HEP Software Development in China

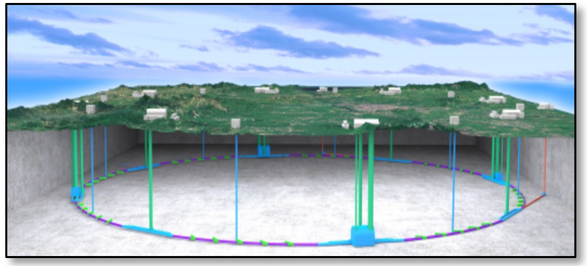
Xingtao HUANG

Shandong University

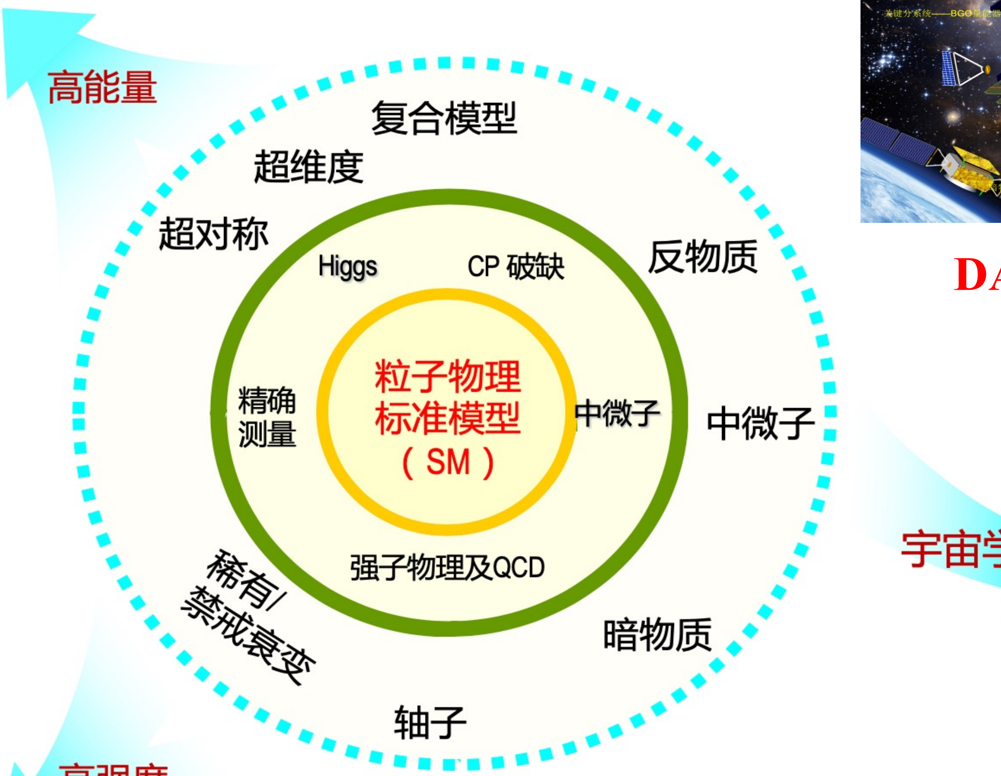
Outline

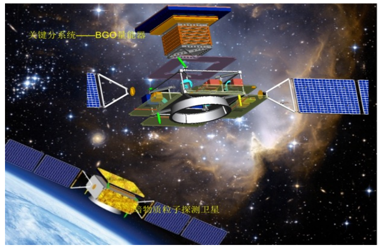
- Introduction
- Offline software of BESIII、CEPC、JUNO
- Offline software of STCF
- Summary

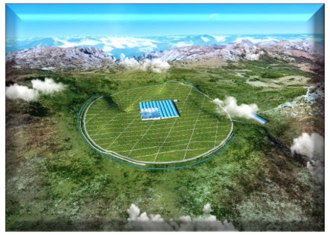
China based HEP Experiments



CEPC



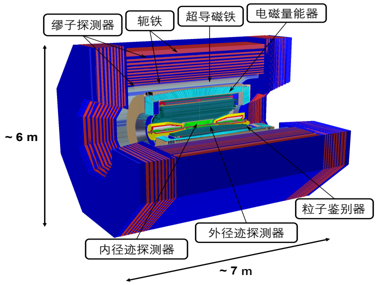




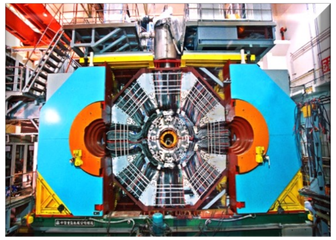
DAMPE

LHAASO

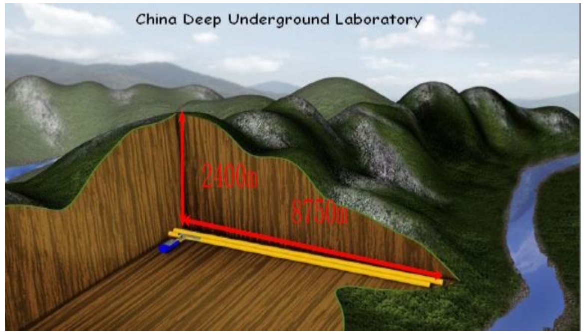
Daya Bay, JUNO



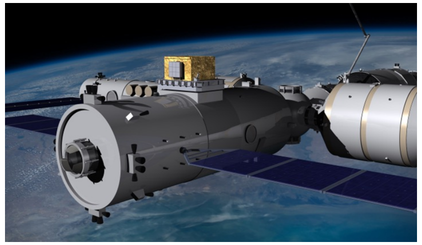
STCF



BESIII



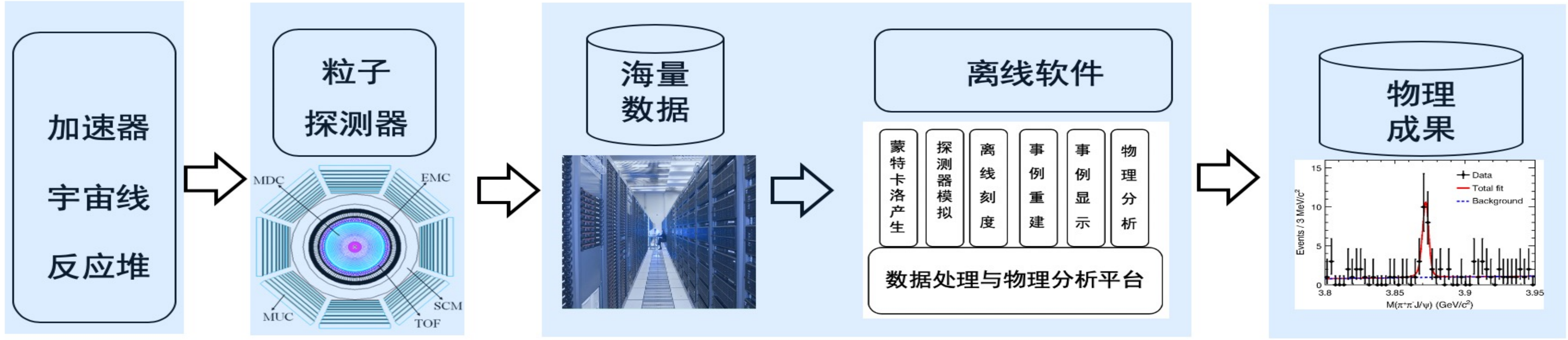
CJPL



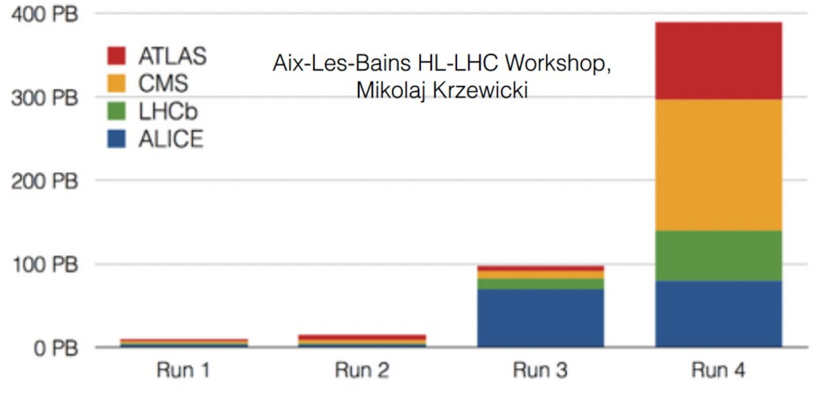
HERD

Key Technologies for HEP Experiments

Particle Source **Detector** **Computing and Software** **Physics Analysis**



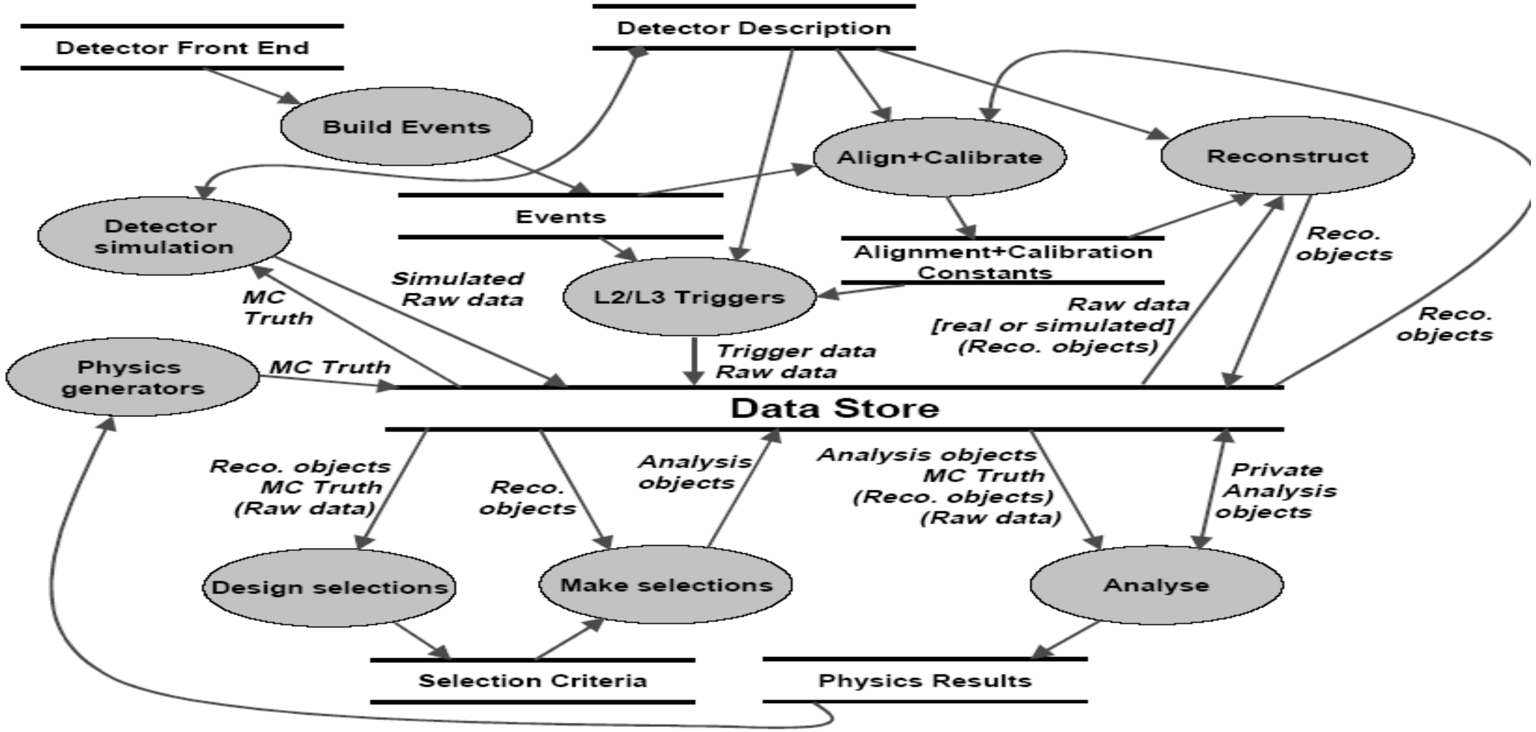
Experiments	Operation	Data Volume (PB/Year)	Total Data Volume（PB）
BESIII	2008-2028	0.5	10
LHAASO	2019-2030	1.2	12
JUNO	2024-2032	2.0	20
STCF	-	~ 300	-
CEPC	-	1.5-3（HF） 500- 5000（ZF)	-



RAW data to offline for LHC experiments

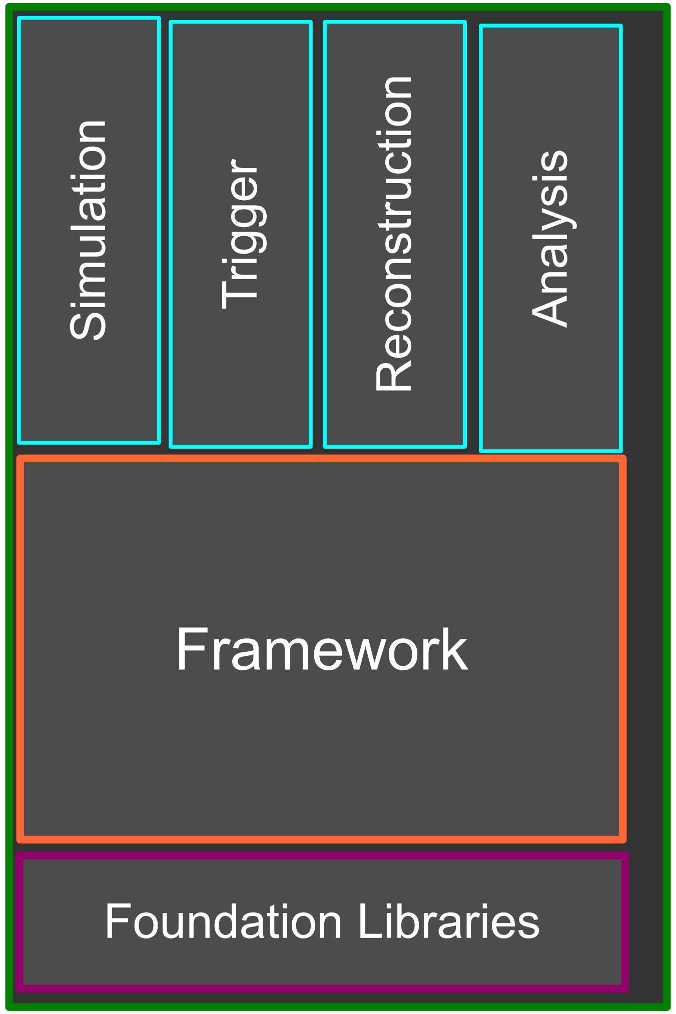
Offline Data Processing and Software

- Key features:
  - Massive amount of experimental data
  - Complex data processing processes
  - The experiment is large-scale, long-term, and carried out in international cooperation
  - High performance and more flexible software system are required



Technical Options

- **New technical challenges**
  - How to develop large scale software which can be run on diverse hardware like CPU /GPU /FPGA/ other accelerators
  - How to take advantage of novel technologies such as AI and Quantum Computing technologies to cope with the tasks that involve complicated data processing
- **Technology survey and our choices**
  - Some important HEP software packages are chosen as the cornerstone
    - Software framework: GAUDI/SNiPER
    - Simulation: GEANT4
    - Analysis: ROOT
    - Detector Description: DD4hep
    - Event Data Model: EDM4hep
  - Join the efforts to resolve the common challenges confronted by future HEP experiments (STCF，CEPC，FCC，ILC …)
  - At the same time, develop some specific software to meet the experiment’s needs



Several Popular Software Frameworks

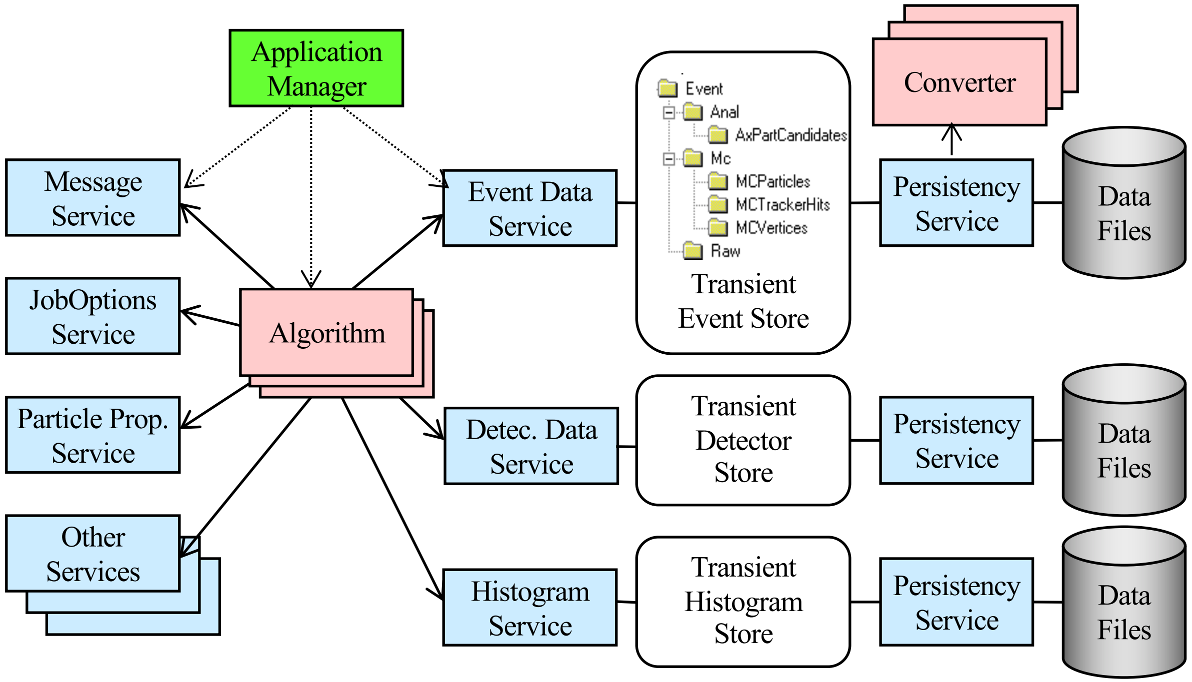
- ROOT: flexible and powerful, need more manpower for some service development
- ART: optimized for high intensity physics experiments and a little complex
- Gaudi: very popular for collider physics experiments, very comprehensive but a bit heavy
- SNiPER: lightweight and very flexible, suitable for both non-collider and collider experiments

	ROOT	ART	Gaudi	SNiPER
Job configuration	**Root script**	**FHiCL**	**Python /TXT**	**Python**
Learning curve	**S**	**L**	**L**	**S**
Adoption	**Phenix** **Alice**	**Mu2e,** **NOvA,** **LArSoft**	**ATLAS** **BESIII** **DYB**	**JUNO** **LHAASO** **STCF** **HERD** **nEXO**

Gaudi Framework

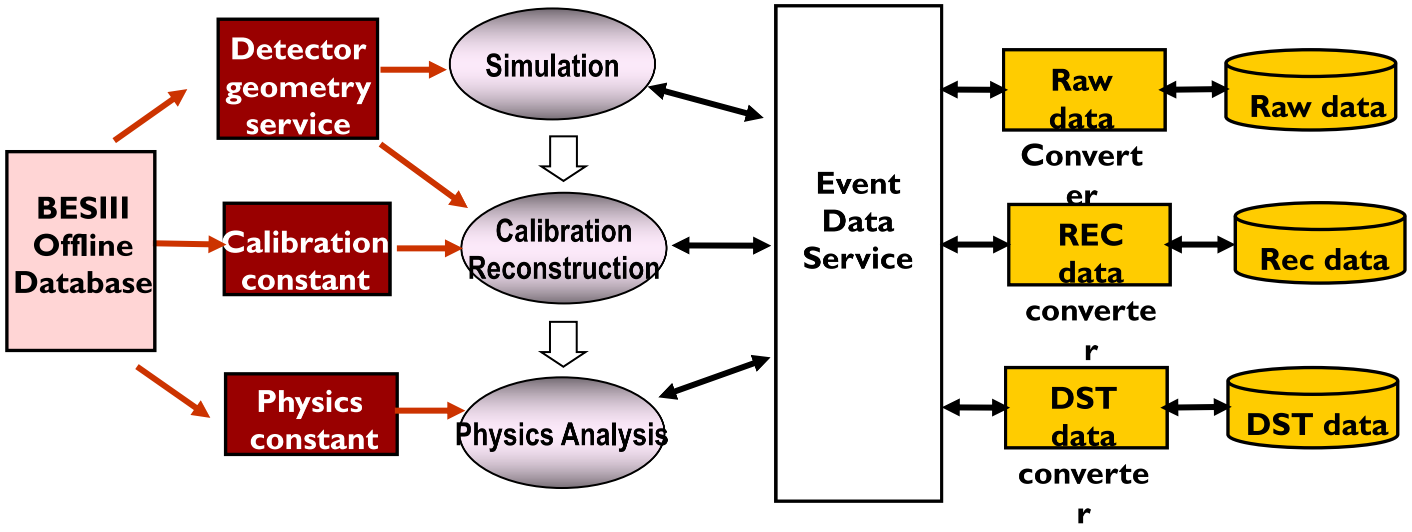
- Originally developed by LHCb and used by LHCb, ATLAS, HARP, GLAST, BESIII, Daya Bay etc.
- It provides a skeleton of applications into which developers plug in their code and provides most of the common functionality.

