Parton showers beyond leading logarithmic accuracy

Taming the accuracy of event generators, CERN, 29 June 2020

Frédéric Dreyer

UNIVERSITY OF OXFORD

based on arXiv:2002.11114 (to appear in PRL)

with Mrinal Dasgupta, Keith Hamilton, Pier Monni, Gavin Salam & Gregory Soyez

Core questions on event generators

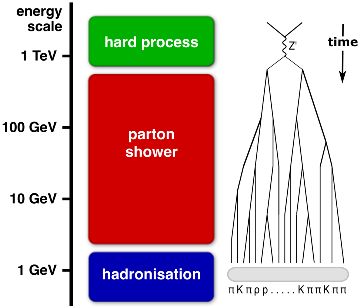
- Why is improving event generators important for precision measurements at the LHC?
- What decisions and physics go into designing a parton shower?
- How do we assess the accuracy of a parton shower?
- Can we systematically achieve next-to-leading logarithmic accuracy for broad range of observables?

Why are event generators important?

- Provide realistic and generic all-purpose simulations used in broad range of analyses with various cuts and observable choices.
- LHC collisions probe wide range of scales at which physics needs to be described accurately.
- Event generators are relied on for many different purposes, e.g.
  - background and signal estimates for new physics searches
  - uncertainty estimates for standard model measurements
  - phenomenology studies of new tools and observables
  - training data for machine learning models

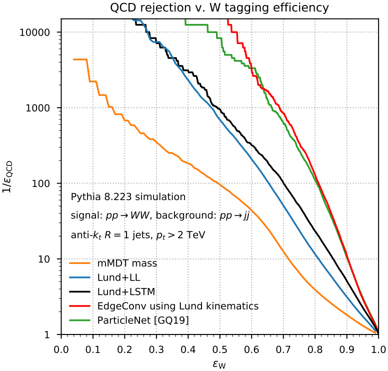
What goes into simulating a high-energy collision?

- 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?



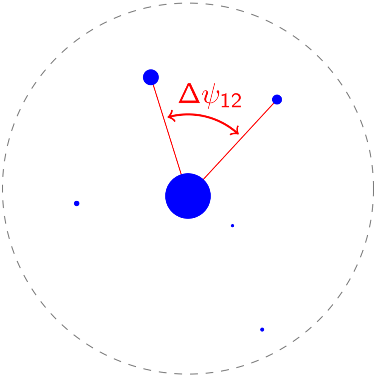
What information is hidden in an event?

- There has been much progress in using jet substructure to design efficient taggers for boosted $H/W/Z$ bosons.
- Recent state-of-the-art tools use machine learning models trained on Monte Carlo data.
- These methods are only as good as the training data they rely on.



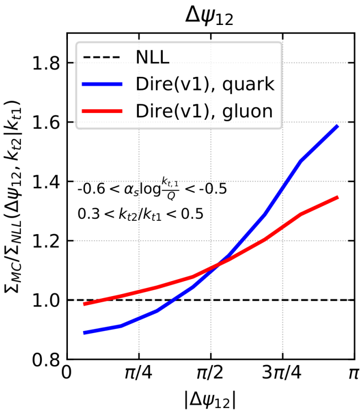
Azimuthal structure in jets

- Consider the angle $\Delta\psi_{12}$ between the hardest emissions in a jet.



Azimuthal structure in jets

- Consider the angle $\Delta\psi_{12}$ between the hardest emissions in a jet.
- Large deviations from NLL in current dipole showers.
- Difference between quark and gluon distributions might lead to models that learn to discriminate based on a feature that doesn't exist in real data!

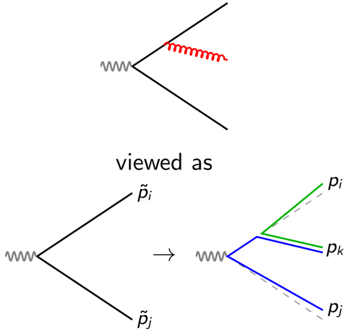


Basic picture of dipole showers

- 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.

