Optimization and Outer Tracker Extension of the Cellular Automaton Algorithm for CMS Phase-2 HLT Tracking

Jan Gerrit Schulz
on behalf of the CMS Collaboration

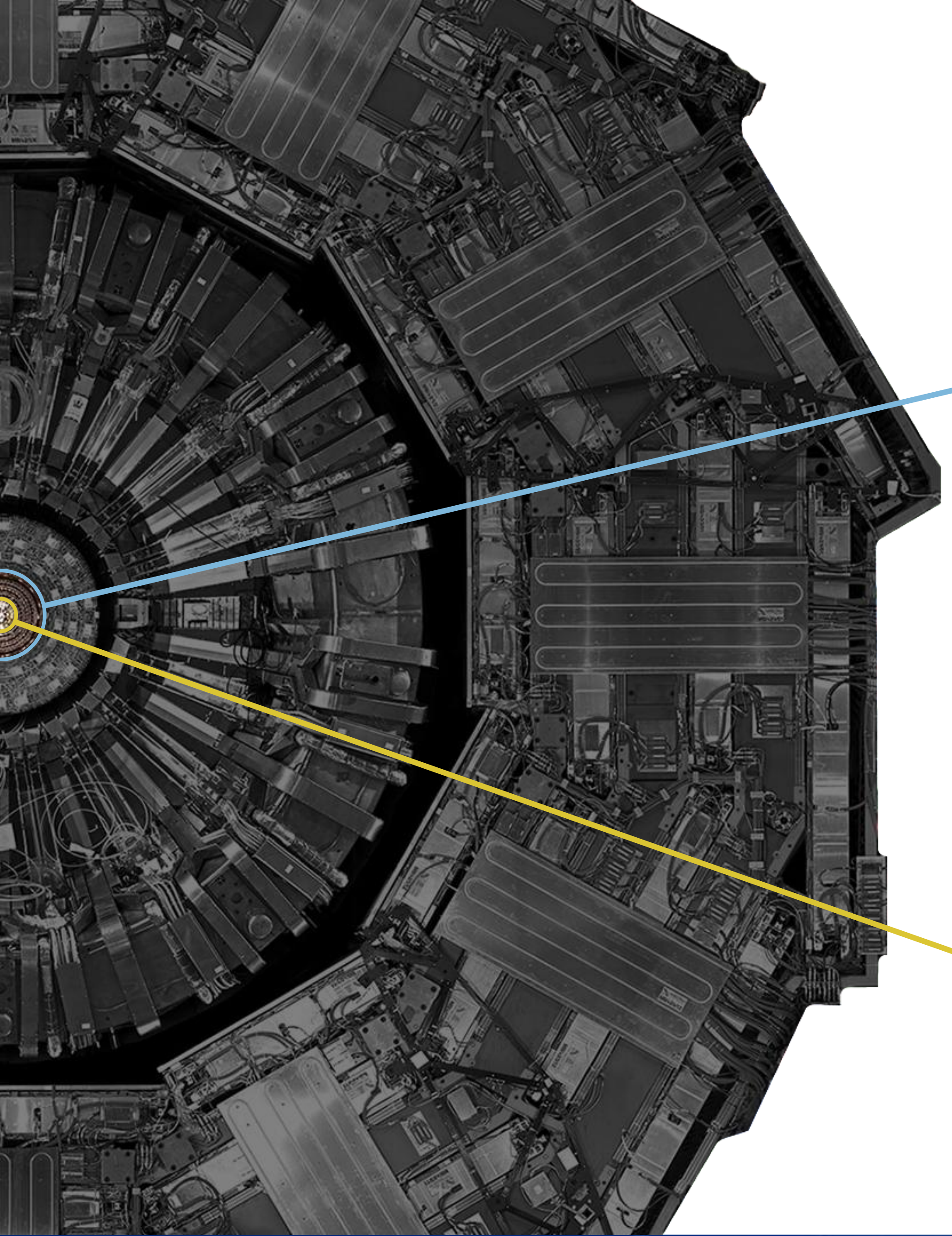
CERN, Switzerland
RWTH Aachen University, Germany

November 11th, 2025 • Connecting the Dots • Tokyo, Japan

CERN | NexTGen Next Generation Triggers | CMS

future detector
for HL-LHC

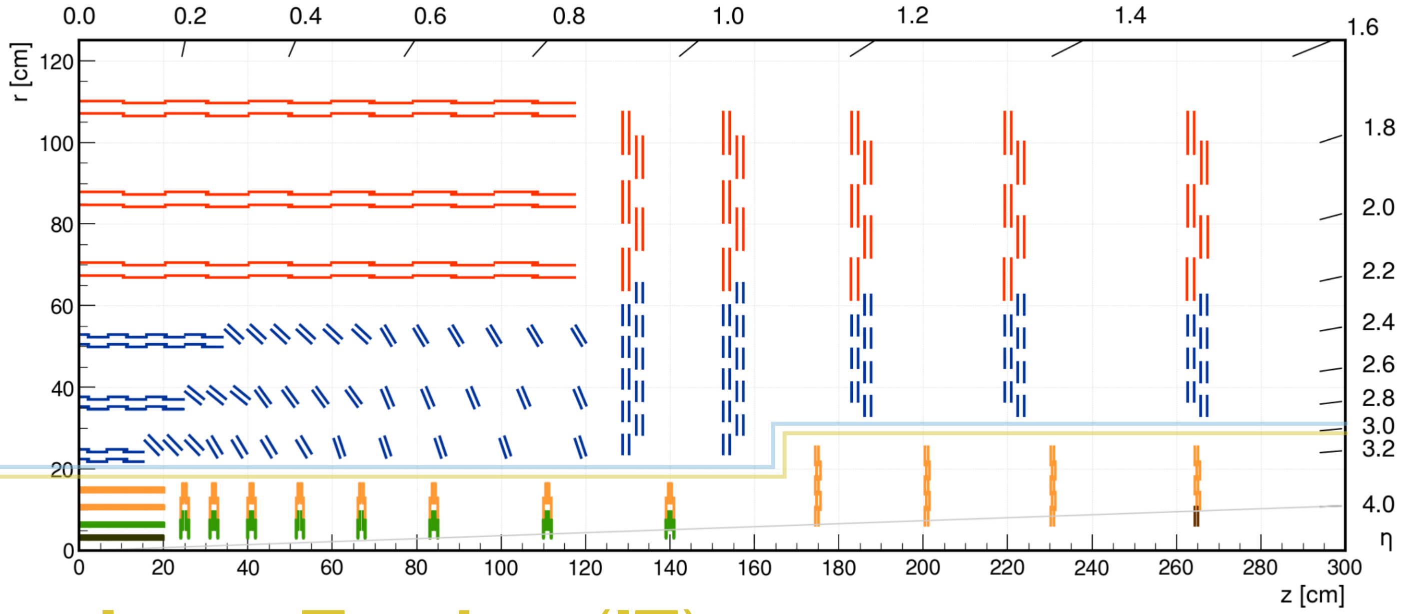
CMS Tracker in Phase-2



Outer Tracker (OT)

Double-sided silicon modules in 6 barrels + 2×5 disks

Pixel-strip (PS) modules + two strip (2S) modules



Inner Tracker (IT)

Pixel modules (■,■,■)

with ~2B pixels in 4 barrels + 2×12 disks

Tracking in the HLT of CMS Phase-2

At the High-Level Trigger (HLT), tracking aims for:

- ❑ **high efficiency**
- ❑ **low fake rate**
- ❑ **good track quality**
- ❑ **fast execution time**
  foreseen HLT event rate of 750 kHz after L1-Trigger

Legacy configuration: baseline in [CMS-TDR-022](CMS-TDR-022)
→ **Iterative tracking** approach:

Initial step

Seeds = pixel tracks (4 hits from IT)	Full track building (IT+OT) starting from seeds	Track fitting + selection

↓

High-p_T triplet step to recover inefficiencies of initial step

Seeds = pixel hit triplets (3 hits from IT)	Full track building (IT+OT) starting from seeds	Track fitting + selection

Pixel Tracking in the HLT of CMS Phase-2

Pixel tracks play **crucial** role in HLT

- should be **extremely fast tracking** usable **for…**
  - …**seeding of (slower) full tracking on selected HLT paths**
  - …**many other paths where full tracking will not run**
- need again: high efficiency, low fake rate, good quality

e.g. directly used for HLT muons (see Luca's talk in ~1h)

basis for pixel vertices, used for:

- High-purity selection,
- Tau reconstruction,
- Potential trimmed tracking, …

Initial step

Seeds = pixel tracks (4 hits from IT)	Full track building (IT+OT) starting from seeds	Track fitting + selection

High-p_T triplet step to recover inefficiencies of initial step

Seeds = pixel hit triplets (3 hits from IT)	Full track building (IT+OT) starting from seeds	Track fitting + selection

BUT: in barrel, **~10% of tracks hit less than 4 pixel sensors** (see later) and **3-hit tracking with PU 200 is unmanageable** (fakes, memory)

Improvements for Phase-2 pixel tracking needed

How to Pixel Track? – Algorithms in CMS

Legacy

- Algorithm used in the Legacy baseline for iterative HLT tracking
- Based on cellular automata
- Runs fully on CPU
- Is limited exactly 4 hits per pixel track for time performance reasons

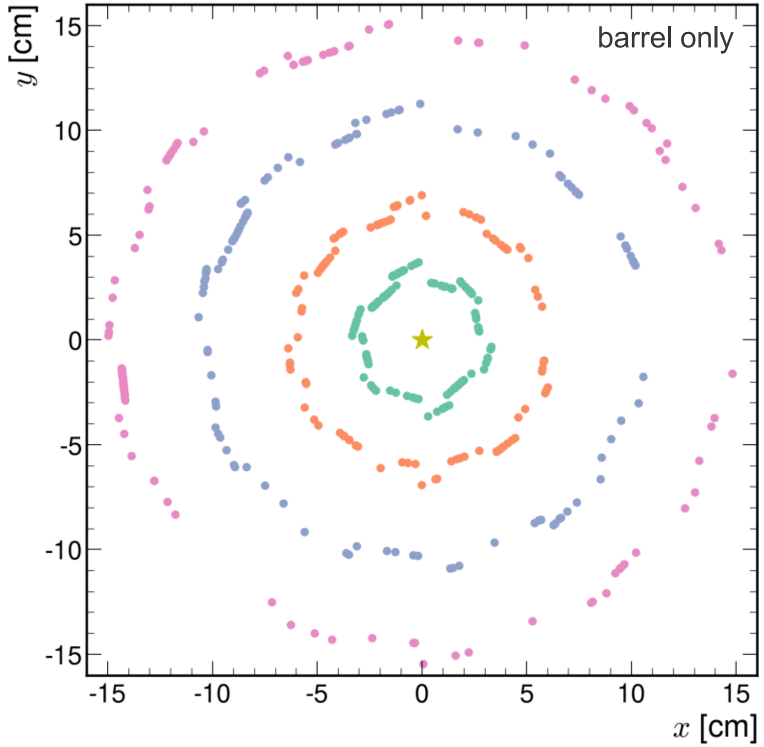
Patatrack

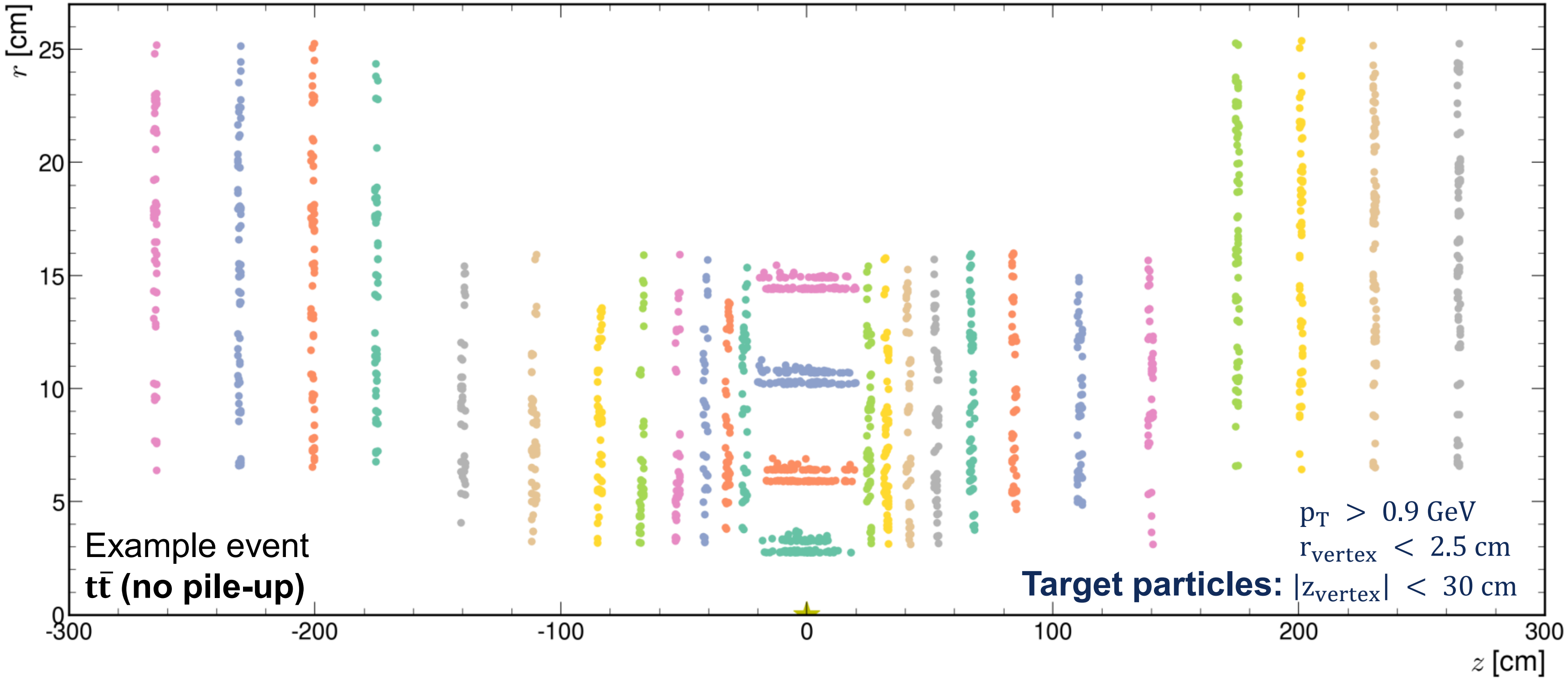
- Algorithm used in HLT Run-3, early Phase-2 tests (see CMS-DP-2021-013)
- Also based on cellular automata
- Heterogeneously implemented with portability library alpaka → run on GPU
- Not limited in number of hits per track

→ due to clear advantages, **optimization** for HLT Phase-2 is fully **concentrated on Patatrack**

How to Pixel Track? – The Patatrack Algorithm

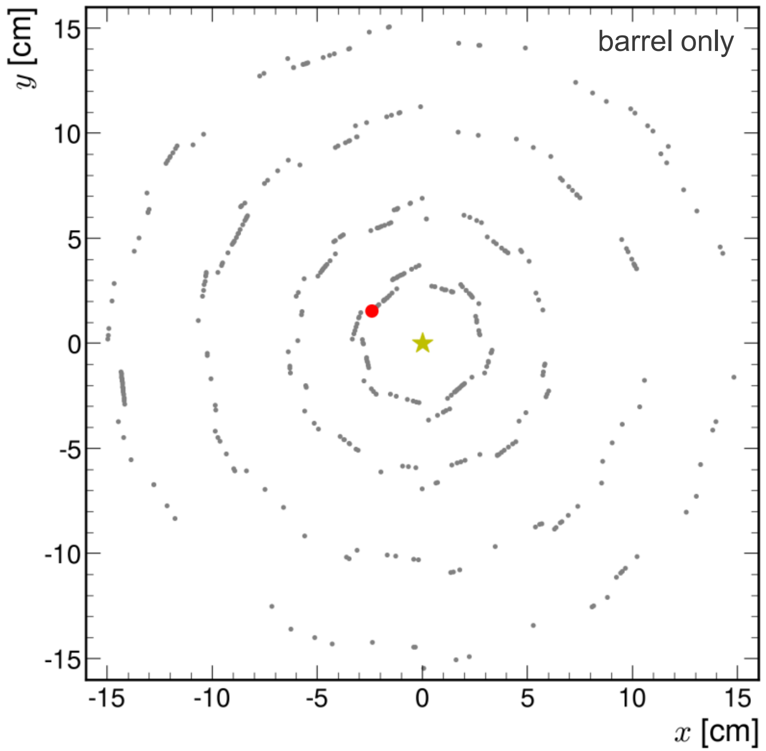
RecHits → *connect to* → **Doublets** → *connect to* → **N-tuplets** → *fit to* → **Pixel Tracks**

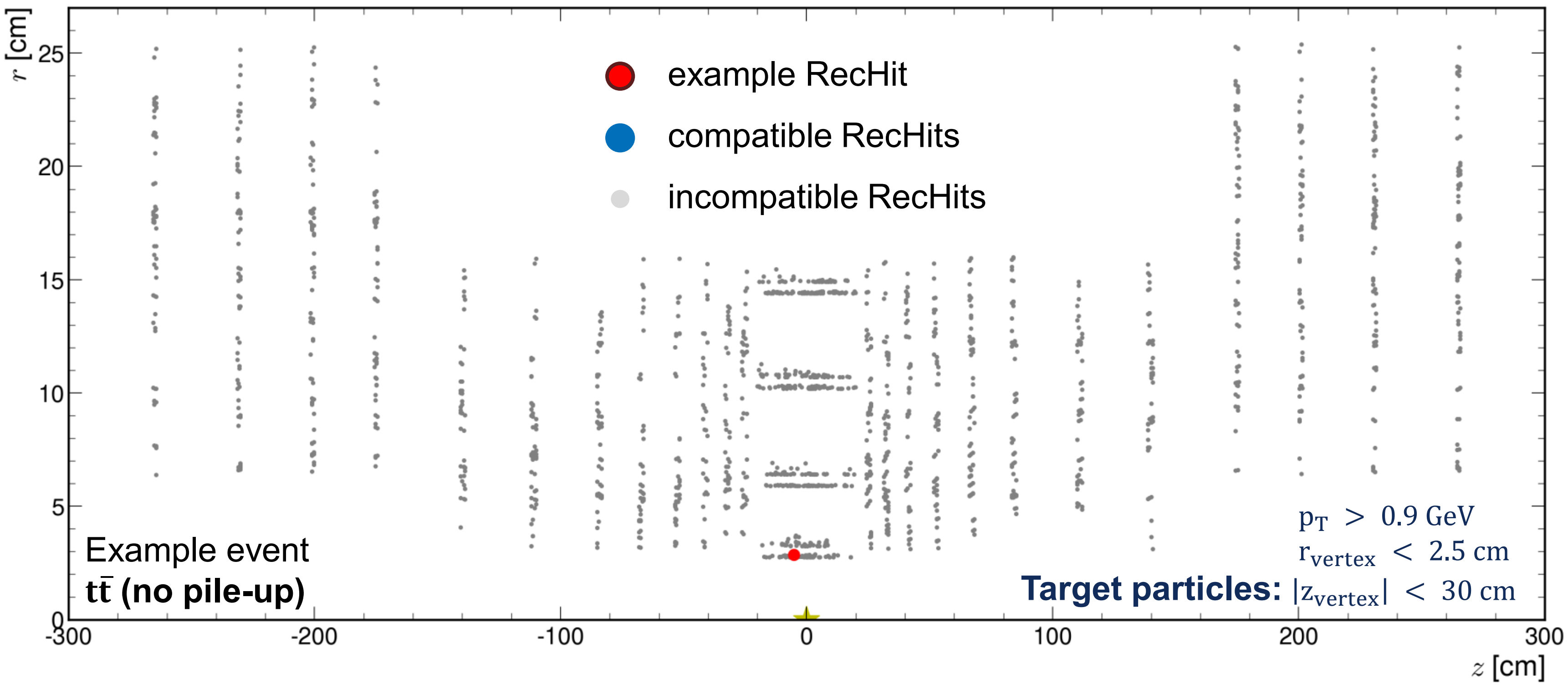




How to Pixel Track? – The Patatrack Algorithm

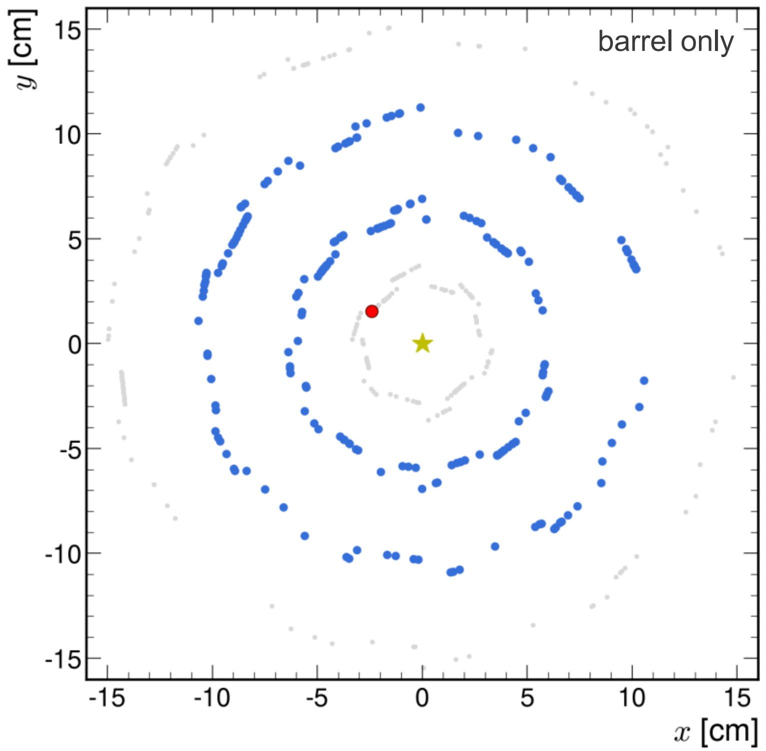
RecHits *connect to* **Doublets** *connect to* **N-tuplets** *fit to* **Pixel Tracks**

