Elimination of negative weights in MC event samples

Andreas Maier

6 September 2023

J. R. Andersen, A. Maier, D. Maître arXiv:2303.15246
J. R. Andersen, A. Maier Eur.Phys.J.C 82 (2022) 5, 433
J. R. Andersen, C. Gütschow, A. Maier, S. Prestel Eur.Phys.J.C 80 (2020) 11, 1007

What are event weights?

Leading-order cross sections

Example: prediction for dijet production cross section

1. Relate to partonic cross section

$$\sigma_{2\text{ jets}} \stackrel{\text{LO}}{=} \sigma_{2\text{ partons}}$$

2. Simulate partonic scattering events with weights w_i
   - Computed from scattering matrix elements + PDF + phase space factor
   - Weights proportional to probability: $w_i > 0$
   - Sum of weights gives the cross section:

$$\sigma_{2\text{ partons}} = \sum_i w_i$$

What are negative event weights?

Next-to-leading-order cross sections

Example: prediction for dijet production cross section

1. Relate to partonic cross section

$$\sigma_{2\text{ jets}} \overset{\text{NLO}}{=} \sigma_{2\text{ partons}} + \sigma_{3\text{ partons}}$$

2. Simulate partonic scattering events

$$\sigma_{2\text{ partons}} = \sum_i w_i$$

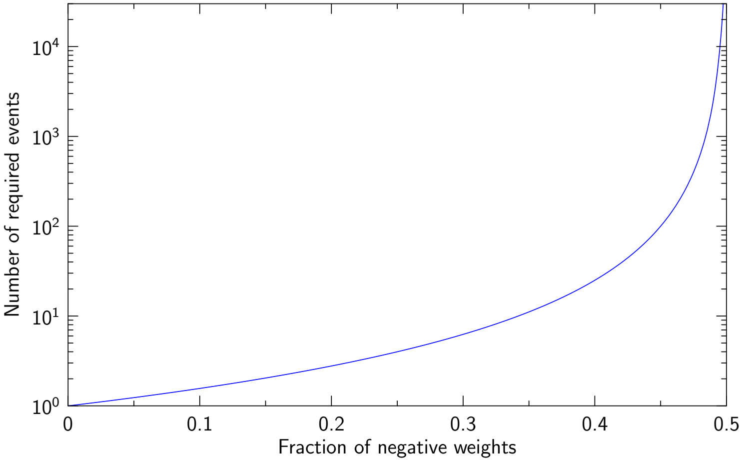
$$\sigma_{3\text{ partons}} = \sum_j w_j$$

$\sigma_{2\text{ partons}}, \sigma_{3\text{ partons}}$ not separately observable:

Events weights can be either positive or negative

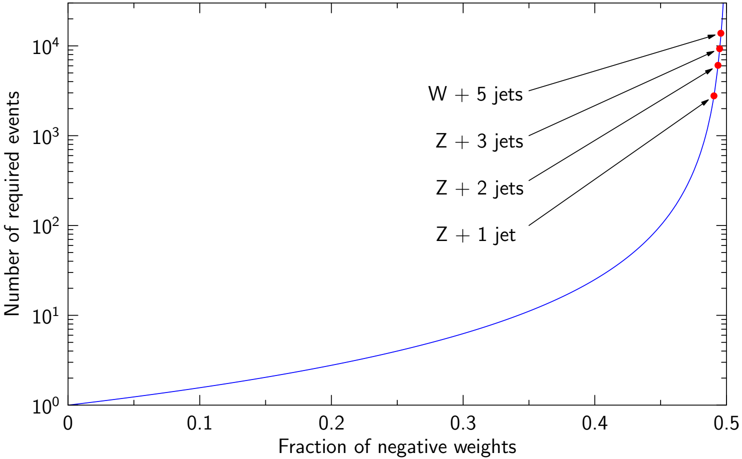
Why are negative event weights a problem?

Number of required events to reach given accuracy:



Why are negative event weights a problem?

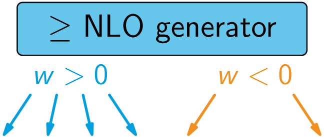
Number of required events to reach given accuracy:

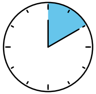


[BlackHat + Sherpa 2013 + 2017]

Why are negative event weights a problem?

