A Level-1 Tracking Trigger for the CMS Upgrade using stacked silicon strip detectors and advanced pattern recognition technologies

Gaëlle Boudoul

IPN-Lyon

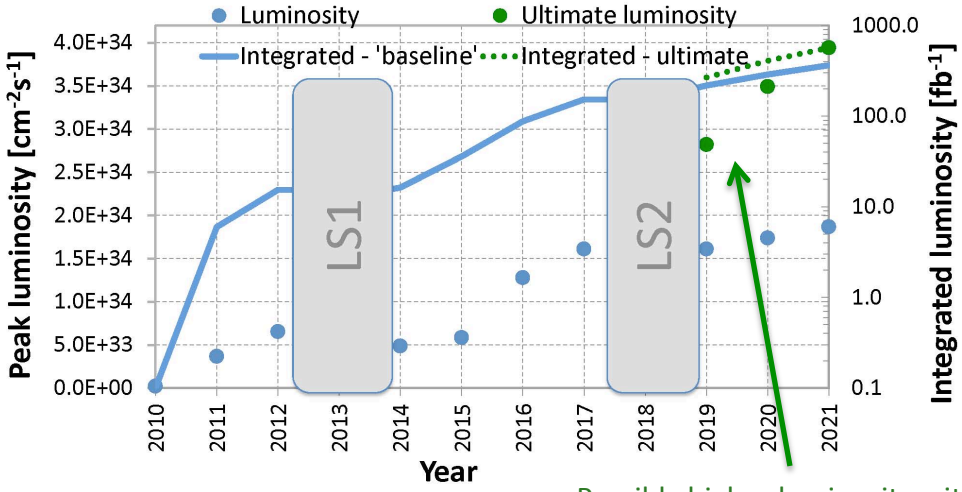
On behalf of the CMS Collaboration

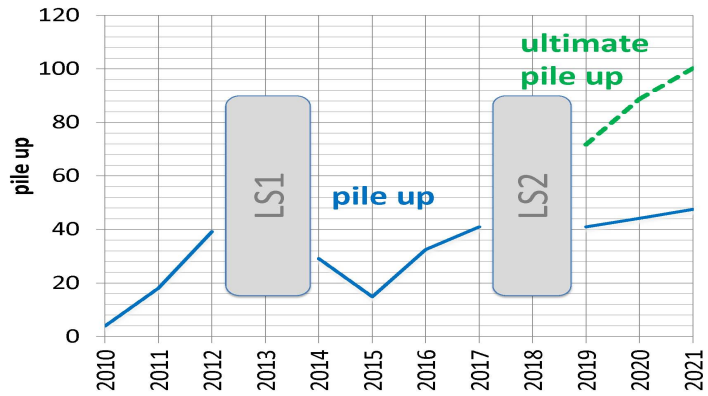
WIT2012 – Workshop on Intelligent Trackers (Pisa-Italy)

Outline

- Motivation for upgrade
- Motivation for L1 Tracking Trigger for CMS Tracker Upgrade
- Tracker Reduced Readout
- Tracker Layouts and L1 Tracking concepts
- Performance and Triggering capabilities: First Results on Barrel-Endcap Geometry
- Conclusions

Motivation for upgrade: LHC luminosity and Pile-up





M. Zimmerman Fermilab workshop Nov 2011

Possibly higher luminosity with 50ns bunch spacing after injection upgrade

2022: Long Shutdown 3 → High luminosity project (HL-LHC)

Prepare for operation at 5.10^{34} cm^{-2} sec^{-1}

Prepare for an integrated luminosity of $3000fb^{-1}$ ($200\ fb^{-1}$ to $300\ fb^{-1}$ /yr)

50 vs 25ns operation offers flexibility, but pile-up affect performance and upgrade detector designs,

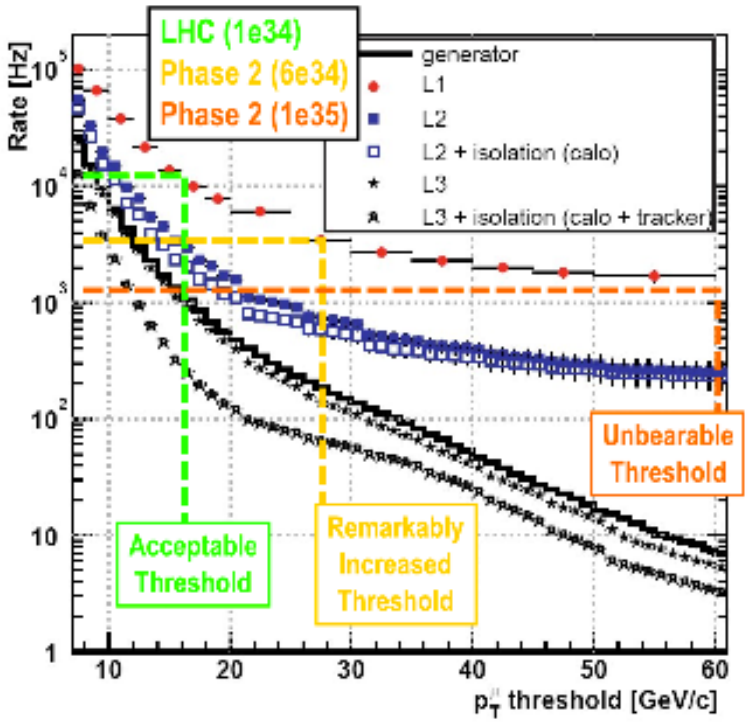
200 PU in phase2 if 5 10^{*34} Hz/cm^2 at 50ns spacing

Increase of the LHC luminosity well above 10^{34} $cm^{-2}s^{-1}$ requires subtancial upgrade of the CMS Tracking:

- Physics potential require full performance detector
- Higher radiation resistance (instantaneous and integrated level)
- Higher readout granularity (keep channel occupancy at an adequate level)
- Ability to contribute information for the L1 trigger

Motivation for a L1 Tracking Trigger

- Trigger system needs to maintain an output rate at 100kHz (despite the factor 5 increase in luminosity and the 5 to 10 times larger pile-up)
  - Impossible to achieve using only information from tl detectors
  - Flattening of the L1 rates as a function of the pt
    - Increasing the thresholds will not help



