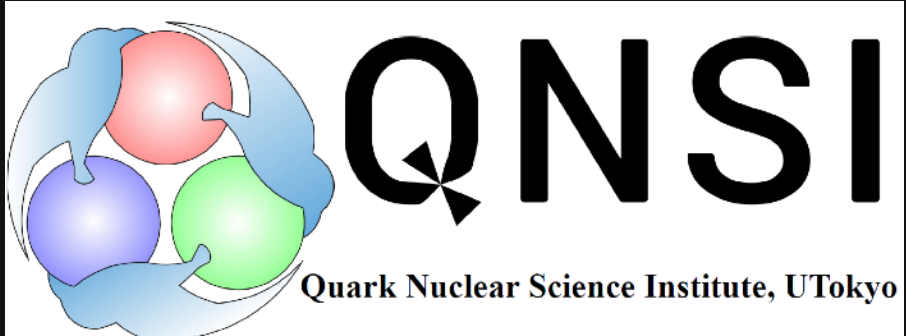
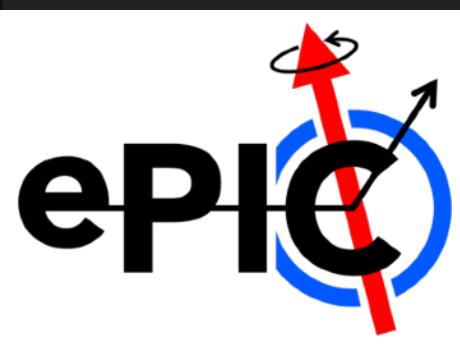


Streaming reconstruction



Quark Nuclear Science Institute, The University of Tokyo
Takuya Kumaoka, Taku Gunji



Tokyo City University
Marco Meyer-Conde, Oliver Partridge

Today's topics

Time splitting and pre-event identification

- Updated the time splitting algorithm (add calorimeters).
- Estimated the pre-event identification quality for some kinds of physics events with the 10x100 GeV backgrounds simulation.

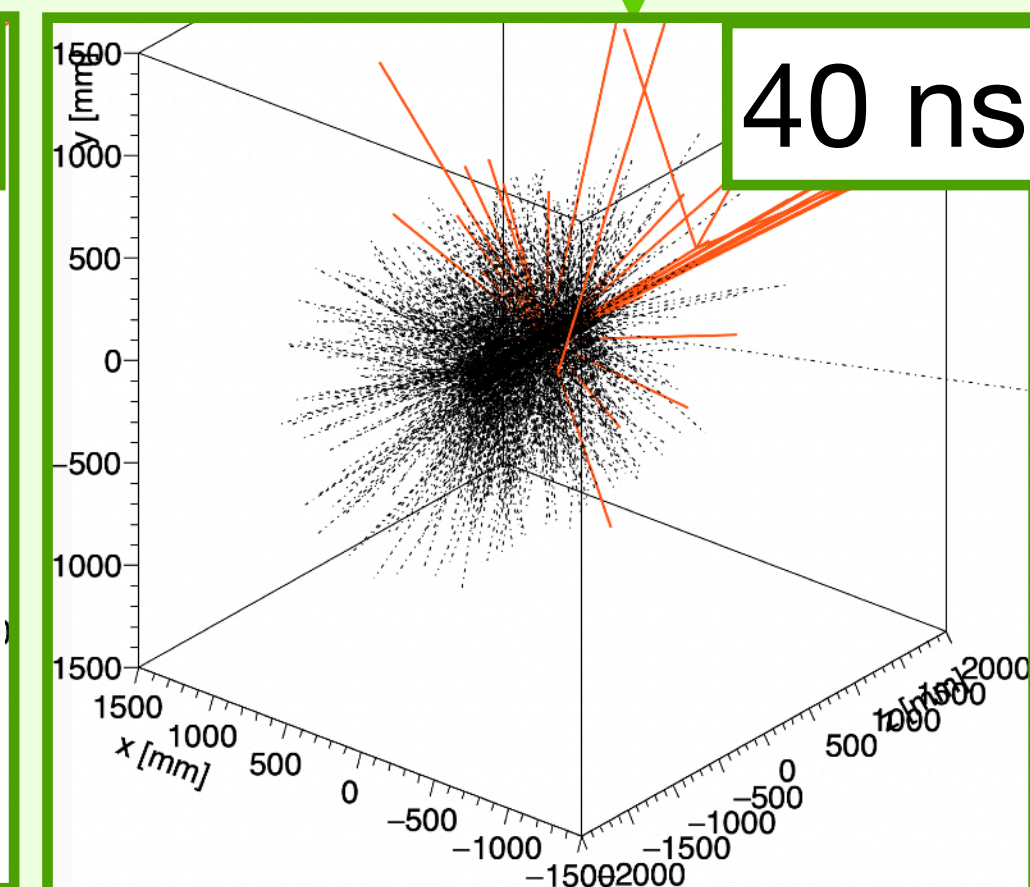
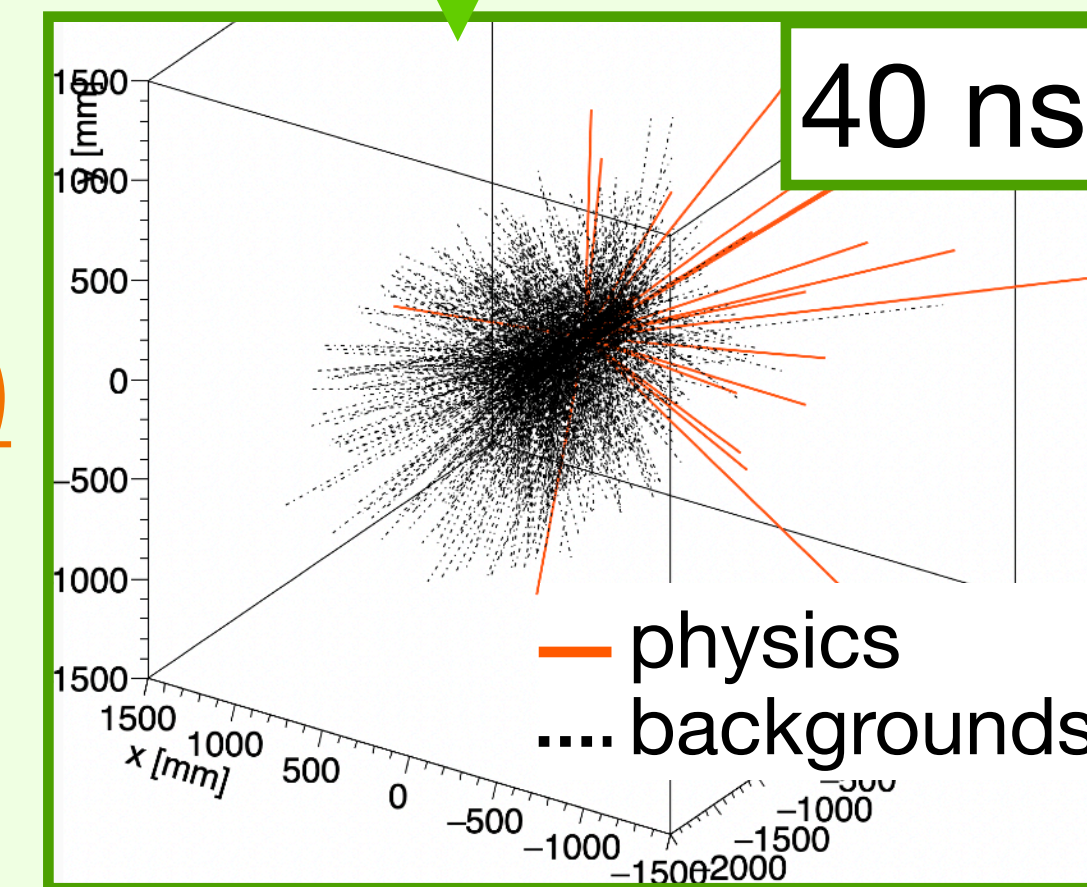
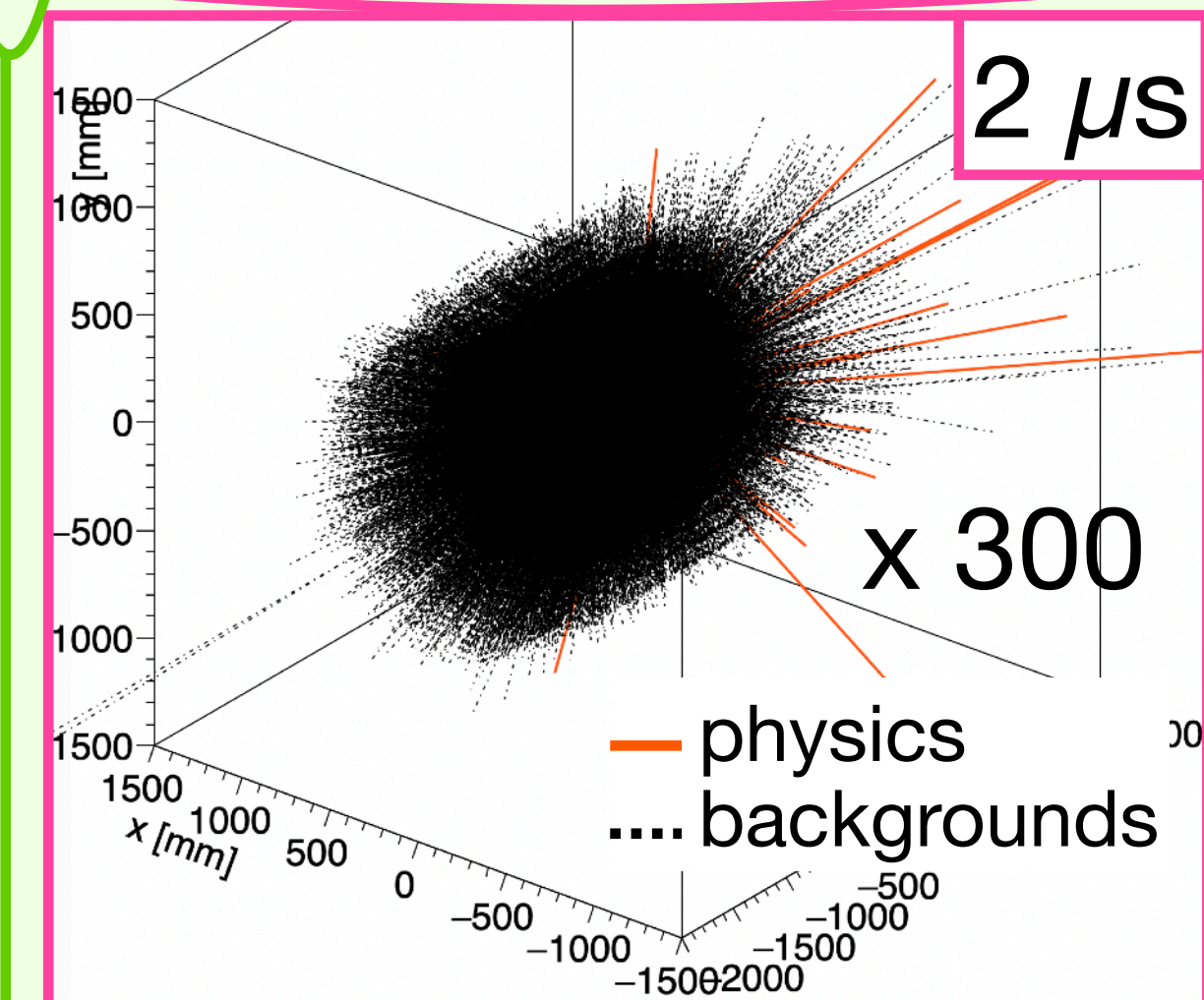
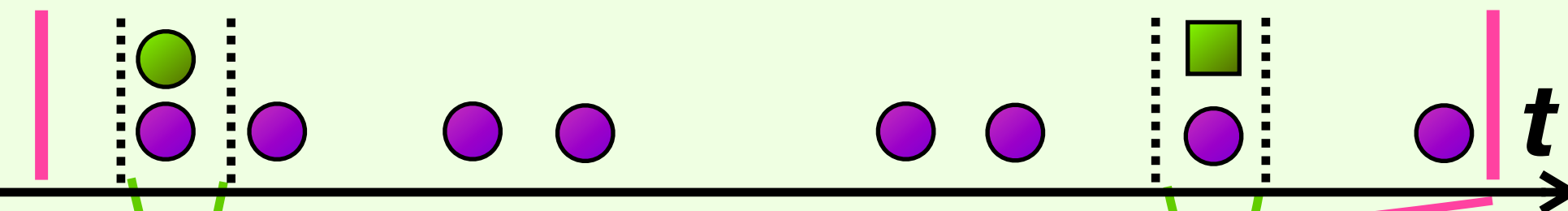
Event identification and event classification using ML

- Started to make a event identification and event classification using ML.
- Got the first some results.
- Implement the algorithm into ElCrecon.

Purpose of this study

- Time-frame (TF) ~ 0.6 ms / super TF $\sim \times 1000$ TFs
- Acquire DIS events at 500 kHz (@ 18×275 GeV²)
- Expected background rate (dominated by Synchrotron radiation) : 3.3 GHz (@ 18×275 GeV²)
- + detector noises
- 300 DIS events/TF and background $\sim 2.0 \times 10^6$ events/TF
- **Essential to deploy streaming reconstruction to select physics event and reject background**
- + **With triggerless readout, events must be identified in software.**

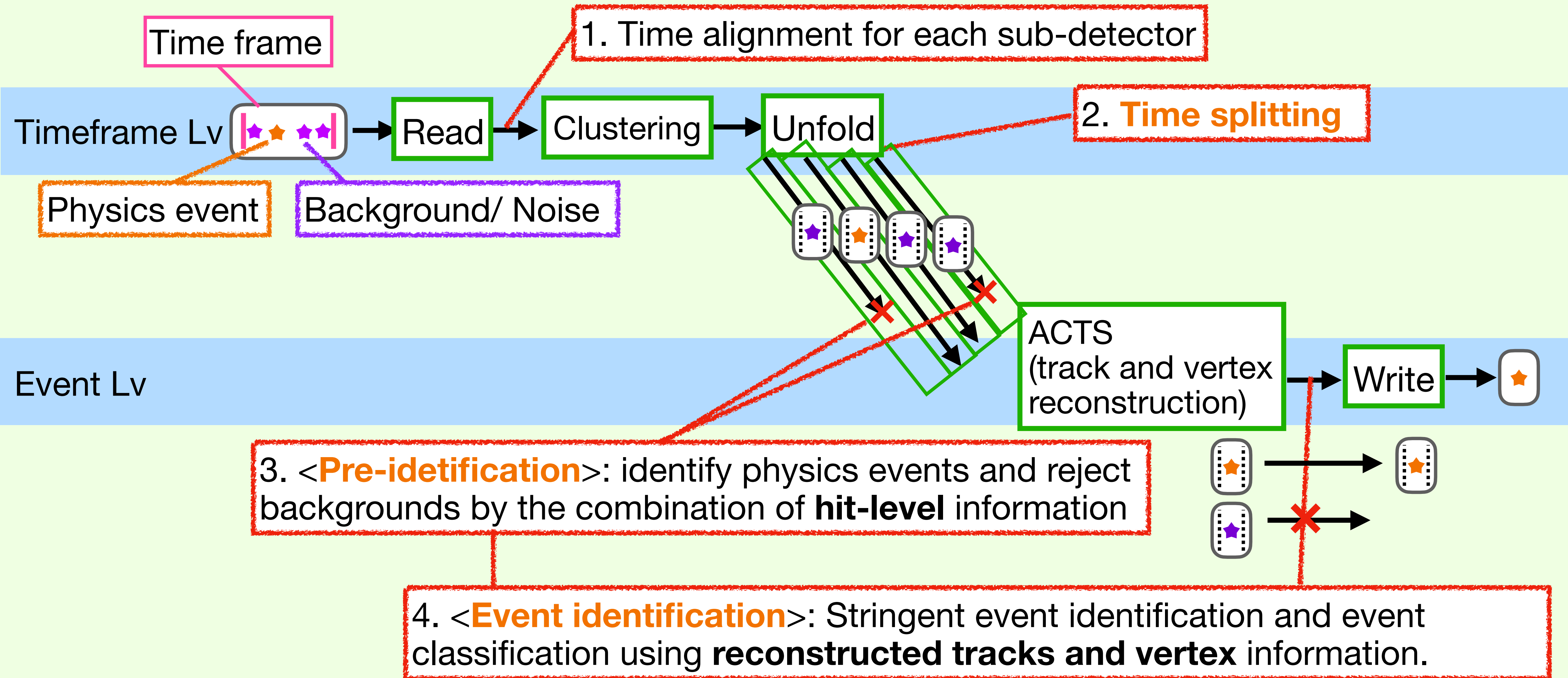
time-frame



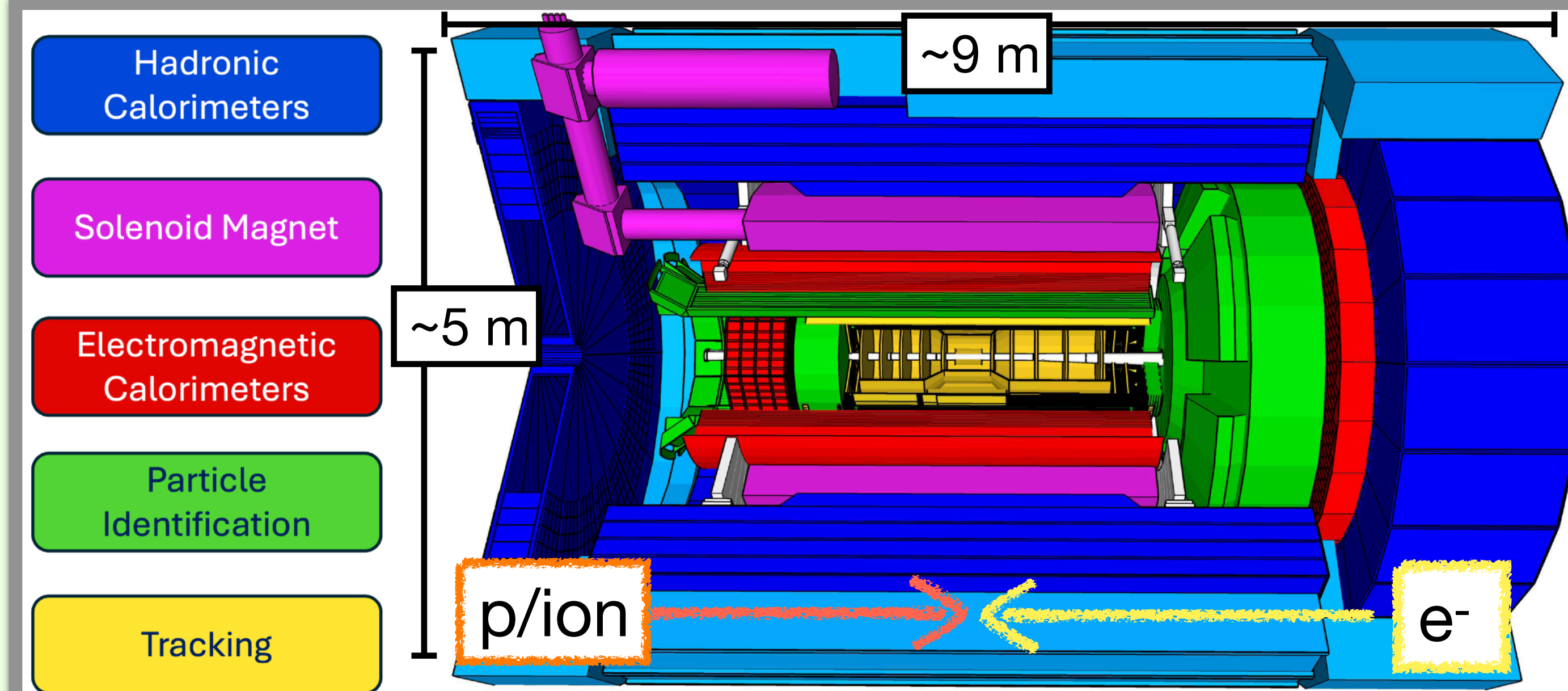
Purpose

Development of streaming event reconstruction and algorithms for event classification (ElCrecon)

Online Event Identification System



ePIC (electron-Proton/Ion Collider) detector / Time resolution



Calorimetry: (ECal: ~20 ns)

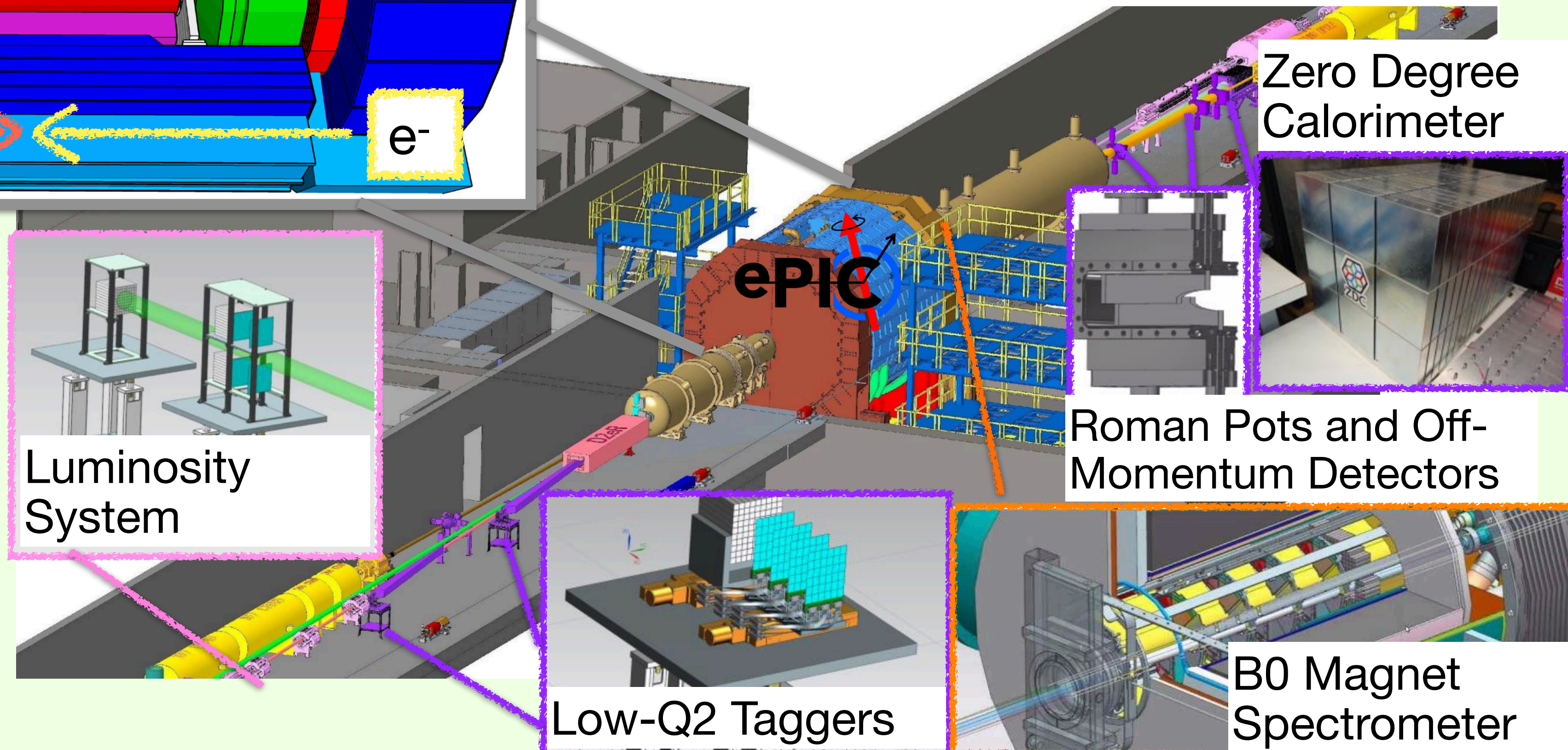
- Imaging Barrel EMCal
- PbWO4 EMCal in backward direction
- Finely segment EMCal+HCal in forward direction
- Outer HCal (sPHENIX re-use)
- Backwards HCal (tail-catcher)

Tracking:

- New 1.7 T solenoid
- Si MAPS tracker (~2 μ s)
- MPGDs (~10 ns)

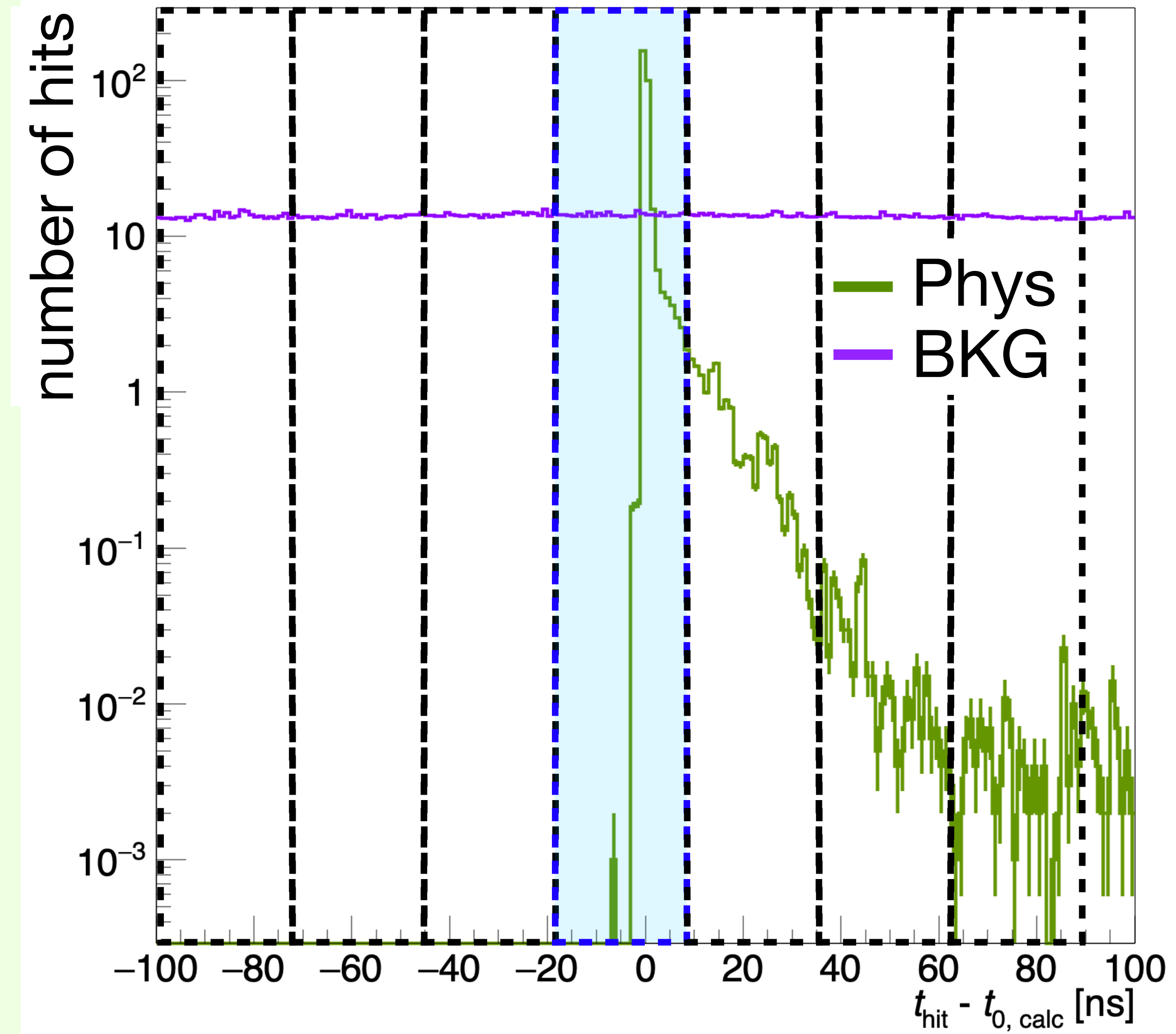
PID:

- hpDIRC
- pfRICH
- dRICH
- AC-LGAD (~30 ps TOF, B0)

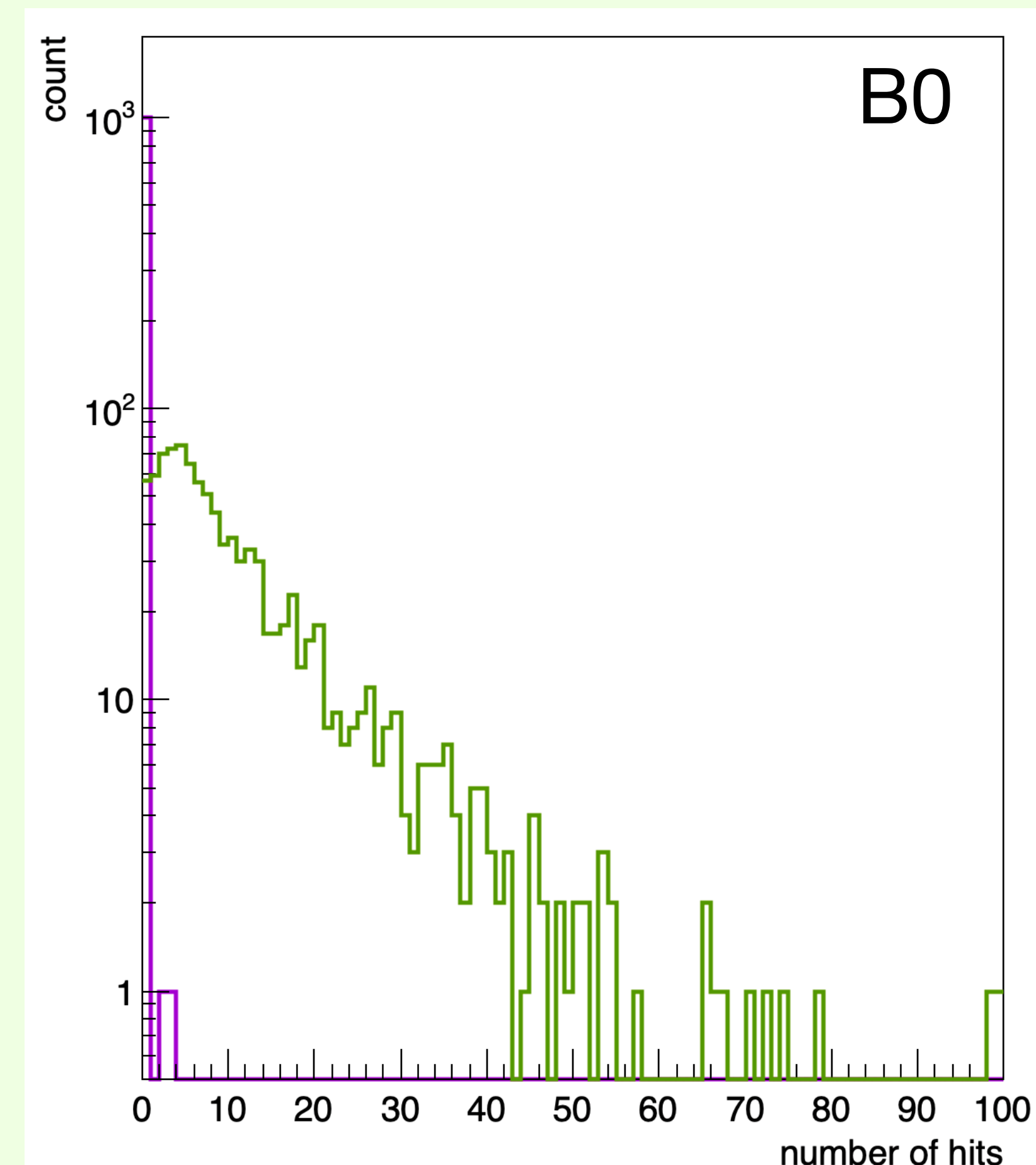
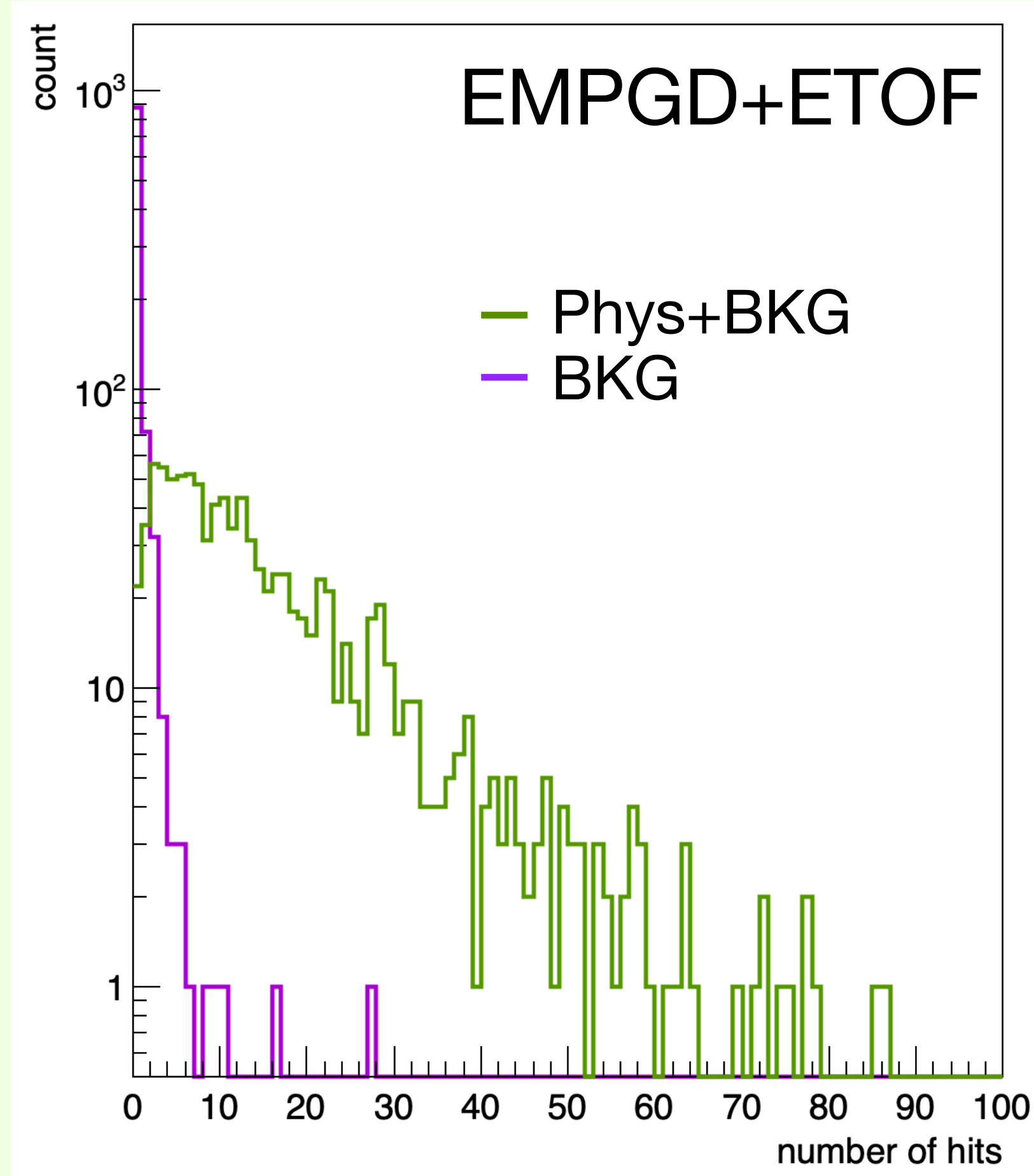
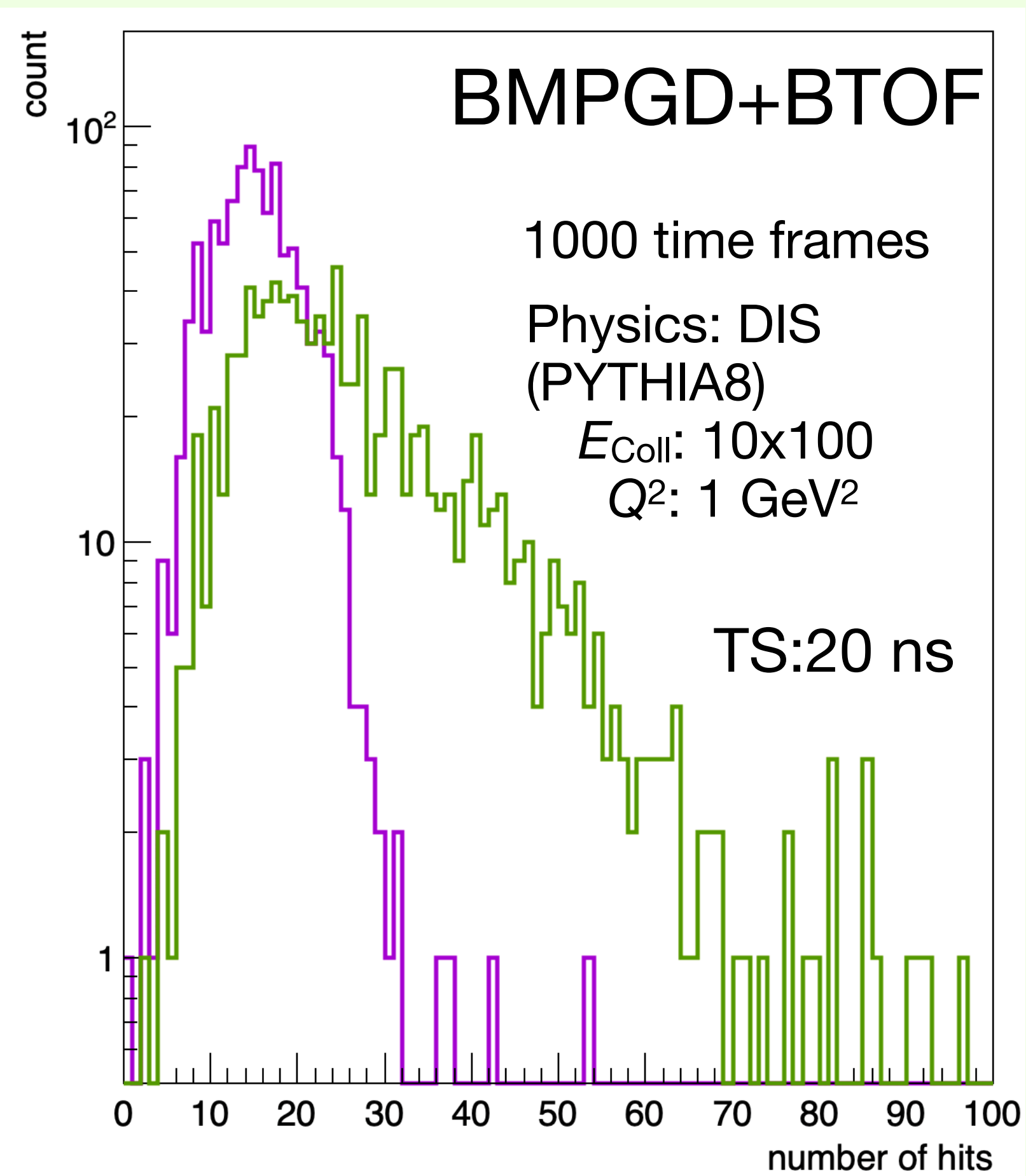


New pre-event identification algorithm (ver.4)

1. Scan time bands for a time frame.
2. Count the number of hits of each sub-detector without MPAS in the time band.
3. Judge the then number of hits over the threshold.
4. If over the threshold, trigger the time region (time slice).



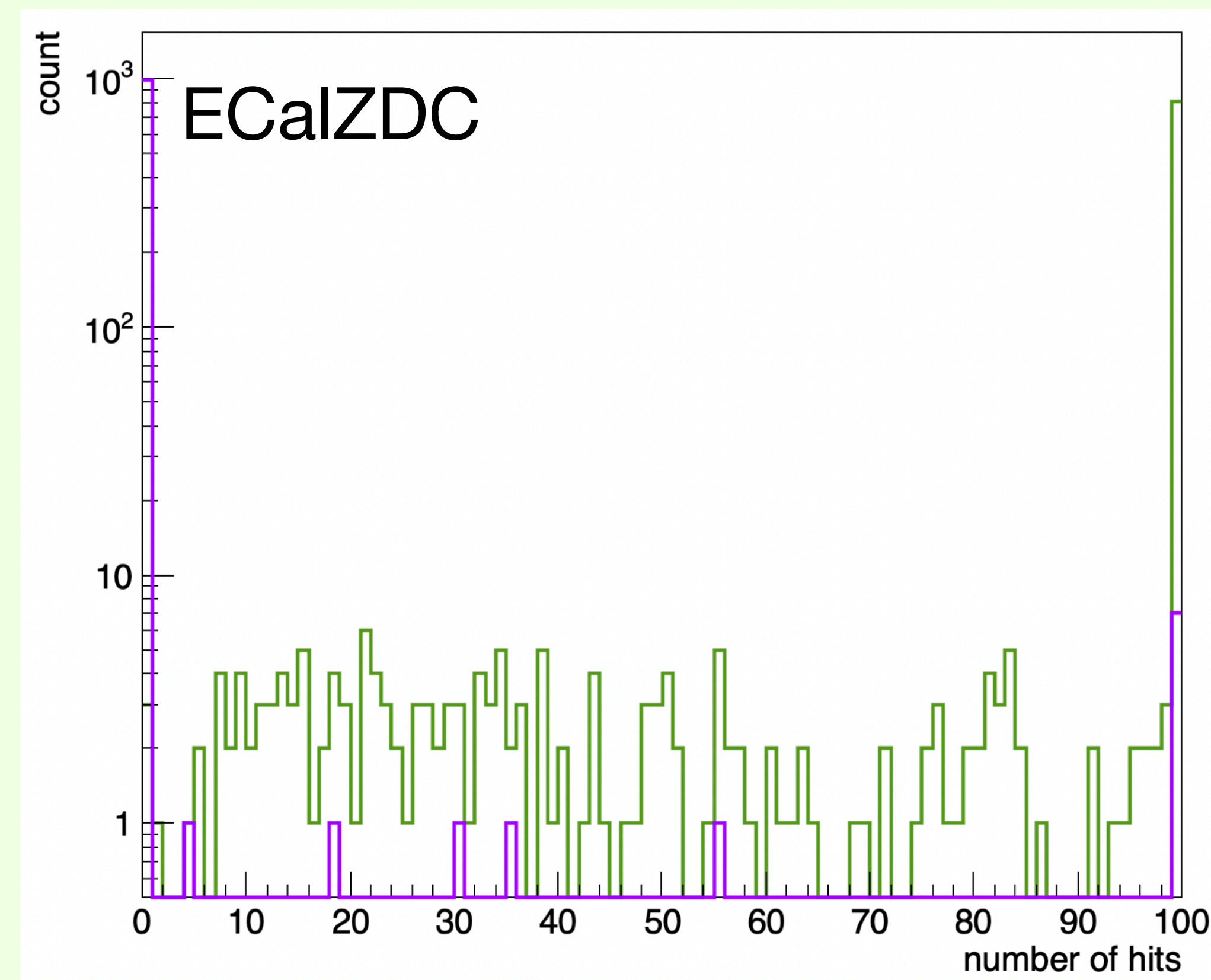
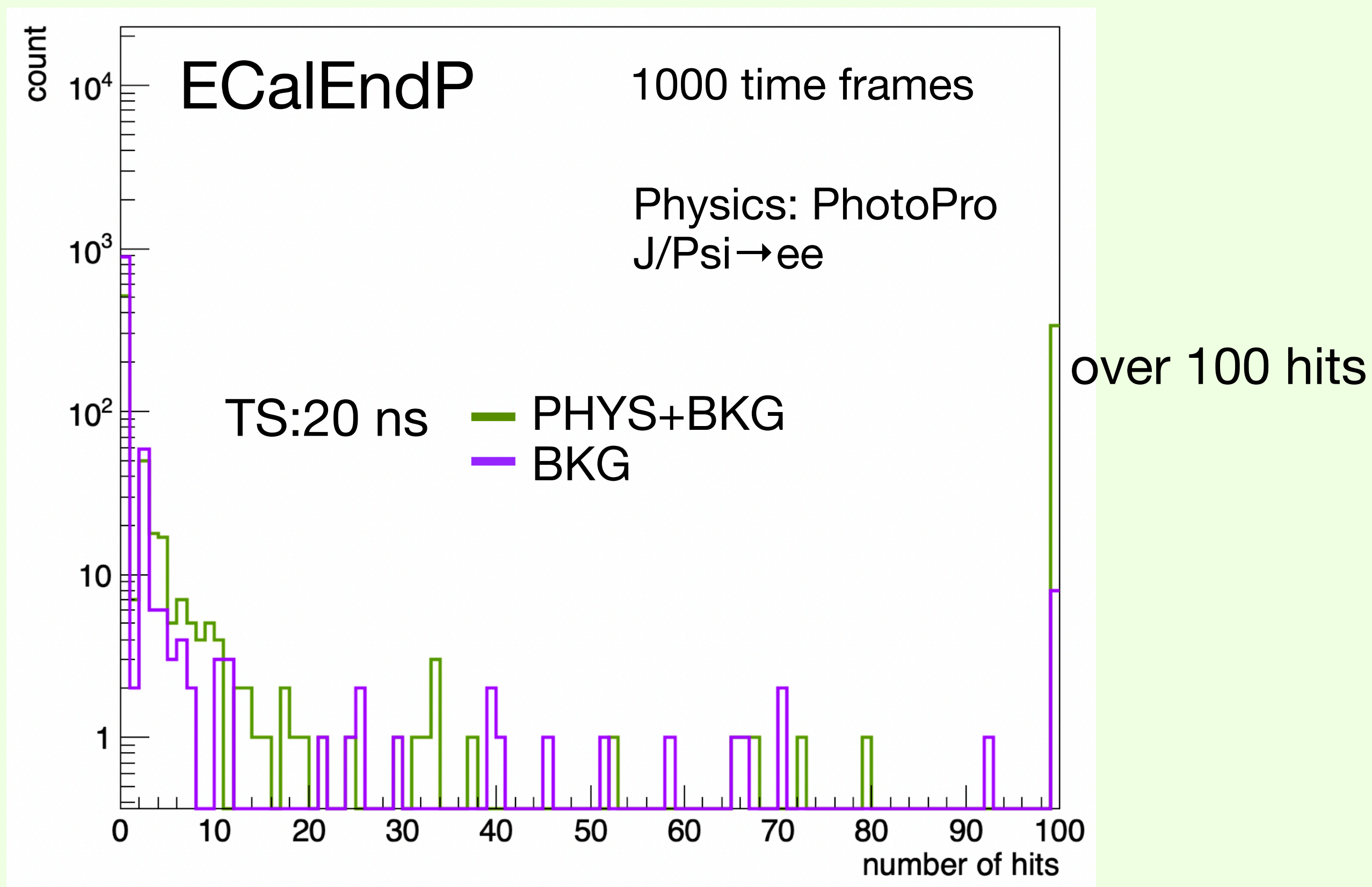
S/N for each detector sets



→ EMPGD+ETOF: num of hits > 2 || B0: num of hits > 0

S/N for each detector sets

PhotoPro J/Psi ee

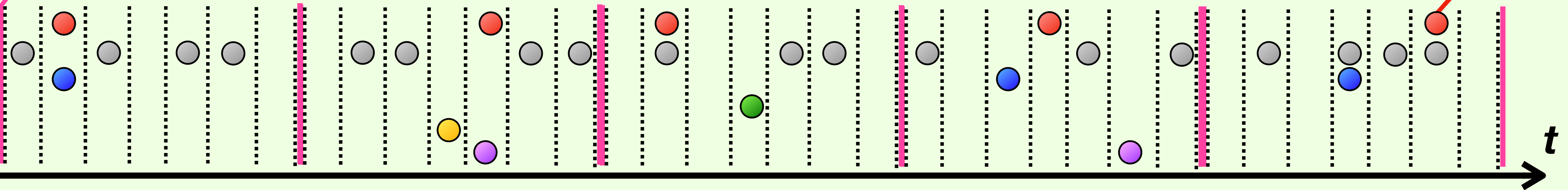


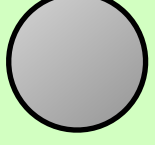
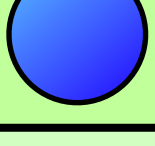
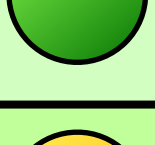
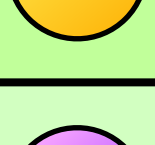
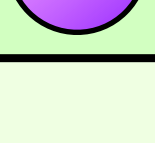
→ ECalEndP: num of hits > 10 || ZDC: num of hits > 10

Simulation for Trigger Quality Evaluation

time-frame (2 μ s: 500 kHz): 1000 events)

physics events (500 kHz)



	DIS NC 10x100 $Q^2 > 1 \text{ GeV}^2$ or 100 GeV^2 SIDIS, DDIS, EXCLUSIVE(photo production J/Psi)	500 kHz
	Synchrontron Radiation	36.6 GHz
	Electron bremsstrahlung radiation	3.18 MHz
	Electron Touscheck scattering (intrabeam dattering)	233.5 kHz
	Electron Coulomb scattering processes	29.6 kHz
	Proton beam gas interactions	31.9 kHz

Evaluate event-identification efficiency in a **blind-analysis** setting, without prior knowledge of event positions in the continuous data stream.

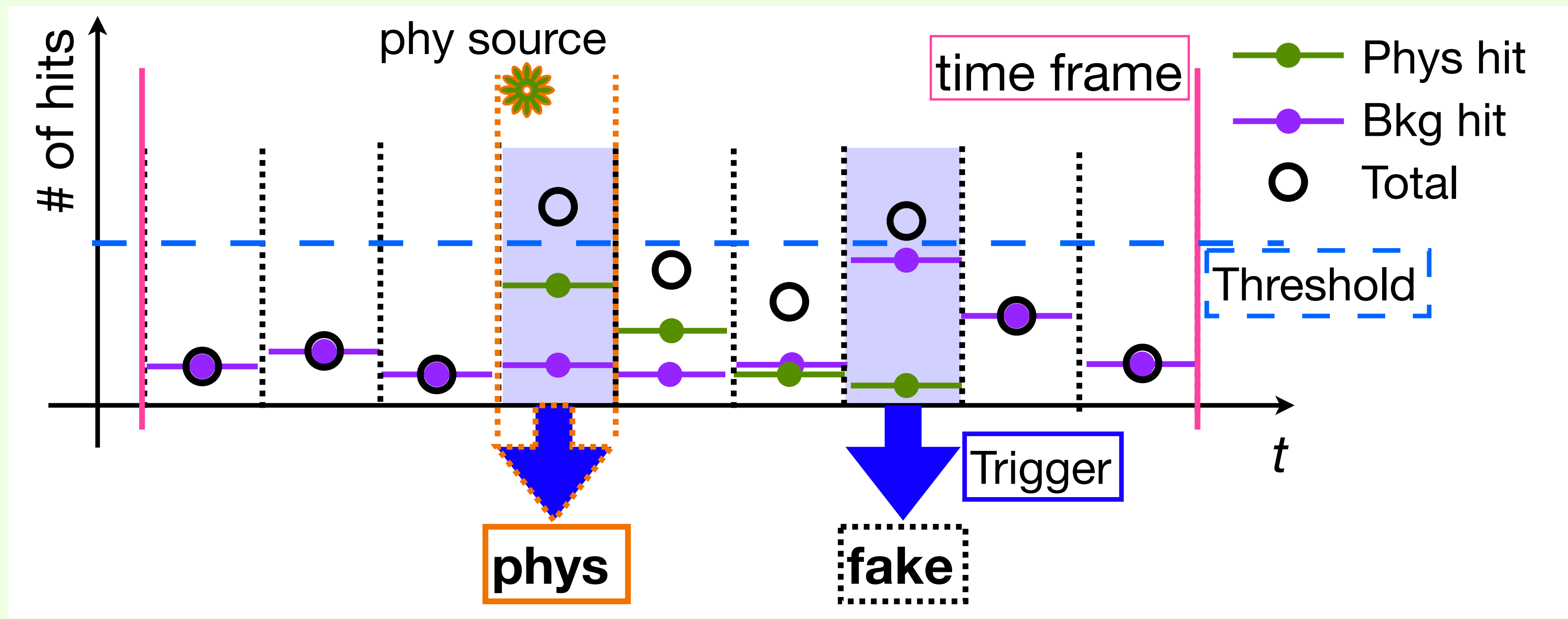
Pre-identification Efficiency Definition

The triggered time slice containing the production time of a physics

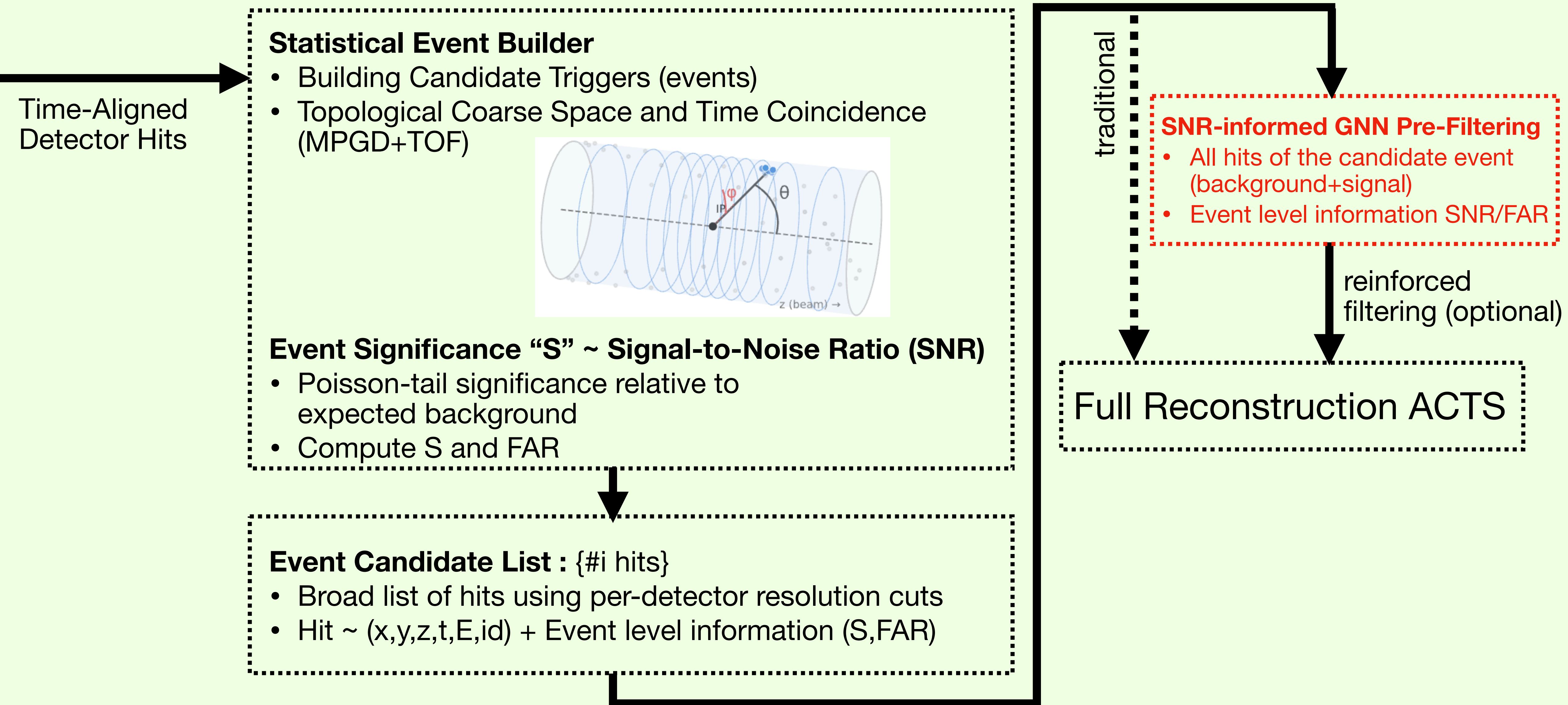
→ **✓ physics trigger**

→ **✗ fake triggers** (Even if it includes the hits which coming from the physics events)

$$\text{Efficiency (\%)} = \frac{\text{Number of triggers of physics or fake}}{\text{Number of physics events}} \times 100$$

