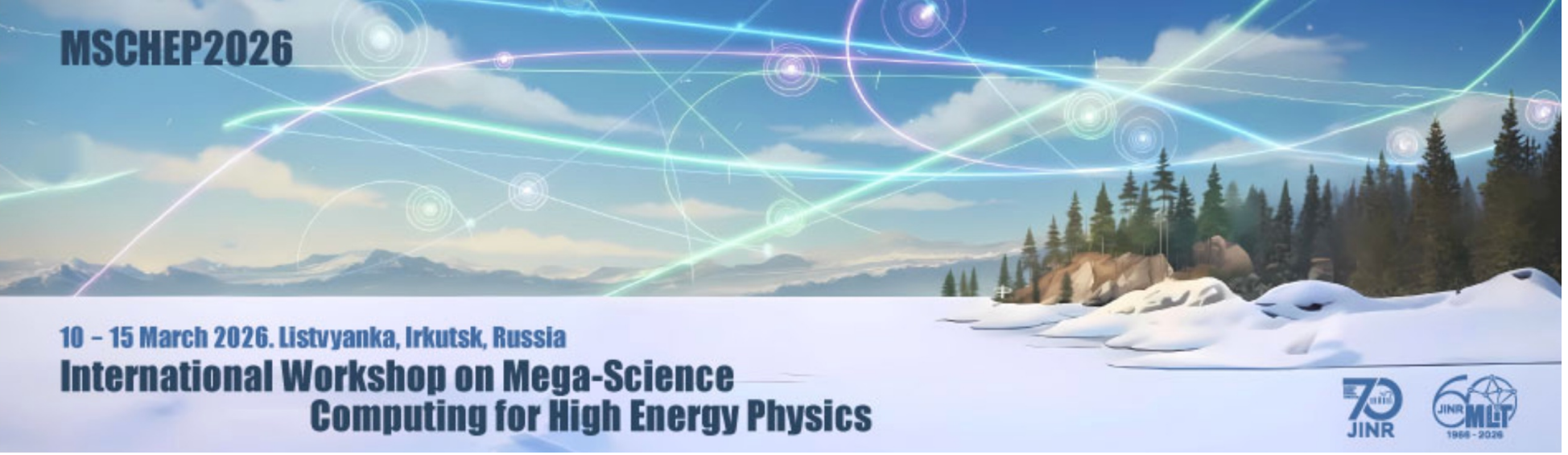


超级陶粲装置
Super Tau-Charm Facility

STCF fast simulation framework

Xiaoshuai Qin
Shandong University

On behalf of the fast simulation working group

Mar.10-16, 2026

Outline

- Introduction to STCF project
- Fast simulation framework
- Performance compared with full simulation
- Further prospects

Super Tau-Charm Facility

- Huge data statistic
  - The STCF will produce a data sample about a factor of

    100 larger than that of the present τ-charm factory(BESIII)
  - Raw data processing chain (~110 PB/year)
  - MC data production (at the similar level of raw data)
- Variety of physics topics
  - XYZ physics, $e^+e^- \to \pi^+\pi^- J/\psi$
  - Hyperon CP symmetry, $J/\psi \to \Lambda\bar{\Lambda}$
  - Collins effect
  - Charmed hardon, K^0 system, τ physics
- Background
  - Expect to have O(10^{11}) level background events for 1ab^{-1} data.
- Huge CPU/storage consumption

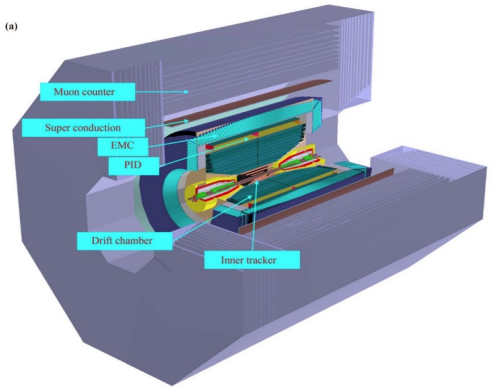
Basic design:

Peak luminosity:

$0.5\times10^{35}\text{cm}^{-2}\text{s}^{-1}$@$E_{cm}$=4GeV

Luminosity per year: 1.4ab^{-1}

Energy region: E_{cm} =2–7GeV



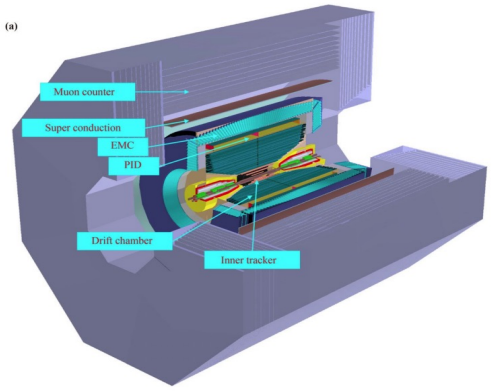


Fast simulation for STCF

- Goals of physics simulation
  - Feasibility studies for a dedicate physical topic (physical sensitivity, background etc.)
  - Optimize detectors (efficiency, resolution etc.)
  - Comparison to past/existing/future experiments

- Simulation Tools
  - Full Simulation (For signal and standard for fast simulation.)
  - **Fast Simulation (For the background study, sensitivity estimation, etc.)**

Basic design:
Peak luminosity:
$0.5\times10^{35}cm^{-2}s^{-1}$@$E_{cm}$=4GeV
Luminosity per year: $1.4ab^{-1}$
Energy region: E_{cm} =2–7GeV



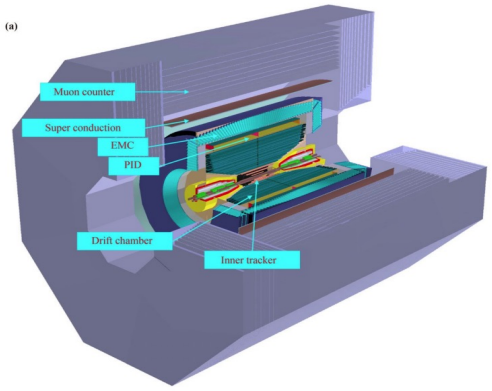


Fast simulation for STCF

- Basic ideas for Fast simulation:
  - Fast and small storage capacity (inclusive MC samples)
  - Flexible for detector responses/parameters
  - Convenient and friendly for users

- Dedicated for pre-study of STCF project
  - Fast simulation framework developed long before the project
  - Event data management.

