WP4: fast tracking device

Nicola Neri

INFN- Sezione di Milano

Cagliari, 1 Dec 2017

Outline

- Introduction
- Real-time tracking
- Artificial retina algorithm
- Detector prototype with embedded tracking capabilities
- 4D real-time tracking
- Summary

Trigger strategy at LHCb

- CPU-based trigger is already a big challenge for LHCb upgrade at $L=2 \cdot 10^{33}\,cm^{-2}\,s^{-1}$ in 2020
- Trigger decision based on tracking information of the entire event

	CPU time[ms]	
Tracking Algorithm	No GEC	GEC = 1200
VELO tracking	2.3	2.0
VELO-UT tracking	1.4	1.3
Forward tracking	2.5	1.9
PV finding	0.40	0.38
Total @29 MHz		5.6
Total	6.6	5.4

Upgrade HLT

VELO tracking

VELO-UT
$p_T > 200$ MeV/c

Forward reco
$p_T > 500$ MeV/c

PV finding

Trigger cuts to
reduce rate to 1 MHz

Muon ID

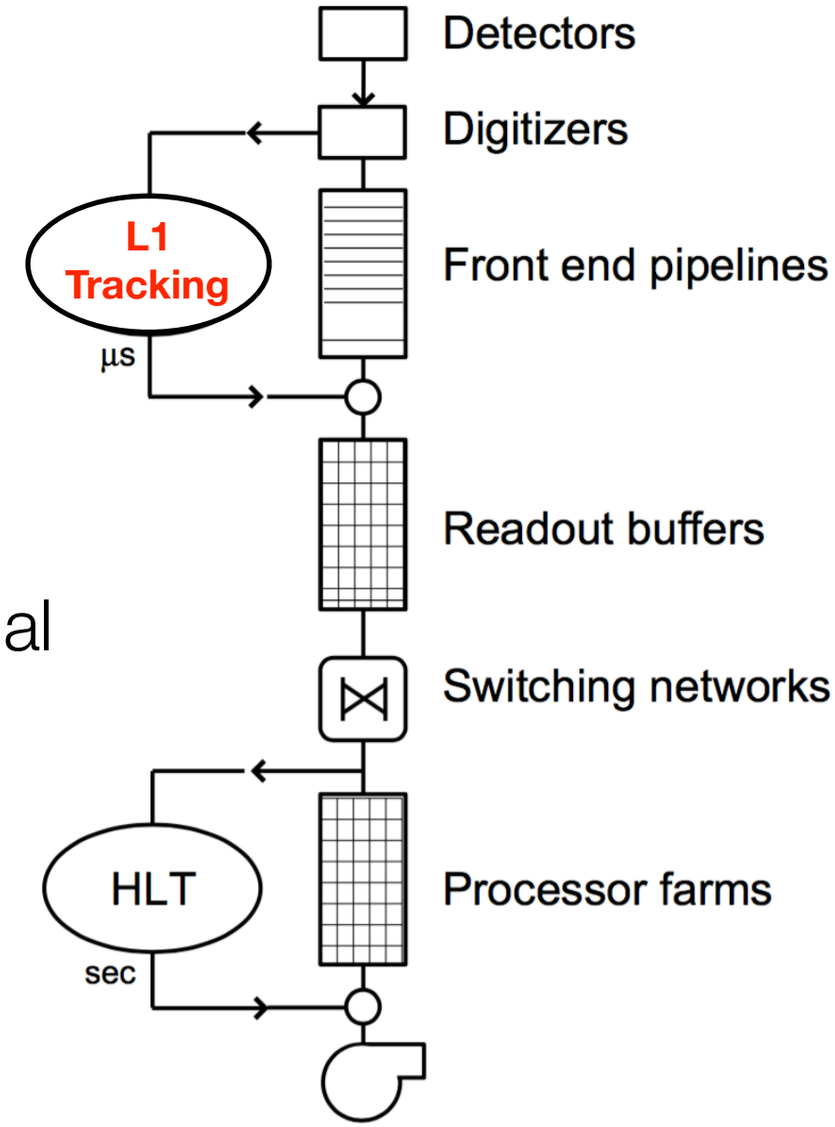
Simplified Kalman Fit

Online RICH PID

- No obvious solution for CPU-based trigger at high luminosity $L=2 \cdot 10^{34}\,cm^{-2}\,s^{-1}$

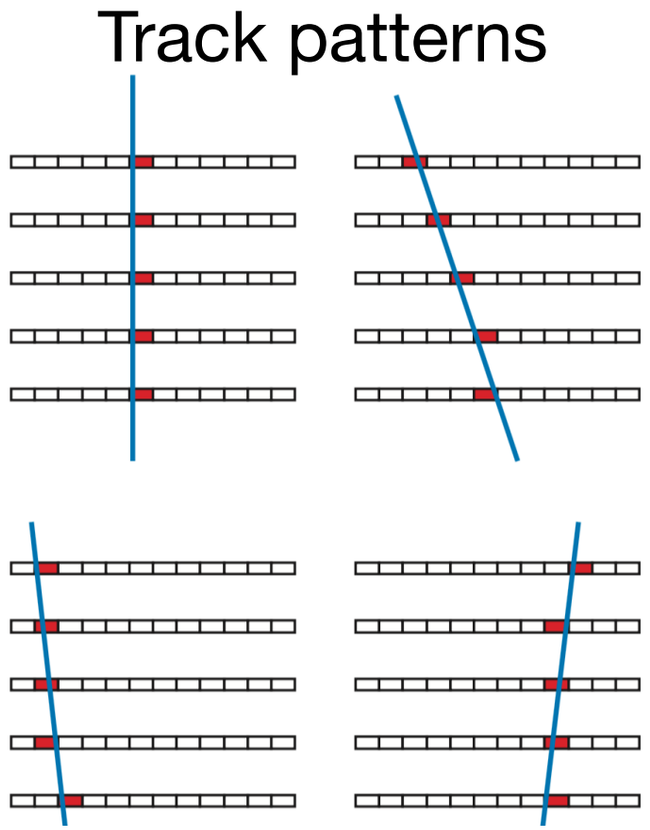
Real-time tracking for HL-LHC

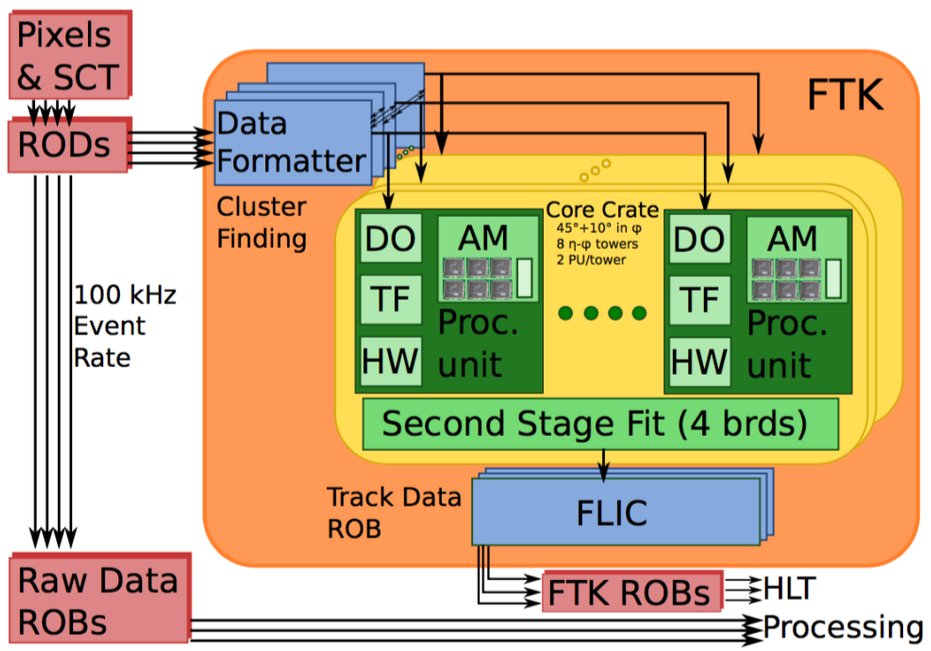
- Full exploitation of HL-LHC requires new detectors and trigger systems
- L1 trigger decisions based on tracking information are crucial:
  - reduce data rate to a sustainable level
  - maintain good efficiency and purity for signal events
- Real time tracking is extremely challenging at LHC: 40MHz throughput, large flow of data Tbit/s, short latency ≲1μs
- Necessary to find innovative solutions



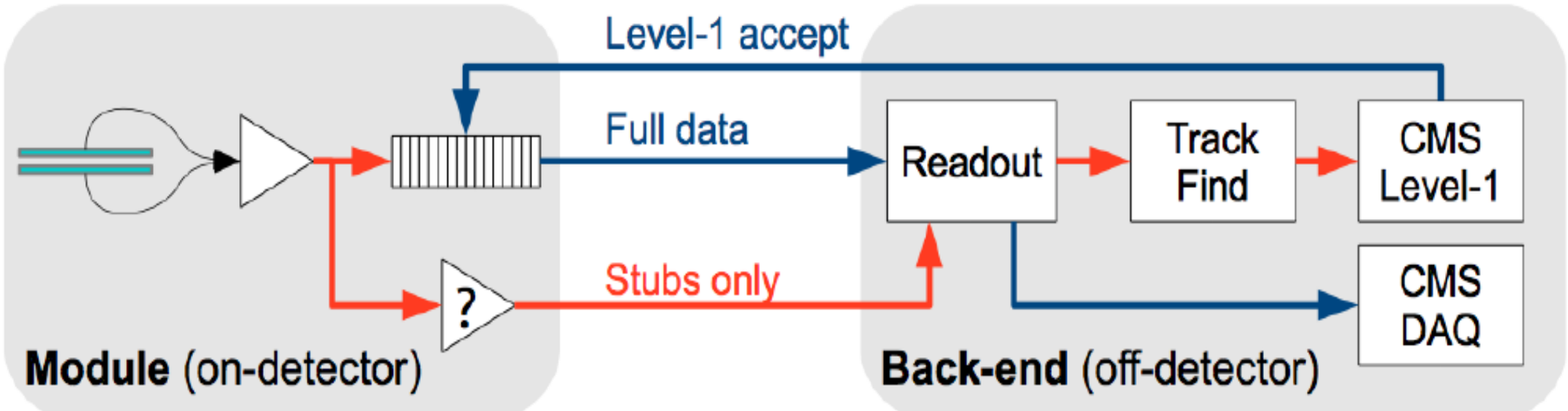
Existing fast track finders

- Track pattern recognition without combinatorics
  - parallel matching of hits to pre-calculated track patterns, track parameters from linearised fit
  - use custom ASICs: Associative Memory (AM), based on content-addressable memory (CAM)
- First use in CDF experiment: SVT, latency 10μs and input rate 30 kHz
- FTK device in ATLAS use similar concept. Latency ~50μs and input rate 100 kHz





CMS phase-II trigger scheme



L1 Latency 10 µs
(20 µs in option)
L1A rate ≤1 MHz
HLT rate ≤10 kHz

- Use stubs information at early stage of processing to reduce data rate and fast tracking for L1 trigger

