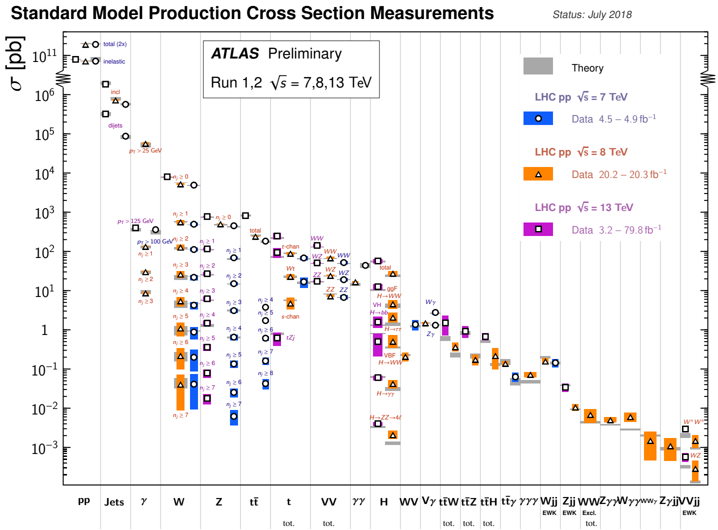
Parton showers for the precision era of the LHC

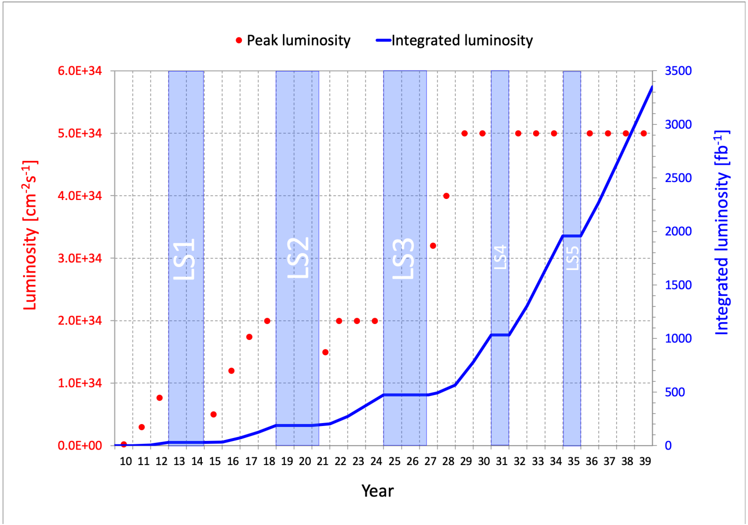
Alexander Karlberg

Seminar Fundamentale Wechselwirkungen
Albert-Ludwigs-Universität Freiburg

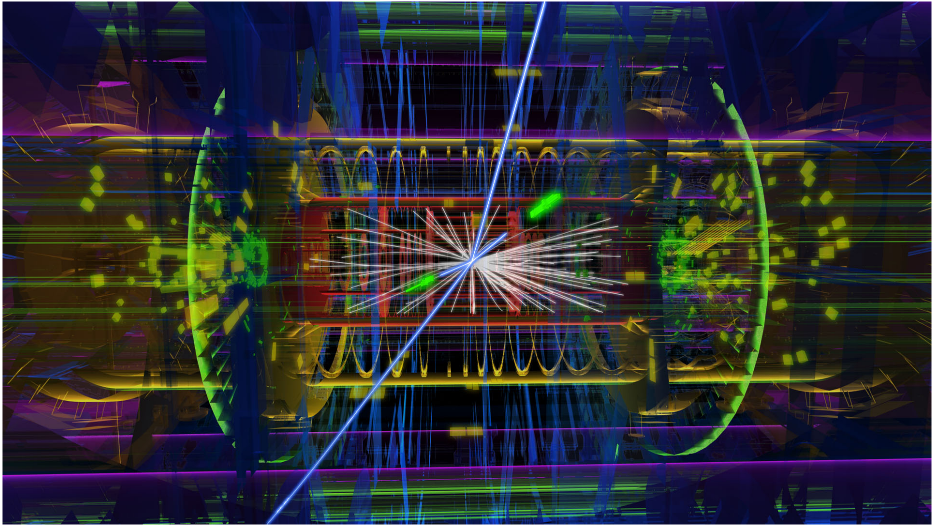
The precision era of the LHC



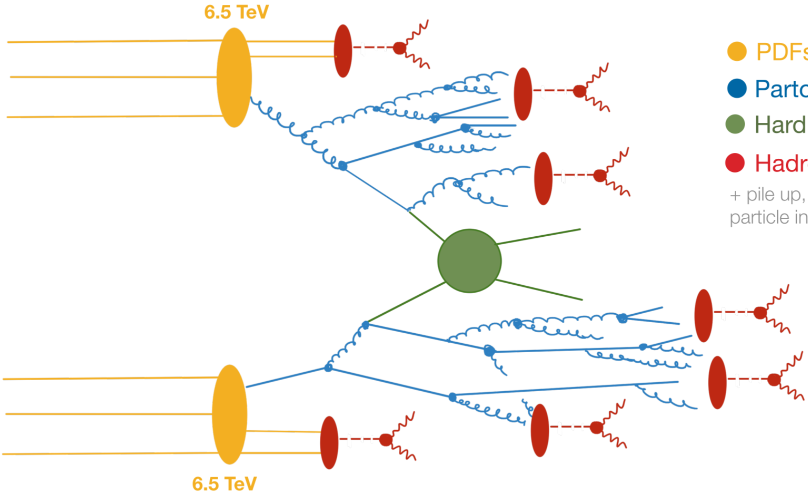
The precision era of the LHC



The LHC: A messy environment



Anatomy of an LHC collision



- PDFs / beam remnants
- Parton shower $\mathcal{O}(1-100)$ **GeV**
- Hard scattering $\mathcal{O}(0.1-1)$ **TeV**
- Hadronisation $\mathcal{O}(1)$ **GeV**

+ pile up, underlying event, multiple-particle interactions (MPI)…

courtesy M. van Beekveld

The ubiquitous Parton Shower

Pythia 8

An introduction to PYTHIA 8.2 #1

Torbjörn Sjöstrand (Lund U., Dept. Theor. Phys.), Stefan Ask (Cambridge U.), Jesper R. Christiansen (Lund U., Dept. Theor. Phys.), Richard Corke (Lund U., Dept. Theor. Phys.), Nishita Desai (U. Heidelberg, ITP) et al. (Oct 11, 2014)

Published in: *Comput.Phys.Commun.* 191 (2015) 159-177 • e-Print: 1410.3012 [hep-ph]

pdf links DOI cite 4,050 citations

Herwig 7

Herwig++ Physics and Manual #1

M. Bahr (Karlsruhe U., ITP), S. Gieseke (Karlsruhe U., ITP), M.A. Gigg (Durham U., IPPP), D. Grellscheid (Durham U., IPPP), K. Hamilton (Louvain U.) et al. (Mar, 2008)

