UNIVERSITY OF BRISTOL / LAL ORSAY

Assessments

Renato Quagliani
June 9, 2015

1 OVERVIEW OF WORK DONE

A description of the work done up to now will be described. It consists in two major blocks. The first one is related to the analysis work and it has been done between June 2014 and October 2014 , period in which i was finishing the master degree in Ferrara Unviersity and i get the NPAC master 2 at Paris Sud University. The second one is related to the work done in the context of the preparation for the upgrade of the LHCb detector consisting in the improvements of the *Seeding* tracking algorithm which actually becomes the development of new algorithm , the *Hybrid Seeding*. Before discussing the work done in this domain a small introduction about the tracking and the upgrade of the *LHCb* detector is mandatory.

1.1 INDTRODUCTION TO LHCB AND TO THE UPGRADE SCINTILLATING FIBRE TRACKER (SCIFI).

The *LHCb* detector [1] aims at search indirect signatures of New Physics through quantum loop induced processes through the measurements of strongly suppressed Standard Model processes. In this context precision plays a fundamental role. Up to now, in *LHCb* and other *LHC* experiments, no strong deviations from the Standard Model (SM) have been observed and even if *LHCb* has already provided the world's best measurements in some channels, it is still limited from improving them by statistics rather than systematics. Due to this, the upgrade of the *LHCb* forward spectrometer is mandatory in order to collect *O(100)* times more data and reduce the statistics uncertainty of a factor 10 in order to be comparable with the theoretical one [2] [3].

The *LHCb* detector is designed to be a single-arm forward spectrometer aiming to detect particles and their decay products, and as the b in the name of the detector suggests, it is designed to study particles containing b and c quarks which are produced strongly boosted in the forward and backward (lost) directions for symmetric energies of colliding protons.[1] The detailed description of the *LHCb* detector can be found in [1]. The data taking at *LHCb* during 2011 and 2012 at LHCb are mainly determined by few steps:

- Interesting events are selected by the *L0 Trigger* which is implemented at the hardware level aiming to reduce the 40 MHz bunch crossing rate to 1MHz making use of estimations and measurements of the signature of particles having high E_T, p_T through the muon stations and the calorimeters. The main reason why this is done is because the read-out system of some of the detector for Run-I and Run-II can afford an incoming rate of 40MHz. In the upgrade infact, all the read-out will be substituted and the *L0* trigger will be replaced by a software one.
- *High Level Trigger*: It consists in an *Online* software trigger where the full reconstruction of tracks is performed. It's at this level that the tracking algorithms are run. After this step , data are stored and an *Offline* reconstruction is also performed before providing usable object for data analysis. [2]
- *LHCb* luminosity is kept constant and it's reduced wrt other *LHC* experiments of 2 orders of magnitudes. This reduction in luminosity is achieved thanks to the *Luminosity Levelling* mechanism which avoid head-on collisions separating beams perpendicularly to the collision plane. This reduction of luminosity is mandatory since the main studies on b and c hadrons require an extraordinary precise reconstruction of the production vertex of the $b\bar{b}$ pairs, so, the VErtex LOcator (*VELO*) can be placed at very small distance from the interaction point limiting problems coming from radiation damages. [3]

Regarding the upgrade, a brief description on what it consists is mandatory. A small recap is given in table 1.1 where ν stands for the average number of visible interaction per bunch crossing.

[1] The relevant process infact is the quark-gluon fusion which can be obtained in proton proton collisions only with strongly asymmetric PDFs.

[2] During the Run-I, the seeding algorithm (called *PatSeeding*) in the *HLT* was run making use of the left-over hits coming from the *Forward* algorithm. During Run-I the *Seeding* was used in the online reconstruction in tandem with the *Forward* as described before while in the offline reconstruction it was run as a *Standalone* algorithm. For the Run-II online and offline reconstruction gives more or less the same results, but, anyway, the seeding algorithm remains one of the most important one.

