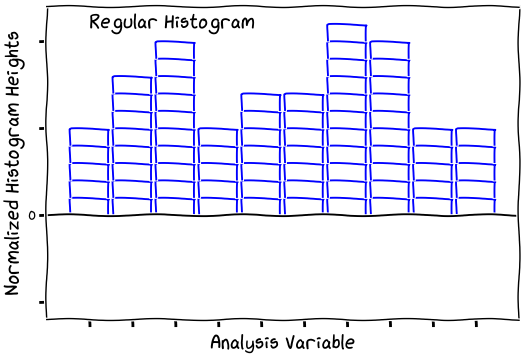
Report No.: FERMILAB-SLIDES-24-0100-ETD

ARCANE Reweighting:
A Solution to the Negative Weights Problem in Collider Monte Carlo

Prasanth Shyamsundar, Fermilab Quantum Division

DPF–Pheno 2024 (May 13, 2024)
University of Pittsburgh and Carnegie Mellon University, Pittsburgh

What is the Negative Weights Problem in Collider Monte Carlo?



Histogram height:

$$\widehat{F}_B(b) = \frac{1}{N}\sum_{i=1}^{N} W_i\,\delta_K(B_i, b)$$

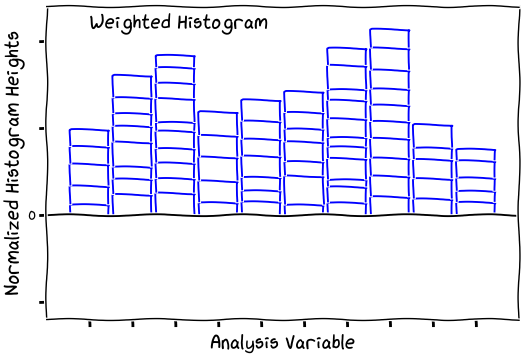
Histogram error estimate:

$$\widehat{\sigma}_B(b) = \frac{1}{\sqrt{N}}\sqrt{\frac{1}{N}\sum_{i=1}^{N} W_i^2\,\delta_K(B_i, b)}$$

$$\frac{\widehat{\sigma}_B(b)}{\widehat{F}_B(b)} = \frac{\sqrt{\sum_{i=1}^{N} W_i^2\,\delta_K(B_i, b)}}{\sum_{i=1}^{N} W_i\,\delta_K(B_i, b)}$$

- Individual events can have negative weights—artifact of our theory formalisms.
- The distribution of *any* physical (IR-safe) observable will be non-negative.
- Weight variance is bad. Variance from negative weights cannot be unweighted away.

What is the Negative Weights Problem in Collider Monte Carlo?



Histogram height:

$$\widehat{F}_B(b) = \frac{1}{N}\sum_{i=1}^{N} W_i\,\delta_K(B_i, b)$$

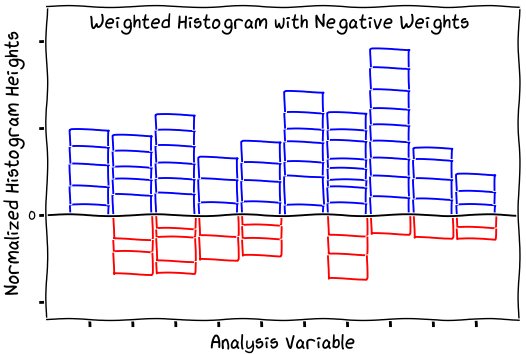
Histogram error estimate:

$$\widehat{\sigma}_B(b) = \frac{1}{\sqrt{N}}\sqrt{\frac{1}{N}\sum_{i=1}^{N} W_i^2\,\delta_K(B_i, b)}$$

$$\frac{\widehat{\sigma}_B(b)}{\widehat{F}_B(b)} = \frac{\sqrt{\sum_{i=1}^{N} W_i^2\,\delta_K(B_i, b)}}{\sum_{i=1}^{N} W_i\,\delta_K(B_i, b)}$$

- Individual events can have negative weights—artifact of our theory formalisms.
- The distribution of *any* physical (IR-safe) observable will be non-negative.
- Weight variance is bad. Variance from negative weights cannot be unweighted away.

What is the Negative Weights Problem in Collider Monte Carlo?



