Universidade de Santiago de Compostela

Facultade de Física
Departamento de Física de partículas

Traballo de fin de grao

Search for the decay $B_s^0 \to \phi \overline{K}^{*0}$

Author: Sofía Otero Ugobono
Tutor: Abraham Antonio Gallas Torreira

Abraham Antonio Gallas Torreira, Profesor do Departamento de Fśica de Partćulas da Universidade de Santiago de Compostela,

INFORMA:

Que o presente traballo, titulado **Busca da desintegración** $B_s^0 \to \phi \overline{K}^{*0}$, foi realizado por **Sofía Otero Ugobono** baixo a súa dirección no Departamento de Física de Partículas da Universidade de Santiago de Compostela como Traballo de fin de Grao.

E para que así conste, asina en Santiago de Compostela, 4 de setembro de 2014.

Abraham Antonio Gallas Torreira

Resumo

Neste traballo realízase un estudo da canle de desintegración $B_s^0 \to \phi\overline{K}^{*0}$ cos datos recollidos polo experimento LHCb do CERN nos anos 2011 e 2012. Cunha luminosidade de 1 fb^{-1} e unha enerxía no centro de masas de 7 TeV no 2011 e 2 fb^{-1} e 8 TeV no 2012, a cantidade de sucesos desta desintegración medidos foi de $39,5 \pm 8,2$ en 2011 e 55 ± 11 en 2012, cunhas significancias de $5,70\sigma$ e $5,52\sigma$, respectivamente. Ademais, levouse a cabo o cálculo do cociente de ramificación da desintegración e da razón entre este e o cociente de ramificación da desintegración $B_d^0 \to \phi K^{*0}$, obtendo $\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0}) = (0,88 \pm 0,15) \cdot 10^{-6}$ e $\frac{\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})}{\mathcal{B}(B_d^0 \to \phi K^{*0})} = 0,088 \pm 0,014$. Estes resultados compáranse cos valores establecidos polo modelo estándar e cos obtidos anteriormente pola colaboración do LHCb. Esta é a mellor medida desta magnitude feita ata agora, considerando unicamente as incertezas de orixe estatístico.

Resumen

En este trabajo se realiza un estudio del canal de desintegración $B_s^0 \to \phi\overline{K}^{*0}$ con los datos recogidos por el experimento LHCb del CERN en los años 2011 y 2012. Con una luminosidad de 1 fb^{-1} y una energía en el centro de masas de 7 TeV en 2011 y 2 fb^{-1} y 8 TeV en 2012, la cantidad de sucesos de esta desintegración medidos fueron $39,5 \pm 8,2$ en 2011 y 55 ± 11 en 2012, con unas significancias de $5,70\sigma$ y $5,52\sigma$, respectivamente. Además, se llevó a cabo el cálculo del cociente de ramificación de la desintegración y de la razón entre este y el cociente de la desintegración $B_d^0 \to \phi K^{*0}$, obteniendo $\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0}) = (0,88 \pm 0,15) \cdot 10^{-6}$ y $\frac{\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})}{\mathcal{B}(B_d^0 \to \phi K^{*0})} = 0,088 \pm 0,014$. Estos resultados se comparan con los valores establecidos por el modelo estándar y por los obtenidos anteriormente por la colaboración LHCb. Ésta es la mejor medida de esta magnitud hecha hasta ahora, considerando únicamente las incertidumbres de origen estadístico.

Abstract

The purpose of this study is analysing the decay channel $B_s^0 \to \phi\overline{K}^{*0}$ using data samples collected by the LHCb at CERN during the 2011 and 2012 runs. With a luminosity of 1 fb^{-1} and a centre of mass energy of 7 TeV in 2011 and 2 fb^{-1} and 8 TeV in 2012, the number of events of this decay was of 39.5 ± 8.2 in 2011 and 55 ± 11 in 2012, with significances of 5.70σ and 5.52σ, respectively. Furthermore, the branching fraction of this decay and the ratio between it and the branching fraction of $B_d^0 \to \phi K^{*0}$ were calculated, resulting in $\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0}) = (0.88 \pm 0.15) \cdot 10^{-6}$ and $\frac{\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})}{\mathcal{B}(B_d^0 \to \phi K^{*0})} = 0.088 \pm 0.014$. These results are compared to those established by the standard model and to those obtained previously by the LHCb collaboration. This is the best measurement of this magnitude achieved so far, taking into account solely the statistical uncertainties.