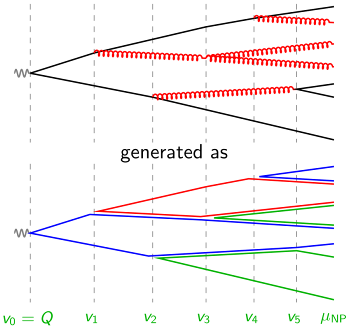


Dipole shower evolution

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

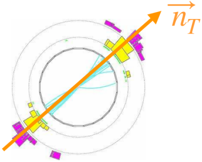


What is the accuracy of a parton shower?

- 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 $\alpha_s L \sim 1$

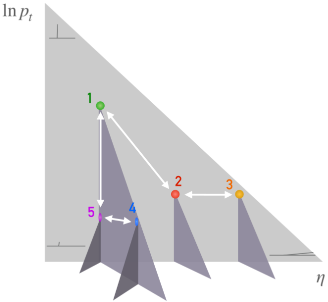
$$\sigma(1 - T < e^{-L}) = \sigma_0 \exp[\underbrace{L g_1(\alpha_s L)}_{\text{LL}} + \underbrace{g_2(\alpha_s L)}_{\text{NLL}} + \underbrace{\alpha_s g_3(\alpha_s L)}_{\text{NNLL}} + \dots]$$

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

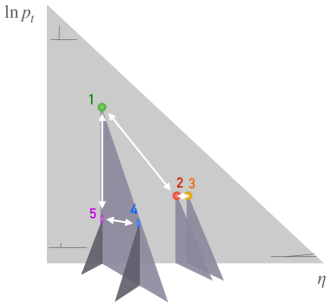
Achieving NLL accuracy

- 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 $(d_{12}, d_{23}, \cdots \gg 1)$



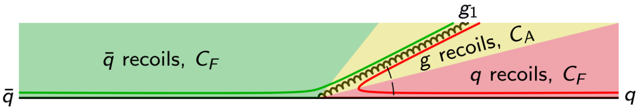
Achieving NLL accuracy

- 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 ($d_{23} \sim 1$)



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

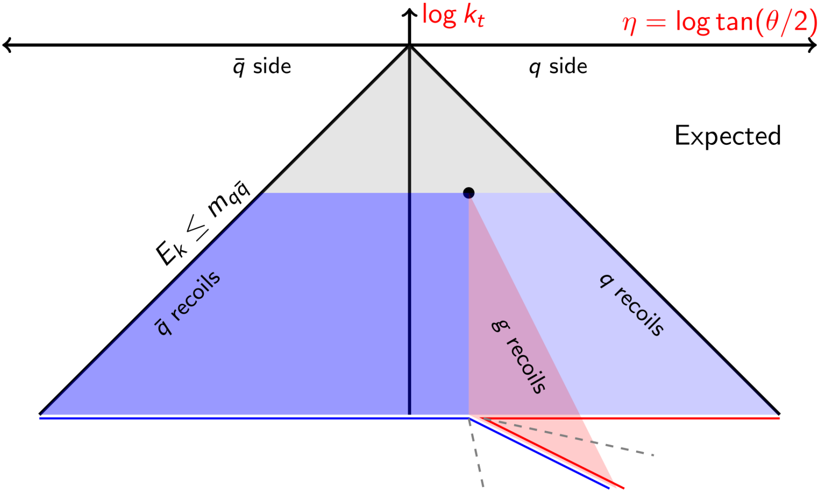
- How to associate colour and transverse recoil to dipoles?



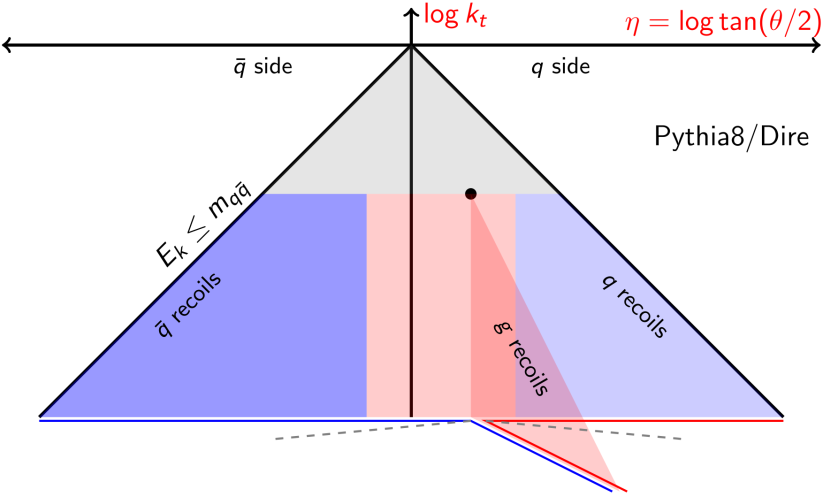
- The choice of evolution variable (transverse momentum, angle, . . .)

