FEDERAL UNIVERSITY OF RIO DE JANEIRO

PHYSICS INSTITUTE

Contribution to the Analysis of the $B_c^+ \to K^+K^-\pi^+$ Decay in the LHCb Experiment

Julian Stenzel Martins

Masters Dissertation presented to the Physics Graduate Program at the Physics Institute of the Federal University of Rio de Janeiro - UFRJ, as part of the necessary requisites to obtain a Master of Science (Physics) degree.

Advisor: Leandro Salazar de Paula

Co-Advisor: Adlène Hicheur

Rio de Janeiro

March 2016

S379c Stenzel Martins, Julian

Contribution to the Analysis of the $B_c^+ \to K^+K^-\pi^+$ Decay in the LHCb Experiment / Julian Stenzel Martins – Rio de Janeiro, 2016.

178f.

Orientador: Leandro Salazar de Paula

Coorientador: Adlène Hicheur

Dissertação (mestrado) - Universidade Federal do Rio de Janeiro, Instituto de Física, Programa de Pós-Graduação em Física, 2016.

1. Física do B. 2. Méson Bc. 3. Modelo Padrão. 4. LHC. 5. LHCb. I. Salazar de Paula, Leandro, orient. II. Hicheur, Adlène, coorient. III. Contribution to the Analysis of the $B_c^+ \to K^+K^-\pi^+$ Decay in the LHCb Experiment

Contribution to the Analysis of the $B^+_c \to K^+K^-\pi^+$ Decay in the LHCb Experiment

Julian Stenzel Martins

Leandro Salazar de Paula
Adlene Hicheur

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Física, Instituto de Física, da Universidade Federal do Rio de Janeiro – UFRJ, como parte dos requisitos necessários à obtenção do título de Mestre em Ciências (Física).

Aprovada por:

.....
Leandro Salazar de Paula
(Presidente e Orientador)

.....
Adlene Hicheur
(Co-Orientador)

.....
Andre Sznadjer

.....
Arthur Marques Moraes

Rio de Janeiro, RJ – Brasil
Março de 2016

Resumo

Contribution to the Analysis of the $B_c^+ \to K^+K^-\pi^+$ Decay in the LHCb Experiment

Julian Stenzel Martins

Orientador: Leandro Salazar de Paula

Coorientador: Adlène Hicheur

Resumo da Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Física do Instituto de Física da Universidade Federal do Rio de Janeiro - UFRJ, como parte dos requisitos necessários à obtenção do título de Mestre em Ciências (Física).

Os estudos dos decaimentos do méson B_c^+ para estados finais contendo $K^+K^-\pi^+$, seja por diagramas de aniquilação direta ou diagramas contendo mésons D^0 intermediários, são potencialmente úteis para sondar as previsões do Modelo Padrão. Poucas medições já foram feitas em relação ao méson B_c^+, por conta de sua baixa probabilidade de produção em aceleradores de partículas. A observação deste tipo de decaimento requer energia de colisão e luminosidade que só se tornaram possíveis com a construção do LHC.

Esta dissertação detalha os passos da análise do decaimento $B_c^+ \to K^+K^-\pi^+$, que procura encontrar um sinal mensurável do méson B_c^+ usando dados coletados pelo detector LHCb entre os anos 2010 e 2012. A seleção de eventos foi concluída e o modelo de ajuste completo foi determinado.

Palavras-chave: Física do B, méson B_c, Modelo Padrão, LHC, LHCb.

Abstract

Contribution to the Analysis of the $B_c^+ \to K^+K^-\pi^+$ Decay in the LHCb Experiment

Julian Stenzel Martins

Advisor: Leandro Salazar de Paula

Co-Advisor: Adlène Hicheur

Abstract of the Masters Dissertation presented to the Physics Graduate Program at the Physics Institute of the Federal University of Rio de Janeiro - UFRJ, as part of the necessary requisites to obtain a degree as a Master of Science (Physics).

The studies of the B_c^+ meson's decays to final states containing $K^+K^-\pi^+$, occurring through either direct annihilation diagrams or diagrams containing an intermediate D^0 meson, are potentially useful to probe Standard Model predictions. However, very few measurements regarding the B_c^+ meson have been performed, due to its low probability of being produced in particle accelerators. A measurement of these decays is only possible with the collision energies and the unprecedented amount of statistical data collected at the LHC.

This dissertation details the steps in the $B_c^+ \to K^+K^-\pi^+$ analysis that attempts to perform a measurement of the B_c^+ signal, using data collected by the LHCb detector from 2010 to 2012. The event selection is completed and the full fit model is determined, in preparation for a final measurement.

Keywords: B physics, B_c meson, Standard Model, LHC, LHCb.

Acknowledgments

It is no exaggeration to say that the two years of study and work that culminated in the writing of this dissertation were the hardest of my life. I was pushed beyond what I previously thought I was capable of and in retrospect, I can happily say that the experience has made me grow to new heights as both a person and a professional. However, none of it would have been possible without the contributions and support of individuals to whom I am now, and always will be, incredibly grateful.

First off, I would like to thank my wonderful parents, Kristine and Jorge. Your support, advice, and love helped me over these two years in ways that I can never fully express. I am incredibly lucky and truly thankful to be your son, and I am aware of that every single day.

In the hardest moments of this course, no one made more of a difference to my spiritual well-being than my amazing, gorgeous and nerdy girlfriend, Naiara. Your two months by my side were more valuable to me than you will ever know. No matter how far away we are from each other, I carry you in my heart every day, and with you I have the strength to overcome any obstacle. Mozi, I love you more than anything in the world, believe it.

What can I say to Bruno, Let, and Nandes? I am so fortunate to have had the opportunity to share my home with you for those all too brief months. Bruno, for years you have been so much more than just a cousin to me. You are my oldest and best friend, and I love you as I have loved few other people in my life. I can never thank you for all you have done. Let, even though we didn't know each other as well as I would have liked before sharing an apartment, those months were more than enough for me to realize what an adventurous, happy, and responsible person you are, and I am truly glad to have

been able to get to know you. Nandes, you have gone from a person who terrified me over the Internet to one of the people who I most enjoy talking to. I am so grateful for your intelligence, humor, and our hours of FIFA, which always helped after the bad days. I want to thank the three of you, from the bottom of my heart, for the most magical months of my life. With you, I became a new and better person, and I will never forget our incredible adventure for as long as I live.

To Anderson, Oscar, and Daniela, thank you for always helping whenever I needed it. You all saved me many times over, and made me feel comfortable in a place where I felt like I would never belong.

Max and Molly, even though you are cats and probably have little to no conception of a world outside the apartment that is your domain, your presence helped me through terrible times. Although you obviously will never know or understand my gratitude, I still feel that I must register it here, for my own peace of mind. You are more than animals, you are members of my family, invaluable companions with the amazing ability of bringing me joy in even the hardest times. Thank you, Mom, for bringing them into my life.

To Leandro and Adlène, no words will ever truly convey my gratitude. Without your unending patience and goodwill, none of this would have been remotely possible. Thank you both for always answering my ridiculous amount of questions with clarity and candor. Thank you for pushing me whenever it was necessary, and for understanding when I was having trouble. Thank you for being integral parts of this, the greatest accomplishment of my life, and for helping me take the next step. This period was incredibly difficult and frustrating for me, and at times I felt truly hopeless. However, I wouldn't trade it for anything, and for that, again, thank you.

Last but not least, I would like to acknowledge the CAPES, CNPq and FAPERJ for their financial support throughout these two years of my masters course.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The Standard Model of Particle Physics has described with great success the known fundamental particles and their interactions for more than 40 years. However, the Standard Model is not a wholly complete theory, as it fails to provide explanations for certain phenomena. For this reason studies are currently being undertaken to fully understand the Model's limitations, in order to help drive the development of extensions to the theory.

One of the useful avenues for testing the predictions of the Standard Model are studies of rare particle decays. The LHCb Experiment was designed for the studies of B mesons, which are highly unstable particles containing the b-quark, one of the heaviest quarks known. Although these kinds of mesons are produced very rarely, they can be reconstructed with great precision in the LHCb detector due to their decay characteristics. The B_c meson is of particular interest to probe Standard Model predictions, as it is the highest-mass particle of the B family and very few measurements exist as yet for its decay processes.

This study describes a statistical analysis, that used data collected by the LHCb Collaboration during the Large Hadron Collider's Run I, which took place from 2010 to 2012. In this dissertation, the B_c^+ meson's annihilation decays to $K^+K^-\pi^+$ final states are discussed, and the full selection process and fit model determination is detailed, in preparation for a measurement attempt. Chapter 2 presents a historical and theoretical overview, and closes with the motivation behind the study of B_c decays. In Chapter 3, the

Large Hadron Collider and LHCb detector are described in detail. Chapter 4 discusses all the steps of the analysis leading up to (but not including) a final measurement, and the final conclusion to the dissertation is shown on Chapter 5.

Chapter 2

Theoretical Overview

This chapter provides a theoretical introduction and background for the topics that are relevant to the subject of this dissertation. In section 2.1, a historical background and review of the Standard Model is given, followed by a brief discussion on the importance of B-physics and the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ decay itself (section 2.2).

2.1 The Standard Model

The Standard Model (SM) of Particle Physics is a theory that was developed and worked upon continually in the second half of the twentieth century by both theoretical and experimental particle physicists. It regards the *electromagnetic*, *weak*, and *strong* nuclear interactions, and also serves to classify and organize all known subatomic particles. The Standard Model is one of the most successful theories in all of Physics, having enormous and significant triumph in explaining and accurately predicting decades' worth of experimental results.

However, it is not a wholly complete theory of fundamental interactions, leaving some unexplained and poorly understood phenomena. The *gravitational* interaction is notably absent from the Standard Model, as the full theory of gravitation as per Einstein's General Relativity has proven challenging to incorporate into the Model's framework. The theory also does not account for *Dark Matter/Energy*, the non-zero mass of neutrinos,

and doesn't provide any manner of calculating the masses of quarks and leptons. A number of theoretical speculations have emerged that attempt to explain these deficiencies, which are referred to as Physics *Beyond* the Standard Model (BSM). While some of these new theories include extensions of the current Standard Model (Supersymmetry is a well-known example), others, such as Superstring theory, are entirely new. Even though no direct experimental evidence yet exists to support these theories, they are currently one of the most active and commented-upon areas of research in both theoretical and experimental physics. The details of this section were extracted primarily from [1, 2, 3], and references for mentioned discoveries are located throughout the text.

2.1.1 Historical Background

The Standard Model's inception and development are of fundamental importance to modern particle physics, and were consequences of a series of theories and experimental observations that began more than a hundred years ago[1]. Discussing all the achievements that led to (or contributed to) the Model is impractical and not in the scope of this dissertation, so only the most important and famous of these will be mentioned.

The concept of subatomic particle physics can be said to have been born in the closing years of the nineteenth century. The notion that matter was composed of small and indivisible particles called *atoms* had been well-known for almost a hundred years (due to John Dalton's work in chemistry [4]) when Joseph John Thomson showed in 1897 [5] how cathode rays were comprised of negatively-charged particles that were much smaller than atoms. He called these particles *corpuscles* (later permanently renamed to *electrons)*, and they were the first subatomic particles to be discovered and identified[2]. Thomson believed that electrons were constituents of atoms which, being neutral particles, would have to have some source of positive charge in order to offset the electrons. Thomson proposed the *plum pudding* model, which would be shown to be incorrect a little over a decade

[1] Or thousands of years ago, if Democritus of Abdera and the Greek atomists are considered.
[2] For which Thomson would be awarded the Nobel Prize in Physics nine years later.

later.

In 1908, Hans Geiger and Ernest Marsden performed, at Ernest Rutherford's behest, a now-famous scattering experiment [6] in which α-particles were fired at a thin sheet of gold foil. The results of the experiment are well-known: most of the α-particles passed through the foil without the slightest perturbation, and the few that were scattered did so at large angles. Rutherford would conclude [7] that the scattered particles had encountered small and heavy objects, which led to the realization that the *nucleus* of the atom occupied a very small portion of the atom's total volume, and all the positive charge was located there, separated from the surrounding electrons (Rutherford atomic model). Thus the second subatomic particle, positively-charged and located in the nucleus of atoms, would be identified as the *proton* by Rutherford himself.

A few years later in 1913, Niels Bohr would propose his own atomic model in which electrons circled (or "orbited") the nucleus at certain distances, being held there by the electromagnetic Coulomb force[3]. Bohr would combine his model with the still-young quantum theory to calculate the hydrogen spectrum in a way that successfully agreed with experimental results [8, 9, 10], but his orbital model was incomplete. Considering only the presence of protons and electrons, the *masses* of atoms made little sense, and were poorly understood. This would remain a mystery for another eighteen years, until James Chadwick proved the existence of the *neutron* in 1932[4], a neutral particle with the same mass as the proton, that was also located in the atomic nucleus [11].

This "classical" trio of subatomic particles would be joined by the *photon* in 1923, after Arthur Compton's experiments confirming light's particle-like behavior in the subatomic realm [12]. His experiment (and the resulting famous scattering formula) put a close to a decades-long re-kindling of the debate on the nature of light, initiated by Max Planck's *quantization* of electromagnetic radiation in 1900 [13, 14] and expanded upon by Albert

[3]Bohr would receive a Nobel Prize in Physics eight years later, partly for this reason.
[4]For which he would also receive a Nobel Prize in Physics, in 1935.

Einstein's groundbreaking work on the *photoelectric effect* in 1905[5] [15]. The field of particle physics was emerging, and a mystery on the nature of the atomic nucleus would set the stage for the numerous discoveries that would occur in the following decades.

At this point, the three-particle architecture of the atom presented a problem concerning the stability of the nucleus. If all atomic nuclei were composed of only protons and neutrons, classical electromagnetic repulsion would attempt to separate the protons in the nucleus from each other. The apparent stability of the nucleus indicated that there must be another interaction at work that was more powerful than the electromagnetic interaction at those small distances, and it would be eventually called the *strong* interaction. Hideki Yuwaka proposed a theory on the strong interaction in 1935 [16], with the interaction occurring through a short-range field that was mediated by a massive force-carrying particle, of which there was no experimental evidence.

In 1937, studies of cosmic rays undertaken by two separate groups (Neddermeyer-Anderson [17] and Street-Stevenson [18]) provided evidence of particles that were similar to Yukawa's proposed force carrier. However, there were some key differences, mainly regarding how the particle did not interact strongly with the atomic nucleus, something that should occur if the particle were actually the mediator of the strong force. The Bristol-based group under Cecil Powell, which included Giuseppe Occhialini and César Lattes, would identify Yukawa's particle in 1947 [19], upon coming to the conclusion that cosmic rays carried two different particles. The first of these would be called the *muon*, and it was lighter and had a longer lifetime than the second discovered particle, which fit Yukawa's predictions[6]. Yukawa's particle was the *pion* (π), and although it was not actually the mediator of the strong interaction, it stands as the first *meson* to be discovered. The muon, on the other hand, behaved in every way like a heavy-mass electron, and was the first *lepton* to be identified after the electron (while protons and

[5]Planck would receive the Nobel Prize in 1918 for his pioneering work in quantum mechanics, and Einstein would receive his own Nobel Prize in 1921 for his paper on the photoelectric effect.

[6]And for which he would receive a Nobel Prize in 1949.

neutrons were the first *baryons*, see subsection 2.1.2).

A few years earlier in 1928, Paul Dirac proposed the equation that now bears his name [20], which described free electrons with energy given by the relativistic energy formula. Dirac's equation had the peculiar (and problematic) feature of allowing a solution with *negative* energy for every solution with positive energy. He attempted to explain this with his well-known *negative-energy sea* theory, in which all possible negative energy states were already "occupied" by electrons that, upon being promoted to a positive energy state, would leave a "hole" in the sea. This hole would represent an entirely different particle, one with positive energy and charge. It soon became apparent that this positively-charged particle would necessarily have the same mass as the electron, and a particle with these characteristics was not known at the time.

Still studying cosmic rays, Carl Anderson discovered in 1932 a particle with exactly the properties that Dirac required, which would be called the *positron* [21]. Richard Feynman and Ernst Stueckelberg's more modern (and popularly accepted) interpretation of Dirac's negative energy solutions [22, 23] formulated the "holes" as positive-energy solutions of a *different* particle than the electron. The newly-discovered positron satisfied this explanation, which in fact introduced the fundamental idea that every particle had its own *antiparticle* with equal mass but opposite quantum numbers (of which charge is an example)[7]. Therefore the positron was in fact the *antielectron*, and the *antiproton* and *antineutron* would be discovered later in Berkeley, in 1955 [24] and 1956 [25]. Particles (matter) and antiparticles (antimatter) cannot coexist, and annihilate each other upon contact producing photons and energy. The observable universe is matter-dominated and the reason for this matter-antimatter *asymmetry* remains a mystery to this day, being in fact one of the main goals of the LHCb experiment, but it is not included in the scope of the analysis which is the subject of this dissertation.

Another physics milestone in the 1930s was the proposal of the *neutrino* by Wolfgang

[7]It is also important to note that some neutral particles are *their own* antiparticles, such as the photon.

Pauli, due to problems in the determination of electron energy in nuclear beta decays, where measured electron energies varied from the values predicted using simple conservation of energy. Pauli suggested that the electron's energy spectrum was due to the presence of another particle, one that was electrically neutral, had very small mass, and that possessed the energy that was missing from the energy conservation prediction. Neutrinos rarely interact with matter, and direct experimental evidence of their existence was only provided in the 1950s by Cowan and Reines [26], while the *antineutrino* was proved to exist as a separate particle (i.e. the neutrino wasn't its own antiparticle) in 1962 by Lederman, Schwartz, Steinberger et al [27].

In 1947, the known elementary particles were protons, neutrons, electrons, muons, and pions, and apart from the muon (whose role was still not well understood), they all fit into existing theory and were well accepted. This would change at the end of the year with Rochester and Butler's cloud chamber experiments showing the existence of a neutral massive particle [28] that decayed into two pions. Initially called "V-particles" due to the shape of their decay pion tracks in the cloud chamber photographs, these particles would soon be called *kaons*. They were the first in a wave of new particles that started to be discovered in the 1950s[8], and that were collectively referred to as "strange" particles. The strange particles were produced via the strong interaction but decayed via the *weak* interaction, unlike the particles that came before them, which prompted Kazuhiko Nishijima [29] and Murray Gell-Mann [30] to independently propose the existence of a new quantum property (like charge or spin) for these particles. Gell-Mann famously coined the name *strangeness* (S) for this property, which is not conserved in any particle decay mediated by the weak interaction.

The flood of new particles produced and identified in the first accelerator years showed that the realm of elementary particles was potentially vastly populated, and in 1961

[8]Mainly due to the use of the first particle accelerators, which allowed production and observation of particles in a laboratory setting.

Gell-Mann introduced a way to organize the known particles according to their quantum characteristics [31]. In a manner not dissimilar to Mendeleev's periodic table, Gell-Mann's *Eightfold Way* sorted particles into geometric arrays in which different values of strangeness and charge lay upon different lines of the arrays. The discovery of the Ω^- particle in 1964 [32] laid to rest doubts on the model's accuracy, as its existence and quantum properties had all been predicted by Gell-Mann years before.

In the same year as the Ω^-'s discovery, Gell-Mann and George Zweig independently proposed an explanation for the Eightfold Way's geometric organization [33, 34, 35] the *quark* model, that states that particles that interact via the strong force (*hadrons*) are composed of smaller particles called quarks, that have three types (or *flavors*). These three flavors were initially called *up* (u), *down* (d), and *strange* (s), they all had their own values of electric charge (that were fractions of the electron charge) and strangeness, and also had corresponding *antiquarks*. The quark model's *composition rules* for the known elementary particles were quite simple. Baryons (such as protons, neutrons, the Ω^-, for example) are composed of three quarks[9], while mesons (pions, kaons, among others) are composed of a quark-antiquark pair. Using the composition rules, particles can be re-expressed as their constituent quarks, for example: uud is a proton, $u\bar{d}$ is a π^+, and $s\bar{u}$ is a K^-.

However, the quark model was not without its problems, facing uncertainty in the years following its introduction due to numerous experiments' failure to "see" individual quarks in an experimental setting. Early *deep* inelastic scattering experiments performed in the late 1960s attempted to ascertain the interior composition of protons in a manner similar to Rutherford's strategy with the atom, and the results themselves were also similar [36]. Most of the incident particles passed through the proton with no interaction, while a few were scattered at large angles. Therefore, the positive charge of the proton was also concentrated in a small fraction of its volume, but in this case the experimental

[9] Antibaryons are likewise composed by three antiquarks.

evidence suggested that there were *three* concentrations of charge, not just one, which was predicted by the model.

Another issue with the quark model was due to Pauli's famous exclusion principle, which at this time had already been expanded from only electrons to incorporate all *fermions*, such as quarks. Some particle's quark compositions make little sense in view of the exclusion principle[10], which led to Moo-Young Han and Yoichiro Nambu's proposal of *color* charge for quarks [37]. In this formulation, quarks can have one of three "colors": *blue*, *green*, and *red*, and any naturally occurring elementary particle would have a *colorless* quark composition. Colorless cases include mixtures of three same-flavor quarks of different colors and quark-antiquark pairs (which can have color charges of blue-*anti*blue, for instance), as well as any combination of these two cases.

The Standard Model of Particle Physics' current formulation was finalized soon after experimental confirmation of the existence of quarks, and the following decades saw the discovery of new leptons and quark flavors that were predicted by the model, such as Burton Richter and Samuel Ting's independent discoveries of the J/ψ meson in 1974 [38, 39] that indicated the existence of the *charm* (c) quark, and the *tau* lepton's (a massive "cousin" of the electron) discovery the following year by Martin Perl *et al* at the Stanford Nuclear Accelerator Center (SLAC) [40], which added to its worldwide acceptance. The current six-quark configuration was completed when the fifth quark flavor, the *bottom* (b), was found in 1977 upon discovery of the Υ meson [41], and the sixth (and final, so far) flavor, the *top* (t), was found in 1995 [42] when particle accelerators finally achieved the energies necessary to produce them.

The final elements in the list of discoveries that led to (and further completed) the Standard Model are the *gauge bosons*, the elementary particles that serve as mediators of the four fundamental interactions. The *gluon*, mediator of the strong interaction, was proposed by Gell-Mann in 1962 [43] and observed in 1978 by the PLUTO Collaboration

[10]The Ω^- for instance is composed of three s quarks.

at Hamburg [44]. Only five years later, Sheldon Glashow, Steven Weinberg and Abdus Salam's proposal of the *electroweak* interaction [45, 46], that attempted to unify the electromagnetic and weak forces, was shown to be correct upon the discoveries in 1983 of the $W^{\pm}$ [47] and Z [48] bosons at CERN, which completed the Standard Model's group of gauge bosons[11].

2.1.2 Summary of the Standard Model

Symmetries play a fundamental role in modern physics. Noether's theorem implies that symmetries in a physical system are related to the conservation of a physical property in that system, and the symmetries themselves can be classified into two types. *Global* symmetries are valid in all points of spacetime, while *local* symmetries have different symmetry transformations related to their spacetime coordinates. Furthermore, symmetries can be *continuous* if the physical system's invariance holds under continuous change of its geometry, or *discrete* if the invariance only holds in certain configurations. These types of symmetries are present in modern physical theories[12], and local discrete symmetries are especially important due to their role in the derivation of *gauge* theories.

Gauge theories are a general type of field theory, in which the theory's Lagrangean is *invariant* under a certain group of local transformations. *Quantized* gauge fields are the basis for all physical theories that successfully describe the dynamics of elementary particles, where the *quanta* of the fields are entirely different kinds of particles called *gauge* bosons. Quantum ElectroDynamics (QED) was the first quantum gauge theory to describe the interactions of charged fermions[13], and it is based on the most simple symmetry group, U(1), having therefore one gauge field and one associated gauge boson, which are the electromagnetic field and the photon.

In the 1950's, Chen-Ning Yang and Robert Mills proposed a way to describe the fields

[11] The photon, mediator of the electromagnetic interaction, was of course already well-known.

[12] Special relativity is based on Lorentz symmetry, and is an example of a well-known physical theory which is based on a global symmetry.

[13] Electrons, in this case.

of protons and neutrons using a gauge theory, based on *isospin* symmetry[49]. Yang-Mills theory uses special unitary groups (SU(N)) to describe elementary particle behavior, and it was used successfully in the formulation of Quantum ChromoDynamics (QCD), the quantum gauge theory for the strong interaction. The *quarks* which are described by QCD possess color charge in lieu of electric charge, which is used to define a local symmetry using the SU(3) gauge group. The gauge field for QCD thus contains eight gauge bosons, called gluons. Yang-Mills theory is also used for the weak interaction, which uses a third gauge group, SU(2), and possesses three unique gauge bosons of its own. These three quantum gauge theories and their internal symmetries are the framework for what is called the Standard Model of particle physics.

The Standard Model was established in the 1960s and 1970s and is one of the most successful theories in all of physics. It is a Yang-Mills gauge quantum field theory that contains the internal symmetries of the SU(3) $\times$ SU(2) $\times$ U(1) unitary product group, and therefore encompasses the strong, electromagnetic and weak interactions. The Model also describes all the fundamental particles, which are divided into three groups[14]: leptons, quarks, and the aforementioned gauge bosons.

Leptons are fermions, and they are particles whose main characteristic is that they do not "feel" the strong interaction at all (due to their lack of color charge). The six existing leptons[15] are further classified into two different types based on their charge, and form three *generations* of lepton pairs, each containing one charged and one neutral particle. Being fermions, leptons obey Fermi-Dirac statistics and the Pauli exclusion principle, and also possess the broad characteristics of mass and spin.

The charged leptons are referred to as *electron*-like, and the first and most famous of these is evidently the electron (e^-) itself. Its neutral pair is the electron neutrino (ν_e), and together they form the first lepton generation. The other two generations are pairs

[14]The much sought-after Higgs particle stands alone as a technical "fourth" group, apart from the other three which contain multiple particles each.

[15]Considering that each lepton has a respective *anti*lepton, technically there are *twelve* leptons in all.

of similar nature: the second generation is composed of the muon (μ^-) and the muon neutrino (ν_μ), while the third is constituted by the tau (τ^-) and the tau neutrino (ν_τ). Each of the charged leptons is alike in every way except for their masses, which increase in subsequent generations. Therefore the tau holds the highest mass, and both it and the muon are unstable particles that rapidly decay. The neutrinos, on the other hand, are all extremely lightweight particles that interact very little with matter, making precise measurements of their masses a challenging task. Table 2.1 summarizes the six leptons and their quantum numbers.

Flavor	Mass ($\mathrm{GeV/c^2}$)	Charge	L_e	L_μ	L_τ
e	0.00051	-1	1	0	0
ν_e	$< 7 \times 10^{-9}$	0	1	0	0
μ	0.10566	-1	0	1	0
ν_μ	< 0.0003	0	0	1	0
τ	1.77682	-1	0	0	1
ν_τ	< 0.03	0	0	0	1

Table 2.1: Summary of the three lepton generations and their quantum properties which, after charge, are the electron, muon, and tau numbers, respectively.

Quarks are the only elementary particles to have color charge, and for this reason they are also the only particles that interact via all four fundamental forces. They share certain similarities with leptons, such as a three-generation organizational scheme and the fact that both types of particles are fermions, but in most other ways they differ greatly. Each quark flavor can have one of three "colors" and therefore each quark generation, which is composed of two flavors, has a total of 6 separate quarks. These particles are also unique in the sense that they cannot be observed in isolation. Due to color confinement, quarks are held within the particles that they form, called hadrons. Since particles must be *colorless*, color charge also dictates the possible ways to combine quarks, which further

classifies the hadrons into baryons and mesons.

Baryons are composed of three quarks with three different colors[16](*anti*baryons are likewise composed of three antiquarks with three different anticolors), the main examples of which are protons and neutrons. Mesons, such as pions and kaons, contain a quark-antiquark pair, which must also possess the same respective color charge in order to engender a color-neutral particle. They are easily produced in high-energy interactions and are *all* unstable (unlike baryons), rapidly decaying into leptons and photons. Another notable difference between baryons and mesons is due to the spin which they gain from their constituent quarks: baryons, having three, are fermions; while mesons are *bosons*, obeying entirely different quantum statistics. Table 2.2 summarizes the quark generations and their quantum numbers, in which the central mass values are extracted from [50].

Flavor	Mass (GeV/c^2)	Charge	U	D	S	C	B	T
d	$0.0048^{+0.0005}_{-0.0003}$	$-1/3$	0	-1	0	0	0	0
u	$0.0023^{+0.0007}_{-0.0003}$	$2/3$	1	0	0	0	0	0
s	0.095 ± 0.005	$-1/3$	0	0	-1	0	0	0
c	1.275 ± 0.025	$2/3$	0	0	0	1	0	0
b	4.18 ± 0.03	$-1/3$	0	0	0	0	-1	0
t	173.21 ± 0.87	$2/3$	0	0	0	0	0	1

Table 2.2: Summary of the three quark generations and their quantum properties which, after charge, are "upness", "downness", strangeness, charm, beauty, and truth, respectively.

As quarks and leptons are both fermions, and they make up all known matter. Both types of particles can interact via the weak force, and a property of this force brings up key characteristics of these particles, and also of our observable universe. This property is the violation of *parity* symmetry, proposed by Tsung Dao Lee and Chen Ning Yang in 1956 [51] and confirmed the following year by Chien-Shiung Wu et al through the study of

[16]By analogy with light, where white can be obtained by combining blue, green and red.

beta decays [52]. In essence, only left-"handed" particles participate in weak interactions, and the three quark and lepton generations are thus organized into left-handed *doublets* and right-handed *singlets*, as is represented below.

$$\begin{pmatrix} \nu_l \\ l^- \end{pmatrix}_L (l^-)_R \begin{pmatrix} q_u \\ q_d \end{pmatrix}_L (q_u)_R \, (q_d)_R \tag{2.1}$$

R and L represent right and left-handed chirality, l and ν represent the lepton-neutrino pair of a lepton generation, and q_u and q_d represent the quark pair of a quark generation. A noteworthy aspect of this representation is the lack of a right-handed singlet for neutrinos, as these kinds of particles[17] are still hypothetical.

The Standard Model's third particle group are the gauge bosons, the mediators of the electromagnetic, weak, and strong forces. Being bosons, these particles have integer spin values and thus obey Bose-Einstein statistics. The multiplicity of the gauge bosons for each interaction is based on the special unitary groups that describe them, and therefore there are a different number of mediators for each interaction.

As the electromagnetic interaction's gauge group is U(1), it only contains one gauge boson, the photon. Its unique quality of having zero mass (which is a necessity as per special relativity) has a simple consequence for the electromagnetic gauge field and its interaction, which is *infinite* range. The weak interaction's gauge group is SU(2) and has three gauge bosons, of which two are charged and one is neutral: W^+, W^-, and Z^0. These are the only force carriers that possess mass, which justifies the weak interaction's short range. Furthermore, the existence of charged and neutral gauge bosons for the same force allows for two different types of weak interactions (called *current* interactions): the *charged*-current interaction that is mediated by one of the $W^\pm$ bosons (of which nuclear beta decay is a famous example) and the *neutral*-current interaction mediated by the Z^0.

[17] Right-handed neutrinos would *only* interact via the gravitational interaction, and are thus called *sterile* or *inert* neutrinos.

Having no charge or flavor quantum numbers of any kind, the Z^0-mediated interactions only alter the *momentum* of particles, and it does not participate in any decay interactions. This characteristic of the Z^0 is due to it being its own antiparticle, similarly to the photon and unlike the $W^{\pm}$, which are each other's antiparticle.

The strong force is the most complex of the three interactions, with the highest number of gauge bosons (eight) due to the SU(3) group. These bosons are called *gluons* and they carry color charge[18], which are always in pairs of color-anticolor of different flavors (i.e. blue-antired). Gluons, like quarks, are color-confined and cannot be seen in isolation. In fact, color confinement is a crucial aspect of the strong interaction, since the gluon is a massless particle. Intuitively this would indicate that the strong force, like electromagnetism, has infinite range, which is not the case. Color confinement does not allow gluons to be exchanged between quarks at a range larger than (roughly) the diameter of an atomic nucleus, above which the increasingly higher energy requirements make spontaneously generating a quark-antiquark pair from vacuum the more favourable option. Table 2.3 summarizes the Standard Model's five gauge bosons and their quantum numbers.

The concept of gauge invariance is paramount to the determination of the properties of gauge bosons. For example, local gauge invariance of the U(1) group is broken by the addition of a mass term to its Lagrangean, which implies in the massless property of its gauge boson, the photon. The electromagnetic and weak interactions are currently considered by the *electroweak* theory to be manifestations of the same force, unifying their unitary groups as SU(2) $\times$ U(1). However, according to Yang-Mills theory this unification leads to the existence of four *massless* gauge bosons for the electroweak interaction. This is clearly not the case, as the $W^{\pm}$ and Z^0 particles, gauge bosons for the weak interaction, have high masses. The weak interaction's non-infinite range also contradicts the existence

[18]Therefore they not only mediate the strong interaction, but also "feel" it themselves, leading to gluon-gluon interactions and even gluons interacting with themselves.

Gauge Boson	Mass (GeV/c^2)	Charge
γ (photon)	0	0
W^-	80.385	-1
W^+	80.385	$+1$
Z^0	91.188	0
g (gluon)	0	0

Table 2.3: Summary of the gauge bosons, separated by interaction. The gluons are further divided by the color charge they carry.

of massless force mediators, as there is no effect comparable with the strong interaction's color confinement to prevent it. A solution to this conundrum was formulated in the 1960's by numerous physicists [53, 54, 55, 56, 57, 58, 59], and is known as the Englert-Brout-Higgs-Guralnik-Hagen-Kibble mechanism, or, more simply, the *Higgs* mechanism.

The Higgs mechanism successfully explained how the gauge bosons of the weak interaction acquire their relatively large masses (while the photon does not) and was a key element in the unification of the weak and electromagnetic interactions. Later it was expanded to also describe the process through which fermions acquire mass, being today one of the most important concepts contained in the Standard Model. Simply stated, the Higgs mechanism is due to a new quantum field which permeates all of space, called the *Higgs field*. As it is a quantum gauge field, it possesses its own gauge boson called the *Higgs boson*. The gauge symmetry of the SU(2) $\times$ U(1) group that describes the electroweak interaction is unbroken at high energies (above 100 GeV), and at this energy scale, its four gauge bosons are massless, as described by Yang-Mills theory. Below this energy threshold, there is a *spontaneous* breaking of local gauge symmetry, which is the cornerstone of the Higgs mechanism. The Higgs field breaks down the SU(2) $\times$ U(1) symmetry and interacts with the three gauge bosons pertaining to the weak force's SU(2)

group, making them massive. The electromagnetic force's U(1) subgroup, on the other hand, remains unaffected and its gauge boson stays massless, as is required.

While a more detailed and mathematical description of the Higgs mechanism and spontaneous symmetry breaking are beyond the scope of this dissertation (for more in-depth information, see [60]), they are worth mentioning, as the mechanism has been expanded to also provide masses to other elementary particles. Fermions also interact with the Higgs field, gaining their own masses. In the Standard Model, the Higgs mechanism provides the masses of leptons and quarks, where the mass of "heavier" particles (such as the tau lepton and the bottom quark) is due to their relatively strong interaction with this field. The Higgs boson itself, whose existence helps confirm the veracity of the Higgs mechanism, is the final elementary particle encompassed by the Standard Model. Its tentative discovery at the Large Hadron Collider at CERN in 2012 was an enormous success for the theory. As it stands, the Model is composed of 61 particles: 36 quarks, 12 leptons, 12 gauge bosons, and finally the Higgs boson.

2.2 Motivation for the Analysis

The bottom quark is of particular interest for the analysis that is the subject of this dissertation. One of the Large Hadron Collider beauty (LHCb) experiment's main goals is the study of the existing asymmetry between matter and antimatter in the observable universe. This imbalance is one of the most widely investigated problems in modern physics, and at its heart is *Charge-Parity* (CP) violation, the simultaneous breaking of charge conjugation symmetry (exchanging a particle with its antiparticle) and parity symmetry (inversion of spatial coordinates). In a CP-symmetric universe equal amounts of matter and antimatter would have been generated by the Big Bang, which would then consequently *annihilate* each other, producing copious amounts of radiation. A scenario in which one type predominantly survives indicates that matter and antimatter obey different sets of laws, thus violating CP symmetry.

The Standard Model contains numerous sources of CP violation, and in this context the most important of these are quark flavor-changing weak decays. In the 1960s before the existence of the third quark generation, Nicola Cabibbo introduced the *Cabibbo matrix*[61], whose elements represented the probability of quarks changing from one flavor to another through weak decays. However, the experimental observation in 1964 of CP violation in weak decays illustrated a problem with the four-quark model, as the elements of the Cabibbo matrix did not allow this type of violation. However, *three* quark generations would allow a CP-violating complex phase, and in 1973 Makoto Kobayashi and Toshihide Maswaka proposed the existence of the third generation, and modified the Cabibbo matrix accordingly[62]. The new matrix, now with nine elements representing weak interaction quark couplings and allowing CP violation, is called the *Cabibbo-Kobayashi-Maskawa* (CKM) matrix.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.2}$$

On the left side of the equation are the weak interaction doublet partners of up-type quarks[19]. The first matrix term on the right is the CKM matrix itself, composed of V_{xy} elements that are related to the probability of a decay transition from a quark of flavor x to one of flavor y (or vice-versa), and the second term is simply a vector of mass eigenstates of the down-type quarks. The complex phase which is necessary for CP-violation only appears with a minimum of three quark generations, and it affects particularly strongly the matrix elements pertaining to the decays of bottom quarks (which also possess small values, in comparison with the other matrix elements), making them objects of interest for the study of CP-violation. Furthermore, the bottom quark is the *heaviest* quark that can form hadrons as its heavier partner, the top quark, has a lifetime that is too short to

[19]The up, charm and top are *up*-type quarks, while the down, strange and bottom are *down*-type.

permit hadronization. Due to the rich spectrum of *b-hadrons* that are available for study[20] and the relative ease in experimentally identifying bottom quarks because of their decays' unique characteristics, *B-physics* has become in recent decades a major field for research into CP violation and physics beyond the standard model.

Because of the large interest in studying B-mesons, two particle accelerators specifically designed to produce them were built in the 1990s. These colliders, referred to as *B-factories*, were built for the use of the *Belle* [63] and *BaBar* [64] experiments, respectively located at the High Energy Accelerator Research Organization (KEK) in Japan and the Stanford Linear Accelerator Center (SLAC) in the USA. Both experiments have since produced noteworthy results, and notable examples are: the first observation of CP-violation in a B-meson system [65] and the precise measurements of the CKM matrix's quark mixing elements $|V_{ub}|$ and $|V_{cb}|$ [66], both at the Belle experiment, and the first detection of the lowest-energy particle of the *bottomonium* family[21] [67] at the BaBar experiment.

The potential of B-physics was such that in 1998 the technical proposal of the LHCb experiment was submitted to (and approved by) CERN with the express goal of becoming the next generation heavy flavor experiment [68], due to higher energy thresholds and a more modern and specialized detector. While the existing B-factories made great progress in their first decade of operation, the energy constraints of their accelerators limit the statistics of B-mesons that they can produce for study. As it is located at the Large Hadron Collider, the LHCb experiment is able to take advantage of larger energies to produce much higher statistics of B-mesons, such as the B^+ and B^0, than can be obtained at the B-factories. Furthermore, rare particles such as the B_s^0 and the $B_c^\pm$ can *only* be produced in sufficient quantity for study at LHC energies, currently making the LHCb the only source for numerous decays of interest in B-physics. Allied with its detector's highly

[20] Λ_b^0, Ω_b^-, $B^\pm$, and $B_c^\pm$ are some examples.
[21] Particles containing a bottom-antibottom quark pair.

efficient event reconstruction and flexible trigger (see Section 3.2), the LHCb experiment is uniquely prepared to both perform high-precision measurements of Standard Model parameters and search for New Physics. One of the particles of interest for studies of this kind, and the most relevant for this dissertation, is the $B_c^{\pm}$ meson.

The $B_c^{\pm}$ meson possesses a number of properties that make it an attractive subject of study. It is currently the only known charmed bottom meson, being composed of a c-b quark pair. As it is a meson, one member of the pair must be an antiquark, and the charge of the B_c is determined by this: a charm-antibottom ($c\bar{b}$) quark pair represents the positively-charged meson, while a anticharm-bottom ($\bar{c}b$) pair represents its negatively-charged antiparticle. As it is composed by quarks of the second and third generations, the $B_c^{\pm}$ is the particle with the highest mass of the entire b-quark family, being more massive than even the b-baryons (see Table 2.4). This makes its decays attractive for study, as there are few other decay processes that can contribute significant components to its mass window (see section 4.2). Rare decays, such as those of the $B_c^{\pm}$, allow for *probing* of mechanisms beyond the Standard Model, as many of their decay parameters have yet to be precisely measured and compared to predictions.

The $B_c^{\pm}$ in particular can decay in a number of significant ways. It was initially discovered in 1998 by the CDF Collaboration at Fermilab [69], through studies of semi-leptonic decays of the form $B_c^{\pm} \to J/\psi l^{\pm} \nu$, where $l^{\pm}$ represents either electrons or muons, and ν represents a neutrino. Using data recorded by LHCb in Run 1 of the LHC, many $B_c^{\pm}$ decays were observed and published from 2012 to 2014. Two examples of decay observations made in this period are the $B_c \to J/\psi K^+ K^- \pi^+$ [70] and $B_c^+ \to B_s^0 \pi^+$ [71], both at LHCb. In particular, the weak decays of the B_c meson are of interest because both the c and b quark decay channels can be observed, where the c-quark decays are dominant. More information on the interest in studying the B_c meson can be found in [72].

For this dissertation, a *charmless* b-hadron decay is studied. These decays involve a

Particle	Mass (MeV/c^2)	Quark Composition
B^+	5279.29 ± 0.15	$u\bar{b}$
B^0	5290.61 ± 0.16	$d\bar{b}$
B_s^0	5366.79 ± 0.23	$s\bar{b}$
B_c^+	6275.1 ± 1.0	$c\bar{b}$
Λ_b^0	5619.51 ± 0.23	udb
Ξ_b^0	5791.8 ± 0.5	usb
Ξ_b^-	5794.4 ± 1.2	dsb
Ω_b^-	6048.0 ± 1.9	ssb

Table 2.4: Selected particles in the b-quark family. Top four are mesons, bottom four are baryons.

flavor-changing currents, which are mediated by the $W^{\pm}$ (for charged currents, like $b \to u$) and Z^0 (for neutral currents, like $b \to s$ or d) gauge bosons, and can be used to study potential deviations from Standard Model predictions. Charmless b-hadron decays can have different numbers of particles in their final states (body-multiplicity), and in this case a decay to a three-body hadronic final state is considered, of the kind $B_c^{\pm} \to h^{\pm}h^{\mp}h^{\pm}$. In decays of this kind without any charmed particles in the final state, the b-quark's decay occurs through a charged flavor-changing current, proceeding through $\bar{b}c \to W^+ \to u\bar{q}$ annihilation diagrams (see Figure 2.1), where q in this case is either a d or u quark. Direct annihilation contributions for the B_c also have the attractive quality of being more visible

in decay channels ($\approx 10\%$) than for other B-mesons, such as the B^+.

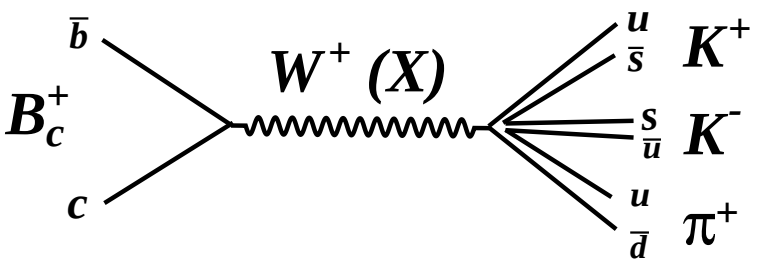

Figure 2.1: Tree Feynman diagram for the $B_c^+ \to K^+K^-\pi^+$ channel.

In order to take advantage of optimal sensitivity in LHCb, the final state for this dissertation's decay only contains charged particles with lifetimes large enough to be measured by the LHCb detector. Final states with three tracks are a simple way to reconstruct charged *mother* particles, and the mother particle being a meson, these final states can be either fully mesonic ($h'^+h'^-h^+$) or containing a proton-antiproton pair ($p\bar{p}h^+$), where the mesons are either kaons or pions. The phase space of these decays can contain other processes as well, such as b-quark charmless decays with intermediate resonances (such as $B_c^+ \to D^0(K\pi)h^+$), and $c \to q$ transitions (like $B_c^+ \to B_q^0(h_1^+h_2^-)h^+$). For this particular study, the chosen final state is $K^+K^-\pi^+$, occurring through annihilation processes or the aforementioned intermediate resonance decays, such as $B_c^+ \to D^0K^+$ (see Figure 2.2). At first, the main objective is to see whether isolating any significantly measurable B_c signal for this channel, using the statistics provided by Run 1 of the LHC, is possible. If so, the second goal is to measure the B_c's branching fraction (which is still unknown) through the use of control channels of well-established decays, and compare with the Standard Model predictions. Any deviations could indicate the presence of new physics, such as a BSM charged particle in place of the intermediate W^+ gauge boson. Even if a signal is

not measured with current data, the road will have been paved for future attempts using statistics from LHC's Run II or beyond.

