Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Dipartimento di Fisica e Astronomia
Corso di Laurea Magistrale in Fisica

Preliminary studies for a future LHCb Upgraded Vertex-Locator using 4D track reconstruction

Relatore:

Prof. Angelo Carbone

Presentata da:

Tommaso Fulghesu

Correlatrice:

Dott. Serena Maccolini

Anno Accademico 2019/2020

Contents

Abstract

This thesis reports a preliminary study aimed at finding realistic solutions for the future Vertex Locator of the LHCb experiment in the harsh scenario imposed by the High Luminosity phase of the Large Hadron Collider (LHC). One of these possible solutions to cope with higher multiplicity is the introduction of the time associated to each pixel. The addition of the fourth coordinate, the time, will improve the efficiency of track and primary vertices reconstruction and reducing the rate of fast tracker. The impact of the time information is discussed in the present work, through simulations. The fast simulation developed simulates the Upgrade-I detector geometry and the deposit of charge inside each pixel. It has been validated with a full, official, LHCb simulation. Differently from the latter, the fast simulation has the advantages to be more fast and versatile, to cope with higher rates of Upgrade-II conditions. This project is done in collaboration with the INFN project TimeSpot, which proved that silicon detectors with a time resolution of $\mathcal{O}(30\,\mathrm{ps})$ are feasible [1]. Because of the high data-rate and the demand for real-time devices, FPGA solution are investigated as candidates to perform clustering at read-out level, due to their high predisposition to perform easy and repetitive processes in a highly parallelized form. Since FPGAs receives in input binary data, a format called raw banks is implemented in the fast simulation. Each raw bank encodes active pixels information of a detector sensor. Once these information are processed, FPGAs will reconstruct tracks from the hits in parallel way. A research is conduced in order to evaluate the clustering efficiency of FPGA-based algorithm in terms of the CPU one.

Introduction

The main aim of the LHC [2], the largest and highest-energy particle collider in the World, is to allow physicists to test the predictions of different theories of Particle Physics, in particular the ones beyond the Standard Model. To give a proof of many and many theories relative to dark matter and super-symmetries, the LHC has to be improved in terms of the energy reachable by collisions which happen inside and, at the same time, of the precision in measurements made. One of the key parameter of the accelerator is the luminosity, which is proportional to the number of collisions per unit of time. The LHCb experiment [3], one of the four main detector of the LHC, investigates on why Nature prefers matter with respect to antimatter: it is mainly dedicated to the study of parameters which characterize CP violation and the decays of charm and beauty hadrons. To cope with the higher luminosity released by the LHC, the LHCb experiment has to undergo several changes in its design [4]. Every 25ns, corresponding to a frequency of 40 MHz, two beams of protons cross one another in the so-called interaction point. In each interaction point detector, new and more powerful electronic components have to replace the older ones to reconstruct the decay process in a high-multiplicity tracks environment.

The LHCb experiment has to undergo several changes to fully exploit the HiLumi-LHC potential, even its components are subject to change. In particular a component of the tracking system, the Vertex Locator, is considered for the purposed studies. The different Upgrade VELO are presented and the structural modifications are shown, in particular between Upgrade-I and Upgrade-II. The improvements necessary to satisfy the Upgrade-II conditions brenching out in two different direction: the addition of time information to associate to each pixel and the removal of the radio-frequency foil as module shield. To test these options, a new tool faster and more easy-to-modify with respect to the official simulation has been developed and validated: the fast simulation. It has the

aim at testing the influence of time information on tracking performances. Moreover, a study, based on the fast simulation, has been conducted on the FPGAs, instead of CPUs read-out. The thesis is organized in five different chapters, as follows:

- Chapter 1 describes the physics of the LHCb experiment. A section is also dedicated to the studies of the Upgrade-II physic cases is also added.

- Chapter 2 illustrates the composition of the Large Hadron Collider and its future plan. More detail is reserved to one of its experiment, the LHCb one.

- Chapter 3 contains a specific part of the LHCb detector, the Vertex Locator (VELO), which is part of the tracking system.

- Chapter 4 describes the development of the fast simulation. This is validated with the official LHCb one under Upgrade-I conditions. The time information is added to face the higher multiplicity of the Upgrade-II scenario. Furthermore, this chapter contains the description of how data are managed in the simulation performed and future projects.

Chapter 1

The LHCb physic case

The following chapter aims at giving to the reader a general idea of the LHCb experiment physics. A brief introduction to the physics case for the Upgrade-II is presented. More in details, two key CP-violation measurements to probe during this future phase are treated and explained.

1.1 CP violation and Quark-mixing matrix

The main research field at LHCb regards studies associated with the violation of the CP-symmetry, i.e. the violation of the laws of physics for charge (interchange of particles and anti-particles) and parity (spatial coordinates inversion) symmetries simultaneously. CP-violation, discovered in 1964 by the study of the τ-puzzle (neutral kaon system), is at the base of the weak interactions and is one of the main explanation for the dominance of matter over anti-matter [5].

The full lagrangian of the Standard Model is composed by three parts:

$$\mathcal{L}_{SM} = \mathcal{L}_{kinetic} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa} \tag{1.1}$$

The kinetic term describes the dynamic of the spinor fields and, in presence of interaction,

it could be expressed through the covariant derivative as:

$$\mathcal{L}_{kinetic} = i\bar{\psi}(D^{\mu}\gamma_{\mu})\psi, \tag{1.2}$$

where $\bar{\psi} = \psi^{\dagger}\gamma^0$. Since only the left component of the quark doublets is involved in weak charged current interactions, the lagrangian becomes:

$$\mathcal{L}_{kinetic} = i\bar{u}_{iL}^I\gamma_{\mu}\partial^{\mu}u_{iL}^I + i\bar{d}_{iL}^I\gamma_{\mu}\partial^{\mu}d_{iL}^I - \frac{g}{\sqrt{2}}\bar{u}_{iL}^I\gamma_{\mu}W^{-\mu}\bar{d}_{iL}^I - \frac{g}{\sqrt{2}}\bar{d}_{iL}^I\gamma_{\mu}W^{+\mu}\bar{u}_{iL}^I + \dots, \tag{1.3}$$

using $W^+ = \frac{1}{\sqrt{2}}(W_1 - W_2)$, $W^- = \frac{1}{\sqrt{2}}(W_1 + W_2)$ and I for the flavour states representing the interaction eigenstates.

Since the weak-interaction mediator bosons W and Z aquire masses through the spontaneous symmetry breaking, the Higgs field is added to the Lagrangian. In order to provide mass through Higgs mechanism, the Standard Model allows the quark flavour mixing [6]. The 3×3 unitary matrix that describes this mixing is also known as *Cabibbo-Kobayashi-Maskawa* matrix. The CKM-mechanism is the origin of CP violation, and earned Kobayashi and Maskawa the Nobel price in 2008, *"for the discovery of the origin of the broken symmetry which predicts the existence of at least three families of quarks in nature"* [7].

The interaction between the Higgs scalar field and the fermions is given by the so-called Yukawa lagrangian:

$$\mathcal{L}_{Yukawa} = -(Y_{ij}\bar{\psi}_{Li}\phi\psi_{Rj}) + h.c., \tag{1.4}$$

with the Higgs isospin doublet

$$\phi(x) = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{1.5}$$

and Y_{ij} are 3×3 arbitrary complex matrices that give rise to coupling between different families (quark mixing).

After the spontaneous symmetry breaking. ϕ acquires a value (the vacuum expectation value) correspond to

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{1.6}$$

Then, the mass terms for fermions becomes:

$$\mathcal{L}_{Yukawa}^{quark} = -(M_{ij}^d\bar{d}_{Li}^I d_{Rj}^I + M_{ij}^u\bar{u}_{Li}^I u_{Rj}^I) + h.c + \text{interaction terms}. \tag{1.7}$$

8

To obtain the proper mass term, the matrices M^d and M^u should be diagonalized. Using unitary matrices V (such that $V_L^\dagger V_L = \mathbb{1}$), one has $M_{diag}^f = V_L^f M^f V_R^{f\dagger}$, with $f = u, d$. Thus, the lagrangian becomes:

$$\mathcal{L}_{Yukawa}^{quark} = -(\bar{d}_{Li}(M_{ij}^d)_{diag}d_{Rj} + \bar{u}_{Li}(M_{ij}^u)_{diag}u_{Rj}) + h.c, \tag{1.8}$$

and the eigenstates d and u are now mass eigenstates, arised from the absorbtion of the matrices V by the interaction eigenstates.

$$d = (V^d)_{ij}d_j^I,$$
$$u = (V^u)_{ij}u_j^I.$$

As a result, the charged-current $W^\pm$ interactions couple to the physical u_{Li} and d_{Li} according to the Cabibbo-Kobayashi-Maskawa mixing matrix

$$V_{CKM} = (V_L^u V_L^{d\dagger}) = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{1.9}$$

Where each element of the CKM matrix defines the strength of interaction [8]. The transition from the *down* type quarks to the *up* time quarks is described by V_{ud} and, vice versa, the *up* type quarks to the *down* time quarks is described by V_{ud}^*.
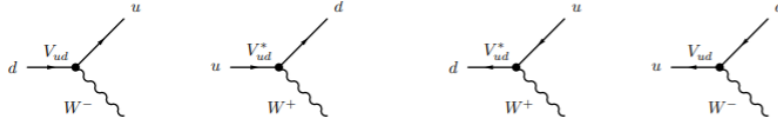

Figure 1.1: Charged currents with V_{ud} and V_{ud}^* that define the strength of interactions.

The CP violation becomes evident in the complex Yukawa coupling

$$\mathcal{L}_{Yukawa} = -(Y_{ij}\bar{\psi}_{Li}\phi\psi_{Rj} + Y_{ij}^*\bar{\psi}_{Rj}\phi^\dagger\psi_{Li}). \tag{1.10}$$

Under CP operation the spinor fields transform as

$$CP(\bar{\psi}_{Li}\phi\psi_{Rj}) = \bar{\psi}_{Rj}\phi^\dagger\psi_{Li}$$

and $\mathcal{L}_{Yukawa}$ remains unchange if $Y_{ij} = Y_{ij}^*$. The kinetic term has similar behaviour and it is possible to conclude that the total Standard Model lagrangian remains unchanged

if $V_{ij} = V_{ij}^*$. The complex nature of the CKM matrix is the origin of CP violation in the Standard Model. This matrix, by unitarity, can be parametrized by four parameters (three mixing angles and the CP-violating mixing phase). Using the Wolfenstein parametrization [9] it is possible to express the CKM matrix in terms of A, ρ and η

$$V_{CKM} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{1.11}$$

where:

$$A = \frac{\sin\theta_{23}}{\sin^2\theta_{12}}, \quad \lambda = \sin\theta_{12}, \quad , \rho = \frac{\sin\theta_{13}\cos\delta}{\sin\theta_{12}\sin\theta_{23}}, \quad \eta = \frac{\sin\theta_{13}\sin\delta}{\sin\theta_{12}\sin\theta_{23}}. \tag{1.12}$$

The off-terms of diagonal are relatively small, which implies that weak interaction of quarks from different generation are suppressed, in particular between first and third families.

1.1.1   Unitary triangle

From the six orthogonality relations it is possible to build six unitary triangles. Let's consider the relation:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0, \tag{1.13}$$

which is particulary relevant for B-decays. By dividing the three sides by $|V_{cd}V_{cb}^*|$, it is possible to obtain the famous unitary triangle shown if Figure 1.2.
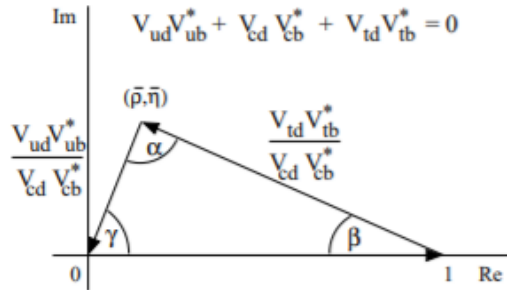

Figure 1.2: Unitary triangle represented in the complex plane.

The apex of the triangle is given by

$$\bar{\rho} + i\bar{\eta} = \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}, \tag{1.14}$$

with $\bar{\rho}$ and $\bar{\eta}$ that can be expressed through Wolfenstein parameters ρ and η

$$
\begin{aligned}
\bar{\rho} &= \rho \left(1 - \frac{1}{2}\lambda^2\right) + \mathcal{O}(\lambda^4), \\
\bar{\eta} &= \eta \left(1 - \frac{1}{2}\lambda^2\right) + \mathcal{O}(\lambda^4),
\end{aligned}
\tag{1.15}
$$

and the angles of the unitary triangle are defined as:

$$\alpha = \arg\left[\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right], \quad \beta = \arg\left[\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right], \quad \gamma \arg\left[\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right] \tag{1.16}$$

Measurements of the triangle elements can be provided by the studies on neutral kaons and B-meson systems.

1.2 Physic cases for Upgrade II

The LHCb Upgrade-II is being designed as the flavour physics experiment for the High Luminosity LHC era, described in Section 2.1.1. It will allow to determine, with an order of magnitude higher precision, or better, the unitary triangle parameters, in particular the angle γ. Discrepancies among the various measurements that can be made of the sides and angles of the unitary triangle would be strong indications of physics beyond the SM. For this reason there would be required high precision measurements, as those expected from LHCb Upgrade-II. Furthermore, Upgrade II will also allows a wide-ranging set of CP violation measurements in charm and beauty decays and lepton universality tests in $b \to c\, l^- \bar{\nu}_l$ processes.

1.2.1 Time dependent CP-violation measurements in High Luminosity era

Many of the most important observables to test the SM in the B system are intrinsically related to the phenomenology of flavour oscillation: a $B^0_{(s)}$ meson can mix into

its antiparticle $\bar{B}^0_{(s)}$ and vice-versa, creating a loop process that can be influenced in its transition amplitudes by unknown physics.

The mixing results in physical states, with defined masses and lifetimes, are:

$$B^0_{(s)H} = pB^0_{(s)} + q\bar{B}^0_{(s)} \quad \text{and} \quad B^0_{(s)L} = pB^0_{(s)} - q\bar{B}^0_{(s)} \tag{1.17}$$

with p and q complex coefficients that satisfy the relation $|p|^2 + |q|^2 = 1$.
The two physical states, labelled with H and L to distinguish the heavier one and the lighter one, have mass difference $\Delta m_d(s) = m_{B^0_{(s)H}} - m_{B^0_{(s)L}}$ and width difference $\Delta\Gamma_d(s) = \Gamma_{B^0_{(s)L}} - \Gamma_{B^0_{(s)H}}$.

If the flavour eigenstates $B^0_{(s)}$ and $\bar{B}^0_{(s)}$ decay into a general final state f, their decay-time-dependent rates depends on those quantities and are proportional to some mixing parameters, such as the transition amplitude. All those parameters can, for certain choices of final state f, be related to angles of the unitary triangle; in other cases they provide null tests of the SM. Further precise test that can be done with Upgrade-II can regard the assumption of CPT symmetry conservation, fundamental in the SM, and the one in the decay-time-dependent analyses that $\left|\frac{q}{p}\right| = 1$, *i.e.* absence of CP violation in mixing [10].

Measurements of mass and width differences

The values of $\Delta m_{d(s)}$ and $\Delta\Gamma_{d(s)}$ strongly affect the sensitivity with which the parameters of mixing, such as the amplitude, can be measured. In the B^0_s system, the large value of Δm_s means that excellent vertex resolution is necessary to resolve the flavour oscillations and hence to determine these parameters. Potential improvements to the vertex resolution may therefore improve the sensitivity to decay-time-dependent CP-violation effects in the B^0_s system beyond that from simple scaling with luminosity.

Assuming that good flavour tagging performance can be maintained, the measurements of both $\Delta m_{d(s)}$ and Δm_s [11] can be improved, reducing in this way uncertainties in CP-violation measurements and providing a strong constraint on the length of one side of the unitary triangle.

Different is the discussion relative to the width difference $\Delta\Gamma$ because $\Delta\Gamma_s$ is relatively

large, while $\Delta\Gamma_{d(s)}$ can be considered negligibly small in most analysis. The size of $\Delta\Gamma_s$ allows to determine either through flavour-tagged or untagged analyses of B_s^0 decays the transition amplitude, while it is usually unmeasurable in B^0 system The decay-time-dependent angular analysis of $B_s^0 \rightarrow J/\psi\phi$ allows measurement of $\Delta\Gamma_s$ simultaneously with CP-violation parameters. Therefore, improved knowledge of $\Delta\Gamma_s$ will be obtained together with measurements of $\phi_s^{c\bar{c}s}$, as discussed in next paragraph.

Measurements of ϕ_s from $B_s^0 \rightarrow J/\psi\phi$

Measurements of decay-time-dependent CP asymmetries in the B_s^0 system using $b \rightarrow c\bar{c}s$ transitions are sensitive to the CKM phase $\beta_s \equiv \arg\left[-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right]$. Neglecting gluonic and electroweak penguin loop [12], the experimentally observable is the phase $\phi_s^{c\bar{c}s} = -2\beta_s$, which has a precise SM prediction of -36.4 ± 1.2 mrad based upon global fits to experimental data [13]. Deviations from this value would be a clear sign of physics beyond the SM.

The most statistically sensitive measurement $\phi_s^{c\bar{c}s}$ is given by the flavour-tagged angular analysis of the $B_s^0 \rightarrow J/\psi\phi$ decay [14]. Angular analysis is necessary to disentangle the interfering CP-odd and CP-even components in the final state, which arise due to the relative angular momentum between the two vector resonances. In addition, there is a small ($\simeq 2\%$) CP-odd K^+K^-, for the ϕ decay, S-wave contribution that must be accounted for. Results of $\phi_s^{c\bar{c}s}$ analysis in terms of $\Delta\Gamma_s$, determined from fits of $B_s^0 \rightarrow J/\psi\phi$ data and, in the case of $B_s^0 \rightarrow J/\psi\pi^+\pi^-$ LHCb, of data, are shown in Figure 1.3a. Measurements are consistent with SM prediction [14], but there is still possibility of other new physics contribution of $\mathcal{O}(10\%)$, for this reason improving experimental precision will be essential to have good control over possible hadronic beyond-the-SM physics effect [15][16][17].
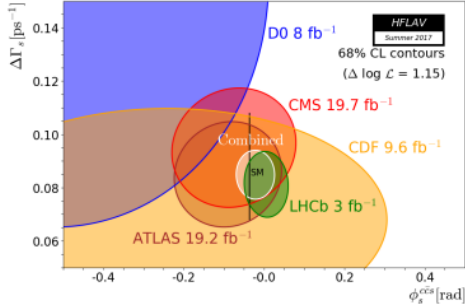
(a) Fit results from various experiments of ϕ_s in terms of $\Delta\Gamma_s$.
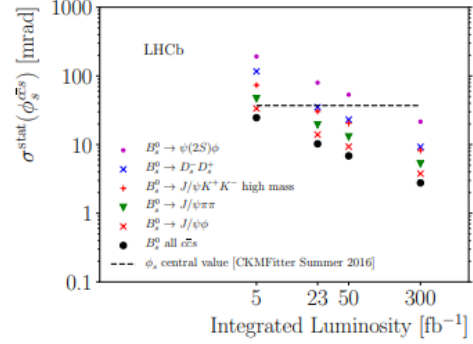
(b) Statistical precision measurements on ϕ_s from three different modes of B_s^0 meson decay

Figure 1.3: Current experimental status on ϕ_s

It is very important, in order to improve the average precision and to perform powerful consistency check of the SM, to have multiple independent high-precision measurements. This is guaranteed by Upgrade-II through studies of independent CP-violation effects for each polarisation state in the $B_s^0 \to J/\psi\phi$. Additional information can be obtained from $B_s^0 \to J/\psi K^+ K^-$ decays with $K^+ K^-$ invariant mass above the $\phi(1020)$ meson [16]. Among other channels, competitive precision can be obtained from $B_s^0 \to J/\psi\pi^+\pi^-$ decays, which have been found to be dominated by CP-odd component [17]. Figure 1.3b shows the statistical uncertainties related to $\phi_s^{c\bar{c}s}$ with integrated luminosity for individual decay modes and for their combination.

