DOCTORAL THESIS

Search for CP Violation in Baryon Decays and Radiation Damage of the LHCb VELO

Author:
Gediminas Sarpis

Supervisor:
Prof. Chris Parkes

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy

in the

Department of Physics and Astronomy in the School of Natural Sciences
Faculty of Science and Engineering

2020

Blank page

Contents

Word count: 31883

Blank page

List of Figures

List of Tables

Blank page

Blank page

THE UNIVERSITY OF MANCHESTER

Abstract

Faculty of Science and Engineering
Department of Physics and Astronomy in the School of Natural Sciences

Doctor of Philosophy

Search for CP Violation in Baryon Decays and Radiation Damage of the LHCb VELO

by Gediminas Sarpis

A search for *CP* violation in baryonic four-body $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays is performed using the data collected by the LHCb detector during Run 1 (2011-2012) and Run 2 (2016-2017) of the LHC. The collected data corresponds to $7\,\text{fb}^{-1}$ of integrated luminosity and contains approximately 36 thousand signal candidates. The measurement is done using two independent methods - Triple Product Asymmetries and the energy test. The Triple Product Asymmetries method is sensitive to the *P*-odd type of *CP* violation and is based on the binned and integrated measurements of *P*-odd asymmetry variables. The energy test is a model independent technique, that can search for *P*-even and *P*-odd *CP* violation. The technique is based on a comparison of the distribution of the events in the decay phase space. Neither of the methods find a deviation from CP conservation hypothesis with a significance higher than 3σ. Parity violation is measured with both methods with significance higher than the 5σ observation threshold.

Studies of the radiation damage of the LHCb VELO silicon-strip detector are presented. The data taken during special regular VELO performance tests, called Charge Collection Efficiency scans, is studied. Multiple VELO performance aspects are investigated, such as the bad strip detection and exclusion algorithms, Charge Collection Efficiency and Cluster Finding Efficiency. An investigation of unforeseen tracking efficiency reduction, caused by the Second Metal Layer effect, is carried out. The time dependence of the radiation damage effects of the VELO subdetector are presented. A novel procedure of beneficially annealing the VELO is performed and presented. The high voltage tests of the VELO are performed during the last runs of the LHC, testing the detector beyond its design limitations. The listed studies allowed the evolution of the radiation damage effects of the VELO sensors to be seen. In particular, approximately 2% of VELO strips are found to be bad, but the number is stable throughout the years. The required bias voltage of the sensors is seen to rise with increasing irradiation, as predicted by the Hamburg model. The Second Metal Layer effect evolution in time is quantified. The effect starts early in Run 1 with initial irradiation, but stabilizes for the rest of the VELO operation until the end of Run 2.

Blank page

Declaration of Authorship

Candidate name: Gediminas Sarpis
Faculty: Faculty of Science and Engineering
Thesis title: Search for CP Violation in Baryon Decays and Radiation Damage of the LHCb VELO

This work represents the combined efforts of the author and his colleagues in the LHCb collaboration. Some of the content has been published elsewhere and/or presented to several audiences. No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Signed:

Date:

Blank page

Copyright Statement

Blank page

Seneliui Aleksui

Acknowledgements

This, too, shall pass

Throughout my PhD, I have remembered the words above on many occasions. It has been my mantra in the times of frustration and uncertainty, yet also in the moments of great joy. I repeat them now and want to begin by thanking the World for not ending...

The words "large international collaboration" do little justice to the personal subjective experience of being a part of one. True, the LHCb collaboration is large. It is evident by the number of brilliant minds I have met. The "international" part is quite self-explanatory, during my PhD I was lucky enough to live in the UK, France and Lithuania. While working at CERN I would cross international borders 2-4 times a day on average. This is truly exciting and unique work. Yet, I think the most important characteristic of the LHCb is encoded in the word "collaboration".

Many times I was in awe of the selfless dedication towards a common goal my LHCb colleagues showed. I was also lucky enough to be a part of those precious moments of discovery, when the excitement was in the air and verbal reports were replaced by happy faces, filled with mutual understanding that we have just been a part of something bigger than any individual.

I must start by thanking my PhD supervisor Chris, who is not only the best supervisor one could hope for, but also exhibits almost a Zen-master quality of calming a restless mind when needed, and waking up a lazy one, when the time comes.

As often happens in life, good things come in pairs. Hence, the second person I thank is Chris B. It is hard to quantify how much of this thesis would not exist without your help and instruction (especially on all things Python). It is much easier for me to estimate with high degree of certainty that I made a friend. Let Mont Blanc be but a start of our adventures.

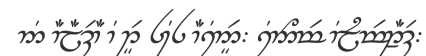
Next, I want to thank all great LHCb colleagues from Manchester. I thank Will, Adam and Lucia for showing me what to aspire in order to become a great postdoc. I also thank Gulio, George, Marco and Eva, for your help at the beginning and throughout my PhD. I thank Dónal, for all of the adventure tales that I am so jealous of. Apart from helping with the physics side of things, all of you have made me feel very welcome!

My LHCb colleagues aside, I thank Ste, who is a statistics guru par excellence and a great listener too, as well as Nico, who always supported me during lunch at R1. From my times at CERN, I thank Baptiste for amazing dinners.

A special thanks goes to Paula, Tomasz and Shanzhen, who have introduced me to a weird and exciting world of particle detectors and helped me make a contribution to the VELO group. The hardware side of things was at times scary, but always exciting!

I want to thank my family. To my parents Viktoras and Ramona, I thank you for your support throughout my life and especially my PhD. Who knows who I would be if not for "The Universe in a Nutshell" and "Science and the Universe" in your bookshelves. To Mindaugas and Viktorija, I am happy to have such nice siblings to play and travel with (or occasionally do physics).

There are no words in any language to describe the relationship I have with Darius and Ineta. I could call you my best friends, but that presupposes hierarchy and social conventions. You are the two forming "principles" of all that can be considered "me". Ineta (*MDLR*), you remain my standard of Joy, as well as number one person in my life by the amount of letters exchanged. Dariau,



Sasha, it feels like we are getting closer to demystifying conciousness. Yet, whenever we get a glimpse of it, it disappears. I must admit, this exercise, which has been going for years, is one of the most entertaining activities I've ever done. Maybe that is the answer, thank you for sharing the weird and wonderful experience of being present with me.

During the first year of my PhD in Manchester I was a part of an extremely rare phenomena. Seven seemingly different people formed an excited strongly bound resonance. I remain perplexed by the supernatural miracle of **Kiaunės**. Rūta, yet again you leave me speechless. The only thing I know for certain is that you are one of the very few aerospace engineers who know the answer to the riddle "Houston, we have a problem. Why are there two Suns in the sky?". Lukai, after a decade of procrastination, you were the muse who inspired me to pick up my guitar. The New Years eve in France will always be... A lovely day! Ugne, many of the things I want to say must be omitted to maintain plausible deniability, but I can safely say, that you constantly show me and others what a true friendship looks like. Deimante and Tomai, so many years have passed and yet we still find each other living in the same flat. I take it as a sign of significant compatibility of minds, that leaves the residual cosiness when we are together.

Finally, I thank my only true Long Term Attachment - my beloved wife, Judy. You fit in all of the cat-egories. You were by my side at CERN, read every word of my thesis, climbed mountains with me and belonged to all of the circles of friends listed here.

Judita, while I was trying to figure out why matter is, you constantly reminded me why it does.
Miau.

Blank page

Preface

The field of High Energy Physics offers a plethora of topics on which exciting measurements and analyses can be performed. Previous measurements have allowed us to gain deep insights into the fundamental structure and laws of the Universe. Discovery of four fundamental forces, hundreds of particles composed of different quark combinations, the Higgs boson and CP violation in the quark sector are just a few examples. It is clear that in order to continue exploring the world of Particle Physics further, huge colliders, such as the LHC, need to be built. These machines are capable of unprecedented collision energy and allowed us to explore the "Energy Frontier" of Particle Physics, creating new heavy fundamental and composite particle states. On the other hand, new discoveries can be made by increasing the precision of the measurements of known particle parameters, such as coupling strengths and masses. In order to achieve this, the initial state of particle collisions needs to be measured and controlled with extremely high accuracy. This is referred to as the "Intensity frontier". Naturally, both of these research directions present new economic, engineering and computing challenges. These challenges can only be overcome by big international collaborations and continuous accumulation of knowledge and expertise over numerous analyses, software improvements and technological refinements.

This thesis presents a physics analysis using data from the LHCb experiment and work done on the radiation damage of the LHCb VELO detector. The thesis is structured as follows:
The First Chapter gives a brief theoretical overview. The Standard Model of Particle Physics is presented with a specific focus on the *CP* violation. The chapter concludes with a review of previous *CP* violation searches in baryonic decays.
The Second Chapter introduces the LHCb detector at the LHC, CERN. Emphasis is given on the VELO subdetector. The chapter concludes with a brief note on the ongoing LHCb upgrade.
The Third Chapter presents the studies of the VELO radiation damage. A brief overview of general principles of silicon-based detectors and their radiation damage is given. This is followed by the measurements of the VELO performance and its evolution with increasing exposure to radiation.
The Fourth Chapter presents and compares different methods for CP violation searches.
The Fifth Chapter presents the main physics analysis of this thesis. A search for *CP* violation in $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays is carried out and results from two different methods used are given.
The Sixth Chapter concludes this thesis. This chapter recaps the main results of the thesis and gives an outlook on the future work in the areas covered by this thesis.

The work presented in this thesis has been done in collaboration with colleagues from the LHCb. In particular, the physics analysis presented in Chap. 5 has been carried out by a number of proponents from the University of Manchester and Universitá degli

Studi di Milano and INFN Milano, Italy. A natural split of effort occurred and work on the Triple Product Asymmetries measurement has been done by the proponents from Milan, while the work on the energy test method has been done by the proponents from Manchester, including the author of this thesis. The Milan group has also worked on the previous analysis of this channel, published in Ref. [53]. The analysis presented in Chap. 5 supersedes the previous analysis but uses the same selection for Run 1 and a similar selection for Run 2, which was also performed by the Milan group. The details on the data selection and Triple Product Asymmetry measurement can be found in Ref. [125]. However, for completeness and comparison, a short description of the techniques and results of the Triple Product Asymmetries analysis are also presented here.

The radiation damage studies and the monitoring of the LHCb VELO detector is also a team effort. Many people from the VELO group contributed to the ideas, software and measurements of the VELO performance and its dependence on the radiation damage. In particular, the author of this thesis belongs to a lineage of researchers who studied the CCE Scan data throughout the years. The work on the radiation damage is documented in the theses of A. Webber [148], J. Harrison [109] and S. Chen [22]. The analysis of the bias voltage of VELO sensors was also performed by A. Davis [1] and W. Barter [13], who was also responsible for the beneficial annealing of the VELO. The main contribution of the author focused on the bad strip removal algorithms and the CFE studies, allowing quantification of the Second Metal Layer Effect. The studies of the changes to sensor depletion voltages and Hamburg model prediction are briefly presented for completeness.

The author of this thesis fist joined the LHCb collaboration in 2015 as a University of Glasgow MSci student. The author worked on two projects during the two final years of the MSci. The first project was a *CP* violation measurement in $B^0_d \to \pi^+\pi^-$ decays, focusing on the analytical expression of the normalization for the probability density function for these decays. The second project, undertaken in the final year of the MSci, was the first unbiased life-time measurement of Λ_c^+ and Ξ_c^+ decaying into three-body final state, $pK\pi$. The author has also been involved in numerous teaching, outreach and interdisciplinary projects opportunities. These works are not described here.

Chapter 1

Theory

Particle Physics is one of the most active and exciting fields in fundamental science. Huge experimental apparatus, such as the LHC, allows the collection of high energy particle collision data with ever increasing yields. This allows us to perform precision measurements and searches for new fundamental and composite particles. The theoretical models that describe these particle states and their interactions are becoming more and more elaborate. In some cases, the theoretical predictions agree with the experimental results with unprecedented accuracy. An astounding example of this is the measurement of the electron magnetic moment $(g-2)$, which agrees with the theoretical prediction within the relative uncertainty of 10^{-10} [108], making it the single most precise measurement of any field of science and a testament to human ingenuity.

The state of the art paradigm of theoretical particle physics is quantum field theory (QFT). The most popular and successful QFT description of the fundamental particle interactions is the Standard Model of Particle Physics (SM). The SM defines a list of fundamental particles and three fundamental forces responsible for the interaction of such particles. The fourth known fundamental force is gravity which has proven very difficult to be described as a QFT. However, the masses of the fundamental particles are extremely small, hence the force of gravity between them can safely be neglected.

The SM is an extremely accurate and predictive tool. However, it is clear that the SM is not the ultimate description of nature as it fails to explain multiple observed phenomena, such as gravity, dark matter and neutrino oscillations that imply non-zero masses of neutrinos. For this reason, the search for Beyond Standard Model (BSM) phenomena defines the physics programme of the new decade.

This chapter gives a brief introduction to the Standard Model, followed by a description of the role of symmetries in particle physics and the CPT theorem [1]. Afterwards, *CP* violation is discussed from both theoretical and experimental perspectives. The search for *CP* violation in $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays is the main physics analysis of this thesis, which is described in Chap. 5 [2].

[1] Here CPT is a combined symmetry of Charge conjugation, Parity tranformation and Time reversal, discussed later in the text

[2] Throughout this thesis the use of natural units ($c = \hbar = 1$) and the inclusion of charge conjugate processes is assumed unless otherwise stated.

1.1 Standard Model

The Standard Model is a combination of several QFTs, namely Quantum Electrodynamics (QED) and Quantum Chromodynamics (QCD), that describe the Electroweak and Strong interactions. The incorporation of the Higgs Mechanism completed the Standard Model by explaining how charged leptons and massive bosons obtain their mass. The work leading to the current formulation of the SM was finalized in mid-1970s in large part by Sheldon Glashow [103], Steven Weinberg [149] and Abdus Salam [139], earning them the Nobel Prize in 1979.

The main goal of a particle physics theory is to explain the building blocks of matter and their interactions. In the Standard Model, this is achieved by grouping particles into two categories - fermions and bosons. The fermions are matter particles with spin $\frac{1}{2}$ that follow Fermi-Dirac statistics (e.g. obey the Pauli exclusion principle). Bosons, on the other hand, have integer spins, follow Bose-Einstein statistics and are the mediator particles responsible for the fundamental interactions of fermions.

Fermions are further split into quarks and leptons. Leptons contain the three charged massive leptons (e^-, μ^-, τ^-), as well as three corresponding neutrinos $(\nu_e, \nu_\mu, \nu_\tau)$. The charged leptons interact electromagnetically and via the weak interaction, while neutrinos carry no electric charge and can only interact via weak decays, making them very difficult to detect. The discovery of neutrino oscillations proved that neutrinos have very small but non-zero mass, the origin of which is yet to be understood [86].

The quark sector is composed of 6 different quarks; three of these (u, c, t) have electric charge of $+\frac{2}{3}$, while the other three (d, s, b) have $-\frac{1}{3}$. Quarks of the same electric charge differ only by their mass. For this reason, the quarks with $+\frac{2}{3}$ $(-\frac{1}{3})$ charge are sometimes called up-type (down-type) quarks. Apart from interacting electromagnetically and via the weak interaction, quarks also carry an additional charge, called the colour charge, hence they can interact via the strong interaction.

All 12 fermions can be split into three families (or generations). Each family contains a charged lepton, a corresponding neutrino and a pair of up/down-type quarks. The families only differ by their mass. The fermions of the second and third generation are more massive and decay into less massive counterparts. For this reason, all stable baryonic matter in the universe is only composed of the first family of quarks and fermions. The Standard Model is agnostic on the origin of fermion family hierarchy or the number of fermion generations. This is one of the biggest unanswered questions in physics.

The other components of the Standard model are the bosons. Most bosons have spin 1 and are thus called gauge bosons, while the Higgs boson has spin 0 and is called a scalar boson. Bosons are force carrier particles responsible for the interactions of fermions. The photon (γ) is a massless boson responsible for the electromagnetic interaction. Z^0 and $W^\pm$ bosons are massive particles that carry the weak nuclear force. Massless gluons mediate the strong interaction. They do not have an electric charge but carry double-colour charge of QCD, which allows for eight different charge combinations. Finally, the Higgs boson is a massive chargeless boson, which emerges as a consequence of the Higgs mechanism

responsible for particles acquiring their masses. The list of Standard Model particles and their interaction graph can be seen in Figs. 1.1.

Mathematically the Standard Model is represented as a product of three unitary groups, representing the local gauge symmetries according to which the particle fields transform,

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{1.1}$$

The $SU(3)_C$ term corresponds to the quantum chromodynamics (QCD) that describe the strong interaction. The $SU(2)_L \otimes U(1)_Y$ defines the electroweak force. By the Noether's theorem [129], each symmetry has a corresponding conserved charge. In the SM, the charges arising from the invariance under gauge transformations also have a corresponding gauge boson field, which is in the adjoint representation of the corresponding group.

1.1.1 Quantum Chromodynamics

QCD introduces a charge with three distinct values according to the $SU(3)_C$ symmetry. By analogy to the RGB colour scheme, this charge is called the colour charge and can have one of three colour values: red, green or blue. The antiparticles have anti-colour charge, which is called anti-red, anti-green and anti-blue.

The generators of the Lie algebra of the $SU(3)$ group are the eight Gell-Mann matrices. These matrices are traceless, hermitian and linearly independent. In nature, each Gell-Mann matrix corresponds to a bosonic quantum field which couples to the colour charge. The gauge bosons mediating the strong interaction are called gluons. Gluons are doubly-colour charged and come in eight varieties, each having one unit of colour and one unit of anti-colour.

Gluons interact with quarks and other gluons to form strongly bound states. Such states are colour neutral and can consist of multiple bound quarks [101]. There are many known two- and three-quark states. The two-quark states are called mesons and are made up of quark-antiquark pairs with opposite colour charge (examples include pions, kaons, B and D mesons). The three-quark states are made up of three quarks or three anti-quarks in a colour neutral arrangement (for example proton, neutron and Λ_b^0). More exotic states made up of four [24, 43] or five [59, 56] quarks have been discovered. More recently, with LHCb having made major contributions to these exotic hadrons it is also possible that there might exist even more exotic states of more quarks, or states consisting of only gluons, called glueballs.

Another peculiar property of QCD is the quark confinement. It is impossible to separate bound quark states and thus observe an isolated quark. This happens because the interaction energy in the gluon field grows with increasing distance between quarks until it is more energetically favourable for quark-antiquark pairs to appear. This process is called hadronisation.

QCD is a very successful field theory, describing the interactions of quarks and gluons. The use of QCD to explain the high multiplicity of composite quark states was one of the

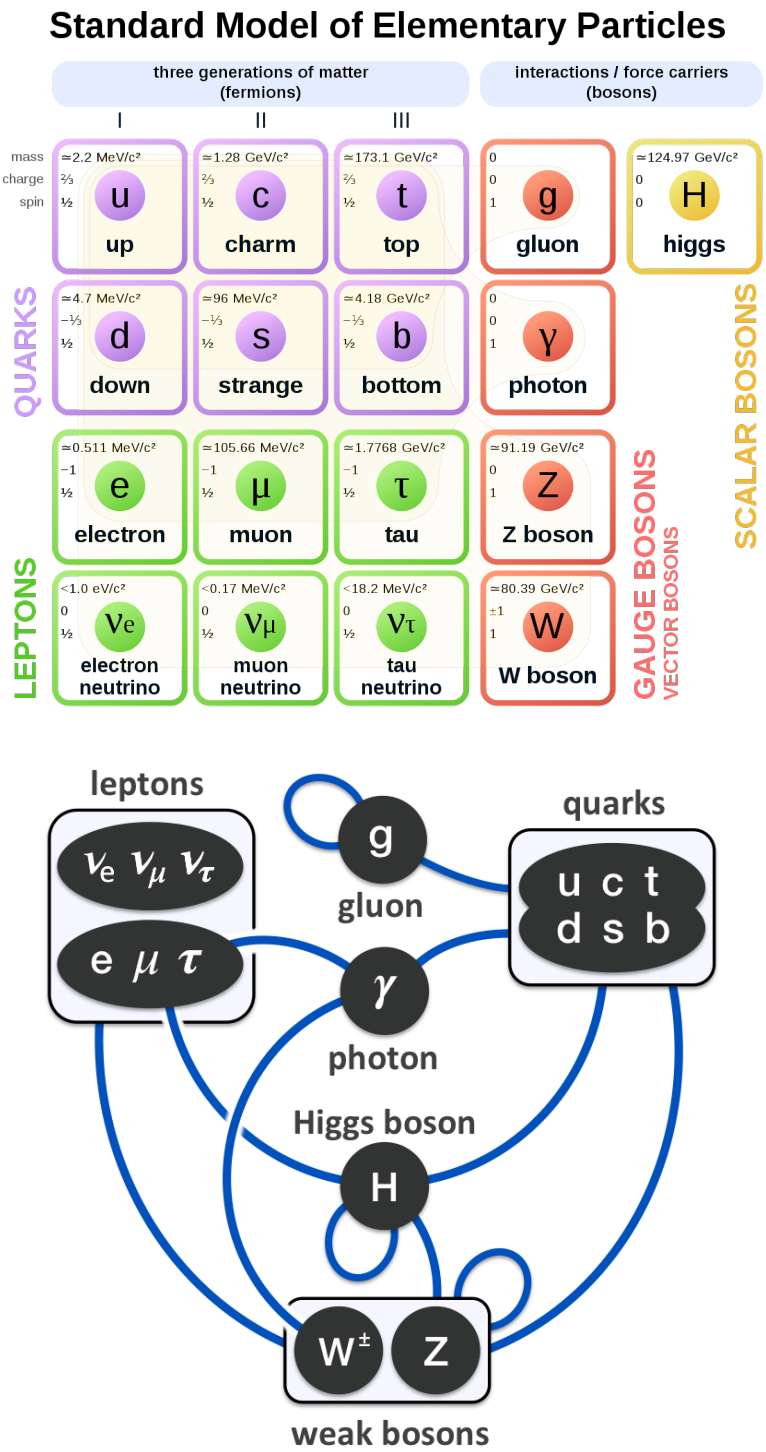


Figure 1.1: (Top) A structured list of Standard Model particles. (Bottom) A graph showing standard model couplings and self-couplings. Reproduced from [90, 95].

biggest triumphs in particle physics. However, there are still many unanswered questions in QCD and much work to be done in order to obtain an even more accurate theory of the strong interaction.

1.1.2 Electroweak Interaction

The other term of the SM - $SU(2)_L \otimes U(1)_Y$ corresponds to the electroweak interaction. The $SU(2)_L$ symmetry introduces three gauge bosons W_1, W_2, W_3 that couple to weak isospin - T, while the $U(1)_Y$ adds another gauge boson - B, which couples to the hypercharge - Y. However, none of these bosons actually exist in nature. Instead, a process called spontaneous symmetry breaking occurs at low energy scales. This proceeds via the Higgs mechanism, when a complex scalar $SU(2)_L$ doublet, the Higgs field, breaks the SM symmetry as follows:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \rightarrow SU(3)_C \otimes U(1)_Q \tag{1.2}$$

This separates the electroweak interaction into electromagnetism and the weak force. The observable EM and weak gauge bosons are constructed as linear combinations of the $W_{1,2,3}$ and B bosons:

$$W^{\pm} = \frac{1}{\sqrt{2}}(W_1 \pm W_2) \tag{1.3}$$

$$\begin{pmatrix} \gamma \\ Z \end{pmatrix} = \begin{pmatrix} cos\theta_W & sin\theta_W \\ -sin\theta_W & cos\theta_W \end{pmatrix} \begin{pmatrix} B \\ W^3 \end{pmatrix} \tag{1.4}$$

where θ_W is the weak mixing angle proportional to the couplings of $SU(2)_L$ and $U(1)_Y$.

The new $W^{\pm}$ and Z^0 bosons couple to three out of four degrees of freedom of the Higgs field and become massive. The photon (γ) does not interact with the Higgs field and hence remains massless. This is due to the fact that a special linear combination of the weak hypercharge (Y) and the third component of the weak isospin (T_3),

$$Q = T_3 + \frac{1}{2}Y_W \tag{1.5}$$

does not couple to the Higgs field. Hence, it leaves an unbroken $U(1)_Q$ symmetry, which corresponds to Electromagnetism and the familiar electric charge Q. Furthermore, the fourth remaining degree of freedom of the Higgs field does not couple to other gauge boson fields and obtains a non-zero vacuum expectation value (vev). This introduces a new scalar boson particle - the Higgs boson. The Higgs boson is massive (125 GeV), carries no electric or colour charge, but has a weak isospin ($-\frac{1}{2}$) and weak hypercharge ($+1$). The charged fermions obtain their masses through the interactions with the Higgs boson.

In 2012, the existence of the Higgs boson was experimentally confirmed by the ATLAS [30] and CMS [38] collaborations. This has completed the Standard Model of particle

physics. However, there are observed phenomena that the Standard Model fails to explain, hence it can not be the complete theory of nature. One of these phenomena is the observed universal matter-antimatter imbalance, presented in the next section.

1.2 CP Violation

It is widely accepted that the Big Bang was a very symmetric phenomenon. This means that approximately 13.8 billion years ago [85] equal amounts of matter and antimatter emerged from the Big Bang. In this early stage of the universe, matter and antimatter particles annihilated. The signature of this can still be observed in the Cosmic Microwave Background Radiation. However, a small fraction ($\mathcal{O}(10^{-9})$) of baryon excess emerged, which now creates all of the matter in the observable universe. The evidence for the absence of antimatter in our universe has also been measured experimentally (e.g. by the AMS experiment [15]).

The phenomenon creating this asymmetry between baryons and antibaryons is called baryogenesis. It has been shown, that baryogenesis can occur if three so called Sakharov conditions are met [138].

1. Baryon number is violated
2. C and CP symmetries are violated
3. Interactions out of thermal equilibrium

These conditions are sufficient to produce baryon asymmetry. However, no baryon number violating decays have been observed yet and the magnitude of *CP* violation within the Standard Model is orders of magnitude too small to explain the observed universal asymmetry [84]. For this reason, studies of *CP* violation are essential in order to understand the nature of this asymmetry and gain insight into alternative *CP* violating processes.

1.2.1 Symmetries in Physics and CPT Theorem

Arguably, the most important concept in modern physics is the one of symmetry. This has been formalised by Emmy Noether in her famous theorem that links the differentiable symmetries of the action of a physical system with corresponding conservation laws [129]. Such symmetries can be continuous, like translational or rotational symmetries, corresponding to conservation of momentum and angular momentum respectively. Other types of symmetries include permutation symmetry, which leads to particles of the same type being indistinguishable, or unitary symmetries, like $SU(3)_C$, $SU(2)_L$ and $U(1)_{EM}$, that have been described in Sec. 1.1, and lead to conservation of colour, isospin and electric charges respectively.

Another very important type of symmetry is that of the discrete symmetries, for example, the three symmetries that form the combined CPT symmetry. Here C stands

for charge conjugation, which reverses the signs of electric charge, magnetic moment and the colour charge of the particle, i.e. swaps it to its antiparticle. The electromagnetic interactions are invariant under C transformation.

The other term of the combined CPT symmetry is the Parity (P). It transforms the right-handed spacial coordinate system into a left-handed one and vice versa, $\mathcal{P} : (x, y, z) \mapsto (-x, -y, -z)$. Under a parity transformation, certain physical variables change their sign, while others do not. In the former case, the variables are said to be P-odd, while in the latter they are said to be P-even. Examples of P-even variables include invariant mass, angular momentum and the magnetic field B, while examples of P-odd variables include helicity, velocity and momentum. A quantum state is described by a wave function which can also be P-even or P-odd, or have terms that are either P-even or P-odd. Constructing P-odd variables and subsequently measuring them might allow calculation of the P-odd quantum mechanical amplitudes of a given quantum state. This approach is used in the physics analysis described in the Chap. 5, where the measurement of P-odd triple-product variable allows a search for CP violation in the P-odd amplitudes of the decay.

The final term T in the CPT stands for time reversal. This transformation reverses the arrow of time and hence the signs of all time dependent quantities (such as velocities and momenta). The combined CPT symmetry is believed to be an exact symmetry of the universe. This is also required by the CPT theorem [142] which states that any Lorentz invariant local QFT must obey it. The violation of CPT symmetry would lead to many testable phenomena (for example the difference between the masses of particles and antiparticles) none of which have been observed [119].

The individual terms of the CPT symmetry can be violated, as long as there is cancellation conserving the CPT combination. Such violations of individual symmetries of CPT have indeed been observed. Parity violation was first observed in the β-decays of ^{60}Co by C.S. Wu [152]. This experiment showed that the weak interaction maximally violates parity. Almost at the same time, another team led by L.M. Lederman have observed the C violation in the muon and antimuon decays [100]. It was believed that CP violation could still be conserved until the proof of its violation was obtained in 1964 by Cronin and Fitch when analysing kaon decays [25].

After the initial discovery of CP violation, many analyses have been performed in different meson and baryon decay channels. The first observation of CP violation in the B^0 meson system was done by the Belle and Babar collaborations in 2001 [36, 33]. Recently, the LHCb collaboration has discovered CP violation in the D^0 meson system, which is the first observation of CP violation in charm quarks and the first in an up-type quark system [57].

All of the mentioned discoveries of CP violating decays are potentially consistent with the predictions of the Standard Model. The situation for the charm quark system is not yet clear with some papers arguing that it cannot be accommodated in the SM [114], while others claiming that it can [99]. In the SM, CP violation can arise via the CKM mechanism (described in the following section). However, the amount of CP violation in the SM and observed in these measurements is orders of magnitude too small to explain the observed

universal asymmetry [84]. For this reason, the interesting and rapidly evolving terrain of *CP* violation studies are paramount in understanding other possible *CP* violating effects beyond the standard model.

1.2.2 CKM Mechanism

CP violation exists in the standard model due to the fact that the weak force does not couple to the mass eigenstates of quarks but rather their linear combinations. The flavour changing quark transitions are described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [117]. This matrix also predicted the third generation of quarks. According to the CKM matrix, the weak decays, mediated by the $W^{\pm}$ bosons, change the up-type quarks into down-type quarks and vice versa. Each quark has a different probability of changing its flavour. These probabilities are the CKM matrix elements and are represented below.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.6}$$

Here the d', s', b' represent the weak eigenstates that mix into mass eigenstates d, s, b on the right. The $|V_{ij}|^2$ gives the probability of quark with flavour i to transition into a quark with flavour j. This implies that the CKM matrix has to be unitary, i.e. $V_{CKM}^{\dagger} V_{CKM} = 1$, where $V_{CKM}^{\dagger}$ is the Hermitian conjugate of the CKM matrix and describes antiquark transitions. Imposing the unitarity constraints, the CKM matrix is left with four degrees of freedom - three rotation angles $\theta_{12}, \theta_{13}, \theta_{23}$ and a complex phase δ which is the source of *CP* violation in the Standard Model. With these four parameters in mind, the CKM matrix can be expressed as a product of three rotation matrices, which obtains the following form,

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{-i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{-i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{-i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{-i\delta} & c_{23}c_{13} \end{pmatrix} \tag{1.7}$$

Here the s_{ij} and c_{ij} are the shorthands for $\sin\theta_{ij}$ and $\cos\theta_{ij}$.

It is known experimentally that the rotation angles of the CKM matrix differ in magnitude significantly ($s_{13} \ll s_{23} \ll s_{12} \ll 1$). This hierarchy allows us to reparametrise the CKM matrix in the so called Wolfenstein parametrisation [151]. In this parametrisation $s_{12} = \lambda, s_{23} = A\lambda^2, s_{13}e^{i\delta} = A\lambda^3(\rho - i\eta)$. All of the new four parameters, A, λ, ρ and η, have a magnitude $\mathcal{O}(1)$ and allows us to separate different orders of the CKM elements. The CKM matrix in the Wolfenstein parametrisation up to order λ^3 is given below,

$$V_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{1.8}$$

It can be seen that the CP-violating complex phases only appear in the terms of $\mathcal{O}(\lambda^3)$. This encodes the magnitude of CP violation in the Standard Model as being small.

The magnitudes of individual CKM matrix elements encode the probability of quark transitions and can be obtained by combining measurements from many different decay channels. The current values of these elements are summarised in the matrix below, taken from Ref. [83],

$$|V_{CKM}| = \begin{pmatrix} 0.97446 \pm 0.00010 & 0.22452 \pm 0.00044 & 0.00365 \pm 0.00012 \\ 0.22438 \pm 0.00044 & 0.97359^{+0.00010}_{-0.00011} & 0.04214 \pm 0.00076 \\ 0.00896^{+0.00024}_{-0.00023} & 0.04133 \pm 0.00074 & 0.999105 \pm 0.000032 \end{pmatrix} \tag{1.9}$$

1.2.3 Unitarity Triangle

A popular graphical representation of the CKM matrix measurements are the so called Unitarity Triangles. The unitarity of the CKM matrix implies a set of 12 constraints that can be expressed in the form like the example below,

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \tag{1.10}$$

This equation describes a triangle in the complex plane, with each of the three terms in the equation representing a side. The sides of this triangle can be normalised by setting the term $V_{cd}V_{cb}^* = 1$. The resulting triangle is shown in Fig. 1.2. A small area near the angle α represents the current experimental uncertainties on the CKM matrix element measurements. If the unitarity condition does not hold, it means that the unitarity triangle is not closed. This would be a clear evidence of physics beyond the Standard Model. Similarly, deviations from the SM could be found by comparing the measurements of angles made in different particle decay modes. Some modes are sensitive primarily to SM tree diagrams, while others involve loop diagrams and can more readily be influenced by new physics.

1.2.4 Types of CP violation

In the Standard Model CP violation can arise in three different ways. The first is called "Direct CP violation" which happens when the decay amplitude for a particle is different to that of its antiparticle ($\Gamma[M \to f] \neq \Gamma[\overline{M} \to \bar{f}]$). This is the only mechanism of CP violation in the decays of charged particles. Direct CP violation was established in kaons by the NA48 and KTEV experiments [81, 36] and has been seen in the B system. LHCb observed direct CP violation in the B_s system [44]. The recent LHCb observation of CP in the charm system by LHCb is also of this form [57].

The second type of CP violation is CP violation in the mixing of neutral particles. This happens because the mass eigenstates of certain neutral mesons are superpositions of flavour eigenstates. For example, for a neutral hadron N the two mass eigenstates N_+ and

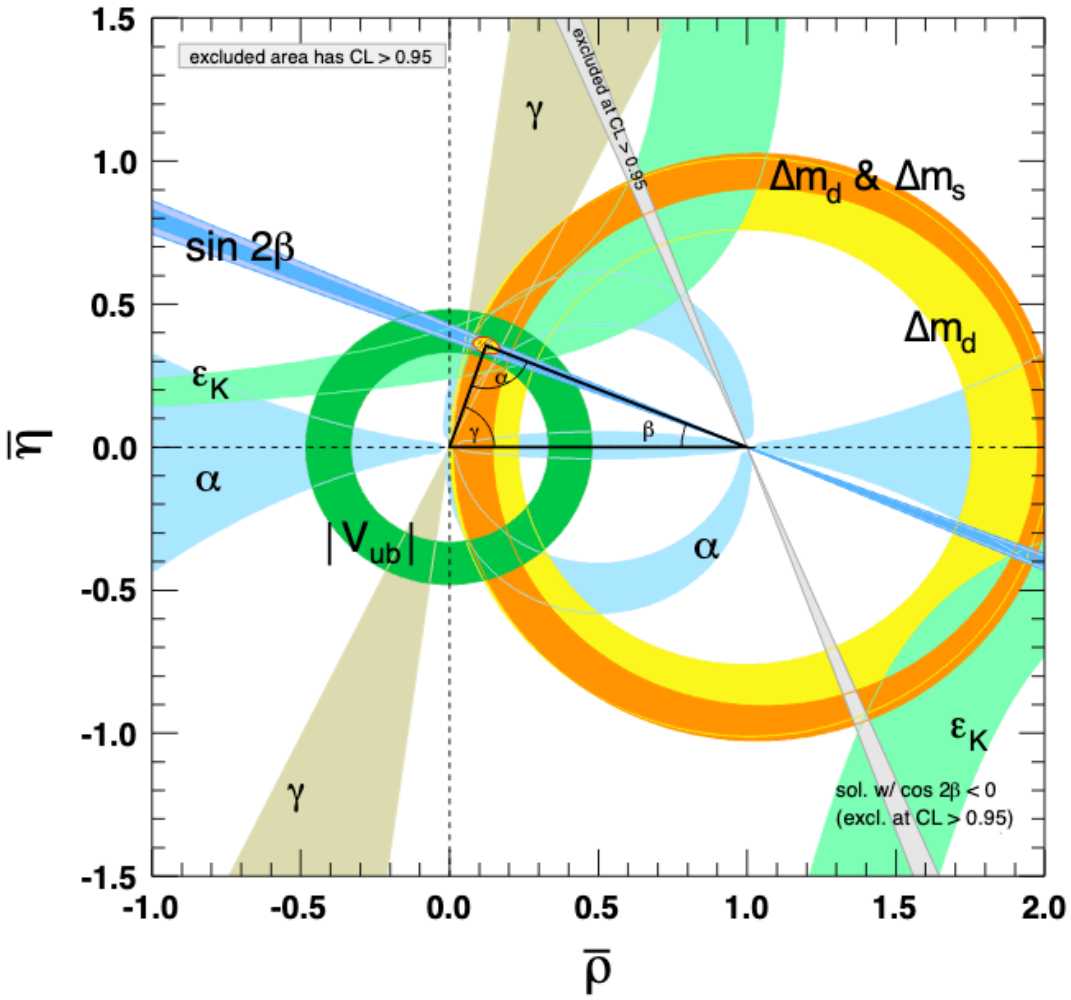


Figure 1.2: Constraints of the unitarity triangle in the $\bar{\rho}, \bar{\eta}$ plane. The $\bar{\rho}$ and $\bar{\eta}$ are a commonly used unitarity triangle coordinates that can be expressed as certain ratios and products of the CKM matrix elements. The small shaded area at the apex of the triangle corresponds to 95% CL. Reproduced from [83].

N_- are

$$|N_+\rangle = p|N^0\rangle + q|\bar{N}^0\rangle \tag{1.11}$$

$$|N_-\rangle = p|N^0\rangle - q|\bar{N}^0\rangle, \tag{1.12}$$

where N^0 and $\bar{N}^0$ are the flavour eigenstates with p and q being the complex amplitudes satisfying $|p|^2 + |q|^2 = 1$. When a neutral meson propagates in space, it oscillates between its particle and antiparticle states. If the transition probabilities to these states are different for particle and antiparticle (i.e. $|\frac{p}{q}| \neq 1$) the process will violate *CP* symmetry.

The final type of *CP* violation is the interference between direct *CP* violation and mixing *CP* violation amplitudes. In this case, the decay amplitudes of a neutral hadron N^0 into a final state f without mixing, $N^0 \to f$, interfere with the decay amplitudes of N^0 into the same final state but with mixing, $N^0 \to \bar{N}^0 \to f$. This can only happen if both N^0 and $\bar{N}^0$ can decay into the same final state (e.g. *CP* eigenstate). For this type of *CP* violation to occur the following condition needs to be present,

$$arg(\lambda_f) + arg(\lambda_{\bar{f}}) \neq 0, \tag{1.13}$$

where $\lambda_f \equiv \frac{q}{p}\frac{\bar{A}_f}{A_f}$ and A_f is the decay amplitude.

CP violation in mixing or in the interference between direct *CP* violation and mixing is sometimes called "indirect *CP* violation". Indirect *CP* violation was the type observed first in the kaon system [25] and B systems [32]. However, it has not yet been observed in the charm system.

1.3 Amplitude Analysis

The decays of short-lived particles can be described by quantum mechanical amplitudes (sometimes called "Feynman amplitudes"). These describe the probability of transition of the physical system from an initial state to a final state and are usually denoted $\mathcal{M}(i \to f)$. Accessing the decay amplitudes is very useful as it allows calculation of particle lifetimes and branching fractions into specific decay modes.

Let us consider a weak decay $\mathcal{M}(i \to f)$ and its conjugate process $\overline{\mathcal{M}}(\bar{i} \to \bar{f})$. The decay amplitudes for these processes can be denoted A_f and $\overline{A}_{\bar{f}}$. The complex terms appearing in the amplitudes will appear as complex conjugates in the *CP*-conjugate amplitude. This means that the phases of such amplitudes will have opposite signs. In the Standard Model, such a complex phase is precisely the phase of the CKM matrix responsible for *CP* violation, hence it is called the "weak phase". Other phases do not change their sign under *CP* transformation and usually come due to strong interactions, hence they are called "strong phases". This allows us to expand a decay amplitude by writing out every contribution to it. In a simple two amplitude case this would be,

$$A_f = |a_1|e^{i(\delta_1+\phi_1)} + |a_2|e^{i(\delta_2+\phi_2)}, \tag{1.14}$$

$$\overline{A}_{\overline{f}} = |a_1|e^{i(\delta_1-\phi_1)} + |a_2|e^{i(\delta_2-\phi_2)}. \tag{1.15}$$

Here $|a_i|$'s correspond to the magnitudes of individual contributions to the total decay amplitude, ϕ_i's are the weak phases and δ_i's are the strong phases. Individual contributions might correspond to different intermediate resonances of the decay. In the above example, the *CP* asymmetry can be expressed as,

$$\mathcal{A}_f = -\frac{2|a_1 a_2|\sin\delta_2 - \delta_1 \sin\phi_2 - \phi_1}{|a_1|^2 + |a_2|^2 + 2|a_1 a_2|\cos\delta_2 - \delta_1 \cos\phi_2 - \phi_1}. \tag{1.16}$$

The amplitude model of a decay can be expressed in many different conventions. One useful convention is to split the amplitudes into terms that are *P*-even and *P*-odd. These will also be $\hat{T}$-even and $\hat{T} - odd$, where $\hat{T}$ is the motion reversal operator that reverses the momentum and spin directions [93].

CP violation can arise due to the contributions from *P*-even or *P*-odd terms in the amplitude model and can hence be split into *P*-even *CP* violation or *P*-odd *CP* violation. In the case of *P*-even *CP* violation, one needs to measure the asymmetry in the *P*-even observables, such as the decay rates. A simple example of a measurement for *P*-even *CP* violation is the S_{CP} method described in Chap. 4.

In a three-body decay only the *P*-even amplitudes are present and hence only *P*-even *CP* violation can occur. However, in a four-body decay, such as the case in the decays described in Chap. 5, *P*-odd amplitudes can be present. In order to probe *P*-odd *CP* violation the difference in the parity violation between different flavour particles needs to be measured. This can be achieved by constructing asymmetry variables that are sensitive to the *P*-odd (*T*-odd) amplitude terms. An example of such variable is a Triple Product Asymmetry - $a_{CP}^{\hat{T}\text{-odd}}$, used in the physics analysis in Chap. 5, where triple products are the products of final state momenta (i.e. $C_{\hat{T}} \equiv \vec{p}_1 \cdot (\vec{p}_2 \times \vec{p}_3)$).

$$a_{CP}^{\hat{T}-odd} \propto \sin(\phi_{even} - \phi_{odd})\cos(\delta_{even} - \delta_{odd}) \tag{1.17}$$

The Triple Product Asymmetry vanishes if $\phi_{even} - \phi_{odd} = 0$ or π. Similarly it vanishes if $\delta_{even} - \delta_{odd} = \frac{\pi}{2}$ or $\frac{3\pi}{2}$. This makes it complementary to the usual $\mathcal{A}_{CP} \propto \sin\delta_1 - \delta_2 \sin\theta_1 - \theta_2$, which vanishes if either $(\delta_1 - \delta_2)$ or $(\theta_1 - \theta_2)$ are equal to 0 or π. In essence, these two asymmetries are probing different regions of the decay phase-space.

It is possible to construct an amplitude model with terms that modify the values of different observables. This way, the model can be evaluated against the experimental measurements which would allow a better understanding of the underlying decay dynamics. Building and testing such an amplitude model constitutes an amplitude analysis. An example of an amplitude analysis of the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decay described in Chap. 5

using angular distributions of the decay was given in Ref. [92]. The amplitude model of a specific decay topology can contain terms arising due to the existence of certain resonances and interference between them.

1.4 Review of CP violation in baryons

The study of *CP* violation is an exciting and important part of particle physics. However, all of the experimental observations of *CP* violation have been made in meson decays, while no *CP* violation has been observed in the decays of baryons. This leaves a wide and unexplored field for the studies of *CP* violation.

Previous measurements of baryonic channels have been limited by the small number of baryons produced in the experiments prior to the LHC. The only measurement of direct *CP* violation in beauty baryons was performed by the CDF collaboration with the statistical uncertainty of $\mathcal{O}(0.1)$ [37]. The state of these measurements are rapidly changing with the large datasets of baryonic decays provided by the LHCb detector at the LHC. The LHCb detector has measured the *CP* asymmetries is a number of baryonic channels: $\Lambda_b^0 \to pK_S^0\pi^-$ [77], $\Lambda_b^0 \to J/\psi\, p\pi^-$, $\Lambda_b^0 \to J/\psi\, pK^-$ [63], $\Lambda_b^0 \to \Lambda K^+\pi^-$, $\Lambda_b^0 \to \Lambda K^+K^-$ [64], $\Lambda_b^0 \to pK^-\mu^+\mu^-$ [61], $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ [72]. Also, a ΔA_{CP} measurement has been performed in the $\Lambda_c^+ \to ph^+h^-$ channel, where h is either a kaon or a pion [40].

An interesting family of baryonic decays are the charmless four-body beauty decays, such as $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$, $\Lambda_b^0 \to pK^-\pi^+\pi^-$. These decays propagate through $b \to u\bar{u}d$ and $b \to s\bar{s}d$ transitions that are topologically similar to the B_0 mesonic decays, where *CP* violation has been established. Furthermore, the four-body final state allows for *P*-odd amplitude contributions and hence, *P*-odd *CP* violation in the decays.

Firm theoretical predictions on the amount of *CP* asymmetry expected in a specific hadronic decay channel are not currently possible. However, the evidence for *CP* violation obtained in the previous LHCb analysis [53] has sparked some interest among the theorists. For example, Ref. [92] claims that certain P-odd asymmetries in four-body Lb decays are expected to vanish. While Ref. [113] predicts *CP* asymmetries of up to 20% in two-body $\Lambda_B^0 \to pK^{*-}$ decays. The significant number of resonances in the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ channel mean that large local *CP* asymmetries might be induced. Furthermore, its comparatively large branching ratio mean that it is arguably the best B baryon channel in which to search for *CP* violation. An analogous situation in B mesons is $B^+ \to 3h$ decays, where h is either a kaon or a pion. These decays are fully hadronic with large branching ratios and the potential for significant interfering resonances. In this case asymmetries have been observed by LHCb reaching up to $\mathcal{O}(50\%)$ in some regions of the decay phase space [54, 41, 60]. This surprising result was unexpected and motivated the creation of new theoretical models (e.g. Charm Rescattering proposed in Ref. [16]).

A previous LHCb measurement has found evidence for *CP* violation in the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays at just over the 3σ level [53]. The analysis presented in Chap. 5 constitutes an update on the previous measurement with a significantly larger data sample and is published in Ref. [68]. This result supersedes the previous result.

Chapter 2

The LHCb Experiment

The current chapter briefly describes the accelerator complex at CERN, focusing on the Large Hadron Collider (LHC) and its four main experiments. This is followed by a more detailed overview of the LHCb detector, which was used to collect data for the physics analysis described in Chap. 5. Special emphasis is given to the description of the VELO subdetector. The radiation damage studies of the VELO are desribed in Chap. 3. Afterwards, a brief overview of LHCb Software architecture is given, followed by a note on the currently ongoing LHCb Upgrade work.

2.1 Large Hadron Collider

The Large Hadron Collider (LHC) [97] is a 27 km synchrotron at CERN (Geneva, Switzerland). It produces proton-proton or heavy ion collisions, used for the particle physics research by the four main LHC experiments: ATLAS, CMS, LHCb and ALICE [31, 39, 78, 27]. These experiments have different designs and physics applications. The ATLAS (A Toroidal Lhc ApparatuS) [31] and CMS (Compact Muon Solenoid) [39] are general-purpose detectors with a varied physics programme, including searches for BSM physics and new heavy paricles. The ALICE (A Large Ion Collider Experiment) [27] is dedicated for heavy ion (PbPb) collisions and studies of the Quark Gluon Plasma. The LHCb (LHC beauty) [78] detector was originally designed as a heavy beauty and charm hadron factory, which focuses on *CP* violation searches and rare decays, but now has a broad programme including such topics as QCD, ion collisions, fixed-target collisions, electroweak, long-lived particles and hadron spectroscopy. A detailed description of the LHCb detector follows in Sec. 2.2.

The LHC is installed in the tunnel previously used by the LEP accelerator [127] and consists of 1232 dipole and 392 quadruple magnets that keep the beams in their circular orbits as well as focus them. These superconducting magnets operate at the temperature of 1.9K, which is possible because of an advanced liquid-helium based cooling system. The magnets create a field of up to 8.3 T and accelerate protons up to a centre-of-mass energy of 14 TeV [1]. In order to reach this energy, the protons go through a multi-step acceleration chain, shown in Fig. 2.1. Each step allows for higher energy to be achieved,

[1] In 2011 $\sqrt{s} = 7$ TeV was used, with an increase to 8 TeV in 2012 and to 13 TeV in 2015-2018

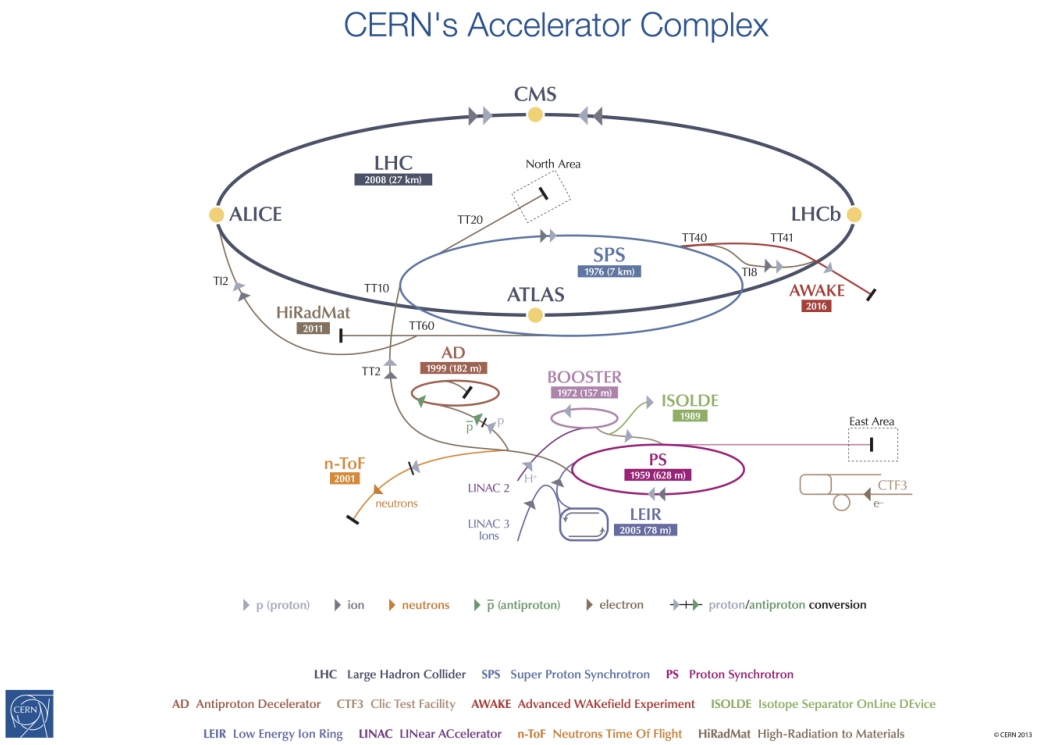


Figure 2.1: The diagrammatic representation of the accelerator complex at CERN, used to boost protons for the LHC. Reproduced from [107].

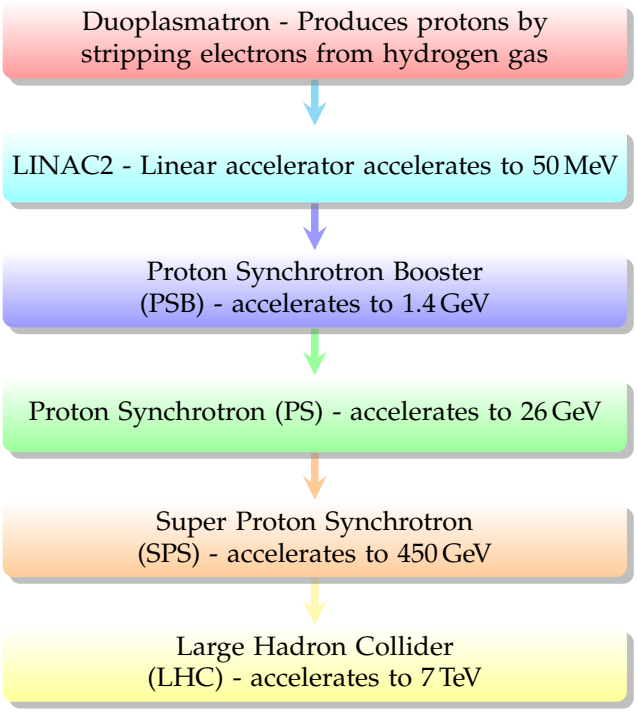


Figure 2.2: Different acceleration steps required to reach the LHC centre-of-mass energy of 14 TeV.

as shown in the diagram 2.2.

The LHC beam has the design luminosity of $10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ and consists of 2808 bunches with approximately 1.1×10^{11} protons per bunch. Each bunch is 53 mm long and has spacing equivalent to 24.95 ns between collisions. The rate of bunch crossings is 40.08 MHz. This makes LHC the most powerful accelerator to date with an unprecedented luminosity, creating interesting challenges for the detector design and especially data readout.

One way to control the number of collisions at an interaction point is luminosity levelling. This is particularly important for the LHCb detector because it is designed to operate at lower luminosities than the general purpose detectors. At the LHCb, luminosity levelling is achieved by controlling the mean number of proton-proton interactions (μ). This is done by offsetting the beams in the horizontal direction. The beam offset is then reduced throughout the fill in order to have a constant instantaneous luminosity. This leads to 1-2 primary vertices per bunch crossing, which is much smaller compared to 20 at ATLAS and CMS. The instantaneous luminosity over a single LHC fill is shown in Fig. 2.3. During the two runs of the LHC, the LHCb detector collected over $9\,\mathrm{fb}^{-1}$ of integrated luminosity. The integrated recorded luminosity for different years of LHCb operation can be seen in Fig. 2.3.

2.2 LHCb Experiment

The LHCb experiment is designed for heavy flavour physics analyses at the LHC. These analyses include searches for *CP* violation, precision measurements of the CKM matrix parameters, heavy hadron spectroscopy, rare decays and lepton flavour universality measurements. Notable discoveries done by the LHCb collaboration include the observation of *CP* violation in charm decays [18], discovery of a pentaquark state [58], discovery of the doubly-charmed Ξ_{cc}^{++} baryon [62], observation of rare $B_S^0 \to \mu^+\mu^-$ decay [115] and many others.

The design of the LHCb detector is optimized for precision measurements of charm and beauty hadrons [78]. In the proton-proton collisions, the b and c quarks are mainly produced through gluon fusion with a large forward boost and can be detected at small angles from the beam pipe as shown in Fig. 2.4. To provide the best coverage for these particles, LHCb is built as a single arm forward spectrometer, covering the pseudorapidity [2] range of $2 < \eta < 5$ (15 to 300 mrad). This geometry also has the added benefit of reducing the material budget and allowing for the readout electronics to be placed outside of active detector area. This in turn allows for excellent momentum resolution of the detected particles.

[2] Pseudorapidity is defined as $\eta = \frac{1}{2} log(\frac{|\mathbf{p}|+p_z}{|\mathbf{p}|-p_z})$

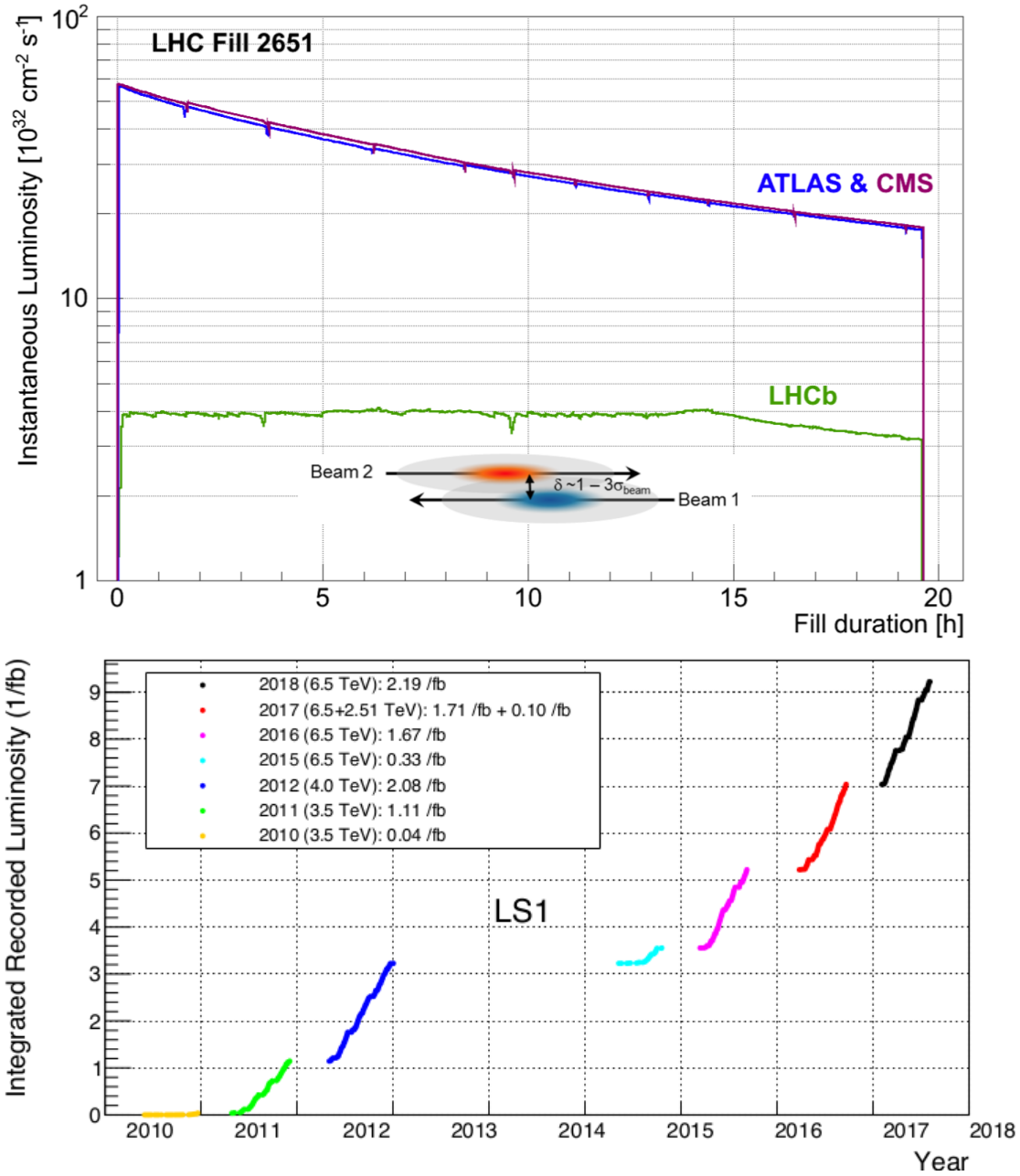


Figure 2.3: (Top) The change in instantaneous luminosity over a single LHC fill for ATLAS, CMS and LHCb. The luminosity of LHCb remains constant within 5% by changing the beam overlap, which is depicted at the bottom of the figure. (Bottom) The integrated luminosity recorded by the LHCb detector in different years of data taking. Reproduced from [46].

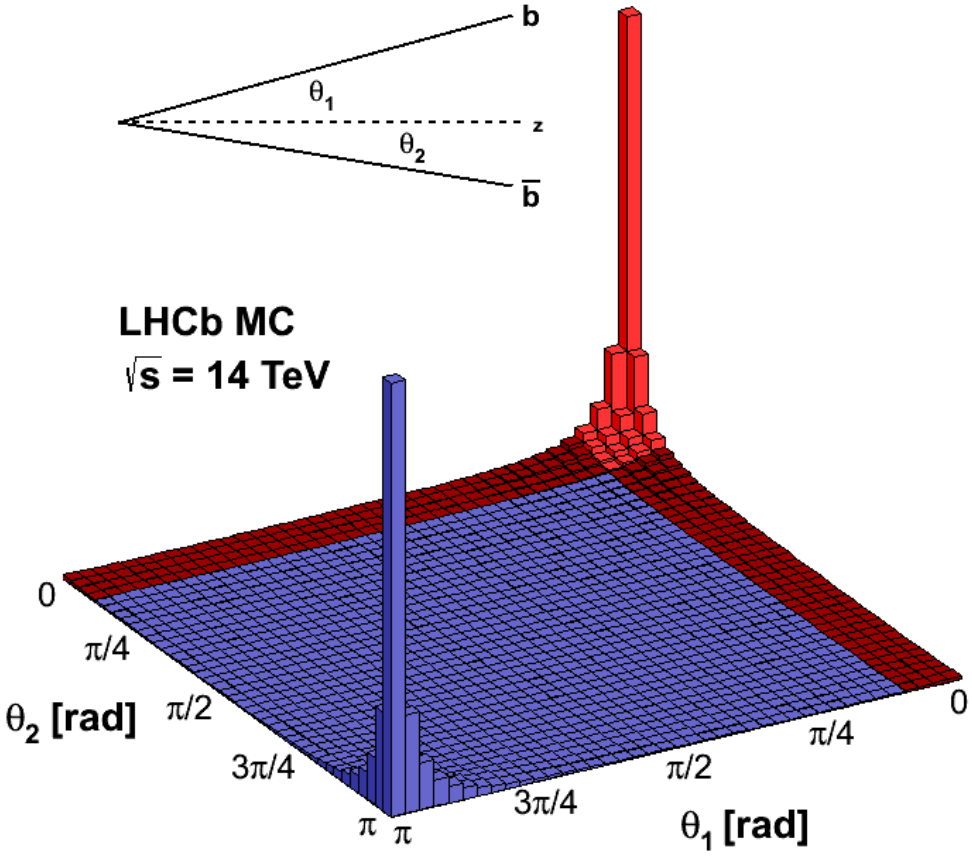


Figure 2.4: Rapidity distribution for $b\bar{b}$ production at the LHC. The red region shows the LHCb angular acceptance, which provides a good coverage of the events. Reproduced from [96].

The LHCb detector consists of multiple subdetectors. The interaction point is surrounded by the Vertex Locator (VELO), followed by the first Ring Imaging Cherenkov Detector (RICH), first silicon tracker detector - Tracker Turicensis (TT), bending magnet, three tracker stations, each consisting of Inner Tracker (IT) and Outer Tracker (OT), second RICH, first muon station, Scintillating Pad Detector (SPD), PreShower (PS), electromagnetic and hadronic calorimeters (ECAL and HCAL) and 4 more muon stations. The layout of the LHCb detector is shown in Fig. 2.5. Further explanation of the design and functions of the LHCb subsystems is given in the following sections.

