Jets and their substructure

Giovanni Stagnitto

University of Zurich

Swiss National Science Foundation

LHCP 2023, Belgrade, 22-26.05.2023
What are jets?

*Naive definition:* collimated bunch of hadrons flying roughly in the same direction

*Proper definition:* a collection of hadrons defined by means of a jet algorithm

“Jet [definitions] are legal contracts between theorists and experimentalists”
MJ Tannenbaum
The $k_t$ algorithm and its siblings

$$d_{ij} = \min(p_{ti}^{2p}, p_{tj}^{2p}) \frac{\Delta y^2 + \Delta \phi^2}{R^2}$$
$$d_{iB} = p_{ti}^{2p}$$

$p = 1$ \hspace{1cm} $k_t$ algorithm


$p = 0$ \hspace{1cm} Cambridge/Aachen algorithm

Y. Dokshitzer, G. Leder, S. Moretti and B. Webber, JHEP 08 (1997) 001

$p = -1$ \hspace{1cm} anti-$k_t$ algorithm

MC, G. Salam and G. Soyez, arXiv:0802.1189

NB: in anti-kt pairs with a hard particle will cluster first; if no other hard particles are close by, the algorithm will give perfect cones
\[ d_{ij} = \min(p_{t_i}^{2p}, p_{t_j}^{2p}) \frac{\Delta y^2 + \Delta \phi^2}{R^2} \quad d_{iB} = p_{t_i}^{2p} \]
\[ d_{ij} = \min(p_{ti}^{2p}, p_{tj}^{2p}) \frac{\Delta y^2 + \Delta \phi^2}{R^2} \quad d_{iB} = p_{ti}^{2p} \]

\[p = 0 \quad \text{Cambridge/Aachen algorithm}\]

Cambridge/Aachen: iteratively recombine the closest pair

Particularly useful when looking “inside” the jet…
Jet substructure in one slide
from Simone Marzani

- the two major goals of the LHC
- search for new particles
- characterise the particles we know
- jets can be formed by QCD particles but also by the decay of massive particles (if they are sufficiently boosted)
- how can we distinguish signal jets from background ones?

\[ p_t \gg m \]

- the final energy deposition pattern is influenced by the originating splitting
- hard vs soft translates into 2-prong vs 1-prong structure
- picture is muddled by many effects (hadronisation, underlying event, pileup)
- two-step procedure:
  - grooming: clean the jets up by removing soft radiation
  - tagging: identify the features of hard decays and cut on them
Disclaimer

Substructure of jets is a very broad topic, with a lot of recent developments …
Disclaimer

... a standard topic, with a dedicated textbook!

Simone Marzani
Gregory Soyez
Michael Spannowsky

Looking Inside Jets


Up-to-date version on the arXiv (1901.10342)
Disclaimer

I will focus on a single tool, adopted in a wide range of applications, the **Lund plane** [Z. Phys. C43 (1989) 625] a way of depicting the pattern of QCD radiation, inside a jet or in the whole event

In particular, I will show some examples of:

1) analytic calculations
2) machine learning techniques
3) heavy quark studies based on the Lund plane.

Let’s first define it!
For each step of the declustering, record the variables:

\[ \Delta_{ab} = \sqrt{(y_a - y_b)^2 + (\phi_a - \phi_b)^2}, \quad k_t = p_{tb} \Delta_{ab} \]

and add a point in the primary, secondary, etc. planes. We associate a **kinematic structure to a high-energy jet**.
Lund plane & analytics
Lund jet plane density

Simplest observable, defined on the primary Lund plane:

\[ \rho(\Delta, k_t) = \frac{1}{N_{\text{jet}}} \frac{d n_{\text{emission}}}{d \ln k_t d \ln 1/\Delta} \]

At LO: \( \rho_i \simeq \frac{2\alpha_s(k_t)C_i}{\pi}, (C_q = C_F, C_g = C_A) \)

Clear separation between QCD regimes!
Infrared and collinear safe (if we consider a “pixel”)

\( \sqrt{s} = 14 \text{ TeV}, \rho > 2 \text{ TeV} \)
Pythia8.230(Monash13)
Lund Plane density measurements

ATLAS
\[ \sqrt{s} = 13 \text{ TeV}, \ 139 \text{ fb}^{-1}, \ p_{T} > 675 \text{ GeV} \]

