniques based on reconstructing Lund diagrams/planes
   mimics angular ordering, separate different physical effects (e.g. $k_{t,\mathrm{cut}}$ reduces NP effects)

Take-home message from this talk: connections between substructure and PSR

PSR → substructure

- parton showers have helped designing many substructure tools
- boosted jets ⇒ resummations

substructure → PSR

can design substructure variables sensitive to specific parton-shower/resummation effects

⇒ connected to several recent parton-shower developments

⇒ connected to several QCD measurements at the LHC

+ connections with machine-learning approaches from pQCD

Conclusions

Basic "substructure" facts

1. Substructure tools are now well establishes
2. Many interesting techniques based on reconstructing Lund diagrams/planes
   mimics angular ordering, separate different physical effects (e.g. $k_{t,\mathrm{cut}}$ reduces NP effects)

Take-home message from this talk: connections between substructure and PSR

PSR → substructure

- parton showers have helped designing many substructure tools
- boosted jets ⇒ resummations

substructure → PSR

can design substructure variables sensitive to specific parton-shower/resummation effects

⇒ connected to several recent parton-shower developments
⇒ connected to several QCD measurements at the LHC

+ connections with machine-learning approaches from pQCD

Showcases the need to develop PS&R and the role of substructure

Backup

Construct the Lund tree in practice: use the Cambridge(/Aachen) algorithm

Main idea: Cambridge(/Aachen) preserves angular ordering

e^+e^- collisions

1. Cluster with Cambridge ($d_{ij} = 2(1-\cos\theta_{ij})$)
2. For each (de)-clustering $j \leftarrow j_1 j_2$:

$$\eta = -\ln\theta_{12}/2$$
$$k_t = \min(E_1, E_2)\sin\theta_{12}$$
$$z = \frac{\min(E_1,E_2)}{E_1+E_2}$$
$$\psi \equiv \text{some azimuth},\ldots$$

Jet in pp

1. Cluster with Cambridge/Aachen ($d_{ij} = \Delta R_{ij}$)
2. For each (de)-clustering $j \leftarrow j_1 j_2$:

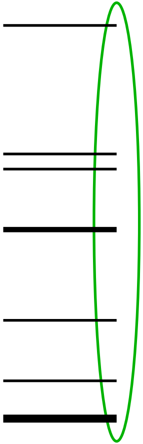
$$\eta = -\ln\Delta R_{12}$$
$$k_t = \min(p_{t1}, p_{t2})\Delta R_{12}$$
$$z = \frac{\min(p_{t1},p_{t2})}{p_{t1}+p_{t2}}$$
$$\psi \equiv \text{some azimuth},\ldots$$

Primary Lund plane

Starting from the jet, de-cluster following the "hard branch" (largest E or p_t)

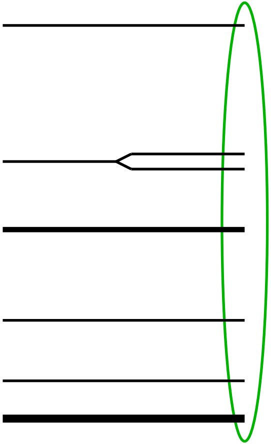
The Lund plane(s) representation: C/A (de)-clustering

use Cambridge/Aachen to iteratively recombine the closest pair



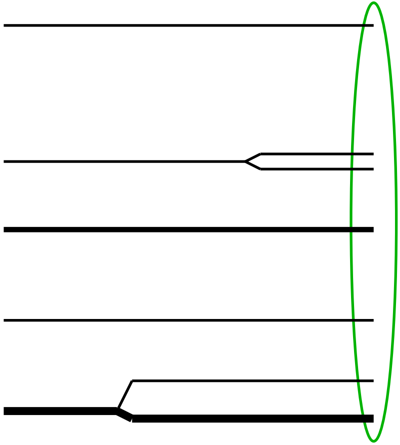
The Lund plane(s) representation: C/A (de)-clustering

use Cambridge/Aachen to iteratively recombine the closest pair



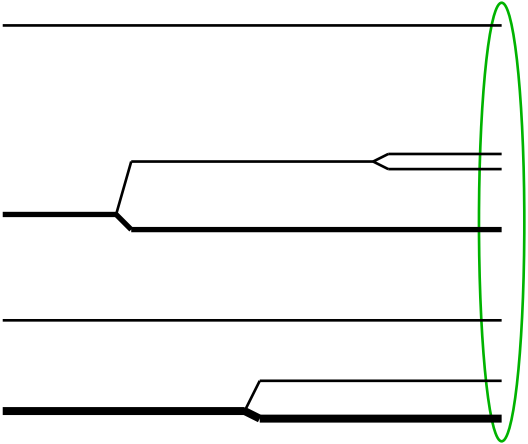
The Lund plane(s) representation: C/A (de)-clustering

