On the reduction of negative weights in MC@NLO

Paolo Torrielli
Università di Torino and INFN

Taming the accuracy of event generators, 03/07/2020

Based on
Frederix, Frixione, Prestel, PT, 2002.12716
Outline

- Negative weights.
- Negative weights in MC@NLO.
- A new NLO+PS matching: MC@NLO-$\Delta$.
- Some results.
- Outlook / open issues.
Negative weights (I)

- Cross section beyond LO not positive definite locally in phase space.

- Not a conceptual problem, just a reduction in generation efficiency.

- Main effect: larger event samples for a given target MC accuracy.

- Generate $N = N_+ + N_-$ unweighted (up to a sign) events

  $N_- = f N$ with negative weight, $N_+ = (1 - f)N$ with positive weight, $0 \leq f < 0.5$.

  $$\sigma = \omega \left( N_+ - N_- \pm \sqrt{N_+ + N_-} \right) = \omega \left( (1 - 2f)N \pm \sqrt{N} \right).$$

- Same thing for positive-definite generation

  $$\sigma = \omega' \left( M \pm \sqrt{M} \right).$$
Negative weights (II)

\[ \sigma = \omega \left( (1 - 2f)N \pm \sqrt{N} \right) = \omega' \left( M \pm \sqrt{M} \right). \]

- To reach the same MC error
  \[ N = \frac{M}{(1 - 2f)^2}. \]

- For example: with \( f = 30\% \), which may occur for complicated processes, \( N = 6.25 \times M \).

- In experiments a significant amount of time is spent after event generation (detector simulation, ...).

- Reduction of negative weights and sample size at the price of some increase in generation time can still be beneficial.
\[ d\sigma_{\text{MC@NLO}} = d\Phi_B \, d\Phi_{\text{rad}} \left[ \frac{B + V}{\int d\Phi_{\text{rad}}} + K_{\text{PS}} \right] \mathcal{F}_{\text{PS}}^{(n)} + d\Phi_B \, d\Phi_{\text{rad}} \left[ R - K_{\text{PS}} \right] \mathcal{F}_{\text{PS}}^{(n+1)} \]

\[ = d\sigma^{(S)}(\Phi_B, \Phi_{\text{rad}}) \mathcal{F}_{\text{PS}}^{(n)} + d\sigma^{(H)}(\Phi_B, \Phi_{\text{rad}}) \mathcal{F}_{\text{PS}}^{(n+1)} \]

- \( \mathcal{F}_{\text{PS}}^{(j)} = \) shower spectrum starting from \( j \)-body kinematics

- \( S \) events. Showered by the PS starting from Born \((n\text{-body})\) kinematics \( \mathcal{F}_{\text{PS}}^{(n)} \)

- \( H \) events. Showered by the PS starting from real \((n+1\text{-body})\) kinematics \( \mathcal{F}_{\text{PS}}^{(n+1)} \)

- \( K_{\text{PS}} = \) Monte Carlo counterterm: \( \mathcal{O}(\alpha_S) \) expansion of the shower emission probability

\[ d\Phi_{\text{rad}} \, K_{\text{PS}} \propto D(q, \mu_1, \mu_2) \sum_c \sum_{\ell \in c} \frac{dq^2}{q^2} \frac{dz}{2\pi} \frac{\bar{\alpha}_S(q^2)}{P(z)} \]

\( q^2, z \) shower variables (depend on colour flow \( c \), line \( \ell \)); \( D \) dampening profile.
Negative weights in MC@NLO: $\mathbb{S}$ events

$$d\sigma^{(S)}(\Phi_B, \Phi_{rad}) \mathcal{F}_{PS}^{(n)} = d\Phi_B d\Phi_{rad} \left[ \frac{B + V}{\int d\Phi_{rad}} + K_{PS} \right] \mathcal{F}_{PS}^{(n)}$$

$\mathbb{S}$ events have Born kinematics, but $d\sigma^{(S)}$ has support in the full $n + 1$-body phase space.

