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

Motivation

Understanding high-energy collisions requires a description of physics across a wide range of scales (from $\mathcal{O}(\Lambda_{QCD})$ to $\mathcal{O}(\text{TeV})$)

This talk

1. Lund diagrams as a (historical) conceptual tool for parton showers and resummations
2. promoting to a practical tool for jet physics
3. (Brief) overview of the wide range of applications
4. More extensive description of quark/gluon discrimination

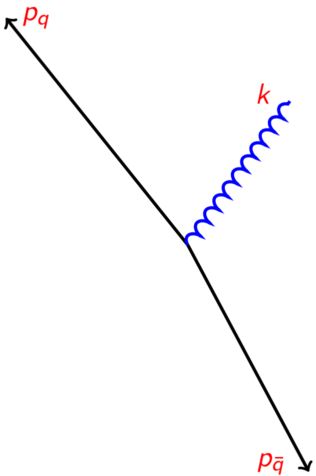
Basic observation

Exploring widely different scales $\leftrightarrow$ exposing the soft and collinear divergences of QCD
Obvious connections with parton shower and resummations

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:

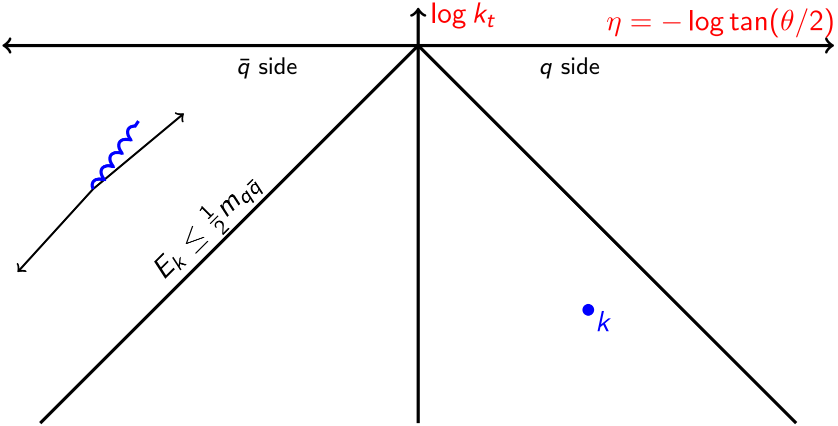
- Rapidity: $\eta = \frac{1}{2}\log\frac{z_q}{z_{\bar{q}}}$
- Transverse momentum: $k_\perp$
- Azimuth ϕ

In the soft-collinear approximation

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

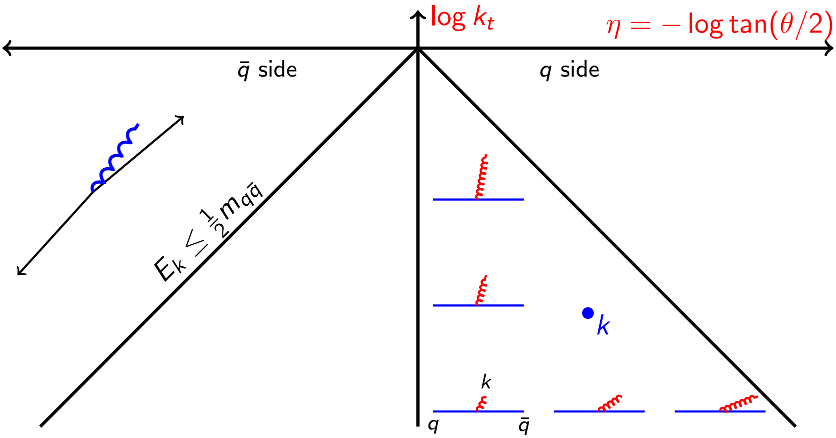
Basic features of QCD radiations: the Lund plane

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

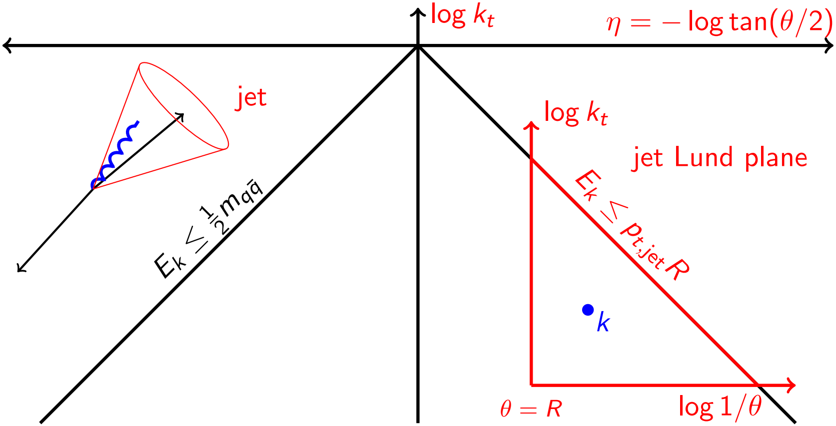


Basic features of QCD radiations: the Lund plane

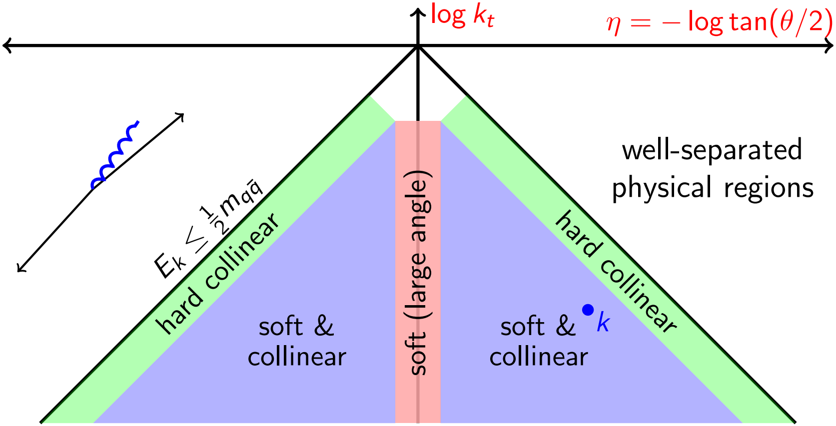
Lund plane: natural representation uses the 2 "log" variables η and $\log k_\perp$



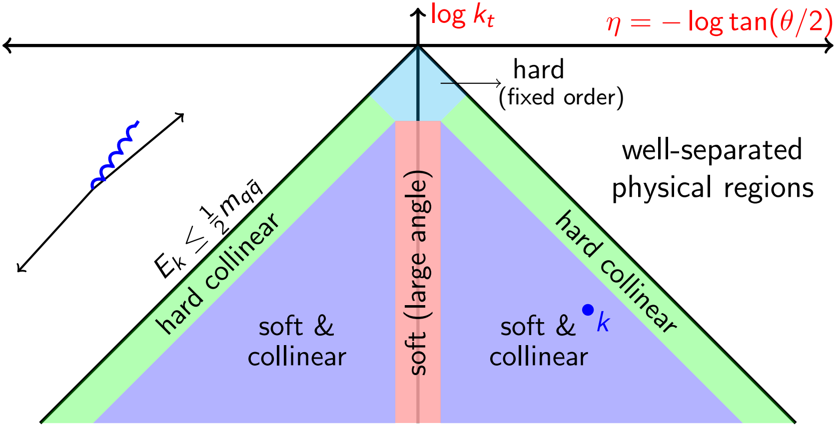
Lund plane: natural representation uses the 2 “log” variables η and $\log k_\perp$



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

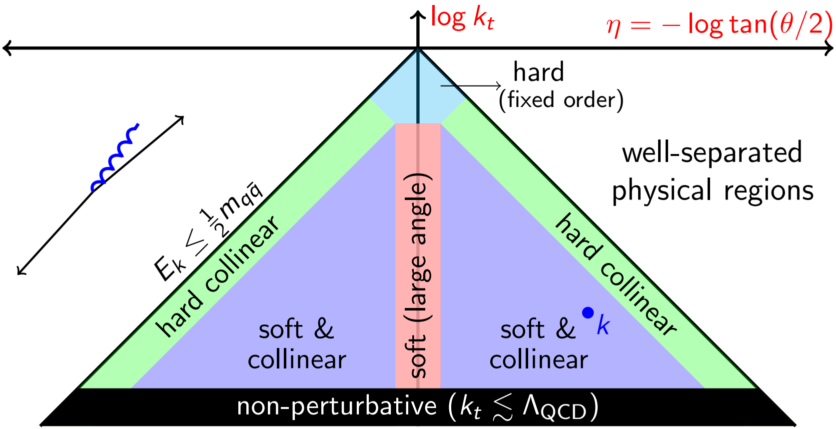


Lund plane: natural representation uses the 2 "log" variables η and $\log k_\perp$

