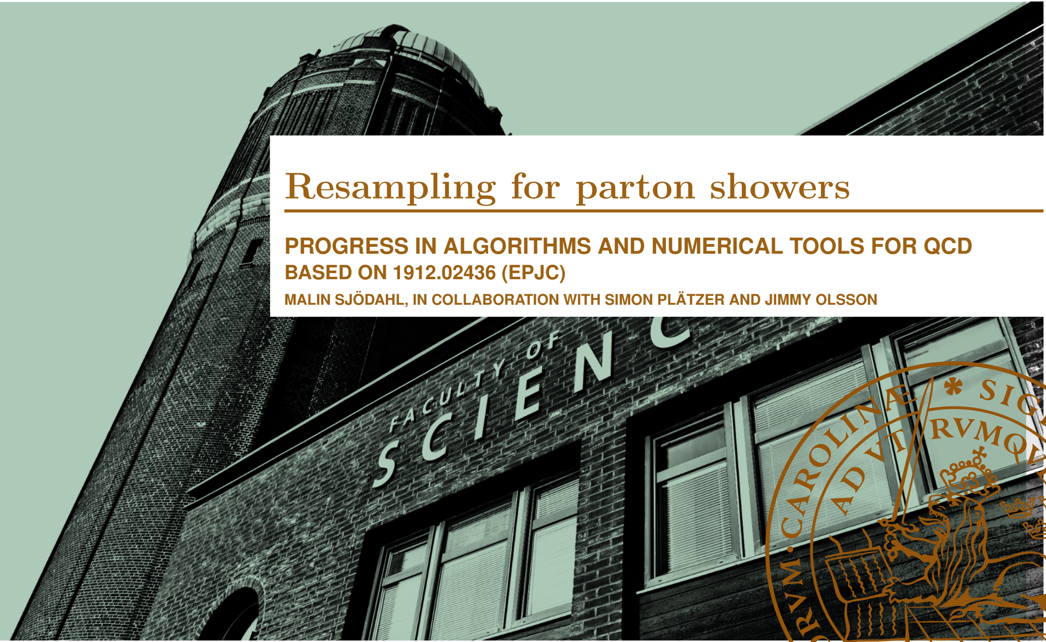
Lund University

# Resampling for parton showers

## PROGRESS IN ALGORITHMS AND NUMERICAL TOOLS FOR QCD

### BASED ON 1912.02436 (EPJC)

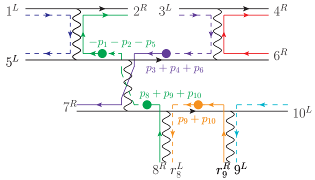
MALIN SJÖDAHL, IN COLLABORATION WITH SIMON PLÄTZER AND JIMMY OLSSON



# Introduction

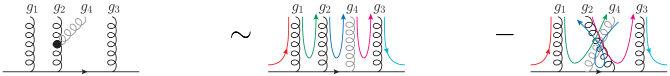
# Background

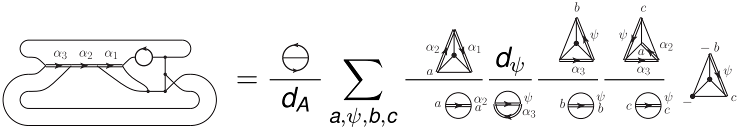
- Chirality flow



$$\sim \; [15]\langle 64\rangle[10\,9]\big(\langle r_9 9\rangle[9 r_8]+\langle r_9 10\rangle[10 r_8]\big)$$
$$\big([33]\langle 37\rangle+[34]\langle 47\rangle+[36]\langle 67\rangle\big)\big(-\langle 89\rangle[91]\langle 12\rangle$$
$$-\langle 8\,10\rangle[10\,1]\langle 12\rangle-\langle 8\,10\rangle[10\,5]\langle 52\rangle\,\big)$$