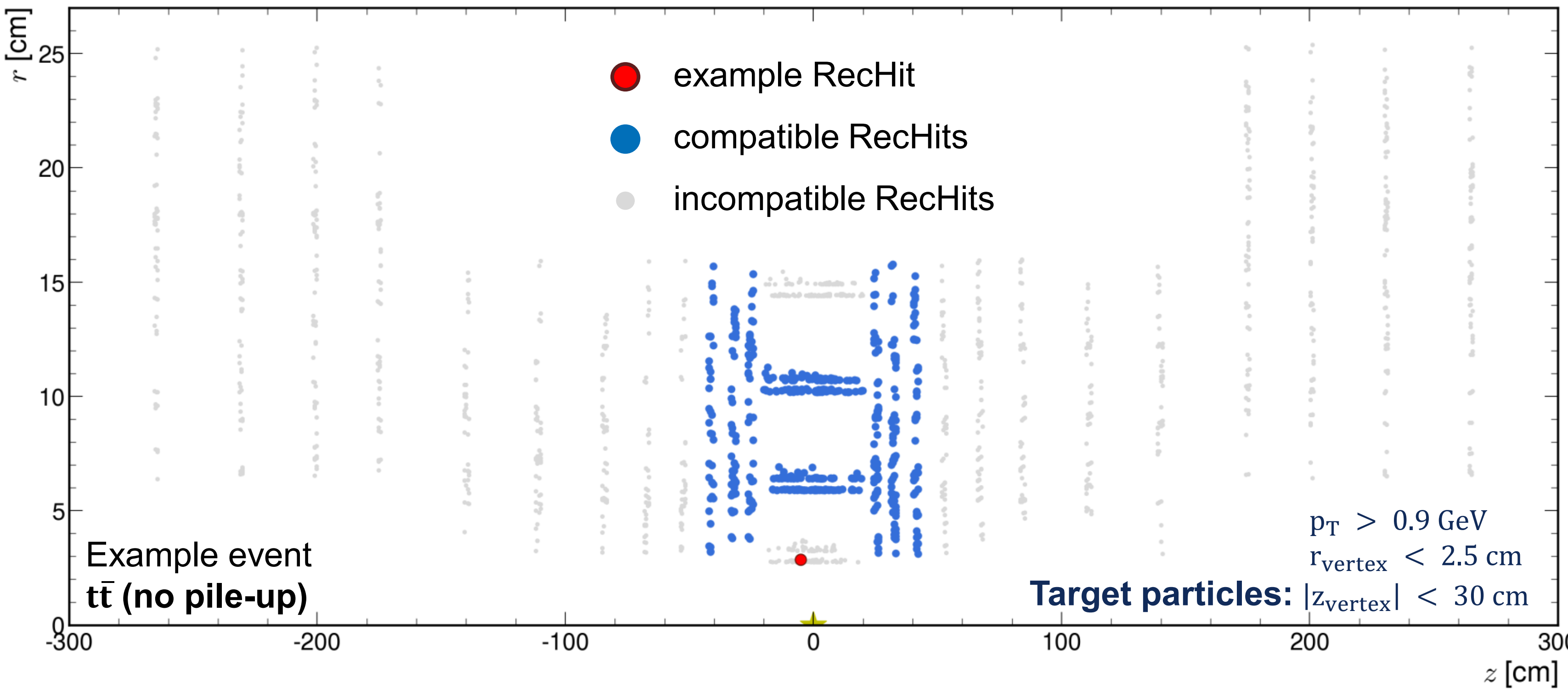




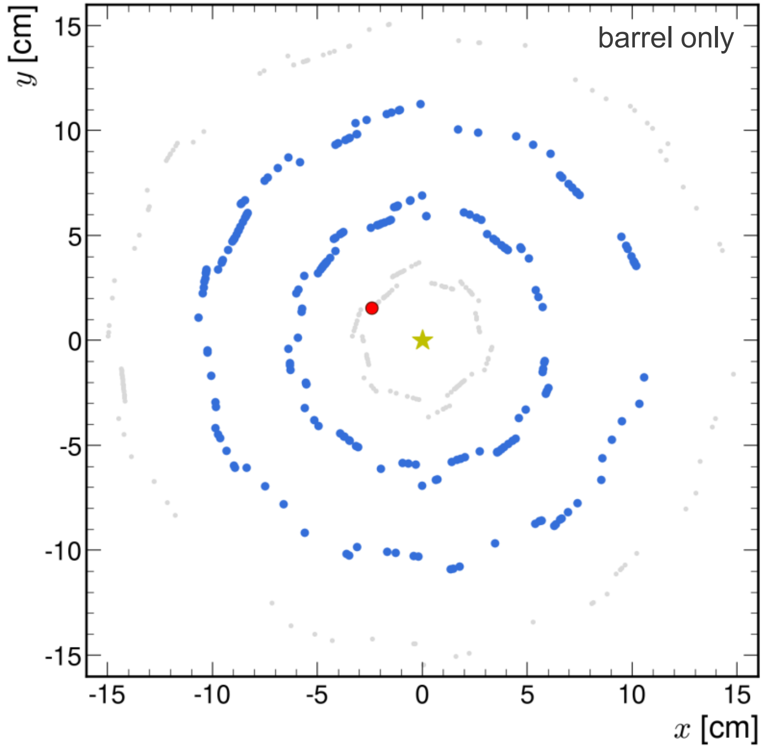
How to Pixel Track? – The Patatrack Algorithm

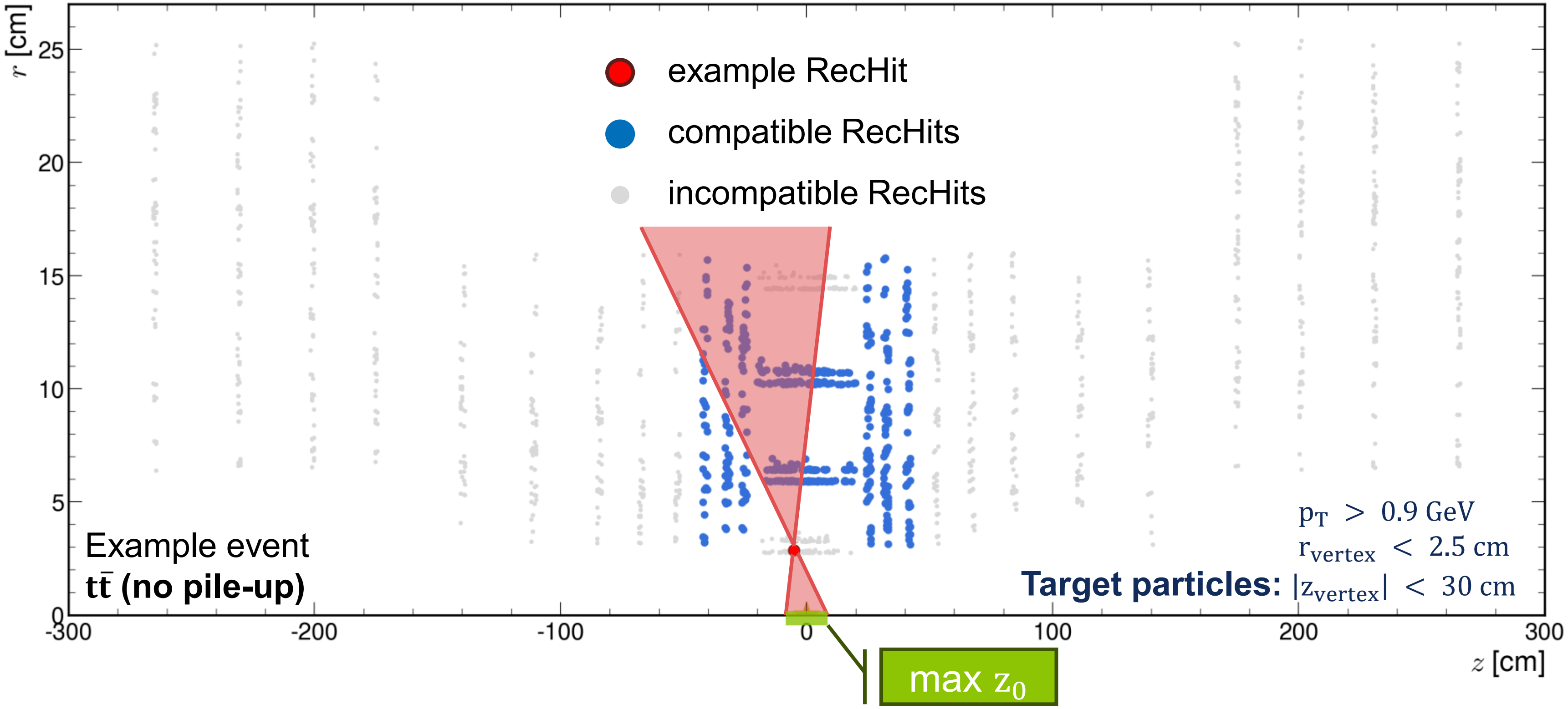
RecHits *connect to* **Doublets** *connect to* **N-tuplets** *fit to* **Pixel Tracks**





How to Pixel Track? – The Patatrack Algorithm

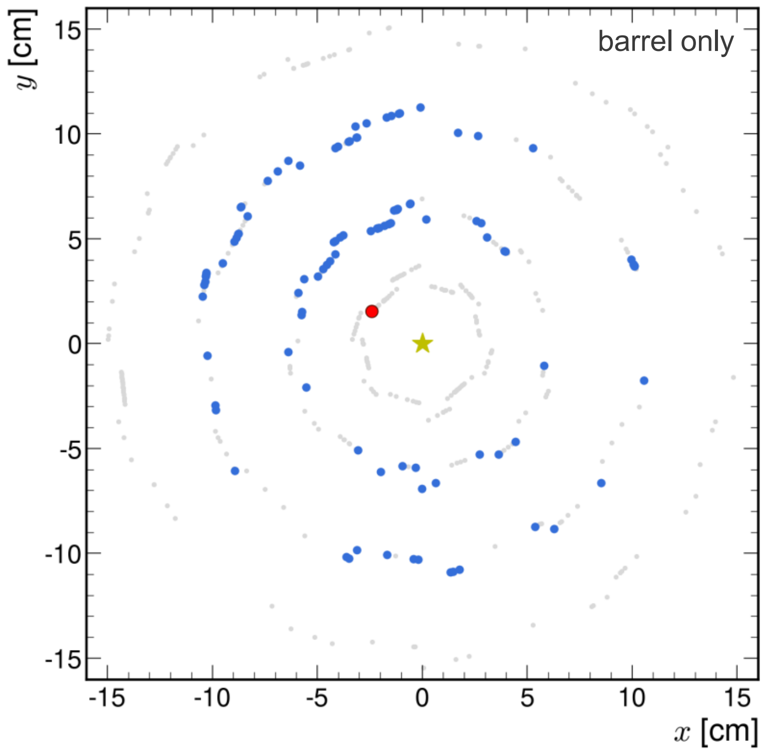


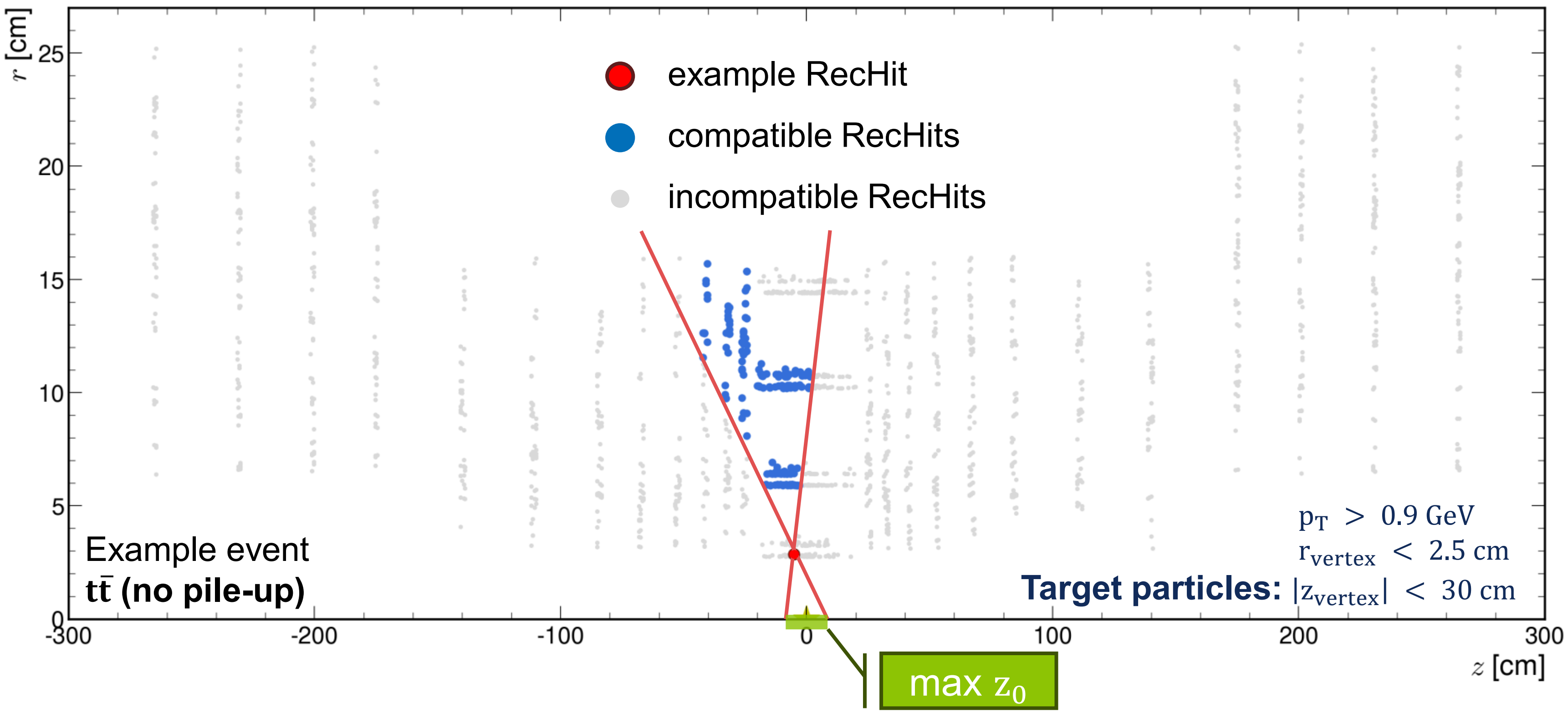


Target particles: $p_T > 0.9$ GeV, $r_{vertex} < 2.5$ cm, $|z_{vertex}| < 30$ cm

How to Pixel Track? – The Patatrack Algorithm

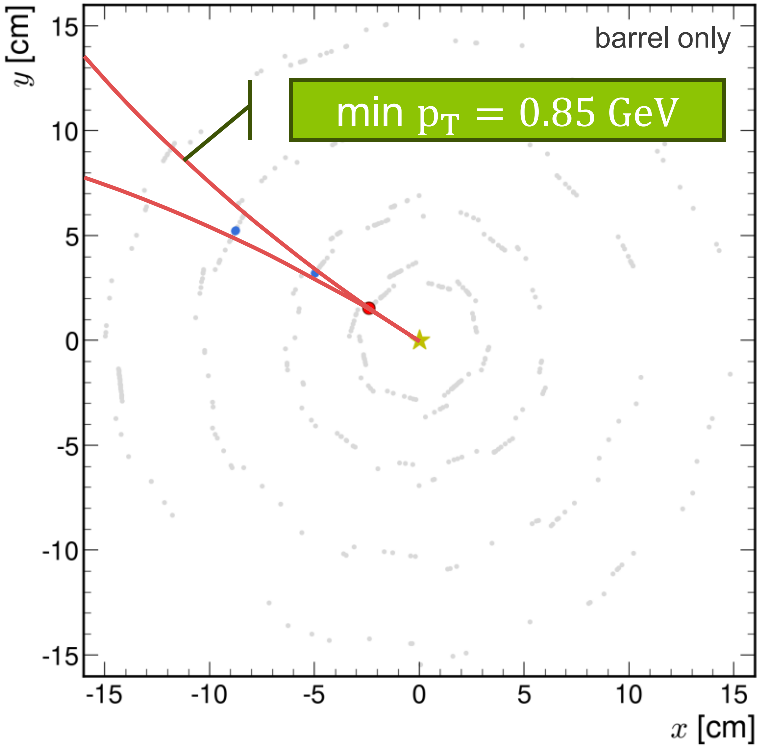
RecHits → *connect to* → **Doublets** → *connect to* → **N-tuplets** → *fit to* → **Pixel Tracks**

