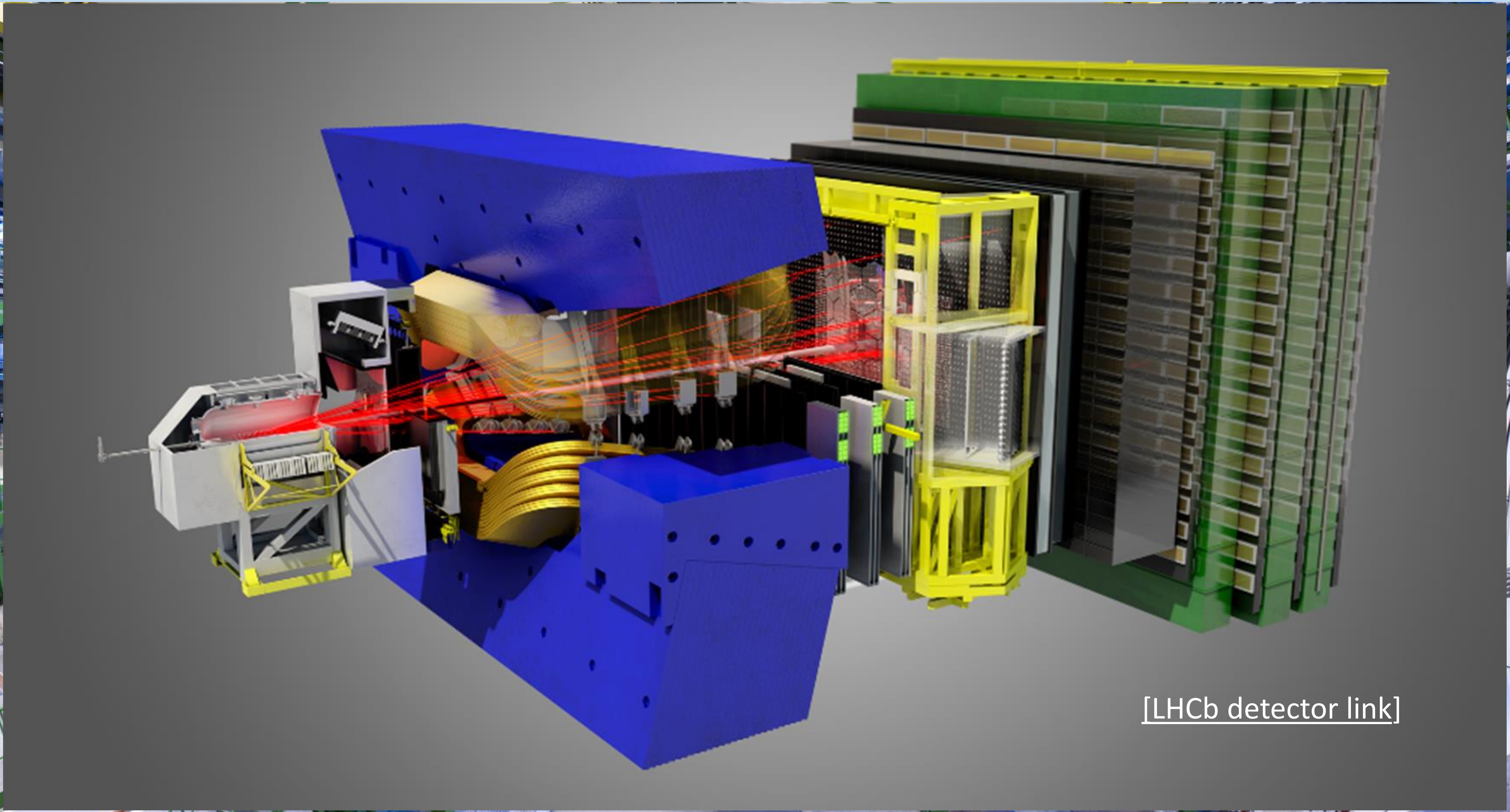
Standalone track reconstruction and matching algorithm for GPU based High Level Trigger at LHCb

Brij Kishor Jashal, Instituto de Física Corpuscular (IFIC), Valencia
Christina Agapopoulou, Lukas Calefice, Arthur Hennequin, Louis Henry,
Arantza Oyanguren, Lorenzo Pica, Jiahui Zhuo
On behalf of LHCb colaboration

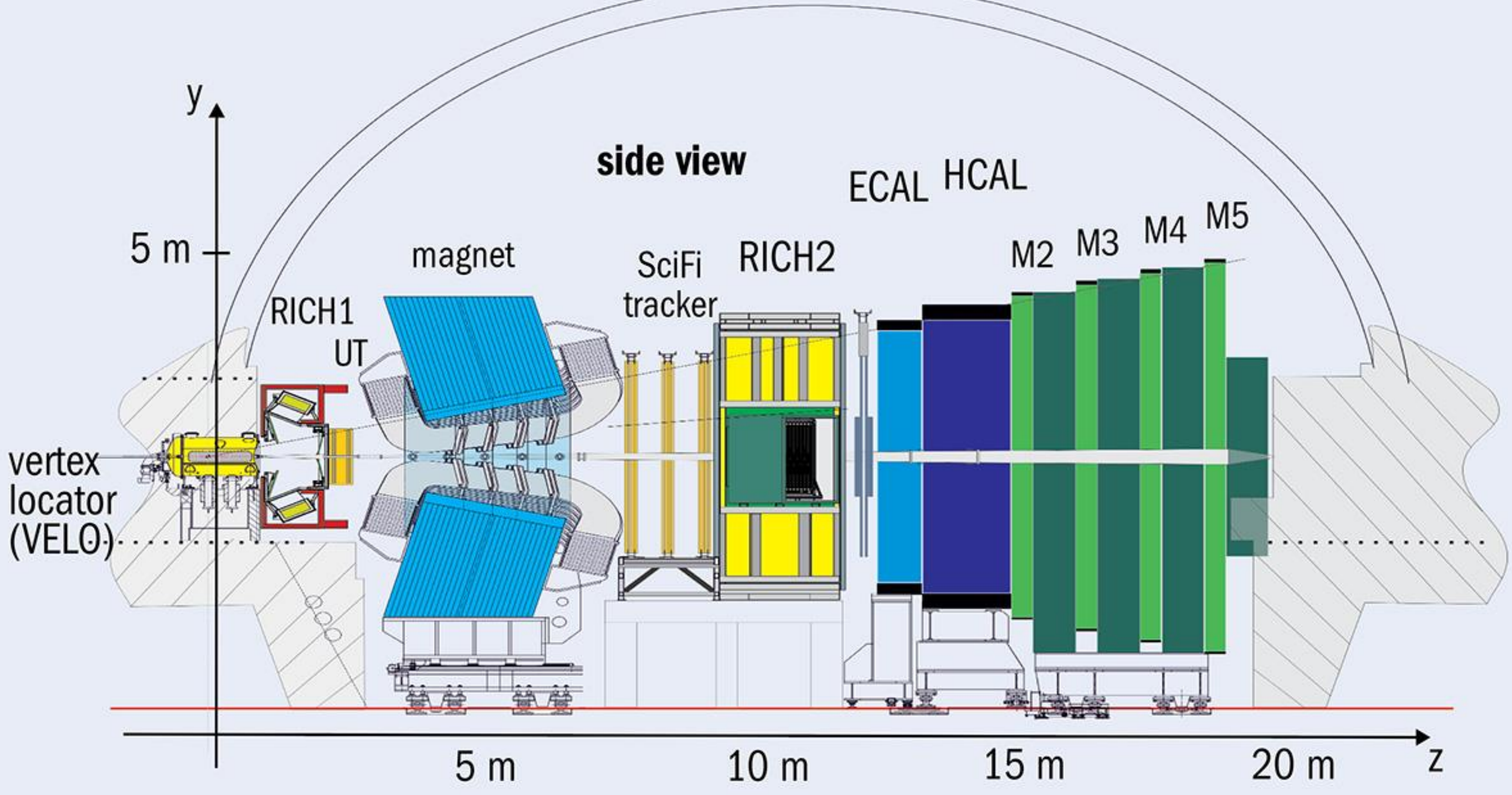
Introduction: LHCb detector



[LHCb detector link]

Introduction: LHCb detector

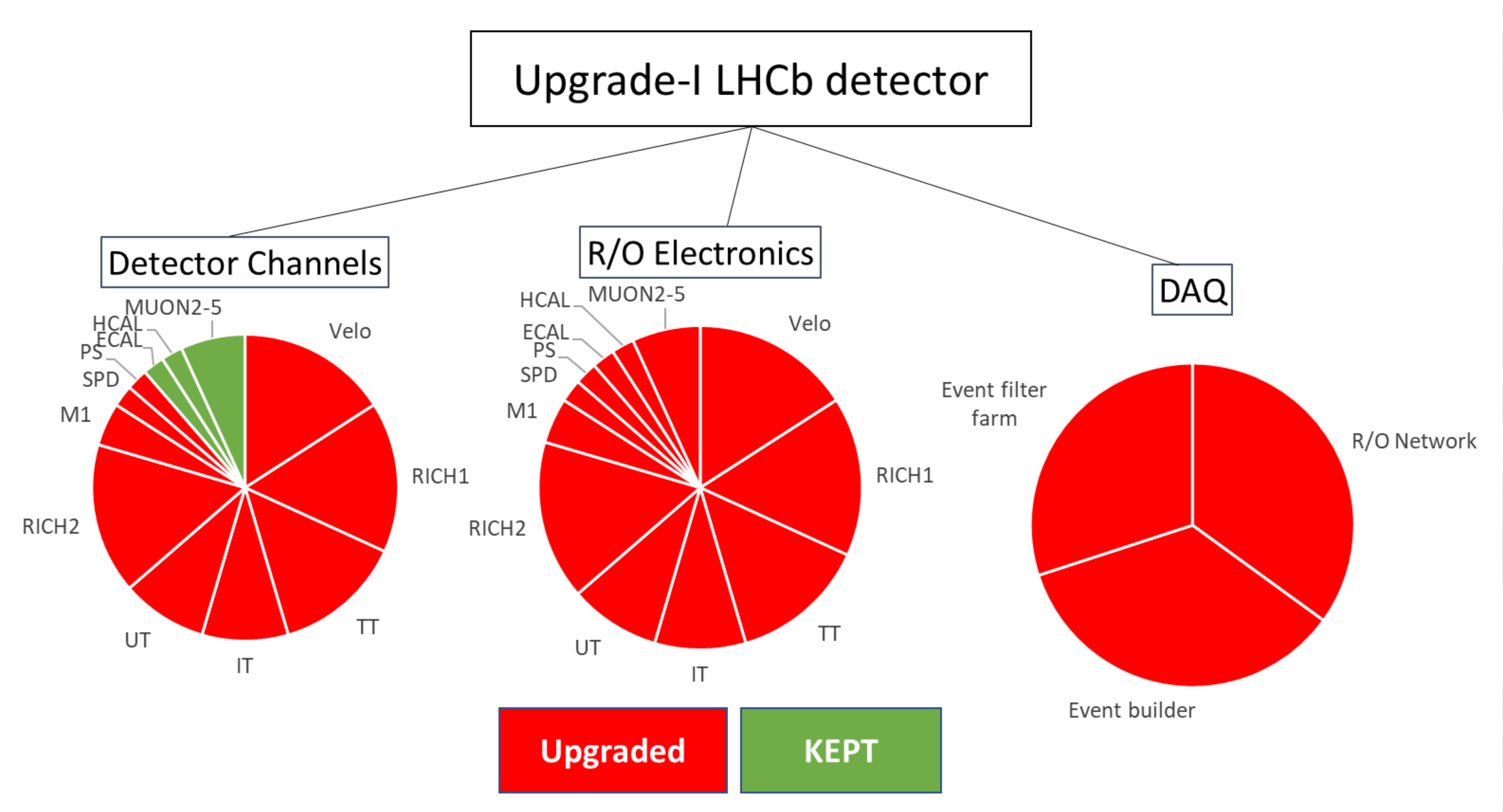
[CERN Courier: LHCb's momentous metamorphosis]



Upgraded LHCb detector for Run-3

- **5x** higher instantaneous luminosity 2x$10^{33} cm^{-2} s^{-1}$
- **10x** per unit time signal yield
- **6x** more pileup

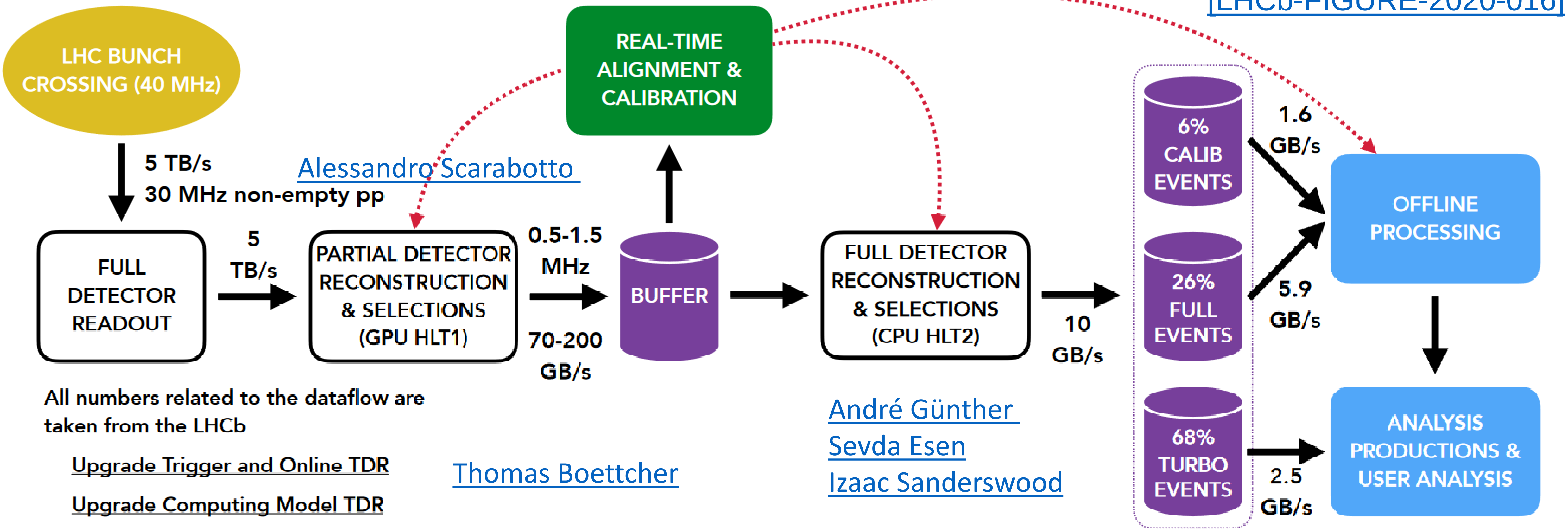
[LHCB-TDR-018]



Introduction: LHCb RTA data flow.

