Simulation truth for optimal CMS pixel tracking

Jan Gerrit Schulz on behalf of the CMS Collaboration

CERN, Switzerland | RWTH Aachen University, Germany

Introduction

The Patatrack pixel tracking algorithm [1], which is used in the CMS experiment, is an example of a tracking algorithm with many parameters ($O(100)$). Relying purely on automated parameter tuning with only the final pixel tracks as an objective can have several drawbacks:

- ❖ potentially long convergence time
- ❖ no guarantee for optimal outcome
- ❖ no further insight into the algorithm

→ makes it difficult to even define optimal performance

In this work [2], we explore an extended approach by developing simulation truth objects that correspond to the various intermediate objects built in the Patatrack reconstruction:

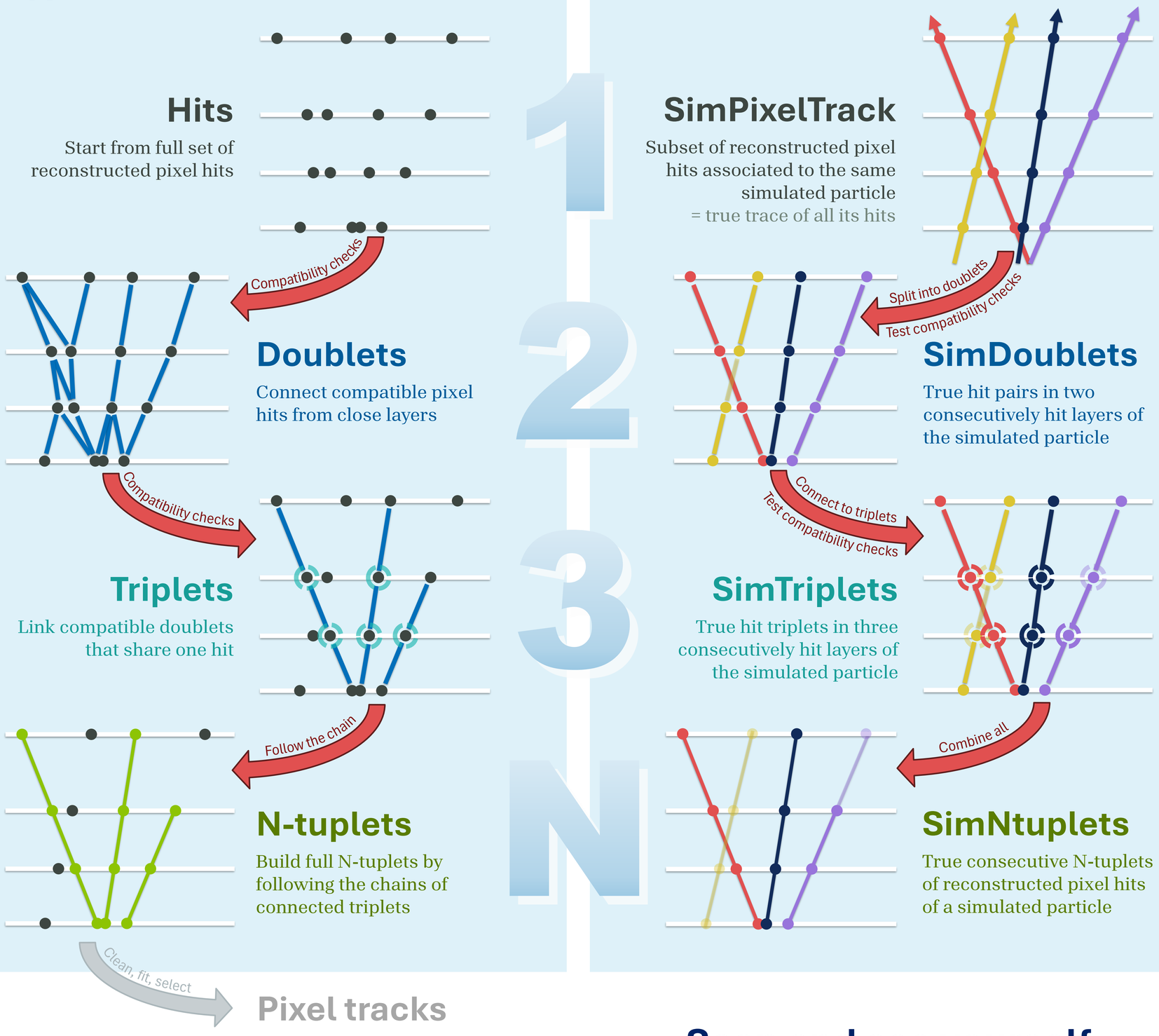
- ✓ in-depth analysis of algorithm's performance
- ✓ insights, such as direct identification of issues that lead to miss of individual particles or hits
- ✓ optimize parameters with more control
- ✓ deeper understanding → improve algorithm itself