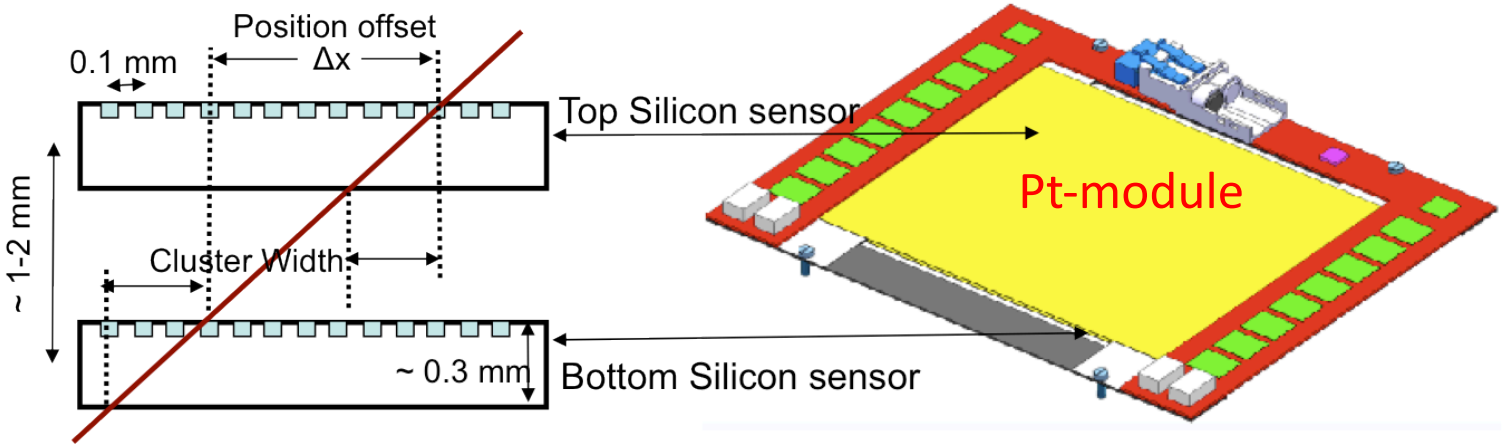
- The tracker data will have to provide trigger tracks used at the L1 stage to allow:
  - Precise transverse momentum measurement of muons for application of a sharp threshold around 10-20 GeV/c.
  - Matching and isolation for electron and tau identification and veto for photons.
  - Association to proper interaction vertices to reduce rate of accidentals due to pile-up.

Tracker-Trigger: how to reduce data volume

- About 90% of the charged particles produced in minimumBias collisions have Pt<1GeV/c^2 and ~97% have Pt < 2GeV/c^2

  ➔**Reduced readout of hits associated with tracks above ~2 GeV/c^2**

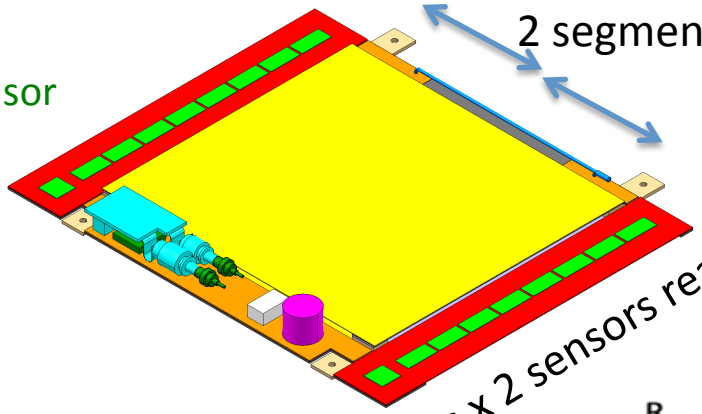
- Two complementary steps to reduce data volume (both based on the track deviation in the plan perpendicular to the CMS magnetic field):
  - **STEP1: Cluster width approach**
    - Preselection of hits according to their cluster width (CW)
      - CW is proportional to the radial distance of the sensor from the IP and inversly proportional to the Pt
  - **STEP2: Stacked tracker**: correlation between preselected hits in nearby sensors
    - Exploit track direction of flight measurement to reconstruct 'track stubs' above a given Pt threshold.

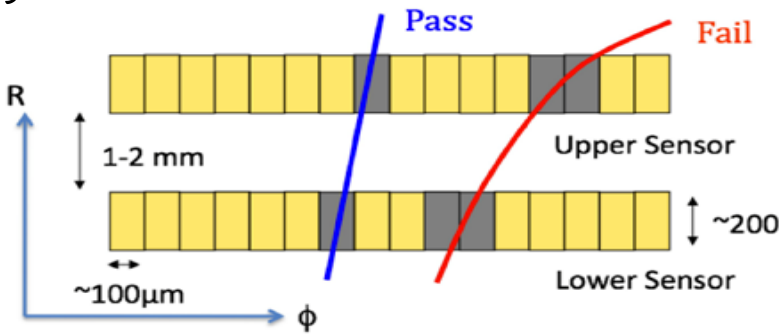


Detector Modules

2Strip Pt Module

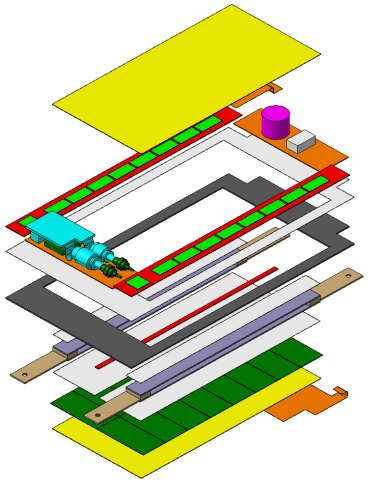
- 2 closely space strip sensor
- Pitch 90 µm
- Strip length:46.3 mm
- Pt information (trigger)





Pixel+Strip Pt module

- Pitch 99 µm
- Strip length 21 mm
- Pt information (trigger)
- z information



correlate selected hits in two closely separated sensors to discriminate between high and low Pt tracks

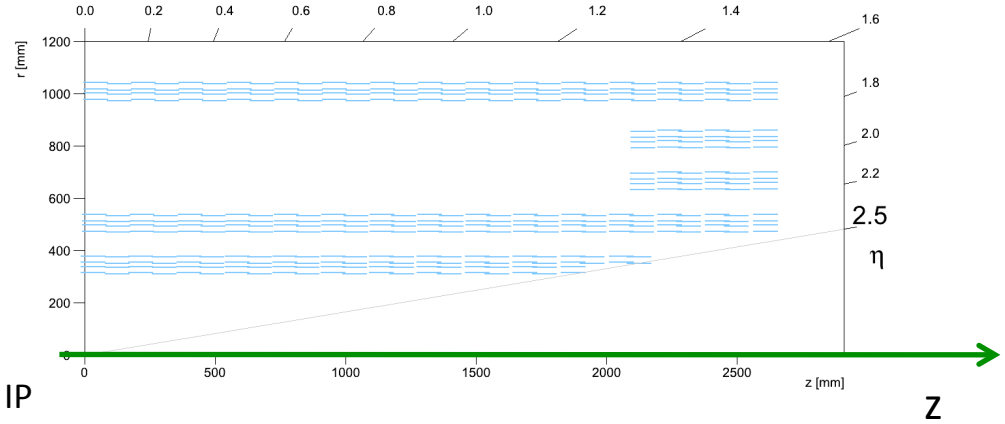
Two complementary Tracker layouts
I) Long Barrel geometry

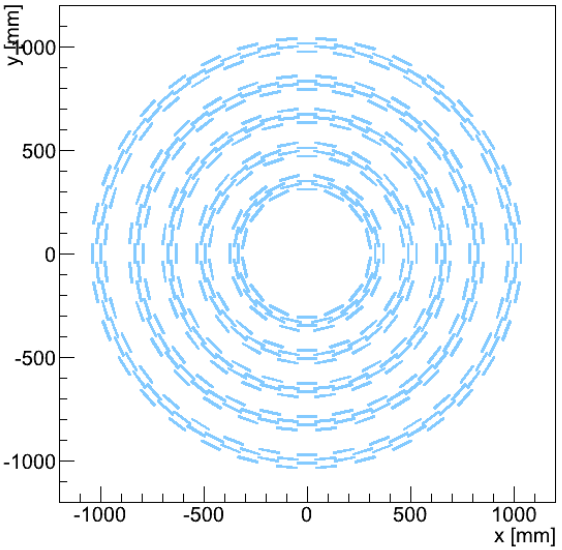
Dedicated talk (E. Salvati) this conference

Long closely spaced layers of Pt Modules arranged in Double Stacks separated by ~4 cm

(Full long Barrel outer tracker of stacked triggering layers)

Cluster positions in z direction provided (pixelated modules).



