Positivity in QCD predictions beyond leading order

Gavin Salam

Rudolf Peierls Centre for Theoretical Physics & All Souls College, Oxford









Science and Technology Facilities Council THE ROYAL SOCIETY





Melissa van Beekveld NIKEHF



Mrinal Dasgupta Manchester



Basem El-Menoufi Monash





Pier Monni CERN



GPS Oxford



Nicolas Schalch Oxford

PanScales A project to bring logarithmic understanding and accuracy to parton showers



Silvia Ferrario Ravasio CERN



Ludovic Scyboz Monash



Keith Hamilton Univ. Coll. London

Alba Soto-Ontoso

Granada



Jack Helliwell Monash



Alexander Karlberg CERN



Grégory Soyez IPhT, Saclay



Silvia Zanoli Oxford



Frédéric Drever



Rok Medves

Emma Slade



European Research Council Established by the European Commission

ERC funded 2018-2024















Logarithmically-accurate and positive-definite NLO shower matching

arXiv:2504.nnnnn



Melissa van Beekveld NIKEHF



Silvia Ferrario Ravasio CERN



Alexander Karlberg CERN



GPS Oxford



Jack Helliwell Monash



Monash



Alba Soto-Ontoso Granada



Grégory Soyez IPhT, Saclay



Silvia Zanoli Oxford





Logarithmically-accurate and positive-definite NLO shower matching

arXiv:2504.nnnnn



Melissa van Beekveld NIKEHF



Silvia Ferrario Ravasio CERN



Alexander Karlberg CERN



GPS Oxford



Jack Helliwell Monash



Monash



Alba Soto-Ontoso Granada



Grégory Soyez IPhT, Saclay



Silvia Zanoli Oxford





Why not just plain (N)NLO?

 $\sigma = \sigma_0 + \alpha$

LO

incredibly powerful, get scattering cross-sections from first few orders of perturbative expansion in the strong coupling α_s

$$s_s\sigma_1 + \alpha_s^2\sigma_2 + \cdots$$

NLO NNLO



What kind of contributions do we get at NLO?



KITP, March 2025

Divergences are present in both real and virtual diagrams.

They arise when an emission has a small energy ($E \ll 1$) or a small angle ($\theta \ll 1$).

In dim-reg, this brings $1/\varepsilon^2$ for each order in α_s .





LO (2-particle) tree-level event

with weight 1.00000 px, py, pz, E = -1.32 -1.38 -49.96 50.00 px, py, pz, E = 1.32 1.38 49.96 50.00



LO (2-particle) tree-level event with weight 1.00000 px, py, pz, E = -1.32 -1.38 -49.96 50.00 px, py, pz, E = 1.32 1.38 49.96 50.00

LO event $(q\bar{q})$



LO (2-particle) tree-level event with weight 1.00000 px, py, pz, E = -1.32 - 1.38 - 49.96 50.00px, py, pz, E = 1.32 1.38 49.96 50.00

NLO (3-particle) tree-level event

with weight 893.22103, multiplying (alphas/2pi) px, py, pz, E = -1.60 -1.75 -49.87 49.93 px, py, pz, E = 1.31 1.36 49.25 49.29 px, py, pz, E = 0.30 0.39 0.62 0.79

LO event $(q\bar{q})$

NLO event, with real emission ~ LO event + extra soft gluon and large positive weight



LO (2-particle) tree-level event with weight 1.00000 px, py, pz, E = -1.32 - 1.38 - 49.96 50.00px, py, pz, E = 1.32 1.38 49.96 50.00

NLO (3-particle) tree-level event with weight 893.22103, multiplying (alphas/2pi) px, py, pz, E = -1.60 - 1.75 - 49.87 49.93px, py, pz, $E = 1.31 \quad 1.36 \quad 49.25 \quad 49.29$ px, py, pz, E = 0.30 0.39 0.62 0.79

NLO (2-particle) virtual subtraction event with weight -84.49299, multiplying (alphas/2pi) px, py, pz, E = -1.32 - 1.38 - 49.96 50.00px, py, pz, E = 1.32 1.38 49.96 50.00

NLO (2-particle) virtual subtraction event with weight -808.58646, multiplying (alphas/2pi) px, py, pz, E = -1.61 - 1.75 - 49.94 50.00px, py, pz, $E = 1.61 \quad 1.75 \quad 49.94 \quad 50.00$

