Physics and Computing Performance of reconstruction algorithms for the GPU High Level Trigger 1 of LHCb

Dorothea vom Bruch*[1]

on behalf of the LHCb collaboration

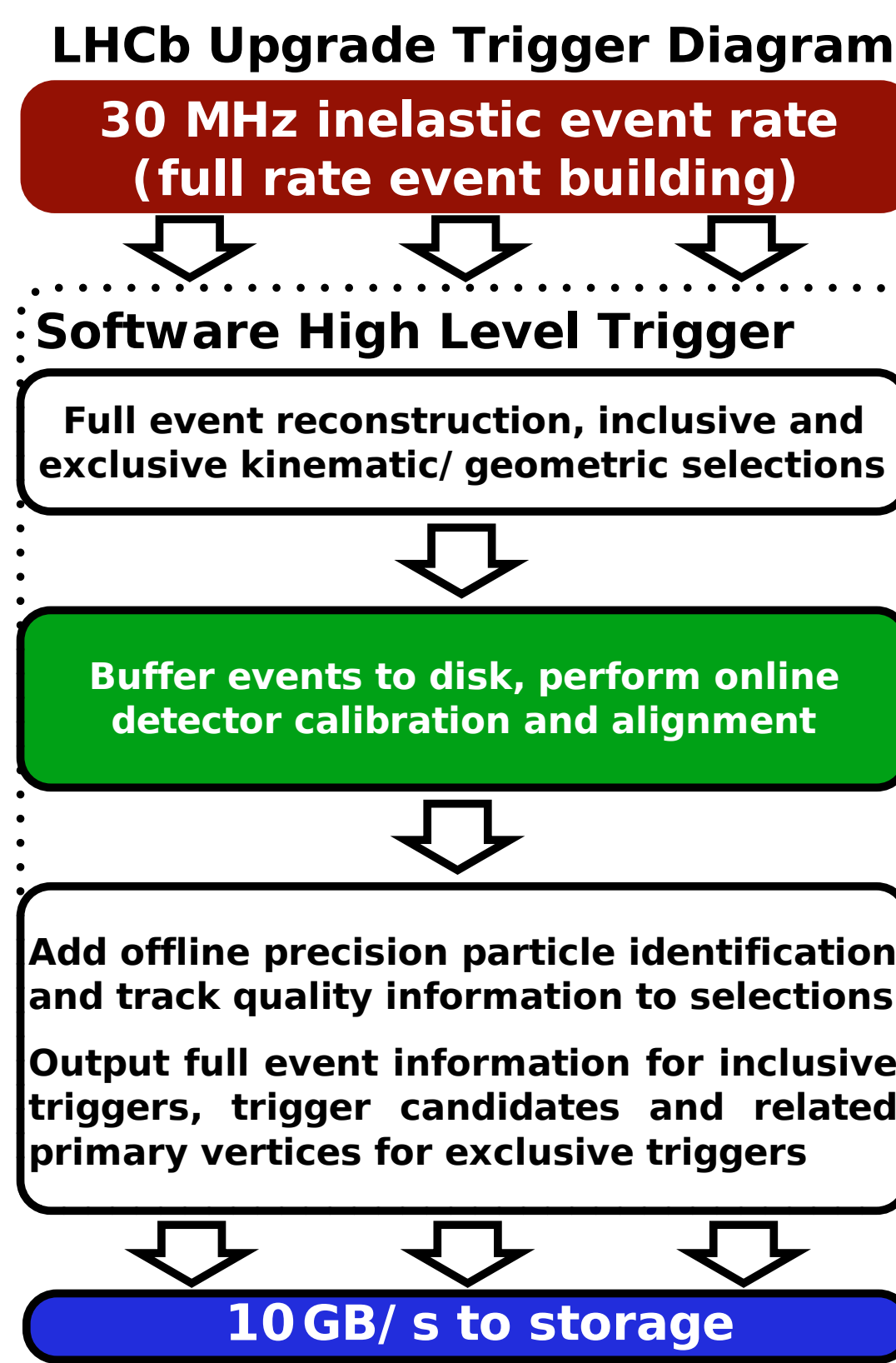[1] LPNHE, Sorbonne Université, Paris Diderot Sorbonne Paris Cité, CNRS/IN2P3, Paris, France