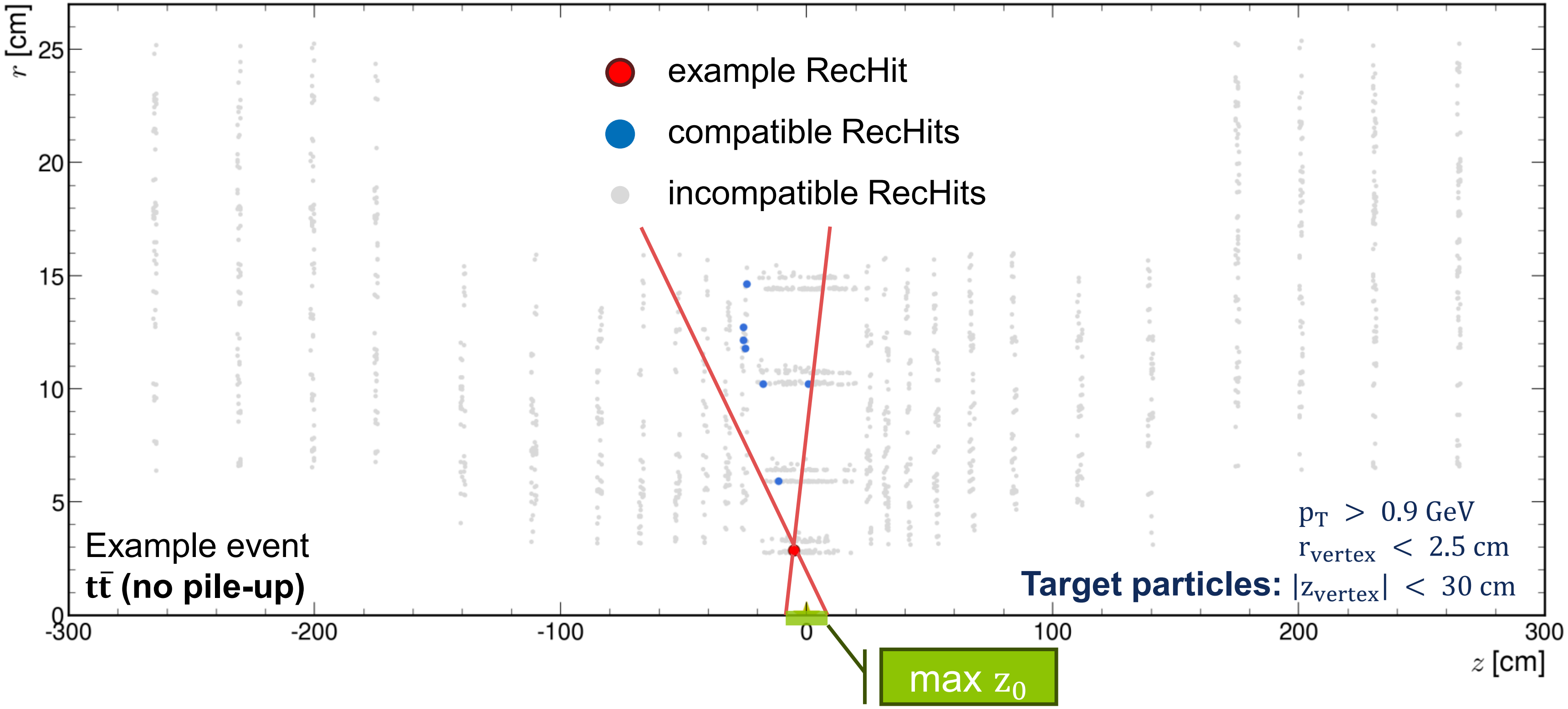




How to Pixel Track? – The Patatrack Algorithm

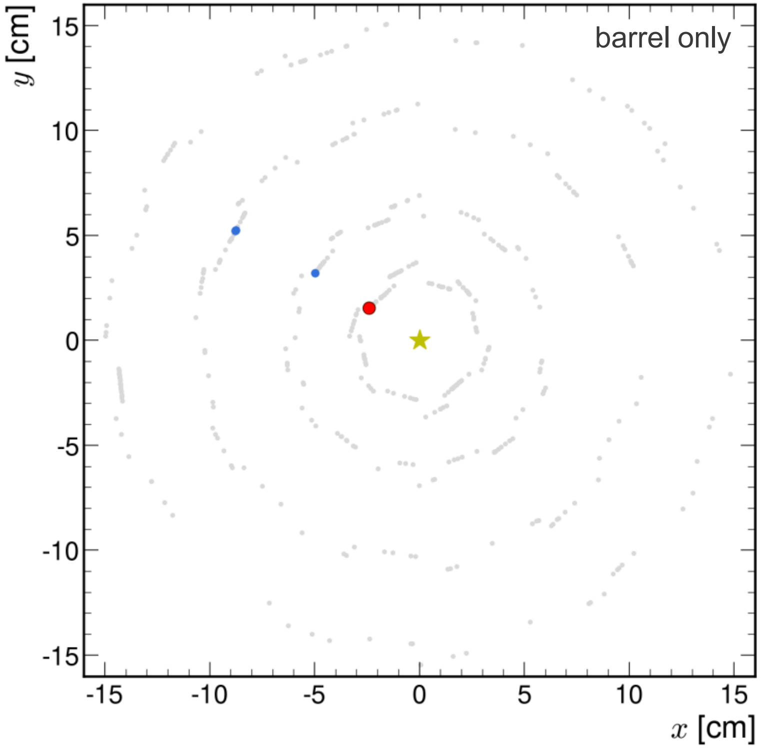
RecHits *connect to* **Doublets** *connect to* **N-tuplets** *fit to* **Pixel Tracks**

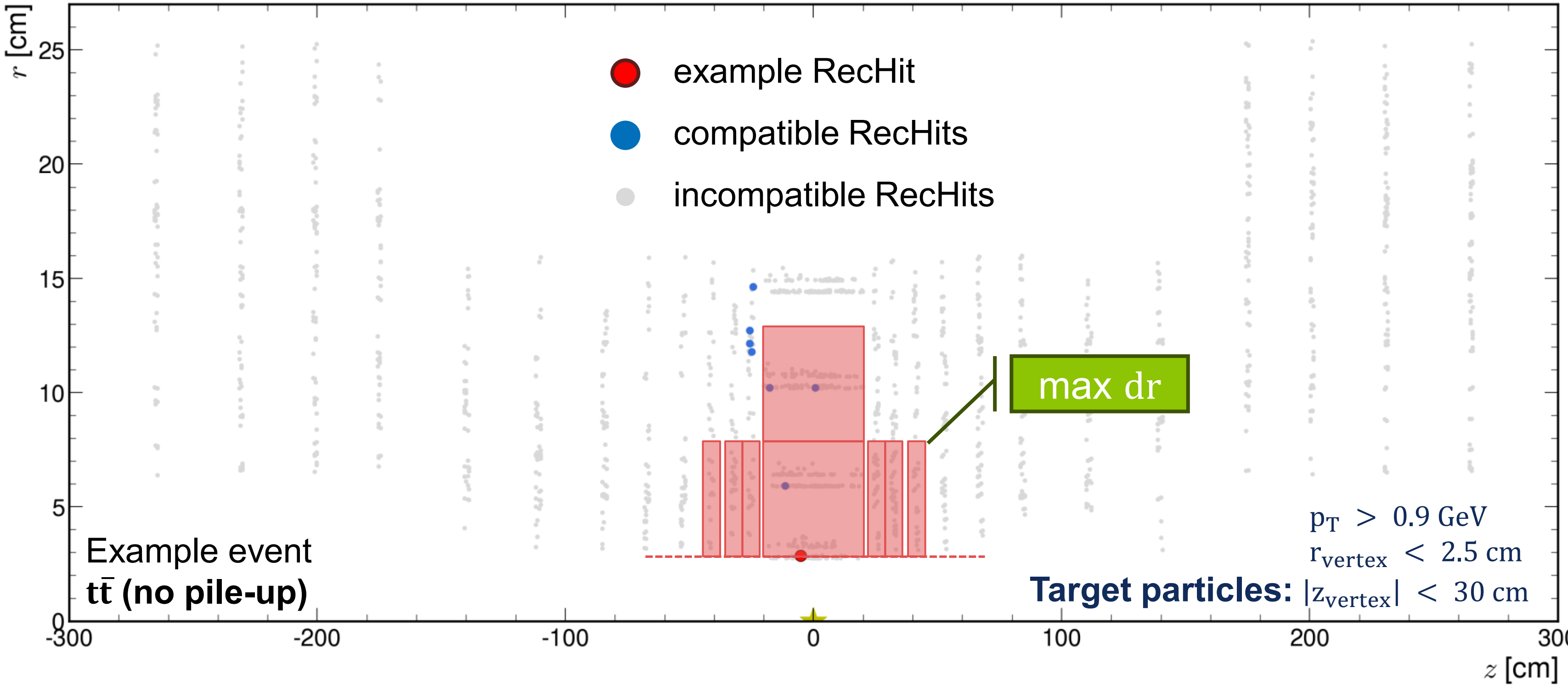




How to Pixel Track? – The Patatrack Algorithm

RecHits *connect to* **Doublets** *connect to* **N-tuplets** *fit to* **Pixel Tracks**

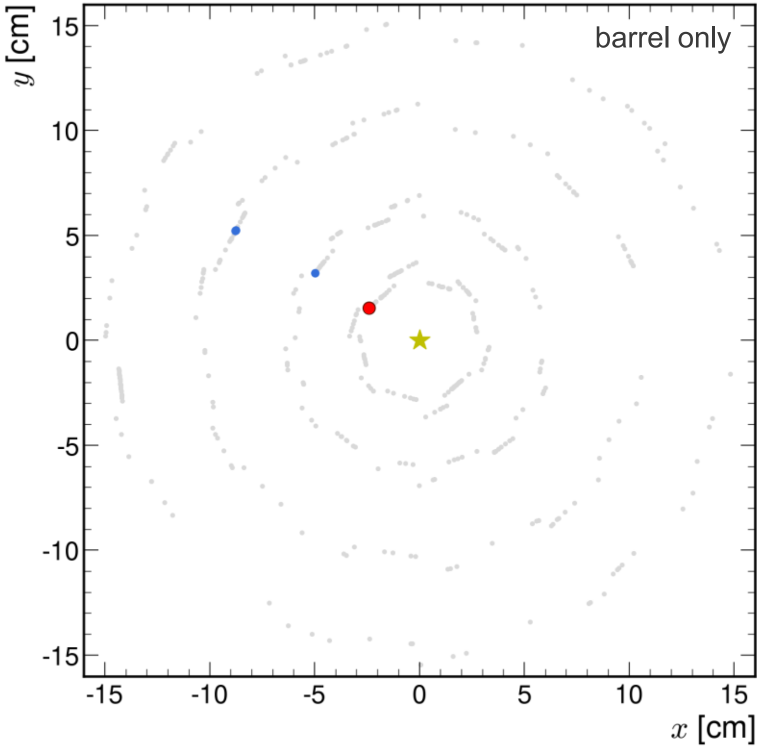


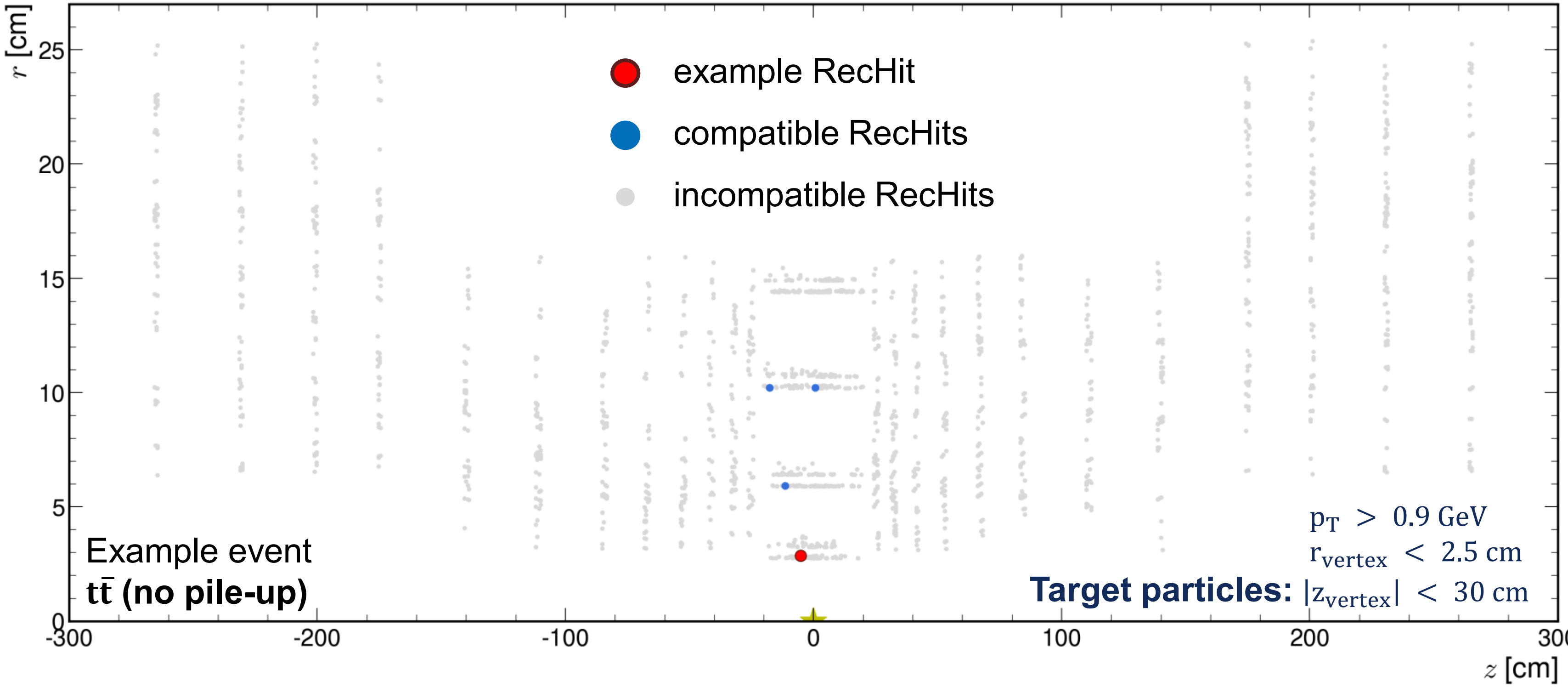


Target particles: $p_T > 0.9$ GeV, $r_{vertex} < 2.5$ cm, $|z_{vertex}| < 30$ cm

How to Pixel Track? – The Patatrack Algorithm

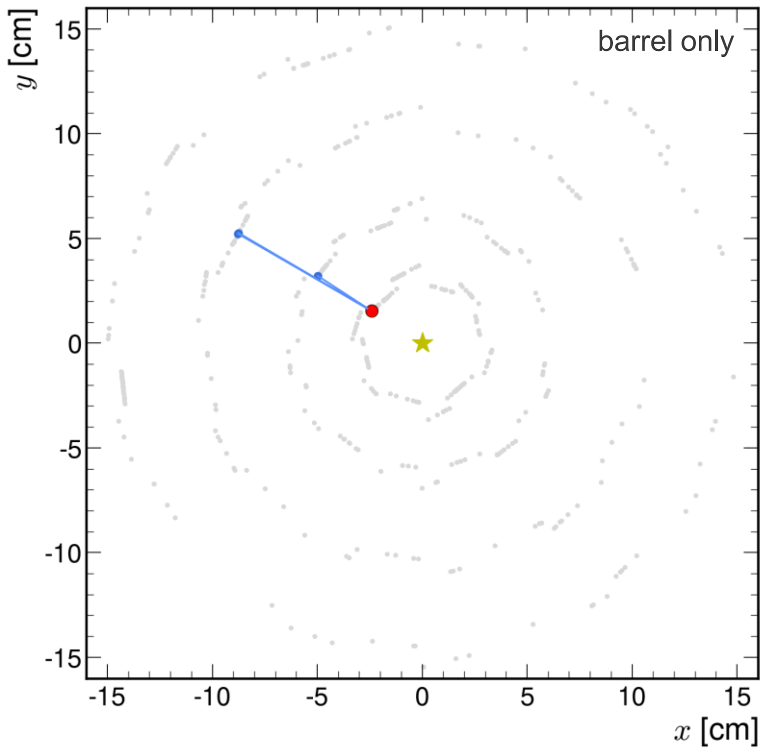
RecHits *connect to* **Doublets** *connect to* **N-tuplets** *fit to* **Pixel Tracks**

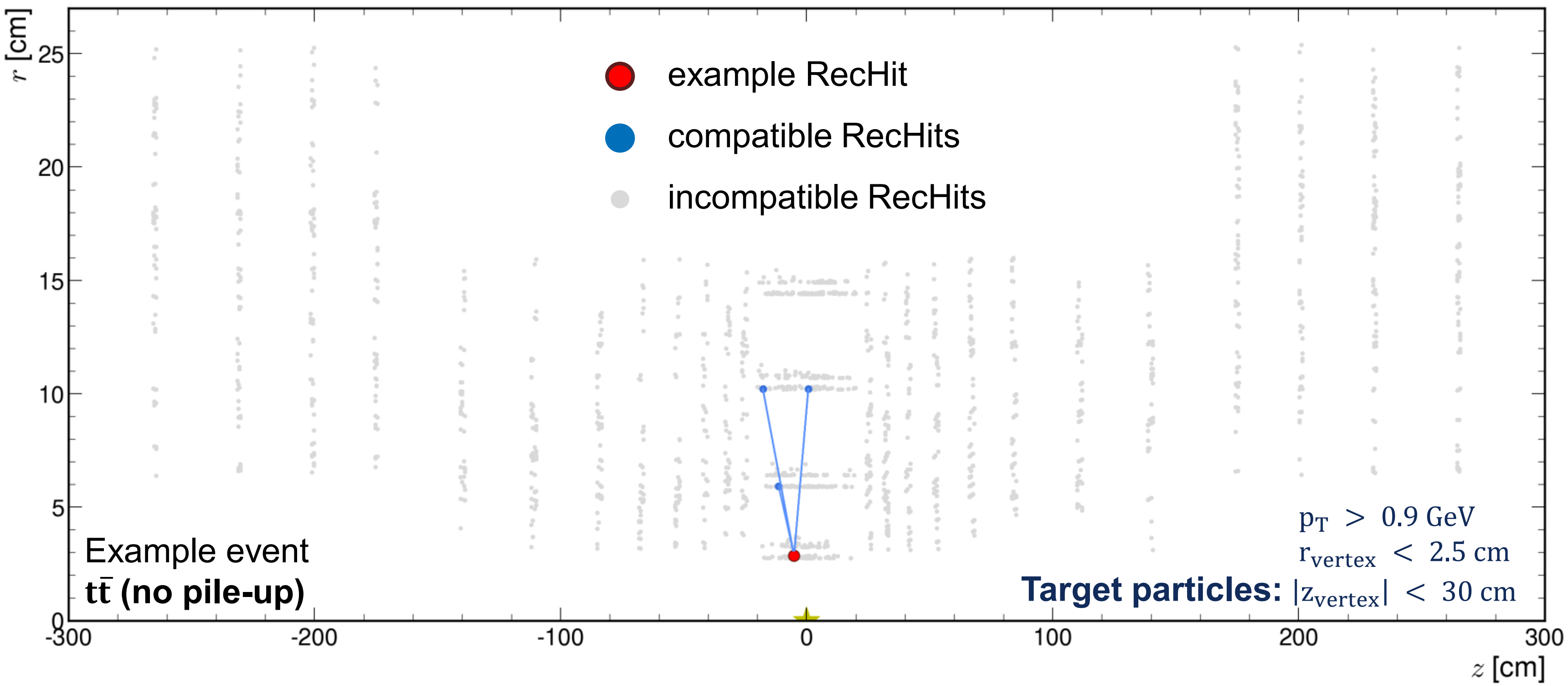




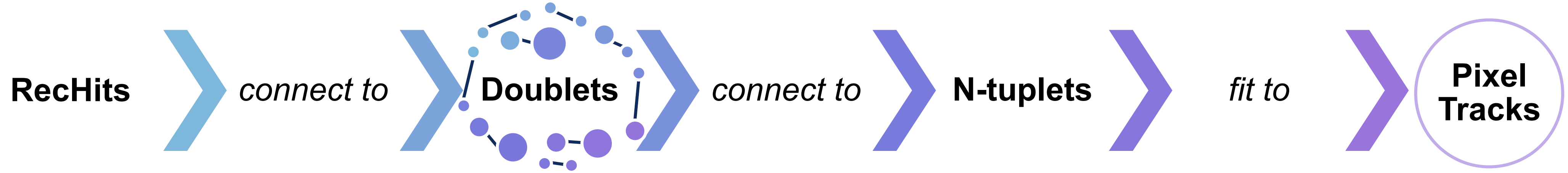
How to Pixel Track? – The Patatrack Algorithm

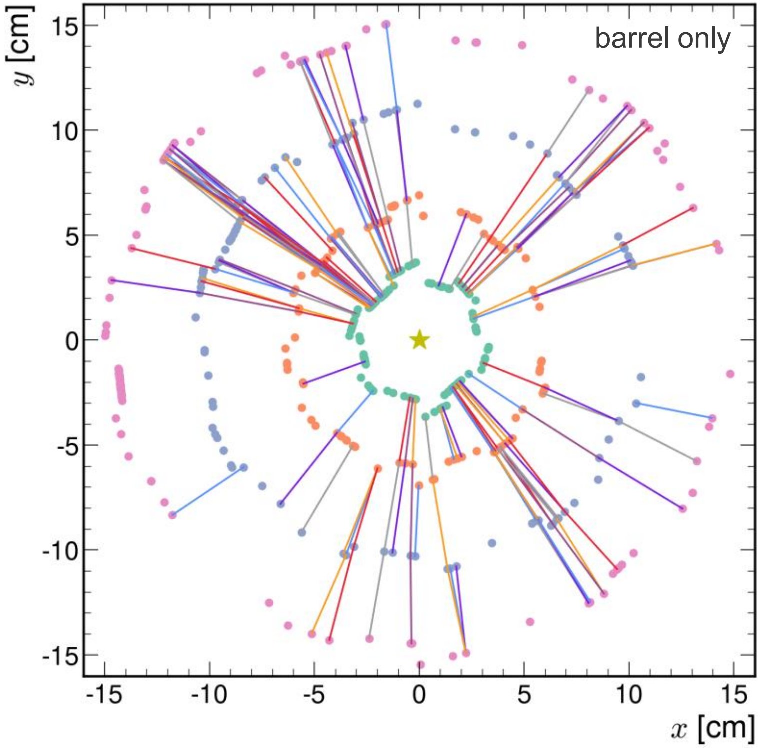
RecHits *connect to* **Doublets** *connect to* **N-tuplets** *fit to* **Pixel Tracks**

