UNIVERSITÀ DEGLI STUDI DI MILANO

FACOLTÀ DI SCIENZE E TECNOLOGIE

Corso di Laurea Magistrale in Fisica

A STUDY FOR THE MEASUREMENT OF THE Λ BARYON ELECTROMAGNETIC DIPOLE MOMENTS IN LHCb

Relatore: Prof. Nicola NERI
Correlatore: Dott.ssa Elisabetta SPADARO NORELLA

Tesi di Laurea di:
Alessandro DE GENNARO
Matricola 933289
Codice P.A.C.S.: 14.20.-c

Introduction

Electric and magnetic dipole moments of particles are sensitive to physics within and beyond the Standard Model. In this thesis, sensitivity studies for the measurement of the Lambda baryon electromagnetic dipole moments based on pseudo experiments will be performed. In addition, the possibility of a first measurement using data collected with the LHCb detector will be explored.

Contents

Chapter 1

Flavour physics and CP symmetry violation

This chapter explores the theoretical framework for the rest of the thesis. Section 1.1 provides a basic introduction to the Standard Model of Particle Physics and flavour physics in particular; Section 1.2 delves into the inner workings of discrete symmetries in quantum physics; Section 1.3 discusses the relevance of electromagnetic dipole moments of elementary particles as a test for CP violation and CPT symmetry; finally, Section 1.4 introduces the main physics motivation for this thesis, the study of dipole moments of the Λ^0 baryon and the proposed measurement technique.

1.1 The Standard Model of Particle Physics

Ever since Democritus' philosophy of atomism, one of the driving desires behind mankind's advancements in the fields of natural science has been to reduce reality to its basic components. While one can convincingly argue that we may never fully understand what has come to be known as the quantum world, the Standard Model of Particle Physics (Standard Model, or SM, for short) [1] is as close as physics has to offer to a comprehensive theory of the building blocks of matter and energy.

In addition to predicting a number of then-unknown particles discovered in later years, the Standard Model has shown remarkable consistency against high precision tests, especially in the better known electroweak sector [2]. Despite this, it would be a mistake to call it *complete*, even if only for the three fundamental forces it covers. Many experimental evidences, some of which will be discussed in the following pages, have already opened cracks in the model, and many more may emerge in the future; one of the recurring topics of this chapter will thus be the need for physics Beyond the Standard Model (BSM).
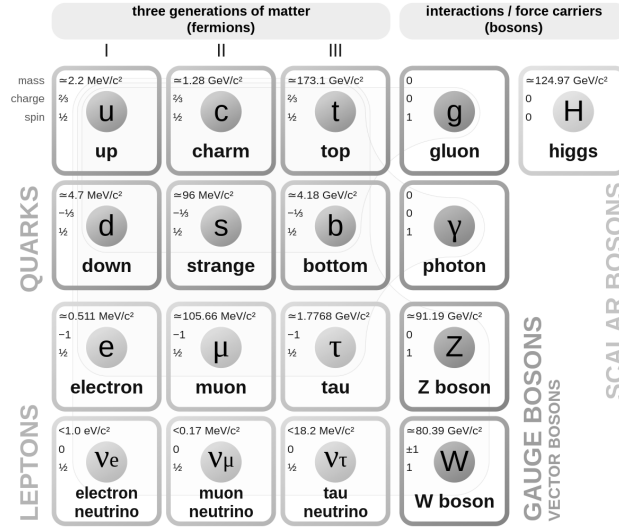
Figure 1.1: The seventeen currently known elementary particles of the Standard Model. Antiparticles are not depicted.

1.1.1 Elementary particles

Intuitively, a particle is said to be *elementary* when no substructure can be probed. A century of efforts in the fields of nuclear, quantum, and high energy physics has whittled down the spectrum of matter to just seventeen unique fundamental particles, colloquially known as the *particle zoo* and depicted in Figure 1.1.

Each particle is joined by an *antimatter particle* (*antiparticle* for short), a companion of opposite charge identified by the prefix *anti-*, e.g. antimuon for the muon; the only exception to this naming convention is the electron, whose antiparticle, for historical reasons, is known as positron. While often omitted for the sake of brevity, antiparticles are elementary particles in every respect, distinct from their partners (bar neutral gauge bosons and the Higgs boson, which are their own antiparticles) and related to them through the transformation of charge conjugation (see Section 1.2.2).

Leptons

Leptons are fermions (half-integer spin particles) not sensitive to the strong nuclear interaction. There are currently six *flavours* of leptons grouped in three generations: each generation comprises a *charged* lepton (electron, muon, tauon) and a *neutral* lepton (electron neutrino, muon neutrino, tauon neutrino).

All charged leptons have a charge of $-e$, where e is the elementary positive charge, and their masses range from $\approx 0.5\,\mathrm{MeV}$ for the electron to over $1.7\,\mathrm{GeV}$ for the tauon [3]. By contrast, as the names suggest, all neutrinos are electrically neutral and are assumed massless in the Standard Model[1]; this implies that their only meaningful interactions happen through the weak nuclear force, which grants them their characteristic evasiveness to most particle detectors.

Quarks

Much like leptons, quarks are also fermions existing in three generations. The main difference from the former category is that quarks, besides interacting through weak and electromagnetic forces, are also susceptible to the strong nuclear forces; this allows them to bind together in composite states known as *hadrons*, which are classified as *baryons* (states of three quarks) and *mesons* (states of one quark and one antiquark)[2].

Quarks can be classified as *up-type* (up, charm and top quarks) and *down-type* (down, strange and bottom quarks): up-type quarks have a fractionary charge of $+\frac{2}{3}e$, whereas down-type quarks have a charge of $-\frac{1}{3}e$. All quarks also have one of three *color* charges (red, green or blue), while antiquarks similarly have one of three *anti-color* charges (antired, antigreen or antiblue). A combination of all three colors/anti-colors or a combination of a color and its matching anticolor produces *colorless* particles, a property shared by all observed quark composite states.

Unlike leptons, quarks are impossible to observe directly: according to the phenomenon of *color confinement*, the energy of the interaction field between two color charges being pulled apart increases with their distance until it becomes high enough to create a quark-antiquark pair. This process of *fragmentation* develops many times over in such a way that the final observable state is entirely composed of colorless particles. For this reason, high energy physics experiments such as LHCb do not detect free quarks, instead observing cone-shaped streams of hadrons known as *hadronic jets*.

Gauge bosons and fundamental interactions

In quantum field theory, the interaction between two fields is implemented through the exchange of an intermediary particle known as *force carrier*. In the Standard Model all force carriers are vector (spin 1) bosons known as *gauge bosons*. The name is owed to the *gauge principle* used to introduce them: the

---

[1]The observation of flavour oscillation in solar neutrinos shows that neutrinos do in fact have non-zero, albeit very small, mass [4].

[2]As recently as 2003, evidence has surfaced for the existence of exotic hadrons composed of four (*tetraquarks*) [5] and five quarks (*pentaquarks*) [6].

localization of a global continuous symmetry group provides the free fermion Lagrangians with interaction terms with the proviso that one or more bosonic fields are introduced.

The gauge principle accounts for the implementation of three fundamental interactions along with their gauge bosons: the *strong nuclear force* with its massless gluon, responsible for the binding of both quarks inside baryons and nucleons inside atomic nuclei; the *electromagnetic force* mediated by the massless photon, the importance of which should be known from everyday life; and the *weak nuclear force* with two massive $W^{\pm}$ and Z bosons, the source of many subnuclear processes such as β radioactivity.

The latter two forces share a unified description in the Glashow-Weinberg-Salam theory as a single *electroweak interaction* and are introduced via localization of a $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ symmetry group, the first related to the conservation of weak isospin in left-handed chirality states and the second to the conservation of hypercharge. Quantum chromodynamics (QCD), the theory of the strong nuclear force, is based on a separate $\mathrm{SU}(3)_C$ symmetry acting on the three-dimensional space of color charges.

There are no gauge bosons nor gauge theories associated to the fourth known fundamental force, gravity. Since every attempt to reconcile the general theory of relativity with quantum mechanics has failed so far, gravity is presently excluded from the Standard Model. This doesn't affect SM predictions at the subatomic level on account of the remarkably low intensity of said force, over 30 orders of magnitude lower than the weak interaction.

The Higgs boson

The Higgs boson is one of the latest additions to the Standard Model, being proposed in 1964 [7] and observed in 2012 by the ATLAS [8] and CMS [9] collaborations. Its introduction solved perhaps the most insidious SM shortcoming at the time: gauge theories, which the model was built on, only worked under the assumption that all particles involved were massless, whereas the local invariance would fall apart (*gauge breaking*) when adding a free mass term.

The Higgs field accounts for mass generation of the weak bosons $W^{\pm}$ and Z via the Brout-Englert-Higgs mechanism resulting from the spontaneous electroweak symmetry breaking; elementary fermions also gain mass through a distinct, Yukawa-like interaction with the field.

1.1.2 Flavour physics

A reader unfamiliar with SM terminology may find amusing the use of the word *flavour* to refer to what have been so far presented as different kinds of particles altogether. However quirky, the lexical choice highlights a defining feature: flavour, much like the degree of sweetness in a recipe, can change [10].

As often happens in particle physics, the rules are somewhat easier for leptons. For a given generation $\ell = (e, \mu, \tau)$, one can define a lepton family number L_ℓ as the difference between the number of particles and antiparticles of said generation, charged leptons and neutrinos alike:

$$L_\ell := n(\ell^-) - n(\ell^+) + n(\nu_\ell) - n(\bar{\nu}_\ell). \tag{1.1}$$

For all three generations, L_ℓ is conserved in every interaction except neutrino oscillations.

Quarks are not as straightforward. A similarly defined quark flavour number, such as the so-called *topness* (or *truth*)

$$T := n(t) - n(\bar{t}), \tag{1.2}$$

is preserved through EM and strong interactions, but can change when the state undergoes a weak charged interaction, i.e. a weak interaction mediated by the charged gauge bosons $W^\pm$. Weak interactions for quarks can be accurately described if we assume that the weak eigenstates (d', s', b') of down-type quarks, i.e. the weak isospin doublet partners to up-type quarks, are related to the free mass eigenstates (d, s, b) through a rotation:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.3}$$

In this notation, the probability for a quark of flavour i to change into a quark of flavour j as a result of a weak charged interaction is proportional to $|V_{ij}|^2$.

The unitary rotation matrix is known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix V_{CKM}. The moduli of its components up to the third decimal place, according to the most recent estimates [3], are

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \approx \begin{pmatrix} 0.974 & 0.224 & 0.004 \\ 0.221 & 0.987 & 0.041 \\ 0.008 & 0.039 & 1.013 \end{pmatrix}. \tag{1.4}$$

A full definition of the CKM matrix requires four independent parameters. Particularly useful for the following sections is the standard parameterization

with three angles θ_{12}, θ_{23}, θ_{13}, expressing the mixing between different quark generations, and a complex phase δ_{13}. Defining $s_{ik} := \sin\theta_{ik}$ and $c_{ik} := \cos\theta_{ik}$, V_{CKM} can be written as

$$V_{\mathrm{CKM}} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \quad (1.5)$$

The phase δ_{13} is known as the CP-violating phase. To fully understand what it means and its role in particle physics, however, a digression into discrete symmetries is needed.

1.2   Discrete symmetries and CP violation

In quantum mechanics, a system described by a Hamiltonian $\hat{\mathcal{H}}$ is *symmetric* under a transformation $\hat{S}$ if the two operators commute, i.e.

$$\left[\hat{\mathcal{H}}, \hat{S}\right] = 0. \quad (1.6)$$

Symmetries are of great relevance in physics on account of Noether's theorem, which establishes a relationship between the *continuous* symmetry of a system and a corresponding conservation law; the emphasis is on the requirement of continuity, meaning the related transformation changes the system «in a smooth way», much like a rotation does. An example of this principle has already been presented earlier in this thesis: the three symmetry groups employed in SM gauge theories all imply the conservation of a specific charge, be it weak isospin, hypercharge, or color.

This section will instead delve into *discrete* symmetries [11], which do not share said «smoothness» property. The absence of a Noether-like theorem for this class of transformations does not detract from their importance in physics: as will be shown, the three symmetries we'll focus on have a remarkable influence on many fields of study.

1.2.1   Parity inversion

The *parity inversion* (or simply *parity*) transformation $\hat{P}$ flips the sign of the three spatial coordinates:

$$\hat{P} : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \to \begin{pmatrix} -x \\ -y \\ -z \end{pmatrix}. \quad (1.7)$$

Its action on a quantum $|\psi(\vec{x}, t)\rangle$ is therefore

$$\hat{P}|\psi(\vec{x}, t)\rangle = |\psi(-\vec{x}, t)\rangle. \tag{1.8}$$

For parity eigenstates (known as parity-defined states) a parity quantum number can be introduced; such states may have a P-parity eigenvalue $\eta_P = +1$ (parity-even states) or $\eta_P = -1$ (parity-odd states). Since by definition $\hat{P}^2 = \mathbb{1}$, where $\mathbb{1}$ is the identity operator, these are the only two allowed P-parity values.

A similarly dichotomous behaviour is observed on both scalar and vector quantities commonly used in classical physics, such as momenta and electromagnetic fields. Parity-even scalar quantities are called *true scalars* or just *scalars* (e.g. energy), whereas parity-odd ones are called *pseudoscalars* (e.g. helicity). Likewise, vector quantities are either *polar vectors* (e.g. angular momentum) or *axial vectors* (e.g. linear momentum).

As far as is currently known, gravity, electromagnetic and strong nuclear interactions conserve parity. The same cannot be said for the weak interaction, the P-violating properties of which were first proven in the 1956 Wu experiment on the ^{60}Co β^- decay [12].

1.2.2 Charge conjugation

The transformation of *charge conjugation* $\hat{C}$ changes the sign of all electric charges:

$$\hat{C} : q \to -q. \tag{1.9}$$

It should be readily apparent that $\hat{C}$ has close ties with the concept of antimatter. The action of charge conjugation turns a quantum state into its antimatter partner, inverting the sign of all flavour quantum numbers in the process:

$$\hat{C}|\psi\rangle = |\bar{\psi}\rangle. \tag{1.10}$$

For single-particle systems, the only possible $\hat{C}$ eigenstates are particles that are their own antiparticle, like the photon, for which a C-parity $\eta_C = \pm 1$ is defined by analogy with the P-parity eigenvalue.

Unlike in the case of parity, there isn't a single breakthrough experiment credited for showing that the weak interaction is not C-symmetric: it was known from theory that parity violation in a weak process, when observed under certain conditions, would also imply a violation of charge conjugation, with such a violation being confirmed shortly after Wu's results [13]. As for the other three fundamental interactions, no evidence of C symmetry violation has surfaced so far.

1.2.3   Time reversal

Perhaps the most intuitively named of the three discrete symmetries discussed here, *time reversal* $\hat{T}$ does exactly what it promises:

$$\hat{T} : t \to -t. \tag{1.11}$$

The action of time reversal on a quantum state is represented by an *antiunitary* (unitary and antilinear) operator, which implies a complex conjugation on top of the time reversal itself:

$$\hat{T} |\psi(\vec{x}, t)\rangle = |\psi^*(\vec{x}, -t)\rangle \tag{1.12}$$

There are a number of arguments for the antiunitarity of $\hat{T}$, the most straightforward being that it prevents final states with negative energy.

Once more, gravity, strong and electromagnetic forces are T-symmetric, whereas the weak nuclear force isn't. However, as will be explained shortly, this knowledge isn't the result of a direct experiment, instead exploiting a side effect of the CPT theorem.

1.2.4   CP symmetry and violation

The sequential combination of C, P and T transformations, commonly designated as CPT symmetry, plays a key role in the foundations of quantum physics. As well as being the only combination of said transformations still observed to be a symmetry of physical laws, the *CPT theorem* states that any Lorentz-invariant local quantum field theory must be CPT-symmetric. Because a violation of the CPT symmetry would imply the collapse of the modern quantum physics framework, it is generally accepted that a T-violating process must also be a CP-violating process. This bears an important consequence on the study of discrete symmetry violations: because of the self-evident hindrances in building a time-reversed experimental setup outside of trivial cases, every test of T violation becomes by necessity a test of CP violation.

CP symmetry is also an interesting field of study in and of itself [14]. While C and P symmetries are maximally violated by the weak interaction, CP isn't; this is exemplified by the chirally left-handed neutrino, which possesses a CP-partner (the right-handed antineutrino) despite lacking both a P-partner (the right-handed neutrino) and a C-partner (the left-handed antineutrino).

The subject of CP violation is also closely tied to another long-standing dilemma in both particle physics and cosmology: the observed asymmetry between matter and antimatter in our Universe. A perfectly CP-symmetric system would produce a roughly equal number of particles and antiparticles, which would annihilate one another and yield an empty Universe; our very

existence implies a primordial imbalance that resulted in baryogenesis and therefore some degree of CP violation.

Since gravity, EM and strong interactions all individually conserve parity and charge conjugation, it stands to reason that they are also CP-symmetric, leaving the weak nuclear interaction as the only possible CP-violating fundamental force. Experiments conducted over the last 50 years have found that, despite CP still being preserved in most weak processes, some select interactions show evidence of CP violation.

The first indirect discovery came in 1964 by Christenson et al. [15], who observed the long-lived neutral kaon two-pion decay $K_L^0 \to \pi^+\pi^-$ with a branching ratio of $\approx 10^{-3}$ over all charged modes. This result could only be explained by assuming that the K_L^0 weak eigenstate is a mixture of both $\eta_{CP} = \pm 1$ eigenstates, with the ability to oscillate between the two. A more direct evidence was found in 1999 by the KTeV collaboration at Fermilab [16] via the observation of differing decay rates in $K_{L/S}^0 \to \pi^+\pi^-$ against $K_{L/S}^0 \to \pi^0\pi^0$ channels, and CP violation in weak processes was definitely established in the early 2000s via studies on B mesons decays conducted in so-called «B-factories» such as BaBar at SLAC [17] and Belle at KEK [18].

Despite the significant number of experimental evidences collected ever since, the extent of known CP-violating processes is several orders of magnitude below what is expected from cosmological estimates. The matter-antimatter imbalance at the time when the Universe cooled below the pair production threshold temperature can be quantified through the *baryon asymmetry parameter* [19], computed as the difference between the densities of baryons and antibaryons divided by their sum:

$$\eta := \frac{n_B - n_{\bar{B}}}{n_B + n_{\bar{B}}}. \tag{1.13}$$

While this parameter cannot be directly measured at the present time, we can approximate it by noting that almost no antimatter currently exists in the Universe ($n_{\bar{B}} \approx 0$) and almost all of the original matter will have annihilated into photons ($n_B + n_{\bar{B}} \approx n_\gamma$):

$$\eta \approx \frac{n_B}{n_\gamma}. \tag{1.14}$$

Both of these quantities can be probed by studying the intergalactic medium and the cosmic microwave background, finding $\eta_{\mathrm{obs}} \approx 10^{-10}$ [20]. As for the Standard Model prediction, all SM sources of CP violation arise from quark mixing, more specifically from the δ_{13} complex phase mentioned in the parameterization (1.5) of the CKM matrix. Computing the baryon asymmetry parameter with this knowledge leads to a much lower $\eta_{\mathrm{SM}} \approx 10^{-20}$ [19].

New sources of CP violation are therefore required to match the observed value, with a promising field being the search for intrinsic electromagnetic dipole moments [14].

1.3   Electromagnetic dipole moments

1.3.1   EDMs

The electric dipole moment (EDM) $\vec{\delta}$ is the measure of a system's *polarity*, i.e. the spatial separation of positive and negative charges within the system. For the simplest of the relevant charge configurations, a dipole of point charges $\pm q$ separated by a distance r, the EDM is expressed as

$$\vec{\delta} = q\vec{r}, \tag{1.15}$$

where the displacement vector $\vec{r}$ points from the negative charge to the positive one.

It's hardly a feat of imagination to theorize that a composite particle like the neutron could acquire an EDM, even if the three quarks inside it cannot be thought of as a system of charges in the classical sense. It may be less intuitive that elementary, point-like particles such as electrons and quarks can also gain one due to quantum effects resulting in the creation and destruction of virtual particles (so-called *loops* in higher order Feynman diagrams).

For a spin-$\frac{1}{2}$ particle, its EDM is written in Gaussian units as [21]

$$\vec{\delta} = d\frac{\mu_B}{2}\vec{s}, \tag{1.16}$$

where d is a dimensionless quantity referred to as *gyroelectric factor*,

$$\mu_B = \frac{e\hbar}{2mc}, \tag{1.17}$$

is the particle magneton, c is the speed of light in a vacuum, m is the particle mass and

$$\vec{s} = 2\frac{\langle \vec{S} \rangle}{\hbar} \tag{1.18}$$

is the spin polarization vector, related to the average value of the spin $\vec{S}$ divided by the reduced Planck constant $\hbar$.

When the particle crosses an external electric field $\vec{E}$, its EDM will polarize by changing the direction of the spin. This introduces an energy term in the system's Hamiltonian with the form

$$H_{\text{EDM}} = -\vec{\delta} \cdot \vec{E}. \tag{1.19}$$

We can now check how the term (1.19) behaves when acted upon by some of the discrete transformations outlined in Section 1.2. The behaviour of spin $\vec{S}$, hence of the spin-related EDM $\vec{\delta}$, can easily be shown to be that of polar vectors, i.e. parity-even. By contrast, the electric field $\vec{E}$ is an axial vector, i.e. parity-odd, which makes H_{EDM} a parity-odd pseudoscalar:

$$H_{\text{EDM}} \xrightarrow{\hat{P}} -H_{\text{EDM}}. \tag{1.20}$$

When considering time reversal $\hat{T}$, the situation is specular: the EDM $\vec{\delta}$ flips its sign, whereas the electric field $\vec{E}$ remains unchanged, implying

$$H_{\text{EDM}} \xrightarrow{\hat{T}} -H_{\text{EDM}}. \tag{1.21}$$

The above result in particular contains a crucial piece of information: assuming the validity of the CPT theorem, a Hamiltonian containing the EDM's interaction term (1.19) can only be CP-symmetric if the average (or *permanent*) EDM of the particle is zero.

It follows that, for a particle to have a permanent EDM, CP symmetry must be violated in some measure[3]. As explained in Section 1.2.4, the only known source of CP violation within the Standard Model is the complex phase δ_{13} in quark mixing, which may give a small contribution to the EDMs of point-like particles such as electrons ($\delta_e \lesssim 10^{-40}\,e\,\text{cm}$ [22]) and quarks ($\delta_q \lesssim 10^{-34}\,e\,\text{cm}$ [23]) via beyond-tree-level diagrams. Composite particles such as baryons are accorded some leeway on account of their finite size: the weak interaction between quarks inside the neutron, for instance, contributes to a possible EDM up to $\delta_n \lesssim 10^{-31}\,e\,\text{cm}$ [24].

In all cases, the predicted SM contributions are orders of magnitudes below the sensitivity reached by current generation experiments. For all intents and purposes, the observation of a non-zero permanent EDM in a baryon would imply the discovery of a BSM source of CP violation.

1.3.2 MDMs

The magnetic dipole moment (MDM) $\vec{\mu}$ of a system can be interpreted as the measure of how intense a torque the system experiences when crossing a magnetic field $\vec{B}$:

$$\vec{\tau} = \vec{\mu} \times \vec{B}. \tag{1.22}$$

---

[3]This line of reasoning only applies to systems that are parity eigenstates. Water molecules are notoriously polar, but their EDMs do not violate any fundamental symmetry because the molecule's ground state is a superposition of parity-even and parity-odd eigenstates.

Unlike the case of EDMs, the extension of MDMs from classical to quantum physics is less extreme, as long as one acknowledges the affinity between angular momentum and a particle's intrinsic spin. A classical rotating body with charge q, mass m and angular momentum $\vec{L}$ gains an MDM in the form

$$\vec{\mu} = \frac{q}{2m}\vec{L}, \tag{1.23}$$

assuming charge and mass are identically distributed. A very similar relation holds for a non-classical, point-like spin-$\frac{1}{2}$ particle [21]:

$$\vec{\mu} = g\frac{\mu_B}{2}\vec{s} \tag{1.24}$$

Here g is the dimensionless *gyromagnetic factor* accounting for the transition from classical to quantum physics, whereas μ_B and $\vec{s}$ are the same particle magneton and spin polarization vector defined in equations (1.17) and (1.18) respectively.

Similary to EDMs, MDMs also induce a spin rotation when subjected to a magnetic field $\vec{B}$, introducing a Hamiltonian term

$$H_{\mathrm{MDM}} = -\vec{\mu} \cdot \vec{B}. \tag{1.25}$$

Under parity and time reversal transformations, the MDM $\vec{\mu}$ behaves in the same way as the EDM $\vec{\delta}$, both being dependent on the particle's spin $\vec{S}$ (odd under $\hat{P}$, even under $\hat{T}$). In contrast with $\vec{E}$, however, the magnetic field $\vec{B}$ behaves in the *same* way as $\vec{\mu}$, effectively cancelling out their signs when the Hamiltonian (1.25) is acted upon:

$$H_{\mathrm{MDM}} \xrightarrow{\hat{P}} H_{\mathrm{MDM}}, \tag{1.26}$$

$$H_{\mathrm{MDM}} \xrightarrow{\hat{T}} H_{\mathrm{MDM}}. \tag{1.27}$$

Factoring in the CPT theorem, this result entails that a non-zero intrinsic MDM for fundamental particles does not imply CP violation. For this reason, measurements of MDMs are instead used as precision tests of the CPT theorem, since their values should not change between a particle and its antimatter partner.