Contents

1 Introduction

One of the most interesting unsolved problems in physics today is baryon asymmetry, i.e. the asymmetry between matter and antimatter in our universe, which is made exclusively of matter. $\mathcal{CP}$ violation could be key to explain this asymmetry, however, the size of $\mathcal{CP}$ violation predicted by the Standard Model (SM) is not sufficient to account for the observed baryon asymmetry. One way to address this discrepancy is to postulate the existence of new elementary particles and interactions beyond the SM that act as sources of $\mathcal{CP}$ violation. Such particles and interactions could be associated to very high energy scales, which would explain why they have not yet been detected. Consequently, direct detection of this particles and interactions is not in the least simple, nevertheless, these could manifest themselves as small deviations of certain observables from their SM predictions. For this reason the precise study of $\mathcal{CP}$ violation is of the utmost importance. Specifically the analysis of $\mathcal{CP}$ asymmetries in flavour changing neutral current processes represents a crucial test of the SM. Examples of such decays are both channels examined in this study. The main decay channel, the centre of this study, is: $B_s^0 \to \phi\overline{K}^{*0}$. The secondary channel, fundamental to calculate observables associated to the main channel, is: $B_d^0 \to \phi K^{*0}$. What makes these processes so valuable for the search of New Physics[1] are the quark transitions that take place. According to the SM the $\mathcal{CP}$ violation in this transitions should be extremely small, rendering the measurement impossible. Hence, measuring a non-zero value would imply the presence of physics beyond the SM.

The current study is based on data collected by the LHCb experiment at CERN in 2011 and 2012. The objective of this investigation is analysing the decay $B_s^0 \to \phi\overline{K}^{*0}$, determining its branching fraction $\left(\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})\right)$ and the ratio between the branching fractions of the two channels analysed in this study $\left(\frac{\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})}{\mathcal{B}(B_d^0 \to \phi K^{*0})}\right)$. The experimental values obtained are compared to those predicted by the SM and with the only other experimental results published so far [1]. The theoretical values for the branching fraction of the decay $B_s^0 \to \phi\overline{K}^{*0}$ vary slightly according to the calculation method used. Calculations based on the quantum chromodynamics (QCD) factorisation framework predict a value of $(0.4^{+0.5}_{-0.3}) \cdot 10^{-6}$ [2] while in perturbative QCD a value of $(0.65^{+0.33}_{-0.23}) \cdot 10^{-6}$ [3] is obtained. As for the ratio $\frac{\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})}{\mathcal{B}(B_d^0 \to \phi K^{*0})}$, the theoretical value predicted by the SM is $(0.043^{+0.075}_{-0.046})$ [2]. On the other hand, the only experimental values of this magnitudes hitherto published are:

$$\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0}) = \left(1.10 \pm 0.24(stat) \pm 0.14(syst) \pm 0.08\left(\frac{f_d}{f_s}\right)\right) \cdot 10^{-6}$$

$$\frac{\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})}{\mathcal{B}(B_d^0 \to \phi K^{*0})} = 0.113 \pm 0.024(stat) \pm 0.013(syst) \pm 0.009\left(\frac{f_d}{f_s}\right)$$

[1] Physics beyond the SM.

2 The Experiment

2.1 Large Hadron Collider

The Large Hadron Collider (LHC) is the largest particle accelerator in the world. Built at the CERN[2] accelerator complex in Geneva, Switzerland, it is located at a mean depth of 100 m in the 27-km tunnel where the LEP[3] used to be. Two proton beams travel in opposite directions inside the accelerator. Each beam consists of 2808 bunches of $1.15{\cdot}10^{11}$ protons each. With a maximum energy of 7 TeV per proton[4], when the beams collide the centre-of-mass energy ($\sqrt{s}$) of the proton pairs colliding is 14 TeV. For the beams to achieve this energy first a single beam is produced and passed through several pre-accelerators. These are stated below, starting with the first one, which is connected to the proton source, and ending with the last one, just before the particles are injected into the LHC: *LINAC 2, Proton Synchrotron Booster* (PSB), *Proton Synchrotron* (PS), *Super Proton Synchrotron* (SPS). When coming out of the SPS and before entering the LHC the now 450-GeV beam is split into two. Inside the LHC ring both beams are accelerated up to 7 TeV each. By means of superconducting magnets a 8.3 T magnetic field is produced, keeping the circular trajectory of the proton beams. The magnets are placed inside a cryostat of superfluid He at 1.9 K to ensure their proper performance. Just as important for the operation of the LHC as the refrigeration of the magnets, is that the ultra-high vacuum inside the tubes where the beams travel is preserved.