Retina INFN project

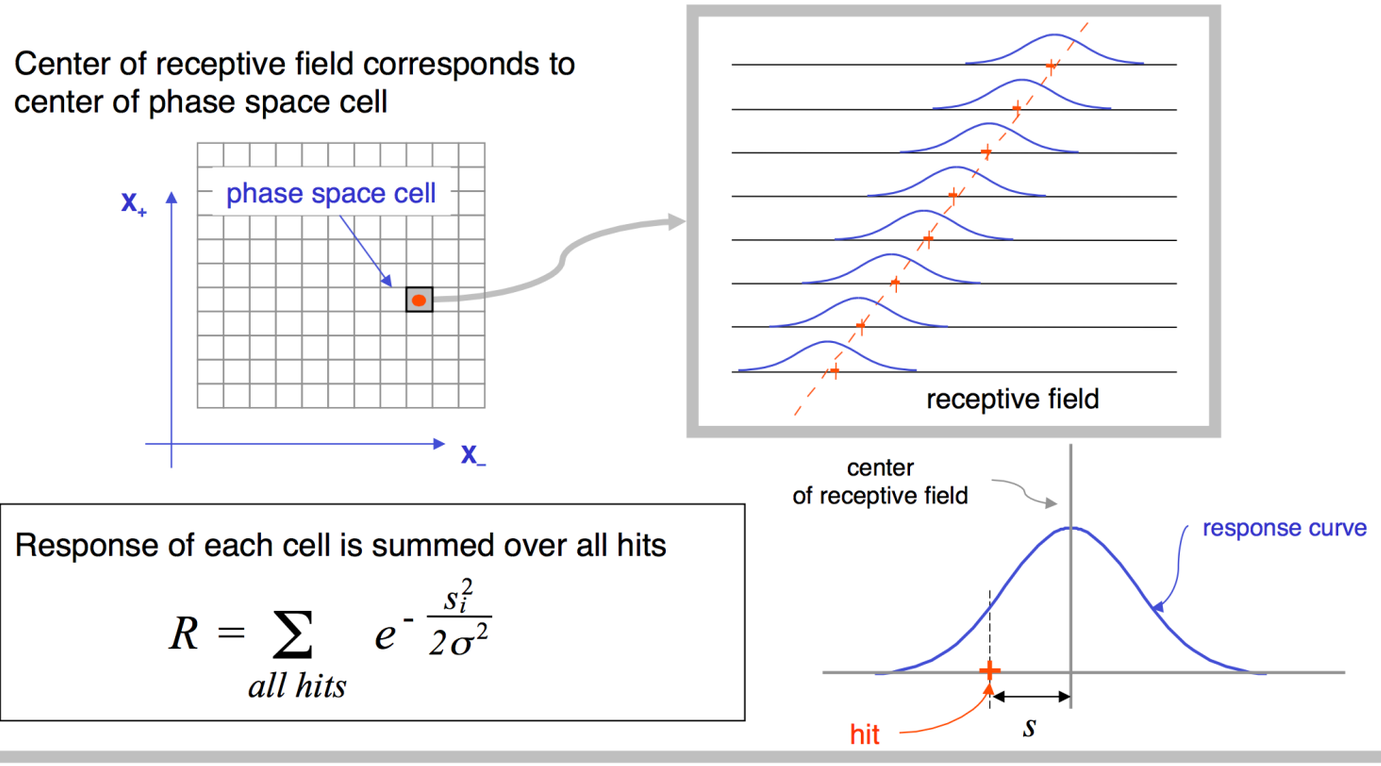
- INFN-Retina R&D project started in 2015. Milano and Pisa groups involved
- Develop hardware prototype of a real time tracking device for intensive tracking applications (1-100 Giga tracks/sec), *e.g.* HL-LHC experiments
- Main deliverables:

  - Real time tracking detector prototype for test beam
  - Fast track finding system compatible with large DAQ framework for test with simulated data at 40 MHz event rate

Artificial retina algorithm

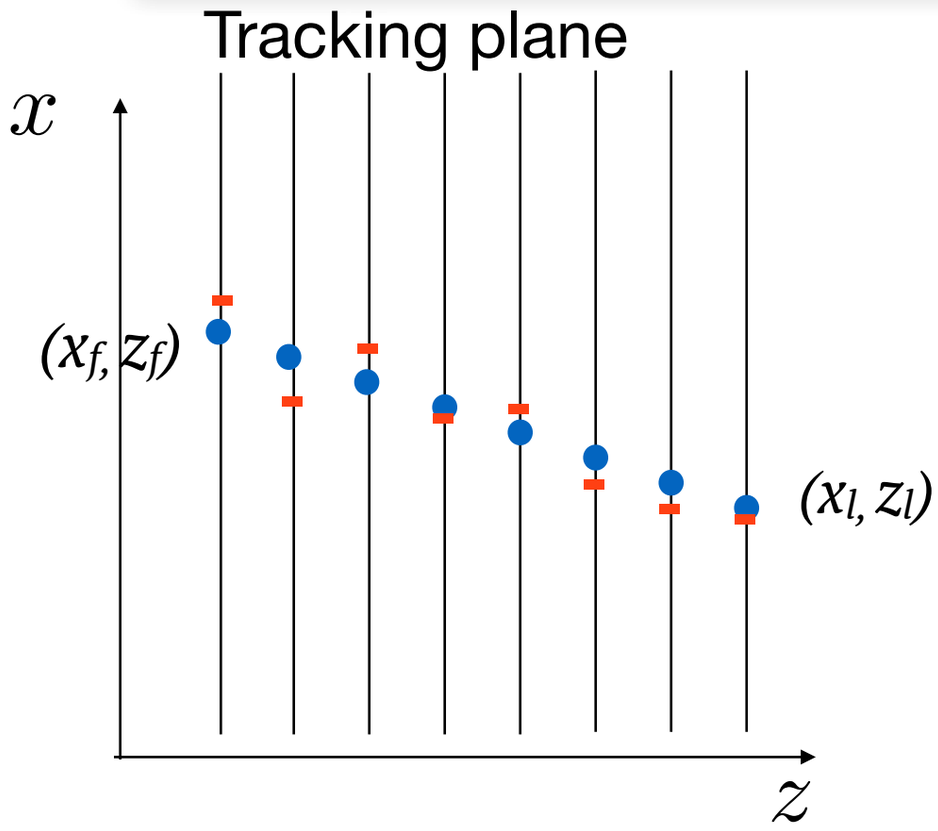
- Basic algorithm for fast track finding

L. Ristori, "An artificial retina for real-time track finding" [NIM A453 (2000) 425-429]



- Inspired by mechanism of visual receptive fields
  - massive parallelisation and analog response of track receptors (*R*)
  - pattern recognition and track fit by interpolation of *R* values

Track identified by retina algorithm



S_i distance between cluster and track receptor

cluster of hits

track receptors

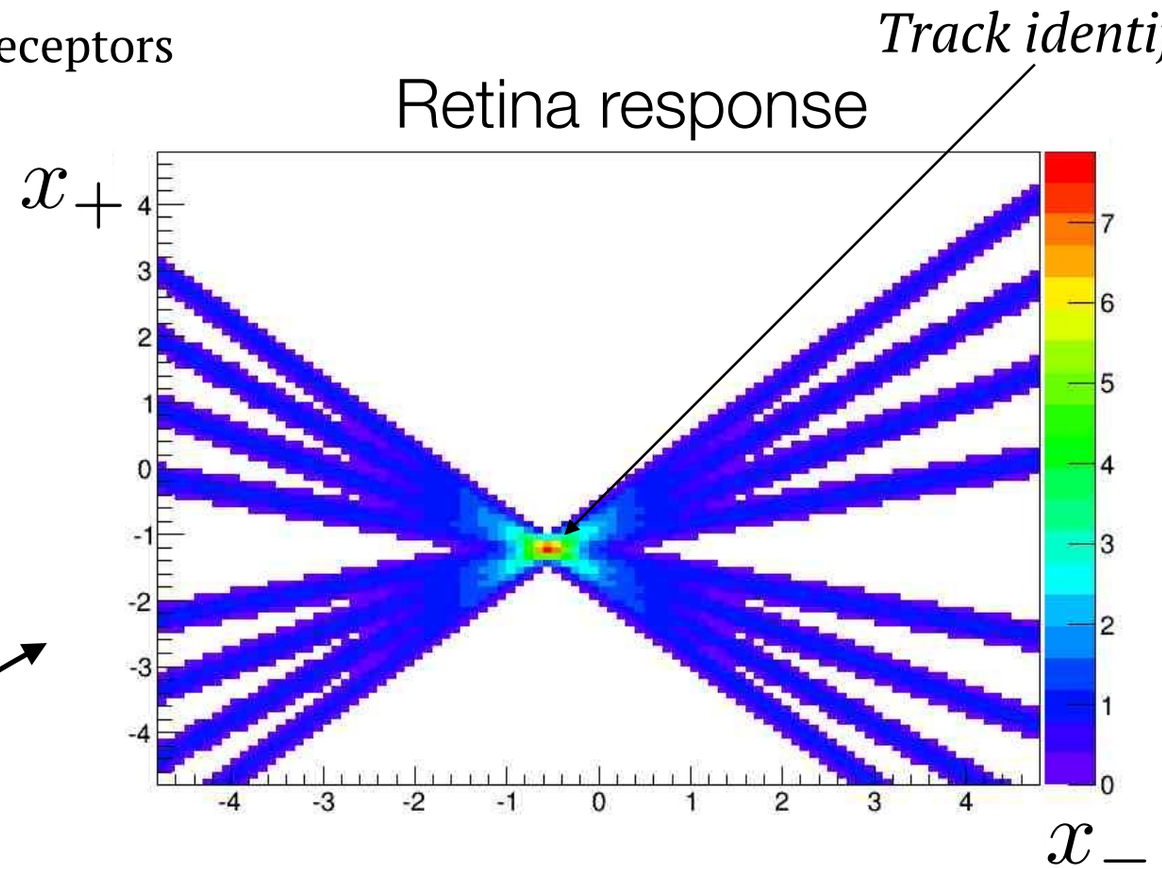
2D track: $x(z) = x_+ + x_- \dfrac{z - z_+}{z_-}$

Track parameters $x_\pm = \dfrac{x_l \pm x_f}{2}$

Excitation of the cellular units

$$R = \sum_i \exp\left(-\frac{s_i^2}{2\sigma^2}\right) \quad \text{if} \quad s_i < 2\sigma$$

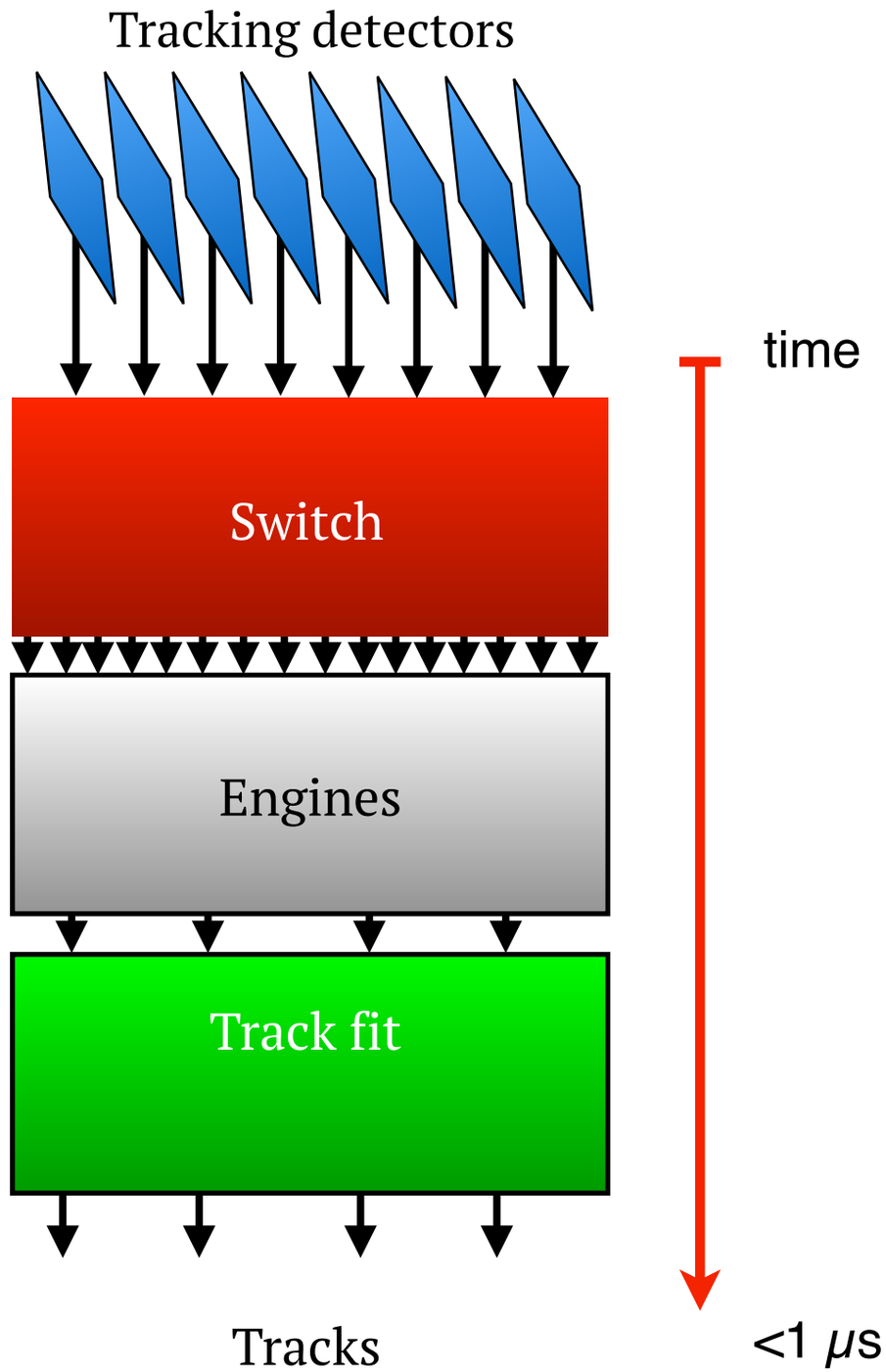
$$R = 0 \quad \text{if} \quad s_i > 2\sigma$$