Chapter 2

The LHCb detector at the LHC

LHCb is one of the main experiment at the most famous particle accelerator in the world, the Large Hadron Collider, at CERN. After a short description of the hadron collider, this chapter is dedicated to the description of the LHCb detector. In particular in Section 2.2.4, different scenarios for the future project of LHC, called HiLumi LHC (HL-LHC) are considered.

2.1 The Large Hadron Collider

The LHC [2] was successfully commissioned in 2010 and started delivering proton-proton collisions with different centre-of-mass energy[1] and luminosity[2]. With the challenge to test the predictions of different theories of particle physics, the operational program of LHC was divided in *Runs*, during which particles circulate and collides inside the accelerator, alternated by *Shutdowns*, periods dedicated to the maintenance and the improvement of the structure.

The LHC accelerator ring, in Figure 2.1 represents the last acceleration stage, after sev-

---

[1]The energy measured in the center-of-mass reference frame, so the reference frame centered in the weighted average position of all the masses

[2]A measurement of the number of collisions that can be produced in a detector per cm^2 and per second.

eral particle accelerators. At the beginning, protons are obtained by ionizing hydrogen gas with an electric discharge and are accelerated by the *Linear Accelerator 2* (LINAC2) up to 50 MeV. Then protons are then injected into a sequence of three synchrotrons (*Proton Synchrotron Booster*, *Proton Synchrotron* and *Super Proton Synchrotron*), where, at the end, the protons reach energy of 450 GeV. Finally, the beam is split and injected into the LHC where two separate beams[3] travel in clockwise and anticlockwise directions and are accelerated up to different TeV. After all the acceleration stages have been performed the two beams collide in four interaction points where the main LHC experiments are located: ALICE [18], ATLAS [19], CMS [20], LHCb [3]. Each beam consist of 2808 bunches[4] of protons distant 25 ns, with resulting bunch crossing frequency of 40 MHz. Each bunch contains nominally 1.15×10^{11} protons.
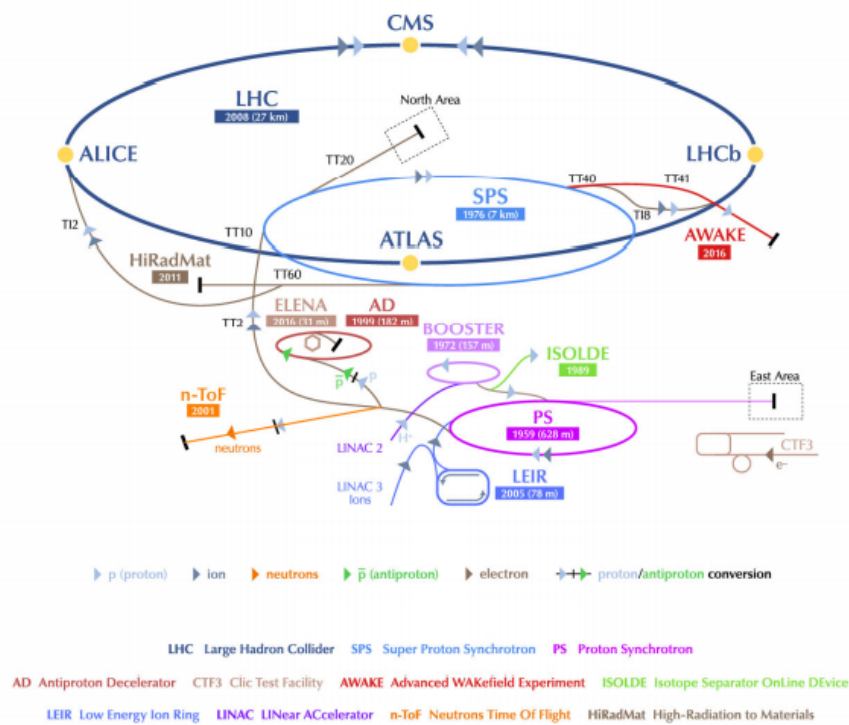

Figure 2.1: Schematic view of the CERN accelerator complex.

---

[3]Ensamble of particles which circulates on one side of the accelerator

[4]Agglomerates of particles which travel along stable orbits

2.1.1 High luminosity phase at LHC

During Run 1 *pp* collisions were studied at a centre-of-mass energy of $\sqrt{s} = 7$ TeV in 2010-2012, and at $\sqrt{s} = 8$ TeV from April 2012 until the end. The Run 2 started in June 2015 and ended in December 2018. During this period, the centre-of-mass energy reached the value of $\sqrt{s} = 13$ TeV, with a luminosity (2×10^{34} cm^{-2}s^{-1}) doubled with respect to nominal designed value.

Run 3, instead, expected to start in 2021, will have an increase in energy upon to $\sqrt{s}$ = 14 TeV, keeping the same luminosity as in Run 2. The statistical gain in running the accelerator without a luminosity increase will become marginal and if on one hand this would lead to doubling the statistics from the Run 2 to the Run 3, on the other hand the time necessary to reduce by half the statistical errors in measurement results Would be ten years starting from the end of Run 2. For this reason a new project, called *High Luminosity LHC* or *HL-LHC* is scheduled to start in 2026, until 2029. During this period, an increasing of luminosity by a factor five with respect to nominal value is expected through an upgrade of LHC major components and the use of crab cavities at the interaction regions.

The crab cavities are able to "tilt" the proton bunches in each beam, maximising their overlap at the collision point [21]. In this way every single proton in the bunch is forced to pass through the whole length of the opposite bunch, which increases the probability that it will collide with another particle, as represented in Figure 2.2. After being tilted, the motion of the proton bunches appears to be sideways – just like a crab [22].

During Run 4 and the High Luminosity phase, a peak luminosity between 5 and 7.5 $\times$ 10^{34}cm^{-2}s^{-1} is expected, while this value has to be exceeded during Run 5 (2031-2034) by at least one half [23].
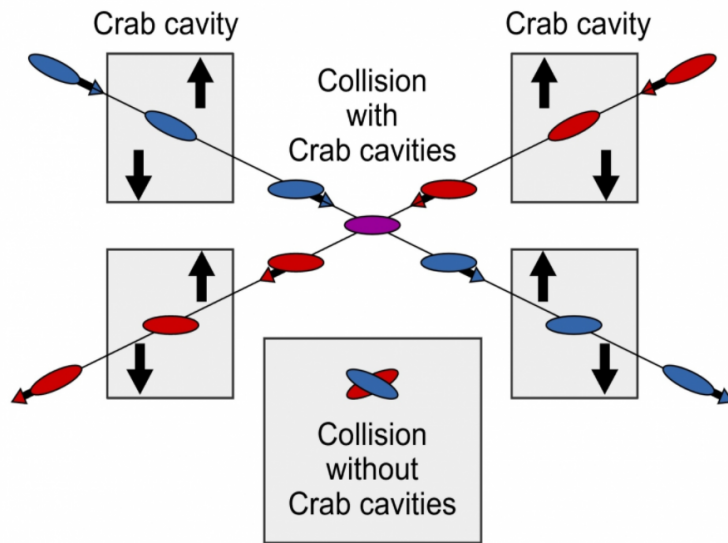
Figure 2.2: At the moment, the counter-propagating (red and blue) proton bunches meet with a crossing angle (bottom). Crab cavities will administer a transverse RF kick that tilts the bunches so that they appear to move sideways, causing them to collide head-on (purple) at the collision point.

The luminosity plan of LHCb experiment, represented in figure 2.3 from 2010 to 2037, is a bit different. Indeed in this case the High Luminosity phase will start with the beginning of Run 5 instead of at the end of Run 4. The LHCb experiment makes use of the beam levelling to reduce the instantaneous luminosity below the values delivered by the LHC accelerator.
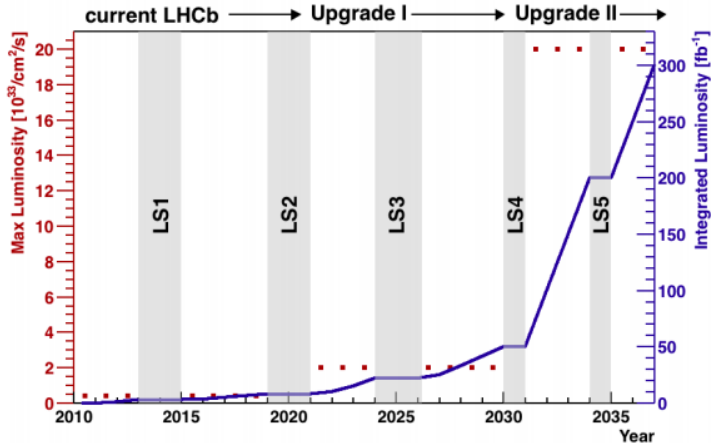
Figure 2.3: LHCb luminosity plan from 2010 to 2037. Red dots are the value measured or predicted of instantaneous luminosity, while the solid blue line represents the measured or expected integrated luminosity [24].

2.2 The LHCb experiment

LHCb is one of the four main experiment investigating the hadron collision at the Large Hadron Collider (LHC) at CERN. Its primary purpose is the search for indirect evidences of new physics in CP violation and rare decays of beauty and charm hadrons. The detector is a single-arm spectrometer designed for high precision heavy flavour physics measurement.

Heavy flavour hadrons, like charm and beauty hadrons, are produced in large amount and mostly at small angles. Because of its particular geometry, the detector covers a pseudo-rapidity range of $2 < \eta < 5$, where the pseudo-rapidity is defined as

$$\eta = -\log \left[\tan \left(\frac{1}{2} \theta \right) \right], \tag{2.1}$$

where θ denotes the angle with respect to the beam axis. The choice of the detector geometry is justified by the fact that at high energies heavy-flavour hadrons, like charm and beauty hadrons, are produced in large amount within the angular acceptance. Moreover,

19

both the particle and the anti-particle are produced in the same forward (or backward) cone. The layout of the LHCb spectrometer is shown in Figure 2.4.

The detector provides an excellent primary and secondary vertex reconstruction; it includes a high-precision tracking system consisting of a silicon-strips vertex detector surrounding the collision region (VELO), a silicon-strip detector located upstream of a dipole magnet, and three stations of silicon-strip detectors and straw drift tubes placed downstream of the magnet. Moreover, LHCb has two ring-imaging Cherenkov (RICH) detectors, an electromagnetic and a hadronic calorimeter for particle identification. Muons are identified by a system composed of alternating layers of iron and multiwire proportional chambers.
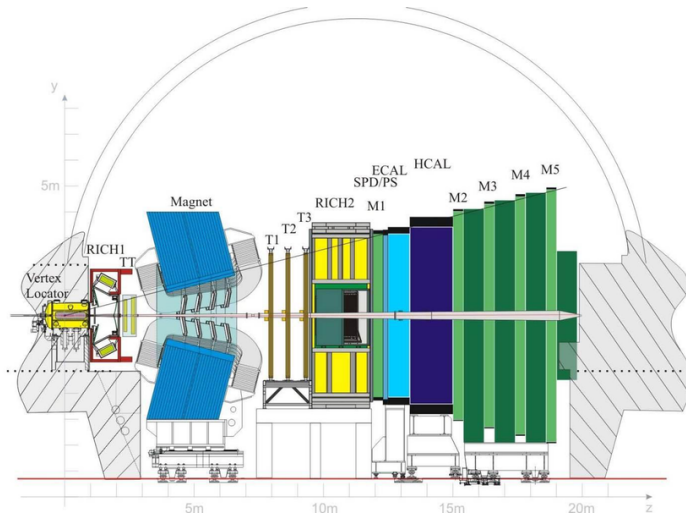
Figure 2.4: Profile of the LHCb detector: The right-handed coordinate system adopted has the z axis along the beam, and the y axis along the vertical [25].

2.2.1 Tracking system

The tracking system provides the measurement of the momentum of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 MeV/c and the measurement of the impact parameter with a resolution of $(15 + 29/p_T)$ µm, where p_T is the component of the momentum transverse to the beam, in MeV/c [25]. It is composed by different parts:

- The Vertex Locator (VELO) [26], near the interaction region, with the aim of reconstructing primary and secondary vertices. The detector is described more in details in Section 3

- The Trigger Tracker(TT), placed upstream the magnet in order to provide reference segments used to combine the track reconstructed in the tracking stations with those reconstructed in the VELO.

- Three tracking stations, placed after the magnet, composed by two different detectors with different technologies: the Inner Tracker (IT) and the Outer Tracker(OT). The first one has finer segmentation and will lead to more precise measurements, but will cover only 2% of detector acceptance. The remaining part is covered by the OT.

- The dipole magnet, which provides an integrated field of 4 T m along the vertical direction, allows to estimate the charged particles momenta by measuring the track trajectory curvature in the horizontal plane.

2.2.2   Particle identification

LHCb has an high efficiency in particle identification through its two ring-imaging Cherenkov (RICH) detectors, a system of calorimeters and five muon stations [27].

Different types of charged hadrons, mainly kaons and pions, with momenta in a range from 2 to 100 GeV/c, are distinguished using the information from Cherenkov radiations.

Photons, electrons and hadrons are identified by mean of the energy they release in the calorimeter system consisting of scintillating-pad detectors (SPD), preshower detectors (PS), an electromagnetic (ECAL) [28] and a hadronic calorimeter (HCAL) [29]. The ECAL and the HCAL are the main components of the calorimeter system. The ECAL provides measurements of the energy released by particles through electromagnetic processes once they crossed the material. It is a sampling calorimeter separated in independent modules. The ECAL consists on alternating layers of scintillating material alternated with lead absorbers. When a particle passes through the absorber it produces a shower of electrons and photons originated from bremsstrahlung and pair production

mechanisms. The electrons interact and deposit their energy in the scintillating material; thus, the converted photons travel to the photomultipliers where the signal is collected and recorded. The HCAL structure is very similar to the ECAL one with the difference of thicker layers and the use of iron instead of lead for the absorber layer. The HCAL measures the energy of particles that interact by strong nuclear interactions.
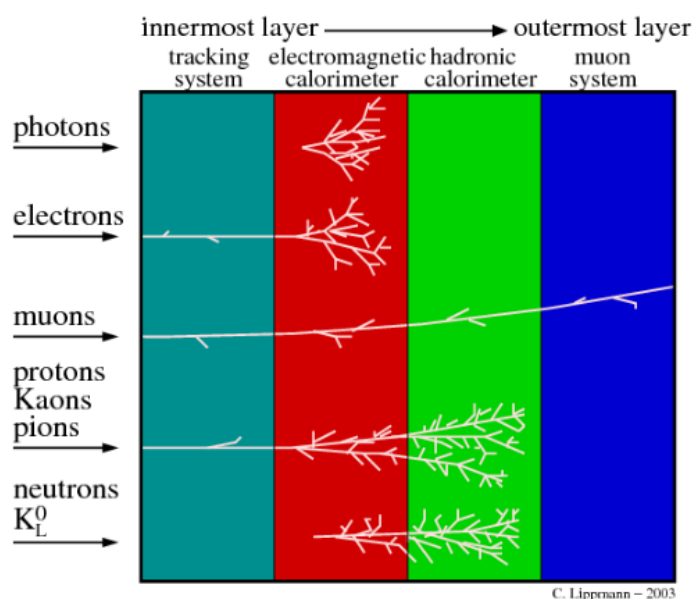

Figure 2.5: Different components originated from the particle collision. Only charged particles leave tracks inside the tracking system. All particles, except neutrinos and muons, creates showers inside the calorimeters. Finally muons lose almost all of their energy and they are stopped inside the muon detector.

Muons are identified by a system composed of five alternating layers of iron and multi-wire proportional chambers. Only three of these stations are used for transverse momentum measurements, while last two provides binary information whenever the particle passed the absorber material or not. For trigger purposes a 5-hit coincidence in all the muon stations allows to identify high-momentum muons. To form a muon track, hits are combined though a straight line extrapolated to the interaction region. Given the hits of the muon track and the reconstructed momenta of the matched track the likelihood is estimated and used for particle identification.

A schematic representation of different signatures of particles is shown in Figure 2.5 [30].

2.2.3   Trigger system

At the moment, the frequency of bunch crossings at the interaction point inside LHCb is 40 MHz, however only a fraction of particles generated from the collision are inside the detector acceptance. The event rate that it is possible to record by offline computing is only 2 MHz and the trigger [31] has to provide the highest efficiency for within the allowed rate.

The system is structured into two levels:

- The *Level 0* (L0). It reduces the rate to 10MHz and it is based on the fact that particles originated from a B decay have a higher transverse momentum than particles coming directly from the primary proton-proton interaction. The *Level 0* is based on calorimeters and muon system, where the required rate can be reached. Furthermore, it uses two dedicated silicon layers of the VELO to perform a simplified vertex reconstruction. In this way, it is possible to improve the rejection of pp interaction and the reconstruction of B meson decays.

- The *High Level Trigger* (HLT). It has access to all the information inside the detector and confirms the separation and identification of particles with high-p_T, using the regions around the candidate direction. Through the information coming from the VELO, the High Level Trigger bases the separation between particles on the impact parameter to the proton-proton interaction vertex. The mesons B have, therefore, long life and they usually decades 1 cm far from the interaction point. Two types of selections are applied at the tracking level. The inclusive selections has the aim to collect the resonances decays and it is useful for calibration. The exclusive ones provide the highest possible efficiency for fully-reconstructed B decays through the separation between B-mesons and everything else based on mass and vertex quality.

2.2.4   High luminosity scenarios at LHCb

The goal of Upgrade II is to be able to operate at an instantaneous luminosity of $1 - 2 \times 10^{34}$ cm^{-2}s^{-1} and to collect an integrated luminosity of 300 fb [32]. The luminosity delivered at LHCb from HL-LHC will essentially depend, once fixed the main beam parameters (such as number of bunches, bunch population, filling scheme and so on) on the minimum β^* and on the crossing angles achievable at the interaction point. The Upgrade-II project at LHCb proposes studies regarding the main parameters of LHCb for given values of HL-LHC beam parameters.

