Department of Physics and Astronomy
University of Heidelberg

Master thesis

in Physics

submitted by

Pallav Mohan
born in Delhi, India

Jan 2021

Analysis of the Decay at LHCb

$\mathbf{B^+ \to \psi(2S)(\to \mu^+\mu^-)\pi^+ k_s^0(\to \pi^+\pi^-)}$

This Master thesis has been carried out by Pallav Mohan
at the
Physikalisches Institut Heidelberg
under the supervision of
Prof. Dr. Ulrich Uwer

Abstract

The thesis aims at measuring the Branching Ratio for the decay channel $B^+ \rightarrow \psi(2S)\pi^+K_s^0$ using $B^+ \rightarrow J/\psi(1S)\pi^+K_s^0$ as reference channel. The analysis has been carried out using Run 1 dataset of the LHCb experiment where 3 fb^{-1} of proton-proton collision data was recorded, 1 fb^{-1} in 2011 at center-of-mass energy 7 TeV and 2 fb^{-1} in 2012 at center-of-mass energy 8 TeV. The Branching ratio was measured at (7.28 $\pm$ 0.01 (stat. + sys.)) $\times 10^{-4}$ where the first uncertainty is statistical, the second is systematic uncertainty. The results are compatible with the previous measurements.

Zusammenfassung

Die Arbeit zielt darauf ab, das Verzweigungsverhältnis für den Zerfallskanal $B^+ \rightarrow \psi(2S)\ pi^+K_s^0$ mit $B^+ \rightarrow J/\psi(1S)\pi^+K_s^0$ als Referenzkanal zu messen. Die Analyse wurde unter Verwendung des Datensatzes von Run 1 des LHCb Experiments durchgeführt, in dem 3 fb^{-1} Proton-Proton-Kollisionsdaten aufgezeichnet wurden, 1 fb^{-1} im Jahr 2011 bei Kollisionsenergie 7 TeV und 2 fb^{-1} im Jahr 2012 bei Kollisionsenergie 8 TeV. Das gemessene Verzweigungsverhältnis beträgt (7.28 $\pm$ 0.01 (stat. + Sys.)) $\times 10^{-4}$, wobei die erste Unsicherheit die statistische die zweite die systematische und die dritte die durch der Unsicherheit des Verzweigungsverhältnisses des Kontrollkanals beschreibt. Die Ergebnisse sind mit den vorherigen Messungen kompatibel.

Contents

List of Figures

Introduction

The analysis deals with the study of two closely related decay channels $B^+ \to \psi(2S)K_s^0$ and $B^+ \to J/\psi(2S)K_s^0$ carried out using the Run 1 dataset of the LHCb experiment. Simulated events (Monte Carlo) were used throughout the thesis as a proxy for signal. The analysis was carried out using tools like sWeighting, re-Weighting and PID calibration tools. Decision tree's were also trained in order to further minimise the combinatorial background populating the data sample using Machine Learning tools like the BDT's (Boosted Decision Tree's). The decision tree's were configured in order to minimise the overtraining.

The Structure of the thesis is as follows: Chapter 1 deals with the theoretical framework of Standard Model, mainly focusing on B-Physics and CP violation and ends with presenting open ended questions in Standard Model. LHCb accelerator complex and its components relevant for the analysis are discussed in detail in Chapter 2. Analysis tools and strategy and preselection technique is described in Chapter 3 along with a dedicated section on the Monte Carlo. In Chapter 4 decision trees are trained and tested using gradient boosting option and the output MVA variable is later used to suppress the combinatorial background surrounding the signal events. In Chapter 5 Branching Ratio is calculated and final concluding remarks are given at the end.

1 The Standard Model

In this chapter we will lay the theoretical underpinnings that are essential for the analysis done in this thesis. One of the main reason of studying the channel, $B^+ \to \psi(2S)\, K^{*+}$ is that it will turn up in other measurements, where a precise knowledge of its Branching ratio (BR) and direct CP asymmetry (A_{CP}^{dir}) could be used to constrain its contribution in the fit. Charged B mesons are the best candidates for direct CP violation as there is no contribution from mixing (cf. section 1.4.1). And these B decays (semi-leptonic) [1] provide tests for both strong and weak interaction as the the decay Amplitude depends on strong and weak phase. Standard Model predicts small differences in BR for B^+ and B^- mesons. Therefore large differences in BR (cf. Equation 1.6), would be evidence of new Physics.

1.1 Introduction to Standard Model

The Standard Model or SM for short, is a culmination of more than 400 years of Physics, providing an elegant description of reality. This vast body of knowledge aims at explaining the universe around us. At its core lie the basic physical laws which describe the fundamental particles and dictates how the matter particles and everything[1] in the universe interacts. As per our current understanding of the Universe, merely 4.6% of it is described by the Standard Model which is also known as the ordinary matter but a better part of the Universe

[1] with the exception of gravity. As its not incorporated into the SM (*cf. section 1.5*).

(95%) is still unaccounted[2] for by the Standard Model constituting of the dark matter (approx. 24%) and dark energy (approx. 71.4%) [2].

Standard Model is a relativistic quantum field theory describing the interaction of quarks and leptons along with the force carriers also known as the gauge bosons. Quarks and leptons collectively known as fermions, have spin 1/2 while bosons carry integer spin. Quarks and Leptons naturally fall into doublets :

$$\begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix}$$

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix}, \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix}, \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}$$

with all the up type quarks carrying +2/3 of the electric charge, e and down type quaks carrying –1/3 of the electric charge. Whereas e, μ and τ carry –1 charge and neutrinos carry no charge. Besides the aforementioned quarks and leptons along with the gauge bosons, scalar Higgs particle (Spin 0) has been recently added to the SM, which is an excitation of the Higgs field. The coupling of the SM particles with this Higgs field is what is responsible for providing mass to the SM particles. All the particles of the SM are listed in Table 1.1. Quarks unlike leptons carry an additional degree of freedom called colour charge. Colours come in 3 flavours – r (red), g (green) and b (blue). Finally the SM has an exact copy of itself but with antiparticles. The antimatter particles of quarks and leptons carry opposite electric charge and colour charge but exact same mass and spin as their counterparts. The force carrying particles are the gauge bosons they are the photon γ for the electromagnetic force, $W^{\pm}$ and Z for the weak force, and 8 gluons (g) which carry the strong force. All fermions come in three generations with each generation having similar properties except for mass, each new generation being heavier than the previous one. For instance τ is 3477 times heavier than e and μ is about 200 times heavier than the electron.

Yang and Mills in 1930 came up with gauge theory for explaining the hadrons, it was before Quantum Chromo Dynamics or QCD was invented. This theory turns out to be central in explaining the SM. Standard Model is a gauge theory associated with the Lie group

[2]For more on open questions related to SM refer to section 1.5.

$$SU(3) \bigotimes SU(2) \bigotimes U(1)$$

and the SM Lagrangian is invariant under the gauge transformation of this group. This invariance results in existence of the gauge bosons or the force carriers with the caveat that these gauge bosons need to be massless. Then the problem arises with the existence of massive gauge bosons in the weak theory i.e. $W^{\pm}$ and Z bosons. This problem was later solved by electroweak symmetry breaking [3] which predicted the existence of yet another massive scalar particle called the Higgs, which was recently discovered at the LHC.

	First generation	**Second generation**	**Third generation**	**Electric charge**
	Fermions			
Quarks	Up (0.0022 GeV)	Charm (1.27 GeV)	Top (172.76 GeV)	2/3
	Down (0.0047 GeV)	Strange (0.093 GeV)	Bottom (4.18 GeV)	–1/3
Leptons	Electron (0.51 MeV)	Muon (105.66 MeV)	Tau (1776.86 MeV)	–1
	Electron neutrino (<1.1 eV)	Muon neutrino (<1.1 eV)	Tau neutrino (<1.1 eV)	0

Gauge Bosons (force carriers)	Strong Force	Gluons (zero mass)	0
	Weak Force	$W^{\pm}$ (80.4 GeV)	± 1
	Weak Force	Z (91.2 GeV)	0
	Electromagnetic Force	Photon (zero mass)	0
Scalar Boson	Higgs (125.18 GeV)		0

Table 1.1: The twelve fundamental fermions (up) divided into quarks and leptons and gauge bosons (down) of the Standard Model. Masses taken from [4].

Quantum Electrodynamics or QED [3, 5], which is the quantum theory of electromagnetism which describes the interaction of photon which is the the force carrier, with the fermions, is a theory associated with the symmetry group U(1). On the other hand weak force which is described by the symmetry group SU(2), is responsible for β–decay, nuclear fusion and decay of mesons and hadrons, $W^{\pm}$ and Z are the force carries for the weak theory. Even tough weak force appears to be weak it is really not, the crucial point is that the weak force appears to be feeble w.r.t electromagnetic force only due to its limited range but not because of its intrinsic strength and thats the case because the force carriers for the weak interaction are massive compared to the massless gauge boson (γ) for QED. It was later discovered that at short enough distances or high enough energies the two forces really have same intrinsic strength and thus are unified under electroweak theory, described by $SU(2)\bigotimes U(1)$ group. Note that the weak interaction only take place for left–handed chiral particles states and right–handed chiral antiparticles states[3] and it is the only place in the Standard Model where partiy is not conserved. QCD on the other hand is the theory describing the interactions of quarks and gluons. It is a SU(3) gauge theory with 8 gluonic fields $G^a_\mu(x)$ where a = 1, ..., 8, which represent the 8 gluons of the QCD. Since the generators of the QCD are 3×3 matrices the wavefunction ψ has an additional degree of freedom called the colour charge [5]. Quarks carry this colour charge which comes in 3 flavours – r (red), b (blue), g (green) and antiquarks carry the opposite colour charge as quarks namely – $\overline{r}$ (antired), $\overline{b}$ (antiblue) and $\overline{g}$ (antigreen). These gluons which are the massless gauge bosons for QCD themselves carry this colour charge and they combine with non zero colour charge particles namely q and $\overline{q}$ to form colour neutral states[4].

Symmetry plays a big role in field theory. Symmetry [2] is defined as a transformation of the system or a function, under which it remains invariant. Groups which are the basic building blocks of symmetry were first identified by Sophus Lie, who came up with groups describing continuous symmetry, today these groups are known as Lie groups. Symmetries can be discrete or continuous, the key principle is that the symmetries dictate the physical laws. Which can be seen using Noether's theorem[5], and when applied to α_{QED}[6],

[3] It is a result of V–A structure of the weak Interaction. For more detailed discussion refer to [5].

[4] As per colour confinement, no coloured objects can propagate freely since the strong force which unlike other forces increases as the separation between two quarks increases which results in massive amount of energy getting stored within the gluonic field between the two quarks. And as the separation increases further the energy within the field is large enough to create a new q$\overline{q}$ pair

[5] Works for continuous symmetries only.

[6] α_{QED} or QED Lagrangian which is given as : $\overline{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}$

results in existence of conserved charge.

1.2 CP Violation in SM

CP Symmetry refers to the fact that the laws of Physics should be invariant if particles are interchanged with antiparticles and the spatial coordinates are inverted. Violation of this symmetry is referred to as CP violation [5] which was first observed in 1964 in neutral Kaon decays. It is the mechanism which explains our matter dominated universe. For most part the idea that there are pockets of the Universe dominated by antimatter to equalise the number of matter particles can be rejected as there would be some astronomical observation verifying huge bursts of radiation at the intersection of matter and antimatter. The dominance of matter over antimatter is believed to have arisen at the very early stages of our universe. During that time universe could be thought of as undergoing the process :

$$\gamma + \gamma \rightleftharpoons p + \overline{p}$$

As the temperature reduced, the rate of forward reaction also dropped and so did the rate of backward reaction as the baryonic and antibaryonic density reduced as a result of inflation. After the "freeze out" calculations done considering no CP violations predict :

$$n_b = n_{\overline{b}} \sim 10^{-18}\, n_\gamma$$

where n_b refers to the number density of baryons and $n_{\overline{b}}$ refers to the number density of antibaryons. But this is in contradiction with the matter dominance seen in the Universe[5] :

$$\frac{n_b - n_{\overline{b}}}{\gamma} \sim 10^{-9}$$

which means that for every 10^9 antibaryons there are $10^9 + 1$ baryons. To explain this Sakharov in 1967 formulated three conditions which must be satisfied in the early universe [6] :

1. Baryon number violation
2. CP violation
3. Departure from thermal equilibrium, since in thermal equilibrium number density would have balanced out

1.2.1 Origin of CP Violation

P in CP refers to Parity transformation. Under parity transformation all the spatial coordinates of a wavefunction $\psi(\overrightarrow{x})$ are inverted such that :

$$x \to -x,\ y \to -y,\ z \to -z,\ \text{and}\ t \to t,$$

The overall Parity [3, 5] of the particle is the product of the particles intrinsic parity and parity under reflection (i.e. parity transformation). The SM Lagrangian for a charged current is given by $\overline{f}\gamma^{\mu}P_L f V_{\mu}$ where $P_L = (1-\gamma^5)/2$ is the left handed chiral projection operator and V_{μ} represents the gauge boson. Action of P_L on the Lagrangian gives two terms, one is a vector term, γ^{μ} and the other term is the axial vector term, $\gamma^{\mu}\gamma^5$ (opposite parity). Had there been only axial vector term or the vector term in the Lagrangian there would have been no Parity violation but since the two terms are not only present but also interfere, it leaves the Weak Interactions[7] with maximally violating the Parity operation.

C in CP refers to charge conjugation. It refers to transforming a particle wavefunction into an antiparticle wavefunction and if a certain process happens with same probability before and after charge conjugation, then it is said to exhibit C symmetry. Weak sector of the Standard Model maximally violates the C symmetry as the weak vertex only couples with left handed fermions and right handed antifermions.

The combined CP would lead a left handed particle into a right handed anti particle. As under CP transformation helicity[8] changes[9] from left to right and vice versa and the momentum vector transforms as, $\overline{p} \to -\overline{p}$. Until 1950's it was thought that even though P and C are individually not conserved but the

[7] Note that Parity is conserved in case of QED and QCD.

[8] It is the normalised component of particle spin in the direction of particle momentum.

[9] for more detailed discussion of this topic refer to chapter 21 of [3]

combined CP would be conserved for the weak interactions. However experiments done in 1960's confired that CP symmetry was broken. Even though the process involving CP and their CP conjugates both occur but their probability of occurance is not the same. These processes differ by a factor of 1 in a 1000 and this small difference in probability of occurance is what is known as CP violation.

CPT is thought to be an exact symmetry of the quantum field theory, where T refers to time reversal $t \to -t$. CP asymmetry implies T asymmetry, which would imply $\langle \psi' | H | \psi \rangle \neq \langle T\psi | THT' | \psi' \rangle$ where H is the Hamiltonian. In other words $THT^{-1} \neq H$ which means that H is not real and CP is violated. Another way to write the charged current interaction is :

$$(u\ c\ t)\gamma^{\mu} P_L V \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

where V is the CKM matrix, more on that refer to section 1.3. This matrix can incorporate the CP violation in its phase angle, η.

Finally, note that the weak interactions in the quark sector isn't the only source of CP violation in SM, in principle strong interactions can also violate CP symmetry but there hasn't been any decisive experimental study showing that. The third source of CP asymmetry can be in the lepton sector which can be explained via PMNS[10] matrix. In this thesis we will be focusing on CP violation only in the quark sector.

1.3 CKM Matrix

The coupling strength of weak interaction with quarks and leptons differ from each other. Furthermore the strength of interaction between ud and us vertex also differs. This was explained by Cabibbo back when only 3 quarks were known to exist. Cabibbo gave a hypothesis which explained this observation and it predicted the existence of charm quark. According to his hypothesis [5, 7] weak eigenstates of weak Hamiltonian differs from the mass eigenstates. Today we know that 3 generations of quarks exist, therefore this idea can be naturally extended to 3 generations of quarks :

[10] for more on this refer to [5]

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.1}$$

where the entries in the left column represent the weak eigenstates and the entries in the right column represent the mass eigenstates and the 3×3 matrix in the middle is the CKM matrix (V_{CKM}) and it contains three Euler rotation angles and 1 phase. Note that in order for V to be complex there needs to be atleast 3 families of quarks any less and the matrix would have real entries.

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \times \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta'} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta'} & 0 & c_{13} \end{pmatrix} \times \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{1.2}$$

where $s_{ij} = \sin\phi_{ij}$ and $c_{ij} = \cos\phi_{ij}$. V_{CKM} is analogous to PMNS matrix in the lepton sector. Relative strength of interaction between different quarks depend on the matrix element associated with the interaction. The "smalleness" of the rotation angle between the weak and mass eigenstates lead to a near diagonal nature of the CKM matrix and being so highly diagonal in nature gives us clues as to why flavour changes outside generations are so suppressed.

Another useful form of this matrix is via Wolfenstein parametrisation, where the matrix is written in terms of 4 parameters λ, A, ρ, η.

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \sim \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} \tag{1.3}$$

For CP to be violated in the quark sector,η, which is the complex phase in the matrix, must be non zero. As can be seen the parameter η appears only at two places in the V_{CKM} and in this thesis we will exclusively focus on B-meson systems to study the CP violating effects.

1.4 B Physics

The decay of b-quark is explained by the weak and the physical mass eigenstates not being equal and being related by a small rotation angle as was explained in the previous section 1.3. This explains the decay of b-quarks to lighter c- and u- quarks[11], outside of its SU(2) doublet. The decay rate [3, 7] depends on the relevant V_{CKM} factor and is proportional to the square of that factor and for b– decay which highly favour c– quark over u– as ($|V_{cb}^2|/|V_{ub}^2| \sim 105$)[12] [8]. Therefore b– decay rate $\Gamma_b \sim V_{cb}^2$ where $V_{cb}^2 \approx 0.0016$, makes the decay width of b-quarks quite small. The small width implies longer lifetime ($\approx 10^{-12}$s) and that makes it easier to tag and study b-quarks. However if a b-quark is produced at the Primary Vertex (PV), it will hadronize and given the long lifetime of b-mesons, they travel a few millimetres in the detector before decaying into the Secondary Vertex (SV). Figure 1.1a refers to the tree level decay of the Signal Channel ($B^+ \to \psi K^{*+}$). There are no Flavor Changing Neutral Current (FCNC) transitions at tree level in SM. These transitions are introduced at one loop level or through penguin processes but are highly suppressed due to the Glashow–Iliopoulos–Maiani (GIM) mechanism.[13] Figure 1.1b indicates one such process for the Signal Channel, these gluonic penguin processes are highly suppressed and only contribute $\mathcal{O}(\sim 0.5\% - 1\%)$ [9]. The interference of the amplitude between the penguin processes and the tree level processes, results in direct CP violation[14] and which can lead to decay rate asymmetries in B^+ and B^- mesons.

CP violation is not restricted to the kaon systems it has also been observed in heavy neutral mesons and in particular $B^0(\bar{b}d) \rightleftharpoons \overline{B}^0(b\bar{d})$ system. Since neutral B-mesons are massive, they have over 400 decay modes but B^0 and $\overline{B}^0$ share relatively common few decay modes which can interfere. This results in explaining the interference between B^0 and $\overline{B}^0$ with a small angle β. Now as previously discussed there is no reason for the mass eigenstates to same as flavour eigenstates [5] as can be seen in equation 1.10.