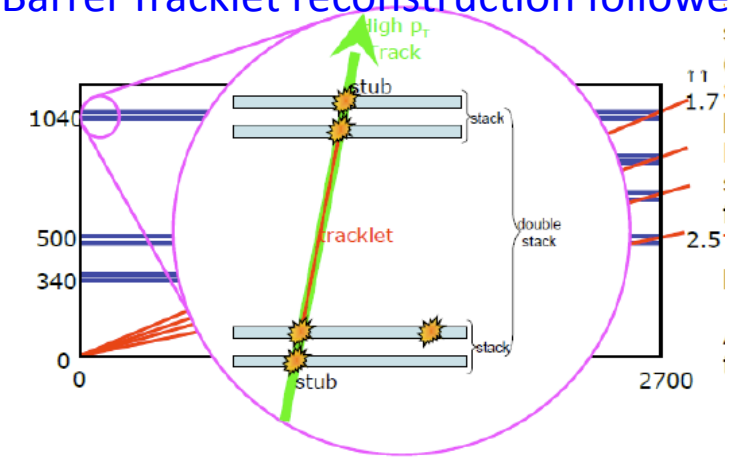


XY section of the barrel

From Hits to L1 Tracks:

Super Layers => Stubs -> Tracklets -> L1 Tracks

Long Barrel Tracklet reconstruction followed by propagation to next super layer



L1 tracking precision potential

pT resolution 2% @ 10 GeV in forward

Tracking precision

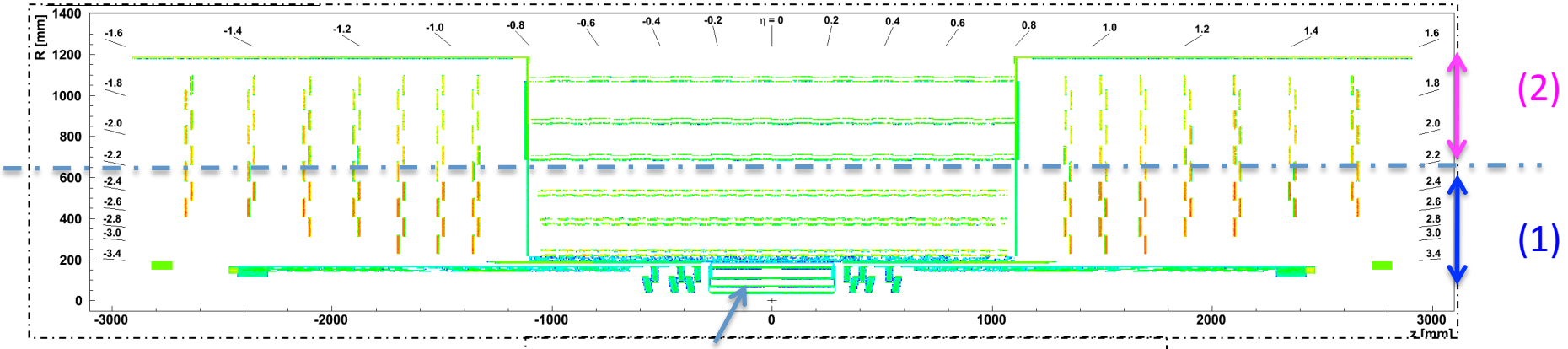
pT resolution 2.1% @ 10 GeV

pT resolution 5% @ 100 GeV

Two complementary Tracker layouts
II) Barrel-EndCap Geometry

This talk

The "Barrel-EndCap" design comprises 6 barrel layers and 7 endcap disks composed of rings.



Inner pixel Detector : 4 layers- 3disks

The inner part (1) is populated by Pixel-Strip stacked (PS) modules .

The outer part (2) is populated by Strip stacked (2S) modules.

The number of endcap disk is optimized for tracking performance.

Different spacings between the two sensors of the Pt modules:
0.8mm in the outer barrel (2S)
1.6 and 2.6mm in the inner barrel (PS)
4.0, 2.6 and 1.2mm in the outer end-cap (2S)
4.0mm in the inner end-cap (PS)

L1 tracking precision potential
- pT resolution 4% @ 10 GeV in forward

Tracking precision
- pT resolution 1.4% @ 10 GeV
- pT resolution 3% @ 100 GeV

From Hits to L1 Tracks (Barrel-EndCap geometry)

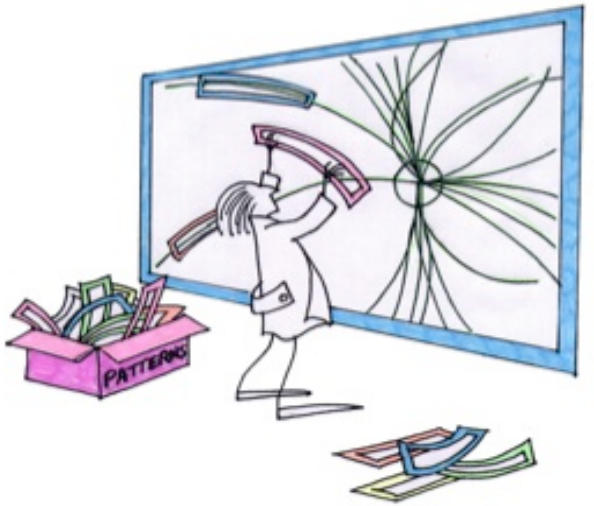
Independent Pt layers => Stubs -> L1 Tracks

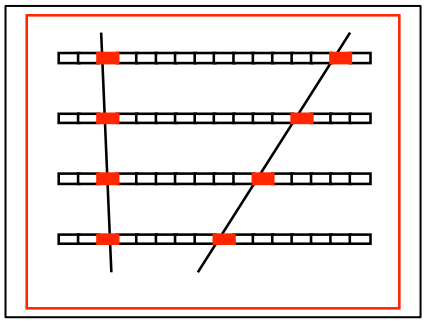
Barrel -EndCap track reconstruction from pattern comparison using Associative Memories

Pattern Matching using Associative memory
(M.Dell' Orso and L. Ristori: initial idea in 1985)

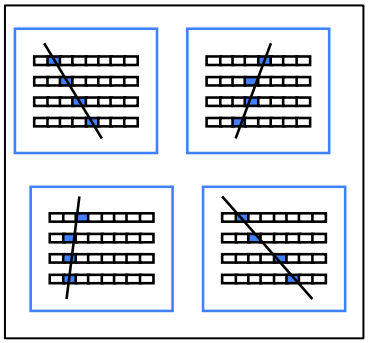
Hardware-based pattern recognition for fast track triggering has been very successful for HEP (CDF SVT)