[3] The decrease from the maximal designed luminosity of *LHC* of $10^{34}\mathrm{cm}^{-2}s^{-1}$ to the *LHCb* one $10^{32}\mathrm{cm}^{-2}s^{-1}$ permit to reduce the average number of inelastic collisions from 27 to 0.53 and it allows to reconstruct with extraordinary precision the primary vertices.

	Current *LHCb*	Upgrade *LHCb*
Trigger	Hardware(*L0*) + Software(*HLT*)	Only Software
Read-Out rate	*L0*: 40 *MHz* to 1*MHz* for readout	40 *MHz* Full software trigger for every 25 ns bunch crossing
Luminosity	$4\times10^{32}\mathrm{cm}^{-2}s^{-1}$	$2\times10^{33}\mathrm{cm}^{-2}s^{-1}$
ν	2	7.6

Table 1.1: Main differences between actual LHCb and the Upgrade

The reason why the *L0* trigger will be removed in the upgrade is mainly because lot of analysis looking for deviations from Standard Model at *LHCb* are limited by statistics and a fixed 1 *MHz* readout at the upgrade running condition will be too limiting. In order to reach the physics goals the *LHCb* detector will be upgraded and the installation of the new detectors and the new read-out is expected to happen during the long shutdown 2 in 2018/2019. Each subdetectors will be replaced as shown in Fig. 1.1.

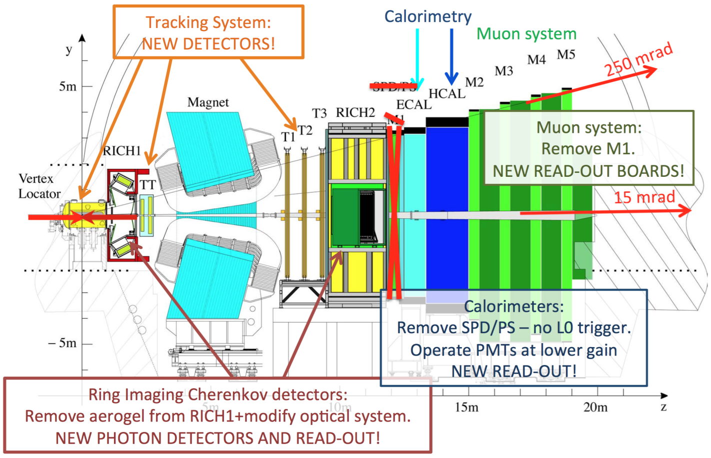

Figure 1.1: LHCb detector modifications for the upgrade. The Vertex Locator (*VELO*), the Tracker Turicensis(*TT*) and the 3 T-Stations are the main resposible of the tracking at LHCb. The Particle Identification is ensured by the two ring cherenkove detectors (*RICH1* and *RICH2*). The particles energy measurement is provided by the hadronic and the electromagnetic calorimeters (HCAL and ECAL respectively). The properties of the muons are then mainly determined by the Muon System.

The actual *VELO* will be replaced by a lightweight hybrid pixel detector capable of 40 MHz readout at the upgrade luminosity which is 5 times greater than the actual one. The trackers downstream and upstream the magnet will be also completely replaced. In particular, the tracker downstream the magnet which is now made by the so-called Inner tracker (*IT*) and the Outer Tracker (*OT*) will be replaced by the scintillating fibre tracker (*SciFi*) detector to cope for the higher luminosity, higher occupancy and the 40 *MHz* read-out.

The current tracker system downstream the magnet implement two different technologies: the Outer Tracker system is based on gaseous straw tube detector for a global resolution on the bending plane (*x-z*) of around 200 μm while the Inner Tracker is made of Silicon microstrip detectors (also employed for the *Upstream Tracker* upgrade). The main reason why it's necessary to replace the *IT* and *OT* is related to the occupancy being too high in upgrade conditions and the fact that the electronics for them was designed for a 1MHz read-out rate. On top of that, the upgrade phase of LHCb is designed to collect an integrated luminosity greater than 50 fb^{-1} and the detector itself is required to be resistent to the corresponding radiation damages.