Parameter	Unit	Lumi scenario		
Target leveled lumi	10^{34} cm^{-2}s^{-1}	**1.0**	**1.5**	**2.0**
β^*	m	1.5	1.5	1.5
Crossing plane		V	V	V
Magnet polarity		$\pm$	$\pm$	$\pm$
External crossing angle	μrad	320	320	320
Crossing angle at IP	μrad	419	419	419
Virtual (Peak) luminosity	10^{34} cm^{-2}s^{-1}	1.79	1.79	1.79
Leveled pile-up		28	42	50.3
Long. RMS luminous region (start)	mm	44.7	44.7	44.7
Peak line pile-up density (start)	mm^{-1}	0.25	0.37	0.44
Eff. line pile-up density (start)	mm^{-1}	0.15	0.20	0.20
RMS luminous time (start)	ns	0.186	0.186	0.186
Peak time pile-up density (start)	ns^{-1}	21.2	21.2	21.2
Fill duration	h	8.0	-	7.9
Leveling time	h	3.6	1.3	0
Integ. lumi. at LHCb	fb^{-1}/y	**42.5**	**49.9**	**51.0**
Integ. lumi. at ATLAS/CMS	fb^{-1}/y	257.5	-	256.4

Figure 2.6: HL-LHC parameters and Luminosity Scenario for LHCb Upgrade-II, with different leveled luminosities and dipole polarities for a vertical crossing plane. The values provided assume standard HL-LHC beams parameters and duty cycle. [23].

In the particular scenario treated in Table 2.6, a vertical crossing plane is implemented: a horizontal crossing angle can be used at injection and the crossing plane can be rotated before establishing collisions in LHCb. The vertical crossing allows identical interaction point (IP) characteristics (Luminosity, pile-up, and size of the beam spot) for each detector magnet polarity to be achieved. In this way significant samples with both magnet polarities can be collected and used in order to simplify the study of some sources of systematic uncertainties in CP-violation measurements.

24

Energy deposition

In order to operate at high luminosity and guarantee same or better efficiency in energy deposition through different machine components, additional elements need to be added to the machine layout at each side of the interaction point:

- a TAS (Target Absorber) to protect the inner triplet from quenching, and to limit its radiation dose;

- a TAN (Target Absorber Neutrals) to shield the recombination dipoles from high-energy neutral particles;

- a TCL (Target Collimator Long) to protect the cold magnets in the matching sections from collision debris.

Integrated luminosity

LHCb Upgrade II aims to bring the total data collected by LHCb to $300\,\mathrm{fb}^{-1}$ or more. Assuming the phase composed by two runs each of three years separated by one year of technical stop, this will require about $50\,\mathrm{fb}^{-1}$ of integrated luminosity collected per year. The predictions for LHCb in the standard HL-LHC operational scenario are reported in Reference [33], where it is assumed an operational period for proton-proton physics collision of 160 days/year.

Three different luminosity scenarios were proposed in order to achieve the expected integrated luminosity. The maximum instantaneous luminosity scenario, with peak at $2 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, is disfavoured by the machine as it pushes parameters to their limits and increases the pile-up and occupancy that must be accommodated by the experiment design for a relatively modest gain in luminosity. The lowest instantaneous luminosity scenario ($1 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$) may achieve the target integrated luminosity of $300\,\mathrm{fb}^{-1}$ with seven years of operation. The most favoured scenario is the middle one, with a maximum levelled luminosity of $1.5 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. This luminosity will correspond to an average pile-up of 42 visible pp interactions per bunch crossing and more than 2000 charged particles per bunch crossing produced on average inside the LHCb acceptance.

These conditions will put extreme requirements on detectors, read-out and reconstruction algorithms. One crucial aspect for the LHCb physics programme is to maintain an efficient and clean association of beauty and charm decay vertices to the corresponding proton-proton collision vertices. Simulation studies show that excellent timing resolution in the tracking detectors, such as the VELO, is going to play a crucial role in this. It is expected that a timing resolution of the order of 50 ps per track will allow to keep the fraction of mis-associated vertices at the level of a few percent.

Chapter 3

The Vertex Locator detector

Vertex reconstruction is a fundamental requirement for the LHCb experiment, in fact displaced secondary vertices are a distinctive feature of b-hadron decays. The VErtex LOcator (VELO) provides precise measurements of track coordinates for reconstruction of primary vertices. It allows to identify primary vertices and secondary vertices when particles have large impact parameter with respect to the matched interaction point. This guarantees to achieve the aimed measurements, the accurate reconstruction of beauty and charm hadrons decays.

3.1 Original VErtex LOcator design

The VELO [26] was originally composed by a series of 21 planar tracking silicon stations and operated with this design until the end of Run-2. Each station consisted of a double layer of silicon micro-strip sensors, one measuring the azimuthal angle and one measuring the radial distance from the beam axis, as represented in Figure 3.1. To allow for an overlap between the left and right halves, in order to cover the full azimuthal acceptance and for alignment issues, the detectors in the left and right halves were displaced by 1.5 cm along the beam axis. The two halves are protected from the beam radiation by the *RF foil*, which is an aluminium shield. It separates the machine vacuum from the secondary vacuum in which the modules are located.

27
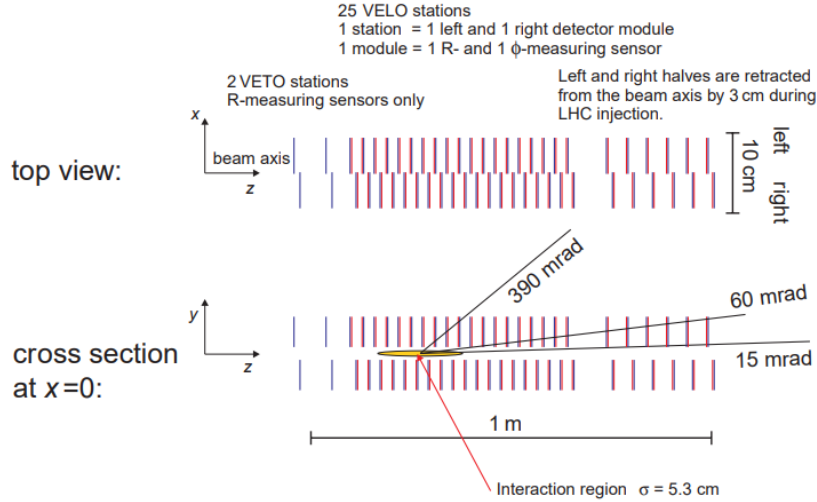
Figure 3.1: Silicon stations along the beam axis. In the top figure one has the sight of the VELO from above, indicating the overlap between the left and right detector halves. In the bottom one, the cross section of the setup at $x = 0$ along the beam axis shows also the nominal position of the interaction area ($\pm 2\sigma$). The three lines indicate the maximum and minimum angular coverage of the VELO and the average angle of tracks in minimum bias events respectively [26].

The VELO was equipped with 50 silicon sersors with circular shape, shown in Figure 3.2. Each sensor, composed by 2048 channels, allowed to evaluate azimuthal (R) and radial (ϕ) particle directions and its particular geometry guarantees a coverage of $\sim 182°$ in both directions. This allowed to obtain good tracking performances and it was possible to decouple the routing of the signals from the strip geometry and to move the electronics as far as possible out of the acceptance, using a double metal layer.
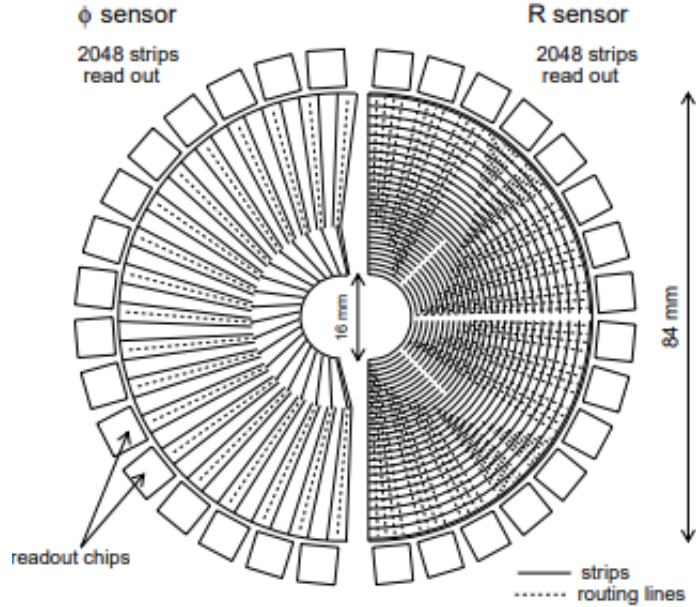
Figure 3.2: Schematic view of a radial and azimuthal measuring sensor.

3.2   The VeloPixel for Upgrade-I

With the Upgrade-I, the number of stations is increased to 26. Each of them is divided into two L-shaped silicon pixel detector modules [34]. The module contains twelve VeloPix ASICs[1], which consist of a matrix of 256×256 pixels of 55×55 µm^2 each. A total number of 41 million of pixels enhance the track reconstruction speed and precision. The pixel structure is described more in detail in Section 4.2.2.

Two groups of three ASICs are mounted with an overlap of 220 µm in vertical direction and the same is done in horizontal direction, in order to ensure full coverage for angled tracks. As the inner horizontal tile is located on the same side as the outer vertical tile, the latter is displaced by 1 mm with respect to the inner vertical tile, resulting in a small acceptance gap. The inner tiles of left and right modules form a square

---

[1]The application-specific integrated circuit is an integrated electronic circuit chip customized for a particular use

acceptance "hole", with the boundaries of the pixel cells closest to the beam being located at $x, y = \pm 5.1$ mm. The hybrid components could be approximated by a 450 µm thick layer of kapton and a 30 µm thick layer of copper, and the GBTx chip can be modelled as a 1 mm thick silicon block with the same lateral dimensions as the VeloPix (see Figure 3.3a). A transverse profile of upgraded modules in a station is represented in Figure 3.3b.
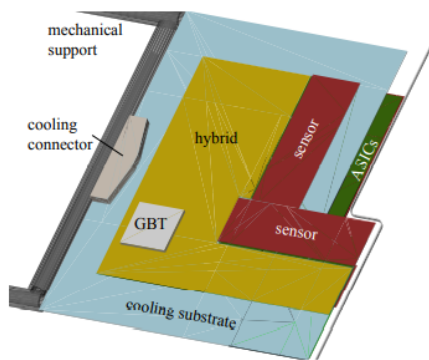
(a) Layout of the components in a VELO module,including a cross section of the RF foil at the z-position. [35]
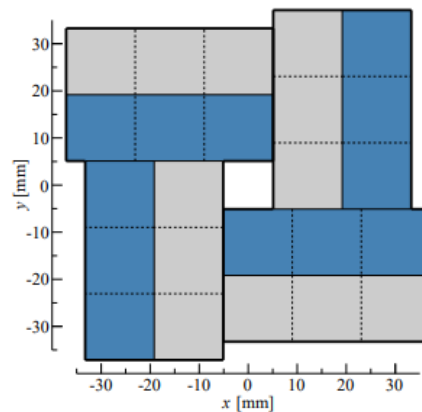
(b) Schematic layout in the transverse xy plane of each VELO station [36].

Figure 3.3: Schematic view of VeloPixel layout and its components

The Upgrade LHCb VeloPixel silicon vertex detector is capable of 40 MHz read-out at a luminosity of 2×10^{33} cm^{-2}s^{-1} [35]. The new architecture is planned to be installed and used by the beginning of the LHC Run 3, after the LS2 shutdown. A summary of some basic parameters of the Upgrade VELO is presented in Table 3.1.

The major developments obtained can be synthesized as follow:

- improvement of the digital architecture of the binary read-out chip;

- optimisation of the module layout and foil shape, through realistic pixel geometry in the simulation;

- introduction of microchannel cooled hybrid pixel modules, through endurance test at high pressure of this technology.

Number of modules	52
Number of ASICs per module	12
Number of ASICs (total)	624
Number of sensors	208
Number of pixels	41M
Number of optical links	1664
Sensor thickness	$200\,\mu m$
Pixel dimensions	$55 \times 55\,\mu m^2$
Position of the first station upstream	$-289\,mm$
Position of the last station downstream	$751\,mm$
Total active area	$1243\,cm^2$
Peak total data rate	$2.85\,Tbit/s$

Table 3.1: System parameters of the Upgrade VELO [34]

.

To allow a read-out at 40 MHz keeping the same (or even improving) its physics performance, it is necessary to replace completely the silicon sensors and electronics with new hybrid pixel sensors, based on new radiation hard ASIC, called *VeloPix*, capable of coping with the data rates. The Upgrade VELO guarantees a cooling system based on circulation of evaporating CO_2 within miniature channels.

Left and right modules are disposed in two different configurations depending on the activity of LHCb: during the data acquisition phase, with stable beam, a fully closed configuration is reached, while in other circumstances, when the VELO is fully open, the two modules are separated by 6 cm vacuum space. A schematic layout of the Upgrade VELO is represented in Figure 3.4. The z positions of the modules have been changed in order to reach similar acceptance given the smaller module size and smaller distance from the beam line to the first measured point.
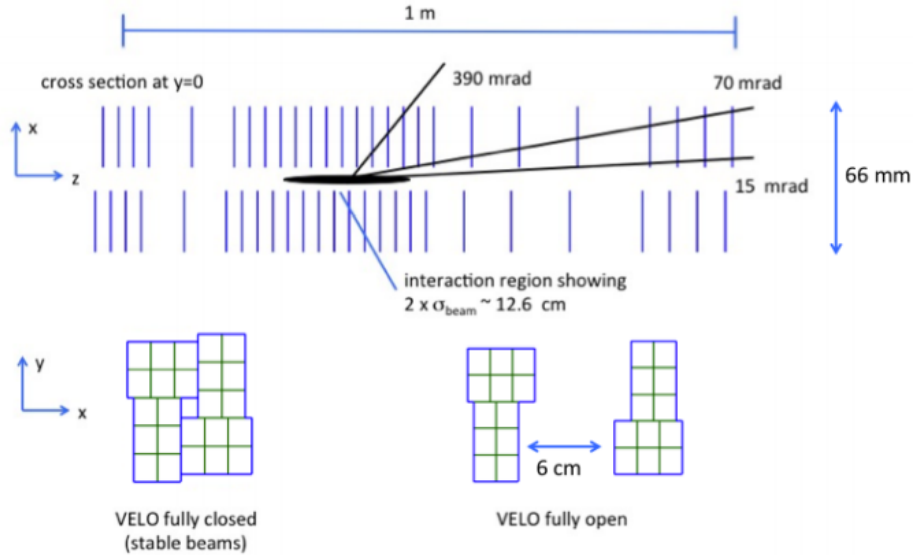
Figure 3.4: Layout of the VeloPixel. In the top figure there is a representation of the cross section at y=0 of the stations, similar to the nominal one in Figure 3.1, with changed z position. The figure below represents the two configurations of the modules in the *xy* plane [34].

Changing the modules shapes carries a changing even in the RF foils, which reduces its thickness to 250 µm, with respect to the 300 µm of nominal value thickness, and follows the L-shape of the modules.

3.3   The VELO Upgrade-II

The larger luminosity expected for Upgrade-II will ask a huge improvement to the read-out system in order to cope with a rate approximately 10 times bigger. If the geometry and triggerless read-out is maintained as it is at present, the ASIC would have to tolerate rates of 8 Ghits/s, with up to 500 khits/s in the hottest pixels; the radiation damage would rise by a factor 7.5, necessitating periodic replacements of modules, and the data output from the detector would exceed 30 Tbit/s, a number which could rise with the

inclusion of extra information from finer pitch pixels and precise time stamps. Even the techniques for real-time pattern recognition have to be improved in order to cope with the increased occupancy. Moreover, reduction of the material in the VELO, such as the remotion of the RF foil, will limit the amount of secondary particles which do not come from the proton-proton vertex and have a strong impact on the detector performances.

The higher luminosity leads, in the VELO region, to a larger combinatorics of the number of vertices inside the VELO and so a more challanging primary and seconday vertices reconstruction, in particular for long-lived particles. Indeed, because of the increased amount of primary vertices (around 50), time-dependent measurements on B-hadrons decays and their proper tracks reconstruction becomes more difficult. To cope with ambiguity when attempting to extrapolate back to a primary vertex candidate, several changes has to be applied to the detector. In the following paragraphs, two different modifications to the Upgrade VELO are described and, during this project, the addition of time as processing information is considered.

Adding time information

A possible solution to face the high luminosity challenges can be obtained by adding to the detector the capability to measure the time to distinguish between vertices. Associate time to each pixel, would transform the VELO into a true 4D tracking device. This is challenging due to the small area available in the pixel Front-End ASIC read-out electronics on top of the already huge Upgrade-II data rates. However a goal of 50 ps or better seems possible on the timescale of Upgrade-II. The timestamp will be present as a series of bits that encodes the information on the time that is associated to a pixel, i.e. the time during which a pixel is hit by a particle. This timestamp leads to cut down the number of combinatorics and this could result determinant for the quality and speed of the pattern recognition algorithm. Even those tracks which fall out of the LHCb acceptance but are reconstructed in the VELO would contribute to the primary vertex reconstruction, boosting statistical power, and the measurement would occur close to the interaction point, reducing time of flight uncertainties.

Removing RF foil

Recently simulation studies has shown that removing the RF foils will improve the detector performances and the impact parameter resolution. The reason is the large amount of low momentum particles arising from the interaction of prompt particles with its matter. Moreover, the RF foils increase dimensions of pixels and, as consequence, of the material scattering effect. However, a solution for shielding the electronics from radio-frequency pick up must be found: it could be the adoption of wires, or maybe ultra-thin foils.

3.3.1 The TIMESPOT collaboration

The TIME and SPace real-time Operating Tracker (TIMESPOT)[37] is an INFN project started in 2017 with the aim of provide tracking alternatives able to cope with the high luminosities expected by the harsh scenario of Hi-Lumi LHC. The collaboration with TimeSpot for the development of the fast simulation produced interesting studies on possible improvements on the Upgrade-II VELO.

More in details, the TimeSpot work package[38] in which the project is inserted in is dedicated to the elaboration of fast-tracking algorithms and devices. Two are the topics elaborated on:

- the introduction of FPGAs[2] for clustering processes at the read-out level, amply studied in Section 4.4.1;

- the application of geometry alternatives to the traditional sensor ones, like the introduction of trench pixels, discussed more in detail in Section 4.3.2.

In the first case TimeSpot studies lead to the development of an highly parallel algorithm that makes use of both precise timing ans space information of the particle hits in the detector. The architecture implemented makes the system modular and suitable

---

[2]The Field Programmable Gate Array is an electronic hardware device that allows to be configured by a customer. The FPGA is usually programmable throgh an hardware description language (HDL), similar to that used for an ASIC.

for the application on multiple FPGA boards. The FPGA technology allows for high performances in specific tasks, low latency and highly pipelined implementations, keeping the possibility of a programmable logic. In the second case, the TimeSpot initiative guided to the ideation of an innovative 3D sensor prototype, characterised by cylindrical electrodes penetrating deep into the bulk material, perpendicularly to the surface. The structure shows a very good radiation hardness because it is able to decouple the charge carrier drift distance from the sensor thickness. Moreover, the fact that the charge carriers are collected perpendicularly to the sensor thickness minimises time uncertainties due to their diffusion. The sensor geometry results then optimized in terms of timing performance.

Chapter 4

Fast Simulation for Upgrade-II

In this Chapter, the original thesis work is presented. It consists in the development of a code to perform am advanced simulation of the VELO detector in the Upgrade-II scenario. The software developed can simulate the response of the detector, including the information on time, given by a hypothetical sensor, like the one developed by the TimeSpot Collaboration. In addition, the code mimics a read-out system, producing the so-called raw banks, that can be used as input to test real-time fast reconstruction solutions, like the one implemented on FPGAs.