[11] As they are kinematically allowed.
[12] $V_{cb} \approx 0.041$ and $V_{ub} \approx 0.004$ [5].
[13] It ensures flavor conservation in tree level processes.
[14] more on that in section 1.4.1.

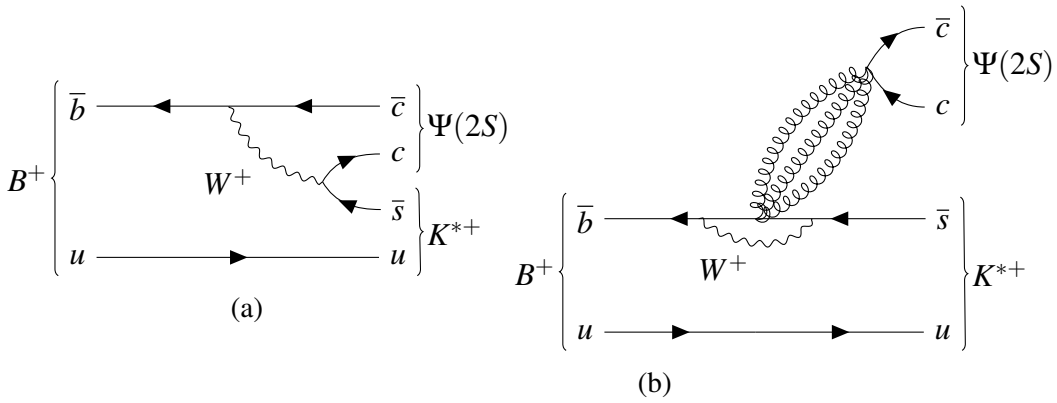


Figure 1.1: Decay of $B^+ \to \psi(2S)(\to \mu^+\mu^-) + K^{*+}(\to K_s\pi^+)$ with contribution from (a) tree level decay (dominant) and (b) Penguin decay (suppressed).

1.4.1 CP Violation in B-meson system

CPV can be primarily observed in 2 distinct ways [5, 7] :

1. Direct CP violation : if the decay width of $B^0 \to f$ is not the same as for $\overline{B}^0$ to its CP conjugate $\overline{f}$ then CP is said to be violated directly.

$$\Gamma(B^0 \to f) \neq \Gamma(\overline{B}^0 \to \overline{f})$$

2. Indirect CP violation : can occur either through mixing or through interference due to oscillation. Mixing occurs because the CP flavour eigenstates are not the same as physical mass eigenstates but are related by a small angle β. Incase of interference the events in the decay $B^0 \to f$ can be populated by oscillation terms $B \to \overline{B}^0 \to f$.

In this discussion we will restrict to only B-meson mixing.

Direct CP Violation

Direct CP violation in Kaon systems is calculated by measuring the quantity[15] $\mathrm{Re}\left\{\varepsilon'/\varepsilon\right\}$. Charged B mesons [7] on the other hand can only violate CP di-

[15] $\mathrm{Re}\left\{\varepsilon'/\varepsilon\right\} = (1.65 \pm 0.26) \times 10^{-3}$ [5] It's a small effect relative to $K^0 \leftrightarrow \overline{K^0}$ mixing which is the dominant source of CP violation in Kaon systems.

rectly as there is no contribution from mixing. As previously discussed in section 1.4 the interference in the decay Amplitude is caused by contributions from tree level decay and penguin processes which can cause CP asymmetries[16].

Incase of B^+ and B^- mesons following relation holds between their decay amplitude :

$$\frac{|\overline{A}|}{A} \neq 1$$

For B^+ decay ($B^+ \to f$), the decay Amplitude consists of strong (ϕ) and weak phase (δ). Making the total decay amplitude to be :

$$A_f = |A_1|e^{\phi_1+\delta_1} + |A_2|e^{\phi_2+\delta_2} \tag{1.4}$$

Similarly for the B^- decay the total decay Amplitude is given by :

$$\overline{A_f} = |A_1|e^{-\phi_1+\delta_1} + |A_2|e^{-\phi_2+\delta_2} \tag{1.5}$$

Given that the weak phase is CP odd, it changes sign in equation 1.5 whereas the strong phase which is CP even, does not. The direct CP asymmetry for the charged B mesons is measured by finding the yield of A_{CP}^{dir} ,which is given as :

$$A_{CP}^{dir} = \frac{\Gamma(B^- \to f^-) - \Gamma(B^+ \to f^+)}{\Gamma(B^- \to f^-) + \Gamma(B^+ \to f^+)} \tag{1.6}$$

Measured charged asymmetries for the decay Control channel and Signal channel are : $A_{CP}^{dir} = 0.048 \pm 0.029\,(\text{stat.}) \pm 0.016\,(\text{sys.})$ and $A_{CP}^{dir} = -0.077 \pm 0.207\,(\text{stat.}) \pm 0.051\,(\text{sys.})$ respectively [1].

[16] $\Gamma(B^+ \to f^+) \neq \Gamma(B^- \to f^-)$

***Addendum: B-Meson Mixing**

In this treatment[17] of B meson mixing in $B^0 - \overline{B}^0$ system we assume the mixing probabilities to be the same as in:

$$P(B^0 \to \overline{B}^0) = P(\overline{B}^0 \to B^0) \tag{1.7}$$

The physical states of the $B^0 - \overline{B}^0$ system are the eigenstates of the Hamiltonian of the system and given that $B^0 - \overline{B}^0$ system share very few common decay modes, the interference between the decay can be neglected leaving us with off diagonal decay terms in the Hamiltonian to be zero i.e. $\Gamma_{12} = \Gamma_{21}^* \approx 0$. The Hamiltonian of the system can be written as :

$$H \approx \begin{pmatrix} M - \frac{1}{2}\Gamma & M_{12} \\ M_{12}^* & M - \frac{1}{2}\Gamma \end{pmatrix} \tag{1.8}$$

where the term M_{12} is due to the box diagram and the corresponding physical eigenstates associated with the Hamiltonian are :

$$|B_L\rangle = \frac{1}{\sqrt{1+\|\xi\|^2}}\left(\left|B^0\right\rangle + \xi\left|\overline{B}^0\right\rangle\right) \text{and} \, |B_H\rangle = \frac{1}{\sqrt{1+\|\xi\|^2}}\left(\left|B^0\right\rangle - \xi\left|\overline{B}^0\right\rangle\right) \tag{1.9}$$

where B_L and B_H are the light and heavy states respectively and where ξ is given by :

$$\xi = \left(\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}\right)^{\frac{1}{2}} \approx \frac{M_{12}^*}{|M_{12}|}$$

Therefore equation 1.9 can be written as :

$$|B_L\rangle = \frac{1}{\sqrt{2}}\left(\left|B^0\right\rangle + e^{-i2\beta}\left|\overline{B}^0\right\rangle\right) \text{ and } |B_H\rangle = \frac{1}{\sqrt{2}}\left(\left|B^0\right\rangle - e^{-i2\beta}\left|\overline{B}^0\right\rangle\right) \tag{1.10}$$

[17] for more detailed discussion of this topic refer to section 14.6.1 of [5] and section 9.6.1 of [7]

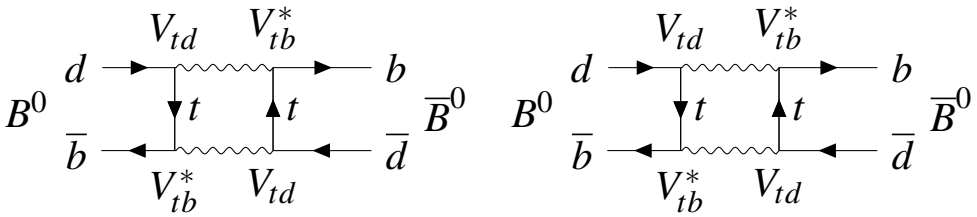


Figure 1.2: Dominant box diagram for $B^0 \leftrightarrow \overline{B}^0$ *mixing* [5].

Note that the main contributors to the mixing in the box diagram (Figure 1.2) is by virtual t-quarks.

The flavour asymmetry after time t is given by :

$$A(t) = \frac{P(B^0 \to B^0) - P(B^0 \to \overline{B}^0)}{P(B^0 \to B^0) + (B^0 \to \overline{B}^0)} = \frac{\cos(\Delta m t)}{\cosh(\frac{\Delta \Gamma t}{2})} \tag{1.11}$$

where Δm is given by :

$$\Delta m = m(B_H) - m(B_L) \propto |(V_{td} V_{tb}^*)^2| \tag{1.12}$$

CP violation in B-meson mixing is a very small effect in $B^0 - \overline{B}^0$ system $\mathcal{O}(10^{-4})$. Since $V_{tb} \sim 1$ equation 1.12 implies that mass difference of heavy and light physical eigenstates $\sim \|V_{td}^2\|$. Therefore the measurement of Δm provides a way to determine $\|V_{td}\|$.

1.5 Open Questions in SM

There are quite a few unanswered questions in Standard Model which hints towards new Physics, some of which are listed below [2, 5, 10]:

1. Gravity is a force which is completely neglected in the Standard Model because of its feeble strength compared to other forces in particle interactions. Note that the other three coupling constants become nearly equal as we probe at higher and higher energies and at length scales $\mathcal{O}(10^{-32})$

m the so called grand-unified length. Unification of gravity with other forces is still an open problem in theoretical physics.

2. Why does SM have 3 generations of fermions, it is something that is not explained by SM along with the almost diagonal nature of the CKM matrix. The complex phase η of the CKM matrix and the PMNS matrix account for the CP violating effects in the quark and lepton sector respectively. The CP violating effects in lepton sector are yet to be explored. Additionally the amount of observed CP asymmetry in the universe is much larger compared to what is accounted for in the Standard Model.

3. Neutrino oscillations hint towards massive neutrinos and the fact that right-handed neutrinos are sterile and dont participate in SM interactions could form potential dark matter candidates.

4. Dark Matter which is said to be formed of weakly interacting massive particle (or WIMP) is yet to be observed directly or indirectly in experiments. These WIMP's are massive particles only interacting weakly. The elusive nature of WIMP's make them really hard to detect.

5. Experiments in QCD indicates no CP violating effets even though there is no reason for CP to be conserved in QCD. This is known as the Strong CP problem.

6. Hierarchy Problem : simply stated refers to the huge disparity between the weak length scale $\mathcal{O}(10^{-18})$ m and the planck length scale $\mathcal{O}(10^{-35})$ m. According to QFT the tendency of virtual particles is to make the weak length as short as possible $\sim$ planck length but the two length scales differ by a factor $\mathcal{O}(10^{17})$. This problem provides us with hints for new Physics beyond SM.

2 The LHCb experiment

This chapter we will lay the groundwork for the LHCb experiment and the key detector components and features that are relevant for the analysis presented in the thesis.

2.1 The Large Hadron Collider

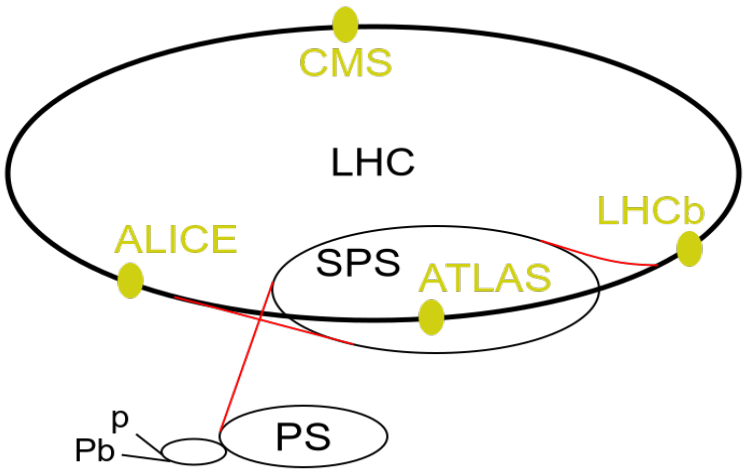


Figure 2.1: The LHC accelerator complex [11]

The LHC or the Large Hadron Collider [12] is a 26.7 Km long Synchrotron located in a underground tunnel near Geneva, Switzerland (see Figure 2.1). It

is a machine of many superlatives with technological complexity pushed to the extreme [2] and it happens to be the world's most powerful particle accelerator. Its main aim is to accelerate particles hurling in opposite direction at nearly the speed to light, to make them collide and study the remnants of the debris which resulted from the collision of the particles. The physics goal at LHC is to explore the uncharted territory of Physics beyond the Standard Model and in turn answer some of the open questions in Particle Physics (*cf.* section 1.5). LHC consists of 4 large detectors (ATLAS, CMS, ALICE and LHCb) and 2 small ones (TOTEM and LHCf), which record the particles traces (a.k.a tracks) and other useful parameters to evaluate the nature of the particle. Due to the High luminosity at the LHC, enormous amount of data gets generated and through the use of sophisticated trigger system (*cf.* section 2.6) which works like a multi-layered 'sieve' only the interesting candidates make it through to be looked at in the offline data analysis.

During the Run 1 (2010-2012) [12] , in the proton mode, centre of mass energy of 8 TeV was achieved. Before the proton beam can be injected into the main LHC ring, the beam is pre-accelerated to 450 GeV by Proton Synchrotron Booster, the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS). The proton beam consists of 2800 bunches with the bunch-to-bunch separation of 25 ns which equals the crossing frequency of 40 MHz. During normal run conditions, beams do not collide head on but they glance at a small angle $\sim 150 - 200\ \mu rad$ in order to avoid unwanted collisions. Note that the LHC dipole magnets are responsible for maintaining the circular trajectory of the beam in the beam pipe [2]. The 4 T–m dipole in the LHCb spectrometer gets its polarity reversed every few weeks during the run time in order to reduce the systematic uncertainties which can occur due to left right asymmetry of the detector. For the LHCb, the experiments are planned at nominal luminosity of $2 \times 10^{32} cm^{-2} s^{-1}$ which equals 10^{12} b-events in one year of run time. Also note that the integrated at LHCb was steadily increased from 1.0 fb^{-1} in 2011 to 2.0 fb^{-1} in 2012.

Following the Run 1 [10] there was a long shutdown period (LS1) which lasted for 2 years and during that time LHC was prepared to reach higher luminosity and center of mass energy. During Run 2 (2015-2018) the center of mass energy of 13 TeV was attained. Integrated luminosity at LHCb was increased again during Run2 from 1.6 fb^{-1} in 2016 to 1.7 fb^{-1} in 2017 to finally 2.1 fb^{-1} in 2018. After Run 2 there was another shut down period and LHC is prepared to reach its maximum designed luminosity and centre of mass energy of 14 TeV.

2.2 LHCb experiment

LHCb or Large Hadron Collider beauty experiment [2, 13] is one the 4 major experiments at LHC, dedicated to the study of heavy b- and c- quarks and their CP violating effects. Unlike ATLAS and CMS it works as a forward spectrometer pertaining to the fact that heavy quarks in proton-proton collisions are mainly produced in the forward region. The layout of the detector geometry is shown in Figure 2.2. Using the right-hand coordinate system, the z-axis points towards the beam axis and the y-axis point upwards. Due to high pseudorapidity η[1] at the LHCb experiment, the beam pipe structure is rather delicate as the collision debris of interest is measured at large pseudorapidity range, $1.6 < \eta < 4.9$ [14].

CP violation in the quark sector does not account for the overall observed CP asymmetry that we see in the universe. There are suggestions that CPV (*cf.* section 1.2.1) in the lepton sector during the early evolution of the universe could have been responsible for the observed CP asymmetry [5]. Future experiments at Belle II and LHCb will be conducted at even higher precision to probe the CP violation in the quark sector. It is quite likely that the yet undiscovered, beyond the Standard Model phenomena are responsible for the observed asymmetry. The study of Λ_b^0 done at LHCb confirmed the existence of Pentaquarks. The detector at LHCb also measures, the production cross section, spectroscopic measurements and study of heavy flavoured b- and c-quarks.

The displacement of the B decay vertex from the primary vertex helps tag b-events [7]. However due to large boost in the forward direction, reconstructing the decay trajectory back to the primary vertex can be very challenging especially if there are more than 1 primary vertices and it is one of the reason for keeping the luminosity at LHCb relatively low compared to other experiments at LHC. Precise measurements of Primary and decay vertices, excellent momentum resolution of the particles and techniques to do particle identification (PID) are the tools needed to study the decay of heavy flavoured hadrons, the implementation of these methods are discussed in upcoming sections.

[1] $\eta = -\ln\tan\theta/2$ where θ is the spherical angle.

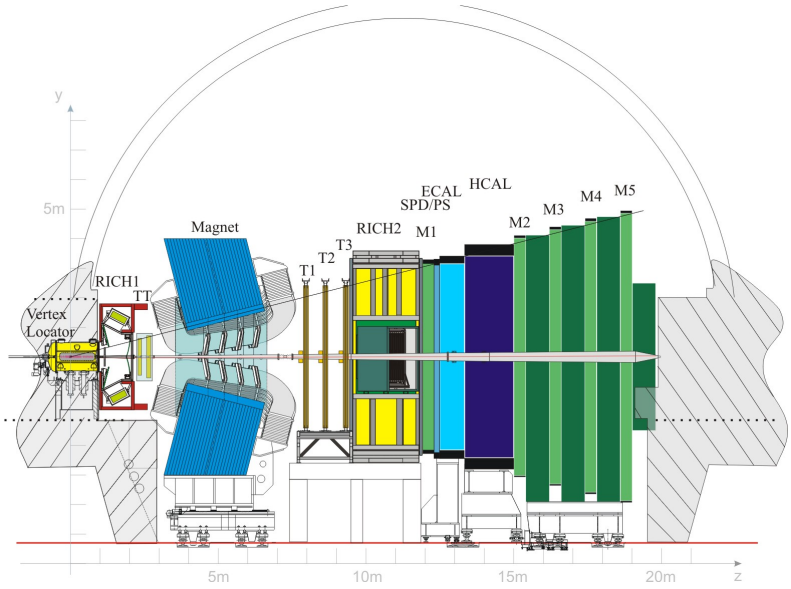


Figure 2.2: View of LHCb and its detector components [14]

2.3 RICH Detector

Ring-Imaging Cherenkov or RICH, detector forms a crucial component of the LHCb detector. Its primary role consists of charged hadron identification of pions (π), proton (p) and kaons (K) [15], with pion being the most abundant charged particle produced in the hadronic interactions. Charged hadron identification helps in reduction of combinatorial background and is also helpful in flavour tagging and thus is crucial for flavour physics experiments.