use Cambridge/Aachen to iteratively recombine the closest pair



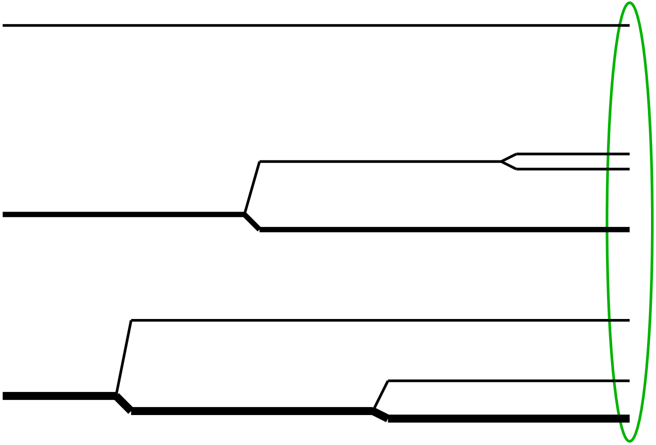
The Lund plane(s) representation: C/A (de)-clustering

use Cambridge/Aachen to iteratively recombine the closest pair



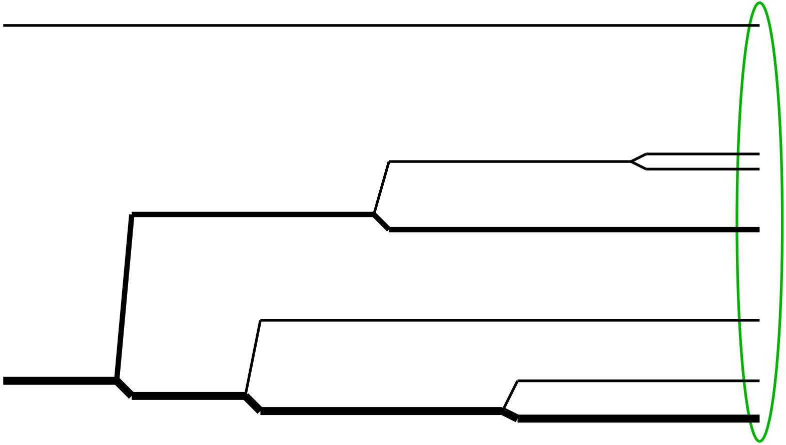
The Lund plane(s) representation: C/A (de)-clustering