To evaluate the impact of the harsh scenario described in Section 3.3 to the detector performances, a few changes are tested one-by-one using Upgrade-I detector as a baseline [39]. In this way it is possible to test the impact of the detector modifications with pre-existing pattern recognition algorithms and compare them to the baseline. The official LHCb simulation aims at reproducing the same layout and geometry of the Upgrade VELO and, once validated, to treat the effects brought by these two different changes:

- introduction of a timestamp per hit, with precision $\mathcal{O}(50\,\mathrm{ps})$;

- removal of the RF foil or reduction of its thickness

The studies on the improvements of the Upgrade-II are obtained using two different frameworks:

- a full and official LHCb simulation;

- a fast and parametric simulation of the response of the detector and track reconstruction.

In particular this thesis focuses on studying the addition of time information to the pixel sensors. A fast simulation is developed, to have more flexibility relative to the geometric parameters of the detector, sensors resolutions and reconstruction algorithms [40]. The fast simulation has been validated in all of its part by the full one.

In the following chapter it is provided a detailed discussion on how the fast simulation has been developed. The main differences between the two simulations are shown. Different parameters are used to validate the fast simulation with the full one under Upgrade-I condition, such as the deposit of charge and the tracks reconstruction performances. Then, the changes implemented to face the Upgrade-II challenging scenario are discussed, in particular sensor developed by TimeSpot collaboration [1] is included. Because of the high rate of data acquisition required, the fast simulation has been developed to manage data in a particular kind of format, called *raw bank*, useful for FPGA studies. A successful clustering algorithm[41], already developed at LHCb, based on raw banks is implemented in the fast simulation and has to be tested yet. The final section of the chapter is dedicated to the future prospects that has the project. First, the application of the clustering algorithm to the FPGAs and the evaluation of the performances obtained in comparison with the CPU ones. Then, the track reconstruction has to be evaluated for the fast simulation, basing on a TimeSpot analysis of the performances for 4D tracking FPGA-based algorithm [42]. Finally, an additional measurement can be performed for the reconstruction of primary vertices.

4.1 Differences between the two simulations

Both simulations adopt as input the results of particle collision simulation obtained through `GAUSS` and `PYTHIA`. The LHCb package (`GAUSS v53r2`) is used to study relatively large samples ($\sim 20k$ events) in the context of the vertex detector, while the simulation of pp interaction at 14 TeV is obtained through `PYTHIA` [43], a program used

for event generation in high-energy physics. Moreover, the decays of unstable particles are described by the program `EVTGEN` [44], while their final state radiation is generated through `PHOTOS` [45]. Finally, to describe the interaction between particles and the detector, `GEANT4` [46] was used. The detector layout adopted is the one described in [47]. In order to evaluate the beam parameters in terms of the luminosity required, as represented in Table 2.6, an algorithm for both time and space is adopted. The resulted emmitance of 2.5 µm came from a luminous region of 15 µm transverse to the beam, 44 mm in the longitudinal direction and 180 ps in time. The luminosity, used in the simulation, is $1.5 \times 10^{34}\,\mathrm{cm^{-2}s^{-1}}$ [39] with 2400 colliding bunches (7.5 times Upgrade-I), resulting in an average number of visible interactions per bunch crossing of about 44. The main differences between fast and full simulations are discussed in the following paragraphs.

Full simulation

Differently from the fast simulation, the full simulation uses two different frameworks of the `GAUDI` project [48], `BOOLE` [49] for reading the output of `GAUSS` simulation and `BRUNEL` [50] for the track reconstruction.

All the information, included time, that regards the passage of a charged particle through the VELO are saved in `MCHits` during the simulation phase. Then, they are used to create clusters, useful basic structure for the tracking algorithm to study the impact of detector resolutions. As advantage, this simulation guarantee a nearly complete description of the interaction of charged particle with the upgraded VELO. Unfortunately, it is a bit restrictive as changes of the detector geometry require a significant amount of work.

Fast simulation

The fast and parametric simulation adopted, exactly as the full one, `PYTHIA` and `GAUSS` for the event simulation and the interaction with the material using the Upgrade I conditions. The simulation, validated with the full one, has implemented approximately the RF foils and contains a charge deposit simulation described more in detail in Section 4.2.1. The main aim of the fast simulation is to test several different layouts in an easy way

and to perform studies relative to the possibility of introduction of the time coordinate associated to the passage of the particle through the detector or the reduction/removal of RF foils. Before start describing how the fast simulation is made, it is possible to access to the developed software through the link:

https://gitlab.cern.ch/acarbone/vpfastsim/tree/master

4.2 Fast simulation validation under Upgrade-I conditions

The fast simulation is based on the detector geometry of the Upgrade-I (see Section 3.2). The numbering of the modules follows the z-positions, i.e. the z-positions of the modules increase from module 0 to module 51. The ones in the half at positive x (*left modules*) have even numbers, while the modules in the half at negative x (*right modules*) have odd numbers. Each module is composed by eight sensors, which are also numbered consecutively (from 0 to 207), with sensors $0 - 3$ being located on module 0, sensors $4 - 7$ on module 1 and so on.

The sensor contains 256×768 pixels, divided into three chips (matrices of 256×256 pixels) as represented in Figure 4.1. Each pixel has a pitch of 55 µm in the local x and y directions.
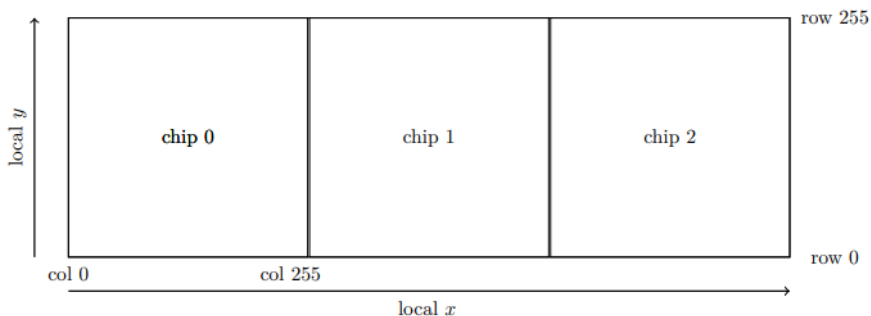
Figure 4.1: Representation of a sensor divided into three chips in terms of local $x - y$ coordinates. [36]

39

A particle which crosses a pixel produces a charge deposit as described in Section 4.2.1. This charge is used as parameter to validate the fast simulation. An additional test on the accuracy of the developed software is guaranteed by the study of the performances of tracking reconstruction, in comparison with the ones obtained with the full simulation.

4.2.1   Passage of charged particles through matter

When a particle crosses a material, it is generally characterized by a loss of energy through different interactions and a deflection from its incident direction. These effects are results of inelastic collisions with the atomic electrons of the material and elastic scattering from nuclei.

Energy Loss of Charged Particles

When particle loses energy inside the material, it is able to ionize its atoms, transferring enough energy to allow the removal of one or more electrons from their external shells. If the energy lost by the particle is not enough to remove the outermost electron, the target atom is excited. As consequence, it emits a photon through de-excitation [51].

The mean energy loss through ionization by a moderate relativistic ($0.1 \leq \beta\gamma \leq 1000$) heavy particle with charge z and velocity β in a medium with atomic number Z, mass number A is well-described by the Bethe-Bloch equation

$$\left\langle -\frac{dE}{dx} \right\rangle = K z^2 \frac{Z}{A} \frac{1}{\beta^2} \left[\frac{1}{2} \ln \frac{2 m_e c^2 \beta^2 \gamma^2 W_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right]. \qquad (4.1)$$

It is called *mass stopping power* and it is measured in $\mathrm{MeVg^{-1}\,cm^2}$. At low velocity ($\beta\gamma \sim 0.1$), the particle has a velocity comparable with the atomic electron one and in this case the energy loss is governed by a factor $1/\beta^2$. For higher velocities relativistic effects took place and, after a point of minimum, called *MIP* (Minimum Ionizing Particle), a logarithmic rise begins, as represented in Figure 4.2. The last factor in Equation 4.1 is the *density correction*, due to the polarization of the atoms in the material once the particle crosses it. Polarized atoms act as shield for furthest atoms, reducing in this way the long range interaction of particles. This effect is larger for large $\beta\gamma$.
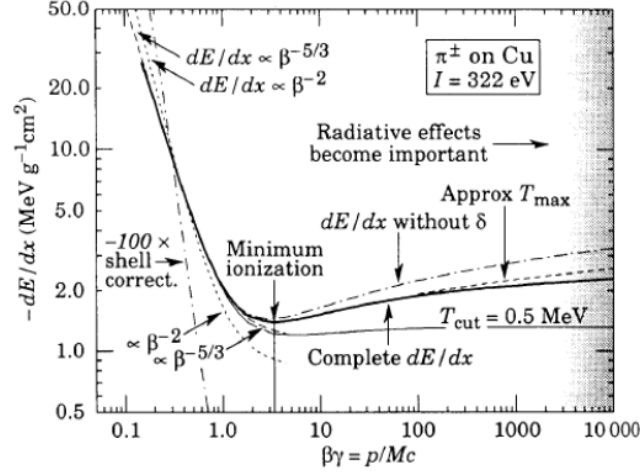
Figure 4.2: Mass stopping power in terms of the velocity of the particle crossing the material [51].

The energy loss rate $-\frac{dE}{dx}$ (with the minus sign to represent the fact that the energy is loss and not gained by the particle) does not depend on the thickness of the material [52].

The number of electrons produced in silicon by the passage of a charged particle through the material depends on the energy loss of the particle itself, which is simulated according to a Landau distribution. The energy loss expressed in Equation 4.1 is a mean quantity, which means that for the same path of the same particle, at fixed momentum we can obtain different values for repeated measurements, due to statistical fluctuations. The total energy loss is then a statistical process. The most probable energy-loss in detectors with moderate thickness x is given by

$$\Delta_p = \xi \left[\ln \frac{2mc^2\beta^2\gamma^2}{I} + \ln \frac{\xi}{I} + j - \beta^2 - \delta(\beta\gamma) \right], \tag{4.2}$$

where $\xi = \frac{K}{2}\frac{Z}{A}z^2\frac{x}{\beta^2}$ (in MeV for x in gcm^{-2}) and $j = 0.200$ [53]. Differently from the energy loss rate, Δ_p/x scales as $a\ln x + b$. Thus, for the Landau distribution, the mean is not the value with higher probability, because the tail makes mean greater than most probable value. The fact that the probability distribution of the energy loss is asymmetric is the reason why this phenomenon is called *energy loss straggling*. The

distribution function for the energy deposit by 500 MeV pions in silicon, where the ionization energy is 3.6 eV is represented in Figure 4.3a for different layer thickness. For very thick absorbers the Landau distribution is less skewed and approaches to a Gaussian with a tail at high energies, while it usually fails in the description of energy loss in thin absorbers, such as TPC gas chamber [54]. Figure 4.3b shows the most probable energy loss scaled with respect to the MIP ($388\,\mathrm{eV\mu m^{-1}}$) in terms of the velocity of the particle for several detector thickness.
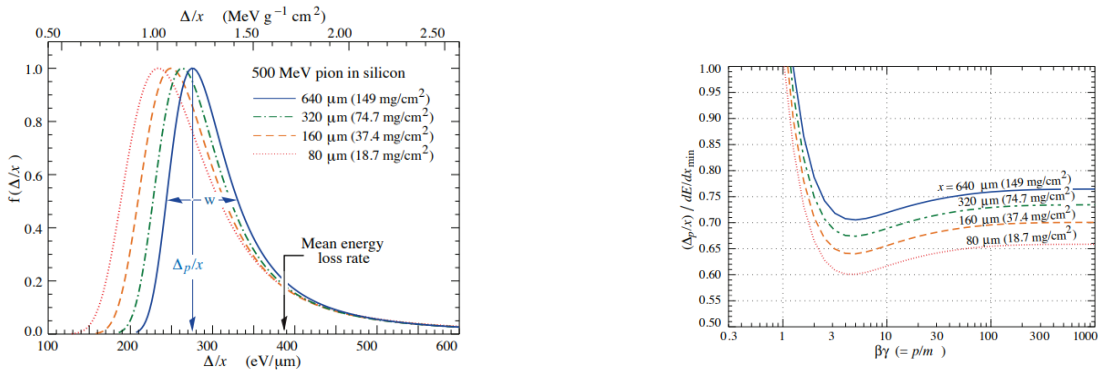
(a) Energy loss probability distribution in silicon for 500 MeV normalized to unity at the most probable value (Δ_p/x). w represents the FWHM.

(b) Most probable value of energy-loss scaled to the mass stopping power at the minimum in terms of $\beta\gamma$.

Figure 4.3: Energy loss representations for several silicon detector thickness [53].

Multiple scattering

When a particle crosses a medium is deviated from its original direction by small angles (see Figure 4.4), mostly due to Coulomb scattering from material nuclei. For small-angles scatters, the distribution follows a Gaussian, as expected from the central limit theorem, while more "hard" scatters produce a non-Gaussian tail [55].

It is possible to evaluate the distribution of angular deviation with respect to the original direction approximated as Gaussian distribution [56], with a RMS width given by:

$$\theta_0 = \theta_{plane}^{RMS} = \frac{13.6\mathrm{MeV}}{\beta c p} z \sqrt{\frac{x}{X_0}} \left[1 + 0.038 \ln \left(\frac{x z^2}{X_0 \beta^2} \right) \right] \tag{4.3}$$

where βc is the velocity of incident particle, p is its momentum, z is its charge and x/X_0 is the number of radiation length[1] crossed by the particle inside the material.
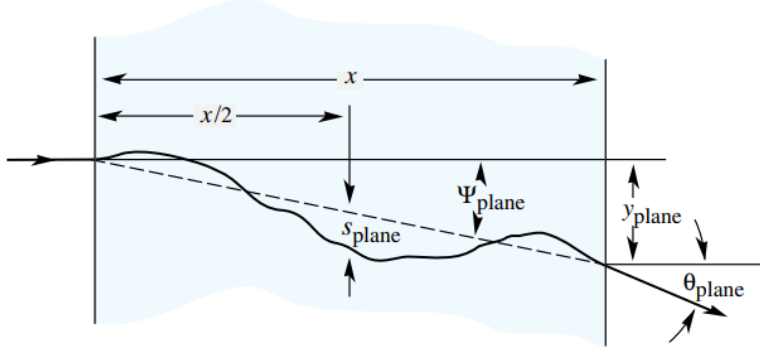
Figure 4.4: Schematic representation of deviations through small angles of incident particle passing through a medium with thickness x [56].

The multiple scattering, in addition to an angular deviation θ, brings a vertical shift y, as described in Figure 4.4. Those quantities have a correlation coefficient $p_{y\theta} = \sqrt{3}/2 \sim$ 0.87. For the Monte Carlo generation of a joint $(y_{plane}, \theta_{plane})$ distribution it is more convenient to use two Gaussian random variables z_1 and z_2, with zero mean and unitary variance, and then set

$$y_{plane} = z_1 x \theta_0 \frac{(1 - p_{y\theta}^2)^{1/2}}{\sqrt{3}} + z_2 p_{y\theta} x \frac{\theta_0}{\sqrt{3}},$$

$$\theta_{plane} = z_2 \theta_0.$$

(4.4)

The second term for y_{plane} represents the displacement that would have occurred for a deflection θ_{plane} at the point $x/2$.

4.2.2 Deposition of charge inside the pixel

The fast simulation implements the deposit of charge in the material through the energy loss of the particle traversing the VELO modules material, presented in Equation 4.2.

---

[1]The radiation length (measured in gcm^{-2}) represents the mean distance over which an electron, after the loss of energy through bremsstrahlung, has a fraction $1/e$ of its initial energy.

The total path crossed by a particle is divided in intervals of 5 µm, as shown in Figure 4.5, and for each interval the number of electrons released is randomly extracted from a Gaussian distribution with a $\frac{1}{n^2}$ contribution for the Landau tail. Even the dispersion of charge in neighbouring pixels and diffusion effects on z-direction are considered.
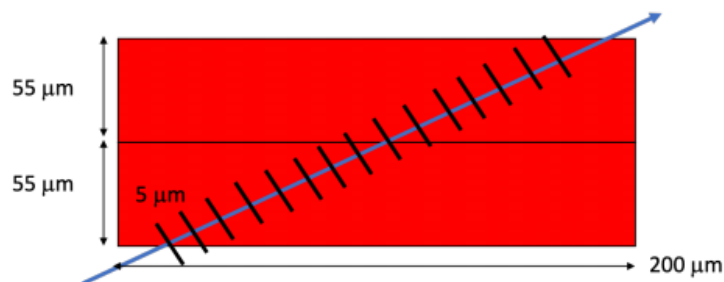
Figure 4.5: Schematic view of the passage of a particle through two adjacent pixels (that are 200 µm thick and have an effective area of 55 µm × 55µm)

The charge deposit of fast simulation is validated with the full one, as represented in Figure 4.6. In Figure 4.6a it is represented the number of electrons released per interval (every 5 µm), with a peak around 350 e^-/interval. Figure 4.6b shows the distribution of the number of electrons collected by each pixel.
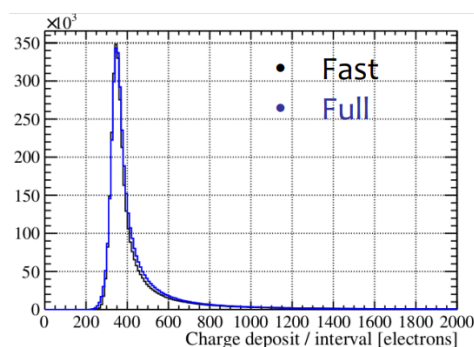
(a) Distribution of number of electrons produced by ionization in 5µm.
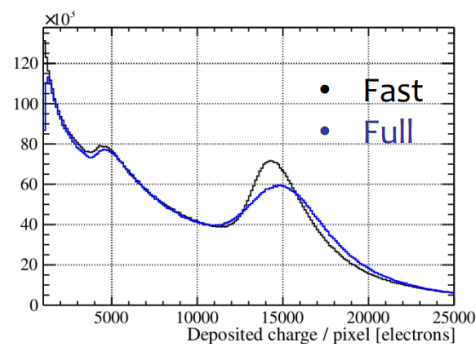
(b) Distribution of electrons coming from the ionization of silicon pixel by the passage of a particle.

Figure 4.6: Charge deposit by the passage of a particle inside the silicon modules

It is possible to observe a more pronounced peak at 15000 electrons, corresponding to

44

the case when a particle completely crosses the pixel without depositing energy to any of its neighbours. The peak at 5000 electrons represents instead the case when a particle crosses two pixels while the large amount of low deposited energy is given by particles crossing more pixels and other diffusion effects.