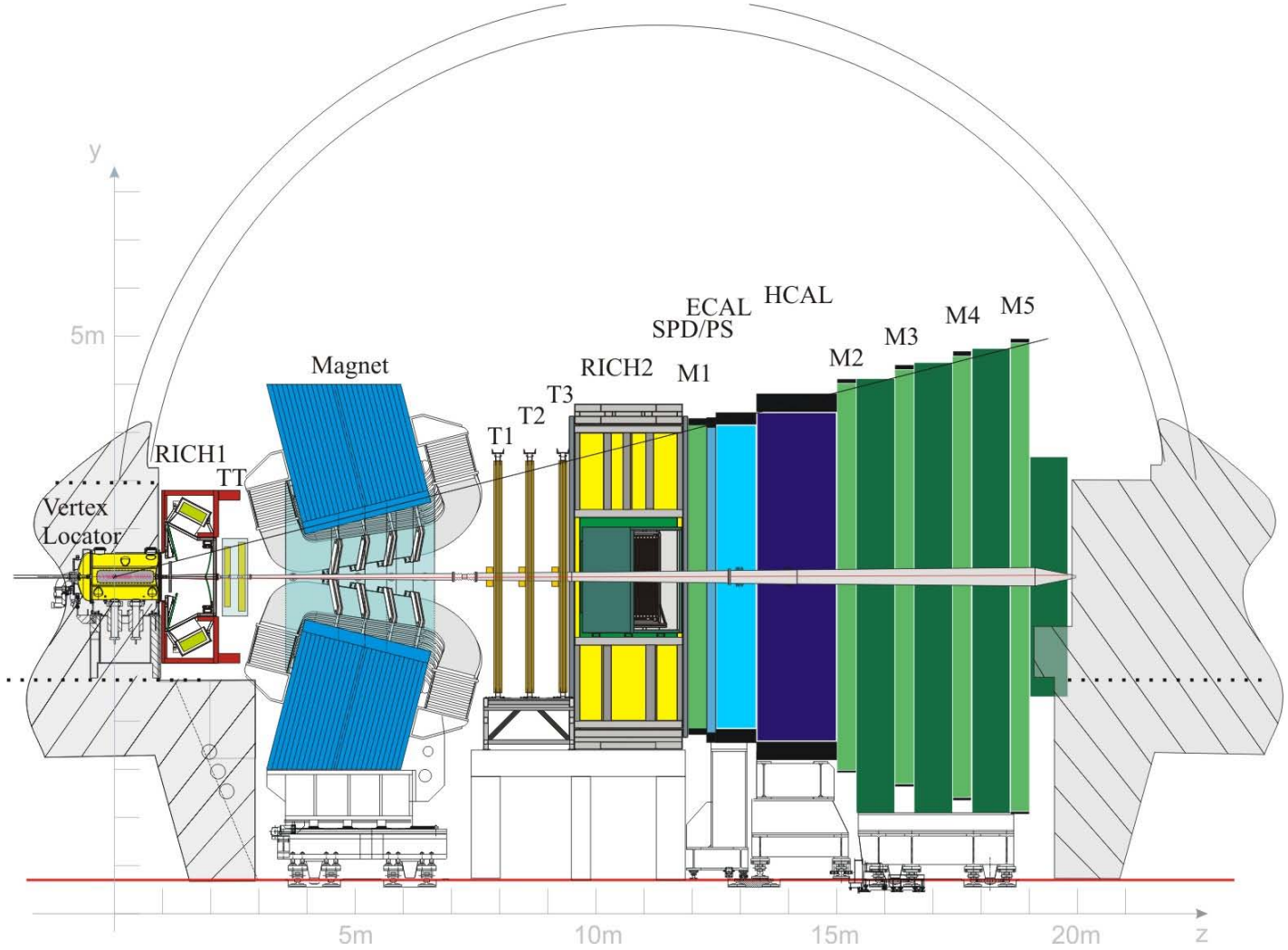


Figure 2.5: The schematic representation of the LHCb detector, showing different subsystems. Reproduced from [45].

2.2.1 Vertex Locator (VELO)

The subdetector surrounding the proton-proton interaction region at the LHCb is the Vertex Locator (VELO) [79]. It is designed to measure the locations of primary vertices from proton collisions and secondary decay vertices from short lived particle decays. The VELO is based on silicon-strip technology and has two movable halves consisting of 21 modules each, arranged along the beam axis as shown in Fig. 2.6.

Each semicircular module consists of two sensors, allowing measurement of the radial and angular position of a particle intercept at a precision of 13 μm in the transverse plane and 71 μm along the beam axis. Most of the sensors use n^+ implants on n-doped silicon while a pair of the most upstream sensors use n^+ implants on p-doped silicon to test the technology. Each sensor contains 2048 strips embedded in the 300 μm-thick silicon bulk for charge collection. The layout of these strips corresponds to either R or Φ measurement configuration.

In an R sensor, the strips are split into four 45° regions (zones). Each zone contains 512 strips that are arranged as arcs of a circle and vary in pitch linearly from 38 μm closest to the beam to 102 μm at the far edge of the sensor. The Φ sensors have two "fan-shaped" regions. The inner region contains 683 strips, while the outer one contains 1365 strips. This helps to reduce occupancy and prevents the strips from having a large pitch at the far edge. The Φ sensor strips are not perfectly radial but have a 10° stereo-angle inclination for the inner region and 20° for the outer region, allowing for better pattern recognition.

Due to the use of two different geometries for the VELO sensors, different solutions for charge readout connections to the electronic system had to be implemented. The strips that are further away from the far edge of the sensor need to be connected to the readout chain via an additional conductive routing lines, referred to as the second metal layer. In the Φ sensors, the routing lines are parallel to the strips while in the R sensors they are orthogonal. The layout of strips and second metal layer is shown in Figs. 2.7.

Another unique feature of the VELO is that the two detector halves can be moved closer to the beam or away from it. This is essential because during the proton beam injection and ramp-up the beam is unstable and can deviate from its nominal trajectory. For this reason, before the beam reaches required stability the VELO is kept in an "open" configuration and the closest detector component is 35 mm from the beam. When the beams stabilize, the VELO is closed around the precise location of the beam, which differs from fill to fill. In the "closed" configuration the closest part of the VELO is only 7 mm away from the beam, while the closest active part is 8.3 mm away. This made the VELO the closest subdetector system to the proton beam at the LHC. Furthermore, when the VELO is closed, the neighbouring sensors overlap by approximately 1.5 mm, covering the full acceptance and allowing for better alignment.

There is no beam pipe inside the VELO and the detector is in vacuum. The VELO modules are separated from this primary beam vacuum by a 300 μm-thick aluminium box, which prevents the RF pickup from the proton beam. For this reason, the VELO has to be kept in a secondary vacuum. Being in a secondary vacuum also prevents any possible contamination of the primary vacuum from outgassing of the modules. Furthermore,

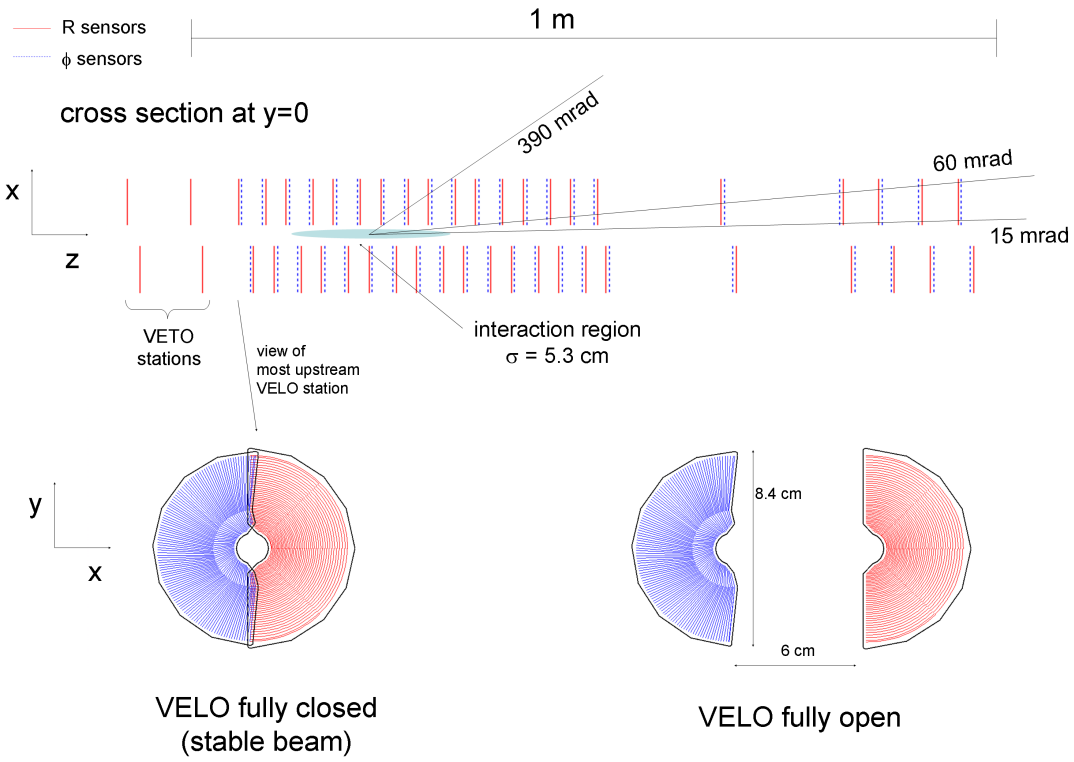


Figure 2.6: (Top) Diagram showing the positions of VELO modules along the beam axis with interaction region highlighted in blue. (Bottom) Example of two VELO modules in open (right) and closed (left) configurations. Reproduced from [79].

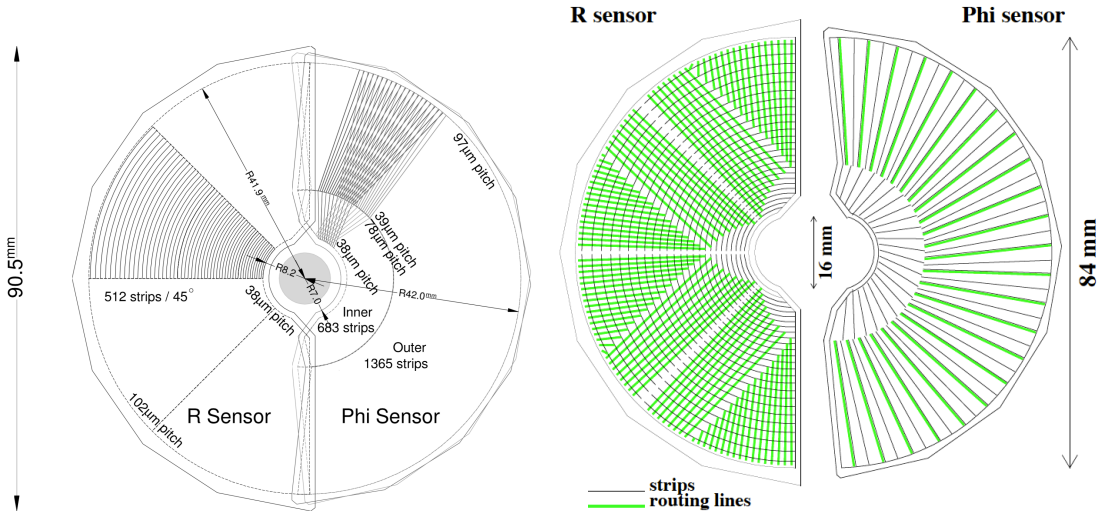


Figure 2.7: The layout of conducting strips (left) and the routing lines, referred to as the second metal layer (right), for R and Φ VELO sensors. Reproduced from [78, 79].

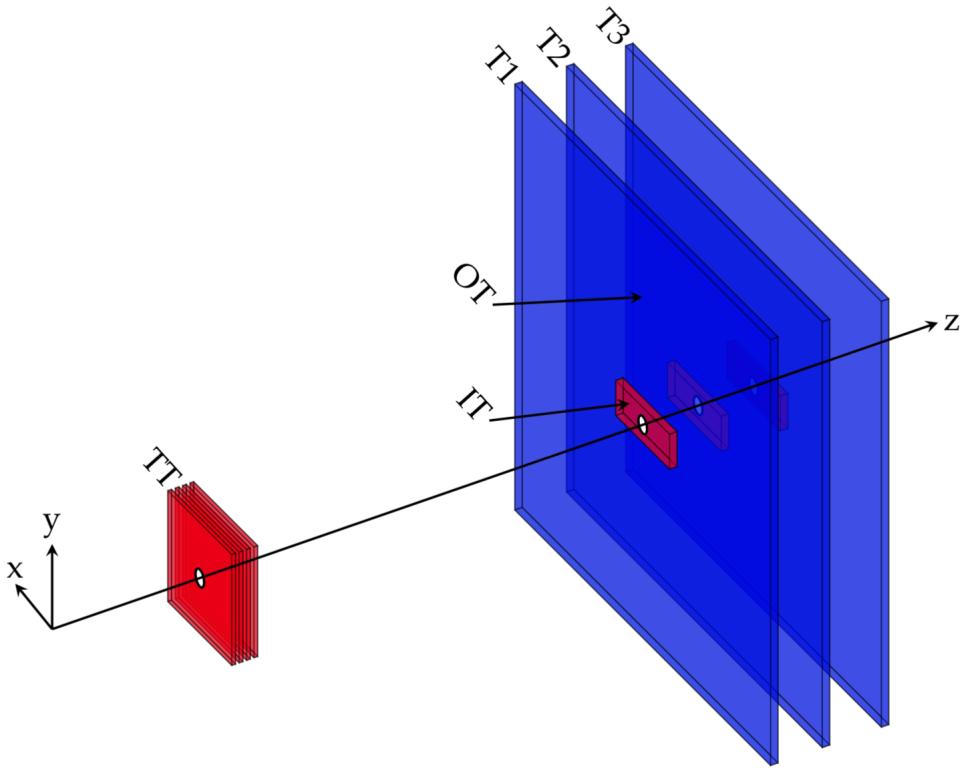


Figure 2.8: The layout of LHCb Tracking Stations. The TT is situated before the magnet and the IT is situated in each of the three tracking stations after the magnet. Together, the TT and IT form the Silicon Tracker (ST), shown in red. The IT is surrounded by OT, shown in blue. Reproduced from [109].

the VELO modules are cooled by a CO_2 system to about $-10°C$ during operation. The average temperature of the VELO during Run 1 was approximately $-7°$C.

Due to its close proximity to the beam, the VELO is exposed to an extreme radiation environment. During the first two runs of the LHC, VELO sensors were under the fluence of up to $6.5 \times 10^{14}\,\mathrm{cm}^{-2}$ 1 MeV neutron equivalents ($1\,\mathrm{MeV}n_{eq}$). In this environment significant radiation damage of the sensors is expected. The analysis of VELO radiation damage and its effects on physics performance is described in Chap. 3.

2.2.2 Tracking Stations

The LHCb tracking system is comprised of the VELO and the tracking stations situated downstream from the interaction point. The tracking stations contain the Tracker Turicensis (TT), which is situated after the VELO and RICH, upstream from the LHCb bending magnet. There are three more tracking stations downstream from the magnet, each containing an Inner Tracker (IT) and an Outer Tracker (OT). The tracking stations are shown in Fig. 2.8.

The TT is a silicon micro-strip detector which uses p^+ implants on n-doped silicon sensors and covers the full LHCb acceptance. It measures 157.2 cm in width and 132.4 cm in height. The TT contains four sensor layers, each with the width of 500 µm. The layers have long silicon strips, with the longest being 37 cm. The strip pitch is 183 µm. In the

first and fourth layer the strips are vertical. The second and third layers are rotated by a stereo-angle of -5° and $+5^\circ$ respectively. This allows the TT to have a hit resolution of approximately 53 μm.

The LHCb bending magnet is placed downstream from the TT. The magnet is a warm dipole (non-superconducting) and is able to bend the trajectories of charged particles in the x plane. This allows measurement of the particle's momentum from the curvature of the tracks in the magnetic field. The design of the magnet covers the full LHCb acceptance and provides an integrated field of 4 Tm and a peak strength of 1.1 T.

The tracking system downstream from the magnet contains three tracking stations (T1, T2, T3). Each of these have an Inner Tracker (IT), which uses similar silicon technology as the TT. For this reason, the TT and IT collectively comprise the Silicon Tracker (ST) project [121]. However, the IT only covers the regions closest to the beam pipe, where the particle flux is highest. Each of the three tracking stations contain four IT boxes, which in turn contain four layers of silicon. Similar to the TT, the second and third layers are rotated by a stereo-angle of -5° and $+5^\circ$. The sensors in each layer use 320 μm-thick silicon for the top and bottom modules and 410 μm-thick for the side modules.

The IT is surrounded by the Outer Tracker (OT) [66], which uses different technology from the ST system. The OT is a straw-tube detector which allows for precise measurement of particles' momenta and provides tracking via the drift-time measurement. The 4.9 mm inner diameter straw tubes contain a mixture of Argon, CO_2 and O_2 gases (with 70/28.5/1.5 ratio). This allows a drift-coordinate resolution of 205 μm and drift time of 50 ns. However, this time is twice as long as the minimum time between LHC bunch crossings, hence spillover effects occur. Just like the ST system, the OT contains four layers with second and third being rotated by $\pm 5^\circ$.

The tracking system is able to achieve momentum measurements with relative uncertainty of 0.5% for low momentum particles and 1.0% for high momentum particles at 200 GeV [46].

2.2.3 Ring Imaging Cherenkov Detector (RICH)

Ring Imaging Cherenkov Detectors (RICH) [105] are placed on both sides of the LHCb magnet. These detectors are filled with a gaseous medium which has a refractive index (η) greater than 1. For this reason, fast particles travelling through the RICH will travel faster than the local speed of light and emit Cherenkov radiation. The Cherenkov photons will be emitted in a cone, with the angle proportional to the velocity of the particle [3].

Combining the information about the particle's velocity with the information about its momentum from the LHCb tracking system allows assignment of a mass hypothesis for the particle (PID). Thus, it allows identification of different particle species. The most abundant particles in the LHCb are pions, hence all detected particles are compared against the hypothesis of them being a pion. This comparison is quantified by assigning a difference in log-likelihood value for a specific particle ID when compared to a pion

[3] Cherenkov angle is calculated as $\theta_c = arccos(\frac{c}{\eta v})$, where v is the velocity of the particle in the RICH medium

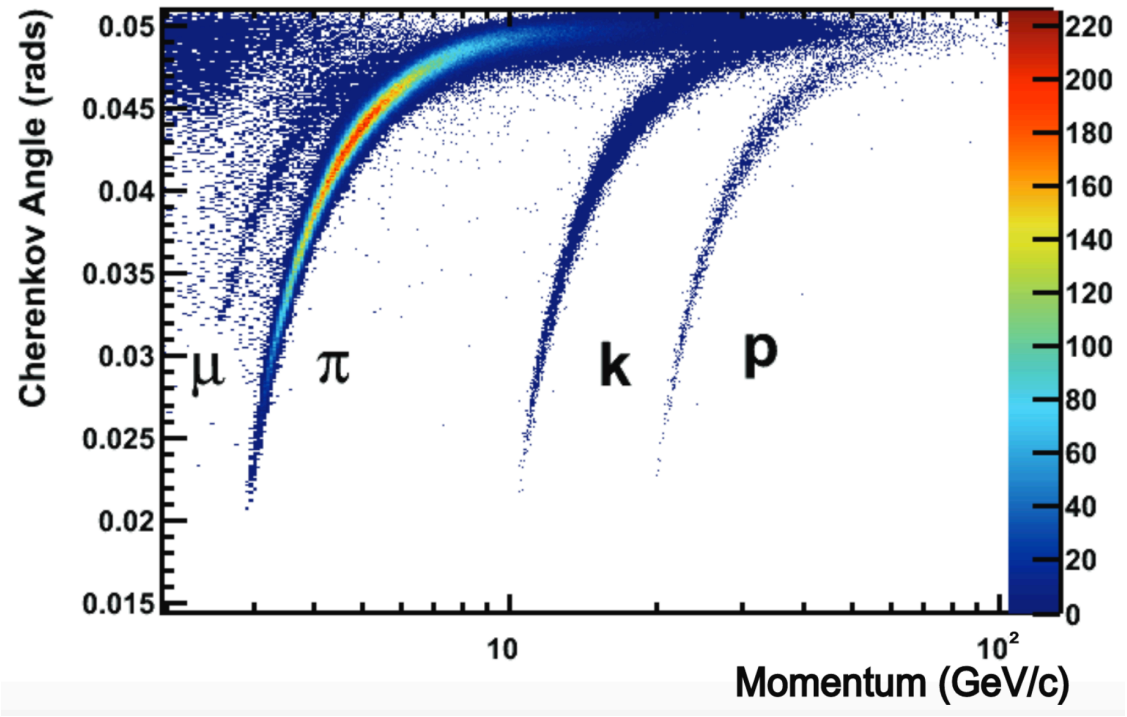


Figure 2.9: Reconstructed Cherenkov angle as a function of track momentum in the C_4F_{10} radiator. The different bands correspond to different particle species and give a clear separation. Reproduced from [105].

hypothesis. The different particle species can be seen by plotting the track momentum against the Cherenkov angle, as shown in Fig. 2.9.

In order to assign the PID hypothesis, the particle's velocity needs to be estimated. This depends on the particle's momentum and the medium used in the RICH detector. The RICH1 detector is situated between the VELO and the TT, upstream from the bending magnet. It is filled with C_4F_{10} gas and covers the momentum range between 2 and 40 GeV. The RICH1 provides full coverage of the LHCb acceptance. The RICH2 is placed after the tracking stations, downstream from the magnet. It is filled with CF_4 gas, which allows it to cover momenta between 15 and 100 GeV. Since the RICH2 is focusing on high momentum particles, it covers a high pseudorapidity range ($\eta < 3$). This is due to the fact that high momentum particles usually have high forward boost and travel at low angle from the beam pipe, while some low momentum particles are bent out of the LHCb acceptance by the magnet.

Both RICH detectors use primary spherical mirrors that reflect the Cherenkov light onto the secondary planar mirrors. This way the emitted photons are reflected out of the LHCb acceptance and into the Hybrid Photon Detectors (HPDs), where the circles formed by the Cherenkov cone projected on a plane can be reconstructed. Since the detectors are close to the LHCb magnet, the fringe magnetic field distorts these circles and further corrections have to be applied.

2.2.4 Calorimeters

The LHCb calorimeters are positioned downstream from the RICH2. They allow measurement of the deposited energy of electrons, photons and hadrons. However, the energy of charged particles can typically be obtained from momentum measurements, attained from the combined output of the tracker system. For this reason, the calorimeter system is not used in many analyses involving charged particles. Though for electrons it is used to add the additional energy from emitted photons. The main purpose of the LHCb calorimeters is to act as a first trigger stage (described in Sec. 2.2.6). Another use of calorimetry is the identification and reconstruction of neutral particles, such as photons.

The LHCb calorimetry system is based on the detection of scintillating light. This light is emitted inside a scintillator material when it is exposed to ionising radiation i.e. intercepted by a high energy particle. The photons from the scintillation travel through wavelength shifting fibres and are detected by Multianode Photomultiplier Tubes (MaPMTs).

The first layer of the calorimeters is the Scintillating Pad Detector (SPD), which only interacts with charged particles, thus distinguishing between charged and neutral particle species. It is followed by a 15 mm-thick lead converter and the PreShower (PS) calorimeter that discriminates between photons and neutral hadrons. The third layer is the Electromagnetic Calorimeter (ECAL), which measures the energy deposited by electrons, photons and photons emitted from neutral pions. The final component of the calorimetry system is the Hadronic Calorimeter (HCAL), which measures the energy of hadronic showers and is mainly used for triggering, although in jet based analyses (e.g. QCD electroweak) it is also used for neutral hadrons.

All four calorimeters have finer segmentation around the beam and courser segmentation further away from it. The ECAL and HCAL are designed as "shashlik" detectors with alternating layers of scintillating tiles and absorber material (lead for ECAL and iron for HCAL). The ECAL is able to contain the full electromagnetic shower and has energy resolution of $\sigma_E/E = (8.2 \pm 0.4)\%/\sqrt{E/\text{GeV}} \oplus (0.87 \pm 0.07)\%$. The HCAL is too short to contain the full hadronic shower and and is used for triggering. The energy resolution of the HCAL is: $\sigma/E = (69 \pm 5)\%/\sqrt{E} \oplus (9 \pm 2)\%$.

2.2.5 Muon System

The final subdetector of the LHCb detector is the muon system [65], which contains five stations (M1-M5). The first muon station (M1) is placed before the calorimeters, while four remaining stations are placed after the HCAL. All muon stations use multi-wire proportional chamber (MWPC) technology with highest granularity closest to the beam pipe. However, due to a large particle flux, the inner-most region of the M1 station also uses a triple gas electron multiplier (GEM) technology.

The first three muon stations allow estimation of the muon transverse momentum, while the last two stations are used to identify penetrating muons. Iron absorber plates of 80 cm thickness are placed after each of the four stations behind the calorimeters (M2-M5),

thus selecting the penetrating muons. The muon momentum required to reach M5 is at least 6 GeV. The muon system is used in the first stage hardware trigger, described in Sec. 2.2.6, as well as for muon reconstruction in the offline analysis.

2.2.6 Trigger

The bunch-crossing rate at the LHC can reach up to 40 MHz. However, due to the bunch filling scheme, the bunch-crossing at LHCb reaches 30 MHz. Out of all of the collision events at LHCb, only a small fraction are interesting and have a clean signal. Furthermore, it would be impossible to save all of the events to storage due to readout system and permanent storage limitations. To filter out and save the events of interest, a three-stage trigger system is employed [42].

In the first stage, the LHCb uses a hardware trigger, called L0 (Level 0). This trigger is based on the output of calorimeters and the muon system and contains three inclusive conditions. The first condition triggers on charged hadron signatures on the HCAL and requires the transverse energy deposited to be higher than 3.5 GeV. The second condition triggers on a muon signal in the muon system by searching for hits that are consistent with originating from the same track. Afterwards, it applies a minimum transverse momentum requirement, thus rejecting muons originating from pion or kaon decays. The muon momentum can be estimated by calculating the expected energy loss of a muon traversing the muon system. The last condition triggers on photons, neutral pions or electrons/positrons using the ECAL.

The L0 trigger reduces the event rate to approximately 1 MHz. This rate is still too high to be fully stored, hence the events are passed to the software based trigger. The first software based trigger is called HLT1 (High Level Trigger) and performs partial event reconstruction. The HLT1 reconstructs primary vertices and estimates the impact parameter (IP - Distance between the point of closest approach of a track and primary vertex). Events with low IP or no hits in the muon stations are discarded. This way, the event rate is further reduced to 30 kHz in Run 1 (80 kHz in Run 2).

After the HLT1, the events are passed to the second software trigger stage, called HLT2. The HLT2 trigger performs full event reconstruction and checks it against a list of trigger conditions for different physics analyses. If the event matches any of the trigger lines, it is saved to permanent storage. The event storage rate is 5 kHz in Run 1 and 12.5 kHz in Run 2. The stored events undergo finer reconstruction in centralized processing campaigns called “stripping”, which produces output files of a size more suitable for use in offline analysis.

2.3 LHCb Software

The software infrastructure of LHCb is modular and based on C, C++ and Python packages. Each package contains a well defined scope of use cases and is regularly updated. The packages are developed in a way to allow communication between different packages.

This is achieved by assuring that the output of a specific package is readable by multiple packages further down the chain.

All of the LHCb software packages are based on an experiment independent Gaudi framework [11] which is specifically designed for HEP data processing. Within Gaudi, the LHCb specific software modules are responsible for different functions. The Brunel module is used for offline reconstruction. In Run 2 this is also done by the Moore module, to allow for reconstruction with Turbo Trigger. After the offline reconstruction, the data is further processed by DaVinci, which allows production of ROOT [21] tuples - a conveniently sized and structured data for the front-end analyst.

In order to perform sensitivity studies, test different physics models or detector simulations, a sophisticated Monte Carlo simulation is required. The simulated data for LHCb applications is obtained using the Gauss module [26]. Gauss uses auxiliary modules and libraries to perform a full simulation of the physics event. This simulation chain includes full detector description, particles' interactions with matter, the underlying QCD scattering event etc. Having a fast, reliable and realistic simulation is paramount for almost all particle physics analyses.

The LHCb software environment is constantly being updated. The big data challenges need to be overcome by developing efficient programs that exploit all of the state of the art. Many different options are considered, including the use of GPUs, FPGAs and Machine Learning. Apart from the hardware upgrades mentioned here, the data acquisition and event reconstruction software is also being improved and optimized. The software upgrades are based on modern programming paradigms of multi-threaded data processing and machine learning. These upgrades utilize many years of experience from the LHCb collaboration.

2.4 LHCb Upgrade

The LHC finished its second physics run in December 2018, thus starting the Long Shutdown 2 (LS2), which is expected to take until 2021. The LHC accelerator complex will be upgraded, as well as all of the LHC based experiments. The currently ongoing ATLAS [29] and CMS [123] upgrades are relatively minor compared with the full-scale of these detectors, their major changes will happen during the LS3. The first major scale changes to detector systems are occurring to the ALICE [28] and LHCb [134] detectors.

In the case of the LHCb detector, almost all of the subsystems will be replaced, essentially creating a whole new detector. This profits from the huge improvements in detector technology during the last decade, as well as replacements required for the ageing of the old detector due to the radiation damage. These upgrades are required because of the increase in instantaneous luminosity of the LHC. During Run 3 and Run 4 the LHCb detector is expected to collected approximately $50\,\text{fb}^{-1}$ of integrated luminosity, which is five-fold the luminosity collected in all of the previous years of running.

The new LHCb design is shown in Fig. 2.10. The biggest change of the upgrade is the 40 MHz readout ability by all of the LHCb detector electronics. The VELO will be

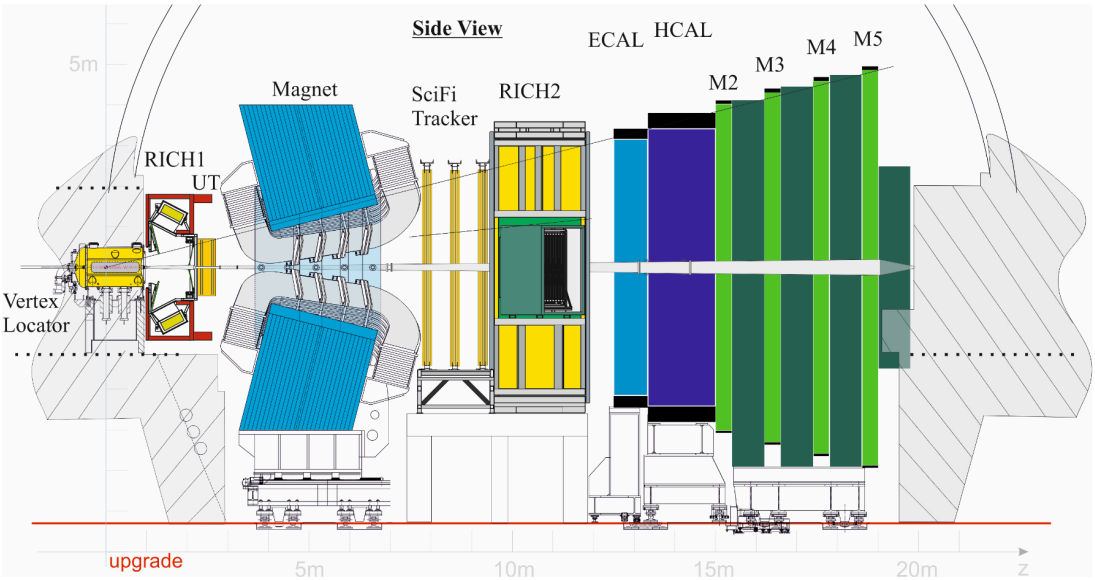


Figure 2.10: The diagram showing subsystems of the upgraded LHCb detector. Reproduced from [134].

entirely replaced with a new subdetector based on rectangular sensors that use silicon-pixel technology with 55x55 µm pixels and a low-mass micro-channel silicon substrate [49]. The RICH 1 detector will be entirely changed. Both of the RICH detectors will use new mirrors, optimizing the performance for higher luminosity operation [47]. The HPDs of the RICH detectors will be replaced by Multianode Photon Multiplier Tubes (MaPMTs) that offer higher resolution. The TT will be replaced by the Upstream Tracker (UT), which is based on silicon-strip technology with shorter strips and finer segmentation in some regions. The three tracking stations, containing IT and OT will be replaced by a large scale detector SciFi using 11,000 km of 250 µm scintillating fibre. The PS, SPD and M1 subdetectors will be removed, as they are no longer needed for triggering. The front-end of ECAL, HCAL and remaining muon stations will remain the same, but the electronics will be upgraded to satisfy the 40 MHz readout requirement.

Apart from the detector hardware upgrades, the Data Acquisition system (DAQ) will also be entirely replaced [48]. A major challenge for the LHCb upgrade is the ability to select and write interesting collision events to permanent storage. This has to be achieved without sacrificing the quality of each individual event, while maximising the output rate at the same time. In Run 3, a "Real-Time Analysis" (RTA) paradigm will be applied. Here "real-time" refers to the time between the physical collision event and the time the corresponding signal is saved to permanent storage.

Furthermore, after the HLT1 processing, the data will be saved in the disk buffer (see Fig. 2.11). This data can be used to run online alignment and calibration asynchronously from the software trigger reconstruction. This is done at the start of each LHCb run. If the run parameters are observed to have changed after the calibration, the new alignment and calibration constants are saved to a conditions database. The online alignment has already been implemented during Run 2 [42].

During Run 1 and Run 2 the data corresponding to the collision events at the LHCb

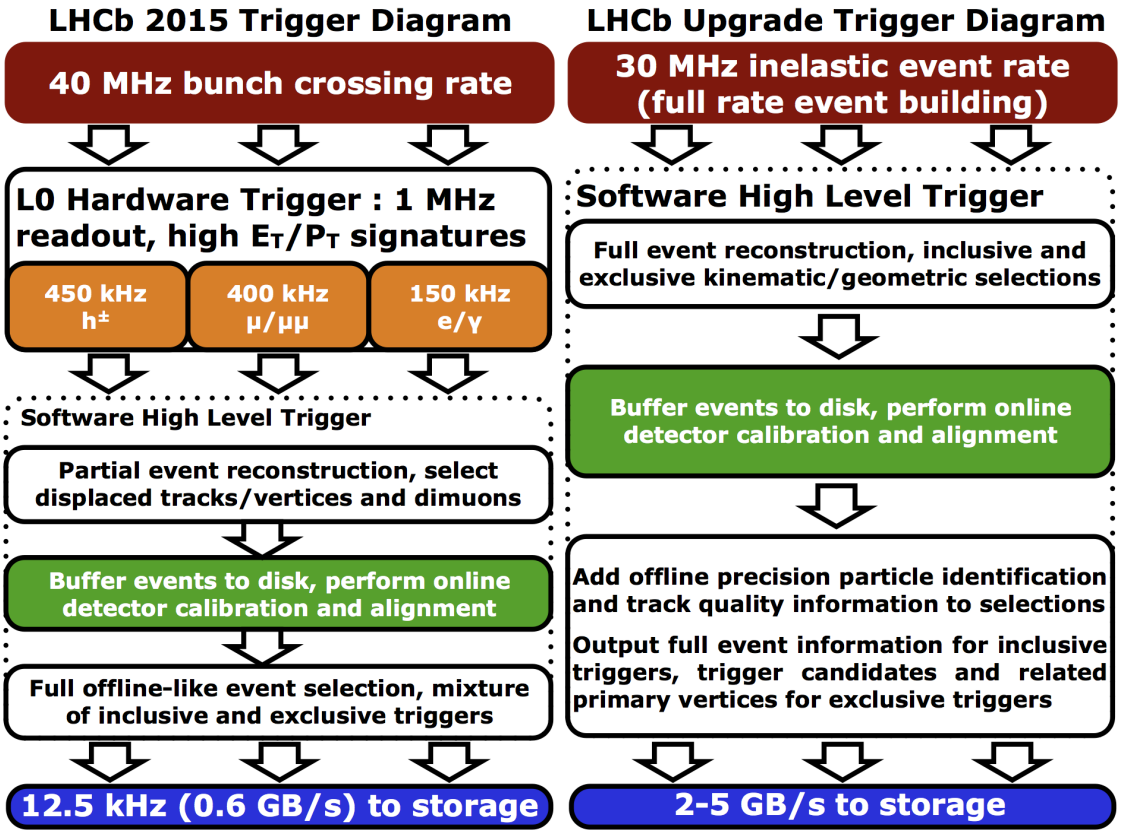


Figure 2.11: The diagrams of different trigger stages used in Run 2 (left) and Run 3 (right). As can be seen, the L0 hardware trigger is removed and the trigger is fully software based after the LHCb upgrade. Reproduced from [42].

had to be processed by the multi-stage trigger system, described in Sec. 2.2.6. The first trigger stage was a hardware trigger called the L0. However, it was seen that the L0 trigger would be the main bottleneck going to higher luminosity. The L0 trigger introduced a large inefficiency in the trigger chain which was affecting the beauty and charm hadron decays, but was small for muonic decays [42]. During Run 2 it was shown that a fully software trigger is capable of reconstructing the events with quality on par with the offline event reconstruction. Furthermore, with the increased instantaneous luminosity available after the upgrade, most of the events would trigger the L0 response, hence making it useless in the event discrimination. This motivates the major change to the trigger in the LHCb upgrade. The hardware trigger will no longer be used during Run 3. All of the events will be selected by a purely software trigger, which will provide a full event reconstruction and sufficient discrimination for channels of interest. The diagrams comparing the trigger stages in Run 2 and Run 3 are shown in Fig. 2.11.

Another major change in the event processing after the LHCb upgrade will be the use of a different event information persistence model. Each fully reconstructed event takes approximately 69kB of space in the permanent storage. However, most of the information stored is not useful for physics analysis. For example, it might be enough to have access to the momenta of the decay products of a heavy hadron decay candidate and

the positions of all primary vertices in that particular event. This allows reconstruction of the underlying decay and dissociation of the decay products from other vertices. Saving only this information to storage saves almost an order of magnitude of disk space. However, certain analyses might require more information about the decay (e.g. other hadrons originating from the same primary vertex). A middle ground between these two extremes is called the "selective persistence model" and is illustrated in Fig. 2.12. In Run 3, most of the events will be saved without retaining the full event information. This will allow higher throughput and event yields, while not compromising the quality of the data or the scope of possible physics analyses.

The above mentioned hardware and software upgrades are vital for the efficient running of the LHCb detector in Run 3. During this period, the instantaneous luminosity of the LHC will be five times larger than that of Run 2. The detector is expected to collect $50\,\text{fb}^{-1}$ of data during this period and have significant improvements in the trigger efficiency for many modes, which will lead to a huge improvement in statistical uncertainties of many important physics measurements.

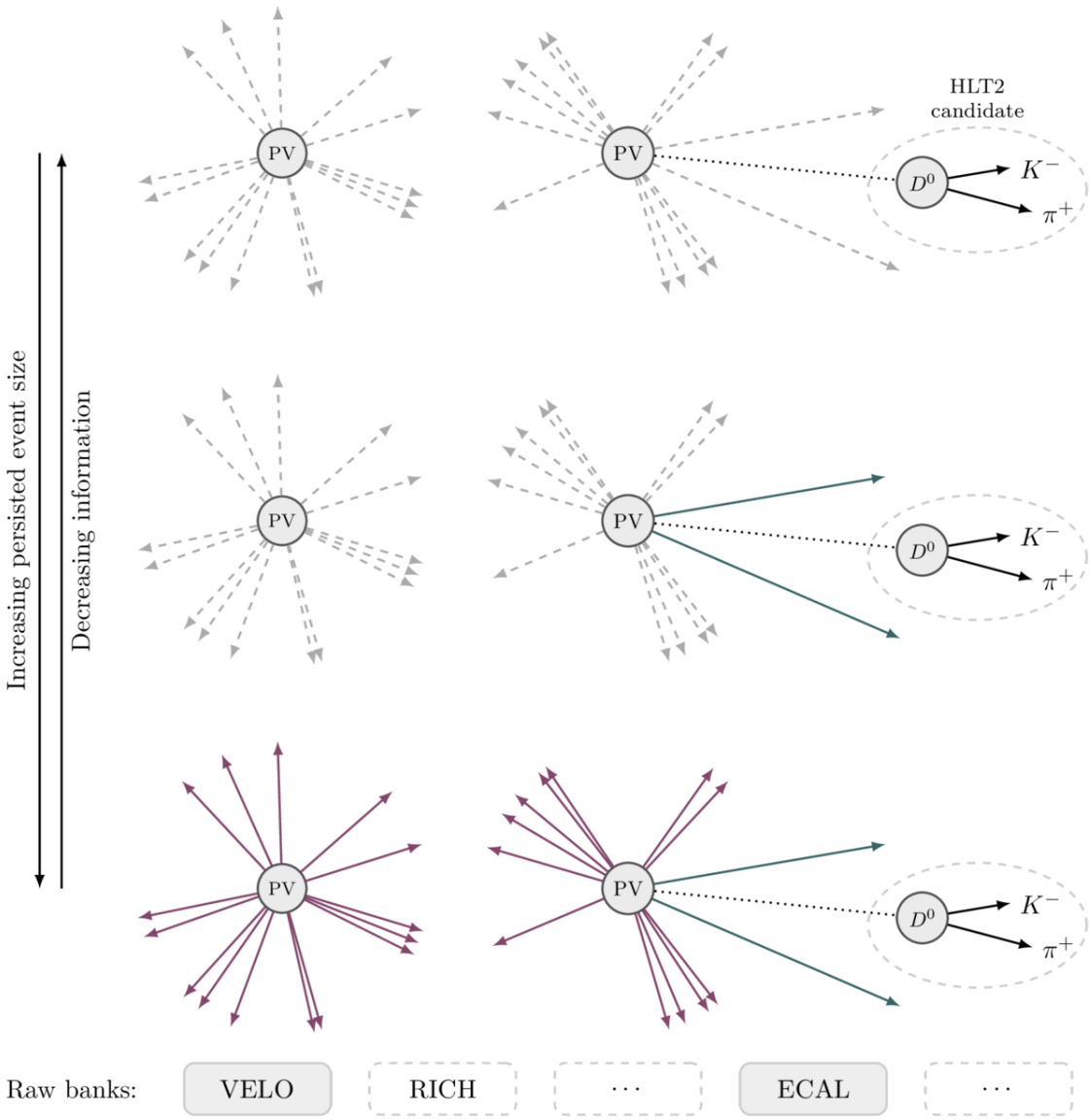


Figure 2.12: The diagram shows the same reconstructed $D^0 \to K^-\pi^+$ event with different levels of object persistence. Objects marked with solid lines are saved to storage after the event reconstruction. The top model saves only the final state momenta and primary vertices of the event, which is the minimal amount of information required to reconstruct the decay. The middle model additionally saves the objects that can be freely selected (e.g. charged kaons and pions originating from the same primary vertex as the D^0 candidate). The bottom model saves full information of the decay and takes up the most space in storage. Reproduced from [2].

Chapter 3

Radiation Damage of the VELO Subdetector

Many particle physics analyses depend on precise particle track reconstruction and vertexing. This is paramount when studying heavy hadron decays (such as in the analysis presented in Chap. 5). The hadrons created during the proton-proton collisions at the LHC travel only up to a couple of millimetres before decaying into more stable particles. In order to accurately measure the positions of primary and secondary decay vertices, silicon detectors are employed. Modern silicon detectors provide high precision, while using a minimal amount of material which particles need to traverse in order to be detected. This is referred to as the "material budget" and is an important consideration when designing a particle detector. Reducing the material budget minimises the probability of particles scattering or interacting with the detector. The active element of a detector can be formed from a thin, self-supporting and homogeneous silicon wafer. However, due to the complexity and high quality requirements, silicon detector technology is expensive. For this reason, the silicon detectors are usually only used in the regions with highest fluence and occupancy. For example, the VELO detector at the LHCb is based on silicon-strip technology and surrounds the interaction region. It is situated only 8.2 mm away from the proton beam. This radiation environment causes novel challenges and requires a continuous monitoring of the VELO in order to assess the radiation damage and its impact on the detector's performance.

This chapter describes studies of the VELO radiation damage. A brief review of the principles of particle detection in silicon is given is Sec. 3.1. The mechanisms of radiation damage in silicon are described in Sec. 3.1.1. This is followed by the description of the VELO subdetector in Sec. 3.2 with a special emphasis on the strip layout and the electronics readout chain. The specialized Charge Collection Efficiency (CCE) scans of the VELO are presented in Sec. 3.3. The analysis of the data obtained from the CCE scans is split into CCE analysis and Cluster Finding Efficiency (CFE) analysis which are presented in sections 3.4 and 3.5 respectively.

3.1 Operational Principles of Silicon Detectors

The most common choice of material for semiconductor detectors is silicon. This is due to the fact that silicon is the second most abundant element on Earth (after oxygen) and a large amount of research has been done on industrial applications of it. Natural silicon usually exists in the form of oxides, such as sand (SiO_2), but very pure silicon crystals can be produced via a couple of different chemical processes [128, 133].

Silicon has the atomic number 14. The 14 electrons are occupying three shells with (2,8,4) arrangement. The four electrons in the outer shell are the valence electrons. A silicon atom can form a covalent bond with four other atoms to complete the outer shell thus forming a crystalline lattice structure.

In a single atom electrons will occupy well defined discrete energy levels. However, in a lattice made of many atoms the electron orbitals will overlap with a small difference in energy. This energy split arises due to the Pauli exclusion principle, which prevents any two electrons from having the same quantum numbers. A crystal is made of many atoms and the energy levels are so close to each other that they can be considered a continuous band. This band is called the energy band.

At the temperature of absolute zero the energy band occupied by the highest energy electrons is called the valence band. However, at higher temperatures the electrons can occupy a higher energy band, called the conduction band. The electrons in the valence band are strongly bound to the atoms and are not mobile, while the electrons in the conduction band can freely move in the lattice.

The energy between the valence and conduction bands is called the band gap energy - E_g. Depending on the band gap energy, different materials can be classified into conductors, semiconductors or insulators, as shown in Fig. 3.1. The valence and conduction bands overlap in conductors (metals) and allow free charge flow, while they are separated by a large band gap inside insulators (e.g. glass with $E_g \approx 9\,\text{eV}$). At a temperature of absolute zero silicon acts as an insulator but the band gap energy is small enough for electrons inside a higher temperature silicon to transfer from the valence to the conduction band. In the latter case, silicon becomes semiconducting. The band gap energy is temperature dependent, as the thermal motion of atoms can change the interatomic spacing in the lattice. For room temperature silicon the band gap energy is 1.12 eV [116]. However, at -10°C (the usual operating temperature of the VELO) the band gap energy is expected to be 1.21 eV [23], while the measured value is closer to 1.16 eV [111].

When silicon is semiconducting, the electron that is able to transfer to the conducting band is referred to as a negative free charge carrier. In the lattice, the absence of an electron behaves much like a free positive charge carrier and is referred to as a "hole". The number of free charge carriers in the semiconductor modifies its conductive properties and depends on temperature. Pure silicon at room temperature has conductivity which is several orders of magnitude smaller than that of conductors. However, the number of free charge carriers can be increased by introducing atoms of other elements into the silicon lattice. This is called doping and can increase the conductivity significantly.

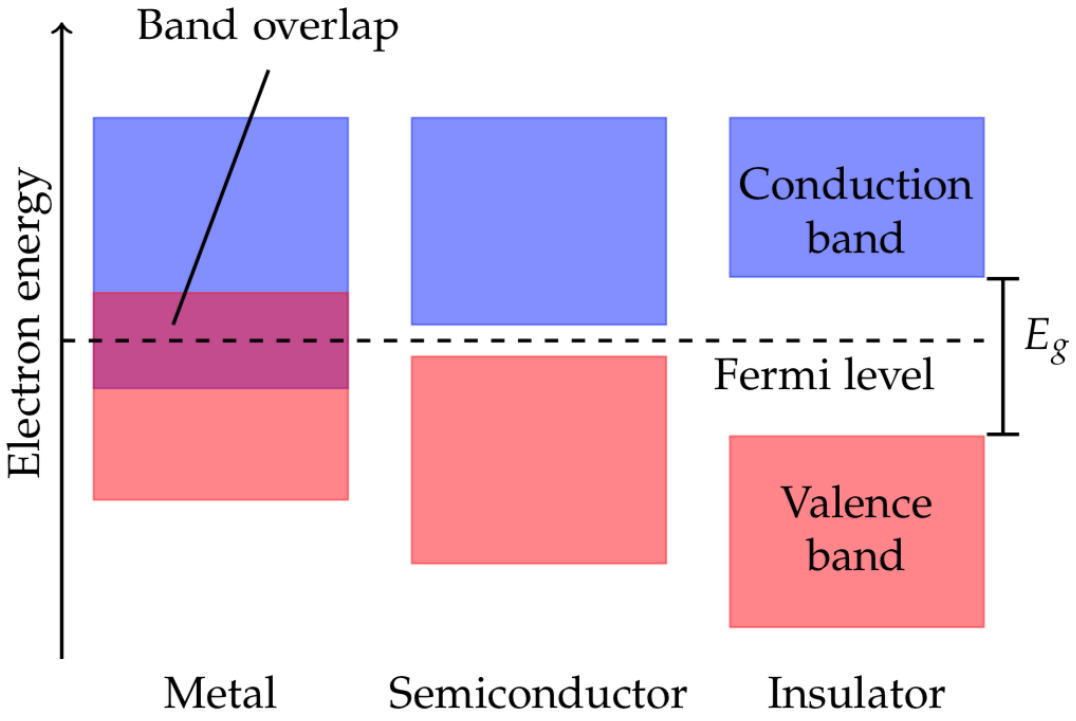


Figure 3.1: An illustration of different classes of materials depending on the band gap energy. The conductors (metals) have overlapping conduction and valence bands, allowing for easy charge transfer. Insulators have a large band gap energy, preventing free charge movement. Semiconductors have a relatively small band gap, allowing for some electrons to transfer to the conduction band due to thermal activity. The Fermi level, denoted by the dashed line, is a hypothetical energy level which would have a 50% probability of being occupied at thermal equilibrium. Reproduced from [110].

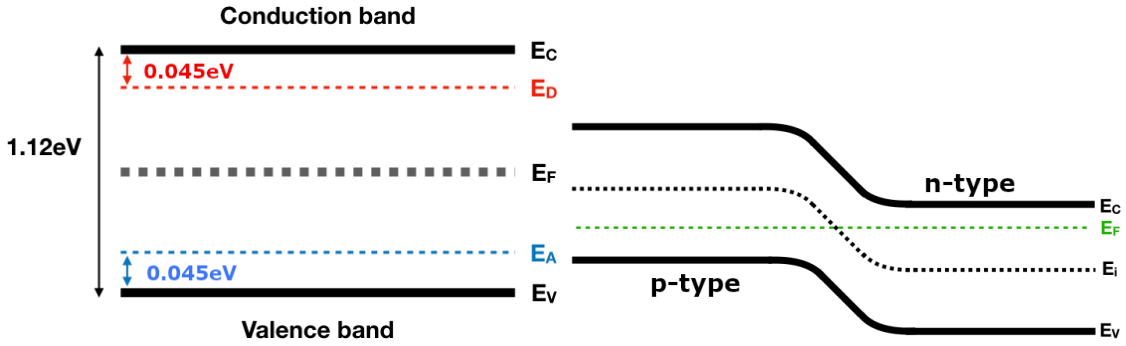


Figure 3.2: (Left) A diagram showing energy bands for doped silicon. The new E_D and E_A bands represent the additional energy levels due to the introduction of donor and acceptor doping respectively. (Right) The diagram showing energy bands of a silicone diode $p-n$ junction. In thermal equilibrium, the Fermi level must be constant across the material. The Fermi level in undoped silicon is represented by the dashed E_i band. Reproduced from [148].

The dopant atoms have either one additional or one less electron in their outer shell. The ones with an additional electron are called "donors" as they bond with four silicon atoms and the extra electron can easily transfer to the conduction band. Donor dopants increase the number of free charge carriers and the conductivity of the semiconductor. A commonly used donor dopant is phosphorus. A silicon doped with donor dopants forms an n-type semiconductor. The dopants with a missing electron in the outer shell are called "acceptors". The introduction of acceptor dopants to the silicon crystal increases the number of holes in the lattice. Silicon doped with an acceptor dopant forms a p-type semiconductor.

Effectively, the dopants change the structure of silicon energy levels. This happens by either introducing additional energy levels near the valence band, when an acceptor dopant is used, or the conduction band, when a donor dopant is used. The modified energy band structure is shown in Fig. 3.2.

A silicon diode is a combination of p- and n-type silicon semiconductors. The interface between these two types of semiconductors is called a $p-n$ junction and is shown in Fig. 3.3. The free charge carriers diffuse across this junction and combine together, forming a depletion region with no free charges. An electric field permeates the depletion region due to the ionized dopant atoms. This creates a potential V_{bi} (for "built-in") which is opposite to the diffusion direction of the holes. Because of this potential difference, the free charge carriers cannot diffuse further. Due to the thermal excitation of the electrons a small reverse current flows through the junction. This current is called a "leakage current" and can be reduced by lowering the temperature of the silicon detector. For example, the average temperature of the VELO in Run 1 was approximately $-7°$C [79]. This current also increases with radiation damage, giving an indirect way to estimate the amount of radiation the detector was exposed to.

The depletion region can be asymmetric and extends further to one of the sides. This happens when one of the sides of the junction is doped more. Such an arrangement is called "abrupt" and is used in silicon detectors, where a smaller highly-doped implant is

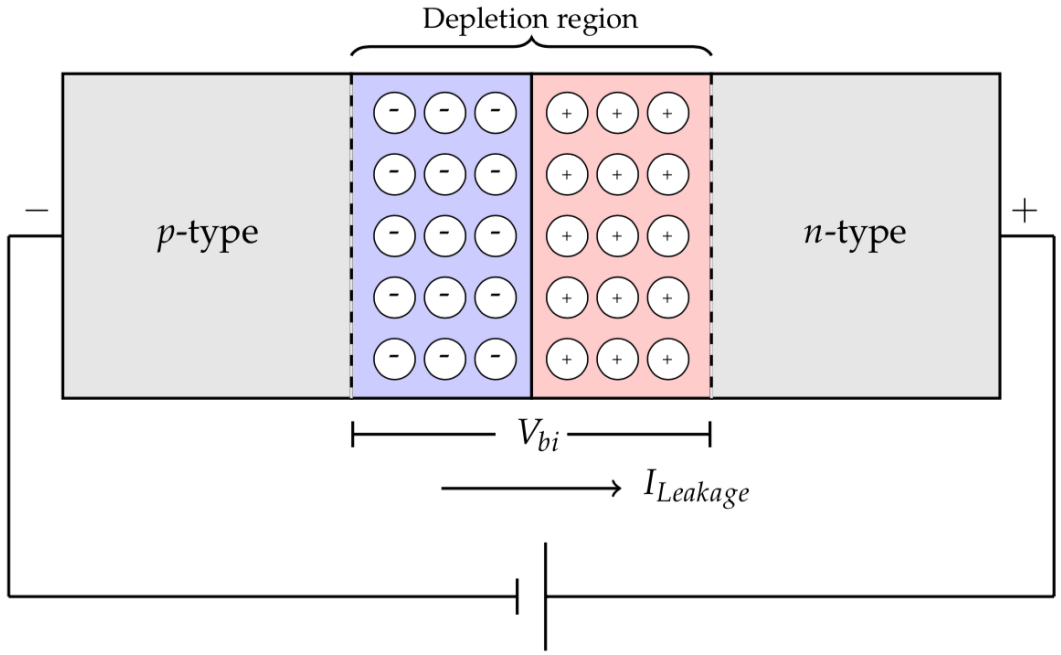


Figure 3.3: An illustration of a semiconductor diode with a symmetric $p-n$ junction. The depletion region is formed by the diffusion of negative and positive charge carriers (electrons and holes) which results in a small potential difference (built in voltage). Also a small reverse leakage current runs across the junction due to the thermal motion of electrons. Reproduced from [110].

joined with a low-doping bulk. The band gap diagram for a $p-n$ junction is shown in Fig. 3.2.

The depletion region can be further extended by applying a reverse bias voltage to the diode, as shown in Fig. 3.3. This pulls the free charges away from the edge of the depletion region. For optimal physics performance, the silicon diode should be fully depleted. This can be achieved by applying a required voltage, called the "depletion voltage" - V_{dep}. The value of the depletion voltage is proportional to the effective doping concentration, N_{eff},

$$V_{dep} + V_{bi} = \frac{q_0}{2\epsilon_r\epsilon_0}|N_{eff}|d^2, \tag{3.1}$$

where q_0 is the electron charge, d is the depth of the depletion region and $\epsilon_r\epsilon_0$ is the permittivity of silicon. The effective doping concentration is defined as the difference between the doping concentrations of donor, N_D, and acceptor, N_A,

$$N_{eff} = |N_D - N_A|. \tag{3.2}$$

For a typical silicon sensor the depletion voltage is a lot higher than the built-in voltage, which thus can be neglected. However, the depletion voltage itself depends on the irradiation of the silicon. Usually, the depletion voltage is increased throughout the detector operation in order to keep the sensor fully depleted over the active detector area. This sets a natural exploitation lifetime of a silicon sensor, as at some point the depletion

voltage is high enough to cause the "breakdown" effect. During the breakdown, the current rapidly increases with the increasing voltage. The silicon essentially becomes conducting, allowing the current to run across. This can damage the detector. For this reason, silicon detectors might have to be operated without being fully depleted towards the end of their exploitation, which will reduce their Charge Collection Efficiency (CCE).

The measurement of CCE is based on the way the signal is obtained in a silicon detector. The electrons and holes drift to the electrodes and induce charge on them. This effect is explained by Shockley–Ramo theorem [143]. Apart from allowing quantification of the performance of the detector, CCE can also be used as an indirect way to monitor sensor radiation damage. When a charged particle intercepts a silicon sensor, it deposits some of its energy by creating electron-hole pairs. To create one electron-hole pair a particle needs to deposit an average of 3.6 eV [6]. The mean energy loss of minimally ionizing particles produced in LHCb, while traversing the silicon crystal lattice of the VELO sensor is about 388 eV / μm. For example, a charged particle with a perpendicular incidence on one layer of the VELO module, which is 300 μm thick, will deposit a mean of 116 keV energy, or 32300 electron-hole pairs [22]. If the sensor is not fully depleted and the particle intercepts the sensor outside the depletion region, the electron-hole pairs will either recombine or drift in a random path. Due to the timing constraints, such electron-hole pairs are not useful for particle detection. This behaviour will cause a reduction in the CCE.

3.1.1 Radiation Damage in Silicon

The VELO sensors are exposed to large amounts of radiation caused by high energy particles. The particles produced in proton-proton collisions of the LHC traverse the silicon of the VELO sensors and create electron-hole pairs which get read out as signal. Apart from that, they also interact with the silicon nuclei and create deformations in the crystalline lattice through recoiling effects. Some examples of such defects are shown in Fig. 3.4.

A nucleus displaced from its lattice position is called an interstitial. The pair of the interstitial and the corresponding vacancy in the lattice is called a Frenkel pair. In silicon, approximately 25 eV of energy has to be transferred to the atom to form such a pair in the lattice [146]. These pairs usually combine due to the thermal motion of the atoms. This happens to approximately 60% of Frenkel pairs and can happen to as many as 95% of them in highly disordered cluster defects [141]. The remaining defects in the lattice interact in many different ways, forming more complex defect structures, such as divacancies or clusters of defects.

The defects in the lattice have an impact on the macroscopic electrical properties of the silicon. This happens because they introduce additional energy levels in the silicon band-gap. One effect is the increase in the leakage current over the silicon diode. This is due to the presence of extra free charge carriers - electrons elevated to the conduction band via the energy levels introduced due to the defects. Leakage current has been shown to increase linearly with fluence, hence they are important measures of the radiation damage.

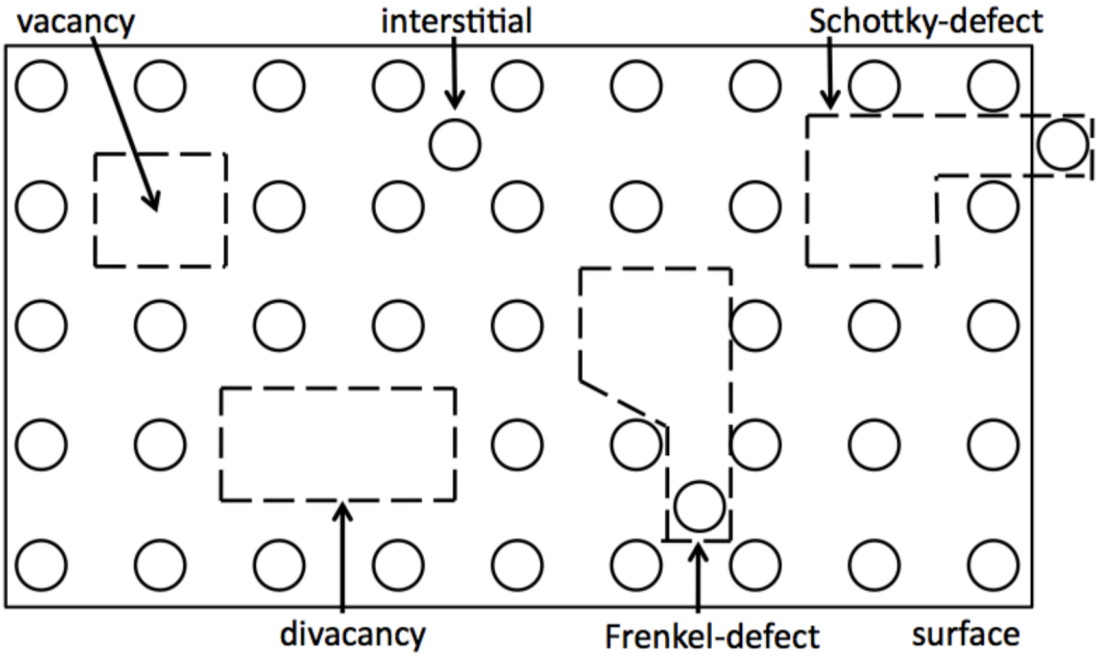


Figure 3.4: Some of the basic defects caused by radiation in silicon. Impurity atoms in the lattice can interact with these defects to form complex point defects. Also basic defects can form more complicated clusters of defects. Reproduced from [10]

After irradiation of the silicon sensor the depletion voltage of the sensor also changes. This effect is caused by the radiation induced effects behaving similarly to doping atoms. For example, extra energy levels introduced near the conduction band make the atoms easy to ionize, thus introducing a positive space charge in the silicon. On the other hand, if the new energy levels are close to the valence band, a negative space charge is introduced similarly to the effect of acceptor doping. These effects will cause the overall depletion voltage to change. The n-type silicon depletion voltage dependence on the fluence is shown in Fig. 3.5. Initially, the depletion voltage is reduced. This happens due to the inactivation of the donor dopants, which is caused by the formation of complex defects and removal of the donor atoms from the lattice. The depletion voltage reaches a value close to zero. At this point an effect called the "type-inversion" occurs and the n-type silicon starts behaving as p-type. After the type-inversion, the depletion voltage starts rising. This is caused by the introduction of the acceptor defects and the increase in the concentration of holes. In the Fig. 3.5 this happens at the fluence of approximately $10^{12} n_{eq}\mathrm{cm}^{-2}$.

