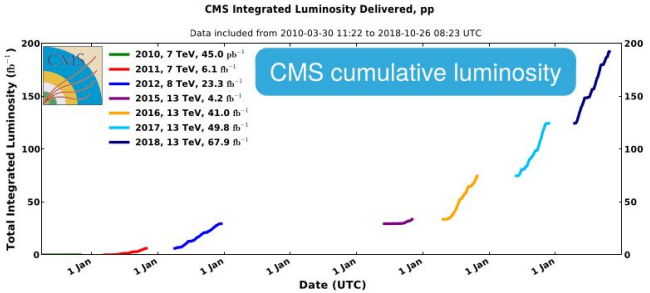
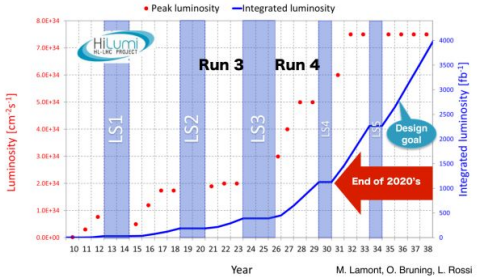
Overview of jet substructure

Frédéric Dreyer
5 July 2022

Challenges at the LHC

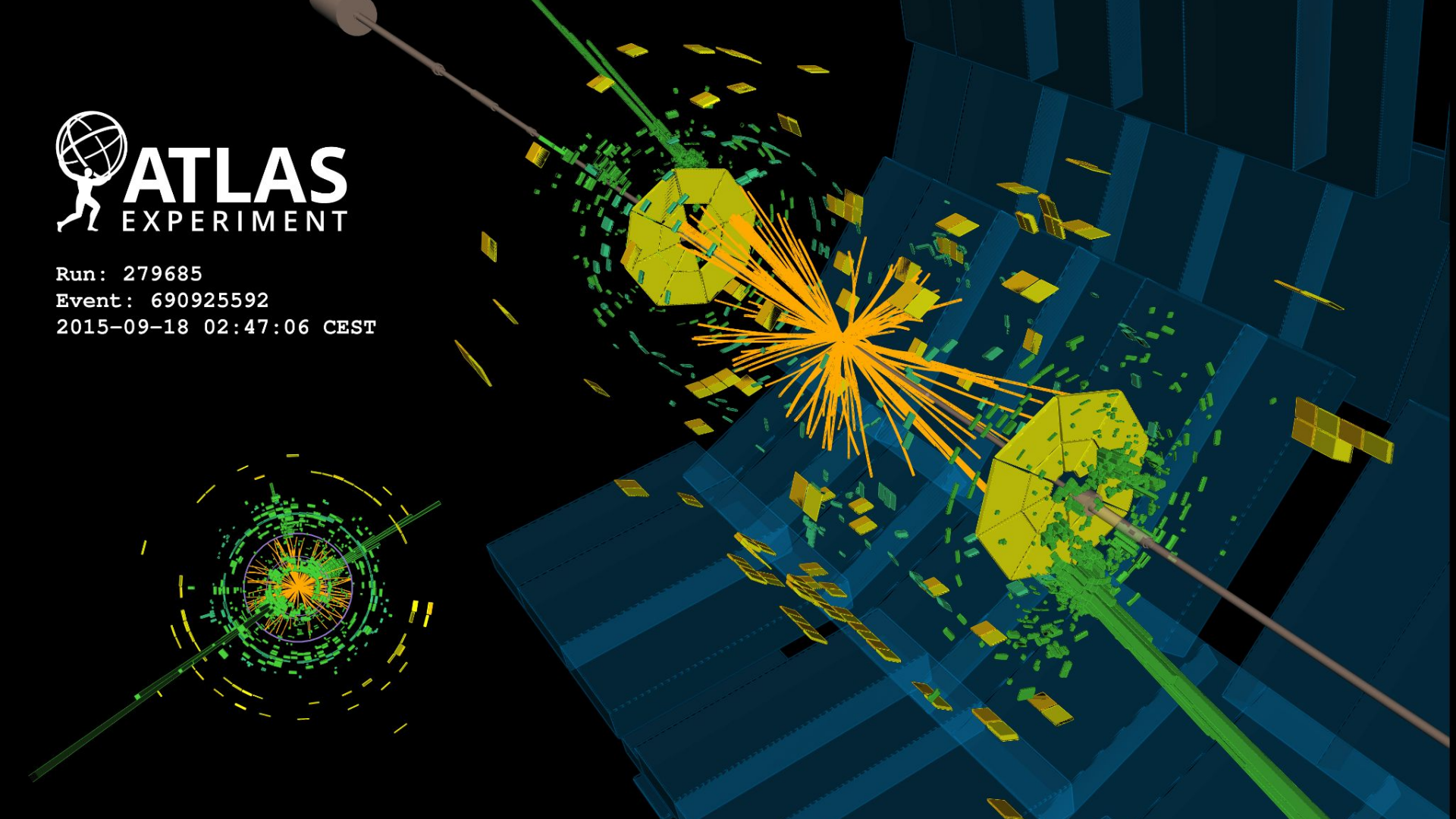
- LHC offers access to a whole qualitatively new set of interactions, Yukawas couplings, which can be probed at precision over a wide range of momenta.
- Extremely broadband new-physics search machine, with ~ 1k channels across several orders of magnitude in momentum scales.
- Accurate predictions and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





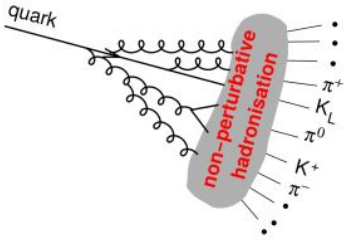
ATLAS and CMS		
Run 3	Run4	HL-LHC total
300 fb^{-1}	1 ab^{-1}	3 – 4 ab^{-1}

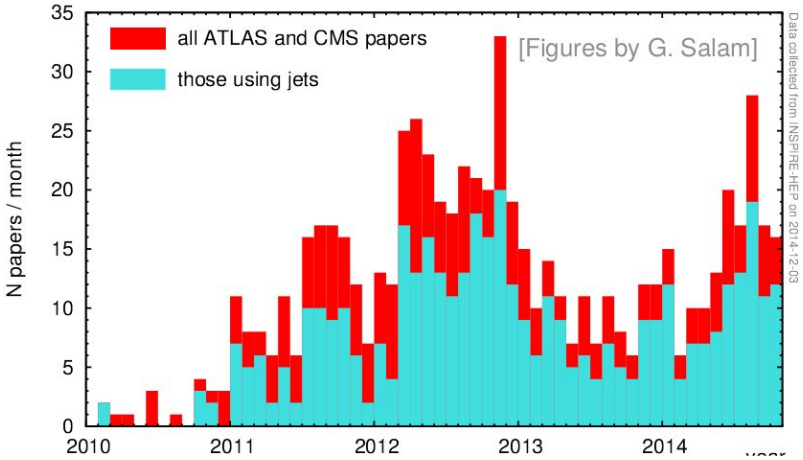
LHCb		
Run 3	Run4	HL-LHC total
23 fb^{-1}	50 fb^{-1}	300 fb^{-1}


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

Jets

- Because of color confinement, quarks and gluons shower and hadronise immediately into collimated bunches of particles.
- Hadronic jets emerge from a number of processes:
  - Scattering of partons inside colliding protons
  - Hadronic decay of heavy particles
  - Radiative gluon emission from partons
- Jets are prevalent at hadron colliders, and used in 2/3 of ATLAS and CMS analyses





Jet algorithm

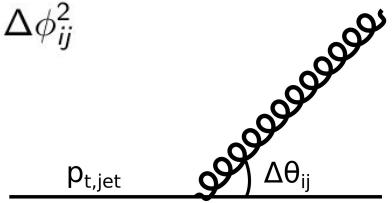
A jet algorithm maps final state **particle momenta** to **jet momenta**.

$$\underbrace{\{p_i\}}_{\text{particles}} \Longrightarrow \underbrace{\{j_k\}}_{\text{jets}}$$

Requires an external parameter, the **jet radius R**, which specifies up to which angle separate partons are recombined into a single jet.