1.3.3   Measurement of EDMs and MDMs

For the purposes of this thesis, the measurement of EDMs and MDMs of a particle is performed by exploiting the precession of spin in an electromagnetic field [25]. In the laboratory frame, a neutral particle flying with velocity $\vec{\beta}$

through homogeneous electromagnetic fields $\vec{E}$ and $\vec{B}$ experiences a precession of the non-relativistic spin polarization vector $\vec{s}$ described by the equation

$$\frac{d\vec{s}}{dt} = \vec{s} \times \vec{\Omega}, \qquad \vec{\Omega} := \vec{\Omega}_{\text{EDM}} + \vec{\Omega}_{\text{MDM}}. \tag{1.28}$$

The angular velocity vector $\vec{\Omega}$ is itself the sum of two contributions due to the respective intrinsic dipole moments of the particle:

$$\vec{\Omega}_{\text{EDM}} = \frac{d\mu_B}{\hbar}\left(\vec{E} - \frac{\gamma}{\gamma+1}(\vec{\beta}\cdot\vec{E})\vec{\beta} + \vec{\beta}\times\vec{B}\right), \tag{1.29}$$

$$\vec{\Omega}_{\text{MDM}} = \frac{g\mu_B}{\hbar}\left(\vec{B} - \frac{\gamma}{\gamma+1}(\vec{\beta}\cdot\vec{B})\vec{\beta} - \vec{\beta}\times\vec{E}\right). \tag{1.30}$$

Assuming $\vec{E} = 0$, as will be the case in the experimental setup employed in this work, the angular velocity simplifies to

$$\vec{\Omega} = \frac{\mu_B}{\hbar}\left[d\vec{\beta}\times\vec{B} + g\left(\vec{B} - \frac{\gamma}{\gamma+1}(\vec{\beta}\cdot\vec{B})\vec{\beta}\right)\right]. \tag{1.31}$$

An analytical solution of the above system of equations is possible under the approximation that the precession of the particle spin depends only on the integrated value of the magnetic field $\vec{B}$ along the particle's flight path l. In the absence of field gradients, dictated by the homogeneity requirement, such integrated field $\vec{D}$ can be written as

$$\vec{D} := \int_0^l dl'\, \vec{B}(\vec{r}_0 + \hat{\beta}l') \approx \langle\vec{B}\rangle\, l, \tag{1.32}$$

where $\hat{v}$ labels the normalized vector of $\vec{v}$. The time evolution of the spin polarization vector $\vec{s}$ ($\vec{s}(0) = \vec{s}_0$) is

$$\vec{s}\,(t) = \left(\vec{s}_0\cdot\hat{\Omega}\right)\hat{\Omega} + \left[\vec{s}_0 - \left(\vec{s}_0\cdot\hat{\Omega}\right)\hat{\Omega}\right]\cos\left(\left|\vec{\Omega}\right|t\right) + \left(\vec{s}_0\times\hat{\Omega}\right)\sin\left(\left|\vec{\Omega}\right|t\right). \tag{1.33}$$

From an experimental point of view, measurement of time isn't trivial; instead, one can efficiently measure the flight length of an unstable particle $l = \beta ct$ during its lifetime. The equation describing the spin precession as a function of l has a very similar form to (1.33):

$$\vec{s}\,(l) = \left(\vec{s}_0\cdot\hat{\Phi}\right)\hat{\Phi} + \left[\vec{s}_0 - \left(\vec{s}_0\cdot\hat{\Phi}\right)\hat{\Phi}\right]\cos\left|\vec{\Phi}\right| + \left(\vec{s}_0\times\hat{\Phi}\right)\sin\left|\vec{\Phi}\right|, \tag{1.34}$$

with

$$\vec{\Phi} = \frac{\mu_B}{|\vec{\beta}|\hbar c}\left[d\vec{\beta}\times\vec{D} + g\left(\vec{D} - \frac{\gamma}{\gamma+1}(\vec{\beta}\cdot\vec{D})\vec{\beta}\right)\right] \tag{1.35}$$

and $\vec{D}$ defined as in (1.32).

Equation (1.34) provides a way to measure the values of EDMs and MDMs for neutral particles by studying the change in polarization after their flight through a magnetic field. For unstable particles, the polarization at the time of decay can be inferred in a fairly straightforward way from the angular distribution of their products. Conversely, theoretical knowledge or measurement of the original spin polarization $\vec{s}_0$ of the particle are both far from easy tasks in a general setting.

1.4 Proposal of a measurement of the $\mathbf{\Lambda^0}$ electromagnetic dipole moments with the LHCb detector

The Λ^0 baryon, also historically known as the Λ^0 hyperon[4] and sometimes labelled only as Λ, is a spin-$\frac{1}{2}$ baryon with (u, d, s) valence quarks. As the first identified baryon beyond the two nucleons, it played a key role in the discovery and christening of the strange quark: its mass of $\approx 1116\,\mathrm{MeV}/c$ [3], the lightest among s-bearing baryons, meant its only viable decay channels were mediated by the flavour-changing weak interaction, giving the Λ^0 a much longer half-life than expected; this property was dubbed *strangeness*, a name later inherited by the new quark that indirectly caused it.

The Λ^0 baryon is a prime candidate to probe CP violation. Unlike in the case of the prospective discovery of a neutron EDM[5], a non-zero Λ^0 baryon EDM could not be explained by any phenomena within the Standard Model and would therefore imply the existence of BSM physics.

1.4.1 Highlights from previous measurements of the $\mathbf{\Lambda^0}$ dipole moments

Among the first meaningful measurements of the Λ^0 electromagnetic dipole moments were conducted in the early seventies at CERN near Geneva, Switzerland, by means of nuclear emulsion experiments [27] [28]. The experimental

---

[4]A *hyperon* is a baryon with one or more strange quarks, but no heavier quarks. The nomenclature emerged in the period following the discovery of the strange quark, when no further quarks besides the first three were known; nowadays, the term is rarely used.

[5]Quantum chromodynamics allows for a CP-violating term proportional to the QCD vacuum angle θ. Current measurements of the neutron EDM [26] constrain $\theta \lesssim 10^{-10}$, a fine-tuning suppression known as the *strong CP problem*; nevertheless, experimental discovery of a non-zero neutron EDM could be traced back to this term and would not necessarily require the introduction of new physics.
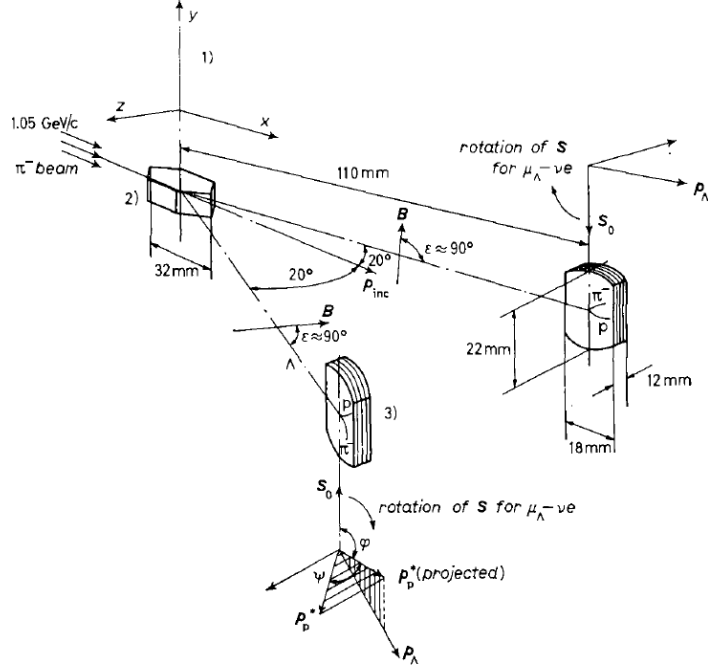
Figure 1.2: Experimental layout used for the 1971 Λ^0 EDM/MDM measurements [27].

layout is sketched in Figure 1.2: Λ^0 baryons are produced via reaction

$$\pi^- + p \rightarrow \Lambda^0 + K^0, \tag{1.36}$$

running a $1.05\,\text{GeV/c}$ π^- beam from the CERN Proton Synchrotron accelerator into a fixed polyethylene target. The production cross section peaks at $0.8\,\text{mb}$ and the resulting Λ^0 have $\approx 100\%$ transverse polarization over a large angular production region in the reaction reference frame. The Λ^0 angular acceptance was restricted to $18°$–$22°$, near the maximum production angle of $21°$, to amplify the signal-to-background ratio; this corresponds to Λ^0 with $|\vec{p}| \in [500\,\text{MeV/c}, 800\,\text{MeV/c}]$. Traversing a $20\,\text{T}$ transverse pulsed magnetic field, Λ^0 baryons become spin-polarized as described in Section 1.3.3 before decaying into a $p\pi^-$ pair. The final polarization direction is probed through the angular distribution of the above decay products, detected by $1.2\,\text{mm}$ thick Ilford K5 nuclear emulsion stacks. The measured Λ^0 magnetic dipole moment, improving an order of magnitude over the previous world average value, was

$$\mu_\Lambda = (-0.66 \pm 0.07)\,\mu_N, \tag{1.37}$$

with μ_N being the nuclear magneton [27]. The Λ^0 EDM was measured at

$$\delta_\Lambda = (-5.9 \pm 2.9) \times 10^{-15}\,e\,\text{cm}, \tag{1.38}$$

Figure 1.3: Side view of the experimental apparatus (*top*) and perspective illustration of the coordinate system (*bottom*) of the Fermilab setup for the Λ^0 EDM/MDM measurements [29].

the 95% confidence level upper limit being found at $\delta_\Lambda \lesssim 10^{-14}$ e cm [28].

These results were improved during the late seventies and early eighties with the neutral hyperon spectrometer at Fermilab in Batavia, Illinois [29] [30]. Figure 1.3 depicts the experimental setup: the incident 300 GeV proton beam impacts on a beryllium target with an angle determined by the M1 upstream magnet; outgoing particles are focused by a brass collimator and cross the M2 magnetic field, serving the dual purpose of purging the beam of charged particles and triggering the Λ^0 spin precession in the horizontal plane; finally, the $\Lambda^0 \to p\pi^-$ decay products within a defined volumetric acceptance are detected by the C1–C6 multi-wire proportional chambers. With a much lower Λ^0 initial polarization ($\approx 8\%$ on average) with respect to the CERN emulsion experiments, the Fermilab team was nonetheless able to best their results owing to a tenfold increase in Λ^0 statistics and precise measurement of the integrated magnetic field. The extracted value for the MDM [29] was

$$\mu_\Lambda = (-0.6138 \pm 0.0047)\,\mu_N, \tag{1.39}$$

while the EDM measurement [30] was

$$\delta_\Lambda = (-3.0 \pm 7.4) \times 10^{-17}\,e\text{ cm}, \tag{1.40}$$

setting the upper limit to $\delta_\Lambda < 1.5 \times 10^{-16}$ e cm [3]. At the time of writing, this is the current best limit on the Λ^0 electric dipole moment.

1.4.2 Proposal overview

My work in this thesis is directed towards the prospective measurement of the Λ^0 baryon electromagnetic dipole moments with the LHCb experiment at the Large Hadron Collider (see Chapter 2), exploiting the spin precession technique outlined in Section 1.3.3. The unique features of the LHCb experimental setup and a careful selection of the Λ^0 production channel will allow for significant simplifications of the general equation (1.34) for neutral unstable particles.

The LHCb detector is equipped with a magnetic field directed along the laboratory frame y axis[6] used for tracking purposes. Gradient field effects for the $\vec{B}$ field within the detector acceptance are negligible, being estimated at [25]

$$\frac{\hbar}{2mc}\frac{\beta\gamma}{\gamma+1}\frac{|\nabla B|}{B} \approx 7.4 \times 10^{-16}, \tag{1.41}$$

with $B := |\vec{B}|$. Assuming production near the beam collision point, the Λ^0 baryon's average mean life of $\approx 2.6 \times 10^{-10}\,\mathrm{s}$ [3] allows a sizeable number of them to traverse at least part of the LHCb magnetic field region (roughly ranging $z = 2.5\,\mathrm{m}$–$7.95\,\mathrm{m}$) before decaying. Spin precession can thus be measured, provided we know both initial and final polarization states.

1.4.3 Polarization measurements

The problem of the Λ^0 initial polarization measurement is circumvented by selecting Λ^0 produced through the weak decay of the bottom baryon Λ_b^0

$$\Lambda_b^0 \to J/\psi\,(\to \mu^+\mu^-)\,\Lambda^0\,(\to p\pi^-), \tag{1.42}$$

as well as its charge-conjugate[7]

$$\bar{\Lambda}_b^0 \to J/\psi\,(\to \mu^+\mu^-)\,\bar{\Lambda}^0\,(\to \bar{p}\pi^+). \tag{1.43}$$

Parity violation in this decay produces Λ^0 with almost 100% longitudinal polarization [31], meaning that the original polarization is aligned to the Λ^0 momentum in the Λ_b^0 helicity frame (see Figure 1.5 and related discussion).

The nature of LHCb as a forward detector implies that Λ^0 baryons will mostly fly along the laboratory frame z axis, and therefore the initial polarization can be written as $\vec{s_0} = (0, 0, s_0)$. Equation (1.34) for the Λ^0 spin

---

[6]The remainder of this chapter assumes the standard right-handed LHCb coordinate system, see Section 2.2.

[7]For the sake of brevity, charge-conjugate notation will be omitted in the rest of this thesis except where relevant to the topic at hand.

precession after the magnetic field region can thus be simplified assuming a field $\vec{B} = (0, B_y, 0)$:

$$\vec{s} = \begin{cases} s_x = -s_0 \sin \Phi \\ s_y = -s_0 \frac{d\beta}{g} \sin \Phi \\ s_z = s_0 \cos \Phi \end{cases}, \tag{1.44}$$

with

$$\Phi = \frac{D_y \mu_B}{\beta \hbar c} \sqrt{d^2 \beta^2 + g^2} \approx \frac{g D_y \mu_B}{\beta \hbar c}, \tag{1.45}$$

$\beta := \left| \vec{\beta} \right|$ and

$$D_y := D_y(l) = \int_0^l dl' B_y. \tag{1.46}$$

Note from equation (1.44) that a non-vanishing intrinsic EDM introduces a s_y component to the final polarization, the MDM precession of which would otherwise be confined to the xz plane.

The polarization after the magnetic field can be probed by studying the angular distribution of the $\Lambda^0 \rightarrow p\pi^-$ decay products. The expected angular distribution for said decay [25] [32] [33] is

$$\frac{dN}{d\Omega'} = 1 + \alpha \vec{s} \cdot \hat{k}, \tag{1.47}$$

where $\Omega' := (\theta', \phi')$ is the solid angle in the Λ^0 helicity frame (see Figure 1.4a), corresponding to the momentum direction of the proton and pointing along the unit vector

$$\hat{k} = \begin{pmatrix} \sin \theta' \cos \phi' \\ \sin \theta' \sin \phi' \\ \cos \theta' \end{pmatrix}, \tag{1.48}$$

whereas $\alpha \approx 0.732$ [3] is the decay asymmetry parameter. The combined measurements of the initial polarization (from the momenta of Λ^0 produced via decays (1.42) and (1.43)) and the final polarization (from angular distribution (1.47)) allow for a study of Λ^0 electromagnetic dipole moments based on the single components of the precession (1.44).

Deviations from this simplified treatment ought to be considered when taking into account the different relevant frames of reference, three of which are sketched in Figure 1.4:

- the laboratory frame S_L (Figure 1.4c), with the z axis along the proton beam and the y axis along the vertical coordinate;
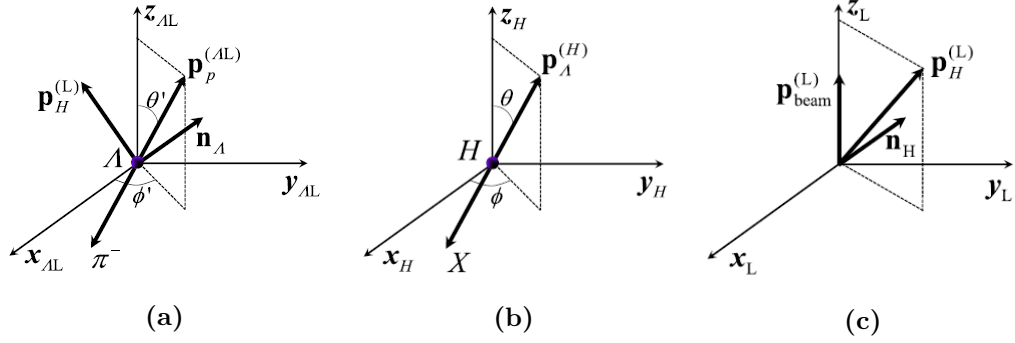
Figure 1.4: Frames of reference for the $\Lambda_b^0 \to J/\psi \, (\to \mu^+\mu^-) \, \Lambda^0 \, (\to p\pi^-)$ decay [25]: *(a)* the Λ^0 helicity frame $S_{\Lambda L}$; *(b)* the Λ_b^0 (heavy hadron) helicity frame S_H; *(c)* laboratory frame S_L. $\vec{p}_H$ is the Λ_b^0 momentum, $\vec{p}_\Lambda$ is the Λ^0 momentum (corresponding to solid angle (θ, ϕ) in the S_H frame) $\vec{p}_p$ is the proton momentum (corresponding to solid angle (θ', ϕ') in the $S_{\Lambda L}$ frame), $\vec{p}_{\mathrm{beam}}$ is the proton beam momentum, while $\vec{n}_\Lambda$ and $\vec{n}_H$ are the normals to the Λ^0 and Λ_b^0 production planes respectively.

- the heavy hadron Λ_b^0 helicity frame S_H (Figure 1.4b), with the z axis given by the Λ_b^0 momentum in S_L and the x axis parallel to the normal to its production plane;

- the two Λ^0 helicity frames $S_{\Lambda L}$ (Figure 1.4a) and S_Λ. These are functionally the same frame of reference, the difference being that the z axis is defined in the direction of the Λ^0 momentum in S_L and S_H respectively.

The polarization given by the equation of motion derived in Section 1.3.3 refers to the comoving rest frame of the Λ^0 (also known as the *canonical frame)*, related to the S_L frame by a Lorentz boost. By contrast, Equation (1.47) for the angular distribution is computed with the solid angle Ω' in the particle helicity frame $S_{\Lambda L}$. Canonical and helicity frames are related by a rotation angle, meaning that $\vec{s_0}$ is not fixed to be perpendicular to $\vec{B}$, as assumed in the solution of system (1.31). This effect arises in the case of Λ^0 not directed along the S_L z axis and is expected to be negligible in the single-arm geometry of the LHCb detector.

More significant is the Wick rotation (see Figure 1.5), owing to the orientation discrepancy between S_Λ frame (where the Λ^0 has the maximal longitudinal polarization) and $S_{\Lambda L}$ (where the angular distribution of Λ^0 decay products is measured). This phenomenon introduces a dilution effect to the precession measurement: a Λ^0 with polarization $\vec{s_0} = s_0 \hat{z}_\Lambda$ in the S_Λ frame gains in the

Figure 1.5: Sketch of the Λ_b^0 production at the primary vertex (PV) and its decay into Λ^0 in the three S_H, S_Λ and $S_{\Lambda L}$ frames of reference, as seen from the S_L yz plane [25]. Particle momenta are represented as *solid* lines for S_L, *dash-dotted* lines for S_H. The Λ^0 polarization vector (*thick arrows on the right*) is aligned to the S_Λ z axis and rotated by Wick angle α in $S_{\Lambda L}$. *Short-dashed* lines trace the $p_\Lambda^{(L)}$ back to the $z = z_{PV}$ plane, identifying the apparent production point of the Λ^0 in S_L (impact parameter). These points correlate with the Λ^0 helicity angle θ computed in S_H, also pictured in S_L as θ_L. By virtue of (1.49), this imprints a correlation between Λ^0 impact parameter and Wick rotation of its polarization vector, pictorially highlighted by the *thick arrows on the left*.

$S_{\Lambda L}$ frame a transverse component of magnitude $s_0 \sin \alpha$, where

$$\sin \alpha = \frac{m_\Lambda}{m_H} \frac{\left| \vec{p}_H^{(L)} \right|}{\left| \vec{p}_\Lambda^{(L)} \right|} \sin \theta \tag{1.49}$$

and θ is the Λ^0 helicity angle, i.e. the angle formed by the Λ^0 momentum in S_H with respect to the frame z_H axis [34]. As seen from Figure 1.5, the helicity angle θ is related to the impact parameter with respect to the primary vertex in the S_L frame [35]. Since θ determines the magnitude of Wick angle α, this relation can be exploited to select ensembles of Λ^0 with similar initial polarization, so that the Wick effect on the particles within a specific ensemble can be neglected.

1.4.4   Closing remarks

The LHCb tracking dipole magnet provides an integrated field $D_y \approx \pm 4 \, \mathrm{T\,m}$ [36], allowing for a maximum precession angle of $\Phi_{\max} \approx \pm \frac{\pi}{4}$ for Λ^0 baryons

traversing the entire region and reaching about 70% of the maximum s_y component in equation (1.44).

With the 8 fb^{-1} integrated luminosity provided by the LHC Run[8] 1 and 2 data collected with the LHCb detector and a global event reconstruction efficiency of $\varepsilon = 0.2\%$, the attainable sensitivity on the Λ^0 gyroelectric factor d has been estimated at $\sigma_d \approx 1.5 \times 10^3$, to be compared with the current best limit of 1.7×10^{-2}. With the Run 3 projected luminosity of $50\,\text{fb}^{-1}$ and an efficiency of $\varepsilon = 1\%$ achievable with the upgraded LHCb trigger system (see Section 2.4), this limit can further be improved to $\approx 3 \times 10^{-4}$ [25]. A precision measurement of the gyromagnetic factor g for Λ^0 and $\bar{\Lambda}^0$ baryons could also serve as a further precision test of the CPT theorem.

---

[8]*Run* is the conventional term for data-taking periods at LHC, see Section 2.1.

Chapter 2

The LHCb experiment

2.1 The Large Hadron Collider

At the moment of writing, the Large Hadron Collider (LHC for short) is the largest and most powerful particle collider in the world. When the LHC was first approved by the European Organization for Nuclear Research (CERN) in 1994, it was originally going to be built with an initial center-of-mass collision energy $\sqrt{s} = 10$ TeV, with the plan to upgrade it to 14 TeV at a later stage; however, after negotiations with nonmember states, in 1996 the CERN council approved the construction at 14 TeV energy in one stage [37].

LHC is located at the CERN laboratory near Geneva, Switzerland, housed in the underground tunnel previously occupied by the LEP experiment. Its structure, sketched in Figure 2.1, consists of two counterrotating rings hosting beams for particle-particle collisions (mainly protons, but LHC is also used for ion collisions).

Four main experiments are stationed at the LHC ring intersection points: ATLAS and CMS are high-luminosity experiments focused on general purpose proton-proton collisions; ALICE is optimized for lead-on-lead collisions with lower center-of-mass energy and luminosity compared to the former two; finally, LHCb is dedicated on the study of b hadrons and will be the focus of the rest of this chapter. Beyond the above four, a number of small-scale, more specialized experiments also work with LHC, such as TOTEM, MilliQan and MoEDAL.

The LHC schedule alternates data collection phases (*Runs*) with maintenance work phases (*Long Shutdowns*, LS for short); while the shutdowns are designed for consolidation and improvement of the collider itself, mainstay experiments usually take advantage of the hiatus to upgrade their detectors in both hardware and software. Run 1 began in 2010 and continued until early 2013, during which period the LHC provided a center-of-mass energy of $\sqrt{s} = 7$–8 TeV. After the 2-year-long interruption for LS1, operations resumed

Figure 2.1: Layout of the Large Hadron Collider with its four main experiments [37].

in 2015 for Run 2 with an upgraded $\sqrt{s} = 13\,\text{TeV}$. The collider was once again shut down at the end of 2018 (LS2) and work began towards the High Luminosity (HiLumi) upgrade of the LHC. After suffering delays due to the COVID-19 pandemic, Run 3 is currently scheduled to start in the second quarter of 2022 with $\sqrt{s} = 14\,\text{TeV}$, finally reaching its maximum collision energy by design.

2.2 The LHCb experiment and detector

LHCb (the b stands for *beauty*[9]) is a single-arm detector designed to study heavy-flavour physics at the LHC, with the main objective of providing precision measurements of CP violation and rare decays of b and c hadrons [39].