Enormous challenges in implementing pattern recognition for tracking trigger at HL-LHC, due to much higher occupancy (pile up), event rates and high granularity

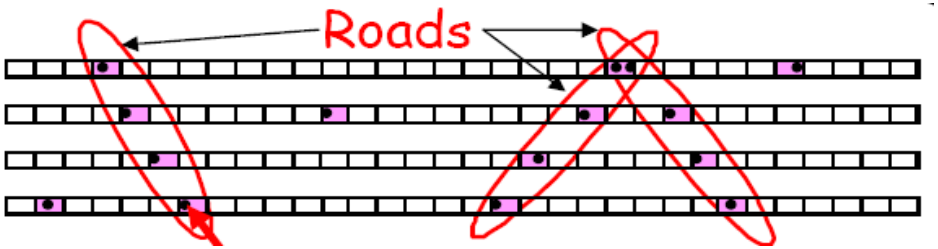




The Event



The Pattern Bank



Find low resolution track candidates first (roads)

Then perform track fitting downstream

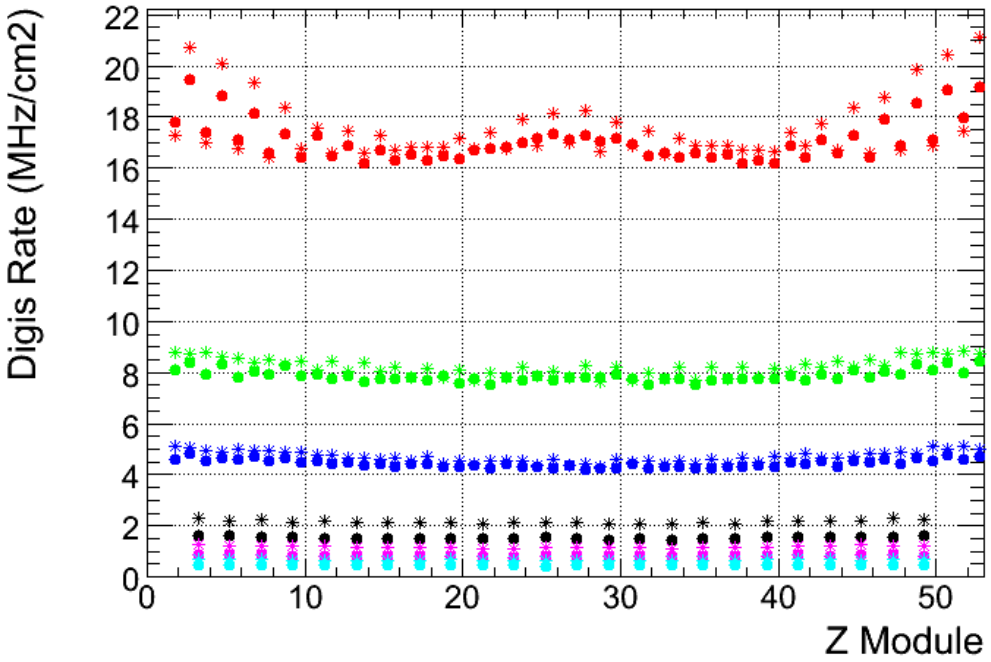
Performance and Triggering capabilities
First Results on Barrel-Endcap Geometry

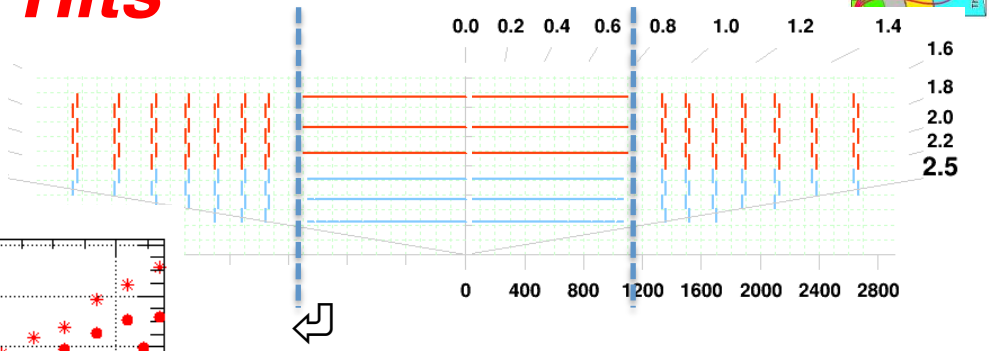
- Full Simulation of the Barrel-EndCap Geometry with Material Budget (**Only the BARREL part is considered for the moment**)

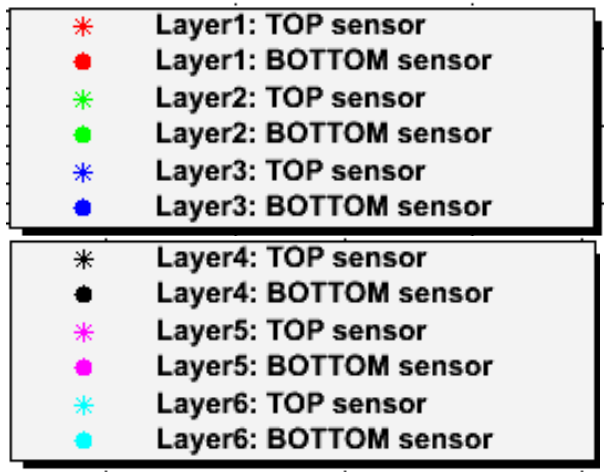
- MinimumBias sample @14TeV - 200 PU events - 50 ns bunch spacing (20MHz)

- First Results will be presented for
  - Digi hits
  - Clusters (selective readout step1)
  - Track Stubs (selective readout step2)
  - Turn-on Curves (efficiency)
  - Tracks at L1 using using Associative Memories