Real time alignment, talk by Florian Reiss

[LHCb-FIGURE-2020-016]



Alessandro Scarabotto

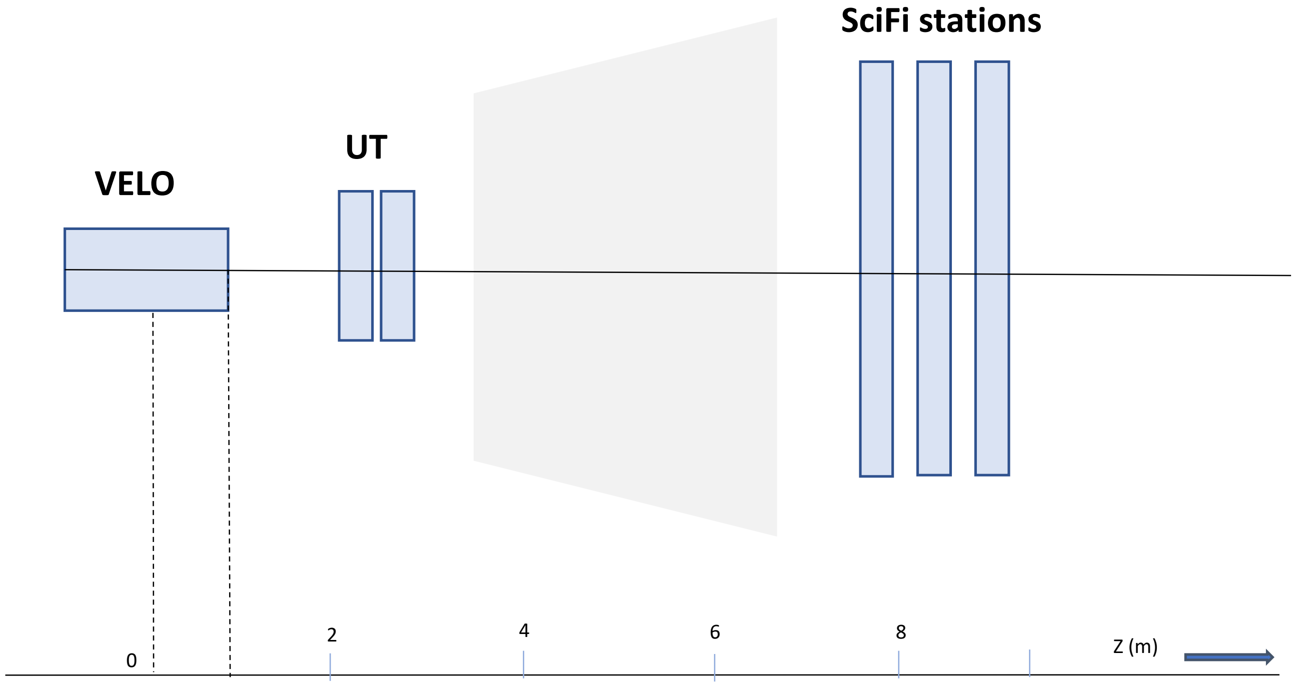
Thomas Boettcher

André Günther
Sevda Esen
Izaac Sanderswood

All numbers related to the dataflow are taken from the LHCb

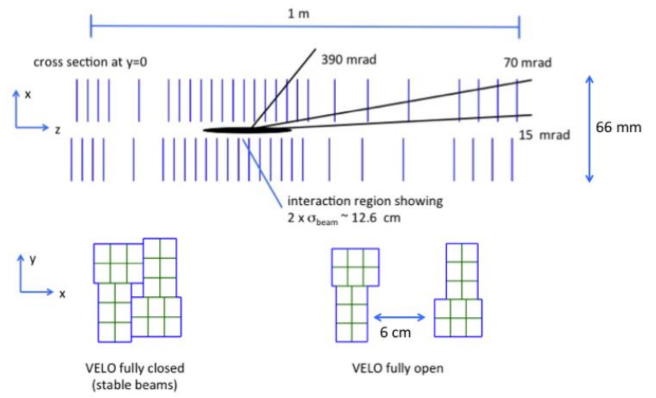
- Upgrade Trigger and Online TDR
- Upgrade Computing Model TDR

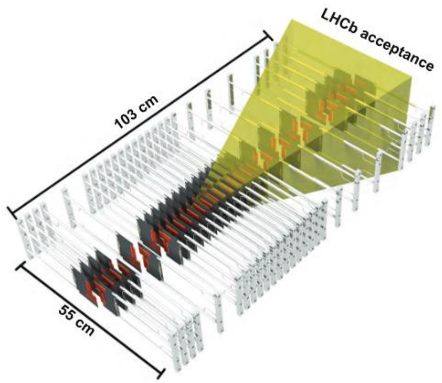
Introduction: Tracking system

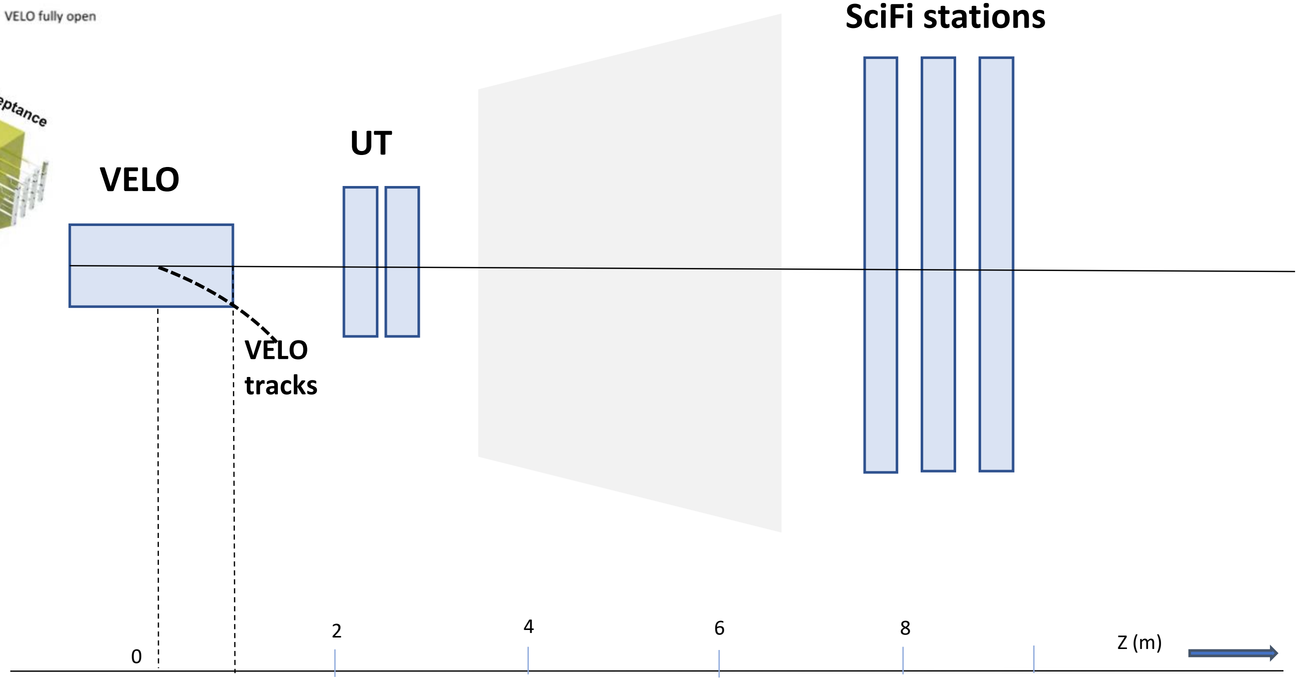


Introduction: Tracking system

VELO: Silicon pixel detector, 52 planes

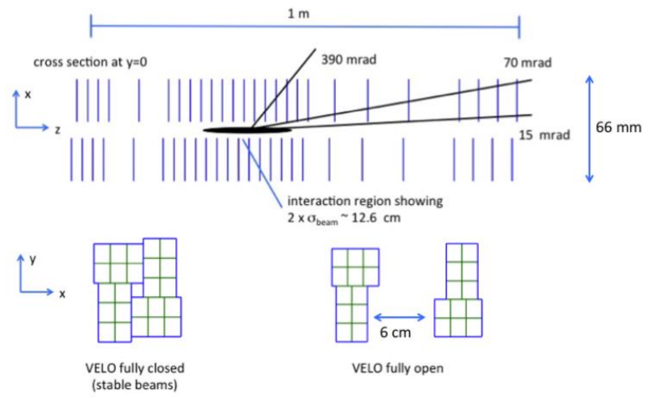




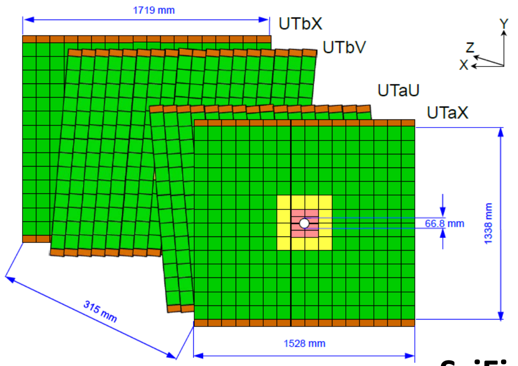


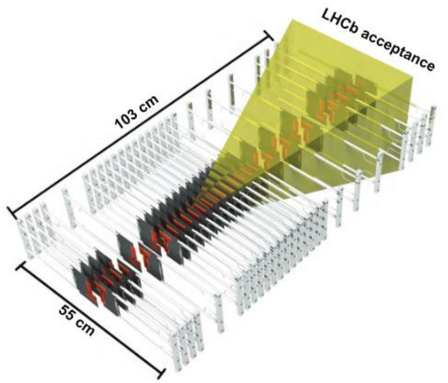
Introduction: Tracking system

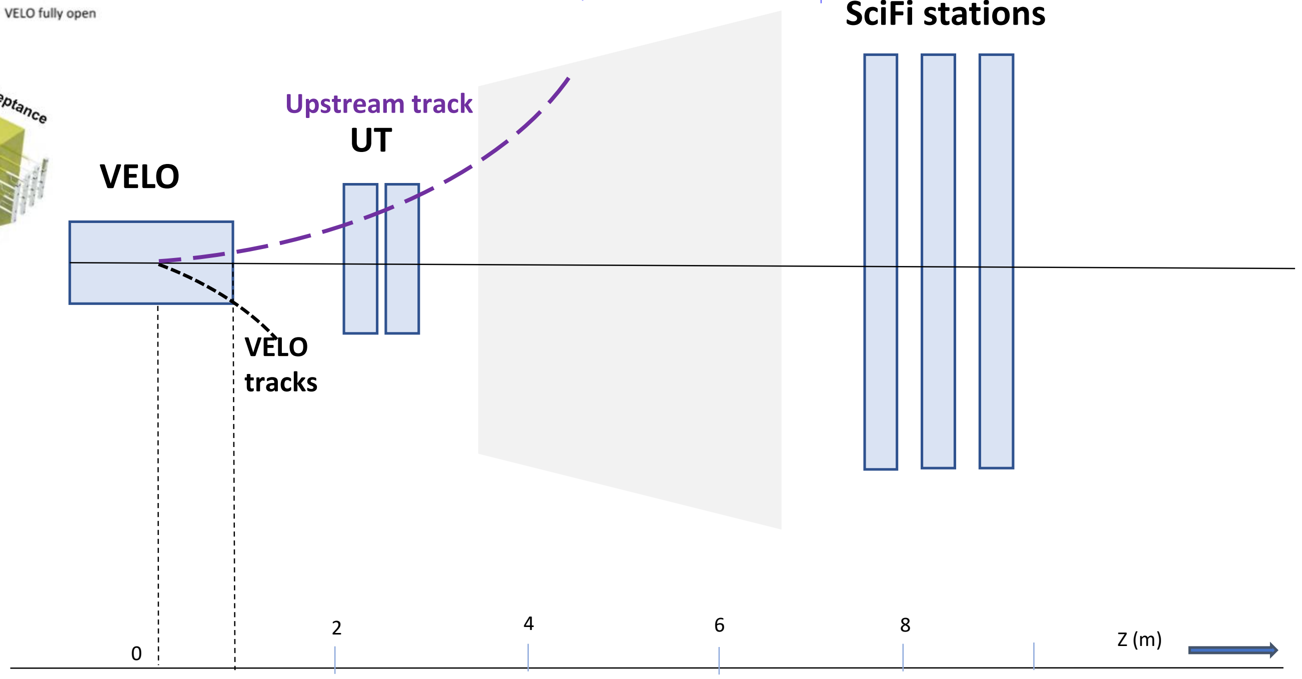
VELO: Silicon pixel detector, 52 planes



UT: Silicon strip detector, 4 planes

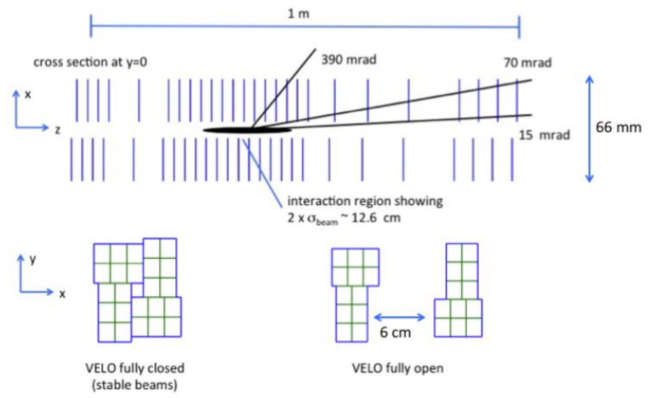


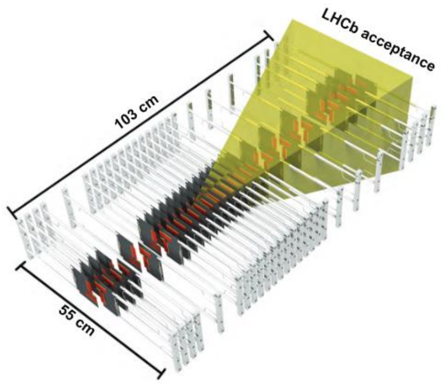




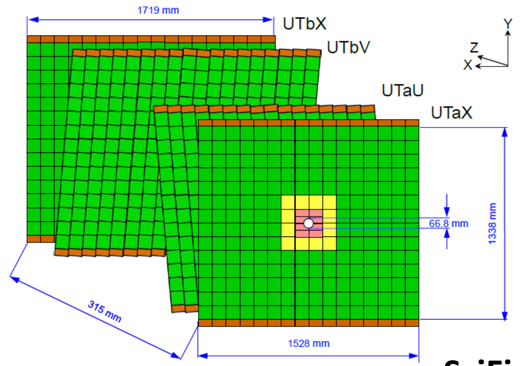
Introduction: Tracking system

VELO: Silicon pixel detector, 52 planes

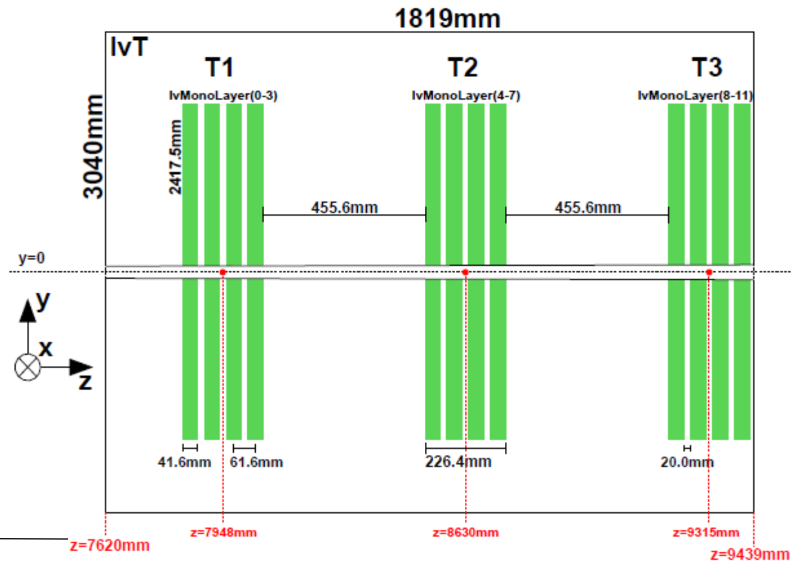


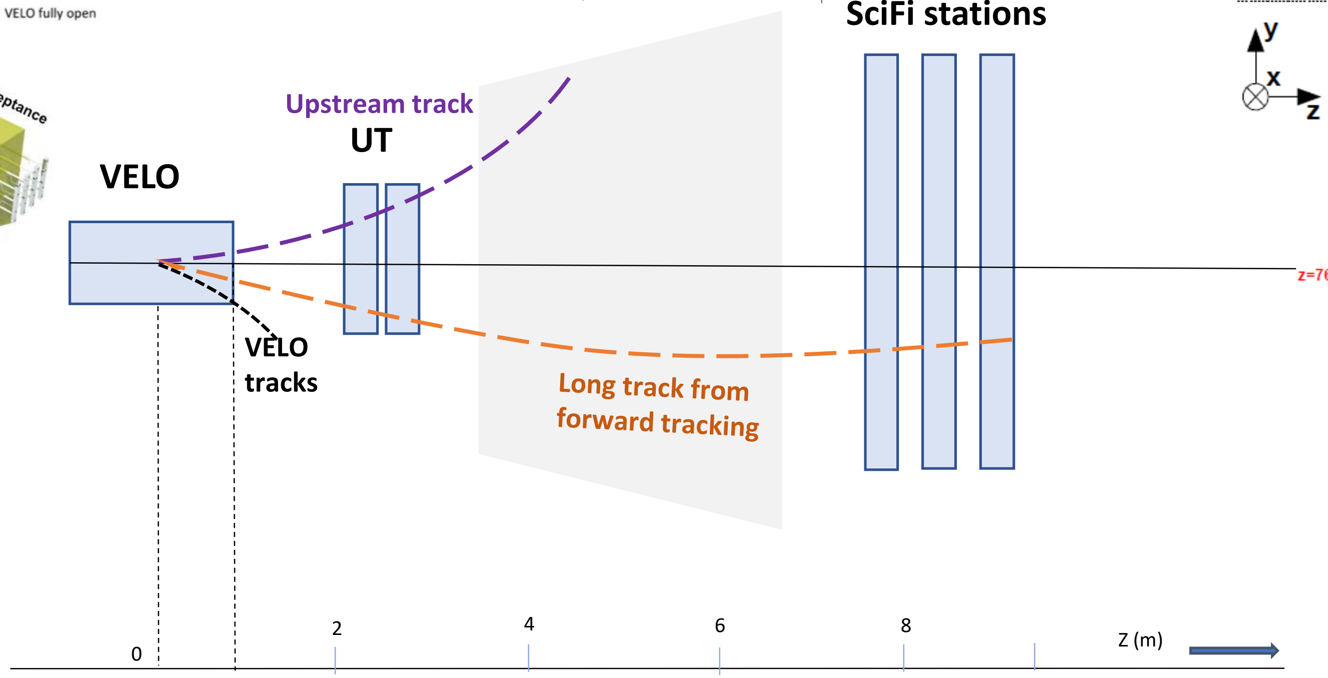


UT: Silicon strip detector, 4 planes



SciFi: Scintillating Fibre detector
(12 planes of 2x2.5 m long))