Motivation

- From 2021: software-only high level trigger (HLT)
- Significant computing challenge
- Potential option: use GPUs

R&D standalone project: Allen

- Run full first stage of HLT (HLT1) on GPUs
- Process thousands of events in parallel
- Exploit data-parallelism within events
- One GPU has to process 30/60 k events/s

LHCb Upgrade Trigger Diagram

30 MHz inelastic event rate (full rate event building)

Software High Level Trigger

Full event reconstruction, inclusive and exclusive kinematic/ geometric selections

Buffer events to disk, perform online detector calibration and alignment

Add offline precision particle identification and track quality information to selections

Output full event information for inclusive triggers, trigger candidates and related primary vertices for exclusive triggers

10 GB/ s to storage

Infrastructure of Allen:

- Custom memory manager for GPU memory
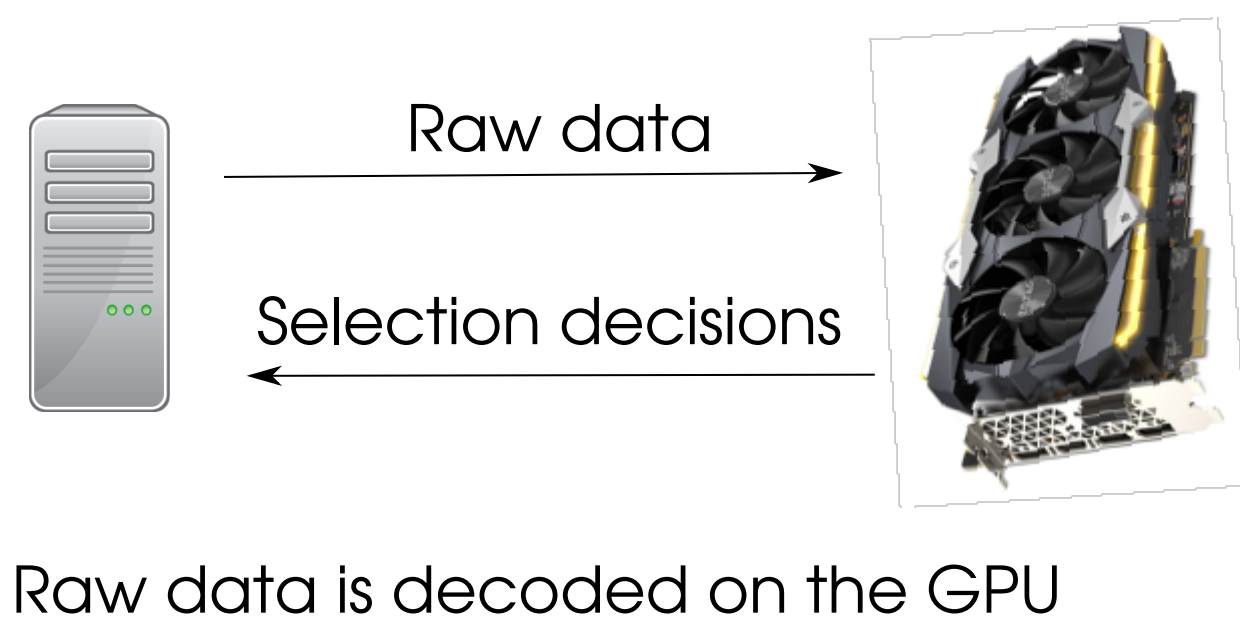- Static scheduler
- Physics performance checks

Link to Allen gitlab repository

Tasks of Allen:

- Run full HLT1 chain
- Reduce data rate by factor 30 based on single and two-track selections

---

Data flow

Event builder node

Raw data

Selection decisions

Raw data is decoded on the GPU

Summary

- First project running a **full HLT** on GPUs for a HEP experiment
- All decoding, clustering and tracking **algorithms** are **implemented in CUDA**
- The **current sequence** of full Velo, primary vertices, full UT and SciFi decoding runs on a V100 at **112 kHz**

6. Muon: 4 multi wire proportional chambers

Muon ID

- Extrapolate SciFi tracks to muon stations, find closest hits
- Decide if track originates from a muon
  Work in progress