Design two new showers with different recoil: PanLocal and PanGlobal

Transverse recoil bondaries

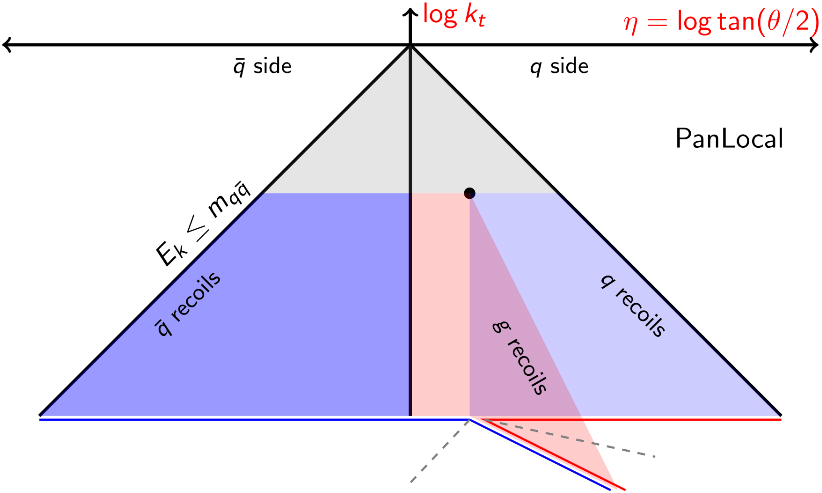


Transverse recoil bondaries



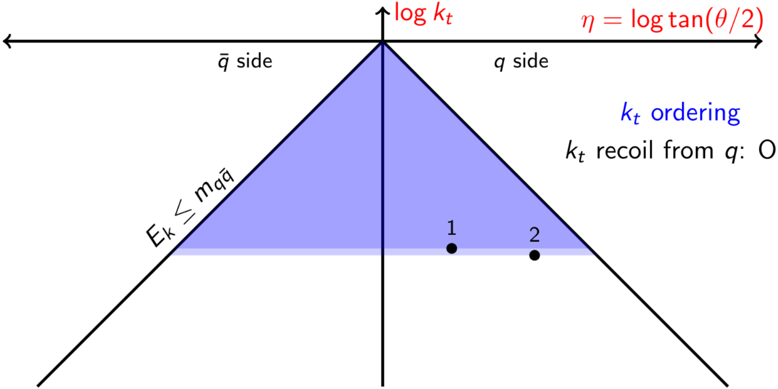
- transverse recoil assigned to end that is closer in angle in dipole c.o.m. frame

Transverse recoil bondaries



- transverse recoil assigned to end that is closer in angle in event c.o.m. frame