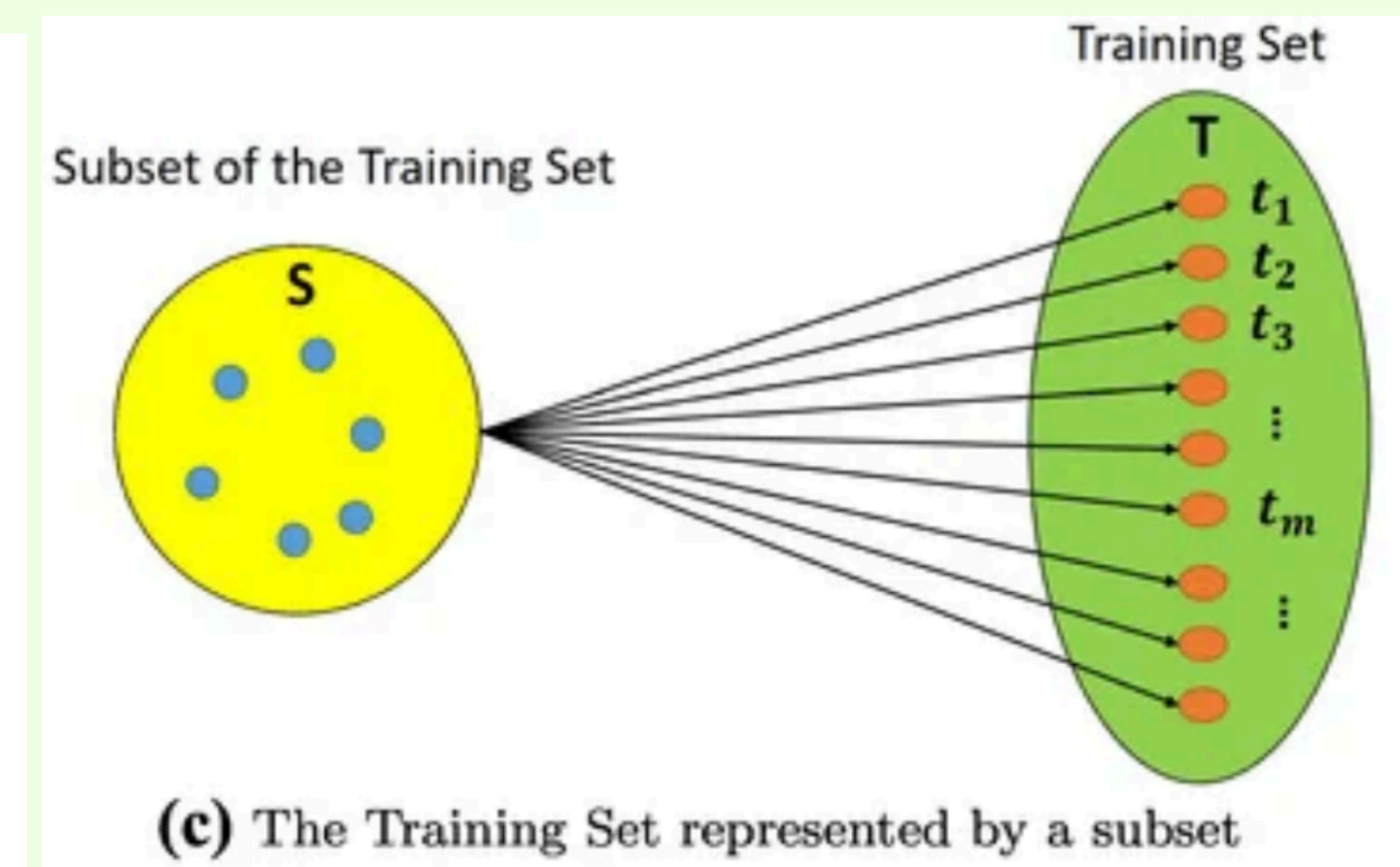
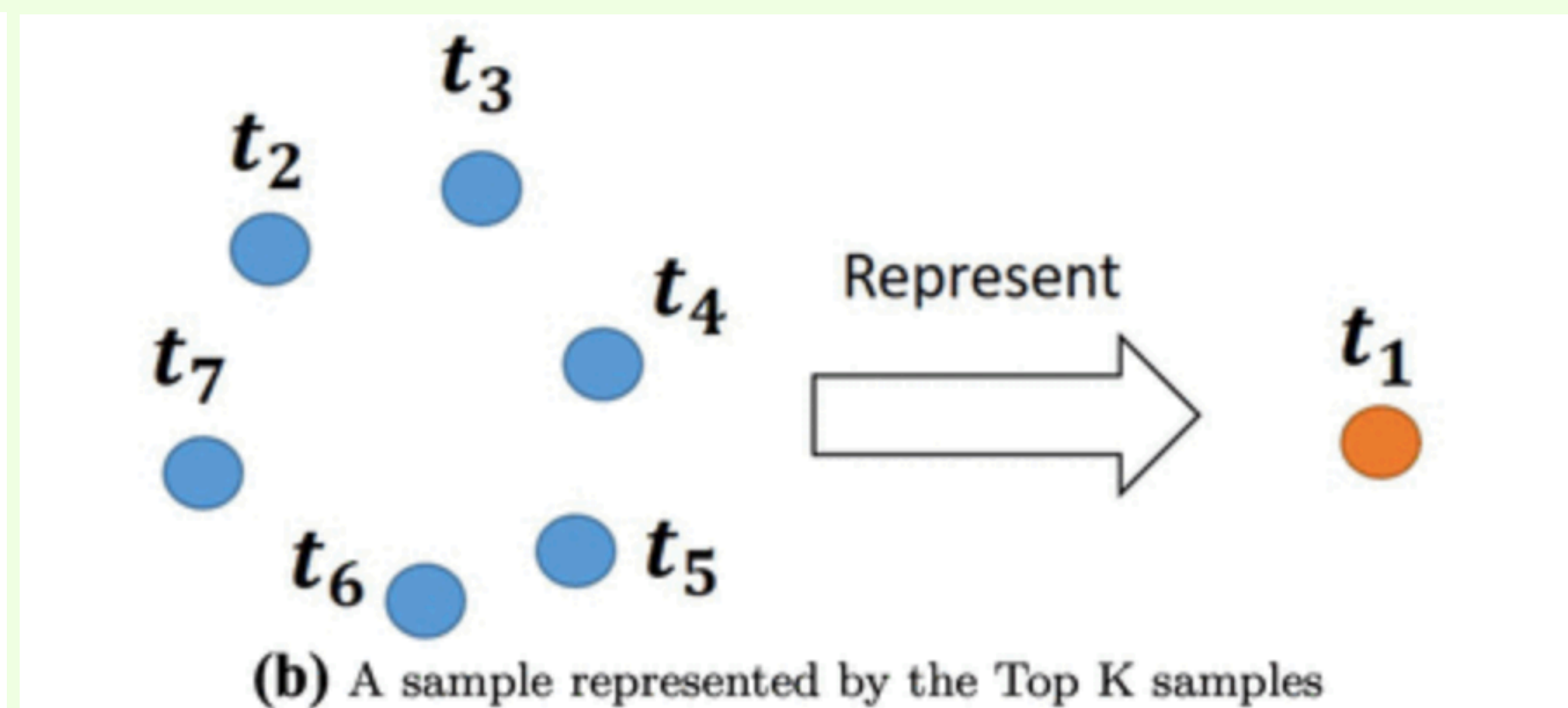
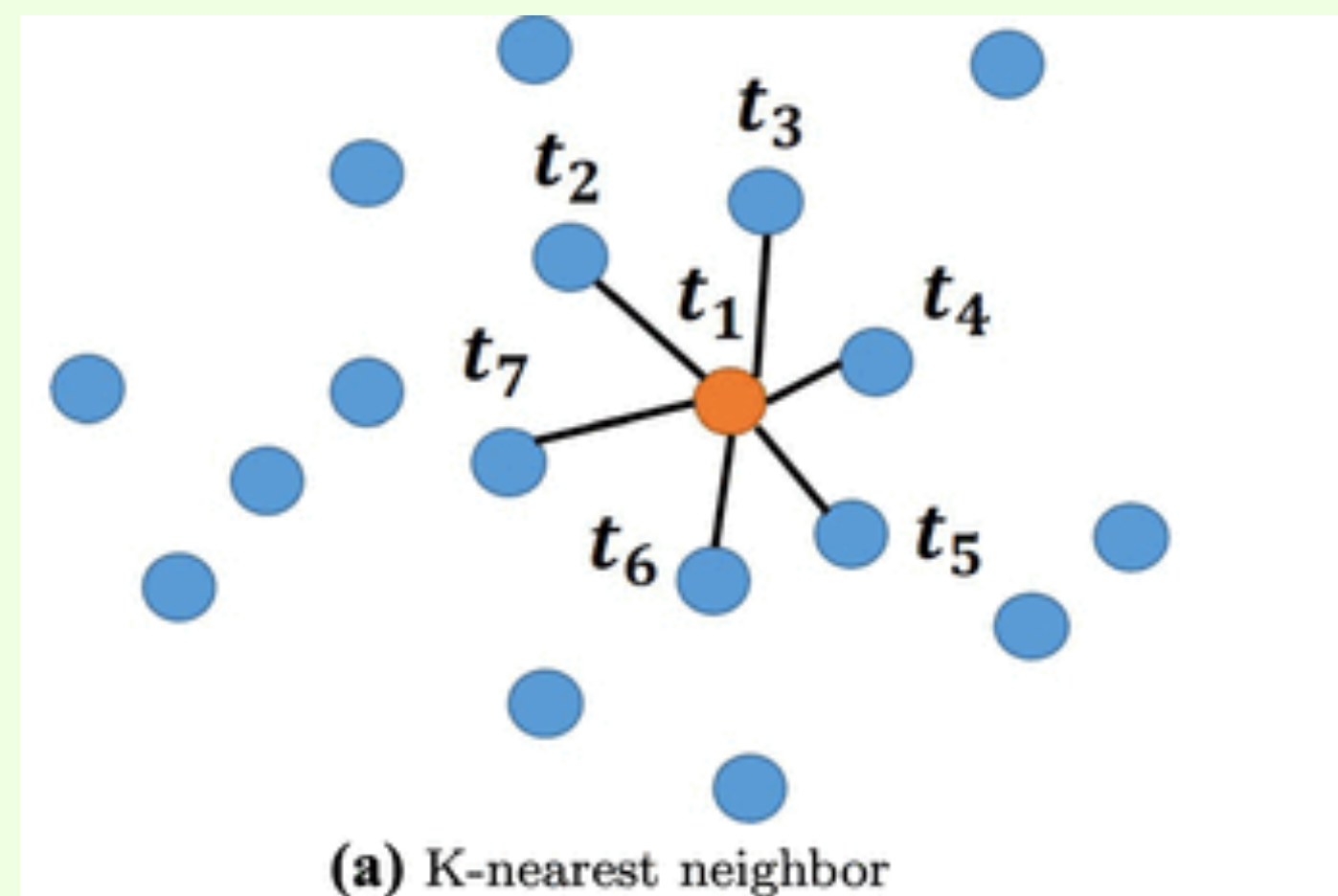


Statistical Event Builder, including SNR-informed GNN Pre-filtering



Models / Setup

GNN Layer ~ kNN (dynamic graph) + EdgeConv (static graph)



- ONNX is **only** capable of importing static graphs
- **kNN** : k-NearestNeighbors is dynamic graph (flexible number of input hits)
“nanoflann” lightweight header library: “KDTree Space + kNN algorithm” in C++

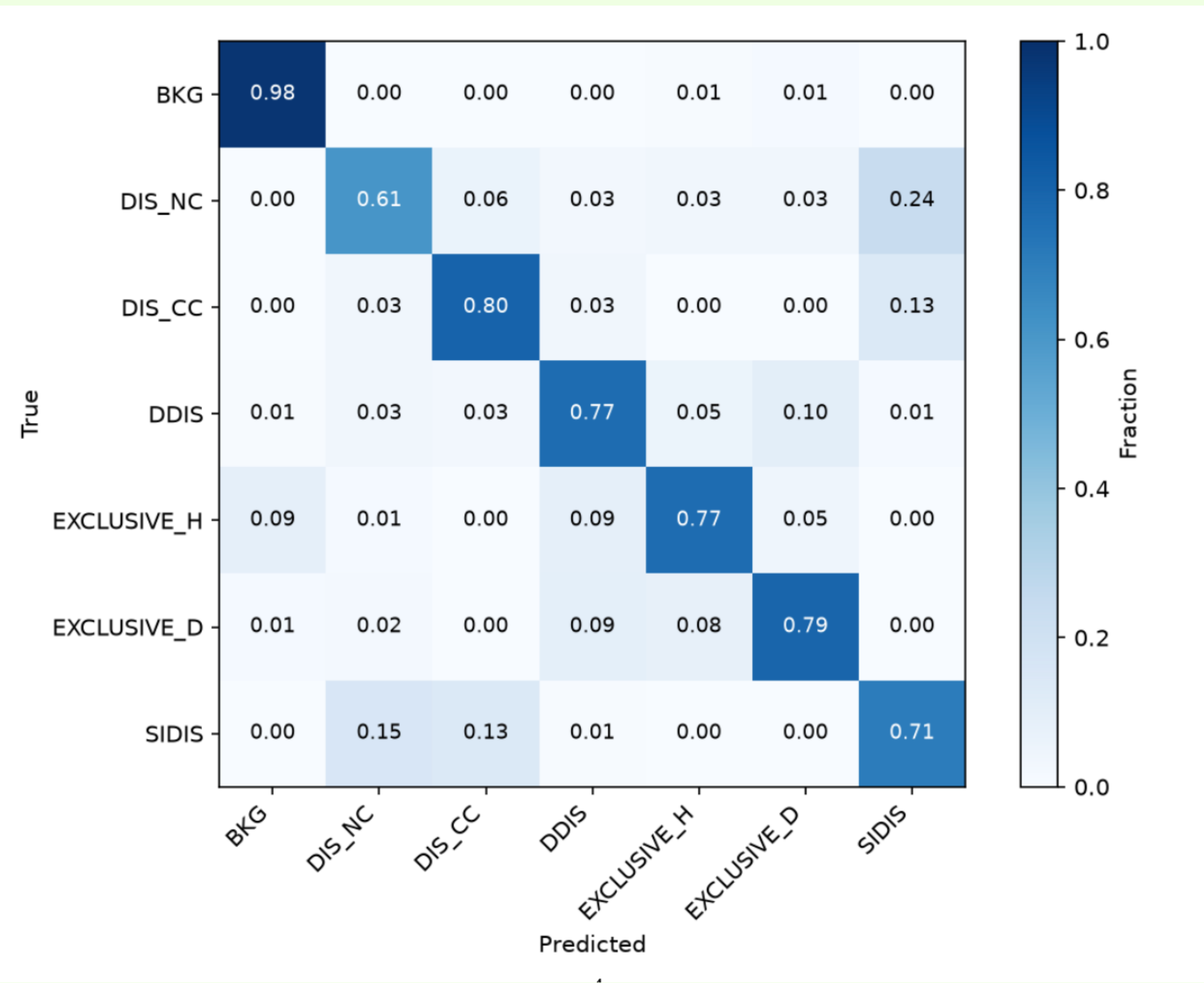


<https://github.com/jlblancoc/nanoflann>

- **EdgeConv**: hit-to-hit connection to be trained
Goal: find patterns and reject background noise

Initial performances before optimization

Confusion Matrix (initial trial)



Physics classes

#	Class	Description
0	BKG	Machine backgrounds (SynRad, Touschek, Coulomb, BeamGas, Brems)
1	DIS_NC	DIS neutral current (scattered e ⁻ visible in backward ECal)
2	DIS_CC	DIS charged current (forward jets + missing-E)
3	DDIS	Diffractive DIS (rapidity gap + intact proton in RPs)
4	EXCLUSIVE_H	Hard exclusive non-diffractive (DVCS, DVMP, DEMP, TCS)
5	EXCLUSIVE_D	Diffractive vector-meson (J/Ψ, ρ, φ; 2-body decay)
6	SIDIS	~ photoproduction for Q~[0,1]

Classes to be improved:

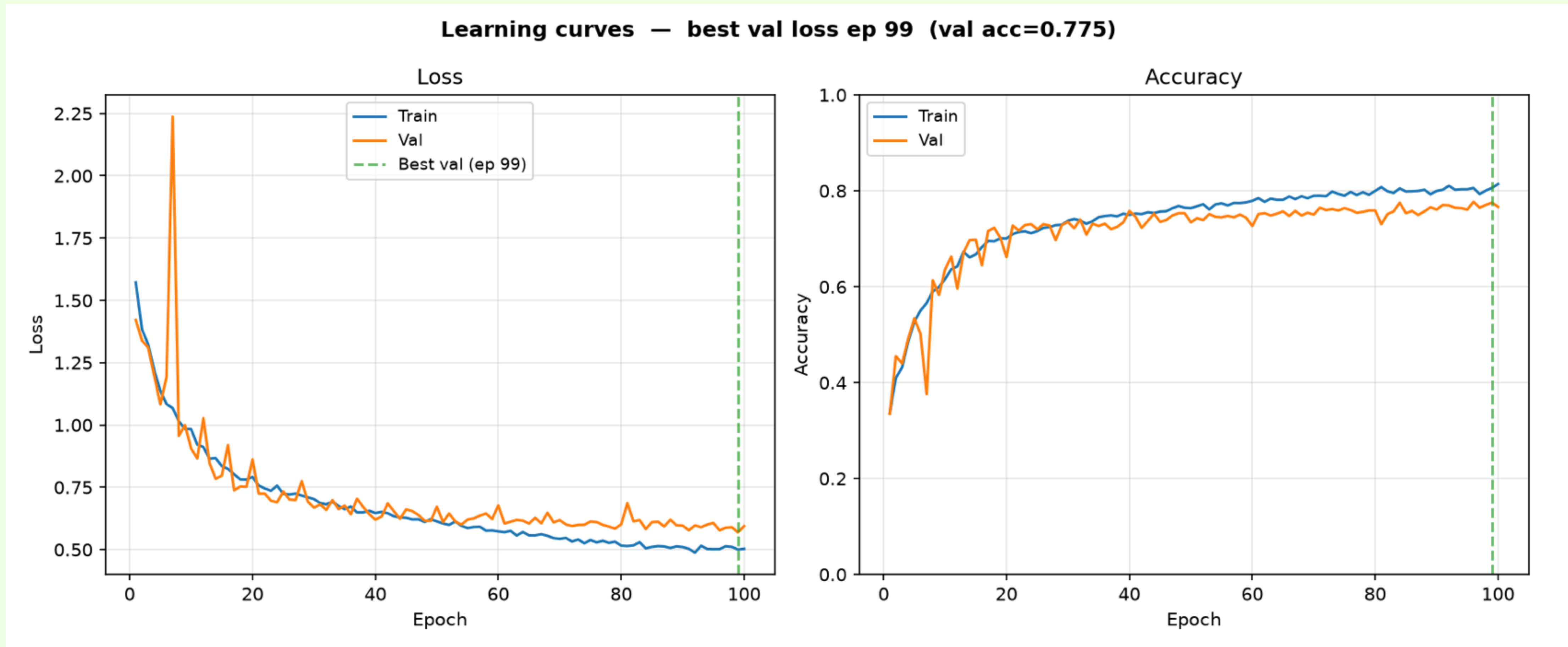
- exclusive to be separated
- sidis to be renamed

Oliver Partridge’s flash talk
(AI Town Hall 16/07 ~ 11h)

Curriculum learning to be presented tomorrow :

- learning from “no noise”
- learning step by step: reaching very high noise

Initial performances before optimization



- **Aiming at +90% accuracy** in high noise conditions (vacuum case)
- **GNN Model Output:**
 - (1) physics event classification
 - (2) pile-up event separation
 - (3) hit-map distribution (physics vs. bkg separation)

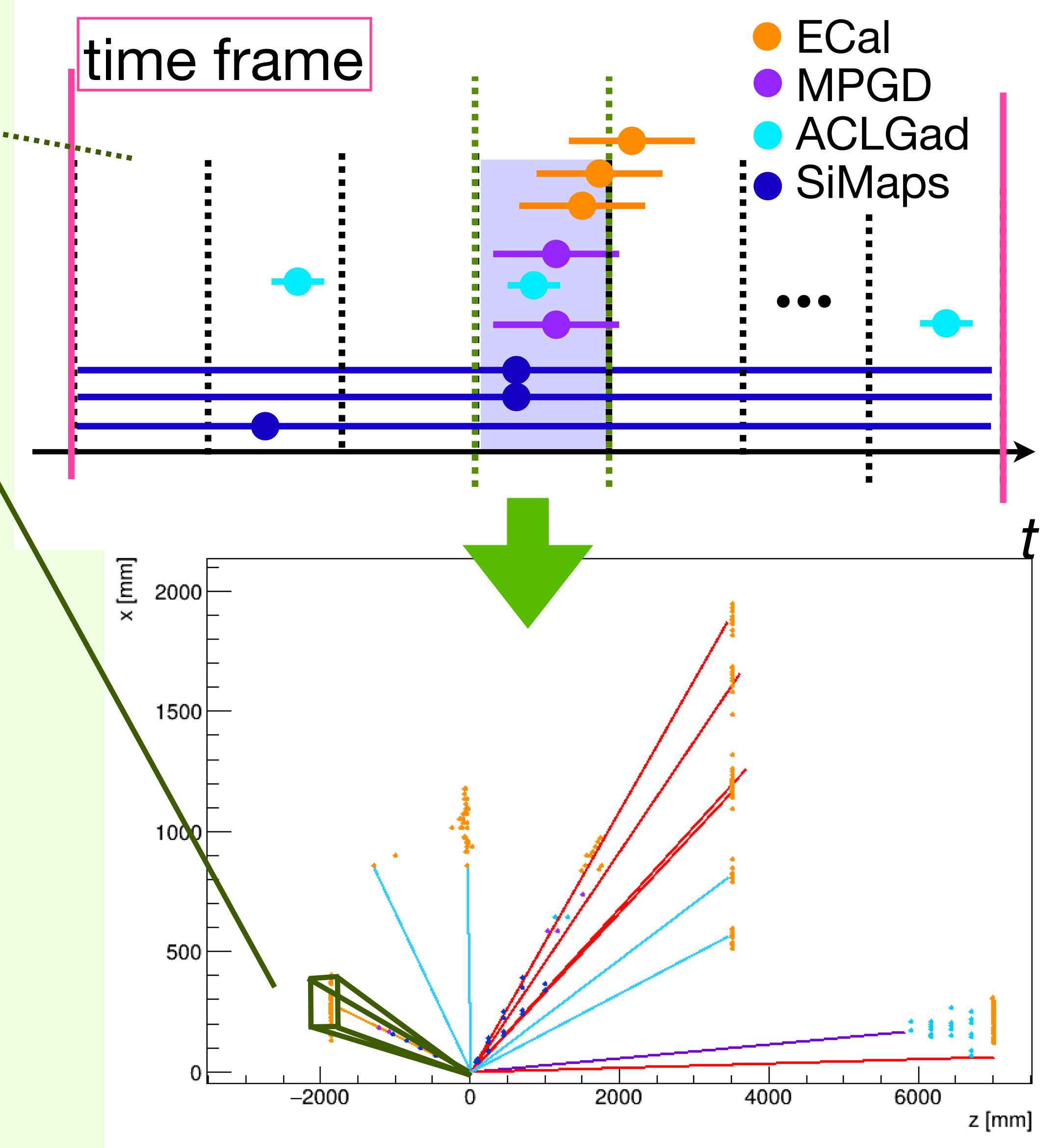
Outlook

- Implement the current algorithm into **the main branch**. (early Aug)
- Develop a **machine learning** approach for event identification.
 - PR for EICRecon (Integrated ONNX+EventBuilder). (late July)
 - Implement the event classification.
- Apply **multi-threading**. (late Aug)
- Explore **GPU** acceleration. (Oct?)
- Enhance **detector digitization** and **include detector noise**. (???)
- Validate the quality of the extracted physics events. (Check the tracking quality)

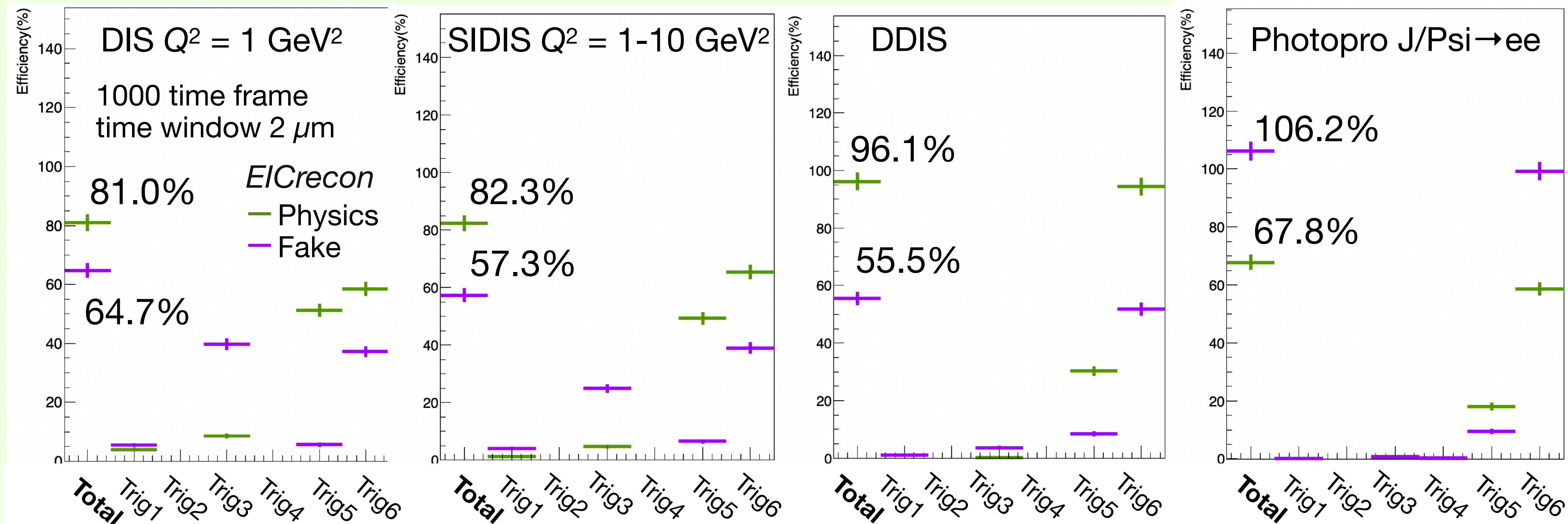
Backup Slides

New pre-event identification algorithm (ver.5)

1. Scan time bands for a time frame.
2. Count the number of hits in the each region without MPAS in the time band.
(region: $\eta \times \phi = 10 \times 10$ for barrel, end caps)
 - Backward ECal hits x10 [sTrig1] (+ track match [sTrig2])
 - Barrel ECal hits x10 [sTrig3] (+ track match [sTrig4])
 - Forward ECal hits x10 [sTrig5] (+ track match [sTrig6])
 - B0 tracks [sTrig7] x 4hits
 - ZDC ECal [sTrig8] x 50hits
3. Judge the then number of hits over the threshold.
4. Require combine the each region triggers.
 - $cTrig1 = (sTrig1 \parallel sTrig3 \parallel sTrig5) > 1 \ \&\& \ sTrig7$
 - $cTrig2 = (sTrig2 \parallel sTrig4 \parallel sTrig6) > 1 \ \&\& \ sTrig7$
 - $cTrig3 = (sTrig1 \parallel sTrig3 \parallel sTrig5) > 1 \ \&\& \ sTrig8$
 - $cTrig4 = (sTrig2 \parallel sTrig4 \parallel sTrig6) > 1 \ \&\& \ sTrig8$
 - $cTrig5 = (sTrig1 \parallel sTrig3 \parallel sTrig5) > 3$
 - $cTrig6 = (sTrig2 \parallel sTrig4 \parallel sTrig6) > 2$
5. If over the threshold, trigger the time region.