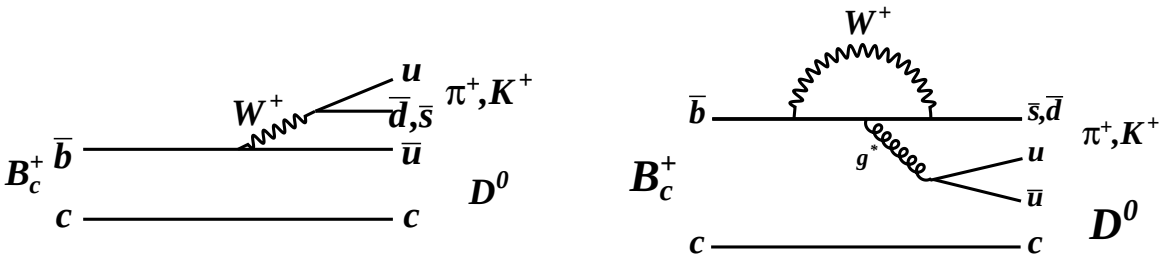

Figure 2.2: (left) Tree and (right) time-like penguin Feynman diagrams for $B_c^{\pm} \to D^0 h^{\pm}$ decays.

Chapter 3

The LHCb Experiment

The LHCb is one of the experiments currently taking place at the LHC (Large Hadron Collider). This chapter gives detailed information on the working conditions of the LHC accelerator and the LHCb detector (section 3.1). However, since the accelerators, detectors and software used in these experiments are constantly being worked on and upgraded, the details presented in this chapter will represent the state of the LHC and LHCb in the years from which the data used in this study was taken: 2011 and 2012. Sections 3.2 and 3.3 detail the LHCb experiment's detector and hardware systems, section 3.4 contains information regarding the *online* event selection system (*trigger*), and section 3.5 will present the LHCb experiment's *software* systems.

3.1 The Large Hadron Collider

CERN (Conseil Européen pour la Recherce Nucléaire) is an international research organization located near Geneva, Switzerland, which currently hosts the largest accelerator complex ever designed. This complex is composed of a series of accelerator machines, the most powerful of which is called the *Large Hadron Collider*, or LHC. The LHC is the largest and most powerful particle accelerator in history, and was designed to accelerate and collide beams of protons. While initially projected to accelerate these beams up to a center of mass energy of $\sqrt{s} = 14$ TeV, in 2011 and 2012 these energies were $\sqrt{s} = 7$ and

$\sqrt{s} = 8$ TeV, respectively.

The protons inside the beams are produced through hydrogen (H_2) gas ionization, and they undergo acceleration in a series of stages through the use of the aforementioned accelerators (see Figure 3.1), before being finally injected into the LHC. The first acceleration stage occurs in the LINAC (*LINear ACcelerator*) where the beams acquire a maximum energy of 50 MeV. Afterwards, the beams undergo three more acceleration stages in different machines, in the following sequence: PSB (*Proton Synchrotron Booster*), PS (*Proton Synchrotron*) and SPS (*Super Proton Synchrotron*), which accelerate the beams up to energies of 1.4 GeV, 26 GeV and 450 GeV, respectively. These are not the highest possible obtainable energies, rather they are the energies at which these accelerators operated when they were used for data production. Some of these machines were, at one time or another, the most powerful particle accelerators on the planet, and in this way they continue to be useful today as preliminary accelerators for the newer generation of colliders.

The final acceleration stage occurs in the LHC itself, which receives the proton beams in groups (called *bunches*) that contained, in 2011 and 2012, 1.45×10^{11} and 1.7×10^{11} protons each, respectively. The final proton beams involved in the collection of the data used in this study were comprised of 1380 bunches.[1] The beams are injected into the LHC's two tubes in opposite directions, forming two different beams which are then accelerated to their final energies of 3.5/4 TeV each. After this final acceleration stage, the beams circulate in the LHC, and may last up to 20 hours while data is measured and recorded.

In order to achieve these energies, the LHC requires considerable physical dimensions. The largest segment of the accelerator has a circumference of 27km, and is located simultaneously in France and Switzerland. The accelerator itself is underground, in a tunnel located at depths that vary from 50 to 150 meters. The underground placement

[1]However, the number of bunches per beam varies according to the LHC experiment, and these numbers will be presented along with the experiments themselves, later on in this section.

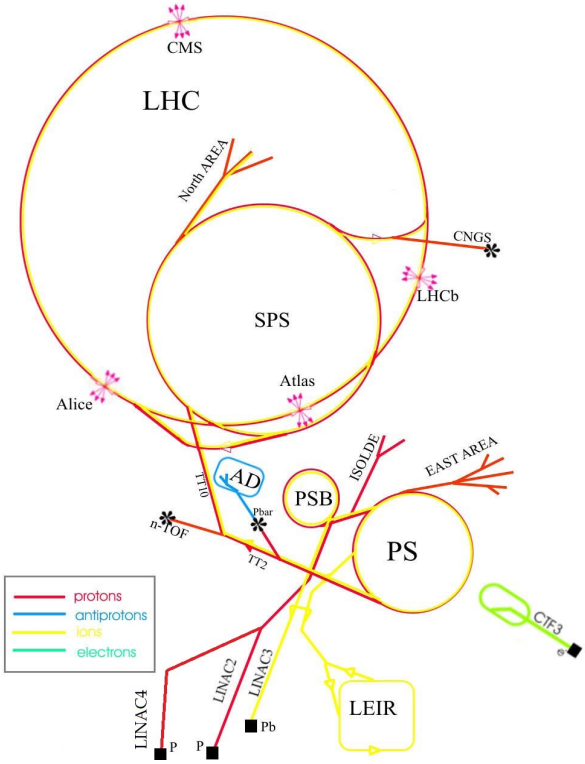

Figure 3.1: Current LHC accelerator complex. Shows the linear and circular accelerators, and positions of major experiments (magenta arrows).

.

of the accelerator is necessary because it shields the data from contamination (referred to as *"noise"*) coming from radiation and especially cosmic rays, and also protects the surrounding populace from any major accelerator malfunctions.

The machine itself has two internal tubes, containing the LHC's accelerated proton beams. These beams orbit in opposite directions inside the tubes, and they are confined inside the tubes and guided in their trajectory by magnets located throughout the full extension of the accelerator. The magnets used for this purpose require superconductors to function and are thermally regulated, being kept at a temperature of 1.9K.

The orbiting beams in the LHC tubes are kept separate, and only interact in four points along the circumference of the accelerator, where the four largest CERN detectors are located (Fig. 3.2). These detectors are used by the four main LHC experiments:

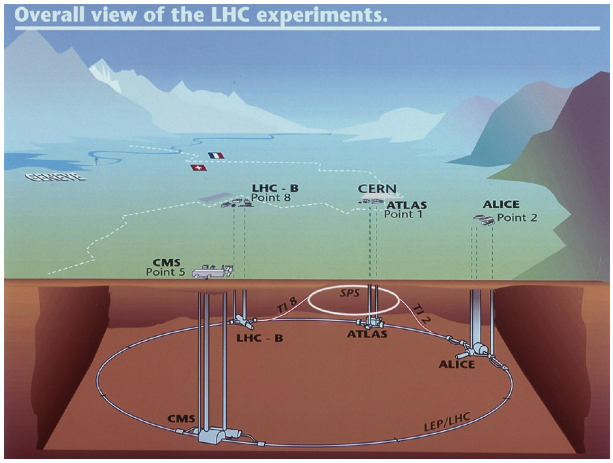

Figure 3.2: Schematic representation of the LHC complex, showing the relative locations of its four largest experiments.

.

ALICE, ATLAS, CMS and LHCb. ALICE (*A Large Ion Collider Experiment*) is, as the name suggests, dedicated to observing the collisions of heavy ions, and has the objective of collecting information about the behavior of QCD (*Quantum ChromoDynamics*) in extreme energy conditions. CMS (*Compact Muon Solenoid*) and ATLAS (*A Toroidal Lhc ApparatuS*) are the LHC's two general-purpose detectors and their associated experiments have multiple objectives. One of the most important of these objectives, and which was achieved in 2012 to much media attention, was finding the Higgs boson. Other goals include studying the properties of the Higgs and Higgs-like bosons, studying the Standard Model of Particle Physics, and searching for signs of supersymmetry and extra dimensions beyond the Standard Model.

The last of the four largest LHC experiments is the LHCb (*Large Hadron Collider beauty*) experiment, which uses a detector specially designed to study the interactions and rare decays of the b-quark, as well as the CP (*Charge-Parity*) violation that occurs in these decays. The LHCb experiment has a variety of objectives related to heavy-flavor quark (mainly beauty and charm) physics, such as the study of jets, charm, and electroweak/QCD physics. Only the LHCb detector's specifications and working condi-

tions are of interest in this study, and its performance and general characteristics will be discussed in the following paragraphs and in Section 3.2.

The final important detail regarding LHC data is the luminosity, which measures the number of collisions produced by the accelerator. This performance is characterized by the *instantaneous* and *integrated* luminosities. The instantaneous luminosity represents the rate of collisions, while the integrated luminosity represents the total number of collisions, thus being the most common way to present the quantity of recorded data in a certain time period. The luminosity itself is the physical quantity that relates the rate of a certain event R with its cross-section σ:

$$R = \mathcal{L}\sigma \tag{3.1}$$

The LHC was projected to function with an instantaneous luminosity of $\mathcal{L} \approx 10^{34} cm^{-2} s^{-1}$. The LHCb experiment, however, operated with instantaneous luminosities of $\mathcal{L} = 3 \times 10^{32} cm^{-2} s^{-1}$ in 2011 and $\mathcal{L} = 4 \times 10^{32} cm^{-2} s^{-1}$ in 2012. The total integrated luminosity recorded by the LHCb experiment during this period is expressed in equation 3.2, while Figure 3.3 shows the evolution of the LHCb experiment's integrated luminosity for the three years of Run 1.

$$\mathcal{L}_{int} = \int \mathcal{L} dt \approx 3.0 \text{ fb}^{-1} \tag{3.2}$$

The LHCb experiment requires a smaller instantaneous luminosity than most other LHC experiments (ALICE being an exception). This is due to the need, in flavor-precision physics, to properly resolve decay vertex structures, which is significantly complicated by large event pileup. In order to reduce event pileup, detector occupancy and radiation effects, the LHCb detector possesses real-time luminosity control, which is based on adjusting the beam transversal overlap at the LHCb interaction point. In this fashion, the instantaneous luminosity is reduced and the LHCb physics yield is maximized through

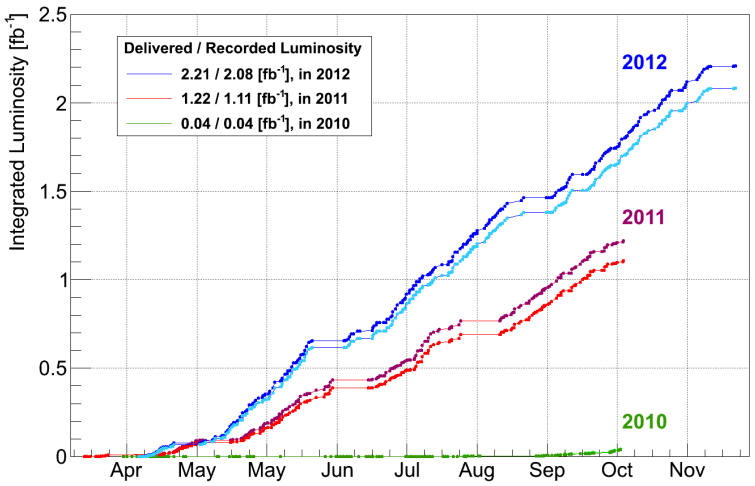

Figure 3.3: Integrated luminosity produced by the LHC (dark red, blue and green) and recorded by LHCb (light red, blue and green), for the entirety of Run 1.

.

optimal event pileup. More information on the LHCb detector's precision luminosity control can be seen in [73].

3.2 The LHCb Detector

The LHCb detector is a single-arm forward spectrometer, designed for the study of particles containing b or c quarks. The detector covers an angular region from 10 mrad to 300 and 250 mrad in the horizontal and vertical planes, respectively, and is shown in Figure 3.4.

The LHCb detector's geometrical design is due to the $b\bar{b}$ pair's characteristic of only being produced at small angles in respect to the proton beam during high-energy collisions, which is illustrated in Figure 3.5.

The detector includes a high-precision tracking system consisting of a silicon-strip vertex sub-detector surrounding the interaction region, a large-area silicon-strip sub-detector located upstream of a dipole magnet with a bending power of about 4 Tm, and three sta-

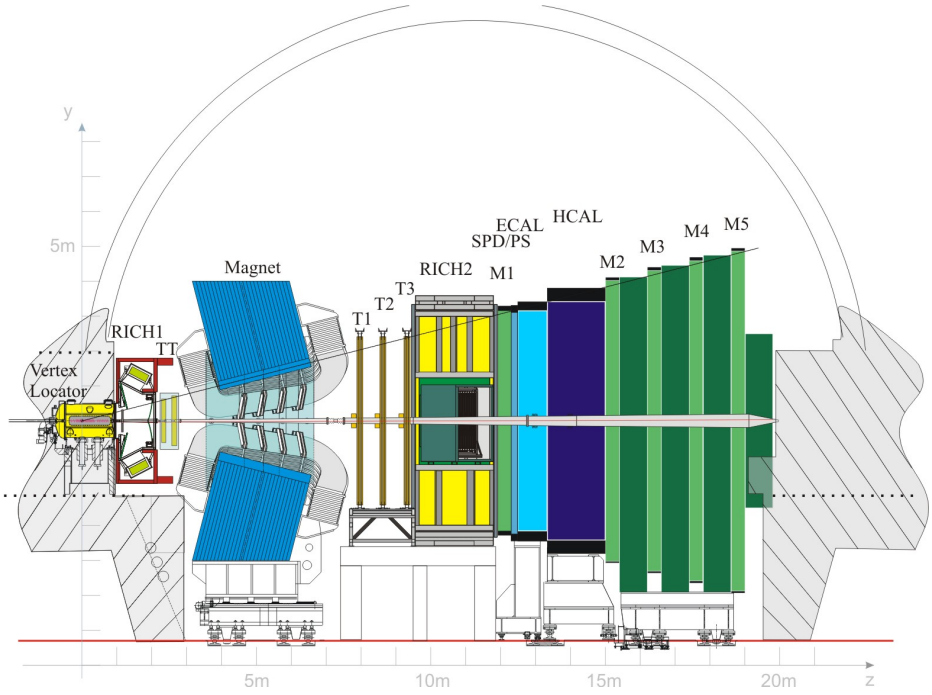

Figure 3.4: Side view of the LCHb detector and its sub-detectors, with vertical y and horizontal z axes.

.

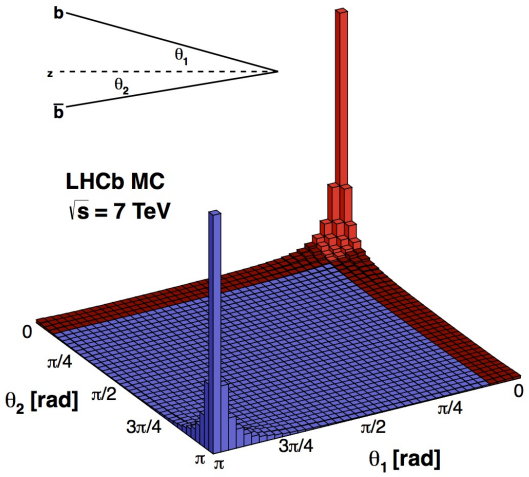

Figure 3.5: Monte Carlo simulation showing $b\bar{b}$ pair production at small angles during high-energy collisions. The x and y axes represent the angles between the $b(\bar{b})$-hadrons and the direction of the proton beams.

.

tions of silicon-strip sub-detectors and straw drift tubes placed downstream of the magnet. The tracking system provides a measurement of the momentum of charged particles with a relative uncertainty that varies from 0.5% at low momentum to 1.0% at 200 GeV/c. The minimum distance of a track to a primary interaction vertex (PV), known as the impact parameter, is measured with a resolution of $(15 + 29\ p_T^{-1})\ \mu m$, where p_T is the component of the momentum transverse to the beam, in GeV/c. Different types of charged hadrons are distinguished using information from two ring-imaging Cherenkov (RICH) detectors. Photons, electrons and other hadrons are identified by a calorimeter system consisting of scintillating-pad and preshower detectors, an electromagnetic calorimeter and a hadronic calorimeter. Muons, on the other hand, are identified by a system composed of alternating layers of iron and multiwire proportional chambers.

The following sections contain further information on the subsystems of the LHCb detector, and borrow mainly from [74] and [75], which contain fully detailed descriptions of the detector.

3.3 Detector Systems

3.3.1 Tracking System

The LHCb detector must be capable of not only measuring particle trajectories and momenta, but also of precisely separating primary and secondary vertices of decays. In order to achieve this, the detector utilizes a tracking system composed of a dedicated vertex detector known as the VELO (*VErtex LOcator*), discussed in the next subsection, and four tracking stations located in different areas of the detector. The first tracking station is called TT (*Tracker Turicensis*, see Section 3.3.1), and is located between RICH-1 (see Section 3.3.4) and the LHCb dipole magnet (see Section 3.3.2). The other three stations, T1-T3 (which contain both the *Inner* and *Outer* Trackers, see Section 3.3.1), are located over a 3 meter span between the magnet and RICH-2, as illustrated in Figure 3.6.

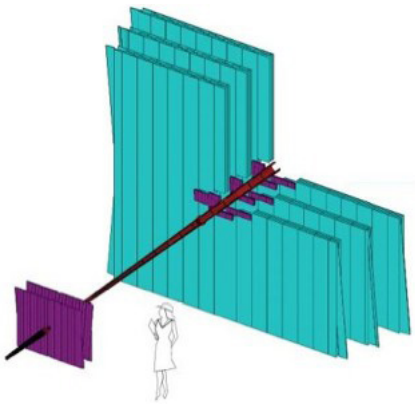

Figure 3.6: Representation of tracking stations of the LHCb detector. Present in the figure are tracking stations (purple) and the Outer Tracker (green).

Vertex Locator (VELO)

The function of the Vertex Locator sub-detector is to precisely measure track coordinates close to the LHC's beam interaction point and identify displaced secondary vertices which are a characteristic of B-hadron decays. B-mesons, however, have short mean lifetimes, on the order of one picosecond (ps)[2], which only allows them to travel distances on the order of milimeters before decaying into other particles. Therefore, the 42 silicon detector elements (separated in two halves) that compose the VELO are placed at a distance of only 5mm from the proton beam's interaction point, which puts the detectors in the path of the proton beams themselves.

In order to protect the detectors from any damaging effects of the high-energy beams and prevent any interference to the beams themselves, the VELO system has the capability to regularly change the locations of the detector elements between the measuring (physics) position of 5mm and the retracted position of 35 mm. Figure 3.7 shows the VELO's silicon module's arrangement along the beam direction and their fully open and closed positions.

The VELO system is positioned in order to cover the angular acceptance of the downstream detectors, being able to detect particles with a pseudorapidity[3] in the range

[2]The B_c meson in particular has a mean lifetime of approximately half a picosecond.

[3]Commonly used in experimental particle physics, pseudorapidity (represented by η) describes the

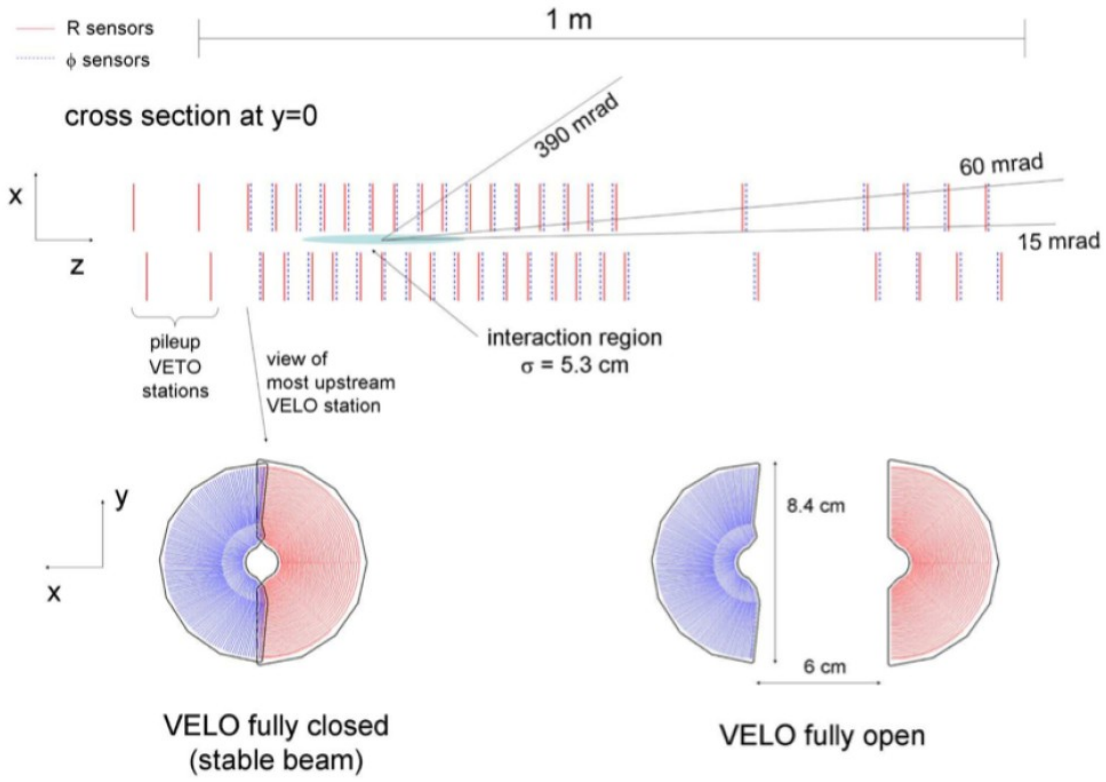

Figure 3.7: (top) Cross-section in the (x, z) plane of the VELO silicon sensors, in the fully closed physics position. (bottom) The front face of the VELO's first modules, in (left) closed and (right) open positions.

$1.6 < \eta < 4.9$ and emerging from primary vertices in the range $|z| < 10.6$ cm. Moreover, the VELO is designed to use cylindrical geometry ($r\phi$ coordinates), which enables faster track and vertex reconstruction than simple rectilinear schemes. Each of the VELO's $220\mu m$-wide modules (Figure 3.8) is designed to provide all 3D spatial information necessary for these reconstructions, and therefore possesses two distinct sensors.

The first of these is the ϕ-sensor (or ϕ-measuring sensor) which provides information on the beam's azimuthal coordinates. The sensor is composed of radial strips that are separated into inner ($r < 17.25$mm) and outer ($17.25 < r < 41.90$mm) regions, each region having a different number of strips. The inner region contains 683 5.9 mm-long strips, while in the outer region there are 1365 $24.9mm$-long strips. Taking both regions into account, each sensor is composed of 2048 strips, providing the system with approximately 180000 measurement channels.

The second sensor present in every module of the VELO system is the R-sensor (or r-measuring sensor), which provides information on the radial distance from the beam axis. The final required z-coordinate is provided by each sensor plane's position within the experiment. Each module in the system is capable of covering a azimuthal acceptance of $0 < \phi < 182°$. In the fully closed position, the modules are superimposed and cover the entire $0 < \phi < 2\pi$ acceptance range. More information on the VELO system and it's performance can be found in [76] and [77].

Tracker Turicensis

The Tracker Turicensis (formerly known as the Trigger Tracker), or TT, is a silicon microstrip sub-detector located between the RICH 1 and the Magnet in the main LHCb detector (see Figure 3.4). One of the two sub-detectors in the Silicon Tracker system, the TT is a 150cm wide and 130cm high tracking station, with an active area of 8.4 m^2 and 143360 readout strips, each of which can be up to 38 cm in length, allowing the TT

angle between a particle's momentum and the beam axis.

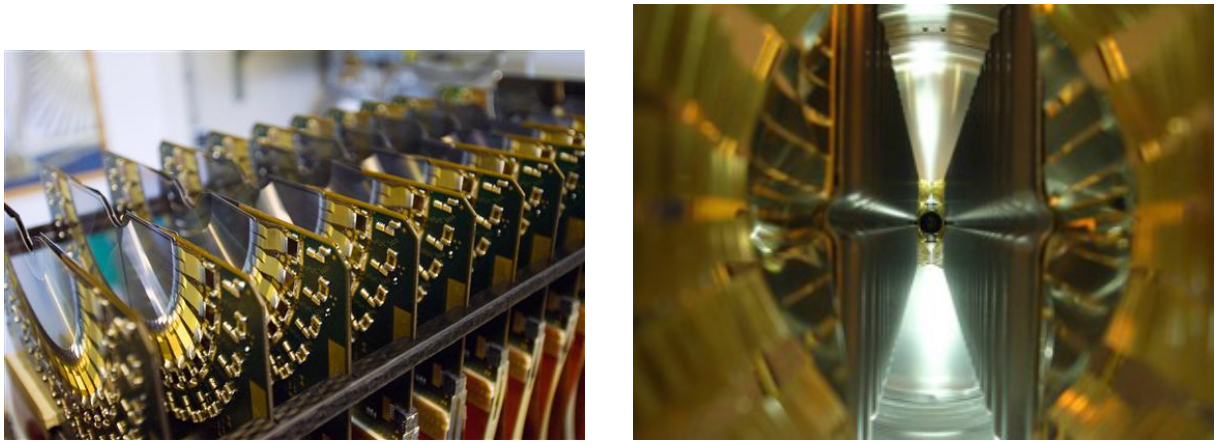

Figure 3.8: VELO system modules. The left image shows an outside view of the modules, while the right image looks into the system, in its closed position.

to cover the full acceptance of the experiment. The TT possesses four detection layers, which are comprised of a series of modules. Each layer has an $x-u-v-x$ arrangement (see Figure 3.9), in which the first and last layers have vertical strips, while the second and third layers have strips that are rotated by $-5°$ and $+5°$, respectively.

In order to aid in track reconstruction, the TT's four layers are arranged into pairs, $x-u$ and $v-x$, which are separated by approximately 27cm along the LHC beam axis and housed in a thermally and electrically insulated detector volume. This volume maintains an internal temperature of under 5°C, and is constantly flushed with nitrogen, to avoid condensation. The basic component of the TT's detection layers (see Figure 3.10) is a half-module that covers half the height of the LHCb acceptance, consisting of a row of seven silicon sensors. The regions above and below the LHC beam pipe are each covered by one of these half-modules, while the regions to the sides of the beam pipe are covered by rows of seven (first two layers) or eight (last two layers) 14-sensor long full modules. These full modules are assembled from two half-modules, and cover the full height of the LHCb acceptance. The TT is capable of providing a spatial resolution of 50μm on the position of every single *hit*. More information on the Tracker Turicensis can be found in Section 5.2 of [74].

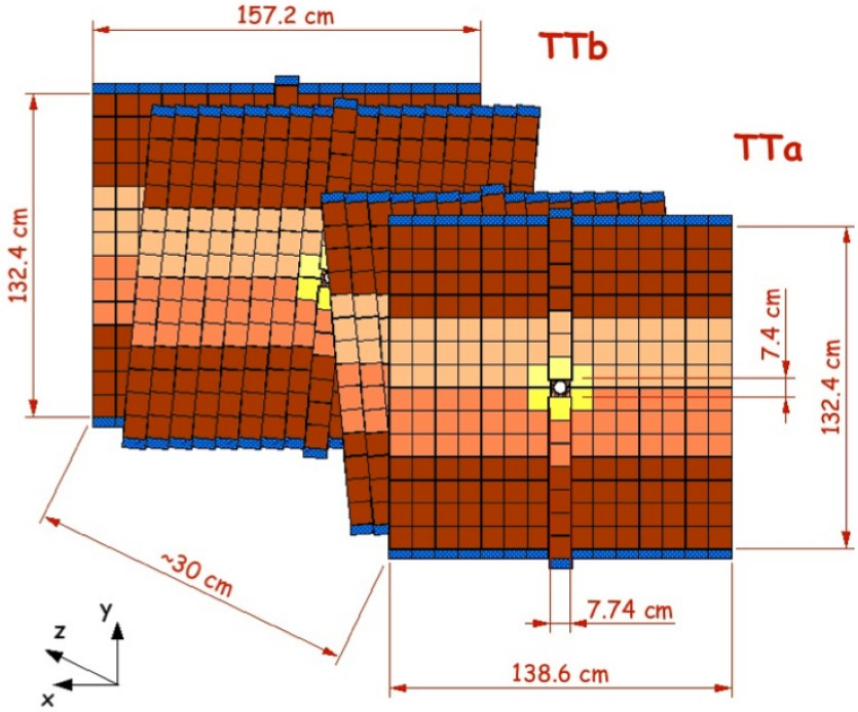

Figure 3.9: The TT's $x - u - v - x$ geometry, and the dimensions of the TTa and TTb, which contain the first two and last two layers, respectively.

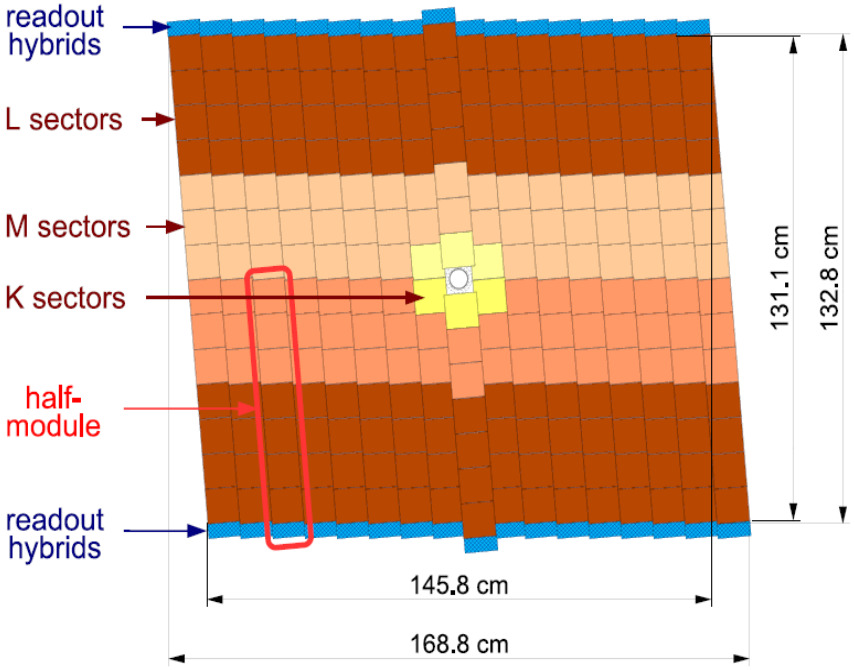

Figure 3.10: Layout of the third (v) TT detection layer. The half-module is marked in red, while different readout sectors have different shadings.

Inner Tracker

The Inner Tracker (or IT) system is similar to the TT in many respects. It uses silicon microstrips, the same four-layer $x-u-v-x$ configuration, and each of its stations consists of detector volumes that are insulated and thermally regulated in the same way. However, the IT system is comprised of three stations (T1, T2 and T3) instead of one, which are arranged around the LHC beam pipe, and they do not cover the full LHCb acceptance. The IT covers a 120cm wide and 40cm high cross-shaped region in the center of the three stations, which are located downstream of the detector's magnet (see Figure 3.4). Even though this represents only 2% of the full area covered by all the tracking stations, approximately 20% of the particle trajectories that pass through the stations are detected by the IT system.

Each IT station is made up of four detector volumes containing four layers of silicon sensors, organized in groups of seven modules. The modules have slightly different sensor configurations, depending on their location. Modules above or below the beam pipe contain only one 320 μm-wide sensor, while modules to either side of the beam pipe contain two 420 μm-wide sensors. In all other respects the sensors are identical, and IT's spatial resolution is approximately 50 μm per hit, the same as the TT's. Figure 3.11 shows the configuration of the IT detector volumes surrounding the LHC beam pipe, and more information regarding the Inner Tracker system can be found in [78].

Outer Tracker

The silicon sensors that are used in the entirety of the TT system and in the IT area of the T1-T3 tracking stations, while costly, offer precise spatial resolution, which is why they are positioned close to the beam pipe where the density of particles is highest. In the areas further from the beam pipe, where particle density is smaller and precise spatial resolution is therefore easier (and less expensive) to achieve, trajectories are detected by

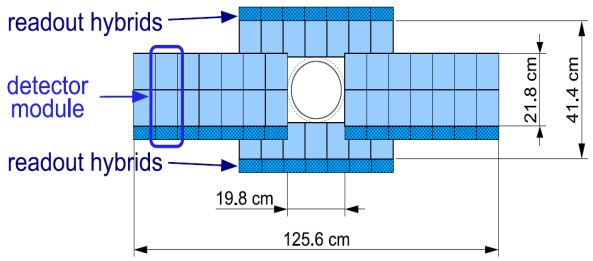

Figure 3.11: Layout of an IT station's x detection layer, surrounding the central LHC beampipe.

the Outer Tracker (OT), a *drift-time*[4] detector that uses thousands of gas-filled straw tubes (each with a diameter of 4.9mm) to track charged particle trajectories and measure their momentum over a large acceptance area. Each of the T1-T3 stations has its own OT straw tube component (see Figures 3.6 and 3.12).

The OT system is an array of straw tube modules, each of which contains two staggered layers of drift-tubes, using the same $x-u-v-x$ configuration as the IT and TT. Each double layer is composed of 4608 tubes, and all three tracking stations combined possess approximately 55000 tubes. The tubes each have a wire in the center and are filled with a mixture of Argon (70%) and CO_2 (30%) gases. A potential difference between the central wire and the inner walls is maintained in each tube, so that when a charged particle passes through the tube, the gas is ionized and produces electrons which are accelerated towards the wire, and the particle's trajectory is determined by the time it takes for the electrons produced to reach the wire (see Figure 3.13).

Straw tube technology is less expensive than silicon detectors, making it possible for the OT system to cover a much larger area surrounding the beam pipe: each T1-3

[4]Drift-time refers to the time interval between the passage of a particle through the detector and the detector's measurement of said particle.

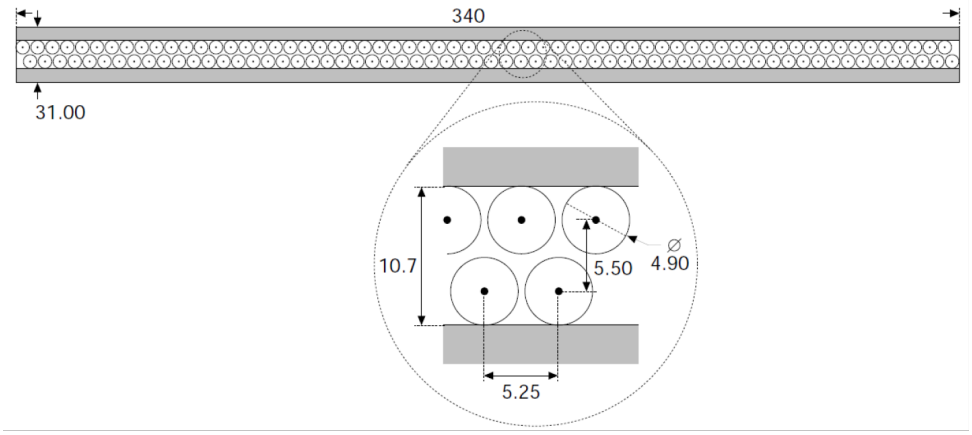

Figure 3.12: Cross-section of a straw-tube module (dimensions in mm).

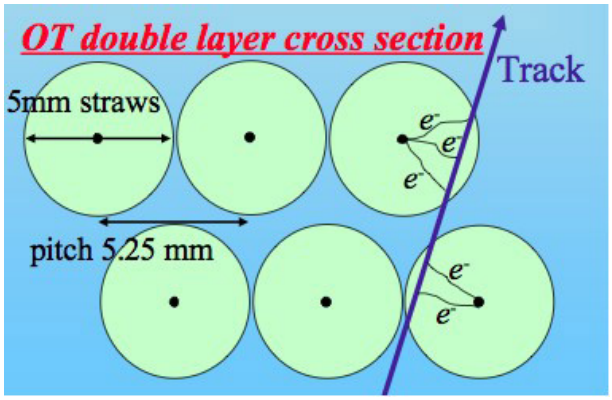

Figure 3.13: Schematic of a charged particle's passage through an OT module, with ionization-produced electrons inside the affected straw tubes.

station has a total active area of 5971×4850 mm^2. However, since drift-time detectors by definition have a larger delay between the passage of a particle and a measurement when compared to silicon detectors, the spatial resolution of the OT system is not as precise as the resolution offered by the TT and IT systems. The drift time of the OT system is under 50ns, offering a drift-coordinate resolution of 200 μm. On the other hand, the mass resolution of the system is very precise, which is necessary in order to accurately determine the invariant mass of reconstructed b-hadrons. More information on the Outer Tracker system can be seen in [79].

3.3.2 Magnet

Located between the TT and T1-3 tracking stations (see Figure 3.4), the LHCb detector's dipole magnet is used to measure the momentum and electric charge of particles (see Figure 3.14). It is a "warm" magnet [5] capable of fulfilling the LHCb experiment's contrasting requirements of simultaneously maintaining a magnetic field level inside the RICH detector's envelope (see Section 3.3.4) of less than 2 mT while also producing the highest possible field in the region between the VELO and the TT system. Furthermore, the magnet is also able to invert its polarity, which is fundamentally important in studying systematic effects in asymmetry measurements.

The main components of the magnet are two identical, conical saddle-shaped coils, which are placed mirror-symmetrically to each other in the magnet yoke. These 27-ton coils deliver a maximum magnetic field of 1.1T, and an integrated magnetic field of 4Tm for tracks of 10m lengths (see Figure 3.15). Further information on the LHCb detector's dipole magnet can be found in [80] and in Chapter 4 of [74].

[5] As opposed to a super-conducting magnet, which was originally proposed but not built due to prohibitively high cost.

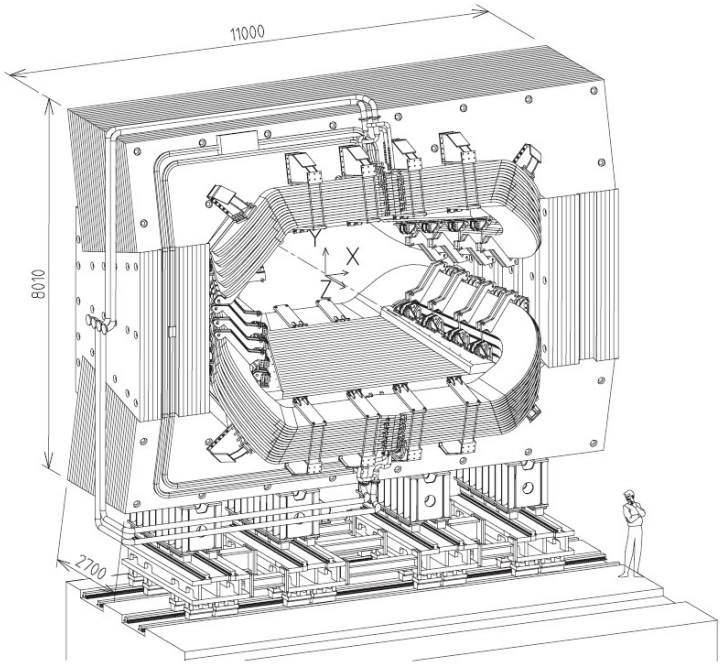

Figure 3.14: Scale view of the LHCb magnet (units in mm). In this representation, the interaction point is behind the magnet.

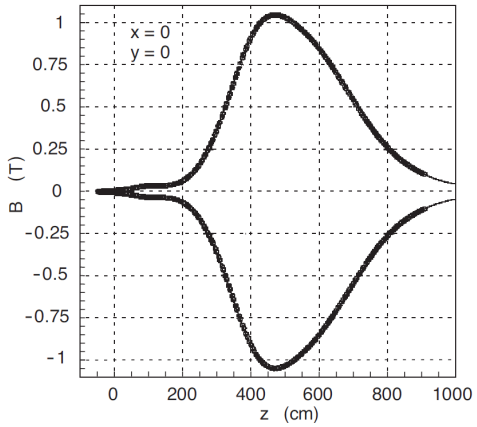

Figure 3.15: Generated magnetic field along the z axis.

3.3.3 Calorimeters

The LHCb's calorimeter system is composed of four components: SPD (*Scintillator Pad Detector*), PS (*PreShower*), ECAL (*Electromagnet CALorimeter*) and HCAL (*Hadronic CALorimeter*), and is used both to measure the transverse energy and momentum of particles and to identify photons, electrons and hadrons for the first *trigger* level (L0, see Section 3.4.1). The system makes measurements 4 μs after interacting with incident particles and is capable of precisely determining their positions, which are fundamental for reconstructing π^0 and prompt photons with good accuracy. These particles (and efficient measurement of high transverse momentum in general) are important for the physics program of the LHCb, since they contribute to the flavor tagging of B-meson decays. Calorimeters provide the main way of identifying and measuring particles that have no electrical charge, and they accomplish this by blocking the passage of the particles and measuring the energy they lose. Electrons interact at high energies through *bremsstrahlung*[6], i.e. through the emission of photons. The emitted photons subsequently create electron-positron pairs that emit more photons, creating an *electromagnetic shower*, and Hadron interactions at high energies likewise produce *hadronic showers*. The ECAL and HCAL sub-detectors are designed to be capable of forcing their respective particles to deposit all their energy, which is why their lengths are expressed as functions of the radiation length X_0. Both calorimeters function in similar ways, using scintillating material that emits photons upon particle interaction, which are then collected and have their wavelength modified by *Wavelength-Shifter* (WLS) fibers for easier detection and measurement through photomultipliers (see Figure 3.16).

SPD and PS

The SPD and PS, placed next to each other and before the main section of the ECAL, are responsible for identifying neutral and charged particles that cause showers and initiating

[6]German for "braking radiation" or "deceleration radiation".

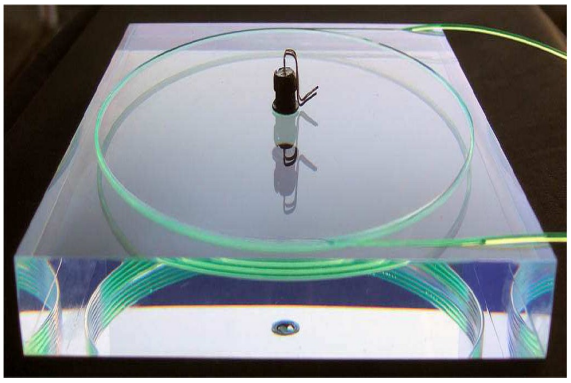

Figure 3.16: Individual scintillator pad showing the WLS fiber layout.

the showers themselves, respectively. The SPD/PS detector consists of two almost identical planes of rectangular scintillator pads of high granularity, offering a total of 12032 detection channels. There is a 15mm lead converter with a thickness of $2.5X_0$ between the planes. The SPD's sensitive area is 7.6m wide and 6.2m high, and due to the detector's angular acceptance, the PS's dimensions are approximately 0.45% larger.

ECAL

The LHCb detector's electromagnetic calorimeter is located after the PS and 12.5m from the interaction point, and is composed of alternating layers of passive and active materials. The passive materials are responsible for initiating electromagnetic showers and reducing particle energy, while the active material is used to measure how many layers the incident particles are able to penetrate. This alternating configuration uses *shashlik* technology, with layers of 4mm-wide scintillating and 2mm-wide lead tiles. When all 66 layers are taken into account, this represents a 42cm-thick (or $25X_0$) barrier, capable of completely stopping photons and electrons and measuring their energy. The energy resolution has a dependency on the energy of the incident particle:

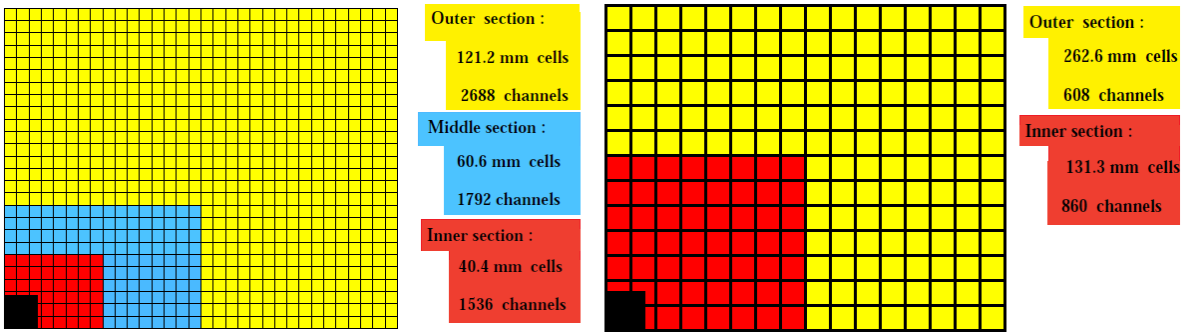

Figure 3.17: Lateral segmentation of the SPD/PS and ECAL (left) and the HCAL (right), showing one quarter of the detector front face. Left figure dimensions are for the ECAL.

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \bigoplus b \tag{3.3}$$

Where $a = 0.09\sqrt{GeV}$ (shower statistical uncertainty), $b = 0.008$ (systematic uncertainty), and $\bigoplus$ represents a quadrature sum.

