The Lund Jet Plane

BOOST, Paris, 18 July 2018

Frédéric Dreyer

based on arXiv:1807.04758

with Gavin Salam & Gregory Soyez

Identifying boosted objects

Two main approaches to identify boosted decays:

1. Manually constructing substructure observables that help distinguish between different origins of jets.
2. Apply machine learning models trained on large input images or observable basis.

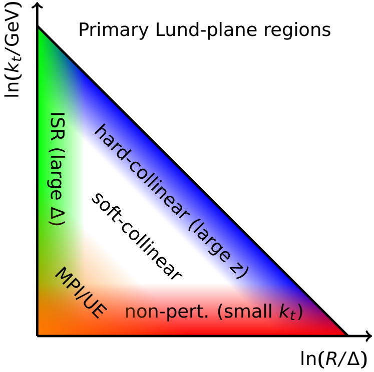
Aim of this talk: present a method bridging some of the gap between these two techniques.

THE LUND PLANE

Lund diagrams

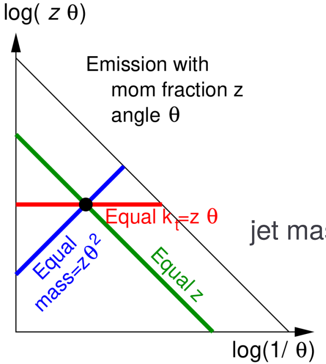
- Lund diagrams in the $(\ln z\theta, \ln\theta)$ plane are a very useful way of representing emissions.
- Different kinematic regimes are clearly separated, used to illustrate branching phase space in parton shower Monte Carlo simulations and in perturbative QCD resummations.
- Soft-collinear emissions are emitted uniformly in the Lund plane

$$dw^2 \propto \alpha_s \frac{dz}{z}\frac{d\theta}{\theta}$$

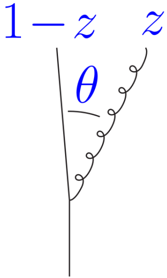


Lund diagrams

- Features such as mass, angle and momentum can easily be read from a Lund diagram.



$$\text{jet mass} \equiv \frac{m^2}{p_t^2 R^2} \approx z_1 \theta_1^2$$



- Substructure algorithms can often also be interpreted as cuts in the Lund plane.

Studying jets in the Lund plane

Lund diagrams can provide a useful approach to study a range of jet-related questions

- First-principle calculations of Lund-plane variables.
- Constrain MC generators, in the perturbative and non-perturbative regions.
- Brings many soft-drop related observables into a single framework.
- Impact of medium interactions in heavy-ion collisions.
- Boosted object tagging using Machine Learning methods.

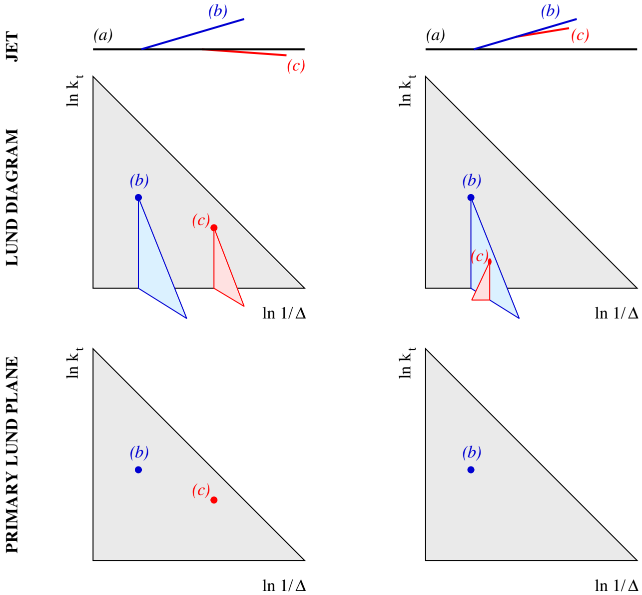
We will use this representation as a novel way to characterise radiation patterns in a jet, and study the application of recent ML tools to this picture.

Lund plane representation

To create a Lund plane representation of a jet, recluster a jet j with the Cambridge/Aachen algorithm then decluster the jet following the hardest branch.

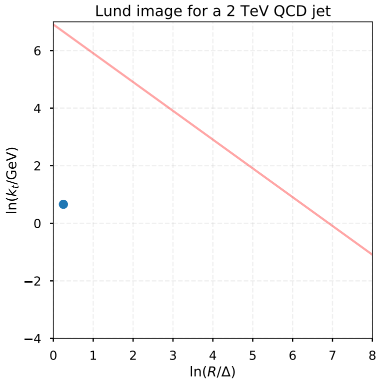
1. Undo the last clustering step, defining two subjets j_1, j_2 ordered in p_t.
2. Save the kinematics of the current declustering
$$\Delta \equiv (y_1 - y_2)^2 + (\phi_1 - \phi_2)^2, \quad k_t \equiv p_{t2}\Delta,$$
$$m^2 \equiv (p_1 + p_2)^2, \quad z \equiv \frac{p_{t2}}{p_{t1}+p_{t2}}, \quad \psi \equiv \tan^{-1}\frac{y_2 - y_1}{\phi_2 - \phi_1}.$$
3. Define $j = j_1$ and iterate until j is a single particle.

