A Positive Resampler for Monte Carlo Events with Negative Weights

Jeppe R. Andersen (Durham), Christian Gütschow (UCL, ATLAS)
Andreas Maier (DESY, Zeuthen), Stefan Prestel (Lund)

June 4, 2020

Overview of Talk

1. Definition of the problem
2. A solution
3. Examples

The Straightforward Interpretation of Born Level

At Born level, there is a clear interpretation of the evaluation of the matrix element point by point in phase space as the local contribution to the cross section

$$\sigma = \int |\mathcal{M}|^2 \text{ Flux } \mathrm{dPS}$$

$|\mathcal{M}|^2|$ is positive definite. Clear interpretation of the **probability for each phase space point** to arise.

Furthermore LL parton showers generate positive weights only.

Possible to convey the outcome of the calculation as **probabilities of momentum configurations.**

The Problem at Higher Orders

The calculations of perturbative corrections are less straightforward[1] to interpret in terms of probabilities for momentum configurations.
Schematic pure NLO:

$$\sigma = \int (|\mathcal{M}_n^0|^2 + 2\mathrm{Re}(\mathrm{M}^0{}_n\overline{\mathrm{M}}_n^1))\ \text{Flux}\ \mathrm{dPS_n} + \int |\mathcal{M}_{n+1}|^2\ \text{Flux}\ \mathrm{dPS_{n+1}}$$

Cancellation between different phase spaces, subtractions terms, flux factor can be negative... Contribution from each phase space depends on the organisation of the cancellation of divergences (subtraction terms).

It is of course easy enough[1] to perform these calculations; the interpretation in terms of probabilities is lost, and the **results conveyed in terms of histograms**.

[1] difficult

The Problem at Higher Orders

The dependence on the specifics of the calculation is complicated further when matching to a parton shower - **additional sources of negative weights** are introduced not just by the method of matching for each multiplicity (MC@NLO, POWHEG), but also when merging multiple such calculations (MEPS@NLO, UNLOPS).

The amount of negative weights depend on the details of such procedures, and can be adjusted through some of the arbitrary/allowed choices of each formalism [see e.g. Frederix, Frixione et al.arXiv: 2002.12716]

The Problem at Higher Orders

The Problem at Higher Orders

Why negative weights could be considered a problem:

1. Phase space integration: The phase space sampler is normally optimised to sample according to the cross section, because the MC variance is the square of the weights. However, the uncertainty on a bin with $(90 - 80) = 10$ events is 13 events! Need a different phase space optimisation to reduce the uncertainty in regions with large cancellations
2. Sometimes the expensive part of the computing comes after the cross section is calculated, e.g. in detector simulation of the events. Frustrating/impossible to spend $9 + 8$ times the compute power that the cross section warrants.

Negative weight events no more unphysical than the positive weight events that they cancel. An indication of the complexity of the calculation, rather than a (separate) problem.

Restoring a Physical Interpretation of the Event Weight

Aim: **cluster events** into neighbourhoods, define their weight to be the **average contribution to the cross section** in that neighbourhood. This can then be followed by an **unweighting** procedure, the **reduces the number of events in the sample, without affecting the statistical variance.** *Positive Resampler* arXiv: 2005.09375

> *The effective cross section is given by $\sigma = \sigma_p - \sigma_n$, where σ_p is the contribution from events with positive weights, and σ_n that from negative weights. The Monte Carlo variance associated with the sample is $s_p^2/N_p + s_n^2/N_n$, where s_p^2, s_n^2 is the variance of the integrand. If we replace the sample with one that has N_s positively weighted events, the variance of this new sample is s_s^2/N_s. If therefore $N_s/(N_p + N_n)$ is similar to s_s/s_p or s_s/s_n, the variance can be unchanged.*

How High a Dimension to Choose

Idealistic case (i.e. the limit of infinitely narrow bin widths) where a differential distribution in the observable $\mathcal{O}_1$ is calculated with both positive and negative weight events in a f-body phase space region Ω_f. The differential distribution in the observable $\mathcal{O}_1$ is then constructed as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\mathcal{O}_1} = \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega_f}\,\frac{\mathrm{d}\Omega_f}{\mathrm{d}\Omega_n}\,\frac{\mathrm{d}\Omega_n}{\mathrm{d}\mathcal{O}_1},$$

where

$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega_f}$ signifies the cross section calculated in terms of the final state momenta.