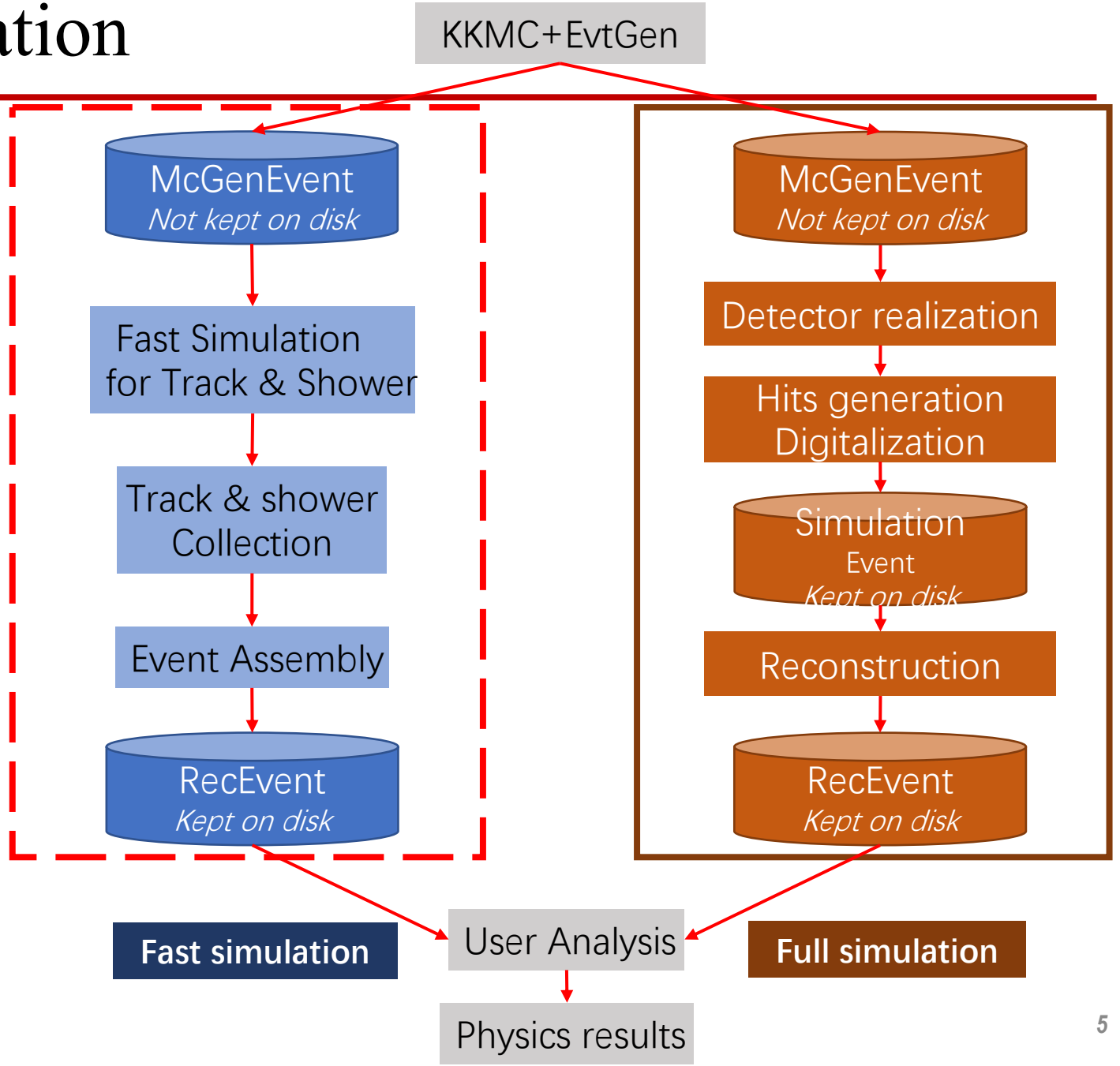
Basic design:
Peak luminosity:
$0.5\times10^{35}cm^{-2}s^{-1}$@$E_{cm}$=4GeV
Luminosity per year: $1.4ab^{-1}$
Energy region: E_{cm} =2–7GeV





Scheme for Fast Simulation

- Generator: keep McGenEvent in storage
- Fast simulation for charge and neutral tracks (resolution, efficiency etc.)
- Random seed management for repeating analysis.
- Same analysis procedure as full reconstruction for users
- Expect to have O(10^{11}) level background events for 1ab^{-1} data.
- Apply this scheme to reduce CPU and storage consumption

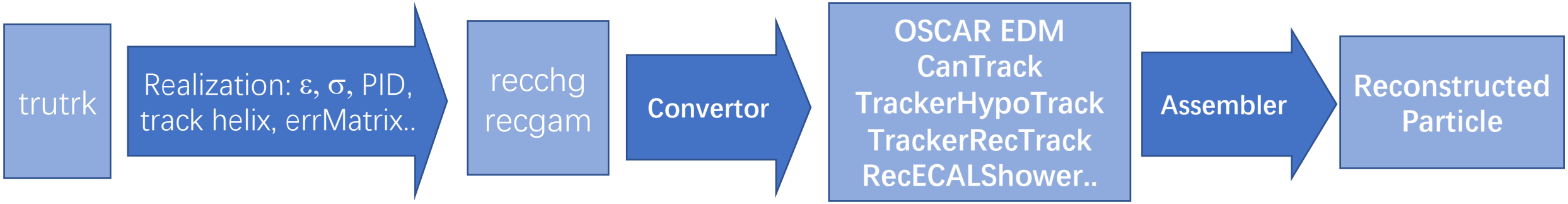


What is missed in fast simulation

- No detailed detector geometry，no particle transportation based on Geant4
  - No information of hits
  - No information of decaying in flight
  - No gamma conversion
  - No precise extrapolation or ExtTrack
- No measurements from subdetectors(but deposited ECAL shower energy is realized)
- PID decision is realized，but no detailed measurements from sub-detectors like dE/dx, RICH likelihood, MUD likelihood

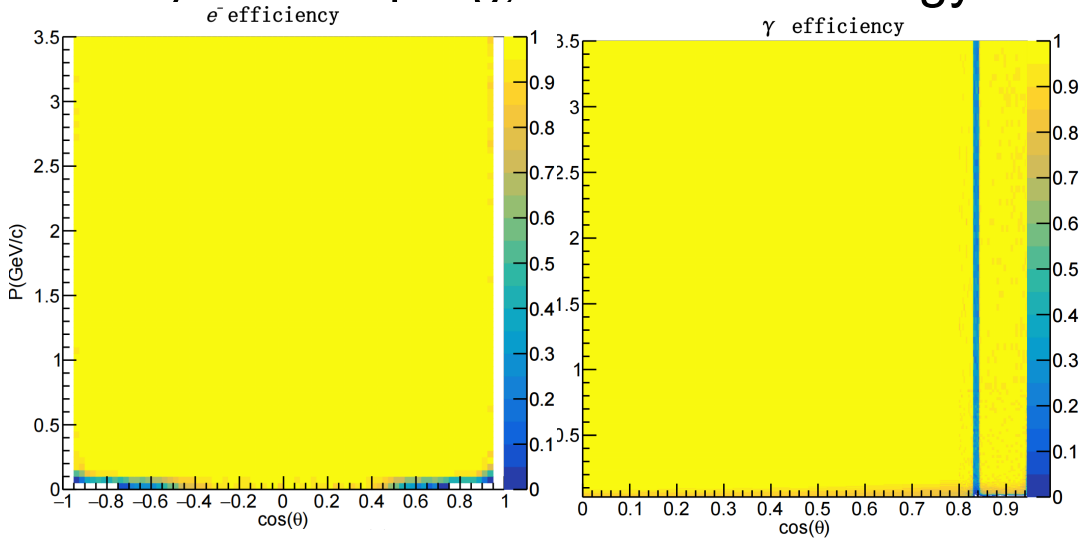
Data flow

- Truth information:
  - Particle type, momentum, vertex, decay chain
  - Data type: struct
- Information after reconstruction
  - Realization of tracking efficiency, sub-detector efficiency, momentum/energy and spatial resolution, decay vertex or generating point
  - Helix parameter and error matrix for tracks
  - Binned parameterization with respect of $p, \cos(\theta)$
- Track and shower assembling
- Event assembling

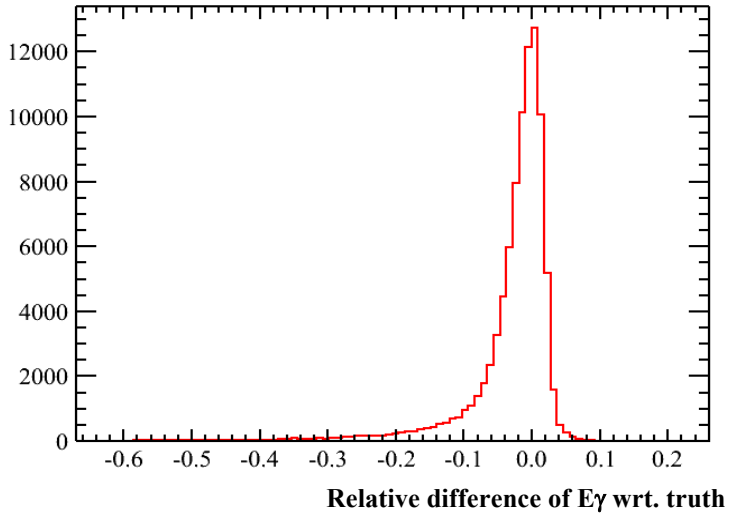


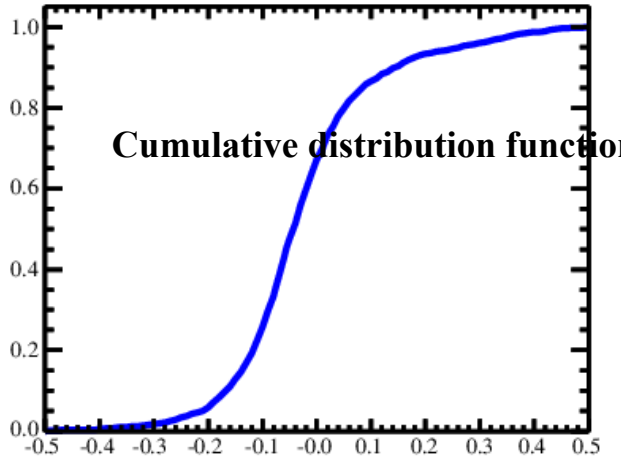
Realization for track and shower — efficiency and resolution