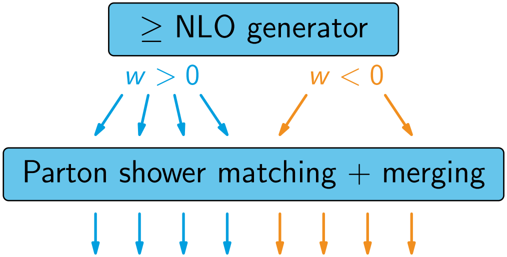
Event simulation chain

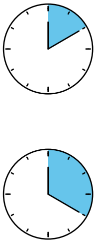




Why are negative event weights a problem?

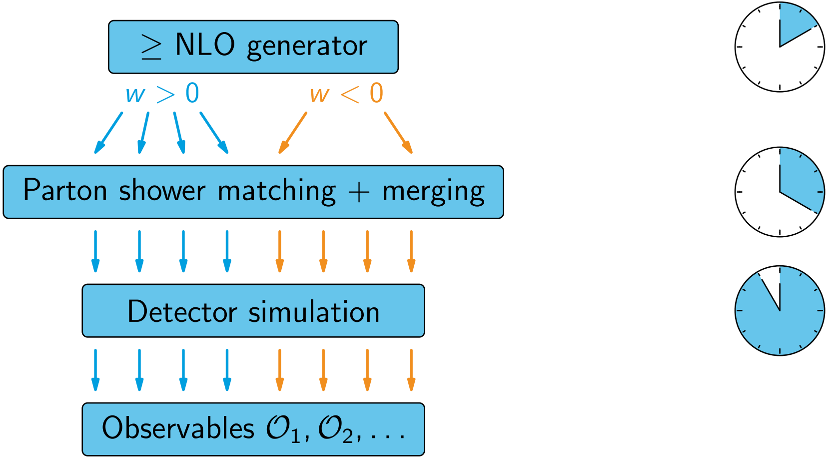
Event simulation chain





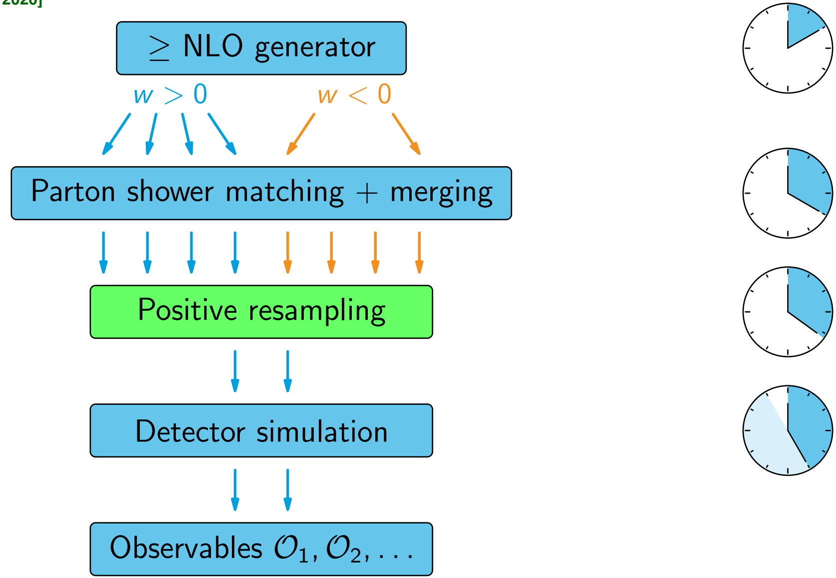
Why are negative event weights a problem?