Another effect caused by the radiation damage is the drop in CCE. This happens when a defect, located near valence or conduction bands, traps a free charge carrier and delays the charge transfer. If this delay is longer than signal readout time of the detector, it might reduce the charge collection. After irradiation the CCE of a silicon sensor, as well as the signal to noise ratio, is reduced.

Monitoring the electric properties of the silicon sensors, such as the depletion voltage and leakage currents, allows estimation of the sensor radiation damage. This in turn

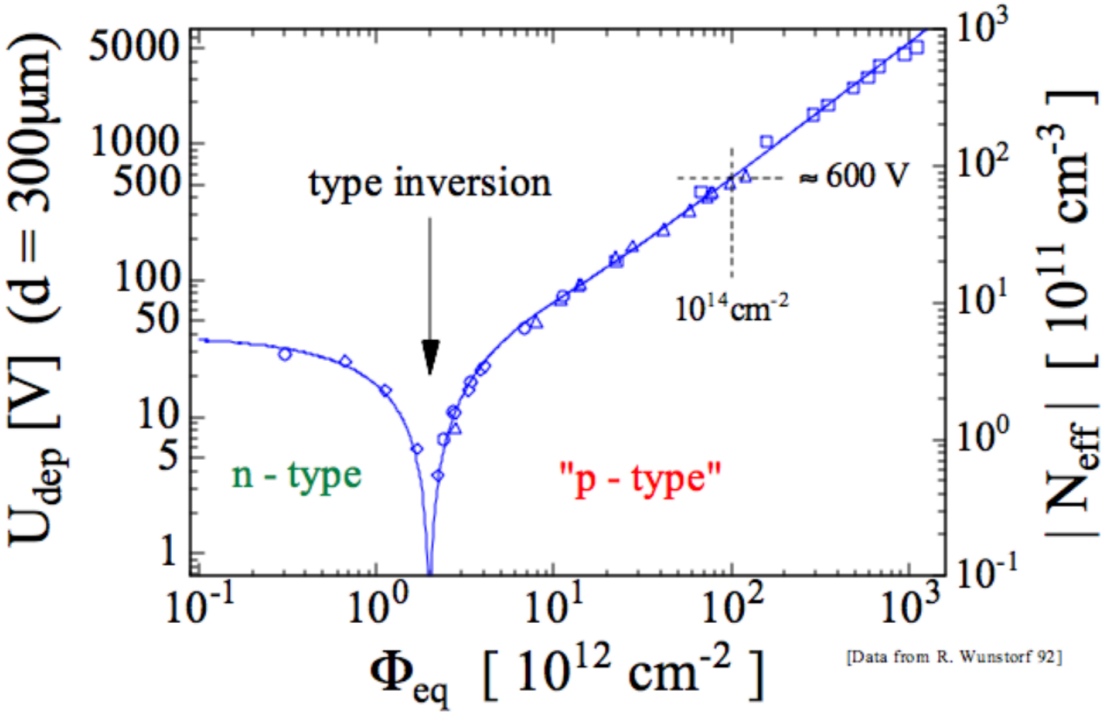


Figure 3.5: Depletion voltage dependence on fluence for *n*-type silicon. Reproduced from [126].

allows prediction of the future performance of the sensor and informs the choice of the optimal operation parameters, especially the bias voltage. The studies of VELO radiation damage are given in Sec. 3.3 and the following sections.

3.1.2 Annealing

The defects in the silicon lattice will move and interact due to the thermal excitations in an effect called "annealing". It will initially cause the electrical properties of the silicon sensor to revert to their previous values before the irradiation. For example, the leakage current will decrease, and the depletion voltage of an *n*-type silicon will increase prior to the type-inversion or decrease after the type-inversion. This is a beneficial effect that can be used to boost the detector's performance. However, after some time "reverse annealing" starts. This happens when the initial recombination of simple defects stops and the defects form complex defect structures. The reverse annealing causes the silicon to become more *p*-type and increases its depletion voltages after the type-inversion, which is undesirable. To stop the reverse annealing, the silicon sensors are usually operated at a low temperature. However, a careful temperature control over the operation time of the detector allows exploitation of the beneficial annealing. This has been done to the VELO detector and is described in Sec. 3.4.2.

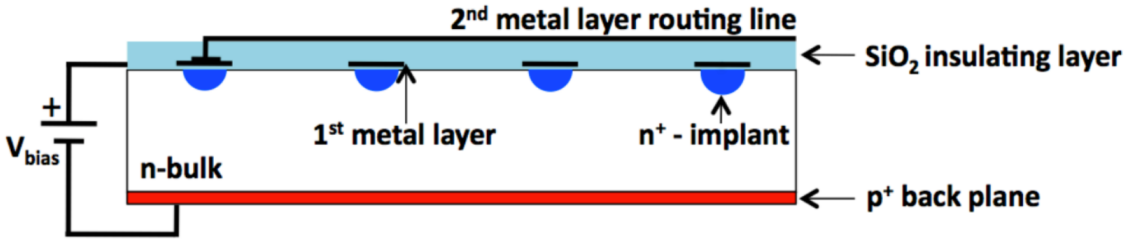


Figure 3.6: A cross-section of an n-type VELO sensor. The aluminium strips (first metal layer) are positioned above the n^+ implant and run perpendicular to the page. The readout routing lines (second metal layer) run horizontally. Reproduced from [80].

3.2 VELO Strips and Electronics

The VELO subdetector has been described in Sec. 2.2.1. This section provides a further overview of the microscopic properties of the VELO, such as the strip and routing line layout. A brief explanation of the VELO readout electronics and data formats is also given.

3.2.1 VELO Strips

The 82 VELO sensors are 300 µm thick silicon semicircles with an n^+-on-n structure. They are referred to as the n-type sensors. There are also two experimental n^+-on-p (p-type) sensors. The studies of n^+-on-p R-type sensor are presented in Sec. 3.6. The n-type sensors consist of n-type bulk with an n^+ implant which is heavily doped. The structure of an n-type VELO sensor is shown in Fig. 3.6. The charge readout is enabled by narrow aluminium strips that run along the n^+ implants, these are referred to as the first metal layer. The strips collect the deposited charge from the depletion region of the silicon. The silicon in the VELO sensors is enriched with oxygen in order to make it more radiation hard. The oxygen concentration in VELO sensors is above 10^{17} cm^{-3}.

There are two types of VELO sensors - R type and Φ type. The R-type sensors have the strips arranged in concentric circles and allow for the measurement of radius of the particle intercept. The Φ sensors have a "fan-like" design and measure the polar angle of the intercept. The types of VELO sensors are further described in Sec. 2.2.1 and shown in Fig. 2.7. For an R-type sensor, the strip-width increases from the innermost to the outermost region in the range of $11 - 38$ µm. The strip pitch is defined as the distance between the centres of the neighbouring strips. It increases as a function of radius from 42 µm to 102 µm. The pitch for a Φ-type sensor differs for the inner and outer regions that are separated at 17 mm radius. In the inner region, the strip pitch varies linearly from 38 µm to 78 µm at the radius of 17 mm. In the outer region, it increases from 39 µm at the region boundary to 97 µm at the outer edge of the sensor. Some VELO parameters are summarised in Table 3.1.

Due to the geometric constraints of the VELO sensors, the charge readout from the R-sensor strips and the strips of the inner region of Φ sensors has to be done via conductive routing lines, referred to as the second metal layer. These are approximately 11 µm thick

Parameter	Value
Silicon thickness	300 µm
Strip pitch	$40 - 120$ µm
Strip width	$11 - 38$ µm
Routing line width	~ 11 µm
Inner silicon edge to beam axis	7 mm
Radial distance of active strips from beam axis	$8.2 - 42$ mm
Sensors position along beam-axis	-300 to 750 mm
Oxygen enhancement	$> 1 \times 10^{17}$ cm^3

Table 3.1: The VELO sensor design parameters. The sensor position along the beam axis is given relative to the beam interaction region. Reproduced from [80].

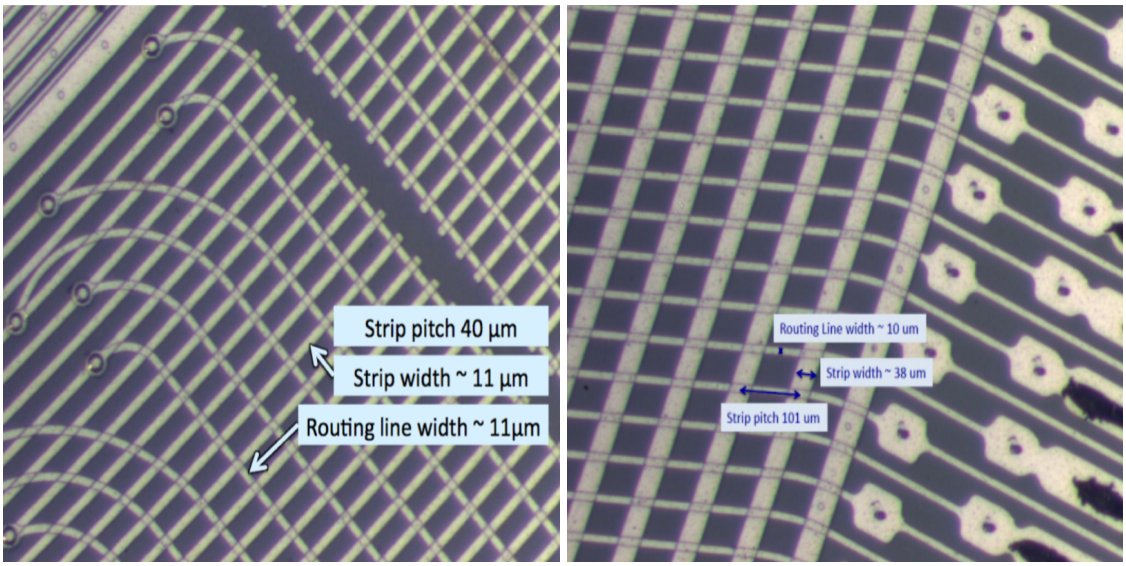


Figure 3.7: Photographs of inner (left) and outer (right) regions of the *R*-type sensor. Each strip is connected to a routing line for charge readout. The charge is transferred to the outer edge of the sensor where it is collected by the readout electronics. Reproduced from [22].

and insulated from the strips by a 3.8 ± 0.3 µm thick layer of SiO_2. For the *R*-type sensors the routing lines run perpendicular to the strips, while for the Φ-type sensors it was possible to run them parallel to the strips. The layout of the routing lines is given in Fig. 2.7. Photographs of the VELO strips are shown in Fig. 3.7.

3.2.2 VELO Electronics

Each of the VELO sensors have 2048 strips for charge readout. The charge collected by these strips and transmitted via the routing lines is read by 16 Beetle chips, located on the front-end electronics. Each Beetle chip is connected to 128 strips via four analogue links to 32 strips. The Beetle chip is an ASIC specifically designed for the VELO and the Inner Tracker of the LHCb detector [122]. These chips store analogue signals from up to 160 events. If the first level of the hardware trigger (L0) sends a response, the analogue data from the Beetle chips is sent to the TELL1 FPGA boards [106]. The TELL1 boards

perform digitization of the data with their ADC converters. One ADC count in a TELL1 corresponds to approximately 440 electrons collected [130]. The output from TELL1 boards is sent to the trigger computing farms (HLT1 and HLT2) for further processing and event reconstruction.

The design of the electronics introduces some unavoidable offset to the signal. For this reason, the the ADC count will be non-zero even if no charge is collected by the sensor. Each channel is offset by a "pedestal", which is adjusted to be in the middle of the ADC range. This allows signal to be collected even if the common mode is negative. The common mode noise is caused by signal fluctuations that are shared by all strips of one 32-strip link. Both, the pedestal and the common mode need to be subtracted from the ADC signal. This is done by the TELL1 boards. Furthermore, noise fluctuations of 2-3 ADC counts occur in each channel. All of these effects are corrected when identifying the signal from the VELO sensors.

3.2.3 VELO DAQ Data Formats

In order to save storage space, the majority of the data read by the VELO sensors is saved in a Zero-Suppressed (ZS) format. This format only saves the ADC values from the channels with counts above a certain threshold, corresponding to the signal. These charge signals are converted into VELO "clusters" that are defined as the data from one or more adjacent strips with ADC count above a threshold. This processing step is performed by the TELL1 boards. However, it is important to also have access to data from the strips that do not pass the ADC threshold. A small fraction of data with the full ADC readout of every strip is saved. This data is referred to as Non-Zero-Suppressed (NZS). This data is especially important for noise, Charge Collection Efficiency and Cluster Finding Efficiency studies. The CCE and CFE studies relating to the VELO radiation damage are described in the following section.

3.3 Charge Collection Efficiency Scans

The electric parameters of the silicon will change over the course of its use due to irradiation. In particular, the bias voltage needed to fully deplete the sensor will change. This has to be taken into account for optimal performance of the detector. For this reason, the depletion voltage, Charge Collection Efficiency (CCE), Cluster Finding Efficiency (CFE) and the number of bad strips need to be monitored throughout the detector's lifetime.

At LHCb the depletion voltages of the VELO sensors have been measured prior to the installation. This has been done by using the "CV method" [145] which involves direct measurement of the sensor capacitance at different bias voltages. A range of different depletion voltages have been obtained as a result of testing the VELO sensors, with a maximum at approximately 65V. However, after the installation it is no longer possible to perform such direct measurements. Instead, the "Effective Depletion Voltage" (EDV) has to be estimated from the proton-proton collision data and the ADC signal response of the VELO sensors during a so called CCE scan.

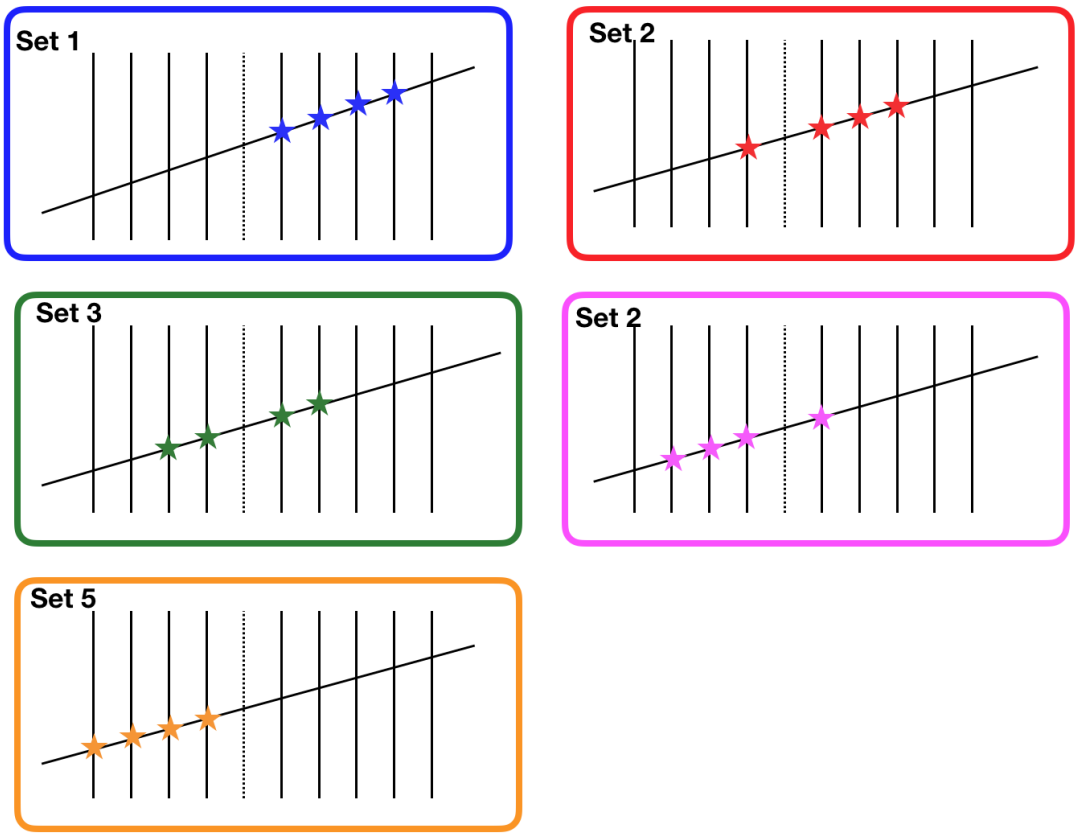


Figure 3.8: An illustration of five different track extrapolation categories. The sensor-under-test is shown as a dashed line. The first category is a set of hits on four down-stream sensors from the sensor-under-test. These hits allow to extrapolate the particle intercept position on the sensor-under-test. Other categories have differing number of neighbouring sensors left and right of the sensor-under-test. Reproduced from [87].

During a CCE scan every fifth VELO sensor is operated at a different bias voltage and excluded from the tracking algorithms. These are called "sensors-under-test". The process is repeated for different bias voltages and different sets of sensors. To measure the charge deposited at a specific location of the sensor-under-test, one of five track extrapolation patterns is used, as shown in Fig. 3.8. Using different patterns allows to test the corner sensors and mitigate possible extrapolation issues in the track reconstruction algorithm.

The CCE scan is a dedicated test that has to be taken during the LHC beam time. For this reason, the CCE scans have usually been performed during technical runs of the LHC to minimise the loss of physics data. There were nine CCE scans performed during the Run 1 and 16 during the Run 2 of the LHC. These are listed in Tables 3.2 and 3.3 respectively. Consequently, LHCb has one of the largest and most consistent data sets available to study radiation damage in silicon sensors.

Due to the time constraints of the CCE scans a limited amount of data can be collected for each set of sensors and each value of the bias voltage. For this reason, only a small number of bias voltages have been tested. Initially this number was seven, while it later got increased to 13. The 13 voltage steps had to be modified as the VELO sensors aged

CCE Scan Date	$\sqrt{s}$(TeV)	Run	Delivered Luminosity [fb^{-1}]
2010	7	77499	0
2011	7		
March		87147, 87153	0.04
July		95936	0.4
September		101359	0.8
October		104107	1.2
2012	8		
April		111267	1.2
July		120006, 120010	1.9
September		129494, 129519	2.7
2013	8	135605	3.4

Table 3.2: The list of CCE scans taken during Run 1. The dates of the scans and the centre-of-mass energy of the LHC are given as well as the approximate total delivered integrated luminosity at the time of the scan. Reproduced from [148].

CCE Scan Date	Run	Delivered Luminosity [fb^{-1}]
2015		
3rd June	153587, 153601, 153602	3.47
5th July	156887	3.47
20th October	166262	3.70
20th November	168083, 168092 (168139, 168257)	3.83
2016		
24th April	173040	3.84
4th August	181247 + 181311	4.81
25th October	185508	5.71
2017		
5th June	192712	5.76
23rd September	199420	6.70
27th October	201210	7.28
21st November	202688	7.63
2018		
17th April	205963	7.73
26th June	210717	8.55
23rd September	215358	9.74
25th September	215591	9.77
2nd December	219089 219090 219091 219092 219093 219095 219106 219109 219113 219117 219121	10.20

Table 3.3: The list of CCE scans taken during Run 2. The dates of the scans and the corresponding LHCb run numbers are given as well as the approximate total delivered integrated luminosity at the time of the scan. All of the scans have been taken with the LHC running at 13 TeV centre-of-mass energy.

Date	Voltage steps
2010	0, 10, 30, 50, 70, 100, 150
2011	0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150
April 2012	0, 10, 20, 30, 40, 50, 60, 70, 80, 100, 120, 150
July 2012	10, 20, 30, 40, 50, 60, 70, 80, 100, 120, 150, 170, 200
2015	20, 40, 50, 60, 70, 80, 100, 120, 150, 170, 200, 225, 250
January 2017	20, 40, 60, 80, 100, 120, 150, 170, 200, 225, 250, 280, 320
May 2018	40, 60, 80, 100, 120, 150, 170, 225, 250, 280, 320, 360, 400

Table 3.4: List of different bias voltage steps used in CCE scans throughout different operation periods of the VELO detector.

and accumulated more radiation damage. Some low voltage steps were skipped to save time, while additional high voltage steps were included. The dates of these changes and the corresponding voltage steps are given in Table 3.4. The last change in the CCE Scan voltages was proposed by the author of this thesis, where the 20V and 200V steps were removed and 360V and 400V steps included.

During the CCE scan, both ZS and NZS data formats are taken (described in Sec. 3.2.3). This is particularly valuable as the NZS data allows a full study to be performed. The data from these scans has been used in CCE studies described in the following section. Apart from estimating the EDV (Sec. 3.4), the CCE Scan data also allows measurement of the CFE and investigation of bad VELO strips. These studies are presented in Sec. 3.5.

3.4 Charge Collection Efficiency

The main purpose of the CCE scans is the estimation of the Effective Depletion Voltage of the VELO sensors. This allows the operating voltage of the sensors to be kept above the required EDV for them to be fully depleted. For example, during Run 1, the VELO sensors were operated at 150V. However, for the n^+-on-n sensors, the EDV will rise after the type-inversion and then a higher bias voltage is required. In order to estimate the EDV, the data from the CCE Scan is used.

A set of tracks intercepting the sensors-under-test is investigated. The clusters, constructed from the charge deposited on the strips nearest to the extrapolated track position, are identified. The search window for formed clusters also include two strips on both sides of the intercepted strip. By measuring the charge deposited in the clusters corresponding to a large number of particle tracks, the ADC distribution for each sensor at each voltage step can be obtained. This distribution is fitted with a Landau function convolved with a Gaussian. The fit is used to obtain the Most Probable Value (MPV) of the ADC distribution, as shown in Fig. 3.9. The Effective Depletion Voltage (EDV) is defined as the 80% of the MPV of the corresponding pedestal-subtracted ADC distribution. This 80% value is chosen as it provides the best agreement between the EDV and the measurements of depletion voltage prior to the VELO installation and irradiation.

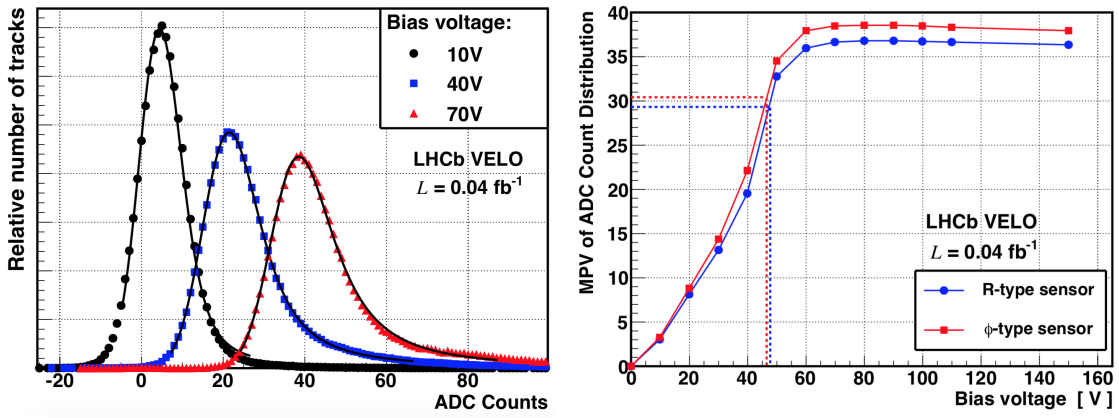


Figure 3.9: (Left) The pedestal subtracted ADC distributions for an R-type sensor at three different bias voltages. (Right) The MPV of the fit to the ADC distribution vs. bias voltage. The dashed lines represent the ADC that is 80% of the maximum value, and the corresponding EDV. Reproduced from [80].

Different regions of the VELO sensors are exposed to different amounts of fluence. For this reason, each sensor is split into five radial partitions. The fluence is approximately proportional to $r^{-1.75}$ [80]. This allows the radial partitions to be chosen in such a way that going from inner partition to outer would amount to approximately half the fluence. The EDV of each sensor has been measured throughout the operation of the LHCb detector. The fluence is determined by scaling an estimated fluence for 1 fb^{-1}. The calculation uses GEANT simulation and the NIEL tables [80]. The scaling factors take into account the difference in particle multiplicity between 7, 8 and 13 TeV centre-of-mass energies and approximately corresponds to the ratio of the logarithm of the energies. A general time trend of the EDV for different VELO sensors and their radial partition is shown in Fig. 3.10.

The EDV evolution in time for the inner region of the VELO sensors is shown in Fig. 3.11. As can be seen from the trend, the EDV of the n^+-on-n VELO sensors decreased for a short duration at the beginning of the operation. This was prior to the type-inversion, which happened at the fluence of approximately $10 - 20 \times 10^{12}$ 1 MeV n_{eq} cm^{-2}. Afterwards, the EDV increased, except for a short period of beneficial annealing which happened between 2011 and 2012, when the insulation of the cooling system was changed.

3.4.1 Hamburg Model Prediction

The overall evolution of the depletion voltage of the silicon can be approximately predicted using the Hamburg Model [126]. This model takes into account the temperature and fluence of the sensor to estimate the depletion voltage evolution in time. The model is based on three different effects. A constant radiation damage term increases the depletion voltage. A beneficial annealing term reduced the voltage initially but is overtaken by a reverse annealing component, which dominates later. The parameters used in the Hamburg model are listed in Table 3.5. More details on the Hamburg model can be found in Ref. [126].

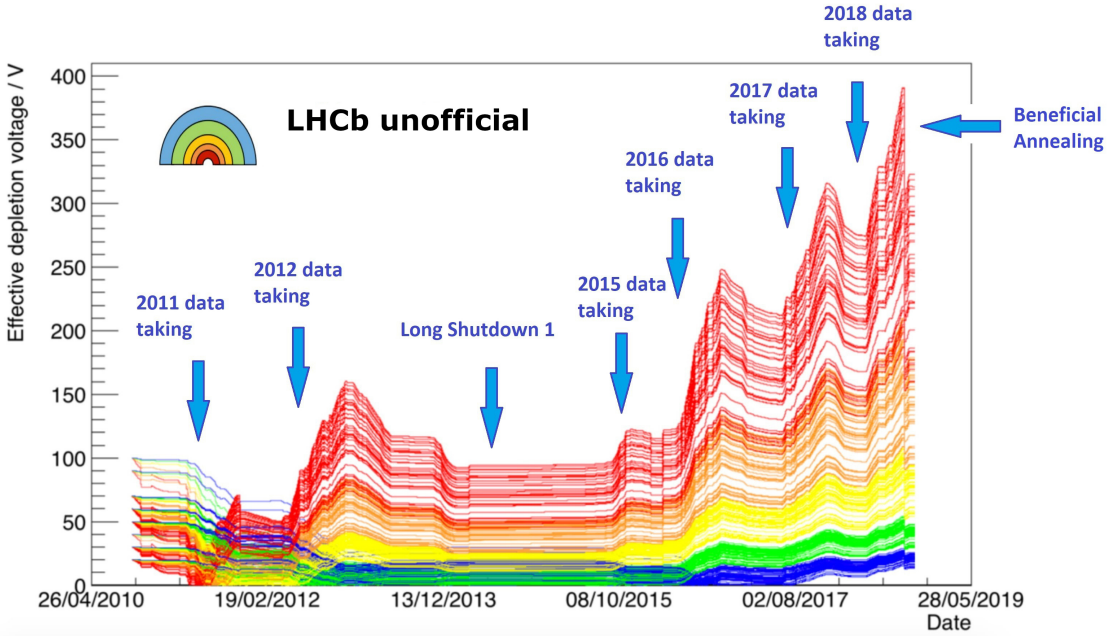


Figure 3.10: EDV evolution throughout Run 1 and Run 2 of the LHCb operation. Each line corresponds to a specific radial partition of a single VELO sensor. Different coloured lines correspond to different radii, as shown in the "rainbow" illustration. Extended based on the work of the author of Ref. [22].

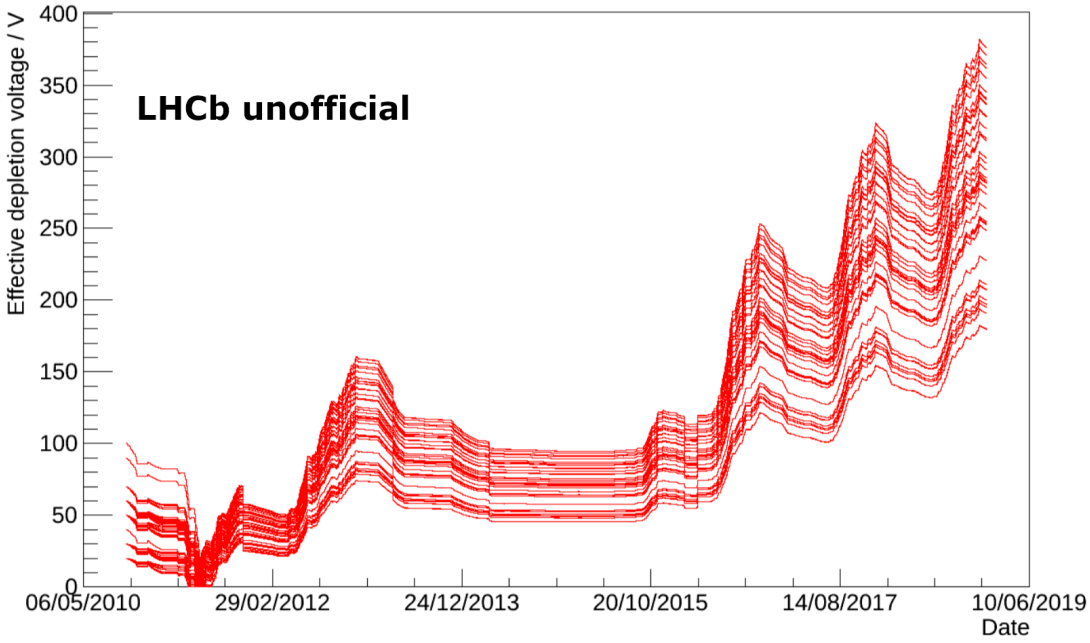


Figure 3.11: EDV evolution throughout Run 1 and Run 2 of the LHCb operation for the inner region of the VELO sensors. Each line corresponds to a specific VELO sensor. Reproduced from [22].

Parameter	Description	Value
k_b	Boltzmann constant	$8.6174 \times 10^{-5}\,\mathrm{eV}/K$
q_e	Elementary charge	$1.6022 \times 10^{-19}\,C$
ϵ_r	Relative permittivity of silicon	$11.9\,F/m$
ϵ_0	Permittivity of free space	$8.85419 \times 10^{-14} F/\,\mathrm{cm}$
G_a	Silicon oxygen enrichment	$1.4 \times 10^{-2}\,\mathrm{cm}^{-1}$
E_A	Activation energy of silicon	1.086 eV
k_{0a}	Frequency factor for beneficial annealing	$3.6 \times 10^{13}\,s^{-1}$
g_Y	Oxygen enrichment	$7.4 \times 10^{-2}\,\mathrm{cm}^{-1}$
Φ_{sat}	Fluence factor of the reverse annealing term	2.1×10^{14}
E_y	Activation energy	1.3252 eV
K_{0y}	Frequency factor for reverse annealing	$4.9 \times 10^{15}\,s^{-1}$
G_c	Introduction of acceptors from proton irradiation	$5.3 \times 10^{-3}\,\mathrm{cm}^{-1}$
c	The exponent factor of the stable damage component	8.56×10^{-14}
d	Sensor thickness	300 µm

Table 3.5: A list of parameters used in the Hamburg model estimation of the EDV dependence on fluence. Based on [126].

The CCE scan data and estimated EDV values have been compared against the Hamburg model prediction. An agreement is seen between the two after the type-inversion, as shown in Fig. 3.12.

The Hamburg model prediction can be used to estimate the EDV of VELO sensors in the future. This allows the operating voltages of the sensors to be changed to ensure full depletion. Towards the end of Run 2 of the LHC, it has been observed that the EDV of many VELO sensors was close to their operation voltages. In this case, the sensors might be underdepleted and the CCE can be reduced. To avoid that, a proposal to raise the operating voltages of the VELO sensors has been put forward. In the end, most of the VELO sensors operated at 350V, while some have been set to as high as 400V.

3.4.2 Beneficial Annealing of the VELO Detector

One of the benefits of the accurate measurement of the VELO radiation damage effects and continuous temperature control was the opportunity to exploit beneficial annealing of the silicon sensors and thus reduce the EDV required to operate them. Due to almost constant cooling of the sensors, the silicon has not reached the reverse annealing phase, which meant that it could still be beneficially annealed by increasing its temperature in a carefully controlled manner. This was desirable, because towards the end of Run 2, the VELO sensors were operating with high bias voltages (320V to 350V, compared to the maximum design voltage of 500V). Beneficially annealing the sensors provided a way to reduce the bias voltages. To achieve this, the VELO was kept at room temperature for 2 days (starting on 18.09.2018), during a technical stop in its final year of operation. During this time the temperature were increased from $-30°$C to $25°$C. Two CCE scans were taken approximately a month before the annealing and a week after it. A reduction in EDV

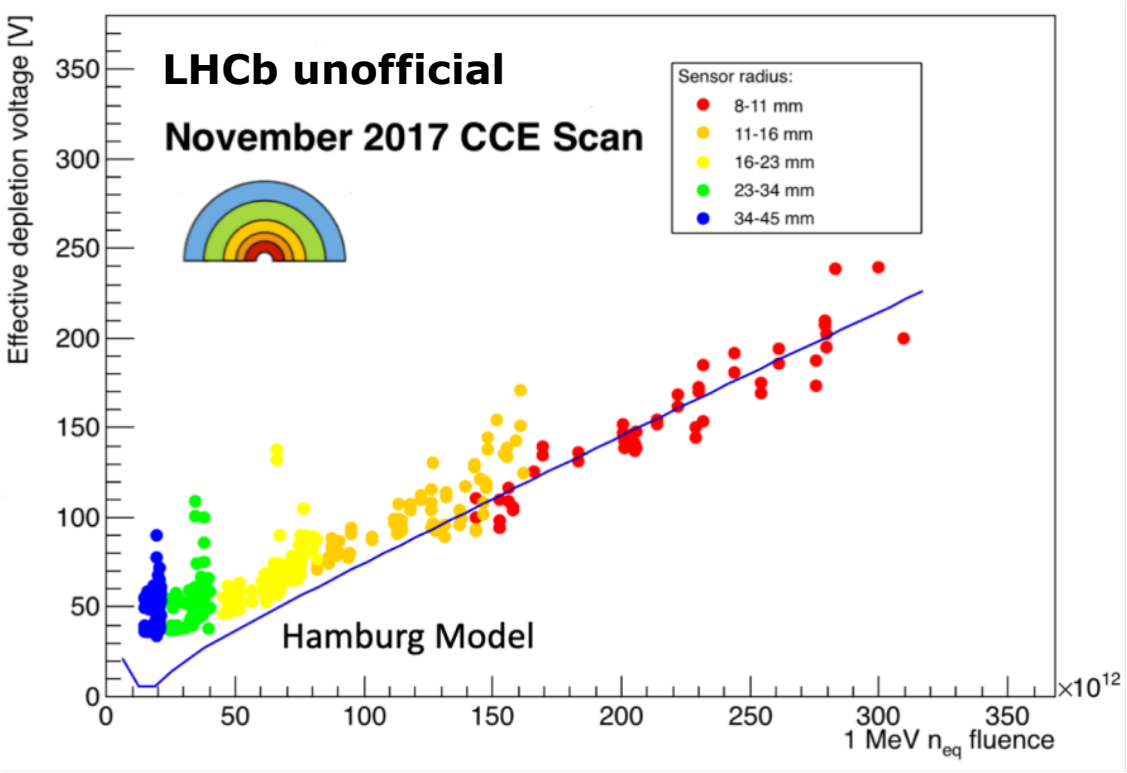


Figure 3.12: The EDV dependence on fluence. Different colours correspond to different radii, as shown in the "rainbow" illustration. The red dots correspond to the inner and most irradiated parts of the sensors. Hamburg model prediction is overlaid and shows an agreement with the observed trend. Based on the work of the author of Ref. [22].

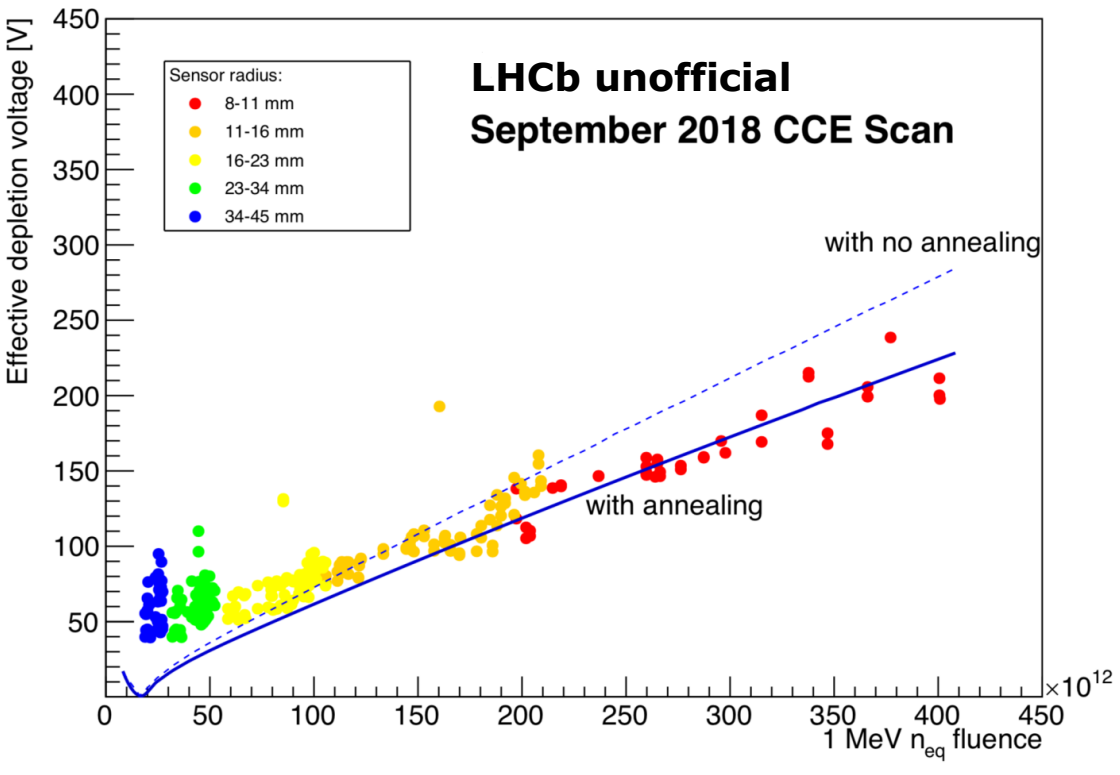


Figure 3.13: The EDV dependence on fluence using the data of a CCE scan taken after the intentional beneficial annealing of the VELO. Different colours correspond to different radii, as shown in the "rainbow" illustration. The red dots correspond to the inner and most irradiated parts of the sensors. Hamburg model prediction with an annealing period is shown as a solid blue line and shows a reduction in EDV when compared to the no annealing hypothesis (dashed line). Based on the work of the author of Ref. [13].

of $70 - 80$V was observed. The comparison of the Hamburg model prediction with and without the intentional beneficial annealing can be seen in Fig. 3.13. The reverse annealing which was supposed to start after the beneficial annealing stage was of no concern, as the sensors would no longer be used after the Run 2. To the knowledge of this thesis author, the intentional beneficial annealing in order to improve the detector performance is unprecedented in an operational experiment in HEP. The proposal for the beneficial annealing of the VELO has been put forward in Ref. [13] and presented by this thesis author in Ref. [140].

3.5 Cluster Finding Efficiency

The VELO track reconstruction algorithm relies on finding the corresponding clusters on the VELO sensors. A cluster is defined as one or more adjacent strips with ADC signals above certain thresholds. In order to efficiently reconstruct the particle tracks, the Cluster Finding Efficiency (CFE) of the VELO sensors has to be as high as possible. Prior to installation and irradiation the mean CFE of the VELO sensors was 99.97% [79].

It is expected that the CFE will be reduced with increasing radiation damage of the sensors or failures in electronic channels. For this reason, the measurements of VELO sensor CFE has been done throughout the operation of the LHCb detector. The method for measuring the CFE is similar to the one for CCE, described in Sec. 3.4. A sensor-under-test is removed from tracking algorithms, the track intercept is found by interpolating from hits from neighbouring sensors and a search for a corresponding VELO cluster is carried out on the intercepted strip and five strips on each side of it. The CFE is defined as the ratio between the number of tracks with the corresponding clusters found and the total number of studied tracks. Usually the CCE scan data has been used to measure the CFE. This allows the CFE to be measured for different voltages. However, the CFE method is not reliant on the NZS data and can be run on generic physics samples. This allows the use of samples with much higher statistics and tracking of the CFE evolution with much greater granularity. In particular, it is possible to measure the CFE of the whole sensor in small 2D regions, which allows regions of inefficiency to be found.

Measuring the CFE allows study of the performance impact of the radiation damage and identification of individual poorly performing sensor strips and electronic channels. The analysis of the bad strips of the VELO is presented in the following section. Also, after irradiation a notable drop in the CFE at large radius of the *R*-type sensors has been observed in Run 1 [80]. This unexpected effect is described in Sec. 3.5.2.

3.5.1 Bad Strips

Each of the VELO sensors has 2048 strips for charge collection from the n^+ implants. A strip can stop producing the ADC signal or produce a poor signal. If a strip is not producing a signal even when intercepted by a particle, it is referred to as a "dead strip". If a strip is producing a signal which is significantly higher than the average of neighbouring strips, it is referred to as a "noisy strip". Both dead and noisy strips will reduce the CFE of a sensor. Prior to irradiation the bad strips have been identified with a couple of different methods. A visual inspection was carried out before the installation of the VELO sensors and physically damaged strips identified. After the installation, strips with no signal or signal four times higher than the average of 30 neighbouring strips have been identified as bad. With the exclusion of bad strips, the estimation of the CFE shows a sizeable improvement, as can be seen in Fig. 3.14. The removal of bad strips from the analysis is useful as it allows the study of additional effects that can be masked by the effect of individual bad strips.

However, it has been noted that the bad strip finding algorithm might not be robust in some cases. For this reason, new algorithms had to be created. A simple algorithm is based on the normalized frequency of strip hits. With a large dataset of tracks each strip should be able to have a large number of signal responses, corresponding to hits by intercepting particles. However, when the hit map for a VELO sensor is plotted, areas with low numbers of hits or without hits can be seen, as shown in Fig. 3.15.

The strip hit frequency can be used to find dead and noisy strips. To achieve this, a large data set of tracks is used and normalized strip-hit frequencies are calculated for

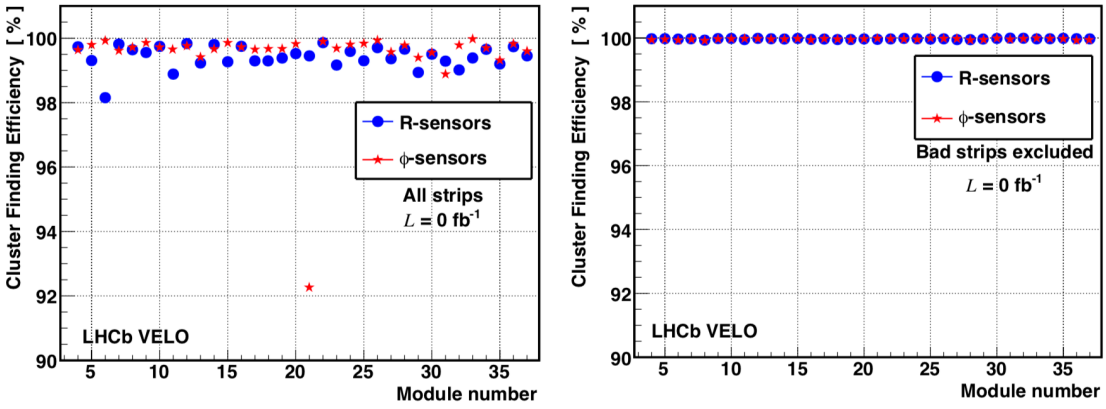


Figure 3.14: The CFE of VELO sensors prior to irradiation. Figure on the left shows the CFE for different sensors with bad strips, while figure on the right shows the CFE after bad strip removal. The Φ sensor on VELO module 16 has a broken Beetle chip and shows a low CFE before the malfunctioning strips are removed from the CFE measurement. Reproduced from [148].

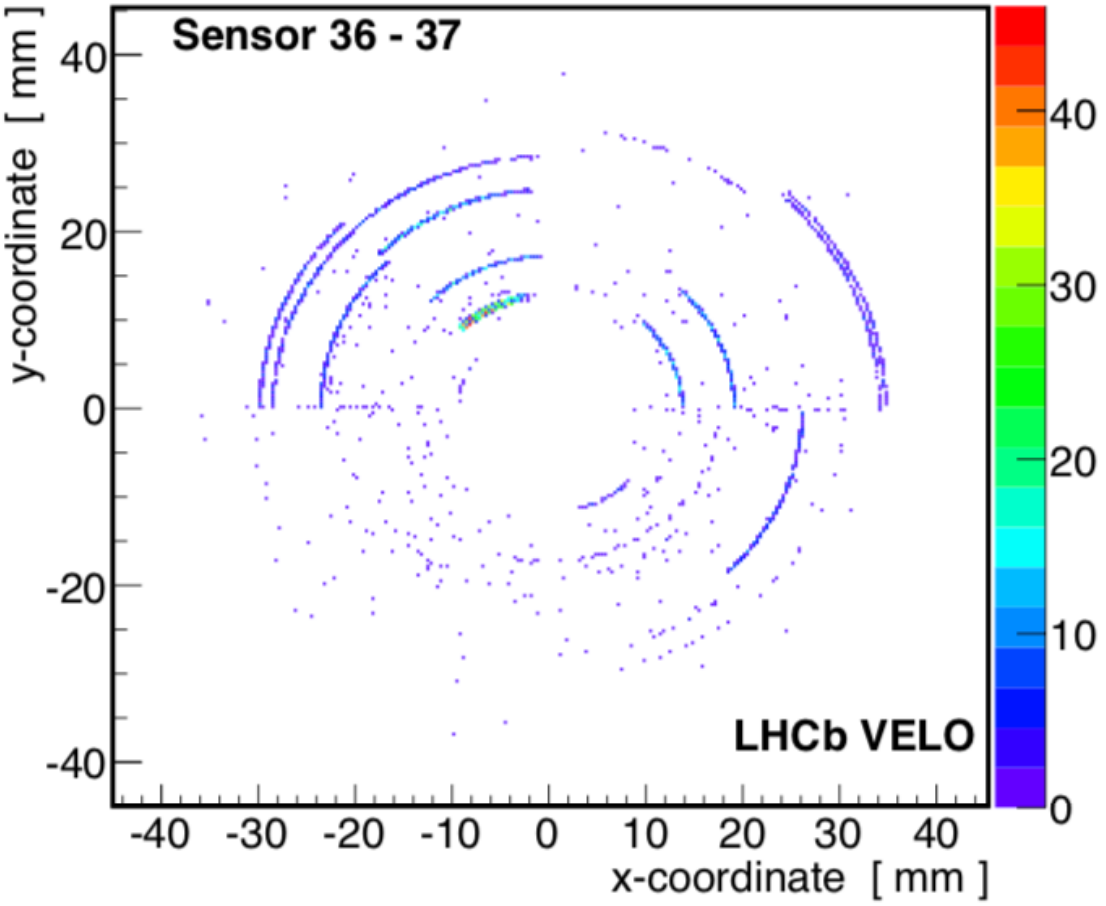


Figure 3.15: A 2D map of two VELO sensors (R36 and R37), showing areas where no clusters are formed. Bad strips can be clearly identified as line segments in the map with high inefficiency. The strips terminate at the interface between four *R* zones. The coloured bar shows the number of tracks in the region without corresponding clusters. Reproduced from [148].

each strip. The frequency is plotted against the strip ID (unique number identifying each strip) and fitted with line segments. The fit is motivated by the fact that in each of the four zones of R sensors strips are situated in semicircular segments radially outwards and the particle flux is linearly decreasing with radius. The first and fifth segments are inclined upwards. This is because of a subtlety in the geometry of VELO R-type sensors. Each R-type sensor has diagonal cut-outs in the innermost region. The strips in these cut-outs vary in length and are expected to have lower hit frequency. The hit frequency is compared to the neighbouring strip hit frequencies by calculating the residual, which is defined as the difference between the frequency value and the value of the fitted line segment at that point. The residuals are fitted with a Gaussian distribution and any outliers (residuals with values at least 3σ away from the Gaussian mean) are marked as bad. The procedure applied on two different VELO sensors can be seen in Fig. 3.16.

The "frequency method" allows the identification of the majority of bad strips. The application of this method on the last CCE Scan taken in Run 2 has yielded 1.35% of bad strips for R-sensors. However, due to the specifics of VELO clustering algorithms, some dead or noisy strips can appear as functional. This happens during the Common-mode suppression stage, where the common electrical noise of 32-strip groups is subtracted from the individual strip ADC signal. Hence a dead strip can appear to have a non-zero signal, or a noisy strip can have a reduced noise and not pass the threshold for detection. To identify these strips a more fundamental approach is needed.

A complementary approach of bad strip removal was created, referred to as the "RMS method". The new approach is based on the underlying shape of the pedestal-subtracted ADC distribution of each strip. This decouples the bad strip identification from the condition of neighbouring strips and allows detection of dead or noisy strips that would otherwise be missed by the frequency method. An example of the pedestal-subtracted ADC distribution projections for all strips of a single VELO sensor is shown in Fig. 3.17. As can be seen, the 32-strip groups fluctuate a lot, while the fluctuations are smaller within the groups. This corresponds to the Beetle chip connections and the common mode.

The RMS method first requires a Gaussian convolved with a Landau to be fit to the ADC distributions of each strip and fit parameters to be checked against a set of indicators. The three indicators of a bad strip are that it either has the $RMS < 2.25$, $Kurtosis < 50$ or $Skewness < 0.2$. These indicators allow to check whether the ADC distribution of the strip has an expected shape, i.e. is not too sharp, has expected tails and is not too wide. The ADC distribution of a healthy VELO strip is expected to be highly asymmetric and have a long tail towards the large ADC values, hence the skewness of such a strip should be high. An example of bad strips identified by the RMS method is shown in Fig. 3.18. As can be seen, the strips are clearly showing different ADC behaviour from their neighbouring ones.

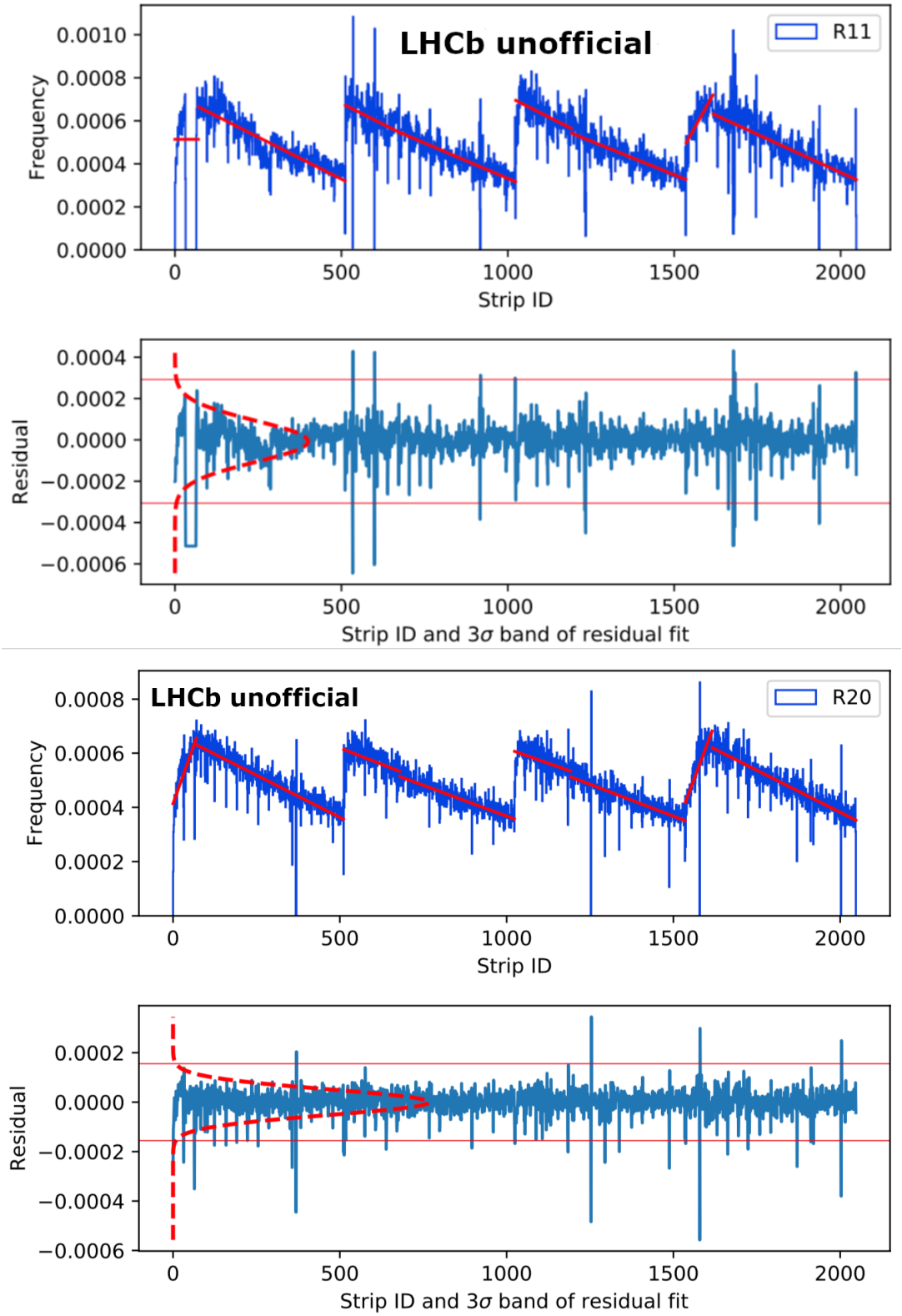


Figure 3.16: The frequency method applied on two VELO sensors - R11 (top) and R20 (bottom). For sensor R11, a large region with strips not producing signal can be seen for low strip IDs, this corresponds to a malfunctioning Beetle chip. The sensor R20 shows healthy set of strips with a small number of dead and noisy strips. The 2018 September CCE scan data is used for the study.

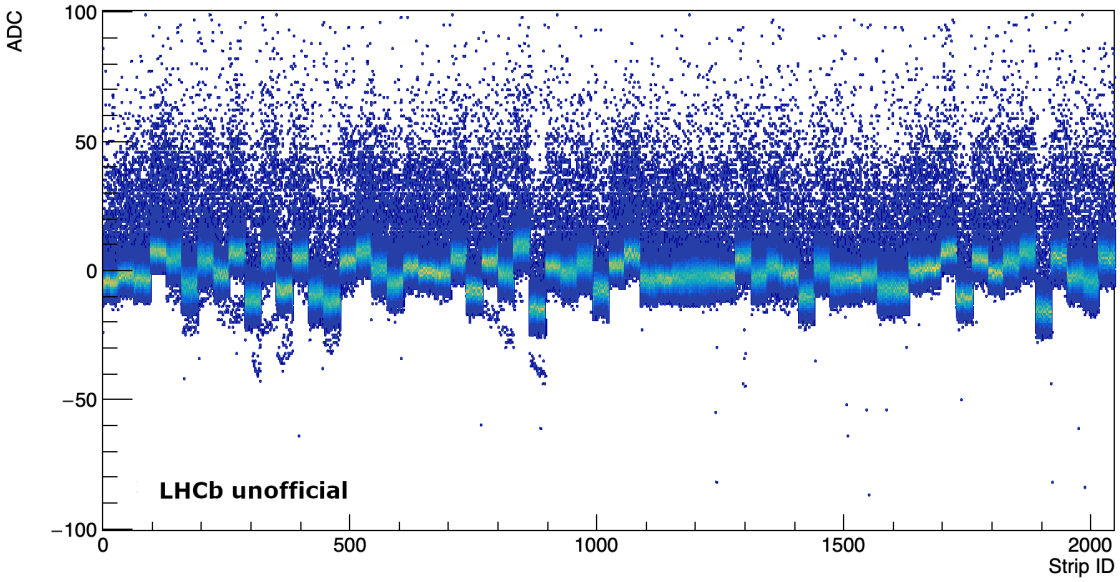


Figure 3.17: The projected ADC distributions of all strips for VELO sensor R35. The fluctuations observed correspond to groups of 32 strips connected to the same Beetle chip output and hence having the same common mode noise. Healthy strips have a large tail towards the increasing ADC, as can be seen in the top-half of the plot. The 2018 September CCE scan data is used for the study.

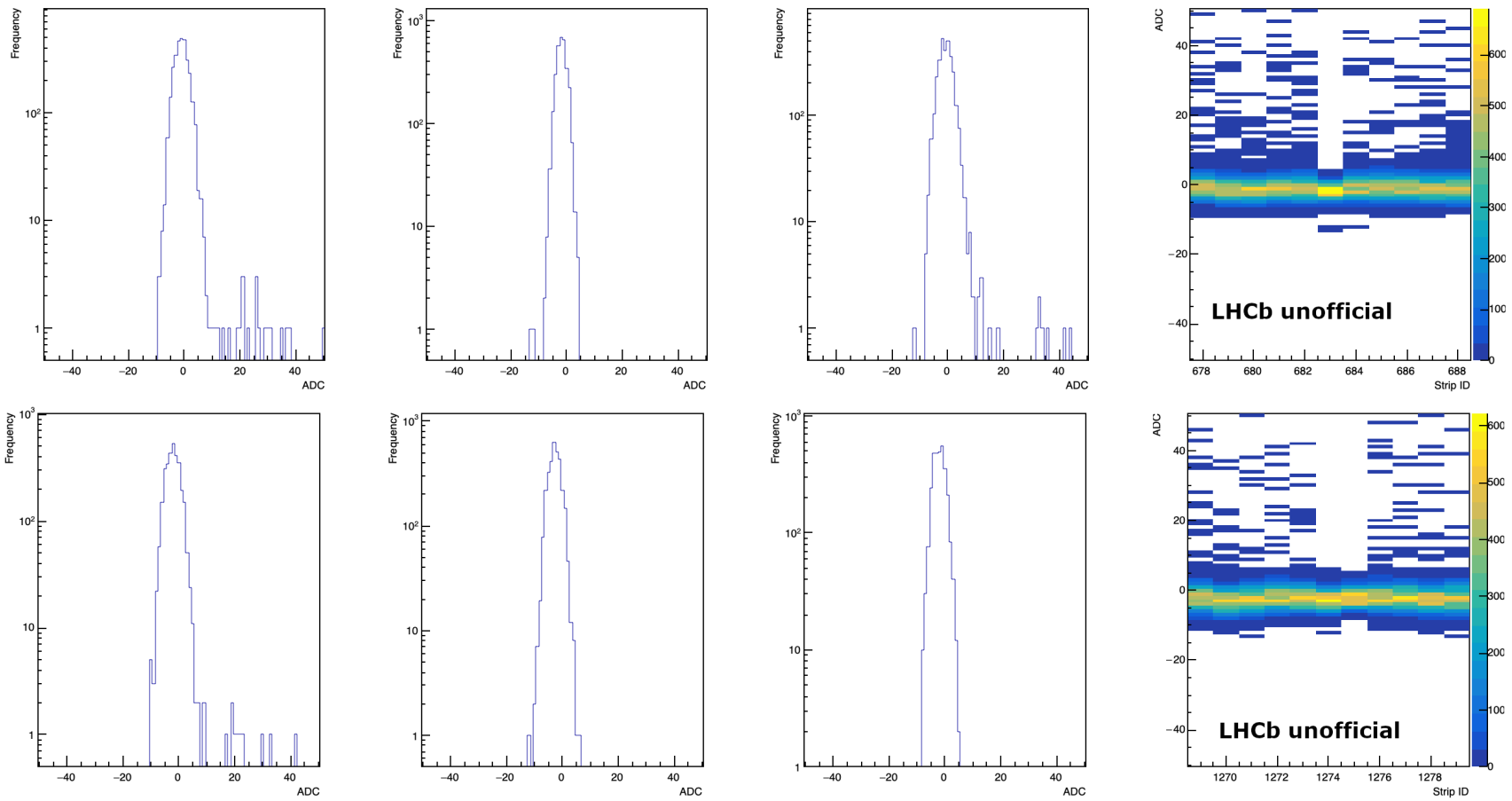


Figure 3.18: The RMS method applied on the VELO sensor R35. Plots show the ADC distribution of the strip identified as bad in the center and the distributions of the neighbouring strips next to it. The last panel of the plots shows the projected ADC distribution for the identified and neighbouring strips. The 2018 September CCE scan data is used for the study. Based on the work of [1].

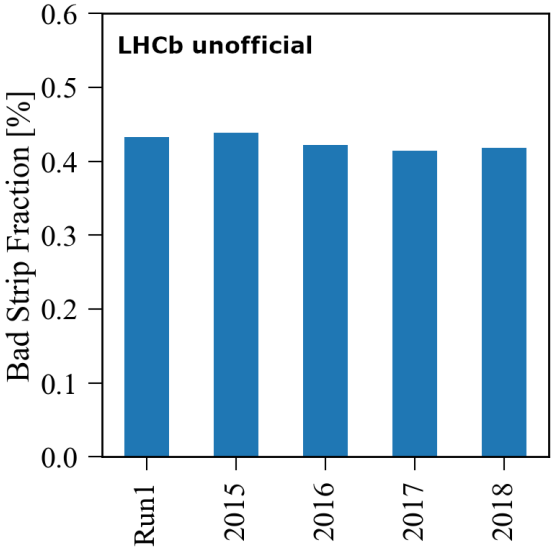


Figure 3.19: The fraction of bad strips found by the RMS method over different years of VELO operation. Based on the work of [1].

It is useful to monitor the number of bad strips identified as a function of time as a measure of performance. The fraction of bad strips found by the RMS method for different years of LHCb operation is shown in Fig. 3.19. No significant degradation of the VELO strips has been observed. It can be seen that the fraction of bad strips was relatively small and almost constant throughout the operation of the VELO. The implementation of this method has been done in collaboration with the author of Ref. [1].

The two methods of bad strip removal are combined and a complete bad strip list for the VELO *R*-sensors is produced. Using the data from the last CCE scan of Run 2, the total fraction of strips identified as bad in the *R*-type sensors is approximately 2%. In order to further mitigate the effects not corresponding to the actual radiation damage, it was decided to include the five innermost and outermost strips from the CFE calculation. This prevents a fake reduction in CFE, caused by the particles being scattered out of the sensor acceptance but the extrapolated track hit position incorrectly intercepting a VELO strip. With the removal of edge strips, the total fraction of excluded strips reaches 2.2%.