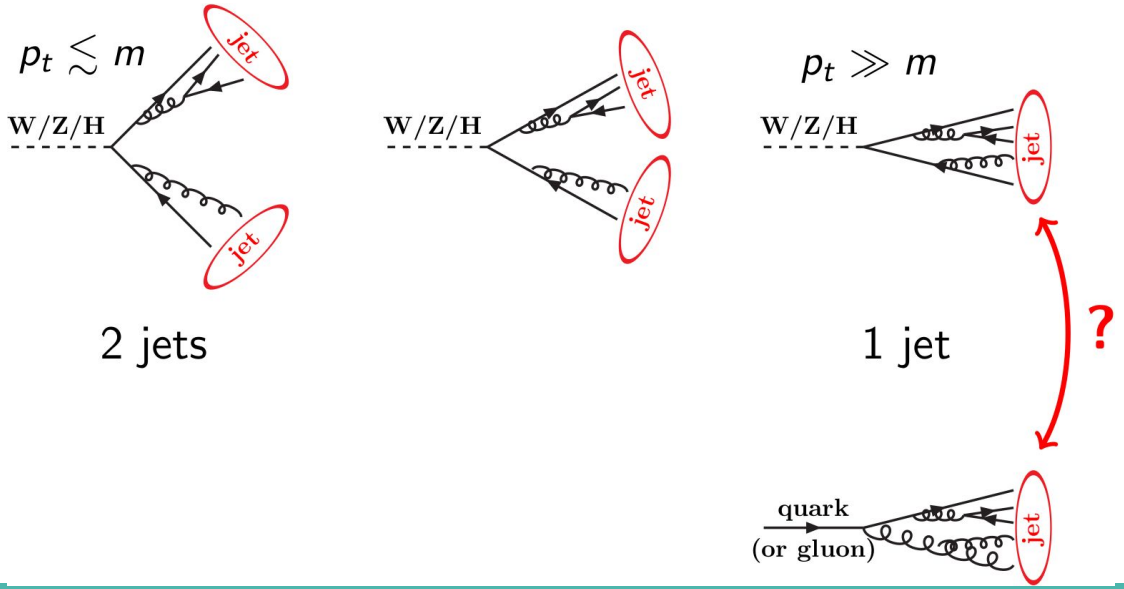
Recursively cluster particles that are closest in a metric defined by

$$d_{ij} = \min(p_{t,i}^{2p}, p_{t,j}^{2p}) \frac{\Delta\theta_{ij}^2}{R^2} \qquad \Delta\theta_{ij}^2 = \Delta y_{ij}^2 + \Delta\phi_{ij}^2$$



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.

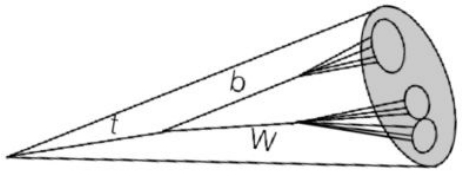


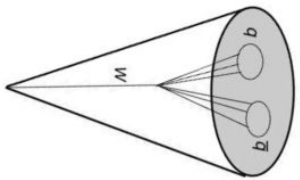
Jet radius

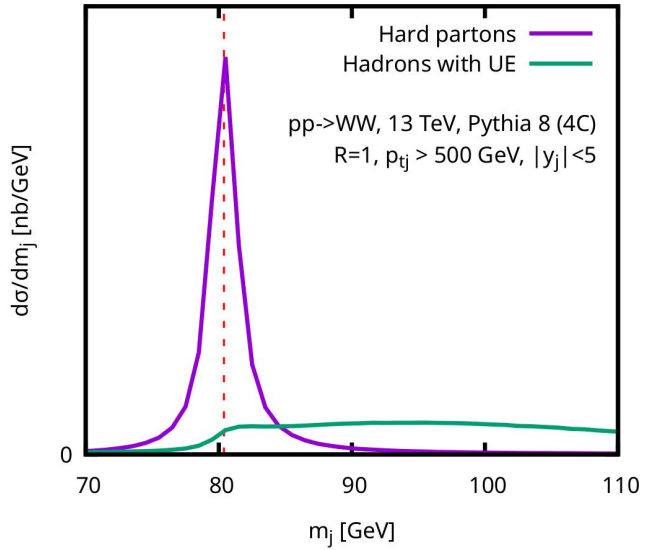
- The radius parameter in the equation roughly controls the size of the jet
  - Complementary information to the algorithm
- Typical choice for small-R jets: R=0.4 (ATLAS and CMS)
  - Used for “standard” QCD jets by most experimental analyses
  - Idea is roughly to contain a light quark or gluon in such a jet
- Typical choice for large-R jets: R=0.8 (CMS) or R=1.0 (ATLAS)
  - Used for “boosted” jets by most experimental analyses
  - Aim is to contain a hadronically decaying particle (W, Z, H, top, etc)

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.







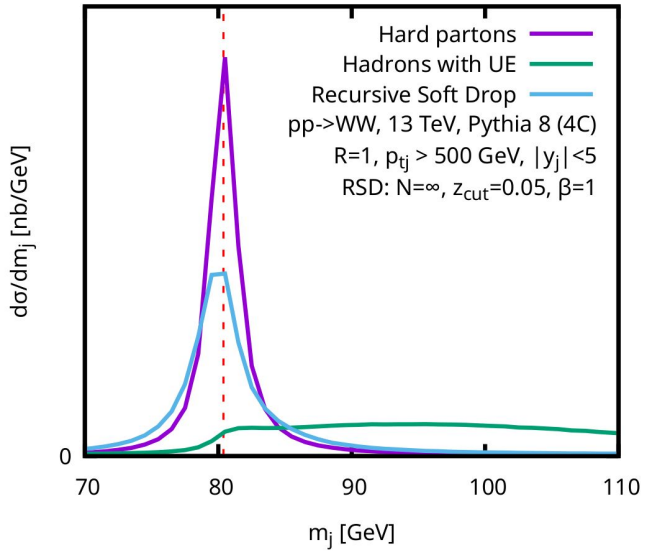
Jet grooming: (Recursive) Soft Drop/mMDT

- Mass peak can be partly reconstructed by removing unassociated soft wide-angle radiation (grooming).
- Recurse through clustering tree and remove soft branch if

$$\frac{\min(p_{t,1}, p_{t,2})}{p_{t,1} + p_{t,2}} > z_{\text{cut}} \left(\frac{\Delta R_{12}}{R_0} \right)^{\beta}$$

- Define (z_g, θ_g) pair for branch above cut

[Dasgupta, Fregoso, Marzani, Salam JHEP 1309 (2013) 029]
[Larkoski, Marzani, Soyez, Thaler JHEP 1405 (2014) 146]
[Dreyer, Necib, Soyez, Thaler JHEP 1806 (2018) 093]
[Mehtar-Tani, Soto-Ontoso, Tywoniuk Phys.Rev.D 101 (2020) 3, 034004]



Jet grooming: common tools

Trimming

- Take jet with radius R
- Reclusters components into smaller subjets with radius $R_{sub} < R$
- Keep subjets that satisfy $p_{t,\,sub} > z_{cut}\, p_{t,\,jet}$

[Krohn, Thaler, Wang, JHEP 1002 (2010) 084]

Pruning

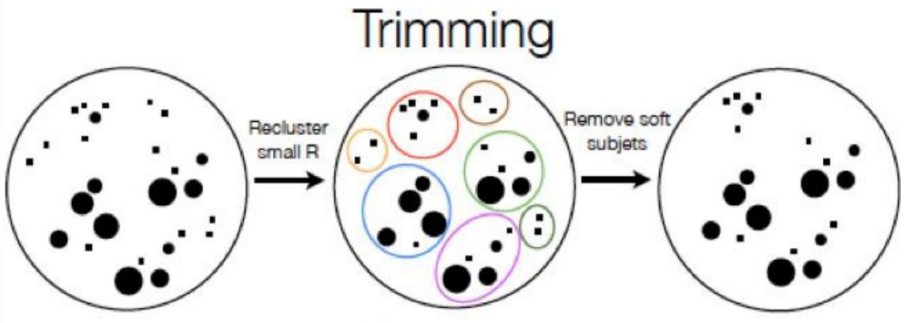
- Define pruning radius $R_{prun} = R_{cut}\ 2\ m / p_t$
- For every step of clustering $j_1 + j_2 \rightarrow j_{12}$, check:
  - Wide-angle: $\Delta R_{12} > R_{prun}$
  - Soft: $\min(p_{t1}, p_{t2}) < z_{cut}\, p_{t,\,jet}$
- If either condition fails, eliminates softer subjet
- If both pass, continue clustering

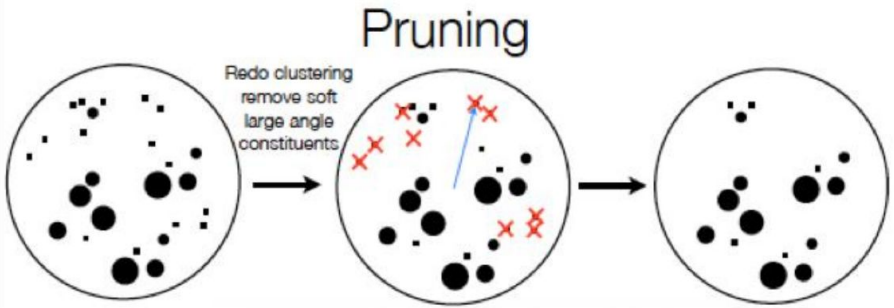
[Ellis, Vermilion, Walsh, PRD81 (2010) 094023]

