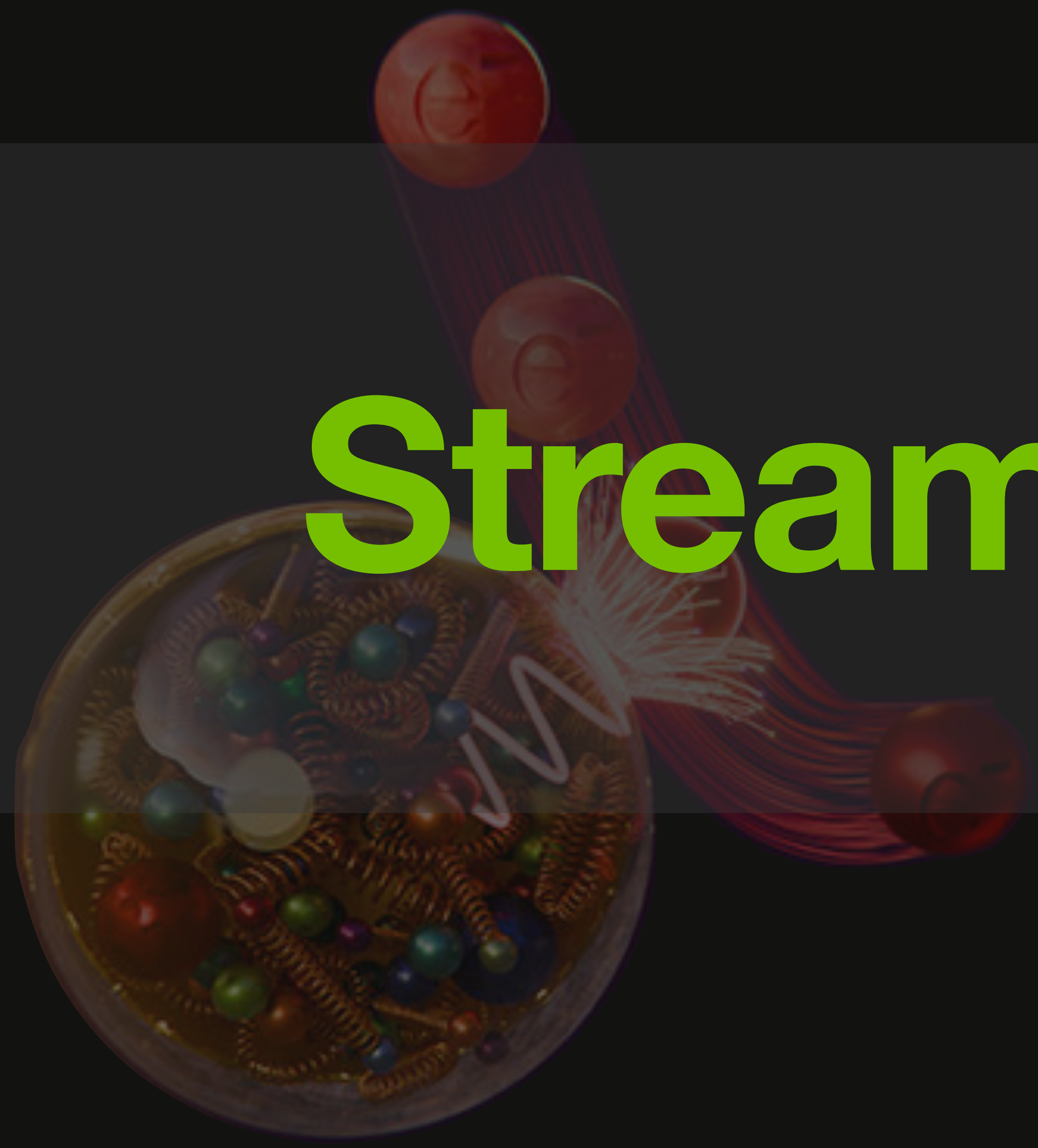
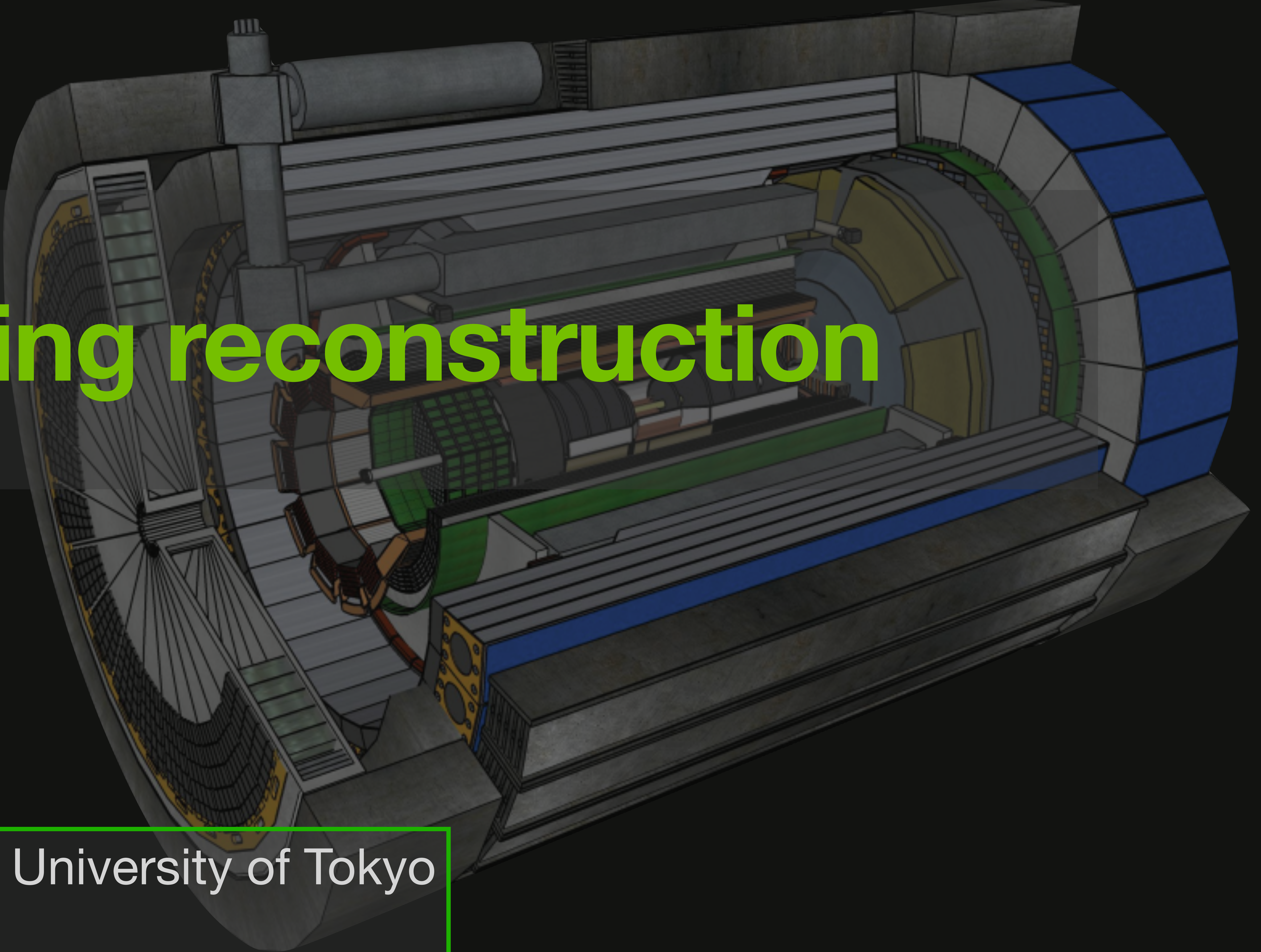




Streaming reconstruction

Quark Nuclear Science Institute, The University of Tokyo
Takuya Kumaoka

EIC-ePIC DAQ Software & Computing Model

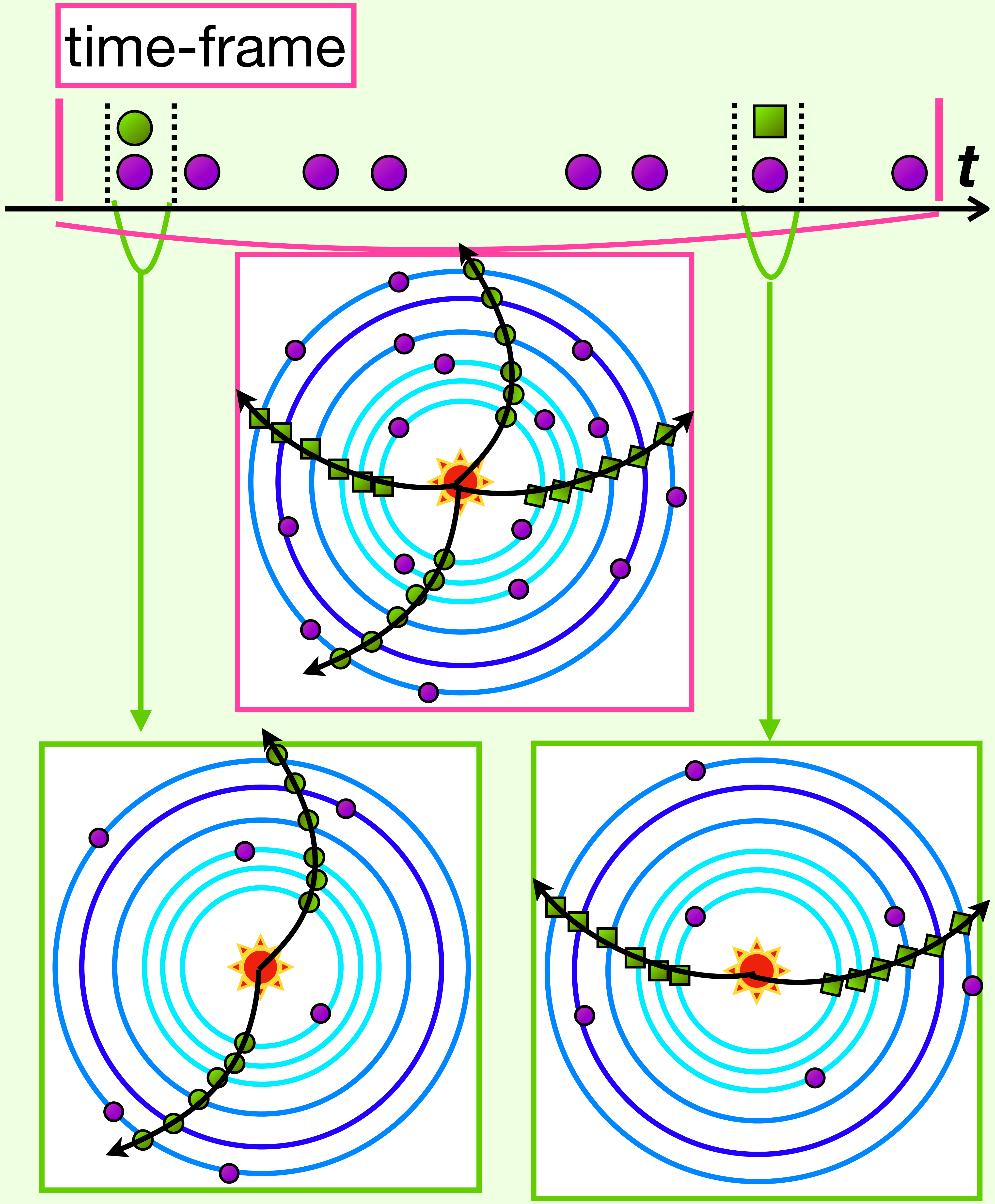
Purpose of this study

Acquire unbiased physics events at 500 kHz (@ 18x275 GeV^2)

Expected background rate (dominated by Synchrotron radiation) : 3.3 GHz(@ 18x275 GeV^2)

→ **Essential to deploy streaming reconstruction to select physics event and reject background**

(10 Tbps → 100 Gbps)

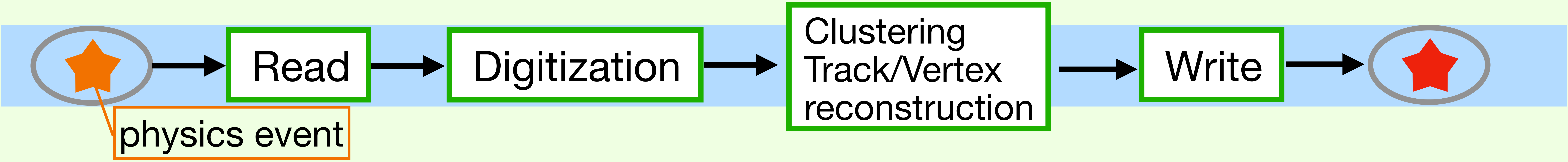


Purpose

Development of streaming reconstruction and algorithms for event selection

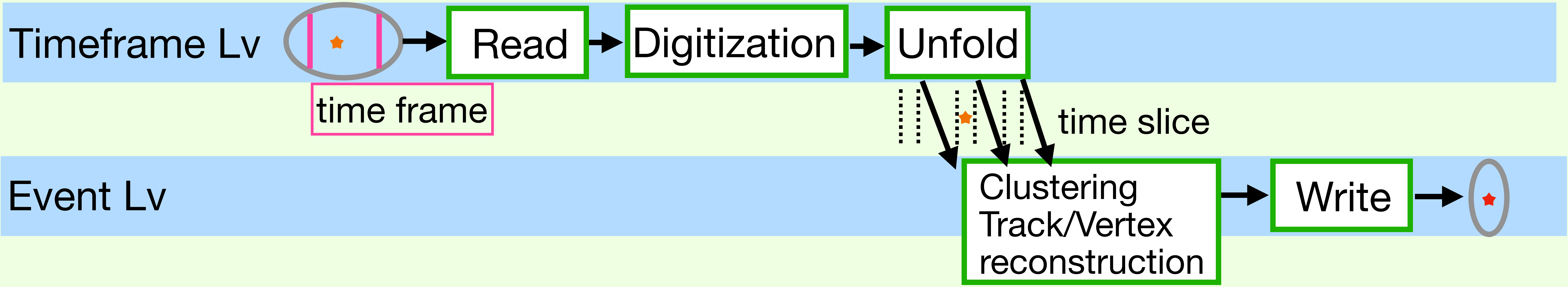
JANA2 Data Flow

Event based topology



Time frame based topology

※ Unfold: time frame is splited into time slices



Splitting a single frame (time frame) into multiple frames (time slices) for parallel processing requires an **additional data stream.**

→The **unfolder** splits a fame and transfers them to a different stream level

Online Event Filtering System

In 2025, I developed **1-3** baseline frameworks. (Of course, there's still plenty of room for improvement.)

Current status of development

1. Implementation of the pre-filter algorithm in EICrecon

2. Validation that the algorithm works through a blind analysis

3. Performance test for the pre-filter

Sub-detectors used for the pre-filter

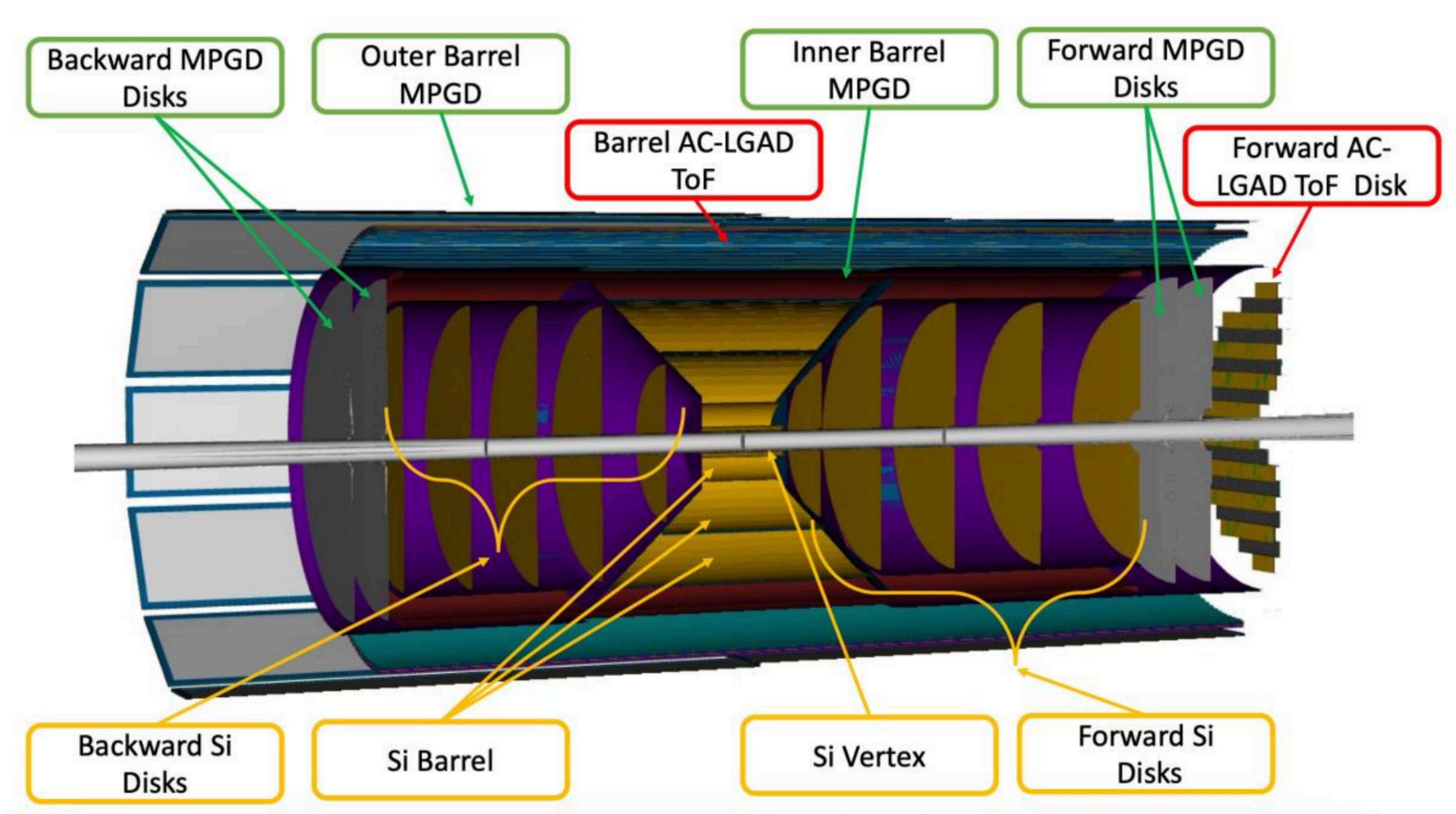
Need some sub-detectors's hit coincidence to extract a physics event from a time frame

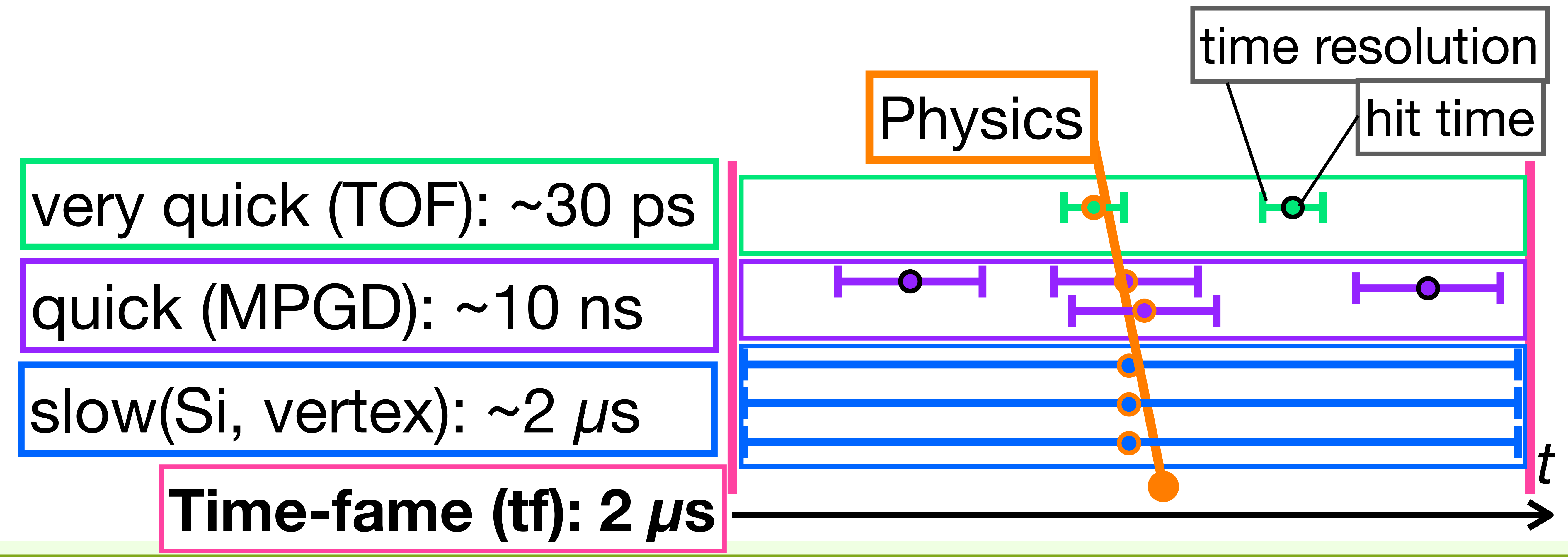
→ Use tracker detectors (**ToF**, **MPGD** and **Silicon (MAPS)**) for the event filter

