Quarks and gluons in the Lund plane

Gregory Soyez, with Frederic Dreyer and Adam Takacs
based on arXiv:2112.09140

IPhT, CNRS, CEA Saclay

FCC Physics Workshop, February 7-11 2022

1. A few word about quark v. gluon tagging
2. The Lund plane representation of jets (and events)
3. Using Lund variables to tag quarks and gluons
   - An approach using perturbative QCD
   - An approach using Deep Learning
4. Applications and results
   - Validation (in the collinear limit)
   - Boosted jets at the LHC
   - a (preliminary) FCC-*ee* application

Quarks v. gluons

Main topic of this talk

Discriminate quark-initiated jets from gluon-initiated ones

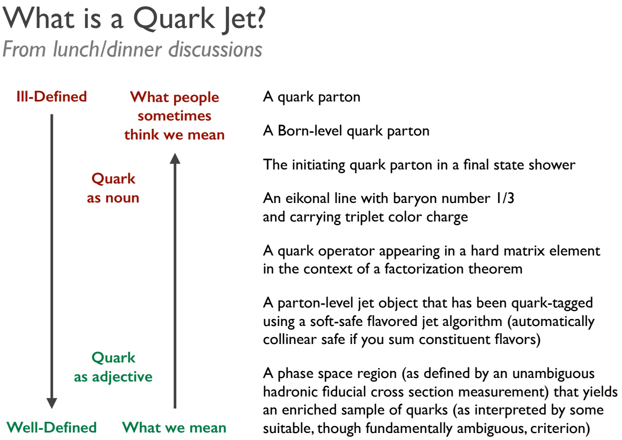
generic tool with many applications:

coll	q	g
any	BSM	QCD
hh	VBF H	$gg \to H$
ee	$Z \to q\bar{q}$	$H \to gg$

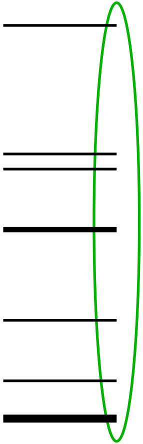
Intrinsic limitation:

- q and g not really well defined
  e.g. a "gluon jet" is different in $q\bar{q} \to Zg$ or in $gg \to gg$
- cross-section in q/g-enriched region is well-defined

see e.g. arXiv:1704.03878

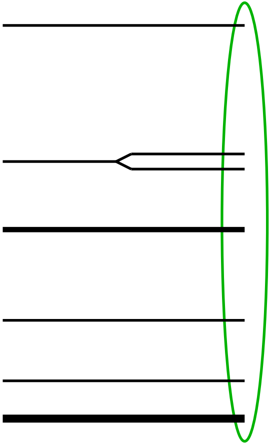


use Cambridge/Aachen to iteratively recombine the closest pair



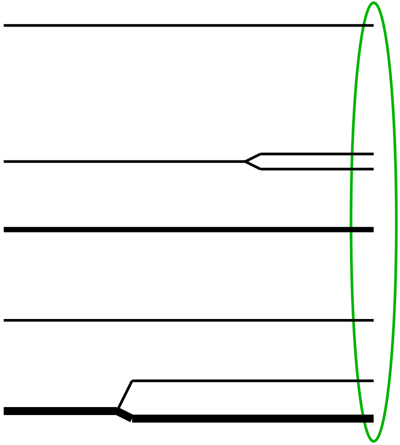
The Lund plane(s) representation (1/3)

use Cambridge/Aachen to iteratively recombine the closest pair



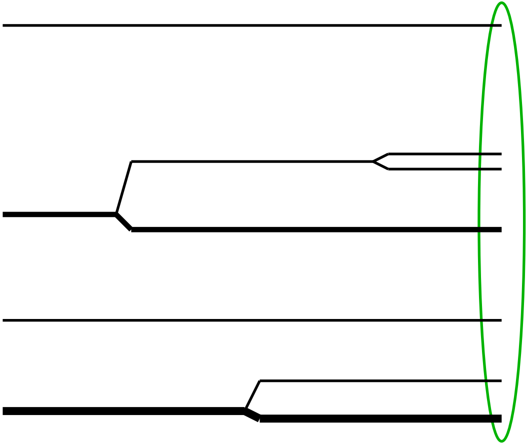
The Lund plane(s) representation (1/3)