$\frac{\mathrm{d}\Omega_f}{\mathrm{d}\Omega_n}$ encodes e.g. the jet clustering and is the Jacobian for the f-body phase space into the n-body phase space ($n < f$) that the observable depends on.

$\frac{\mathrm{d}\Omega_n}{\mathrm{d}\mathcal{O}_1}$ is traditionally included in the calculations by the binning of the n-body phase space in terms of $\mathcal{O}_1$.

How High a Dimension to Choose

We have for any other observable $\mathcal{O}_2$

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\mathcal{O}_2} = \frac{\mathrm{d}\Omega_f}{\mathrm{d}\mathcal{O}_2} \left\{ \left(\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega_f} \frac{\mathrm{d}\Omega_f}{\mathrm{d}\Omega_n} \frac{\mathrm{d}\Omega_n}{\mathrm{d}\mathcal{O}_1} \right) \frac{\mathrm{d}\mathcal{O}_1}{\mathrm{d}\Omega_f} \right\}.$$

All reference to the cross section is in $(\cdots)$, so cancellation of $+/-$ needs to be done just once. The rest is binning distributions.

How High a Dimension to Choose

Can it be done at all? Well, if one needs to describe just one distribution, then obviously it can be done by using the "neighbourhoods" of the bins.

If one needs to describe two distributions of n and m bins respectively, one can use the outer product of $n \times m$ bins. And so on.

Or one could bin in the four-momenta of all quantities in Ω_n that enter in the analyses.

So it can clearly be done. The question is how difficult[2] is it?

[2]easy

Specific Example

LHC@14TeV, UNLOPS merged calculations of W+0,1,2jets@NLO.
Standard Rivet analyses of `MC_WINC` and `MC_WJETS`
Distributions chosen for the Positive Resampler:

a) *W boson transverse momentum.* in bins of 5 GeV.
b) *Parton shower evolution variable t.* $\frac{d\sigma}{d\log t}$ in $\sqrt{N}$ bins, where N is the number of events in each sample.
c) *Total cross section.* All events are contained in a single bin.

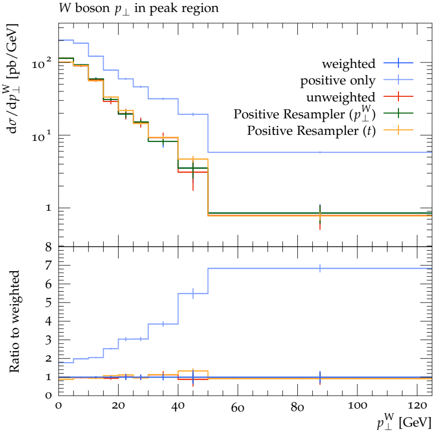
Demonstrate the viability of positive resampling independently of the concrete analysis.
In the next step, we apply the following procedure to each bin in the chosen distribution.

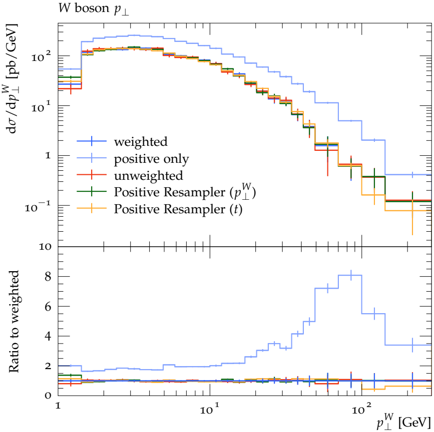
1. *Change weights.* We turn negative-weight into positive-weight events, preserving the height of each bin exactly.
2. *Partial unweighting.*
3. *Bin restoration.* We rescale the weights of all events in the bin such that the original bin height is restored.

