Probing fundamental interactions at the LHC with machine learning

Seminar in Theoretical Particle Physics, Zurich, 30 November 2021

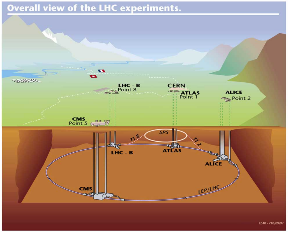
Frédéric Dreyer

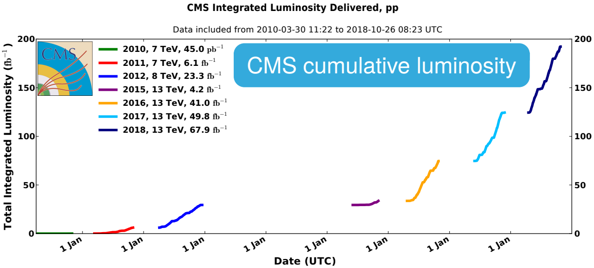
UNIVERSITY OF OXFORD

THE ROYAL SOCIETY

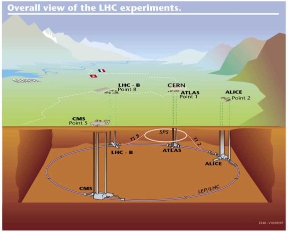
Physics at the high energy frontier

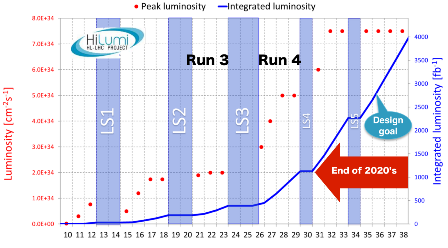
- LHC offers access to a whole qualitatively new set of interactions, Yukawas couplings, which can be probed at precision over a wide range of momenta.
- Extremely broadband new-physics search machine, with ∼1k channels across several orders of magnitude in momentum scales.
- Accurate predictions and optimized algorithms are required to make sense of noisy data spanning orders of magnitude in energy.





- LHC offers access to a whole qualitatively new set of interactions, Yukawas couplings, which can be probed at precision over a wide range of momenta.
- Extremely broadband new-physics search machine, with ∼1k channels across several orders of magnitude in momentum scales.
- Accurate predictions and optimized algorithms are required to make sense of noisy data spanning orders of magnitude in energy.

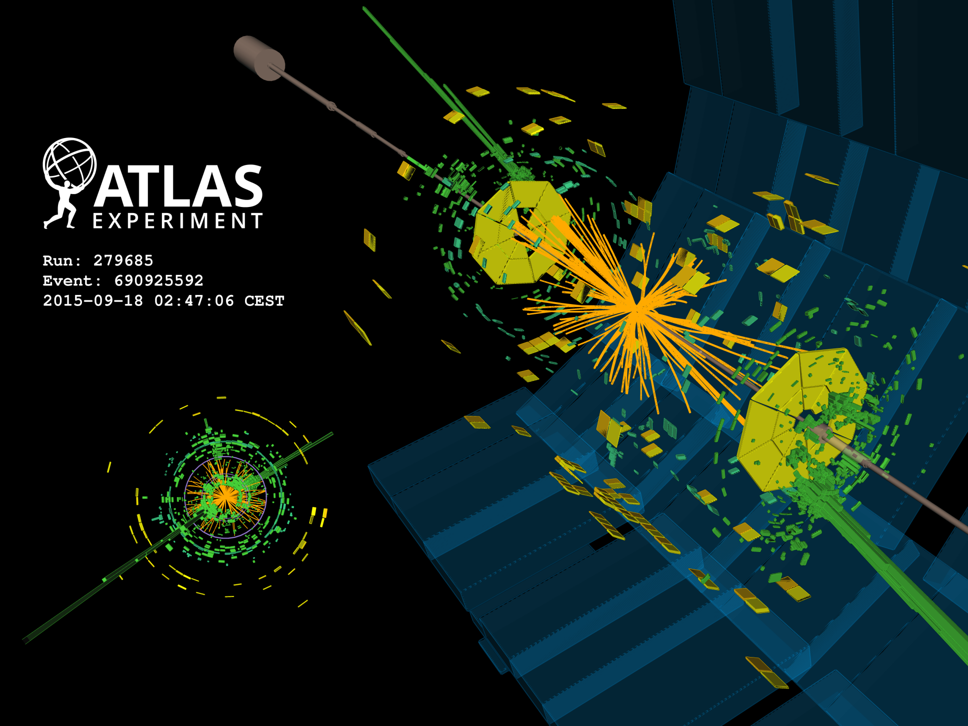




M. Lamont, O. Bruning, L. Rossi

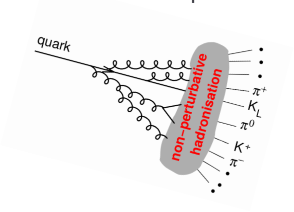
ATLAS and CMS		
Run 3	Run4	HL-LHC total
300 fb^{-1}	1 ab^{-1}	3 – 4 ab^{-1}

LHCb		
Run 3	Run4	HL-LHC total
23 fb^{-1}	50 fb^{-1}	300 fb^{-1}


ATLAS
EXPERIMENT
Run: 279685
Event: 690925592
2015-09-18 02:47:06 CEST

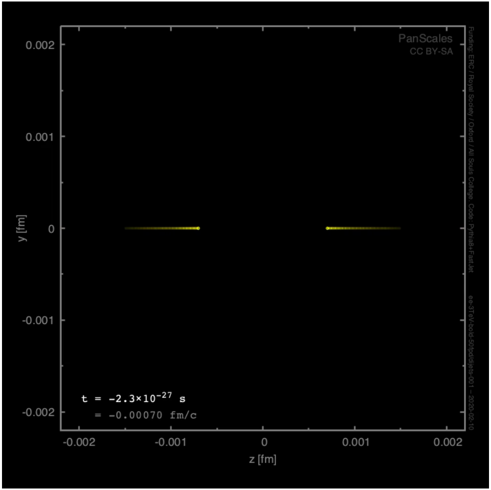
Jets as proxies for quarks and gluons

Because of color confinement, partons shower and hadronise immediately into collimated bunches of particles.



Jets are defined through a sequential recombination algorithm

$$\underbrace{\{p_i\}}_{\text{particles}} \overset{\text{jet alg.}}{\Longrightarrow} \underbrace{\{j_k\}}_{\text{jets}}$$

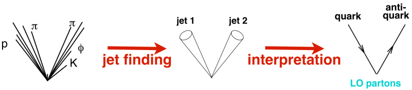


— incoming beam particle
— intermediate particle
— final particle

Jets are prevalent at hadron colliders.

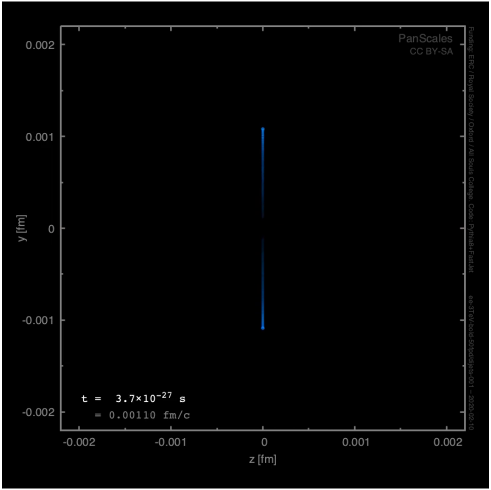
Jets as proxies for quarks and gluons

Because of color confinement, partons shower and hadronise immediately into collimated bunches of particles.