At four locations around the accelerator ring the beams are slightly deviated for the collisions between protons to occur. Located at each of these places of the tunnel are the main experiments: ATLAS (*A Toroidal LHC Apparatus*), CMS (*Compact Muon Solenoid*), ALICE (*A Large Ion Collider Experiment*) and LHCb (*Large Hadron Collider beauty*). ATLAS and CMS are two general-purpose detectors that share the same scientific goals, albeit their design and operation is rather different. Both investigate a wide range of physics, from studying the properties of the Higgs boson to searching for extra dimensions or particles that could make up dark matter. ALICE is designed to study the strong interactions in the quark-gluon plasma. To achieve the required extreme energy densities it is necessary to collide atomic nuclei, the collisions performed for this experiment are Pb-Pb and Pb-p. Finally, the LHCb is oriented specifically to the study of $\mathcal{CP}$ violation[5] and flavour symmetry

[2] *Organisation européenne pour la recherche nucléaire.*

[3] *Large Electron-Positron Collider.*

[4] The LHC was designed to eventually reach those energies, so far in 2011 and 2012 it reached 3.5 TeV and 4 TeV, respectively. The LHC is expected to attain an energy of 6.5 TeV in 2015, after a two-year period of maintenance and upgrades.

[5] $\mathcal{CP}$ violation is the non-invariance of fundamental interactions under the combined transformation of charge conjugation ($\mathcal{C}$) and parity ($\mathcal{P}$). Charge conjugation implies the transformation of a particle into its antiparticle, whilst parity transformation entails the inversion of the spatial coordinates of the given particle. Without $\mathcal{CP}$ violation matter and antimatter would behave in the same way. It should be noted that the only fundamental force that violates $\mathcal{CP}$ symmetry is the weak force.

breaking[6] mainly in b-hadron decays.

2.2 LHCb

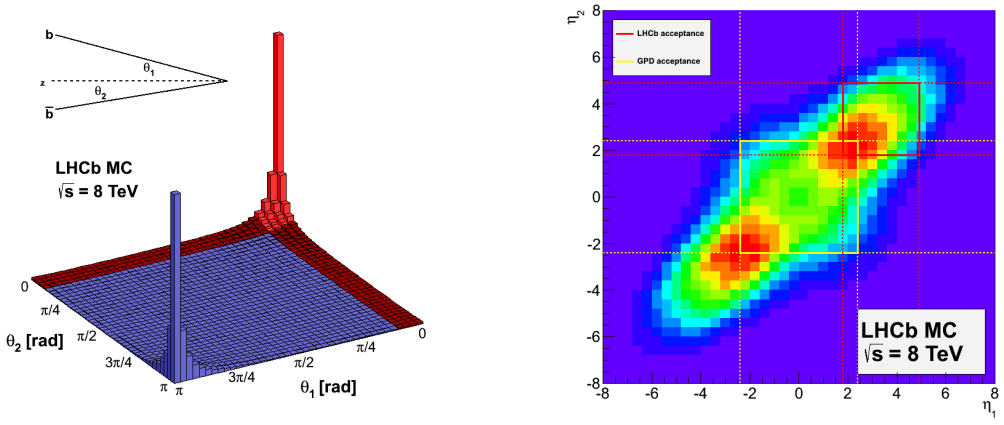
This detector has an specific purpose, the study of rare decays and $\mathcal{CP}$ violation of B mesons. B mesons are those made up of a quark b and a quark u, d, s or c. The aim of this study is to shed light on the existent asymmetry between matter and antimatter in our universe. This type of analysis requires the precise reconstruction of particle tracks. Especially necessary is the precise reconstruction of the primary vertex (PV), i.e. the point of collision between protons to form a b hadron, and the secondary vertex (SV), i.e. the point where the b hadron decay takes place. Since the proton-proton collisions occur in the intersection of two bunches, there is a possibility that multiple collisions happen almost simultaneously making them indistinguishable to the spectrometer. As a consequence, the greater the number of PVs the greater the difficulty of identifying them. Thus the increase in luminosity, although positive in the sense that more b hadrons are produced, has a negative impact in the precise detection of vertices. In order to achieve an optimum relation between the amount of PVs and the ability to discriminate between them the nominal luminosity of the LHCb must be smaller than that in the LHC. Specifically in 2011 the integrated luminosity was of 1 fb^{-1} an in 2012 of 2 fb^{-1}.

2.2.1 Detector characteristics

The LHCb is a single-arm spectrometer with an angular coverage from 10 mrad to 300 mrad in the horizontal plane and to 250 mrad in the vertical plane. Unlike the other detectors in the LHC the LHCb does not surround the beam but covers a low-angle region. This is due to the fact that $b\bar{b}$ pairs are predominantly produced within a low-angle forwards-and-backwards cone (see Figure 1a).

Figure 2 shows a diagram of the LHCb spectrometer, formed by a series of different sub-detectors indicated and briefly described below. Before, it should be noted that one of the most important pieces of the spectrometer is the magnet. The purpose of this magnet is curving the trajectories of charged particles so that their momenta can be measured. The total magnetic field produced aught to be of 4 Tm to attain a 0.4% precision for momenta of up to 200 GeV.

[6]Flavour-changing processes. Flavour is a property that distinguishes different species of particles. In the SM there are 6 quark flavours (*up* (u), *down* (d), *strange* (s), *charm* (c), *beauty* or *bottom* (b) and *truth* or *top* (t)).