RICH detectors work on Cherenkov Principle, which states that when a charged particle traverses through a dielectric medium of refractive index η [5] and if the velocity of that particle is greater than speed of light in that medium, $v > c/\eta$ then in doing so, the particles of the medium get ionised and in the process of returning to the state of equilibrium, the medium particles emit Cherenkov photons and this photon signal is read by Hybrid Photon Detector (HPD). The identification of the charged particle is achieved by the measurement of the Cherenkov angle, θ_c which is given by :

$$\cos\theta_c = c/\eta v$$

There exists a very strong inverse correlation [15] between particle momentum and the polar angle, at which they are produced. With large momentum

particles being mainly produced at low polar angle and that is the reason why LHCb has 2 RICH detectors, RICH 1 and RICH 2 as can be seen in Figure 2.2. RICH 1 is located right after the Vertex Locator (VELO) (*cf.* section 2.4) around the interaction point and covers low momentum range (2-40 GeV/c) with angular acceptance of 25-300 mrad. RICH 2 on the other hand is located downstream of the magnet and covers high momentum tracks which are not affected by the magnetic field (15-100 GeV/c) and has an angular acceptance of 15-120 mrad [14]. For more on RICH detectors refer to [16].

2.4 Vertexing & Tracking

Vertexing [10, 17] refers to the reconstruction of the particle interaction vertex which constitute of primary interaction vertex and the secondary decay vertex. While tracking on the other hand, refers to the reconstruction of the trajectory traversed by a particle inside a detector. The tracking system consists of 4 tracking station, Tracker Turicensis (TT) which is situated upstream of the magnet and 3 tracking stations (T1-T3) situated downstream of the magnet and a Vertex Locator (VELO) which is situated right next to the interaction region. TT consists of silicon microstrip detectors. Now the outer part of three stations (T1-T3) away from the beam pipe is the Outer Tracker (OT) which consists of straw-tubes while the inner region of (T1-T3) stations is called the Inner Tracker (IT) which also consists of silicon microstrip detector. Note that the magnetic field is directed along the y-axis such that particles get deflected along the x-axis. To overcome the detector induced asymmetries the polarity of the magnetic field is periodically switched.

The primary task of VELO [18, 19] is the reconstruction of the primary vertex (PV), secondary vertex and the track reconstruction. Vertex reconstruction is of fundamental importance to the LHCb experiment. As b- and c- decay vertices are displaced from the primary interaction vertex which makes them easier to tag. VELO (see Figure 2.3) also provides precise track coordinates of b- and c- hadrons which help in reconstructing the interaction vertices. This information is then used in Flight distance measurements and computation of the decay time of the particles, which are required for the measurements of the lifetime and time dependent measurements in the oscillating $B_s^0 - \bar{B}_s^0$ system. It is also possible to extract information about the Impact Parameter (IP) of these particles which can also be used for flavour tagging. The measurements from the VELO are also fed to the High Level Trigger (HLT, see section 2.6) which reduces the event rate from 1MHz to a few kHz.

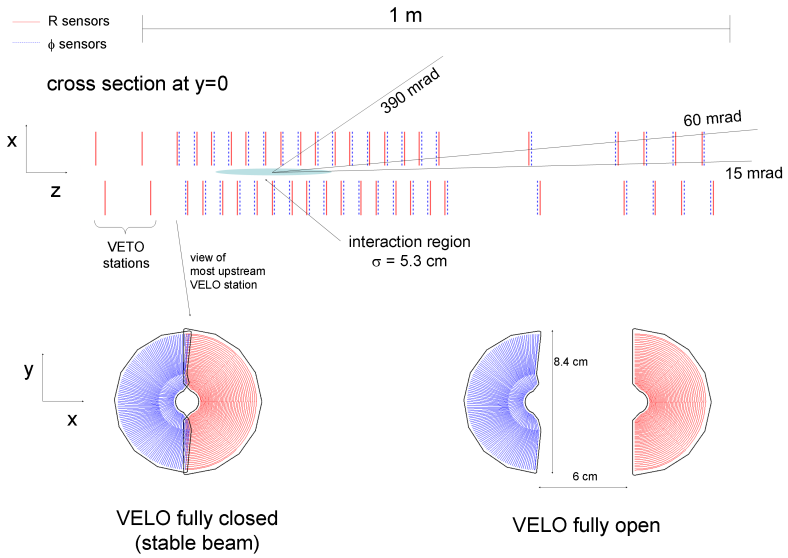


Figure 2.3: Cross section in the (x,z) plane of the VELO silicon sensors at y=0, with the detector in the fully closed position. The front face of the first module is also illustrated in both the closed and open positions [14].

Due to the large boost, b- vertices are displaced from the PV which makes the task of reconstructing them quite challenging especially if the number of PV's is large[2][7]. These contraints demand high cluster efficiency from the VELO for efficient pattern recognition and fake track rejection [18]. Velo [14] has to detect particles in the pseudorapidity range $1.6 < \eta < 4.9$ and in order to accomplish fast and efficient reconstruction of particle tracks and vertices, R and Φ sensors are employed. VELO consists of several silicon modules (21 stations) stacked along the beam pipe with only a separation $\sim$ 7mm from the beam. Each module contains R and Φ sensor (see Figure 2.3). The Φ sensor provides information about the azimuthal angle around the beam and the R sensor gives information about the radial distance r, from the beam axis. The information about the third coordinate is extracted from the position of each sensor along the detector. Due to the close proximity with the beam pipe, the VELO sensors undergo harsh radiation damage which shorten's their lifetime to 3 years [19] after which they need to be replaced.

Tracker Turicensis (TT) [20] which is located between VELO and the dipole magnet plays a crucial role in track reconstruction along with the three outer stations (T1-T3) forming the outer part of the tracking system, consisting of

[2]It is one of the reasons of keeping luminosity at LHCb low compared to ATLAS and CMS.

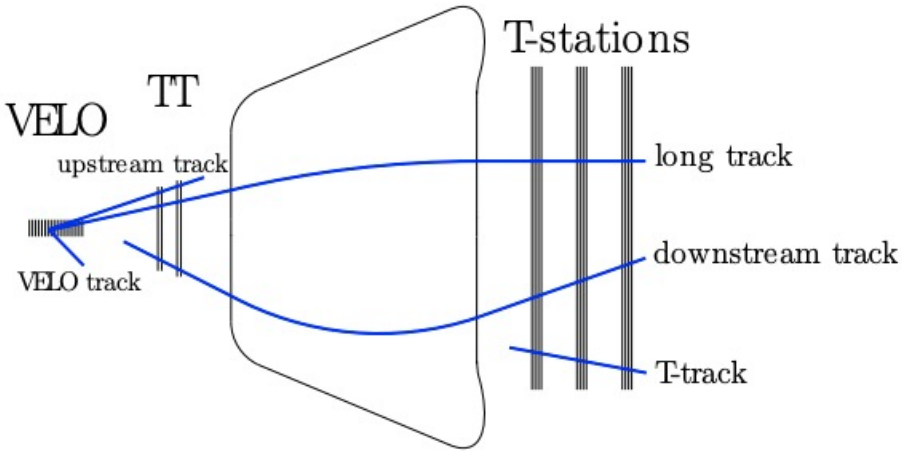


Figure 2.4: Different Type of Tracks within LHCb [20].

straw tubes. Particles traversing through these stations ionise the gas within the straw-tube which generates a current which then travels to the center of the tube (Anode) and then the drift time of these electrons is used as a gauge to measure the distance of the particle trajectory to the wire. With the help of reconstruction algorithms and hits in the all 4 stations (TT, T1–T3), 3 types of tracks can be identified (see Figure 2.4). Long tracks which traverse the complete tracking system are the most important tracks as their momenta are very precisely known. Upstream tracks which are low momenta tracks and also have low momentum resolution, only make it through to the TT station. These upstream tracks are used as background in RICH PID algorithms. Downstream tracks are of long lived particles which make it through TT station and T Stations (T1–T3). In the decay for example: $B^+ \to \Psi(2S)\pi^+ K_s$, K_s decays outside the VELO acceptance.

2.5 Particle Identification

Particle identification (PID) [14] is used in charged hadron identification, flavour tagging, reduction of combinatorial background (by suppressing a similar looking decay product) and trigger system[3]. RICH 1 which contains flurobutane

[3]by reducing 40 MHz event rate down to 2 KHz, which can then be written to tape.

(C_4F_{10}) and aerogel as radiator, provides PID for particles that lie within its angular acceptance and momentum range (2-40 GeV/c). While RICH 2 which is located downstream of the Tracking Station (T3), contains CF_4 gas as radiator and provides PID for particles with momentum in the range 15-100 GeV/c.

Muon system plays a vital role in the LHCb trigger[4] system providing information for high p_T muon trigger (Level–0) and muon identification for HLT (High Level Trigger, *cf.* section 2.6). Muon system consists of 5 rectangular stations (M1–M5) placed perpendicular to the beam axis comprising of 1380 chambers and covering an area of 435 m^2. With M1 station situated upstream of the calorimeters, which helps improve p_T measurements in the trigger. While the stations (M2–M5) are located downstream of the calorimeters. The downstream muon stations (M2–M5) are coated with thick iron absorbers ($\sim 80cm$) to filter out high energy muons. The minimum momentum threshold for muon to cross through all 5 stations is around $\sim 6GeV/c$. Additionally stations M1–M3 have high spatial resolution with main function to determine muon p_T and muon tracks while stations M4–M5 have a low spatial resolution with their main purpose to identify highly penetrating muons.

Along side the muon stations is the Calorimeter system consisting primarily of Electromagnetic CALorimeter (ECAL) and Hadronic CALorimeter (HCAL) along with auxiliary units like the PS/SPD detectors. SPD or Scintillating Pad Detector is used in detection of charged particles and is placed upstream of the PS (Preshower) detector. PS detectors are used for the detection of electromagnetic showers which is followed by HCAL, which in particular, consists of iron and scintillating tiles that work as absorber and active material respectively. Not only are the calorimeters useful in the PID of electrons, photons and hadrons along with the measurement of their energy and position but its output is also fed to the first level trigger (L0).

Information from all sub detectors namely two RICH detectors, ECAL, HCAL and muon systems is combined to detect particles of type (e, μ, π, K and p). Photons and neutral pions (π^0) are detected in the ECAL since π^0 decays into two photons ($\pi 0 \rightarrow \gamma\gamma$) and this can be seen as two distinct electromagnetic showers in the ECAL. PID for RICH consists of using log likelihood estimation[5] for a given track to have produced a hit on the RICH photodetector. The pixel track combination is matched with the cherenkov radiation and the likelihood is compared by varying the particle hypothesis of each track in turn (e, μ, π, K and p). For a given track to be considered as a muon, it must register

[4] trigger is discussed in detail in section 2.6

[5] the algorithm starts by assuming all particles to be pions, since pions are the most abundant particles produced in the pp collisions.

hits in atleast 2-4 muon stations depending on momentum. For further enhancing the selection purity of muon tracks, the log-likelihood difference between muon and pion hypothesis is determined coupled with information from RICH and calorimeter system.

2.6 Trigger System

Given that LHCb operates at lower luminosity compared to other detectors (as was discussed in previous section). The number of visible[6] interactions in the LHCb spectrometer are $\sim$ 10 MHz. The LHCb trigger system [14] then reduces this rate to $\sim$ 2 KHz. The trigger system at LHCb is fed information from TT and T1–T3 stations, calorimeter system and muon stations along with SPD/PS and the pile up system. The trigger system comprises of 2 levels, first is the L0 or first level trigger, which is a hardware level trigger and it takes information from the Pile–up system[7], Calorimeter Triggers[8] and Muon Trigger[9] line (see Figure 2.5) and this information is then fed to the DU (Decision Unit) which then collects all the information and makes a final decision. The output rate of L0 is limited to 1MHz.

The general decay features [10] of weakly decaying heavy mesons exhibit similar decay topology for instance as shown in Figure 2.6. An easy method to find displaced Secondary Vertex (SV) is to search for tracks with large IP (impact parameter) and refers to the IP of the f_1 w.r.t PV as shown again in figure 2.6. Tracks coming from displaced SV will have large IP compared to tracks that are produced in primary pp interactions. Due to the large mass of b– hadrons, the decay products of b– hadron ends up with large transverse[10] momentum (p_T) and energy (E_T). Therefore the L0 trigger line consists of :

- Large energy (E_T) in calorimeters for hadrons, electrons and photons.
- And, a large p_T for the muons in the muon chamber.

[6] An interaction that produces 2 charged particles, combined with the hits in the VELO and (T1–T3) stations are deemed as visible interactions.

[7] The Pile–up system in the VELO computes the number of primary interactions per bunch crossing.

[8] It calculates the number of tracks by computing the energy stored in the calorimeters and the number of hits in SPD.

[9] Looks for two muons with largest p_T

[10] Most of the tracks generated at the PV from the pp collisions come from strong interactions, which do not have large p_T in the resulting particles.

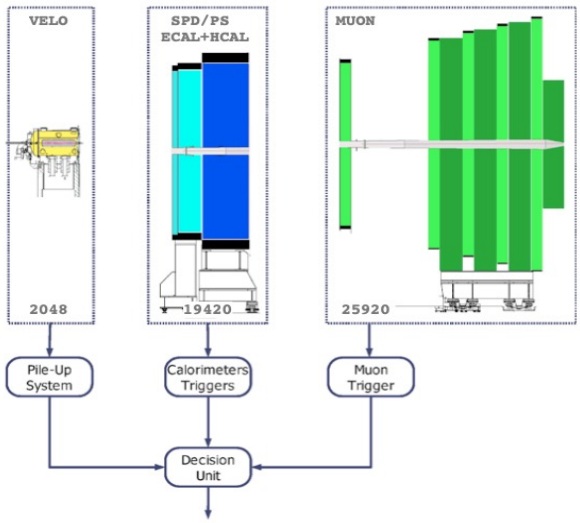


Figure 2.5: Overview of L0 Trigger. Pile–up system receives 2048 channels from the VELO, the Calorimeter Triggers are fed 19420 channels from all of the sub-detectors in the calorimeter system and Muon trigger is fed 25920 channels from the Muon chamber. All these channels are finally processed by the DU to make a final decsion [14].

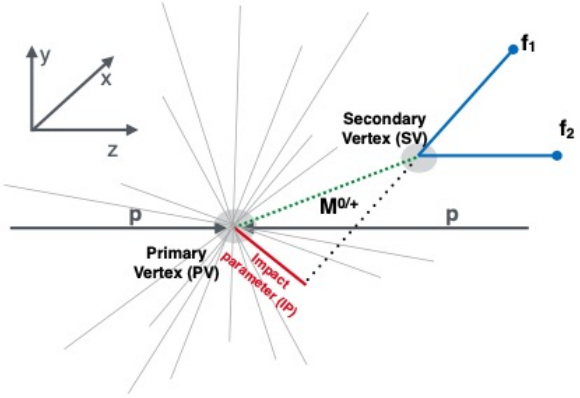


Figure 2.6: Generic decay of a $M^{0/+}$ hadron containing a heavy quark [10].

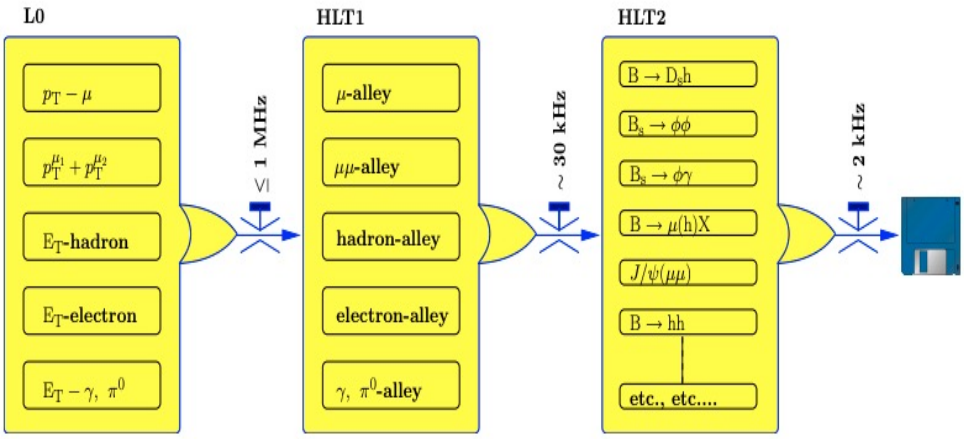


Figure 2.7: Flow diagram of the trigger sequence [14].

The particles are identified as electrons, photons or hadrons based on the information coming from SPD, PS and the calorimeter system. All this information once fed to L–0 DU, it makes a decsion of whether to keep a certain event or to discard it. This completes the Level–0 trigger.

Events passing the L0 are send to the second level of the trigger system called the High Level Trigger (HLT). It is fully software based, consisting of C++ applications that run on Event Filter Farm (EFF) which consists of approximately 2000 computing nodes. HLT reduces the rate of accepted events down to 2 KHz (see Figure 2.7) which implies that HLT rejects a bulk of the uninteresting event candidates that were selected by L0. Events that pass the HLT are processed by the reconstruction algorithm and sub-detector alignment calibration methods and this is referred to as offline reconstruction and selection. Note that HLT is further subdivided into 2 stages HLT 1 and HLT 2.

HLT 1 is responsible for the reconstruction of particles in VELO and T stations. It selects tracks which pass the selection criteria of high p_T and/or large IP and it reduces the rate to $\sim$ 30 KHz. HLT 2 performs an offline reconstruction of the b– hadron final states using the inclusive and exclusive trigger algorithms which aim to further reduce the rate to 2 KHz which is then written to the storage.

3 Analysis Tools & Strategy

This chapter introduces the analysis strategy employed for studying the decay: $B^+ \to \psi K^{*+}$. Different methods like Preselection, sWeighting and Re–weighting along with the PID calibration techniques for the Monte Carlo, are used to enhance the discriminatory power of the variables. Later, the BDT's are trained and pure signal sample is studied and signal and background yields are extracted for efficiency calculation.

3.1 Preselection Strategy

Due to similar kinematic properties the channel $B^+ \to J/\psi K^{*+}$ acts a reference channel to our signal channel: $B^+ \to \psi K^{*+}$. Plots in figure 3.9 and Appendix B reflect just how similar the two channels are. K^{*+} mesons which are produced at the B^+ decay[1] vertex, decays into K_s^0 and π^+. Due to small momentum of the π^+ it is referred to as soft pion (π_s^+) but unless explicitly stated this π_s^+ would be referred to as simply π^+. K_s^0 decays weakly and has a lifetime of ~ 8.9 ns [4] and it travels a distance of few centimeters in the detector[2] before it decays into a pair of oppositely charged pions π^+ and π^-. So the decay chain can be written as: $K^{*+} \to K_s^0(\to \pi^+\pi^-)\pi^+$ and the decay topology can be seen in Figure 3.1.

[1] B^+ mesons are produced at the primary vertex in pp collision.

[2] K_s^0 Flight Distance (FD) is sufficiently long for its decay vertex to be resolved, where FD refers to the distance a particle travels from a given decay vertex to another.