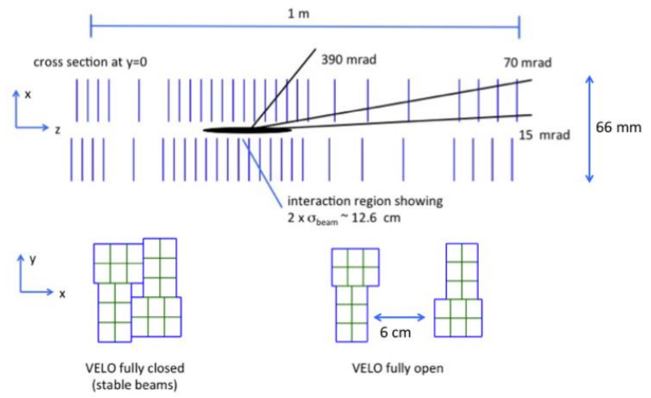




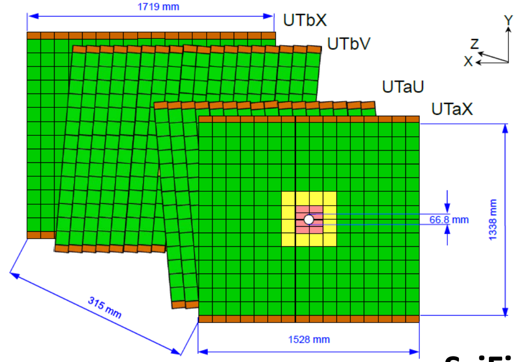
- 3 stations (T1, T2, T3) with 4 layers each in x-u-v-x configuration
- u and v-layers are tilted by a stereo angle of 5° in the vertical plane
- Two halves per layer with 5 modules (6 for T3) with 8 scintillating fibre mats

Introduction: Tracking system

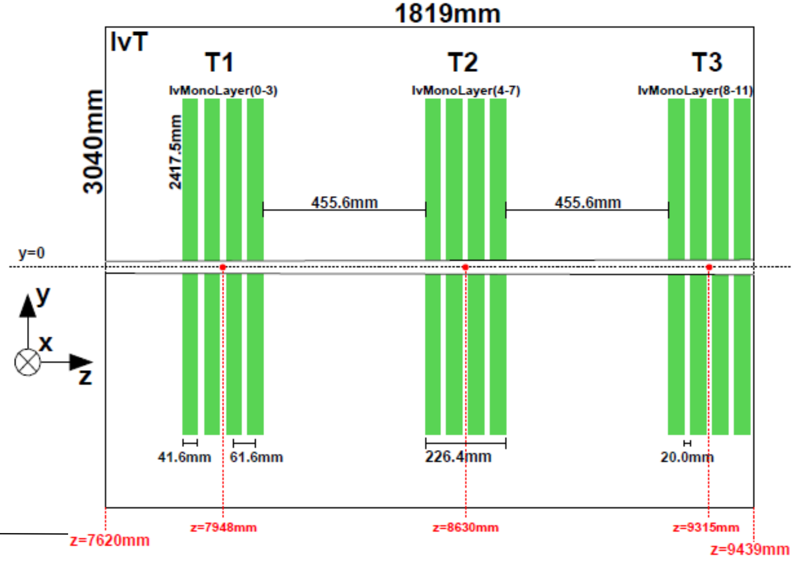
VELO: Silicon pixel detector, 52 planes

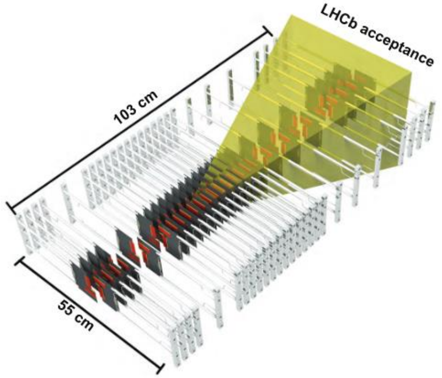


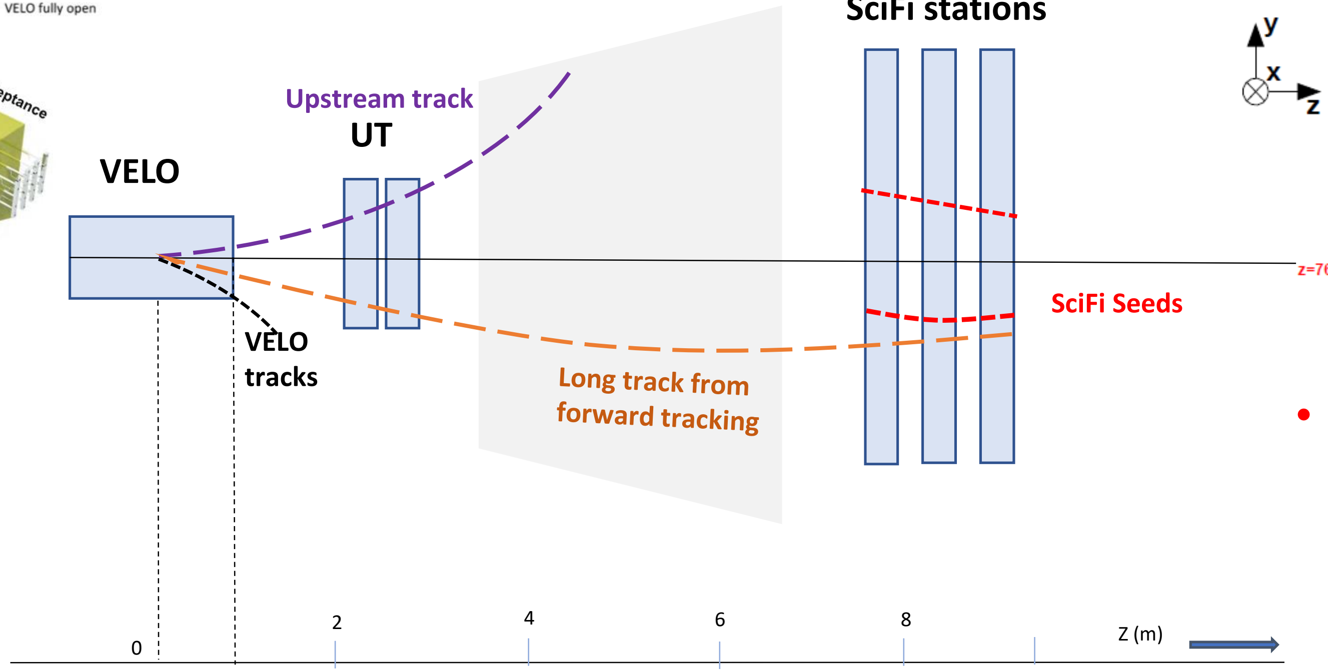
UT: Silicon strip detector, 4 planes



SciFi: Scintillating Fibre detector
(12 planes of 2x2.5 m long))



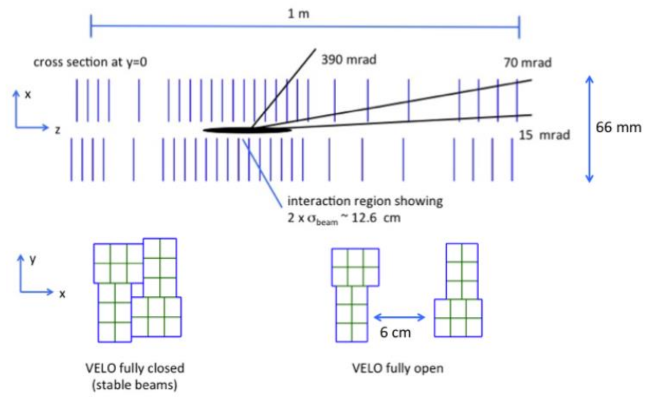




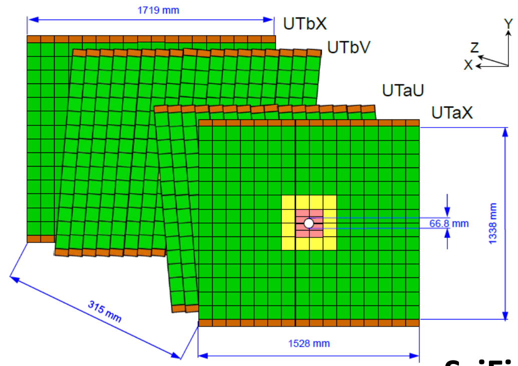
- Standalone SciFi Seeding at HLT1

Introduction: Tracking system

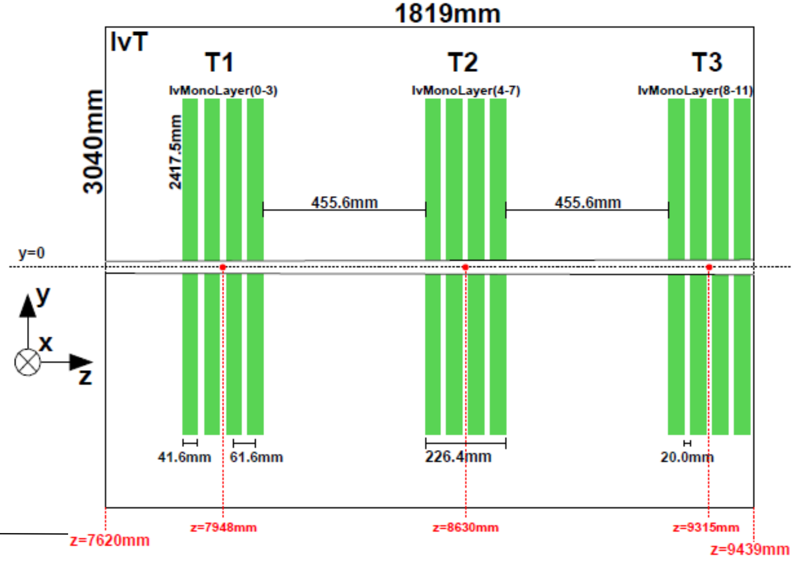
VELO: Silicon pixel detector, 52 planes

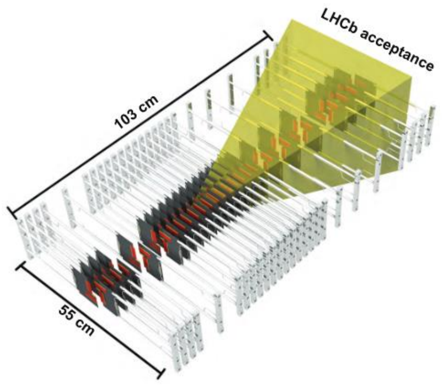


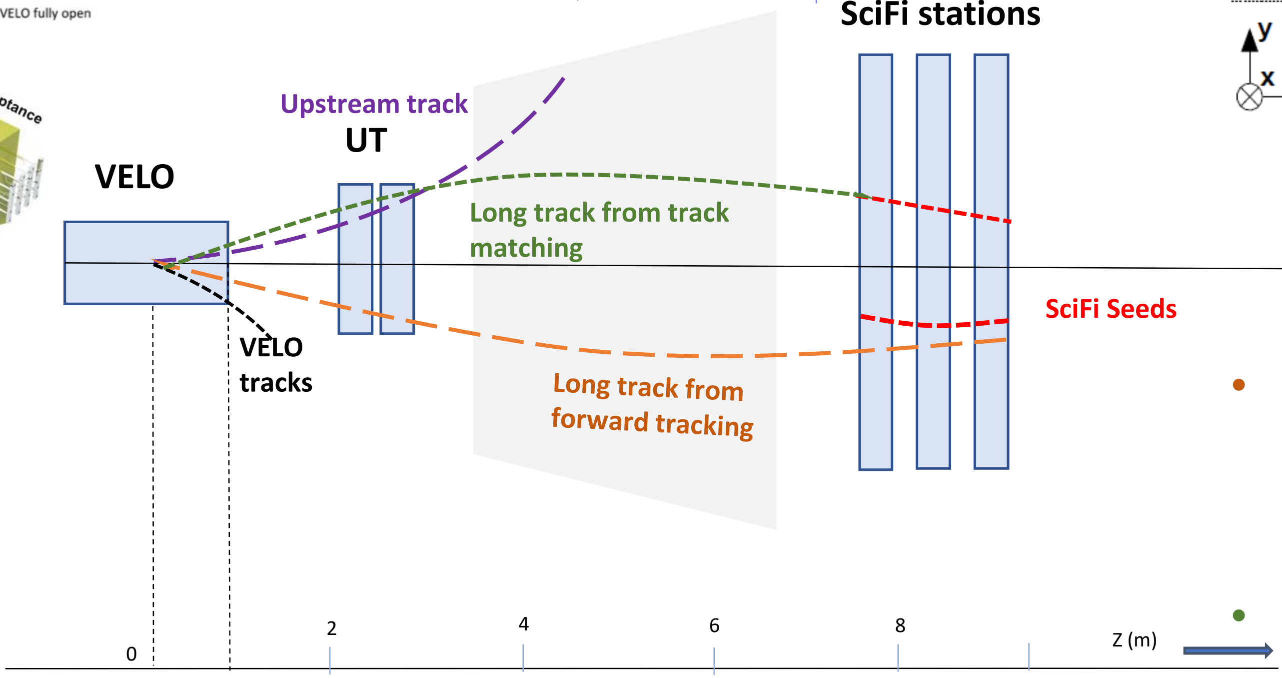
UT: Silicon strip detector, 4 planes



SciFi: Scintillating Fibre detector
(12 planes of 2x2.5 m long))



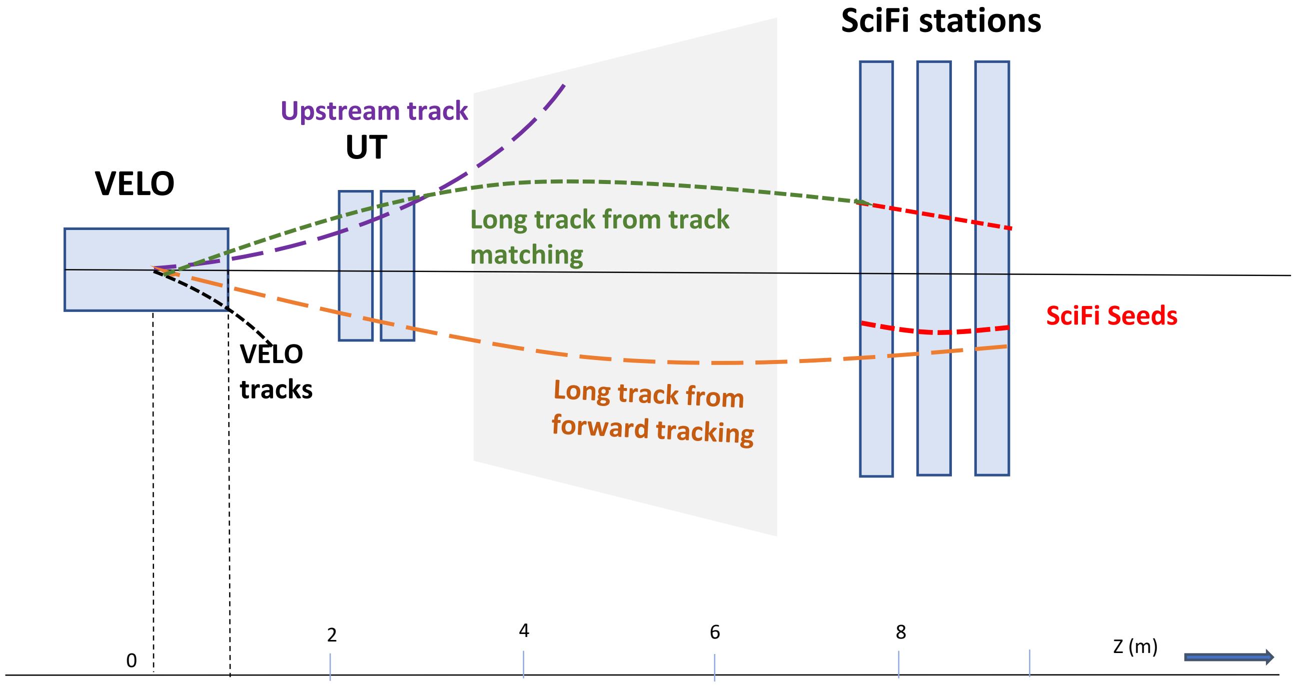




- Forward tracking: Velo tracks are extended to UT and then to SciFi to create long tracks (Baseline HLT1)
- VeloSciFi Track matching: SciFi seeds are matched with Velo tracks to create long tracks. (UT hits can also be added)