Published in: *Eur.Phys.J.C* 58 (2008) 639-707 • e-Print: 0803.0883 [hep-ph]

pdf links DOI cite 2,644 citations

Sherpa

Event generation with SHERPA 1.1 #1

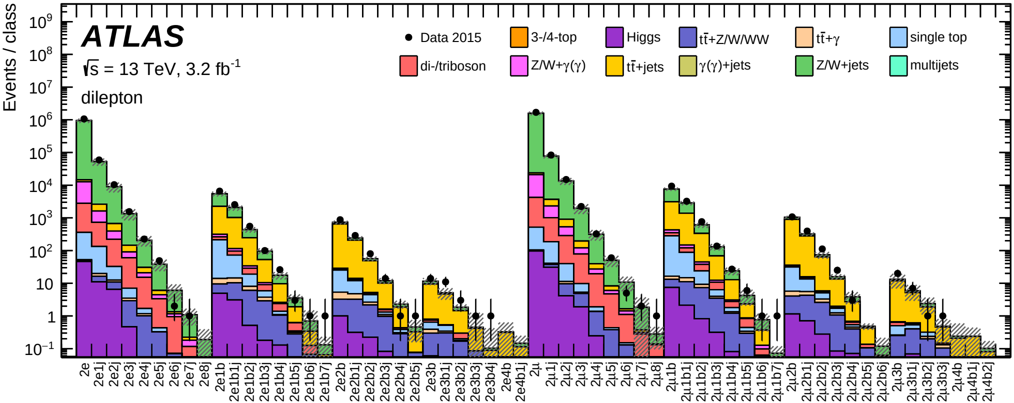
T. Gleisberg (SLAC), Stefan. Hoeche (Zurich U.), F. Krauss (Durham U., IPPP), M. Schonherr (Dresden, Tech. U.), S. Schumann (Edinburgh U.) et al. (Nov, 2008)

Published in: *JHEP* 02 (2009) 007 • e-Print: 0811.4622 [hep-ph]

pdf links DOI cite 3,386 citations

Parton Showers enter one way or another in almost 95% of all ATLAS and CMS analyses. Collider physics would not be the same without them.

The ubiquitous Parton Shower



[1807.07447]

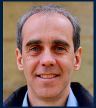
European Physical Society
High Energy and Particle Physics Division

The **2021 High Energy and Particle Physics Prize of the EPS** for an outstanding contribution to High Energy Physics is awarded to **Torbjörn Sjöstrand and Bryan Webber** for the conception, development and realisation of parton shower Monte Carlo simulations, yielding an accurate description of particle collisions in terms of quantum chromodynamics and electroweak interactions, and thereby enabling the experimental validation of the Standard Model, particle discoveries and searches for new physics.

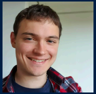
Torbjörn Sjöstrand: founding author of Pythia
Byran Webber: founding author of Herwig (with Marchesini†)

The PanScales collaboration

Oxford


Gavin Salam


Melissa van Beekveld


Rok Medves


Frederic Dreyer


Ludo Scyboz


Jack Helliwell

UCL


Keith Hamilton


Rob Verheyen

Manchester


Basem El-Menoufi

CERN


Mrinal Dasgupta


Gregory Soyez


Pier Monni


AKB


Alba Soto-Ontoso

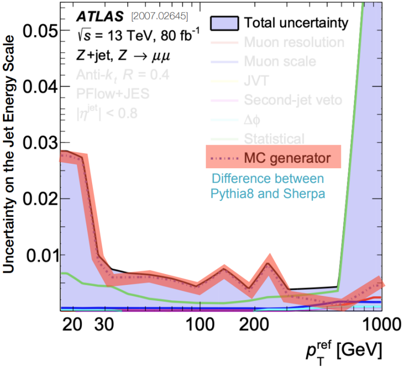

Silvia Ferrario Ravasio

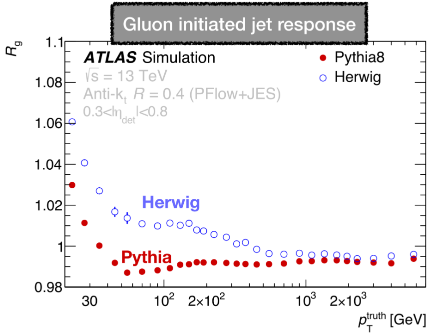
Movie

http://panscales.org/videos.html

Differences matter!

Jet energy calibration uncertainties feed in to all jet analyses at the LHC





Differences amongst MC generators is the dominant uncertainty

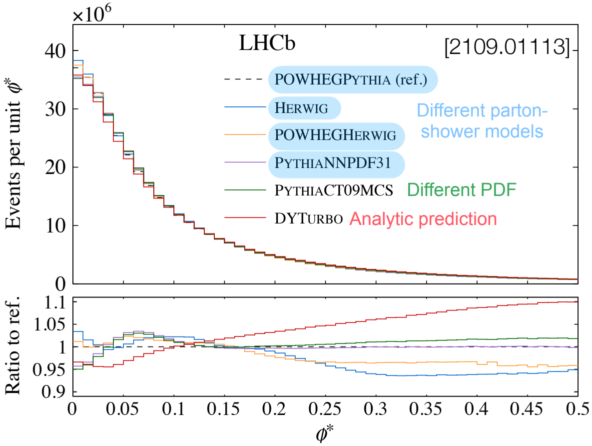
But differences matter…

Consider measurement of W boson mass

Measurements of p_T^Z in $Z/\gamma^* \to l^+l^-$ decays used to validate the MC predictions for p_T^W

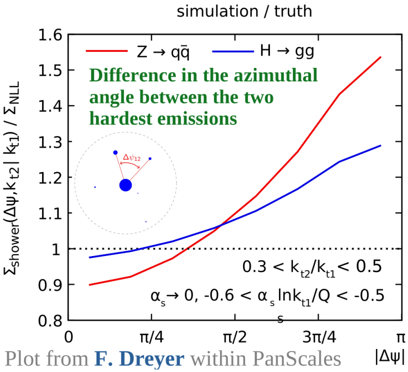
The envelope of shifts in m_W originating from differences in these shower predictions is the dominant theory uncertainty (11 MeV)

$$m_W = 80354 \pm 23_{\text{stat}} \pm 10_{\text{exp}} \pm 17_{\text{theory}} \pm 9_{\text{PDF}} \text{ MeV}$$

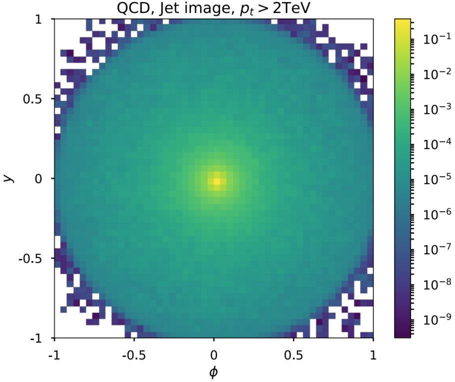


$$\phi^* = \frac{\tan((\pi - \Delta\phi)/2)}{\cosh(\Delta\eta/2)} \sim \frac{p_T^Z}{m_{ll}}$$ [1009.1580]