- Track parameters obtained by interpolation of R values around maximum

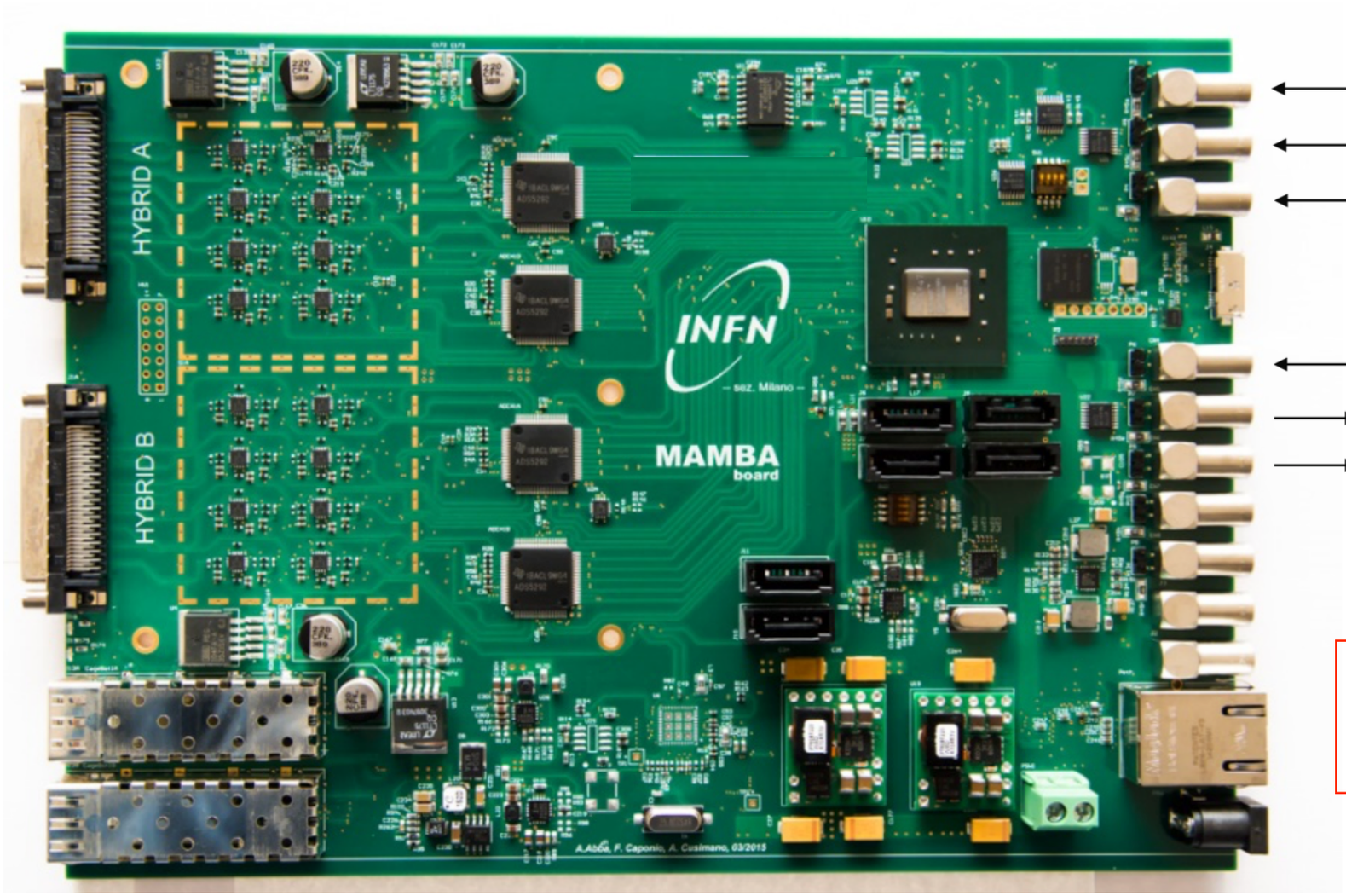
Artificial retina architecture

- **Three main blocks:**
  - Switch: delivers in parallel the hits from the detectors to only appropriate cellular units
  - Engine: block of cellular units for parallel calculation of the weights
  - Track fit: interpolation of adjacent cell weights for track parameter determination
- **Main differences with AM approach:**
  - only relevant data reach the processing units (engines). Data processing starts already in the switch while data is transmitted
  - retina algorithm provides analog response contrarily to AM “yes/no” pattern matching



Mamba board

- Readout 8 detectors in 1x mode (300 kHz) or 2 detectors in 4x mode (1.1 MHz)



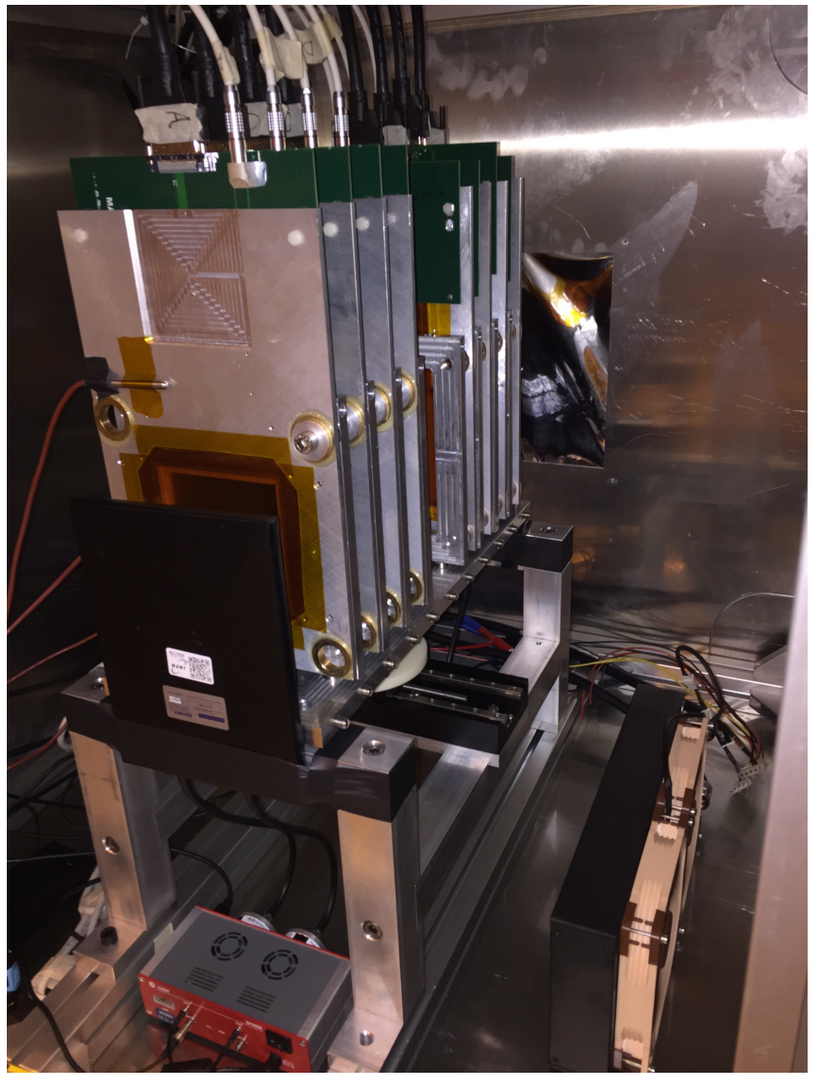
← SCINTILLATOR 1 / COINCIDENCE
← SCINTILLATOR 2
← TELESCOPE CLOCK
← TELESCOPE RESET
→ ACCEPTED TRIGGER
→ BUSY

Mamba board
by Nuclear Instruments

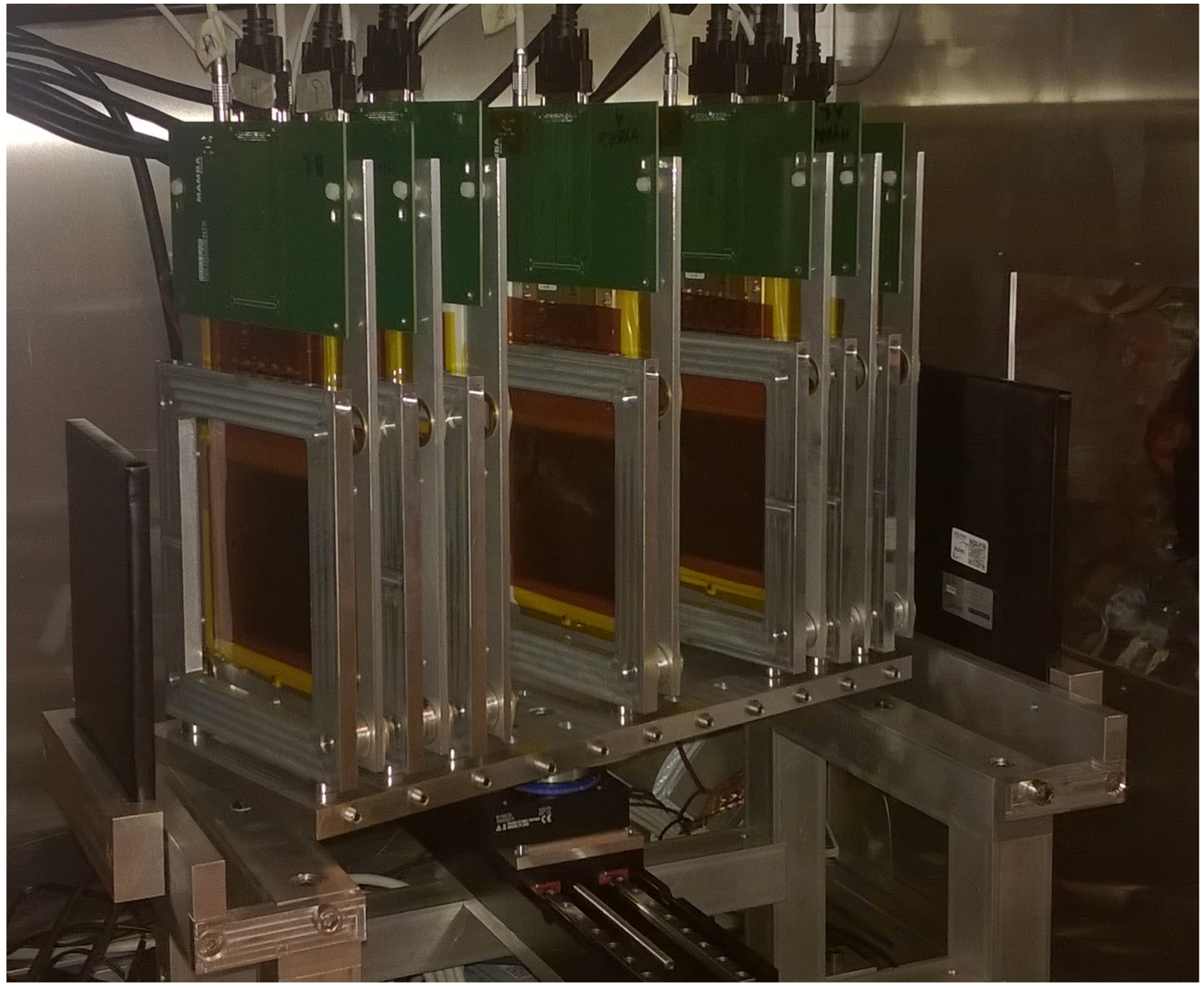
Telescope on beam at SPS

- Telescope tested on 180 GeV/c proton beam
- Rotation angle wrt beam axis: 0, 2, 4, 8, 16, 20 degree

