STCF Workshop Software Sessions

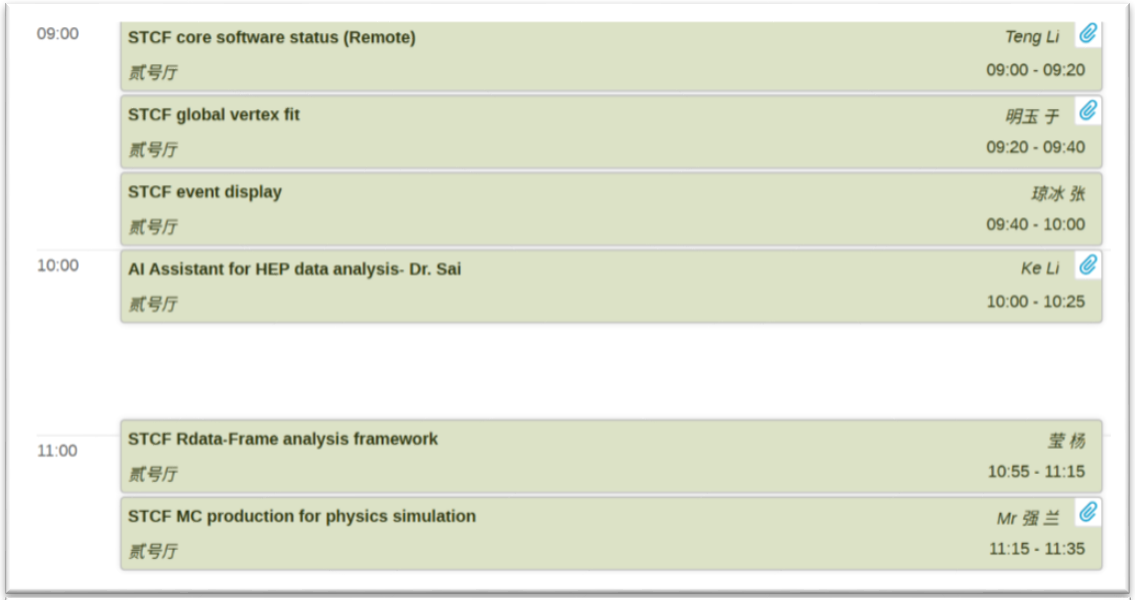
Take-aways and Highlights

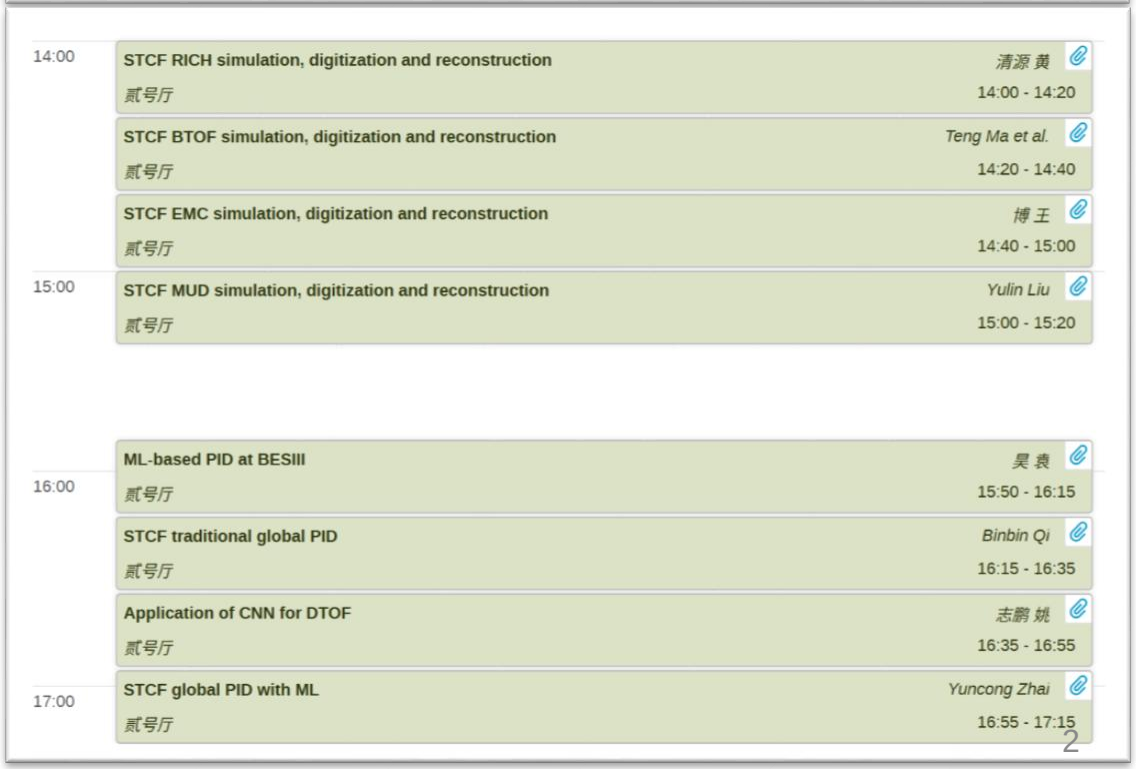
Xingtao Huang (SDU), **Xiaocong Ai** (ZZU)

2025 超级陶粲装置研讨会, 2025年7月5日，湘潭

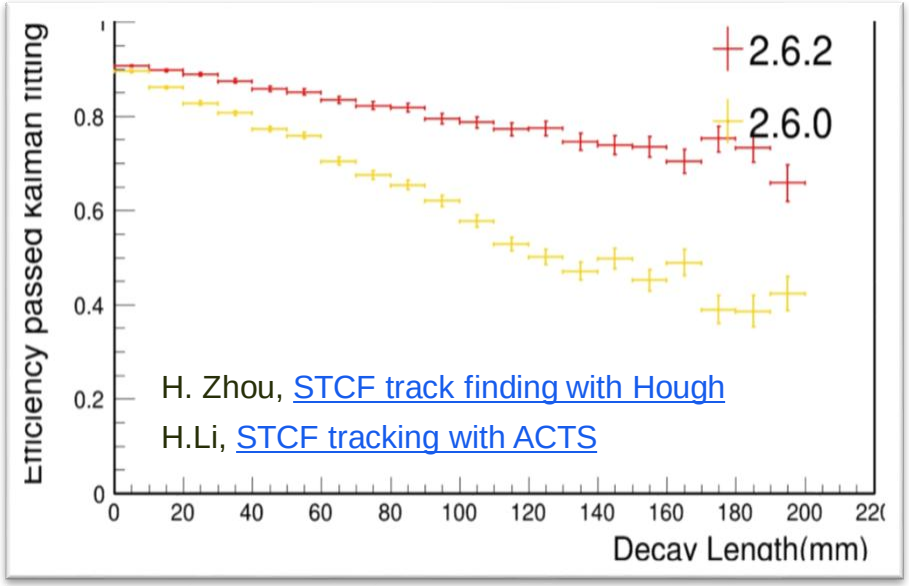
Statistics

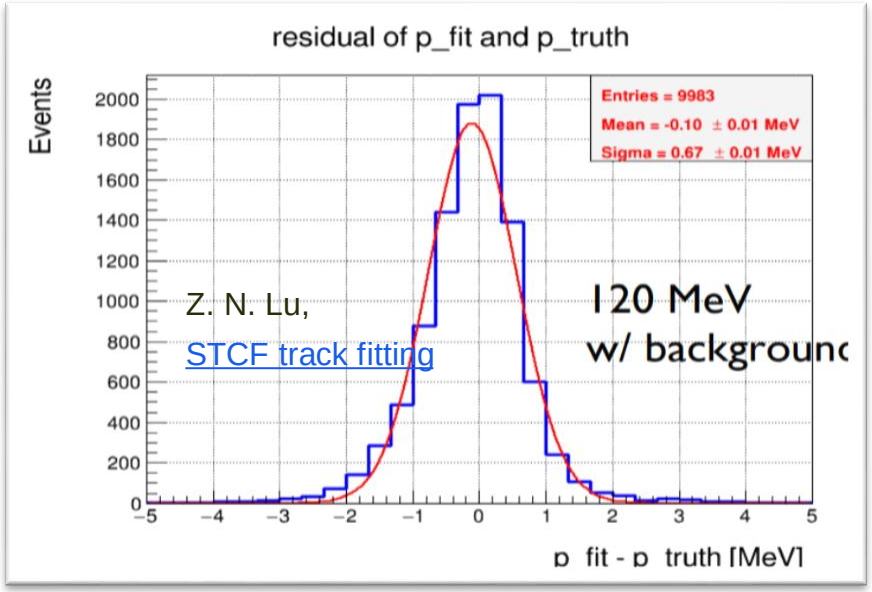
- **21** talks in software sessions
  - 17 about STCF, 3 about BESIII, 1 about LLM for analysis
  - Covering **tracking/photon** (5), **vertexing** (1), **PID** (8), **calibration** (3) , **analysis** (2), **software framework** (1), **event display** (1) , and **sample production** (1)