Unlike the other three main experiments at LHC, LHCb has a forward-optimized geometry shown in Figure 2.2, with an angular acceptance of 10–300 mrad in the bending plane and 10–250 mrad in the non-bending plane. Such a layout, more reminiscent of fixed target experiments than beam colliders,

---

[9]Before settling on the names *top* and *bottom* for the third generation of quarks, the names *truth* and *beauty* were among those proposed. While they never gained enough momentum in the scientific community, echoes of the failed nomenclature are still present in heavy quark vocabulary, for instance in the alternative name *truth* for the *topness* flavour number mentioned in Section 1.1.2, as well as in the official name for the LHCb experiment.
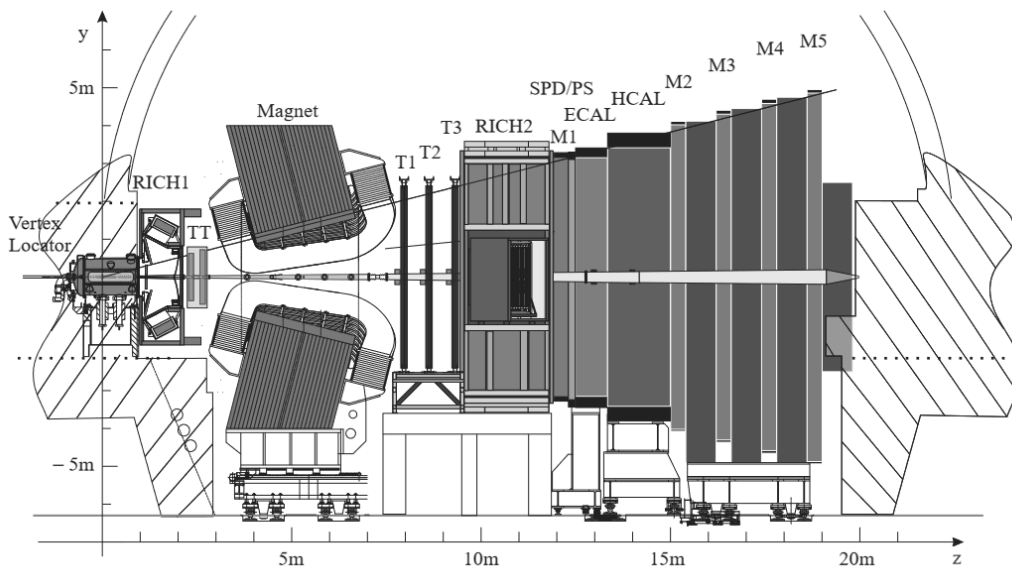
Figure 2.2: Side view of the LHCb detector used for LHC Runs 1 and 2 [38].

is motivated by the fact that $b\bar{b}$ pairs produced at high energies are usually collimated in the same forward/backward cone. A more in-depth look at the tracking and particle identification systems will be taken in Sections 2.2.1 and 2.2.2 respectively. The standard LHCb coordinate system, used as reference for the rest of this thesis, is a right-handed system centered on the beam interaction point, with the z axis along the the beam pipe and y axis directed vertically upwards.

By all accounts, LHCb has been an incredibly successful experiment, reaching its 600th publication by the end of 2021 and leading to state-of-the-art results both within and outside of its intended research framework. Among the main results achieved by the LHCb Collaboration in the field of CP symmetry violation are world-class precision measurements of heavy quark mixing [40] [41], the first observation of CP violation in the charm sector [42] and competitive measurements of the CKM unitarity triangle parameters [43]. LHCb has also been active in the study of rare b-hadron decays [44], as well as in conventional and exotic hadron spectroscopy [45], particularly concerning evidences of new pentaquark states [46]. Finally, LHCb has the distinction of being the only current LHC experiment with the ability to use a fixed-target setup by exploiting the SMOG internal gas target [47], originally developed for precise luminosity measurements; ever since 2015, the adoption of noble gas targets has allowed the Collaboration to obtain b-physics results involving nuclear collisions [48].
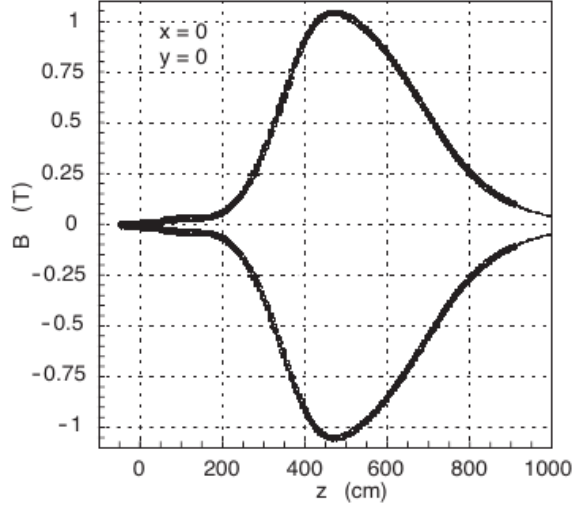
Figure 2.3: LHCb magnetic field along the z axis [49].

2.2.1   Tracking

In order to measure the momenta of charged particles through their bending curve, LHCb employs a dipole magnet [49] consisting of two trapezoidal coils bent at 45° on the two transverse sides, seen in Figure 2.2 around $z \approx 5\,\text{m}$ (the magnet is placed so that the line connecting the centers of the pole faces crosses $z = 5.3\,\text{m}$).

This magnet provides an integrated field of $\int B dl \approx \pm 4\,\text{T m}$ for $10\,\text{m}$ tracks[10]. Most of this field is contained in the $z \in [2.5, 7.95]$ m region, with a small fraction ($\int B dl \approx 0.12\,\text{T m}$) upstream of $z = 2.5\,\text{m}$. The field map along z, measured with a precision of 4×10^{-4}, is shown in Figure 2.3 for $x = y = 0$. Dishomogeneities in the xy plane for fixed z are estimated at $\lesssim 6\%$ within the LHCb detector acceptance.

VELO

As the name suggests, the VErtex LOcator (VELO) system [50] is designed to provide precision measurements of charged tracks near the beam interaction point, in order to correctly reconstruct detached secondary vertices typical of b- and c-hadron decays.

The VELO detector comprises 42 silicon modules along the beam direction, each consisting of a pair of half discs measuring the radial and azimuthal track

---

[10]The $\pm$ sign is due to the fact that the magnet operates alternatively in up and down polarities, inverting the sign of the magnetic field.

Figure 2.4: Front view diagram of the VELO detector in fully closed (*left*) and fully open (*right*) configurations [50].
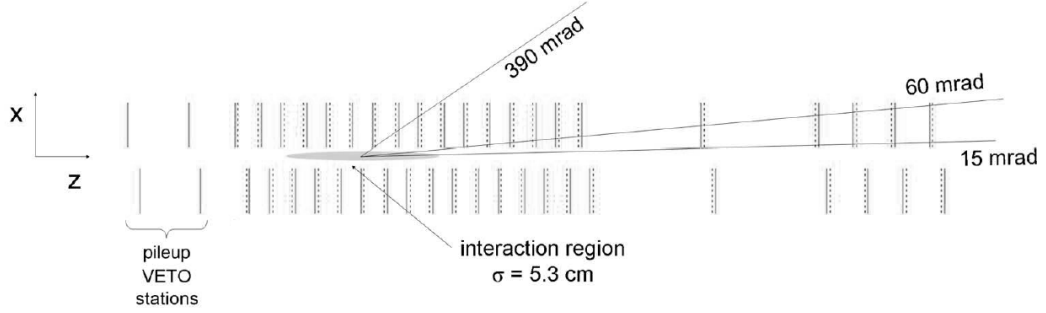

Figure 2.5: Cross-section diagram of the fully closed VELO detector in the xz plane at $y = 0$ (top view). Radial sensors are depicted as *solid* segments, azimuthal sensors as *dashed* segments [50].

coordinates respectively. These modules cover the $1.6 < \eta < 4.9$ positive pseudorapidity range, as well as some negative pseudorapidity portion to improve primary vertex reconstruction, and are able to detect particles emerging from primary vertices with $|z| < 10.6\,$cm. Due to high risk of radiation damage during beam injection from the Super Proton Synchrotron (SPS) into LHC, these modules can be retracted by 3 cm in so-called *fully open* configuration, whereas during collision phase the VELO operates in *fully closed* configuration (see Figure 2.4).

Figure 2.5 shows the xz plane cross-section of the VELO modules; the two halves of the detector are z-shifted by 1.5 cm to ensure full azimuthal acceptance, resulting in the partial overlap seen in fully closed configuration. Four radial-only *pile-up sensors*, part of the Level-0 hardware trigger system (see Section 2.3), are placed upstream to help veto multiple-interaction events.

Figure 2.6: Front view of the third TT layer (different readout sectors are labeled with different shadings) [39].

Tracker Turicensis

The Tracker Turicensis (TT) [51], formerly known as Trigger Tracker, is a $150\,\mathrm{cm} \times 130\,\mathrm{cm}$ tracking station located just upstream of the dipole magnet. Its placement serves the main purpose of tracking low-momentum particles ($|\vec{p}| \lesssim 1.5\,\mathrm{GeV/c}$) that would otherwise be bent out of the detector by the magnet without reaching the T stations.

The TT consists of four readout layers of silicon microstrip sensors arranged in a *x-u-v-x* configuration (vertical in the first and last layers, rotated by a stereo angle of $\mp 5°$ in the second and third layer respectively) for a total active area of $\approx 8.4\,\mathrm{m}^2$. A $200\,\mathrm{\mu m}$ strip pitch ensures a single-hit resolution $\lesssim 50\,\mathrm{\mu m}$.

The third TT layer is depicted in front view in Figure 2.6. The basic unit of a layer is the *half module*, covering half the LHCb height acceptance. Each half module consists of a row of seven sensors bonded together to form either three or two *readout sectors*. Modules near the beam pipe are of the former category, with four sensors bonded in the L sector, two in the intermediate M sector and a single sensor for the K sector closest to the beam (4–2–1 modules); other modules forgo the K sector and bond the spare sensor in the M sector (4–3 modules). Front-end readout hybrids, one for each sector, are placed at the L-end of the half modules, outside of the detector acceptance, connected directly to the L sector and indirectly to the M and K sectors via Kapton flex cables.

Figure 2.7: Top (*left*) and front (*right*) views of a T tracking station [52]. IT and OT are labeled with lighter and darker shades of grey respectively. Dimensions are given in cm; for the top view, lateral dimensions are not to scale.

Figure 2.8: Front view of an *x* detector layer in the T2 Inner Tracker [39].

T stations

The three T stations, labeled as T1–T3, are the last line of defense for LHCb tracking purposes, covering the $z \approx 7.7$–9.4 m region downstream of the dipole magnet [52]. Each T station is composed of an Inner Tracker for the region near the beam pipe and an Outer Tracker for the farther regions, as sketched in Figure 2.7.

The Inner Tracker (IT) [52] shares many similarities with the TT design, being developed in conjuction with it under the common Silicon Tracker (ST) project. Sporting the same four layers of silicon microstrips in *x-u-v-x* config-uration, it covers a comparatively smaller 120 cm $\times$ 40 cm cross-shaped surface (see Figure 2.8) for a total active area of ≈ 4 m^2, less than half the TT. As a

Figure 2.9: Side view diagram of the LHCb tracking systems for LHC Runs 1 and 2 with sketched examples of the main track classification categories.

consequence, individual modules only include one or two sensors connected to the readout hybrids via a pitch adapter.

The much larger Outer Tracker (OT) [53] is a drift detector consisting of an array of Ar/CO_2 straw-tube modules. Each module contains two layers of straw tubes with $4.9\,$mm inner diameter, ensuring a $50\,$ns drift time and $200\,\mu$m spatial resolution. Within a single T station, said modules are arranged in four layers in x-u-v-x configuration (see Figure 2.7a) with $\pm 5°$ vertical tilt for u and v layers respectively. The OT covers the entire $300/250$ mrad LHCb detector acceptance.

Track classification and the problems with T tracks

Overall, the LHCb tracking system has very high efficiency, besting 96% in the momentum range $|\vec{p}| = 5$–200 GeV/c for tracks crossing all three detector stations (VELO, TT and T1–T3) [54]. However, not all particles enjoy this luxury: low momentum particles ($|\vec{p}| \lesssim 1.5\,$GeV/c) are unable to reach the T stations due to the sharp magnet bending curve, while daughters of longer-lived particles with $c\tau \gtrsim 30\,$cm will miss the VELO and possibly even the TT detector.

Thus, in spite of the great efficiency, it's useful to define track categories in the LHCb working environment depending on what hits were recorded in which detectors:

- *Long* tracks are reconstructed from hits in the VELO and T stations, with TT hits being optional. They are the most commonly used tracks

for physics analysis in LHCb, exploiting the full performance of the detector tracking system; the long distance between VELO and T detectors means that only stable or quasi-stable charged particles can aim for this classification.

- *Upstream* tracks are reconstructed from hits in the VELO and TT detectors. As previously discussed, low-momentum particles usually fall in this category due to inability to reach the T stations.

- *VELO* tracks are reconstructed from hits in the VELO detector alone. These are often large-angle or backward tracks, valuable to correctly identify the primary vertex of the event.

- *Downstream* tracks are reconstructed from hits in the TT and T1–T3 detectors. This is the most common category to study long-lived particles (LLP) decaying after the VELO, such as Λ^0 baryons and K_S^0 mesons.

- *T* tracks are only reconstructed from hits in the three T stations. Since missing both VELO and TT is a rare occurrence, these tracks largely come from LLPs with $c\tau \gtrsim 5\,\mathrm{m}$ decaying after the dipole magnet.

Sketches of tracks satisfying the above requirements are depicted in Figure 2.9.

Among the four, T tracks have seen the least use over LHC Runs 1 and 2. The main reason is related to their key property of only having measured constraints after the magnet: in order to reconstruct their origin vertex, T tracks have to be extrapolated through several meters of intense magnetic field, an operation that has ripercussions in terms of both accuracy and timing. On top of that, the further a charged particle is produced after the dipole magnet, the less the residual magnetic field will be able to imprint a significant bending radius, negatively affecting the momentum resolution at T station level.

Over the ten years of detector operation there has been little physics incentive to fix or mitigate these issues, as most LLPs relevant to LHCb can already be studied using Downstream tracks. The Λ^0 EDM/MDM measurement proposal outlined in Section 1.4 is one of the atypical cases where T tracks are downright essential, since the technique hinges on the measurement of the spin precession of the baryon traversing the magnetic field before decaying in the $p\pi^-$ pair. Over the course of the following chapters, I'll delve into more detail on the approaches adopted to overcome the problems associated with T tracks in view of competitive measurements of the Λ^0 electromagnetic dipole moments.

Figure 2.10: Top view of the RICH 1 (*left*) and RICH 2 (*right*) detectors [39].

2.2.2 Particle identification

While tracking outgoing particles is obviously of paramount importance for physics analysis, knowledge of *what* particles are being tracked is also crucial. The ability to distinguish protons, pions and kaons is of particular interest at LHCb due to its research objectives in CP violation and *b* physics, requiring precise flavour tagging and physical background rejection. For the above reasons, a complex ecosystem of detectors dedicated to particle identification (PID) is in place.

RICH

Roughly 90% of pions, protons and kaons from *B* meson decays have momentum in the 2–150 GeV/c range [54]. Since their momentum distributions depend on the polar angle of production, LHCb employs two Ring Imaging CHerenkov (RICH) detectors [55] to cover the full momentum range for these particles.

The RICH 1 detector, sketched in Figure 2.10a, is located upstream of the dipole magnet, wedged between the VELO and TT tracking detectors. The detector exploits the different spectra of Cherenkov angles as a function of momentum for different kinds of particles. During Run 1, RICH 1 used two radiator materials: an aerogel layer ($n = 1.03$) and a C_4F_{10} gas layer ($n = 1.0014$). This allowed RICH 1 to perform π/K identification in the 1–60 GeV/c range. Due to occupancy problems, the silica aerogel radiator providing identification in the low momentum range $|\vec{p}| \lesssim 10\,\text{GeV/c}$ was removed for Run

Outer section :
121.2 mm cells
2688 channels
Middle section :
60.6 mm cells
1792 channels
Inner section :
40.4 mm cells
1536 channels

Outer section :
262.6 mm cells
608 channels
Inner section :
131.3 mm cells
860 channels

(a)   (b)

Figure 2.11: Front view of the lateral segmentation of SPD/PS and ECAL (*left*) and HCAL (*right*) calorimeters [39]. Only a quarter of the detector is depicted. Dimensions are given for the ECAL in the left figure.

2; since the kaon Cherenkov threshold in C_4F_{10} is $\approx 9.7\,\mathrm{GeV/c}$, they can still be identified by operating RICH in so-called *kaon veto mode*, whereby kaons are marked by the lack of Cherenkov light [54] [56]. RICH 1 covers from 25 mrad (lower limit imposed by the beryllium beam pipe section) up to the full LHCb acceptance.

Acting as complement to its partner, RICH 2 (Figure 2.10b) operates downstream of the T tracking stations and is optimized for a high momentum range, providing PID from $\approx 15\,\mathrm{GeV/c}$ up to and beyond $100\,\mathrm{GeV/c}$. Its lower limit of acceptance is ≈ 15 mrad, dictated by the required clearance of 45 mm around the beam pipe.

Calorimeter

The LHCb calorimeter system [57] serves the dual purpose of identifying hadrons, electrons and photons and measuring their energies. Its design follows the standard high energy physics approach of an electromagnetic calorimeter (ECAL) for the detection of electrons and photons, followed by a hadronic calorimeter (HCAL) for the detection of charged and neutral hadrons.

Placed at 12.5 m from the beam interaction point, the ECAL employs a shashlik layout[11], alternating layers of absorber (2 mm thick lead) and sampler (4 mm thick polystyrene scintillator tiles) perpendicular to the beam axis. Due to the steep dependence of hit density from the distance from the beam pipe, the calorimeter adopts a variable cell size and is segmented in three distinct sections outlined in Figure 2.11a. The ECAL is approximately $25X_0$ long,

---

[11]The nomenclature references the *shashlik*, or *šašlyk*, a traditional meat dish consisting of skewers threaded with alternating pieces of meat, fat and vegetables. The dish is popular throughout the Caucasus and Central Asia regions, including the former Soviet Union, where the shashlik calorimeter technology was first developed.

with X_0 being the radiation length; this allows for the full containment of electromagnetic showers from high energy photons, which is of paramount importance for energy resolution.

Electron detection is particularly tricky due to the significant pion background, both of the charged and neutral variety. To combat this, two ancillary detectors are located upstream of the ECAL proper: the scintillator pad detector (SPD) selects charged particles to veto π^0, while the preshower detector (PS) rejects $\pi^\pm$. Collectively, the SPD/PS detectors consist of two scintillator pads enclosing a 15 mm thick lead plate with a 7.6 m × 6.7 m sensitive area. Transverse segmentation is designed to projectively match the ECAL segmentation down to the individual cell size.

The HCAL is a sampling calorimeter as well, employing iron as absorber and scintillating tiles as active material. In contrast to the ECAL and SPD/PS detectors, however, the scintillating tiles run parallel to the beam axis, interspersed with 1 cm thick layers of iron; meanwhile, the longitudinal structure alternates scintillating tiles with iron spacers, both of length $\lambda_I \approx 20$ cm, λ_I being the hadron interactiorn length in steel. The transvese segmentation of the detector, sketched in Figure 2.11b, envisages one less section and a comparatively larger cell size than ECAL, owing to the differing sizes of electromagnetic and hadronic showers. Since competitive hadron energy resolution does not require full containment of the shower, the HCAL only extends for ≈ 5.6 interaction lengths.

In all four subdetectors, scintillating light is conveyed through wavelength-shifting fibres to photomultiplier tubes for conversion and magnification; due to the lower light yield of HCAL modules, their phototubes operate at a higher gain.

Muon system

The final components of the PID system are the five muon stations M1–M5 [58], providing trigger and limited tracking for muons in LHCb. Stations M2–M5 are placed downstream of the calorimeter system, separated from each other by 80 cm of iron; these absorber layers select muons on the basis of penetration, with a 6 GeV/c momentum threshold required to cross the fifth station. The lone M1 station precedes the calorimeter system with the goal of improving transverse momentum measurement. The muon system provides acceptance in the 20–306 mrad region in the bending plane and 16–258 mrad region in the non-bending plane, in line with the global LHCb acceptance.

A side view diagram of the muon system is depicted in Figure 2.12: each station is divided in four R1–R4 regions with increasing distance from the beam pipe. While transverse spatial resolution progressively worsens in outer
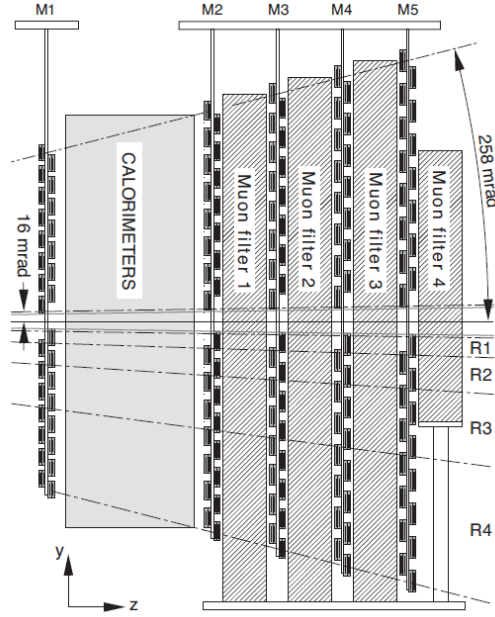
Figure 2.12: Side view diagram of the muon system [39].

regions, the growing influence of large angle multiple scattering means it would be limited anyway.

The most sensitive area is the R1 region of the M1 station, since the large particle flux imposes strict limits on radiation hardness to prevent ageing effects during the LHC projected lifetime. For this reason the M1-R1 region alone employs gas electron multiplier foils, while the remainder of the muon system consists of multi-wire proportional chambers with a $Ar/CO_2/CF_4$ gas mixture.

Overall, the five stations combined cover a total area of $435\,\mathrm{m}^2$. Stations M1–M3, by virtue of their high spatial resolution along the x coordinate, are used to determine the direction of the candidate muon track and compute the transverse momentum with $\approx 20\%$ resolution; stations M4–M5 have lower performance on this front and their contribution mainly consists in the identification of highly penetrating particles.

2.3 The LHCb data flow

Considering the complexity of the LHCb detector environment, as seen in Section 2.2, it should come to no surprise that the data elaboration process is equally multifaceted. Figure 2.13 sketches the data flow approach during LHC Run 2; while a full discussion of the mechanics is beyond the scope of

Figure 2.13: Diagram of the LHCb Run 2 data flow.

this thesis, this section aims to provide a basic understanding of the different steps in order to grasp key concepts relevant to the following work.

2.3.1 Trigger

The trigger system [59] provides the first triage of all data recorded by the LHCb detector. LHC collides proton-proton bunches at a nominal 40 MHz rate, with $\approx 1\%$ resulting in $b\bar{b}$ events of interest for LHCb. Furthermore, only $\approx 15\%$ of these events will produce a reconstructible b hadron (i.e. with all decay products within the detector acceptance) [54], and studies on topics such as CP violation are likely to require decays with small ($\lesssim 10^{-3}$) branching ratios. Peak writing speeds for data storage are in the order of a few kHz, making it impossible to save all information even forgoing the high costs this would entail in terms of storage space. The LHCb trigger system therefore has to skim out the vast majority of uninteresting events with high efficiency, and it needs to be fast about it.

The trigger system employed during LHC Runs 1 and 2 can be broken down into three distinct phases, or *levels*. First comes Level-0 (L0), which is implemented directly on hardware via custom-made electronics. Working synchronously with the 40 MHz bunch-crossing rate, the L0 trigger is only able to read three parts of the LHCb detector independently: the VELO pile-up radial modules, used to reject events with multiple primary proton-proton interactions; the calorimeter trigger (L0-Calorimeter), which selects hadron, photon and electron candidates; and the muon trigger (L0-Muon), which obviously selects muons.

The L0-Calorimeter trigger is based on information from all four subdetectors of the calorimeter system (SPD, PS, ECAL and HCAL) and uses it to compute the transverse energy E_T deposited by incoming particles. Events with a large number of charged tracks are vetoed based on the number of hits in the SPD to manage the limited computation time allotted for the subse-

quent trigger levels. Based on the E_T measurement, the trigger builds hadron, electron and photon candidates.

Each of the four L0-Muon trigger processors selects the two highest p_T tracks from its assigned quadrant among candidates crossing all five muon stations. The single muon trigger sets a threshold on the highest transverse momentum p_T of the pair, while the dimuon trigger does so on the product of p_T of both candidates.

After combining all information, the L0 trigger outputs at a maximum rate of 1 MHz fixed by front-end electronics, the majority being used for muon and hadron triggers (electrons and photons only take up $\approx 15\%$ of the L0 output rate). These data are then sent to the event filter farm (EFF), where the software-based high level trigger (HLT) algorithms, implemented in the Moore application, process them to further reduce the rate for storage: the first stage (HLT1) gets the rate down to 100 kHz, while the second (HLT2) outputs at roughly 12.5 kHz. Both HLTs are divided in independently operating *trigger lines*, each line consisting of specific selection instructions for a determined class of events.

HLT1 reconstructs Long charged particles with $p_T > 500$ MeV. First it combines VELO hits to form straight line tracks, then it looks for ≥ 3 TT hits in a region around the straight line extrapolation of the VELO tracks. The small upstream portion of magnetic field allows for momentum determination with a 20% resolution used to reject low-p_T tracks. After that, the tracks are extrapolated at the T stations, looking for hits in IT and OT on one side of the straight line VELO-TT extrapolation (depending on the charge estimate). Finally, all tracks are fit with a Kalman filter with a simplified detector geometry. At this stage, particle identification is only possible for muons on account of the tight timing constraints.

Owing to the rate reduction performed by HLT1, HLT2 is able to reconstruct the entire event: reconstruction of charged tracks with $p_T > 80$ MeV is performed using all tracking sub-detectors (see Section 2.2.1), along with reconstruction of neutral clusters and implementation of the full particle identification system (Section 2.2.2).

