university of groningen
faculty of science and engineering

Asymmetries Between Matter and Antimatter Creation in Λ_b Decays in high q^2 region

David Vetter

University of Groningen

Asymmetries Between Matter and Antimatter Creation in Λ_B Decays in high q^2 region

Bachelor's Thesis

To fulfill the requirements for the degree of
Bachelor of Science in Physics
at University of Groningen under the supervision of
Dr. Ann-Kathrin Perrevoort (Van Swinderen Institute, University of Groningen)
and
Dr. Lorenz Willmann (Van Swinderen Institute, University of Groningen)

David Vetter (s5245672)

July 14, 2025

Contents

Acknowledgments

I would like to express my upmost gratitude to a number of great individuals whom have helped me throughout this bachelor thesis. Before that, I would also like to thank all of the colleagues at the KVI institute who have made the working environment a very pleasurable experience.

Firstly, I cannot express my gratefulness enough to my supervisor Dr. Ann-Kathrin Perrevoort. Thank you, for consistently answering any sort of question or concern that I had. Thank you for always being incredibly detailed with your explanations and guidance. Thank you for you the organization, and the late e-mails on the weekends. I sincerely appreciate all of the work you have put in to make my experience a great one. The commitment you have demonstrated from the very beginning until the very end to ensure that everything was working seamlessly is incredibly appreciated. It was a true pleasure working with you.

I also want to express my gratitude to Dr. Mick Mulder. Albeit not being my supervisor, you were always there to answer any question. I want to truly thank you for all the explanations and insights you have given me throughout my thesis. I have learned a tremendous amount because of you. I want to also thank you for staying patient throughout your explanations while Tomáš and I stared at you with confused faces. Thank you for everything, you are one of the main reasons why I enjoyed this bachelor project so much. I wish you great success with your future postdoc.

My many thanks go out to all of the LHCb Groningen Collegues, but especially I want to thank Tomáš Timko, Jorge Gil-Nagel, and Juriën Michèl Schut for the insightful conversations. You all made the several hard working hours a lot more enjoyable. I also want to wish all of you great success in your future endeavors.

Abstract

The decays $\Lambda_b^0 \to \Lambda^0\mu^+\mu^-$ and $\bar{\Lambda}_b^0 \to \bar{\Lambda}^0\mu^+\mu^-$ are analyzed using data from the LHCb experiment at CERN, focusing on the high-q^2 (squared dimuon invariant mass) region ($q^2 > 14.3\ \mathrm{GeV}^2/c^4$) to search for signs of CP — combined charge conjugation (C) and parity (P) symmetry — violation (CPV) through matter–antimatter asymmetry. The presence of an asymmetry beyond the expected CPV from the CKM matrix in weak interaction, would indicate physics beyond the Standard Model (BSM). The analysis was performed on proton–proton collision data from Run 2 (years 2016-2018) at $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of approximately 6 fb^{-1}. The measured raw asymmetry between Λ_b^0 and $\bar{\Lambda}_b^0$ decays was $(5.31 \pm 8.33)\%$, corresponding to a statistical significance of 0.421σ, concluding that the result is has no statistical significance. This result is consistent with Standard Model predictions and prior studies, providing no evidence for CPV or BSM effects in this observable. To improve the sensitivity of future analyses, increased integrated luminosity is recommended to reduce statistical uncertainties. Additionally, the implementation of a boosted decision tree tailored to this decay, along with selection criteria for the lifetime of Λ_b^0 and Λ^0 baryons, could enhance background suppression.

1 Introduction

The *Standard Model of particle physics* (SM), developed in the latter half of the twentieth century, has been essential for describing fundamental particles and their interactions. It has successfully managed to unify electromagnetic, weak and strong nuclear forces within a single theoretical framework and has demonstrated remarkable accuracy in its predictions across a high range of energies, many of which have been confirmed through high-precision experiments [1, 2]. Examples of this include the discoveries of the W and Z bosons [3, 4] and later the Higgs boson in 2012 [5, 6]. These were major milestones that affirmed the SMs validity. Despite its success, the SM does not account for several observed phenomena, such as the nature of dark matter, the matter-antimatter asymmetry of the universe, and the unification between the gravitational force with the aforementioned forces [7]. These limitations suggest that the SM is not the definitive theory of fundamental physics but rather a great approximation for our current understanding. Therefore, testing the limits of the SM, especially for decays which are sensitive probes for physics Beyond the Standard Model (BSM), is a valuable objective for high-energy physics.

A possible method to test for BSM physics is through the study of violations of discrete symmetries such as *parity (P), charge conjugation (C)*, and the combined *CP* symmetry. The SM incorporates CP violation (CPV) via weak interactions — mediated by the massive vector boson $W^{\pm}$ and Z — through the unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix. However, this mechanism is insufficient to explain the observed matter-antimatter asymmetry of the Universe, as originally outlined by Sakharov's condition for baryogenesis [8]. CPV was first observed in the neutral kaon system in 1964 [9], and more recently, the LHCb collaboration managed to provide the first evidence for CPV in baryonic decays, such as $\Lambda_b^0 \to \Lambda^0 K^0 K^-$ [10]. This discovery accentuates the importance of searching for CPV effects beyond mesonic decays. Unlike mesons, the presence of spin-$\frac{1}{2}$ baryonic particles in both initial and final states introduce nontrivial angular distributions and kinematic correlations that can be sensitive to new sources of CPV [11]. In particular, decays such as the $\Lambda_b^0 \to \Lambda^0 \mu^+ \mu^-$ are experimentally accessible, making them promising candidates for such tests. Although there is no expected CPV beyond the CKM contribution within the SM for this decay, any observed asymmetry beyond the expected CKM contributions in the production or decay rates of baryons and antibaryons could serve as an indication of BSM contributions [12].

The search for CPV in baryonic decays requires a capable experimental setup, able to reconstruct rare heavy-flavor processes with high precision. The LHCb experiment at CERN, one of the four main detectors at the LHC is explicitly designed for this purpose, specializing in the study of matter-antimatter differences in beauty quark decays [13]. It features forward geometry and excellent vertex resolution, high efficiency muon identification, and powerful particle identification that allows for reconstruction of hadrons [13, 14]. The decay $\Lambda_b^0 \to \Lambda^0 \mu^+ \mu^-$ is well suited for such studies due to its baryonic final state, and the dimuons, which LHCb can reconstruct with high precision. In this thesis, proton-proton collision data collected during Run 2 of the LHC (2016–2018) at $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of approximately 6 fb^{-1} [13], will be analyzed using Python in combination with the *uproot* library, enabling access to ROOT files outside of traditional CERN software. The main objective is to model the decay and determine the relative frequency of this process for matter and antimatter, and to identify possible asymmetries that could indicate CPV or contributions from physics BSM. Giving rise to my research question:

Can raw asymmetries between matter and antimatter in high-q^2 regions reveal deviations from Standard Model predictions?

2 Theory

This section will outline the concepts relevant to the analysis performed for this thesis. It covers an overview of the Standard Model (SM) of particle physics, followed by a discussion on discrete symmetries, in particular, the aforementioned charge conjugation (C), parity (P), and their combination (CP). In addition to the implications of their violation. Lastly the Λ_b^0 decay will be described, highlighting its role as a probe for potential physics beyond the Standard Model (BSM).

2.1 The standard model of particle physics

The Standard Model (SM) of particle physics is a well-established theoretical framework that describes and unifies three of the four known fundamental interactions in nature: electromagnetic, weak, and strong forces, as well as describing the elementary particles that mediate and participate in these interactions [1]. Matter and antimatter are composed of fermions, particles with half-integer spin that follow the Pauli exclusion principle. These are divided into two categories: quarks and leptons, each organized into three generations. Each generation is a set of particles with similar properties, but different rest masses. The higher generations are more massive and more unstable, causing them to often decay into particles of a lower generation. Quarks combine to form hadrons, including baryons and mesons, while leptons include charged particles: electrons, muons, and taus, in addition to their associated neutrinos. The interactions between these fermions are mediated by gauge bosons: photons for the electromagnetic force (described by quantum electrodynamics), $W^{\pm}$ and Z bosons for the weak force, and gluons for the strong force — described by quantum chromodynamics (QCD). This is visually explained Figure 1.

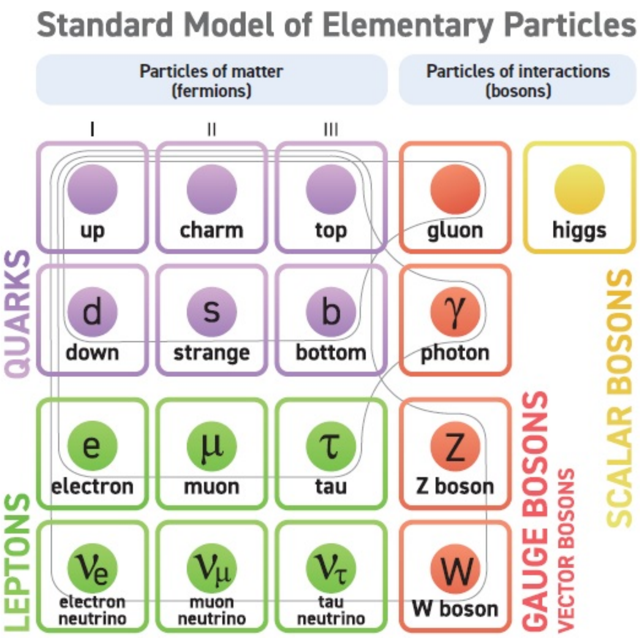


Figure 1: Schematic of elementary particles and their properties in the Standard Model [15].

These interactions are governed by a symmetry structure, described by the gauge group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, which defines the SM's theoretical foundation [16]. Each component of this group corresponds to one of the three fundamental forces: QCD for the strong interactions ($SU(3)_C$), and the electroweak theory for the unified description of electromagnetic and weak forces ($SU(2)_L \otimes U(1)_Y$). The most recent addition to the SM is the Higgs boson, discovered in 2012, which arises from the spontaneous breaking of the electroweak symmetry and is responsible for giving mass to the $W^{\pm}$ and Z bosons, while leaving the photon massless [5, 6].