- Key Features
  - All components with well defined "interfaces": Algorithm, Data Object , Service, Converter ……
  - Clear separation between "algorithms" and "data" ("persistent data" and "transient data")
  - Clear separation between "persistent data" and "transient data"

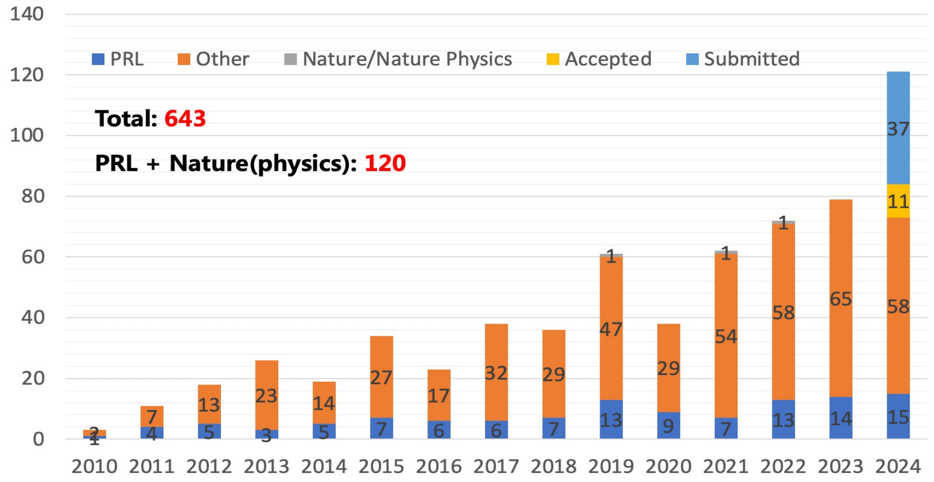


BESIII Offline Software System (BOSS)

- BESIII offline software system (BOSS) is an offline data processing software system
  - Developed based on GAUDI framework providing event data service and constants data service
  - Developed Simulation, reconstruction/calibration, and data analysis algorithms for data processing and physics analysis,
  - External Libs: Geant4, ROOT, GDML, MySQL, ……
  - Operation System: Cent OS 7, GCC 4.9.3 -> Alma Linux 9.3, GCC 13

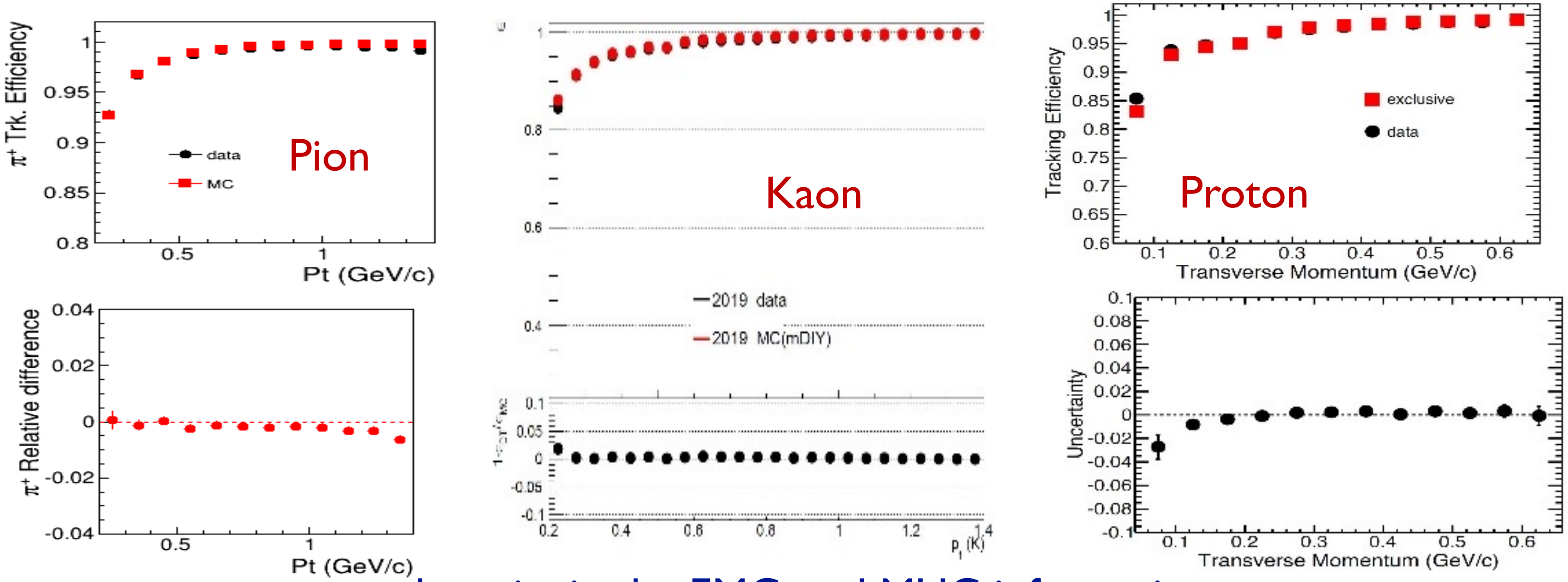


Publications by Sep. 5 2024)

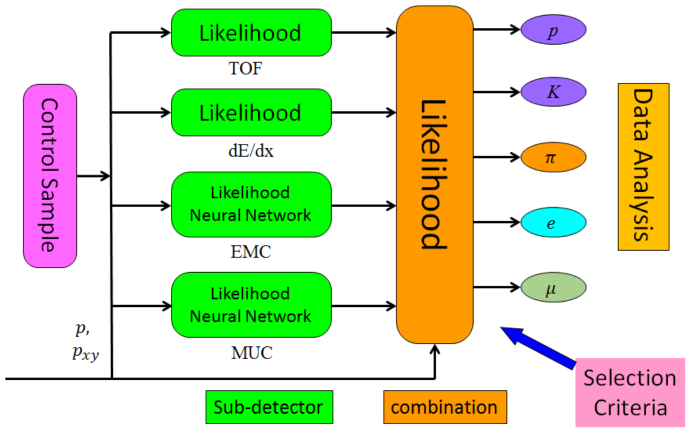


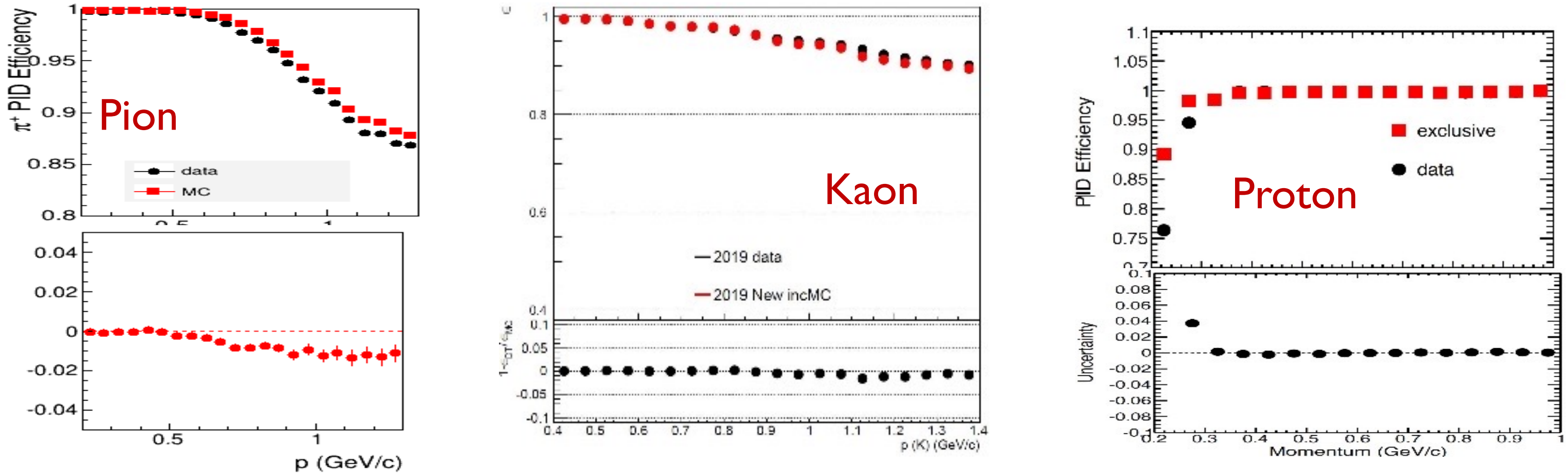
Tracking and PID

- Track finding
  - Pattern Matching
  - Conformal Transform
  - Hough Transform
- Track fitting
  - least square method
  - Kalman filter method



- PID: combine TOF information, dE/dx measurements, energy deposits in the EMC, and MUC information.

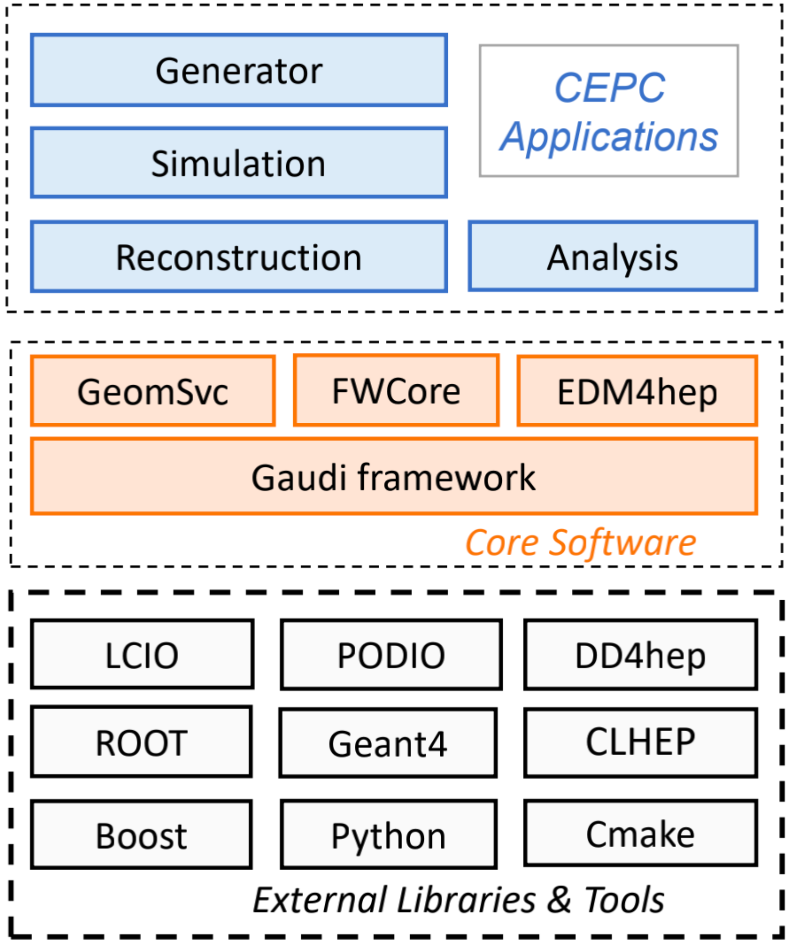




CEPC Software (CEPCSW)

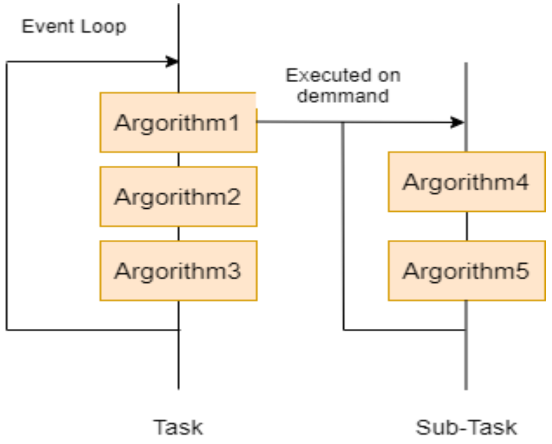
- **CEPCSW software structure**
  - Core software
  - Applications: simulation, reconstruction and analysis (see talks given by Weidong and Shengshen)
  - External libraries
- **Core software**
  - Gaudi/Gaudi Hive: defines interfaces to all software components and controls their execution.
  - EDM4hep: generic event data model
  - K4FWCore: manages the event data
  - DD4hep: geometry description
- **CEPC-specific framework software: generator, Geant4 simulation, beam background mixing, fast simulation, machine learning interface, etc**
- **Now build the offline system of the experiment to support detector R&D and physics potential studies (RefDetTDR)**

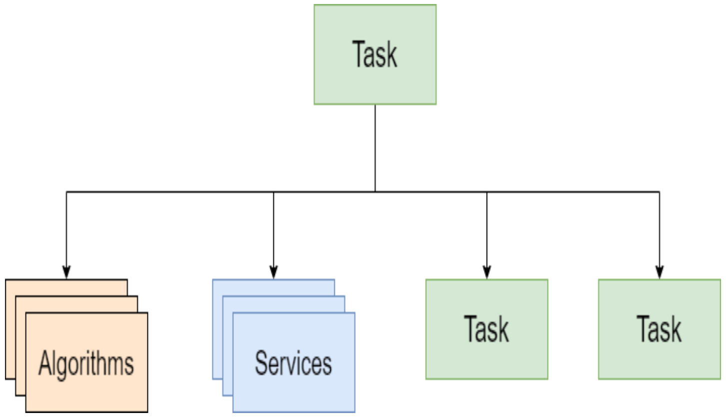
https://github.com/cepc/CEPCSW

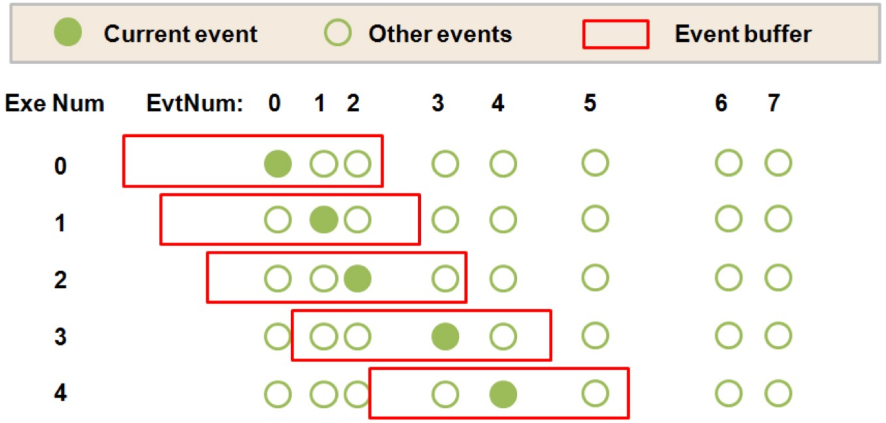


SNiPER Framework

- SNiPER(Software for Non-Collider Physics Experiment) : a general-purpose software framework
  - Light weighted: only dependent on Boost. Python
  - Highly modularized and extensible
  - Originally developed for JUNO, but also adopted by several experiments: LHAASO, STCF, HERD, nEXO, etc.
- Key components
  - Algorithm: an unit of event data proceeding
  - Service: an unit for common functions that can be called by users, anywhere when necessary.
  - Task: a lightweight application manager to assemble specific algorithms, service as well as sub-tasks.!

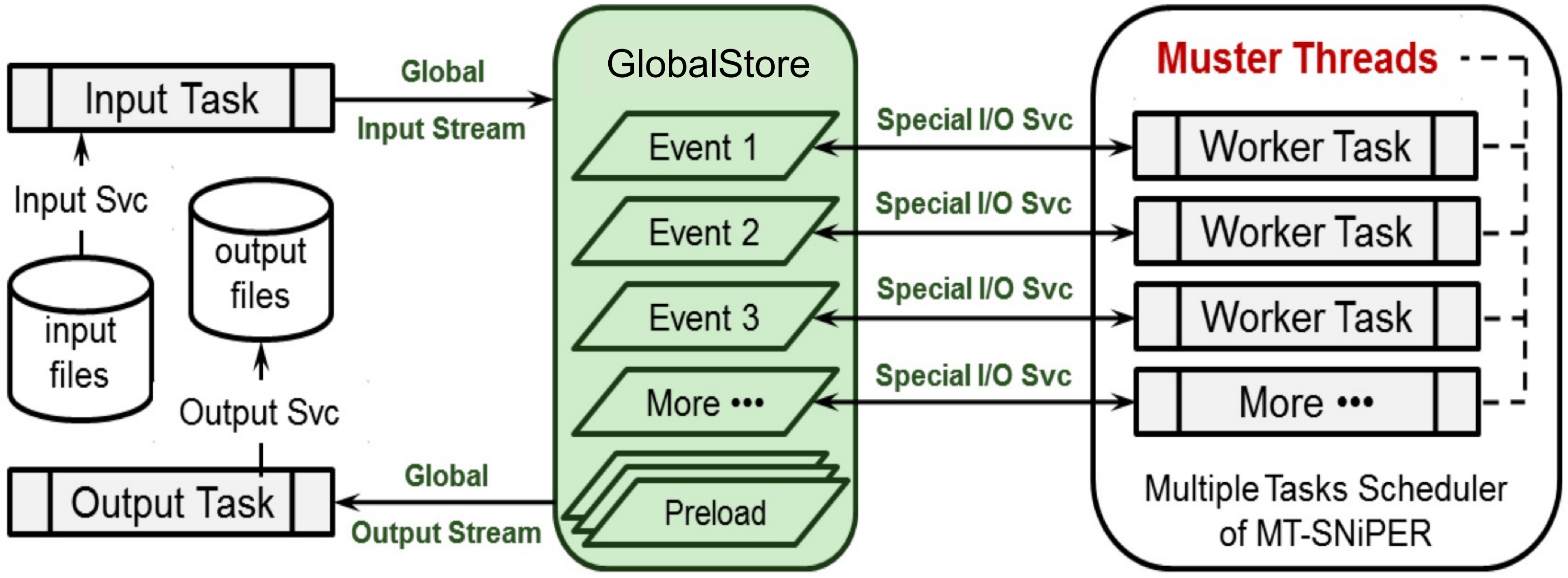






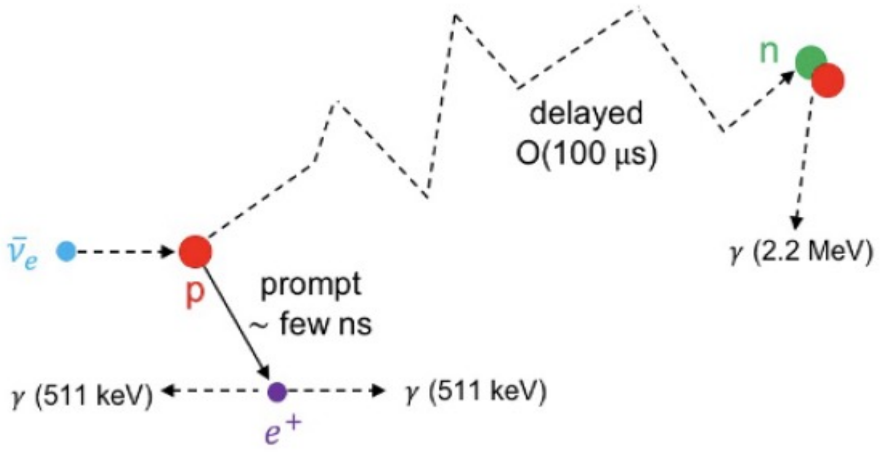
Parallelism in MT-SNiPER

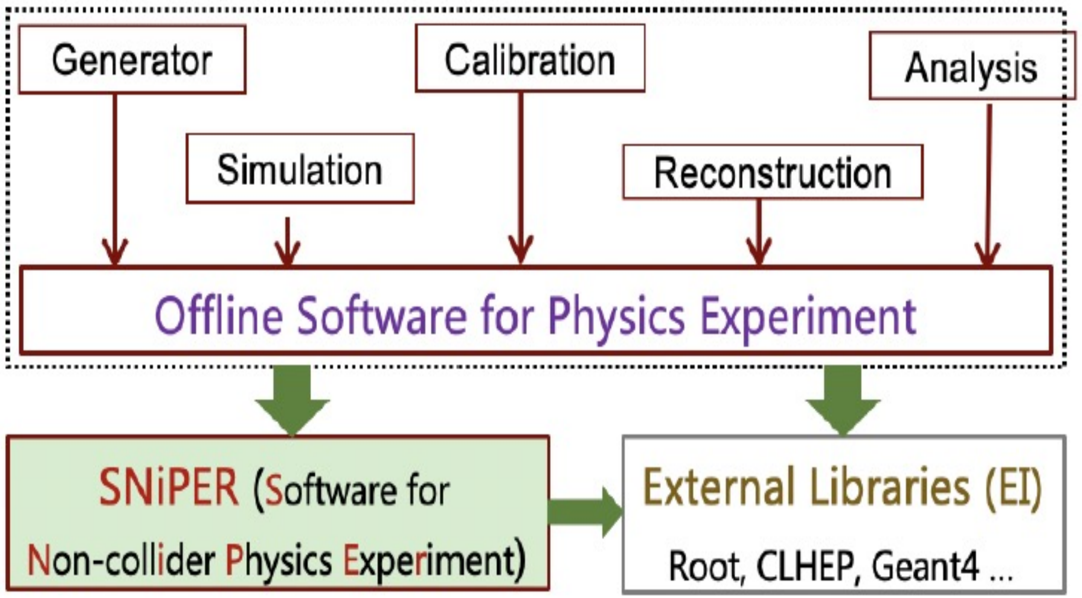
- SNiPER provides simple interfaces for building multithreaded applications
  - Based on Intel TBB
  - SNiPER Muster (Multiple SNiPER Task Scheduler) works as a thread pool/scheduler
  - Data I/O is binded to dedicated I/O thread for flexibility
  - A Global Store is developed to support multithreaded event data management
  - Application code is mostly consistent for serial and parallel execution



JUNO Offline Software (JUNOSW)

- Developed based on SNiPER framework since 2012
- Support event splitting and hit-level event mixing
- Support time-correlated analysis (event buffering)
- Support multi-threaded sim., recon., and analysis
- Support event correlation at different processing stages
- Integration of machine learning methods
- The JUNO data processing and analysis chain was completed and has been used by several rounds of M.C. data challenge, and has been used for processing commissioning data this year

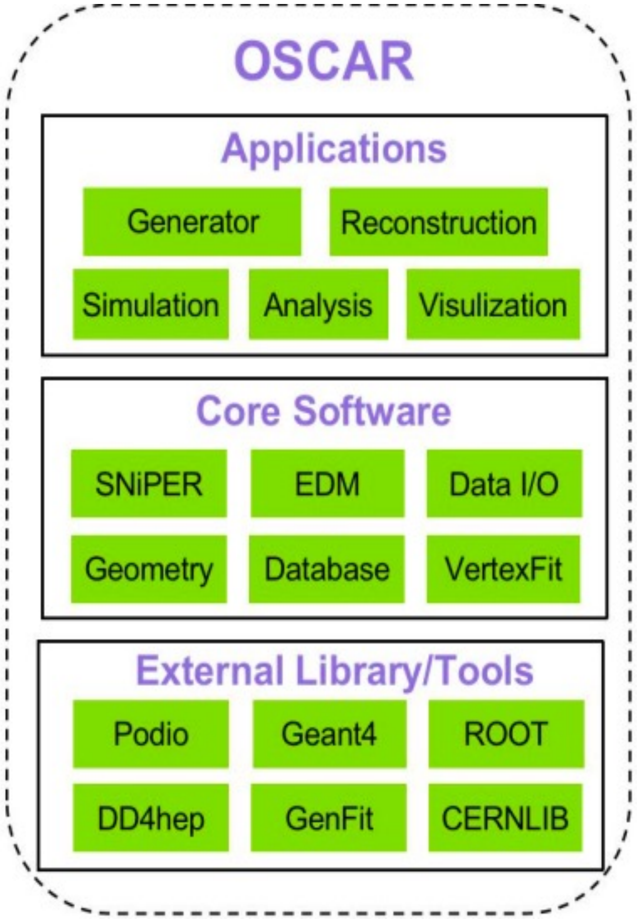




STCF Offline Software System

More details in Teng Li' talk (Nov. 21):
STCF core software and analysis tools

- The Offline Software of Super Tau-Charm Facility （OSCAR） is designed for detector design, MC data production and physics analysis at STCF：
  - Applications: STCF specific application software
  - Core software: common platform for the offline software
  - External libraries and tools
- Core software are developed for common functionalities
  - Event loop control (sequentially and concurrently)
  - Detector data and event data management
  - Common tools for data analysis
  - To support efficient parallel and heterogeneous computing
- Some applications are ported from BESIII