The adopted solution for the upgrade is the *Scintillating Fibre Tracker*. The active and light-transport (also wavelenght shifter) material for the detector are the fibres themselves and the read-out is provided by arrays of silicon photomultipliers. The pitch of a single channel is designed to be equal to 250 μm and each module of the detector consist of 5 (6 in central region) closed packed fibres ($\Phi\sim250\mu m$) layer.

The *SciFi* is placed in the downstream region with respect to the magnet (*B* field lines direction are along the y direction) consists of 3 stations. Each of them is composed by four layers of scintillating fibres detectors oriented in the so-called stereo x-u-v-x configuration.[4]. The stereo configuration allows the reconstruction of tracks using their projection in the bending x-z plane and the evaluation of the y information is obtained through the combined measurement of the u-v layers.

Each layer is divided in two halves (roughly y>0 and y<0) equipped at y=0 by a mirror to improve the light yield in the higher occupancy area (the region close to the beam pipe).

A skematic view of how a single station look like is given in 1.2 and it will consists of around 10.000 *Km* of fibres and 560.000 read-out channels distributed over the 12 layers (6 x, 3u, 3v).

The requirements for the *SciFi* to satisfy the physics program of *LHCb* can be listed in few concise points:

- The hit efficiency has to be ~ 99%
- The noise cluster rate should be less than 10%, reason why the *SiPM* read-out has to be cooled

[4]The x - layer contains fibres running perpendicularly to the x-z plane while the u(v) layers are exactly the same as the x-layers but rotated by a "stereo" angle of +5(-5) degree in order to reconstruct also the y-information of the tracks.

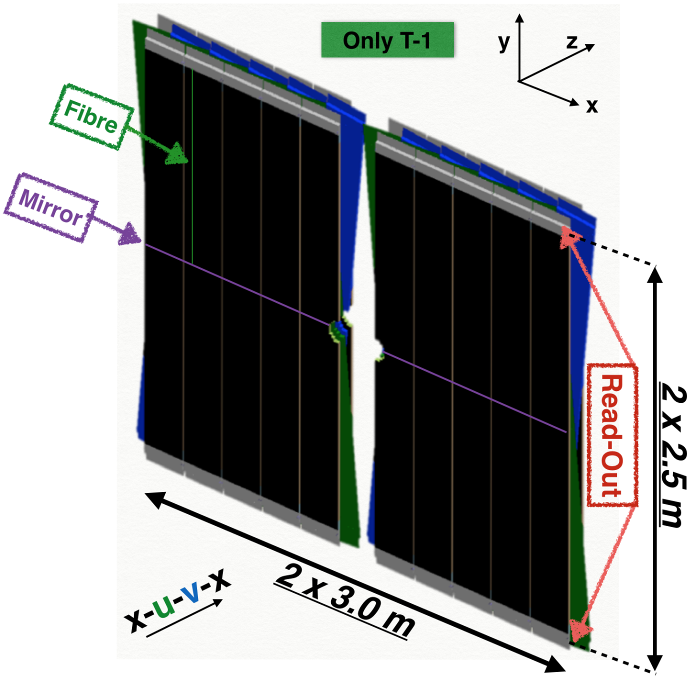

Figure 1.2: Sketch of a single station of the SciFi tracker. The Station is made of 4 different layers in x-u-v-x configuration, the central region is mirrored and the read-out is done at the edges of the layers far from the beampipe.

- The position resolution in the bending plane has to be ~ $100\mu m$ and the material budged has to be as reduced as possible $\frac{X}{X_0} < 1\%$ per detector layer
- The readout has to be performed at 40 MHz
- The tracker has to be efficient for at least an integrated luminosity of 50 fb^{-1}
- The *SciFi* tracker has to replace the actual one, so additional goemetrical constraints must be satisfied.

1.2 Pattern recognition at LHCb

The track reconstruction at *LHCb* is performed in different steps. The idea is to firstly reconstruct tracks collecting list of hits from sub-detectors following some given criteria (track model in a subdetector, expected behavior of tracks in between one detector and another). Each of the found track is then stored in a temporary container and a global fit and finer treatment is performed afterwards (*Kalman Filter* and *Ghost and Clone killer*). We will then refer to the first step as *pattern recognition* and the second one as the *final fit*.