use Cambridge/Aachen to iteratively recombine the closest pair

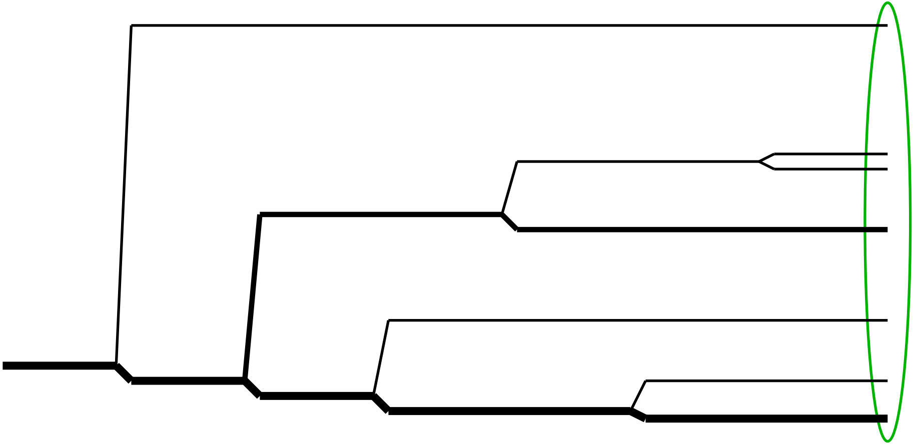


The Lund plane(s) representation: C/A (de)-clustering

use Cambridge/Aachen to iteratively recombine the closest pair

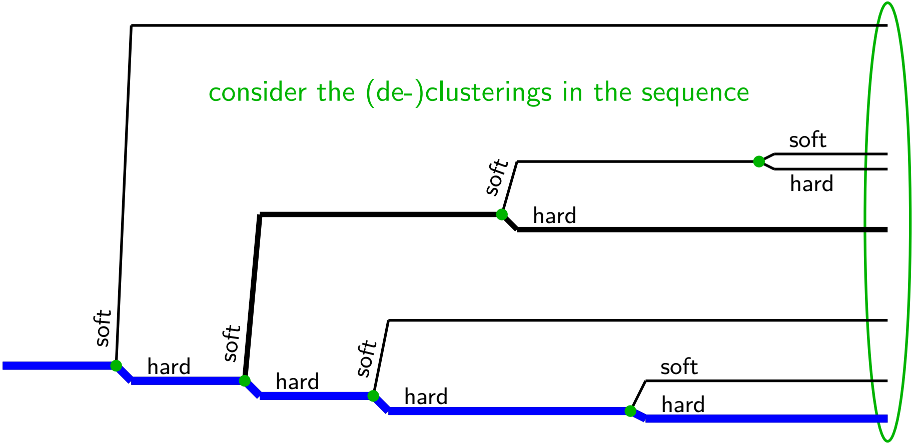


use Cambridge/Aachen to iteratively recombine the closest pair

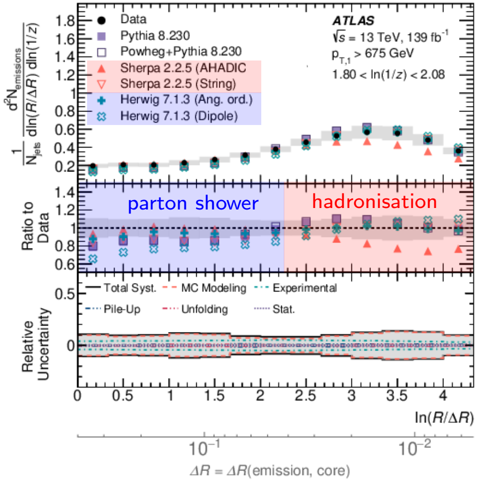


The Lund plane(s) representation: C/A (de)-clustering

use Cambridge/Aachen to iteratively recombine the closest pair

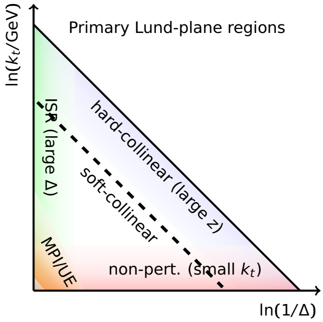


Note: conceptually the largest-energy (p_t or z) branch $\equiv$ emissions from the "leading parton"

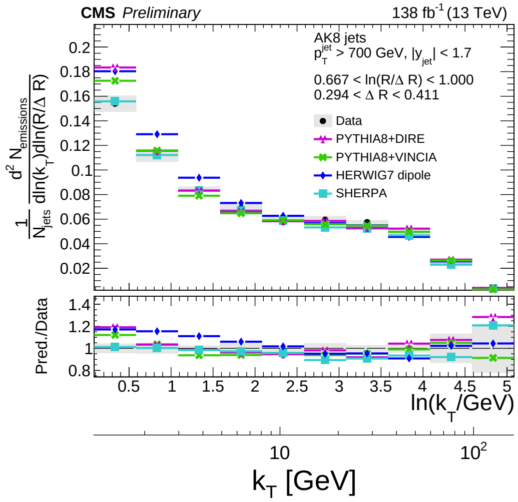


“standard” data vs. Monte Carlo comparison

Recall that different Lund regions are sensitive to different physics:



Obvious comparisons MC vs. data (2/2)



Large spread between Monte Carlo generators also observed by CMS

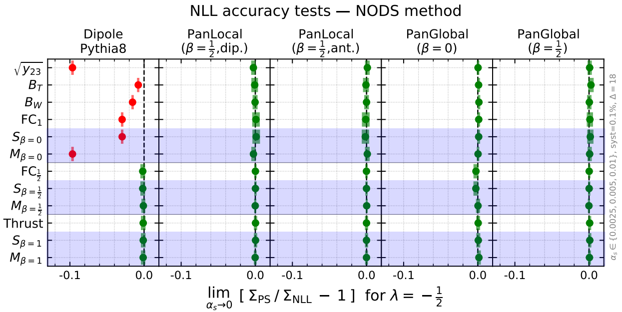
see CMS-PAS-SMP-22-007 for additional comparisons (scales, tunes, ...)