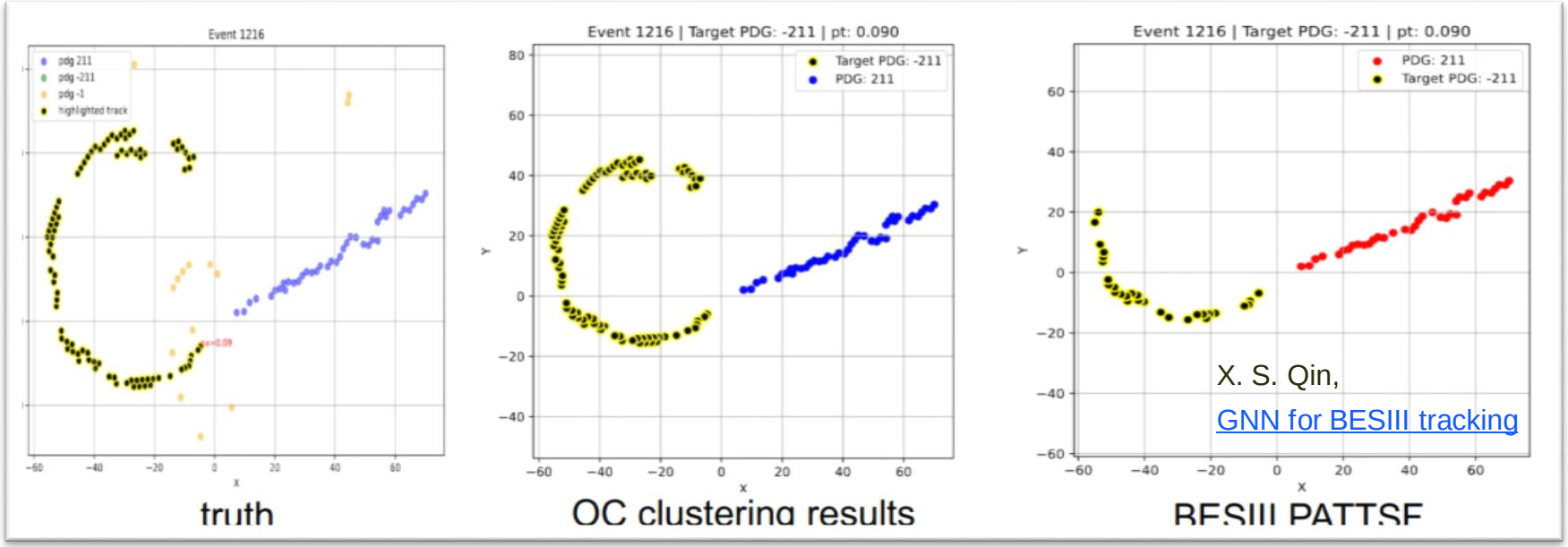


Tracking highlights





Decent **tracking eff.** for **long-lived particles** and **track parameters** for even **low p_T curling** tracks at STCF

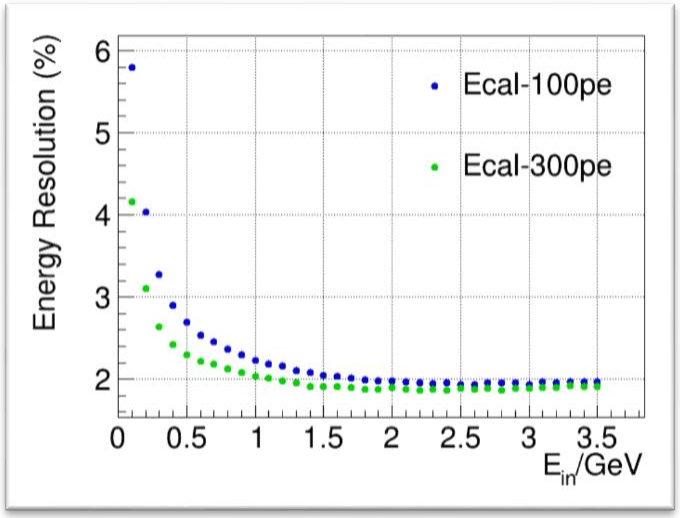


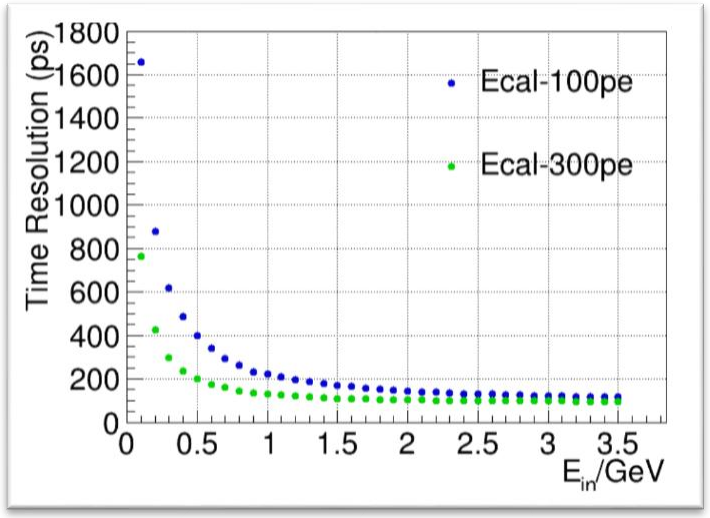
GNN-based **Object Condensation (OC)** explored, showing good **track finding performance** for **curling tracks** at BESIII

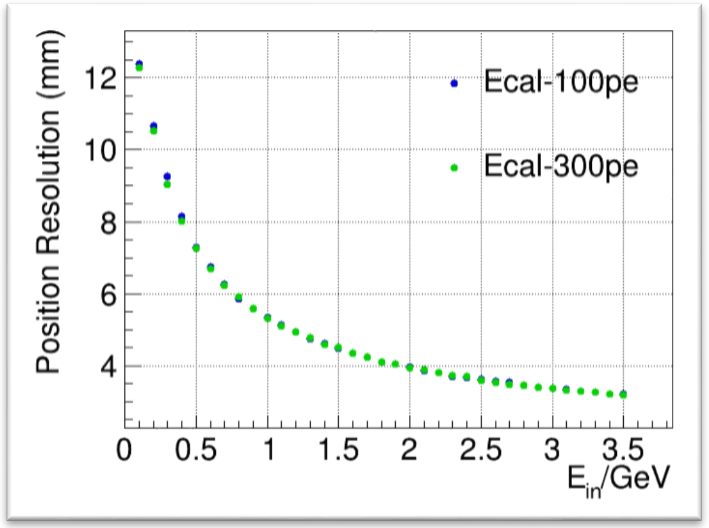
Photon/π^0 performance highlights

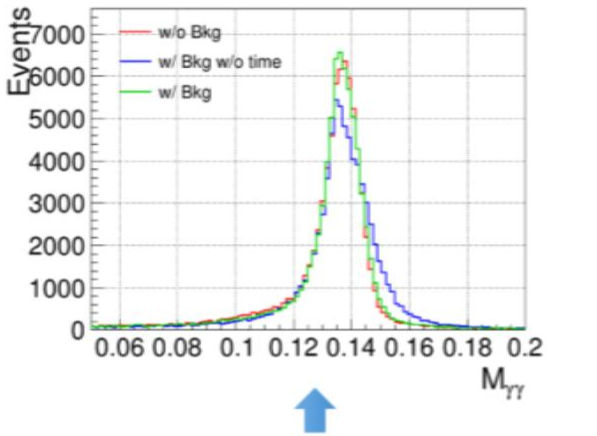
B. Wang,

STCF EMC simulation, digitization and reconstruction



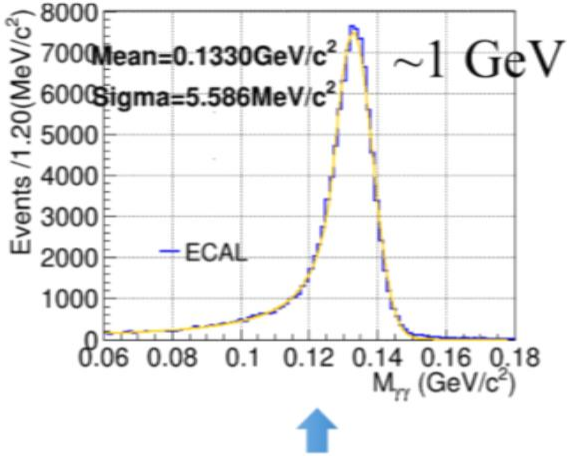




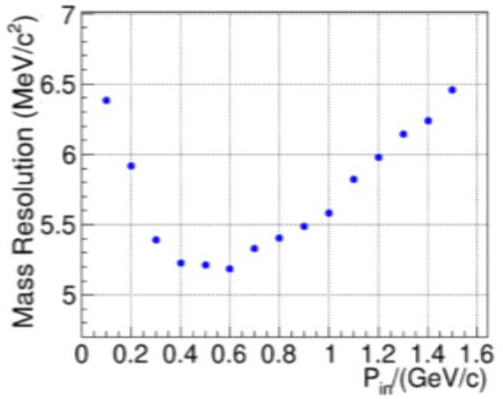


$-10ns < T_{seed} < 10ns$

π^0 with the minimum ΔM



Truth matched π^0



- ✓ $J/\psi \to \rho\pi$ sample
- ✓ $\sigma_{\pi^0} = 5.6$ MeV/c² @ 1 GeV
- ✓ With background