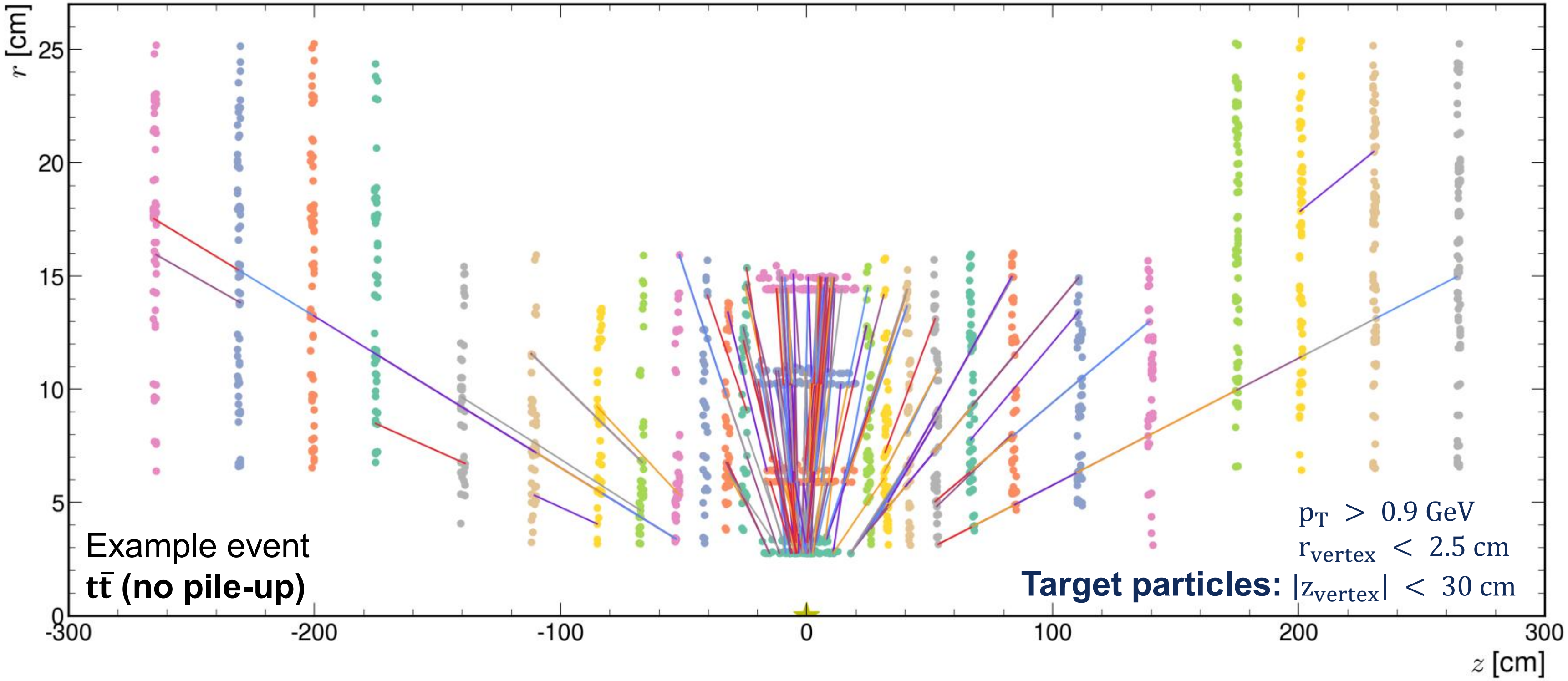




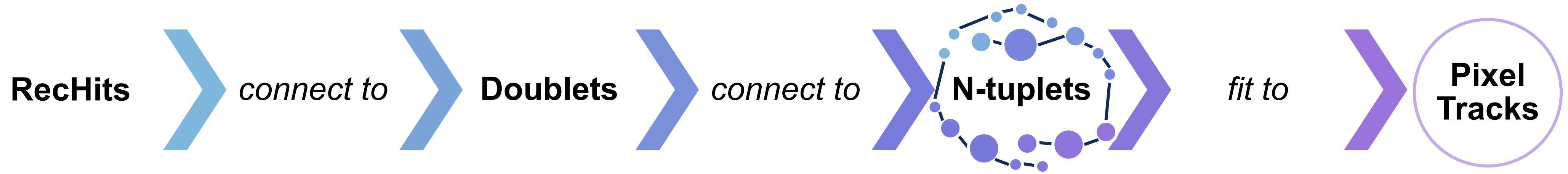
How to Pixel Track? – The Patatrack Algorithm

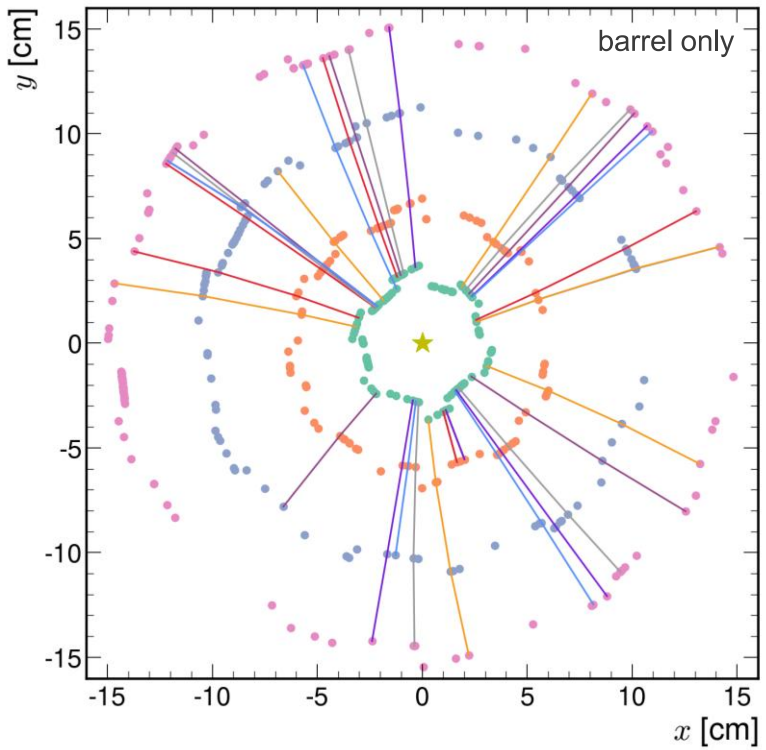


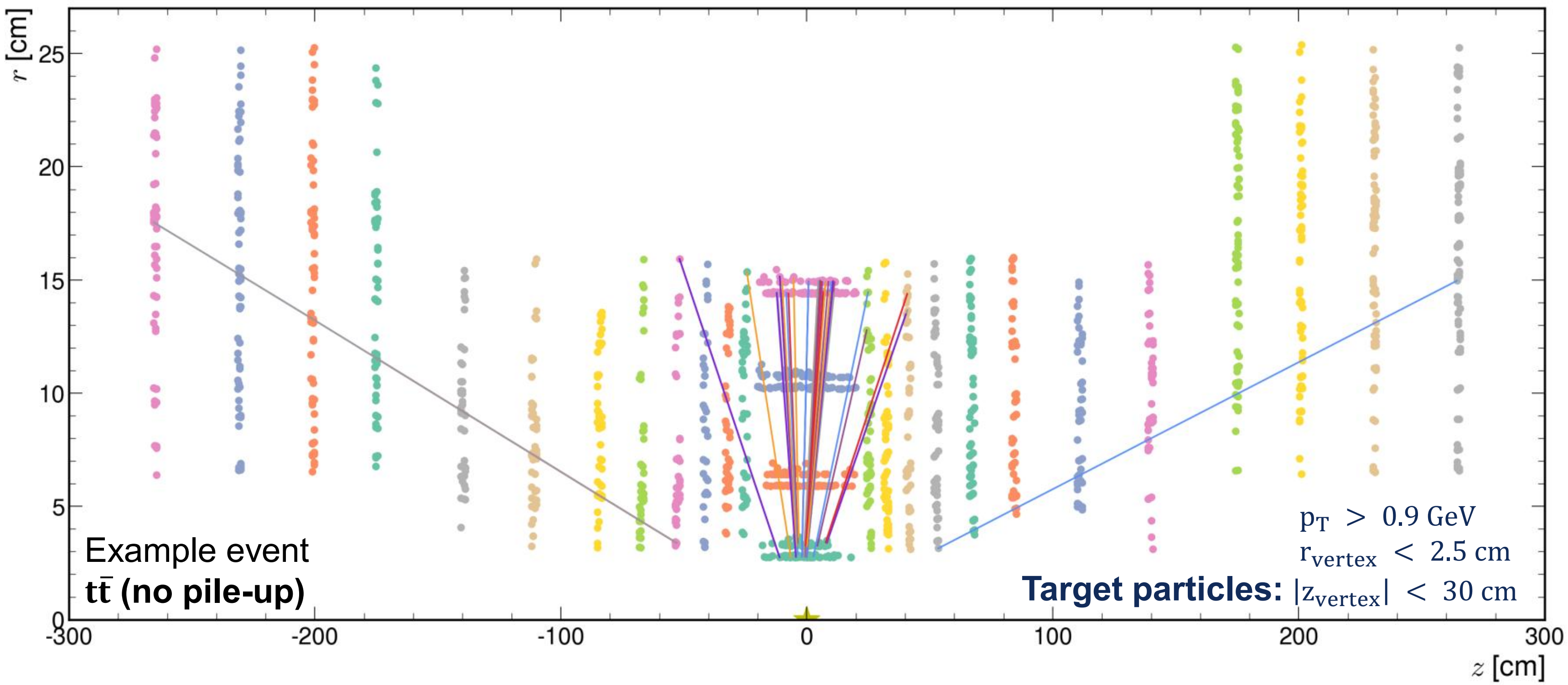




How to Pixel Track? – The Patatrack Algorithm

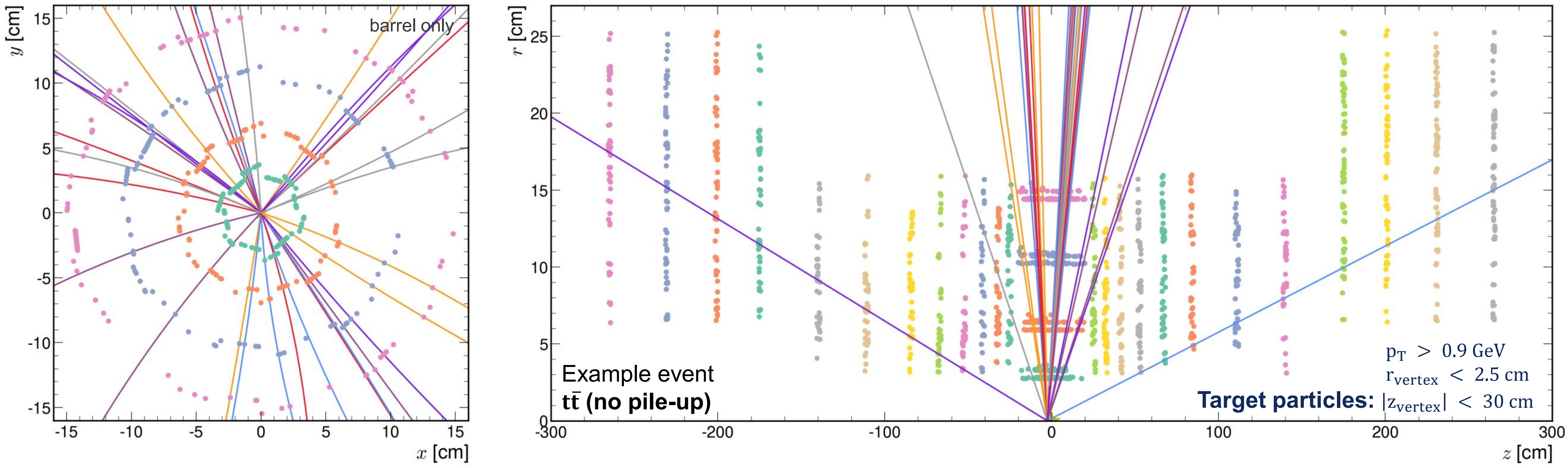






How to Pixel Track? – The Patatrack Algorithm

RecHits *connect to* **Doublets** *connect to* **N-tuplets** *fit to* **Pixel Tracks**



Improving Patatrack Pixel Tracking

Development of **new simulation truth** targeting performance evaluation and optimization of Patatrack (for more details, see poster later)

Tuned the existing configuration

- Establish correct layer connections
- Improve cut parameters for compatibility checks
- Extended set of compatibility checks
- Implemented a *preliminary* track selection

Key issue resolved with further developments

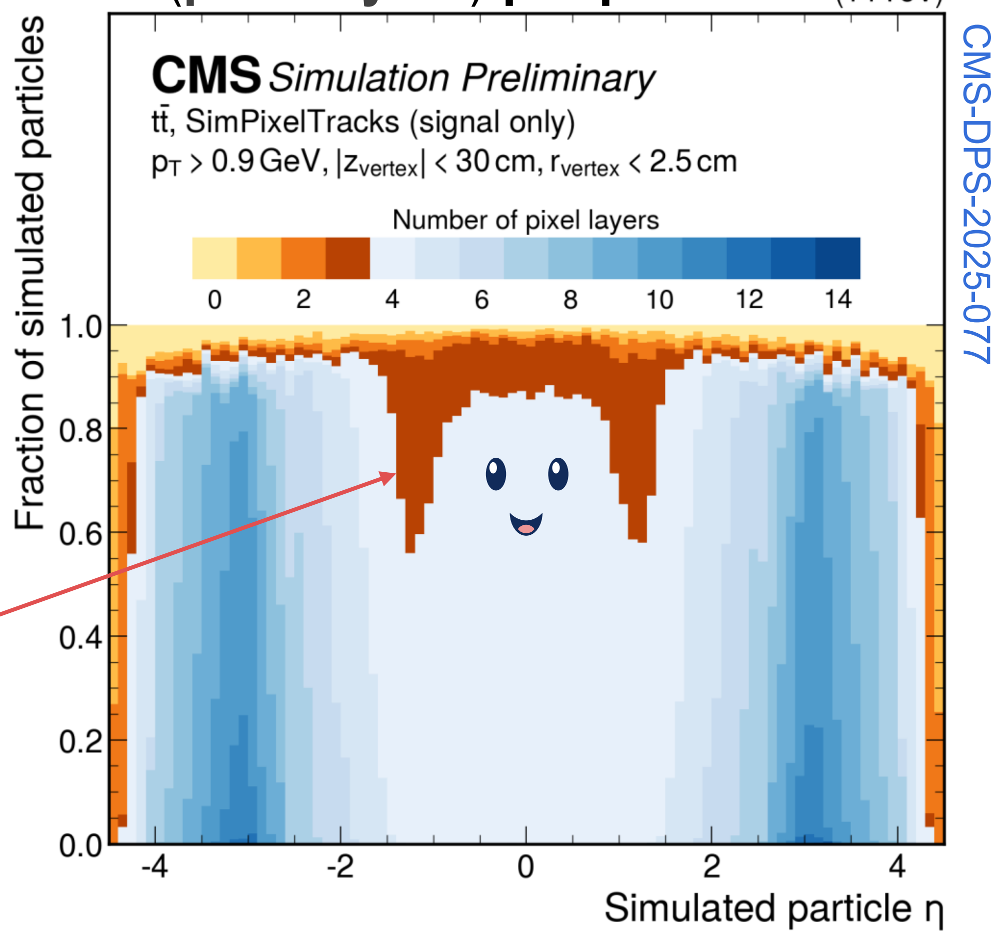
- **No full 4-layer coverage in** the barrel + **transition** from barrel to endcap (require 4 hits for clean pixel tracks)
- **Main reason for second tracking iteration with triplet seeds**

Require less hits per track or rely on second iteration
(disfavored due to exploding fake rate)

?

Collect more hits
(next slide)

N(pixel layers) per particle



Extending Patatrack to the Outer Tracker

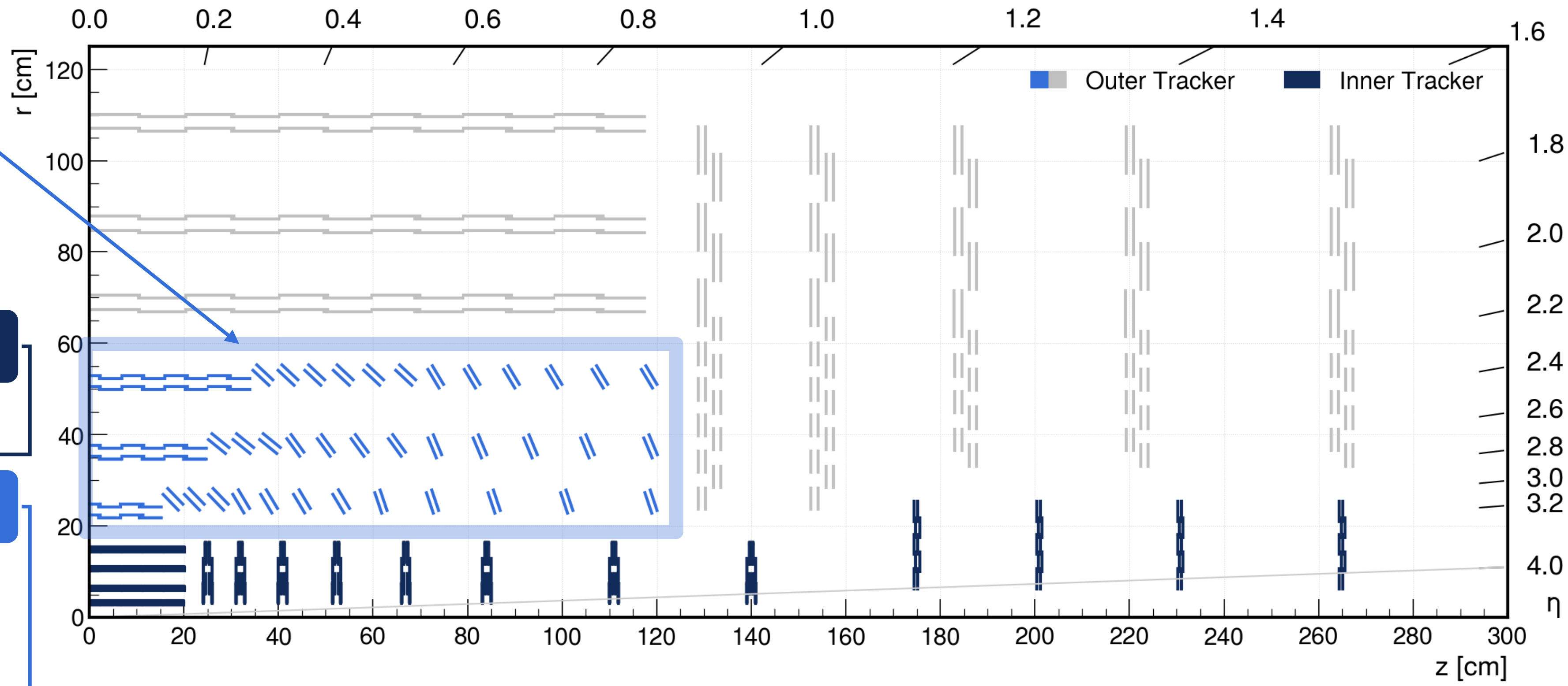
- ❖ Add **hits from macro-pixel part of the PS modules in the first 3 OT barrel layers** to recover efficiency
- ❖ 2 configurations presented:

Patatrack (IT, ≥4 hits)

- Using **only the IT** as before

Patatrack (IT + 3 OT layers, ≥4 hits, ≥2 IT hits)

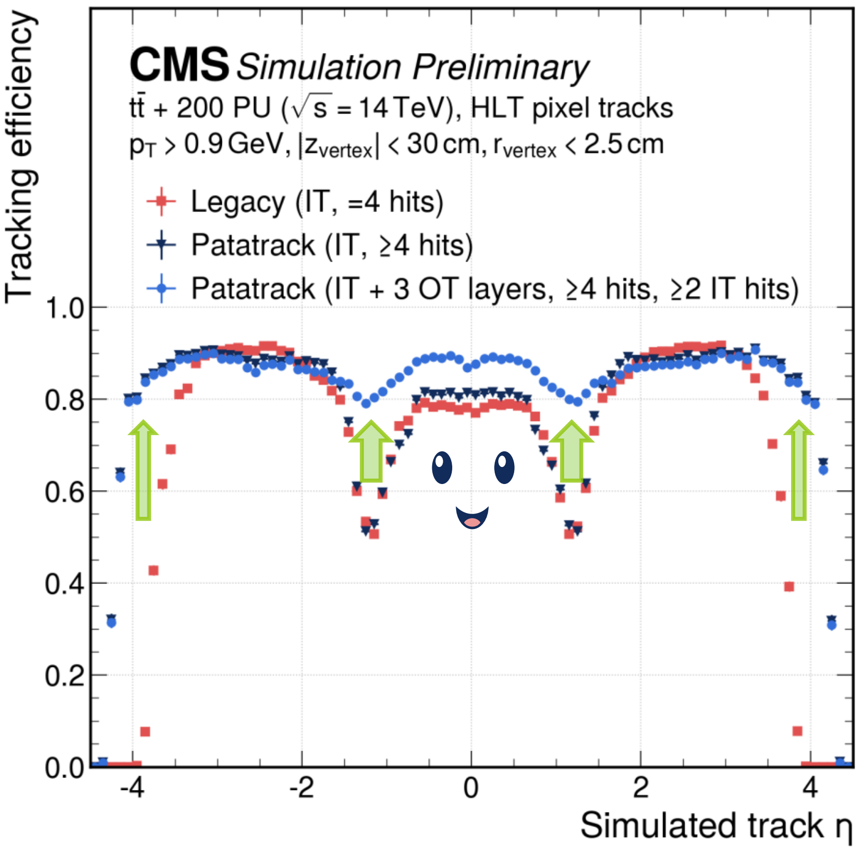
- Using **IT** + **additional OT layers**
- Longer tracks allow for tighter track selection
- High-purity tracks with as few as 2 IT hits



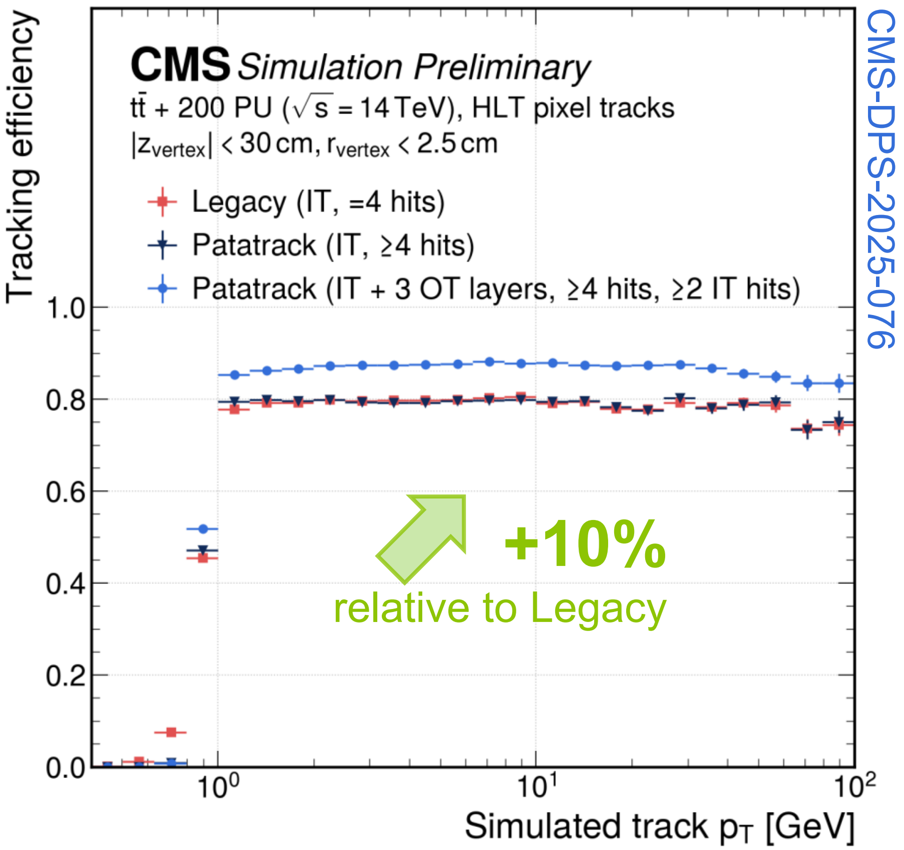
Pixel Tracking **efficiency**

definition here

Efficiency vs η



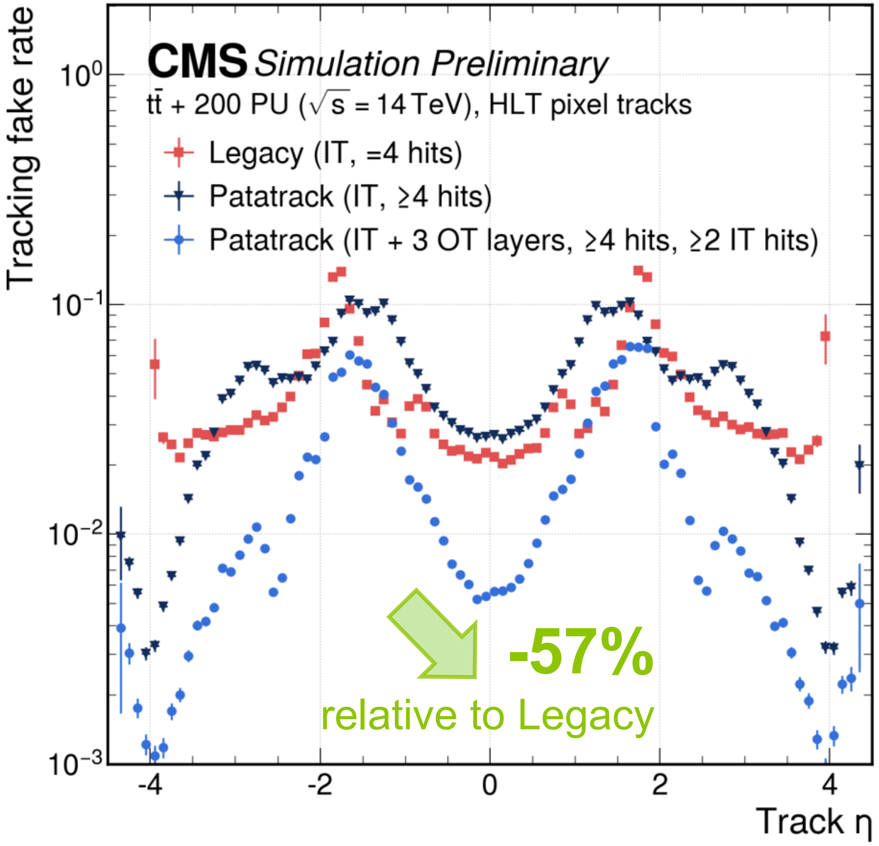
Efficiency vs p_T



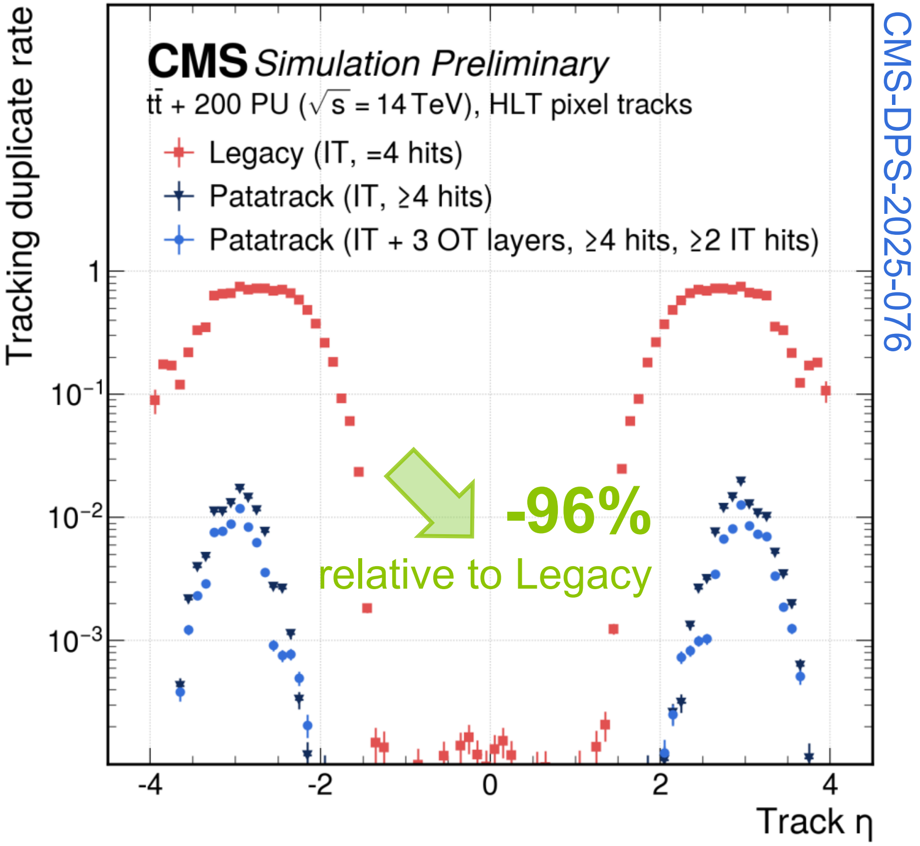
- **Legacy**
  - Low efficiency in barrel + transition to endcap
- **Patatrack (IT only)**
  - Low efficiency in barrel + transition to endcap
  - Extended efficiency at large $|\eta| > 3.5$
  - Overall efficiency similar to Legacy
- **Patatrack with OT-extension**
  - great boost in performance in barrel + transition to endcap
  - best efficiency overall: >5% more across full p_T band ($p_T >$ 0.9 GeV)
  - **Recover inefficiencies due to tracks with ≤3 IT hits**

Pixel Tracking **fakes & duplicates**

Fake rate



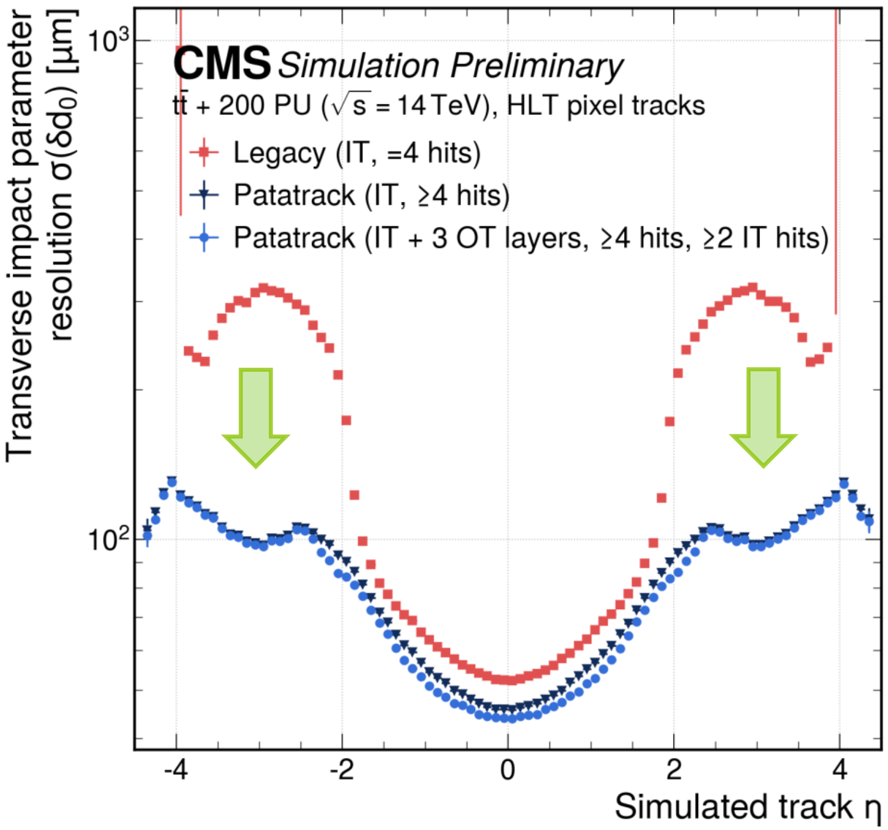
Duplicate rate



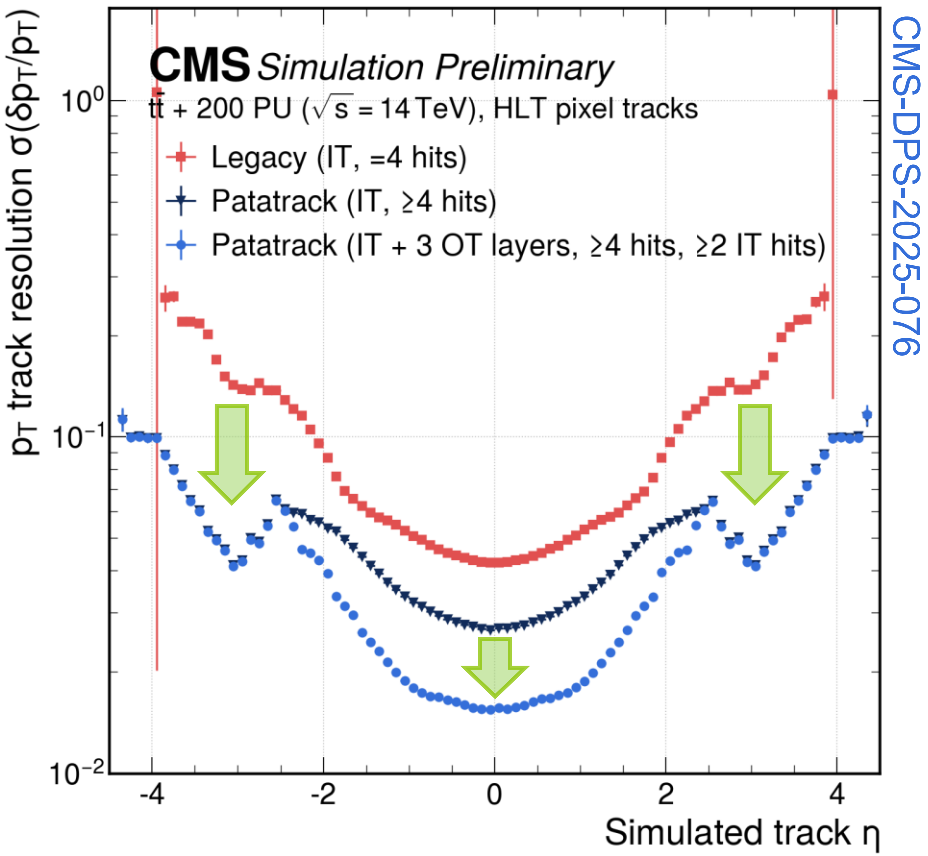
- **Legacy**
  - ❖ Very high duplicate rate in endcaps due to 4-hits per track limitation
- **Patatrack (IT only)**
  - ❖ Fake rate similar level as Legacy
  - ❖ Very low duplicate rate
- **Patatrack with OT-extension**
  - ❖ Lowest duplicate rate
  - ❖ Lowest fake rate due to possibility for tighter selection requirements thanks to longer tracks

Pixel Tracking **resolutions**

Impact parameter resolution



p_T resolution



- **Legacy**
  - Poor track quality due to limited number of hits and no proper fit
- **Patatrack (IT only)**
  - Improved track quality compared to Legacy
- **Patatrack with OT-extension**
  - Additional quality improvement for $|\eta| < 2.5$ because of additional OT hits
  - Especially p_T resolution improved thanks to larger sagitta

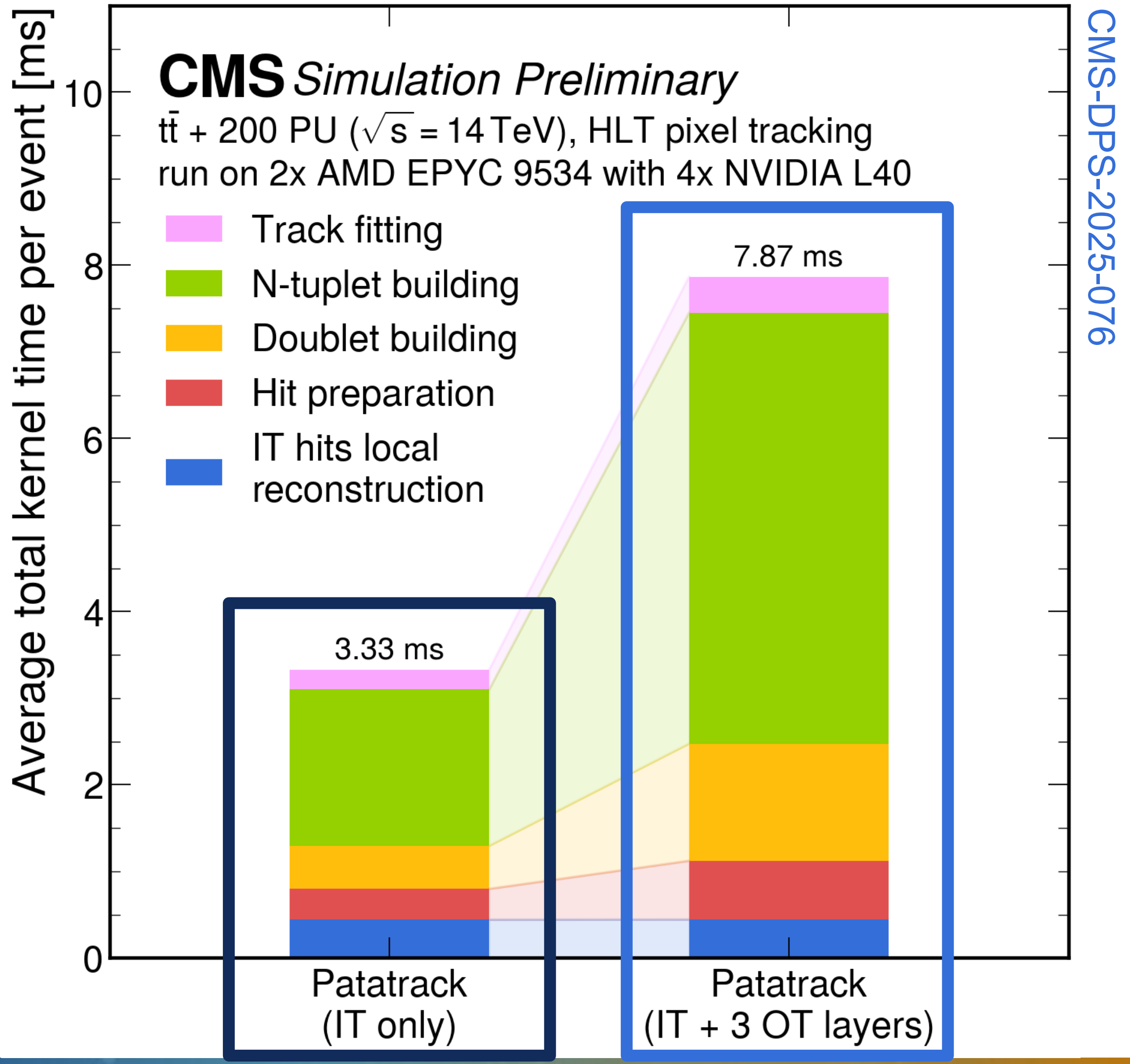
Pixel Tracking **GPU time**

Legacy
does not run on GPU

Patatrack (IT only)
baseline GPU time on the reference machine
~3.3 ms per event → very fast