NLO (2-particle) virtual finite event

with weight 2.66667, multiplying (alphas/2pi) px, py, pz, E = -1.32 - 1.38 - 49.96 50.00px, py, pz, E = 1.32 1.38 49.96 50.00

LO event $(q\bar{q})$

NLO event, with real emission ~ LO event + extra soft gluon and large positive weight

NLO event, "virtual" correction ~ LO event and large negative weight



event weights are ~ probabilities

- real life doesn't have negative probabilities
- real life doesn't have (near-)divergent probabilities
- > you can evade these problems in perturbation theory if you ask very limited kinds of questions, i.e. nearly always summing real & virtual divergences (infrared safe observable, single momentum scale)*
- but experiments don't limit themselves to those kinds of questions

* though there can still be nasty surprises, cf. Chen et al 2102.07607, GPS & Slade 2106.08329









Key innovation of 2002-'04: correct or replace first step so that perturbative expansion of hard process + parton-shower is equivalent to the true NLO.

Frixione & Webber: MC@NLO hep-ph/0204244

Nason: **POWHEG** hep-ph/0409146

[>7500 citations; these methods used also in Sherpa, Herwig]











Key features of MC@NLO and POWHEG events

- ► MC@NLO and POWHEG methods, supplemented with parton showers + hadronisation models, provide NLO-accurate realistic hadron-level events
- they avoid the problem of (near) divergent event weights
- \blacktriangleright instead the event weights are just ± 1

This is a big advantage over "pure" NLO

But the event sample doesn't quite look like a true physical event sample, because there are still some negative weights

Are negative weights a problem?

Given fraction f of negative-weight events, to reach the same statistical error as for Nunit positive-weight events, you need to generate a larger number of events,

(1

E.g. for f = 15 % this doubles the required number of events.

$$\frac{N}{-2f)^2}$$

KITP, March 2025

NLO state of the art [POWHEG]



More complex processes tend to have higher fractions of negative weights. Mitigation options (e.g. folding, Nason 0709.2085) often trade off negative weight fraction v. generation time.

Problem usually worse for NNLO event generation







	step-0 (s) (grid setup)	step-1 (s) (integration)	step-2 (s) (generation)	negative S weights
$pp \rightarrow e^+e^-$				
default	1	14	147	7.1%
$2 \times 2 \times 1$ folding	1	33	258	2.1%
$4 \times 4 \times 1$ folding	1	114	781	1.8%
Born spreading	113	30	189	2.0%
$pp \rightarrow W^+ j$				
default	10	604	2013	24.2%
$2 \times 2 \times 1$ folding	10	1265	5160	13.2%
$4 \times 4 \times 1$ folding	7	2803	16020	9.0%
Born spreading	355	645	2226	18.8%

Frederix & Torrielli 2310.04160

wall times for 1 million events on a 12-core i7-8700K @ 3.7 GHz desktop machine





Some LHC experiments' statements on negative weights and machine learning

- settings, but at LO."
- potentially impossible"
- ► ATLAS <u>2412.15123</u>: "Since XGBoost [ML framework] cannot handle negativeweight events, the absolute value of each event weight is used."

► ATLAS <u>2211.01136</u>: "To avoid the use of negative weights present in the nominal NLO sample in the training of the multivariate discriminant used to separate SM $t\bar{t}t\bar{t}$ events from background [...], a sample was produced with similar generator

 \blacktriangleright CMS <u>2411.03023</u>: "However, the binary cross-entropy given by Eq. (2), can become negatively unbounded for negative event weights, making the classification task



other work trying to reduce negative weight fractions (+ further refs below)

K. Danziger, S. Höche and F. Siegert, Reducing negative weights in Monte Carlo event generation with Sherpa, 2110.15211.

J. R. Andersen and A. Maier, Unbiased elimination of negative weights in Monte Carlo samples, Eur. Phys. J. C 82 (2022) 433, [2109.07851].

J. R. Andersen, A. Maier and D. Maître, Efficient negative-weight elimination in large high-multiplicity Monte Carlo event samples, Eur. Phys. J. C 83 (2023) 835, [2303.15246].

J. R. Andersen, A. Cueto, S. P. Jones and A. Maier, A Cell Resampler study of Negative Weights in Multi-jet Merged Samples, 2411.11651.

B. Nachman and J. Thaler, Neural resampler for Monte Carlo reweighting with preserved uncertainties, Phys. Rev. D 102 (2020) 076004, [2007.11586].

within statistical uncertainties." [2303.15246]

- E.g. "We have demonstrated that the fraction of negative event weights in existing large high-multiplicity samples can be reduced by more than an order of magnitude, whilst preserving predictions for observables

are we doing our (perturbative QFT) job properly if we can't deliver guaranteed positive predictions?



