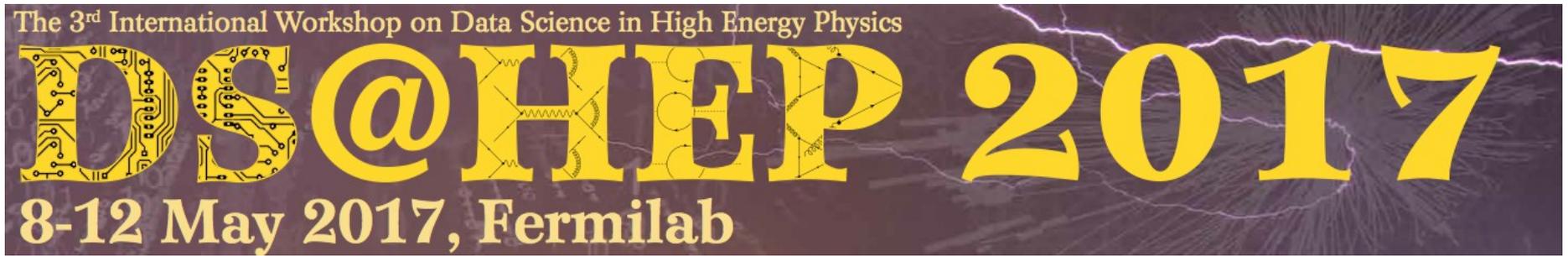




Charged Particle Tracking Hands-On



Dustin Anderson, Steve Farrell,
Dorian Kcira, Jean-Roch Vlimant



Content



- Hands-on Logistics
- Experimental setup
- Charged particle trajectories
- Charged particle tracking
 - Algorithms in use
 - Other approaches



Hands-on Logistics



Schedule



- Monday 1h30
- Wednesday 1h20 + 1h30
- Thursday 1h20 + 1h30
- Friday
 - Wrap up presentation



Register



- Add your name and github username on this googledoc
 - <https://docs.google.com/spreadsheets/d/1s3QIJvgrfyKDS>
 - We will update it with the machine name and the gpu to be set
- Navigate to the hub url provided
 - You will have to allow the github app access
- Start a session
 - This will spawn a jupyter session on the machines at caltech
- Register to the slack channel
 - https://join.slack.com/dshep2017/shared_invite/MTgwNz



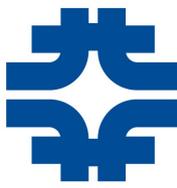
Setup



- Open a new terminal in jupyter (new : top right)
→ git clone
<https://github.com/HEPTrkX/heptrkx-dshep17.git> to
get the handson material
- Set `CUDA_VISIBLE_DEVICES` to the integer between
0 and 7 in first cell of NB
- Tracking hands-on are in the **hands-on/** directory
- The data is imported from **/inputdata/**
- Basic tutorials are available under **tutorial/**
→ Chances are that `root_numpy` won't work at this time

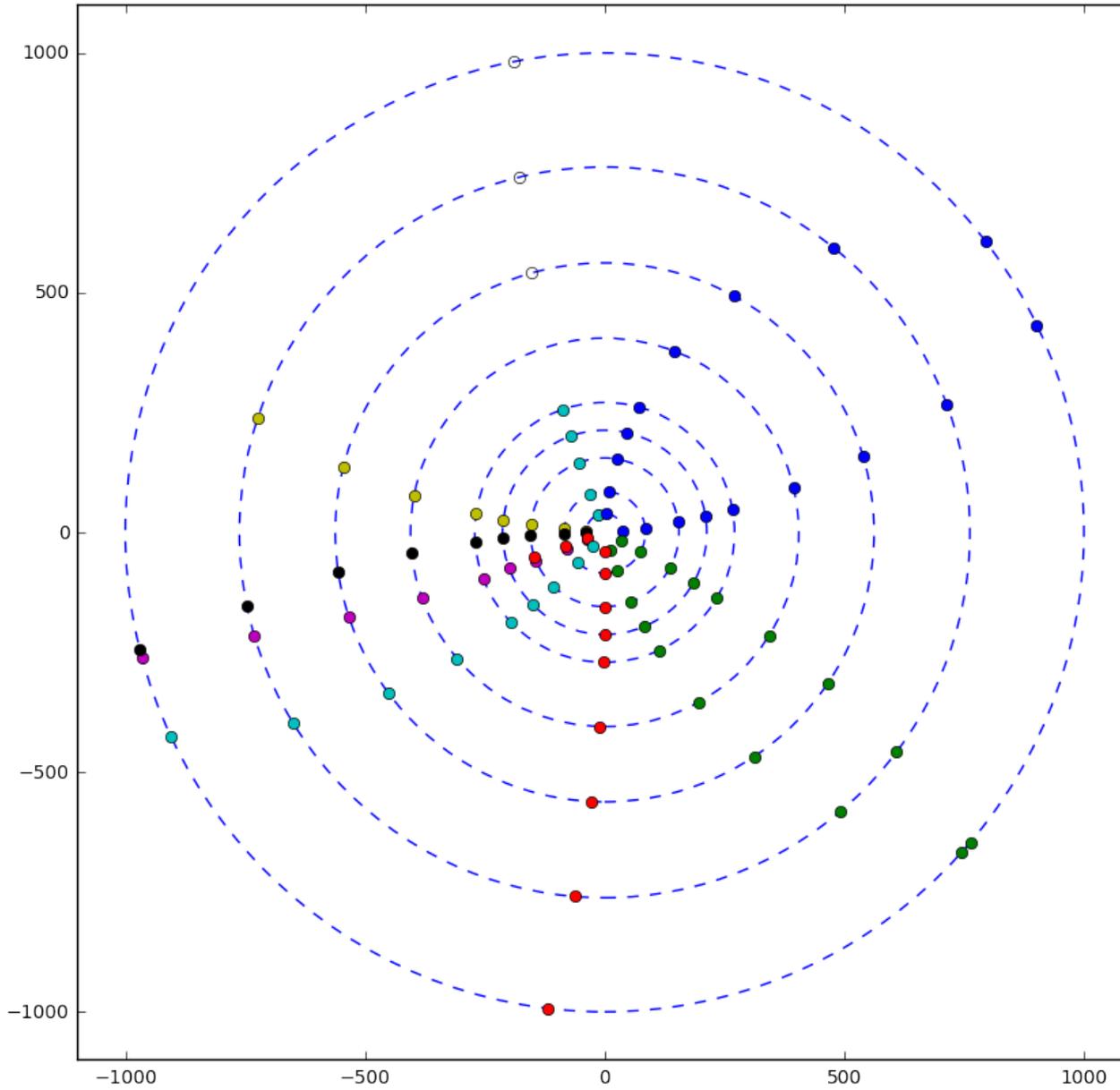


Dataset



- Kudo to Yetkin Yilmaz, David Rousseau, Balazs Kegl, Isabell Guyon, Mikhail Hushchyn from the RAMP challenge during <https://ctdwit2017.lal.in2p3.fr/>
- Simple 2D geometry
 - ~4k events (*event_id* index)
 - Distribution of tracks with poisson distribution with mean 10 (*cluster_id* index)
 - Flat pT distribution between 300(100) MeV and 1GeV
 - 9 layers (*layer* index)
 - Granular in phi (*iphi* index)
 - 2D hits (*x,y* global position)
- **Challenge in preparation**, only simplified generator can be used for now

Dataset





Hands-on

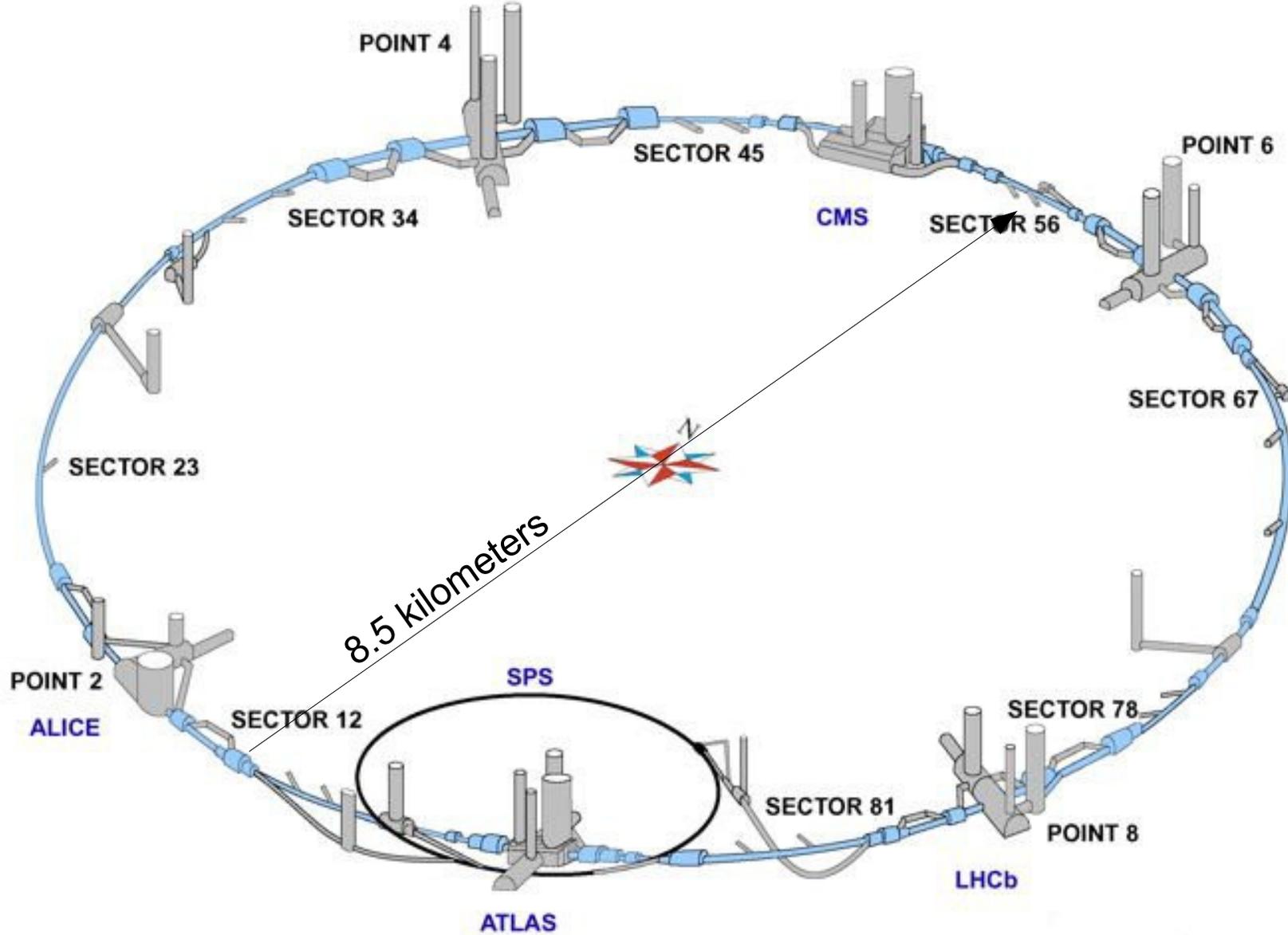
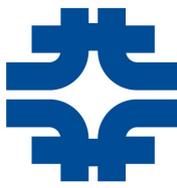


- Get people started, for those who need to get started.
- Brainstorm on way to attack the problem
- Simple starting kit model
- Track candidate prediction with convolutional neural nets
- Track parameters prediction with CNN and LSTM
- Hit association prediction with sequence to sequence
- ...

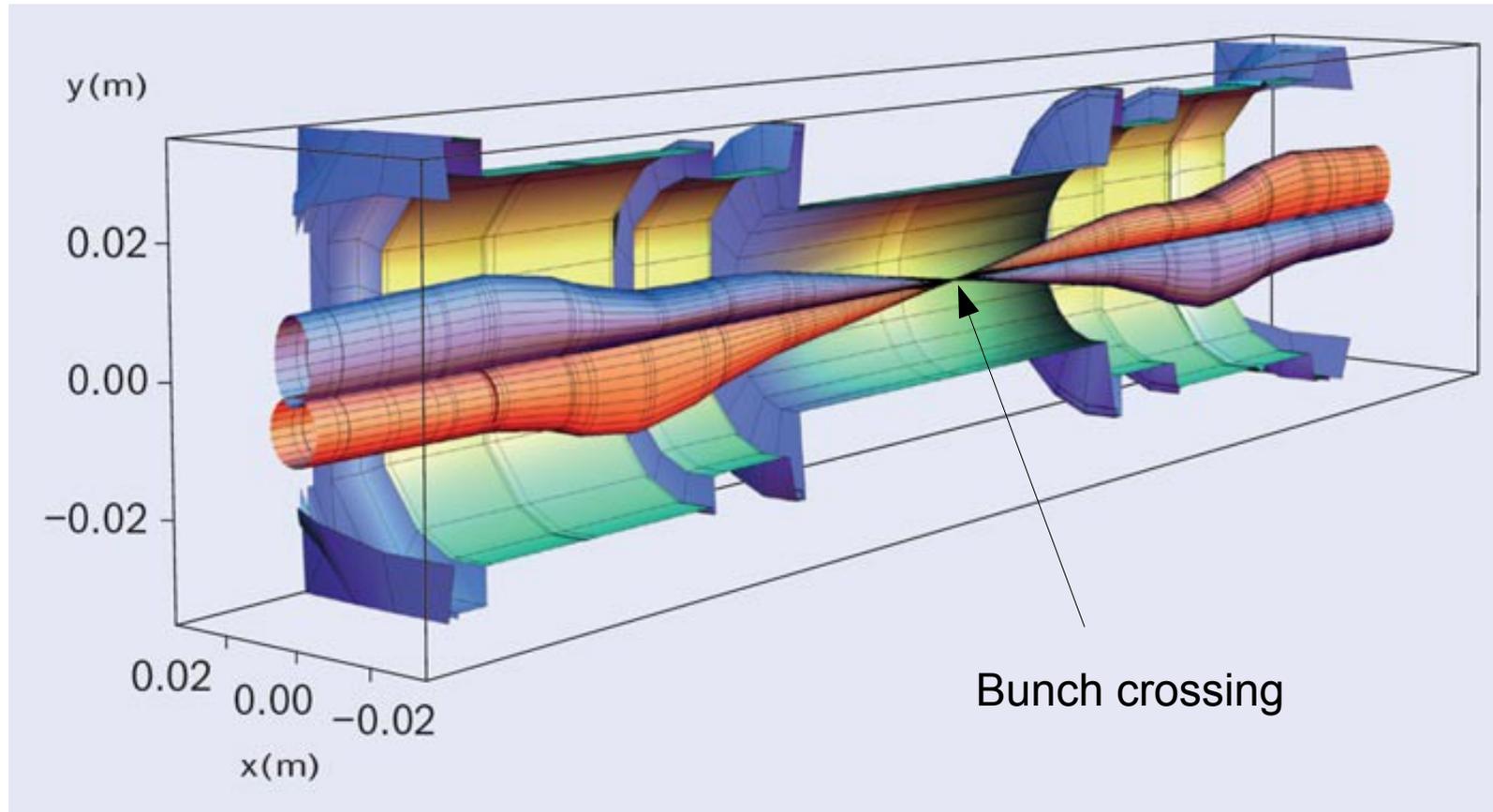
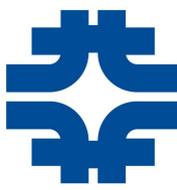


