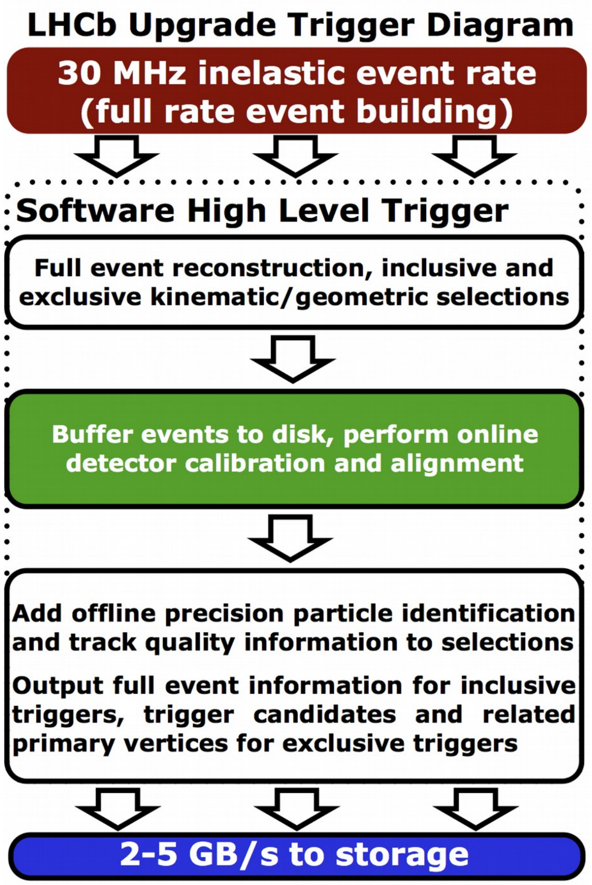
Stand-alone reconstruction and trigger with the SciFi detector

Louis Henry
Milano, 04/10/2019

Introduction: the LHCb trigger system

- Rates of charm and beauty hadron production unmanageable by hardware trigger
  - Need fast event reconstruction
- Additionally, Run 2 has shown that it is possible to reach offline levels of quality at trigger stage
  - Need best event reconstruction
- Current trigger scheme in the upgrade will then follow two steps:
  - **Fast** rec. (HLT1) of long tracks to reduce rate from **30 MHz**
  - **Best** rec. of **all** tracks at **~1 MHz** rate.
- This poses two challenges to any analysis:
  - Can we run the seeding at the Fast level to have a specific selection of our events?
  - We need to make sure we can write trigger lines in the Best level, or else our tracks are lost (no full event saved anymore).



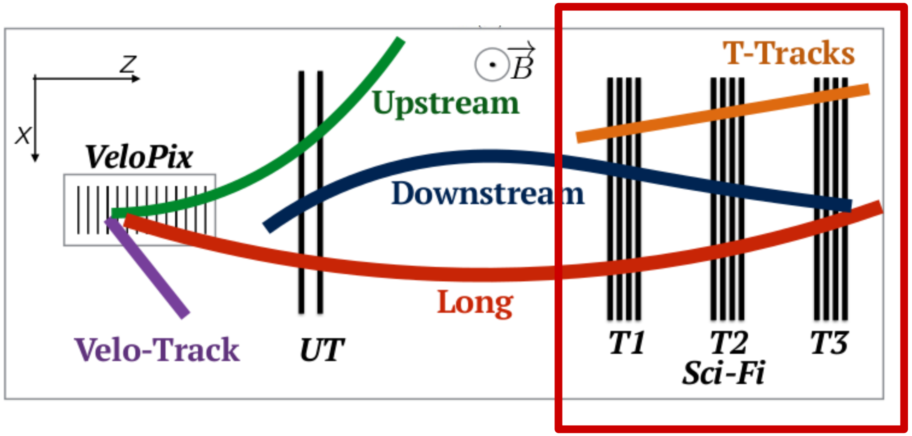
We need to speed-up HLT2-level reconstruction to ensure we can even perform the analysis