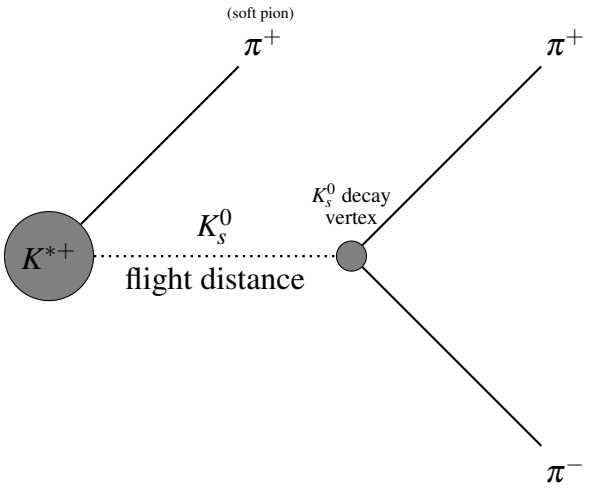


Figure 3.1: Schematic view of the decay topology of K^{*+} into $\pi^+ K_s^0$

The preselection cuts were determined by comparing the data sideband[3] for the $m(B^+)$ in the range 5600 to 6000 MeV and comparison plots against the Monte Carlo (*cf.* 3.3) were generated as can be seen in Appendix A (Figure 5.2 and Figure 5.3). The list of preselection variables is given in Table 3.1 and Table 3.2. These cuts are then later applied to the invariant mass distribution of the B^+. Fitting is performed on the candidates that pass this preselection criteria which results in signal and background yield, which is then used in the calculation of figure of merit[4] and preselection efficiency[5] ε. Preselection lends both of these channels to be relatively pure. Loose constraints are placed on transverse momenta[6] of the final state and intermediate particles to get rid of large amount of combinatorial background (*cf.* section 4.1), as combinatorial background mainly consists of lower momenta than signal. Although solely kinematic constraints on the particles does not yield an optimum selection, the efficiencies can be further enhanced by using PID variables. PIDK($(\pi^+) < 10$) cut was used in preselction. Note that PID variables need to be corrected using PID calibration technique (*cf.* section 3.3.4). Variable cuts are listed in Table 3.1 and Table 3.2. High purity was obtained by applying these few preselection cuts on both the channels. Further suppression of the combinatorial

[3]which acts as background proxy and is used for comparison with the simulated events.

[4]$FOM = N_S/\sqrt{N_S+N_B}$ where N_S and N_B refers to the number of signal and background events respectively.

[5]Variables for preselection are choosen such that they are uncorrelated or close to being uncorrelated.

[6]Transverse momenta or P_T refers to the particle momenta perpendicular to the beam axis: $P_T = \sqrt{P_x^2 + P_y^2}$

background was done by training BDT's (refer to section 4.2).

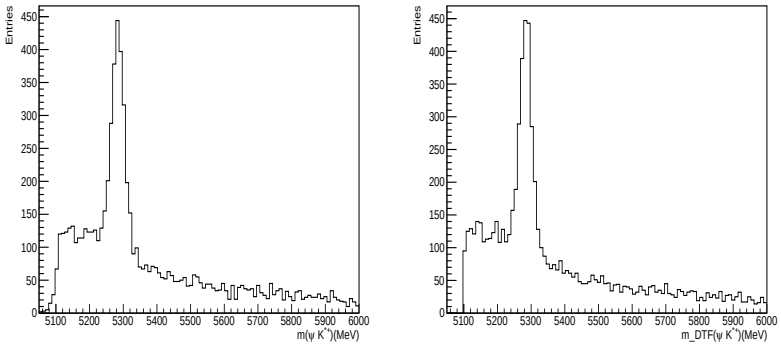
For a larger signal and background separation, throughout the analysis $m(B^+)$ variable is substituted by the B^+ Decay Tree Fitter ($m\_DTF(B^+)$) variable. Comparison plots between the two variables can be seen in Figure 3.3 and Figure 3.5, these plots also suggest that these two variables are very much alike. DTF algorithm [21] takes into account the whole decay chain and parametrizes it into particle's vertex position, flight distance and momentum and performs a fit on these parameters using relevant contsraints and these fits are performed using a Kalman filter [22].

3.1.1 Selection of $B^+ \to \psi(2S)K^0_s\pi^+$ candidates

The decay $B^+ \to \psi K^{*+}$ forms the signal channel for the analysis with B^+ decaying into ψ and $K^{*+}(892)$. $\psi(3686)$t meson exists in 2S resonance state and it decays into a pair of oppositely charged muons (μ^+ and μ^-). The dataset for the channel is taken from the stripping line B2XMuMuLine. Figure 3.2a shows the invariant B^+ mass which comes directly from the stripping whereas Figure 3.2b refers to the B^+ Decay Tree Fitter. Preselection cuts are listed in Table 3.1 which as mentioned previously were obtained by comparing data sideband with the Monte Carlo simulations (*cf.* Figure 5.2, Appendix A). Additionally a two sided mass cut on ψ is placed in the mass range $3640 < \mathrm{m}(\psi) < 3730$, in order to further reduce the background events populating the data sample. Figure 3.3 shows the result of applying these selection cuts on the Decay Tree Fitter which yields a high purity of $B^+ \to \psi K^{*+}$ candidates.

Particle	**Cut**
K^{*+}	$P_T > 700$
B^+	$P_T > 1300$
K^0_s	$P_T > 400$
ψ	$3640 < m_\psi < 3730$
π^+	PIDK < 10

Table 3.1: Preselection cuts for $B^+ \to \psi(2S)K^0_s\pi^+$



(a) (b)

Figure 3.2: Comparison plot for signal channel with no preselection cuts applied of (a) Invariant B^+ mass distribution and (b) B^+ Decay Tree Fitter ($m\_DTF\,[0]\,(B^+)$).

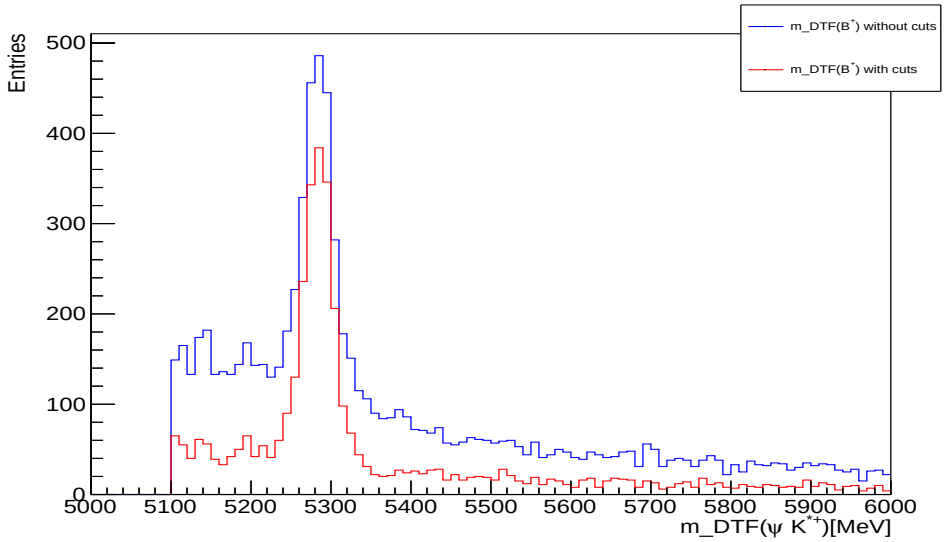


Figure 3.3: Comparison of $m\_DTF\,[0]\,(B^+)$ with and without the preselection cuts (*cf.* Table: 3.1) applied.

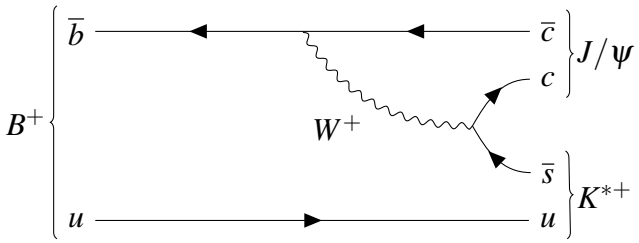


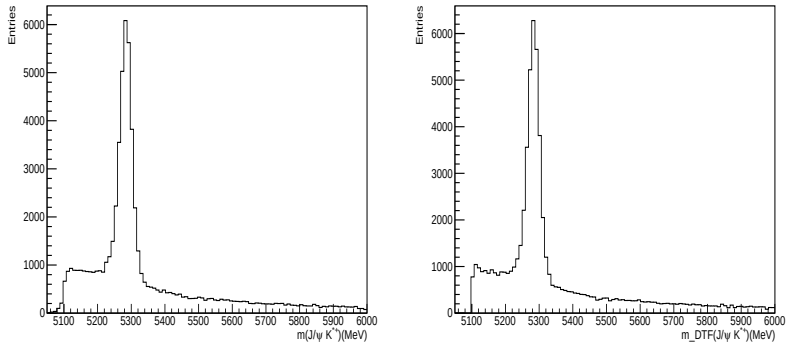
Figure 3.4: Lowest order Feynman diagram corresponding to the decay $B^+ \rightarrow J/\psi(1S)K^{*+}$

3.1.2 Selection of $B^+ \rightarrow J/\psi(1S)K_s^0\pi^+$ candidates

The decay $B^+ \rightarrow J/\psi 1SK^{*+}$ is used as a control channel for the analysis with the lowest order Feynman diagram shown in Figure 3.4. The $J/\psi(3096)$ exists in 1S resonance state and decays into a pair of muons (μ^+ and μ^-). The preselection cuts are listed in Table 3.2 and were obtained by comparing Monte Carlo events with the data sideband for $m(B^+)$ and can be seen in Figure 5.3 (Appendix A.). Similar to $B^+ \rightarrow \psi K^{*+}$ channel, a two sided mass cut on J/ψ is placed in the mass range $3050 < \mathrm{m}(J/\psi) < 3140$ in order to reduce the combinatorial background. Applying these cuts results in a relatively pure sample as can be seen in Figure 3.6. Comparison between the invariant B^+ mass distribution and B^+ Decay Tree Fitter for this channel can be seen in Figure 3.5.

Particle	**Cut**
K^{*+}	$P_T > 700$
B^+	$P_T > 1300$
K_s^0	$P_T > 400$
J/ψ	$3050 < m_{J/\psi} < 3140$
π^+	PIDK < 10

Table 3.2: Preselection cuts for $B^+ \rightarrow J/\psi(1S)K_s^0\pi^+$



(a) (b)

Figure 3.5: Comparison plot for reference channel with no preselection cuts applied (a) Invariant B^+ mass distribution and (b) B^+ Decay Tree Fitter ($m\_DTF\,[0]\,(B^+)$).

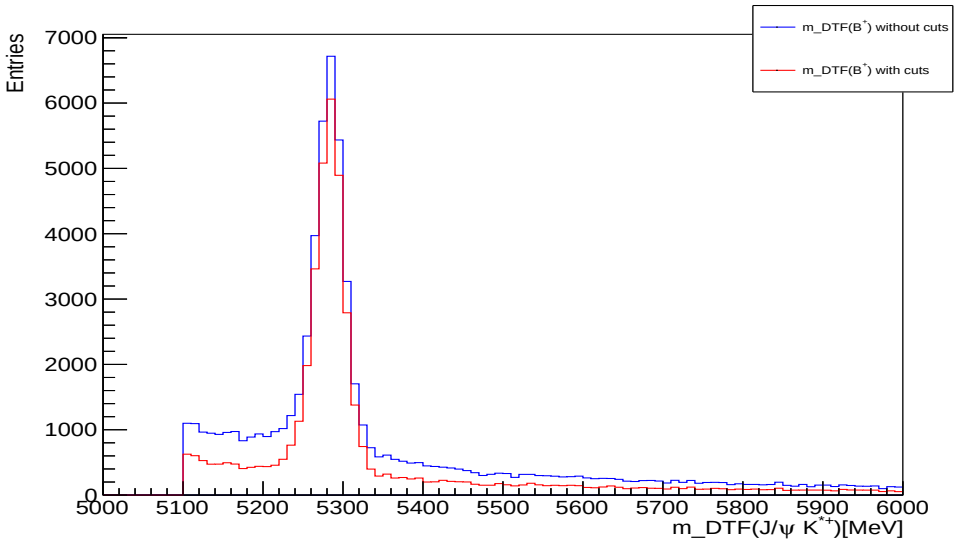


Figure 3.6: Comparison of B^+ Decay Tree Fitter with and without the preselection cuts (*cf.* Table: 3.2) applied.

3.2 Signal Selection

The data sample consists of background and signal component. The sources of these components can in most cases be unknown. A sample variable can

be classified as discriminating or a control variable on the basis of whether its underlying sources are known or unknown. In case of discriminating variable the underlying sources are known, whereas incase of control variable this is not the case. A statistical tool sPlot [23] can be used to unfold these sources from a given discriminating variable. It relies on performing a fit on the discriminating variable using the maximum likelihood approach (discussed in the next section) and extract the sWeights from these fits which help in establishing the sources of signal and background events in a discriminating variable. These sWeights[7] (sP) can be either positive or negative depending on whether the contribution is from a signal or background source respectively.

This method is then used in reconstructing the sources for the given control variable without any prior knowledge of its sources. The underlying assumption behind this method is that the control variable are uncorrelated with the discriminating variable. More formally, the sPlot formalism is designed to reconstruct true distribution (M_n(x)) for a control variable x, while keeping all the signal events and rejecting all the background events where M_n(x) is solely determined from the PDF's of the discriminating variable. In the analysis $m\_DTF(B^+)$ is used as the discriminating variable.

The quality of the fit can be gauged by comparing the expected distribution of the control variable (Monte Carlo) with the sWeighted data sample as can be seen in Figure 3.10 which shows comparison plots for the control channel. Note that the plots shown in Figure: 3.8a, 3.8b and 3.8c are further corrected and are discussed in detail in a later section (*cf.* section 3.3).

***Addendum: Likelihood Method**

The data sample comprises of varies species of events which include signal components (ie. interesting sources of events) and background component (uninteresting events from sources that accompany the signal). The likelihood method [23] works by maximizing $\mathscr{L}$ as given in equation:

$$\mathscr{L} = \sum_{e=1}^{N} ln\{\sum_{i=1}^{N_s} N_i f_i(y_e)\} - \sum_{i=1}^{N_s} N_i \tag{3.1}$$

[7]These sWeights are calculated using the equation: $sP_n(y_e) = \frac{\sum_{j=1}^{N_s} V_{nj} f_j(y_e)}{\sum_{k=1}^{N_s} N_k f_k(y_e)}$, for more on this refer to [23]

where,

- N is the total number of events in data sample.
- y is the set of discriminating variables.
- f_i is the PDF of an i^{th} discriminating variable.
- x is the set of control variables which do not appear in above expression.

And these parameters can be determined by maximizing this log-likelihood.

3.2.1 Parameterization of the fit

The data contains contributions from both signal and background components and this section outlines the method to classify sample data as signal or background on statistical basis. The data B^+ invariant mass distribution for $B^+ \to \psi K^{*+}$ and $B^+ \to J/\psi K^{*+}$ candidates are fitted using maximum likelihood method. The fit model is defined for the two channels separately but the background component which is dominated by the combinatorial background (*cf.* section 4.1) in both the channels is described by a common decaying exponential function:

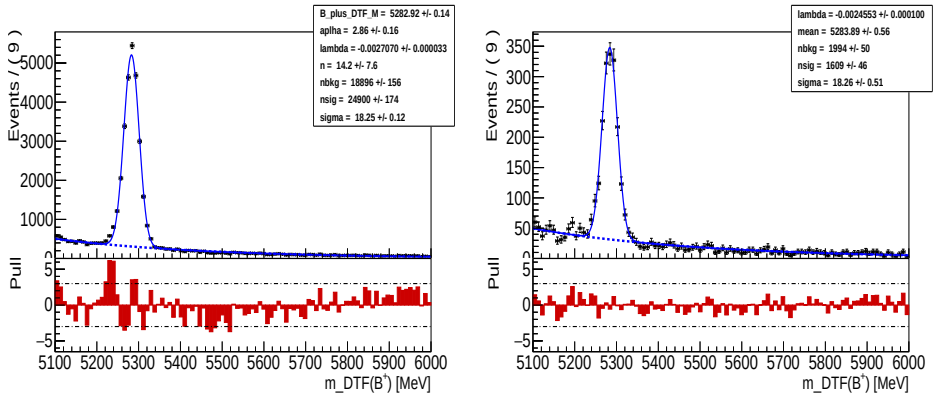
$$f_{bkg} = \lambda \exp\{-\lambda x\}$$

whereas the signal component incase of $B^+ \to J/\psi K^{*+}$ is best described by a crystal ball function [24]:

$$f_{sig}(x;\alpha,n,\bar{x},\sigma) = \begin{cases} \exp\left\{-\frac{(x-\bar{x})^2}{2\sigma^2}\right\}, & \text{for } \frac{x-\bar{x}}{\sigma} > -\alpha \\ A.(B-\frac{x-\bar{x}}{\sigma}), & \text{for } \frac{x-\bar{x}}{\sigma} \leq -\alpha \end{cases}$$

where,

$$A = \left(\frac{n}{|\alpha|}\right)^n \exp\left(-\frac{|\alpha|^2}{2}\right),$$
$$B = \frac{n}{|\alpha|} - |\alpha|,$$
$$N = \frac{1}{\sigma(C+D)}.$$



(a) (b)

Figure 3.7: Fit to the B^+ mass distribution for (a) Reference channel ($B^+ \to J/\psi K^{*+}$) and (b) Signal channel ($B^+ \to \psi K^{*+}$) shown along with the pull plots.

Therefore the Total PDF for the Control channel can be written as:

$$PDF(f_{sig}, f_{bgk}) = N_{sig}.f_{sig} + N_{bkg}.f_{bkg} \tag{3.2}$$

while for the Signal channel, $B^+ \to \psi K^{*+}$ the signal component is described using a gaussian:

$$f'_{sig} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu^2)}{2\sigma^2}\right)$$

Therefore the Total PDF for the control channel can be written as:

$$PDF'(f'_{sig}, f_{bgk}) = N_{sig}.f'_{sig} + N_{bkg}.f_{bkg} \tag{3.3}$$

The fits done on the data sample for both the channels and are shown in Figure 3.7 and the bottom plots in the same figure show the pull distribution which represents the goodness[8] of the fit. All the free parameters are established from the fit itself and the resulting yields are listed in Table 3.3 along with the measured Figure of Merit (FOM) and preselection efficiency (ε_{presel}).

[8]The difference between the data points and fit the function.

Channel	**N_S**	**N_B**	**FOM**	**$\varepsilon^{sig}_{presel}$**
$B^+ \to \psi K^{*+}$	1535.88	1903.38	26.19	0.919
$B^+ \to J/\psi K^{*+}$	23757.40	18029.40	116.22	0.945

Table 3.3: Yields extracted from the 1–D fits to both the channels along with the calculation of figure of merit (FOM) and preselection efficiency ($\varepsilon^{sig}_{presel}$). N_S and N_B refer to the number of signal and background events respectively in the $\pm$ 2 σ region from the signal peak.

3.3 Monte Carlo

Monte Carlo (MC) events are simulated events created using framework like PYTHIA or DAVINCI. MC tries to recreate the experimental conditions and performance as they were during the time of particle collision. The event statistics for the two decays channels can be seen in Table 3.4 along with the year of the MC sample and the magnetic polarity of the samples.