The pattern recognition at *LHCb* is done depending on the path the track goes through, so it's based on the datector's hit content as shown in Fig. 1.3.

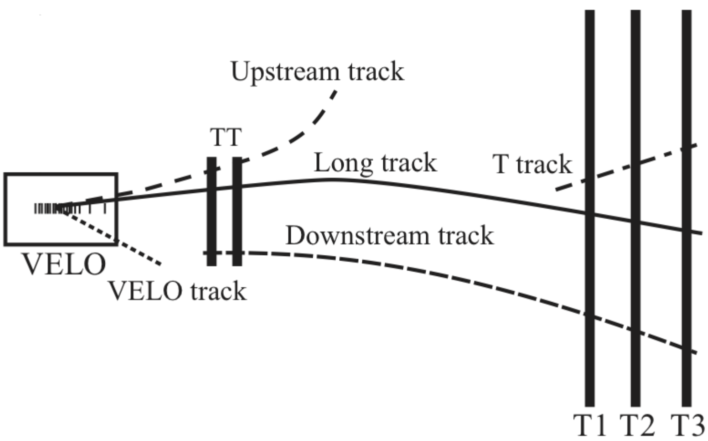

Figure 1.3: Track type at *LHCb*. Velo tracks are basically straight lines since the magnetic field is almost 0 in that region. Tracks are mainly bended in the x-z [5]plane between the Tracker Turicensis (Upstream Tracker for the upgrade) and the T-Stations which is composed by the Inner Tracker(IT) and the Outer Tracker(OT) (Run-I and Run-II), while for the upgrade the stations will be replaced by the Scintillating Fibre tracker (SCIFI).

[5] z- is the beam axis direction and the y axis is the B field line direction.

In the tracking system of *LHCb* each track type is reconstructed by a dedicated algorithm and a schematic layout of how it is done is given in table 1.2.

Track type	Used detector	Algorithm(s)	Input tracks	Output tracks
Velo Tracks or Velo-Segment	Velo	Velo algorithm	/	Velo
Seed Tracks or T-Tracks	T-Stations (SciFi in upgrade)	Seeding algorithm	Allow the possible usage of the leftover hits of forward. If Not: Standalone Algo	Seed
Long Tracks (1)	Velo + TT + T-Stations (TT → UT T-Stations → SciFi in upgrade)	1) Forward tracking: Search in T-Stations knowing Velo-Segment (adding also TT) 2)Matching algorithm: Merge T-Tracks with Velo-Segment 3)BestSelector= Forward+Matching	Velo Velo and Seed	Long
Downstream Tracks	T-Stations and TT (SciFi and UT)	Downstream algorithm: Use T-Tracks and add TT (UT upgrade) hits	Seed	Downstream
Upstream Track	Velo and TT	Upstream algorithm: Use Velo segment and add TT(UT upgrade) hits	Velo Container	Upstream

Table 1.2: Tracking algorithm at LHCb

All the tracks produced by the algorithms provided in Table 1.2 are processed afterwards by the *Kalman Filter* which fit the tracks assigning a sort of *univoque* χ^2 to the track (each algorithm infact has an internal track model parametrisation resultin in different χ^2 computation) taking into account the *B(x,y,z)* field map and the material budget computing for instance corrections due to multiscattering.

From a more technical point of view, in the *Kalman Filter* and in the *LHCb* framework each track is defined by a vector of track state $\left(x, y, t_x, t_y, \frac{q}{p}\right)_z^T$ $(t_x = \frac{dx}{dz}, t_y = \frac{dy}{dz})$ which is then propagated by a 5x5 matrix through the detector considering the interactions and the B-Field map [6]. So, the main goal of the listed algorithm in Table 1.2 is to provide a preliminary set of tracks composed by compatible hits in a given sub-detector and only at the end of the algotithms the found tracks are converted into a vector of track states, which can be handled by the *Kalman Filter*.