Experimental Setup

The Large Hadron Collider LHC

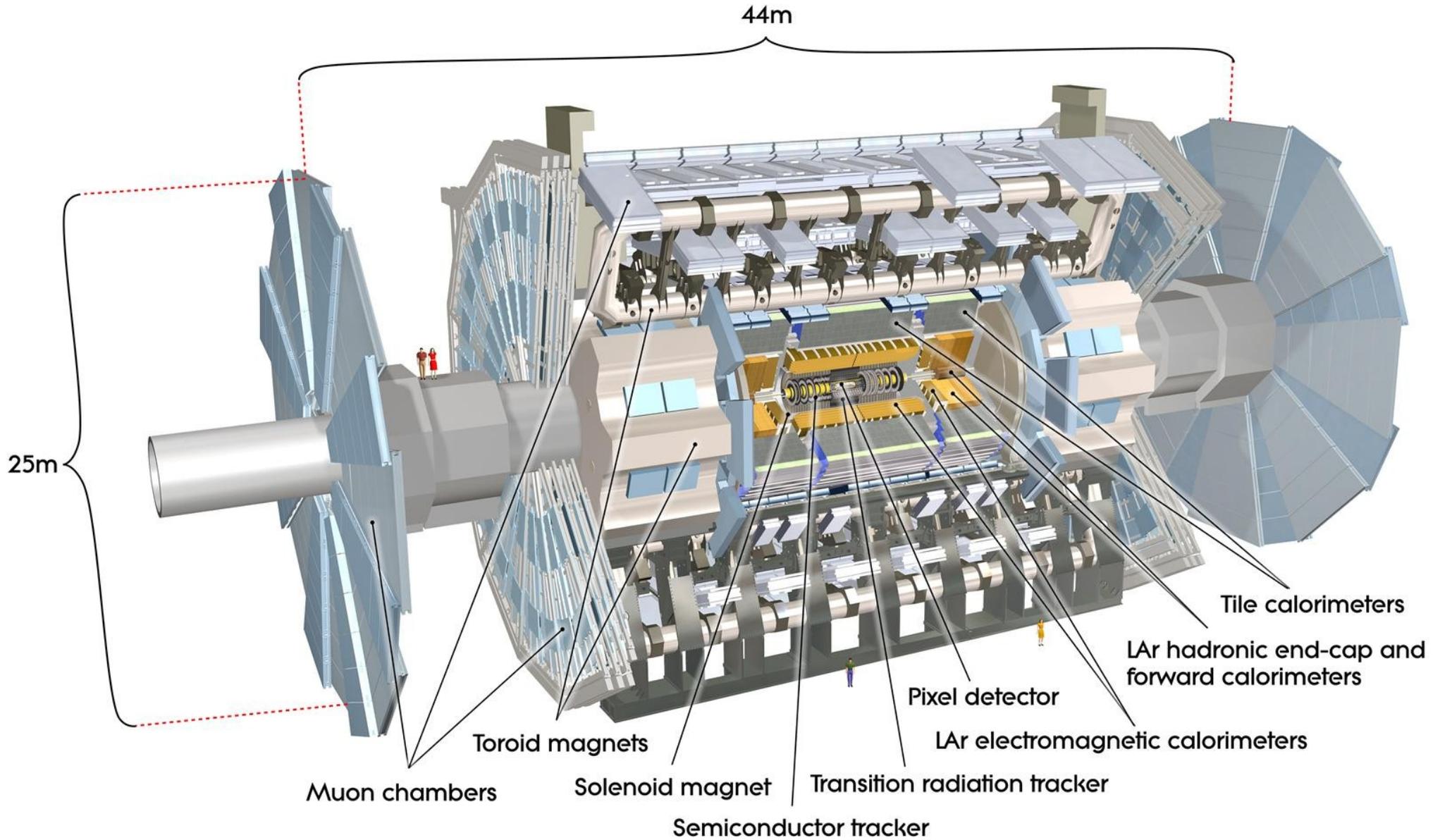


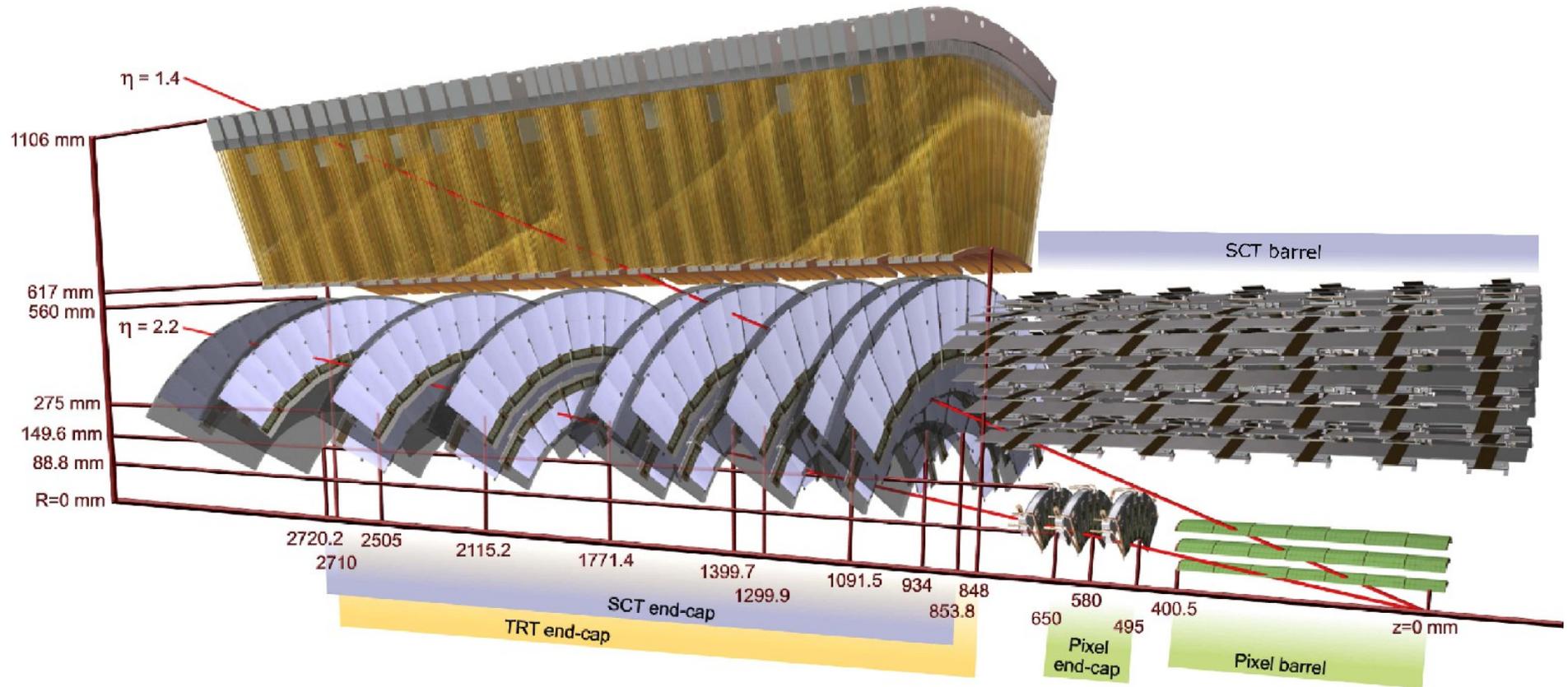
Collision at the LHC



- 10^{11} protons per bunch
- Bunch crossing every 25 ns (40MHz)
- Average number of proton-proton interaction per bunch crossing in ATLAS-CMS : 25-45

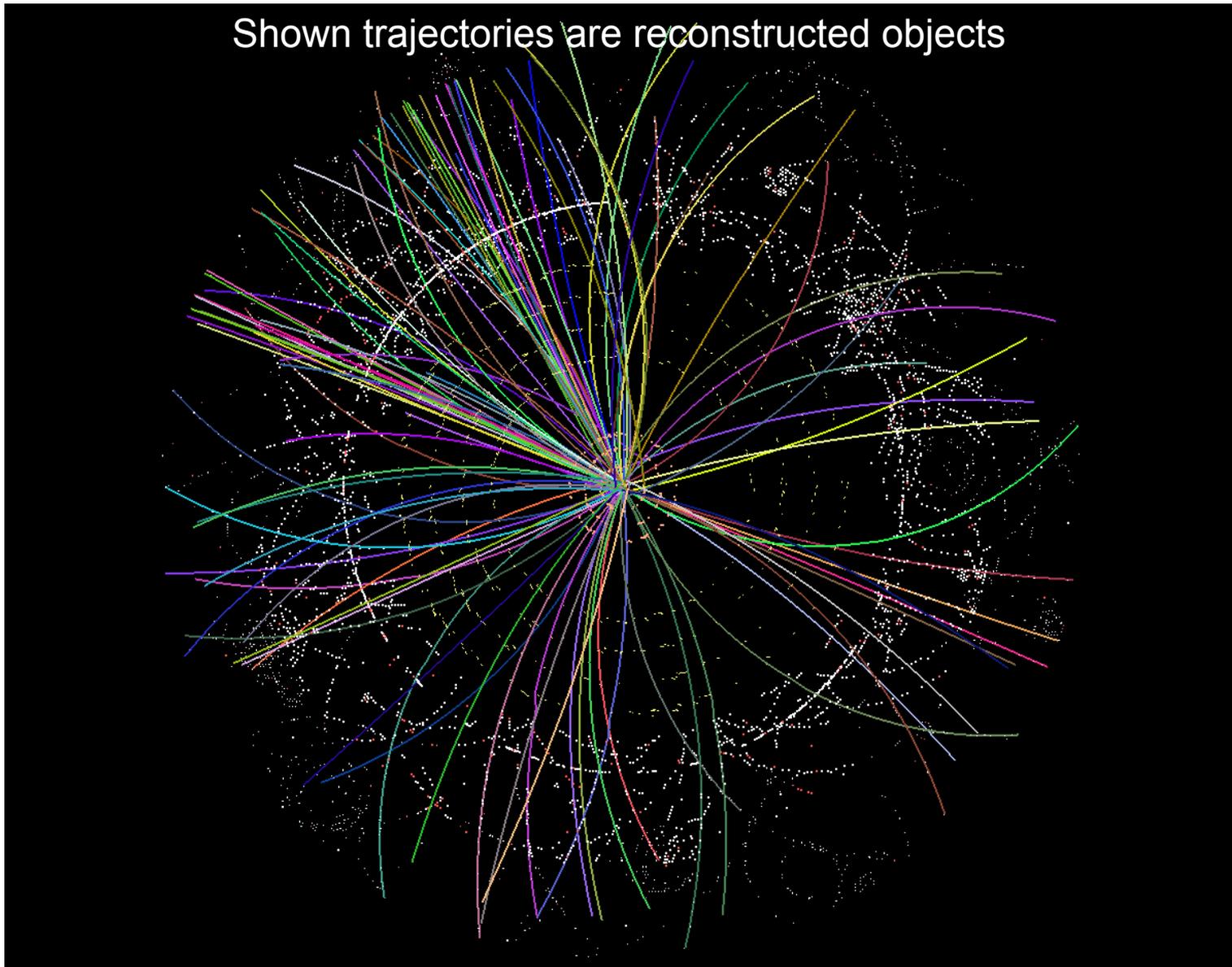
ATLAS





**2T solenoid magnetic field
along the beam line**

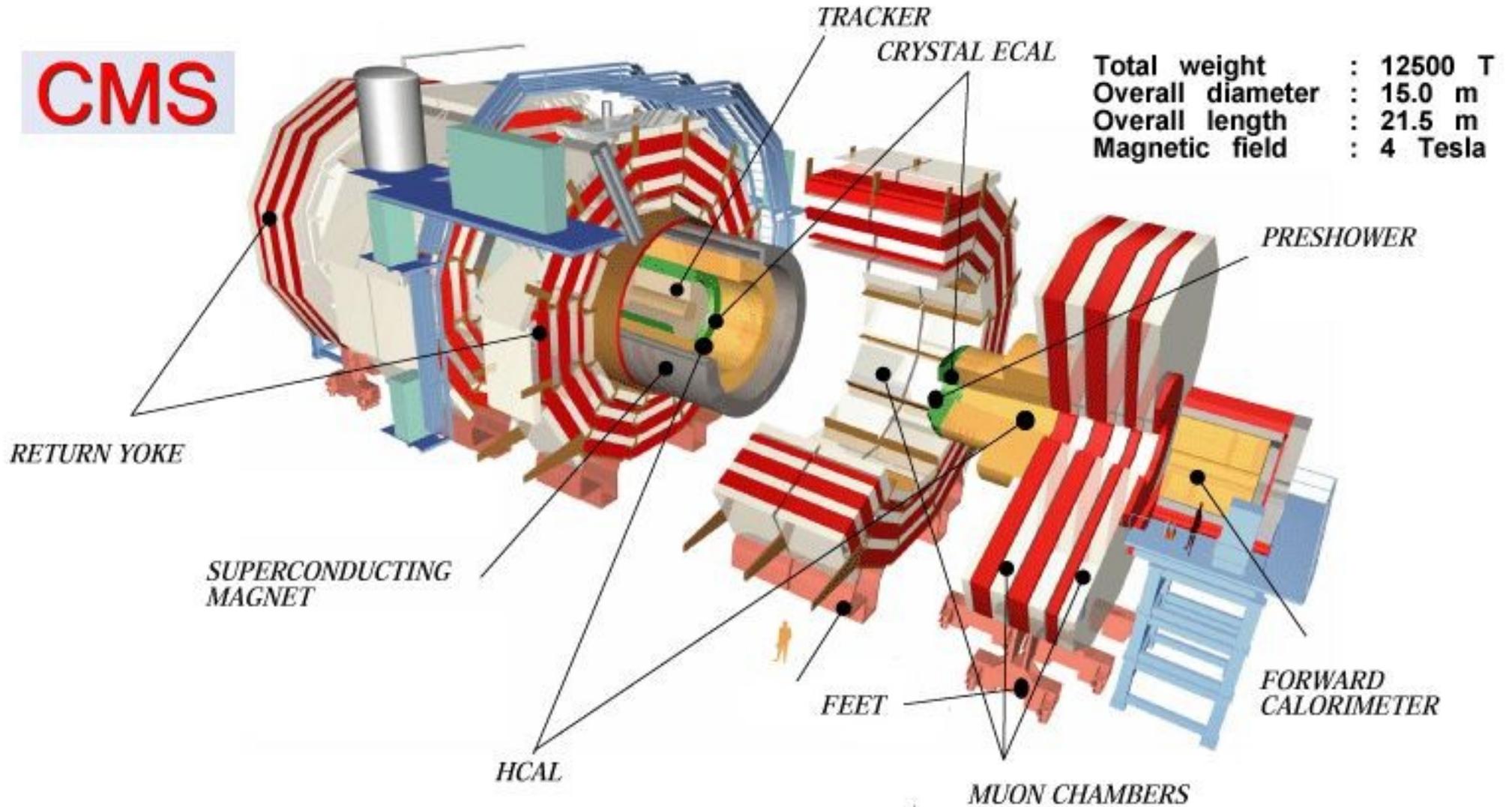
<http://atlas.cern/discover/detector/inner-detector>