		Year		Total
		2011	2012	
Signal Channel (Events)	Mag Up	50,553	35,327	180,502
	Mag Down	59,741	34,881	
Control Channel (Events)	Mag Up	–	–	18,869
	Mag Down	18,869	–	

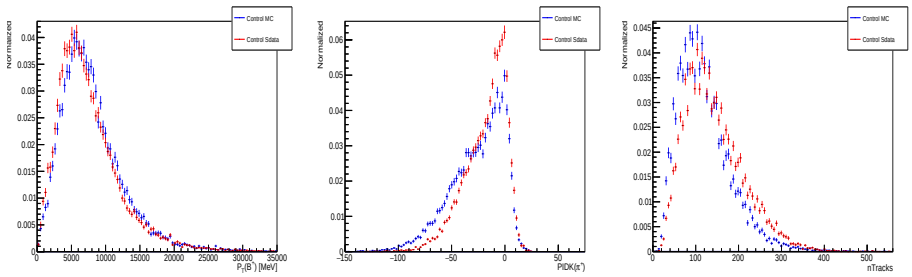
Table 3.4: Statistics of the simulated events from the Run 1 at the LHCb for the channels: $B^+ \to J/\psi K^{*+}$ and $B^+ \to \psi K^{*+}$

Monte Carlo's used in the analysis were Truth Matched [25]. Monte Carlo truth Matching refers to applying a cut variable on the MC sample called BackgroundCategory. The MC events are ranked according to their BackgroundCategory. Value of 10 and 0 represent true MC distribution with no background contamination. Value of 50 and above usually represent background events (combinatorial background).

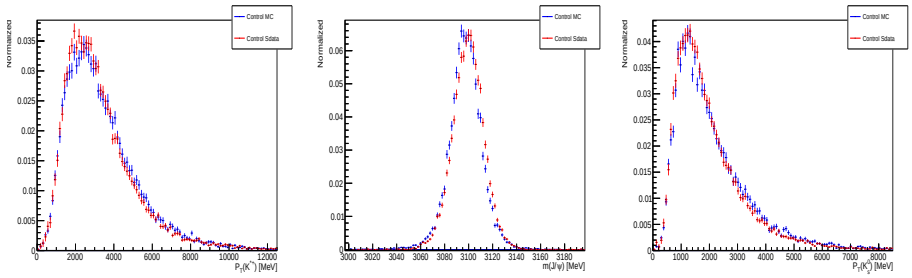
3.3.1 Comparing: Monte Carlo & Data

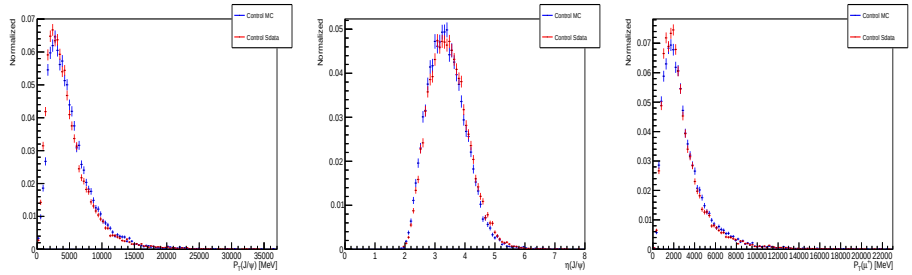
Figure 3.10 shows the comparison plots between the sWeighted data sample and the truth Matched Monte Carlo. These plots give a visual inspection that

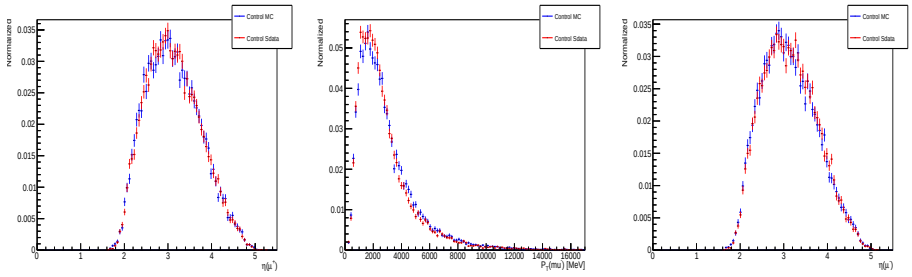
the sWeighted data and MC events are similar. The plots 3.8a, 3.8b and 3.8c are from variables: $BP_T(B^+)$, $PIDK(\pi^+)$ and nTracks which show discrepancy between Truth Matched Monte Carlo and sWeighted data sample. These 3 variables will be later corrected for using methods like PID calibration and MC re-weighting.

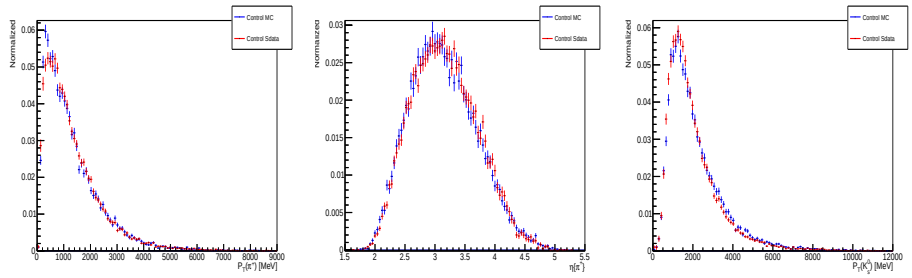

Control MC
Control Sdata
Normalized
$P_T(B^+)$ [MeV]
PIDK(π^+)
nTracks

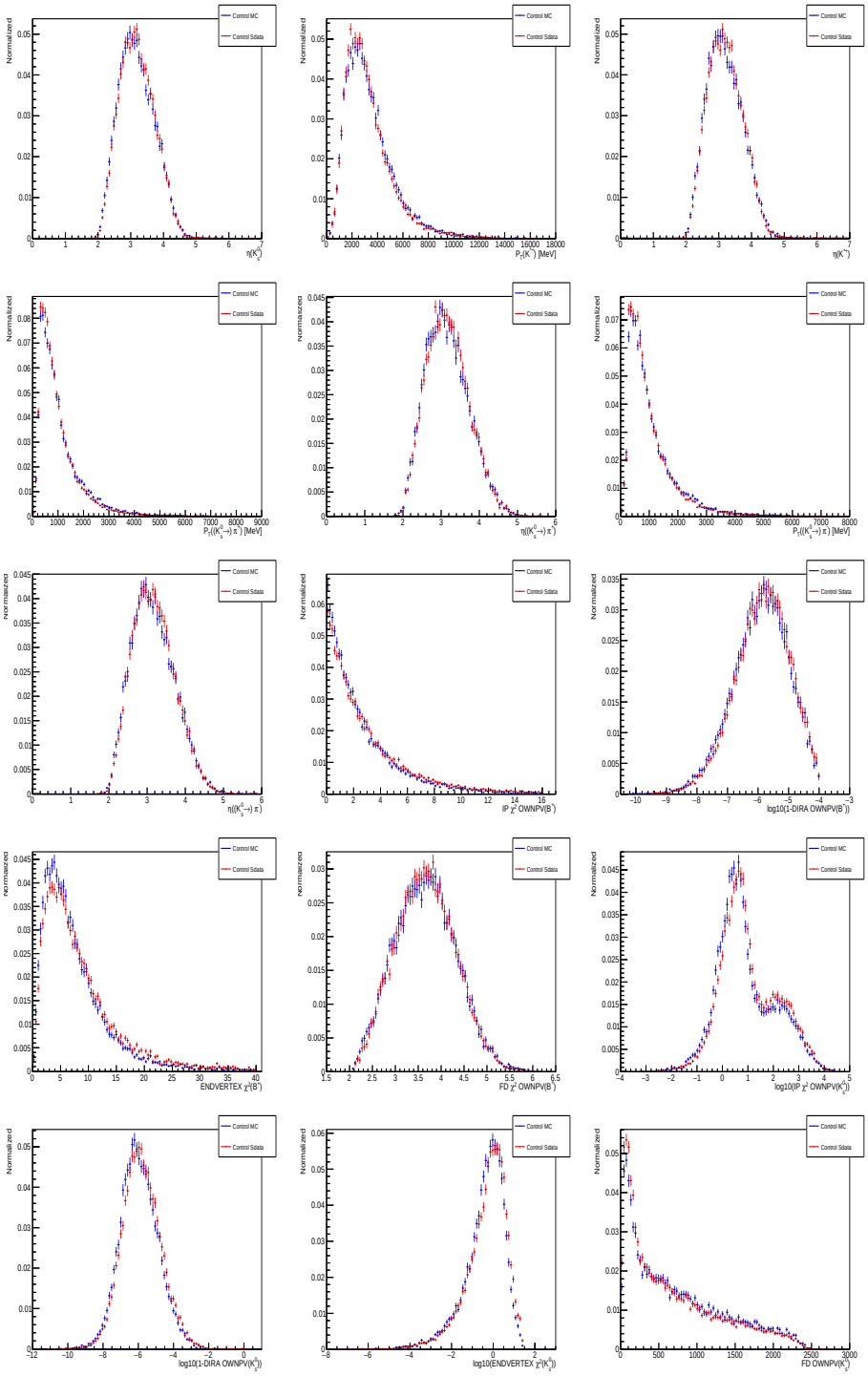
(a) (b) (c)


Control MC
Control Sdata
Normalized
$P_T(K^+)$ [MeV]
m(J/ψ) [MeV]
$P_T(K^+_2)$ [MeV]


Control MC
Control Sdata
Normalized
$P_T(J/\psi)$ [MeV]
$\eta(J/\psi)$
$P_T(\mu^+)$ [MeV]


Control MC
Control Sdata
Normalized
$\eta(\mu^+)$
$P_T(mu)$ [MeV]
$\eta(\mu)$


Control MC
Control Sdata
Normalized
$P_T(\pi^+)$ [MeV]
$\eta(\pi^+)$
$P_T(K^0_S)$ [MeV]


Control MC
Control Sdata
Normalized
normalized

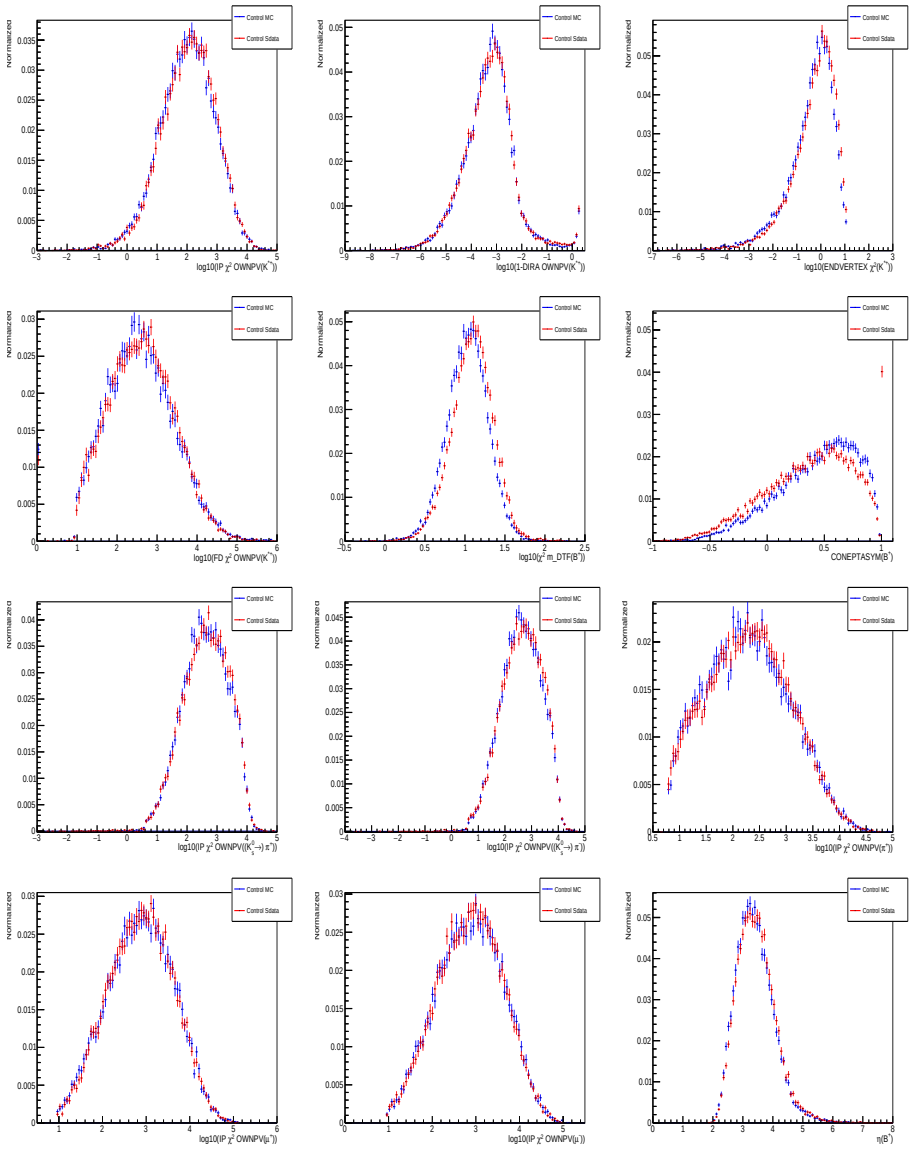


Figure 3.10: Comparison plots for the Control channel between the Sweighted dataset with the Truth Matched Monte Carlo.

3.3.2 Comparing: Signal Monte Carlo & Control Monte Carlo

These plots shown in Figure 3.11 show a good agreement between the two Monte Carlo's (one from Control channel and other from Signal channel) and hence establishing that the two decay channels are rather similar and can be used in the analysis.

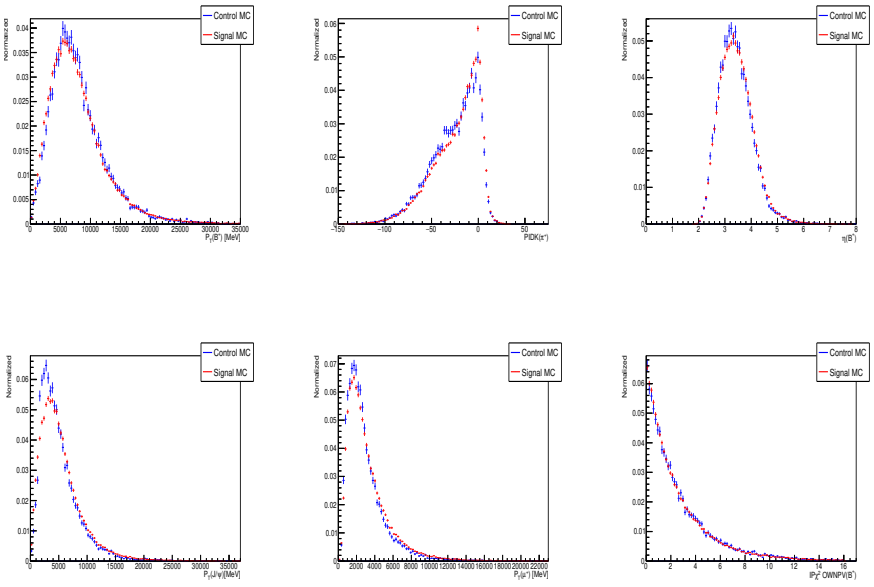


Figure 3.11: Simulated events (Truth Matched Monte Carlo) comparison between Signal and Control channel (for more plots *cf.* Appendix B).

3.3.3 Re-weighting

Re-weighting method provides a good agreement between the data and the Monte Carlo, as stated earlier the variables $P_T(B^+)$ and nTracks are typically re-weighted. The re-weighted ratio can be seen at the bottom plot in Figure 3.12. These ratios are obtaining by diving two normalised histogram's, sWeighted data with Truth Matched Monte Carlo. These ratios are then applied to the data to obtain a good agreement between the two.

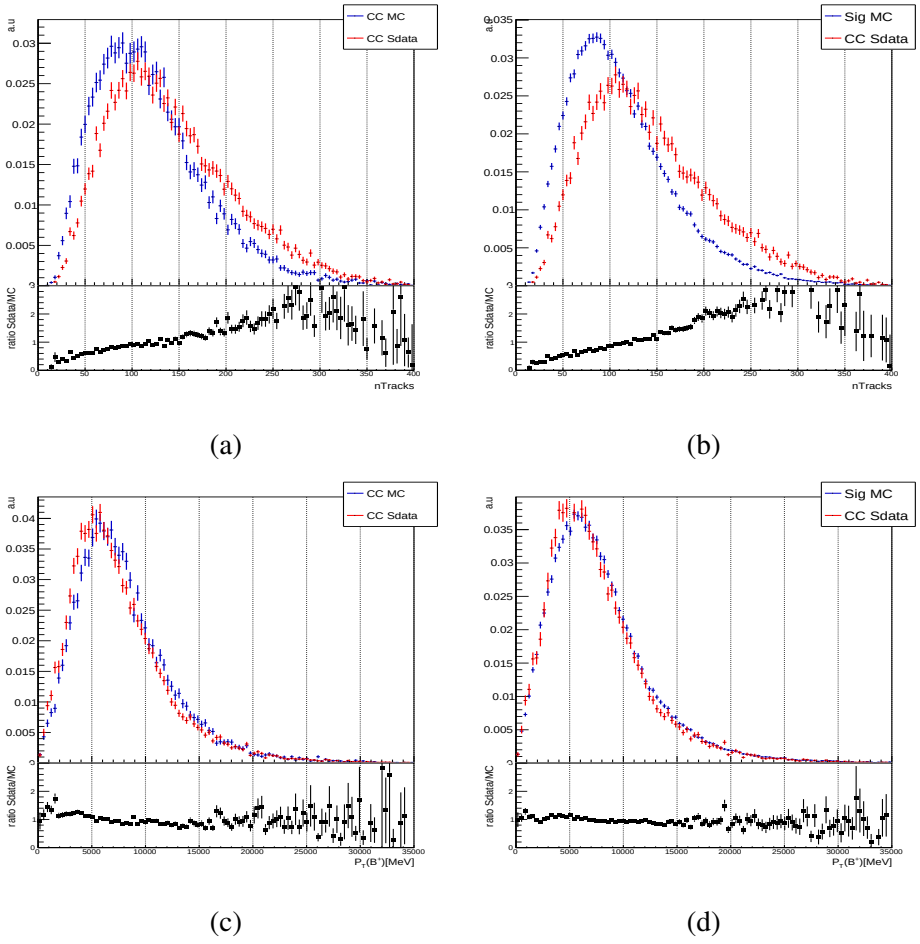


Figure 3.12: Weighted plots for Simulated events obtained by dividing nTracks distribution of sWeighted data by the Truth Matched Monte Carlo (Plots: (a) & (b)) and similarly for the plots: (c) & (d) the re–weight ratio for the Monte Carlo is obtained for transverse momentum of the B^{+}.