3 stages of NLO event generation

- 2. Generate real radiation, e.g. extra gluon, with correct real matrix element
- 3. Let a parton shower generate all remaining perturbative emission

1. Generate "Born" event, e.g. $q\bar{q} \rightarrow Z$, with an overall NLO-correct normalisation



3 stages of NLO event generation

- 1. Generate "Born" event, e.g. $q\bar{q} \rightarrow Z$, with an overall NLO-correct normalisation
- 2. Generate real radiation, e.g. extra gluon, with correct real matrix element
- 3. Let a parton shower generate all remaining perturbative emission



* and also the KrkNLO [Jadach et al <u>1503.06849</u>] and MAcNLOPS [Nason & GPS, <u>2111.03553</u>] methods

+ $\mathcal{O}(\alpha_s)$	$\mathcal{O}(\alpha_s)$
Oevent nalisation	Generation of first emission
e negative	can be negative
e negative	always positive

KITP, March 2025



3 stages of NLO event generation

- 1. Generate "Born" event, e.g. $q\bar{q} \rightarrow Z$, with an overall NLO-correct normalisation
- 2. Generate real radiation, e.g. extra gluon, with correct real matrix element
- 3. Let a parton shower generate all remaining perturbative emission



* and also the KrkNLO [Jadach et al <u>1503.06849</u>] and MAcNLOPS [Nason & GPS, <u>2111.03553</u>] methods

KITP, March 2025



NLO Born normalisation, inclusive over subsequent branching

Born + NLO norm. = Born + 1-loop virtual + real(NLO) (LO + NLO) (LO) $\bar{B}(\Phi_{\rm B}) = B_0(\Phi_{\rm B}) + V(\Phi_{\rm B}) + \int R(\Phi_{\rm B}, \Phi_{\rm rad}) d\Phi_{\rm rad},$



relative order α_s





NLO Born normalisation, inclusive over subsequent branching



relative order α_s





How does it work in practice?

- \blacktriangleright Choose a Born phase space point Φ_R randomly

$$\bar{B}(\Phi_{\rm B}) = B_0(\Phi_{\rm B}) + V(\Phi_{\rm B}) + C_{\rm int}(\Phi_{\rm B}) + \int [R(\Phi) - C(\Phi)] d\Phi_{\rm rad},$$

 \blacktriangleright accept with probability $|B|/\max$, event weight is sign of B

> Instead of accurately evaluating the $d\Phi_{rad}$ integral, choose a random real phase space point Φ_{rad} and use that to get a "single-point Monte Carlo" estimate for the integral

relative order α_s



 $\bar{B}(\Phi_{\rm B}) = B_0(\Phi_{\rm B}) + V(\Phi_{\rm B}) + C_{\rm int}(\Phi_{\rm B}) + \int [R(\Phi) - C(\Phi)] d\Phi_{\rm rad},$



Gavin Salam

each point \equiv one Born phase space choice and one radiation phase space choice

Plotted as a function of one of the 3 real radiation phase space variables

> Result should be $1 + \mathcal{O}(\alpha_s)$

But sometimes coefficient in front of α_s is large and result is negative







 $\bar{B}(\Phi_{\rm B}) = B_0(\Phi_{\rm B}) + V(\Phi_{\rm B}) + C_{\rm int}(\Phi_{\rm B}) + \int [R(\Phi) - C(\Phi)] d\Phi_{\rm rad},$



reweighting of integration variables can help*

But still some negativeweight events and reweighting not always easy or successful

* recent proposals can be viewed as doing something similar: "Born spreading" [Frederix & Torrielli, 2310.04160] and "ARCANE" [Shyamsunda, <u>2502.08052</u>, 2502.08053]



Key question

Can you evaluate the following integral fast & reliably, and be both positive and NLO accurate

$\bar{B}(\Phi_{\rm B}) = B_0(\Phi_{\rm B}) +$

$$\int [R(\Phi) - C(\Phi)] \ d\Phi_{\rm rad},$$

KITP, March 2025

Could we just discard the negative events?

raw distribution of weights



Gavin Salam

How do you guarantee that what you are discarding is genuinely beyond the NLO order you're trying to control?