2.2 CP violation and weak interactions

In particle physics, the main symmetries are parity (P), charge conjugation (C), and time reversal (T), collectively known as CPT. This combined symmetry is a fundamental symmetry in nature which cannot be violated. Additionally, the individual symmetries are conserved in both strong and electromagnetic interactions. However in the weak interaction, it has been observed that P, C, and CP are violated. CP violation (CPV) is introduced through the Cabibbo-Kobayashi-Maskawa (CKM) matrix.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1}$$

The CKM is a unitary matrix which describes the probability of transition between quark flavors in weak interactions. The elements in the CKM matrix, V_{ij} represent the probability of the transition from an up-type quark i (u, c, t) , to a down-type quark j (d, s, b). These transitions are solely mediated by the charged $W^{\pm}$ bosons, and can occur in both *Flavor Changing Charged Currents* (FCCC) and *Flavor Changing Neutral Currents* (FCNC). In this paper, the focus is set on FCNCs, which are transitions in which a quark changes flavor but not its type. This is illustrated in Figure 2.

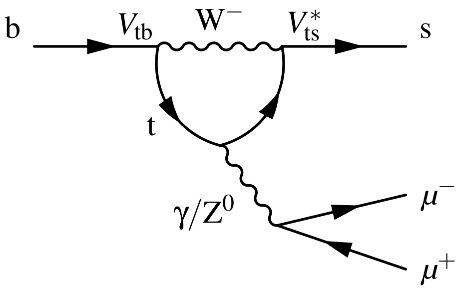


Figure 2: Lowest order Feynman diagram of an FCNC interaction via electroweak penguin diagram.

In this paper, the $\Lambda_b^0 \to \Lambda^0\mu^+\mu^-$ decay at a quark-level is an example of a FCNC, where the transition is — as visualized in Figure 2 — a $b \to s$ transition. This decay cannot be described by a tree-level diagram and must instead be explained by a penguin diagram, which is rarer due to the presence of virtual particles. [17]. Moreover, the decay may exhibit CP violation due to the quark transitions governed by the CKM matrix. However, the CP violation in this decay channel is expected to be extremely small, likely similar to the suppressed asymmetries observed in neutral B-meson systems, given the shared underlying $b \to s$ transition. For reference, the Standard Model predicts semi-leptonic CP asymmetries of $a_{sl}^d = -(5.1 \pm 0.5) \times 10^{-4}$ for B_d^0 mesons, and $a_{sl}^s = +(2.2 \pm 0.2) \times 10^{-5}$ for B_s^0 mesons [18].

2.3 The Λ_b^0 decay

Flavor physics represents a key area in the search for physics BSM. In particular, rare particle decays are of great interest, as they are strongly suppressed within the SM and thus more sensitive to potential contributions from BSM processes. Among these, FCNC interactions are especially promising, since they are forbidden at tree level in the SM and can only occur through higher-order loop diagrams. Mesonic FCNC, especially of the B meson have been analyzed in depth to search BSM physics [19,20]. Meanwhile, baryonic rare decays have not been explored to the same extent. However, baryonic decays have several distinct advantages over mesonic decays, such as being able to involve diquark system as a spectator leading to different hadronic physics. Hence, providing crucial complementary information for the mesonic processes [21]. The research sector has therefore started to investigate the baryonic decays further, especially the Λ_b baryon due to its heavy b quark.

It is for this reason that the $\Lambda_b^0 \to \Lambda^0\mu^+\mu^-$ decay is researched. It involves a FCNC decay in the form of $b \to s$ as mentioned in the previous section. Adding to the penguin diagram in Figure 2, the Feynman diagram for the entire decay is seen in Figure 3. Due to the rarity of this decay, similar channels such as $\Lambda_b^0 \to \Lambda^0(\to p^+\pi^-)J/\psi(\to \mu^+\mu^-)$ are also studied to improve our understanding of the $\Lambda_b^0 \to \Lambda^0\mu^+\mu^-$ process and support the analysis.

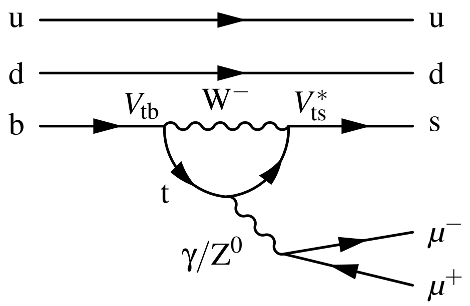


Figure 3: Lowest order Feynman diagram of the $\Lambda_b \to \Lambda\mu^+\mu^-$ decay.

The advantages of the $\Lambda_b^0 \to \Lambda^0 J/\psi$ decay mode is that it is less rare, see Table 1. The secondary decay $\Lambda^0 \to p^+\pi^-$ enables clearer distinction between matter and anti matter, as the $\bar{\Lambda}_b$ decay's secondary decay becomes $\bar{\Lambda}^0 \to p^-\pi^+$, helping differentiate with greater ease between the Λ_b and $\bar{\Lambda}_b$ decays, enabling the calculation of the raw asymmetry between them.

Decay mode	**Branching fraction**	q^2 **region**
$\Lambda_b^0 \to \Lambda\, J/\psi$	$(3.08 \pm 0.26 \pm 0.12) \times 10^{-4}$	–
$\Lambda_b^0 \to \Lambda^0\mu^+\mu^-$	$(6.9 \pm 1.3) \times 10^{-7}$	$16.0 < q^2 < 20.0\ \mathrm{GeV}^2/c^4$

Table 1: Relevant decay modes and branching fractions for Λ_b^0 decays [1,22].

Particle	**Quark content**	**Mass (MeV)**	**Lifetime (s)**	**Charge**
Λ_b^0	udb	5619.60 ± 0.17	$(1.42 \pm 0.03 \pm 0.01) \times 10^{-12}$	0
Λ	uds	1115.683 ± 0.006	$(2.632 \pm 0.020) \times 10^{-10}$	0
J/ψ	$c\bar{c}$	3096.900 ± 0.006	$(7.2 \pm 0.1) \times 10^{-21}$	0
μ^-	–	$105.6583755 \pm 0.0000023$	2.1969811×10^{-6}	–
p	uud	$938.27208816 \pm 0.00000029$	Stable	+
π^-	$u\bar{d}$	139.57039 ± 0.00018	$(2.6033 \pm 0.0005) \times 10^{-8}$	–

Table 2: Properties of relevant particles in the Λ_b^0 decay [1].

Additional properties of the relevant particles for this thesis are presented in Table 2. Helping to understand the Feynman diagram in Figure 3. Additionally, the masses of the Λ_b^0 and pion particles are going to be relevant for the results.

2.3.1 Reconstruction of Λ_b^0 candidates

A crucial step in this analysis is the reconstruction of Λ_b^0 candidates from their final-state decay products. At the LHCb detector, only the stable, charged particles, such as muons, protons, and pions are directly measured. This is done based on their charge, momentum, and responses in the detector. The decay topology of the Λ_b^0 baryon is illustrated in Figure 4. It decays into a Λ^0 baryon and a pair of oppositely charged muons. The Λ^0 subsequently decays into a proton and a pion. The four-momentum of the Λ^0 can be reconstructed from its decay products:

$$p_{\Lambda^0} = p_p + p_\pi$$

This is then combined with the four-momenta of the muons to reconstruct the full Λ_b^0 candidate:

$$p_{\Lambda_b^0} = p_{\Lambda^0} + p_{\mu^+} + p_{\mu^-}$$

From this, the invariant mass is calculated to identify signal candidates consistent with the known Λ_b^0 mass. The displaced vertices of the Λ^0 and dimuon system from the primary vertex (PV) help to suppress background from prompt decays.

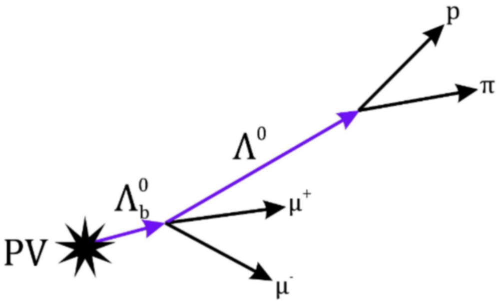


Figure 4: Illustration of the decay topology for the Λ_b^0 baryon, starting from the primary vertex (PV) and ending with the final-state decay products.

3 The LHC and LHCb

3.1 The LHC

The Large Hadron Collider (LHC) [23] is the world's largest and most powerful particle accelerator, located at CERN near Geneva, Switzerland. It consists of a 27 km circular ring equipped with superconducting magnets and accelerating structures designed to boost protons to extremely high energies. During Run 2, which spanned from 2015 to 2018, the LHC operated at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ [24]. Among its four main detectors, the LHCb (Large Hadron Collider beauty) experiment is uniquely optimized for studying the decays and CP violation of hadrons containing b and c quarks. Unlike the general-purpose detectors ATLAS and CMS, LHCb is designed as a forward spectrometer, which allows it to efficiently collect and analyze the products of heavy flavor hadron decays.

3.2 The LHCb detector

The LHCb detector [25] employs a unique forward spectrometer design that covers the pseudorapidity range ($2 < \eta < 5$) most favored for $b\bar{b}$ production in high-energy proton-proton (pp) collisions [26]. This asymmetric geometry significantly enhances the collection efficiency of events involving b-hadrons.

Throughout its operation, LHCb has seen continuous upgrades to improve its performance, particularly in light of increasing luminosity demands. These upgrades have enhanced tracking resolution, improved particle identification, and increased the data acquisition capability of the detector. The detector itself consists of multiple subsystems, each optimized for a specific role in reconstructing particle decays. These include precision tracking detectors for measuring particle trajectories, dedicated systems for identifying hadrons, electrons, and muons, and calorimeters for energy measurement. Together, they enable detailed reconstruction of decay vertices, particle types, and kinematics, which are essential for comparing experimental results with theoretical predictions. An overview of the LHCb detector layout is illustrated in Figure 5.
The innermost subdetector of the LHCb experiment is the Vertex Locator (VELO), which plays a crucial role in reconstructing the decay vertices of short-lived particles. Located immediately around the interaction point, the VELO allows for the precise measurement of the displacement between secondary decay vertices and the primary vertex. It consists of 21 semi-circular silicon modules aligned along the beam (z) axis. Each module is equipped with sensors that measure either the radial distance r or the azimuthal angle ϕ in the plane perpendicular to the beamline. Charged particles travels across these modules, producing hits in the silicon strips. By combining successive hits across multiple modules, tracks of charged particles can be reconstructed. These tracks are then extrapolated back to their point of origin, enabling precise vertex reconstruction and the identification of short-lived particles such as the Λ_b^0.