- Color, parton showers

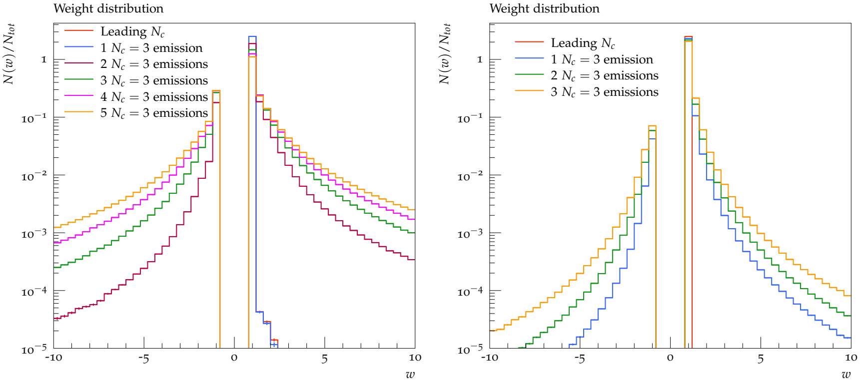




# Motivation



## Weights

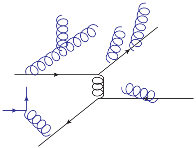



Weight distribution for $e^+e^-$ (left) and $pp$ collisions (right) depending on the number of $N_c = 3$ emissions allowed 1912.02436 (MS, SP J. Thorén, JHEP)

# Parton showers

# Parton showers

- Adds more and more **softer** emissions to a **hard** matrix element. Order emissions in some “shower time” like $k_\perp$, virtuality or angle.



- The splitting kernel diverges for small momentum fractions $x$ of initial the particles momentum, for example

$$P_{g\to gg}(x) = C_A\left[\frac{1-x}{x} + \frac{x}{x-1} + x(1-x)\right]$$

- Different emissions compete against each other, the highest scale “wins”
- Use the Sudakov veto algorithm to pick emission
- At leading order and leading color, all contributions to the emission probability are positive, and we have no reason to introduce weights.

# Parton showers

- Given a starting scale $Q$, the scale $q$ of the next emission and some additional splitting variables $z$, we have the distribution

$$\frac{\mathrm{d}S_P(q|Q)}{\mathrm{d}q\,\mathrm{d}z} = \Delta_P(\mu|Q)\delta(q-\mu) \qquad \leftarrow \text{no emission}$$
$$+\Delta_P(q|Q)P(q,z)\theta(Q-q)\theta(q-\mu) \quad \leftarrow \text{next emission at } q,$$

  where
  - $P$ is a splitting kernel
  - $\mu$ is the infrared cutoff
  - $\Delta_P(q|Q)$ is the Sudakov factor

$$\Delta_P(q|Q) = \exp\left(-\int_q^Q dp \int dz\, P(p,z)\right)$$

- Our goal is to simulate the process

$$Q \overset{S_P}{\rightsquigarrow} (\tilde{q}_1,\tilde{z}_1) \overset{S_P}{\rightsquigarrow} (\tilde{q}_2,\tilde{z}_2) \overset{S_P}{\rightsquigarrow} (\tilde{q}_3,\tilde{z}_3) \overset{S_P}{\rightsquigarrow} \cdots$$

# The Sudakov algorithm



Assume that $R(q)$ is an overestimate of the splitting kernel $P(q)$. Start from some higher scale $Q$ and find the next scale $q$ (see for example MS&SP, 1108.6180 EPJP)

1) solve $rnd = \Delta_R(q|Q)\theta(q-\mu)$ for $q$
2) If $q = \mu$, return $\mu$
3) else, return $q$ with probability $P/R$
4) if $q$ is not returned update $Q$ to $q$, and continue from start

In this simple case, the weights of all events can be kept at unity $\rightarrow$ no issue with weight spreading

LUND UNIVERSITY

# Parton showers

- Today parton showers attempt to go beyond leading order/beyond leading log, and are matched and merged to hard matrix elements. This introduces negative contributions to the probability to emit.
- Even at leading order, color suppressed contributions give rise to negative contributions.
- To deal with negative contributions weights are introduced, and the weighted Sudakov algorithm is used. (SP, Bellm et. al. 1605.08256, PRD)
- On top of that variations in $\alpha_s$, or generally variations in splitting kernel can be accommodated by using weights. (SP, Bellm et. al. 1605.08256, PRD)
- $\rightarrow$ Good reasons to have weights in parton showers.

# Weights in showers

# The weighted Sudakov algorithm

Start from some higher scale and a weight $w$, find next scale, weight (and other splitting variables): (SP, Bellm et. al. 1605.08256, PRD)

1) Make a trial splitting according to $S_R$ for some (any!) $R \neq P$
2) If the return scale is $q = \mu$, there is no emission and the cut-off scale $\mu$ is returned with weight $w$
3) Else, the trial splitting variables $q, z$ are accepted with probability $\epsilon$ and the return weight is

$$w \leftarrow w \times \frac{1}{\epsilon} \times \frac{P}{R}$$

4) Otherwise the splitting is rejected and the weight becomes

$$w \leftarrow w \times \frac{1}{1-\epsilon} \times \left(1 - \frac{P}{R}\right)$$

Then, continue from start, but with $Q = q$.