Telescope aligned with beam axis



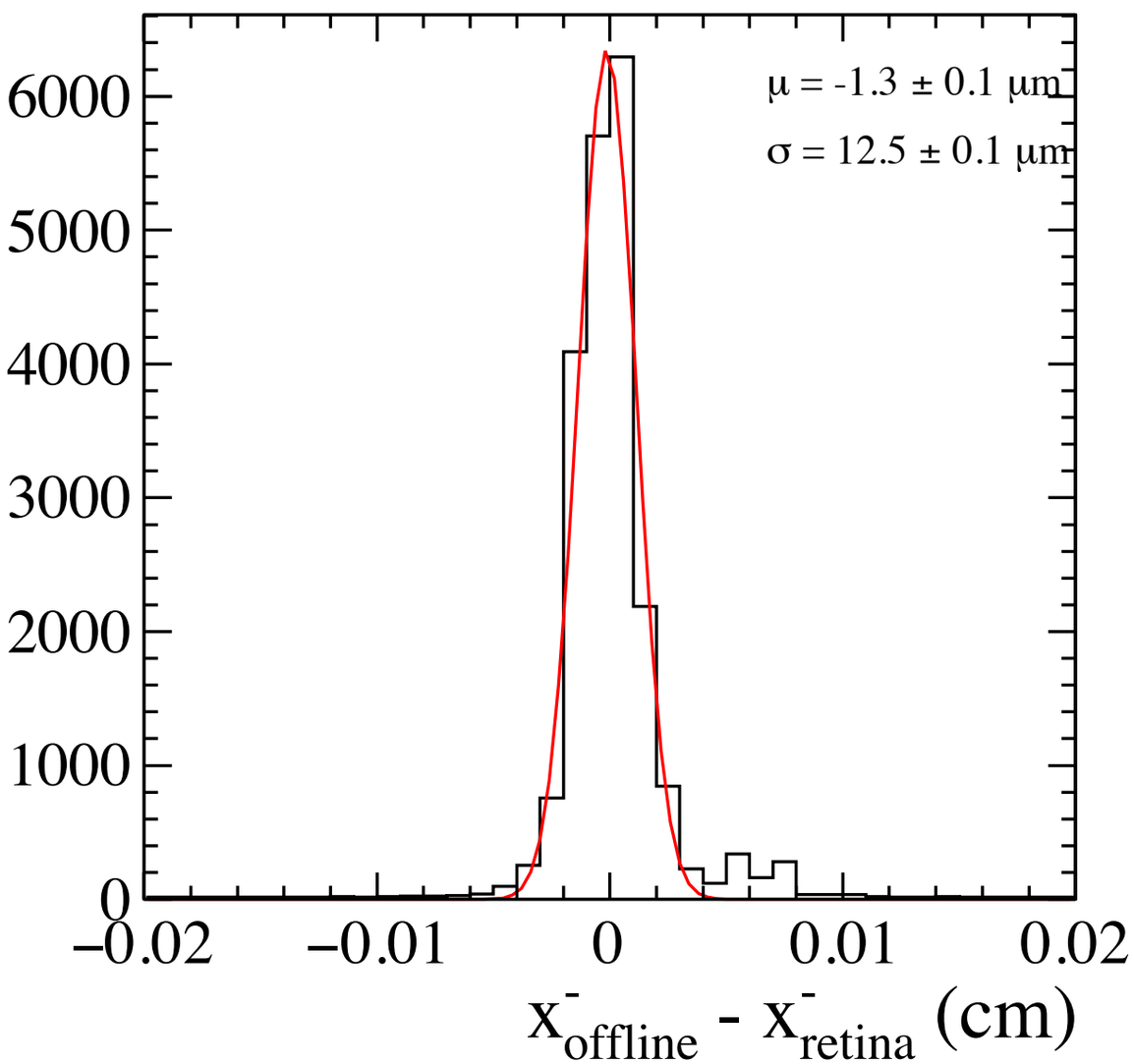
Telescope rotated wrt beam axis

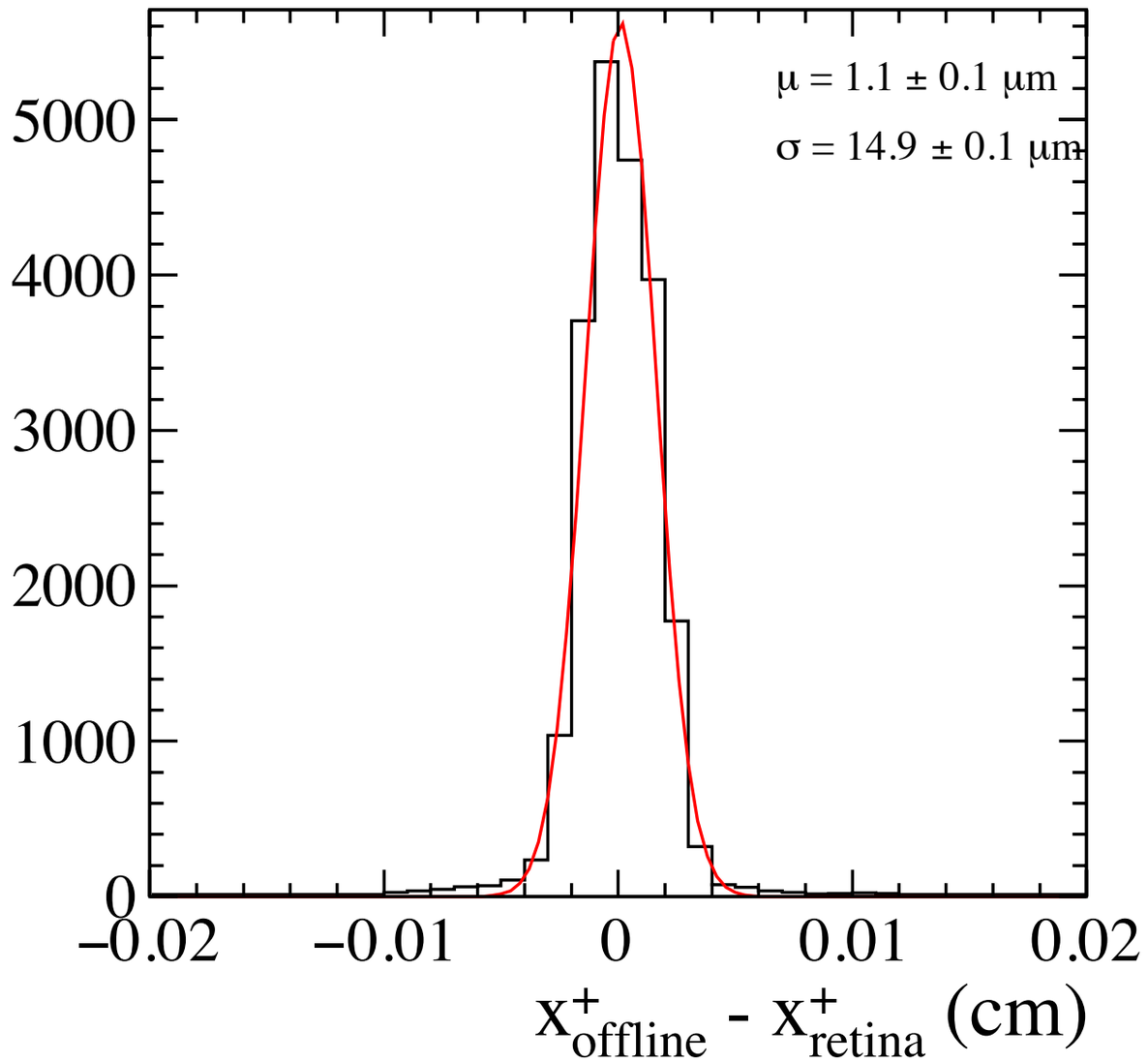


Track residuals: offline - retina

- It works! Offline-Retina track parameter residual are peaked at zero

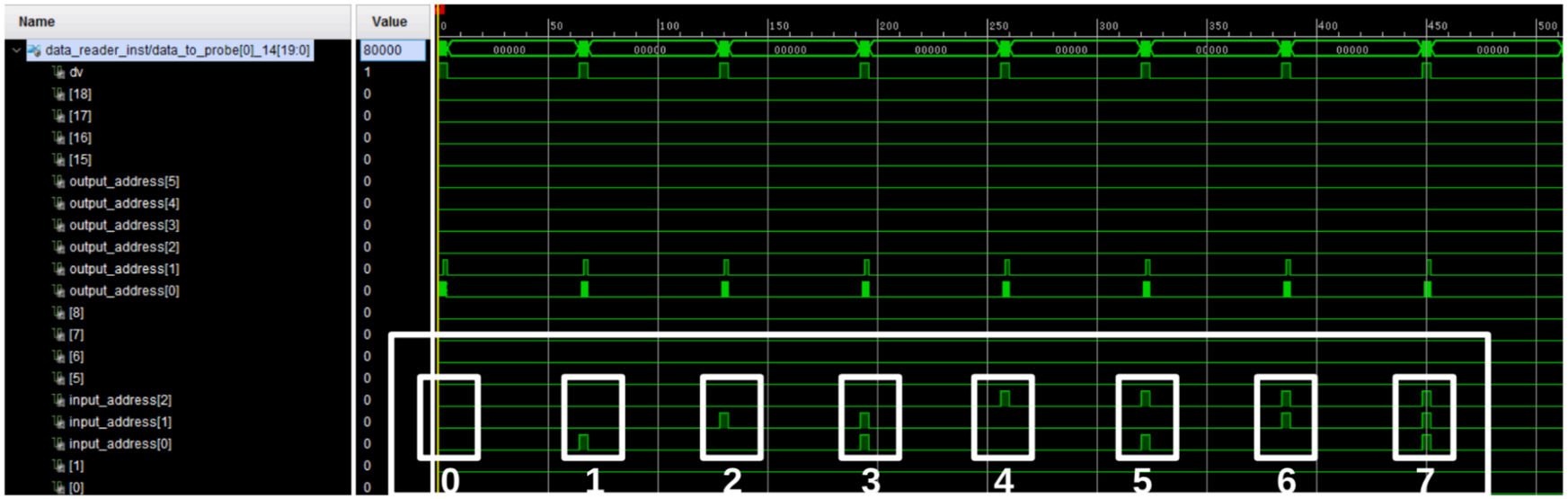
N. Neri *et al.*, Nucl. Instrum. Meth. A 845, 607-611 (2017)





Work in progress

- Test with simulated data on latest generation FPGA
- simulate response of a sector of a realistic detector (e.g. LHCb VELO at high luminosity) and evaluate the performance of the system (throughput, efficiency, latency, etc.)

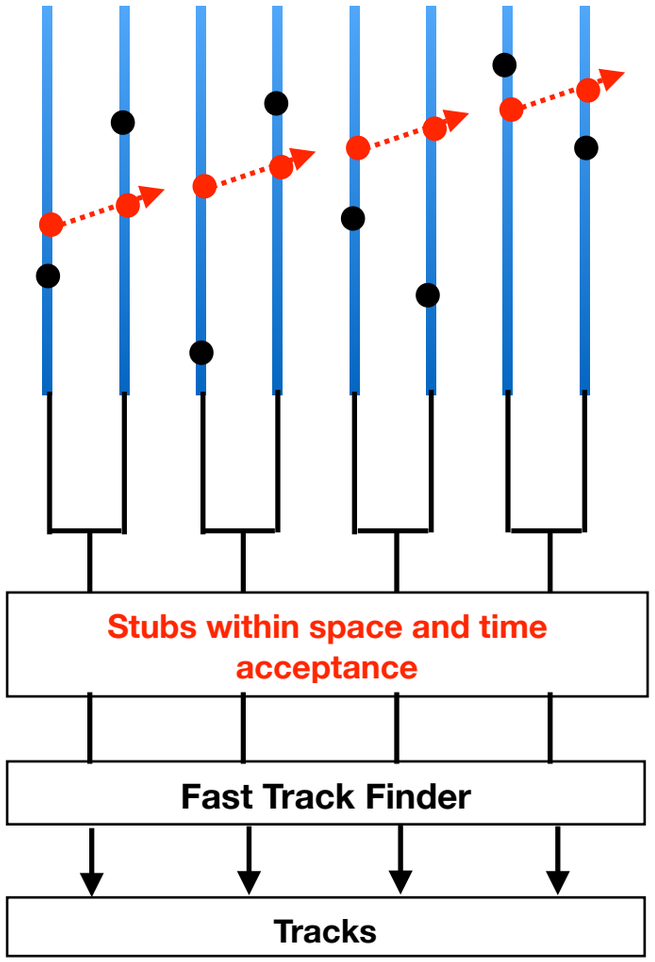


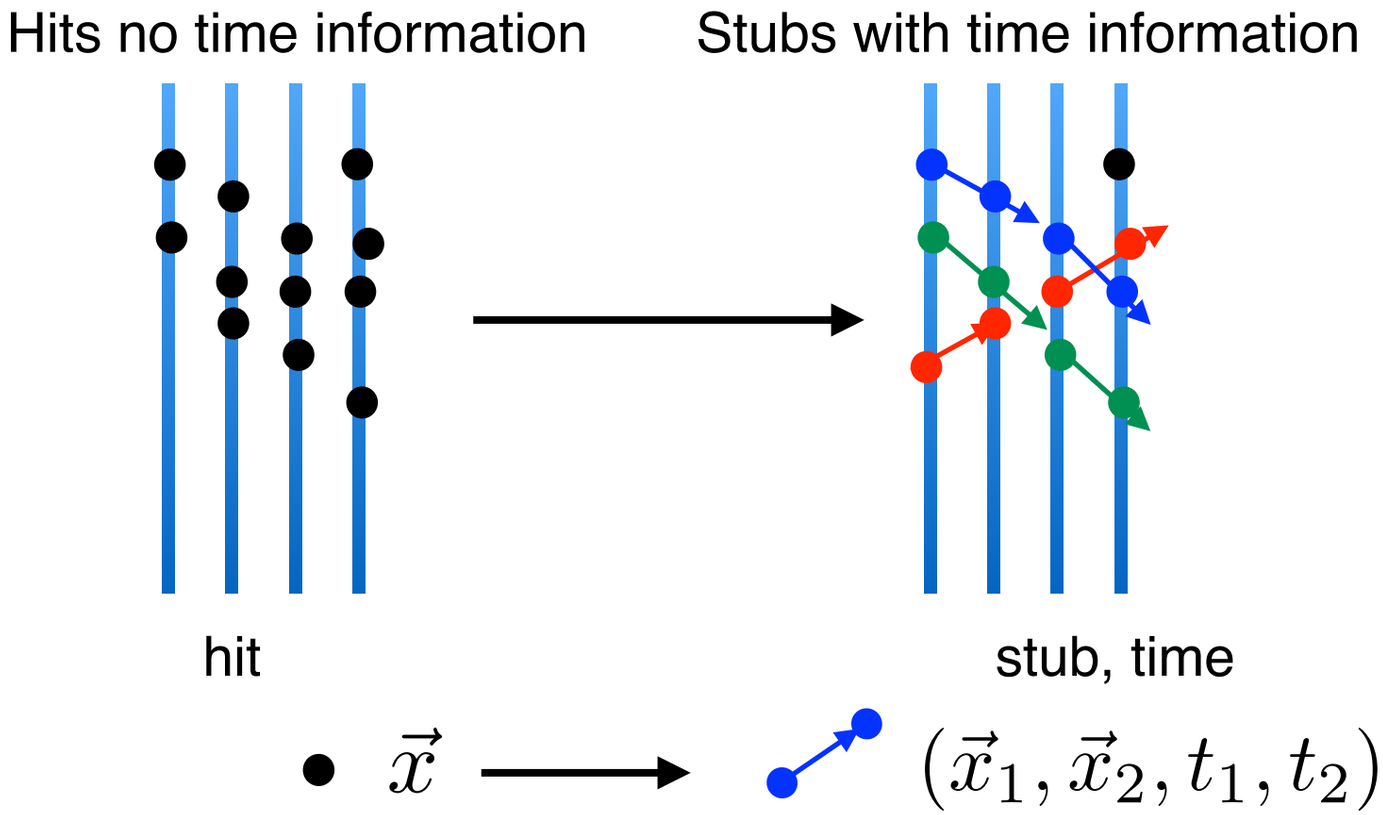
- test on Xilinx Virtex UltraScale xcvu095-ffvc2104-3-e
- switch proved to work at 480 MHz clock speed

