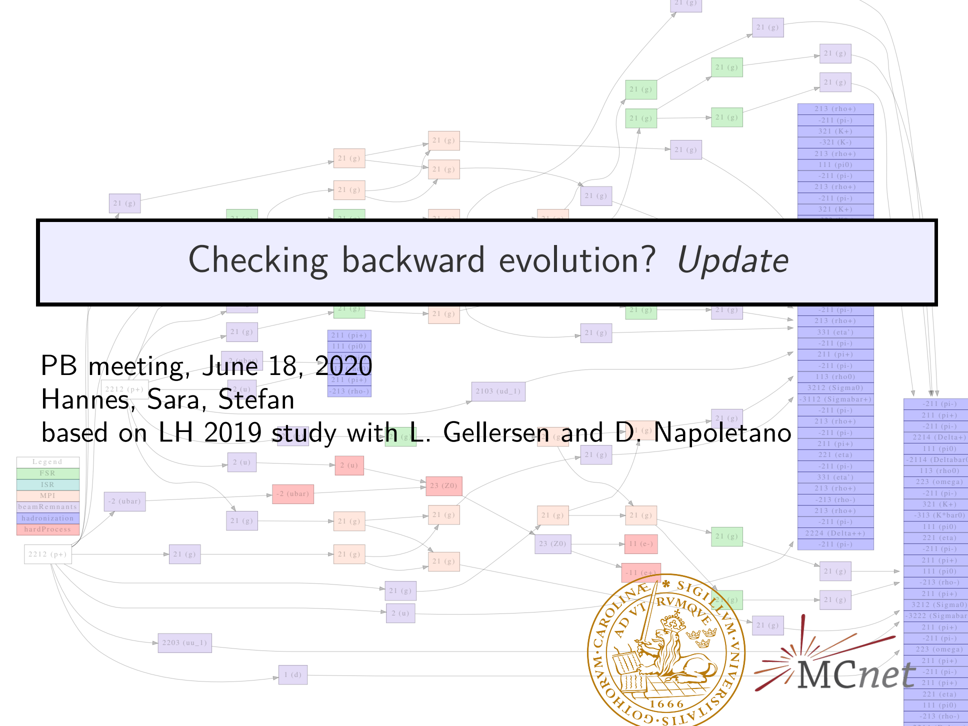


Checking backward evolution? *Update*

PB meeting, June 18, 2020

Hannes, Sara, Stefan

based on LH 2019 study with L. Gellersen and D. Napoletano

Precision MC generators require a solid understanding which assumptions to trust.

Any self-respecting generator will “amend” DGLAP evolution to avoid soft double-counting.

But how bad are the consequences? Do inconsistencies point to improvements?

[Note: Derivation following arXiv:1506.05057]
DGLAP evolution equation:

$$\frac{df_a(x,t)}{d\ln t} = \sum_{b=q,g} \int_x^1 \frac{dz}{z} \frac{\alpha_s}{2\pi} \left[P_{ba}(z)\right]_+ f_b(x/z,t) \; ,$$

where P_{ab} = regularized kernels. Assume: Write using unregularized kernels $\hat{P}_{ab}$, restricted to all but an ε-environment around the pole, plus an endpoint:

$$P_{ba}(z,\varepsilon) = \hat{P}_{ba}(z)\,\Theta(1-z-\varepsilon) - \delta_{ab}\,\frac{\Theta(z-1+\varepsilon)}{\varepsilon} \sum_{c=q,g} \int_0^{1-\varepsilon} d\zeta\,\zeta\,\hat{P}_{ac}(\zeta)$$

w/o momentum conservation, $\varepsilon \to 0$ is possible, allowing identification of $\left[P_{ba}(z)\right]_+$ as the $\varepsilon \to 0$ limit of $P_{ba}(z,\varepsilon)$.

DGLAP II *(repetition)*

With this, the DGLAP evolution equation becomes:

$$\frac{1}{f_a(x,t)}\,\frac{df_a(x,t)}{d\ln t} = -\sum_{c=q,g}\int_0^{1-\varepsilon} d\zeta\,\zeta\,\frac{\alpha_s}{2\pi}\hat{P}_{ac}(\zeta) + \sum_{b=q,g}\int_x^{1-\varepsilon}\frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,\hat{P}_{ba}(z)\,\frac{f_b(x/z,t)}{f_a(x,t)}\;.$$

The first term is the “virtual” part, the second a “spectrum” contribution.

(Note: We can argue about the precise form of the “virtual” part, but that’s not really important here. The relevant point is $\varepsilon \ll 1$.)