Digi Hits





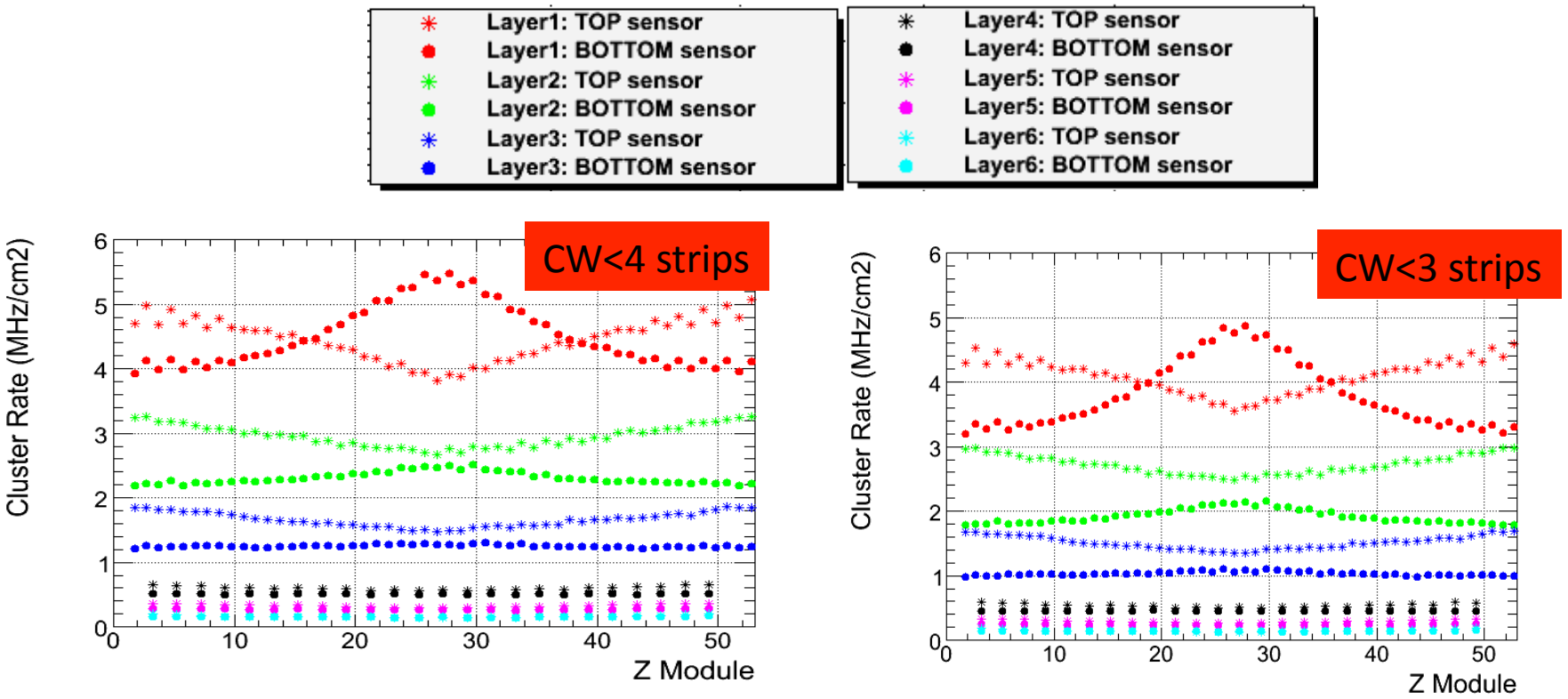


The rates (MHz/cm2) are given as a function of the position of the modules along z axis for bottom and top sensors separately.

Rates are decreasing with radius as expected

Rates on the top sensor are higher compared to the bottom rates -> Hits coming from secondaries particles produced in between.

Reduced Readout Step1: Cluster Rates

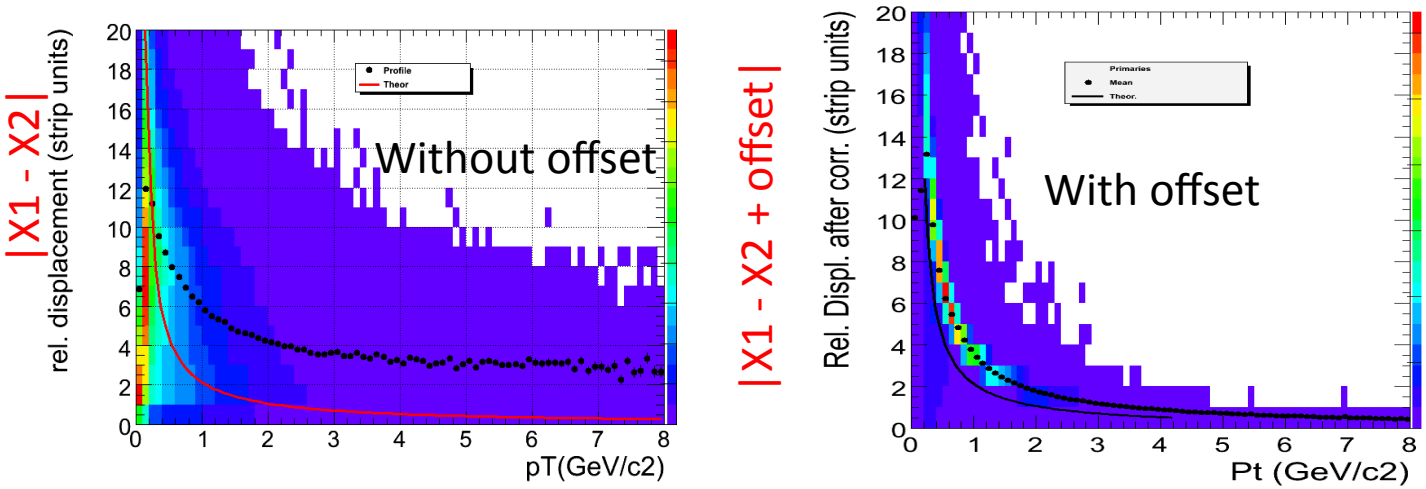


Two different cuts on cluster width (CW) lead to a significant impact on the initial cluster rates:
Cut1 : CW<4 -> cluster data reduction = 14-16% (depending on the layer)
Cut2 : CW<3 -> cluster data reduction = 27-28% (depending on the layer)

These cuts should be optimized to reduce the rates without affecting the final efficiency when applying the 2nd step (stub building)

Reduced Readout Step2: Track Stubs

- The position of the preselected clusters is defined by the center of the cluster with half strip resolution
- Correlated hits in nearby sensors are stubs (X1=position of the bottom sensor / X2 = position of the top sensor).
- Cluster Position Correction (offset) according to their location on the bottom sensor is needed in order to take into account the flatness of sensors:



Position correlation in a Pt-module

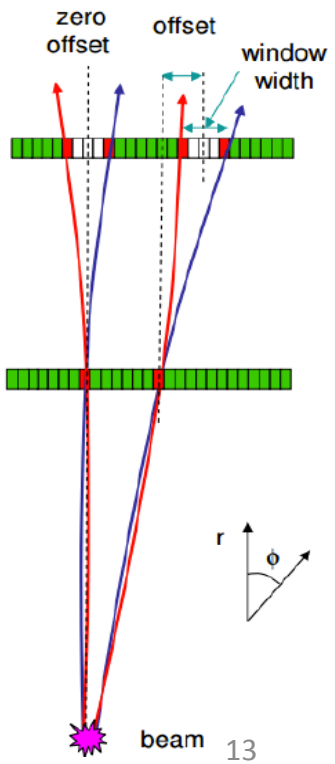
-For each cluster on the bottom sensor, a matching window centered on its position (+offset) is defined on the top sensor.

|X1 - X2 + offset| < Diffmax

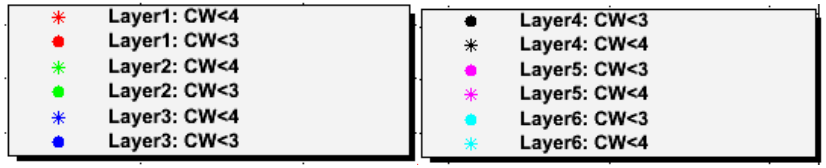
Diffmax corresponds to the threshold (window size according to the expected Pt cut