Particles that decay beyond the VELO are tracked by the rest of the LHCb tracking system, which includes the Tracker Turicensis (TT) upstream of a dipole magnet, and three downstream tracking stations (T1, T2, T3). The dipole magnet can operate in two field polarities: Magnet Up (MU) and Magnet Down (MD), which bend charged particles in opposite vertical directions. This polarity is regularly alternated to help cancel out asymmetries from the detector. Charged particles first leave hits in the TT and then bend in the magnetic field, depending on their momentum and charge. As

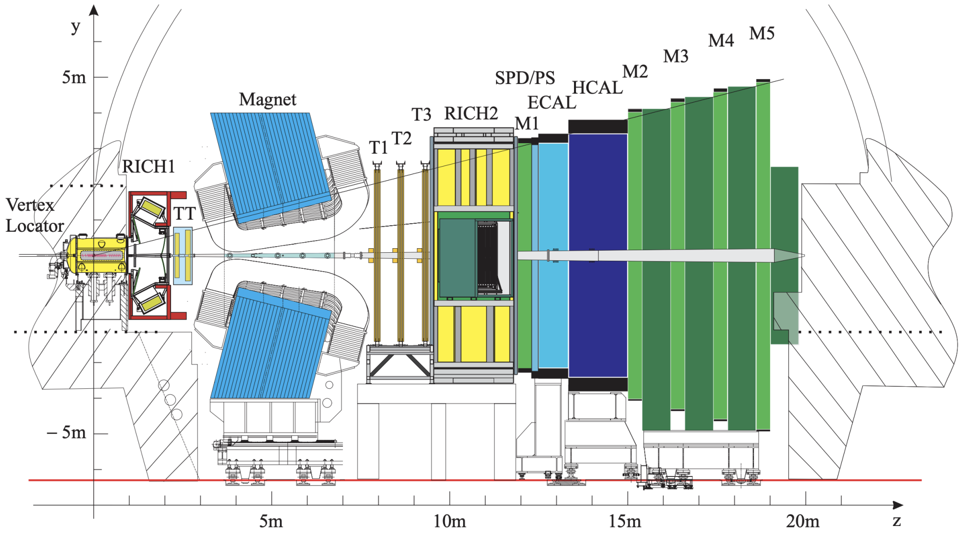


Figure 5: Schematic of LHCb detector layout as per Run 2 [25]

they continue, they produce further hits in the downstream tracking stations. The full track can be reconstructed using this information, providing both the trajectory and momentum of the particle. Tracks that include VELO and downstream detector hits are known as long tracks (LL), while tracks originating outside the VELO are reconstructed from only the TT and T-stations and are called downstream tracks (DD), see Figure 6. In the context of the $\Lambda_b^0 \to \Lambda^0(\to p\pi^-)J/\psi(\to \mu^+\mu^-)$ decay, the decay products of the Λ^0 baryon can fall into either category depending on whether its decay occurs inside or outside the VELO. As such, data for the analysis are usually split into LL and DD categories to account for these differences in reconstruction. In this analysis they will be combined, this will be further discussed in the Results section.

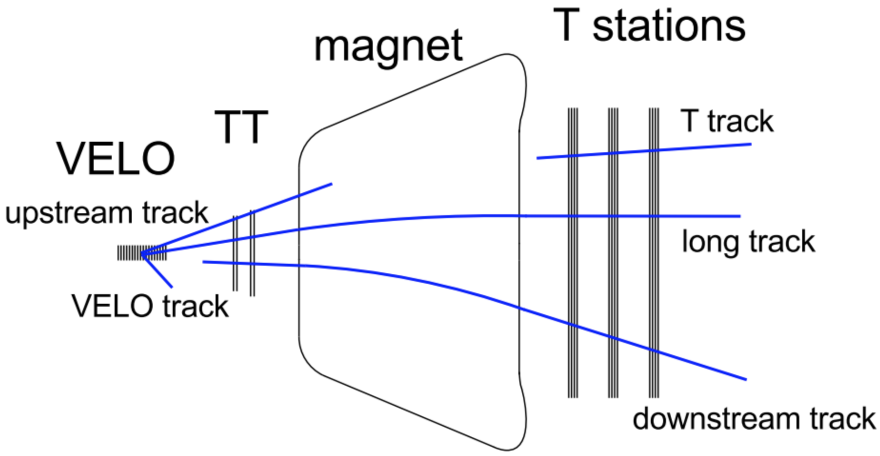


Figure 6: Reconstruction of track types in tracking detector at LHCb [26]

3.3 Particle Identification in LHCb

There are four main detector systems responsible for particle identification (PID): The two RICH detectors, the calorimeter systems, and the muon system. These detector systems present information about the particle, in addition to the reconstruction of momenta and trajectory of a particle, it successfully identifies particles.

Among the PID systems, the two Ring Imaging Cherenkov (RICH) detectors main function is to identify different species of charged hadrons, such as pions, kaons, and protons. Positioned between the tracking stations, the RICH detectors operate based on the Cherenkov effect: when a charged particle travels through a medium at a speed greater than the phase velocity of light in that medium, it emits Cherenkov radiation at a characteristic angle. This Cherenkov angle is directly related to the particle's velocity. By combining this velocity information with momentum measurements obtained from the tracking system, the mass of the particle can be obtained. This allows the RICH system to provide separation of particle types across a broad momentum range. However, in certain overlapping regions of momentum, the Cherenkov angles for different particles can be similar, which may lead to a misidentification of the particle. These misidentifications between hadrons are common, such as misidentifying kaons with pions. This can contribute to background in analyses and affect precision measurements, especially for the Λ_b^0 decay analyzed in this thesis.

The calorimeter system at LHCb is composed of two distinct components: the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL). Both are positioned downstream of the RICH detectors and measure the energy of the particles. These calorimeters consist of alternating layers of dense absorber material and scintillators. When a particle enters the calorimeter, it initiates a cascade of secondary particles (electromagnetic or hadronic showers) depending on the nature of the incident particle. As these showers pass through the scintillating material, they produce flashes of light, which are then collected by photodetectors. The total amount of detected light is proportional to the particle's energy and serves as the basis for identifying the type of particle.

The ECAL is optimized for the identification of electrons and photons, which produce well-contained electromagnetic showers. The HCAL, with thicker absorber material, is suited for detecting hadrons such as pions and protons, which undergo strong interactions. By analyzing the energy deposition patterns in both calorimeters, it is possible to discriminate between particle types. Muons, due to their minimal interaction with matter, pass through both calorimeters with little energy loss, distinguishing them from hadrons and electrons. Instead, muons are primarily identified in the final PID system: the dedicated muon chambers, located at the back of the detector, which provide independent confirmation of a particle's identity through a separate tracking system.

4 Analysis and Results

This chapter presents the methodology and outcomes of the data analysis conducted to investigate the raw asymmetry between Λ_b^0 and $\bar{\Lambda}_b^0$ decays in the high-q^2 region. It outlines the selection criteria applied to suppress background and enhance signal purity, the reasoning for combining different track types to increase statistical precision, and the modeling strategies used to extract signal yields. The chapter concludes with the calculation of the raw asymmetry and its statistical significance, addressing the central research question regarding potential matter-antimatter imbalance at LHCb.

To analyze the data collected at the LHCb detector, the `uproot` package [27] in Python was used. `Uproot` is specifically designed to interact with ROOT files, the standard data format used at CERN. The analysis was performed on proton-proton collision data collected during Run 2 of the LHCb (2016–2018) at $\sqrt{s} = 13$ TeV, corresponding to an integrated luminosity of approximately 6 fb^{-1} [13]. Although Run 2 includes data from 2015, that year was excluded from the analysis due to its relatively low statistics. As a result, only data from 2016 to 2018 were considered. Furthermore, events corresponding to both long tracks (LL) and downstream tracks (DD) were combined to increase the available statistics, this decision and its implications are discussed in the following section.

The core of the analysis involves modeling both the signal and background components in the reconstructed invariant mass spectrum of Λ_b^0, and extracting the number of signal events separately for Λ_b^0 and $\bar{\Lambda}_b^0$, as well as for each magnetic field configuration (Magnet Up and Magnet Down). The final quantity of interest is the raw asymmetry between Λ_b^0 and $\bar{\Lambda}_b^0$ decays, obtained by combining the MU and MD samples to maximize statistical significance.

4.1 Selection criteria applied to Λ_b^0 candidates

A vital step in analyzing the Λ_b^0 mass spectrum is the suppression of background from misidentified particles. To achieve this, specific selection criteria (cuts) are applied to the dataset to isolate appropriate Λ_b^0 candidates.

Before applying these selections, the reconstruction of the full decay chain of the Λ_b^0 candidate is performed using the decay tree fitter (DTF) [28]. This algorithm reconstructs the decay under specific constraints, such as requiring that the Λ_b^0 originates from the primary vertex (where the proton-proton collision occurs). Additionally, it fixes the mass of the intermediate Λ^0 baryon to its known value from the Particle Data Group (PDG). Using the DTF mass improves both spatial and momentum resolution, resulting in a sharper invariant mass peak and better separation from background. This is especially important in regions with low signal-to-background ratios, such as the high-q^2 region.

One important selection involves the pseudorapidity, η, a commonly used spatial variable that describes the angle of a particle relative to the beam (z) axis. The LHCb detector is designed to cover the pseudorapidity range $2 < \eta < 5$ [13], so the data is restricted to this interval. Events outside this range fall beyond the detector's acceptance and therefore cannot be reliably reconstructed or analyzed.

To further reduce the background, selections in the transverse momentum, p_T, were applied. A minimum threshold of $p_T > 2\,\text{GeV}/c$ was imposed on the Λ_b^0 candidates. This cut suppresses low-momentum background that is more likely to originate from random combinatorial sources—random combinations of particles that can be misidentified—or secondary interactions with detector material. This ensures that the selected candidates lie within regions where the trigger and reconstruction efficiencies are well understood and is consistent with the thresholds commonly used in the LHCb trigger system. The choice of this specific threshold is also consistent with previous studies [29], which applied a similar cut for consistent candidate selection in Λ_b^0 decays.