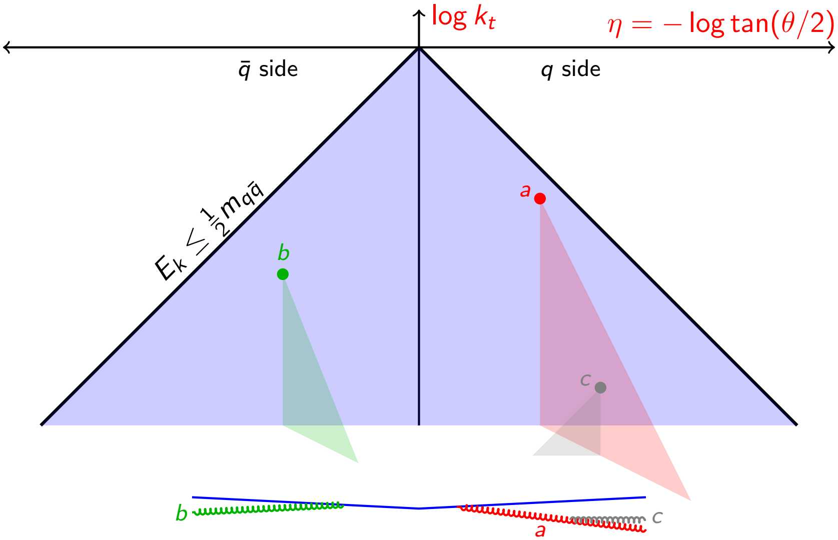


Basic features of QCD radiations: the Lund plane

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



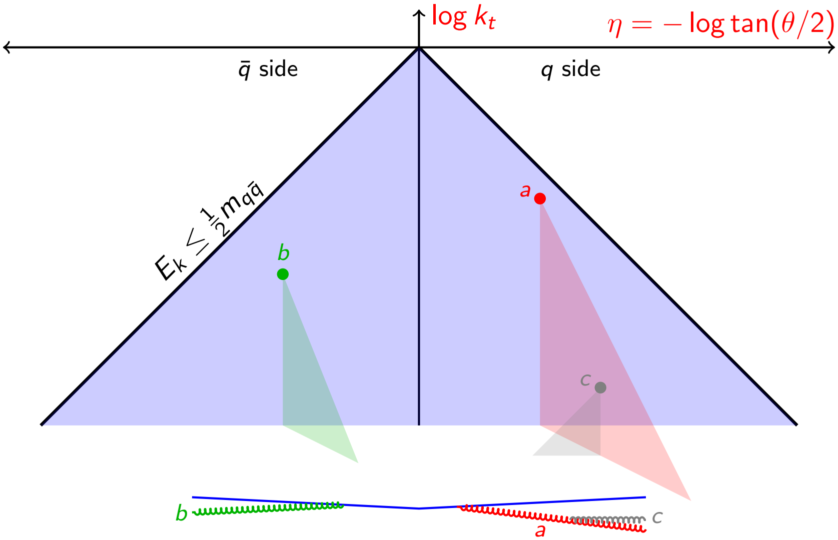
Multiple emissions in the Lund plane



Each emission spawns its own plane

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

Multiple emissions in the Lund plane



Each emission spawns its own plane

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

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

Take-home message #1

Lund diagrams represent (multiple) radiation across scales

Set of nice properties:

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

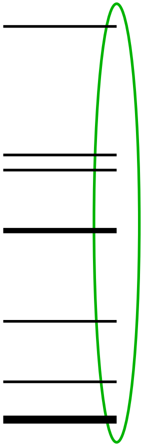
Lund planes:

promoting Lund diagrams to a practical tool

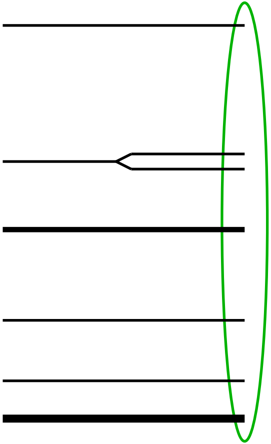
For simplicity, take a high-p_t LHC jet (similar for full e^+e^- events)

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

use Cambridge/Aachen to iteratively recombine the closest pair

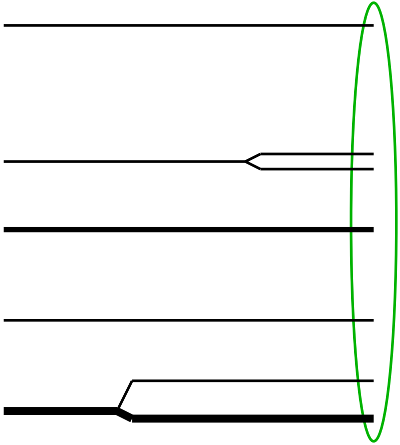


use Cambridge/Aachen to iteratively recombine the closest pair



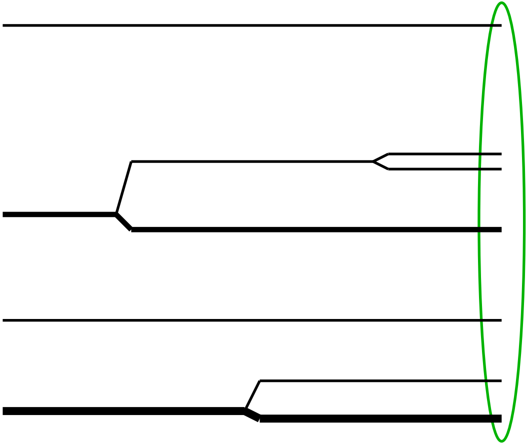
The Lund plane(s) representation (1/3)

use Cambridge/Aachen to iteratively recombine the closest pair

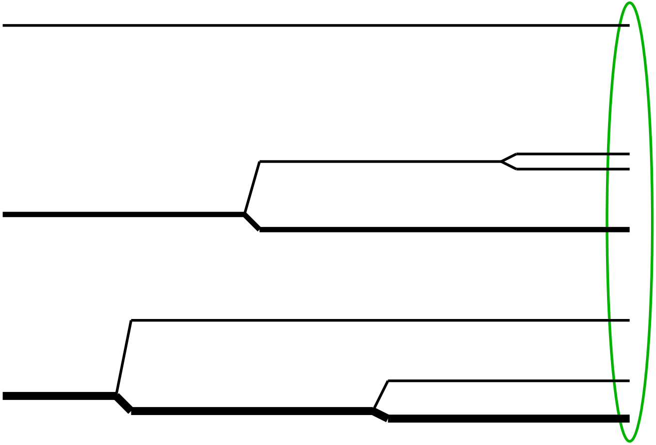


The Lund plane(s) representation (1/3)

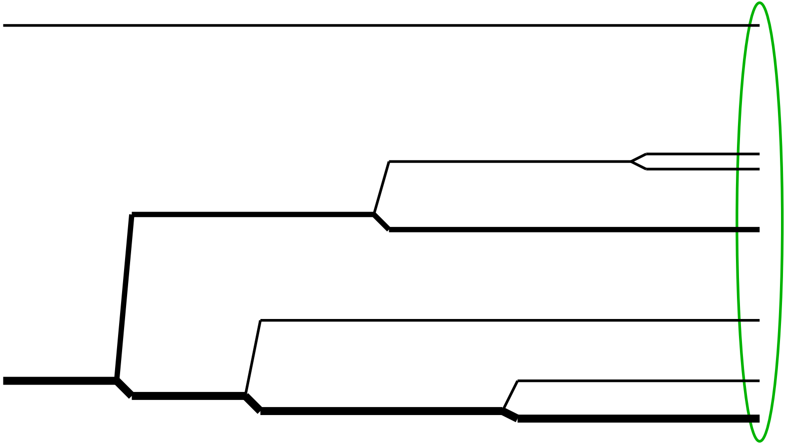
use Cambridge/Aachen to iteratively recombine the closest pair



use Cambridge/Aachen to iteratively recombine the closest pair

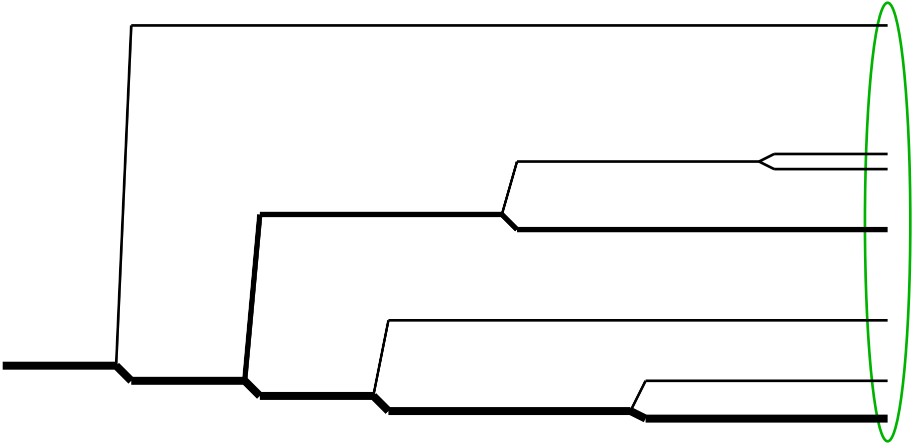


use Cambridge/Aachen to iteratively recombine the closest pair



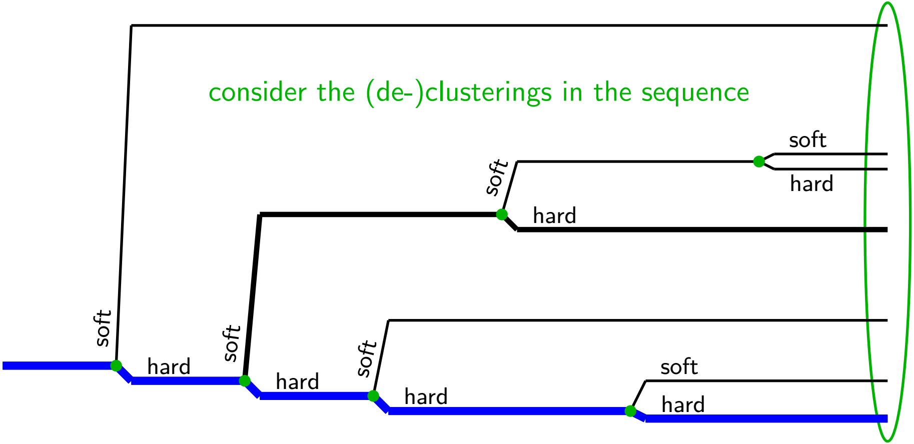
The Lund plane(s) representation (1/3)

use Cambridge/Aachen to iteratively recombine the closest pair



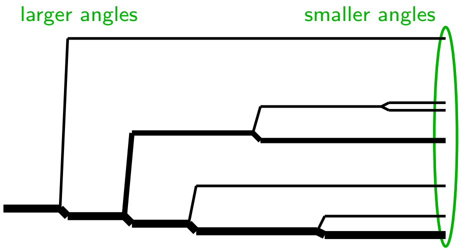
use Cambridge/Aachen to iteratively recombine the closest pair

consider the (de-)clusterings in the sequence



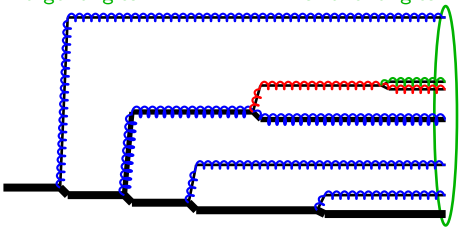
Note: conceptually the largest-energy (p_t or z) branch ≡ emissions from the "leading parton"