Locally in the $n + 1$-body phase space their weight can be negative.

Negative $\mathbb{S}$ weights can be reduced by folding [Nason, 0709.2085] (used in POWHEG-BOX [Alioli, et al., 1002.2581])

First integrate $\mathbb{S}$ short-distance cross section over $d\Phi_{rad}$, and then generate Born phase space, i.e.

$$d\sigma^{(S)}(\Phi_B, \Phi_{rad}) \mathcal{F}_{PS}^{(n)} \rightarrow \mathcal{F}_{PS}^{(n)} \int d\Phi_{rad} \frac{d\sigma^{(S)}(\Phi_B, \Phi_{rad})}{d\Phi_{rad}}$$
Negative weights in MC@NLO: $S$-event folding

- In practice

\[
\mathcal{F}_{PS}^{(n)} \int d\Phi_{rad} \frac{d\sigma^{(S)}(\Phi_B, \Phi_{rad})}{d\Phi_{rad}} \sim \mathcal{F}_{PS}^{(n)} \sum_{i=1}^{n_\xi} \sum_{j=1}^{n_y} \sum_{k=1}^{n_\phi} \frac{w_{ijk}}{n_\xi n_y n_\phi} \frac{d\sigma^{(S)}(\Phi_B, \xi_i, y_j, \phi_k)}{d\Phi_{rad}}
\]

- At a fixed Born phase-space point ($\Phi_B$) one generates $n_\xi \times n_y \times n_\phi$ radiative configurations ($\xi, y, \phi$ are the FKS variables for the radiative phase space).

Possible using the MINT integrator \cite{Nason,0709.2085}.

- The more the one averages over radiative variables, the more the negative contributions are reduced (dominated by Born).

- Price to pay: increase running time in generation of $S$ events (typically a factor up to $n_\xi \times n_y \times n_\phi$).

- This is a purely technical aspect, i.e. no change in the matching formula.
Negative weights in MC@NLO: $H$ events

$$d\sigma^{(H)}(\Phi_B, \Phi_{rad}) \mathcal{F}^{(n+1)}_{PS} = d\Phi_B d\Phi_{rad} \left[ R - K_{PS} \right] \mathcal{F}^{(n+1)}_{PS}$$

- Sketch for $p_T$ of the Born-level system

Before showering (unphysical)

After showering (physical)

MC@NLO spectrum
Negative $H$ weights come mainly from small/intermediate $p_T$ configurations.

This region is PS-dominated: efficiently filled by $S$ events after showering.

In this region, $H$ events have little impact on shapes of distributions, mostly affect normalisation.

MC@NLO is allowing many negative $H$ events there at short-distance level, which are eventually totally compensated by the shower.

A new matching scheme, MC@NLO-$\Delta$, to reduce this problem in the first place.
MC@NLO-Δ (I)

- Main reasoning behind MC@NLO-Δ: suppress $R - K_{PS}$ at small $p_T$.

- Suppression factor $0 \leq \Delta \leq 1$, with support in the $n+1$-body phase.

- $\Delta$ designed not to spoil any of the MC@NLO accuracy properties.

- $\Delta$ constructed with sole PS information, can be used to enrich the NLO – PS cross talk.
\[ d\sigma_{\text{MC@NLO}} = d\sigma^{(S)}(\Phi_B, \Phi_{\text{rad}}) F_{PS}^{(n)} + d\sigma^{(H)}(\Phi_B, \Phi_{\text{rad}}) F_{PS}^{(n+1)} \]

\[ \Delta \rightarrow 0 \text{ in the S/C region, to dampen } H \text{ events there.} \]

\[ \Delta \rightarrow 1 \text{ in the hard region, to preserve exact NLO matrix-element information there.} \]
Construction of $\Delta$ factor

- $\Delta$ is constructed as the product of no-emission probabilities associated with all QCD legs present at Born level

\[
\Delta = \prod_{i=1}^{n} \Pi_i(q, \mu)
\]

\[
= \prod_{i=1}^{n} F_i \exp \left\{ -\frac{1}{N_{\ell_i}} \sum_{\ell_i} \sum_j \int_{q^2}^{\mu^2} \frac{dt}{t} \frac{\bar{\alpha}_S(t)}{2\pi} \int_{\epsilon(t,\ell_i)}^{1-\epsilon(t,\ell_i)} dz \frac{1}{2} P_{ji}(z) \right\}.
\]

- Suppression factors for all potential Born-level radiators.