Jets are defined through a sequential recombination algorithm

$$\underbrace{\{p_i\}}_{\text{particles}} \overset{\text{jet alg.}}{\Longrightarrow} \underbrace{\{j_k\}}_{\text{jets}}$$

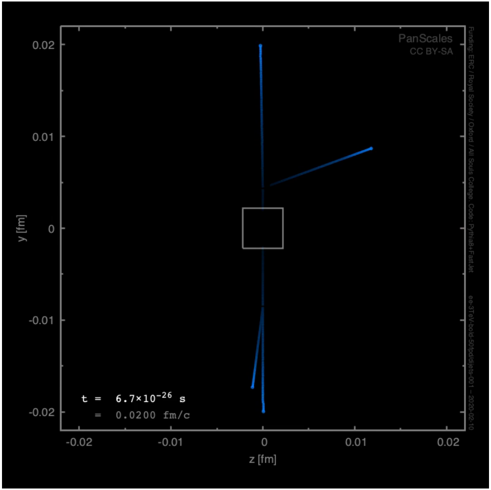


— incoming beam particle
— intermediate particle
— final particle

Jets are prevalent at hadron colliders.

Jets as proxies for quarks and gluons

Because of color confinement, partons shower and hadronise immediately into collimated bunches of particles.



Jets are defined through a sequential recombination algorithm

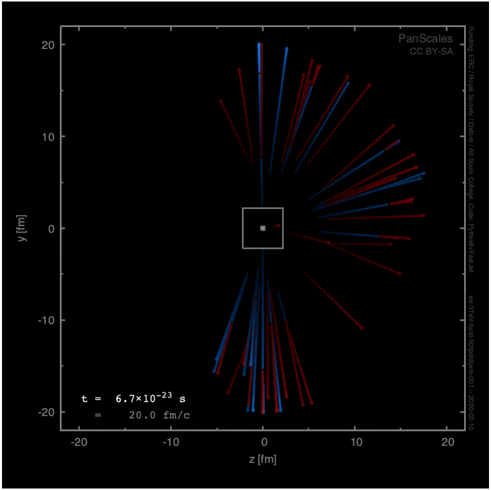
$$\underbrace{\{p_i\}}_{\text{particles}} \overset{\text{jet alg.}}{\Longrightarrow} \underbrace{\{j_k\}}_{\text{jets}}$$

— incoming beam particle
— intermediate particle
— final particle

Jets are prevalent at hadron colliders.

Jets as proxies for quarks and gluons

Because of color confinement, partons shower and hadronise immediately into collimated bunches of particles.



Jets are defined through a sequential recombination algorithm

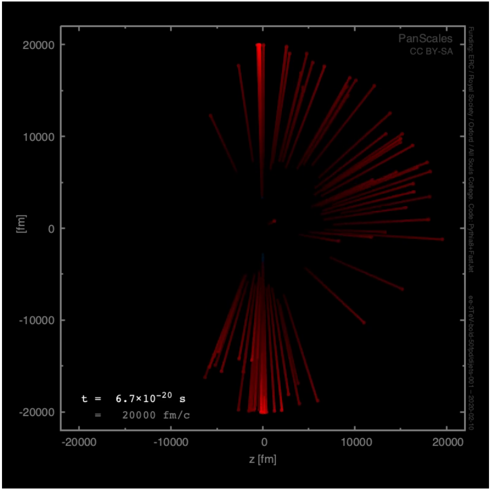
$$\underbrace{\{p_i\}}_{\text{particles}} \overset{\text{jet alg.}}{\Longrightarrow} \underbrace{\{j_k\}}_{\text{jets}}$$

— incoming beam particle
— intermediate particle
— final particle

Jets are prevalent at hadron colliders.

Jets as proxies for quarks and gluons

Because of color confinement, partons shower and hadronise immediately into collimated bunches of particles.



Jets are defined through a sequential recombination algorithm

$$\underbrace{\{p_i\}}_{\text{particles}} \overset{\text{jet alg.}}{\Longrightarrow} \underbrace{\{j_k\}}_{\text{jets}}$$

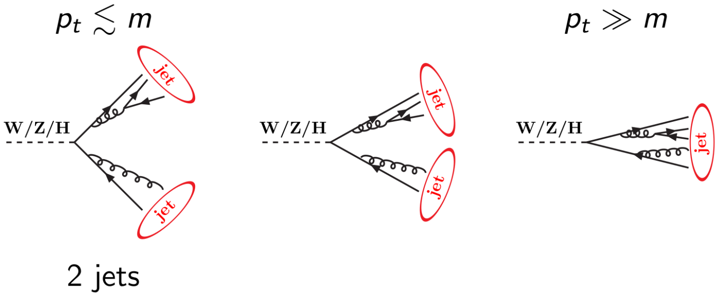
— incoming beam particle

— intermediate particle

— final particle

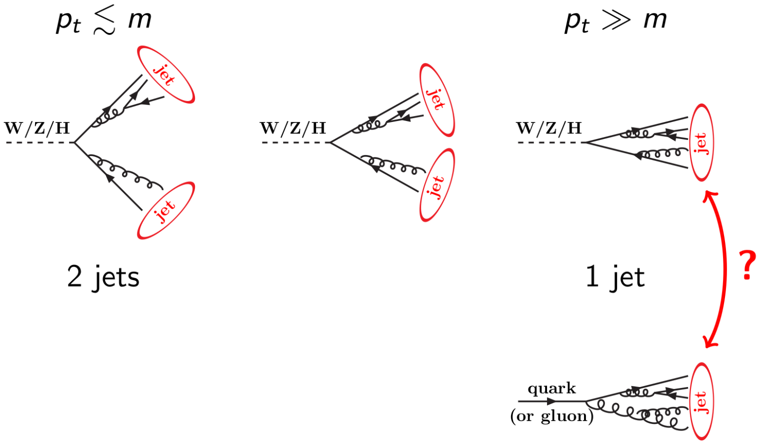
Jets are prevalent at hadron colliders.

- At LHC energies, EW-scale particles (W/Z/t…) are often produced with $p_t \gg m$, leading to collimated decays.
- Hadronic decay products are thus often reconstructed into single jets.

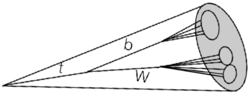


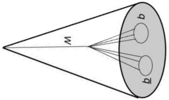
Boosted objects at the LHC

- At LHC energies, EW-scale particles (W/Z/t…) are often produced with $p_t \gg m$, leading to collimated decays.
- Hadronic decay products are thus often reconstructed into single jets.

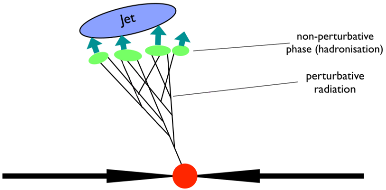


- Many techniques developed to identify hard structure of a jet based on radiation patterns.
- In principle, simplest way to identify these boosted objects is by looking at the mass of the jet.



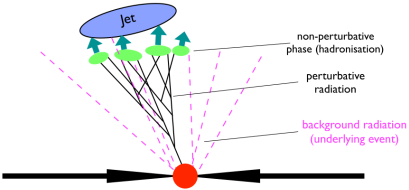


- Many techniques developed to identify hard structure of a jet based on radiation patterns.
- In principle, simplest way to identify these boosted objects is by looking at the mass of the jet.
- But jet mass distribution is highly distorted by QCD radiation and pileup.