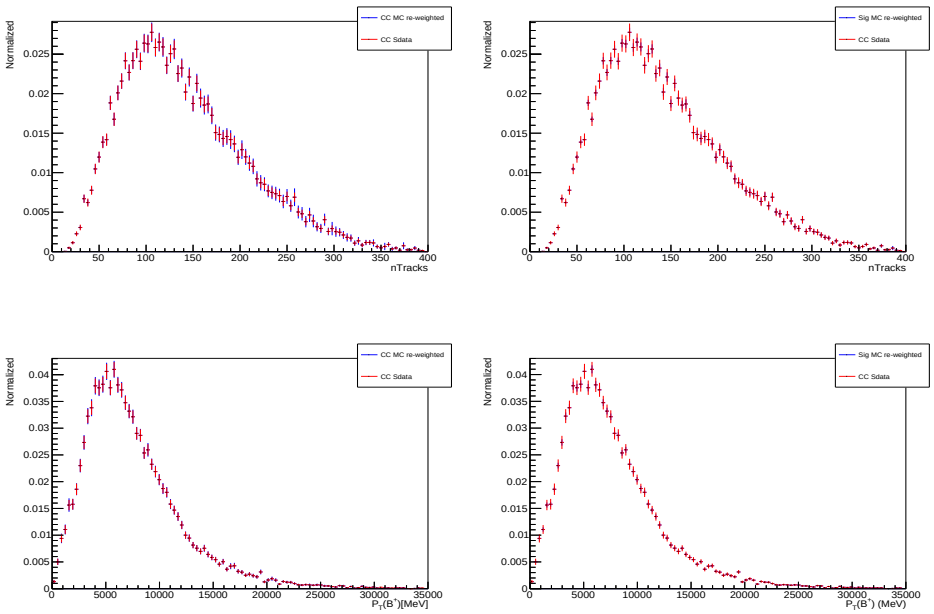


Figure 3.13: Comparison plots showing corrected Monte Carlo with the re–weights applied to nTracks and $P_T(B^+)$ distribution.

3.3.4 PID Calibration

Figure 3.14 shows the PID plots for π^+. The plots represent the corrected and uncorrected PID plots.

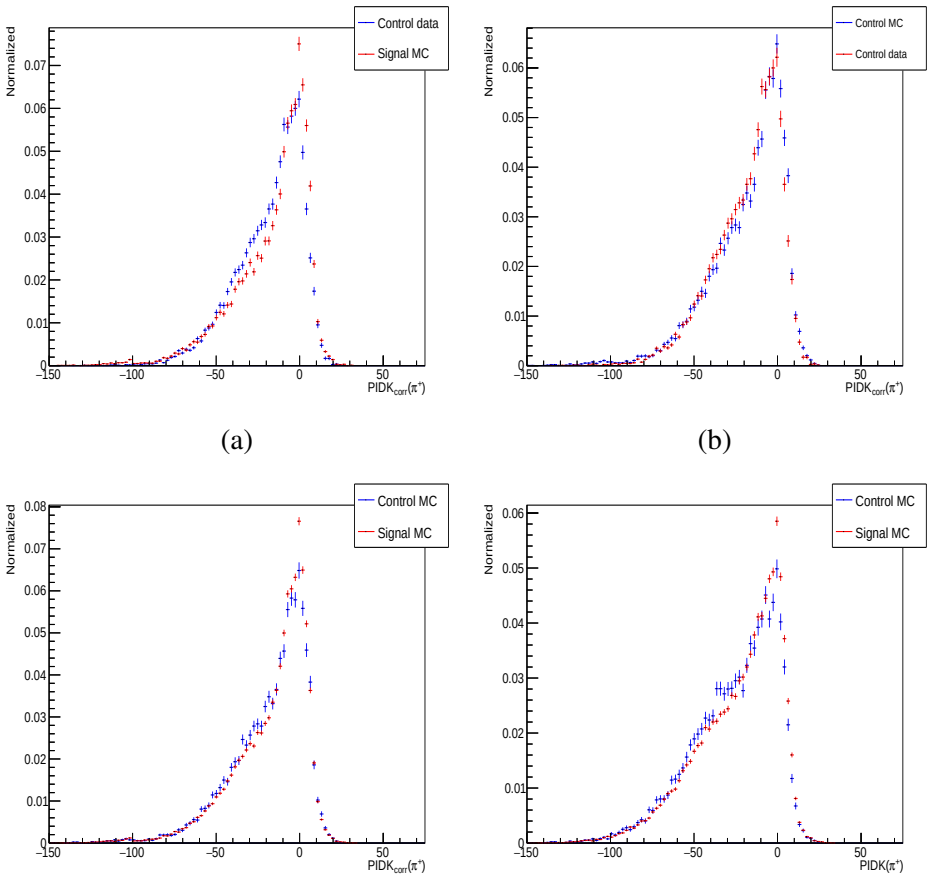


Figure 3.14: Pions (π^+) PIDK correction plots with the corrections applied to the simulated events for both the channels.

4 Multivariate Analysis, Training BDT's

This chapter introduces the types of background that are essential for the analysis. Later on, Boosted Decision Tree's or BDT's are introduced and trained and tested. In the analysis gradient boosting was implemented with the bagging option to train the decision tree's in order to suppress the combinatorial background from the data sample and limit the classifier output from getting overtrained.

4.1 Background

In the analysis mainly 2 types of backgrounds are considered which include physical background coming from random particles and misidentified particles which are given wrong flavour tag.

1. Combinatorial Background: Random particles which do not originate from $B^+ \to J/\psi K^{*+}$ or $B^+ \to \psi K^{*+}$ decay. Since these particles are independent of the Signal and reference channel decay particles, hence they reside throughout the B-meson mass spectrum right from lower end of the mass tail to the upper end. For instance, combinatorial background from ψ decay which is combined with unrelated K_s^0 candidate is one such source of this type of background. Also note that this background can be significantly reduced with a trained BDT but even after that a small residue of this background would still be present.

2. Random pion: Pions are the most abundantly produced particles at the

LHCb detector, these random pions can easily mimic the soft pions (π_s^+) produced from the B-meson decay, in the decay chain $B^+ \to J/\psi\pi^+K_s^0$ or $B^+ \to \psi\pi^+K_s^0$. Note that these pions exhibit very different kinematic properties as compared to the B^+ daughter particles for the Reference and Signal channel. Additionally these random pions can also mimic one or both of the the final state muons in the decay channels.

3. Ghost Pion: Unrelated hits in the detector that are reconstructed as coming from a particle which does not exist[1]. These particle tracks on average come from low P_T pions. And as these hits are unrelated, the inferred trajectory and hence the the ghost pion charge is usually opposite of that of the original pion.

4.2 Introduction to Boosted Decision Tree's

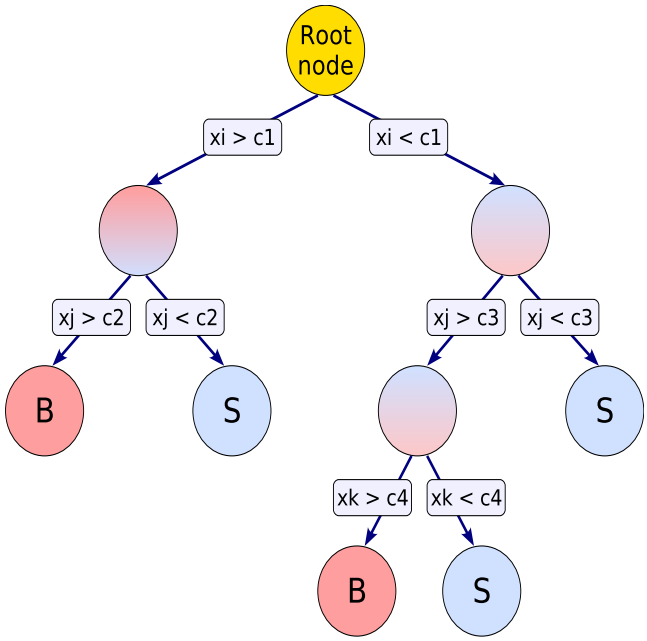


Figure 4.1: Schematic view of Boosted Decision Tree, starting with the Root Node. Each node having a binary split as either B (background) or S (signal) using the input variable. The decision is made on the basis of majority of events that end up at a given node [26].

[1] As these tracks cannot be matched with a true Monte Carlo particle [14].

A common MVA (Multivariate Analysis) tool used in analysis at LHCb is called the BDT or Boosted Decision Tree [26]. Where a root node is consecutively split into binary nodes and at every successive step a new layer is added to the decision tree, the splitting of the node continues until a decision is made of whether to classify a given event as background or signal (*cf.* Figure 4.1). The BDT algorithm uses the input variables or the discriminating variables during the training phase, and determine a cut value that maximises the separation between the signal and background. The tree grows larger as more input variables are added.

During the training phase the Monte Carlo sample was used as a proxy for signal and m_DTF(B^+) sideband data from 5600 MeV to 6000 MeV was used as background proxy to train the BDT classifier. This classifier can combine multiple weak training input variables with low separation power (between signal and background) and combine multiple such decision trees to create a forest and yield a strong classifier[2] and therein lies the power of the BDT method. The boosting method applied in the analysis is called as gradient boosting[3] as it is much less likely to be influenced with the statistical fluctuations. The classifier learns to separate background and signal events in multidimensional phase space and combines the output into a single variable called the classifier output. Post training phase this classifier output can then be applied to a data with an unknown composition of signal and background component.

MVA Configuration

All the BDT configuration options are manually selected in order to yield a strong classifier output with little to no overtraining and these options are listed in Table 4.1. During the training and testing of the BDT the data is split randomly into two equal halves in order to obtain an optimal results from the BDT. These BDT's can easily suffer from overtraining[4] where they can learn to pick up statistical fluctuations in data and in order to reduce overtraining BoostType chosen to be Gradient. The number of trees in a forest (NTrees) are set to 170 and maximum depth of individual decision tree (MaxDepth) is set to 2. Convergence is attained when 2.5% of training events have reached at a given node[5]. Additionally, UseBaggedGrad is set to True where a bagging criteria is implemented on the decision tree. Bagging denotes a resampling technique

[2]This process of combining multiple weak classifiers into a strong one, is called as boosting.
[3]Hence the name BDTG.
[4]Large number of nodes in BDT is mainly the reason why BDT's are prone to overtraining.
[5]Classification of a given event is done by reaching a majority vote.

where a trained event has a possibility of replacement, meaning that the same event is allowed to be picked several times [26].

Option	Value
BoostType	Gradient
NTrees	170
MinNodeSize	2.5%
UseBaggedGrad	True
BaggedSampleFraction	0.5
nCuts	20
MaxDepth	2

Table 4.1: Configuration options used for booking the Decision Tree (for the Training and Testing phase).

Input Variables

The input variables are chosen to enhance the separation power of the BDTG response variable and are listed in Table 4.2 where they are ranked according to their power of separation.

The following variables are used in the training process:

- Transverse momentum (P_T) of pion (π^+), K short (K_S^0) and B^+.
- Logarithm of 1–DIRA of B^+, where DIRA refers to the the cosine of the Direction Angle and is defined as the angle between the particle's momentum and the direction vector between the primary and secondary vertex and for a signal particle its usually close to 1. Note that the logarithm was chosen to get a more flatter distribution, as BDT's can handle flat distributions better than spiking one's.
- Impact parameter χ^2 of B^+ defined as the difference of the χ^2 of the primary vertex with and without the tracks of interest. Note that the signal events usually have a small value for this parameter.
- Flight Distance of B^+ and K_S^0 (*cf.* Figure 3.1), refers to the distance a particle travels from a given decay vertex to another.

Rank	Variable	% Sepration
1	$P_T(pi^+)$	34.12
2	$ln(1-DIRA(B^+))$	32.64
3	$P_T(K_S^0)$	30.55
4	$P_T(B^+)$	22.77
5	$IP\chi^2(B^+)$	16.18
6	$FlightDistance(B^+)$	13.71
7	$P_TConeAsym(B^+)$	8.31
8	$FlightDistance(K_S^0)$	1.94

Table 4.2: List of variables trained using BDTG method of TMVA, variables are listed in accordance with their power of separation (given in percent).

- Cone Asymmetry P_T [27] of B^+ is defined as the number of charged tracks that pass through a cone around a given particle's flight direction and do not contribute to its reconstruction. The quantity is given as:

$$P_TConeAsym(B^+) = \frac{\sum_i P_T - \sum_j P_T}{\sum_i P_T + \sum_j P_T} \tag{4.1}$$

where i is refers to all the B^+ candidates in the decay $B^+ \to \psi K^{*+}$ while j refers to all the other tracks within the cone.

Figure 4.2 provides the distribution of the input variables trained using BDTG, and it provides a visual inspection of the variables and their power of separation. Correlation matrices for signal and background samples are computed with TMVA framework and are shown in Figure 4.3. MVA poses a distinct advantage as compared to a cut based selection criteria, as any possible correlation between the input variables can be taken into account. Note that using input variables which are less correlated yield a strong classifier output with larger separation power between signal and background. Input variables trained by BDTG in the analysis are satisfyingly independent of each other (*cf.* Figure 4.3).

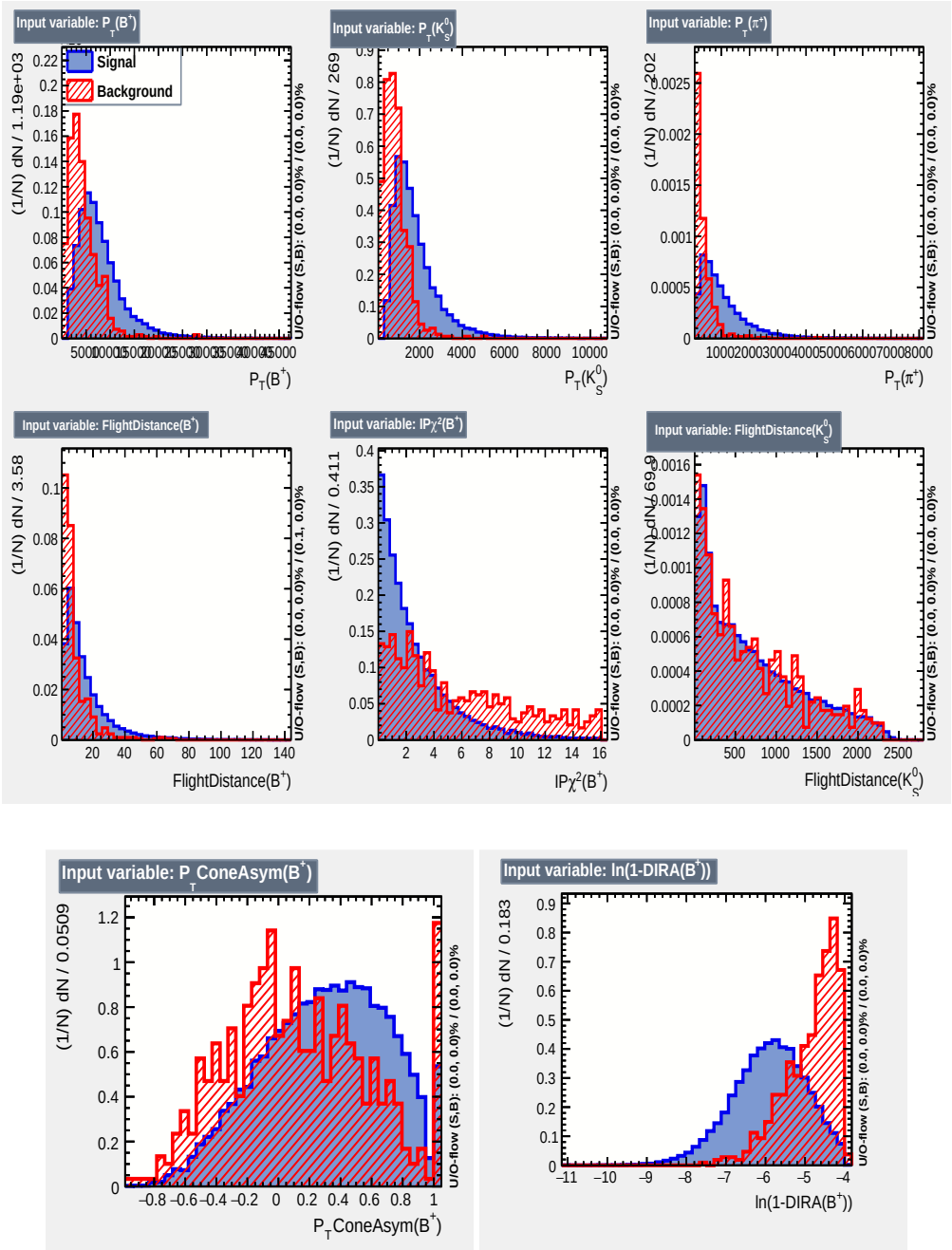
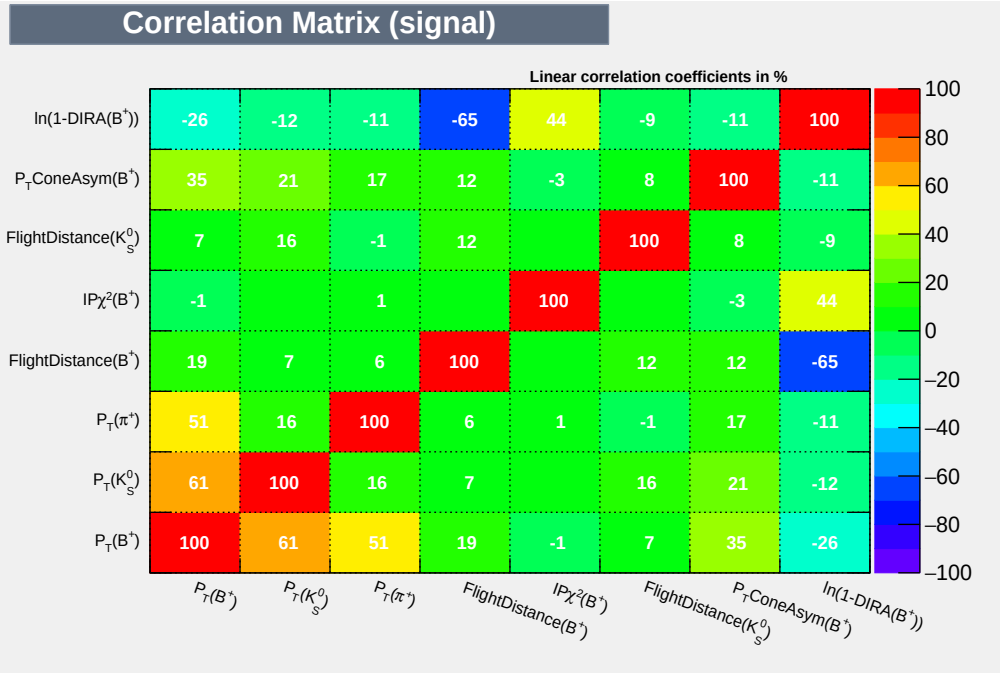


Figure 4.2: Distribution of the MVA input variables for the decay channel, $B^+ \to \psi K^{*+}$. Monte Carlo is used as the Signal (blue) proxy and the m_DTF(B^+) sideband from 5600 MeV to 6000 MeV is used as the background (red) proxy.