1.3 Hybrid Seeding Algorithm

Since no real data is available at the upgrade condition, a huge effort must be done to reproduce the real data taking conditions and what the real output of the detector would be. This task is obtained through different steps in the simulation. Without going into details we can make a short summary of how the *LHCb* software is working in order to provide, for a given algorithm the final efficiencies. In this document efficiencies will be given for a simulated sample of $B_s \to \phi\left(K^+K^-\right)\phi\left(K^+K^-\right)$ event.

- The *Gauss* application is responsible of the physics simulation, starting from the $pp \to b\bar{b}$ production (package doing that is called *Pythia*) given the 25 *ns* bunch crossing, the $\sqrt{s} = 14$ *TeV* centre-of mass energy collision and the ν^7*=7.6* and the subsequent $B_s \to \phi\left(K^+K^-\right)\phi\left(K^+K^-\right)$ (package *EvtGen*) decay chain are provided. Inside *Gauss* also the interaction of the produced final states in the detector is simulated by the *Geant* package. At this level of the simulation the so-called *MCHits* are produced, corresponding to the energy deposition in the active material of the detector. A true particle is then given by a certain set of *MCHits*.

- The read-out, the *SiPm* response , the light attenuation in the fibres and the *Clusterization* together with the front-end boards (data - encoding) is simulated by the *Boole* application. The output of this step are the *Clusters* containing binary information about the corresponding position in the tracker.

- Once the digitization is done, the resulting objects that pattern recognition algorithm can handle are the so called *PrHit* (in the SciFi context) and making use of them one can finally reconstruct tracks. The *PrHit* are obtained converting the binary information of the *Clusters* into a geometrical information usable for pattern recognition. The track reconstruction is done inside the *Brunel* application.

- In each of the previous steps one can store the *Linkers* which allows to figure out the list of *MCHit* associated to a true particle which are finally converted to cluster (so into a *PrHit*). Thanks to them one can define the *reconstructible*[8] tracks and compare them with the output of the pattern recognition algorithm (*reconstructed* tracks) in order to provide tracking efficiencies numbers.

[6] The formalism is quite similar to accelerator physics.
[7] Average number of visible interaction per bunch crossing
[8] For the SciFi a reconstructible track is defined if it leaves at least 1 x-layer hit in each of the three stations together with at least one hit in the u or v layer for each of the three stations

The most important figures of merit classifying a pattern recognition algorithm are the following:

- *Reconstruction efficiency* : how many of the tracks we find in the algorithm are actually tracks we expect to reconstruct? For instance,for analysis requiring n final states, the yield will be reduced by a factor ϵ^n.
- *Ghost rate* : how many tracks we reconstruct are associated to pesudo-random combination of hits?. This is important because we don't want to recontruct fake tracks and in addition we don't want to feed the *Kalman Filter* with too many fake tracks for timing constraints.
- *Timing* : how fast we are finding the tracks. This is particualry important since in the upgrade the tracking algorithms has to be run online at an incoming rate of *40 MHz*.

  The work done from October 2014 to now is encoded in this framework.

  At the beginning i was involved in the debugging of the simulation code, putting my hands in the detector description, the simulation of the clusterisation and the encoding/decoding of the clusters (basically the *Boole* and *Brunel* applications and packages). Some problems and inefficiencies in the software were infact not well understood passing from the default detector description (the one of the technical design report [**?**]) to the new one where bigger gaps between channels were simulated, together with a new shape of detector layers, and a more realistic model of the clusterisation were implemented.

  In that context my effort was fundamental (i took part of the *DILBERT* task force for the software fixing and bug correction) to recover part of the loss in efficiency and we finally provide a stable version of the simulation of the detector from which the development of a new pattern recognition algorithm becomes possible.

  In the meanwhile i also took part to the *Test Beam* data analysis aiming to measure the radiation lenght of the fibres doing light yield measurements scanning along the active material at different distances from the read-out and also evaluating the dependency of the cluster size against track incident angles. I also analyse some of the data from the test beam to study the cluster size properties as a function of the incident angle of tracks in the detector in order to tune the simulation to be the most reliable possible. [**?**] [**?**]. The results of how the efficiencies for the seeding algorithm change in time will be shown at the end of this section to highlight my contribution to the upgrade *SciFi simulation and reconstruction* working group. I presented regularly my work in the internal *SciFi simulation and reconstruction* working group and i presented once the results of the group in the parallel session of the *LHCb week*.