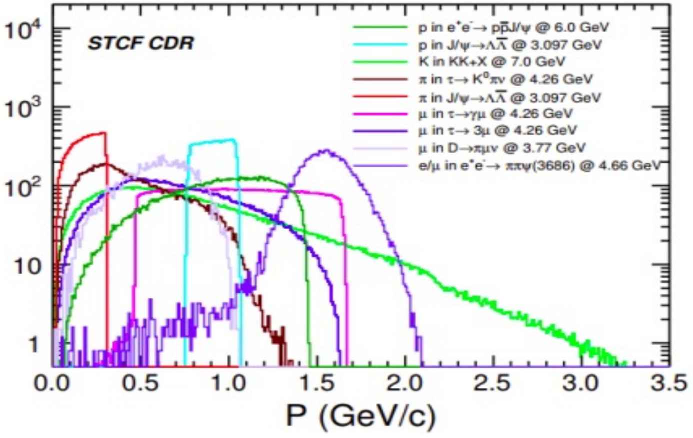


W.H. Huang et al 2023 JINST 18 P03004

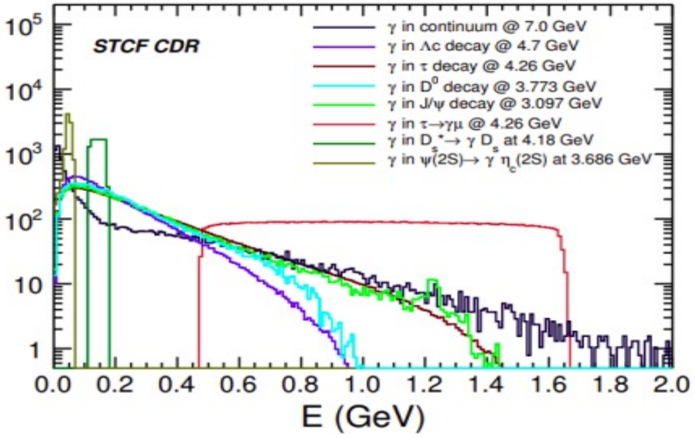
Physics requirements

- **To achieve physics goals:**
  - Higher event rate： 400 kHz
  - Big data volume: ~30 GB/s
  - final particles
    - e, μ, K, π, proton (p up to 3.5 GeV/c, most with p < 2 GeV/c, lots of particles with p < 400 MeV/c)
    - γ (energy coverage from 25 MeV to 3.5 GeV); KL, neutron (up to 1.6 GeV)

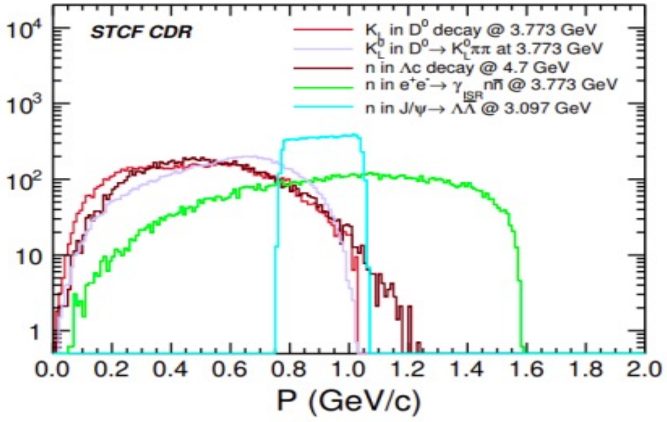
Component	Num. of channels	Readout time window	Event size (B)	Total (B/s)
ITK (Silicon)	50M	500 ns	14300	5.72G
ITK (μRWELL)	10552	500 ns	17232	6.89G
MDC	11520	1 μs	20400	8.16G
PID (RICH)	518400	500 ns	15600	6.24G
PID (DTOF)	6912	500 ns	7380	2.95G
EMC	8670	500 ns	15000	6.00G
MUD	41280	500 ns	262	105M
Total(Silicon)	50.6M	–	72.9k	29.2G
Total(μRWELL)	594k	–	75.9k	30.4G



Charged particle momentum

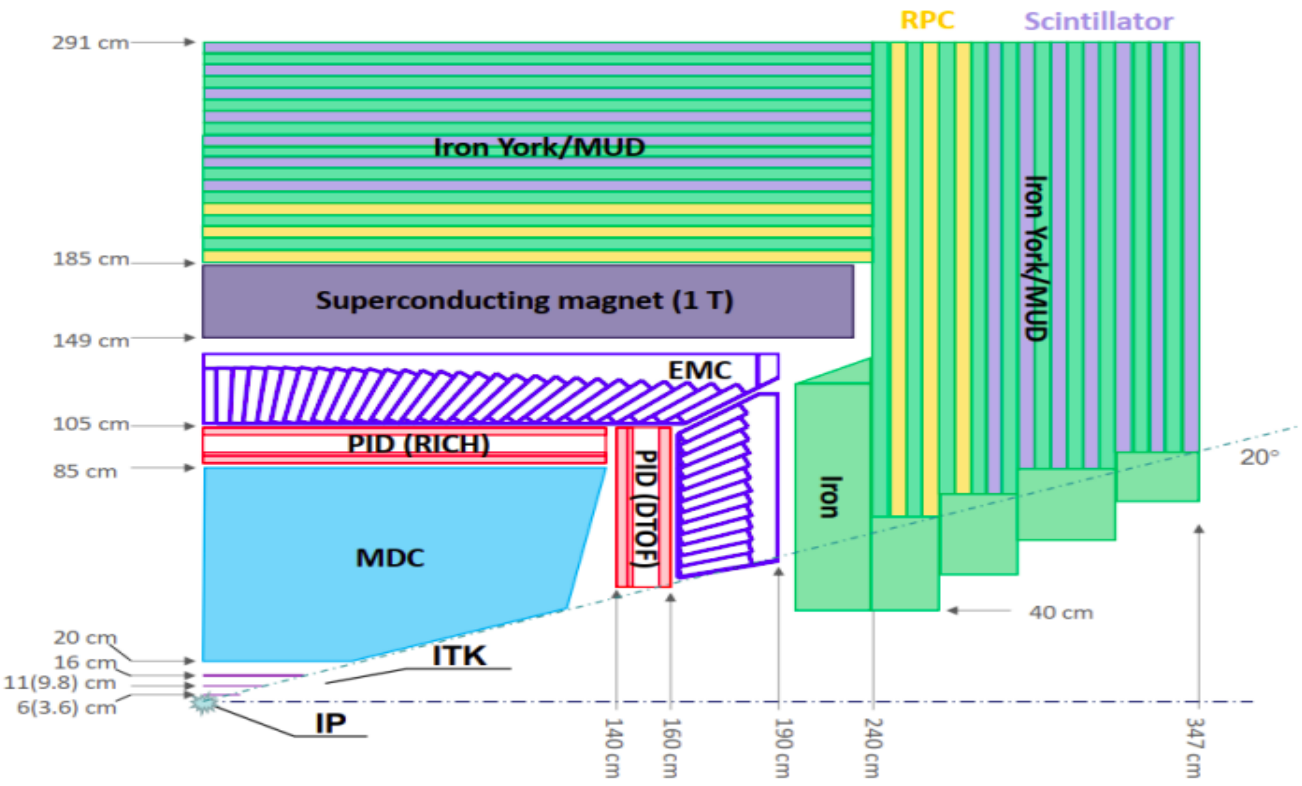


Photon energy



KL and neutron momentum

Physics requirements



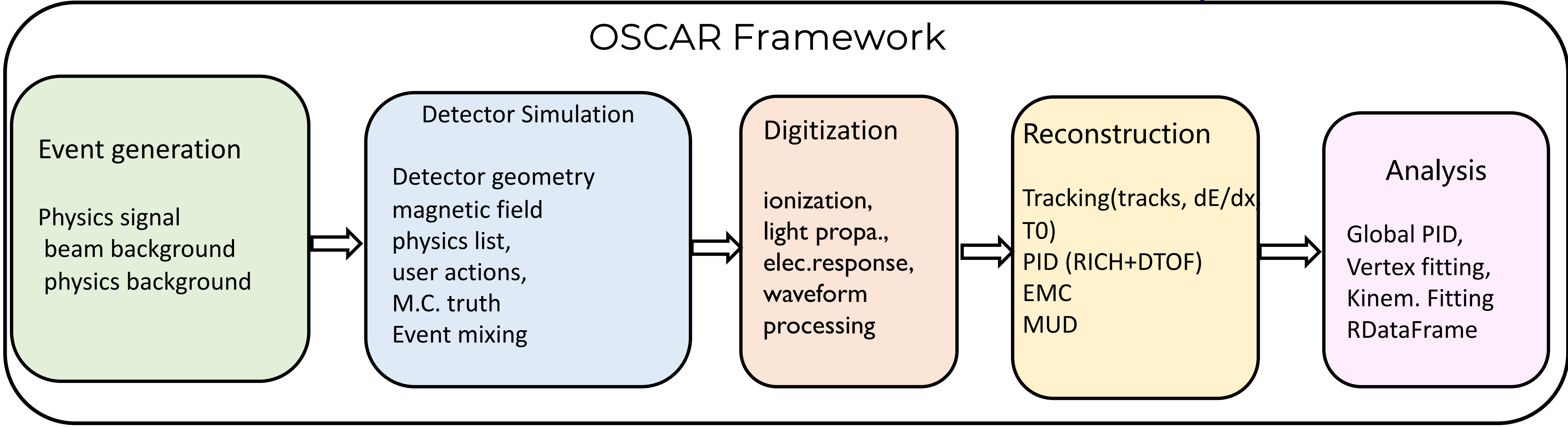
Performance requirements

- ITK:
  - Material < $0.01X_0$, σ_{xy} < 100 μm
- MDC:
  - Material < $0.05\ X_0$
  - σ_{xy} < 130 μm, $\sigma(p)/p$ < 0.5% at 1 GeV/c
  - dE/dx resolution < 6%
- PID:
  - $3\ \sigma\ \ \pi/K$ separation, PID efficiency > 97% up to 2 GeV
- EMC:
  - σ_E < 2.5%, σ_{pos} ~ 4 mm, σ_t ~ 300 ps at 1 GeV
- MUD:
  - μ efficiency > 95% above 0.7 GeV with $\pi \to \mu$ misidentification rate <3%

- Good efficiency and resolution for tracks (photon)
- Powerful PID between charged (neutral) particles

Status of STCF Offline Software

- Dedicated Core Software System including the underlying framework, DM system, GMS etc. are developed
- Full chain of detector simulation and physics object reconstruction has been built
- Physics analysis with full detector simulation and reconstructed objects is supported
- Parallelized simulation and reconstruction based on MT-SNiPER is ready



Status of STCF Offline Software

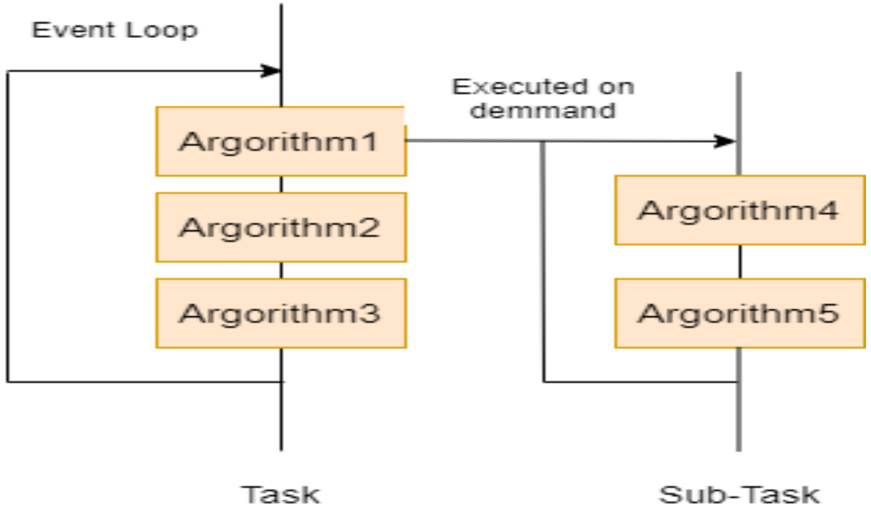
Finalised · Working, under optimization · Under development · Unknown/Not started

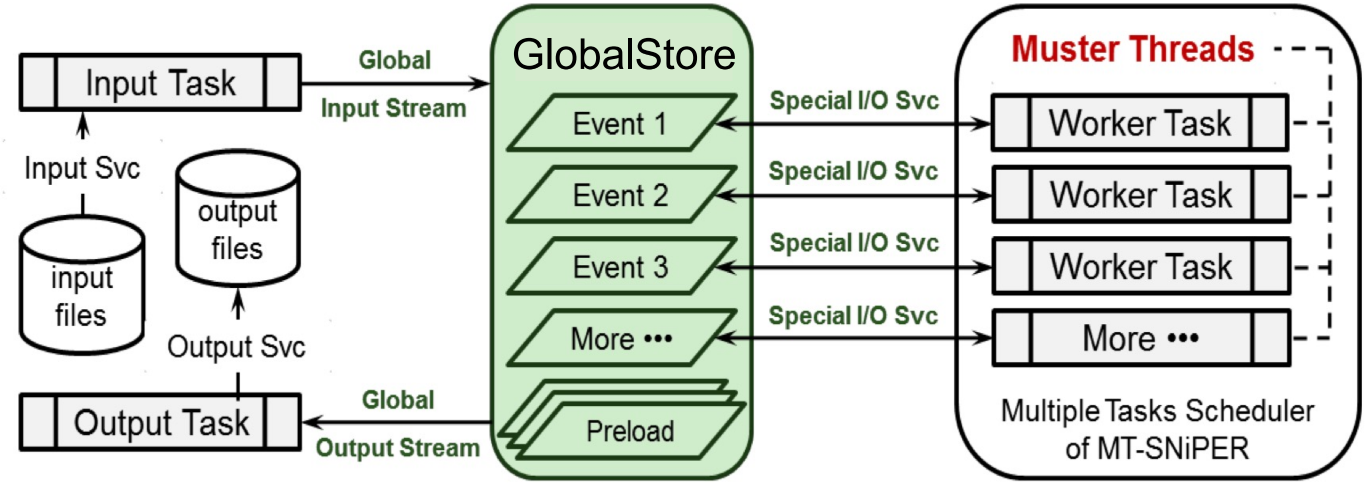
<table>
<tr><th rowspan="2"></th><th rowspan="2">Simulation</th><th rowspan="2">Digitization</th><th colspan="2">Reconstruction</th><th colspan="3">Analysis (W/O bkgs)</th></tr>
<tr><th>On top of Digitization</th><th>On top of Bkgs + Digitization</th><th>Global PID Charged</th><th>Global PID Neutral</th><th>Vertex/Kinematic Fit</th></tr>
<tr><td>ITK</td><td>Finalised</td><td>Working, under optimization</td><td rowspan="2">Working, under optimization</td><td rowspan="2">Working, under optimization</td><td rowspan="7">1) Single particles
Finalised
2) Physics Processes
Working, under optimization</td><td rowspan="7">1) Single particles
Working, under optimization
2) Physics Processes
Unknown/Not started</td><td rowspan="7">Finalised</td></tr>
<tr><td>MDC</td><td>Finalised</td><td>Working, under optimization</td></tr>
<tr><td>RICH</td><td>Finalised</td><td>Working, under optimization</td><td>Working, under optimization</td><td>Working, under optimization</td></tr>
<tr><td>DTOF</td><td>Finalised</td><td>Working, under optimization</td><td>Finalised</td><td>Finalised</td></tr>
<tr><td>BTOF</td><td>Finalised</td><td>Working, under optimization</td><td>Under development</td><td>Under development</td></tr>
<tr><td>ECAL</td><td>Finalised</td><td>Working, under optimization</td><td>Finalised</td><td>Finalised</td></tr>
<tr><td>MUD</td><td>Finalised</td><td>Under development</td><td>Working, under optimization</td><td>Working, under optimization</td></tr>
</table>

Underlying framework: SNiPER

More details in Teng Li' talk (Nov. 21): STCF core software and analysis tools

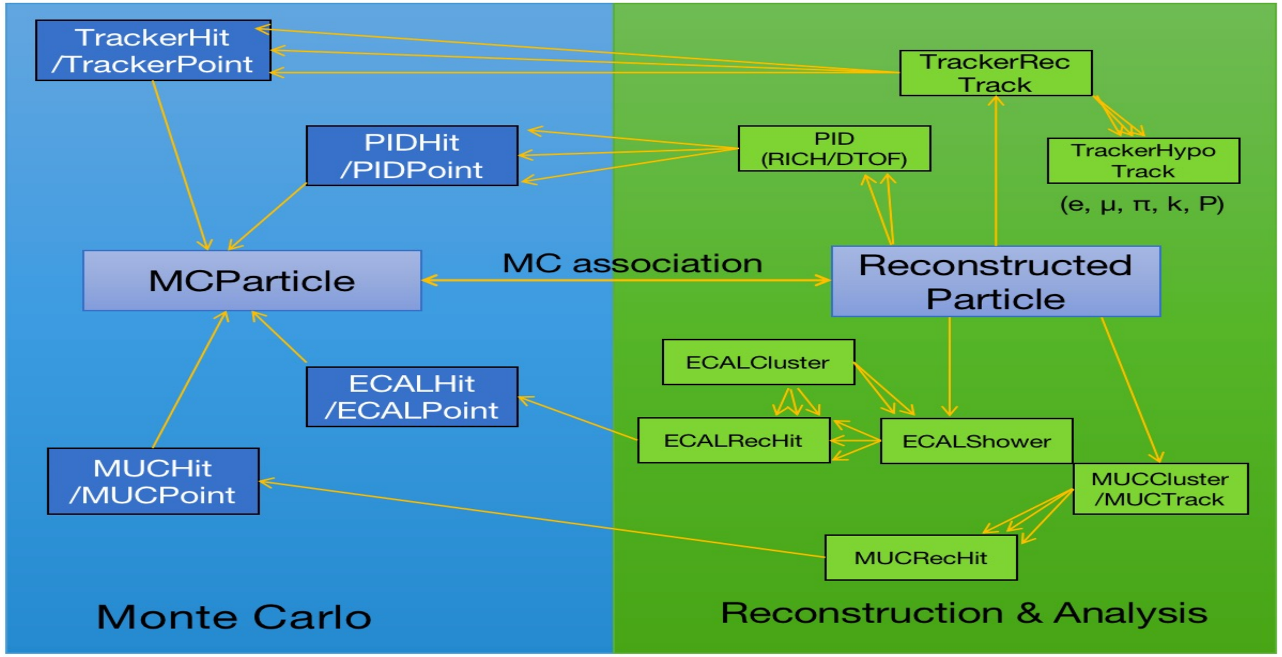
- Developed since 2012, very lightweight and flexible, supporting both non-collider experiments and collider experiments
- Providing basic functions of event loop control, application interface, job configuration, logging etc.
- Providing simple interfaces for building multi-threaded applications, thus supporting both serial and parallel event processing with extension of SNiPER, Muster
- Adopted by JUNO (neutrino), LHAASO (cosmic ray), nEXO (neutrinoless double beta decay) and HERD (dark matter)

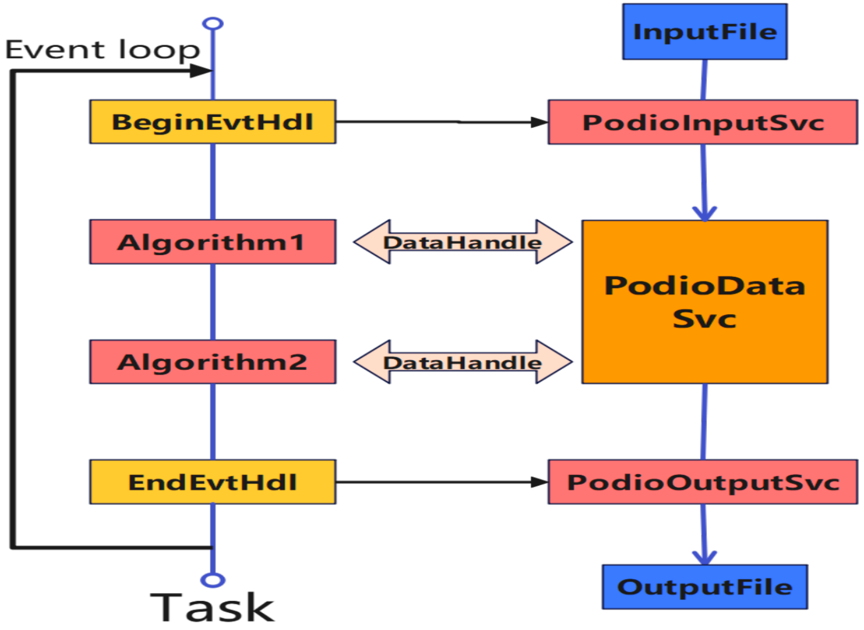




Event Data Model (EDM) and Serial Event Data Management

- Very crucial and taking significant effect on performance of the event processing
- Developed STCF EDM based on podio, which is also adopted by EDM4hep
  - Define event data and relationship between data object in YAML file
  - Re-use MCParticle and ReconstructedParticle from EDM4hep as the core index
  - Good support for multithreading
- Extended SNiPER data management system based on podio