Trigger Efficiency (Blind Analysis)



Trigger 1: $(\text{ECal} > 5 + \text{Trk} > 0) > 0 \ \&\& \ B0 > 4$

Trigger 2: $(\text{ECal} > 5 + \text{Trk} > 0) > 0 \ \&\& \ ZDC > 4$

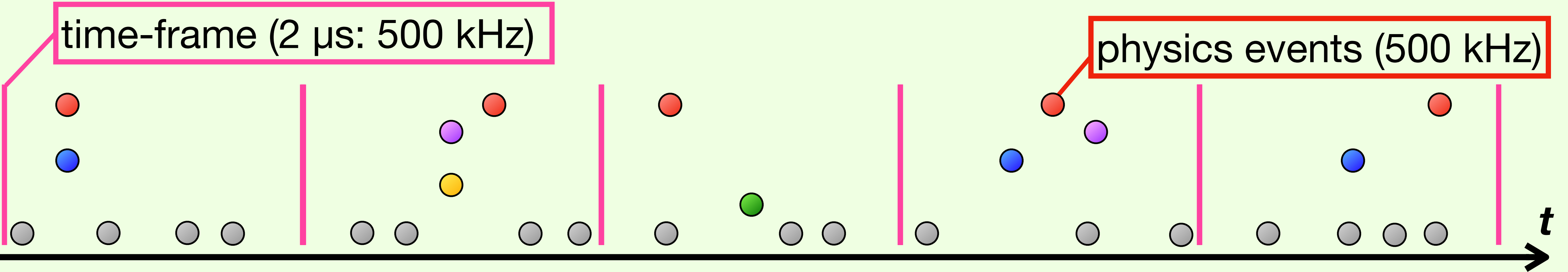
Trigger 3: $(\text{ECal} > 10) > 0 \ \&\& \ B0 > 4$

Trigger 4: $(\text{ECal} > 10) > 0 \ \&\& \ ZDC > 4$

Trigger 5: $(\text{ECal} > 5 + \text{Trk} > 0) > 1$

Trigger 6: $(\text{ECal} > 10) > 2$

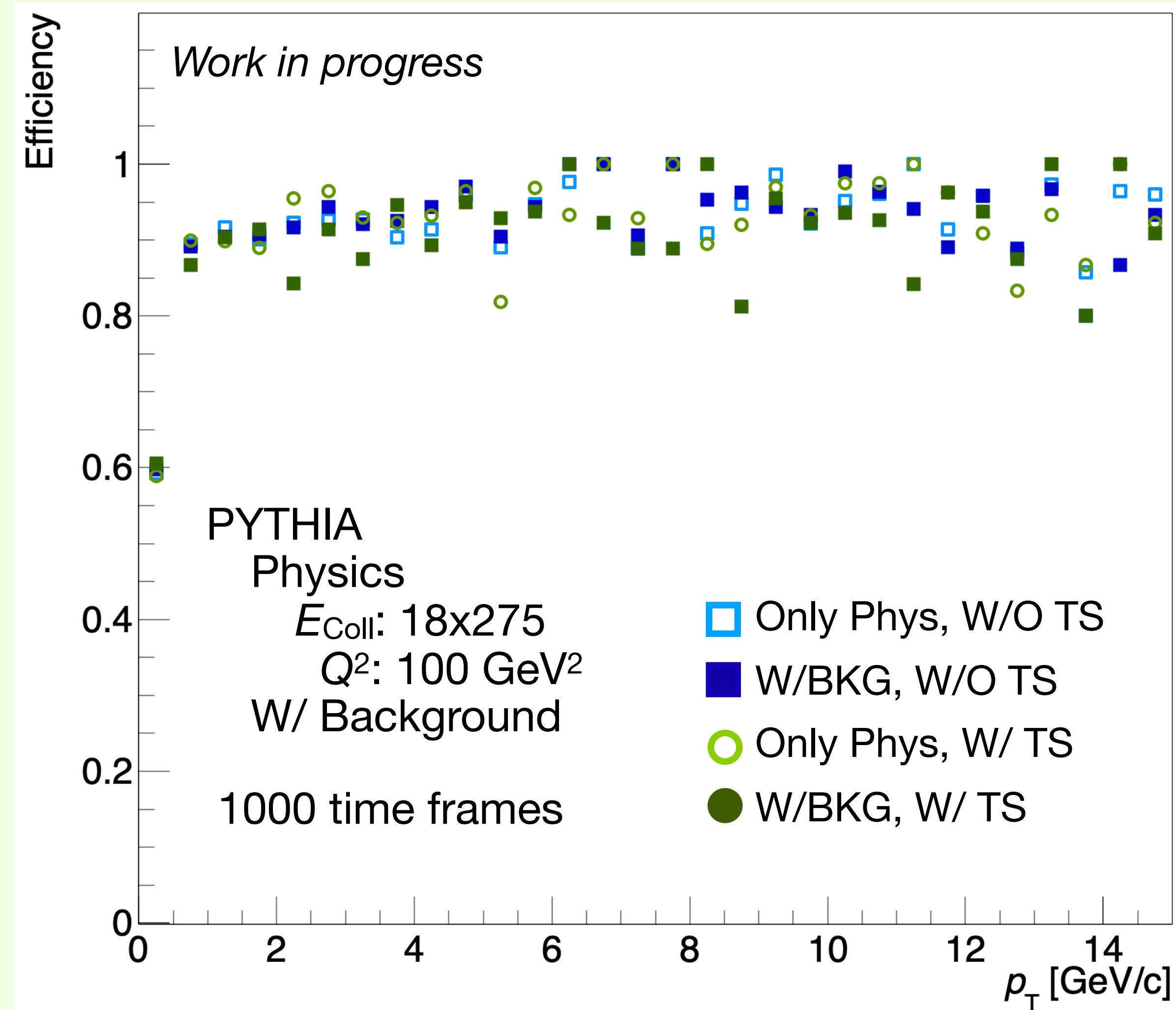
- It could reduce the fake trigger rates, but the physics event efficiency was not enough.
- Not yet optimize the trigger parameters
- Search for what types of event missed and recover these events.



rates in kHz	Old	New	Vacuum
● Total ep	500 kHz	500 kHz	
● hadron beam gas	22.5 kHz	31.9 kHz	10000Ahr
		342.8 kHz	100Ahr
● e beam gas (Bremsstrahlung scattering)	317 kHz	3177.3 kHz	10000Ahr
● e beam gas (Coulomb losses - w/ collimators)	0.72 kHz	29.56 kHz	10000Ahr
● e intrabeam (Touschek losses - w/ collimators)	1.3 kHz	233.5 kHz	
● electron SR	14 MHz	36.6 GHz	

Tracking performance in the triggered events

- The **efficiency** of W/ time slice were mostly consistent with W/O ones.
- The **purities** of W/ time slice W/O ones.
- **Track/vertex χ^2** and track p_T become a little better.
- **Process time** become faster using the time splitting (**-10%** / 1000 time frames (for large number of hits events: **-50%**))



(※ This results are **only for $Q^2=100 \text{ GeV}^2$** , and background is previous version (less))

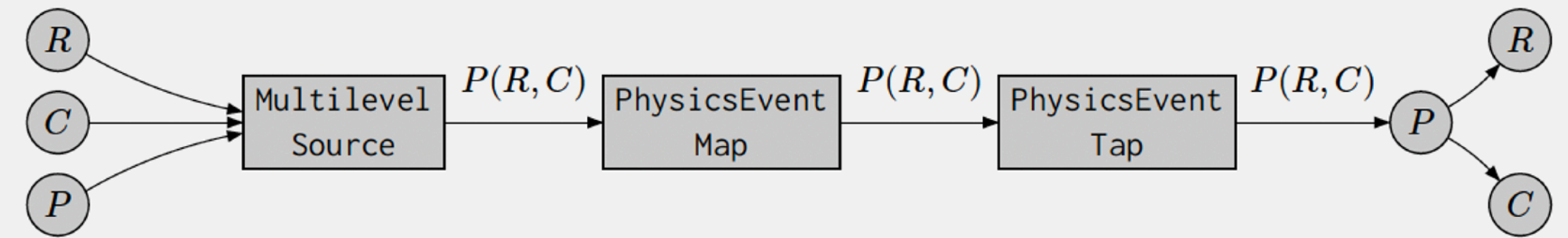
ElCrecon framework

- Base framework: JANA2

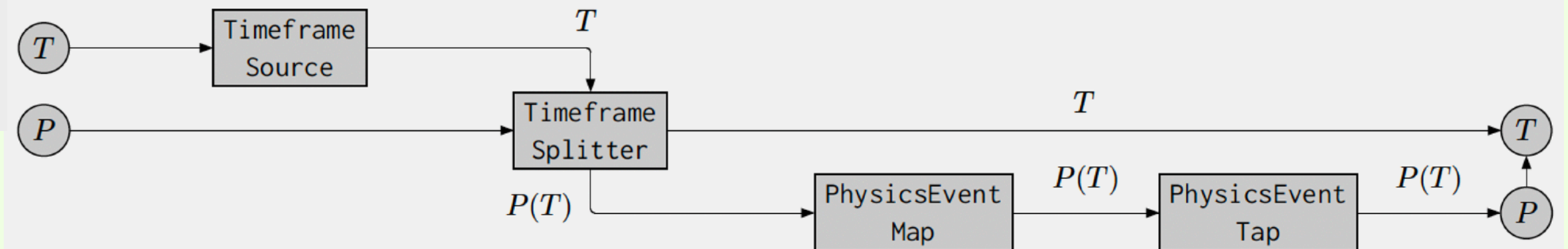
General purpose software framework for future experiments developed by Jefferson Lab (JLab) (used by GlueX experiment)

Multi-threaded JANA2 framework provides a component-level hierarchical decomposition of data boundaries into Run, Timeframe, PhysicsEvent, and Subevent levels.
→ Essential for streaming processing

Introducing multilevel sources



EICrecon timeframe splitting



- Data Model: PODIO (Plain Old Data Input/Output)

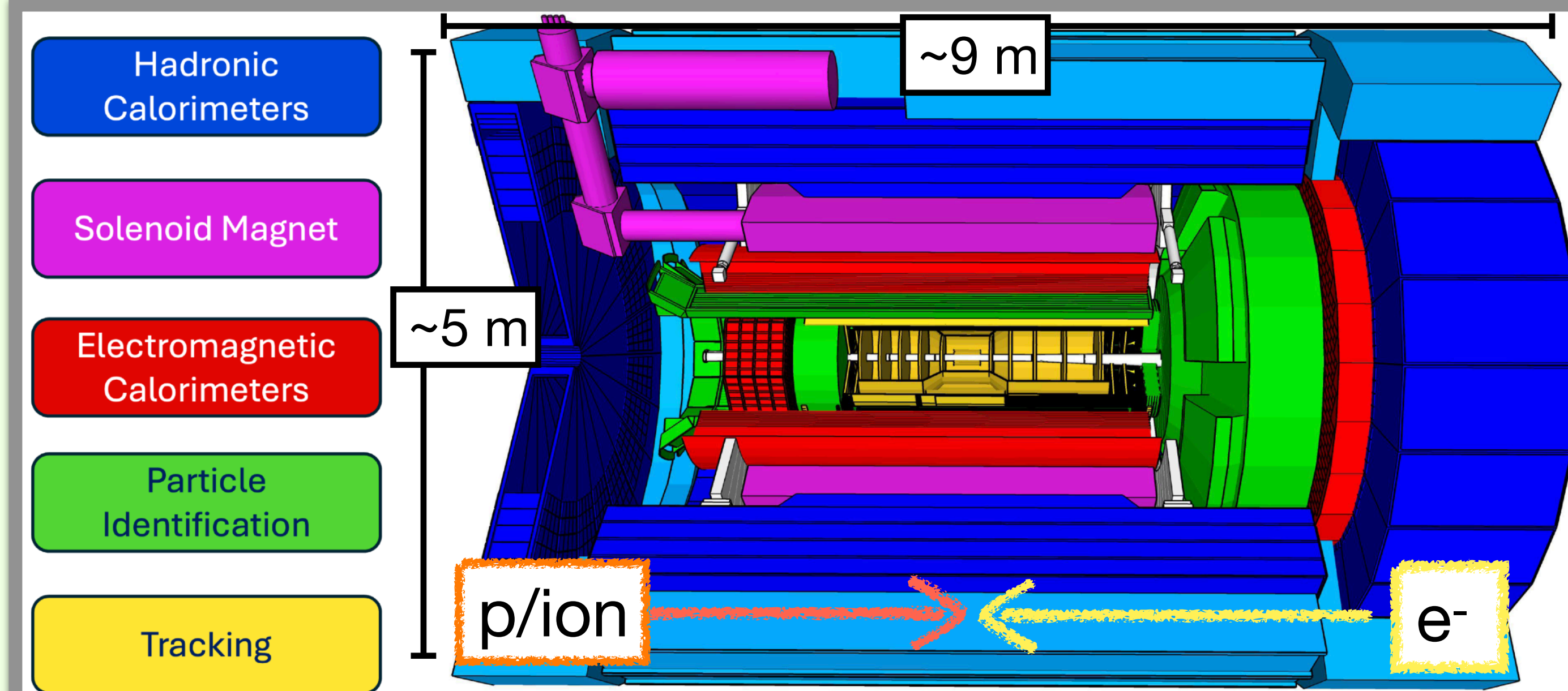
API data driven model developed by CERN

Based on yaml: edm4eic ([git](#)) ← edm4hep ([git](#)) ← podio ([page](#))

- Tracking: ACTS (A Common Tracking Software)

A modular, experiment-independent tracking library developed at CERN

ePIC (electron-Proton/Ion Collider) detector

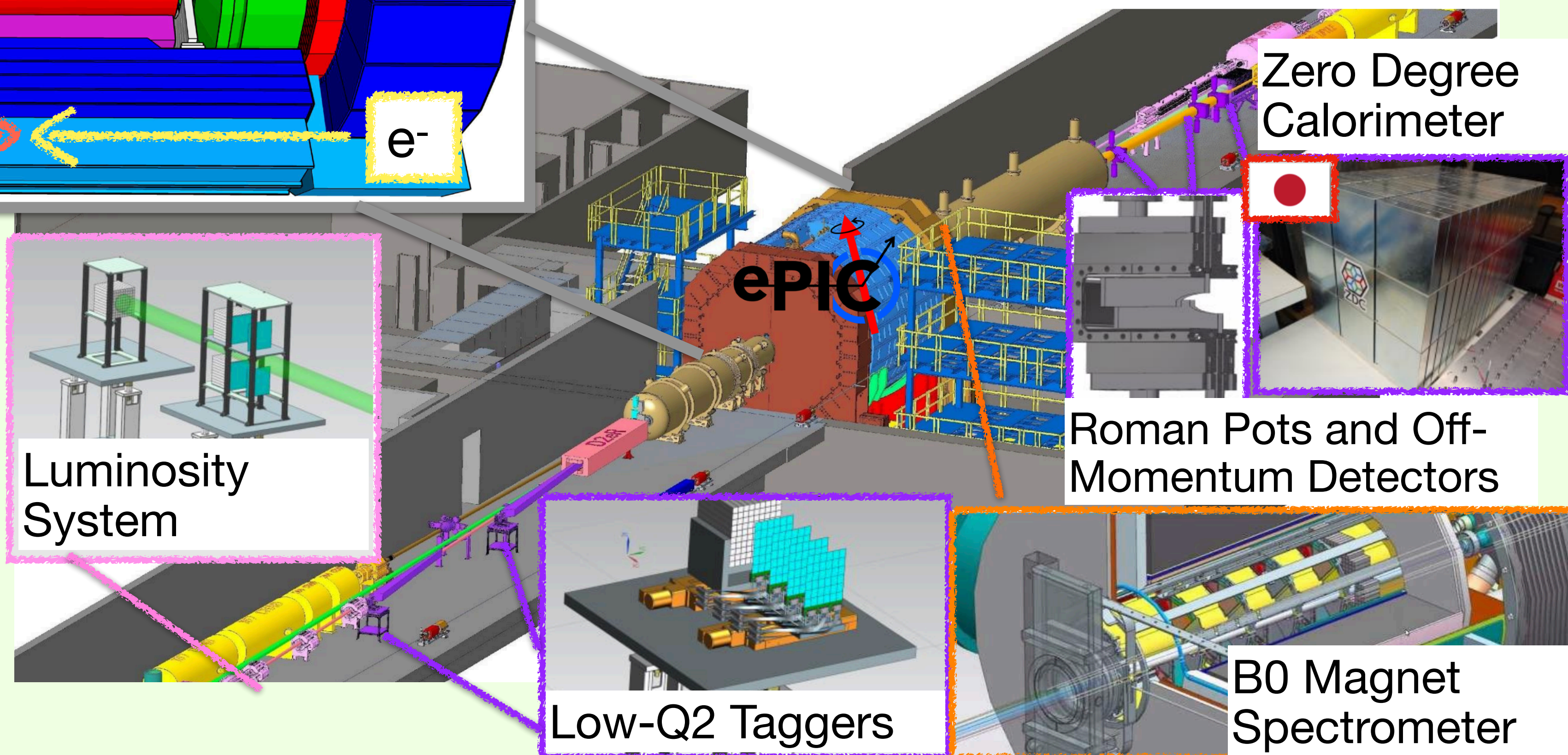


Calorimetry:

- Imaging Barrel EMCal
- PbWO₄ EMCal in backward direction
- Finely segment EMCal+HCal in forward direction
- Outer HCal (sPHENIX re-use)
- Backwards HCal (tail-catcher)

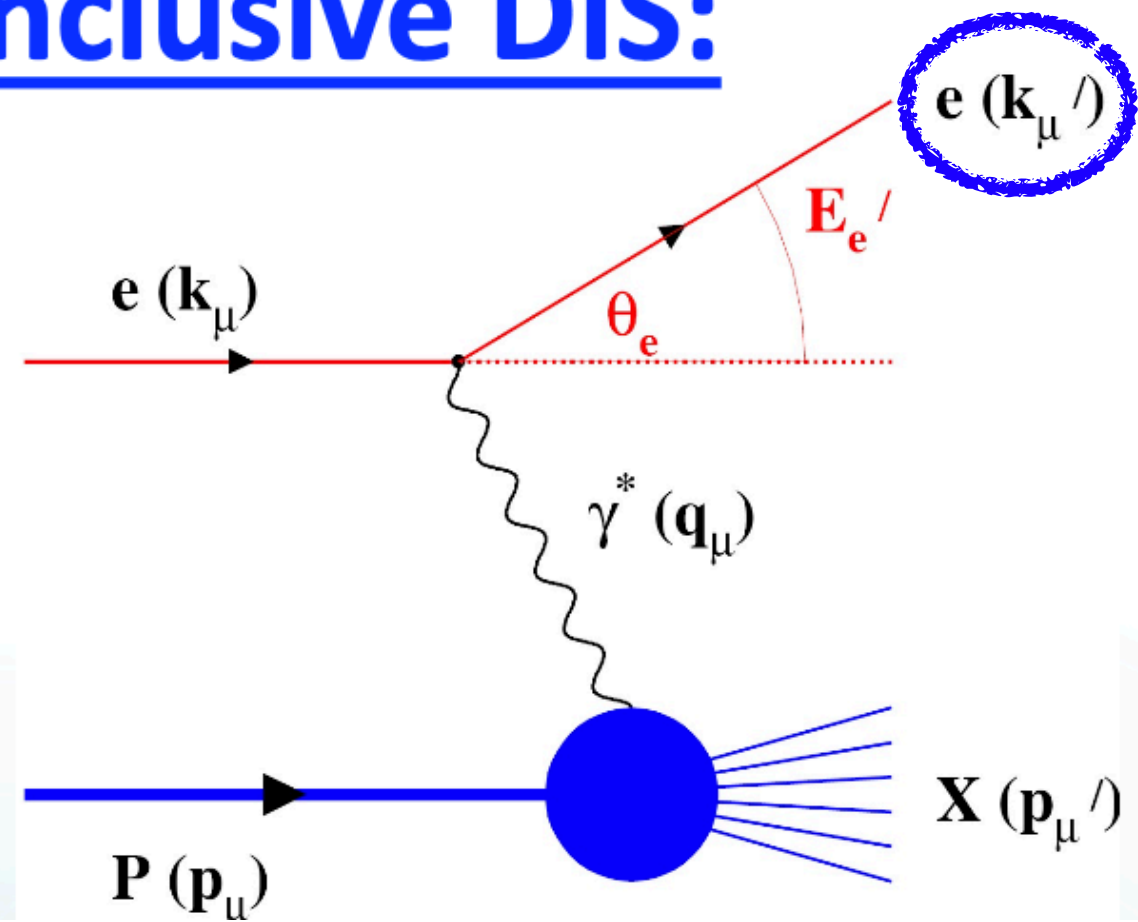
- ## Tracking:
- New 1.7 T solenoid
 - Si MAPS tracker
 - MPGDs (μ RWELL/ μ Megas)

- ## PID:
- hpDIRC
 - pfRICH
 - dRICH
 - AC-LGAD (~30ps TOF) 



ePIC Detector Concept

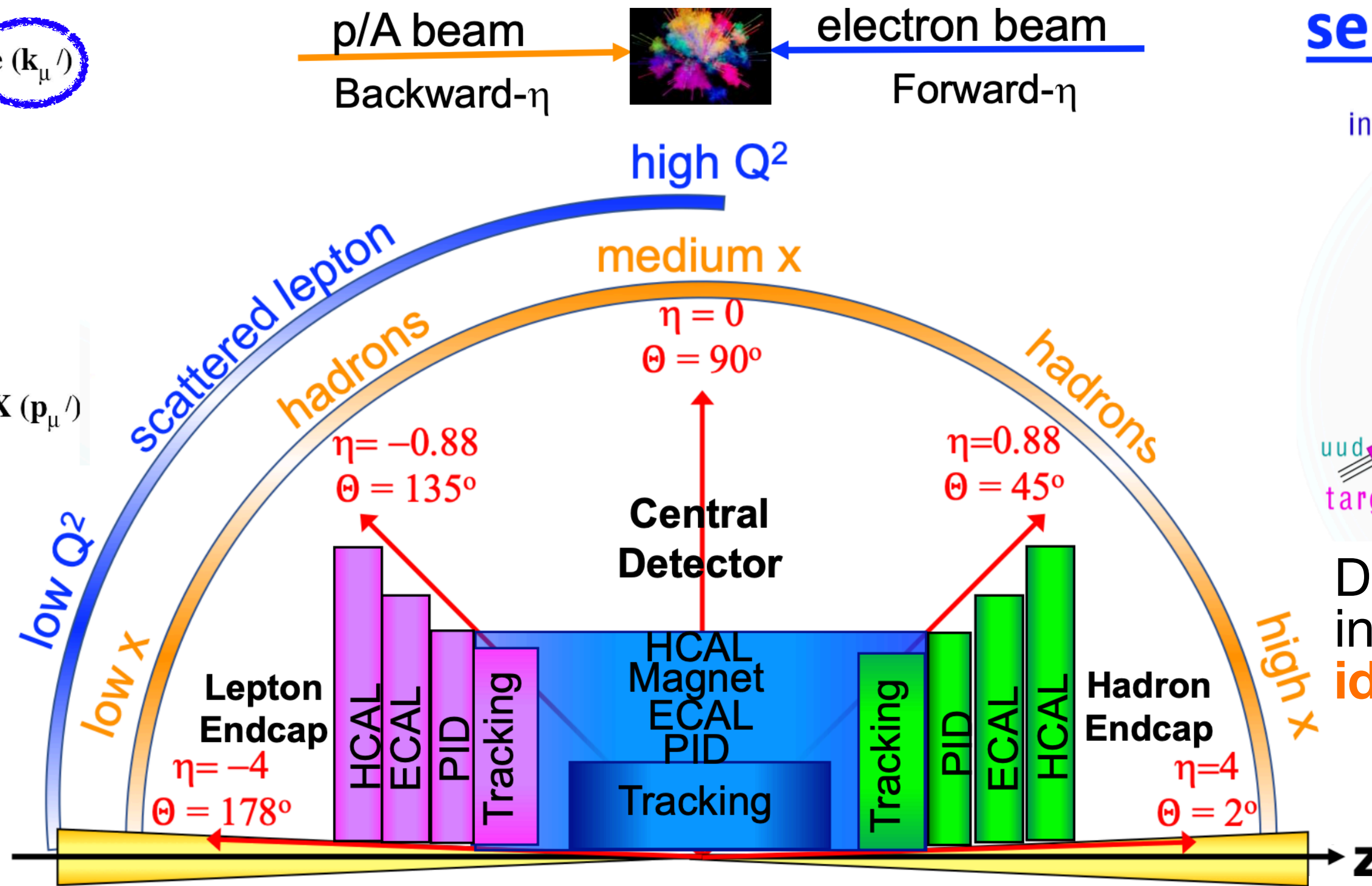
inclusive DIS:



Detect only the **scattered lepton** in the detector

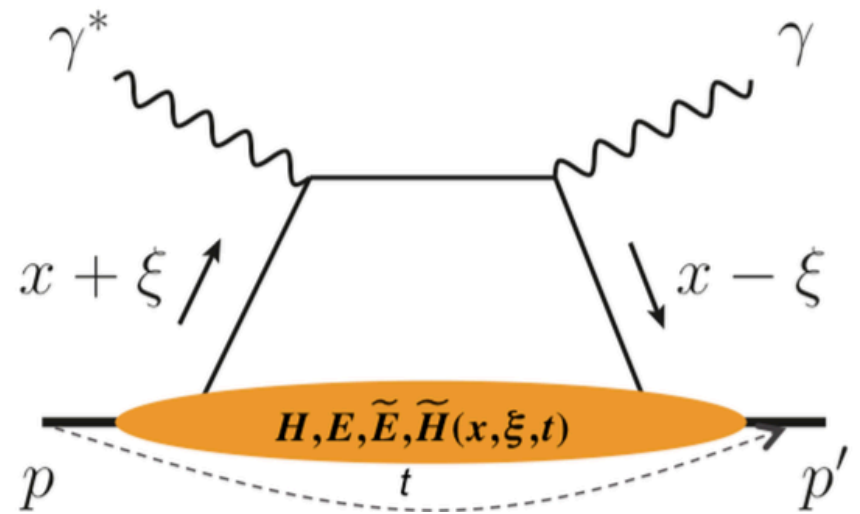
Very low Q^2
scattered lepton
Bethe-Heitler
photons for
luminosity