Another key variable used in this selection is q^2, the square of the dimuon invariant mass. Since this thesis focuses on the high-q^2 region, defined as $q^2 > 14.3\,\text{GeV}^2/c^4$, which lies above the $\psi(2S)$ resonance, this threshold was chosen to ensure consistency with prior research [30]. Events below this value are therefore excluded from the analysis. This selection is illustrated in Figure 7, where a distinct peak corresponding to the $\psi(2S)$ resonance can be seen just below the applied cut. This cut allows the analysis to concentrate on the high-q^2 regime, where potential deviations from Standard Model predictions may manifest differently compared to lower q^2 regions.

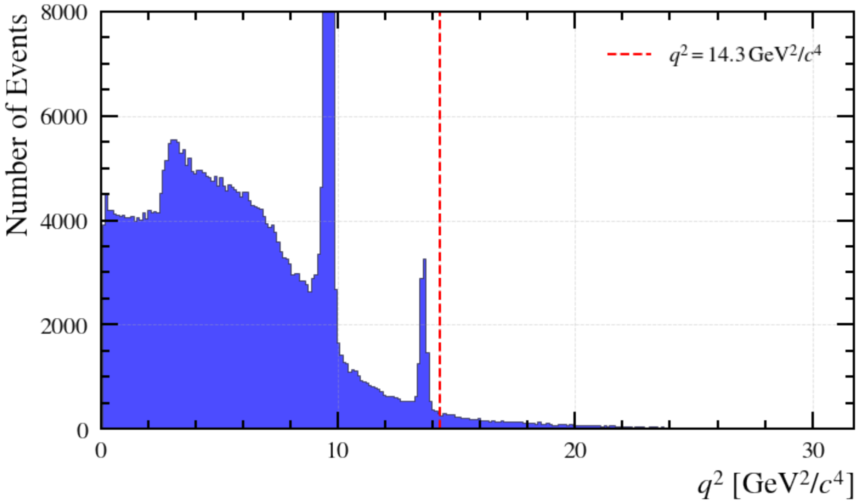


Figure 7: Distribution of the dimuon invariant mass squared (q^2) for selected $\Lambda_b^0 \to \Lambda^0\mu^+\mu^-$ candidates. The vertical dashed red line at $q^2 = 14.3\,\text{GeV}^2/c^4$ marks the lower boundary of the high-q^2 region. The prominent peak just before the threshold corresponds to the $\psi(2S)$ resonance.

Another background that can contaminate the Λ_b^0 candidate sample originates from the decay $B^0 \to K_S^0\mu^+\mu^-$, where the K_S^0 subsequently decays into two pions. In certain cases, one of these pions can be misidentified as a proton, leading the $K_S^0 \to \pi^+\pi^-$ decay to be reconstructed as a $\Lambda^0 \to p^+\pi^-$ decay. This misidentification creates a false Λ_b^0 candidate and contributes to a small bump in the invariant mass spectrum slightly below the true Λ_b^0 mass. To suppress this specific background, a K_S^0 mass veto is applied by reconstructing the invariant mass of the pion-proton pair under the assumption that both tracks are actually pions, using the following formula:

$$m_{K_S^0} = \sqrt{(E_\pi + E_p)^2 - (\vec{p}_\pi + \vec{p}_p)^2}, \quad \text{with} \quad E_p = \sqrt{m_\pi^2 + |\vec{p}_p|^2} \tag{2}$$

where $m_{K_S^0}$ is the invariant mass, E_π is the pion energy, $\vec{p}_\pi$ is the pion momentum, $\vec{p}_p$ is the proton momentum, E_p is the energy of the proton treated as a pion, and m_π is the pion mass.

If the resulting invariant mass lies close to the known K_S^0 mass, the event is excluded. Specifically, candidates with a reconstructed K_S^0 mass in the window $470 < m_{K_S^0} < 520\,\text{MeV}/c^2$ are vetoed. This approach effectively removes events that likely come from misidentified K_S^0 decays, thereby improving the accuracy of the Λ_b^0 candidates. The pion mass, m_π is given by 139.57 MeV/c^2 [1].

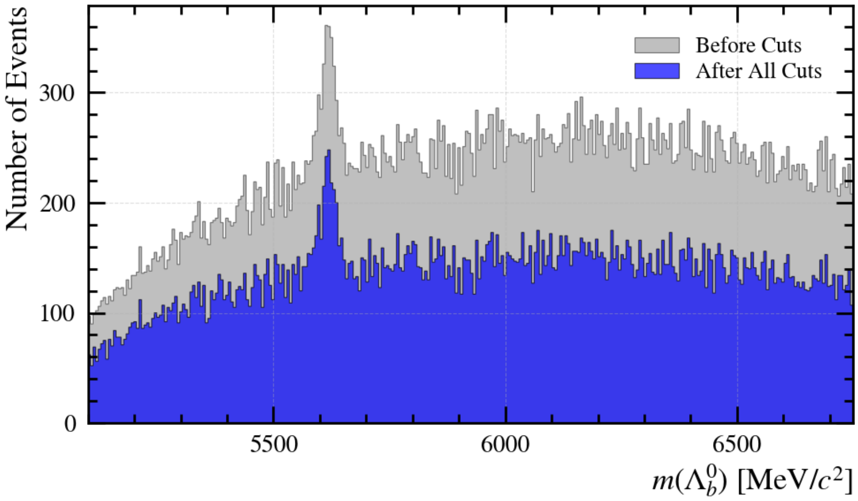


Figure 8: Reconstructed Λ_b^0 candidate mass before and after selection cuts. Gray: events with only $q^2 > 14.3, \text{GeV}^2/c^4$; Blue: events after applying q^2, η, p_T, and K_S^0 veto cuts.

Applying the selection criteria to the data results in a significantly lower combinatorial background, visualized in Figure 8. It must be taken into account that the combinatorial background is also present in the peak and is not necessarily removing signal, although the peak gets slightly reduced due to the cuts.

Lastly, for the final results, the separation between matter and antimatter is vital to answer the research question: Can asymmetries in the raw production between matter and antimatter in high-q^2 regions reveal deviations from Standard Model predictions? To distinguish between Λ_b^0 and $\bar{\Lambda}_b^0$ candidates in the dataset, particle identification was performed based on the charge of the daughter particles from the Λ^0 decay. Specifically, the sign of the proton and pion IDs was used: a Λ^0 decay (and by extension a Λ_b^0 candidate) is identified when the proton candidate has a positive charge and the pion candidate has a negative one. Conversely, the opposite charge configuration corresponds to a $\bar{\Lambda}_b^0$ candidate.

This charge-based separation ensures a clean distinction between matter and antimatter candidates, which is essential for any measurement of raw asymmetries. By applying this criterion along with the full set of cuts—including those on q^2, pseudorapidity, transverse momentum, and a K_S^0 veto—the final dataset is well-defined for investigating potential asymmetries in the high-q^2 regime. All the previously mentioned cuts can be summarized in Table 3.

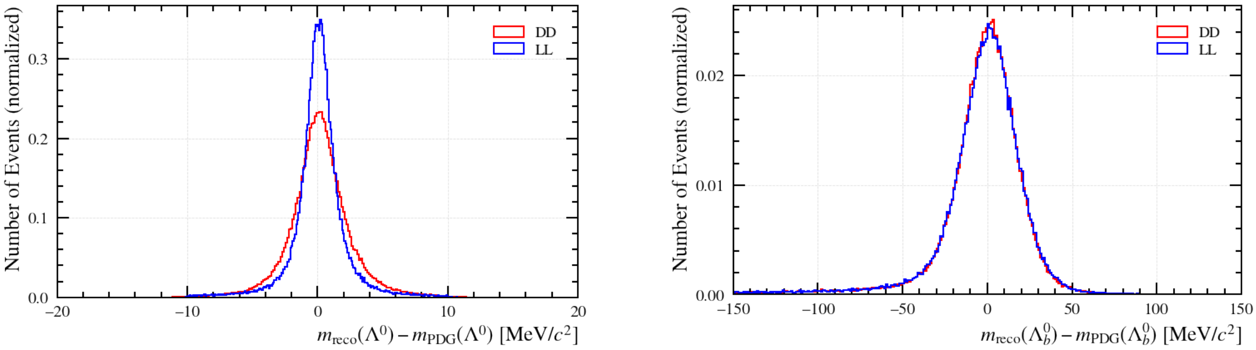
Variable	**Cut**	**Explanation**
q^2	> 14.3 GeV$^2/c^4$	high-q^2 region
η	(2, 5)	detector acceptance
p_T	> 2 GeV/c	suppresses combinatorial background
$m_{K_s^0}$	(470, 520) MeV/c^2	eliminating peaking background
π_{ID}	> 0 or < 0	differentiation between matter and antimatter
p_{ID}	> 0 or < 0	differentiation between matter and antimatter
$m_{\Lambda_b^0}$	(5100, 6800) MeV/c^2	region of interest

Table 3: Summary of all selected cuts applied to the data during the analysis.

4.2 Combination of LL and DD datasets

A recurring challenge in the analysis was the limited number of events in the high-q^2 region. To increase statistical power, both Long (LL) and Downstream (DD) track types were considered for inclusion. However, before combining these categories, it was necessary to verify that LL and DD tracks yield comparable reconstruction performance.

To assess this, the mass resolution of the Λ^0 and Λ_b^0 baryons was studied using simulated signal events. The Λ^0 particle, reconstructed from a proton and pion, is particularly sensitive to LL/DD differences and thus serves as a good benchmark for comparing reconstruction quality. The resolution was calculated through the mass residuals $(m_{\mathrm{reco}} - m_{\mathrm{PDG}})$. This was plotted separately for LL and DD candidates, as shown in Figure 9a. The resolution of the full Λ_b^0 mass was also examined in Figure 9b.



(a) Λ^0 mass resolution

(b) Λ_b^0 mass resolution

Figure 9: Normalized distribution of the (a) Λ^0 mass resolution and (b) Λ_b^0 mass resolution for Long (LL) and Downstream (DD) tracks using simulated signal events. The resolution is defined as the difference between the reconstructed mass and the corresponding known value from the Particle Data Group (PDG) [1] for (a) Λ^0 and (b) Λ_b^0.

As seen in Figure 9a, there is a small difference in the Λ^0 mass resolution between DD and LL tracks. However, this difference is moderate and does not significantly impact the overall resolution. More importantly, Figure 9b shows that the Λ_b^0 mass resolution is substantially broader than that of the Λ^0, indicating that the dominant contribution to the Λ_b^0 resolution comes from the dimuon system. Since the Λ^0 mass is tightly constrained in the decay tree fit (DTF) using its known PDG value, minor differences between LL and DD tracks are largely corrected.