(a) (b)

Figure 1: (1a) b and $\overline{b}$ production angle. (1b) Comparison between detector acceptances for $b\overline{b}$ pairs in the LHCb and in general-purpose detectors (GPD).

- **VELO** *(Vertex Locator)***:** this detector allows for the precise detection of primary and secondary vertices. Due to the importance of this detector for the experiment, a full description of it can be found in section 2.2.3.

- **RICH** *(Ring Imaging Cherenkov)* **System:** this system comprises two different detectors: RICH-1 (before the magnet, Figure 2) and RICH-2 (after the magnet). Through the Cherenkov effect these detectors can measure the velocity of particles. Joining the value for the velocity and the momentum (obtained thanks to the tracking system) it is possible to calculate the mass of the particle. Knowing the mass and charge of the particles it is possible to identify them. The RICH detectors work by registering emissions of Cherenkov radiation. The emitted radiation forms a cone of light and its shape is related to the particle's velocity.

- **Tracking System:** consistent of one detector located upstream of the LHCb dipole magnet, the Tracker Turicensis (TT), and three downstream of the magnet (Tracking Stations T1, T2, T3). This configurations permits the measurement of the position of particles in several points along their trajectory, which makes it possible to determine the radius of curvature of the path caused by the presence of a magnetic field. The value of the radius is used to calculate the momenta of the particles and the sense of rotation to determine the sign of the charge.

  Two detector technologies are employed: the TT and the Inner Tracker[7] (IT), formed by the area around the beam pipe in stations T1-T3, use silicon microstrip detectors; the Outer Tracker, which covers the largest fraction of sensitive area in stations T1-T3, is made up of straw-tube drift chambers.

[7]Due to the detection technology employed, the combination of the TT and the IT is called Silicon Tracker.

- **Calorimeter System:** the calorimeter system possesses several layers: the Scintillating Pad Detector (SPD), the Pre-Shower Detector (PS), the Electromagnetic Calorimeter (ECAL) and the Scintillating Tile Iron Plate Hadron Calorimeter (HCAL). Each one of these sub-detectors fulfils a specific function, as a whole they enable the identification of electrons and hadrons as well as the measurement of their energy and location.

  The SPD and PS detectors design is identical, nonetheless, their tasks are different and so is their electronics. Since there is a lead layer of 12 mm the SPD can identify whether or not the incident particles are electrically charged. On the other hand, the PS determines the electromagnetic character of the particles.

  Meanwhile, the ECAL has among its functions the accurate reconstruction of the kinematic parameters of neutral particles and the measurement of the transverse momenta of electrons, photons and π^0. Lastly, the HCAL measures the transverse momentum of hadrons.

- **Muon Detection System:** the interest in muon detection lies in the presence of muons in the final states of plenty of B-meson decays, as well as in the ruling out of their presence in the case of muon-free decays. The muon detection system is composed of five rectangular stations, each one slightly bigger than the previous and covering a combined area of 435 m^2. Each station contains chambers filled with a combination of three gases: CO_2, Ar and CF_4. The passing muons ionise the gas atoms causing a cascade of ionisation which is then collected by the array of wires inside the chamber. The current produced through those wires indicates the passage of a muon.