Pixel track reconstruction vs Simulation truth objects

Patatrack pixel tracking [1] based on cellular automata → allows high parallelization + fast execution on GPUs.

Algorithm builds tracks by connecting pixel hits in multi-step process:

To monitor and evaluate performance, Patatrack's intermediate objects are built from simulation truth. Truth objects = true connections between reconstructed hits that originate from the same simulated particle:

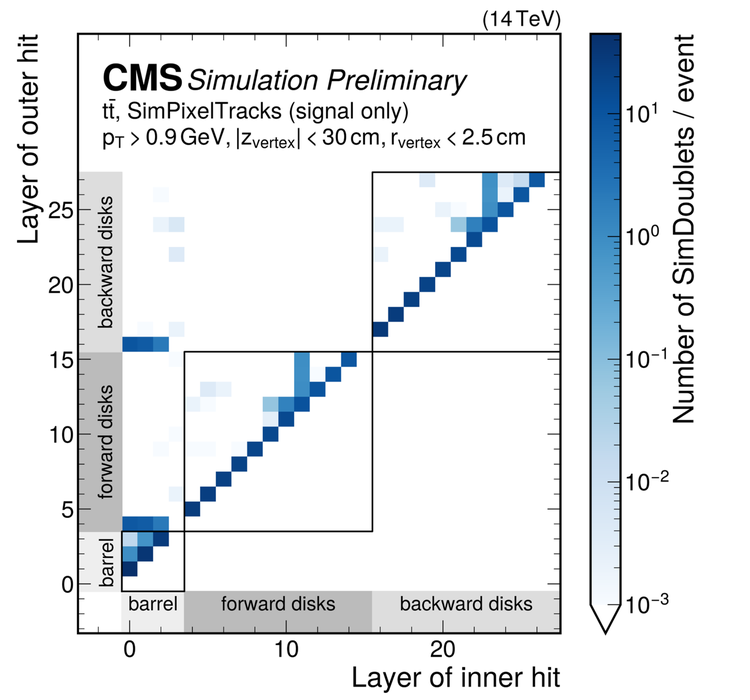


Performance analysis

These simulation truth objects of simulated signal particles can be used to analyze the various reconstruction steps individually.