\frac{\ln(1/z)}{\ln(1/(z + \rho^\text{core})/\rho^\text{emission})}

\begin{align*}
\ln(1/z) & = 10^{-2} \\
\ln(1/(z + \rho^\text{core})/\rho^\text{emission}) & = 10^1
\end{align*}

\begin{align*}
\Delta R & = \Delta R(\text{emission, core})
\end{align*}

ALICE Preliminary
pp \( \sqrt{s} = 13 \text{ TeV} \)
Charged-particle jets anti-\( k_T \), \( R = 0.4 \)
\( \Delta \eta < 0.5 \), \( 0.25 < \rho^\text{ch} < 120 \text{ GeV/c} \)

\begin{align*}
\ln(k_T/\text{GeV}) & = 0.1 \\
\ln(\rho^\text{ch}/\text{GeV}) & = 0.1
\end{align*}

ALICE-PUBLIC-2021-002

CMS Preliminary
\( 138 \text{ fb}^{-1} (13 \text{ TeV}) \)
AK4 jets
\( p_T^\text{ch} > 700 \text{ GeV} \), \( \Delta \eta < 1.7 \)

1.2
0.8
0.4
0.2
1.2

CMS PAS SMP-22-007
Lund Plane density & MCs

Up to 20-30% difference between Monte Carlo generators in different slices of the plane

Ability of the Lund jet plane to isolate physical effects

→ useful input to both perturbative and non-perturbative model development and tuning
Lund Plane density at all-orders

[Lifeon, Salam, Soyez (2007.06578)]

Logarithmically dominant terms with structure:

\[ \alpha_s^{n+1} \ln^m \Delta \ln^{n-m} z, \quad 0 \leq m \leq n, \quad z = \frac{k_t}{p_{t,\text{jet}} \Delta} \]

Their resummation requires to deal with:

- Running coupling corrections (numerically dominant)
- Hard-collinear logarithms (can change flavour)
- Soft effects (large-angle emissions)
- Clustering logarithms

Non-perturbative estimated through Monte Carlo

Matching to fixed-order NLO

Clear separation of contributions

Non-perturbative  Resummation  NLO

Good agreement with ATLAS data in several slices of the plane
Defined as (average) number of Lund declustering
(in the full tree) with $k_t \geq k_{t,\text{cut}}$

Computed up to NNDL, with $L = \ln(Q/k_{t,\text{cut}})$

$$\langle N(\alpha_s, L) \rangle = \langle N(\alpha_s, 0) \rangle \left[ h_1(\alpha_s L^2) + \sqrt{\alpha_s} h_2(\alpha_s L^2) + \alpha_s h_3(\alpha_s L^2) + \ldots \right]$$

Novel method: recycle DL results with insertions of NDL or NNDL genuine ingredients

Black dots \( \propto \alpha_s L^2 \)
Blue dots \( \propto \alpha_s L \), Red dots \( \propto \alpha_s \)
Lund multiplicity at LEP...
[Medves, Soto-Ontoso, Soyez (2205.02861)]

Lund multiplicity at LEP

\[ e^+ e^- \rightarrow Z \rightarrow \text{jets} \]
\[ \sqrt{s} = 91.2 \text{ GeV} \]

OPAL Cambridge multiplicity

\[ \langle N_{\text{Cam}} \rangle \]

\[ \frac{\text{ratio to data}}{10^{-4} \quad 10^{-3} \quad 10^{-2} \quad 10^{-1}} \]

\[ \text{rel. uncert.} \]

\[ \frac{\text{pert.}}{10^{-4} \quad 10^{-3} \quad 10^{-2} \quad 10^{-1}} \]

\[ \text{NP} \quad \text{total} \]
... and at the LHC

[Medves, Soto-Ontoso, Soyez (2212.05076)]

Counting the mean number of subjects per anti-$k_t$ jet with relative $k_t \geq k_{t,\text{cut}}$

Resummation up to NNDL in $L = \ln(p_\perp R/k_{t,\text{cut}})$:

- Universal ingredients from $e^+e^-$ event-wide result
- Presence of jet radius impacts the large-angle components starting at NDL, with a process dependence (e.g. $Z$ + jets or dijets)
- Additional presence of experimental fiducial cuts used for the jet analysis in a collider environment.