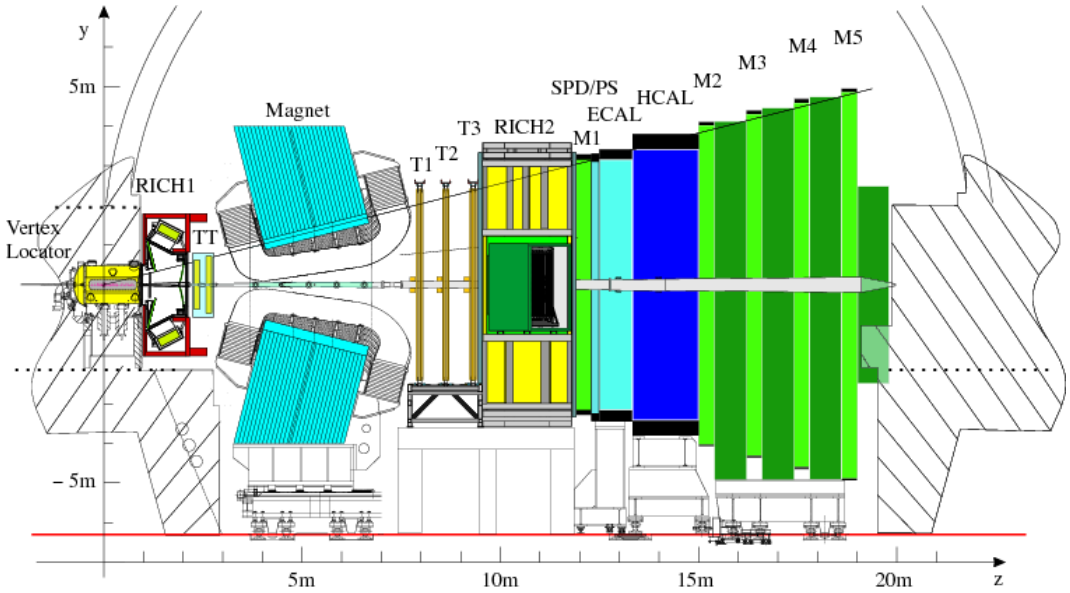


Figure 2: Layout of LHCb detector on the transverse plane

2.2.2 Trigger

The bunch crossing frequency inside the LHCb is of 40 MHz. The frequency of crossings with visible[8] interactions is 10 MHz of which 15 kHz are events with all of their particles coming from a B-meson decay. Furthermore the production rate of interesting B-meson decays is only of about a couple hundred Hz.

The 10 MHz rate of visible interactions is impossible to analyse and save since the storage capacity available is, obviously, not unlimited. Accordingly, it is necessary to reduce the event rate to about 2 kHz (offline computing capacity), but attempting to lose the least amount possible of interesting events. The Trigger is responsible for the reduction of the amount of data recorded. The trigger is organised in two levels, the *Level-0* (L0) and the *High Level Trigger* (HLT):

Level Zero (L0)

Particles from a B decay have a higher transverse momentum with respect to the particle beam axis (p_T) than particles coming directly from the primary proton collision. The L0 exploits this characteristic in order to reduce the event rate to 1 MHz.

To perform this task the L0 collects data from those sub-detectors capable of identifying high-p_T particles (the calorimeters and the muon detection system). In addition, it uses two dedicated silicon layers of the VELO to perform a simplified vertex reconstruction, which allows the rejection of events events with multiple proton-proton interactions.

High Level Trigger (HLT)

The HLT is a software with access to all the information existent about the events that already passed through the L0. It executes a reconstruction algorithm in a farm of 1000 16-core computers. The HLT is divided into two sub-levels: the HLT1 with an output rate of a few tens of kHz and the HLT2 which outputs the 2 kHz that are stored for later analysis.

The HLT1 confirms the candidate particle of the L0 by revising the data from the other sub-detectors, especially from the tracking system and the VELO. This way it is possible to select particles through another one of the characteristic properties of particles from B decays: the significant impact parameter[9] (IP) to the proton-proton interaction vertex. Owing to the fact that B mesons possess a relatively long half-life their flight distance from the p-p collision to the point where their decay occurs is of around 1 cm, resulting in a high IP. If a candidate particle does not meet the requirements it is discarded, this reduces the event rate significantly [10]. This event rate allows the HLT2 to run a complete reconstruction

[8] An interaction is defined as visible when it produces at least two charged particles with sufficient hits in the VELO and T1–T3 so as to reconstruct them.

[9] The impact parameter of a particle to a specific vertex is defined as the minimum distance between the vertex and the direction outlined by the velocity of the particle.

[10] In 2011, for example, the output rate of the L0 was of 870 kHz, while the output for the HLT1 was of 43 kHz.

of the events by using tracks from the VELO as a basis for the rest of the tracking. Another indicative of B decays that is searched for is the occurrence of displaced vertices away from the primary proton-proton interaction. Within the HLT2 two types of selections are applied: one focused on decays of resonances which are useful for calibration, the other aims to provide the highest possible efficiency in the full reconstruction of B decays of interest, using all the available information of the event