The noise of the VELO Upgrade-I silicon sensors is assumed to behave like a Gaussian of $\sigma = 150$ electrons. For this reason a pixel is activated with a minimum charge of 1000 electrons. This threshold is much more than 5σ of the background and so can be largely considered signal.
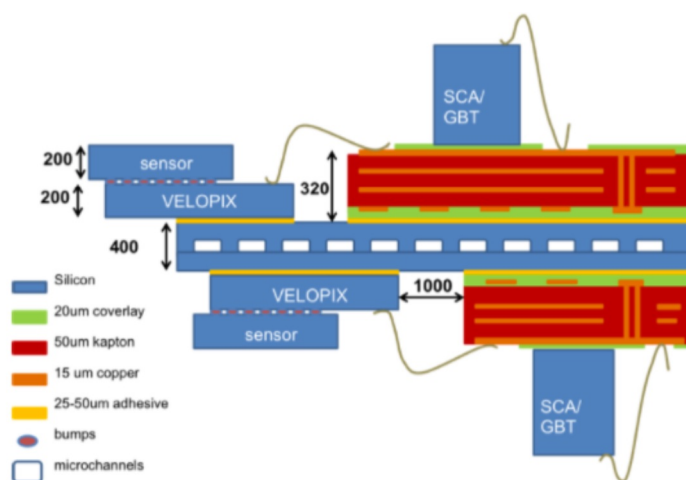
Figure 4.7: Trasverse profile of the silicon layer.

Fast simulation keeps in consideration even the multiple scattering on RF foils, making some approximations on the geometry of the detector. The RF foil is considered as an aluminium cylinder with radius 5 mm and width 350 µm, while each silicon layer has a total thickness of 800 µm (see Figure 4.7).

The multiple scattering is implemented inside the fast simulation as an option that can be switched off in case of necessity to perform studies of ideal cases. The introduction of multiple scattering lead to reduction of tracking and primary vertices reconstruction performances, but allowed to have a more realistic simulation. A proof of the validation for the multiple scattering implementation is provided in Figure 4.8, where the resolution of the impact parameter is plotted in terms of the inverse of the transverse momentum. It is possible to notice that in the case of fast simulation without the multiple scattering

implementation, the impact parameter resolution results, as expected, independent by the tranverse momentum. Multiple scattering covers a crucial role in the determination of this dependence. Random deviations, carried by the scattering with material nuclei on the crossing particle, have root mean square

$$\sigma_{IP} = \frac{13.6\text{MeV}}{\beta c p_T} z \sqrt{\frac{X}{X_0}} \left(1 + 0.038 \ln \frac{X}{X_0} \right).$$ (4.5)

The correct dependence of the impact parameter resolution with $1/p_T$ is reproduced when the multiple scattering is included in the fast simulation.
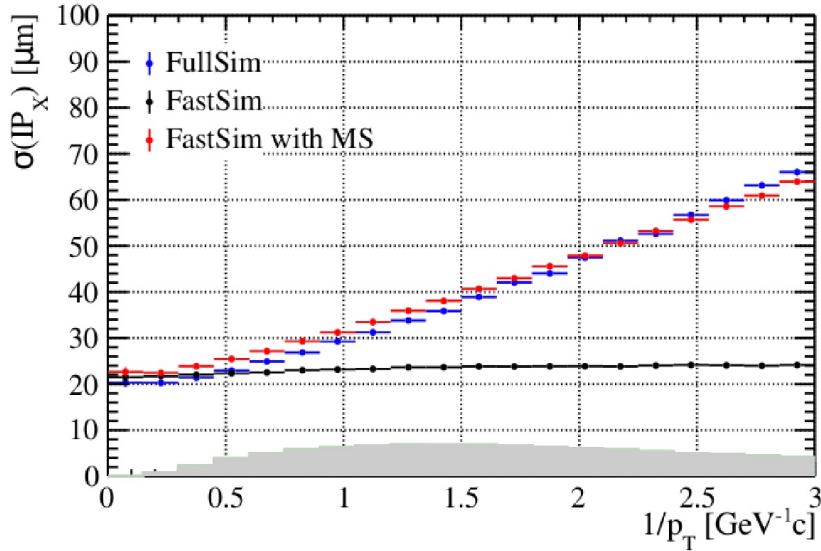
Figure 4.8: Resolution on the impact parameter vs $1/p_T$.

4.2.3   Clustering

The process of clustering correspond to group neighbouring pixels on a sensor. A cluster is defined as an isolated group of eight-way connected active pixels, for this reason diagonal neighbours are treated in the same way as horizontal and vertical neighbours. Thus, a pixel fully-contained inside the sensor has eight neighbours, a pixel on the edge of the sensor has five and a corner pixel has three. The net effect of clustering is the creation of a bitmap representation (2D array of bits) of a sensor composed by active (1) and non-active (0) pixels.
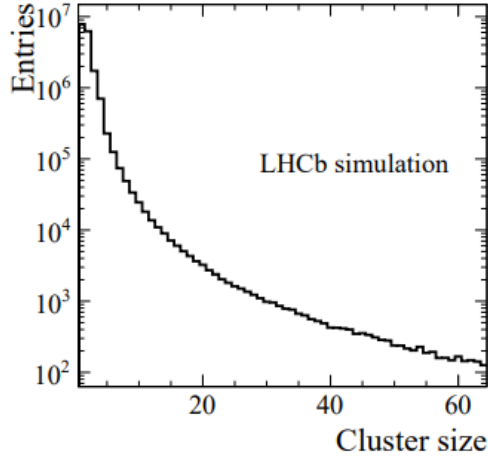
Figure 4.9: Distribution of cluster sizes for 10000 events

Clusters can have arbitrary size, however usually the dimension of clusters is relatively small. For 10000 events, the most of created clusters have size smaller than five (see Figure 4.9).

An example of clustering algorithm is presented as follows. A cluster is created starting from a single active pixel. The algorithm searches for any other neighbouring active pixels and if it founds any, then the founded pixel is added to the cluster. The algorithm continues recursively until no other neighbouring active pixels are found. Once the cluster is formed, it will represent an valid hit for the pattern recognition. For the position of this hit, the mean of x and y coordinates of each single pixel that forms the cluster is used. If even the time coordinate is considered (see Section 4.3.1), the time associated to an hit is given by the mean of the time of every pixel in the cluster. Another example of clustering algorithm is presented in Section 4.4.1.

4.2.4 Track reconstruction

The pattern recognition allows to reconstruct tracks starting from the hits resulted from clustering. The main algorithm of track reconstruction is based on looking for pairs of hits in neighbouring stations which are compatible with track slopes $\left|\frac{dx}{dz}\right| < 0.4$ and

$\left|\frac{dy}{dz}\right| < 0.4$. By an extrapolation in the upstream direction, the closest hit is seached inside a window of position prediction and it is added if it passes a cut on the maximal scattering angle, as represented in Figure 4.10. If the time information is included in the simulation, it represents an additional information for the pattern recognition process. In this case, indeed, in order to reconstruct a real track, the distance between the hits candidates has to be compatible with the distance travelled by a particle with velocity c. For tracks reconstructed using three hits, it is required that all of them are unused by other tracking candidates. The tracking algorithm implemented for the TimeSpot project, described in Reference [42] consists in the identification of doublets of particle hits in adjacent detector planes and the precise timing information is used to suppress fake hit combinations. The candidates for pattern reconstruction are identified in parallel from groups of stubs, couples of hits in adjacent sensors, with similar parameters. Differently from the one previously described, this algorithm requires at least four hits for the proper reconstruction of a track. Further details on this algorithm are reported in Section 4.4.2.
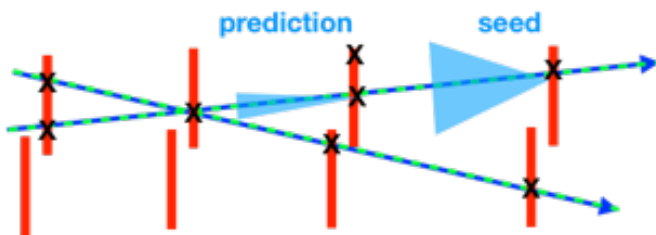
Figure 4.10: Schematic of the track reconstruction by looking the presence of any hit in the upstream direction

Tracking system of fast simulation has been validated with the full one, as shown in Figure 4.11 using a sample of 15000 events. The reconstructed tracks from fast simulation are represented with dots, while the line is the expected distribution from full simulation. neri
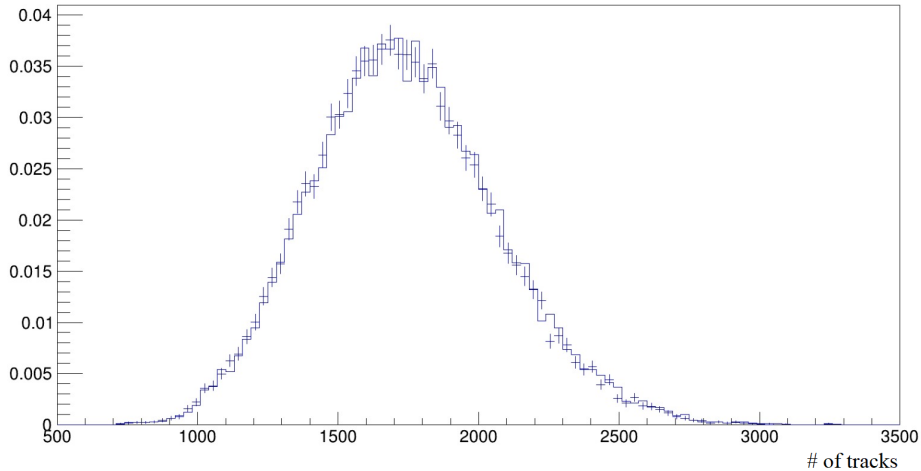
Figure 4.11: Validation of fast simulation tracking through the good agreement of reconstructed tracks for fast (dots) and full (line) simulation with 15000 events.

Tracking performances

The performances of pattern recognition are evaluated through standard figure of merit such as the efficiency and the ghost rate. The efficiency of tracking is given by

$$\epsilon = \frac{N_{reconstructed}}{N_{reconstructible}}, \tag{4.6}$$

where a track is considered reconstructible if there are clusters associated to it on three or more modules. This parameter can be calculate for several sub-sample of particles, some of them are the VELO and long particles. The VELO particle is defined as a particle with pseudo-rapidity $2 < \eta < 5$ and which is not an electron. Moreover it is reconstructible as a *long particle* if it is a VELO particle with components which satisfy $12 < \left|\frac{p_x}{p_z}\right| < 300$ and $12 < \left|\frac{p_y}{p_z}\right| < 250$. A particle is considered as reconstructed if at least 70% of the (cluster) measurements on a track are associated to the Monte Carlo particle.

One of the main source of inefficiency is represented by the *ghost track*, or fake track. They cannot be associated to any simulated particles. The ghost rate is evaluated as

$$R_{ghost} = \frac{N_{non-assignedtracks}}{N_{reconstructedtracks}}. \tag{4.7}$$

49

Ghost tracks have less then 70% of clusters associated to a MC particle. In Table 4.1, efficiency and ghost tracks are compared between fast and full simulations. The validation of fast simulation for track reconstruction is presented also in Figure 4.12, where efficiency and ghost rates are plotted as function of the number of primary vertices. The performances obtained are slightly different: the efficiency of VELO particles is a bit overestimated, while the ghost rate is underestimated for the fast simulation with respect to the full one.

Simulation	$\epsilon_{VELO}[\%]$	$R_{GHOST}[\%]$
U-I Fast Simulation	98.0	0.50
U-I Full Simulation	97.8	0.44

Table 4.1: Table with values of efficiency and ghost rate for VELO particles for fast simulation, without considering the RF foils, and for full simulation.
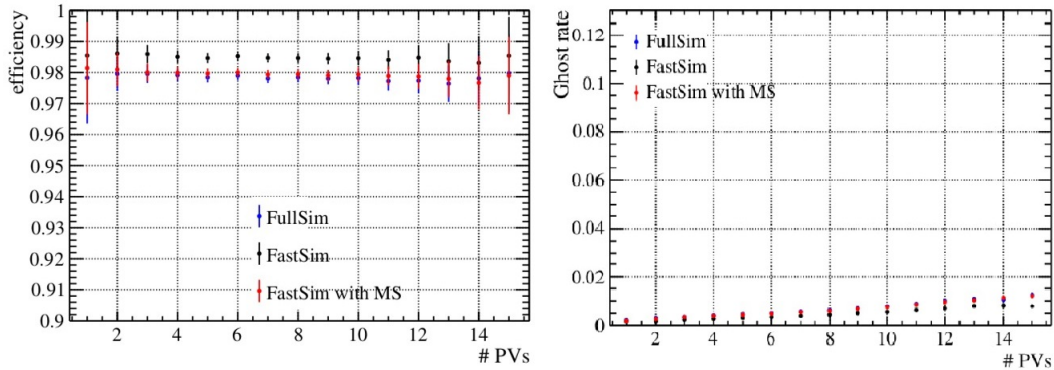
Figure 4.12: Validation of fast simulation performances (black dots) with respect to the full ones (blue dots). The tracking efficiency (on the left) of the fast simulation is slightly overestimated, while its ghost rate (on the right) is a bit underestimated.

4.3 From Upgrade-I to Upgrade-II

The fast simulation aim at testing the contribution in terms of performances provided by the addition of time information to each pixel sensors. A particular kind of sensor,
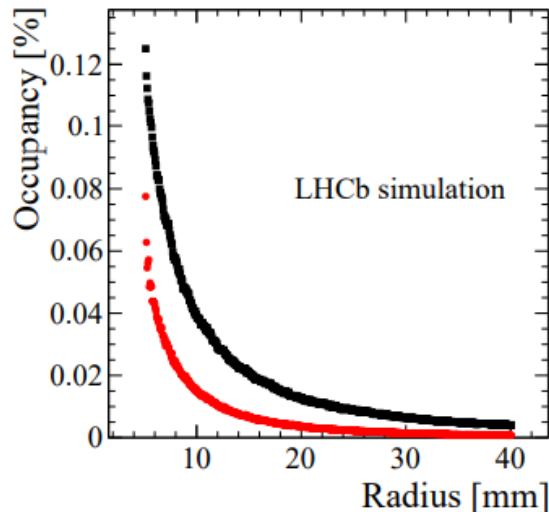
Figure 4.13: Pixel occupancy (black squares) and cluster occupancy (red circles) as function of radius for the station closest to the nominal interaction point.

developed by TimeSpot collaboration, is also included to face the Upgrade-II challenging scenario. The developed software manages data in a format designed in order to produce binary outputs, that can be eventually send to FPGAs for the clustering and track finding. A clustering algorithm, based on a 3D FPGA-based one, is included in the simulation.

The Upgrade-I conditions change drastically passing to the Upgrade-II, as discussed in Section 3.3. The operational conditions require an improvement of the already challenging ones present in the current VELO by a factor of 7.5. The rate of events to deal with is approximately ten times bigger. The several changes between the two Upgrade VELO can be understand just analyzing as example the occupancy in both cases.

The occupancy represents the fraction of particles present in a given area of the detector. The occupancy is generally higher for regions near the interaction point, around the module 15. Furthermore, the occupancy of the VELO decreases with the distance of the beam axis, as shown in Figure 4.13 for the Upgrade-I.

Figure 4.14a and Figure 4.14b represent histograms of simulated occupancy in terms of

the z position for all the particles and only VELO particles[2], respectively, using a sample of 100 events under the Upgrade-II conditions.
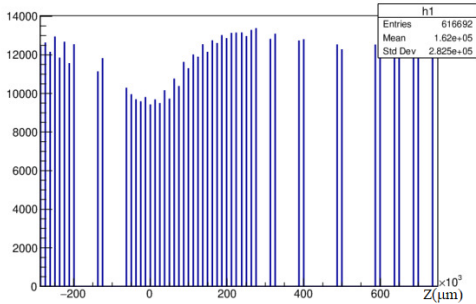
(a) Distribution of the occupancy vs z for all particles.
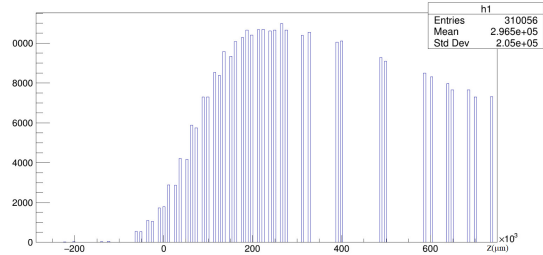
(b) Distribution of the occupancy vs z only for VELO particles.

Figure 4.14: Simulated occupancy in terms of the modules position. The peak corresponds to the most occupied module, near the interaction point.

An example of the transversal plane occupancy for the module 15 for an event of the fast simulation is reported in Figure 4.15a and in Figure 4.15b for Upgrade-II scenario. The increased occupancy in the Upgrade II VELO involves, as consequence, a less efficient track reconstruction, due to the increased number of combinatorics and tracks to be reconstructed. Data reported in Table 4.2 are relative to official LHCb simulations in both cases of Upgrade-I and Upgrade-II. A possible solution to this problem could be the introduction of time associated to the pixel, as previously reported in Section 4.2.4.

Upgrades	n_{tracks}	$\epsilon_{VELO}[\%]$	$R_{ghost}[\%]$
U-I	215	98.1	0.5
U-II	1690	95.4	4.2

Table 4.2: Comparison of the performance of tracking algorithms between Upgrade-I and Upgrade-II conditions. Performances evaluated in the Upgrade-II case are obtained with a simulation (full) that already includes the reduction of RF foil to a thickness of 150 µm

.

---

[2]A VELO particle is a non-electron with $2 < \eta < 5$ that activates at least 3 pixels.
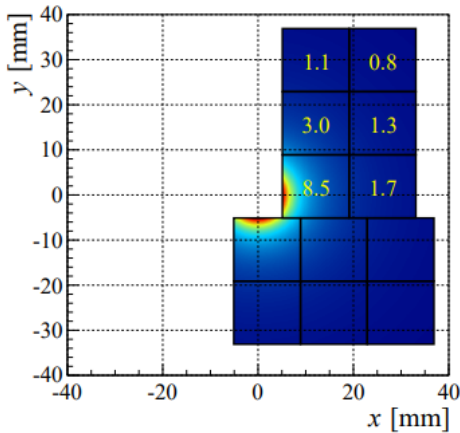
(a) Mean number of particles crossing ASIC per event in the Upgrade-I case.
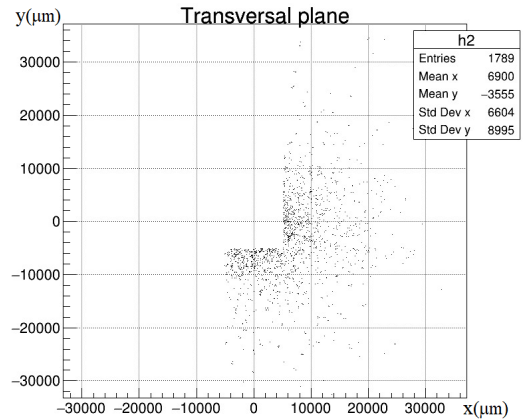
(b) Representation of the active pixels for the transversal module closest to the interaction point using fast simulation.

Figure 4.15: VELO occupancies in terms of the radius distance to the beam axis for Upgrade-I (left) and Upgrade-II (right).

4.3.1 Time Information

In order to efficiently reconstruct and distinguish tracks and face to the increased number of primary vertices a study relative to the use of the time information is conduced. The introduction of a fourth coordinate could lead to an improvement in tracking reconstruction. The effects on tracking performances related to this enhancement are discussed more in details in Section 4.3.1.

