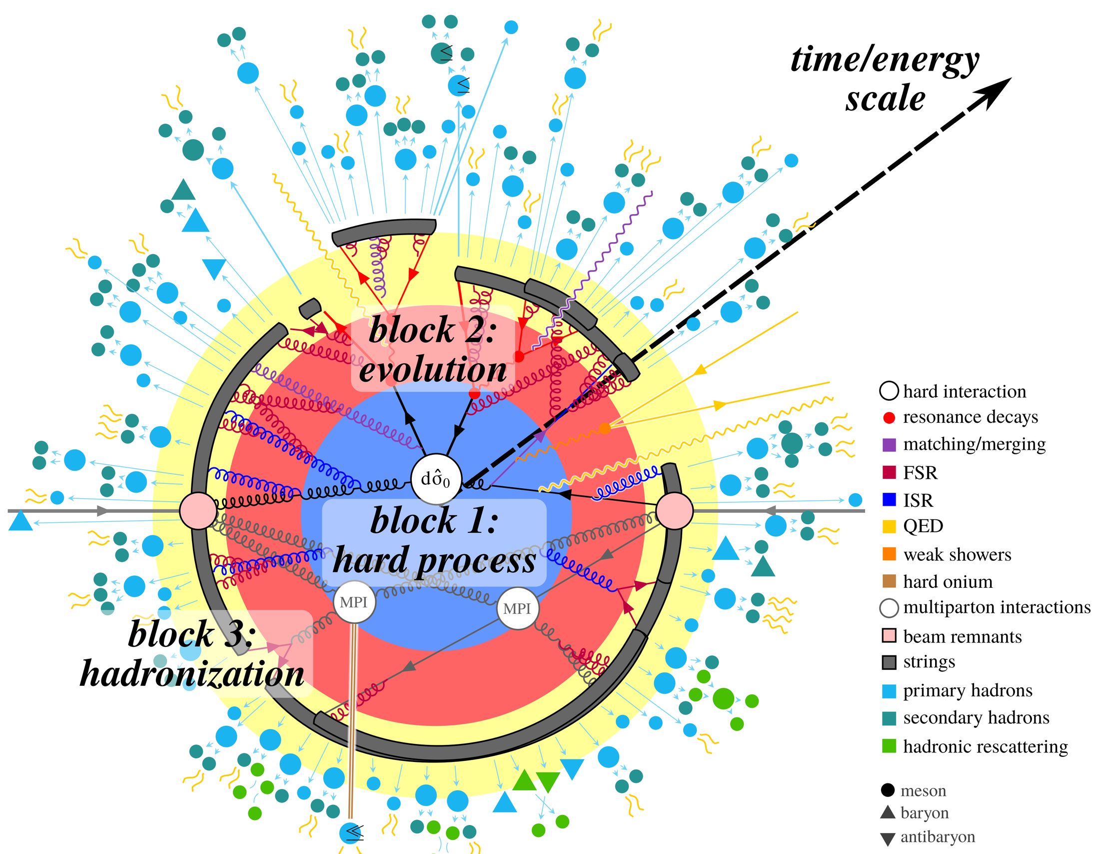
Monte Carlo reweighting for hadronization uncertainties

Christian Bierlich, Philip Ilten, Tony Menzo, Stephen Mrenna, Manuel Szewc, *Michael K. Wilkinson*, Ahmed Youssef, Jure Zupan, *i.e.*, the MLHAD team