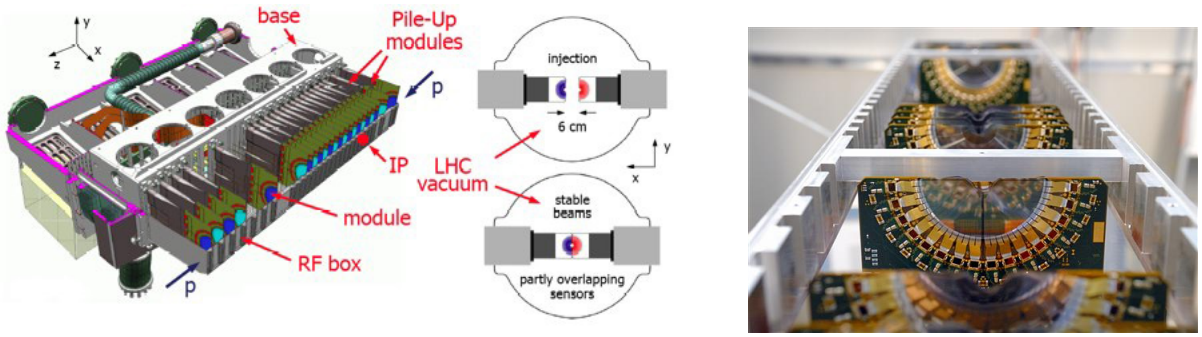
2.2.3 VELO

It is inside the VELO (see scheme in Figure 3a) that the collisions for the LHCb experiment between the LHC beams occur. As mentioned before, the function of the VELO is identifying among all the particles produced in one collision the vertices of production and decay of B mesons. This task is not at all simple given that the half-life of B mesons is rather short, although long in comparison to other particles. This means that the flight of the particle takes place close to the beam, making it exceptionally hard to differentiate this from other particles. Due to its design the VELO has a mean resolution for primary vertices of 42 μm along the beam direction and 10 μm perpendicularly. Also, it is capable of achieving an IP resolution of 20 μm and a flight distance resolution[11] of between 220μm and 370μm.

The VELO features 21 detection stations, each formed by two semicircular silicon sensors, a type R and a type ϕ, with a radius of approximately 42 mm and a thickness of 0.3 mm (see Figure 3b). At the centre of each sensor there is an 8-mm opening that lets the LHC beam circulate unimpeded. Type R sensor measure the radial coordinate via azimuthal strips. Type ϕ sensors measure the angular coordinate using radial strips. Charged particles produced by the collisions traverse the silicon detectors generating electron-hole pairs, the produced current is detected using application-specific electronics.

Since the central opening of the VELO is smaller than the aperture required by the LHC during beam injection, the structure of the sub-detector must be retractable (see Figure 3a). By retracting the detector during injection, VELO-beam interference and beam-caused damages are avoided. Finally, given the proximity to the beam needed the VELO must be inside the vacuum tube of the LHC. To preserve the integrity of the main vacuum system of the LHC, the sensors are separated from the beam by an aluminium foil.

[11] In this case the resolution depends greatly on the decay channel under study.



(a) (b)

Figure 3: (3a) Layout overview of the VELO. (3b) Image of the VELO modules.

3 Physical Motivation

3.1 Interest of the decay channels $B_s^0 \to \phi\overline{K}^{*0}$ and $B_d^0 \to \phi K^{*0}$

As explained in section 1, the consistency of the standard model can be tested through these type of processes. Moreover, being loop-mediated B-meson decays(also known as *penguin*[12] decays), i.e. decays where a quark changes its flavour but not its charge, makes them excellent candidates for the discovery of physics beyond the SM. In particular, for these decays transitions between the third and second generation of quarks[13] ($\overline{b} \to \overline{s}$) and between the third and first generation ($\overline{b} \to \overline{d}$) take place. Since the SM predicts such a tiny $\mathcal{CP}$ violation in the $\overline{b} \to \overline{s}$ transition, its observation would point to new physics beyond the SM.

Feynman diagrams (see Figure 4) yield the decay amplitude for a given decay process. The amplitude is related to the probability of occurrence of said decay and depends significantly on the particles involved in the process. Since the resultant *daughter* particles, product of both decays ($\phi\overline{K}^{*0}$ in the case of the B_s^0 and ϕK^{*0} for the B_d^0), have spin 1, in the final state three different polarization states are available and all of them must be taken into account when calculating the amplitude:

$$A = \sum_k A^{(k)} \tag{1}$$

[12]Term coined by J. R. Ellis to refer to certain Feynman diagrams where a quark changes flavour through a one-loop process.

[13]Both quarks and leptons can be organised into sets of particles called *generations* or *families*. The lightest and most stable particles constitute the first generation while the others form the second and third generations.

In the case of quarks the generations are: $\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix}$

It is worthy of note that quarks are also divided into two types:*u-type*, comprising the u, c and t quarks; *d-type*, comprising the d, s and t quarks.

Three quarks contribute to the internal loop: $\overline{u}$, $\overline{c}$ and $\overline{t}$. Due to the large difference in mass in relation to the rest of the quarks, the $\overline{t}$ quark contribution is dominant. Ergo, the amplitude is given by:

$$A^{(k)}_{B_s^0 \to \phi\overline{K}^{*0}} = A^{(k)}_s V^*_{tb} V_{ts}, \tag{2}$$

$$A^{(k)}_{B_d^0 \to \phi K^{*0}} = A^{(k)}_d V^*_{tb} V_{td}, \tag{3}$$

where $A^{(k)}_s$ and $A^{(k)}_d$ are coefficients that include the kinematic and strong-interaction-related terms. Additionally, the V^*_{tb}, V_{ts} and V_{td} factors are CKM (Cabibbo–Kobayashi–Maskawa) matrix elements, that represent the relation between the quarks in each vertex of the diagram, each with a different flavour. For a detailed description of the CKM matrix, see section 3.2.