Machine learning and jet sub-structure



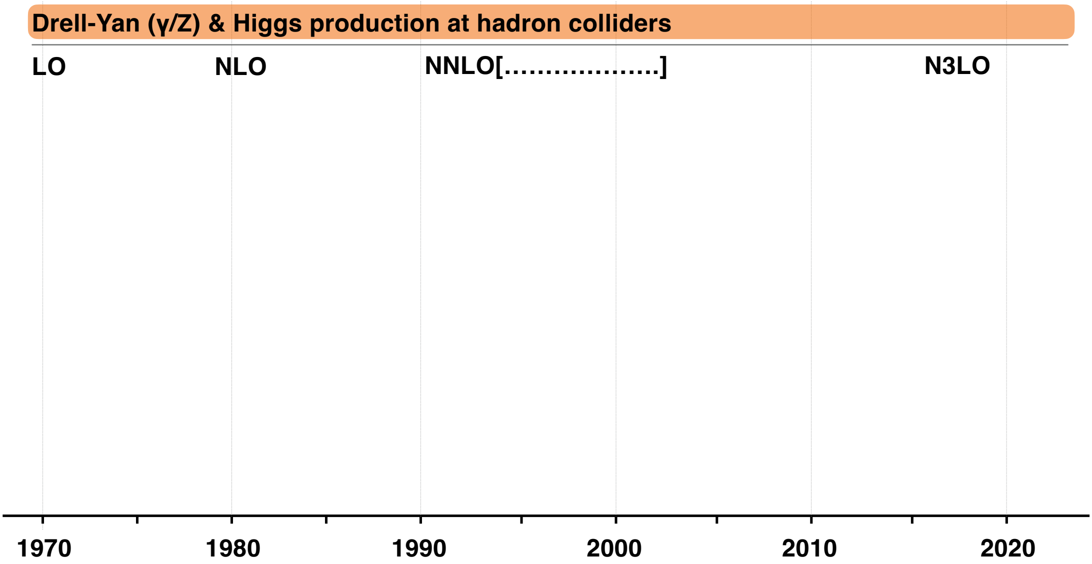
Plot from **F. Dreyer** within PanScales



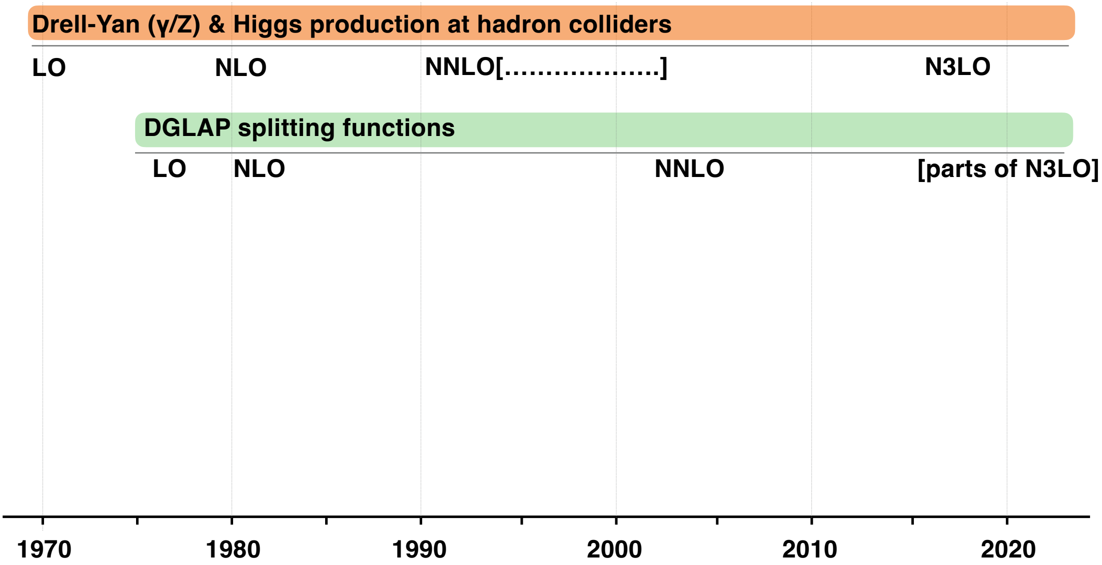
[1511.05190]

Machine learning might learn un-physical "features" from MC $\rightarrow$ can significantly impact the potential of new physics searches.

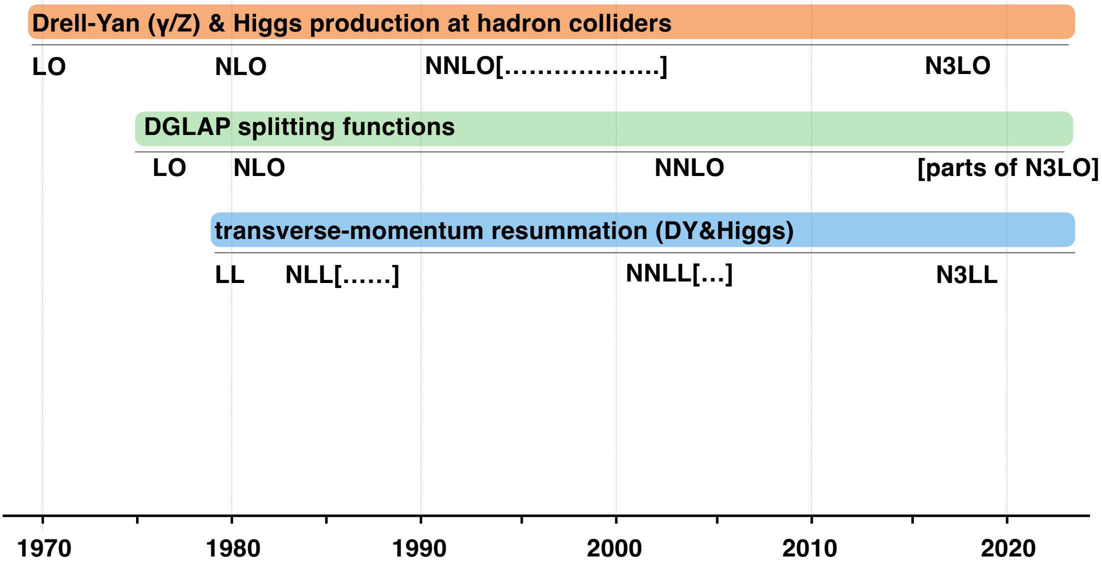
selected collider-QCD accuracy milestones