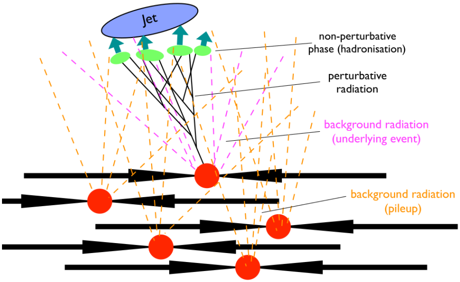
Boosted objects at the LHC

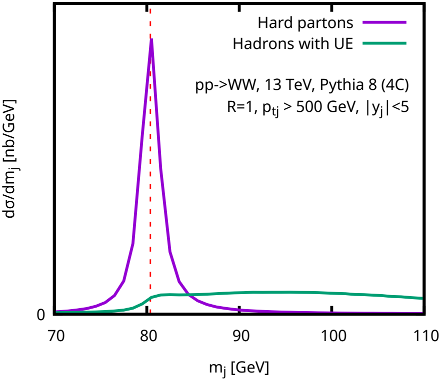
- Many techniques developed to identify hard structure of a jet based on radiation patterns.
- In principle, simplest way to identify these boosted objects is by looking at the mass of the jet.
- But jet mass distribution is highly distorted by QCD radiation and pileup.



Boosted objects at the LHC

- Many techniques developed to identify hard structure of a jet based on radiation patterns.
- In principle, simplest way to identify these boosted objects is by looking at the mass of the jet.
- But jet mass distribution is highly distorted by QCD radiation and pileup.





Identifying boosted objects

Two main avenues to study boosted decays:

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

Aim: New methods bridging the gap between these two approaches.

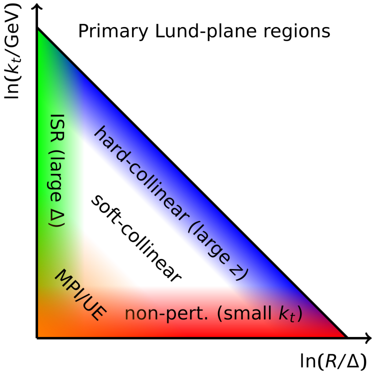
We will introduce the Lund plane representation of jets and use it as a framework to tackle a range of ML-based problems

BOOSTED OBJECTS & 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

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

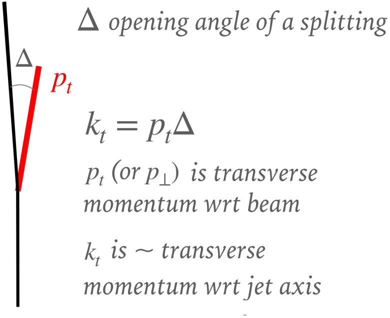


[Andersson et al, Z.Phys. C43 (1989) 625]
[**FD**, Salam, Soyez, JHEP 1812 (2018) 064]

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

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

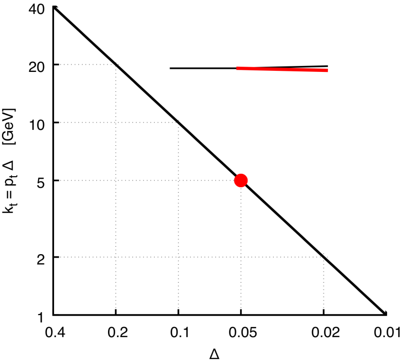


[Andersson et al, Z.Phys. C43 (1989) 625]
[**FD**, Salam, Soyez, JHEP 1812 (2018) 064]

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

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



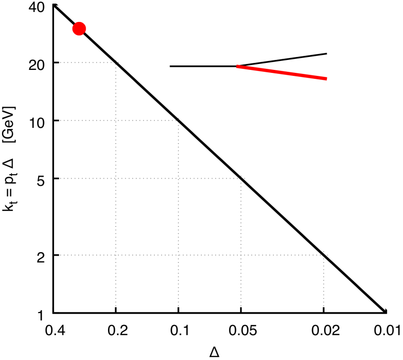
[Andersson et al, Z.Phys. C43 (1989) 625]

[**FD**, Salam, Soyez, JHEP 1812 (2018) 064]

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

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



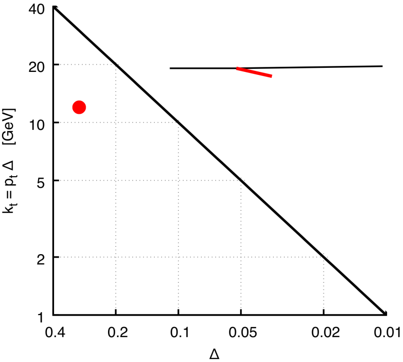
[Andersson et al, Z.Phys. C43 (1989) 625]

[**FD**, Salam, Soyez, JHEP 1812 (2018) 064]

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

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



[Andersson et al, Z.Phys. C43 (1989) 625]

[**FD**, Salam, Soyez, JHEP 1812 (2018) 064]

Lund plane representation

To create a Lund plane representation of a jet, use the (Cambridge/Aachen) clustering sequence of the jet to associate a unique Lund tree to each jet.

1. Undo the last clustering step, defining two subjets j_1, j_2 ordered in transverse momentum.
2. Save the kinematics of the current declustering step i as a tuple $\mathcal{T}^{(i)} = \{k_t, \Delta, z, m, \psi\}$

$$\Delta \equiv (y_1 - y_2)^2 + (\phi_1 - \phi_2)^2, \quad k_t \equiv p_{t2}\Delta,$$
$$m^2 \equiv (p_1 + p_2)^2, \quad z \equiv \frac{p_{t2}}{p_{t1}+p_{t2}}, \quad \psi \equiv \tan^{-1}\frac{y_2-y_1}{\phi_2-\phi_1}.$$

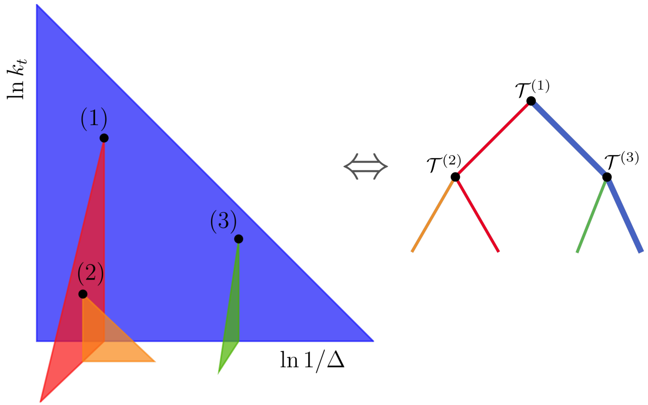
3. Repeat this procedure on both j_1 and j_2 until they are single particles.

Cambridge/Aachen clustering: pairwise recombination of particles with smallest Δ separation.

[**FD**, Salam, Soyez, JHEP 1812 (2018) 064]

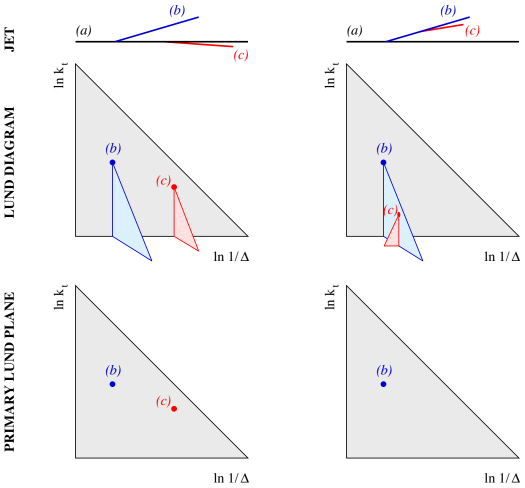
Lund plane representation

- Each jet is thus mapped onto a tree of Lund declusterings from its clustering sequence.
- Primary sequence of hardest transverse momentum branch is of particular interest for measurements and visualisation.