Results

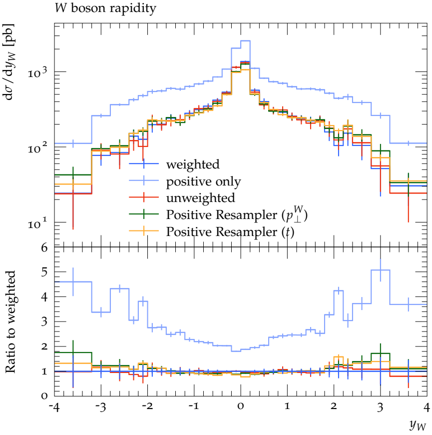
Sample	Number of events	Passing Cuts
weighted	5.3M	195k
positive only	3.2M	121k
unweighted	1.5M	52k
Positive Resampler(t)	659k	25k
Positive Resampler($p_\perp^W$)	33k	33k

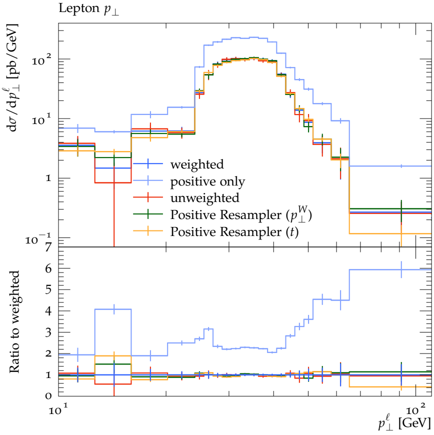
Results, Inclusive W



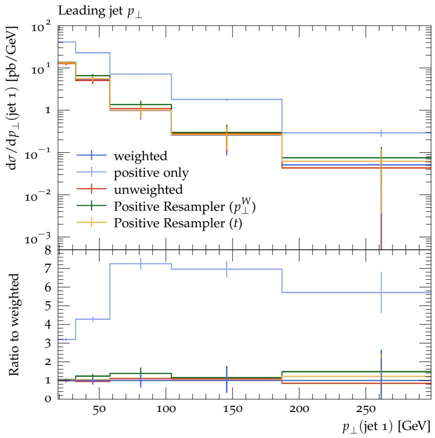


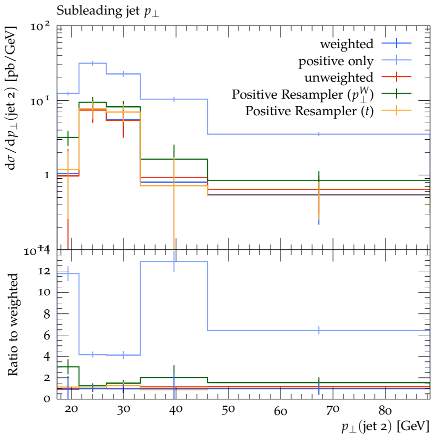
Results, Inclusive W



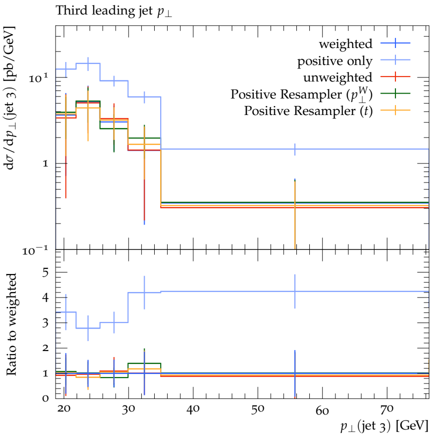


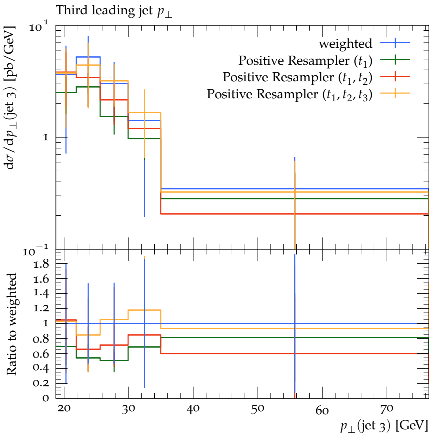
Results, W+Jets





Results, W+Jets





Conclusions

Proof by construction of the possibility of turning any event sample with weights both positive and negative into a sample where a probabilistic interpretation can be given for each configuration to arise.

The analysis given here is just an example, but a complicated one. It is straightforward[3] to generalise.

[3] It really is