  Parallelize over events, SciFi tracks and stations
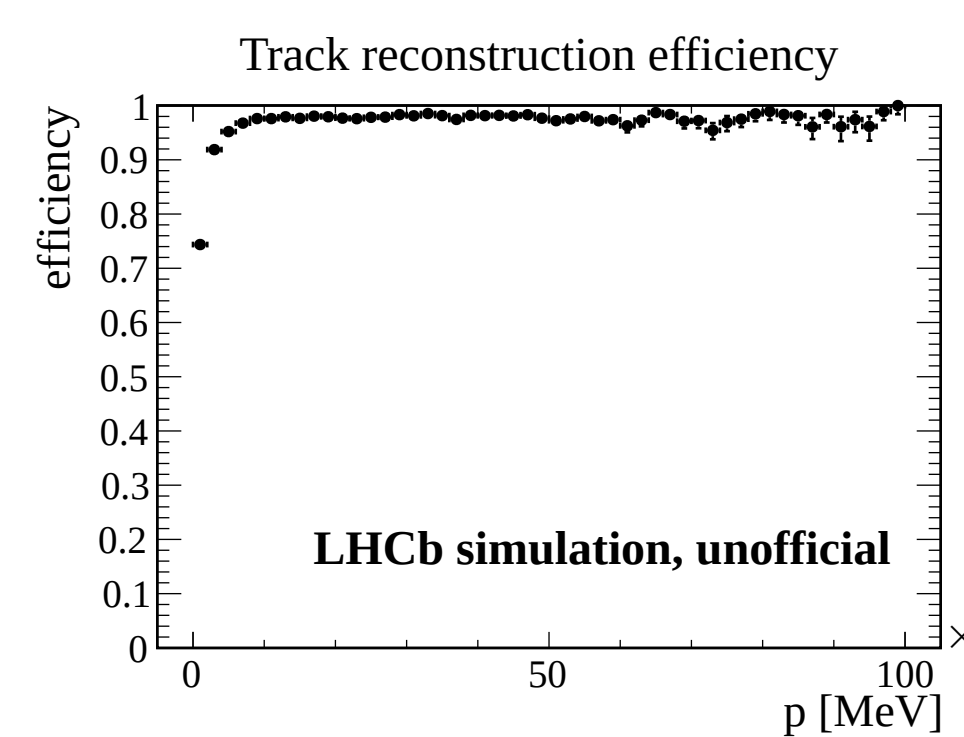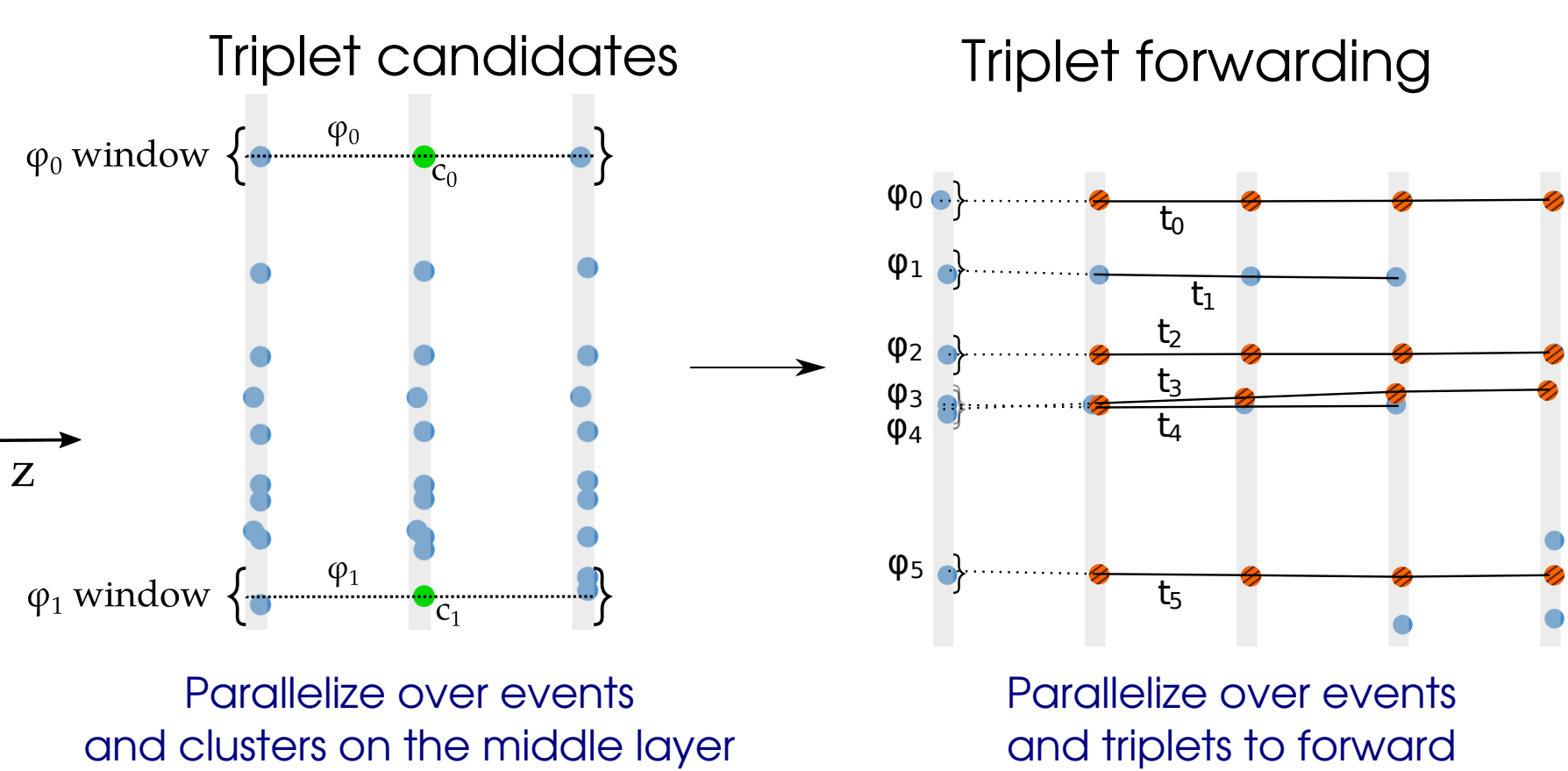
1. Velo: Pixel detector, 26 planes
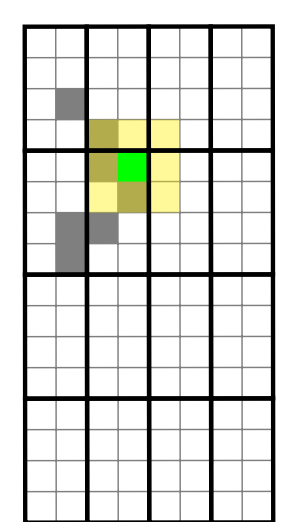