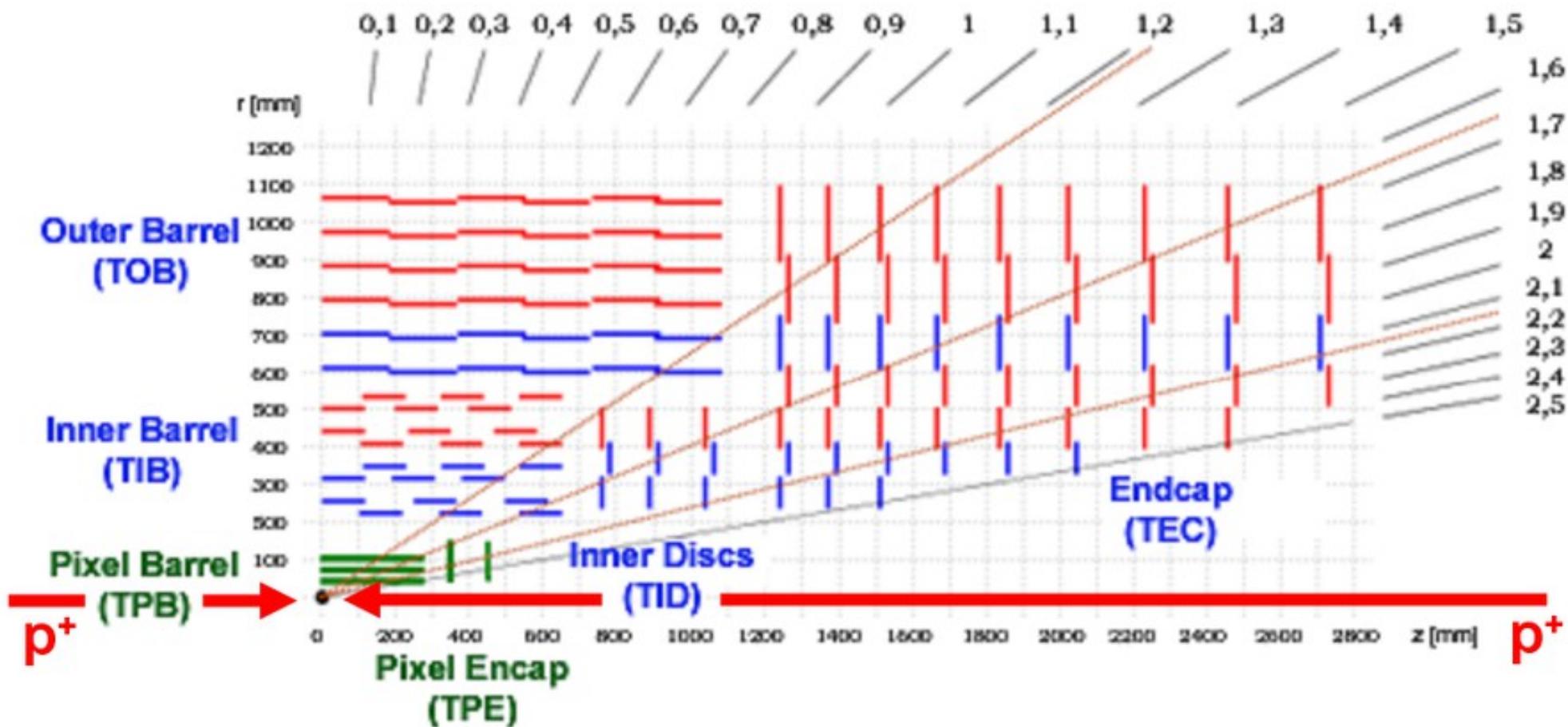


CMS



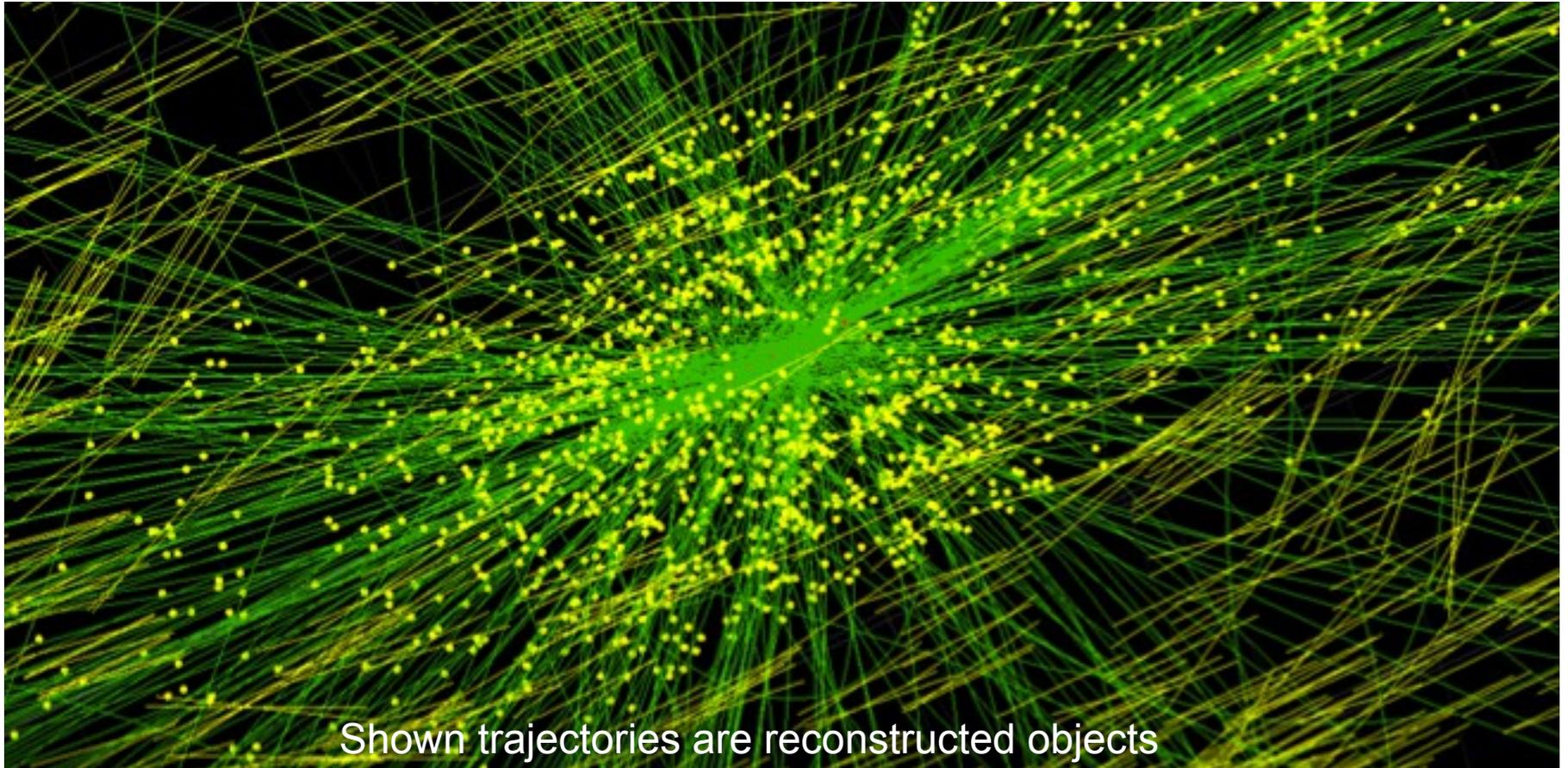
CMS

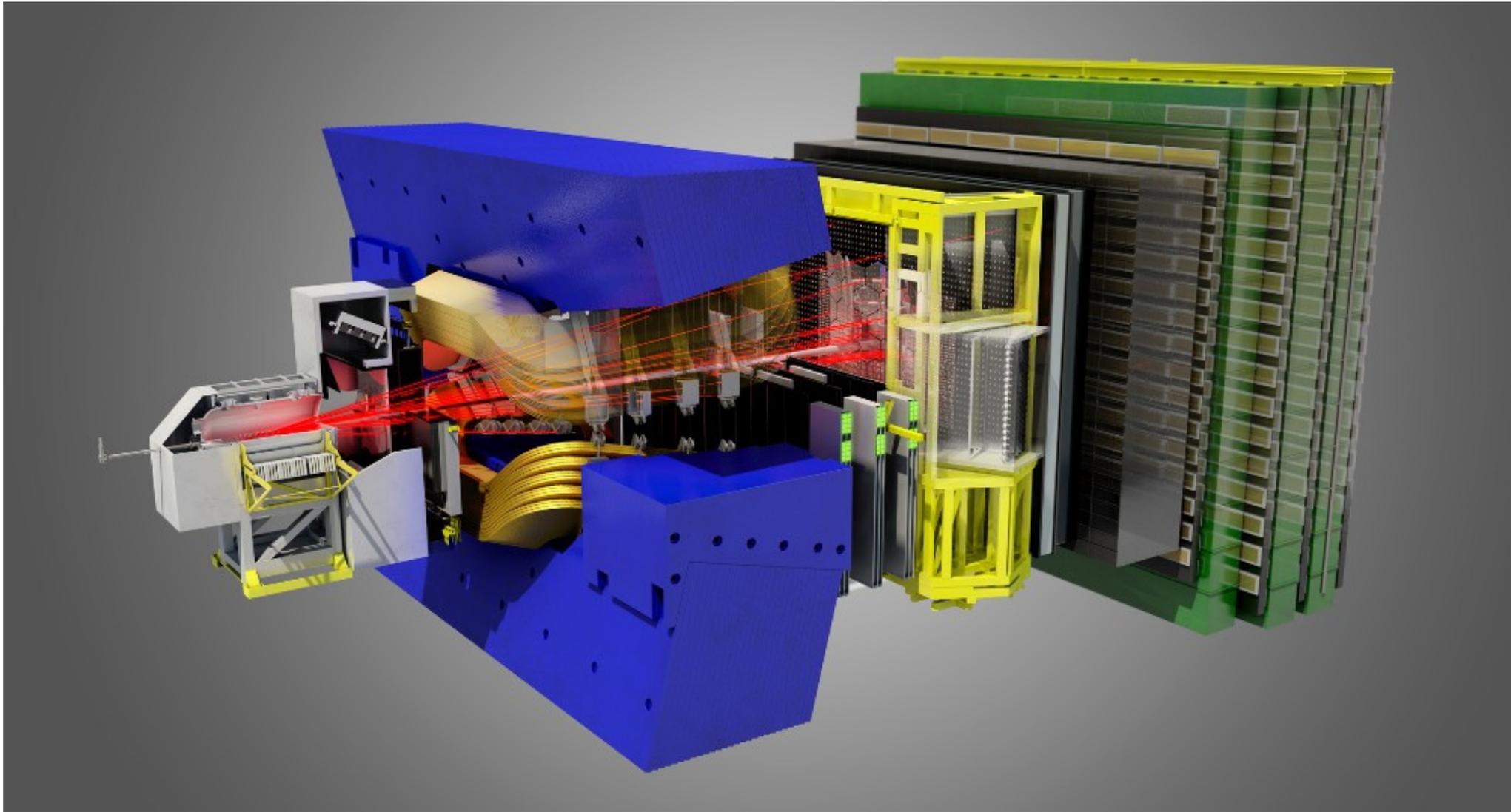




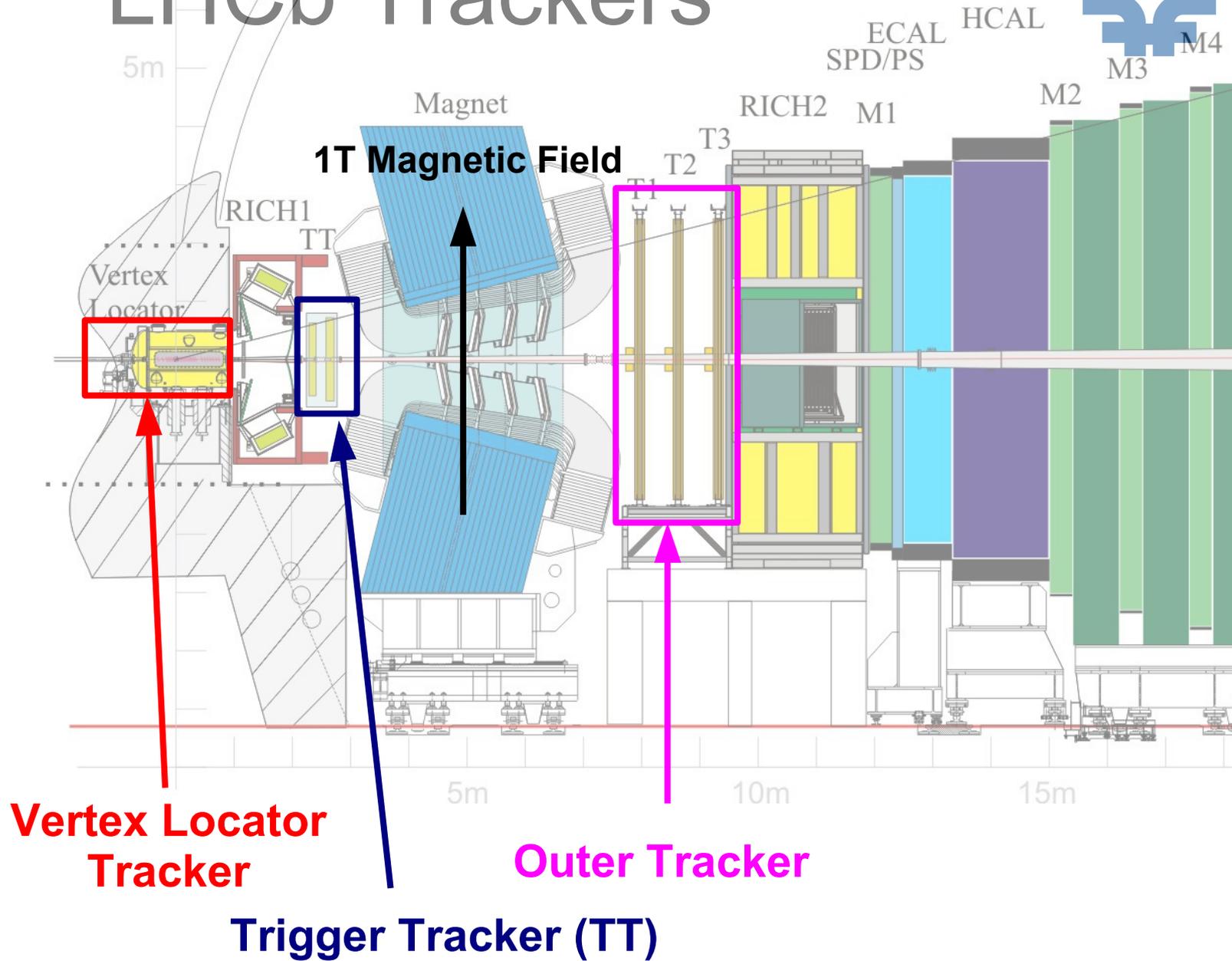
3.8T solenoid magnetic field along the beam line

<http://cms.web.cern.ch/news/tracker-detector>





LHCb Trackers



**Vertex Locator
Tracker**

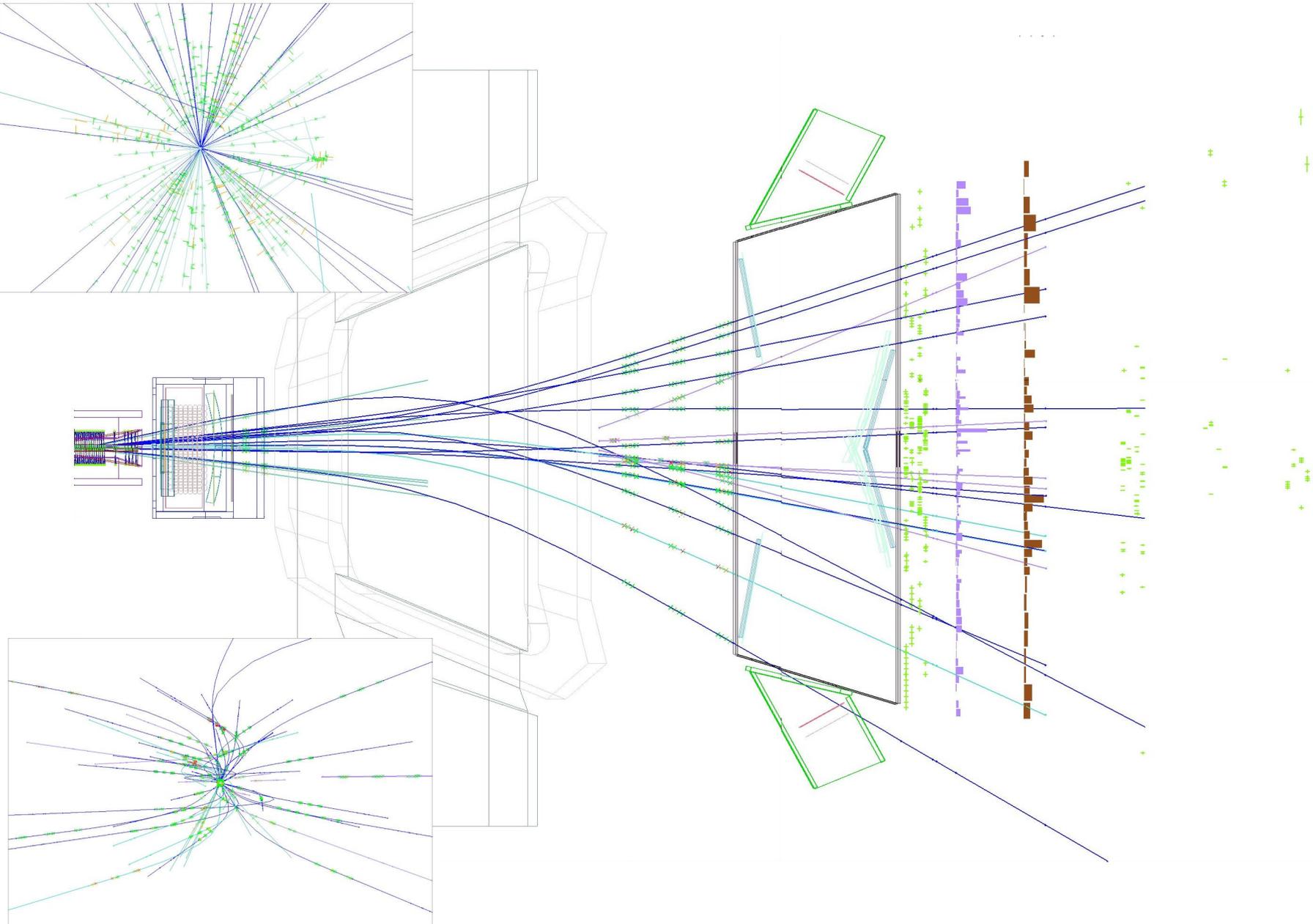
Outer Tracker

Trigger Tracker (TT)

<https://lhcb-public.web.cern.ch/lhcb-public/en/Detector/VELO-en.html>
<https://lhcb-public.web.cern.ch/lhcb-public/en/Detector/Trackers-en.html>