Histogram height:

$$\widehat{F}_B(b) = \frac{1}{N}\sum_{i=1}^{N} W_i\,\delta_K(B_i, b)$$

Histogram error estimate:

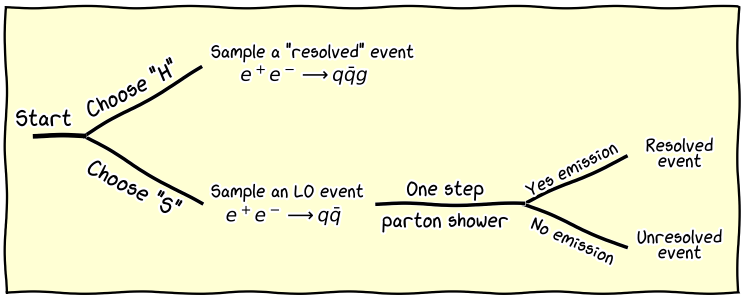
$$\widehat{\sigma}_B(b) = \frac{1}{\sqrt{N}}\sqrt{\frac{1}{N}\sum_{i=1}^{N} W_i^2\,\delta_K(B_i, b)}$$

$$\frac{\widehat{\sigma}_B(b)}{\widehat{F}_B(b)} = \frac{\sqrt{\sum_{i=1}^{N} W_i^2\,\delta_K(B_i, b)}}{\sum_{i=1}^{N} W_i\,\delta_K(B_i, b)}$$

- Individual events can have negative weights—artifact of our theory formalisms.
- The distribution of *any* physical (IR-safe) observable will be non-negative.
- Weight variance is bad. Variance from negative weights cannot be unweighted away.

Origin of the Negative Weights Problem

- Different pathways within an event generator can lead to practically identical events (for the purposes of subsequent simulation steps and analyses).
- These different pathways can provide positive or negative contributions.
- Example: $\mathbb{H}$- and $\mathbb{S}$-type events in MC@NLO.



Generator flowchart for MC@NLO $e^+e^- \to q\bar{q} + 1\,\text{jet}$ events

Phrasing the Problem in Terms of Latent Variables

$$F_V^{\mathrm{ph}}(v) = \int_{\mathcal{H}} \mathrm{d}h \, F_{(V,H)}^{\mathrm{ph}}(v, h)$$

- V: Event attributes that are "visible" to the subsequent simulation stages.
- H: Latent or hidden event attributes that are not used by subsequent stages. Example: A variable indicating whether the event was of $\mathbb{H}$- or $\mathbb{S}$- type.
- Integration wrt h represents an aggregation or marginalization over the different pathways that lead to the same V.
- Even if $F_V^{\mathrm{ph}}(v)$ is guaranteed to be non-negative, $F_{(V,H)}^{\mathrm{ph}}(v, h)$ could be negative.
- Sidenote: In practice, $F_V^{\mathrm{ph}}(v)$ itself could be negative for some values of v.

- **Traditional solution approach:** Create an alternative $F_{(V,H')}^{\mathrm{ph}'}$ with less negative weights problem. **But this is difficult in practice.**

ARCANE Reweighting

- Idea: Modify $F^{\text{ph}}_{(V,H)}$ as follows:
$$F^{\text{arcane}}_{(V,H)}(v,h) \equiv F^{\text{ph}}_{(V,H)}(v,h) + G_{(V,H)}(v,h)$$
where $G_{(V,H)}(v,h)$ integrates wrt h to 0.
- Question: How to sample weighted events from $F^{\text{arcane}}_{(V,H)}$?
- An option: Sample some events from $F^{\text{ph}}_{(V,H)}$ and some events from $G_{(V,H)}$.
But, this won't reduce the negative weights problem.
- ARCANE's approach: Incorporate $G_{(V,H)}$ with an additive reweighting:
$$\boxed{W^{\text{arcane}} \equiv W + \frac{G_{(V,H)}(V,H)}{P^{\text{mc}}_{(V,H)}(V,H)}}$$
This is the core idea of ARCANE reweighting.