Therefore, the combination of LL and DD tracks is well-justified. It improves statistical precision while maintaining consistent mass resolution across track types, ensuring no degradation of the final observable.

4.3 Data Modeling

To extract the raw asymmetry between matter (Λ_b^0) and antimatter ($\bar{\Lambda}_b^0$) decays, an accurate modeling of the reconstructed mass distribution is essential. The modeling process is divided into three main components: the signal model, the background model, and the combined model used to fit the invariant mass spectrum of the Λ_b^0 candidates.

4.3.1 Signal Modeling

To model the signal, Monte Carlo (MC) simulated data was used. Although the same reconstruction algorithms are applied as in real data (misreconstruction can still occur), MC provides the advantage that the true origin of each particle is known. This makes it possible to build a clean and reliable template for the expected signal shape in data, using only events that are known to be genuine signal.

The distribution of the reconstructed Λ_b^0 mass from the MC simulation exhibits a Gaussian-like peak with asymmetric tails. To model this shape, three candidate functions were considered: a Gaussian, a single-sided Crystal Ball, and a double-sided Crystal Ball (DSCB) function. Each was fitted to the MC data, with the DSCB providing the best fit based on the reduced χ^2 value. The DSCB function — widely used in high-energy physics — combines a Gaussian core with power-law tails on both sides. This allows it to model deviations from a purely Gaussian shape caused by detector resolution effects or reconstruction imperfections. Its functional form is:

$$f(m;\mu,\sigma,\alpha_L,n_L,\alpha_R,n_R) = \begin{cases} A_L\left(B_L - \frac{x-\mu}{\sigma}\right)^{-n_L}, & \text{if } \frac{x-\mu}{\sigma} < -\alpha_L \\ \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right), & \text{if } -\alpha_L \le \frac{x-\mu}{\sigma} \le \alpha_R \\ A_R\left(B_R + \frac{x-\mu}{\sigma}\right)^{-n_R}, & \text{if } \frac{x-\mu}{\sigma} > \alpha_R \end{cases} \tag{3}$$

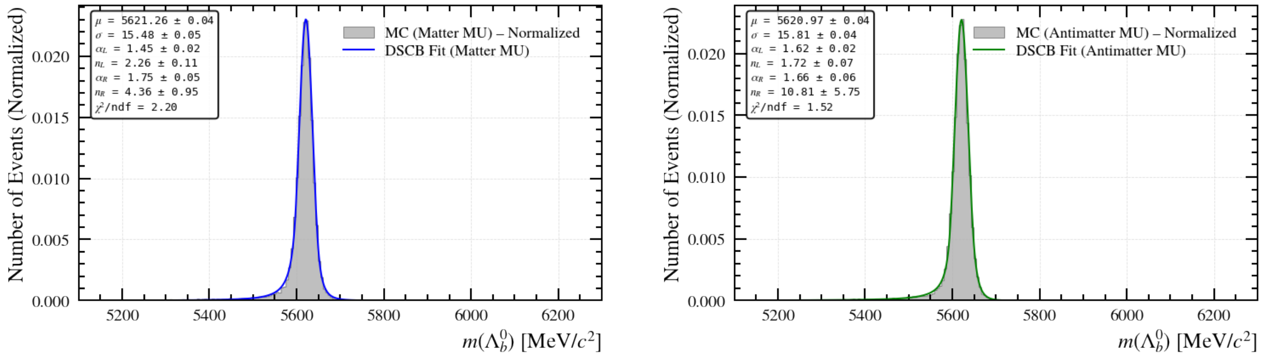
where:

$$B_L = \frac{n_L}{|\alpha_L|} - |\alpha_L|, \qquad A_L = \left(\frac{n_L}{|\alpha_L|}\right)^{n_L} \exp\left(-\frac{|\alpha_L|^2}{2}\right),$$

$$B_R = \frac{n_R}{|\alpha_R|} - |\alpha_R|, \qquad A_R = \left(\frac{n_R}{|\alpha_R|}\right)^{n_R} \exp\left(-\frac{|\alpha_R|^2}{2}\right).$$

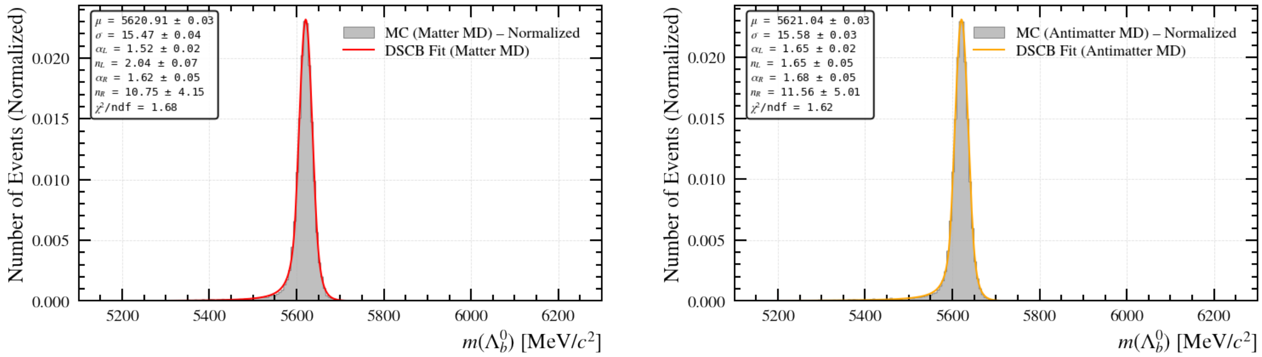
With m being the Λ_b^0 mass , μ the mean and σ is the standard deviation of the gaussian function, α_L and α_R are the left and right transition points, n_L and n_R are the left and right tail parameters, A_L and A_R are normalization constants, and B_L, B_R are shape parameters.

Since the Double-Sided Crystal Ball is used as a probability density function (PDF), its integral over the full domain must equal 1, and hence, it is normalized. Therefore, it was fitted to the normalized MC simulation data. This fitting was performed separately for four categories: Matter selection with Magnet Up (MU), Matter selection with Magnet Down (MD), Antimatter selection with Magnet Up, and Antimatter selection with Magnet Down. The signal model is illustrated in Figure 10.



(a) Matter with Magnet Up (MU)

(b) Antimatter with Magnet Up (MU)



(c) Matter with Magnet Down (MD)

(d) Antimatter with Magnet Down (MD)

Figure 10: Fits to signal MC data using the double-sided Crystal Ball function for the four data categories: matter/antimatter and Magnet Up/Down.

The parameters extracted from the DSCB fits shown in Figure 10 provide an accurate description of the signal shape for each data category (Matter/Antimatter and MU/MD). These fitted parameters will be used as fixed inputs (apart from the mean and standard deviation) for the signal component in the combined fit to the real data, as discussed in Section 4.3.3, more clearly shown in Table 4.

Parameter	Matter MU	Matter MD	Antimatter MU	Antimatter MD
μ [MeV/c^2]	5621.26 ± 0.04	5620.91 ± 0.03	5620.97 ± 0.04	5621.04 ± 0.03
σ [MeV/c^2]	15.48 ± 0.05	15.47 ± 0.04	15.81 ± 0.05	15.58 ± 0.03
α_L	1.45 ± 0.02	1.52 ± 0.02	1.62 ± 0.02	1.65 ± 0.02
n_L	2.26 ± 0.11	2.04 ± 0.07	1.72 ± 0.07	1.65 ± 0.05
α_R	1.75 ± 0.05	1.62 ± 0.05	1.66 ± 0.06	1.68 ± 0.05
n_R	4.36 ± 0.95	10.75 ± 4.15	10.81 ± 5.75	11.56 ± 5.01

Table 4: Fitted parameters of the Double-Sided Crystal Ball (DSCB) function, with uncertainties.

4.3.2 Background Modeling

The background in the invariant Λ_b^0 mass distribution is primarily composed of combinatorial background — random combinations of tracks that mimic signal decays. To model this background, a sidebands method was employed. This approach involves selecting regions of the mass spectrum that lie well outside the expected signal peak, where contributions from true Λ_b^0 decays are negligible. By analyzing these sideband regions, the shape and characteristics of the background can be determined. This procedure enables an accurate estimation of the background contribution in the signal region, as illustrated in Figure 11.

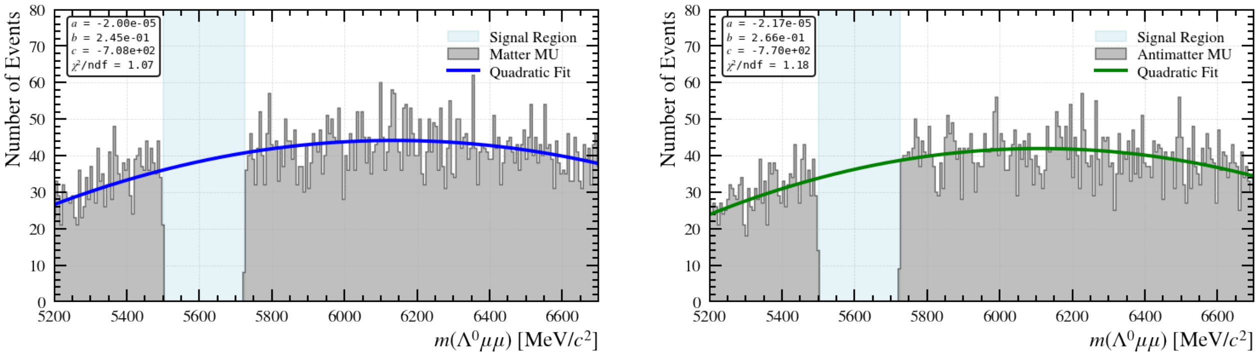
The sidebands were established by setting a veto region around the Λ_b^0 mass peak. Specifically, the signal region $(5500, 5725)\,\mathrm{MeV}/c^2$ was excluded to ensure that the background model was constructed solely from data outside the signal peak. This approach minimizes signal contamination in the background fit. To describe the background distribution, a second-degree polynomial was chosen in the form:

$$f(m) = ax^2 + bx + c \tag{4}$$

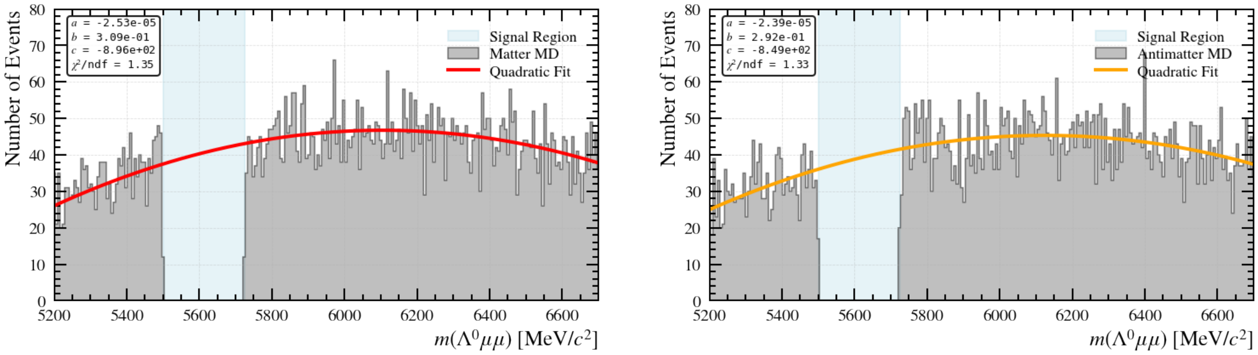
The choice of a quadratic function was motivated by the observed curvature of the background in the sidebands, as well as prior studies of backgrounds in high-q^2 regions in Λ_b^0 decays [30]. The goodness of each fit was evaluated using the reduced chi-square (χ^2/ndf) value, providing a quantitative measure of how well the function matched the background shape in each of the four data categories.

Parameter	Matter MU	Matter MD	Antimatter MU	Antimatter MD
a [10^{-5}]	-2.00 ± 0.20	-2.53 ± 0.21	-2.17 ± 0.19	-2.39 ± 0.22
b	0.245 ± 0.024	0.309 ± 0.025	0.266 ± 0.023	0.292 ± 0.026
c	-708 ± 70	-896 ± 73	-770 ± 67	-849 ± 76

Table 5: Fitted parameters of the quadratic background function for each data category. The Coefficient a has been rescaled to units of 10^{-5} for readability.



(a) Matter with Magnet Up (MU) (b) Antimatter with Magnet Up (MU)



(c) Matter with Magnet Down (MD) (d) Antimatter with Magnet Down (MD)

Figure 11: Quadratic fit to the sideband regions of the invariant mass distribution for each of the four data categories: matter/antimatter and Magnet Up/Down.

4.3.3 Combined signal and background model

This section is crucial to obtain the number of signal events for all data categories. To perform this fit, the invariant mass distribution was modeled using a combined function consisting of a signal component and a background component. The signal shape was described using a Double-Sided Crystal Ball (DSCB) function, while the background was modeled with a second-order polynomial. The background parameters were previously obtained (see Table 5) by fitting the sideband regions of the mass spectrum and were kept fixed during the final fit. Similarly, the DSCB tail parameters (α_L, n_L, α_R, n_R) were obtained (see Table 4) from fits to Monte Carlo simulated signal data and were also fixed. This approach reduces the number of free parameters in the final fit to just three: the signal yield normalization (N_{sig}), the Gaussian mean (μ), and the width (σ). Fixing the background and DSCB tail parameters improves the stability and robustness of the fit, especially in regions with limited statistics.

It is important to emphasize that the signal shape consists of a probability density function (PDF), as discussed in Section 4.3.1. Since the integral over all its domain is equal to one, it does not account for the absolute number of signal events in the dataset. Therefore, when fitting the PDF to real data, a normalization factor is required. This factor is introduced as a free parameter in the fit and corresponds to the total number of signal events, denoted as N. However, because the fit is performed on a binned histogram, the actual number of events depends on the bin width. To correctly interpret N_{sig}

as the number of signal events, N, it must be divided by the bin width. The number of signal events, N, is then used to calculate the raw asymmetry, which will be promptly discussed.

Firstly the number of signal events for each of the four data categories (matter MU/MD, and antimatter MU/MD) was calculated from the combined model shown in Figure 12. The parameters obtained from the fits to each data category are summarized in Table 6.

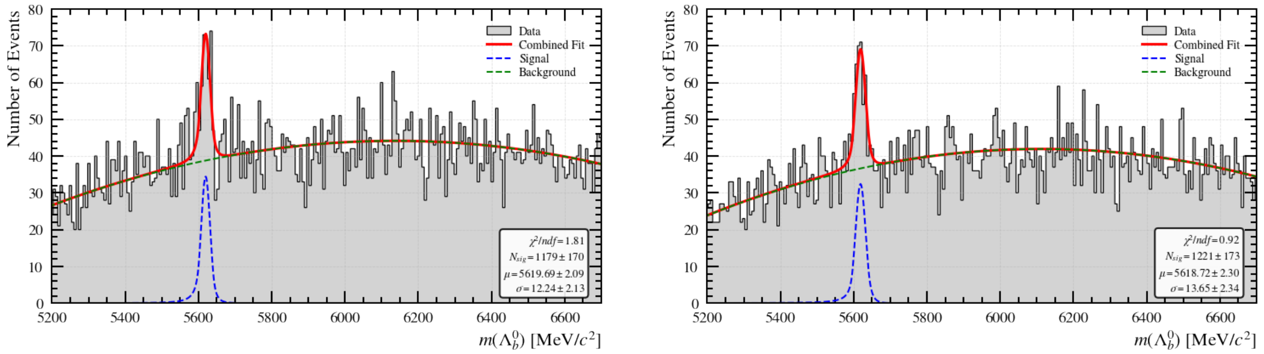
Category	N_{sig}	μ [MeV/c^2]	σ [MeV/c^2]	N
Matter MU	1170 ± 170	5619.7 ± 2.09	12.2 ± 2.13	185 ± 26.7
Matter MD	1220 ± 207	5620.9 ± 3.84	19.4 ± 4.11	191 ± 32.4
Antimatter MU	1220 ± 173	5618.7 ± 2.30	13.6 ± 2.34	191 ± 27.2
Antimatter MD	1280 ± 196	5621.4 ± 3.18	17.6 ± 3.25	201 ± 30.7

Table 6: Fit results for the normalization factor, N_{sig}, peak position μ, and width σ for each data category. The number of signal events, N, was obtained by dividing N_{sig} by the histogram bin width, to account for binning dependence.

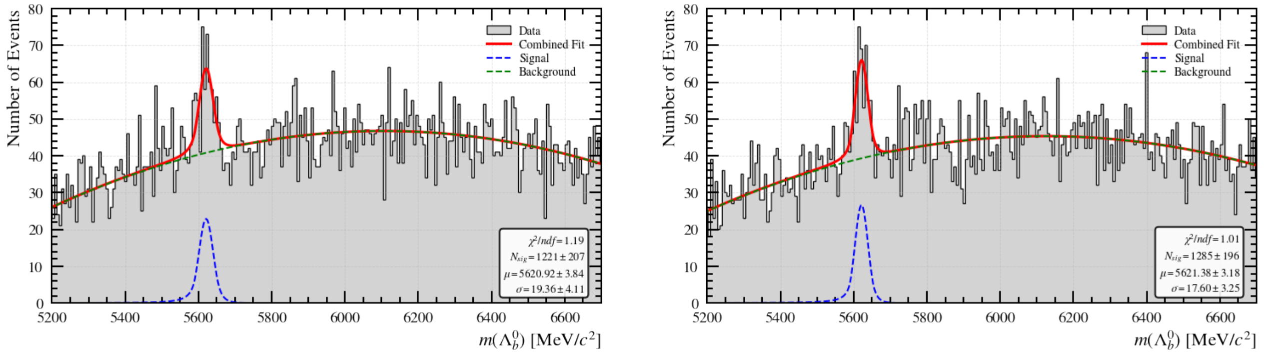
When comparing the signal yields between the MU and MD polarization modes, it is important to account for differences in the integrated luminosity across the datasets. Although both MU and MD datasets span the same data-taking periods, their respective luminosity weights differ slightly due to variations in detector conditions and trigger configurations. To allow for a fair comparison, the raw number of signal events N in each category is divided by the corresponding effective luminosity, as shown in Table 7. This normalization procedure removes the dependence on luminosity, making the yields directly comparable between polarizations.

Datasets	**Signal Events**	**Weighted Signal Events**	**Eff. Lum. Weight**	**Lum. [pb^{-1}]**
Matter MU	187.4 ± 27.1	602.7 ± 87.1	0.3110	2705.6
Matter MD	191.8 ± 32.5	611.7 ± 103.5	0.3136	2728.4
Antimatter MU	194.6 ± 27.6	625.9 ± 88.9	0.3110	2705.6
Antimatter MD	201.3 ± 30.7	641.9 ± 97.8	0.3136	2728.4

Table 7: Raw and luminosity-normalized signal yields for each dataset category. The effective luminosity is computed using the normalized weights from each year of data-taking. The Luminosity value for each magnet polarization was obtained by summing the 2016, 2017, and 2018 contributions.



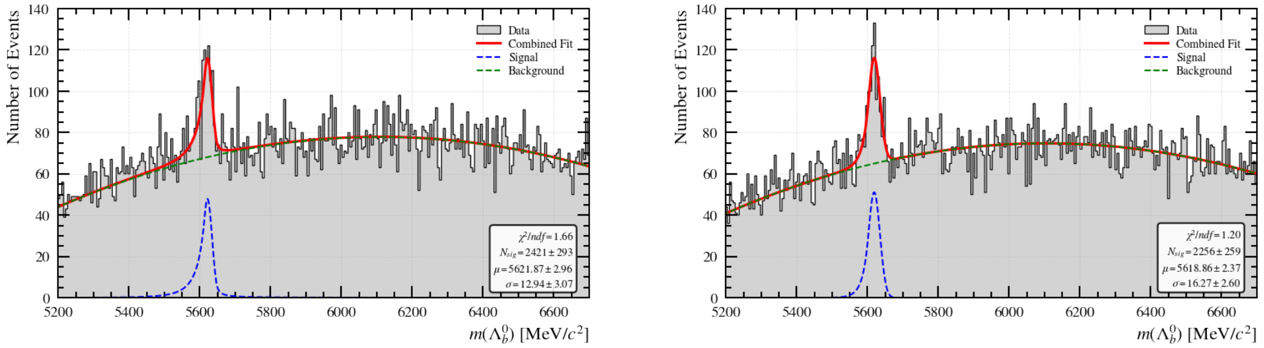
(a) Matter with Magnet Up (MU) (b) Antimatter with Magnet Up (MU)



(c) Matter with Magnet Down (MD) (d) Antimatter with Magnet Down (MD)

Figure 12: Combined signal and background fit for each of the four data categories: matter/antimatter and Magnet Up/Down. The signal is represented by the dotted blue line, the background by the dotted green line, and the combined by the red line. N_{sig} representing the signal normalization factor. The mean, (μ), standard deviation, σ, and normalization factor, N_{sig}, of the signal were set as free parameters. The remaining parameters are fixed and obtained from Table 4 and 5.