DGLAP III *(repetition)*

Defing the Sudakov form factor Δ and the no-emission probability Π,

$$\begin{aligned}
\Delta_a(t_1,t_0) &= \exp\left\{ -\int_{t_1}^{t_0}\frac{dt}{t}\sum_{c=q,g}\int_0^{1-\varepsilon} d\zeta\,\zeta\,\frac{\alpha_s}{2\pi}\hat{P}_{ac}(\zeta)\right\} \\
\Pi_a(t_1,t_0;x) &= \exp\left\{ -\int_{t_1}^{t_0}\frac{dt}{t}\sum_{b=q,g}\int_x^{1-\varepsilon} \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,\hat{P}_{ba}(z)\,\frac{f_b(x/z,t)}{f_a(x,t)}\right\}
\end{aligned}$$

the DGLAP equation can be rewritten as

$$f_a(x,t)\Delta_a(t,\mu^2) = f_a(x,\mu^2)\,\Pi_a(t,\mu^2;x)\ .$$

The assumptions going into this are

1. P are the DGLAP kernels
2. $\varepsilon \ll 1$

DGLAP IV *(repetition)*

The DGLAP equation can be rewritten as

$$f_a(x,t)\Delta_a(t,\mu^2) = f_a(x,\mu^2)\,\Pi_a(t,\mu^2;x)\ .$$

The assumptions going into this are

1. We use DGLAP kernels (for DGLAP evolution)
2. $\varepsilon \ll 1$

The evolution can be produced by multiple iteration[1]. Parton showers rely on sampling the spectrum (i.e. $\ln(\Pi)\otimes\Pi$) to produce evolution. No other backward evolution schemes have passed the test of time.

Most showers do not use vanilla DGLAP kernels (due to having to avoid soft double-counting), nor have $\varepsilon \ll 1$ in all regions of phase space, due to momentum conservation constraints. Angular ordering tends to lead to larger values of ε.

How much is the relation violated in practise, for common showers?

[1] cf. arXiv:hep-ph/0312355

Note that the function

$$D_a(t, \mu^2; x) = \frac{f_a(x, \mu^2)}{f_a(x, t)} \Pi_a(t, \mu^2; x) \ .$$

should be x-independent for all a, and every combination t, μ^2.

Note that we don't really need to agree that $D_a = \Delta_a$, nor do we need to worry about the precise value of D_a yet.
Since the value of D_a depends heavily on t, μ^2 and the guts of the PS (ordering, etc.), it's best to normalize

$$d(t, \mu^2; x_i) = \frac{D(t, \mu^2; x_i)}{\sum_j D(t, \mu^2; x_j)} = \frac{\frac{f_a(x_i, \mu^2)}{f_a(x_i, t)} \Pi_a(t, \mu^2; x_i)}{\sum_j D(t, \mu^2; x_j)} \ ,$$

to compare the x-independence for various t on equal footing

We want to check the x-independence of

$$d(t,\mu^2;x_i) = \frac{D(t,\mu^2;x_i)}{\sum_j D(t,\mu^2;x_j)} = \frac{\frac{f_a(x_i,\mu^2)}{f_a(x_i,t)}\,\Pi_a(t,\mu^2;x_i)}{\sum_j D(t,\mu^2;x_j)}\ ,$$

Expectations:

1. Large separation between t,μ^2 gives long evolution, which could lead to small inconsistencies accumulating.
2. Usually $\varepsilon = \varepsilon(\bar{t}, m_{\mathrm{D}})$, where m_{D} is the dipole mass. For Pythia

$$\varepsilon_{\mathrm{Pythia}} = \frac{\sqrt{t_{\mathrm{Pythia}}}}{m_{\mathrm{D}}}\left(\sqrt{1+\frac{t_{\mathrm{Pythia}}}{4\,m_{\mathrm{D}}^2}} - \frac{\sqrt{t_{\mathrm{Pythia}}}}{2\,m_{\mathrm{D}}}\right)$$

Smaller m_{D} thus allows ε to be closer to 1, and violate the assumptions more severely.

We want to check the x-independence of

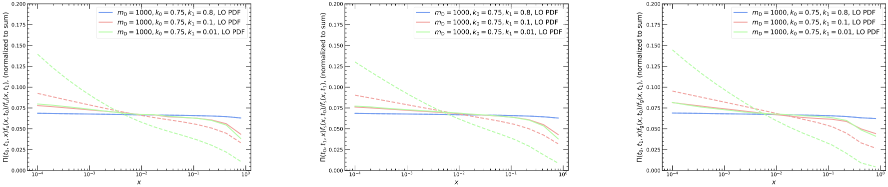
$$d(t, \mu^2; x_i) = \frac{D(t, \mu^2; x_i)}{\sum_j D(t, \mu^2; x_j)} = \frac{\frac{f_a(x_i, \mu^2)}{f_a(x_i, t)} \Pi_a(t, \mu^2; x_i)}{\sum_j D(t, \mu^2; x_j)} \; ,$$