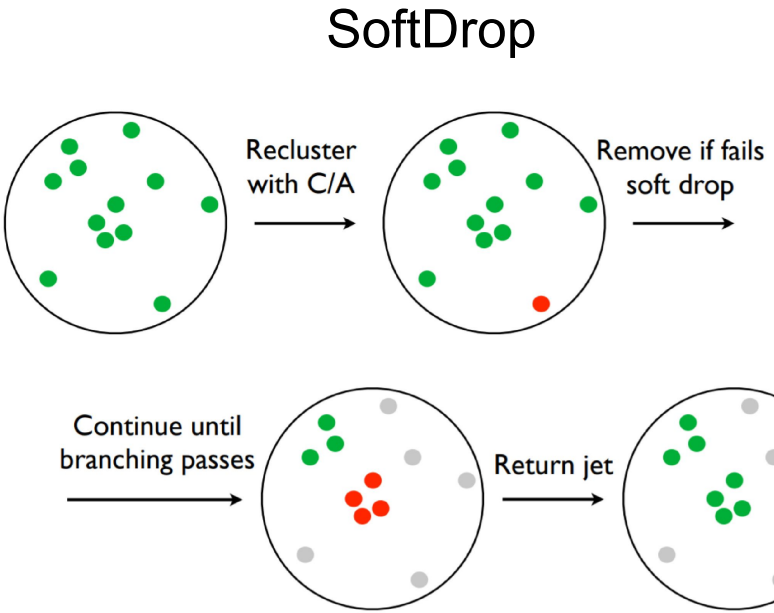
(Recursive) Soft Drop / mMDT

- Decluster jet $j_{12} \rightarrow j_1 + j_2$
- Check condition $\min(p_{t1}, p_{t2})/p_{t,jet} > z_{cut}(\Delta R_{12}/R)^{\beta}$
  - z_{cut}, β: tunable values
- If condition fails, the softer subjet is removed
- If passes, stops recursion
- For $\beta=0$, it is mMDT

[Dasgupta, Fregoso, Marzani, Salam JHEP 1309 (2013) 029]
[Larkoski, Marzani, Soyez, Thaler JHEP 1405 (2014) 146]
[Dreyer, Necib, Soyez, Thaler JHEP 1806 (2018) 093]
[Mehtar-Tani, Soto-Ontoso, Tywoniuk Phys.Rev.D 101 (2020) 3, 034004]


Trimming
Recluster
small R
Remove soft
subjets


Pruning
Redo clustering
remove soft
large angle
constituents


SoftDrop
Recluster
with C/A
Remove if fails
soft drop
Continue until
branching passes
Return jet

Groomed jet mass

- Connection between **measurements** and **calculations**
- **Jet mass** is one of the simplest observables
- **Grooming** eliminates part of UE contamination
- Modified MassDrop Tagger and SoftDrop can be calculated analytically to high precision
- **Needs to be resummed at all orders, matched to fixed-order**

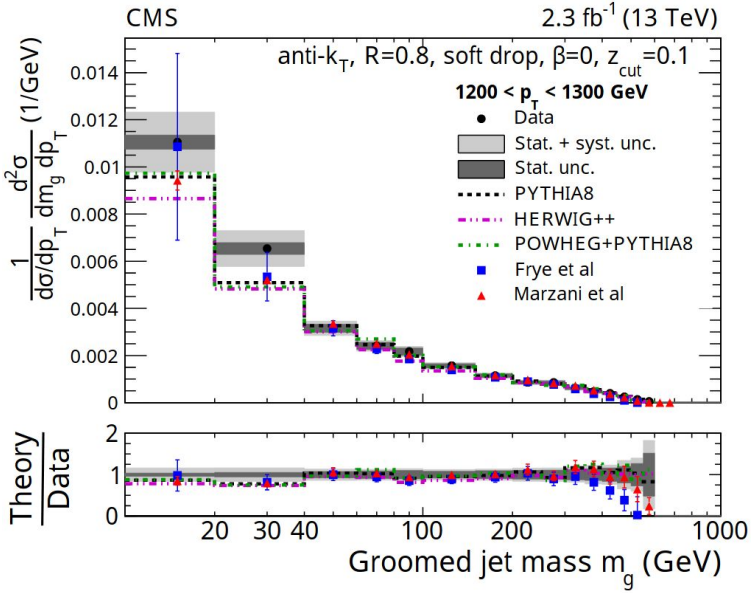
Groomed jet mass

- Various interesting QCD structures emerging
  - For mMDT it becomes $[\alpha_s\, f(z_{cut})\, \log(1/\rho)]^n$ at leading-log
  - Finite z_{cut} introduce a flavour changing matrix structure
- Compare with experiment → needs a matching procedure:

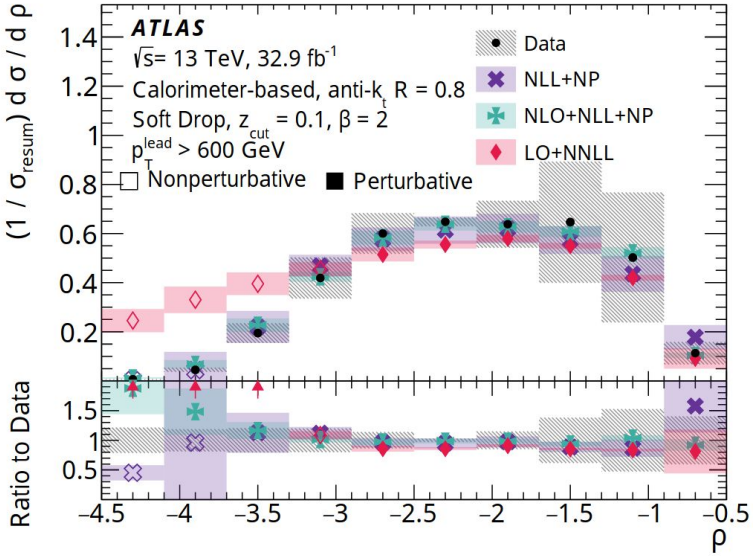
Resummation of large logs ← $\underbrace{N^k LL}_{\text{small } \rho} + \underbrace{N^m LO}_{\text{large } \rho}$ → **Fixed order** at ~$O(\alpha_s)$

- Calculations done with different theoretical approaches
  - NLL + LO for $z_{cut} \ll 1$ — Frye, Larkoski, Schwartz, Yan (2016)
  - LL + NLO for all z_{cut} — Marzani, Soyez, Schunk (2017)
  - Inclusive jets version — Kang, Lee, Liu, Ringer (2018)

Groomed jet mass



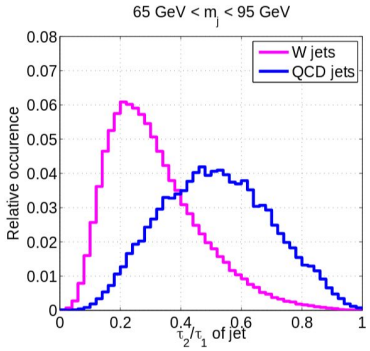
CMS measurements with mMDT
CMS-PAS-SMP-16-010

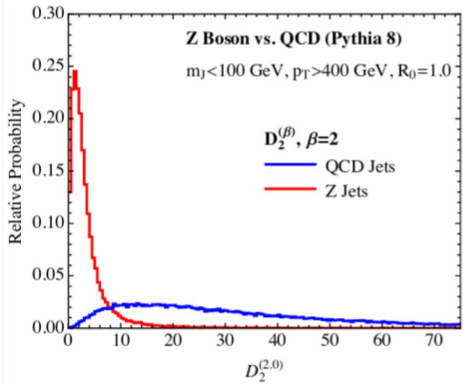


ATLAS measurements with SoftDrop
CERN-EP-2019-269

Identifying jets with substructure observables

- Variety of observables have been constructed to probe the hard substructure of a jet (V/H/t decay lead to jets with multiple hard cores).
- Radiation patterns of colourless objects (W/Z/H) differs from quark or gluon jets.
- Efficient discriminators can be obtained e.g. from ratio of N -subjettiness or energy correlation functions.