Patatrack (IT only + 3 OT layers)
kernel times increase to ~7.9 ms → still very fast

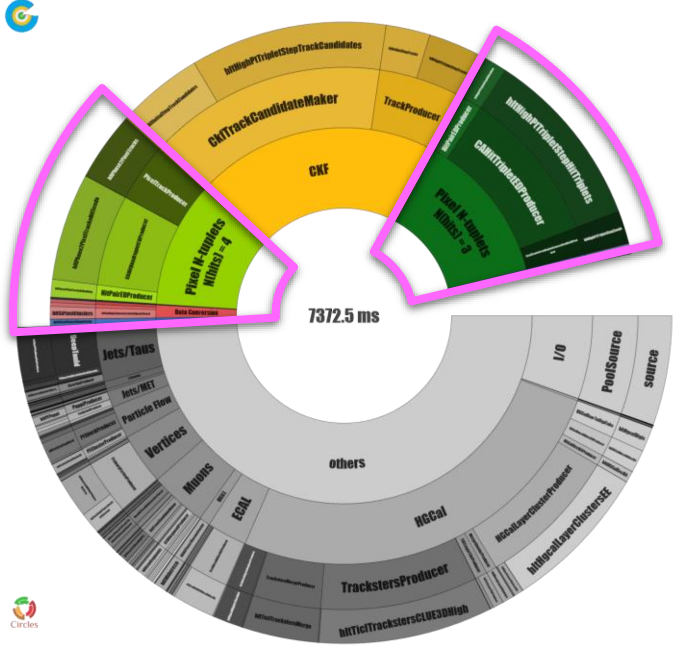
Execution time (all kernels)



Execution time: big picture

Execution time (full HLT)

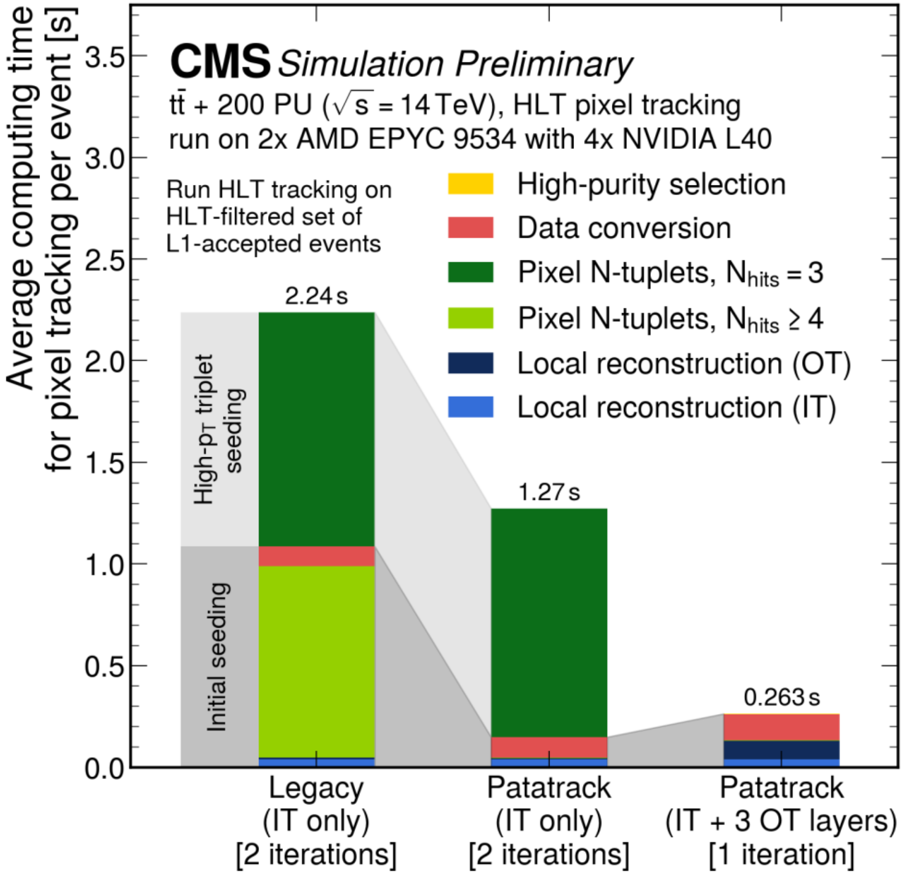
CMS-DPS-2025-076



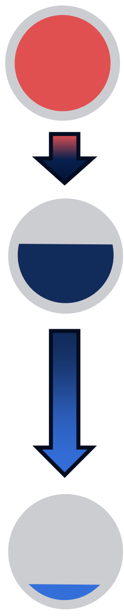
Legacy configuration
[2 tracking iterations]

Full pie chart	full **HLT**	~7.4 s
Colored part	full **tracking**	~3.5 s
Pink part	total track **seeding**	2.24 s

Execution time (track seeding)



CMS-DPS-2025-076



Legacy

Patatrack (IT only) **-43%** relative to Legacy

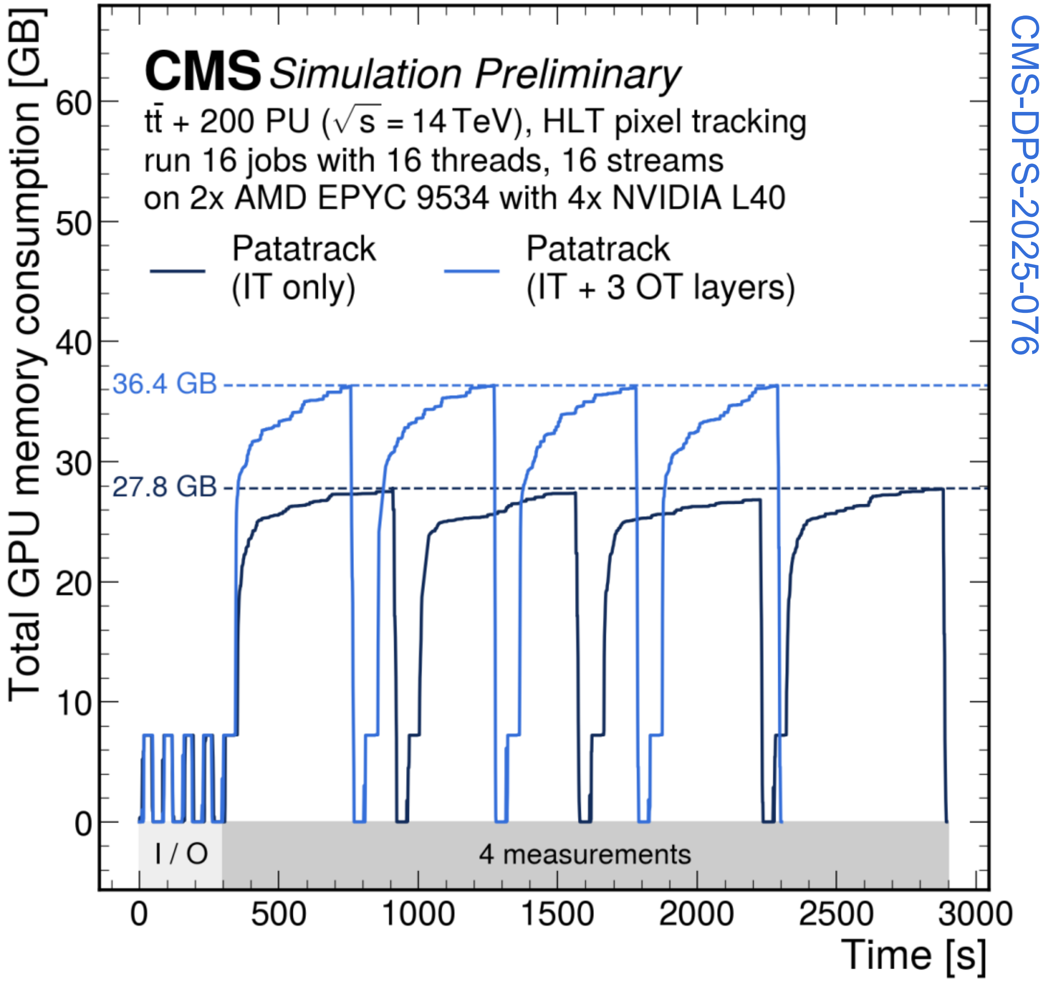
Greatly **speedup first seeding iteration** *(CPU to GPU)*

Patatrack (IT + 3 OT) **-88%** relative to Legacy

Replace two iterations by a single one with OT-extended pixel seeds *(as clean track seeds with as few as 3 IT hits can already be build with quadruplets)*

Extension could save execution time ~2.5 s
(2 s seeding + 0.5 s building+fitting)

Pixel Tracking **GPU memory**



Legacy
does not run on GPU

Patatrack (IT only)
baseline GPU memory of ~130MB per thread

Patatrack (IT only + 3 OT layers)
increase in GPU memory of ~30%

The extension is affordable in terms of GPU memory

Each job runs on 1000 events per measurement.
16 jobs × 16 threads
→ (up to) **256 events processed simultaneously on four L40 GPUs**

Conclusion

Heterogeneous version of CMS Phase-2 pixel tracking has been optimized and extended to first 3 barrel layers of the Outer Tracker

Resulting in:

- **Higher overall efficiency** of initial track seeds
- **Lower fake and duplicate rates** of the same
- **Better track parameter estimates** of the seeds
- Moderate increase in GPU memory
- **Greatly reduced overall tracking time** due to possibility to replace two iterations with single one with fast, clean, high-quality seeds produced on GPU

References

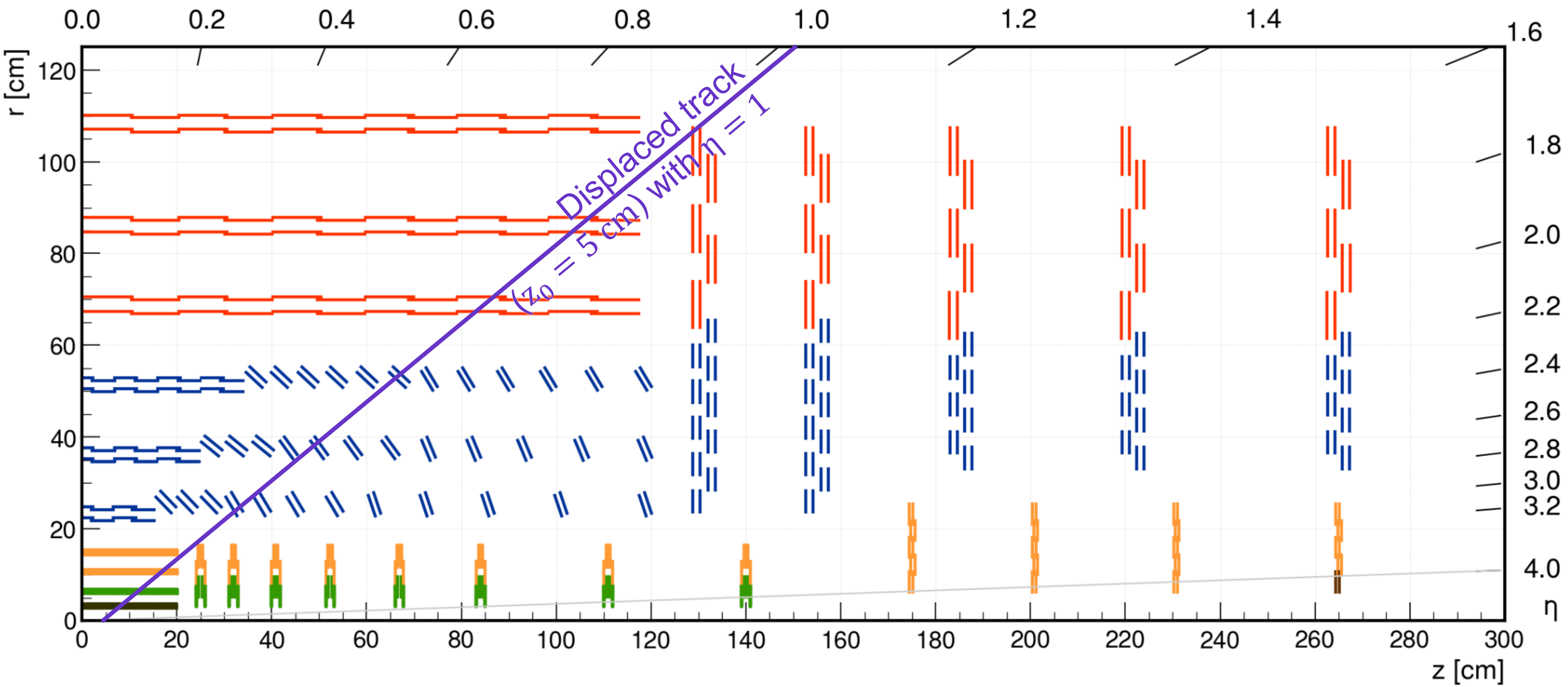
CMS-DP-2025-076

CMS-DP-2025-077

This work has been funded by the Eric & Wendy Schmidt Fund for Strategic Innovation through the CERN Next Generation Triggers project under grant agreement number SIF-2023-004.

Backup

Origin of the ghost



More performance

CMS-DP-2025-076

Tracking performance metrics from CMS-DP-2025-076

Standard tracking performance metrics such as efficiency, fake or duplicate rate are based on the comparison and matching of the reconstructed track collection to the simulated particle collection. A **track is associated to a simulated particle** if more than 75% of its reconstructed hits originate from the simulated particle. A track that does not meet this requirement with any simulated particle is considered a misidentified or **fake track**.

Definition of efficiency: Efficiency is defined as the **fraction of simulated target particles that have an associated track**. The target particles are a subset of all simulated particles and are defined as the charged particles originating from the signal (hard scattering) vertex with some additional phase-space selections:

- A transverse momentum, p_T, of more than 0.9 GeV.
- A production vertex with radial displacement to the center of the CMS detector, r_{vertex}, of less than 2.5 cm.
- A production vertex with absolute longitudinal displacement to the center of the CMS detector, $|z_{vertex}|$, of less than 30 cm.

Those cuts are explicitly mentioned in all subsequent efficiency plots in this document. Note that the efficiency plots over one of those three variables do not include the respective selection of that parameter.

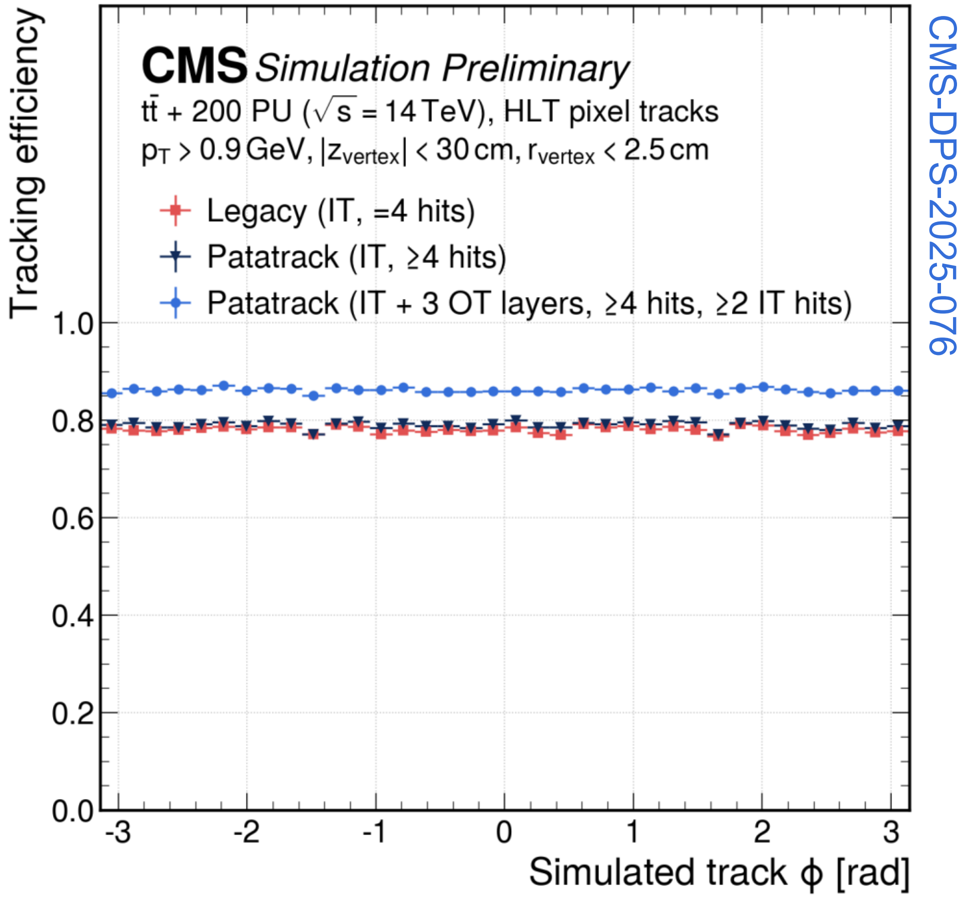
Definition of fake rate: Fake rate is defined as the fraction of **number of fake tracks over the total number tracks**.

Definition of duplicate rate: Duplicate rate is defined as the **number of tracks** that are **associated to simulated particles with multiple** associated **tracks over the total number of tracks**.

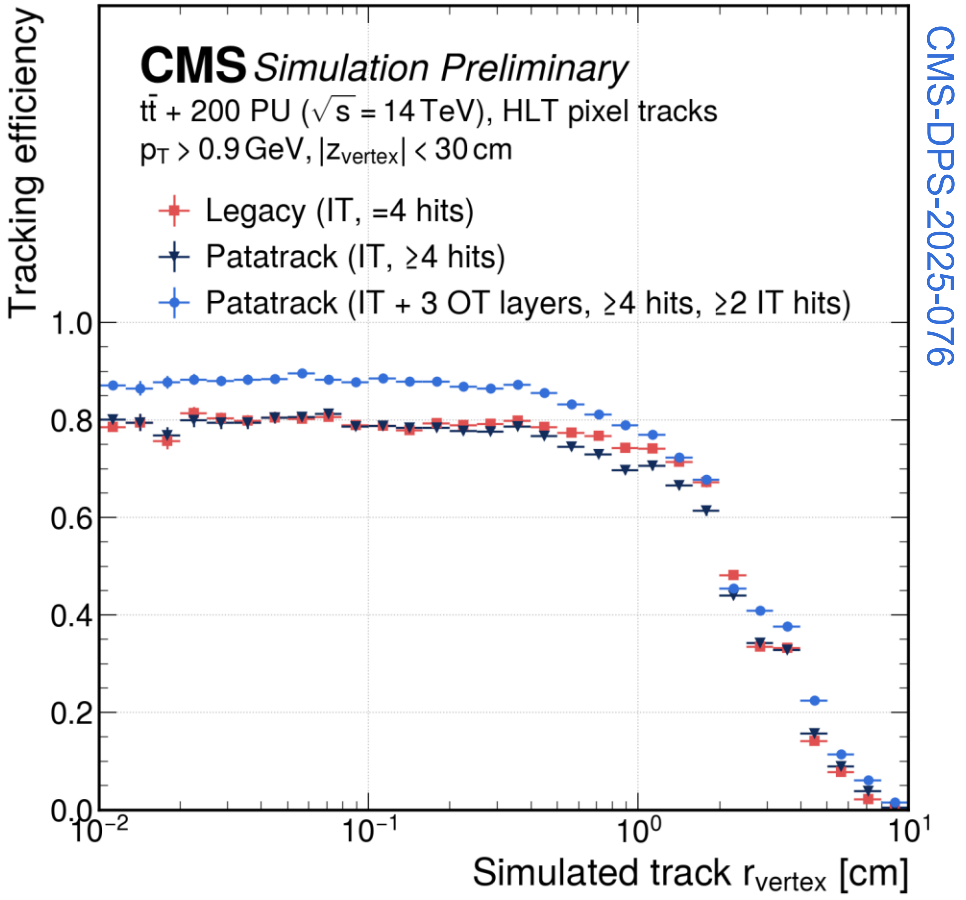
The physics performance of the pixel tracks are measured with 5000 events of simulated $t\bar{t}$ interactions with on average 200 additional pileup interactions at a center of mass energy of 14 TeV. The CMS Phase-2 detector geometry is used and simulated under ideal conditions, i.e., excluding any module failure scenarios.