+

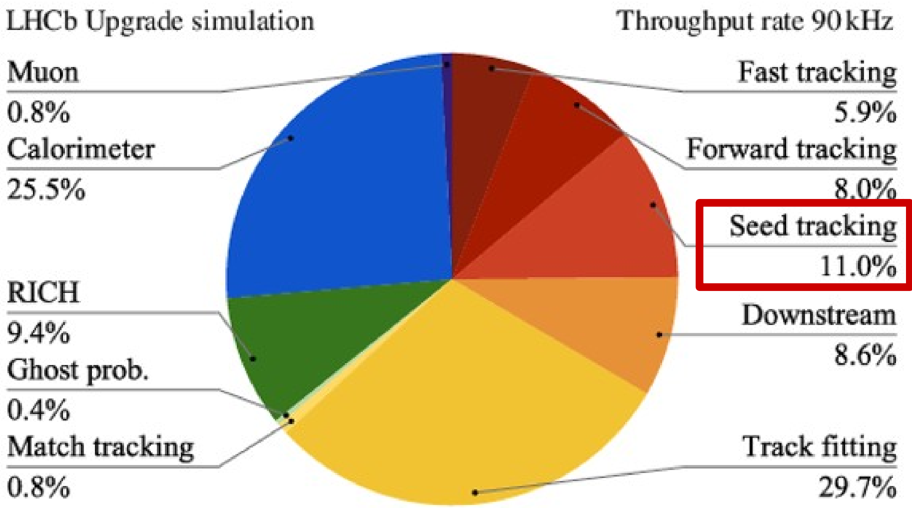
If we could have access to Λ in HLT1, we could boost statistics

Introduction: track types and trigger

- T-tracks = tracks with hit only in the forward-most tracker (SciFi, for Scintillating Fiber).
- They are necessary for:
  - Building some long tracks
  - Building "downstream tracks"
  - For electron PID
  - **By themselves → Λ EDM/MDM measurement**
    - See Salvatore's slides



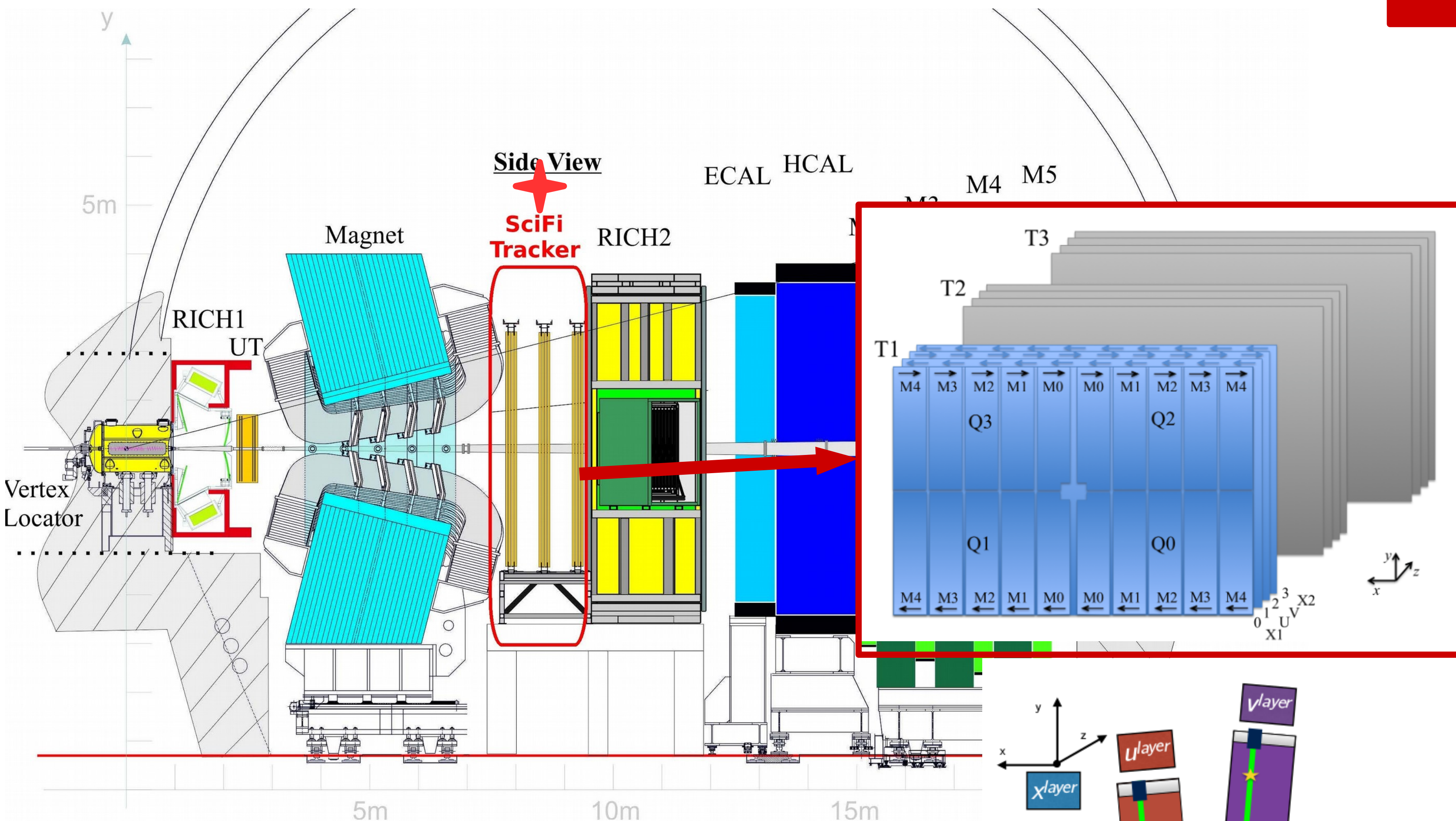
- Reconstruction of these segments needs to be improved and made faster anyway.