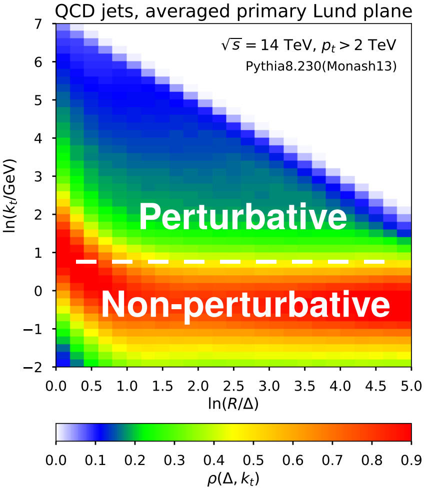
Lund plane representation

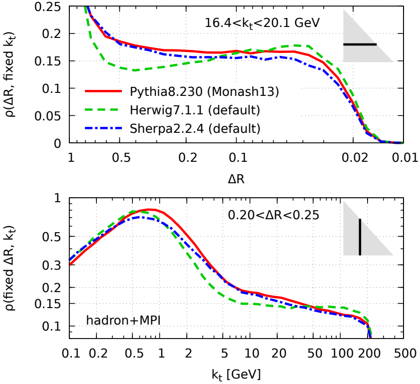
- Each jet is thus mapped onto a tree of Lund declusterings from its clustering sequence.
- Primary sequence of hardest transverse momentum branch is of particular interest for measurements and visualisation.



Jets as Lund images

Average over declusterings of hardest branch for 2 TeV jets.





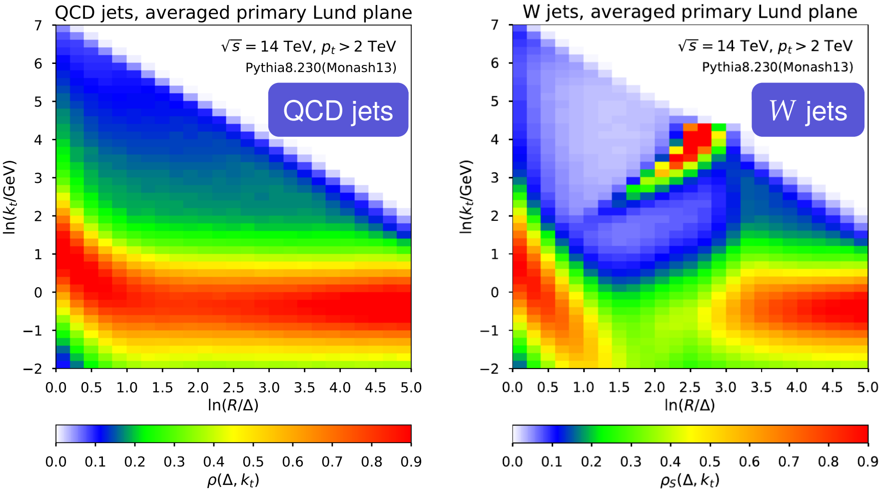
[**FD**, Salam, Soyez, JHEP 1812 (2018) 064]

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

- Non-perturbative region clearly separated from perturbative one.

Jets as Lund images

Average over declusterings of hardest branch for 2 TeV jets.

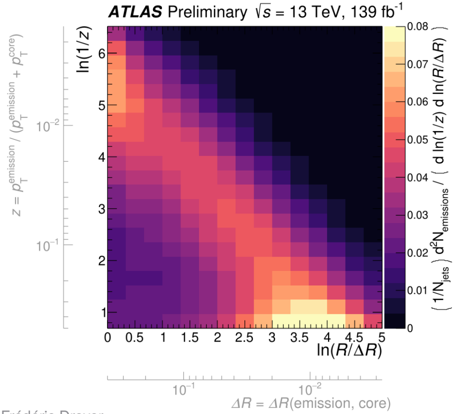


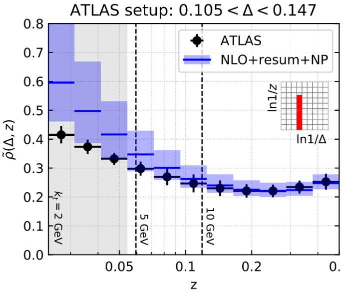
- Hard splittings visible along the diagonal line with jet mass $m = m_W$.

Measurement of the primary Lund plane

Lund images provide an opportunity for experimental measurements and comparisons with theory

- Lund plane can be predicted analytically, and the calculation is systematically improvable.
- Can be compared to data and used e.g. for α_s extractions.





[Lifson, Salam, Soyez, JHEP 10 (2020) 170]

Tagging jets in the Lund Plane

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

Several different approaches:

- A log-likelihood function constructed from a leading emission and non-leading emissions in the primary plane.
- Primary Lund plane as an input to CNN and LSTM.
- Full Lund plane as input to graph networks.

As a concrete example, we will take dijet background, with WW and $t\bar{t}$ signal events.

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).

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

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

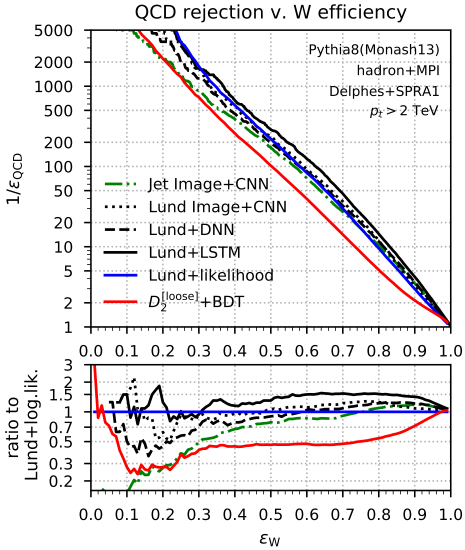
$$\mathcal{L}_{n\ell}(\Delta, k_t; \Delta^{(\ell)}) = \ln\left(\rho_S^{(n\ell)} / \rho_B^{(n\ell)}\right)$$

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

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

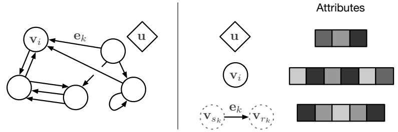
Boosted W tagging

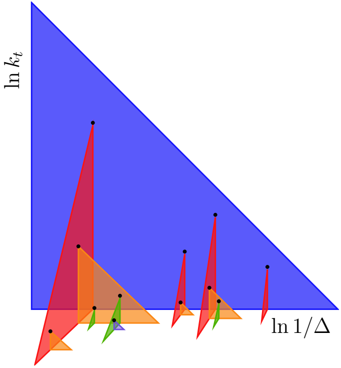
- LL approach already provides substantial improvement over best-performing substructure observable.
- LSTM network achieves even better results than those obtained with LL or older ML methods.
- Large gain in performance, particularly at higher efficiencies.



[**FD**, Salam, Soyez, JHEP 1812 (2018) 064]

- Performance can be improved further by taking secondary/tertiary Lund planes into account, particularly relevant for top tagging.
- Treat each declustering of the Lund tree as a node on a graph.





Many promising applications of graphs, e.g.