LHCb Visualization



Shown trajectories are reconstructed objects

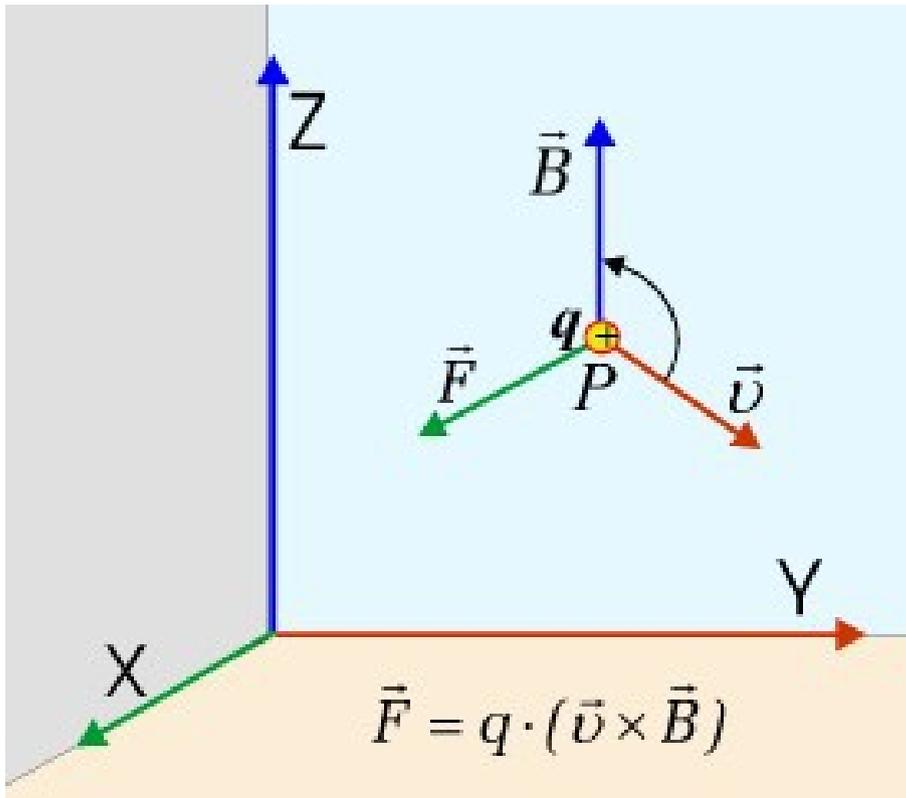
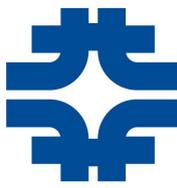


A Charged Particle Journey

First order effect : electromagnetic elastic interaction of the charge particle with nuclei (heavy and multiply charged) and electrons (light and single charged)

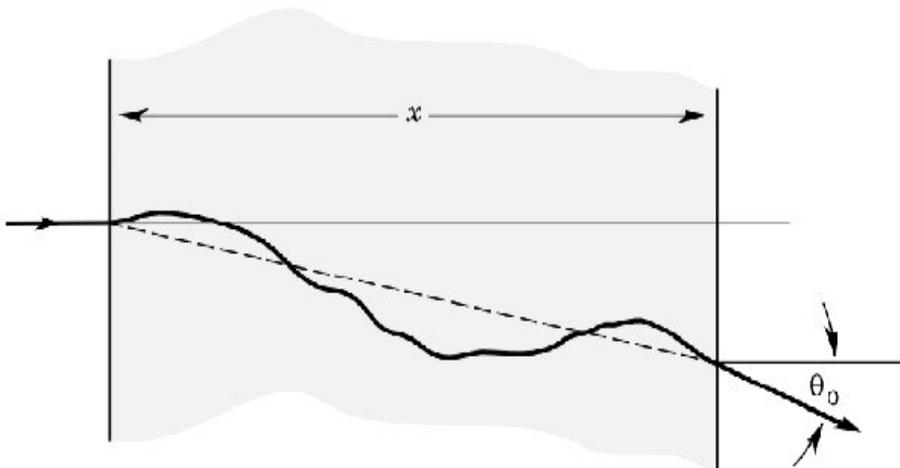
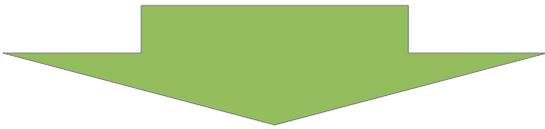
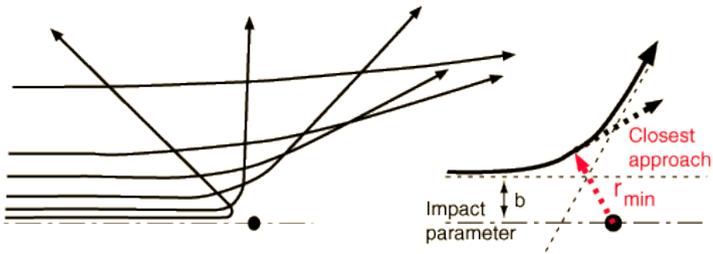
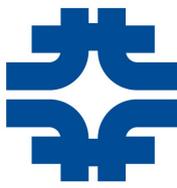
Second order effect : inelastic interaction with nuclei.

Magnetic Field



- Magnetic field B acts on charged particles in motion : Lorentz Force
- The solution in uniform magnetic field is an helix along the field : 5 parameters
- Helix radius proportional to the component of momentum perpendicular to B
- Separate particles in dense environment
- Bending induces radiation : bremsstrahlung
- The magnetic field has to be known to a good precision for accurate tracking of particle

Multiple Scattering



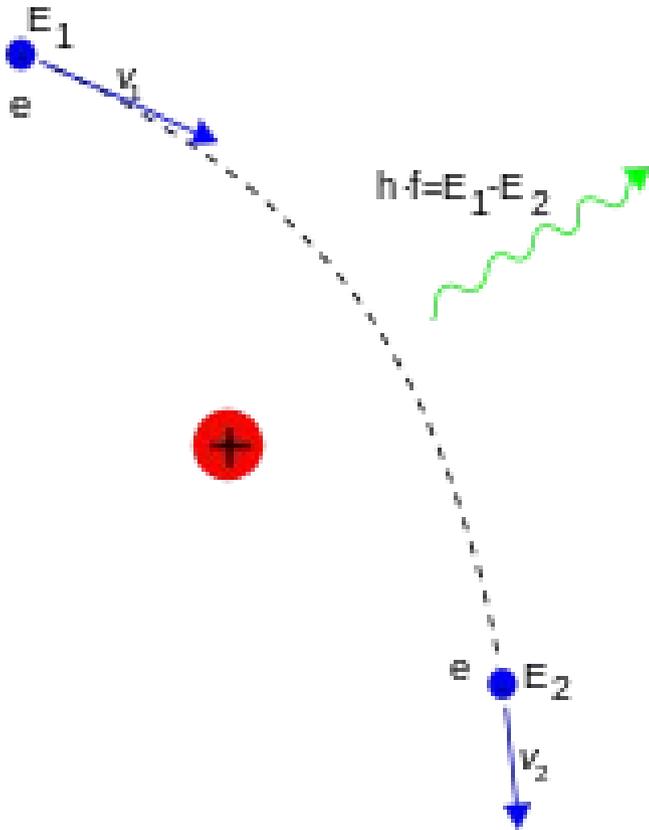
- **Deflection on nuclei** (effect from electron are negligible)
- Addition of scattering processes
- Gaussian approximation valid for substantial material traversed

Gaussian Approximation

$$\theta^2 = \left(\frac{13.6 \text{ MeV}}{\beta c p} \right)^2 * \frac{x}{X_0}$$

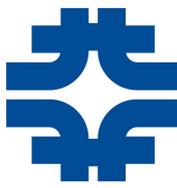
- β - particle velocity
- ρ - material density
- P - particle momenta

Bremsstrahlung



- Electromagnetic radiation of charged particles under acceleration due to nuclei charge
- Significant at low mass or high energy
- Discontinuity in energy loss spectrum due to photon emission and track curvature
- Can be observed as kink in the trajectory or presence of collinear energetic photons

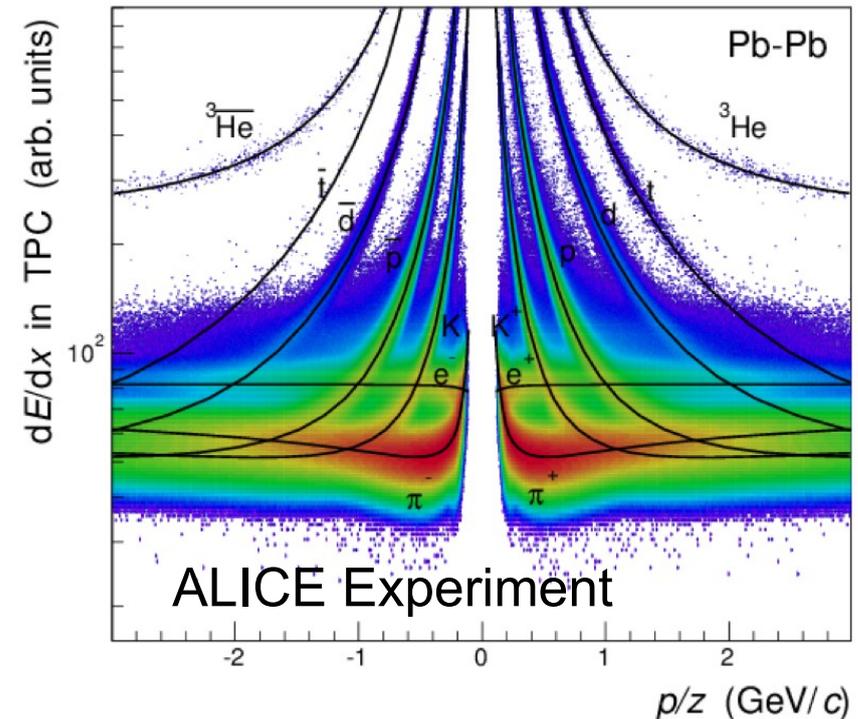
Energy Loss



- Momentum transfer to electrons when traversing material (effect of nuclei is negligible)
- Energy loss at low momentum depends on mass : can be used as mass spectrometer

$$dE / dx = k_1 \frac{Z}{A} \frac{1}{\beta^2} \rho \left(\ln \left(\frac{2m_e c^2 \beta^2}{I(1-\beta^2)} \right) - \beta^2 - \frac{\delta}{2} \right)$$

β - particle velocity
 ρ - material density
 Z - atomic number of absorber
 A - mass number of absorber
 I - mean excitation energy
 δ - density effect correction factor - material dependent and β dependent





Summary on Material Effects

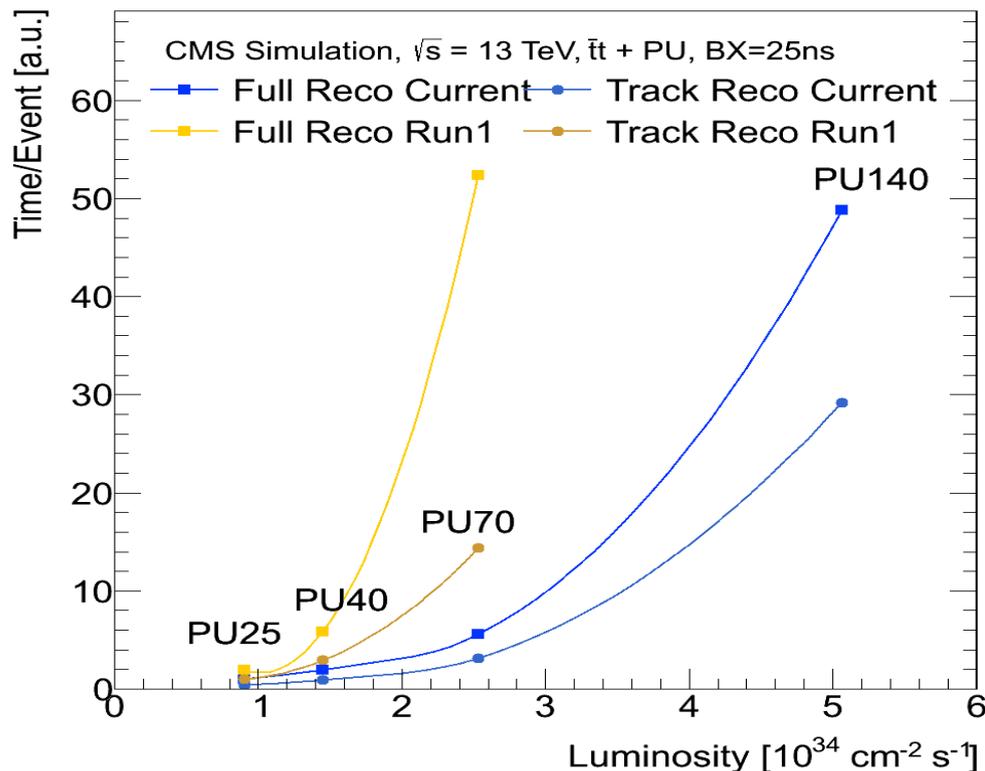


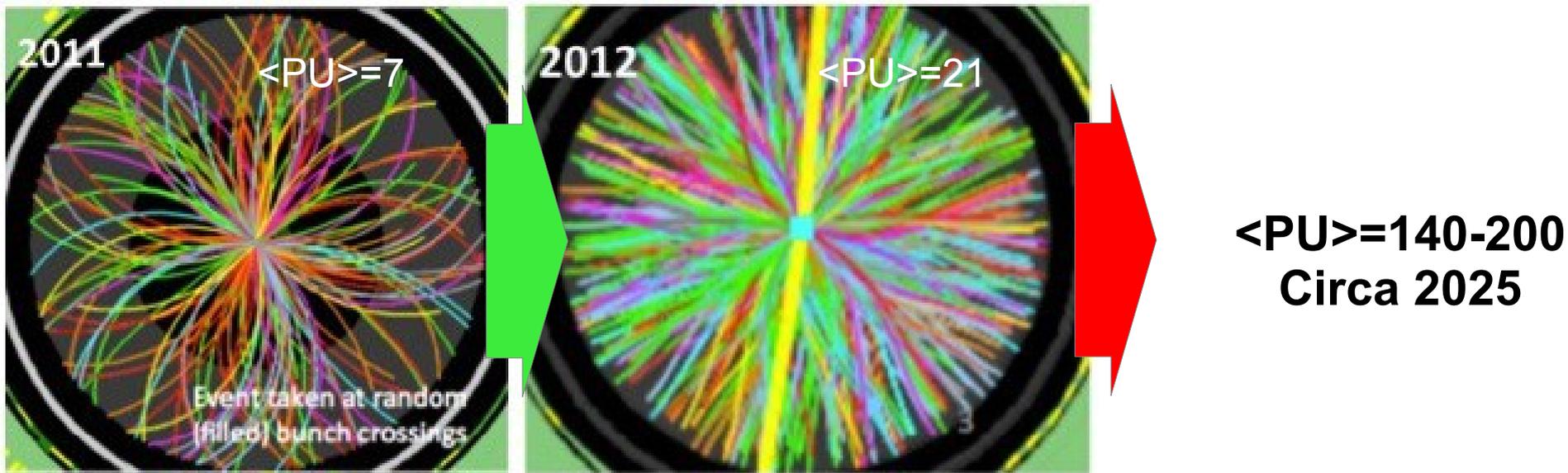
- Collective effects can be estimated statistically and taken into account in how they modify the trajectory
- Bremstrahlung and nuclear interactions significantly distort trajectories



High Luminosity LHC The Challenge

- Charged particle track reconstruction is one of the **most CPU consuming task** in event reconstruction
- **Optimizations (to fit in computational budgets) mostly saturated**
- Large fraction of CPU required in the HLT. **Cannot perform tracking inclusively at CMS and ATLAS. Online tracking strategy for LHCb.**





- CPU time extrapolation into HL-LHC era far **surpasses growth in computing budget**
- **Need for faster algorithms**
- Approximation allowed in the trigger