- Stronger suppression where the PS is expected to radiate more.

- $F_i = f_i(x, q) / f_i(x, \mu)$ for $i$ in the initial state, $F_i = 1$ in the final state.

- Gluons enter two dipoles: we take a (PS-driven) weighted average of the two contributions where the contribution with the smaller scale $q$ has larger weight.

- One needs enforce $\Delta = 1$ if $q > \mu$, or in a PS dead zone.

- $\Delta$ provided numerically by the PS, for each phase-space point.
Choice of scales entering the $\Delta$ factor (I)

$$\Delta = \prod_{i=1}^{n} \Pi_i(q, \mu)$$

- ‘Starting’ scales $\mu$ are typical hard scales associated with the kinematics of the process.

- Pick randomly a Born-level colour flow, depending on its relative contribution to $K_{PS}$.

- Compute reference scales $M_{ab}^2 = (\bar{k}_a + \bar{k}_b)^2$, with $\bar{k}$ = Born-level momenta $ab$ colour-connected partons in the picked flow.

- One shower starting scale per dipole end: $\mu_{ab} = D^{-1}(r_{ab}, \mu_1, ab, \mu_2, ab)$

  $r_{ab} =$ flat random numbers, and $\mu_{i,ab} = f_i M_{ab}$

  $f_{1,2} = O(1)$ factors to assess shower-scale systematics.
Choice of scales entering the $\Delta$ factor (II)

\[ \Delta = \prod_{i=1}^{n} \Pi_i(q, \mu) \]

- ‘Stopping’ scales $q$ must be $q \to 0$ for soft/collinear radiation ($\Delta \to 0$), and $q \approx \mu$ for hard radiation ($\Delta \to 1$).

- We passe Born kinematics and colour flow + real kinematics to the PS.

- PS returns $q_{ab} =$ shower variable associated with the real radiation occurring from radiator $a$, colour-connected to $b$; in general $q_{ab} \neq q_{ba}$.
Remarks

- Parton shower called at run time (and not only after event generation) to get information on
  - no-emission probabilities $\rightarrow$ pre-tabulated shower Sudakovs
  - target scales $q_{ij}$

- Cross-talk may become beneficial in MadGraph5_aMC@NLO: may lead to abandon hard-coded $K_{PS}$ and gain flexibility.

- LH events can be naturally endowed with two scales per colour dipole $\mu_{ij}$ for $S$ events, $q_{ij}$ for $H$ events.

- Especially relevant for complex multi-leg processes, where one shower starting scale per event may be suboptimal.

```
<event>
  ....
  <scales muf='1.0E+01' mur='1.0E+01' ... scalup_a_b='X' ...>
  </scales>
</event>
```
Results - setup

- LHC at 13 TeV.

- NNPDF2.3 with $\alpha_S(M_Z) = 0.119$.

- Central $\mu_R, \mu_F = H_T/2 = \frac{1}{2} \sum_i \sqrt{m_i^2 + p_{T,i}^2}$.

- MadGraph5_aMC@NLO interfaced with PYTHIA8.