Event simulation chain

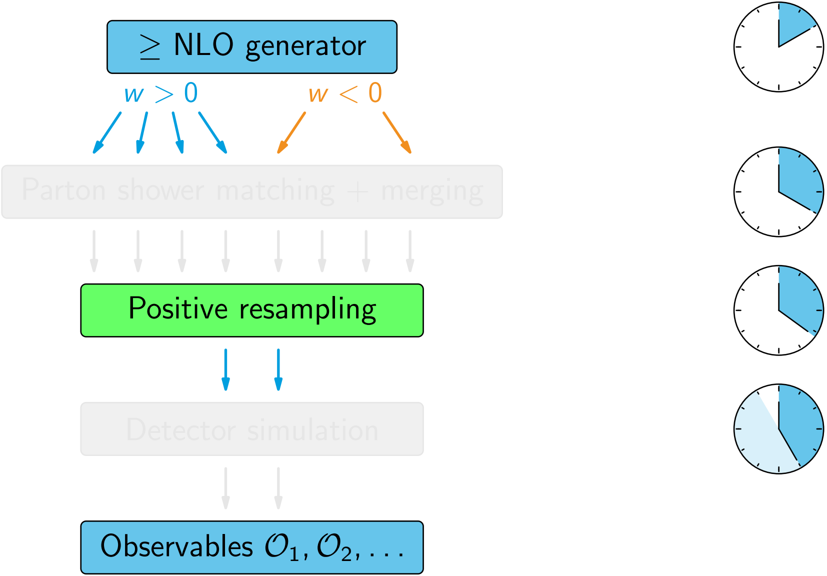


Why are negative event weights a problem?

[Andersen, Gütschow, Maier, Prestel 2020]

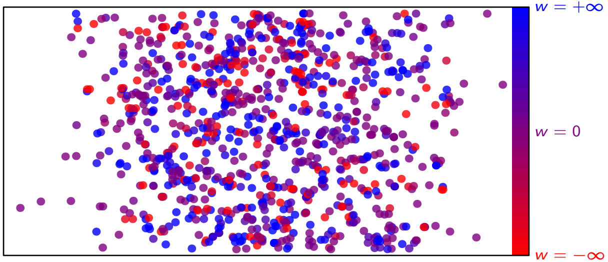


Why are negative event weights a problem?



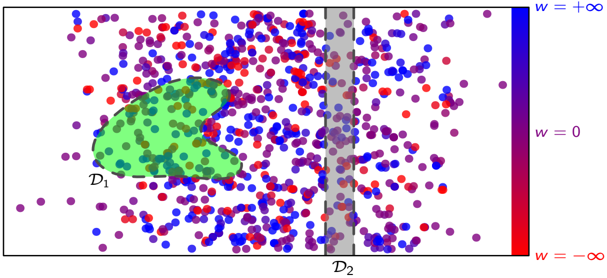
Observables

Events in 2D projection of phase space:



Observables

Events in 2D projection of phase space:



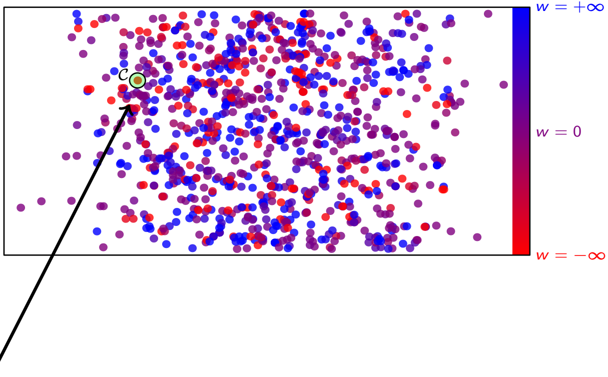
Observables $\mathcal{O}$:

- Select region $\mathcal{D}$ in phase space $\geq$ experimental resolution
- $\mathcal{O} = \sum_{i \in \mathcal{D}} w_i \geq 0$ with sufficient statistics

e.g. histogram bins

Redistribute weights without affecting any observable

Cell resampling

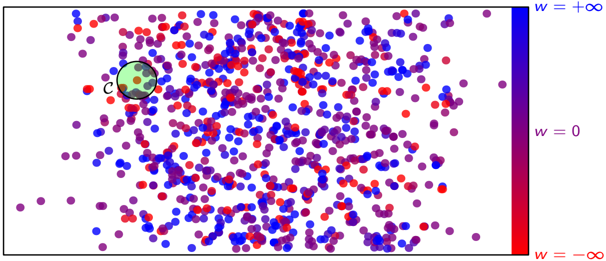


Cell resampling:

Repeatedly

1. Choose seed event with $w < 0$ for cell $\mathcal{C}$

Cell resampling

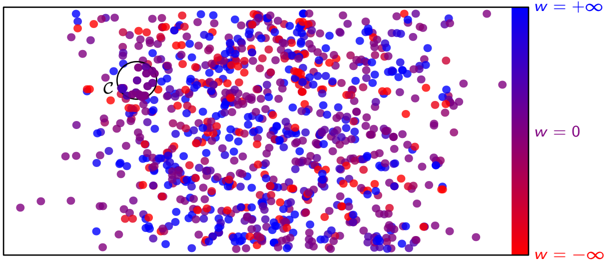


Cell resampling:

Repeatedly

1. Choose seed event with $w < 0$ for cell $\mathcal{C}$
2. Iteratively add nearest event to cell until $\sum_{i \in \mathcal{C}} w_i \geq 0$

Cell resampling

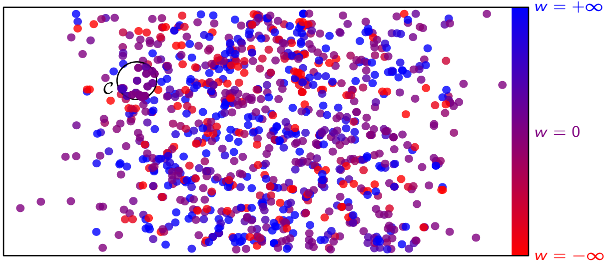


Cell resampling:

Repeatedly

1. Choose seed event with $w < 0$ for cell $\mathcal{C}$
2. Iteratively add nearest event to cell until $\sum_{i \in \mathcal{C}} w_i \geq 0$
3. Redistribute weights: $w_i \to w = \frac{\sum_{j \in \mathcal{C}} w_j}{\text{\# events in } \mathcal{C}} \geq 0$

Cell resampling



Cell resampling:

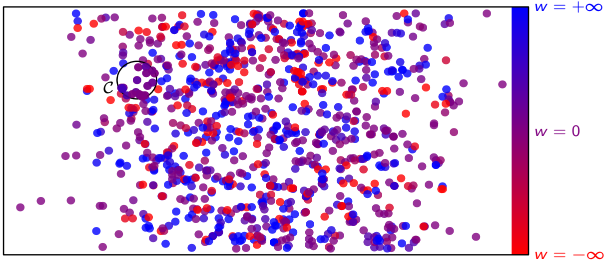
Repeatedly

1. Choose seed event with $w < 0$ for cell $\mathcal{C}$
2. Iteratively add nearest event to cell until $\sum_{i \in \mathcal{C}} w_i \geq 0$
3. Redistribute weights: $w_i \to w = \frac{\sum_{j \in \mathcal{C}} w_j}{\text{\# events in } \mathcal{C}} \geq 0$

Sufficient statistics: cell size $\ll$ experimental resolution

Otherwise: limit cell size, accept $w < 0$

Cell resampling



Cell resampling:

Repeatedly

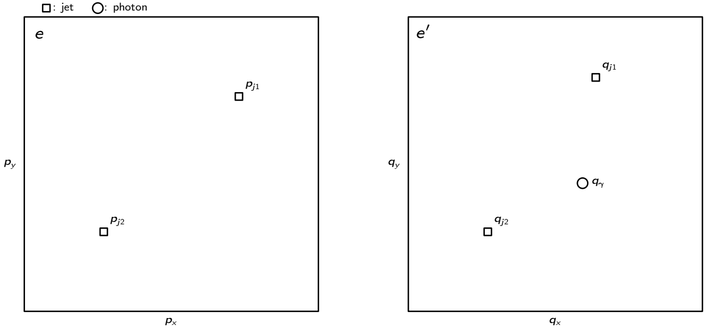
1. Choose seed event with $w < 0$ for cell $\mathcal{C}$
2. Iteratively add nearest event to cell until $\sum_{i \in \mathcal{C}} w_i \geq 0$
   What does "nearest" mean?
3. Redistribute weights: $w_i \to w = \frac{\sum_{j \in \mathcal{C}} w_j}{\text{\# events in } \mathcal{C}} \geq 0$