Algorithms In Use



Event Reconstruction

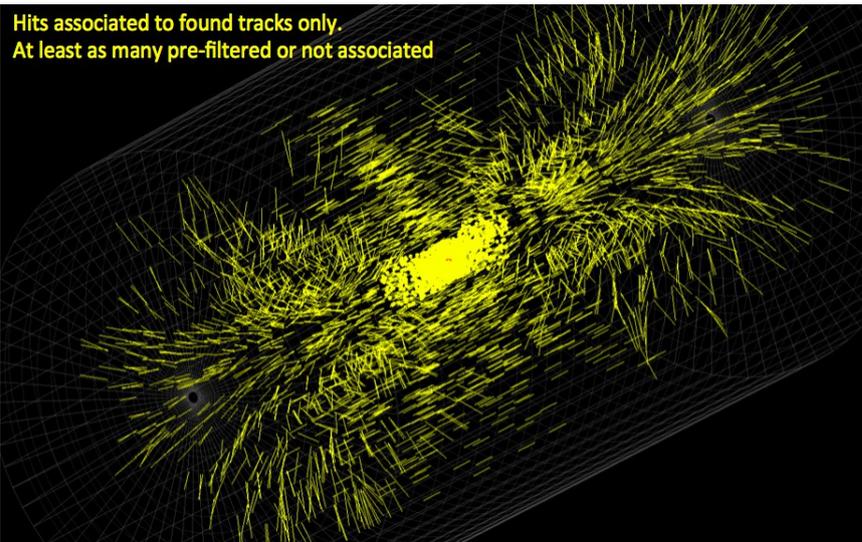


From individual measurements in sub-detectors to kinematics and properties of particles created in collisions

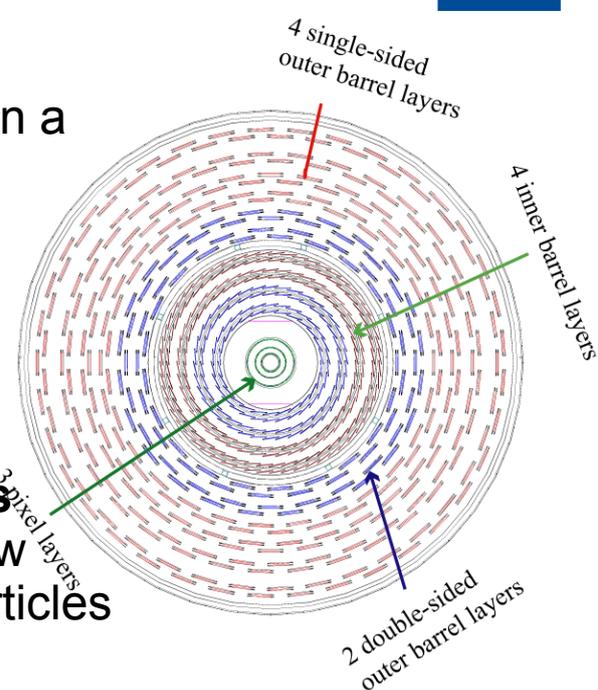
In a Nutshell



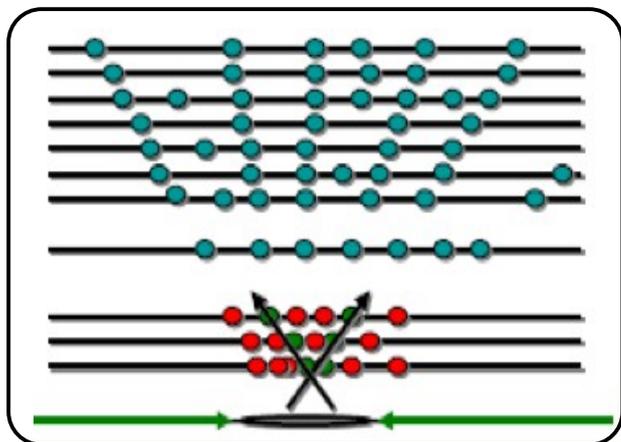
Hits associated to found tracks only.
At least as many pre-filtered or not associated



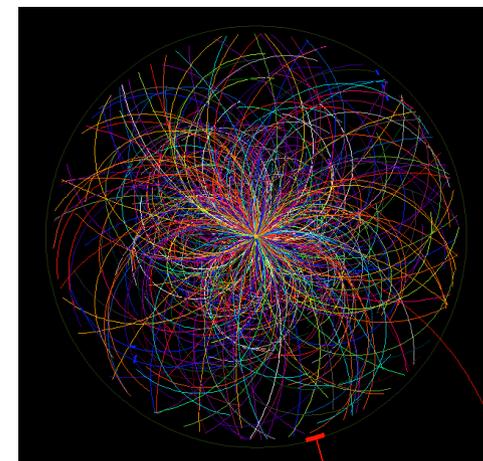
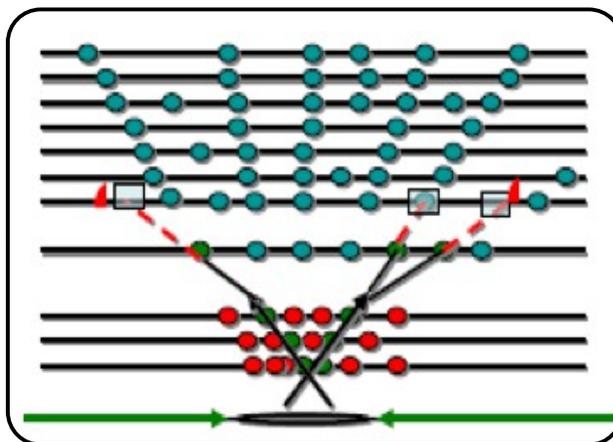
- Particle trajectory bended in a solenoid magnetic field
- Curvature is a proxy to momentum
- Particle ionize silicon pixel and strip throughout several concentric layers
- **Thousands of sparse hits**
- Lots of hit pollution from low momentum, secondary particles



Seeding



Kalman Filter



- **Explosion in hit combinatorics** in both seeding and stepping pattern recognition
- **Highly time consuming task** in extracting physics content from LHC data

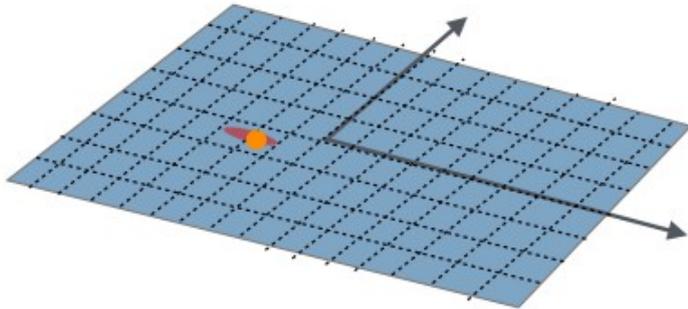
In a Nutshell

- Several Times
- 
- Hits preparation
 - Seeding
 - Pattern recognition
 - Track fitting
 - Track cleaning

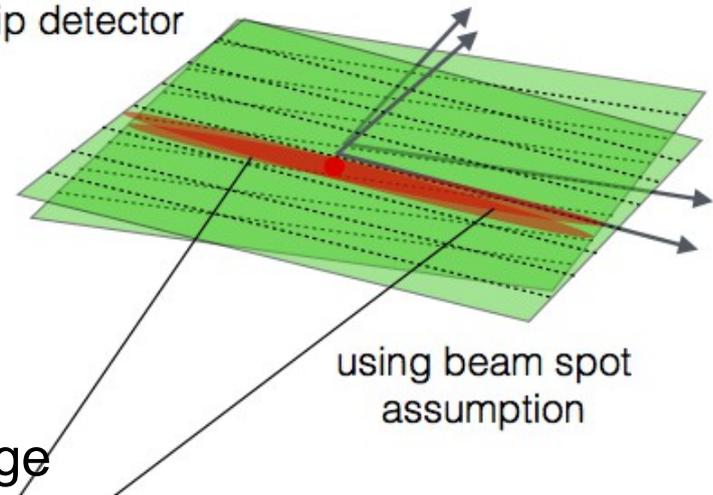
Hit Preparation



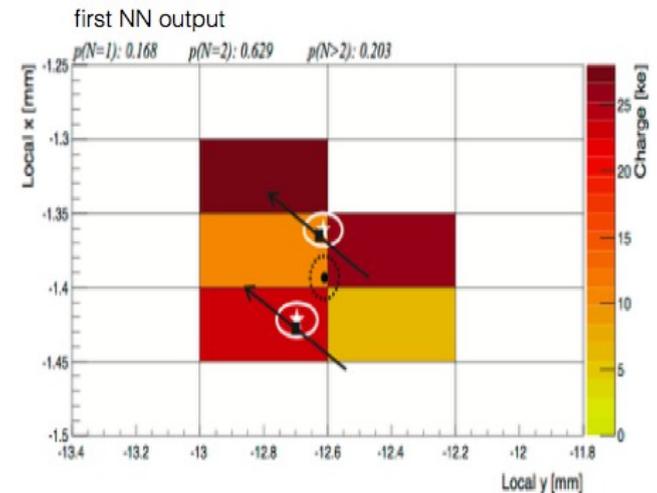
pixel detector



strip detector

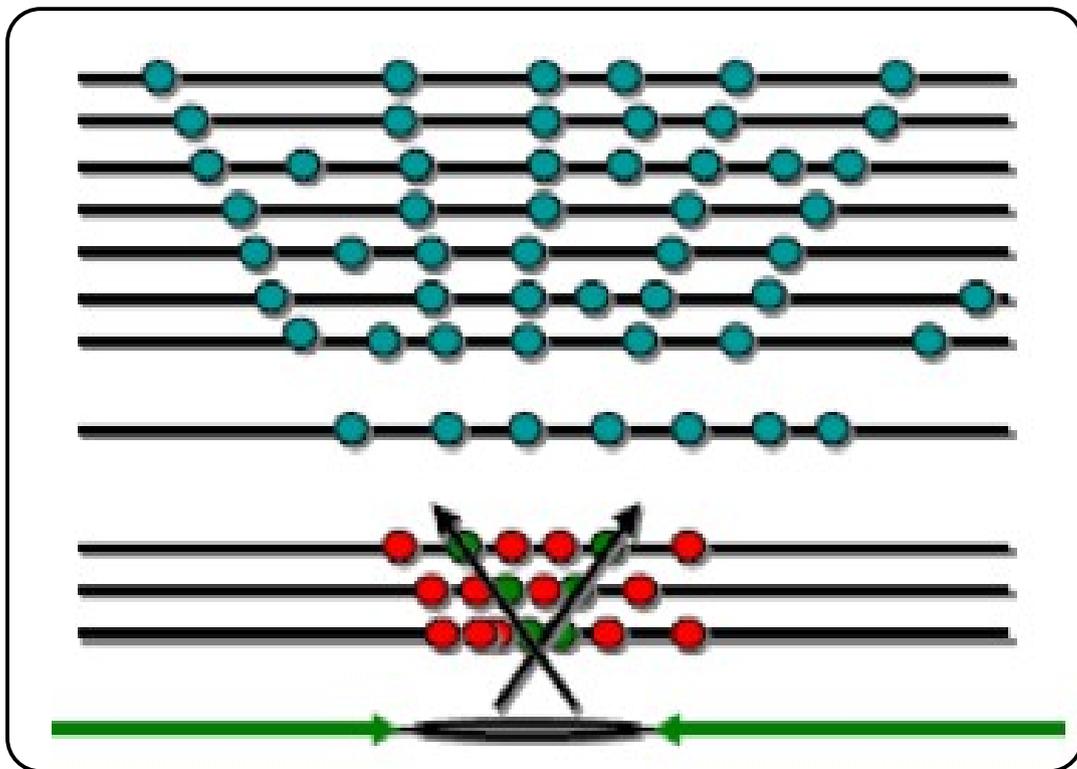


- Calculate the hit position from barycenter of charge deposits
- Use of neural net classifier to split cluster in ATLAS
- Access to trajectory local parameter from cluster shape
- Remove hits from previous tracking iterations
- HL-LHC design include double layers giving more constraints on the local trajectory parameters

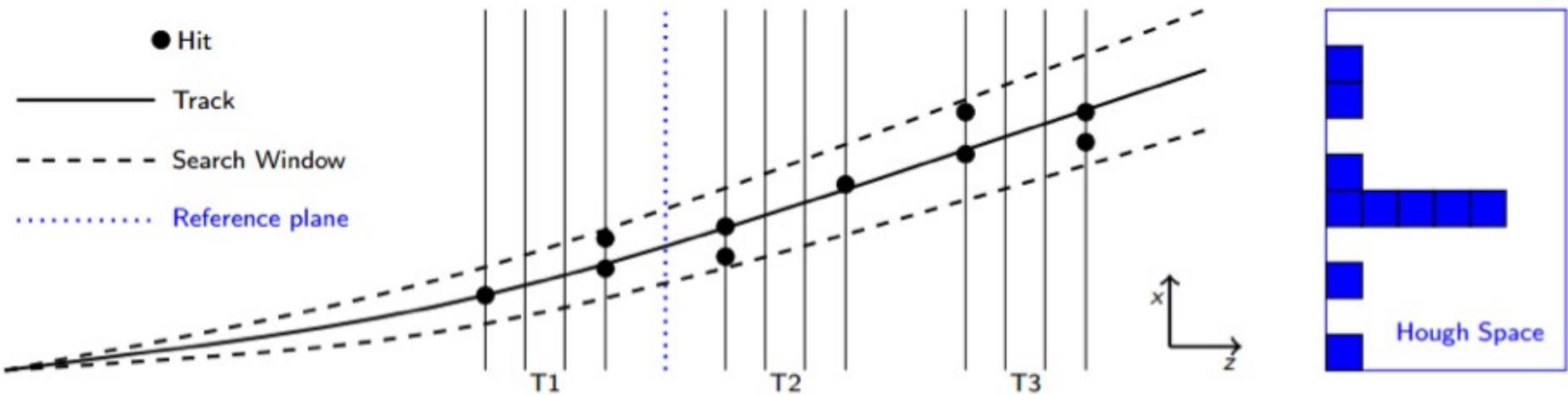
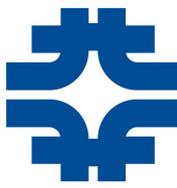


Example of cluster split

Seeding

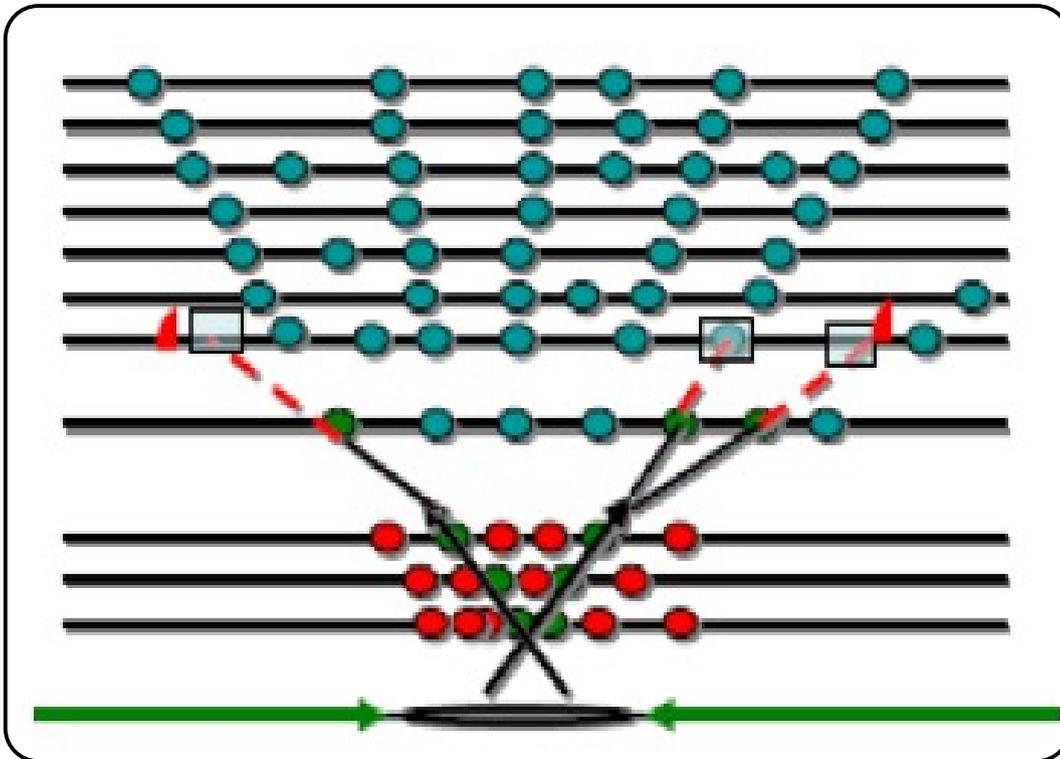


- Combinatorics of 2 or 3 hits with tight/loose constraints to the beam spot or vertex
- Seed cleaning/purity plays an important role in reducing the CPU requirements of subsequent steps
 - Consider pixel cluster shape and charge to remove incompatible seeds
- Initial track parameters from helix fit

