Workfest Announcement

Optimization of Monte Carlo production for physics analysis

Monte Carlo simulation for physics analysis can be very **expensive** in terms of **computing resources** and **wall time**.

When planning a large Monte Carlo production campaign, it's important to know **how long** it may take and, if necessary, leverage **smart techniques** to reach the required statistical precision.

Several options are available, from **event weighting** to populate low-efficiency and low-cross-section regions, to **fast Monte Carlo** based on an AI model of the detector acceptance.

During this Workfest session, we will explore these ideas with the help of collaborators who have tackled these kinds of problem before. We invite everyone who feels Monte Carlo simulation for their analysis may be too expensive to bring their use case and explore with us how these alternative options could be applied.

Contributions

- Efficient simulated data in particle physics: event generation, reweighting and fits

  Derek Glazier

- Parnassus for CLAS12: an ML surrogate for detector simulation and reconstruction

  Tyler Hellstern

- RG-D MC: simulation requirements

  Mariana Tenorio Pita

- Open discussion of analysis use cases

  All Workfest participants

Topics discussed

- How long is it going to take?
  - gemc + coatjava ≈ 1–2 s CPU per event: 1B events ≈ 0.3M core-hours
  - OSG: ~200M events/day, shared among all users
- How many events are really needed?
  - MC statistical uncertainty negligible vs. the total uncertainty, not necessarily Nrec ⩾ Ndata
- Filtering generated events outside the CLAS12 acceptance before gemc
- Event weighting to populate low cross-section regions (e.g. Q^2)
- Fast MC: AI models of the detector acceptance and the coatjava fast MC

Main takeaways

- There is room to optimize MC production
- All the options discussed generated interest, from AI-based fast MC to event weighting and pre-filtering
- A practical approach to implementing pre-filtering was discussed
- The coatjava fast MC should be completed