The comparison of CFE for each VELO sensor prior to and after the bad strip removal is shown in Fig. 3.20. As can be seen, the scatter of the points is significantly reduced after the application of the bad strip algorithm. This allows sensitivity to additional effects as discussed in the following section. The remaining trends in points are likely due to the interpolation and extrapolation methods of the tracks and to the track angles covered in particular sensors. This is discussed in the following sections.

The bad strip algorithms have been applied on data sets corresponding to the CCE scans taken at different years of VELO operation. This allows checking of whether the bad strip list is evolving in time. This could happen if for a specific CCE scan a group of strips was noisy, while it performed well afterwards. An example showing the bad strip match between different years is shown in Fig. 3.21.

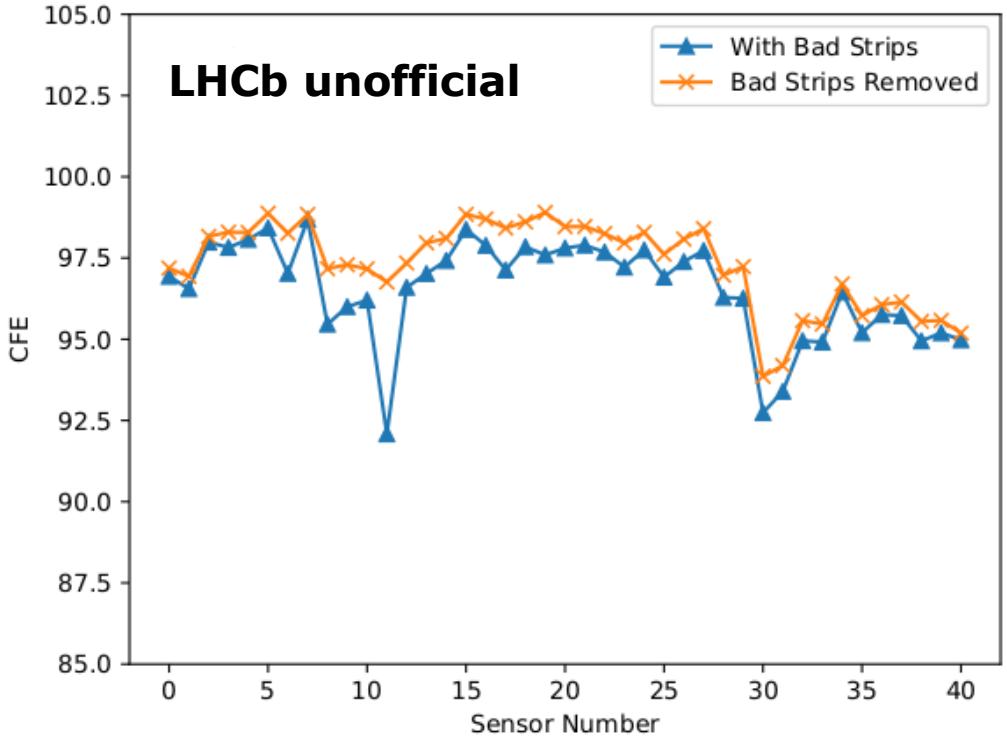


Figure 3.20: The updated measurement for the CFE of different VELO *R*-sensors. The measurement uses the last CCE scan data taken at the end of Run 2 (2018 September). Most sensors operated at 320V, while a small subset operated at 350V bias voltages in order to be fully depleted. The CFE prior to the bad strip removal is shown as a blue line. A clear improvement in the CFE is seen after removing the bad strips, as shown by the orange line.

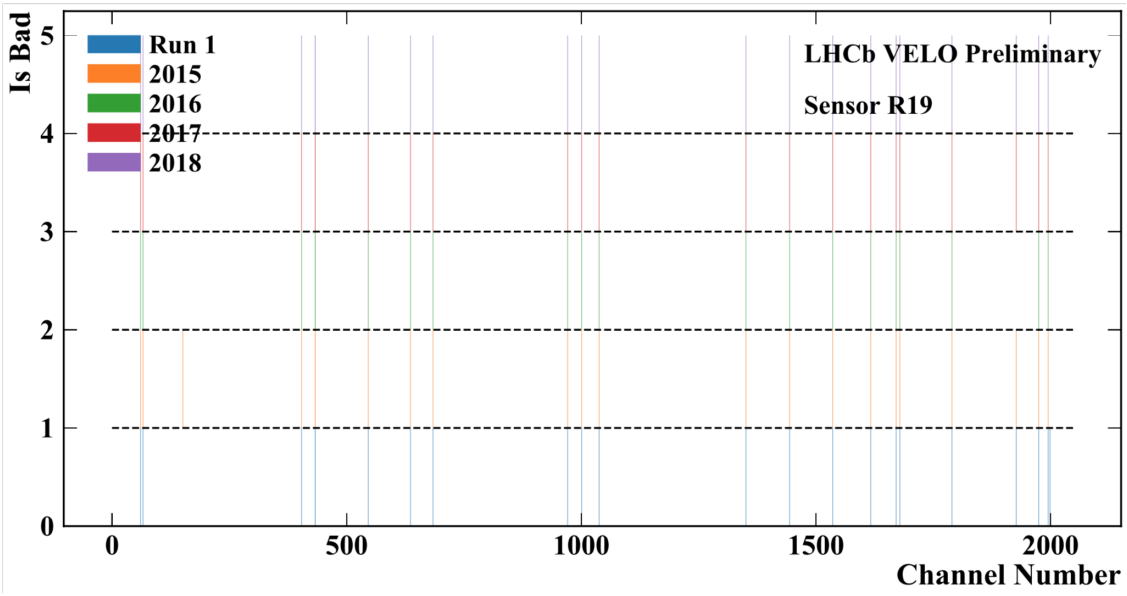


Figure 3.21: A comparison of five different bad strip lists for a single R19 sensor. The lists correspond to the CCE scans taken at different years. Each row corresponds to a specific year. Lines show the strips identified as bad. Most strips are identified as bad throughout the different years of data taking. Some strips have been marked as bad only in one year, while others did not meet the criteria for being bad for a specific year. This could be caused by the calibration of VELO and different running conditions. Based on the work of [1].

3.5.2 Second Metal Layer

The CFE calculation does not rely on the limited amounts of existing NZS data and can use any physics data (ZS) from the LHCb. This allows measurement of the CFE of the VELO with fine granularity (e.g. creation of 2D efficiency maps of the sensors). After initial irradiation, an unexpected drop in CFE has been observed at large radius of the *R*-sensors, as shown in Fig. 3.22. However, the reduction of the CFE does not seem to continue proportionally with increasing irradiation and is believed to have stabilized. Further investigation of this effect found a direct correlation between the regions with low efficiency and the layout of the routing lines on an *R*-type VELO sensor, as shown in Fig. 3.23. For this reason the effect is called the "Second Metal Layer" effect.

The Second Metal Layer effect can be further investigated by looking at the CFE dependence on the track distance to the nearest strip and routing line. As can be seen in Fig. 3.24, the effect is strongest for tracks that hit the sensor close to a routing line and furthest away from a strip. The results shown in Fig. 3.24 are obtained from the Run 2 (2018) CCE Scan data. For comparison, the same measurement is done using Run 1 (2013) data and shown in Fig. 3.25. As can be seen, the trends for Run 1 and Run 2 studies are very similar. Similar effects have been observed in the ATLAS and CMS experiments [120, 137].

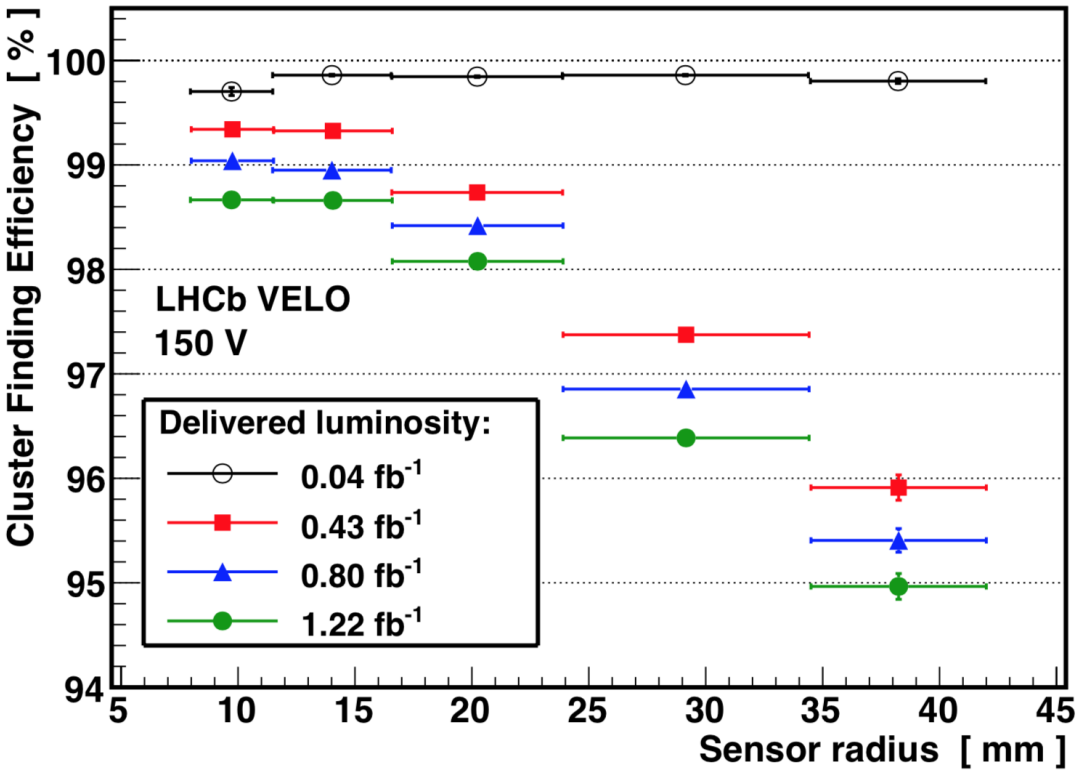


Figure 3.22: The CFE of an *R*-type VELO sensor as a function of radius. A sudden drop in efficiency is visible at high radius with a small amount of irradiation. Afterwards, the efficiency drop is slow. This is caused by the Second Metal Layer effect. Reproduced from [148].

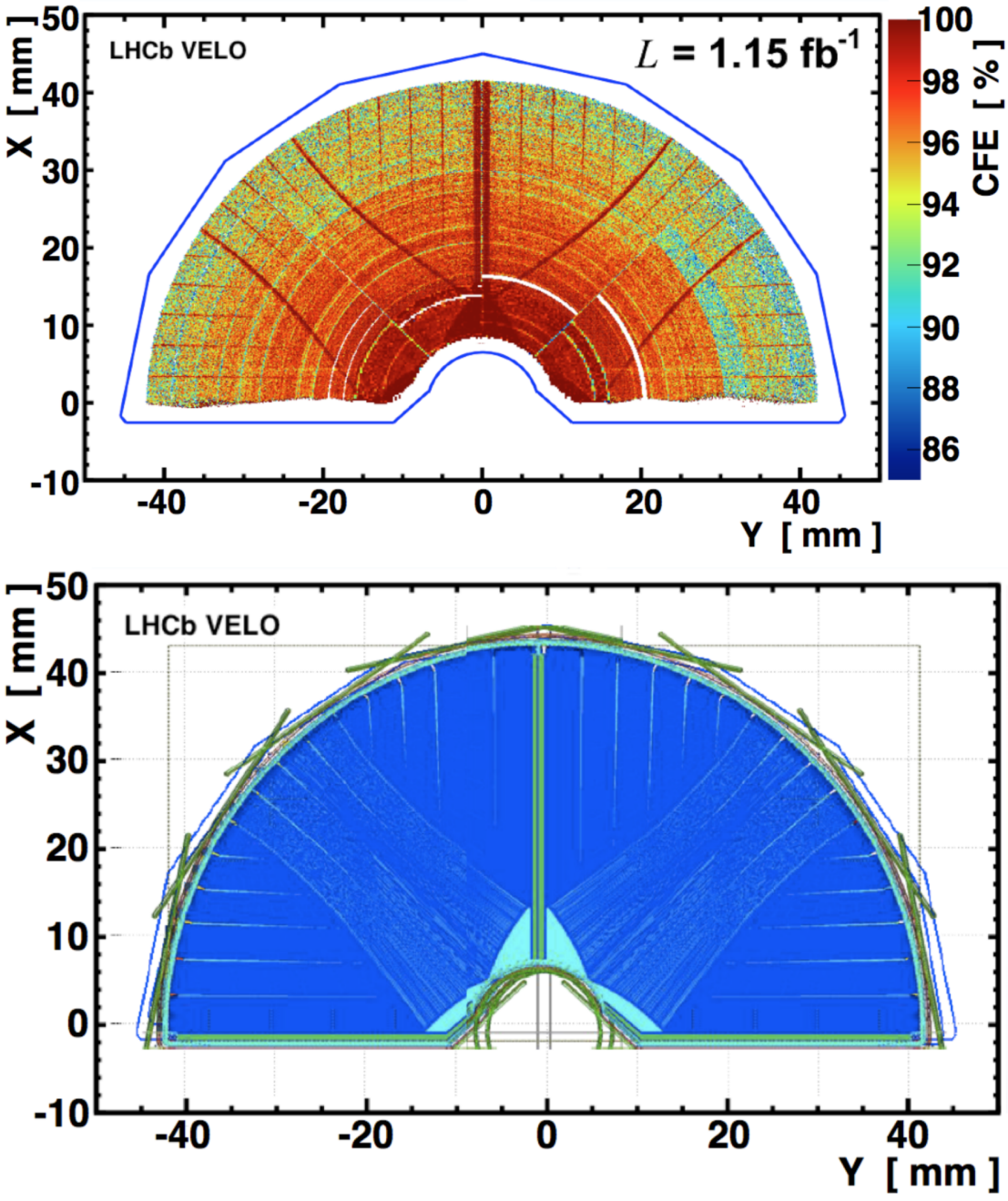


Figure 3.23: (Top) The 2D CFE map of a single *R*-type VELO sensor. (Bottom) The layout of the routing lines (Second Metal Layer). The regions with low efficiency line up with the routing line layout. Reproduced from [148].

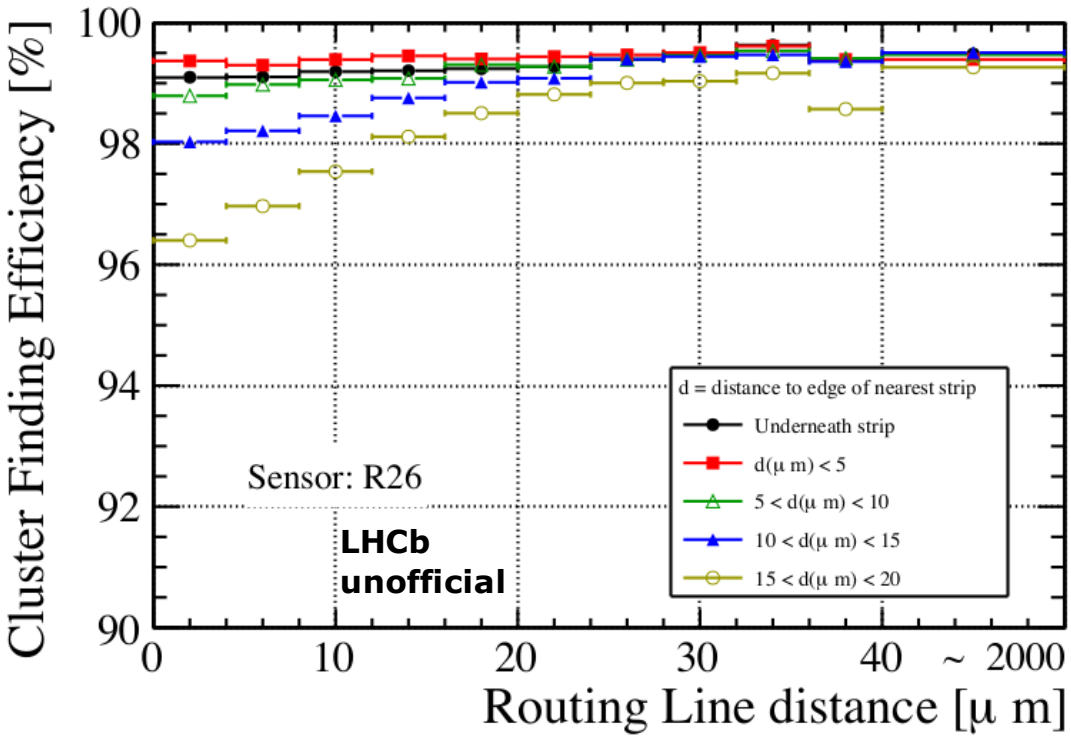


Figure 3.24: The CFE as a function of the track intercept distance to the routing line. Five different categories of tracks are shown, for different distance to the edge of the nearest strip. The lowest efficiency is observed for the tracks close to the routing line and far away from the nearest strip. This supports the charge induction on the routing lines hypothesis. The 2018 September CCE scan data is used for the study. The results are shown for a single sensor (R26) with the nominal operating voltage of 320V.

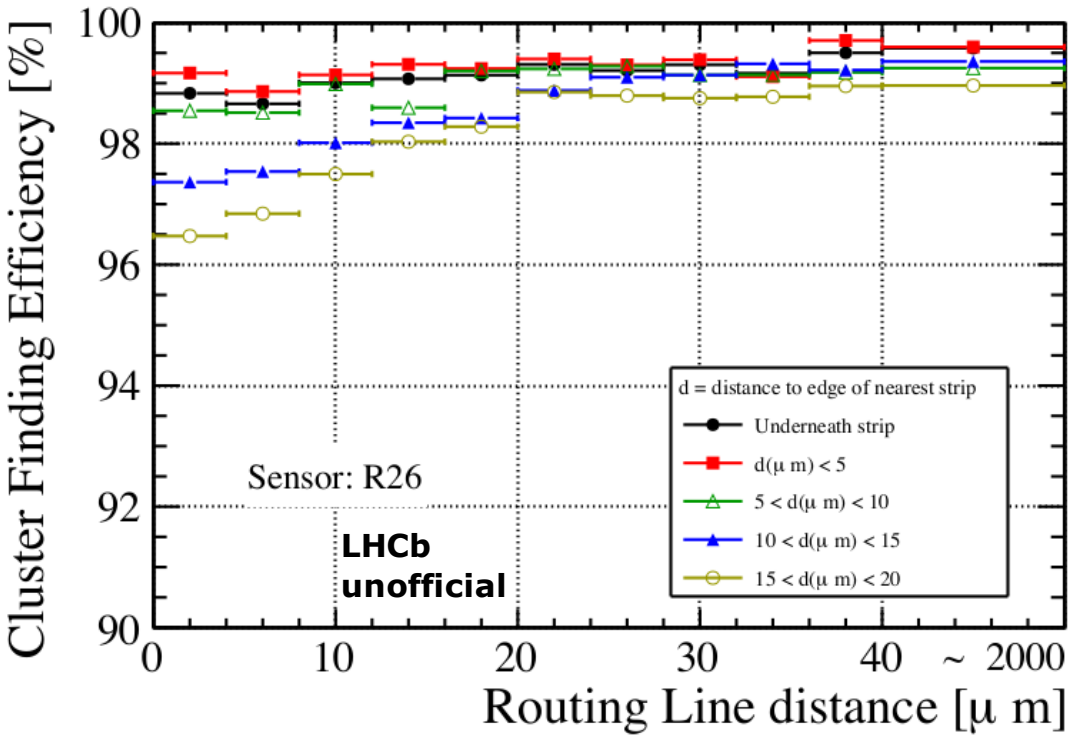


Figure 3.25: The CFE as a function of the track intercept distance to the routing line. Five different categories of tracks are shown, for different distance to the edge of the nearest strip. The lowest efficiency is observed for the tracks close to the routing line and far away from the nearest strip. This supports the charge induction on the routing lines hypothesis. The Run 1 (2013) CCE scan data is used for the study. The results are shown for a single sensor (R26) with the nominal operating voltage of 150V.

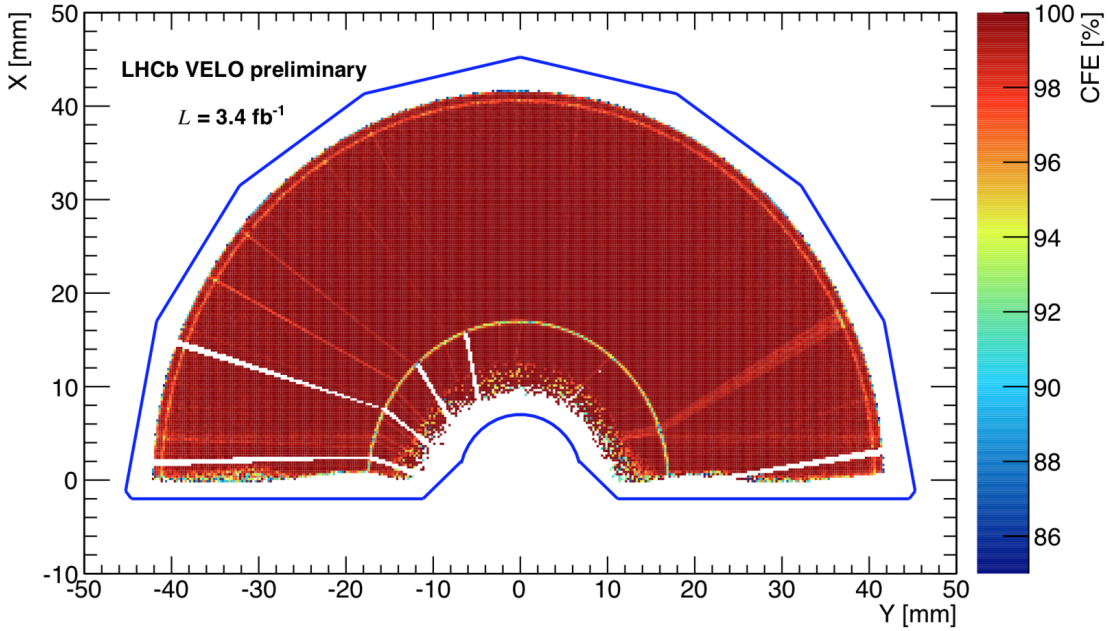


Figure 3.26: A 2D efficiency map of a Φ-sensor. The sensor exhibits no efficiency drop at large radius, characteristic of the second metal layer effect. This is due to a different routing line layout, allowing the outer section strips to run parallel to the routing lines. A line with low efficiency at approximately 17 mm is the interface between inner and outer regions of the sensor. Reproduced from [110].

The findings presented above clearly show that charge pick-up is occurring to the second-metal layer routing lines, thus decreasing the charge collected on the strips (first metal layer). A device simulation has been performed to attempt to replicate the effect [98]. The second metal layer of the sensor are separated from the bulk-silicon by a SiO_2 layer of around 4 µm. A p-isolation layer is used beneath the SiO_2 to separate the n^+ strips; this layer traps holes. A negatively charged electron inversion layer consequently builds up beneath this, acting as a conductive layer. This layer assists in the shielding of the routing lines. However, after irradiation surface damage effects can occur leading to fewer electrons in the layer and reduced carrier mobility in the layer, both effects potentially causing less shielding. Consequently charge may be picked up on the metal layer of the routing lines rather than the metal layer of the strips. The simulation was performed using the software Synopsys TCAD [153] to simulate the electric field and induced charge in the device. The simulation is based on realistic device parameters and uses a surface damage model described in Ref. [132]. This model reproduces broadly the features that are observed here.

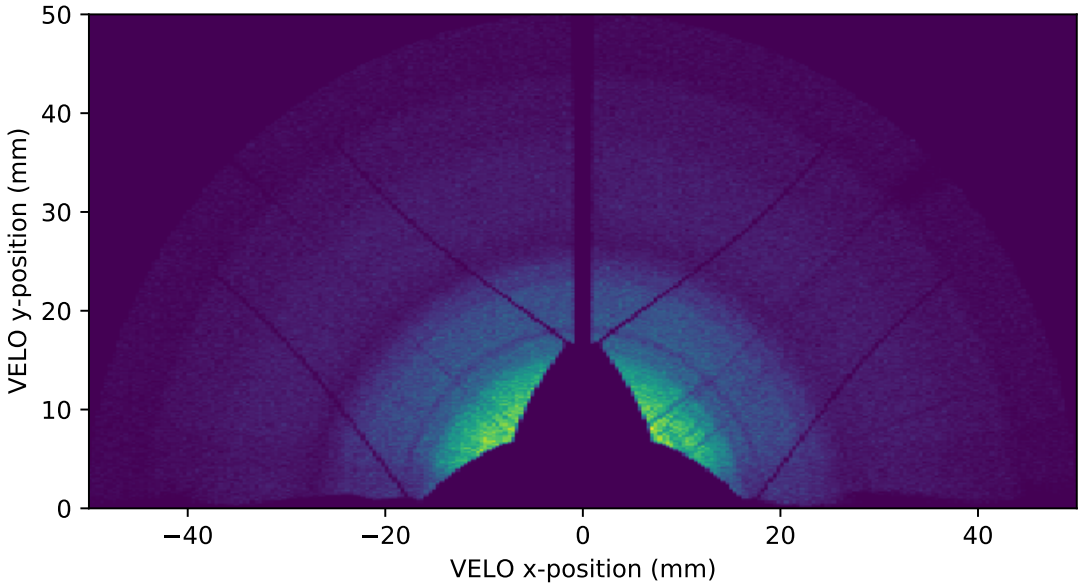
In the *R*-sensors, most routing lines are intercepting strips at large radius of the sensor (see Fig. 2.7). Hence, it explains why the CFE reduction is highest at large radii. For a comparison, the second metal layer effect is not present in the Φ-sensors as the routing lines from the inner region of the sensor run parallel to the strips of the outer region, as shown in Fig. 3.26.

Another study has been performed to directly compare the efficiency in the regions

with and without routing lines. This is possible because of an intricacy of the the VELO *R*-sensor geometry. The innermost part of the sensor does not require routing lines, as the strips within that part are already connected with routing lines that remain distant from the region. This is best seen in two self-imaging plots in Fig. 3.27. These plots show maps of track intercepts in the areas with and without routing lines, hence are "negatives" of each other.

The CFE can be calculated in the regions with and without the routing lines, to see a direct comparison of the two regions. The region not covered by the routing lines only extends to approximately 14 mm, hence the comparison is only possible in that region. A clear drop in efficiency is seen for the sections with routing lines in them, as shown in Fig. 3.28.

In order to quantify the Second Metal Layer effect and track its evolution in time it has to be decoupled from other effects. For example, the CFE also decreases if a number of strips on the sensor are malfunctioning. The removal of such strips from the CFE measurement has been described in Sec. 3.5.1. Also, the CFE may depend on the track category used (see Fig. 3.8). This can occur if a sensor is separated from its neighbouring sensor by a larger distance. An example of such gap in the VELO is between sensors 31 and 32, as shown in Fig 2.6. In this case, extrapolating the hit on sensor from the hits on the four downstream sensors (first category) will yield low CFE. This is because many lower momentum particles may be scattered out of the five-strip search window and form a cluster further away from the expected track intercept point. To mitigate this and similar effects, a software package was written which allows the monitoring of different aspects of *R*-sensor CFE. The software produces a panel of figures, showing the CFE vs Voltage, CFE for different track categories, CFE for different zones (quarters) of the sensor, the track residuals in linear and log scale (number of strips between the intercept point and cluster) as well as a 2D plot of inefficiency. The plots comparing the mentioned properties of a single *R*-sensor before and after bad strip removal are shown in Fig. 3.29.



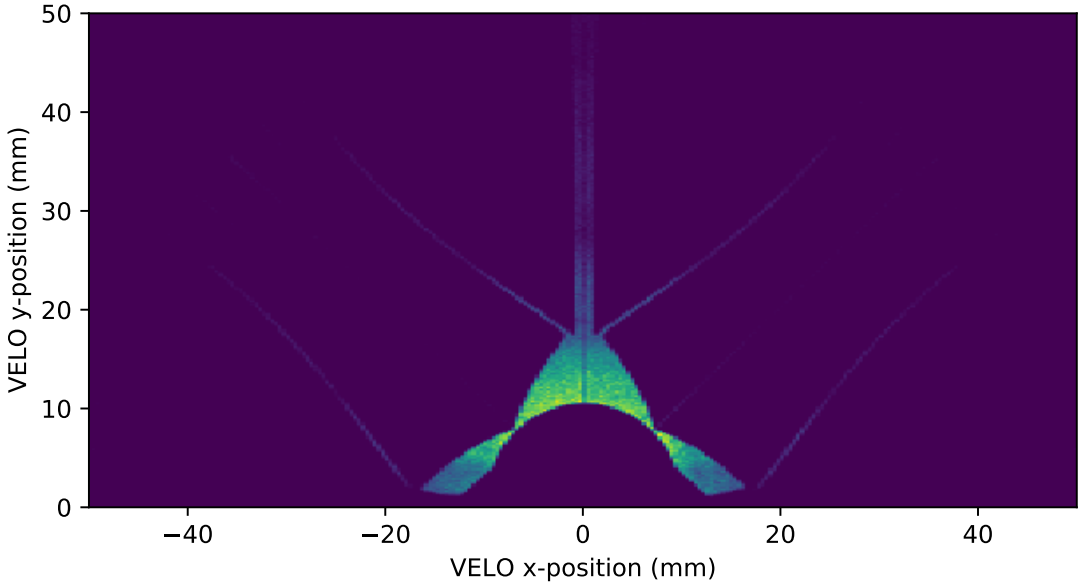


Figure 3.27: (Top) The heat-map of track intercepts of an *R*-sensor showing the layout of the routing lines. Bottom, the inverse of the previous image, showing the track intercept heat-map without routing lines in the region.

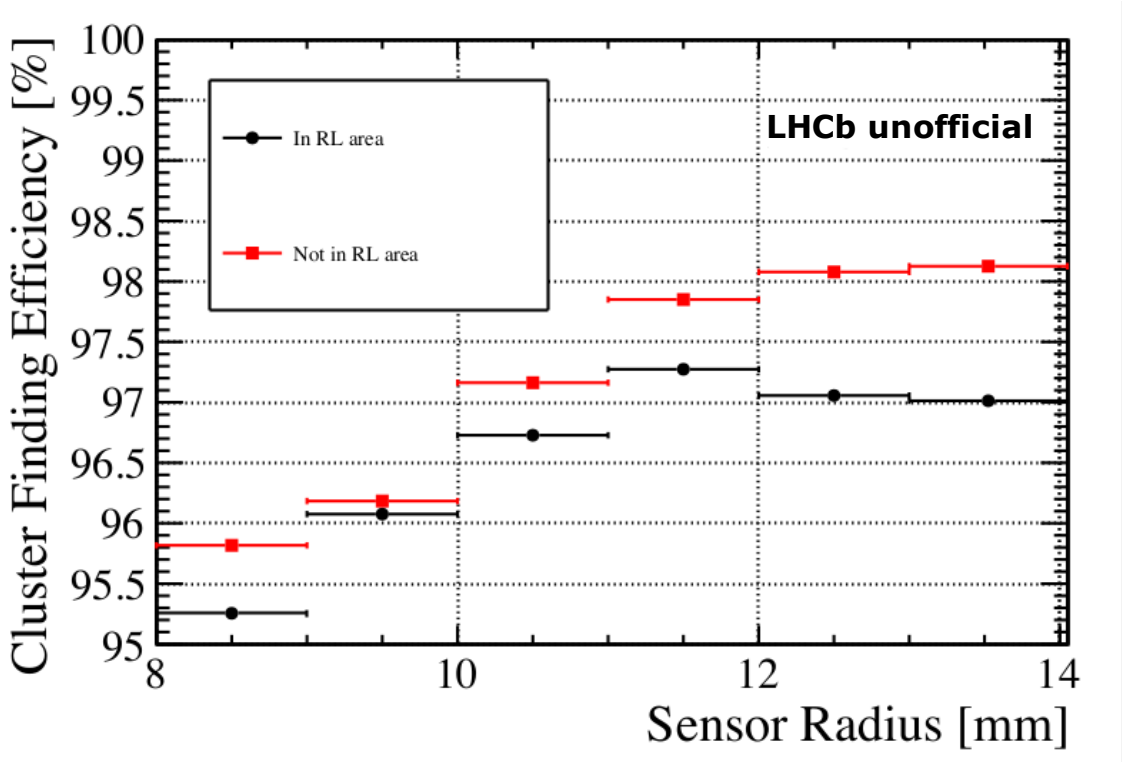


Figure 3.28: The CFE vs radius in the inner part of an *R*-sensor for track hitting the area with and without routing lines in it. As can be seen, the areas without the routing lines have higher efficiency in all radial regions. The study used the data from the final CCE Scan of 2018. The results are averaged over all *R*-sensors at their nominal operating voltage of 320V (with a small subset of sensors at 350V).

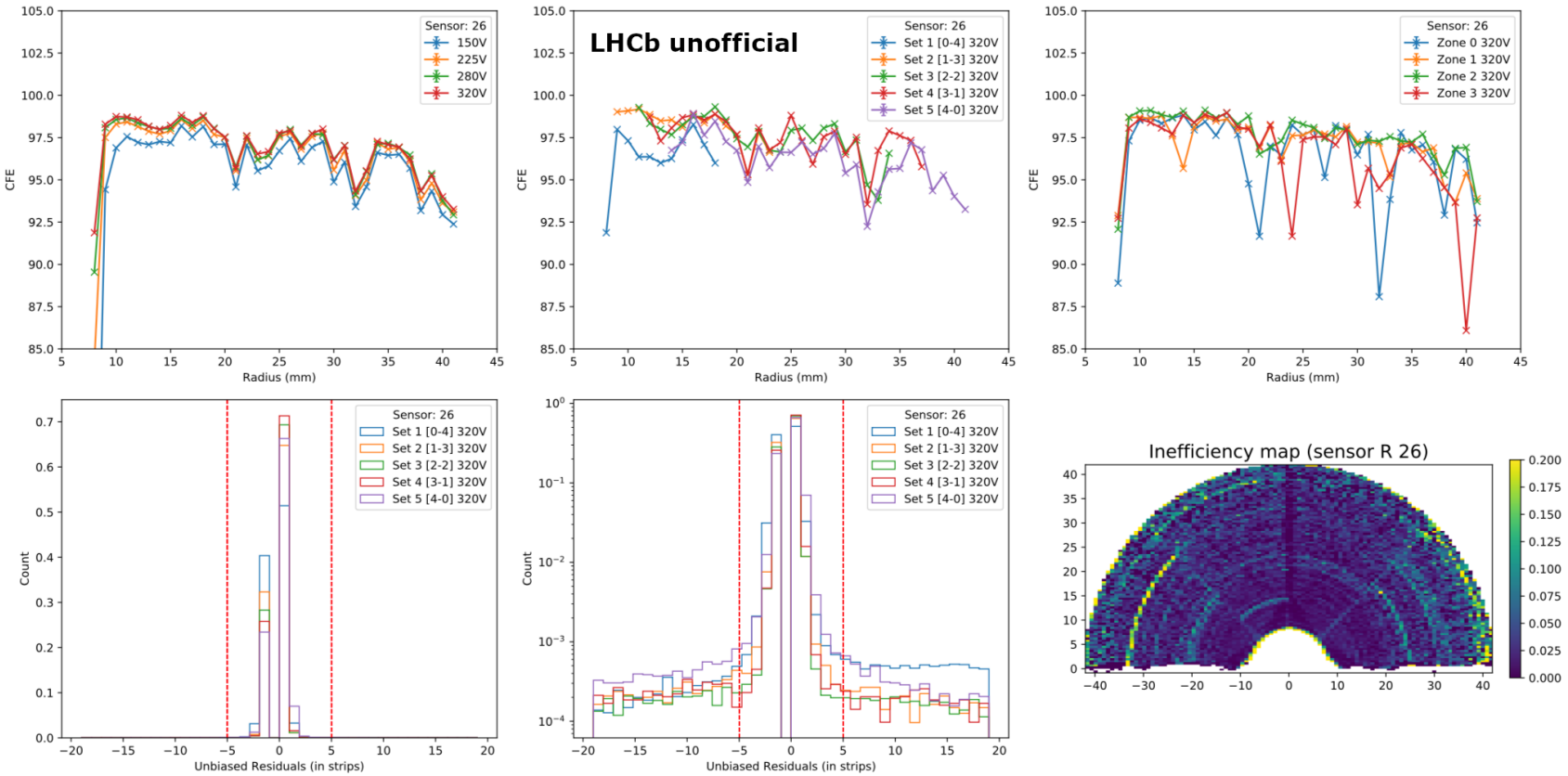


Figure 3.29: A panel of the CFE monitoring plots for R26 sensor. The shown plots show the trends prior to the bad strip removal. Top row of the plots show: The CFE vs radius for four different voltages, the CFE vs radius for five different categories (some trends terminate due to low statistics). The CFE vs radius for four different zone (quarters). The first two plots in the bottom row show: distance from the cluster to the extrapolated track intercept in number of strips on linear and log scale (the red dashed lines show the search window of the CFE method). The bottom right plot shows a 2D inefficiency map.

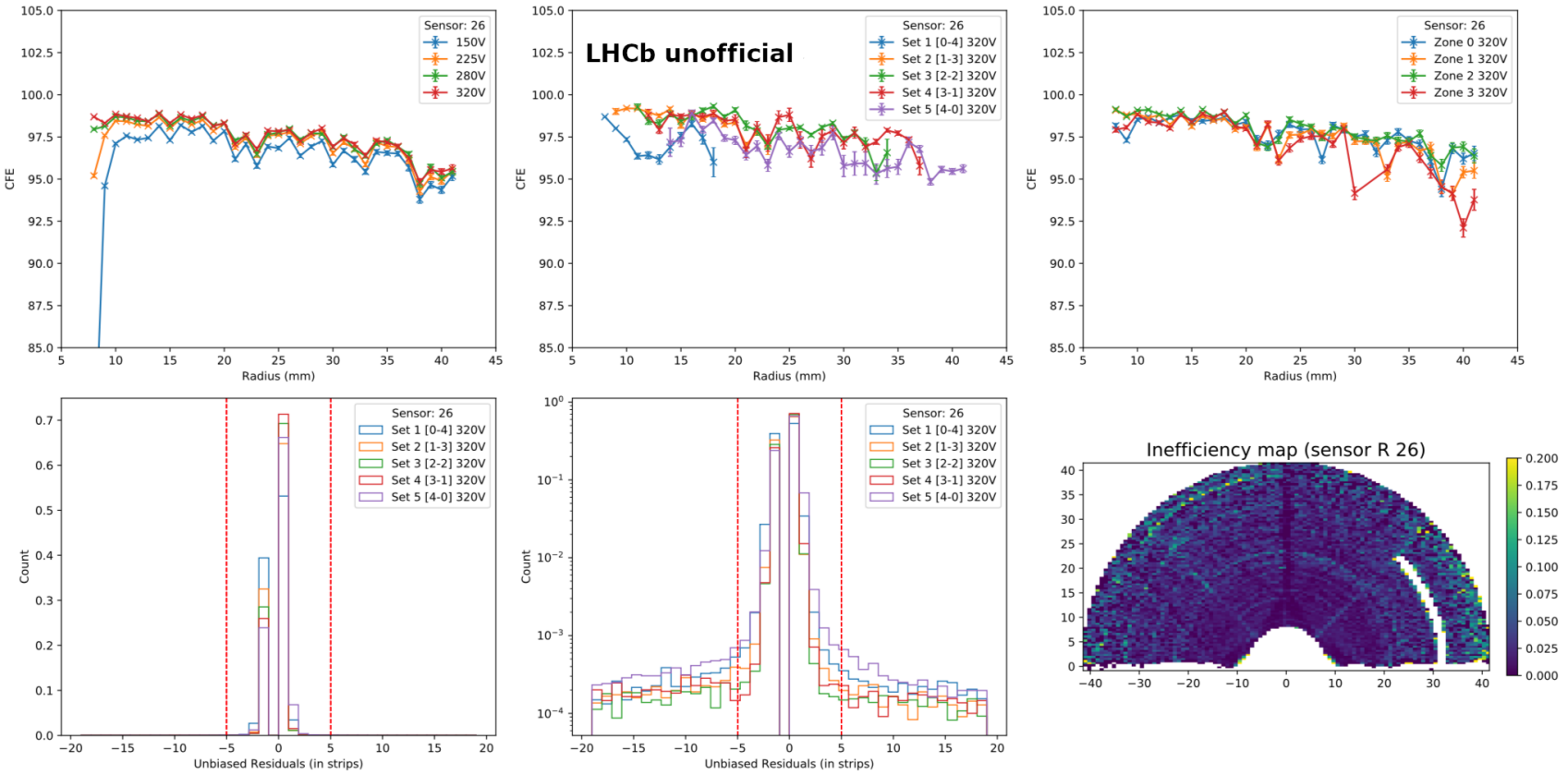


Figure 3.30: A panel of the CFE monitoring plots for R26 sensor after the bad-strip removal. As can be seen, the highly inefficient regions (shown in yellow) disappear after the bad strip removal. The white region on the inefficiency map corresponds to a group of noisy strips that were removed, even though they did not appear as inefficient.

The CFE can also depend on the angle of the track. This is caused in part by the CFE algorithm, requiring a track on 4 neighbouring sensors. There are very few tracks that are parallel to the beam (acute incidence) at large sensor radius. For this reason, many CFE measurements lack statistics for such angle categories. Similarly, for the sensors at the end of the detector, the track intercepting the innermost region cannot be reconstructed as they are outside the acceptance of the detector. For this reason, the inner region shows low efficiency with high uncertainty (five innermost strips have been removed from the calculation, reducing this effect).

The average CFE of all VELO *R*-type sensors vs the angle of incidence is shown in Fig. 3.31. The results are separated for the regions with and without routing lines. A clear reduction can be seen for the tracks with low angle of incidence in the routing line region at low radius.

Taking all of the above mentioned considerations into account, one can track the evolution of the second metal layer effect throughout different periods of LHCb operation. The CFE dependence on integrated luminosity of the four most downstream sensors of the VELO is shown in Fig. 3.32. The bad strip removal algorithms are applied on the dataset. Also, only tracks belonging to a stable track category and intercepting the sensors at large radius ($r > 35\,\text{mm}$) and near a routing line ($d < 200\,\mu\text{m}$) are considered. For comparison. The bottom plot shows the same calculation for tracks away from the routing line. Clear improvement in the CFE can be observed. Similar plots for four central VELO sensors (R24-27) are given in Fig. 3.33

The CFE is seen to drop with initial irradiation and then stabilize. However, the trend for non-routing line regions does not change significantly with increased luminosity. This is an indication, that the reduction in CFE is indeed correlated to the presence of routing lines and caused by the Second Metal Layer effect. Also, this is compatible with the surface damage hypothesis. After the initial reduction, the CFE was stable and did not significantly reduce the performance of the VELO detector.

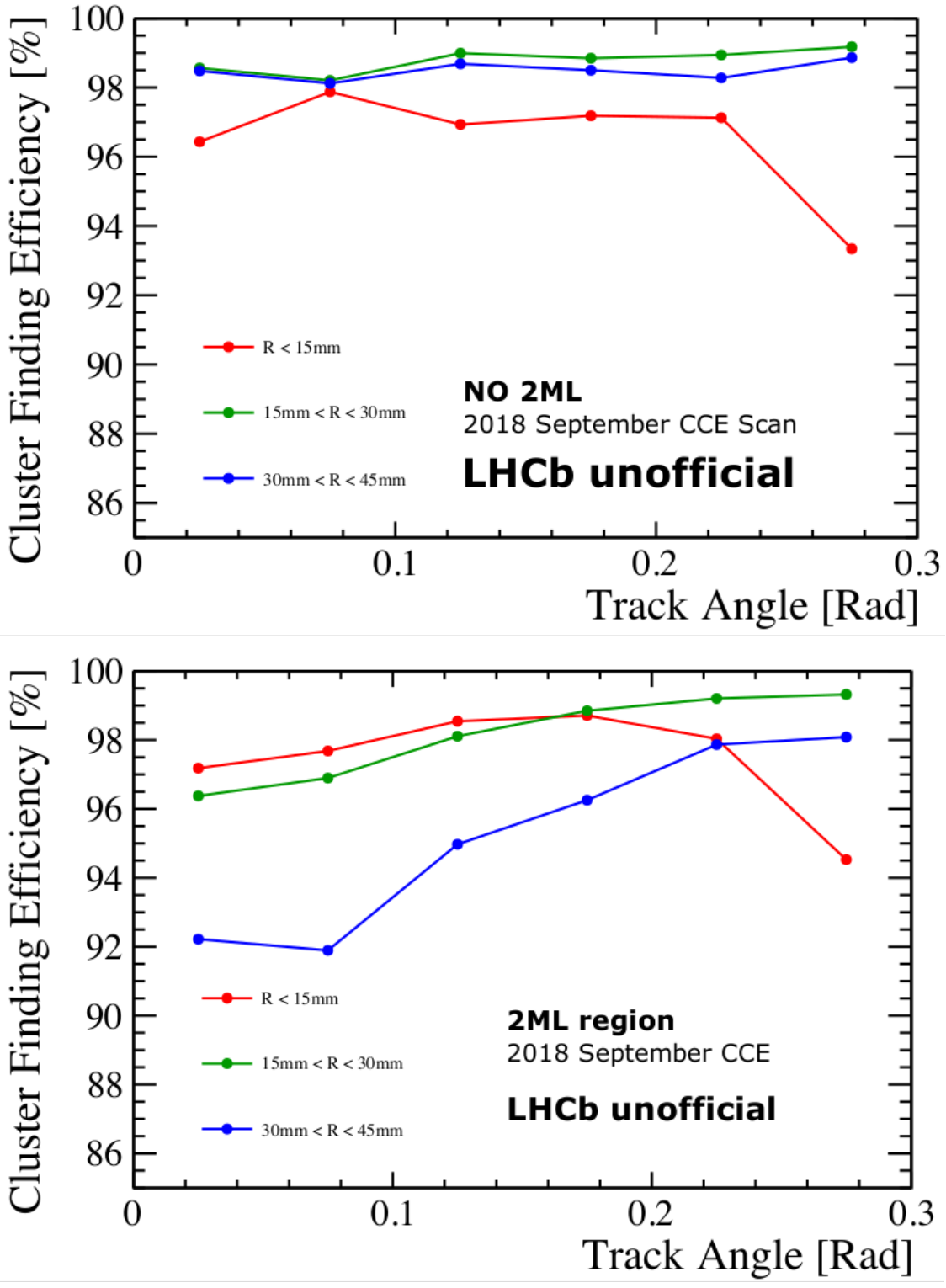


Figure 3.31: The average of the CFE of all VELO *R*-sensors from the 2018 September CCE scan data. Top plot shows the CFE vs track angle for 3 different radial intervals in the region without the routing lines (No 2ML). Bottom plot shows the same measurement in the region with routing lines (2ML). A clear reduction in the CFE is observed for low radius and low track angle tracks. The study used the data from the final CCE Scan of 2018. Sensors operated at their nominal voltages of 320V (350V for a small subset).

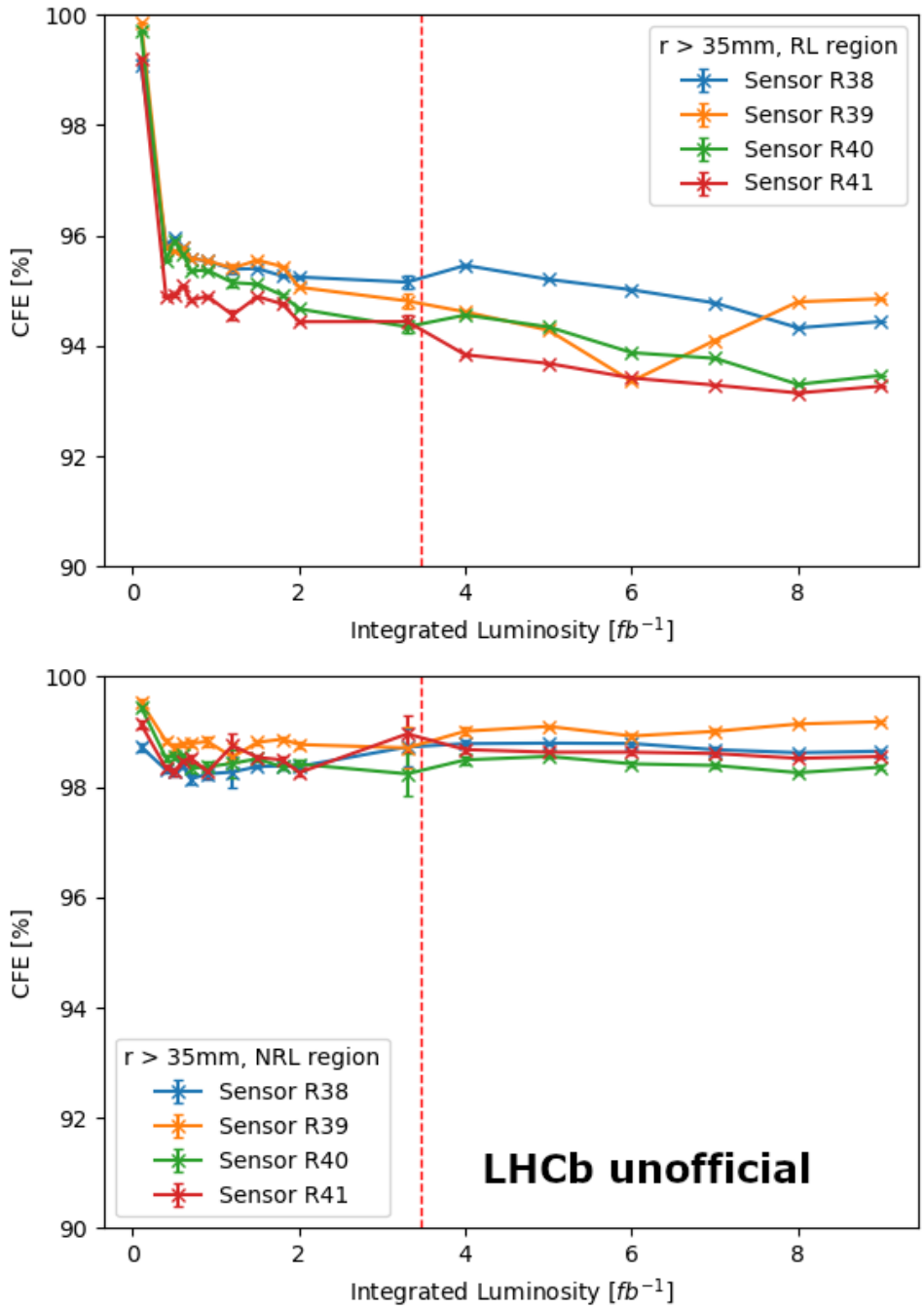


Figure 3.32: The CFE dependence on luminosity for VELO sensors at the far downstream end of the VELO (R38-R41). Only the tracks intercepting the sensors at large radius ($r > 35$mm) are considered. (Top) CFE measured from the tracks intercepting near the routing lines ($d < 200\,\mu$m). (Bottom) CFE measured from the tracks away from the routing line. The vertical dashed line shows the start of Long Shutdown 1.

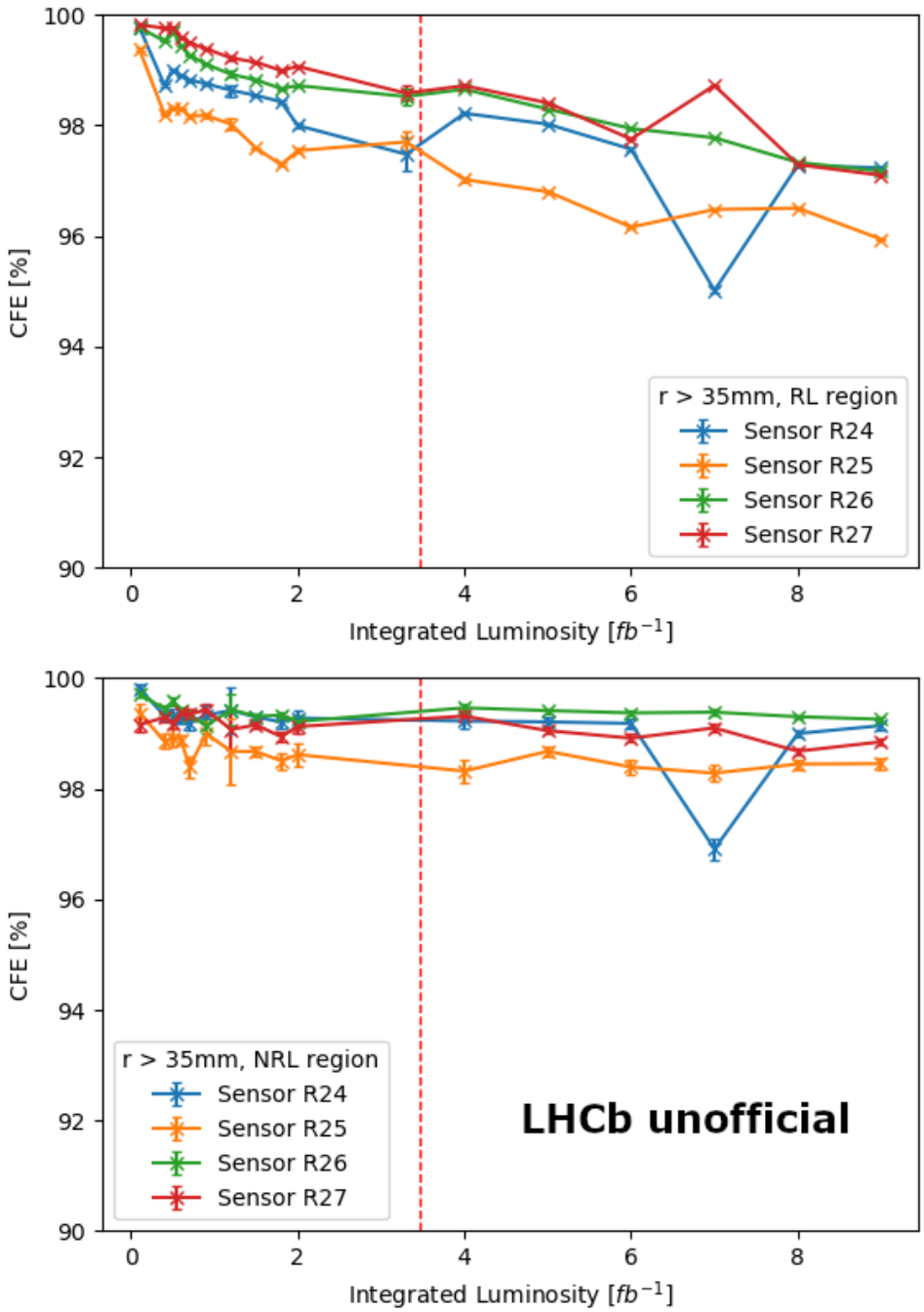


Figure 3.33: The CFE dependence on luminosity for VELO sensors at the center of the VELO (R24-R27). Only the tracks intercepting the sensors at large radius ($r > 35$mm) are considered. (Top) CFE measured from the tracks intercepting near the routing lines ($d < 200\,\mu$m). (Bottom) CFE measured from the tracks away from the routing line. The vertical dashed line shows the start of Long Shutdown 1.

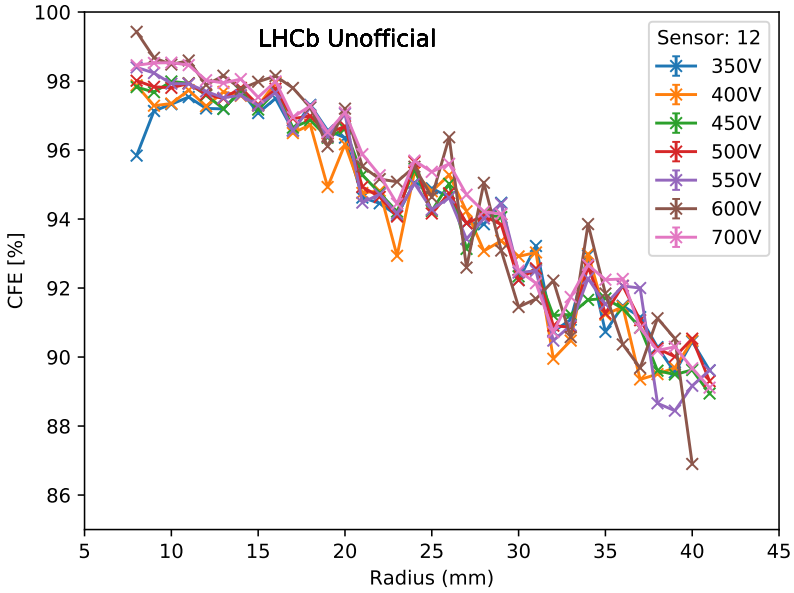


Figure 3.34: CFE vs radius for a VELO sensor R12. The blue line shows the results with the nominal operating voltage of the sensor. The observed reduction in CFE at high radius is believed to be due to the second metal layer effect.

3.5.3 Final Runs of LHCb

At the end of Run 2 (2018) most VELO sensors were operating at 350V. It was suspected that the most irradiated regions of the sensors might be under-depleted. Multiple tests (IV and CCE scans) were performed at voltages of up to 400V. During the last two LHC fills of Run 2 a special test has been performed. The voltage of VELO sensors was incrementally increased up to 700V, while the design voltage was only 500V. This was an attempt to study the thermal runaway of the sensors and allowed measurement of the CFE dependence on the depletion voltage. This test was carried out during the last LHC fills of the VELO operation period, which allowed measurement of the detector performance past its design limitations. If the sensors or system were damaged during these tests it would be no loss since they will be replaced by the Upgrade for Run 3. All sensors operated well and a slight increase in the efficiency was observed up to 600V. Beyond that, issues were seen in some sensors, causing efficiency drops. The CFE of a single sensor (R12) for different voltage values is shown in Fig. 3.34. The most irradiated region (first points on the left of the plot) clearly shows an increase in efficiency from 350V to 400V. A comparison of the CFE for different voltages (350V, 450V and 700V) is shown in Fig. 3.35. The drop in efficiency at higher radius is due to the second metal layer effect.

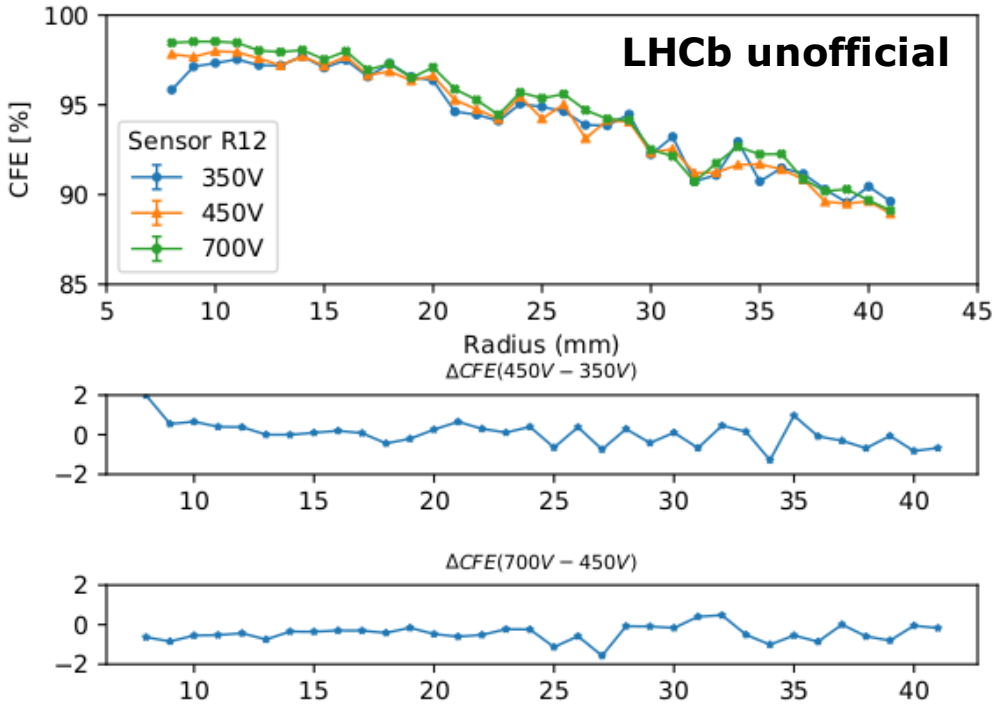


Figure 3.35: (Top) CFE vs radius for a VELO sensor R12. (Bottom) Difference of CFE for different voltages. The nominal voltage of 350V exhibits the lowest CFE in the small radius region.

3.6 The CCE and CFE studies of the n^+-on-p sensors

During the VELO construction, it became possible to produce a small number of sensors with the n^+-on-p technology. This was a novelty at the time and was utilised with some risk as less testing had been performed on sensors with this technology. Also, not a lot of testing has been done on the sensors with this technology. In order to investigate the possible use of n^+-on-p sensors in the future, two such sensors were installed in a single module of VELO (R0 and Phi64). This module is the most upstream module, and consequently would limit any negative effect if its performance was not as good as expected. The module was fully operational throughout the running of the LHCb detector.

A simple linear least-squares fit is applied on the EDV vs Fluence distributions for the n^+-on-p and n^+-on-n sensors using data obtained from the September 2018 CCE scan. Each set of points from the three outermost regions of each sensor are fitted independently with a straight line. As the uncertainties on the points are initially unknown the fit is performed with arbitrary uncertainty on each point. For the n^+-on-n fits are performed on 40 sensors. The standard deviation of the distribution of gradients from these fits is calculated, and used as the uncertainty on the single n^+-on-p fit. For the n^+-on-n fits the corresponding uncertainty on the mean of the gradients is used. The result of the determination for the n^+-on-n sensors is $(3.10 \pm 0.17) \times 10^{-13}\ V/(1\,\text{MeV}\ n_{eq}\ \text{cm}^{-2})$. For the the n^+-on-p sensor the result is $(5.57 \pm 1.08) \times 10^{-13}\ V/(1\,\text{MeV}\ n_{eq}\ \text{cm}^{-2})$. Consequently the extracted gradients are compatible at the 2.3σ level, and while the n^+-on-p value is somewhat larger than the n^+-on-n the rise in EDV with fluence is thus compatible at the precision of this test.

Furthermore, the CFE of the n^+-on-p R-type sensor can be measured. The CFE vs radius of the R0 sensor for different bias voltage values is shown in the Fig. 3.37. It is consistent with the CFE for other sensors and shows no major reduction in performance. The Second Metal Layer effect is also visible in the n^+-on-p sensor, as shown in Fig. 3.38. However, unlike the downstream sensors (see Fig. 3.32), the decrease in the CFE for R0 sensor is more gradual and reaches a constant at approximately $2\,\text{fb}^{-1}$. The CFE of the sensor drops further by approximately 1% between $7\,\text{fb}^{-1}$ and $9\,\text{fb}^{-1}$ (end of Run 2). This suggests that the n^+-on-p behaviour differs from the n^+-on-n sensors, however no strong conclusions can be drawn as only one n^+-on-p sensor is available for comparison.