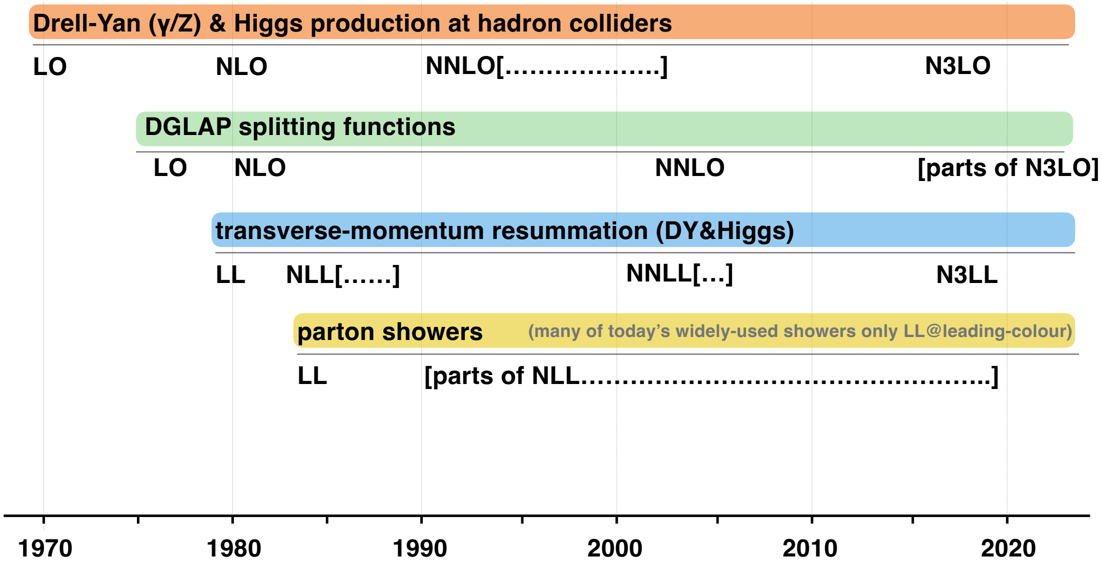
selected collider-QCD accuracy milestones



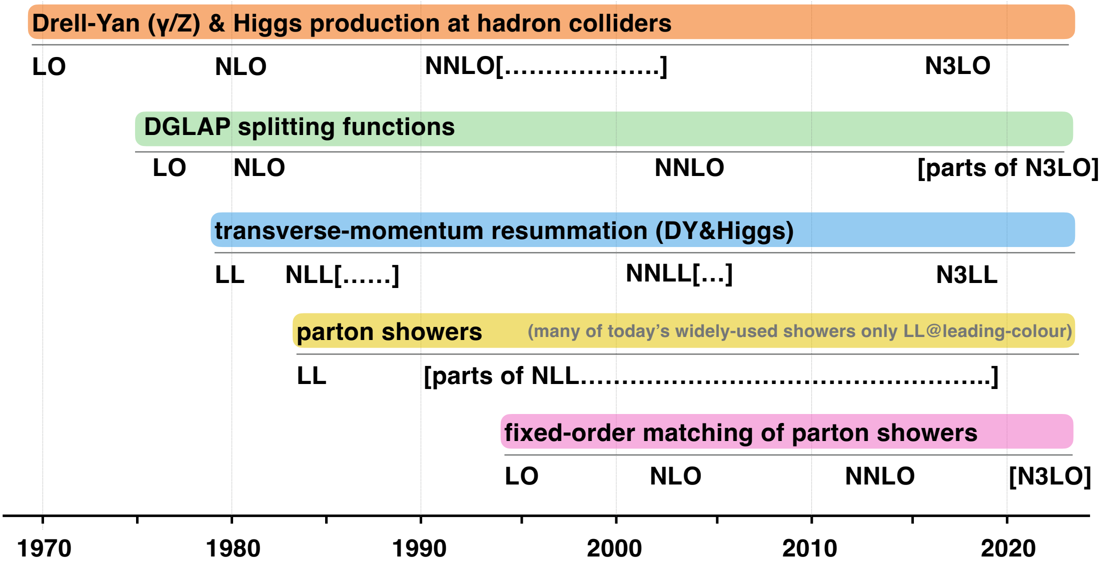
selected collider-QCD accuracy milestones



selected collider-QCD accuracy milestones



selected collider-QCD accuracy milestones



A Parton Shower in a nutshell

In one line: A Parton Shower is an iterative stochastic algorithm that takes n particles and maps them to to $n+1$ particles.

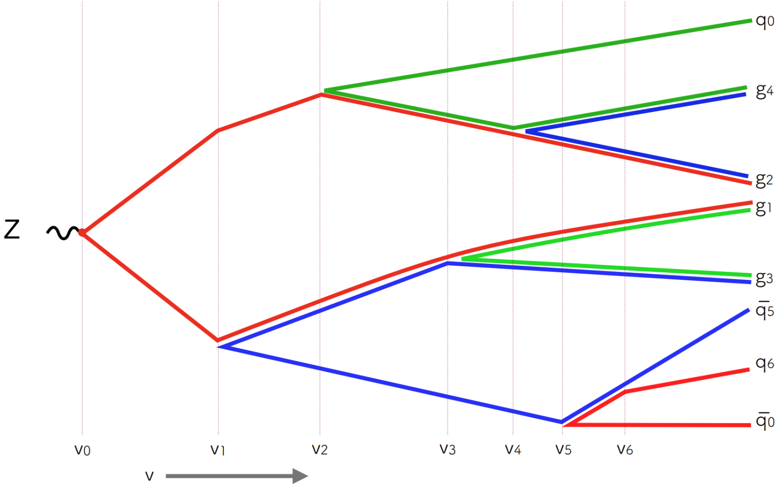
In order to do so one needs:

- A kinematic ordering variable, v, so that every phase space point is only reached once (and a cut-off $v_{cut} \sim \Lambda_{\mathrm{QCD}}$)
  - $\rightarrow$ Standard dipole showers take $v \sim k_T$ but many sensible choices exists
- A recoil map $\{p_n\} \rightarrow \{p_{n+1}\}$ to ensure momentum conservation and on-shellness of final-state particles
  - $\rightarrow$ Typically either local (only splitting dipole takes recoil) or global (all partons take recoil)
- An evolution equation governing the probability for a splitting $\widetilde{ij} \rightarrow ijk$ to take place

$$\mathrm{d}\mathcal{P}_{\widetilde{ij} \rightarrow ijk} \sim \frac{\alpha_S}{\pi} \mathrm{d}\ln v \, \mathrm{d}\bar{\eta} \, \frac{\mathrm{d}\phi}{2\pi} \left[g(\bar{\eta}) z_i P_{ik}(z_i) + g(-\bar{\eta}) z_j P_{jk}(z_j) \right] \tag{1}$$

! Mostly fixed by QCD

A Parton Shower in a nutshell



courtesy G. Salam

Accuracy of Parton Showers

How do you even define the accuracy of an algorithm as described above?