- Detector responses extracted from fully simulated single particle MC
- Efficiency via sampling, momentum/energy and spatial resolution via smearing



Efficiency sampling





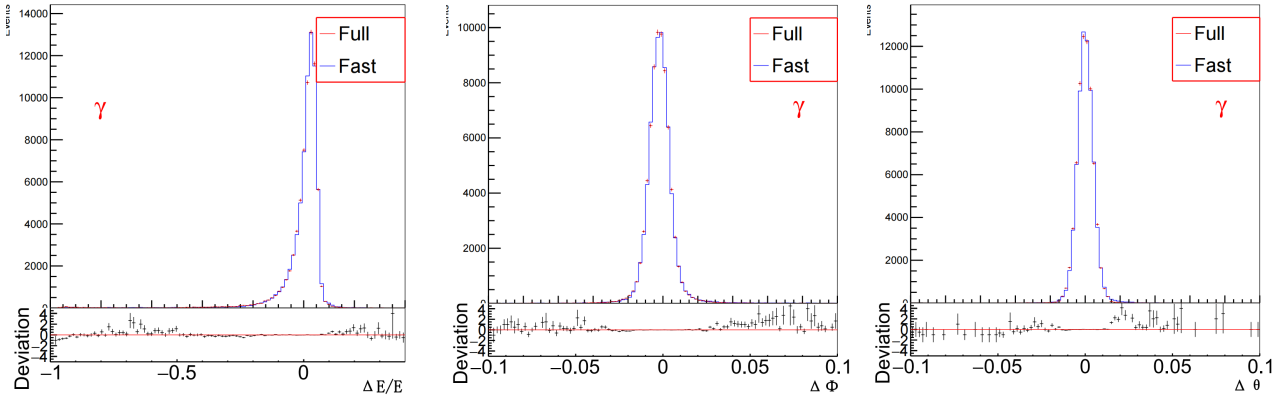
Resolution via smearing(CDF)

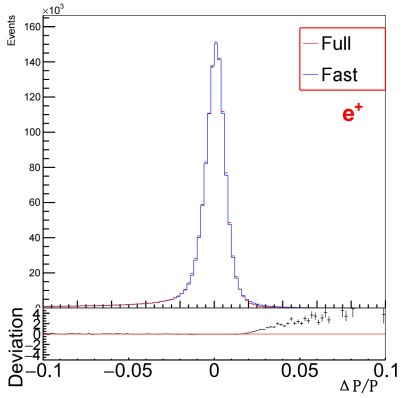
Realization of track helix parameters

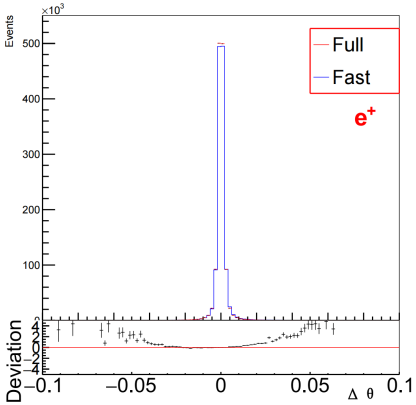
- Covariant matrix elements
  - $d_\rho, \phi_0, \kappa, d_z, tg\lambda$
  - parameterization binning with respect of $p, \cos(\theta)$
  - correlations need to be considered
- Diagonal elements: cumulative distribution function
- Non-diagonal elements:
  - correlation partially using modified form: $cov'_{03} = \frac{cov(d_0,d_z)}{\sqrt{cov(d_0,d_0)cov(d_z,d_z)}}$
  - d_0: POCA distance to IP on x-y plane.
  - d_z: POCA distance to IP on z axis.

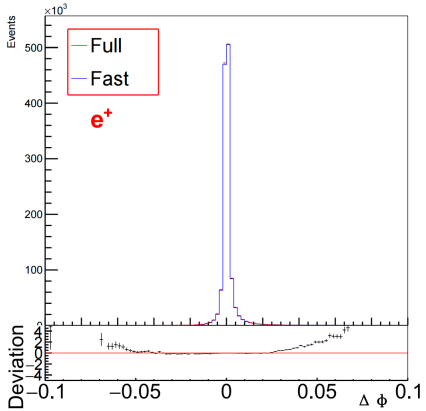
Resolution validation: single photon/charged track

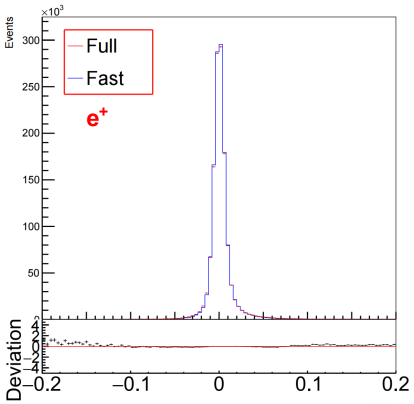
- Energy, momentum, vertex reconstruction resolution