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



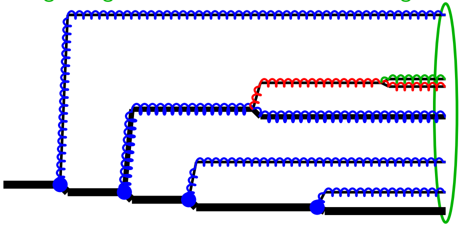
- closely follows our beloved angular ordering

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

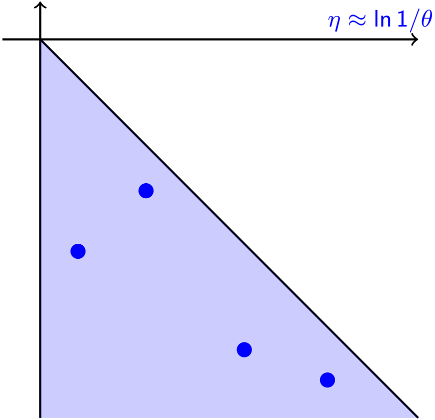


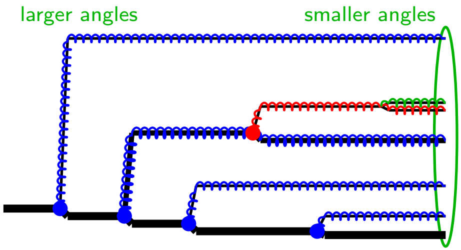
- closely follows our beloved angular ordering
- i.e. mimics partonic cascade

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

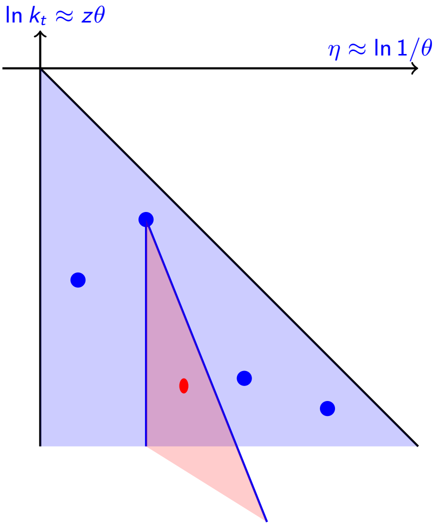


- closely follows our beloved angular ordering
- i.e. mimics partonic cascade
- can be organised in Lund planes
  - primary

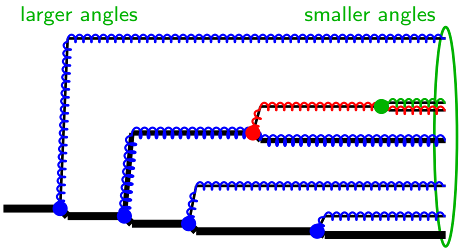




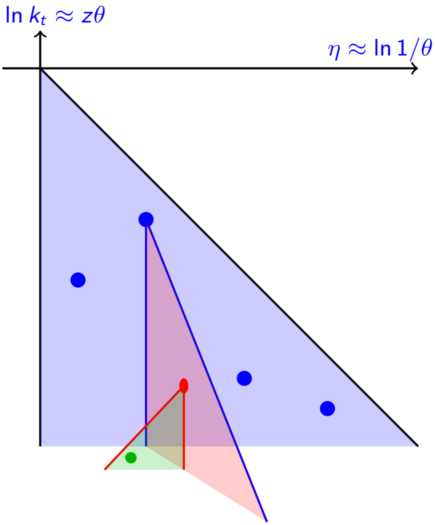
- closely follows our beloved angular ordering
- i.e. mimics partonic cascade
- can be organised in Lund planes
  - primary
  - secondary



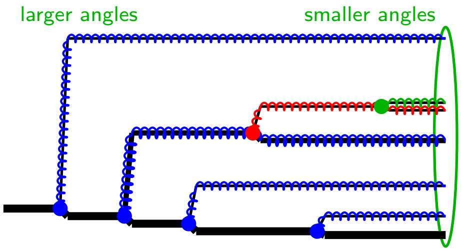
The Lund plane(s) representation (2/3)



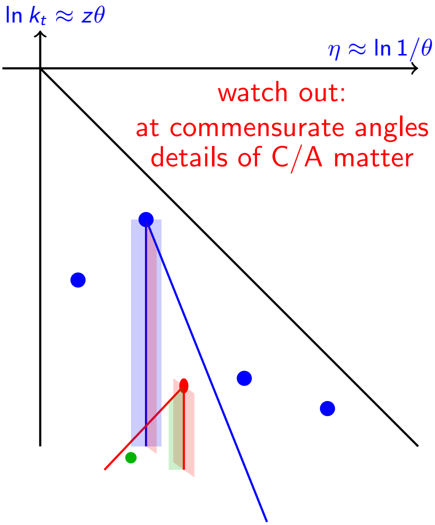
- closely follows our beloved angular ordering
- i.e. mimics partonic cascade
- can be organised in Lund planes
  - primary
  - secondary
  - ...



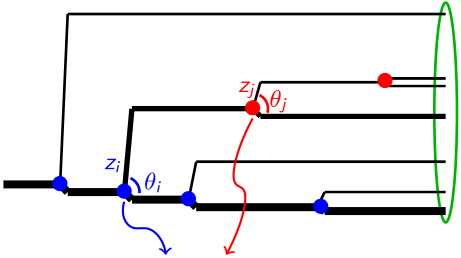
The Lund plane(s) representation (2/3)



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



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



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

for jets in pp: (similar for ee events)

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

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

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

Two different Lund ($\mathcal{L}$) structures

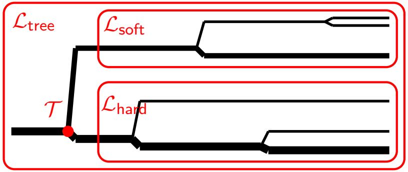
"primary plane" (follow hard branch)

$\mathcal{L}_{\text{prim}} \equiv \{\mathcal{T}_i\}$

OR

full (de-)clustering tree

$\mathcal{L}_{\text{tree}} \equiv \{\mathcal{T}, \mathcal{L}_{\text{hard}}, \mathcal{L}_{\text{soft}}\}$



Note: branchings with $k_t > t_{t,\min} \Rightarrow$ perturbative

Take-home message #2

Central observation

For a jet (or a *ee* event) one can **construct** a structure that captures the properties of Lund diagrams

The rest of this talk covers several applications:

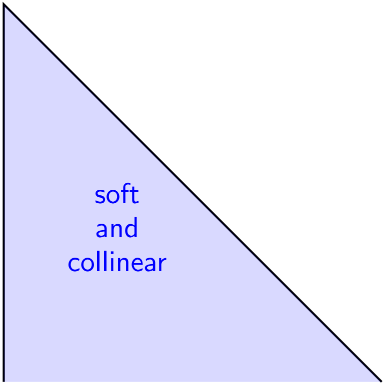
- ✓ Calculations (and measurements)
- ✓ Monte-Carlo developments
- ✓ Tagging (incl. machine learning and **quark/gluon discrimination**)

Application #1: QCD calculations

Average number of emission at given k_t, ΔR:

[A. Lifson, G. Salam, GS, arXiv:2007.06578]

$$\rho = \frac{1}{N_{\text{jets}}} \frac{d^2 N}{d \ln \Delta R \, d \ln k_t}$$



- Double-logarithmic behaviour:

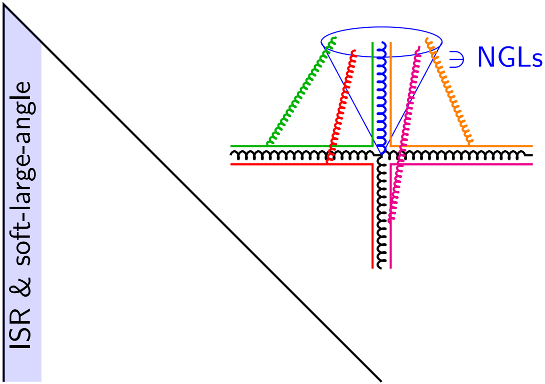
$$\rho = \frac{2\alpha_s(k_t) C_R}{\pi}$$

Primary Lund plane multiplicity

Average number of emission at given k_t, ΔR:

[A. Lifson, G. Salam, GS, arXiv:2007.06578]

$$\rho = \frac{1}{N_{\text{jets}}} \frac{d^2 N}{d \ln \Delta R \, d \ln k_t}$$



- Double-logarithmic behaviour:

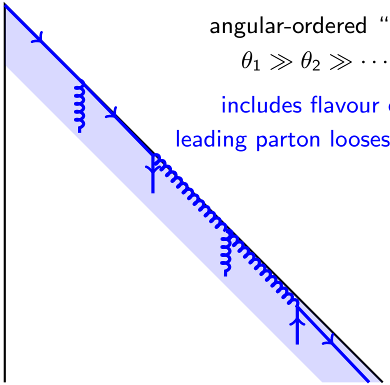
$$\rho = \frac{2\alpha_s(k_t) C_R}{\pi}$$

- Single-log calculation including
  - ✓ Running-coupling (trivial)
  - ✓ ISR+large angle

Average number of emission at given k_t, ΔR:

[A. Lifson, G. Salam, GS, arXiv:2007.06578]

$$\rho = \frac{1}{N_{\text{jets}}} \frac{d^2 N}{d \ln \Delta R \, d \ln k_t}$$



- Double-logarithmic behaviour:

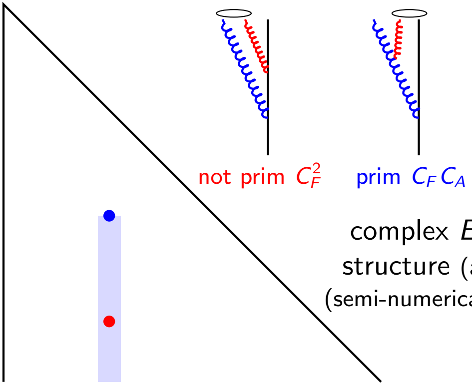
$$\rho = \frac{2\alpha_s(k_t) C_R}{\pi}$$

- Single-log calculation including
  - ✓ Running-coupling (trivial)
  - ✓ ISR+large angle
  - ✓ Hard-collinear branchings

Primary Lund plane multiplicity

Average number of emission at given k_t, ΔR: [A. Lifson, G. Salam, GS, arXiv:2007.06578]

$$\rho = \frac{1}{N_{\text{jets}}} \frac{d^2 N}{d \ln \Delta R \, d \ln k_t}$$



complex E-ordered structure (akin NGLs)
(semi-numerical treatment)

- Double-logarithmic behaviour:

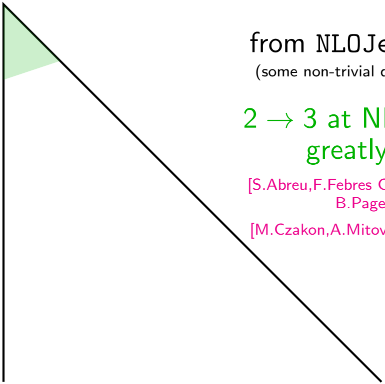
$$\rho = \frac{2\alpha_s(k_t) C_R}{\pi}$$

- Single-log calculation including
  - ✓ Running-coupling (trivial)
  - ✓ ISR+large angle
  - ✓ Hard-collinear branchings
  - ✓ Clustering effects

Primary Lund plane multiplicity

Average number of emission at given k_t, ΔR: [A. Lifson, G. Salam, GS, arXiv:2007.06578]

$$\rho = \frac{1}{N_{\text{jets}}} \frac{d^2 N}{d \ln \Delta R \; d \ln k_t}$$



from NLOJet++
(some non-trivial details)

$2 \to 3$ at NNLO would greatly help!

[S.Abreu,F.Febres Cordero,H.Ita, B.Page,V.Sotnikov,2102.13609]

[M.Czakon,A.Mitov,R.Poncelet,2106.05331]

- Double-logarithmic behaviour:

$$\rho = \frac{2\alpha_s(k_t) C_R}{\pi}$$

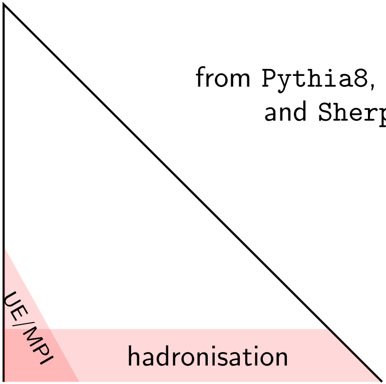
- Single-log calculation including
  - ✓ Running-coupling (trivial)
  - ✓ ISR+large angle
  - ✓ Hard-collinear branchings
  - ✓ Clustering effects
- + Matching to NLO ($\sim$ top)

Primary Lund plane multiplicity

Average number of emission at given k_t, ΔR:

[A. Lifson, G. Salam, GS, arXiv:2007.06578]

$$\rho = \frac{1}{N_{\text{jets}}} \frac{d^2 N}{d \ln \Delta R \; d \ln k_t}$$

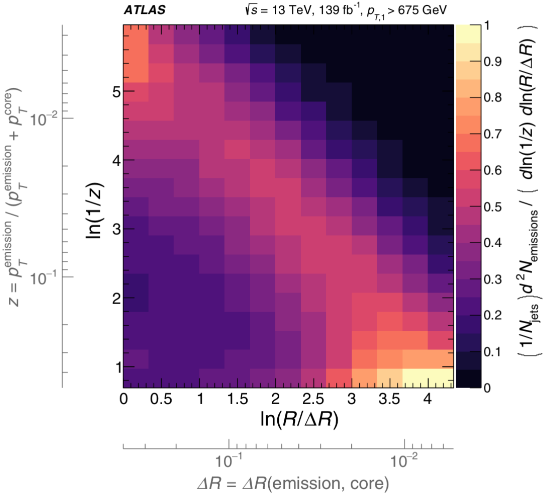


from `Pythia8`, `Herwig7` and `Sherpa2`

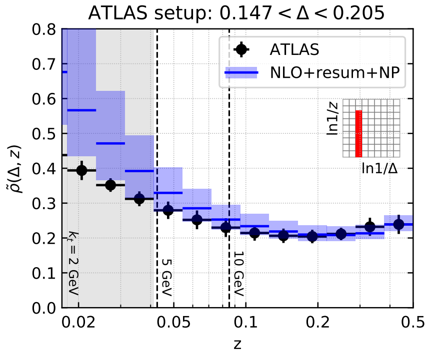
- Double-logarithmic behaviour:

$$\rho = \frac{2\alpha_s(k_t) C_R}{\pi}$$

- Single-log calculation including
  - ✓ Running-coupling (trivial)
  - ✓ ISR+large angle
  - ✓ Hard-collinear branchings
  - ✓ Clustering effects
- + Matching to NLO ($\sim$ top)
- + NP corrections ($\sim$ bottom)



[ATLAS, 2004.03540]



- good agreement (particularly for $k_t \gtrsim 5$ GeV)
- commensurate exp.&th. uncert.
- Can we get α_s from this?

Lund multiplicity (1/2)

Lund multiplicity

count the (average) number of Lund declusterings (in the full tree) with $k_t \geq k_{t,\text{cut}}$

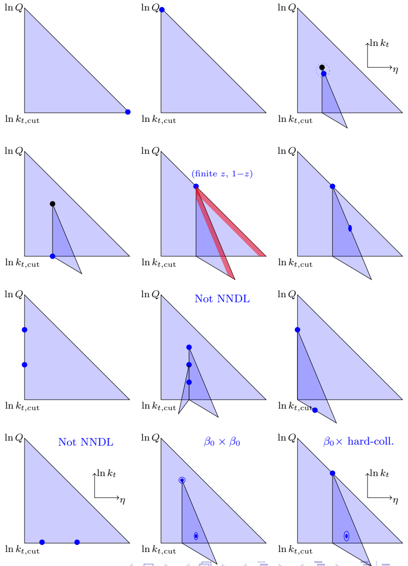
All-order structure ($L = \ln \frac{Q}{k_{t,\text{cut}}}$):

$$\langle N^{\text{LP}}(L, \alpha_s)\rangle = \underbrace{h_1(\alpha_s L^2) + \sqrt{\alpha_s} h_2(\alpha_s L^2)}_{\text{Since 1992}} + \underbrace{\alpha_s h_3(\alpha_s L^2)}_{\text{New NNDL!!}} + \dots$$

[R. Medves, A. Soto, GS, 2205.02861]

$$2\pi h_3^{(q)} = D_{\text{end}}^{q\to qg} + \left(D_{\text{end}}^{g\to gg} + D_{\text{end}}^{g\to q\bar{q}}\right)\frac{C_F}{C_A}(\cosh\nu - 1) + D_{\text{hme}}^{qqg}\cosh\nu + \frac{C_F}{C_A}\left[(1-c_\delta)D_{\text{pair}}^{q\bar{q}}(\cosh\nu - 1) + \left(K + D_{\text{pair}}^{gg} + c_\delta D_{\text{pair}}^{q\bar{q}}\right)\frac{\nu}{2}\sinh\nu\right]$$
$$+ C_F\left[\left(\cosh\nu - 1 - \frac{1-c_\delta}{4}\nu^2\right)D_{\text{clust}}^{(\text{prim})} + (\cosh\nu - 1)D_{\text{clust}}^{(\text{sec})}\right] + \frac{C_F}{C_A}\left[D_{\text{e-loss}}^{g}\frac{\nu}{2}\sinh\nu + \left(D_{\text{e-loss}}^{q} - D_{\text{e-loss}}^{g}\right)(\cosh\nu - 1)\right]$$
$$+ \frac{C_F}{C_A}\frac{\pi^2\beta_0^2}{8C_A}\left[3\nu(2\nu^2-1)\sinh\nu + (\nu^4 + 3\nu^2)\cosh\nu\right] + \frac{C_F}{2}\Big\{(B_{gg} + c_\delta B_{gq})^2\nu^2\cosh\nu + 8\left[2c_\delta B_{gg} - 2c_\delta B_q - (1-3c_\delta^2)B_{gq}\right]B_{gq}\cosh\nu$$
$$+ \left[4B_q(B_{gg} + (2c_\delta+1)B_{gq}) - (B_{gg} + c_\delta B_{gq})(B_{gg} + 9c_\delta B_{gq})\right]\nu\sinh\nu + 4(1-c_\delta^2)B_{gq}^2\nu^2 + 8\left[2c_\delta B_q - 2c_\delta B_{gg} + (1-3c_\delta^2)B_{gq}\right]B_{gq}\Big\}$$
$$+ \frac{C_F}{C_A}\frac{\pi\beta_0}{2}\Big\{(B_{gg} + c_\delta B_{gq})\nu^3\sinh\nu + \left[2B_q - 2B_{gg} + (6-8c_\delta)B_{gq}\right]\nu\sinh\nu + 2(B_q + B_{gg} + B_{gq})\nu^2\cosh\nu - 4(1-c_\delta)B_{gq}(2\cosh\nu - 2 + \nu^2)\Big\}$$

Side product: NNDL Cambridge multiplicity for $y_{\text{cut}} = k_{t,\text{cut}}^2$



Lund multiplicity (1/2)