Throughout the years of LHCb operation, the two n^+-on-p sensors operated with no issues. The sensor performance and radiation damage was compatible with the n^+-on-n sensors. The n^+-on-p technology is used in a lot of the currently ongoing LHC detector upgrades [29, 123]. In particular, the new VELO detector will use silicon-pixel sensors with the n^+-on-p design [49].

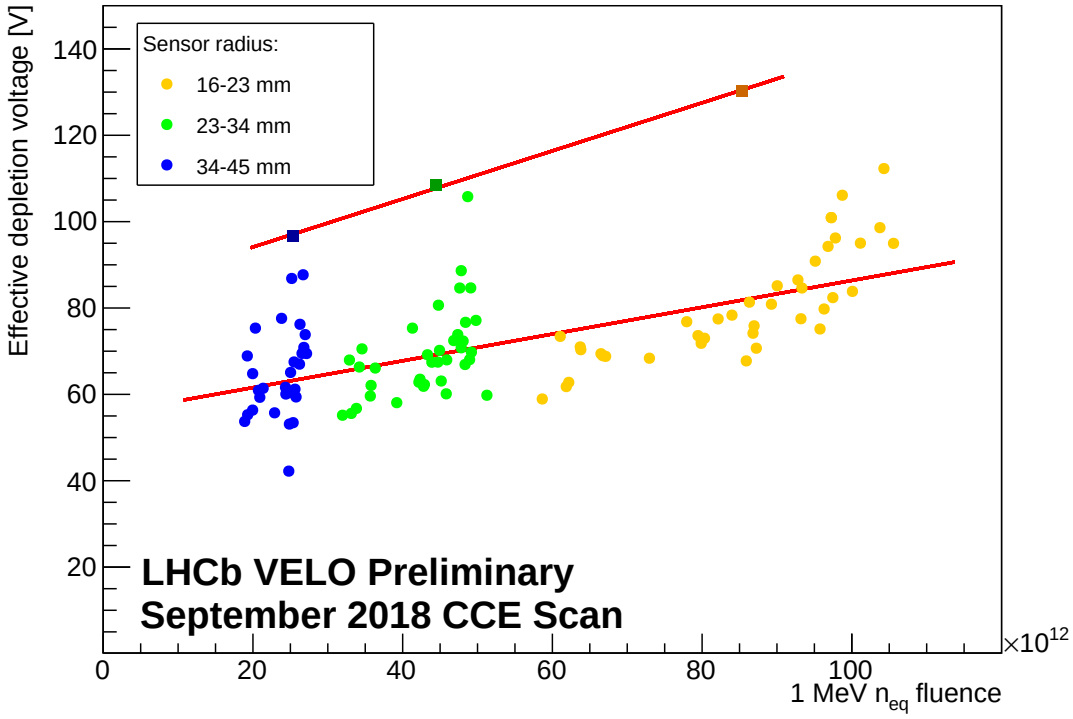


Figure 3.36: The EDV dependence on fluence of all VELO sensors for three radial regions. The R0 n^+-on-p sensor measurements are shown with darker squares. Simple linear fit is performed for n^+-on-p and n^+-on-n sensors and shown as red lines.

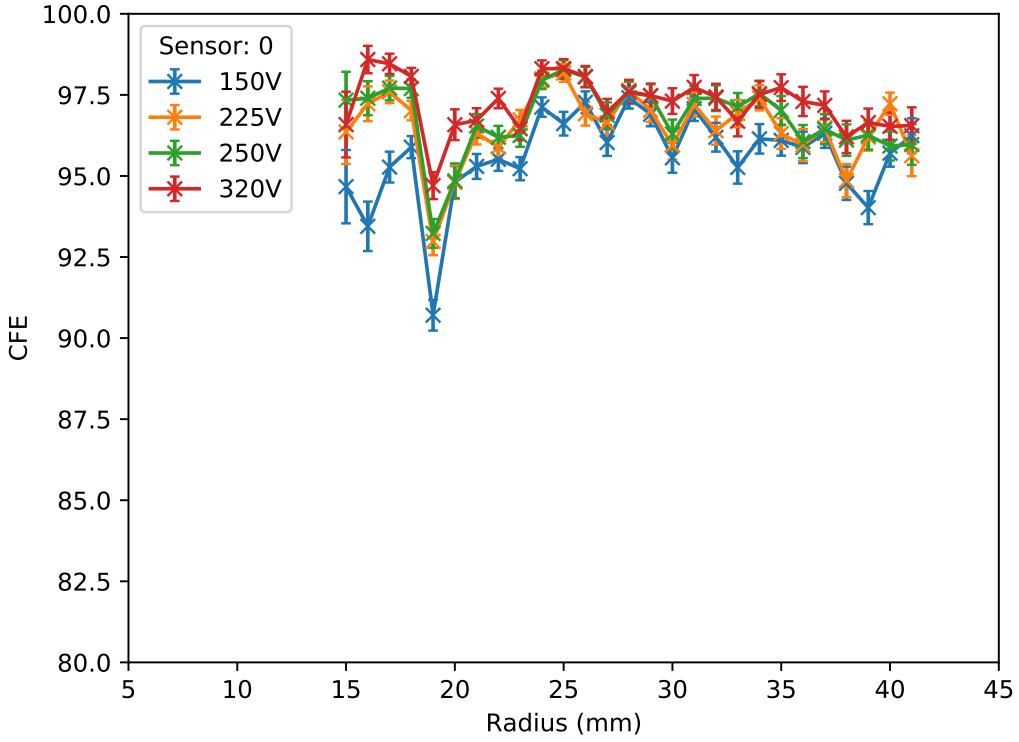


Figure 3.37: The CFE vs radius for different bias voltages of the R0 sensor using the n^+-on-p technology. The CFE trends are similar to the ones of n^+-on-n sensors. The last CCE scan of Run 2 data is used for the CFE measurement.

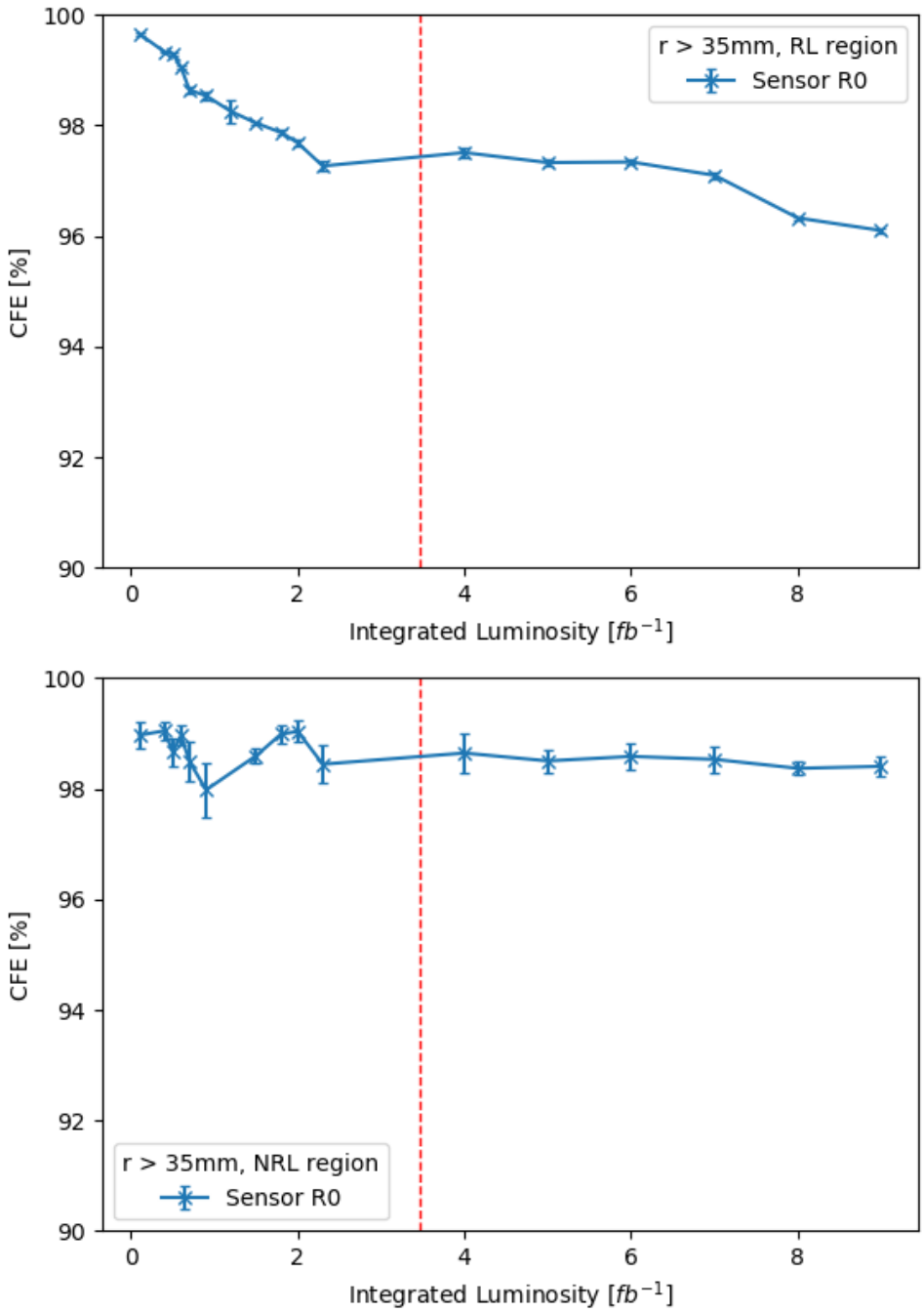


Figure 3.38: The CFE dependence on luminosity for the R0 VELO sensor which uses the n^+-on-p technology. Only the tracks intercepting the sensors at large radius ($r > 35$mm) are considered. (Top) CFE measured from the tracks intercepting near the routing lines ($d < 200\,\mu$m). (Bottom) CFE measured from the tracks away from the routing line. The vertical dashed line shows the start of Long Shutdown 1.

3.7 Conclusions

The radiation damage of the VELO subdetector has been studied using the CCE Scan data collected in Run 1 and 2 of the LHC. The CCE and CFE of the VELO silicon-strip sensors have been measured in order to quantify the radiation damage and track its evolution during different data taking periods. At the end of Run 2, most VELO sensors operated at 350V. The voltage steps used in the CCE scans have been changed accordingly, introducing 360V and 400V and removing 20V and 200V.

The agreement between the observed EDV dependence on fluence and the Hamburg model prediction has been investigated. The data and the model are shown to be broadly in agreement. Using guidance from the model, beneficial annealing was conducted on the VELO towards the end of Run 2. As a consequence, the EDV of some VELO sensors is measured to have dropped by 70-80V. To the knowledge of this thesis author, this may be the first instance of intentional beneficial annealing for the purpose of increasing the performance of a working silicon detector.

The CFE of the VELO sensors has been measured for different subsets of the CCE Scan data. Two bad-strip removal algorithms have been devised and implemented to isolate the CFE reduction effects from the actual radiation damage. An effect of charge loss to the second metal layer routing lines is observed. The Second Metal Layer effect has been studied and its evolution in time investigated. A special high-voltage test has been performed at the end of the VELO operation. An increase in the CFE has been observed with increased bias voltage. The VELO sensors have been tested with voltages beyond their maximum design voltage. This test constituted the final stage of the VELO operation. Afterwards, the VELO was removed in order to be replaced by the new VELO upgrade detector.

Chapter 4

CP Violation Search Methods

In the Standard Model of particle physics, *CP* violation arises from a non-vanishing complex phase term in the CKM matrix [118]. Beyond Standard Model particles might contribute to *CP* violation, thus studying it provides an indirect probe for New Physics.

A number of model independent direct *CP* violation search techniques exist. One of the most commonly used methods is the binned χ^2 approach, where the relative event densities in the phase space of a particular decay are compared [17]. Another binned method, called Triple Product Asymmetries (TPA), is based on the measurement of the final state particles' momenta triple products. This technique has been used in the previous *CP* violation search in $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays [53]. Both of these methods are discussed below. There have also been many applications of unbinned techniques, including the nearest neighbours approach [73] and angular moments of the cosine of the helicity angle [34]. The analysis described in Chap. 5 uses an unbinned, model independent technique known as the energy test to search for *CP* violation in $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays. The energy test has been used by the LHCb collaboration previously (first applications in three-body decays [71] and four-body decays [74]). A detailed description of the energy test technique is given in this chapter.

4.1 Binned χ^2

The simplest *CP* violation search technique is the binned two-sample χ^2 test, usually referred to as S_{CP} [34]. This technique is sensitive to local *CP* violation and neglects any global effects. It is based on measuring the statistical significance, which is obtained by calculating the difference in number of events in bins of samples X and $\overline{X}$. In the case of X and $\overline{X}$ being the particle and anti-particle samples this constitutes a test for *P*-even *CP* violation (different types of *CP* violation are discussed in Chap. 1).

$$S^i_{CP} = \frac{N^i(X) - \alpha N^i(\overline{X})}{\sqrt{N^i(X) + \alpha^2 N^i(\overline{X})}}, \quad \alpha = \frac{N_{tot}(X)}{N_{tot}(\overline{X})} \tag{4.1}$$

where $N^i(X)$ and $N^i(\overline{X})$ denote the numbers of X and $\overline{X}$ candidates in the i^{th} bin. For large N^i values, Gaussian uncertainties may be assumed. The ratio between total X and $\overline{X}$ event yields is denoted α and introduced to account for possible global production or detection

asymmetries. The small α^2 significance correction factor in the denominator varies in literature [34, 73]. The form given above is recommended.

In the case of no CP violation, the S^i_{CP} follows a Gaussian distribution with $\mu = 0$ and $\sigma = 1$. The value of χ^2 is computed from $\chi^2 = \sum (S^i_{CP})^2$ and together with the number of degrees of freedom ($N_{bins} - 1$ in this case) allows direct calculation of the p-value, which shows the compatibility of the measurement on data in comparison to the no CP violation hypothesis.

The χ^2 method is easy to implement and does not require a lot of computing resources. The choice of binning scheme used in this method needs to be made considering the available data yield and expected localization of CP violating effects in the phase space of the decay under investigation. A sufficient number of events per bin is required in order to avoid a reduction in sensitivity, yet using finer binning might be beneficial, as the contribution from local effects will be enhanced. Different analyses have used this method with various numbers of bins. For example, an initial application used $\mathcal{O}(10^3)$ bins [34, 17], while later analyses used $\mathcal{O}(10 - 10^2)$ bins [70, 73].

This method was used to successfully observe CP violation in $B^+ \to h^+h^+h^-$ decays, where h is either a pion or a kaon [52, 51], where $\mathcal{O}(10^2)$ bins were used with their placement motivated by the resonant structures in this decay. Adaptive binning was used to keep similar numbers of events in each bin.

4.2 Triple Product Asymmetries

Another approach to search for the P-odd CP violation is Triple Product Asymmetries [94, 92] which utilises the $\hat{T}$ operator, sometimes called "naive time reversal" or "motion reversal". This unitary operator reverses the momentum and spin three-vectors without exchanging initial and final state particles (as done by anti-unitary T operator). A group of measurable quantities that are $\hat{T}$-odd (i.e. change sign under $\hat{T}$ transformation) are the scalar triple products of final state particle momenta or spin three-vectors: $C_{\hat{T}} \equiv \vec{v}_1 \cdot (\vec{v}_2 \times \vec{v}_3)$, where $\vec{v}_i$ are measured in the rest frame of the parent particle. The triple products are used to define $A_{\hat{T}}$ and $\overline{A}_{\hat{T}}$ asymmetry variables

$$A_{\hat{T}} = \frac{N(C_{\hat{T}} > 0) - N(C_{\hat{T}} < 0)}{N(C_{\hat{T}} > 0) + N(C_{\hat{T}} < 0)}, \qquad \overline{A}_{\hat{T}} = \frac{\overline{N}(-\overline{C}_{\hat{T}} > 0) - \overline{N}(-\overline{C}_{\hat{T}} < 0)}{\overline{N}(-\overline{C}_{\hat{T}} > 0) + \overline{N}(-\overline{C}_{\hat{T}} < 0)}, \quad (4.2)$$

where N and $\overline{N}$ are the numbers of particle and antiparticle decays. $\overline{A}_{\hat{T}}$ is constructed form charge conjugate final state particles momenta or spins. These asymmetry variables are both $\hat{T}$-odd and P-odd, hence they allow the measurement of the P-odd contributions to CP violation only. They also might be affected by final state interactions that could introduce fake asymmetry. For this reason, it is necessary to to measure the difference between $A_{\hat{T}}$ and $\overline{A}_{\hat{T}}$. This factorises the weak and strong CP violating phase contributions.

The P and CP violating observables are thus defined as

$$a_P^{\widehat{T}\text{-odd}} = \frac{1}{2}\left(A_{\widehat{T}} + \overline{A}_{\widehat{T}}\right), \qquad a_{CP}^{\widehat{T}\text{-odd}} = \frac{1}{2}\left(A_{\widehat{T}} - \overline{A}_{\widehat{T}}\right), \tag{4.3}$$

where the $1/2$ term is added to normalize the weak phase [104]. A significant deviation from zero in these observables would imply P or CP violation. Before measuring these asymmetries, the phase space of the decay under investigation is split into bins, thus enhancing the contribution from local effects. By construction, this method is largely insensitive to production and detection asymmetries [75].

The Triple Product Asymmetries technique was used in the analysis, which provided the first evidence for CP violation in baryons [53]. The analysis of this channel with more data and two methods (Triple Product Asymmetries and energy test) is described at length in Chap. 5.

4.3 Energy Test

While the CP violation search techniques presented in Sec. 4.1 and Sec. 4.2 are based on splitting the decay phase space into smaller partitions (bins) and measuring certain variables for each subset of the data, there is also a group of methods which avoids binning altogether. This might be preferable and in certain cases may lead to a higher sensitivity to CP violation. One such technique, called the energy test, introduced in Refs. [7, 8] uses the distance between points distributed in a multi-dimensional space to obtain correlation between two samples of data.

The intrinsic design of the energy test lends itself perfectly to CP violation studies. The first proposal to apply this technique in Dalitz plot analyses was given in Ref. [150]. This technique was already applied in CP violation searches in three and four-body decays [69, 74]. The description of the energy test follows Refs. [12, 131], where more details can be found.

The energy test allows for an evaluation of the probability that two samples originate from the same underlying statistical population. This is achieved by calculating a test statistic T, which is proportional to the average distances between points belonging to two samples and populating a multi-dimensional space. In the context of CP violation, the samples are the particle and anti-particle events populating the phase space of the corresponding decay, spanned by a set of physical coordinates (the discussion of the choice of coordinates is given in Sec. 4.3.5). The test statistic is defined as

$$T = \sum_{i,j>i}^{n} \frac{\psi_{ij}}{n(n-1)} + \sum_{i,j>i}^{\overline{n}} \frac{\psi_{ij}}{\overline{n}(\overline{n}-1)} - \sum_{i,j}^{n,\overline{n}} \frac{\psi_{ij}}{n\overline{n}}, \tag{4.4}$$

where the first and second terms correspond to a weighted average distance between matter and antimatter events respectively, while the third term corresponds to the weighted average distance between events of opposite flavours. The indices i and j refer to a pair of events. The n and $\overline{n}$ are the corresponding matter/antimatter yields in the data and allow

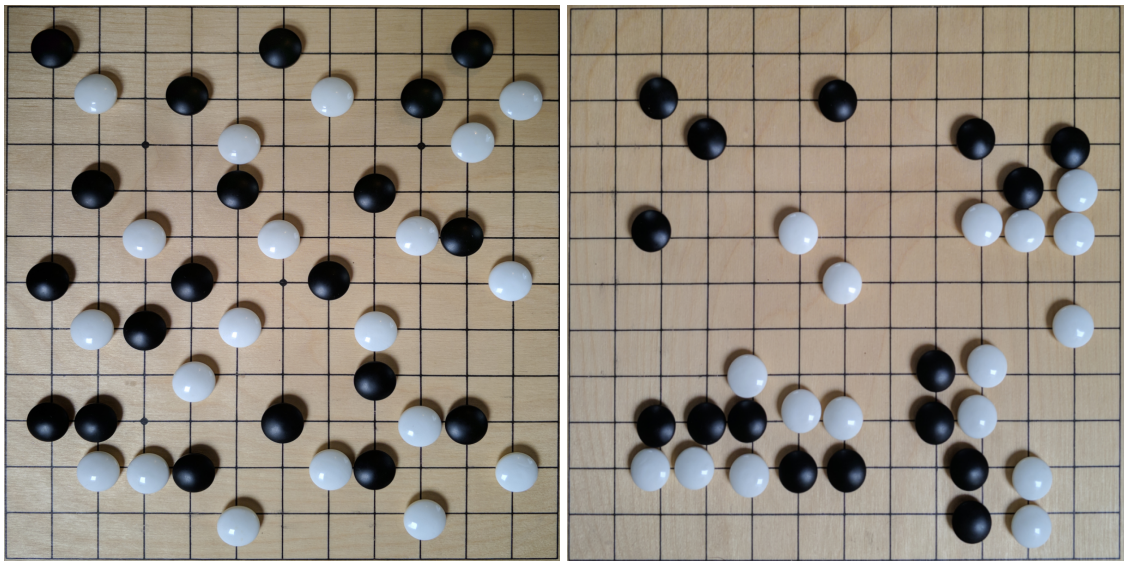


Figure 4.1: An illustration showing two distributions of oppositely charged particles. If the distributions are identical (left), the energy test would yield the T value close to zero. If, however, the distributions are different (right), the T value will be larger than zero, showing that the distributions are not consistent with coming from the same underlying population. The grid in the illustration has no analogous meaning, as the energy test is an unbinned technique.

the normalization of the three terms in order to remove the impact of global asymmetries in the samples that might originate from detection or production effects. The weighting function ψ_{ij} should depend on the distance between candidates in phase space. If it is independent of the distance between candidates, then $T = 0$. Furthermore, it should decrease with increasing Euclidean distance d_{ij} between events i and j. This way the sensitivity to local effects is enhanced. The distance value depends on the choice of coordinates used to define the phase space. This choice is discussed in Sec. 4.3.5. A Gaussian function is chosen[1], defined as $\psi_{ij} \equiv \psi(d_{ij}) = e^{-d_{ij}^2/2\delta^2}$, where δ is a tunable parameter defining the effective radius in phase space within which test sensitivity to local asymmetries is enhanced. This is analogous to a the choice of bin-width (number of bins) in binned analyses. The choice of δ and its impact on the sensitivity is described in Sec. 4.3.1. If the two samples are identical (there is no *CP* violation), the T value will be close to zero, while any significant deviation from zero would mean the existence of asymmetry in the two sample as shown in the illustration 4.1.

The distribution of the test statistic T is not know a priori, hence the significance of the measured value needs to be assessed by comparing it with the distribution of T values from permutation samples. These samples are constructed by randomly assigning flavour to each event, thus making the samples *CP*-symmetric. The p-value and significance can be calculated by counting the number of T values from permutations, that are greater than

[1]The original suggestion was to use $\psi_{ij} \equiv 1/|\Delta\vec{v}|$, in which case the expression for T obtains the form of the electric potential of a system of particles with opposite charges, where a non-identical distribution of charges would lead to a non-zero electrostatic energy of the system. This is the etymology of the name of the energy test method.

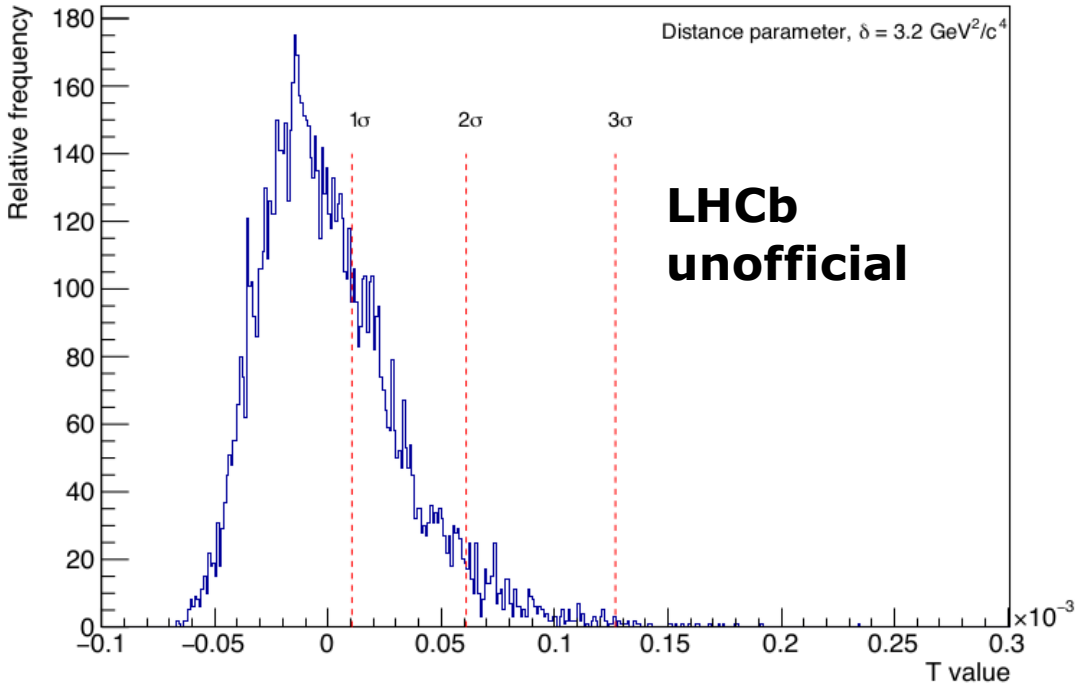


Figure 4.2: T value distribution, obtained from permutations of real data. Red lines show $1\sigma, 2\sigma$ and 3σ significance limits. Reproduced from [14].

the T value observed in data as shown in Fig. 4.2.

In the case of a highly significant result, a large number of permutations is required (possibly $\mathcal{O}(10^8)$), which might lead to impractical computational time requirements, especially if the data samples are large. A way to overcome this limitation, known as the Scaling Method, is described in Sec. 4.3.2.

4.3.1 Sensitivity of the Energy Test

The sensitivity of the energy test method depends on several factors. First of all, it strongly depends on the choice of coordinates that span the phase space of the decay. Choosing the optimal set of coordinates is discussed in the Sec. 4.3.5. Secondly, the sensitivity depends on the choice of distance parameter δ. This dependence is complex and not described analytically. In most general terms, δ should not be smaller than the distance resolution, but not too large to dilute local effects. The optimal value of delta might change depending on many factors, for example: the overall size of the phase space (defined by the mass of the parent particle), the signal yield, width of the dominant resonance contributions in the decay channel, etc.

Multiple studies using the simulated data (see Sec. 5.4) have been performed to find the optimal choice of δ. For example, two different toy models, where CP violation is included in a different resonance are shown in Fig. 4.3. It can be seen that in certain cases the sensitivity dependence has an extremal value of δ, while in others it behaves asymptotically, motivating larger value of δ to be chosen. Since the behaviour of delta is

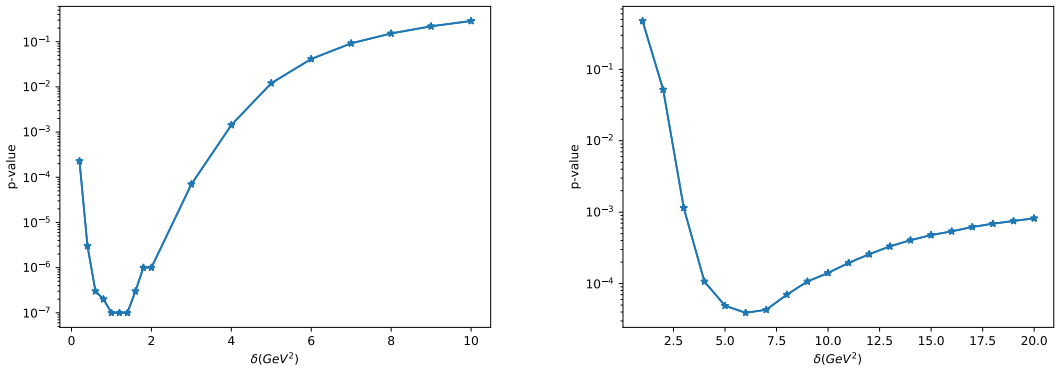


Figure 4.3: The dependence of p-value on δ in two simplified simulation models where *CP* violation is included via Δ^+ resonance (left) and a_1 resonance (right). Different minima and asymptotic behaviour is observed in these models.

strongly dependent on the way *CP* violation is included in the simulated sample it was decided to adopt a more general, model-independent and data-driven approach, where several different δ values are used to give a good coverage of the phase space. The specific values are chosen considering the mean distance between events in the phase space as well as mean distance between 600 nearest neighbours. This number of nearest neighbours is chosen because in a binned analysis with local *CP* asymmetry of 20% one would need at least 600 events per bin, to achieve 5σ (observation threshold) significance.

4.3.2 Scaling Method

The energy test requires running a large number of permutations in case of a significant result. In particle physics, datasets commonly contain $\mathcal{O}(10^4 - 10^6)$ events, meaning that permuting the dataset millions of times might not be feasible. To alleviate this problem, Refs. [7, 8] suggest the use of a Generalized-Extreme-Value (GEV) function to fit the T distribution. The corresponding p-value is then calculated from the fraction of the integral above the observed T value. This method has been applied in [69, 74]. However, due to concerns of overestimating the significance and incomplete understanding of the GEV function behaviour, this method is depreciated.

A new suggestion to deal with the need for many permutations is presented in Ref. [12] and is based on a newly discovered property of the T value distribution, allowing the use of a small subset of the data. It has been empirically shown in Ref. [12, 102] that in the case, where same events are not reused to calculate distances between pairs of events increasing the sample size by a factor k will scale each of the three terms in the T expression (eq. 4.4). This means that for sufficiently large samples, the distribution of kT is independent of the sample size. This scaling property can be used to create T distribution under a null hypothesis (e.g. no *CP* violation), by using a small fraction ($\mathcal{O}(10^2)$) of events and multiplying the resulting T values by the ratio between the full dataset size and the subset size. This saves up a lot of computing time as the energy test only needs to be run

on the full dataset once for the actual T value of the data, while for permutations it can be run on a small subsample of the data, which can be quickly repeated millions of times. The scaling property has been used for the first time in the analysis described in Chap. 5.

4.3.3 Visualization of Asymmetries

The energy test yields a single p-value, corresponding to the probability that the two samples used originate from the same underlying population, i.e. the probability that there is no *CP* violation in the decay under investigation. In case a significant asymmetry is found in the samples, it is essential to understand its physical sources. This can be done by doing a full amplitude analysis, but for a four-body decay it might be currently unobtainable. The construction of the energy test allows for visualisation of the contribution to the T value from different regions of phase space. This is physically motivated, as a large contribution to the *CP* violation is expected to arise from the main resonances existing in the decay. The events corresponding to such resonances would be localized in the phase space and could be seen in 1D or 2D projections of the multi-dimensional phase space.

In order to visualize the asymmetries one needs to calculate the contribution to the T value from each individual event. This quantity, called T_i and $\overline{T}_i$ for different flavour events, is the sum of all pairwise distances involving a specific event i. The T_i is expressed as follows:

$$T_i = \frac{1}{2n\,(n-1)} \sum_{j \neq i}^{n} \psi\left(d_{ij}\right) - \frac{1}{2n\bar{n}} \sum_{j}^{\bar{n}} \psi\left(d_{ij}\right). \tag{4.5}$$

$$\bar{T}_i = \frac{1}{2\bar{n}\,(\bar{n}-1)} \sum_{j \neq i}^{\bar{n}} \psi\left(d_{ij}\right) - \frac{1}{2n\bar{n}} \sum_{j}^{n} \psi\left(d_{ij}\right). \tag{4.6}$$

the first term in both expressions is the sum over distances from the event i to the events of the same flavour and the second term is the sum over distances to the opposite flavoured events. The total T value is simply $T = \sum_i T_i + \sum_i \overline{T}_i$.

A unique significance can be assigned to each event by comparing its T_i value against the distribution of all T_i values as shown in Fig. 4.4. It is advisable to only consider positive T_i, related to the amount of asymmetry coming from particular events.

Once the contribution to total significance is assigned to each event, it is possible to visualize these contributing regions by plotting projections of the phase space, as shown in Fig. 4.5. It is then possible to infer which resonances contributed the most, to the result. This is an important input, when trying to obtain a full amplitude description of the decay.

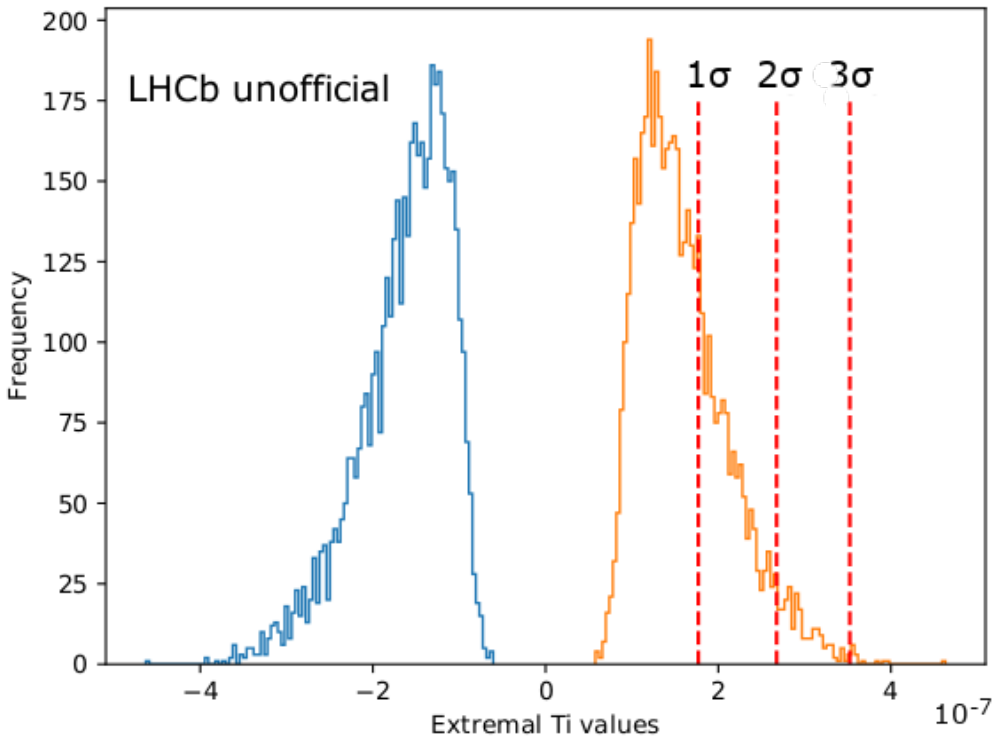


Figure 4.4: The distribution of minimal and maximal T_i values for a particular decay. Red lines show $1\sigma, 2\sigma$ and 3σ significance limits.

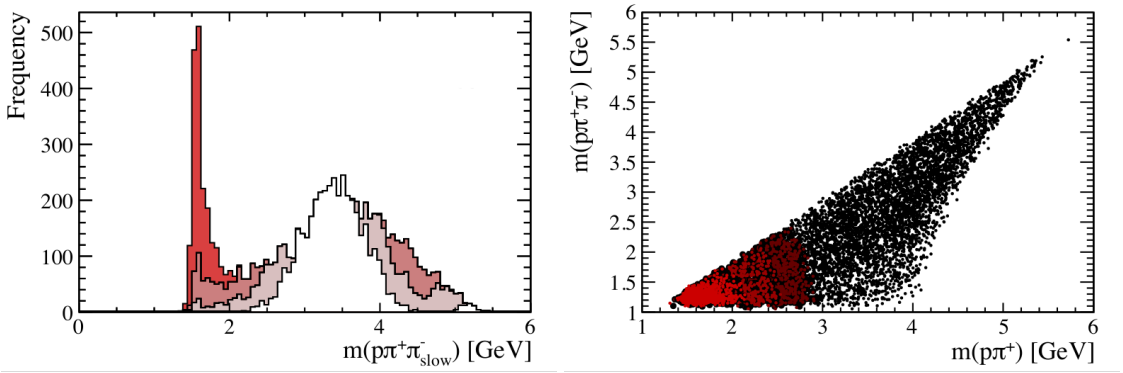


Figure 4.5: Phase space projections showing regions of different contribution to the overall significance using simulated data with *CP* violation included in via a specific resonance. The 1D mass projection (Left) show $1\sigma, 2\sigma$ and 3σ regions visualized as increasingly darker shades of red. The 2D mass plot (Right) show a cluster of event that with large contribution to the overall significance.

4.3.4 Background Subtraction Technique

In certain particle decays, *CP* violation effects might be small. In order to observe such effects, the methods used should be sensitive to small variations across the phase space. This also causes a complication. The method can become sensitive to the effects mimicking *CP* violation. For example, the dataset might have a significant background component, coming from partially reconstructed decays or misidentified particles. Such background reduces the purity of the sample. Furthermore, it is possible, that the background fraction will be different for matter and antimatter events. This can be caused by global detection or production asymmetries. When searching for *CP* violation in two samples of different purities, it is possible to overestimate the result, leading to a false claim of the discovery of asymmetry.

The energy test also requires different sample purities to be taken into account. When calculating *CP* symmetric permutations, the difference in purities is not reproduced. Hence, the distribution of T values for the permutations will not represent the null hypothesis T distribution of real data. To remedy this, a special background subtraction technique is proposed in Ref. [131].

The technique uses a representative background sample (e.g. high mass side band) to subtract background contributions to the T value. The T value expression is changed accordingly. Compared to its original form given in eq. 4.4, the new expression obtains terms subtracting background-background and background-signal contributions

$$\begin{aligned} T = &\frac{1}{2w(w-1)}\Big(\sum_{i}^{n}\sum_{j\neq i}^{n}\psi_{ij} - \frac{2b}{b_s}\sum_{i}^{n}\sum_{j}^{b_s}\psi_{ij} + \frac{b(b+1)}{b_s(b_s-1)}\sum_{i}^{b_s}\sum_{j\neq i}^{b_s}\psi_{ij}\Big) \\ &+ \frac{1}{2\bar{w}(\bar{w}-1)}\Big(\sum_{i}^{\bar{n}}\sum_{j\neq i}^{\bar{n}}\psi_{ij} - \frac{2\bar{b}}{\bar{b}_s}\sum_{i}^{\bar{n}}\sum_{j}^{\bar{b}_s}\psi_{ij} + \frac{\bar{b}(\bar{b}+1)}{\bar{b}_s(\bar{b}_s-1)}\sum_{i}^{\bar{b}_s}\sum_{j\neq i}^{\bar{b}_s}\psi_{ij}\Big) \\ &- \frac{1}{w\bar{w}}\Big(\sum_{i}^{n}\sum_{j}^{\bar{n}}\psi_{ij} - \frac{\bar{b}}{\bar{b}_s}\sum_{i}^{n}\sum_{j}^{\bar{b}_s}\psi_{ij} - \frac{b}{b_s}\sum_{i}^{b_s}\sum_{j}^{\bar{n}}\psi_{ij} + \frac{b\bar{b}}{b_s\bar{b}_s}\sum_{i}^{b_s}\sum_{j}^{\bar{b}_s}\psi_{ij}\Big), \end{aligned} \tag{4.7}$$

where w and $\overline{w}$ are the number of signal events in the main samples, b and $\bar{b}$ are the number of background events in the main samples, while b_s and $\bar{b}_s$ denote the number of background events in the background samples. Also terms with $j < i$ are included, which is balanced by an extra $1/2$ factor for simplicity. If the background sample accurately describes the background contribution to the main sample, this altered expression provides an unbiased T value estimate regardless of the individual sample purities. More details can be found in Ref. [131].

4.3.5 Choice of Coordinates

A decay of a pseudo-scalar particle M into n pseudo-scalar particles $(A, B, C...)$ $M \to ABC...(n)$ can be described by n four-vectors p^{μ}_{A}, p^{μ}_{B}, p^{μ}_{C}..., and consequently $4n$ parameters. The known masses of the identified final state particles $A, B, C...$ remove n degrees of freedom. E, p conservation removes an additional four degrees of freedom. The system

$\Lambda_b^0 \to p\pi^-\pi^+\pi^-$	Two-body masses	Three-body masses
Unphysical	$\pi_s^-\pi_f^-$	
Physical	$p\pi_s^-$, $p\pi_f^-$, $p\pi^+$, $\pi^+\pi_s^-$, $\pi^+\pi_f^-$	$p\pi^+\pi_s^-$, $p\pi^+\pi_f^-$, $p\pi_s^-\pi_f^-$, $\pi^+\pi_s^-\pi_f^-$
Selected	$p\pi^+$, $\pi^+\pi_s^-$, $p\pi_s^-$	$p\pi^+\pi_s^-$, $\pi^+\pi_s^-\pi_f^-$

Table 4.1: Table listing all possible two and three-body invariant mass combinations for the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decay. The selected combinations are used in the energy test as a part of the analysis set out in Chap. 5.

can be rotated freely around all spatial axes, removing a further three degrees of freedom. Hence, $3n-7$ degrees of freedom remain. Consequently a three-body decay phase space is fully described by the two variables conventionally used in Dalitz plot analyses. The phase-space of a four-body decay can be fully described by five parameters [131].

A common choice of coordinates to describe the phase space of a decay are the invariant masses of the decay-product combinations. In a three-body decay, the phase space is fully spanned by two invariant mass pairs. Also the phase space is flat in m^2, making it a regular choice to define the distances between events. However, four-body decays are more complicated. First of all, five variables are required to fully describe the decay phase space. Also, the phase space is no longer flat in m^2, so this choice is not necessarily optimal.

In a four-body decay, one can construct six invariants s_{ij} for each of the decay product pairs. These are defined as $s_{ij} = (p_i + p_j)^2 = m_i^2 + m_j^2 + 2(p_i p_j)$. Linear combinations of these can be used to construct three body combinations of decay products $(p_i + p_j + p_k)^2 = s_{ij} + s_{ik} + s_{jk} - m_i^2 - m_j^2 - m_k^2$. This gives a total choice of 10 invariant mass combinations. An example of such combinations for a specific four-body decay $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ can be seen in Table 4.1. Analysis of these decays is described in detail in Chap. 5.

Another complication common to multi-body decays is the presence of identical particles in the final state. In the case of $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays, two negative pions are in the final state. These need to be uniquely defined, otherwise the sensitivity to P-even CP violation vanishes, while the sensitivity to P-odd CP violation might be drastically reduced. In the CP violation search analysis outlined in Chap. 5, the same-sign pions are ordered by their momenta in the Λ_b rest frame. The one with the higher momentum is labelled "fast pion" (π_f^-), while the other one is labelled "slow pion" (π_s^-), as can be seen in Table 4.1.

The non-physical combination refers to the two-body doubly-charged meson decay, where no known resonances have this topology. The specific selection of five invariant mass combinations is motivated by the dominant resonances in this channel. E.g. $p\pi^+$ is the doubly-charged combination, where the peaking structure can clearly be seen in Fig. 4.6, which arises due to Δ^{++} resonance contribution. In a similar way $p\pi_s^-$, $p\pi^+\pi_s^-$ should pick out other Δ and N^* resonances, while $\pi^+\pi_s^-$, $\pi^+\pi_s^-\pi_f^-$ should contain mesonic a_1

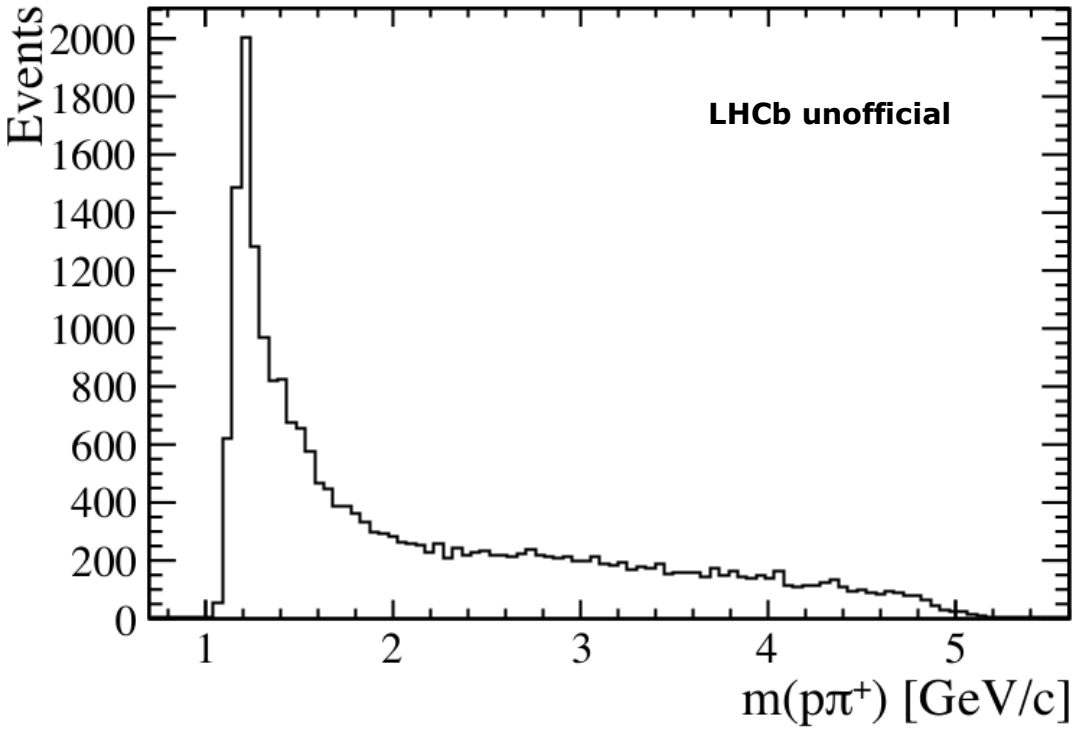


Figure 4.6: Two-body invariant mass distribution for $p\pi^+$ candidates reconstructed by the LHCb detector, used as one of the coordinates to define the phase space of $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decay. A peak at 1232GeV/c^2 can be seen, which corresponds to the Δ^{++} resonance contribution [82]. This mass combination can be used to measure distances between events in the energy test. Reproduced from [14].

and ρ contributions.

It was decided to use the invariant mass squares of the selected combinations to calculate the distances between events for the energy test. This is a common choice in three-body analyses and is a familiar metric to many analysts. However, sensitivity studies were done exploring the use of invariant masses or a combination of helicity angles and masses. No significant change in sensitivity was observed.

4.3.6 Different Versions of Energy Test

It is possible to test the samples for *CP* violation by comparing the data split by the event flavour. Using invariant masses, their squares, or a combination of masses and helicity angles as the phase space coordinates, allows the samples to be tested for *P*-even *CP* violation, as all of these variables are even under parity transformation. The *P*-even decay amplitudes are proportional to the sine of the strong-phase difference, while *P*-odd decay amplitudes are proportional to the cosine of the strong phase difference [94]. *P*-odd *CP* violation arises due to the interference of *P*-even and *P*-odd decay amplitudes. In order to test for *P*-odd *CP* violation, one needs to compare parity odd quantities. An example of such a quantity is the triple product of final state momenta $C_{\hat{T}} \equiv \vec{p}_A \cdot (\vec{p}_B \times \vec{p}_C)$ (the same as used in Sec. 4.2). Here, $p_{A,B,C}$ are the momenta of the final state particles in the parent

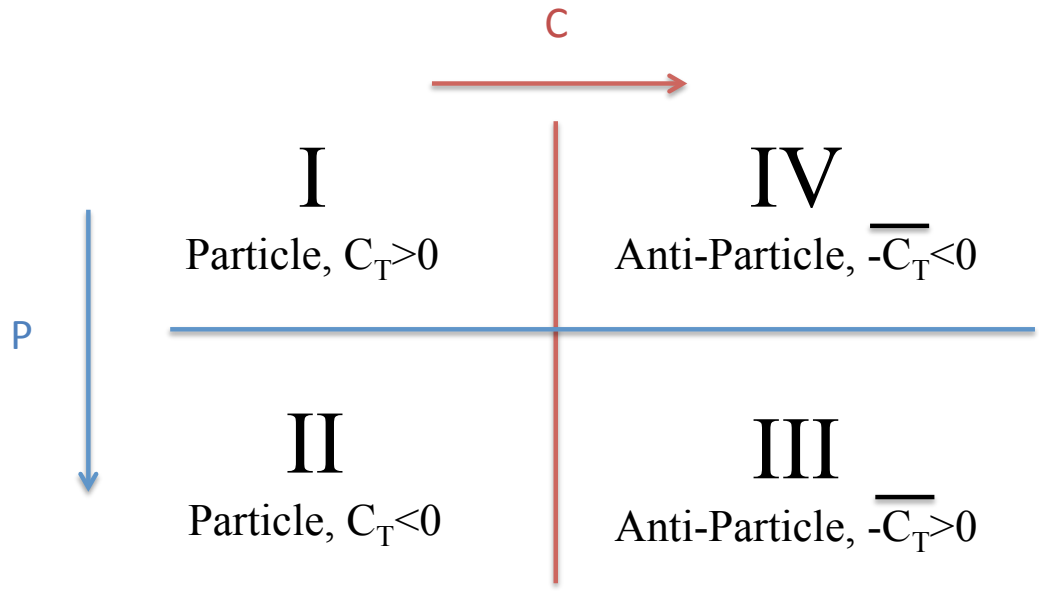


Figure 4.7: Symmetry transformation relationships of the four data subsamples. Samples [I] and [III] are related by *CP* transformation, so are samples [II] and [IV]. Reproduced from [131].

particle rest frame. Applying a *CP* transformation on the triple product changes its sign: $CP(C_T) = -C(C_T) = -\overline{C}_T$. Hence, the corresponding triple product for an antimatter event will have the opposite sign. This allows the full dataset to be split into four subsets, using the flavor of the event and the sign of the triple product. These subsamples are related by C and P transformations, as shown in Fig. 4.7.

$$[I]\ \Lambda_b(C_T > 0),\ [II]\ \Lambda_b(C_T < 0),\ [III]\ \overline{\Lambda}_b(-\overline{C}_T > 0),\ [IV]\ \overline{\Lambda}_b(-\overline{C}_T < 0) \tag{4.8}$$

It is then possible to apply the energy test in a couple of different versions. This is achieved by combining relevant subsamples listed in eq. 4.8. For example, combining subsamples [I + II] and comparing against the combination [III + IV] constitutes the *P*-even *CP* violation test as this is simply comparing all matter events versus all antimatter events. In order to test for *P*-odd *CP* violation, one needs to compare combinations of samples [I + IV] versus [II + III]. This compares all events with the same C_T sign against opposite signed events. In the *P*-odd version of the energy test, the contribution from *P*-even *CP* violation cancels, hence these tests can be considered orthogonal and can be applied on the same dataset.

One caveat of the *P*-odd energy test version is the fact that an extra step needs to be added in the sample selection. The test statistic T (eq. 4.4) of the energy test contains normalization terms that cancel out the overall global asymmetry in the two samples, that might occur due to detection or production effects. However, combining data subsamples by their C_T sign mixes matter and antimatter events. This means that the global asymmetries might not cancel and, in the worst case, get enhanced. Such an effect would mimic

CP violation and could lead to a false claim of discovery. In order to counteract this, before splitting the events by C_T sign the matter and antimatter subsamples are equalized by randomly discarding events from the bigger sample, until the samples have equal yields. Only then the events are split further by the C_T sign and *P*-odd version of the energy test is carried out. The *P*-odd energy test was applied for the first time in Ref. [74].

Finally, it is also possible to compare samples [I + III] versus [II + IV], which constitutes a parity violation test. The subsamples I and III as well as the subsamples II and IV are related through *CP* transformation. The comparisons of the subsamples [I + III] and [II + IV] are essentially checking two instances of parity transformation (I vs II) and (III vs IV) and hence tests for parity violation. This version of the energy test is applied for the first time in the analysis set out in Chap. 5 and yields an observation of parity violation in beauty baryons.

The ability to run the energy test in different configurations is a unique property of the test. For example, the triple product asymmetries method, described in Sec. 4.2 is only sensitive to *P*-odd *CP* violation, while the energy test can be run in independent *P*-even and *P*-odd versions, as well as to test for *P* violation.

4.4 Conclusions

Three different methods for *CP* violation searches have been introduced. Methods sensitive to *P*-even and *P*-odd types of *CP* violation were discussed. The use of *P*-odd (*T*-odd) triple product variables was described. A self contained description of the energy test method was given, introducing new developments, such as the scaling property, suggestions for the choice of distance parameter δ and the option to test for parity violation. The choice of coordinates to define the decay phase space was discussed.

Chapter 5

Search for CP violation in $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays

This chapter describes the search for CP violation in the baryonic $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays. The analysis uses two methods: the energy test and the Triple Product Asymmetries, described in Chap. 4. This work is based on the publication in Ref. [68] and the corresponding internal LHCb document [14]. Section 5.1 gives the motivation for this analysis and discusses the previous measurement performed on this channel. Sections 5.2, 5.3 and 5.4 describe the data used for the analysis, as well as its modelling and simulation. This is followed by Section 5.5, describing the sensitivity studies. Sections 5.6 and 5.7 describe the estimation of uncertainties and cross-checks carried out. The results for both methods are given in Section 5.8.

5.1 Introduction and Previous Measurements

CP violation is well established in the Kaon [25] and Beauty [9, 36] systems. Recently, LHCb observed CP violation in the Charm system [57]. However, all of these observations have been made in meson decays. To date, no CP violation has been observed in the baryonic systems. Furthermore, the Standard Model predicts asymmetries of up to 20% in certain b-baryon decays [113]. Also, CP violation in baryons is one of the Sakharov conditions for baryogenesis [138], required to explain the matter-antimatter asymmetry in the Universe, as described in Chap. 1. Analysis of these decays may lead to the first observation of CP violation in baryons, thus refining the CKM mechanism further or leading to a better understanding of other sources of CP violation.

The search for CP violation in $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ charmless beauty baryon decays is described below. The $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays proceed via tree $b \to ud\bar{u}$ and penguin $b \to du\bar{u}$ transitions as shown in Fig. 5.1. Both of these topologies are proportional to the Wolfenstein CKM parameter λ^3. Consequently, the diagrams will contribute with a similar magnitude, and high interference effects between them could be expected. This may give rise to the CP violation in these decays. Furthermore, these decays have a rich resonance structure, enhancing the sensitivity to CP violation. Finally, Beyond Standard

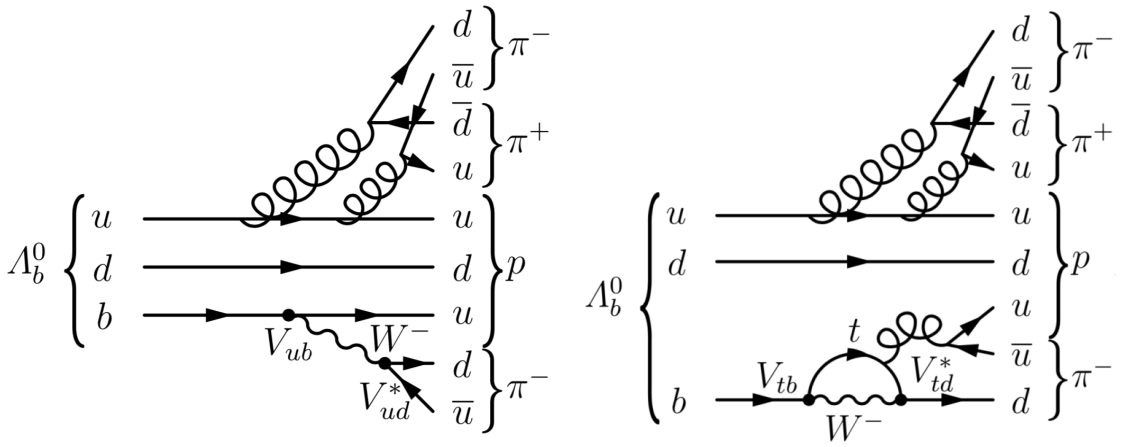


Figure 5.1: Tree (left) and penguin (right) diagrams for $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays. The tree and penguin diagrams have similar magnitude and *CP*-violating effects can potentially arise from the interference of amplitudes with different weak phases.

Model particles might contribute to the loop diagrams and give access to indirect New Physics measurements.

A previous measurement of this channel has been performed on the LHCb Run 1 data using Triple Product Asymmetries method [53]. It yielded the first evidence of *CP* violation in baryons with 3.3σ significance. The analysis used two binning schemes: one motivated by the presence of expected resonances in these decays ($\Delta(1232)^{++}$ and $\rho(770)^0$), the other using a homogeneous binning in the angle between the decay planes - $|\Phi|$ (discussed in Sec. 5.5.1). Using the second binning scheme a 3.5σ deviation from *CP* conservation hypothesis was observed, while combining both binning schemes gives the reported result of 3.3σ. It was dominated by the fourth angular bin, as can be seen in the Fig. 5.2.

The new analysis reported here uses a larger dataset. It repeats the previous measurement using the Triple Product Asymmetries with the binning where the effect was found. Furthermore, two new binning schemes for the Triple Product Asymmetries method are used. In addition, the new analysis adds the use of the energy test method (described in Sec. 4.3). This chapter describes both techniques but particularly concentrates on the energy test. This method has not been applied on baryonic decay measurements in the past. It allows searches of both *P*-even and *P*-odd *CP* violation, whereas the Triple Product Asymmetries are only sensitive to *P*-odd amplitudes (different types of *CP* violation are described in Sec. 1.2.4).

During Run 2 the LHCb detector was able to collect a dataset of $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays which is four times larger than the Run 1 data set used in the previous analysis. If the central value of the effect observed in Run 1 does not change, the increased signal even yield would reduce the statistical uncertainty and lead to an expected significance of the measurement above the 5σ discovery threshold.

Both the energy test and the Triple Product Asymmetries are "observation tools" providing a *p*-value which corresponds to the probability of the measurement being compatible with *CP* conservation hypothesis. In the case of a significant deviation from

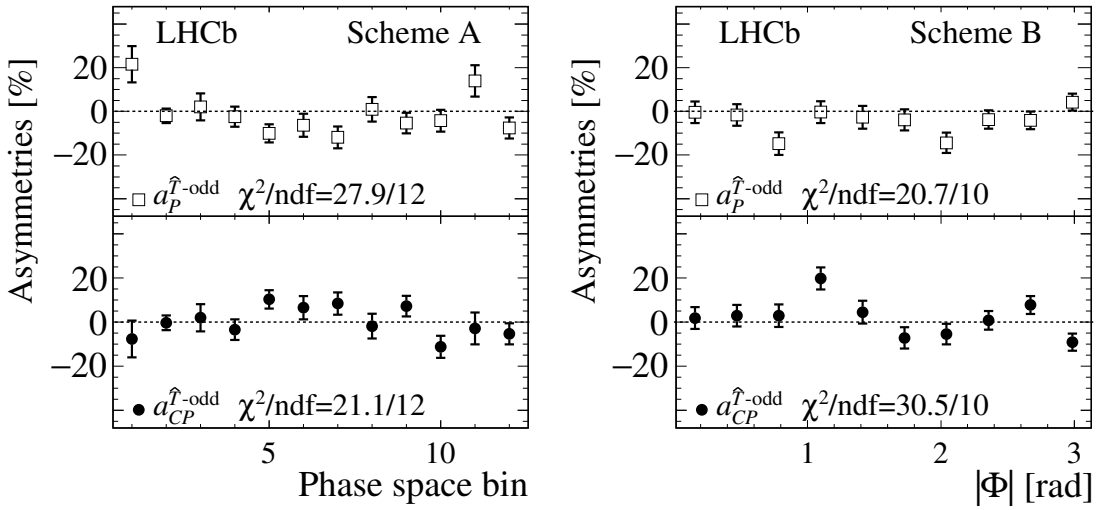


Figure 5.2: The Triple Product Asymmetries results for two binning schemes (A and B) for the previously published LHCb Run 1 analysis. The *P*- and *CP*-violating asymmetries are shown as boxes and circles respectively. The error bars represent one standard deviation, calculated as the sum in quadrature of the statistical uncertainties. The values of the χ^2/ndf are quoted for the *P*- and *CP*-conserving hypotheses for each binning scheme. Reproduced from Ref. [53].

this hypothesis further studies would have to be performed in order to understand the source and nature of the *CP* violation. For example, it would be beneficial to visualize the phase space regions where the biggest contributions to the *CP* violation has been observed. This is a natural and simple extension of the energy test method, described in Sec. 4.3.3. Furthermore, a full amplitude analysis (described in Sec. 1.3) would be invaluable in a complete understanding of the underlying mechanisms responsible for the *CP* violation in this channel.

5.2 Data Selection

This analysis uses data collected by the LHCb detector during Run 1 (2011-2012) and Run 2 (2015-2017) data taking periods at the centre-of-mass energies of 7 TeV (2011), 8 TeV (2012) and 13 TeV (2015, 2016 and 2017) corresponding to 6.9 fb^{-1} of integrated luminosity. In this section, various data selection requirements are listed, followed by the descriptions of control and background datasets.

5.2.1 Online Selection

During the proton-proton collisions, the particle multiplicity inside the LHCb detector is very high [124]. A sizeable fraction of primary vertices have more than 20 charged tracks. Some of these tracks do not originate from the signal decay and thus constitute different backgrounds. For this reason, the signal candidates need to be selected, which is achieved by using the LHCb trigger system (described in Sec. 2.2.6). The signal events have to pass

multiple trigger requirements. The hardware trigger (L0) requires the energy deposited in the hadronic calorimeter to be above 3.5 GeV [124]. If this condition is passed, the first software trigger (HTL1) performs minimal track reconstruction and applies further requirements on the transverse momentum and impact parameter of the track candidate. This ensures that high momentum tracks with good separation from the primary vertex are selected. The selected events then undergo full reconstruction in the second software trigger (HLT2), which adds extra constraints based on b-decay topology. The events that pass the trigger system requirements are saved to permanent storage.

The signal candidates used in this analysis are selected from the LHCb storage using centralised *stripping* campaigns [1]. A set of selection criteria is applied to the data. These criteria are optimized to pick out $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decay candidates by assuring that a good quality of the reconstructed primary vertex and final state particle tracks is obtained. The requirements are listed in Table 5.2.

This analysis uses charmless four-body $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays. The charm resonances existing in these decays need to be removed by applying veto cuts. The resonances will appear as peaks in the the reconstructed invariant mass distributions of different combinations of the Λ_b^0 decay products. These peaks are fitted with the corresponding probability density functions and the events inside a 3σ band around the resonance peaks are vetoed. Vetoes are applied on long-lived charmed particles: $\Lambda_c^+, D^0, D^+, D_s^+$.

Furthermore, the events are removed if the Λ_b^0 candidate is reconstructed under mis-ID hypothesis of K or π (e.g. assigning $K(\pi)$ mass to the $p(K)$ track candidate). Poorly reconstructed charmed resonances due to $\pi \to \mu$ mis-ID, such as J/ψ are removed as well. The cross-feed background from B^0 and B_S^0 decays is reduced by vetoing events with ϕ and $K^{*0}(892)$ resonances, where the proton is misidentified as a kaon or a pion. No long lived particles with light quarks that have the flight distance of $c\tau \approx \mathcal{O}(1\text{cm})$, such as Λ or K_s^0, are reconstructed. This is done by applying a vertex cut $\chi^2_{vtx} < 20$ in the stripping line selection. Strongly decaying resonances are part of the signal. The vetoed resonances are listed in Table 5.1.