HCAL

The LHCb detector's hadronic calorimeter is placed right after the ECAL, 13.3m beyond the interaction point. It has considerable dimensions, being 8.4m tall, 6.8m wide, and 1.65m thick, and functions in a similar fashion to the ECAL, i.e. using alternating layers of passive and active materials. In the HCAL, 16mm-thick iron tiles are used instead of lead for the passive material, in order to be capable of stopping hadrons, which are heavier than photons and electrons. There are three other main differences between the two calorimeters. The first is the scintillating tile orientation of the HCAL, which runs parallel to the beam axis instead of perpendicularly, which is the case in the ECAL (see Figure 3.18). The second is the granularity, which is smaller for the HCAL due to the larger width of hadronic showers, in comparison to electromagnetic showers; and the third is the different segmentation of the HCAL, which only has an inner and outer region (see

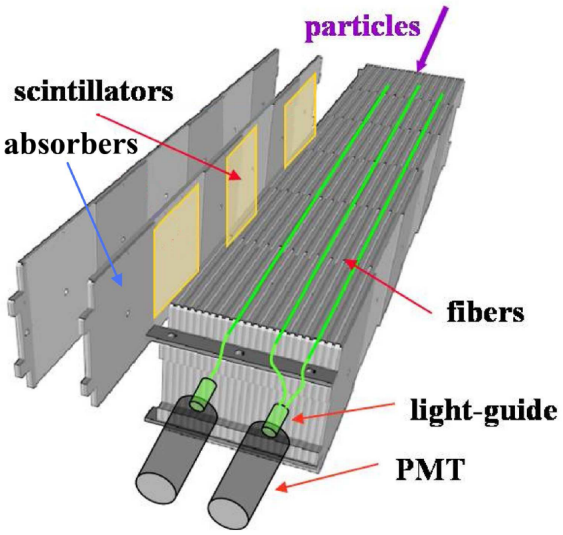

Figure 3.18: Schematic of HCAL internal cell structure.

Figure 3.17). The HCAL's energy resolution is also similarly calculated:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \bigoplus b \tag{3.4}$$

Where $a = 0.69\sqrt{\text{GeV}}$ (shower statistical uncertainty) and $b = 0.09$ (systematic uncertainty). More information on the LHCb detector's calorimeters can be found in [81] and in Section 6.2 of [74].

3.3.4 Ring Imaging Cherenkov (RICH) Detectors

Charged particles, when travelling through a dielectric medium at velocities greater than the speed of light in that medium ($v > c/n$), polarize and excite the molecules in the medium, which subsequently emit photons. The result is that the charged particle emits a cone of light in a phenomenon known as *Cherenkov* radiation. Since the radiation cone's emission angle θ is explicitly related to the particle's velocity:

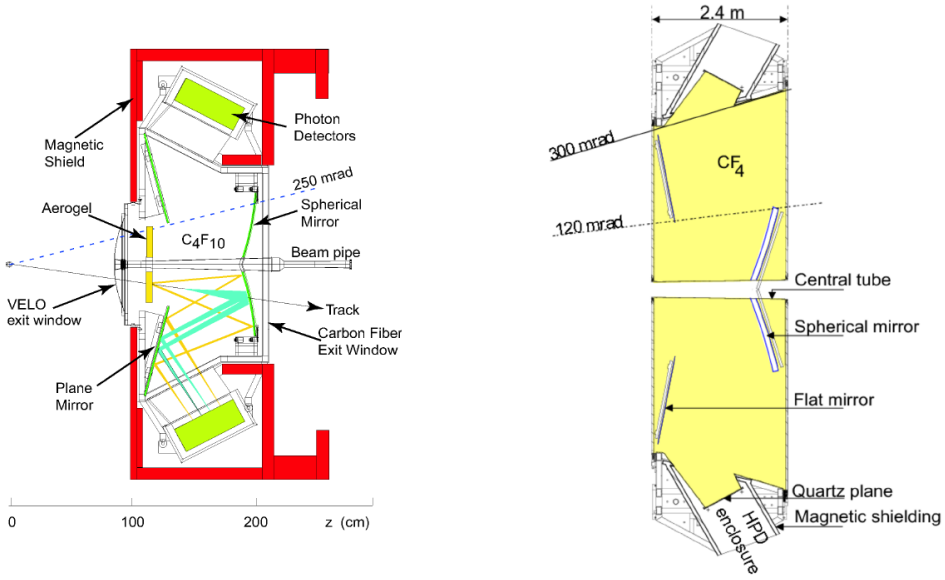

Figure 3.19: Schematic representation of the (left) RICH-1 and (right) RICH-2 sub-detectors.

$$cos\theta = 1/\beta n \tag{3.5}$$

where β is the ratio of v to the speed of light c, knowing the angle of Cherenkov radiation emission equates to knowing the speed of the emitting particle. The LHCb detector contains two Cherenkov radiation-measuring sub-detectors, RICH-1 and RICH-2 (see Figure 3.19).

Two RICH detectors are necessary to cover the full momentum range, and these are placed in different places after the interaction point (see Figure 3.4) using different radiative mediums (see Figure 3.20). RICH-1 is located upstream of the magnet between the VELO and TT, approximately 1m from the interaction point. It covers the full LHCb acceptance (see Section3.2) and low momentum charged particle range of $\approx$ 1-60GeV/c, and uses aerogel and C_4F_{10} radiators. RICH-2 is located downstream of the magnet, between the T1-3 stations and the first muon chamber (see next section), approximately 10m from the interaction point. It utilizes a CF_4 radiator and covers the momentum range

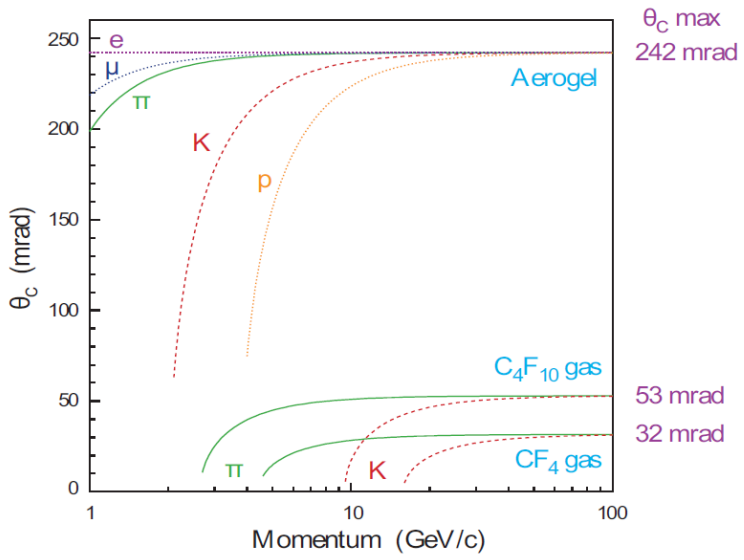

Figure 3.20: Cherenkov angle versus particle momentum for the three different RICh radiators.

from $\approx$ 15GeV/c to more than 100GeV/c, through a more limited angular acceptance of $\pm$15mrad to $\pm$120mrad (horizontal) and $\pm$100mrad (vertical), which is the range where higher-momentum particles are produced.

Measurements of Cherenkov radiation from the RICH detectors are used for particle identification. The information they provide on charged particle speeds are combined with momentum and trajectory measurements coming from the tracking system in order to calculate the particle's mass and charge, which then determines its identity. For the LHCb experiment, the RICH detectors help identify protons, kaons and pions, which are commonly found in the final states of b-hadron decays, such as the decay that is the subject of this study. More information on the LHCb's RICH detectors can be seen in [82] and in Section 6.1 of [74].

3.3.5 Muon System

The final LHCb sub-detector is called the muon system, and its main purpose is to identify muons (μ), an absolutely crucial function due to the presence of these particles in the final

states of many important CP-sensitive b-hadron decays[7]. It provides information for the high-p_T muon trigger (see Section 3.4), and muon identification for the high-level trigger (HLT) and offline analysis. Furthermore, muons are important for quark flavor-tagging, being used in both CP assymetry and oscillation measurements.

The LHCb's muon system is composed of five rectangular stations (see Figure 3.21), M1-5, which are placed along the beam axis (see Figure 3.4). M1 is placed 12.1m after the interaction point and is separated from the other four stations by the electromagnetic and hadronic calorimeters, being used for the trigger system's p_T measurements. The other four stations M2-5 are placed at the end of the detector, respectively at 15.2m, 16.4m, 17.6m and 18.8m after the interaction point. Interwoven between the final four stations are 80cm thick iron absorbers, used to select muons. The geometry of the five stations is scaled according to the distance from the interaction point, meaning each station has slightly larger transverse dimensions than the one before it. In total, the system is composed of 1380 chambers (276 in each station) and covers a total area of 435 m^2.

Each station is organized into four regions R1-4. The linear dimensions of the regions and their segmentations scale in the ratio 1:2:4:8, in order to maintain roughly the same detector occupancy in all regions. The detectors in the stations provide space point measurements of muon tracks, delivering binary (yes/no) information, which is measured by dividing the chambers into *logical pads*. The dimensions of the pads define the *x,y* resolution (see Figure 3.22), and each station's four regions have different resolutions that are larger (i.e. less precise) the farther away they are from the beam pipe.

Almost all of the muon systems' stations use Multi-Wired Proportional Chambers (MWPC), which are filled with a gaseous mixture ($Ar(40\%), CO_2(50\%)$ and $CF_4(10\%)$) and have a system of anode wires between two metal cathode plates. Their function is similar to the straw tubes used in the tracking system: a potential difference is applied

[7] $B_s^0 \to J/\psi(\mu^+\mu^-)\phi$ and $B^0 \to J/\psi(\mu^+\mu^-)K_s^0$ being two well-known examples.

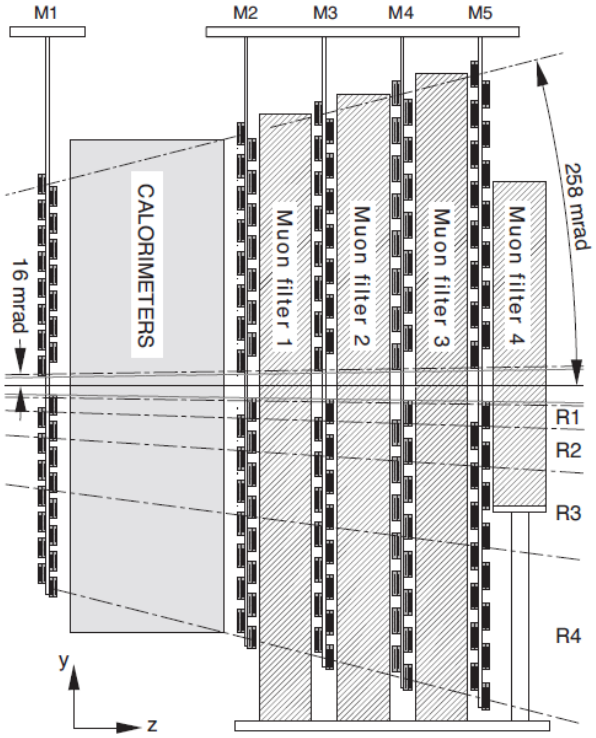

Figure 3.21: Side view of the muon system, with the calorimeters separating M1 from M2-5. The four iron filters are also shown.

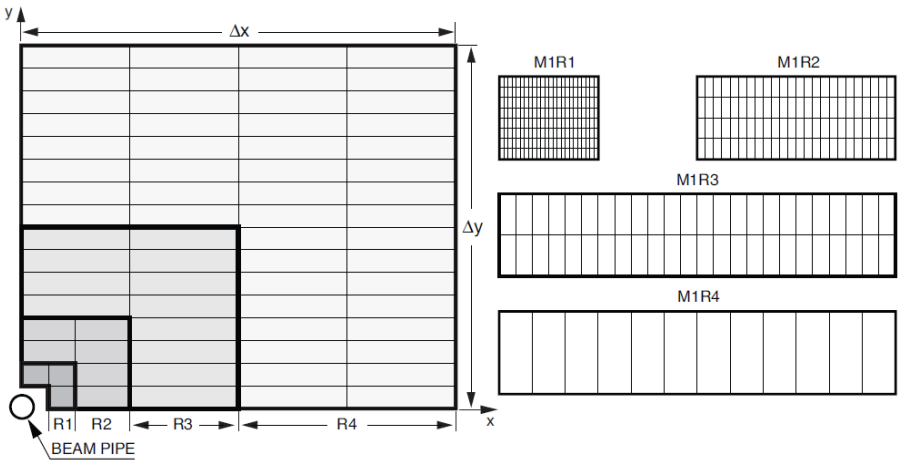

Figure 3.22: Left: front view of a muon station quadrant in which each rectangle represents a chamber. Right: division into logical pads of R1-4 in M1.

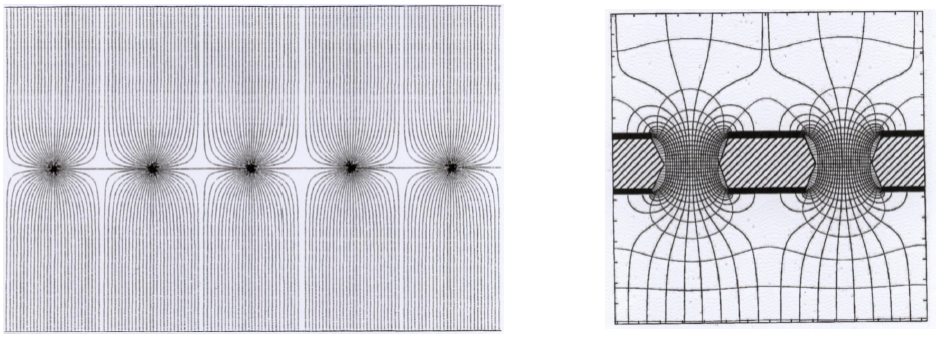

Figure 3.23: Field lines for (left) MWPC central wires and (right) GEM plate holes.

between the wires and metal plates, and upon interaction with a charged muon the gas mixture is ionized. The subsequent emitted ionization electrons are collected by the wires and are used to measure the *hit* position.

MWPC's have a limit of data acquisition, which is why they aren't used in the first region of the first muon station. R1 is closest to the beam pipe while M1 is located before the calorimeter system, meaning that this specific region expects a much higher particle rate. Therefore, this region is the only one to use a unique detector chamber system, called the triple Gas Electron Multiplier (triple-*GEM*). Each chamber is filled with the same gaseous mixture as the MWPC's, and is composed of two conducting metal walls (functioning as cathode and anode), with three thin GEM plates in the middle. Each GEM plate is filled with holes, in which strong electric fields are applied. Incident charged particles ionize the gas, and the emitted electrons pass through the holes in the GEM plates, creating a discharge which is then collected at the anode wall. The electric field lines for the MWP and triple-GEM chambers can be seen in Figure 3.23, and more information on the LHCb detector's muon system can be found in [83] and in Section 6.3 of [74].

3.4 Trigger Systems

The LHCb experiment operates using a collision rate of approximately 16 MHz[8], which is less than half the LHC's proton beam crossing rate of 40 MHz. This is due to the LHCb's requirement of smaller event pileup, which was mentioned in the beginning of this chapter. However, even this smaller collison rate is more than the detector's hardware is capable of processing and storing for later analysis, meaning that possibly meaningful events may be lost. For this reason, the LHCb detector possesses a dedicated *trigger* system whose function is to select which produced events are recorded during data-taking. This selection is optimized using both the topology of B-meson decays, whose lifetimes are long enough to form decay vertices out of the interaction point, and particles with high p_T, due to the B-meson's large invariant mass.

The *trigger* system is divided into two levels that have different functions: the Level-0 trigger (L0) and the High Level Trigger (HLT), which is further divided into two sub-levels. The former is activated at the hardware level and uses dedicated electronic systems, which take into account information from the VELO, muon stations and calorimeters to make individual event selections/rejections in $\approx 4\mu$s. The latter operates at the software level using C++ algorithms running simultaneously on 1000 computers[9]. Being applied at the software level, the HLT is much more flexible than the L0, having more time to make decisions and being much easier to modify.

The following subsections are based mainly on [84] and Chapter 7 of [74], where more information on the LHCb detector's trigger systems can be found.

3.4.1 Level-0 (L0)

The L0 trigger takes its information from three sources: the calorimeter trigger, the muon trigger, and the pile-up system. The information they provide is combined in the Level-0

[8]This represents only an *average* collision rate, since the time between subsequent collisions varies.
[9]Referred to as a *processor farm*.

Decision Unit (L0DU), which is then responsible for the final decision on whether or not to accept an event. The approximately 16 MHz crossing rate used by the LHCb experiment is reduced to approximately 1 MHz by this first trigger level.

L0 Calorimeter Trigger

The Level-0 calorimeter trigger system looks for high E_T particles, which may be electrons, photons (γ), neutral pions (π^0), or hadrons. This search is accomplished by adding the E_T of $2x2$ cells and selecting those *clusters*[10] with the largest values of E_T. Using information from the SPD, PS, ECAL and HCAL calorimeters (see Figure 3.24), each cluster selected is then identified as an electron (L0Electron), photon (L0Photon), or hadron (L0Hadron), as per the following conditions: if a cluster is measured in the HCAL, it is classified as a hadron; if it is measured in the ECAL and the corresponding HCAL cluster isn't activated, the cluster can be a photon or an electron. In this case, the final classification is made according to the amount of *hits* in the SPD.

These SPD *hits* are also used to veto certain events. A large number of SPD cells with a *hit* provide a measure of charged track multiplicity, and the hit is rejected if the number is too large. This rejection is done to avoid accepting events that are too complex and therefore consume over-large processing times.

L0 Muon Trigger

The L0 Muon trigger is composed of a system of processors installed in each quadrant of the LHCb detector's five Muon Stations (see Figure 3.25). The objective of this trigger level is to select the two muons with the largest transverse momentum p_T in each quadrant, allowing each event to have up to eight muon candidates. The system searches for *hits* defining a straight line that originates at the interaction point and passes through the five Muon Stations. An event is accepted for any combination of two conditions, based on a p_T

[10]Calorimeter clusters are groups of cells which are activated through interaction with an incident particle.

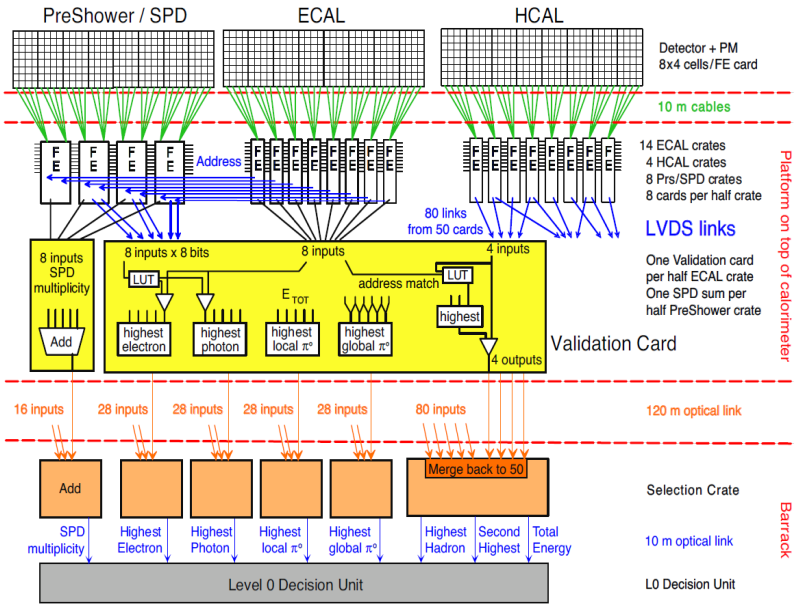

Figure 3.24: Architecture of the Level-0 calorimeter trigger.

threshold: if any single muon candidate possesses p_T above the threshold (L0Muon); or if the combination of the largest and second-largest p_T candidates are above the threshold.

L0 Pile-Up System

The L0 Trigger level's Pile-Up system is composed of two 300μm-thick silicon planes, which are perpendicular to the beam-line and located upstream of the VELO. Each plane consists of overlapping radial-sensors, similar to those used in the VELO modules. The planes measure the quantity of primary vertices in an event and reject it if this quantity is higher than a predefined limit.

3.4.2 High-Level Trigger (HLT)

The second trigger level is the High-Level Trigger, or HLT, and its main purpose is to further reduce the rate of data transfer that passes the L0 trigger, in order to reach a transfer rate small enough to record. The L0 level output of 1 MHz is reduced to approximately 3 KHz by the HLT (see Figure 3.26), which uses a C++ application

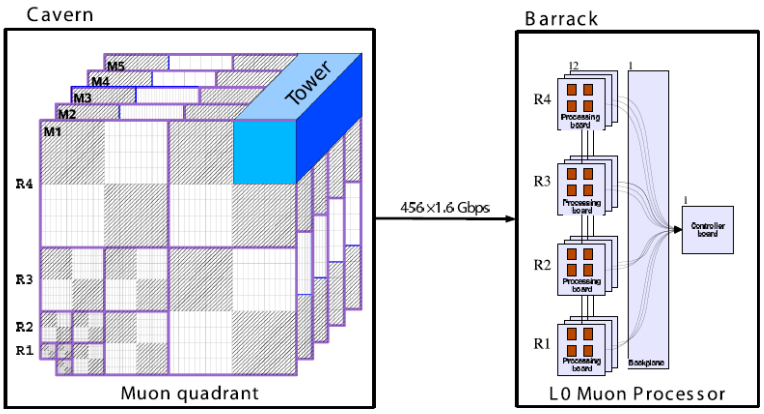

Figure 3.25: Achitecture of the L0 Muon system, showing the processing boards of each Muon Station quadrant.

executed in every CPU of a specialized computing farm known as the *Event Filter Farm*. Each application of the HLT has access to all the information on a given event that passes the L0 level, and being software-based, the HLT is highly flexible and can be upgraded over time.

The HLT is divided into two stages, called HLT1 and HLT2, which have different purposes. HLT1 searches for high p_T particles that *don't* originate from the primary vertex. HLT2 completes the (complete or partial) reconstruction of the event through two types of selections, *inclusive* or *exclusive*. The exclusive selections use tracking performance data, while the inclusive selections select partial b-hadron decays to known final states, such as ϕX and $J/\psi X$, among others.

3.5 Software Systems

The trigger system delivers raw subdetector data, which can not be immediately analyzed without first being interpreted and reconstructed by LHCb's dedicated software systems. These software systems (or *packages*) are based on the *Gaudi*(see [85]) framework, and are capable of data analysis (*DaVinci*), event reconstruction (*Brunel*) and simulations

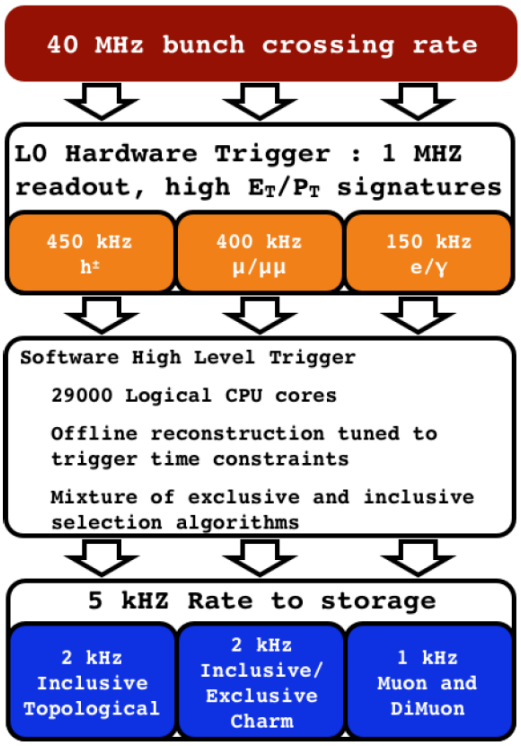

Figure 3.26: Map of data flow through the trigger steps.

(*Gauss*), among others. These packages are further detailed below, while the following subsections will focus on the LHCb's track reconstruction and stripping software.

- *Gauss* (see [86]) is the LHCb program that simulates what happens in the LHCb detector during data-taking. This allows us to understand the detector's experimental conditions and performance, while also providing Monte Carlo simulations of decay channels of interest. This is fundamental in many decay analyses, including the analysis that is the subject of this dissertation. The *Gauss* simulation program is composed of two separate phases using three dedicated applications: *Pythia* (see [87], [88] and [89]), *EvtGen* (see [90]) and *Geant4* (see [91] and [92]).

  The first phase is the *Generator* Phase, where the Pythia and EvtGen applications generate the proton-proton collisions, produced particles and their subsequent decays. The second is the *Simulation* Phase, where the detector and its experimental setup are simulated in the Geant4 application. The result of these two stages is simulated information, that attempts to "mimic" in the most accurate way possible not only the particles that are produced at the interaction point, but also the LHCb detector's performance in measuring them.

- *Brunel* (see [93]) is the LHCb's event reconstruction application, which is capable of processing either real data coming from the LHCb detector or the output of the LHCb's dedicated simulation packages. The application creates *protoparticles*, which only contain information on particle momentum and trajectory, using information coming from detector *hits*. As output, the Brunel application generates reduced *Data Summary Tapes* (rDST's), which are datasets that only contain primary information on particle identification and their reconstructed tracks (see next subsection). These datasets are then submitted to the *stripping* process, which is discussed in subsection 3.5.2.

- The *DaVinci* (see [94]) package is the physics analysis software used in the LHCb

experiment. It is capable of event selection, vertex and unstable particle reconstruction. Its output information is saved in data files (*ntuples*, *trees*, and histograms are common examples) that are subsequently analyzed using the `ROOT` (see [95]) framework.

3.5.1 Track Reconstruction

Correct particle identification is fundamental for proper event reconstruction. There are only a few particles that can be directly detected by particle detectors: protons, neutrons, pions, kaons, photons, electrons and muons, and the first step to determining the momentum of these particles is to accurately reconstruct their *tracks*, or trajectories, through the detector, from the VELO to the calorimeters. The combination of particle identification (provided by the sub-detectors) and correct determination of their momentum (from their tracks) allows for the discovery of numerous other variables and quantities of interest.

A single measured event in the LHCb detector can possess more than 100 individual tracks, which may not even pass through the entire spectrometer, due to its unique geometrical design. This makes correct track reconstruction a complicated affair, and individual tracks are thus categorized into 5 different types (see Figure 3.27), based on how they are detected by the tracking system:

- *Long Tracks* pass through all sub-detectors in the tracking system: VELO, TT and the tracking stations. Since these are the tracks with the largest amount of information (being therefore well-measured), they are very important and useful for decay reconstructions.

- *Upstream Tracks* only pass through the VELO and TT stations, and generally are created by particles with lower momentum, which have their trajectories "bent" by the LHCb magnet's field out of the detector's acceptance. Since these tracks only pass through part of the tracking system, their momentum resolution is poor,

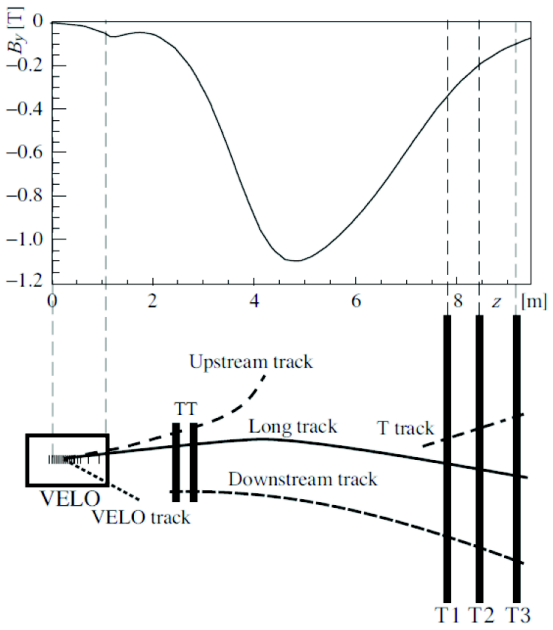

Figure 3.27: (bottom) Schematic of different track types and (top) y-component of the LHCb magnet's magnetic field through the entirety of the tracking system.

but they can also be used for decay reconstructions, flavor tagging and the study of background *noise* in the RICH1, one of the sub-detectors that these tracks pass through.

- *Downstream Tracks* only pass through the TT and Tracking stations, and are usually products of particles that decay outside of the VELO's acceptance.

- *VELO Tracks* are only measured in the VELO, usually being large angle tracks (which quickly leave the detector acceptance) or backward tracks (which move in the opposite direction as the detector layout). These tracks are useful for primary vertex reconstruction.

- *T Tracks* are only measured in the Tracking stations, and are mainly produced in secondary interactions that occur after the main collisions at the interaction point, and are used in the RICH2 sub-detector.

While the specific decay algorithm is different for each type of track, they all use the same fitting strategy, known as the *Kalman Filter* (see [96]), which is designed to fit tracks in systems that take track measurements at different times. During the fitting process, new measurements can be added to the *Kalman Filter*, progressively increasing fit quality. The *Kalman Filter* also accounts for multiple scattering and energy loss due to particle interaction with sub-detector materials, and the quality of the reconstructed tracks is measured by both the χ^2 and *pull* distributions of track parameters.

The *Kalman Filter* is bi-directional, and the fitting process starts at the farthest possible position from the interaction point and develops in the opposite direction to the particle's trajectory, until it reaches the closest possible position. After reaching this point, the fitting process reverses direction and updates the fit. This process may occur multiple times for each track.

Of the five different types of tracks, only the Long and Downstream Tracks are used for any actual physical analysis of a decay, since they are the only tracks that interact with Tracking system sub-detectors placed before and after the LHCb magnet, and therefore are the only tracks that can be used to determine momentum for event reconstruction. However, since Long Tracks have more sub-detector information than Downstream Tracks, their reconstruction efficiency is larger.

On average, the efficiency for Long Track particles with momentum above 10GeV/c is $\approx 94\%$, while the momentum resolution varies from $\delta p/p = 0.35\%$ to $\delta p/p = 0.55\%$ for momentum values above 10GeV/c and below 140GeV/c, respectively. For Downstream Tracks with momentum above 5GeV/c, the same efficiency is $\approx 80\%$ and the resolution momentum is 0.43%.

3.5.2 *Stripping*

The reduced Data Summary Tapes that are generated by the Brunel application (for real or simulated collisions and decays) and automatically recorded and saved are only referred

to as *reduced* because of the small amount of information they possess on individual particles (only reconstructed tracks and particle identification). However, even with this small amount of information, rDST's represent a volume of data that is too large to be directly analyzed, and they must still be submitted to a final pre-selection phase in order to form complete DST's. This pre-selection phase is known as the *stripping* phase, in which the data in the rDST's is further catalogued using a series of loose cuts that are specific to different decays and studies, such as B-meson decay channel cuts, or inclusive cuts to select displaced decay vertexes due to high invariant mass, for example.

Due to the various different cuts and requirements for each channel, the *stripping* phase cuts are organized into *stripping lines* that are subsequently grouped into *streams*, and each *stream* is used for decay channels that share common characteristics. The details pertaining to the *stripping line* used for the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ analysis will be discussed in section 4.1.2.

The *stripping* phase produces analyzable data, and recording, generating, and supplying the data that is needed by CERN's multiple experiments requires considerable computing power (see Figure 3.28). This computing power is accomplished through the *Worldwide LHC Computing Grid* (WLCG), or simply *Grid* (see [97]). The *Grid* is an international computer network that automatically spreads the computing tasks required by different analyses and users to computers located across the planet. This allows researchers in locations without dedicated computing banks to carry out potentially complex data analysis remotely.

The trigger phase is carried out in the *online* system at CERN, but all data production having to do with the reconstruction and *stripping* stages are spread across different regions, organized into *Tiers*. The only *Tier*-0 region is CERN itself, where main processing occurs and from which raw data is distributed to the seven *Tier*-1 regions. There are also *Tier*-2 regions[11], which exclusively produce Monte Carlo simulations.

[11] An example of which is the *Centro Brasileiro de Pesquisas Físicas* (CBPF), located in Rio de Janeiro,

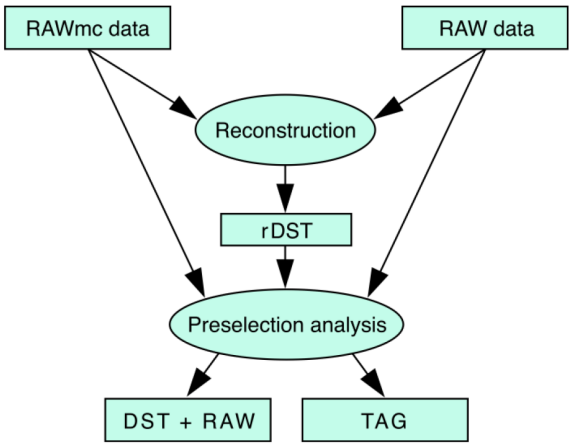

Figure 3.28: Data flow in LHCb software, from *raw* real/simulated data to fully analyzable Data Summary Tapes.

Brazil.

Chapter 4

The Search for $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$

This chapter will detail the steps of the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ analysis. The main source for the information contained in the following sections is the LHCb note for the analysis, the current version of which can be found in [98]. Section 4.1 will summarize the data/simulation samples used and the steps of the selection process. The determination and application of the *fit model* is described in Section 4.2.

It is important to note that in analyses of this kind, all stages of data selection are performed using data sets in which the *signal region* is excluded. This is referred to as *blinding*[1], and it is done to avoid any bias by the researchers. No information on the signal region is permitted during blinding, and the analysts must seek to fully understand and optimize the contributions to the regions immediately above and below the signal region (called *side-bands*) after the selection process is fully completed before *unblinding* is allowed by LHCb Collaboration reviewers.

It is also worth mentioning that the B_c's b quark *fragmentation fraction* has never been measured, and consequently its value is still unknown. Therefore, the measurement that will be performed in this analysis is the *ratio* of the product of the fragmentation and branching fractions between two *different* channels. One of these channels is $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$, while the other must be a channel whose fractions have already been measured and established. Channels that are used in this way are referred to as *control* channels,

[1]The analysis itself is also called a *blind analysis*.

and they allow for normalization of the observed signal to a known, measured quantity. In this way, depending on the significance of the final measurement, the fragmentation fraction for the B_c can be inferred, or its upper limit determined.

The determination of the optimal fit strategy for this analysis is discussed in Section 4.3, and the blinded fit tests are detailed in Section 4.4. The possible control channels for this analysis are discussed in Section 4.5. The acceptances of the simulation samples are shown after the discussion on the control channels, in Section 4.6, and Section 4.7 closes the chapter with discussions on the efficiencies, the systematic uncertainties pertaining to the study, and the final steps of the analysis strategy *after* permission is given for unblinding.

4.1 Selection

As was discussed in section 2.2, the objective of this analysis is to observe the B_c meson's decay to the three-body $KK\pi$ final state. This final state can be achieved in a number of different ways, and in this case the search is for the direct *annihilation* process. However, the data samples which are used in the analysis are not completely specific to certain decay channels, meaning that many channels with the same final state can be included in the same sample. Separating the the desired annihilation process from all other contributions is the first goal of the analysis, and it is accomplished through the *selection* process, which is arguably the most important step in any high-energy physics analysis based on data samples.

As the objective of an analysis such as this one is the isolation of a rare signal, the selection process' goal is to reduce as much as possible any events that do *not* come from the desired decay process (called *noise*), while at the same time preserving as much as possible the amount of desired signal candidates. However, most strategies that seek to eliminate noise have the side effect of also affecting signal events. Because of this, there are two relevant quantities that the selection process seeks to maximize, and the first is

purity, the ratio of the number of signal events and the *total* amount of events (signal or otherwise) in a data sample. The higher the purity value, the easier it is to significantly observe the desired signal. The second quantity is the *efficiency*, which is the ratio of signal events before and after the selection process. However, it is usually challenging to maximize both quantities simultaneously, and in the context of very rare decays that use data with limited statistics, it is desirable to sacrifice efficiency in order to obtain the highest possible purity.

The *performance* of the selection is characterized by these two quantities, and its optimization is an essential stage of the analysis. As calculating the purity and efficiency separately can be difficult and inefficient, the optimization of the selection is accomplished through the use of a *Figure of Merit*. The FoM is a quantity that is a function of a number of parameters, such as efficiency and number of events, and the main way of establishing the best selection configuration is the maximization of this quantity (see section 4.1.5).

This section will first describe the samples used and the chosen trigger/preselection configuration (subsections 4.1.1 and 4.1.2), followed by the multivariate analysis (subsection 4.1.3), particle identification (subsection 4.1.4), and the optimization of the final selection (subsection 4.1.5). More technical information regarding these subsections can be found in Appendix A.

4.1.1 Data and Simulation Samples

This analysis used the full data sample collected by LHCb in Run I of the LHC, which took place in 2011 and 2012. This represents an integrated luminosity of approximately 3.0 fb^{-1}, separated as follows:

- 2011 data: 559.86 ± 9.57 pb^{-1} (MagDown) and 418.23 ± 7.15 pb^{-1} (MagUp), total: 978.09 ± 16.72 pb^{-1}.

- 2012 data: 991.26 ± 11.50 pb^{-1} (MagDown) and 999.42 ± 11.59 pb^{-1} (MagUp), total:

1990.68 ± 23.09 pb^{-1}.

Numerous 2011 and 2012 Monte Carlo simulations were also used, representing not only the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ signal, but also many background contributions coming from cross-feeds and partially reconstructed events, as well as the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ control channel (see Section 4.5). Detailed information on all the simulated samples can be seen in Tables A.1 and A.2.

4.1.2 Triggers and Pre-selection

The data recorded by the LHCb detector is organized into reduced Data Summary Tapes (as mentioned in subsection 3.5), that are then submitted to an additional pre-selection stage called the *stripping* phase, further cataloguing the data using loose cuts that are specific to each decay study. However, before the stripping phase the decay information collected by the detector must first pass a series of trigger decisions. In the case of the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ analysis, as the final state is composed of three hadrons, trigger decision requirements are applied at all trigger levels (L0, HLT1, and HLT2) to select hadronic decays. It is important to note that trigger decisions that *accept* an event can be classified into two types: *Trigger on Signal* (TOS) if the trigger objects associated with the *signal* are enough to trigger the event, and *Trigger Independent of Signal* (TIS) if the event could have been triggered by trigger objects that are *not* associated with the signal. A TIS-type decision is desirable, as it carries less bias associated with the chosen signal. Specifically, the high-level trigger lines are requested to be Hlt1TrackAll TOS [99] for the HLT1 decision and Hlt2Topo(2- or 3-Body) TOS [100] for the HLT2 decision. At the L0 level events are permitted to be selected by either the hadron line (L0 Hadron TOS) or selected independently of the signal by any L0 line (L0 Global TIS). Tables 4.1 and 4.2 show the selection criteria for the high-level trigger lines used.

The next step is the stripping phase, that further specifies and prepares the data for analysis using loose kinematical and topological cuts that are performed for a series of

Selection Criterion	AllL0 Line
L0 Channel	All Physics Triggers
VELO Track IP	$> 125\mu$m
VELO Track Hits	> 9
Missed VELO Track Hits	< 3
Track Momentum	> 12.5GeV
Track p_T	> 1.5GeV
Track χ^2	< 3
Track IP χ^2	> 36

Table 4.1: Values of selection cuts used in HLT1 decision. *Missed hits* refers to the difference between the number of hits associated with the VELO track and the number of expected hits, according to the track's direction and its first measured point.

different variables (21 in all). These cuts attempt to select, for each event, two oppositely-charged kaons and a pion that are produced at the same secondary vertex (SV)[2] and that also have invariant mass in the B_c mass *window*. The invariant mass is calculated using energy and momentum measurements of the identified final state particles, through relativistic kinematical equations.

$$\begin{aligned} E_{B_c} &= E_K + E_K + E_\pi \\ \vec{p}_{B_c} &= \vec{p}_K + \vec{p}_K + \vec{p}_\pi \\ m_{B_c} &= \frac{1}{c^2}\sqrt{E_{B_c}^2 - (p_{B_c})^2 c^2} \end{aligned} \tag{4.1}$$

The B_c meson has a relatively short lifetime, and its flight distance (distance between the primary and secondary vertices) is directly related to this time. Therefore, the strip-

[2]The secondary vertex marks the location where the B_c meson decays, and the location where the meson itself is produced is called the *primary* vertex.

Quantity	Selection Criteria
All input particle transverse momenta	$p_T^{min} > 500$MeV
All input particle momenta	$p^{min} > 5$GeV
All input particle track χ^2 values	track $\chi^2_{max} < 5$
All input particle IP χ^2 values	IP $\chi^2_{min} > 16$
B candidate corrected mass	$4\text{GeV}/c^2 < M_{corrected} < 7\text{GeV}/c^2$
Hardest daughter momentum	$p_T^{max} > 1.5$GeV
Best daughter track χ^2	Track $\chi^2_{min} < 3$
Sum of daughter transverse momenta	$\sum p_T > 4, 4.25$GeV (2, 3-body)
Sum of daughter IP χ^2	$\sum \text{IP}\chi^2 > 100, 150$ (2, 3-body)
Particle/Particle and Particle/n-body DOCA	DOCA < 0.15 mm
B candidate flight distance χ^2	FD$\chi^2 > 64$
B candidate signed flight distance	FD > 0
Prompt D veto	$M > 2.5\text{GeV}/c^2$ or 2, 3-body IP $\chi^2 > 16$

Table 4.2: Selection criteria used in the HLT2 topological lines.

ping cuts for this analysis take this into account by accepting more events with lower flight distances and lifetimes. Many other variables are used in the stripping lines, and are related to particle track momentum, the minimum distance between tracks (DOCA[3]), particle identification probability (ProbNN variables), track and B_c candidate *impact parameter* (the smallest distance between a track and its vertex), and finally the *quality* of reconstructed variables, which is represented by the χ^2 quantity. The data samples used by this analysis were selected by the `StrippingBc2hhh_kkpi_exclLine` stripping *line* that contains all performed cuts, which can be seen on Table 4.3. Not listed in the table

[3]Distance Of Close Approach

is the additional requirement that all selected tracks *not* be muons, using information from the muon subdetector.

After the *online* application of the trigger lines and stripping requirements, the simulation and data samples are produced with their own *offline* trigger variables. These variables are used by the researchers to perform offline trigger cuts, which are specific to each analysis and can be modified if necessary. The cut values chosen for this study are shown in Table 4.4.

While they are important aspects of the analysis, it is worth mentioning that the definition of the trigger lines, stripping requirements, and offline trigger cuts was performed by the LHCb Collaboration, and not by the author of this dissertation. The existing trigger lines were studied, and the chosen line was the one most adequate for this study. The following subsections detail the steps that are specific to this analysis.

4.1.3 Multi-Variate Analysis

The B_c signal is contained in the data samples, hidden underneath much more numerous contributions that need to be excluded in order for any signal observation to be possible. The most important noise contribution is the combinatorial background, and it is reduced primarily through a second set of cuts on kinematical and topological variables. Many different quantities can be used to reduce noise, but individual variables do not have large *separation power*, and their cuts are thus not very efficient at achieving this rejection. In fact, it is impossible to achieve the required purity and efficiency using cuts on only one variable, and therefore methods that combine many different variables and their cuts are necessary. Many successful attempts to perform event classification simultaneously with multiple variables were made by high-energy physics collaborations, and many different methods were developed. These methods were organized into a C++ package called the Toolkit for *Multivariate* Analysis (TMVA) [101], which today is automatically implemented into the `ROOT` framework and specifically designed for high-energy physics

Variables	Stripping cuts
Tracks p_T	> 0.25 GeV/c
Tracks p	> 2.5 GeV/c
Tracks IPχ^2	> 1
Tracks χ^2/n.d.f.	< 3
`ProbNNghost(track)`	< 0.5
`ProbNNk(Kaon)`	> 0.2
`ProbNNpi(pion)`	> 0.15
Maximum DOCA between tracks	< 0.2 mm
p_T of the highest-p_T track	> 1.5 GeV/c
Sum of p_T of tracks	> 4.5 GeV/c
Sum of p of tracks	> 22 GeV/c
Sum of IPχ^2 of tracks	> 200
$B_c^\pm$ candidate $M_{KK\pi}$ window	$[6.0, 6.5]$ GeV/c^2
$M_{KK\pi}$ window for control channels	$[5.1, 5.5]$ GeV/c^2
$B_c^\pm$ candidate IP χ^2	< 10
$B_c^\pm$ candidate P_T	> 1 GeV/c
Distance from SV to any PV	> 1.5 mm
Secondary Vertex χ^2	< 20
$B_c^\pm$ candidate pointing angle	> 0.9998
$B_c^\pm$ Flight Distance χ^2	> 150
Number of long tracks in the event	< 200

Table 4.3: Stripping selection criteria for $B_c^\pm$ decays to $K^\pm K^\mp \pi^\pm$.

Offline Trigger Variable	Cut Value
`B_L0HadronDecision_TOS2`	1
`B_L0Global_TIS2`	1
`B_Hlt1TrackAllL0Decision_TOS2`	> 0
`B_Hlt2Topo2BodyBBDTDecision_TOS2`	> 0
`B_Hlt2Topo3BodyBBDTDecision_TOS2`	> 0
`B_Hlt2Topo4BodyBBDTDecision_TOS2`	> 0

Table 4.4: Values of selection cuts used for the offline trigger variables.

applications. The MVA techniques are based on supervised learning algorithms, that use *training* events with a known output to "teach" the algorithm to produce a function. This function, if trained correctly, is then capable of generalizing from the training data to new situations without bias.

In this case, the variables with highest separation power contained in the data sets are used as input to produce a single function, that by itself has more separation power than the variables used to train it. This allows for noise rejection using cuts on only one "variable", specifically generated for that purpose. The TMVA package automatically selects variables with the highest desired separation, and in this case 23 were chosen to be used as input. This constitutes a set of variables in which there is a hierarchy of separation power that does not exceed two orders of magnitude, i.e. the *less* separating variable has a separation power of at least 1% of the *most* separating variable. Of these 23 variables, 9 are related to single-track quantities (χ^2 per degree of freedom, ghost[4] probabilities and momenta), 7 are associated with the B_c candidate (p, p_T, vertex χ^2, impact parameter, impact parameter χ^2, flight distance, and cosine of the angle between the flight vector and momentum), and 7 deal with track combinations (pointing p_T, p_T^{max}, p_T^{med}, impact parameter of the track with p_T^{max}, sum of impact parameters, sum of impact parameter χ^2, maximum distance of closest approach between any pair of tracks). Some of the variables are quite highly correlated, which is usually something to be avoided, but they are all found to positively impact the discriminating power of the MVA function, and are therefore all kept in the training process. More detailed descriptions of the 23 variables, the table ranking them by separation power, and the correlation coefficient matrices can be seen in A.2.