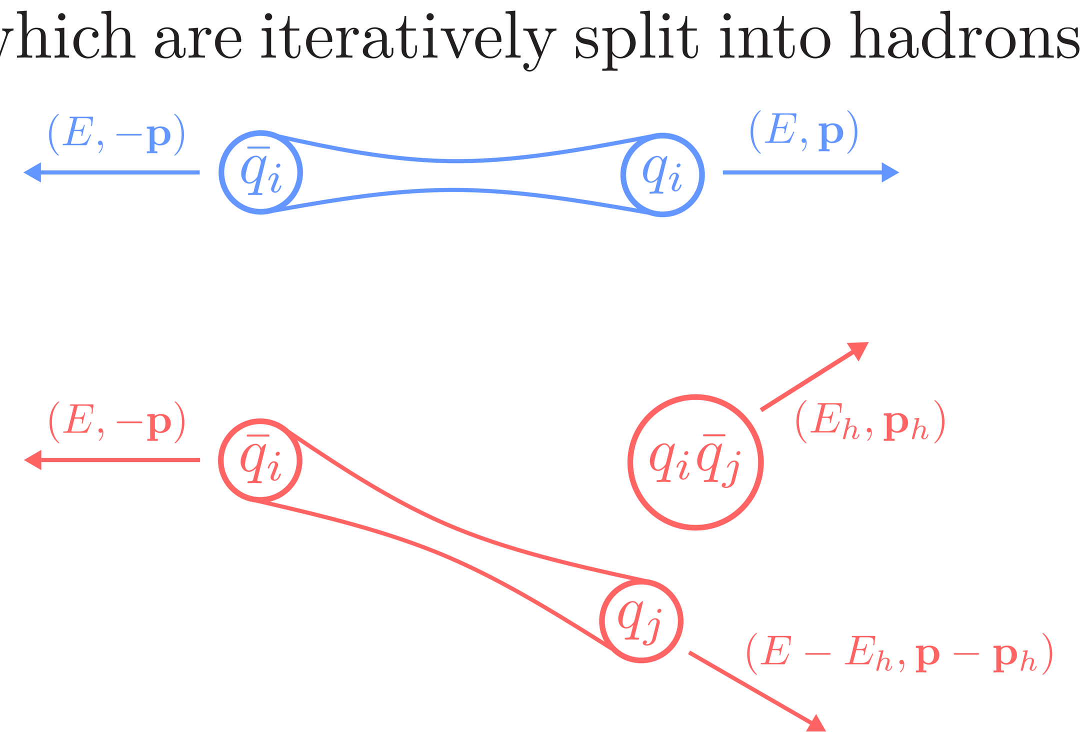
Hadronization

Monte Carlo event generators simulate full particle collisions in three blocks:



1. *hard process* - initial high energy interaction (perturbative)
2. *evolution* - perturbate to nonperturbative scale $\sim 1\,\mathrm{GeV}$
3. *hadronization* - combine quarks and gluons into hadrons via phenomenological models

In the string model of hadronization (implemented in PYTHIA), partons are bound by the strong force into strings, which are iteratively split into hadrons:



Modified accept-reject

A probability distribution $P(z)$ can be sampled when a reliable overestimate $\widehat{P}$ is known. In the standard accept-reject algorithm, for a trial value of z,

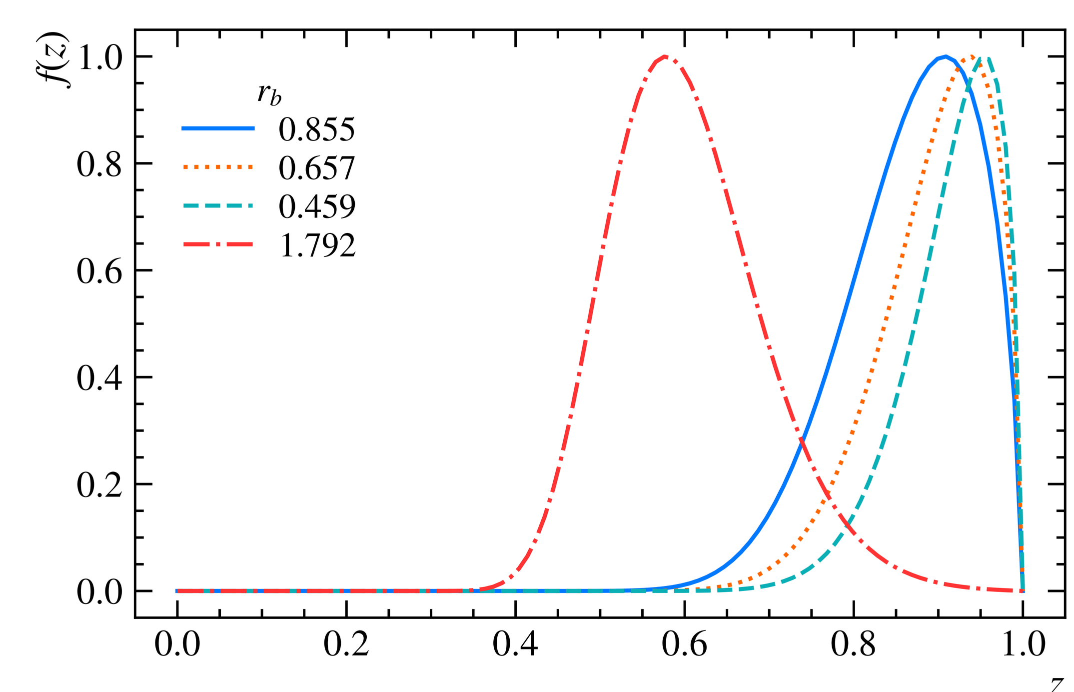
$$P_{\text{accept}}(z) \equiv \frac{P(z)}{\widehat{P}} \leq 1\,.$$