-A stub is produced if a cluster position is found in the window

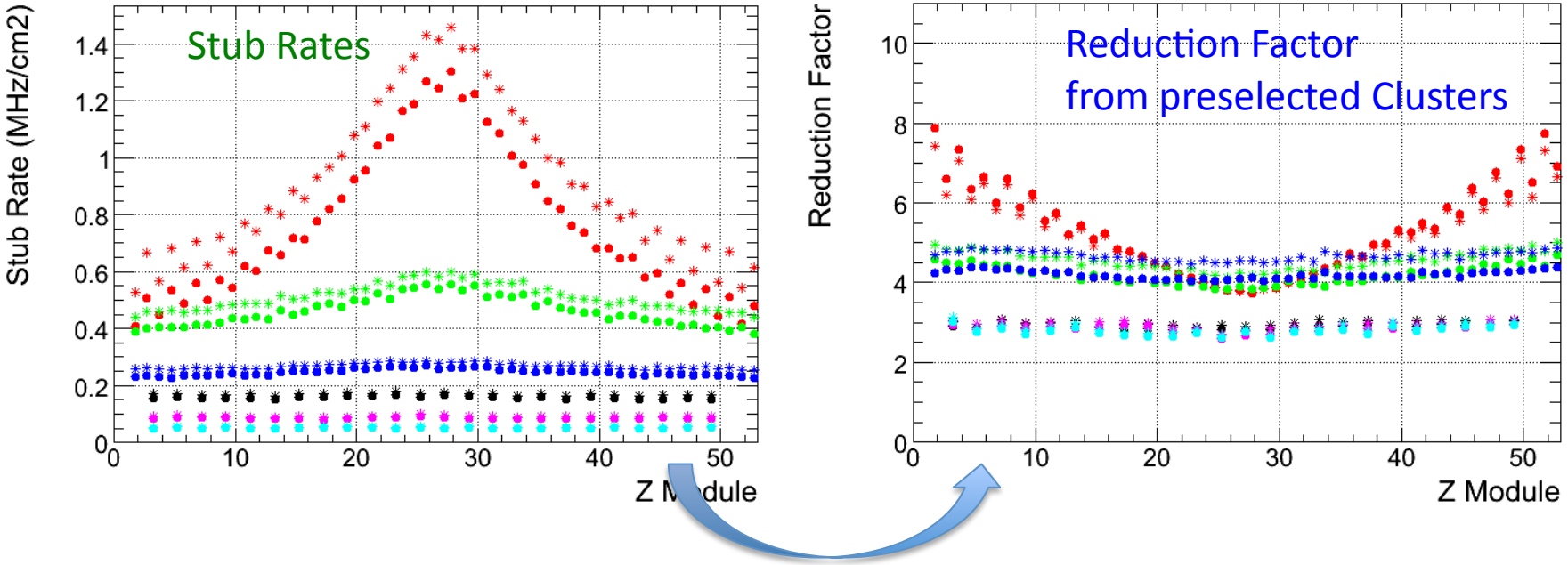
-A cluster can form ONE stub at maximum.



Track Stubs Rates and Reduction factor



Window width=6 strips



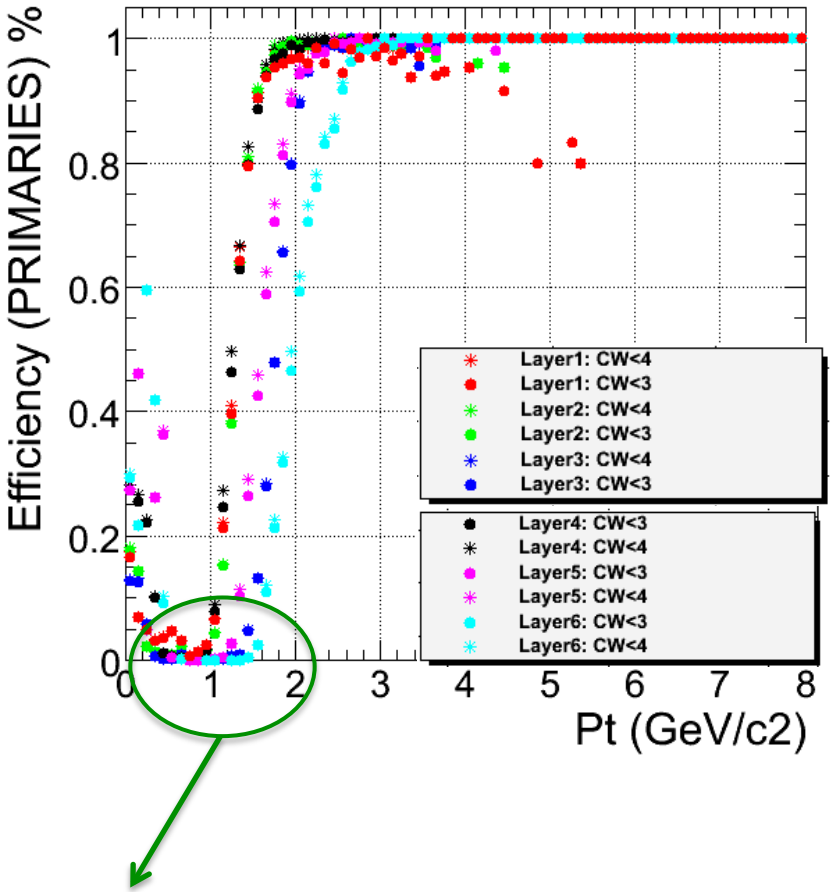
Like clusters, the rates are reduced when applying a tighter cut on the CW

The reduction factor from preselected clusters is higher for inner layers

The cut (Ww width) is the same on every layers: It should certainly be optimized according to the radius and sensor spacing.

Track Stubs: Turn on curves

Stub production efficiency (physical processes associated to primary vertices only)



Same window cut -> different thresholds depending on the radius

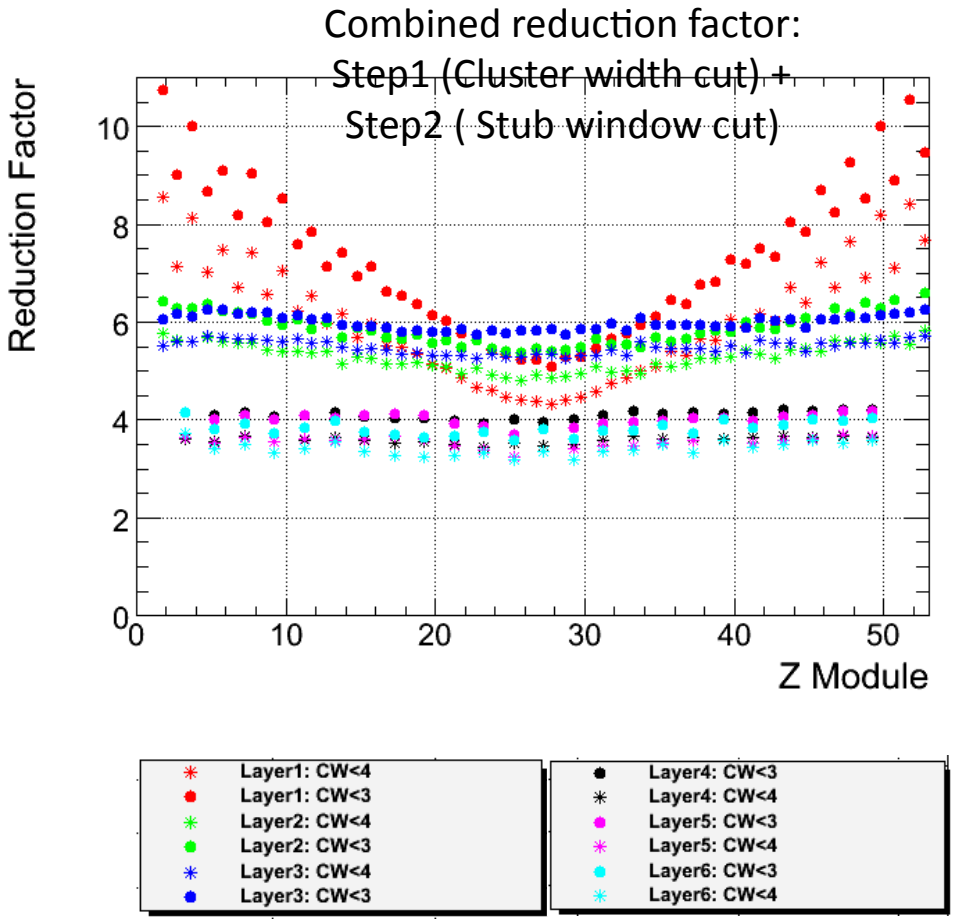
Effect of the cut (window size) is clearly visible, the efficiency is dropping dramaticcaly below 1-2.5GeV depending on the layer.