2.3.2   From reconstruction to analysis

The data flow reaches a fork after HLT2. While the trigger output can be used for data analysis, the hectic timing requirements for said phase mean that only information on the decay of interest for the related HLT2 trigger line is reconstructed, leaving the rest as raw data.

The *full stream* branch rectifies this by performing a slower-paced, offline re-reconstruction of the full decay tree controlled by the Brunel application.

Resulting data are saved in Data Summary Tape (DST) format, with a single event occupying $\approx 150\,\text{kB}$ of space. These DST files undergo the *stripping* process, carried out by the DaVinci application: this applies dedicated loose selection algorithms (*stripping lines*) and groups events in *streams* (the dimuon stream for $\mu^+\mu^-$ events, for instance) to ease access for data analysis.

The resulting DST files, also referred to as *full* DSTs to distinguish from *reduced* DSTs output by Brunel, can then be processed by the end user through DaVinci to apply their own filter algorithms and extract ROOT nTuples[12] suitable for physics studies.

During Run 1, the full stream path was the only one available, because HLT2 performed a significantly worse reconstruction than Brunel even on the implemented decays. This changed in Run 2 with the introduction of the parallel *Turbo stream* [61], i.e. the storage of HLT2 output for direct usage by analysts through DaVinci. The EFF upgrade conducted during the LS1 increased the HLT2 allotted computation time enough to match HLT2 and Brunel performances; the key difference between the full and turbo streams is that the latter only saves information on the decay of interest for the related HLT2 trigger line to save storage space, preventing future offline reconstruction of the full decay tree by Brunel.

This change was motivated by the increase in center-of-mass collision energy from $8\,\text{TeV}$ to $13\,\text{TeV}$: the higher *b*- and *c*-production cross-sections mean that more events of interest for LHCb are produced in Run 2 despite the same bunch-crossing frequency. The Turbo stream allows more saved data to be ready for analysis, while the slower Brunel offline process fulfills the need for reconstruction of decays outside of those implemented in HLT2 trigger lines.

2.3.3   Monte Carlo simulations

The comparison between experimental results and theory predictions is a critical aspect of physics analysis. In the case of high-energy physics experiments, the latter rely on the correct simulation of events from collision to detector interaction.

In LHCb, the production of Monte Carlo (MC) data is controlled by the Gauss application, which is in charge of coordinating the several cogs of the simulation machine: proton-proton collisions are simulated via MC generator software such as PYTHIA [62] and POWHEG [63]; the decay of generated

---

[12]ROOT [60] is an C++ open-source data analysis toolkit developed developed at CERN and popular within the high-energy physics community. A key feature of ROOT is the `TTree` class (*tree* for simplicity), a C++ object container organized in independent *branches*, or *columns*, and optimized for large data sets. The `TNtuple` class is a `TTree` with float variables only.

particles is described by EvtGen [64], while the Geant4 toolkit [65] simulates the propagation and interaction with the material using a detailed modelization of the LHCb detector.

As the final step, the simulated hits from the virtual detector in Geant4 are digitized using the Boole application, which aims to mimic the real output from the LHCb detector. This allows simulated data to be processed with the same software as the real one, as described in the earlier paragraphs of this section.

2.4 LHCb detector upgrade for Run 3

During LHC Runs 1 and 2, the LHCb experiment has collected $\approx 8\,\text{fb}^{-1}$ of data. As much as this has allowed the LHCb Collaboration to achieve impressive results in the heavy-flavour sector and beyond, as touched upon in Section 2.2, many ongoing analyses are still limited by low statistics. For this reason, during the LS2, work has been carried out to upgrade the detector in software and hardware with the goal of reaching a $50\,\text{fb}^{-1}$ integrated luminosity by the end of Run 3 [66].

The main feature of this upgrade will be the removal of the L0 hardware trigger in favour of a fully software trigger [67]. After the performance match between HLT2 and Brunel achieved in Run 2, the high level trigger will see an even more central stage in Run 3, being tasked with full reconstruction of all events of interest in the context of a five-fold increase in operational luminosity (up to $2 \times 10^{33}\,\text{cm}^{-2}\,\text{s}^{-1}$). To prevent trigger yield saturation with such an increase in luminosity, the LHCb readout rate will be augmented from the current 1 MHz to the LHC bunch crossing rate of 40 MHz.

Many parts of the LHCb detector will concurrently be upgraded or partially rebuilt, as seen in Figure 2.14, with the most significant changes concerning the tracking system. The upgraded VELO, also known as VELO Pixel, will replace the silicon microstrip technology with hybrid pixel modules with integrated CO_2 cooling, ensuring better radiation hardness, impact parameter resolution and tracking times [68]. Upstream of the magnet, the Upstream Tracker (UT) will take over from the TT, featuring four high granularity silicon microstrip layers with improved coverage of the LHCb angular acceptance. Finally, both inner and outer trackers of the T1–T3 stations will be replaced by the Scintillating Fiber Tracker (SFT or SciFi), based on 2.4 m plastic scintillating fibers read out by silicon photo-multipliers [69]. This global upgrade is projected to reduce reconstructed fake tracks by as much as 70%, drastically shortening trigger timing [66].

The PID system will also undergo modifications, albeit less substantial

Figure 2.14: Side view of the upgraded LHCb detector for future usage in LHC Run 3 [66].

ones [70]. Current hybrid photon detectors used in the RICH system cannot be disentangled from the embedded readout electronics operating at 1 MHz and will thus be replaced by commercial multianode photomultipliers; the optical layout of RICH1 will also be updated to spread gas rings over the entire detector plane, reducing particle occupancy problems[13]. Whereas the calorimeter system will largely stay the same, muon station M1 will be removed and additional shielding will surround the beam pipe in the M2 region to improve radiation hardness.

2.5   Data used for this thesis

In the following chapters of this thesis, I will employ data collected during LHCb Run 2, corresponding to an integrated luminosity of $6\,\mathrm{fb}^{-1}$, to conduct studies related to the measurement of the Λ^0 baryon electromagnetic dipole moments pursuing the approach described in Section 1.4. The exclusive decay chosen for the analysis is the Λ_b^0 channel

$$\Lambda_b^0 \to J/\psi\,(\to \mu^+\mu^-)\,\Lambda^0\,(\to p\pi^-), \qquad\qquad (1.42\text{ revisited})$$

---

[13]This is an issue specific to RICH1, being much closer to the interaction point than its downstream counterpart.

exploiting the $\approx 100\%$ longitudinal polarization of the Λ^0. This comparatively rare channel (BR $\approx 10^{-5}$) is selected at HLT level via the inclusive $J/\psi \rightarrow \mu^+\mu^-$ detached[14] trigger line. The need to measure spin precession in the magnetic field restricts the available Λ^0 to those decaying after the LHCb dipole magnet, implicitly requiring its decay products $p\pi^-$ to be T tracks.

The $\tau \approx 1.5 \times 10^{-12}\,$s mean life of the Λ_b^0 [3] places its decay vertex well inside the VELO detector; in conjunction with the high efficiency reconstruction of muons in LHCb, this justifies a Long track requirement in their case. The Λ_b^0 vertex accuracy afforded by the J/ψ half of the decay chain enables us to impose kinematic constraints on the Λ_b^0 half, partially offsetting some of the problems with T tracks I have pointed out in Section 2.2.1.

Along with Run 2 data, simultaded samples of (1.42) events are used to study the effect of the different steps of signal selection requirements. These have been generated with PYTHIA [62] using the LHCb-specific configuration implemented in Gauss [71], and have been digitized following the procedure detailed in Section 2.3.3.

---

[14]The *detached* requirement adds a selection criterium on the muon impact parameters, enforcing their incompatibility with the lowest-χ^2 primary vertex.

Chapter 3

Λ_b^0 and Λ^0 decay vertex reconstruction

This chapter details my work towards the improvement of the vertex reconstruction process for decays involving T tracks. Section 3.1 delves into a deep study of the vertexing process at LHCb and the two algorithms employed in this thesis; Section 3.2 introduces the problem of low vertexing efficiency for the decay of interest $\Lambda_b^0 \rightarrow J/\psi \, \Lambda^0$; Section 3.3 presents my efforts in the characterization of the non-converged events in search for the root cause of the vertexing falure; finally, Section 3.4 proposes my solution to improve the signal yield through partial recovery of non-reconstructed events.

3.1   Vertex reconstruction algorithms at LHCb

3.1.1   Vertex Fitter algorithm

The Vertex Fitter (VF), implemented as part of the LoKi analysis toolkit, is the main vertexing algorithm used for the reconstruction of the Λ_b^0 decay.

Under VF formalism, each daughter particle is represented by a 7-dimensional

vector[15]

$$\vec{p} = \begin{pmatrix} \vec{r} \\ \vec{q} \end{pmatrix} = \begin{pmatrix} r_x \\ r_y \\ r_z \\ p_x \\ p_y \\ p_z \\ E \end{pmatrix}, \tag{3.1}$$

containing its 4-momentum $\vec{q}$ computed at the *reference point* $\vec{r}$. This parameter vector has an associated covariance matrix V, which can be written in block structure as

$$\begin{pmatrix} V_r & V_{rq} \\ V_{rq}^T & V_q \end{pmatrix}. \tag{3.2}$$

It is also convenient to identify its formal inverse matrix $G := V^{-1}$, which has an analogous block form:

$$\begin{pmatrix} G_r & G_{rq} \\ G_{rq}^T & G_q \end{pmatrix} = \begin{pmatrix} V_r & V_{rq} \\ V_{rq}^T & V_q \end{pmatrix}^{-1} \tag{3.3}$$

Taking the daughter particles as inputs, the Vertex Fitter will output the best fit value $\vec{x}$ for the common origin vertex, along with its covariance matrix C and the χ^2 to evaluate the goodness of fit.

The algorithm builds the decay tree from the bottom-up via a «leaf-by-leaf» approach, fitting one vertex at a time (e.g. $J/\psi \to \mu^+\mu^-$, $\Lambda^0 \to p\pi^-$) and then moving upwards (e.g. $\Lambda_b^0 \to J/\psi\,\Lambda^0$). This process is blind to the downstream leaves and only considers kinematic information of the immediate daughter particles, without accounting for momenta and mass constraints.

Iterating paradigm

The basic unit of recursion of the Vertex Fitter is the *iteration*: the algorithm is set to repeat the vertexing process until either a convergence condition is satisfied (see later) or the fit reaches the set number of allowed iterations, 10 by default. In the latter case, a non-convergence error is thrown and the candidate event is discarded.

---

[15]This chapter assumes the standard right-handed LHCb coordinate system, see Section 2.2.

At the beginning of each iteration, the final vertex covariance matrix C_n^{i-1} from the previous iteration[16] is scaled down by a factor $s^2 = 10^{-4}$:

$$C_0^i = C_n^{i-1} \times s^2. \tag{3.4}$$

The algorithm then performs a *proper transportation*, a dedicated routine in which all daughter particles are extrapolated to the z component of the current (tentative) position of the common production vertex $\vec{x}_n^{i-1}$.

As mentioned in Section 2.2.1, extrapolation using T tracks is a sensitive affair: unlike the case for other track types, no constraints are available besides the downstream measurement performed by the T tracking stations, meaning the tracks have to be propagated through several meters while accounting for the intense and non-homogeneous LHCb magnetic field. For this analysis, said extrapolation was performed via numerical solution of the track propagation equations using an approach based on the Runge-Kutta (RK) method [72] [73].

Step

Within an individual iteration i, denoted by a superscript, the Vertex Fitter algorithm proceeds by *steps* denoted by subscripts, with each step k coinciding with the addition of the k-th daughter particle.

Given information on the vertex position $\vec{x}_{k-1}$ obtained using the first $k-1$ particles, track k is added through the following recursive procedure. First the inverse vertex covariance matrix is updated:

$$C_k^{-1} = C_{k-1}^{-1} + V_{rk}^{-1}, \tag{3.5}$$

where the reference point inverse covariance matrix V_{rk}^{-1} has been updated at the beginning of the iteration through the proper transportation phase.

If C_k^{-1} can successfully be inverted, the algorithm updates the current best estimate of the common origin vertex:

$$\vec{x}_k = C_k \left[C_{k-1}^{-1} \vec{x}_{k-1} + V_{rk}^{-1} \vec{r}_k \right]. \tag{3.6}$$

To conclude the step, the vertex χ^2 is updated to the account for the new position:

$$\begin{aligned} \chi_k^2 = \ &\chi_{k-1}^2 \\ &+ (\vec{r}_k - \vec{x}_k)^T V_{rk}^{-1} (\vec{r}_k - \vec{x}_k) \\ &+ (\vec{x}_k - \vec{x}_{k-1})^T C_{k-1}^{-1} (\vec{x}_k - \vec{x}_{k-1}) \end{aligned} \qquad . \tag{3.7}$$

---

[16]The subscript n identifies the final step number, see later.

Seeding

As one can observe, the procedure described above requires, at each step, both a previous estimated vertex position $\vec{x}_{k-1}$ and an associated inverse covariance matrix C_{k-1}^{-1}. In particular, step $k = 1$ demands the existence of $\vec{x}_0$ and C_0^{-1}.

For iterations $i > 1$, such roles are handily filled by the final vertex computed during the previous iteration. For the purpose of providing the first step of the first iteration with these values, at the beginning the algorithm tries to extract a *vertex seed*, a first estimate of the decay vertex position, through a dedicated procedure depending on decay topology and properties of particles involved.

In the case of interest of the $\Lambda^0 \to p\pi^-$ two-body decay, said procedure is a simplified step of the Kalman filter:

$$C_0^{-1} = V_{r1}^{-1} + V_{r2}^{-1} \tag{3.8a}$$

$$\vec{x}_0 = C_0 \left(V_{r1}^{-1}\vec{r}_1 + V_{r2}^{-1}\vec{r}_2 \right) \tag{3.8b}$$

Subscripts 1 and 2 as used above refer to the two daughter particles in the decay (i.e. proton and pion).

Termination and smoothing

The two VF convergence conditions are both based on comparisons between the vertex position computed at the end of the current iteration with the one from the previous iteration, with convergence being called if either one of them is satisfied.

The first condition is placed on the absolute distance between the vertices:

$$\left\| \vec{x}_n^i - \vec{x}_n^{i-1} \right\| \leq d_1 \tag{3.9}$$

where $d_1 = 1\,\mu\text{m}$ by default. The second condition, by far the more commonly satisfied one when reaching convergence, is a condition on vertex distance «in χ^2 units»:

$$\left(\vec{x}_n^i - \vec{x}_n^{i-1} \right)^T C_n^{i\,-1} \left(\vec{x}_n^i - \vec{x}_n^{i-1} \right) \leq d_2 \tag{3.10}$$

with $d_2 = 0.01$. While condition (3.9) can be satisfied at any point in the vertexing process, (3.10) convergence additionally requires $i > 1$, thereby excluding the very first iteration.

When convergence is reached, the algorithm applies a smoothing process: for each daughter particle, the reference point $\vec{x}_k$ is fixed to the final vertex position $\vec{x}_n^i$ and the momentum $\vec{q}_k$ is updated accordingly as

$$\vec{q}_k = \vec{q}_n^i - V_{rq_k} V_{rk}^{-1} \left(\vec{r}_k - \vec{x}_k \right) \tag{3.11}$$

Finally comes the evaluation of the relevant covariance matrices. The vertex covariance matrix C is obviously fixed at C_n^i; the algorithm also computes for each entry the correlation matrix $E_k := \mathrm{corr}\,(\vec{x}, \vec{q}_k)$ between the vertex position and the particle momentum

$$E_k = -F_k C, \tag{3.12}$$

and the particle momentum covariance matrix

$$D_k = V_{q_k} - V_{rq_k} V_{rk}^{-1} V_{rq_k}^T + F_k C F_k^{-1}, \tag{3.13}$$

with

$$F_k = -V_{rq_k} V_{rk}^{-1} \tag{3.14}$$

being an auxiliary matrix.

Mother particle creation

Assuming the found vertex is inside the LHCb fiducial volume, the fit is validated and a χ^2 is determined by taking the last step value from (3.7) and adding the χ^2 from any short-lived daughter particle. Degrees of freedom (DOFs) for χ^2 reduction are computed as follows:

- each track contributes 2 DOFs;

- each ρ^+-like particle[17] contributes 2 DOFs;

- each sub-vertex contributes 3 DOFs plus further DOFs from the downstream decay tree;

- the sum total is reduced by 3.

A mother particle is subsequently created using the (3.1) representation with reference point $\vec{x}_{\mathrm{mother}}$ fixed to the new-found vertex coordinates. Its 4-momentum is computed as a simple sum of the 4-momenta of its daughters extrapolated at the vertex:

$$\vec{q}_{\mathrm{mother}} = \sum_{k \in \mathrm{daughters}} \vec{q}_k. \tag{3.15}$$

The parameter vector covariance matrix (3.2) is determined as follows:

$$V_r^{\mathrm{mother}} = C, \tag{3.16}$$

---

[17]A ρ^+-like particle is a particle resulting from the combination of 1 long-lived particle and ≥ 2 photons. The category identifier is owed to the topology of the $\rho^+ \to \pi^+ \pi^0$ decay with $\pi^0 \to \gamma\gamma$.

$$V_q^{\text{mother}} = \sum_{k \in \text{daughters}} \left[D_k + \sum_{\substack{j \in \text{daughters} \\ j \neq k}} \left(F_k C F_j^T + F_j C F_k^T \right) \right], \qquad (3.17)$$

$$V_{rq}^{\text{mother}} = \sum_{k \in \text{daughters}} E_k, \qquad (3.18)$$

with D_k, E_k and F_k for each daughter resulting from (3.13), (3.12) and (3.14) respectively.

Finally, the mother particle measured mass M_{mother} is computed as magnitude of 4-vector $\vec{q}_{\text{mother}}$ in the $(-,-,-,+)$ metric

$$M_{\text{mother}} = \sqrt{E_{\text{mother}}^2 - p_{x\,\text{mother}}^2 - p_{y\,\text{mother}}^2 - p_{z\,\text{mother}}^2}, \qquad (3.19)$$

Its associated uncertainty is defined as

$$\sigma_M^{\text{mother}} = \sqrt{\frac{1}{4 M_{\text{mother}}^2} v^T H v}, \qquad (3.20)$$

with

$$v := \frac{dM_{\text{mother}}^2}{d\vec{q}} = \begin{pmatrix} -2 p_{x\,\text{mother}} \\ -2 p_{y\,\text{mother}} \\ -2 p_{z\,\text{mother}} \\ 2 E_{\text{mother}} \end{pmatrix} \qquad (3.21)$$

and

$$H := \sum_{k \in \text{daughters}} V_{q_k}. \qquad (3.22)$$

3.1.2 Decay Tree Fitter algorithm

While the leaf-by-leaf approach adopted by the Vertex Fitter is fast, it brings alongside it the significant drawback of forgoing upstream information when fitting the downstream branches of a decay. This is especially notable for decays like $K_S^0 \to \pi^0 \pi^0 \to \gamma\gamma\gamma\gamma$, where the final state has no tracks to form a vertex with. Even in the relatively more traditional case of the $\Lambda_b^0 \to J/\psi\, \Lambda^0$ decay, however, the VF algorithm still limits our options. In particular, it prevents the placing of *mass constraints* on mother particles, where the fit fixes the invariant mass of the $p\pi^-$ pair to the PDG value for $m(\Lambda^0)$, for instance.

To combat this problem, all reconstructed events in this analysis undergo a refit process based on the Decay Tree Fitter (DTF) algorithm [74] first developed in BaBar. This algorithm takes the entire decay chain as input and
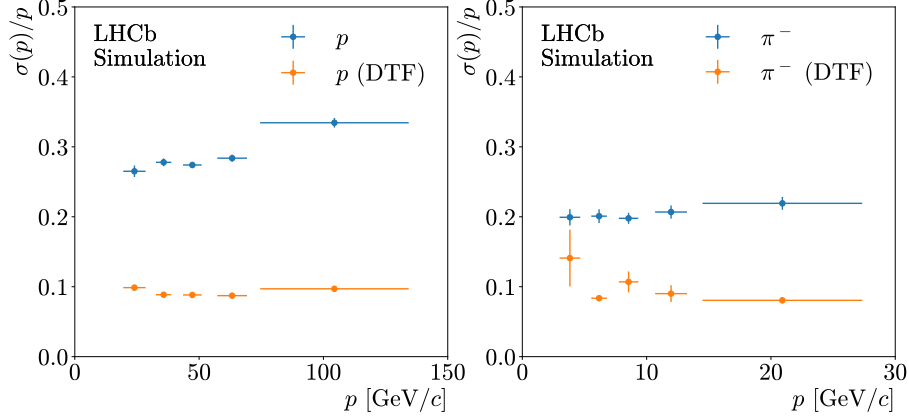
Figure 3.1: Momentum resolution.

allows to place mass constraints on $p\pi^-$ and $\mu^+\mu^-$ invariant masses to match $m(\Lambda^0)$ and $m(J/\psi)$ respectively.

While this step introduces another filtering process and related efficiency to account for[18], it proves invaluable for our physics motivations as it mitigates the most problematic drawback of T track usage, momentum resolution. As shown in Figure 3.1, using both J/ψ and Λ^0 mass constraints improves $\vec{p}$ resolution from 20–30%[19] to $\approx 10\%$.

3.2 Reconstruction efficiency of the Λ_b^0 and Λ^0 decays

To compute the vertex reconstruction efficiency for the Λ_b^0 decay chain, it is useful to conceptualize our event selection as a five step process:

1. reconstruction of associated tracks for all charged daughter particles;

2. reconstruction of the three decay vertices (Λ^0, J/ψ and Λ_b^0);

3. preliminary selections based on kinematic variables to filter out most background (see Section 4.1);

---

[18]DTF convergence efficiency with the double mass constraint is relatively uneven across the $z_\Lambda^{\mathbf{vtx}}$ spectrum, starting at $\approx 50\%$ for $z_\Lambda^{\mathbf{vtx}} \approx 6\,\mathrm{m}$ and growing up to $\approx 85\%$ for $z_\Lambda^{\mathbf{vtx}} \approx 7.5\,\mathrm{m}$. Comparatively, vertex reconstruction is still the primary cause of event loss in this analysis.

[19]Pion momentum resolution is higher because the pion receives a smaller fraction of the Λ^0 momentum, thereby having a larger bending curve than the proton and allowing for better momentum measurement at the T stations.

Figure 3.2: Reconstruction efficiency of simulated $\Lambda_b^0 \to J/\psi \ (\to \mu^+\mu^-) \ \Lambda^0 \ (\to p\pi^-)$ events as function of the z component of the true Λ^0 decay vertex. Given that $J/\psi \to \mu^+\mu^-$ events are reconstructed as part of the trigger step, the low efficiency is attributed to failure in reconstructing Λ^0 and Λ_b^0 decay vertices.

4. Decay Tree Fitter refit with appropriate mass constraints for the analysis at hand (usually J/ψ and Λ^0);

5. further selections applied to events passing all previous steps. Detailed in Chapter 4, these include a physical background veto and signal selection via a trained multivariate classifier.

For the purposes of this section, we are interested in the first two steps (track and vertex reconstruction).

Efficiencies are computed with respect to *reconstructible* particles, a flag attributed during the simulation process based on the number of *hits* (charged clusters with defined positions) in specific modules of the LHCb detector. A track is said to be reconstructible as VELO track with hits in ≥ 3 VELO modules, while it's reconstructible as T track with ≥ 1 hits in both the x and stereo layers of each T station. If these conditions are satisfied simultaneously, the track qualifies for reconstructibility as Long track [75].

At Monte Carlo level, a track is deemed to be *reconstructed* if it can be successfully matched to at least one MC particle; for T and Long tracks, this is true if at least 70% of the hits from the respective relevant detectors for reconstructibility are shared between reconstructed and true track. For Λ_b^0 events with a true $z_{\text{vtx}}^\Lambda \in [6.0\,\text{m}, 7.6\,\text{m}]$, this results in a track reconstruction efficiency in the 60% to 80% range.

When considering how many of these reconstructed charged particles pass the vertex reconstruction (*vertexing*) process, the computed efficiency is much lower. Figure 3.2 plots the resulting Λ_b^0 vertexing efficiency through the whole

true $z_{\mathrm{vtx}}^{\Lambda}$ spectrum, showing that said efficiency never manages to get past the 50% threshold. This means that over half of our candidate signal events is lost during the second step of the five step selection process.

While available information does not distinguish between the three individual vertexing phases (J/ψ, Λ^0 and Λ_b^0), we can make some reasonable assumptions. Being Long tracks, muons and antimuons have well reconstructed momentum with constraints across the LHCb detector; for this reason their influence on the vertexing efficiency dip is considered negligible. Furthermore, the rare usage of T tracks for physics analysis in LHCb suggests that problems are likelier to arise in the $\Lambda^0 \to p\pi^-$ vertexing and then cascade into the $\Lambda_b^0 \to J/\psi\,\Lambda^0$ reconstruction.

For the above reasons, in the following sections I'll focus on the $\Lambda^0 \to p\pi^-$ decay to search for issues and solutions, with the goal of improving signal yield.

3.3 Characterization of non-converged events

[@todo]

The full simulated $\Lambda_b^0 \to J/\psi\,\Lambda^0$ dataset constructed for studies on the measurement of the Λ^0 electromagnetic dipole moments omits a lot of technical information on the decays, the retention of which would make storage and quick access impractical. Furthermore, a lot of the work in this and the next sections required direct changes to the vertexing algorithm for debugging and event recovery, which in turn required data to be reprocessed. For these reasons, the studies in the following sections were conducted on a smaller sample of [@todo:how_many] Λ_b^0 decays.