- Equivalent to angularities/EECs:

$$S_\beta = \sum_{i\in\mathcal{L}} E_i\, e^{-\beta\eta_i}$$

$$M_\beta = \max_{i\in\mathcal{L}} E_i\, e^{-\beta\eta_i}$$

  - ✓ subjets allows for the use of "max"
  - ✓ sum≠max at NLL
  - ✓ can be defined in *pp*



[M.Dasgupta,F.Dreyer,K.Hamilton,P.Monni,G.Salam,GS,2002.11114]
[K.Hamilton,R.Medves,G.Salam,L.Scyboz,GS,2011.10054]

- N-subjettiness-like: sum excluding the N largest

$$\tau_N^{\beta,\text{Lund}} = \sum_{i\in A_N} E_i\, e^{-\beta\eta_i} \qquad \text{with} \qquad A_N = \text{argmin}_{X\subset\mathcal{L},|\mathcal{L}\setminus X|=N-1}$$

  - ✓ Could replace sum by max (likely gaining a simpler resummation structure)
  - ✓ Could be defined on the primary plane only

Interesting Lund-plane(s)-based observables

Many Lund-based observables potentially interesting/measurable at the LHC

Lund densities

- already proven useful
- potential extensions (e.g. multiplicities)
- heavy quarks (e.g. b jets)
  dead cone is a relatively small phase-space, but $b \sim$ light over large region
- other processes? $Z + j$? top quarks?

$\Delta\Psi_{12}$

Sensitivity to log accuracy and spin correlations

More generally: probes correlations between 2 emissions

expect subleading effects (compared to above asymptotic studies)

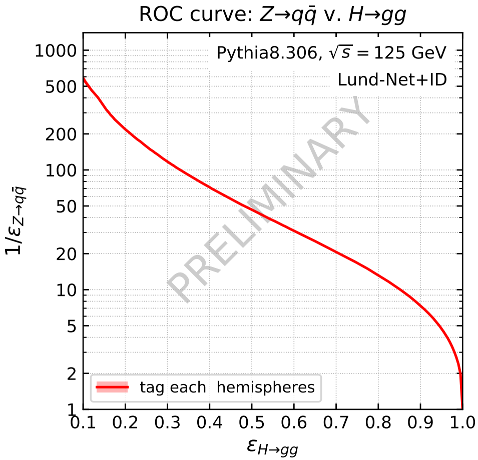
Others?

Large flexibility to

- (re-)interpret existing tools
  (grooming, angularities, N-subjettiness, ...)
- design taylored observables
  (measurements, MC constraints, heavy ions, ...)

Towards full-event tagging

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

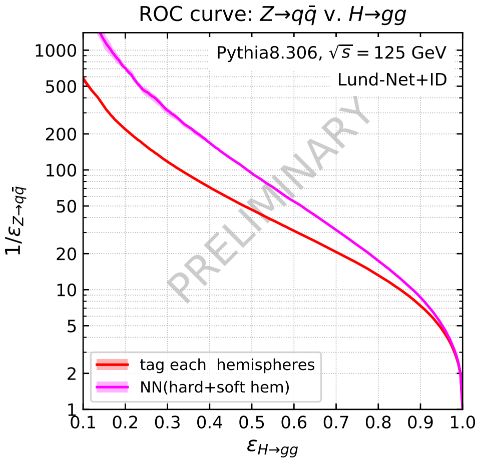


observed performance:

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

  full event clearly worse that (jet)^2

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



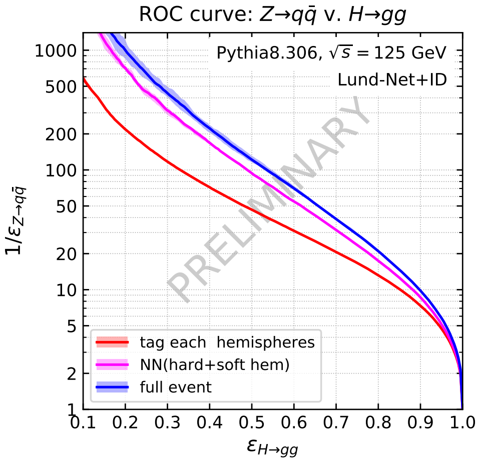
observed performance:

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

  clear performance gain

Towards full-event tagging

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

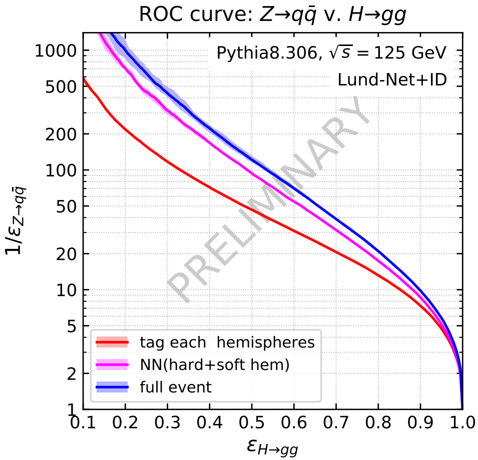


observed performance:

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

Towards full-event tagging

$e^+e^- \to Z \to q\bar{q}$ v. $e^+e^- \to H \to gg \qquad (\sqrt{s} = 125$ 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)?