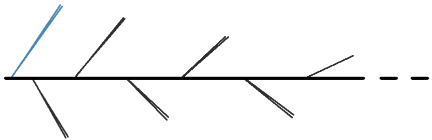
Lund plane representation



Lund representation of a jet

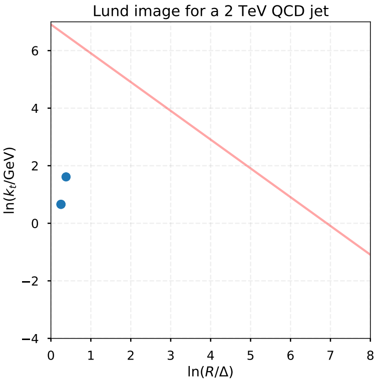
- Each jet has an image associated with its primary declustering.
- For a C/A jet, Lund plane is filled left to right as we progress through declusterings of hardest branch.
- Additional information such as azimuthal angle ψ can be attached to each point.

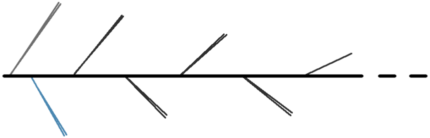




Lund representation of a jet

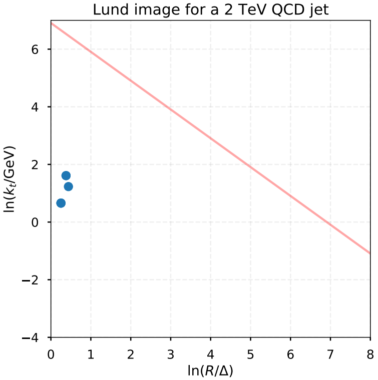
- Each jet has an image associated with its primary declustering.
- For a C/A jet, Lund plane is filled left to right as we progress through declusterings of hardest branch.
- Additional information such as azimuthal angle ψ can be attached to each point.

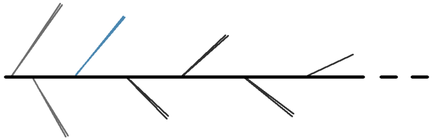




Lund representation of a jet

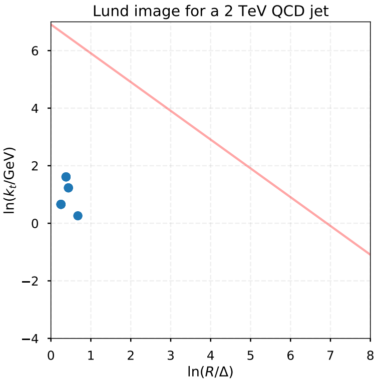
- Each jet has an image associated with its primary declustering.
- For a C/A jet, Lund plane is filled left to right as we progress through declusterings of hardest branch.
- Additional information such as azimuthal angle ψ can be attached to each point.

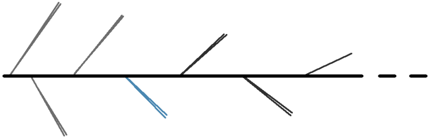




Lund representation of a jet

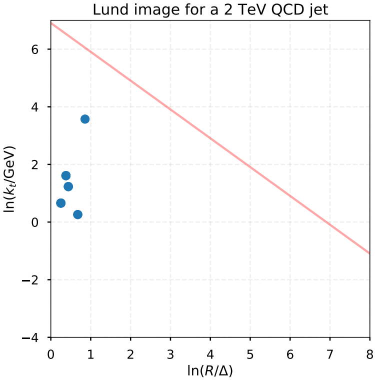
- Each jet has an image associated with its primary declustering.
- For a C/A jet, Lund plane is filled left to right as we progress through declusterings of hardest branch.
- Additional information such as azimuthal angle ψ can be attached to each point.

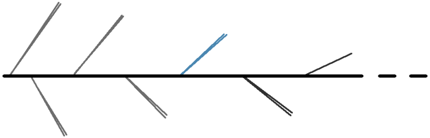




Lund representation of a jet

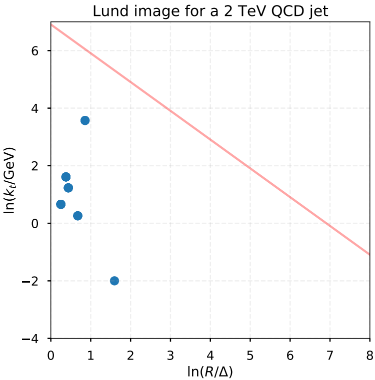
- Each jet has an image associated with its primary declustering.
- For a C/A jet, Lund plane is filled left to right as we progress through declusterings of hardest branch.
- Additional information such as azimuthal angle ψ can be attached to each point.