3.3.1 Behaviour through VF iterations

The VF process reaches convergence if either condition (3.9) or (3.10) is satisfied, i.e. if the vertex position estimated at iteration i and the one from iteration $i - 1$ are «close enough» either in absolute distance or χ^2 distance, up until $i_{\mathrm{max}} = 10$. This is predicated on the principle that the algorithm refines its vertex estimate after each iteration, homing in on the candidate vertex with the lowest $\tilde{\chi}_{\mathrm{vtx}}^2$.

Such a behaviour is not found in non-converging (henceforth also known as *failed*) events. This can be seen by increasing $i_{\mathrm{max}} = 100$, which causes a negligible $\approx 2\%$ increase in converged events. It follows that, for the vast majority of missing events, failure of convergence is not a product of low computation time and must instead result from some internal malfunction of the vertexing process.
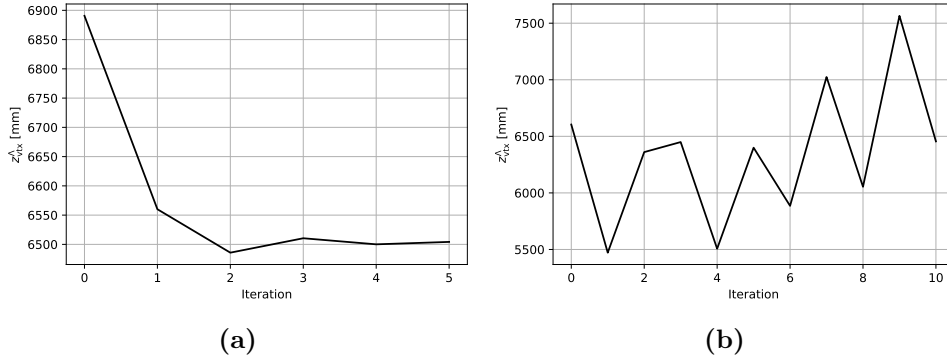
(a)

(b)

Figure 3.3: Best values of tentative $\Lambda^0 \to p\pi^-$ vertex z component at the end of each Vertex Fitter iteration for examples of converged *(a)* and failed *(b)* events.

This is readily apparent when studying the vertex positions throughout the iterating process for examples of converged and failed events of simulated signal. Figure 3.3 compares the values of $z_{\mathrm{vtx}}^{\Lambda}$, the z component of the $\Lambda^0 \to p\pi^-$ decay vertex, as reconstructed by the VF in iterations $i = 0$ to 10 ($i = 0$ being the starting seed). Figure 3.3a (converged) exhibits the expected behaviour, with the algorithm refining its vertex estimate after every iteration and finally converging as early as $i = 5$. By contrast, Figure 3.3b (failed) presents an *oscillating behaviour* of $z_{\mathrm{vtx}}^{\Lambda}$, swinging by as much as $1\,\mathrm{m}$ in the opposite direction after an iteration completes. While a particularly tricky instance of the first type of event may potentially benefit by an increased $i_{\max}$, no amount of allotted computations can lead the second type to convergence.

We can gain more insight into the nature of this oscillation by approaching the problem from a more «geometrical» point of view. Figure 3.4 shows the topology of the converged event from Figure 3.3a in the bending xz and non-bending yz planes. While the points of closest distance between the tracks in the two planes are not exactly matched, the algorithm manages to converge on a vertex position that is reasonably close to the true position from the generated decay.

The behaviour for the non-converged event from Figure 3.3b, shown in Figure 3.5, is drastically different. The dotted line following the best vertex estimate through iterations in the bending plane (Figure 3.5a) shows that the algorithm gravitates *around* the point of closest distance between proton and pion tracks ($z \approx 6.25\,\mathrm{m}$), never outright choosing it as candidate. On the other hand, in the yz plane (Figure 3.5a) the two tracks cross at a much more downstream point ($z \approx 7.5\,\mathrm{m}$), which the Vertex Fitter also largely ignores during the iteration process.

It may be tempting to attribute the failed convergence in this event to the comparatively larger gap between proton and pion tracks at their point of
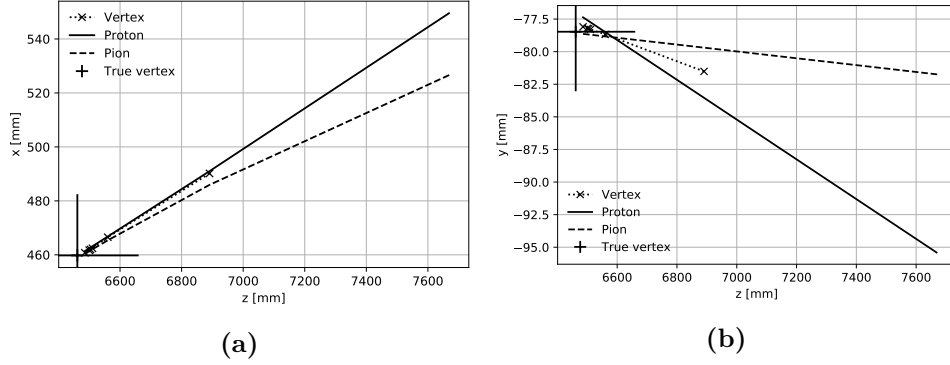
(a)   (b)

Figure 3.4: $\Lambda^0 \to p\pi^-$ decay topology for an instance of converged event, projected in the *xz (a)* and *yz (b)* planes. The temporary best vertex locations chosen by the Vertex Fitter at the end of each iterations are marked with *diagonal crosses* and linked by the *dotted line* in iteration order (the second iteration best vertex is connected to the first, the third to the second and so on). The reconstructed vertex, while not explicitly represented, can be identified by the cluster of vertex point markers. At the beginning of each iteration, daughter particles are extrapolated at the z of the best vertex of the previous iteration: proton (*solid line*) and pion (*dashed line*) tracks are drawn joining the respective reference points after transportation in order of decreasing z (i.e. not in iteration order). The true vertex is marked by the *large Greek cross.*

closest distance in the bending plane. However, the VF algorithm has shown to be capable of bridging an imperfect track extrapolation: the converged event shown in Figure 3.6 has a closest track distance of some $40\,\text{cm}$, an order of magnitude greater than the track distance seen in Figure 3.5. Despite the large gap, the VF is able to recognize $z \approx 3\,\text{m}$ as the best candidate for the vertex, even if said point is far removed from the actual $\Lambda^0 \to p\pi^-$ decay vertex.

a

We can make a more convincing remark by analyzing performance in the yz plane, where tracks *don't* bend. Figure **??** shows that, in the converged case, the yz track crossing is z-aligned with the closest xz distance point. This doesn't happen for the failed event: while Figure **??** shows that yz tracks cross almost coinciding with the true vertex position, we have already pointed out that this is $50\,\text{cm}$ short of the xz closest distance point. Convergence failure for the event in Figures **??** and **??** can thus be interpreted through the lens of *conflicting information*: the best vertex candidate has different z_{vtx} in the xz (with magnetic field) and yz (without magnetic field) planes, and the VF algorithm flip-flops between the two.

For didactic purpose, the analysis in this section has focused on just one event. All the emerged patterns are however commonplace throughout the
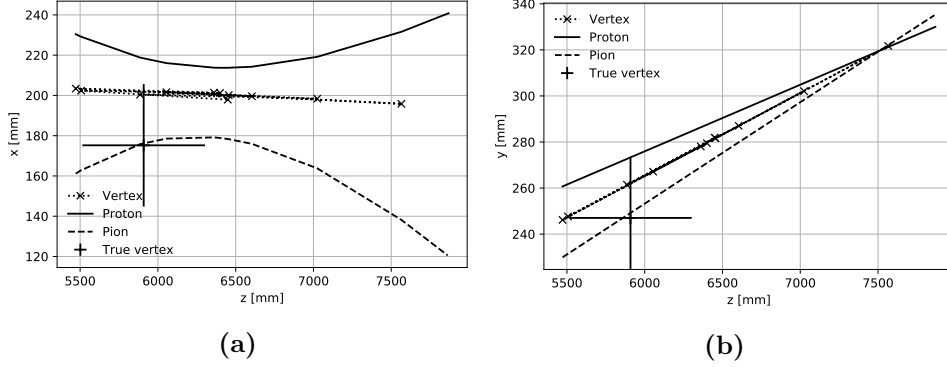
Figure 3.5: $\Lambda^0 \to p\pi^-$ decay topology for an instance of non-converged event, projected in the xz (a) and yz (b) planes. Notation follows the one in Figure 3.4.
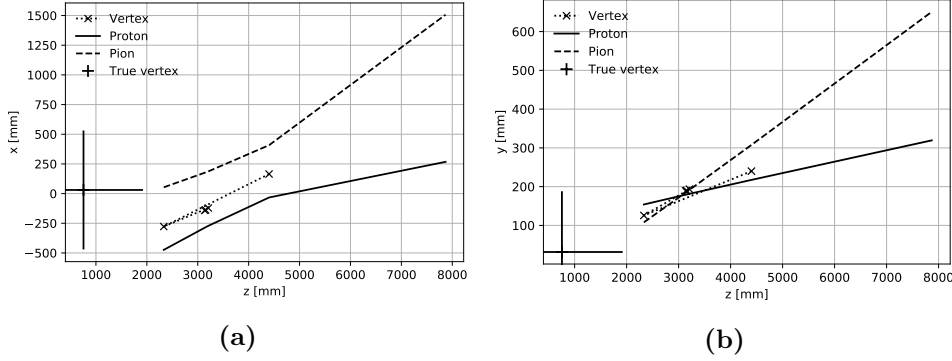

Figure 3.6: $\Lambda^0 \to p\pi^-$ decay topology for an instance of converged event with large track gap in the xz plane, projected in the xz (a) and yz (b) planes. Notation follows the one in Figure 3.4.

failed events I have examined, with the oscillating vertex behaviour in particular being a constant in almost all of them. While every Λ_b^0 vertexing failure being the fault of xz and yz track mismatch would be a reckless conclusion, I have been able to use these findings, along with other from the following paragraphs, to devise a partial solution in Section 3.4.

3.3.2 Decay kinematics before and after interaction with the detector

To gather information on the possible sources of the oscillating behaviour outlined in Section 3.3.1, I investigated the collective properties of a sample of $\Lambda_b^0 \to J/\psi\,\Lambda^0$ decays in search of significant discrepancies between converged and failed events.
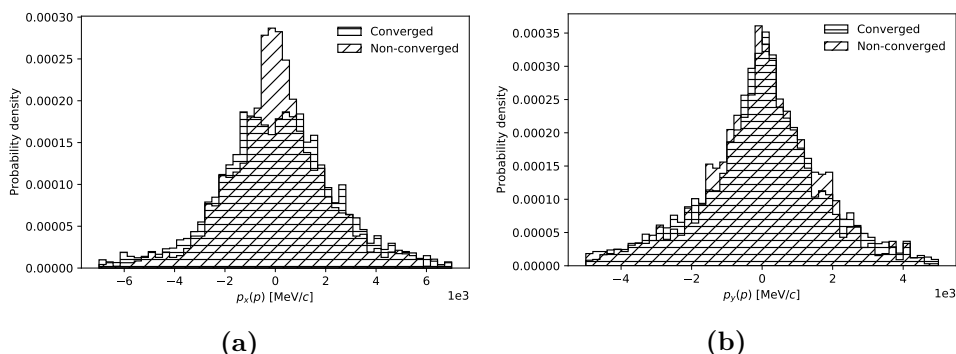
Figure 3.7: Normalized distributions of proton protoparticles p_x *(a)* and p_y *(b)* components. Distributions of events with converging vertex are depicted with *horizontal hatching*, non-converging with *diagonal hatching*.

No difference between the two categories emerged at Monte Carlo level when considering the true values of basic kinematic variables such as daughter particle momenta or decay vertices locations of the three unstable particles Λ_b^0, Λ^0 and J/ψ. Moreover, there doesn't seem to be a critical decay geometry that triggers the vertexing issues; for instance, there is no evidence that $\Lambda^0 \to p\pi^-$ decays lying largely in the xz plane, a setup quite unfriendly to the VF algorithm (see Section 4.1.1), have any disproportionate representation amongst non-converged events. This rules out the hypothesis that the failure of vertex fit convergence be caused by particle characteristics at production, such as protons and pions with unusually low transverse momentum p_T.

I also considered the possibility that kinematic asymmetries between converged and failed events may arise after interaction with the LHCb detector; one example of this would be the case where failed events correlate with hits in a specific sections of the IT or OT trackers. To this end I studied the event *protoparticles*, data structures created during the LHCb event reconstruction process with the intent of encapsulating all relevant information available for the associated particle: particle identification from the RICH and muon detectors, results from calorimeter hits and track information. Tracks themselves contain details on momentum in relation to a certain reference point which, for stable charged particles, corresponds to the position of first measurement.

Figure 3.7a depicts the only major discrepancy between converged and failed events I found in this process: non-converged Λ_b^0 decays have a proton protoparticle p_x distribution concentrated in $p_x \approx 0$, while converged decays show even dispersion of the central peak in the $[-1\,\mathrm{GeV}, 1\,\mathrm{GeV}]$ region. This is to be contrasted with p_y, the other component of the protoparticle transverse momentum, whose two distributions are almost overlapping (Figure 3.7b). While statistics in the reduced samples are too low to say for sure,

pions do not appear to show similar discrepancies in their protoparticles.

Assuming p_x measurement at detector level is not biased, this result suggests that protons from failed events have generally lower p_x when they reach the T stations; this would not be apparent from p_T distributions at the point of production, since the final p_x value depends on the bending radius imprinted by the magnet (which in turn depends on momentum and distance traveled). Such low values of p_x are potentially more susceptible to poor measurement by the T stations, which in turn affects the overall momentum resolution based on the particle bending radius in the xz plane.

3.3.3   Decay kinematics after extrapolation

As will be later discussed in Sections 3.4 and 4.1.1, reconstruction of the Λ^0 vertex is affected by a significant positive bias of median value $\mu_{\frac{1}{2}}(z_\Lambda^{\text{reco}} - z_\Lambda^{\text{true}}) \approx$ 43 cm. In spite of such a discrepancy, the standard modus operandi for kinematics-at-vertex analyses usually compares the true momenta (evaluated at the true vertex) with the reconstructed momenta (evaluated at the reconstructed vertex).

For this section I have followed a different approach, instead opting to transport via Runge-Kutta extrapolator the reconstructed p and π^- at the true $\Lambda^0 \to p\pi^-$ vertex position, injected from Monte Carlo information. Since the extrapolator takes raw protoparticles as inputs, this process bypasses any smoothing applied during the fit process and, given an observable f (particle reference points and momenta, for instance), it allows for a comparison between the true value f_{true} and the RK-extrapolated value f_{RK} at the actual z_{vtx}^Λ, circumventing the effect of vertex bias. Any potential mismatch will be normalized in terms of *residuals*

$$\text{res}(f) := \frac{f_{\text{RK}} - f_{\text{true}}}{\sigma_f^{\text{RK}}}, \tag{3.23}$$

with σ_f^{RK} being the uncertainty computed by the RK algorithm. Assuming a correct estimation of such uncertainties, we expect the residuals to follow a standard normal distribution.

Figure 3.8 shows p_z residuals for proton and pions extrapolated at the Λ^0 true decay vertex, juxtaposing converged and non-converged event distributions. The first takeaway is that VF-reconstructed events have a remarkably non-gaussian distribution, with a slightly positive mean and asymmetric tails. This is particularly apparent in the case of the proton, where for $p_z^{\text{RK}} < p_z^{\text{true}}$ the RK algorithm systematically underestimates the error. More relevant for this chapter is Figure 3.8b specifically, which highlights the fact that pion tracks from non-converged events generally sport a strong negative bias on p_z.
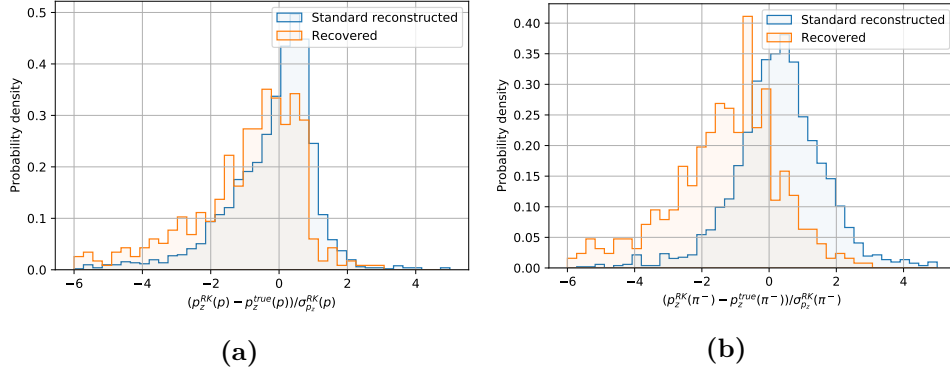
(a) (b)

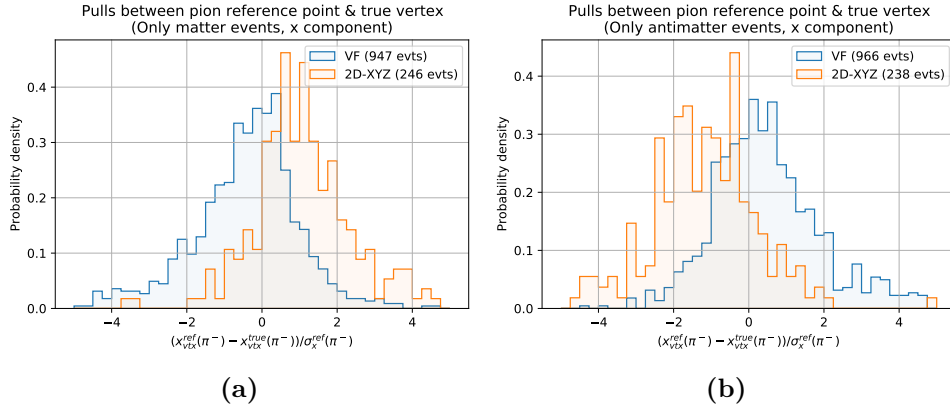**Figure 3.8:** Left right

(a) (b)

Figure 3.9: Left right

This behaviour can also be observed from a different perspective by considering the pion reference point x residuals[20], plotted in Figure 3.9. For this part of the analysis, we forgo the established omission of charge-conjugate notation and separate events in matter ($\Lambda^0 \to p\pi^-$) and antimatter ($\bar{\Lambda}^0 \to \bar{p}\pi^+$) events. Since all events in this sample have magnet polarity $M_{\text{pol}} = +1$, i.e. magnetic field flowing towards the positive direction of the y axis, matter and antimatter daughter particles will bend in opposite directions due to charge inversion.

Comparing Figures 3.9a and 3.9b, x components of pion reference points in non-converged events show opposite bias when extrapolated at the Λ^0 vertex. Such a behaviour is easily understood in terms of transportation: tracks are extrapolated from downstream protoparticles towards decreasing z; if the mag-

---

[20]The Runge-Kutta extrapolator transports a particle up to a specified z coordinate. As a consequence, x and y reference points gauge the discrepancy between track extrapolation and the true position of the MC particle at said z, with $z_{\text{ref}}^{\text{RK}} \approx z_{\text{ref}}^{\text{true}}$ within extrapolator tolerance.

netic bending effect in the zx plane is overstated (if tracks «bend too much»), the resulting x reference point at true $z_{\mathrm{vtx}}^{\Lambda}$ will have an offset with respect to the actual origin vertex of the particle with polarity-depdendent sign (cf. Figure **??**).

Excessive bending is exactly what is expected of a track with correctly estimated p_x and p_y, but underestimated p_z. Furthermore, a lower p_z will affect the closest distance point between proton and pions in the zx plane, but will have a much lower impact on the track crossing point in the yz plane, supporting the conflicting information interpretation for vertex oscillation given in Section 3.3.1. Considering the information gathered so far, the two most plausible causes of failed convergence appear to be wrong extrapolation by the Runge-Kutta algorithm, perhaps triggered by lower protoparticle p_x observed in Section 3.3.2, and/or poor measurement from T stations resulting in p_z underestimation of protons and pions.

3.4 Recovery of non-converged events through refit with rescaled uncertainties

As outlined in Section 3.3.1, vertexing failures in the $\Lambda^0 \rightarrow p\pi^-$ decay can be attributed to candidate vertices in different planes providing conflicting information. In order to circumvent this phenomenon, my proposal for the recovery of these failed events involves performing a refit process with a slightly altered version of the standard Vertex Fitter, designed to give more importance to tracks lying on a specific plane.

The obvious choice for said plane is the yz plane, since extrapolation of these tracks does not need to be concerned with magnet bending and is therefore expected to be less prone to error. However, this would penalize events with poor yz protoparticle reconstruction, for instance events with parallel or diverging tracks in said plane. To maximize the recovery efficiency of my solution, I have elected to perform three separate refits on non-converging events, prioritizing $yz \rightarrow xz \rightarrow xy$ planes in this order. In a worst-case scenario, this would quadruple the time complexity of the vertexing process; in practice, half of all events converge with the standard VF, and about $\approx 15\%$ more converge after the first refit attempt (yz plane).

Considering the yz plane as an example, we can prioritize available information for tracks in this plane by artificially increasing the uncertainty σ_x for the x component of the candidate vertex position $\vec{x}$. At each step in a VF iteration, $\vec{x}$ is updated according to (3.6). Uncertainties enter the computation through three terms:

1. C_{k-1}^{-1}, inverse $\vec{x}$ covariance matrix computed at the previous step (or

previous iteration);

2. $V_{r_k}^{-1}$, inverse covariance matrix of reference point $\vec{r}_k$, computed at the true transport of particle k;

3. C_k, current $\vec{x}$ covariance matrix, inverted from the matrix sum of the previous two terms as in (3.5).

Ergo, the best approach to increase σ_x while minimizing additional computation time is to act on the individual components of C_{k-1}^{-1} and $V_{r_k}^{-1}$ just before the (3.5) sum.

Assuming gaussian uncertainties, a standard three-dimensional covariance matrix will have the form

$$C = \begin{pmatrix} \sigma_x^2 & \rho_{xy}\sigma_x\sigma_y & \rho_{xz}\sigma_x\sigma_z \\ \rho_{xy}\sigma_x\sigma_y & \sigma_y^2 & \rho_{yz}\sigma_y\sigma_z \\ \rho_{xz}\sigma_x\sigma_z & \rho_{yz}\sigma_y\sigma_z & \sigma_z^2 \end{pmatrix}, \tag{3.24}$$

where $\rho_{ij} := \mathrm{corr}(i,j)$. Its inverse matrix is written as

$$C^{-1} = \frac{1}{K} \begin{pmatrix} \frac{1-\rho_{yz}^2}{\sigma_x^2} & \frac{\rho_{xz}\rho_{yz}-\rho_{xy}}{\sigma_x\sigma_y} & \frac{\rho_{xy}\rho_{yz}-\rho_{xz}}{\sigma_x\sigma_z} \\ \frac{\rho_{xz}\rho_{yz}-\rho_{xy}}{\sigma_x\sigma_y} & \frac{1-\rho_{xz}^2}{\sigma_y^2} & \frac{\rho_{xy}\rho_{xz}-\rho_{yz}}{\sigma_y\sigma_z} \\ \frac{\rho_{xy}\rho_{yz}-\rho_{xz}}{\sigma_x\sigma_z} & \frac{\rho_{xy}\rho_{xz}-\rho_{yz}}{\sigma_y\sigma_z} & \frac{1-\rho_{xy}^2}{\sigma_z^2} \end{pmatrix}, \tag{3.25}$$

with

$$K := \frac{\det C}{\sigma_x^2\sigma_y^2\sigma_z^2} = 1 + 2\rho_{xy}\rho_{xz}\rho_{yz} - \rho_{xy}^2 - \rho_{xz}^2 - \rho_{yz}^2. \tag{3.26}$$

Going back to the yz plane example, we increase σ_x by introducing a multiplicative factor $s_x < 1$ for relevant covariance matrix components as follows:

$$\begin{aligned} C_{xx}^{-1} &= C_{xx}^{-1} \times s_x^2, \\ C_{xy}^{-1} &= C_{yx}^{-1} = C_{xy}^{-1} \times s_x, \\ C_{xz}^{-1} &= C_{zx}^{-1} = C_{xz}^{-1} \times s_x, \end{aligned} \tag{3.27}$$

with other components left as is. This process is also applied to $V_{r_k}^{-1}$ and replicated at each step of the refit algorithm, which I'll refer to as σ_x-*rescaled*. Similarly, the σ_y-rescaled and σ_z-rescaled algorithms prioritize planes xz and xy respectively; their extension from (3.27) should be trivial. For the remainder of this section, I'll also refer to their sequential application $\sigma_x \to \sigma_y \to \sigma_z$ as the σ-*rescaled refit process*, unless otherwise stated.