The $\mathcal{CP}$-conjugated decay channels would have the same Feynman diagrams changing particles for antiparticles. Meanwhile, the CKM matrix elements are interpreted as representative of transformations. V_{ba} represents the transformation of an initial quark "a" into a final quark "b", while V^*_{ba} represents the transformation of an antiquark "a" into an antiquark "b". Besides, the following relation is satisfied: $V^*_{ba} = V_{ab}$. Taking this into consideration, together with the fact that the coefficients $A^{(k)}_{s,d}$ are invariant under $\mathcal{CP}$ conjugation (except for a factor $\eta^{(k)} = \pm 1$ and a phase), it becomes apparent that: $\left|A^{(k)}_{B_s^0 \to \phi\overline{K}^{*0}}\right| = \left|A^{(k)}_{\overline{B}_s^0 \to \phi K^{*0}}\right|$ and $\left|A^{(k)}_{B_d^0 \to \phi K^{*0}}\right| = \left|A^{(k)}_{\overline{B}_d^0 \to \phi\overline{K}^{*0}}\right|$. This means that the decay probabilities are the same for both $B^0_{s,d}$ mesons and $\overline{B}^0_{s,d}$ mesons, i.e. there is no $\mathcal{CP}$ violation.

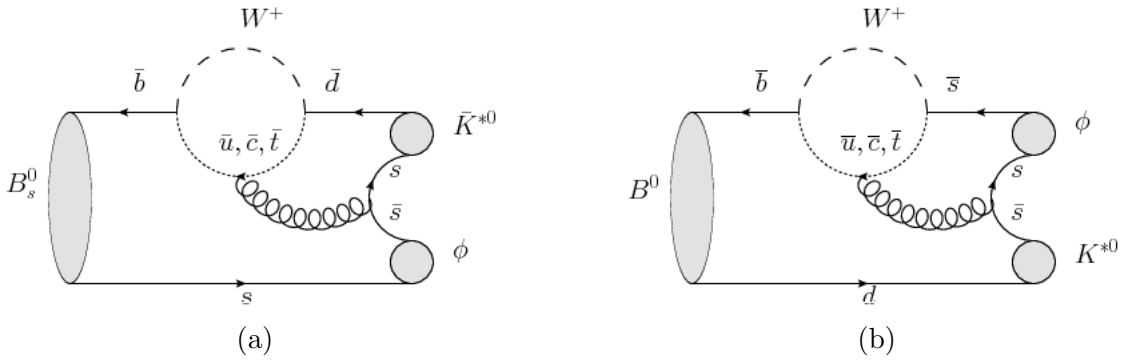


Figure 4: Feynman diagrams for the decays:(4a) $B^0_s \to \phi\overline{K}^{*0}$; (4b) $B^0 \to \phi K^{*0}$.

3.2 CKM matrix

The decays under study in this dissertation are mediated by the weak interaction, since it is the only interaction that allows for flavour transitions between quarks. Depending on the generations involved, the transition is likely to happen or not. The information concerning these probabilities is represented by way of a matrix where each element is directly related to the probability of a quark experiencing a flavour transformation. This matrix is known as the Cabibbo–Kobayashi–Maskawa (CKM) matrix. The most general expression of it is:

$$V_{CKM} \equiv \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \tag{4}$$

Where each element $|V_{ij}|^2$ represents the probability for a u-type quark of flavour j to transform into a d-type quark of flavour i. Flavour transitions show a clear preference for transformations within the same generation. The CKM matrix elements take into account the existent tendencies on flavour transitions, which are directly related to the intensity of the weak interactions between quarks. The latest determinations [4] of the modulus of the CKM matrix elements are listed below (5). Furthermore, a graphical representation of these transition probabilities is presented in Figure 5.

$$|V_{CKM}| = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix} \tag{5}$$

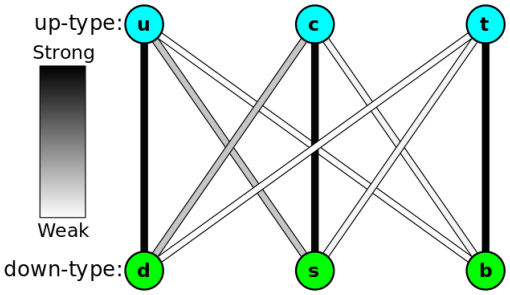


Figure 5: Graphical representation of the transition probabilities between different quark flavours. The line intensity is given by the corresponding CKM matrix element.

3.3 Branching fractions

The branching fraction of a decay ($\mathcal{B}$) is the ratio between the number of particles decaying via a given channel and the total amount of decaying particles. The primary aim of this study is determining the branching fraction for the $B_s^0 \to \phi\overline{K}^{*0}$ and $B_d^0 \to \phi K^{*0}$ decays, in order to compare the experimental values with the theoretical predictions and thus test the order of magnitude of the CKM matrix elements. Further, the value of $\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})$ is obtained in this paper.