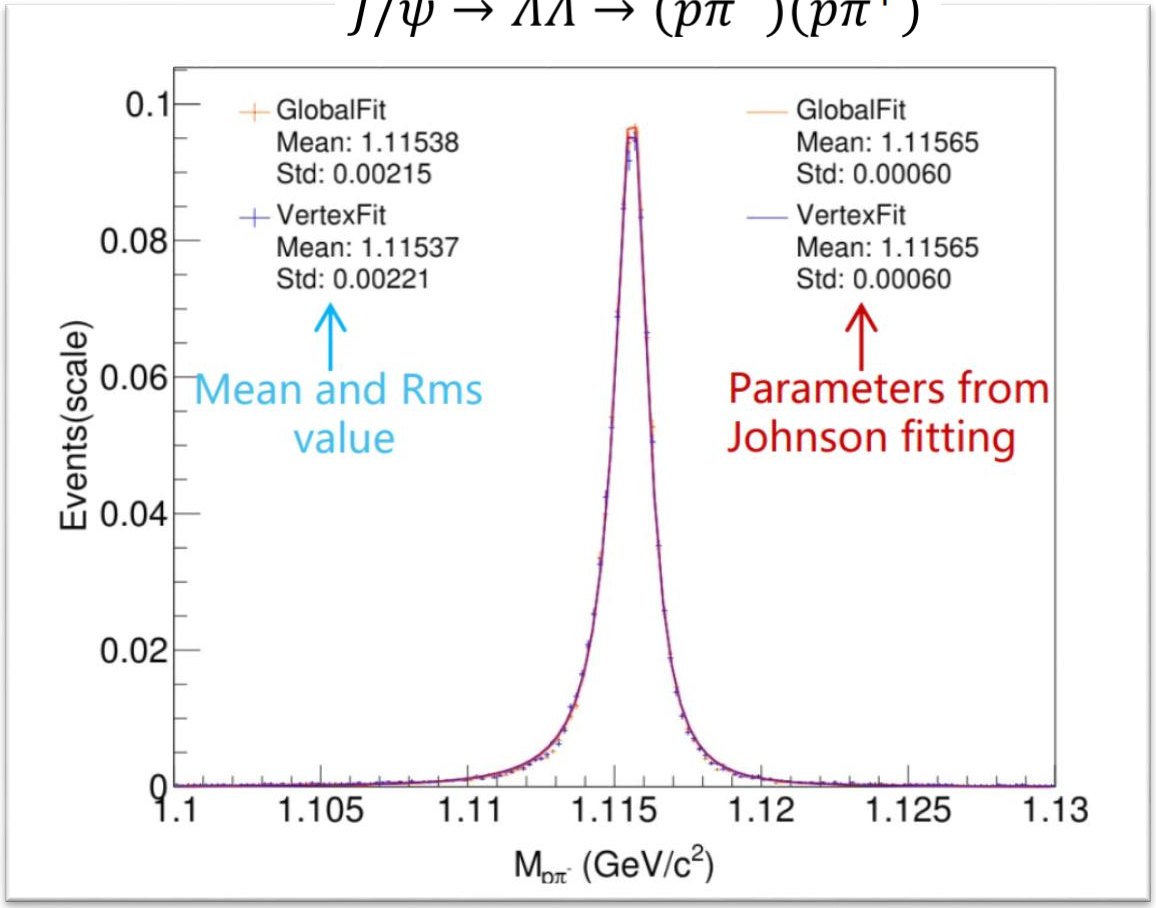
Solid performance for both single **photon** and π^0 at STCF

Vertexing highlights

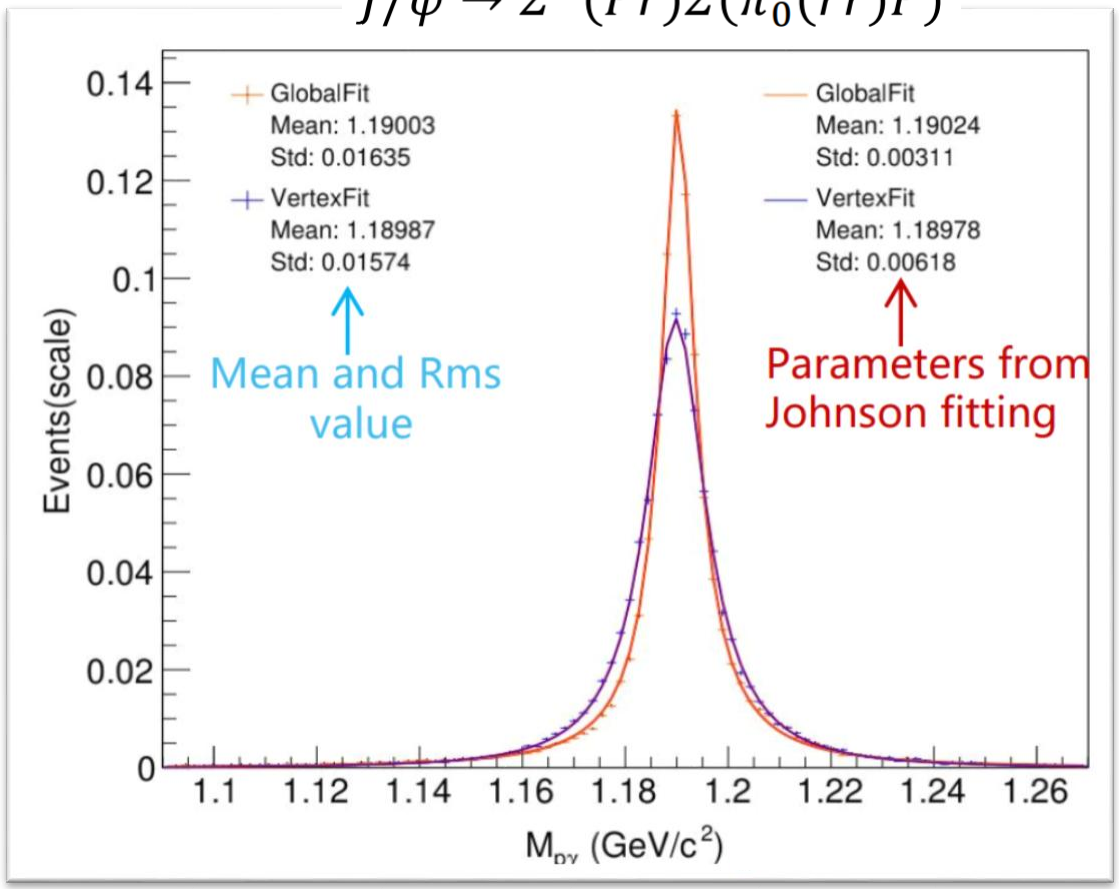
M. Y. Yu,
STCF global vertex fit

- New advanced global vertex fit implemented at STCF, improving both **resolution** and **efficiency**

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

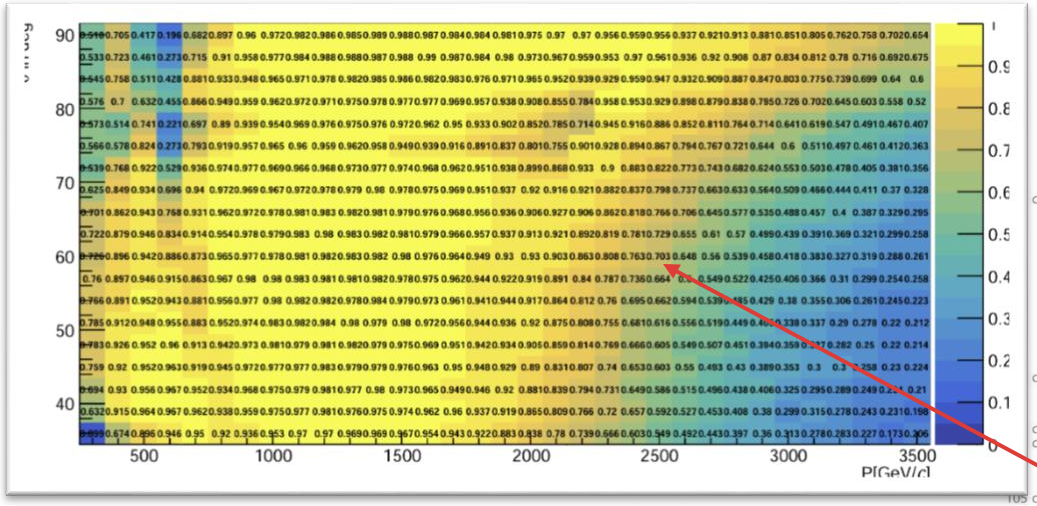


$J/\psi \to \Sigma^+(Pr)\overline{\Sigma}(\pi_0(rr)\overline{P})$

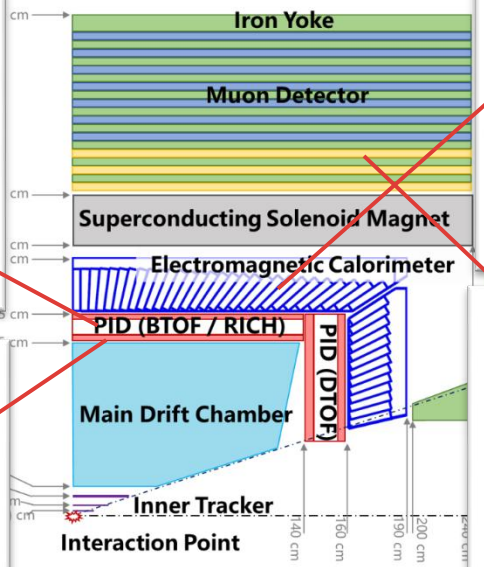


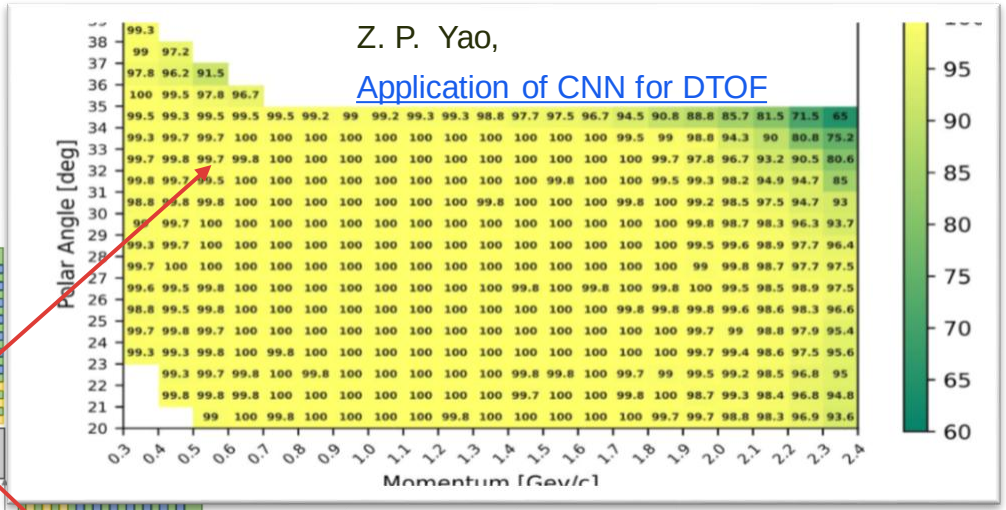
PID highlights

- Mature **likelihood**-based and **CNN/BDT**-based techniques to **exploit/optimize** individual detectors at STCF!