Sufficient statistics: cell size $\ll$ experimental resolution

Otherwise: limit cell size, accept $w < 0$

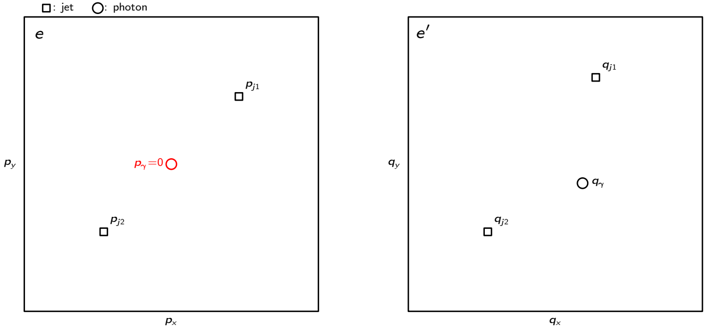
Distances in phase space

Example



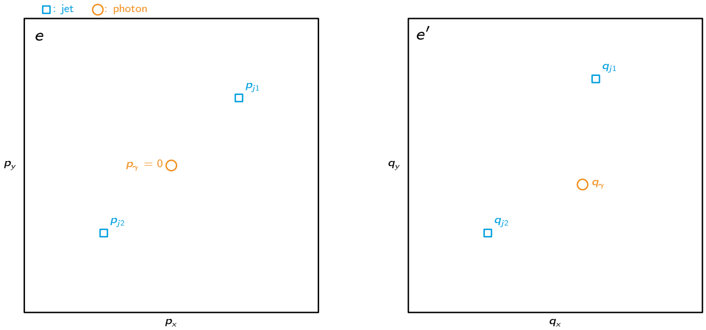
Distances in phase space

Example



Distances in phase space

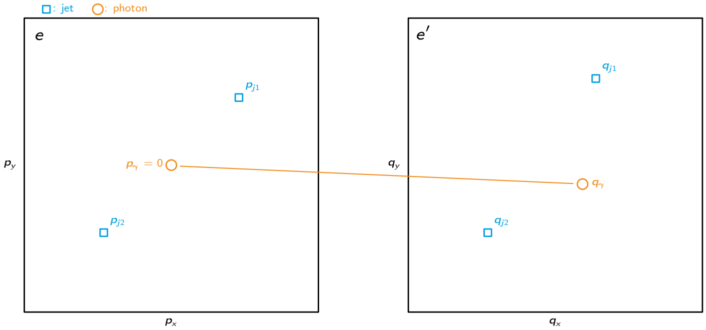
Example



$$d(e, e') = d(s_j, s_j') + d(s_\gamma, s_\gamma')$$

Distances in phase space

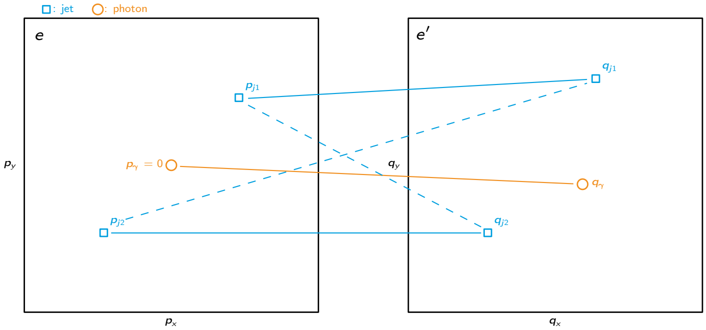
Example



$$
\begin{aligned}
d(e, e') &= d(s_j, s_j') + d(s_\gamma, s_\gamma') \\
&= d(s_j, s_j') + d(p_\gamma, q_\gamma)
\end{aligned}
$$

Distances in phase space

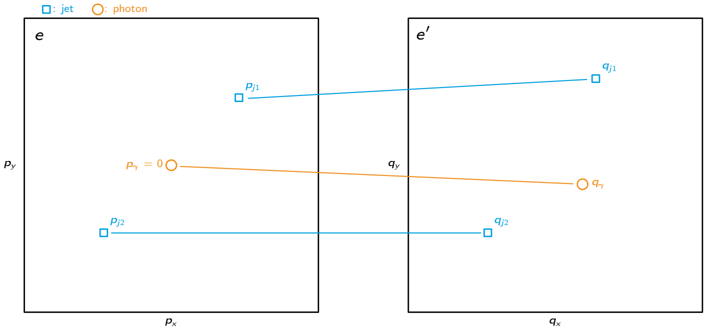
Example



$$d(e, e') = d(s_j, s_j') + d(s_\gamma, s_\gamma')$$
$$= \min\left[d(p_{j1}, q_{j1}) + d(p_{j2}, q_{j2}), d(p_{j1}, q_{j2}) + d(p_{j2}, q_{j1})\right] + d(p_\gamma, q_\gamma)$$