use Cambridge/Aachen to iteratively recombine the closest pair

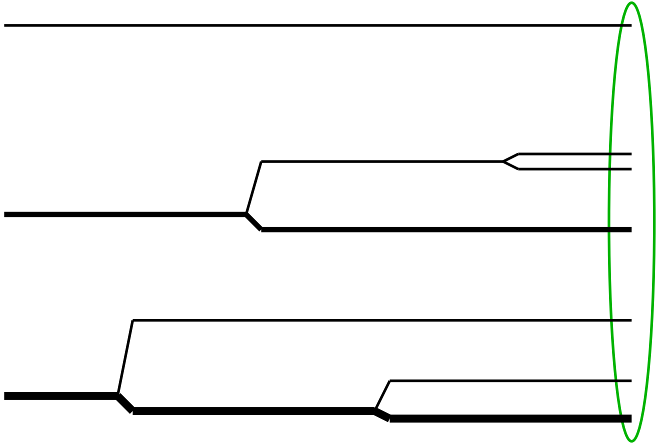


The Lund plane(s) representation (1/3)

use Cambridge/Aachen to iteratively recombine the closest pair

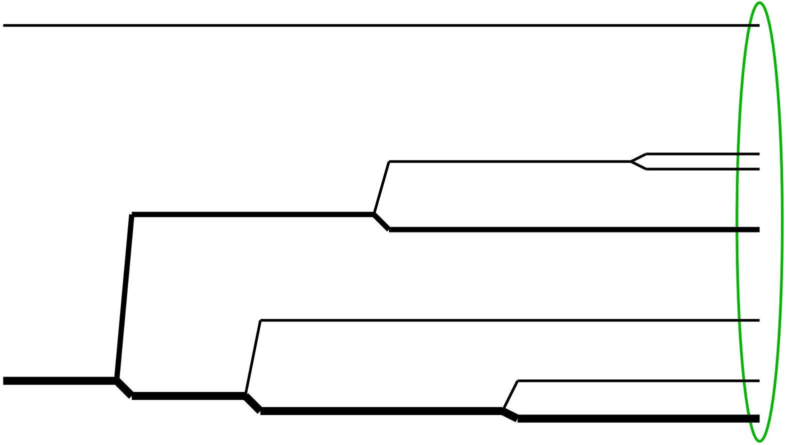


use Cambridge/Aachen to iteratively recombine the closest pair



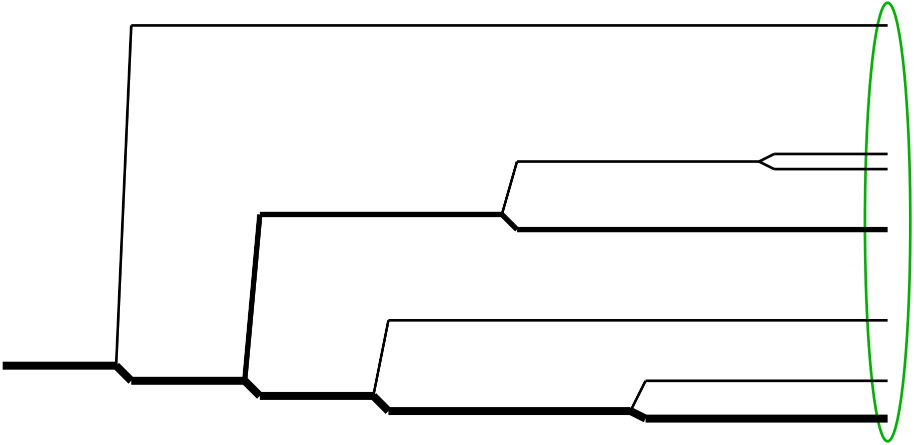
The Lund plane(s) representation (1/3)