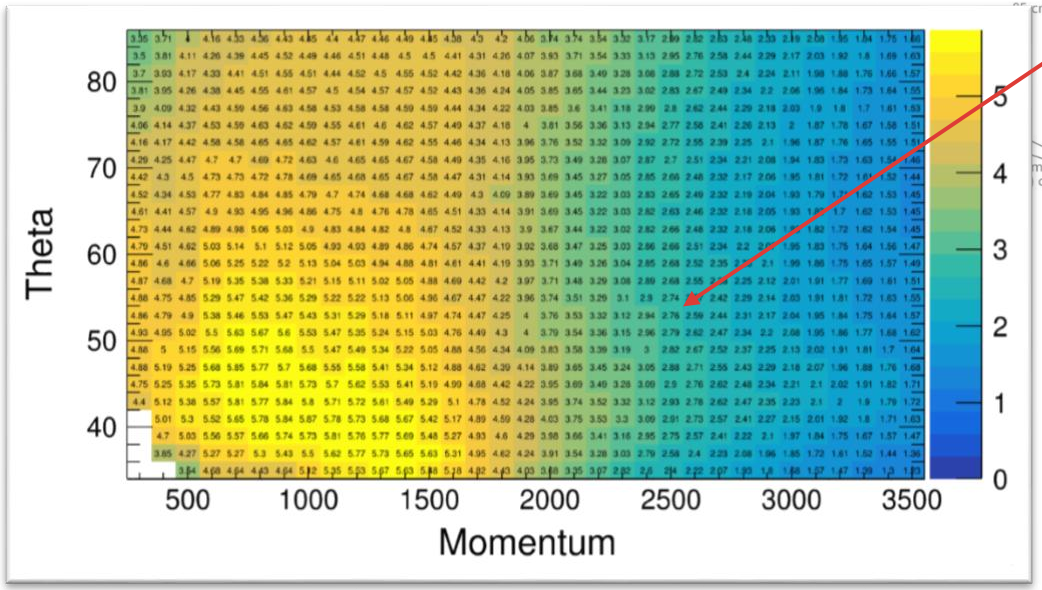


Q. Y. Huang,
STCF RICH simulation,
digitization and reconstruction

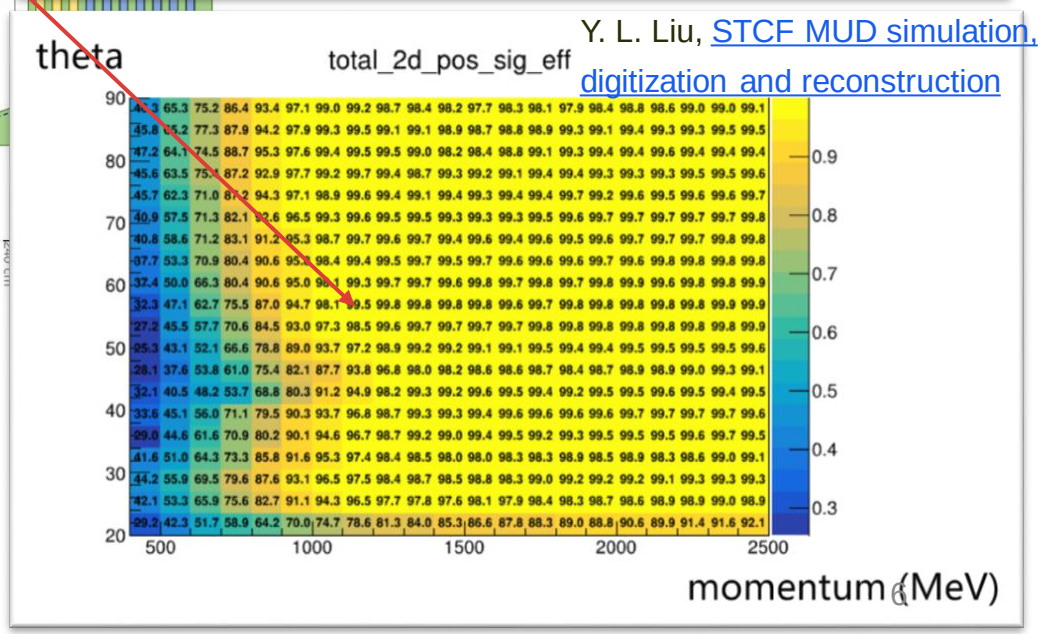




Z. P. Yao,
Application of CNN for DTOF



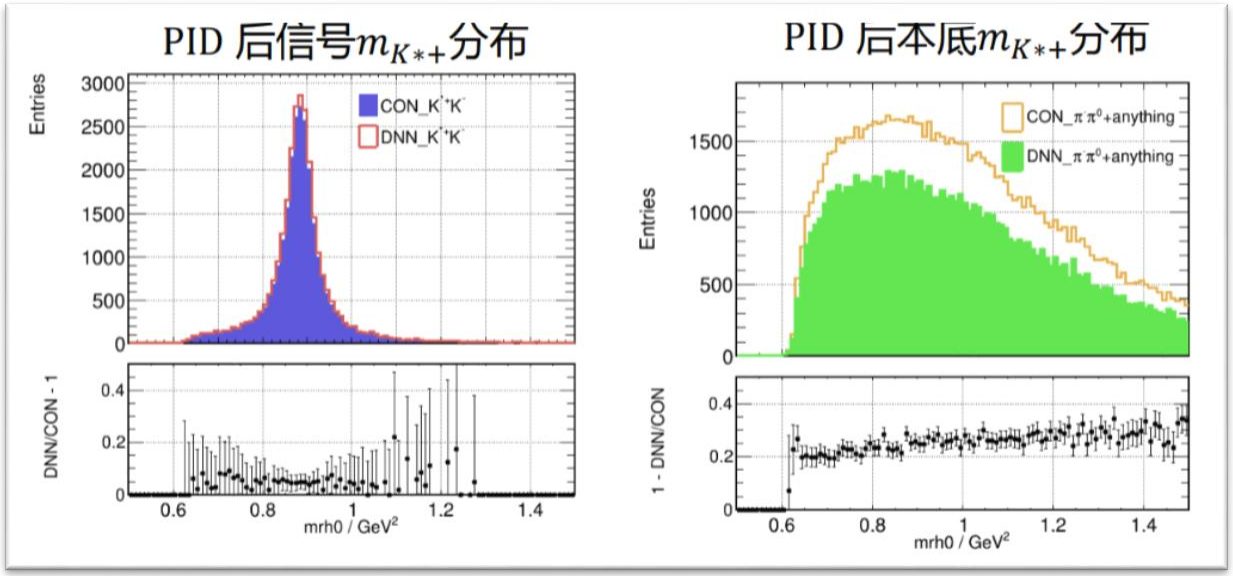
T. Ma,
STCF BTOF simulation,
digitization and reconstruction