Luminosity Detector Low Q^2 -Tagger



exclusive DIS

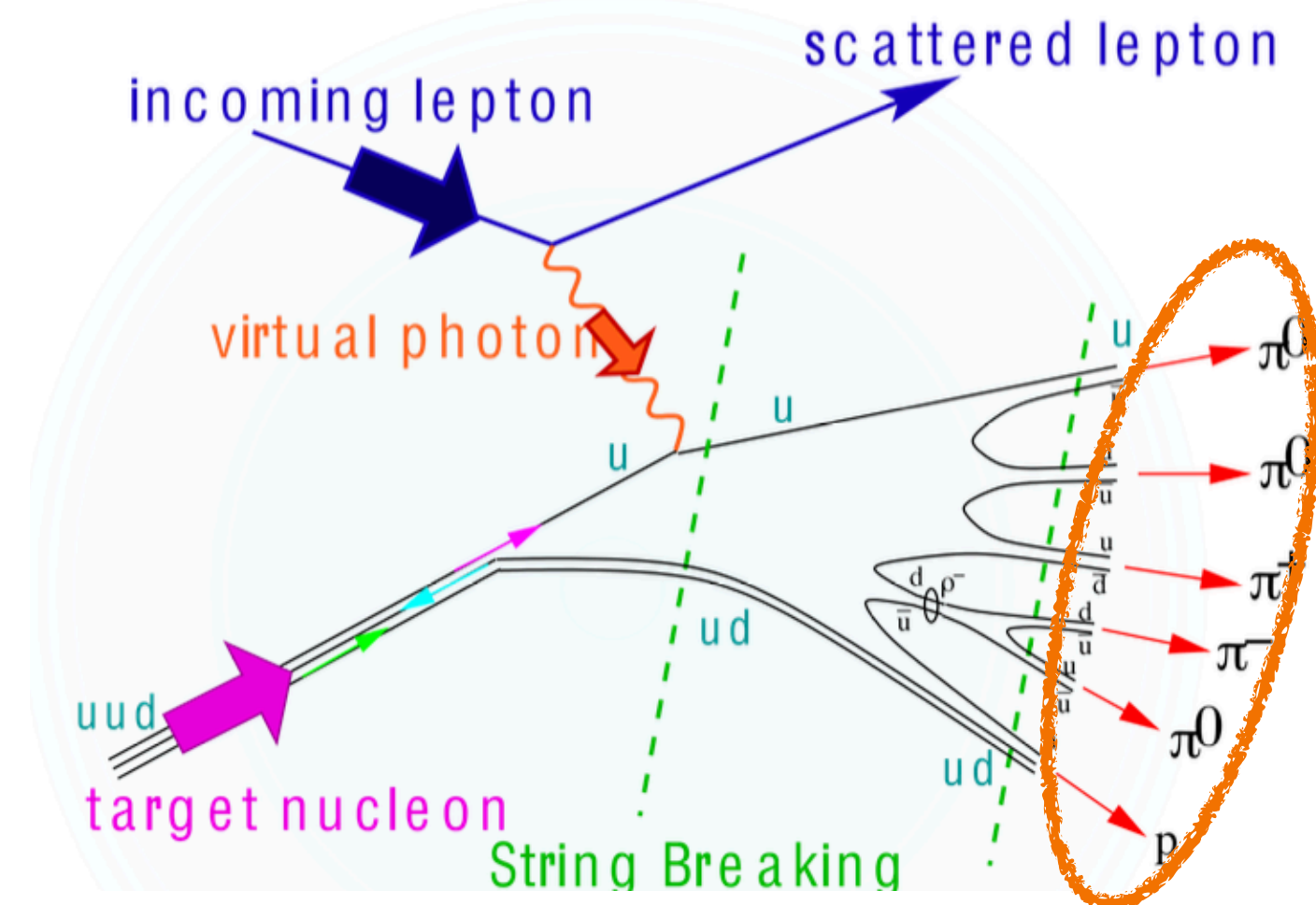
Detect the scattered lepton,
identify produced hadron/
jets and **target remnants**



ZDC

Forward Tracking

semi-inclusive DIS

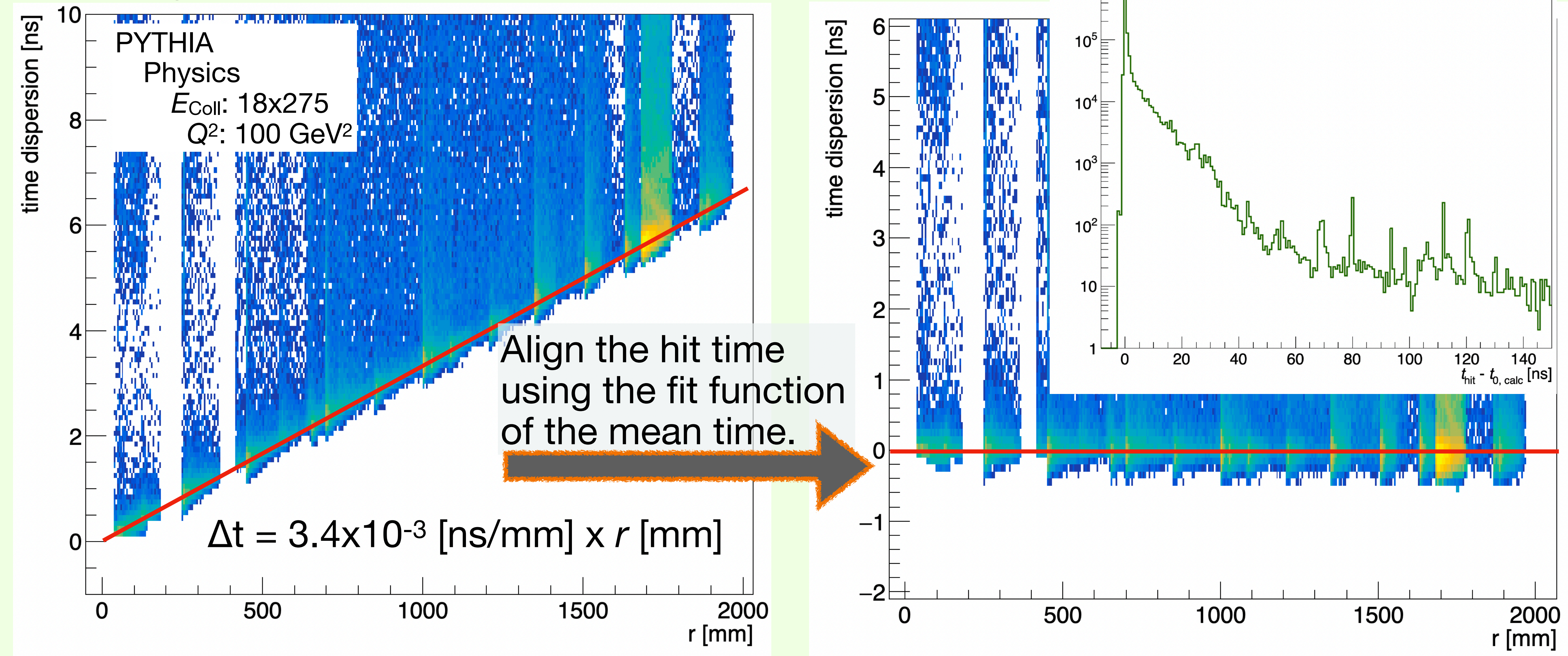


Detect the scattered lepton
in coincidence with
identified hadrons/jets

Particles from nuclear breakup and from diffractive reactions

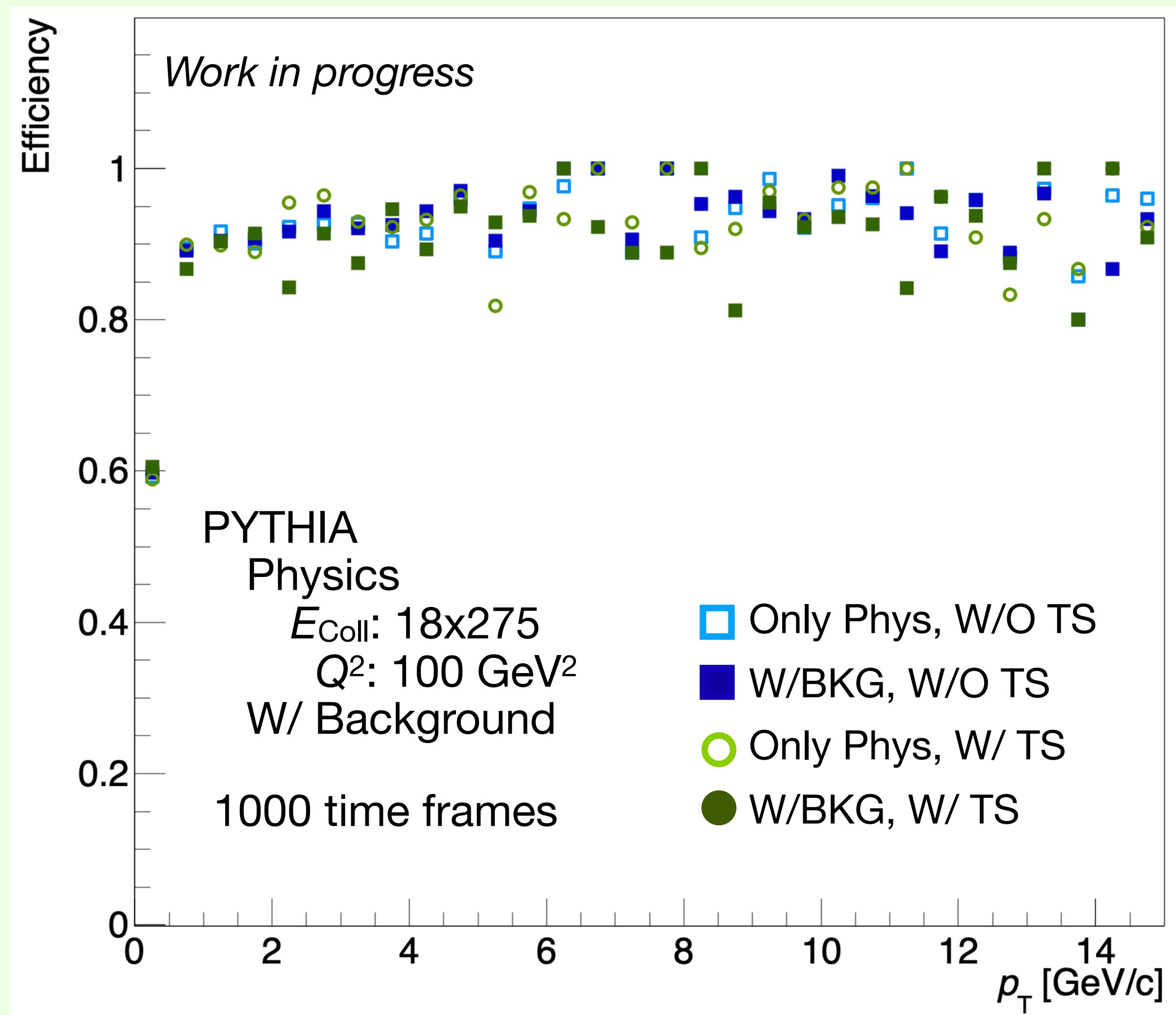
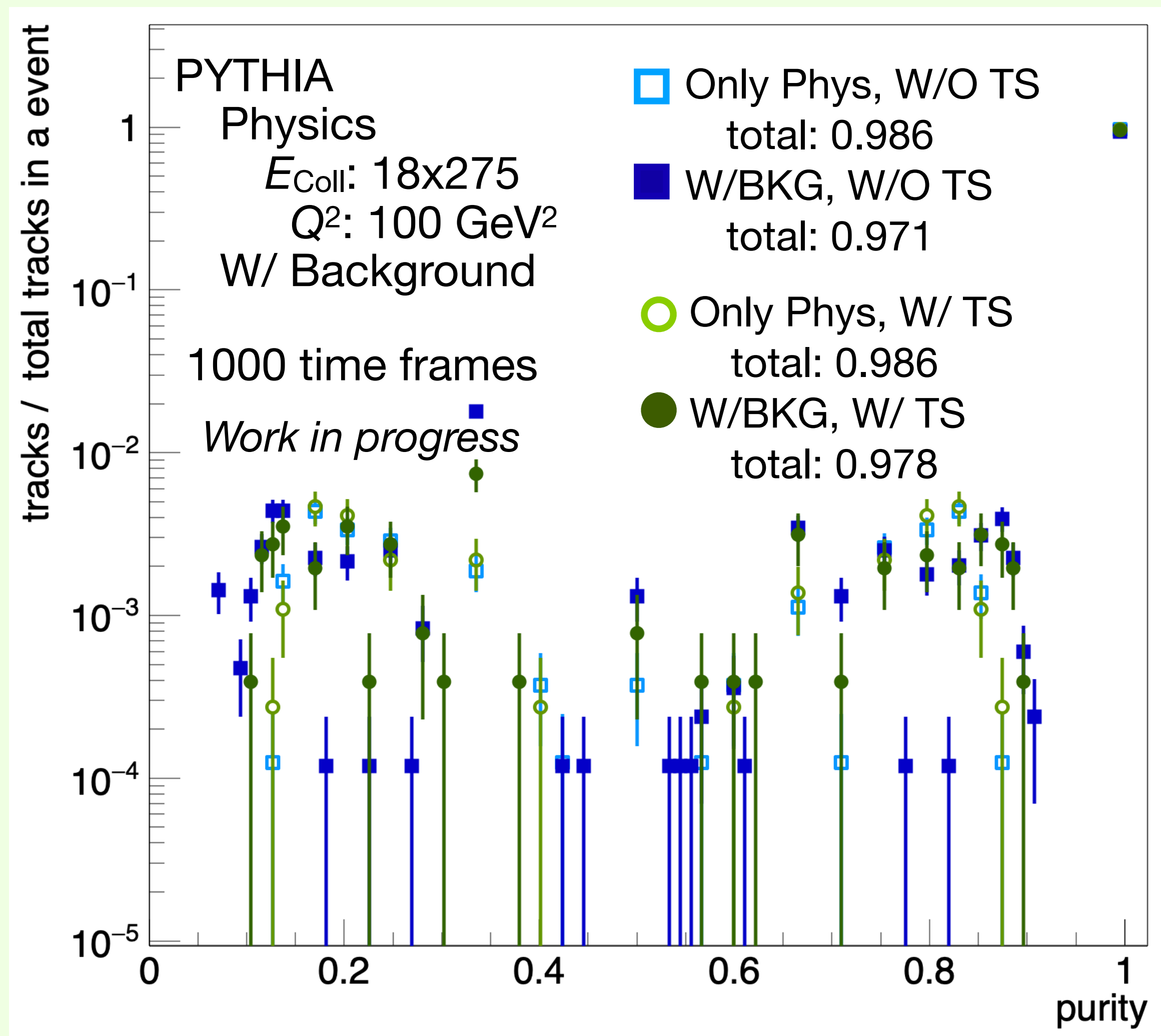
Hit time dispersion of DIS events

Time alignment of DIS events



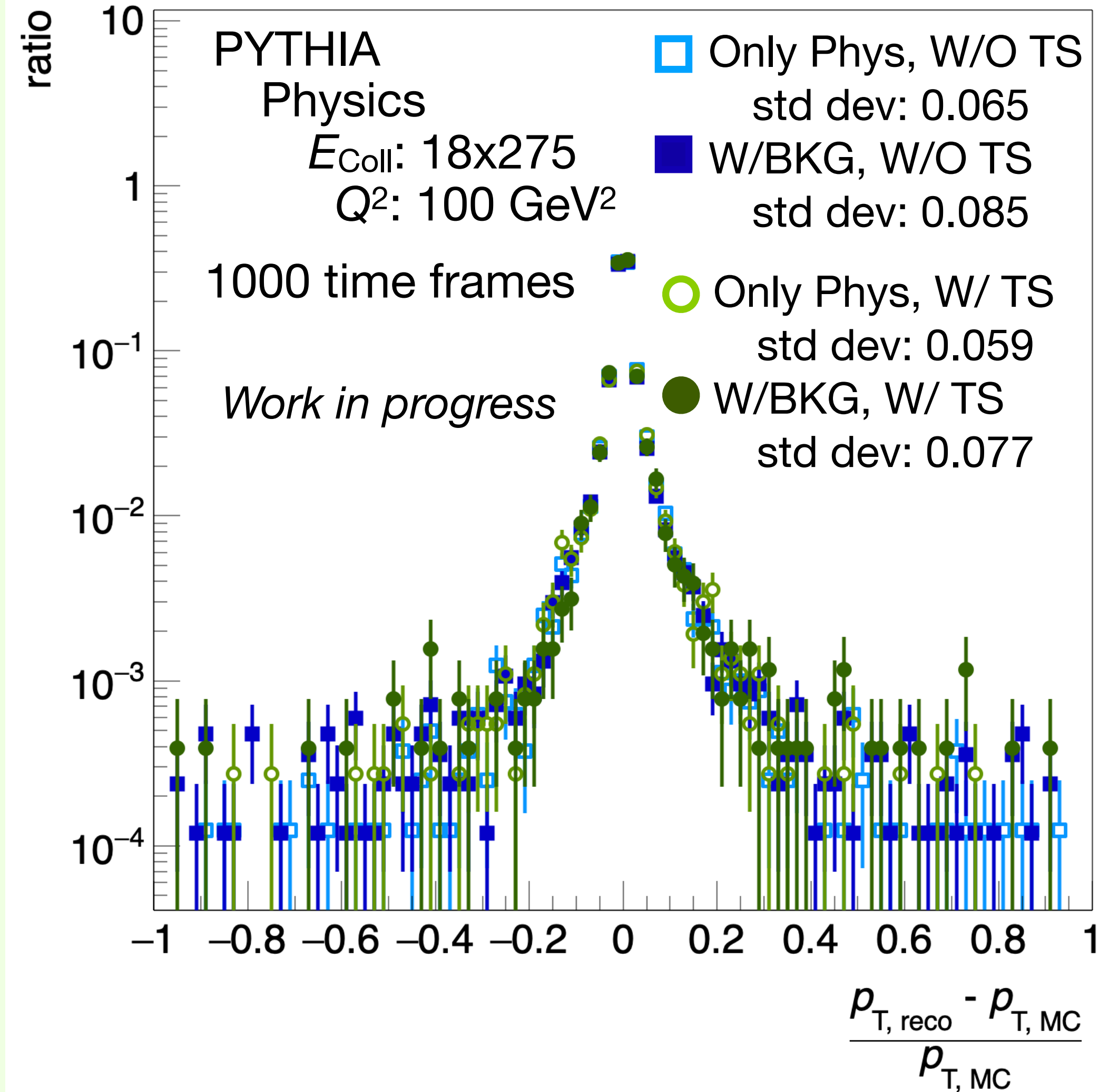
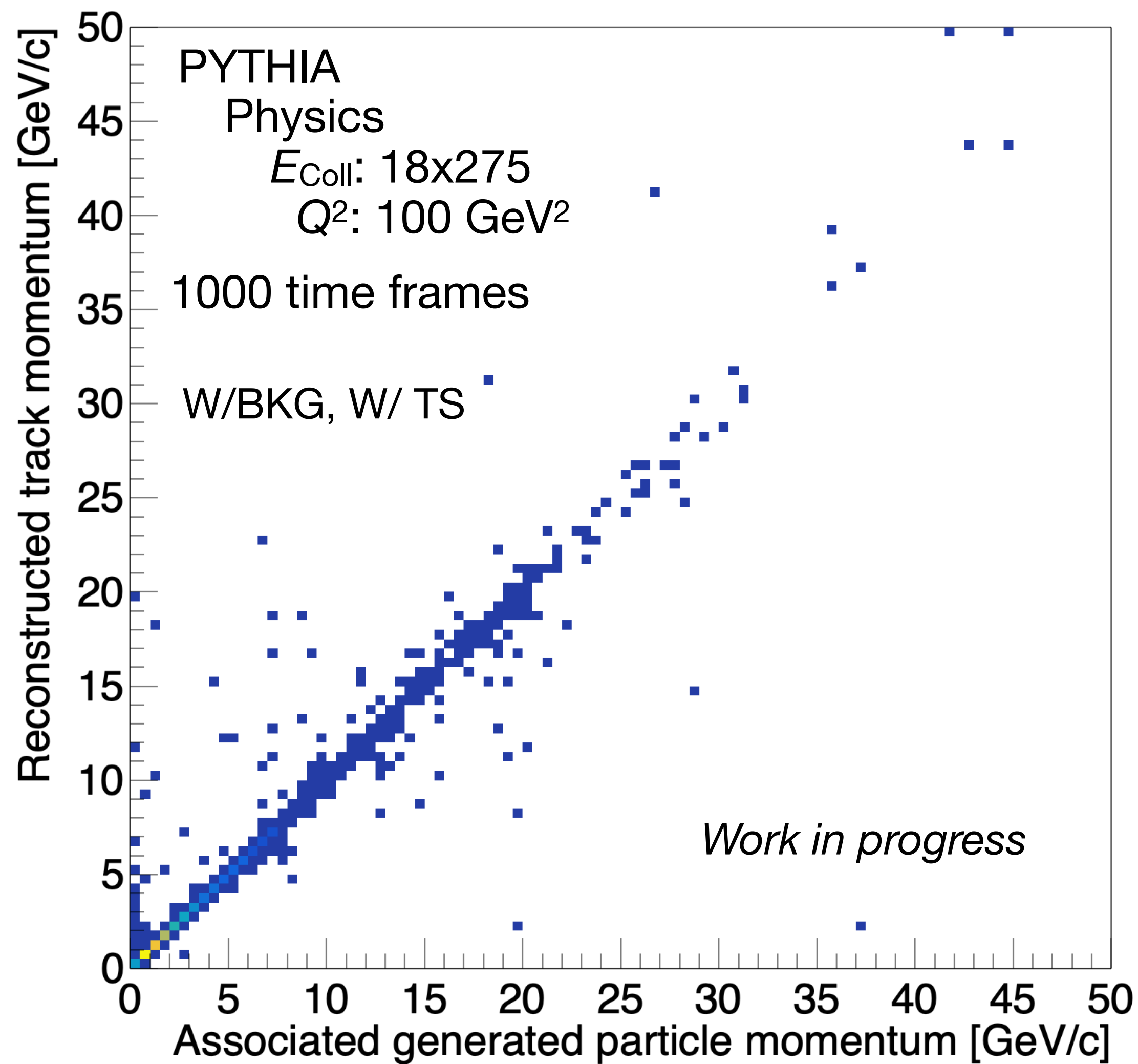
Fit the mean hit time for each region.

Track Purity and Efficiency Evaluation



The purities of W/ time slice were a **little better** than W/O ones.
The efficiency of W/ time slice were mostly consistent with W/O ones, but a **little worse**.

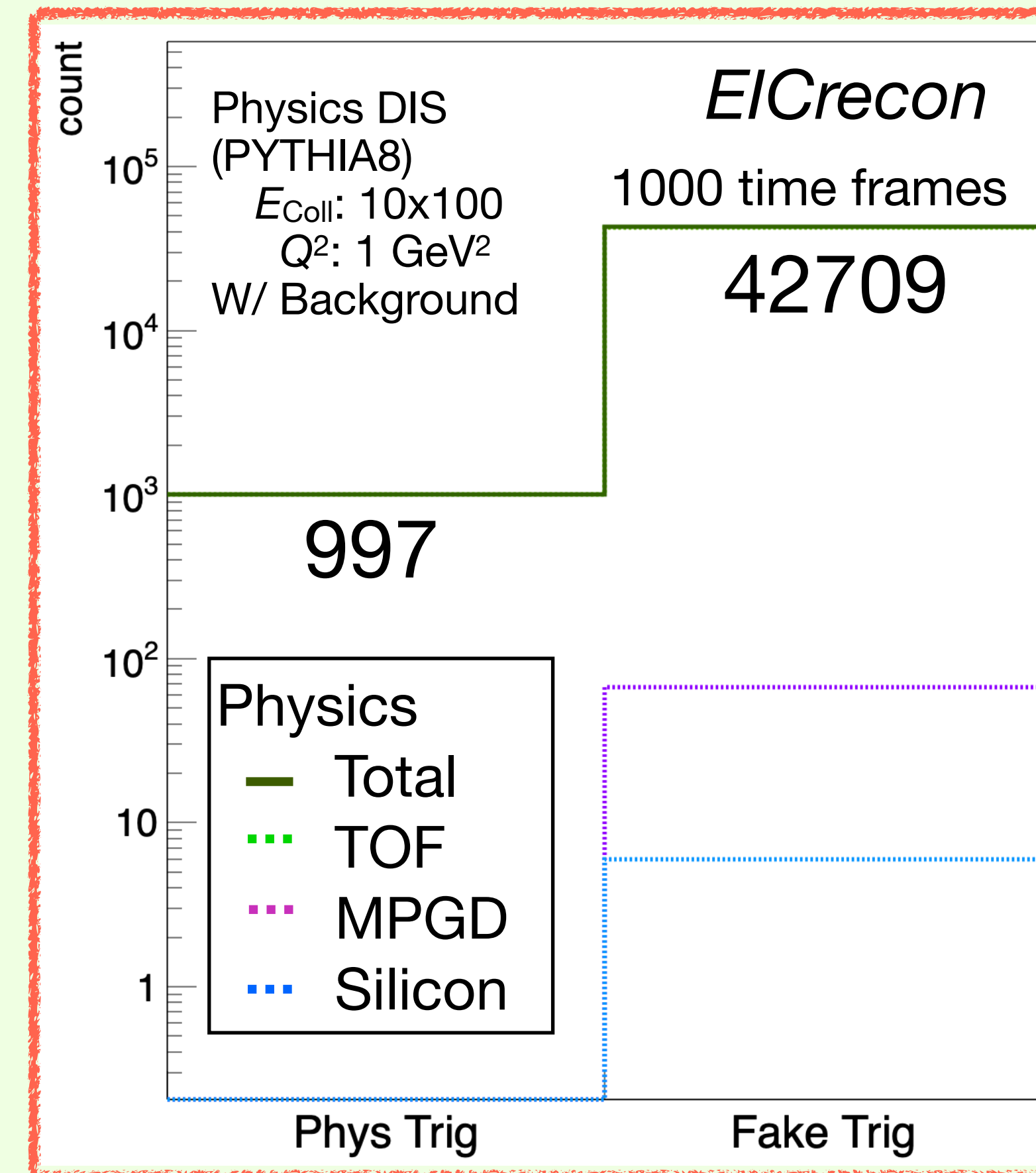
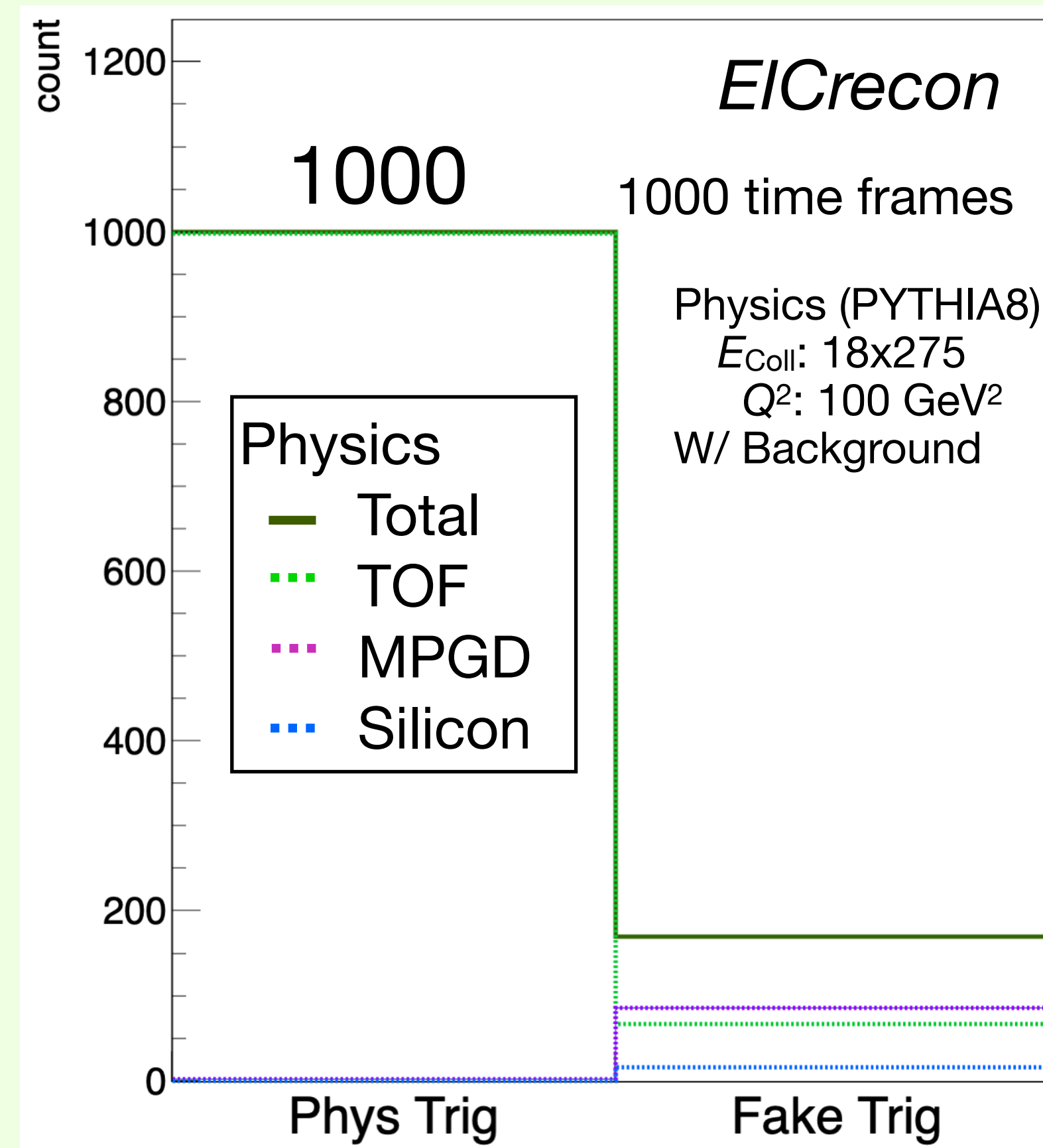
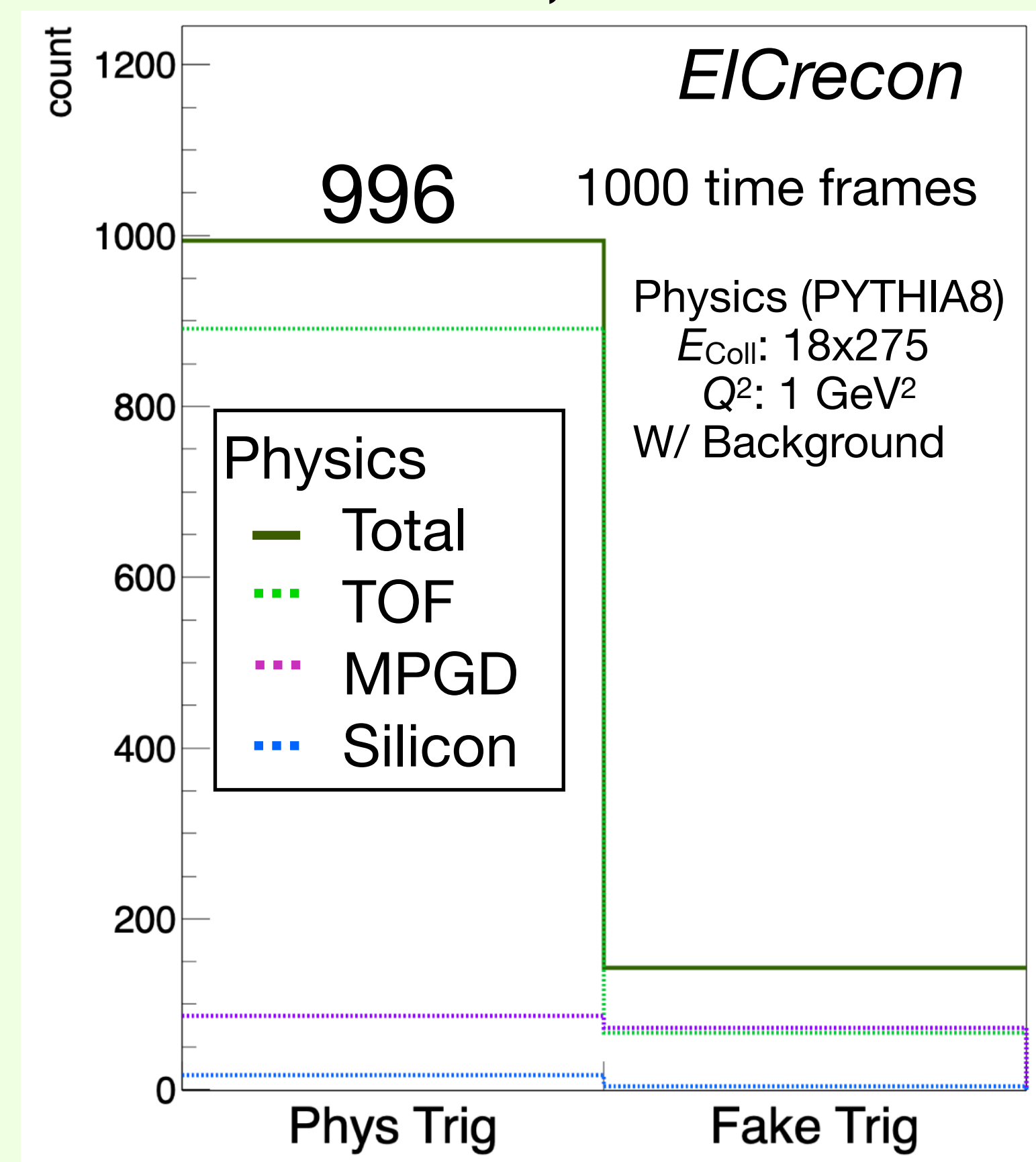
Track p_T Resolution