When applying perturbation theory to total cross sections, it is easy to talk about the accuracy (LO, NLO, NNLO, ...)

$$\sigma = \sum_n c_n \alpha_S^n \tag{2}$$

Similarly for logarithmically enhanced observables we may talk about their logarithmical accuracy (LL, NLL, NNLL, ...)

$$\sigma(\mathcal{O} < e^L) = \sigma_{tot} \exp\left[\frac{1}{\alpha_S} g_1(\alpha_S L) + g_2(\alpha_S L) + \alpha_S g_3(\alpha_S L) + \cdots\right] \tag{3}$$

when $\alpha_S << 1, \quad \alpha_S L \sim -1$.

But both of these equations are *observable* dependent.

Accuracy of Parton Showers

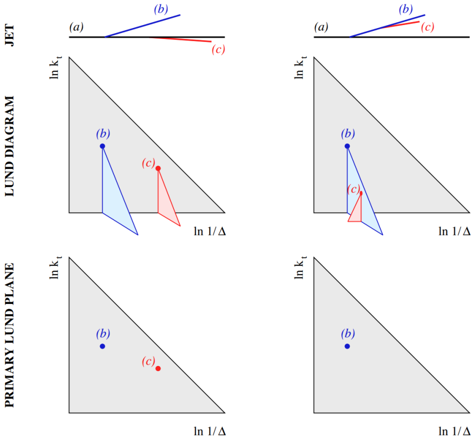
At colliders we can ask arbitrary questions about an event. The same is true for parton showers (+ hadronisation), e.g.

- Number (multiplicity) of particles in event (or jet)
- Energy in detector slice
- Angular distributions inside jets
- Even if we don't ask, machine learning might...

We therefore need to establish how to determine the logarithmic accuracy with which a parton shower can make predictions.

To do so we need to introduce the *Lund Plane* (B. Andersson et al (1989) & F. Dreyer et al. [1807.04758])

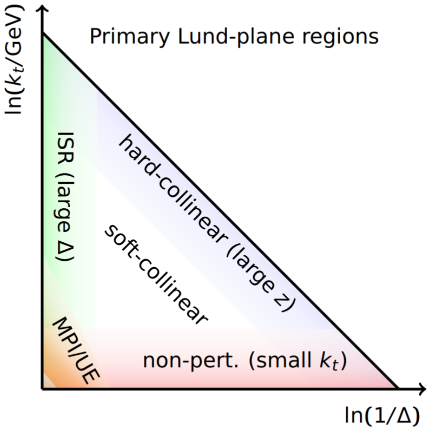
The Lund Plane



[1807.04758]

- Cluster the event with the Cambridge/Aachen algorithm, producing an angular ordered clustering sequence.
- Decluster the last clustering and record the transverse momentum and the opening angle of the declustering (plus other kinematics).
- Iterate along the hardest branch after each declustering to produce the *primary* Lund Plane.
- Following the softer branch produces the secondary, tertiary, etc Lund Plane.
- One can impose cuts easily on the declusterings (e.g. that they satisfy $z > z_{\text{cut}}$)

Logarithms in the Lund Plane



[1807.04758]

- The emission probability in the Lund Plane is then
$$\mathrm{d}\rho \sim \alpha_s \,\mathrm{d}\ln k_T \,\mathrm{d}\ln\theta$$
- Hence emissions that are well-separated in *both* directions are associated with *double logarithms* of the form $\alpha_s^n L^{2n}$
- Emissions separated along one direction are associated with *single logarithms* of the form $\alpha_s^n L^n$
- Emissions that are close in the Lund Plane are associated with a factor α_s^n
- We are now ready to state the PanScales NLL criteria for Parton Showers

NLL accurate Parton Showers

Fixed Order Matrix Element Condition

- Shower must reproduce fixed order n-particle matrix elements when emissions are well-separated in the Lund Plane, ie when the cross section is logarithmically enhanced.
- Supplement this with unitarity, 2-loop running and correct cusp anomalous dimension

Resummation Condition

- Shower must reproduce known NLL analytical resummations
- Global event shapes
- Multiplicity
- Non-global observables (slice observables), technically at leading single log (SL).

[1805.09327] & [2002.11114]

NLL accurate Parton Showers

Fixed Order Matrix Element Condition

- Fairly straightforward. Generate n emissions with your shower and compare to either factorised matrix elements (numerically very stable) or a full matrix element in some kinematic limit.
- Be careful to cover the collinear/soft phase space.

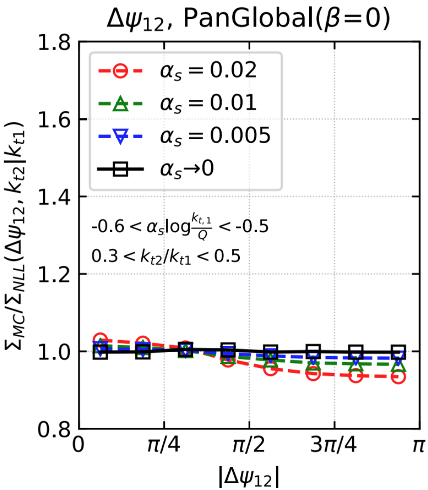
Resummation Condition

- This in general is trickier for 2 reasons:
- Requires the existence of NLL analytical results.
- Can't just compare

$$\frac{\Sigma^{\mathrm{PS}}(\alpha_S L)}{\Sigma^{\mathrm{NLL}}(\alpha_S L)} = \frac{\Sigma^{\mathrm{PS}}(\alpha_S L)}{\sigma_{tot} \exp\left[\frac{1}{\alpha_S} g_1(\alpha_S L) + g_2(\alpha_S L)\right]}$$

  as the shower in general induces spurious higher order terms.
- How do we disentangle spurious "NNLL" terms from genuine NLL violations?

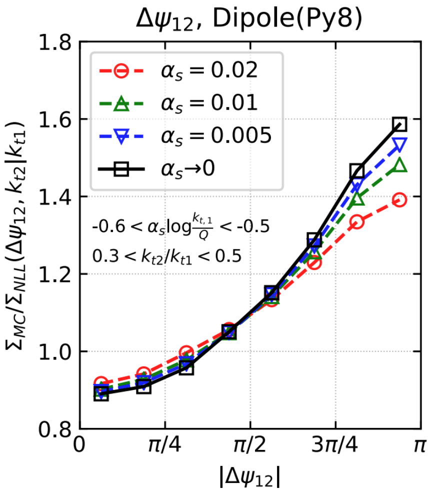
NLL tests



[2002.11114]

- Run the full shower with a specific (finite) value of $\alpha_S = \alpha_S(Q)$ and measure your favourite observable (that you can resum to NLL
- Take the ratio to NLL and see that it is not flat.
- To see if there is an NLL mistake reduce α_S while keeping $\alpha_S L$ fixed, ie include more collinear and soft emissions.
- Genuine NLL effects are $(\alpha_S L)^n$ and are therefore unchanged. NNLL on the other hand goes as $\alpha_S(\alpha_S L)^n$ and should therefore vanish.
- Go as small in α_S as possible and extract $\alpha_S \to 0$.
- Now is it flat?