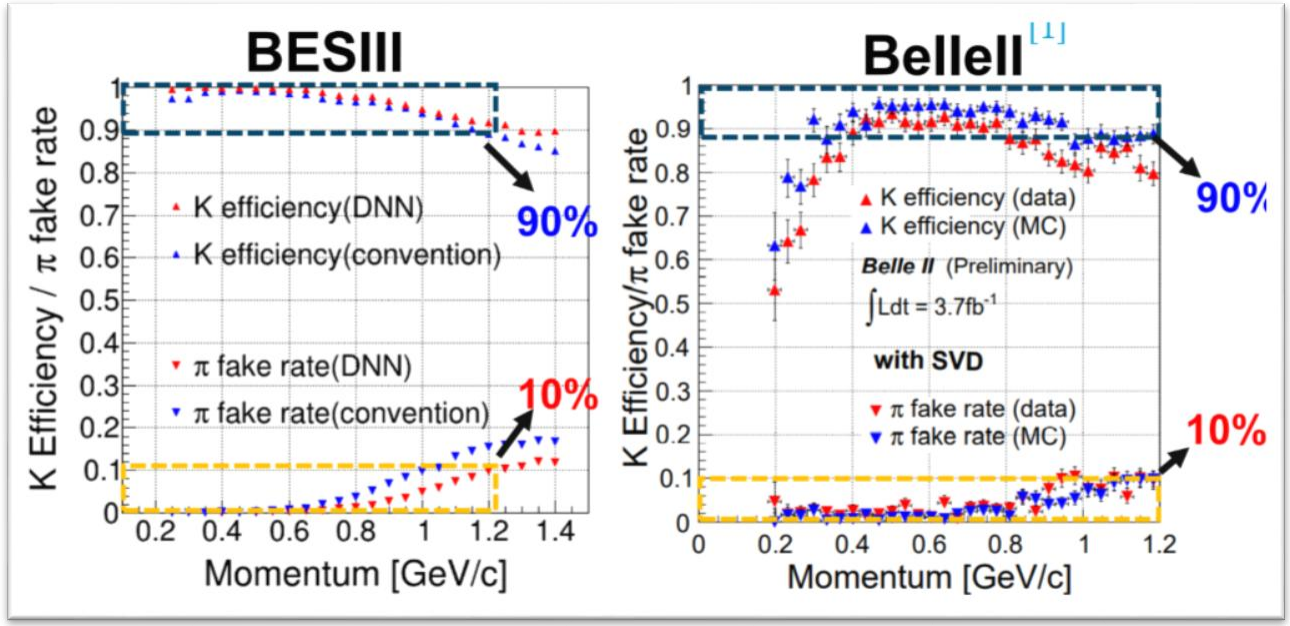


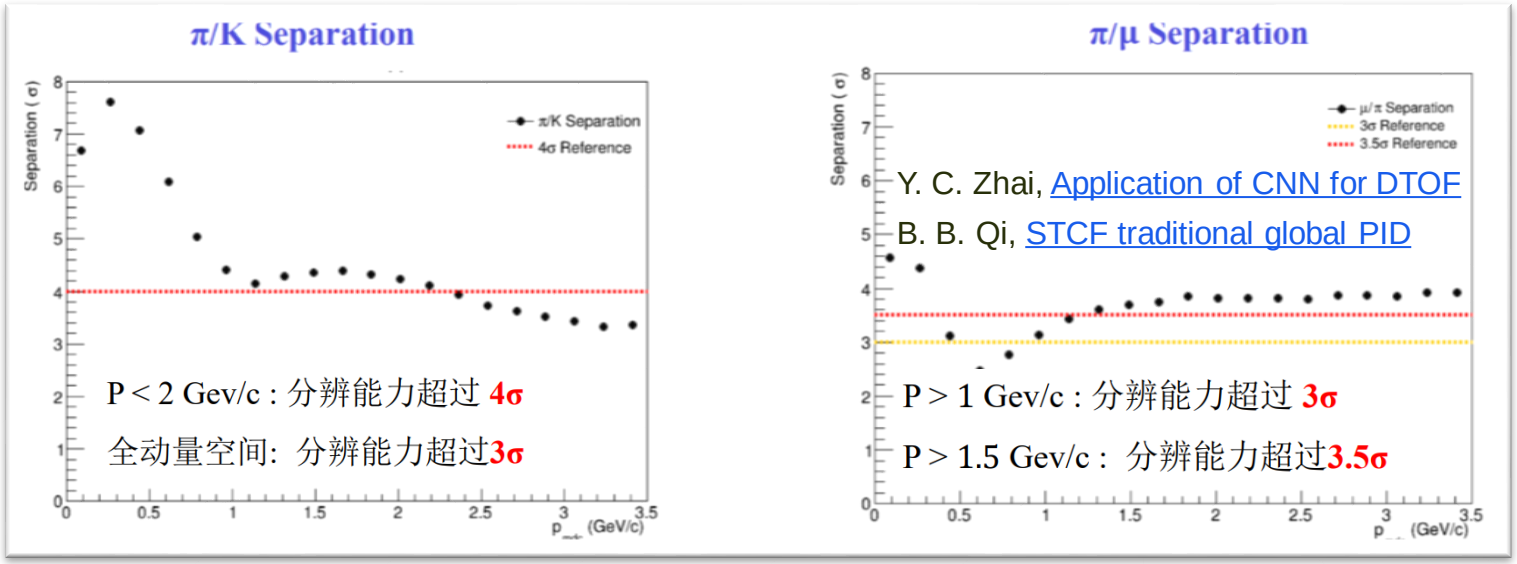
Y. L. Liu, STCF MUD simulation,
digitization and reconstruction

PID highlights

H. Yuan, ML-based PID at BESIII



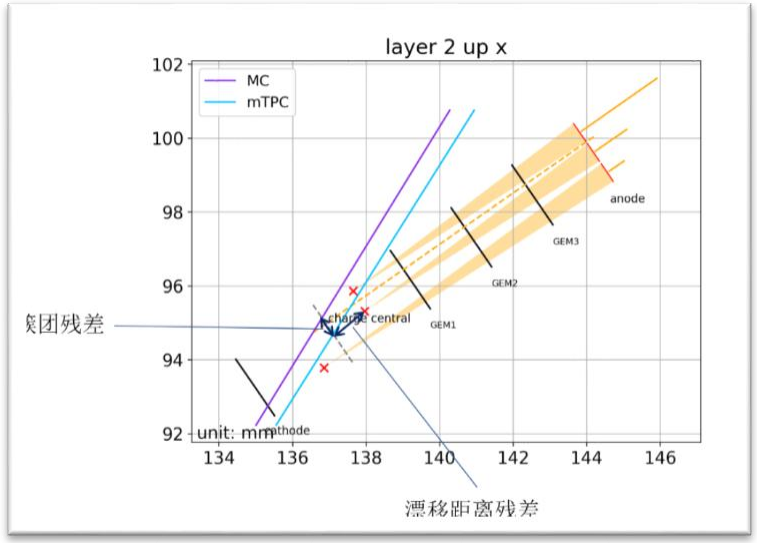


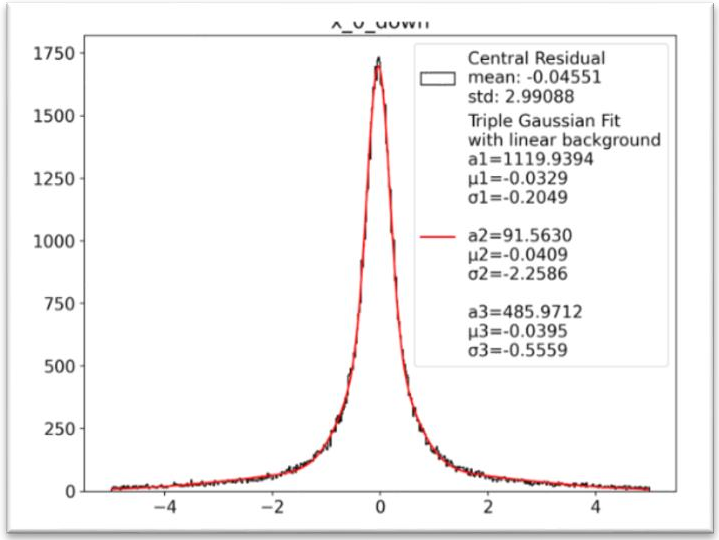


- **Global DNN** provides significant **better efficiency, lower fake rate** and **reduced syst.** compared to conventional method at BESIII
- **Global BDT** provides promising performance at STCF, consolidated by **traditional counterpart**

Calibration highlights

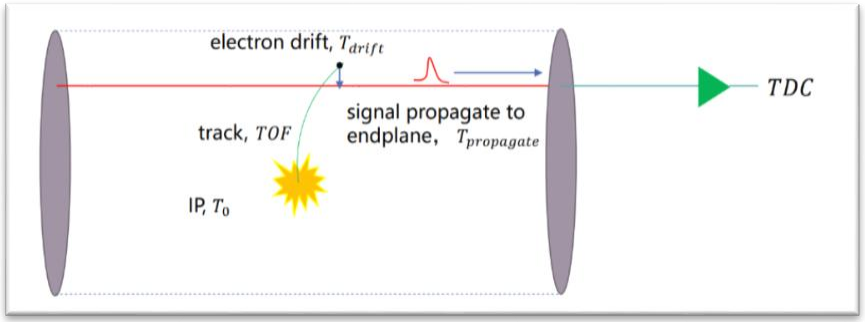
D. Jiang, BESIII CGEM microTPC calibration





~500 um spatial resolution using **microTPC calibration** at BESIII CGEM

- T-Q is the dominant contribution to time resolution

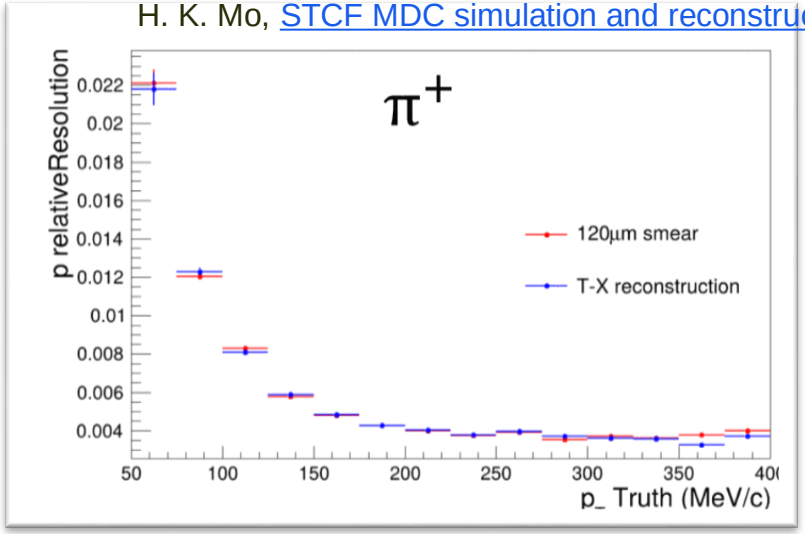


Z. Z. Liang, STCF DTOF T0 reconstruction

物理过程	质心能量	T0重建效率	筛选500MeV径迹	筛选400MeV径迹	筛选300MeV径迹
$e^+e^- \to e^+e^-$	4 GeV	**99.60%**	99.75%	99.75%	99.73%
$e^+e^- \to \mu^+\mu^-$	4 GeV	**98.48%**	98.53%	98.53%	98.53%
$e^+e^- \to \pi^+\pi^- J/\psi \to \pi^+\pi^- e^+e^-$	3.686 GeV	**98.83%**	99.37%	99.37%	99.18%
$e^+e^- \to \pi^+\pi^- J/\psi \to \pi^+\pi^- K^+K^-$	4 GeV	**97.97%**	96.84%	96.81%	97.59%
$e^+e^- \to \pi^+\pi^- J/\psi \to \pi^+\pi^- \mu^+\mu^-$	3.686 GeV	**98.68%**	99.43%	99.43%	99.29%
$e^+e^- \to \pi^+\pi^- J/\psi \to \pi^+\pi^- p^+\bar{p}^-$	4 GeV	**95.92%**	89.17%	89.19%	92.30%