4D real-time tracking

4D fast tracking pixel detector proposal

- Rad-hard pixel detector with precise space and time information for 4D tracking
- Detectors with embedded tracking capabilities providing track segments “stubs”





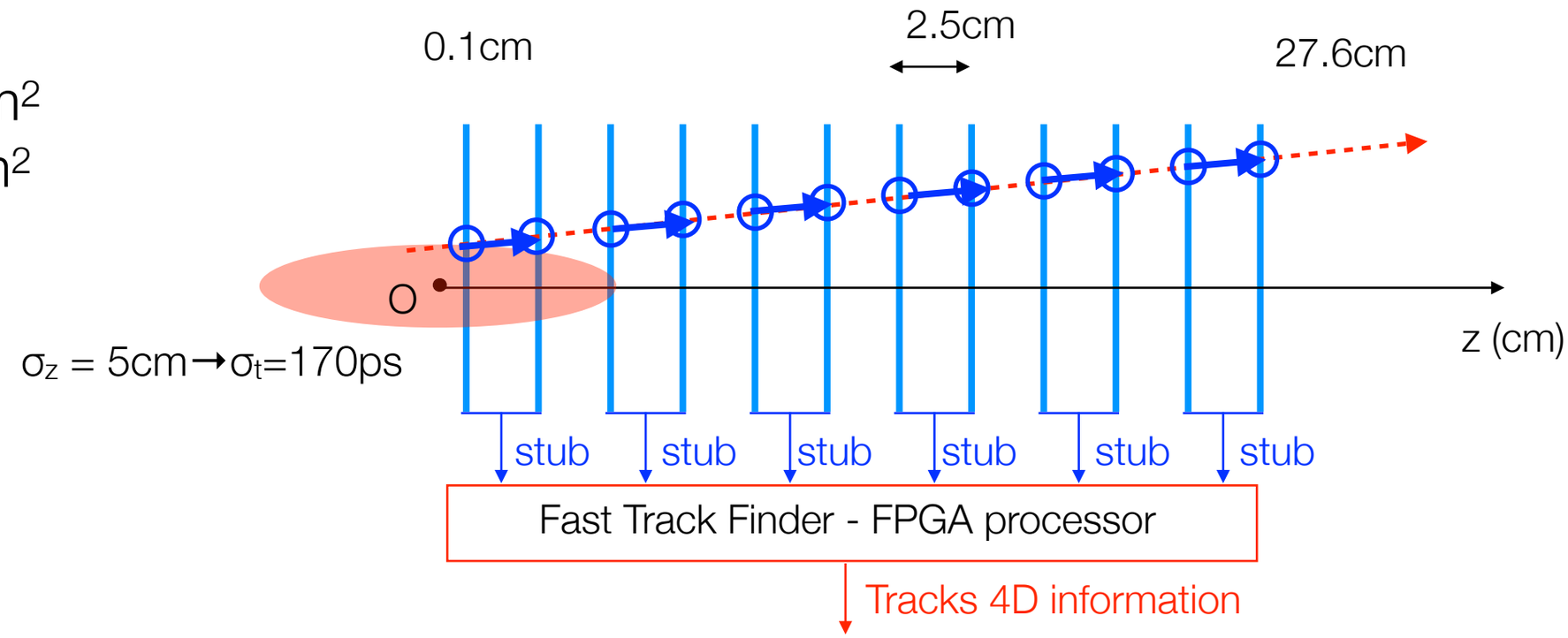
- Devices with embedded tracking capabilities in real time. Relieve workload of online trigger (save CPU time)
- Real-time track reconstruction can help selecting events and reducing data size (save bandwidth and disk space)

4D fast tracking simulations

N. Neri *et al.*, JINST 11 (2016) no.11, C11040

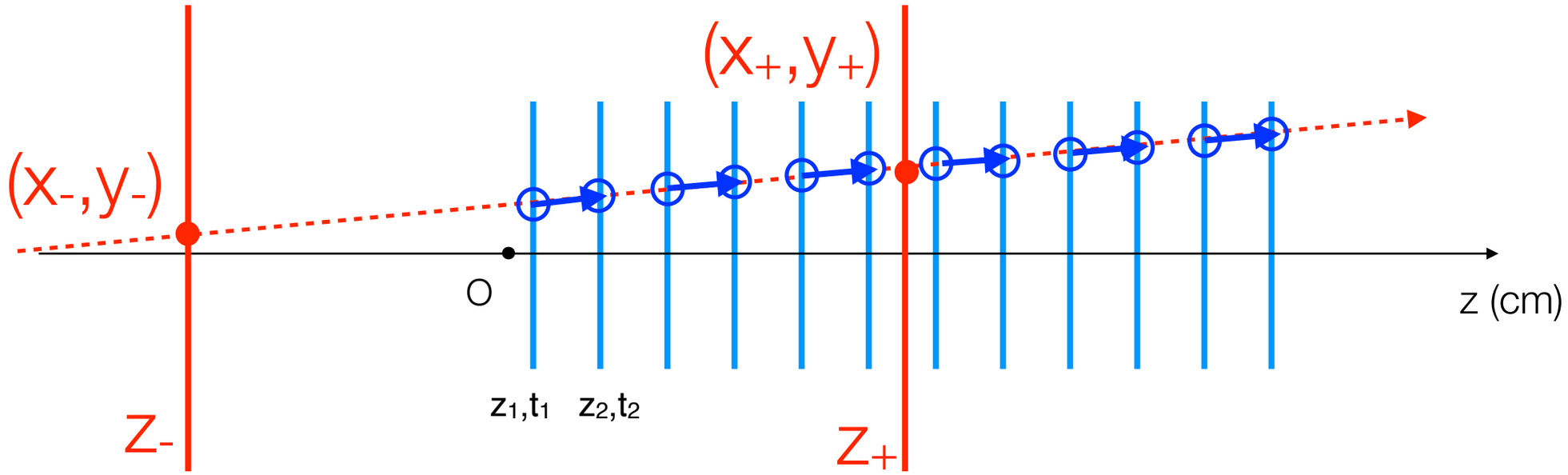
- Simple case: 12 layer telescope in forward region
- At lumi=$10^{34}cm^{-2}s^{-1}$: pileup~40 and ~1200 tracks/event

Sensor area = 6x6cm^2
pixel size = 55x55μm^2
thickness = 200 μm
time res σ_t=30 ps



- Reconstruct stubs at early stage, e.g. in FPGA: apply acceptance cuts (time, direction), reduce data flow and simplify track reconstruction
- Provide stubs in input to Fast Track Finder FPGA processor for real-time 4D track reconstruction using space and time information

Stub based fast tracking approach



- Stubs are projected to reference planes (in red) and a (x_-,y_-), (x_+,y_+) pair identifies a 3D track. No further processing required, see CERN-LHCb-PUB-026
- Time of the stub: $t_{stub}=(t_1+t_2)/2$, velocity $v = \frac{t_2 - t_1}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}}$
- Select stubs within space and time acceptance

$$\left|\frac{dx}{dz}\right|, \left|\frac{dy}{dz}\right| < 0.3 \quad \mathrm{IP}_{xy} < 1\mathrm{mm} \quad |z| < 10\mathrm{cm} \quad |v/c - 1| \leq 4\sqrt{2}\sigma_t c/\Delta z$$

- Measure 3D track parameters and time of the track at the origin, t_0

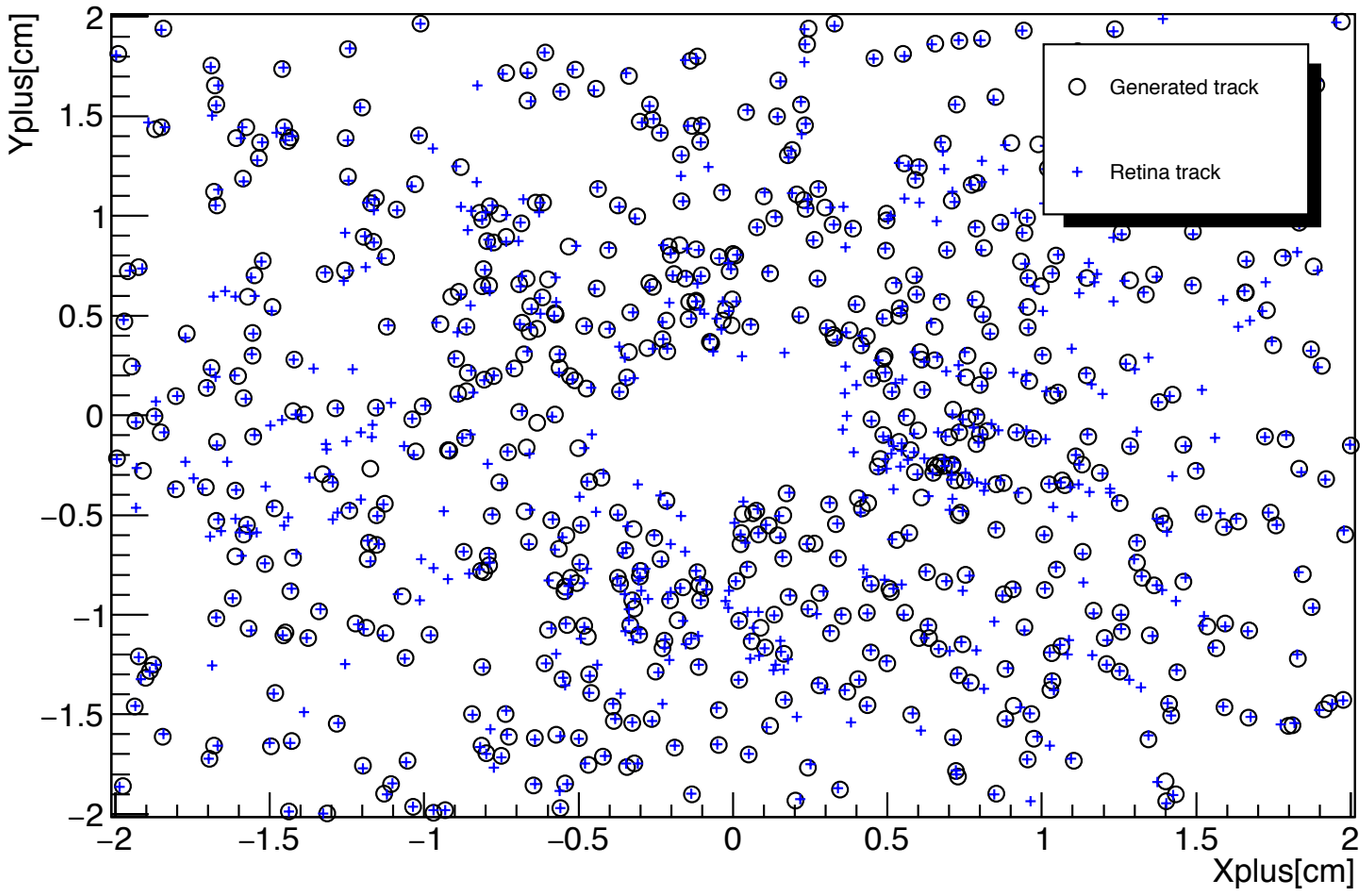
4D fast track finding algorithm

- Artificial retina algorithm adjusted to use stubs info:

$$W = \sum_{all\ \ stubs}^{N} \exp\left(-\frac{s_i^2}{2\sigma^2}\right) \qquad W_N = W/N$$

$$s_i = |(x_+, y_+)_{\mathrm{stub},i} - (x_+, y_+)_{\mathrm{engine}}|$$

Distance of the stub projection from the track receptor in the reference plane (x_+,y_+)