# Weights in parton showers



- Weights are introduced to treat the emission probability
- ... as well as the no-emission probability
- Since we have many (no-)emissions in parton shower, they multiply, and large weights can arise.
- $\rightarrow$ Would be good to have a way of reducing the spreading of weights
- Standard solution, reweighing (unweighting):
  Example: If a weight $w$ is sufficiently small, keep it only with probability $w$, and if kept put to 1.
- A problem with this is that it only deals with small weights, large weights are left untreated.

# Statisticians solution: Sequential importance sampling with resampling



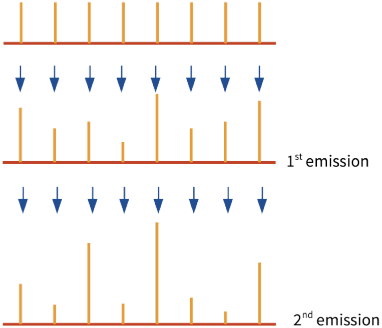
- The growing weight distribution problem within sequential Monte Carols is a standard problem for a mathematical statistician. It appears in many simulation contexts: signal processing, robotics, autumatic control theory, financing, computational biology.
- It has a standard solution, resampling: Do many simulations in parallel, for example run many showers in parallel. Repeatedly, as the simulation proceeds (for example when the 10 000 parton showers are running and accumulating weights), resample, i.e., among the (say) 10 000 simulations (here showers) randomly select, in proportion to the weight 10 000 instances of the simulation (showers), and if selected keep the weight to 1. (N. Gordon, D. Salmond, and A. F. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation", IEE Proc. F, Radar Signal Process. 140 (1993), pp. 107–113.)

# Resampling

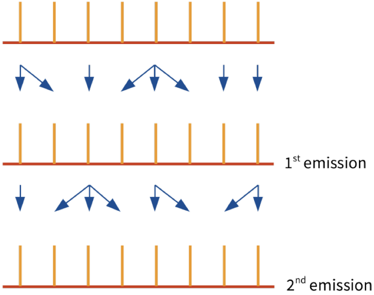


- This means that some simulations (showers) will be duplicated and others will be forgotten. The duplicated showers will then evolve differently, but keep sharing a part of their history.

Standard:




Resampling:

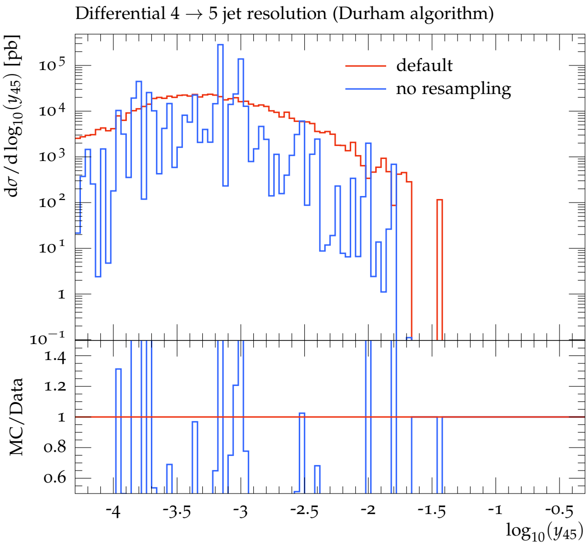


# Implementing resampling

# Let's try it out

Take a simple parton shower, and look at some distribution. Then destroy the statistical convergence beyond recognition be repeatedly using the weighted Sudakov algorithm.

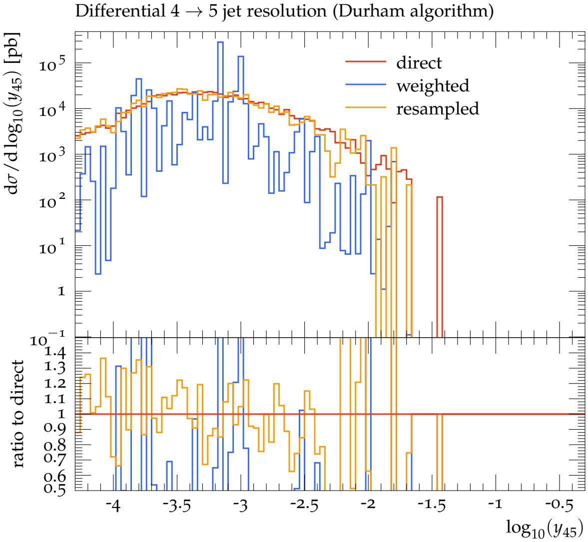
(Here, the Durham 4 $\to$ 5 jet resolution scale, shower, with $\epsilon = 0.5$. Shower from S. Höche, Tutorial on Parton Showers CTEQ, 2015.).