Lund multiplicity

count the (average) number of Lund declusterings (in the full tree) with $k_t \geq k_{t,\text{cut}}$

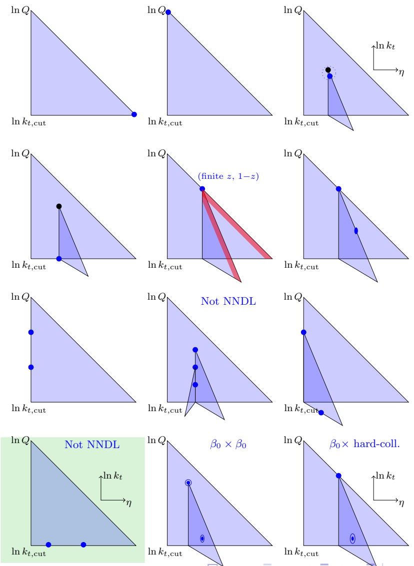
All-order structure ($L = \ln \frac{Q}{k_{t,\text{cut}}}$):

$$\langle N^{\text{LP}}(L,\alpha_s)\rangle = \underbrace{h_1(\alpha_s L^2) + \sqrt{\alpha_s}\, h_2(\alpha_s L^2)}_{\text{Since 1992}} + \underbrace{\alpha_s h_3(\alpha_s L^2)}_{\text{New NNDL!!}} + \dots$$

[R. Medves, A. Soto, GS, 2205.02861]

$$\begin{aligned}
2\pi h_3^{(q)} = {} & D_{\text{end}}^{q\to qg} + \left(D_{\text{end}}^{g\to gg} + D_{\text{end}}^{g\to q\bar q}\right)\frac{C_F}{C_A}(\cosh\nu - 1) + D_{\text{hme}}^{qqg}\cosh\nu + \frac{C_F}{C_A}\left[(1-c_\delta)D_{\text{pair}}^{q\bar q}(\cosh\nu-1) + \left(K + D_{\text{pair}}^{gg} + c_\delta D_{\text{pair}}^{q\bar q}\right)\frac{\nu}{2}\sinh\nu\right] \\
& + C_F\left[\left(\cosh\nu - 1 - \frac{1-c_\delta}{4}\nu^2\right)D_{\text{clust}}^{(\text{prim})} + (\cosh\nu-1)D_{\text{clust}}^{(\text{sec})}\right] + \frac{C_F}{C_A}\left[D_{\text{e-loss}}^{g}\frac{\nu}{2}\sinh\nu + \left(D_{\text{e-loss}}^{q} - D_{\text{e-loss}}^{g}\right)(\cosh\nu - 1)\right] \\
& + \frac{C_F}{C_A}\frac{\pi^2\beta_0^2}{8C_A}\left[3\nu(2\nu^2-1)\sinh\nu + (\nu^4+3\nu^2)\cosh\nu\right] + \frac{C_F}{2}\Big\{(B_{gg}+c_\delta B_{gq})^2\nu^2\cosh\nu + 8\left[2c_\delta B_{gg} - 2c_\delta B_q - (1-3c_\delta^2)B_{gq}\right]B_{gq}\cosh\nu \\
& + \left[4B_q(B_{gg}+(2c_\delta+1)B_{gq}) - (B_{gg}+c_\delta B_{gq})(B_{gg}+9c_\delta B_{gq})\right]\nu\sinh\nu + 4(1-c_\delta^2)B_{gq}^2\nu^2 + 8\left[2c_\delta B_q - 2c_\delta B_{gg} + (1-3c_\delta^2)B_{gq}\right]B_{gq}\Big\} \\
& + \frac{C_F}{C_A}\frac{\pi\beta_0}{2}\Big\{(B_{gg}+c_\delta B_{gq})\nu^3\sinh\nu + \left[2B_q - 2B_{gg} + (6-8c_\delta)B_{gq}\right]\nu\sinh\nu + 2(B_q+B_{gg}+B_{gq})\nu^2\cosh\nu - 4(1-c_\delta)B_{gq}(2\cosh\nu - 2 + \nu^2)\Big\}
\end{aligned}$$

No "long-distance effect" $\Rightarrow$ simpler than k_t



Lund multiplicity (2/2)

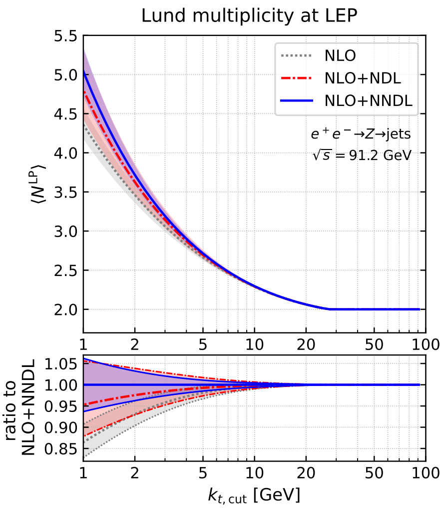
[R. Medves, A. Soto, GS, 2205.02861]

NNDL Matched to NLO

- Clear effect of resummation
- Clear effect compared to NDL (incl. uncert)

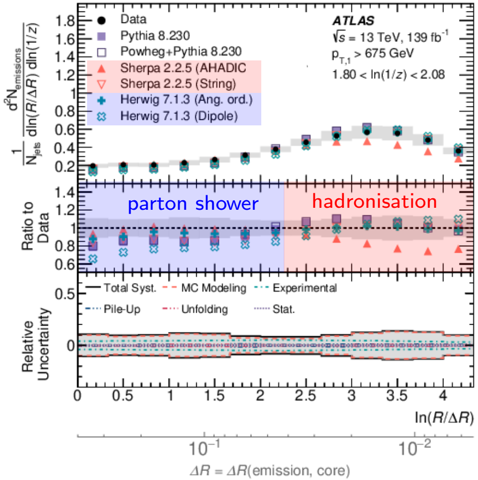
Several questions

- LEP (ALEPH) measurement?
  see. e.g. Y.Chen *et al.* 2111.09914
- Upgrade to LHC jets?
- Can it lead to an α_s measurement?
- NNLO? N^{3}DL?



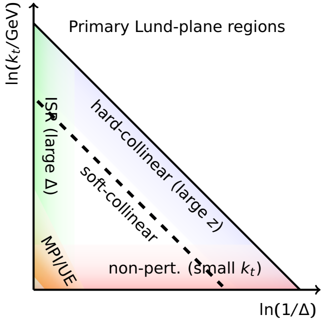
Application #2: MC development

Obvious comparisons



"standard" data vs. Monte Carlo comparison

Recall that different Lund regions are sensitive to different physics:



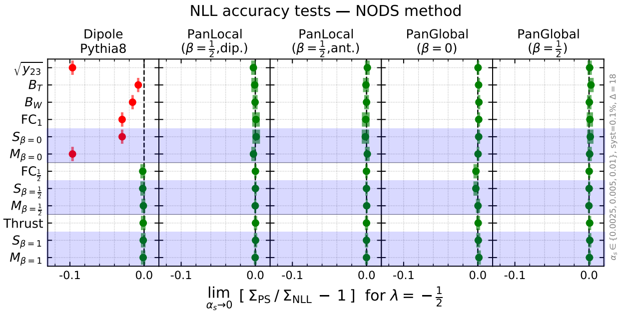
Revisiting "standard" substructure observables [skip]

- 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]

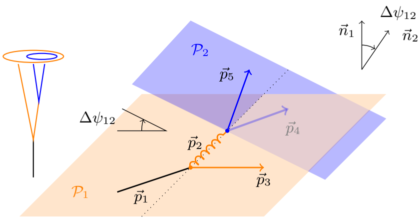
- *N*-subjettiness-like: sum excluding the *N* largest

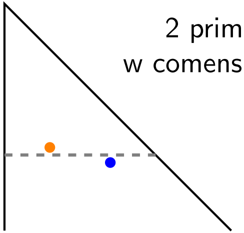
$$\tau_N^{\beta,\text{Lund}} = \sum_{i\in A_N} E_i\, e^{-\beta\eta_i} \qquad \text{with} \qquad A_N = \text{argmin}_{X\subset\mathcal{L},|\mathcal{L}\setminus X|=N-1}$$

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

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

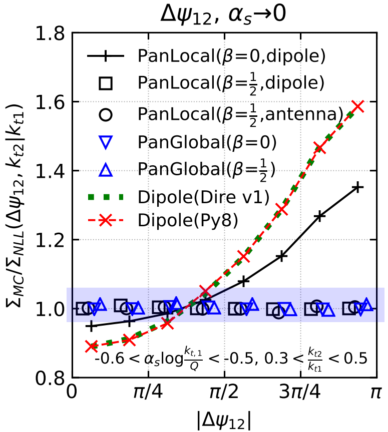
Azimuth between 1st and 2nd prim. declust.