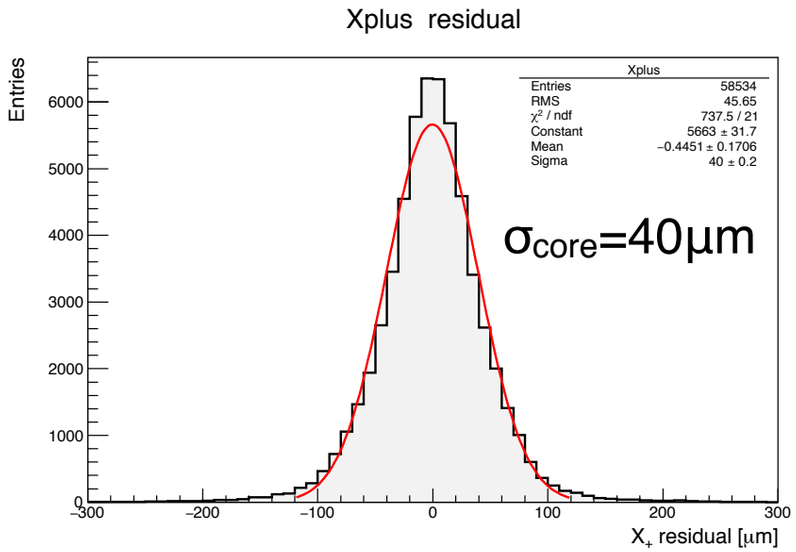
A local maximum of the artificial retina response W, identifies a track

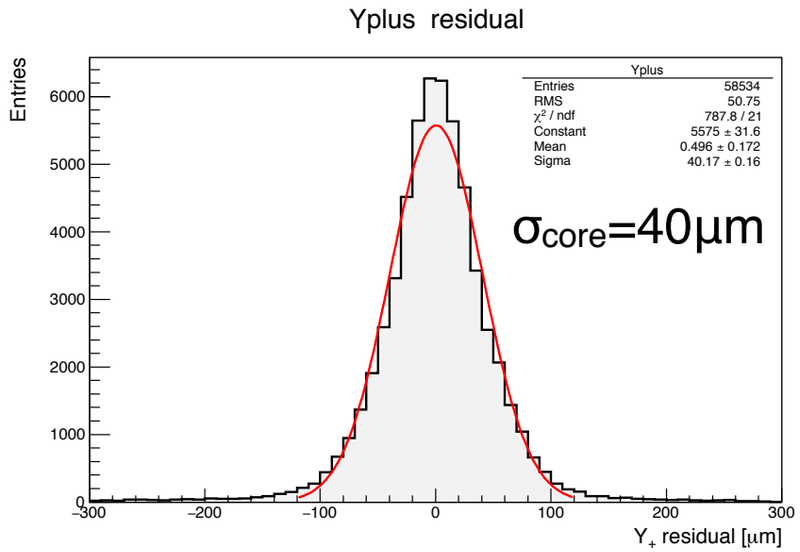
x_+,y_+ track parameters are calculated by interpolation of W values around maximum response

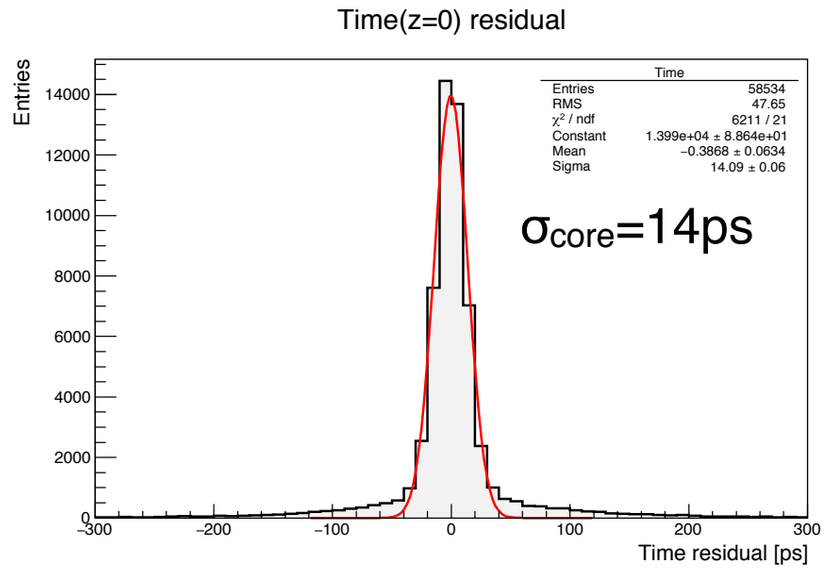
x_-,y_- track parameters and time of the track at the origin t_0 are determined by the engine with maximum W value

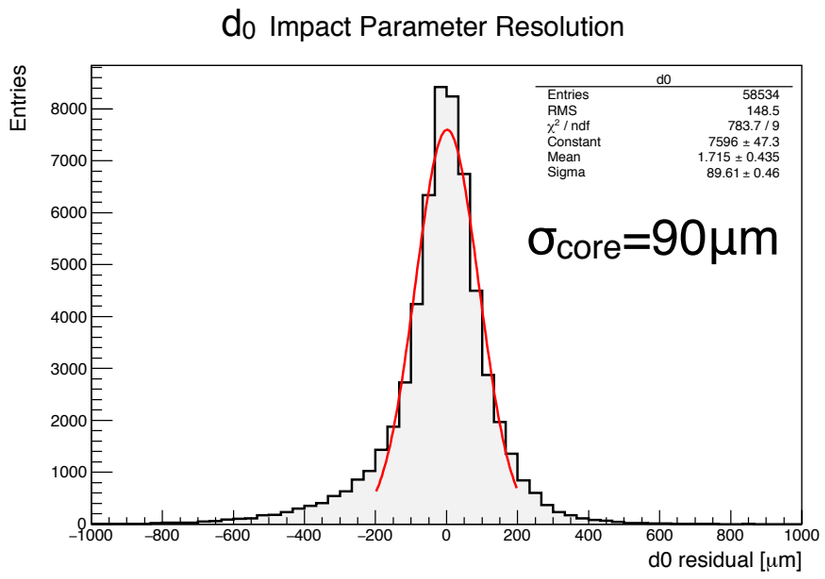
Resolution on track parameters

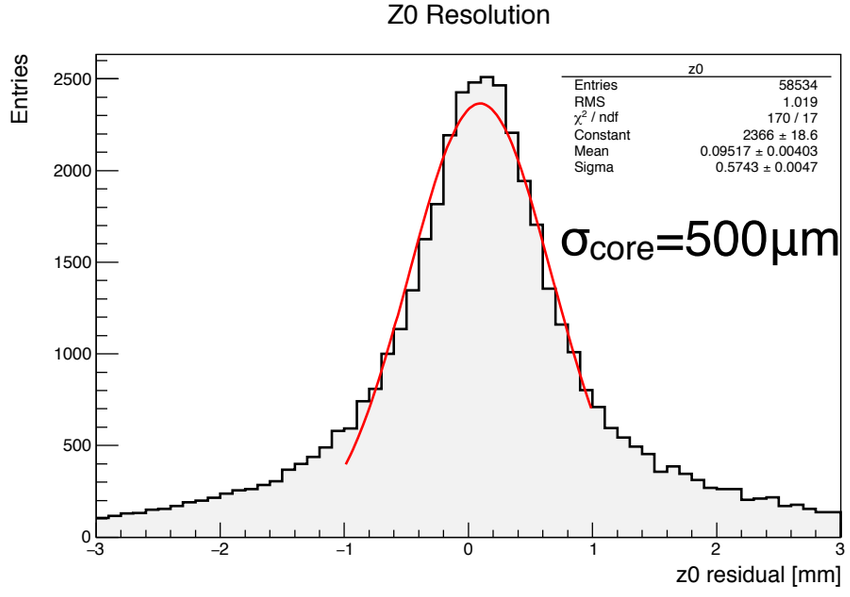
- Events with 40 pp interactions, 1200 tracks per event











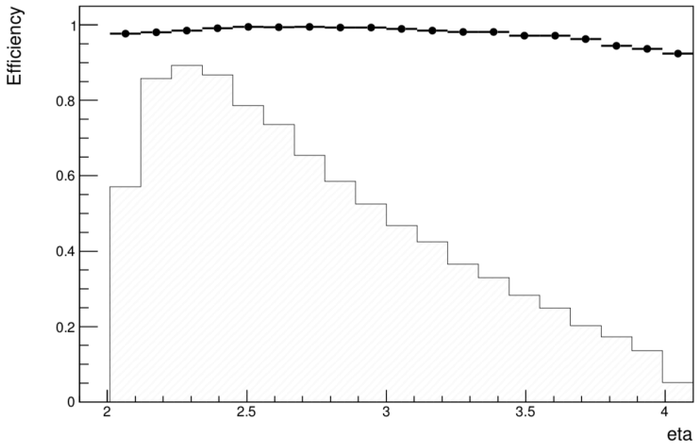
Online track reconstruction efficiency

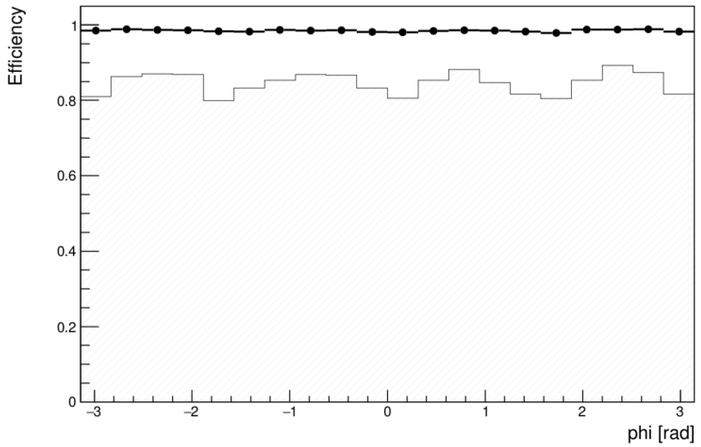
- Track efficiency vs track parameters: d_0, z_0, ϕ, η

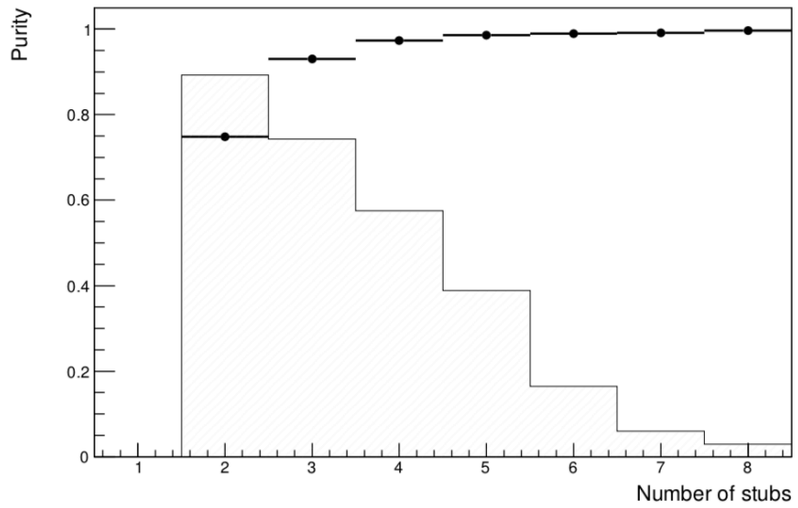
  d_0, z_0 (distance of closest approach to z axis), ϕ azimuthal angle, η=pseudo rapidity