When training a MVA function, the input data needs to clearly discriminate between signal and background events, in order to properly teach the algorithm what kinds of

[4]Pseudo-random combinations of hits, formed by the reconstruction algorithms and that do not represent real particles.

events it needs to select. Therefore the signal events are represented by a flat phase space $B_c^+ \to K^+K^-\pi^+$ three-body simulation, while the background events are taken from the side-bands in the data samples, defined as ranges in the $m_{K^+K^-\pi^+}$ invariant mass. Since the signal region is excluded from the data sets, these ranges correspond to $m_{K^+K^-\pi^+} = [6, 6.15]$ GeV/c^2, $[6.35, 6.5]$ GeV/c^2. Graphs generated by the TMVA showing signal and background distributions of the input variables can be seen in Figures 4.1 and 4.2.

With clearly defined signal and background events, the MVA's calculation of the separation power for a given variable is thus:

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(\hat{y_S}(y) - \hat{y_B}(y))^2}{(\hat{y_S}(y) + \hat{y_B}(y))^2} dy \tag{4.2}$$

where $\hat{y_S}$ and $\hat{y_B}$ are the signal and background probability density functions for a given variable y. In principle there is no difference between background input data from the lower and higher-mass side-bands, but in this case, due to observed fluctuations in the lower side-band that could possibly represent peaking structures[5] (see Figure 4.3), only events in the upper-mass side-band were used in the MVA training.

There are many functions that the MVA training can produce, and for this analysis, the performance of 5 different ones were compared: *Boosted Decision Tree* (BDT) [102], *Fisher* [103], *Likelihood Principal Component Analysis* (LikelihoodPCA) [104], *Likelihood* [105], and *Multi Layer Perceptron* paradigm (MLPCE). Graphs generated by the TMVA showing background rejection versus signal efficiency curves for each MVA technique can be seen in Figure 4.4. Optimally, the chosen MVA technique should maximize both background rejection and signal efficiency, and in this case the BDT technique is clearly seen to be the best choice. When compared to linear methods the BDT has noticeably

[5]Contributions coming from other decay channels that are incorrectly identified in the reconstruction process.

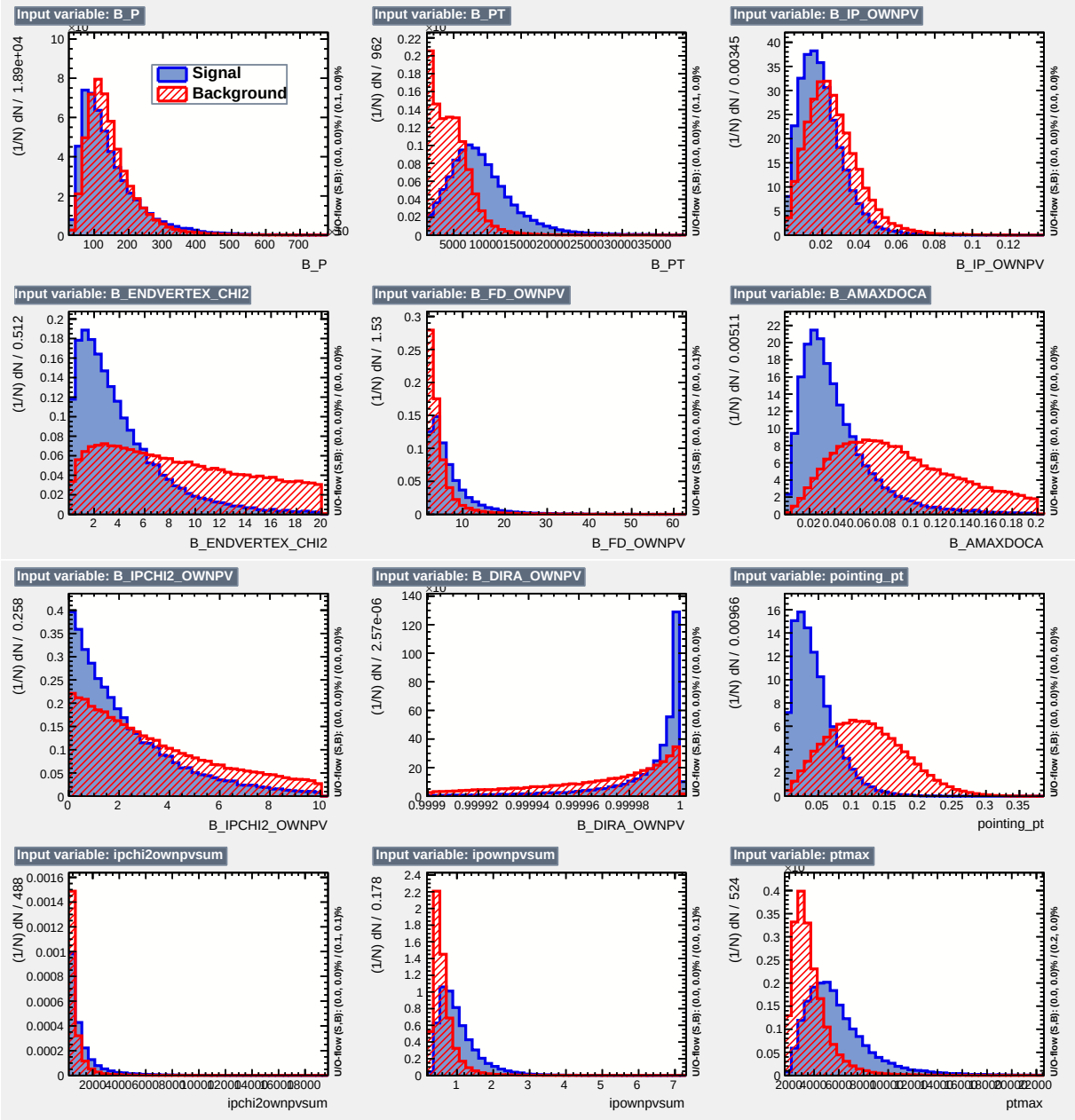

Figure 4.1: First part of the distributions of the input variables for the multivariate analysis training. The signal is represented in plain blue and the background in hatched red.

better performance, and in comparison with the MLPCE technique the BDT performs similarly but with much less computing time for the training. Checks were performed after the training to make sure there was no bias due to overtraining or sample specificity,

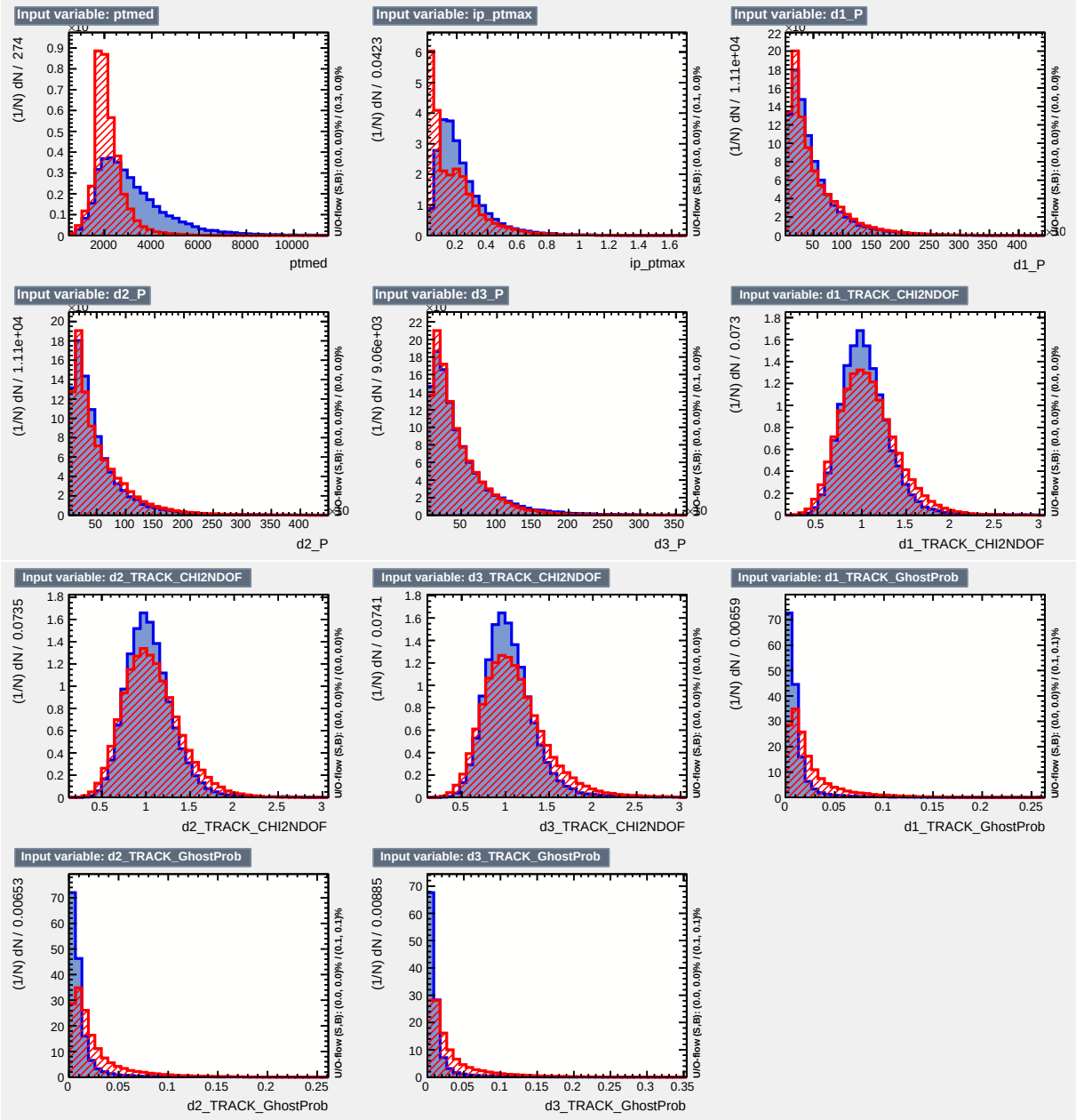

Figure 4.2: Second part of the distributions of the input variables for the multivariate analysis training. The signal is represented in plain blue and the background in hatched red.

and are detailed in subsection A.2.3.

After the overtraining check, the final two BDT shapes (one for each year) are obtained using full statistics (both side-bands), and then compared when applied to the data and

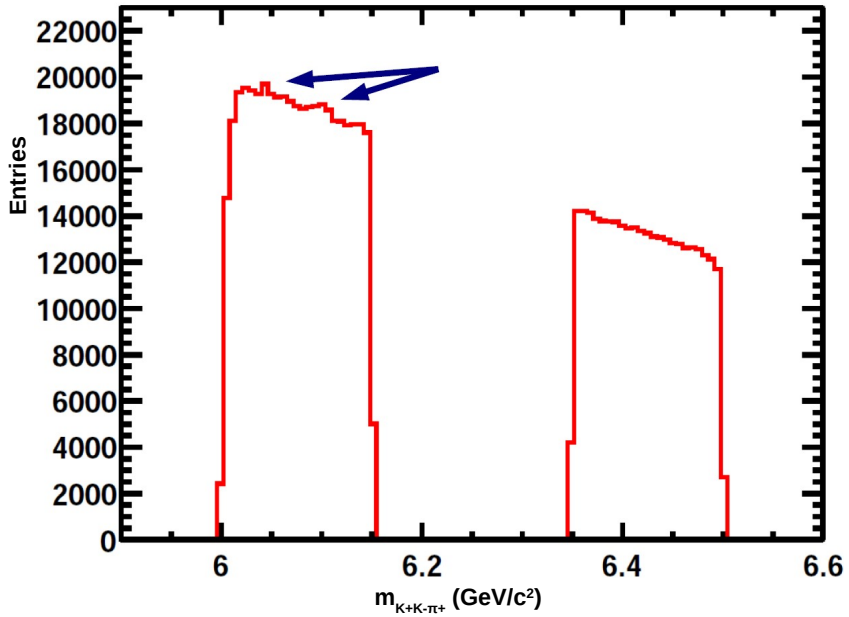

Figure 4.3: Distribution of side-band data before any MVA selection. Arrows indicate possible peaking structures in the lower side-band region.

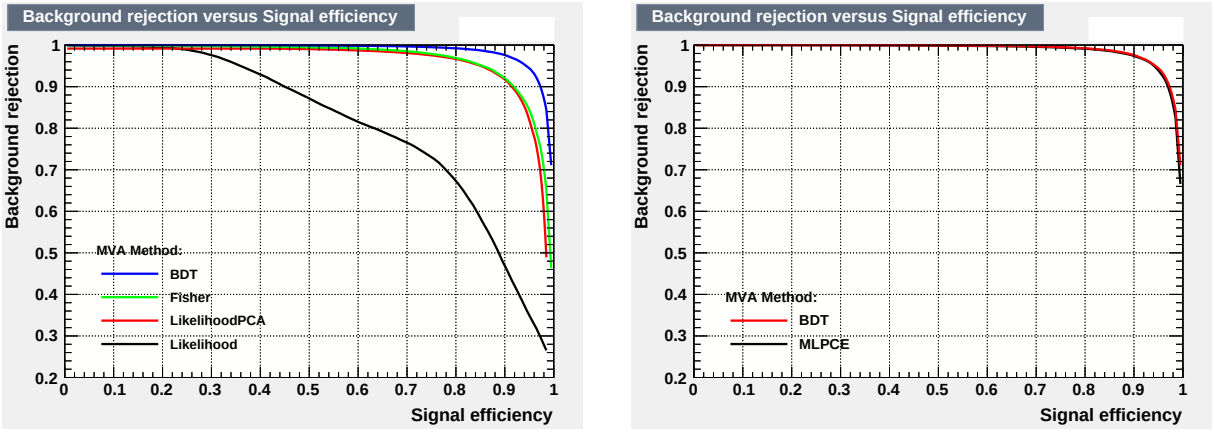

Figure 4.4: Background rejection vs. signal efficiency curves for (left) BDT and linear methods and (right) BDT and Multi Layer Perceptron paradigm.

Monte Carlo samples for both years (BDT_{2011} applied to year 2012 and vice versa). The BDT shapes for each year are not expected to differ in a significant way, and the result of the comparison is shown in Figure 4.5. Cuts on the final BDT variables are successful in reducing unwanted background contributions, but cuts that are too *tight* (i.e. at relatively

high values of the BDT) may accomplish this at the cost of also rejecting signal events. The optimal cut is determined through a simultaneous study of the BDT and particle identification variables, which are discussed below.

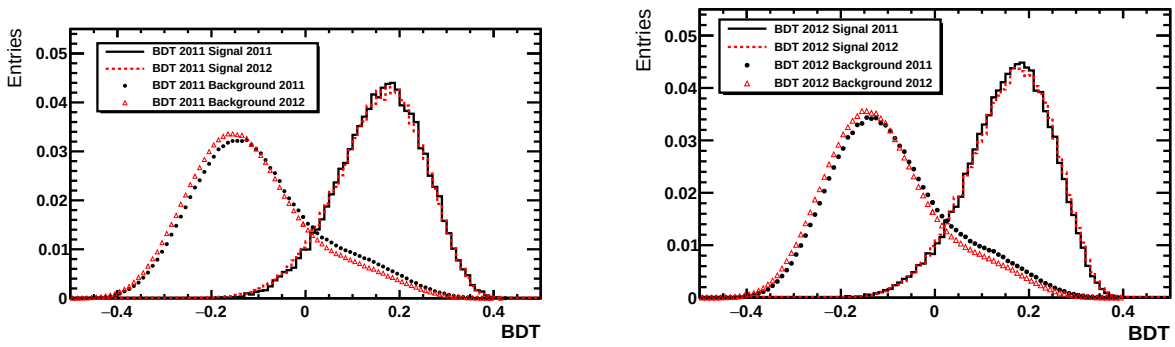

Figure 4.5: Superimposition of (left) BDT_{2011} and (right) BDT_{2012} for both 2011 and 2012 simulation and data samples, represented by red and black, respectively.

4.1.4 Particle Identification

As mentioned previously, the data and simulation samples that are used in the analysis contain a series of kinematical and topological variables, with values for each entry in the sample. However, these are not the only kinds of variables contained in the samples. The final tool used in the selection process are *particle identification* (PID) variables, that use information from the LHCb detector subsystems to give values of probability on the particle identity of any given entry. These variables allow for *offline* optimization of PID cuts, and are widely used in high-energy physics analyses.

There are two different kinds of PID variables, and both were tested in this analysis. The first kind are called *Difference in Log Likelihood* (`DLL`) variables, which are produced using particle identification information taken from the muon, RICH, and calorimeter systems of the LHCb detector (see 3.3). This information is in the form of *likelihoods* whose logarithms are then added linearly, forming a combined set. The variable then describes the probability that a given track belongs to a particle of certain mass, relative

to the hypothesis that the track belongs to a pion[6], as pions are the most commonly produced particles in hadronic detectors.

The second kind are the `ProbNN` variables, which are newer and use more data than the `DLL` variables, adding information from the tracking system with the information from the RICH, calorimeter, and muon detector. This data is then combined with a MVA, and the final `ProbNN` variables are obtained through training using simulation samples. Since they are trained, the performance of the obtained `ProbNN` variables depends on the simulation samples used. In this case, the MCTuneV2 version of the variables are used.

4.1.5 Optimization

As the final state for this analysis includes both pions and kaons, the criterion for their identification is studied separately, and for each class of particle, a range of PID variable cuts are tested and *simultaneously* optimized with cuts in the BDT variable. This optimization is accomplished through the use of the *Punzi criterion*[106] as a figure of merit:

$$\frac{\epsilon_S}{a/2 + \sqrt{B}} \tag{4.3}$$

In this equation, ϵ_S is the product of the BDT cut and PID efficiencies, B is the expected number of background events in the signal region $m_{K^+K^-\pi^+} = [6.15, 6.35]$ GeV/c^2, and a is the number of sigma for the signal significance ($a = 3$ or 5, representing evidence and observation levels, respectively). The determination of the other quantities is detailed below.

- **BDT Cut Efficiency** - Determined using signal simulation samples. The BDT variable is added to the samples and various cuts are performed. For each cut, the

[6]For instance, `DLL` $(\mu - \pi)$ expresses how likely a chosen track belongs to a particle with the mass of a muon, relative to the hypothesis that the same track belongs to a particle with the mass of a pion.

BDT efficiency is the ratio between the number of events that "pass" the cut and the total number of events (before any cut is performed). In general, the BDT efficiency decreases as the cut value increases.

- **PID Efficiency** - The previously discussed PID variables are also present in the Monte Carlo simulation samples, but they are not used for the efficiency calculation. This is because the PID variables are not well described by the simulations, varying from those found in data. Instead, PID efficiencies are obtained through the use of a series of specific decay channels whose reconstruction is possible using only kinematical information, and therefore does *not* rely on particle identification. These channels are called *calibration samples*, as their decays can be used to study other processes with similar characteristics. The efficiencies of PID variable cuts can therefore be estimated for simulated samples, and for this purpose the `ROOT` framework is equipped with the `PIDCalib` package [107]. The `PIDCalib` package produces variables specific to both the simulation sample and the desired PID configuration, and these variables, once added to the simulation samples, contain the PID efficiency for *each entry*. The final PID efficiency for each BDT cut is then the mean of the PID efficiencies for all the entries that pass the cut. For this analysis, the $D* \to D^0(K\pi)\pi$[7] calibration sample was used.

- **B** - The expected number of background events in the signal region is determined using the upper side-band of the data samples. For each BDT cut, the distribution of surviving events is fitted using a linear function (see Figure 4.6), which is then properly normalized [8]. The normalized function is then integrated in the signal region, giving an estimation of the number of background events for each BDT cut.

[7] There are many other decay modes that can be used for calibration, $K^0_S \to \pi^+\pi^-$ and $\Lambda^0 \to p\pi^-$ are two examples.

[8] As event distributions are shown in histograms, in this case the normalization is due to *bin width*.

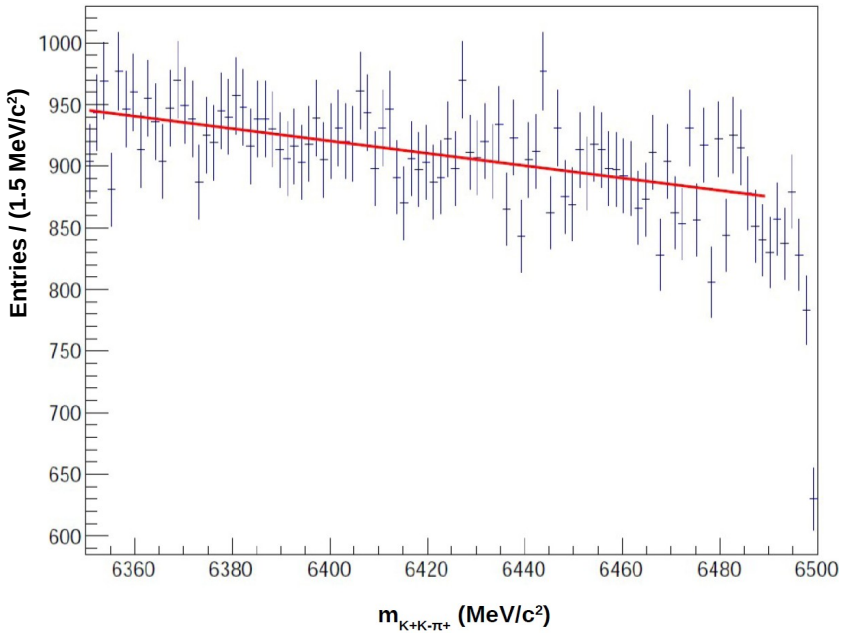

Figure 4.6: Example of an upper mass side-band distribution following a BDT-PID combination cut, and the fitted linear function (in red). The sudden drop-off at the right side of the distribution is an artificial edge effect, which is why the fit does not take it into account.

The optimization process is accomplished by testing various BDT cuts and PID configurations, and finding which BDT-PID combination *maximizes* the Punzi figure of merit. The kaon identification was studied first, by *freezing* the pion PID variable at `DLL` $(K - \pi) < -5$ and testing four `ProbNNk(kaon)` $> x$ cuts, with $x = 0.2$, 0.3, 0.4, and 0.5. Simultaneously, 11 BDT $> x$ cuts are analyzed, with $x \in [0, 0.2]$. Both 2011 and 2012 simulation and data samples were studied, but the 2011 data samples in particular have smaller statistics than those for 2012, and at tight cuts the background estimation is unreliable at best. Therefore, the *frozen pion* study's conclusions were taken from the 2012 test results. The Punzi criterion values of all 44 BDT-PID combinations for 2012, with both $a = 3$ and 5, are shown in Figure 4.7. Table 4.5 shows the BDT-PID combinations that supplied the highest Punzi criterion value, for both years and values of a. From these results, the final PID variable cut for kaons is chosen to be `ProbNNk(kaon)`

> 0.4.

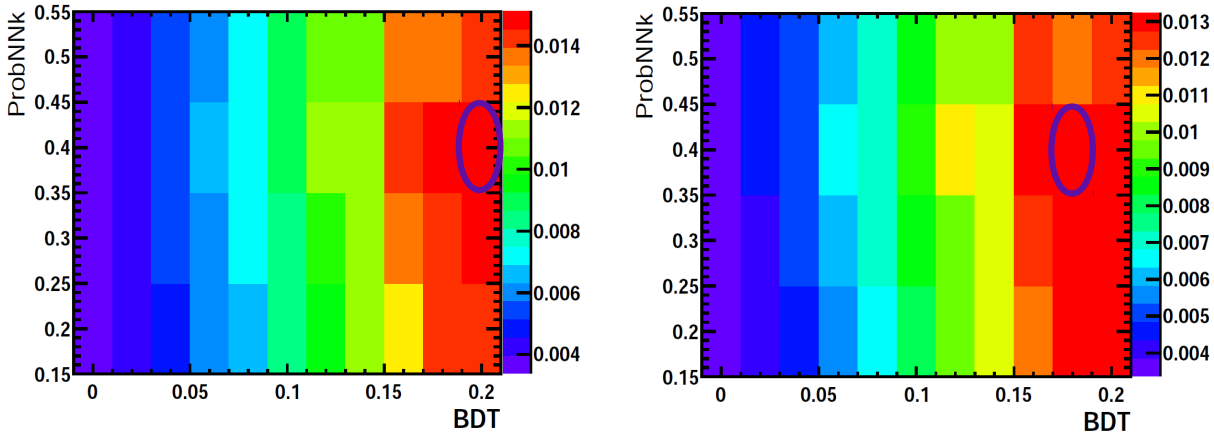

Figure 4.7: Variation of the frozen pion study's 2012 Punzi figure of merit as a function of the BDT and PID requirements, for (left) a=3 and (right) a=5. The color scale on the right indicates the values of the figure of merit for each cell, and the highest-valued cell is circled in purple.

FoM	Year 2011	Year 2012
Punzi a = 3	BDT = 0.20 `ProbNNk` = 0.3	0.20 0.4
Punzi a = 5	BDT = 0.20 `ProbNNk` = 0.3	0.18 0.4

Table 4.5: Positions of the maximum of Punzi figures of merit for the frozen pion study.

With the determination of the optimal kaon PID cut, the pion PID configuration is studied next. Two cuts are analyzed for `DLL` $(K-\pi) < z$, with $z = 0$ and -5, and five cuts are analyzed for `ProbNNpi(pion)` $> y$, with $y = 0.15$, 0.2, 0.3, 0.4, and 0.5. The BDT is analyzed in the same fashion as the frozen pion study, with 11 cuts ranging from 0 to 0.2. Also similarly to the pion study, the lack of statistics for the background estimation with tight cuts for year 2011 makes its results less reliable. The values of the Punzi criterion for all 77 BDT-PID combinations of 2012, with $a = 3$ and 5, are shown in Figure 4.8, and

Table 4.6 shows the optimal configurations. All highest-value combinations have BDT values close to the tightest considered cut (BDT > 0.2), and the reason for not testing even tighter configurations is the amount of events that pass the cuts. At BDT > 0.2 the number of surviving events is minimal, making it difficult to fit properly and extrapolate any physically meaningful number of events in the signal region.

From the results of these two identification studies, the final BDT-PID combination is chosen to be:

- **BDT** > 0.18
- `ProbNNk(kaon)` > 0.4
- `ProbNNpi(pion)` > 0.5

This configuration represents the optimal *tight* cut scenario. However, in order to seek maximum statistical significance in the final result, two more BDT bins in the region BDT $\in [0.04, 0.18]$ will be considered for a *simultaneous* fit (see subsection 4.3.2).

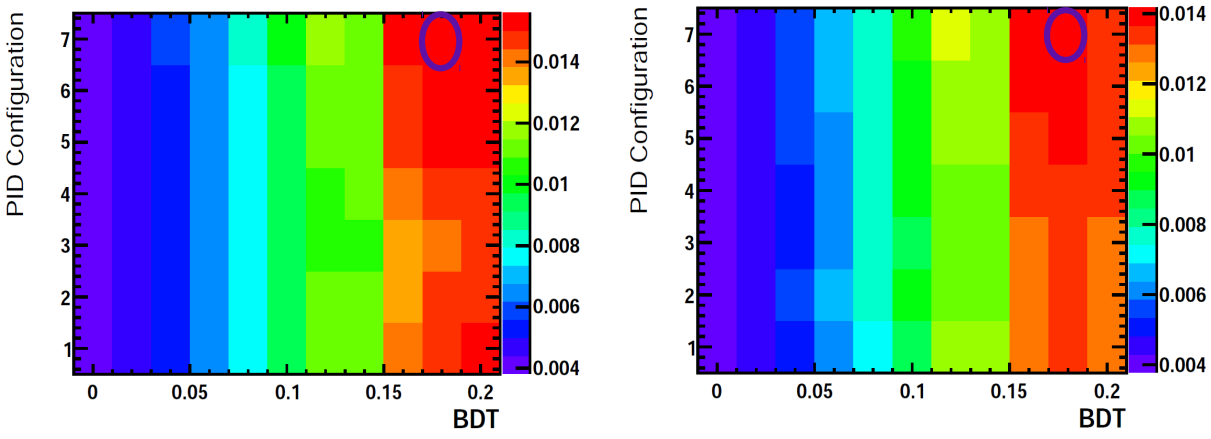

Figure 4.8: Variation of the frozen kaon study's 2012 Punzi figure of merit as a function of the BDT and PID requirements, for (left) a=3 and (right) a=5. The color scale on the right indicates the values of the figure of merit for each cell, and the highest-valued cell is circled in purple. PID Configurations 1 and 2 are for the `DLL` variable, while 2-7 are for the ProbNNpi variable.

FoM	Year 2011	Year 2012
Punzi a = 3	BDT = 0.20 `ProbNNpi` = 0.2	0.18 0.5
Punzi a = 5	BDT = 0.20 `ProbNNpi` = 0.2	0.18 0.5

Table 4.6: Positions of the maximum of Punzi figures of merit for the frozen kaon study.

4.2 Fit Model

With the final BDT and frozen PID variable cuts, the following stage in the analysis is the determination of the *fit models* for all the components present in the B_c mass window, in order to construct a fit *probability density function* (PDF) capable of fully modelling the post-cut mass distribution. This means finding mathematical functions that best describe the expected distribution of events ("shapes") for these components, which include the signal, combinatorial background, cross-feeds, and partially reconstructed decays. The most optimal fit model is determined through the maximum likelihood method, which is implemented through the *RooFit Toolkit for Data Modeling* [108], a collection of packages incorporated into the `ROOT` framework, which automatically attempts to fit distributions[9] using user-specified functions. The ranges of the function's parameters are also specified, and the fitting process is composed of the RooFit package's attempts to vary the values of these parameters in such a way as to make the PDF match the data distribution as well as possible, using the maximization of the likelihood as a criterion. Studies are performed to discover which components actually have contributions to the mass window, and those that do are modeled and added to the final fit function. In the following subsections, each component is discussed separately.

[9]RooFit also automatically performs other convenient functions during fitting, such as normalization.

4.2.1 $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm} signal$

The shape of the signal component is provided by the simulation samples, and it is modelled using a *Double Crystal Ball* (DBC) function. The Crystal Ball function consists of a Gaussian core portion and an exponential tail, and the DBC is simply two of these functions used simultaneously, in order to properly describe asymmetric tails. Each Crystal Ball function is described as:

$$f(m;\mu,\sigma,\alpha,n) \propto \begin{cases} \exp(-\frac{(m-\mu)^2}{2\sigma^2}) : \frac{m-\mu}{\sigma} > -\alpha \\ A.(B - \frac{m-\mu}{\sigma})^{-n} : \frac{m-\mu}{\sigma} \leq -\alpha \end{cases} \tag{4.4}$$

where $A = (\frac{n}{|\alpha|})^n.\exp(\frac{|\alpha|^2}{2})$, $B = \frac{n}{|\alpha|} - |\alpha|$, μ is the mean and σ the width of the Gaussian core portion, n is the exponent of the exponential tail, and α the transition point at which the function "switches" from the exponential to the Gaussian regimes. Each Crystal Ball function in the DBC has its own parameter values, with the only exception being μ, which in this case is chosen to be equal for both functions.

Figure 4.9 shows the Double Crystal Ball fits (in linear and logarithmic scales) to the 2011 and 2012 simulated samples for the $m_{K^+K^-\pi^+}$ invariant mass, as well as the *pulls*[10] of the fits. Tables 4.7 and 4.8 show the parameters for both fits, where $frac(1/2)$ represents the ratio between the tail exponents (n) of each Crystal Ball function.

4.2.2 Combinatorial Background

The combinatorial background is the simplest component to model, as its shape is conventionally described by a simple offset exponential function $\propto \exp(c.(m - 6.079))$.

[10]The pull is the difference between the value of the fit function at a certain point and the value of the actual data at that point. Therefore, small pulls are desirable, as they indicate a good-quality fit.

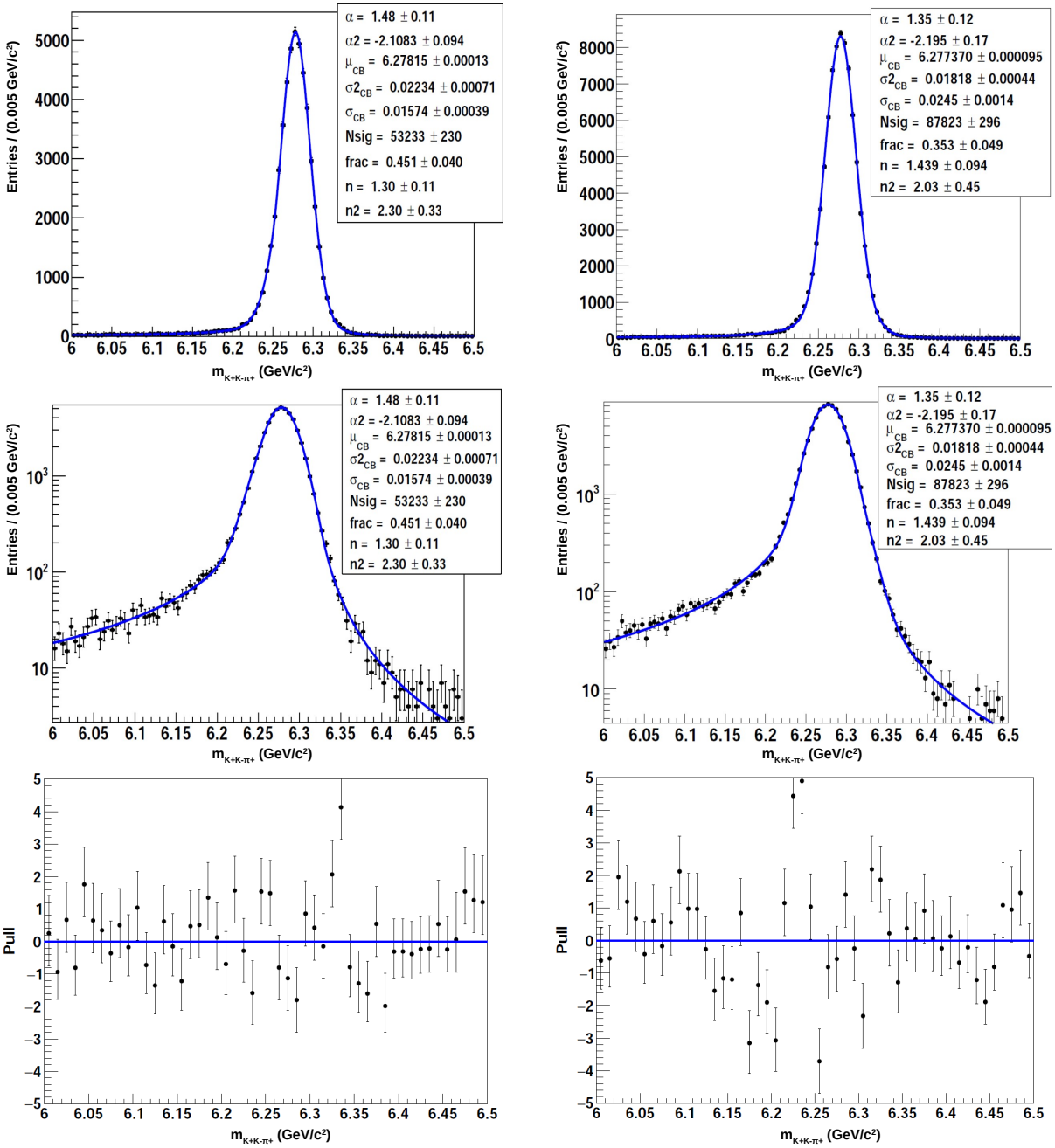

Figure 4.9: Double Crystal Ball mass fits to (left) 2011 and (right) 2012 simulation samples, with (top) linear and (middle) logarithmic scales, while (bottom) shows the pulls. The relatively large pulls present for the 2012 fit are acceptable, as the signal shape is in general very well described.

Parameter	Value
μ_S	$6.27817 \pm 0.00014\text{GeV}/c^2$
σ_1^S	$0.01580 \pm 0.00056\text{GeV}/c^2$
σ_2^S	$0.02164 \pm 0.00059\text{GeV}/c^2$
n_1	1.37 ± 0.11
n_2	2.96 ± 0.42
α_1	1.40 ± 0.11
α_2	-1.9846 ± 0.091
$frac(1/2)$	0.428 ± 0.047

Table 4.7: Parameters of the Double Crystal Ball signal model, for year 2011.

Parameter	Value
μ_S	$6.277370 \pm 0.000095\text{GeV}/c^2$
σ_1^S	$0.0245 \pm 0.0014\text{GeV}/c^2$
σ_2^S	0.01818 ± 0.00044
n_1	1.439 ± 0.094
n_2	2.03 ± 0.45
α_1	1.35 ± 0.12
α_2	-2.19 ± 0.17
$frac(1/2)$	0.353 ± 0.049

Table 4.8: Parameters of the Double Crystal Ball signal model, for year 2012.

4.2.3 $B_c^\pm \to K^\pm K^\mp K^\pm$ cross-feed

Cross-feeds are decay channels whose final states are *similar* to that of the decay being studied. In this case, the cross-feeds under consideration are the $B_c^\pm$'s decays to $K^\pm K^\mp K^\pm$, $K^\pm \pi^\mp \pi^\pm$, and $\pi^\pm \pi^\mp \pi^\pm$. Their contributions to the $m_{K^+K^-\pi^+}$ mass window are due to one or two of their final state particles being *mis*-identified, which is a somewhat common occurrence in many decays, due to the inherent complexity and uncertainty of the particle track reconstruction and identification process. In the case of the $B_c^\pm \to K^\pm K^\mp K^\pm$ cross-feed, the third kaon is incorrectly identified as a pion. As the

kaon has a higher mass than the pion, when this misidentification occurs the shape of the cross-feed is *shifted* to a lower mass, appearing slightly to the left of the signal component in this case.

The contribution of the $K^{\pm}K^{\mp}K^{\pm}$ final state coming from annihilation processes is expected to be Cabbibo-suppressed by the CKM matrix parameters related to the changes in flavor between first and second generation quarks, $\frac{|V_{us}|^2}{|V_{ud}|^2}$. In order to study this contribution, a $B_c^{\pm} \to K^{\pm}K^{\mp}K^{\pm}$ simulation sample is produced and a PID efficiency *pollution* map (seen on Table 4.9) is inferred through `PIDCalib` mis-identification (misID) rates of simulated events that pass our chosen PID cuts. In other words, `PIDCalib` produces a variable expressing the efficiency of our final PID cuts for each entry in the simulated samples, as in 4.1.5. However, seeing how this is a wholly incorrect and unwanted final state, this efficiency is called the misID rate, and the smaller it is, the better.

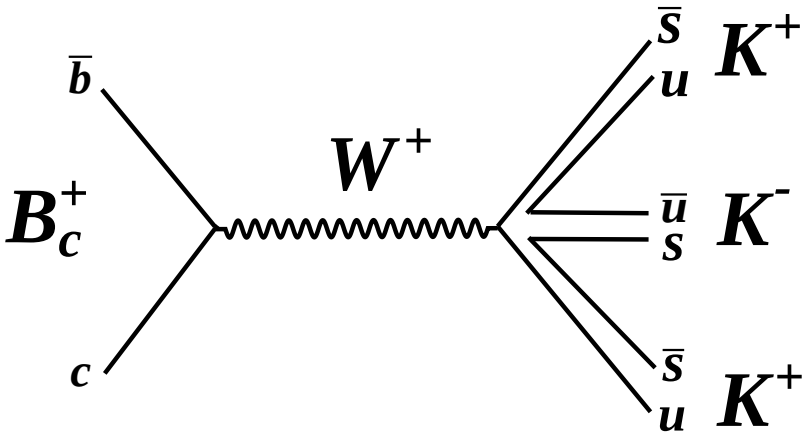

Figure 4.10: Tree Feynman diagram for the $B_c^+ \to K^+K^-K^+$ channel.

The efficiency map shows that this channel engenders a small remnant after PID selection, which may impact the final B_c mass distribution. Therefore, this cross-feed is modelled using the same model and methodology as 4.2.1, and the Double Crystal Ball

fits to the 2011 and 2012 simulated $B_c^{\pm} \to K^{\pm}K^{\mp}K^{\pm}$ samples can be seen in Figure 4.11, where the distributions have been fully reweighed by the misID rates. This means that the number of entries in each histogram bin have been multiplied by the misID percentage of that bin, more accurately representing the actual number of events that pass the PID selection, as well as the correct final shape of the distribution.

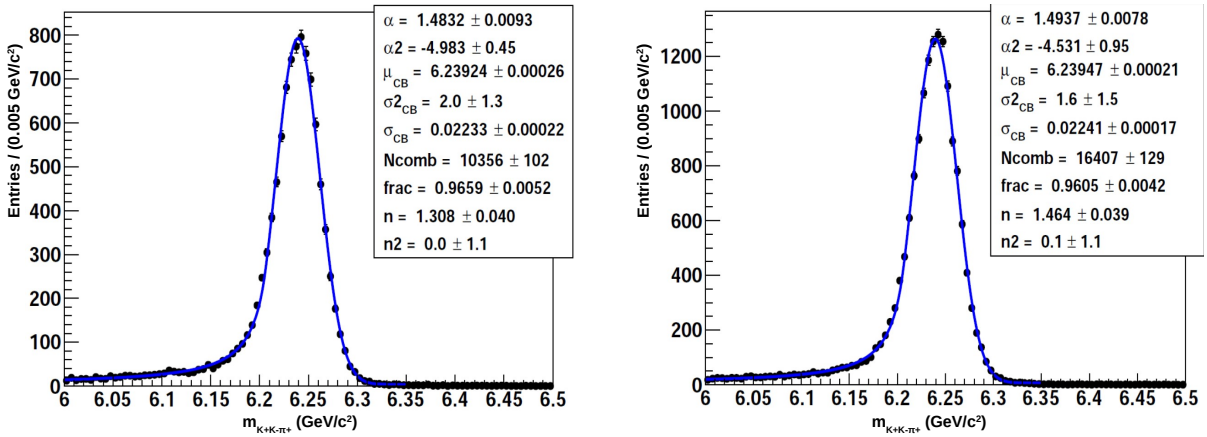

Figure 4.11: Double Crystal Ball function invariant mass fits to $B_c^{\pm} \to K^{\pm}K^{\mp}K^{\pm}$ simulated events, for (left) 2011 and (right) 2012 samples. Both distributions have been reweighed by the misID rates.

4.2.4 $B_c^{\pm} \to K^{\pm}\pi^{\mp}\pi^{\pm}$ cross-feed

As in the previous section, this cross-feed contribution is due to mis-identification of one of the final state particles, which in this case wrong identification of the first pion as a kaon. Its annihilation contribution is also expected to be Cabbibo-suppressed by $\frac{|V_{us}|^2}{|V_{ud}|^2}$, however this same final state is also obtainable through the $B_c^{\pm} \to D^0(K^{\pm}\pi^{\mp})\pi^{\pm}$ charm decay (see Section 4.2.6). A PID efficiency pollution study (Table 4.9) is also performed for this channel, where it is seen that the final PID selection engenders a remnant approximately an order of magnitude smaller than the $B_c^{\pm} \to K^{\pm}K^{\mp}K^{\pm}$. Nevertheless, this cross-feed is modelled anyway for both documentation and to test against the possible $B_c^{\pm} \to D^0(K^{\pm}\pi^{\mp})\pi^{\pm}$ contribution, using simulation samples and the same model and

methodology as 4.2.1. As the lower-mass pion is incorrectly identified as a kaon, the distribution for this cross-feed is shifted in the opposite direction as the $B_c^{\pm} \to K^{\pm}K^{\mp}K^{\pm}$, appearing slightly to the *right* of the signal component.

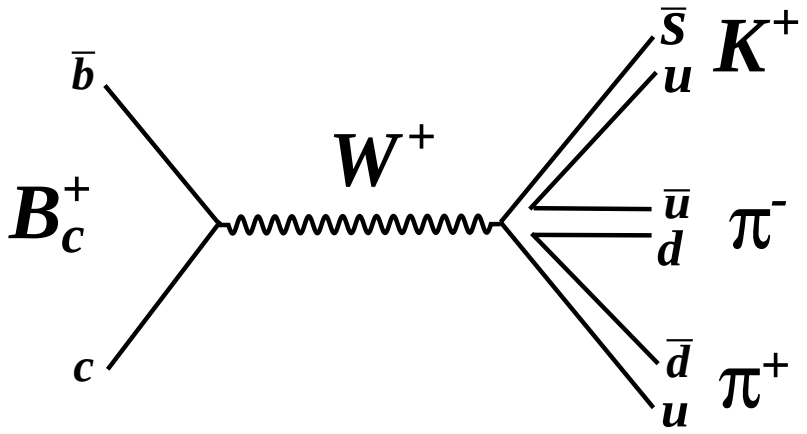

Figure 4.12: Tree Feynman diagram for the $B_c^+ \to K^+\pi^-\pi^+$ channel.

The Double Crystal Ball fits to the 2011 and 2012 simulated $B_c^{\pm} \to K^{\pm}\pi^{\mp}\pi^{\pm}$ samples can be seen in Figure 4.13. As was done in 4.2.3, the mass distributions have been reweighed with the misID rates.

4.2.5 $B_c^{\pm} \to \pi^{\pm}\pi^{\mp}\pi^{\pm}$ cross-feed

The $B_c^{\pm} \to \pi^{\pm}\pi^{\mp}\pi^{\pm}$ channel is the third and final possible cross-feed contribution, and unlike the previous two, it is *not* expected to be Cabbibo-suppressed by $\frac{|V_{us}|^2}{|V_{ud}|^2}$. In this case, as two pions are simultaneously mis-identified as kaons, the distribution of the cross-feed would be shifted to an even higher mass as the previous cross-feed channel, appearing near the right edge of the mass window. A study of the PID efficiency pollution (Table 4.9) shows that the final PID cuts are successful at reducing this channel to a truly negligible amount. For this reason, the channel is *not* modelled, as its minimal contribution would be hidden (or *absorbed*) by the combinatorial background.