If an event generated with $P(z)$ has weight w, the weight w' for an event generated with $P'(z)$ can be calculated by tracking *all* the trial z values:

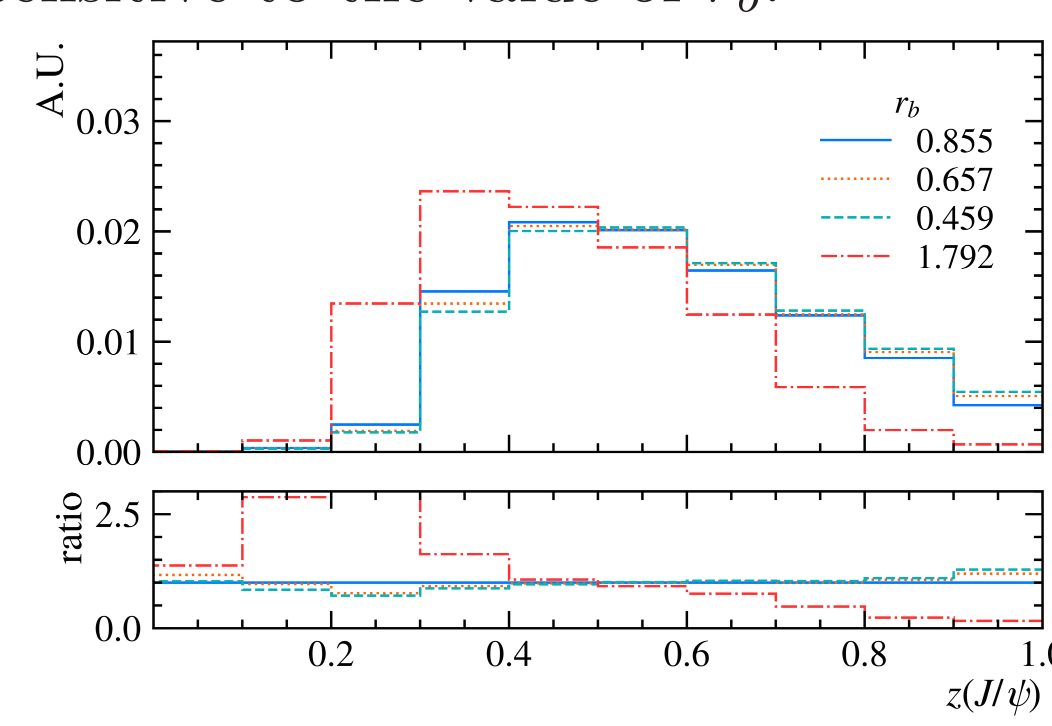
$$\frac{w'}{w} = \prod_{i \in \text{accepted}} \frac{P'_{i,\text{accept}}(z)}{P_{i,\text{accept}}(z)} \times \prod_{j \in \text{rejected}} \frac{P'_{j,\text{reject}}(z)}{P_{j,\text{reject}}(z)}.$$

Parameter uncertainties

In the string model, the probability $f(z)$ to emit a hadron with a fraction of the string's longitudinal lightcone momentum z depends on parameters fixed from fits to experimental data, *e.g.*, r_b:

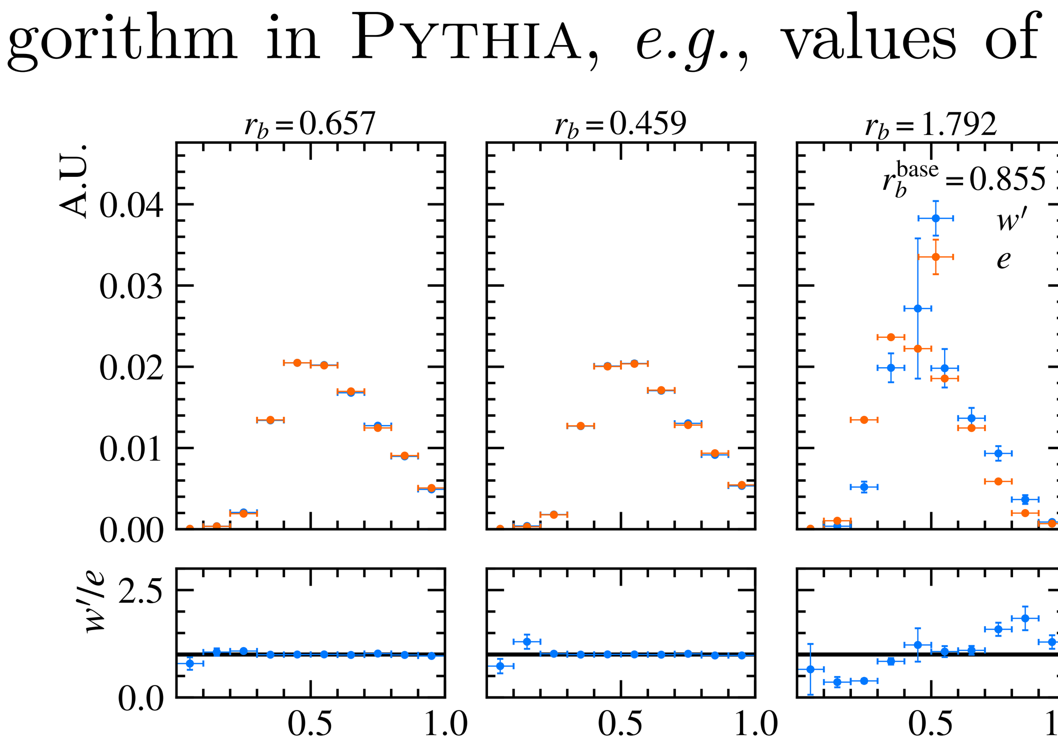


The uncertainties are typically evaluated by simulating collisions with different parameter values and comparing to relevant observables in experimental data; *e.g.*, the ratio of the transverse momentum of a J/ψ meson to that of the jet in which it is found, $z(J/\psi)$, is sensitive to the value of r_b:



Replicating distributions

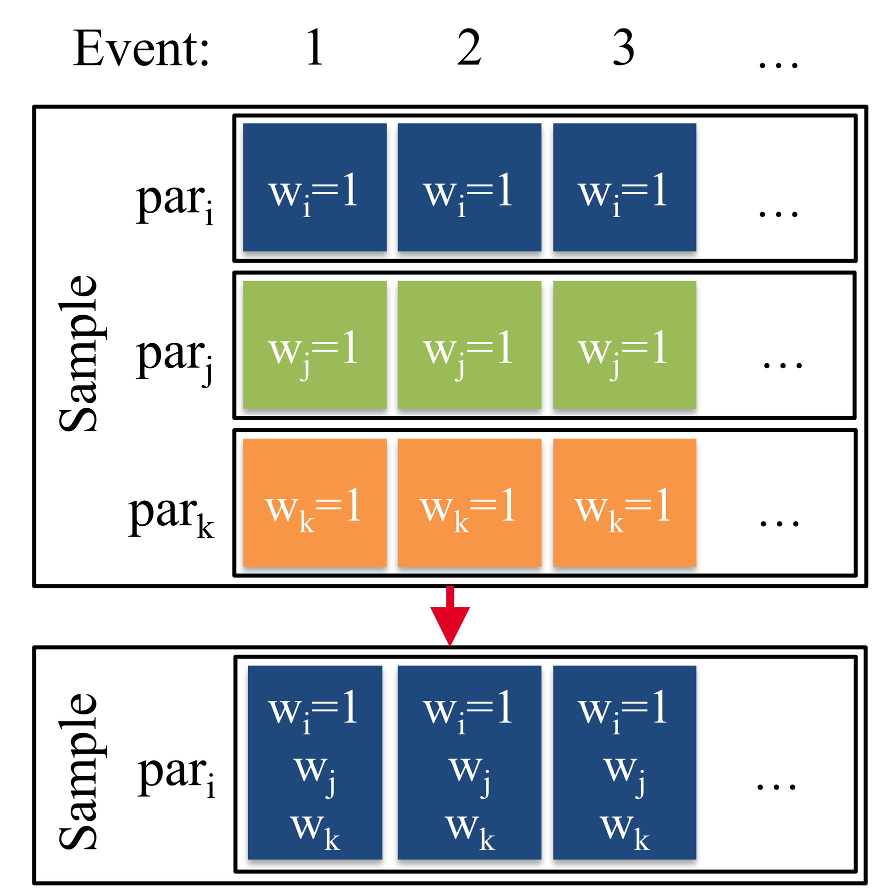
Simulation generated with one set of hadronization parameter values can be reweighted to match another by implementing the modified accept-reject algorithm in PYTHIA, *e.g.*, values of r_b:



The e distributions have been generated with r_b set to the given values; the w' distributions have been reweighted from $r_b = r_b^{\text{base}}$ using the modified accept-reject algorithm, implemented in PYTHIA.

Reweighting

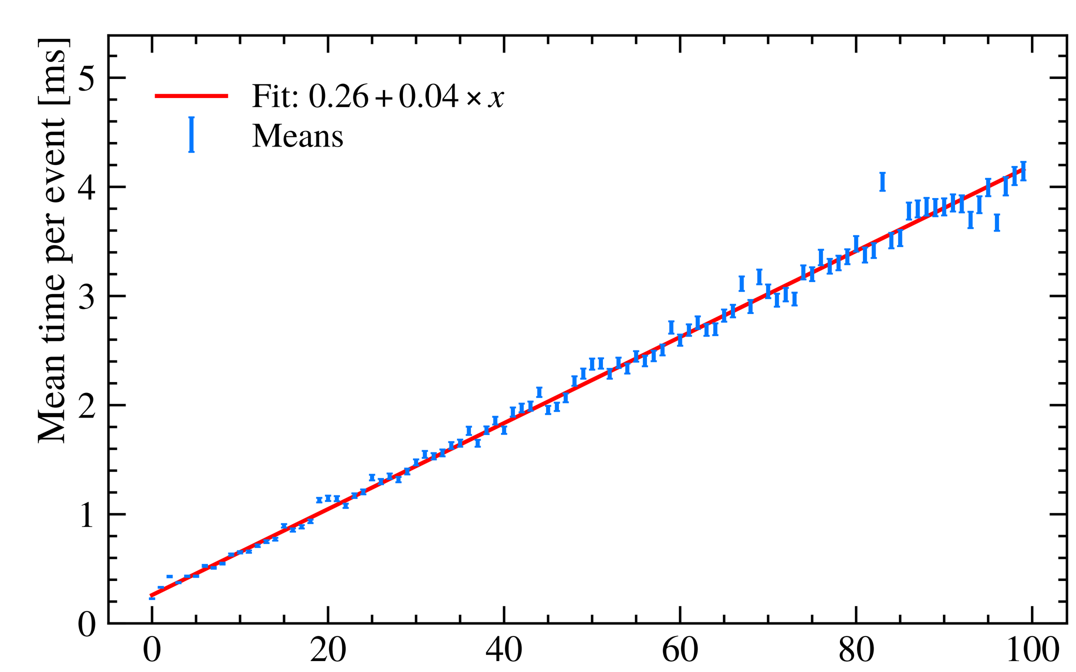
Generating simulation for each input parameter value is computationally expensive, especially when including detector effects. Reweighting calculates alternative weights for each event so that a single simulation can be used to estimate the effect of different parameter values:



This is conceptually straightforward but can be technically challenging.

Improving timing

The average time required to generate a single event using the modified accept-reject algorithm scales linearly as a function of the number of alternative values considered for parameter a:



In this example, generating an additional value of a using the modified accept-reject algorithm is $\approx 6.5\times$ faster than not using it; were detector simulation included, the speedup would be far greater, $\mathcal{O}(1000)$.

Publication

These findings have been submitted to SciPost, and the pre-print is available as `arXiv:2308.13459`.