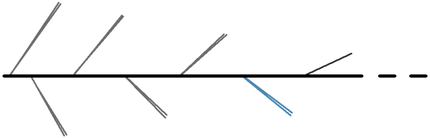




Lund representation of a jet

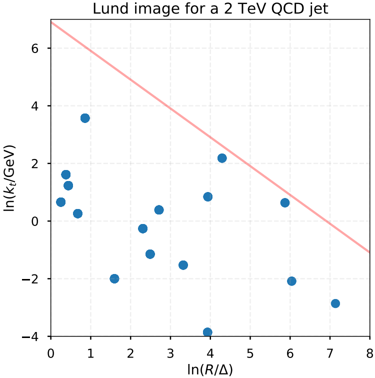
- Each jet has an image associated with its primary declustering.
- For a C/A jet, Lund plane is filled left to right as we progress through declusterings of hardest branch.
- Additional information such as azimuthal angle ψ can be attached to each point.

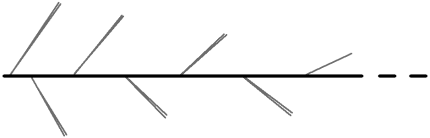




Lund representation of a jet

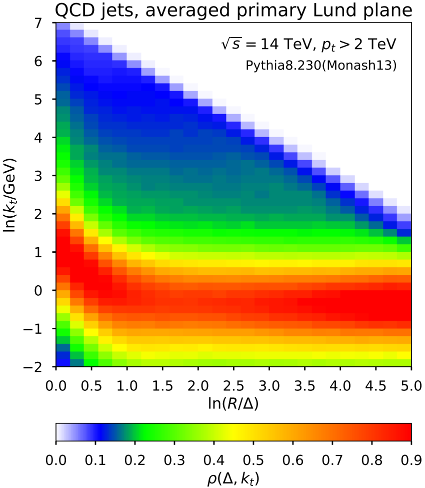
- Each jet has an image associated with its primary declustering.
- For a C/A jet, Lund plane is filled left to right as we progress through declusterings of hardest branch.
- Additional information such as azimuthal angle ψ can be attached to each point.

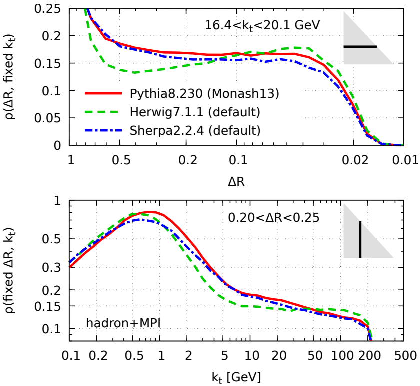




Jets as Lund images

Average over declusterings of hardest branch for 2 TeV QCD jets.

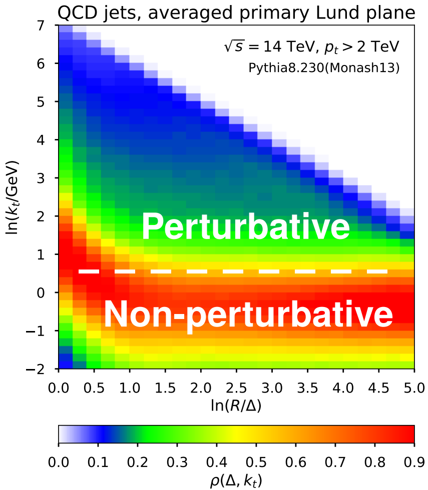


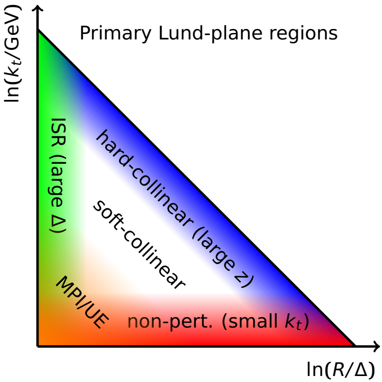


$$\rho \sim 2C_F \frac{\alpha_s(k_t)}{\pi}$$

Jets as Lund images

Average over declusterings of hardest branch for 2 TeV QCD jets.

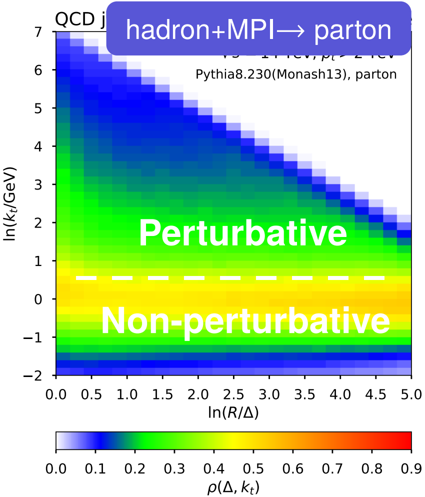


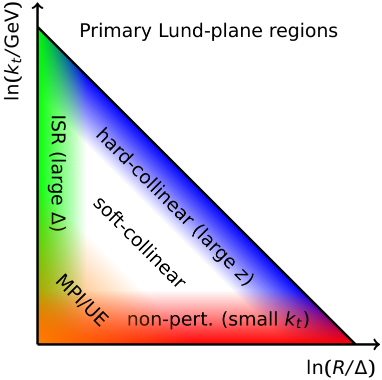


Non-perturbative region clearly separated from perturbative one.