- No hadronisation, no underlying event, no QED showers.
### Results (I)

|                  | MC@NLO     |         | MC@NLO-Δ   |         |
|------------------|------------|---------|------------|---------|
|                  | 111  221   | 441     | Δ-111  Δ-221| Δ-441   |
| $pp \rightarrow e^+e^-$ | 6.9% (1.3) | 3.5% (1.2) | 3.2% (1.1) | 5.7% (1.3) | 2.4% (1.1) | 2.0% (1.1) |
| $pp \rightarrow e^+\nu_e$ | 7.2% (1.4) | 3.8% (1.2) | 3.4% (1.2) | 5.9% (1.3) | 2.5% (1.1) | 2.3% (1.1) |
| $pp \rightarrow H$         | 10.4% (1.6)| 4.9% (1.2)| 3.4% (1.2)| 7.5% (1.4)| 2.0% (1.1)| 0.5% (1.0)|
| $pp \rightarrow Hb\bar{b}$ | 40.3% (27) | 38.4% (19) | 38.0% (17) | 36.6% (14) | 32.6% (8.2)| 31.3% (7.2)|
| $pp \rightarrow W^+j$      | 21.7% (3.1)| 16.5% (2.2)| 15.7% (2.1)| 14.2% (2.0)| 7.9% (1.4)| 7.4% (1.4)|
| $pp \rightarrow W^+t\bar{t}$ | 16.2% (2.2)| 15.2% (2.1)| 15.1% (2.1)| 13.2% (1.8)| 11.9% (1.7)| 11.5% (1.7)|
| $pp \rightarrow t\bar{t}$  | 23.0% (3.4)| 20.2% (2.8)| 19.6% (2.7)| 13.6% (1.9)| 9.3% (1.5)| 7.7% (1.4)|

- ▶ 111, 221, 441, are the folding parameters $n_\xi$, $n_y$, $n_\phi$.
- ▶ Number in brackets is the ‘relative cost’ defined as $1/(1 - 2f)^2$.
- ▶ $\Delta$ + folding quite effective for $t\bar{t}$, or $Wj$.
- ▶ $Hb\bar{b}$ still challenging. Ongoing further investigation.
Results (II)

- $p_T(Born \text{ system}) = \text{maximally sensitive to matching systematics.}$

- Same shape at small $p_T$, $\mathcal{O}(5\%)$ difference in the matching region, same shape and normalisation at high $p_T$.

- Difference compatible with systematics effects from shower-scale variations (not shown).

- Folding does not affect distributions, just statistics.
Results (III)

- Similar pattern for gluon fusion Higgs production as for DY.
- For $t\bar{t}$ up to 30% difference in the matching region.
Results (IV)

- $Wt\bar{t}$: asymptotic regime approached very slowly, shower effects at hundreds of GeV.

- MC@NLO-$\Delta$ converges to NLO faster than MC@NLO.

- MC@NLO’s single shower scale may be suboptimal w.r.t. multiple scales based on the kinematics of the single dipoles.
Results (V)

- **MC@NLO vs MC@NLO-Δ:** positive (brown) and negative (orange) $H^+$ events and showered results (black) in $t\bar{t}$ production.

- **Solid:** $(f_1 = 0.1, f_2 = 1.0)$, dotted: $(f_1 = 0.1, f_2 = 0.55)$, dashed: $(f_1 = 0.55, f_2 = 1.0)$. 

---

Paolo Torrielli

On the reduction of negative weights in MC@NLO
Outlook

- Reduction of negative weights in MC@NLO achieved by folding for $S$ events, by modifying matching prescription for $H$ events: MC@NLO-$\Delta$.

- Some physics and technical benefits of MC@NLO-$\Delta$.
  - Reduction of negative weights and size of event samples
  - Better scale assignments (one scale per dipole end)
  - Reduced sensitivity to PS in the soft limit
  - Enhanced flexibility in the MadGraph5.aMC@NLO implementation

- Some drawbacks: folding and MC@NLO-$\Delta$ may increase the running time.

- I consider it as a first step of a more general revision of MC@NLO formalism, and of its implementation in MadGraph5.aMC@NLO.

- Work in progress / open questions.
  - Optimise implementation to reduce running times
  - Other ways to tackle the problem with less drawbacks? A couple of ideas to be investigated.
  - Revision of the matching scheme: a chance not only to reduce negative weights but also to include subleading terms?
  - What is the formal logarithmic accuracy of MC@NLO / MC@NLO-$\Delta$?

Thank you for your attention