use Cambridge/Aachen to iteratively recombine the closest pair

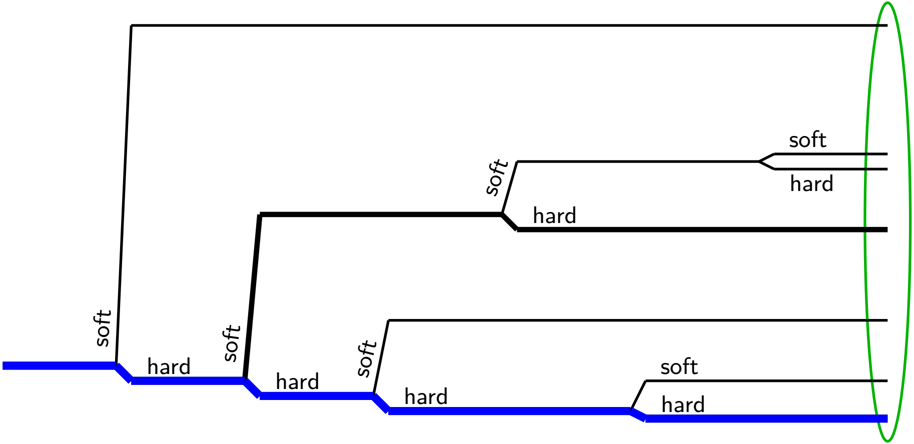


The Lund plane(s) representation (1/3)

use Cambridge/Aachen to iteratively recombine the closest pair

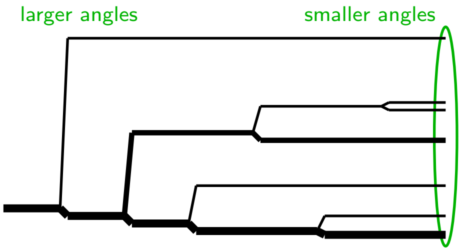


use Cambridge/Aachen to iteratively recombine the closest pair



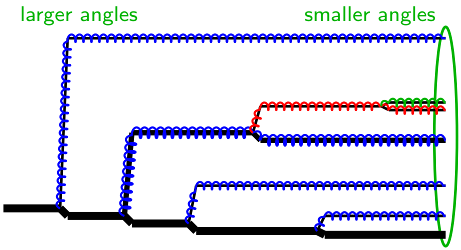
E.g.: conceptually the largest-energy (p_t or z) branch $\equiv$ emissions from the "leading parton"

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

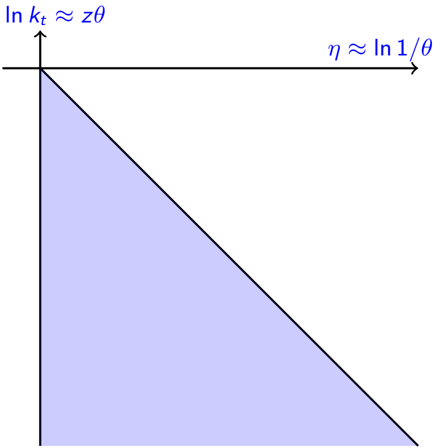


- closely follow our beloved angular ordering

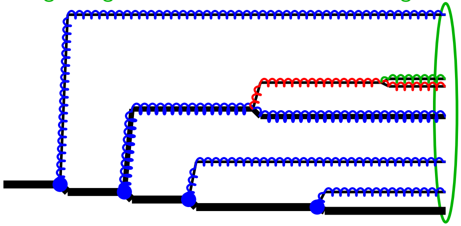
The Lund plane(s) representation (2/3)



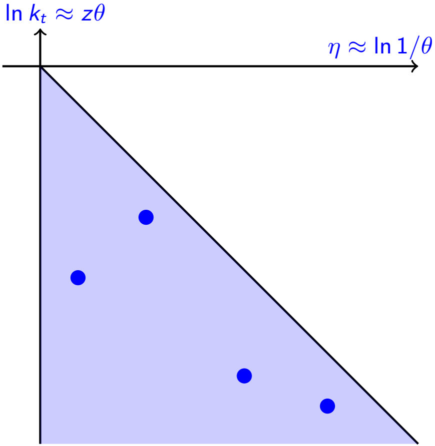
- closely follow our beloved angular ordering
- i.e. mimics partonic cascade
- can be organised in Lund planes