Distances in phase space

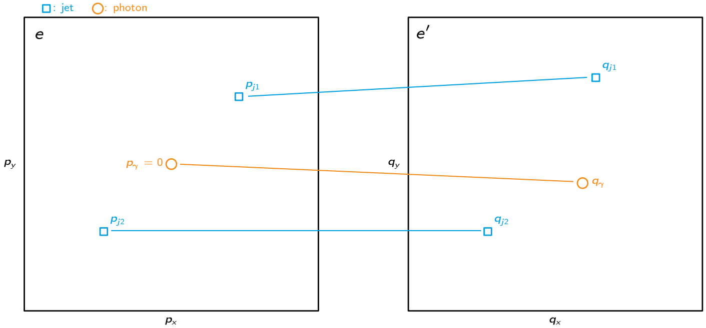
Example



$$
\begin{aligned}
d(e, e') &= d(s_j, s_j') + d(s_\gamma, s_\gamma') \\
&= d(p_{j1}, q_{j1}) + d(p_{j2}, q_{j2}) + d(p_\gamma, q_\gamma)
\end{aligned}
$$

Distances in phase space

Example



$$d(e, e') = d(s_j, s'_j) + d(s_\gamma, s'_\gamma)$$

$$\stackrel{\tau=0}{=} |\vec{p}_{j1} - \vec{q}_{j1}| + |\vec{p}_{j2} - \vec{q}_{j2}| + |\vec{p}_\gamma - \vec{q}_\gamma|$$

Distances in phase space

Concrete implementation

jets   electrons

1. Collect all infrared-safe objects in event e into sets $\{ s_1 , s_2 , \dots , s_T \}$

$$d(e, e') = \sum_{t=1}^{T} d(s_t, s'_t)$$

Distances in phase space

Concrete implementation

jets → s_1, electrons → s_2

1. Collect all infrared-safe objects in event e into sets $\{ s_1, s_2, \ldots, s_T \}$

$$d(e, e') = \sum_{t=1}^{T} d(s_t, s'_t)$$

2. Objects in s_t have four-momenta $(p_1, \ldots \ldots, p_P)$

   Objects in s'_t have four-momenta $(q_1, \ldots, q_Q, 0, \ldots, 0)$

$$d(s_t, s'_t) = \min_{\sigma \in S_P} \sum_{i=1}^{P} d_t(p_i, q_{\sigma(i)})$$

Distances in phase space

Concrete implementation

1. Collect all infrared-safe objects in event e into sets $\{ s_1 , s_2 , \ldots , s_T \}$ (jets: s_1; electrons: s_2)

$$d(e, e') = \sum_{t=1}^{T} d(s_t, s'_t)$$

2. Objects in s_t have four-momenta $(p_1 , \ldots \ldots\ldots , p_P)$

   Objects in s'_t have four-momenta $(q_1 , \ldots , q_Q, 0, \ldots , 0)$

$$d(s_t, s'_t) = \min_{\sigma \in S_P} \sum_{i=1}^{P} d_t(p_i, q_{\sigma(i)})$$

Distances in phase space

Concrete implementation

jets, electrons

1. Collect all infrared-safe objects in event e into sets $\{ s_1 , s_2 , \ldots , s_T \}$

$$d(e, e') = \sum_{t=1}^{T} d(s_t, s_t')$$

2. Objects in s_t have four-momenta $(p_1 , \ldots \ldots\ldots , p_P)$

   Objects in s_t' have four-momenta $(q_1 , \ldots , q_Q , 0, \ldots , 0)$

$$d(s_t, s_t') = \min_{\sigma \in S_P} \sum_{i=1}^{P} d_t(p_i, q_{\sigma(i)})$$

Efficient minimisation: Hungarian algorithm [Jacobi 1890]

Distances in phase space

Concrete implementation

1. Collect all infrared-safe objects in event e into sets $\{ s_1 , s_2 , \ldots , s_T \}$ (jets: s_1; electrons: s_2)

$$d(e, e') = \sum_{t=1}^{T} d(s_t, s_t')$$

2. Objects in s_t have four-momenta $(p_1 , \ldots \ldots\ldots , p_P)$
   Objects in s_t' have four-momenta $(q_1 , \ldots , q_Q , 0, \ldots , 0)$