This talk is about two new algorithms implemented at HLT1 level on GPUs

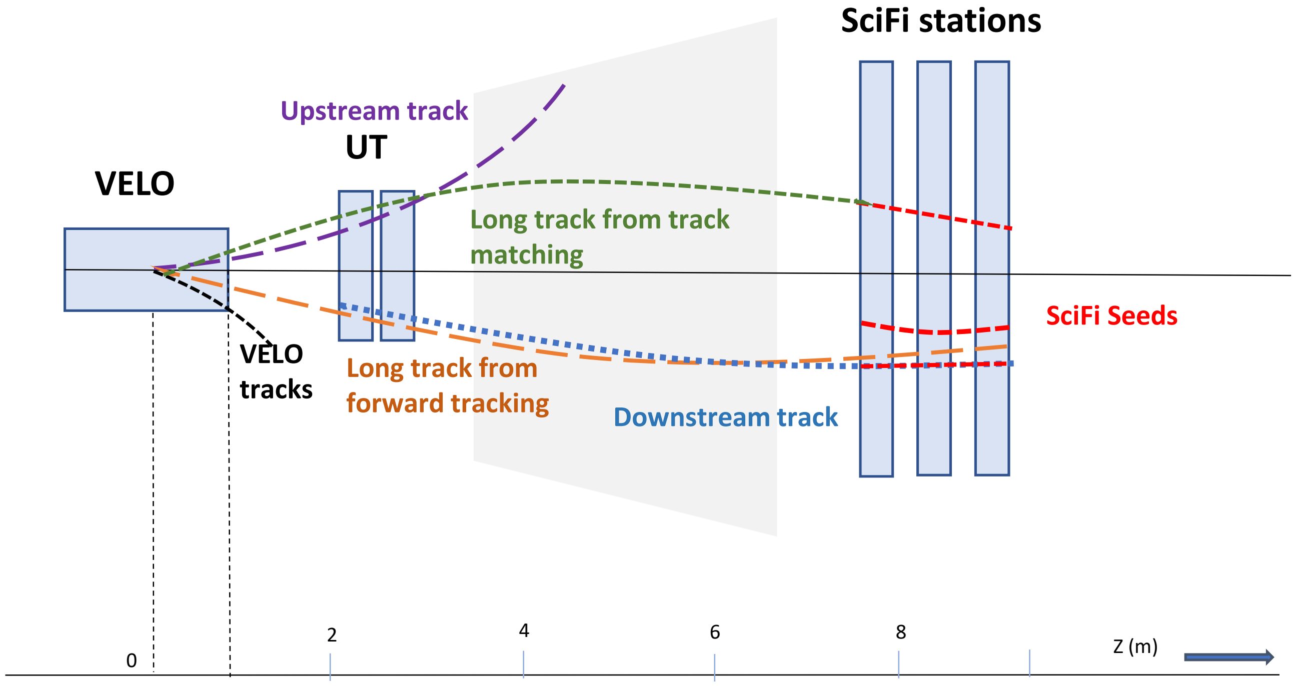
- Standalone SciFi Seeding
- VeloSciFi Track matching (Long tracks) (alternate to current forward tracking)



SciFi seeding

SciFi Seeds:

- +Velo tracks → produce Long tracks (Alternative to current forward tracking)
- +UT hits → produce Downstream tracks
- Help in ECAL PID (matching a track to a cluster helps distinguish electrons from photons).
- Extending physics reach of LHCb – See talk by Izaac Sanderswood



- Standalone SciFi seeding algorithm in HLT1
  - ✓ 12 layer with ~450 hits each (average)
  - ✓ Residual magnetic field and 3 stations spread over 1.8 meters
  - ✓ Fibers 2.7 meters long

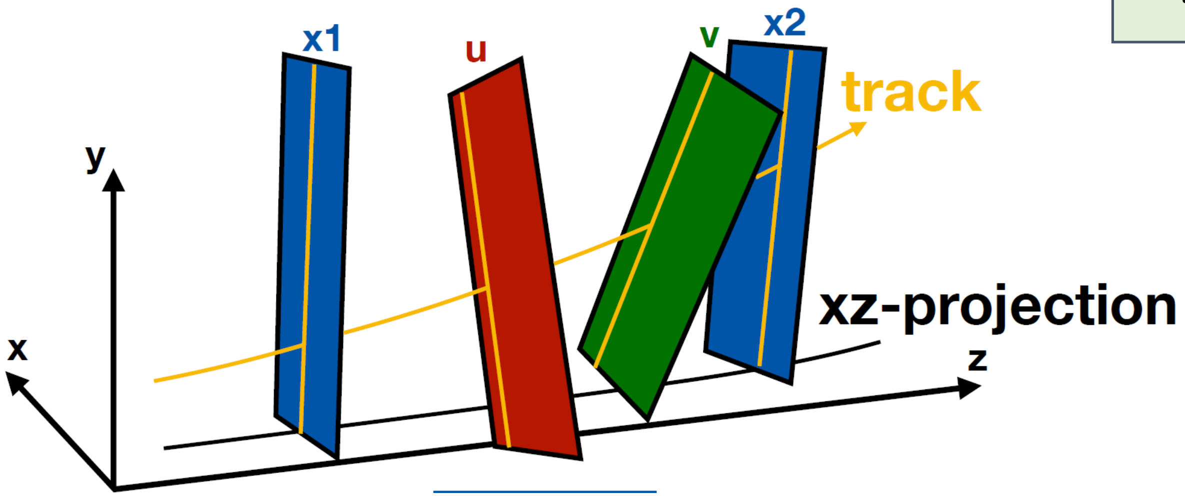
- Two cases optimized for varying initial layers to account for hit inefficiencies

- Each iteration consists of two main components of algorithm

  Seeding_XZ

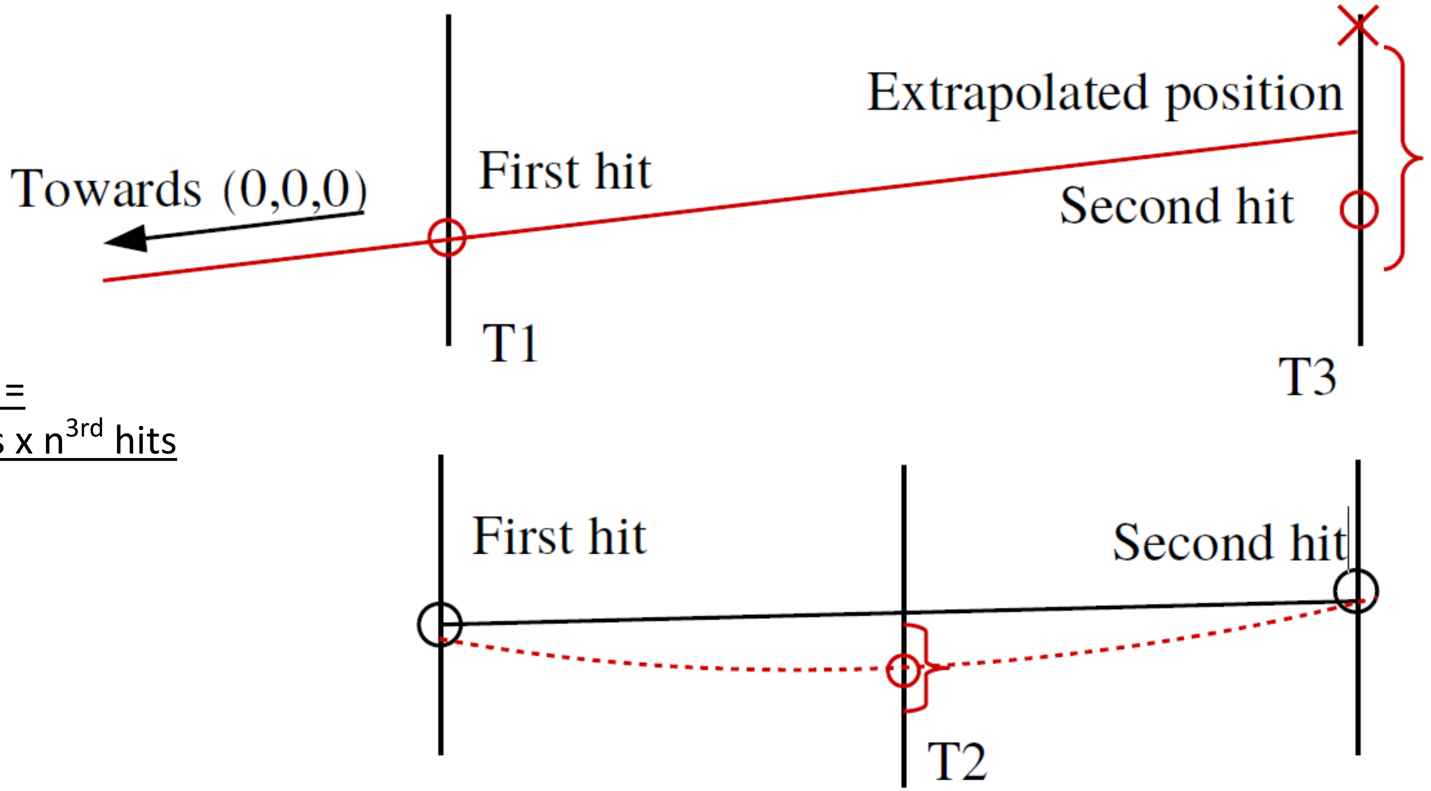
  Seeding_confirmTracks

SciFi hits
⇩
SciFi Seed_XZ
⇩
Add UV hits to confirmTracks
⇩
SciFi seeds



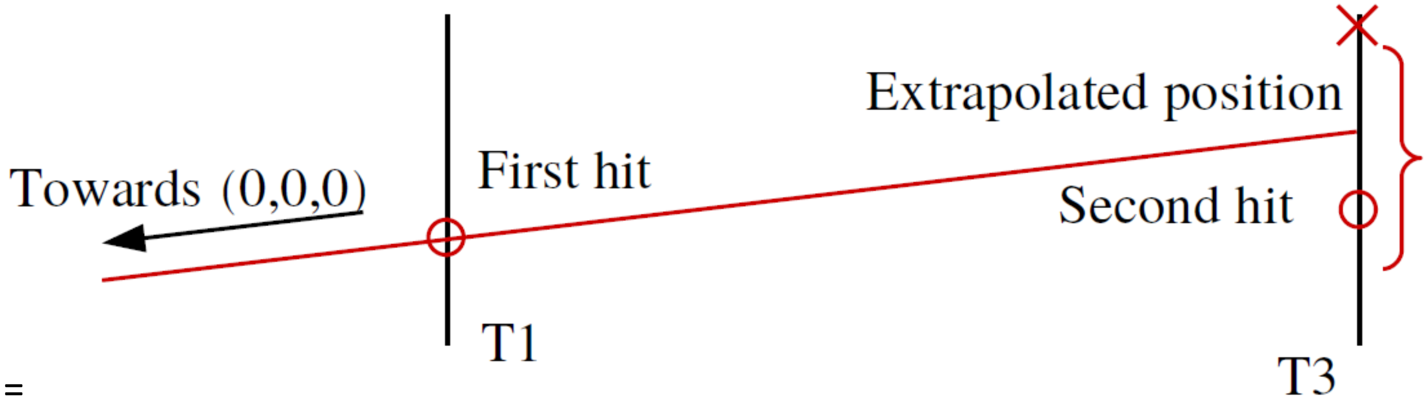
SciFi seeding: *Seed_XZ*

SciFi Hits ⇨ Doublets searched in T1 & T3, (P > 3Gev/c & origin at 0,0,0) ⇨ for each doublet charge-momentum estimation, narrower window for 3rd hit ⇨ 3 hit combinations with hits from T2 station, taking bending into account

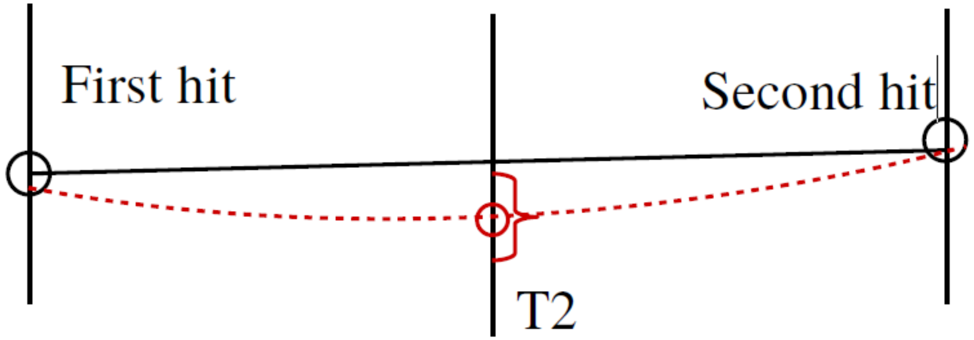


XZ combinatorics =
n^{1st} hits x n^{2nd} hits x n^{3rd} hits

SciFi Hits → Doublets searched in T1 & T3, (P > 3Gev/c & origin at 0,0,0) → for each doublet charge-momentum estimation, narrower window for 3rd hit → 3 hit combinations with hits from T2 station, taking bending into account → 3 hit + Parabola with cubic correction estimation on MC to predict positions in the remaining layers → Add remaining hits, accept candidates with at least two additional hits from remaining layers i.e min 5 out of 6 possible hits → Build 3x3 matrix for track fit with all contributing hits, reject track candidates with $\chi^2/ndof$ worse than a fixed value (<6) → Seed_XZ candidates



<u>XZ combinatorics =</u>
<u>n^{1st} hits x n^{2nd} hits x n^{3rd} hits</u>



Seed_XZ candidates

- Still need to remove ghosts (still around 50-60% of XZ segments)
- Need to measure parameters of the track

Seeding_confirmTracks: Add U/V hits (6 hits)

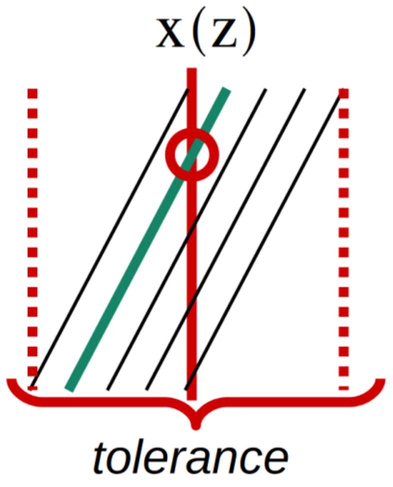
1. For each XZ segment, open a very large tolerance window in y for initial UV-layer .
    - Many different fibres (black) cross a given x(z) hypothesis (red).
    - Each fibre crosses that hypothesis at a different y, giving a y measurement.