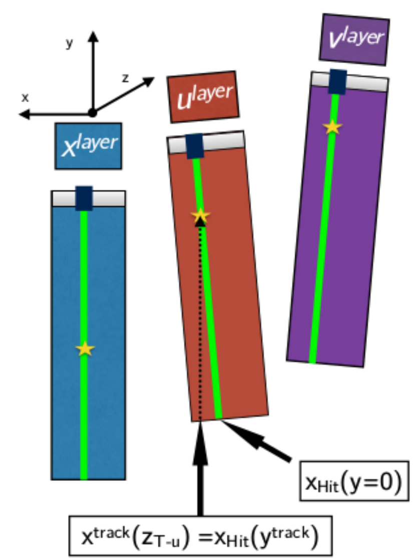
Needs to get down for LHCb programme

Getting stand-alone T-tracks at trigger level is not obvious. This talk will be about the reconstruction, not the selection (much further down the road)

The SciFi detector: overview

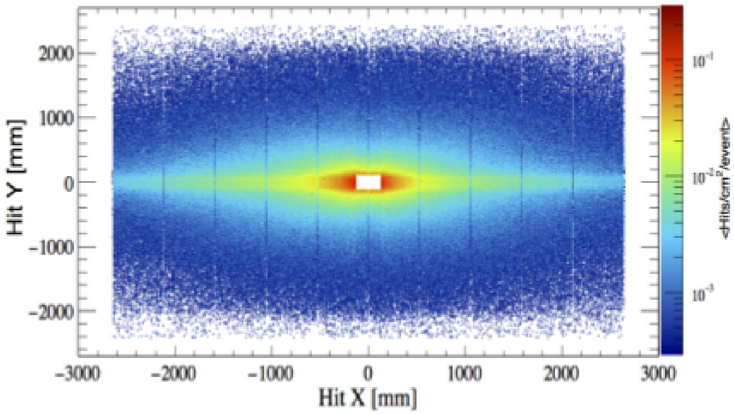


3 **stations** composed of 4 **layers** in x-u-v-x geometry.
X layer: vertical, dy/dx = 0.
U/V layer: vertical, dy/dx = +/- 5º
(B field along y → traces are ~ straight in y)

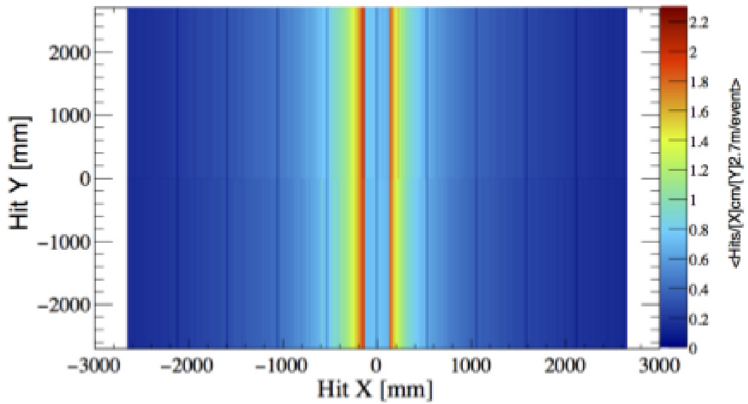


The SciFi detector: overview

- Upgrade to scintillating fibers, each 2.5m long.
  - No information along the fiber