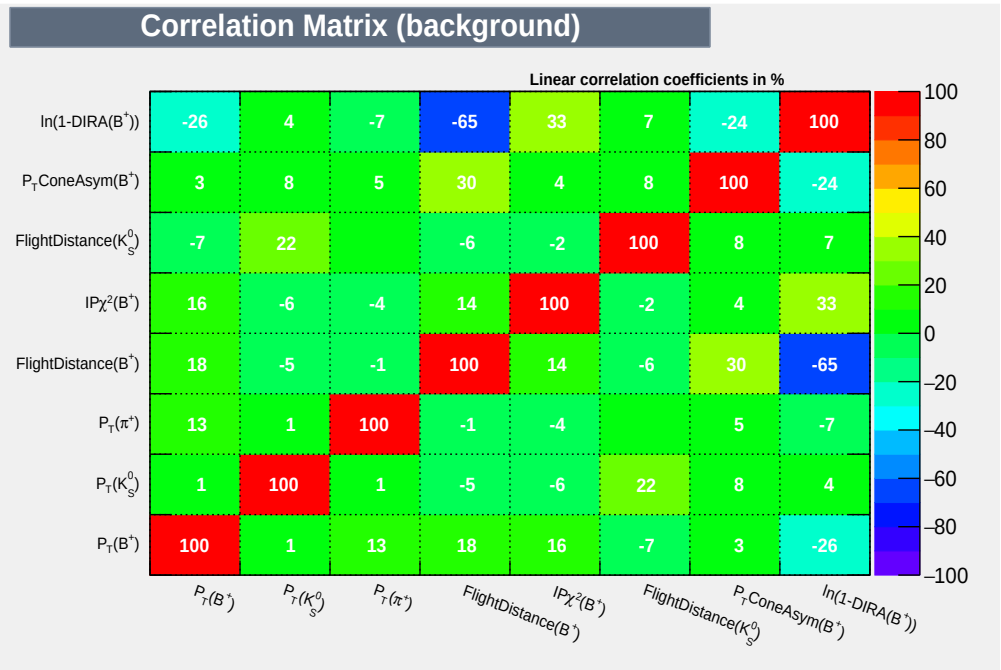


Figure 4.3: Correlation coefficients between the input variables after training and testing of the BDTG.

4.3 MVA Output

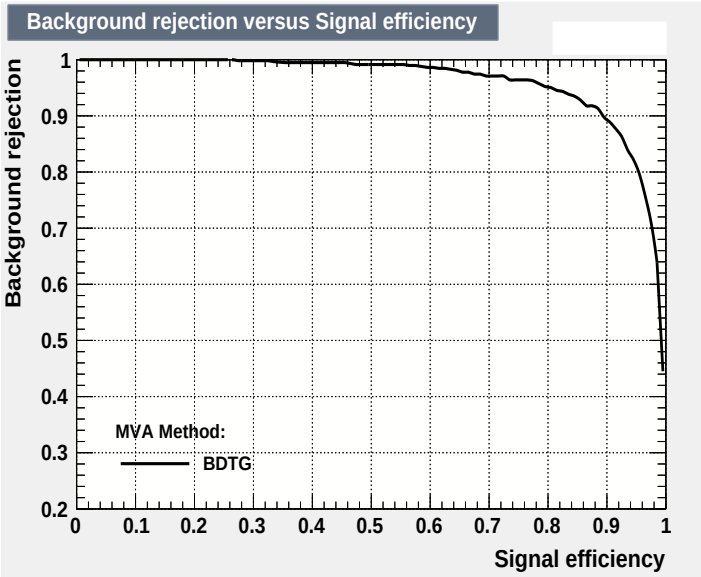


Figure 4.4: ROC (Receiver Operating Characteristic) curve for BDTG. Displaying background rejection versus signal efficiency.

The BDTG output is between -1 and +1 with -1 being purely background like and +1 being purely signal like event. The performance of a trained BDT can be gauged by looking at ROC (Receiver Operating Characteristic) curve (see Figure 4.4) which is a plot between signal efficiency and background rejection $(1-\varepsilon_{bkg})$. Note that the background rejection worsen's with higher signal efficiency. The overall area under the ROC curve provides a good estimate on how well the classifier performed in separating signal and background.

As previously mentioned BDT's are prone to overtraining. Where overtraining refers to when MVA algorithm interprets statistical fluctuations on training sample as a part of the variable distribution. A common way to test for overtraining is by splitting the data into two equal halves (randomly) into training and testing phase. Then comparing the MVA response for signal and background for the two phases. Figure 4.5 overlays the testing and training results for the given BDTG classifier and considering that there is no significant deviation in the two signal plots, it can be concluded that no overtraining occurred. However due to lack of background events used during training phase a slight overtraining can be seen in the background sample.

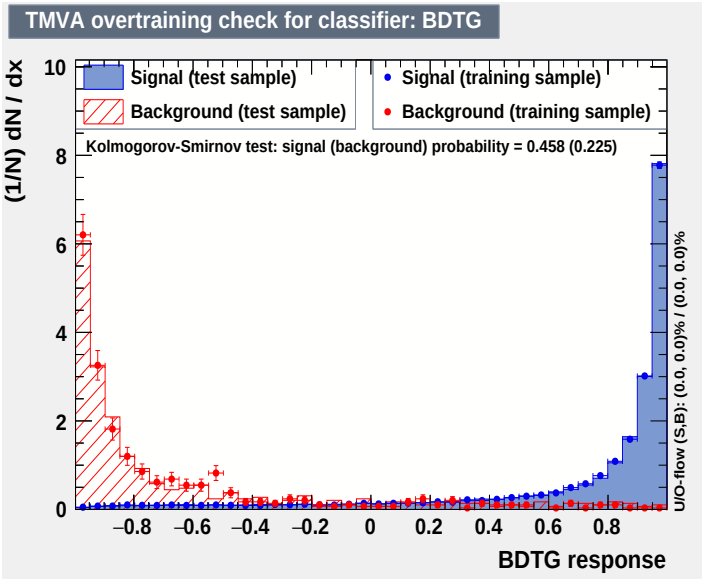


Figure 4.5: BDTG output response plot for training and testing sample. Kolmogorov-Smirnov test [28] provides a quantitative test for overtraining between the overlaid plots.

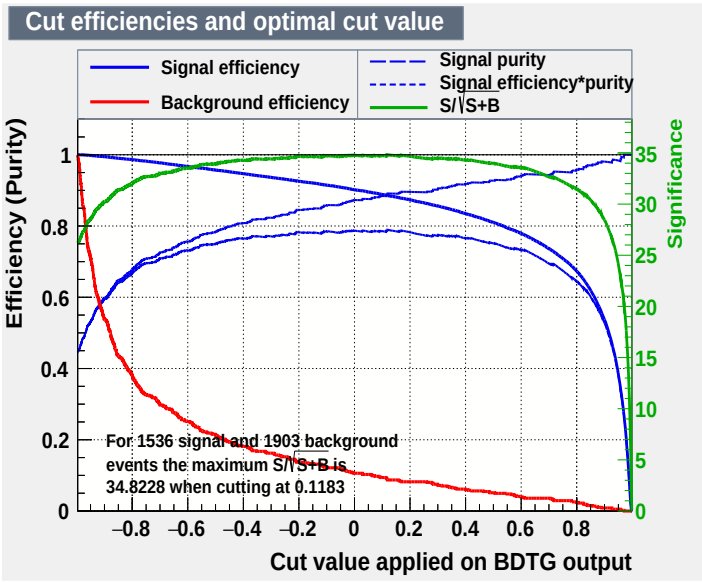


Figure 4.6: Optimisation of the BDTG cut value in order to maximise the figure of merit. Signal and Background yields taken from Table 3.3.

BDT's were trained to get rid of majority of the combinatorial background while keeping most of the signal events intact. Next the optimal working point or the BDTG cut value was determined. Preselection event statistics were taken from Table 3.3 for the Signal channel in order to evaluate a cut value for the BDTG. Using the event statistics and BDTG response, figure of merit, $(S/\sqrt{S+B}) = 34.82$ and BDTG cut value > 0.1183 were calculated by the MVA classifier as can be seen in Figure 4.6.

5 Branching Ratio Measurements

In this chapter we finally calculate the Branching Ratio for the Signal Channel using the Run 1 data.

5.1 Signal Yield Extraction

Preselection cuts were applied to the data from both the channels, along with the BDTG cut > 0.1183. The candidates that passed the cut were fitted using the same models as described before while performing the fits during the preselection stage. The signal component for the decay $B^+ \to \psi K^{*+}$ is described by a gaussian and the signal component for the decay of $B^+ \to J/\psi K^{*+}$ is described by a crystal ball function [24], the background component for both decays is described by a decaying exponential function (*cf.* Section 3.2.1). The resulting fit plots are shown in Figure 5.1. The pull distribution plots show the goodness of the fit. The amount of combinatorial background that has been suppressed can also be seen in Figure 5.1.

The yields exacted from these fits is displayed in Table 5.1. Where $N_{S|BDT}$ refers to the number of signal events from the whole fit with BDTG and preselection cuts applied to m_DTF(B^+) distribution. Similarly $N_{B|BDT}$ refers to the background events after applying BDTG and preselection cuts to m_DTF(B^+) distribution. Figure of merit is calculated using:

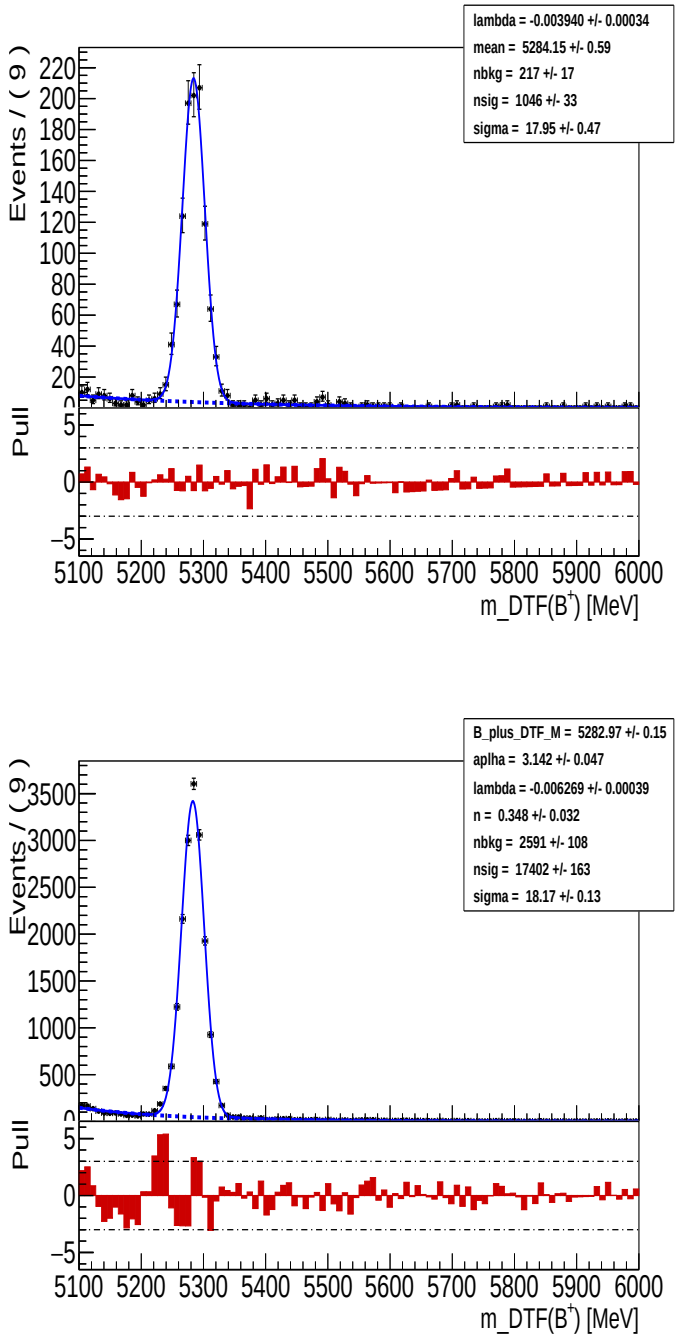


Figure 5.1: Fits to B^+ mass distribution after applying the BDT cut to Signal Channel (top) and Reference Channel (bottom). Pull plots can be seen at the bottom of each plot. The yields are represented in the Table 5.1

Channel	$N_{S\|BDT}$	$N_{B\|BDT}$	FOM	ε^{sig}_{BDT}	$\varepsilon^{sig}_{presel}$
$B^+ \to \psi K^{*+}$	1046 ± 33	217 ± 17	29.43	0.575	0.919
$B^+ \to J/\psi K^{*+}$	17402 ± 163	2591 ± 108	123.07	0.658	0.945

Table 5.1: Yields extracted from the 1–D fits to B^+ mass distribution (*cf.* Figure 5.1) to both Signal ($B^+ \to \psi K^{*+}$) and Control ($B^+ \to J/\psi K^{*+}$) channel after applying the BDTG cut. Figure of merit (FOM) and BDT efficiency (ε^{sig}_{BDT}) along with preselection efficiency ($\varepsilon^{sig}_{presel}$) for the Monte Carlo are listed in the table.

$$FOM = \frac{N_S}{\sqrt{N_S + N_B}} \tag{5.1}$$

where N_S refers to total number of signal events and N_B refers to total number of background events.

5.2 Determining Efficiency

Efficiencies are calculated on the Monte Carlo sample using the same cut criteria as was done for the data sample. The BDT efficiency for the Monte Carlo is defined as:

$$\varepsilon^{sig}_{BDT} = \frac{N_{mc}|(BDTG > 0.1183\ applied + TruthMatched + Presel.\ cuts\ applied)}{N_{mc}|(Truth\ Matched)} \tag{5.2}$$

where the numerator refers to the number of Monte Carlo events passing the 3 fold cut criteria and the denominator refers to the number of Truth Matched Monte Carlo events. A very similar expression for preselection efficiency for the Monte Carlo also exist, which is defined as:

$$\varepsilon^{sig}_{BDT} = \frac{N_{mc}|(TruthMatched + Presel.\ cuts\ applied)}{N_{mc}|(Truth\ Matched)} \tag{5.3}$$

where this time no BDTG cut is applied to the Monte Carlo sample.

5.3 Computing Branching Ratio

The Branching Ratio (BR) for the control channel is $\mathcal{BR}(B^+ \to J/\psi\pi^+K_S^0)^{PDG} = (1.44 \pm 0.08) \times 10^{-3}$ [4] while that of the Signal Channel is $\mathcal{BR}(B^+ \to \psi\pi^+K_S^0)^{PDG} = (6.7 \pm 1.4) \times 10^{-4}$ [4]. The Branching ratio's of the two channels is related through:

$$\frac{\mathcal{BR}(B^+ \to \psi K^{*+})}{\mathcal{BR}(B^+ \to J/\psi K^{*+})} = \frac{\varepsilon^{cc}_{presel}}{\varepsilon^{sc}_{presel}} \frac{\varepsilon^{cc}_{BDT}}{\varepsilon^{sc}_{BDT}} \frac{N^{sc}_{S|BDT}}{N^{cc}_{S|BDT}} \tag{5.4}$$

where superscript "sc" is used to denote signal channel and superscript "cc" is used to denote control channel or reference channel. The calculation of the BR is shown below:

$$\mathcal{BR}(B^+ \to \psi K^{*+}) = \mathcal{BR}(B^+ \to J/\psi K^{*+}) \frac{\varepsilon^{cc}_{presel}}{\varepsilon^{sc}_{presel}} \frac{\varepsilon^{cc}_{BDT}}{\varepsilon^{sc}_{BDT}} \frac{N(B^+ \to \psi K^{*+})}{N(B^+ \to J/\psi K^{*+})} \tag{5.5}$$

$$\mathcal{BR}(B^+ \to \psi K^{*+}) = (1.44 \pm 0.08) \times 10^{-3} \left(\frac{0.945}{0.919}\right) \left(\frac{0.658}{0.575}\right) \left(\frac{1046 \pm 33}{17402 \pm 163}\right)$$

$$\mathcal{BR}(B^+ \to \psi\pi^+K_S^0) = (1.02 \pm 0.01) \times 10^{-4} \tag{5.6}$$

The reported value in Equation 5.6 is corrected by a factor $\mathcal{R}$, where $\mathcal{R}$ is the ratio of $\mathcal{BR}$ for the decay $\psi(2S) \to \mu^+\mu^-$ with respect to $J\psi(1S) \to \mu^+\mu^-$. As it turns out that decay of J/ψ into two muons is 7.14 times more likely than ψ decaying into two muons. Therefore $\mathcal{BR}(B^+ \to \psi K^{*+})$ now becomes:

$$\mathcal{BR}(B^+ \to \psi K^{*+}) = (1.02 \pm 0.01) \times 10^{-4} \times \mathcal{R} \tag{5.7}$$

$$\mathcal{BR}(B^+ \to \psi K^{*+}) = (7.14 \pm 0.01) \times 10^{-4} \tag{5.8}$$

where 0.01×10^{-4} is the sum of statistical and systematic uncertainty.

Outlook & Final Conclusion

Some of the sources of errors are listed below which might have affected the analysis one way or another:

- The final fit on the control channel data sample was not nearly as good as it was for the signal channel data sample (*cf.* Figure 5.1). The lower mass spectrum of the B^+ incase of control channel is not perfectly fitted. This is due to partially reconstructed background events caused due to mismodelling.
- Monte Carlo samples that were used were not perfect as there can be discrepancy between a real signal event and a Monte Carlo event.
- We might have failed to consider some background that might creep into the data sample.
- More input variables could have been used to train the BDT to get a more robust classifier output.
- In the analysis Run 2 dataset was not considered, which could have yielded more accurate results
- Systematic uncertainties were not considered in the analysis which might have yielded better results.
- Material in the detector is not perfectly described by Monte Carlo.

The analysis was mainly focused on in computing the Branching Ratio for the decay channel $B^+ \to \psi(2S)K^{*+}$. As this decay channel closely resembles the one with J/ψ resonance, the decay channel $B^+ \to J/\psi(2S)K^{*+}$ was therefore used in the analysis as a reference channel. The Branching Ratio, $\mathcal{BR}(B^+ \to \psi K^{*+})$ was calculated to be $(7.14 \pm 0.01(stat. + sys.)) \times 10^{-4}$ where the first uncertainty is statistical, the second is systematic. The results were found to be compatible with the the previous measurements.

.

Bibliography

[1] B. Aubert *et al.*, "Measurements of branching fractions and charge asymmetries for exclusive b decays to charmonium," *Physical Review Letters*, vol. 94, no. 14, Apr 2005. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.94.141801

[2] G. F. Guidice, *A Zeptospace Odyssey*. Oxford University Press, 2010.

[3] G. Kane, *Modern Elementary Particle Physics*, 2nd ed. Cambridge University Press, 2017.

[4] P. Zyla *et al.*, "Particle data group," 2020. [Online]. Available: HTTP://PDG.LBL.GOV

[5] M. Thomson, *Modern Particle Physics*. Cambridge University Press, 2017.

[6] A. D. Sakharov, "Baryon asymmetry of the universe," *Sov. Phys. Usp.*, no. 34417, 1991.

[7] H. Schopper, *Particle Physics Reference Library*, H. Schopper, Ed. Springer, 2020, vol. 1.