Table 5.1: List of vetoed resonances and the corresponding vetoed mass window. The misidentified particles are highlighted. Rejection fraction for $b \to c$ resonances is reported. Reproduced from [14].

Resonance	Veto region [GeV/c^2]	Rejection fraction
$\Lambda_c^+ \to p\pi^+\pi^-_{slow}$	2.253-2.307	2.6%
$\Lambda_c^+ \to p\pi^+\pi^-_{fast}$	2.255-2.304	
$D^0 \to \pi^+\pi^-_{slow}$	1.826-1.881	2.6%
$D^0 \to \pi^+\pi^-_{fast}$	1.842-1.873	
$J/\psi \to \boldsymbol{\mu}^+\boldsymbol{\mu}^-$ ($\mu \to \pi^+, \pi^-_{fast}$)	3.061-3.129	2.0%
$J/\psi \to \boldsymbol{\mu}^+\boldsymbol{\mu}^-$ ($\mu \to \pi^+, \pi^-_{slow}$)	3.054-3.139	
$\Lambda_c^+ \to p\mathbf{K}^-\pi^+$ ($K^- \to \pi^-_{fast}$)	2.24-2.31	3.2%
$\Lambda_c^+ \to p\mathbf{K}^-\pi^+$ ($K^- \to \pi^-_{slow}$)	2.24-2.31	
$K^{*0}(892) \to \mathbf{K}^+\pi^-_{fast}$ ($K^+ \to p$)	0.84-0.96	-
$K^{*0}(892) \to \mathbf{K}^+\pi^-_{slow}$ ($K^+ \to p$)	0.84-0.96	-

[1] *StrippingXb2phhhLine* in the BHADRON stream is used

Table 5.2: Stripping selection criteria for `StrippingXb2phhhLine` stripping line. Reproduced from [14].

	Selection criteria	**Variable definition**	**Accepted values**
proton	p_{T}	Transverse momentum	$> 250\,\mathrm{MeV}/c$
	p	Total momentum	$> 1.5\,\mathrm{GeV}/c$
	χ^2_{IP} (PV)	Difference in the vertex-fit χ^2 of a PV reconstructed with and without the considered track	> 16
	χ^2/ndf	Track fit quality	< 3
	Ghost. Prob.	Probability, that the reconstructed track does not correspond to a physical particle	< 0.4
	ProbNNp	Track probability being associated to a proton based on a neural network response of RICH data	> 0.05
π	p_{T}	Transverse momentum	$> 250\,\mathrm{MeV}/c$
	p	Total momentum	$> 1.5\,\mathrm{GeV}/c$
	χ^2_{IP} (PV)	Difference in the vertex-fit χ^2 of a PV reconstructed with and without the considered track	> 16
	χ^2/ndf	Track fit quality	< 3
	Ghost. Prob.	Probability, that the reconstructed track does not correspond to physical particle	< 0.4
Comb. cut	m($p\pi\pi\pi$)	Four-body invariant mass	$< 6.405\,\mathrm{GeV}/c^2$
	m($pKKK$)	Four-body invariant mass	$> 5.195\,\mathrm{GeV}/c^2$
	$\sum_{daug.} p_{\mathrm{T}}$	Sum of transverse momenta of final state particles (daughters)	$> 3.5\,\mathrm{GeV}/c$
	p_{T}	Transverse momenta	$> 1.5\,\mathrm{GeV}/c$
	χ^2_{DOCA}	Distance of closest approach χ^2 of any two final state particles	< 20
Λ_b^0	χ^2_{vtx}	Λ_b^0 candidate's vertex χ^2	< 20
	$\chi^2_{VD}(PV)$	Λ_b^0 candidate's flight distance χ^2 w.r.t. best PV	> 50
	χ^2_{IP} (PV)	Λ_b^0 candidate's impact parameter χ^2 w.r.t. best PV	< 16
	cos(DIRA)	Cosine of the Λ_b^0 candidate's pointing angle	> 0.9999

5.2.2 Offline Selection

The offline selection uses data driven methods and is not reliant on simulation samples. The Run 1 selection uses identical criteria as in the previous analysis of this channel [53]. The Run 2 selection is outlined below.

After the veto cuts (see Table 5.1) the selection procedure employs Multivariate Analysis (MVA) techniques to obtain better signal and background separation. This is done using the TMVA package [112]. A kinematically similar high-yield data sample of $\Lambda_b^0 \to pK^-\pi^+\pi^-$ decays is used for training the Boosted Decision Tree (BDT) classifier [19]. A sample of background events is obtained from the high-mass sideband (HMSB) region of $5.85\,\text{GeV}/c^2 < m(pK^-\pi^+\pi^-) < 6.4\,\text{GeV}/c^2$ on the $\Lambda_b^0 \to pK^-\pi^+\pi^-$ invariant mass spectrum and is used for background training. An additional veto cut is applied on the B_S^0 masses to remove $B^0_{(s)} \to h^+h^-h^+h^-$ background with h^+ misidentified as a proton since this background could mimic signal kinematic distributions.

A preliminary particle identification cut is applied on all tracks, $PID_h > 0.2$ where h can be a proton, kaon, pion or muon. The PID variable for a proton is defined as

$$PID_p = ProbNNp \cdot (1 - ProbNNK) \cdot (1 - ProbNN\pi) \cdot (1 - ProbNN\mu), \tag{5.1}$$

where $ProbNN_{p,K,\pi,\mu}$ is the probability obtained from a neural network response for a specific particle type based on the information from the RICH detectors. Equivalent definitions exist for kaon, pion and muon. The preselected data is fitted using a fit model described in Sec. 5.3. The signal component is estimated from the fit using the *sPlot* technique [135]. The extracted signal and background distributions are used to train the BDT.

The BDT works by taking a number of input variables and creating a list of binary decisions (e.g. cuts on kinematic quantities). Each decision is assigned a node in a tree-graph. The algorithm searches for a set of nodes that optimize a desired outcome. In High Energy Physics this is usually defined by a figure-of-merit which quantifies the separation between signal and background. Furthermore, the BDT non-trivially combines the input variables into a single variable, referred to as the BDT response or BDT output. The specific set of nodes (a path in the decision tree-graph) correspond to a unique value of this BDT output variable. Hence, applying a cut on the BDT output allows optimal signal and background separation to be obtained.

The BDT algorithm requires some input data to optimize the output. This is called BDT training. Usually the real data is randomly split into three subsamples of equal size and used to train the BDT classifier. The first subsample is used for training and the second for validating the selection which is then applied on the third subsample to provide the desired signal-background discrimination. This is repeated by changing the functions of the three subsamples. This way the training uses real data and introduces no bias. The variables used for the BDT training are ranked by counting the frequency of their use in the BDT nodes and weighting each node-split by the square of the separation gain

achieved and number of events in the node. The list of BDT input variables and their significances is given in Table 5.3.

The BDT selection yields 90% of signal efficiency and 85% of background rejection when optimizing for the usual figure-of-merit: $S/\sqrt{S+B}$. The optimal significance is estimated by calculating the signal and background efficiency (ϵ_S, ϵ_B). The corresponding yields $S = \epsilon_S S_0$ and $B = \epsilon_B B_0$ are determined using the fit results before the BDT cuts, where S_0 and B_0 are estimated from the invariant mass distribution in the $[5.5 - 5.7]\,\mathrm{GeV}/c^2$ mass window. The optimal BDT cut is found: $BDT_{output} > 0.268$.

After the BDT selection, PID cuts are applied in order to suppress combinatorial and cross-feed backgrounds $\Lambda_b^0 \to pK^-\pi^+\pi^-$ and $B^0 \to K^+\pi^-\pi^+\pi^-$. These backgrounds are kinematically similar to the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ signal. The PID efficiencies are determined from the vetoed resonance sample $\Lambda_b^0 \to \Lambda_c^+(\to p\pi^+\pi^-{}_{slow})\pi^-{}_{fast}$. This allows the significance $\mathcal{S}$ and purity $\mathcal{P}$ for the signal data samples to be calculated. The relevant expressions are given in Eq. 5.2

$$\mathcal{S} = \frac{\varepsilon_S(\mathrm{PID}) \cdot N_S}{\sqrt{\varepsilon_S(\mathrm{PID}) \cdot N_S + \varepsilon_B(\mathrm{PID}) N_B}}, \tag{5.2}$$

$$\mathcal{P} = \frac{\varepsilon_S(\mathrm{PID}) \cdot N_S}{\varepsilon_S(\mathrm{PID}) \cdot N_S + \varepsilon_B(\mathrm{PID}) \cdot N_B},$$

where $\varepsilon_S(\mathrm{PID})$ and $\varepsilon_B(\mathrm{PID})$ are the efficiencies for signal and background respectively, for a specific PID cut. N_S and N_B are the numbers of signal and background events obtained from the signal fits after the BDT cuts. These variables are calculated in the Λ_b^0 mass window of $[5.5, 5.7]\,\mathrm{GeV}/c^2$. This yields the final PID selection of $PID_p > 0.3$ and $PID_\pi > 0.4$. Further details on the data selection can be found in Refs. [125, 14].

Table 5.3: Discriminating variables used in the BDT. Reproduced from [14].

Particle	Variable	Description
Λ_b^0	Lb_Cone_PTAsym_1	$\frac{p_{T\,\Lambda_b^0} - \sum p_{\mathrm{T}}}{p_{T\,\Lambda_b^0} + \sum p_{\mathrm{T}}}$, where $p_{T\,\Lambda_b^0}$ is the Λ_b^0 transverse momentum and $\sum p_{\mathrm{T}}$ is the sum of all charged long tracks p_{T} in a cone around the Λ_b^0 candidate with a radius $R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 1$, where η is the pseudorapidity and ϕ is the azimuthal angle
	Lb_LogIPChi2	$\log(\chi^2_{IP})$, χ^2_{IP} is the χ^2 difference between primary vertex fits when the Λ_b^0 candidate is added or excluded from the fit
	Lb_LogACosDIRA_OWNPV	$\log\left(\arccos(\mathrm{DIRA})\right)$, DIRA is the cosine of the direction angle, *i.e.* the angle between the particle momentum and the vector from the primary to the secondary vertex
	DTF_VCHI2NDOF	$\chi^2_{\mathrm{vtx}}/\mathrm{ndf}$ is the χ^2 of the vertex fit per number of degrees of freedom
p	proton_pt	transverse momentum of the proton
	proton_pz	momentum of the proton along the beam line
$h^-h^+h^-$	Log_hm_IP_OWNPV	$\log(\mathrm{IP}_{\mathrm{OWNPV}})$, $\mathrm{IP}_{\mathrm{OWNPV}}$ is the impact parameter of the track with respect the primary vertex
	Log_hm2_IP_OWNPV	
	Log_hp_IP_OWNPV	

	Variable	Relevance
	proton_pt	2.0×10^{-1}
	Lb_LogACosDIRA_OWNPV	1.8×10^{-1}
	proton_pz	1.1×10^{-1}
	Log_hm2_IP_OWNPV	1.1×10^{-1}
	DTF_VCHI2NDOF	9.5×10^{-2}
	Log_hp_IP_OWNPV	9.4×10^{-2}
	Lb_Cone_PTAsym_1	8.4×10^{-2}
	Log_hm_IP_OWNPV	8.0×10^{-2}
	Lb_LogIPChi2	5.8×10^{-2}

5.2.3 $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$ Control Channel

A sample of $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$ decays is selected and used as a control channel for systematic uncertainty studies and cross-checks (described in Sec. 5.6 and Sec. 5.7). Additional selection criteria are applied on the stripping line. These are listed in Table 5.4.

Table 5.4: Additional selection criteria for the $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$ control channel. The criteria include Neural Network (NN) responses for different particle types. Reproduced from [14].

variables	Definition	values
PID_p	NN response for the particle being a proton	> 0.05
PID_K	NN response for the particle being a kaon	> 0.05
PID_π	NN response for the particle being a pion	> 0.05
$\text{m}(\Lambda_c^+)$	Invariant mass of Λ_c^+ candidate	$m(pK^-\pi^+) \in [2260, 2300]$ MeV

The $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$ decays proceed through subsequent $b \to c$ and $c \to s$ transitions. Due to this double transition, the *CP* violation in this channel is expected to be negligible in the Standard Model. Hence, this channel can be used to test for possible experimental bias and in other control checks. The triple product for the cross-channel is defined using the Λ_c^+ decay product momenta in the Λ_b^0 rest frame $C_{\hat{T}} \equiv \vec{p}_p \cdot (\vec{p}_{K^-} \times \vec{p}_{\pi^-})$ and analogously for $\overline{C}_{\hat{T}}$.

5.2.4 Selection for the Energy Test

The presence of background dilutes the sensitivity of the energy test method. Also, in the case of a significant background fraction and if the background distributions are different for matter and antimatter samples a fake *CP* asymmetry might be measured, as described in Sec. 4.3.4. To reduce the presence of background, only events from a tighter mass window are considered. Different mass regions are studied and listed in Table 5.5. The chosen mass window corresponds to a 2.5σ band around the Λ_b^0 mass peak and maximizes the product of efficiency and purity.

This is tested by mixing various proportions of the simulated data, with *CP* violation introduced, and HMSB background events to obtain the efficiency and purity values from Table 5.5. The energy test is applied on these samples and the sensitivity of it is evaluated by "scanning" through various choices of the distance parameter δ as shown in Fig. 5.3. The background fraction in the signal region is small and consistent between matter and antimatter samples. For this reason, the background subtraction technique, described in Sec. 4.3.4, does not need to be employed.

Table 5.5: The efficiency and purity associated with different mass windows, as determined using the combined Run 1 and Run 2 datasets. The first four rows have mass windows given in terms of the resolution of the Λ_b^0 mass peak. The fit model is discussed in Sec. 5.3. Reproduced from [14].

Mass Window [GeV/c^2]	Purity	Efficiency	Efficiency × Purity
5.583-5.654 (2σ)	0.851	0.909	0.774
5.574-5.662 (2.5σ)	0.820	0.947	0.776
5.565-5.671 (3σ)	0.788	0.964	0.760
5.547-5.689 (4σ)	0.731	0.979	0.712

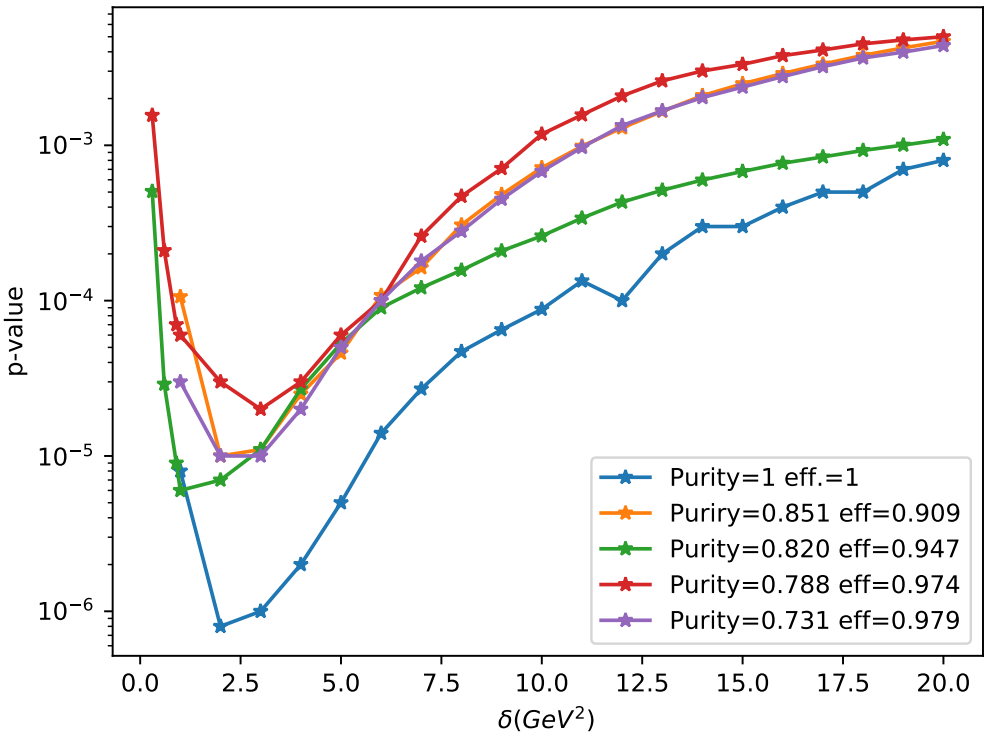


Figure 5.3: The dependence of p-value on the choice of the distance parameter δ in different purity and efficiency scenarios. Simplified simulation data, described in Sec. 5.4, is used. The chosen working point corresponds to the green line.

5.3 Fit Model

The shape of the signal decay invariant mass distribution is determined from the simulation and modelled with a sum of two Crystal Ball (CB) functions [144] with the same mean and width parameters. The signal candidates are truth-matched in the simulation by requiring that the four reconstructed Λ_b^0 decay products are either directly produced from a Λ_b^0 decay or via strong intermediate resonances. The obtained fit parameters are then used to fit the real data sample. The main background arises from three different sources, each of which is modelled separately.

- **Partially reconstructed decays:** produce background which is localized in the low invariant mass region. These decays include $\Lambda_b^0 \to p\pi^+\pi^+\rho^-$, with $\rho^- \to \pi^-\pi^0$, and $\Lambda_b^0 \to p\pi^+\pi^-K^{*-}$, with $K^{*-} \to K^-\pi^0$, where π^0 is not reconstructed. These candidates appear as low-mass shoulder-like shapes in the signal invariant mass distribution.

- **Cross-feed background:** mainly comes from four-body decays of Λ_b^0 or B^0, where one of the decay products has been misidentified. In particular $\Lambda_b^0 \to pK^-\pi^+\pi^-$ and $B^0 \to K^+\pi^-\pi^+\pi^-$ channels, where the kaon is misidentified as a proton or pion, contribute to this background. This background is reduced by using PID cuts and resonance vetoes (e.g. $K^{*+}(892)$ with alternative mass hypothesis for the proton candidate). This background could have a non-zero *CP* asymmetry because it contains *CP*-violating B decay events and has a similar mass distribution to the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ signal.

- **Combinatorial background:** arises due to the random combinations of charged tracks looking similar to the expected signal event. This background is reduced by applying track quality cuts and selecting tracks that are compatible with originating from the displaced vertex of Λ_b^0 decay.

The partially reconstructed background is modelled empirically using an Argus function [5] convolved with a Gaussian resolution function. Cross-feed background components are parametrised with kernel-estimated probability density functions [89] by modelling the invariant mass distributions of simulated $\Lambda_b^0 \to pK^-\pi^+\pi^-$ and $B^0 \to K^+\pi^-\pi^+\pi^-$ decays under the wrong mass hypothesis of the final state particles. The systematic uncertainties of modelling the simulated shapes are taken into account as a fit model systematics, discussed in Sec. 5.6. Finally, the combinatorial background shape ($\mathcal{P}_{comb}$) is modelled with an exponential function as shown in Eq. 5.3.

$$\mathcal{P}_{comb} = A \exp\left(-\lambda_{comb} \cdot m\right). \tag{5.3}$$

where A is a normalization factor and λ_{comb} is a shape parameter that is allowed to vary free in the fit.

The signal and background models are used to obtain an unbinned maximum likelihood fit to the $m(p\pi^-\pi^+\pi^-)$ invariant mass spectrum. The tail of the signal distribution (α and n parameters of the two CB), as well as the fraction of the two CB functions, are fixed to the values obtained from fitting the simulated data, while the μ and σ of the signal Gaussian core are allowed to vary freely.

The B^0 and Λ_b^0 decay cross-feed background yields are estimated from real data. The signal invariant mass spectrum are fitted for $B^0 \to K^+\pi^-\pi^+\pi^-$ and $\Lambda_b^0 \to pK^-\pi^+\pi^-$ candidates by assigning an alternative mass hypothesis to proton or one of the negative pions. This yields 5834 ± 303 events of $\Lambda_b^0 \to pK^-\pi^+\pi^-$ and 2109 ± 154 of $B^0 \to K^+\pi^-\pi^+\pi^-$ cross-feed background events in the signal region.

The signal fits are shown in Fig. 5.4, the relevant parameters are listed in Table 5.6. Analogous fits are performed for the $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$ control channel, yielding 425767 ± 1593 events. Further details on the signal and control channel fit models can be found in Refs. [14, 125].

Table 5.6: The results for the fit to data of $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ candidates. Signal model parameters are fixed in the fit from the simulation results except for the mean μ and the σ value. The coefficients $c_{Part.\ reco}$ and $p_{Part.\ reco}$ refer to the partially reconstructed background shape and control the low mass slope and the exponent of the corresponding function respectively. Reproduced from [14].

Variable	Value
Part. reco bkg	6944 ± 361
Comb. bkg	23746 ± 617
$B^0 \to K^+\pi^-\pi^+\pi^-$	2108 ± 153
$\Lambda_b^0 \to pK^-\pi^+\pi^-$	6414 ± 230
$\Lambda_b^0 \to p\pi^-\pi^+\pi^-$	27630 ± 209
$\mu(\Lambda_b^0)(\mathrm{MeV}/c^2)$	5618.1 ± 0.1
$\sigma(\mathrm{MeV}/c^2)$	17.7 ± 0.1
$c_{Part.\ reco}$	0.8 ± 10.9
$p_{Part.\ reco}$	0.26 ± 0.14
$\lambda_{comb}(c^2/\mathrm{GeV})$	-4.39 ± 0.13
Λ_b^0 signal model	Value (fixed from simulation)
n_1	2.54
α_1	1.48
n_2	2.40
α_2	-1.85
f_{CB_1}	0.48

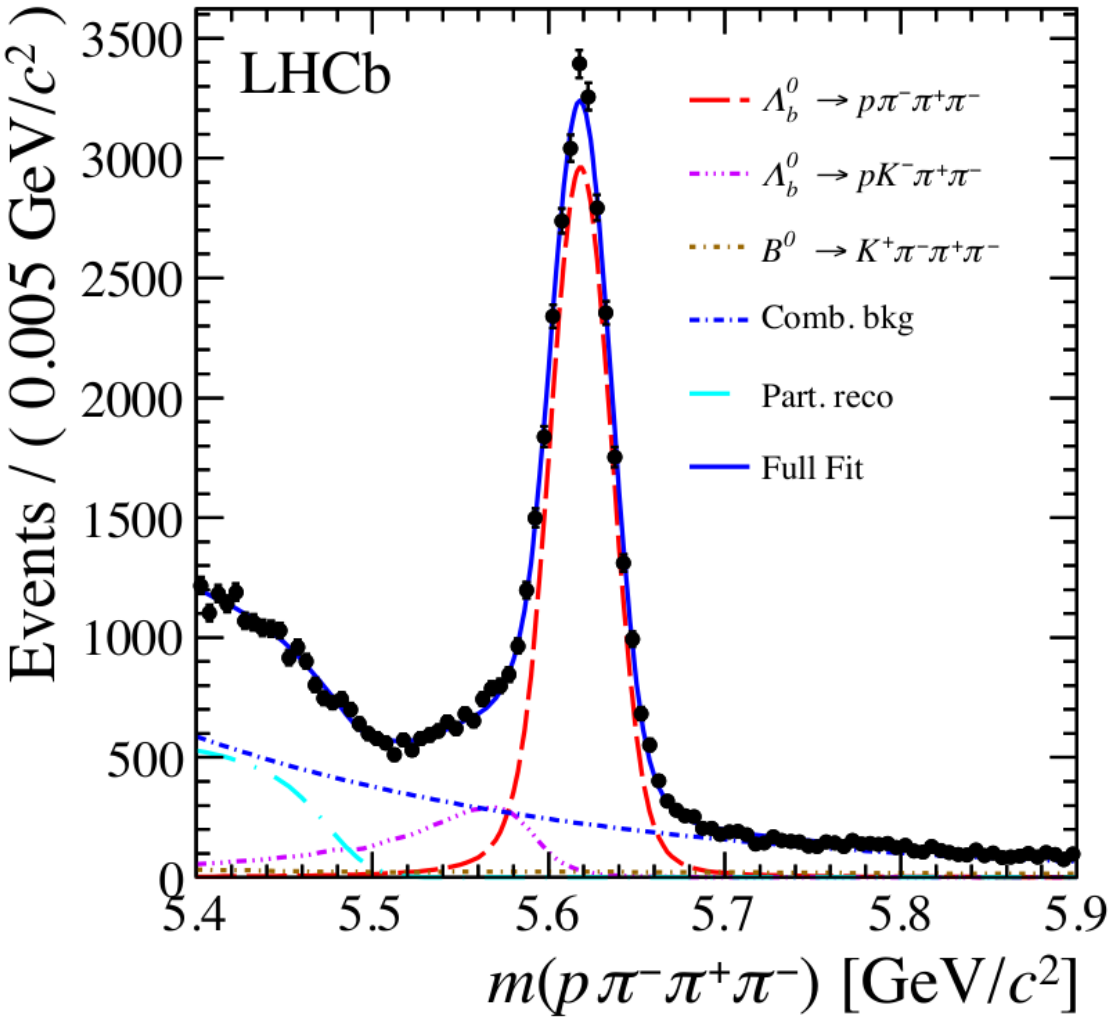


Figure 5.4: Fit to the reconstructed invariant mass distribution for $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ signal candidates using the combined Run 1 and Run 2 data samples. Reproduced from [14].

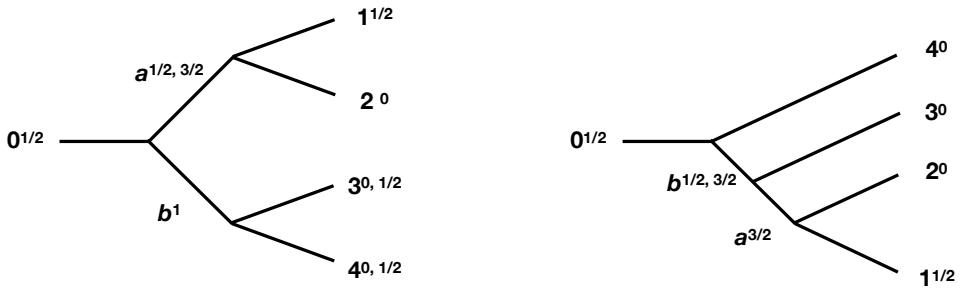


Figure 5.5: The spin-1/2 initial state particle decaying into a four-body final state with quasi-two-body topology (left) and cascade topology (right).

5.4 Simulation

The $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decay is complex and has many contributing resonances. No amplitude model exists for this decay and previous simulation does not match the invariant mass distributions observed in the data. While neither of the methods used in the analysis of this channel have explicit model dependence, a simplified model is developed. This allows simulation samples for sensitivity studies to be created.

The Λ_b^0 baryon can decay into a four-body final state via many different intermediate resonances. Several different decay topologies are available. One such topology is a quasi-two-body decay, where intermediate resonances decay into pairs of final state particles (e.g. $\Lambda_b^0 \to p\pi^-\rho(\to \pi^+\pi^-)$). Another possible topology is a "cascade" of decays. In this case, the initial state particle decays into a pair of stable and unstable decay products, while the unstable resonance subsequently undergoes secondary two-body decay until the set of final state particles emerge (e.g. $\Lambda_b^0 \to \pi^-\Delta^+(\to \pi^-\Delta^{++}(\to \pi^+p))$). This topology is of particular interest because large contributions to P-odd CP violation are expected to arise in the interference of different mass-degenerate spin-states of the main resonances [92]. The examples of quasi-two-body and cascade topologies are shown in Fig. 5.5. The list of resonances that are likely to contribute to the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays is given in Table 5.7.

The topologies outlined above were simulated using RapidSim [88] and TensorFlow [3] packages. RapidSim allows for an easy decay phase space simulation but does not include resonance interference. A cocktail of resonances was chosen to adequately match the data invariant mass distributions. This allows P-even CP violation to be introduced by changing the relative fractions of resonance contributions. The proportions of resonances included are listed below:

- 45% $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ non-resonant decays filling the phase space uniformly
- 23% $\Lambda_b^0 \to \pi^-\Delta^+(\to \pi^-\Delta^{++}(\to \pi^+p))$
- 20% $\Lambda_b^0 \to (a_1^- \to (\rho \to \pi^+\pi^-)\pi^-)p$

Topology	Dominant resonances
$\Lambda_b^0 \to (N^{+*} \to (\Delta^{++} \to p\pi^+)\pi^-)\pi^-$	$N^{+*}(1520)$, $N^{+*}(1535)$, $N^{+*}(1650)$, $N^{+*}(1675)$, $N^{+*}(1680)$, $N^{+*}(1700)$, $N^{+*}(1710)$, $N^{+*}(1720)$, $N^{+*}(1875)$, $N^{+*}(1900)$, $N^{+*}(2190)$, $\Delta^{++}(1232)$
$\Lambda_b^0 \to (N^{+*} \to p(\rho \to \pi^+\pi^-))\pi^-$	$N^{+*}(1720)$, $N^{+*}(1875)$, $N^{+*}(1900)$, $\rho^0(770)$
$\Lambda_b^0 \to (N^{+*} \to p(f^0 \to \pi^+\pi^-))\pi^-$	$N^{+*}(1535)$, $N^{+*}(1650)$, $N^{+*}(1675)$, $N^{+*}(1680)$, $N^{+*}(1700)$, $N^{+*}(1875)$, $N^{+*}(1900)$, $f^0(500)$
$\Lambda_b^0 \to (a_1^- \to (\rho \to \pi^+\pi^-)\pi^-)p$	$a_1^-(1260)$, $\rho^0(770)$
$\Lambda_b^0 \to p\pi^-\pi^+\pi^-$	Non-resonant

Table 5.7: List of dominant resonances and corresponding topologies in the existing in the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decay [82].

Table 5.8: Helicity amplitude decomposition for the $\Lambda_b^0 \to (N^{*+} \to (\Delta^{++} \to p\pi^+)\pi^-)\pi^-$ "cascade" topology decay. Here the ϕ_A and θ_A are the azimuthal and polar angles of the A particle's momentum. Only the $A_\pm$ and $B_\pm$ helicity amplitudes encode the weak decay and could contain the CKM phase. The definitions of the helicity amplitudes $A_\pm$, $B_\pm$ and $b_{i\pm}$ can be found in Ref. [136]. Reproduced from [14].

$\sqrt{2}$	$Re((A_+^*B_+ + A_-^*B_-)(b_{1+}^*b_{3+} + b_{1-}^*b_{3-}^*))$	$(1+3\cos^2\theta_p)$	$\cos\theta_{\Delta^{++}}$	
$1/2$	$(\lvert B_+\rvert^2 + \lvert B_-\rvert^2)(\lvert b_{3+}\rvert^2 + \lvert b_{3-}\rvert^2)$	$(1+3\cos^2\theta_p)$		
$9/4$	$(\lvert A_+\rvert^2 + \lvert A_-\rvert^2)(\lvert b_{2+}\rvert^2 + \lvert b_{2-}\rvert^2)$	$\sin^2\theta_p$	$\sin^2\theta_{\Delta^{++}}$	
$1/4$	$(\lvert A_+\rvert^2 + \lvert A_-\rvert^2)(\lvert b_{1+}\rvert^2 + \lvert b_{1-}\rvert^2)$	$(1+3\cos^2\theta_p)$	$(1+3\cos^2\theta_{\Delta^{++}})$	
$-3\sqrt{2}/2$	$Re((A_+^*B_+ + A_-^*B_-)(b_{2+}^*b_{3+} + b_{2-}^*b_{3-}))$	$\sin 2\theta_p$	$\sin\theta_{\Delta^{++}}$	$\cos\phi_p$
$-3/2$	$(\lvert A_+\rvert^2 + \lvert A_-\rvert^2)Re(b_{1+}^*b_{2+} + b_{1-}^*b_{2-})$	$\sin 2\theta_p$	$\sin 2\theta_{\Delta^{++}}$	$\cos\phi_p$
$3/2$	$(\lvert A_+\rvert^2 + \lvert A_-\rvert^2)Re(b_{1+}^*b_{2-} + b_{1-}^*b_{2+})$	$\sin^2\theta_p$	$\sin^2\theta_{\Delta^{++}}$	$\cos 2\phi_p$
$-3\sqrt{2}/4$	$Im((A_+^*B_+ - A_-^*B_-)(b_{2+}^*b_{3+} + b_{2-}^*b_{3-}))$	$\sin 2\theta_p$	$\sin 2\theta_{\Delta^{++}}$	$\sin\phi_p$
$-3/2$	$(\lvert A_+\rvert^2 - \lvert A_-\rvert^2)Im(b_{1+}^*b_{2+} + b_{1-}^*b_{2-})$	$\sin 2\theta_p$	$(1-3\cos^2\theta_{\Delta^{++}})\sin\theta_{\Delta^{++}}$	$\sin\phi_p$
$3\sqrt{2}/2$	$Im((A_+^*B_- - A_-^*B_-)(b_{2+}^*b_{3-} + b_{2-}^*b_{3+}))$	$\sin^2\theta_p$	$\sin^2\theta_{\Delta^{++}}$	$\sin 2\phi_p$
$-9/4$	$(\lvert A_+\rvert^2 - \lvert A_-\rvert^2)Im(b_{1+}^*b_{2-} + b_{1-}^*b_{2+})$	$\sin^2\theta_p$	$\sin\theta_{\Delta^{++}}\sin 2\theta_{\Delta^{++}}$	$\sin 2\phi_p$

- 12% $\Lambda_b^0 \to p\pi^-\rho(\to \pi^+\pi^-)$

This empirical cocktail reproduces some of the main features in the data but fails to describe finer details of the distributions, as can be seen in Fig. 5.6. While this allows to simulate P-even CP violation it has no access to P-odd amplitudes. This simple simulation generates only the sum of Breit-Wigner resonances, hence a more sophisticated approach is needed. The TensorFlow package is employed as a Monte Carlo generator for simulation, allowing explicit definition of Breit-Wigner line shapes [20] for decaying resonances and full amplitude model of the decay. An example of an amplitude model for the "cascade" topology is provided by the authors of Ref. [92] and given in Table 5.8.

This amplitude model is based on the Helicity formalism [136]. The terms in the model correspond to helicity amplitudes that are connected to different angular distributions of the decay. The model allows interference between spin-1/2 and spin-3/2 N^{*+} resonances

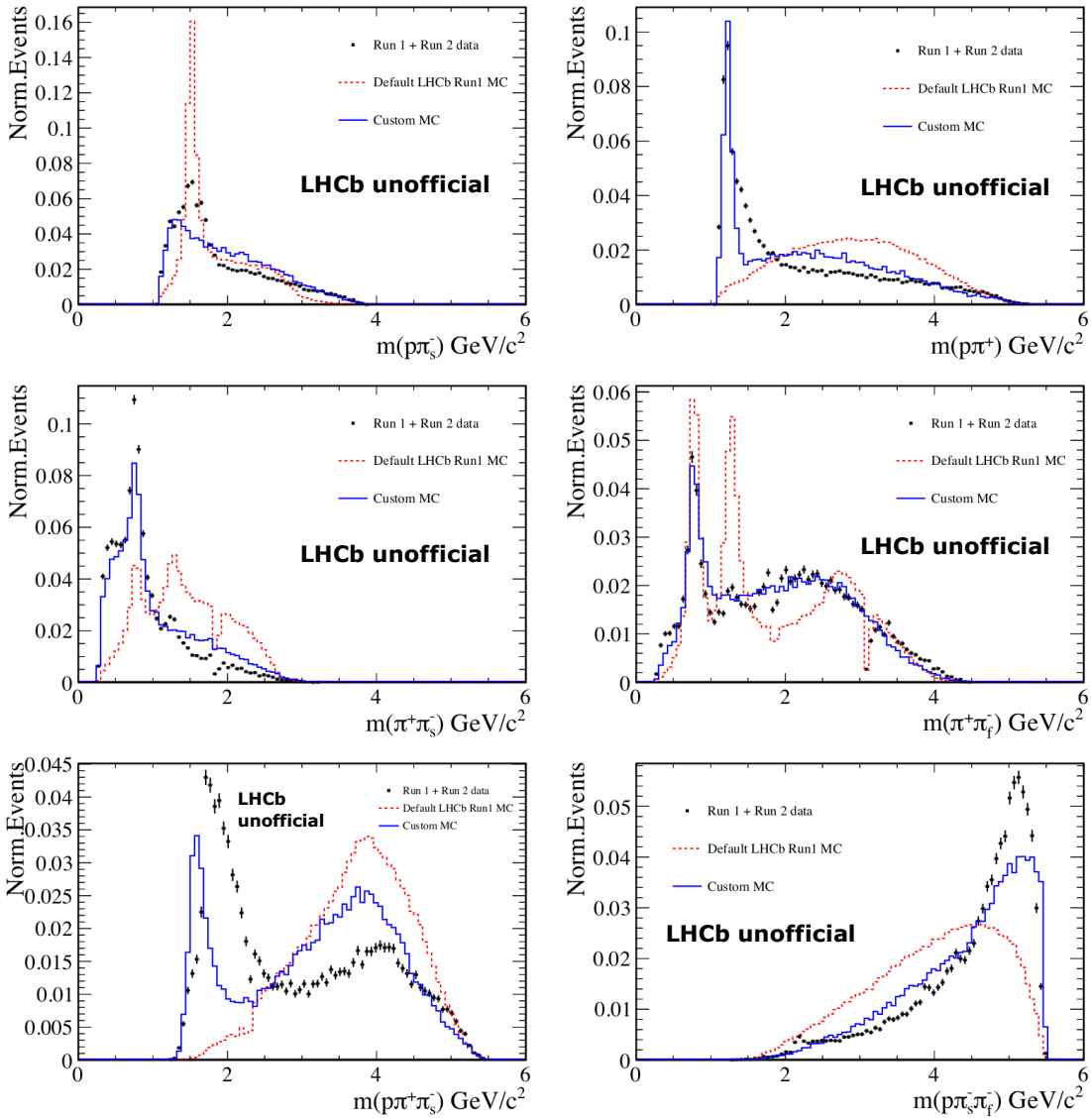


Figure 5.6: Invariant mass distributions for different combinations of Λ_b^0 decay products. Real data is shown in black, the overlaid *ad hoc* resonance cocktail is shown in blue and previous LHCb simulation is in red. The cocktail reproduces some of the main features observed in the data allowing it to be used in sensitivity studies.

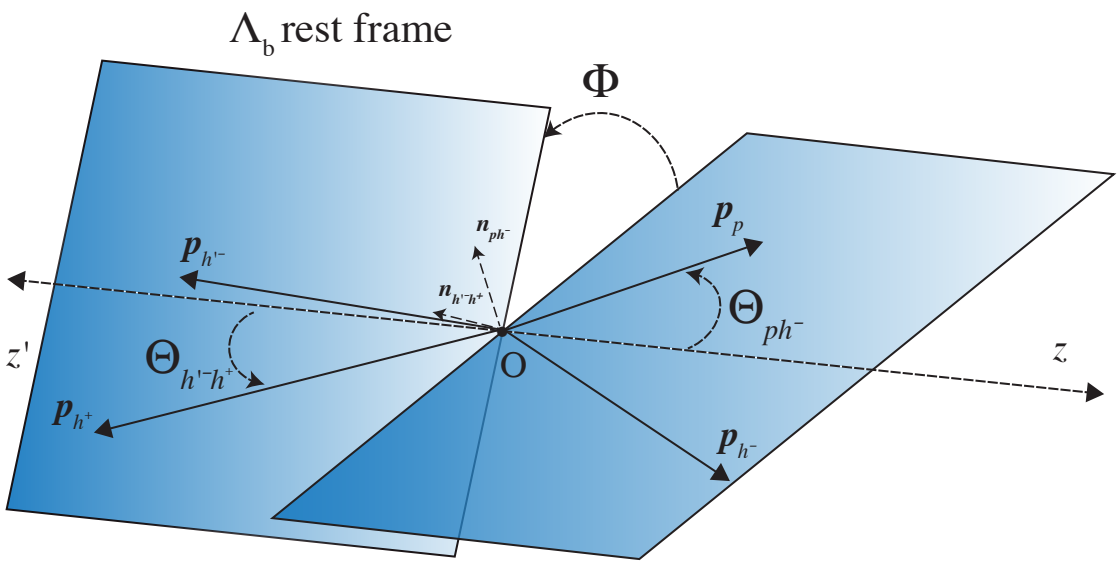


Figure 5.7: Diagram showing the decay products of Λ_b^0 with the corresponding angular variables. The angle Φ is the angle between decay planes formed by quasi-two-body decay product pairs. It is used in the definition of binning for the Triple Product Asymmetries method. Reproduced from [53].

to be simulated and thus introduces *P*-odd *CP*-violating effect in the simulated data. It is particularly important for the application of the Triple Product Asymmetries method, as the evidence presented in the previous analysis [53] was observed in the angle $|\Phi|$ between the decay planes (Fig. 5.7), which is a *P*-odd variable.

5.5 Sensitivity Studies

5.5.1 Choice of Binning for the Triple Product Asymmetries

The sensitivity of the Triple Product Asymmetry measurement depends on the choice of binning. Previous analysis of this channel [53] used two binning schemes. The first (scheme *A*) was based on the invariant masses of the decay product pairs forming the dominant resonances (e.g. $m(p\pi^+)$ for $\Delta^{++}(1232)$ resonance) and used 12 bins. The second binning scheme (*B*) used 10 bins of homogeneous binning in the angle between decay planes

$$|\Phi| = |\arccos(\hat{n}_{p\pi^-{}_s} \times \hat{n}_{\pi^+\pi^-{}_f})|, \tag{5.4}$$

where $\hat{n}_{ab}$ is the vector normal to the plane, defined by the momenta of particles a and b in the Λ_b^0 rest frame. The angle Φ is defined in the interval $(-\pi, \pi)$ and is shown in Fig. 5.7. The deviation from the *CP* violation hypothesis was mainly driven by the fourth bin of the second scheme (B), as can be seen in Fig. 5.2. When combined, the two binning schemes yielded the reported 3.3σ deviation from the *CP* conservation hypothesis, which constituted the first evidence of *CP* violation in the baryon sector.

One of the tasks in the current analysis is to improve the binning for the Triple Product Asymmetries method, thus providing better sensitivity to the *CP* violation. Since the

previous measurement of this channel, there have been improvements in the theoretical understanding of charmless four-body decays. In particular, discussions with the theory community gave insights into the nature of *CP* violation in the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays [91, 92].

The contributions to *P*-odd *CP* violation are expected to only arise in certain decay topologies where there is a possibility of interference between resonances with opposite parity. One such topology is the previously mentioned "cascade" decay $\Lambda_b^0 \to \pi^-\Delta^+(\to \pi^-\Delta^{++}(\to \pi^+ p))$, where N^{*+} resonances with different spins can interfere. For example, if $N^{*+}_{1/2}$ and $N^{*+}_{3/2}$ states have almost equal mass, the interference between S and P waves is possible. The amplitudes for this decay are given in Table 5.8.

Furthermore, the *P*-odd *CP* violation is expected to vanish in the $\Lambda_b^0 \to (a_1^- \to (\rho \to \pi^+\pi^-)\pi^-)p$ decay if it is dominant in a certain region of the phase space and is not interfering with other resonances. Indeed, the events decaying through the a_1^- resonance are mainly localized in the high mass region of $m(p\pi^+\pi^-_{slow})$ invariant mass combination with minor contributions from other resonances. For this reason, an additional cut of this mass variable is applied to remove the a_1^- resonance contribution and focus on the region populated by various N^{*+} resonances. The chosen mass cut, $m(p\pi^+\pi^-_{slow}) < 2.8$, removes approximately 45% of the signal events but is expected to increase the overall sensitivity by avoiding dilution effects. The excluded subsample is used for a cross-check, showing that no *CP* violation is measured.

The use of the simplified amplitude model, described in Sec. 5.4, allows *P*-odd *CP* violation to be introduced that resembles the effect seen in the previous analysis [53]. A specific choice of amplitudes, given in Table 5.9, ensures that the asymmetry is introduced in the local regions of phase space, while it cancels out if measured over all phase space. This is illustrated in Fig. 5.8. It can be seen that using the azimuthal angle of the proton in the Δ^{++} rest frame - $|\Phi_p|$ gives an increase in the sensitivity to *CP* violation when compared to the previously used angle between the decay planes $|\Phi|$. This is due to the fact that $|\Phi_p|$ is directly related to the "cascade" topology decay amplitudes, where *CP* violation is introduced, while using the angle Φ dilutes the sensitivity to such effects.

Taking these new insights into account allows new binning schemes to be defined for the Triple Product Asymmetries method. The previous homogeneous 10-bin binning in the angle Φ is repeated for comparison (referred to as binning scheme *B*, to be consistent with Ref. [53]). Two subsets of this binning are defined by applying the low $m(p\pi^+\pi^-_{slow}) < 2.8\,\text{GeV}/c^2$ invariant mass cut to focus on N^{*+} resonance contributions and reduce possible dilution from a_1 events with negligible expected *CP* violation. The binning in the angle $|\Phi|$ of the subset of data passing this cut is referred to as binning B_2, while the excluded region is binning B_1 [2]. The second binning scheme is designed to enhance the sensitivity to *CP* violation coming from the specific N^{*+} "cascade" topology and uses bins of proton helicity angles. As for the first binning, it has two subsets A_1 and A_2, where A_1 is the binning of the data excluded by the $m(p\pi^+\pi^-_{slow}) < 2.8\,\text{GeV}/c^2$

[2] The naming convention is chosen to reflect the existence of a_1 resonance in the A_1 and B_1 binning

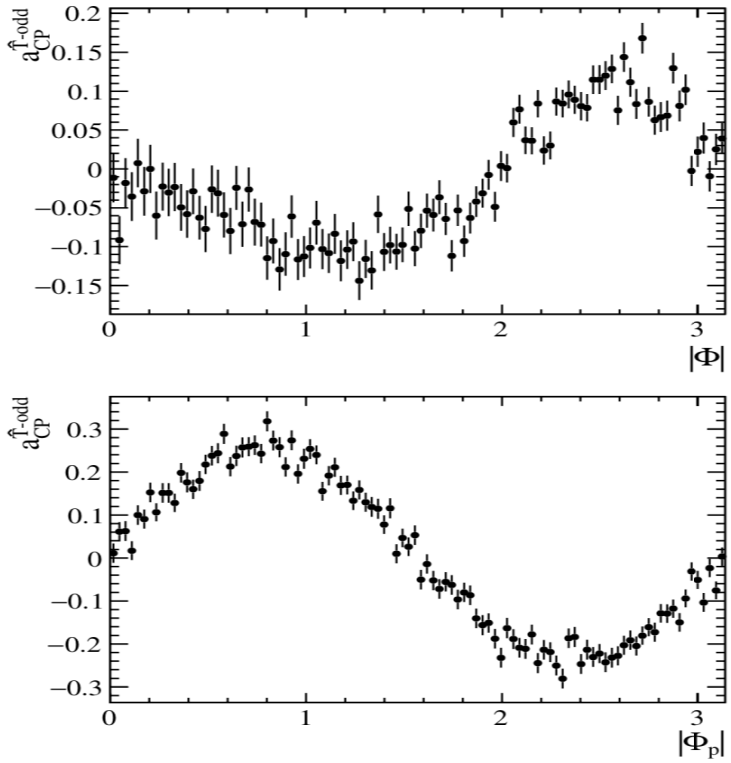


Figure 5.8: The Triple Product Asymmetries vs angular variables measured on simplified simulation data. (Top) *CP* asymmetry vs the angle between the decay planes Φ used in previous analysis. (Bottom) An improved sensitivity is seen if the azimuthal angle of the proton in Δ^{++} rest frame- $|\Phi_p|$ is used. Reproduced from [14].

	A_+	A_-	B_+	B_-	$b_{1+}/-b_{1-}$	$b_{2+}/-b_{2-}$	$b_{2+}/-b_{2-}$
Matter	1	0	i	i	0	1	1
Antimatter	1	0	i	$-i$	0	1	1

Table 5.9: The specific choice of amplitudes used to include CPV in the simulated sample.

invariant mass cut, and A_2 is the selected data. This binning scheme is defined in Table 5.10.

Before unblinding the data it was decided that only the results from binning schemes A_2 and B_2 will be quoted as the published results of this analysis for the Triple Product Asymmetries measurement. This is because these binning schemes are expected to have the highest sensitivity to the *CP* violation.

Table 5.10: Definition of the new binning (scheme A) based on the proton helicity angles. Reproduced from [14].

Bin number	Polar angles	Azimuthal angles
1	$\theta_p \in [0, \pi/4]$, $\theta_{\Delta^{++}} \in [0, \pi/4]$ $\theta_p \in [\pi/2, 3\pi/4]$, $\theta_{\Delta^{++}} \in [\pi/2, 3\pi/4]$	$\lvert\Phi_p\rvert \in [0, \pi/2]$
2	$\theta_p \in [0, \pi/4$, $\theta_{\Delta^{++}} \in [\pi/4, \pi/2]$ $\theta_p \in [\pi/2, 3\pi/4]$, $\theta_{\Delta^{++}} \in [3\pi/4, \pi]$	$\lvert\Phi_p\rvert \in [0, \pi/2]$
3	$\theta_p \in [0, \pi/4]$, $\theta_{\Delta^{++}} \in [\pi/2, 3\pi/4]$ $\theta_p \in [\pi/2, 3\pi/4]$, $\theta_{\Delta^{++}} \in [0, \pi/4]$	$\lvert\Phi_p\rvert \in [0, \pi/2]$
4	$\theta_p \in [0, \pi/4]$, $\theta_{\Delta^{++}} \in [3\pi/4, \pi]$ $\theta_p \in [\pi/2, 3\pi/4]$, $\theta_{\Delta^{++}} \in [\pi/4, \pi/2]$	$\lvert\Phi_p\rvert \in [0, \pi/2]$
5	$\theta_p \in [\pi/4, \pi/2]$, $\theta_{\Delta^{++}} \in [0, \pi/4]$ $\theta_p \in [3\pi/4, \pi]$, $\theta_{\Delta^{++}} \in [\pi/2, 3\pi/4]$	$\lvert\Phi_p\rvert \in [0, \pi/2]$
6	$\theta_p \in [\pi/4, \pi/2]$, $\theta_{\Delta^{++}} \in [\pi/4, \pi/2]$ $\theta_p \in [3\pi/4, \pi]$, $\theta_{\Delta++} \in [3\pi/4, \pi]$	$\lvert\Phi_p\rvert \in [0, \pi/2]$
7	$\theta_p \in [\pi/4, \pi/2]$, $\theta_{\Delta^{++}} \in [\pi/2, 3\pi/4]$ $\theta_p \in [3\pi/4, \pi]$, $\theta_{\Delta^{++}} \in [0, \pi/4]$	$\lvert\Phi_p\rvert \in [0, \pi/2]$
8	$\theta_p \in [\pi/4, \pi/2]$, $\theta_{\Delta^{++}} \in [3\pi/4, \pi]$ $\theta_p \in [3\pi/4, \pi]$, $\theta_{\Delta^{++}} \in [\pi/4, \pi/2]$	$\lvert\Phi_p\rvert \in [0, \pi/2]$
9	$\theta_p \in [0, \pi/4]$, $\theta_{\Delta^{++}} \in [0, \pi/4]$ $\theta_p \in [\pi/2, 3\pi/4]$, $\theta_{\Delta^{++}} \in [\pi/2, 3\pi/4]$	$\lvert\Phi_p\rvert \in [\pi/2, \pi]$
10	$\theta_p \in [0, \pi/4]$, $\theta_{\Delta^{++}} \in [\pi/4, \pi/2]$ $\theta_p \in [\pi/2, 3\pi/4]$, $\theta_{\Delta^{++}} \in [3\pi/4, \pi]$	$\lvert\Phi_p\rvert \in [\pi/2, \pi]$
11	$\theta_p \in [0, \pi/4]$, $\theta_{\Delta^{++}} \in [\pi/2, 3\pi/4]$ $\theta_p \in [\pi/2, 3\pi/4]$, $\theta_{\Delta^{++}} \in [0, \pi/4]$	$\lvert\Phi_p\rvert \in [\pi/2, \pi]$
12	$\theta_p \in [0, \pi/4]$, $\theta_{\Delta^{++}} \in [3\pi/4, \pi]$ $\theta_p \in [\pi/2, 3\pi/4]$, $\theta_{\Delta^{++}} \in [\pi/4, \pi/2]$	$\lvert\Phi_p\rvert \in [\pi/2, \pi]$
13	$\theta_p \in [\pi/4, \pi/2]$, $\theta_{\Delta^{++}} \in [0, \pi/4]$ $\theta_p \in [3\pi/4, \pi]$, $\theta_{\Delta^{++}} \in [\pi/2, 3\pi/4]$	$\lvert\Phi_p\rvert \in [\pi/2, \pi]$
14	$\theta_p \in [\pi/4, \pi/2]$, $\theta_{\Delta^{++}} \in [\pi/4, \pi/2]$ $\theta_p \in [3\pi/4, \pi]$, $\theta_{\Delta^{++}} \in [3\pi/4, \pi]$	$\lvert\Phi_p\rvert \in [\pi/2, \pi]$
15	$\theta_p \in [\pi/4, \pi/2]$, $\theta_{\Delta^{++}} \in [\pi/2, 3\pi/4]$ $\theta_p \in [3\pi/4, \pi]$, $\theta_{\Delta^{++}} \in [0, \pi/4]$	$\lvert\Phi_p\rvert \in [\pi/2, \pi]$
16	$\theta_p \in [\pi/4, \pi/2]$, $\theta_{\Delta^{++}} \in [3\pi/4, \pi]$ $\theta_p \in [3\pi/4, \pi]$, $\theta_{\Delta^{++}} \in [\pi/4, \pi/2]$	$\lvert\Phi_p\rvert \in [\pi/2, \pi]$

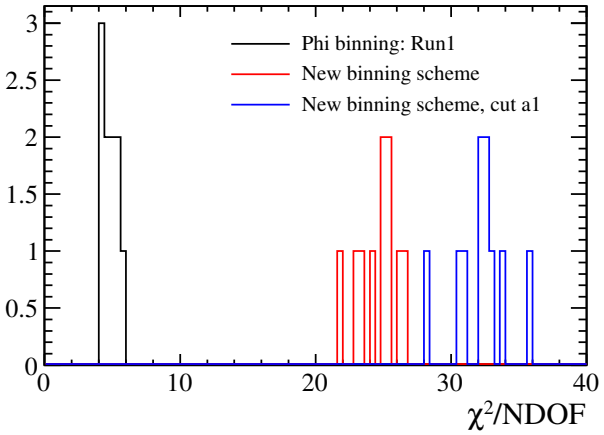


Figure 5.9: Diagram showing the χ^2 sensitivity comparison of three different binning schemes using the data from simplified simulation. The homogeneous angle $|\Phi|$ binning used in Run 1 analysis is shown in black (scheme B). The new binning in proton helicity angles, scheme A, defined in Table 5.10, is shown in red. The new binning (Scheme A_2) with the $m(p\pi^+\pi^-{}_{slow}) < 2.8$ mass cut applied is show in blue. Reproduced from [14].

5.5.2 Assigning Significance to the Triple Product Asymmetries Results

Once the Triple Products Asymmetries method is applied on the data and the results from either binned or integrated measurements are obtained, the significance of the results needs to be calculated. This is done by performing a χ^2 test, as in the previous analysis [53]. First of all, each bin is fitted to extract the asymmetries. All of the individual fits use the same fit model, outlined in Sec. 5.3, but are performed independently, since different phase space regions might have different kinematic distributions for signal and background. The $\chi^2 \equiv X^T V^{-1} X$ variable is constructed, where X is the array of $a_{CP}^{\widehat{T}\text{-odd}}$ values for each bin and V^{-1} is the covariance matrix, defined as the sum of statistical and systematic error matrices, which are discussed in Sec. 5.6. An average systematic uncertainty is assigned to each bin. The statistical uncertainties are not correlated among different bins, while maximal correlation is assumed for systematic uncertainties.

Simulations show that the distribution of $a_{CP}^{\widehat{T}\text{-odd}}$ values is Gaussian. The distribution of χ^2 values is obtained from pseudo-experiments by running permutations. The fraction of χ^2 values obtained from the permutations that are higher than the χ^2 value observed in the data allows to calculate the p-value and hence the significance of the measurement.

The different binning schemes are compared by applying them on simulated data, with CP violation introduced in the N^{*+} resonance. The comparison shows an improvement in sensitivity with this simulation of the new over the previous binning over the previous binning, which improves further if a_1 region is cut out. This comparison is shown in Fig. 5.9.

5.5.3 Energy Test Sensitivity Optimization

Multiple factors can enhance or reduce the sensitivity of the energy test, as described in Sec. 4.3. In the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays, the existence of two negative pions requires the introduction of a pion-ordering scheme. The chosen scheme orders the negative pions by the magnitude of their momenta in Λ_b^0 rest frame. This choice is also equivalent to the ordering by $m(p\pi^+\pi^-)$ invariant mass combination. This combination contains the N^{*+} resonances, hence its use in ordering negative pions also enhances the separation of N^{*+} resonances that occupy the low $m(p\pi^+\pi_{slow})$ mass range from other resonances, like the a_1, that occupy the high mass range. Alternative pion ordering schemes have been considered but no significant improvement to the sensitivity of the test was observed.

The second factor that greatly influences the sensitivity of the energy test is the choice of phase space coordinates. Many alternative sets of such coordinates exist. For example, invariant mass combinations of the decay products, their squares, or a mixture of masses and helicity angles of the decay can be used to define the decay phase space. A popular choice of the squares of invariant mass combinations is chosen to define the phase space as the alternatives do not give a significant increase in sensitivity. The different coordinate choices are compared by using a simplified simulation with the *CP* violation introduced in the *P*-odd amplitudes (see Table 5.9). Three different alternatives are investigated: invariant masses, invariant masses squared and a combination of invariant masses and helicity angles, with required normalization. The results of this comparison are shown in Fig. 5.10.

The five invariant-mass-square combinations chosen are:

- $m^2(p\pi^+)$ - contains the $\Delta^{++}(1232)$ resonance
- $m^2(\pi^+\pi^-_{slow})$ - contains the $\rho(775)$ resonance
- $m^2(p\pi^+\pi^-_{slow})$ - contains N^{*+} resonances
- $m^2(p\pi^+\pi^-_{fast})$ - contains $a_1^-(1260)$ resonance
- $m^2(p\pi^-_{slow})$

After the pions are ordered and phase space coordinates defined, the energy test requires a choice of the distance parameter δ. In decay modes where a good quality amplitude model exists, this can be used for optimising the chosen value [69]. However, no accurate amplitude model exists for the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decay. For this reason the optimal choice of δ is not known analytically, hence a more general data-driven approach is taken. In the current analysis three different δ values are chosen based on the mean distances between the events in the decay phase space:

1. Median distance between events in the phase space: $\mathbf{13}\,\mathrm{GeV}^{\mathbf{2}}$
2. Mean distance to the 600^{th} nearest neighbour: $\mathbf{2.7}\,\mathrm{GeV}^{\mathbf{2}}$
3. Mean distance to the 600^{th} nearest neighbour in the $m(p\pi^+\pi^-_s) < 2.8\,\mathrm{GeV}$: $\mathbf{1.6}\,\mathrm{GeV}^{\mathbf{2}}$

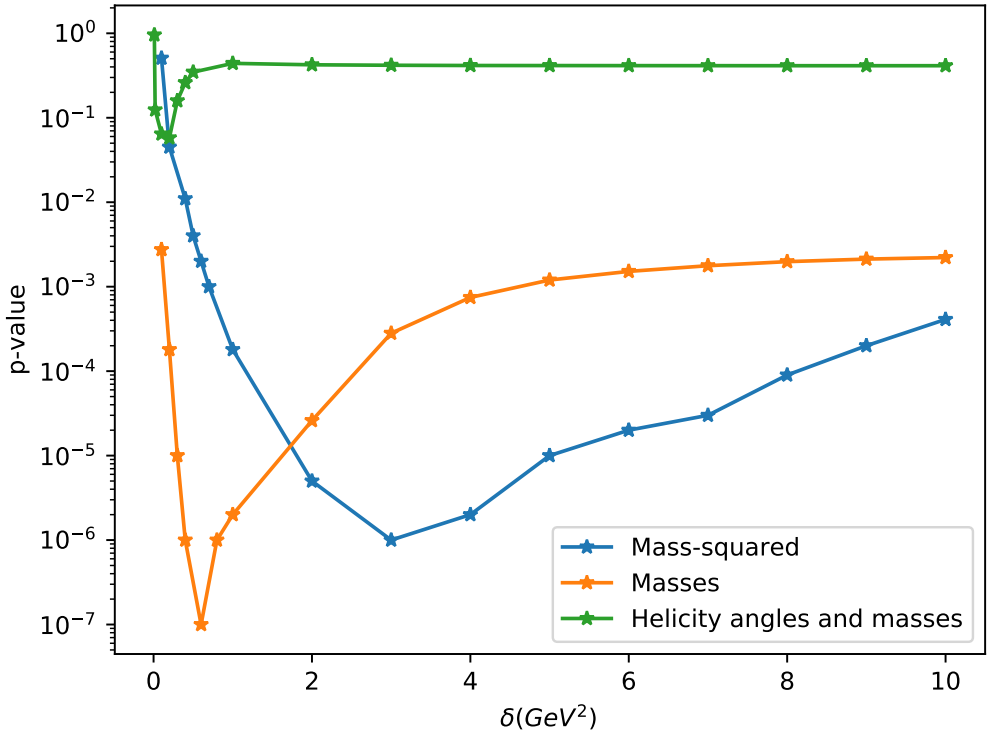


Figure 5.10: The sensitivity dependence on the value of the distance parameter δ for different choices of phase space coordinates. A single simplified simulation dataset is used for all of the choices of coordinates. Clear differences in sensitivity can be seen depending on the choice of coordinate system.

The 600 nearest neighbours are chosen because of an analogy with binned analyses. Assuming a large magnitude of *CP* violation ($\sim$ 20%) in the local regions of the phase space, one would require 600 events in the bin to be sensitive to this amount of asymmetry at the significance of 5σ.

Furthermore, these choices provide a good coverage of the overall size of the phase space, as can be seen in the event-distance distributions in Fig. 5.11. The first choice of a large value for $\delta = 13\,\text{GeV}^2$ would be sensitive to *CP* violation if the sensitivity of the measurement increased with higher value of δ. Such asymptotic behaviour has been observed in some of the simplified simulations. The second and third choices focus on the regions of phase space where events with the same resonance structures cluster together.