ToF
- Very high time resolution (~ 30 ps)
- One layer / There are uncovered region

MPGD
- High time resolution (~ 10 ns)
- Wide covered region
- Two layers

Silicon (MAPS)
- 2 μs time window (~ time frame length)
- Wide covered region
- Multiple layers

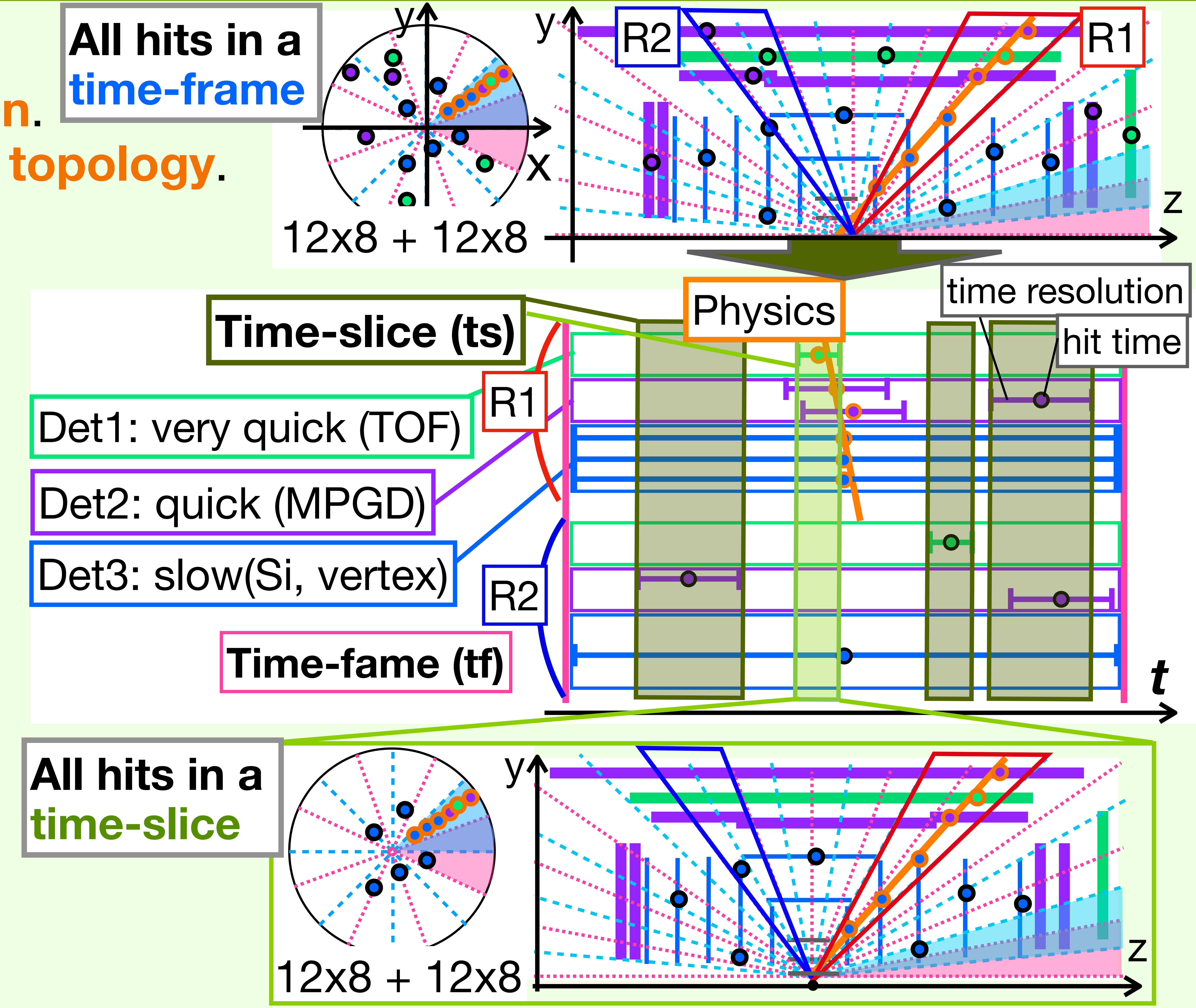




Pre-filter Algorithms

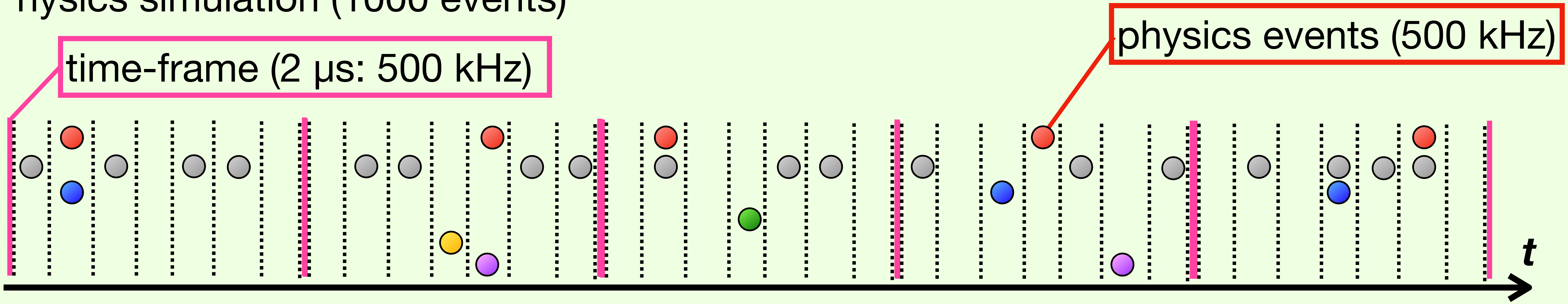
Key:
- Consider the detector's **time resolution**.
- Use coincidences in both **time** and **hit topology**.

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.



Simulation Setup

Physics simulation (1000 events)



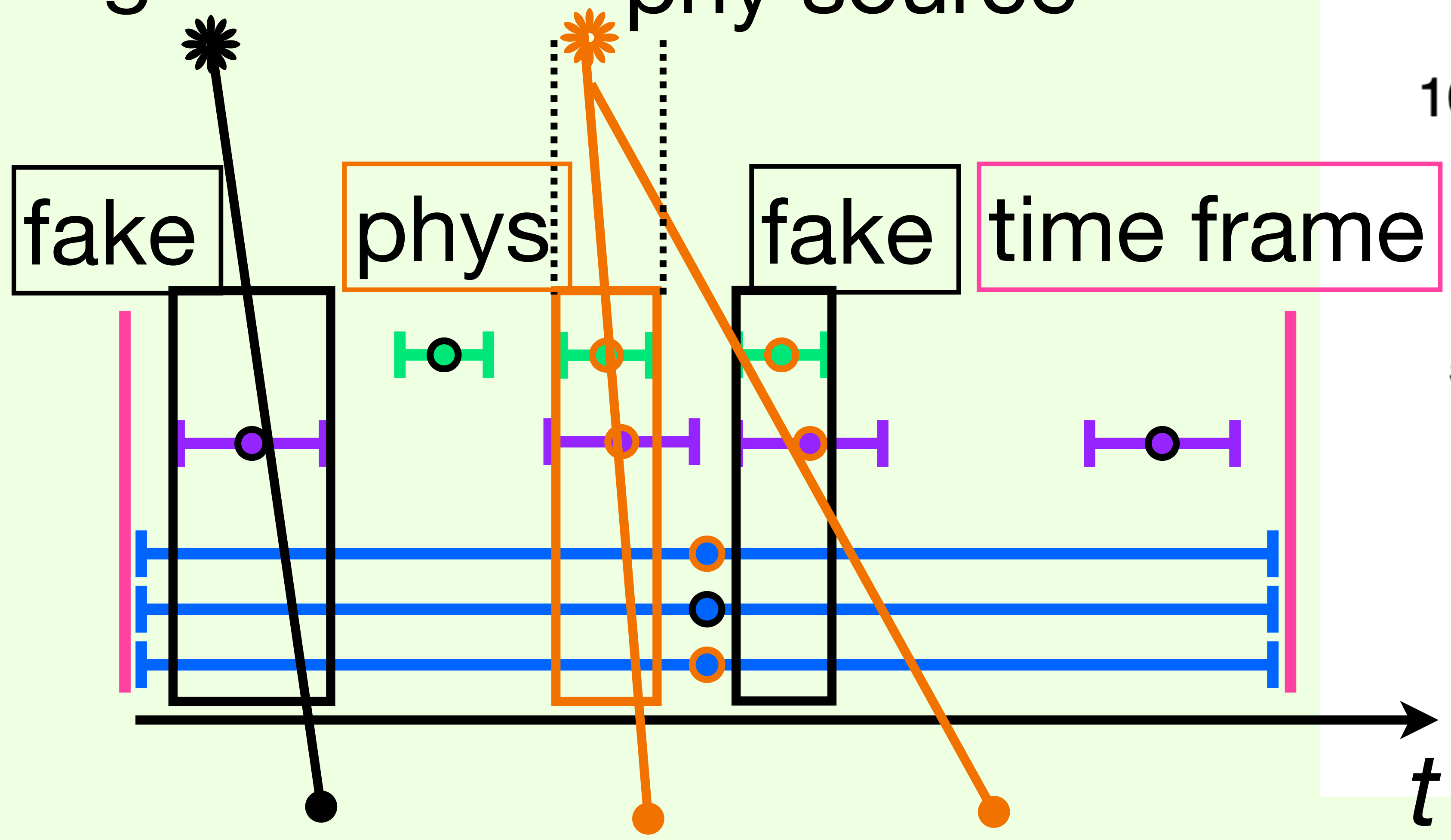
	DIS NC 18x275 Q^2 > 1 GeV^2 or 100 GeV^2 (Deep inelastic scattering neutral current)	500 kHz
	Synchrontron Radiation	14 MHz
	Electron bremsstrahlung radiation	317 kHz
	Electron Touscheck scattering (intrabeam dcattering)	1.3 kHz
	Electron Coulomb scattering processes	0.72 kHz
	Proton beam gas interactions	22.5 kHz

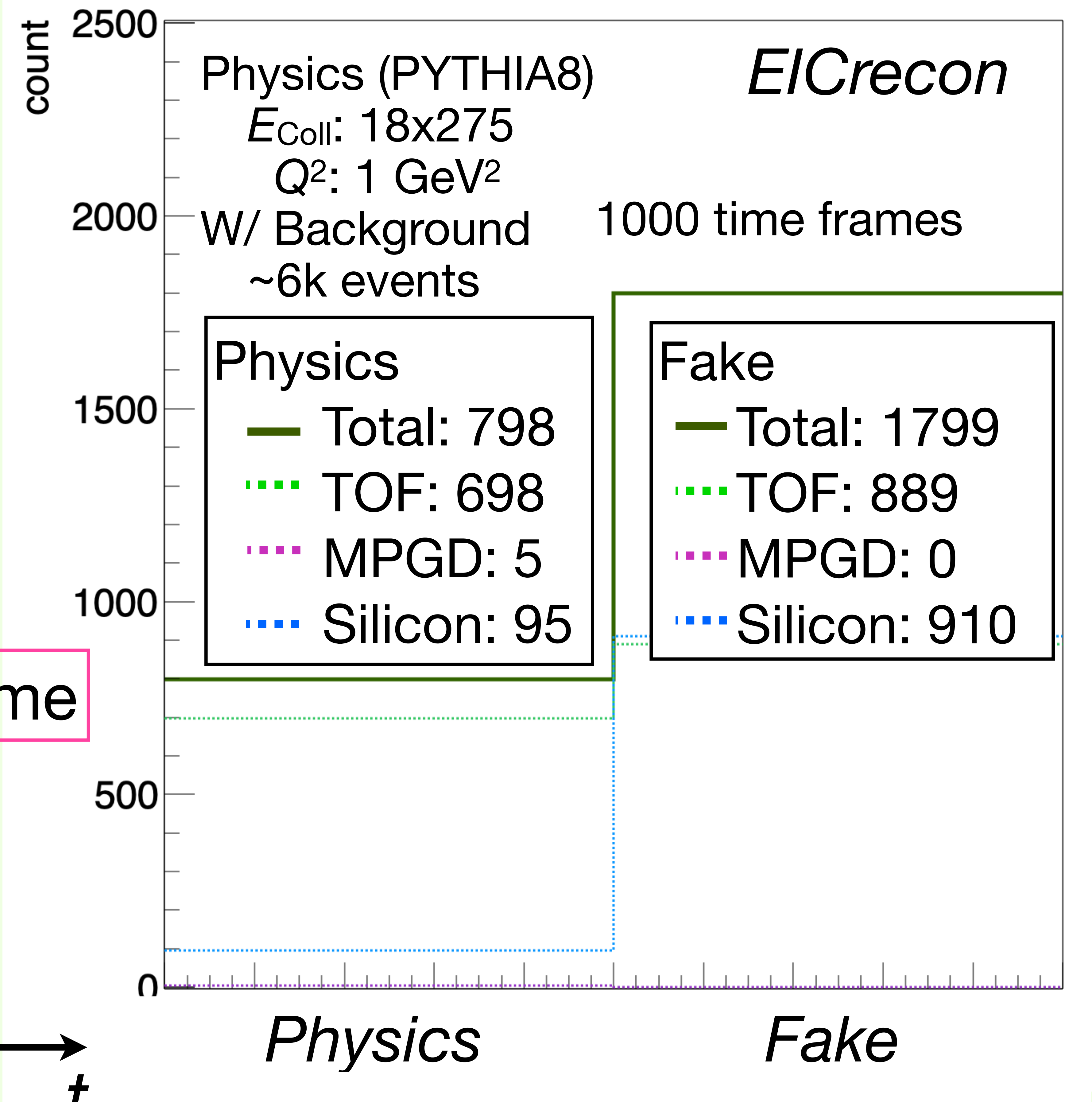
※ Very specific case to easy trigger (high Q^2).
In the real case, the high Q^2 events are lower than 500 kHz.

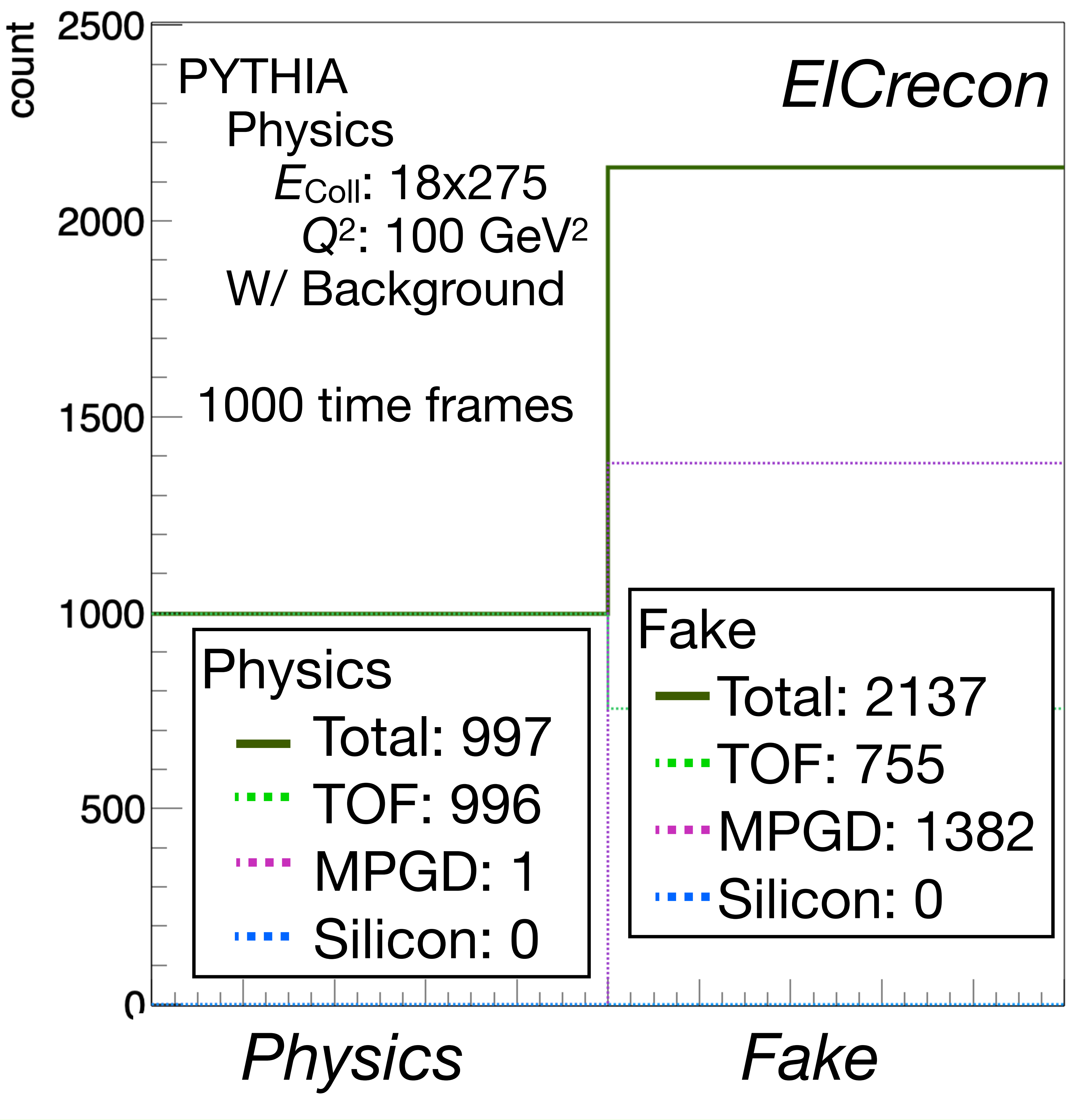
※ Recently we found the synchrotron radiation should be **14 MHz → 3.3 GHz**.

First Event Filter Efficiency

Time slice containing the production time of a physics
→ ✔ **physics trigger**
→ ❌ **fake triggers**







For Q^2 = 100 GeV^2 events, the first event filter efficiency is over **99%**.
→ Looks nice.

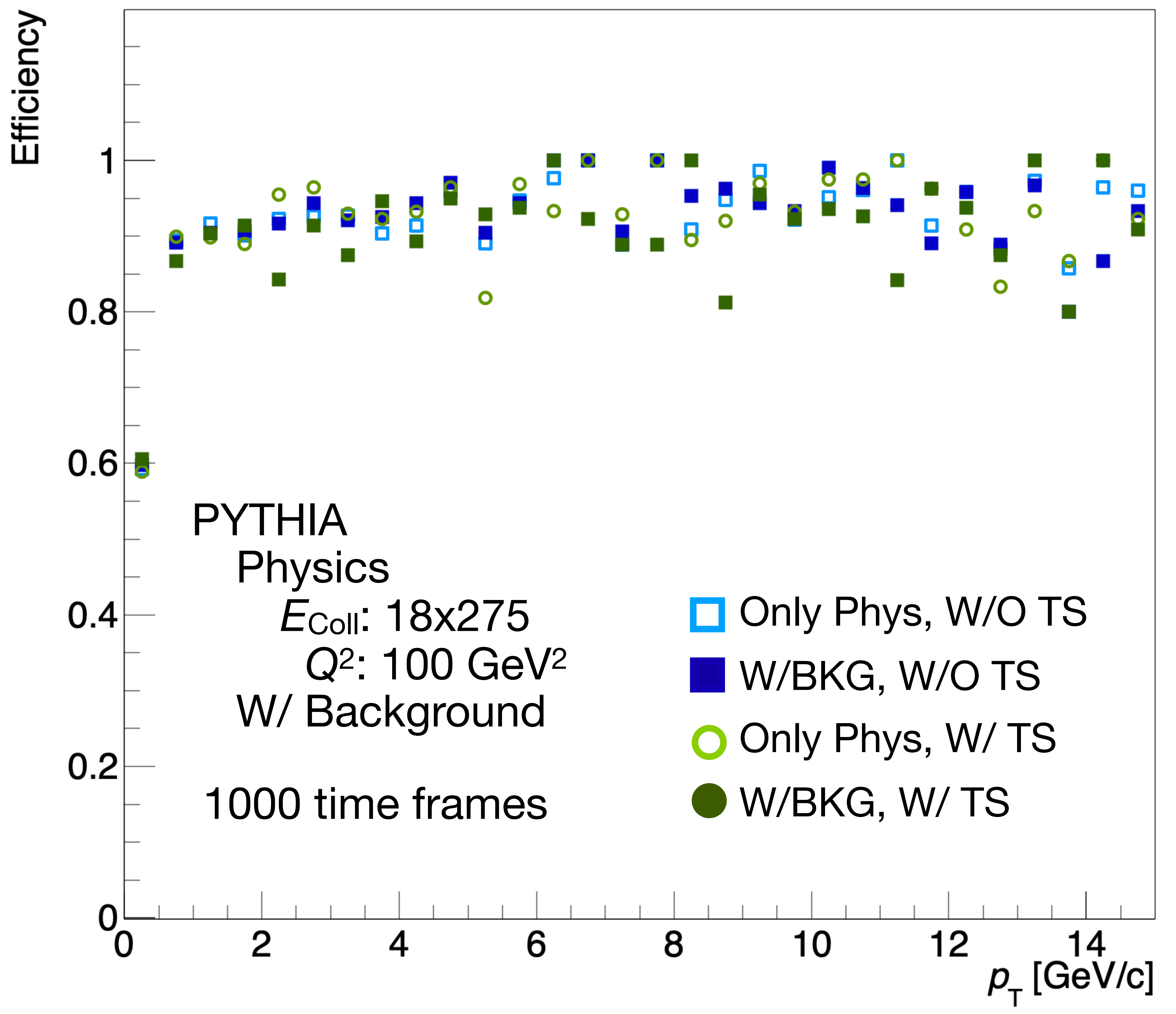
For Q^2 = 1 GeV^2 events, the first event filter efficiency is about **80%**.
→ This quality is not enough. → What type of events are missed? → Lower than Q^2 = 1 GeV^2?