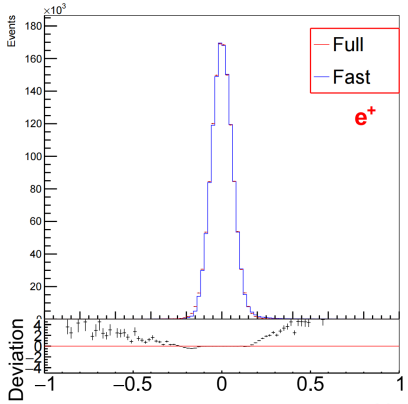






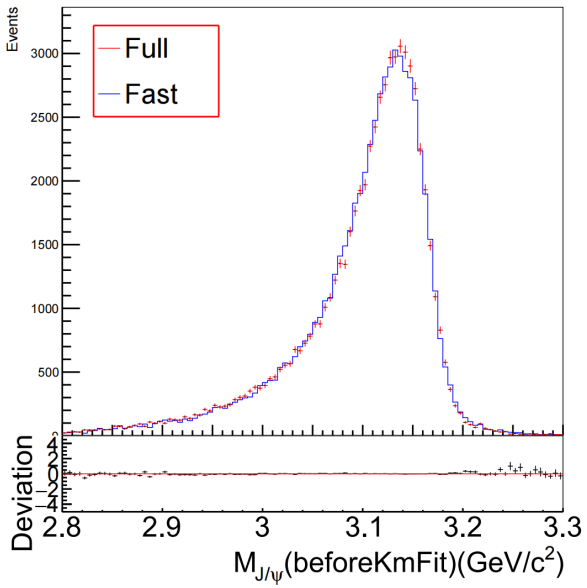


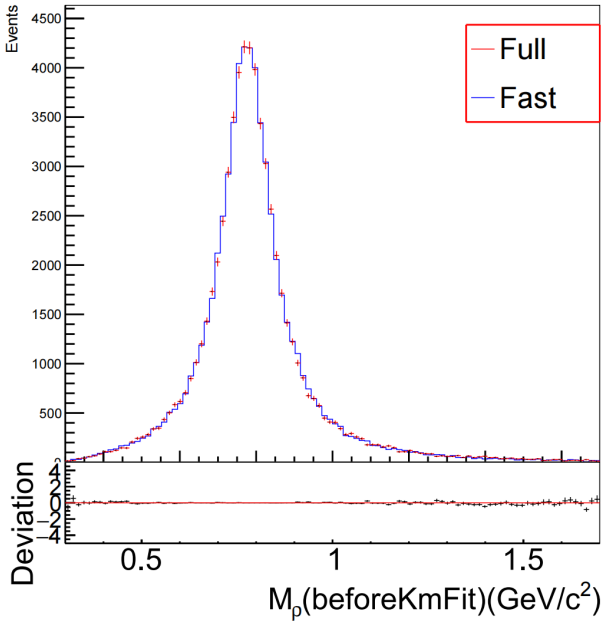


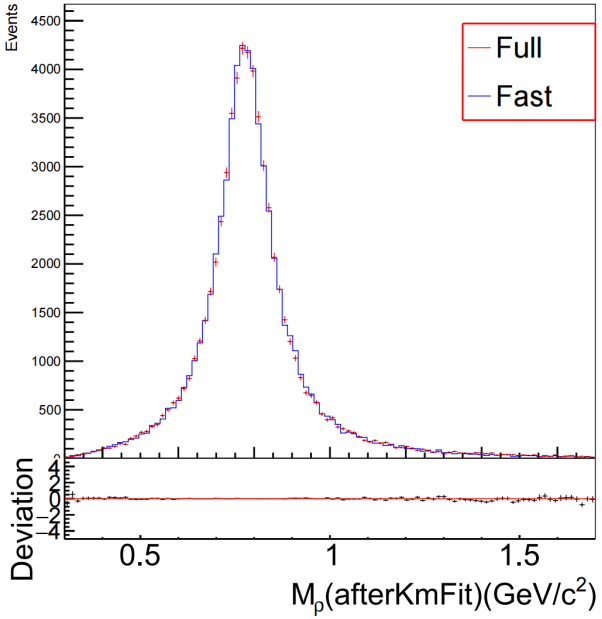


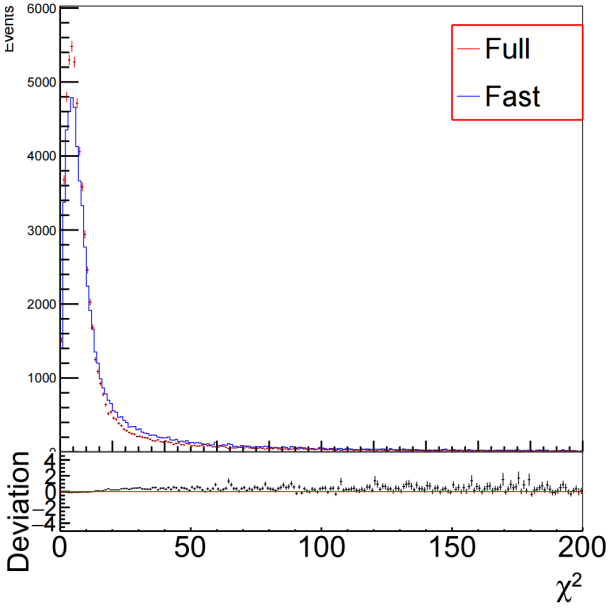
Validation from different processes

- $J/\psi \to \rho\pi$
  - track, photon fast simulation
  - kinematic fitting
  - generally works fine



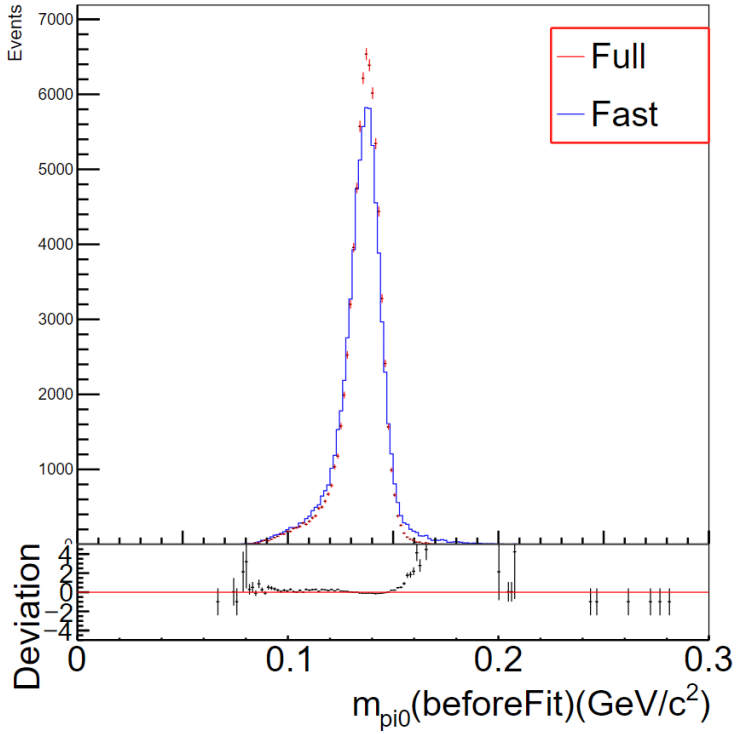


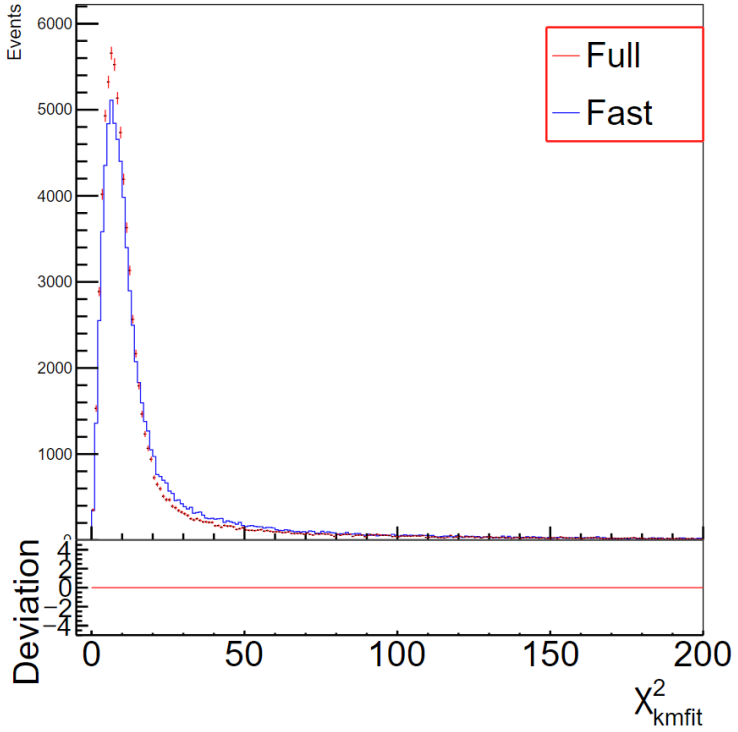


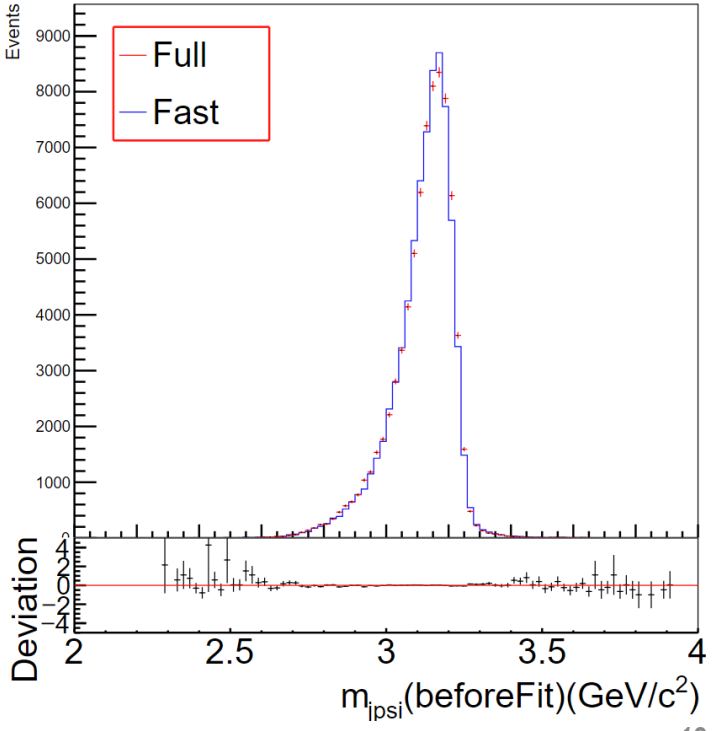


Validation from different processes

- $J/\psi \to \gamma\pi^0$
  - photon fast simulation
  - π^0 fast simulation
  - kinematic fit