Pixel Tracking **efficiency**

Efficiency vs ϕ

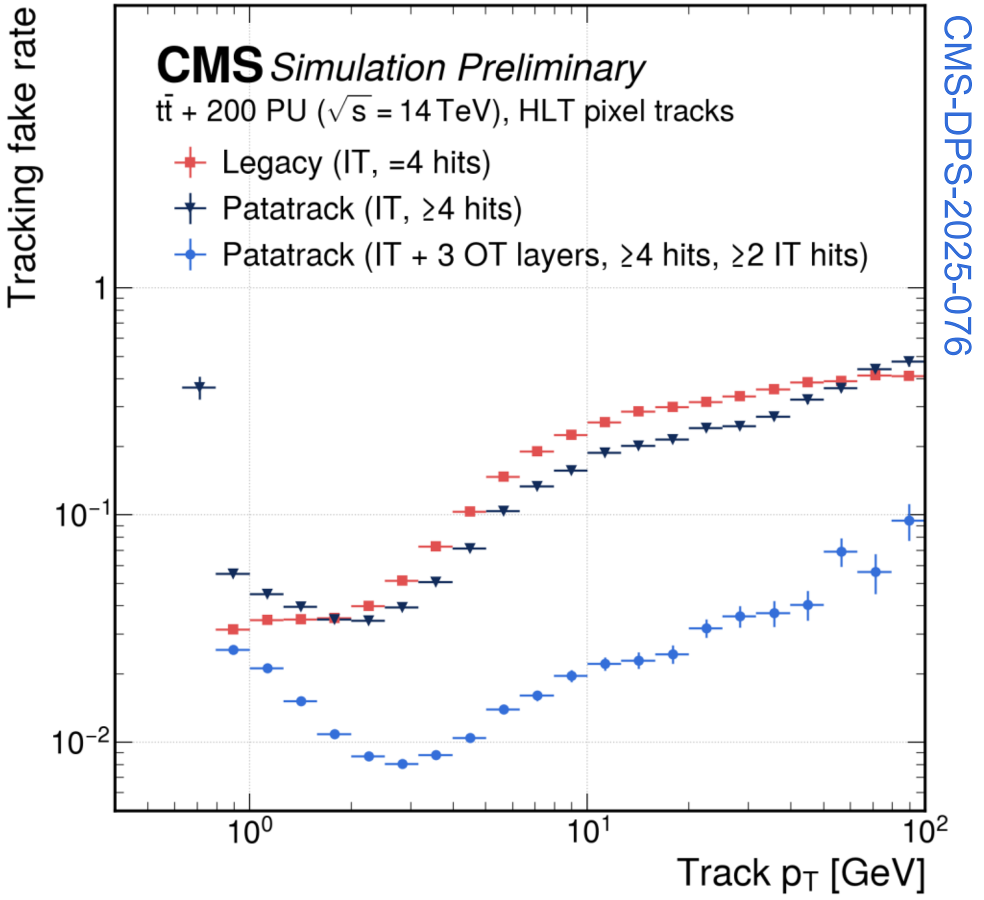


Efficiency vs r_{vertex}

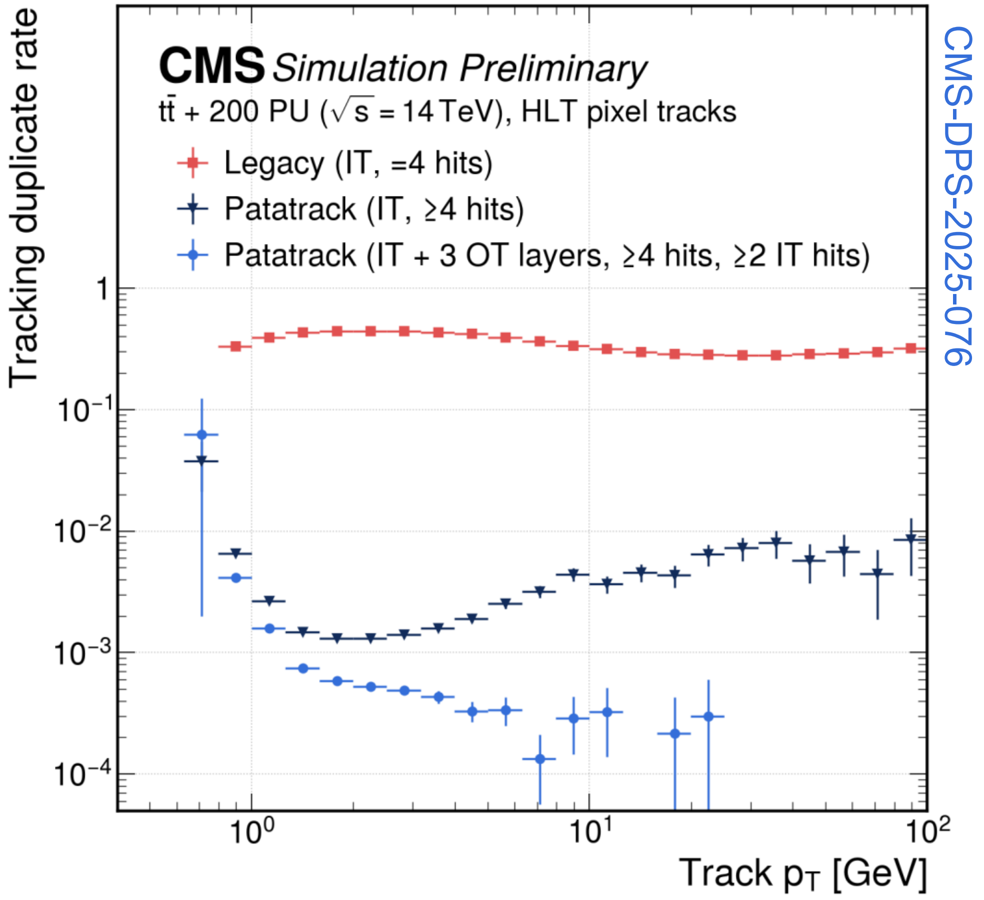


Pixel Tracking **fake & duplicates** vs p_T

Fake rate

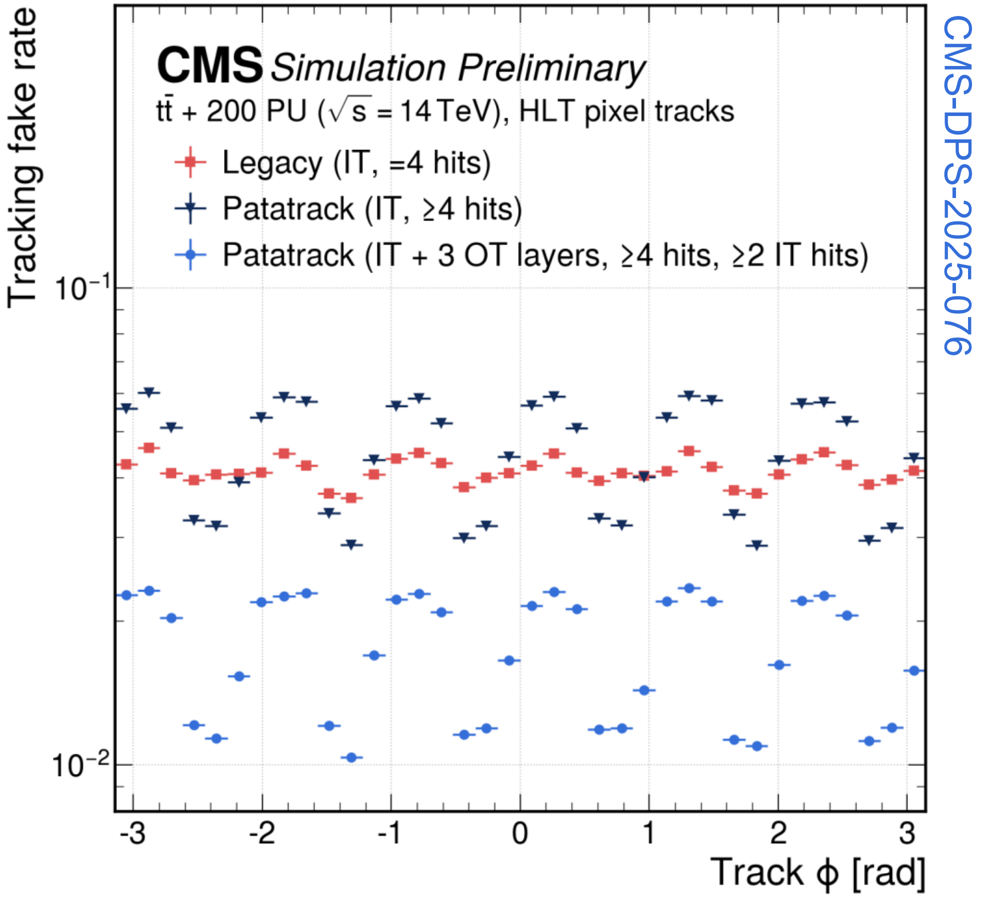


Duplicate rate

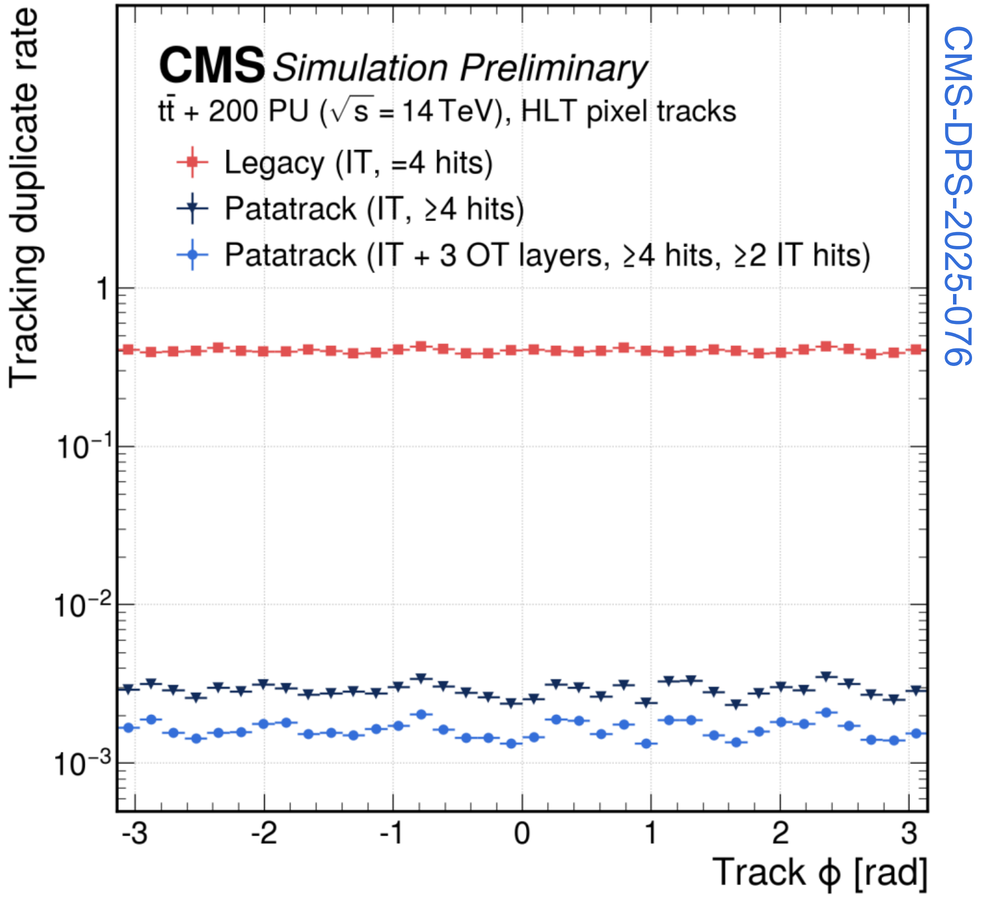


Pixel Tracking **fake & duplicates** vs ϕ

Fake rate

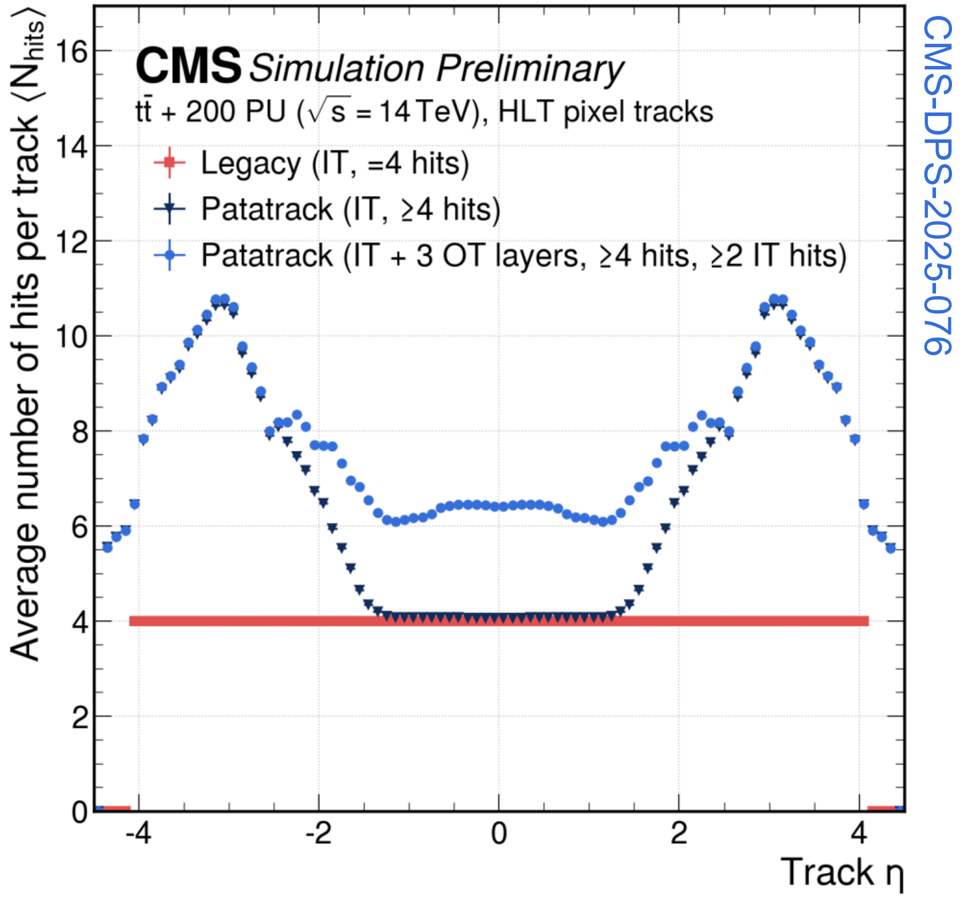


Duplicate rate

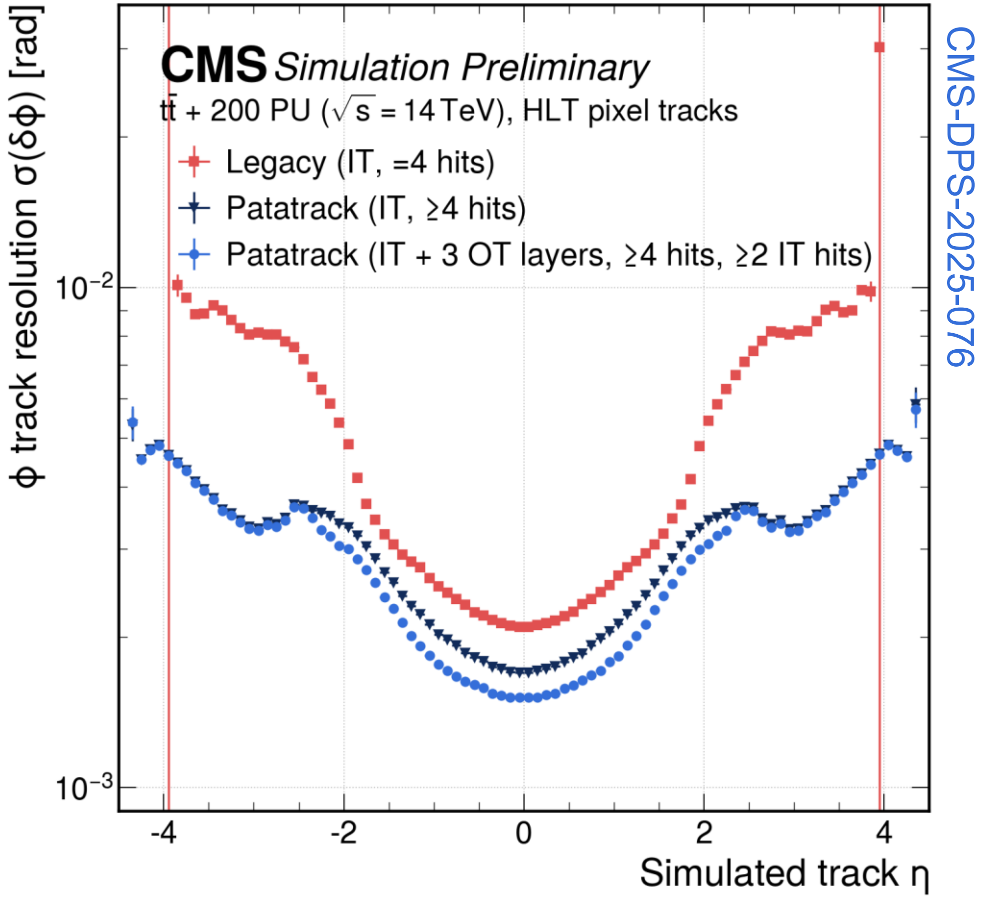


Pixel Tracking **hits | resolution**

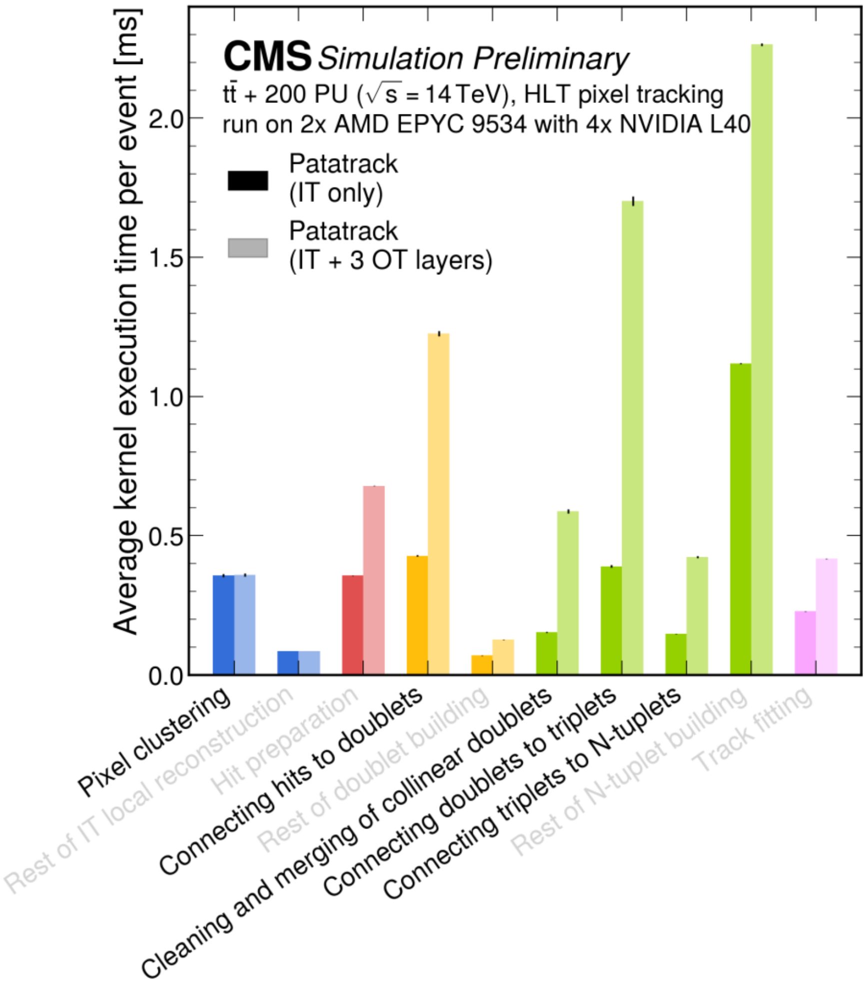
Number of hits



ϕ resolution



Pixel Tracking **kernel times**

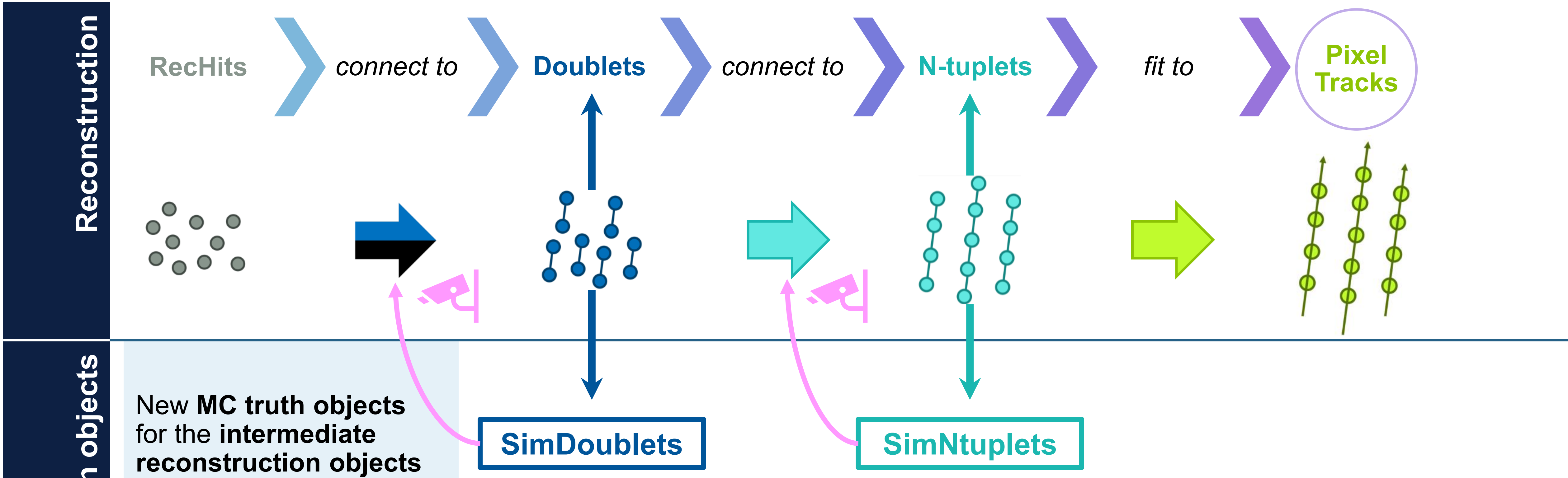


CMS-DPS-2025-076

SimPixelTrack optimization

CMS-DP-2025-077 | poster

SimPixelTracks – Developed new MC truth



New **MC truth objects** for the **intermediate reconstruction objects**

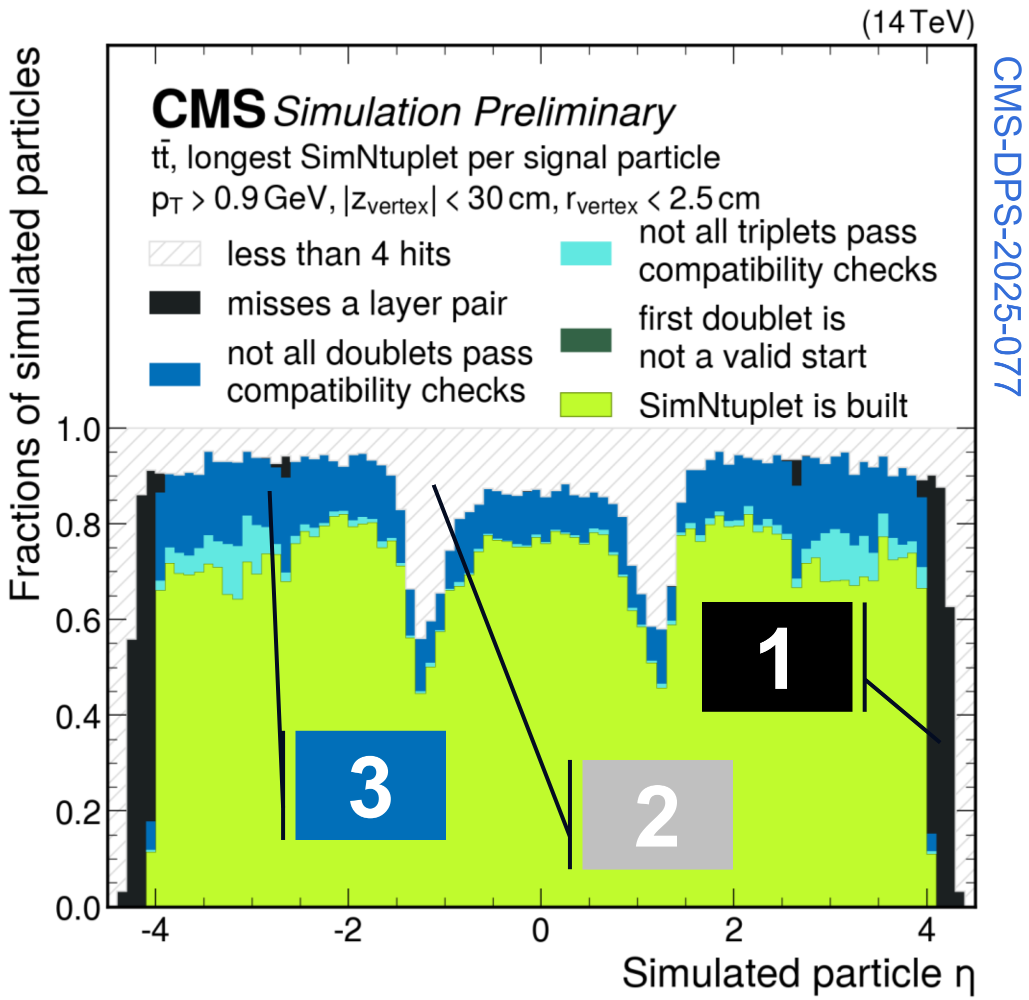
→ **true RecHit connections of simulated particles**

→ **more control when optimizing**

→ Enables detailed analysis of existing **parameters** for **building** and connecting **doublets**

→ Analyze at **which** reconstruction **step** we **lose efficiency** and **why**

SimPixelTracks – Investigate Patatrack efficiency



Not enough RecHits in the Pixel detector:

≤ 3 RecHits

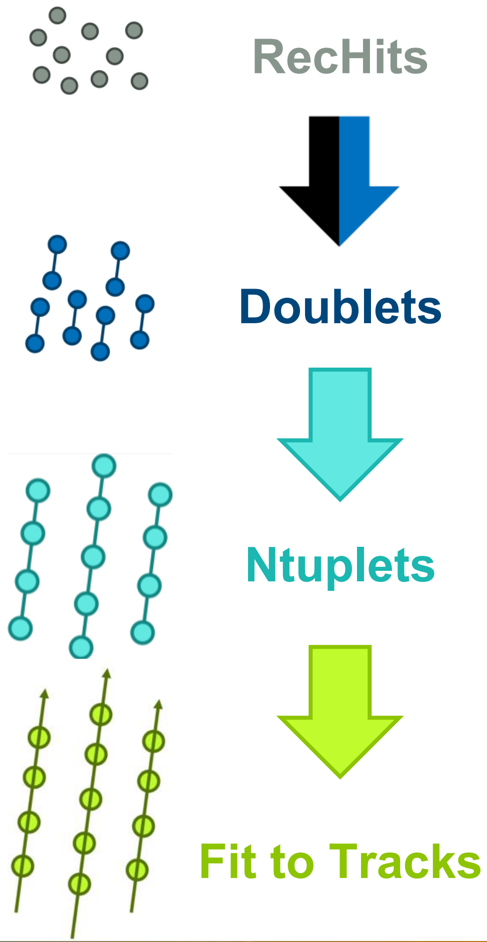
Sim**Doublet not built**…