Other Time splitting performances

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

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

→ Need to apply low Q^2 events / more high background simulation.



Outlook map

5. Test online chain (E0 - E1 orchestraton::connect detector/rucio) (**- Sep?)**

1. Test for **various** physics events (**low Q^2**) (**- March)**

3. Apply **multi-threads** (or GPU) (**- May**)

6. Apply detector response (???)

2. Up date the pre-filter using **all sub-detectors** (**- March)**

4. Implement "**Event filter**": Stringent event filter and event identification using **reconstructed tracks and vertex** information. (**- Jun)**

7. Add calibration stream line (**???)**

Outlook1: physics simulations

Physics simulation [git]: All ready is prepared.

BACKGROUNDS	Update background configurations (#5)
DDIS	Add DDIS configuration for rapgap3.310-1.0 ep noRad 10x...
DIS	Add BeAGLE1.03.02-2.0 DIS eAu configurations for 10x10...
EXCLUSIVE	Add lAger3.6.1-1.0 DVMP JPsi MuMu configurations for 10
SIDIS	Add DIJET dataset configuration for pythia6.428-dijet-v1..
SINGLE	Collect timing data for single neutrons
SR	nevents needs gzip

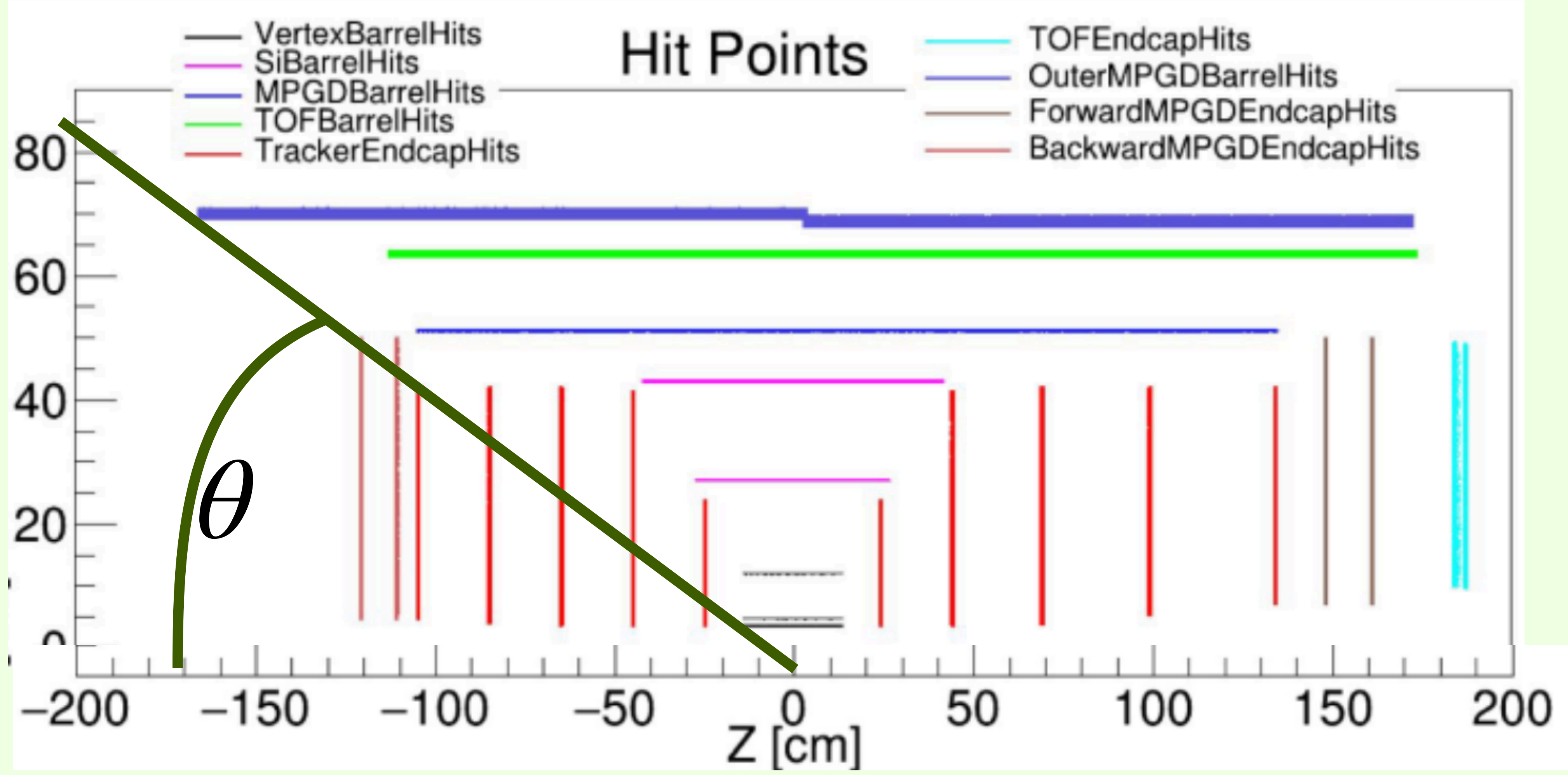
EXCLUSIVE physics

DDVCS	1.1.6-1.0 EpIC DDVCS and 1.3.0-1.0 eSTARlight Photoproduction JPSI...
DEMP	Add DEMPgen-1.2.4
DIFFRACTIVE_JPSI	1.1.6-1.0 EpIC DDVCS and 1.3.0-1.0 eSTARlight Photoproduction JPSI...
DIFFRACTIVE_PHI	New diffractive phi and diffractive rho samples (#16)
DIFFRACTIVE_RHO	New diffractive phi and diffractive rho samples (#16)
DVCS	Adding DVCS 10x130 version 1.1
DVMP	Add lAger3.6.1-1.0 DVMP JPsi MuMu configurations for 10x130, 10x2...
OMEGA	More and smaller chunks for EXCLUSIVE
PHOTOPRODUCTION_JPSI	1.1.6-1.0 EpIC DDVCS and 1.3.0-1.0 eSTARlight Photoproduction JPSI...
RHO	More and smaller chunks for EXCLUSIVE
SPECTROSCOPY	Split spectroscopy files

Most physics simulations are prepared.
→ Various DIS physics simulations have been prepared (W/ background, some beam energy)
→ **Need simulations of EXCLUSIVE/SIDIS for eicrecon.**

Outlook2: use all detector

The current trigger algorithm uses **only tracking detectors**, but it misses low-Q^2 events because the emitted particles do not produce hits.



Tracking detector

Silicon MAPS: ~2 μs [ref, p11]
MPGD: ~10 ns [ref, p11]
TOF: ~30 ps[ref, p11]

PID

hpDIRC: ~50 ps [ref, sec2]
dRICH: ~20 ps [ref, p2]
pfRICH / HRPPD / LAPPD: ~50 ps [ref, sec2]

Calorimeter

Barrel Imaging Calorimeter: ~ 3.25 ns [ref]
Backward/Forward EMCal, Barrel/Forward/Backward HCal: ?? s
Zero Degree Calorimeter: ~30ps [ref]

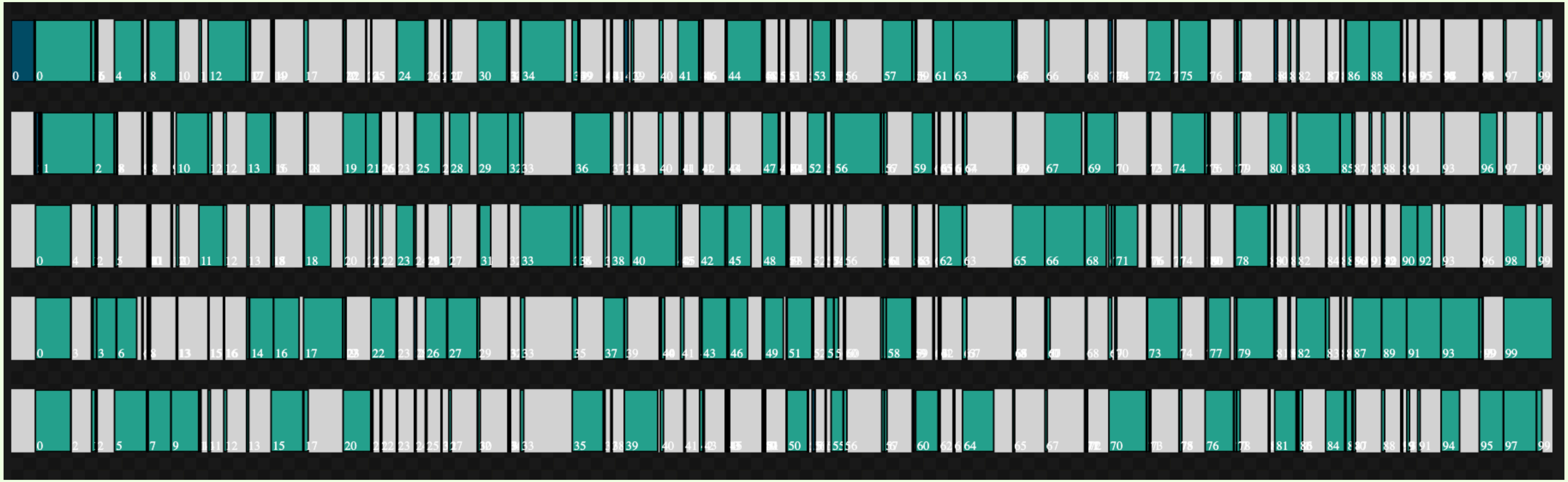
Far-forward

Roman Pots / Off-Momentum Detectors / B0: 30 ps [ref]

Outlook3: nthread test

nthread=1: 456.223 s / 100 time frame

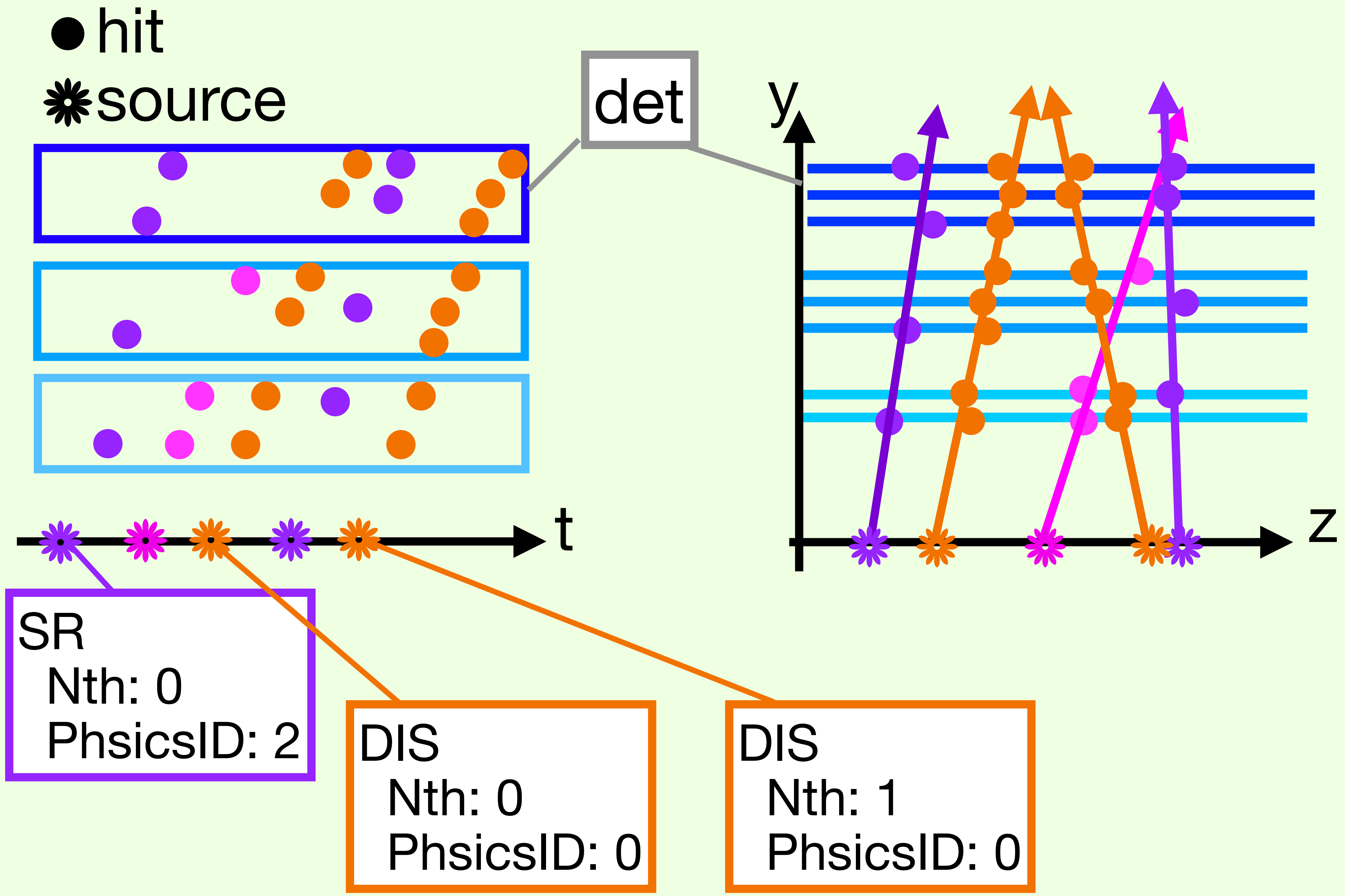
nthread=5: 541.52 s / 100 time frame



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

New Source Data Type Proposal

Currently we are missing some event source information (source time, numbering, Q^2…).
We have MC particles information and it has generation status to distinguish physics and background. However, this status cannot distinguish different sources of same type source.
→The proposing data type enable to distinguish sources and give valuable information.
And it is helpful for both MC and reconstruction (a little duplicate reco vertex).



✹ EventInfo ← NEW
- Nth
- CollisionTiming
- CollisionEnergy
- CollisionPosition
- CollisionAngle
- Q2
- PhysicsID

background →

[DIS, EXCLUSIVE, SR, etc…
0 1 2 …]

● DetRecHits
- ⋮
- EventAssociation ← NEW

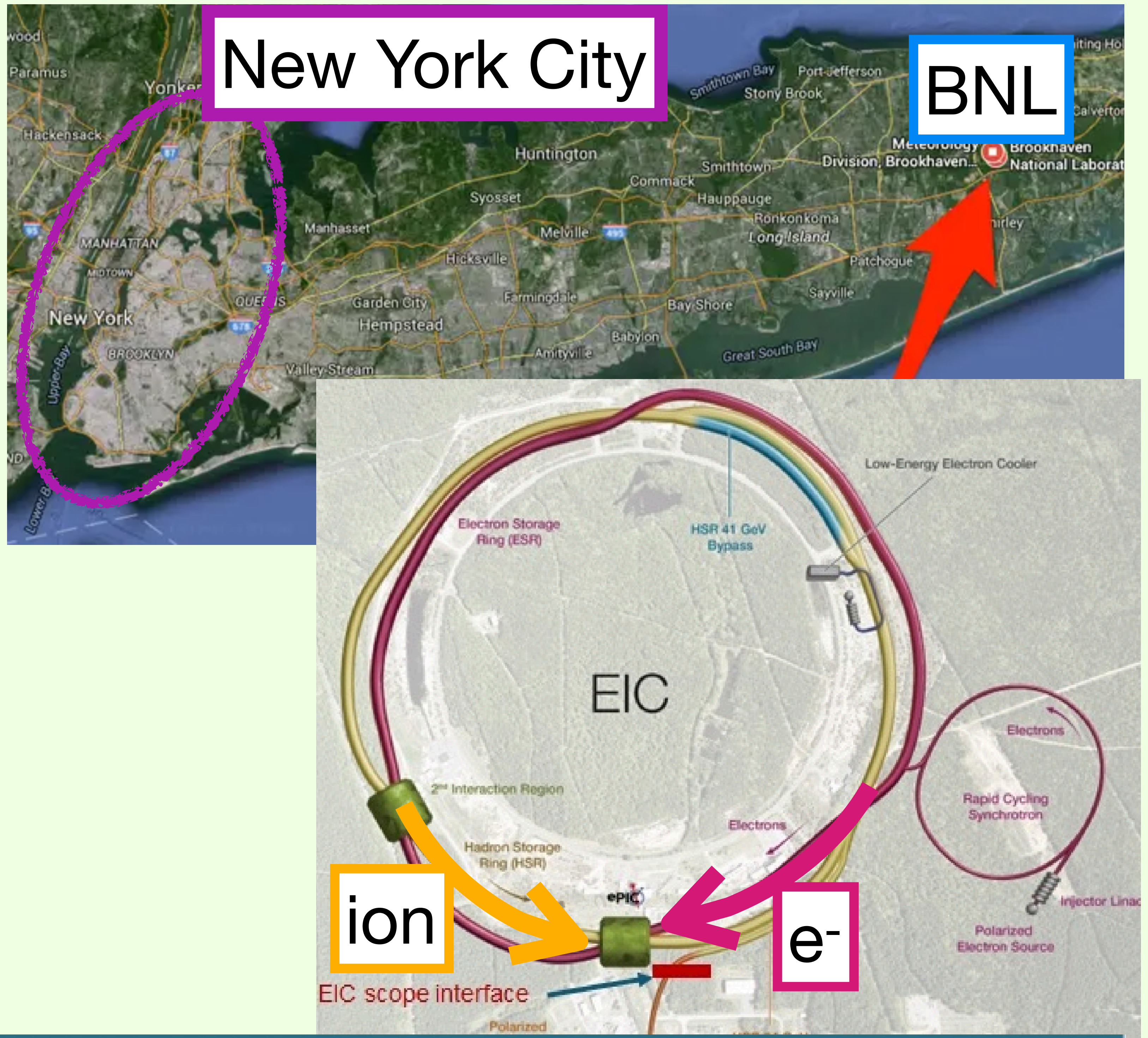
Backup Slides

Electron-Ion Collider

EIC (Electron Ion Collider) (2035-)

New accelerator of Brookhaven National Laboratory(BNL)

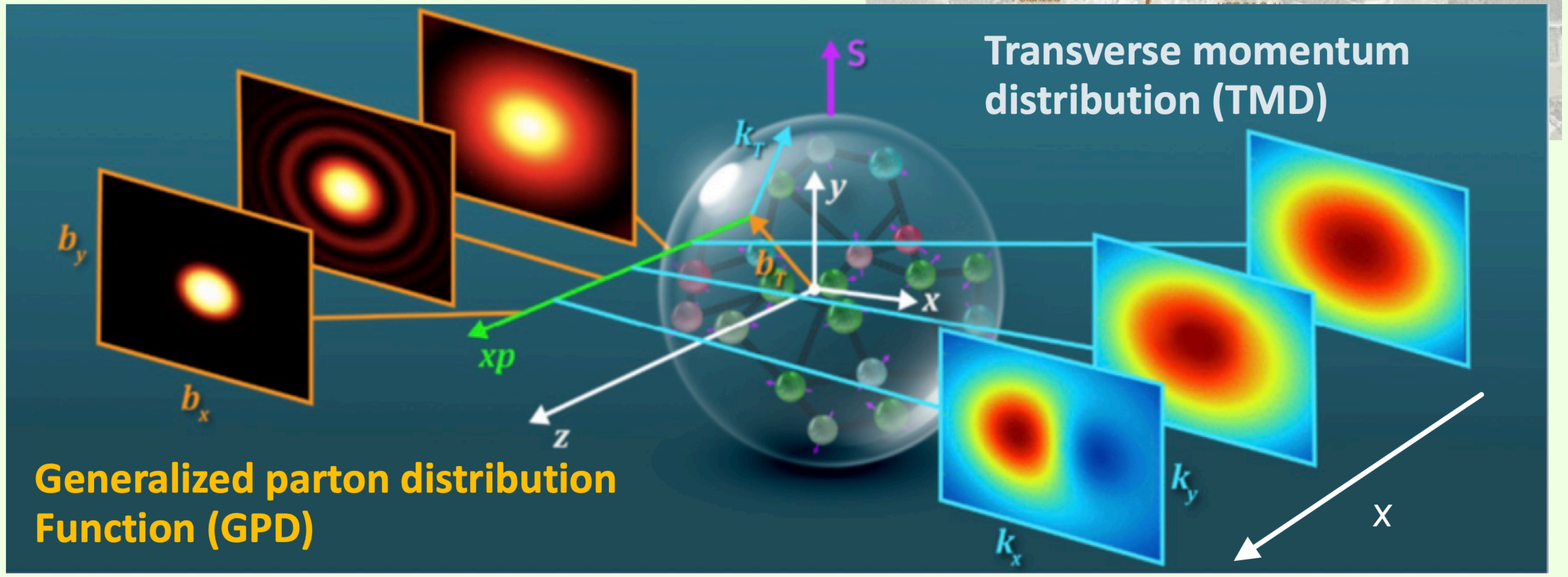
- **Characteristics**:
  - Electron beam **polarization**(80%) 5-18 GeV
  - Proton beam **polarization**(70%) 40-275 GeV
  - Ion beam: d/He/…/Au/U, 40-110 GeV/u
  - Central Mass Energy: $\sqrt{s}$ = 29–140 GeV
  - 100-1000 x HERA luminosities:10^{33}-10^{34} $cm^{-2}s^{-1}$