# Let's try it out



Take a simple parton shower, and look at some distribution. Then destroy the statistical convergence beyond recognition be repeatedly using the weighted Sudakov algorithm, then add resampling. It works like magic!




LUND UNIVERSITY

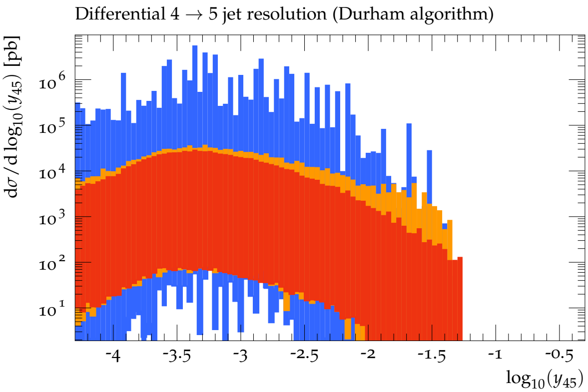
# Conclusion and outlook

# Conclusion and outlook

- Weights needed in parton showers
- Resampling works well for sequential importance sampling
- For our test shower, resampling worked like magic!
- Further improvement can likely be obtained by finding the optimal number of resamplings, keeping differently many showers in different steps, resampling only when the “effective sample size” reaches a certain value (statistics is poor enough)
- I strongly believe in this method

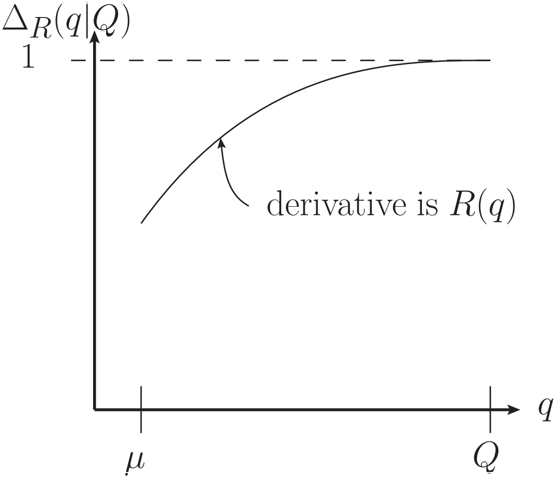
# Backup: Envelope of 300 runs





# Backup: The Sudakov factor





# Backup: Adaptively triggered resampling, effective sample size



Introduce the coefficient of variation

$$C_V = N \sum_{i=1}^{N} \left( \frac{w_i}{\sum_{l=1}^{N} w_l} - \frac{1}{N} \right)^2$$

and the effective sample size

$$ESS = \frac{N}{1 + C_V^2}$$

resample only when the effective sample size falls below a some given value

# Backup: Variance of sample



The variance, for uncorrelated events with mean $\overline{y}$,

$$s^2 = \frac{1}{N(N-1)} \sum_{l=1}^{N} (y_i - \overline{y})^2 = \overline{y}^2 \left(1 - \frac{N}{N-1}\right) + \left(\frac{N}{N-1}\right) \frac{1}{N^2} \sum_{i=1}^{N} y_i^2$$

is modified as

$$\overline{v}^2 \left(1 - \left(\frac{N}{N-1}\right)^{n+1}\right) + \left(\frac{N}{N-1}\right)^{n+1} \frac{1}{N^2} \sum_{i=1}^{N} v_i^2$$

where

$$v_i = \sum_{l\text{:originally from ancestor i}} w_l h(\xi_l)$$

for some observable $h(\xi_l)$ of the event $\xi_l$. (A. Lee and N. Whiteley, Variance estimation in particle filters, Biometrica 105 (2018), no. 3)

# Backup: Durham jet-resolution parameter



The plots show $y_{45}$, where

$$y_{ij} = \frac{2\min(E_i^2, E_j^2)}{Q^2}(1 - \cos\theta_{ij})$$

with $E_{i/j}$ being the energy of the particle (jet) $i/j$, $\theta_{ij}$ is the angle between them and $Q^2$ is the squared center-of-mass energy.
This is the a measure of the relative energy scale at which an event starts to be counted as a 5-jet event rather than a 4-jet event.
(Catani et. al. PLB)