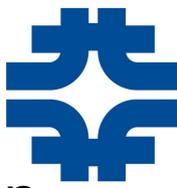


- Project hits within a search window onto a reference plane
- Find clusters of hits
- Done in inner and outer tracker of LHCb thanks to low density of hits

- Use of the Kalman filter formalism with weight matrix
- Identify possible next layers from geometrical considerations
- Combinatorics with compatibles hits, retain N best candidates
- No smoothing procedure
- Resilient to missing modules
- Hits are mostly belonging to one track and one track only
- Hit sharing can happen in dense events, in the innermost part
- Lots of hits from low momentum particles



Kalman Filter



$$K_k = C_{k|k-1} H_k^\top (V_k + H_k C_{k|k-1} H_k^\top)^{-1}$$

$$p_{k|k} = p_{k|k-1} + K_k (m_k - H_k p_{k|k-1})$$

$$C_{k|k-1} = (I - K_k H_k) C_{k|k-1}$$

H_k is the projection matrix

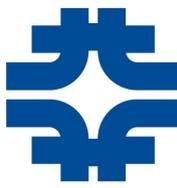
V_k is the hit covariance matrix

$p_{i|j}$ is the trajectory state at i given j

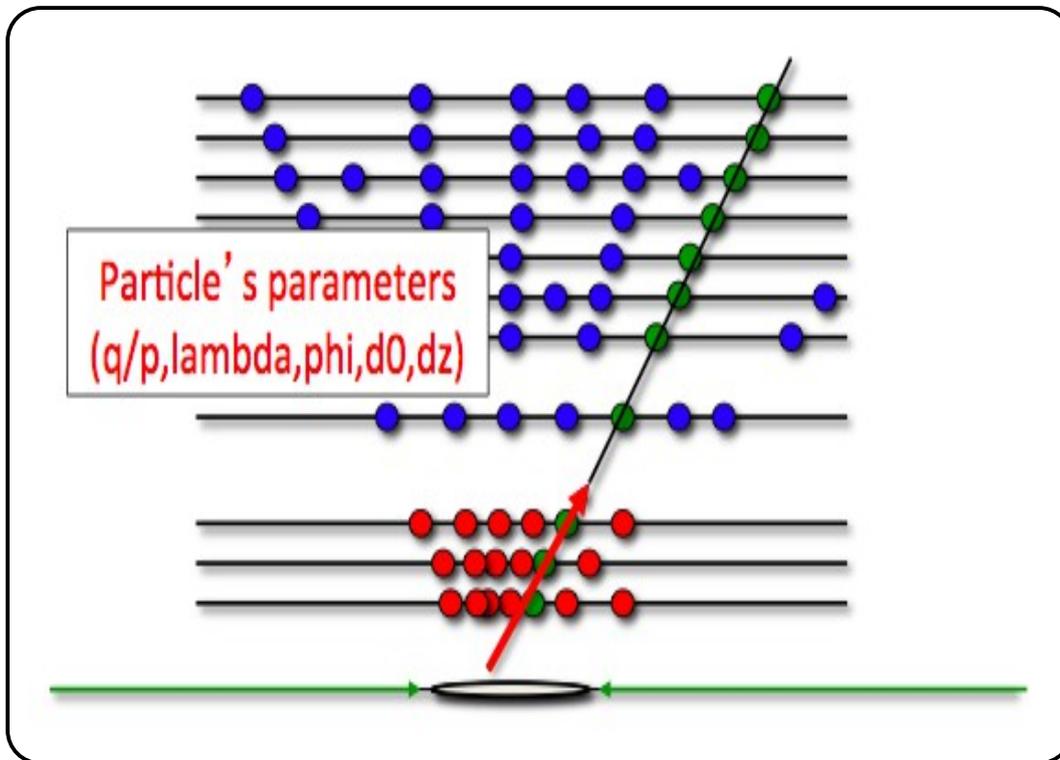
$C_{i|j}$ is the trajectory state covariance matrix at i given j

- Trajectory state propagation done either
 - ✓ Analytical (helix, fastest)
 - ✓ Stepping helix (fast)
 - ✓ Runge-Kutta (slow)
- Material effect added to trajectory state covariance
- Projection matrix of local helix parameters onto module surface
 - Trivial expression due to local helix parametrisation
- Hits covariance matrix for pixel and stereo hits properly formed
 - × Issue with strip hits and longitudinal error being non gaussian (square)

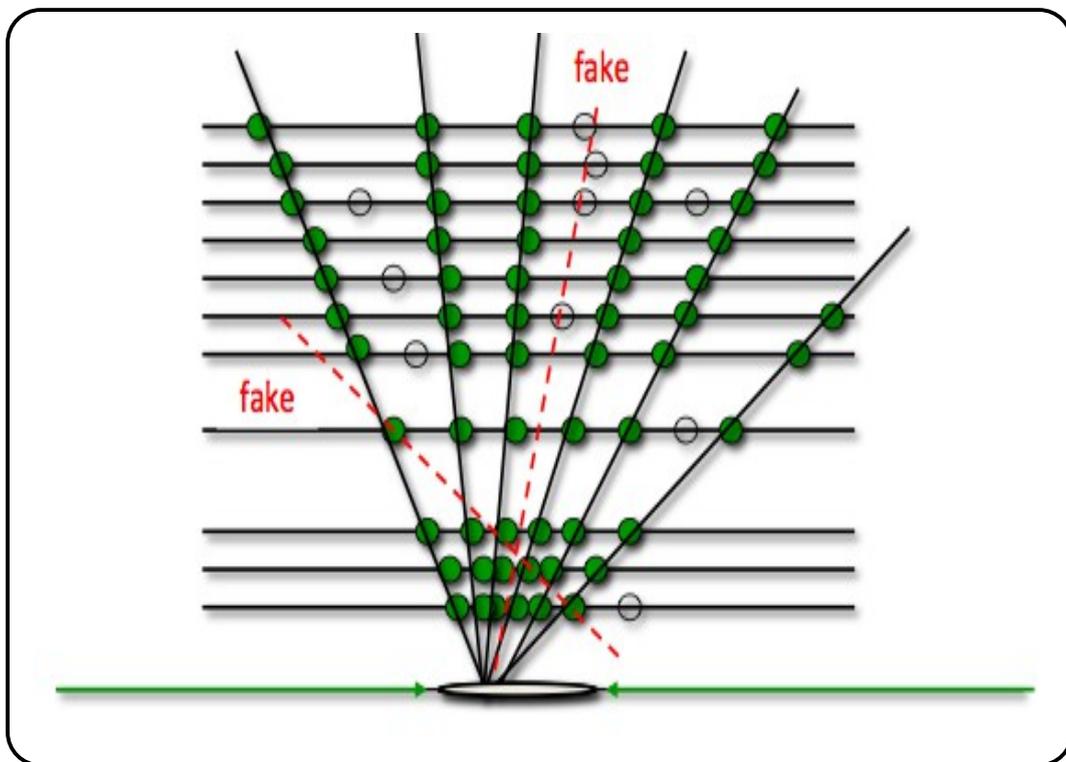
Track Fitting



- Use of the Kalman filter formalism with weight matrix
- Use of smoothing procedure to identify outliers
- Field non uniformity are taken into account
- Detector alignment taken into account



- Track quality estimated using ranking or classification method
→ Use of MVA
- Hits from high quality tracks are removed for the next iterations where applicable
- Efficiency should





Performance



- It is not mandatory to reconstruction 100% of all tracks
 - Tracks within jets are crucial
 - Tracks in very dense jet helps with resolution and identification
- Tracking in the high level trigger could suffer performance degradation in favor of speed-up
- Predominance of low momentum tracks from overlapping collisions
 - No need to reconstruct them all
 - Would be good enough to reject the hits

Other Approaches

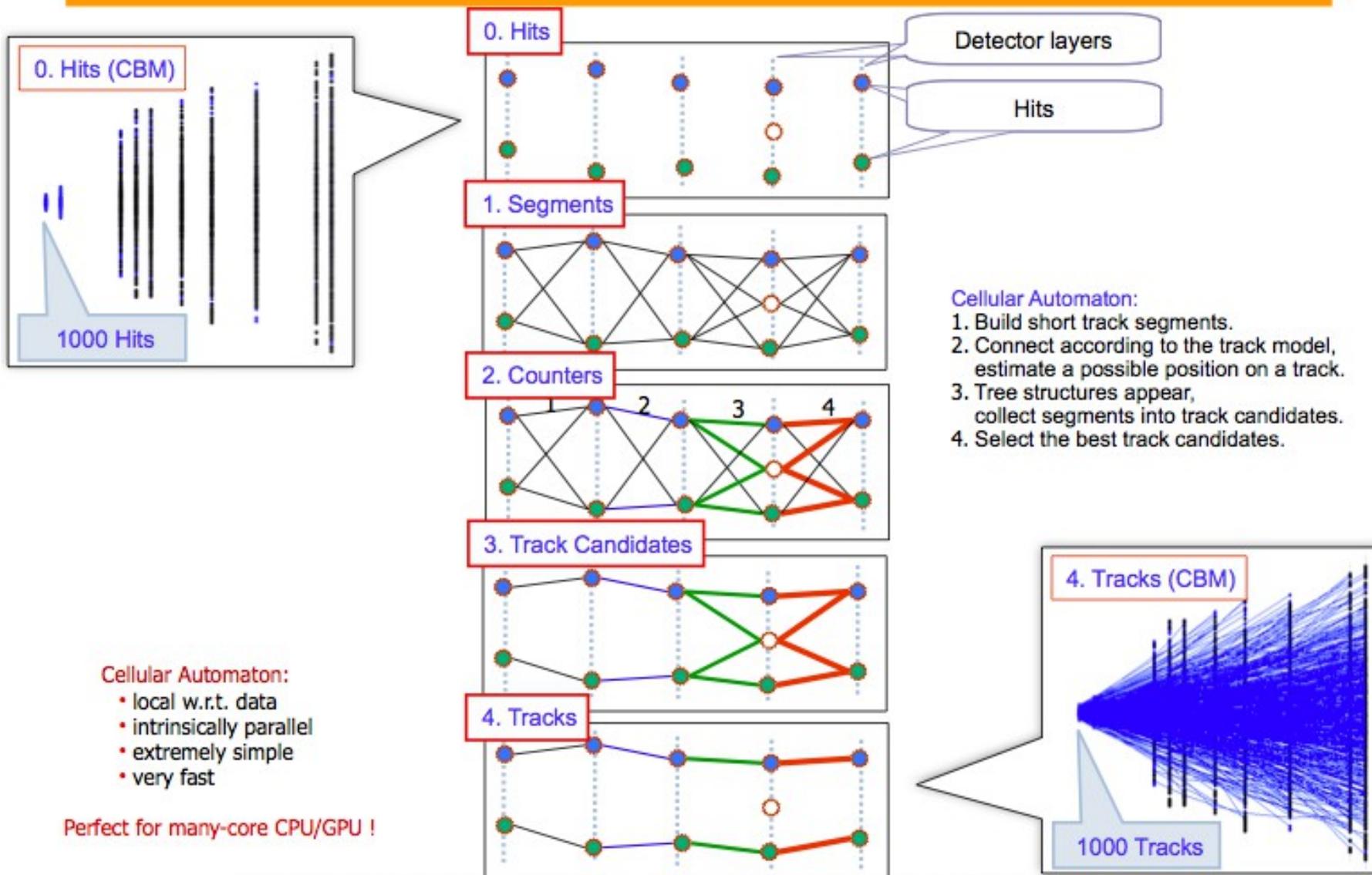
Material from
Connecting The Dots

<https://indico.hephy.oeaw.ac.at/event/86/>
<https://ctdwit2017.lal.in2p3.fr/>

CA at CBM



Cellular Automaton (CA) Track Finder



Useful for complicated event topologies with large combinatorics and for parallel hardware

Associative Memory



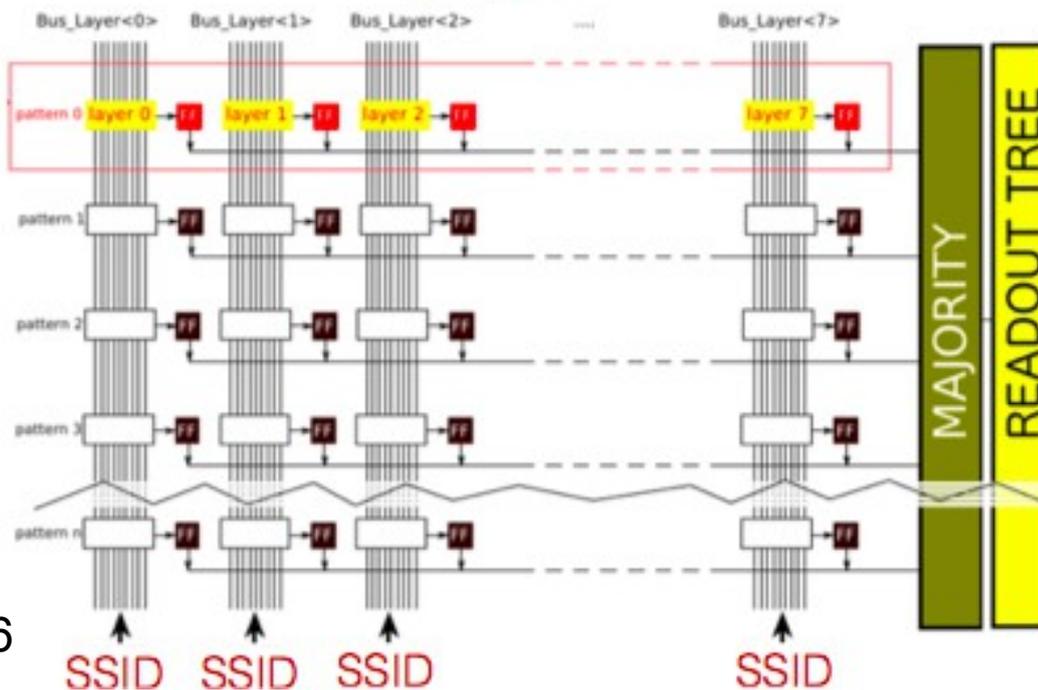
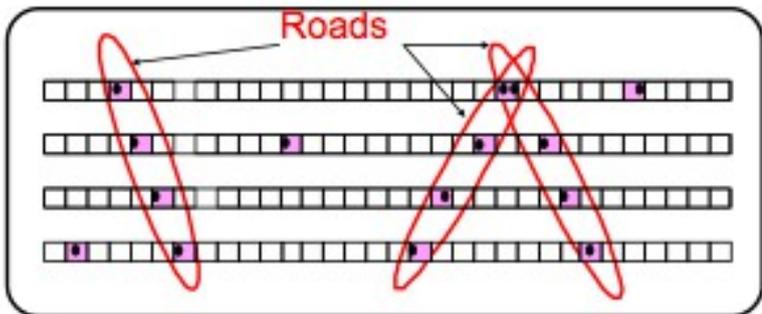
Main feature: avoid hit combinatorics

Configuration

- The combinations of resolution-downgraded stubs (SSIDs) generated by the tracks are stored inside the AM chip

Operation

- The SSIDs of each tracker layers are sent to the AM input buses
- The matched road addresses are sent as AM chip output



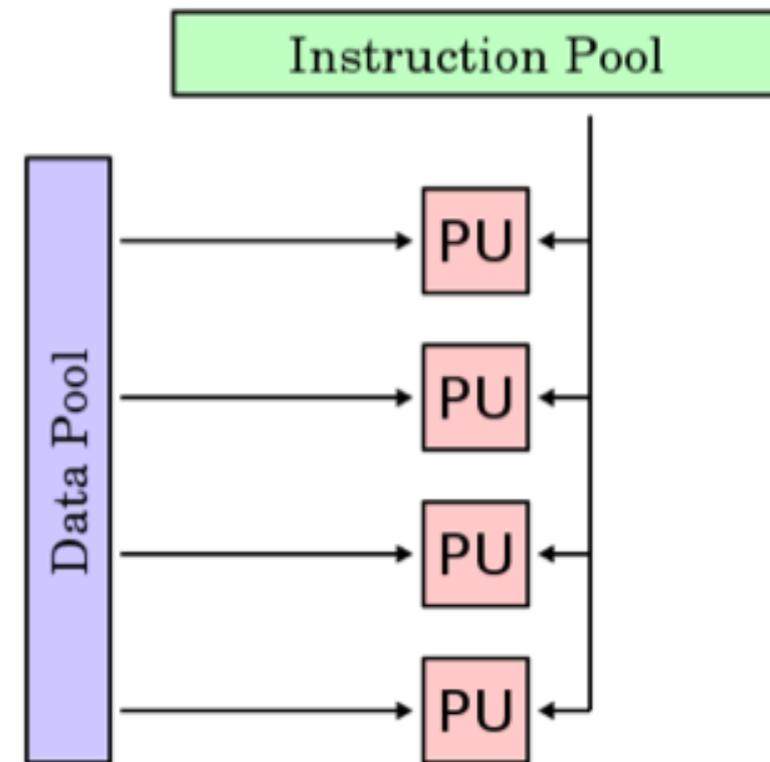
- KF tracking cannot ported in straightforward way to run in parallel
- Need to exploit two types of parallelism with parallel architectures