What you think you should see



What you really see

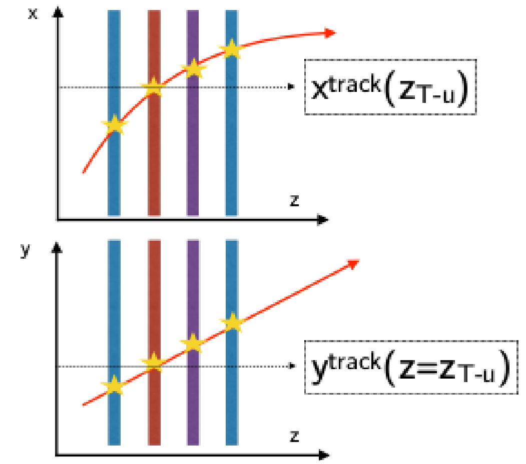
- Need to account for missing hits; residual magnetic field → tracks are not straight.

	Spacial resolution	Hit efficiency	Field
VeloPixel	12 μm	> 99%	Negligible
SciFi	42-100 μm	~ 99%	Significant

Much different environment than VELO reconstruction, which is the benchmark for “fast” reconstruction

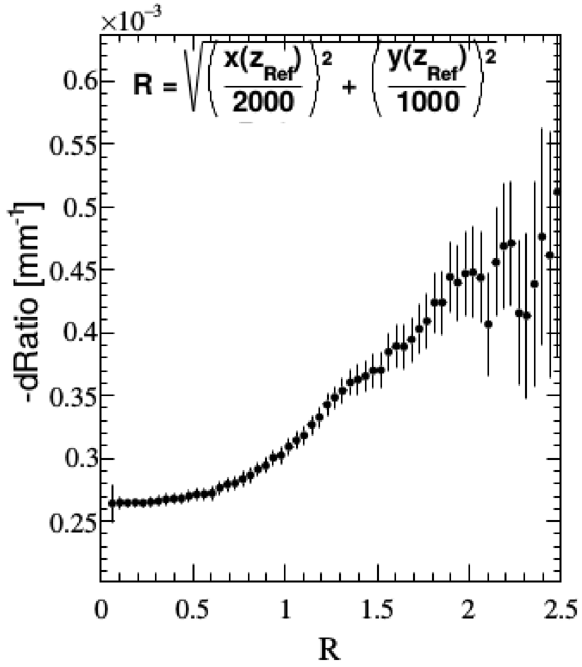
Choosing the track model

- Reconstructing a track gives information about its **(x,y) position as a function of z** and its **momentum** (through curvature).
- Relies heavily on chosen track model. In SciFi, $B_y > B_x$, and so the chosen track model is:



$$x(z) = a_x + b_x dz + c_x dz^2(1 + d_{\text{ratio}} dz)$$
$$y(z) = a_y + b_y dz$$

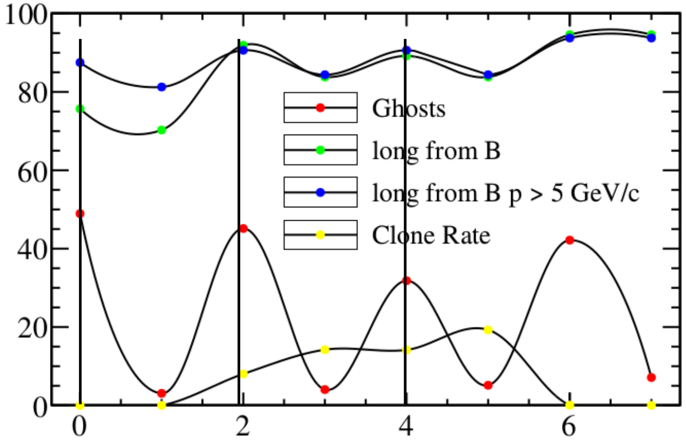
- Note: this c_x parameter also serves as a stand-alone momentum measurement! (~1/P)
- $d_{\text{ratio}} \rightarrow$ cubic correction, depending on stray fields and so track position in the (x,y) plane.
  - Possible to account for this during the full fit.
  - Very simulation-dependent!