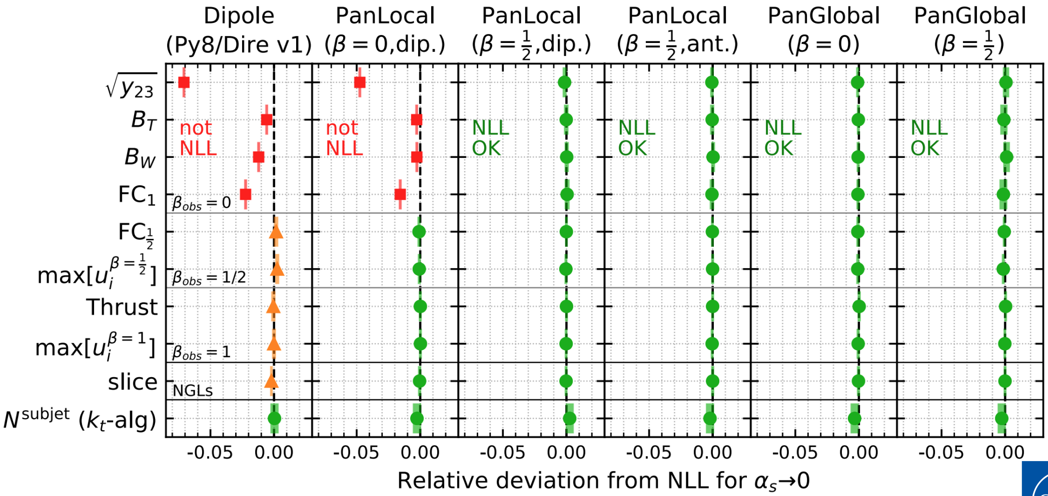
NLL tests



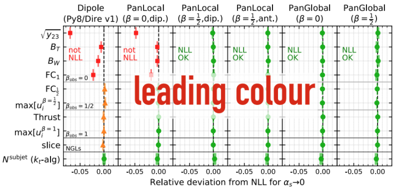
[2002.11114]

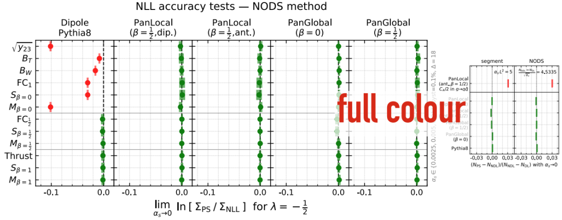
- Run the full shower with a specific (finite) value of $\alpha_S = \alpha_S(Q)$ and measure your favourite observable (that you can resum to NLL
- Take the ratio to NLL and see that it is not flat.
- To see if there is an NLL mistake reduce α_S while keeping $\alpha_S L$ fixed, ie include more collinear and soft emissions.
- Genuine NLL effects are $(\alpha_S L)^n$ and are therefore unchanged. NNLL on the other hand goes as $\alpha_S(\alpha_S L)^n$ and should therefore vanish.
- Go as small in α_S as possible and extract $\alpha_S \to 0$.
- Now is it flat?

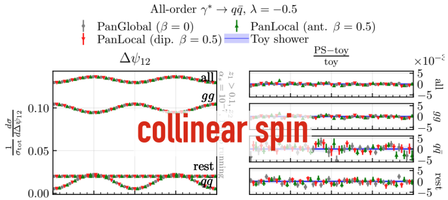
NLL tests summary

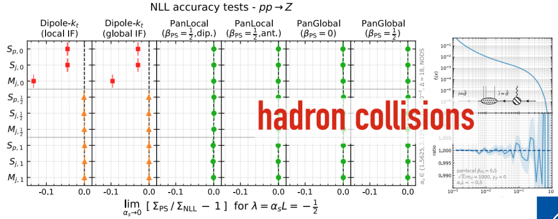


NLL tests summary









Part II: NLO matching

Status of matching

Biggest improvement in event generator accuracy in the last 20 years is due to the inclusion of hard matrix element corrections through (N)NLO matching.[1]

The two by far most used NLO matching schemes are MC@NLO (additive) and POWHEG (multiplicative).

In a nut-shell they work by either replacing or modifying the hardest emission in the parton shower with one which is accurate at $\mathcal{O}(\alpha_S)$.

In past hard to address interplay between logarithmic accuracy and fixed-order accuracy due to lack of the former in showers.

A lot of work done within PYTHIA, HERWIG, and SHERPA to make sure the hand over from the hard matrix element generator (HEG) to the shower is handled correctly.

[1] There is a separate discussion on the impact of merging which will not be addressed here.

Matching and log accuracy

There are two reasons why this may be of interest to study.

The first one is that with controlled logaritmic accuracy in the shower, we may wish to ensure that the matching does not break the logaritmic accuracy.

Furthermore, mathcing is a necessary ingredient to go beyond NLL accuracy.

In general we expect matching to contribute at the NNDL ($\alpha_S^n L^{2n-2}$) and NSL level ($\alpha_S^n L^{n-2}$).

Concretely in e^+e^- event shapes have an NNDL resummation structure given by

$$\Sigma_{\text{NNDL}}(\alpha_S, \alpha_S L) = \left(1 + C_1 \frac{\alpha_S}{2\pi}\right) \exp\left(L g_1(\alpha_S L) + g_2(\alpha_S L)\right) .$$

And hence matching is all you need to get NNDL (for these observables).

Matching on one slide

The aim of any matching procedure is to ensure that the first (hardest) emission of the shower reproduces the full α_S structure of QCD, not just the logarithmically enhanced terms.

MC@NLO: Aims to do this by *supplementing* the shower events by a set of “hard” events that correct the shower. Does not touch the shower at all, but needs information about how the shower generates its hardest emission.

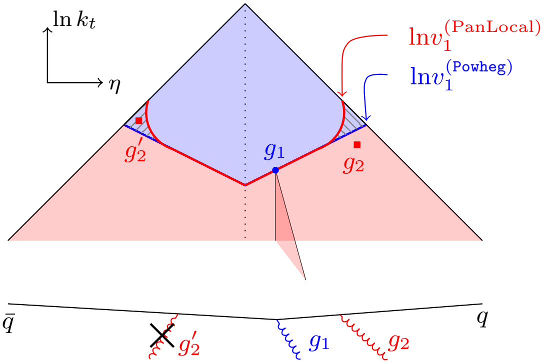
POWHEG: *Replaces* the hardest emission of the shower with its own. Does not need to know anything about the shower, but the shower evolution needs to be modified to fix any mismatch between POWHEG’s evolution variable and the shower’s.

A number of variations exist but these two by far the most succesful so far.