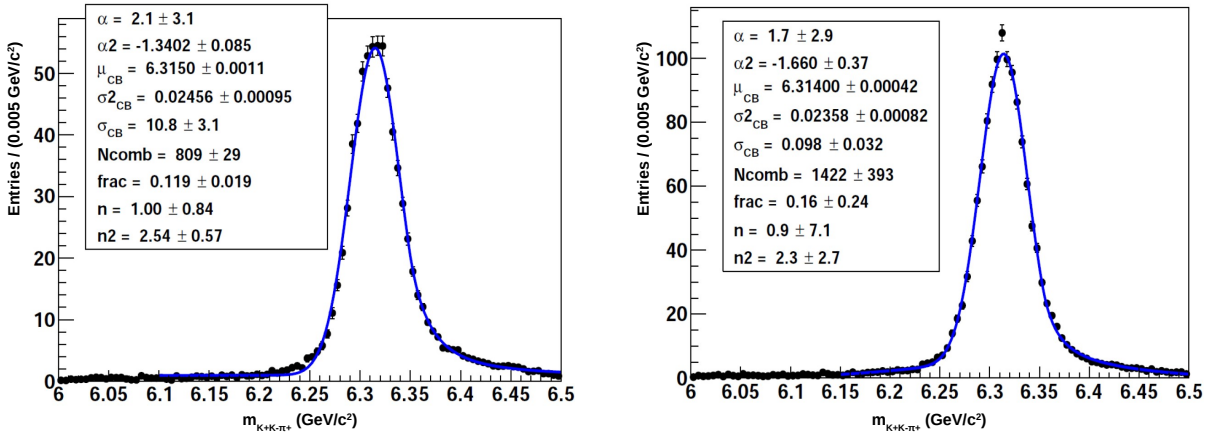

Figure 4.13: Double Crystal Ball function invariant mass fits to $B_c^{\pm} \to K^{\pm}\pi^{\mp}\pi^{\pm}$ simulated events, for (left) 2011 and (right) 2012 samples.

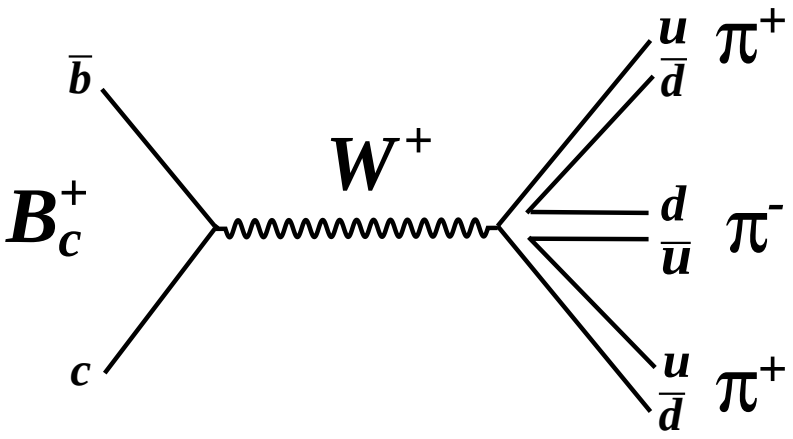

Figure 4.14: Tree Feynman diagram for the $B_c^+ \to \pi^+\pi^-\pi^+$ channel.

4.2.6 Contribution/Pollution from Charm

There are two charm decays whose contributions are important to note in this study: $B_c^{\pm} \to D^0(K^{\pm}\pi^{\mp})\pi^{\pm}$ and $B_c^{\pm} \to D^0(K^{\pm}\pi^{\mp})K^{\pm}$, which are themselves the subject of a dedicated LHCb study (see [109]) and whose diagrams can be seen in Figure 2.2. These decay processes have a D^0 meson as an intermediate resonance particle, and are therefore

Channel	PID Efficiency / misID Rate
2011 $B_c^\pm \to K^\pm K^\mp \pi^\pm$	53%
2012 $B_c^\pm \to K^\pm K^\mp \pi^\pm$	52%
2011 $B_c^\pm \to K^\pm K^\mp K^\pm$	12%
2012 $B_c^\pm \to K^\pm K^\mp K^\pm$	11%
2011 $B_c^\pm \to K^\pm \pi^\mp \pi^\pm$	1.8%
2012 $B_c^\pm \to K^\pm \pi^\mp \pi^\pm$	2%
2011 $B_c^\pm \to \pi^\pm \pi^\mp \pi^\pm$	0.07%
2012 $B_c^\pm \to \pi^\pm \pi^\mp \pi^\pm$	0.09%

Table 4.9: Approximate PID efficiencies for simulation samples and cross-feed simulation samples, after chosen PID cuts. The first two rows show the efficiency for correct identification, all the others show the % of cross-feed channels that pass the PID cuts.

not annihilation processes. However, their final states are similar to that of the $B_c^\pm \to K^\pm K^\mp \pi^\pm$, and they may contribute part of the signal in the $K^\pm \pi^\mp \pi^\mp$ and $K^\pm K^\mp \pi^\pm$ channels, as well as the $K^\pm \pi^\mp \pi^\mp \to K^\pm K^\mp \pi^\pm$ cross-feed[11]. It is worthwhile mentioning, however, that the misID rate (see Table 4.9 shows that the $K^\pm \pi^\mp \pi^\mp$ is suppressed in this case. In particular, the $B_c^\pm \to D^0(K^\pm \pi^\mp) K^\pm$ is also relevant to this analysis, as it is a possible *control channel* (see Section 4.5).

4.2.7 $B_c^\pm \to K^\pm K^\mp \pi^\pm \pi^0$ Partially Reconstructed Background

The final decay processes that can contribute to the $m_{K^+K^-\pi^+}$ mass window are 4-body decays of the form $hhh\pi^0$, where h represents kaons and pions, and the π^0 particle is "lost" during the reconstruction process. For this reason, these contributions are called *partially reconstructed* backgrounds, and they appear in the mass window due to their reconstructed final state being $KK\pi$. The possible contributions from four of these channels have been

[11] Where the first pion is mis-identified as a kaon.

studied, although only two are expected to have any significant contributions. As was done with the cross-feed channels, all the distributions of the partially reconstructed backgrounds have been weighted with their mis-identification rates.

The first of these is the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}\pi^0$ channel, whose expected peaking background is located in the lower side-band of the $m_{K^+K^-\pi^+}$ mass window. A fit is performed for this channel using an Argus shape (created by the ARGUS Collaboration, see [110]) convolved with a Gaussian function, and the results can be seen in Figure 4.15.

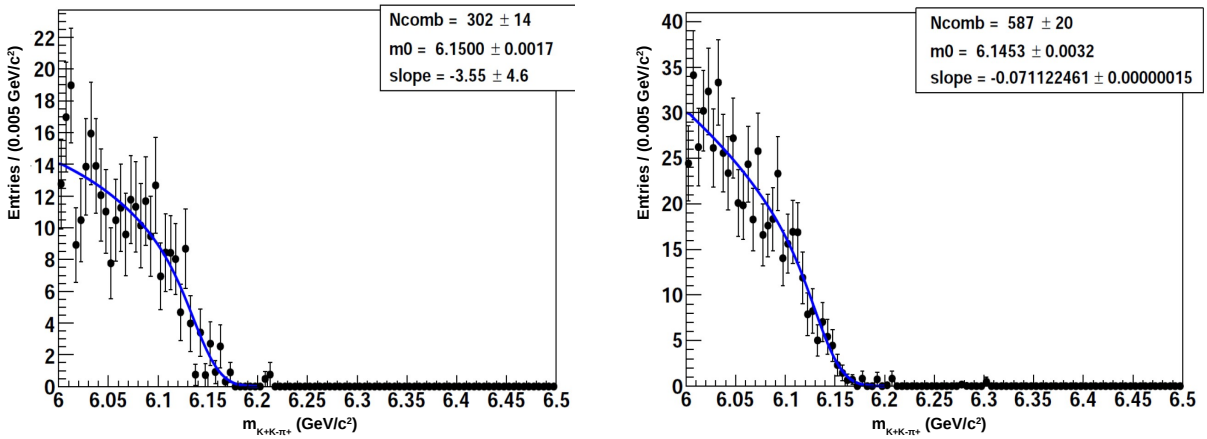

Figure 4.15: Argus⊗Gaussian invariant mass fits to $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}\pi^0$ simulated events, for (left) 2011 and (right) 2012 samples.

4.2.8 $B_c^{\pm} \to K^{\pm}K^{\mp}K^{\pm}\pi^0$ Partially Reconstructed Background

The $B_c^{\pm} \to K^{\pm}K^{\mp}K^{\pm}\pi^0$ channel is the second partially reconstructed channel which is expected to contribute a peaking background in the lower side-band of the $m_{K^+K^-\pi^+}$ mass window. As its shape is similar to that of the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}\pi^0$ channel, this channel was fit with the same model and methodology as in subsection 4.2.7, and the results can be seen in Figure 4.16. However, during blinded fit tests of the entire $m_{K^+K^-\pi^+}$ mass window, it was seen that the addition of this channel did not impact the fit quality. As the expected yield for this channel is smaller than that of the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}\pi^0$ channel and its peaking background's location is in the same region of the lower side-band, this

channel's contribution is absorbed by both the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}\pi^0$ and the combinatorial background. Therefore, this channel is disregarded in the final fit.

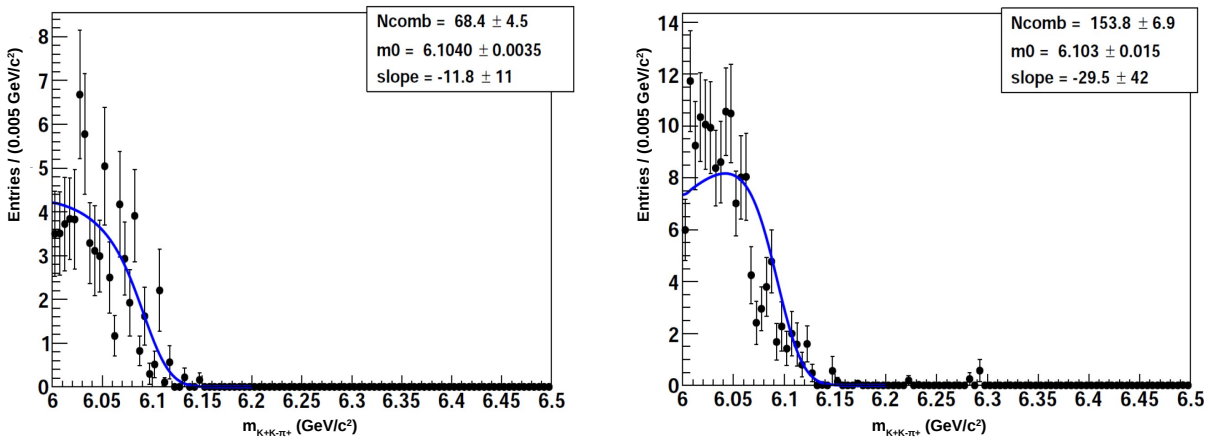

Figure 4.16: Argus⊗Gaussian invariant mass fits to $B_c^{\pm} \to K^{\pm}K^{\mp}K^{\pm}\pi^0$ simulated events, for (left) 2011 and (right) 2012 samples.

4.2.9 $B_c^{\pm} \to K^{\pm}\pi^{\mp}K^{\pm}\pi^0$ and $B_c^{\pm} \to \pi^{\pm}\pi^{\mp}K^{\pm}\pi^0$ Partially Reconstructed Backgrounds

The $B_c^{\pm} \to K^{\pm}\pi^{\mp}K^{\pm}\pi^0$ and $B_c^{\pm} \to \pi^{\pm}\pi^{\mp}K^{\pm}\pi^0$ channels are not expected to have any meaningful contribution to the $m_{K^+K^-\pi^+}$ mass window, as their yields are negligible after the selection process. Furthermore, their distributions, which can be seen in Figures 4.17 and 4.18, are broad and extend across the $m_{K^+K^-\pi^+}$ window. For these reasons, these two channels end up being absorbed by either the combinatorial background or the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}\pi^0$ component, and therefore no fit is performed for them, and they are disregarded in the final mass fit.

4.3 Expectations and Fit Strategy

4.3.1 Expectations

For annihilation decay processes that occur through $\bar{b}c \to W$, it is difficult to predict an expected yield due to the imprecise, if not unavailable, theoretical predictions and

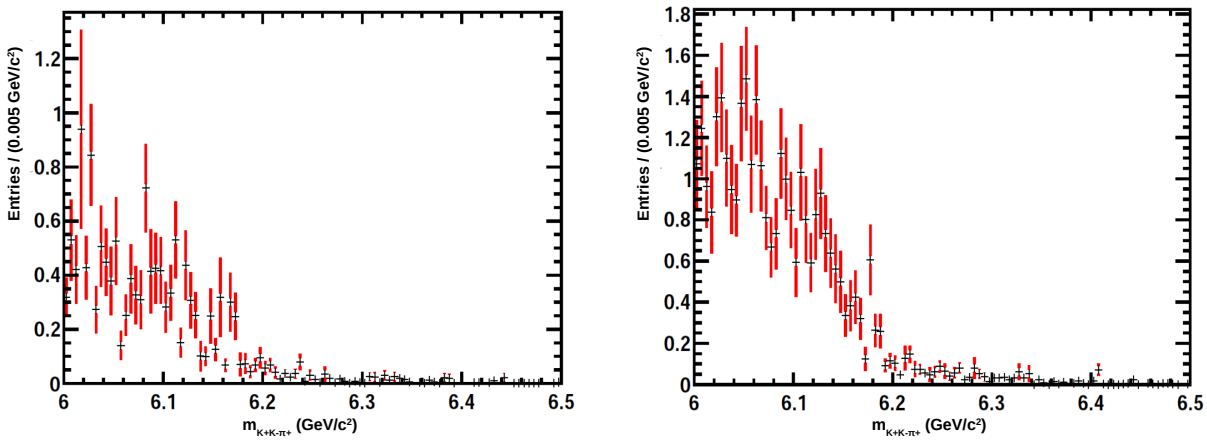

Figure 4.17: $B_c^{\pm} \to K^{\pm}\pi^{\mp}\pi^{\pm}\pi^0$ mass distributions, for 2011 (left) and 2012 (right) simulated events.

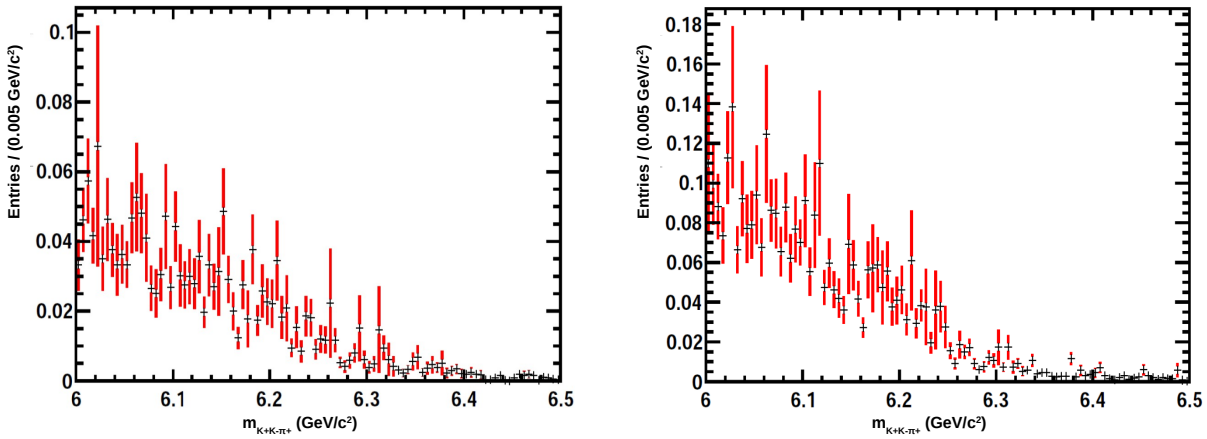

Figure 4.18: $B_c^{\pm} \to \pi^{\pm}\pi^{\mp}\pi^{\pm}\pi^0$ mass distributions, for 2011 (left) and 2012 (right) simulated events.

knowledge on the b-quark's fragmentation fraction into B_c^+. Any predictions are at this point little more than educated guesswork, but the order of magnitude of the number of expected events can be inferred from the LHCb analysis on the $B_c^{\pm} \to p\bar{p}\pi^{\pm}$ decay [111], which shares numerous similarities with the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$.

In particular the $B_c \to J/\psi\pi$ decay is of interest. Its branching fraction is $\approx 10^{-3}$, while the $J/\psi \to p\bar{p}$ process' is $\approx 2 \times 10^{-3}$. The combined order of magnitude for the $B_c \to J/\psi(p\bar{p})\pi$ decay's branching fraction is thus $\approx 10^{-6}$. For the $J/\psi \to KK$ process,

the rate is even smaller: 2.37×10^{-4}, giving an order of magnitude of $\approx 10^{-7}$ to the rate of the $B_c \to J/\psi(KK)\pi$ decay. This rate is compatible with predictions (using KK^* decays) of the branching fractions of mesonic channel annihilation diagrams, which can be anywhere between 10^{-8} and 10^{-6}.

However, the $B_c \to J/\psi(p\bar{p})\pi$ decay was *not* observed using the LHC's Run I data. This makes an observation of the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ annihilation decay very unlikely, as its branching fraction is expected to be similar to that of the $B_c \to J/\psi(KK)\pi$ process, which is itself already an order of magnitude *smaller* than that of the $B_c \to J/\psi(p\bar{p})\pi$ decay. Therefore for the statistics produced in Run I, no more than $O(10)$ events with $K^{\pm}K^{\mp}\pi^{\pm}$ final states are expected, coming mostly from the $B_c^{\pm} \to D^0(K^{\pm}\pi^{\mp})K^{\pm}$ decays.

4.3.2 BDT Divisions and the Simultaneous Fit

The optimal selection configuration described in subsection 4.1.5 includes a BDT cut > 0.18, which represents nearly 30% of the BDT shape area (or BDT cut efficiency). This type of cut is referred to as a "tight" selection scenario, which is defined through its maximization of the Punzi criterion. However, there is an alternative way to extract a signal, in which a *simultaneous* fit is performed in adjacent BDT bins, who all have approximately the same area (and therefore approximately the same signal yield). The simultaneous fit therefore requires the definition of component yields and shapes for each BDT bin. In order to deal with the proportionally large amount of background events, the entire region BDT > 0.04, which represents approximately 80% of the total BDT area, is chosen for this exercise[12]. The three bins with similar areas to be used in the simultaneous fit are defined as: $0.04 <$ BDT < 0.12, $0.12 <$ BDT < 0.18, and $0.18 <$ BDT.

[12]The *entire* BDT area is not used because below the 0.04 threshold, there is no sensitivity for signal extraction.

4.3.3 Toy Monte Carlo and the Choice of Fit Strategy

The tight cut and simultaneous fit scenarios are two possible strategies for signal extraction, and in order to evaluate which is the most statistically sensitive method for this case, a *Toy Monte Carlo* (Toy MC) study is performed. This is done through numerous (in this case, 1000) simulated experiments, where the number of signal and background events are pre-defined, and which are done with and without the addition of a signal component. After the simulated experiments, distributions are generated for component yields, pulls, and most importantly, $-log(\mathcal{L})$ for signal sensitivity, for both of these configurations. The mean for each likelihood distribution is obtained through a fit using a Gaussian function, and for each test the significance of the signal component is obtained through:

$$\sigma = \sqrt{-2 \times log(\frac{\mathcal{L}_{bkg}}{\mathcal{L}_{sig+bkg}})} \tag{4.5}$$

Where $\mathcal{L}_{sig}$ and $\mathcal{L}_{bkg}$ are the means of the likelihood distributions for the simulations with and without the addition of the signal component, respectively. In the Toy MC exercise done for this analysis, as the objective is to simply compare the significance of each method, only the shapes of the signal and combinatorial background components are used, and they are taken from the fits shown in Section 4.2.

For the number of signal events used in the simulations, three signal yield cases are considered inside the BDT > 0.04 region, following the discussion in the previous subsection. These yields represent respectively the realistic, cautiously optimistic and unrealistically optimistic scenarios: $N_S = 10$, 25, and 50 events. These yields represent the total number of events in the entire BDT > 0.04 region, so for the simultaneous fit each BDT bin has 1/3 of the total yield. The tight fit scenario, on the other hand, only has a signal yield of 1/3 of the total number of events, as it is akin to only fitting the BDT > 0.18 bin.

The number of background events is estimated in a similar fashion as in subsec-

tion 4.1.5, i.e. through a fit to these side-bands. In this case, the same exponential function used to model the combinatorial background is used to fit both side-bands [13], which is then normalized and integrated in both the blinded signal region and the entire $m_{K^+K^-\pi^+}$ mass window to obtain the estimated yields. The exponential fit extrapolations to the side-bands for data samples in which the signal region is blinded can be seen in Figure 4.19, while the background yields are described in Table 4.10. The Gaussian-fitted distributions of yields and pulls for signal events can be seen in Section A.3, for both fit methods.

The results of the Toy MC significance tests are shown in Table 4.11, which shows that the simultaneous fit method offers similar or better statistical significances than the tight cut scenario, especially in the most realistic case of 10 signal events. Therefore, the chosen fit strategy for this analysis is the simultaneous fit to adjacent BDT bins.

BDT bin	Year 2011	Year 2012	All
$[0.04, 0.12]$ Signal Region	1547.7	4738.59	6286.5
Full fit range	3947.6	12068.60	16016.7
$[0.12, 0.18]$ Signal Region	135.3	553.17	694.7
Full fit range	373.6	1416.28	1799.2
> 0.18 Signal Region	35.6	86.27	122.2
Full fit range	97.2	226.71	324.3

Table 4.10: Estimated background yields.

4.4 Blind Fits

Following the determination of the most statistically sensitive fit strategy, the next step is a series of blind fit tests to the full data samples, using the component models defined in subsection 4.2, and in which the signal region is blinded during the fit. This is an

[13] Previously, for simplicity only the upper side-band was fitted with a linear function.

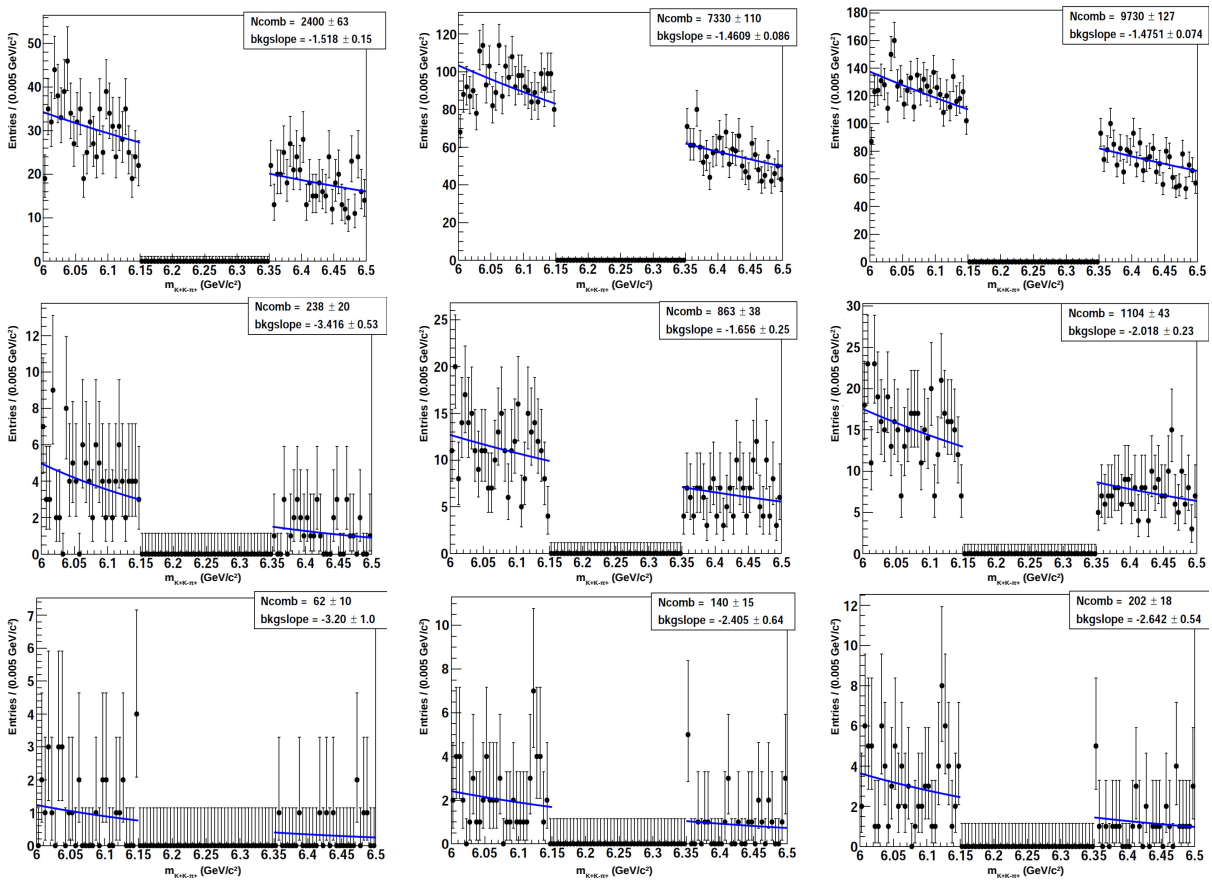

Figure 4.19: Exponential mass fits to (left) 2011, (center) 2012, and (right) combined side-band data samples for (top) BDT $\in [0.04, 0.12]$, (middle) BDT $\in [0.12, 0.18]$, and (bottom) BDT > 0.18.

important point to clarify, as to this point the data samples used all had the events in the signal region *excluded*. Now the *complete* data sets are used, which is necessary for the fit tests, as the fit must now take the signal component into account. If data samples without any events in the signal region are used, the fitting process will not have any events *to fit* in this region, and any produced results will not be sensible. However, the signal region is still blinded to the researchers, who have no access to any quantitative or visual information from the fit. Therefore, even though the events in the signal region are taken into account during the fit, they are not shown in the produced graphs of the results.

Expected overall signal yield →	10	25	50
2011	1.068	1.273	1.581
2012	1.342	1.334	1.631
All	1.136	1.572	2.519
2011	1.149	1.249	1.661
2012	1.543	1.249	1.691
All	1.200	1.549	2.546

Table 4.11: Expected average statistical significances (in σ) against fit method and expected signal yields. Tight BDT cut and Simultaneous fit methods are shown in the second and third lines, respectively.

This section thus presents the blind fit tests performed in the $[6.0, 6.5]$ GeV/c^2 mass range, for three spectrum configurations, two of which have cuts performed on the $m_{K^+K^-}$ and $m_{K^-\pi^+}$ invariant mass: full spectrum (no cuts); B regions ($m_{K^+K^-}$ and $m_{K^-\pi^+}$ < 5 GeV/c^2) excluded; and with both B and D^0 regions excluded (the D^0 region being $m_{K^-\pi^+}$ values within the $[1.834, 1.894]$ GeV/c^2 mass window). The invariant mass cuts are used to eliminate the possibility of events coming from the decays of B^0, $B^\pm$, and D^0 particles.

The blind fit tests of the three different spectrum configurations are all performed using the same methodology, starting with only the signal and combinatorial background components, and including the relevant cross-feed and partially reconstructed components (Section 4.2) in succession. All fits use the combined 2011 and 2012 data samples, and are performed over the full mass range, using the simultaneous fit method. Furthermore, all the fits have fixed signal, cross-feed, and partially reconstructed shapes, allowing only the background slope parameter of the combinatorial component and the event *yields* to fluctuate. However, as the yields of the cross-feed components are related to the signal yield [14], they are treated by Gaussian constraints based on the misID rates. The cross-

[14] As the cross-feeds result from B_c meson decays, the yields for these components only make sense if they are smaller than the signal yield itself.

feed yields are allowed to fluctuate within a range of values which is represented by a Gaussian function, whose mean and σ depend on the product of the signal yield and the ratio of the misID rate (for that specific cross-feed channel) and the PID efficiency.

Tables 4.12 and 4.13 summarize all the tests performed, showing the values of $-log\mathcal{L}$ and χ^2 per degree of freedom[15] in all the blind fits. In both tables, (a) represents signal and combinatorial background components, (b) represents the $K^{\pm}K^{\mp}\pi^{\pm}\pi^0$ partially reconstructed component, and (c) and (d) represent the $K^{\pm}K^{\mp}K^{\pm}$ and $K^{\pm}\pi^{\mp}\pi^{\pm}$ cross-feed components, respectively. This exercise has the objective of showing the relevance of each fit component, depending on the cuts and vetoes in the B region and D^0 band.

It is also important to mention that the values of $-log\mathcal{L}$ and χ^2 per degree of freedom can only be compared between different component combinations in the *same spectrum configuration*. Different spectrum configurations have different numbers of events (due to the invariant mass cuts), and therefore it makes no sense to compare their values of $-log\mathcal{L}$ and χ^2. Only within a specific spectrum configuration does a comparison of these quantities have any physical significance. Figures 4.20 and 4.21 show the blind fit tests for the full spectrum configuration. The blind fit tests for the other configurations are very similar (visually speaking) and are not included, as their truly relevant information is contained in the aforementioned tables. From the tables it can be seen that all components contribute positively to the values of minimum $-log\mathcal{L}$, and that the fit itself is of acceptable quality. In both figures, the components are represented as per the following:

- **Solid red line** - The full, combined PDF model;
- **Dashed blue line** - Combinatorial background;
- **Dashed magenta line** - $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}\pi^0$ partially reconstructed component;
- **Dashed green line** - $B_c^{\pm} \to K^{\pm}K^{\mp}K^{\pm}$ cross-feed component

[15]Both of these quantities are allowed to be printed in the blind fits, as they are related to fit quality and give no information regarding signal yield or significance.

- **Dashed yellow line** - $B_c^{\pm} \to K^{\pm}\pi^{\mp}\pi^{\pm}$ cross-feed component

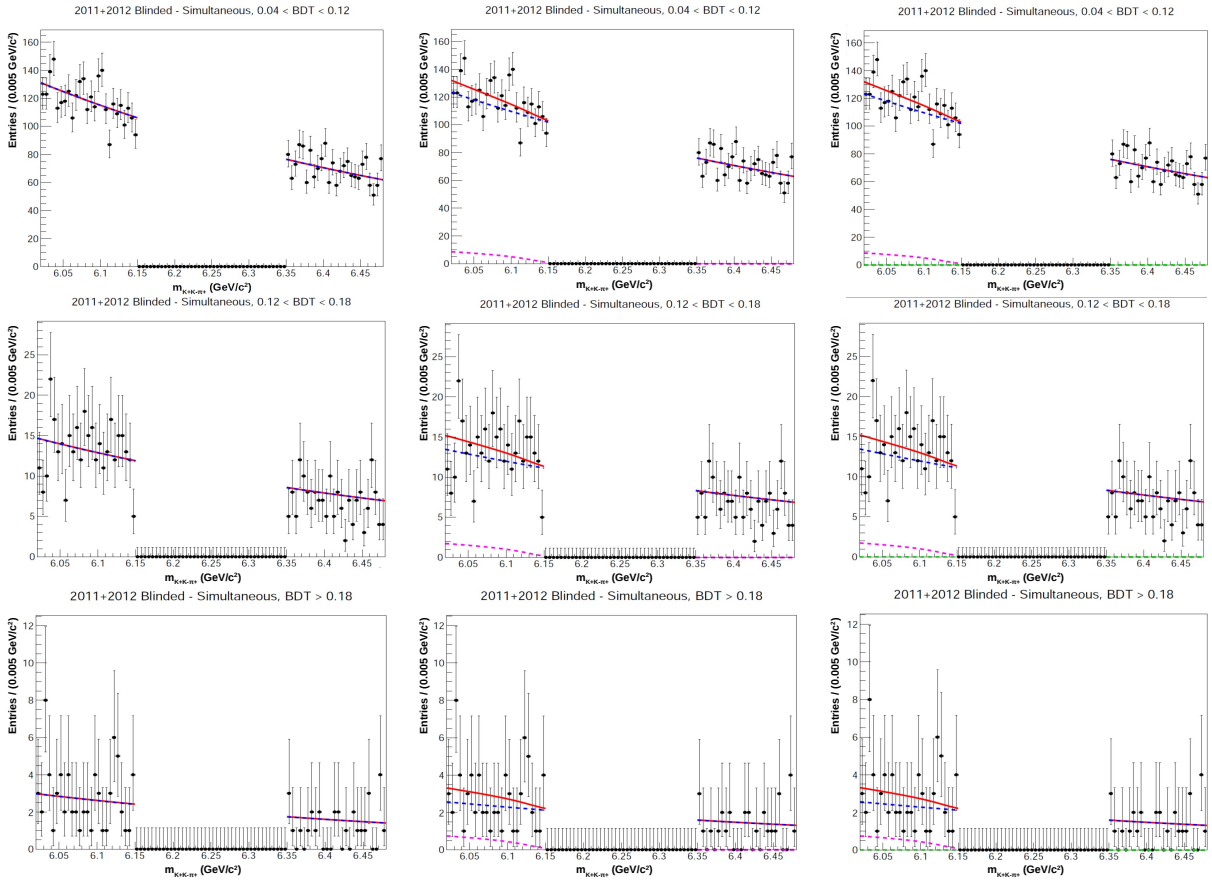

Figure 4.20: Full mass blind fit tests performed using the full spectrum and combined 2011+2012 data, using signal and background components only (left); signal, background and partially reconstructed components (middle); signal, background, partially reconstructed and $K^{\pm}K^{\mp}K^{\pm}$ cross-feed components (right); for BDT bins (top) $[0.04, 0.12]$, (middle) $[0.12, 0.18]$ and (bottom) $> [0.18]$.

4.5 Control Channel Fits

4.5.1 Fit to $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ Signal

With the interest of normalizing any B_c signal to a *known* measured quantity, the $B^+ \to K^+K^-\pi^+$ decay is the immediate first candidate as a control channel, as it possesses the same final state, similar topology, and was the subject of a dedicated LHCb study (see

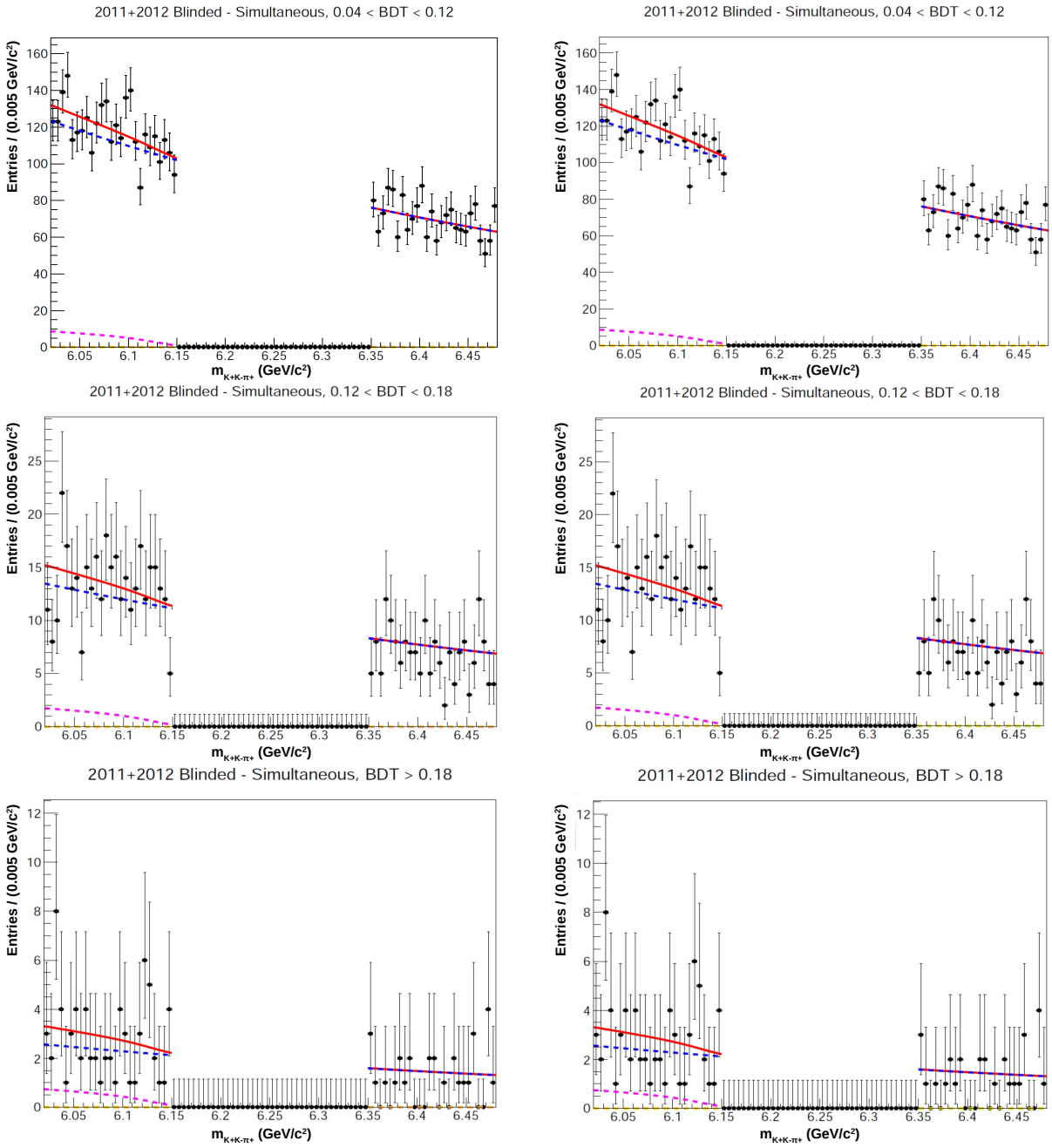

Figure 4.21: Full mass blind fit tests performed using the full spectrum and combined 2011+2012 data, using signal, background, partially reconstructed and $K^{\pm}\pi^{\mp}\pi^{\pm}$ cross-feed components (left); signal, background, partially reconstructed, $K^{\pm}K^{\mp}K^{\pm}$ and $K^{\pm}\pi^{\mp}\pi^{\pm}$ cross-feed components (right); for BDT bins (top) $[0.04, 0.12]$, (middle) $[0.12, 0.18]$ and (bottom) $>[0.18]$.

Components	No Dalitz Cuts	mKK < 5 GeV/c^2	mKπ < 5 GeV/c^2	mKπ and mKK < 5 GeV/c^2	mKπ and mKK < 5 GeV/c^2 and D0 Veto
(a)	-71784.2	-51450.2	-64161.8	-44108.8	-43645.7
(a)+(b)	-71785.9	-51452.3	-64163.2	-44110.4	-43647.3
(a)+(b)+(c)	-71790.2	-51456.6	-64167.8	-44114.7	-43651.9
(a)+(b)+(d)	-71791	-51457.4	-64168.3	-44115.5	-43652.4
All components	-71795.3	-51461.7	-64172.6	-44119.8	-43656.7

Table 4.12: Values of minimum NLL returned by the full spectrum blind fit tests, for various Dalitz mass cuts (columns) and components used in the fit (lines). Here (a) represents signal and combinatorial background components, (b) represents the $K^{\pm}K^{\mp}\pi^{\pm}\pi^0$ partially reconstructed component, (c) and (d) represent the $K^{\pm}K^{\mp}K^{\pm}$ and $K^{\pm}\pi^{\mp}\pi^{\pm}$ cross-feed components, respectively.

Components	No Dalitz Cuts	mKK < 5 GeV/c^2	mKπ < 5 GeV/c^2	mKπ and mKK < 5 GeV/c^2	mKπ and mKK < 5 GeV/c^2 and D0 Veto
(a)	1.01665	1.01461	0.945736	0.977125	0.949031
(a)+(b)	0.995286	0.978098	0.927391	0.948254	0.949031
(a)+(b)+(c)	0.997417	0.979635	0.929261	0.949592	0.921076
(a)+(b)+(d)	0.990329	0.977979	0.927243	0.948109	0.91969
All components	0.997269	0.979469	0.929107	0.949436	0.920923

Table 4.13: Values of χ^2 per degree of freedom returned by the full spectrum blind fit tests, for various Dalitz mass cuts (columns) and components used in the fit (lines). Here (a) represents signal and combinatorial background components, (b) represents the $K^{\pm}K^{\mp}\pi^{\pm}\pi^0$ partially reconstructed component, (c) and (d) represent the $K^{\pm}K^{\mp}K^{\pm}$ and $K^{\pm}\pi^{\mp}\pi^{\pm}$ cross-feed components, respectively.

[112]). The B^+ invariant mass window was included in the stripping line used for this analysis (see subsection 4.1.2) for this reason, and the B^+ signal is estimated from a fit to this mass window using almost the same configuration as in the aforementioned study. The exception is the signal function, which in our case is a Double Crystal Ball instead of a Cruijff function. The B^+ signal is defined over the spectrum, with the exclusion of the D^0 band ($[1.834, 1.894]$ GeV/c^2) in both the $m_{K^+K^-}$ and $m_{K^-\pi^+}$ invariant mass distributions. With the phase space defined this way, the branching fraction for this process is known [50]: $\mathcal{B}(B^\pm \to K^\pm K^\mp \pi^\pm) = (5.0 \pm 0.7) \times 10^{-6}$.

For the selection, the same BDT (variable and) cut > 0.04 as for the B_c^+ region is applied. The chosen PID is the same as is applied in the dedicated study [112]: `ProbNNk(kaon)` > 0.45, `ProbNNpi(kaon)` < 00.5, `ProbNNpi(pion)` > 0.5, and `ProbNNpi(kaon)` < 0.05. For the fit, the following components are used: a Double Crystal Ball function for the signal in which the mean and σ of the first CB are allowed to float; an exponential function for the combinatorial background; two partially reconstructed $K^\pm K^\mp \pi^\pm \pi^0$ backgrounds (decaying from $B^\pm$ and $B_S^\pm$ mesons), both fit with Cruijff functions [113]; and a $B^\pm \to K^\pm \pi^\mp \pi^\pm$ cross-feed component, which is also fit with a Cruijff function. Originally there was also a $B^\pm \to K^\pm K^\mp K^\pm$ cross-feed component, which in this case was found to have a negligible impact on the fit due to efficient PID rejection, and was thus not considered.

The results of the fit are shown in Figures 4.22 and 4.23 (log scale and pulls), while Table 4.14 lists the yields of all components used. A separate fit was performed on the Monte Carlo simulated $B^\pm$ signal, and Table 4.15 compares the resolution parameters from the full fit and the MC signal. For the full fit, the cross-feed component yields are treated in similar fashion as in 4.4, i.e. Gaussian constraints based on misID rates. For this control mode, further details regarding the signal and background modelling performed for the $B^\pm$ fit can be seen in Section A.4, and the acceptances for the $B^\pm \to K^\pm K^\mp \pi^\pm$ simulation samples can be seen in Section A.5.

A further consistency check is performed using signal sWeights generated from the fit via the sPlot tool [114]. Using information from the full fit, covariance-weighted quantities (called sWeights) are generated for every event in such a way that their sum is equal to the total number of events of the type that the weights are generated for - in this case, the signal, and as it is obtained from the fit, this yield already takes into account the BDT and PID selection. Distributions can thus be taken from the data and weighted with the sWeights in order to obtain the shape of the distribution for *only* the signal component. With the sWeights, a data/simulation comparison has been performed for the $B^{\pm}$ P, IP, and IP χ^2 variables, as shown in Figure 4.24. The variable distributions taken from the simulation samples also undergo the full selection, with the PID effect simulated in this case by the same `PIDCalib` method used in subsection 4.1.5. The shapes are in very good agreement, indicating that the $B^{\pm}$ model is consistent.

Component	Yield
$B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ signal	5182.4 ± 226.8
Combinatorial background	16242.6 ± 1150.1
$B^{\pm} \to K^{\pm}K^{\mp}K^{\pm}$ cross-feed	159.6 ± 18.6
$B^{\pm} \to K^{\pm}\pi^{\mp}\pi^{\pm}$ cross-feed	347.6 ± 36.1
$B_s \to$ 4body partially reconstructed	5174.6 ± 788.8
$B \to$ 4body partially reconstructed	150.3 ± 83.8

Table 4.14: Component yields for the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ fit.