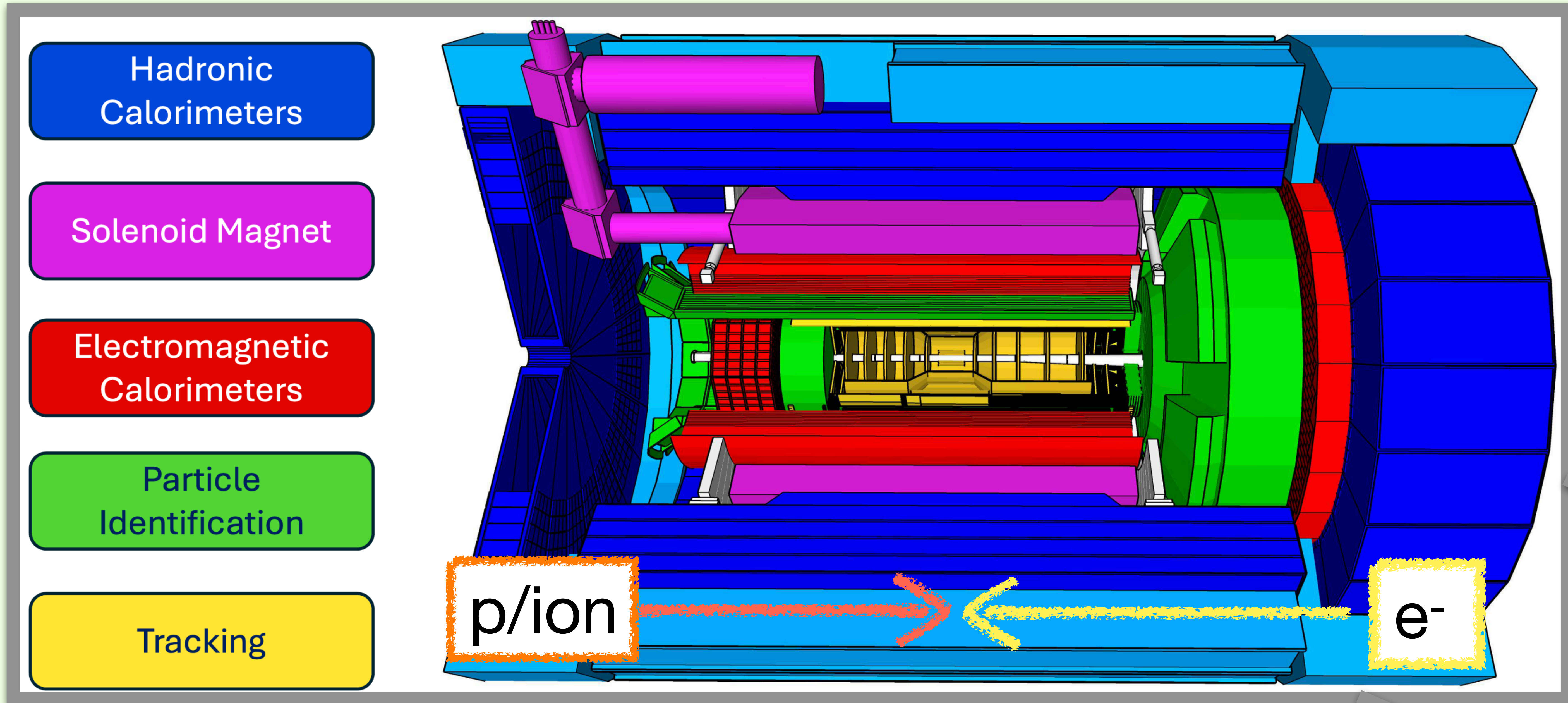
Main Physics Targets:

Nucleon and Nucleus structure using polarized e+p and e+A collisions over the wide beam energy

- How does the **mass** of the nucleon arise?
- How does the **spin** of the nucleon arise?
- What are the emergent properties of dense systems of gluons?



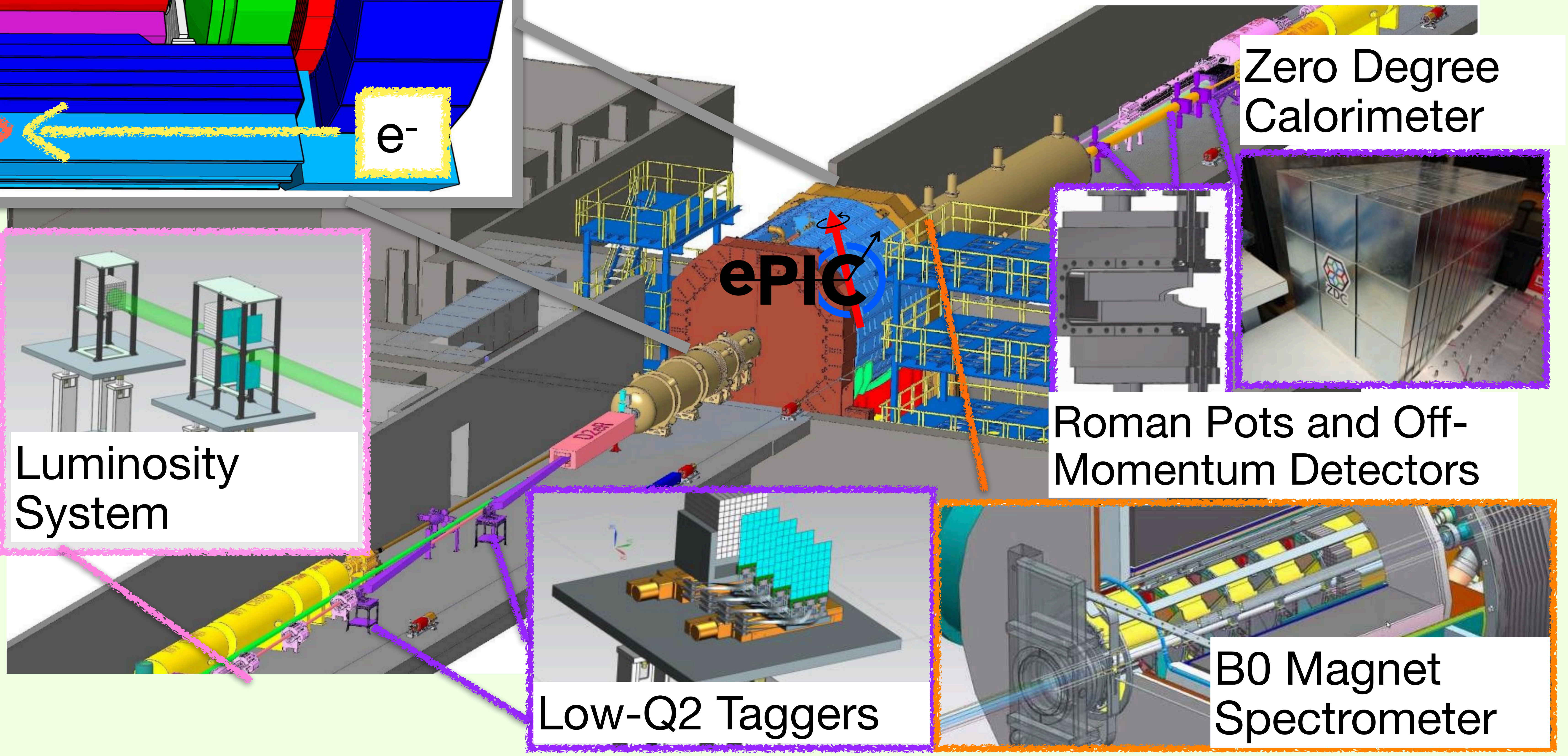
ePIC (electron-Proton/Ion Collider) detector



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)

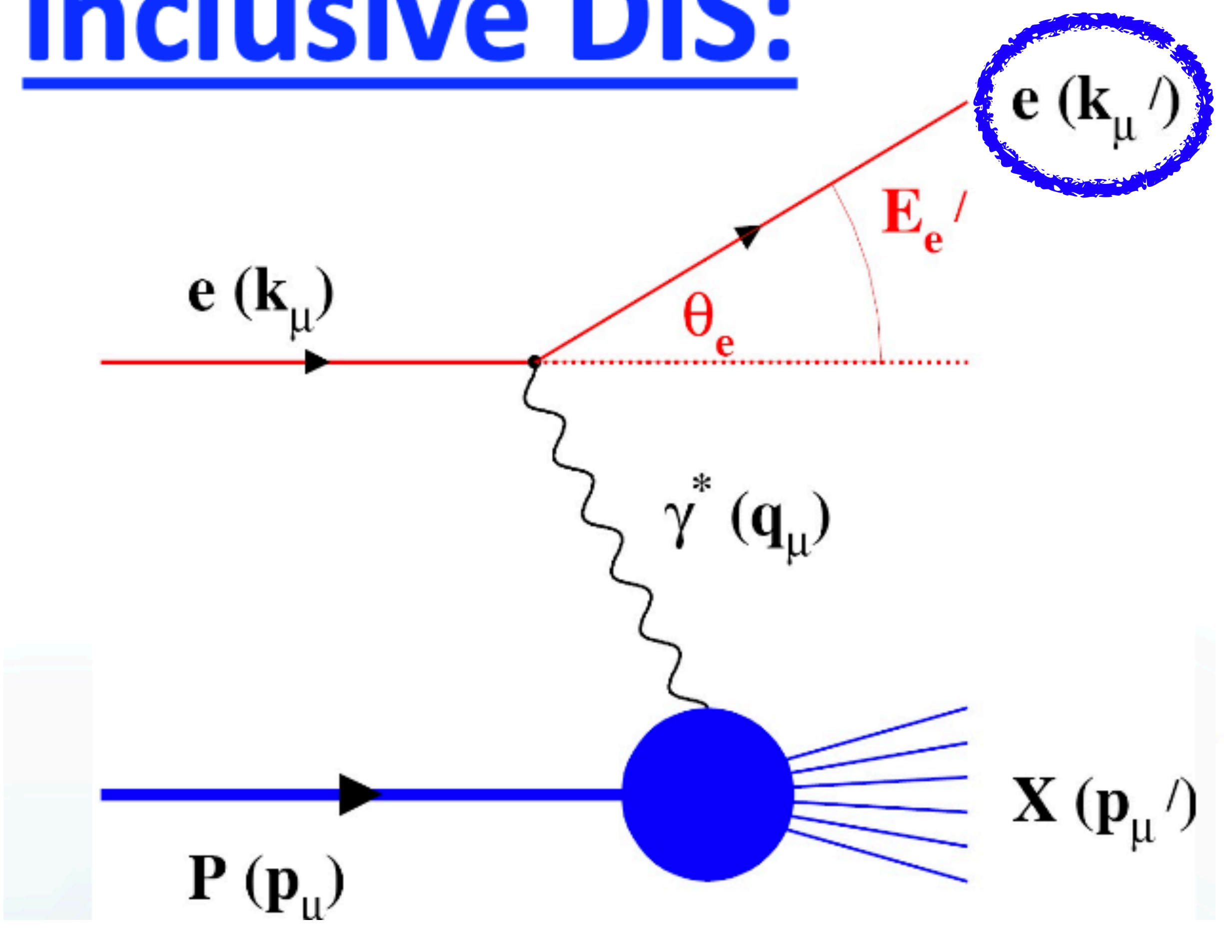
Tracking:
- New 1.7 T solenoid
- Si MAPS tracker
- MPGDs (μRWELL/μMegas)

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



ePIC Detector Concept

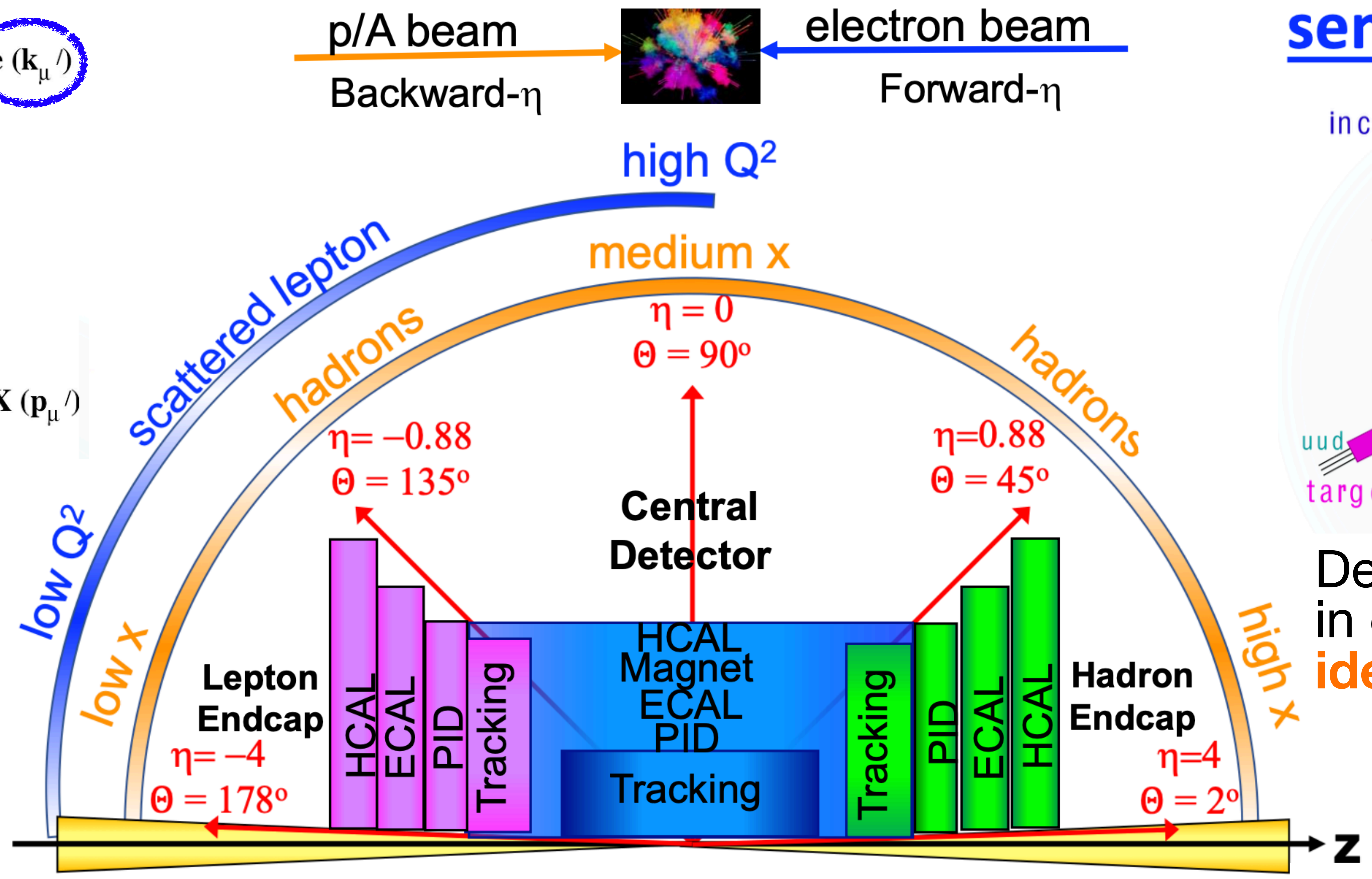
inclusive DIS:



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

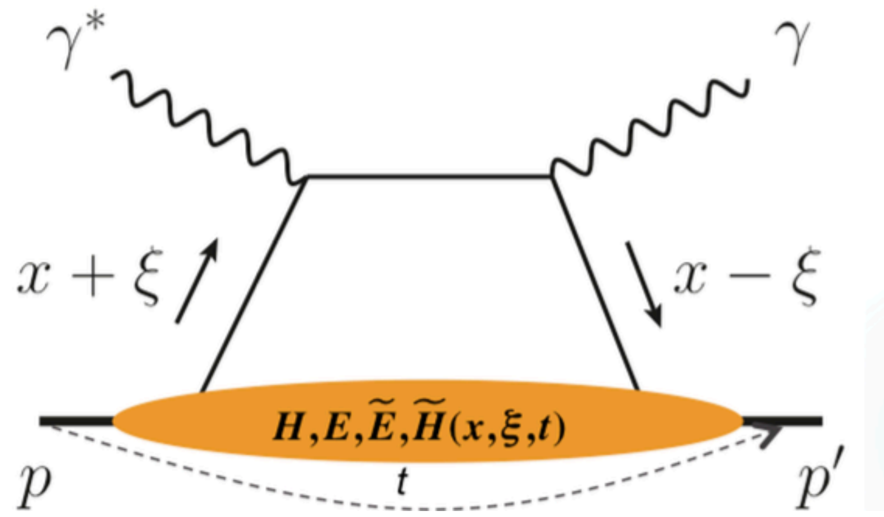
Very low Q^2 scattered lepton Bethe-Heitler photons for luminosity