The time information associated to each pixel is evaluated in the fast simulation as

$$t_{px} = t_{OV} + \frac{1}{c\beta} |z_{px} - z_{PV}| \frac{p}{p_z}. \tag{4.8}$$

It results from two different contributions. The time of flight of the particle to reach the pixel from the primary vertex position and the origin vertex time, determined from a given clock start, which correspond to the initial time of the event. The first term of the equation, the origin vertex time, is determined by `GEANT4` in the simulation, starting from the time of creation of its primary vertex. Unfortunately, the latter is not simulated by `GAUSS` and it is set out of the framework with a Gaussian centered at 0

53

and with $\sigma = 212\,\mathrm{ps}$. The second term of the equation represent the time of flight of the particle until it reaches a given pixel.

In the case of a pixel which is crossed by more than one particle, the first passage time is taken as representing the pixel and the information relative to the crossing time of the second particle is lost. The error involved by this assumption could be neglected considering that almost all the pixels are hit by only one particle (99.17% of the total pixels). Only a small fraction of the pixels (0.8%) are crossed by two different particles and even smaller frequencies characterize pixels which are crossed by more than two particles (0.004% by three particles, 0.002% by four particles and so on).

Effects of time addition to track reconstruction

The introduction of time information associated to each pixels reduces the combinatoric problem and improves the track reconstruction, increasing the efficiency and reducing the ghost rate. In order to consider a realistic use case, it is important to simulate different pixel time-resolution scenario. Table 4.3 shows the efficiency for VELO and long tracks and the ghost rate for different hit time resolution. It is possible to notice that increasing the time window for the resolution will lead a better efficiency for track reconstruction, but even an higher probability of ghost tracks. The efficiency and the ghost rate in terms of the number of primary vertices are shown in Figure 4.16a and 4.16b respectively. The reduction of the time resolution improves the tracking performances.
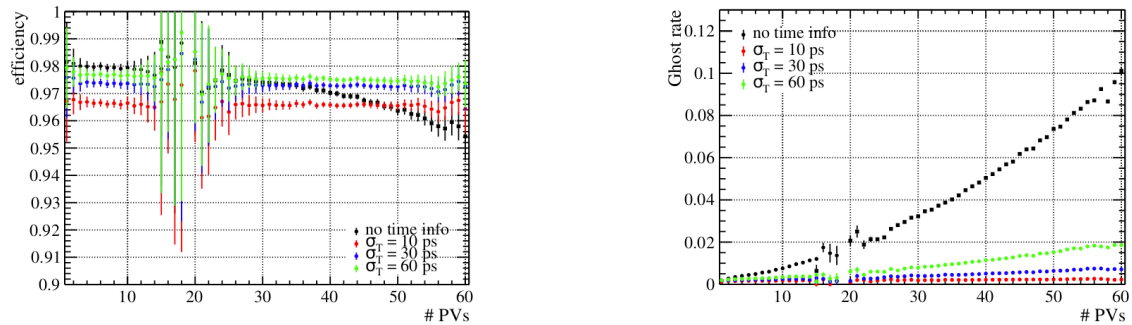
(a) Efficiency in terms of primary vertices for different time resolution.

(b) Ghost rate in terms of primary vertices for different time resolution.

Figure 4.16: Efficiency and ghost rate for different timing resolution.

54

Time Resolution	3 hits			4 hits		
	$\epsilon_{VELO}[\%]$	$\epsilon_{LONG}[\%]$	$R_{GHOST}[\%]$	$\epsilon_{VELO}[\%]$	$\epsilon_{LONG}[\%]$	$R_{GHOST}[\%]$
$\sigma_t = 10\,\mathrm{ps}$	97.5	97.1	0.23	98.5	98.2	0.14
$\sigma_t = 20\,\mathrm{ps}$	97.9	97.6	0.32	98.9	98.7	0.17
$\sigma_t = 30\,\mathrm{ps}$	98.0	97.8	0.43	99.0	98.9	0.20
$\sigma_t = 40\,\mathrm{ps}$	98.1	97.8	0.56	99.0	98.9	0.24
$\sigma_t = 50\,\mathrm{ps}$	98.1	97.8	0.72	99.0	99.0	0.28
$\sigma_t = 60\,\mathrm{ps}$	98.0	97.9	0.90	99.0	99.0	0.33
NO TIMING	97.8	97.7	4.20	98.7	98.9	1.32

Table 4.3: Comparison between tables without considering time (on the left) and adding time information (on the right), with values of efficiency and ghost rate for VELO and long particles for different timing resolution. On the left, values are associated to tracks reconstructed using at least three hits, while on the right the minimum number of hits to make a track is four: in this case performances slightly improve.

4.3.2 Simulation of the TimeSpot sensor

Fast simulation adopted an alternative particular type of pixel for the charge collection, called *trench pixel*, leaving the possibility to switch to the standard pixel described in 3.2. This kind of pixel is implemented in the TimeSpot sensor and its performances are tested in [1]. The 3D silicon sensor has the advantage to collect signal amplitudes which are independent on inter-electrode distance and electrode shape. In this way, a large freedom is left for the optimization of the performances through the geometry customization. Moreover, in order to guarantee the fast time application required by the Upgrade VELO, the sensor has small inter-electrode distance. This distance is indeed directly proportional to the charge collection time, however it is constrained by physical size and pitch of the pixel and pixel capacitance. The pixel studied in the TimeSpot has a $n - p$ doping profile, which guarantees an efficient electron collection at increasing radiation damage. The dimensions of the trenches applied are chosen in order to balance to opposite effects: on one side, a long trench minimizes the weak field area between trenches, on the other small distance between trenches increases pixel capacitance, causing a worsening in time resolution.
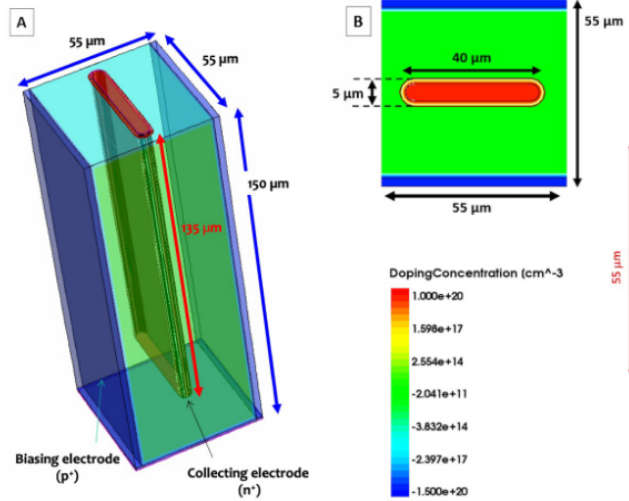
Figure 4.17: Geometry of the pixel of TimeSpot sensor. In figure the trench is deep 135 µm because of the pixel simulated in TimeSpot studies is 150 µm thick [1].

The implementation of this kind of pixel in the fast simulation required few adjustments in its dimension. The optimized trench size for the purposed studies is $5 \times 40\ \mathrm{µm}^2$ in the traversal plane and 180 µm thick. The geometry of a pixel in the fast simulation considers these trench regions as non-active components and the deposited charge is evaluated only in the effective volume. Even the charge threshold to activate a pixel is re-scaled considering only effective area of the pixel. From TimeSpot studies, a $\sigma = 1000$ electrons is predicted for the noise distribution, which means that the threshold has to be set at $3000 - 5000$ electrons, corresponding to 3 and 5 σ respectively. A schematic representation of the trench pixel is reported in Figure 4.17.

4.3.3 Channels in the Upgrade-II VELO

As described in Section 4.2, the total number of pixels contained inside the VELO is 40894464, divided in 52 different modules, each with 4 sensors formed by 3 chips (see Figures 4.1 and 3.3b). Each pixel is identified by a certain *channelID*, which contains inside all the information relative to its position inside the detector.

56

Figure 4.18: Schematic representation of the division of bits necessary to identify a channel

To identify uniquely a channel, a 26-bits word, with the scheme in Figure 4.18 is sufficient. Indeed, 16 bits are necessary to represent the pixel address inside a chip, which is composed by 256×256 pixels, 2 bits are used for the identification of the chip inside a sensor and 8 bits are dedicated to identify the sensor, that ranges from 0 to 207.

Super-Pixel concept

The super-pixel is a group of neighbouring pixels and forms a first, basic, cluster idea (see Section 4.2.3). More in detail, it is composed by eight pixels, enumerated from 0 to 7 as represented in Figure 4.19. The idea of super-pixel arises from the necessity of efficiently read-out the active pixels from the chips in the sensor[36]. It is found the best pixel format to be managed, especially if one considers to exploit FPGAs in the future. Any other format would require more bit and result in a higher data rate. Another advantage provided by the super-pixel is the availability of a first, initial clustering option at the read-out level.
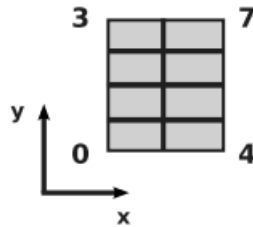
Figure 4.19: Structure of SP. The coordinates are local with respect to a sensor.

In the simulation, the time associated to a super-pixel is set as the time of the first active pixel inside that specific super-pixel.

So, for assumption

$$t_{SP} = t_{px_{first}}. \tag{4.9}$$

This is an approximation which rely on the consideration that most of the SP are composed by pixels activated by the same particle and so share the same time within the technical resolution of the sensor. In order to validate this assumption, a test using the Root Mean Square Error method is performed. The RMSE associated to the time of the super-pixel is evaluated as

$$\text{RMSE} = \sqrt{\frac{\sum_{i=0}^{N_{active}} (t_i - \hat{t})^2}{N_{active}}}, \tag{4.10}$$

where the sum is conduced over all the active pixels inside a super-pixel and $\hat{t}$ is the estimation of the time of the super-pixel, so for the assumption $t_{px_{first}}$. The resulting number of super-pixels which have a non-zero error on time assumption is negligible with respect to the total number of hit super-pixels ($\mathcal{O}(0.3\%)$).

4.3.4   Raw Bank

The raw bank is a format of data saving and handling particularly useful for the reconstruction software and that, at the same time, meets the requirements imposed by the DAQ hardware. Furthermore, the raw bank format offers a flexible and robust solution for data structures required by FPGA, where no particular data order is required and a larger rate of data is processed. For the clustering point of view, raw bank has to reproduce data structure used, as discussed in Section 4.2.3. Thus, a raw bank, with the format represented in Figure 4.20, is constructed for each sensor. To have a complete description of each super-pixel which forms a sensor, the associated word is composed by 36 bits. The number of words inside a raw bank corresponds to the number of active super-pixels in the sensor which is associated to that particular raw bank.
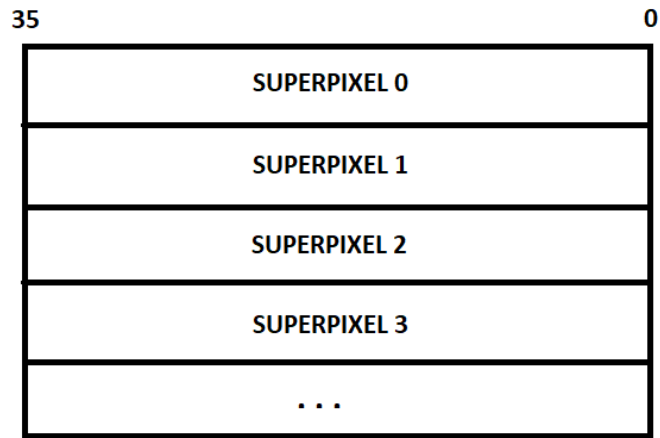
Figure 4.20: Format of a raw bank associated to a given sensor. It contains all the active super-pixels information.

A raw bank is created for each of the 208 sensors. Figure 4.21 shows the structure of the 36-bit word which characterize the super-pixel. Information contained are divided in the following way:

- the least significant byte contains the information relative to the active (1) and non-active (0) pixels inside a super-pixel;

- 15-bits are dedicated to represent row and column of the super-pixel in the sensor geometry;

- other 12-bits encodes time information, as discussed in 4.3.4;

- the most significant bit, called *hint*, indicates if the super-pixel is isolated. It assumes value 0 when there are no neighbouring active super-pixels in the sensor, 1 otherwise.

1	12	8	7	8

Figure 4.21: Schematic view of a word associated to a super-pixel, with the correspondent division in bits of the relative information

The 15 bits required to describe the address of a sensor depend on the position of the sensor itself and on the local orientation of the super-pixel: for sensors which develops horizontally, the super-pixel is disposed with two pixels on the vertical axis and four pixels on the horizontal axis; vice versa for the sensors which develops vertically. In this way, the first kind of sensors have 192×128 super-pixels, so 7-bits are sufficient to represent the rows, while for the second one 128×192 and 7-bits contain super-pixels columns. Further details relative to how fast simulation stores information in raw banks are reported in Section 4.3.5.

Encode time information

In the fast simulation the number of bits required to save the time associated to a super-pixel is 10, since almost all the super-pixels are characterized by a time within 3 ns, as it is possible to see in Figure 4.22a and 4.22b. In the first case, the times associated to the super-pixels are simulated for around 1000 events considering all the particles, while in the second case only VELO particles are considered. If one considers a time resolution of 6 ps, which can be considered as sufficiently fine, the largest times are ~ 500 times it. Hence, 10 bits, that represent $2^{10} = 1024$ values, results even abundant.

Unfortunately, the read-out operation by the detector cannot happen in a time window of few nanoseconds. If we suppose a more reasonable clock cycle of 25 ns for taking data, 12 bits are required for the purpose.
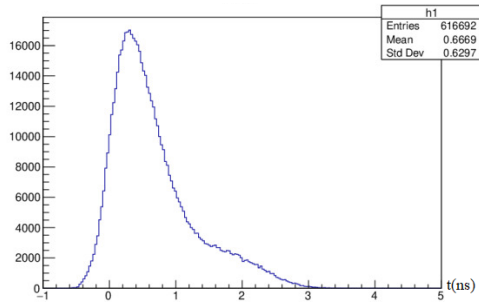
(a) Distribution of the simulated time for all particles.
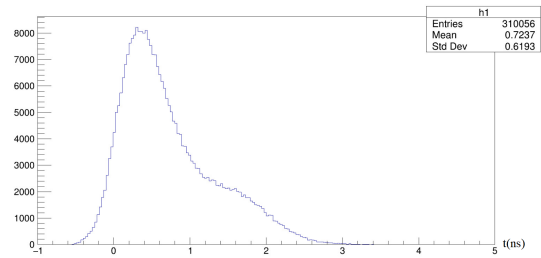
(b) Distribution of the simulated time only for VELO particles.

Figure 4.22: Time associated to a super-pixel with respect to a given origin time of the event.

4.3.5 Implementation of the raw bank on the fast simulation

The program which simulates the raw bank structure in the fast simulation is basically divided into two classes. Initially, the collision simulation parameters are read from an input file and they are stored into a `vector<vector<mcPart>>`, where `mcPart` is the `struct` that defines particles properties. Since there are more than one particles for each event, `vector<mcPart>` collects all particles properties of an event. The outer `vector` runs on all the events contained in the input file.

`VPupgradeI` class contains properties and methods which defines the Upgraded VELO geometry. Here a `map` associates each super-pixel ID to all information related to it, such as the pixel inside which has been crossed and the particles which crossed it. The information are stored in the following structure:

<div align="center">

`vector<map<int,vector<tuple>>>`

</div>

As said, the `map` associates the channel of the super-pixel (`int`) to a set of information (`vector<tuple>`), which characterize it.
Each `tuple` contains:

- ID of the active pixel inside the super-pixel (labelled from 0 to 7);

<div align="center">61</div>

- the particle key, which identify a particle in the input file;

- the released energy for each particle, evaluated as explained in 4.2.1;

- the ID of the layer (from 0 to 51);

- the active pixel position coordinates;

- the time associated to the active pixel, as reported in 4.3.1;

`RawBank` class contains two main methods (`bitmap` and `vpResponse`). The first one associates to each super-pixel ID the word with format presented in 4.3.4. In order to do that the following structure is adopted:

$$\texttt{vector<map<int,string>}.$$

The `map` associates each super-pixel ID (`int`) with the word (`string`), while the outer `vector` runs on all the events. Moreover, `bitmap` method creates for each event a text file (*PixelPosition.txt*), where, for all the active pixels, the following quantities are saved: ID of the layer, coordinates of the pixel and time associated. Even other useful information are implemented and can be displayed, such as the particle frequency for each pixel, the occupancy of the detector and the amount of charge released inside. `vpResponse` method creates a directory for each sensor in the VELO and populates each directory with a file "vpResponse.txt" which represent the raw bank.

4.3.6   Clustering with the raw banks

The input of the clustering algorithm corresponds to the output of the raw banks, in the format presented in Section 4.3.4. The clustering algorithm works decoding the raw bank and creating a bitmap image for each sensor. The advantage of using the super-pixel is evident: it corresponds to a first possible cluster. Moreover, if a super-pixel has a non-active *hint* bit, it means that it is isolated and the only possible clusters are searched inside the super-pixel itself. Because of structure of the super-pixel, the maximum number of clusters that an isolated super-pixel can contains is two. If it is

possible to produce a flag for isolated super-pixels, the clustering reduces to a simple pattern look-up for about half of the clusters (see Section 4.4.1).

Even the time information could be helpful for the clustering algorithm: in order to be added to a cluster, an active pixel has to be near in "time", as well as in space, to other active pixels. If a neighbouring pixel has an associated time enough different with respect to others, it probably means that it was activated by the passage of another particle and it will not be added to the cluster. However, these cases will be extremely rare.

In the Upgrade-I of the LHCb, the entire detector requires an High Level Trigger read-out system, described in Section 2.2.3, that deals with 40 MHz data processing rate. Moreover the future Upgrade-II scenario will be even more challenging: LHCb will operate without any hardware trigger, in order to fully read-out every event by all detectors and reconstruct with *offline quality*, before any software trigger decision [57]. Since the computing resources at disposition are limited, it is necessary to exploit computing devices which are intrinsically parallel, like FPGA and GPU, to address low-level time-consuming tasks at early DAQ stage. In particular, FPGA are well-suited because of their high efficiency in parallel, simple and repetitive processes like clustering algorithms.

The FPGA-friendly clustering algorithm is based on the one described in Section 4.4.1 [41], already implemented at LHCb. Unfortunately, this algorithm has yet to be tested with the addition of time information. More technical descriptions on how the clustering is implemented in the fast simulation are reported in Section 4.3.7.

4.3.7 Implementation of clustering algorithm in the fast simulation

The implementation of the clustering algorithm, based on the one described in Section 4.4.1, in the fast simulation is highly parallelized and particularly suitable for implementation in FPGA. It is structured into two different classes. The directory associated to each sensor, where the method `vpResponse` described in Appendix 4.3.5 writes the raw bank, is open and from the directory name it is evaluated the *sensorID*. From this identification number it is possible to retrieve important information associated to the sensor, like the module to which it belongs to and the *z*-coordinate associated. The file

"vpResponse.txt" is read and each line representing a super-pixel word of the raw bank is analyzed.

If the first bit is 0, the `clustering` method of `VPCluster` class is initialized and finds the clusters for the isolated super-pixel. This method evaluates the global coordinates (x, y, z, t) of the active super-pixel from the 15-bits which identify its address and from the 12-bits which encode time information (see Figure 4.20). Then, it evaluates local and global coordinates (x, y) of the clusters. The maximum number of clusters inside a super-pixel is two because of super-pixel structure. The individuation of clusters inside the super-pixel is obtained, as explained in Section 4.4.1, through the comparison with 8-bits word that characterize super-pixel configuration. In this way, it is possible to perform clustering for the most part of active pixels only at the first clock cycle.