E.g. in one toy example discarding negativeweight events would give a spurious $\alpha_s^{3/2}$ contribution



Core idea: map the integral to an event-by event integer



Gavin Salam

this can be done in a way that the sum over the distribution gives the exact original answer

KITP, March 2025





Gavin Salam







Start with

- ► normalisation $n_R = 1$
- $\succ k_t$ at max allowed value
- <u>Run the following loop:</u>
- From $\ln k_t$ subtract a random amount sampled from $e^{-M/B_0 \ln 1/k_t}$
- ► With probability $r < \frac{|R C|}{M}$
 - ► if $R > C: n_R \rightarrow n_R + 1$,
 - ► if $R < C: n_R \rightarrow n_R 1$







Start with

- ► normalisation $n_R = 1$
- \succ k_t at max allowed value
- <u>Run the following loop:</u>
- From $\ln k_t$ subtract a random amount sampled from $e^{-M/B_0 \ln 1/k_t}$
- ► With probability $r < \frac{|R C|}{M}$
 - ► if $R > C: n_R \rightarrow n_R + 1$,
 - ► if $R < C: n_R \rightarrow n_R 1$







Start with

- ► normalisation $n_R = 1$
- \succ k_t at max allowed value
- <u>Run the following loop:</u>
- From $\ln k_t$ subtract a random amount sampled from $e^{-M/B_0 \ln 1/k_t}$
- With probability $r < \frac{|R C|}{M}$
 - ▶ if $R > C: n_R \rightarrow n_R + 1$,
 - ► if $R < C: n_R \rightarrow n_R 1$







Gavin Salam















- normalisation $n_B = 1$ k_t at max allowed value

 n_{R} then multiplies Born matrix element $B_{0}(\Phi_{R})$ $\langle n_B \rangle$ is exactly equal to $1 + \left[(R - C)/B_0 d\Phi_{\rm rad} \right]$



<u>p:</u>









Robust positivity with NLO accuracy (spurious terms from NNLO)



Gavin Salam

Now if you discard negative-weight events you have a guarantee that you only change the result at NNLO or beyond

Because each decrement of n_R costs a power of α_s



Robust positivity with NLO accuracy (spurious terms from N3LO)



Gavin Salam

Multiply *M*, *R*, *C* in the algorithm by factor *p* (here p = 2) Increment n_B by \pm — Algorithm gives exactly the same $\langle n_R \rangle$ Keeping only positiveweight events changes integral by just α_s^{p+1}

This is a "foundation" algorithm — can be adapted in many ways

\blacktriangleright **ESME** = Exponentiated Subtraction for Matching Events

normalisation simultaneously (more efficient, only a bit more complex)

Algorithm Stream 1 (ESME) Born + NLO rejection

- 1: Generate Born event according to B_C distribution and set $v = v_{\text{max}}$
- 2: while $v > v_{\min} \operatorname{do}$
- generate next v and Φ_2 according to Sudakov with density $\rho(v)d\ln v$, Eq. (3.6) 3:
- generate random number 0 < r < 14:
- if $C(\Phi) > R(\Phi)$ then 5:
 - if $r > C(\Phi)/M(\Phi)$: veto emission
- else if $r > R(\Phi)/M(\Phi)$: return reject event 7:
- else: accept emission and return continue shower, accept event 8:
- else 9:

6:

10:

- if $r > C(\Phi)/M(\Phi)$: veto emission
- else: accept emission and return continue shower, accept event 11:
- 12: return accept event

> Our implementation actually uses a variant that handles real emissions and NLO







Does ESME give the correct answer? \rightarrow Yes



Gavin Salam

Comparison to exact NLO (and to more standard method, here labelled dBNLO)

Ratio to exact NLO

Check of NLO coefficient







Is ESME fast enough?



- positive definite by construction
- \blacktriangleright turns out to be $\sim 4 \times$ faster than fastest of public NLO tools (POWHEG-Boxv2)*
- ► NB: this is a simple process (Drell-Yan) — timing more critical for more complex processes

* with some effort having gone into optimising it











Conclusions

- Positivity can be built into NLO calculations
- Price to pay is a change of higher terms beyond NLO
 - these higher order terms appear to be numerically modest
 - higher order terms anyway present for most observables in other NLO matching methods
 - Inderlying algorithm can be adapted to push them to arbitrary high order, e.g. for NNLO matching
- Underlying algorithms are simple, should be possible for other groups to try them out, also for more complex processes than the ones we studied
- Key step on path to making pQFT predictions simultaneously accurate and physical