Luminosity Detector
Low Q^2-Tagger

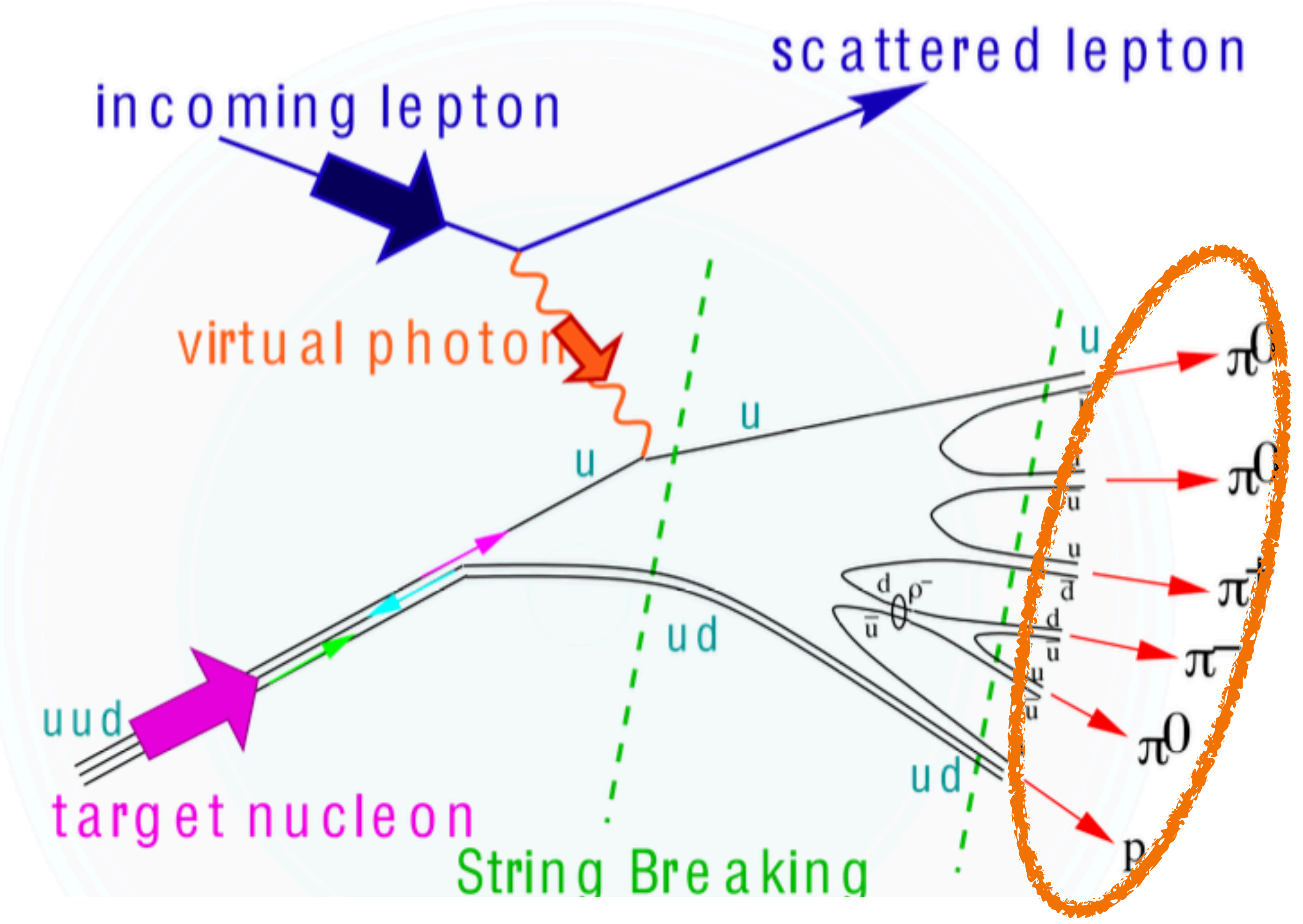


exclusive DIS

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



semi-inclusive DIS



Detect the scattered lepton in coincidence with **identified hadrons/jets**

Particles from nuclear breakup and from diffractive reactions

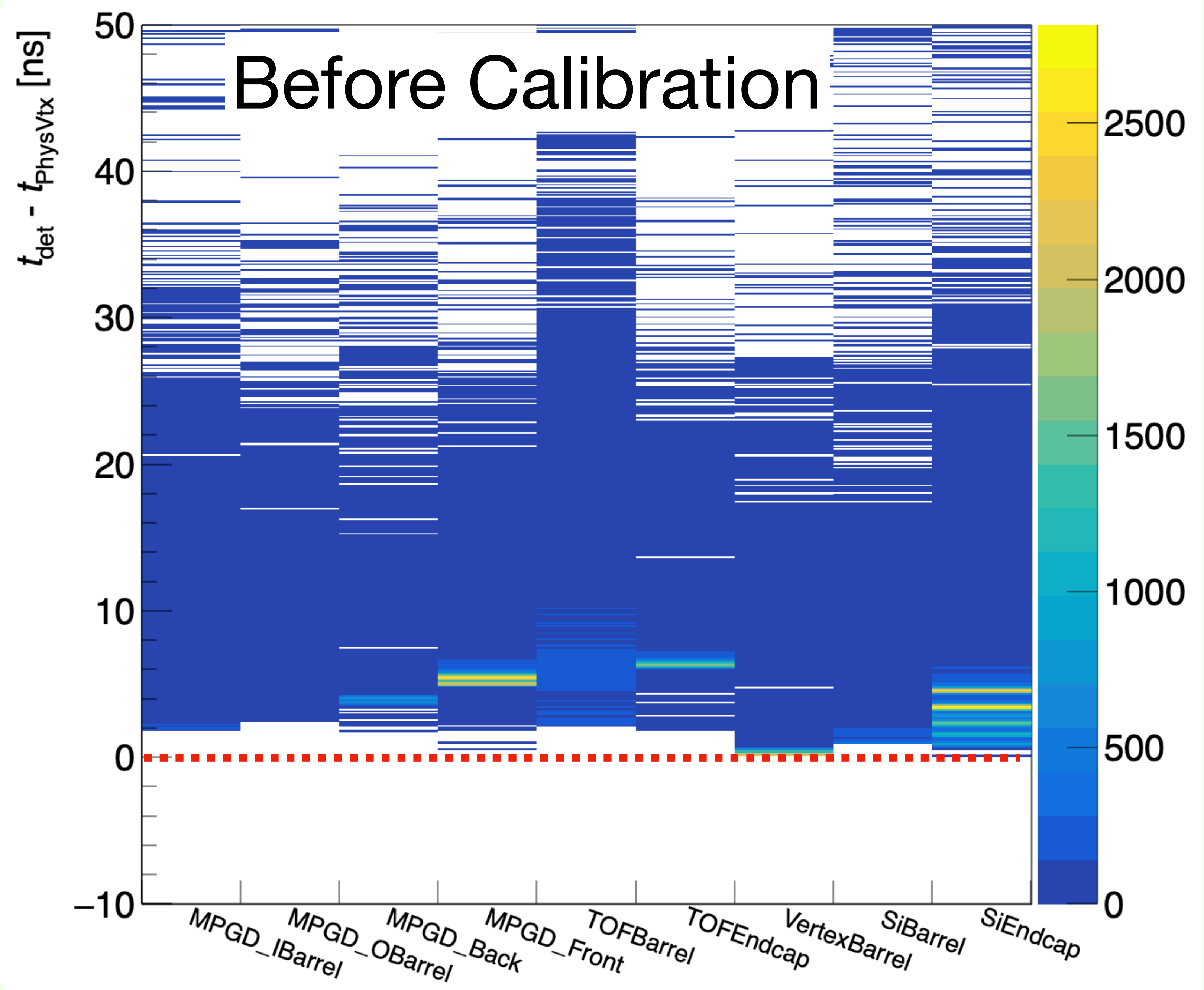
ZDC

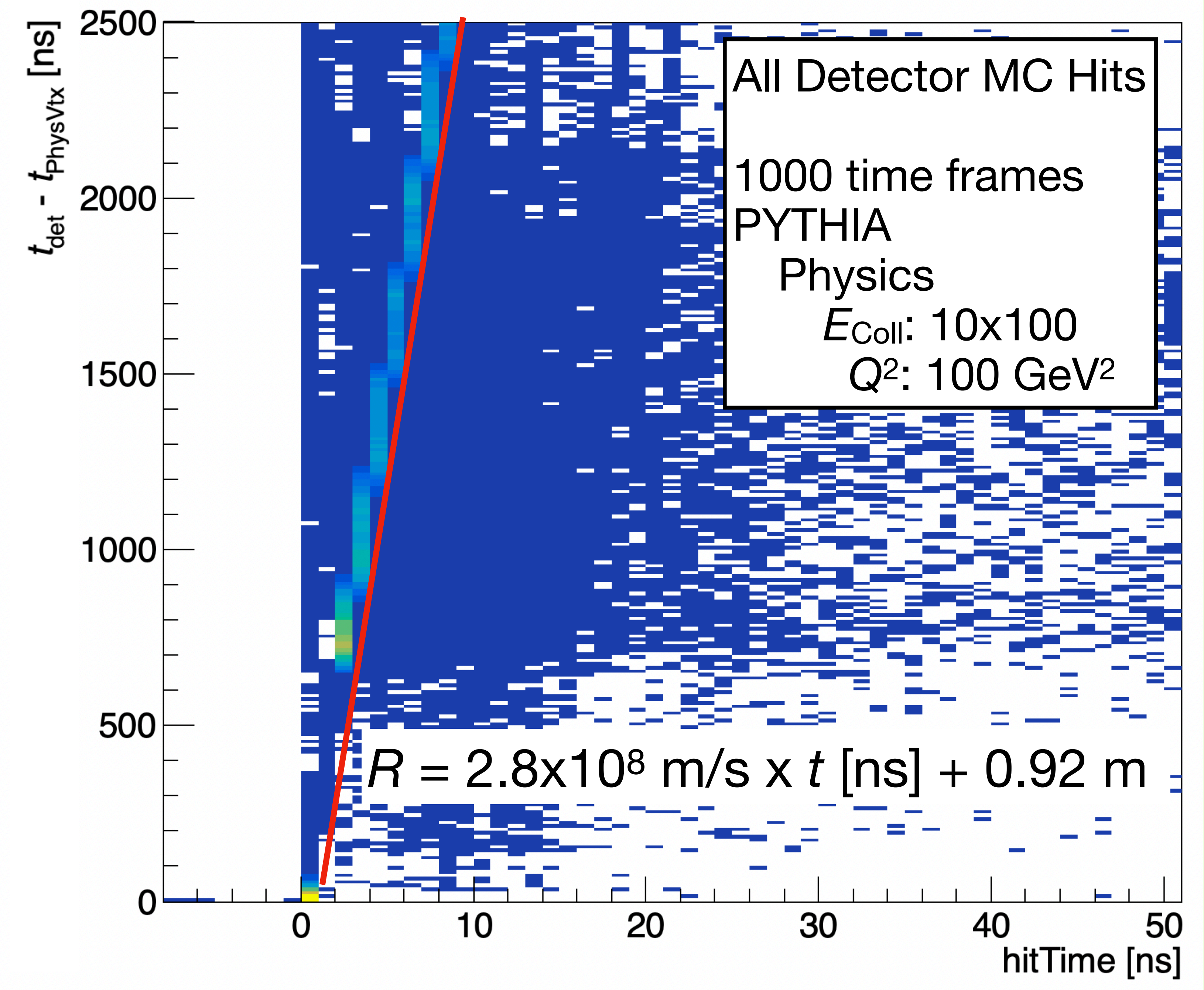
Forward Tracking

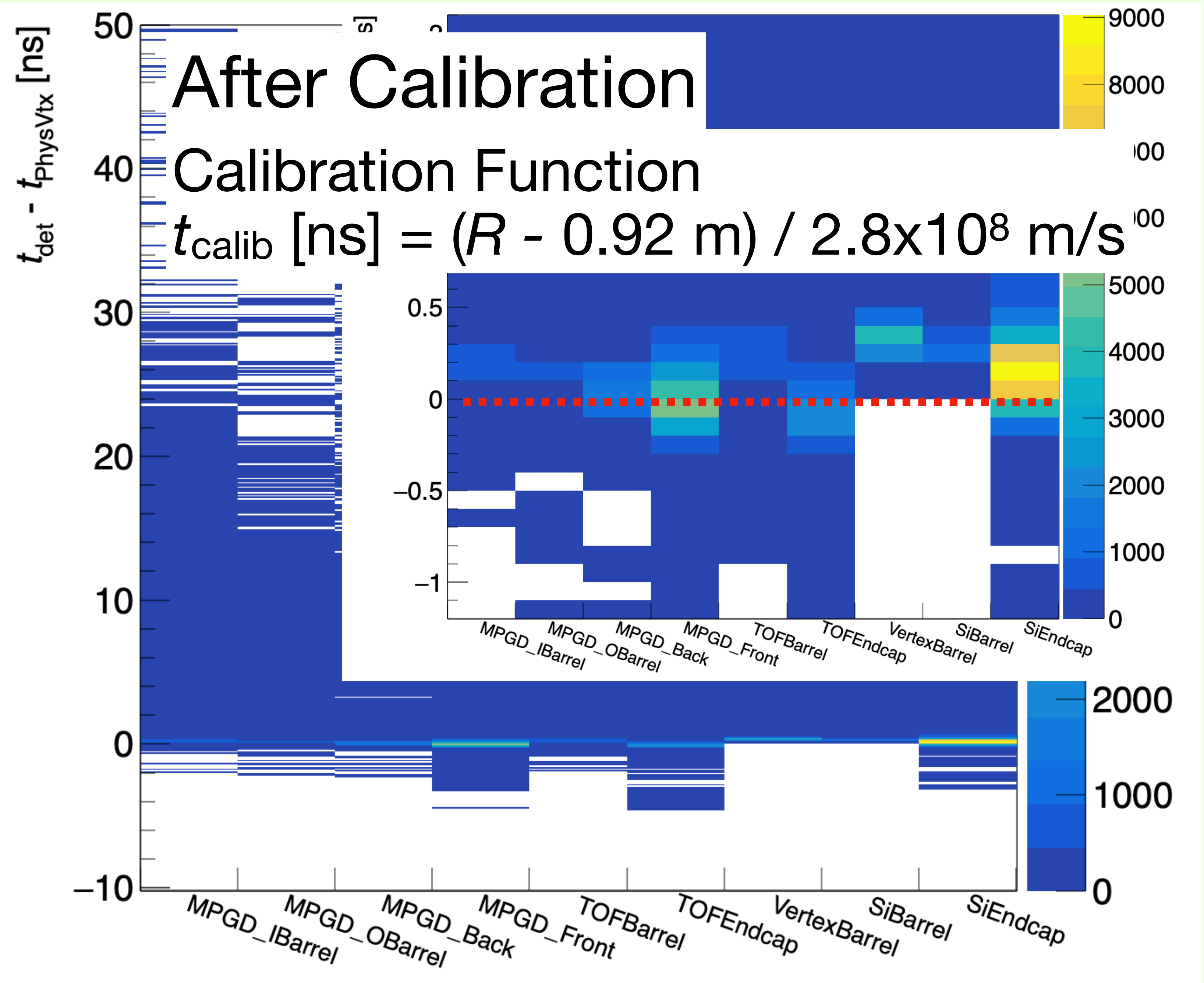
Time Alignment

Create Factory to do time alignment

Original input data **Hit time in a time frame: 13.1356, 1.12963, 1.2371, 1.03809, 1.02581, 24.2638, 10.5798, 9.64017, 1.77883, 1.93878, 1.63005, 1.61152**





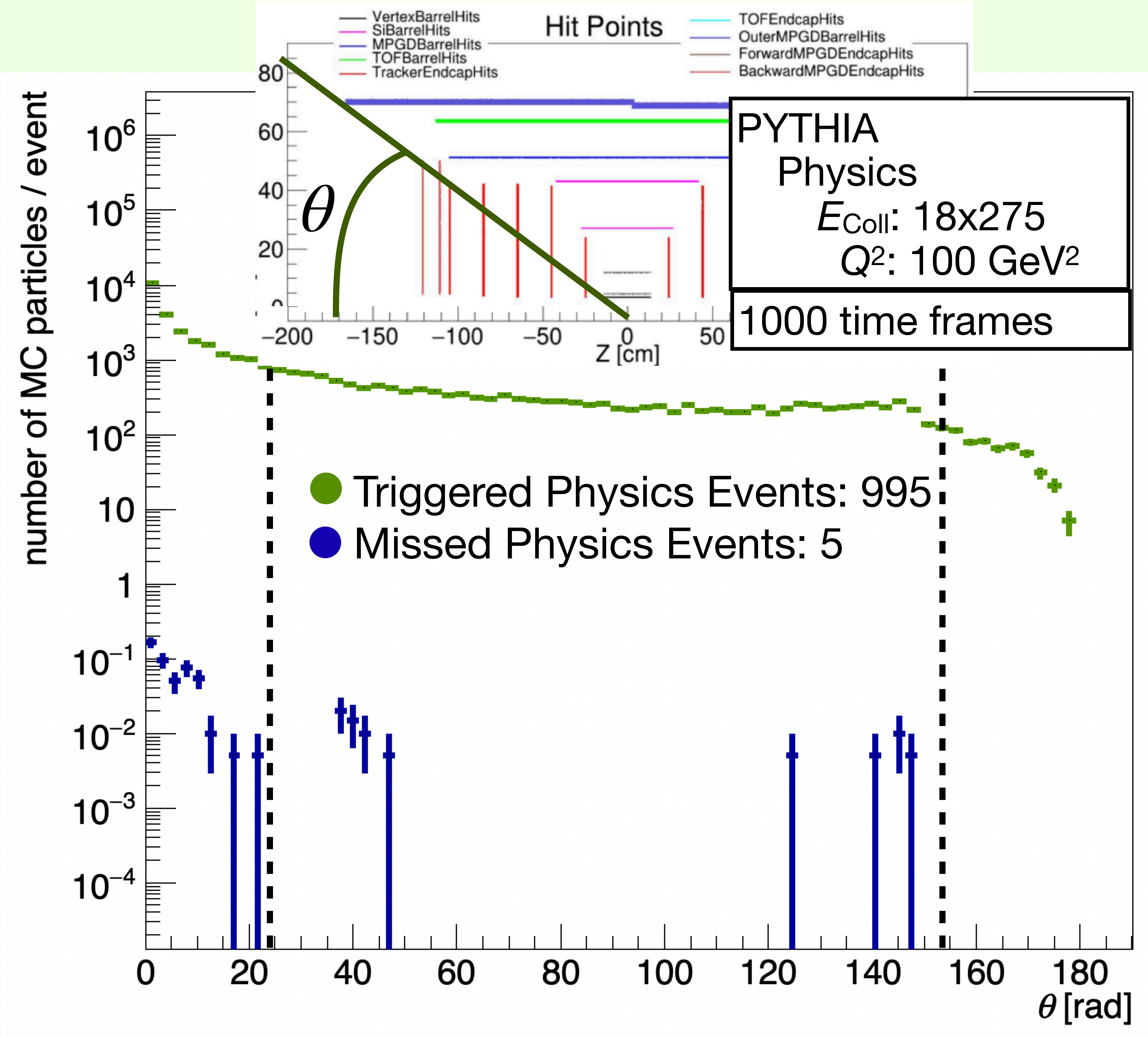


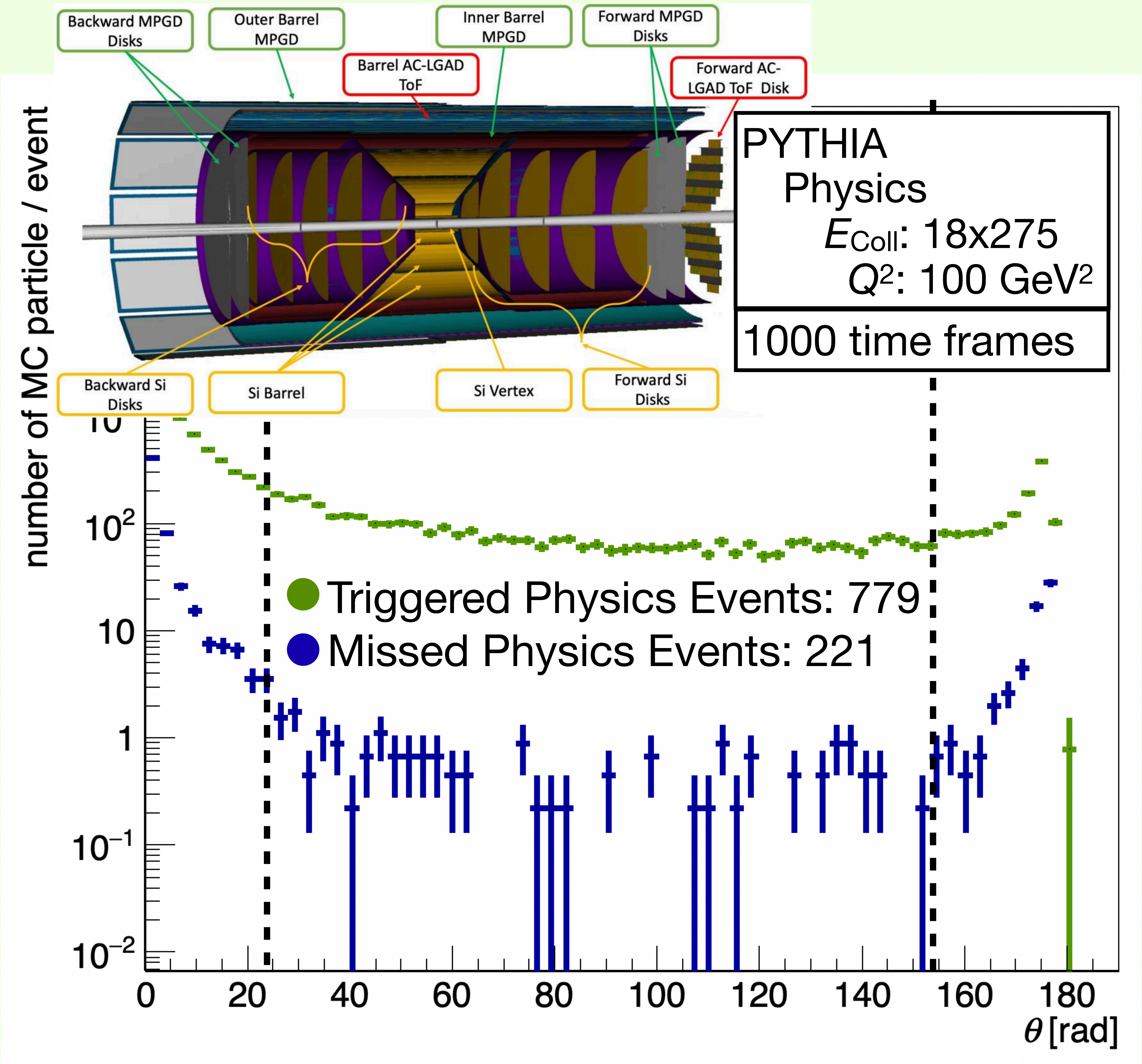
Sort + time alignment **Hit time in a time frame: 0.863942, 0.867679, 0.889219, 0.943039, 1.06582, 1.37656, 1.41053, 1.49302, 1.5067, 1.66467, 1.68391, 1.76771**

It seems work well.

→ **Future plan: Time response of detectors will be included.**

MC particles scattering angle distribution





The number of hits in the vicinity of the barrel region in the missed event is very small.

Total Process Time W/O and W/ TimeSplitter

1000 time-frame

W/O TimeSplitter: **1.7 h** → W TimeSplitter: **1.5 h**

The processing time with the time splitter is **faster** than without it (~10%).

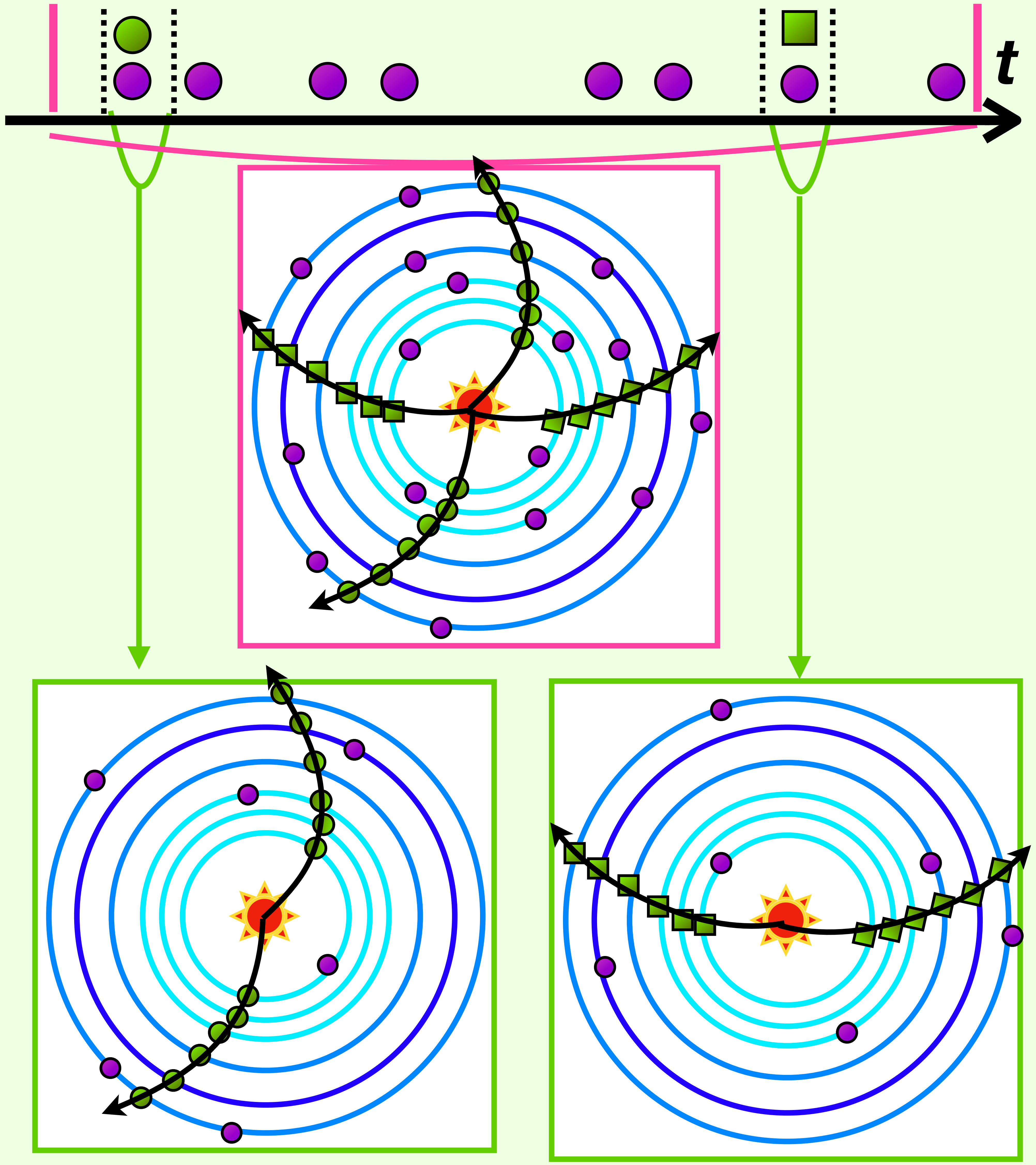
→ This is expected because the splitter reduces background hits and helps the tracking.