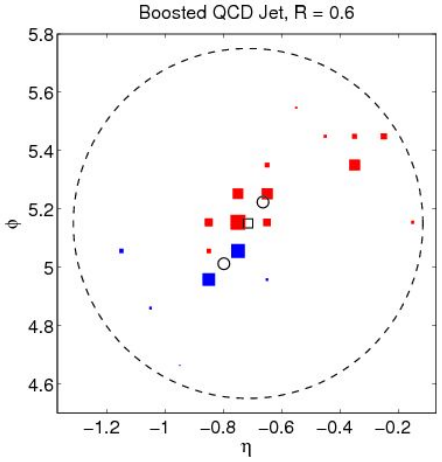
[Thaler, Van Tilburg, JHEP 1103 (2011) 015]

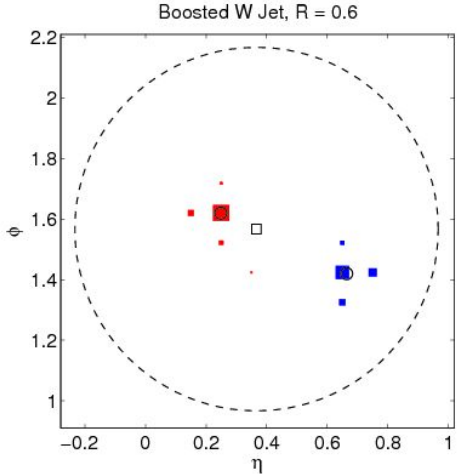
Jet shapes : N-subjettiness

- Measures radiation around N pre-defined axis

$$\tau_N^{(2)} = \frac{1}{p_{t,jet} R^2} \sum_{i \in jet} p_{t,i} \min_{a_i \ldots a_N} (\theta_{ia_1}^2, \ldots, \theta_{ia_N}^2).$$

- Use $\tau_{21} = \tau_2/\tau_1$ for 2-pronged jets and $\tau_{32} = \tau_3/\tau_2$ for 3-pronged jets





[Larkoski, Salam, Thaler, JHEP 1306 (2013) 108]
[Larkoski, Moult, Neill, JHEP 1412 (2014) 009]

Jet shapes : Energy correlation functions

- Measures dispersion through N -point correlation functions, which are sensitive to (N − 1) -prong substructure
- Advantage of not needing pre-defined axis ($z_i = p_{t,i} / p_t$, the momentum fraction)

$$e_2^{(\beta)} = \sum_{1 \le i < j \le n_J} z_i z_j \theta_{ij}^{\beta}$$

$$e_3^{(\beta)} = \sum_{1 \le i < j < k \le n_J} z_i z_j z_k \theta_{ij}^{\beta} \theta_{ik}^{\beta} \theta_{jk}^{\beta}$$

- For 2-pronged jets

$$D_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3}$$

- For 3-pronged jets

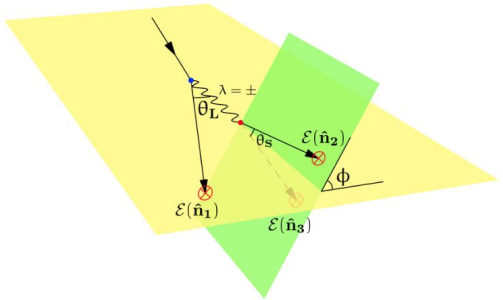
$$C_3^{(\beta)} = \frac{e_4^{(\beta)} e_2^{(\beta)}}{(e_3^{(\beta)})^2}$$

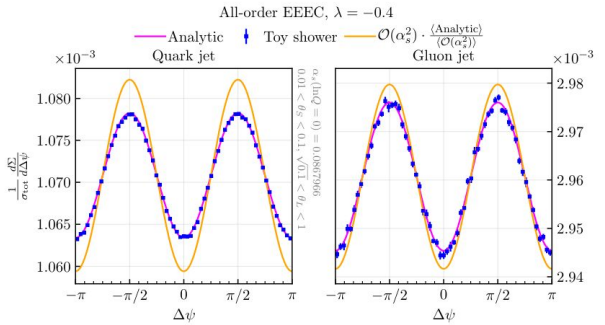
Probing spin effects with energy correlators

- Higher point correlators are particularly interesting for probing QCD, e.g. triple collinear splitting function
- Three-point energy correlator: in the limit $\theta_S \ll \theta_L$ the intermediate gluon is almost on shell and interference between both helicities leads to ϕ dependence
- Interesting probe of quantum interference using jet substructure and of recent implementations of spin correlations in parton showers

$$\frac{1}{\sigma_{\text{tot}}}\frac{d^3\Sigma}{d\Delta\psi d\theta_S d\theta_L} = \left\langle \sum_{i,j,k=1}^{N} \frac{8E_iE_jE_k}{Q^3}\,\delta\left(\Delta\psi - \phi_{(ij)k}\right)\delta\left(\theta_S - \theta_{ij}\right)\delta\left(\theta_L - \theta_{jk}\right)\right\rangle$$

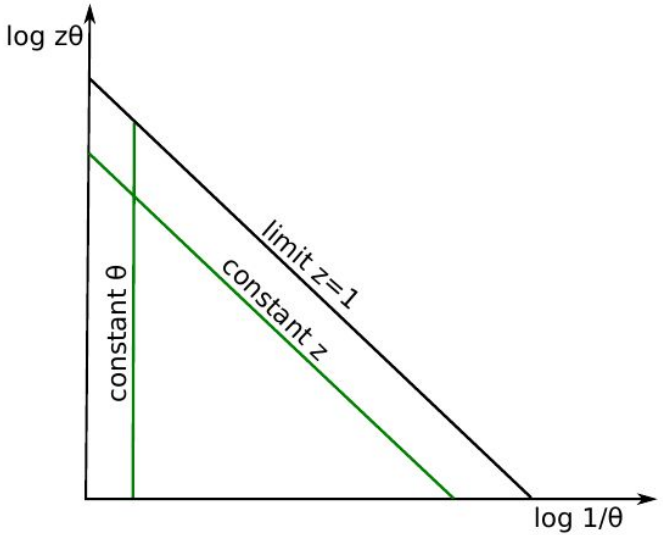
[Chen, Moult, Zhu 2020]
[Karlberg, Salam, Scyboz, Verheyen 2021]

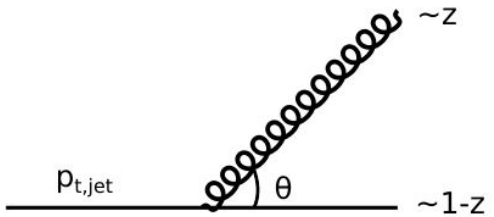




Lund diagram

- Graphical representation of emissions in zθ vs. 1/θ coordinates.
- Used to illustrate branching phase space in parton shower Monte Carlo simulations and in perturbative QCD resummations.

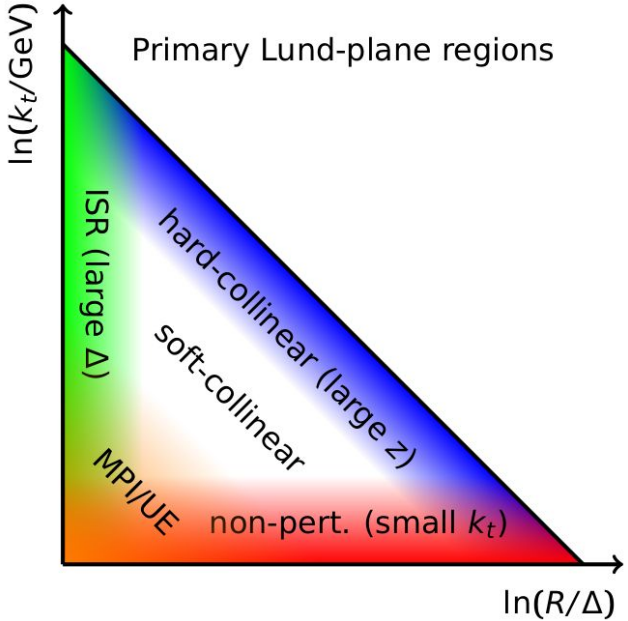


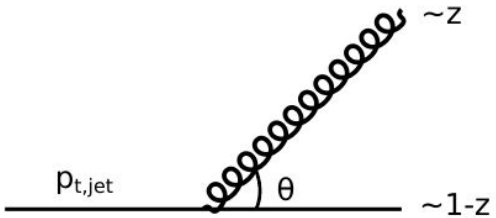


$$dP_{\mathrm{em}} \sim \frac{\alpha_s C_R}{\pi} \frac{d\theta^2}{\theta^2} dz\, p_i(z)$$

Lund diagram

- Soft-collinear emissions are emitted uniformly in the Lund plane
- Different kinematic regimes are clearly separated