Precision calculation has the potential to serve as benchmarks to test and develop MC event generators.
Lund Plane & machine learning
Exploit different structure of primary Lund plane for QCD and signal jets → use density to build likelihood, sequence of declusterings as input to LSTM or DNN, image as input to CNN

Tagging $W$ decay
[Dreyer, Salam, Soyez (1807.04758)]

QCD rejection v. $W$ efficiency

Pythia8(Monash13)
Delphes+SPRA1

$\sqrt{s} = 14$ TeV, $p_t > 2$ TeV

Pythia8.230(Monash13)
A simple one-variable discriminant (color ring) fails in the $H \rightarrow gg$ case, whereas the Lund plane CNN maintains its discrimination power.

Tagging $H$ decays

[Khosa, Marzani (2105.03989)]
[Cavallini, GS et al. (2112.09650)]

Use CNN trained on primary Lund plane images of $Z(H \rightarrow b\bar{b}/gg)$ and $Zb\bar{b}/Zjj$
Exploit the full Lund plane

[Dreyer, Qu (2012.08526)]

Lund tree as input to a graph NN (GNN), dubbed **LundNet**
State-of-the-art performances

---

**QCD rejection - W tagging efficiency**

- Pythia 8.223 simulation
- Signal: $pp \rightarrow WW$, background: $pp \rightarrow jj$
- anti-$k_t$, $R = 1$ jets, $p_t > 500$ GeV

---

**QCD rejection - Top tagging efficiency**

- Pythia 8.223 simulation
- Signal: $pp \rightarrow t\bar{t}$, background: $pp \rightarrow jj$
- anti-$k_t$, $R = 1$ jets, $p_t > 500$ GeV
Two variants:

LundNet-3: trained on $(\ln k_t, \ln \Delta, \ln z)$

LundNet-3: trained on $(\ln k_t, \ln \Delta, \ln z, \ln m, \psi)$

$m$ invariant mass of the pair

$$\psi = \tan^{-1}\left(\frac{y_b - y_a}{\phi_b - \phi_a}\right),$$

azimuthal angle around subject’s axis

LundNet-5 more performant, but LundNet-3 is more resilient to non-perturbative effects

Is LundNet-5 is extrapolating some information on emissions below the $k_{t,\text{cut}}$?
Quark/gluon discrimination

[Dreyer, Soyez, Takacs (2112.09140)]

Optimal discriminant: likelihood ratio

\[ \mathbb{L} = \frac{p_g(\mathcal{L})}{p_q(\mathcal{L})}, \text{ with } \mathcal{L} \text{ the Lund primary tree or full tree} \]

can be calculated analytically up to single logs

Gain in performance when considering the full tree (better kinematics and treatment of correlations)

Lund + ML models have better performance than analytics: what are they learning?
We can work in a setup in which the analytic approach corresponds to the exact likelihood-ratio discriminant (similarly to [Kasieczka, Marzani, Soyez, GS (2007.04319)]):

- events generated in the strong strong-angular-ordered limit

\[ \alpha_s \to 0 \] → ML gives same performance

Moving progressively to the single-logarithmic asymptotic limit, \( \alpha_s \to 0 \) at fixed as \( \alpha_s \ln(Q/k_t,\text{cut}) \), the difference between the two approaches reduces.
We can work in a setup in which the analytic approach corresponds to the exact likelihood-ratio discriminant (similarly to [Kasieczka, Marzani, Soyez, GS (2007.04319)]):

events generated in the strong strong-angular-ordered limit → ML gives same performance

Moving progressively to the single-logarithmic asymptotic limit, $\alpha_s \to 0$ at fixed as $\alpha_s \ln(Q/k_t,\text{cut})$, the difference between the two approaches reduces.

Gain in performance for ML come from effects that are not fully under control (subleading effects beyond single logarithms, large-angle soft emissions, non-perturbative effects)
Tagging $b$-jets

[Fedkevych, Khosa, Marzani, Sforza (2202.05082)]

Study in the boosted region (where $b$-tagging performance usually degrades)

- CNN on primary Lund plane
- DNN is a combination of angularities (single-variable discriminants)
- JetFitter and IP3D are low-level algorithms based on charged particle track reconstruction
- DL1 is a high-level tagger, combining low-level ones

Lund plane CNN has performances similar to dedicated $b$-tagging algorithms
Lund Plane & heavy quarks
Dead-cone effect
[ALICE (2106.05713) Nature 605 (2022)]

The technique introduced in [Cunqueiro, Ploskon (1812.00102)] is based on a C/A declustering sequence, following the $D^0$ dead-cone effect.