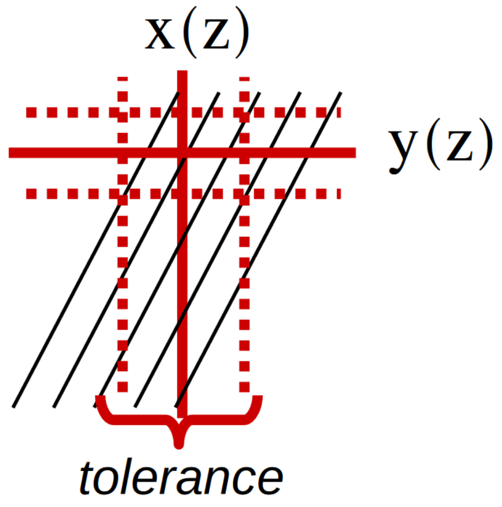
2. For all hits collected in that window, determine the corresponding y and $t_y \rightarrow$ *y(z).*

3. For all other layers, open a much narrower tolerance window around that hypothesis update t_y at each hit found.

4. Accept the tracks with at least 4 hits in UV

5. Fit the candidates with a linear model in Y and keep the ones with best χ^2

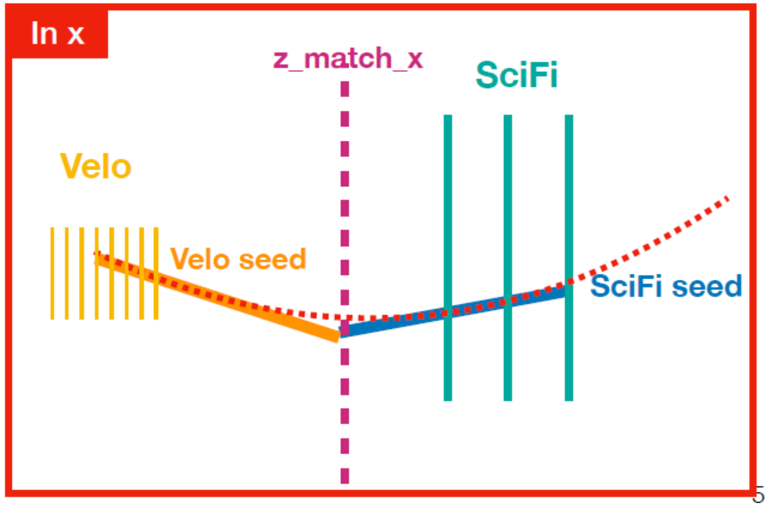


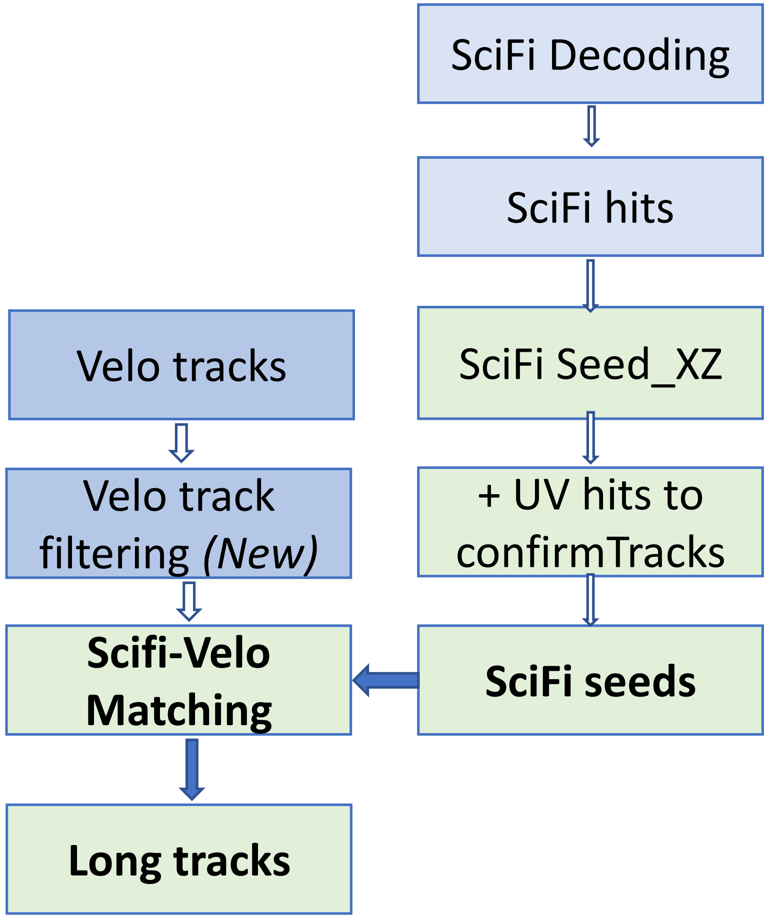


Track Matching (SciFi + Velo)

Algorithm approach:

- SciFi seeds O(80) matched to Velo tracks O(100) without UT hits
- Start with SciFi seeds and parallelize over SciFi loops
- Velo/SciFi seeds extrapolated to magnet as lines ("Kink" approximation)
- Magnetic field and bending in *y* is parametrized with simulation
- Minimal requirements on χ^2 and slope to reduce combinatorics and fakes
- Clone killing: tracks that share a VELO track are compared and only the one with the best χ^2 is kept

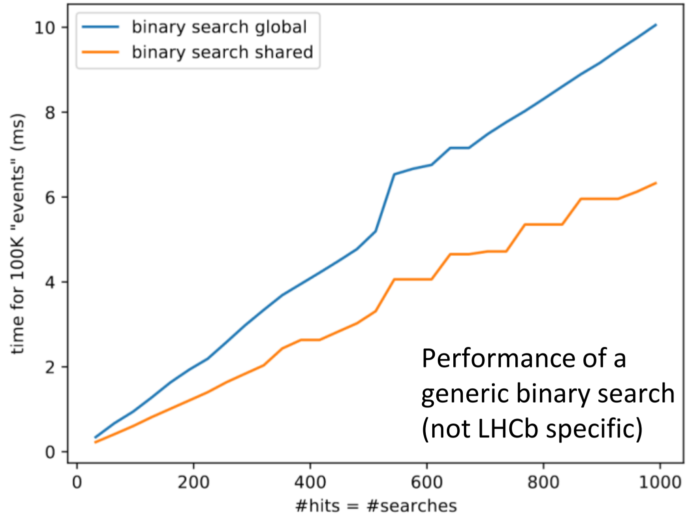


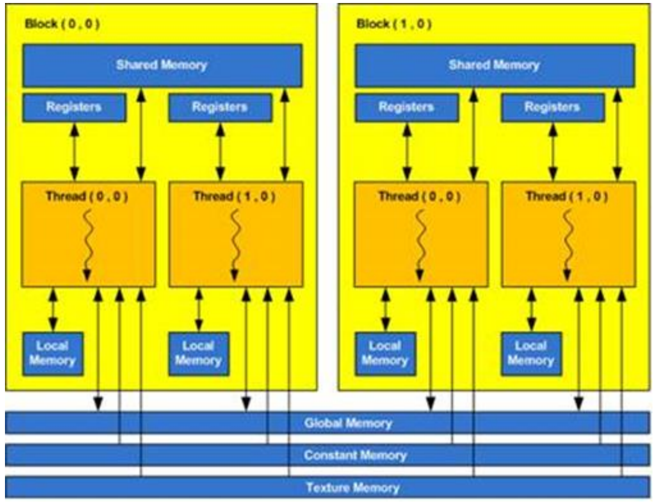


GPU implementation

GPU implementation features for both SciFi seeding and Track Matching

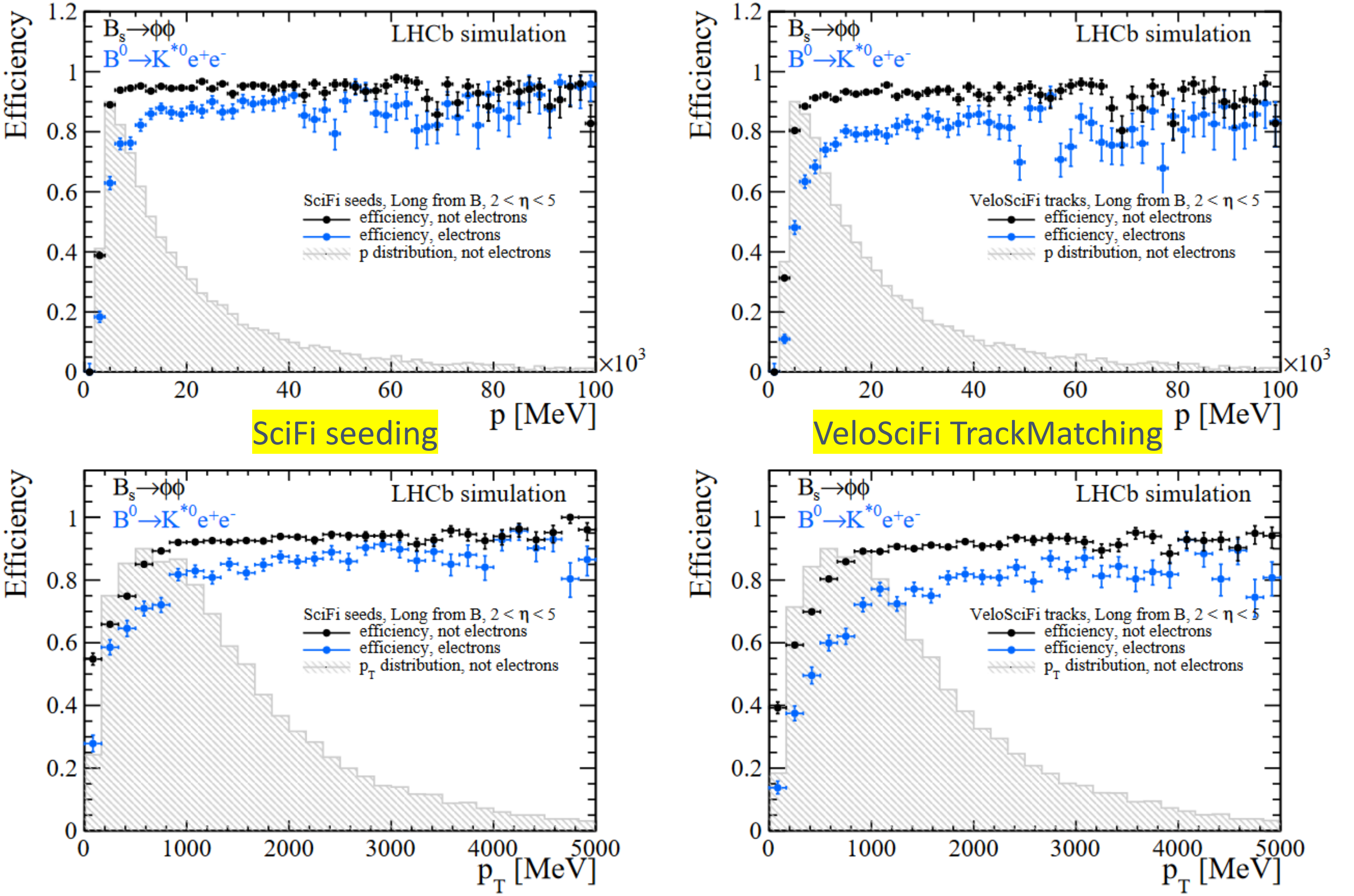
- ✓ Event level parallelization from framework
- ✓ Minimized global memory usage
  - ✓ SciFi hits caching preferred in shared memory of GPU which is faster as compared to global memory
- ✓ For each XZ track in parallel, collect UV hits in 2 different layers
- ✓ Binary search implementation for making pair candidates from hits
- ✓ Optimal block size on A5000 (64 KB shared + 32 KB L1 configuration)
  - ➢ 8kb of shared memory per block => 4 warps per block (128 threads)
- ✓ Simplified approach in the confirm_track step to add UV information as compared to HLT2
- ✓ For clone killing using a shared memory voting algorithm,
  - ✓ For each track candidate in parallel, compute a score based on χ^2 and make decision for rejection of clones





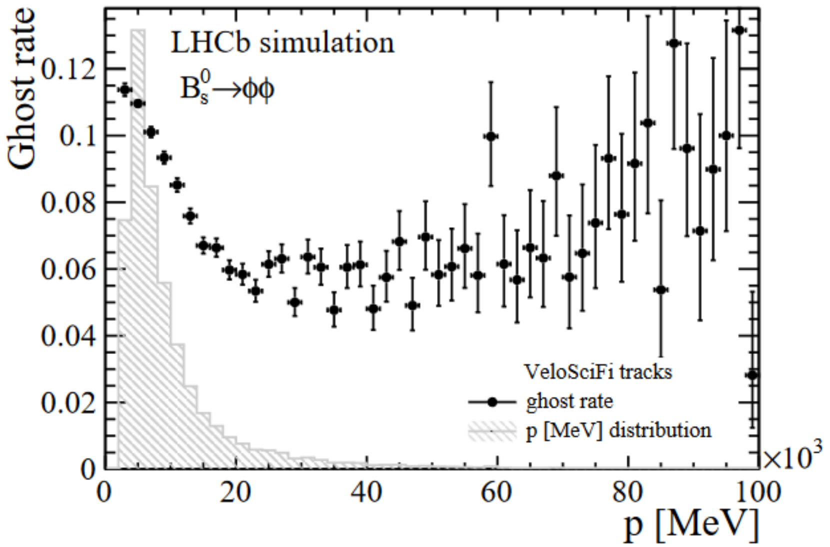
Efficiency

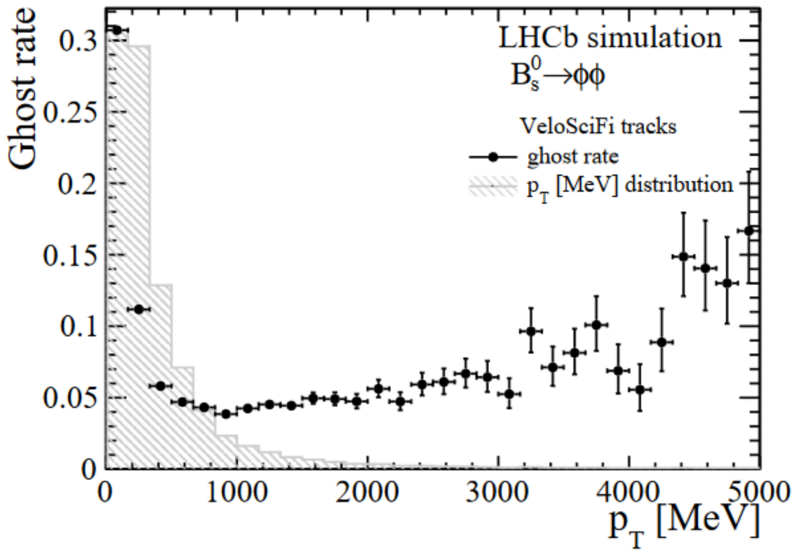
- long tracks from B decays (p > 5GeV): ~83% (92%) –
- long electrons from B decays (p > 5GeV): ~ 70% (75%) –
- Increased efficiency w.r.t the baseline approach without cuts on p_T.