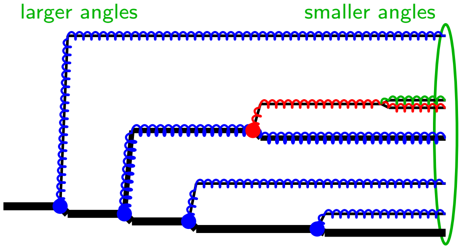
The Lund plane(s) representation (2/3)



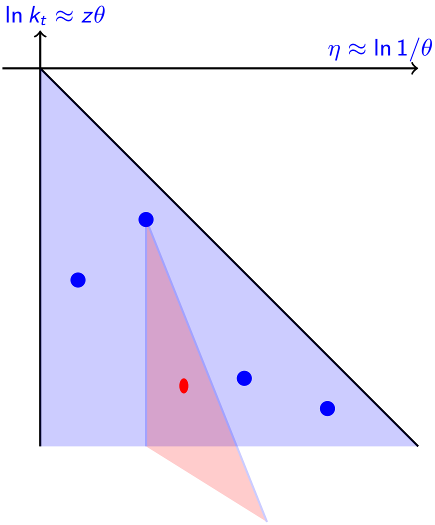
- closely follow our beloved angular ordering
- i.e. mimics partonic cascade
- can be organised in Lund planes
  - primary

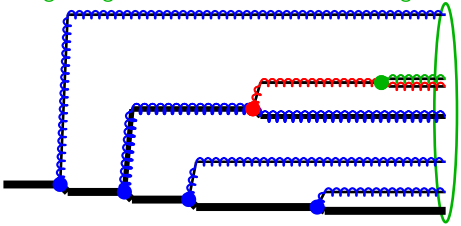


The Lund plane(s) representation (2/3)

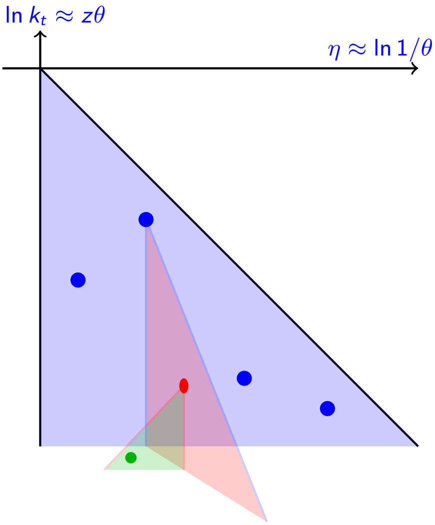


- closely follow our beloved angular ordering
- i.e. mimics partonic cascade
- can be organised in Lund planes
  - primary
  - secondary

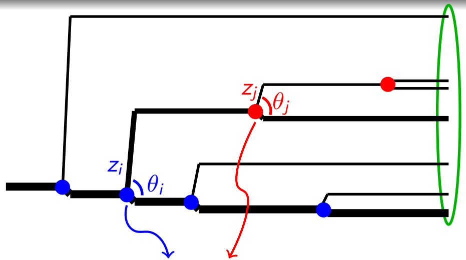




- closely follow our beloved angular ordering
- i.e. mimics partonic cascade
- can be organised in Lund planes
  - primary
  - secondary
  - ...