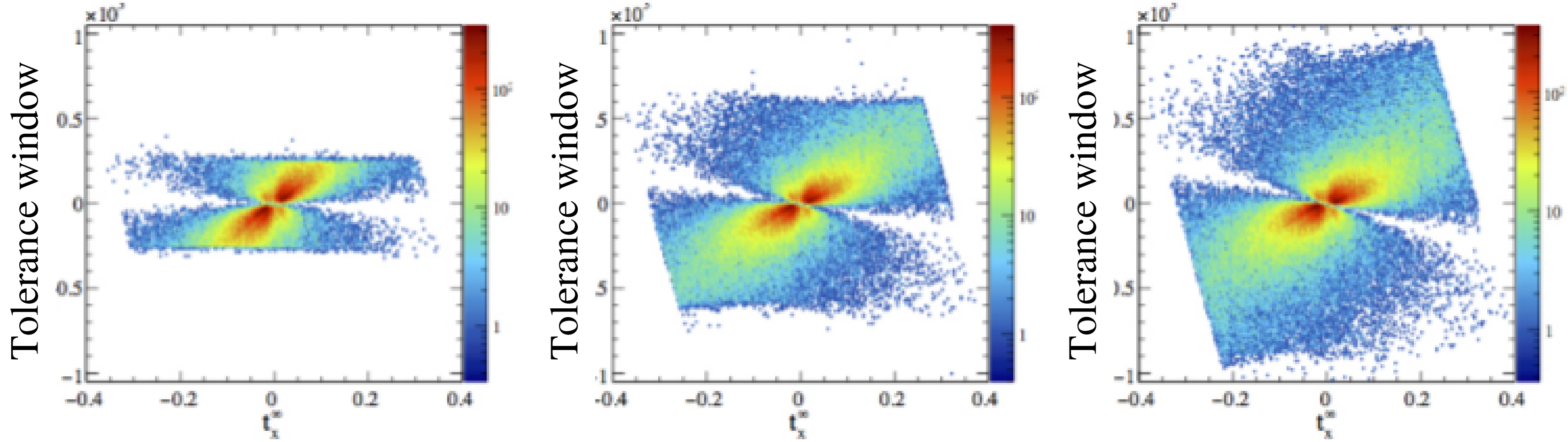
- Possibilities for improvement:
  - Parabolic y model

Stand-alone reconstruction: principle

- Stand-alone reconstruction of a given track will follow the following steps:
  - Looking for hits in x-stations to build a XZ candidate from pairs of hits in first/last layer.
  - Look for compatible hits in U/V layers and build a full candidate
- These steps are repeated 3 times ("**cases**") with different first and last layer (to cover for hit inefficiency) and wider momentum windows.
  - At the end, a recovery step and a clone removal
  - Right: 0 (X-candidate case 0), 1 (adding UV case 0).



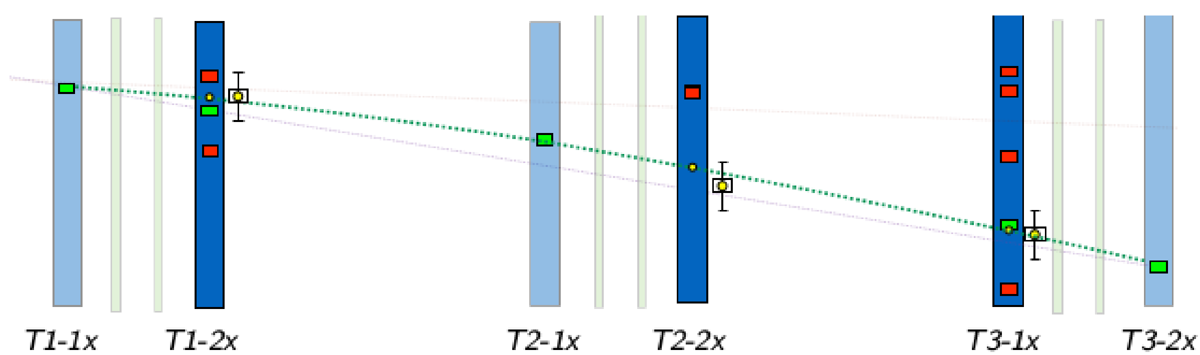
- Below: effect of momentum cuts on window in last layer:
  - Left: p > 5 GeV/c; Middle: p > 2 GeV/c; Right: p > 1.5 GeV/c



Seeding first treats easy (= high-momentum) tracks and then goes to smaller p.