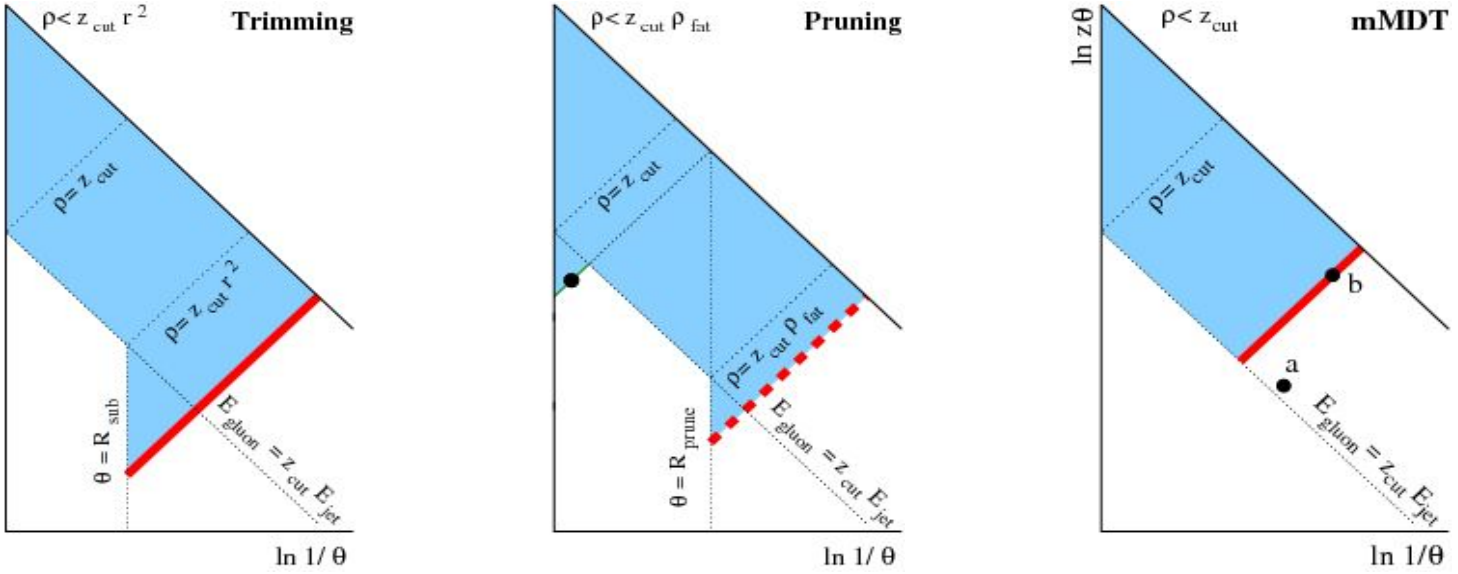




$$dP_{\mathrm{em}} \sim \frac{\alpha_s C_R}{\pi} \frac{d\theta^2}{\theta^2} dz\, p_i(z)$$

Grooming in the Lund plane

- Grooming eliminates kinematic region dominated by NP effects



[Dasgupta, Fregoso, Marzani, Salam JHEP 1309 (2013) 029]

Lund plane representation of jets

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 $\mathcal{T}^{(i)} = \{k_t, \Delta, z, m, \psi\}$

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

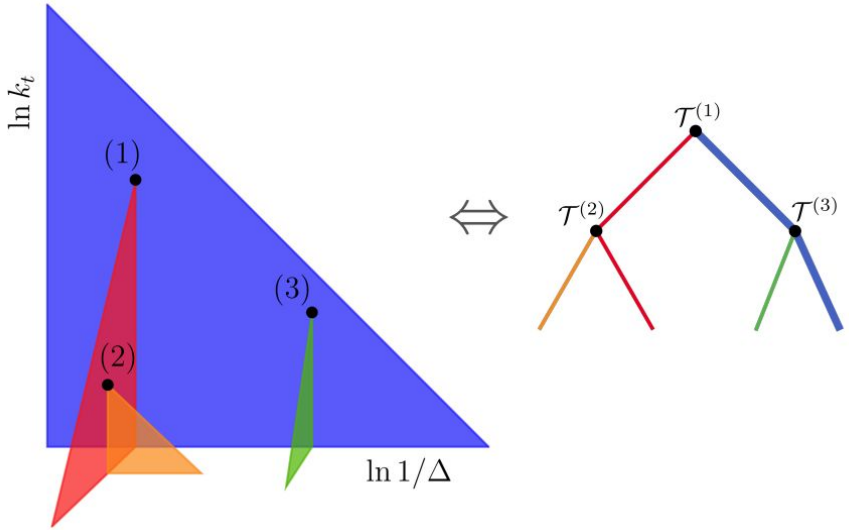
3. Repeat this procedure on both j_1 and j_2 until they are single particles.

[Dreyer, Salam, Soyez, JHEP 1812 (2018) 064]

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

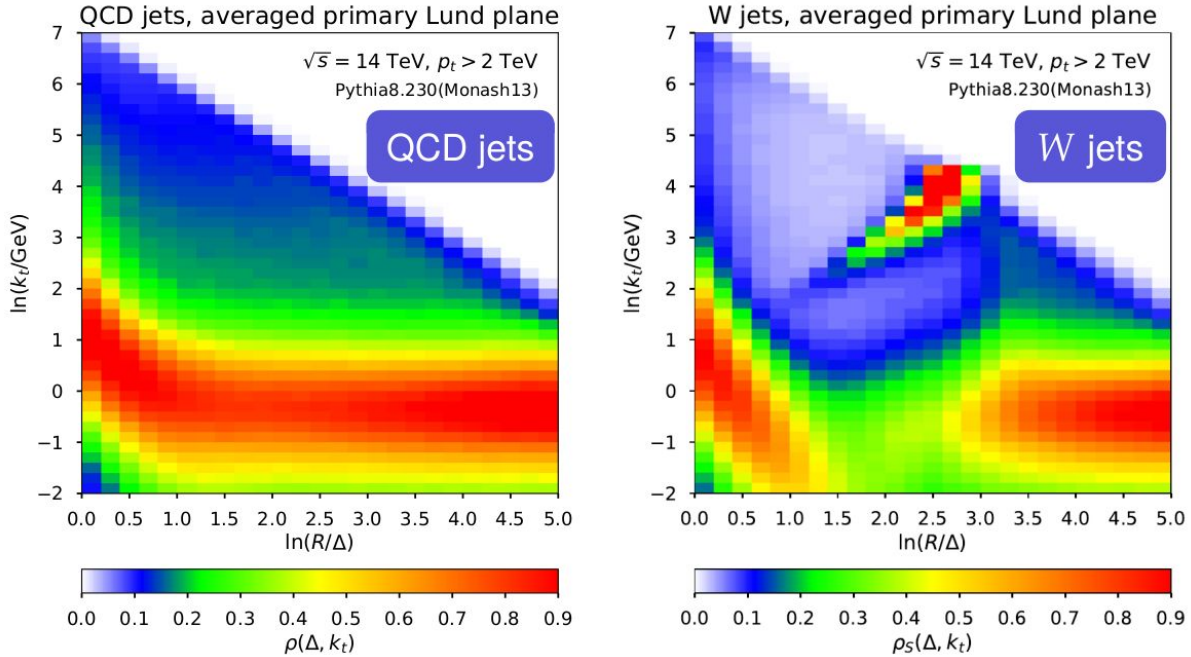
Lund plane representation

- 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.



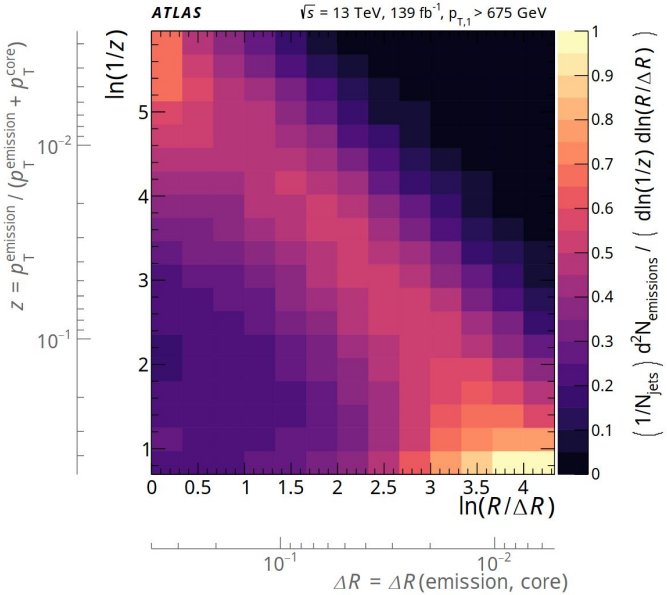
Jets as Lund images

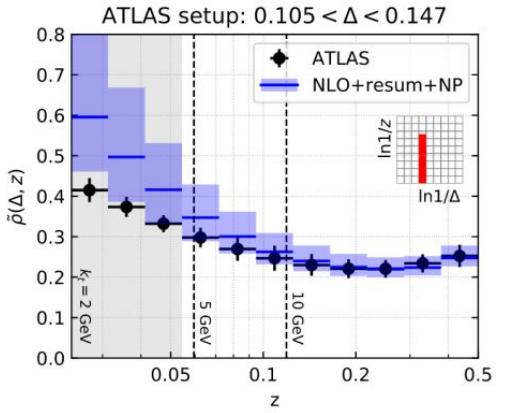
- Hard splittings visible along the diagonal line with jet mass $m = m_W$.