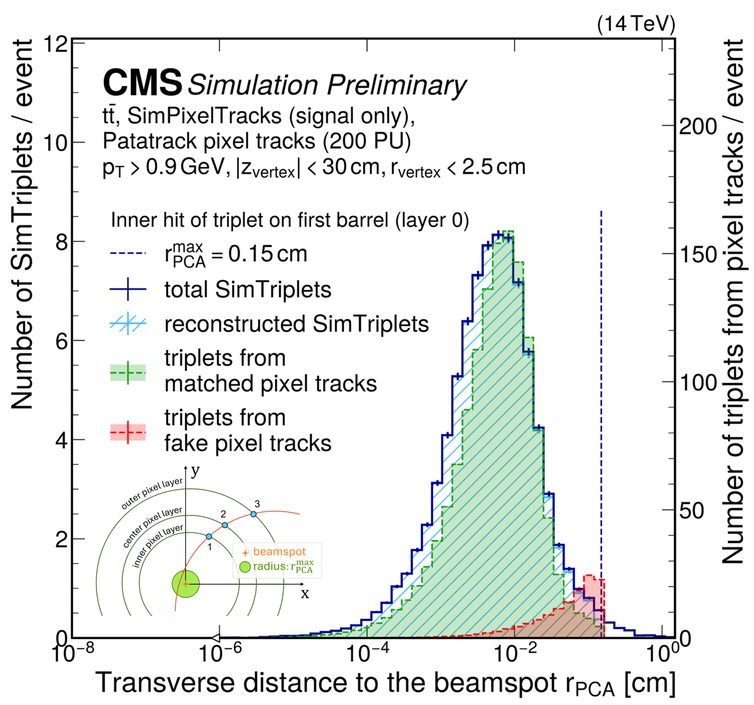
Layer pairs

Choose layer pairs for doublet building according to occupancy of true **SimDoublets** [2]



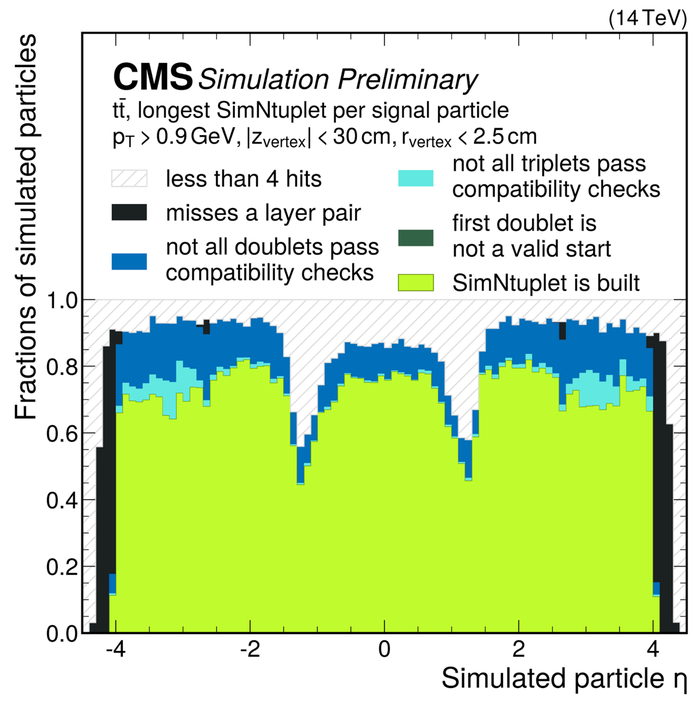
Compatibility checks

Set cuts for doublet/triplet building according to true distributions of **SimDoublets**/**SimTriplets** [2]



Where do we lose particles or hits?

Find reason for lost efficiency or low track quality: check at which reconstruction step the longest **SimNtuplet** per simulated signal particle is lost [2]



Results and conclusions

Work presented was successfully used to analyze and optimize Patatrack pixel tracking for Phase-2 of the CMS High-Level Trigger [3].

Identified and addressed items:

- ✓ missing layer pairs for doublets
- ✓ misconfigured compatibility checks (adjusted according to true signal distributions)
- ✓ issue of insufficient number of hits + investigation of a potential extension of the algorithm to the Outer Tracker (which was then implemented in the reconstruction)
- ✓ development of useful new compatibility checks

Scan and see yourself

References and Acknowledgements

[1] A. Bocci et al. (2020) Front. Big Data, 3.

[2] CMS Collaboration (2025) CMS-DP-2025-077.

[3] CMS Collaboration (2025) CMS-DP-2025-076.

Supported by the Eric & Wendy Schmidt Fund for Strategic Innovation (grant agreement SIF-2023-004).