Masked clustering
- Find cluster seeds
- Load only neighbouring pixels of a seed, use 8-bit mask to find cluster
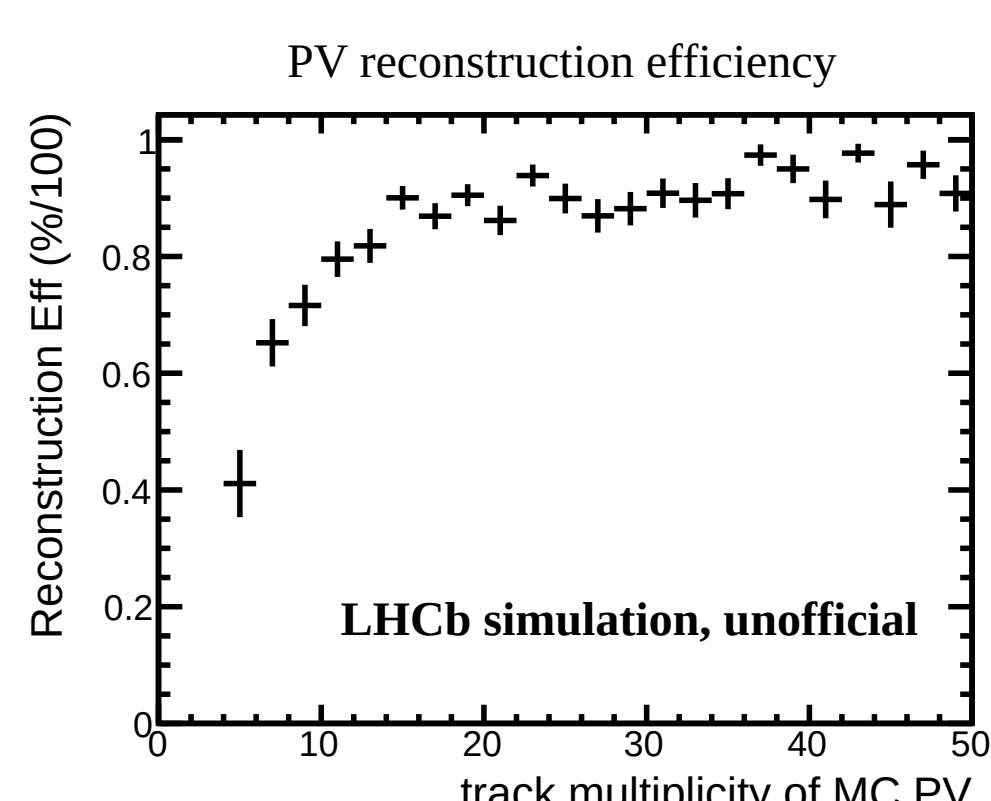
Parallelize over events and within raw banks