>98% TOF-based **event T0** reco efficiency in most processes at STCF

H. K. Mo, STCF MDC simulation and reconstruction



Full MDC tracking methodology (using drift T) is in place at STCF, providing comparable performance to toy smearing method

Analysis techniques highlights

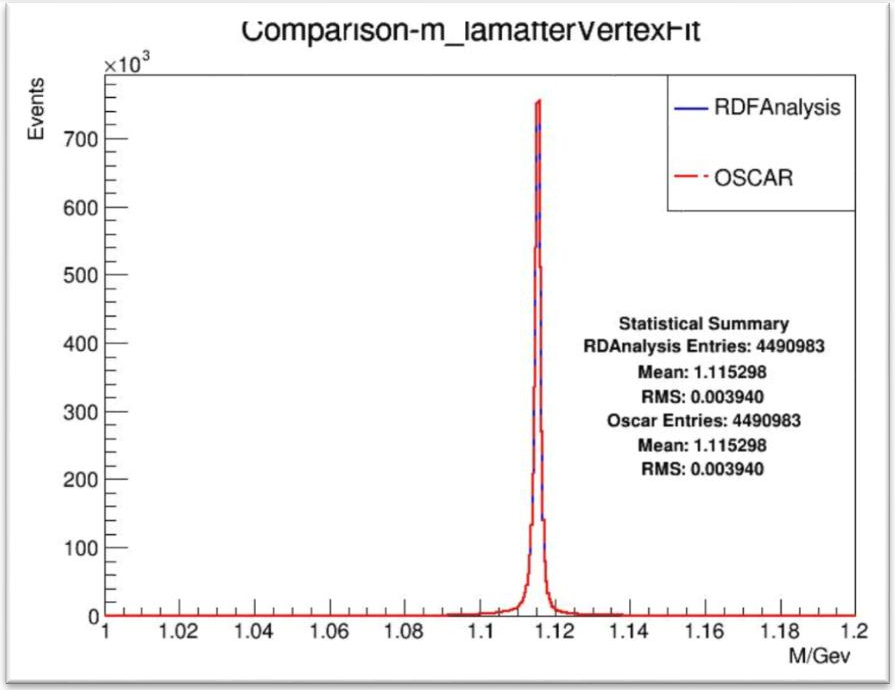
Y. Yang, STCF Rdata-Frame analysis framework

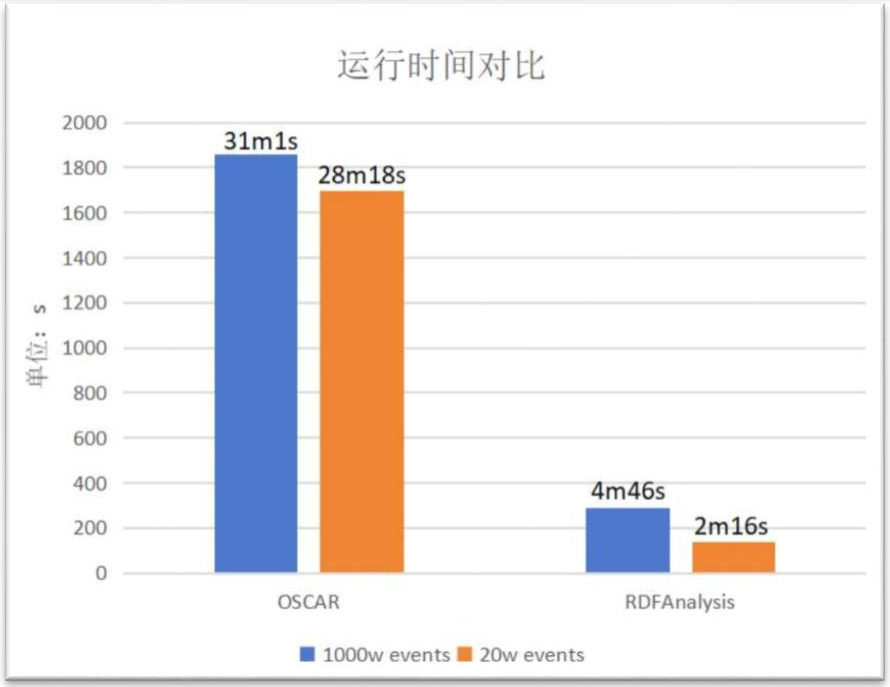
```
# 1. set input or output
input_file = "/lzufs/user/yangying/RecLam1000w/5000files/reclamphsp262new_2933.root"
output_file = "/lzufs/user/yangying/outputtry/analysis_results.root"
nthread=8
# 2. set
particle_config = {
    'particle1_mass': 0.938272,    # select  particle1 mass
    'particle2_mass': 0.139570,      # select  particle2 mass
    'particle1_ID': 4,    # select  particle1 ID
    'particle2_ID': 2,      # select  particle2 ID
    'min_tracks': 4,              # number of charged track
    'min_positive_charge': 2,     # number of positively charged tracks
    'min_negative_charge': 2      # number of negatively charged tracks
}

# output
output_vars = ["RDFAm_lamaftervfit","RDFAm_lamafterkfit","RDFAm_lambeforefit","chi2","RDFAantilam_decaylength","RDFAlam_decaylength","RDFAm_antilambeforefit","RDFAm_an
tilamaftervfit","RDFAm_antilamafterkfit"]
```

Root::RdataFrame based analysis at STCF implemented and validated

- Much **faster** by exploiting parallelization!