SciFi seeding

VeloSciFi TrackMatching

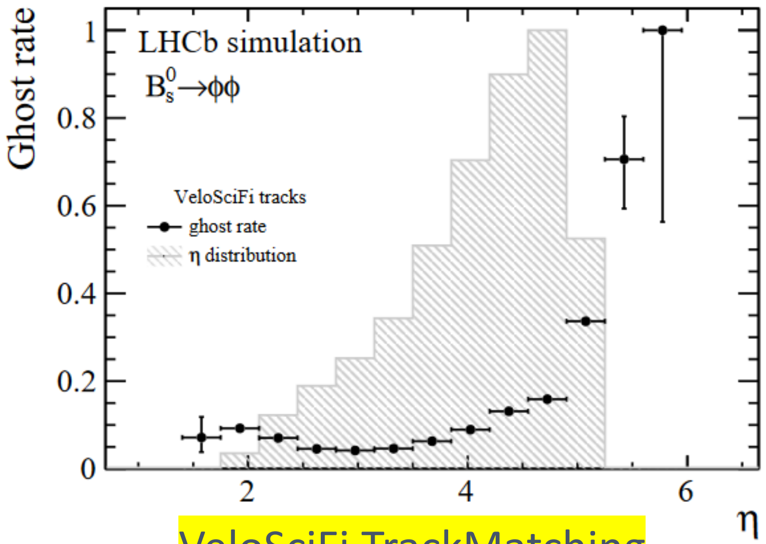




Ghost rate

~9% for all reconstructed tracks

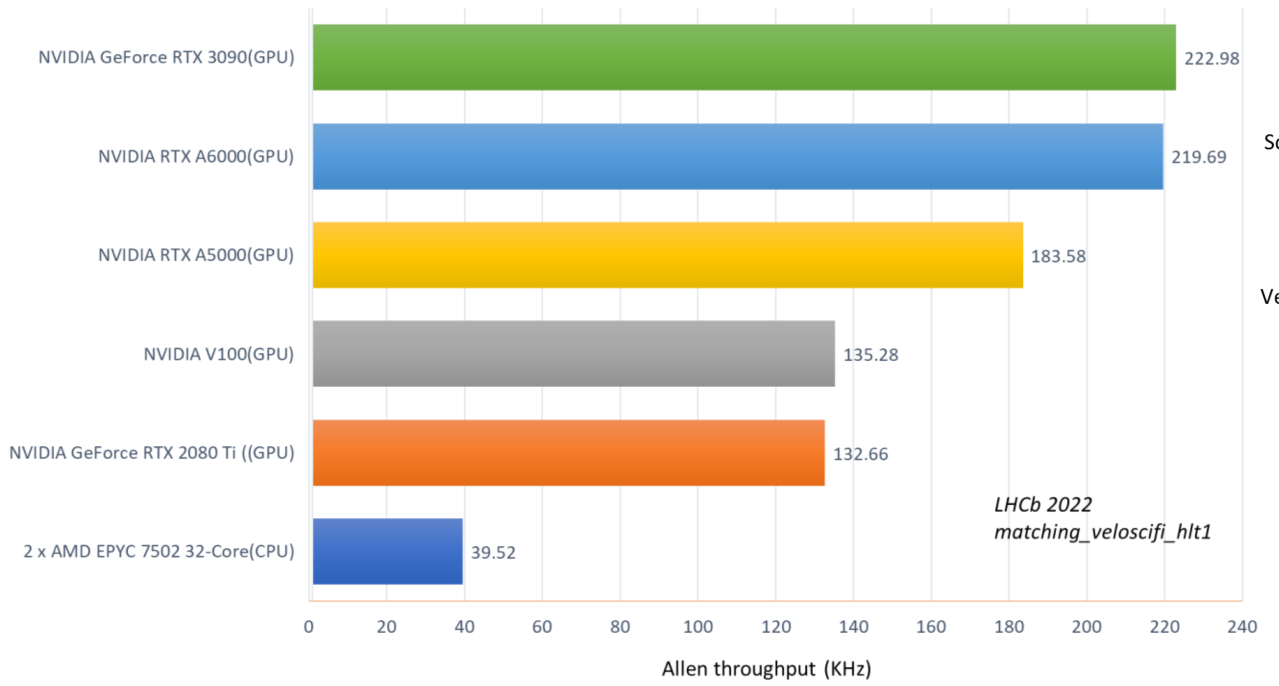
~5% for p > 3GeV, P_T > 0.5GeV

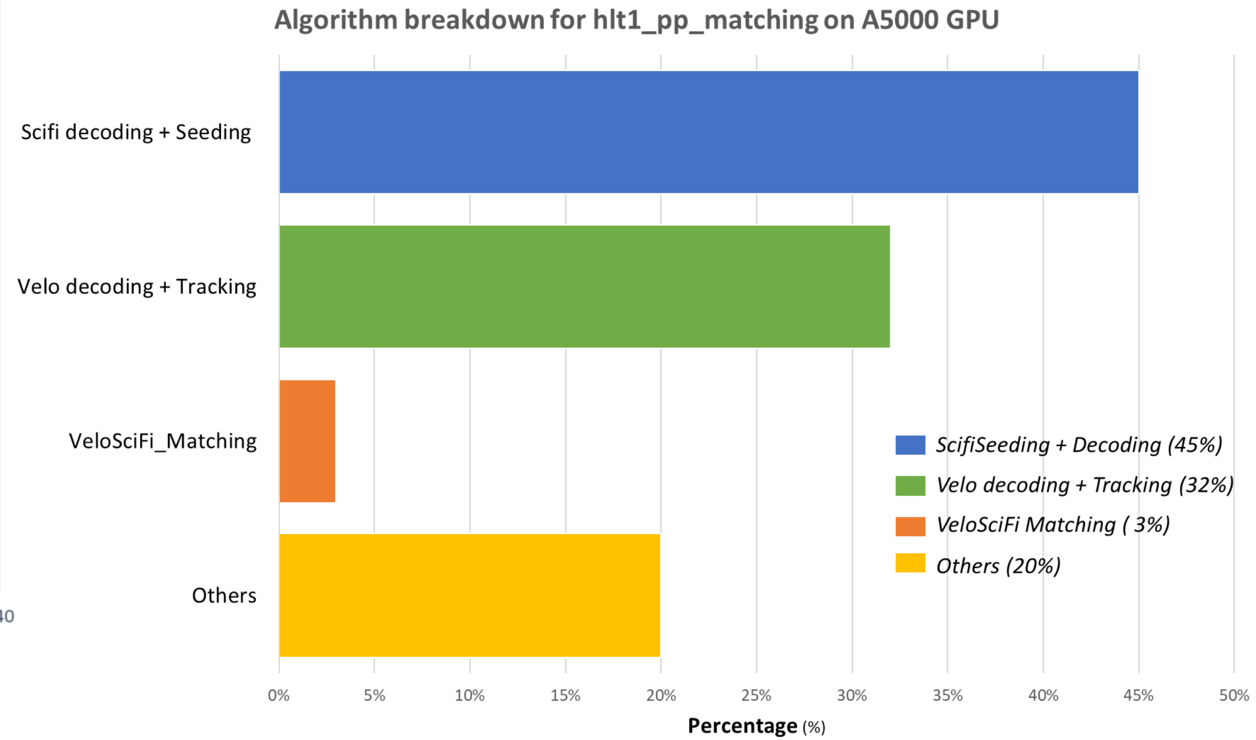


VeloSciFi TrackMatching

Performance (Throughput)

[LHCb-FIGURE-2022-010]

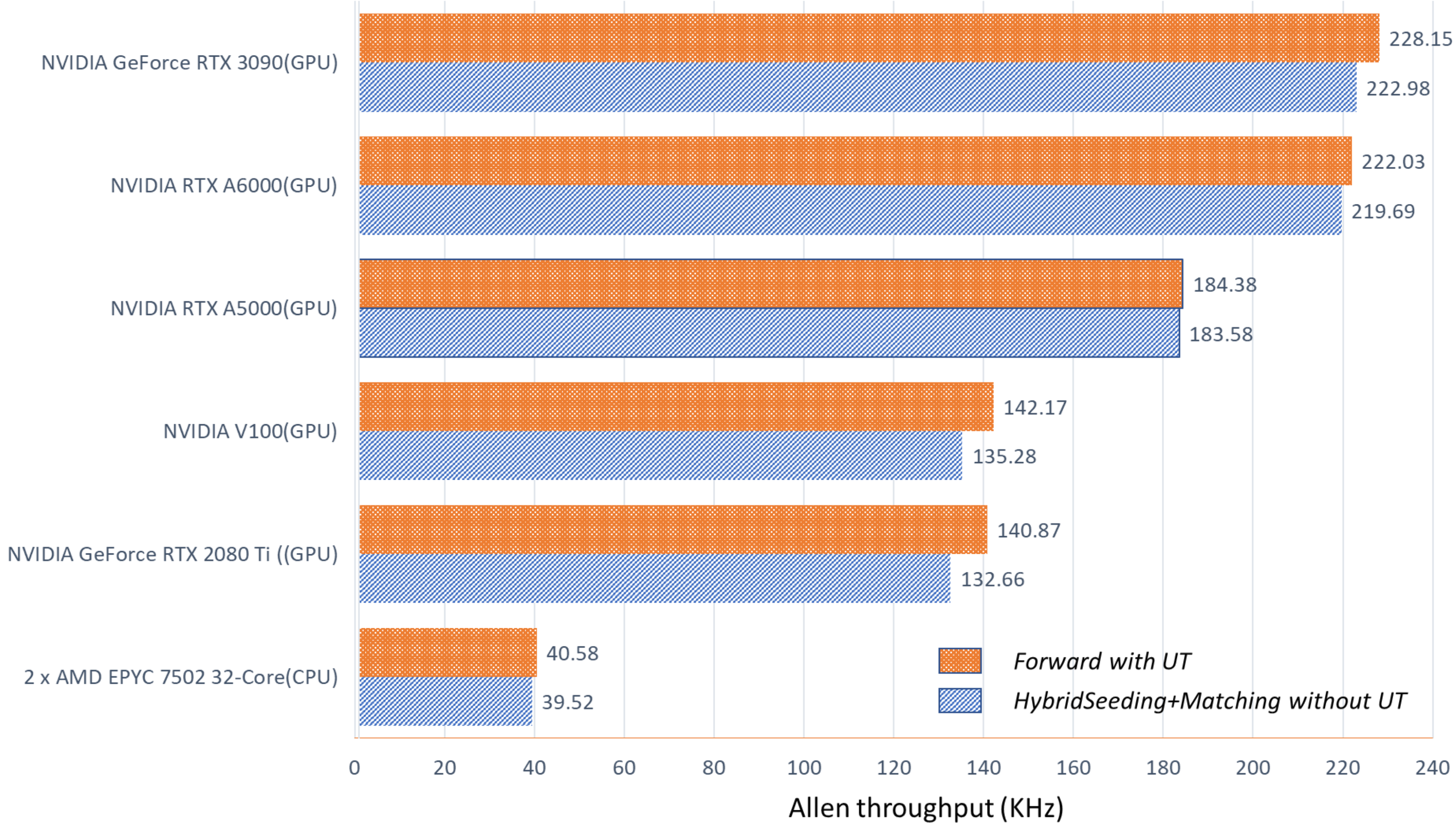




Performance (Throughput)

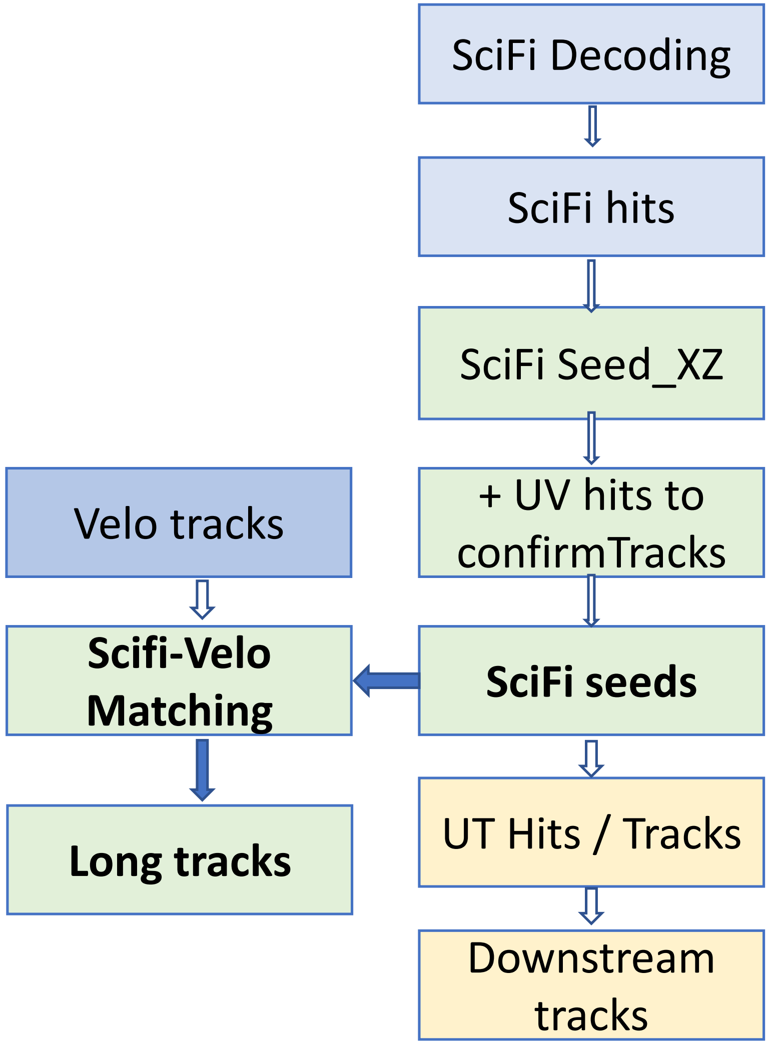
[LHCb-FIGURE-2022-010]

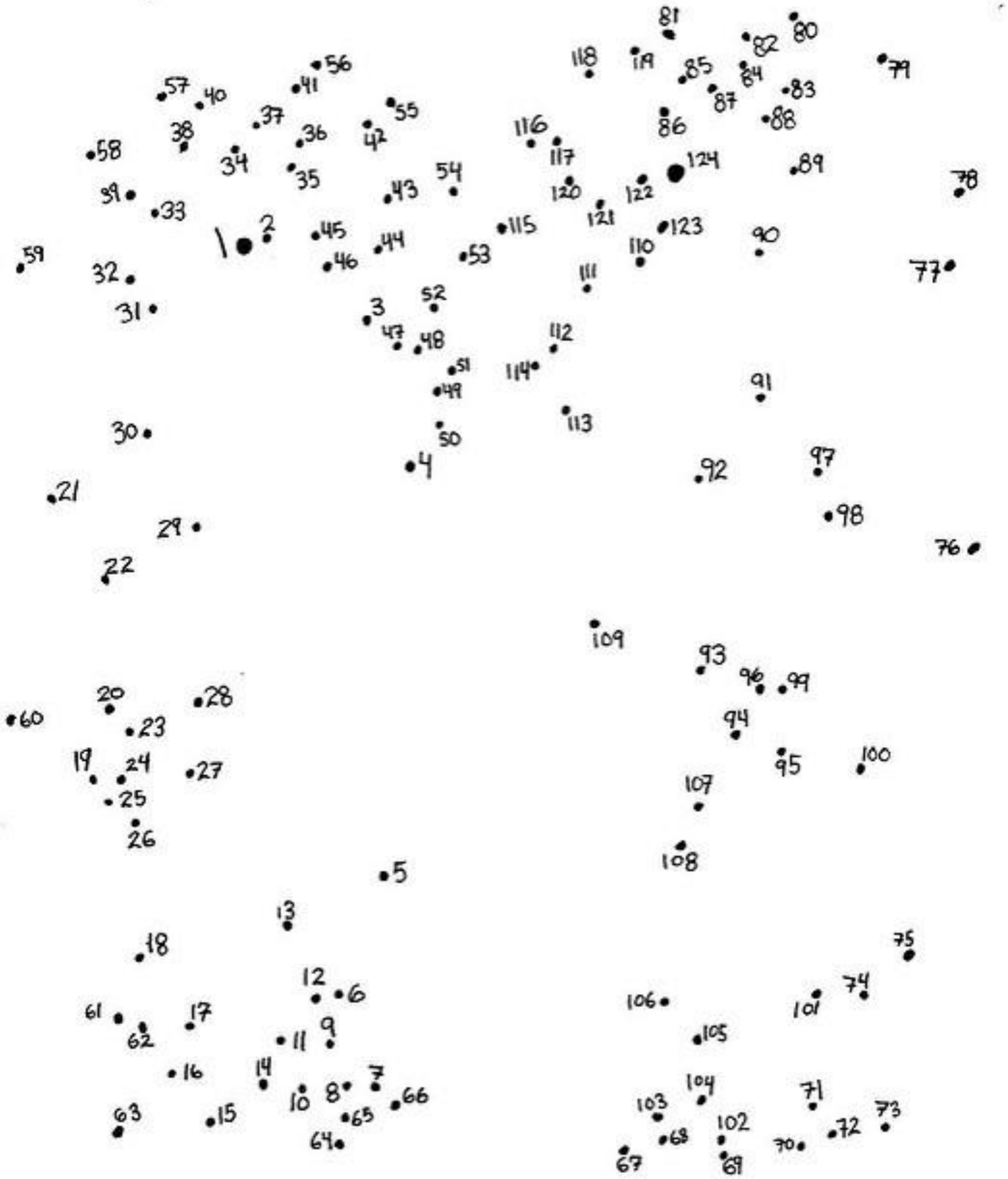
For details of HLT1 see talk by Alessandro Scarabotto



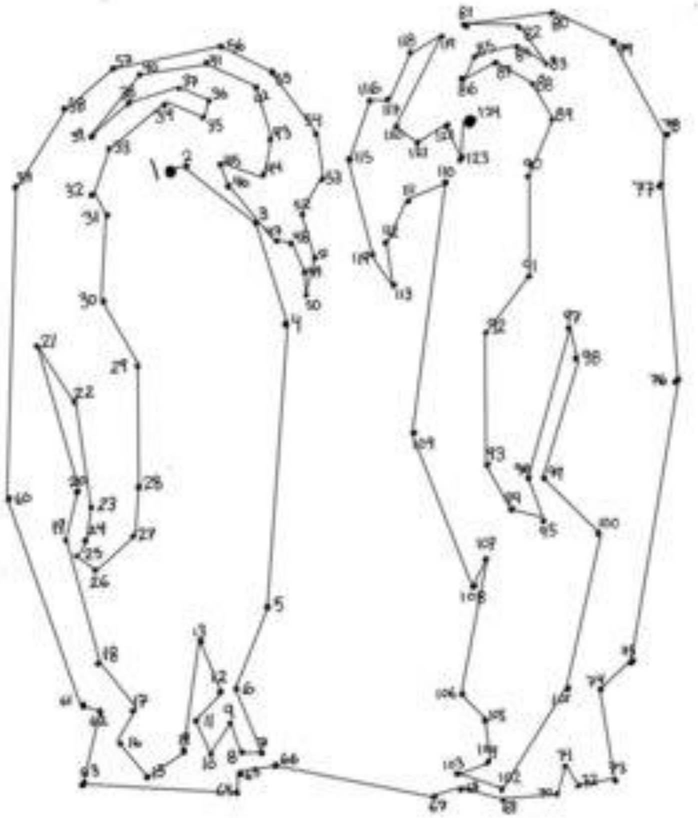
Summary

- Presented a new GPU-based Scintillating Fibre seeding algorithm for HLT1 at LHCb.
- An alternative long track reconstruction based on SciFi seeds + Velo track matching.
- Comparable throughput with Forward-with-UT, similar efficiencies at high momentum, without any hard cut at low p_T.
- Now we have SciFi seeds in HLT1 which are basic building blocks for downstream tracking.