The global positions, retrieved through the method `spCoordinates`, of neighbouring active pixels and the IDs (from 0 to 7) of the active pixels inside are, instead, saved into the following structure:

$$\texttt{vector<pair<hitCoord,vector<int>>>,}$$

called `VeloCache`. The `hitCoord` is a *struct* which contains super-pixel coordinates (x, y, z, t), the `vector<int>` has the active pixel channels associated to the given super-pixel and the outer `vector` stores all the neighbouring super-pixels of the raw bank.

`ClusterAlgorithm` class aims at providing the hits global coordinates also for the super-pixels stored in the `VeloCache`.
More in detail, `makeMatrix` method finds any possible link between super-pixels in two clock cycle. As described in Section 4.4.1, the first super-pixel initializes the *reading matrix* and, at the first clock thick, all the neighbouring active super-pixels in 8-directions are searched and added to the matrix. At the second thick, eventual remaining active super-pixels are added, in order to complete the structure illustrated in Figure 4.24. In order to optimize the algorithm, every super-pixel which is added to the matrix is simultaneously erased from the cache.
`makeMatrix` returns a `vector<vector<hitCoord>>`, where the outer vector refers to all possible matrices instantiated in the raw bank, while the inner one contains global coordinates of the active pixels inside each matrix.

Once the *reading matrices* are formed, they're analyzed by the `makeCluster` method, which calls in its body three other functions:

- `searchSeed`, that runs over all the active pixels stored for the matrix, searching if any match with given patterns. The current implementation defines a seed pixels as an active pixel surrounded by five non-active pixels. Further and more detailed analysis has to been provided in order to choose the most efficient definition for the seed pixel.

- `buildCandidate`, which creates a 3×3 matrix of pixels starting from the seed pixels. A 3-bits flag is associated to the cluster built and it allows to understand if the candidate is isolated or if it has any neighbouring active pixels. In the second case, the candidate is modified in its dimension and a new flag is evaluated.

- `makeClusterFromCandidate`. Once the candidate is read, all the clusters inside are formed in a similar way to the one adopted in `clustering` method for the isolated super-pixel case.

4.4   Future prospects

The future steps of the project aims to testing the fast simulation developed for the Upgrade-II, applying the clustering algorithm described in Section 4.2.3 on FPGAs. The results have to be compared with the ones obtained with the CPUs and the tracking performances are going to be evaluated using the FPGA-based tracking algorithm, developed by TimeSpot[42] described in Section 4.4.2. Further analysis relative to the reconstruction of primary vertices can be performed and used to evaluate the impact of the addition of the time information in the Upgrade-II scenario.

4.4.1   FPGA-friendly clustering algorithm

The algorithm described has already been developed at LHCb for the Upgrade-I [58] and it is implemented in similar way in the fast simulation, as described in Section 4.3.7, with

the addition of time information. The 4D-clustering based on raw banks is suitable to be tested with FPGA, because of his highly parallelizable structure. The performances discussed in the following refers to the 3D-clustering algorithm for the Upgrade-I [41].

Since most of the clusters contain few pixels (see Figure 4.9), it is reasonable to suppose that a significant fraction of clusters is contained in a super-pixel. If a super-pixel has non-active *hint* bit, i.e. has no neighbouring active super-pixel, it is defined *isolated* and it is possible to represent with a flag of 8-bits its configuration, as shown in Figure 4.23. Reconstruct this kind of cluster is extremely fast and requires very small amount of FPGA resources. Associate a flag from 0 to 255 (since there are 256 possible combination of active pixel inside a super-pixel, as shown in Figure 4.23) for isolated super-pixels will accelerate the clustering mechanism to be almost twice as fast. This one-clock cycle individuation of clusters inside isolated super-pixels is based on the memorization on the FPGA of a 8-bits look up table, which represents all possible different configurations of the active pixels. After the consultation, the algorithm return local x and y coordinates with respect to super-pixel position.
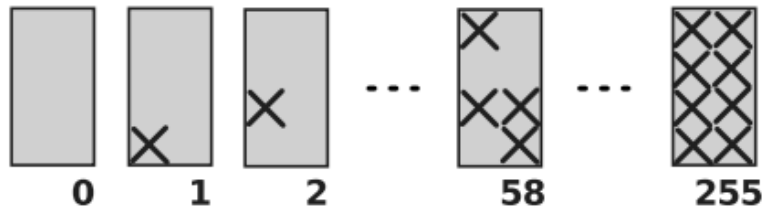
Figure 4.23: All possible combination of active pixels inside a super-pixel. It is possible to store in 8-bits information relative to active pixels inside the isolated super-pixel.

Finding clusters for super-pixels with active *hint* pixel is a bit more complex. The algorithm is based on the RETINA architecture [59], a massively parallel architecture for fast track finding, very suitable for implementation in FPGA devices. According to this algorithm, an elementary block, called *cellular processor* performs operation over the input data, stores the results and compares them with neighbouring cellular processors. The cellular processor stores the status of a single pixel and, once the status of neighbouring cellular processors has been read, it may find a cluster [60]. Because of the large number of pixels, of which only a small fraction is activated ($< 2\%$) inside the

VELO, the algorithm is modified, creating chain of small matrix with dimension 5×3 super-pixels, corresponding to 10×12 pixels. Until they're not initialized, these matrices don't represent any particular region of the VELO: once an active super-pixel is found, it initializes a matrix and it is placed at the center of the matrix, evaluating its coordinates and the coordinates of neighbouring super-pixels. Thus, other active super-pixel inputs with correspondent coordinates are searched and, in case of match, the pixels status is used to fill the right position in the matrix, otherwise the super-pixels are passed on to the next matrix in the chain (see Figure 4.24).
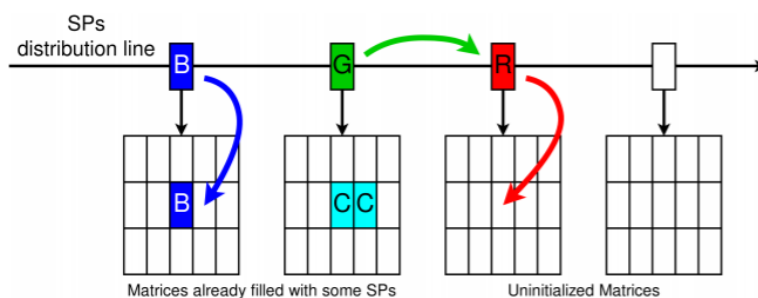
Figure 4.24: Schematic representation of super-pixels which flow along a matrix chain. At every clock cycle, all matrices move the super-pixel input. The (blue) first super-pixel fills the matrix, while the (green) second super-pixel looks for its coordinates correspond to the neighbours of the blue ones. If it is an active neighbouring super-pixel, it is added to the matrix, otherwise it (red) initializes a new matrix.
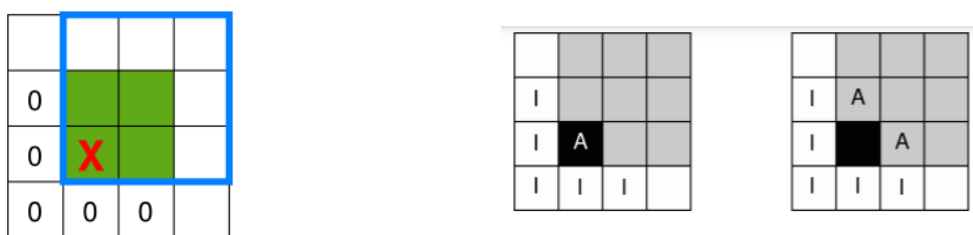
(a) Position of the seed pixel (red cross) for a given pattern

(b) Patterns associate to clusters. A and I indicates active and inactive pixels.

Figure 4.25: Seed pixel for clustering algorithm

After processing all input super-pixels, each cellular processor operates in parallel for

every existing matrix, checking for neighbouring pixels. The fundamental element in the search for clusters is the *seed*. It is a pixel that satisfies particular patterns, like in the example presented in Figure 4.25a, where the seed pixel is an active one surrounded by inactive pixels in the lower triangular region. Other examples of pattern for the recognition of the seed pixel that can be used are illustrated in Figure 4.25b and discussed more in detail in Section 4.3.7.

For every seed pixel a 3×3 cluster candidate with given topology, like the one presented in Figure 4.26, is build. To this candidate is associated a 3-bit word, called *flag* (see Table 4.4), that represents the status of the candidate and it is helpful to continue the clustering algorithm.

Figure 4.26: On the left: To the build candidate is associated a flag which represents its condition, starting from the information that one has about the seed pixels and the neighbouring pixels. On the right: Cluster topology starting from the seed pixel represented in Figure 4.25a.

Meaning	Flag
Isolated	101
Overflow	100
Self-contained and Edge	011
Self-contained and not Edge	010
not Self-contained and Edge	101
not Self-contained and not Edge	101

Table 4.4: Meaning of the 3-bits cluster flags

A cluster is flagged as:

- *Isolated*, if it is reconstructed from an isolated super-pixel;

- *Overflow*, if no enough matrices are instantiated for the active super-pixels, in this case the candidate is treated as isolated;

- *Self-contained*, if all the surrounding pixels are inactive;

- *Edge*, if the candidate has at least one side on the edge of the matrix
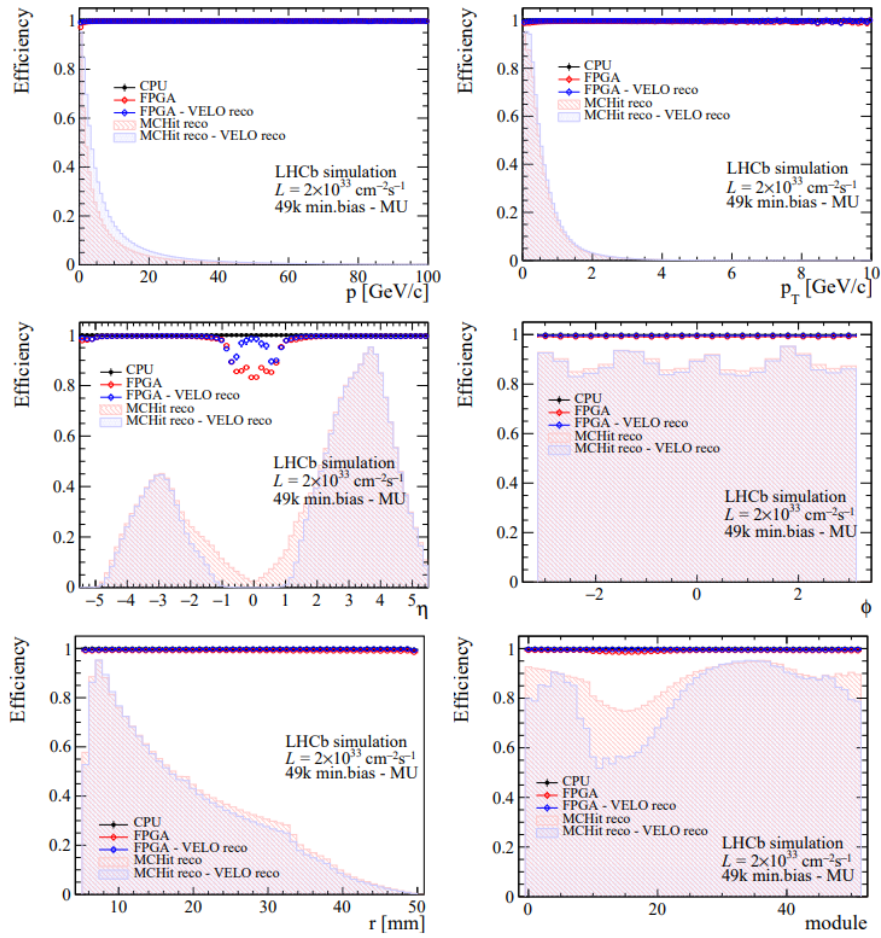
Figure 4.27: Cluster efficiencies for FPGA and CPU-based algorithm as function of various kinematic variables.

69

Clustering efficiency

The efficiency of the FPGA-based clustering method, based on LHCb programs `Rec` and `Moore`, is defined as:

$$\epsilon = \frac{N_{MC_{linked}}}{N_{MC}} \tag{4.11}$$

where the numerator in 4.11 represents the number of MC hits with linked reconstructed cluster, while N_{MC} is the total number of reconstructible MC hits. The efficiency of CPU-based algorithm is unitary by definition.

A comparison between clustering efficiencies for both FPGA (red-blue) and CPU (black) is represented in terms of radius and module number of the hit and of η, ϕ, p and p_T of the corresponding tracks in Figure 4.27. Generally the FPGA cluster finder is a bit less efficient with respect to the CPU algorithm, with difference of inefficiency between the two of $\sim 0.87\%$ on average. This inefficiency rate difference between FPGA and CPU reduces to $\sim 0.40\%$ when one considers only VELO particles.
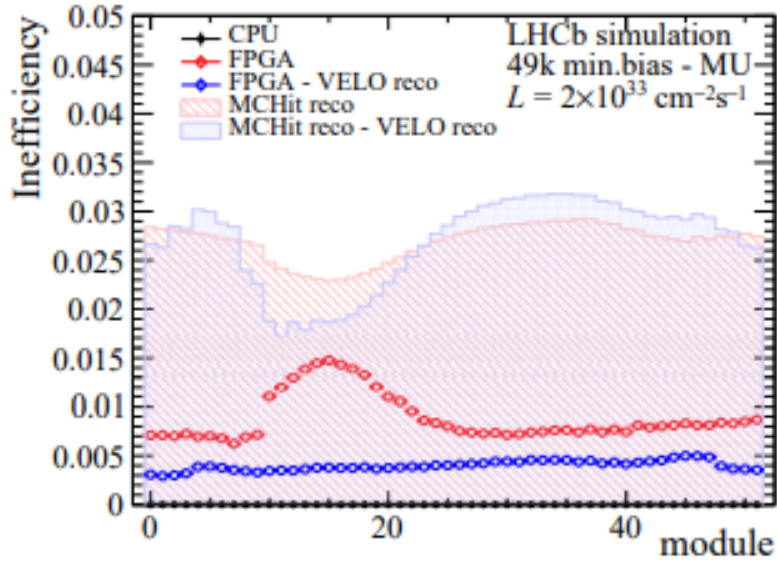
Figure 4.28: Cluster inefficiency for FPGA and CPU-based algorithm in terms of the module number. Near module 15, near the interaction point where the occupancy is larger, the difference between the two inefficiencies is larger

Differently from the CPU algorithm, the FPGA cluster efficiency deteriorates for large clusters and near the nominal interaction point (module 15), where the inefficiency difference reaches 3% (see Figure 4.28). The main source of inefficiences are going to be discussed more in details shortly. In general, the overall results obtained on FPGA-based algorithm show good performances with respect to the CPU ones and this FPGA clustering architecture could be a valid prospect to be applied for the Upgrade-II, in particular considering the advantages relative to gains of time. For this reason, the fast simulation implements raw-bank clustering based on this algorithm and eventual modifications will be applied and studied in future.

Inefficiency sources

Of course some inefficiencies in the clustering algorithm arise if one considers only pixels inside a matrix: in fact, pixels which are on the edges of the matrix cannot start a cluster candidate as seed and clusters which starts from an edge pixel are not reconstructed.
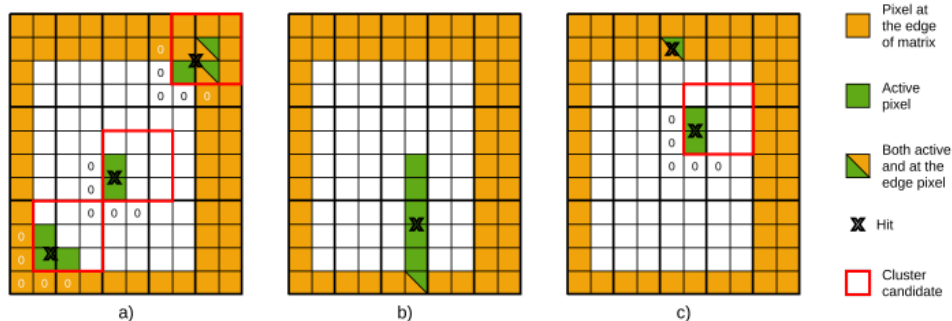
Figure 4.29: Examples of cluster candidates (red squares) and inefficiencies sources

Figure 4.29 shows some examples of clusters that can and cannot be reconstructed. In (a) all the clusters can be reconstructed because none of them starts from the edge of the matrix.(b) shows an example of non-reconstructed clusters, as in (c), because of impossibility to build the cluster candidate.

The FPGA cluster inefficiency depends on the VELO occupancy:

- Larger is the number of active super-pixels present in the VELO, larger is the probability that the cluster candidate starts from an edge of the matrix;

- Larger is the number of active super-pixels, lower is the probability to have isolated super-pixels and so cluster dimensions increase

- Larger is the number of active super-pixels, higher is the probability that the number of matrices initialized are not sufficient to represent the entire bitmap, so the number of overflow super-pixels will grow.

A plot of the cluster inefficiency considering all particles and only VELO particles is represented in Figure 4.30 in terms of the occupancy of the VELO. The CPU inefficiency is always, for definition, zero.
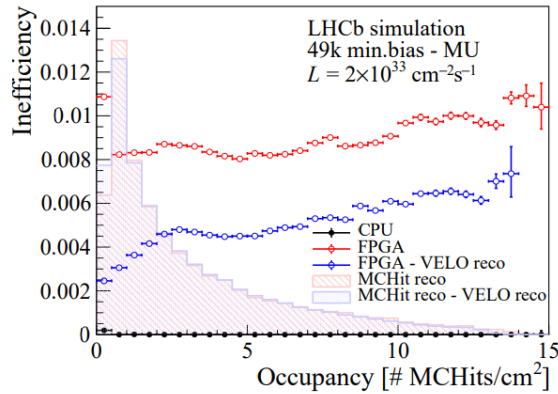
Figure 4.30: FPGA clustering inefficiency in terms of the occupation of the VELO. Increasing the number of active pixels inside the detector leads to a lower efficiency for the FPGA-based clustering algorithm.

More detailed studies have to be conduced in order to reduce the inefficiency sources without losing the rapidity of execution which characterize FPGA clustering algorithm, varying patterns used for the seed pixel research or trying to implement a more efficient algorithm.

72

4.4.2 4D tracking with FPGA

The algorithm here proposed is a 4D real-time track finding algorithm which aims to be considered in the prospect of a future VELO Upgrade-II. This work is within the TimeSpot project[42] and has already been tested in simulation and has being implemented with FPGAs. However, its development is still ongoing and the coordination with the group of simulation is mandatory for its success. It is based on the identification of couples of hits in adjacent sensors, called *stubs*. These stubs are characterized by three track parameters and the stub coordinates.

Since a track is defined by $(x_+, x_-, y_+, y_-, t_0)$, where $+$ and $-$ refer to the sign of the z coordinate, the track parameters are:

$$\begin{aligned}
x(z) &= x_0 + m_x z \\
y(z) &= y_0 + m_y z \\
t(z) &= t_0 + \frac{\sqrt{m_x^2 + m_y^2}}{c} z
\end{aligned} \tag{4.12}$$

where x_0, y_0 and t_0 are the coordinates of particle at $z = 0$ and m_x and m_y are the track slopes.