[Henrion et al. DLPS NIPS '17]

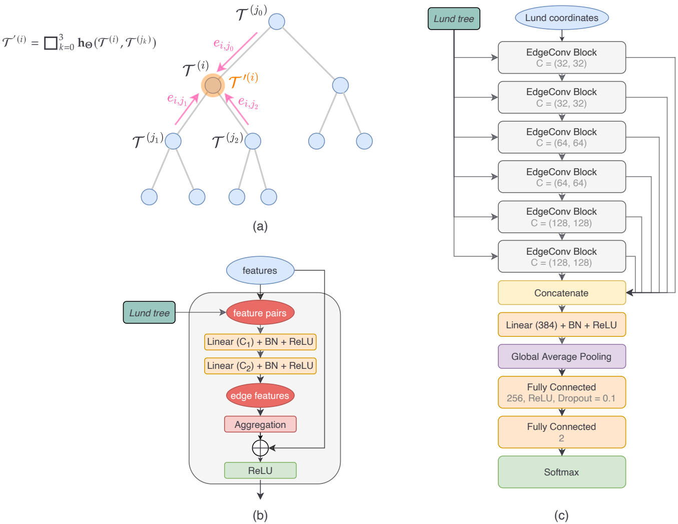
[Martinez et al. EPJP 134 (2019) 7, 333]

[Moreno et al. EPJC 80, 58 (2020)]

[Qu, Gouskos, PRD 101, 056019 (2020)]

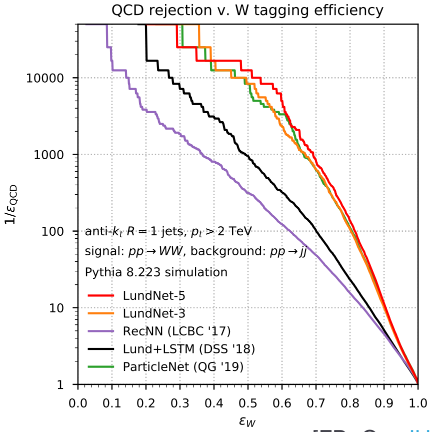
LundNet models

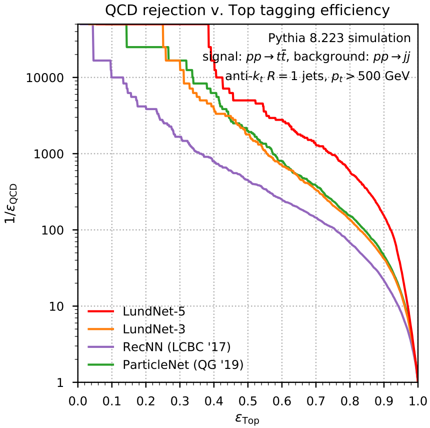
Tuple of kinematic variables as input for each node $\begin{cases} \text{LundNet-5} : (\ln k_t, \ln \Delta, \ln z, \ln m, \psi) \\ \text{LundNet-3} : (\ln k_t, \ln \Delta, \ln z) \end{cases}$



Boosted object tagging with graph networks

- Graph-based methods outperform our previous benchmarks significantly.
- LundNet model provides substantial improvement over ParticleNet and is an order of magnitude faster to train/deploy.

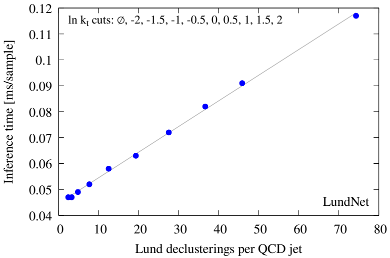




[FD, Qu, JHEP 03 (2021) 052]

Complexity of models

- Direct use of the Lund tree as the graph structure removes the need for a costly nearest-neighbour search.
- LundNet reduces training and inference time by order of magnitude compared to previous graph methods.
- Due to their higher-level kinematic inputs, LundNet takes significantly less epochs to converge to a good solution.
- Training and inference time of the model are reduced as transverse momentum cut is increased and more nodes are removed from input.



	Number of parameters	Training time [ms/sample/epoch]	Inference time [ms/sample]
LundNet	395k	0.472	0.117
ParticleNet	369k	3.488	1.036
Lund+LSTM	67k	0.424	0.131

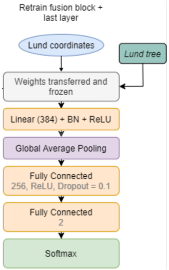
Exploiting universal features of QCD

- Universality of QCD suggests most information learnt in training process is common to different signals and experimental setups
- Can use transfer learning to develop fast and data-efficient jet taggers from existing models.

Consider two models:

- Fine-tuning: retrain all weights with a lower learning rate
- Frozen: keeping the EdgeConv frozen and retraining the final dense layers

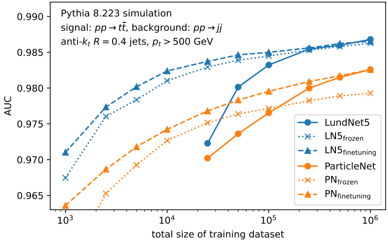
	Training time [ms/sample/epoch]	Total for 10^6 samples [hh:mm:ss]	Total for 10^5 samples [hh:mm:ss]
LundNet5	0.46	03:48:15	00:22:43
$\mathrm{LN5}_{\mathrm{frozen}}$	0.15	01:17:02	00:07:36
$\mathrm{LN5}_{\mathrm{finetuning}}$	0.46	03:48:32	00:22:45
ParticleNet	3.60	30:09:44	02:59:17
$\mathrm{PN}_{\mathrm{frozen}}$	2.16	18:13:21	01:47:37
$\mathrm{PN}_{\mathrm{finetuning}}$	3.60	29:59:46	03:01:04

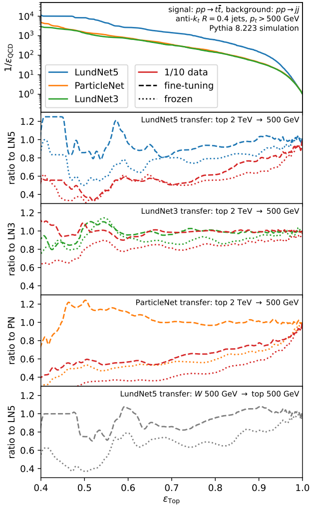


[FD, Grabarczyk, Huss, Monni, in progress]

Efficent jet traggers using transfer learning

- Both models can achieve high performance despite dramatic reduction in training data.
- Can be used to retrain existing taggers with different experimental cuts or even trained on other signals.





Reliable taggers can be obtained with an order of magnitude less data and training time

Understanding what the network is learning

Can we determine what is driving performance of a neural network?

- Consider their application on a simple task where we have first principle understanding.
- Build analytic likelihood-ratio discriminatant for this configuration and compare them with ML models.

We will consider quark/gluon discrimination.

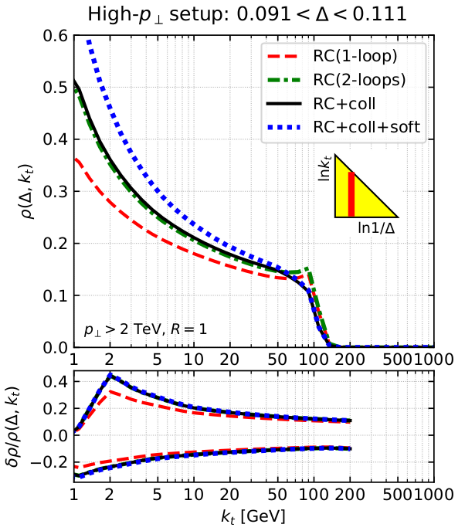
Calculating Lund plane variables

Primary Lund-plane density can be computed to single-logarithmic accuracy for both quarks and gluons.

[Lifson, Salam, Soyez, JHEP 10 (2020) 170]

For given jet with Lund declusterings $\{\Delta_i, k_{t,i}, \ldots\}$ define likelihood ratio

$$\mathbb{L}_{\text{density}} = \prod_i \frac{\rho_g(\Delta_i, k_{t,i})}{\rho_q(\Delta_i, k_{t,i})}$$



Building an analytic q/g discriminant

For a jet with primary declusterings $\{\Delta_i, k_{t,i}, z_i, \ldots\}$ compute the likelihood ratio

$$\mathbb{L}_{\text{primary}} = \frac{p_g(\{\Delta_i, k_{t,i}, z_i, \ldots\})}{p_q(\{\Delta_i, k_{t,i}, z_i, \ldots\})}$$

where $p_{q,g}(\{\Delta_i, k_{t,i}, z_i, \ldots)\}$ is the probability to observe the given set of declusterings if the jet were a quark or a gluon.

$$p_q(\{\Delta_i, k_{t,i}, z_i, \ldots\}) = p^{(\text{final})}(q|q_0) + p^{(\text{final})}(g|q_0)$$
$$p_g(\{\Delta_i, k_{t,i}, z_i, \ldots\}) = p^{(\text{final})}(q|g_0) + p^{(\text{final})}(g|g_0)$$

We can compute all single-logarithms from running coupling and collinear effects.