[8] T. Skwarnicki, "Decays of b quark," *hep-ph*, 1995. [Online]. Available: https://arxiv.org/abs/hep-ph/9512395v1

[9] K. Lingel, T. Skwarnicki, and J. G. Smith, "Penguin decays of b mesons," *Annual Review of Nuclear and Particle Physics*, vol. 48, no. 1, pp. 253–306, Dec 1998. [Online]. Available: http://dx.doi.org/10.1146/annurev.nucl.48.1.253

[10] D. Mitzel, "Search for new physics in rare four-body charm decays at lhcb," Ph.D. dissertation, University of Heidelberg, Demeber,2018.

[11] A. Horvath. Lhc (large hadron collider) and preaccelerator. File:LHC.svg. [Online]. Available: https://commons.wikimedia.org/wiki/File:LHC.svg

[12] O. S. Bruning, P. Collier, and et al., "LHC Design Report Vol.1: The LHC Main Ring," 6 2004.

[13] M. Adinolfi, G. Aglieri Rinella, E. Albrecht, T. Bellunato, S. Benson, T. Blake, C. Blanks, S. Brisbane, N. H. Brook, and et al., "Performance of the lhcb rich detector at the lhc," *The European Physical Journal C*, vol. 73, no. 5, May 2013. [Online]. Available: http://dx.doi.org/10.1140/epjc/s10052-013-2431-9

[14] L. Collaboration, "The LHCb Detector at the LHC," *JINST*, vol. 3, no. LHCb-DP-2008-001. CERN-LHCb-DP-2008-001, p. S08005, 2008, also published by CERN Geneva in 2010. [Online]. Available: https://cds.cern.ch/record/1129809

[15] M. Adinolfi, G. Aglieri Rinella *et al.*, "Performance of the lhcb rich detector at the lhc." *Eur Phys J C Part Fields*, vol. 73, no. 5, p. 2431, 2013, LHCb RICH, Collaboration.

[16] E. Nappi, "Rich detectors," CERN, Switzerland and INFN, Sez. Bari, Bari, Italy, 1999, Lecture Notes.

[17] R. Aaij *et al.*, "Lhcb detector performance," *International Journal of Modern Physics A*, vol. 30, no. 07, p. 1530022, Mar 2015.

[18] ——, "Performance of the lhcb vertex locator," *JINST 9 P09007*, 2014.

[19] Barbosa-Marinho *et al.*, *LHCb VELO (VErtex LOcator): Technical Design Report*, ser. Technical Design Report LHCb. Geneva: CERN, 2001 2001.

[20] R. Aaij *et al.*, "Lhcb detector performance," *International Journal of Modern Physics A*, vol. 30, no. 07, p. 1530022, Mar 2015.

[21] B. Storaci. Decay-tree-fitter. LHCb twiki.

[22] W. D. Hulsbergen, "Decay chain fitting with a kalman filter," *Nucl. Instrum. Meth. A*, vol. 552, pp. 566–575, 2005. [Online]. Available: https://arxiv.org/pdf/physics/0503191.pdf

[23] M. Pivk and F. Le Diberder, "A statistical tool to unfold data distributions," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 555, no. 1-2, pp. 356–369, 2005. [Online]. Available: https://arxiv.org/pdf/physics/0402083.pdf

[24] S. Das, "A simple alternative to the crystal ball function," 2016. [Online]. Available: https://arxiv.org/abs/1603.08591

[25] M. Borsato, M. D. Cian, D. Gerick, and R. Kopecna, "Angular analysis of the decay $b^+ \to k^{*+}\mu^+\mu^-$ with $k^{*+} \to k^0_s\pi^+$ using 9.1 fb^{-1}," April 2020, LHCb-ANA-2018-022.

[26] A. Hoecker, P. Speckmayer *et al.*, "Toolkit for multivariate data analysis with root," p. 158, March 2017.

[27] R. Aaij *et al.*, "Measurement of the branching fraction and cp asymmetry in $b^+ \to j/\rho^+$ decays," *Eur. Phys. J. C*, p. 537, 2018.

[28] G. Marsaglia, W. W. Tsang, and J. Wang, "Evaluating kolmogorov's distribution," *Journal of Statistical Software*, Nov 2010. [Online]. Available: https://www.jstatsoft.org/article/view/v008i18

Appendix A

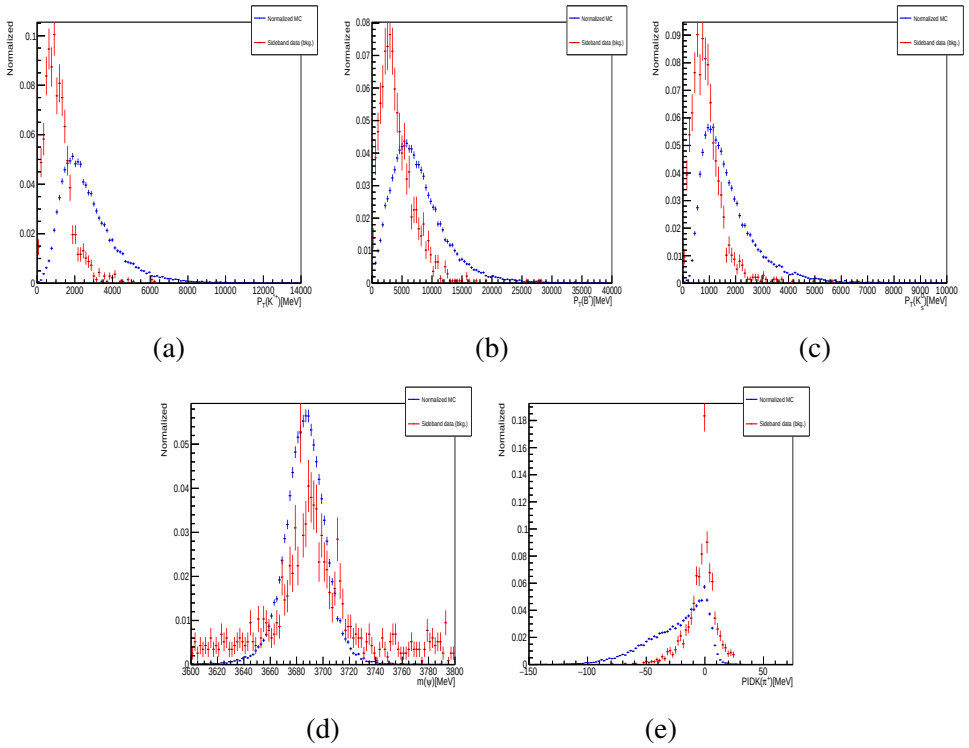


Figure 5.2: Data sideband taken from 5600 to 6000 MeV B^+ mass range and compared with the Truth Matched Monte Carlo for the reference channel: $B^+ \to \psi K^{*+}$

Appendix A

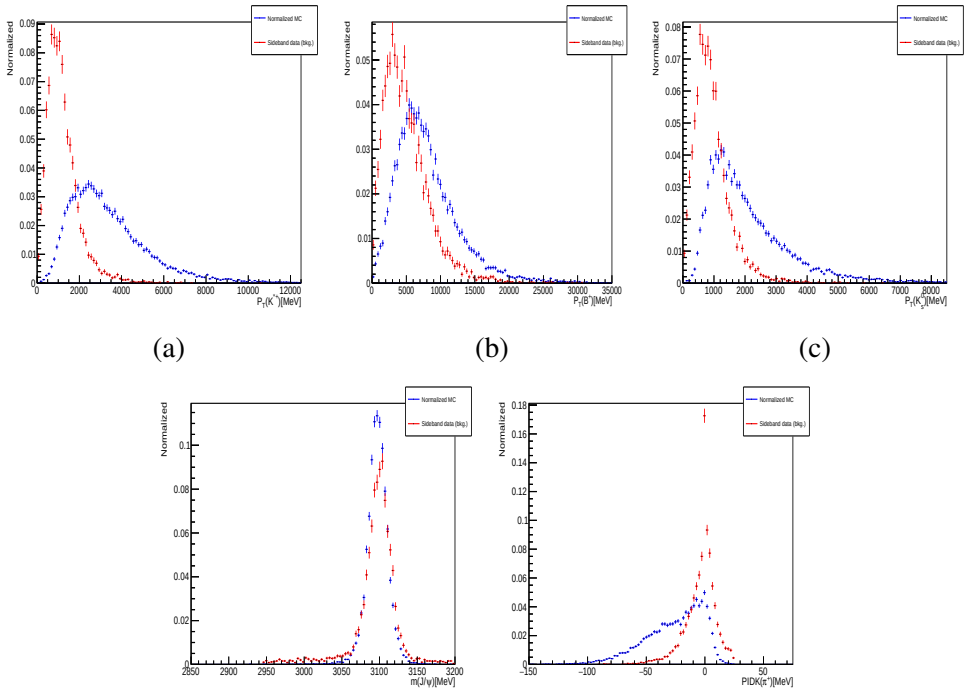
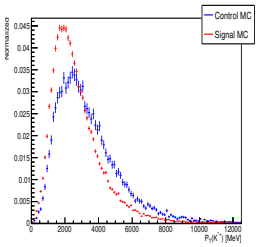
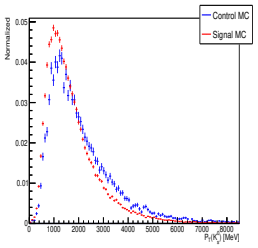
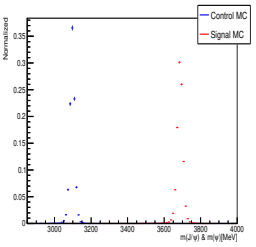


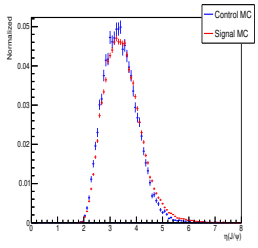
Figure 5.3: Data sideband taken from 5600 to 6000 MeV B^+ mass range and compared with the Truth Matched Monte Carlo for the reference channel: $B^+ \to J/\psi K^{*+}$

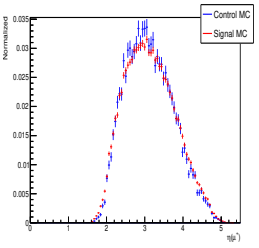
Appendix B

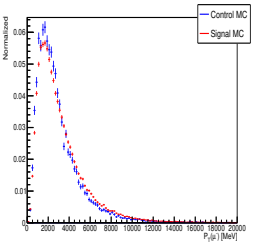

Control MC
Signal MC
normalized
$P_T(K^+)$ [MeV]

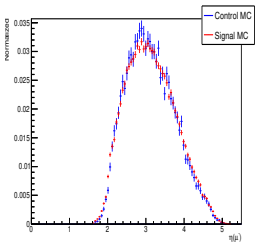

Control MC
Signal MC
Normalized
$P_T(K^+_s)$ [MeV]

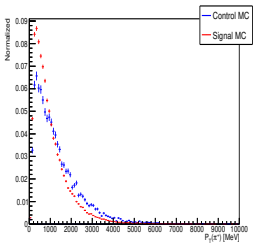

Control MC
Signal MC
Normalized
m(J/ψ) & m(ψ)[MeV]

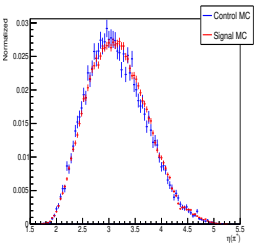

Control MC
Signal MC
Normalized
η(J/ψ)

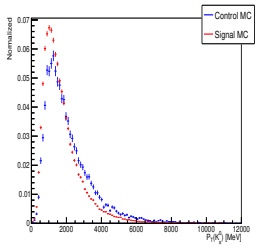

Control MC
Signal MC
Normalized
η(μ⁻)

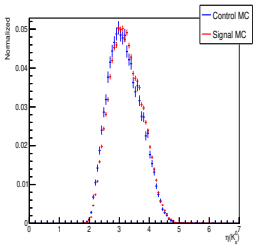

Control MC
Signal MC
Normalized
$P_T(\mu^-)$ [MeV]

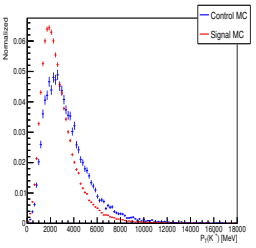

Control MC
Signal MC
normalized
η(μ)

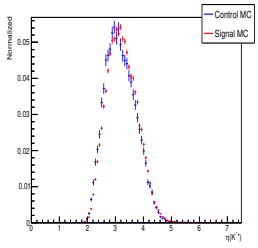

Control MC
Signal MC
Normalized
$P_T(\pi^+)$ [MeV]

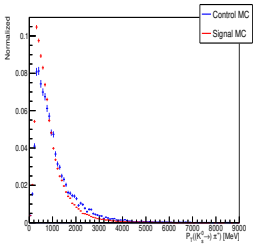

Control MC
Signal MC
Normalized
η(π⁺)

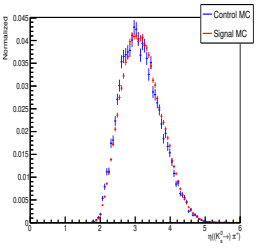

Control MC
Signal MC
Normalized
$P_T(K^+_s)$ [MeV]


Control MC
Signal MC
Normalized
η(K⁺ₛ)

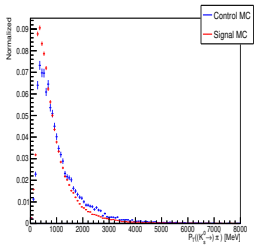

Control MC
Signal MC
Normalized
$P_T(K^+)$ [MeV]

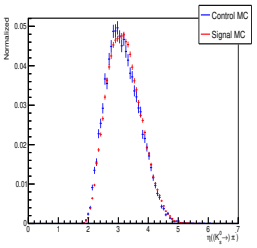

Control MC
Signal MC
Normalized
η(K⁺)

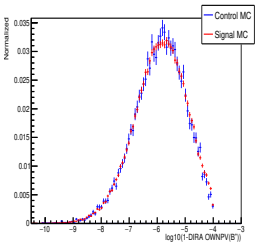

Control MC
Signal MC
Normalized
$P_T(K^+_s \to \pi^+)$ [MeV]

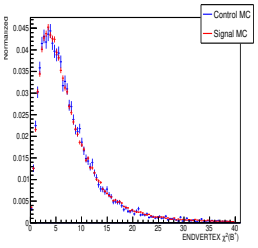

Control MC
Signal MC
Normalized
η(K⁺ₛ → π⁺)

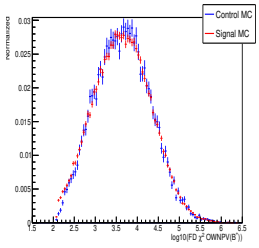
Appendix B

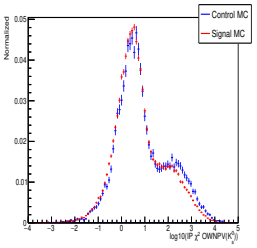

Control MC
Signal MC
Normalized

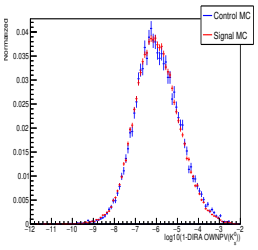

Control MC
Signal MC
Normalized

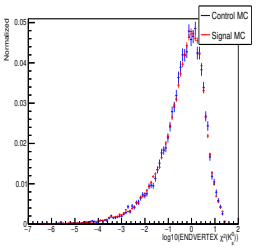

Control MC
Signal MC
Normalized

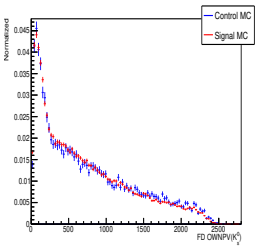

Control MC
Signal MC
Normalized

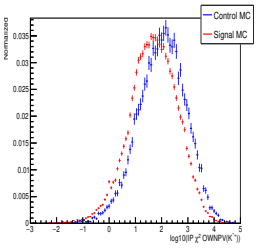

Control MC
Signal MC
Normalized

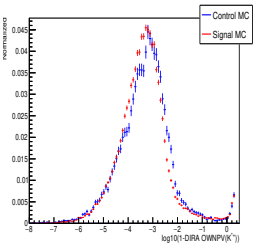

Control MC
Signal MC
Normalized

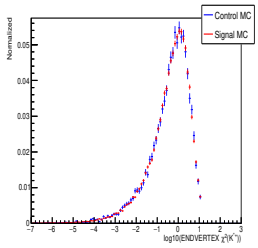

Control MC
Signal MC
Normalized

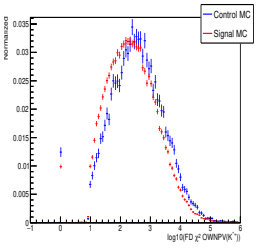

Control MC
Signal MC
Normalized

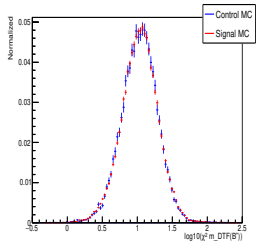

Control MC
Signal MC
Normalized

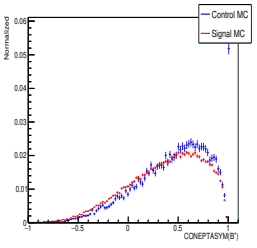

Control MC
Signal MC
Normalized


Control MC
Signal MC
Normalized


Control MC
Signal MC
Normalized


Control MC
Signal MC
Normalized


Control MC
Signal MC
Normalized


Control MC
Signal MC
Normalized

Control MC
Signal MC
Normalized

Control MC
Signal MC
Normalized

Control MC
Signal MC
Normalized

Appendix B

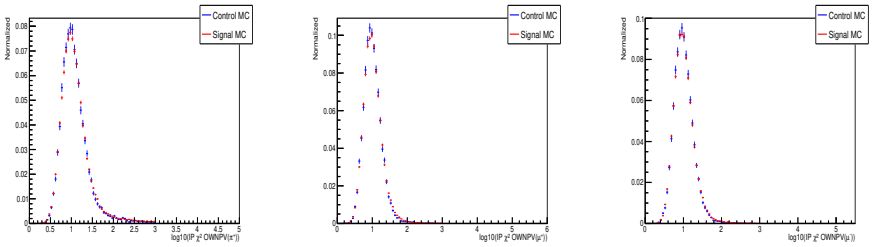


Figure 5.6: Comparison of Monte Carlo events for the reference and signal channel.

Acknowledgements

Firstly I would like to start by thanking my supervisor Prof. Ulrich Uwer for giving me the opportunity to work in his group. Over the past year and a half he has been rather patient with me as I trudged along through the muddy waters. His insightful remarks helped me immensely. This thesis would not have come to fruition had it not been for him, allowing me to work as a master student, for which I am truly grateful.

Next I would like to thank Jascha, a friend in disguise and a through gentleman whose expertise and knowledge got me through to the finish line. There were too many instances to count where his support and understanding helped me overcome seemingly difficult situations.

I would also like to thank LHCb group members at Physikalisches Institut for numerous cake invitations, even though I was trying to cut back on sugar but thats a story for another day. And also Benny for sharing a few laughs during the breaks.

Finally I would like to thank my parents, my younger sister and Nupur for always being there for me during the times I needed them the most.

Eigenständigkeitserklärung:

Heirmit bestätige ich, dass ich die vorliegende Arbeit selbständig verfasst habe und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Unterschrift . . . . . .