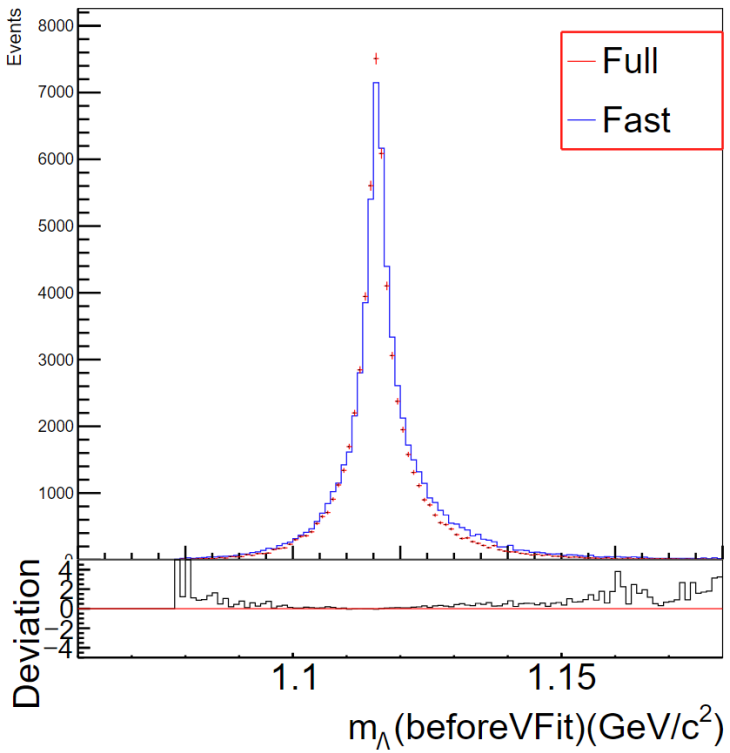


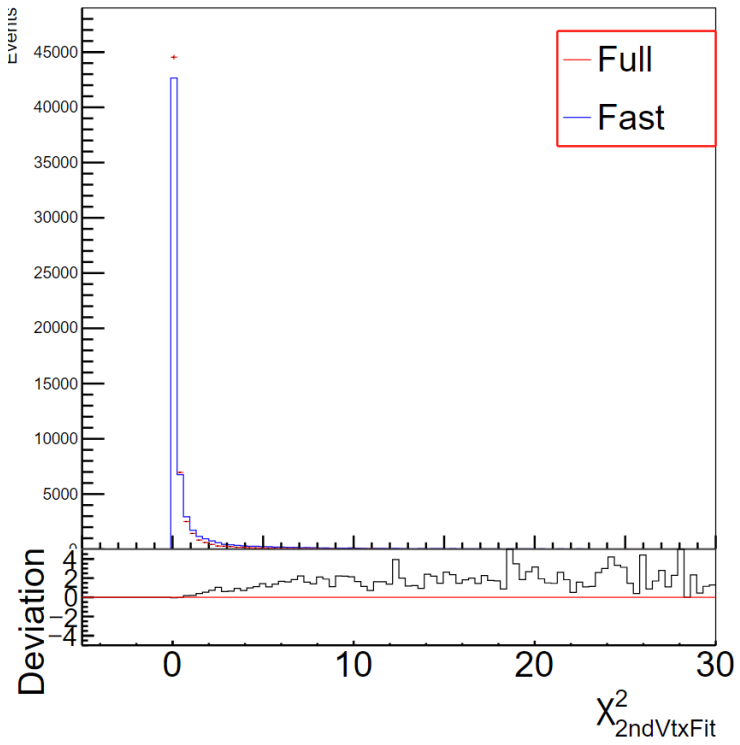


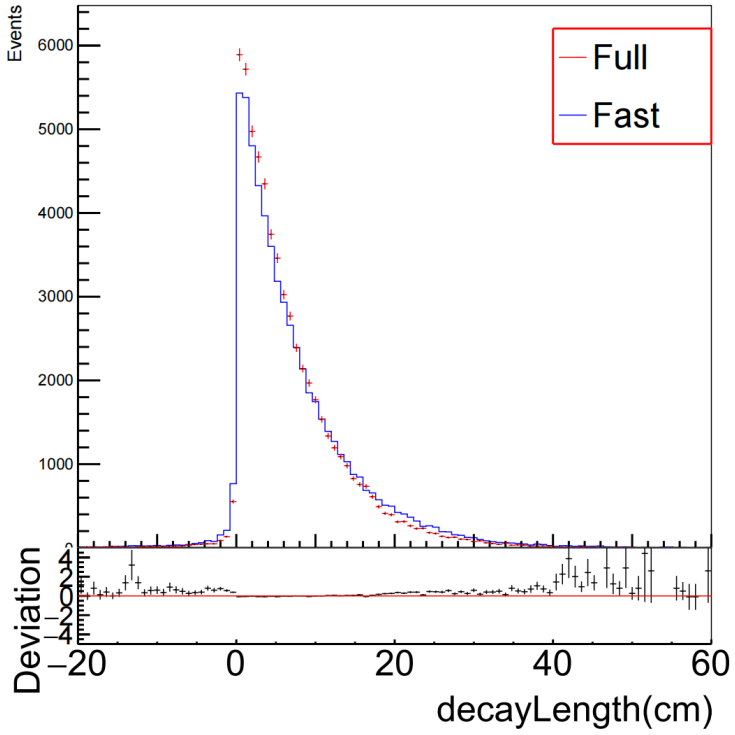


Validation from different processes

- $J/\psi \to \Lambda\bar{\Lambda}$
  - displaced tracks
  - Secondary vertex and vertex fitting



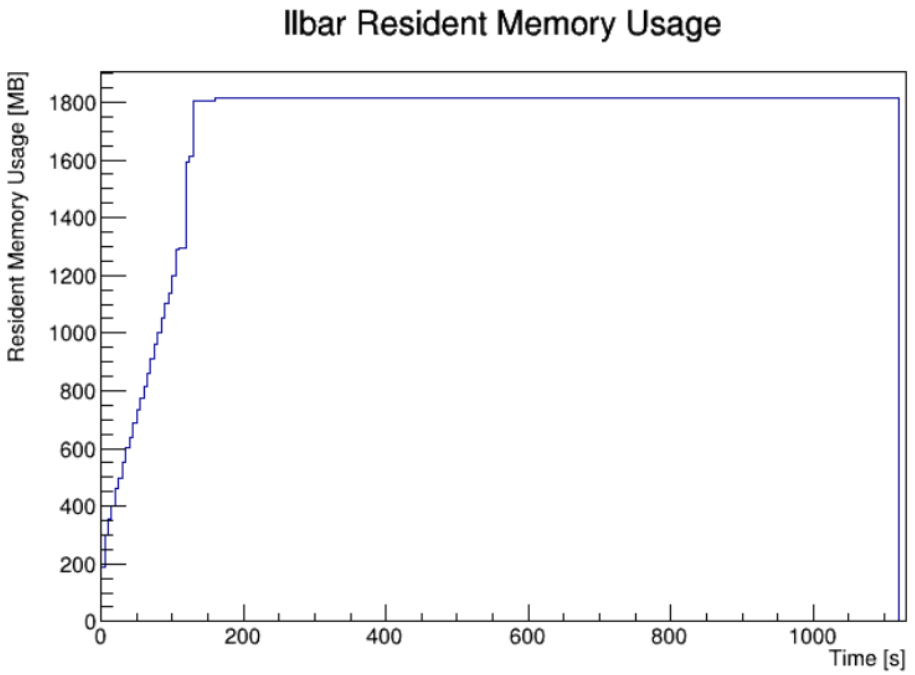




Performance

- Resource consumption reduction:
  - time: about O(10^2) faster
  - storage: more than 10 times smaller
  - memory usage: less than 2G