Single event with large number of hits

W/O TS: **468 s** → W/ TS: **242 s**

→Particularly, this splitting works well for large number of hits.

It is expected it become more faster using multi-threads or GPU.



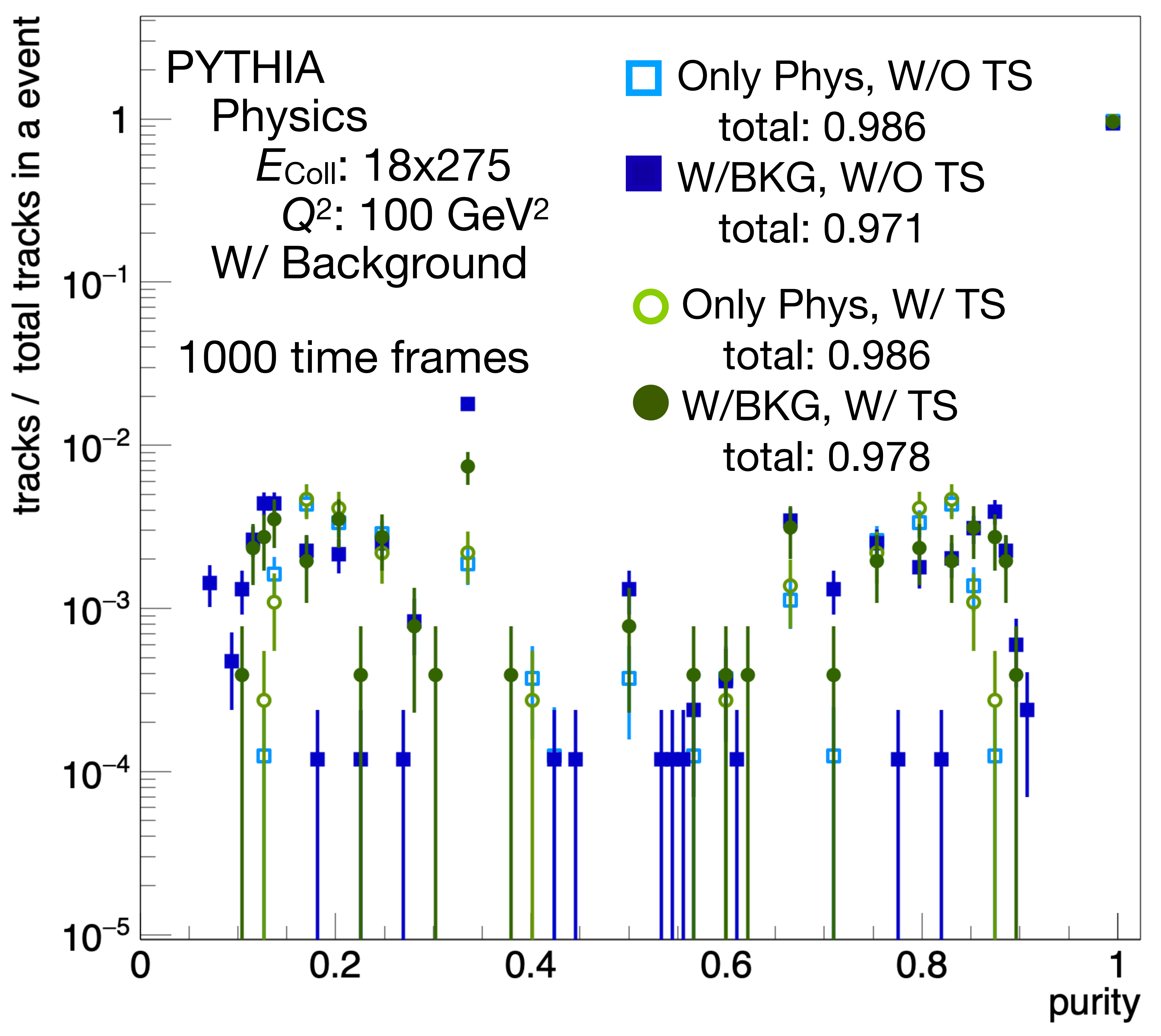
Reconstructed Tracks’ and Vertices’ Parameter Comparison

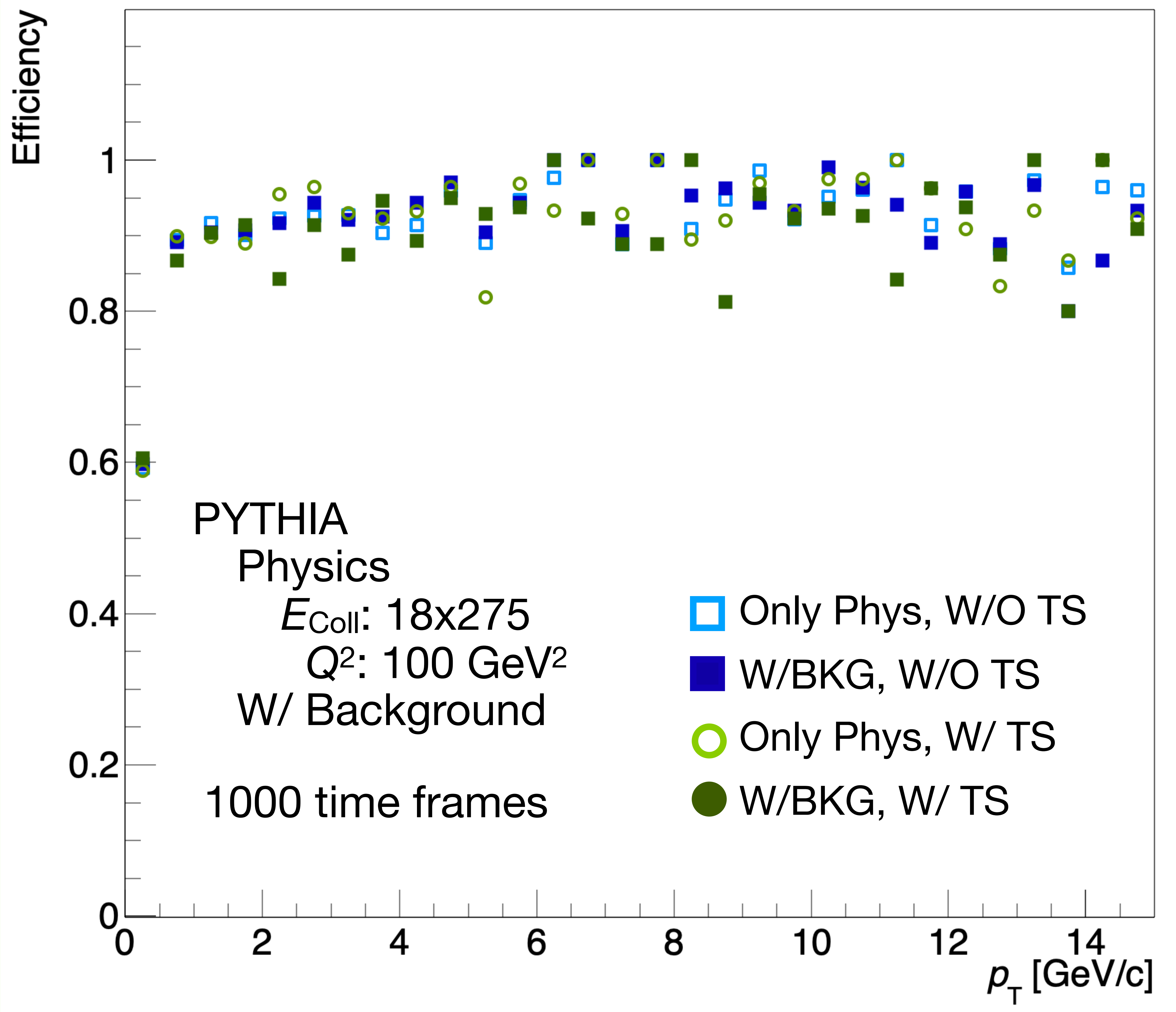
ACTS’s tracking and vertex output comparison

Overall, the results with time splitting are consistent with those without it.

For both the tracking χ^2 and the vertex χ^2, the time-split results show **better** performance than the no-splitting case.

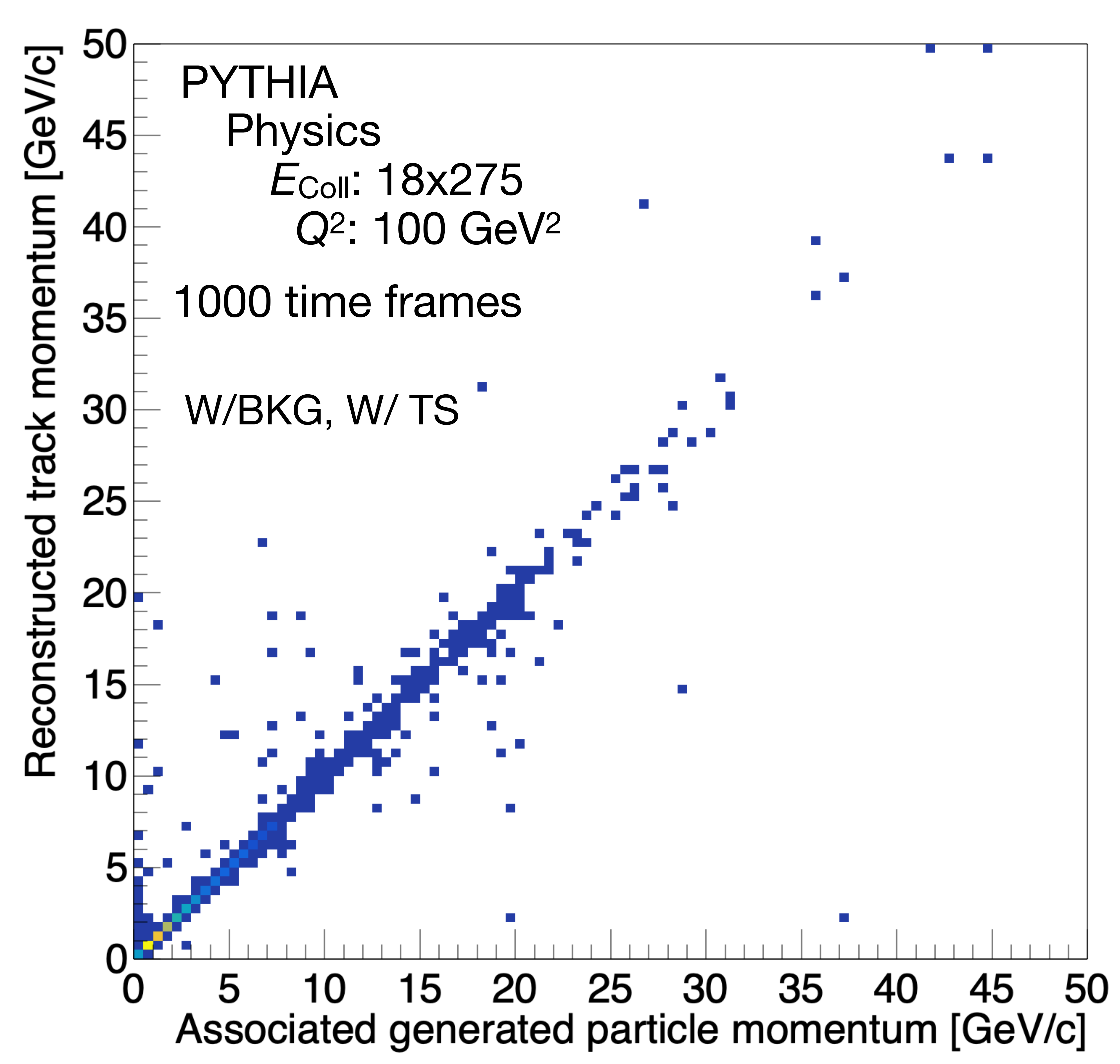
Track Purity and Efficiency Evaluation

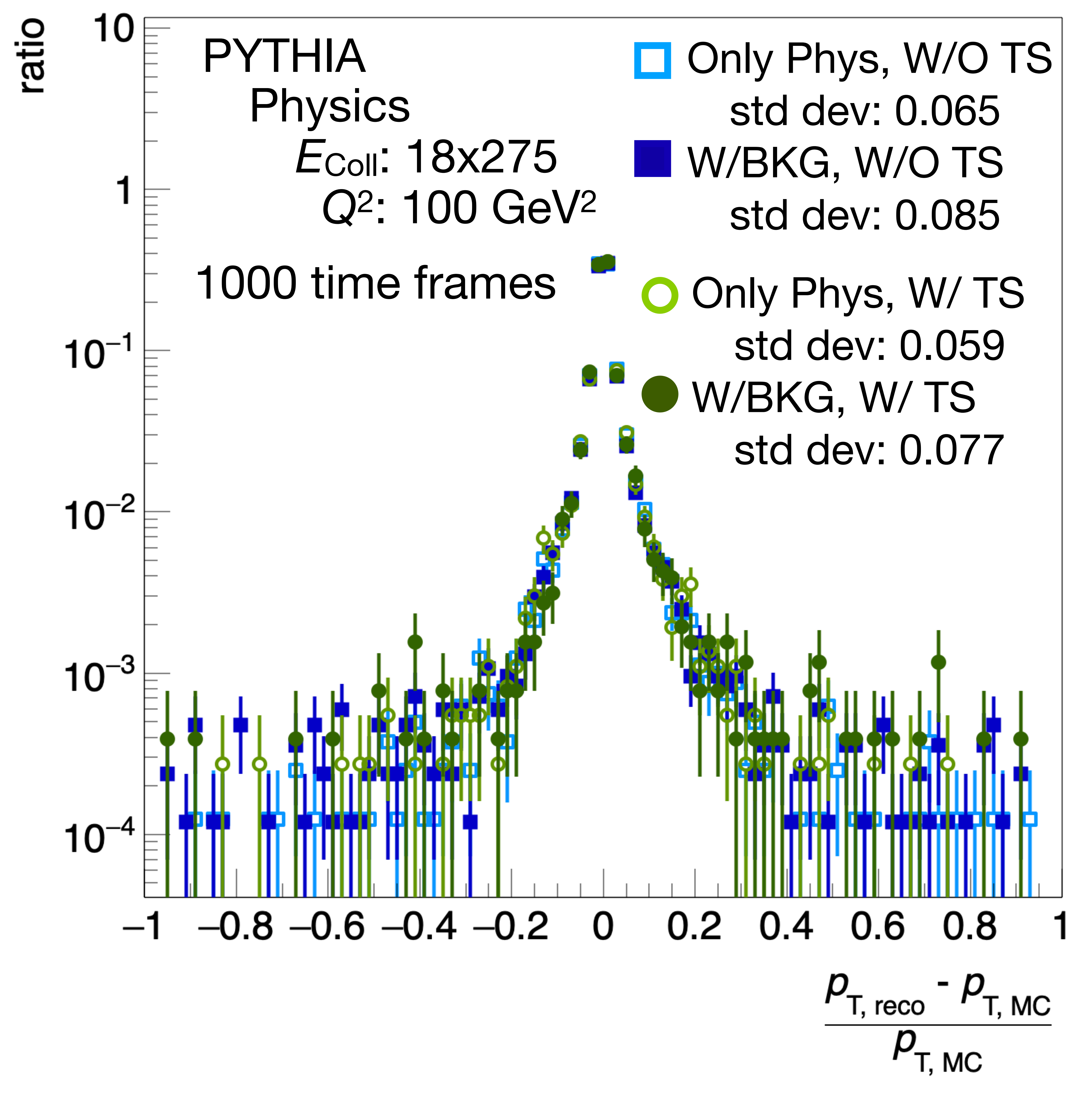




The purities of W/ time slice were a **little better** than W/O ones.
The efficiency of W/ time slice were mostly consistent with W/O ones, but a **little worse**.

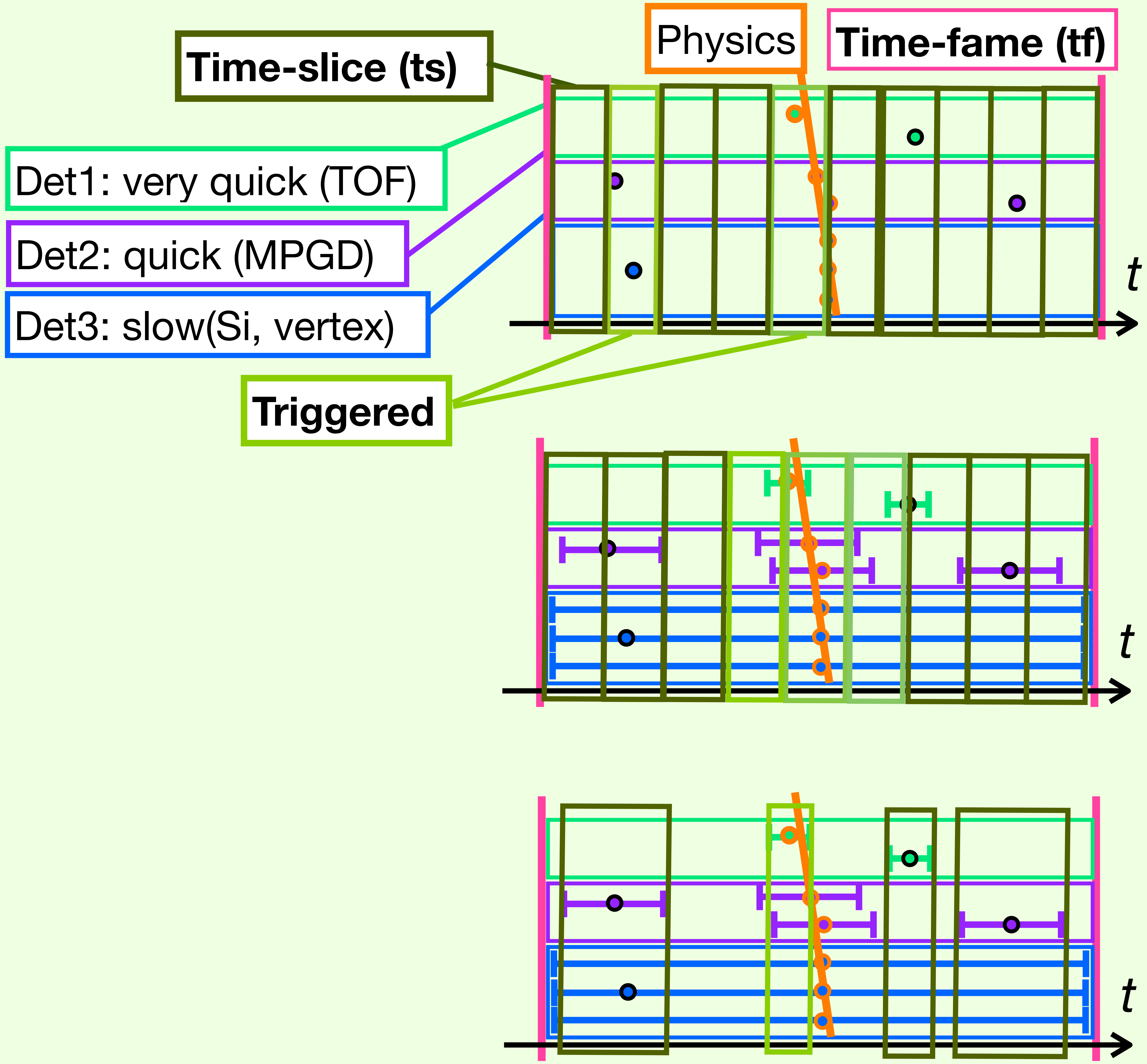
Track p_T Resolution





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.

Time slice algorithms



ver1: Slice a time frame by **constant time** slices
pros: very simple
cons: does not consider the detectors' time resolution.
does not considering hits' topology

ver2: Slice a time frame by **constant time** slices
pros: very simple
considering hits' topology
cons: there are many duplicated time slices including same hits and there are many fake times slices

ver3: Slice a time frame by **detector resolution base**
pros: realistic and sophisticated time slice.
considering hits' topology.
cons: a little complicated.