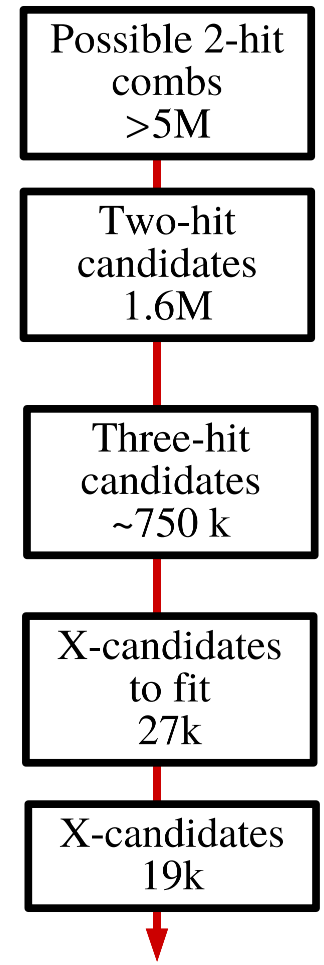
Hybrid seeding: looking for x hits

- Combinatorics are extremely difficult to beat → any bit counts.
- Cascading hypotheses:
  - Take a hit in the first layer.
  - “If it came from the origin, where would it end up?” → projection on the last layer.
  - “Obviously, there is a magnet. In which interval would a particle with a momentum larger than minP end up?” → family of compatible hits in the last layer.
  - “Where would hits in the second station be if it were a line?” → projection on T2.
  - “Wait there is a magnet, and we already know a bit about curvature” → small window
  - For each hit in T2, make a parabola and look for hits in remaining 3 layers. If total > 4, keep candidate and fit it.



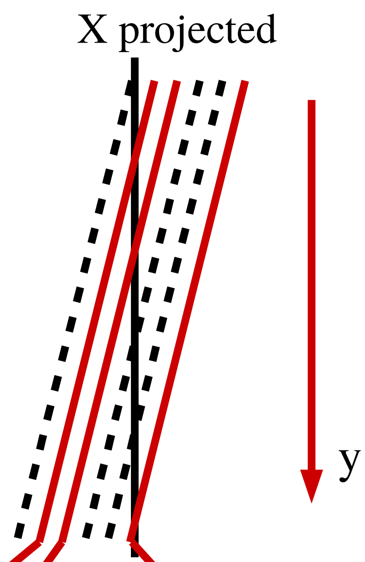
- However, the more hypotheses are made, the more specialised/simulation-dependent the seeding becomes.
  - Ex: “tracks come from the origin” would kill any downstream-track reconstruction.

On 100 events



Hybrid seeding: adding u/v hits

- Each XZ candidate has a projected x position on all 6 U/V layers → each time, several fibers are compatible with that projection
- Collect all compatible hits: they correspond to different projected y coordinates: y(proj) = (xproj – xAt0)/DyDx.
- “Tracks are straight in y and come from the origin” → all U/V hits for a given track candidate should have $t_y = y(z)/z = k$.
- Hough clustering:
  - collect all hits from all layers, put them in a container sorted by t_y.
  - sliding window reading of that container → shows accumulation of hits if there is a track



t_y	0,1	0,2	0,21	0,21	0,22	0,22	0,22	0,26	0,29	0,29	0,31	0,31	0,35	0,36

Hough cluster

- This phase of the tracking is so reliant on the track coming from the origin that there needs to be a dedicated recovery step for downstream tracks.
  - Recovery step: ~10% of the timing → costly.

Hybrid seeding: current state

- Current timing share of the seeding in full reconstruction: 5%
  → **gain of a factor 2 since two months**.
- Still need some gains as the whole trigger is still too slow for normal operations.