Observation of dead-cone effect from measurement of angular distribution

$$R(\theta) = \frac{1}{N^{D^0\text{jets}}} \frac{\text{d}n^{D^0\text{jets}}}{\text{d}\ln(1/\theta)} \bigg/ \frac{1}{N^{\text{inclusive jets}}} \frac{\text{d}n^{\text{inclusive jets}}}{\text{d}\ln(1/\theta)}$$

$\theta$: polar angle in the laboratory frame

- ALICE Data
- PYTHIA 8 LQ / inclusive no dead-cone limit
- PYTHIA 8
- SHERPA LQ / inclusive no dead-cone limit
- C/A reclustered

other parameters: $p_T, \text{inclusive jet}$, $|\Delta R| < 0.5$, $E_{\text{Radiator}}$, $|\Delta \phi| < 2.8 \text{ GeV/c}$, $|\Delta \phi| > 200 \text{ MeV/c}$, $|\Delta \phi| < 10 \text{ GeV}$.
Dead-cone searches in heavy-ion

[Cunqueiro, Napoletano, Soto-Ontoso (2211.11789)]

New groomer (Late-$k_t$), selecting the most collinear splitting above a certain $k_{t,\text{cut}}$

→ suited to heavy-ion environment (reduces the impact of uncorrelated thermal background, typically manifest as fake large angle splittings)
Also LHCb in the game
[slides of Ibrahim Chahrour, on behalf of the LHCb collaboration, DIS2023]

LJP at forward rapidities
Pythia8 Simulations

pp collisions
\( \sqrt{s} = 13 \text{ TeV} \)
\( 2.5 < \eta < 4 \)
\( p_T > 20 \text{ GeV} \)

Light quark jets
Charm jets
Beauty jets

Dead cone at forward rapidities
Pythia8 Simulations

pp collisions
\( \sqrt{s} = 13 \text{ TeV} \)
\( 2.5 < \eta < 4 \)
\( p_T > 20 \text{ GeV} \)

Charm/Light
Beauty/Charm
Beauty/Light

Suppressed collinear radiation = dead cone effect!

Dead cone effect is most prominent for Beauty/Light ratio
Conclusions and outlook

Lund (jet) plane is a **unique tool** for collider phenomenology:

- **Clear separation** of perturbative and non-perturbative regimes
  → extraction of strong coupling constant?
- **Sensitivity to disparate scales**, from few GeV up to several TeV
  → ideal tool for resummation and Parton Showers (PS) studies
- **Observables based on Lund plane amenable to calculability up to high orders**
  → precise comparisons with data and benchmark calculations
- **Lund trees or images as theory-friendly input to machine learning algorithms**
  → good performance and resilience at the same time
Jet substructure techniques are now routinely used in collider phenomenology. This specialised workshop evolves around a recent tool called the Lund Jet Plane(s). The main idea is to use the Cambridge/Aachen clustering technique (i.e. a roughly angular-ordered clustering tree) to associate a kinematic structure, akin to the Lund planes used in resummations and in Monte Carlo generators, to a high-energy jet. This structure can then be used in a wide range of applications. The goal of this workshop is to provide a theoretical and experimental overview of these applications and their connections with other tools in the field. A special emphasis will be put on recent developments and on discussions of future potential directions. This includes the following list of topics:

- Constraints on Monte Carlo generators from Lund plane density measurements
- Tagging of light-quarks vs. gluon
- Boosted V/H/t vs. QCD jets discrimination
- Mass effects in the Lund plane (dead cone, heavy flavor tagging)
- Applications to BSM searches
- Studies of the quark-gluon-plasma in heavy-ion collisions
- Jet substructure measurements (generalised angularities, groomed observables,..)
- Machine learning tools (e.g., LundNet, ParticleNet, GNN)
- Lund-plane observables to constrain parton showers with N^kLL resummation
- Strategies to mitigate quark/gluon fraction issues.
- Possible as extractions with jet substructure.
BACKUP
“CS” = BDT architecture on high-level color-sensitive variables (CS): pull angle $\theta_p$, components of the pull vector $t_\parallel$ and $t_\perp$, color ring $\Theta$ (CR), $D_2$

“CS + LP_{CNN}” = BDT combined with CNN trained on Lund plane images

$$\Theta = \frac{\Delta^2_{ka} + \Delta^2_{kb}}{\Delta^2_{ab}}$$
Quark/gluon discrimination

[Dreyer, Soyez, Takacs (2112.09140)]
Quark/gluon discrimination

[Dreyer, Soyez, Takacs (2112.09140)]
[Dreyer, Salam, Soyez (1807.04758)]