4.5.2 Fit to $B^{\pm} \to D^0(K^{\pm}K^{\mp})\pi^{\pm}$

The fit performed to the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ shown in the previous subsection was performed excluding any charm contributions coming from the D^0 bands, through cuts in the $m_{K^{\pm}K^{\mp}}$ and $m_{K^{\pm}\pi^{\mp}}$ invariant mass distributions. Alternatively, the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ signals

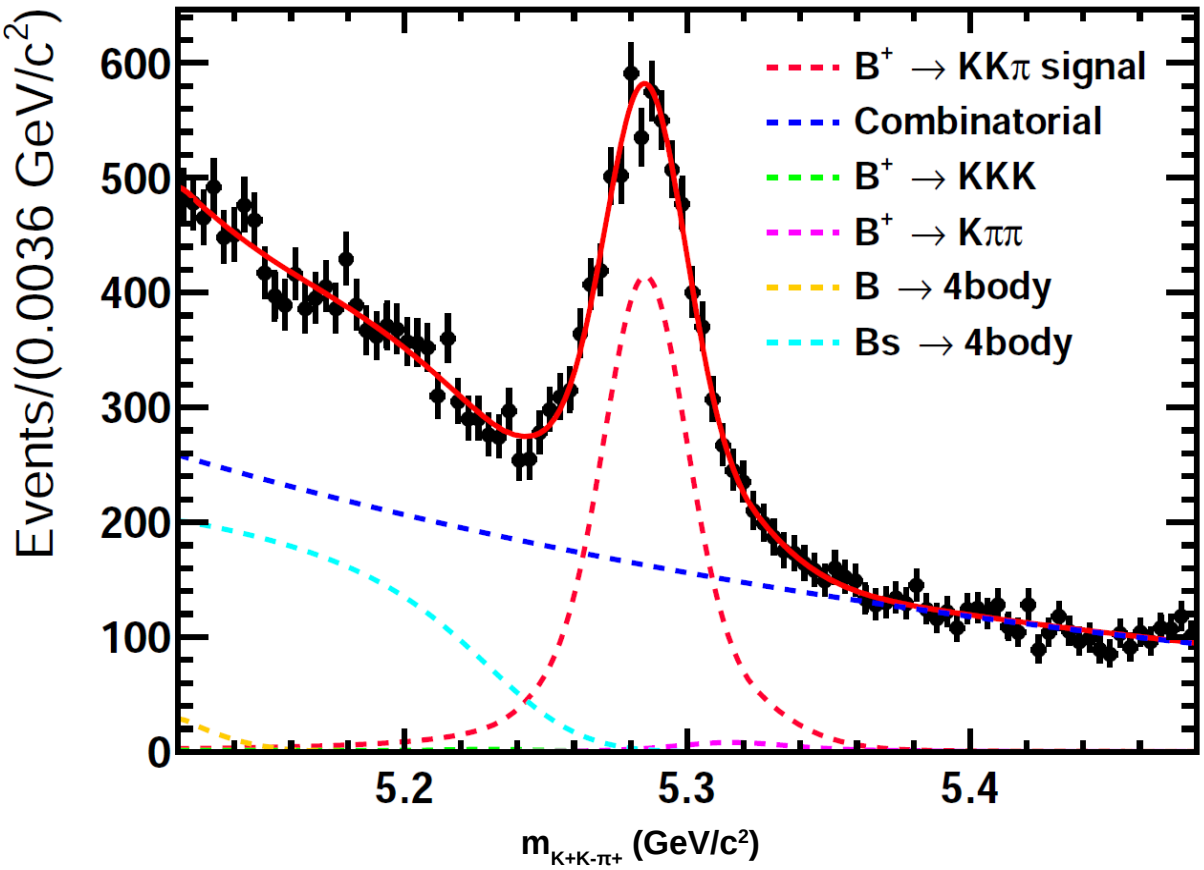

Figure 4.22: Full mass fit to the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ signal using combined 2011+2012 data.

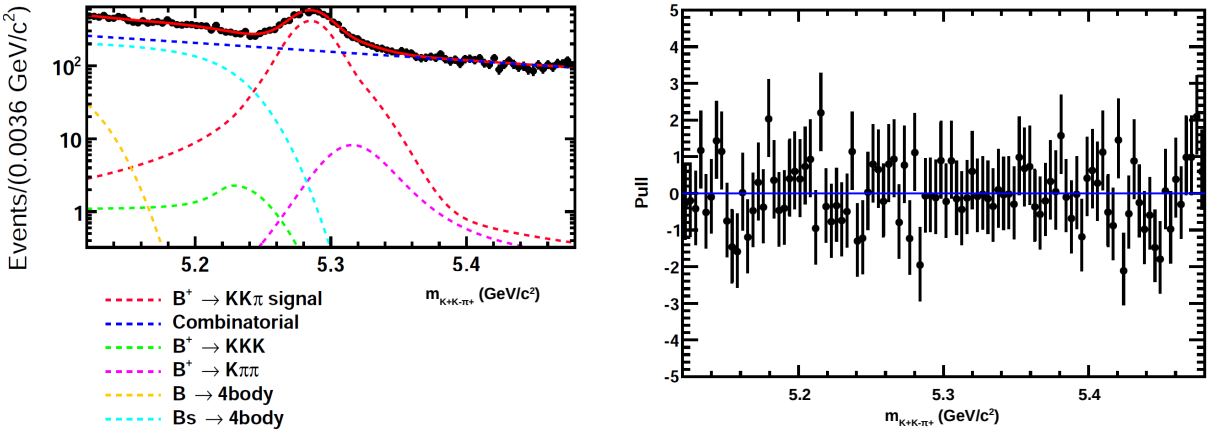

Figure 4.23: (left) Log scale for the full mass fit to the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ signal using combined 2011+2012 data, and (right) pull distribution for the fit.

Parameter	Fitted	MC Expectation
μ_{CB}	5.28365 ± 0.00014	5.280240 ± 0.000082
σ_{CB}	0.0213339 ± 0.0003617	0.02224 ± 0.00057

Table 4.15: Fit parameters for the $B^{\pm}$ signal component, comparing values coming from the full mass fit (second column) and the MC simulation sample fit (third column).

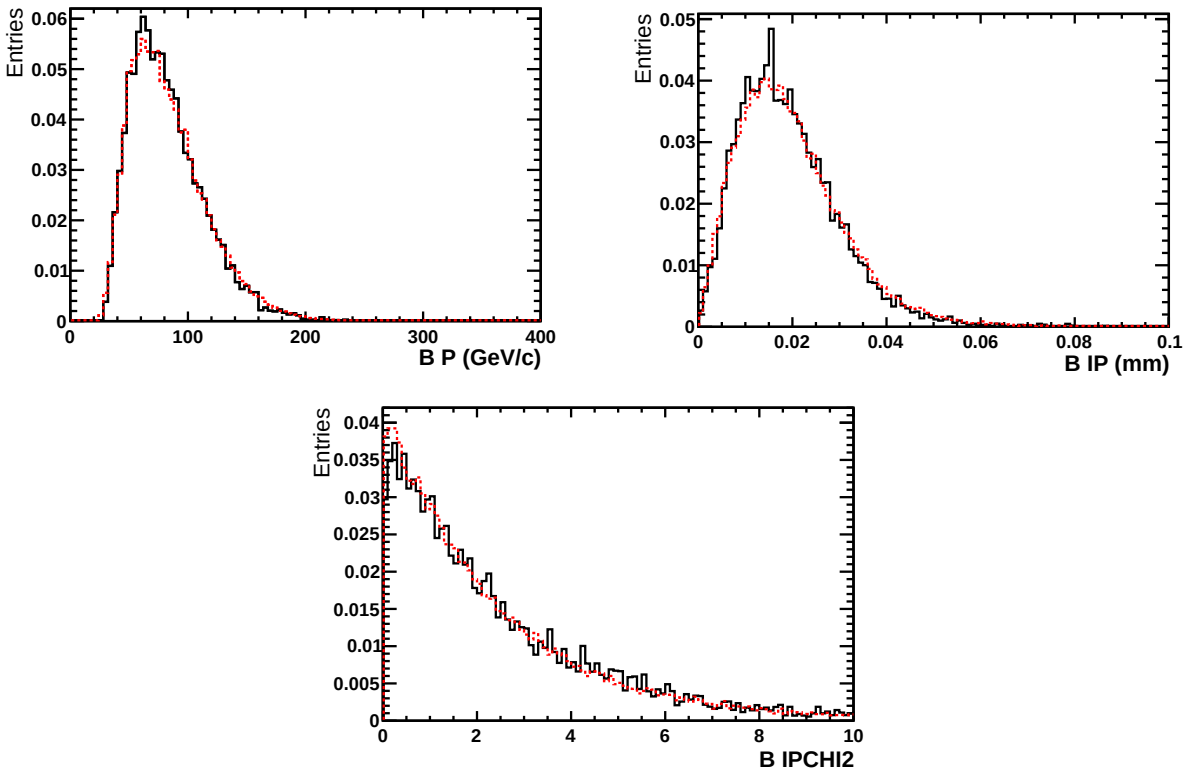

Figure 4.24: $B^{\pm}$ P, IP and IP χ^2 distributions for (solid black line) background subtracted data and (dashed red line) simulation.

coming from the D^0 bands themselves could be used through the following decays:

- $B^{\pm} \to D^0(K^{\pm}\pi^{\mp})K^{\pm}$, with branching fractions $\mathcal{B}(B^{\pm} \to D^0 K^{\pm}) = (3.70 \pm 0.17) \times 10^{-4}$ and $\mathcal{B}(D^0 \to K^{-}\pi^{+}) = (3.93 \pm 0.04)$
- $B^{\pm} \to D^0(K^{\pm}K^{\mp})\pi^{\pm}$, with branching fractions $\mathcal{B}(B^{\pm} \to D^0\pi^{\pm}) = (4.81 \pm 0.15) \times$

10^{-3} and $\mathcal{B}(D^0 \to K^- K^+) = (4.01 \pm 0.07)$ [50]

In particular, it is found that for the $B^\pm \to D^0(K^\pm K^\mp)\pi^\pm$ signal, besides the fact that its branching fraction (1.93×10^{-5}) is significantly higher than for $B^\pm \to K^\pm K^\mp \pi^\pm$ ($(5.0 \pm 0.7) \times 10^{-6}$), it also offers a much cleaner signal (i.e. much higher purity). For these reasons, this channel is considered to be a possible (an possibly even preferable) alternative normalization channel for the $B_c^\pm \to D^0(K^\pm K^\mp)\pi^\pm$ decay.

A fit is therefore performed after selecting the D^0 band in the $m_{K^\pm K^\mp}$ spectrum, and the results (with and without log scales) and pulls can be seen in Figures 4.25 and 4.26. The component yields can be seen in Table 4.16, in which the B_s partially reconstructed channel is not included, as its contribution is largely reduced in the D^0 band. All other components used in the previous $B^\pm$ are present.

Component	Yield
$B^\pm \to K^\pm K^\mp \pi^\pm$ signal	8577.01 ± 109.44
Combinatorial background	1200.55 ± 72.76
$B^\pm \to K^\pm K^\mp K^\pm$ cross-feed	32.79 ± 8.55
$B^\pm \to K^\pm \pi^\mp \pi^\pm$ cross-feed	117.00 ± 43.10
$B \to$ 4body partially reconstructed	1293.76 ± 51.10

Table 4.16: Component yields for the D^0 band in the $B^\pm \to K^\pm K^\mp \pi^\pm$ fit.

4.6 Acceptances

The acceptances for the $B_c^\pm \to K^\pm K^\mp \pi^\pm$ simulation samples are obtained using a similar methodology as in [115]. The two-dimensional acceptance histograms are generated by dividing histograms filled with events that pass each "step" by the histogram filled with the events of the *previous* step. This division is done bin-by-bin, and the result is a

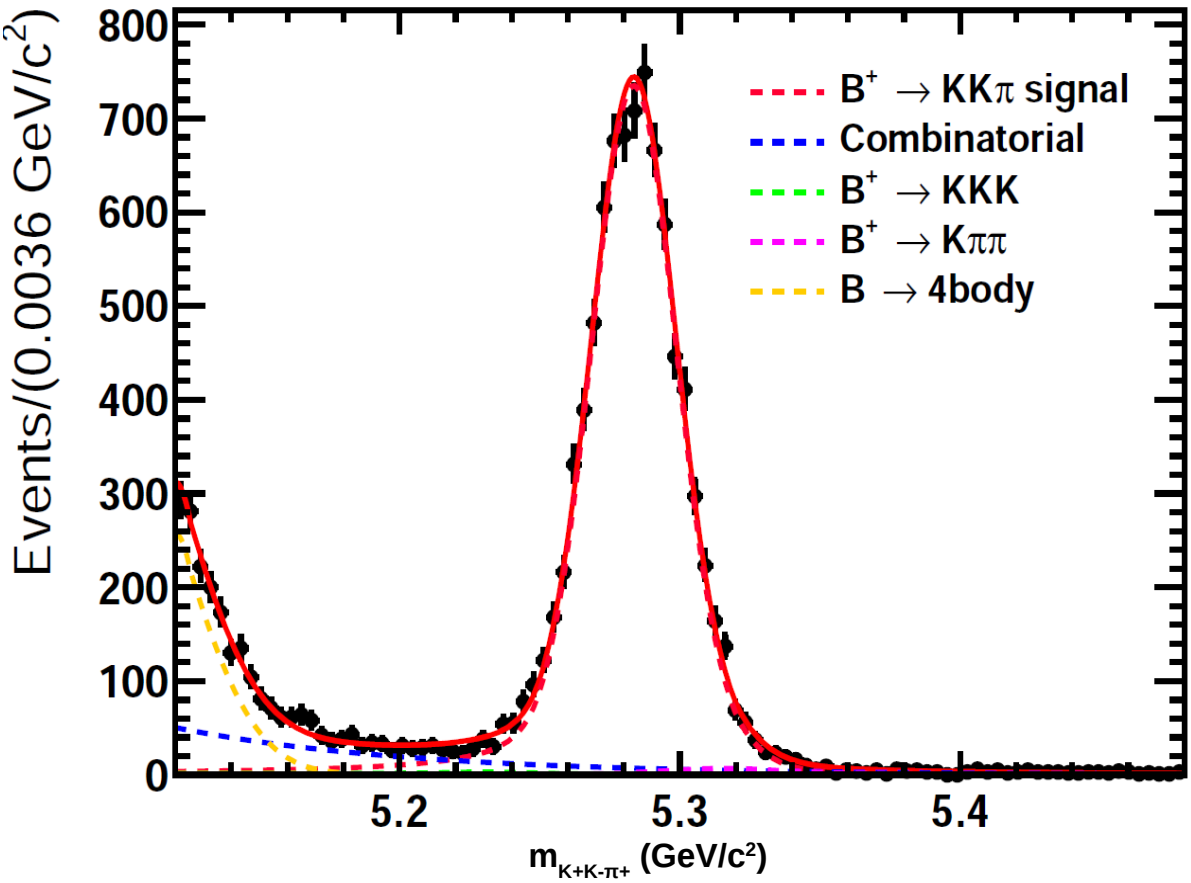

Figure 4.25: Full mass fit to the D^0 band of the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ signal using combined 2011+2012 data.

representation of the efficiency of each step. The final acceptance histograms were then smoothed over by fitting with a two-dimensional fourth-order polynomial:

$$P(x,y) = \sum_{i,j=0}^{i,j=4} a_{ij} x^i y^j \tag{4.6}$$

which represent in total 15 parameters to fit. Here $a_{i,j}$ represents the constant of proportionality of each $x^i y^j$ coordinate pair, while x and y are coordinates in the m^2_{KK} and $m^2_{K\pi}$ invariant mass distributions. This is done in order to both generate histograms without un-physical border fluctuations and to obtain fit functions for the acceptances,

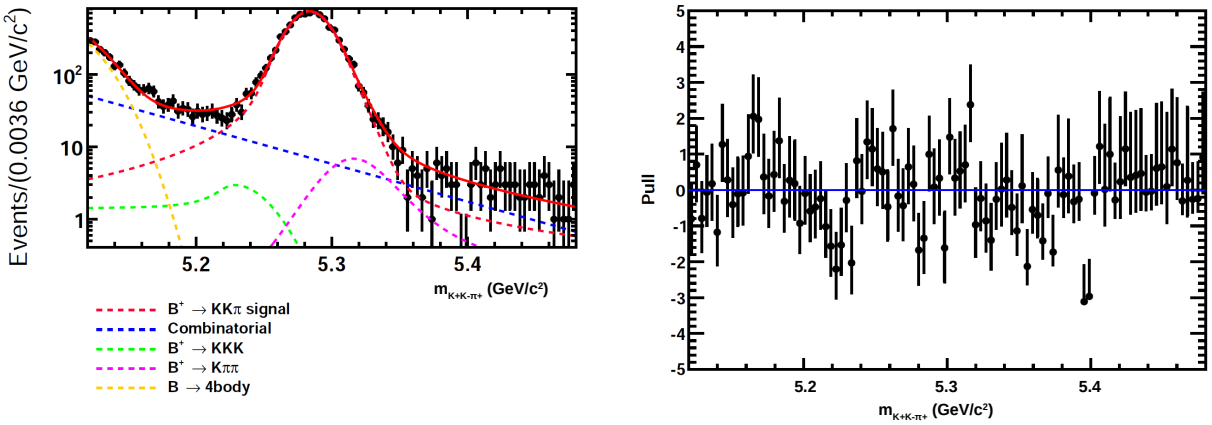

Figure 4.26: (left) Log scale for fit to the D^0 band of the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ signal using combined 2011+2012 data, and (right) pull distribution for the fit.

which allow for the precise determination of efficiencies at any point in the histogram distributions. The effect of triggers, BDT, and PID are studied separately.

4.6.1 Preselection

The preselection acceptance includes the reconstruction from generated events and the stripping requirements. A two-dimensional histogram with the distributions of the m^2_{KK} and $m^2_{K\pi}$ invariant masses was produced from the simulation samples, representing the events *after* reconstruction and stripping. This histogram was divided by a similar histogram generated from the complete generator samples, where no reconstruction or stripping had yet taken place (with the full number of events, as shown in Table A.1). The results are shown in Figure 4.27, where each histogram bin represents the efficiency of the reconstruction and stripping process for a specific invariant mass coordinate pair. The fitted histograms show that the efficiency for this step is, on average, between 6 and 7%.

4.6.2 Triggers

The trigger acceptances are obtained using only the post-reconstruction and stripping simulation samples. A two-dimensional histogram is filled with the m^2_{KK} and $m^2_{K\pi}$ invari-

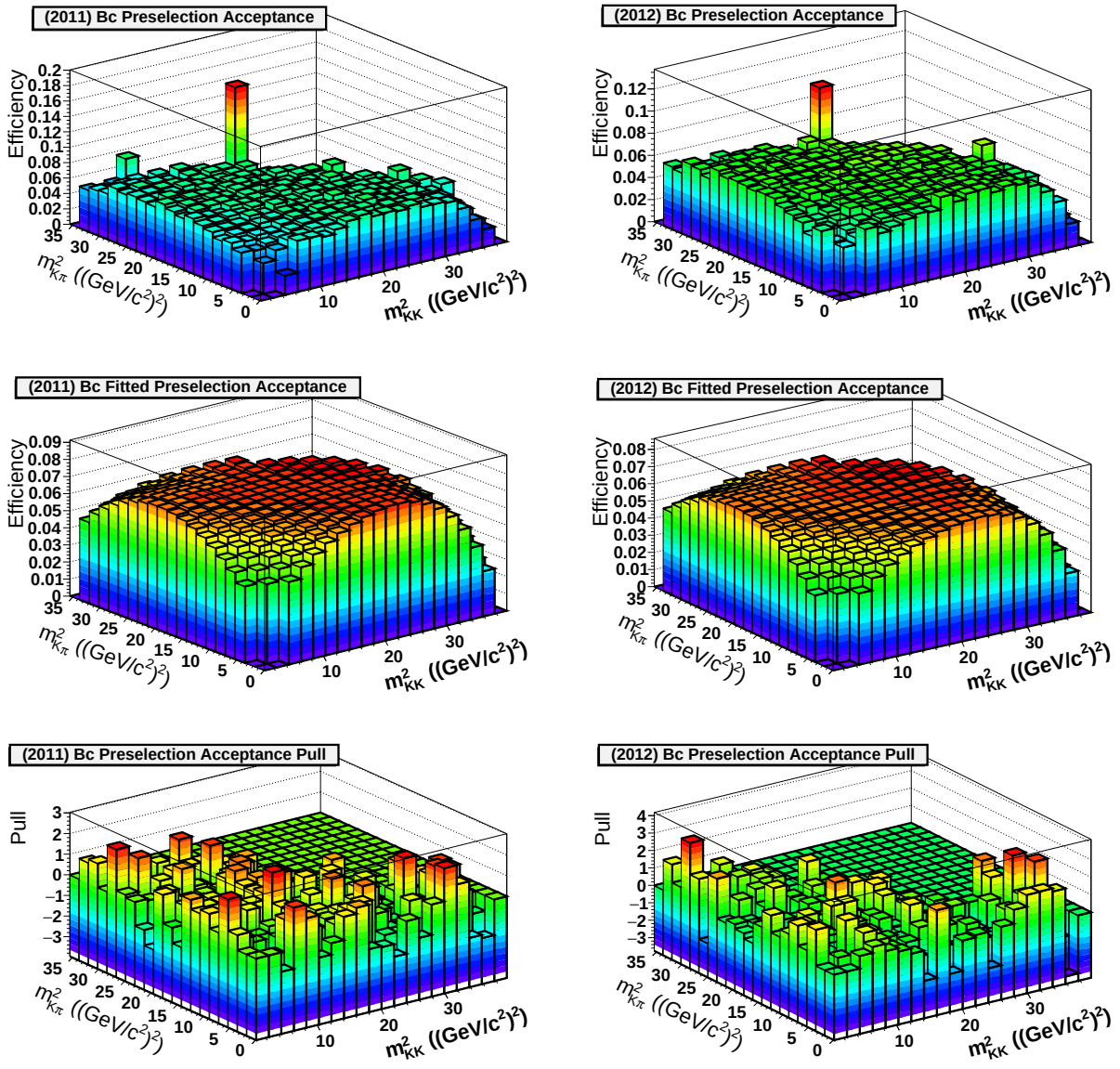

Figure 4.27: (Top) Raw, (middle) fitted and (bottom) pull distributions of the preselection acceptances for (left) 2011 and (right) 2012.

ant mass distributions *after* the offline trigger cuts shown in Table 4.4 are applied. This histogram is then divided by another two-dimensional histogram that is filled by the same mass distributions taken from the simulation samples, this time without any cuts (i.e. the *first* histogram of the previous subsection). The results are shown in Figure 4.28, and in this case each histogram bin represents the efficiency of the offline trigger selection for a

specific invariant mass coordinate pair. The fitted histograms for this step show that the average efficiency of the offline trigger selection is between 30 and 35%.

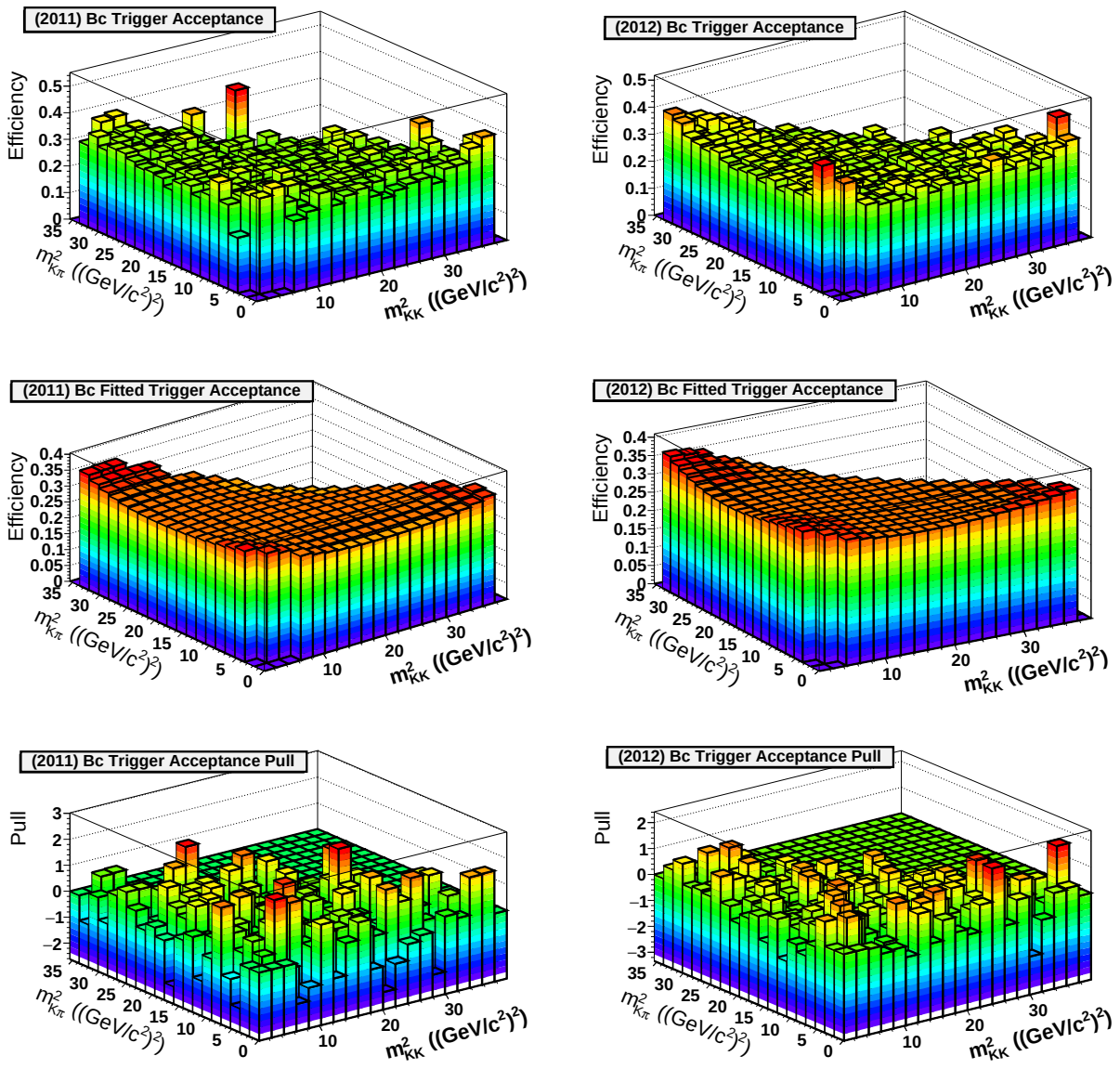

Figure 4.28: (Top) Raw, (middle) fitted and (bottom) pull distributions of the trigger acceptances for (left) 2011 and (right) 2012.

4.6.3 BDT

The acceptances for the BDT selection are obtained in similar fashion as the acceptances for the previous two steps. In this case, two two-dimensional histograms are filled with the distributions of the m^2_{KK} and $m^2_{K\pi}$ invariant masses, separately for both "loose" ($>$ 0.04) and "tight" ($>$ 0.18) BDT (see Section 4.1.3) cut configurations. These histograms are then divided by the invariant mass histogram after only the offline trigger selection. The acceptances for the loose and tight BDT cuts are shown in Figures 4.29 and 4.30, respectively. Here each bin represents the efficiency of the chosen BDT cut for a specific invariant mass coordinate pair. On average, the efficiency of the loose and tight BDT cuts are between 80-90% and 40-45%, respectively.

4.6.4 PID

The final acceptance histograms are obtained through application of the kaon and pion identification cuts discussed in Section 4.1.4. The PID efficiency variable produced by the `PIDCalib` package simulates the effect of the PID cuts, and each bin of the two-dimensional histogram filled with the post-loose BDT cut distributions of the m^2_{KK} and $m^2_{K\pi}$ invariant masses are weighed with this variable. This weighted histogram is then divided by the histogram filled with the post-loose BDT cut invariant mass distributions. The results are shown in Figure 4.31. Here the value of each bin represents the efficiency of the PID variable cuts for a specific invariant mass coordinate pair, which on average is between 40 and 45%.

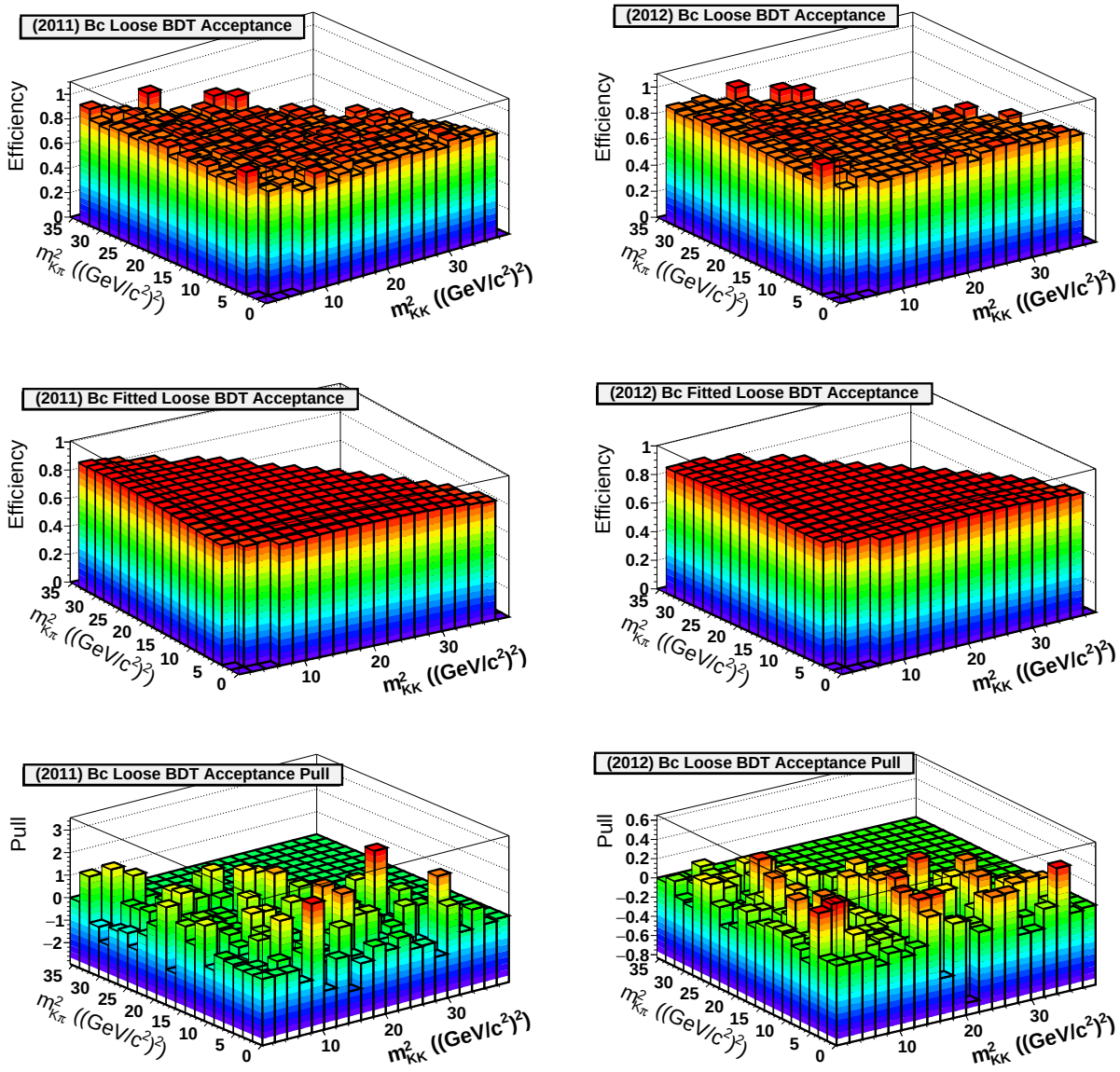

Figure 4.29: (Top) Raw, (middle) fitted and (bottom) pull distributions of the loose BDT acceptances for (left) 2011 and (right) 2012.

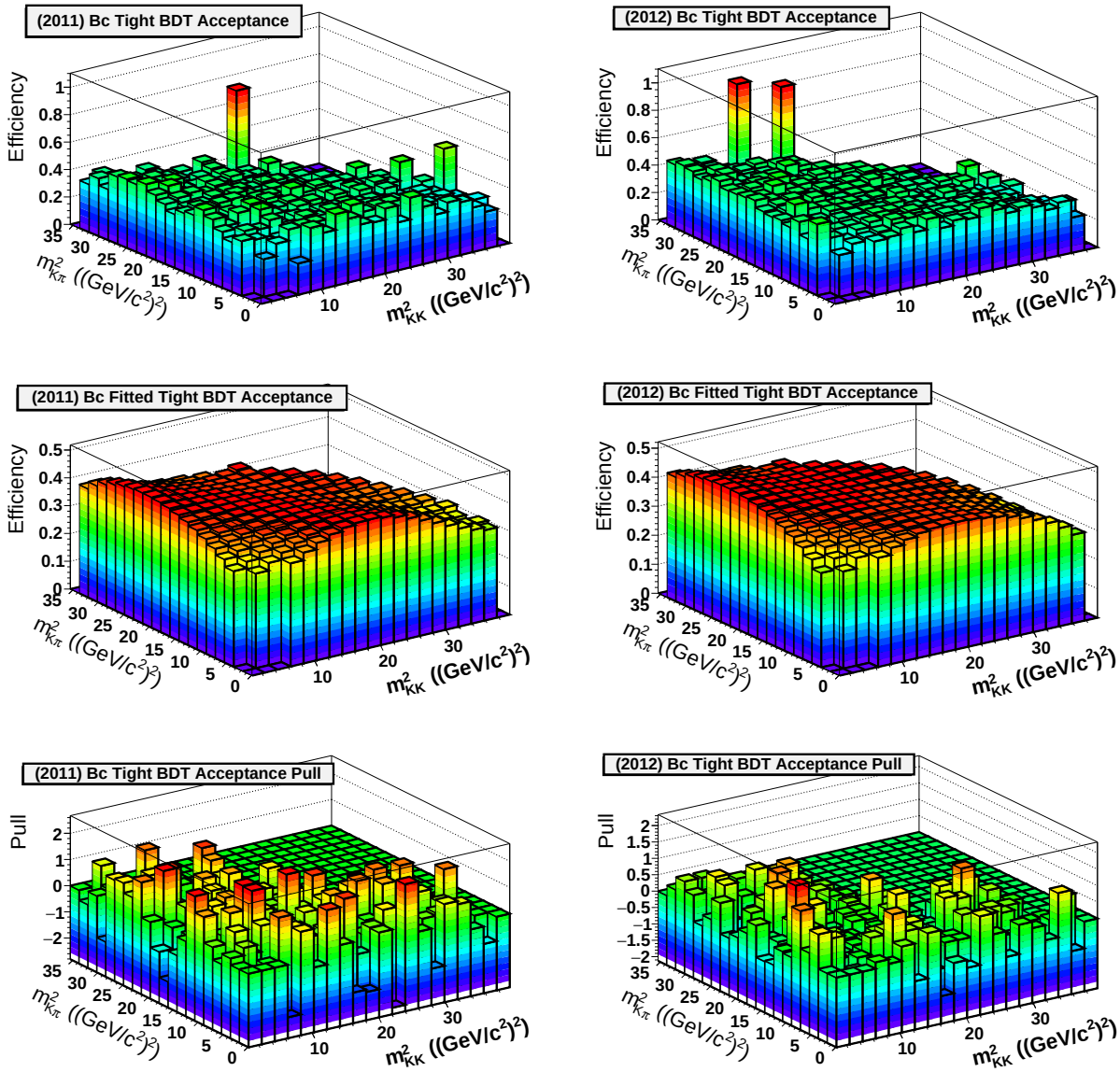

Figure 4.30: (Top) Raw, (middle) fitted and (bottom) pull distributions of the tight BDT acceptances for (left) 2011 and (right) 2012.

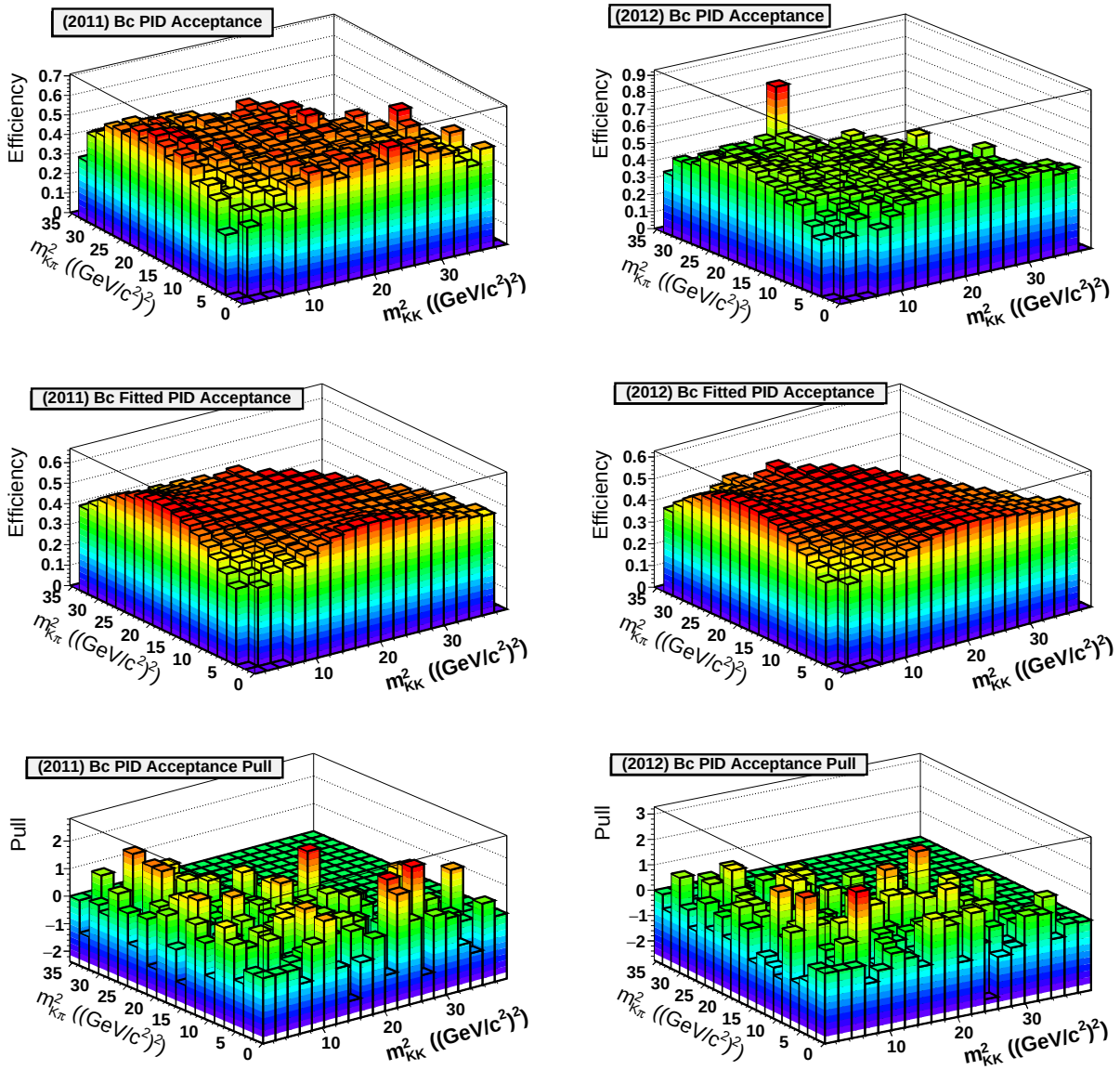

Figure 4.31: (Top) Raw, (middle) fitted and (bottom) pull distributions of the PID acceptances for (left) 2011 and (right) 2012.

4.7 Efficiencies, Systematics and Unblinding

The observable whose measurement is the objective of this analysis is the ratio between the products of the fragmentation and branching fractions for the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ mode and the chosen control channel. For the $B^{\pm}$ control mode, this ratio is expressed as:

$$\frac{f_c \times \mathcal{B}(B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm})}{f_u \times \mathcal{B}(B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm})} = \frac{N_c}{N_u} \times \frac{\epsilon_u}{\epsilon_c} \tag{4.7}$$

where $f_c = f(b \to B_x)$ ($x = u$ for B^+ and c for B_c^+), ϵ_x are the average efficiencies and N_x are the fitted yields. For the control modes including a D^0 intermediate resonance, the equation involves secondary branching fractions:

$$\frac{f_c \times \mathcal{B}(B_c^{\pm} \to [D^0]K^{\pm}) \times \mathcal{B}([D^0] \to K^{\pm}\pi^{\pm})}{f_u \times \mathcal{B}(B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm})} = \frac{N_{D^0}}{N_u} \times \frac{\epsilon_u}{\epsilon_{D^0}} \tag{4.8}$$

Therefore, the measurement of this ratio and subsequent inference of the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ branching fraction requires knowledge of the efficiencies and yields for each decay channel. The efficiencies are estimated using the simulation samples, and will be discussed in the following subsection, while the yields are obtained through the fits to the $m_{K^+K^-\pi^+}$ invariant mass. For the control channels, the yields are already known (see Section 4.5), while for the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$, the yields (if any) will only be known after unblinding is permitted. The analysis strategy following unblinding is discussed in subsection 4.7.3.

As in any physical measurement, there is statistical uncertainty associated with the measuring process. These *systematic* uncertainties are worth mentioning and their sources (and their respective errors) are identifiable, being shown in subsection 4.7.2.

4.7.1 Efficiencies

Efficiencies are calculated for the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ and $B_c^{\pm} \to D^0(K^{\pm}\pi^{\mp})K^{\pm}$ decays, and also for both $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ and $B^{\pm} \to D^0(K^{\pm}K^{\mp})\pi^{\pm}$ control channels. For the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ channel, efficiencies are derived using the fit functions for the acceptances shown in the previous section. In the absence of any information on the signal, a simple average of the efficiencies is calculated for both the region *below* the D^0 spectrum in the $m_{K^-\pi^+}$ invariant mass distribution ($m_{K\pi} < 1.834$ GeV/c^2) and for the full mass distribution, excluding the B region (m_{KK} and $m_{K\pi} < 5$GeV/c^2) and D^0 band from the $m_{K^-\pi^+}$ invariant mass distribution ($1.834 < m_{K\pi} < 1.894$ GeV/c^2). Tables 4.17 and 4.18 show the calculated efficiencies for each step, for both configurations.

Selection type	2011	2012
Reco & stripping	6.29 ± 0.03	5.66 ± 0.02
Triggers/Reco	32.39 ± 0.18	33.67 ± 0.15
$BDT > 0.04$	87.75 ± 0.23	88.20 ± 0.18
PID/Sel	51.29 ± 0.31	49.01 ± 0.24

Table 4.17: Efficiencies (in %) of the different steps, for both years, below the D^0 region in the $m_{K\pi}$ distribution. The efficiencies of the L0, HLT1, and HLT2 trigger levels are 53.24, 77.19, 78.57 and 50.71, 77.87, 83.21 for 2011 and 2012, respectively.

Selection type	2011	2012
Reco & stripping	7.15 ± 0.03	6.37 ± 0.02
Triggers/Reco	32.27 ± 0.18	33.69 ± 0.15
$BDT > 0.04$	89.29 ± 0.23	90.52 ± 0.18
PID/Sel	54.70 ± 0.31	52.19 ± 0.24

Table 4.18: Efficiencies (in %) of the different steps, for both years, excluding the D^0 band from the $m_{K\pi}$ spectrum and B regions from the $m_{K\pi}$ and m_{KK} spectrums. The efficiencies of the L0, HLT1, and HLT2 trigger levels are 52.45, 77.23, 79.57 and 51.92, 76.97, 83.45, for 2011 and 2012, respectively.

For the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$, a *combined* efficiency is calculated by performing a weighted average using the sWeights extracted from the fits shown in Section 4.5.1, in which the $1.834 < m_{K\pi} < 1.894$ GeV/c^2 band is removed, and the efficiencies for each step derived through the fitted $B^{\pm}$ acceptances (see Section A.5). This weighted average follows the harmonic average formula:

$$\epsilon = \frac{\sum\limits_{i} \text{sWeight}_i}{\sum\limits_{i} \frac{\text{sWeight}_i}{\epsilon_i}} \tag{4.9}$$

where ϵ_i is the product of the efficiencies of all steps, for each entry i. With this method, the combined efficiency for $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ is found to be $\epsilon_u = (1.71 \pm 0.12)\%$, where the uncertainty is calculated through the errors in the two-dimensional histogram bins. This efficiency is higher than the combined efficiency for the B_c^+ modes $\approx$ 0.9-1%, due to the larger lifetime of the B^+ ($\tau(B^+) > 2 \times \tau(B_c^+)$), that consequently grant it better triggering and selection capabilities. The efficiencies of the offline trigger variable cuts for the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ channel are shown in Table 4.19.

Selection type	2011	2012
$\to$ L0/Hlt1/Hlt2	53.63/82.10/86.64	47.09/81.43/92.25

Table 4.19: Trigger efficiencies (in %) for both years, for the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ normalization mode.

For the efficiencies of the $B^{\pm} \to D^0(K^{\pm}K^{\mp})\pi^{\pm}$ and $B_c^{\pm} \to D^0(K^{\pm}\pi^{\mp})K^{\pm}$ channels, no acceptance histograms are produced as these processes respectively only occur inside the $1.834 < m_{KK}$ or$m_{K\pi} < 1.894$ GeV/c^2 invariant mass band, and thus the efficiencies for each entry can be expected to be reasonably similar. Therefore, complete two-dimensional histogram representations of these invariant mass distributions are unnecessary, and the

efficiencies are calculated through the simple ratio between the number of events that pass each selection stage and the number of events *before* that stage. Tables 4.20 and 4.21 show the efficiencies for both modes.

Selection type	2011	2012
Reco & stripping	8.43 ± 0.03	7.67 ± 0.02
Triggers/Reco	37.83 ± 0.18	38.07 ± 0.15
$BDT > 0.04$	81.95 ± 0.23	79.95 ± 0.18
PID/Sel	50.76 ± 0.31	48.87 ± 0.24

Table 4.20: Efficiencies (in %) of the different steps, for both years, for the $B^{\pm} \to D^0(K^{\pm}K^{\mp})\pi^{\pm}$ simulation samples. The efficiencies of the L0, HLT1, and HLT2 trigger levels are 54.02, 74.36, 71.72 and 51.61, 74.88, 77.70, for 2011 and 2012, respectively.

Selection type	2011	2012
Reco & stripping	5.90 ± 0.03	5.42 ± 0.02
Triggers/Reco	28.81 ± 0.18	30.03 ± 0.15
$BDT > 0.04$	76.89 ± 0.23	78.53 ± 0.18
PID/Sel	50.76 ± 0.31	48.87 ± 0.24

Table 4.21: Efficiencies (in %) of the different steps, for both years, for the $B_c^{\pm} \to D^0(K^{\pm}\pi^{\mp})K^{\pm}$ simulation samples. The efficiencies of the L0, HLT1, and HLT2 trigger levels are 54.02, 74.36, 71.72 and 51.61, 74.89, 77.70, for 2011 and 2012, respectively.