Measurement of the primary Lund plane

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



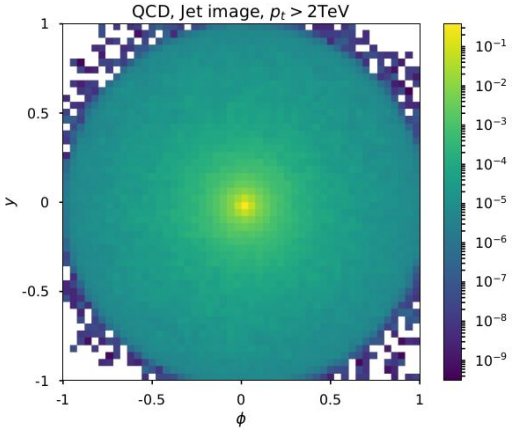


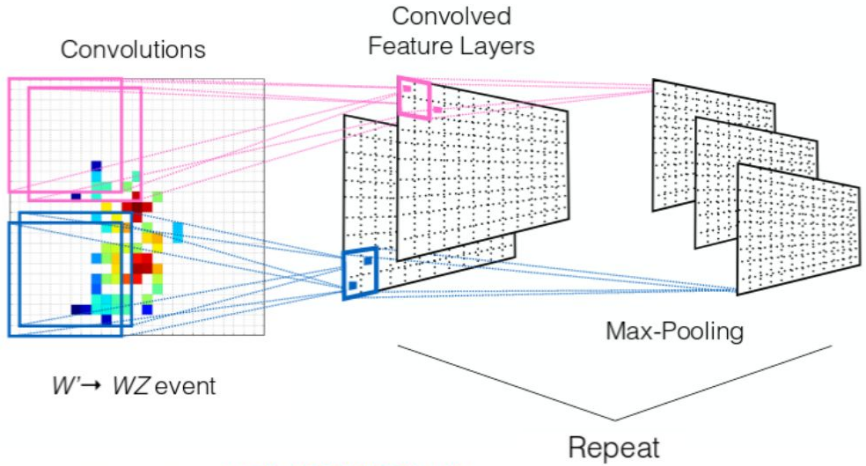
[Lifson, Salam, Soyez, JHEP 10 (2020) 170]

Machine Learning in Jet Physics

Convolutational Neural Networks and Jet Images

- Project a jet onto a fixed $n \times n$ pixel image in rapidity-azimuth, where each pixel intensity corresponds to the momentum of particles in that cell.
- Can be used as input for classification methods used in computer vision, such as deep convolutional neural networks.



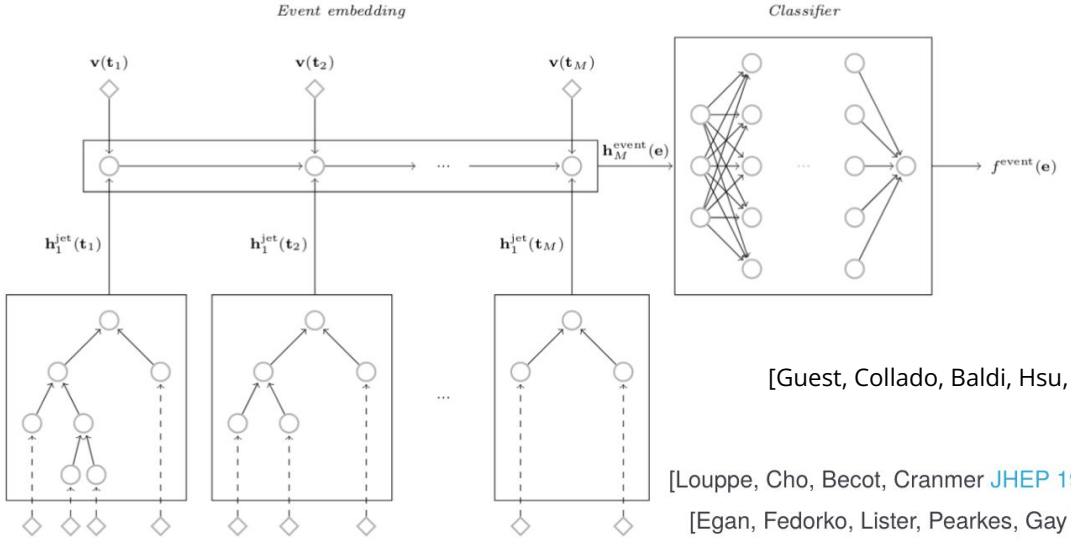


[Cogan, Kagan, Strauss, Schwartzman JHEP 1502 (2015) 118]

[de Oliveira, Kagan, Mackey, Nachman, Schwartzman JHEP 1607 (2016) 069]

Recurrent Neural Networks and clustering trees

- Train a recurrent/recursive neural network on kinematic information of successive declusterings of a jet.
- Techniques inspired from Natural Language Processing with powerful applications in handwriting and speech recognition.



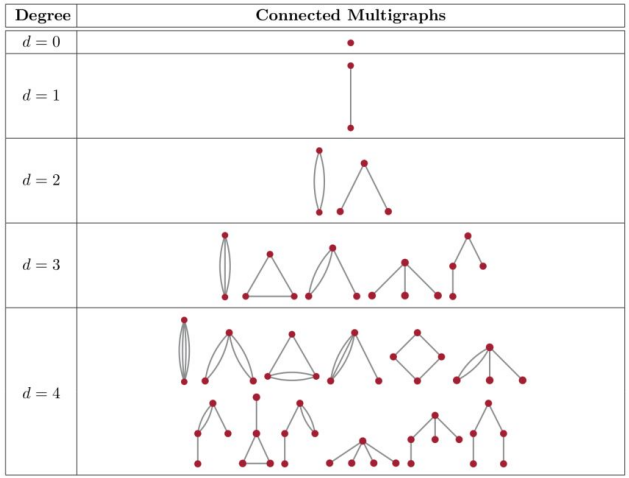
[Guest, Collado, Baldi, Hsu, Urban ,PRD 94 (2016) 11, 112002]

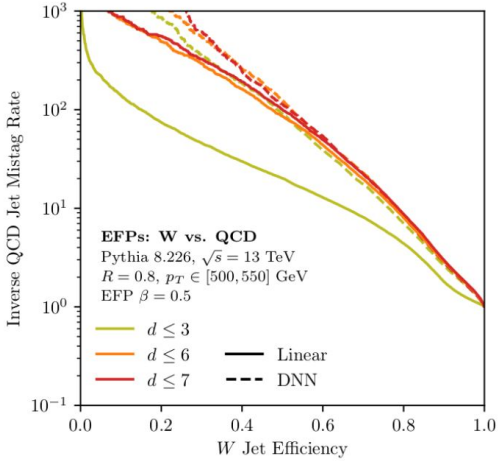
[Louppe, Cho, Becot, Cranmer JHEP 1901 (2019) 057]

[Egan, Fedorko, Lister, Pearkes, Gay 1711.09059]

Observable basis as low-dimensional representation

- Construct an observable basis that encodes the main physical properties of a jet (e.g. set of N -subjettiness ratios, energy flow polynomials, ...).
- Train a dense neural network or use linear methods to build a classifier from these inputs.





[Komiske, Metodiev, Thaler JHEP 1804 (2018) 013]

[Datta, Larkoski JHEP 1706 (2017) 073]

Flow networks

Energy Flow Networks (EFN)
IRC-safe observables
+permutation invariance

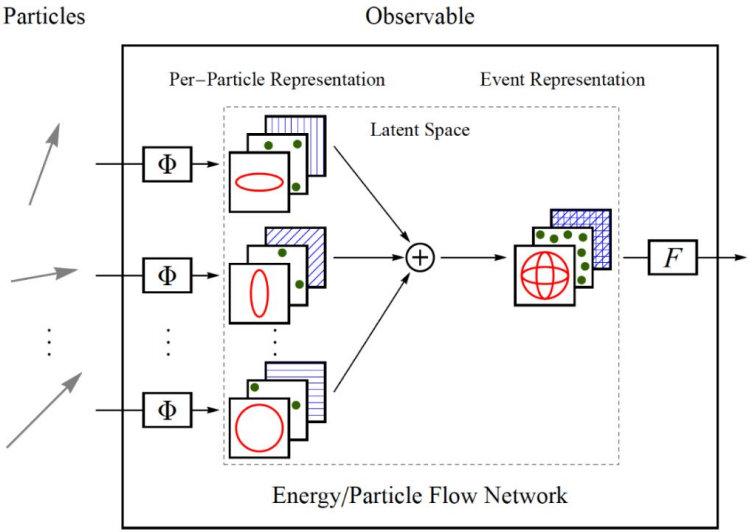
$$\mathrm{EFN} = F\left(\sum_{i=1}^{M} z_i \Phi(y_i, \phi_i)\right)$$

Φ and F parametrized with dense layers.