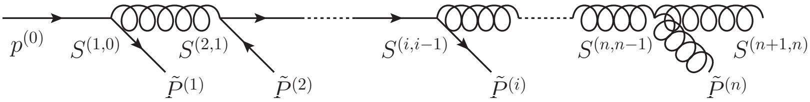
Optimal discriminant at single-logarithmic accuracy

- Computation in the collinear limit where Lund declusterings are strongly ordered in angle $\Delta_1 \gg \Delta_2 \gg \cdots \gg \Delta_n$.
- Construct the quark & gluon probability distribution iteratively from first splitting.

Probabilities after including all Lund declusterings expressed as

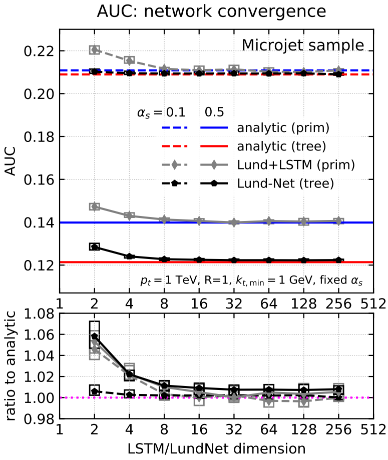
$$p^{(\text{final})} = S^{n+1,n}\tilde{P}^{(n)}S^{n,n-1}\ldots\tilde{P}^{(i)}S^{i,i-1}\ldots\tilde{P}^{(1)}S^{1,0}p^{(0)}$$

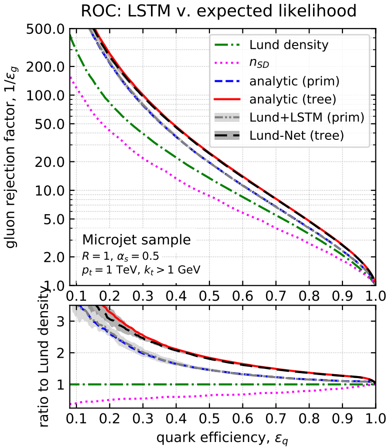
where S is a NLL Sudakov matrix and P a matrix of splitting kernels.



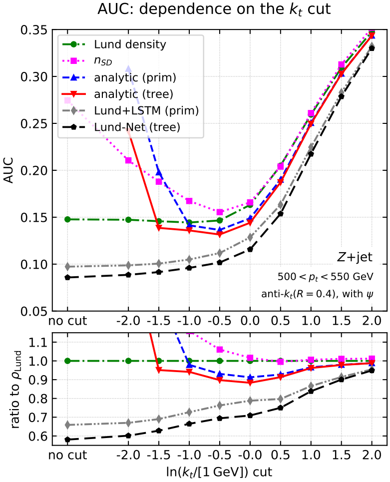
Comparison with pure-collinear parton shower

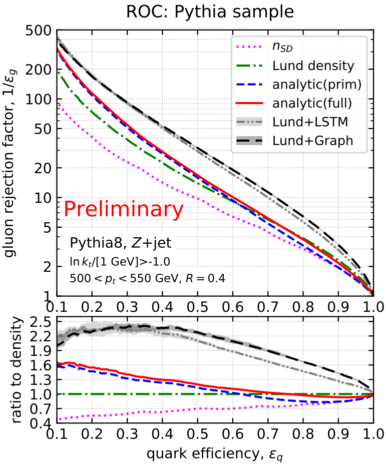
- Compare analytic and deep learning approaches in events generated in the strong-angular-ordered limit.
- In this limit analytic approach is exact and becomes optimal discriminant.





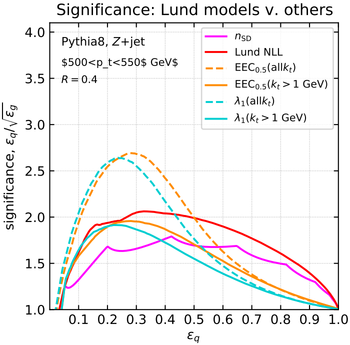
- Applying to Z+jet events generated with Pythia 8: difference in performance, but same qualitative behaviour.

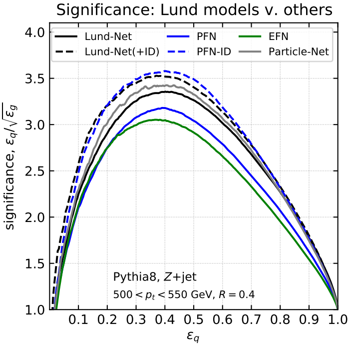




[FD, Soyez, Takacs, in progress]

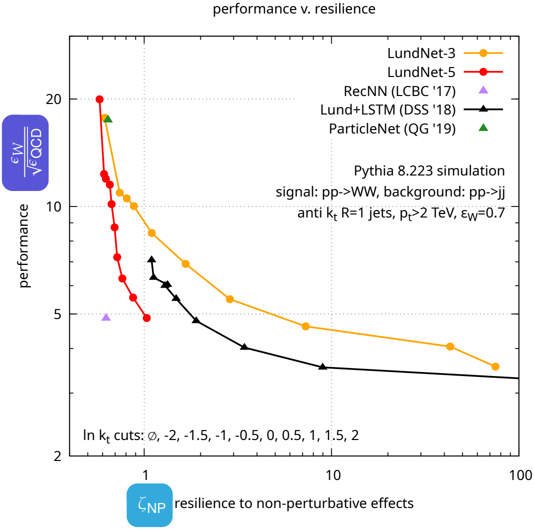
- Comparison of the Lund-plane-based approaches with other analytic and ML models.
- LundNet+ID model achieves marginally higher AUC but PFN-ID has small performance improvement at low signal efficiency.





Robustness to model-dependent effects

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



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

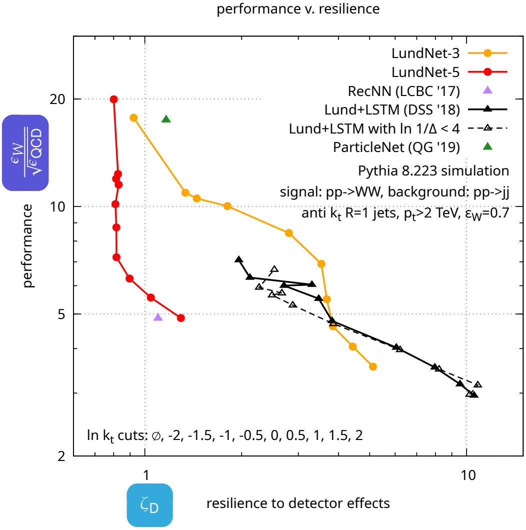
$$\zeta_{NP} = \left(\frac{\Delta\epsilon_S^2}{\langle\epsilon\rangle_S^2} + \frac{\Delta\epsilon_B^2}{\langle\epsilon\rangle_B^2} \right)^{-\frac{1}{2}}$$

$\langle\epsilon\rangle = \frac{1}{2}(\epsilon + \epsilon')$ (c.f. arXiv:1803.07977)

- LundNet-3 performs well even at high resilience.
- Most ML models can reach very good performance but are not particularly resilient to non-perturbative effects.

Robustness to model-dependent effects

- Performance compared to resilience to detector smearing effects.
- Vary Lund plane cut on k_t, which partly reduces sensitivity to detector effects.