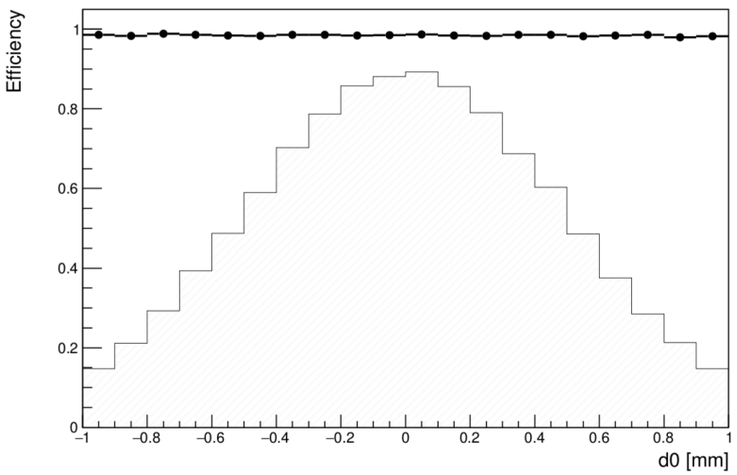
NN, M. Petruzzo in preparation

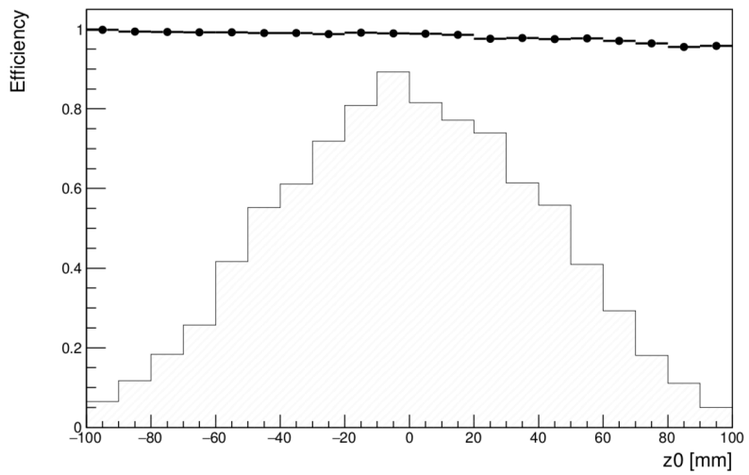
- Track efficiency ~99% and track purity >80% with 1200 tracks per event → track purity ~60% (without time information of the hit)







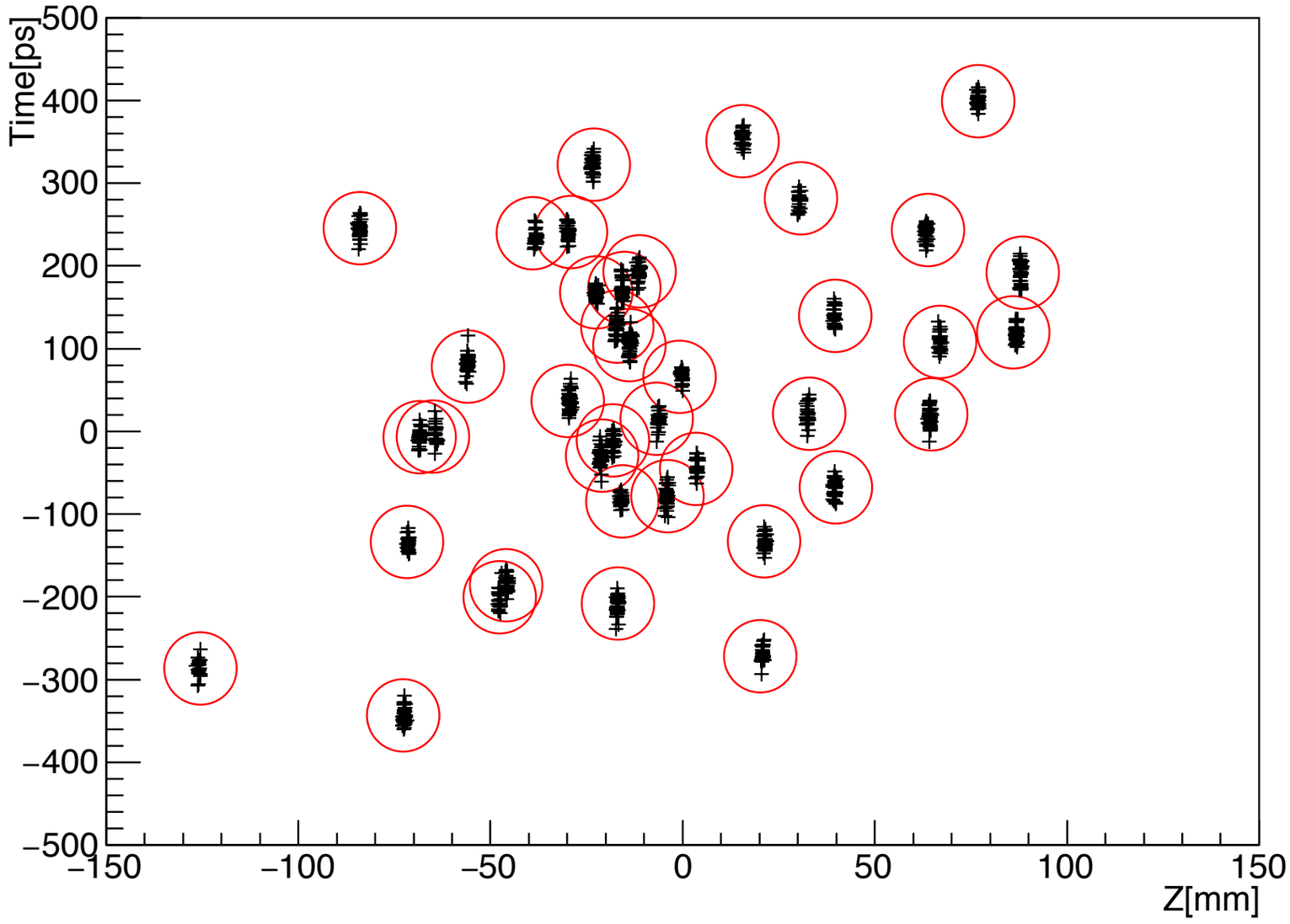




- With 2400 tracks/event, online track efficiency 98,7% and purity 60%
- System is not optimised, room for improve track purity

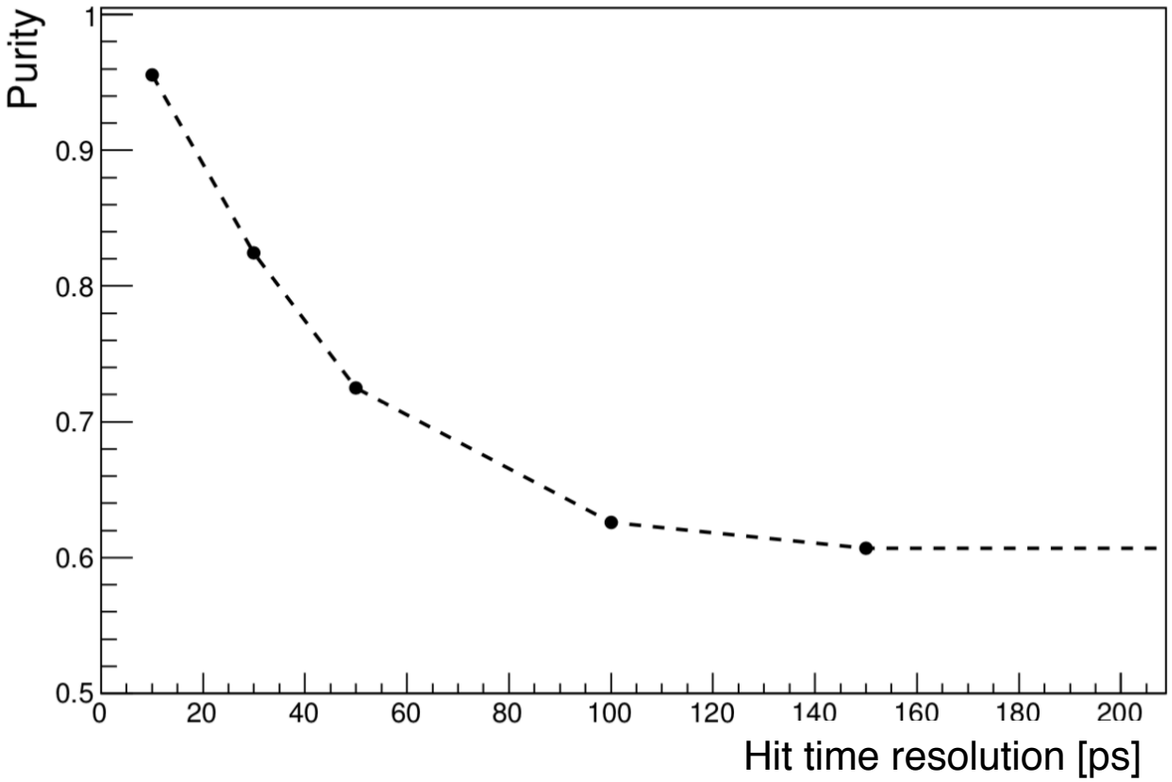
Track association to primary vertex

- Separate tracks in space and time: improved association of tracks to primary vertex
- Track mis-association >10% (no time information) → <1% using precise time information of the hit in offline reconstruction

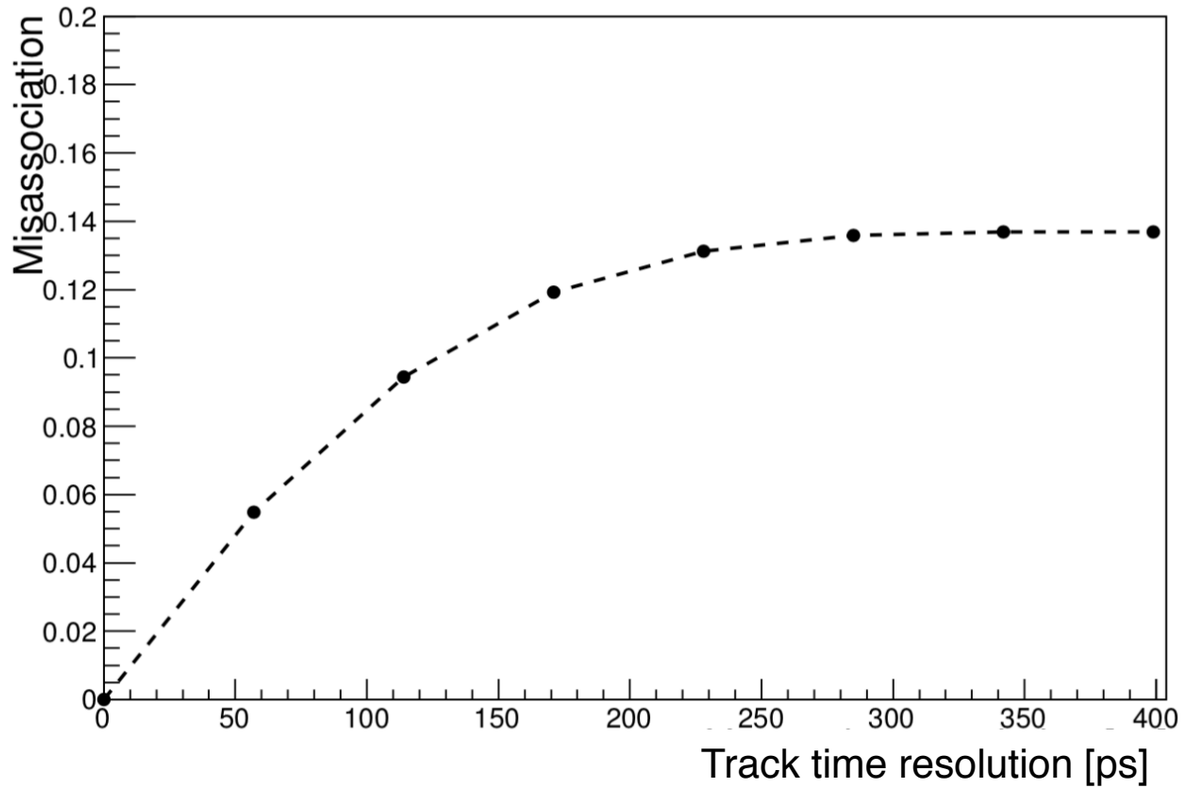


Track purity and vertex mis-association

Purity of real-time reconstructed track



Fraction of mis-associated track to primary vertex



- Clear improvement of performance with time resolution

WP4 activity

People

WP4: Design and implementarion of fast tracking devices				
Resp: N. Neri (MI)		**FT**	**m/pers 018**	**Activity in 2018**
People	**RU**			
A. Carbone	BO	0,2	2,2	design
A. Sidoti	BO	0,2	2,2	design
L. Frontini	MI	0,3	3,3	design/tests
P. Gandini	MI	0,2	3,3	design/tests
N. Neri	MI	0,3	3,3	design/tests
M. Petruzzo	Mi	0,3	2,2	design/tests
Total		**1,5**	**18,7**	

- Preliminary meeting 23 November in Milano
- PhD student (Elisabetta Spadaro Norella, MI) + 1 PostDoc (Jinlin Fu, MI) + INFN staff (Stefano Perazzini, BO) recently joined the WP

Tasks

1. Optimisation of the 4D fast tracking algorithm and tracking detector layout based on high-level simulations (6 months) [Nicola, Marco, Elisabetta, Paolo, altri?]
2. Design of fast track finding device architecture based on a realistic tracking detector and running conditions of the HL-LHC (6 months) [Nicola, Marco, Stefano, Paolo, altri?]
3. Study of the performance of the device based on low-level simulations (6 months) [Marco, Elisabetta, Paolo, altri?]
4. Implementation of the firmware in fast track finding boards based on Bologna (ATLAS) and Milano (Retina) boards. Four boards developed in Bologna will emulate hits of the detector and perform stub reconstruction. The Retina board will receive the stubs and perform the track reconstruction (6 months) [Marco, Elisabetta, Paolo, Stefano, altri?]
5. Emulation of a realist tracking detector at HL-LHC and study of the performance of the device based on simulated data on optimise board to be designed and constructed in Timespot (6 months) [Marco, Elisabetta, Paolo, altri?]
6. Study the performance of the fast track finding device using a multilayer pixel prototype in a test beam (6 months) [All]