4.7.2 Systematics and Corrections

The systematic uncertainties present in this analysis come from a number of different predictable sources. In this case, as the measured observable is a ratio, the systematic uncertainty is essentially the uncertainty on the ratio of the efficiencies $\frac{\epsilon_u}{\epsilon_c}$ (and is therefore called *relative* uncertainty). Therefore, any systematics common to both effiencies will cancel out in the ratio. The sources of systematic uncertainty relevant to this analysis are

discussed below, where it is important to note that the percentage of relative statistical error for some sources is not yet known, requiring first the approval to unblind.

- **PID**: This uncertainty is obtained through the `PIDCalib` package. Due to the precise binning scheme and abundant calibration sample size, the relative PID uncertainty for the PID is estimated to be less than 1%.

- **Limited MC statistics and acceptance**: For the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ decay, the limited sample of simulated events yields a relative uncertainty of 0.6%.

- **Fit model**: With current statistics, the impact of the fit function on the signal extraction is not expected to be substantial. The effect of varying some parameters of the signal and partially reconstructed background functions will be studied *after* unblinding.

- **Generator**: The relative uncertainty is extracted from the generator numbers shown in Table A.1. It is 0.2% for all B_c^+ modes.

- **Correction to B_c^+ lifetime**: The derivation of this uncertainty is explained in detail below, and it reaches 2%.

- **BDT shape**: The fractions of the signal yields measured in each BDT bin (see subsection 4.3.2) will be varied by small amounts after unblinding to infer the impact of the uncertainty on the BDT shape. In the case of an absence of signal, this source of systematics will be discarded.

Furthermore, corrections are applied that take into account known errors in MC simulations. The first of these corrections is for the *lifetime* of the B_c meson. The lifetime used in the simulated samples is 0.454 ps while the current (HFAG) average is 0.507 ± 0.009ps: this means that the B_c efficiency is underestimated. To correct for this, the simulated events that passed the stripping selection are reweighed using correction

factor $\tau_{sim}/\tau_{HFAG} \times \exp(-t/\tau_{HFAG})/\exp(-t/\tau_{sim})$, where τ represents the B_c lifetime and t represents the measured lifetime of each simulated event. The reweighed lifetime and BDT variable distributions are shown in Figure 4.32. The full selection cuts are performed to the reweighed events, and a correction factor is applied. This correction factor is calculated by first fitting the lifetime acceptance (ratio between the lifetime distributions of the post-stripping simulation samples and the pre-stripping generated samples), as shown in Figure 4.33. The function used for the fit is:

$$C(t) = \text{Norm} \times \frac{1}{1 - 1/(1 + (S_1 t)^p - \text{offset})} \times (1 - S_2 t) \tag{4.10}$$

The parameters of the final function used are on Table 4.22. The correction to the stripping efficiency is then calculated by manually generating equal numbers of events with both τ_{sim} and τ_{HFAG} lifetimes, applying the fitted acceptance function to both distributions and then integrating the result. The ratio between the integrated values for both lifetimes provides the efficiency ratio, which is found to be 1.132 ± 0.022. The final check performed is to see whether the time distribution reweighing changes the BDT variable *during* the TMVA training. Figure 4.34 shows the comparison between the default BDT variable and the BDT obtained through training using reweighed signal events: the two shapes are fully compatible.

Parameter	Value
S_1	1.716 ± 0.006
S_2	-0.05302 ± 0.0081
p	5.039 ± 0.038
offset	$-1.621e\text{-}05 \pm 3.73e\text{-}06$
Norm	0.2253 ± 0.0021

Table 4.22: Parameters of the time acceptance fit function.

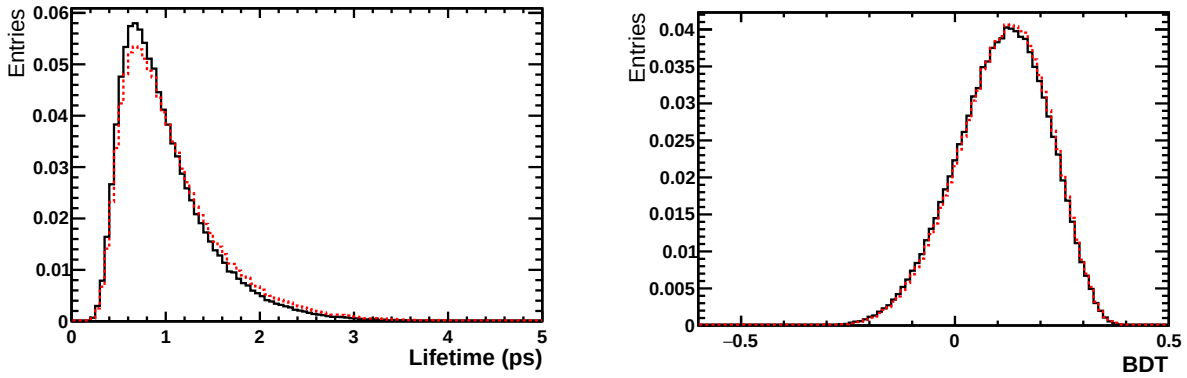

Figure 4.32: (left) True lifetime and (right) BDT variable distributions for (solid black line) unweighted and (dashed red line) lifetime-reweighted stripped $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ events.

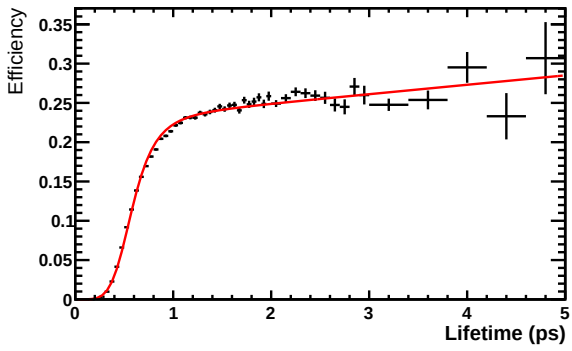

Figure 4.33: Lifetime acceptance for the stripping selection.

The tracking information used to simulate the MC samples also needs correcting. The tracking data/MC correction values are taken from [116], and Figures 4.35 and 4.36 show the maps of the correction factors for $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ and $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$, respectively. sWeights are used to weigh the correction to the $B^{\pm}$ channel, and the ratio between the corrections of the $B^{\pm}$ to $B_c^{\pm}$ channels, for both spectrum configurations, is shown in Table 4.23. The average track correction values for the $B_c^{\pm} \to D^0(K^{\pm}\pi^{\mp})K^{\pm}$ and $B^{\pm} \to D^0(K^{\pm}K^{\mp})\pi^{\pm}$, as well as the sWeighed track corrections to the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$, can be seen in Table 4.24.

The final correction concerns the L0 Hadron TOS trigger efficiency. The efficiency

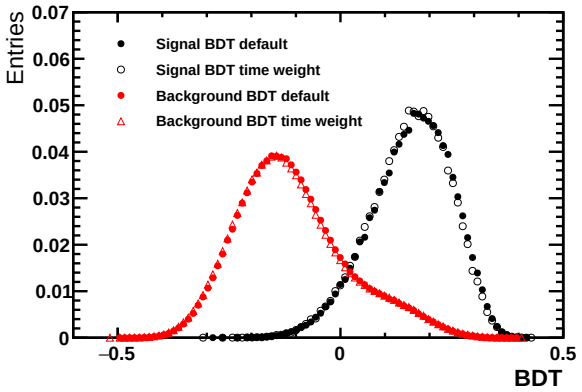

Figure 4.34: Signal and background shapes for default BDT and BDT obtained after reweighing the time distribution of the events.

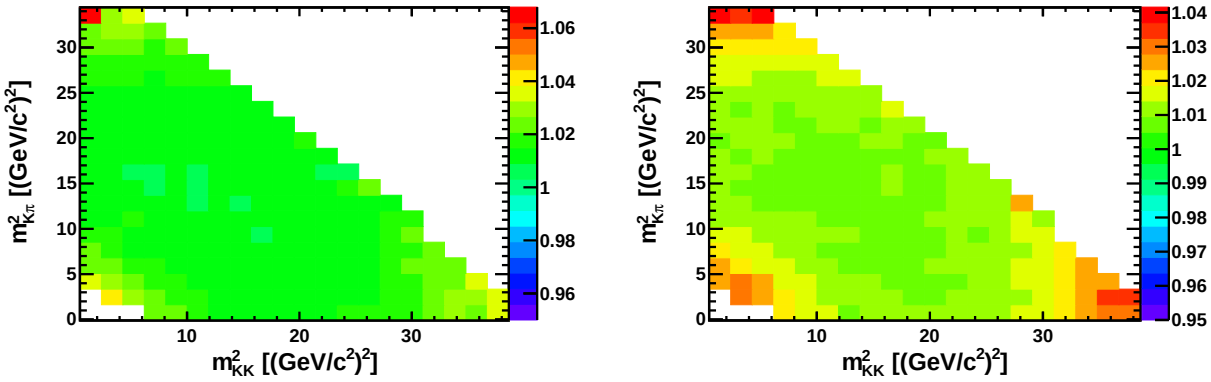

Figure 4.35: Distribution of the (left) 2011 and (right) 2012 tracking efficiency correction factor for $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ decays.

Channel	Year 2011	Year 2012
Below D^0 band in $m_{K\pi}$ distribution	1.00197 ± 0.00004	1.00791 ± 0.00003
Excluding D^0 band and B region	1.00197 ± 0.00004	1.00890 ± 0.00003

Table 4.23: Ratio of sWeighted correction for $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ to average correction for $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$.

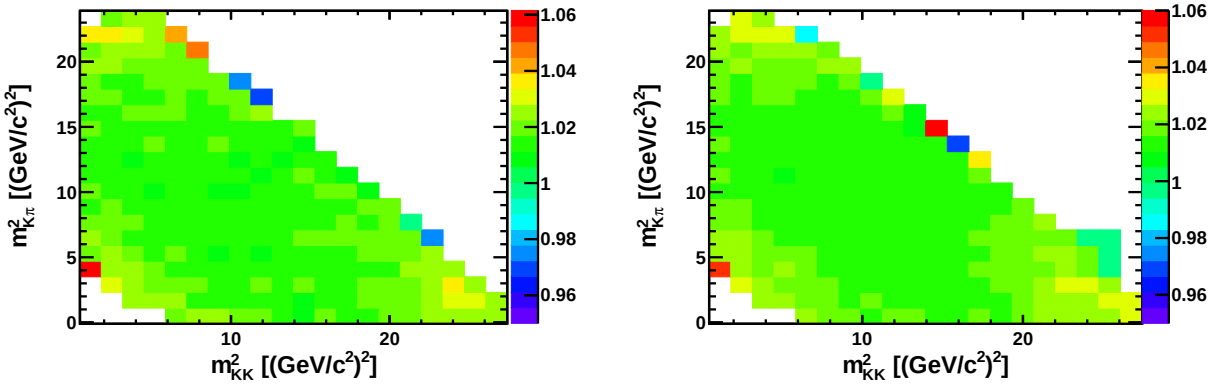

Figure 4.36: Distribution of the (left) 2011 and (right) 2012 tracking efficiency correction factor for $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ decays.

Channel	Year 2011	Year 2012
$B_c^{\pm} \to D^0(K^{\pm}\pi^{\mp})K^{\pm}$	$1.016 \pm 8.345\text{e-}05$	$1.015 \pm 8.163\text{e-}05$
$B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$	$1.020 \pm 2.213\text{e-}05$	$1.020 \pm 2.246\text{e-}05$
$B^{\pm} \to D^0(K^{\pm}K^{\mp})\pi^{\pm}$	$1.018 \pm 9.8\text{e-}05$	$1.016 \pm 9.7\text{e-}05$

Table 4.24: Average track corrections to $B_c^{\pm} \to D^0(K^{\pm}\pi^{\mp})K^{\pm}$, $B^{\pm} \to D^0(K^{\pm}K^{\mp})\pi^{\pm}$, and sWeighted track corrections to $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$.

tables for individual tracks (see [117]) are used to derive the probability that a given B decay passes the L0Hadron TOS requirement, by using the condition that at *least* one track is selected by the trigger. A Dalitz distribution of this probability is generated and compared to the efficiency map used in the simulated samples. The ratio of these two distributions provides the variation of the correction as a function of the position in the Dalitz plane, and this variation map is then used to compute the correction. The ratio of sWeighed correction to $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ to average correction for $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$, for both spectrum configurations, is shown in Table 4.25. The L0 data/MC corrections for

the $B_c^{\pm} \to D^0(K^{\pm}\pi^{\mp})K^{\pm}$, $B^{\pm} \to D^0(K^{\pm}K^{\mp})\pi^{\pm}$, and $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ channels can be seen in Table 4.26.

Configuration	Year 2011	Year 2012
Below D^0 band in $m_{K\pi}$ spectrum	1.085 ± 0.035	1.1085 ± 0.032
Excluding D^0 band and B region	0.859 ± 0.032	0.899 ± 0.017

Table 4.25: Ratio of sWeighted L0 correction for $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ to average L0 correction for $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$.

Channel	Year 2011	Year 2012
$B_c^{\pm} \to D^0(K^{\pm}\pi^{\mp})K^{\pm}$	0.7499 ± 0.1307	0.8131 ± 0.1924
$B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$	0.7890 ± 0.0197	0.7234 ± 0.0123
$B^{\pm} \to D^0(K^{\pm}K^{\mp})\pi^{\pm}$	0.7563 ± 0.0144	0.8100 ± 0.0103

Table 4.26: Average data/MC L0 corrections to $B_c^{\pm} \to D^0(K^{\pm}\pi^{\mp})K^{\pm}$, $B^{\pm} \to D^0(K^{\pm}K^{\mp})\pi^{\pm}$, and sWeighted data/MC L0 corrections to $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$.

4.7.3 Unblinding

Up to this point, every step in the analysis was performed without the researchers having access to any visual or quantitative information from the $[6.15, 6.35]$ GeV/c^2 range of the $m_{K^+K^-\pi^+}$ invariant mass distribution contained in the data samples. After the full selection is completed and the fit model confidently determined, permission to unblind this mass window must come from the LHCb Collaboration reviewers. Once this permission

is granted, the full mass fits are re-done with the full spectrum and the final steps of the study are performed.

As mentioned in Section 4.7, the objective of this analysis is to measure a ratio of the products of the fragmentation and branching fractions, for which a determination of the *signal yield* is necessary. Following unblinding, there are two possible signal yield cases for this analysis. The first is the case where a measurement of signal events is accomplished (see Figure 4.37). Depending on the quantity of measured events and the purity of the final samples, the *significance* (in σ) of the measurement is determined. By convention, significance levels of 3 σ and 5 σ represent "evidence" and "observation" levels, respectively. For this analysis, significances above 2 σ are unlikely for both annihilation decays and processes with intermediate D^0 mesons, so even if there are observed signal events, the only physically significant measurement that will be obtained is an upper limit on the observable ratio.

To accomplish this, based on the selected data and the final fit model, a series of "signal + background vs. background" hypotheses tests are performed for a range of possible values of signal events (N_{sig}). These statistical tests are implemented through the `RooStats` [118, 119] package. `RooStats` is a comprehensive collection of statistical tools which is incorporated into the `ROOT` framework that allows for, among other things, hypothesis testing and the determination of confidence intervals in statistical analyses. For each value of N_{sig}, the `RooStats` package constructs likelihood distributions of the chosen observable. In this case the observable is the aforementioned ratio, and from the likelihood distributions, *p-values* are calculated. The p-value is a measure of significance, using Gaussian standard deviations applied to the likelihood distributions, and p-values for both hypotheses are calculated and compared. The upper limit of the observable ratio is defined as the point where the p-value falls below the 5% level (i.e. above two Gaussian standard deviations), which defines the 95% confidence level (CL). Through the CL method, the upper limit of the observable ratio can be determined.

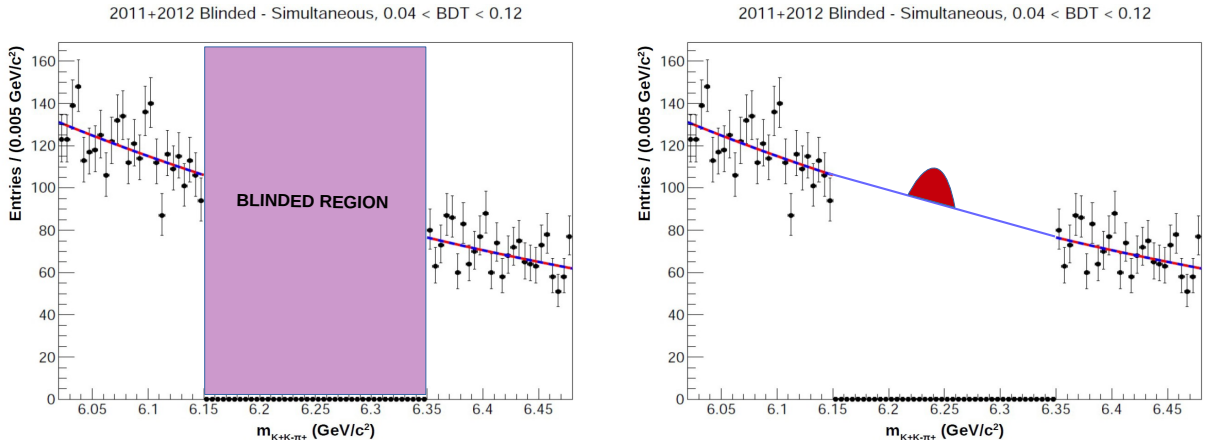

Figure 4.37: Representations of (left) blinded and (right) two post-unblinding scenarios. In the right image the blue line represents the lack of any signal events, while the red "bump" represents a concentration of events in the signal region.

The second case is the complete *lack* of any observable signal events (see Figure 4.37), which means that not enough B_c events passed the selection process in order to be visible among the other background contributions. In this case, it is impossible to ascertain with certainty the quantity of events that *did* survive (if any), but an *upper limit* of this quantity *is* possible to obtain, and therefore an upper limit on the observable ratio can also be determined using the `RooStats` package, in similar fashion to the first signal yield case. Therefore, even if no signal events are measured, a physically significant quantity will still be determined.

In conclusion, even though the final measurement will, in all likelihood, be only an upper limit of the observable ratio, any observation of B_c events will be an accomplishment in itself, due to the rarity of its decay processes and the relatively low statistics used for this study.

Chapter 5

Conclusion

This analysis represents the first search for B_c^+ annihilation decays, performed for $K^+K^-\pi^+$ final states. The objective of this dissertation was to detail the steps of the blind analysis leading up unblinding, when the observation of signal events (if any) will occur.

For the $B_c^+ \to K^+K^-\pi^+$ decay analysis, the full selection has been prepared and applied. Monte Carlo simulation samples of the $B_c^+ \to K^+K^-\pi^+$ decay were produced, with a reconstruction and stripping efficiency of $\approx 6\%$. The signal trigger efficiency applied to the simulation samples was found to be $\approx 33\%$. In a comparison of multivariate techniques, the Boosted Decision Tree method was chosen, as it offers higher signal efficiency and background rejection than the other tested methods. The training of the BDT variable was accomplished using the simulation samples to represent the B_c signal, while the background was represented by the side-bands taken from LHCb data samples. The BDT variable was applied to both simulation and data samples, and the BDT signal efficiency was found to be $\approx 88\%$ for loose cuts (BDT > 0.04) and $\approx 40\%$ for tight cuts (BDT > 0.18). The efficiencies of the particle identification variables were studied using calibration samples processed through the `PIDCalib` package, as the variables contained in the simulation samples themselves are not well simulated, and therefore are not reliable for efficiency studies. Through this method, the PID signal efficiency was found to be $\approx 50\%$. Combining all the selection cuts (with BDT > 0.04), the final estimated signal

efficiency is $\approx 0.9\%$.

In addition, the optimal fit model was determined. In the fit study, the combinatorial background was found to give the largest contribution to the data side-bands. The contributions from other decay channels were studied through simulation samples. The full selection was applied to these samples, in order to determine if their contributions were sizable enough to warrant the inclusion of specific fit functions in the final model. The $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}\pi^0$ partially reconstructed channel was found to have a significant contribution in the lower mass side-band, and it was therefore modeled in the final fit. Two cross-feed channels were also included, $B_c^{\pm} \to K^{\pm}K^{\mp}K^{\pm}$ and $B_c^{\pm} \to K^{\pm}\pi^{\mp}\pi^{\pm}$. Their contributions can not be completely determined before unblinding, as their yields are related to that of the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ decay itself. However, PID efficiencies for these channels were determined using simulation samples, and were found to be $\approx 12\%$ and 2%, respectively. Their addition to the fit model was found to increase the overall fit quality, further warranting their inclusion.

Two possible control channels have also been presented and studied. The $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ and $B^{\pm} \to D^0(K^{\pm}K^{\mp})\pi^{\pm}$ channels have both been studied in previous LHCb analyses, and their fragmentation and branching fractions measured. The fit models for both of these channels were studied and applied, using the full selection utilized for the $B_c^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ decay.

At this point, preparations have been completed for a possible observation of the B_c^+ signal coming from both annihilation decays and processes including an intermediate D^0 meson. An attempt at signal measurement is now possible, and now only the approval of the LHCb Collaboration is required to implement this final step of the analysis. It is important to stress that independently of any actual B_c^+ signal observation, a physically significant measurement will be obtained. Furthermore, the selection performed for this analysis will also be able to be implemented in future studies into these decay channels, using data samples with greatly increased statistics coming from future runs of the LHC.

Bibliography

[1] L. Hoddeson, L. Brown, M. Riordan, and M. Dresden. *The Rise of the Standard Model: Particle Physics in the* 1960*s and* 1970*s*. Cambridge University Press, 1997. 4

[2] W. N. Cottingham and D. A. Greenwood. *An Introduction to the Standard Model of Particle Physics*. Cambridge University Press, 2 edition, 2007. 4

[3] D. Griffiths. *Introduction to Elementary Particles*. Wiley-VCH, 2 edition, 2008. 4

[4] John Dalton. *A New System of Chemical Philosophy*, volume 1. William Dawson & Sons, 1808. 4

[5] J. J. Thomson. Cathode rays. *The Electrician*, 39:104, 1887. 4

[6] Hans Geiger. The scattering of α-particles by matter. *Proceedings of the Royal Society of London A*, 81(546):174–177, 1908. 5

[7] Ernest Rutherford. The scattering of α and β particles by matter and the structure of the atom. *Philosophical Magazine*, 21:669–688, 1911. 5

[8] Niels Bohr. On the constitution of atoms and molecules, part i. *Philosophical Magazine*, 26(151):1–24, 1913. 5

[9] Niels Bohr. On the constitution of atoms and molecules, part ii: Systems containing only a single nucleus. *Philosophical Magazine*, 26(153):476–502, 1913. 5

[10] Niels Bohr. On the constitution of atoms and molecules, part iii: Systems containing several nuclei. *Philosophical Magazine*, 26(155):857–875, 1913. 5

[11] James Chadwick. The existence of a neutron. *Proceedings of the Royal Society of London A*, 136(830):692–708, 1932. 5

[12] Arthur H. Compton. A quantum theory of the scattering of x-rays by light elements. *Physics Review*, 21(5):483–502, 1923. 5

[13] Max Planck. Über eine verbesserung der wienschen spektralgleichung. *Verhandlungen der Deutschen Physikalischen Gesellschaft*, 2:202–204, 1900. 5

[14] Max Planck. Zur theorie des gesetzes der energieverteilung im normalspektrum. *Verhandlungen der Deutschen Physikalischen Gesellschaft*, 2:237, 1900. 5

[15] Albert Einstein. Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt. *Annalen der Physik*, 17(6):132–148, 1905. 6

[16] Hideki Yukawa. On the interaction of elementary particles. *Proceedings of the Physico-Mathematical Society of Japan*, 17(48), 1935. 6

[17] S. H. Neddermeyer and C. D. Anderson. Note on the nature of cosmic-ray particles. *Physical Review*, 51(10):884–886, 1937. 6

[18] J. C. Street and E. C. Stevenson. New evidence for the existence of a particle of mass intermediate between the proton and the electron. *Physical Review*, 52(9):1003–1004, 1937. 6

[19] G.P.S. Occhialini and C.F. Powell. Observations on the production of mesons by cosmic radiation. *Nature*, 162:168–173, 1948. 6

[20] P. A. M. Dirac. The quantum theory of the electron. *Proceedings of the Royal Society of London A*, 117(778):610–624, 1928. 7

[21] Carl D. Anderson. The positive electron. *Physical Review*, 43(6):491–494, 1933. 7

[22] E. C. G. Stueckelberg. Remarque à propos de la création de paires de particules en théorie de la relativité. *Helvetica Physica Acta*, 14:588–594, 1941. 7

[23] Richard Feynman. The theory of positrons. *Physical Review*, 76(76):749, 1949. 7

[24] O. Chamberlain, E. Segrè, C. Wiegand, and T. Ypsilantis. Observation of antiprotons. *Physical Review*, 100(3):947, 1955. 7

[25] B. Cork, G. R. Lambertson, O. Piccioni, and W. A. Wenzel. Antineutrons produced from antiprotons in charge-exchange collisions. *Physical Review*, 104(4):1193, 1956. 7

[26] C. L. Cowan Jr, F. Reines, F. B. Harrison, H. W. Kruse, and A. D. McGuire. Detection of the free neutrino: a confirmation. *Science*, 124(3212):103–104, 1956. 8

[27] G. Danby, J-M. Gaillard, K. Goulianos, L. M. Lederman, N. Mistry, M. Scwhartz, and J. Steinberger. Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos. *Physical Review Letters*, 9(1):36, 1962. 8

[28] G. D. Rochester and C. C. Butler. Evidence for the existence of new unstable elementary particles. *Nature*, 160:855–857, 1947. 8

[29] K. Nishijima. Charge independence theory of v particles. *Physics Today*, 62(8):285, 1955. 8

[30] M. Gell-Mann. The interpretation of the new particles as displaced charge multiplets. *Il Nuovo Cimento*, 4(2):848–866, 1956. 8

[31] M. Gell-Mann and Y. Ne'eman. *The Eightfold Way*. New York, 1964. 9

[32] V. E. Barnes et al. Observation of a hyperon with strangeness minus three. *Physical Review Letters*, 12(8):204, 1964. 9

[33] M. Gell-Mann. A schematic model of baryons and mesons. *Physics Letters*, 8(3):214–215, 1964. 9

[34] G. Zweig. An su(3) model for strong interaction symmetry and its breaking. *CERN Report*, No. 8182, 1964. 9

[35] G. Zweig. An su(3) model for strong interaction symmetry and its breaking: II. *CERN Report*, No. 8419, 1964. 9

[36] J. I. Friedman and H. W. Kendall. Deep inelastic electron scattering. *Annual Review of Nuclear Science*, 22:203–254, 1972. 9

[37] M. Y. Han and Y. Nambu. Three-triplet model with double su(3) symmetry. *Physical Review*, 139(4B), 1965. 10

[38] J. Aubert et al. Experimental observation of a heavy particle j. *Physical Review Letters*, 33(23):1404–1406, 1974. 10

[39] J. Augustin et al. Discovery of a narrow resonance in e^+e^- annihilation. *Physical Review Letters*, 33(23):1406–1408, 1974. 10

[40] M. L. Perl, G. Abrams, A. Boyarski, M. Breidenbach, D. Briggs, F. Bulos, W. Chinowsky, J. Dakin, et al. Evidence for anomalous lepton production in e^+e^- annihilation. *Physical Review Letters*, 35(22):1489, 1975. 10

[41] D. C. Hom et al. Observation of a dimuon resonance at 9.5 gev in 400-gev proton-nucleus collisions. *Physical Review Letters*, 39:252–255, 1977. 10

[42] F. Abe et al. Observation of top quark production in $\bar{p}p$ collisions with the collider detector at fermilab. *Physical Review Letters*, 74(14):2626–2631, 1995. 10

[43] M. Gell-Mann. Symmetries of baryons and mesons. *Physical Review*, 125(3):1067–1084, 1962. 10

[44] PLUTO Collaboration. Jet analysis of the $\gamma(9.46)$ decay into charged hadrons. *Physics Letters B*, 82:449–455, 1979. 11

[45] S. L. Glashow and M. Gell-Mann. Gauge theories of vector particles. *Annals of Physics*, 15(3):437–460, 1961. 11

[46] A. Salam and J. C. Ward. On a gauge theory of elementary interactions. *Il Nuovo Cimento*, 19(1):165–170, 1961. 11

[47] UA1 Collaboration. Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ gev. *Physics Letters B*, 122:103–116, 1983. 11

[48] UA2 Collaboration. Evidence for $z^0 \to e^+e^-$ at the cern pp collider. *Physics Letters B*, 129:130–140, 1983. 11

[49] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Physical Review*, 96:191, 1954. 12

[50] K. A. Olive et al. Review of particle physics. *Chin. Phys.*, C38:090001, 2014. 14, 104, 108

[51] T. D. Lee and C. N. Yang. Question of parity conservation in weak interactions. *Physical Review*, 104(1):254–258, 1956. 14

[52] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Physical Review*, 105(4):1413–1415, 1957. 15

[53] P. W. Anderson. Plasmons, gauge invariance and mass. *Physical Review*, 105:439–442, 1963. 17

[54] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Physical Review Letters*, 13:321–333, 1964. 17

[55] J. Goldstone. Field theories with superconductor solutions. *Il Nuovo Cimento*, 19:154–164, 1961. 17

[56] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global conservation laws and massless particles. *Physical Review Letters*, 13:585–587, 1964. 17

[57] P. W. Higgs. Broken symmetries, massless particles and gauge fields. *Physics Letters*, 12:132–133, 1964. 17

[58] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Physical Review Letters*, 13:508–509, 1964. 17

[59] Y. Nambu. Quasi-particles and gauge invariance in the theory of superconductivity. *Physical Review*, 117:648–663, 1960. 17

[60] S. Weinberg. *The Quantum Theory of Fields, Volume 1: Foundations*. Cambridge University Press, 2005. 18

[61] N. Cabibbo. Unitary symmetry and leptonic decays. *Physical Review Letters*, 10(2):531–533, 1963. 19

[62] M. Kobayashi and T. Maskawa. CP-violation in the renormalizable theory of weak interaction. *Progress of Theoretical Physics*, 49(2):652–657, 1973. 19

[63] Belle Collaboration. *Belle Experiment Website*. `http://belle.kek.jp/`. 20

[64] BaBar Collaboration. *BaBar Experiment Website*. `http://www-public.slac.stanford.edu/babar/`. 20

[65] K. Abe et al. Observation of large CP violation in the neutral b meson system. *Physical Review Letters*, 87(9), 2001. 20

[66] H. Tajima et al. Belle b physics results. *International Journal of Modern Physics A*, 17(22):2967, 2002. 20

[67] B. Aubert et al. Observation of the bottomonium ground state in the decay $\gamma(3s) \to \gamma\eta_b$. *Physical Review Letters*, 101(7), 2008. 20

[68] The LHCb Collaboration. LHCb, a Large Hadron Collider beauty experiment for precision measurements of CP-violation and rare decays. Technical Report CERN-LHCC-1998-004 - LHCC/P4, CERN, Geneva, 1998. 20

[69] F. Abe et al. Observation of the B_c meson in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ tev. *Physical Review Letters*, 81(12):2432–2437, 1998. 21

[70] R. Aaij et al. Observation of the decay $B_c \to J/\psi K^+ K^- \pi^+$. *Journal of High Energy Physics*, 1311, 2013. 21

[71] R. Aaij et al. Observation of the decay $B_c^+ \to B_s^0 \pi^+$. *Physical Review Letters*, 111, 2013. 21

[72] I. P. Gouz et al. Prospects for the B_c studies at LHCb. *Physics of Atomic Nuclei*, 67:1559–1570, 2004. 21

[73] R. Aaij et al. Precision luminosity measurements at LHCb. *Journal of Instrumentation*, 9:P12005, 2014. 30

[74] A. A. Alves Jr. et al. The LHCb detector at the LHC. *Journal of Instrumentation*, 3, 2008. 32, 36, 42, 47, 49, 52, 53

[75] R. Aaij et al. LHCb detector performance. *International Journal of Modern Physics*, A30:1530022, 2015. 32

[76] R. Aaij et al. Performance of the LHCb Vertex Locator. *Journal of Instrumentation*, 9:P09007, 2014. 35

[77] The LHCb collaboration. LHCb VELO (VErtex LOcator). Technical Report CERN-LHCC-2001-011, CERN, Geneva, 2001. LHCb-TDR-005. 35

[78] The LHCb collaboration. LHCb inner tracker. Technical Report CERN-LHCC-2002-029, CERN, Geneva, 2002. LHCb-TDR-008. 39

[79] The LHCb collaboration. LHCb outer tracker. Technical Report CERN-LHCC-2001-024, CERN, Geneva, 2001. LHCb-TDR-006. 42

[80] The LHCb collaboration. LHCb magnet. Technical Report CERN-LHCC-2000-007, CERN, Geneva, 2000. LHCb-TDR-001. 42

[81] The LHCb collaboration. LHCb calorimeters. Technical Report CERN-LHCC-2000-036, CERN, Geneva, 2000. LHCb-TDR-002. 47

[82] The LHCb collaboration. LHCb RICH. Technical Report CERN-LHCC-2000-037, CERN, Geneva, 2000. LHCb-TDR-003. 49

[83] The LHCb collaboration. LHCb muon system. Technical Report CERN-LHCC-2001-010, CERN, Geneva, 2001. LHCb-TDR-004. 52

[84] The LHCb collaboration. LHCb Trigger and Online. Technical Report CERN-LHCC-2014-016, CERN, Geneva, 2014. LHCb-TDR-016. 53

[85] G. Barrand et al. *GAUDI - The Software Architecture and Framework for building LHCb Data Processing Applications*, 2000. 56

[86] LHCb. *Gauss*. `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss/`. 58

[87] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 physics and manual. *Journal of High Energy Physics*, 05:026, 2006. 58

[88] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A brief introduction to PYTHIA 8.1. *Computer Physics Communications*, 178:852–867, 2008. 58

[89] I. Belyaev et al. Handling of the generation of primary events in Gauss, the LHCb simulation framework. *Journal of Physics: Conference Series*, 331:032047, 2011. 58

[90] D. J. Lange. The EvtGen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research*, A462:152–155, 2001. 58

[91] S. Agostinelli et al. Geant4: a simulation toolkit. *Nuclear Instruments and Methods in Physics Research*, A506:250, 2003. 58

[92] John Allison, K. Amako, J. Apostolakis, H. Araujo, P.A. Dubois, et al. Geant4 developments and applications. *IEEE Transactions on Nuclear Science*, 53:270, 2006. 58

[93] LHCb. *Brunel*. `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/brunel/`. 58

[94] LHCb. *DaVinci*. `http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/davinci/`. 58

[95] CERN. *ROOT, a Data Analysis Framework*. `https://root.cern.ch/`. 59

[96] Wouter D. Hulsbergen. Decay chain fitting with a Kalman filter. *Nucl.Instrum.Meth.*, A552:566–575, 2005. 61

[97] CERN. *Worldwide LHC Computing Grid*. `http://wlcg-public.web.cern.ch/`. 62

[98] A. Hicheur, J. Stenzel Martins, and L. de Paula. Study of B_c^+ meson decays to the $K^+K^-\pi^+$ final state (in preparation). CERN-LHCb-ANA-2015-064. 64

[99] Vladimir V Gligorov. A single track HLT1 trigger. CERN-LHCb-PUB-2011-003, 2011. 67

[100] M Williams et al. The HLT2 Topological Lines. CERN-LHCb-PUB-2011-002, 2011. 67

[101] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss. *TMVA 4: Toolkit for Multivariate Data Analysis with ROOT - Users Guide*. CERN. `http://tmva.sourceforge.net/docu/TMVAUsersGuide.pdf`. 70

[102] B. P. Roe, H. Yang, J. Zhu, Y. Liu, I. Stancu, and G. McGregor. Boosted decision trees as an alternative toa rtificial neural networks for particle identification. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 543:577–584, 2005. 73

[103] R. A. Fisher. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7:179–188, 1936. 73

[104] P. D. Wentzell, D. T. Andrews, D. C. Hamilton, K. Faber, and B. R. Kowalski. Maximum likelihood principal component analysis. *Journal of Chemometrics*, 11:339–366, 1997. 73

[105] N. E. Day and D. F. Kerridge. A general maximum likelihood discriminant. *Biometrics*, 23:313–323, 1967. 73

[106] G. Punzi. Sensitivity of searches for new signals and its optimization. In L. Lyons, R. Mount, and R. Reitmeyer, editors, *Statistical Problems in Particle Physics, Astrophysics, and Cosmology*, page 79, 2003. 78

[107] `https://twiki.cern.ch/twiki/bin/view/LHCb/PIDCalibPackage`. 79

[108] *The RooFit Toolkit for Data Modeling*. `http://roofit.sourceforge.net/intro.html`. 83

[109] D. Hill and M. John. Search for $B_c^{\pm} \to D^{(*)}h^{\pm}$ decays. CERN-LHCb-ANA-2015-006. 90

[110] H. Albrecht, R. Glaser, G. Harder, A. Kruger, A. W. Nilsson, A. Nippe, T. Oest, M. Reidenbach, M. Chafer, W. Schmidt-Parzefall, H. Schroder, H. D. Schulz, F. Sefkow, and R. Wurth. Search for hadronic $b \to u$ decays. *Physics Letters B*, 241:278–282, 1990. 92

[111] A. Hicheur. Study of B_c^+ meson decays to the $p\bar{p}\pi^+$ final state. *Physics Letters B*, (arXiv:1603.07037):17. 94

[112] I Bediaga et al. Study of $B^{\pm} \to h^{\pm}h^{+}h^{-}$ decays in 2011 and 2012 data. LHCb-ANA-2014-050. 104, 156

[113] P. del Amo Sanchez et al. Study of $b \to x\gamma$ decays and determination of $|v_{td}/v_{ts}|$. *Physical Review D*, 82:7, 2010. 104

[114] Muriel Pivk and Francois R. Le Diberder. sPlot: A statistical tool to unfold data distributions. *Nucl.Instrum.Meth.*, A555:356–369, 2005. 105

[115] A Hicheur et al. Study of $B^{\pm} \to p\bar{p}h^{\pm}$ decays with 2011+2012 data. CERN-LHCb-ANA-2014-023. 108

[116] `https://twiki.cern.ch/twiki/bin/view/LHCb/TrackingEffStatus`2012S20. 123

[117] `https://twiki.cern.ch/twiki/bin/view/LHCbPhysics/CalorimeterObjects ToolsGroupDOC#L0_Hadron_trigger_efficiencies`. 125

[118] `https://twiki.cern.ch/twiki/bin/view/RooStats/WebHome`. 127

[119] `https://root.cern.ch/viewcvs/branches/dev/roostats/roofit/roostats/doc/usersguide/RooStats_UsersGuide.pdf`. 127

[120] http://lhcb-release-area.web.cern.ch/LHCb-release-area/DOC/gauss/generator/. 143

Appendix A

Supplementary Analysis Information

This appendix consists of important information related to the analysis, but which is not entirely relevant to the main text.

A.1 Simulated Samples and Generator Efficiency

Below are the tables with information on all simulated samples used in the analysis as well as the simulated generator efficiencies for each sample, as briefly discussed in subsection 4.1.1.

The numbers for the simulated samples are obtained after the generator efficiency [120] has been applied, which is mainly related to the geometrical acceptance of the detector.

Decay	Number of events (2011)	Number of events (2012)
Non-resonant $B_c^\pm \to K^\pm K^\mp \pi^\pm$	1006425 (MD) + 1052632 (MU)	2004911 (MD) + 2009211 (MU)
$B_c^\pm \to K^\pm K^\mp K^\pm$	1040948 (MD) + 1030823 (MU)	2092303 (MD) + 2006147 (MU)
$B_c^\pm \to K^\pm \pi^\mp \pi^\pm$	1013926 (MD) + 1061385 (MU)	2042698 (MD) + 2042540 (MU)
$B_c^\pm \to \pi^\pm \pi^\mp \pi^\pm$	1034964 (MD) + 1061097 (MU)	2068599 (MD) + 2060223 (MU)
$B_c^\pm \to K^\pm K^\mp K^\pm \pi^0$	1010984 (MD) + 1001461 (MU)	2021501 (MD) + 2035316 (MU)
$B_c^\pm \to K^\pm K^\mp \pi^\pm \pi^0$	1008162 (MD) + 1037790 (MU)	2009568 (MD) + 2013695 (MU)
$B_c^\pm \to K^\pm \pi^\mp \pi^\pm \pi^0$	1034167 (MD) + 1003223 (MU)	2009749 (MD) + 2044381 (MU)
$B_c^\pm \to \pi^\pm \pi^\mp \pi^\pm \pi^0$	1038087 (MD) + 1027479 (MU)	2015499 (MD) + 2047593 (MU)
$B_c^\pm \to D^0(K^\pm \pi^\mp)\pi^\pm$	505943 (MD) + 520410 (MU)	1000506 (MD) + 1000360 (MU)
Non-resonant $B^\pm \to K^\pm K^\mp \pi^\pm$	269996 (MD) + 253000 (MU)	519250 (MD) + 519999 (MU)

Table A.1: Number of generated events for the different samples. Non resonant implies here in a flat phase space. MU and MD denote Magnet Up and Magnet Down configurations, respectively.

Decay	Generator efficiency (2011/2012) (%)
Non resonant $B_c^\pm \to K^\pm K^\mp \pi^\pm$	$14.336 \pm 0.029 / 14.738 \pm 0.03$
$B_c^\pm \to K^\pm K^\mp K^\pm$	$14.793 \pm 0.030 / 15.201 \pm 0.030$
$B_c^\pm \to K^\pm \pi^\mp \pi^\pm$	$13.942 \pm 0.028 / 14.38 \pm 0.030$
$B_c^\pm \to \pi^\pm \pi^\mp \pi^\pm$	$13.571 \pm 0.028 / 13.952 \pm 0.028$
$B_c^\pm \to K^\pm K^\mp K^\pm \pi^0$	$13.365 \pm 0.017 / 13.833 \pm 0.018$
$B_c^\pm \to K^\pm K^\mp \pi^\pm \pi^0$	$12.871 \pm 0.017 / 13.352 \pm 0.018$
$B_c^\pm \to K^\pm \pi^\mp \pi^\pm \pi^0$	$12.415 \pm 0.016 / 12.865 \pm 0.018$
$B_c^\pm \to \pi^\pm \pi^\mp \pi^\pm \pi^0$	$11.995 \pm 0.016 / 12.441 \pm 0.017$
$B_c^\pm \to D^0(K^\pm \pi^\mp)\pi^\pm$	$14.183 \pm 0.020 / 14.560 \pm 0.030$
Non resonant $B^\pm \to K^\pm K^\mp \pi^\pm$	$16.825 \pm 0.031 / 16.915 \pm 0.031$

Table A.2: Generator efficiencies for the simulated modes (Sim08e/Sim08f/Sim08h, pythia 6, `BcVegPy`).

A.2 Multi-Variate Analysis Variables and Graphs

A.2.1 Discriminating Variables and Separation Power

Rank	Variable	Separation
1	$\text{pointing}_{\text{P}_\text{T}}$	4.493e-01
2	MAXDOCA	3.293e-01
3	$p_T(B)$	3.026e-01
4	p_T^{max}	2.775e-01
5	IP sum	2.421e-01
6	p_T^{med}	2.079e-01
7	B vertex χ^2	2.051e-01
8	$cos(\theta)$	1.853e-01
9	d_3 track GhostProb	1.730e-01
10	d_2 track GhostProb	1.569e-01
11	d_1 track GhostProb	1.564e-01
12	IP of p_T^{max}	1.131e-01
13	B FD	9.900e-02
14	IP χ^2 sum	7.094e-02
15	B IP	6.038e-02
16	B IP χ^2	5.287e-02
17	BP	2.795e-02
18	d_3 track χ^2/NDoF	2.354e-02
19	d_1 track χ^2/NDoF	1.780e-02
20	d_2 track χ^2/NDoF	1.639e-02
21	d_3 P	3.972e-03
22	d_2 P	3.941e-03
22	d_1 P	3.371e-03

Table A.3: Discriminating variables used in the MVA training, sorted according to their decreasing separation power. In this case, $d_1 = K^+$ or K^-, $d_2 = K^-$ or K^+, and $d_3 = \pi^\pm$.

- **pointing**$_{p_T}$ - Variable defined as per the following equation:

$$\text{pointing}_{\text{p}_\text{T}} = \frac{\text{p}_\text{B}\sin(\theta)}{\text{p}_\text{B}\sin(\theta) + \sum_\text{i} \text{p}_\text{T}^\text{i}} \tag{A.1}$$

  Where θ is the angle between the B_c candidate flight vector and its momentum, p_B is the momentum of the B_c candidate, and $\sum_i p_T^i$ is the sum over the transverse

momentum of all the tracks of the decay. In order of separation power, the variables are defined below. Titles of the variables are written in parentheses as they are defined in the data sets.

- MAXDOCA (`B_AMAXDOCA`) - Maximum value of the Distance Of Close Approach (DOCA) variable.
- $p_T(B)$ (`B_PT`) - Transverse momentum of the B_c candidate.
- p_T^{max} (`ptmax`) - Maximum transverse momentum among the three daughter tracks.
- IP sum (`ipownpvsum`) - Sum of the tracks' impact parameters.
- p_T^{med} (`ptmed`) - The second largest transverse momentum among the three daughter tracks.
- B vertex χ^2 (`B_ENDVERTEX_CHI2`) - χ^2 associated with the B_c candidate vertex's reconstruction.
- $cos(\theta)$ (`B_DIRA_OWNPV`) - Cosine of the angle (θ) between the B_c candidate flight vector and its momentum.
- d_3 track GhostProb (`d3_TRACK_GhostProb`) - Probability that the third daughter's track belongs to a ghost particle.
- d_2 track GhostProb (`d2_TRACK_GhostProb`) - Probability that the second daughter's track belongs to a ghost particle.
- d_1 track GhostProb (`d1_TRACK_GhostProb`) - Probability that the first daughter's track belongs to a ghost particle.
- IP of p_T^{max} (`ip_ptmax`) - Impact parameter of the daughter track with highest transverse momentum.