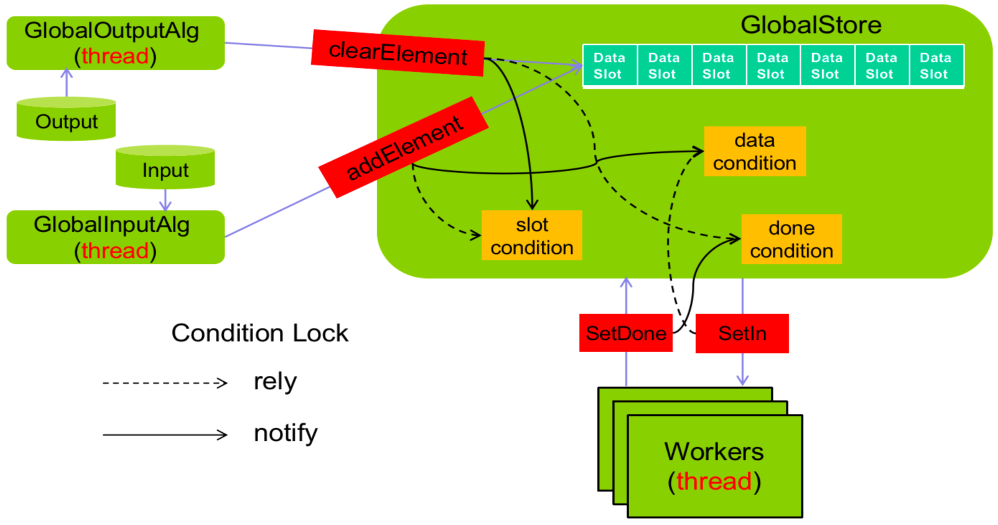


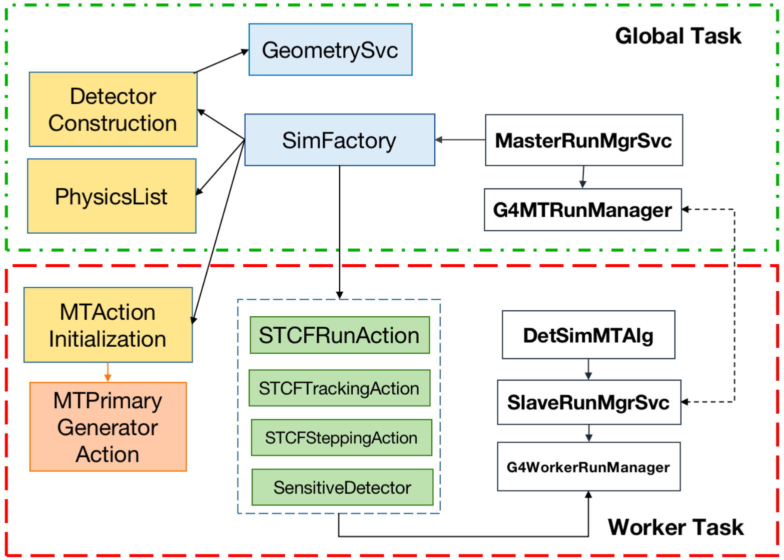


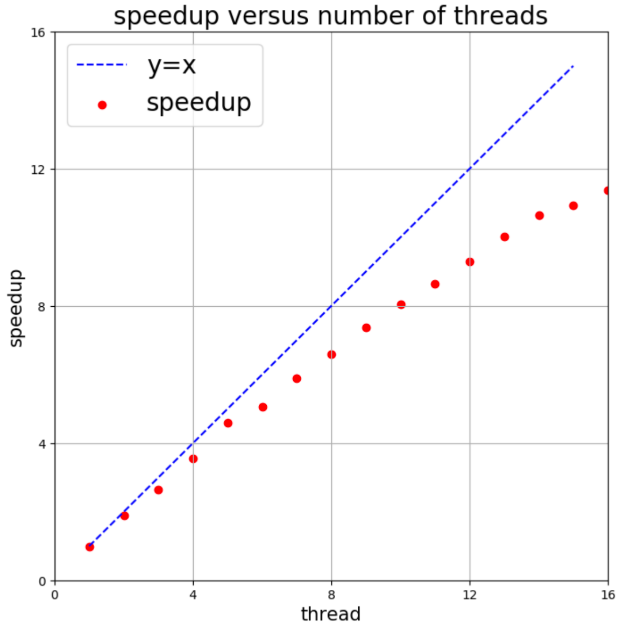
Parallelized Event Data Management

More details in Teng Li' talk (Nov. 21): STCF core software and analysis tools

- Developed parallelized Event Data Management for parallelized event processing
  - GlobalStore to cache multiple events (each within one data slot)
  - Two dedicated threads for input and output respectively
  - SNiPER Muster (Multiple SNiPER Task Scheduler) works as a thread scheduler
- Parallelized detector simulation and reconstruction are developed
- Track-level and algorithm-level parallelism are under R&D



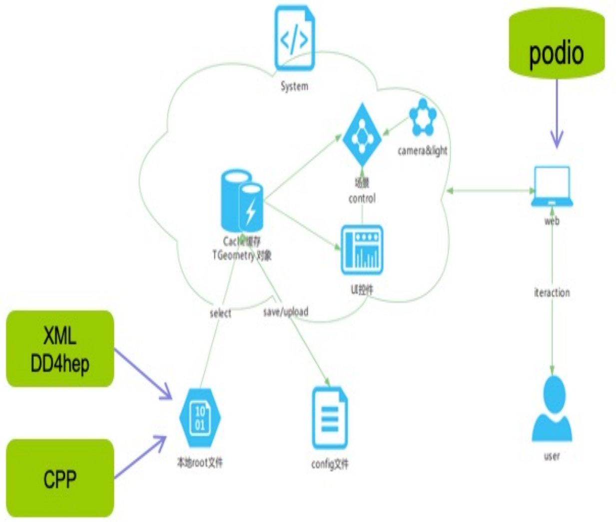


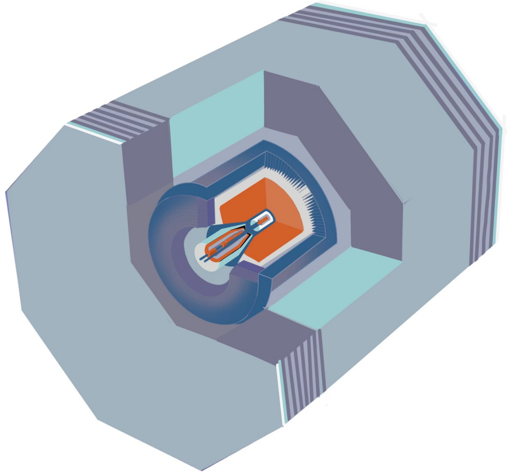


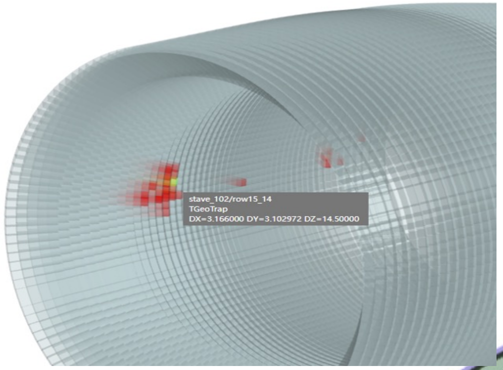
Detector and Event Display

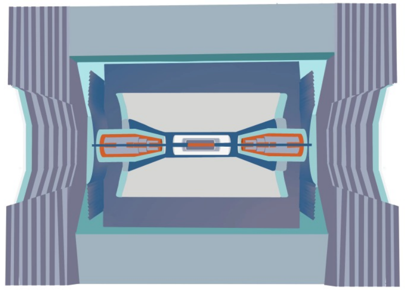
More details in Teng Li' talk (Nov. 21):
STCF core software and analysis tools

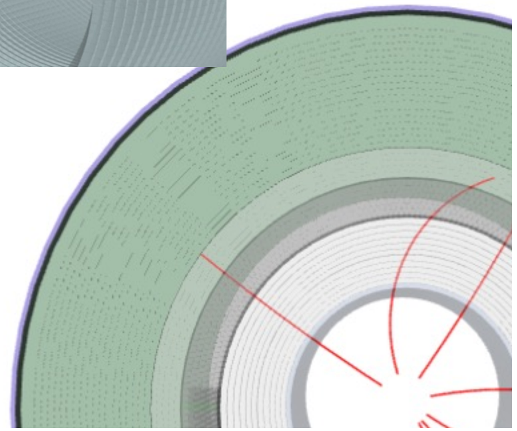
- A common geometry and event display system is being developed to visualize detector geometry, simulated hits and reconstructed particles
  - Based on Web3D technology and the open-source JSROOT
  - 3D engine and graphic library based on Three.JS
  - Using the Vue.js HTML5 development framework to implement the Web interface
  - Beta version released last month







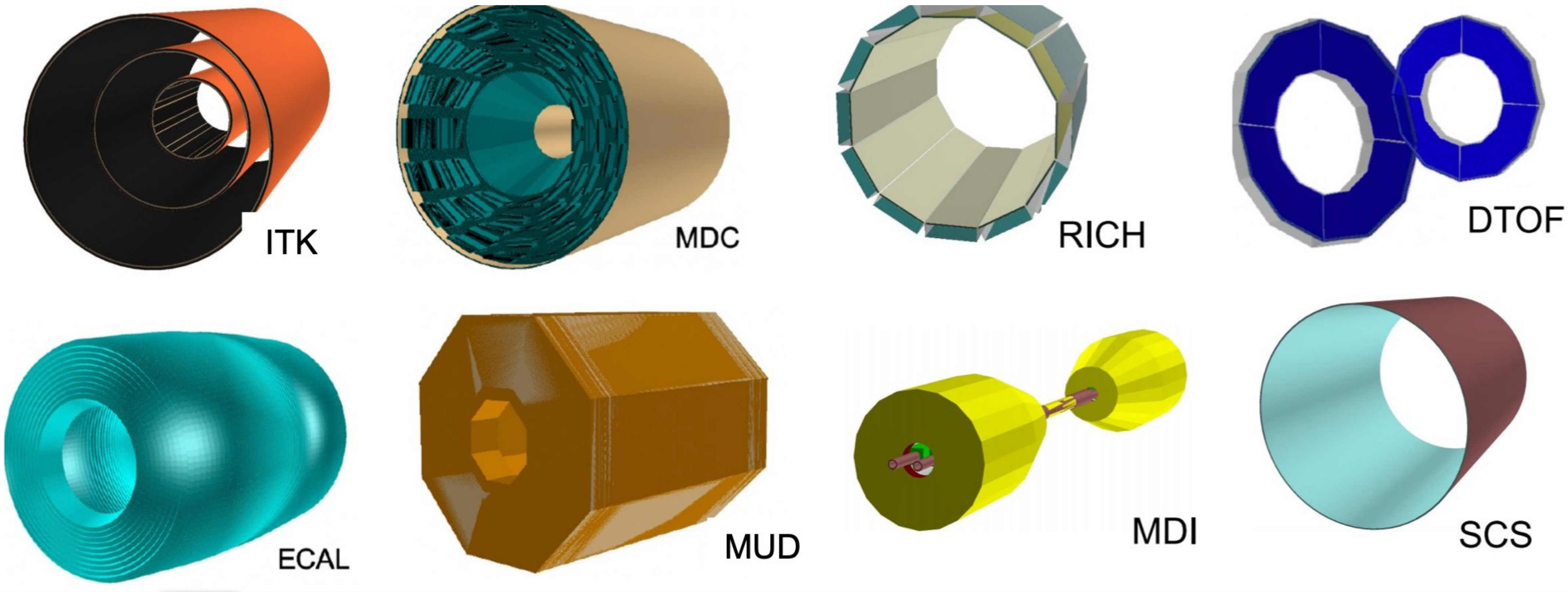




Detector Description

More details in Teng Li' talk (Nov. 21): STCF core software and analysis tools

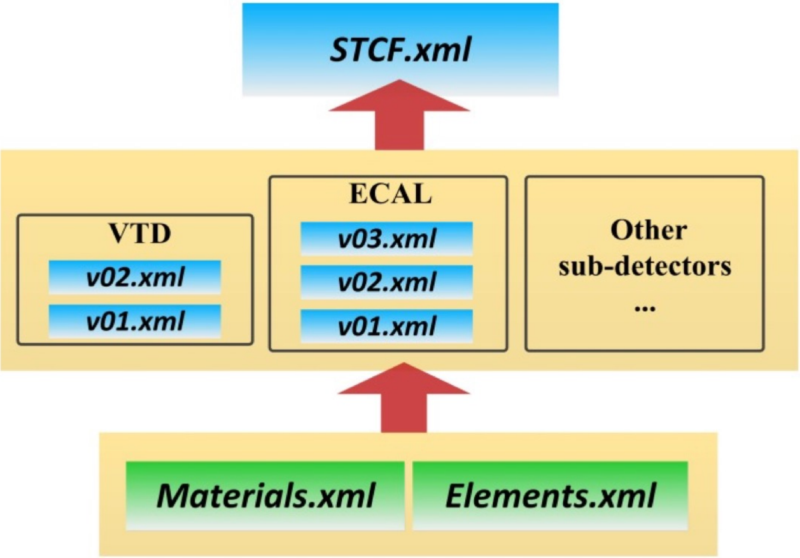
- The full STCF Detector is described with DD4hep
- Each sub-detector is implemented with a single compact file
- Very convenient to optimize detector geometry according to detector experts

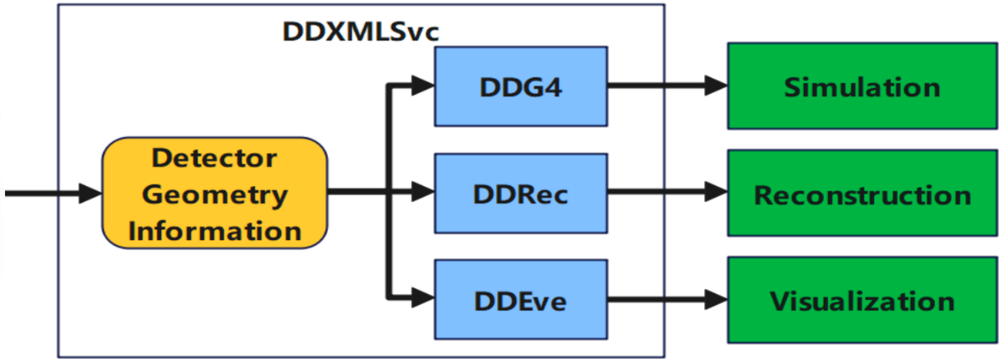


Geometry Management System

More details in Teng Li' talk (Nov. 21): STCF core software and analysis tools

- Developed geometry management system for all applications
  - detector simulation, reconstruction and event display
  - The version number is used for different design options
- Ensure consistent detector information between different applications with a single source of detector description
  - DDG4 for delivering detector geometry to Geant4
  - DDRec for delivering detector geometry to reconstruction algorithms
  - DDXMLSvc: the unified interface to DD4hep, including DDG4 and DDRec

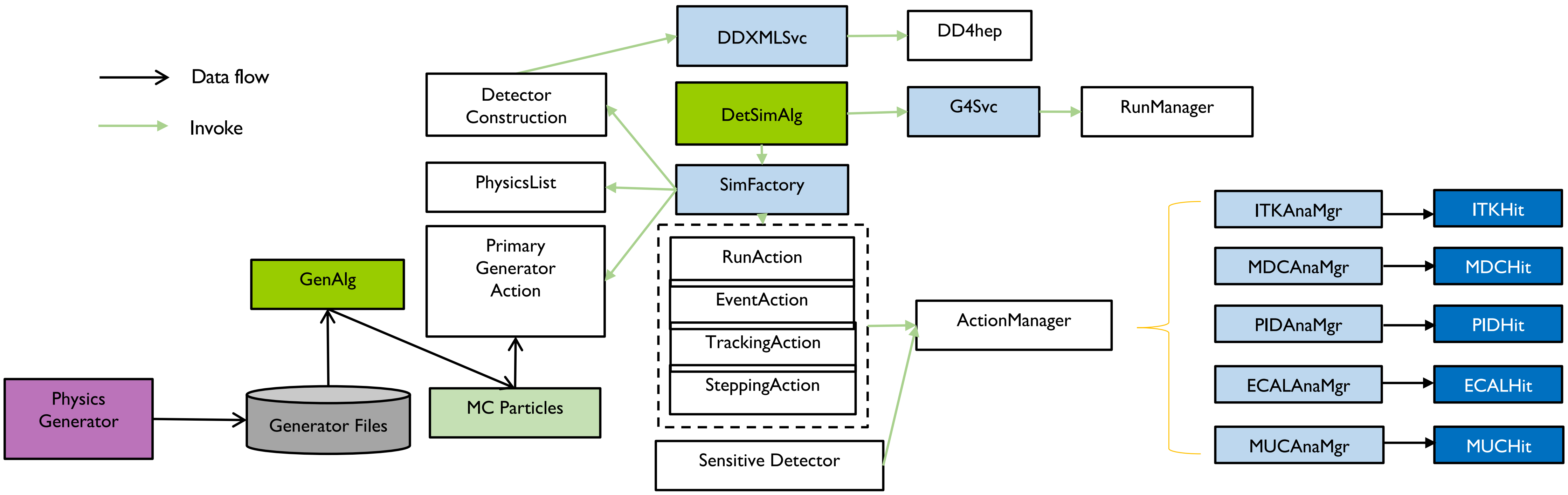




Detector Simulation Chain

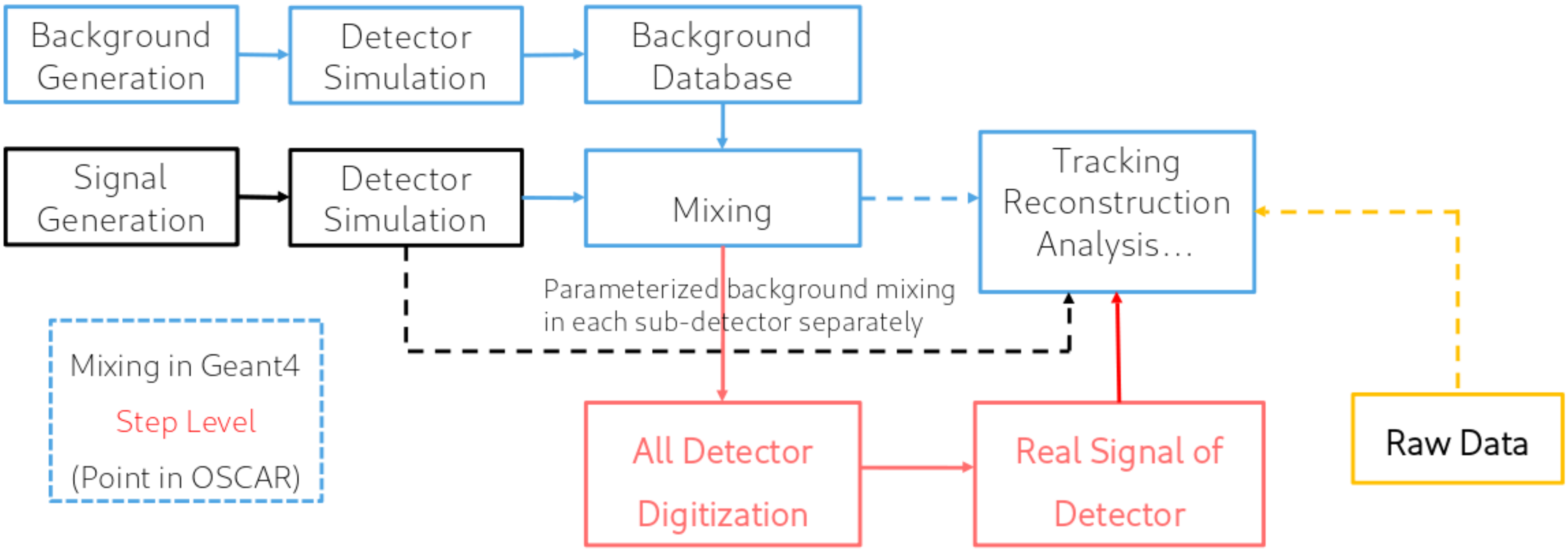
More details in Teng Li' talk (Nov. 21):
STCF core software and analysis tools

- Full chain of detector simulation from generator to simulated information is built
- Providing flexible configuration for different purposes of detector simulation
  - Generator for different physics topics i.e. Babayaga, KKMC, Phokhara, DIAG36, BBBrem …
  - Geometry for different detector design options
  - User actions for recording MC truth information



Schema of Event Mixing

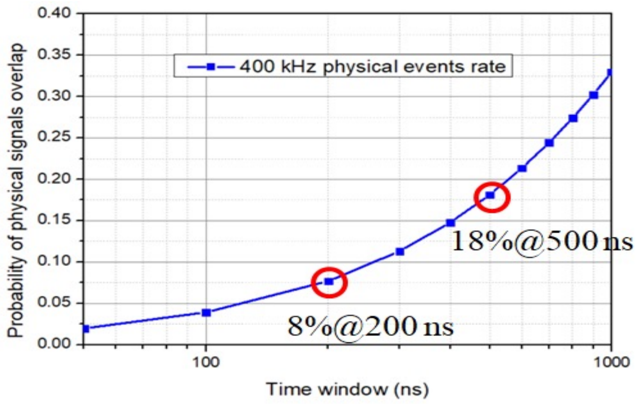
More details in Binbin Qi' talk (Nov. 19):
STCF simulation and digitization

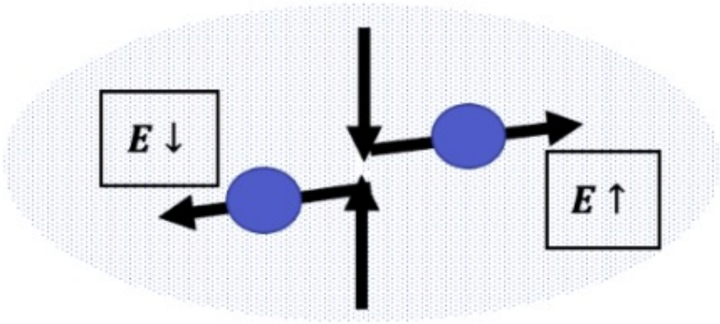


Background Production

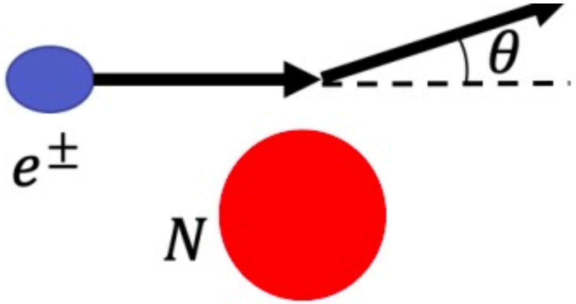
More details in Binbin Qi' talk (Nov. 19): STCF simulation and digitization

- Maximum event rate at STCF: 400 kHz
  - Maximum time window in sub-detectors : 1000 ns
  - Probability of events overlapping in1000 ns: 32.9%
  - Multiple physics events can exist in one event, i.e. pileup
- Main Backgrounds: Touschek, beam-gas, luminosity-related
- Setup background simulation software and produced background data samples