1.4 IMPROVEMENTS APPORTED TO THE ALGORITHM

Several trials in changing the previous algorithm (named here as *Old*) has been done and we finally decide to develop a new one from scratch, once we figured out the limitations of the previous one. The developing of the new one started from very basics things related to track properties, detector occupancy and requirements on tracks.. First of all we need to separate different aspects of the algorithm:

– The way we collect the hits to generate tracks.
– The way we fit the tracks.
– The tolerances we allow in both the fitting and the hit search.
– The requirements we put on tracks for their storage and clone/ghost killing.

. In order to do that, several tools have been developed: one of them, for instance, aims to extract the *Monte-Carlo* information of tracks before the clusterisation step in order to evaluate the track properties and their behavior in the *SciFi* stations.[9] The strategy adopted once the track property was extracted was then the following: instead of using a look-up table for the *B* field information one can directly use the *B*-field effect on tracks parametrising their motion by some fixed constants reflecting the magnetic field behavior.[10]

The main limitations, or problems, of the previous algorithm can be summarized in few points

– It was trying to find everything at once, meaning that it was not disentangling the cases where tracks are easy to finde to the cases where the tracks are harder to find. Harder to find stands for find it paying an high prize in terms of ghost rate.
– The tracks behavior in the *SciFi* tracker was roughtly parametrised.
– The different detector area were treated in the same way even if where we expect a lower efficiency we can loosen the cuts for the track selection.

  The strategy adopted for the track search is to **split the algorithm in nested sub blocks** following the philosophy of the *projective tracking* and *progressive cleaning* of the tracking environment. Since the detector is divided into two halves (*y>0* and *y<0*) a natural choice is to split the algorithm in a first track search in the upper modules of the stations and a second search in the lower half. From *MC* study we figure out that the numebr of tracks going from the upper to the lower half is a negligible fraction of the full set of tracks *O(0.1 %)*. Inside the upper(or lower) half search we run the second main block of the algorithm 3 times whcih we will call here *case* since for each of this run different configurations are applied. Each of this *3 cases* is dependent on the previous one since at the end of each of them (only the first and the second) we mark the hits of the found tracks using some specific criteria, resulting in a progressive cleaning of the tracking

[9] Thanks to that an unofficial tracking algorithm running over true hit position was developped to figure out which is the best track model to use in order to fit the tracks going through the *SciFi* and also to have an estimation of "how much" one can squeeze some search windows without starting to loose to much in efficiency.
[10] The B field in the Sci-Fi tracker region is a fringe field which is not easly parametrisable, but still, some caractheristic behavior can be extracted.

environment moving from one case to the following one. The idea for the first case is to find tracks with high momentum which are easy to find since the kick they get in the bending plane from the *B-field* both in the *SciFi* tracker region and in the propagation between its origin vertex and the *SciFi* is relatively small ($\propto \frac{qB}{p}$). Hits are collected in a relatively small search windows under the infinite momentum assumption. The infinite momentum assumption relies on the fact that very high momentum tracks just goes straight even in the higher magnetic field region, resulting in a linear pointing to the origin vertex. Selecting *MC* hit information and plotting the position of the true hit on the *1st*

1.4.1 Analysis Work : $B^0 \to D^0\overline{D}^0 K^{*0}$ analysis

2 Choice of Thesis Topic

3 Timetable for future

References

[1] LHCb Collaboration , *The LHCb Detector at the LHC, JINST 3(2008) S08005*

[2] LHCb collaboration, *Letter of Intent for the LHCb Upgrade*, CERN-LHCC-2011-001, March 2011.

[3] LHCb collaboration, *Framework TDR for the LHCb Upgrade*, CERN-LHCC-2012-007, May 2012.

[4] LHCb collaboration, *LHCb Scintillating Fibre Tracker Technical Design Report*, CERN-LHCC-2014-001; LHCb TDR 15.