Settings:

- For simplicity, stick to initial-initial color connection (=Drell-Yan-like)
- Pick a value for m_{D}, and then define $t_0 = \sqrt{k_0} m_{\mathrm{D}}$. We fixed $k_0 = 0.75$.
- Define $t_1 = k_1 t_0$. Use $k_1 \in [0.8, 0.1, 0.01]$ as proxy for short, moderate and very long evolution.
- Eventually check various PDFs

...then calculate $D_a(t, \mu^2; x)$ for several values of x (by employing trial showers for "events" with two partons a, b with $(p_a + p_b)^2 = m_{\mathrm{D}}^2$, $x_a = x$, and only allowing radiation from the a-leg)

Plots from LH proceedings, high mass



(a) d-quark evolution (b) s-quark evolution (c) gluon evolution

Figure: x-distribution for different length of parton-shower evolution, for $m_{\mathrm{D}} = 1000$ GeV, leading-order PDF set NNPDF23_lo_as_0119_qed, and for both Pythia (solid curves) and Sherpa (dashed curves)

Planning for extensions: New code strategy

LH plots were generated by

- Code new PS routines to work independent of hard scattering.
- Allow direct user access to core PS routines
- Set up state (with fixed x, m_D) before PS "by hand" internally, to avoid generation bias.
- Calculate $D_a(t, \mu^2; x)$ for a few x values

$\Rightarrow$ **Very intrusive method**

Would be better to have a less intrusive method

- Generate unbiased sample of input events (one incoming parton at variable x_a, one at fixed x_b)
- Feed input to shower
- Use "`UserHooks`" to check t of emissions & fill histogram.

$\Rightarrow$ **Less intrusive, more suitable strategy for checking non-Pythia showers.**

New vs. LH proceedings, high mass

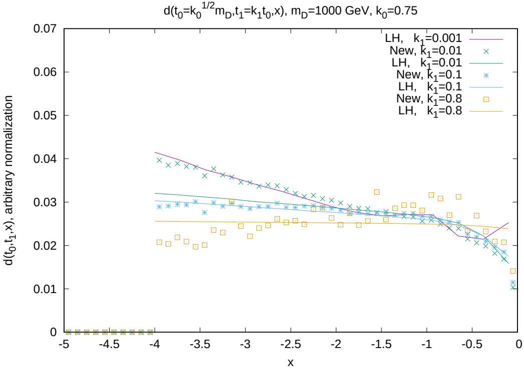


Figure: x-distribution for different length of parton-shower evolution, for $m_{\mathrm{D}} = 1000$ GeV, leading-order PDF set NNPDF23_lo_as_0119_qed, LH vs. new strategy.

$\Rightarrow$ Reasonable agreement? Same trends, but not exactly the same, especially for long evolution at small x.

New vs. LH proceedings, high mass

Maybe it's a residual bias in the input events? $\Rightarrow$ Extract "hand-crafted" hard events from LH study, use as input for new strategy.

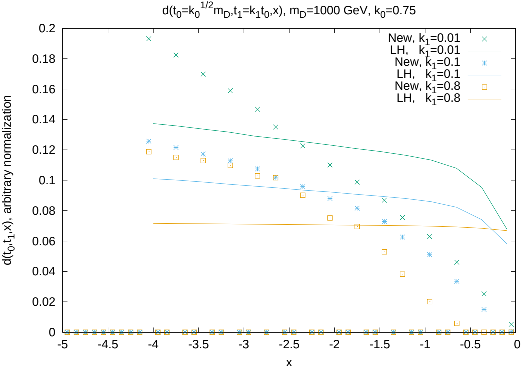


Figure: x-distribution for different length of parton-shower evolution, for $m_{\mathrm{D}} = 1000$ GeV, LH vs. new strategy.

$\Rightarrow$ Well, new idea – new bugs...

Still more to do...

Fun studies, started compaing strategies...

- Consolidated LH and Hannes' previous analysis
- Some small bugs found
- Agreement between LH (intrusive) and new strategy starts getting better, but yet perfect.
- Ideas to check: Input events.

Once new strategy is validated, many things to check:

- Different ordering variables (i.e. phase space constraints)
- Different "conventional" PDFs, parton-branching PDFs
- Different evolution equations (DGLAP, CCFM in Cascade)
- No-splitting operators (i.e. in Deductor) vs. non-Sudakov factor
- What's going on with low-x gluons? Charm?