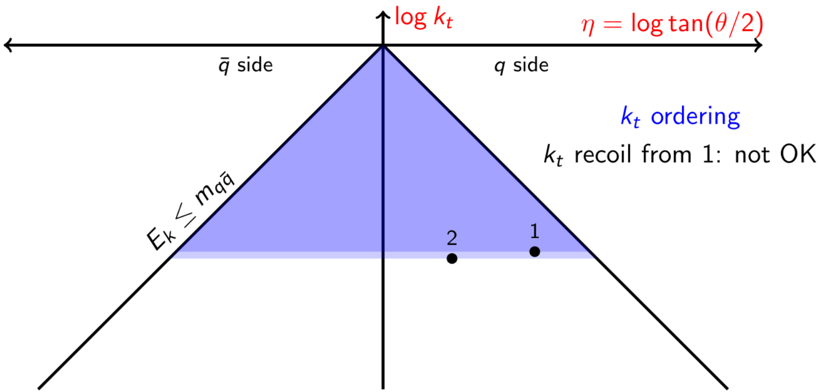
Choice of ordering variable



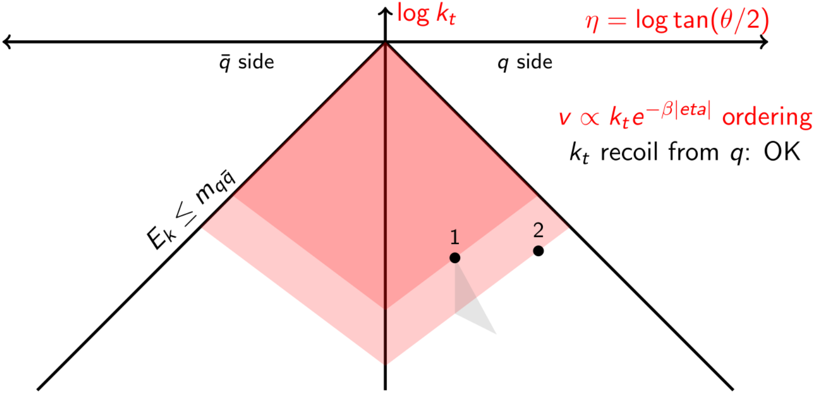
k_t ordering

k_t recoil from q: OK

Choice of ordering variable



Choice of ordering variable



- Use ordering variable intermediate between transverse momentum and angle $0 < \beta < 1$.
- Ensures that emissions with commensurate k_t are produced at successively smaller angles.

Kinematic map of the PanLocal shower

dipole $\tilde{p}_i, \tilde{p}_j \Longrightarrow p_i, p_j, p_k$.

$$p_k = a_k \tilde{p}_i + b_k \tilde{p}_j + k_\perp\,,$$
$$p_i = a_i \tilde{p}_i + b_i \tilde{p}_j - f k_\perp\,,$$
$$p_j = a_j \tilde{p}_i + b_j \tilde{p}_j - (1-f) k_\perp\,,$$

where $f \to 1$ for $k \to i$ and $f \to 0$ for $k \to j$.

$$k_t = \rho v e^{\beta|\bar{\eta}|}\,, \qquad \rho = \left(\frac{s_{\tilde{i}} s_{\tilde{j}}}{Q^2 s_{\tilde{i}\tilde{j}}}\right)^{\frac{\beta}{2}}\,.$$

$$a_k \equiv \sqrt{\frac{s_{\tilde{j}}}{s_{\tilde{i}\tilde{j}} s_{\tilde{i}}}}\, k_t e^{+\bar{\eta}}\,, \qquad b_k \equiv \sqrt{\frac{s_{\tilde{i}}}{s_{\tilde{i}\tilde{j}} s_{\tilde{j}}}}\, k_t e^{-\bar{\eta}}\,,$$

where $s_{\tilde{i}\tilde{j}} = 2\widetilde{p}_i \cdot \widetilde{p}_j$, $s_{\tilde{i}} = 2\widetilde{p}_i \cdot Q$

Partitioning of the dipole occurs at equal angles between the emission and the dipole ends in the event c.o.m. frame

Alternative scheme: global recoil

Alternatively, we can formulate a global recoil scheme, which defines the PanGlobal shower.

Longitudinal recoil is handled by dipole-local map

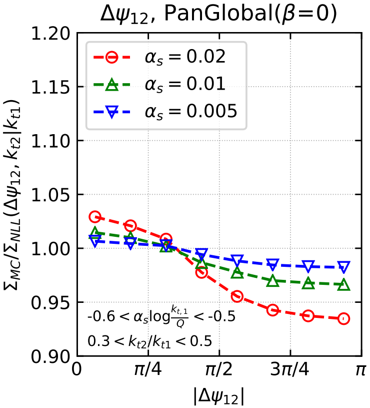
$$\text{dipole } \{\tilde{p}_i, \tilde{p}_j\} \quad \Longrightarrow \quad \begin{aligned} \bar{p}_k &= a_k \tilde{p}_i + b_k \tilde{p}_j + k_\perp \,, \\ \bar{p}_i &= (1 - a_i)\tilde{p}_i \,, \\ \bar{p}_j &= (1 - b_k)\tilde{p}_j \,. \end{aligned}$$

But transverse recoil is distributed across whole event by boost and rescaling.