As proof of concept, I have analyzed the performance of the σ-rescaled refit approach with $s_i = 0.98, i \in \{x, y, z\}$ (corresponding to $\approx 2\%$ increase in vertex

Increased σ	Recovery eff.	$\mu_{\frac{1}{2}}\left[\tilde{\chi}^2_{\text{vtx}}\right]$	$\mu_{\frac{1}{2}}\left[z^\Lambda_{\text{vtx}}\text{ bias}\right]$	$\mu_{\frac{1}{2}}\left[p_z^{\text{DTF}}(p)\text{ bias}\right]$
None (VF)	–	1.0	429 mm	$+1.35\%$
σ_x	62%	4.9	584 mm	-0.84%
σ_y	74%	5.4	635 mm	-0.81%
σ_z	80%	7.8	697 mm	-1.02%
Sequential	100%	6.5	646 mm	-0.93%

Table 3.1: Performance comparison of the three rescaled-σ algorithms with $s_i = 0.98$ (see text for details), contrasted with the performance of their sequential application ($\sigma_x \to \sigma_y \to \sigma_z$) and the standard Vertex Fitter algorithm. Recovery efficiency is defined as the ratio of events reconstructed by a certain flavour over the total number of events recoverable by combining the three algorithms. $\mu_{\frac{1}{2}}$ identifies the median value. Performances for xyz algorithms are computed using all events recovered by the individual flavour, including events recovered by more than one, while values for VF are computed on standard reconstructed events.

uncertainties) on a sample of [@todo:how many] MC-generated $\Lambda_b^0 \to J/\psi\,\Lambda^0$ events, observing a $+25\%$ increase in reconstructed events. As per Figure 3.2, this amounts to about a quarter of all reconstructible events.

I have also run each algorithm individually after the standard VF to gauge their performance on the $n_{\text{tot}}^{\text{reco}}$ recovered events. Results of this test are compiled in Table 3.1. When considering the *recovery efficiency* of each algorithm, defined as the fraction of recovered events converging under said algorithm, i.e.

$$\varepsilon_i^{\text{reco}}\big|_{i\in\{x,y,z\}} = \frac{n_i^{\text{reco}}}{n_{\text{tot}}^{\text{reco}}}, \tag{3.28}$$

the σ_z-rescaled is the better performing one, reaching convergence in $\approx 80\%$ of recoverable events. Despite this, all three σ-rescaled algorithms have a sizable number of «exclusive» events that do not reach convergence under any other variation. Furthermore, while the σ_z-rescaled does recover more events by itself, only $\approx 5\%$ of the total recovered events are exclusive to it, meaning its overall impact on the σ-rescaled refit process is low.

The established $\sigma_x \to \sigma_y \to \sigma_z$ refit order is justified in light of performance evaluation based on the goodness of these fits. As seen again in Table 3.1, the σ_x-rescaled algorithm has comparatively lower vertex $\tilde{\chi}^2$, z^Λ_{vtx} bias and proton p_z bias using the DTF algorithm with J/ψ and Λ^0 mass constraints. Using the σ_z-rescaled algorithm as last resort means only $\approx 5\%$ of recovered events are actually affected by its poor performance.
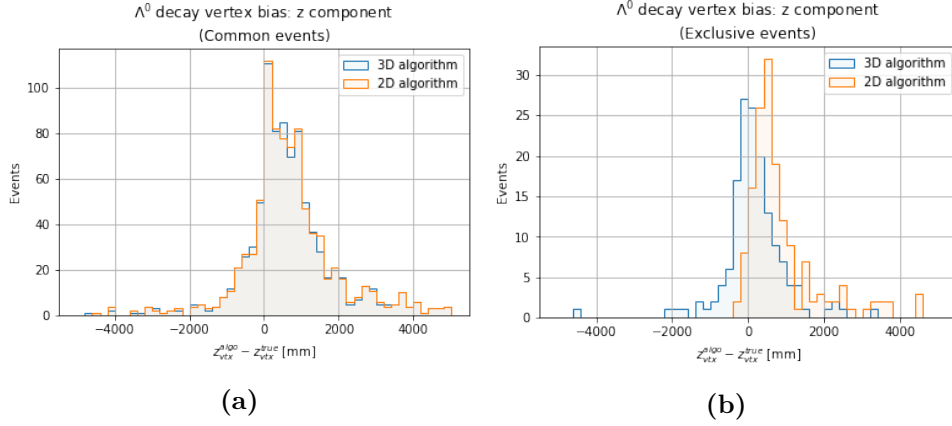
Figure 3.10: Left right

All σ-rescaled algorithms still appear to perform significantly worse than the standard VF on all the above fronts. To investigate this, I have individually run the σ_x-rescaled and standard VF algorithms on all available events and compared their results. Considering the total number of events reconstructed by combining the two, $\approx 74\%$ of them are reconstructed by both, $\approx 13\%$ are only reconstructed by the σ_x-rescaled (these are the σ_x-recovered events from Table 3.1) and a roughly equal $\approx 13\%$ of them are only reconstructed by the Vertex Fitter.

What's more interesting is the performance comparison. Figure 3.10a shows the z_{vtx}^{Λ} bias on events common to VF and σ_x-rescaled. Despite the median 15 cm discrepancy reported in Table 3.1, the two distributions are indistinguishable from the other. The expected difference only emerges when considering events exclusive to each algorithm, as seen in Figure 3.10b.

The differences in z_{vtx}^{Λ} and proton p_z^{DTF} bias are therefore to be understood in terms of what events reach convergence with a specific algorithm, rather than as performance evaluations of the algorithm itself. The σ_x-rescaled algorithm does not intrinsically reconstruct events with larger bias; instead, its yz-centric approach allows it to recover events that throw off the standard VF, and these events tend to have a larger z_{vtx}^{Λ} bias on their best vertex candidate. This tendency is likely the result of the systematic T track p_z underestimation analyzed in Section 3.3.2.

Overall, my σ-rescaled refit process proposal allows for the recovery of an extra $+25\%$ of signal events. While $\tilde{\chi}_{\text{vtx}}^2$ and bias on z_{vtx}^{Λ} are higher compared to events reconstructed via Vertex Fitter (see Figure 3.11 and Table 3.1), there is sufficient evidence pointing towards this being a problem intrinsic to the non-converged events themselves. Their impact on the prospective Λ^0 EDM/MDM measurement will have to be evaluated in dedicated sensitivity
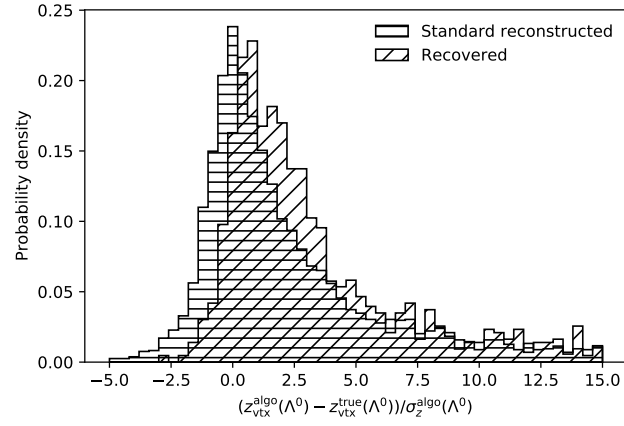
Figure 3.11: A.

studies and possibly incorporated as a source of systematic uncertainty to be accounted for.

Chapter 4

Signal event selection

This chapter outlines the three major steps in the selection of $\Lambda_b^0 \to J/\psi\,\Lambda^0$ events for the study of long-lived Λ^0 baryons in LHCb. Section 4.1 discusses the preliminary kinematic variable selections and includes an in-depth study of evidences of bias in the $\Lambda^0 \to p\pi^-$ vertex reconstruction, particularly its z component; Section 4.2 provides an overview of the veto approach used to account for the main source of physical background, the $\Lambda_b^0 \to J/\psi\,\Lambda^0$ decay; Section 4.3 details my work towards the training, optimization and evaluation of a supervised learning multivariate classifier to perform signal-background discrimination; finally, Section 4.4 brings together all three selection techniques to give proof of $\Lambda_b^0 \to J/\psi\,\Lambda^0$ reconstructibility, fitting the filtered $J/\psi\,\Lambda^0$ invariant mass distributions from simulated signal and Run 2 data.

4.1 Prefiltering

Prefilters, also refered to as *preliminary selections* or simply *pre-selections*, are the foundation of the signal selection process. The main objective of this step is to improve the signal-to-background ratio and reduce the computational workload to analyze data with cuts on kinematic variables.

The applied prefilter criteria are listed in Table 4.1. Efficiencies on signal have been estimated with studies on simulated $\Lambda_b^0 \to J/\psi\,\Lambda^0$ events. The most impactful selection is the one applied to z_{vtx}^Λ, the z component of the $\Lambda^0 \to p\pi^-$ decay vertex; the efficiencies for the left and right cuts as a function of the threshold are shown in Figure 4.1. Since I require Λ^0 to decay after the dipole magnet in order to observe spin precession, the $\approx 32\%$ efficiency of the $z_{\mathrm{vtx}}^\Lambda \geq 5.5\,\mathrm{m}$ cut cannot be avoided. Other selections have a much lower impact on signal, with efficiencies $\gtrsim 80\%$, resulting in a total prefilter efficiency of $\approx 26\%$.

As detailed in Section 2.2.1, a key aspect of my analysis is the employment

Chapter 4

Signal event selection

This chapter outlines the three major steps in the selection of $\Lambda_b^0 \to J/\psi\,\Lambda^0$ events for the study of long-lived Λ^0 baryons in LHCb. Section 4.1 discusses the preliminary kinematic variable selections and includes an in-depth study of evidences of bias in the $\Lambda^0 \to p\pi^-$ vertex reconstruction, particularly its z component; Section 4.2 provides an overview of the veto approach used to account for the main source of physical background, the $\Lambda_b^0 \to J/\psi\,\Lambda^0$ decay; Section 4.3 details my work towards the training, optimization and evaluation of a supervised learning multivariate classifier to perform signal-background discrimination; finally, Section 4.4 brings together all three selection techniques to give proof of $\Lambda_b^0 \to J/\psi\,\Lambda^0$ reconstructibility, fitting the filtered $J/\psi\,\Lambda^0$ invariant mass distributions from simulated signal and Run 2 data.

4.1 Prefiltering

Prefilters, also refered to as *preliminary selections* or simply *pre-selections*, are the foundation of the signal selection process. The main objective of this step is to improve the signal-to-background ratio and reduce the computational workload to analyze data with cuts on kinematic variables.

The applied prefilter criteria are listed in Table 4.1. Efficiencies on signal have been estimated with studies on simulated $\Lambda_b^0 \to J/\psi\,\Lambda^0$ events. The most impactful selection is the one applied to z_{vtx}^Λ, the z component of the $\Lambda^0 \to p\pi^-$ decay vertex; the efficiencies for the left and right cuts as a function of the threshold are shown in Figure 4.1. Since I require Λ^0 to decay after the dipole magnet in order to observe spin precession, the $\approx 32\%$ efficiency of the $z_{\mathrm{vtx}}^\Lambda \geq 5.5\,\mathrm{m}$ cut cannot be avoided. Other selections have a much lower impact on signal, with efficiencies $\gtrsim 80\%$, resulting in a total prefilter efficiency of $\approx 26\%$.

As detailed in Section 2.2.1, a key aspect of my analysis is the employment
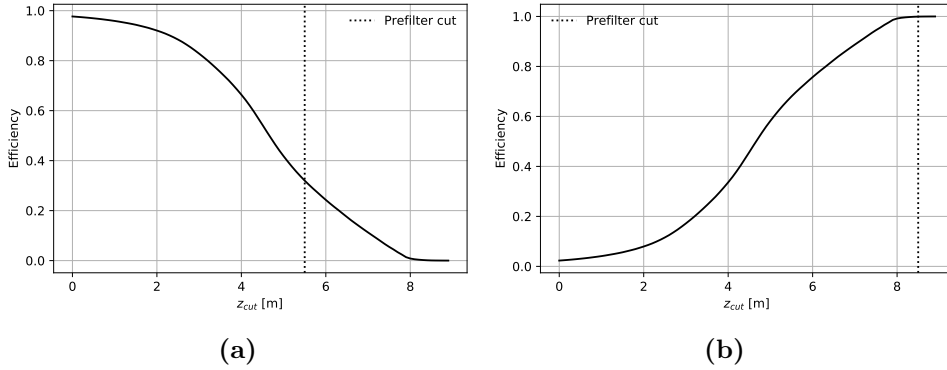
(a)   **(b)**

Figure 4.1: Efficiency of the $z_{\mathrm{vtx}}^{\Lambda} \geq z_{\mathrm{cut}}^{\mathrm{left}}$ *(a)* and $z_{\mathrm{vtx}}^{\Lambda} \leq z_{\mathrm{cut}}^{\mathrm{right}}$ *(b)* prefilter selection criteria on $\Lambda_b^0 \to J/\psi\,\Lambda^0$ simulated signal, as function of the respective thresholds. The *dotted vertical lines* mark the chosen thresholds.

of T tracks for the reconstruction of the Λ^0. The low residual magnetic field for protons and pions produced far from the dipole magnet lowers momentum resolution for the associated tracks down to 20–30%. Resolution can be improved up to $\approx 10\%$ by placing kinematic constraints on $p\pi^-$ and $\mu^+\mu^-$ invariant masses, fixing them to the PDG values of $m(\Lambda^0)$ and $m(J/\psi)$ respectively (these will be henceforth referred to as *mass constraints*). This approach cannot be implemented in the leaf-by-leaf framework of the default Vertex Fitter algorithm for vertex reconstruction; instead each event is refitted with the Decay Tree Fitter (DTF) algorithm in two configurations, single J/ψ and double $J/\psi\,\Lambda^0$ mass constraints (see also Section 3.1.2).

A convergence requirement of the DTF algorithm with the $J/\psi\,\Lambda^0$ mass constraints is therefore added to the prefilter selections. The main drawback of this selection, a steep 45% efficiency on simulated signal events (although raised to 65% when only considering events passing the non-DTF prefilters), is outweighed by the benefits of the improved momentum resolution in the determination of the angular distribution of $\Lambda^0 \to p\pi^-$ decay products.

4.1.1 Reconstruction of Λ^0 decay vertex in prefiltered events

The quality of the $\Lambda^0 \to p\pi^-$ vertex reconstruction affects many aspects of the Λ^0 electromagnetic dipole moments measurement: on top of being fundamental to evaluate how much magnetic field the particle traversed (and thus the extent of spin precession), even the best momentum resolution for protons and pions is worthless if the particles are extrapolated at the wrong point of production. Both $x_{\mathrm{vtx}}^{\Lambda}$ and $y_{\mathrm{vtx}}^{\Lambda}$ are fairly well reconstructed in prefiltered events, with
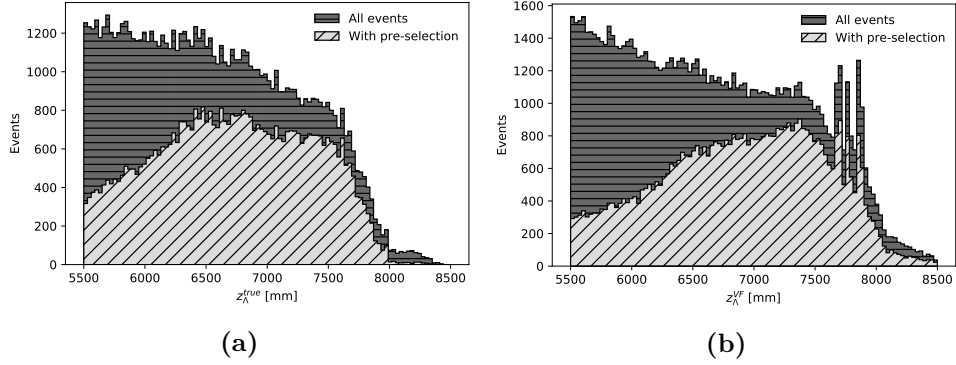
Figure 4.2: Distribution of true *(a)* and reconstructed *(b)* z_{vtx}^{Λ} in simulated $\Lambda_b^0 \to J/\psi\, \Lambda^0$ signal events, without (*dark grey*) and with (*light grey*) prefiltering.

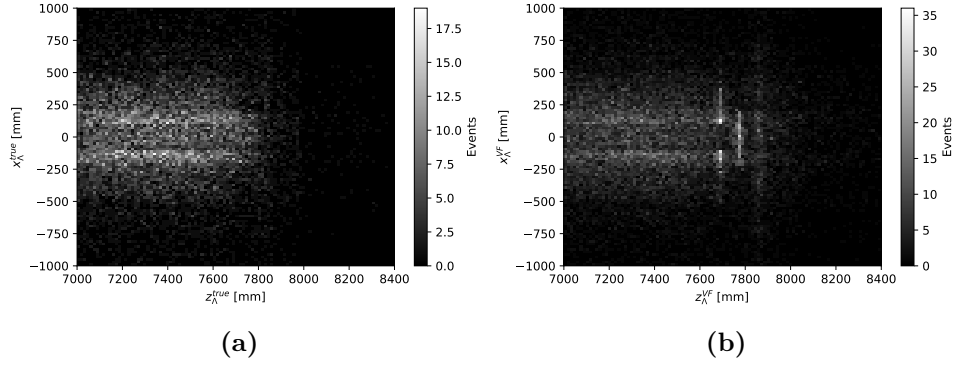**Figure 4.3:** Distribution of simulated $\Lambda_b^0 \to J/\psi\, \Lambda^0$ signal events (prefilters applied) with $z_\Lambda^{VF} \geq 7.0\,\mathrm{m}$, as function of true (*left*) and reconstructed (*right*) x_Λ^{vtx} and z_{vtx}^{Λ}. This corresponds to a top view of true and reconstructed Λ^0 decay vertices.

resolution $\lesssim 1\,\mathrm{cm}$ and no discernible bias. This section will therefore focus on the reconstruction of z_{vtx}^{Λ}.

Figures 4.2a and 4.2b show the distributions of true and reconstructed z_{vtx}^{Λ} respectively for simulated signal events. The most prominent difference between the two is the presence of three peaks in the $[7.5\,\mathrm{m}, 8.0\,\mathrm{m}]$ region of the reconstructed distribution, being found both with and without prefilter selections. The significance of these structures can be inferred by plotting the events as function of z_{vtx}^{Λ} and x_{vtx}^{Λ}, corresponding to a bending plane perspective of the detector. This is shown in Figure 4.3b, highlighting the fact that the peaks in Λ^0 decay vertices have a very precise geometrical location, absent when comparing the true z_{vtx}^{Λ} and x_{vtx}^{Λ} values for the same events (Figure 4.3a). The spatial distribution of the vertices bears a striking resemblance to the layout of a T tracking station (see Figure 2.7a) and z coordinates are consistent with the nominal placement of IT and first OT plane of the T1 station [52].
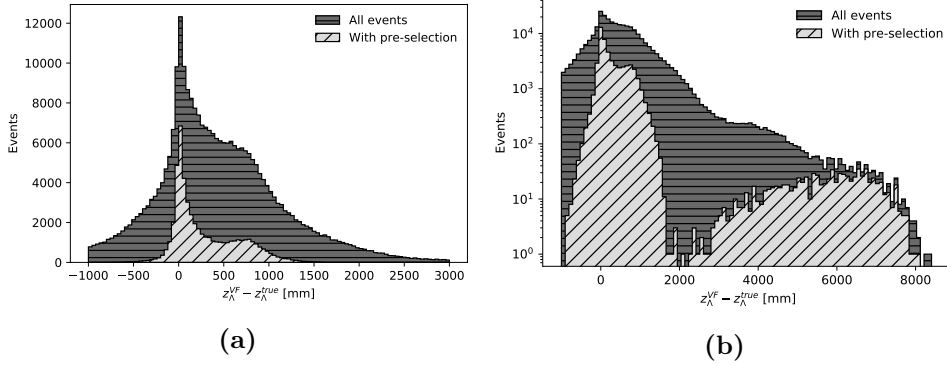
Figure 4.4: Distribution of $z_{\mathrm{vtx}}^{\Lambda}$ bias for simulated $\Lambda_b^0 \to J/\psi\,\Lambda^0$ events in linear *(a)* and logarithmic *(b)* scales, without (*dark grey*) and with (*light grey*) prefiltering.

While dedicated studies are required to gain more insight into the source of these structures, they are assumed to be of minor impact for the purposes of this thesis.

The differing shapes of true and reconstructed $z_{\mathrm{vtx}}^{\Lambda}$ distributions from Figure 4.2 are also evidence of bias effects in the $\Lambda^0 \to p\pi^-$ vertex reconstruction. This is confirmed in Figure 4.4a, showing the distribution of $z_{\mathrm{vtx}}^{\Lambda}$ bias for simulated signal events: the shape is distinctly non-gaussian, with a long tail towards the positive end of the axis counterbalancing the expected ≈ 0 peak. The resulting median bias is $\approx 44\,\mathrm{cm}$.

The positive bias tail can be interpreted as a mistake the vertexing algorithm commits when confronted with a specific decay geometry. When the $\Lambda^0 \to p\pi^-$ decay plane closely aligns with the xz bending plane, the bending induced by the magnet can produce either *opening* or *closing* tracks (depicted in top and bottom diagrams respectively in Figure 4.5. In the latter case the tracks will cross again at $z > z_{\mathrm{vtx}}^{\Lambda}$; if y displacement is sufficiently small, the algorithms converges on this «ghost» vertex instead of the real one.

To test out this hypothesis we define the *horizontality* of a $\Lambda^0 \to p\pi^-$ event as follows:

$$h = \mathrm{sign}\left(\Lambda_{\mathrm{PID}}^0\right)\mathrm{sign}\left(B_y\right)\ \frac{a_y}{|\vec{a}|}, \tag{4.1}$$

where

$$\vec{a} := \vec{p}_p \times \vec{p}_\pi \tag{4.2}$$

is the cross product of proton and pion momenta at production vertex, $\mathrm{sign}\left(B_y\right)$ is the dipole magnet polarity[21] and $\mathrm{sign}\left(\Lambda_{\mathrm{PID}}^0\right)$ is the sign of the PDG Monte

---

[21]The LHCb dipole magnet polarity is reversed roughly twice per month to allow for studies on decay asymmetries [76]. The $B_y > 0$ configuration is conventionally known as *magnet up* polarity, $B_y < 0$ as *magnet down*.

Figure 4.5: Depiction of three $\Lambda^0 \to p\pi^-$ configurations and the associated horizontality values. The horizontal planes in the top and bottom diagrams are aligned to the LHCb xz plane, the vertical plane in the middle diagram to the yz plane.

Carlo particle numbering scheme of the mother particle ($+1$ for Λ^0, -1 for $\bar{\Lambda}^0$) [3].

Decays with $h = \pm 1$ lie exactly on the xz bending plane, $h = -1$ events having closing $p\pi^-/\bar{p}\pi^+$ tracks and $h = +1$ events having opening tracks, while $h = 0$ events lie on the yz plane (see Figure 4.5). A horizontality bias $\Delta h := h_{\text{reco}} - h_{\text{true}} > 1$ thus becomes the signature of a «ghost» vertex $\Lambda^0 \to p\pi^-$ event. As per Figure 4.6a, this issue affects $\approx 25\%$ of reconstructed $\Lambda_b^0 \to J/\psi\,\Lambda^0$ events, most of those being $\Delta h \approx 2$ events (from $h = -1$ to $h = +1$), with almost no event with $\Delta h < -1$.

Isolating $\Delta h \geq 1$ events and studying their z_{vtx}^Λ bias distributions (Figure 4.6b), it becomes clear that they are largely responsible for the high bias observed in Figure 4.4a. Significant asymmetry effects are still visible in the $\Delta h < 1$ distribution, which is still skewed towards positive bias. While not ideal, this is somewhat expected given that the Vertex Fitter algorithm scans for candidate vertices starting from the first measurement position (i.e. the T1–T3 stations) and moving upstream.

Most $\Lambda_b^0 \to J/\psi\,\Lambda^0$ events, even those with ghost vertex reconstruction, still maintain a limited $\lesssim 1.0\,\text{m}$ bias on z_{vtx}^Λ. A smaller substructure with $\geq 2.0\,\text{m}$ bias emerges when plotting the distribution in logarithmic scale, as in Figure 4.4b. Figure 4.7a provides a top view of the Λ^0 decay vertices of these events, showing the distribution of true z_{vtx}^Λ and x_{vtx}^Λ. Most Λ^0 in high bias
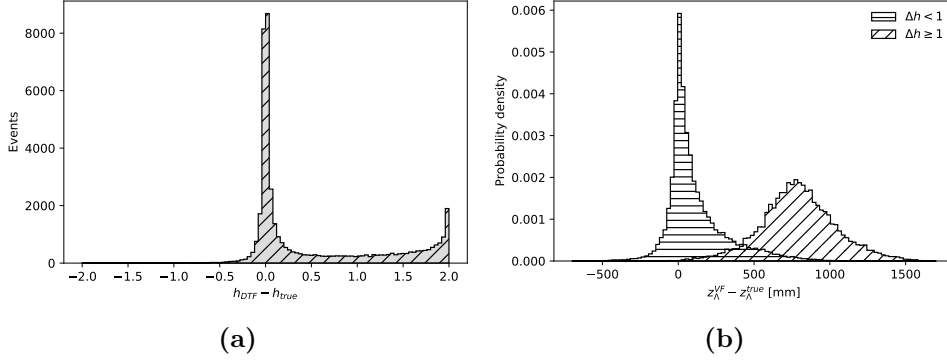
(a)   (b)

Figure 4.6: *(a)* Horizontality bias distribution for simulated $\Lambda_b^0 \to J/\psi \, \Lambda^0$ events (prefilters applied), comparing results from Decay Tree Fitter algorithm with J/ψ and Λ^0 mass constraints with true values. *(b)* Distribution of z_{vtx}^Λ bias for events with horizontality bias < 1 (*horizontal hatching*) and ≥ 1 (*diagonal hatching*).