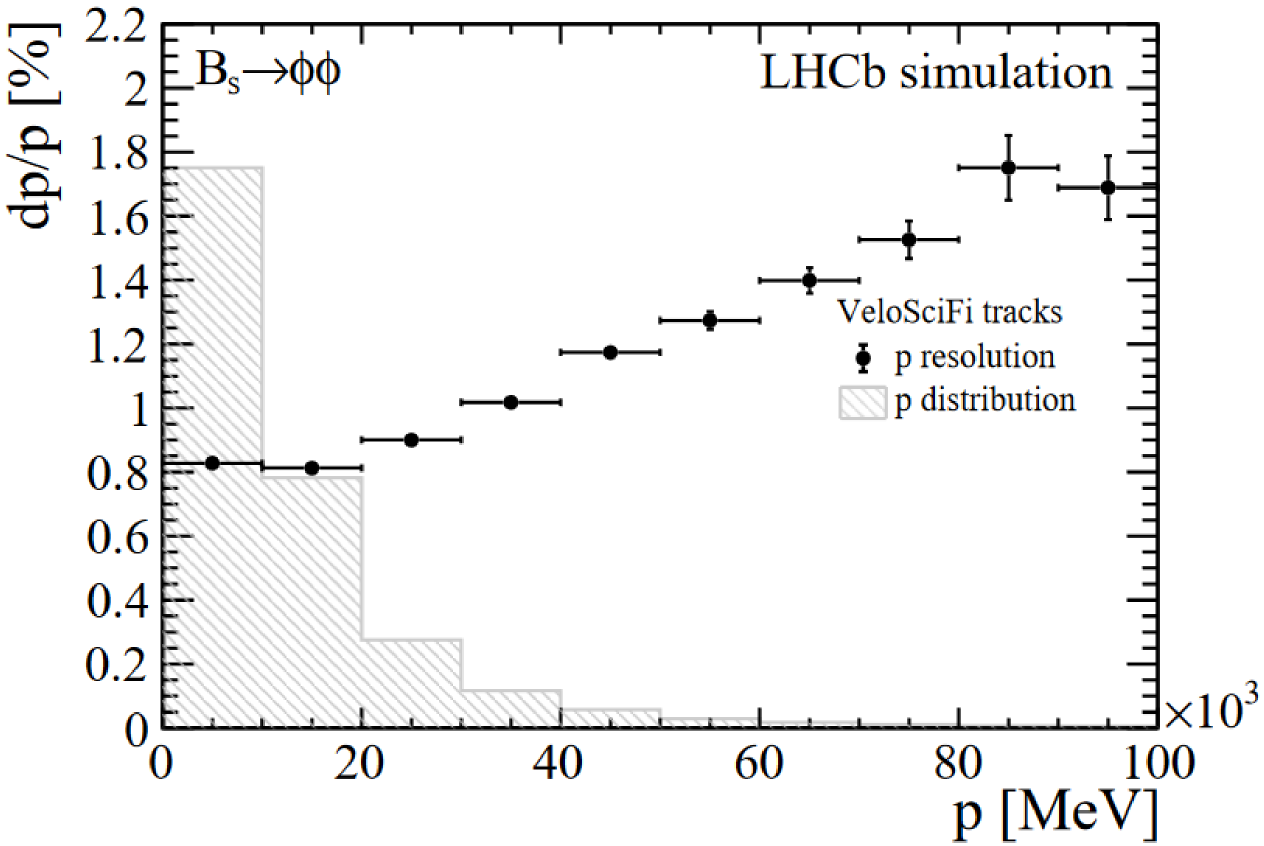




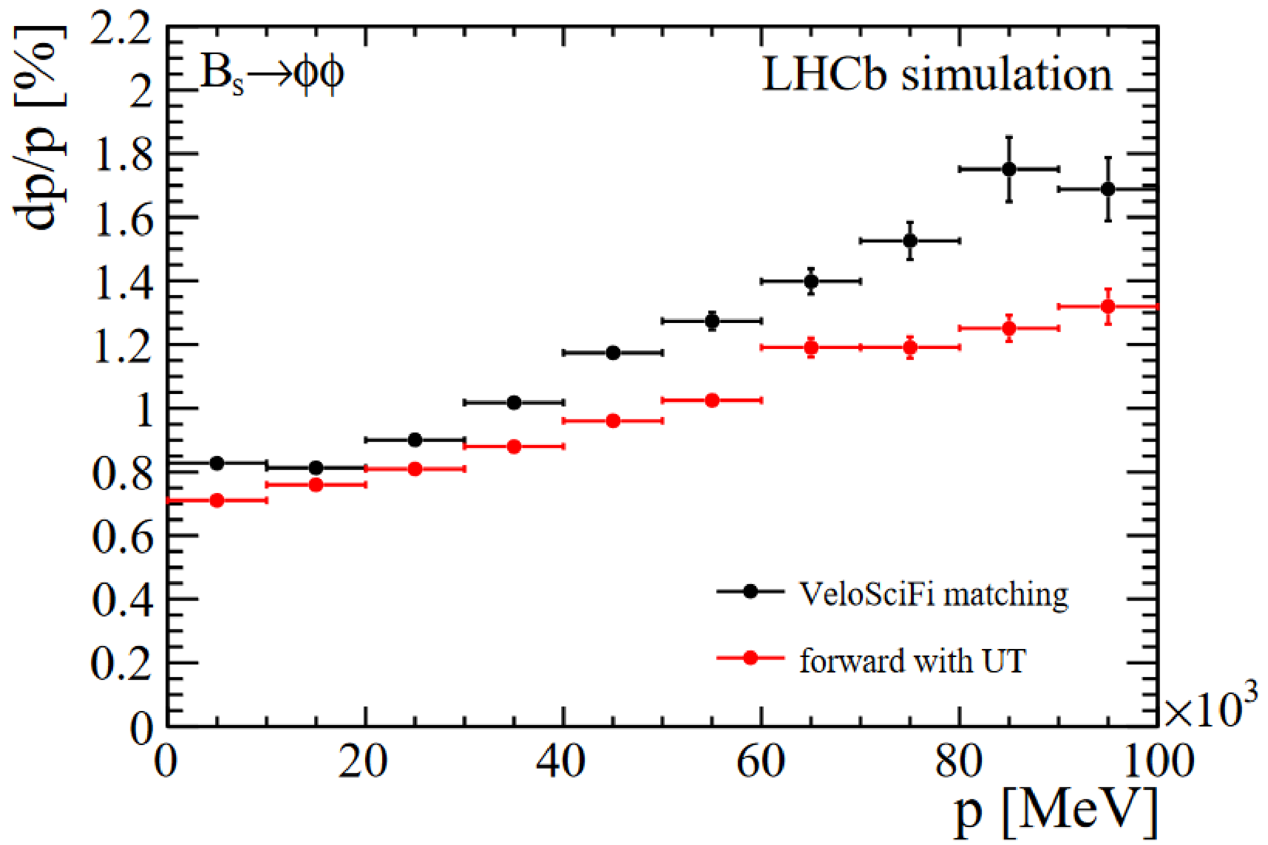
Thank you



Backup: Performance (Momentum resolution) [LHCb-FIGURE-2022-010]

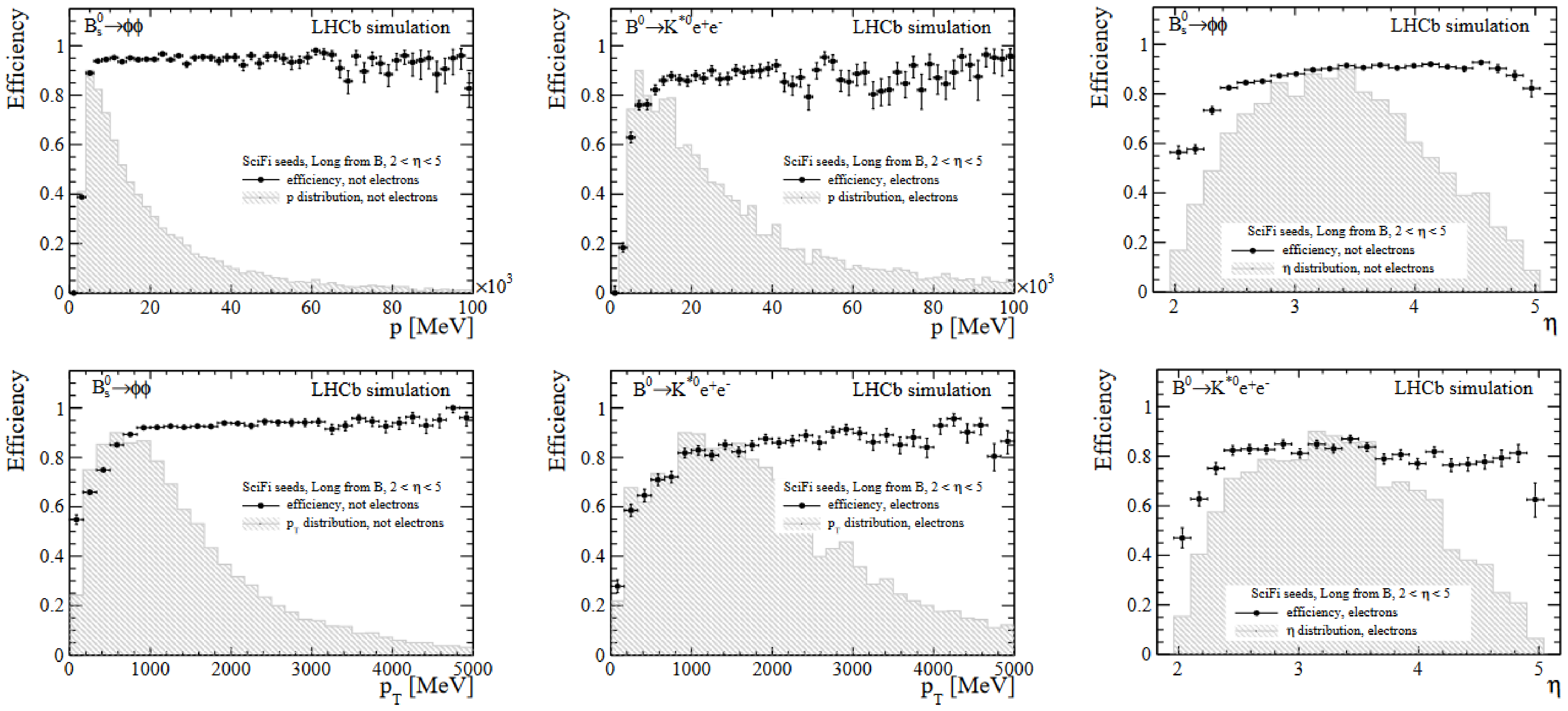


Momentum resolution for HLT1 long tracks from B decays as a function of momentum p. The plot shows reconstructed tracks from 5000 simulated B0s → ϕϕ events where the seeding + matching approach is used for the reconstruction.



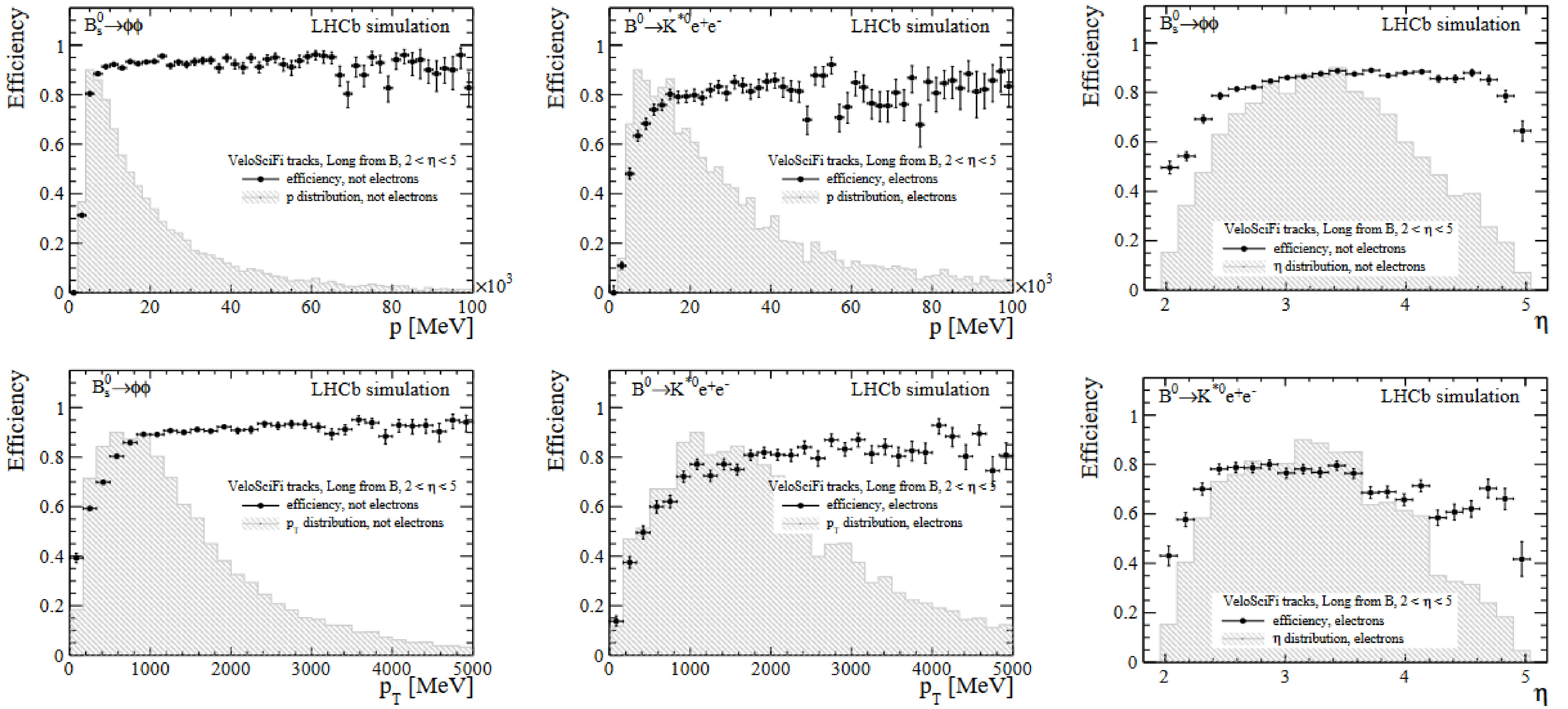
Comparison of the momentum resolution between HLT1 long tracks from the HSM (without UT) (black) and the forward tracking (with UT) (red) from B decays as function of momentum p. The plots are showing reconstructed tracks from 5000 simulated B0s → ϕϕ events. Here the main difference in the high-p region is mainly due to the HSM algorithms running without UT which is the scenario for the data taking in the first year of Run 3

Backup: Performance (Efficiencies SciFi seeds) [LHCb-FIGURE-2022-010]