Average over declusterings of hardest branch for 2 TeV QCD jets.

hadron+MPI⟶ parton

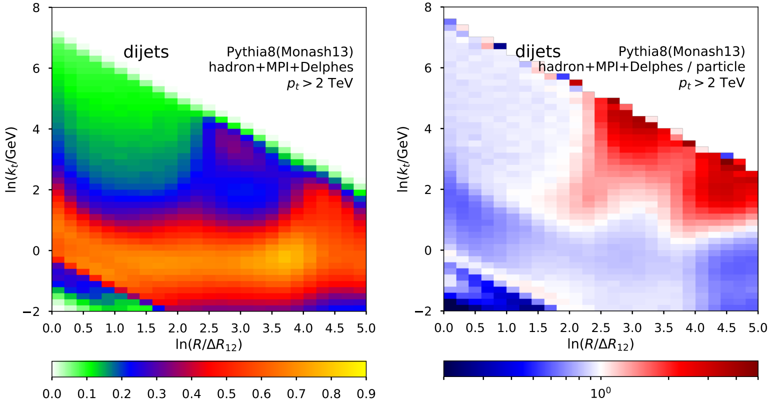




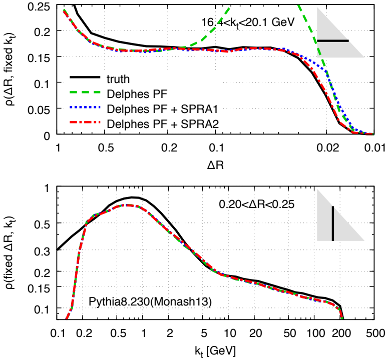
Non-perturbative region clearly separated from perturbative one.

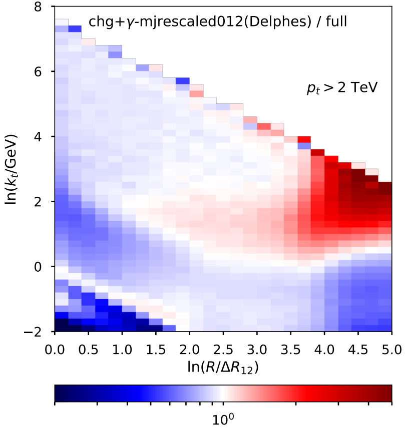
Detector effects

- Detector effects have significant impact on the Lund plane at angular scales below the hadronic calorimeter spacing.
- Two enhanced regions corresponding to resolution scale of HCal and ECal.



Mitigate impact of detector granularity using a subjet particle rescaling algorithm:

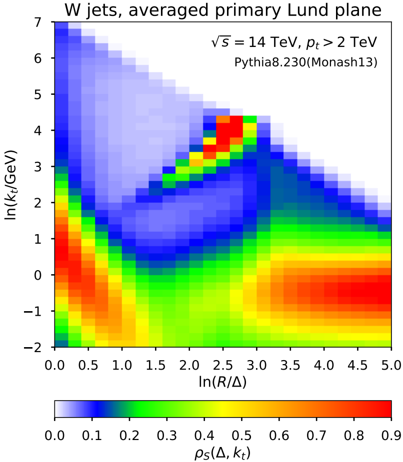




APPLICATION TO BOOSTED W TAGGING

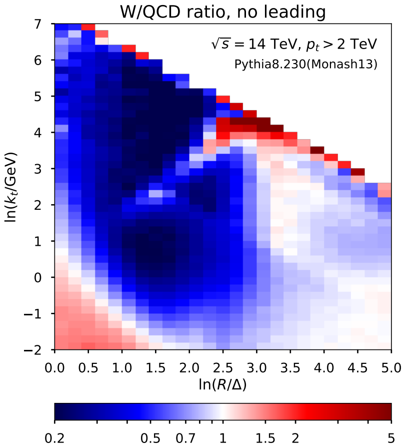
Lund images for QCD and W jets

- Hard splittings clearly visible, along the diagonal line with jet mass $m = m_W$.
- Depletion of events around W peak due to shadow cast by leading emission.



Lund images for QCD and W jets

- Hard splittings clearly visible, along the diagonal line with jet mass $m = m_W$.
- Depletion of events around W peak due to shadow cast by leading emission.



Tagging jets in the Lund Plane

We will now investigate the potential of the Lund plane for boosted-object identification.

Two different approaches:

- A log-likelihood function constructed from a leading emission and non-leading emissions in the primary plane.
- Use the Lund plane as input for a variety of Machine Learning methods.

As a concrete example, we will take dijet and WW events, looking at CA jets with $p_t > 2$ TeV.