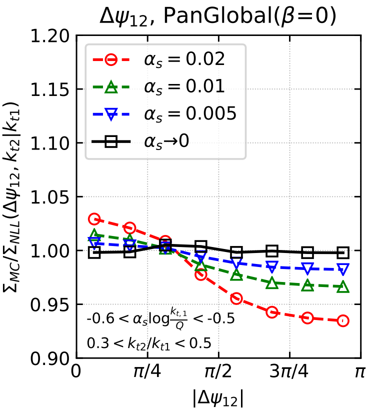
- This scheme works for $0 \leq \beta < 1$, i.e. including for k_t ordering.

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?

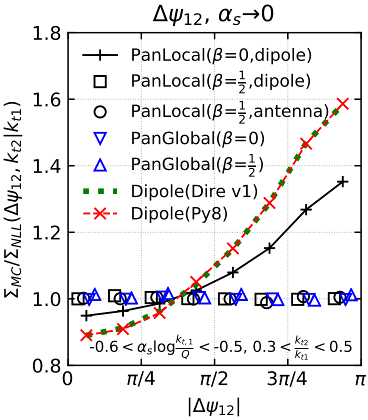


- Run full shower for smaller and smaller values of α_s, keeping $\alpha_s L$ constant
- Ratio to NLL of each distribution deviates from one: because of residual NNLL term or because of NLL mistake?
- Extrapolation $\alpha_s \rightarrow 0$ proves agreement with NLL, here for $\Delta\psi_{12}$ observable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



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



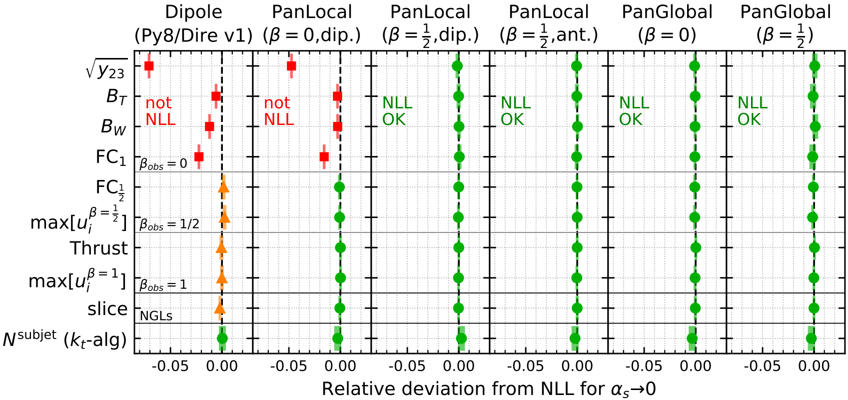
What tests should a NLL shower pass?

NLL accurate shower should pass numerical tests showing agreement with analytic NLL resummation for broad range of observables:

- event shapes sensitive to transverse momentum, e.g. jet broadenings
- event shapes probing $k_t e^{-|\eta|/2}$, e.g. fractional moment $\text{FC}_{1/2}$
- event shapes probing $k_t e^{-|\eta|}$, e.g. thrust
- non-global observables e.g. transverse momentum in a rapidity slice
- jet multiplicity probing the recursive structure of the shower

Systematic assessment of parton shower accuracy

- Pythia 8 deviates from NLL, while PanLocal $0 < \beta < 1$ and PanGlobal $0 \leq \beta < 1$ correctly reproduces global observables, non-global observables and multiplicities.



Orange coding indicates NLL issues at fixed order that are masked at all orders.

CONCLUSIONS

Conclusions

- Parton showers and their accuracies is an essential component for the theoretical and experimental HEP program.
- Formal accuracy of a shower can be defined in terms of comparison with analytic resummations, using systematic numerical checks.
- From simple building blocks, possible to design shower that is NLL accurate, i.e. controlling terms $O(\alpha_s^n L^n)$, for both global and non-global observables.
- Paves the way for further progress: subleading colour, spin correlations, ISR, NNLL, ...