- B FD (`B_FD_OWNPV`) - Flight distance of the B_c candidate.
- IP χ^2 sum (`ipchi2ownpvsum`) - Sum of χ^2 associated with the impact parameters of the tracks.
- B IP (`B_IP_OWNPV`) - Impact parameter of the B_c candidate.
- B IP χ^2 (`B_IPCHI2_OWNPV`) - χ^2 associated with the impact parameter of the B_c candidate.
- BP (`B_P`) - Momentum of the B_c candidate.
- d_3 track χ^2/NDoF (`d3_TRACK_CHI2NDOF`) - χ^2 per degree of freedom associated with the reconstruction of the third daughter's track.
- d_2 track χ^2/NDoF (`d2_TRACK_CHI2NDOF`) - χ^2 per degree of freedom associated with the reconstruction of the second daughter's track.
- d_1 track χ^2/NDoF (`d1_TRACK_CHI2NDOF`) - χ^2 per degree of freedom associated with the reconstruction of the first daughter's track.
- d_3 P (`d3_P`) - Momentum of the third daughter.
- d_2 P (`d2_P`) - Momentum of the second daughter.
- d_1 P (`d1_P`) - Momentum of the first daughter.

A.2.2 Correlation Matrices for Discriminating Variables

Correlation matrices showing the degree of correlation (in %) are generated by the TMVA for both signal and background events, and are shown below for both 2011 and 2012 training.

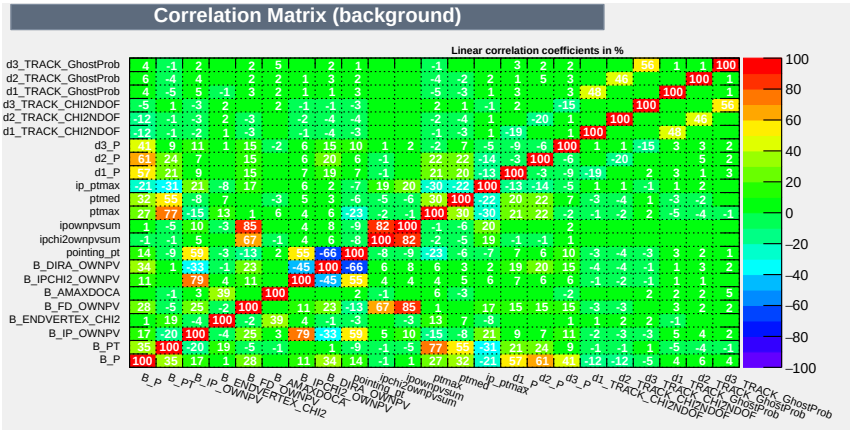

Figure A.1: Background correlation matrix for the 23 variables selected for the 2011 BDT training.

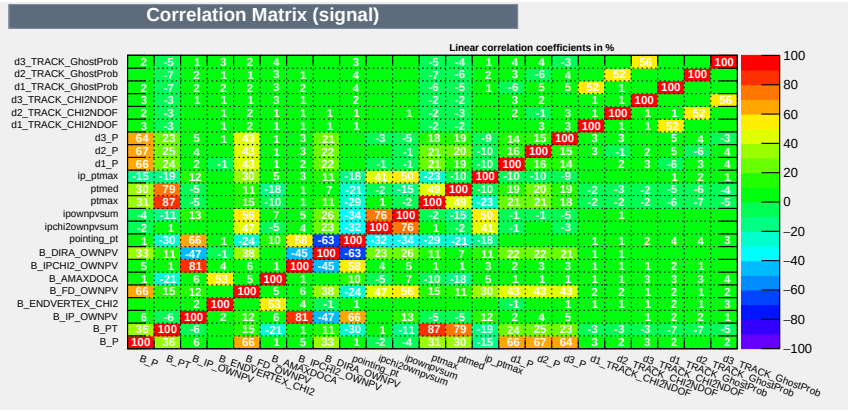

Figure A.2: Signal correlation matrix for the 23 variables selected for the 2011 BDT training.

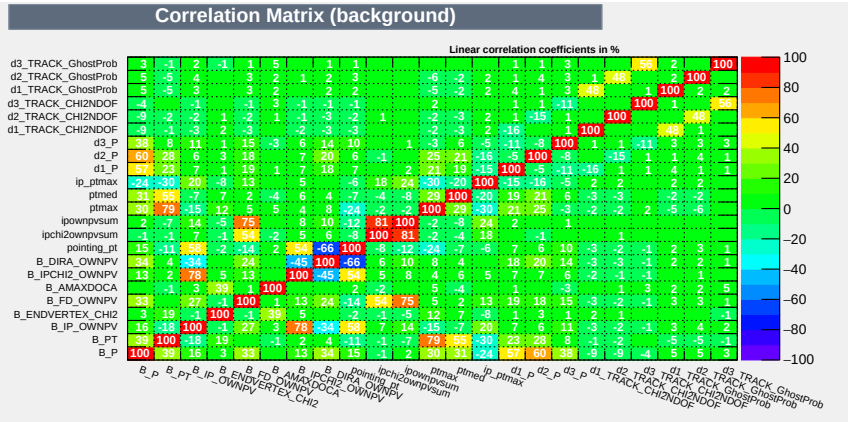

Figure A.3: Background correlation matrix for the 23 variables selected for the 2012 BDT training.

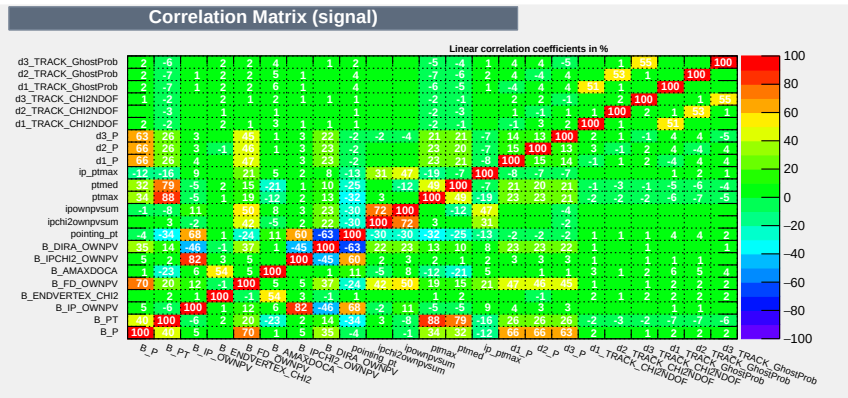

Figure A.4: Signal correlation matrix for the 23 variables selected for the 2012 BDT training.

A.2.3 BDT Variable Overtraining Checks

The first overtraining check is provided by the TMVA itself, comparing test and training samples. Figure A.5 shows the generated graphs with these comparisons, where there is no observed overtraining.

A second check was performed in order to make sure that there were no noticeable biases due to sample specificity. Signal and background samples were divided into different subsamples with the following configurations:

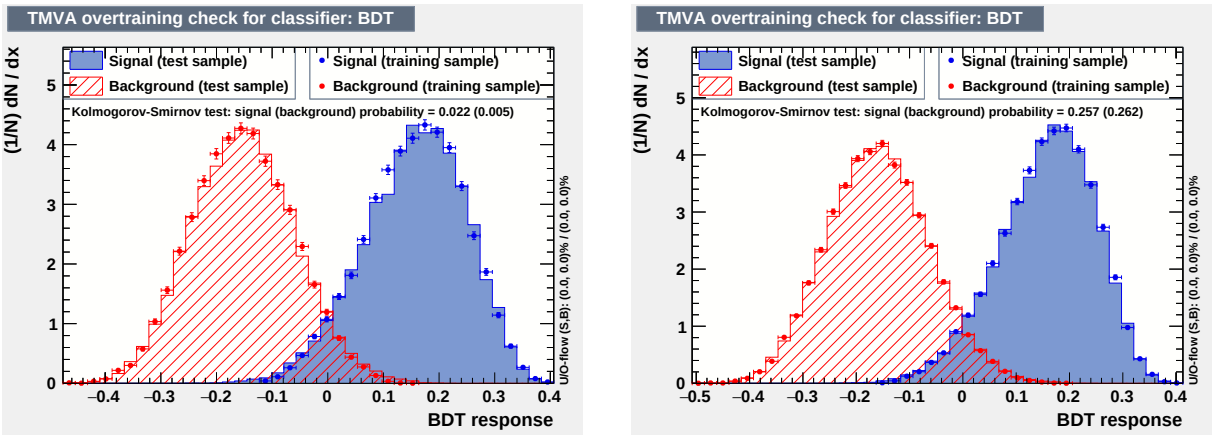

Figure A.5: Overtraining check performed by the TMVA for (left) BDT 2011 and (right) BDT 2012.

- Configuration A: signal represented by events having even event numbers from the signal sample and background from magnet down sample.

- Configuration B: signal represented by events having odd event numbers from the signal sample and background from magnet up sample.

- Configuration C: building samples with events having even event numbers in the full signal and background samples.

- Configuration D: building samples with events having odd event numbers in the full signal and background samples.

For each configuration, a BDT is obtained and tested against the samples of the complementary configuration. No significant overtraining was observed, as shown in Figures A.6 and A.7.

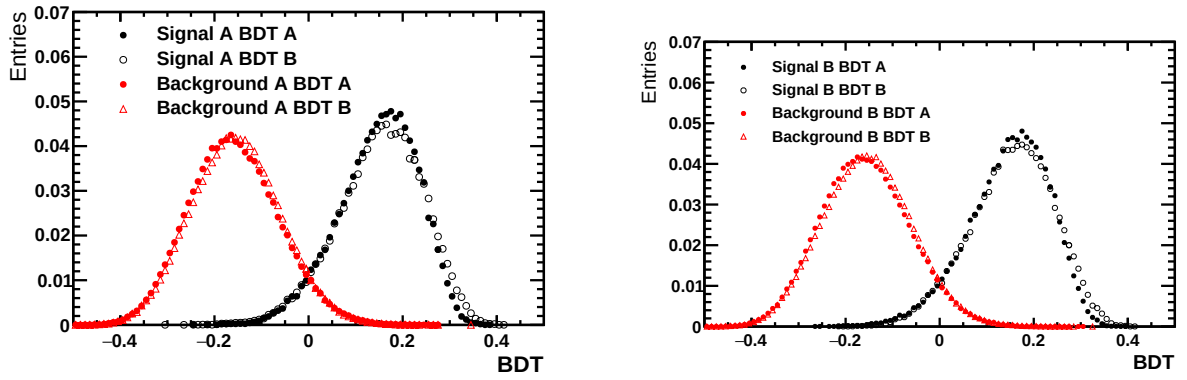

Figure A.6: BDT splitting exercise test sample A (left) and test sample B (right) against trained BDTA and BDTB.

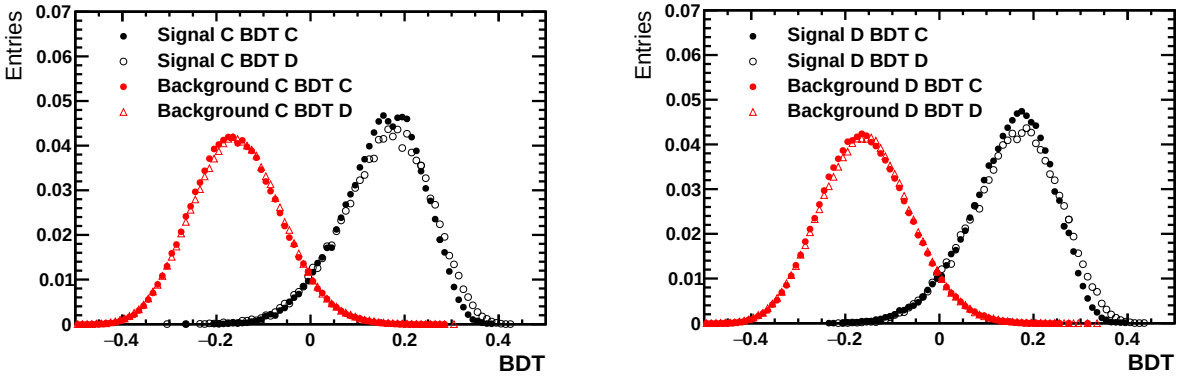

Figure A.7: BDT splitting exercise test sample C (left) and test sample D (right) against trained BDTC and BDTD.

A.3 Toy MC Studies for B_c^+ Signal Extraction

A.3.1 Tight BDT Cut Fit

The distributions of yields and pulls are shown in Figs. A.8. Even for the low signal yield case, $N_S = 10$, there is no significant bias.

A.3.2 Simultaneous fit

The distributions of yields and pulls are shown in Figs. A.9. As was the case for the Tight BDT cut fit, here there is no significant bias, even for the low signal yield case.

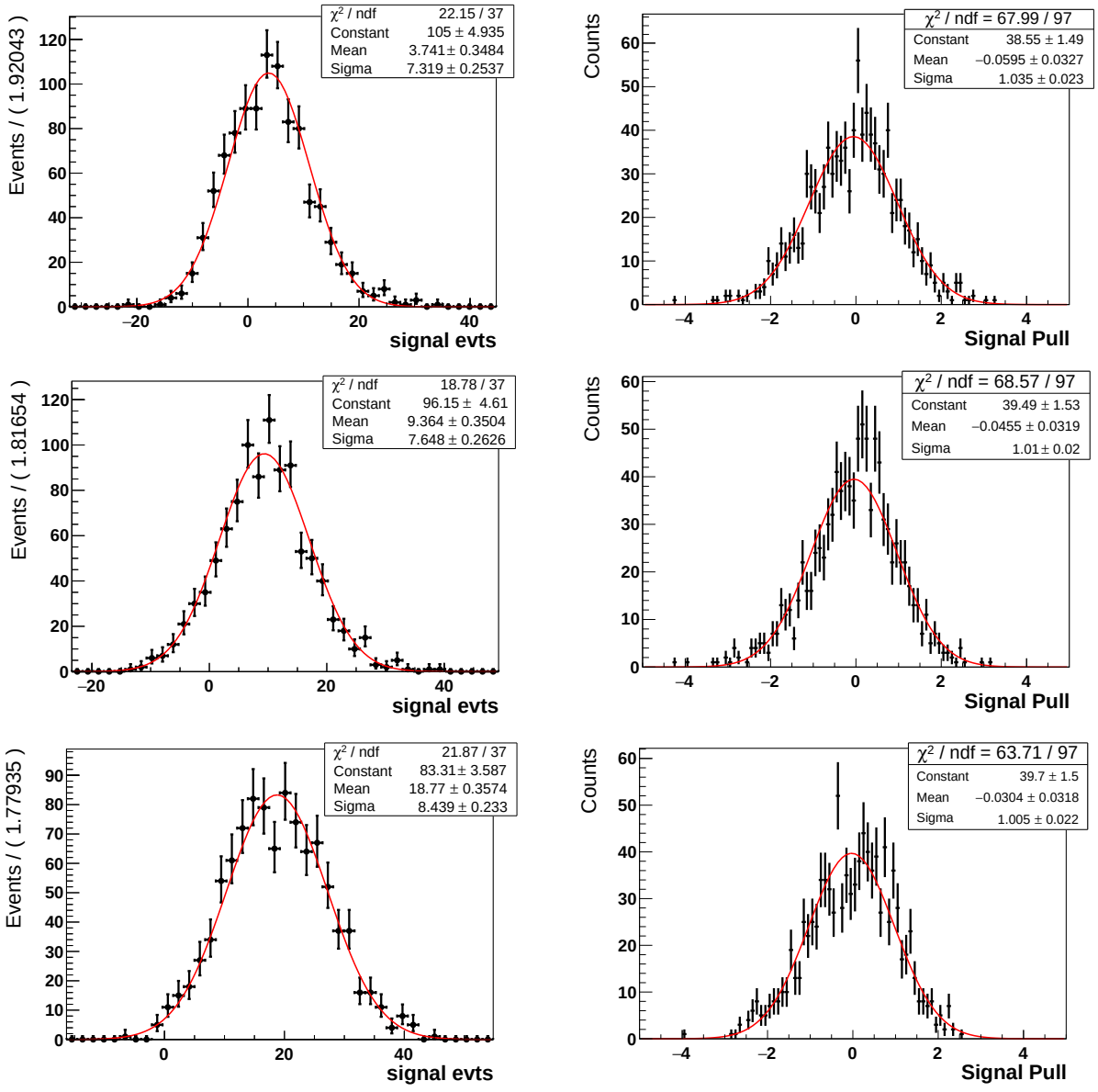

Figure A.8: (Left) yields and (right) pulls toy results for N_S input = (Top) 10, (Middle) 25 and (Bottom) 50 for the tight BDT cut fit.

A.4 $B^{\pm}$ mass fit components

In this appendix, the component models used for the $B^{\pm}$ full mass fit shown in Section 4.5.1 will be detailed, followed by the acceptances for the $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$ simulation samples.

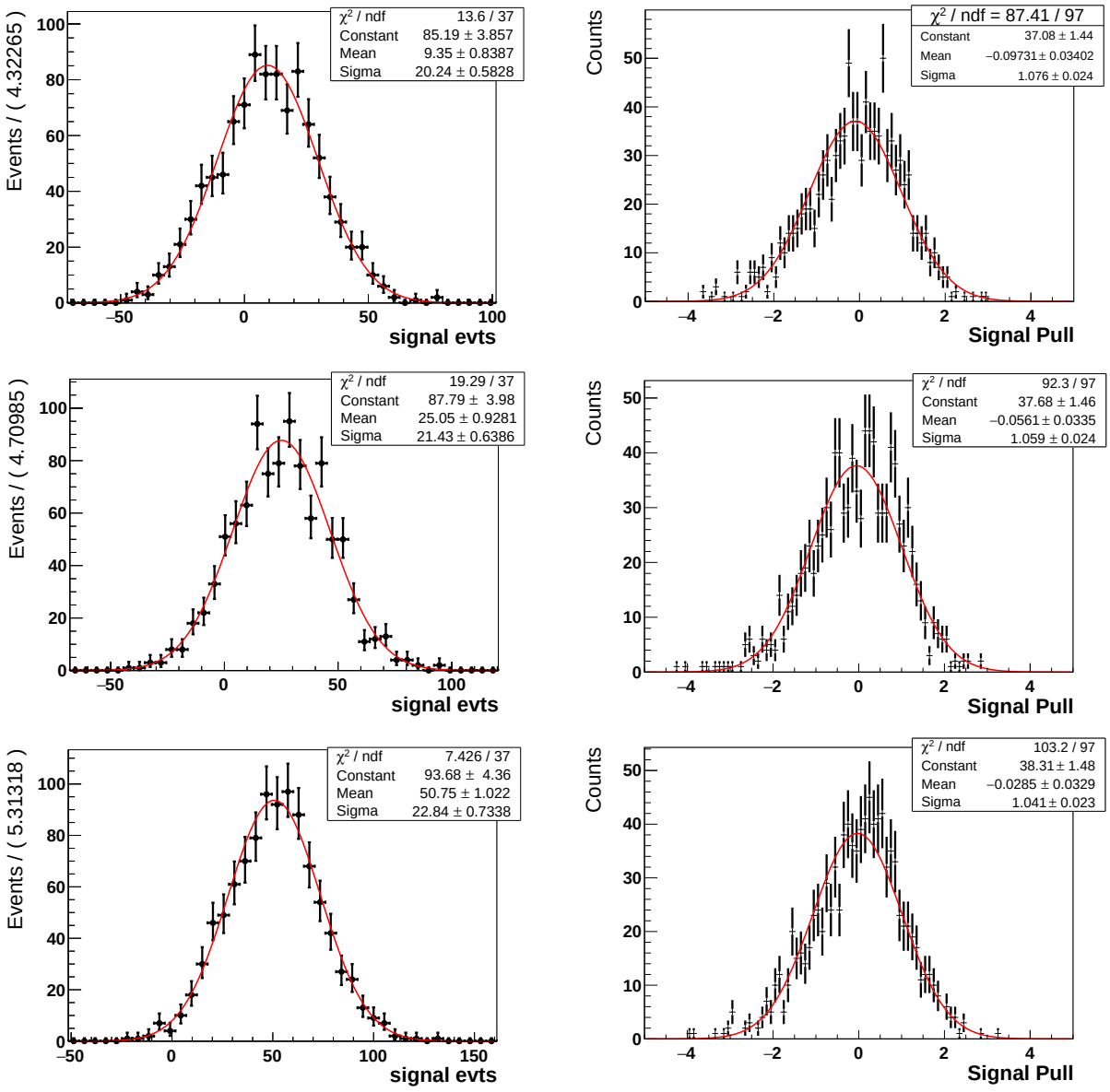

Figure A.9: (Left) yields and (right) pulls toy results for N_S input = (Top) 10, (Middle) 25 and (Bottom) 50 for the simultaneous fit.

A.4.1 $B^{\pm}$ signal model

The $B^{\pm}$ signal is modeled using a Double Crystal Ball function, in a similar fashion to what is shown in Section 4.2.1. The CB function is described as:

$$f(m;\mu,\sigma,\alpha,n) \propto \begin{cases} \exp(-\frac{(m-\mu)^2}{2\sigma^2}) : \frac{m-\mu}{\sigma} > -\alpha \\ A.(B - \frac{m-\mu}{\sigma})^{-n} : \frac{m-\mu}{\sigma} \leq -\alpha \end{cases} \quad \text{(A.2)}$$

The fits to the $B^{\pm}$ signal coming from MC simulation samples are shown in Figure A.10, while the fit parameters are shown in Tables A.4 and A.5.

Parameter	Value
μ_{CB}	$5.28043 \pm 0.00013\text{GeV}/c^2$
σ_{CB}	$0.01410 \pm 0.00044\text{GeV}/c^2$
σ^2_{CB}	$0.0194 \pm 0.0013\text{GeV}/c^2$
n	1.04 ± 0.15
n_2	4.00 ± 0.56
α	1.83 ± 0.15
α_2	-1.750 ± 0.093
$frac(1/2)$	0.568 ± 0.070

Table A.4: Parameters of the Double Crystal Ball $B^{\pm}$ signal model, for year 2011.

Parameter	Value
μ_{CB}	$5.280240 \pm 0.000082\text{GeV}/c^2$
σ_{CB}	$0.02224 \pm 0.00057\text{GeV}/c^2$
σ^2_{CB}	0.01371 ± 0.00033
n	1.41 ± 0.11
n_2	1.00 ± 0.04
α	1.624 ± 0.070
α_2	-2.528 ± 0.061
$frac(1/2)$	0.473 ± 0.040

Table A.5: Parameters of the Double Crystal Ball $B^{\pm}$ signal model, for year 2012.

A.4.2 Combinatorial background

The combinatorial background is modeled by a simple exponential function $\propto \exp(c.(m - 5.080))$.

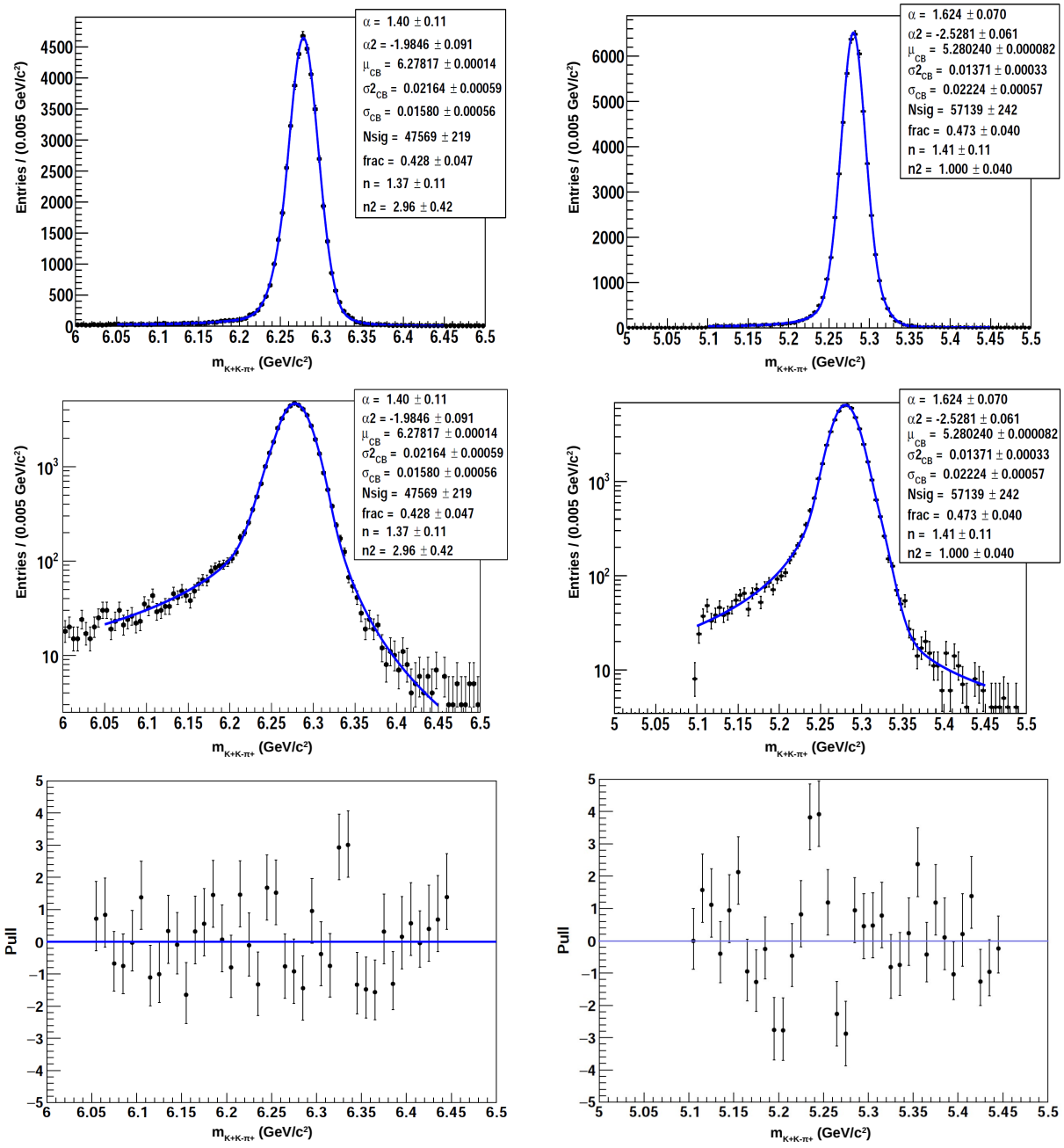

Figure A.10: $B^{\pm}$ signal fits to MC simulation samples, for (left) 2011 and (right) 2012. (top) Fits, (middle) Fits with Log-scale on Y axis, (bottom) pulls.

A.4.3 $B^{\pm} \to K^{\pm}K^{\mp}K^{\pm}$ and $B^{\pm} \to K^{\pm}\pi^{\mp}\pi^{\pm}$ cross-feeds

The $K^{\pm}K^{\mp}K^{\pm}$ and $K^{\pm}\pi^{\mp}\pi^{\pm}$ cross-feed components are both modeled in the same fashion as in (reference note), i.e. using a Cruijff function. The Cruijff function $\mathcal{C}(m)$ is a Gaussian

function with different left-right resolutions (σ_1,σ_2) and non-Gaussian tails (a_1,a_2). The expression for the Cruijff function is:

$$\mathcal{C}(m; m_0, \sigma_1, \sigma_2, a_1, a_2) = \exp\left[\frac{-(m-m_0)^2}{2\sigma_i^2 + a_i(m-m_0)^2}\right] \text{ where } \begin{cases} i = 1 \text{ if } \leq m_0 \\ i = 2 \text{ if } m > m_0 \end{cases} \quad \text{(A.3)}$$

The parameters (which are extracted from [112]) for both channels can be seen on Tables A.6 and A.7.

Parameter	Value
μ_i	$5.2283 \pm 0.0064\text{GeV}/c^2$
σ_{Li}	$0.01505 \pm 0.00738\text{GeV}/c^2$
σ_{Ri}	0.02355 ± 0.00578
a_{Li}	1.33 ± 0.23
a_{Ri}	0.23 ± 0.09

Table A.6: Parameters of the Cruijff function used to model the $B^\pm \to K^\pm K^\mp K^\pm$ cross-feed component.

Parameter	Value
μ_i	$5.3159 \pm 0.0051\text{GeV}/c^2$
σ_{Li}	$0.02011 \pm 0.00345\text{GeV}/c^2$
σ_{Ri}	0.02038 ± 0.01059
a_{Li}	0.14 ± 0.03
a_{Ri}	0.26 ± 0.02

Table A.7: Parameters of the Cruijff function used to model the $B^\pm \to K^\pm \pi^\mp \pi^\pm$ cross-feed component.

A.4.4 Partially reconstructed channels

In the $B^\pm$ fit, two partially reconstructed channels are modeled: $B_S \to$ 4body and $B \to$ 4body. Both channels are modeled in the same way, using a convolution of an Argus shape

and a Gaussian function. The parameters for both channels can be seen on Tables A.8 and A.9.

Parameter	Value
$m0$	$5.24471 \pm 0.00109\text{GeV}/c^2$
slope	-5.005 ± 6.950

Table A.8: Parameters of the convolution function returned by the fitter for the $B_S \rightarrow$ 4body partially reconstructed component.

Parameter	Value
$m0$	$5.12235 \pm 0.03987\text{GeV}/c^2$
slope	-44.48 ± 20.54

Table A.9: Parameters of the convolution function returned by the fitter for the $B \rightarrow$ 4body partially reconstructed component.

A.5 Acceptances for $B^{\pm} \rightarrow K^{\pm}K^{\mp}\pi^{\pm}$

A.5.1 Preselection and reconstruction

Figure A.11 shows the reconstruction and preselection acceptances.

A.5.2 Triggers

See Figure A.12.

A.5.3 BDT

See Figure A.13.

A.5.4 PID

See Figure A.14.

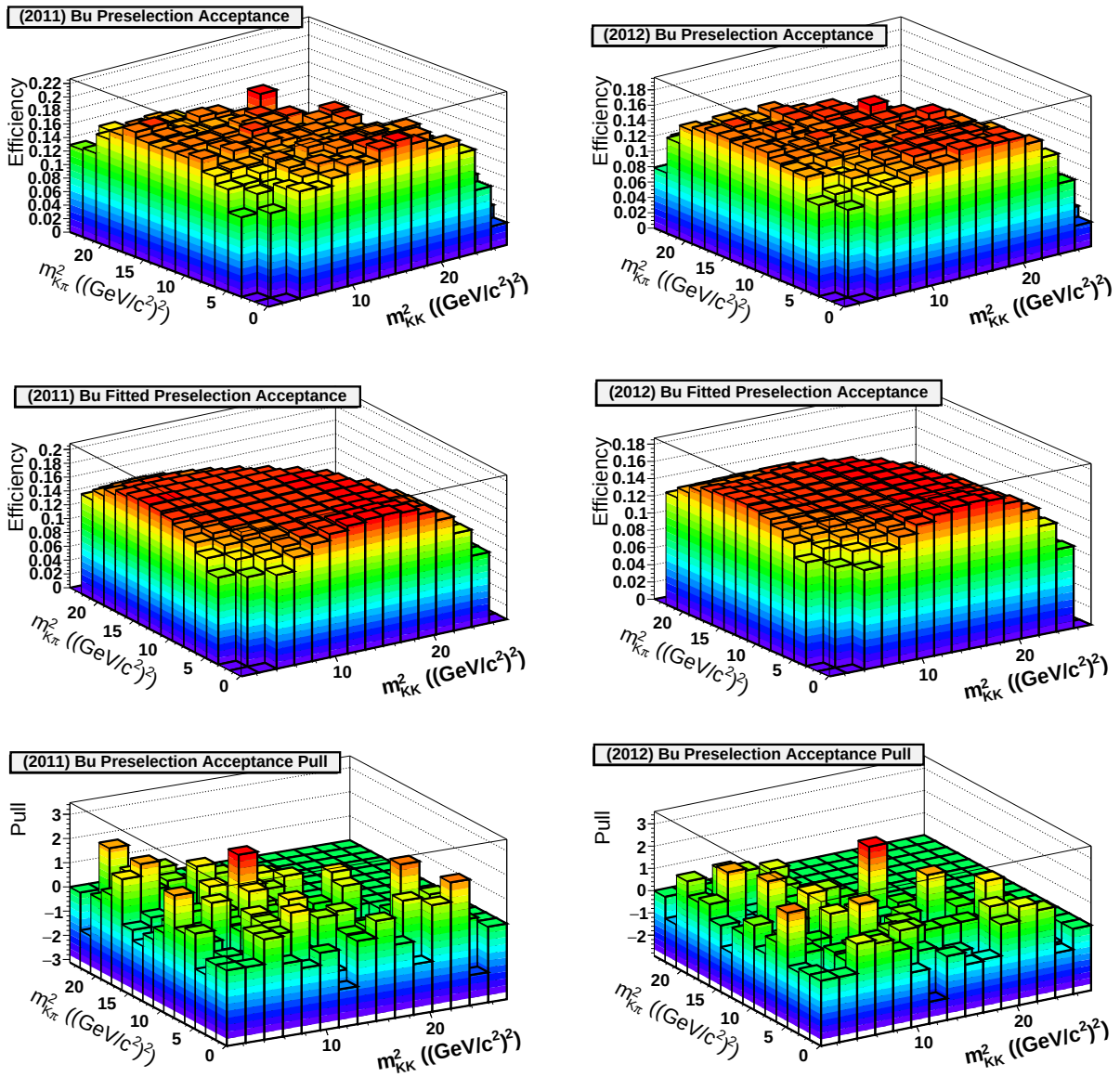

Figure A.11: (Top) Raw, (middle) fitted and (bottom) pull distributions of the preselection acceptances for (left) 2011 and (right) 2012, for $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$.

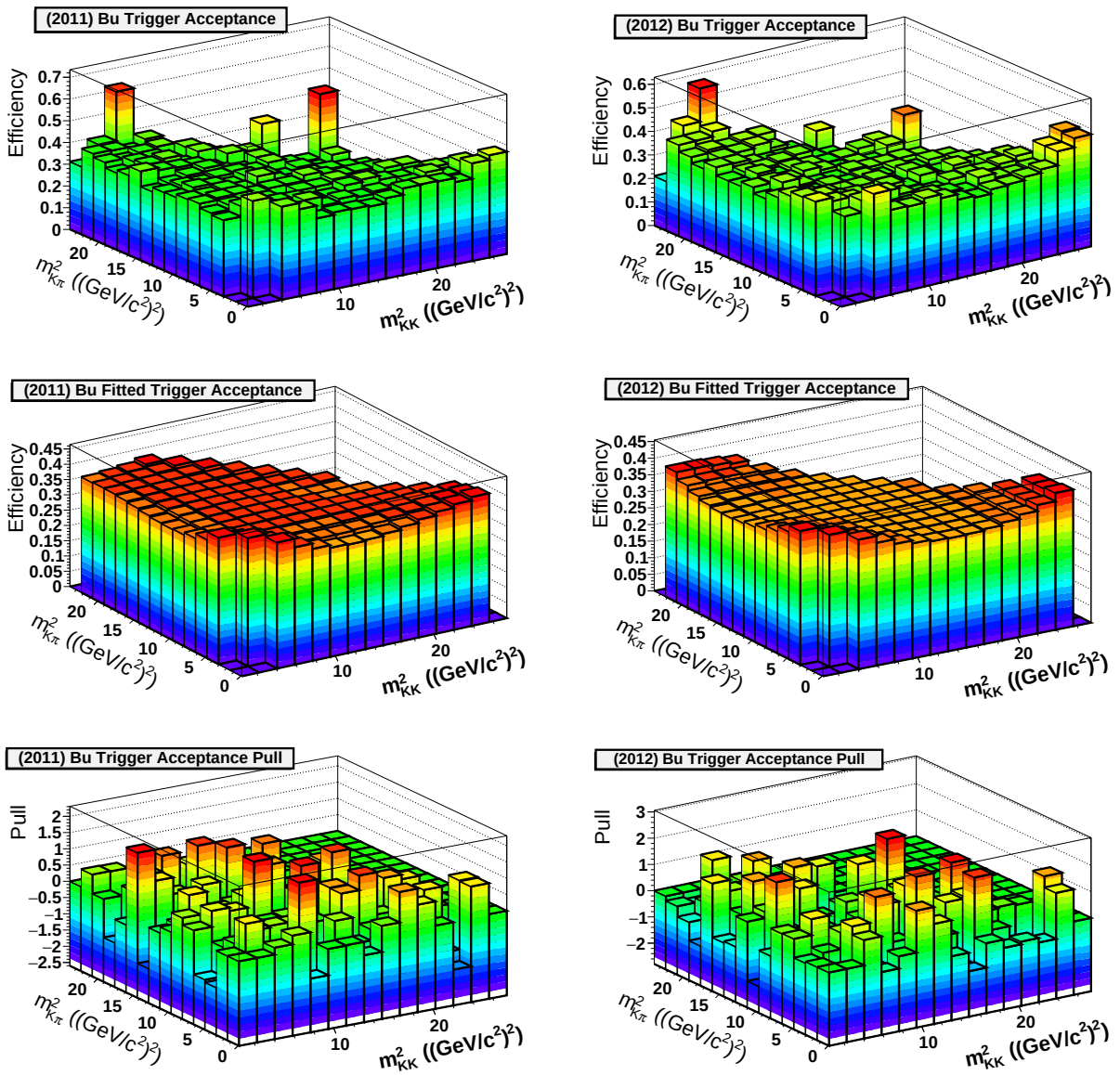

Figure A.12: (Top) Raw, (middle) fitted and (bottom) pull distributions of the trigger acceptances for (left) 2011 and (right) 2012, for $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$.

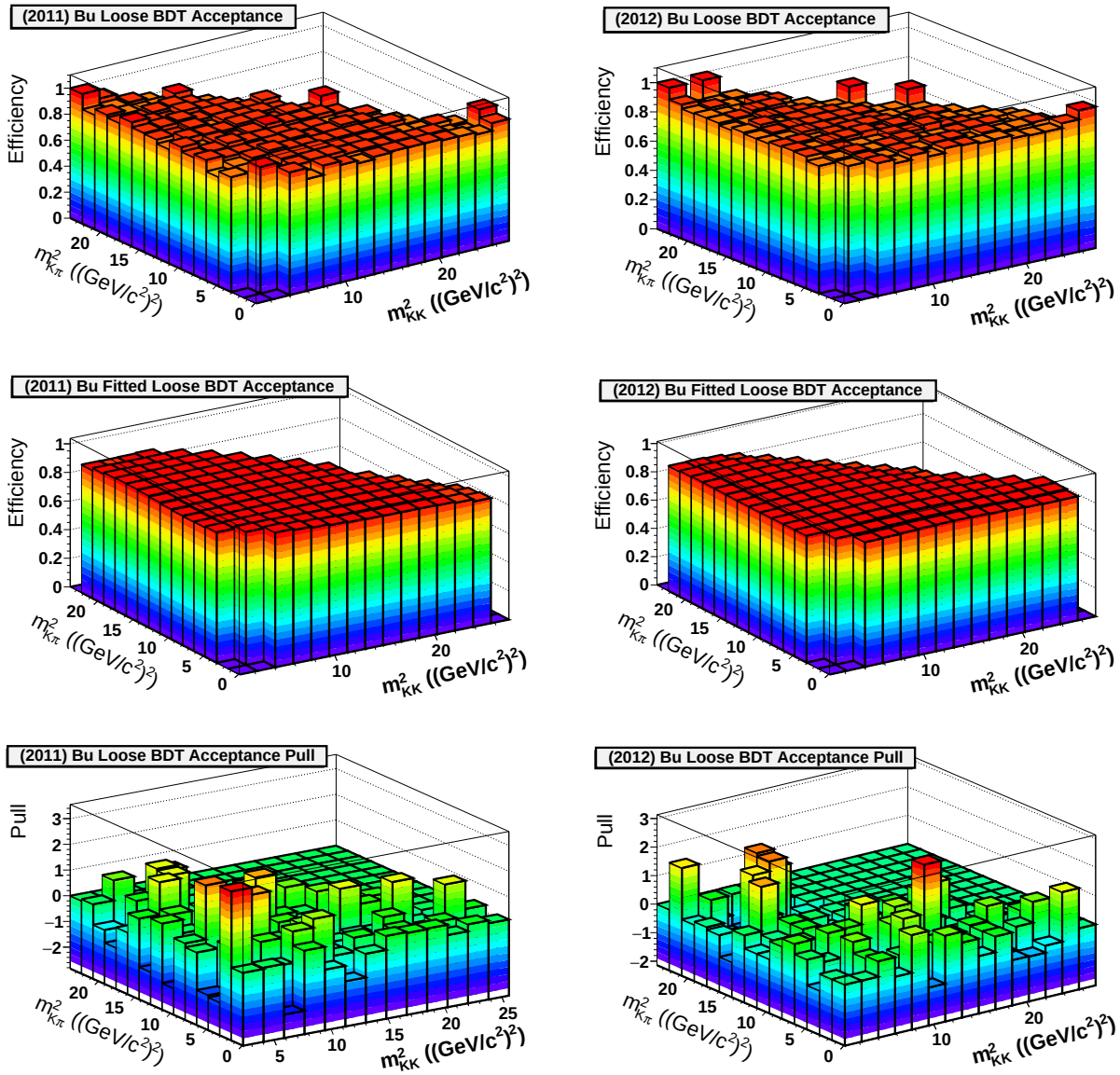

Figure A.13: (Top) Raw, (middle) fitted and (bottom) pull distributions of the loose BDT acceptances for (left) 2011 and (right) 2012, for $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$.

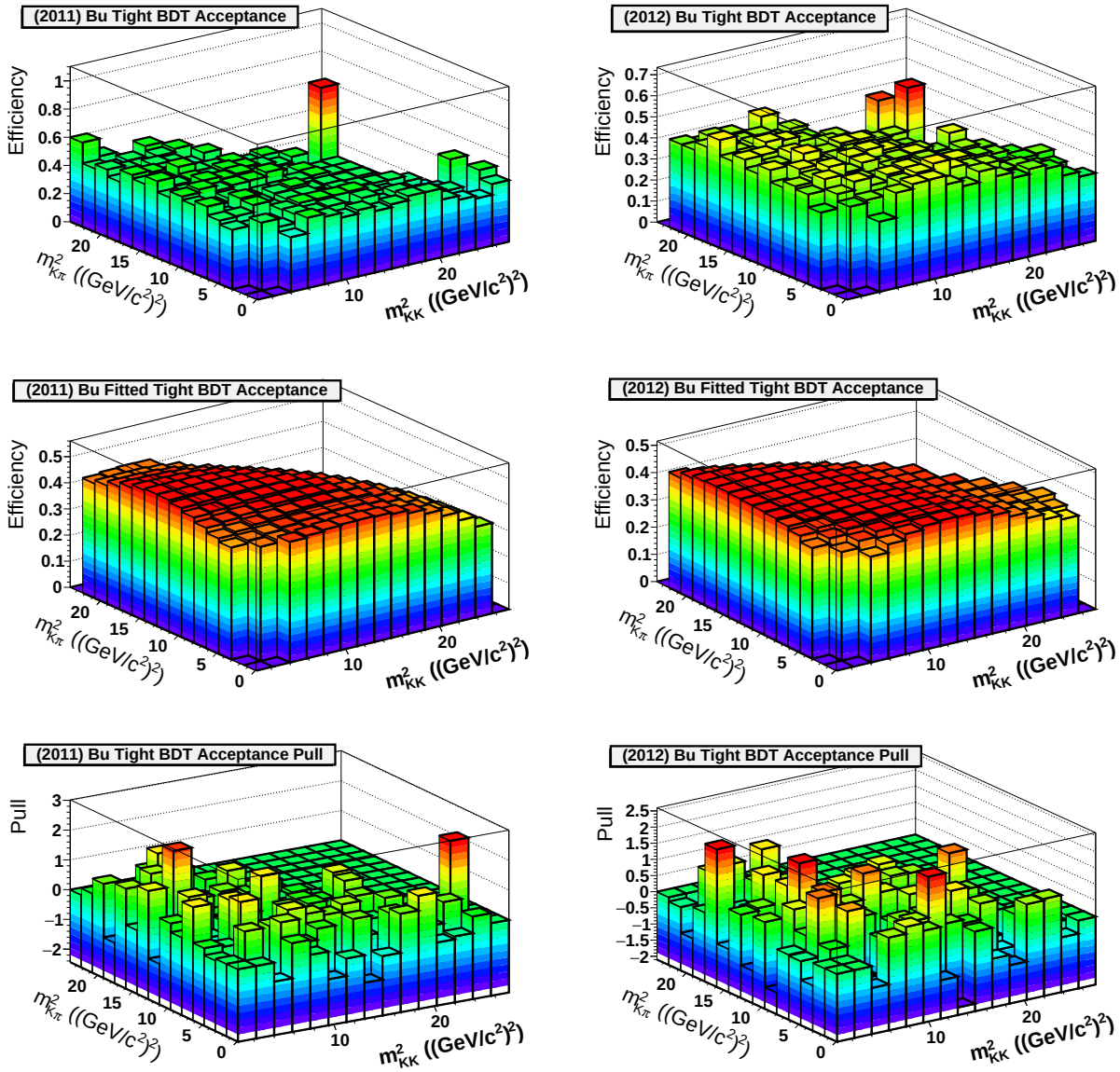

Figure A.14: (Top) Raw, (middle) fitted and (bottom) pull distributions of the PID acceptances for (left) 2011 and (right) 2012, for $B^{\pm} \to K^{\pm}K^{\mp}\pi^{\pm}$.