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



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

Two different structures

"primary plane"
(follow hard branch)

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

OR

full (de-)clustering tree

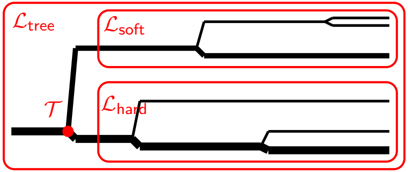
$$\mathcal{L}_{\text{tree}} \equiv \{\mathcal{T}, \mathcal{L}_{\text{hard}}, \mathcal{L}_{\text{soft}}\}$$

for *ee* events: (similar for jets in *pp*)

$$k_t = E_{\text{soft}} \sin\theta,$$

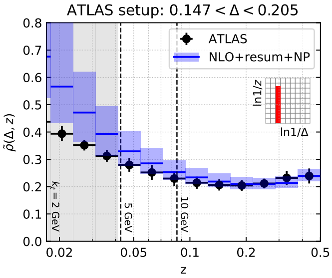
$$z = \frac{E_{\text{soft}}}{E_{\text{parent}}}$$

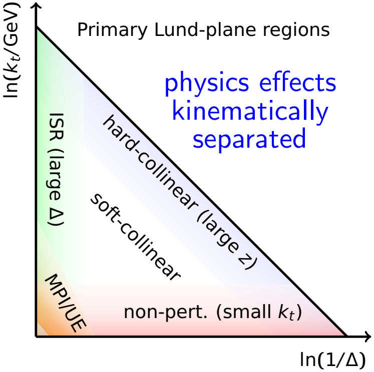
$$\psi \equiv \text{azimuthal angle}$$



(A wide range of) Lund plane applications

- Tagging: q v. g [this talk+[6]], W[1,4], t[4], ...
- Monte Carlo constraints[1]
- pQCD v. data[2,3]
- study quark-gluon plasma[5]
- inputs for Deep Learning[1,4,6]





1 F.Dreyer,G.Salam,GS, arXiv:1807.04758
2 ATLAS, CERN-EP-2020-030
3 A.Lifson,G.Salam,GS, arXiv:2007.06578
4 F.Dreyer,H.Qu, arXiv:2012.08526
5 Alice, ALICE-PUBLIC-2021-002
6 F.Dreyer,GS,A.Takacs, arXiv:2112.09140

Quark/gluon tagging in the Lund plane

Basic approaches

Idea: gluons radiate more than quarks

- shapes (angularities, energy-energy correlations,...)
- multiplicities ($n_{\rm chg}$, Iterated SoftDrop, ...)

Quark/gluon tagging in the Lund plane

Basic approaches

Idea: gluons radiate more than quarks

- shapes (angularities, energy-energy correlations,...)
- multiplicities (n_{chg}, Iterated SoftDrop, ...)

ML approaches

Idea: train a Neural Network

- EFN/PFN (1810.05165)
- ParticleNet (next talk)

Quark/gluon tagging in the Lund plane

Basic approaches

Idea: gluons radiate more than quarks

- shapes (angularities, energy-energy correlations,...)
- multiplicities (n_{chg}, Iterated SoftDrop, ...)

ML approaches

Idea: train a Neural Network

- EFN/PFN (1810.05165)
- ParticleNet (next talk)

This approach/talk

Discriminate quarks and gluons based on Lund declusterings (primaries $\mathcal{L}_{\text{prim}}$, or full tree $\mathcal{L}_{\text{tree}}$)

Idea: best discriminant given by the likelihood ratio

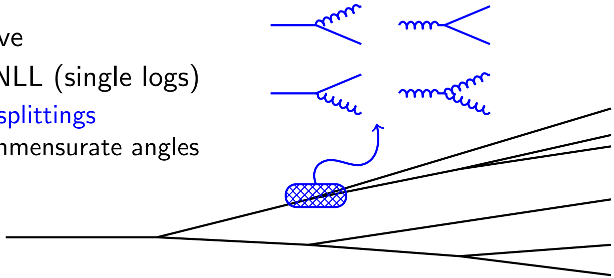
$$\mathbb{L}_{\text{prim,tree}} = \frac{p(g|\mathcal{L}_{\text{prim,tree}})}{p(q|\mathcal{L}_{\text{prim,tree}})}$$

Quark/gluon tagging using perturbative QCD

Approach #1: compute $p(q, g|\mathcal{L})$ in perturbative QCD

Calculation ingredients:

- Consider $k_t \geq k_{t,\text{cut}}$ to stay perturbative
- Resum logs to all orders in α_s, up to NLL (single logs)
  - single logs from "DGLAP" collinear splittings
  - some single logs for emissions at commensurate angles



Comments:

- LL (double logs) equivalent to the Iterated SoftDrop multiplicity
- Paper focused on jets in pp (would apply to FCC-hh) where large-angle and fixed-order corrections are process-dependent (and thus neglected)

Quark/gluon tagging using Machine Learning

Approach #2: input the Lund coordinates in a Neural Network

Primary Lund plane

Ingredients:

- LSTM network
- inputs: $\ln\theta$, $\ln k_t$, $\ln z$, Ψ
- with or without a k_t cut

Comments:

- DNN, CNN tried but less effective
- Could use Attention instead

Full Lund tree

Ingredients:

- Lund-Net (graph) network
- inputs: $\ln\theta$, $\ln k_t$, $\ln z$, Ψ
- with or without a k_t cut

Comments:

- similar to Particle-Net but the Lund structure is simpler
- Could add PDG-ID info

Validation in the collinear limit

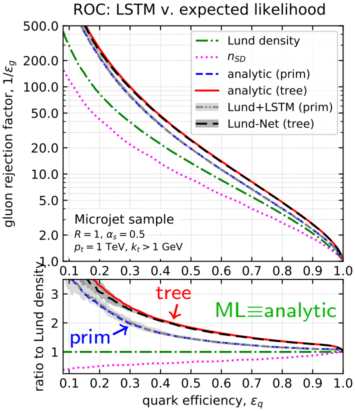
our analytic discriminant is exact/optimal in the dominant collinear limit $\theta_1 \gg \theta_2 \gg \cdots \gg \theta_n$

$\Rightarrow$ ML expected to give the same performance

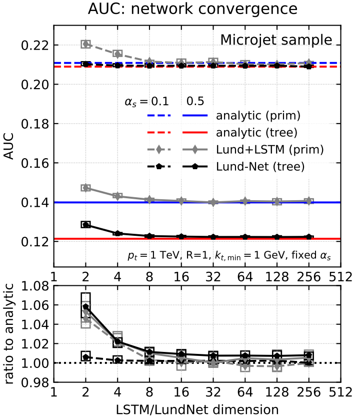
Validation in the collinear limit

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



Microjet
≡
pure-collinear

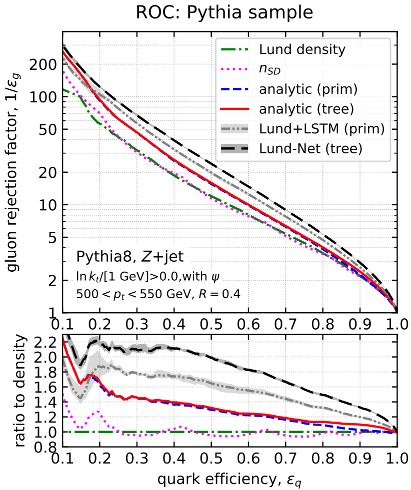


ROC curves agree

Converges for large-enough networks

Jet tagging at the LHC

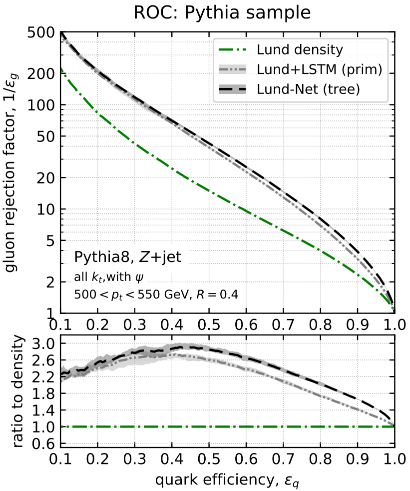
$pp \to Zq$ v. $pp \to Zg$ $\quad (p_t \sim 500\ \text{GeV},\ R = 0.4)$