Particle Flow Networks (PFN)
[1810.05165]
Generalization of EFN beyond IRC safety.
Direct contact with Deep Set frameworks

$$\mathrm{PFN} = F\left(\sum_{i=1}^{M} \Phi(p_i)\right)$$

[Komiske, Metodiev, Thaler JHEP 01 (2019) 121]

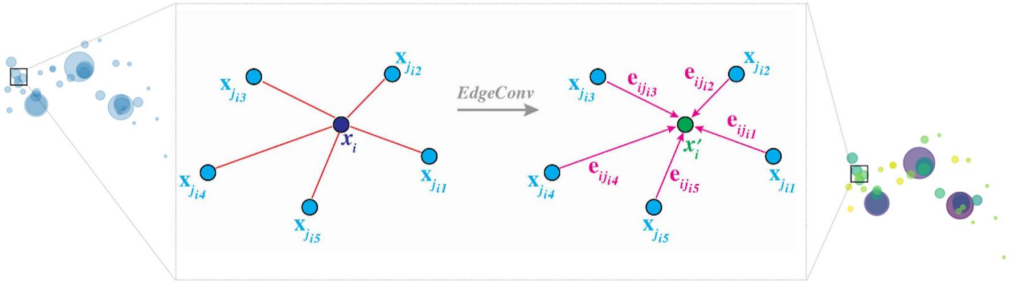


[from S. Macaluso]

Jets as point clouds

- Jet tagging using point clouds,where each jet is treated as an unordered set of constituents.
- Dynamic graph convolutions applied on graph representation of jet constituents

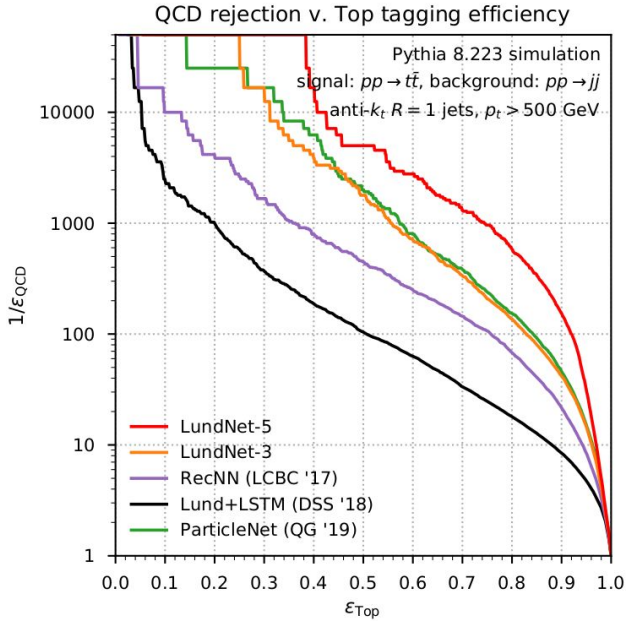
[Hu, Gouskos Phys. Rev. D 101, 056019 (2020)]

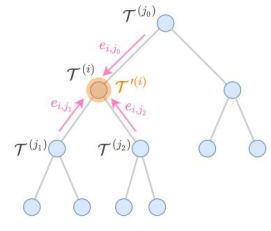


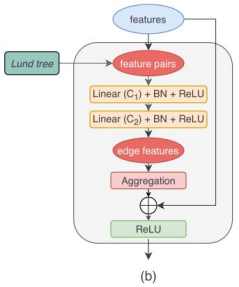
Jet tagging in the Lund plane

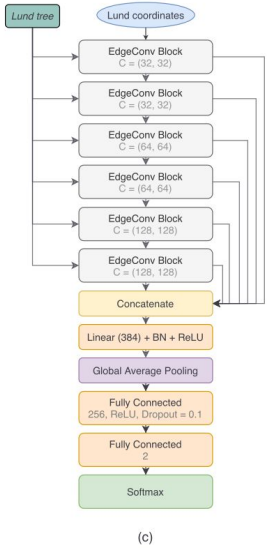
- Graph-based methods relying on Lund plane input can outperform existing benchmarks significantly.

[Qu, Gouskos 2019]
[Dreyer, Qu 2020]









Understanding what the machine learns

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 discriminant for this configuration and compare them with ML models.

We will consider quark/gluon discrimination.

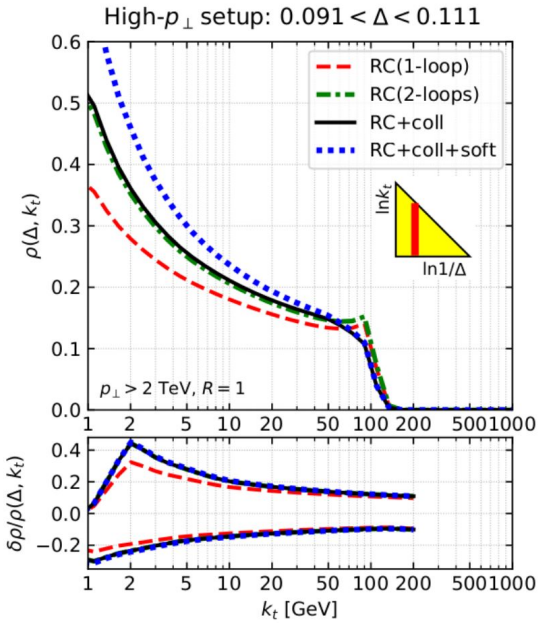
Calculating Lund plane variables

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 $\{\Delta_i, k_{t,i}, \dots\}$ 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 $\{\Delta_i, k_{t,i}, z_i, \dots\}$ compute the likelihood ratio

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

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

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

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

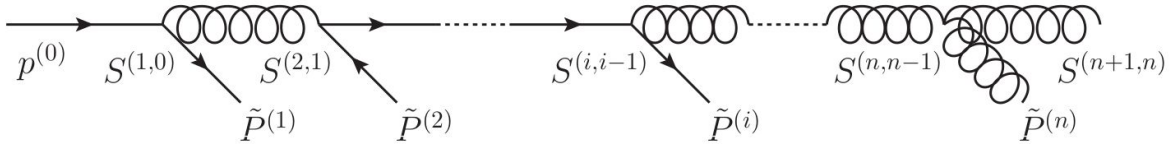
Optimal discriminant at single-logarithmic accuracy

- 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

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

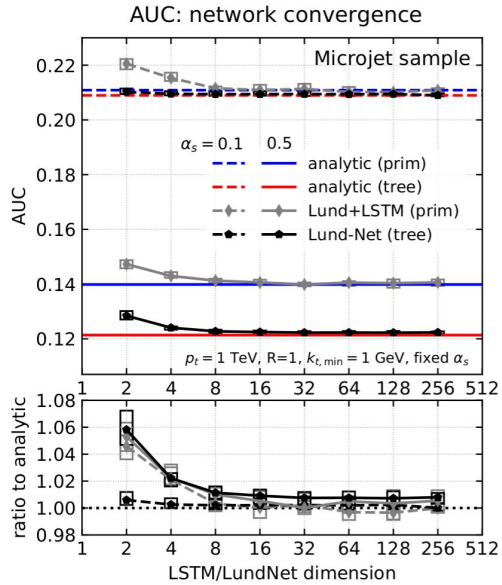
where S is a NLL Sudakov matrix and P a matrix of splitting kernels.

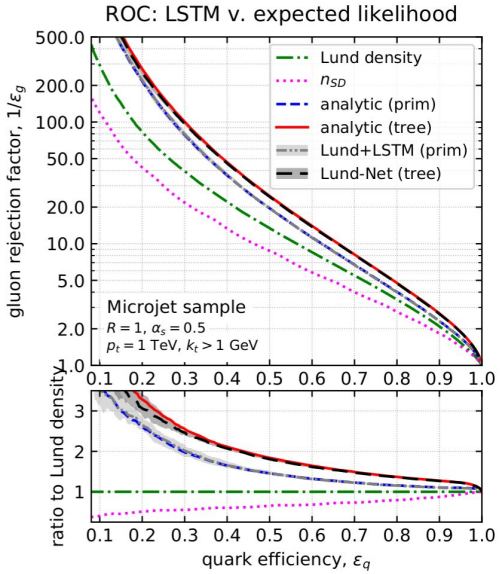


Comparison with pure-collinear parton shower

- 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.