Log-likelihood use of Lund Plane

Log-likelihood approach takes two inputs:

- First one obtained from the “leading” emission, defined as first emision satisfying $z > 0.025$ ($\sim$ mMDT tagger).

$$\mathcal{L}_\ell(m,z) = \ln\left(\frac{1}{N_S}\frac{dN_S}{dm\,dz}\bigg/\frac{1}{N_B}\frac{dN_B}{dm\,dz}\right)$$

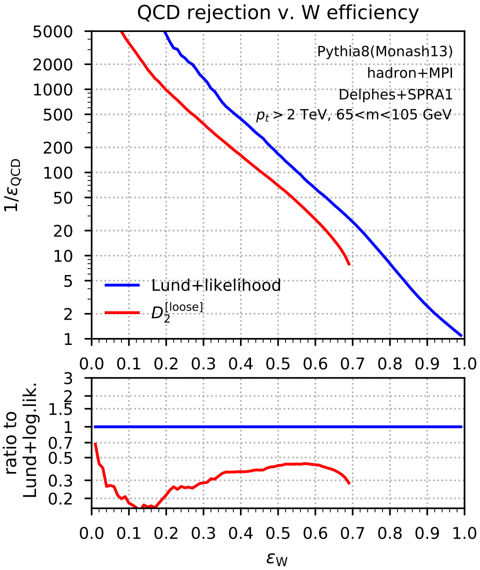
- The second one which brings sensitivity to non-leading emissions.

$$\mathcal{L}_{n\ell}(\Delta, k_t; \Delta^{(\ell)}) = \ln\left(\rho_S^{(n\ell)}\big/\rho_B^{(n\ell)}\right)$$

Overall log-likelihood signal-background discriminator for a given jet is then given by

$$\mathcal{L}_{\text{tot}} = \mathcal{L}_\ell(m^{(\ell)}, z^{(\ell)}) + \sum_{i\neq\ell}\mathcal{L}_{n\ell}(\Delta^{(i)}, k_t^{(i)}; \Delta^{(\ell)}) + \mathcal{N}(\Delta^{(\ell)})$$

- Compare the LL approach in specific mass-bin with equivalent results from the Les Houches 2017 report (arXiv:1803.07977).
- Substantial improvement over best-performing substructure observable.



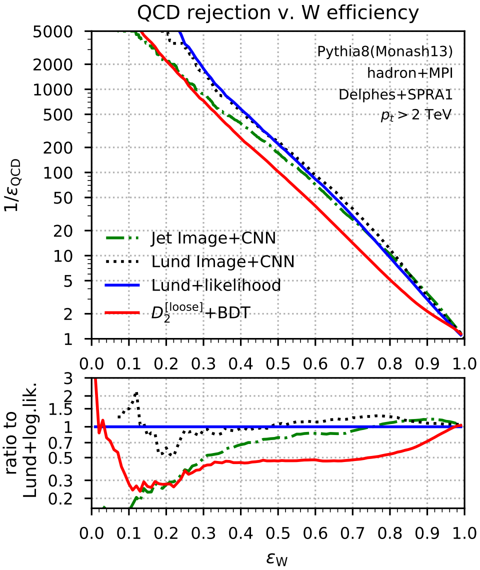
A variety of ML methods can be applied to the Lund plane in order to construct efficient taggers.

We will investigate three approaches:

- Convolutional Neural Networks (CNN) applied on 2D Lund images.
- Deep Neural Networks (DNN) applied on the sequence of declusterings.
- Long Short-Term Memory (LSTM) networks applied on the sequence of declusterings.

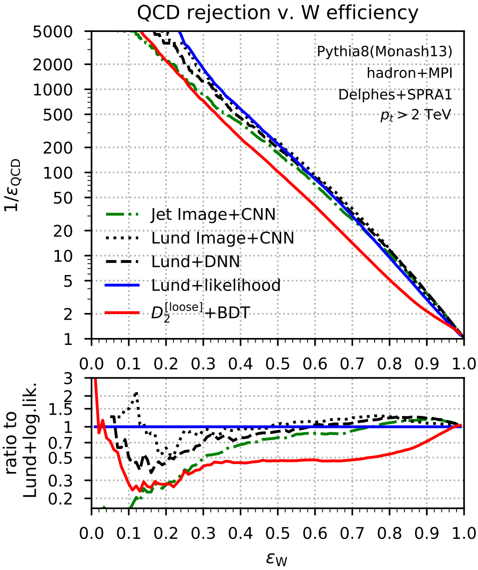
CNN on Lund images for W tagging

- Lund images perform particularly well at high transverse momentum, where $W \to qq$ is most separated on Lund plane.
- Performance on par with LL method, better than regular jet images at low efficiencies.



DNN with the Lund plane

- Applying DNN directly to the sequence of declusterings of the hardest branch.
- Inputs are IRC safe as long as there is a cutoff in transverse momentum.
- Results very similar to previous CNN approach.



Recurrent networks with a Lund plane

- Jets generally associated with a clustering trees, where each node contains similar type of information.
- Particularly well-adapted for recurrent networks, which loop over inputs and use the same weights.
- For each declustering node, we consider the inputs

$$\{\ln(R/\Delta),\, \ln(k_t/\,\mathrm{GeV})\}$$

- In practice, we will use Long Short-Term Memory (LSTM) networks, which can retain dependencies over widely separated points.