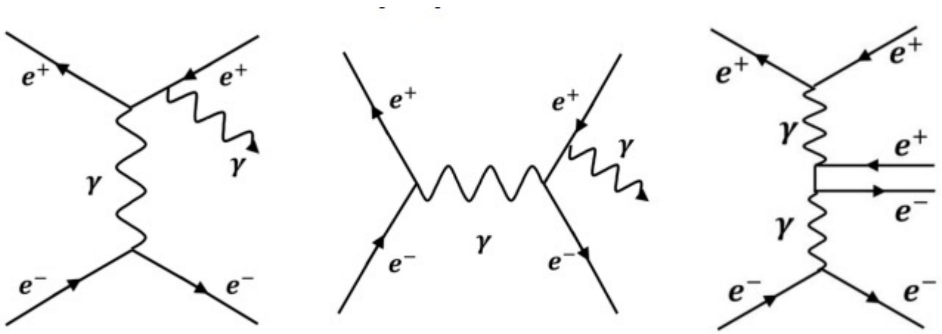




Touschek



Beam-gas

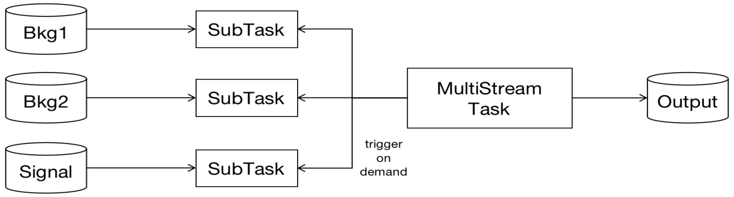


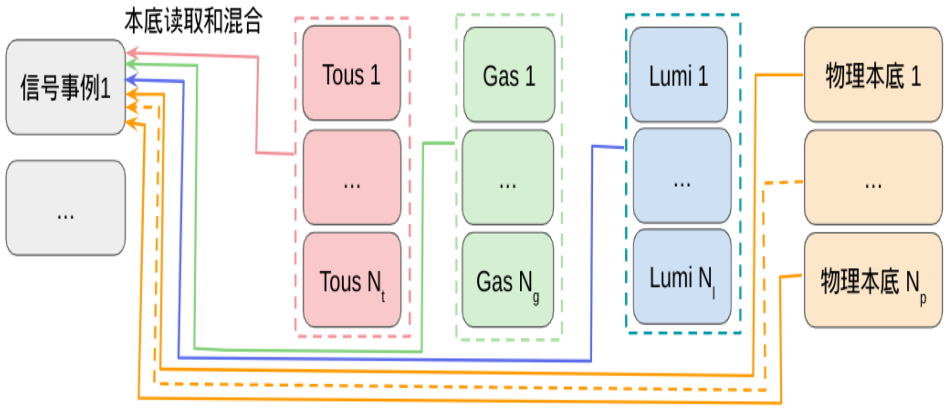
Luminosity-related

Event Mixing Algorithm

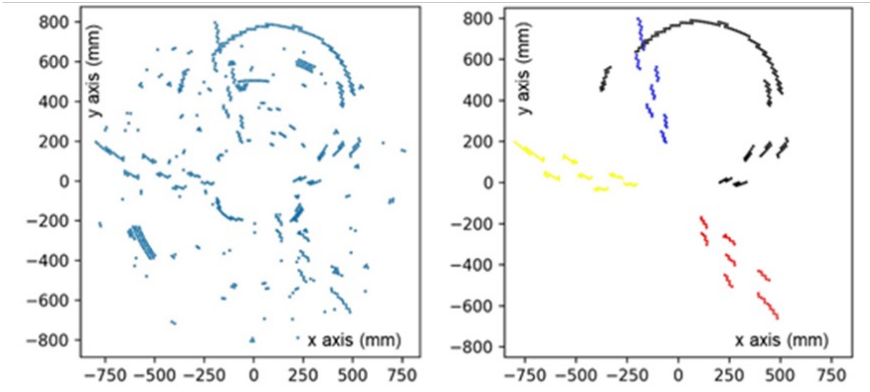
More details in Binbin Qi' talk (Nov. 19): STCF simulation and digitization

- Developed mixing algorithm based on multi-stream functions provided OSCAR
- Providing flexible configuration for event mixing
  - Signal, e.g. $e^+e^- \to \pi^+\pi^- J/\psi(\to \mu^+\mu^-/e^+e^-)$
  - Backgrounds: Touschek, beam-gas, luminosity-related, reading from background database
  - Underlying physics background: $e^+e^- \to anything$ at 4.26 GeV
- Mixed data used for trigger simulation and optimizing of reconstruction algorithms

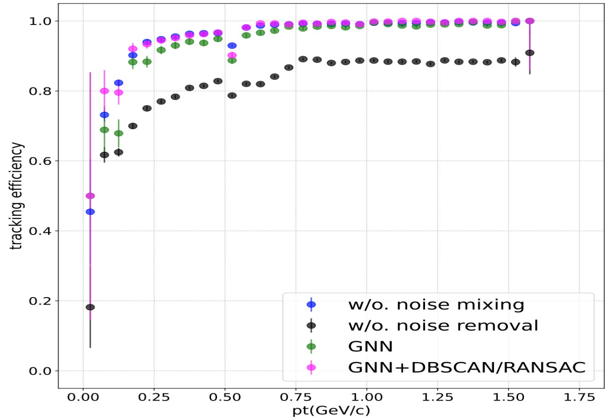




Strategy for event mixing



Trigger simulation



Tracking via GNN noise removal

Digitization

More details in Binbin Qi' talk (Nov. 19):
STCF simulation and digitization

- Developed an unified digitization framework for all sub-detectors on OSCAR
- Each sub-detector implemented its digi. algorithms and produced the digi. information
- The reconstruction algorithms are under tuning with the digi. information as their inputs

Gaseous detectors:

OSCAR raw hit data

- Ionizing
- Electron drift & diffusion
- Avalanche
- Induced signal
- Electronics influence, transmission & noise

OSCAR digi. data

Scintillator detectors:

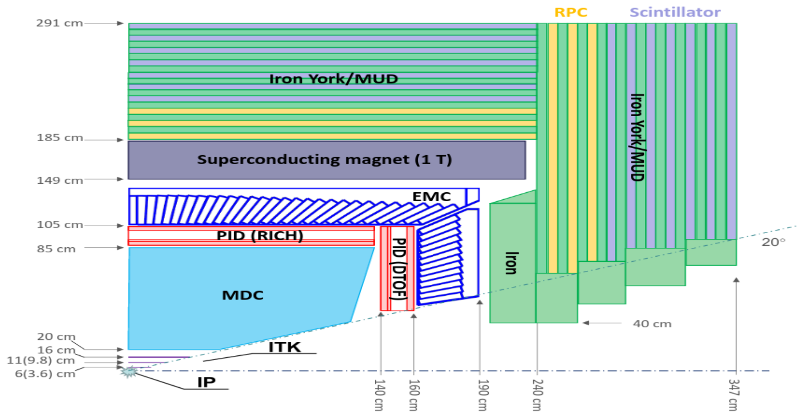
OSCAR raw hit data

- Optical-photon emission and transmission
- Electronics effect
- Waveform generate & fitting
- Peak finding & T Q reconstruction

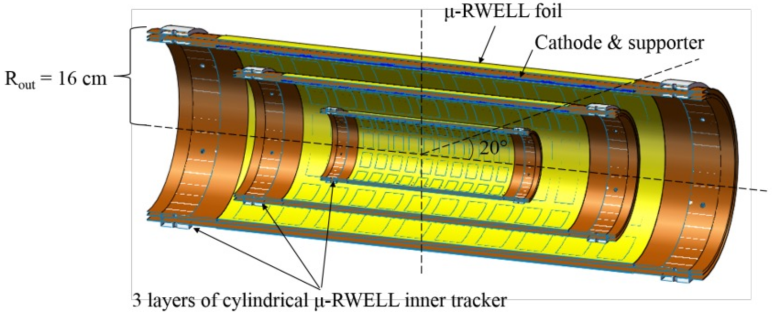
OSCAR digi. data

STCF tracking system

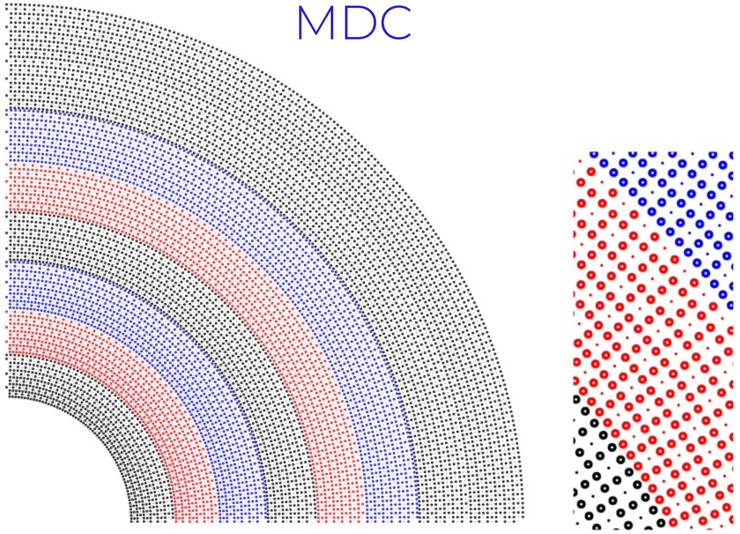
- The baseline tracking system includes uRWELL-based Inner Tracker (ITK) and Main Drift Chamber (MDC)
  - ITK: 3 layers, $\sigma_{r\text{-}\varphi} \times \sigma_z \approx 100$ um x 400 um
  - MDC: 48 layers, $\sigma_{\text{drift dist}} \approx 120\text{~}130$ um



uRWELL-based ITK



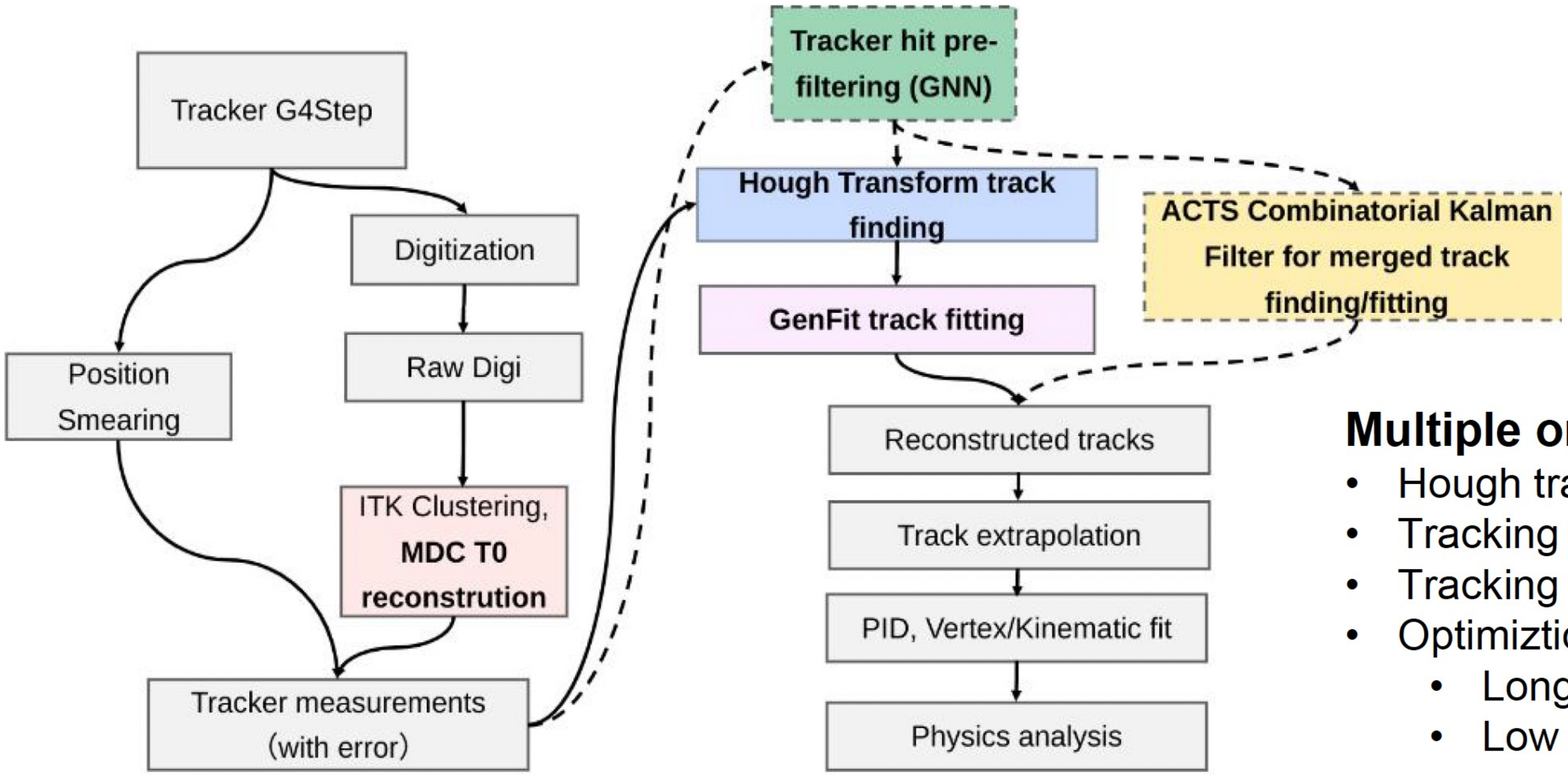
MDC

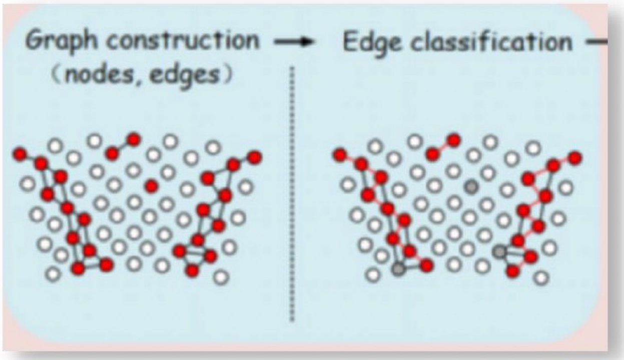


(a) (b)

Figures from STCF CDR (arXiv: 2303.15790)

STCF tracking landscape



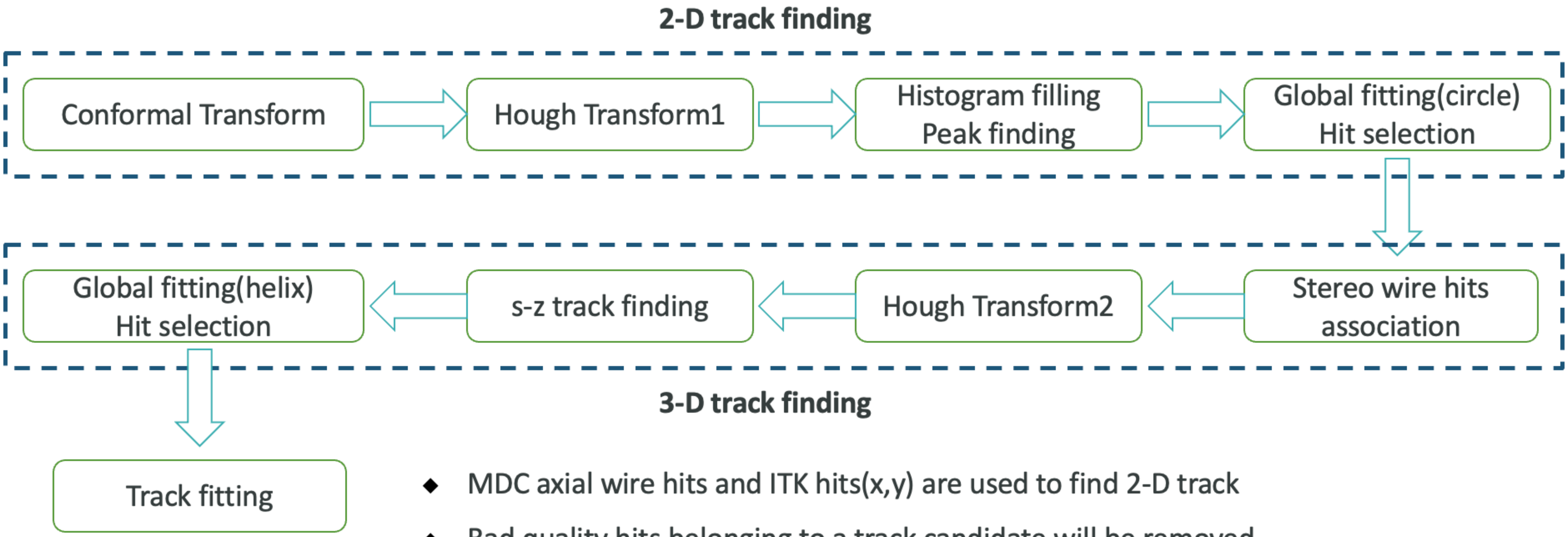


Multiple ongoing studies

- Hough transform as baseline
- Tracking with ACTS
- Tracking with GNN
- Optimiztions
  - Long-lived particles
  - Low momentum particles
  - ...

Tracking based on Hough Transform

More details in Jin Zhang' talk (Nov. 19):
STCF tracking with Hough+GenFi

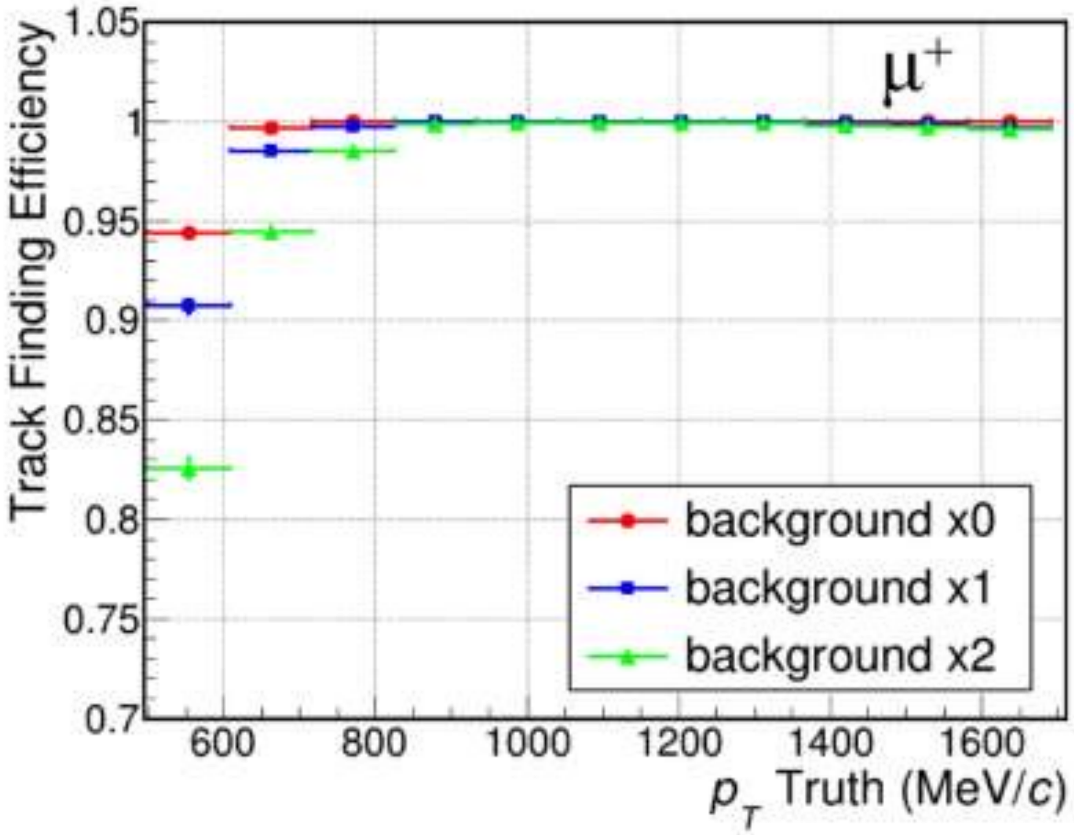


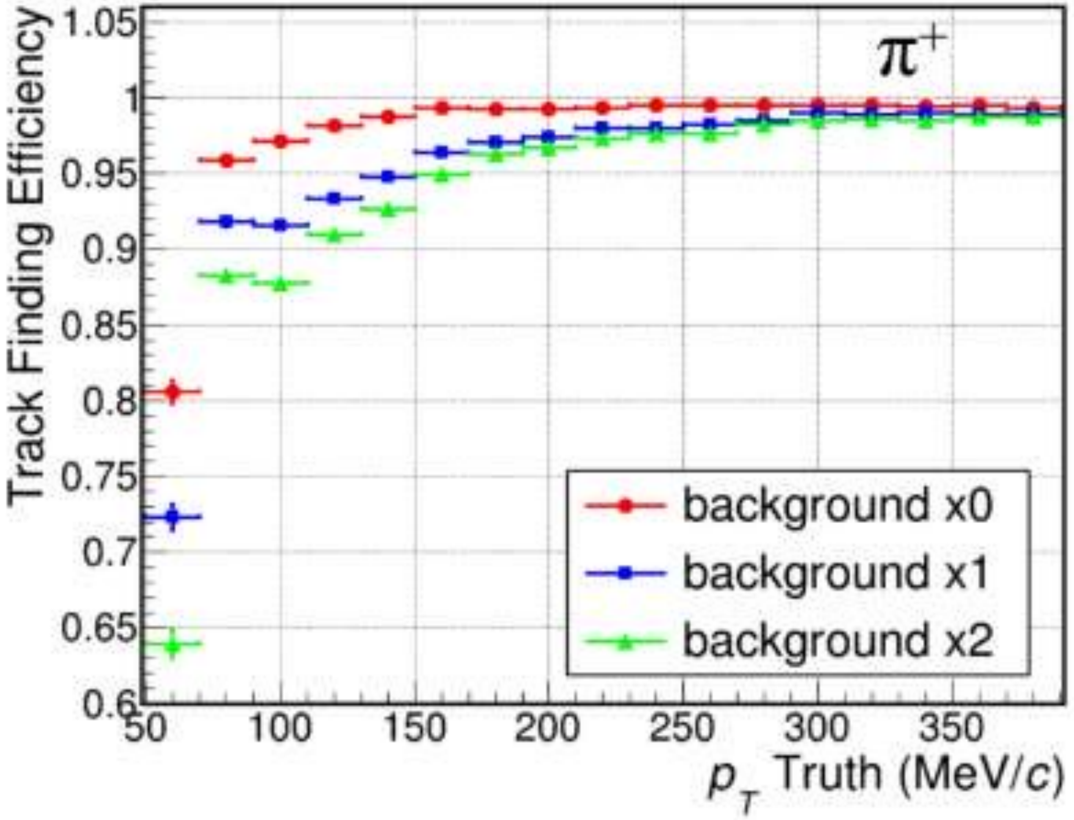
- MDC axial wire hits and ITK hits(x,y) are used to find 2-D track
- Bad quality hits belonging to a track candidate will be removed
- The trajectory is straight in the s-z space → similar to the 2-D tarck finding

Tracking performance of Hough Transform

More details in Jin Zhang' talk (Nov. 19): STCF tracking with Hough+GenFi

- $\mu+$ track finding efficiency is above 95%/90% without/with 1x backgrounds at 500 MeV/c
- $\pi+$ track finding efficiency is above 95%/90% without/with 1x backgrounds at 100 MeV/c