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

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

(c.f. arXiv:1803.07977)

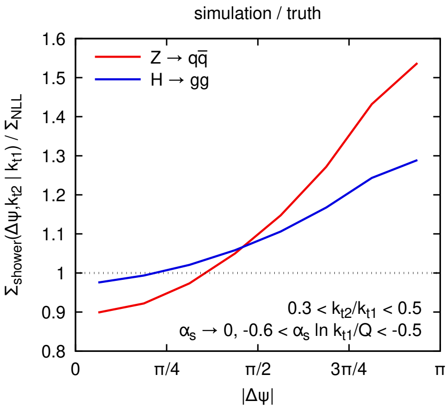
$$\langle\epsilon\rangle = \frac{1}{2}(\epsilon + \epsilon')$$

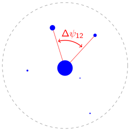
- LundNet-3 performs well even at high resilience.
- Most ML models can reach very good performance but are not particularly resilient to detector effects.

PARTON SHOWER ACCURACY

But what does the machine learn?

- Important limitation stems from the fact that labelled training data is usually obtained from Monte Carlo event generators.
- But parton shower simulations are not perfect tools!





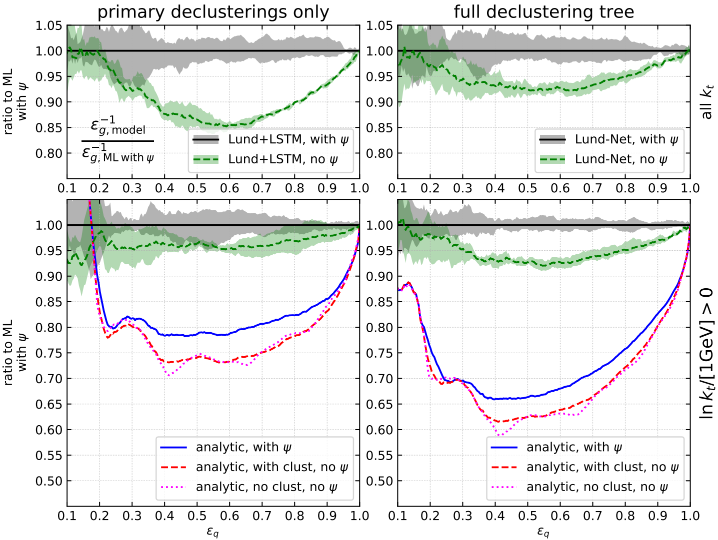
Common dipole showers display quark/gluon differences that should not be there.

- How to be sure ML models are not overfitting unphysical features?

[Dasgupta, **FD**, Hamilton, Monni, Salam, Soyez, Phys.Rev.Lett. 125 (2020) 5, 052002]

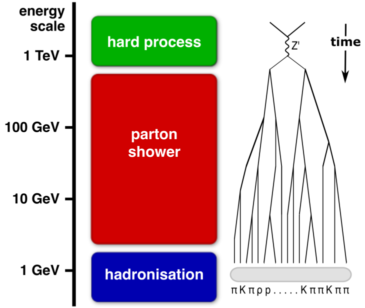
Effect of azimuthal angles

- Lund azimuthal Ψ_i angles have notable impact on discriminating power at intermediate quark efficiencies.

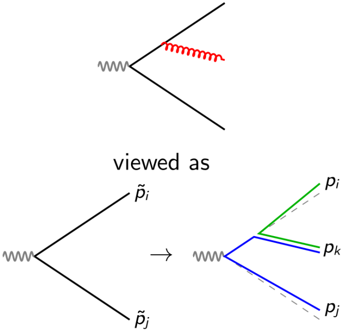


What goes into simulating a high-energy collision?

- LHC collisions probe physics across scales, from hard process at the TeV scale to non-perturbative modelling below the GeV scale.
- Parton showers span several orders of magnitude to provide crucial link between hard interaction and observable particles.
- Multi-scale evolution lead to large logarithms of ratio of scales: to what accuracy are they under control?



- Many showers are dipole/antenna showers where gluon emissions correspond to dipole splittings.
- Squared amplitudes obtained from recursive chain of emissions.



Two key ingredients:

- kinematic mapping $\tilde{p}_i, \tilde{p}_j \rightarrow p_i, p_j, p_k$.
- evolution variable v defining order of emissions.

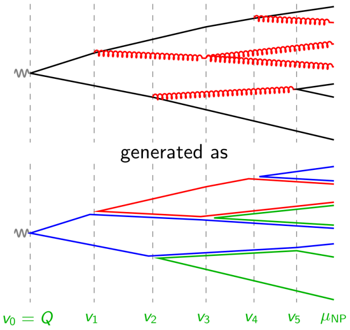
Dipole shower evolution

Evolution from state with n particles to state with $n+1$ is described by

$$\frac{d\mathcal{P}_{n\to n+1}}{d\ln v} = \sum_{\text{dipoles } \{\tilde{i},\tilde{j}\}} \int d\bar{\eta} \frac{d\phi}{2\pi} \frac{\alpha_s(k_t) + K\alpha_s^2(k_t)}{\pi}$$

$$\times \left[g(\bar{\eta}) a_k P_{\tilde{i}\to ik}(a_k) + g(-\bar{\eta}) b_k P_{\tilde{j}\to jk}(b_k) \right],$$

- v is the evolution variable (e.g. k_t in dipole c.o.m. frame)
- $g(\bar{\eta})$ is a function partitioning the dipole using the rapidity of the emission within the dipole (with $g(\bar{\eta}) + g(-\bar{\eta}) = 1$)
- $P_{\tilde{i}\to ik}(z)$ are first-order splitting functions

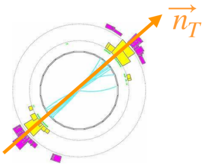


What is the accuracy of a parton shower?

- Parton showers are often referred to as leading logarithmic accurate.
- This means that it generates the correct squared amplitude in limit where both energy and angle of emissions are strongly ordered.
- Distributions can be compared to analytic resummations

For example, Thrust, defined as

$$T = \max_{\vec{n}_T} \frac{\sum_i |\vec{p}_i \cdot \vec{n}_T|}{\sum_i |\vec{p}_i|}$$



we have, for $\alpha_s L \sim 1$

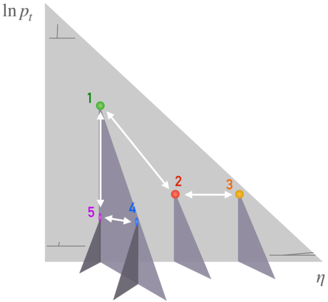
$$\sigma(1 - T < e^{-L}) = \sigma_0 \exp[\underbrace{L g_1(\alpha_s L)}_{\text{LL}} + \underbrace{g_2(\alpha_s L)}_{\text{NLL}} + \underbrace{\alpha_s g_3(\alpha_s L)}_{\text{NNLL}} + \dots]$$

What is the accuracy of a parton shower?