Milestones & deliverables

Milestones

1. Dec. 2018: document with the specifications for the fast track finding device for a realistic tracking detector for the HL-LHC phase
2. Sept. 2019: Test of the performance of the device using emulated tracks on existing boards developed in ATLAS and Retina experiments
3. June 2020: Test of the performance of the device using a emulated data using an optimised board to be developed in Timespot
4. Dec. 2020: Test of the performance of the device on beam using a multilayer pixel detector prototype

Deliverables

1. Dec. 2018: document of specifications for fast track finding device
2. Dec. 2020: Prototype optimised fast track finding board

Timeline

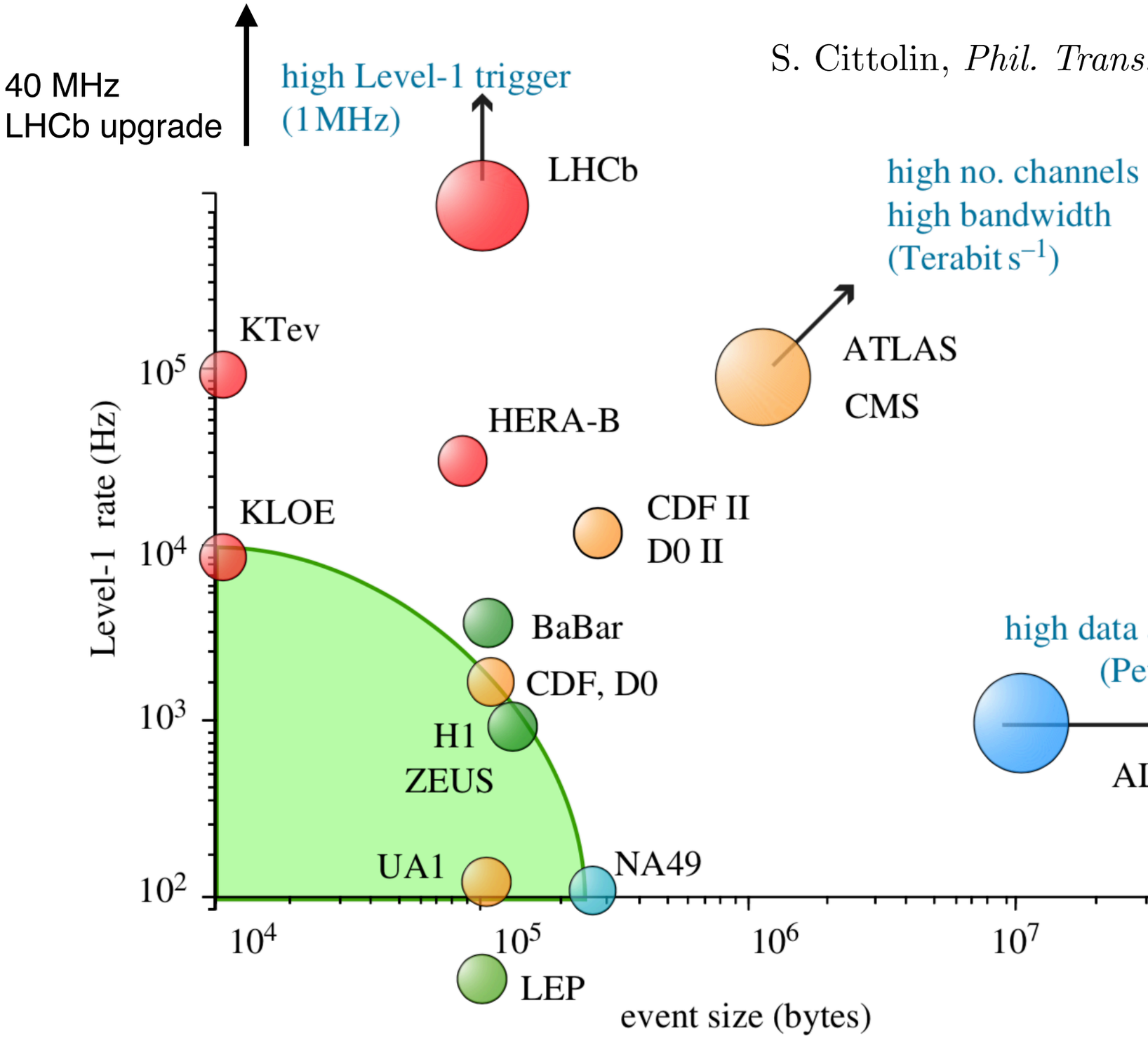
Timeline

Task	2018	2019	2020
Optimisation 4D algorithm and detector layout			
Device architecture			
Low-level simulations			
Test using simulated tracks with existing DAQ and Retina boards			
Test using simulated tracks at high rates with optimised board and track processor			
Integration with the prototype pixel detector and test on beam			

Backup slides

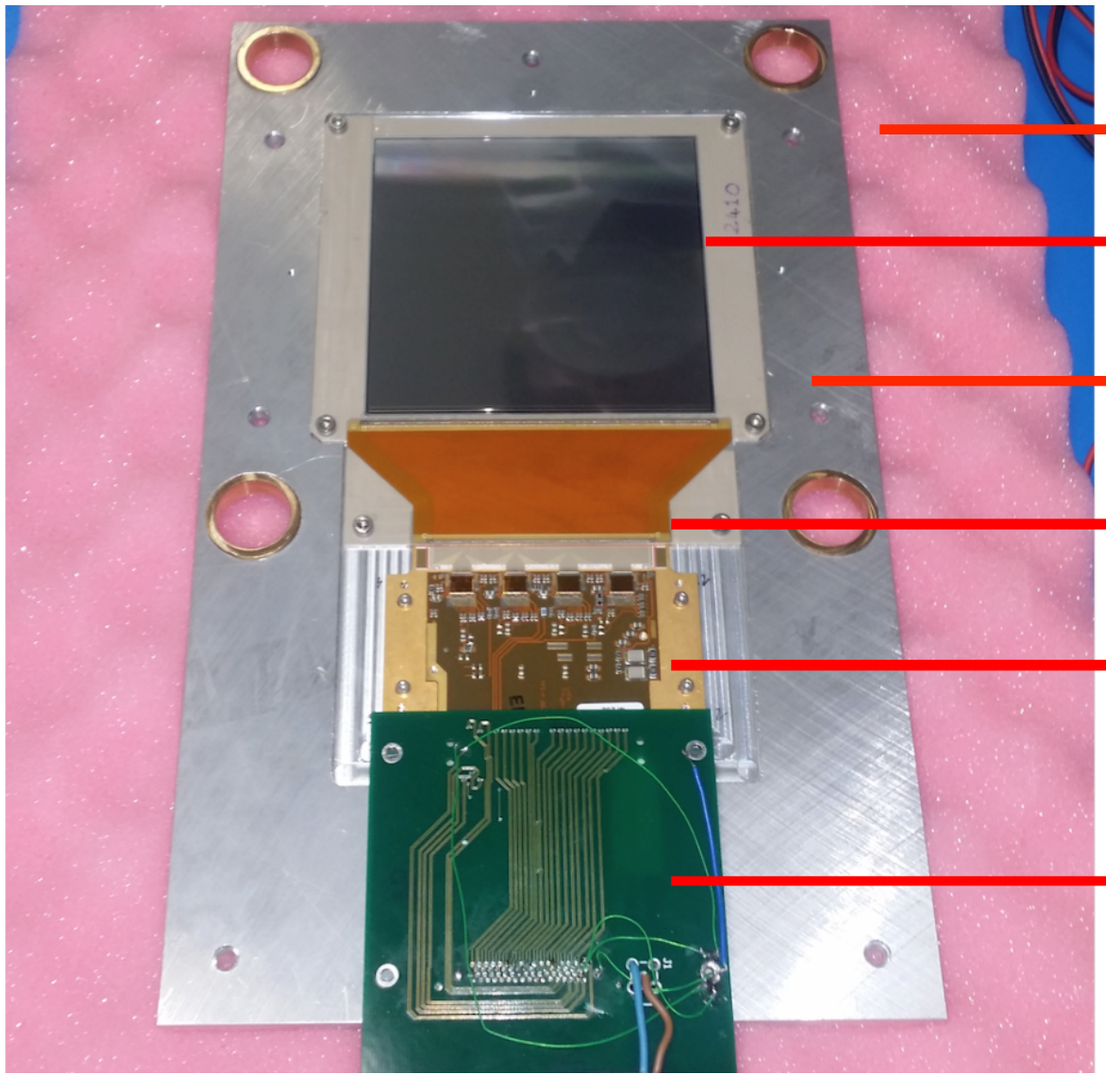
Summary of TDAQ in HEP

S. Cittolin, *Phil. Trans. R. Soc. A* (2012) **370**, 950–964



Telescope module

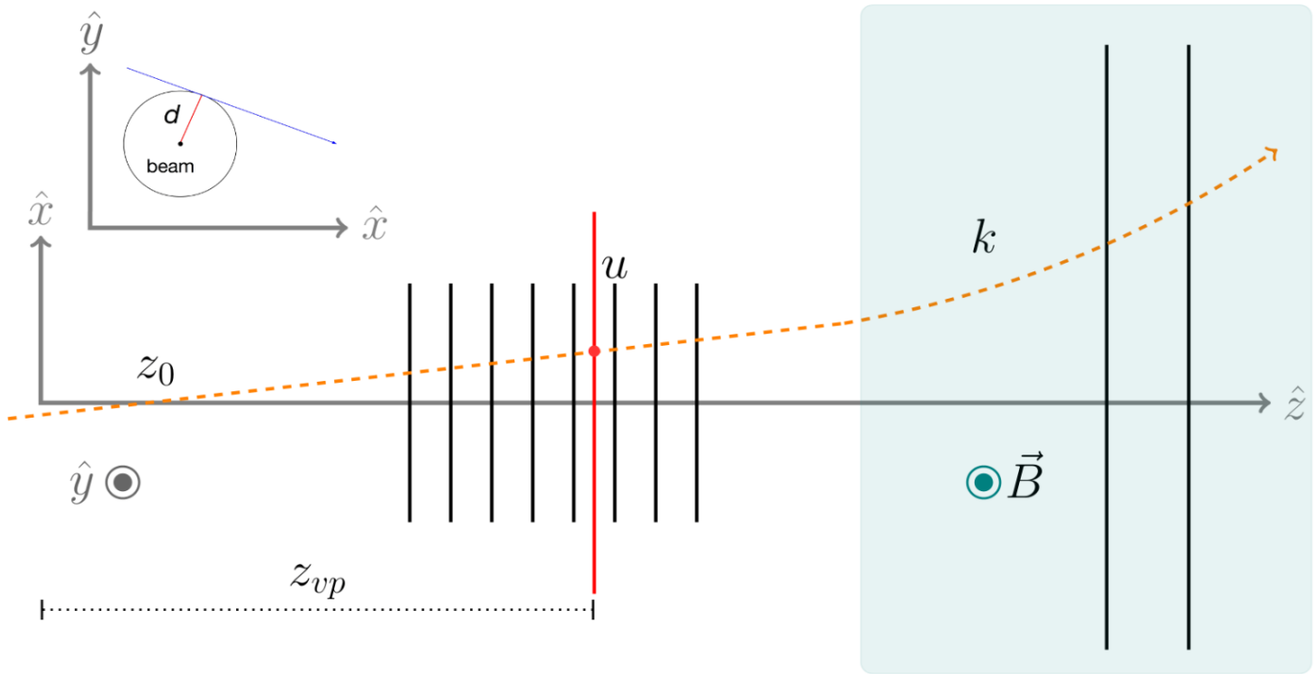
- Single-sided silicon sensors:
  - OB2 STM *p*-in-*n* sensor, 10 cmx10 cm active area
  - 512 strips, 183 μm pitch, 500 μm thickness



Feasibility study for LHC experiments

- The Retina architecture is modular, parallel processing units are scalable. Using adequate FPGA resources can cope with high particle rates and large detectors, e.g. 40 MHz event rate and 300 tracks/event of LHC.
- Delivers 3D tracks with offline-like quality at 40 MHz with <1μs latency
- Case study for the LHCb upgrade simulated and documented here: LHCb-PUB-2014-026, JINST 9 C09001 (2014). Affordable resources and cost (50,000 cells ~ 50 FPGA)

Application in forward spectrometer experiment



- 50 mrad acceptance
  - *O*(100) particles/event
- 8 pixel layers
- 2 silicon strip layers
- ~0.05 T magnetic field
- Pileup: ~8 *pp* events