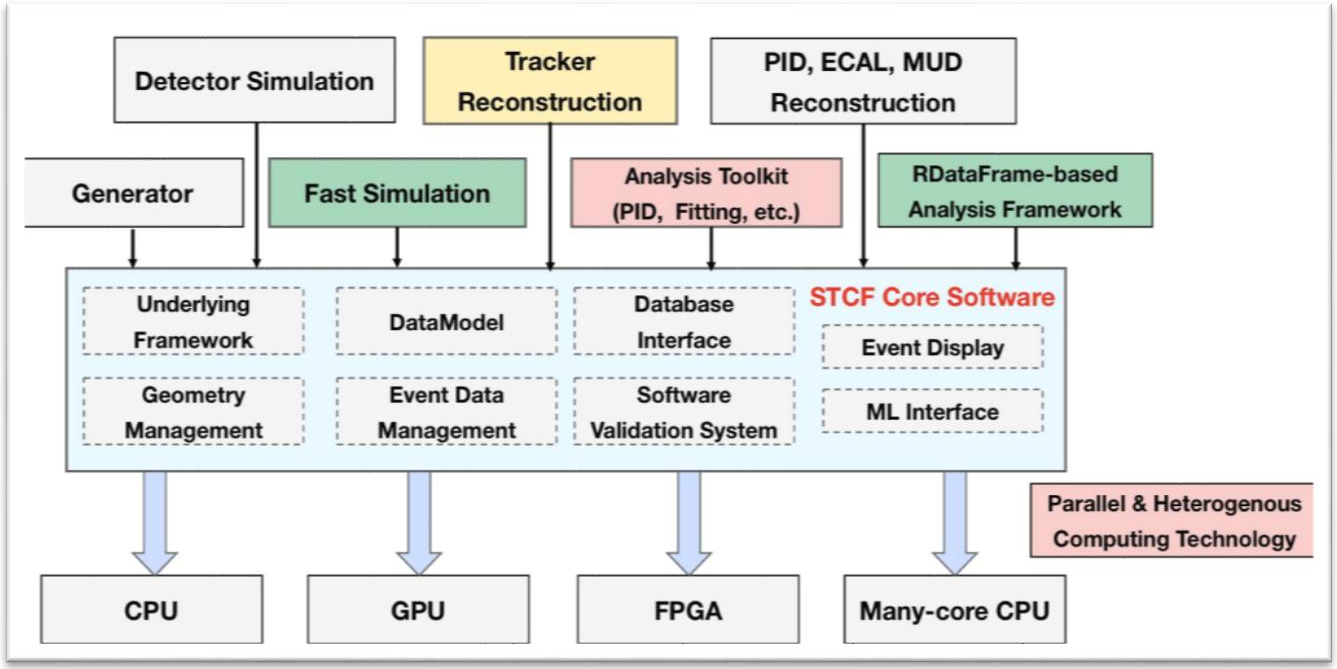


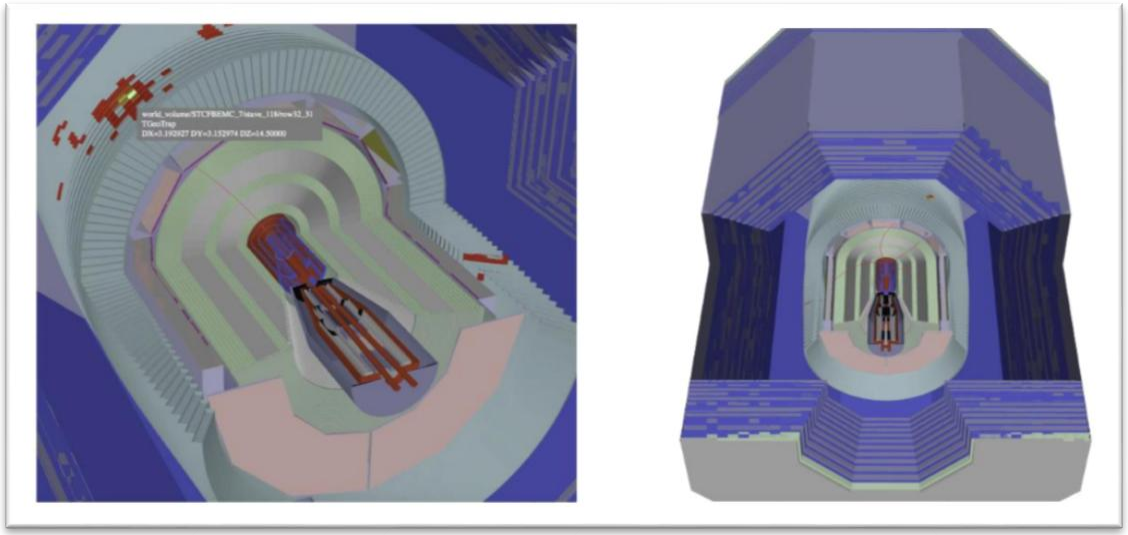
Software framework highlights

T. Li, STCF core software status

Q. B. Zhang, STCF event display

STCF offline software provides mature functionalities for detector simulation, digitization, calibration and reconstruction …, driving ongoing/upcoming **fine physics simulation** and **detector optimization** towards TDR



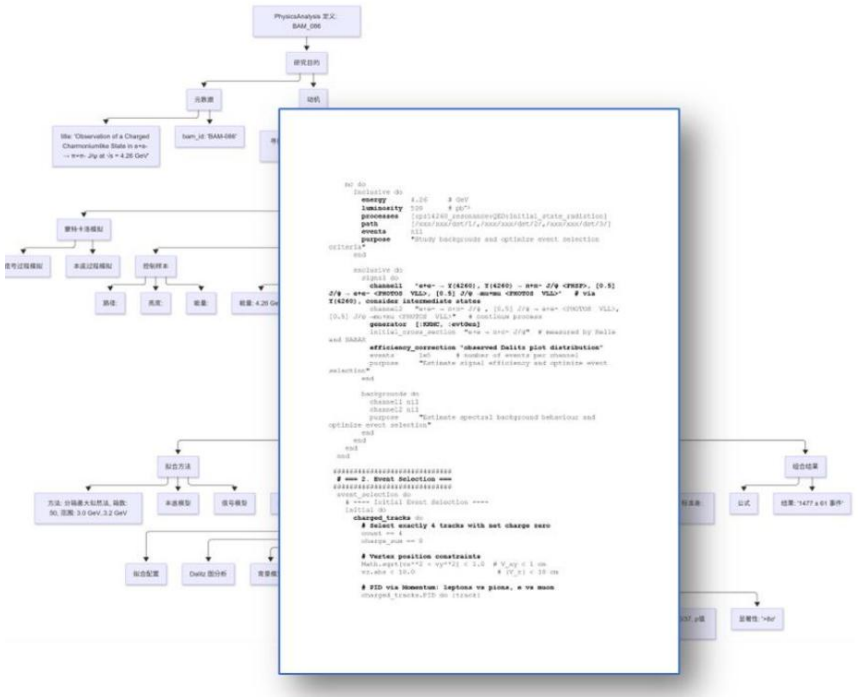


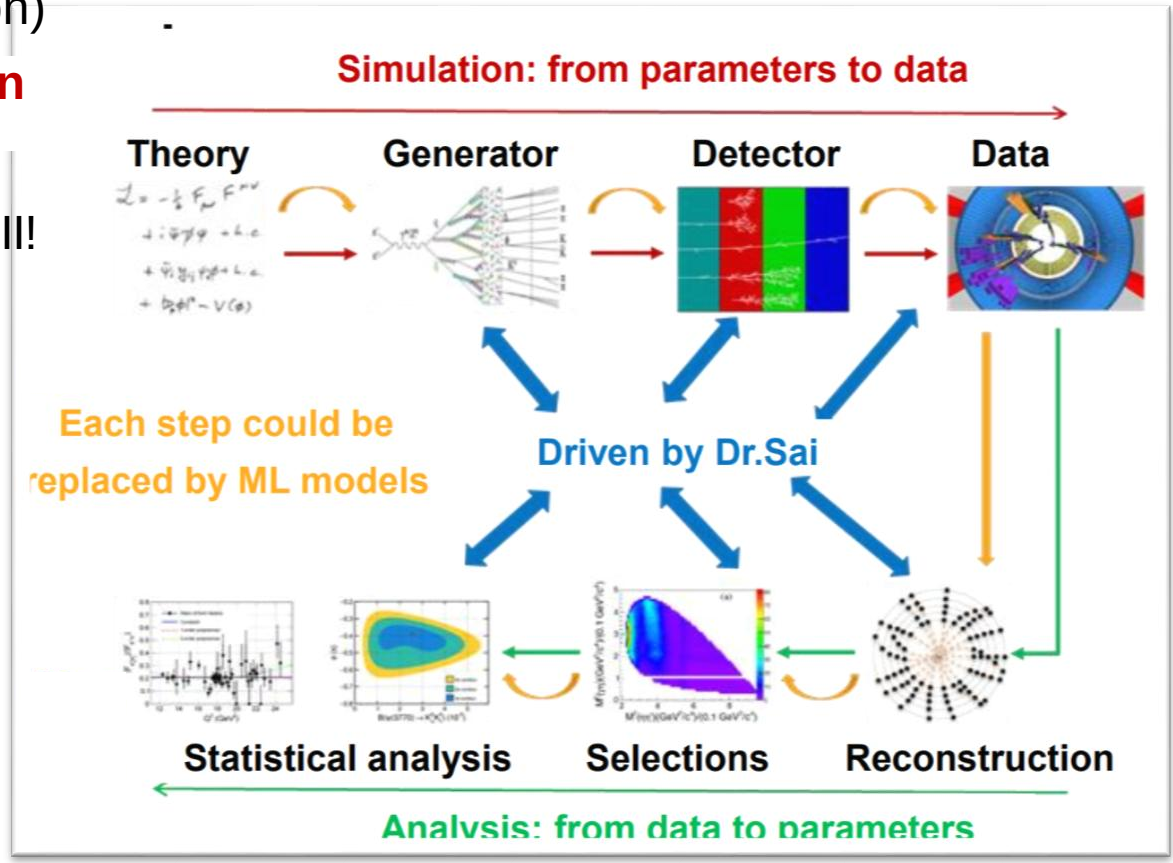
Future: new HEP data analysis paradigm?

Dr. Sai project (based on LLM) aim to **automate** the HEP data analysis

- Dr.Sai V2 is released (automation up to pre event selection)
- Analysis sequence translated (manually now) to a **Domain Specific-Language** (DSL) and fed into LLM
  - Facilitated by >600 physics results published at BESIII!

K. Li, AI Assistant for HEP data analysis- Dr. Sai





Summary

- 21 talks covering basic topics in HEP event processing
- Advanced ML techniques are explored in different chain of event processing at STCF, BESIII…
  - Tracking, PID, data analysis …
  - Complementary to traditional methods in particular phase space

Many thanks to our 21 speakers in software parallel sessions!