(a)   (b)
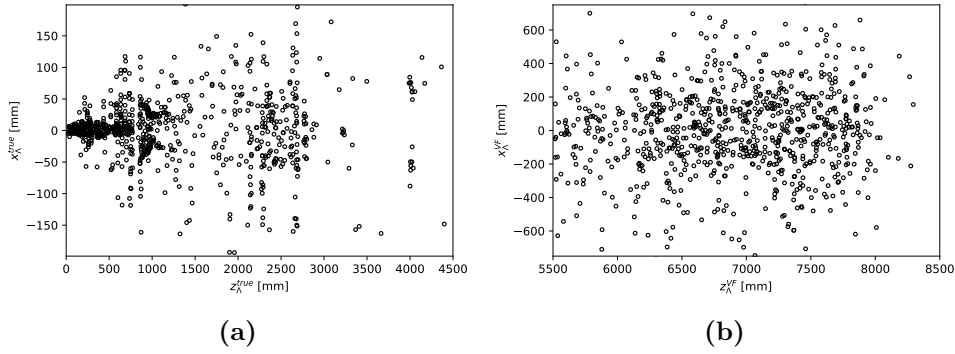
Figure 4.7: Distribution of simulated $\Lambda_b^0 \to J/\psi \, \Lambda^0$ events (prefilters applied) with $z_\Lambda^{\mathrm{VF}} - z_\Lambda^{\mathrm{true}} \geq 2.0\,\mathrm{m}$ as function of true *(a)* and reconstructed *(b)* x_Λ^{vtx} and z_{vtx}^Λ. This corresponds to a top view of true and reconstructed Λ^0 decay vertices.

events decay in the earlier sections of the detector ($z < 3.0\,\mathrm{m}$); the high spatial concentration in specific regions of the xz plane, such as the «wings» around $z \approx 1.0\,\mathrm{m}$, as well as the consistency between the placement of these structures and the location of the different LHCb subdetectors (cf. Figure 2.2), suggest that they may be the result of interaction with the material.

No selection on reconstructed variables is possible to filter this class of events: Figure 4.7b shows that the Λ^0 vertices are reconstructed in seemingly arbitrary positions. Their impact on the overall performance on signal is nevertheless neglectable, since this events amount to roughly 1.7% of the prefiltered simulated sample.
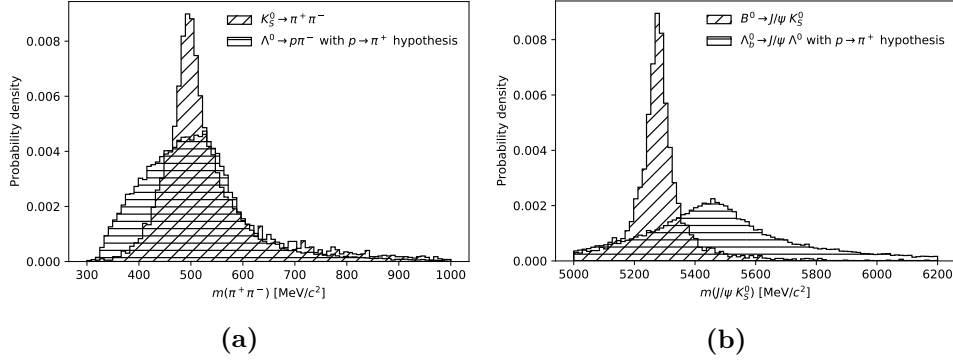
Figure 4.8: Comparison of simulated $m(\pi^+\pi^-)$ *(a)* and $m(J/\psi\ K_S^0)$ *(b)* distributions: $\Lambda_b^0 \to J/\psi\,\Lambda^0$ events with $p \to \pi^+$ mass hypothesis are labeled by *horizontal hatching*, $B^0 \to J/\psi\,K_S^0$ events by *diagonal hatching*.

4.2 Physical background veto

The main source of physical background for the $\Lambda_b^0 \to J/\psi\,(\to \mu^+\mu^-)\,\Lambda^0\,(\to p\pi^-)$ decay is the similar

$$B^0 \to J/\psi\,(\to \mu^+\mu^-)\,K_S^0\,(\to \pi^+\pi^-). \tag{4.3}$$

The final states of the two decays only differ for a $p \leftrightarrow \pi^+$ change. The K_S^0 meson also has a similar mean lifetime to the Λ^0, thus we expect a sizeable number of K_S^0 decaying after the dipole magnet. To top it off, the B^0 mass $(5279.65\,\mathrm{MeV}/c^2)$ is fairly close to that of Λ_b^0 $(5619.60\,\mathrm{MeV}/c^2)$ [3], muddying the waters in invariant mass fits.

As discussed in Section 2.2.2, the LHCb detector employs the two RICH systems to identify and distinguish between protons, pions and kaons; in the case of Λ^0 decaying after the dipole magnet RICH1 contributions is impossible, but information from RICH2 would still be available for the vast majority of the decays at hand. However, this is where the experimental nature of physics analyses with T tracks becomes relevant once again: due to technical issues in the implementation, RICH2 information for particle identification is unavailable for T tracks recorded during LHC Runs 1 and 2[22], making $B^0 \to J/\psi\,K_S^0$ discrimination much more difficult.

To address this issue, events must pass a physical background selection based on the results of a Decay Tree Fitter refit with $p \to \pi^+$ mass hypothesis and kinematic constraints on the J/ψ and K_S^0 invariant masses. The veto rejects an event if the DTF refit converges and $m(J/\psi\ K_S^0)$ is too close in

---

[22]A large effort by the Milan and Valencia LHCb research groups is underway at the time of writing to implement RICH2 information in T track trigger lines for Run 3.
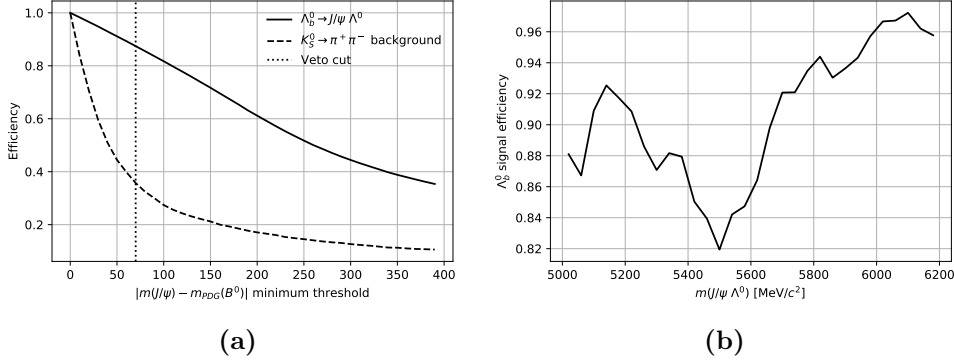
(a)   **(b)**

Figure 4.9: *(a)* Efficiency of physical background veto as a function of the invariant mass discrepancy threshold on simulated signal ($\Lambda_b^0 \to J/\psi \, \Lambda^0$, solid) and background ($B^0 \to J/\psi \, K_S^0$ with proton mass hypothesis, dashed) events. Chosen threshold marked by dotted line. *(b)* Efficiency of the veto on different $m(J/\psi \, \Lambda^0)$ bins for $\Lambda_b^0 \to J/\psi \, \Lambda^0$ signal events.

value to the B^0 PDG mass, i.e.

$$|m(J/\psi \ K_S^0) - m_{\mathrm{PDG}}(B^0)| < m_{\mathrm{thres}}, \qquad (4.4)$$

with tunable threshold m_{thres}.

This approach was tested on simulated $\Lambda_b^0 \to J/\psi \, \Lambda^0$ and $B^0 \to J/\psi \, K_S^0$ samples. The variable $m(J/\psi \ K_S^0)$ was chosen as opposed to $m(\pi^+\pi^-)$ due to the significant overlap between $m(\pi^+\pi^-)$ invariant mass distributions of actual $K_S^0 \to \pi^+\pi^-$ events and $\Lambda^0 \to p\pi^-$ events with $p \to \pi^+$ mass hypothesis (Figure 4.8a); the $m(J/\psi \ K_S^0)$ distributions are comparatively more separate (Figure 4.8b), allowing for a better selection on the variable.

Figure 4.9a shows the veto efficiency on $\Lambda_b^0 \to J/\psi \, \Lambda^0$ signal and $B^0 \to J/\psi \, K_S^0$ physical background as function of the threshold m_{thres}. The chosen value of $m_{\mathrm{thres}} = 70\,\mathrm{MeV}$ retains 87% of signal while rejecting 64% of background; signal efficiency as function of the $m(J/\psi \, \Lambda^0)$ invariant mass is depicted in Figure 4.9b.

4.3   HBDT classifier

Even after prefilter selections, combinatorial background exceeds the signal by a 10^3 factor. This section details my work in training and testing a multivariate classifier to discriminate $\Lambda_b^0 \to J/\psi \, \Lambda^0$ events, in order to lower the signal-to-background ratio to acceptable levels for physics analysis. After taking into consideration a wide range of classifiers, a histogram-based gradient boosting classification tree (also referred to as *histogram-based boosted decision tree*,
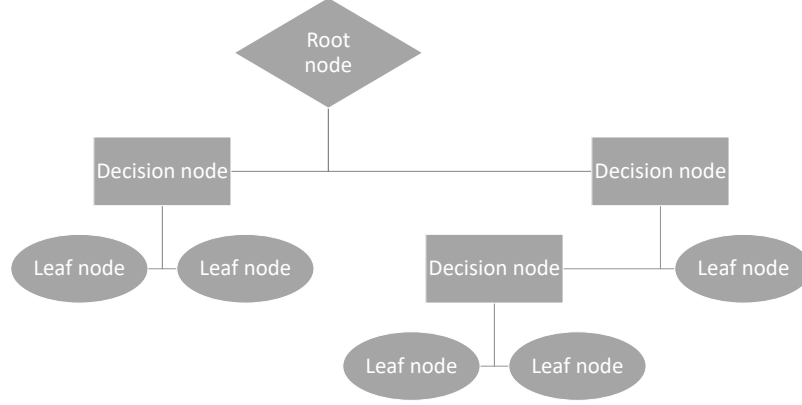
Figure 4.10: Diagram representation of a decision tree classifier.

HBDT for short) was chosen by virtue of its superior performance in studies on sample data [77]. Training, optimization and performance testing of the classifier were made with the Scikit-learn 0.24.2 package [78] for Python 3.6.8 [79].

The basic layout of a binary decision tree classifier is sketched in Figure 4.10: a sequence of decision nodes applies binary conditions according to available information on the individual event and eventually reaches a leaf node associated to the *event score t*, in this case the probability that the event is a $\Lambda_b^0 \to J/\psi\,\Lambda^0$ decay. Performance of the decision tree can be enhanced with *boosting* [80], whereby the ultimate classifier is built as a weighted average of a large number of weaker trees; the result is known as a boosted decision tree (BDT). In the adaptive boosting (AdaBoost) implementation used in this thesis, each tree T_k is trained on a reweighted data sample that prioritizes events misclassified by the T_{k-1} tree. The usage of a histogram-based BDT, arranging input samples into integer-valued bins, allows for much faster estimators when working with large data samples ($n \gtrsim 10^4$).

4.3.1 Training and testing data

Supervised training of a classifier requires signal and background datasets as input to learn to distinguish between the two classes. For signal I used simulated $\Lambda_b^0 \to J/\psi\,\Lambda^0$ events, while for the background I sampled side bands on the left ($[4870.2\,\mathrm{MeV/c^2}, 5020.2\,\mathrm{MeV/c^2}]$) and right ($[6220.2\,\mathrm{MeV/c^2}, 6520.2\,\mathrm{MeV/c^2}]$) of the expected Λ_b^0 peak in the Run 2 data $m(J/\psi\,\Lambda)$ distribution (see Figure 4.11).

I considered two options for the signal-to-background ratio:
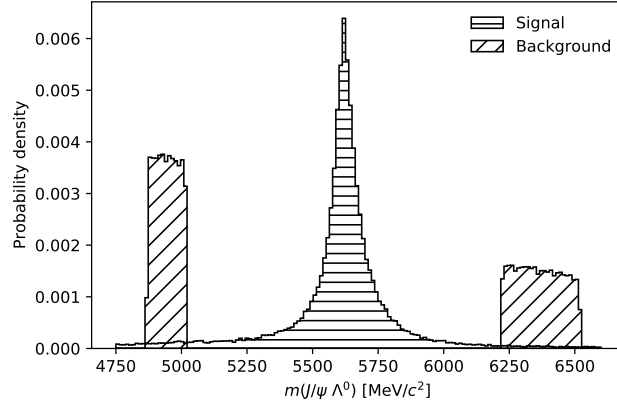
Figure 4.11: Signal (*horizontal hatching*) and background (*diagonal hatching*) data samples used for training the HBDT classifier.

- *balanced* training dataset: $\approx 145\,000$ events, evenly split between signal and background;

- *unbalanced* training dataset: $\approx 73\,000$ signal events, ≈ 3.6 million background events.

The standard approach to machine learning calls for a roughly balanced dataset to train the classifier on signal and background alike. Given the disproportionate signal-to-background ratio in the case at hand, however, I also explored the unbalanced approach, aiming for improved background rejection at the price of (reasonably) subpar signal identification. Classifiers trained with balanced or unbalanced datasets showed very similar performances after score threshold optimization (see Section 4.3.4), the main difference being that balanced HBDTs require harsher score selection criteria; since no reason surfaced to favour either approach, I opted to follow the more conventional path and employed a balanced training dataset. Testing data for performance evaluation was sampled from the same pool as training data in a 1:9 ratio.

The following kinematic variables (also known as *features*) were used for signal discrimination: transverse and longitudinal momenta of p, π^- and J/ψ; coordinates of the $\Lambda^0 \to p\pi^-$ vertex position; Λ_b^0 and Λ^0 pointing angles ξ_p, i.e. the angle between the particle line of flight and its momentum; $\tilde{\chi}^2_{\mathrm{vtx}}$ of the $\Lambda_b^0 \to J/\psi\,\Lambda^0$ and $\Lambda^0 \to p\pi^-$ vertices; the increase $\Delta\chi^2_{\mathrm{PV}}$ of the primary vertex χ^2 when a particle is included in the fit, for Λ_b^0 and Λ^0; the distances χ^2_{dist} between Λ_b^0 and Λ^0 decay vertices and the primary vertex in χ^2 units.

The momenta for proton and pion are computed with J/ψ mass constraints. While the usage of p, π^- momenta from the Decay Tree Fitter with the additional Λ^0 mass constraint was considered, preliminary tests on classifiers

showed a slight degradation of their performance when the double mass constraint was adopted. The standard Vertex Fitter momenta were used for the J/ψ on account of the already great resolution provided by using Long muons.

4.3.2 Hyperparameter optimization

The supervised training process for a classifiers tunes a set of parameters to minimize a given loss function; in the HBDT case, for instance, decision node and tree weights are trained according to the output of a logistic loss function. In contrast, *hyperparameters* are parameters governing the learning process itself and thus require separate optimization.

I selected three hyperparamters for optimization:

- learning rate $L \in [0.003, 0.006, 0.010, 0.015]$, the shrinkage factor for the contribution of each tree in the boosting process.

- maximum number of boosting iterations $n_{max}^{iter} \in [1500, 2500, 5000]$;

- maximum number of leaves for each tree $n_{max}^{leaves} \in [100, 200, 400, 800]$.

Other hyperparameters, such as the maximum number of bins and the minimum number of samples per leaf, were left with the default values of the Scikit-learn implementation.

I performed an exhaustive grid scan of the aforementioned hyperparameter values. Each classifier was evaluated on training data using a stratified k-folds cross validation technique: the training sample is split in $k = 5$ subsamples, the classifier is trained in rotation on $k - 1$ samples and tested with the k-th, and the k results are averaged together. For a given score threshold i, we define *precision* P_i and *recall* R_i in terms of the true/false positive/negative rates: precision

$$P_i = \frac{n_{TP}}{n_{TP} + n_{FP}} \tag{4.5}$$

scores how often the classifier mistakes background for signal, while recall

$$R_i = \frac{n_{TP}}{n_{TP} + n_{FN}} \tag{4.6}$$

scores how well the classifier is able to recognize the signal. Performance of cross-validated classifiers on training and testing sub-samples[23] was evaluated using the *average precision* figure of merit

$$AP = \sum_{i=2}^{n} (R_i - R_{i-1})P_i \tag{4.7}$$

---

[23]Still working within a cross-validation mindset, here *testing data* is used to refer to the k-th fold in the cross-validation. No actual testing data as defined in Section 4.3.1 was used for the hyperoptimization step.
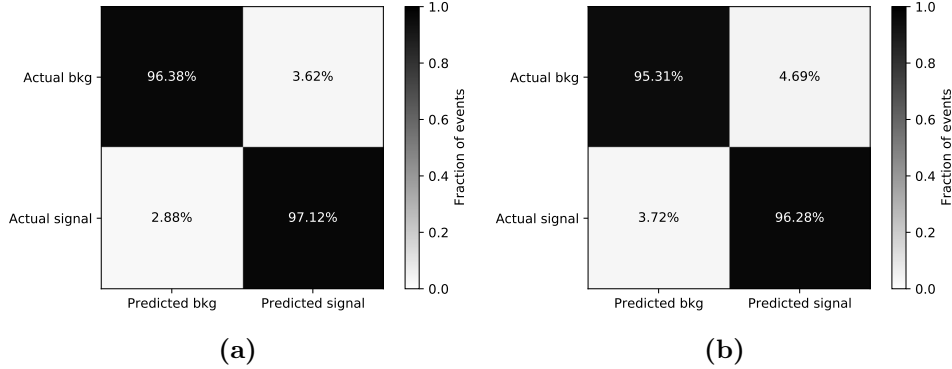
Figure 4.12: Confusion matrices visualizing the performance of the HBDT classifier on training *(a)* and testing *(b)* data samples. Percentages and chromatic scale are normalized to the true event classification: for instance, the top left and top right quadrants of a matrix represent the fraction of true background events reconstructed as background or signal, respectively. Binary classification uses an illustrative response threshold $t_{\text{thres}} = 0.5$.

for a set of n thresholds.

When considering the choice of the best classifier, I aimed to strike a balance between the AP score on testing data, acting as a raw performance grading of the classifier, and the difference in AP score between training and testing samples, which was necessary to prevent overtraining. The HBDT classifier used for the remainder of this thesis, with $L = 0.003$, $n_{\text{max}}^{\text{iter}} = 2500$, and $n_{\text{max}}^{\text{leaves}} = 100$, is the result of this selection process.

4.3.3 Performance test

Performance of the trained HBDT optimized as per Section 4.3.2 was evaluated on the smaller test sample described in Section 4.3.1. Figure 4.12 shows the confusion matrices for training and test data with a score threshold of $t_{\text{thres}} = 0.5$, meaning events with $t > t_{\text{thres}}$ are classified as signal. The matrices summarize the true/false positive/negative rates for the classifier, highlighting the expected slight dip in accuracy when moving from training to test sample.

Comparing the signal and background score distributions (Figures 4.13a and 4.13b respectively) for training and test data samples, we observe a light discrepancy at the end tails. Such a split in favour of performance on training data is usually an early indicator of ovetraining. I investigated this possibility by conducting a two-sample Kolmogorov-Smirnov (K-S) test [81], which computes the probability p that two samples are drawn from the same probability distribution. Conventionally, the K-S is considered passed if $p > 0.05$; this was true for both signal ($p_{\text{sig}} \approx 0.30$) and background ($p_{\text{bkg}} \approx 0.07$), confirming
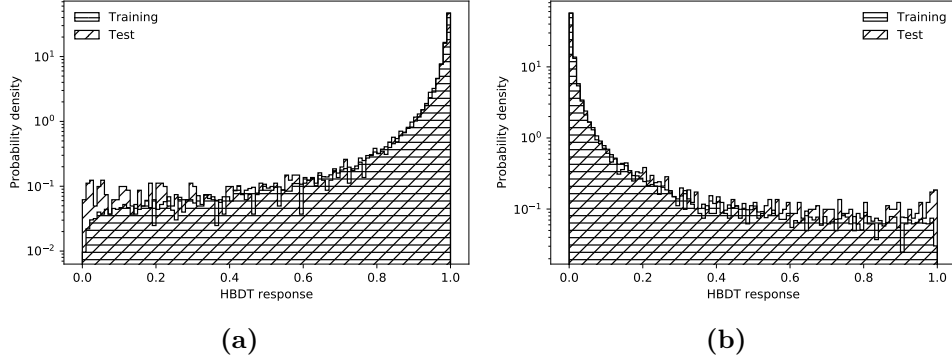
Figure 4.13: Response distribution of the HBDT classifier on signal *(a)* and background *(b)* events. The training sample is represented by *horizontal hatching*, the test sample by *diagonal hatching*.
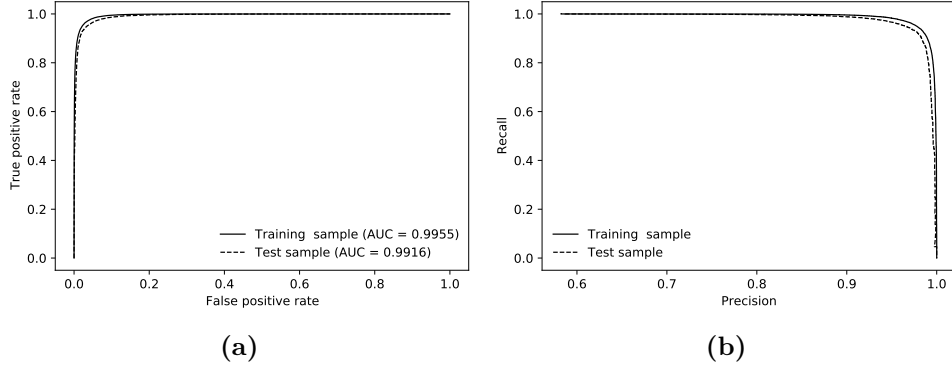

Figure 4.14: Receiving operating characteristic (ROC) curve *(a)* and precision-recall curve *(b)* for the HBDT classifier on training (*solid*) and test (*dashed*) samples. The legend in the ROC plot includes the area-under-curve (AUC) score.

that the classifier is not significantly overtrained.

I assessed the classifier performance at different thresholds using the receiving operating characteristic (ROC) and precision-recall curves. The ROC curve, depicted in Figure 4.14a, plots the true positive versus false positive rates for a set of thresholds left as hidden variables: a random classifier would bisect the plot plane, while a perfect classifier would adhere to the axes and cross the $(0, 1)$ point. The integral below the curve is itself a figure of merit, the area-under-curve (AUC), with better classifiers scoring values closer to unity. The precision-recall curve follows a similar principle to the ROC curve, plotting precision (4.5) and recall (4.6) for different thresholds; in this case, a perfect classifier would cross the $(1, 1)$ point. In both instances, the classifier shows great performance and high consistency between training and test samples.
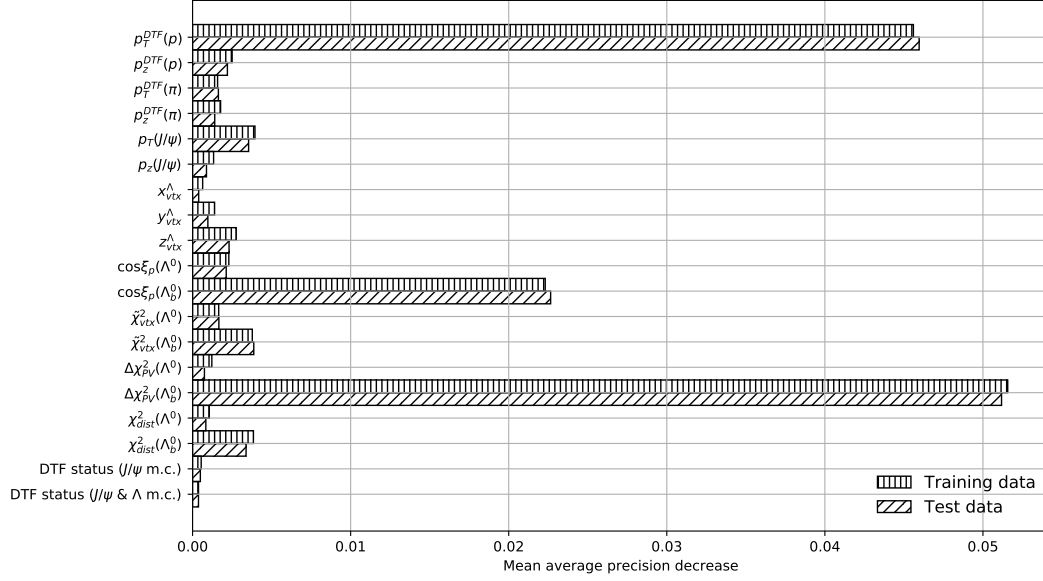
Figure 4.15: Decrease in the HBDT AP score when permutating the values of individual features in training (*vertical hatching*) and testing (*diagonal hatching*) data samples. See the legend of Table 4.1 for details on the feature labels.