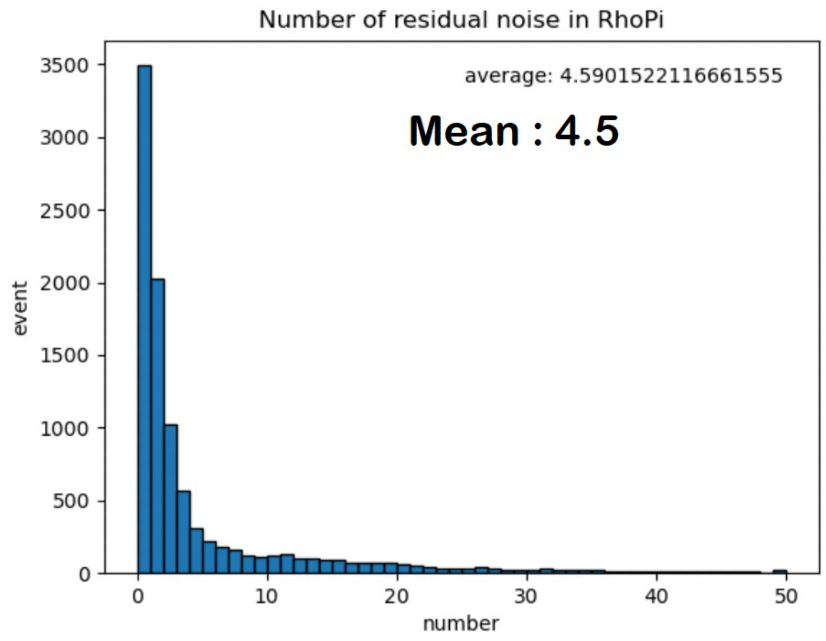




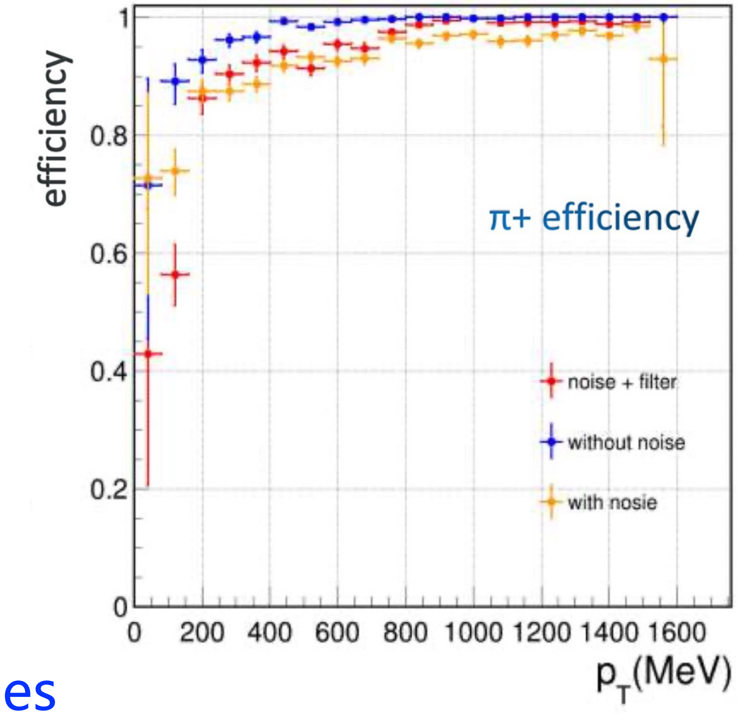
Tracking performance of Hough Transform

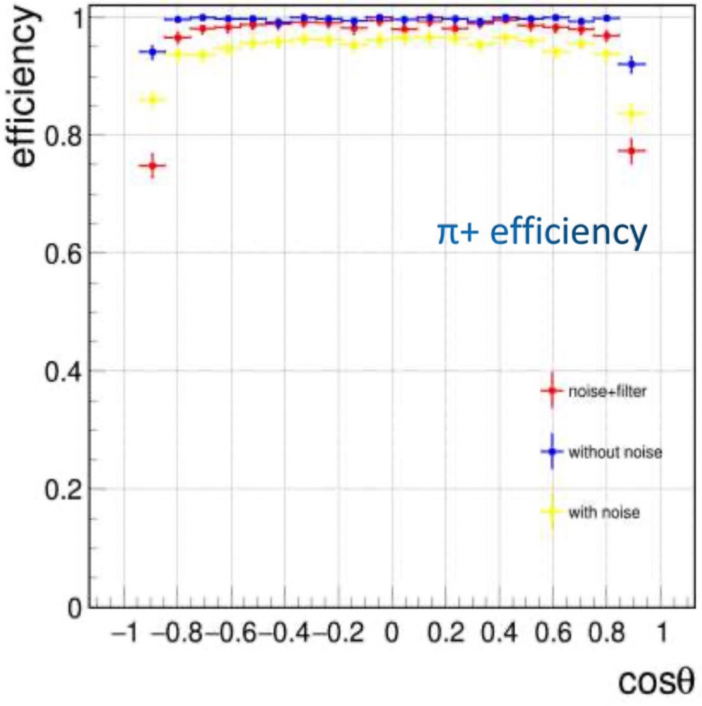
More details in Xiaoqian Jia' talk (Nov. 19): GNN for BESIII/STCF MDC tracking

- π+ track finding efficiency is above 95%/90% without/with backgrounds
- 99% noise hits can be removed by GNN
- The efficiency after GNN filtering noise is significantly improved



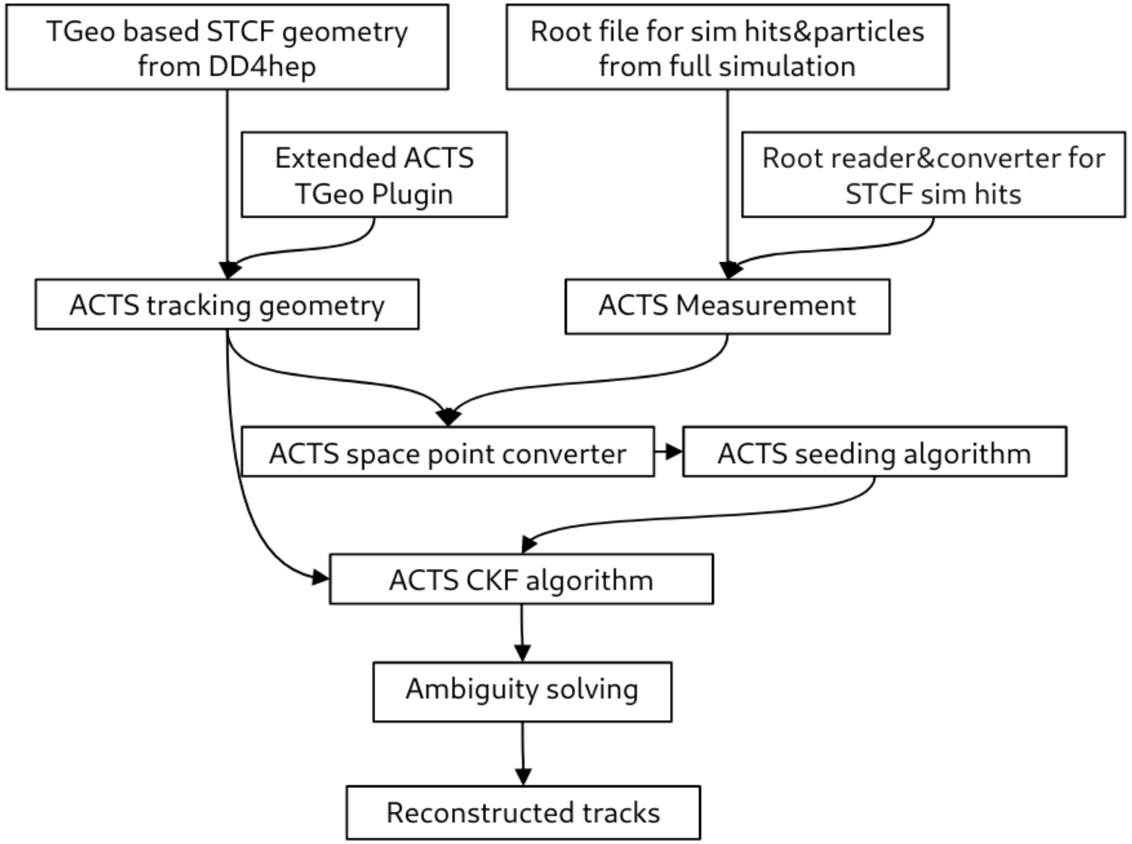
Mixed with 600 random trigger noises





Tracking based on ACTS

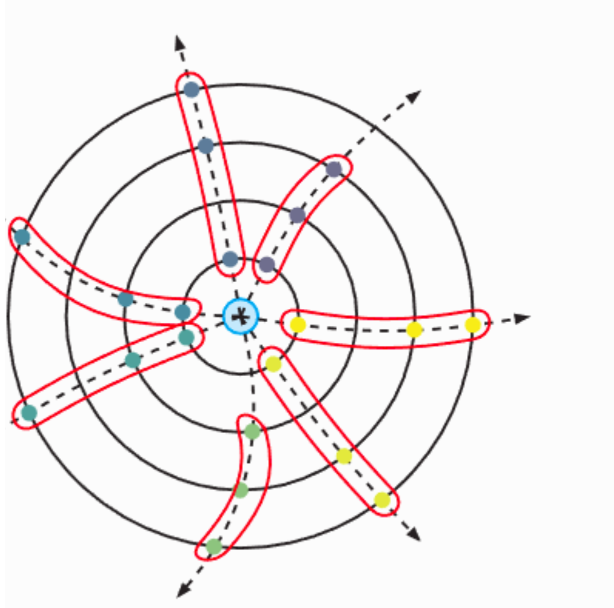
Application strategies of ACTS for STCF

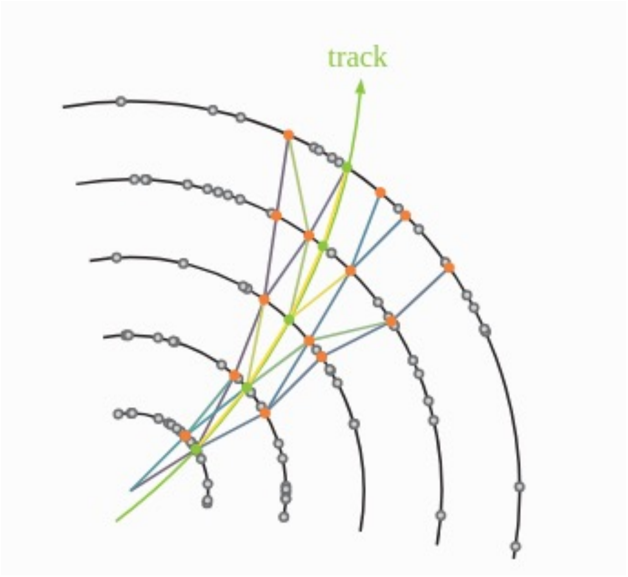


Seeding
(find seeds using hits on ITk layers)

→

Combinatorial Kalman Filter (CKF)
(simultaneous track fitting + finding)



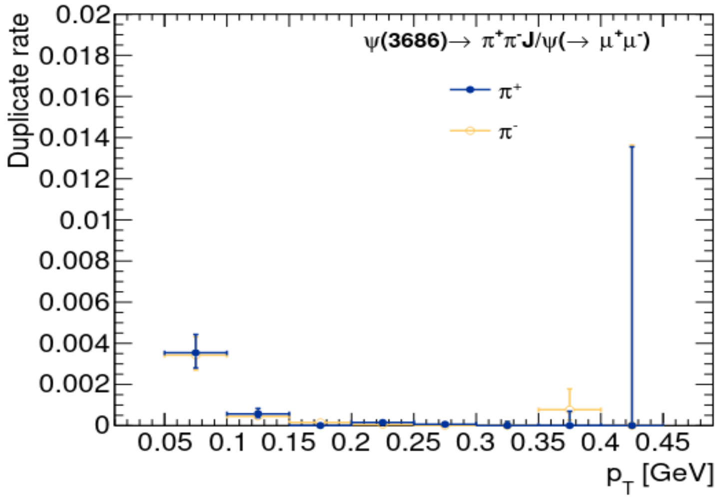


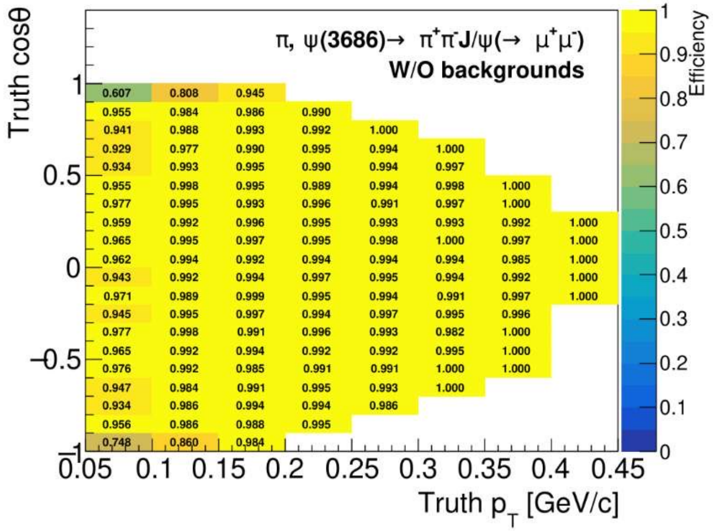
Figures from ACTS [readthedocs]

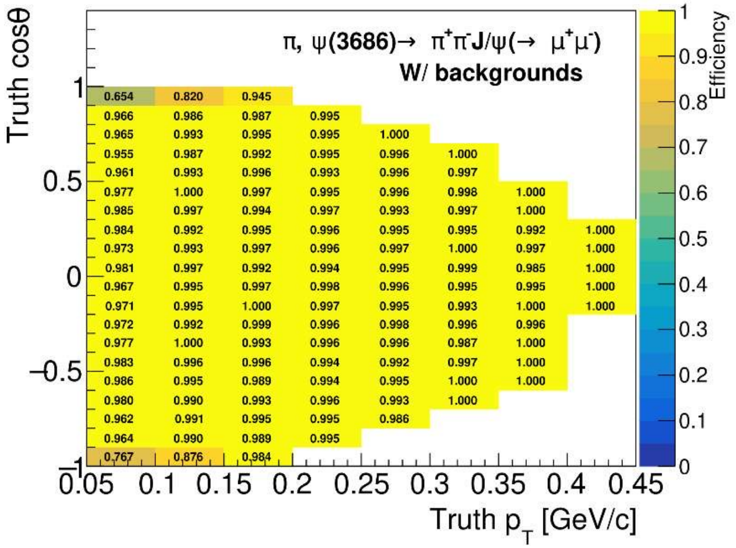
Tracking finding performance with ACTS

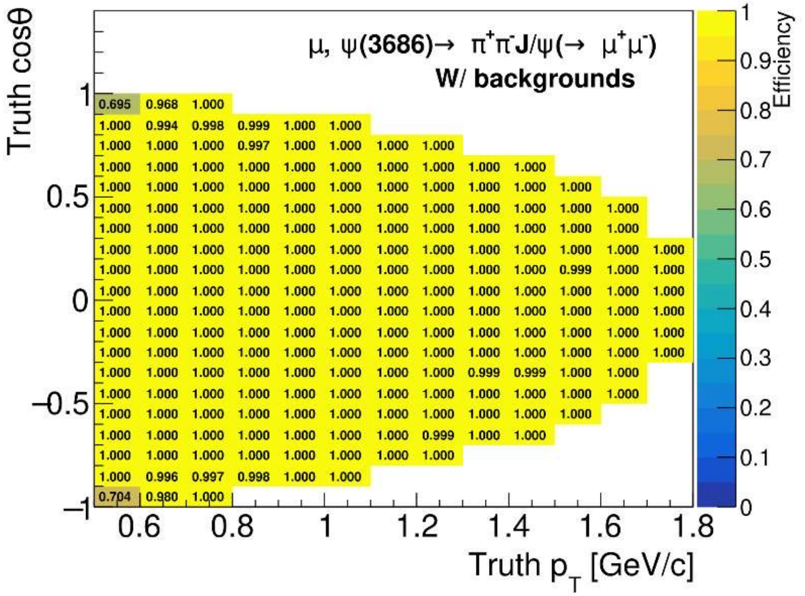
More details in Hao Li' talk (Nov. 19):
STCF tracking with ACTS

- No obvious efficiency loss with 1X backgrounds
- > 96% efficiency of tracks with pT in [50, 100] MeV/c and |costheta|<0.9
- Negligible fake tracks (<0.02%) with pT > 100 MeV/c

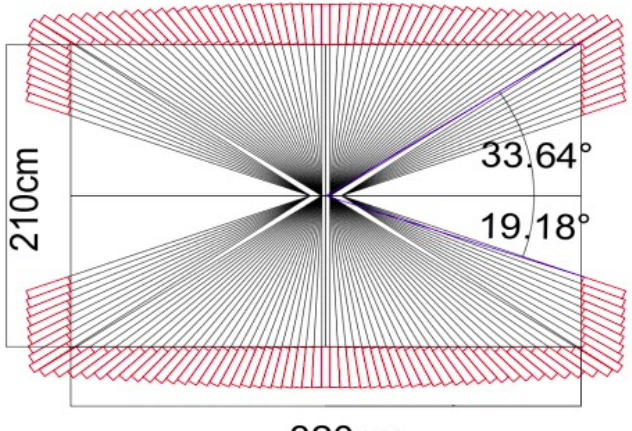






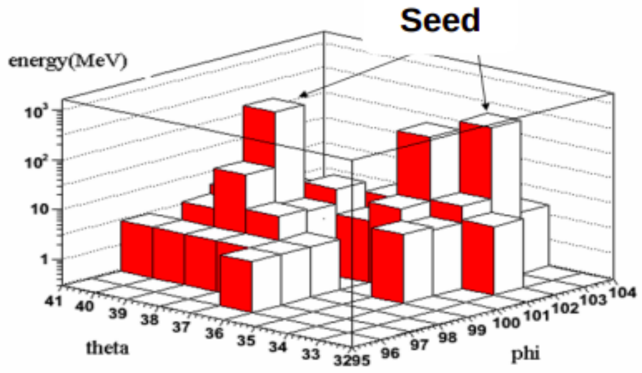


EMC Reconstruction method



Crystal arrangement diagram

Visualized by DD4Hep



π^0 Cluster (two photons)

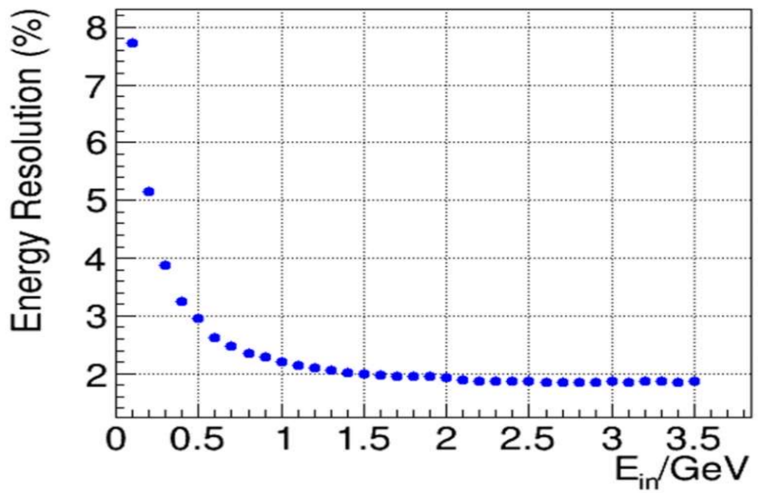
ECALRecAlg

- Cluster search
- Seed search
- Clusters split into Showers
- Calculation and correction of shower energy and position

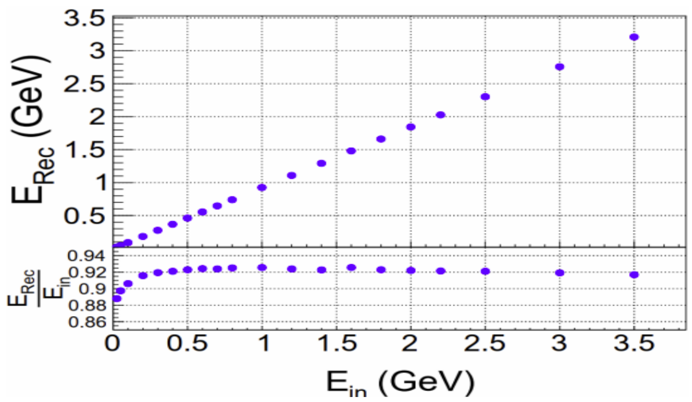
Performance of EMC reconstruction

More details in Bo Wang' talk (Nov. 21): STCF MEC reconstruction

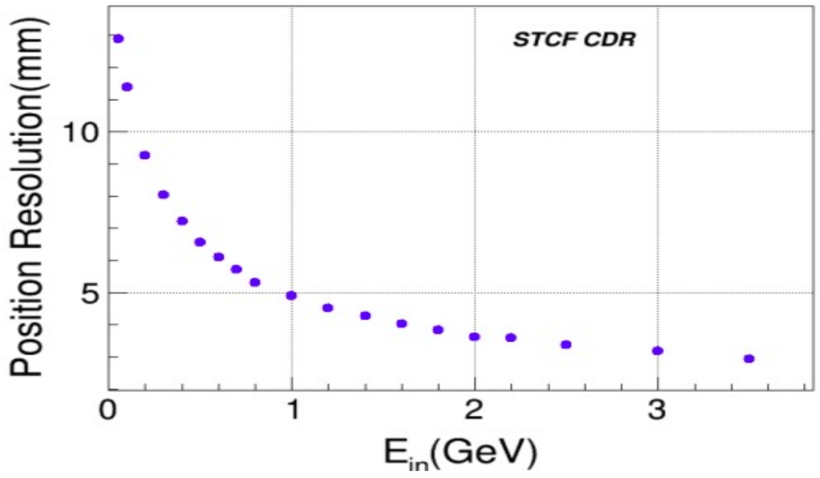
- Energy resolution is 2.2% at 1 GeV



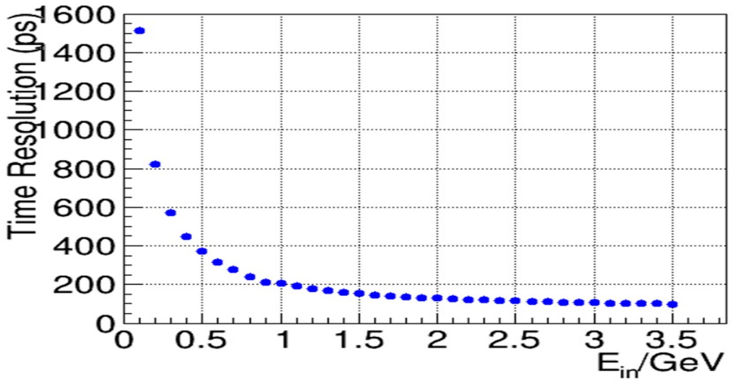
- Good energy linearity between 25 MeV-3.5 GeV



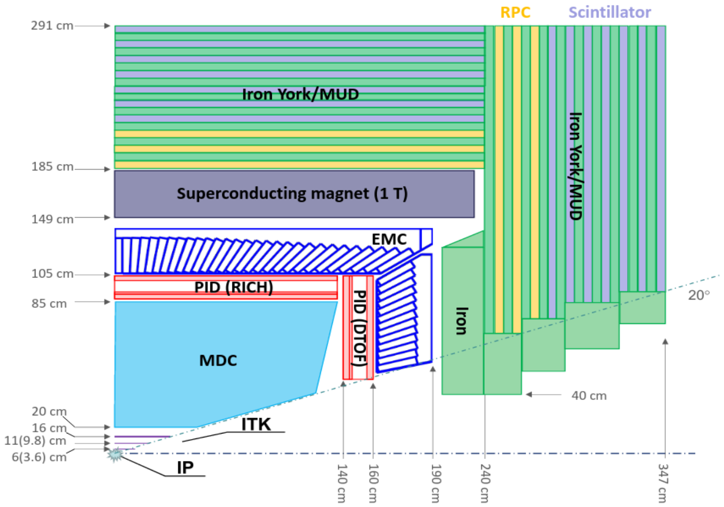
- Position resolution is 4.9 mm at 1 GeV