2 primaries
w comensurate k_t

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



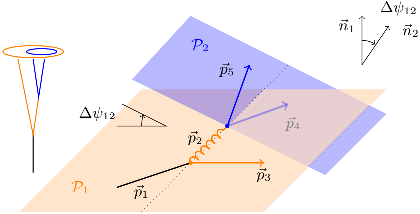
Expected ratio of 1 at NLL

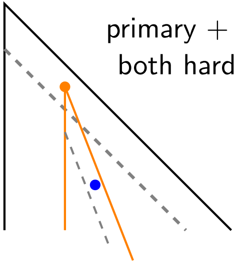
NLL failures for "standard" showers

"New" PanScales shower OK at NLL

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

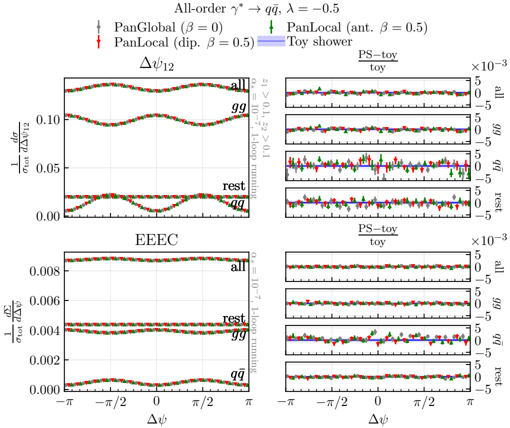
Azimuth between 1st and 2nd prim. declust.





primary + secondary
both hard-collinear

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



Sensitive to (collinear) spin

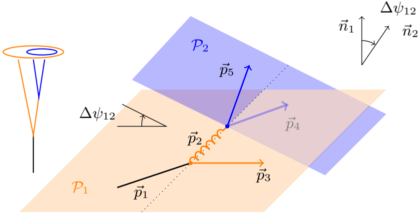
"New" PanScales shower have spin at NLL

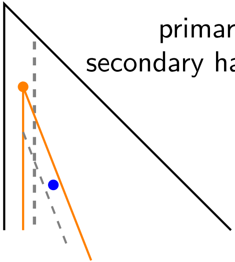
agrees w EEEC from 2011.02492 (EEEC less sensitive)

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

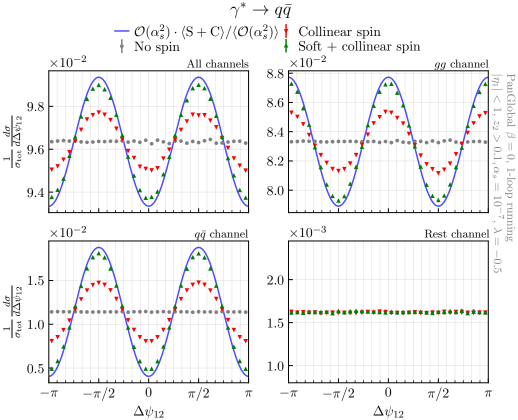
[K.Hamilton,A.Karlberg,G.Salam,L.Scyboz,R.Verheyen,2111.01161]

Azimuth between 1st and 2nd prim. declust.





primary soft
secondary hard-collinear



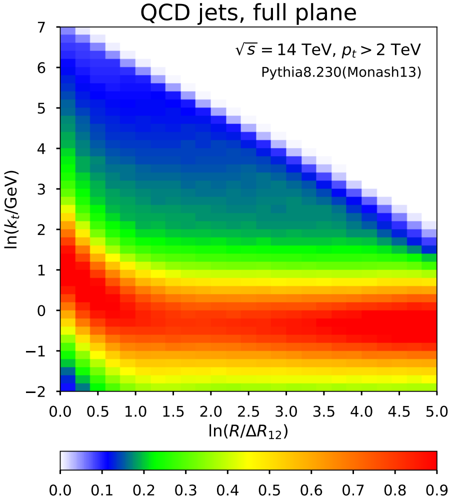
Sensitive to (soft) spin
"New" PanScales shower have spin at NLL
first all-order result

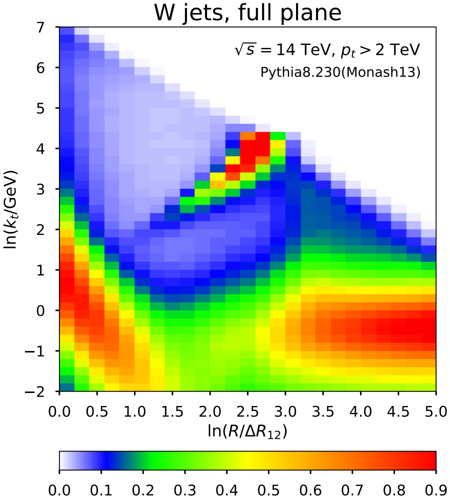
Application #3: Boosted object tagging

(mostly illustrative)

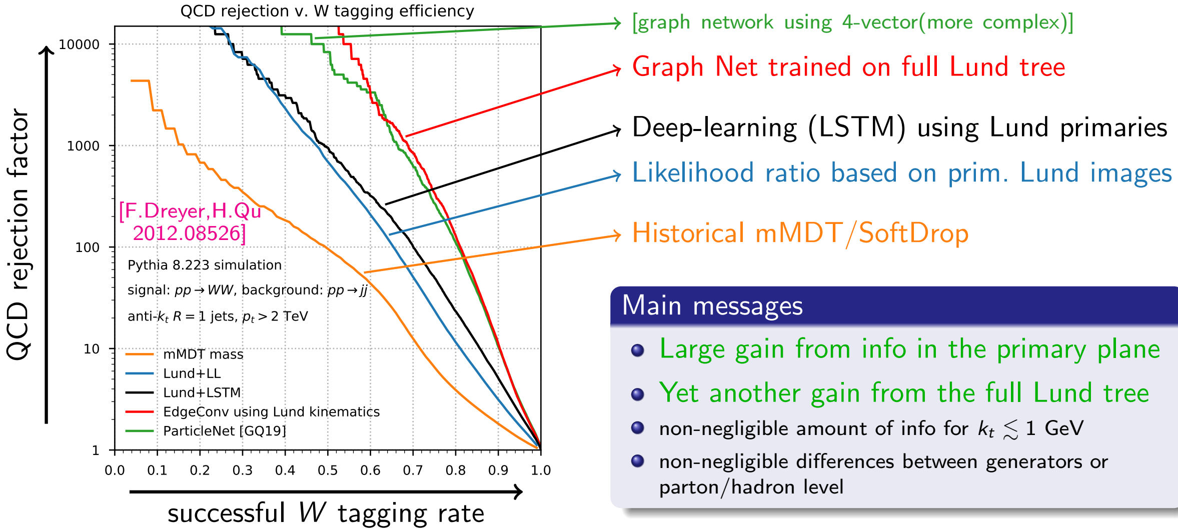
Tagging boosted W bosons (v. QCD jets) [1/2]

Clear potential on a simple image (also: many basic features recognised)





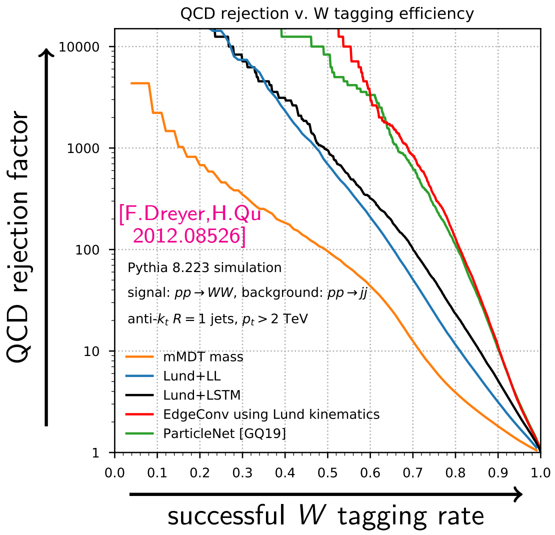
Tagging boosted W bosons (v. QCD jets) [2/2]

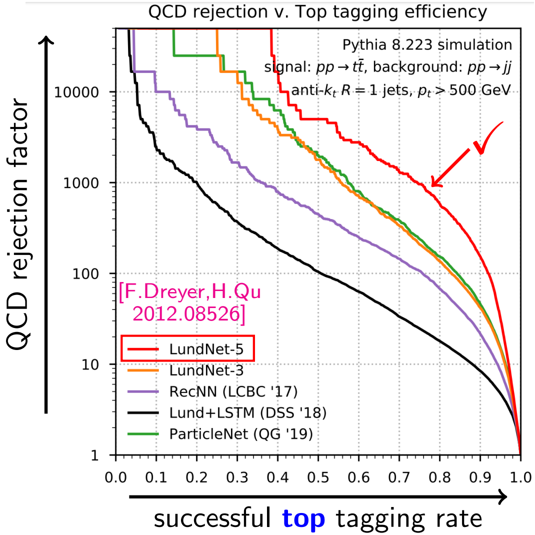


Main messages

- Large gain from info in the primary plane
- Yet another gain from the full Lund tree
- non-negligible amount of info for $k_t \lesssim 1$ GeV
- non-negligible differences between generators or parton/hadron level

Tagging boosted W bosons (v. QCD jets) [2/2]





Take-home message #3

Lund plane variables helpful in all areas of jet substructure

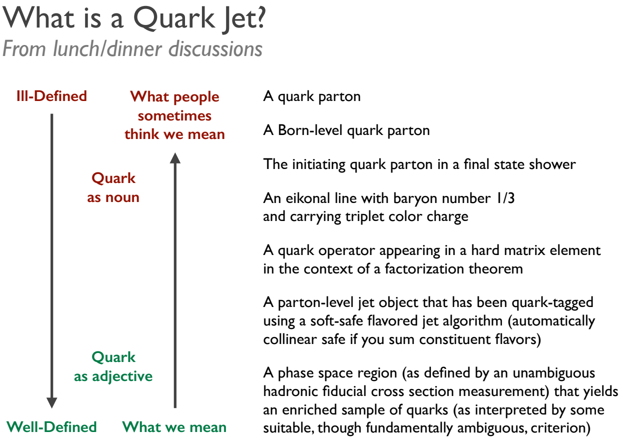
- Variables to test/develop Monte-Carlo generators
- New calculations in (p)QCD and comparisons to data
- Efficient input to Deep-Learning boosted taggers
- Possibilities to craft new observables for a specific purpose

  (Interesting also in heavy-ion collisions)

Quark/gluon discrimination

Goal: using the Lund declustering info (primary or full-tree) can we say if a jet is quark- or gluon-initiated?