In the case of a significant result, the p-values need to be combined to quote a single significance, thus avoiding a possible look-elsewhere-effect. This is due to the fact that the test is applied on the same dataset with three different distance parameters. This can be achieved by defining a new test statistic $Q = p_1 p_2 p_3$, where p_i are the p-values for three different distance scales. The obtained Q-value is compared against Q-values from permutations that have randomly shuffled flavours and hence contain no *CP* violation by construction. The fraction of Q-values in the permutation samples above the value

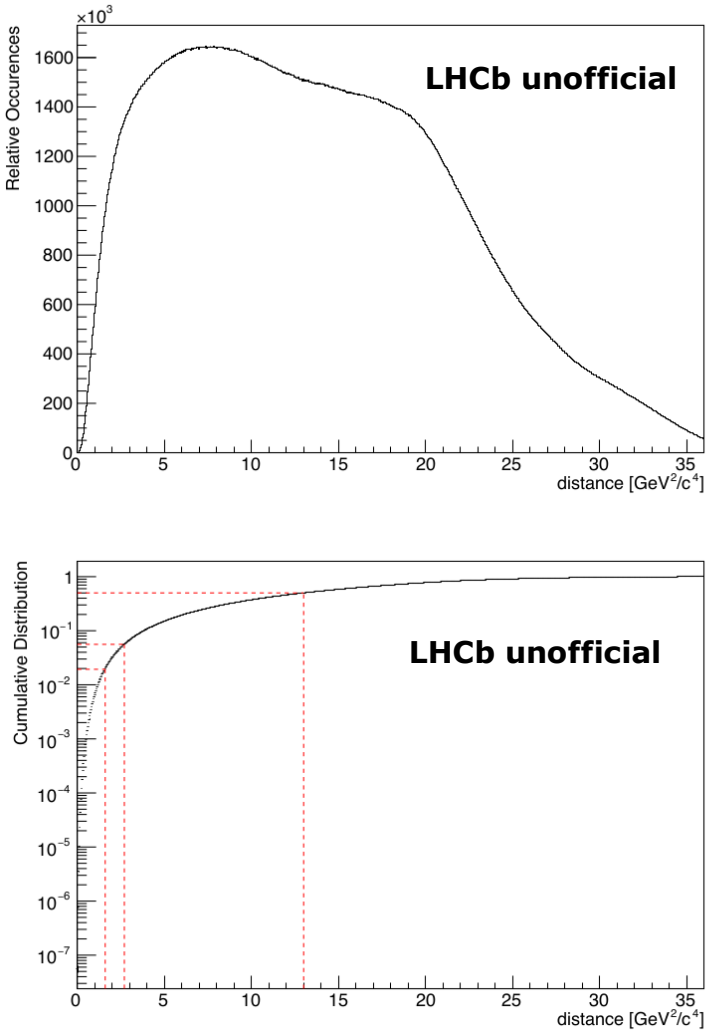


Figure 5.11: The distribution of distances between events in the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decay phase space (top) and the cumulative distances between events (bottom) in real data. This is used to choose the values of the distance parameter δ that are shown as vertical dashed lines in the bottom plot. Reproduced from [14].

observed in the data then defines the significance of the result. Two versions of the energy test are performed, one searching for P-even CP violation and one for P-odd. These two tests are independent, as they search for different effect, and are thus not combined.

5.6 Systematic Uncertainties of the Triple Product Asymmetries Analysis

The Triple Product Asymmetries method has multiple sources of systematic uncertainties that can bias the result and need to be addressed. The first source of uncertainties is the possible experimental bias, which comes from the reconstruction of the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays and detector acceptance effects. To measure the related uncertainties the Cabbibo-favoured $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$ control mode (described in Sec. 5.2.3) is used. The

control mode is expected to have negligible CP violation, hence any non-zero asymmetry in this channel is considered as an experimental bias. The second source of uncertainties is the detector resolution bias. Due to the finite detector resolution events can migrate between the samples with opposite sign of the triple product ($C_{\widehat{T}}$ and $\overline{C}_{\widehat{T}}$). To estimate this effect, a simulation sample is used, which provides a good approximation of the real data. The selection used on simulated data is identical to the one applied on the real data. Finally, the effect of using alternative fit models is also evaluated using pseudo-experiments. All of the Triple Product Asymmetries method systematic uncertainties are summarised in Table 5.11. More information about the treatment of the uncertainties for the Triple Product Asymmetries measurement can be found in Refs. [125, 14].

Table 5.11: Summary of systematic uncertainties for Triple Product Asymmetries measurements of the integrated sample. Reproduced from [14].

Bias	$\Delta A_{\widehat{T}}(\%)$	$\Delta \overline{A}_{\widehat{T}}(\%)$	$\Delta a_{CP}^{\widehat{T}\text{-odd}}(\%)$	$\Delta a_{CP}^{\widehat{T}\text{-odd}}(\%)$
Experimental Bias	± 0.23	± 0.23	± 0.16	± 0.16
$C_{\widehat{T}}$ resolution	± 0.01	± 0.01	± 0.01	± 0.01
Fit Model	± 0.08	± 0.08	± 0.06	± 0.06

5.7 Checks for the Energy Test

The energy test method is by construction insensitive to global asymmetries between the two data sets. However, local asymmetries arising due to detection, acceptance, production, and presence of background might mimic *CP* violation and lead to an incorrect claim of observation. A number of cross-checks have been carried out in order to ensure that the energy test is not sensitive to any of these effects.

A possible source of asymmetry might arise due to the difference in the interaction with the matter of the detector between particles and antiparticles. This can cause a detection asymmetry which might be local if it varies with particle kinematics. This could result in a large observed T-value of the energy test and hence mimic the *CP* violation. The impact of this is estimated by analysing the $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$ control channel, described in Sec. 5.2.3. The results of the control channel analysis are presented in Sec. 5.7.1. This analysis may be conservative, since the kaon detection asymmetries are known to be significantly larger than pion asymmetries [147]. Yet the main source of asymmetry will come from protons contained in both the control and signal channel final states. In the case of such asymmetries, a large T-value would be observed when analysing the control channel. An additional simulation study of induced detection asymmetries is performed and described in Sec. 5.7.2.

The measurements of the control channel would also be sensitive to asymmetries due to the possible difference in the acceptance of Λ_b^0 and $\overline{\Lambda}_b^0$ candidates. This effect should be partially cancelled out by using the data samples with both polarities of the LHCb magnet. However, any residual effect would appear in the control channel as it has the

same momentum requirements applied on the final state particles as the signal channels. Similarly, the production asymmetries of Λ_b^0 and $\overline{\Lambda}_b^0$ might have a kinematic dependence, and again that would also be present in the control channel. However, no evidence of production asymmetries dependant on Λ_b^0 kinematics is found [50].

Apart from detection, acceptance and production effects, the result of the energy test might be biased by the presence of a difference in background in the two samples. This can occur if the background fraction (purity) is different between the two samples or if the background distributions have intrinsically different structure. In order to assess these effects, subsamples with large background components (high mass side band of the signal channel) are analysed. If these samples showed significant differences, this could also apply to the background inside the signal region. The analysis of background samples is presented in Sec. 5.7.3.

These cross-checks provide a way to validate the energy test method and data samples by confirming that if an asymmetry is observed it had to be caused by a real *CP*-violating effect. If these cross-checks are successful in the application of two sample-tests, no systematic uncertainties are assigned. This is the standard procedure in this style of method and has been used, for example, by the LHCb collaboration in previous applications of the energy test [71, 74] and "nearest neighbour" method [73], as well as in applications of the S_{CP} method by LHCb and also by the BaBar collaborations [34, 35, 70, 55, 73].

5.7.1 Control Channel Analysis with the Energy Test

The $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$ channel is described in 5.2.3 and used for various cross-checks of the energy test method. It is kinematically similar to the signal channel, yet it is expected to have negligible *CP* violation. For this reason, the analysis of this channel using the energy test should not result in a significant T-value. The proton momentum distributions are similar for signal and control channels, as can be seen in Fig. 5.13. The coverage provided by the control channel is conservative, as it contains more low-momentum protons where a higher asymmetry is expected [4].

The control channel has much higher yield than the signal channel. The combined Run 1 and Run 2 datasets of the control channel contain over 448,000 events in the signal Λ_b^0 mass range ($5.574 - 5.662\,\text{GeV}/c^2$) as shown in Fig. 5.12. In order to analyse the control channel, it is subsampled 50 times to match the yields expected in the signal channel. Each of these subsamples contain 32000 events of both flavours. The energy test is separately applied on each of the subsamples. As described in Sec. 4.3, the P-even version of the energy test does not require identical yields of matter and antimatter candidates as any global effects are cancelled by the normalization in the T-value calculation. In the P-odd test the flavours are mixed when the two samples to be tested are constructed, hence global effects could be enhanced. This is compensated by equalizing the yields of negative and positive $C_{\hat{T}}$ and $\overline{C}_{\hat{T}}$ subsamples. The kaons in the control channel are treated as pions with the same pion ordering applied. Both P-even and P-odd versions of the energy test are applied with three different distance parameters δ. The distributions of the p-values for all of these test can be found in Fig. 5.14. None of the control channel measurements

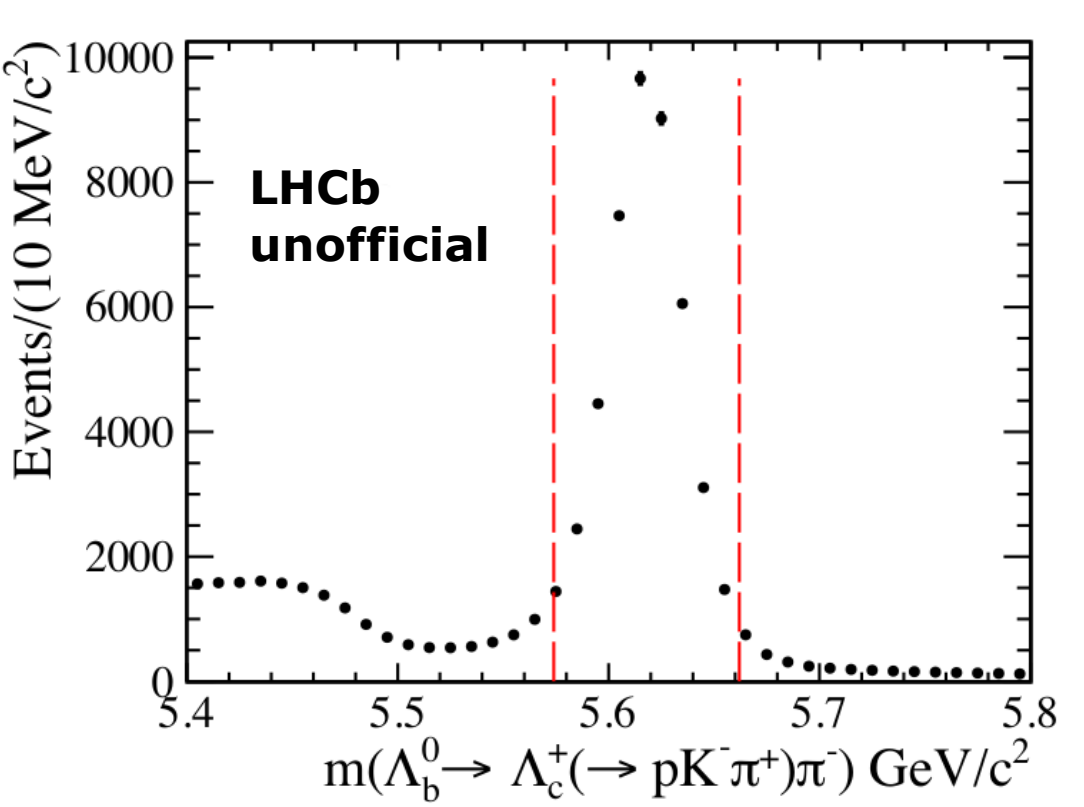


Figure 5.12: The $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$ control channel invariant mass distribution from real data. The mass window used for the energy test cross-check is marked with red dashed lines.

show evidence of asymmetries and all the distributions are consistent with being flat. This confirms that the energy test applied on the signal channel is not expected to be sensitive to production, detection or selection effects.

5.7.2 Detection Asymmetry Analysis

An additional cross-check has been performed in order to confirm that the energy test applied on the signal channel would not be sensitive to a detection asymmetry. As mentioned before, the main contribution to such an asymmetry is expected to come from protons. The proton/anti-proton asymmetries are well understood and quantified in other studies (for example, Refs. [4, 147]).

In this study, a simplistic *CP* symmetric simulation sample is produced using the RapidSim package. The simulated data contains 32 thousand events, which is the expected combined signal yield. A detection asymmetry is introduced by randomly discarding candidate decays, matching the asymmetry found in Ref. [4] as a function of the proton momentum in the lab frame, as shown in Fig. 5.15. The function used to introduce this asymmetry is $A = 1 + 3e^{-p/10\,\text{GeV}}\%$, where p is the proton momentum in the lab frame. This function reasonably accurately describes the asymmetry given in Ref. [4] but is continuous.

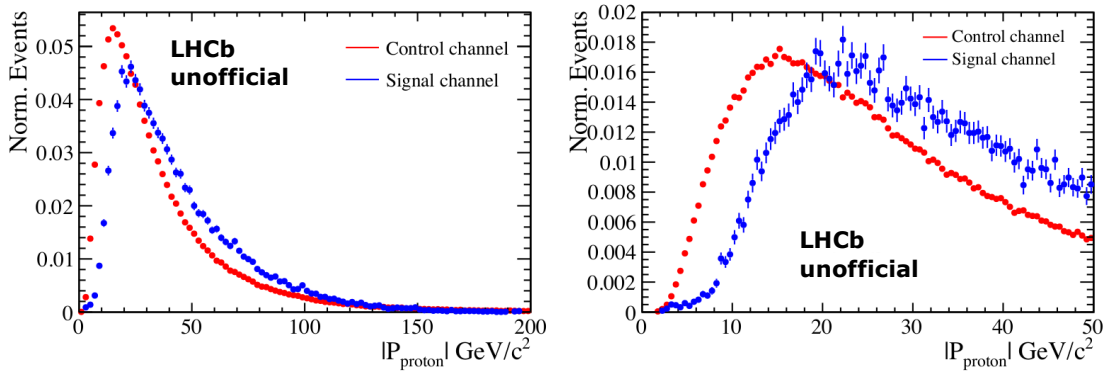


Figure 5.13: The comparison of the absolute value of proton momentum between signal and control channels using the real data.

The energy test has been applied on these simulated samples in both, P-even and P-odd configurations with three different distance scales δ. This has been repeated 30 times and the p-value distributions plotted in Fig. 5.16. No evidence of bias is found and the distributions are consistent with being flat.

Additionally, the asymmetry in the LHCb full-detector simulation is determined as a function of the proton momentum (shown in Fig. 5.17). A linear fit is performed: $A = 0.028 + 0.00012 \times p/\,\text{GeV}\%$. This asymmetry potentially also includes additional effects beyond the simple detection asymmetry set out above, including those associated with the triggering and selection of events. The studies set out above are repeated with this asymmetry: events are discarded in the custom CP-symmetric simulation samples according to this function and the energy test is applied. The distribution of p-values can be seen in Fig. 5.18. As previously, no evidence of bias is found and the distributions are consistent with being flat.

5.7.3 Effects of Background

The asymmetries that might arise due to the background effects are analysed. The data used for the background studies is obtained by taking a subsample of the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays with high invariant mass. This subsample is usually referred to as the high mass side-band (HMSB) region and is expected to be dominated by the combinatorial background. The mass window of $5.75 - 6.1\,\text{GeV}/c^2$ of the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ invariant mass spectrum is chosen. This window, shown in Fig. 5.19, contains approximately the same yield as the expected background component in the signal region. Also the proton momentum distribution does not vary significantly throughout the selected high mass side band region, as shown in Fig. 5.19. This indicates that the background in the high mass side-band should accurately represent the combinatorial background present in the signal region. Any bias observed in the selected background sample would also exist in the background within the signal region. The energy test is applied on the selected high mass side-band data in both P-even and P-odd configurations with three different distance scales δ. No significant p-values are obtained, which is consistent with the combinatorial background being CP-symmetric.

Table 5.12: The resulting p-values obtained by applying the different versions of the energy test on the high mass side-band background data. All results are consistent with *CP* conservation.

δ	1.6 GeV^2	2.7 GeV^2	13 GeV^2
P-even	0.8119	0.7531	0.8952
P-odd	0.3175	0.7556	0.7397

Table 5.13: The resulting p-values obtained when applying the different versions of the energy test on small subset of the $\Lambda_b^0 \to pK^-\pi^+\pi^-$ background, with the yield similar to that expected in the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ signal region. All results are consistent with *CP* conservation.

δ	1.6 GeV^2	2.7 GeV^2	13 GeV^2
P-even	0.1776	0.2168	0.241
P-odd	0.3586	0.6228	0.8093

Furthermore, the signal channel has a small background component (i.e. the purity of the data is high). This background component is consistent between the matter and antimatter samples. For this reason, the background subtraction technique (described in Sec. 4.3.4) does not have to be employed. The results of the background analysis are summarised in Tab. 5.12.

Another source of background is the small contamination coming from $\Lambda_b^0 \to pK^-\pi^+\pi^-$ decays, which arises due to the kaon being incorrectly identified as a pion. This background makes up approximately 10% of the total number of events in the $5.574 - 5.662\,\text{GeV}/c^2$ signal mass window. While the sensitivity to detection asymmetry in this final state has already been discussed in Sec. 5.7.2 and proven not to be significant, the peaking background could contain a true *CP* asymmetry. However, it is expected to be negligible as the contributing tree and penguin diagrams in this decay are proportional to different powers of the Wolfenstein parameter λ of the CKM matrix - $\mathcal{O}(\lambda^4)$ for tree and $\mathcal{O}(\lambda^2)$ for penguin diagrams. Previous LHCb studies have not observed an asymmetry in this channel [76]. However, an additional study is performed by unblinding a small fraction of the $\Lambda_b^0 \to pK^-\pi^+\pi^-$ data, which corresponds to the yield expected to contribute in the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ signal region. The energy test is applied on this subset and no significant p-values are observed, showing that the test is not sensitive to asymmetries that might be present in this background sample. The results are summarised in Tab. 5.13.

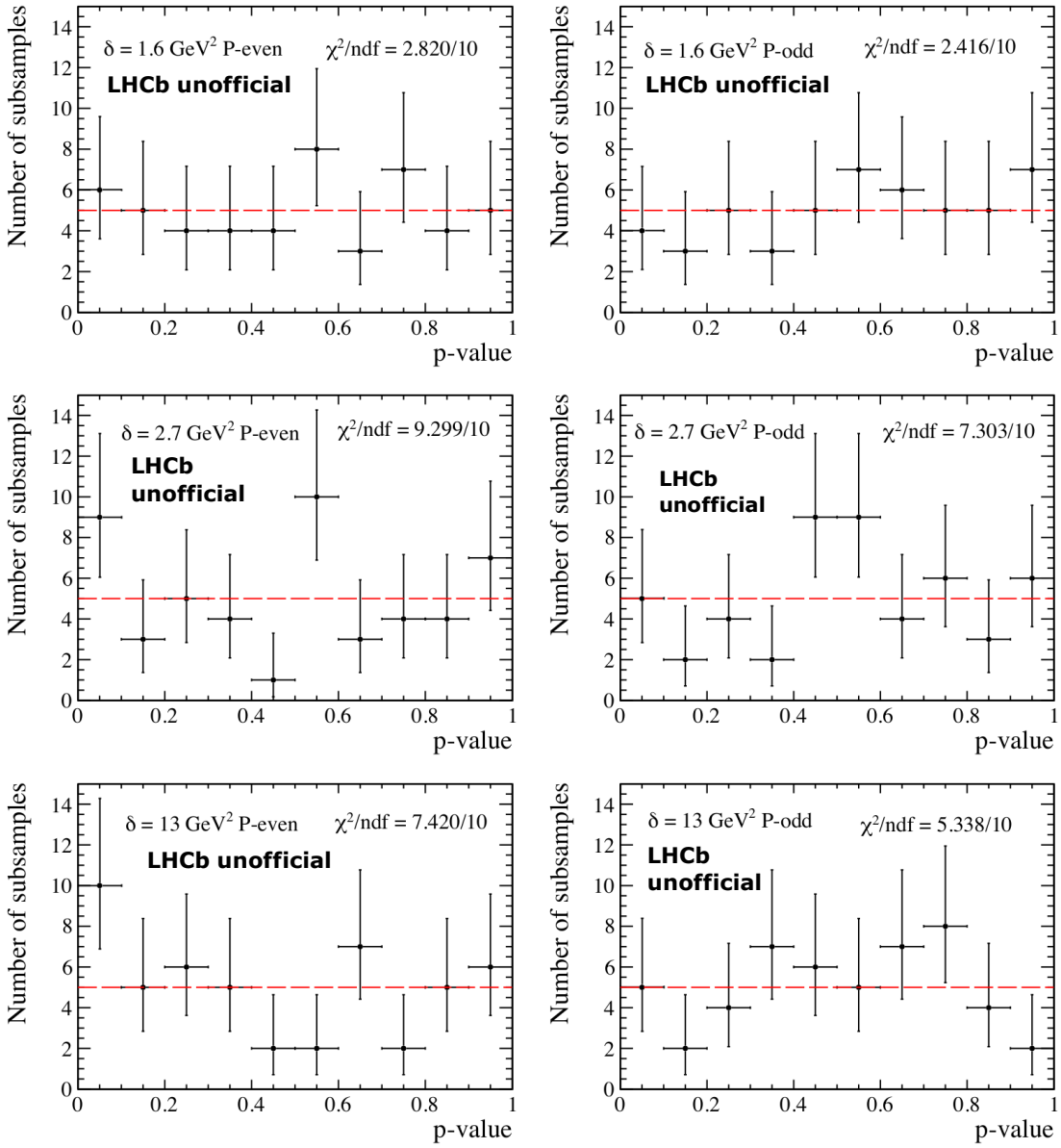


Figure 5.14: Distributions of p-values from the energy test applied to subsamples of the $\Lambda_b^0 \to \Lambda_c^+(\to pK^-\pi^+)\pi^-$ control channel. The top row shows $\delta = 1.6\,\text{GeV}^2/c^4$, the middle shows $\delta = 2.7\,\text{GeV}^2/c^4$ and the bottom row shows $\delta = 13\,\text{GeV}^2/c^4$, for both P-even (left) and P-odd (right) versions of the test. Uncertainties shown here provide 68.3% coverage. All distributions are consistent with Poisson fluctuations from a flat distribution.

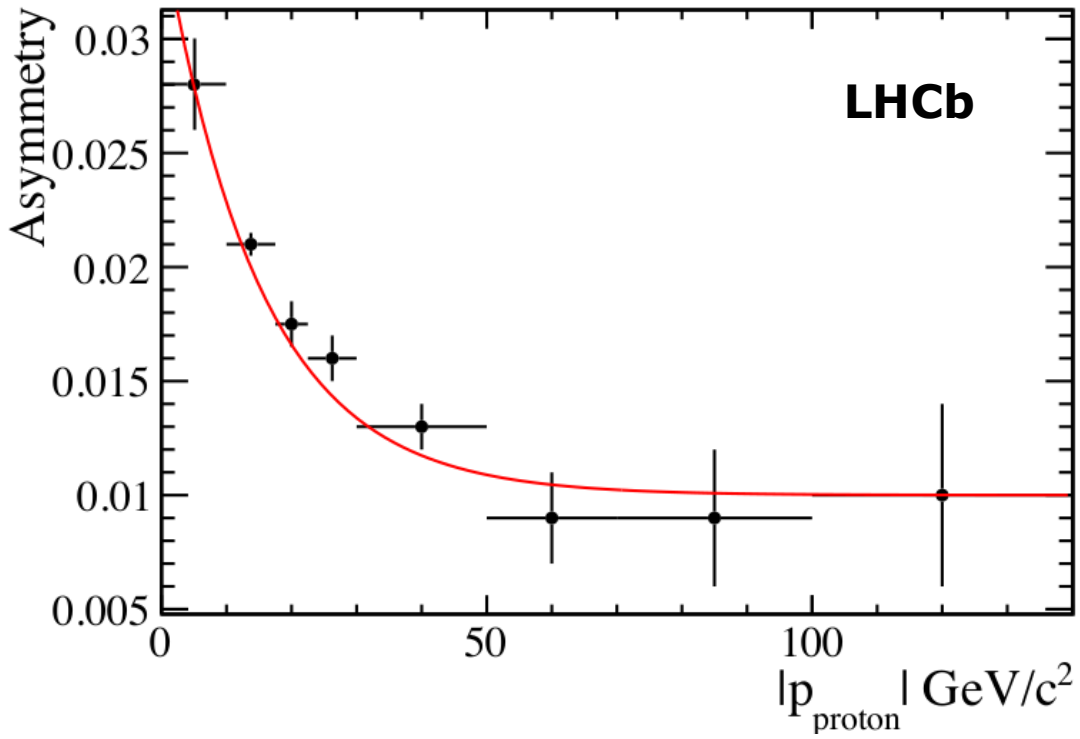


Figure 5.15: Asymmetry between protons and antiprotons as a function of the the absolute value of proton momentum (from Ref. [4]) with the function used to generate asymmetries here overlaid. Reproduced from [14].

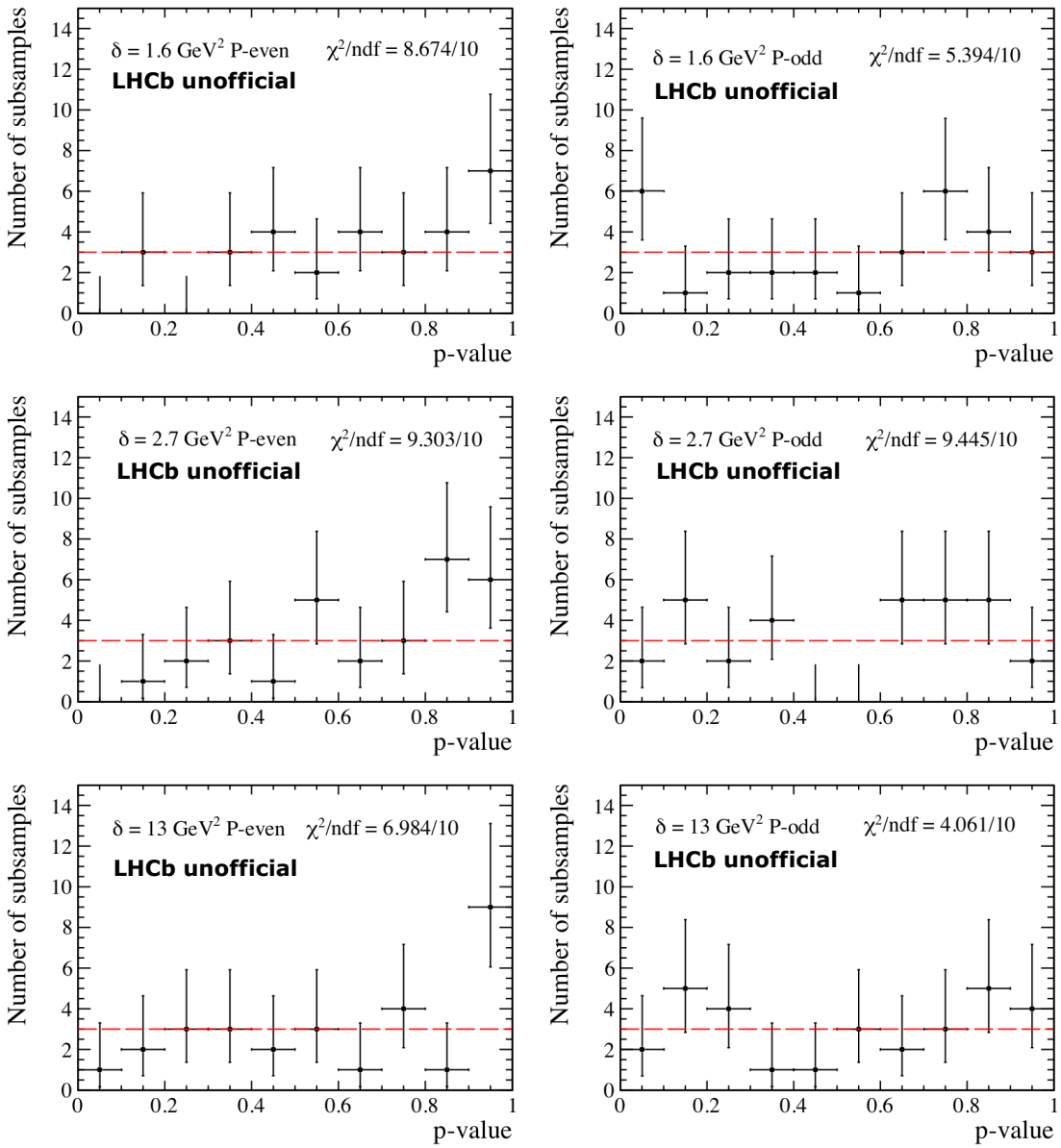


Figure 5.16: Distributions of p-values from the energy test applied to a CP symmetric simulation sample, with a detector asymmetry induced matching that reported in Ref. [4]. The top row shows $\delta = 1.6\,\mathrm{GeV}^2/c^4$, the middle shows $\delta = 2.7\,\mathrm{GeV}^2/c^4$ and the bottom row shows $\delta = 13\,\mathrm{GeV}^2/c^4$, for both P-even (left) and P-odd (right) versions of the test. Uncertainties shown here provide 68.3% coverage. All distributions are consistent with Poisson fluctuations from a flat distribution.

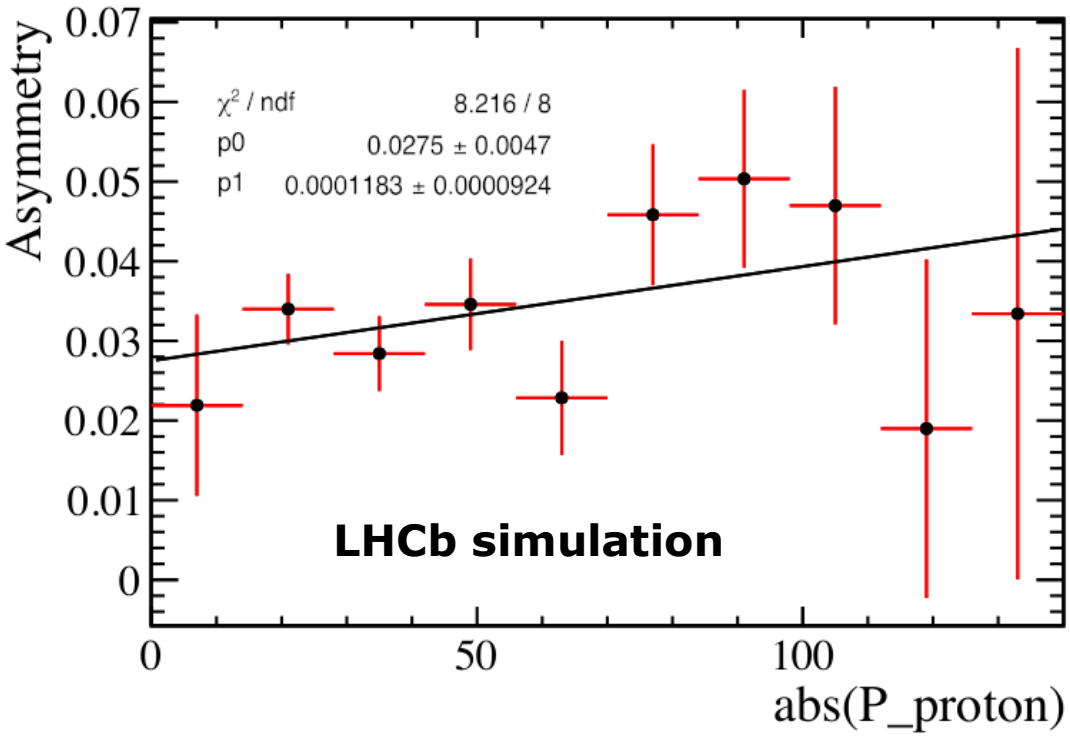


Figure 5.17: Asymmetry between protons and antiproton as a function of the absolute value of the proton momentum in the full LHCb simulation with a linear fit overlaid.

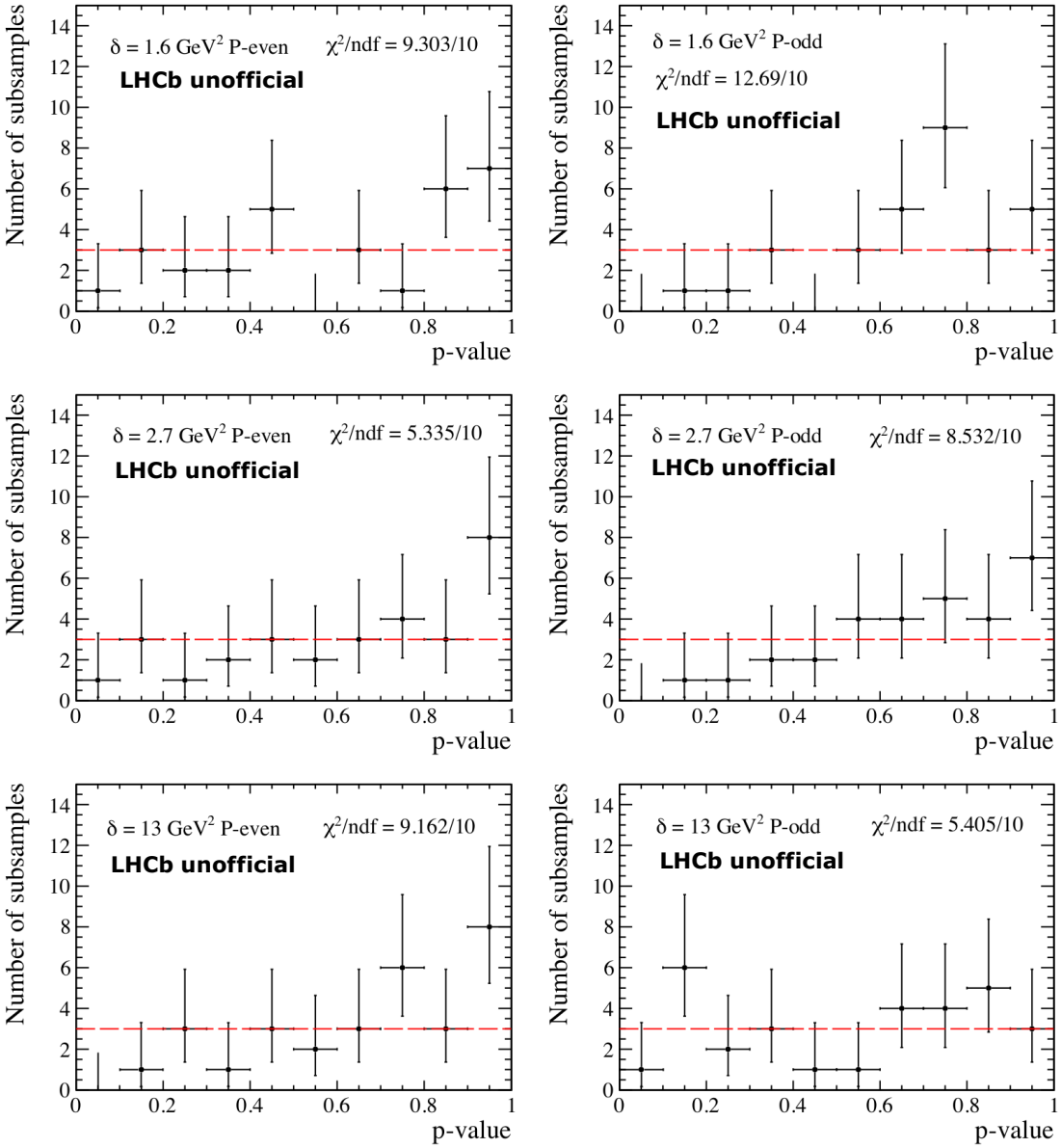


Figure 5.18: Distributions of p-values from the energy test applied to a CP-symmetric simulation sample, with the detector asymmetry in the full LHCb simulated data induced. The top row shows $\delta = 1.6\,\mathrm{GeV}^2/c^4$, the middle shows $\delta = 2.7\,\mathrm{GeV}^2/c^4$ and the bottom row shows $\delta = 13\,\mathrm{GeV}^2/c^4$, for both P-even (left) and P-odd (right) configurations of the test. Uncertainties shown here provide 68.3% coverage. All distributions are consistent with Poisson fluctuations from a flat distribution. Reproduced from [14].

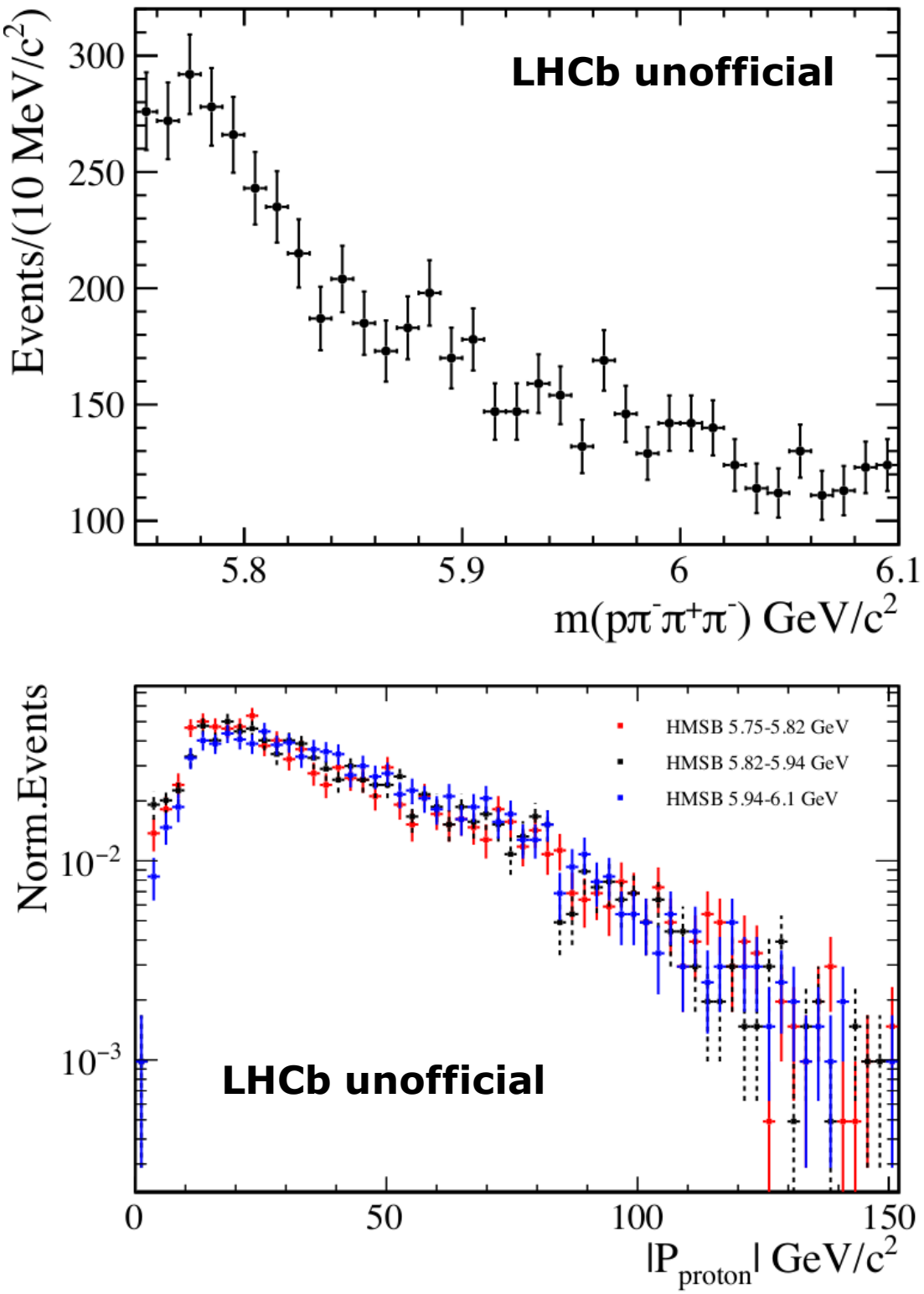


Figure 5.19: The invariant mass distribution of the $m(p\pi^-\pi^+\pi^-)$ data high mass side-band region between 5.75 and 6.1 GeV used for background studies (top). The comparison of the absolute value of proton momenta for three subsets of this mass range (bottom).

5.8 Results

5.8.1 Results of the Triple Product Asymmetries

The search for *CP* violation is carried out using the Triple Product Asymmetries method, described in Sec. 4.2. The measurement is performed in the integrated phase space of the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays, as well as with two different binning schemes (defined in Sec. 5.5.1). The choice of two binning schemes is motivated by the sensitivity studies, described in Sec. 5.5, and is optimized to provide better sensitivity to the simulated *CP* violation sources. However, the true source of *CP* violation in this channel remains unknown. For this reason, the binning scheme, where the effect from the previous analysis was seen, is reused on the new bigger dataset [53]. This also allows direct comparison of the Run 1 result with the current measurement.

The results obtained from the Triple Product Asymmetries measurement on the integrated decay phase space show no deviation from *CP* conservation hypothesis. The statistical and systematic uncertainties of the current measurement are smaller than those of the previous result. The measured integrated asymmetry variables for the current and previous analyses are compared in Table 5.14.

The binned phase space measurements are consistent with *CP* conservation at the level of 0.5 and 2.9σ for the chosen binning schemes A2 and B2 respectively.

The compatibility between Run 1 analysis [53] and the new results is checked by using the previous binning scheme (B). The comparison of the results is shown in Fig. 5.20. As can be seen, the value of the fourth bin, which was driving the previous deviation, has reduced. This reduction in the significance of the measurement is consistent with the previous result being a statistical fluctuation. However, the new result in one of the binning schemes (B2) is 2.9σ, which is marginally consistent with *CP* conservation hypothesis. This result is not driven by a single bin, unlike in the previous analysis.

Additionally, measurement of parity violation is carried out. The integrated measurement shows an asymmetry with 5.5σ significance. The binned measurements result in a deviation from *P* conservation hypothesis of 5.5σ in one of the binning schemes (B1) as well as 5.1σ in the binning used in the previous analysis (B). These results constitute an observation of parity violation in this channel. The $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decay is a weak

Table 5.14: Measurement of the asymmetries from the fit to the full dataset. The first uncertainty is statistical and the second systematic. The results from the previous analysis of this channel are given for comparison [53]. Reproduced from [14].

Asymmetry	Dataset	
	These results (2011-2017)	Run 1 (2011-2012)
$A_{\widehat{T}}$ (%)	$-4.68 \pm 0.99 \pm 0.24$	$-2.56 \pm 2.05 \pm 0.44$
$\overline{A}_{\widehat{T}}$ (%)	$-3.29 \pm 0.99 \pm 0.24$	$-4.86 \pm 2.05 \pm 0.44$
$a_P^{\widehat{T}\text{-odd}}$ (%)	$-3.98 \pm 0.70 \pm 0.17$	$-3.71 \pm 1.45 \pm 0.32$
$a_{CP}^{\widehat{T}\text{-odd}}$ (%)	$-0.70 \pm 0.70 \pm 0.17$	$1.15 \pm 1.45 \pm 0.32$

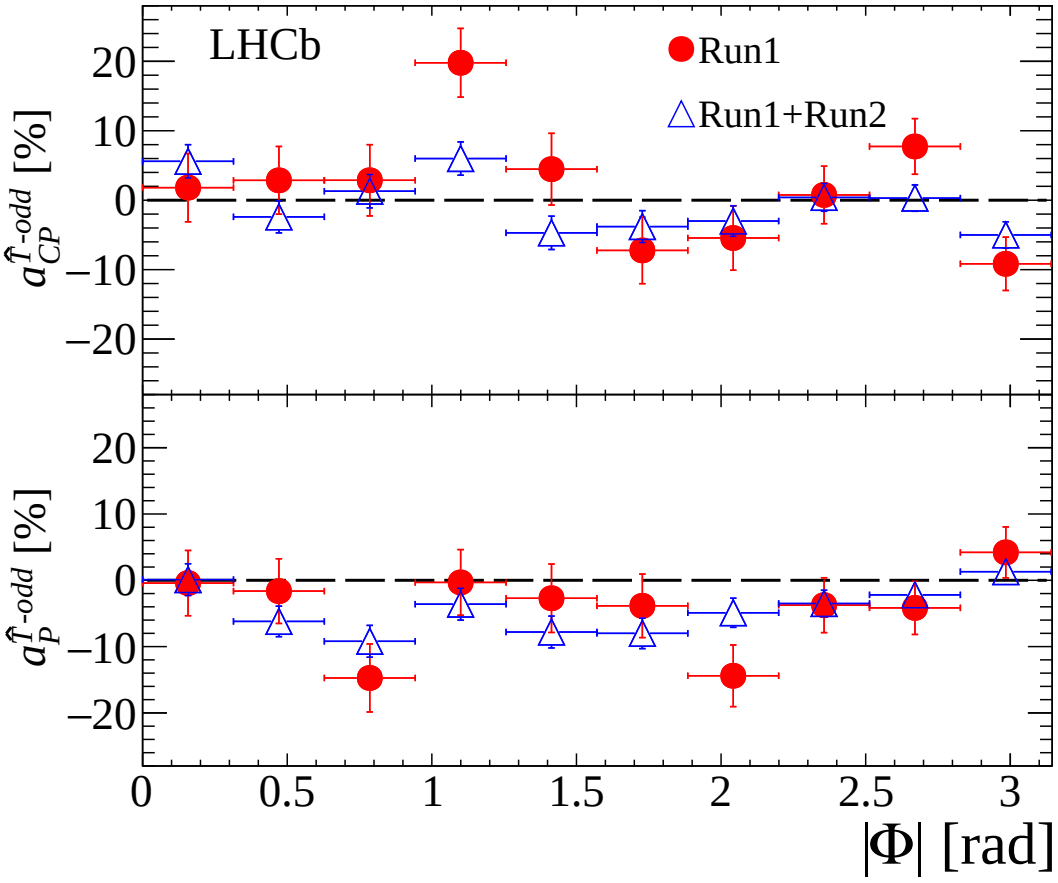


Figure 5.20: Measured asymmetries for binning scheme B with different datasets. The error bars represent the sum in quadrature of the statistical and systematic uncertainties. Reproduced from [68].

decay and is expected to violate parity. However, the level of parity violation expected in a baryon decay is not clear, so this remains an observation of some interest.

The measurements of CP and P asymmetries for all binning schemes are shown in Figs. 5.21 and summarised in Table 5.15. Further details and interpretation can be found in Refs. [125, 14].

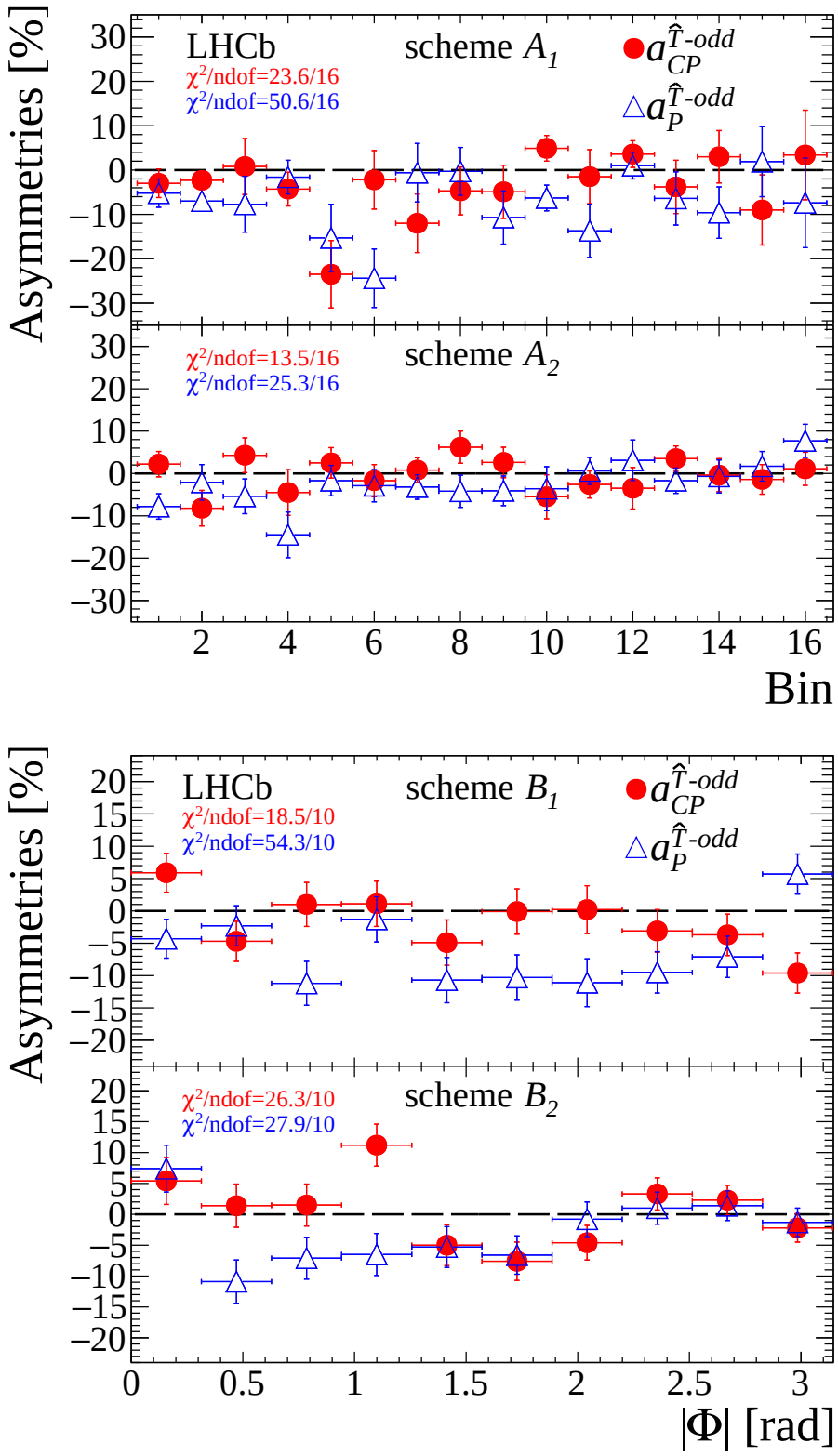


Figure 5.21: Measured asymmetries for the binning schemes A_1 and A_2 (top) and B_1 and B_2 (bottom). The error bars represent the sum in quadrature of the statistical and systematic uncertainties. The χ^2 per ndof is calculated with respect to the null hypothesis and includes statistical and systematic uncertainties. Reproduced from [68].

Table 5.15: The χ^2 and p-values obtained for the CP and P conserving hypotheses for different binning schemes taking into account systematic effects. Reproduced from [14].

Binning scheme	Dominant contribution	Hypothesis	χ^2/ndf	p-value	Significance
$\lvert\Phi\rvert$ (B)		CP conserving	25.20/10	5×10^{-3}	2.8σ
		P conserving	49.38/10	3.47×10^{-7}	5.1σ
$\lvert\Phi\rvert$ (B_1 excluded) (with $m(p\pi^+\pi^-_{slow}) > 2.8\,\text{GeV}$)	$\Lambda_b^0 \to pa_1$	CP conserving	18.49/10	4.72×10^{-2}	2.0σ
		P conserving	54.27/10	4.32×10^{-8}	5.5σ
$\lvert\Phi\rvert$ (B_2) (with $m(p\pi^+\pi^-_{slow}) < 2.8\,\text{GeV}$)	$\Lambda_b^0 \to N^{*+}\pi^-$	CP conserving	26.29/10	3.4×10^{-3}	2.9σ
		P conserving	27.91/10	1.87×10^{-3}	3.1σ
Helicity angles (A_1 excluded) (with $m(p\pi^+\pi^-_{slow}) > 2.8\,\text{GeV}$)	$\Lambda_b^0 \to pa_1$	CP conserving	23.62/16	9.81×10^{03}	1.7σ
		P conserving	50.59/16	1.84×10^{-5}	4.3σ
Helicity angles (A_2) (with $m(p\pi^+\pi^-_{slow}) < 2.8\,\text{GeV}$)	$\Lambda_b^0 \to N^{*+}\pi^-$	CP conserving	13.48/16	6.372×10^{-1}	0.5σ
		P conserving	25.34/16	6.4×10^{-2}	1.9σ

5.8.2 Results of the Energy Test

The energy test is applied on the full data set, described in Sec. 5.2. The CP violation measurements with three different choices of the distance parameter δ ($1.6\,\text{GeV}^2/c^4$, $2.7\,\text{GeV}^2/c^4$, $13\,\text{GeV}^2/c^4$) are carried out in both P-even and P-odd configurations of the test. The resulting six measurements are consistent with CP conservation. The measured p-values correspond to deviations from the CP conservation hypothesis with significances of 3σ or less, hence below the evidence threshold.

In the P-even CP violation energy test configuration, the obtained p-value is at the 3σ observation threshold. However, three different distance scales are used for this test. For this reason, a combined p-value needs to be calculated. This is done by defining a new test statistic $Q = p_1\, p_2\, p_3$, where p_i corresponds to a p-value for a distance scale i. The value of Q observed in data is then compared to the corresponding values from permutations, considering correlations between the different distance scales. The combined p-value for the P-even CP violation energy test configuration is 4.6×10^{-3}, which corresponds to approximately 2.8σ and hence is below the evidence threshold.

The parity violation configuration of the energy test is applied for the first time. The measurement results in an observation of P violation at a significance higher than 5σ with two distance scales δ. The low significance observed with $\delta = 13\,\text{GeV}^2/c^4$ can be interpreted as a dilution of sensitivity to the large value of delta averaging out the contributing small phase space structures. This shows that the three distance scales are indeed probing different phase space features.

The resulting p-values of the different test configurations are summarised in Tab. 5.16.

Table 5.16: The p-values from the energy test for different distances scales and test configurations. Reproduced from [68].

Distance scale δ	$1.6\,\text{GeV}^2/c^4$	$2.7\,\text{GeV}^2/c^4$	$13\,\text{GeV}^2/c^4$
p-value (CP conservation, P even)	3.1×10^{-2}	2.7×10^{-3}	1.3×10^{-2}
p-value (CP conservation, P odd)	1.5×10^{-1}	6.9×10^{-2}	6.5×10^{-2}
p-value (P conservation)	1.3×10^{-7}	4.0×10^{-7}	1.6×10^{-1}

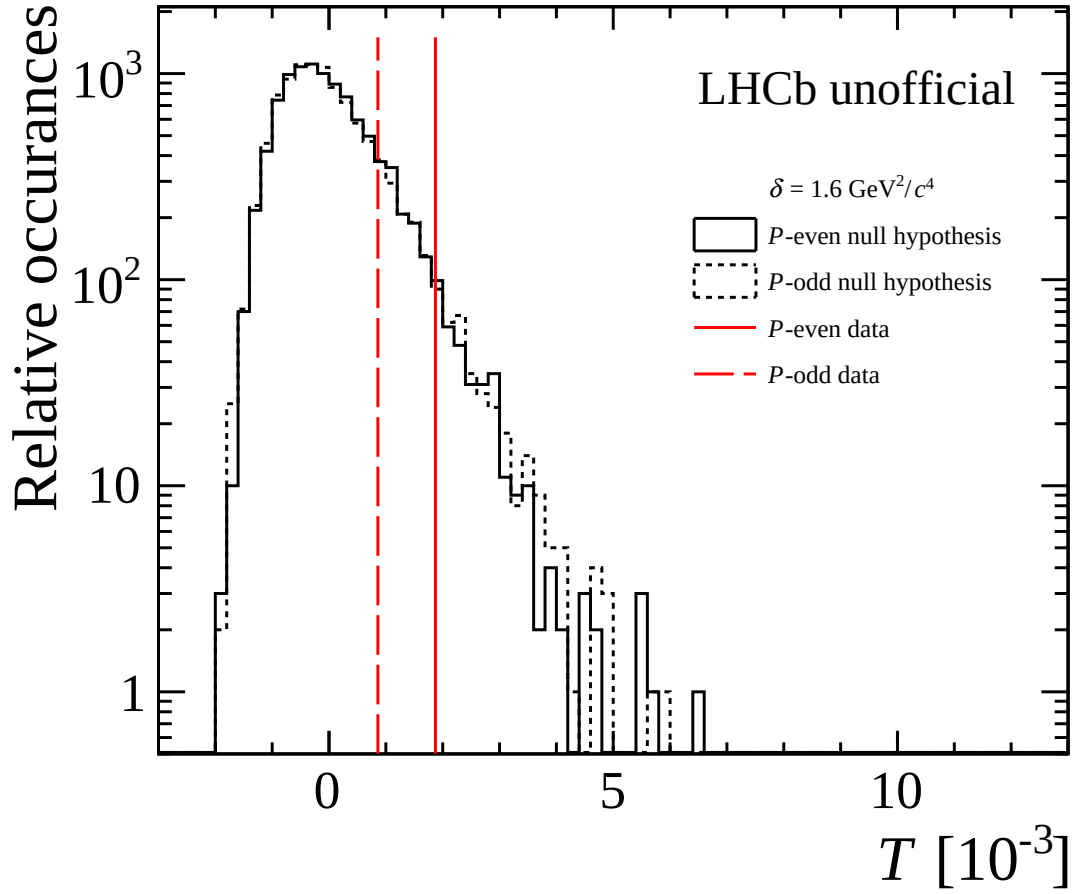


Figure 5.22: The comparison of T-values observed in data with the T-value distribution obtained from CP-symmetric permutations for the distance scale $\delta = 1.6\,\mathrm{GeV}^2/c^4$. The P-even configuration of the energy test is shown as a solid line, and P-odd configuration as a dashed line.

The comparison of T-values observed in data and T-value distributions obtained for permutations for different distance scales δ and test configurations are shown in Figs. 5.22, 5.23, 5.24. A similar comparison of the P violation measurements with three different distance scales is shown in Fig. 5.25.

The theoretical predictions of CP violating asymmetries in this channel go up to 20% [113]. The sensitivity studies described in Sec. 5.5 used models with asymmetries ranging from approximately 5% to 15% and had shown that the test would be sensitive to these effects. The obtained results are consistent with the absence of the effects of similar size in the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays.

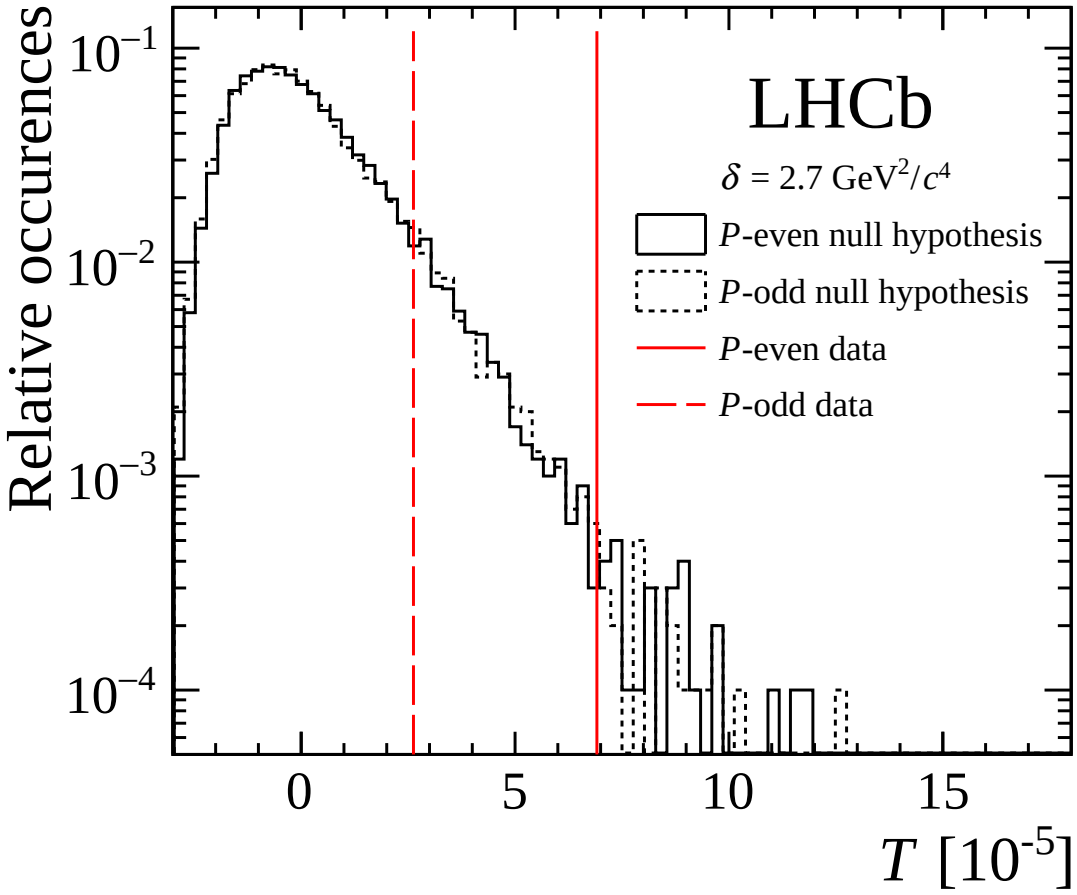


Figure 5.23: The comparison of T-values observed in data with the T-value distribution obtained from CP-symmetric permutations for the distance scale $\delta = 2.7\,\mathrm{GeV}^2/c^4$. The P-even configuration of the energy test is shown as a solid line, and P-odd configuration as a dashed line.

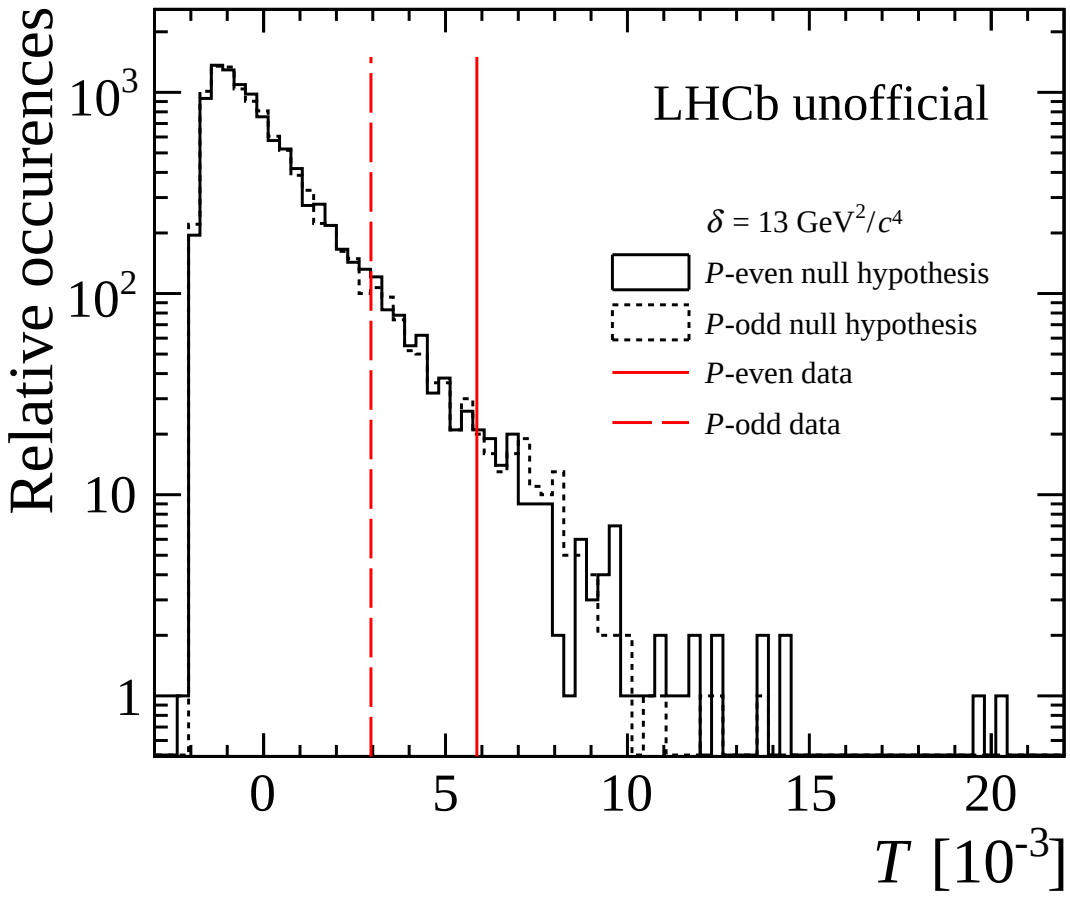


Figure 5.24: The comparison of T-values observed in data with the T-value distribution obtained from CP-symmetric permutations for the distance scale $\delta = 13\,\mathrm{GeV}^2/c^4$. The P-even configuration of the energy test is shown as a solid line, and P-odd configuration as a dashed line.