The stub is formed by all the combinations of hits that is compatible with reconstructible track. Assuming (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2) are the coordinates of the first and second hit of a stub candidate, respectively; the stub coordinates are evaluated as:

$$\begin{aligned}
x_{-,stub} &= \frac{x_1 z_- - x_2 z_-}{z_1 - z_2} \\
x_{+,stub} &= \frac{x_1(z_+ - z_2) - x_2(z_+ - z_1)}{z_1 - z_2} \\
y_{-,stub} &= \frac{y_1 z_- - y_2 z_-}{z_1 - z_2} \\
y_{+,stub} &= \frac{y_1(z_+ - z_2) - y_2(z_+ - z_1)}{z_1 - z_2} \\
t_{0,stub} &= \frac{t_1 + t_2}{2} - \frac{z_1 + z_2}{2c\sqrt{1 + \left(\frac{x_-}{z_-}\right)^2 + \left(\frac{y_-}{z_-}\right)^2}}
\end{aligned} \tag{4.13}$$

with velocity of the particle estimated as $\frac{|\vec{x}_1 - \vec{x}_2|}{t_1 - t_2}$.

Then, the track reconstruction consists on the best fit of all the pattern candidates found. A pattern recognition is based on the projection to a reference plane of multiple stubs associated to a track candidate. The resolution on the track parameters are evaluated fitting the residual distribution of the reconstructed minus the generated parameter with the sum of two Gaussian functions. The value of the resolution is defined as the root mean square of the fitting function. The resulting resolutions of track parameters for Upgrade-II simulation are reported in Table 4.5 and in Figure 4.31 both in the cases of using time information with resolution of 30 ps and without using it.

Resolution	Timing cuts	No Timing cuts
$\sigma_{x_-}(\mu m)$	100.0	73.0
$\sigma_{y_-}(\mu m)$	99.2	72.4
$\sigma_{x_+}(\mu m)$	37.8	35.0
$\sigma_{y_+}(\mu m)$	38.2	35.2
$\sigma_{t_0}(ps)$	45.3	29.1

Table 4.5: Comparison between resolution for track parameters applying or not the timing cuts

The reconstruction efficiency ϵ_{rec} is defined as the ratio between the number of reconstructed tracks and the number of generated tracks that can be reconstructed using the 4D tracking algorithm. In both cases, with and without the time information of the hit, the efficiency is $\epsilon_{rec} = 98.5\%$. The rate of ghost tracks is instead reduced when using the time information. The reconstruction purity P_{rec}, defined as the ratio between the number of well reconstructed tracks and the total number of reconstructed tracks, improves when using the time information, going from $P_{rec} = 60\%$ to $P_{rec} = 82\%$.

The results on tracking performances obtained and discussed rely on an approximated simulated input. It could be interesting to observe the behaviour of the tracking algorithm implemented by TimeSpot using data collected with the fast simulation in the current raw bank format.
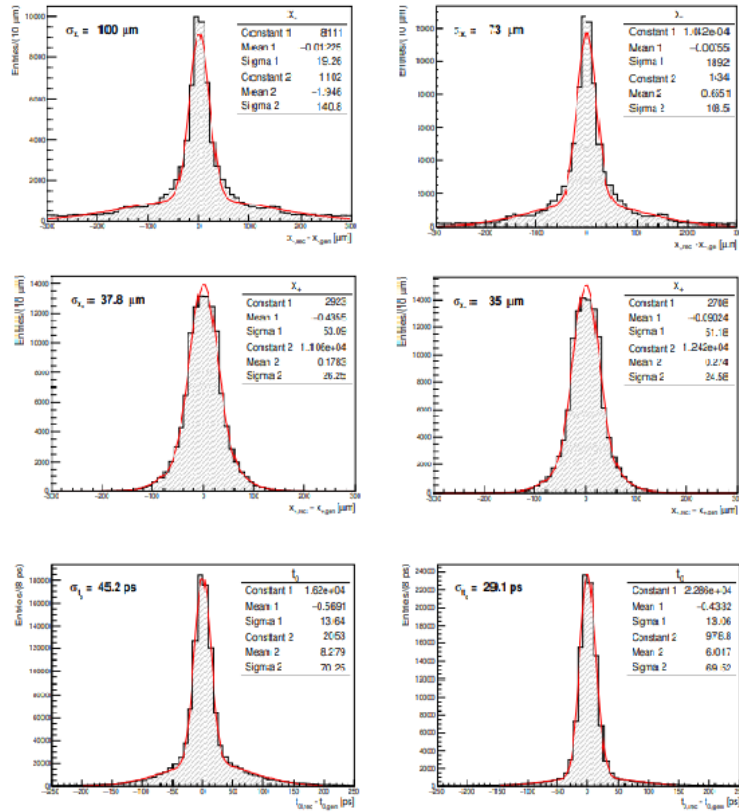
Figure 4.31: Residual distribution minus x_- and x_+ (first for pictures) and t_0 (last two) track parameters for 1200 generated tracks. Images on the left are associated to measurements with time information included, while on the right time is not considered.

4.4.3 Primary vertex reconstruction

The mechanism which allows to reconstruct primary vertices is based on finding the point cf closest approach of a track to the beam axis. The z-coordinate of this point, known as z_{BL}, represents a good approximation for the z position of the vertex.

Usually when primary vertices overlap in space can be distinguished by a time-stamp on tracks enough small (~ 20 ps). The time t_{BL} introduced represents the time measurements of the track at z_{BL} and it is a good approximation for the time of pp interaction.

75

The distribution of t_{BL} and z_{BL} for selected tracks is shown in Figure 4.32.

The 4D primary vertices algorithm is based on the following points:

- find the local maxima in the $t_{BL} - z_{BL}$ plane;

- associate tracks to this maxima;

- iteratively fit the vertices, in order to associates a primary vertex to each track.
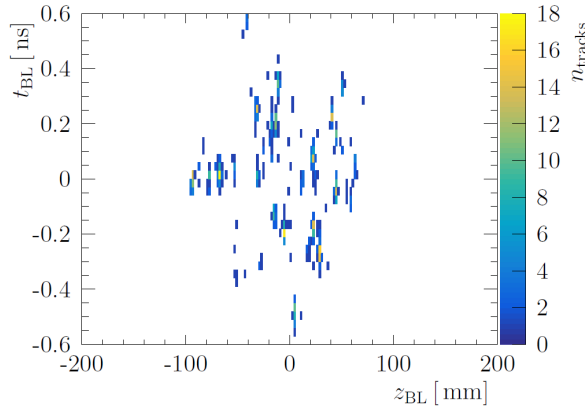
Figure 4.32: t_{BL} vs z_{BL} for selected tracks of Upgrade II. Selected tracks are those which are enough close to the beam (< 0.5 mm) and have uncertainty on z_{BL} enough small (< 0.75 mm)

PV reconstruction efficiency

The efficiency of primary vertices reconstruction is defined as

$$\epsilon = \frac{N_{reconstructed}}{N_{reconstructible}}, \tag{4.14}$$

where the numerator represents the number of primary vertices well-reconstructed and the denominator is the number of vertices with at least four charged tracks in the VELO acceptance. A vertex is defined as *well-reconstructed* if it has at least 75% purity[3].

---

[3]Purity is the fraction of tracks that come from the same true primary vertex

76

Another quantity which represent the performance of primary vertices reconstruction is the *split*, defined as the primary vertices which are actually true vertices that have at least one well reconstructed vertex, but that contains less than 75% of the tracks from this true primary vertex that are associated to it. Table 4.6 represents the different performances of full simulation for Upgrade-I and Upgrade-II conditions. It is possible to notice the deterioration of the performance parameters for primary vertex reconstruction.

Upgrade	Algorithm	Purity[%]	ϵ[%]	$\epsilon_{n>25}$[%]	Split[%]
U-I	`TrackBeamLineVertexFinder`	93.1	82.1	93.5	1.2
	`PatPV4D`	99.2	77.1	90.7	3.7
U-II	`TrackBeamLineVertexFinder`	69.7	25.4	37.6	4.5
	`PatPV4D`	97.2	71.3	86.9	3.0

Table 4.6: Comparison of the performance of full simulation vertexing algorithms for Upgrade-I and Upgrade-II [39].

A study relative to the primary vertex reconstruction with the inclusion of time in the fast simulation has yet to be performed. The primary vertex reconstruction performances obtained could represent an additional parameter to test the impact of time information in the Upgrade-II harsh scenario.

Conclusions

In this thesis, a study regarding a possible improvement of the Vertex Locator for the LHCb experiment in the context of High-Luminosity LHC is presented. Subsequent to the description of the High Luminosity phase at LHC and of the related physics, the geometry of VeloPixel and a future vertex detector for the Upgrade-II phase have been reported. After this introduction, the reader has acquired the fundamental information to understand the (fast) simulation based on VELO Upgrade-I geometry and described in this thesis. A fast simulation is, by construction, more flexible for parametric studies with respect to a full approach, that, offers a good reference point for the validation. The first change with respect to the reproduced Upgrade-I conditions is the addition of time information of a track as seen by pixel sensors, that has been demonstrated with the work of this thesis, brings important improvements in the track reconstruction performances in the Upgrade-II scenario. A possible solution to achieve the goal of having time information in a future detector is proposed by the TimeSpot Collaboration. They are developing a promising silicon sensor, with a time resolution close to 30 ps. In the thesis work, different time resolution values, including the TimeSpot ones, are tested with positive results. The fast simulation proposes, in addition, a possible solution for the read-out, which includes time information in the so-called super-pixel, making clustering and reconstruction less time-consuming in terms of the number of clocks required. A raw bank output format is also implemented, useful for further studies, as real-time clustering and reconstruction on FPGAs hardware. The next steps of the project consist on the application of the clustering algorithm based on raw banks on the FPGA hardware. Thanks to this work, the fast simulation will be used in order to explore different detector geometries, such as different z spacing of the layers and angle of the latter with respect to the beamline, to develop a future VELO detector, capable of facing the extreme

conditions that will be present in the High-Luminosity era. In conclusions, the thesis work here presented shows, trough simulation studies, that the time information in a pixel sensor, with a resolution of about 30 ps, will be necessary to achieve similar performance like the VeloPixel will do in the coming years.

List of Figures

83

List of Tables

Bibliography

[1] L. Anderlini, M. Aresti, A. Bizzeti, M. Boscardin, A. Cardini, G.-F. Dalla Betta, M. Ferrero, G.Forcolin, M. Garau, A. Lai, A. Lampis, A. Loi, C. Lucarelli, R. Mendicino, R.Mulargia, M. Obertino, E. Robutti, S. Ronchin, M. Ruspa, S. Vecchi , "Intrinsic time resolution of 3D-trench silicon pixels for charged particle detection," *Journal of Instrumentation*, 2020. [Online]. Available: https://arxiv.org/pdf/2004.10881.pdf

[2] Brüning, Oliver Sim ; Collier, Paul ; Lebrun, P ; Myers, Stephen ; Ostojic, Ranko ; Poole, John ; Proudlock, Paul, "LHC Design Report," *CERN*, 2004.

[3] LHCb Collaboration, "LHCb Technical Design Report," *CERN*, 2005.

[4] Letizia Diamante, "Transforming LHCb: What's in store for the next two years?" 2019. [Online]. Available: https://home.cern/news/news/experiments/transforming-lhcb-whats-store-next-two-years

[5] A.Carbone, "Lecture Notes," *Università degli studi di Bologna*, 2019.

[6] A. Ceccucci, Z. Ligeti and Y. Sakai, "CKM Quark-Mixing Matrix," *Phys. Rev.*, 2018.

[7] N. Tuning, "Lectures Notes on CP violation," *Nikhef*, 2020.

[8] I.I.Bigi,A.I.Sanda, "CP violation," *CAMBRIDGE UNIVERSITY PRESS*, 2010.

[9] L. Wolfenstein, "Parametrization of the Kobayashi-Maskawa Matrix," *Phys.Rev.*, 1983.

[10] M. Artuso, G. Borissov, and A. Lenz, "CP violation in the B_s^0 system," *Rev.Mod.Phys.*, 2016.

[11] The LHCb Collaboration, "A precise measurement of the B_s^0 meson oscillation frequency," *CERN*, 2016.

[12] Karen Lingel,Tomasz Skwarnicki, James G. Smith, "Penguin decays of B mesons," *Annual Reviews of Nuclear and Particle Science*, 1998.

[13] J.Charles, "Current status of the Standard Model CKM fit and constraints on $\Delta F = 2$ New Physics," *Phys. Rev.*, 2015.

[14] The LHCb collaboration, "Precision measurement of CP violation in $B_s^0 \rightarrow J/\psi K^+ K^-$ decays," *CERN*, 2015.

[15] S. Faller, R. Fleischer, and T. Mannel, "Precision physics with $B_s^0 \rightarrow J/\psi\phi$ at the LHC: The quest for new physics," *Phys. Rev.*, 2009.

[16] The LHCb Collaboration, "Resonances and CP-violation in $\bar{B}_s^0$ and $B_s^0 \rightarrow J/\psi K^+ K^-$ decays in the mass region above the $\phi(1020)$," *CERN*, 2017.

[17] ——, "Measurement of the CP-violating phase ϕ_s in $B_s^0 \rightarrow J/\psi\pi^+\pi^-$," *CERN*, 2014.

[18] ALICE Collaboration, "ALICE Technical Design Report of the Transition Radiation Detector," *CERN*, 2001.

[19] ATLAS Collaboration, "ATLAS TDR," *CERN*, 2012.

[20] CMS Collaboration, "CMS Technical Design Report," *CERN*, 2006.

[21] Rama Calaga, "Crab cavities for High Luminosity LHC," *CERN*, 2019.

[22] Iva Raynova, "Crab cavities: colliding protons head-on," 2017. [Online]. Available: https://home.cern/news/news/accelerators/crab-cavities-colliding-protons-head

[23] Burkhard Schmidt, " The high-luminosity upgrade of the LHC: Physics and technology challenges for the accelerator and the experiments." *Journal of Physics: Conference Series*, 2016.

[24] Apollinari G., Bejar Alonso I., Bruning O., Fessia P., Lamont M., Rossi L., and Tavian L. , "High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report," *CERN*, 2017.

[25] The LHCb Collaboration, "The LHCb Detector at the LHC," *CERN*, 2008. [Online]. Available: https://cds.cern.ch/record/1129809/files/jinst8_08_s08005.pdf

[26] The LHCb collaboration, "LHCb VELO TDR: Vertex locator," *CERN*, 2001.

[27] "LHCb detector." [Online]. Available: https://lhcb-public.web.cern.ch/en/ Detector/Detector-en.html

[28] The LHCb Collaboration, "The LHCb Electronic Calorimeter," *CERN*.

[29] ——, "The LHCb Hadron Calorimeter," *CERN*.

[30] The LHCb collaboration, *CERN*, 2012.

[31] "Trigger system," 2020. [Online]. Available: https://lhcb-public.web.cern.ch/en/ Data%20Collection/Triggers2-en.html

[32] Olaf Steinkamp, "LHCb Upgrades," *CERN*, 2020. [Online]. Available: https: //cds.cern.ch/record/2686662/files/10.1088_1742-6596_1271_1_012010.pdf

[33] E. Metral, "Update of the HL-LHC operational scenarios for proton operation," *CERN*, 2018.

[34] The LHCb Collaboration, "LHCb VELO Upgrade Technical Design Report," *CERN*, 2013.

[35] ——, "Framework TDR for the LHCb upgrade," *CERN*, 2012.

[36] ——, "VP Simulation and Track Reconstruction," *CERN*, 2014. [Online]. Available: https://cds.cern.ch/record/1620453/files/LHCb-PUB-2013-018.pdf

[37] Adriano Lai, "TimeSpot Project - The Challenge." [Online]. Available: https://web.infn.it/timespot/index.php/the-timespot-project-menu

[38] Nicola Neri, "WP4: Fast-tracking algorithms and devices." [Online]. Available: https://web.infn.it/lhcb_milano/wp4-fast-tracking-algorithms-and-devices/

[39] The LHCb collaboration, "VELO upgrade II FTDR," *CERN*, 2020.

[40] A.Carbone, S.Maccolini, "A fast simulation validation towards performance studies for a Velo-U2," *INFN*, 2020.

[41] G. Bassi, L. Giambastiani, F. Lazzari, M. J. Morello, T. Pajero, G. Punzi, "Performance of FPGA-based clustering for the VELO upgrade," *CERN*, 2020.

[42] Marco Petruzzo, "A 4D real-time tracking device for the LHCb Upgrade II," *Università degli studi di Milano*, 2019.

[43] T. Sjostrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 Physics and Manual," *FERMILAB*, 2006. [Online]. Available: https://arxiv.org/pdf/hep-ph/0603175.pdf

[44] D. J. Lange, "The EvtGen particle decay simulation package," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 2001.

[45] P. Golonka, Z. Was, "PHOTOS Monte Carlo: A precision tool for QED corrections in Z and W decays," *The European Physical Journal C*, 2006.

[46] Geant4 collaboration, "Geant4 developments and applications," *IEEE Transactions on Nuclear Science*, 2006.

[47] I. Belyaev, "Handling of the generation of primary events in Gauss, the LHCb 1253 simulation framework," *Journal of Physics*, 2011. [Online]. Available: https://iopscience.iop.org/article/10.1088/1742-6596/331/3/032047/pdf

[48] "Gaudi." [Online]. Available: https://lhcb-comp.web.cern.ch/Frameworks/Gaudi/default.htm

[49] "Boole." [Online]. Available: https://lhcb-comp.web.cern.ch/Digitization/

[50] "Brunel." [Online]. Available: http://lhcb-comp.web.cern.ch/Reconstruction/

[51] W.R.Leo, "Techniques for Nuclear and Particle Physics Experiments," 1994.

[52] D.E. Groom, S.R. Klein , "Passage of Particles Through Matter," *Phys. Rev.*, 2019. [Online]. Available: https://pdg.lbl.gov/2019/reviews/rpp2018-rev-passage-particles-matter.pdf

[53] H. Bichsel, "Straggling in Thin Silicon Detectors," *Rev. Mod. Phys.*, 1988.

[54] H. Bichsel, "A method to improve tracking and particle identification in TPCs and silicon detectors."

[55] H. A. Bethe, "Molière's Theory of Multiple Scattering," *Phys. Rev.*, 1953.

[56] O. I. Gerald R.Lynch, "Approximations to multiple Coulomb scattering."

[57] Mark Williams, "The LHCb VELO Upgrade II," *PIXEL Conference 2018*, 2018.

[58] Federico Lazzari, Giovanni Bassi, Riccardo Cenci,Michael J Morello, Giovanni Punzi, "Real-time cluster finding for LHCb silicon pixel VELO detector using FPGA," *CERN*, 2019.

[59] Alessandro Marchioro, Alex Grillo, "Topical Workshop on Electronics for Particle Physics: Santa Cruz," *Santa Cruz Institute for Particle Physics*, 2017.

[60] W. Baldini, G. Bassi, A. Contu, M. Dorigo, R. Fantechi, F. Lazzari, M.J. Morello, G. Punzi, S. Stracka, G. Tuci, G. Vitali., "Accelerating Real Time Analysis in Run-3 with FPGAs."

Aknowledgements