It can be observed in Table 6 and Figure 12 that the fitted mean μ of the peak signal in all four categories is consistent, and ranges from 5618.72 ± 2.30 MeV/c^2 to 5621.38 ± 3.18 MeV/c^2, which is in agreement with the theoretical Λ_b^0 mass reported by the PDG, 5619.57 ± 0.16 MeV/c^2 [1]. The reason for separating the magnet up (MU) and magnet down (MD) datasets is to identify and account for potential detector-induced asymmetries, such as magnetic field biases that could affect the reconstruction of charged particles differently. The consistency of the fitted means and number of signal events (see Table 6) across both polarities indicates that no significant detector-related biases are present. This provides confidence that the model is effectively representing the data and that the detector response is symmetric with respect to polarity. As a result, it is justified to combine the MU and MD datasets to increase the statistical power of the analysis. The combined result, shown in Figure 13, allows for a more precise comparison of matter and antimatter signal yields. The corresponding fit parameters are provided in Table 8



(a) Matter

(b) Antimatter

Figure 13: Final combined model fits to the invariant mass distributions for (a) matter and (b) antimatter. The signal is represented by the dotted blue line, the background by the dotted green line, and the model fit by the red line.

Category	N_{sig}	μ [MeV/c^2]	σ [MeV/c^2]	N
Matter	2421 ± 293	5621.87 ± 2.96	12.94 ± 3.07	440.1 ± 53.2
Antimatter	2256 ± 259	5618.86 ± 2.37	16.27 ± 2.60	410.2 ± 47.1

Table 8: Fit results for the normalization factor, N_{sig}, peak position μ, and width σ for matter and antimatter. The number of signal events, N, was obtained by dividing N_{sig} by the histogram bin width, to account for binning dependence.

4.4 Raw asymmetry between matter and antimatter

The central objective of this analysis is to investigate potential differences in the production, decay, or detection of Λ_b^0 and $\bar{\Lambda}_b^0$ baryons, which may indicate CP violation or detector-induced asymmetries. To quantify this, the *raw asymmetry* is calculated by comparing the total number of signal events for matter and antimatter candidates, after combining both Magnet Up (MU) and Magnet Down (MD) datasets. The raw asymmetry, denoted as A_{raw}, is defined by the following expression:

$$A_{\text{raw}} = \frac{N(\Lambda_b^0) - N(\bar{\Lambda}_b^0)}{N(\Lambda_b^0) + N(\bar{\Lambda}_b^0)} \tag{5}$$

where $N(\Lambda_b^0)$ and $N(\bar{\Lambda}_b^0)$ correspond to the total number of signal events for matter and antimatter, respectively. This definition provides a direct, model-independent measurement of the relative difference in signal events. Since the analysis incorporates both magnetic polarities, many detector-related asymmetries are expected to cancel out, allowing A_{raw} to serve as a meaningful indicator of any underlying matter–antimatter imbalance in the high-q^2 region.

In contrast to when the magnetic polarities are treated separately, when combining both MU and MD datasets, normalization by the luminosity weights is not required. This is because the respective luminosity factors cancel out in the ratio used to compute the raw asymmetry A_{raw}. Hence, luminosity weighting is not needed for the final results.

After determining the number of signal events for both matter (Λ_b^0) and antimatter ($\bar{\Lambda}_b^0$) decays, using the same method as described in Table 6, the raw asymmetry can be evaluated using Equation 5. By combining the extracted signal yields across polarities, the asymmetry and its statistical uncertainty are computed. The results are summarized in Table 9, which presents both the signal yields and the corresponding raw asymmetry.

Raw Asymmetry	A_{raw} [%]
Matter vs Antimatter	3.51 ± 8.33

Table 9: Total signal yields (N_{sig}) for Λ_b^0 and $\bar{\Lambda}_b^0$, along with the calculated raw asymmetry (A_{raw}) and its statistical uncertainty.

Therefore, the final result illustrating the raw asymmetry between Λ_b^0 and $\bar{\Lambda}_b^0$, as displayed in Table 9, is

$$A_{\text{raw}} = (3.51 \pm 8.33)\%.$$

5 Discussion

This section evaluates the results obtained during the analysis and considers their statistical significance in the context of Standard Model expectations. It also reviews the selection criteria applied throughout the study, identifying areas where improvements or additional selections could enhance signal purity and reduce background contamination. Finally, the methods used to model the data are discussed, with suggestions for how the signal and background description could be refined in future analyses.

5.1 Evaluation of Results

Returning to the introduction, the research question for this thesis was: *Can raw asymmetries between matter and antimatter in high-q^2 regions reveal deviations from Standard Model predictions?*

This was investigated through a detailed analysis of the invariant mass distributions of Λ_b^0 and $\bar{\Lambda}_b^0$ candidates, focusing on the high-q^2 region. By fitting a combined signal and background model to data collected under different polarities (magnet up and magnet down), the number of signal events for both matter and antimatter decays was obtained. These yields were then used to compute the raw asymmetry A_{raw}.

$$A_{\mathrm{raw}} = (3.51 \pm 8.33)\%, \quad \text{for } q^2 > 14.3\ \mathrm{GeV}^2/\mathrm{c}^4$$

The result shows a small positive asymmetry, favoring the production or detection of the Λ_b^0 baryon over the $\bar{\Lambda}_b^0$ antibaryon. However, under the Standard Model (SM), CP violation in this decay channel is expected to be extremely small, likely analogous to the suppressed CP asymmetries observed in neutral B-meson systems of magnitude 10^{-4} to 10^{-5} [18], given the shared $b \to s$ transition. Therefore, any observed asymmetry should ideally be consistent with zero within uncertainties. Deviations from this expectation could potentially signal contributions from physics beyond the Standard Model (BSM), such as new sources of CP violation, or inefficiencies in the detector.

In high-energy physics, a commonly accepted threshold for claiming a discovery for new physics is a significance level of at least 5σ. This corresponds to a very low probability that the observed effect arises from statistical fluctuations. In this analysis, the observed asymmetry of 3.51% with an uncertainty of 8.33% corresponding to a significance of only 0.42σ. This is well below the 5σ threshold. Thus, the result is not statistically significant and cannot be interpreted as evidence of CP violation or beyond the standard model effects. This is additionally in accordance with other studies [29, 31, 32].

5.2 Assessment and improvements to the analysis

A recurring challenge throughout this research was the limited number of events in the high-q^2 region. This posed difficulties for reliable modeling and, by extension, for extracting a precise result. To mitigate this, different datasets (long and downstream track types as well as both magnet polarities) were combined to improve statistical power. A clear improvement for future studies would be to increase the integrated luminosity. This would result in a higher yield of signal events, reducing statistical uncertainties and enabling a more accurate measurements.

5.2.1 Assessment and improvements to the selection criteria

The selection criteria applied in this analysis aimed to reduce combinatorial background effects and retain the signal. Some of these cuts are fundamental to the analysis and require no further improvement. Most notably, the q^2 cut, defined as $q^2 > 14.3\,\mathrm{GeV}^2/c^4$, was essential for focusing the analysis on the high-q^2 region, which is the central scope of this thesis. Similarly, the pseudorapidity (η) cut, restricted to the range $(2,5)$, aligns with the LHCb detector's acceptance and ensures that only well-reconstructed events are included. As such, neither of these cuts are considered for revision.

The transverse momentum (p_T) cut set to $p_T > 2\,\mathrm{GeV}/c$, serves to reduce combinatorial background, as low-p_T tracks are more likely to arise from random combinations rather than actual decay products. While this threshold is reasonable and commonly used, it is somewhat rough and could potentially be optimized further using more targeted background studies. For the mass reconstruction, the Decay Tree Fitter (DTF) mass variable was employed rather than the direct invariant mass. This is a well-justified choice, as the DTF method improves mass resolution by applying vertex constraints and fixing intermediate particle masses, such as the Λ^0, to their known values. This leads to a cleaner peak and better separation between the background and signal.

The charge-based identification of matter versus antimatter candidates using the signs of the proton and pion tracks from the Λ^0 decay is a standard and robust method, leaving little room for improvement. The K^0_S mass veto, implemented to suppress peaking background from $B^0 \to K^0_S\mu^+\mu^-$ decays, was also effective. However, it had minimal impact on the overall dataset due to the limited phase space overlap with the signal region.

To further reduce combinatorial background the implementation of machine learning techniques such as Boosted Decision Trees (BDTs) could be considered. In a related analysis, a BDT is developed with the aim to discriminate against background from $B_s \to \mu^+\mu^-$ decays. Although this BDT is not specifically optimized for the $\Lambda^0_b \to \Lambda^0\mu^+\mu^-$ channel, preliminary tests show its effectiveness in reducing background. Nevertheless, for optimal performance, a new BDT could be trained specifically for this decay mode to differentiate between signal and background events more effectively. Such a machine learning model would take advantage of the decay specific kinematics and topology to suppress the background further without significantly reducing the signal yield.

Another improvement would be applying a cut on the mean lifetime of the Λ^0_b and Λ^0 baryons. Since both particles are long-lived, requiring a minimum decay time or flight distance could help suppress background from short-lived or swiftly produced combinatorial sources. These selection cuts would likely yield a cleaner dataset and improve the precision of the asymmetry measurement.