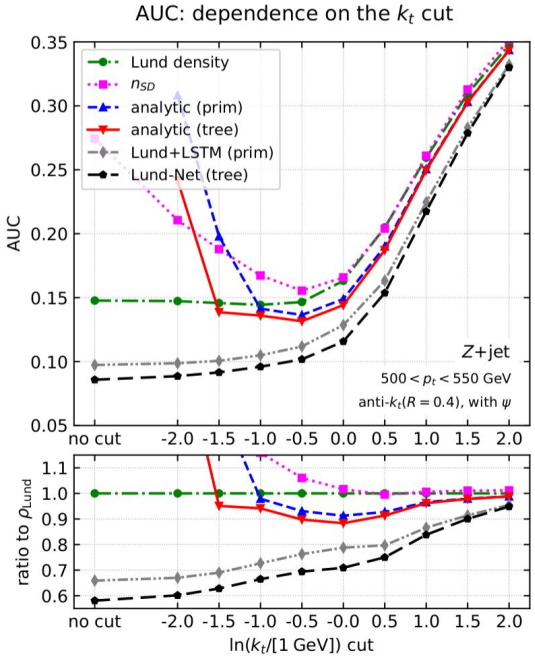
[**FD**, Soyez, Takacs, arXiv:2112.09140]

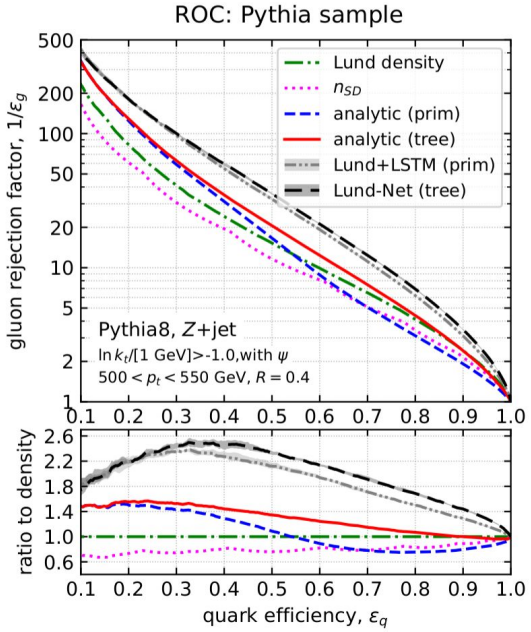




Application to full Monte Carlo

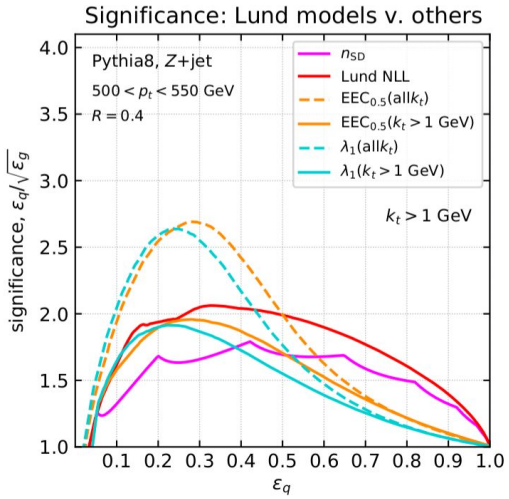
- Applying to Z+jet events generated with Pythia 8: difference in performance, but same qualitative behaviour.

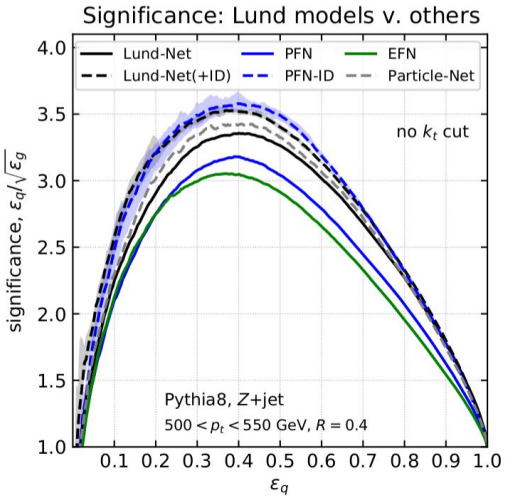




Comparison with other methods

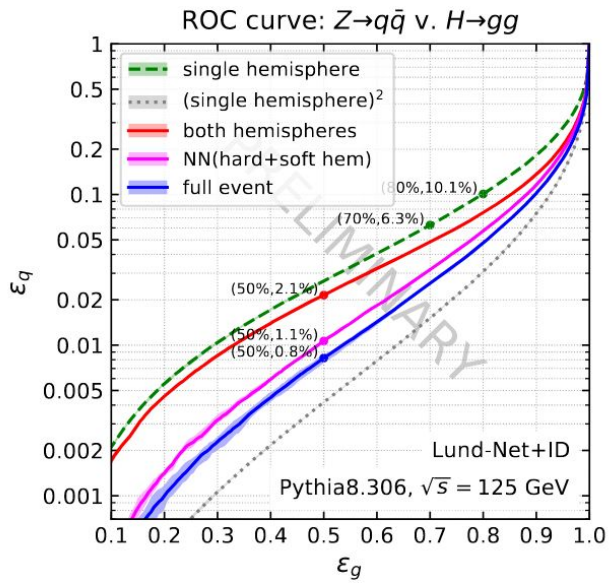
- Comparison of the Lund-plane-based approaches with other analytic and ML models.
- LundNet+ID model achieves marginally higher AUC but PFN-ID has small performance improvement at low signal efficiency.





Higgs tagging at the FCCee

$e^+e^- \to Z \to q\bar{q}$ v. $e^+e^- \to H \to gg$ $(\sqrt{s} = 125$ 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

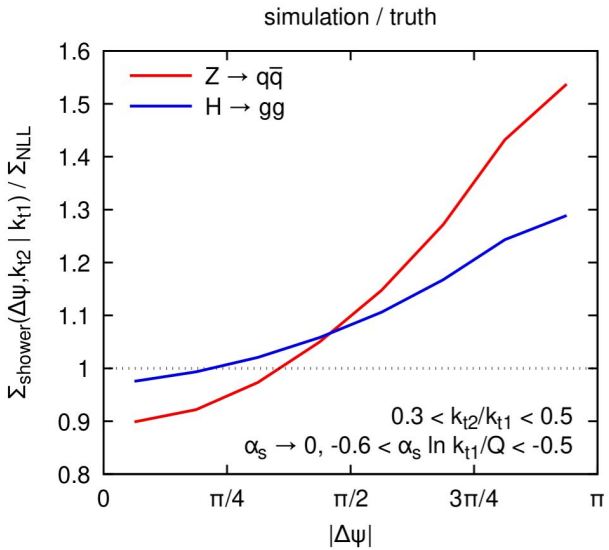
Watch out: numbers to be taken with care

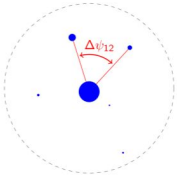
- fixed-order corrections are relevant at large ε_g
- no ISR, no detector effects, ...

[from G. Soyez]

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!



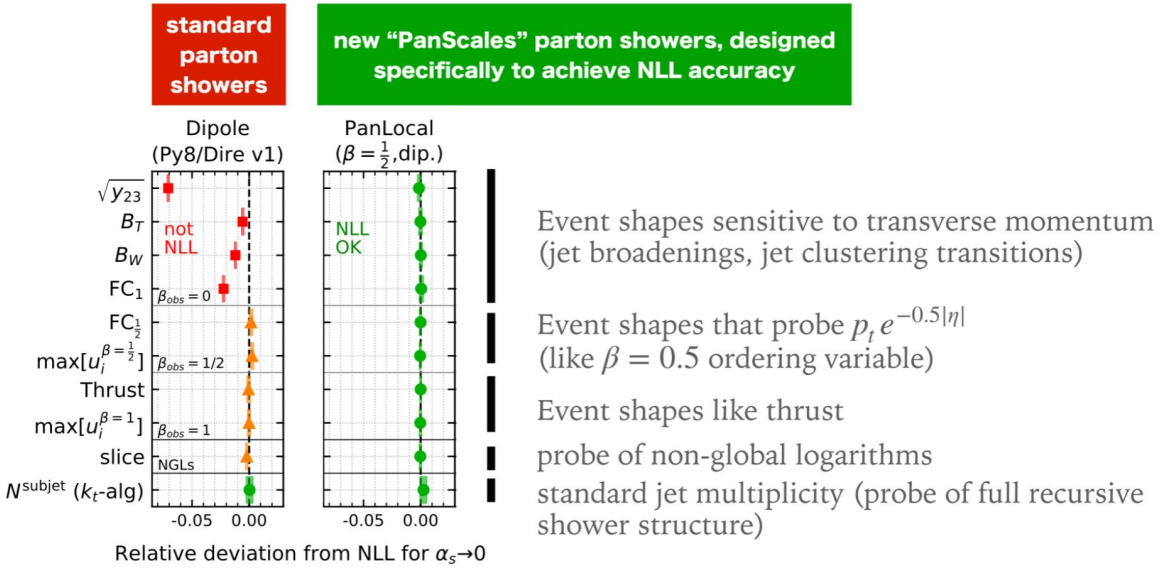


Common dipole showers display quark/gluon differences that should not be there.

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

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

Designing new showers for precision physics



[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.

Conclusions

- Higgs sector and searches for new physics requires us to understand how to relate the fundamental Lagrangian of particle physics with experimental observations.
- Understanding the structure of Higgs potential will require higher precision than what can be achieved today, higher order calculations and improvements to parton showers and their accuracies are an essential step 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.