Parallelize over events and seeds

Pattern recognition: Search by triplet

Triplet candidates

Triplet forwarding

Parallelize over events and clusters on the middle layer

Parallelize over events and triplets to forward

Track reconstruction efficiency

LHCb simulation, unofficial

5. Kalman filter

- Use single precision
- Parametrize transport in magnetic field
- Run w/o smoother

Parallelize over events and SciFi tracks

IP resolution

σ_{VELO} = (25.6±1.0) + (9.5±0.8)/p_T µm
σ_{Kalman} = (23.3±1.1) + (9.5±0.8)/p_T µm

LHCb simulation, unofficial

Momentum resolution

LHCb simulation, unofficial

2. Primary vertex finding

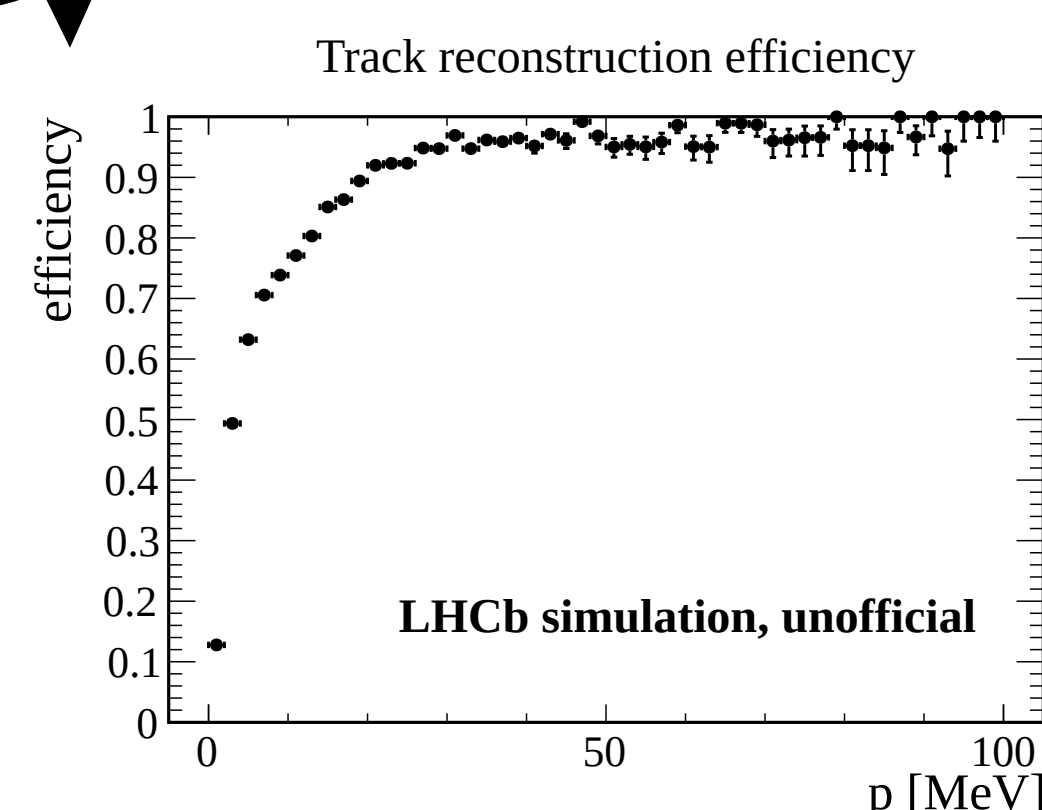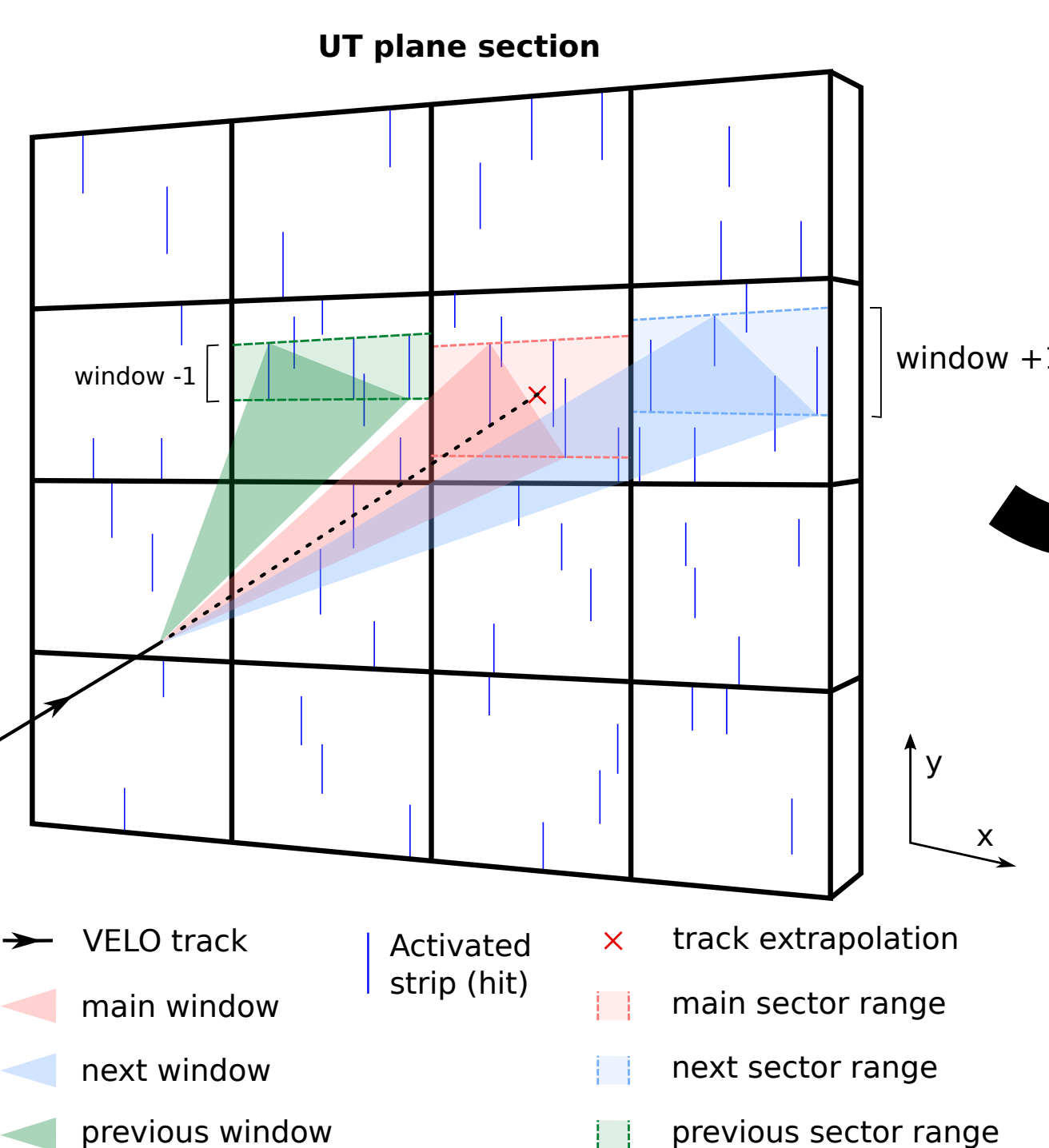- Extrapolate Velo tracks to beamline, find z position of track
- Fill histogram with z-positions
- Find peaks in histogram → seeds
- Vertex fit using weight