The cluster width cut is reducing the rates without affecting noticeably the primary stubs production efficiency.

The window size should now be optimized for each layer separately, in order to maximize the reduction factor while keeping a high efficiency above the cut

The shape at very low Pt (<200MeV) is not fully understood, yet. Most likely due to primary looper tracks

Reduced Readout: Combined Reduction Factor



Several issues for <u>each layer</u> (work in progress):

-78% of tracks are coming from secondary particles (nuclear interactions,…). Secondary vertices can be far from the interaction region which lead to a bad correlation between two clusters

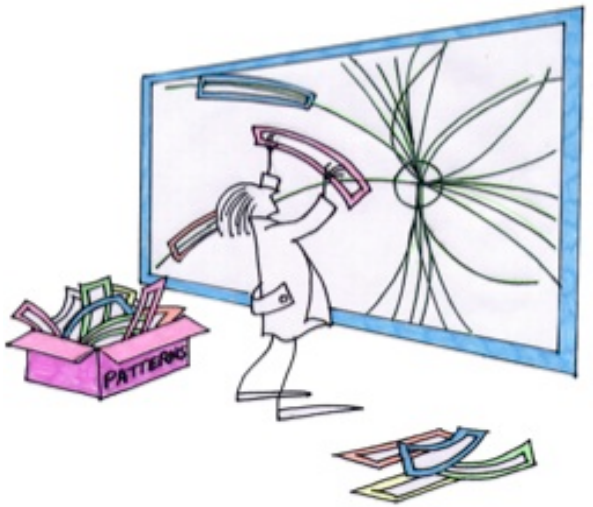
-Impact of secondaries should be carefully evaluated

-The sensor spacing might be refined (currently L1:2.6mm, L2-L3:1.6mm, L4-L5-L6: 0.8mm)

-An appropriate pt cut (i.e Window size) depending on the layer should be defined

In summary: in addition to the track behaviour understanding (loopers, primaries/ secondaries), the optimal set of {sensor spacing / Pt cut} should be derived in order to get the best compromize between a high data reduction factor and a high efficiency above the Pt cut

From Stubs to Tracks at L1 using Associative Memories



Current status for the barrel-EndCap Geometry (INFN-Pisa/IN2P3-Lyon)

- Study the feasibility of a track trigger at L1, based on the same idea than the ATLAS FTK:
  - Rough pattern recognition (*compare active regions with prestored patterns)*
  - Fast track fit (*using hits contained in the matching patterns)*

→ Feasibility studies made by **Giuseppe Broccolo** in 2011 for first version of a barrel-endcap tracker layout (presented at TIPP 2011)

- Develop a flexible software emulation in order to test the feasibility/scaling of such a system with the different tracker geometries foreseen.

- Study in parallel the hardware part of the problem (*ATLAS is working at the HLT, we want to work at L1: 40MHz!).*

Bank Pattern:
First step: dividing the Tracker into sector

Sector size (*in r/φ) is geometrically constrained by the minimal Pt of the track one wants to recover (2GeV/c in our case)*

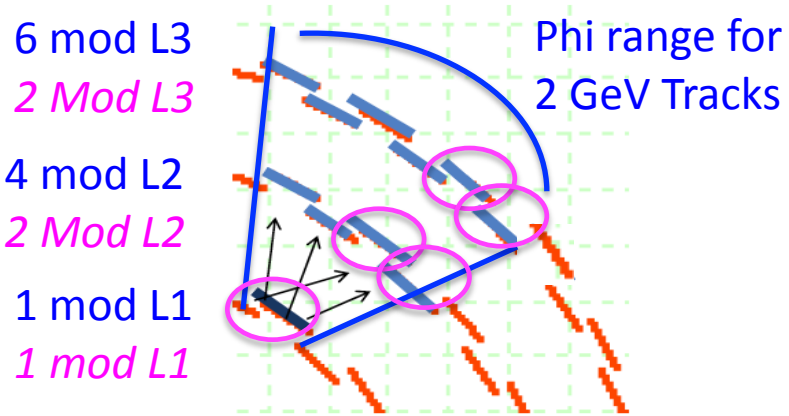
The sector size is also technically constrained by the maximum number of hits which might be fed into the AM card at L1

For the moment we considered only **the barrel part (3 outer layers)**

-From a given hitted module on the first considered layer, one can derive the phi range in terms of number of modules on the next 2 layers
(A super Sector contains 11 modules)

-In order to keep the dimensions of the system as small as possible, (i. e 5 modules per sector) one can divide the combinatorics of possible modules from the supersector into three sectors of 5 modules in r/φ

Super-Sector and Sector
(schematic representation)

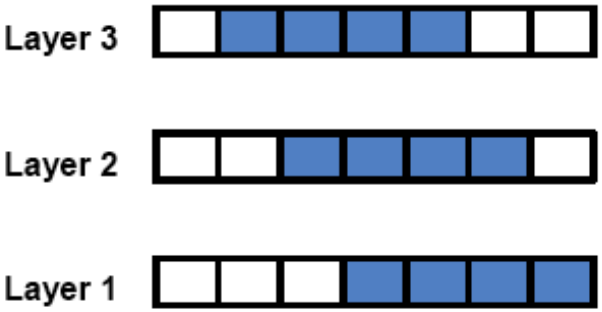


→The size of the reduced sector (5 modules) might involve depending on the technical constraints
→Considering the relative symmetry of the system, it is straightforward, from there, to get the sectors for all the Tracker

Bank Pattern: SuperStrips and Data Format

A pattern is a set of super-strips (one per layer) defining a road into the tracker