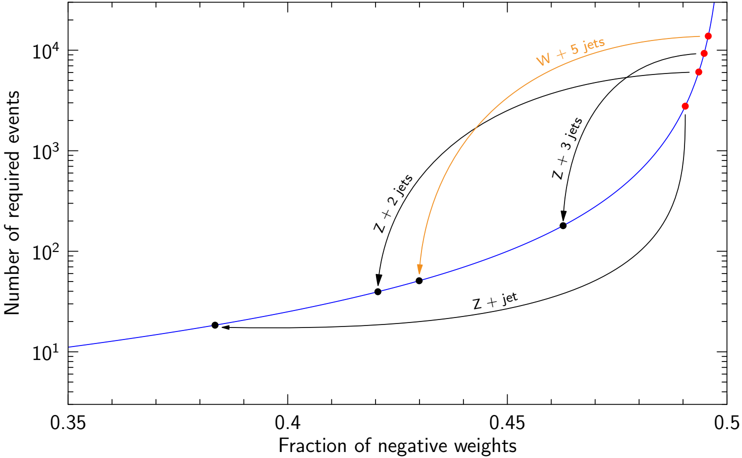
$$d(s_t, s_t') = \min_{\sigma \in S_P} \sum_{i=1}^{P} d_t(p_i, q_{\sigma(i)})$$

3. Choose distance function between particle momenta
   Here: independent of particle type t, do not consider internal structure

$$d_t(p, q) = \sqrt{(\vec{p} - \vec{q})^2 + \tau^2 (p_\perp - q_\perp)^2} \qquad \tau\text{: tunable parameter}$$

Cell resampling

Results

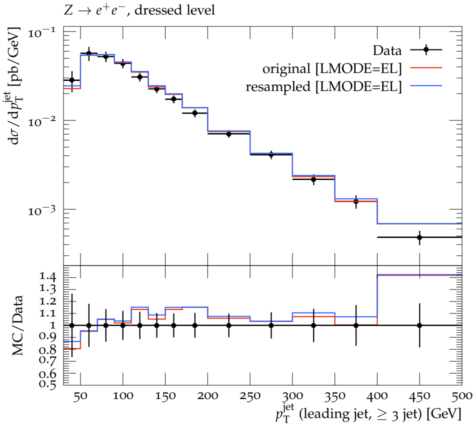


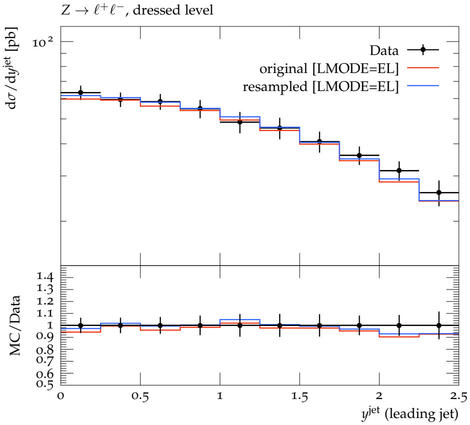
Cell resampling drastically reduces the number of required events

Cell resampling

Results

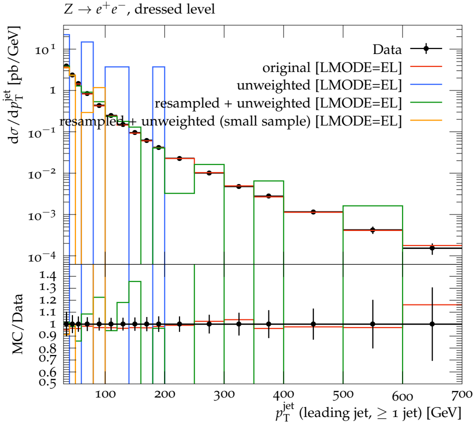
Analysis from ATLAS, Eur. Phys. J. C77 (2017) 361:

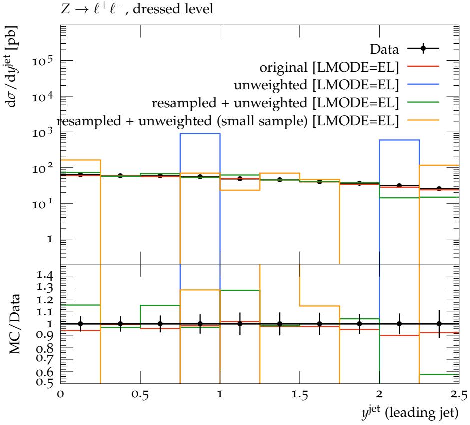




Cell resampling preserves predictions

Unweighting for Z + jet





original: 8.21×10^8 events

unweighted: 320 events

resampled + unweighted: 11574 events

resampled + unweighted (small sample): 320 events

Summary

- Negative event weights lead to slow statistical convergence
- Idea: remove negative weights by smearing over small phase space regions
  - Potential to reduce the number of required events by orders of magnitude
  - Preserves predictions of observables
  - Agnostic with respect to process and observables
  - Automatic improvement with increasing statistics
  - Computationally efficient: ~ 55 CPU hours for one billion events (W + 5 jets)

Summary

- Negative event weights lead to slow statistical convergence
- Idea: remove negative weights by smearing over small phase space regions
  - Potential to reduce the number of required events by orders of magnitude
  - Preserves predictions of observables
  - Agnostic with respect to process and observables
  - Automatic improvement with increasing statistics
  - Computationally efficient: ~ 55 CPU hours for one billion events (W + 5 jets)

Ongoing work:

- Application to parton showered samples ✓ [Andersen, Cueto, Maier, Jones]
- IRC safety with electroweak corrections [Andersen, Maier, Schönherr]
- Systematic estimate of uncertainties
- Integrate into existing workflows
- Guide Monte Carlo event generation? [Andersen, Maier, Maître, Schönherr]

Backup

Distances in phase space

Need distance function $d(e, e')$ between events e, e'

- Essential: $d(e, e')$ small $\Rightarrow$ e, e' look similar in detector or differ only in properties the event generator can't predict
- Desirable: $d(e, e')$ large $\Rightarrow$ e, e' look different in detector

Distances in phase space

Need distance function $d(e, e')$ between events e, e'

- Essential: $d(e, e')$ small $\Rightarrow$ e, e' look similar in detector or differ only in properties the event generator can't predict
- Desirable: $d(e, e')$ large $\Rightarrow$ e, e' look different in detector

Example: infrared safety

- $d(e, e')$ unaffected by collinear splittings with $\Theta \to 0$
- $d(e, e')$ unaffected by soft particles with $p \to 0$

$\Rightarrow$ define distance in terms of infrared-safe physics objects, e.g. jets

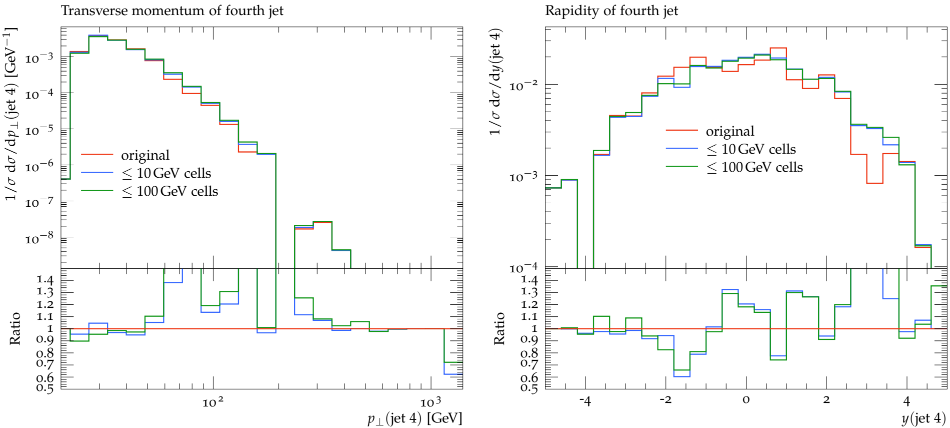
Here: Example for fixed-order (QCD) event generator

Event samples

[BlackHat 2013 + 2017]

Sample	Process	Centre-of-mass energy	# events
Z1	$pp \to (Z \to e^+e^-) + \text{jet}$	13 TeV	8.21×10^8
Z2	$pp \to (Z \to e^+e^-) + 2\ \text{jets}$	13 TeV	5.30×10^8
Z3	$pp \to (Z \to e^+e^-) + 3\ \text{jets}$	13 TeV	1.65×10^9
W5	$pp \to (W^- \to e^-\nu_e) + 5\ \text{jets}$	7 TeV	1.17×10^9

Resampling for W + 5 jets



Nearest-neighbour search