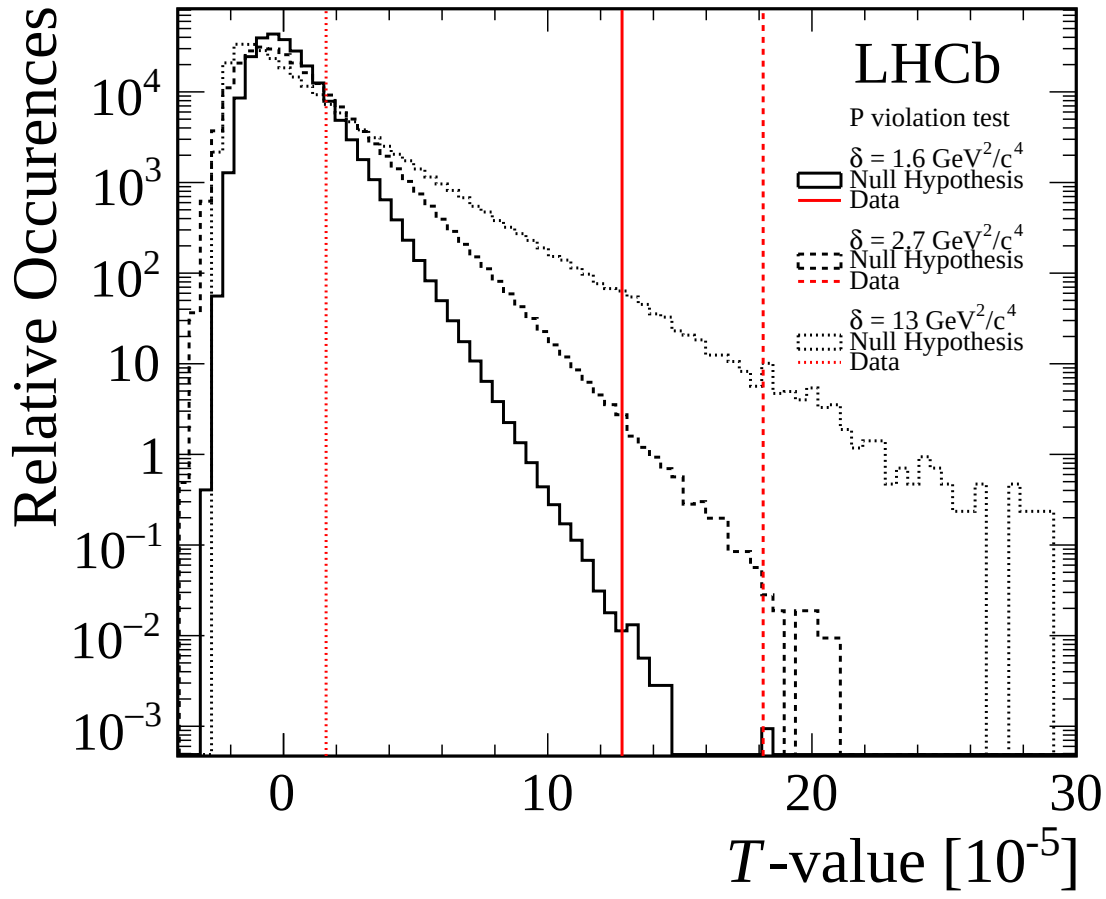


Figure 5.25: The comparison of T-values observed in data with the T-value distribution obtained from permutations for different distance scales in the P violation energy test configuration.

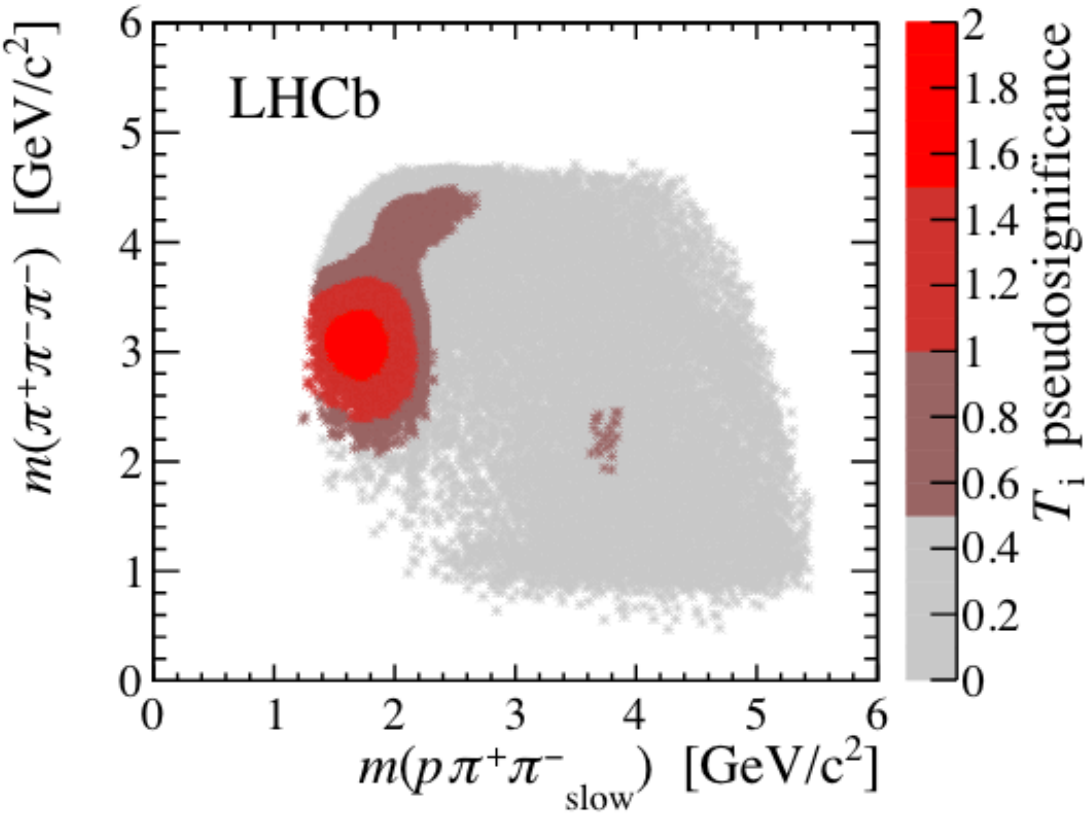


Figure 5.26: The 2D plot showing the results of the visualisation method on the $m(\pi^+\pi^-\pi^+)$ and $m(p\pi^+\pi^-_{slow})$ invariant masses (see Sec. 4.3.3 for details). The bright red regions contribute the most to the obtained T-value.

Additionally, the visualisation technique described in Sec. 4.3.3 is employed to show the phase space regions which contribute the most to the obtained T-values. The results obtained from the visualisation technique is shown in Figs. 5.26, 5.27 and 5.28. It can be seen that the regions of the decay phase space contributing the most to the obtained T-value are the ones containing the dominant resonances in this decay. For example, the peaking structure in Fig. 5.27 with high pseudosignificance corresponds to $\Delta^+(1600)$ and similar resonances, while the highlighted region in Fig. 5.28 contains a_1 and ρ resonances. However, since the overall result is not significant, the pseudosignificances of the visualization technique are hard to interpret and should not be considered as a sign of localized asymmetry. In the scenario of highly significant result this would help to identify the regions with resonances, contributing the most to *CP* violation.

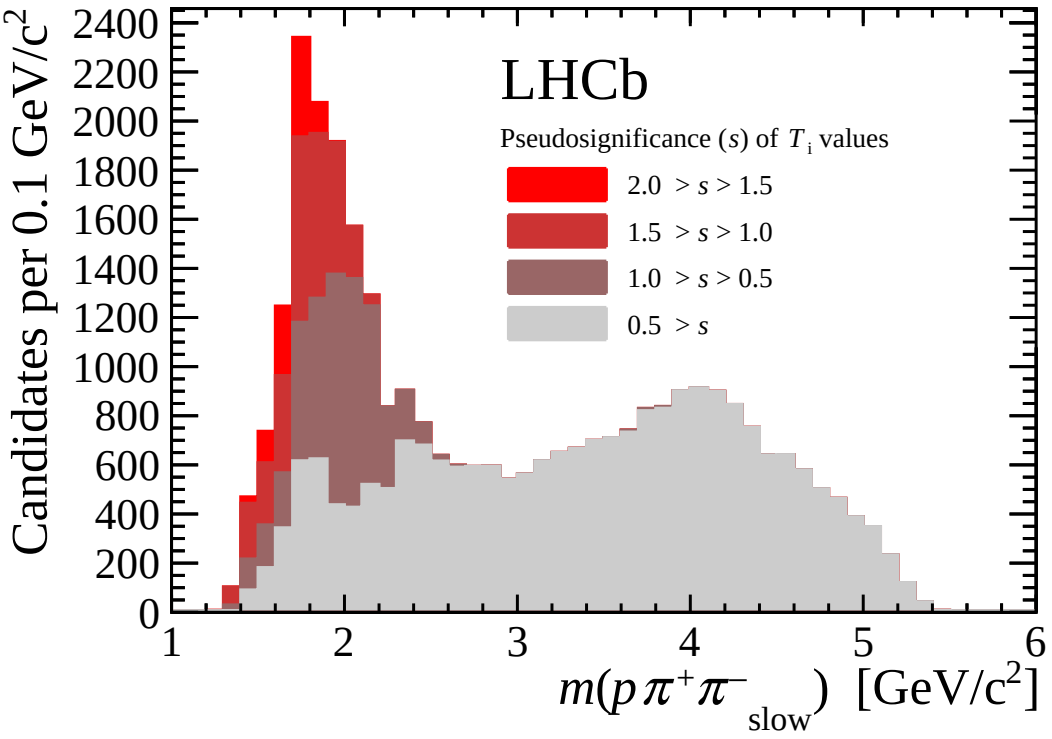


Figure 5.27: The results of the visualisation method on the $m(p\pi^+\pi^-_{slow})$ invariant mass (see Sec. 4.3.3 for details). The bright red regions contribute to the obtained T-value the most.

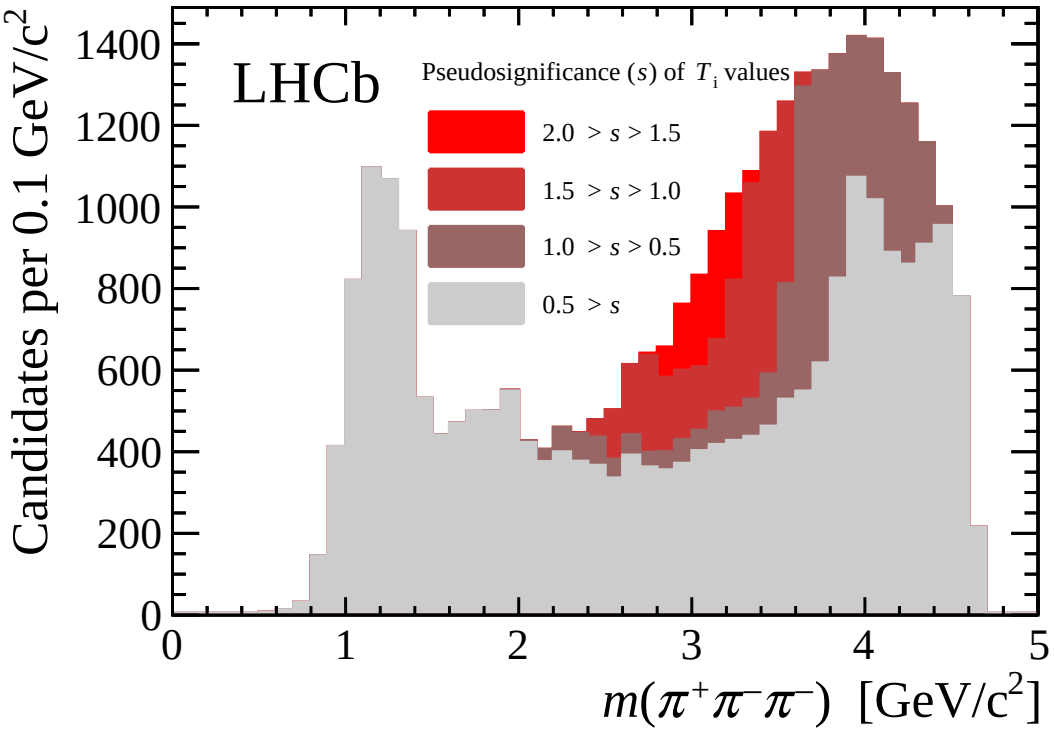


Figure 5.28: The results of the visualisation method on the $m(\pi^-\pi^+\pi^-)$ invariant mass (see Sec. 4.3.3 for details). The bright red regions contribute to the obtained T-value the most.

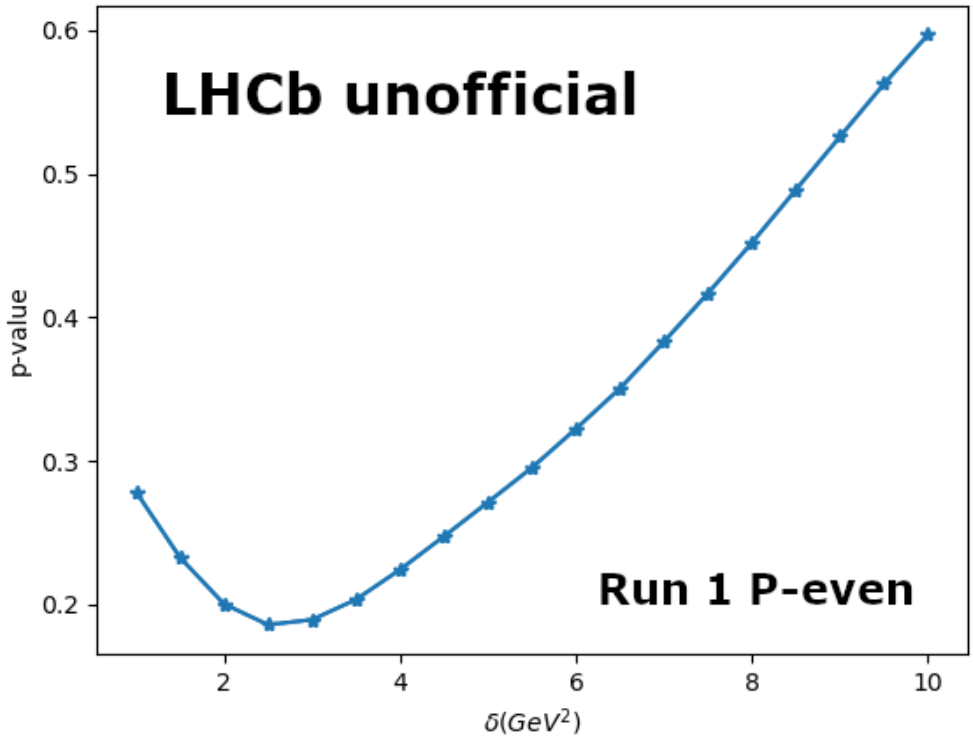


Figure 5.29: The energy test significance dependence on the choice of the distance scale δ. The P-even configuration of the energy test is applied on real Run 1 data sample.

5.8.3 Energy Test Application on Run 1 Dataset

An additional test has been performed after unblinding the full dataset and obtaining the results of the energy test method. As no evidence for CP violation has been obtained with the three selected distance scale δ values, more values could be tested on a subset of the data. This allows to "scan" the sensitivity dependence on the choice of δ. Ten δ values were used in the range $1 - 10\,\mathrm{GeV}^2/c^4$. This method is biasing due to the look-elsewhere effect. If any result was seen it would motivate the choice of δ for the future studies, but they would have to use an independent data set. This is why the study is only performed on the Run 1 sample. This sample will have to be excluded from the future measurements, but the Run 2 sample can still be used.

The P-even and P-odd energy test configurations were applied on the Run 1 data sample (approximately 1/6 of the total signal events). None of the results show high significance, as can be seen in Figs. 5.29, 5.30. While a wide minimum can be seen for the P-even δ trend and a narrow minimum for the P-odd δ trend, these are consistent with random statistical fluctuations and should not be interpreted as real information on the phase space structure. In the light of this, the future studies of this channel with the energy test should be performed with the same values of the distance parameter δ as used in this thesis.

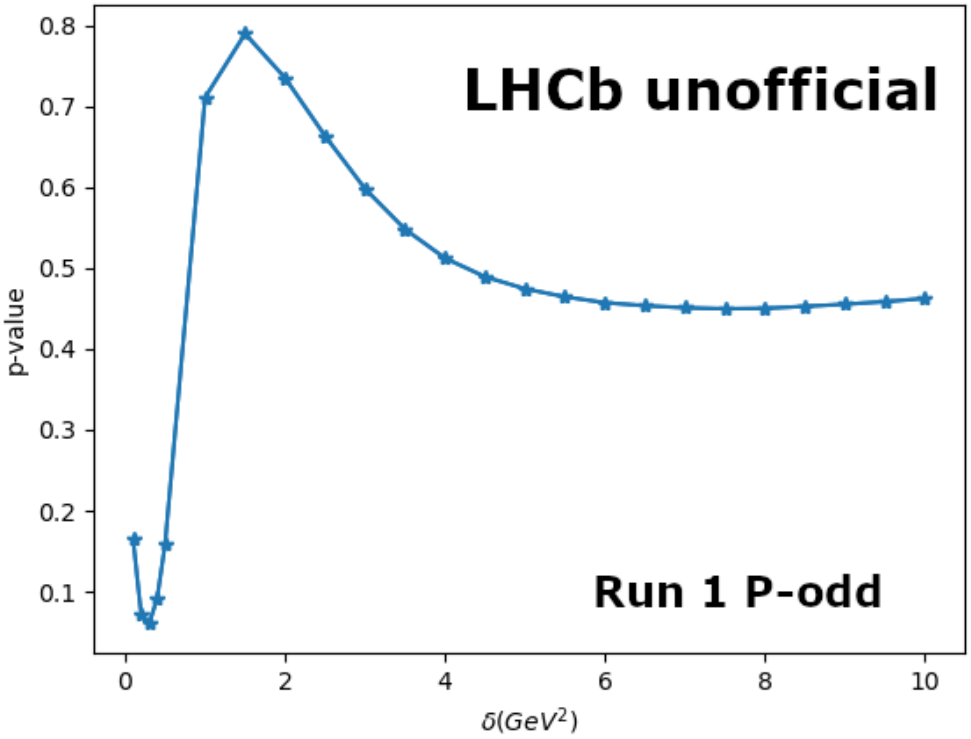


Figure 5.30: The energy test significance dependence on the choice of the distance scale δ. The P-odd configuration of the energy test is applied on real Run 1 data sample.

5.9 Conclusions

A search for *CP* and *P* violation in the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ baryonic decays is performed using two methods: Triple Product Asymmetries and the energy test method. The data used in this analysis has been collected by the LHCb detector in both Run 1 and Run 2 (2011-2017) data-taking periods and corresponds to the integrated luminosity of 6.9 fb^{-1}. There are approximately 36 thousand $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decay candidates, which is a four-fold increase when compared to the previous analysis of this channel [53]. The analysis presented here supersedes the previously published results on Run 1 and significantly expands the scope of the previous study. Both the energy test and the Triple Product Asymmetries method find no significant deviation from the *CP* conservation hypothesis. Furthermore, both methods observe *P* violation in this channel at significance higher than 5σ.

The significance of the effect seen in the previous analysis is reduced, when more statistics are considered. However, the updated analysis shows significance close to 3σ with both methods. This could mean that the new methodology is sensitive to a *CP* violation effect not seen previously. To confirm this hypothesis, more data needs to be taken. The analysed decay channel remains very interesting as with much higher yields to be collected in Run 3 of the LHC it might allow the first observation of *CP* violation in the baryon sector.

Chapter 6

Summary and Outlook

This thesis presents the studies of the radiation damage in the LHCb VELO silicon-strip detector at fluences of up to 6.5×10^{14} 1 MeV n_{eq} cm^2. This thesis is able to present results from the full operational period of the detector across Run 1 and 2 of the LHC. Particular emphasis is based on the analysis of data taken from special CCE scans. This provides a unique body of sensor performance measurements taken in more than 20 CCE scans with 42 *R*-type VELO sensors. The sensors are analysed by dividing each of them into 5 different regions of fluence. Consequently approximately 4200 different measurements are provided. The CCE Scan data is analysed in order to measure the evolution of the EDV, which is compared with the Hamburg model predictions. An approximate agreement between the Hamburg model and EDV of the VELO sensors dependence on fluence is seen. The accurate estimation of the EDV dependence on increasing irradiation and careful temperature control of the VELO detector allowed to beneficially anneal the VELO sensors towards the end of Run 2. To the author's knowledge, this is the first instance of intentional beneficial annealing of silicon sensors of an operating particle detector.

Furthermore, the CCE Scan data is used to estimate the CFE of the VELO sensors. Two methods for rejecting the bad strips are created and applied. The number of bad strips remained relatively low and constant throughout the VELO operation. The RMS method finds approximately 0.4% of dead or noisy strips, while the Frequency method finds around 2% of strips that have significantly higher or lower hit frequency. Different means of isolating separate effects that might reduce the CFE are applied, allowing to isolate the actual radiation damage component. An effect of pick-up of charge on the second metal layer routing lines of the sensor is observed. This leads to a drop in cluster finding efficiency. The evolution of this effect is monitored throughout Run 1 and Run2. For the n^+-on-n sensors this effect is shown to develop quickly, within approximately 0.2 fb^{-1}, and then remain relatively stable throughout the remaining operation over eight years and 9 fb^{-1}. This is consistent with a surface level effect in the silicon sensors.

The VELO detector operated throughout Run 1 and Run 2 of the LHC (2011-2018) without changes to the sensor hardware. During this time, the VELO sensors were exposed to the maximum fluence of approximately 6.5×10^{14} 1 MeV n_{eq} cm^2. The ageing of the sensors due to the radiation damage can clearly be seen in the increase of the EDV as well as a decrease of CFE. This is a bulk radiation damage effect due to damage to the silicon

lattice. At the end of Run 2, most VELO sensors had to be operated at 320V bias voltage in order to remain depleted. However, the VELO detector remained highly efficient in collecting physics data. The analysis of its radiation damage is useful input to the upcoming VELO upgrade, which will collect data in Run 3 and beyond.

The main physics analysis described in this thesis is a search for *CP* violation in four-body baryonic $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays. The sample of these decays is analysed by the LHCb detector in Run 1 and Run 2 of the LHC. The size of the dataset corresponds to approximately $6.9\,\text{fb}^{-1}$ of integrated luminosity or over 36 thousand decay event candidates. Two methods were used to probe the decay phase space for *CP* asymmetry: the Triple Product Asymmetries measurement and the energy test method. The Triple Product Asymmetries method was used in the previous analysis of this channel [53]. In the current analysis, the method uses optimized binning and improves the statistical uncertainty due to an almost sixfold increase in the available signal candidates. No significant deviation from the *CP* conservation hypothesis is found in the phase space of the decay.

The energy test method is applied on the same dataset and run with two different and independent configurations corresponding to *P*-even and *P*-odd *CP* violation measurements. For each configuration three different distance parameters are used. The result with the highest significance is obtained for the *P*-even configuration of the test with the distance parameter value of $2.7\,\text{GeV}^2/c^4$ and corresponds to a 3σ deviation from the *CP* conservation hypothesis. However, when the results for the three distance scales are combined, the total significance of the measurement is reduced to 2.8σ. While the results are lacking significance to claim an evidence for *CP* violation, the tension remains high and the prospect of discovering *CP* violation in baryons with an increased data sample is very tantalising.

Furthermore, both methods are used to search for parity violation in this channel. The Triple Product Asymmetries method find a 5.5σ deviation from the parity conservation hypothesis in the integrated phase space measurement and one of the binned measurements. The energy test is applied to probe for parity violation for the first time. The test finds a significant deviation from the parity conservation hypothesis with two of the three distance scales, corresponding to approximately 5.1σ and 5.3σ for the distance parameter of $1.6\,\text{GeV}^2/c^4$ and $2.7\,\text{GeV}^2/c^4$ respectively.

The radiation damage studies presented in this thesis give important insights on the silicon detector operation in high radiation environments. This contributes to a plethora of ongoing hardware and software improvements that are paramount for a long and successful physics programme of particle physics detectors. The LHCb detector is currently being upgraded. During Runs 3 and 4, it is expected to collect a large amount of data, corresponding to approximately $50\,\text{fb}^{-1}$ of integrated luminosity. This will not only reduce the uncertainties of many previous measurements, but will also allow to test new decay channels, with smaller yields. For example, the physics analysis presented in this thesis is expected to benefit from a much higher signal yield and could lead to the first observation of *CP* violation in baryons. Apart from the $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decay channel

studied here, similar channels can be explored in the future. The methods presented here can be trivially applied on a similar $\Lambda_b^0 \to pK^-K^+\pi^-$ channel. Also, the decays of heavier baryons can be investigated, such as Ξ_b^0 and Ω_b^0. Apart from the *CP* violation studies, the LHCb detector will likely find new heavy composite particle states, measure the CKM matrix parameters with unprecedented precision and hopefully lead to discoveries allowing to probe beyond Standard Model physics [67].

Bibliography

[1] Davis, A., private communication (2019).

[2] R. Aaij et al. "A comprehensive real-time analysis model at the LHCb experiment". In: *JINST* 14.arXiv:1903.01360. 04 (2019), P04006. 15 p. DOI: 10.1088/1748-0221/14/04/P04006. URL: https://cds.cern.ch/record/2665946.

[3] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: https://www.tensorflow.org/.

[4] Jeriek Van den Abeele. "Measuring the Proton Detection Asymmetry at LHCb". In: (2015). https://cds.cern.ch/record/2066952.

[5] H. Albrecht et al. "Search for hadronic b–>u decays". In: *Physics Letters B* 241.2 (1990), pp. 278 –282. ISSN: 0370-2693. DOI: https://doi.org/10.1016/0370-2693(90)91293-K. URL: http://www.sciencedirect.com/science/article/pii/037026939091293K.

[6] R. C. Alig and S. Bloom. "Electron-Hole-Pair Creation Energies in Semiconductors". In: *Phys. Rev. Lett.* 35 (22 1975), pp. 1522–1525. DOI: 10.1103/PhysRevLett.35.1522. URL: https://link.aps.org/doi/10.1103/PhysRevLett.35.1522.

[7] B. Aslan and G. Zech. "New test for the multivariate two-sample problem based on the concept of minimum energy". In: *J. Stat. Comput. Simul.* 75 (2005), pp. 109–119. DOI: 10.1080/00949650410001661440.

[8] B. Aslan and G. Zech. "Statistical energy as a tool for binning-free, multivariate goodness-of -fit tests, two-sample comparison and unfolding". In: *Nucl. Instrum. Meth.* A537 (2005), pp. 626 –636. ISSN: 0168-9002. DOI: 10.1016/j.nima.2004.08.071. URL: http://www.sciencedirect.com/science/article/pii/S0168900204019576.

[9] B. Aubert et al. "Measurement of CP-Violating Asymmetries in B^0 Decays to CP Eigenstates". In: *PRL* 86.12 (2001), pp. 2515–2522. DOI: 10.1103/PhysRevLett.86.2515. arXiv: hep-ex/0102030 [hep-ex].

[10] G. Barbottin and A. Vapaille. "Instabilities in Silicon Devices". In: *Elsevier Science B. V.* (1999).

[11] G. Barrand et al. "GAUDI - A software architecture and framework for building HEP data processing applications". In: *Computer Physics Communications* 140.1 (2001). CHEP2000, pp. 45 –55. ISSN: 0010-4655. DOI: https://doi.org/10.1016/S0010-4655(01)00254-5. URL: http://www.sciencedirect.com/science/article/pii/S0010465501002545.

[12] W. Barter, C. Burr, and C. Parkes. "Calculating p-values and their significances with the Energy Test for large datasets". In: *JINST* 13.04 (2018), P04011. DOI: 10.1088/1748-0221/13/04/P04011. arXiv: 1801.05222 [physics.data-an].

[13] W. J. Barter. "Depletion Voltage Measurements from LHCb". Radiation effects at the LHC experiments and impact on operation and performance. 2018. URL: https://indico.cern.ch/event/695271/contributions/2958676/.

[14] William James Barter et al. "Search for local CP violation in the phase space of the $\Lambda_b^0 \to p\,\pi^-\pi^+\pi^-$ decay". In: (2017). URL: https://cds.cern.ch/record/2257104.

[15] R. Battiston. "The Alpha magnetic spectrometer (AMS): Search for antimatter and dark matter on the International Space Station". In: *Nucl. Phys. Proc. Suppl.* 65 (1998), pp. 19–26. DOI: 10.1016/S0920-5632(97)00970-5. arXiv: hep-ex/9708039 [hep-ex].

[16] I. Bediaga, T. Frederico, and P. Magalhaes. "CP asymmetry from hadronic charm rescattering in $B^\pm \to \pi^-\pi^+\pi^\pm$ decays at the high mass region". In: (Mar. 2020). DOI: 10.1016/j.physletb.2020.135490. arXiv: 2003.10019 [hep-ph].

[17] I. Bediaga et al. "On a *CP* anisotropy measurement in the Dalitz plot". In: *Phys. Rev. D* 80 (9 2009), p. 096006. DOI: 10.1103/PhysRevD.80.096006. URL: https://link.aps.org/doi/10.1103/PhysRevD.80.096006.

[18] Federico Betti. *Observation of CP violation in charm decays at LHCb*. 2019. arXiv: 1905.05428 [hep-ex].

[19] Leo Breiman et al. "Classification and regression trees. Wadsworth". In: *Belmont, CA* (1984).

[20] G. Breit and E. Wigner. "Capture of Slow Neutrons". In: *Phys. Rev.* 49 (7 1936), pp. 519–531. DOI: 10.1103/PhysRev.49.519. URL: https://link.aps.org/doi/10.1103/PhysRev.49.519.

[21] Rene Brun and Fons Rademakers. "ROOT - An object oriented data analysis framework". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 389.1 (1997). New Computing Techniques in Physics Research V, pp. 81 –86. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(97)00048-X. URL: http://www.sciencedirect.com/science/article/pii/S016890029700048X.

[22] S. Chen. "Searches for *CP* violation in multi-body charm decays and studies of radiation damage in the LHCb VELO detector". Presented 04 Dec 2017. 2017. URL: https://cds.cern.ch/record/2302590.

[23] A Chilingarov. "Temperature dependence of the current generated in Si bulk". In: *Journal of Instrumentation* 8.10 (2013), P10003–P10003. DOI: 10.1088/1748-0221/8/10/p10003. URL: https://doi.org/10.1088%2F1748-0221%2F8%2F10%2Fp10003.

[24] S.-K. Choi et al. "Observation of a Narrow Charmoniumlike State in Exclusive $B\pm \to K^{\pm}\pi^{+}\pi^{-}J\psi$ Decays". In: *Physical Review Letters* 91.26 (2003). ISSN: 1079-7114. DOI: 10.1103/physrevlett.91.262001. URL: http://dx.doi.org/10.1103/PhysRevLett.91.262001.

[25] J. H. Christenson et al. "Evidence for the 2π Decay of the K_2^0 Meson". In: *Phys. Rev. Lett.* 13 (4 1964), pp. 138–140. DOI: 10.1103/PhysRevLett.13.138. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.138.

[26] M Clemencic et al. "The LHCb Simulation Application, Gauss: Design, Evolution and Experience". In: *Journal of Physics: Conference Series* 331.3 (2011), p. 032023. DOI: 10.1088/1742-6596/331/3/032023. URL: https://doi.org/10.1088%2F1742-6596%2F331%2F3%2F032023.

[27] ALICE Collaboration. "The ALICE experiment at the CERN LHC". In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[28] ALICE Collaboration. "Upgrade of the ALICE Experiment: Letter Of Intent". In: *J. Phys.* G41 (2014), p. 087001. DOI: 10.1088/0954-3899/41/8/087001.

[29] ATLAS Collaboration. *Letter of Intent for the Phase-I Upgrade of the ATLAS Experiment*. Tech. rep. CERN-LHCC-2011-012. LHCC-I-020. Geneva: CERN, 2011. URL: https://cds.cern.ch/record/1402470.

[30] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (2012), 1–29. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2012.08.020. URL: http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[31] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (2008), S08003–S08003. DOI: 10.1088/1748-0221/3/08/s08003. URL: https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08003.

[32] BaBar Collaboration. "Measurement of Time-Dependent CP Asymmetry in B0 —> c anti-c K(*)0 Decays". In: *Phys. Rev.* D79 (2009), p. 072009. DOI: 10.1103/PhysRevD.79.072009. arXiv: 0902.1708 [hep-ex].

[33] BaBar Collaboration. "Observation of *CP* Violation in the B^0 Meson System". In: *Phys. Rev. Lett.* 87 (9 2001), p. 091801. DOI: 10.1103/PhysRevLett.87.091801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.091801.

[34] BaBar Collaboration. "Search for CP Violation in Neutral D Meson Cabibbo-suppressed Three-body Decays". In: *Phys.Rev.* D78 (2008), p. 051102. DOI: 10.1103/PhysRevD.78.051102. arXiv: 0802.4035 [hep-ex].

[35] BaBar Collaboration. "Search for direct CP violation in singly Cabibbo-suppressed $D^{\pm} \to K^{+}K^{-}\pi^{\pm}$ decays". In: *Phys. Rev.* D87.5 (2013), p. 052010. DOI: 10.1103/PhysRevD.87.052010. arXiv: 1212.1856 [hep-ex].

[36] Belle Collaboration. "Observation of Large *CP* Violation in the Neutral *B* Meson System". In: *Phys. Rev. Lett.* 87 (9 2001), p. 091802. DOI: 10.1103/PhysRevLett.87.091802. URL: https://link.aps.org/doi/10.1103/PhysRevLett.87.091802.

[37] CDF Collaboration. "Measurements of Direct *CP*-Violating Asymmetries in Charmless Decays of Bottom Baryons". In: *Phys. Rev. Lett.* 113 (24 2014), p. 242001. DOI: 10.1103/PhysRevLett.113.242001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.242001.

[38] CMS Collaboration. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[39] CMS Collaboration. "The CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 3.08 (2008), S08004–S08004. DOI: 10.1088/1748-0221/3/08/s08004. URL: https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08004.

[40] LHCb Collaboration. "A measurement of the *CP* asymmetry difference in $\Lambda_c^+ \to pK^-K^+$ and $p\pi^-\pi^+$ decays". In: *JHEP* 03.CERN-EP-2017-316. LHCB-PAPER-2017-044 (2017), 182. 21 p. DOI: 10.1007/JHEP03(2018)182. URL: http://cds.cern.ch/record/2298698.

[41] LHCb Collaboration. "Amplitude analysis of $B^{\pm} \to \pi^{\pm}K^{+}K^{-}$ decays". In: *Phys. Rev. Lett.* 123.23 (2019), p. 231802. DOI: 10.1103/PhysRevLett.123.231802. arXiv: 1905.09244 [hep-ex].

[42] LHCb Collaboration. "Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC". In: *Journal of Instrumentation* 14.04 (2019), P04013–P04013. DOI: 10.1088/1748-0221/14/04/p04013. URL: https://doi.org/10.1088%2F1748-0221%2F14%2F04%2Fp04013.

[43] LHCb Collaboration. "Determination of the $X(3872)$ Meson Quantum Numbers". In: *Phys. Rev. Lett.* 110 (22 2013), p. 222001. DOI: 10.1103/PhysRevLett.110.222001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.110.222001.

[44] LHCb Collaboration. "First observation of *CP* violation in the decays of B_s^0 mesons". In: *Phys. Rev. Lett.* 110.22 (2013), p. 221601. DOI: 10.1103/PhysRevLett.110.221601. arXiv: 1304.6173 [hep-ex].

[45] LHCb collaboration. *LHCb detector*. URL: https://cds.cern.ch/record/1087860.

[46] LHCb Collaboration. "LHCb Detector Performance". In: *Int. J. Mod. Phys.* A30.07 (2015), p. 1530022. DOI: 10.1142/S0217751X15300227. arXiv: 1412.6352 [hep-ex].

[47] LHCb Collaboration. *LHCb PID Upgrade Technical Design Report*. Tech. rep. CERN-LHCC-2013-022. LHCB-TDR-014. 2013. URL: https://cds.cern.ch/record/1624074.

[48] LHCb Collaboration. *LHCb Trigger and Online Upgrade Technical Design Report*. Tech. rep. CERN-LHCC-2014-016. LHCB-TDR-016. 2014. URL: https://cds.cern.ch/record/1701361.

[49] LHCb Collaboration. *LHCb VELO Upgrade Technical Design Report*. Tech. rep. CERN-LHCC-2013-021. LHCB-TDR-013. 2013. URL: https://cds.cern.ch/record/1624070.

[50] LHCb Collaboration. "Measurement of B^0, B_s^0, B^+ and Λ_b^0 production asymmetries in 7 and 8 *TeV* pp collisions". In: *Phys. Lett.* B774 (2017), p. 139. DOI: 10.1016/j.physletb.2017.09.023. arXiv: 1703.08464 [hep-ex].

[51] LHCb Collaboration. "Measurement of *CP* violation in the phase space of $B^\pm \to K^+K^-\pi^\pm$ and $B^\pm \to \pi^+\pi^-\pi^\pm$ decays". In: *Phys. Rev. Lett.* 112 (2014), p. 011801. DOI: 10.1103/PhysRevLett.112.011801. arXiv: 1310.4740 [hep-ex].

[52] LHCb Collaboration. "Measurement of *CP* violation in the phase space of $B^\pm \to K^\pm\pi^+\pi^-$ and $B^\pm \to K^\pm K^+K^-$ decays". In: *Phys. Rev. Lett.* 111 (2013), p. 101801. DOI: 10.1103/PhysRevLett.111.101801. arXiv: 1306.1246 [hep-ex].

[53] LHCb Collaboration. "Measurement of matter-antimatter differences in beauty baryon decays". In: *Nature Physics* 13 (2017), p. 391. DOI: 10.1038/nphys4021. arXiv: 1609.05216 [hep-ex].

[54] LHCb Collaboration. "Measurements of *CP* violation in the three-body phase space of charmless $B^\pm$ decays". In: *Phys. Rev. D* 90.11 (2014), p. 112004. DOI: 10.1103/PhysRevD.90.112004. arXiv: 1408.5373 [hep-ex].

[55] LHCb Collaboration. "Model-independent search for *CP* violation in $D^0 \to K^-K^+\pi^+\pi^-$ and $D^0 \to \pi^-\pi^+\pi^-\pi^+$ decays ". In: *Phys. Lett.* B726 (2013), p. 623. DOI: 10.1016/j.physletb.2013.09.011. arXiv: 1308.3189 [hep-ex].

[56] LHCb Collaboration. "Observation of a narrow pentaquark state, $P_c(4312)^+$, and of two-peak structure of the $P_c(4450)^+$". In: *Phys. Rev. Lett.* 122.22 (2019), p. 222001. DOI: 10.1103/PhysRevLett.122.222001. arXiv: 1904.03947 [hep-ex].

[57] LHCb Collaboration. "Observation of *CP* Violation in Charm Decays". In: *Phys. Rev. Lett.* 122 (21 2019), p. 211803. DOI: 10.1103/PhysRevLett.122.211803. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.211803.

[58] LHCb Collaboration. "Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays". In: *Phys. Rev. Lett.* 115 (7 2015), p. 072001. DOI: 10.1103/PhysRevLett.115.072001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.115.072001.

[59] LHCb Collaboration. "Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays". In: *Phys. Rev. Lett.* 115 (2015), p. 072001. DOI: 10.1103/PhysRevLett.115.072001. arXiv: 1507.03414 [hep-ex].

[60] LHCb Collaboration. "Observation of Several Sources of *CP* Violation in $B^+ \to \pi^+\pi^+\pi^-$ Decays". In: *Phys. Rev. Lett.* 124.3 (2020), p. 031801. DOI: 10.1103/PhysRevLett.124.031801. arXiv: 1909.05211 [hep-ex].

[61] LHCb Collaboration. "Observation of the decay $\Lambda_b^0 \to pK\mu^+\mu^-$ and a search for CP violation". In: *Journal of High Energy Physics* 2017.6 (2017), p. 108.

[62] LHCb Collaboration. "Observation of the Doubly Charmed Baryon Ξ_{cc}^{++}". In: *Phys. Rev. Lett.* 119 (11 2017), p. 112001. DOI: 10.1103/PhysRevLett.119.112001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.112001.

[63] LHCb Collaboration. "Observation of the $\Lambda_b^0 \to J/\psi p\pi^-$ decay". In: *JHEP* 07 (2014), p. 103. DOI: 10.1007/JHEP07(2014)103. arXiv: 1406.0755 [hep-ex].

[64] LHCb Collaboration. "Observations of $\Lambda_b^0 \to \Lambda K^+\pi^-$ and $\Lambda_b^0 \to \Lambda K^+K^-$ decays and searches for other Λ_b^0 and Ξ_b^0 decays to $\Lambda h^+h'^-$ final states". In: *Journal of High Energy Physics* 2016.5 (2016), p. 81.

[65] LHCb Collaboration. "Performance of the LHCb muon system". In: *JINST* 8.LHCB-DP-2012-002. CERN-LHCb-DP-2012-002. LHCb-DP-2012-002 (2012), P02022. 32 p. DOI: 10.1088/1748-0221/8/02/P02022. URL: https://cds.cern.ch/record/1492807.

[66] LHCb Collaboration. "Performance of the LHCb Outer Tracker". In: *Journal of Instrumentation* 9.01 (2014), P01002–P01002. DOI: 10.1088/1748-0221/9/01/p01002. URL: https://doi.org/10.1088%2F1748-0221%2F9%2F01%2Fp01002.

[67] LHCb Collaboration. "Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era". In: (2018). arXiv: 1808.08865.

[68] LHCb Collaboration. "Search for *CP* violation and observation of *P* violation in $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ decays". In: (2019). arXiv: 1912.10741 [hep-ex].

[69] LHCb Collaboration. "Search for CP violation in $D^0 \to \pi^-\pi^+\pi^0$ decays with the energy test". In: *Phys.Lett.* B740 (2015), p. 158. DOI: 10.1016/j.physletb.2014.11.043. arXiv: 1410.4170 [hep-ex].

[70] LHCb Collaboration. "Search for *CP* violation in $D^+ \to K^-K^+\pi^+$ decays". In: *Phys. Rev.* D84 (2011), p. 112008. DOI: 10.1103/PhysRevD.84.112008. arXiv: 1110.3970 [hep-ex].

[71] LHCb Collaboration. "Search for *CP* violation in $D^0 \to \pi^-\pi^+\pi^0$ decays with the energy test". In: *Phys. Lett.* B740 (2015), p. 158. DOI: 10.1016/j.physletb.2014.11.043. arXiv: 1410.4170 [hep-ex].

[72] LHCb Collaboration. "Search for CP violation in $\Lambda_b^0 \to pK^-$ and $\Lambda_b^0 \to p\pi^-$ decays". In: *Physics Letters B* 787 (2018), pp. 124 –133. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2018.10.039. URL: http://www.sciencedirect.com/science/article/pii/S0370269318308104.

[73] LHCb Collaboration. "Search for *CP* violation in the decay $D^+ \to \pi^-\pi^+\pi^+$". In: *Phys. Lett.* B728 (2014), p. 585. DOI: 10.1016/j.physletb.2013.12.035. arXiv: 1310.7953 [hep-ex].

[74] LHCb Collaboration. "Search for *CP* violation in the phase space of $D^0 \to \pi^+\pi^-\pi^+\pi^-$ decays". In: *Phys. Lett.* B769 (2017), p. 345. DOI: 10.1016/j.physletb.2017.05.062. arXiv: 1612.03207 [hep-ex].

[75] LHCb Collaboration. "Search for *CP* violation using *T*-odd correlations in $D^0 \to K^+K^-\pi^+\pi^-$ decays". In: *JHEP* 10 (2014), p. 005. DOI: 10.1007/JHEP10(2014)005. arXiv: 1408.1299 [hep-ex].

[76] LHCb Collaboration. "Search for *CP* violation using triple product asymmetries in $\Lambda_b^0 \to pK^-\pi^+\pi^-$, $\Lambda_b^0 \to pK^-K^+K^-$, and $\Xi_b^0 \to pK^-K^-\pi^+$ decays". In: *JHEP* 08 (2018), p. 039. DOI: 10.1007/JHEP08(2018)039. arXiv: 1805.03941 [hep-ex].

[77] LHCb Collaboration. "Searches for Λ_b^0 and Ξ_b^0 decays to $K_S^0 p\pi^-$ and $K_S^0 pK^-$ final states with first observation of the $\Lambda_b^0 \to K_S^0 p\pi^-$ decay". In: *JHEP* 04 (2014), p. 087. DOI: 10.1007/JHEP04(2014)087. arXiv: 1402.0770 [hep-ex].

[78] LHCb Collaboration. "The LHCb detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[79] LHCb VELO Collaboration. "Performance of the LHCb Vertex Locator". In: *JINST* 9 (2014), P09007. DOI: 10.1088/1748-0221/9/09/P09007. arXiv: 1405.7808 [physics.ins-det].

[80] LHCb VELO Collaboration. "Radiation damage in the LHCb Vertex Locator". In: *JINST* 8 (2013), P08002. DOI: 10.1088/1748-0221/8/08/P08002. arXiv: 1302.5259 [physics.ins-det].

[81] NA48 Collaboration. "A New measurement of direct CP violation in two pion decays of the neutral kaon". In: *Phys. Lett.* B465 (1999), pp. 335–348. DOI: 10.1016/S0370-2693(99)01030-8. arXiv: hep-ex/9909022 [hep-ex].

[82] PDG Collaboration. "Review of particle physics". In: *Chin. Phys.* C40 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.

[83] PDG Collaboration. "Review of Particle Physics". In: *Phys. Rev. D* 98 (3 2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.030001.

[84] PDG Collaboration. "Review of Particle Physics". In: *Phys. Rev.* D98.3 (2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001.

[85] Planck Collaboration. "Planck 2015 results. XIII. Cosmological parameters". In: *Astron. Astrophys.* 594 (2016), A13. DOI: 10.1051/0004-6361/201525830. arXiv: 1502.01589 [astro-ph.CO].

[86] Super-Kamiokande Collaboration. "Evidence for oscillation of atmospheric neutrinos". In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. arXiv: hep-ex/9807003.

[87] P. Collins. *VELO CCE track categories*. 2018.

[88] G. A. Cowan, D. C. Craik, and M. D. Needham. "RapidSim: an application for the fast simulation of heavy-quark hadron decays". In: *Comput. Phys. Commun.* C214 (2017), pp. 239–246. DOI: 10.1016/j.cpc.2017.01.029. arXiv: 1612.07489 [hep-ex].

[89] Kyle S. Cranmer. "Kernel estimation in high-energy physics". In: *Comput. Phys. Commun.* 136 (2001), pp. 198–207. DOI: 10.1016/S0010-4655(00)00243-5. arXiv: hep-ex/0011057 [hep-ex].

[90] MJ Cush. "Standard Model of Elementary Particles". In: (2019). General Photo. URL: https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

[91] Gauthier Durieux. Personal communication. July 2017.

[92] Gauthier Durieux. "CP violation in multibody decays of beauty baryons". In: *JHEP* 10 (2016), p. 005. DOI: 10.1007/JHEP10(2016)005. arXiv: 1608.03288 [hep-ph].

[93] Gauthier Durieux and Yuval Grossman. "Probing CP violation systematically in differential distributions". In: (). arXiv: 1508.3054 [hep-ex].

[94] Gauthier Durieux and Yuval Grossman. "Probing CP violation systematically in differential distributions". In: *Phys. Rev.* D92.7 (2015), p. 076013. DOI: 10.1103/PhysRevD.92.076013. arXiv: 1508.03054 [hep-ph].

[95] "Elementary particle interactions". In: (2008). General Photo. URL: https://en.wikipedia.org/wiki/File:Elementary_particle_interactions.svg.

[96] Christian Elsässer. *$\bar{b}b$ production angle plots*. URL: https://lhcb.web.cern.ch/lhcb/speakersbureau/html/bb_ProductionAngles.html.

[97] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001. URL: https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08001.

[98] A. S. Folkestad. "TCAD simulations of the 2nd metal layer". 2017. URL: https://indico.cern.ch/event/607255/contributions/2682156/.

[99] Enrico Franco, Satoshi Mishima, and Luca Silvestrini. "The Standard Model confronts CP violation in $D_0 \to \pi^+\pi^-$ and $D_0 \to K^+K^-$". In: *Journal of High Energy Physics* (2012). DOI: 10.1007/JHEP05(2012)140. URL: https://doi.org/10.1007/JHEP05(2012)140.

[100] Richard L. Garwin, Leon M. Lederman, and Marcel Weinrich. "Observations of the Failure of Conservation of Parity and Charge Conjugation in Meson Decays: the Magnetic Moment of the Free Muon". In: *Phys. Rev.* 105 (4 1957), pp. 1415–1417. DOI: 10.1103/PhysRev.105.1415. URL: https://link.aps.org/doi/10.1103/PhysRev.105.1415.

[101] Murray Gell-Mann. "A Schematic Model of Baryons and Mesons". In: *Phys. Lett.* 8 (1964), pp. 214–215. DOI: 10.1016/S0031-9163(64)92001-3.

[102] Thomas P. S. Gillam and Christopher G. Lester. "Biased bootstrap sampling for efficient two-sample testing". In: (2018). [JINST13,no.12,P12014(2018)]. DOI: 10.1088/1748-0221/13/12/P12014. arXiv: 1810.00335 [physics.data-an].

[103] Sheldon L. Glashow. "Partial-symmetries of weak interactions". In: *Nuclear Physics* 22.4 (1961), pp. 579 –588. ISSN: 0029-5582. DOI: https://doi.org/10.1016/0029-5582(61)90469-2. URL: http://www.sciencedirect.com/science/article/pii/0029558261904692.

[104] Michael Gronau and Jonathan L. Rosner. "Triple product asymmetries in K, $D_{(s)}$ and $B_{(s)}$ decays". In: *Phys. Rev.* D84 (2011), p. 096013. DOI: 10.1103/PhysRevD.84.096013. arXiv: 1107.1232 [hep-ph].

[105] LHCb RICH Group. "Performance of the LHCb RICH detector at the LHC". In: *Eur. Phys. J.* C73 (2013), p. 2431. DOI: 10.1140/epjc/s10052-013-2431-9. arXiv: 1211.6759 [physics.ins-det].

[106] G. Haefeli et al. "The LHCb DAQ interface board TELL1". In: *Nucl. Instrum. Meth.* A560 (2006), pp. 494–502. DOI: 10.1016/j.nima.2005.12.212.

[107] Julie Haffner. "The CERN accelerator complex. Complexe des accélérateurs du CERN". In: (2013). General Photo. URL: http://cds.cern.ch/record/1621894.

[108] D. Hanneke, S. Fogwell, and G. Gabrielse. "New Measurement of the Electron Magnetic Moment and the Fine Structure Constant". In: *Phys. Rev. Lett.* 100 (12 2008), p. 120801. DOI: 10.1103/PhysRevLett.100.120801. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.120801.

[109] J. Harrison. "Radiation damage studies at the LHCb VELO detector and searches for lepton flavour and baryon number violating tau decays". PhD thesis. University of Manchester, 2014.

[110] Jonathan Robert Harrison. "Radiation damage studies in the LHCb VELO detector and searches for lepton flavour and baryon number violating tau decays". Presented 16 05 2014. 2014. URL: https://cds.cern.ch/record/1712972.

[111] Andrew Hickling et al. *Use of IT (current vs temperature) scans to study radiation damage in the LHCb VELO*. Tech. rep. LHCb-PUB-2011-021. CERN-LHCb-PUB-2011-021. Geneva: CERN, 2011. URL: https://cds.cern.ch/record/1392353.

[112] A. Hoecker et al. "TMVA - Toolkit for Multivariate Data Analysis". In: *ArXiv Physics e-prints* (Mar. 2007). eprint: physics/0703039.

[113] Y. K. Hsiao and C. Q. Geng. "Direct CP violation in Λ_b^0 decays". In: *Phys. Rev.* D91 (2015), p. 116007. DOI: 10.1103/PhysRevD.91.116007. arXiv: 1412.1899 [hep-ph].

[114] Sebastian Jäger et al. "Charming New B-Physics". In: (2019). arXiv: 1910.12924 [hep-ph].

[115] V. Khachatryan et al. "Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data". In: *Nature* 522 (2015), p. 68. DOI: 10.1038/nature14474. arXiv: 1411.4413 [hep-ex].

[116] C. Kittel. *Introduction to solid state physics*. 8th ed. J. Wiley, 2005.

[117] Makoto Kobayashi and Toshihide Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X. DOI: 10.1143/PTP.49.652. eprint: https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf. URL: https://doi.org/10.1143/PTP.49.652.

[118] Makoto Kobayashi and Toshihide Maskawa. "CP Violation in the Renormalizable Theory of Weak Interaction". In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652.

[119] V. Alan Kostelecký and Neil Russell. "Data tables for Lorentz and *CPT* violation". In: *Rev. Mod. Phys.* 83 (1 2011), pp. 11–31. DOI: 10.1103/RevModPhys.83.11. URL: https://link.aps.org/doi/10.1103/RevModPhys.83.11.

[120] Tommaso Lari. "Test beam results of ATLAS Pixel sensors". In: *Semiconductor pixel detectors for particles and X-rays. Proceedings, International Workshop, PIXEL2002, Carmel, USA, September 9-12, 2002*. 2002. arXiv: hep-ex/0210045 [hep-ex]. URL: http://www.slac.stanford.edu/econf/C020909/tlpaper.pdf.

[121] Johan Luisier. "Performance of LHCb Silicon Tracker Detector in the LHC". In: *Physics Procedia* 37 (2012). Proceedings of the 2nd International Conference on Technology and Instrumentation in Particle Physics (TIPP 2011), pp. 851 –858. ISSN: 1875-3892. DOI: https://doi.org/10.1016/j.phpro.2012.04.097. URL: http://www.sciencedirect.com/science/article/pii/S1875389212017798.

[122] S Löchner and M Schmelling. *The Beetle Reference Manual - chip version 1.3, 1.4 and 1.5*. Tech. rep. LHCb-2005-105. CERN-LHCb-2005-105. Geneva: CERN, 2006. URL: https://cds.cern.ch/record/1000429.

[123] J Mans et al. *CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter*. Tech. rep. CERN-LHCC-2012-015. CMS-TDR-10. Additional contact persons: Jeffrey Spalding, Fermilab, spalding@cern.ch, Didier Contardo, Universite Claude Bernard-Lyon I, contardo@cern.ch. 2012. URL: https://cds.cern.ch/record/1481837.

[124] *Measurement of charged particle multiplicities and densities in pp collisions at $\sqrt{s} = 7$ TeV in the forward region*. Data Collection. 2014. DOI: 10.17182/hepdata.63498. URL: https://doi.org/10.17182/hepdata.63498.

[125] A. Merli. "Search for CP violation in the angular distribution of $\Lambda_b^0 \to p\pi^-\pi^+\pi^-$ baryon decays and a proposal for the search of heavy baryon EDM with bent crystal at LHCb". Presented 23 May 2019. 2019. URL: https://cds.cern.ch/record/2687820.

[126] Michael Moll. "Radiation damage in silicon particle detectors: Microscopic defects and macroscopic properties". PhD thesis. Hamburg U., 1999. URL: http://www-library.desy.de/cgi-bin/showprep.pl?desy-thesis99-040.

[127] Stephen Myers. *The LEP Collider, from design to approval and commissioning*. John Adams' Lecture. Delivered at CERN, 26 Nov 1990. Geneva: CERN, 1991. DOI: 10.5170/CERN-1991-008. URL: http://cds.cern.ch/record/226776.

[128] Kozo Nakamura et al. "Formation process of grown-in defects in Czochralski grown silicon crystals". In: *Journal of Crystal Growth* 180.1 (1997), pp. 61 –72. ISSN: 0022-0248. DOI: https://doi.org/10.1016/S0022-0248(97)00206-6. URL: http://www.sciencedirect.com/science/article/pii/S0022024897002066.

[129] E. Noether. "Invariante Variationsprobleme". ger. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1918 (1918), pp. 235–257. URL: http://eudml.org/doc/59024.

[130] A. Papadelis et al. *First study of the VELO detector noise performance at Point 8*. Tech. rep. LHCb-INT-2009-005. CERN-LHCb-INT-2009-005. Geneva: CERN, 2009. URL: https://cds.cern.ch/record/1170205.

[131] Chris Parkes et al. "On model-independent searches for direct CP violation in multi-body decays". In: *Journal of Physics G: Nuclear and Particle Physics* 44.8 (2017), p. 085001. DOI: 10.1088/1361-6471/aa75a5. URL: https://doi.org/10.1088%2F1361-6471%2Faa75a5.

[132] T. Peltola. "Simulation of radiation-induced defects". In: Dec. 2015. DOI: DOI:https://doi.org/10.22323/1.254.0031.

[133] W. G. Pfann, C. E. Miller, and J. D. Hunt. "New Zone Refining Techniques for Chemical Compounds". In: *Review of Scientific Instruments* 37.5 (1966), pp. 649–652. DOI: 10.1063/1.1720273. eprint: https://doi.org/10.1063/1.1720273. URL: https://doi.org/10.1063/1.1720273.

[134] Alessio Piucci. "The LHCb Upgrade". In: *Journal of Physics: Conference Series* 878 (2017), p. 012012. DOI: 10.1088/1742-6596/878/1/012012. URL: https://doi.org/10.1088%2F1742-6596%2F878%2F1%2F012012.

[135] Muriel Pivk and Francois R. Le Diberder. "sPlot: A statistical tool to unfold data distributions". In: *Nucl.Instrum.Meth.* A555 (2005), pp. 356–369. DOI: 10.1016/j.nima.2005.08.106. arXiv: physics/0402083 [physics.data-an].

[136] Jeffrey D. Richman. "An Experimenter's Guide to the Helicity Formalism". In: (1984). URL: http://inspirehep.net/record/202987/.

[137] T. Rohe et al. "Position dependence of charge collection in prototype sensors for the CMS pixel detector". In: *IEEE Transactions on Nuclear Science* vol. 51, no. 3 (2004), pp. 1150–1157. DOI: 10.1109/TNS.2004.829487.

[138] Andrei D Sakharov. "Violation ofCPin variance,Casymmetry, and baryon asymmetry of the universe". In: *Soviet Physics Uspekhi* 34.5 (1991), pp. 392–393. DOI: 10.1070/pu1991v034n05abeh002497.

[139] Abdus Salam. "Weak and Electromagnetic Interactions". In: *Conf. Proc.* C680519 (1968), pp. 367–377. DOI: 10.1142/9789812795915_0034.

[140] G. Sarpis. "Depletion Voltage Measurements from LHCb". Radiation effects in the LHCb VELO Run 1+2. 2019. URL: https://indico.cern.ch/event/769192/contributions/3309454/.

[141] D. K. Schroder. "Semiconductor Material and Device Characterization". In: *John Wiley & Sons* (1990).

[142] Julian Schwinger. "The Theory of Quantized Fields. I". In: *Phys. Rev.* 82 (6 1951), pp. 914–927. DOI: 10.1103/PhysRev.82.914. URL: https://link.aps.org/doi/10.1103/PhysRev.82.914.

[143] W. Shockley. "Currents to Conductors Induced by a Moving Point Charge". In: *Journal of Applied Physics* 9.10 (1938), pp. 635–636. DOI: 10.1063/1.1710367. eprint: https://doi.org/10.1063/1.1710367. URL: https://doi.org/10.1063/1.1710367.

[144] Tomasz Skwarnicki. "A study of the radiative cascade transitions between the Upsilon-prime and Upsilon resonances". DESY-F31-86-02. PhD thesis. Institute of Nuclear Physics, Krakow, 1986.

[145] P. R. Turner. *VELO Module Production - Sensor Testing*. Tech. rep. LHCb-2007-072. CERN-LHCb-2007-072. revised version submitted on 2007-11-06 12:07:30. Geneva: CERN, 2007. URL: https://cds.cern.ch/record/1067138.

[146] V. A. J. van Lint et al. "Mechanisms of radiation effects in electronic materials. Volume 1". In: *NASA STI/Recon Technical Report A* 81 (Jan. 1980), p. 13073.

[147] Mika Vesterinen and Simen Hellesund. *Measuring pion and kaon tracking efficiencies and asymmetries using charm decays with one partially reconstructed track*. Tech. rep. LHCb-INT-2014-041. CERN-LHCb-INT-2014-041. Geneva: CERN, 2014. URL: https://cds.cern.ch/record/1954312.

[148] A. D. Webber. "Radiation damage studies in the LHCb VELO detector and measurement of the flavour-specific asymmetry in semileptonic B-decays". Presented 21 Feb 2013. 2013. URL: http://cds.cern.ch/record/1596046.

[149] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[150] Mike Williams. "Observing *CP* violation in many-body decays". In: *Phys.Rev.* D84 (2011), p. 054015. DOI: 10.1103/PhysRevD.84.054015. arXiv: 1105.5338 [hep-ex].

[151] Lincoln Wolfenstein. "Parametrization of the Kobayashi-Maskawa Matrix". In: *Phys. Rev. Lett.* 51 (21 1983), pp. 1945–1947. DOI: 10.1103/PhysRevLett.51.1945. URL: https://link.aps.org/doi/10.1103/PhysRevLett.51.1945.

[152] C. S. Wu et al. "Experimental Test of Parity Conservation in Beta Decay". In: *Phys. Rev.* 105 (4 1957), pp. 1413–1415. DOI: 10.1103/PhysRev.105.1413. URL: https://link.aps.org/doi/10.1103/PhysRev.105.1413.

[153] Yung-Chun Wu and Yi-Ruei Jhan. "Introduction of Synopsys Sentaurus TCAD Simulation". In: June 2018, pp. 1–17. ISBN: 978-981-10-3065-9. DOI: 10.1007/978-981-10-3066-6_1.