Backup



Algorithm 1 General algorithm to convert NLO subtraction integral to integer

- 1: Set $n_{\rm B} = 1$ and $v = v_{\rm max}$
- 2: while $v > v_{\min} \, \mathrm{do}$
- 3:
- generate random number 0 < r < 14:
- if $r < |R(\Phi) C(\Phi)|/M(\Phi)$ then 5:
- if $R(\Phi) > C(\Phi)$: $n_{\rm B} \to n_{\rm B} + 1$ 6:
 - else: $n_{\rm B} \rightarrow n_{\rm B} 1$
- 8: return $n_{\rm B}$

 n_B then multiplies Born matrix element $B_0(\Phi_B)$ $\langle n_B \rangle$ is exactly equal to $1 + \int \frac{R - C}{D} d\Phi_{\rm rad}$ B_0

order α_s

7:

generate next v and Φ_2 according to Sudakov with density $\rho(v)d\ln v$, Eq. (3.6)

$$d\Phi_{\rm rad} \rightarrow \frac{dv}{v} d\Phi_2$$
$$\Delta(v) = \exp\left[-\int_v^{v_{\rm max}} \frac{dv'}{v'} \rho(v)\right]$$
$$\rho(v) = \int d\Phi_2 J \frac{M(\Phi)}{B_0(\Phi_{\rm B})}$$
$$M(\Phi) \ge \max[R(\Phi), C(\Phi)]$$





Algorithm 1 General algorithm to convert NLO subtraction integral to integer

- 1: Set $n_{\rm B} = 1$ and $v = v_{\rm max}$
- 2: while $v > v_{\min} do$
- 3: generate random number 0 < r < 14: 5: **if** $r < |R(\Phi) - C(\Phi)|/M(\Phi)$ **then** 6: **if** $R(\Phi) > C(\Phi)$: $n_{\rm B} \to n_{\rm B} + 1$ 7: **else**: $n_{\rm B} \to n_{\rm B} - 1$

7:
 else:

$$n_{\rm B} \rightarrow n_{\rm B} - 1$$

 8:
 return $n_{\rm B}$

 n_R then multiplies Born matrix element $B_0(\Phi_R)$ $\langle n_B \rangle$ is exactly equal to $1 + \int \frac{R - C}{D} d\Phi_{\rm rad}$ B_0

order α_s

generate next v and Φ_2 according to Sudakov with density $\rho(v)d\ln v$, Eq. (3.6)

$$d\Phi_{\rm rad} \to \frac{dv}{v} d\Phi_2$$

$$\Delta(v) = \exp\left[-\int_v^{v_{\rm max}} \frac{dv'}{v'} \rho(v)\right]$$

$$\rho(v) = \int d\Phi_2 J \frac{M(\Phi)}{B_0(\Phi_{\rm B})}$$

$$M(\Phi) \ge \max[R(\Phi), C(\Phi)]$$







the stream's default is the total event rate affected.

ESME, R > C



Figure 1: Simple illustration of the different possible actions in the two streams of the ESME algorithm with joint reals and subtractions. The actions are shown separately for the cases $R(\Phi) < C(\Phi)$ (left) and $R(\Phi) > C(\Phi)$ (right). In each case, when summing the two streams, one sees that the "accept evt" action occurs with total weight R/M. One can also verify that the contribution to the total event rate change relative to the B_C normalisation is (R - C)/M. Recall that the default action in stream 1 (2) is to accept (reject) the event if the shower scale reaches v_{\min} — only when the action is different from





Figure 11: PanScales NLL+NLO matched showers, interfaced with Pythia [103], as compared to 13 TeV QED-Born di-lepton data from the ATLAS collaboration [104]. The left-hand plot is for the di-lepton transverse momentum distribution, while the right-hand plot is for the ϕ_{η}^* variable [105], cf. Eq. (5.1). In the Pythia interface, we include Pythia's primordial transverse momentum but not hadronisation, QED effects or multi-parton interactions.







inclusive N3L0 σ uncertaities

KITP, March 2025





"Gold standard" fiducial cross section gives much worse prediction

> Why? And can this be solved?







- The zero
- \mathbb{N}^{3} LO MARCA Killer berturbation
 - theory will be fine

ð

- 0.0

100

80

40

20

$p_{\rm th}$ derivative of acceptance: white = 0

 $\frac{60}{y_{v}} + \frac{6}{2.3} + \frac{0.350}{1.37} + \frac{0.250}{1.52}$

0.5

Higgs rapidity

15

 \mathbf{X}

2.0

MMKO/

Huss et al preliminary of Higgs 202