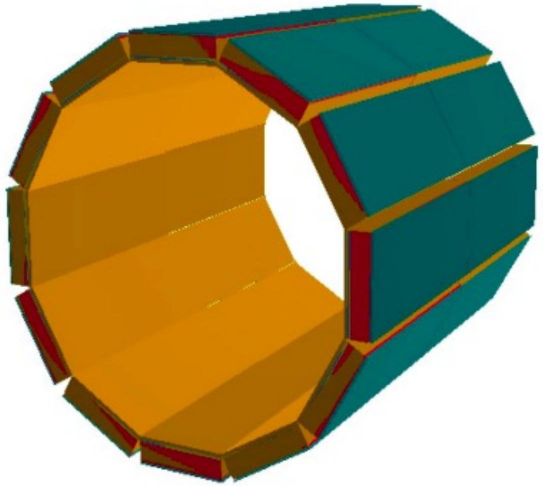


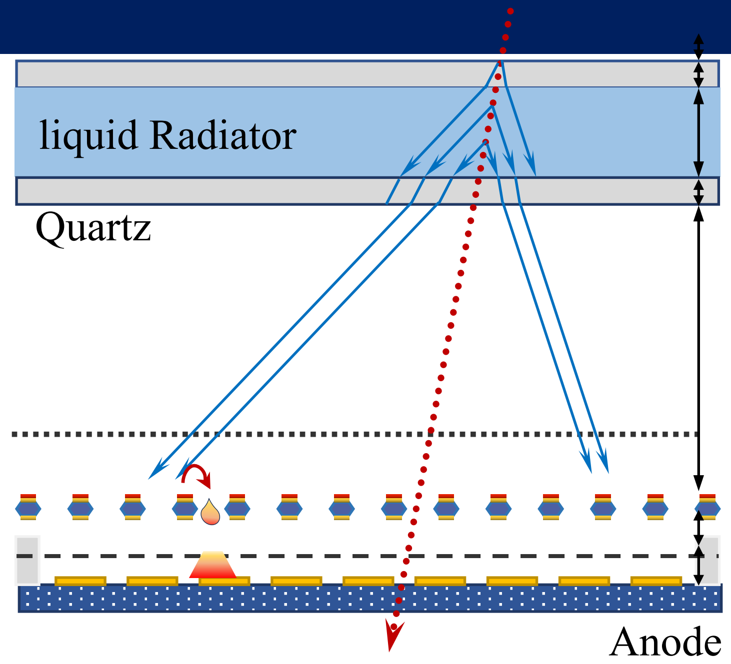
- Time resolution is 200 ps based on waveform fitting

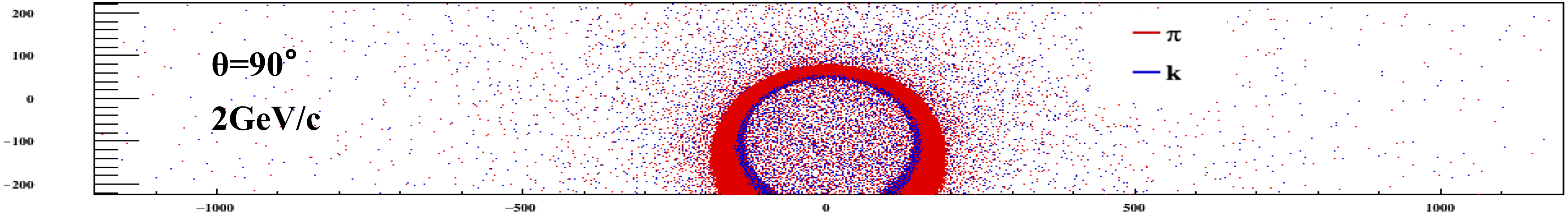


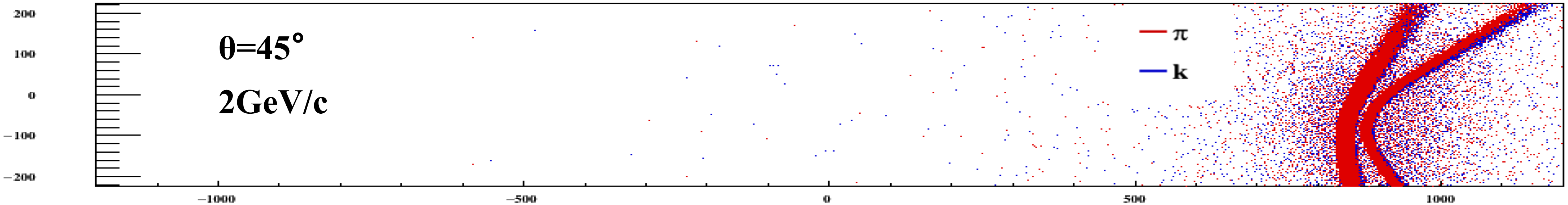
RICH detector











RICH PID performance with likelihood method

More details in Qingyuan' talk (Nov. 20):STCF RICH recon.

- The photon collected in each anode pads follows the **Poisson distribution**

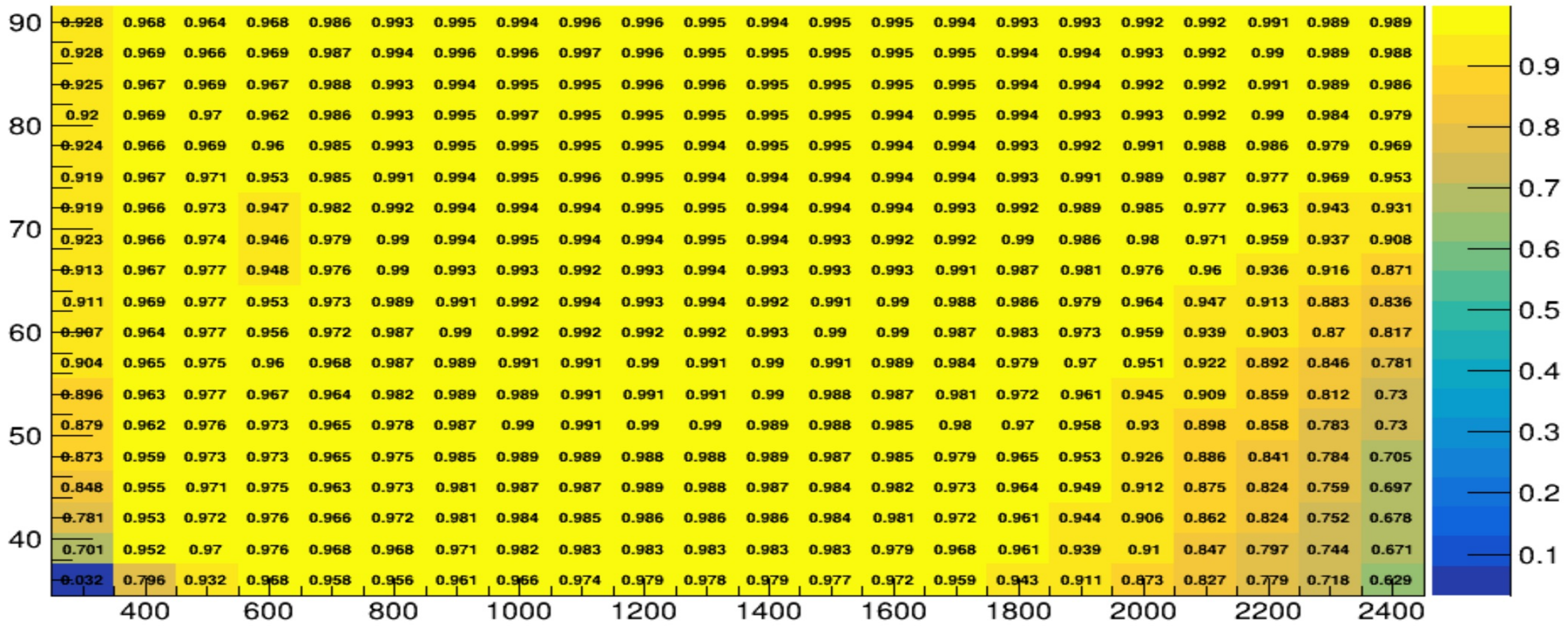
$$pdf_{i,h} = Poisson\,(\underbrace{N_i}_{\text{num of }\gamma\text{ in anode pad}} + \underbrace{10^{-3}}_{10^{-3}\text{ bkg level}}, \underbrace{avg_{i,h}}_{\text{avg of }\gamma\text{ in simulation}} + 10^{-3}),$$

- Likelihood of h hypothesis:

$$\ln \mathcal{L}_h = \Sigma_i^{npads} \ln pdf_{i,h}$$

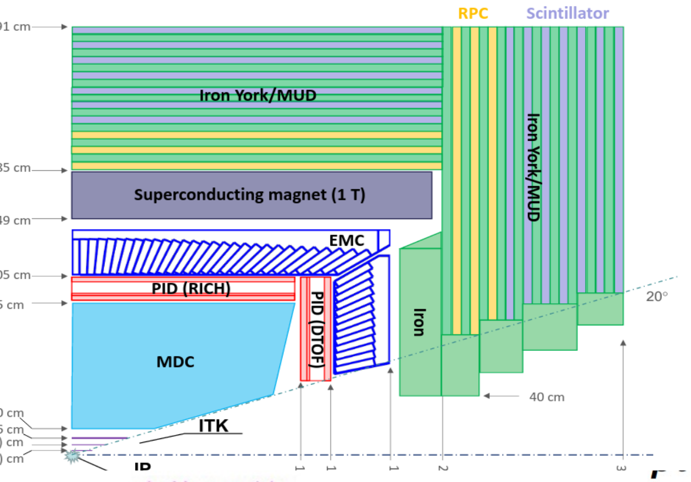
- π, K separation:

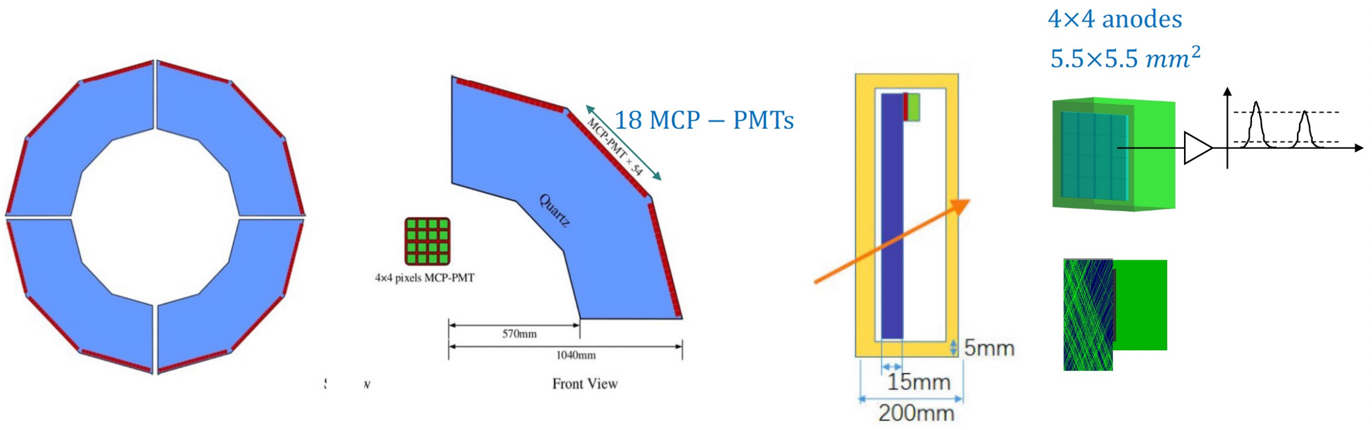
$$DLL = \Sigma_i^{npads} \ln \frac{pdf_{i,\pi}}{pdf_{i,K}}$$

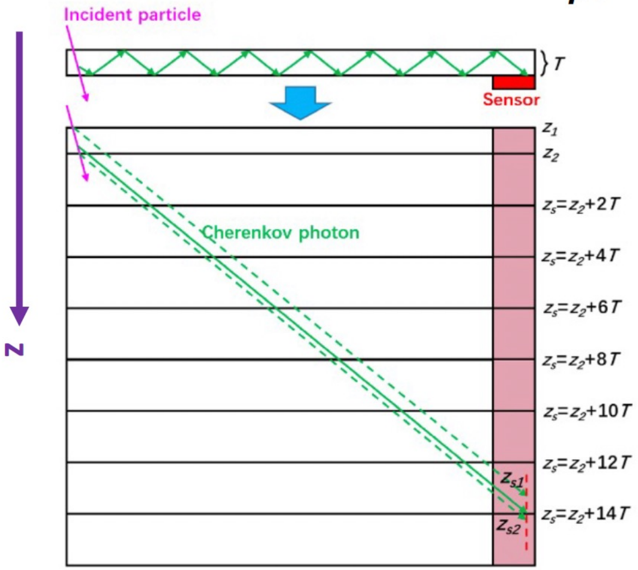


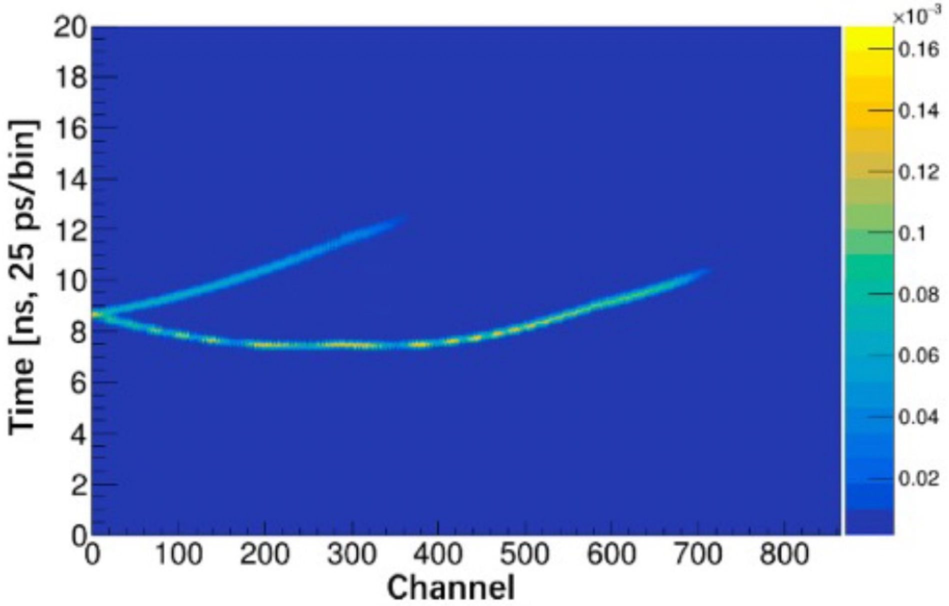
- Pion PID efficiency > 97% in most bins with Kaon mis-identification at 2%

DTOF detector









The simulated TOA versus hit position pattern

DTOF PID performance with Timing method

More details in Wenhui Tian' talk (Nov. 20):STCF DTOF recon.

TOF reconstruction

$$cos\theta_c = \frac{1}{n\beta} = \frac{\vec{v}_t \cdot \vec{v}_p}{|\vec{v}_t| \cdot |\vec{v}_p|} \qquad \begin{cases} \vec{v}_t = (a, b, c) \\ \vec{v}_p = (\Delta X, \Delta Y, \Delta Z) \end{cases}$$

$$LOP = \sqrt{\Delta X^2 + \Delta Y^2 + \Delta Z^2} \Longrightarrow TOF_{rec} = T - \frac{LOP \cdot \overline{n_g}}{c} - T_0$$

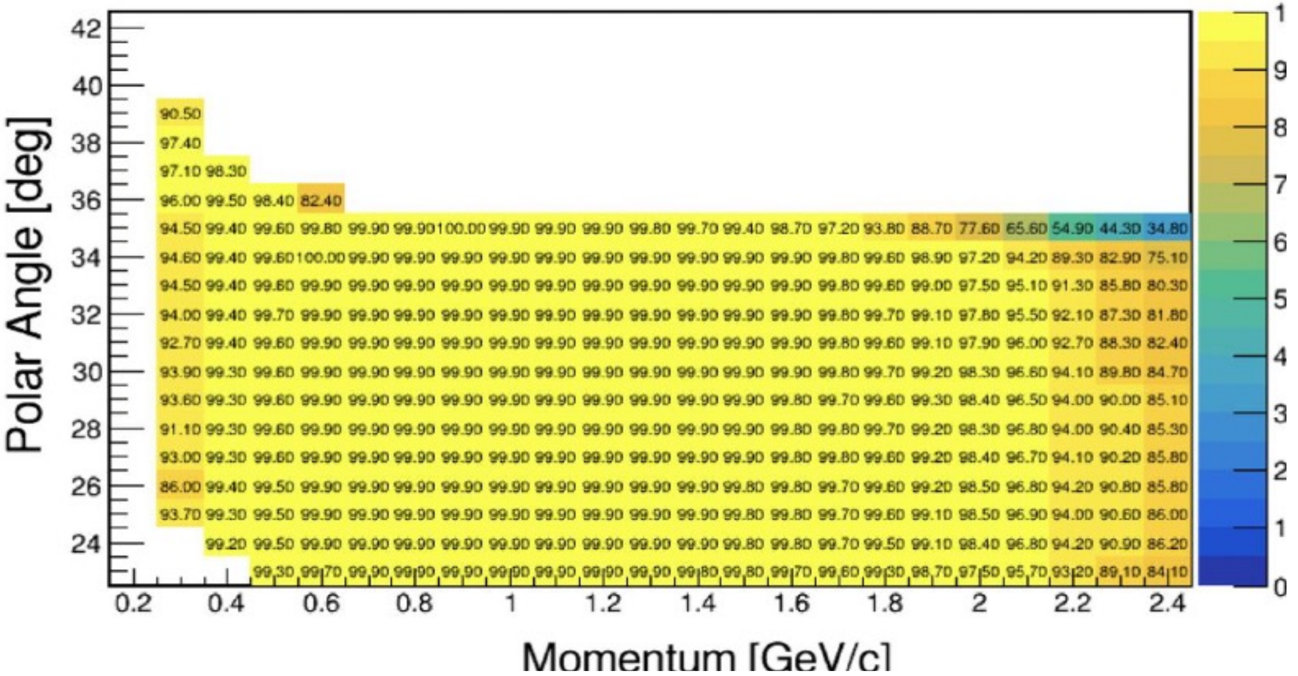
Likelihood construction

$$\mathcal{L}_h = \prod_{i=0}^{N_{p.e.}} f_h(TOF_i)$$

$$f_h(t) = \underbrace{Gaus(TOF_{rec} | TOF_{hypo}, \sigma)}_{\text{signal}} + \underbrace{0.05}_{\text{bkg}}$$

where $TOF_{hypo} = \frac{LOF}{c\beta_{hypo}}$

π efficiency after mixing BKG in different $(|\vec{p}|, \theta)$:
$(K \text{ mis} - \text{ID} = 2\%)$



- Pion PID efficiency > 97% for P =< 2 GeV/c

DTOF PID performance with Imaging method

More details in Wenhui Tian' talk (Nov. 20): STCF DTOF reconstruction

Photon TOA v.s. (x_s, y_s) Reconstruction

$$\cos\theta_c = \frac{1}{n\beta} = \frac{\vec{v}_t \cdot \vec{v}_p}{|\vec{v}_t| \cdot |\vec{v}_p|} \qquad \begin{cases} \vec{v}_t = (a, b, c) \\ \vec{v}_p = (x_s - x_0, y_s - y_0, z_s - z_0) \end{cases}$$

$$z_s = z_2 + 2mT$$

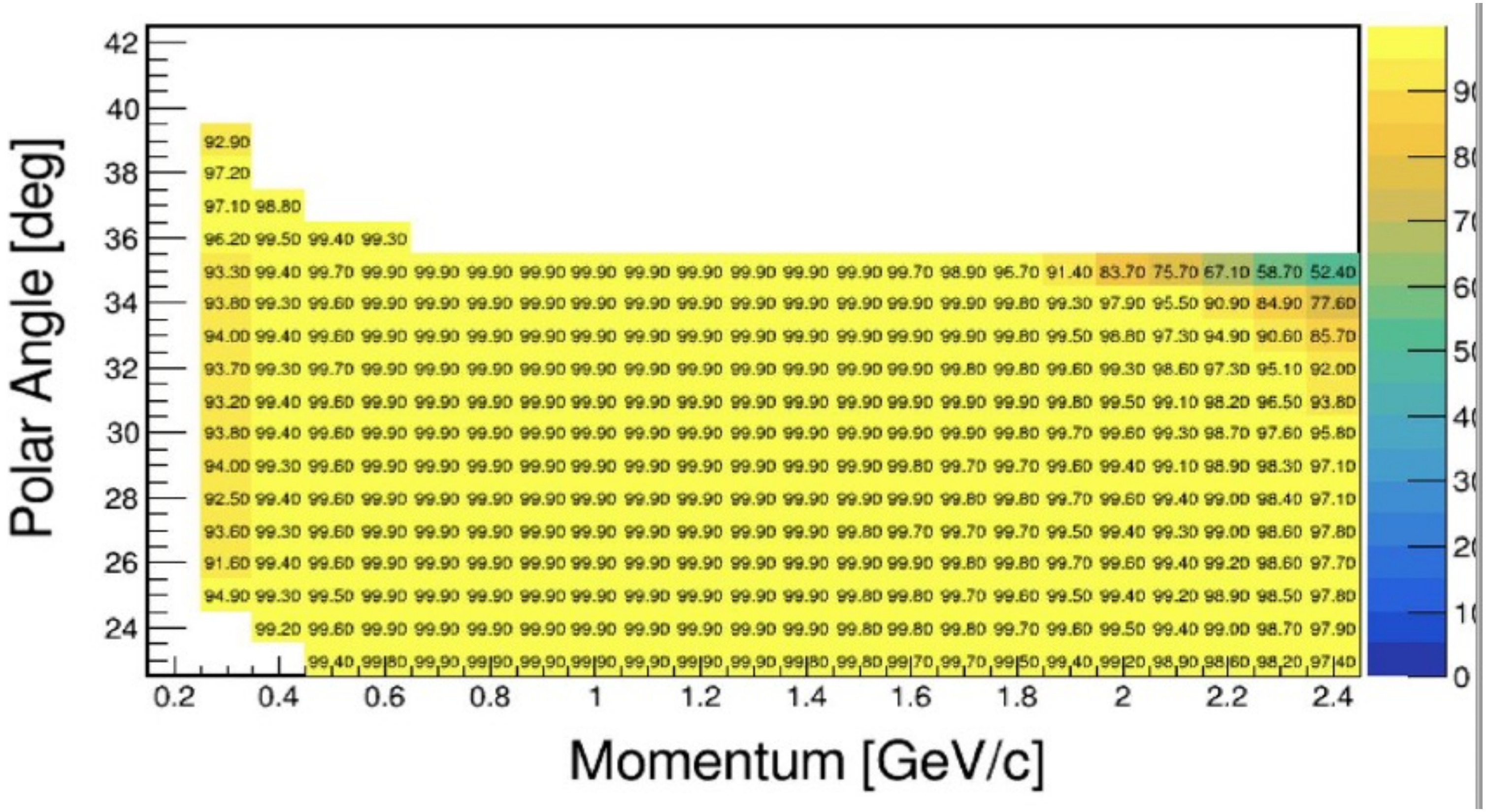
$$(x_s, y_s) \Rightarrow z_e, \phi_c \Rightarrow TOA = TOF + \frac{\Delta LOF_e}{\beta c} + TOP$$