An additional source of background that is not fully removed by the selection criteria is the leakage from resonant decays, such as $\Lambda^0_b \to \Lambda^0\psi(2S)(\to \mu^+\mu^-)$. While the $q^2 > 14.3\,\mathrm{GeV}^2/c^4$ selection is designed to exclude the $\psi(2S)$ resonance, the separation is not perfect. Consequently, the tails from the resonance may extend into the high-q^2 region, resulting in contaminating the non-resonant region. Since the kinematic distributions of the resonant and non-resonant decays are highly similar, it is not possible to distinguish these events using data-driven methods. However, Monte Carlo simulations can be used to estimate the fraction of $\psi(2S)$ decays that leak into the high-q^2 region. By scaling this fraction with the observed number of resonant decays in data, an estimate of the leakage can be obtained and subtracted in future analyses to improve signal purity.

5.2.2 Assessment and improvements to the data modeling

The combined model used in this analysis of a quadratic function for the background and a double sided crystal ball (DSCB) for the signal, proved to be effective for describing the invariant mass distributions across all data categories used. The quality of the fits was assessed using the reduced chi-squared (χ^2/ndf) statistics, which serves as a measure for the goodness of a fit. In particle physics, it is considered the standard way of assessing the goodness of the fit. In all cases, the obtained values of χ^2/ndf were reasonably close to 1, indicating that the chosen functional forms provided a satisfactory description of the data without overfitting.

Nevertheless, a potential improvement would be to make the signal shape parameters (α_L, n_L, α_R, and n_R) free during the fit, rather than fixing them from Monte Carlo simulations. While fixing these parameters ensures stability in lower statistics samples, it could introduce some form of model dependency that could affect the precision of the obtained signal yields. A more dynamic implementation where the code adapts to the data by fitting these parameters within reasonable bounds, could increase flexibility and reduce reliance on external inputs.

Additionally, a further consideration is the choice of a binned fit. While binned fitting is computationally straightforward to implement, it introduces a dependency on the number of bins. This can lead to a loss of information, especially when the number of bins is small or if the bin size is not well optimized. An unbinned fit, by contrast, uses the individual data points and does not from information loss, resulting in a more accurate fit. Implementing unbinned fits could enhance the precision of the signal, particularly in datasets with high statistics or sharper mass peaks.

Lastly, this analysis was performed in Python using the `uproot` library to read the provided ROOT files. More commonly, ROOT [33] is used directly to analyze LHCb data. ROOT provides tools such as `RooFit` [34] and `RooStats` [35], which allow for more powerful statistical modeling, such as more advanced fitting methods, systematic uncertainties, and likelihood profiling. Furthermore, the `RooFit` library specifically, facilitates unbinned fitting. These toolkits could increase the accuracy of the fitting and provide for systematic uncertainties, resulting in more accurate results.

6 Conclusion

To conclude this thesis, the main objective was to analyze the $\Lambda_b^0 \to \Lambda^0\mu^+\mu^-$ and $\bar{\Lambda}_b^0 \to \bar{\Lambda}^0\mu^+\mu^-$ decays in the high-q^2 region ($q > 14.3\ \mathrm{GeV/c^2}$), using data from the LHCb experiment at CERN, in order to test for CP violation (CPV) through matter-antimatter asymmetry. The presence of CPV beyond what is described by the CKM matrix would indicate physics beyond the Standard Model (BSM) or inefficiencies in the detector.

The analysis was performed on proton-proton collision data collected during Run 2 (2016–2018) at a center-of-mass energy of 13 TeV, corresponding to an integrated luminosity of approximately 6 fb^{-1}. Although the results showed a raw asymmetry of 3.51% in favor of the Λ_b^0 baryon production over the $\bar{\Lambda}_b^0$ antibaryon, the statistical uncertainty of 8.33% causes the result to be statistically compatible with no CPV. With a significance of 0.421σ, lying far below the 5σ threshold required to claim evidence for new physics. This result is consistent with Standard Model predictions and previous studies.

To improve the analysis, careful selection criteria were applied, and further enhancements were suggested, such as the development of an optimized boosted decision tree (BDT), and the inclusion of lifetime-based cuts for the Λ_b^0 and Λ^0 baryons. Finally, while the analysis was carried out using the `uproot` library in Python for accessibility, future studies could benefit from using ROOT directly, which is better optimized for high-energy physics workflows and may allow for more precise control over data processing and fitting procedures.

Bibliography

[1] Particle Data Group, S. Navas *et al.*, *Review of particle physics*, Phys. Rev. D **110** (2024) 030001.

[2] ALEPH, DELPHI, L3, OPAL, LEP Electroweak Working Group, *A Combination of preliminary electroweak measurements and constraints on the standard model*, `arXiv:hep-ex/0511027`.

[3] UA1 collaboration, G. Arnison *et al.*, *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at* $\sqrt{s} = 540$ *GeV*, Phys. Lett. B **122** (1983) 103.

[4] L. Di Lella and C. Rubbia, *The Discovery of the W and Z Particles*, Adv. Ser. Direct. High Energy Phys. **23** (2015) 137.

[5] ATLAS collaboration, G. Aad *et al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, `arXiv:1207.7214`.

[6] CMS collaboration, S. Chatrchyan *et al.*, *Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC*, Phys. Lett. B **716** (2012) 30, `arXiv:1207.7235`.

[7] J. Ellis, *Physics Beyond the Standard Model*, Nucl. Phys. A **827** (2009) 187C, `arXiv:0902.0357`.

[8] A. D. Sakharov, *Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe*, Pisma Zh. Eksp. Teor. Fiz. **5** (1967) 32.

[9] 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** (1964) 138.

[10] LHCb collaboration, R. Aaij *et al.*, *Measurement of CP asymmetries in* $\Lambda_b^0 \to ph^-$ *decays*, Phys. Rev. D **111** (2025) 092004, `arXiv:2412.13958`.

[11] X.-G. He, C.-W. Liu, and J. Tandean, *Large CP violation in* $\Lambda_b^0 \to pK^-\pi^+\pi^-$ *and its U-spin partner decays*, `arXiv:2503.24350`.

[12] F.-S. Yu and C.-D. Lü, *New horizon in particle physics: first observation of CP violation in baryon decays*, Sci. Bull. **70** (2025) 2035, `arXiv:2504.15008`.

[13] LHCb collaboration, S. Dekkers *et al.*, *LHCb Upgrade II Scoping Document* , CERN, Geneva, 2024. doi: 10.17181/CERN.2RXP.HDK0.

[14] F. Archilli *et al.*, *Performance of the Muon Identification at LHCb*, JINST **8** (2013) P10020, `arXiv:1306.0249`.

[15] S. R. Shastry, *The standard model of particle physics*, Shaastra Magazine (2022), `https://shaastramag.iitm.ac.in/first-principles/standard-model-particle-physics`.

[16] A. Pich, *The Standard Model of Electroweak Interactions*, in *2010 European School of High Energy Physics*, 1–50, 2012, `arXiv:1201.0537`.

[17] G. Buchalla, A. J. Buras, and M. E. Lautenbacher, *Weak decays beyond leading logarithms*, Rev. Mod. Phys. **68** (1996) 1125, `arXiv:hep-ph/9512380`.

[18] J. Albrecht, F. Bernlochner, A. Lenz, and A. Rusov, *Lifetimes of b-hadrons and mixing of neutral B-mesons: theoretical and experimental status*, Eur. Phys. J. ST **233** (2024) 359, `arXiv:2402.04224`.

[19] M. Misiak, *Rare B-Meson Decays*, in *15th Lomonosov Conference on Elementary Particle Physics*, 301–305, 2013, `arXiv:1112.5978`.

[20] LHCb collaboration, R. Aaij *et al.*, *Measurement of lepton universality parameters in* $B^+ \to K^+\ell^+\ell^-$ *and* $B^0 \to K^{*0}\ell^+\ell^-$ *decays*, Phys. Rev. D **108** (2023) 032002, `arXiv:2212.09153`.

[21] S.-Q. Zhang and C.-F. Qiao, *Rare* Λc *decays and new physics effects*, Phys. Rev. D **110** (2024) 114040, `arXiv:2411.15857`.

[22] L. Greeven, *Decoding beauty: rare baryonic decays SciFi detector commissioning*, PhD thesis, Maastricht University, Netherlands, 2024, doi: 10.26481/dis.20240403lg.

[23] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[24] J. Closier, *End of the first phase and beginning of a new lhcb*, 2018.

[25] LHCb collaboration, A. A. Alves, Jr. *et al.*, *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[26] LHCb collaboration, R. Aaij *et al.*, *Measurement of the track reconstruction efficiency at LHCb*, JINST **10** (2015) P02007, `arXiv:1408.1251`.

[27] J. Pivarski, P. Elmer, and D. Lange, *Awkward Arrays in Python, C++, and Numba*, EPJ Web Conf. **245** (2020) 05023, `arXiv:2001.06307`.

[28] W. D. Hulsbergen, *Decay chain fitting with a Kalman filter*, Nucl. Instrum. Meth. A **552** (2005) 566, `arXiv:physics/0503191`.

[29] LHCb collaboration, R. Aaij *et al.*, *Observation of a* $\Lambda_b^0 - \overline{\Lambda}_b^0$ *production asymmetry in proton-proton collisions at* $\sqrt{s} = 7$ *and* $8\,TeV$, JHEP **10** (2021) 060, `arXiv:2107.09593`.

[30] S. Bouma, *Quantification of backgrounds in the high-*q^2 *region of R(*Λ*)*, Master's thesis, University of Groningen, 2024.

[31] W. K. Lai and A. K. Leibovich, $\Lambda_c{}^+/\Lambda_c^-$ *and* $\Lambda_b^0/\bar{\Lambda}_b^0$ *production asymmetry at the LHC from heavy quark recombination*, Phys. Rev. D **91** (2015) 054022, `arXiv:1410.2091`.

[32] F. Ferrari, *Measurement of* B^0, B_s^0, B^+ *and* Λ_b^0 *production asymmetries in 7 and 8 TeV pp collisions at LHCb*, Nuovo Cim. C **41** (2018) 41.

[33] R. Brun and F. Rademakers, *ROOT - an object oriented data analysis framework*, Nucl. Inst. Meth. in Phys. Res. A **389** (1997) 81, Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996.

[34] W. Verkerke and D. P. Kirkby, *The RooFit toolkit for data modeling*, eConf **C0303241** (2003) MOLT007, `arXiv:physics/0306116`.

[35] L. Moneta *et al.*, *The RooStats Project*, PoS **ACAT2010** (2010) 057, `arXiv:1009.1003`.