Quark v. gluon jets: 0. basic considerations



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

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
- Resum logs to all orders in α_s, up to double logs
  - Each primary radiation comes with a factor $\frac{2\alpha_s(k_t)C_R}{\pi}$
  - Each subsidiary radiation comes with a factor $\frac{2\alpha_s(k_t)C_A}{\pi}$
- Probabilities: $p_{q,g} = \prod_{i\in\text{prim}} \frac{2\alpha_s(k_{ti})C_{F,A}}{\pi} \prod_{i\in\text{others}} \frac{2\alpha_s(k_{ti})C_A}{\pi}$ (up to a negligible Sudakov)
- The ratio largely cancels: $\mathbb{L}_{\text{prim,tree}} = \left(\frac{C_F}{C_A}\right)^{n_{\text{prim}}}$ [C.Frye,A.Larkoski,J.Thaler,1704.06266]
- The optimal discriminant is the primary multiplicity i.e. the Iterated SoftDrop multiplicity

Quark v. gluon jets: I. approach

Optimal discriminant (Neyman–Pearson lemma)

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

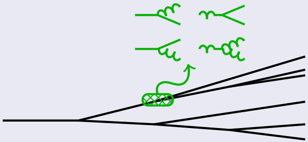
Approach #1

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
- Resum logs to all orders in α_s, up to single logs
  - single logs from "DGLAP" collinear splittings

$$P_q(\mathcal{L}_{\text{parent}}) = S_q(\Delta_{\text{prev}}, \Delta)\left[\tilde{P}_{qq}(z)p_q(\mathcal{L}_{\text{hard}})p_g(\mathcal{L}_{\text{soft}}) + \tilde{P}_{gq}(z)p_g(\mathcal{L}_{\text{hard}})p_q(\mathcal{L}_{\text{soft}})\right]$$

$$p_g(\mathcal{L}_{\text{parent}}) = S_g(\Delta_{\text{prev}}, \Delta)\left[\tilde{P}_{gg}(z)p_g(\mathcal{L}_{\text{hard}})p_g(\mathcal{L}_{\text{soft}}) + \tilde{P}_{qg}(z)p_q(\mathcal{L}_{\text{hard}})p_q(\mathcal{L}_{\text{soft}})\right]$$

  - some single logs for emissions at commensurate angles
    Note: all-order not tractable analytically; we resum any *pair* of commensurate-angle emissions
  - running coupling (in the Sudakov)

Quark v. gluon jets: II. ML validation

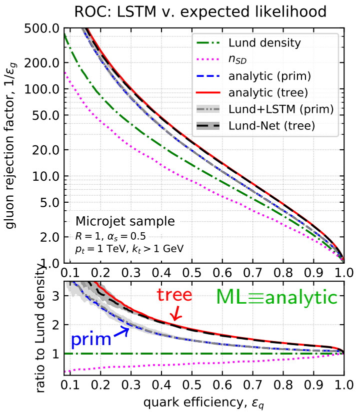
our analytic discriminant is exact/optimal in the dominant collinear limit $\theta_1 \gg \theta_2 \gg \cdots \gg \theta_n$

$\Rightarrow$ 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 $\theta_1 \gg \theta_2 \gg \cdots \gg \theta_n$

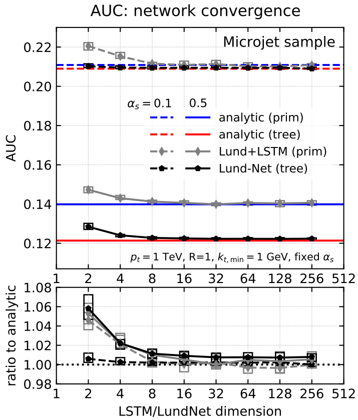
$\Rightarrow$ ML expected to give the same performance



ROC curves agree

Microjet
≡
exact
pure-collinear

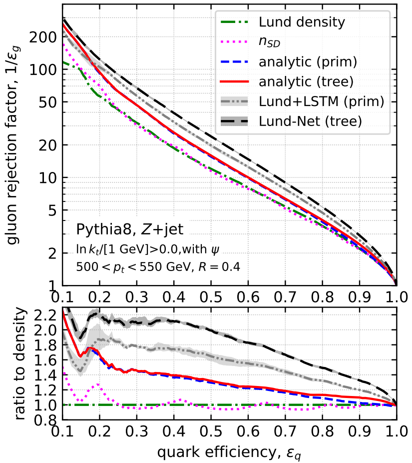
[M.Dasgupta,F.Dreyer G.P.Salam,G.Soyez, 1411.5182]



Converges for large-enough networks

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

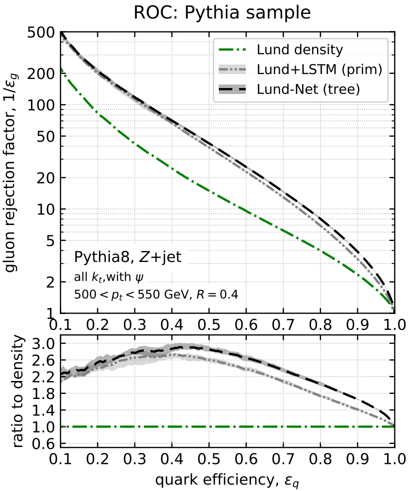
ROC: Pythia sample



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

Quark v. gluon jets: III. performance

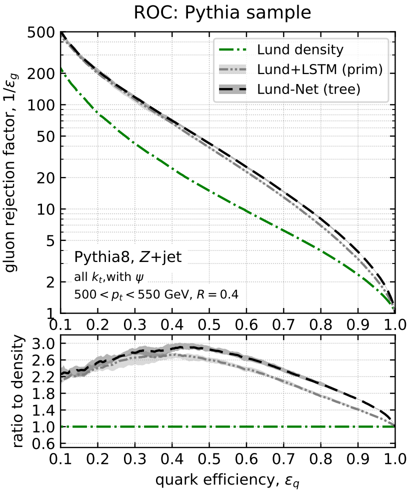
$pp \to Zq$ v. $pp \to Zg$ $\quad (p_t \sim 500$ GeV, $R = 0.4)$



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

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
- Interesting questions:
  - Analytic approach to NP?
  - Apply analytics to other systems ($W/Z/H$, top)

Resilience (1/2)

Question: is your tagger resilient to uncontrolled effects?

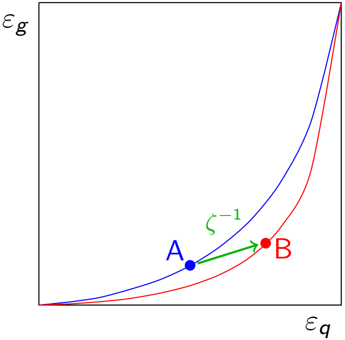
One has:

- 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?

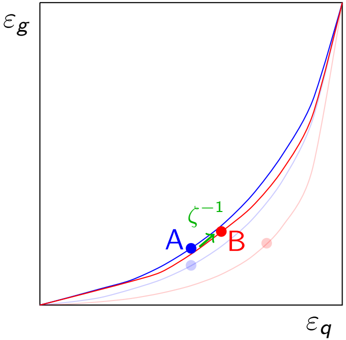
One has:

- 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)

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

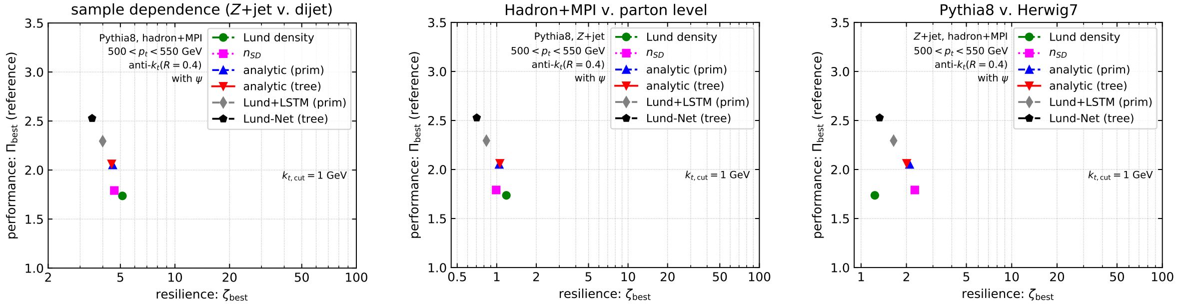
as small as possible.



Less performant
More resilient

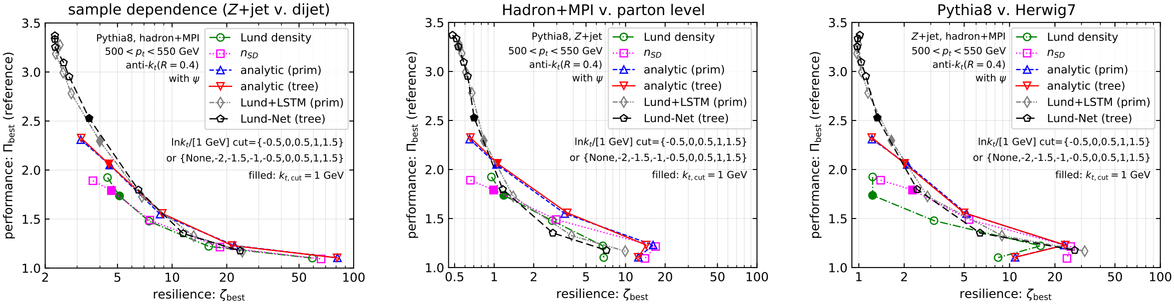
(would probably deserve a study on its own)

Resilience (2/2)



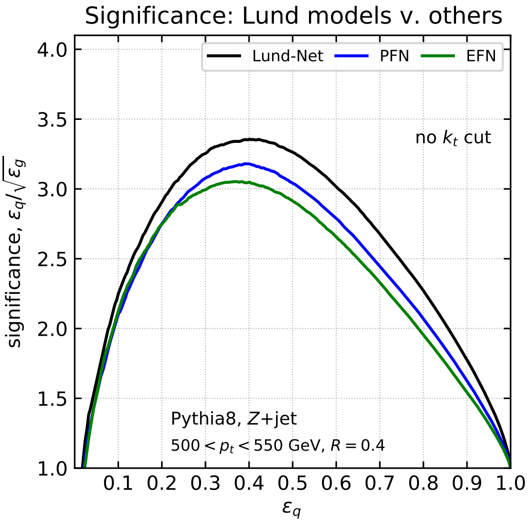
- performance $= \varepsilon_q / \sqrt{\varepsilon_g}$
- working point: $k_{t,\mathrm{cut}} = 1$ GeV, optimal performance (reference: Pythia, hadron+MPI, Z+jet)
- 3 studies: sample (Z+jet v. dijets), NP effects (hadron v. parton), generator (Pythia v. Herwig)
- performance: same ordering as before
-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