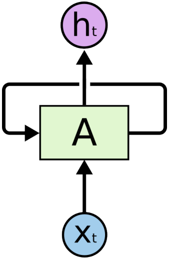
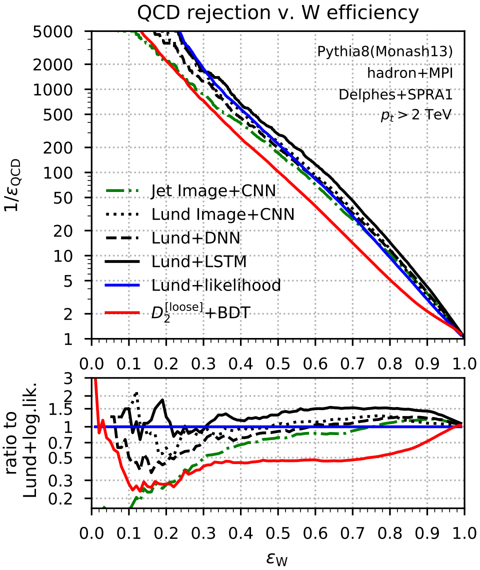


Figure from
http://colah.github.io/posts/2015-08-Understanding-LSTMs/

LSTMs for jet tagging

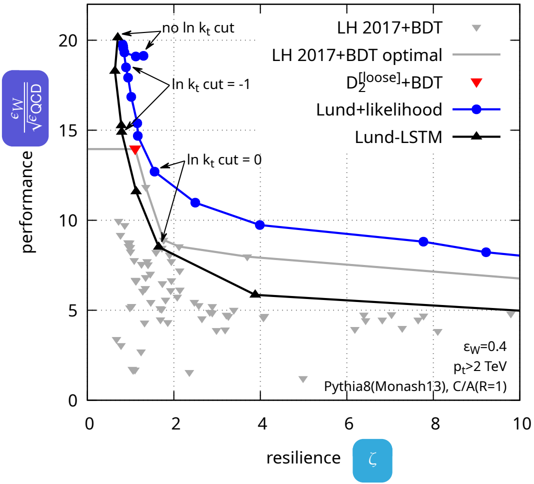
- LSTM network substantially improves on results obtained with other methods.
- Large gain in performance, particularly at higher efficiencies.



Sensitivity to non-perturbative effects

- Performance compared to resilience to MPI and hadronisation corrections.
- Vary cut on k_t, which reduces sensitivity to the non-perturbative region.

performance v. resilience [full mass information]



$\Delta\epsilon = \epsilon - \epsilon'$

$$\zeta = \left(\frac{\Delta\epsilon_S^2}{\langle\epsilon\rangle_S^2} + \frac{\Delta\epsilon_B^2}{\langle\epsilon\rangle_B^2} \right)^{-\frac{1}{2}}$$

$\langle\epsilon\rangle = \frac{1}{2}(\epsilon + \epsilon')$

(c.f. arXiv:1803.07977)

- Lund-likelihood performs well even at high resilience.
- ML approach reaches very good performance but is not particularly resilient to NP effects.

CONCLUSIONS

Conclusions

- Discussed a new way to study and exploit radiation patterns in a jet using the Lund plane.
- Lund kinematics can be used as inputs for W tagging with a range of methods:
  - Log-likelihood function.
  - Convolutional neural networks.
  - Recurrent and dense neural networks.

  Simple LL approach can match performance obtained with recent ML methods.
- While ML can achieve high performance, one needs to mindful of resilience to poorly modeled contributions and systematic uncertainties.

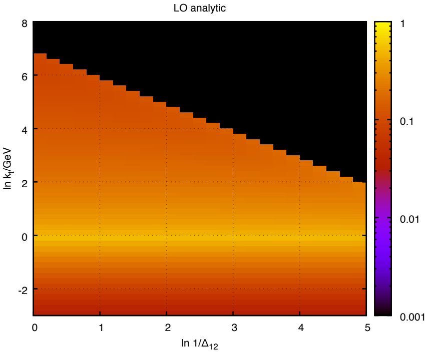
Wide range of experimental and theoretical opportunities brought by studying Lund diagrams for jets. A **rich topic** for **further exploration**.

BACKUP SLIDES

Analytic study of the Lund plane

To leading order in perturbative QCD and for $\Delta \ll 1$, one expects for a quark initiated jet

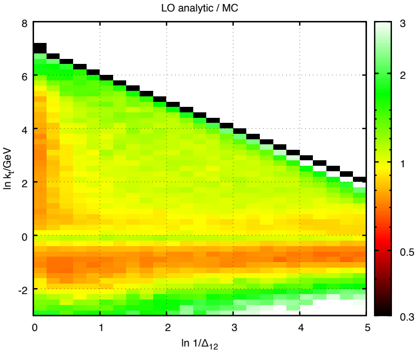
$$\rho \simeq \frac{\alpha_s(k_t) C_F}{\pi} \bar{z} \left(p_{gq}(\bar{z}) + p_{gq}(1-\bar{z}) \right) , \quad \bar{z} = \frac{k_t}{p_{t,\text{jet}} \Delta}$$



- Lund plane can be calculated analytically.
- Calculation is systematically improvable.