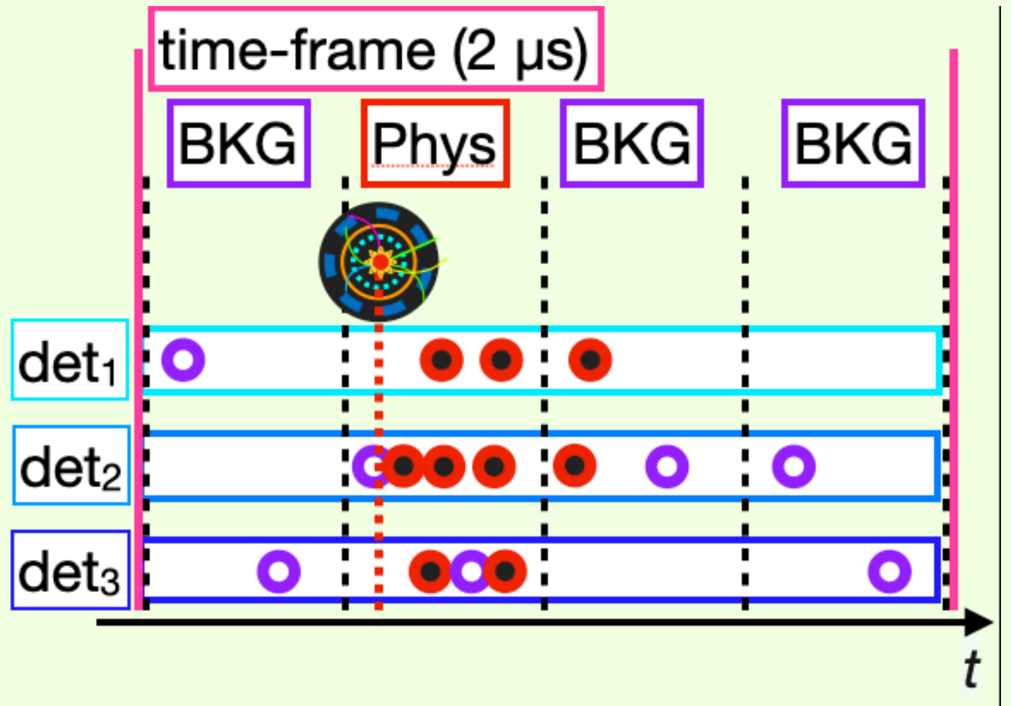
Tracking efficiency for each time slice algorithms

Ver.1
phys:
992/1000
fake: 411

The last collaboration meeting

- All time slice width could capture over 98% physics events.
→ Wider time-window can recover edge events

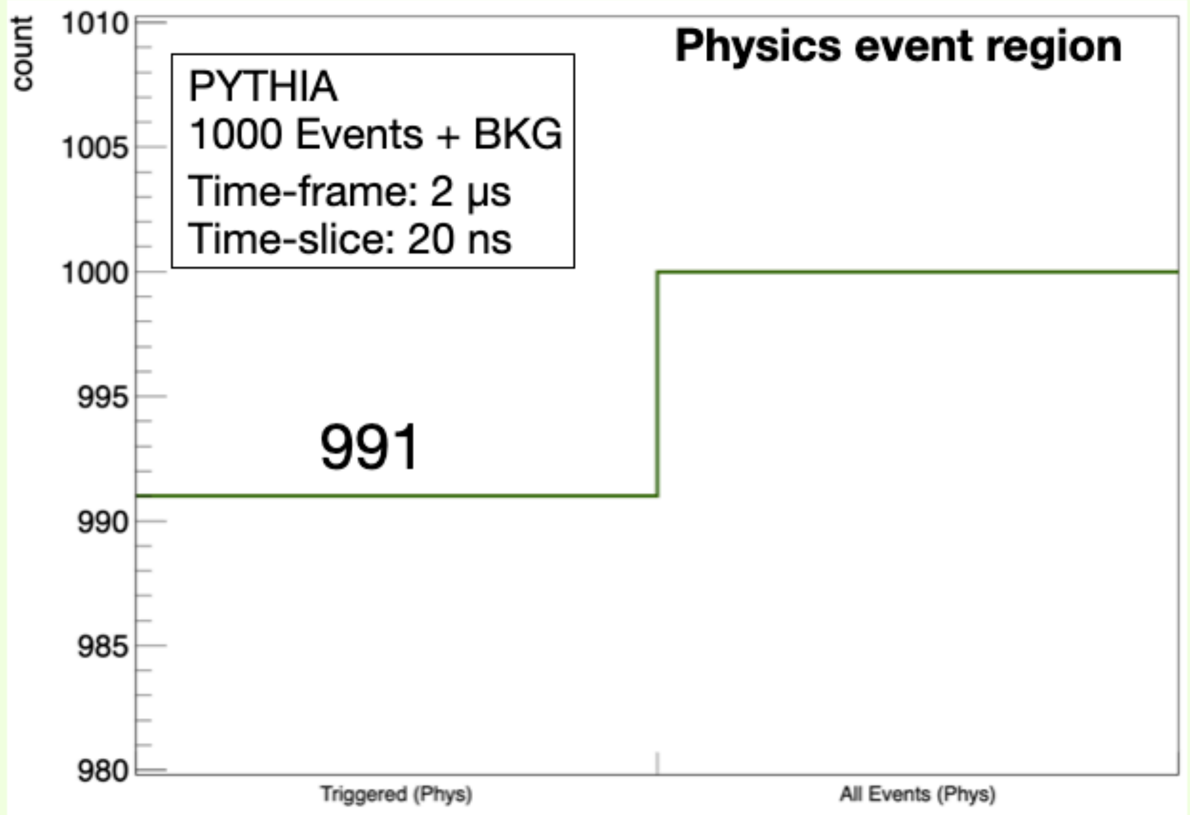
- The background reduction rates for all setting are not significantly changed.
Dominant background event is x28 # of physics.
→ **117/28000 → 0.4% (99.6% rejection)**

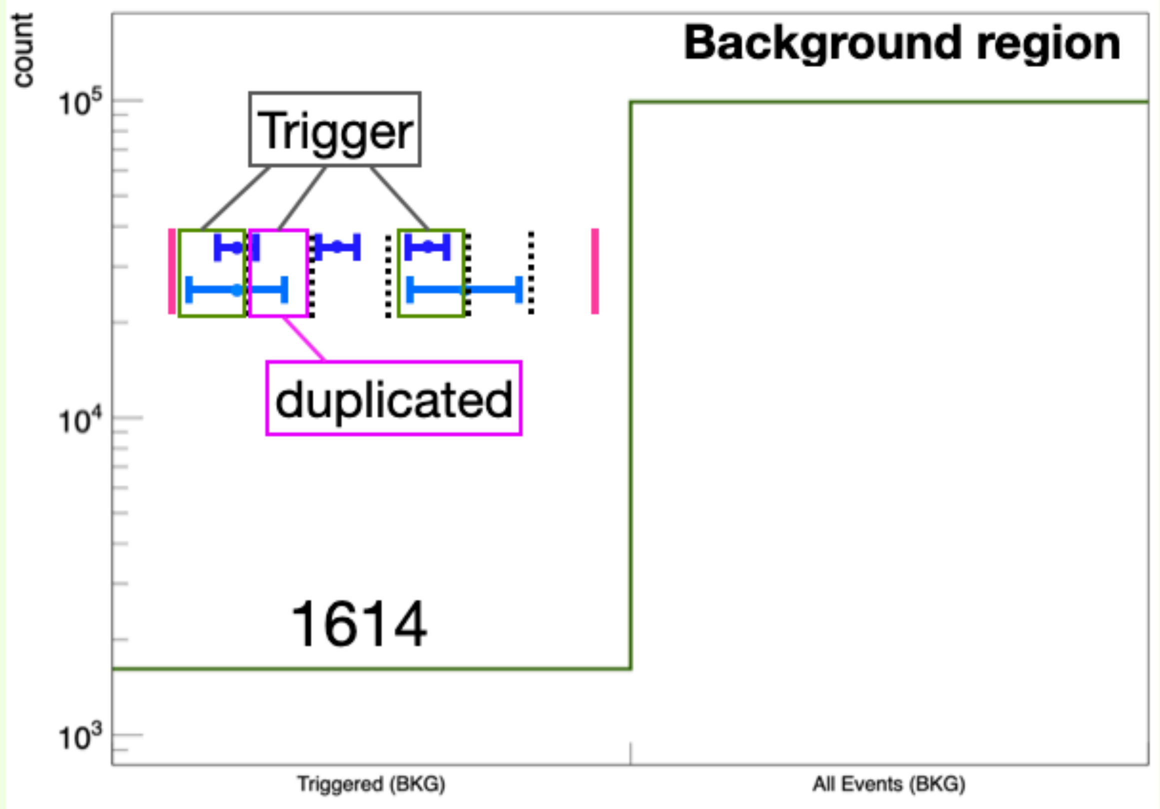


trigger/time slice width [ns]	**5**	**10**	**20**	**200**
Physics trigger rate (trigger/all events)	986/1000 = 98.6%	989/1000 = 98.9%	992/1000 = 99.2%	997/1000 = 99.7%
BKG trigger rate (trigger/all events)	1262/400000 = 0.3%	784/200000 = 0.4%	411/100000 = 0.4%	117/10000 = 1.2%

Ver.2
phys:
991/1000
fake: **10^5**

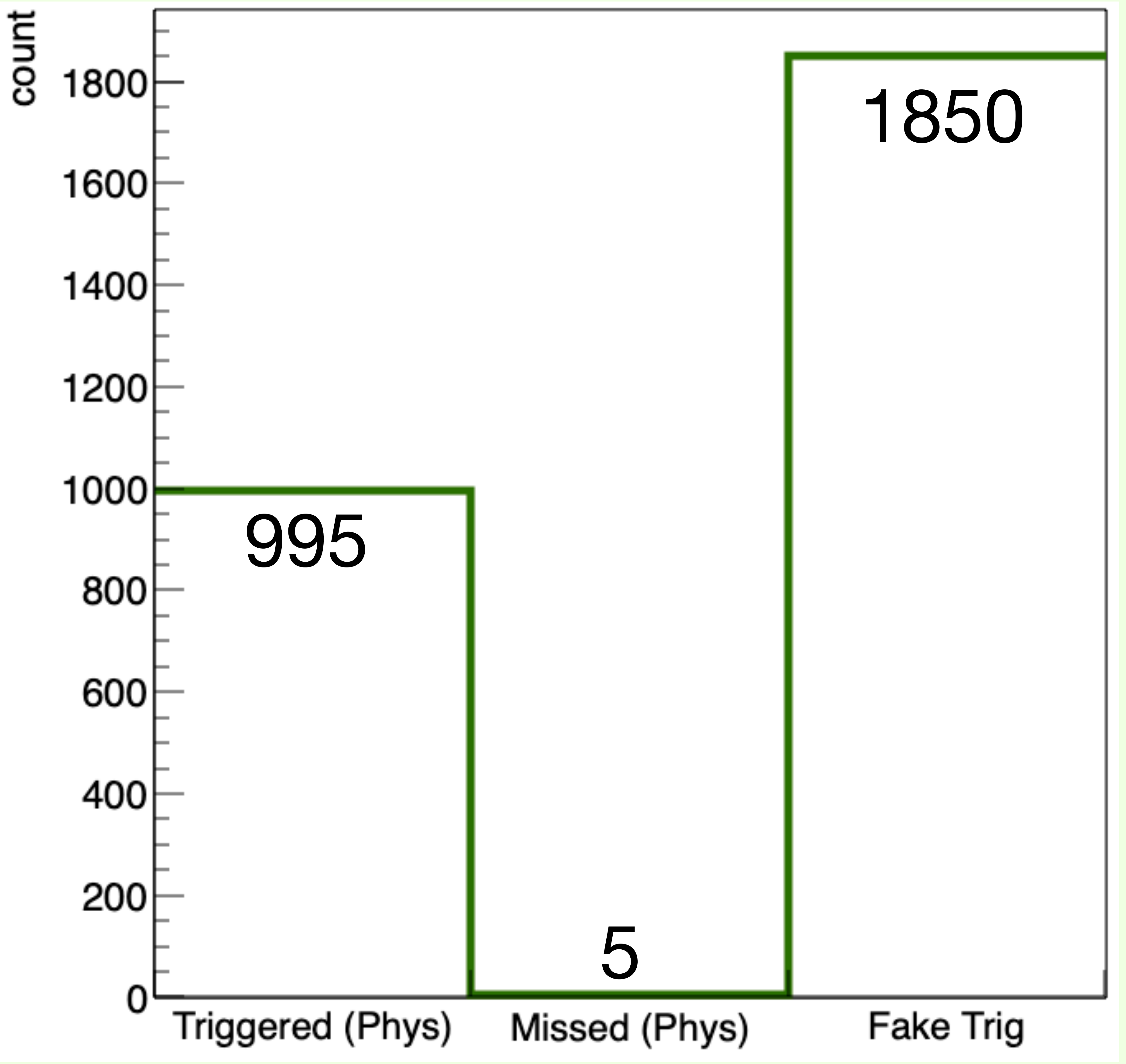
The last SRO meeting





- The efficiency could be kept 99.1%
- Wrong trigger increase.
→ There are duplicated data by considering time resolution
→Need to consider delete such duplicated triggers and more strong selection after reconstruction.

Ver.3



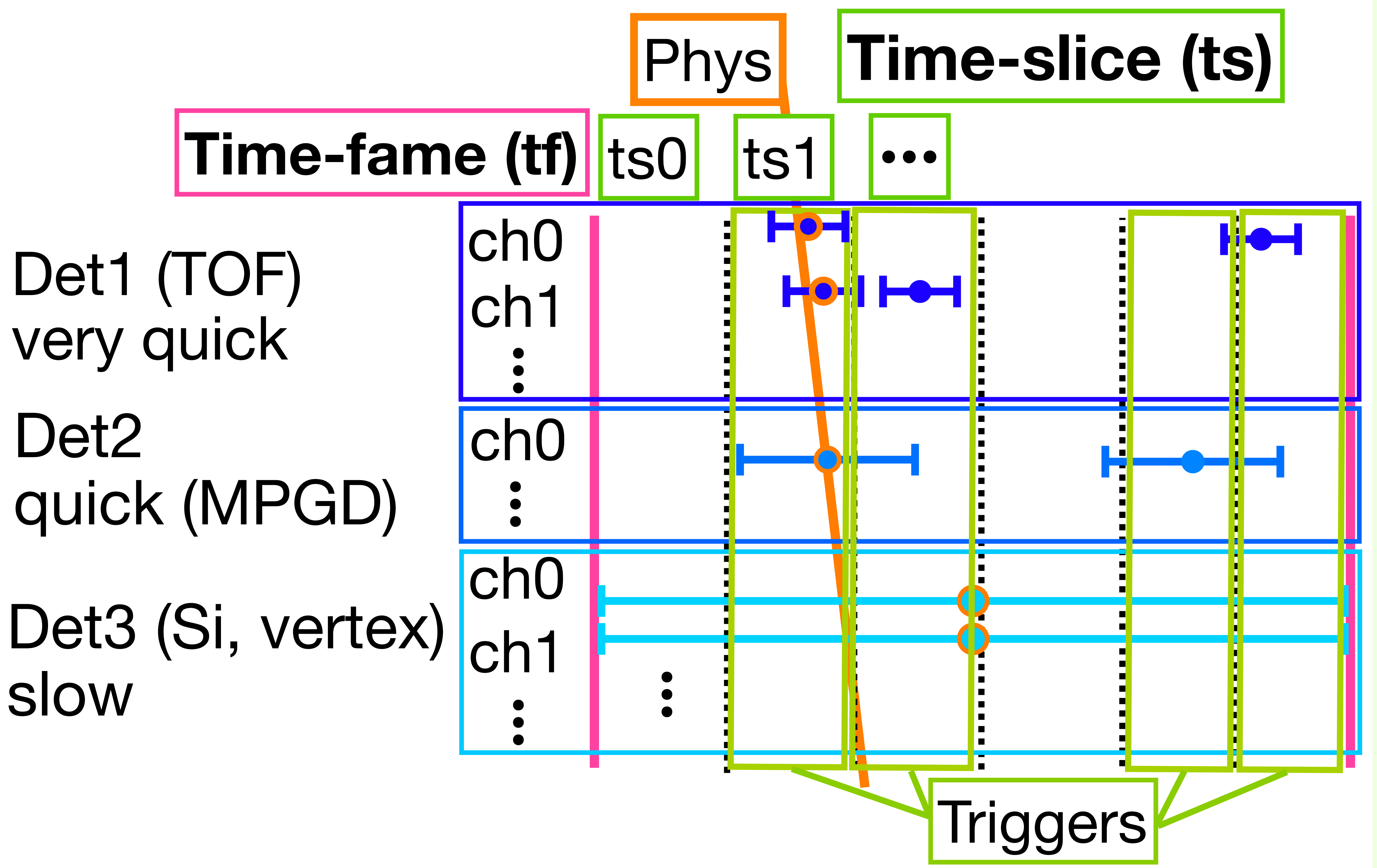
ver.2 algorithm (considering time resolution of detector)

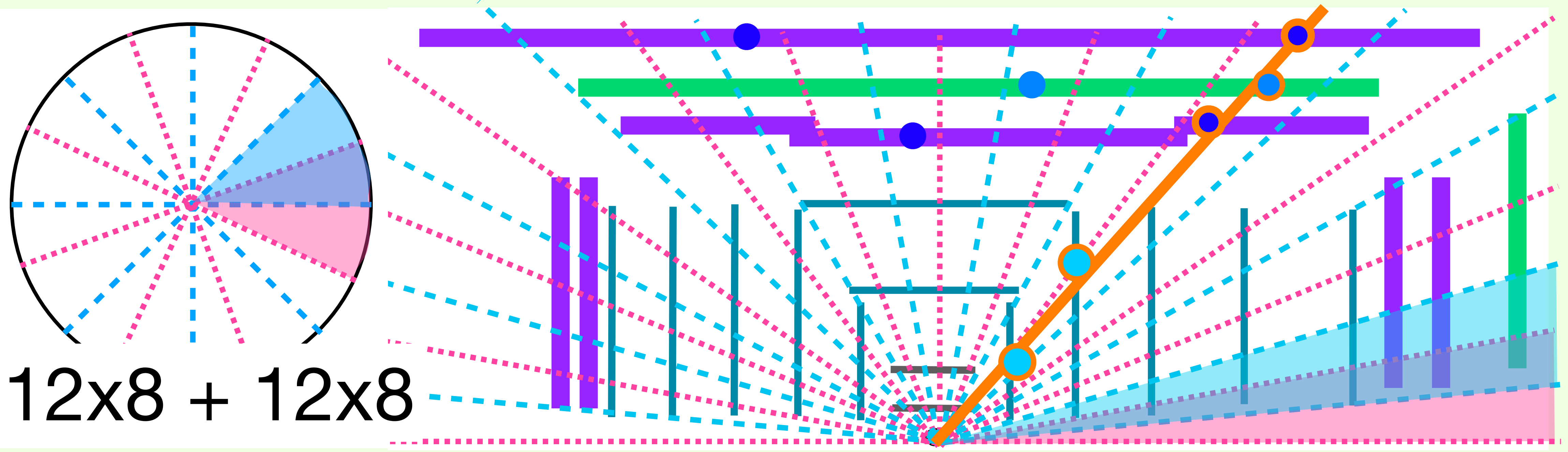
- There is a possibility overwrap time-frame edge by considering hit time resolution.

- Each sub-detector has different time resolution.

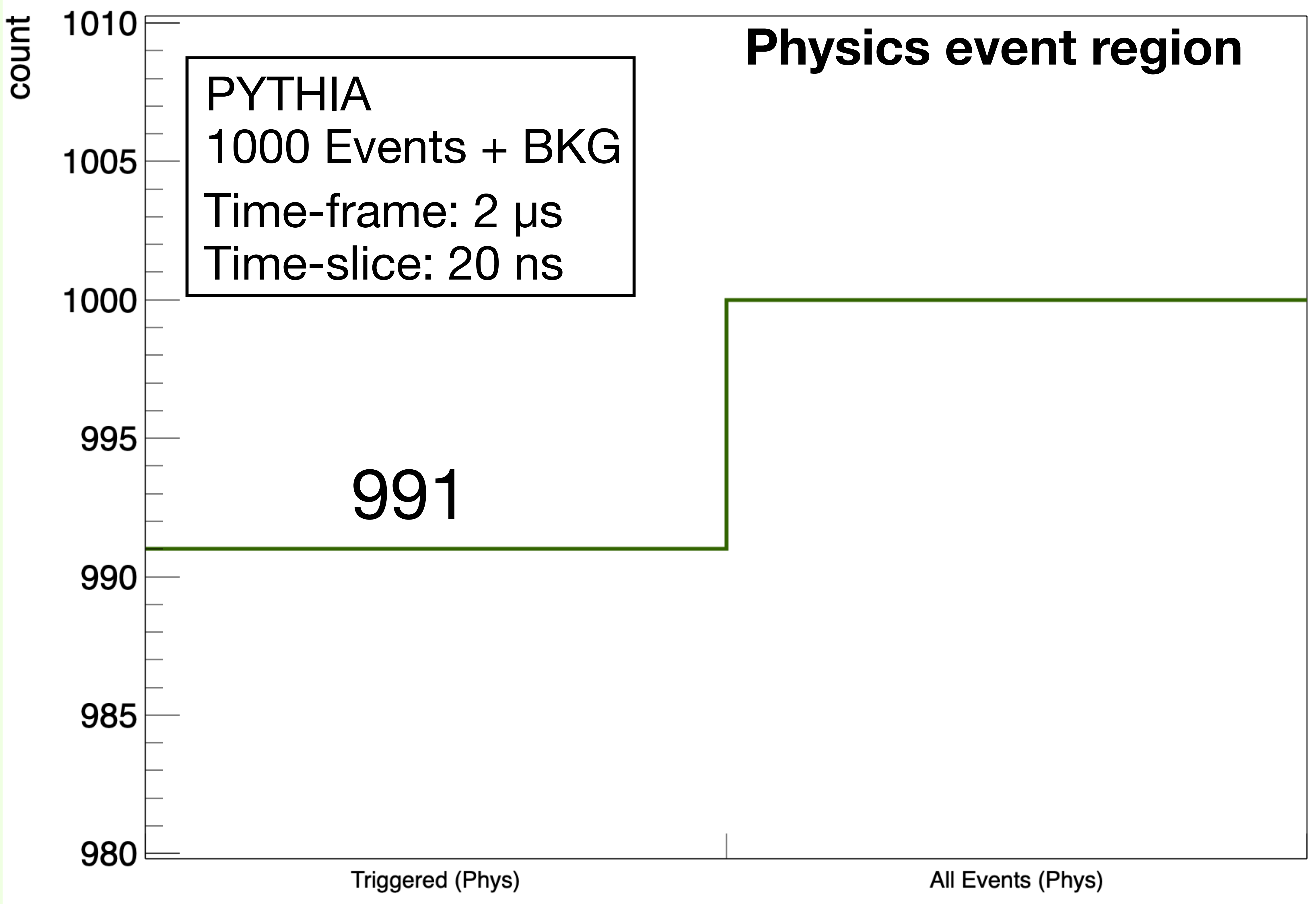
- Silicon detector has large time-resolution (~2 μs).
→ To reduce background, we need to use quick detector like MPGD(~10 ns)/TOF (30 ps).