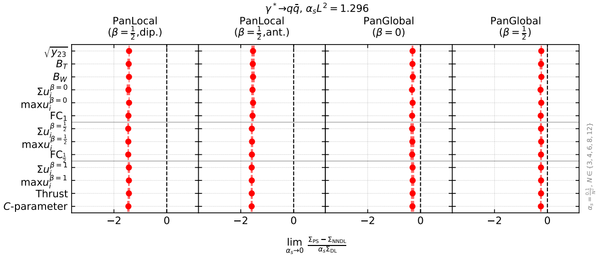
Some caveats

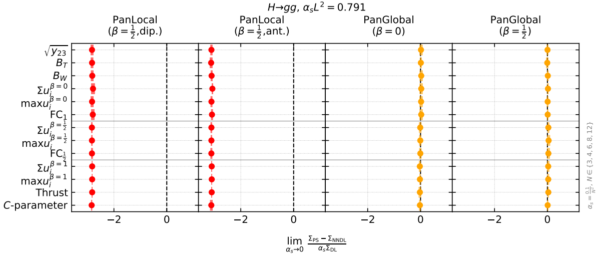
Since MC@NLO does not touch the shower in the IR, one retains NLL almost trivially, even for spin and colour. The only non-trivial bit is to make sure that the true matrix element and the shower effective matrix element are both partitioned in the same way.

The same is not true for POWHEG unless one is very careful with matching evolution contours as well.

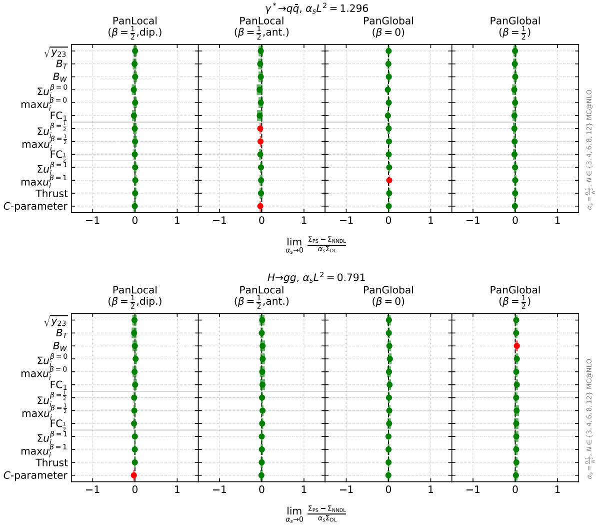


NNDL tests without matching

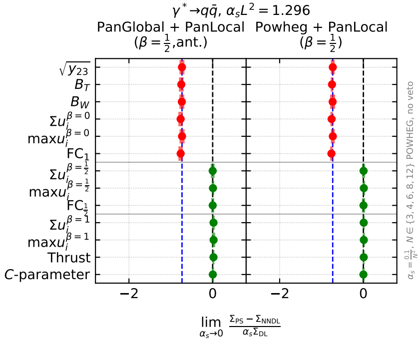


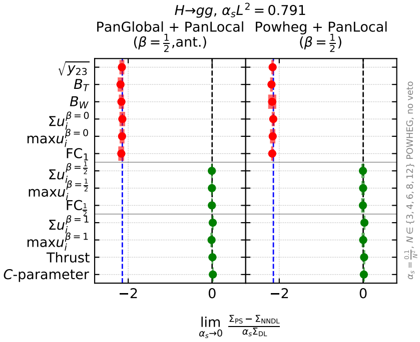


MC@NLO matching

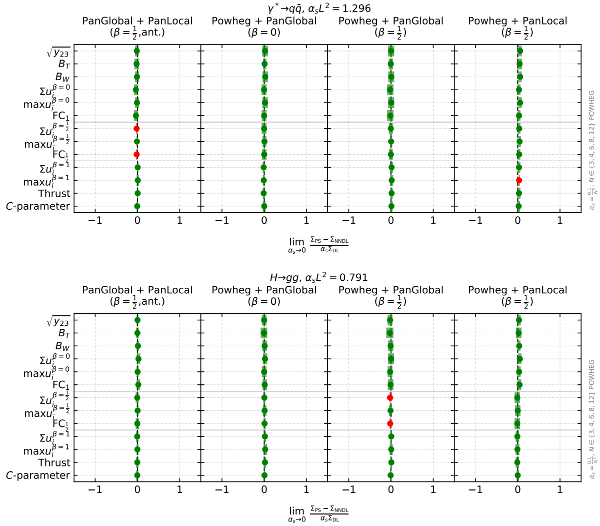


POWHEG matching without vetoes

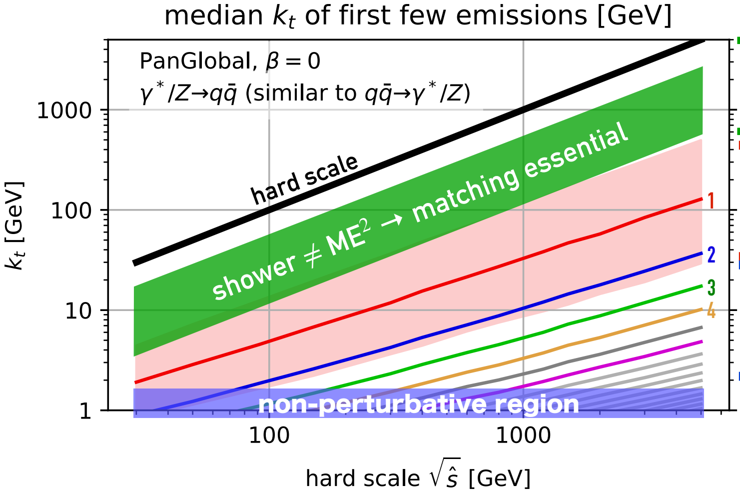




POWHEG matching with vetoes



Where is shower accuracy useful / necessary?



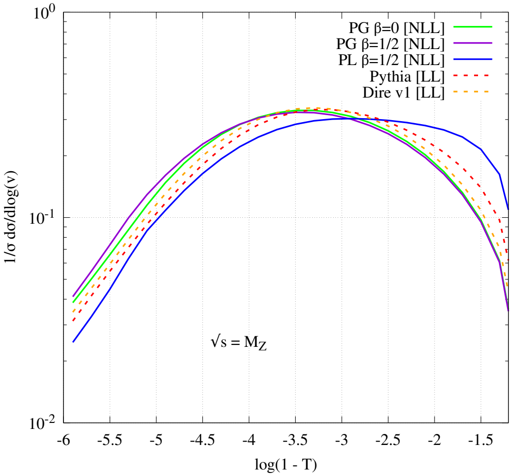
Full matrix-element needed for $k_t \gtrsim 0.1\sqrt{\hat{s}}$?

k_t of first emission (median and 68% interval)

median k_t of 2nd, 3rd, etc. emissions

the shower will be attempting to get all of these "right", together with the virtual corrections

Comparing parton showers



The PanScales showers all have the same formal NLL accuracy.

There can be a significant spread between predictions of the various showers, here shown for Thrust, indicating that spurious NNLL terms are large.

The showers in particular show a larger spread than the two not-NLL showers, Pythia and Dire, do.

Are NLL showers less accurate than LL showers??