To leading order in perturbative QCD and for $\Delta \ll 1$, one expects for a quark initiated jet

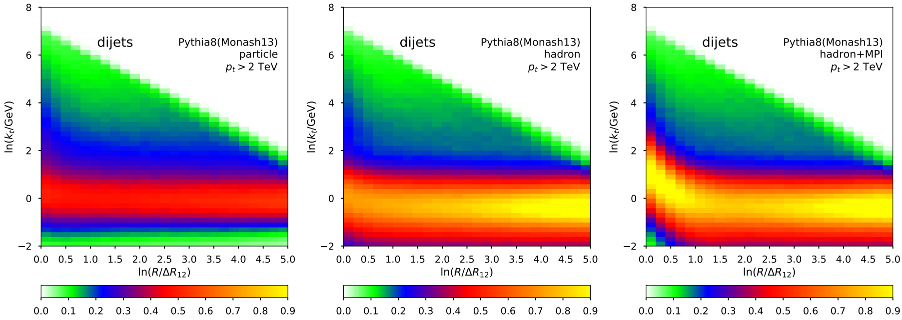
$$\rho \simeq \frac{\alpha_s(k_t) C_F}{\pi} \bar{z} \left(p_{gq}(\bar{z}) + p_{gq}(1-\bar{z}) \right), \quad \bar{z} = \frac{k_t}{p_{t,\text{jet}} \Delta}$$



- Lund plane can be calculated analytically.
- Calculation is systematically improvable.

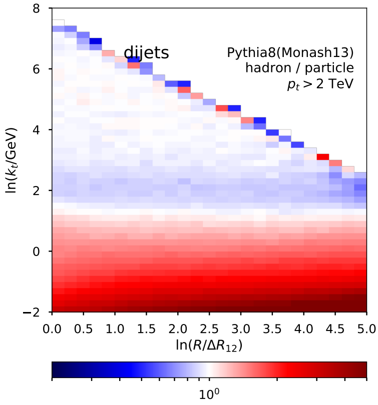
Non-perturbative effects

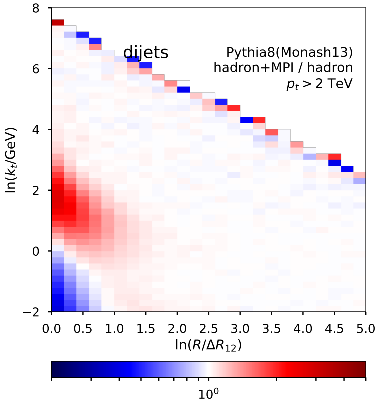
- Hadronisation corrections appear at the bottom of the Lund plane, below $\ln k_t \sim 0.5$.
- Underlying event leads to changes in the large angles region.



Non-perturbative effects

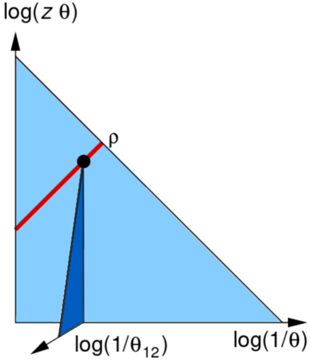
- Hadronisation corrections appear at the bottom of the Lund plane, below $\ln k_t \sim 0.5$.
- Underlying event leads to changes in the large angles region.





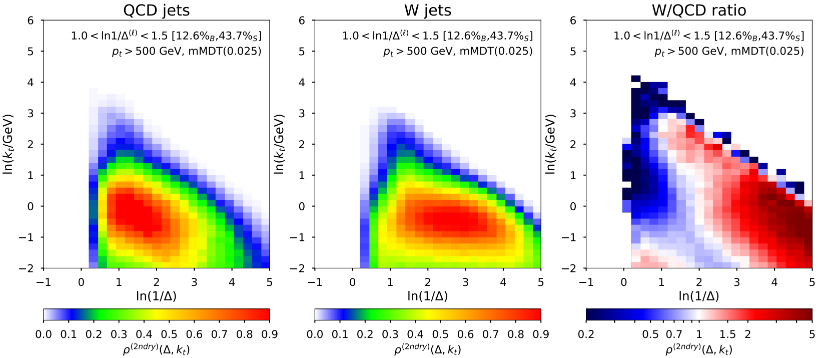
Secondary Lund plane

- Secondary Lund planes are ignored: some information is therefore lost, but still achieves good performance.
- Limitation can be overcome by extending the methods discussed to include secondary planes as inputs, which is relevant at lower p_t's.