- **Vectorization**

- Perform the same operation at the same time in lock-step across different data
- **Challenge: branching** in track building - exploration of multiple track candidates per seed

- **Parallelization**

- Perform different tasks at the same time on different pieces of data
- **Challenge: thread balancing** – splitting the workload evenly is difficult as track occupancy in the detector not uniform on a per event basis



Vectorization

Hough Transform



Hough algorithm

Discretised maximum likelihood optimisation over

$$L(n|\{x_i\}) = \sum_i \int dn \delta(d(n, x_i))$$

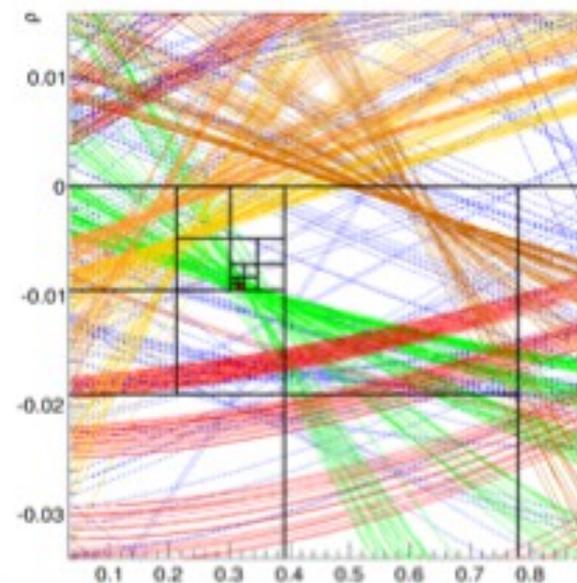
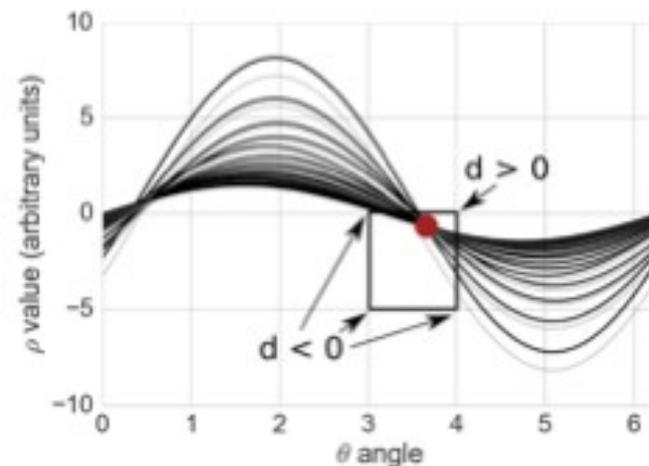
where d is the distance measure of track to hit.
Typically carried out as

- > grid search
- > *Fast Hough* bisecting each dimension

over small volumes dn of the parameter space evaluating only the signs of d on the edges.

Refinements

- > Weighting of hits versus tracks e.g. on distance d or prior distributions
- > Priorisation of search areas
- > Overlapping volumes



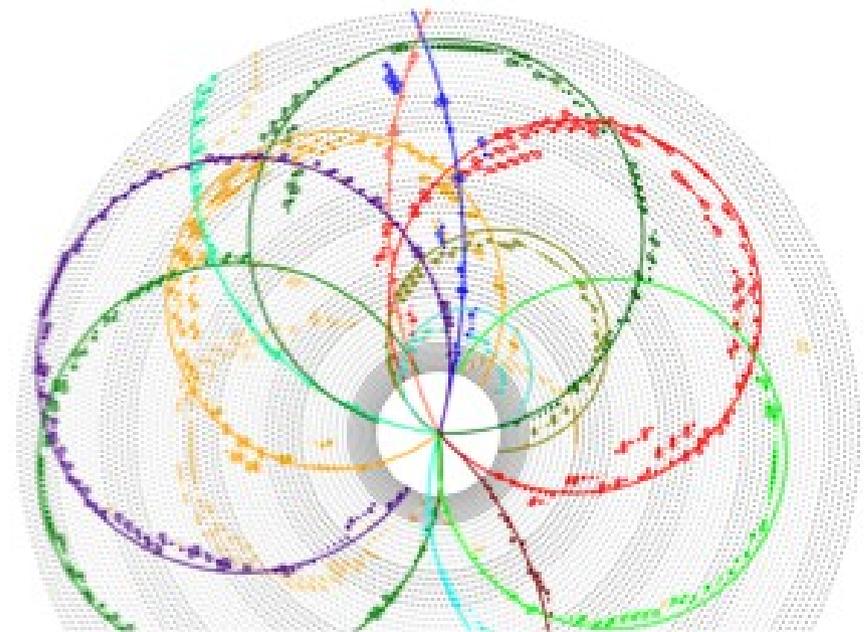
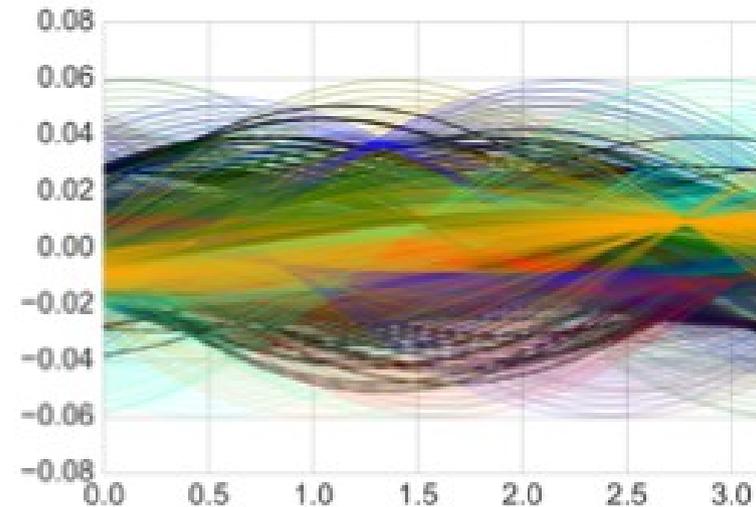
Oliver Frost on behalf of the Belle II collaboration | DESY | 2016-02-22 | Page 4 / 23

Hough Transform



Fast hough search in axial layers

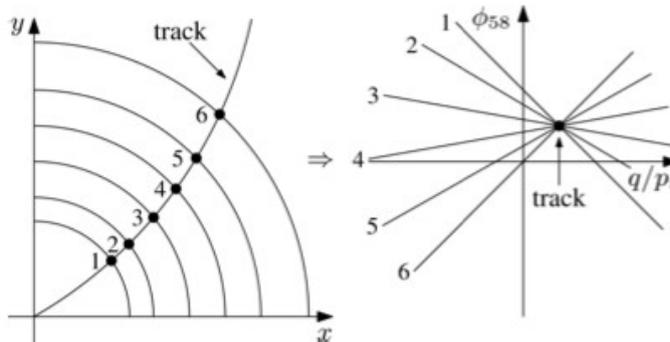
- Working on hits or complete segments
- Multiple passes with priority on high momenta
- Sequential lowering of weight threshold
- Quadtree for structure keeping results of earlier passes
- Sliding bin overlaps pulling towards density centers
- On the fly merging of closeby tracks
- Cleaning and recycling of hits



Oliver Frost on behalf of the Belle II collaboration | DESY | 2016-02-22 | Page 19/23

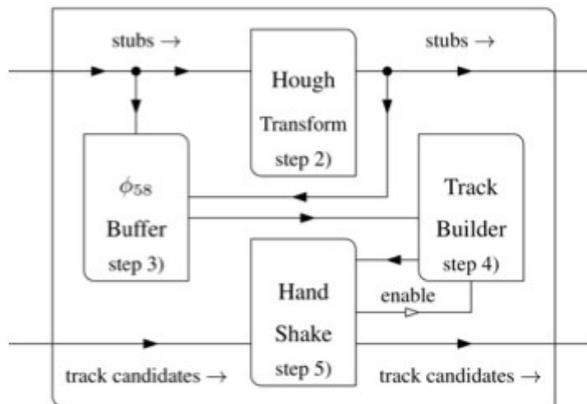
Hough Transform Algorithm

1. Calculate ϕ_{58} (ϕ at r_{58}) for each q/p_T .
2. Fill the stub into appropriate cells in a 32×64 array in $q/p_T \times \phi_{58}$
3. Ignore q/p_T values inconsistent with a stub's bend information (rough p_T estimate).
4. Define cells with stubs in at least 4 or 5 layers as track candidates.
 - i. 4 layer threshold used to cope with barrel-endcap transition region or dead layers



Firmware Implementation - Bin

- Each bin represents a q/p_T column in the HT array



- Hough Transform:
 - Gets ϕ_{58} at left boundary
 - Calculates ϕ_{58} at right boundary
- ϕ_{58} Buffer:
 - Duplicates stubs if it belongs to two cells.
- Track Builder:
 - Sorts stubs in ϕ_{58} cells.
 - Marks ϕ_{58} cells with stubs in at least 4/5 layers.
- Hand Shake:
 - Controls read-out of candidates

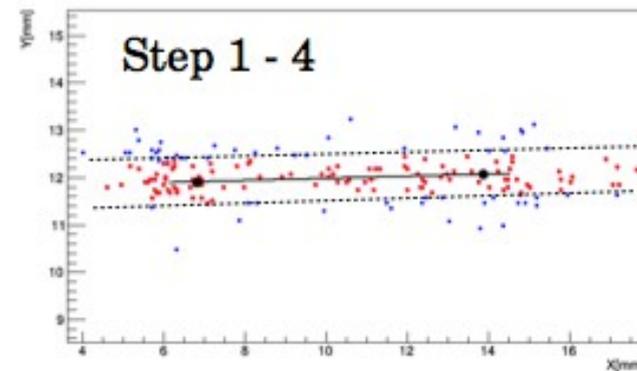
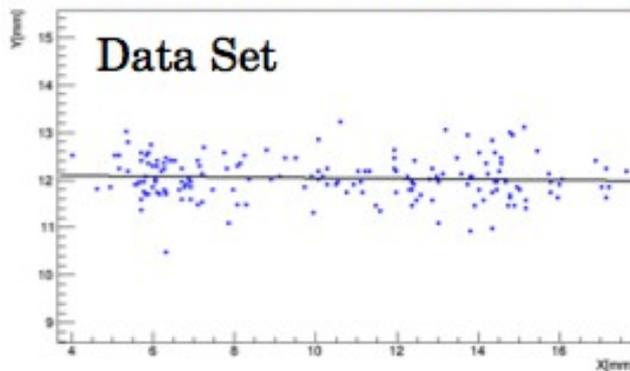


RANSAC

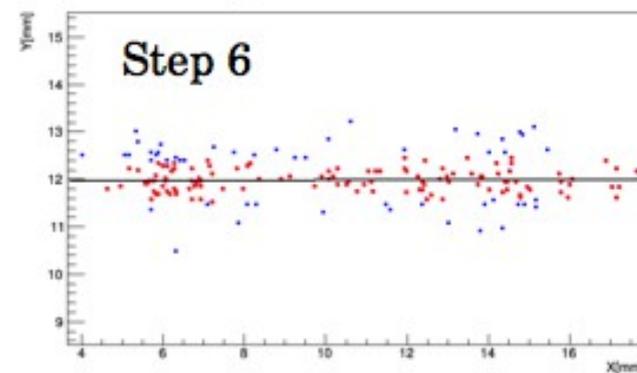
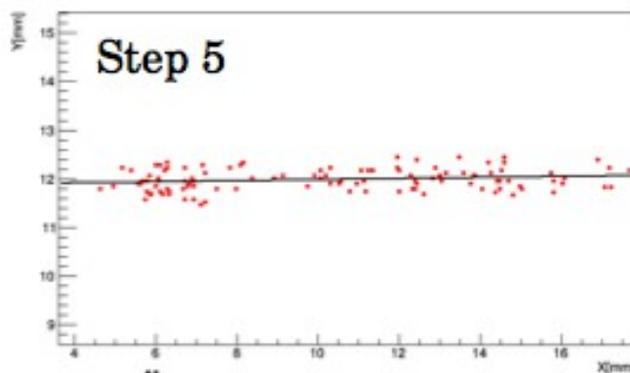


ILC – TPC: RANSAC

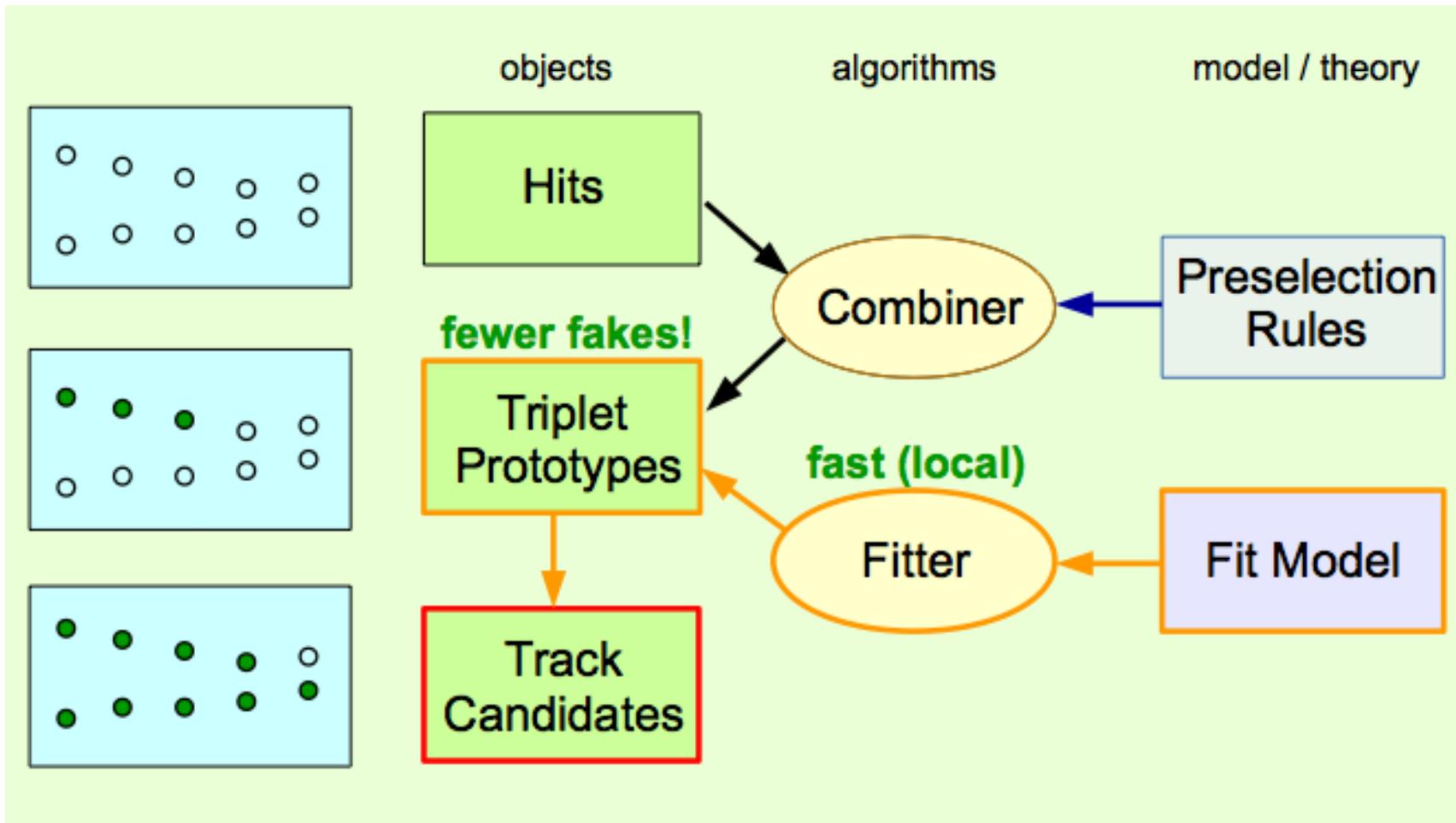
- 1) Choosing two random hits
- 2) Define straight line based on these two hits
- 3) Calculate the distance of all other hits to the line



- 4) Collect those points with the distance less than the tolerance
- 5) If there are enough points in collection then a solution is found. Fit the line using all points in collection.
- 6) Recalculate inliers and outliers points based on this fit and then refit again based on new inlier points.
- 7) Repeat step 1 to 6 N times.
- 8) If more than one solution is found take solution with smallest chi square.