Parallelize over events and Velo tracks

Parallelize over events

Parallelize over events and seeds found

PV reconstruction efficiency

LHCb simulation, unofficial

3. UT: Strip detector, 4 planes

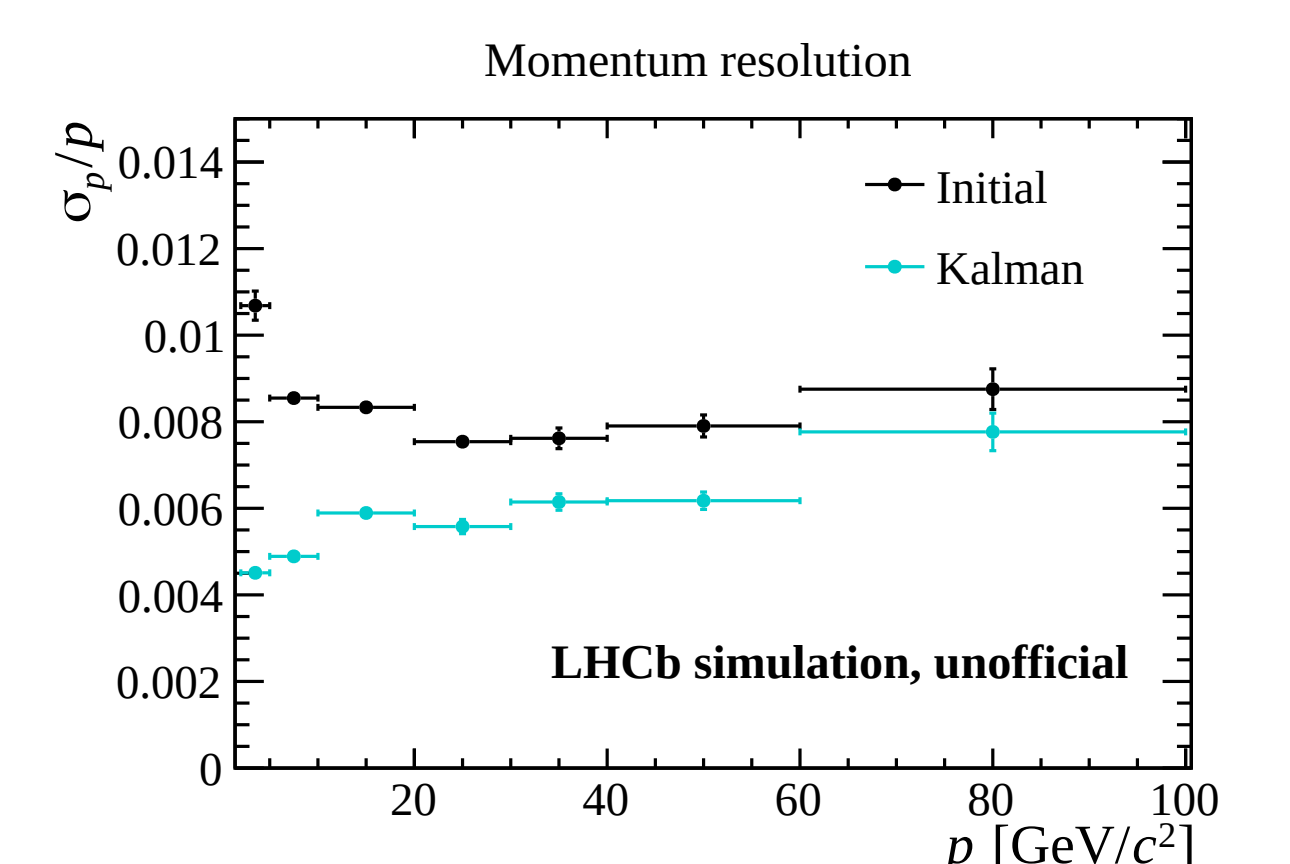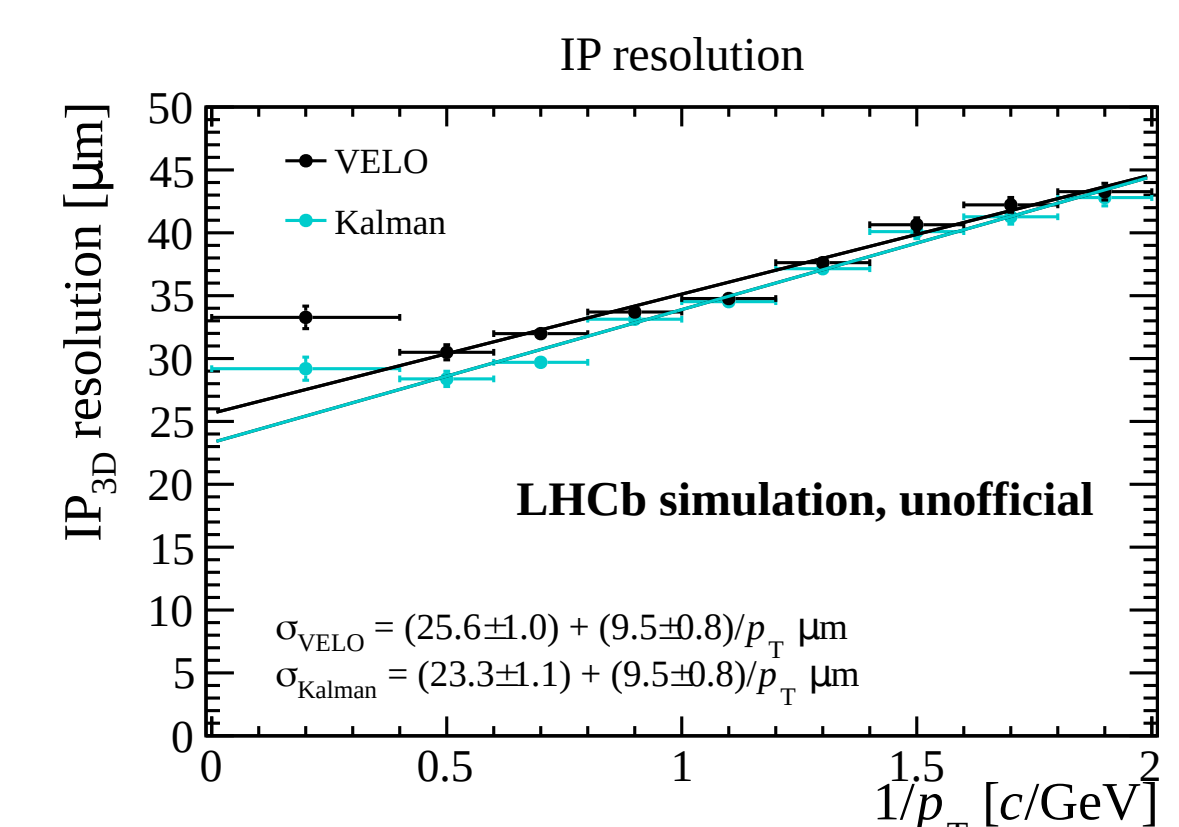### Pattern recognition
- Extrapolate Velo tracks to the UT planes, define search windows
- Form 3-/4-hit tracks
- Obtain momentum estimate from chi2-fit

Parallelize over events and Velo tracks

Parallelize over events, layers and Velo tracks

UT plane section

window -1

window +1

Track reconstruction efficiency

LHCb simulation, unofficial

VELO track
main window
next window
previous window

Activated strip (hit)

track extrapolation
main sector range
next sector range
previous sector range

4. SciFi: 12 planes of 2 x 2.5 m long scintillating fibre arrays

Pattern recognition
- Extrapolate UT tracks to SciFi planes with a parametrization for the magnetic field deflection
- Challenge: single fibre efficiency ~98%

Quadruplets / triplets in last station

x fiber   u fiber

Propagate to other stations

Work in progress