…layers not connected

…compatibility checks not passed

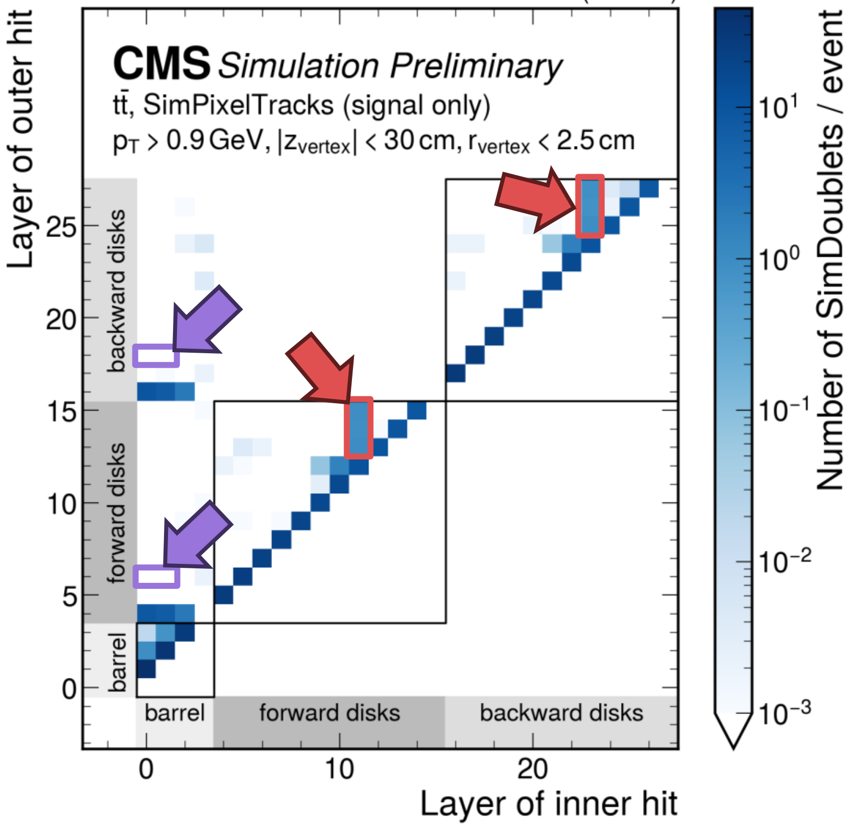
Sim**Ntuplet not built…** doublet connection compatibility check not passed

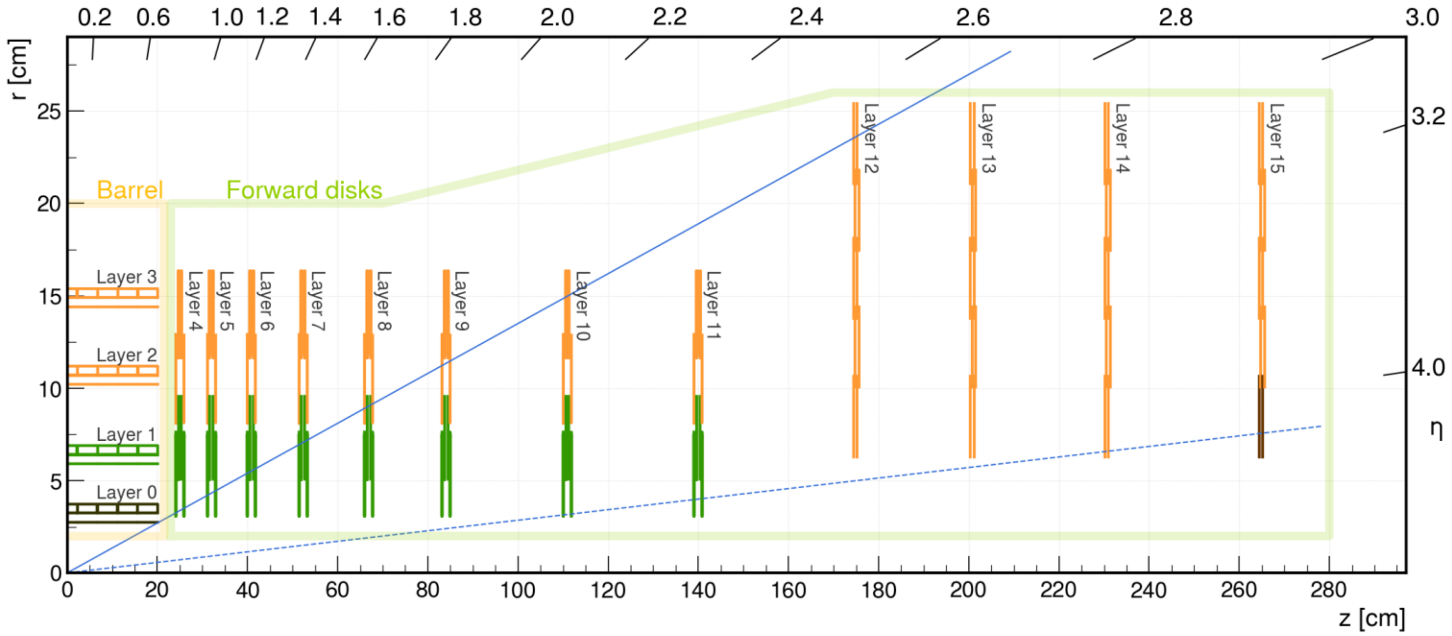
Sim**Ntuplet** is **built** and will be **fitted**



1 Tackling the **doublet losses** by optimizing layer pairs

CMS-DPS-2025-077





2 Use more RecHits

Extending Patatrack to the Outer Tracker

❖ Add **hits from** **macro-pixel part of the PS modules in the first 3 OT barrel layers** to recover efficiency

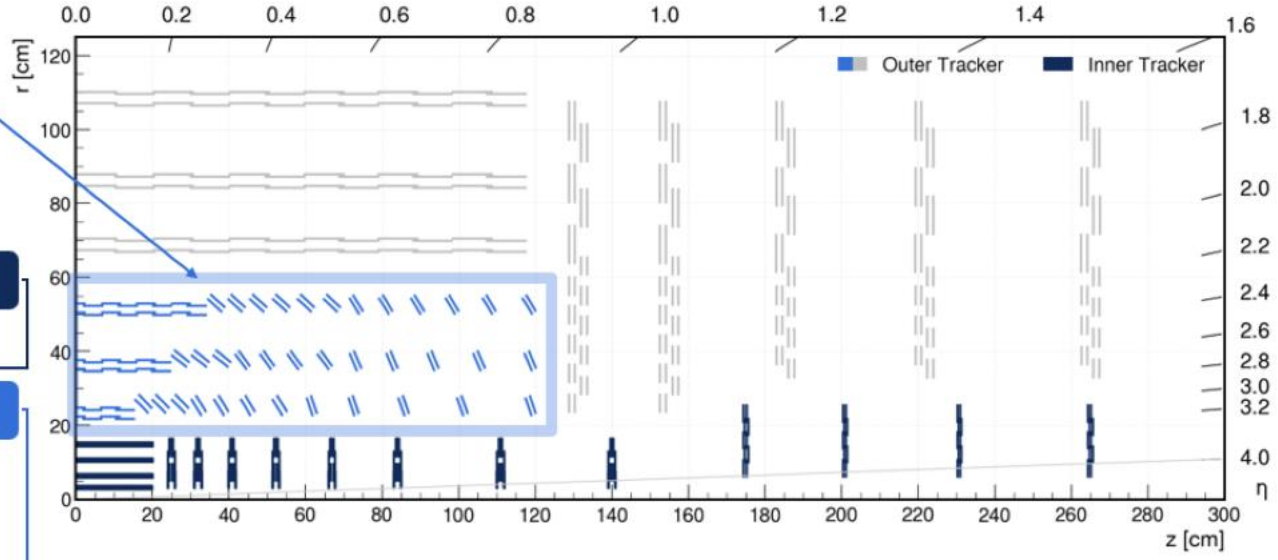
❖ 2 configurations presented:

Patatrack (IT, ≥4 hits)

- Using **only the IT** as before

Patatrack (IT + 3 OT layers, ≥4 hits, ≥2 IT hits)

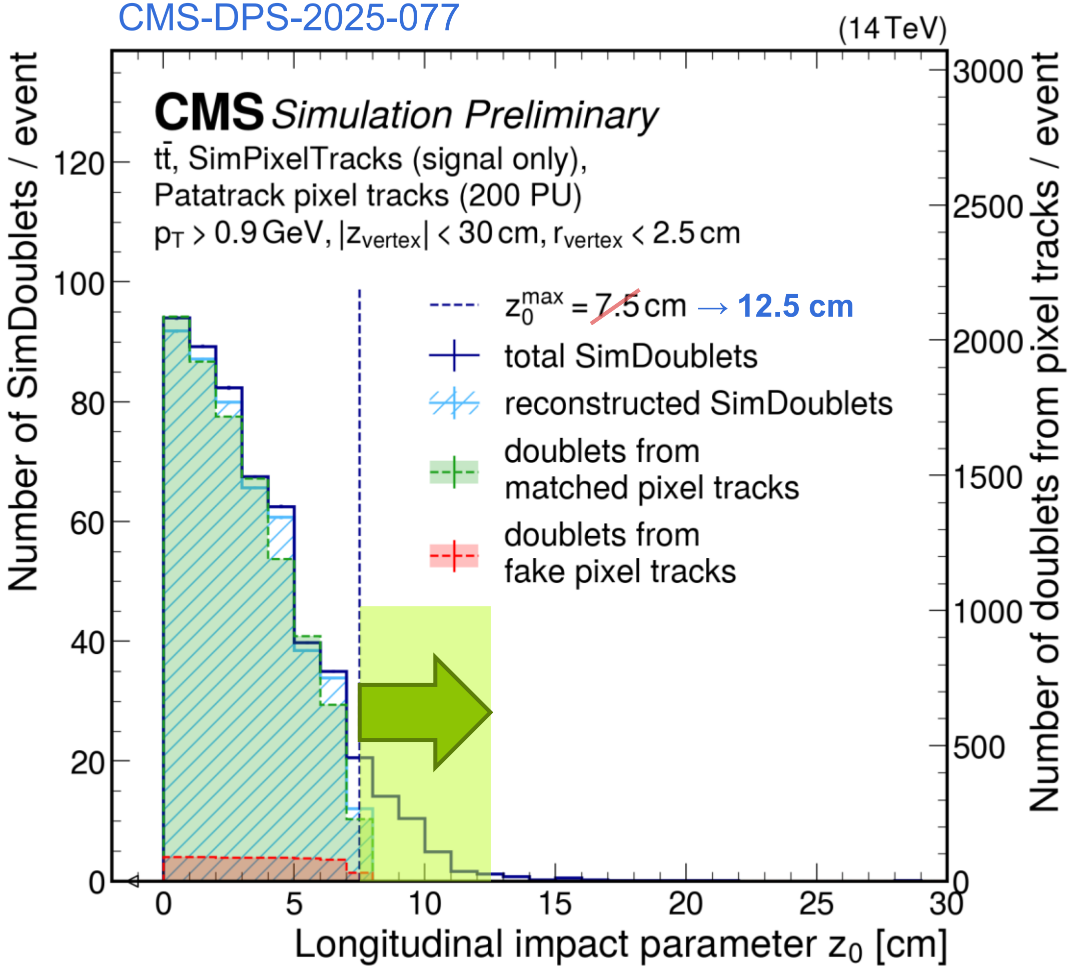
- Using **IT** + **additional OT layers**
- Longer tracks allow for tighter track selection
- High-purity tracks with as few as 2 IT hits



Nov 11, 2025 CTD 25 | Optimized Pixel Tracking for CMS Phase-2 | Jan Schulz 8

3 Optimizing doublet cut parameters

CMS-DPS-2025-077



New tool **SimDoublets** enables accurate **diagnosis** and analysis of **existing cuts**

Precise knowledge on efficiency of individual compatibility checks

Easy identification of problems

Set all cuts directly according **to the true distributions**

CERN

NexTGen

Next Generation Triggers