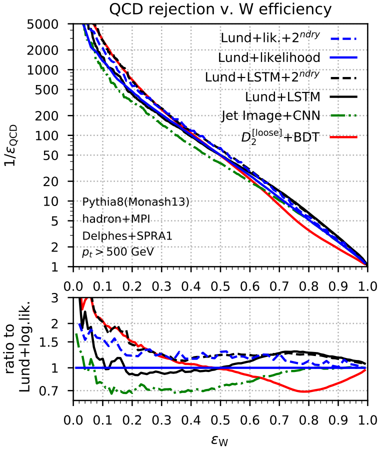
Secondary Lund plane

- Secondary Lund planes are ignored: some information is therefore lost, but still achieves good performance.
- Limitation can be overcome by extending the methods discussed to include secondary planes as inputs, which is relevant at lower p_t's.



Secondary Lund plane

- Secondary Lund planes are ignored: some information is therefore lost, but still achieves good performance.
- Limitation can be overcome by extending the methods discussed to include secondary planes as inputs, which is relevant at lower p_t's.



Subjet-Particle Rescaling Algorithm (SPRA)

Mitigate impact of detector granularity using a subjet particle rescaling algorithm:

- Recluster Delphes particle-flow objects into subjets using C/A with $R_h = 0.12$.
- Taking each subjet in turn, scale each PF charged-particle ($h^\pm$) and photon (γ) candidate that it contains by a factor f_1

$$f_1 = \frac{\sum_{i \in \text{subjet}} p_{t,i}}{\sum_{i \in \text{subjet}(h^\pm, \gamma)} p_{t,i}},$$

  and discard the other neutral hadron candidates.
- If subjet doesn't contain photon or charged-particle candidates, retain all of the subjet's particles with their original momenta.

Recluster the full set of resulting particles (from all subjets) into a single large jet and use it to evaluate the mass and Lund plane.

Log-likelihood use of Lund Plane: leading emission

Log-likelihood approach takes two inputs:

- First one obtained from the "leading" emission.
- The second one which brings sensitivity to non-leading emissions.

Leading emission is determined to be the first emission in the Lund declustering sequence that satisfies $z > 0.025$ ($\sim$ mMDT tagger)

Define a $\mathcal{L}_\ell$ log likelihood function

$$\mathcal{L}_\ell(m, z) = \ln\left(\frac{1}{N_S}\frac{dN_S}{dmdz} \bigg/ \frac{1}{N_B}\frac{dN_B}{dmdz}\right)$$

where the ratio of $\frac{dN_{S/B}}{dmdz}$ is the differential distribution in m and z of the leading emission for signal sample (background) with $N_S(N_B)$ jets.

Log-likelihood use of Lund Plane: non-leading emissions

Non-leading $(n\ell)$ emissions within the primary Lund plane are incorporated using a function

$$\mathcal{L}_{n\ell}(\Delta, k_t; \Delta^{(\ell)}) = \ln\left(\rho_S^{(n\ell)} / \rho_B^{(n\ell)}\right)$$

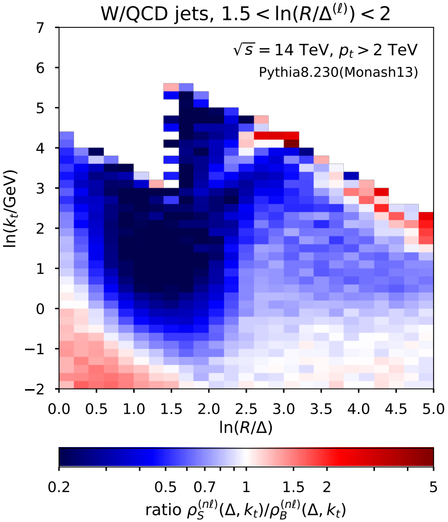
where $\rho^{(n\ell)}$ is determined just over the non-leading emissions,

$$\rho^{(n\ell)}(\Delta, k_t; \Delta^{(\ell)}) = \frac{dn^{(n\ell)}_{\text{emission}}}{d\ln k_t \, d\ln 1/\Delta \, d\Delta^{(\ell)}} \Bigg/ \frac{dN_{\text{jet}}}{d\Delta^{(\ell)}}$$

as a function of the angle $\Delta^{(\ell)}$ of the leading emission.

Log-likelihood use of Lund Plane: non-leading emissions

$\mathcal{L}_{n\ell}$ log-likelihood function in a specific bin.



Log-likelihood use of Lund Plane: full discriminator

Overall log-likelihood signal-background discriminator for a given jet is then given by

$$\mathcal{L}_{\text{tot}} = \mathcal{L}_\ell(m^{(\ell)}, z^{(\ell)}) + \sum_{i \neq \ell} \mathcal{L}_{n\ell}(\Delta^{(i)}, k_t^{(i)}; \Delta^{(\ell)}) + \mathcal{N}(\Delta^{(\ell)})$$

where $\mathcal{N} = -\int d\ln\Delta\, d\ln k_t \left(\rho_S^{(\ell)} - \rho_B^{(\ell)}\right)$.

Each subjet i in the sum brings information about whether it is in a more background-like or signal-like part of the Lund plane.

Optimal discriminator if:

- Leading emission correctly associated with W's two-prong structure.
- Non-leading emissions are independent from each other.
- Emission patterns for those emissions depend only on $\Delta^{(\ell)}$.