The SuperStrip is the main parameter of the system



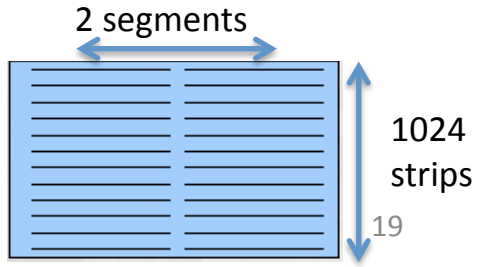
Pattern 1

Geometrically, the super strip is nothing else than a low granularity tracker strip
Technically, what is sent to the AM card is the address of the super-strip in the considered sector

The pattern is then defined with 3 x 12 bits
(3 outer layers considered for the moment)

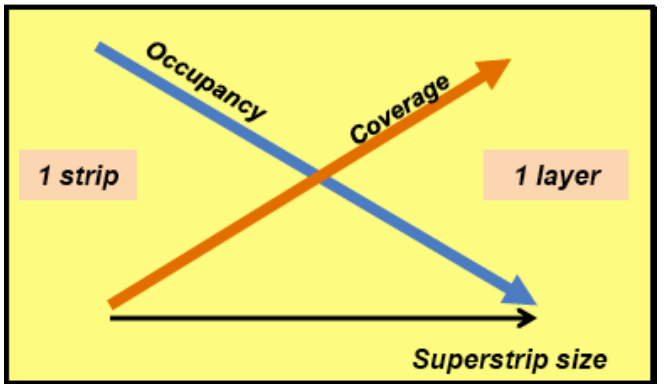
12 bits to code the super strip :
Z module position : 4 bits
Phi position : 1 bit
Segment(*) : 1 bit
Super strip position : 6 bits (*64 possibilities : allows a pitch of ~1.6 mm)*

(*) A sensor on a module is divided into 2 segments of 1024 strips



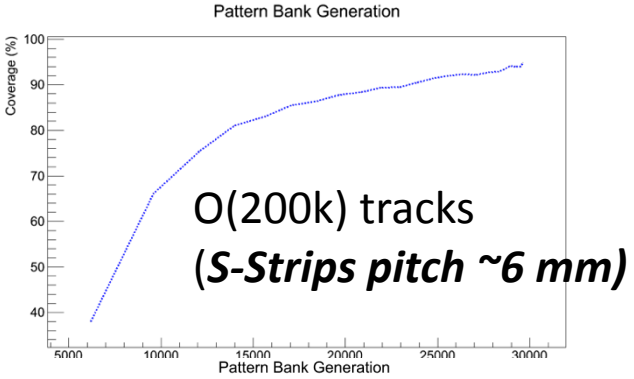
Bank Pattern Generation

Find the balance between the best coverage (*ensure that all interesting tracks are contained into at least one pattern) and the best occupancy (minimize the number of tracks/hits per pattern).*

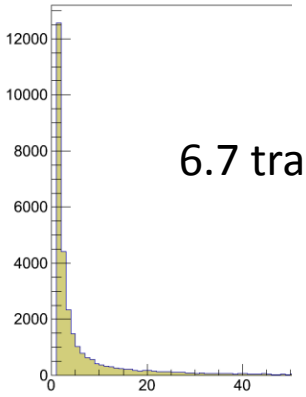


Iterative method: every step (bunch of 10k muon tracks)

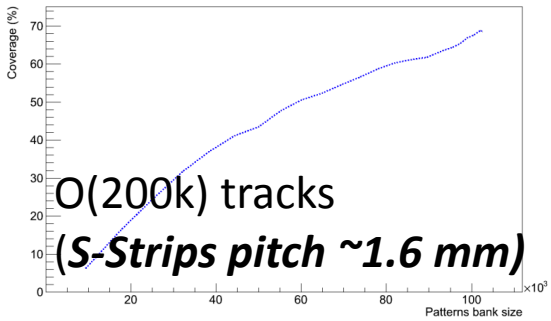
- Generate patterns and add them in the bank
- Compare the number of patterns at iteration i with the one at iteration i-1
- Keep going until the number of patterns is reaching a maximum(full coverage)



O(200k) tracks
(***S-Strips pitch ~6 mm)***



6.7 tracks/pattern in average

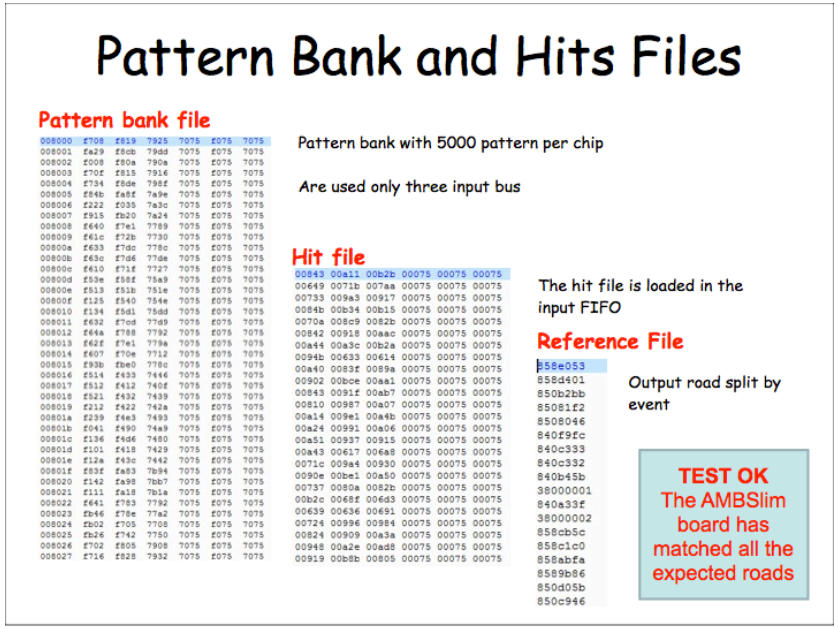


O(200k) tracks
(***S-Strips pitch ~1.6 mm)***

Not enough tracks to fill the bank…
Use a larger events set to create the 1.6 mm pattern bank (***in progress)***

Next Steps and Hardware developments

- Optimize AM bank size and sector dimensioning
  - study bank size as a function of the minimum pT and cluster position resolution
  - Cabling possibilities (configuration should fit hardware constraints)
- Test the patterns banks with Minimum Bias events (including high pile-up) with a dedicated AM emulation software
- Study latency an feasibility with SW and HW setup (IN2P3Lyon-INFNPisa-Perugia)
  - Interface pt-modules with boards prestoring a fixed # of patterns which simultaneously see stubs leaving detector at full speed.
  - Test made on the CDF system using 5K patterns per AM chip with simulated data and that the HW returns the same tracks that pass the selection on simulated AM chip



See talk from Daniel Magalotti this conference

Conclusions

- To exploit HL-LHC project potential, a new generation of tracking devices will be needed, able to provide high-resolution particle trajectories in real time and under unequalled experimental constraints as of today.

- Providing Tracking data for the L1 stands out as a formidable challenge

- The CMS Tracker next generation is still to be designed
  - The knowledge of the expected rates from realistic simulations are essential to design readout circuits and develop pattern recognition hardware.
  - The first simulation results presented today are already exploited as input for selective readout designs and pattern recognition technologies

    -> Strong and combined efforts on R&D and simulations