Comparing parton showers

NO! If we include scale variations this becomes very clear. For showers that have been established to be NLL accurate, for an emission carrying away a momentum fraction z, the emission strength is taken proportional to[2]

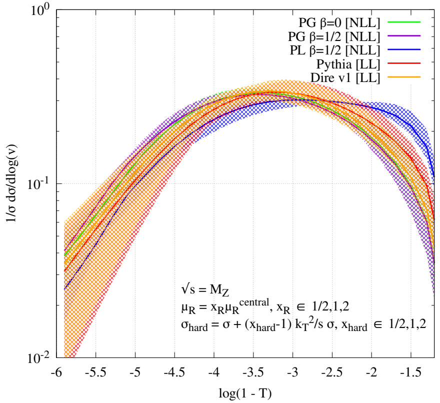
$$\alpha_S(\mu_R)\left(1+\frac{K\alpha_S(\mu_R)}{2\pi}+\frac{2(1-z)\beta_0\alpha_S(\mu_R)}{2\pi}\ln(x_R)\right), \quad \mu_R = x_R\mu_R^{\text{central}}. \tag{4}$$

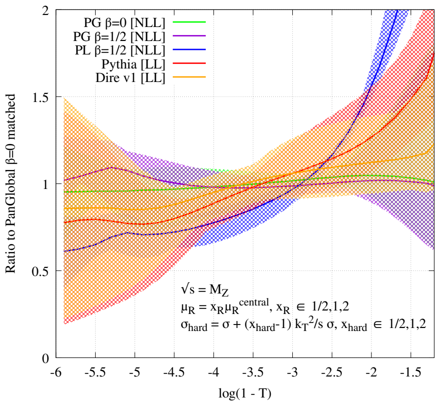
The factor $1-z$ ensures that we only apply the scale compensation in the soft limit, and not the hard where the shower does include all the necessary ingredients. For showers that are not LL we include the term proportional to K (CMW scheme) but omit the $1-z$ term.
In order to assess missing terms in the hard matching region we take the emission strengt proportional to (unless matching that emission)

$$1+(x_{\text{hard}}-1)\frac{k_T^2}{s} \tag{5}$$

[2] Inspired by Mrenna & Skands [1605.08352]

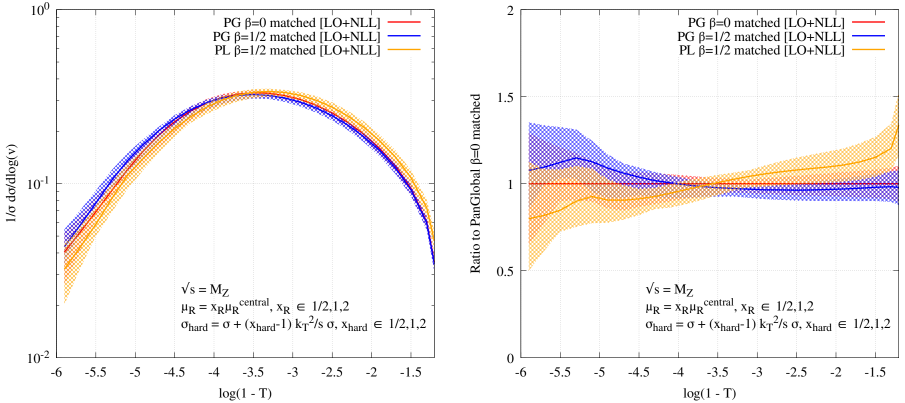
Scale variations in LL and NLL showers





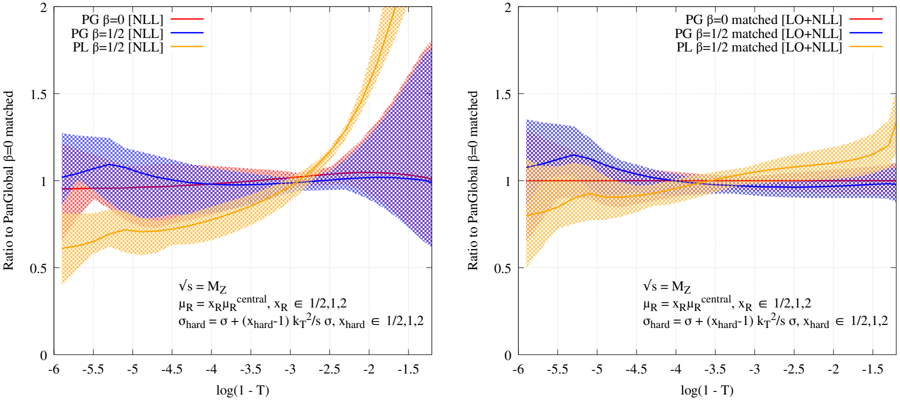
Shower variations reduced significantly in the NLL showers. Showers also almost fully inside LL shower variations. Large discrepancies in hard region expected - can be fixed by matching.

Scale variations in matched NLL showers



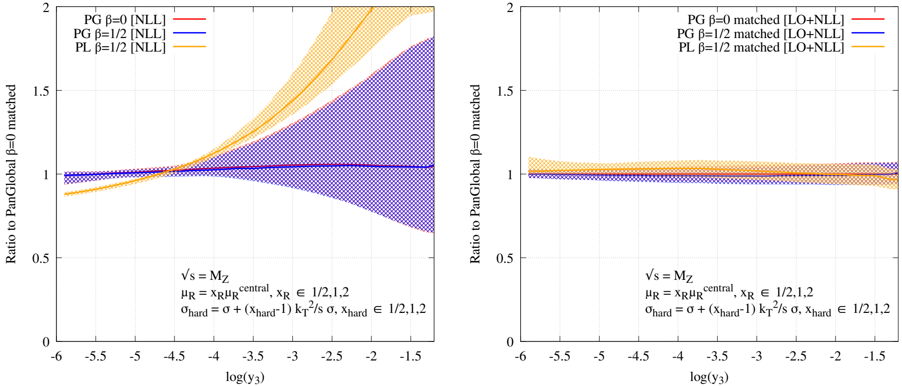
When matching LO matrix elements to the PanScales showers the agreement improves everywhere, not just in the hard region. Probably due to dominance of the first emission. Hard variations significantly decreased.

Scale variations with and without matching



When matching LO matrix elements to the PanScales showers the agreement improves everywhere, not just in the hard region. Probably due to dominance of the first emission.

Scale variations with and without matching for y_3



The agreement after matching is even better for Cambridge y_3. However, clearly scale variations not enough to cover differences between NLL showers.

Conclusions

Parton showers with controlled logarithmic accuracy are emerging.[3]

Such a program is mandatory for precision QCD studies at the LHC and future colliders.

With logaritmic control we can also assign meaningful uncertainties to shower predictions, thereby making them real predictions.

First steps towards NNLL showers are being taken, which will pave the way for unprecedented accuracy in event generator simulations.

Still many developments to come...

[3] See also recent work by Forshaw, Holguin, Plätzer (CVolver), Nagy, Soper (Deductor), Herren, Höche, Krauss, Reichelt, Schönherr (Alaric)