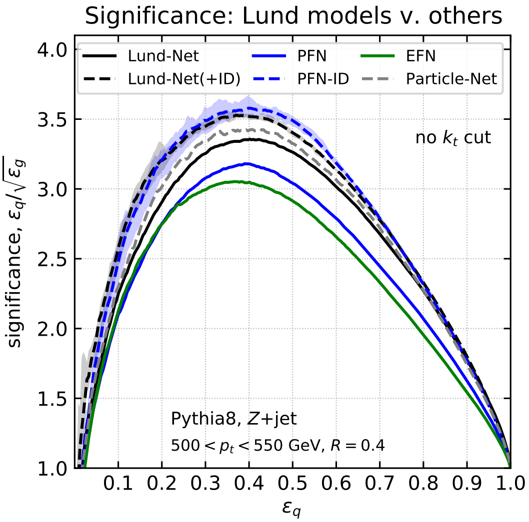


Approaches:

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

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

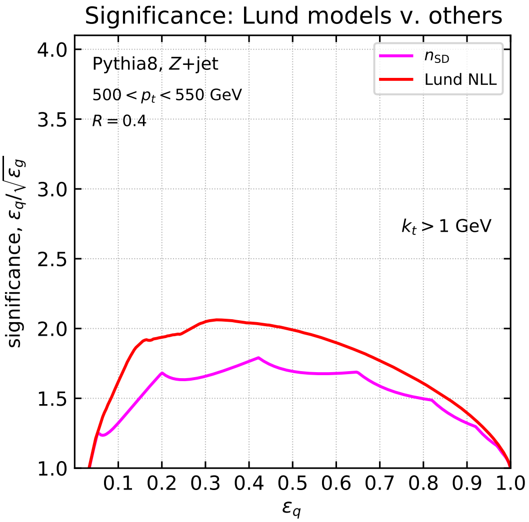
Comparison to other approaches: ML-based



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

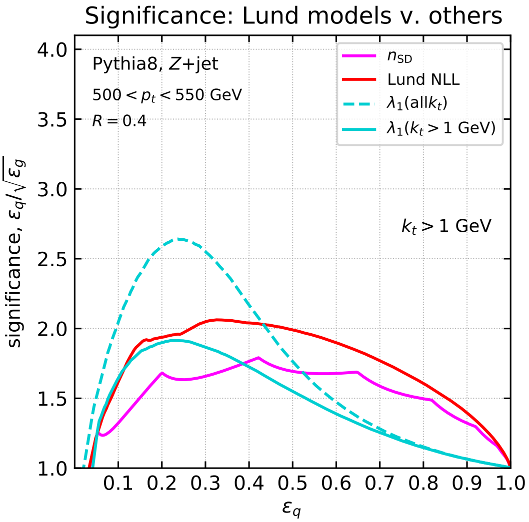


Approaches:

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

▶ clear gain from our analytic approach

Comparison to other approaches: analytics/shapes

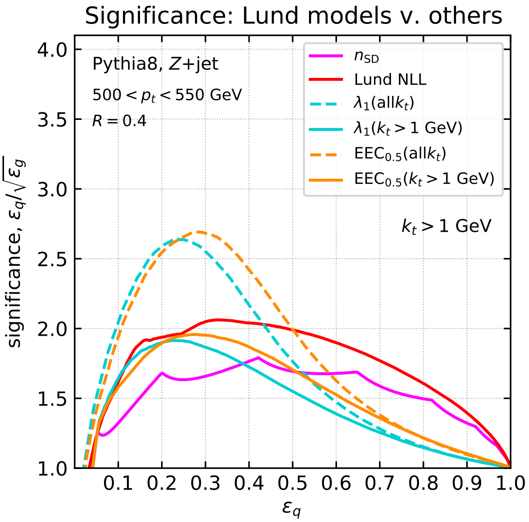


Approaches:

- 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



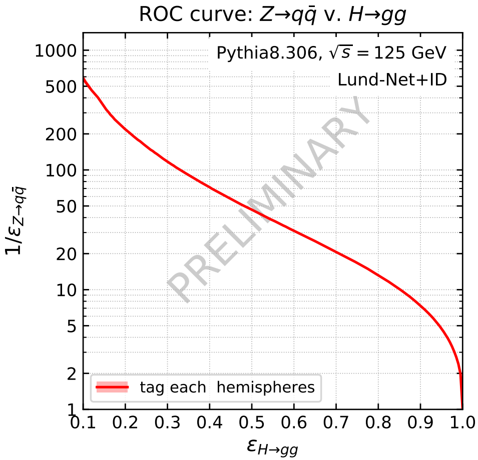
Approaches:

- 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

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

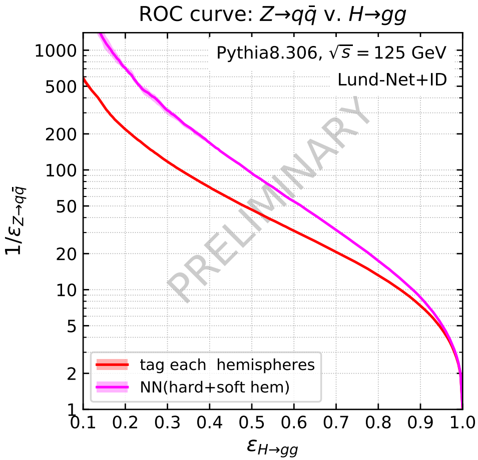


observed performance:

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

  full event clearly worse that (jet)^2

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

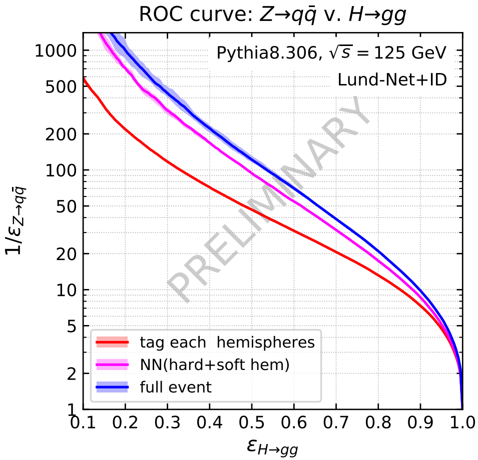


observed performance:

- 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

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

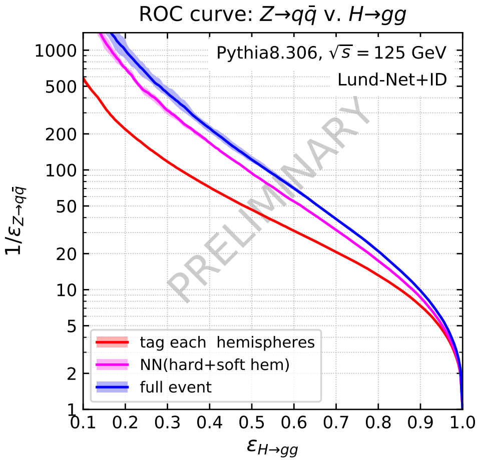


observed performance:

- 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. $e^+e^- \to H \to gg$ $(\sqrt{s} = 125$ 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)?

Conclusions

1. Lund diagrams have helped thinking about resummation and MCs
   Now they can be reconstructed in practice
   - They provide a view of a jet/event which mimics angular ordering
   - They provide a separation between different physical effects
2. Broad spectrum of applications:
   - Wide range of possible (p)QCD calculations
     Main limitation: (non-global) clustering logs; can we apply grooming-like techniques?
   - Large scope for crafting new observables ((p)QCD calculations, MC devel/validation)
   - More connections to deep learning, heavy-ion collisions, ...
3. Quark-gluon tagging:
   - analytic: single-log gives a systematic improvement over ISD multiplicity
   - deep-learning: Lund-Net shows very good performance (also for W and top tagging)

Backup

Promoting to a practical tool

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

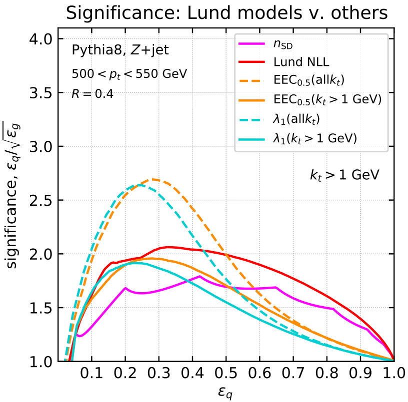
1. Cluster with Cambridge/Aachen $(d_{ij} = \Delta R_{ij})$
2. For each (de)-clustering $j \leftarrow j_1 j_2$:

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

Primary Lund plane

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

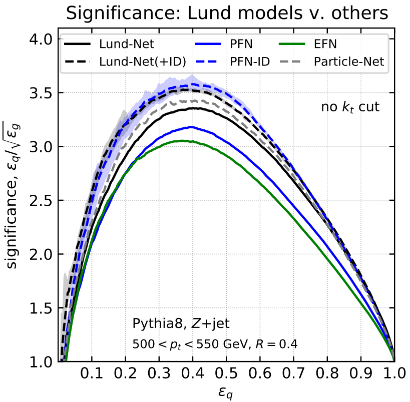
$pp \to Zq$ v. $pp \to Zg$ $\quad (p_t \sim 500$ GeV, $R = 0.4)$



- Analytic approach shows gains for $k_t > 1$ GeV
  (shapes improve at small ε_q by adding smaller k_t)

Quark v. gluon jets: III. performance v. others

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



- Analytic approach shows gains for $k_t > 1$ GeV
  (shapes improve at small ε_q by adding smaller k_t)
- ML performance on par with PFN, slightly better than Particle-Net
  (treatment of PDG-ID could maybe be improved)