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

Jet tagging at the LHC

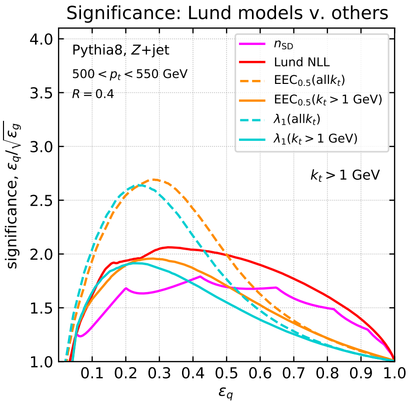
$pp \to Zq$ v. $pp \to Zg \qquad (p_t \sim 500$ GeV, $R = 0.4)$



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

Jet tagging at the LHC

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

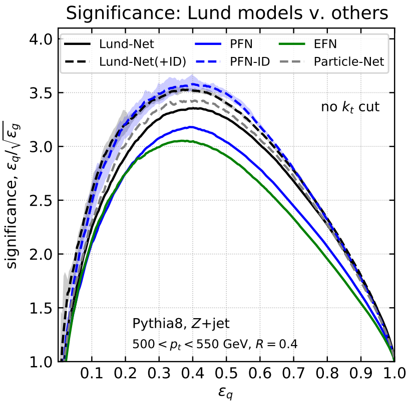


- clear performance ordering:
  1. Lund+ML > Lund analytic > ISD
  2. tree > prim
- larger gains with no k_t cut
- Analytic approach shows gains for $k_t > 1$ GeV

  (shapes improve at small ε_q by adding smaller k_t)

Jet tagging at the LHC

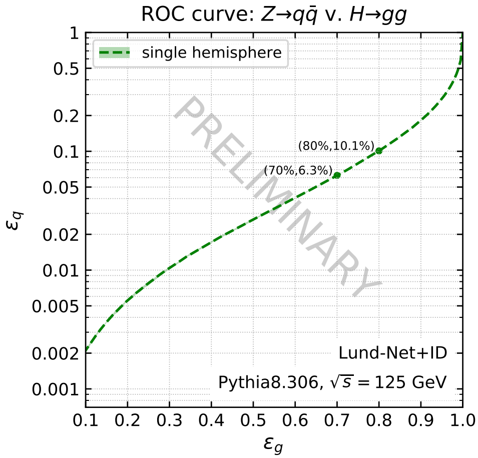
$pp \to Zq$ v. $pp \to Zg \qquad (p_t \sim 500$ GeV, $R = 0.4)$



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

$H \to gg$ at FCC-ee

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



observed performance:

- Lund-Net in single jet/hemisphere
  6% (10%) quark mistag for
  70% (80%) gluon efficiency

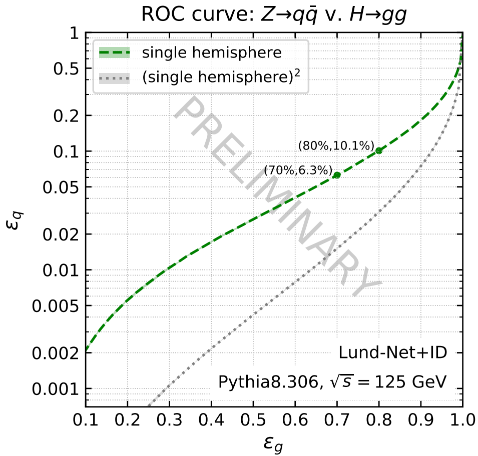
  Not quite the 1% quark mistag in 2107.02686
  (see also David d'Enterria's talk)

  In agreement with numbers from June 2021 FCC week (see also Giovanni Marchiori's talk)