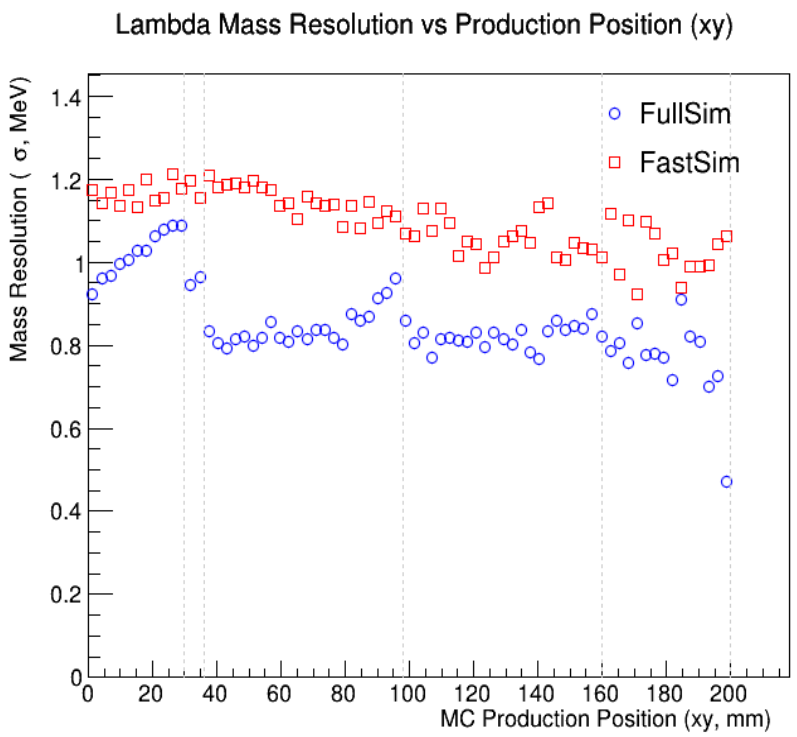
$\pi^+\pi^-\mu^+\mu^-$ (4 tracks)	Full Sim+Rec	Fast Sim
CPU time	~ 300ms(sim) + 800ms(rec) / event	~2ms / event
Event size	~77kb(sim) + 26kb(rec) / event	~ 2kb / event

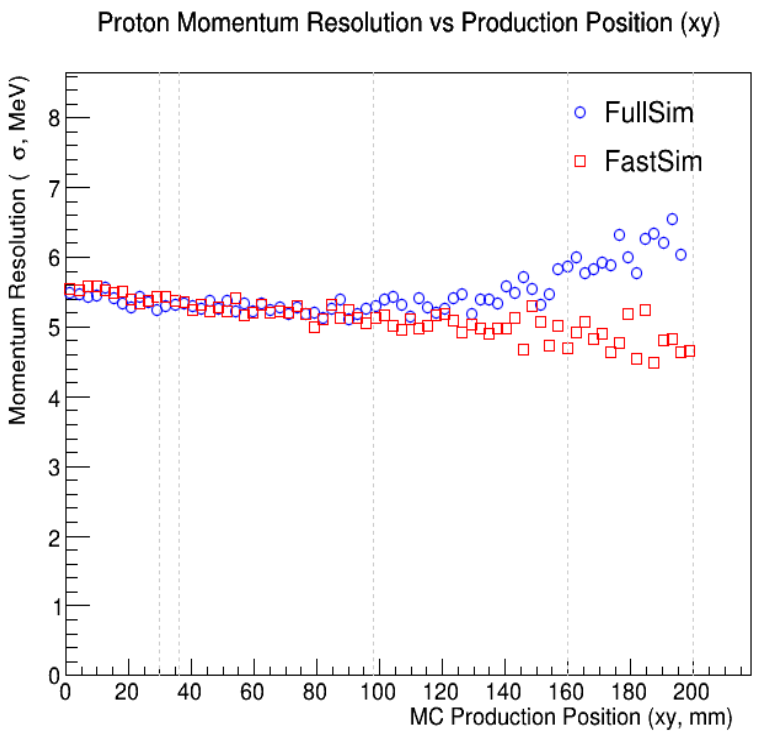


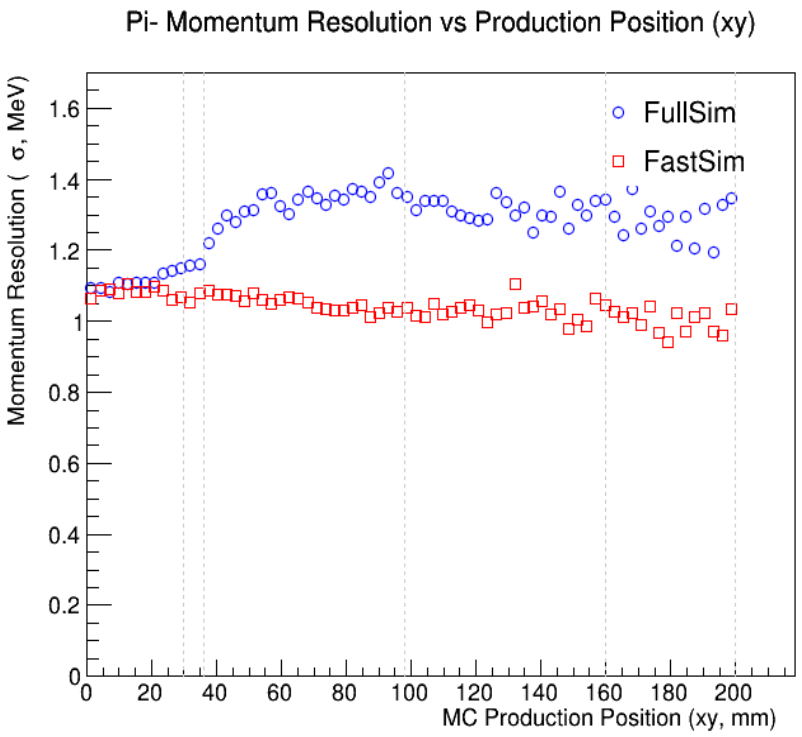
Fast simulation performance related to secondary particles

- Nominal parameterization are based on tracks from interaction point.
- Production position of displaced tracks are far away from IP
- Momentum and mass resolutions are related to production points of tracks