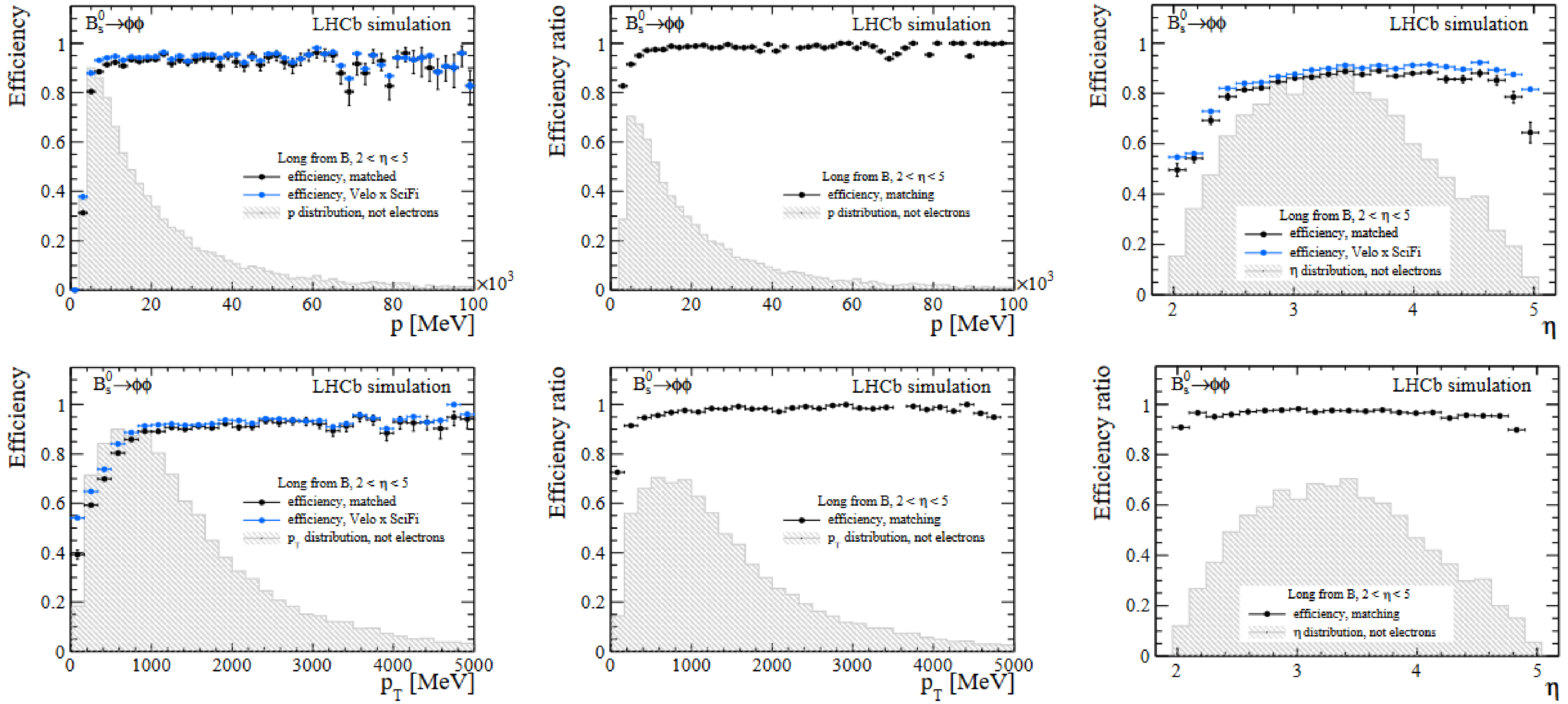
Tracking efficiencies for HLT1 SciFi seeds from B decays as function of momentum p and transverse momentum pT and pseudo-rapidity η. The plots are showing reconstructed SciFi seeds from 5000 simulated B0

Backup: Performance (Efficiencies Scifi + Velo matching) [LHCb-FIGURE-2022-010]

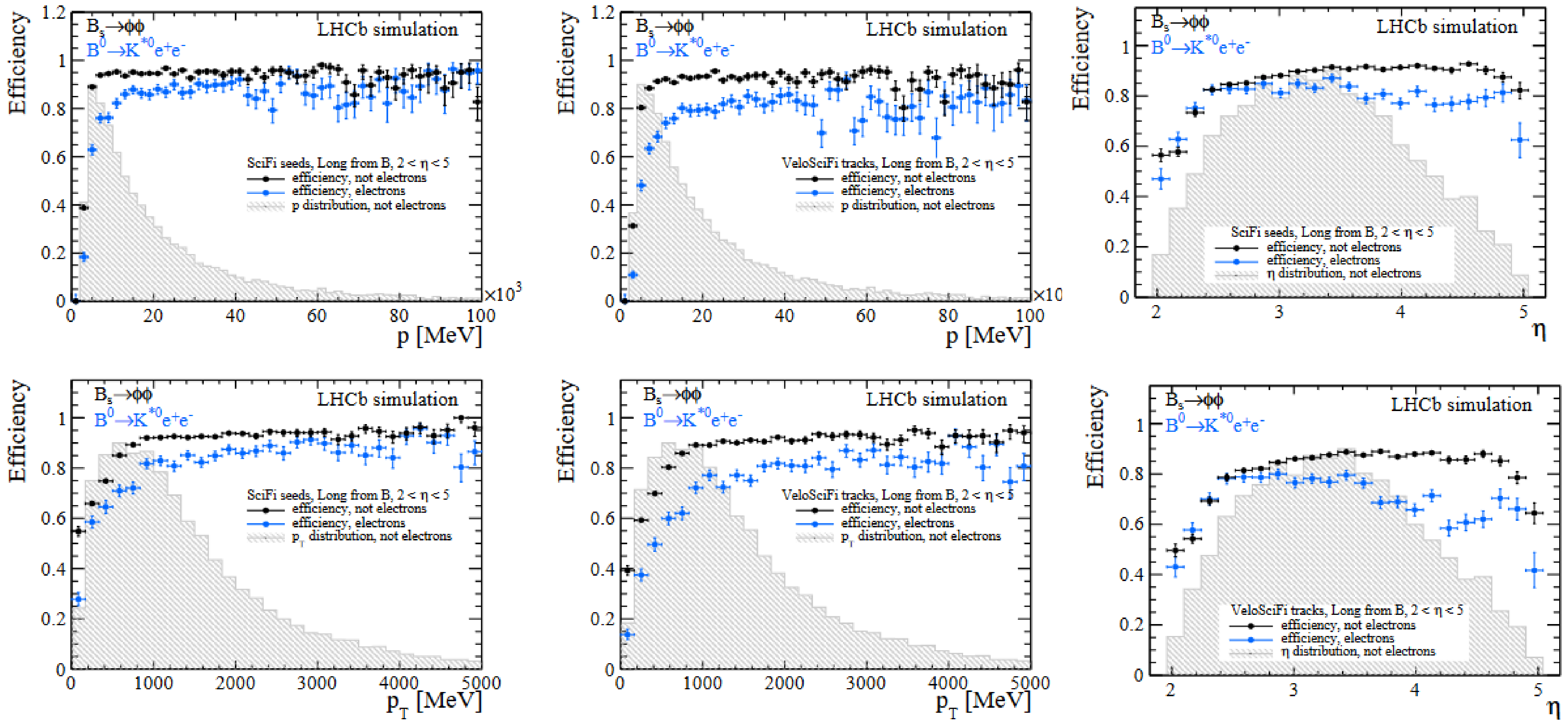


Tracking efficiencies for HLT1 long tracks from B decays as function of momentum p and transverse momentum pT and pseudo-rapidity η. The plots are showing reconstructed tracks from 5000 simulated B0 events where the seeding + matching approach is used for the reconstruction.

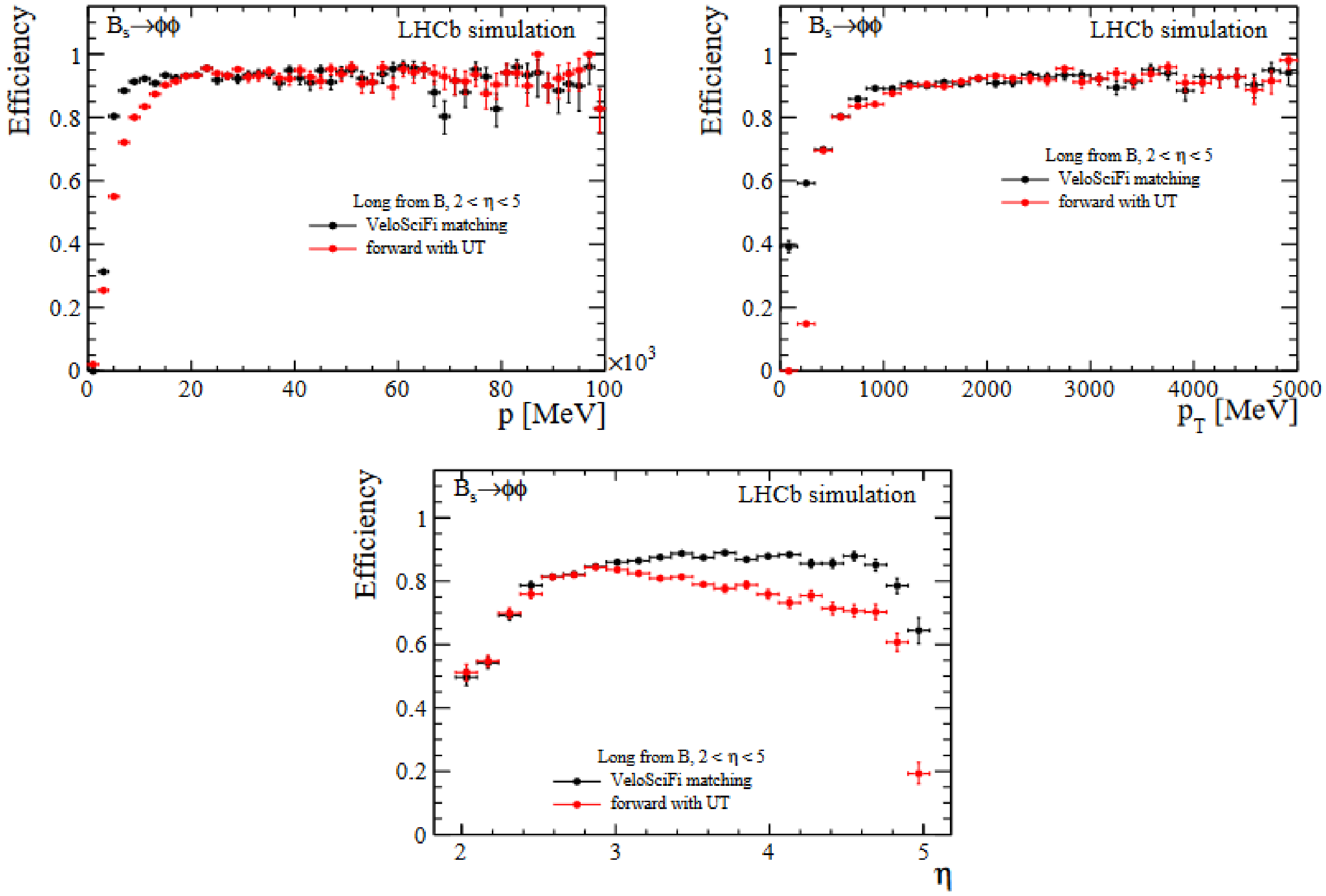
Backup: Performance (Efficiencies comparision of SciFi x Velo vs Matching) [LHCb-FIGURE-2022-010]



Comparison of the tracking efficiency for HLT1 long tracks from B decays with the product of the tracking efficiencies from VELO tracks and SciFi seeds (left) as function of momentum p and transverse momentum pT and pseudo-rapidity η as well as their ratios which is corresponding to the efficiency of the matching (right). The plots are showing reconstructed tracks from 5000 simulated B0s → ϕϕ events where the VELO tracking, SciFi seeding and seeding + matching approach are used for the reconstruction



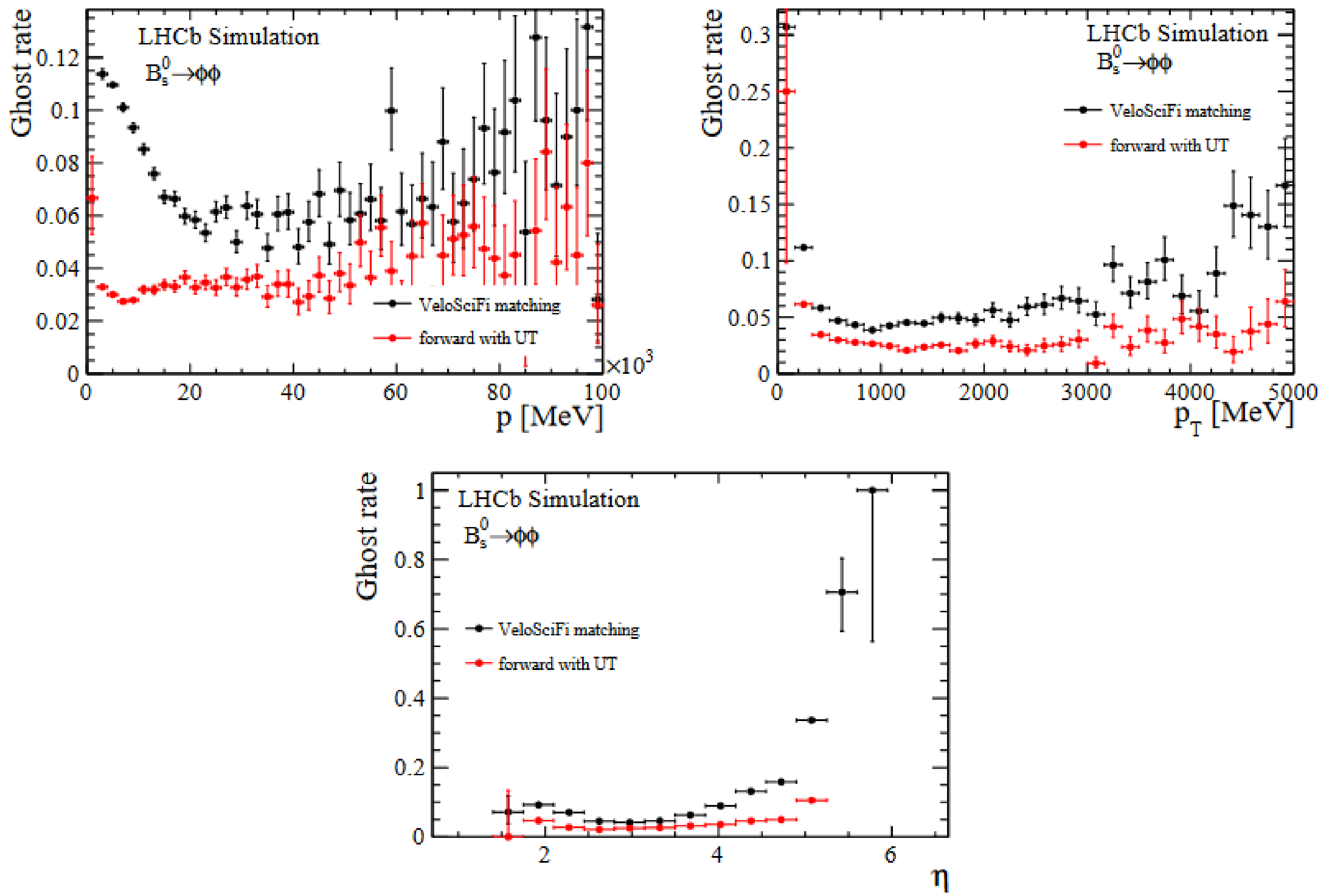
Tracking efficiencies for HLT1 electron (blue) and non-electron (black) SciFi seeds (left) and VeloSciFi tracks (right) from B decays as function of momentum p, transverse momentum pT and pseudo-rapidity η. The plots are showing reconstructed tracks from 5000 simulated B0s → ϕϕ (black) and B0 → K∗0e+e− (blue)



Comparison of the tracking efficiencies between HLT1 long tracks from the Velo-SciFi matching (black) and the forward tracking with UT (red) from B decays as function of momentum p and transverse momentum pT and pseudo-rapidity η. The plots are showing reconstructed tracks from 5000 simulated B0s → ϕϕ events.

Backup: Performance (Ghost rates comparison of long tracks from forward vs Scifi+Velo matching)

[LHCb-FIGURE-2022-010]



Comparison of the ghost rates between HLT1 long tracks from the Velo-SciFi matching (black) and the forward tracking with UT (red) from B decays as function of momentum p and transverse momentum pT and pseudorapidity η. The plots are showing reconstructed tracks from 5000 simulated B0s → ϕϕ events.