- Tracking &vertexing, photon & π^0, PID, analysis etc. has been much consolidated at STCF
- STCF is in full swing for detector optimization and physics simulation

根据STCF一年的亮度1ab^{-1}所估计事例数：

Ecm	Physics goal	Bhabha	Dimu	Digam	Ditau	Two-photon	Hadronic
2.125	Two-photon, light meson	1.6× 10^{12}	2× 10^{10}	7.9× 10^{10}	-	>2× 10^{10}	4.4× 10^{10}
3.097	Hyperon, new physics	8.9× 10^{11}	9× 10^{9}	3.7× 10^{10}	-	>2× 10^{10}	4.5× 10^{12}
3.770	D, $\pi\pi$	5.1× 10^{11}	6.6× 10^{9}	2.5× 10^{10}	2.6× 10^{9}	>3× 10^{10}	2.4× 10^{10}
4.03	D^*, D_s	4.5× 10^{11}	5.8× 10^{9}	2.2× 10^{10}	3.3× 10^{9}	>3× 10^{10}	2.6× 10^{10}
4.26	Y(4260), τ physics	4.0× 10^{11}	5.2× 10^{9}	1.9× 10^{10}	3.5× 10^{9}	>3× 10^{10}	1.5× 10^{10}
4.682	Λ_c	3.3× 10^{11}	4.3× 10^{9}	1.6× 10^{10}	3.4× 10^{9}	>3× 10^{10}	1.5× 10^{10}
6.0	Doubly charmonium	2.0× 10^{11}	2.6× 10^{9}	1.0× 10^{10}	2.4× 10^{9}	>3× 10^{10}	8.4× 10^{9}
7.0	Fragmentation	1.5× 10^{11}	1.9× 10^{9}	7.4× 10^{9}	1.9× 10^{9}	>3× 10^{10}	6.2× 10^{9}
Scaling factor		1× 10^{-4}	1× 10^{-3}	1× 10^{-3}	3× 10^{-1}	<1× 10^{-3}	1× 10^{-3} at J/ψ, 1× 10^{-2} for others
大小合计		~8.4T	~1.3T	~4.0T	~10T	~4.0T	109T

根据事例大小 20kb/evt 估计，事例数估计。

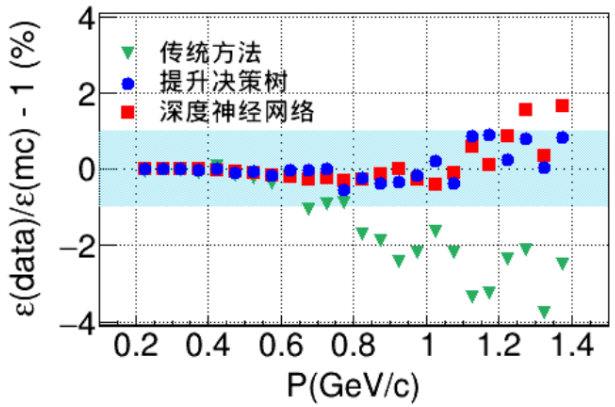
Q. Lan, STCF MC production for physics simulation

Backup

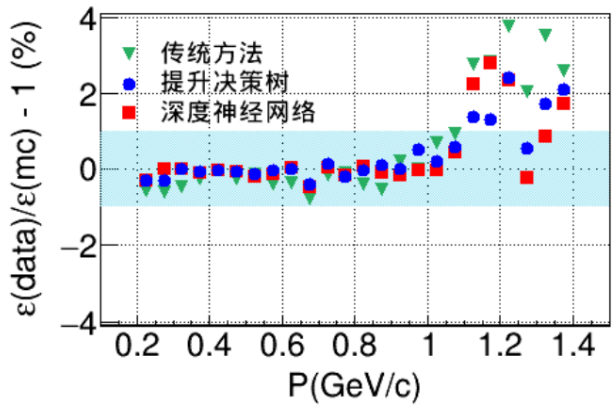
模型性能检查——系统误差

- 系统误差： $\Delta\varepsilon = \frac{\varepsilon(data)-\varepsilon(MC)}{\varepsilon(MC)}$
  - ε：粒子鉴别效率

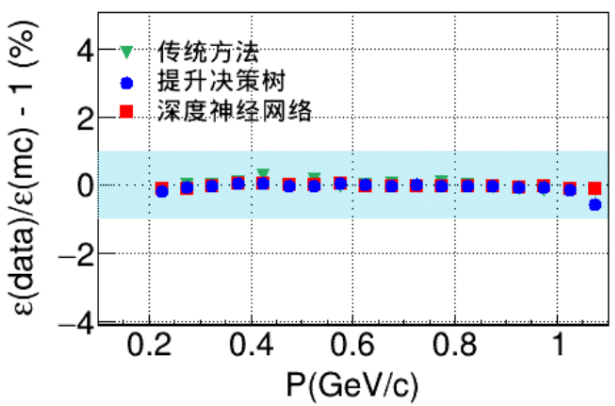
π介子



K介子

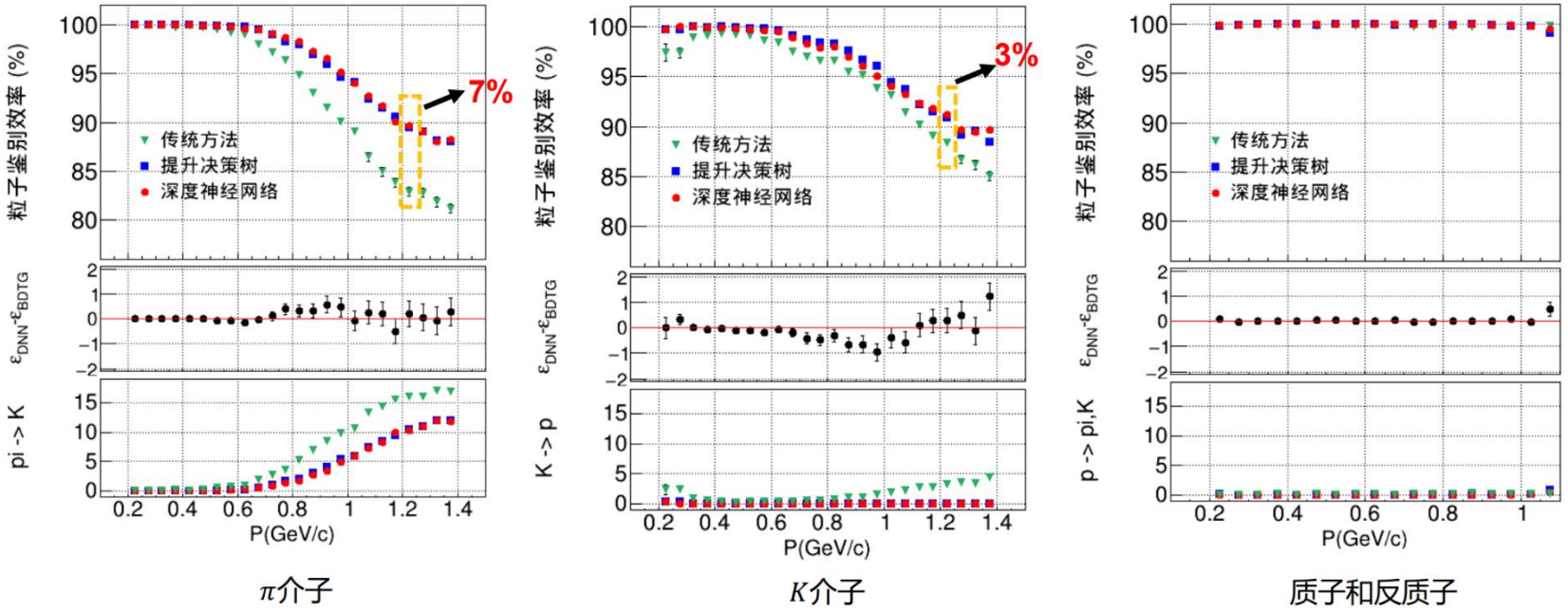


质子和反质子



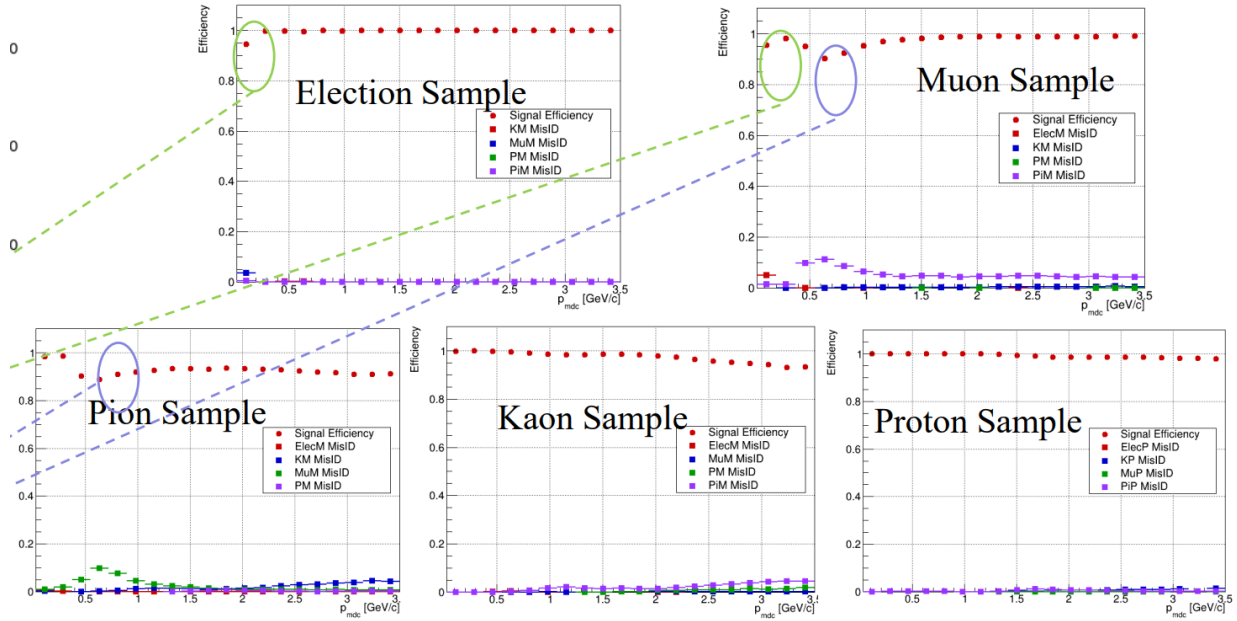
- 深度神经网络的系统误差:
  - π介子和K介子鉴别的系统误差总体小于1%
  - 质子鉴别的系统误差总体小于0.2%

H. Yuan, ML-based PID at BESIII



π介子　　K介子　　质子和反质子

H. Yuan, ML-based PID at BESIII



Y. C. Zhai, Application of CNN for DTOF

B. B. Qi, STCF traditional global PID