$$J/\psi \to \Lambda\bar{\Lambda}$$

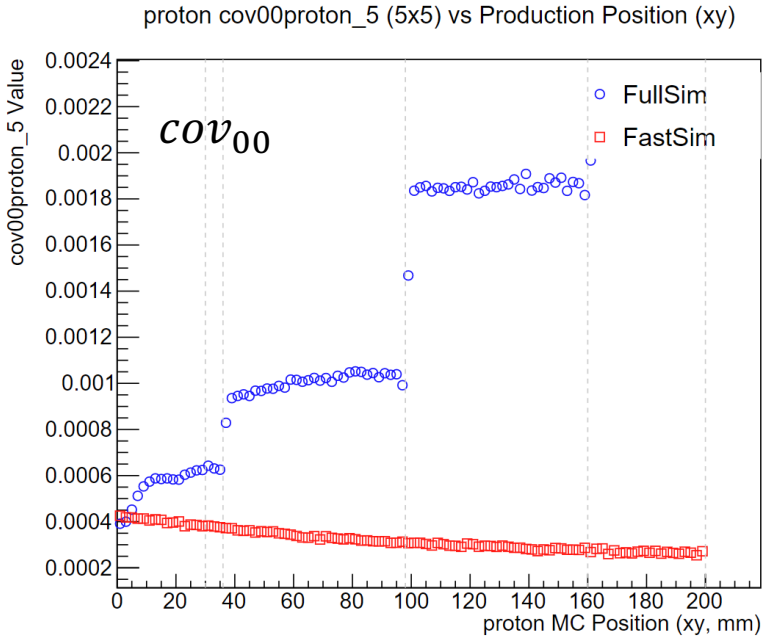


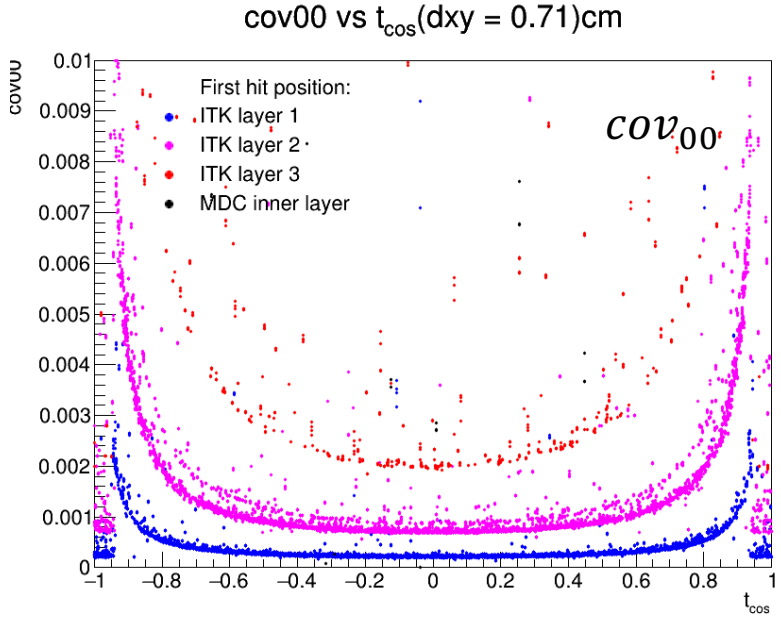


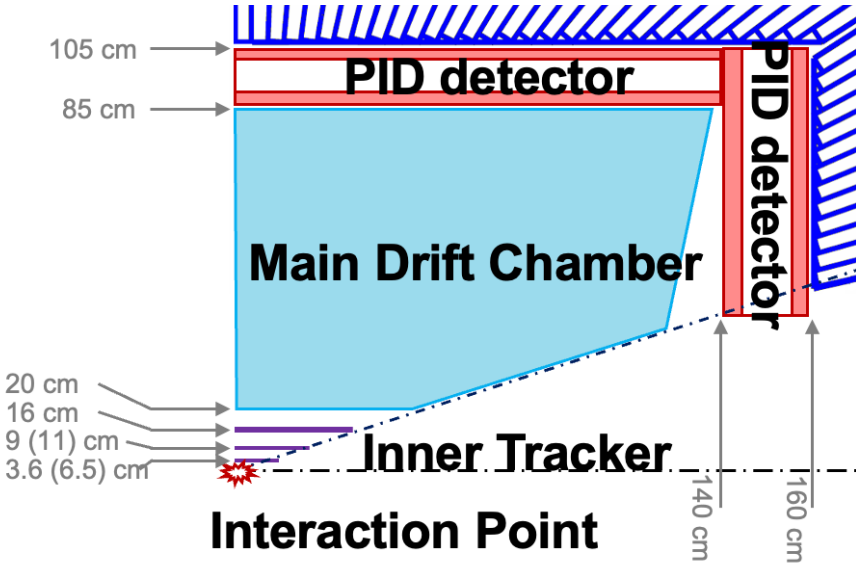


Fast simulation performance related to secondary vertex

- Covariance matrix element of track helix parameters.
- 1st hit used in the reconstructed track locates in different layer of ITK/MDC, the covariant matrix elements varies much.

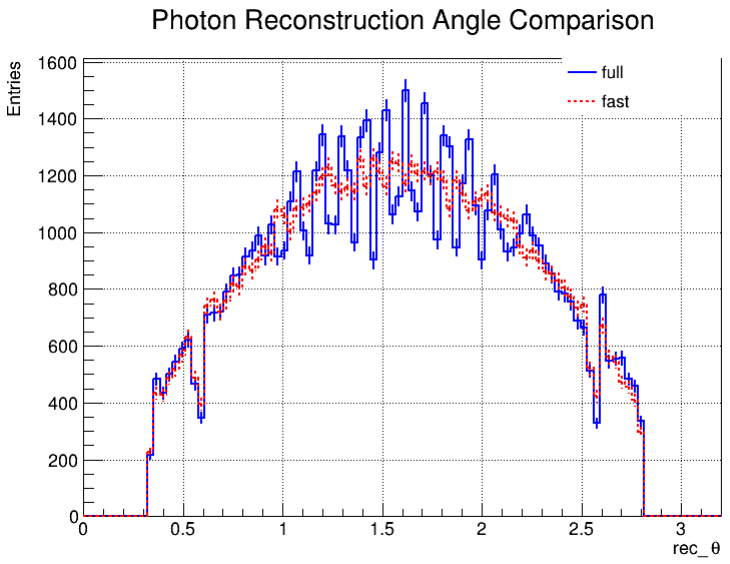




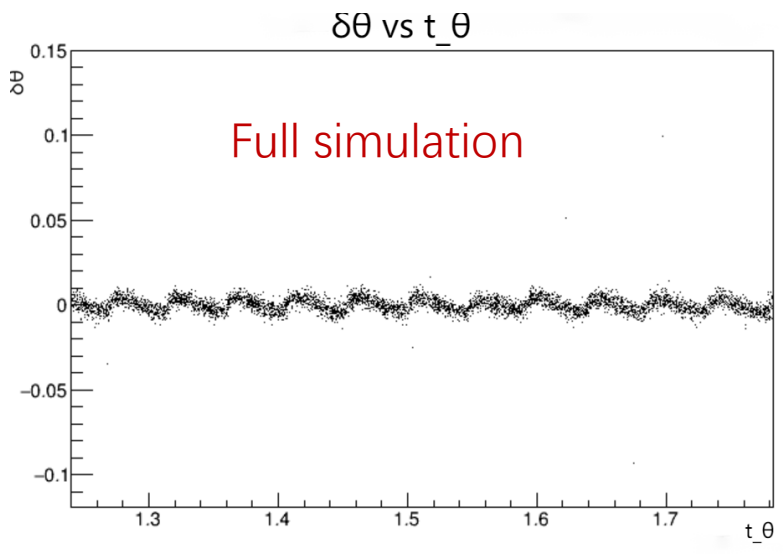


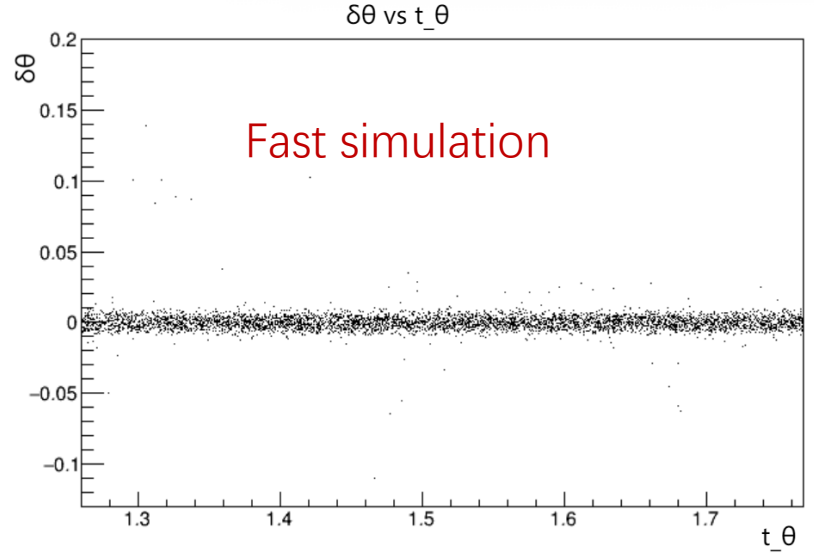
Angular resolution

- Angular resolution varies with respect to different entrance angle on the crystal surface.
- Due to binned parameterization in fast simulation, this fine structure is lost.
- Since the granularity is small, this might not affect data analysis.



Angular residual vs polar angle(truth)