  Watch out: $H \neq HZ$

$H \to gg$ at FCC-ee

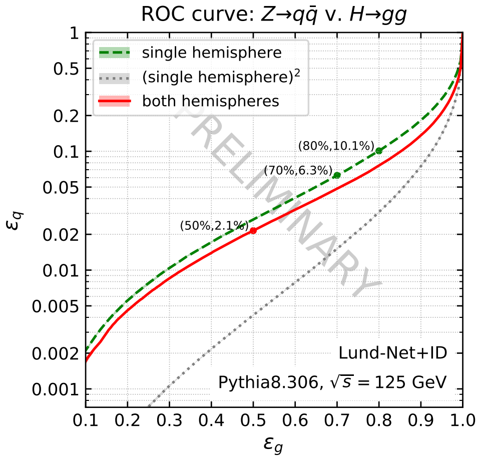
$$e^+e^- \to Z \to q\bar{q} \text{ v. } e^+e^- \to H \to gg \qquad (\sqrt{s} = 125 \text{ GeV, no ISR})$$



observed performance:

- Lund-Net in single jet/hemisphere
- assuming 2 independent tags (not physical)

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



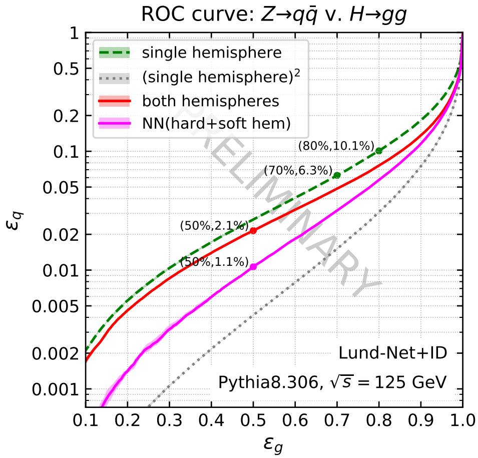
observed performance:

- Lund-Net in single jet/hemisphere
- assuming 2 independent tags (not physical)
- tagging both hemispheres
  i.e. both jets should be tagged

full event clearly worse that (jet)^2

$H \to gg$ at FCC-ee

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



observed performance:

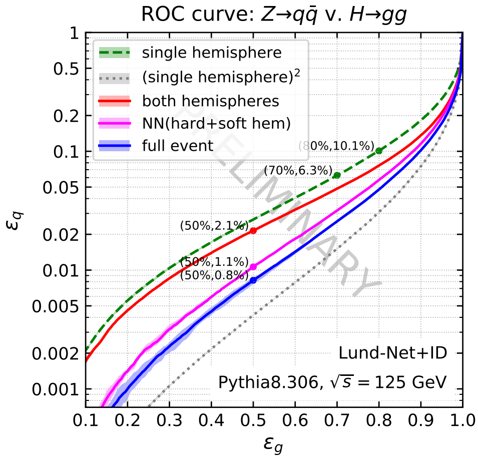
- Lund-Net in single jet/hemisphere
- assuming 2 independent tags (not physical)
- tagging both hemispheres
- double Lund-Net tag

  train separately on hard & soft hemispheres
  use another NN (or MVA) to combine the two

clear performance gain

$H \to gg$ at FCC-ee

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



observed performance:

- Lund-Net in single jet/hemisphere
- assuming 2 independent tags (not physical)
- tagging both hemispheres
- double Lund-Net tag
- Lund-Net for the full event
  Another performance gain

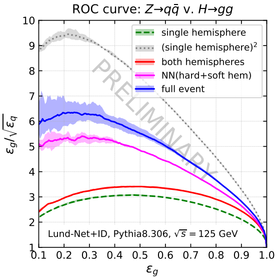
Watch out: numbers to be taken with care

- fixed-order corrections are relevant at large ε_g
- no ISR, no detector effects, ...

Conclusions

1. Quark/gluon discrimination is a complex task with many applications

2. Lund declusterings apply to many domains of jet physics
$\Rightarrow$ braod usefulness for LHC/FCC-*ee*/FCC-*hh*



3. Focus on Lund declusterings for quark/gluon tagging:
   - maximise likelihood ratio either in pQCD or with ML
   - good performance seen at the LHC

4. Preliminary FCC-*ee* study gives
$\lesssim 1\%$ $Z \to q\bar{q}$ mistag rate for $\sim 50\%$ $H \to gg$ signal efficiency
How much better can we do with better theory understanding (higher fixed-order, better parton shower, better tunes, more restrictive tagging, better PID, , ...)?