Likelihood construction

$$\mathcal{L}_h = p_h(N_{p.e.}) \prod_{i=0}^{N_{p.e.}} f_h(x_i, t_i)$$

$$p_h(N_{p.e.}) = \sum_{n=0}^{N} Poisson_h(n, N_e) \times F_{bkg}(N-n)$$

$$f_h(x,t) = S_h(x,t) + const_{bkg}$$

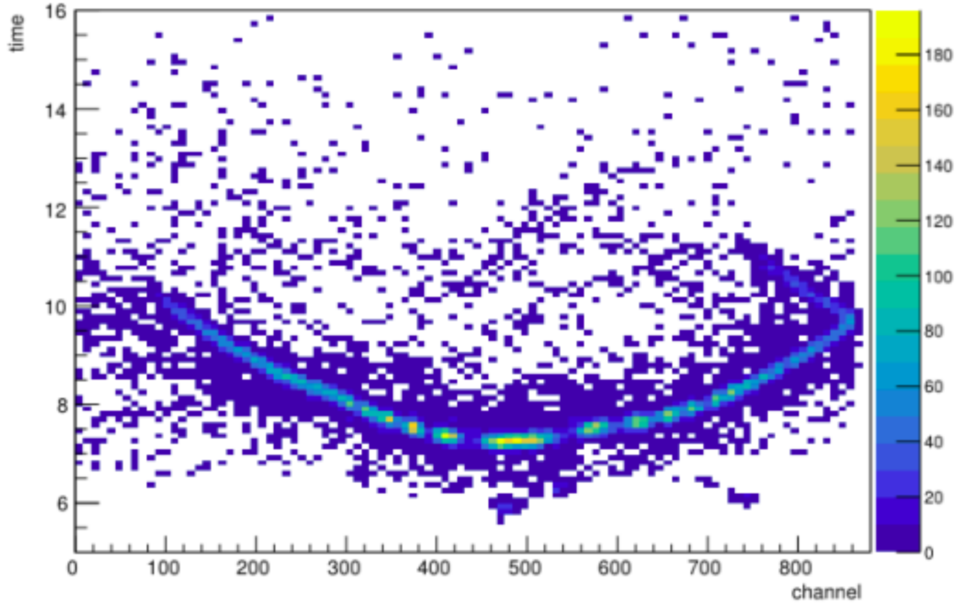


- Pion PID efficiency > 97% in most bins with Kaon mis-identification at 2%

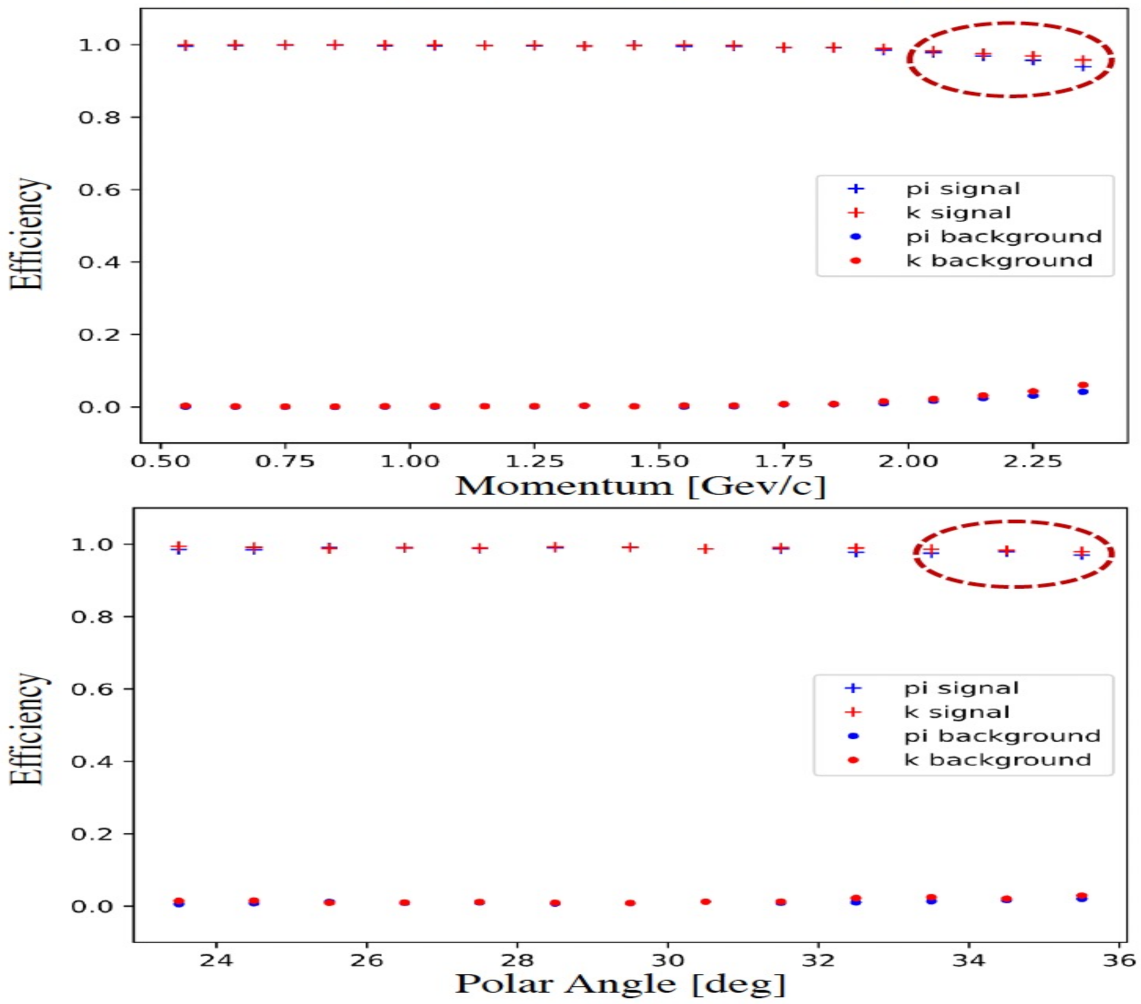
DTOF PID performance with CNN method

More details in Zhipeng Yao' talk (Nov. 20):
STCF RICH/DTOF reconstruction with CNN

- A CNN (based on EfficientNetV2S) which utilizes both timing and spatial information of the DTOF hits is developed
- Efficiency: 99.46% at 2 GeV



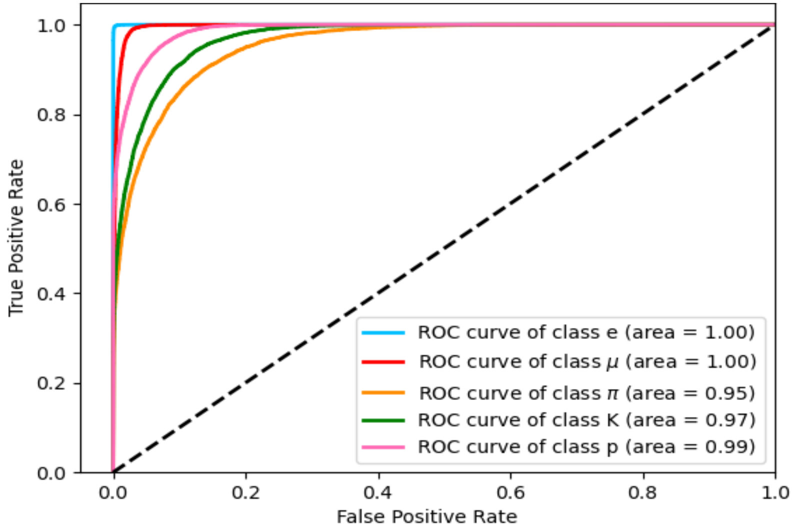
X-label: the hit position of Cherenkov photon by PMT
Y-label: the arrival time of Cherenkov photon by PMT
Value: the number of photons within this bin

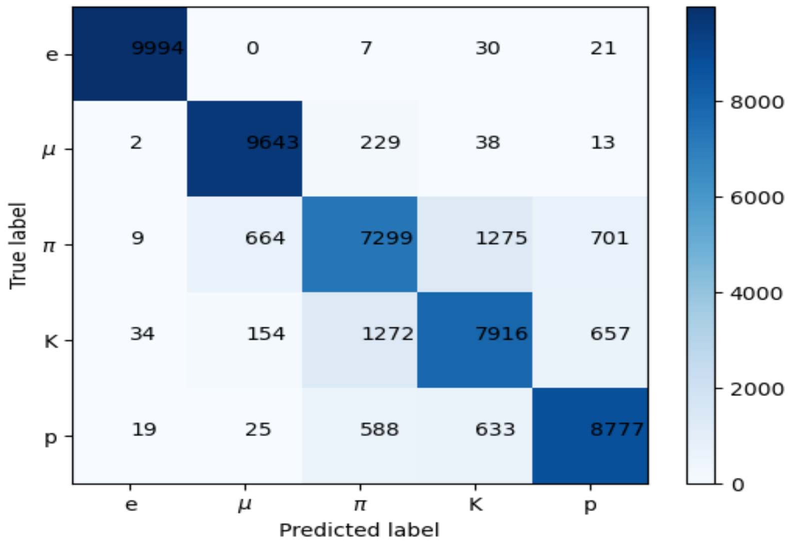


Global PID method

More details in Yuncong Zhai' talk (Nov. 21): STCF global PID with ML

- A global PID algorithm is developed based on ML and selected 45 features from all sub-detector: Tracker/dEdx/RICH/DTOF/ECAL/MUD
- Trained the ML model based on XGBoost for better separation power in different modes: e/μ/π/K/P; π/K/p; e/π/K; π/K; μ/π
- A global neutral particle identifier based on CNN is developed for discrimination of γ/KL/n
- The model and algorithm are provided for physics analysis within OSCAR



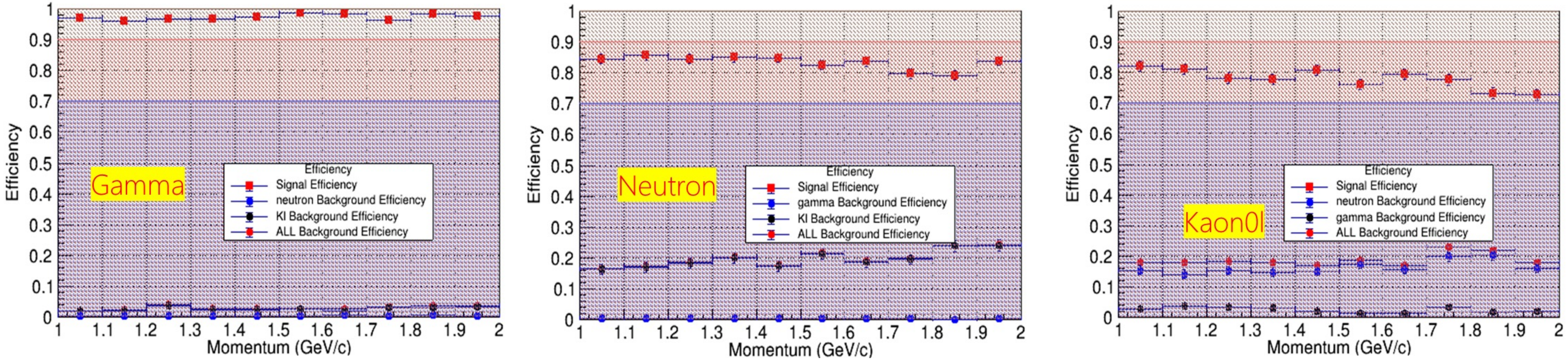


True \ Predicted	e	μ	π	K	p
e	9994	0	7	30	21
μ	2	9643	229	38	13
π	9	664	7299	1275	701
K	34	154	1272	7916	657
p	19	25	588	633	8777

Global PID method

More details in Yuncong Zhai' talk (Nov. 21): STCF global PID with ML

- A global PID algorithm is developed based on ML and selected 45 features from all sub-detector: Tracker/dEdx/RICH/DTOF/ECAL/MUD
- Trained the ML model based on XGBoost for better separation power in different modes: e/μ/π/K/P; π/K/p; e/π/K; π/K; μ/π
- A global neutral particle identifier based on CNN is developed for discrimination of γ/KL/n
- The model and algorithm are provided for physics analysis within OSCAR

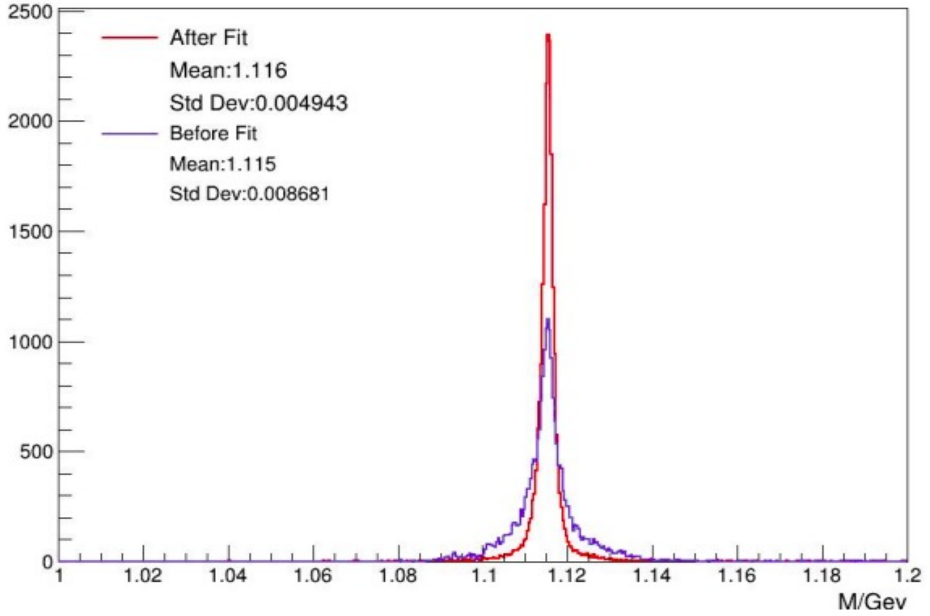


Analysis Tools: Vertex fitting

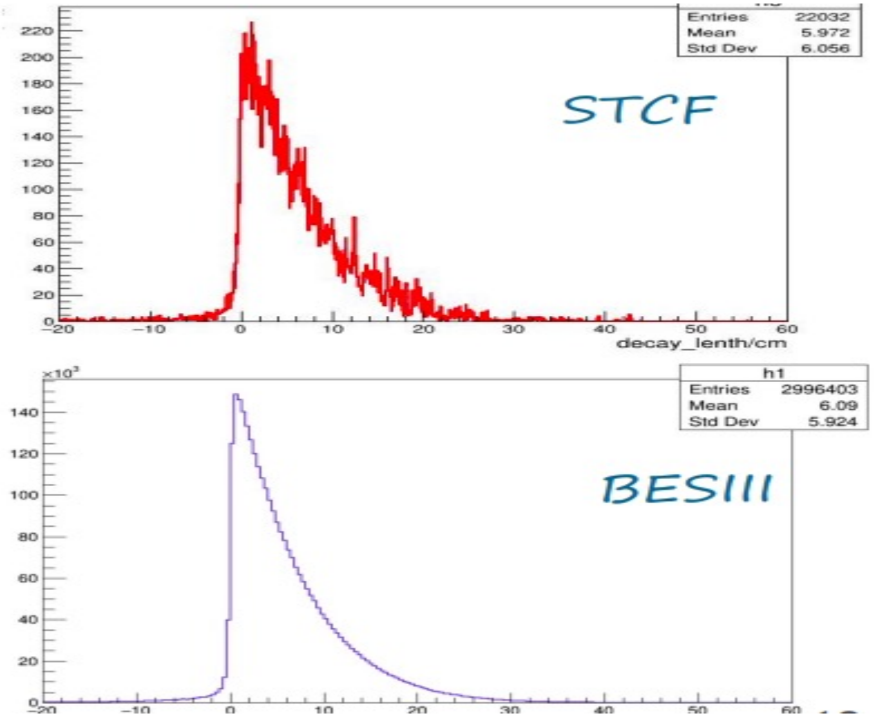
More details in Teng Li' talk (Nov. 21):
STCF core software and analysis tools

- The Vertex fitting and Kinematic fitting tools are developed based on the ones of BESIII after some modifications according to STCF
  - Two methods are used :Lagrange multiplier and Kalman filter
- Both are validated with STCF M.C. data and provided for physics analysis

$$J/\psi \to \Lambda(\to p\pi^-)\bar{\Lambda}(\to \bar{p}\pi^+)$$



Invariant mass of Λ



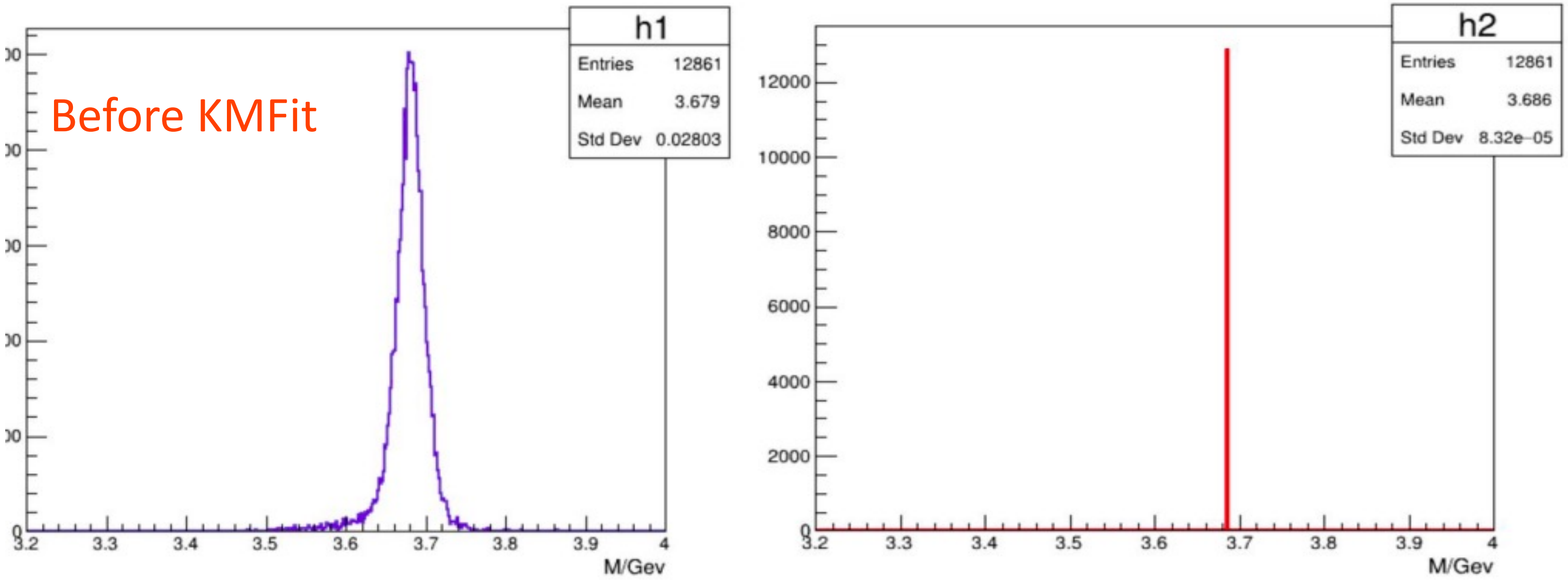
Decay length

Analysis Tools: Kinematic fitting

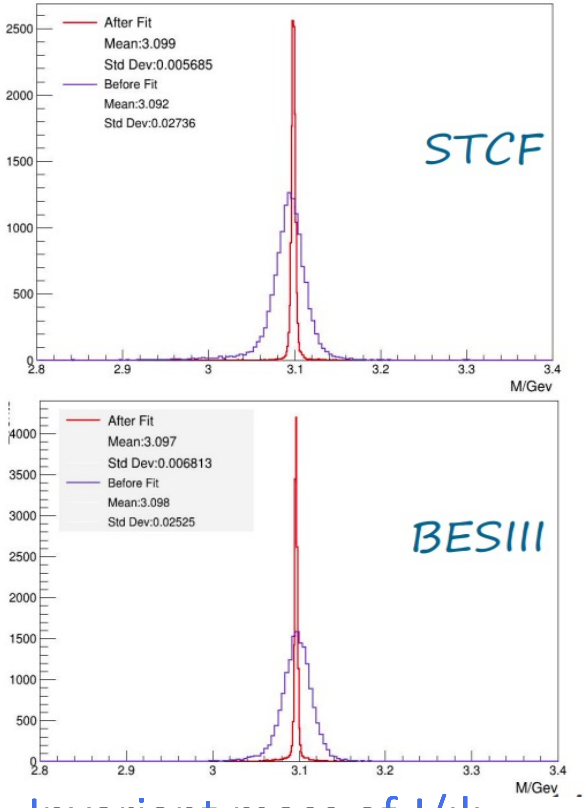
More details in Teng Li' talk (Nov. 21): STCF core software and analysis tools

- The Vertex fitting and Kinematic fitting tools are developed based on the ones of BESIII after some modifications according to STCF
  - Two methods are used :Lagrange multiplier and Kalman filter
- Both are validated with STCF M.C. data and provided for physics analysis

$$\psi(2S) \to \pi^+\pi^- J/\psi(\to \mu^+\mu^-)$$



Invariant mass of ψ(2S)



Invariant mass of J/ψ

Summary

- Offline software system plays important roles in the HEP experiments
- Gaudi based software systems developed to well support Physics analysis or Detector R&D
  - BOSS for BESIII and CEPCSW for CEPC
- SNiPER, a lightweight framework, has been developed by China HEP Team (SDU+IHEP)
  - JUNOSW developed and used for JUNO MC data challenge and detector commissioning
- The STCF offline software system (OSCAR) is being developed based on SNiPER since 2018, and now partially based on Key4hep (podio, dd4hep, edm4hep)
  - The core software is extended to meet requirements of STCF i.e. high luminosity
  - The full chain of detector simulation+ background mixing + digitization+ reconstruction +analysis is built
  - Software performance mostly reaches the CDR requirements and ready to start physics analysis
- Future plan
  - Optimization of recon. algorithms, especially with digitization information after event mixing
  - Improving computing performance to meet high luminosity requirements

Thanks for your attention !