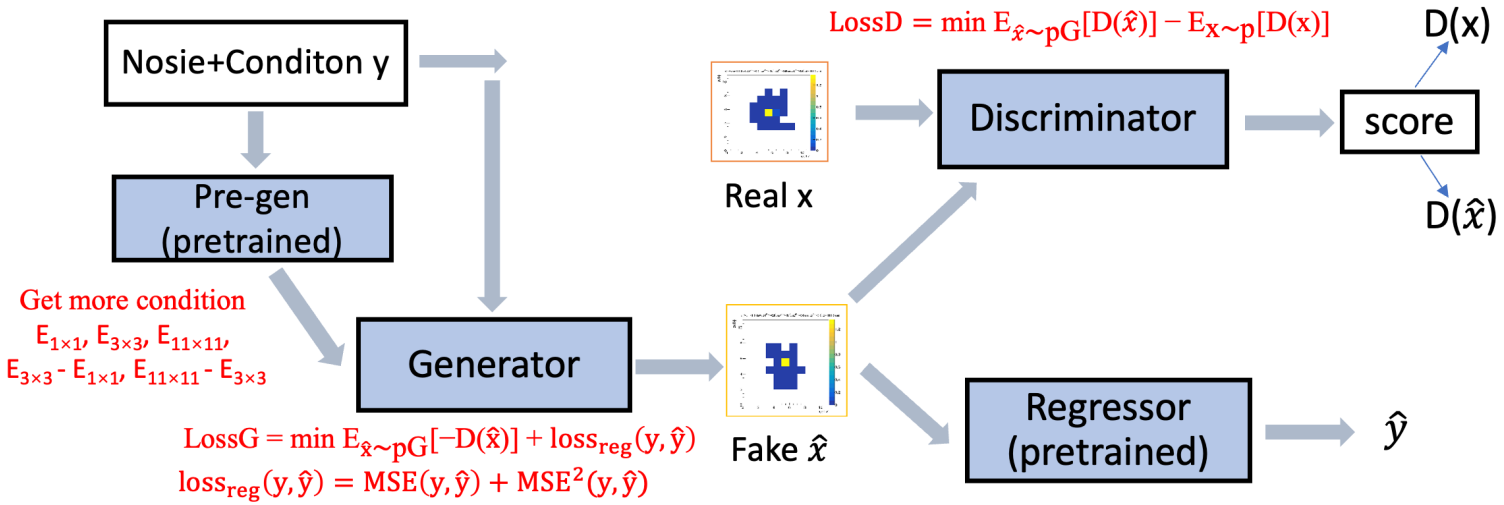




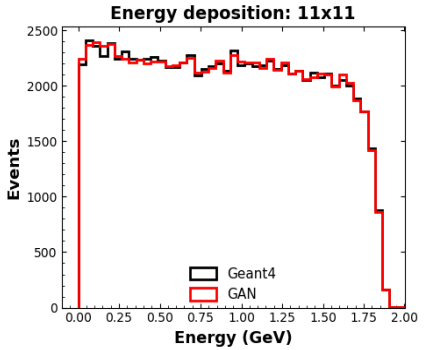
ECAL fast simulation using generative models

Refer to Xiao Yang's talk

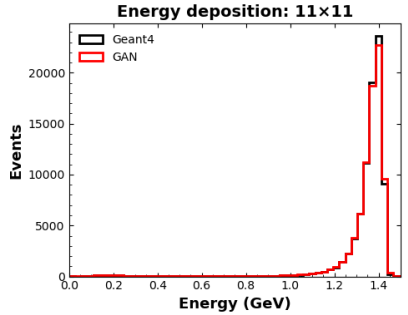
- Researching on hybrid technology of full and fast simulation
- ECAL simulation is the most time consuming part, novel methods like GAN are investigated



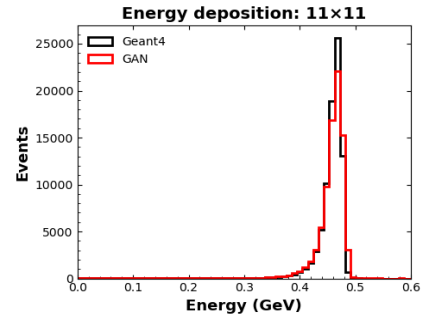
P = 0-2 GeV



P = 1.5GeV



P = 0.5GeV



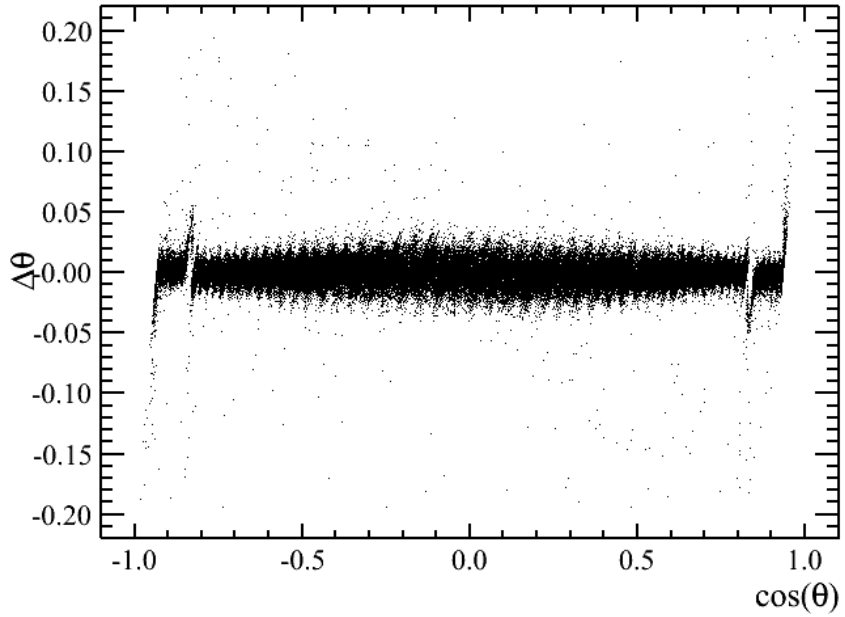
Summary

- Fast simulation framework is developed to cope with challenge of data generation
- Flexible for different detecting response and friendly for physics sensitivity study.
- More features under improvement

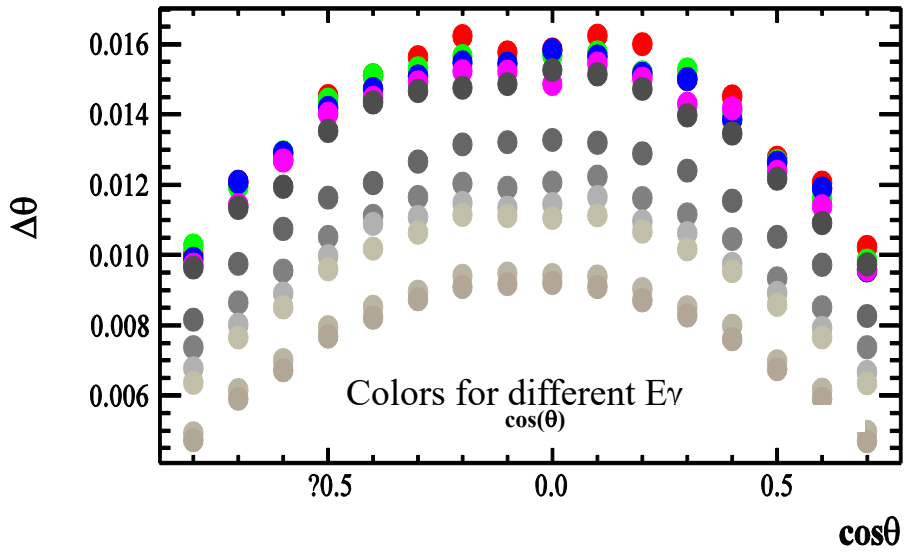
Thank you!

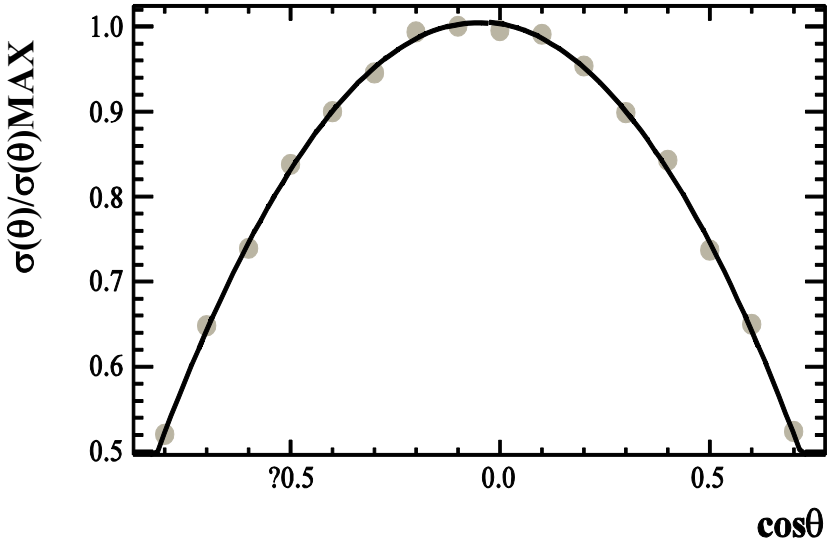
backup

Realization for track and shower — spatial parameterization



- Parametrization
- Can also also be realized via different CDF for momentum and polar angle bins





Parameterization
σ(pt, cos(θ))