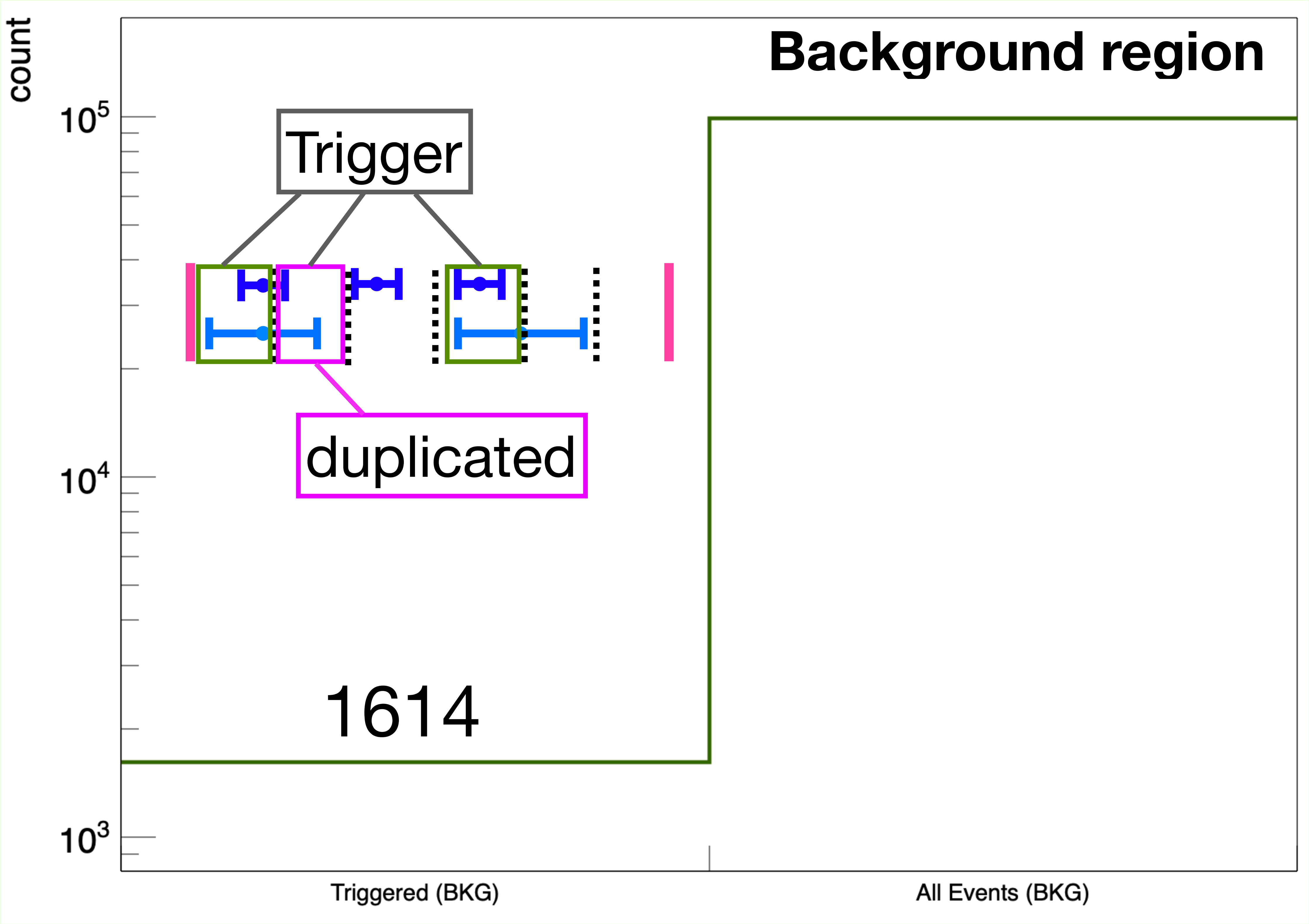
- Request over two hits for one region divided into ~200 (ηxφ = 2x (12x8))





New Trigger Algorithm of Efficiency





- The efficiency could be kept 99.1%
- Wrong trigger increase.

→ There are duplicated data by considering time resolution

→Need to consider delete such duplicated triggers and more strong selection after reconstruction.

New event build algorithm Ver3.0

1. Using the quickest detector (TOF) slices the time-frame.

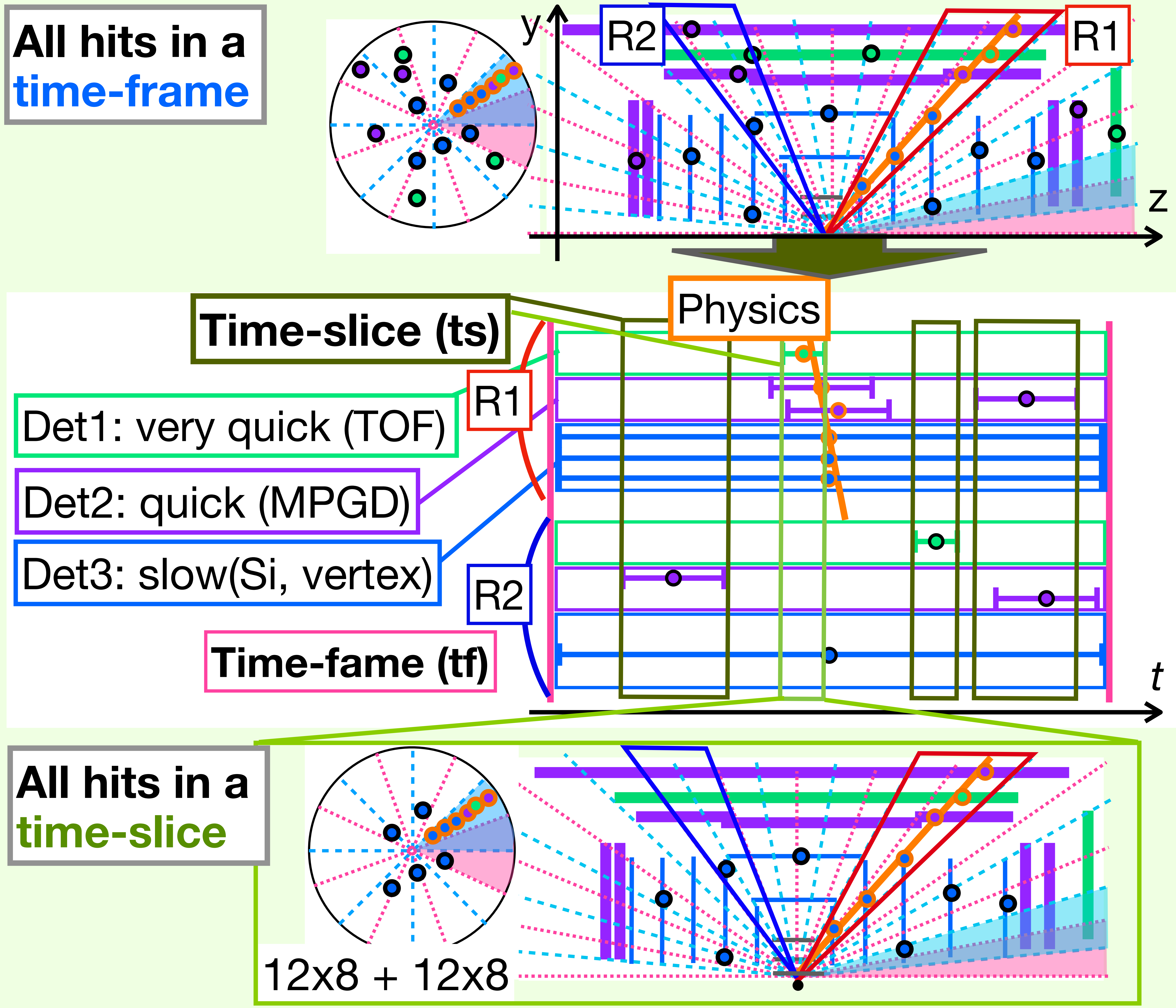
2. Check the topology of the hits.

3. If there are more than three hits in the same region, the time slice is triggered.

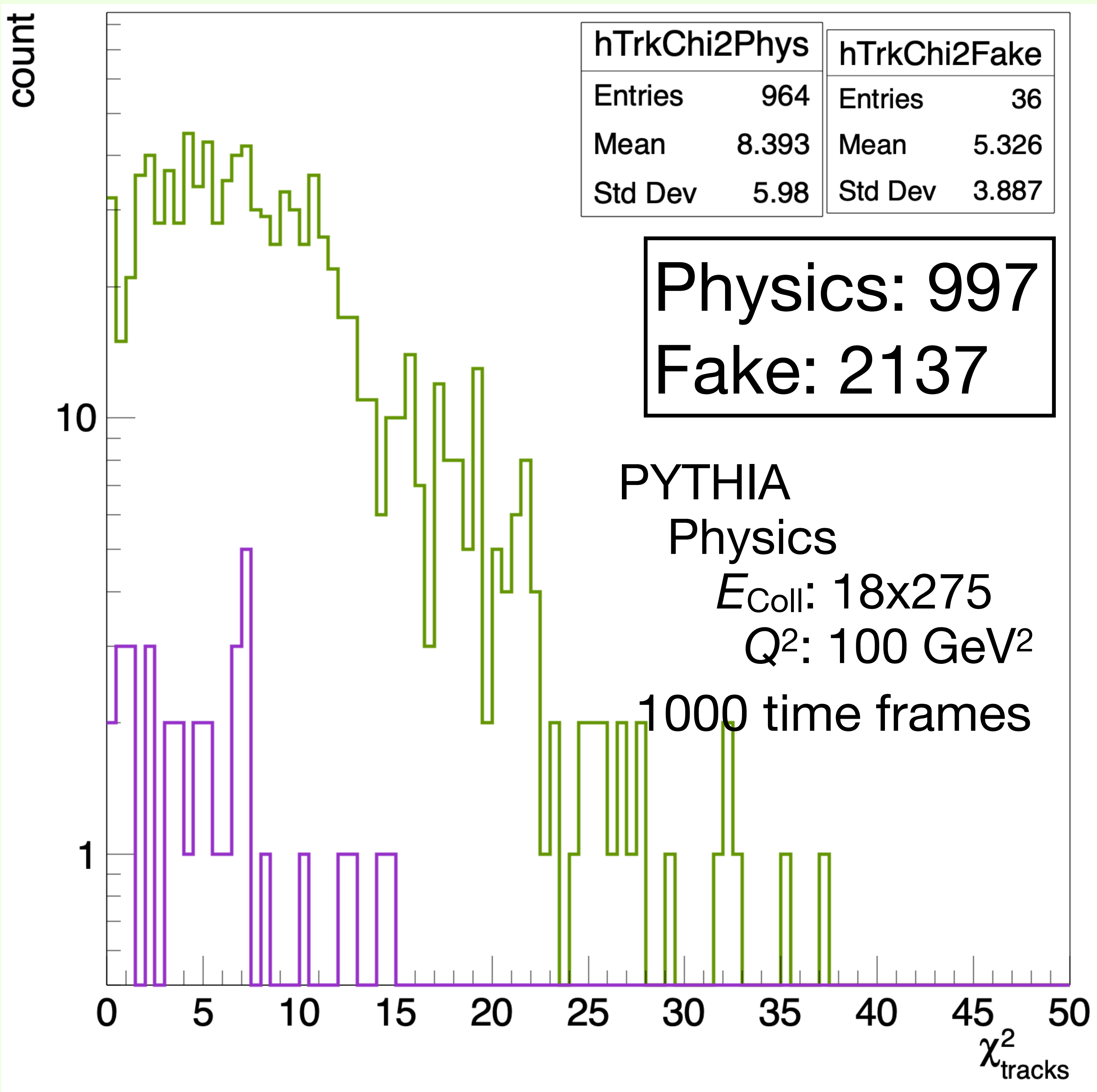
4. Repeat steps 1–3 for all TOF hits.

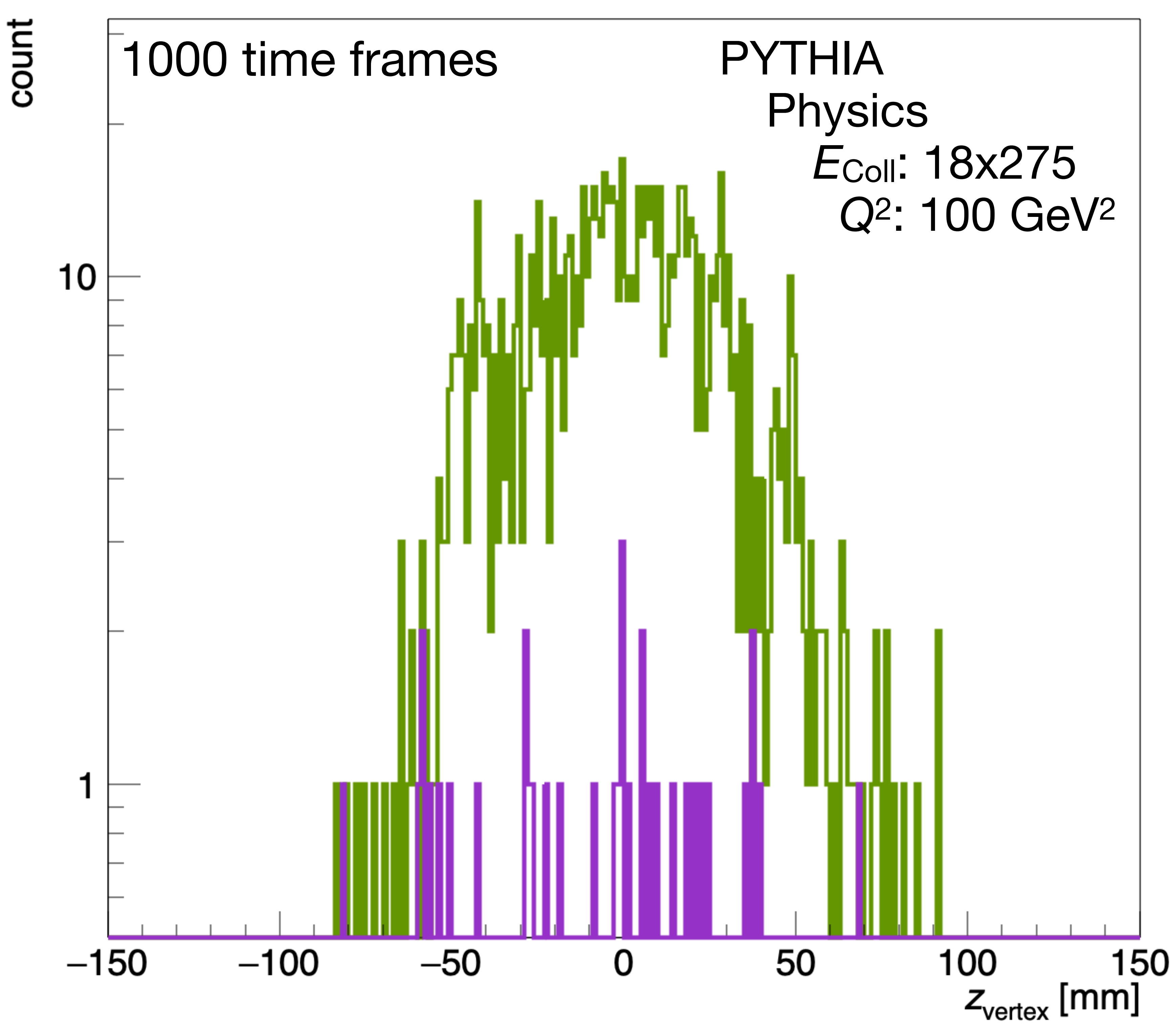
5. If any unused MPGD hits remain, perform the same process using them.

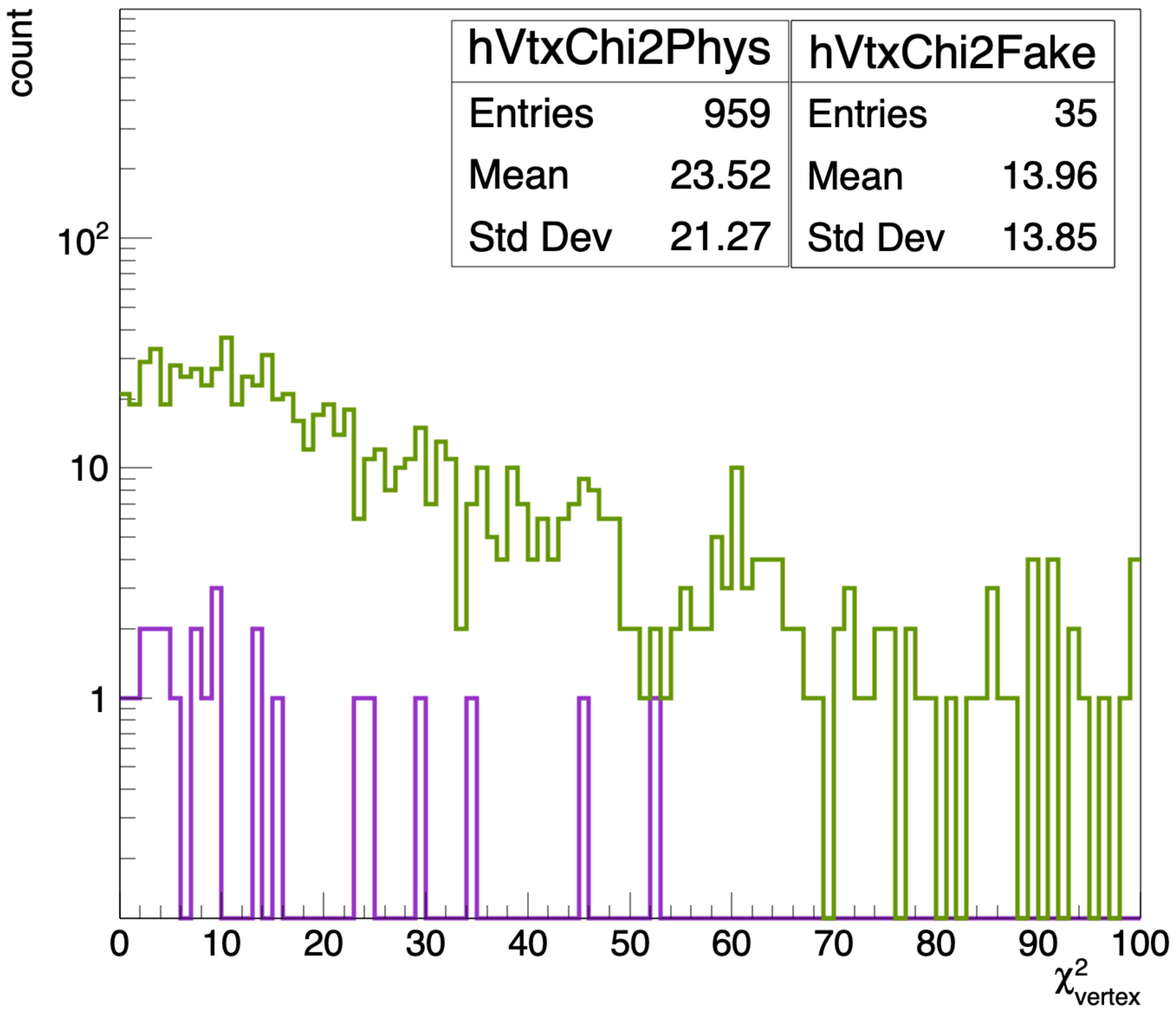
6. If there is no trigger using TOF and MPGD in the time frame, perform the same process based on the Silicon detectors' hits.



Difference between physics and background time slice







- The most fake time slices have only a small number of hits.
- The tendencies of values are not an obvious difference.
- Some time slices do not have a vertex despite the physics time slice.