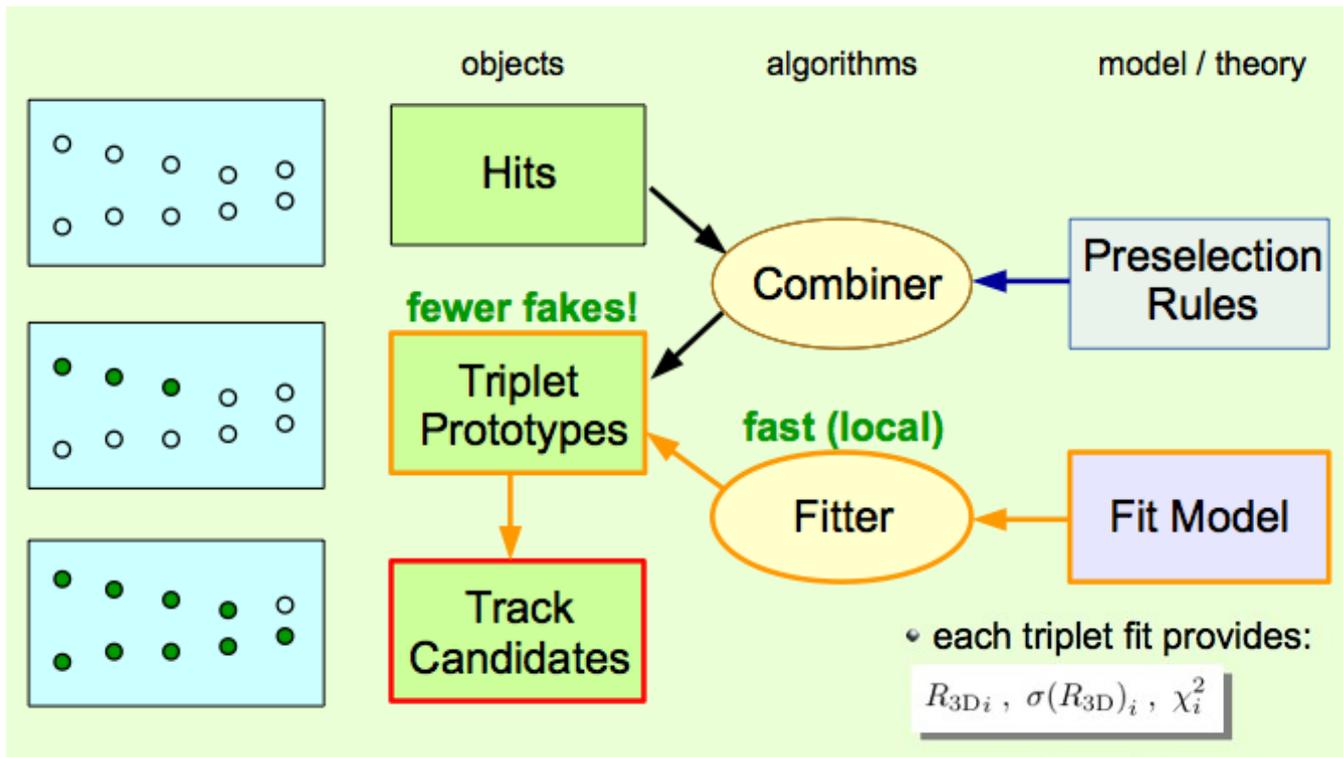


Broken Lines



A.Schöning, Heidelberg University

Broken Lines



- each triplet fit provides:

$$R_{3D,i}, \sigma(R_{3D})_i, \chi_i^2$$

- averaging of R_{3D} (w/o bias correction)

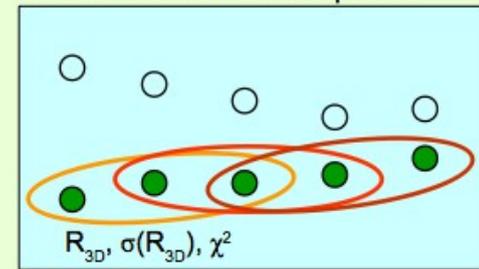
$$\overline{R_{3D}} = \sum_i^{n_{hit}-2} \frac{R_{3D,i}}{\sigma_i(R_{3D})^2} / \sum_i^{n_{hit}-2} \frac{1}{\sigma_i(R_{3D})^2}$$

- combined χ^2

$$\chi_{comb}^2 = \sum_{i=triplet} \chi_i^2 + \frac{(R_{3D,i} - \overline{R_{3D}})^2}{\sigma_i(R_{3D})^2}$$

- extremely simple and fast
- allows to check consistency of triplets "on the fly"
- final selection: $\chi_{comb}^2 < \chi_{cut}^2$

combination of three triplets:

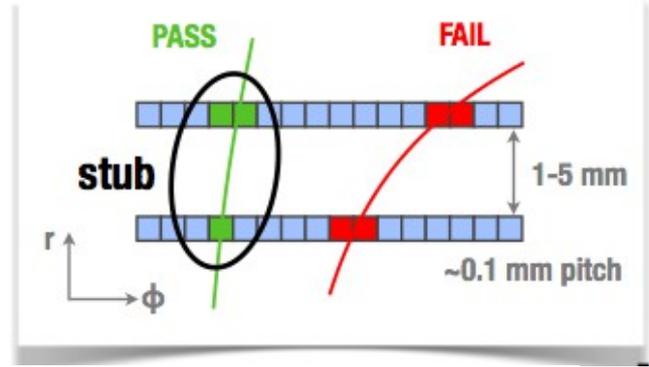


A.Schöning, Heidelberg University

Tracklets

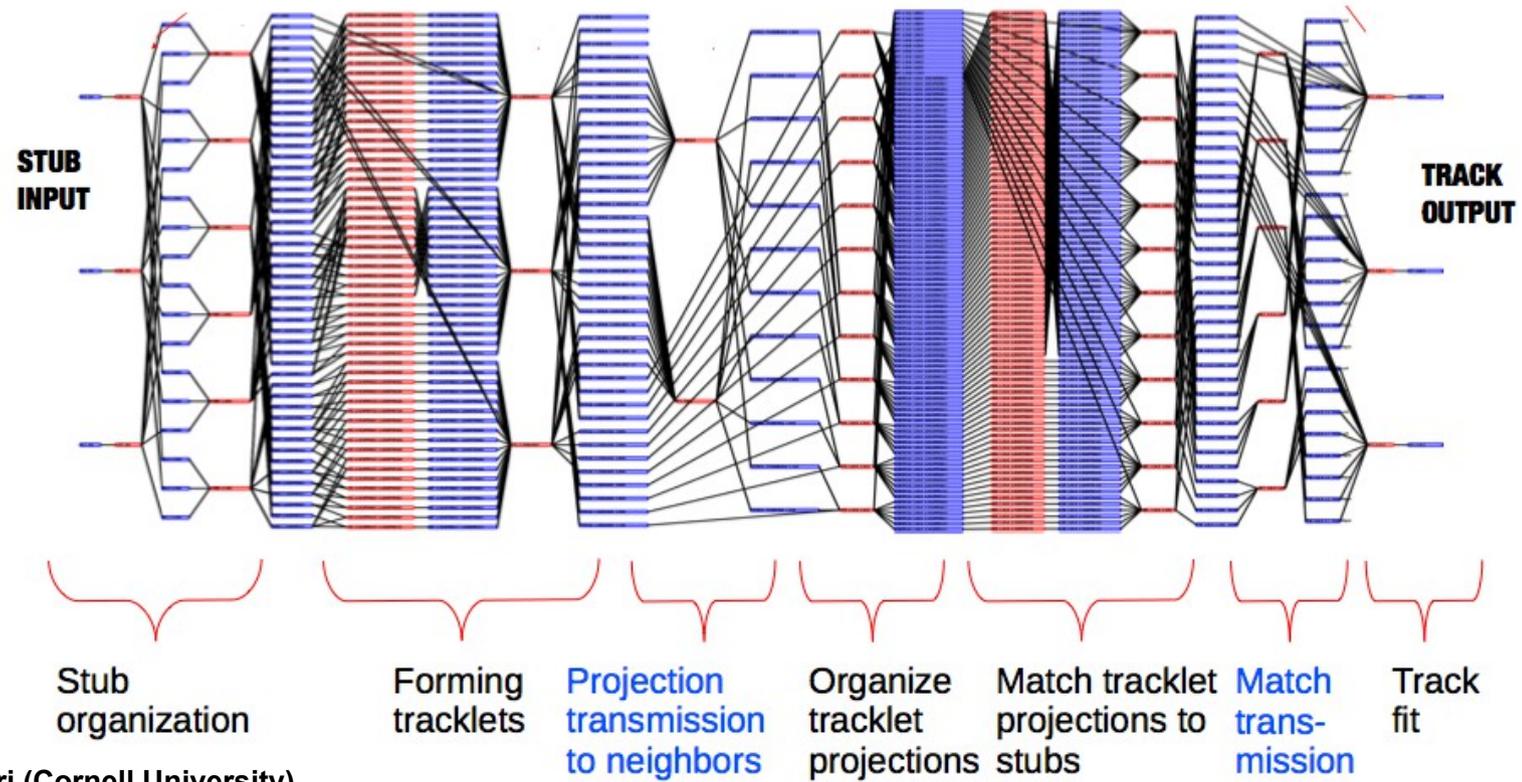


- **Stubs:** Correlated pairs of clusters, consistent with ≥ 2 GeV track
 - ▶ For minbias, rejects $\sim 95\%$ of tracks
 - ▶ Stubs form input to track finding



memories
processing modules

8 processing + 2 transmission steps implements algorithm



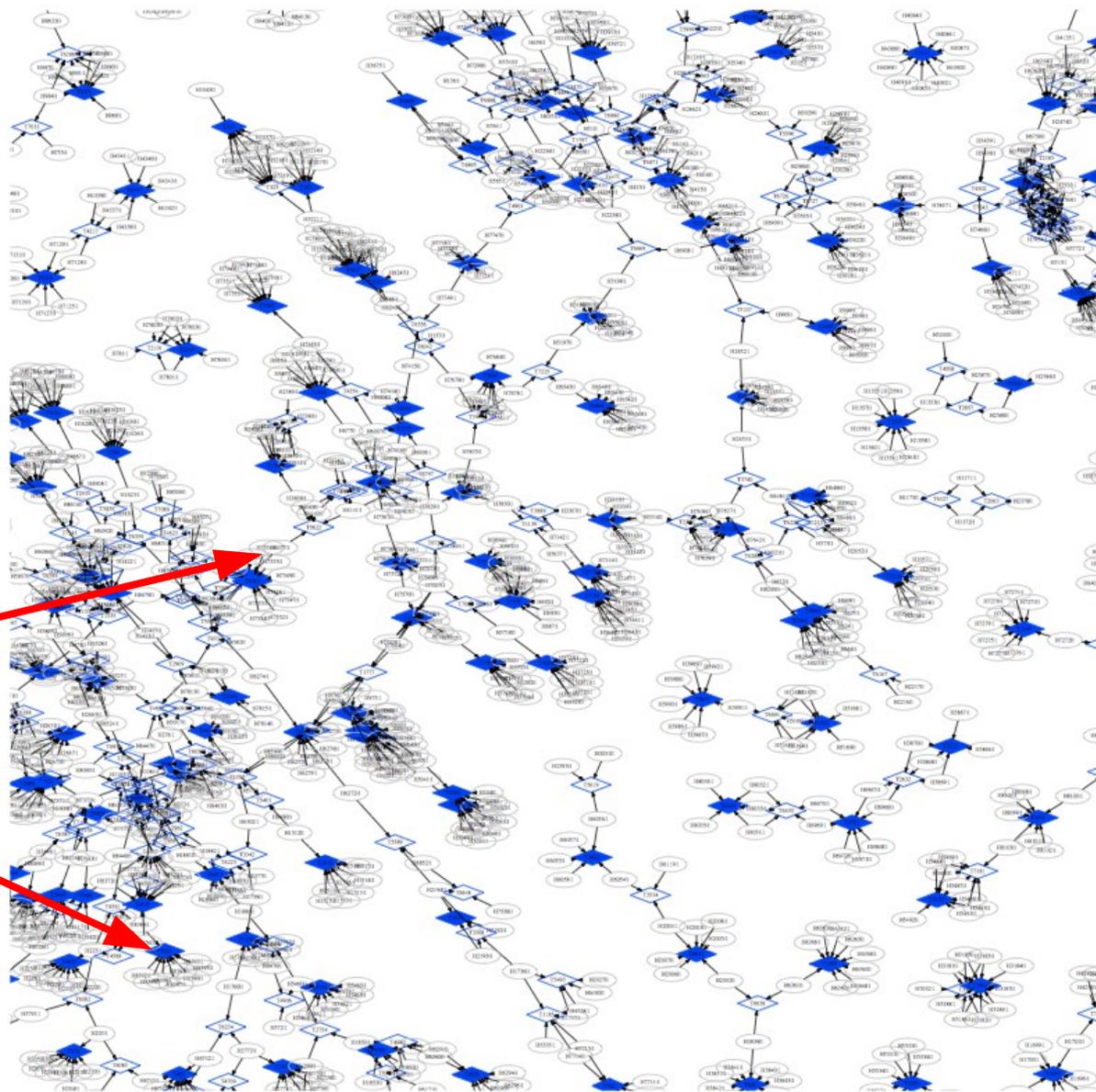
Louise Skinnari (Cornell University)

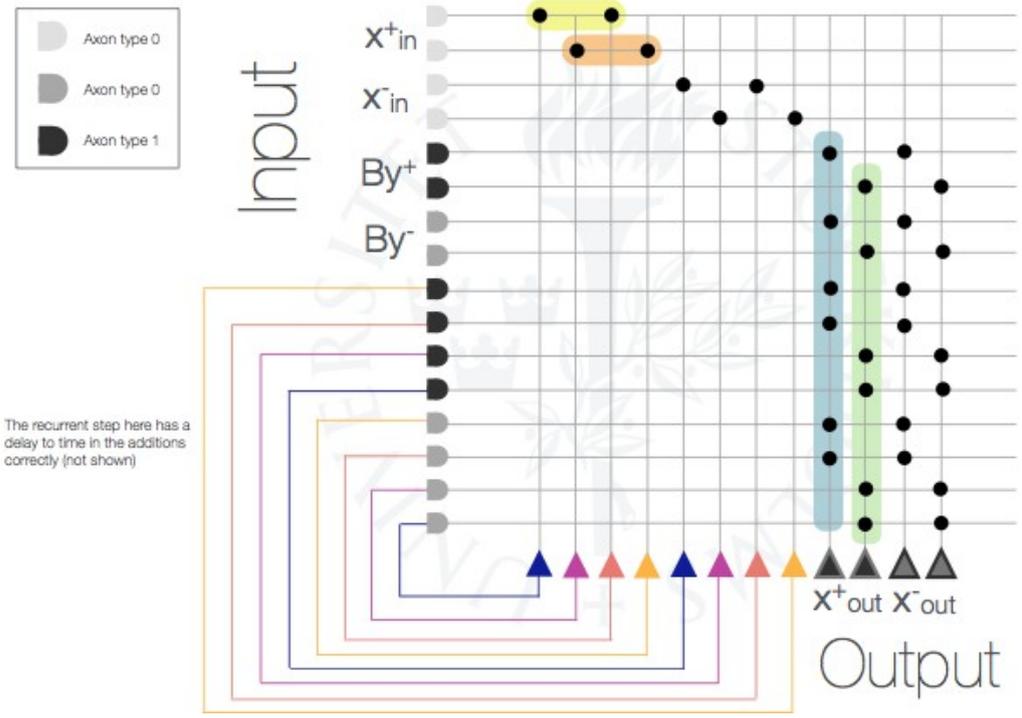
“Disconnecting the Graphs”



Combine hough transform,
image templates and graph
representation to build relations
between track and hits.
Use of decision tree to remove
ambiguities.

Ellipse are hits
Diamonds are tracks

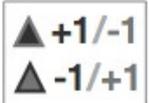




The recurrent step here has a delay to time in the additions correctly (not shown)

Two channels, to deal with positive and negative values as no logic to deal with signed numbers in spikes.

Not shown in this cartoon: logic to combine positive and negative channels after firing.



New computing paradigm
Kalman filter for a simplified tracking
Progress on understanding precision