Good linearity between MC and reconstructed-track p_T .
 The p_T resolution of the time splitting is a little better than W/O time splitting.

Pre-identification Efficiency

- There were problem to count the triggered physics event.
- Fixed them, but there are still some concerns.



- Found the pre-identification can triggered most low Q^2 ($> 1 \text{ GeV}^2$) DIS events.
- In the new bkg simulation, large number of fake triggers were created.
- I am checking the details now.

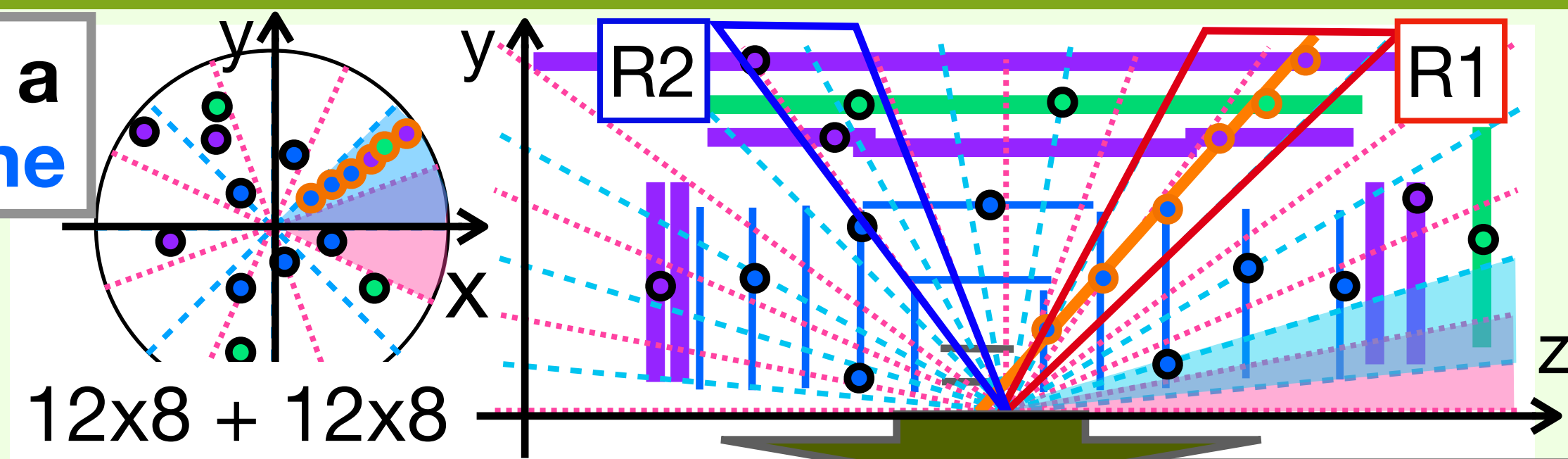
Pre-identification Algorithms

Key:

- **Hit driven** time slice
- Consider the detector's **time resolution**.
- Use coincidences in both **time** and **hit topology**.

1. Using the fastest detector (TOF) slices the time-frame.
2. Check the topology of the hits.
3. If there are more than any three hits in the same region, the time slice is accepted as event candidate.
4. Repeat steps 1–3 for all TOF hits.
5. If any unassociated MPGD hits remain, perform the same process using MPGD hits.
6. if there are no hits in TOF and MPGD in certain time slice, perform the same process based on silicon hits.

All hits in a
time-frame



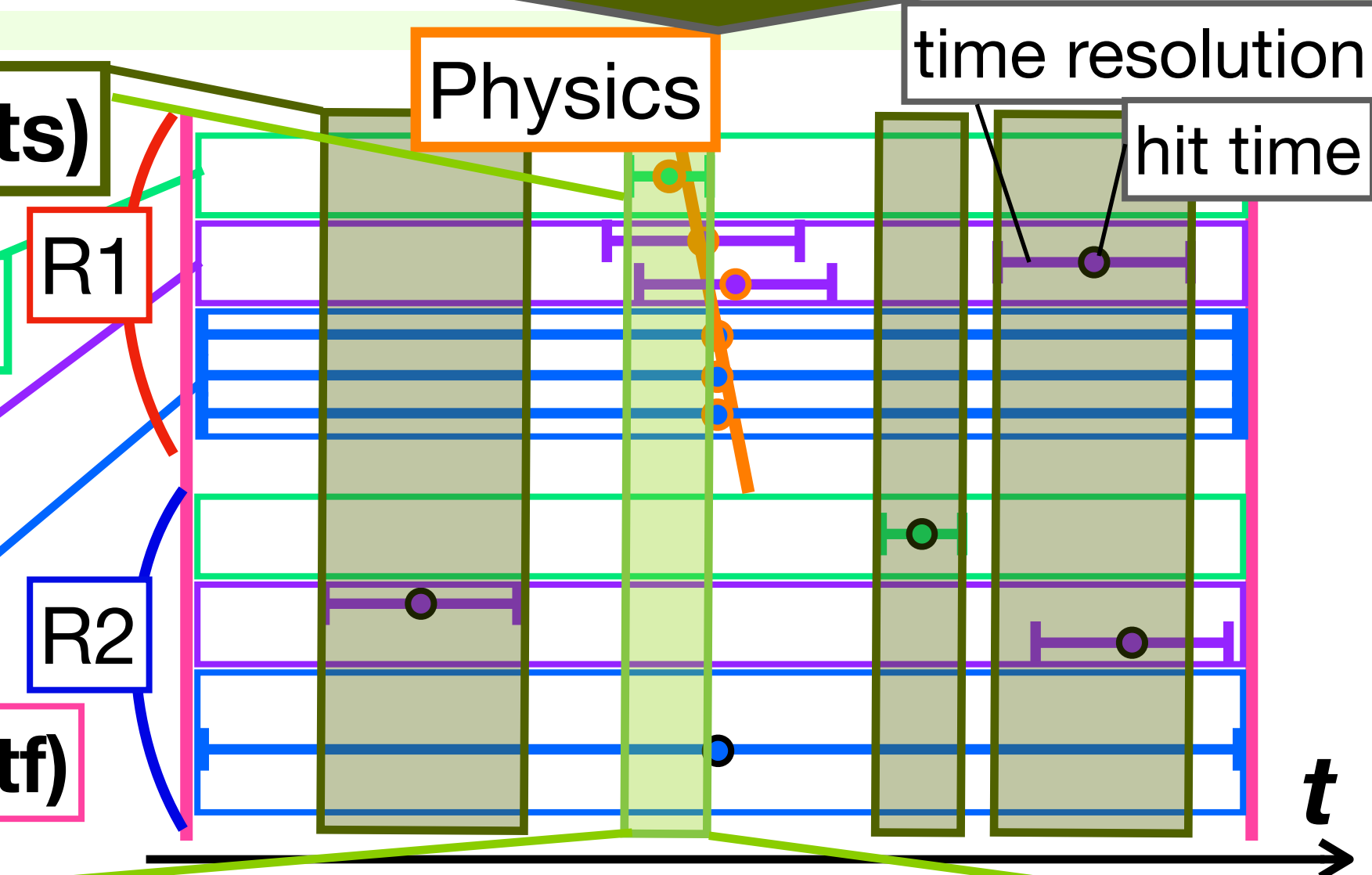
Time-slice (ts)

Det1: very quick (TOF)

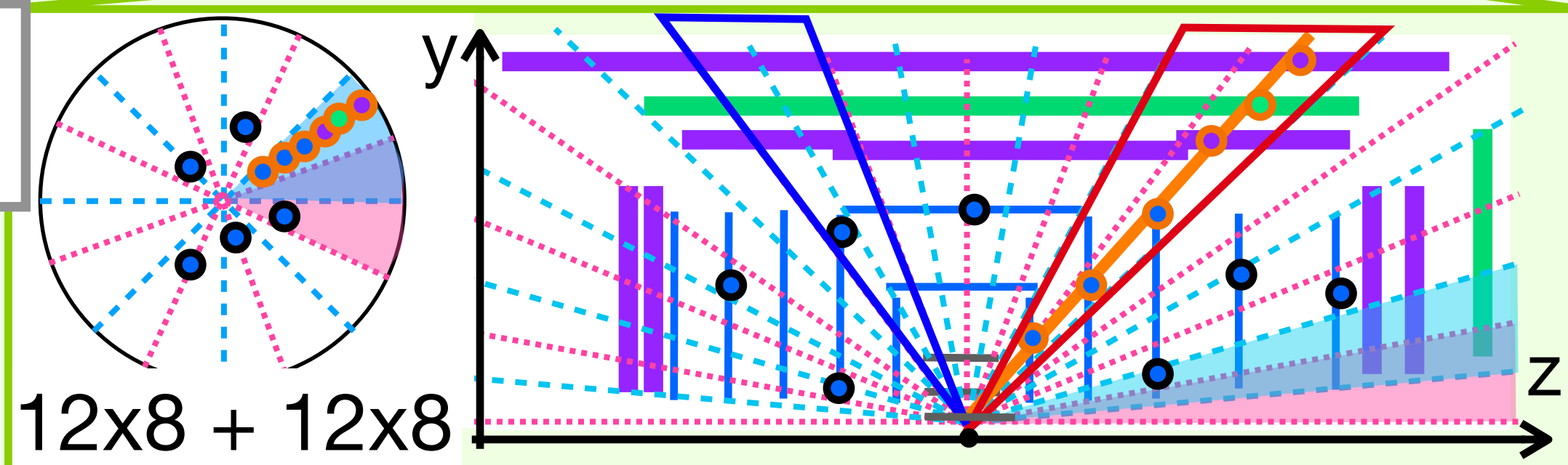
Det2: quick (MPGD)

Det3: slow(Si, vertex)

Time-frame (tf)



All hits in a
time-slice

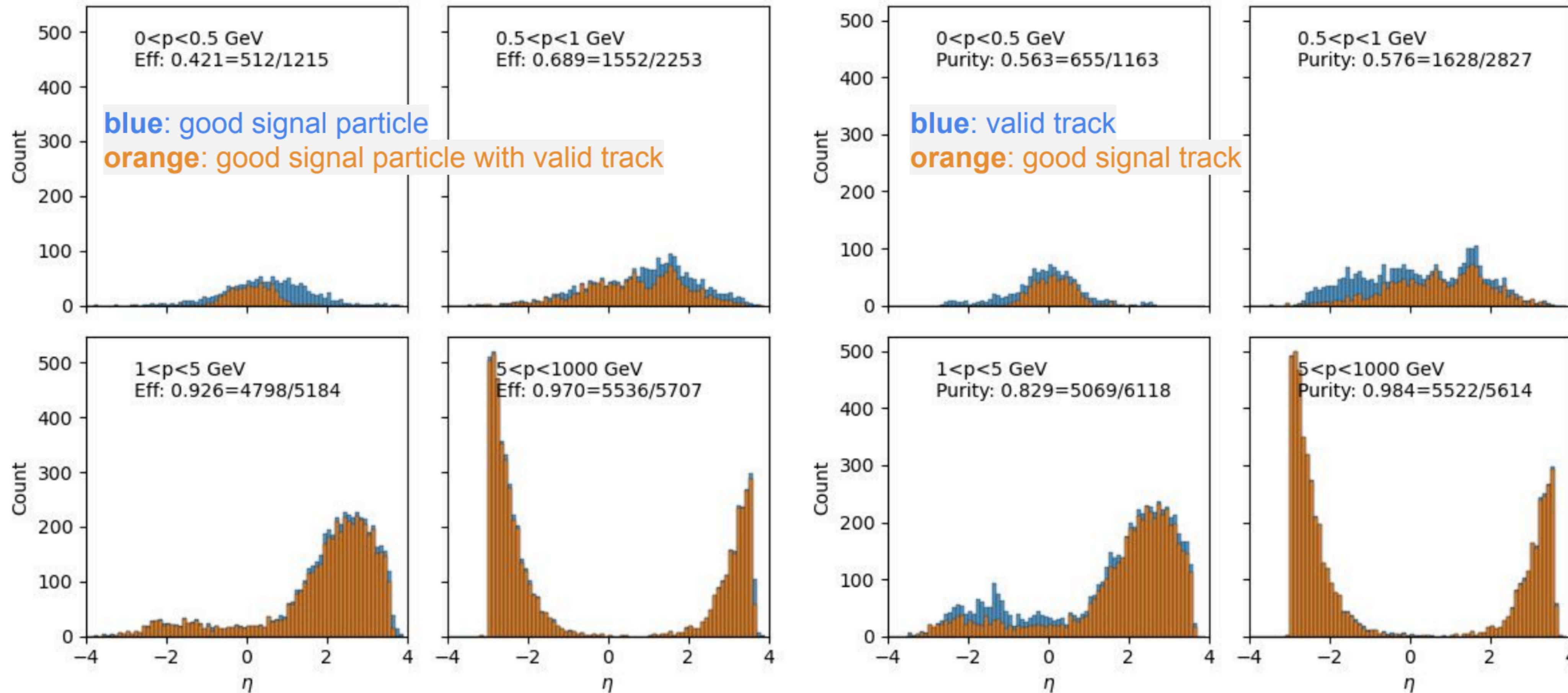


Efficiency and Purity W/ New Background Mixing Simulation

10x275, 10um gold coating, at least 4 hits per track: OK efficiency, OK purity

Efficiency (4 hits) | 10x275, 10um | total=0.863 (12398/14359)

Purity (4 hits) | 10x275, 10um | total=0.819 (12874/15722)



https://indico.bnl.gov/event/31909/contributions/120964/attachments/68635/117876/beampipe_gold_coating_2026-2.pdf

Time resolution of each sub-detector

Tracking detector

Silicon MAPS: $\sim 2 \mu\text{s}$ [[ref, p11](#)]

MPGD: $\sim 10 \text{ ns}$ [[ref, p11](#)]

TOF: $\sim 30 \text{ ps}$ [[ref, p11](#)]

PID

hpDIRC: $\sim 50 \text{ ps}$ [[ref, sec2](#)]

dRICH: $\sim 20 \text{ ps}$ [[ref, p2](#)]

pfRICH / HRPPD / LAPPD: $\sim 50 \text{ ps}$ [[ref, sec2](#)]

Calorimeter

EMCal: $\sim 20 \text{ ns}$ [[ref](#)]

Barrel/Forward/Backward HCal: ?? s

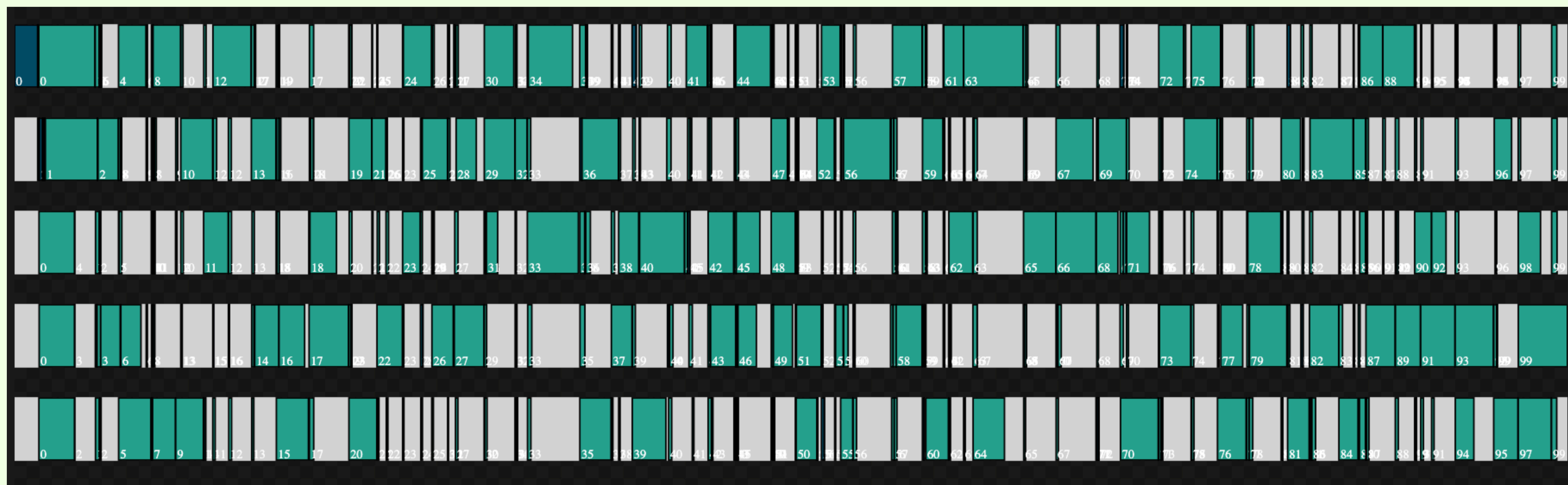
Far-forward (Roman Pots / Off-Momentum Detectors / B0) 30 ps [[ref](#)]

nthread test

nthread=1: 456.223 s / 100 time frame



nthread=5: 541.52 s / 100 time frame



Despite increasing the number of threads correctly, the processing time increased.
→ Need to understand the Nathan's evaluator and the reason.

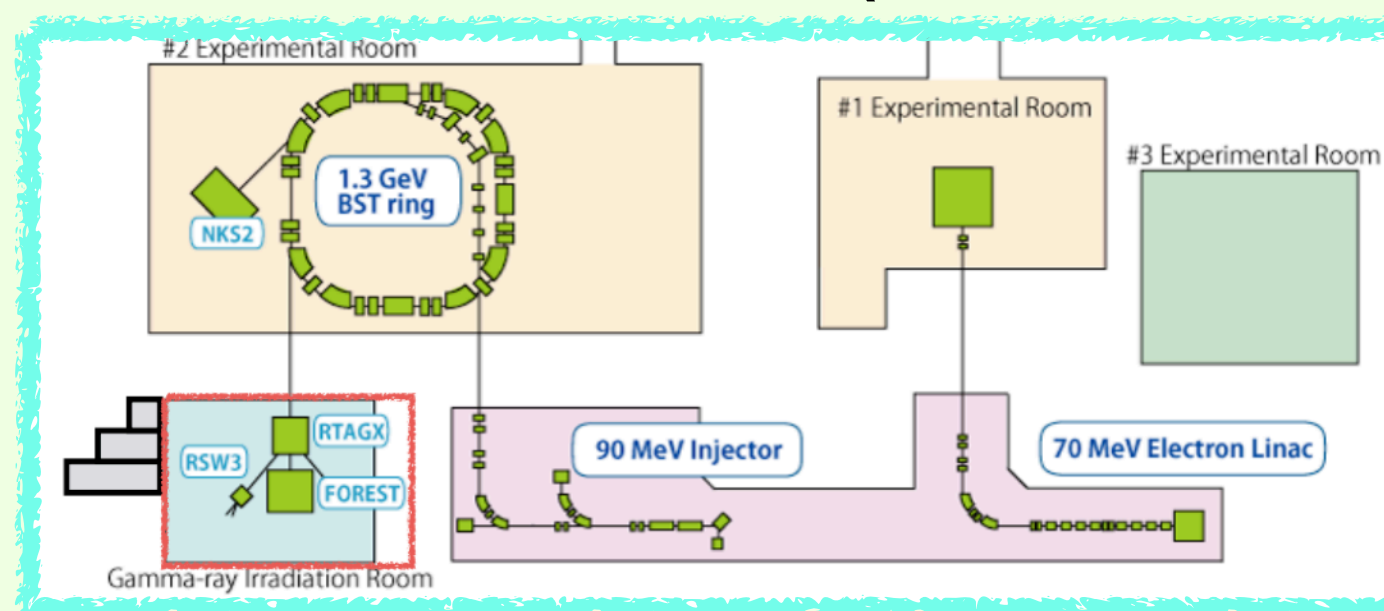
ePIC Streaming DAQ Test with a Local Beam Test

Japan barrel TOF R&D team held a beam test at [RARiS](#) (Tohoku, Japan), 3rd-7th March 2026

GeV- γ beam line:

positron: $< 1 \text{ GeV}/c$, $< 800 \text{ MeV}/c$ ($\sim 30 \text{ Hz}$)

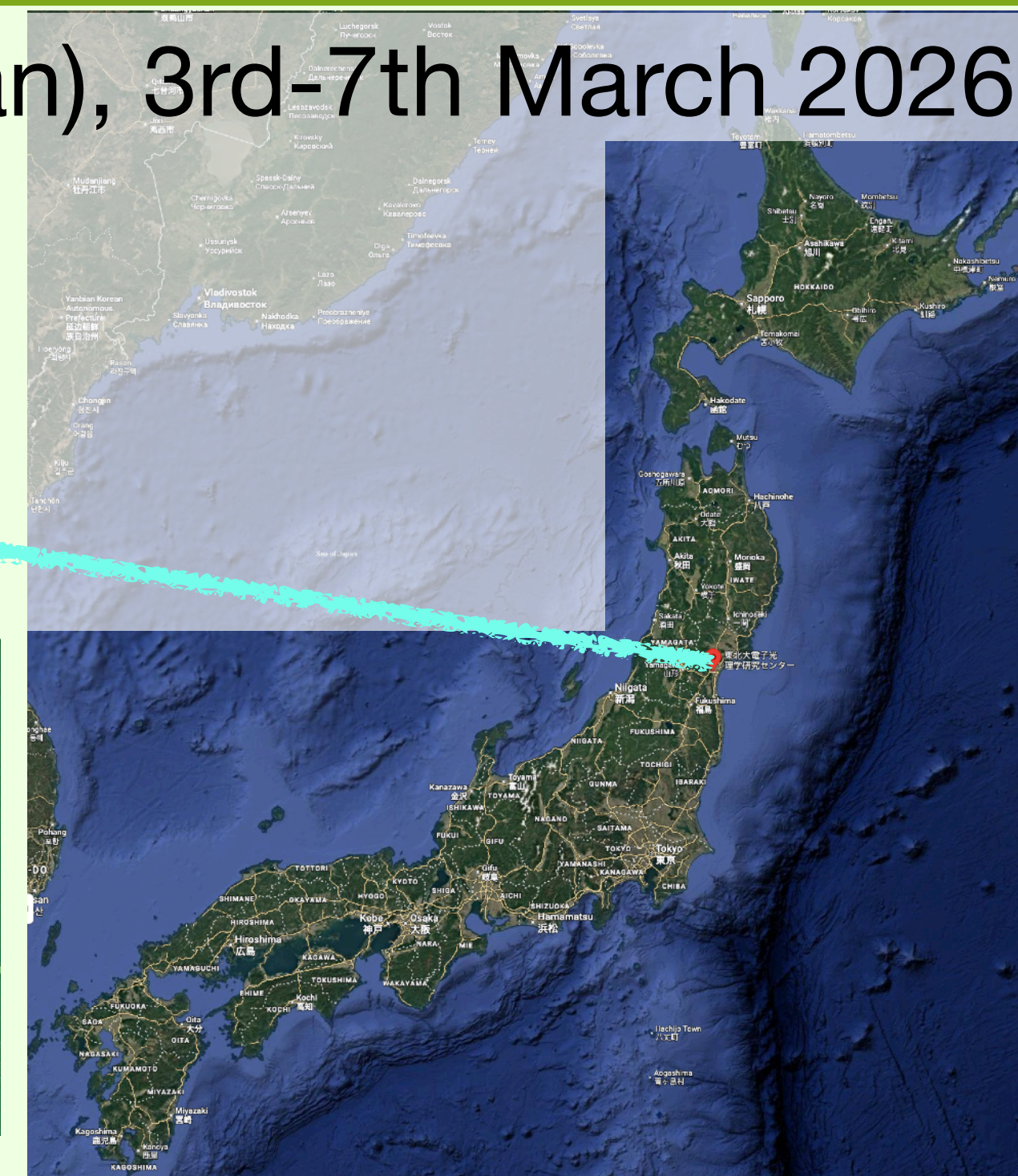
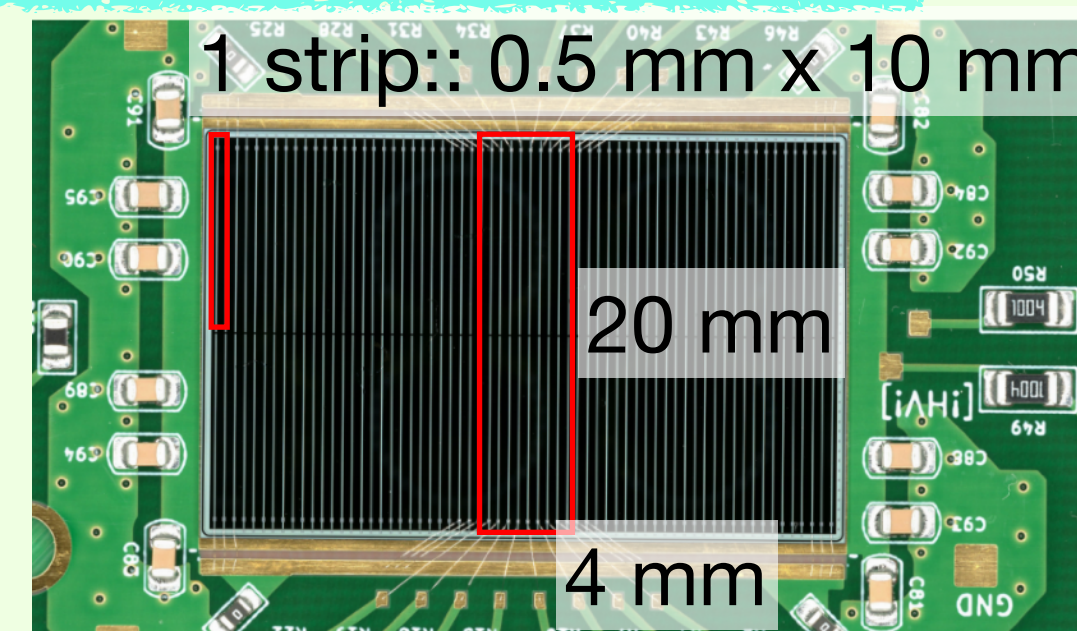
electron: $< 1 \text{ GeV}/c$ ($\sim 30 \text{ Hz}$)