- Are existing dipole showers strictly LL accurate for all observables or better in some contexts?
- For what observables do we achieve a given accuracy with a given parton shower?
- Can we design a parton shower that can systematically achieve NLL accuracy for broad range of observables?
  - global event shapes (Thrust, jet rates, angularities, broadening, ...)
  - non-global observables (e.g. energy in a rapidity slice)
  - multiplicity

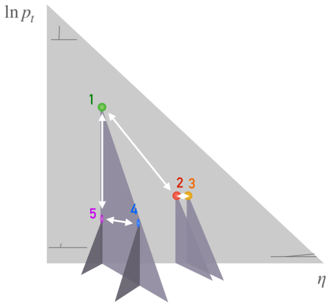
Achieving NLL accuracy

- NLL accuracy requires that the shower generates correct squared amplitude in a limit where every pair of emissions is strongly ordered for at least one logarithmic variable k_t and θ.
- I.e., should reproduce correct effective matrix element squared when all emissions are well separated in Lund diagram ($d_{12}, d_{23}, \cdots \gg 1$)



Achieving NLL accuracy

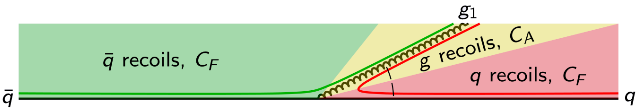
- NLL accuracy requires that the shower generates correct squared amplitude in a limit where every pair of emissions is strongly ordered for at least one logarithmic variable k_t and θ.
- I.e., should reproduce correct effective matrix element squared when all emissions are well separated in Lund diagram ($d_{12}, d_{23}, \cdots \gg 1$)
- allowed to make $O(1)$ mistake when pair of emissions is close ($d_{23} \sim 1$)



Ingredients of a shower

There are two key ingredients in the design of a parton shower

- How to associate colour and transverse recoil to dipoles?



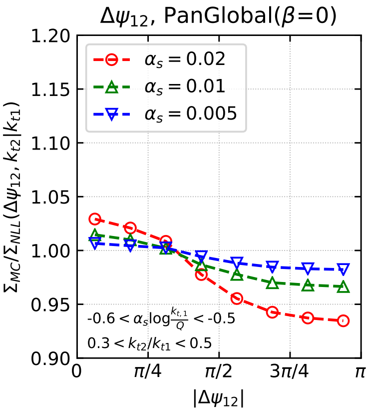
- The choice of evolution variable (transverse momentum, angle, . . .)

Design two new showers with different recoil:

- PanLocal uses ordering variable intermediate between transverse momentum and angle, partitioning dipole in event c.o.m. frame.
- PanGlobal uses k_t ordering but defines global recoil scheme, with longitudinal recoil handled by dipole-local map.

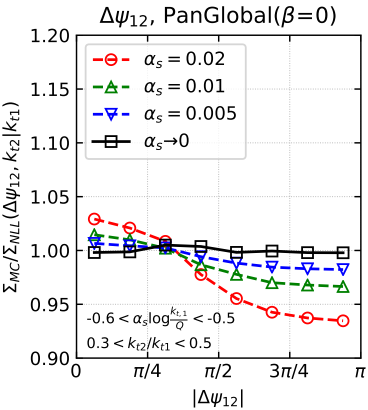
How to probe the accuracy of a shower

- Run full shower for smaller and smaller values of α_s, keeping $\alpha_s L$ constant
- Ratio to NLL of each distribution deviates from one: because of residual NNLL term or because of NLL mistake?

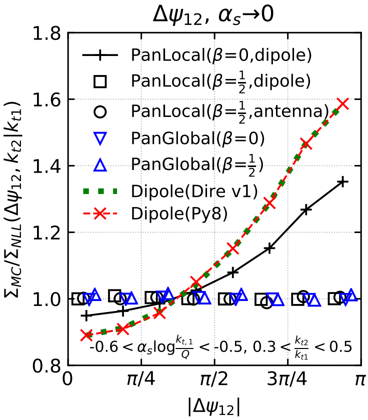


How to probe the accuracy of a shower

- Run full shower for smaller and smaller values of α_s, keeping $\alpha_s L$ constant
- Ratio to NLL of each distribution deviates from one: because of residual NNLL term or because of NLL mistake?
- Extrapolation $\alpha_s \to 0$ proves agreement with NLL, here for $\Delta\psi_{12}$ observable considered earlier

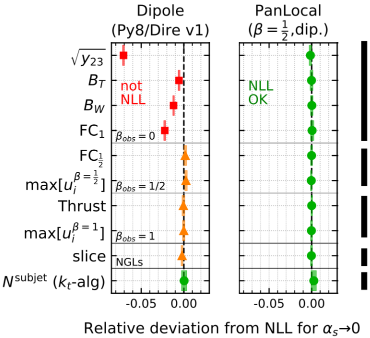


- Run full shower for smaller and smaller values of α_s, keeping $\alpha_s L$ constant
- Ratio to NLL of each distribution deviates from one: because of residual NNLL term or because of NLL mistake?
- Extrapolation $\alpha_s \to 0$ proves agreement with NLL, here for $\Delta\psi_{12}$ observable considered earlier



standard parton showers

new "PanScales" parton showers, designed specifically to achieve NLL accuracy



Event shapes sensitive to transverse momentum (jet broadenings, jet clustering transitions)

Event shapes that probe $p_t e^{-0.5|\eta|}$ (like $\beta = 0.5$ ordering variable)

Event shapes like thrust

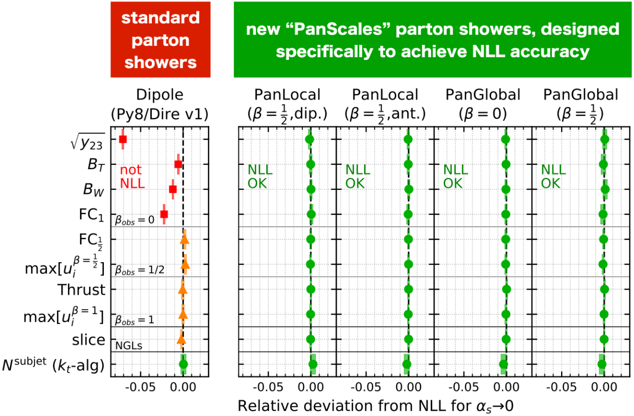
probe of non-global logarithms

standard jet multiplicity (probe of full recursive shower structure)

[Dasgupta, **FD**, Hamilton, Monni, Salam, Soyez, Phys.Rev.Lett. 125 (2020) 5, 052002]

Paves the way for improved simulations with more accurate physical description of perturbative radiation.

Designing new showers for precision physics



All PanScales shower that are expected to agree with NLL pass these tests

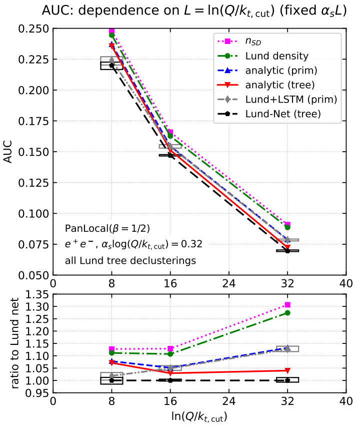
(Standard dipole showers don't)

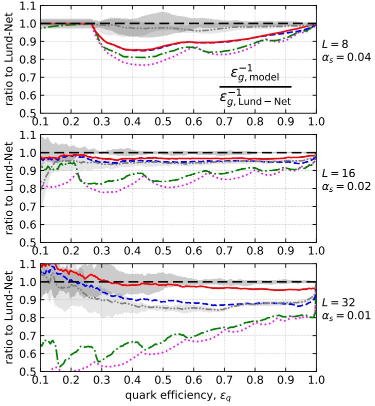
[Dasgupta, **FD**, Hamilton, Monni, Salam, Soyez, Phys.Rev.Lett. 125 (2020) 5, 052002]

Paves the way for improved simulations with more accurate physical description of perturbative radiation.

Asymptotic single-logarithmic limit

- Use training data generated with PanLocal shower and consider limit where subleading effects decrease.
- Analytic and corresponding ML models converge as $\alpha_s \rightarrow 0$.





CONCLUSIONS

Conclusions

- Higgs sector and searches for new physics requires us to understand how to relate with high precision the fundamental Lagrangian of particle physics with experimental observations.
- Exploiting available data to its fullest extent and understanding bias and limitations of machine learning models will be essential steps towards this goal.
- Combination of physical insight and machine learning can lead to substantial impact on our ability to exploit the substructure of jets for searches for new physics.