- Efficiencies:
  - On normal, long tracks: 91.7%
  - On normal, long tracks with P > 5 GeV: 94.9%
  - On tracks from K_S: 91.8%
  - On tracks from Λ: 89%

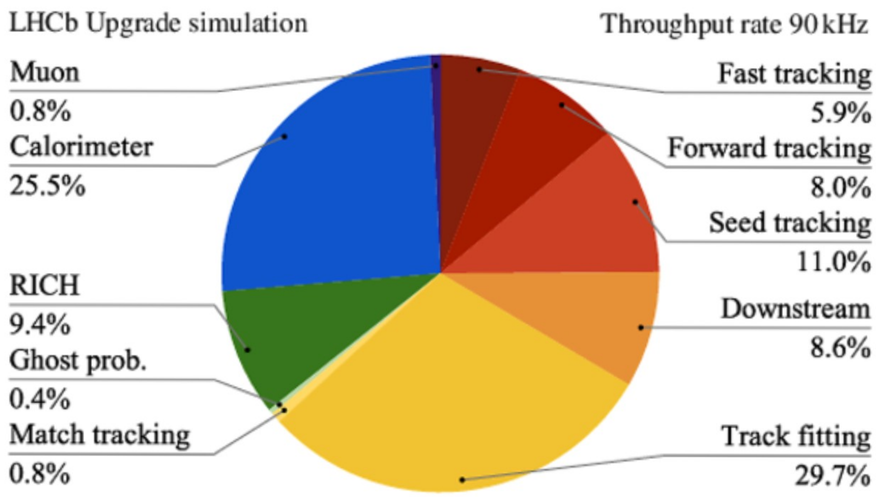
  → Quite good efficiencies on T-tracks from displaced vertices already, still some percents to gain.

- No clear technical obstacle to writing dedicated trigger lines in HLT2 → no roadblock to be able to reconstruct and analyse long-lived particles from K_S and Λ.

Looking for HLT1-compatible reconstruction

- Current reconstruction is barely compatible with HLT2 standards of timing, why even talk about HLT1?
- HLT2 reconstruction needs to be extremely efficient and the highest quality possible.
  - Most of the timing spent on high-hanging fruits
  - Low-hanging fruits (aka 12-hits, high-momentum tracks) are **much** easier to go for.
- Possibility to run a **specialised** seeding much faster
  - High-momentum? No hit inefficiency?
  - HLT1 needs to run at 30k evts/sec/node* running current seeding with only the first case gives roughly 22k evts/sec/node*

  (*: unofficial numbers.)



- Work is advancing in collaboration with other groups, as seeding in the Fast step would greatly benefit downstream reconstruction.
  - Downstream reconstruction → access to Λ baryons and K_S mesons.

Looking for HLT1-compatible reconstruction

- Current work is following 4 directions, each of them with someone working on it.
- **Changing the tuning of the algorithm.**
  - Less simulation-dependent, improve efficiencies and timings.
  - Allow for the creation of extra cases with different momentum thresholds if needed.
- **Changing the Hough Cluster approach**
  - Current Hough cluster needs sorting and then iterative fit with hit removal → takes a third of the timing.
  - U/V hit adding is not very different from VELO reconstruction (straight line) → can draw inspiration.
- **Changing inputs of the algorithm**
  - SciFFit hit classes are being reevaluated in LHCb, as many attributes are layer-dependent (e.g. z position, dyDx).
  - Allows to optimise cache locality.
- **Vectorise the algorithm**
  - Most operations performed are rather simple, costly because of shere combinatorics.
  - Vectorised algorithms deal with combinatorics differently so it could allow to loosen tunings.

Conclusion

- Stand-alone T-track reconstruction is necessary both for the physics programme of LHCb and for our EDM/MDM projects.

- In the upgrade, it is ran at the trigger level, with different challenges:
  - "Best" level: must improve timing by some factors, could improve efficiencies.
    - Current thoughput of the "best" sequence: ~100Hz. Must be 1kHz.
  - "Fast" level: must improve timing by at least a factor 2, possible to setup a specialised version of the reconstruction.
    - Typical throughput needed: 30 kHz, currently a dedicated, unoptimised version is at ~23kHz.

- Two-pronged effort: we need "best" reconstruction to make our analyses even possible in the upgrade… but if we could run reconstruction at "fast" level, much better potential in the upgrade than Run 1 and 2.