Finally, I evaluated the importance of selected features in signal discrimination. This was achieved by permutating the values of each feature at a time and evaluating the average decrease in the baseline AP score of the classifier over ten of such permutations. Results are shown in Figure 4.15 for training and testing samples: in both cases the most discriminating features, causing the largest drops in AP, are the transverse momentum of the proton, the pointing angle of the Λ_b^0 and the increase in the primary vertex χ^2 when including the Λ_b^0 in the fit.

4.3.4   Threshold optimization

The final step in the classifier tuning process was the optimization of the threshold to maximize signal significance

$$\frac{s}{\sqrt{s+b}},\tag{4.8}$$

with s and b being the signal and background event counts after the HBDT filter. A simple way to do it would be to fit the $J/\psi\,\Lambda^0$ invariant mass distribution of filtered data with signal and background functions and extract the corresponding rates via integration; however, this has the undesired side effect of heightening the risk of biasing the selection towards our specific data sample.

A different, more general approach was devised to prevent this. I first identified a «soft» score threshold $t_{\text{thres}} = 0.7$, just high enough for the filtered Run 2 data $m(J/\psi\,\Lambda^0)$ distribution (using values from Decay Tree Fitter with mass constraints on both particles) to show the Λ_b^0 resonance peak. I fit the simulated signal $m(J/\psi\,\Lambda^0)$ distribution (using values from Decay Tree Fitter with mass constraints on both particles) with a double-tailed asymmetric Crystal Ball function. The standard, single-tailed Crystal Ball probability density function is parameterized as

$$f_{\text{CB}}(m; \mu, \sigma, \alpha, n) = \begin{cases} \exp\left[-\frac{1}{2}\left(\frac{m-\mu}{\sigma}\right)^2\right], & \text{if } \frac{m-\mu}{\sigma} \geq -\alpha \\ A \cdot \left(B - \frac{m-\mu}{\sigma}\right)^{-n}, & \text{if } \frac{m-\mu}{\sigma} < -\alpha \end{cases}, \qquad (4.9)$$

with

$$A = \left(\frac{n}{|\alpha|}\right)^n e^{-\frac{\alpha^2}{2}} \qquad (4.10)$$

and

$$B = \frac{n}{|\alpha|} - |\alpha|. \qquad (4.11)$$

Its shape is thus a Gaussian core with a power-law tail on the lower end when m outdistances μ by $\alpha\sigma$. The double-tailed Crystal Ball expands on this concepts by having a Gaussian core with two asymmetric long tails, governed by different power laws:

$$f_{\text{sig}}(m; S, \mu, \sigma, \alpha_1, n_1, \alpha_2, n_2) = N \begin{cases} f_{\text{CB}}(m; \mu, \sigma, \alpha_1, n_1) & \text{if } m \leq \mu \\ f_{\text{CB}}(2\mu - m; \mu, \sigma, \alpha_2, n_2) & \text{if } m > \mu \end{cases}, \qquad (4.12)$$

with normalization

$$N = \frac{S}{\sigma(C_1 + C_2 + D)} \qquad (4.13)$$

itself depending on

$$C_1 = \frac{n_1}{\alpha_1(n_1 - 1)} e^{-\frac{\alpha_1^2}{2}}, \qquad (4.14)$$

$$C_2 = \frac{n_2}{\alpha_2(n_2 - 1)} e^{-\frac{\alpha_2^2}{2}}, \qquad (4.15)$$

and

$$D = \sqrt{2\pi}\left(G_{\text{cum}}(\alpha_2) - G_{\text{cum}}(-\alpha_1)\right), \qquad (4.16)$$

where

$$G_{\text{cum}}(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} dt\, e^{-\frac{t^2}{2}} \qquad (4.17)$$

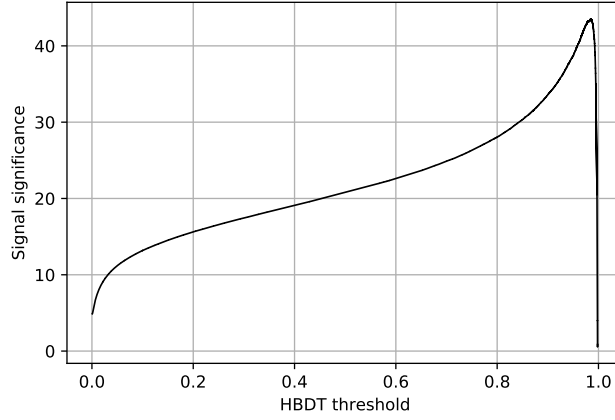is the Gaussian cumulative distribution function evaluated at x.

Figure 4.16: Projected $\Lambda_b^0 \to J/\psi \, \Lambda^0$ signal significance over background as a function of the HBDT response threshold used for selection.

When fitting $m(J/\psi \, \Lambda^0)$ in Run 2 data, the function is the sum of two contributions: signal is modeled with the same double-tailed Crystal Ball from (4.9), with all parameters bar S and σ fixed to their Monte Carlo fit best values; the combinatorial background is fit with a conventional exponential function

$$f_{\mathrm{bkg}}(m; C, a) = Ce^{-am}. \tag{4.18}$$

As will be shown in Section 4.4, the outlined combination of f_{sig} and f_{bkg} provides an excellent description of Run 2 data, as long as sufficient filtering is applied for the Λ_b^0 peak to emerge.

This data fit provides preliminary s_0 and b_0 values, obtained by integrating signal and background fit functions in the $\pm 3\sigma$ region around μ. I used these values to estimate the signal (background) rate s_i (b_i) for events passing a score threshold t_i with associated efficiency ε_i^s (ε_i^b):

$$s_i = s_0 \frac{\varepsilon_i^s}{\varepsilon_0^s}, \tag{4.19a}$$

$$b_i = b_0 \frac{\varepsilon_i^b}{\varepsilon_0^b}. \tag{4.19b}$$

Here ε_0^s and ε_0^b are the signal and background efficiencies associated to the soft threshold with s_0 and b_0 rates. Efficiencies were computed on the test sample and correspond to true/false positive rates evaluated for the ROC curve (Figure 4.14a).

Using the estimated s_i and b_i, I calculated the predicted signal significance (4.8) for each threshold t_i. Figure 4.16 shows the results as a function of the
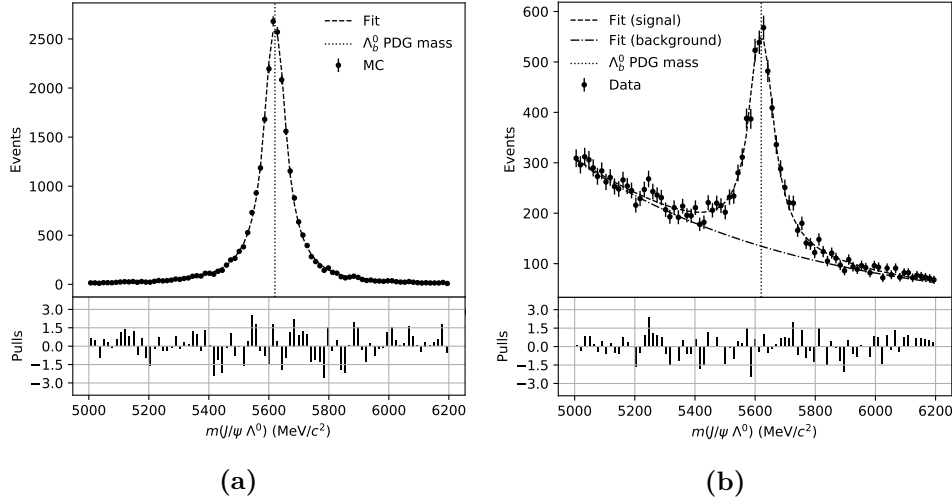
(a)   (b)

Figure 4.17: Fitted $m(J/\psi\,\Lambda^0)$ invariant mass distributions for simulated $\Lambda_b^0 \to J/\psi\,\Lambda^0$ events *(a)* and Run 2 data *(b)* after all selection steps. Signal fit function is *dashed*, background fit function in *(b)* is *dash-dotted*. The PDG value for the Λ_b^0 mass [3] is marked by the *dotted vertical line.* Fit pulls (data-fit discrepancy divided by uncertainty) are shown below the main plots.

score threshold. The best performing threshold was found to be $t_{\text{thres}} = 0.985$, with signal significance ≈ 46 and signal purity

$$\frac{s}{s+b} \approx 0.6. \tag{4.20}$$

4.4   Performance on data

Figures 4.17a and 4.17b show the $m(J/\psi\,\Lambda^0)$ invariant mass distributions of Monte Carlo simulated signal and Run 2 data after applying prefilters, $B^0 \to J/\psi\,K_S^0$ veto and the optimized $t > 0.985$ selection on the HBDT classifier score. To fit the distributions I adopted the same approach outlined in Section 4.3.4: the simulated signal is described with the double-tailed asymmetric Crystal Ball (4.9); the best fit parameter values, except for S and σ, are fixed for the signal function in the Run 2 data fit, while the exponential (4.18) was used to model the combinatorial background.

Binned fit results are listed in Table 4.2. The employed functions provide a great description for both the Monte Carlo simulation and the real data from Run 2, as seen from the fit-data discrepancies consistently below 3σ for each data bin. The resolution of the resonance peak is $\approx 12\%$ worse in real data compared to the simulation, an expected consequence of the underlying combinatorial background. Integrating the functions in the $\mu \pm 3\sigma$ interval,

we estimate a ≈ 3600 signal event yield and ≈ 2400 background event yield. These are going to be the baseline yields for analyses on the Λ^0 electromagnetic dipole moments.

Variable	Unit	Minimum	Maximum
$p(p)$	MeV$/c$	2 000	500 000
$p_T(p)$	MeV$/c$	400	–
$p(\pi^-)$	MeV$/c$	10 000	500 000
z_Λ^{vtx}	mm	5 500	8 500
$p_T(\Lambda^0)$	MeV$/c$	450	–
$m(p\pi^-)$ (Vertex Fitter)	MeV$/c^2$	600	1 500
$m(p\pi^-)$ (combined)	MeV$/c^2$	–	2 000
$m(p\pi^-)$ (measured)	MeV$/c^2$	–	1 500
$\cos\xi_p(\Lambda^0)$	–	0.9999	–
$\Delta\chi^2_{\mathrm{PV}}(\Lambda^0)$	–	–	200
$\chi^2_{\mathrm{dist}}(\Lambda^0)$	–	–	2×10^7
$\chi^2_{\mathrm{vtx}}(\Lambda^0)$	–	–	750
$|m(\mu^+\mu^-) - m_{\mathrm{PDG}}(J/\psi)|$	MeV$/c^2$	–	90
$m(J/\psi\,\Lambda^0)$ (combined)	MeV$/c^2$	4 700	–
$m(J/\psi\,\Lambda^0)$ (Vertex Fitter)	MeV$/c^2$	–	8 500
$|\cos\xi_p(\Lambda_b^0)|$	–	0.99	–
$\Delta\chi^2_{\mathrm{PV}}(\Lambda_b^0)$	–	–	1 750
$\chi^2_{\mathrm{vtx}}(\Lambda_b^0)$	–	–	150

Table 4.1: Prefilter selection criteria applied to simulated $\Lambda_b^0 \to J/\psi\,\Lambda^0$ signal and Run 2 data. The Vertex Fitter invariant mass is computed by the homonymous algorithm; the *combined* invariant mass is computed from the 4-momenta of the daughter particles at the first measurement position, without track extrapolation; the *measured* invariant mass is computed in the same way as the combined mass, but after extrapolation at the reconstructed decay vertex of the mother particle. Angle ξ_p for a particle is computed between the line connecting its origin and decay vertices and the direction of its momentum. $\Delta\chi^2_{\mathrm{PV}}$ is the increase of the primary vertex χ^2 when the particle is included in the fit. χ^2_{dist} is the geometrical distance between the primary vertex and the particle decay vertex in χ^2 units.

Parameter	Unit	Simulation		Data	
		Best value	Error	Best value	Error
S	–	359 200	2 800	64 300	1 900
μ	MeV/c^2	5 620.1	0.5	5 620.1	–
σ	MeV/c^2	36.8	1.0	41.2	1.6
α_1	–	0.985	0.035	0.985	–
n_1	–	2.43	0.11	2.43	–
α_2	–	0.995	0.035	0.995	–
n_2	–	2.71	0.14	2.71	–
C	–	–	–	240 000	40 000
a	$\times 10^{-3}\ \mathrm{MeV}^{-1}c^2$	–	–	1.332	0.029

Table 4.2: Parameter best values and associated uncertainties from $m(J/\psi\,\Lambda^0)$ fits on simulated signal and Run 2 data, after all selection steps. Parameters with no reported error are fixed.

Bibliography

[1] I. J. R. Aitchinson and A. J .G. Hey. *Gauge Theories in Particle Physics: A Practical Introduction. Volume 1: From Relativistic Quantum Mechanics to QED.* CRC Press, 2013.

[2] J. Erler and M. Schott. Electroweak precision tests of the Standard Model after the discovery of the Higgs boson. *Progress in Particle and Nuclear Physics*, 106:68–119, May 2019.

[3] P. A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020.

[4] A. B. McDonald. Sudbury Neutrino Observatory results. *Physica Scripta*, T121:29–32, Jan 2005.

[5] S.-K. Choi et al. (Belle Collaboration). Observation of a narrow charmonium-like state in exclusive $B^{\pm} \rightarrow K^{\pm}\pi^{+}\pi^{-} J/\psi$ decays. *Phys. Rev. Lett.*, 91(26), Dec 2003.

[6] T. Nakano et al. (LEPS Collaboration). Evidence for a narrow $S = +1$ baryon resonance in photoproduction from the neutron. *Phys. Rev. Lett.*, 91:012002, Jul 2003.

[7] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[8] G. Aad et al. (ATLAS Collaboration). Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1–29, Sep 2012.

[9] S. Chatrchyan et al. (CMS Collaboration). Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, Sep 2012.

[10] J. Zupan. Introduction to flavour physics. *arXiv e-prints*, Mar 2019.

[11] M. S. Sozzi. *Discrete symmetries and CP violation: from experiment to theory.* Oxford University Press, 2013.

[12] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, Feb 1957.

[13] R. L. Garwin, L. M. Lederman, and M. Weinrich. Observations of the failure of conservation of parity and charge conjugation in meson decays: the magnetic moment of the free muon. *Phys. Rev.*, 105:1415–1417, Feb 1957.

[14] I. B. Khriplovich and S. K. Lamoreaux. *CP Violation Without Strangeness. Electric Dipole Moments of Particles, Atoms, and Molecules.* Springer, 1997.

[15] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π decay of the K_2^0 meson. *Phys. Rev. Lett.*, 13:138–140, Jul 1964.

[16] A. Alavi-Harati et al. (KTeV Collaboration). Observation of direct CP violation in $K_{S,L} \to \pi\pi$ decays. *Phys. Rev. Lett.*, 83(1):22–27, Jul 1999.

[17] B. Aubert et al. (BaBar Collaboration). Measurement of CP-violating asymmetries in B^0 decays to CP eigenstates. *Phys. Rev. Lett.*, 86(12):2515–2522, Mar 2001.

[18] K. Abe et al. (Belle Collaboration). Observation of large CP violation in the neutral B meson system. *Phys. Rev. Lett.*, 87(9), Aug 2001.

[19] L. Canetti, M. Drewes, and M. Shaposhnikov. Matter and antimatter in the Universe. *New Journal of Physics*, 14(9):095012, Sep 2012.

[20] G. Steigman. Primordial nucleosynthesis in the precision cosmology era. *Annual Review of Nuclear and Particle Science*, 57(1):463–491, Nov 2007.

[21] L. Henry, D. Marangotto, A. Merli, N. Neri, J. Ruiz, and F. Martinez Vidal. Search for new physics via baryon EDM at LHC. 2020.

[22] M. E. Pospelov and I. B. Khriplovich. Electric dipole moment of the W boson and the electron in the Kobayashi-Maskawa model. *Sov. J. Nucl. Phys.*, 53:638–640, 1991.

[23] I. B. Khriplovich. Quark electric dipole moment and induced θ term in the Kobayashi-Maskawa model. *Phys. Lett. B*, 173:193–196, 1986.

[24] I. B. Khriplovich and A. R. Zhitnitsky. What is the value of the neutron electric dipole moment in the Kobayashi-Maskawa model? *Phys. Lett. B*, 109:490–492, 1982.

[25] F. J. Botella, L. M. Garcia Martin, D. Marangotto, F. Martinez Vidal, A. Merli, N. Neri, A. Oyanguren, and J. Ruiz Vidal. On the search for the electric dipole moment of strange and charm baryons at LHC. *The European Physical Journal C*, 77(3), Mar 2017.

[26] J. M. Pendlebury et al. Revised experimental upper limit on the electric dipole moment of the neutron. *Phys. Rev. D*, 92:092003, Nov 2015.

[27] E. Dahl-Jensen, N. Doble, D. Evans et al. A new measurement of the magnetic moment of the Λ-hyperon. *Nuov Cim A*, 3:1–24, May 1971.

[28] G. Baroni, S. Petrera, and G. Romano. An upper limit of the electric-dipole moment of the Λ-hyperon. *Lett. Nuovo Cimento*, 2:1256–1258, Dec 1971.

[29] L. Schachinger et al. Precise measurement of the Λ^0 magnetic moment. *Phys. Rev. Lett.*, 41:1348–1351, Nov 1978.

[30] L. Pondrom et al. New limit on the electric dipole moment of the Λ hyperon. *Phys. Rev. D*, 23:814–816, Feb 1981.

[31] G. Aad et al. (ATLAS Collaboration). Measurement of the parity-violating asymmetry parameter α_b and the helicity amplitudes for the decay $\Lambda_b^0 \to J/\psi\Lambda^0$ with the ATLAS detector. *Phys. Rev. D*, 89:092009, May 2014.

[32] J. D. Richman. An experimenter's guide to the helicity formalism. Technical report, Calif. Inst. Technol., Pasadena, CA, Jun 1984.

[33] T. D. Lee and C. N. Yang. General partial wave analysis of the decay of a hyperon of spin ½. *Phys. Rev.*, 108:1645–1647, Dec 1957.

[34] E. Leader. *Spin in Particle Physics.* Cambridge University Press, 2001.

[35] D. P. Grosnick et al. (The FNAL-E581/704 Collaboration). The design and performance of the FNAL high-energy polarized-beam facility. *Nucl. Instrum. Meth. A*, 290(2):269–292, 1990.

[36] R. Aaij et al. (LHCb Collaboration). LHCb detector performance. *International Journal of Modern Physics A*, 30(07):1530022, Mar 2015.

[37] L. Evans. The Large Hadron Collider. *Annual Review of Nuclear and Particle Science*, 61(1):435–466, 2011.

[38] R. Antunes-Nobrega et al. (LHCb Collaboration). LHCb reoptimized detector design and performance: Technical Design Report. Technical report, Geneva, 2003.

[39] A. A. Alves et al. (LHCb Collaboration). The LHCb detector at the LHC. *JINST*, 3:S08005, 2008. Also published by CERN Geneva in 2010.

[40] R. Aaij et al. (LHCb Collaboration). A precise measurement of the B^0 meson oscillation frequency. *The European Physical Journal C*, 76(7), Jul 2016.

[41] R. Aaij et al. (LHCb Collaboration). Precision measurement of the B^0_S–$\bar{B}^0_S$ oscillation frequency with the decay $B^0_S \rightarrow D^-_S \pi^+$. *New Journal of Physics*, 15, Apr 2013.

[42] R. Aaij et al. (LHCb Collaboration). Observation of CP violation in charm decays. *Phys. Rev. Lett.*, 122:211803, May 2019.

[43] R. Aaij et al. (LHCb Collaboration). Updated LHCb combination of the CKM angle γ. Technical report, CERN, Geneva, Oct 2020.

[44] R. Aaij et al. (LHCb Collaboration). Measurement of the $B^0_s \rightarrow \mu^+\mu^-$ branching fraction and effective lifetime and search for $B^0 \rightarrow \mu^+\mu^-$ decays. *Phys. Rev. Lett.*, 118:191801, May 2017.

[45] R. Aaij et al. (LHCb Collaboration). Study of D_J meson decays to $D^+\pi^-$, $D^0\pi^+$ and $D^{*+}\pi^-$ final states in pp collisions. *Journal of High Energy Physics*, 2013(9), Sep 2013.

[46] R. Aaij et al. (LHCb Collaboration). Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda^0_b \rightarrow J/\psi K^- p$ decays. *Phys. Rev. Lett.*, 115:072001, Aug 2015.

[47] E. Maurice. Fixed-target physics at LHCb. *arXiv e-prints*, Aug 2017.

[48] R. Aaij et al. (LHCb Collaboration). First measurement of charm production in fixed-target configuration at the LHC. *Phys. Rev. Lett.*, 122:132002, Apr 2019.

[49] S. Amato et al. (LHCb Collaboration). LHCb magnet: Technical Design Report. Technical report, Geneva, 2000.

[50] P. R. Barbosa-Marinho et al. (LHCb Collaboration). LHCb VELO (VErtex LOcator): Technical Design Report. Technical report, Geneva, 2001.

[51] J. Gassner, M. Needham, and O. Steinkamp. Layout and expected performance of the LHCb TT station. Technical report, CERN, Geneva, Apr 2004.

[52] P. R. Barbosa-Marinho et al. (LHCb Collaboration). LHCb inner tracker: Technical Design Report. Technical report, Geneva, 2002.

[53] P. R. Barbosa-Marinho et al. (LHCb Collaboration). LHCb outer tracker: Technical Design Report. Technical report, Geneva, 2001.

[54] I. Belyaev, G. Carboni, N. Harnew, C. Matteuzzi, and F. Teubert. The history of LHCb. *The European Physical Journal H*, 46(1), Mar 2021.

[55] S. Amato et al. (LHCb Collaboration). LHCb RICH: Technical Design Report. Technical report, Geneva, 2000.

[56] M. Adinolfi et al. (LHCb RICH group). Performance of the LHCb RICH detector at the LHC. *The European Physical Journal C*, 73(5), May 2013.

[57] S. Amato et al. (LHCb Collaboration). LHCb calorimeters: Technical Design Report. Technical report, Geneva, 2000.

[58] P. R. Barbosa-Marinho et al. (LHCb Collaboration). LHCb muon system: Technical Design Report. Technical report, Geneva, 2001.

[59] R. Antunes-Nobrega et al. (LHCb Collaboration). LHCb trigger system: Technical Design Report. Technical report, Geneva, 2003.

[60] I. Antcheva et al. ROOT — A C++ framework for petabyte data storage, statistical analysis and visualization. *Computer Physics Communications*, 180(12):2499–2512, 2009.

[61] S. Benson, V. Gligorov, M. A. Vesterinen, and J. M. Williams. The LHCb Turbo Stream. *Journal of Physics: Conference Series*, 664(8):082004, Dec 2015.

[62] T. Sjöstrand et al. An introduction to PYTHIA 8.2. *Computer Physics Communications*, 191:159–177, Jun 2015.

[63] S. Alioli, P. Nason, C. Oleari, and E. Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.

[64] A. Ryd et al. EvtGen: A Monte Carlo Generator for B-Physics. Technical report, 5 2005.

[65] S. Agostinelli et al. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[66] A. Piucci. The LHCb upgrade. *J. Phys.: Conf. Ser.*, 878:012012, Jul 2017.

[67] R. Aaij et al. (LHCb Collaboration). LHCb Trigger and Online Upgrade Technical Design Report. Technical report, May 2014.

[68] R. Aaij et al. (LHCb Collaboration). LHCb VELO Upgrade Technical Design Report. Technical report, Nov 2013.

[69] R. Aaij et al. (LHCb Collaboration). LHCb Tracker Upgrade Technical Design Report. Technical report, Feb 2014.

[70] R. Aaij et al. (LHCb Collaboration). LHCb PID Upgrade Technical Design Report. Technical report, Nov 2013.

[71] I. Belyaev et al. Handling of the generation of primary events in Gauss, the LHCb simulation framework. *Journal of Physics: Conference Series*, 331(3):032047, dec 2011.

[72] E. Bos and E. Rodrigues. The LHCb track extrapolator tools. Technical report, CERN, Geneva, Nov 2007.

[73] E. Hairer, G. Wanner, and S. P. Nørsett. *Solving Ordinary Differential Equations I: Nonstiff Problems*, chapter Runge-Kutta and Extrapolation Methods, pages 129–353. Springer Berlin Heidelberg, Berlin, Heidelberg, 1993.

[74] W. D. Hulsbergen. Decay chain fitting with a Kalman filter. *Nucl. Instrum. Meth. A*, 552:566–575, 2005.

[75] P. Li, E. Rodrigues, and S. Stahl. Tracking definitions and conventions for Run 3 and beyond. Technical report, CERN, Geneva, Feb 2021.

[76] M. Vesterinen. Considerations on the LHCb dipole magnet polarity reversal. Technical report, CERN, Geneva, Apr 2014.

[77] L. Pessina. *Studio di decadimenti di barioni Lambda in preparazione alla misura dei momenti di dipolo elettromagnetici a LHCb*. Bachelor's Thesis, Università degli Studi di Milano, Milan, Italy, Oct 2020.

[78] F. Pedregosa et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[79] G. Van Rossum and F. L. Drake. *Python 3 Reference Manual.* CreateSpace, Scotts Valley, CA, 2009.

[80] Y. Coadou. Boosted decision trees and applications. *EPJ Web of Conferences*, 55:02004, 2013.

[81] F. J. Massey. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253):68–78, 1951.