ARCANE Reweighting

$$F^{\text{arcane}}_{(V,H)}(v,h) \equiv F^{\text{ph}}_{(V,H)}(v,h) + G_{(V,H)}(v,h)$$
$$W^{\text{arcane}} \equiv W + \frac{G_{(V,H)}(V,H)}{P^{\text{mc}}_{(V,H)}(V,H)}$$

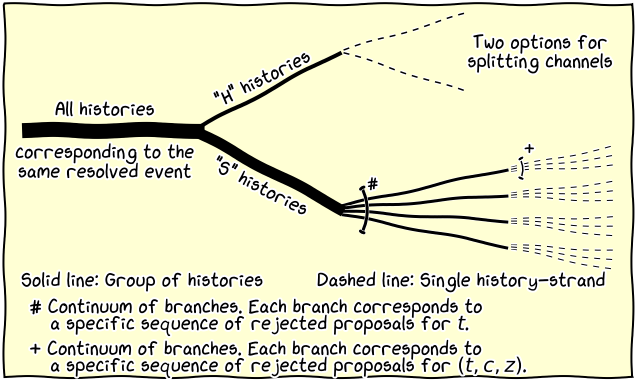
Ingredients of ARCANE Reweighting

- For each event, in addition to V, one needs to track and record H.
- One needs the ability to compute $P^{\text{mc}}_{(V,H)}$ for the different event records.
- **One needs to construct an appropriate redistribution function $G_{(V,H)}$.** $G_{(V,H)}$ can either be engineered or (machine) learned.

"So wait, you haven't actually solved anything yet?"

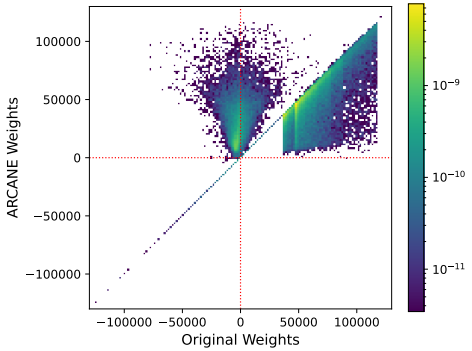
- **My claim:** It is way easier to construct a good $G_{(V,H)}$ than to construct a good alternative $F^{\text{ph}'}_{(V,H')}$ (that is suitable for performing Monte Carlo sampling with).

Sketch of Implementation for MC@NLO $(e^+e^- \longrightarrow q\bar{q} + 1\,\mathrm{jet})$ Events



- **Strategy:** Model $G_{(V,H)}$ as a chain of functions that systematically redistributes 0 across the different histories of a given "visible" event.
- My implementation of $G_{(V,H)}$ just involves performing two 1-dim polynomial fits (and an exploration of some alternative MC histories of each event).

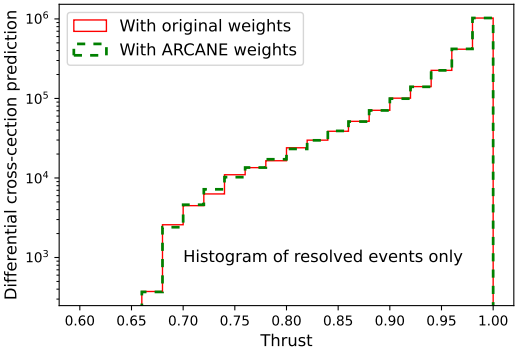
Some Results With MC@NLO $(e^+e^- \longrightarrow q\bar{q} + 1\,\mathrm{jet})$ Events



Weights of resolved events

Negative weights problem is reduced.

Note: Events along diagonal don't have contributions from both $\mathbb{H}$ and $\mathbb{S}$



Histogram of an example observable

Distributions of all "visible" observables will be unaffected.

Summary, What's Next?

Summary

- ARCANE Reweighting Can Avoid Negative Events.
- The technique does not introduce any biases.
- The technique does not require any changes to the matching and merging prescriptions used.
- Events produced by this technique can be treated as being iid.

What's Next?

- Take on more complex and important processes like $(\mathrm{pp} \longrightarrow \mathrm{W} + \mathrm{jets})$ and $(\mathrm{pp} \longrightarrow \mathrm{t\bar{t}} + \mathrm{jets})$.

Thank you!

Acknowledgments

These slides have been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

PS is supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics QuantISED program under the following grants:

- "HEP Machine Learning and Optimization Go Quantum", Award Number 0000240323
- "DOE QuantiSED Consortium QCCFP-QMLQCF", Award Number DE-SC0019219