Detector:

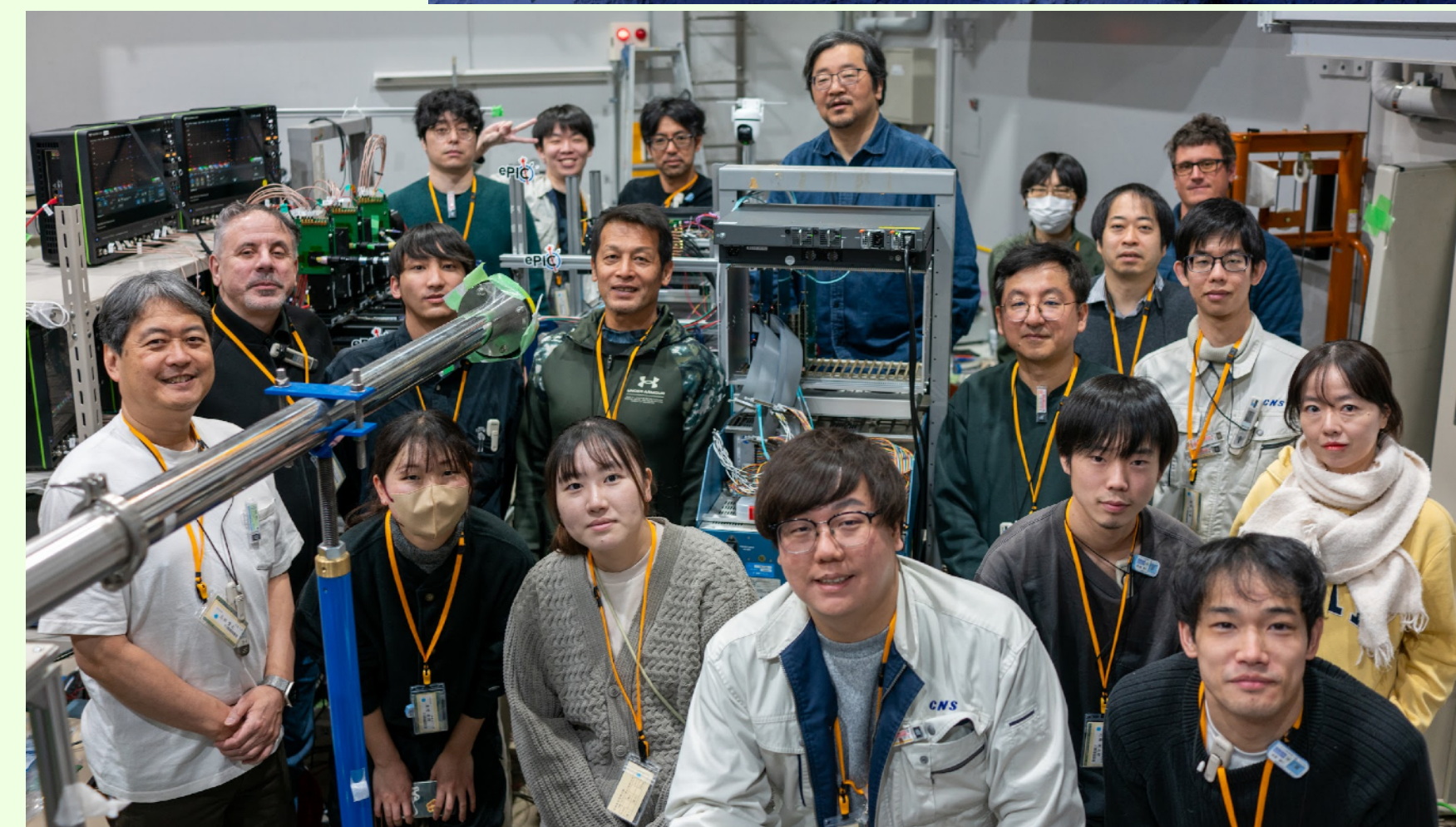
AC-LGAD (AC-coupled Low-Gain Avalanche Diode):

- High time resolution $\sim 30 \text{ ps}$
- High granularity



Beam test

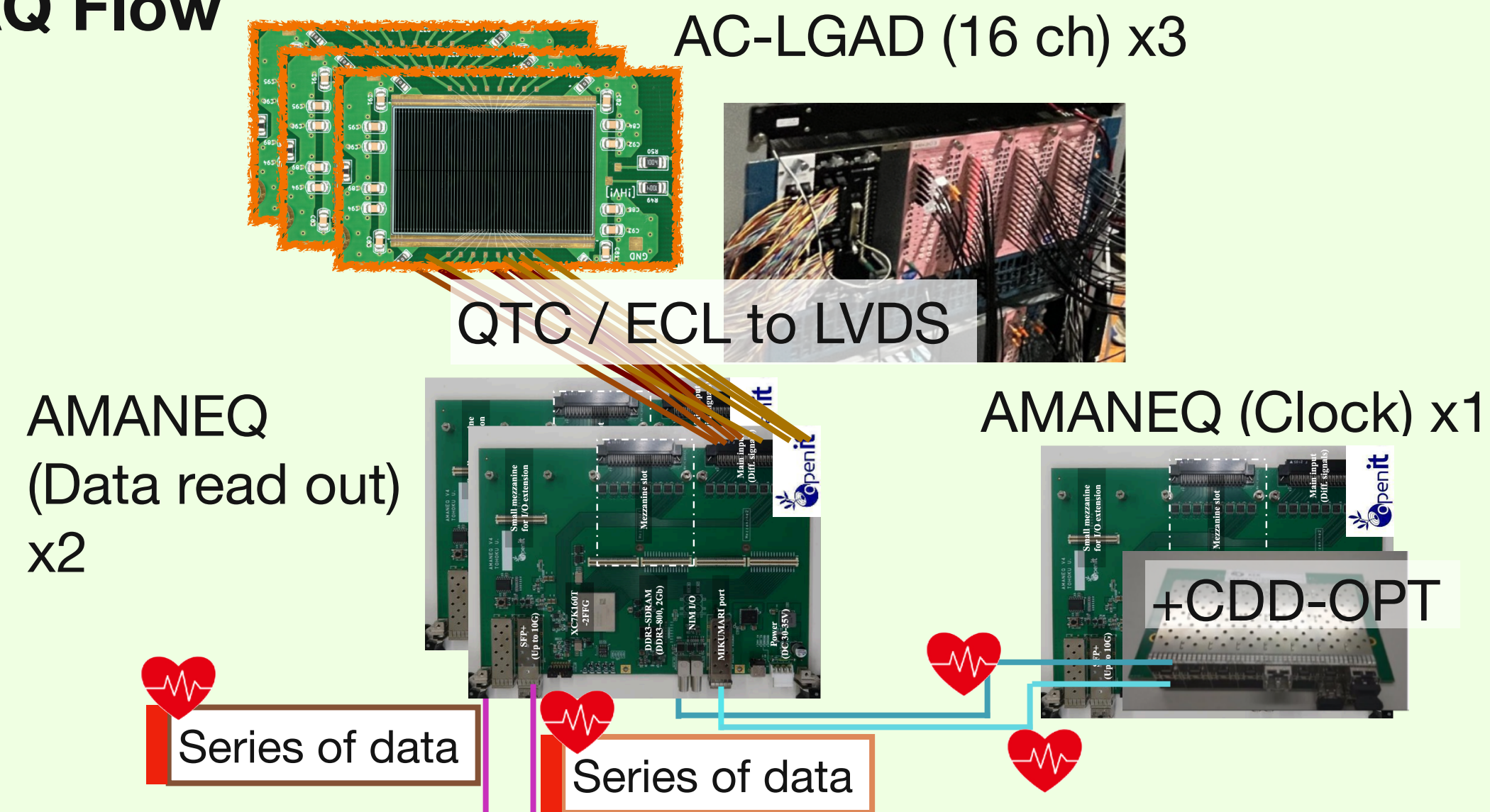
- This beam test was conducted by 23 people.
- We had three parallel lines.
- **line1: mini ePIC streaming test with SPADI using AC-LGAD**
 - N.Kobayashi (RCNP, Osaka Univ), T.Takahashi (RCNP, , Osaka Univ), S. Y. Ryu (RCNP, Osaka Univ), T.Gunji (QNSI, Univ Tokyo), M.Kaneta (Tohoku Univ), K.Ono (Shinshu Univ)
- line2: AC-LGAD performance test using Waveform (Ochroscope)
- line3: RFSoc DAQ test for AC-LGAD performance estimation



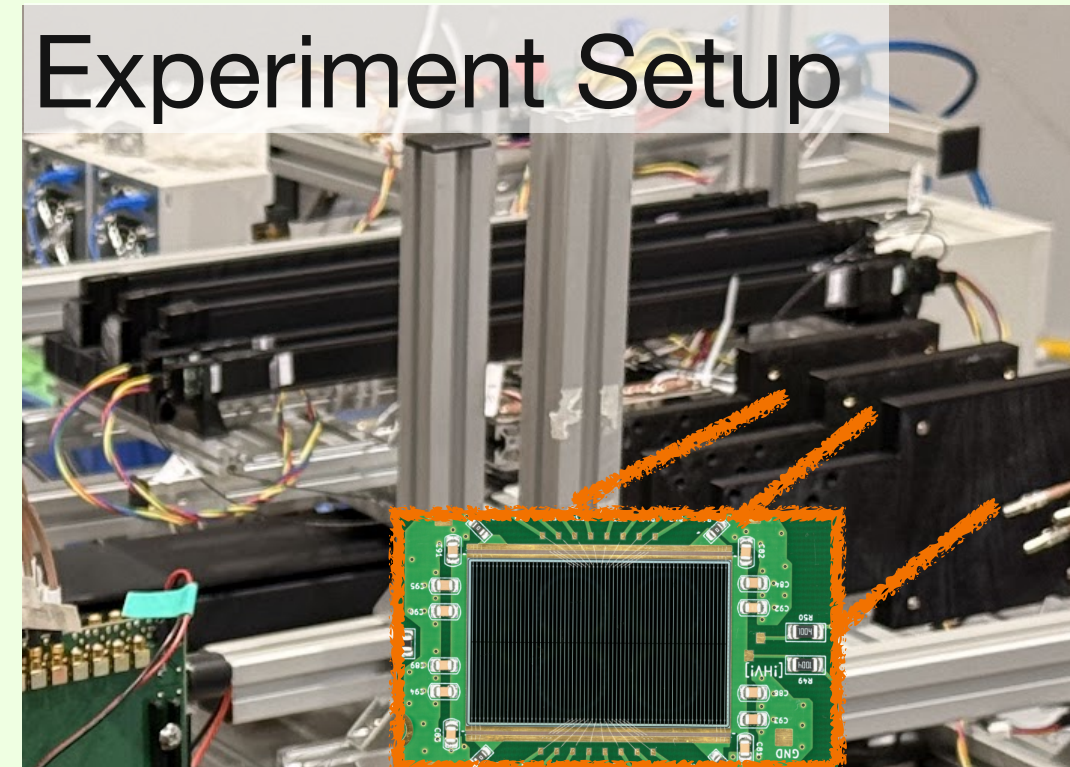
Developed Two Streaming DAQ Lines

Tried mini ePIC streaming orchestration (Echelon0-Echelon2)

DAQ Flow



Experiment Setup



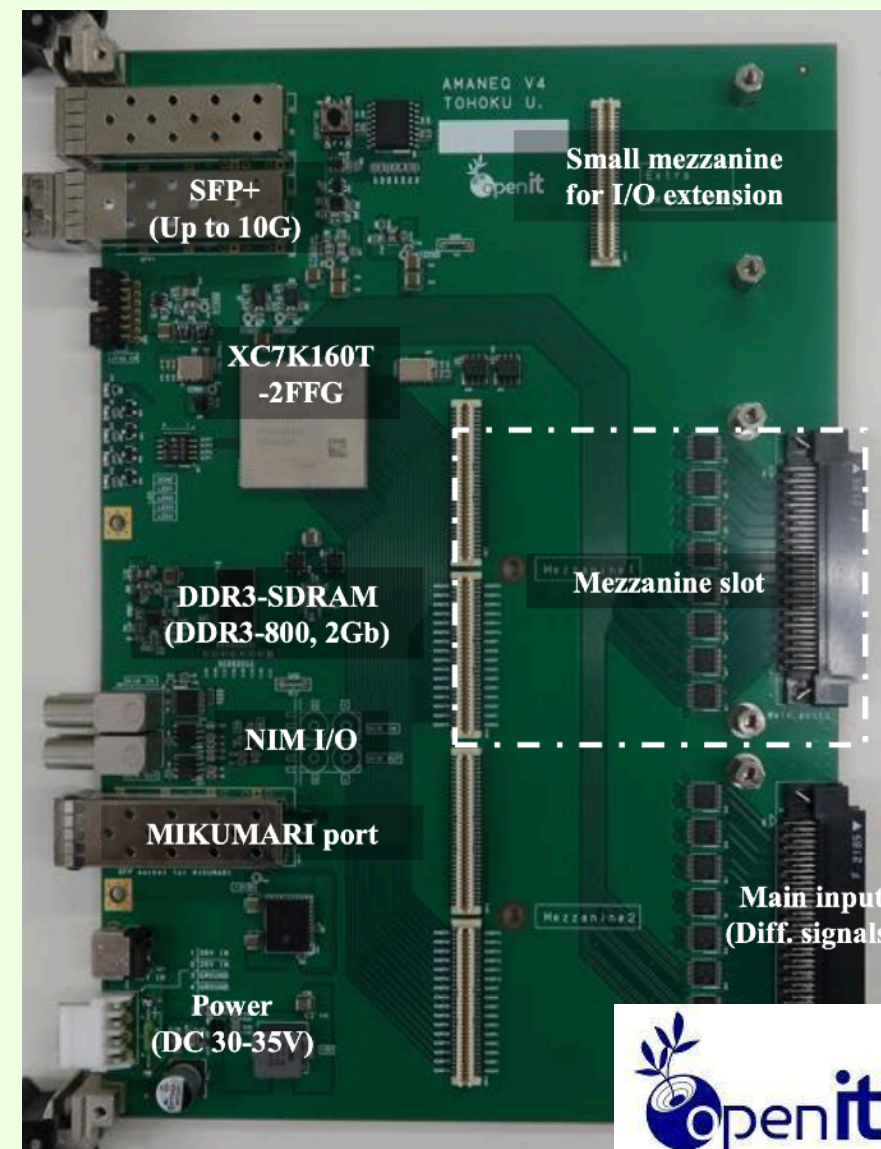
AMANEQ/MIKUMARI Setup



AMANEQ [\[ref\]](#)

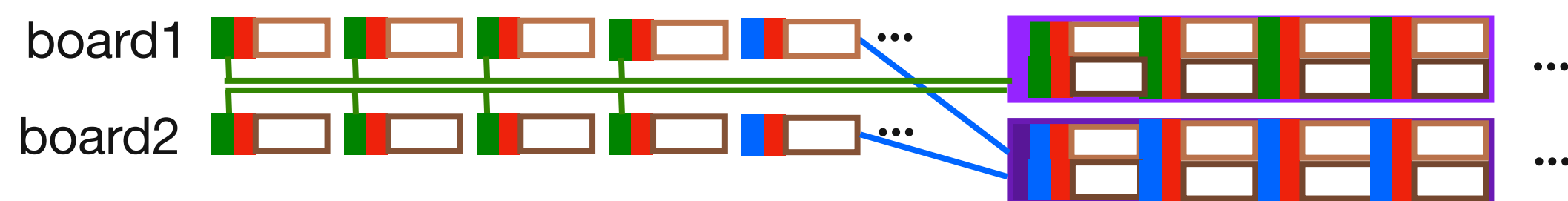
(A main electronics for network oriented trigger-less data acquisition system)

- Kintex7 with speed grade -2
 - Transceiver bandwidth up to 10 Gbps (10 GbE)
- Two mezzanine slot for functionality extension
 - HR-TDC
 - Clock distribution
 - Input-channel extension
 - etc...
- A jitter cleaner (CDCE62002)
- DDR3-SDRAM as a de-randomizer



Desktop PC (NestDAQ) 524 μs

- Sampler
- SubTimeFrameBuilder - TimeFrameBuilder



- Output

1. File sink (EDM4eic) → XRootD → Univ Osaka
2. Send a data as message queue

(3. → GPU logic filter → TimeSlicer → FileSink)

AMANEQ→EDM4eic→ElCrecon

We had not use ElCrecon for actual data not a simulation file produced by npsim.

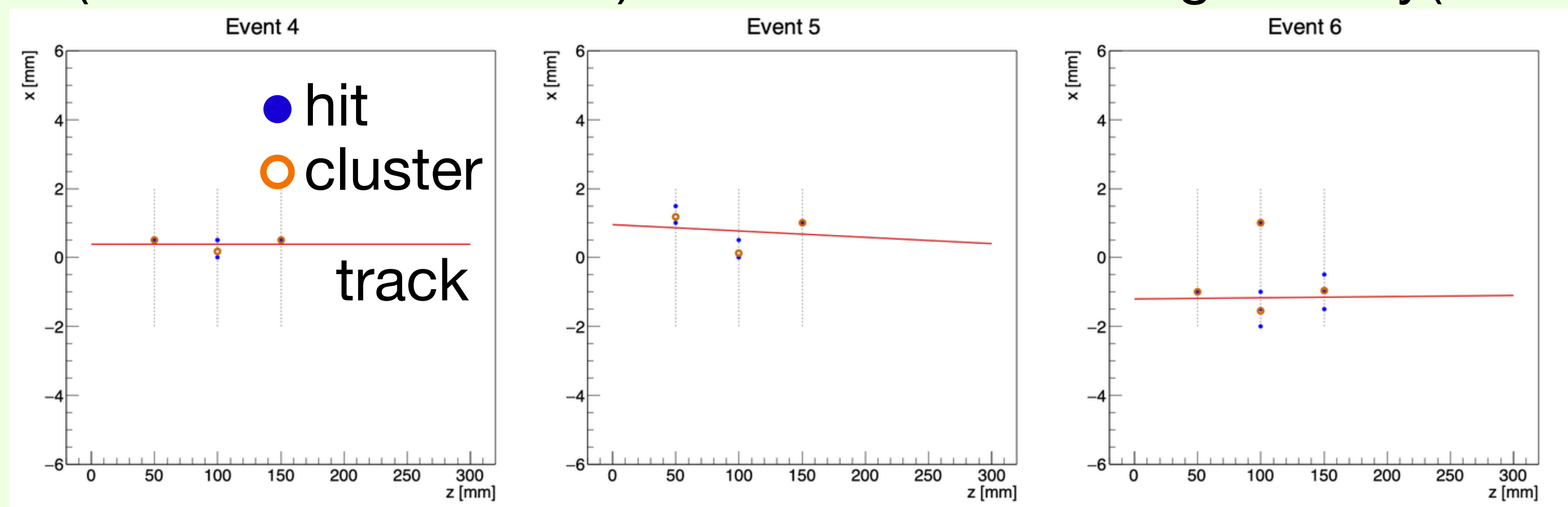
→ Establish to use ElCrecon for a beam test and actual data with SPADI-A DAQ system.

- AMANEQ (TDC) → Sink a file written as edm4eic::RawTrackerHit (N.Kobayashi: RCNP Univ of Osaka) [[git:EDM4eicSink.cxx](https://github.com/EDM4eicSink.cxx)]

- EDM4eic file → ElCrecon (K.Ono: Shinshu Univ, T.Kumaoka: QNSI Univ of Tokyo)

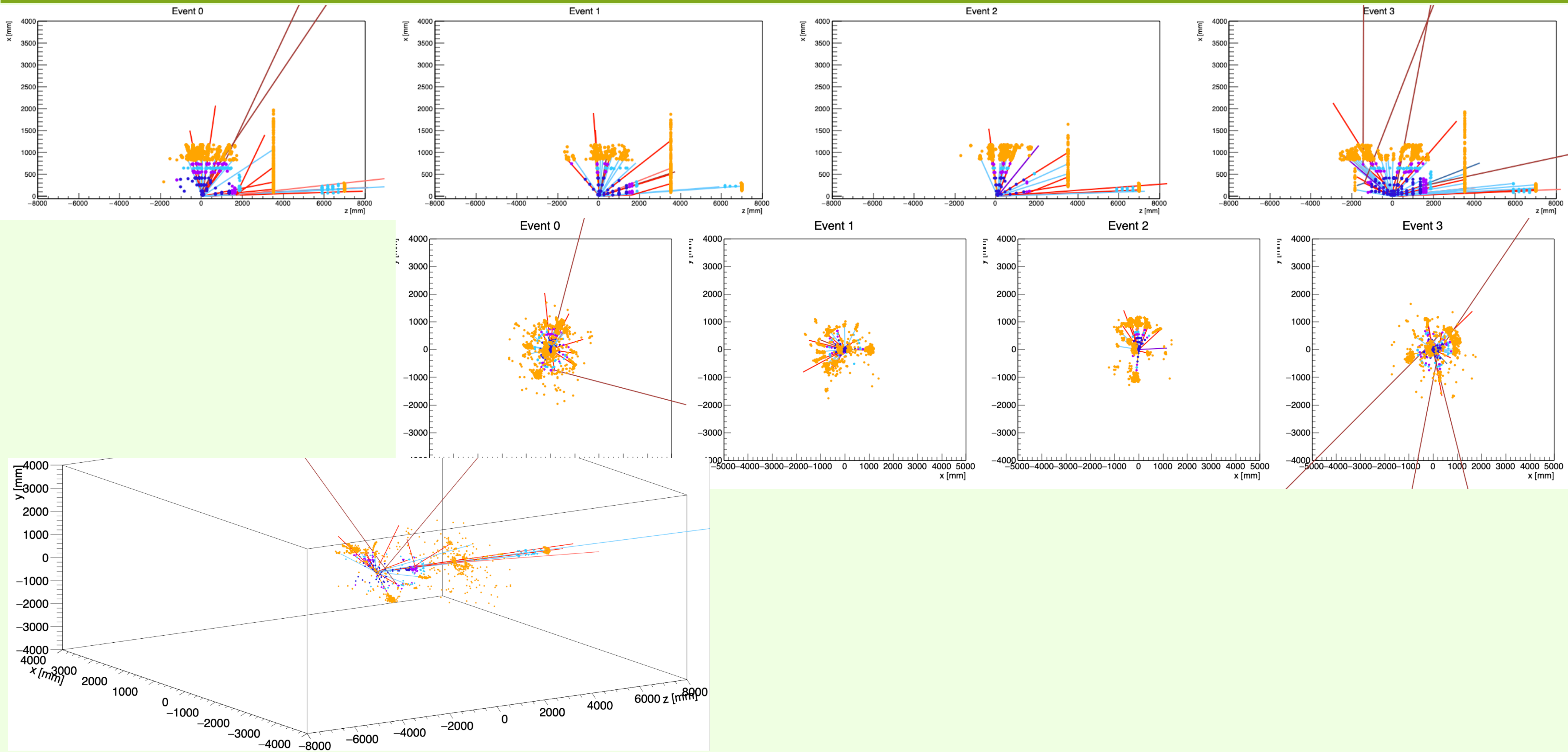
Run the ElCrecon (based on [BTOF Plugin](#)):

Original EventSource (edm4eic::RawTrackerHit)→ LGADHitCalibration_factory (edm4eic::TrackerHit)→ LGADHitClustering_factory(edm4eic::Measurement2D)

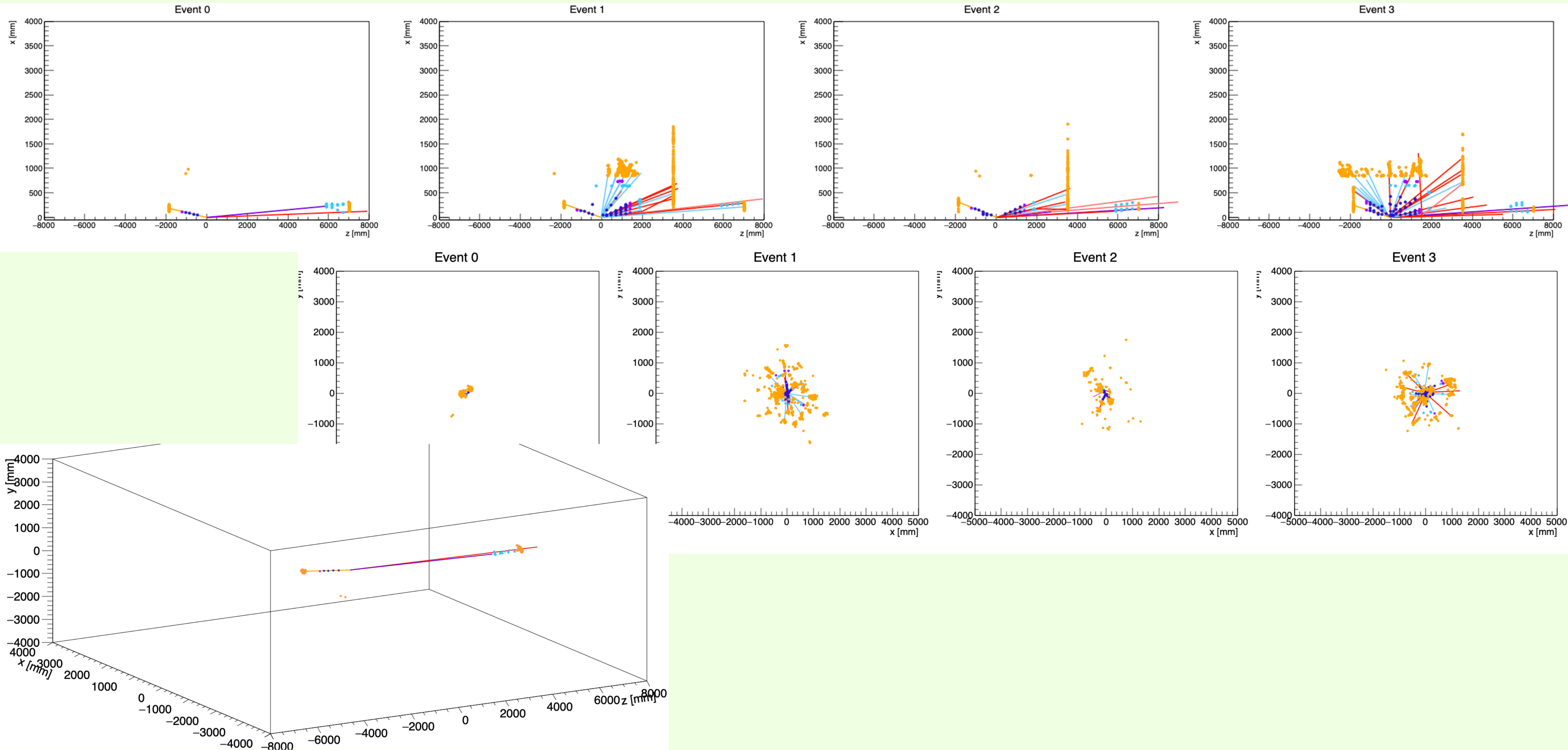


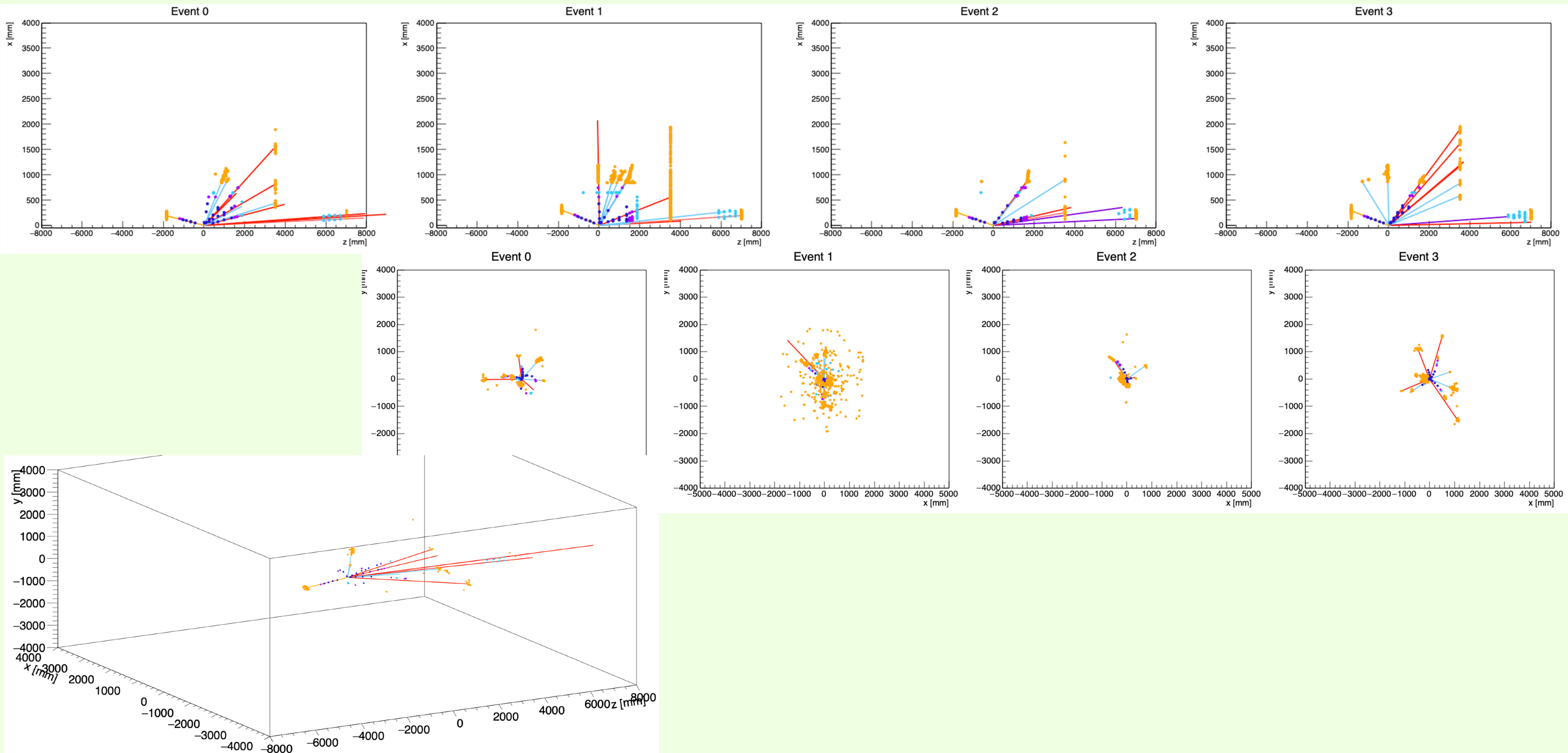
Outlook - Implement the timeframe splitter in this stream.

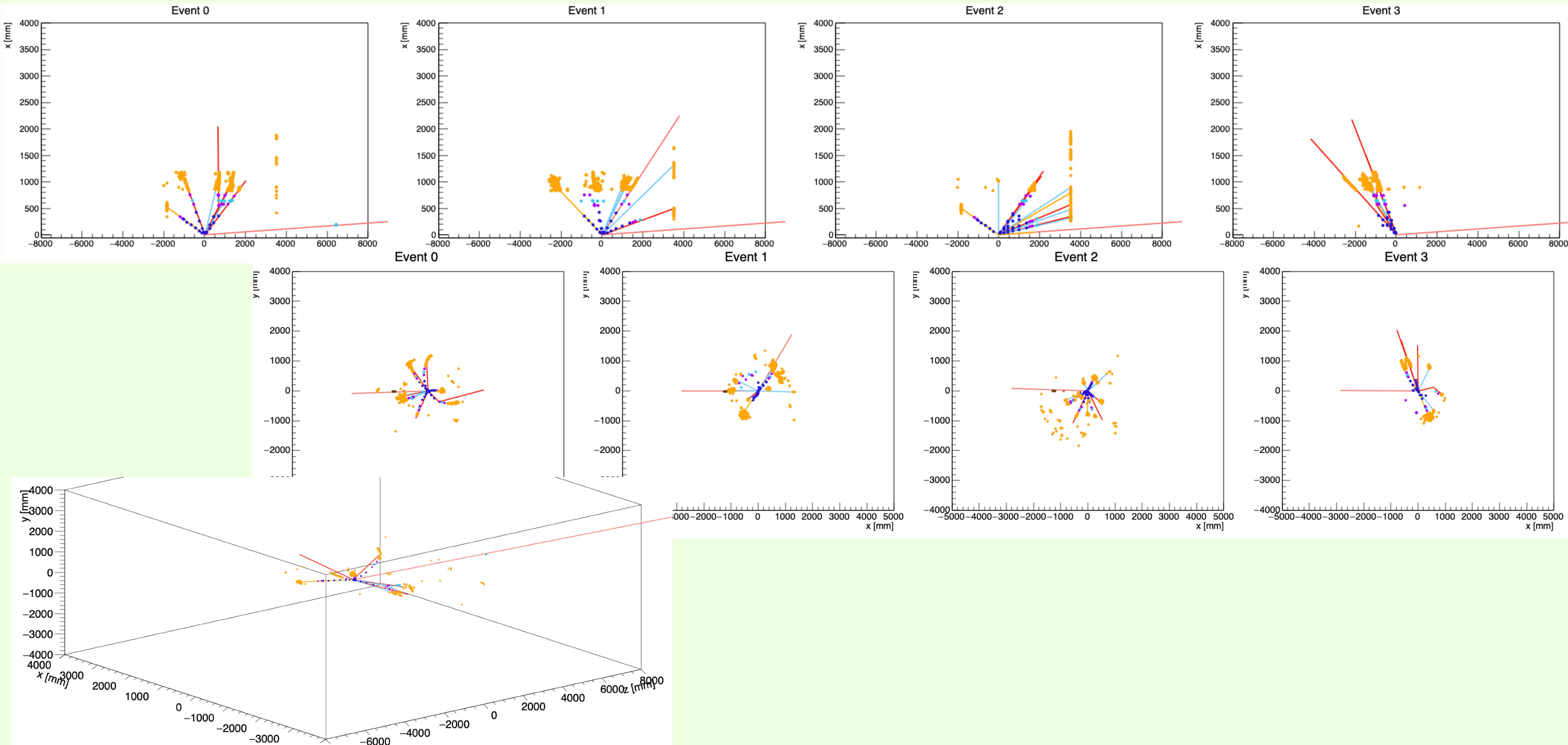
DIS $Q^2 = 100 \text{ GeV}^2$ Event Display



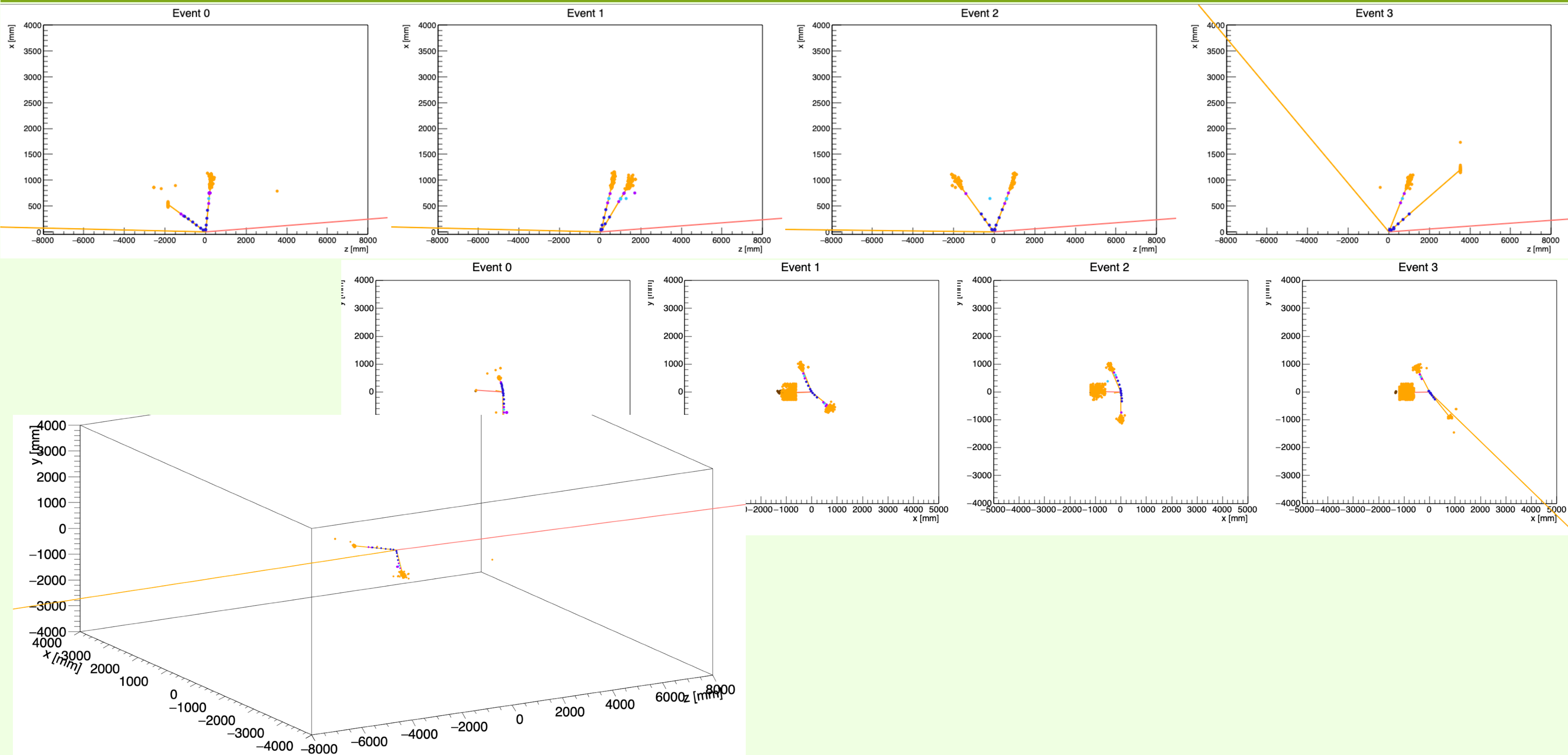
DIS $Q^2 = 1 \text{ GeV}^2$ Event Display



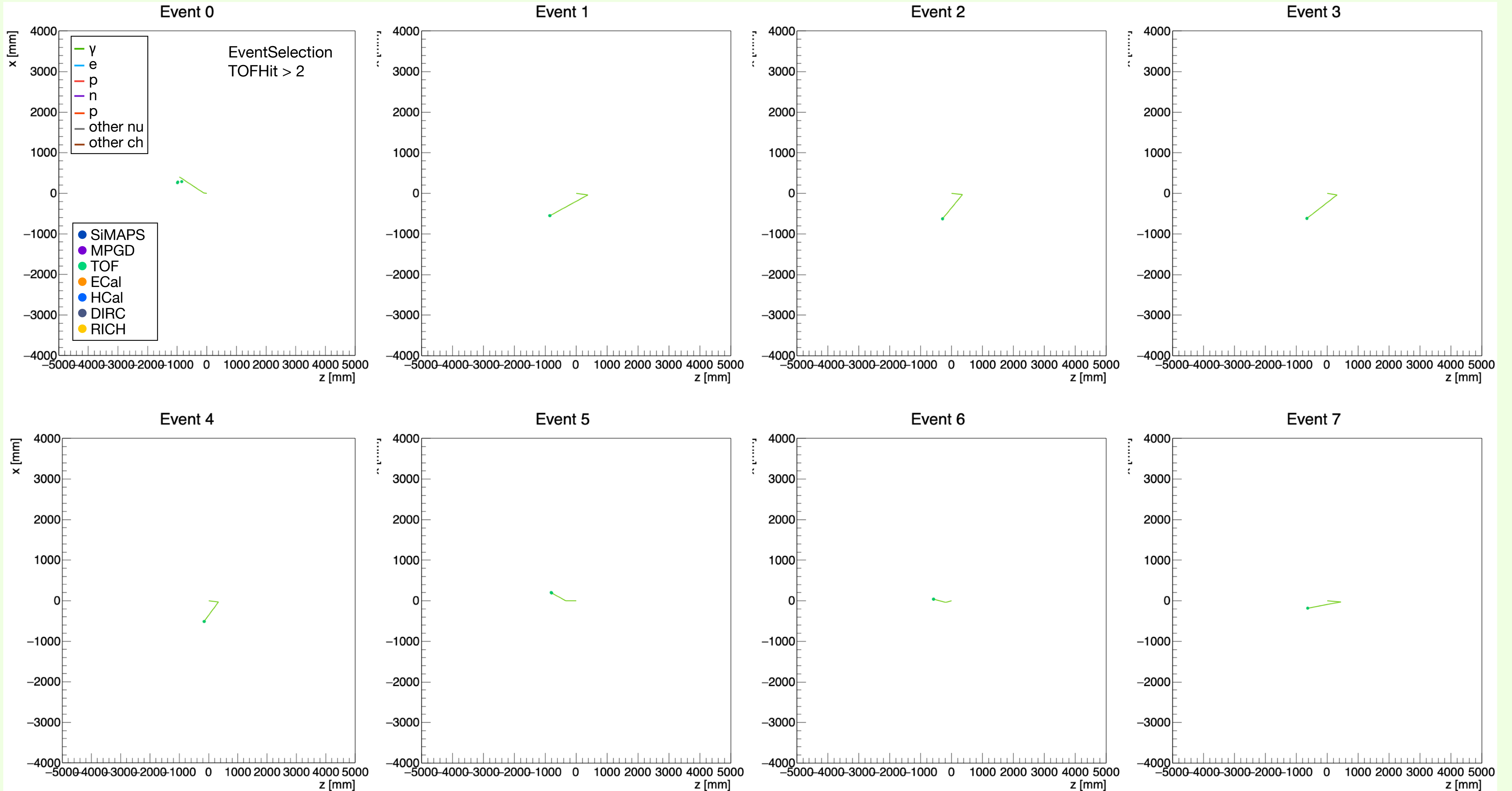




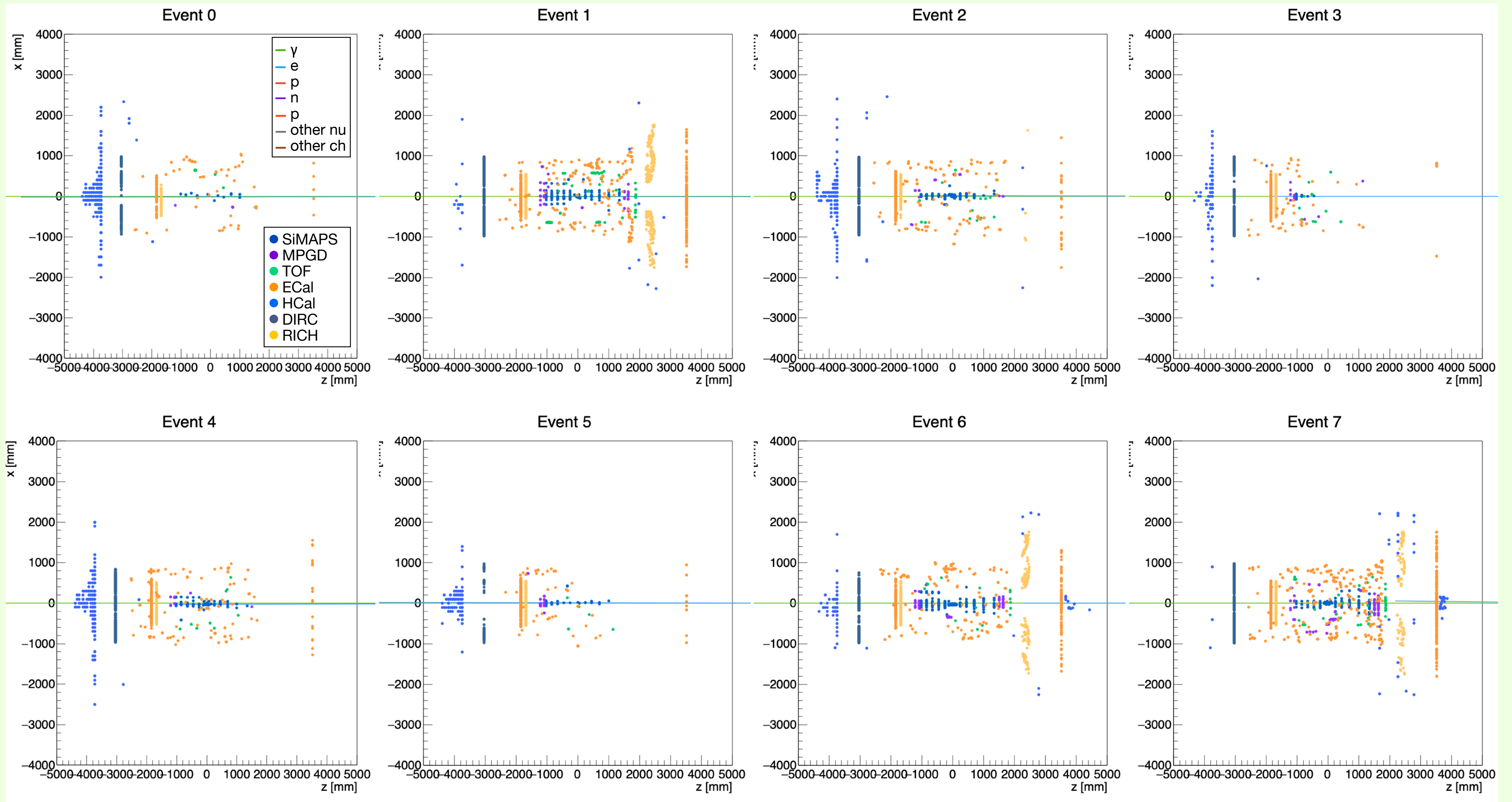
Exclusive Photo-production J/Psi



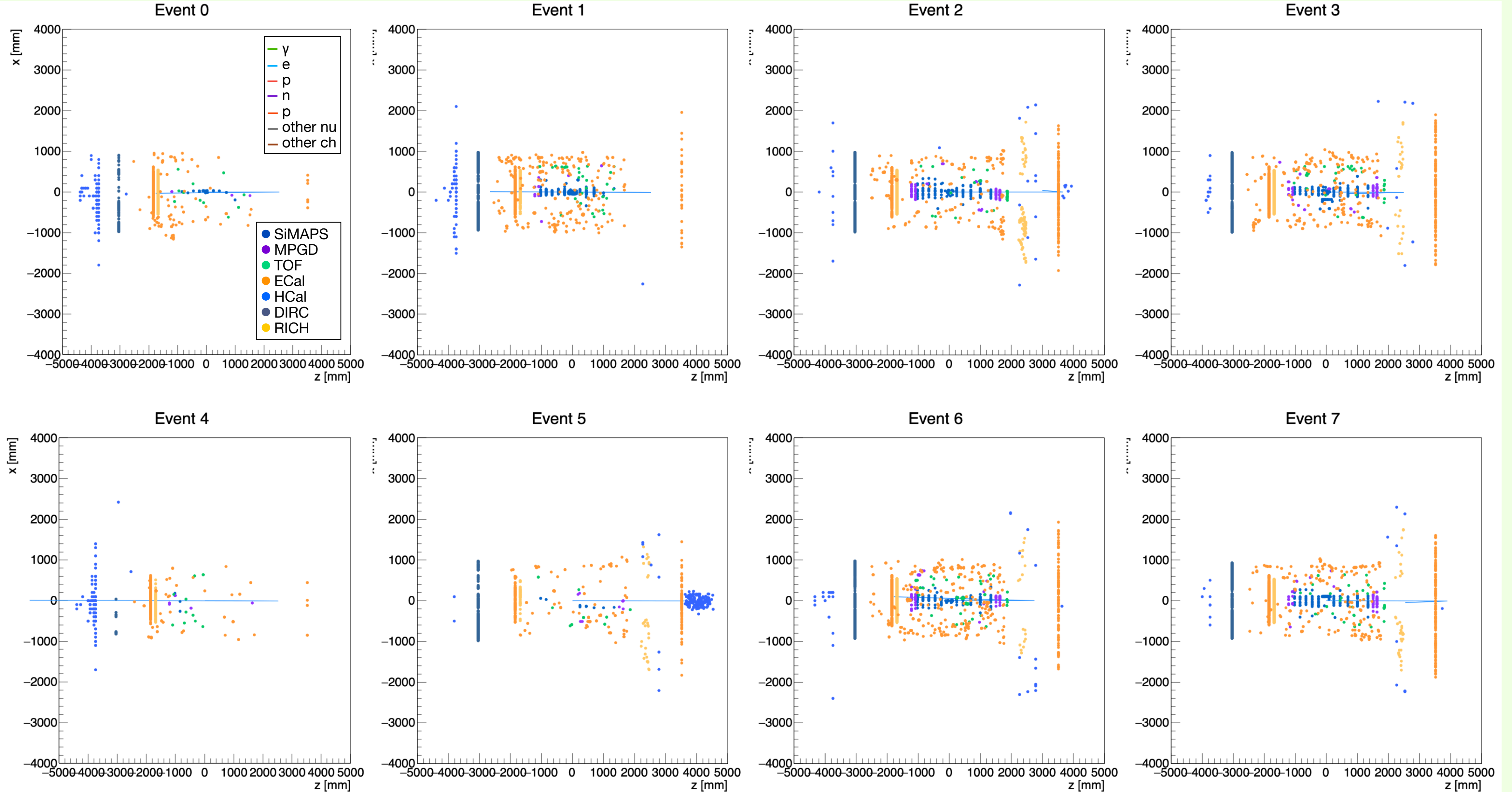
Synchrotron radiation

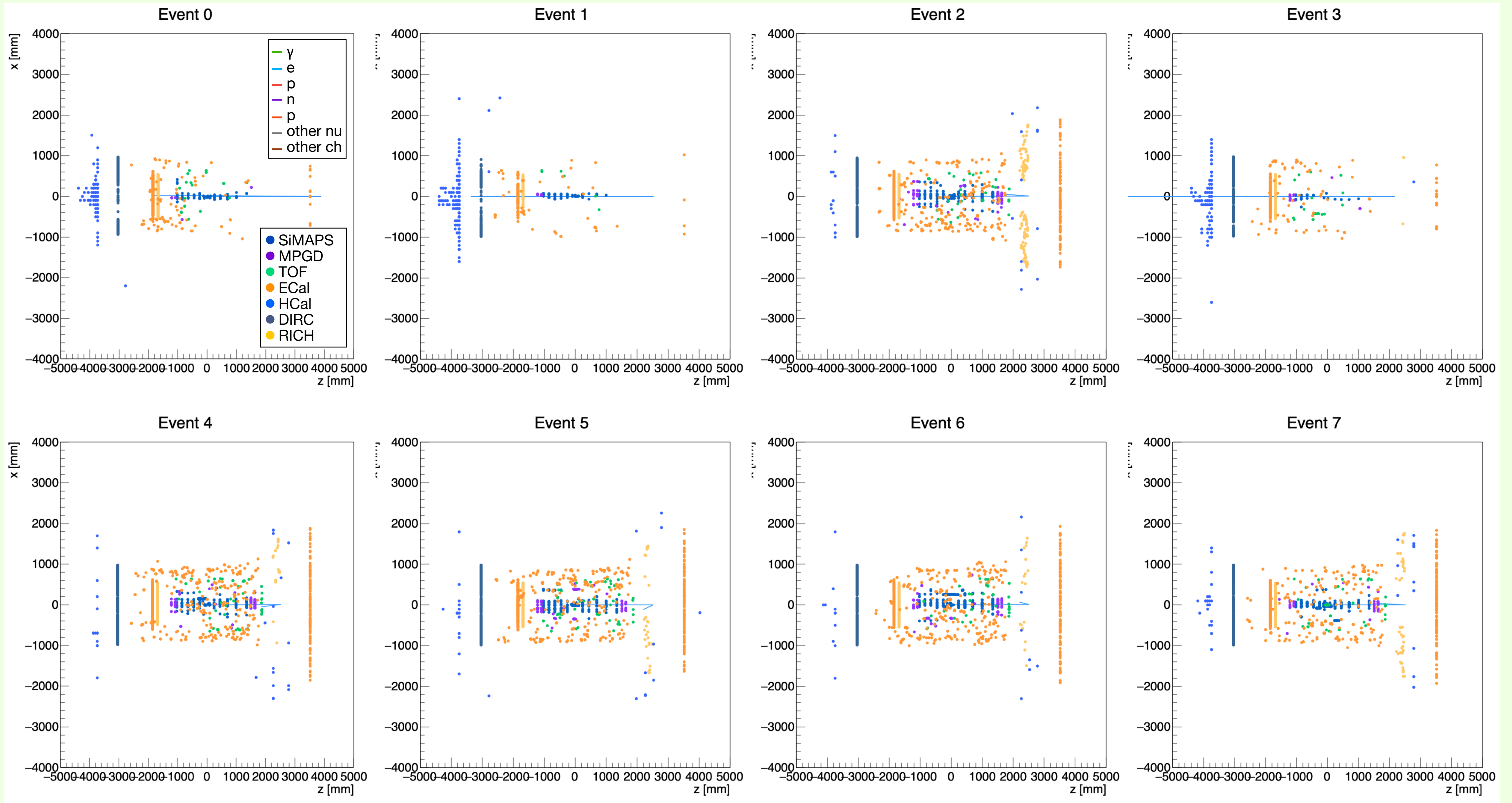


e^- Bremsstrahlung



Coulomb losses





Proton beam gas