3.3.1 $\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})$ calculation

For the purpose of calculating $\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})$, the most accurate value of the branching fraction for the $B_d^0 \to \phi K^{*0}$ is required: $(9.8 \pm 0.6) \cdot 10^{-6}$[4]. Both decays undergo the same selection and share almost identical topologies. Notwithstanding, the different polarizations of the channels may lead to dissimilar angular distributions and therefore affect the detection efficiency. In order to countervail this effect, the following factor is calculated:

$$\lambda_{f_L} = \frac{\epsilon^{B_d^0 \to \phi K^{*0}}}{\epsilon^{B_s^0 \to \phi\overline{K}^{*0}}} = \frac{1 - 0.29 f_L^{B_d^0 \to \phi K^{*0}}}{1 - 0.29 f_L^{B_s^0 \to \phi\overline{K}^{*0}}}, \tag{6}$$

where $\epsilon^{B_d^0 \to \phi K^{*0}}$ and $\epsilon^{B_s^0 \to \phi\overline{K}^{*0}}$ are the reconstruction efficiencies for the $B_d^0 \to \phi K^{*0}$ and $B_s^0 \to \phi\overline{K}^{*0}$ decays, $f_L^{B_d^0 \to \phi K^{*0}}$ and $f_L^{B_s^0 \to \phi\overline{K}^{*0}}$ are the longitudinal polarizations and the factor 0.29 is obtained through simulation. A more detailed calculation of λ_L may be found in [1], the final result being:

$$\lambda_{f_L} = 1.01 \pm 0.06$$

Another factor to be bear in mind is the difference between the production rates of B_s^0 and B_d^0 mesons. The quotient that accounts for this difference is the ratio of hadronisation factors $f_d/f_s = 3.86 \pm 0.22$ [5].

Gathering together all these results, the final expression for the branching fraction is:

$$\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0}) = \lambda_{f_L} \cdot \frac{f_d}{f_s} \cdot \mathcal{B}(B_d^0 \to \phi K^{*0}) \cdot \frac{N_{B_s^0 \to \phi\overline{K}^{*0}}}{N_{B_d^0 \to \phi K^{*0}}}. \tag{7}$$

3.3.2 Relation between $\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})$ and $\mathcal{B}(B_d^0 \to \phi K^{*0})$

As presented earlier in the text, the ratio between branching fractions can be obtained developing the kinematic calculations of the Feynman diagrams. The general expression is:

$$\mathcal{B}(B_q^0 \to VV') = \frac{|V_{tb}^* V_{tq}|^2 \sum_k \int d\phi^{(2)} \left|A_q^{(k)}\right|^2}{2M_{B_q^0}\Gamma_{B_q^0}}, \tag{8}$$

where $\int d\phi^{(2)}$ is the integral in the phase space of V and V'[14], $\Gamma_{B_q^0}$ is the full-width of the meson and $M_{B_q^0}$ its mass. For the B_s^0/B_d^0 ratio to be computable, the values of $A_{s,d}^{(k)}$ must be known, which implies the usage of QCD models. If the $B_s^0 \to \phi\overline{K}^{*0}$ and $B_d^0 \to \phi K^{*0}$ decays were related by a U-spin transformation[15], since the masses of the B_s^0 and B_d^0 are similar, the following approximation would be admissible:

$$\frac{\sum_k \int d\phi^{(2)} \left|A_s^{(k)}\right|^2}{M_{B_s^0}} \approx \frac{\sum_k \int d\phi^{(2)} \left|A_d^{(k)}\right|^2}{M_{B_d^0}}, \tag{9}$$

since the $d \leftrightarrow s$ effect on the phase space would be negligible. Actually, the decay related to B_s^0 through a $d \leftrightarrow s$ transformation is $B_d^0 \to \rho^0 K^{*0}$. However, this is not a pure-penguin decay since it also has a tree-level contribution[16]. This implies that the calculation of the branching fraction is not so simple cannot be described with (8). The approximation in (9) can be used to elude this problem, bearing in mind that it is a very simplified approach, since the assumption of both decays being related through a U-spin transformation is false. The goal is to obtain the order of magnitude rather than the actual value of the ratio. The value obtained through this method is given by eq. (10).

$$\frac{\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})}{\mathcal{B}(B_d^0 \to \phi K^{*0})} \approx \frac{\tau_{B_s^0}}{\tau_{B_d^0}} \frac{|V_{tb}^* V_{td}|^2}{|V_{tb}^* V_{ts}|^2} \approx 0.046 \tag{10}$$

Where the values for τ are those found in the PDG [4]: $\tau_{B_s^0} = (1.516 \pm 0.011)10^{-12}$ s and $\tau_{B_d^0} = (1.519 \pm 0.007)10^{-12}$ s.

[14] V and V' represent vector mesons, i.e. spin-1 mesons.
[15] A U-spin transformation consists in the interchange of s and d quarks.
[16] *Tree* diagrams are those without internal loops.

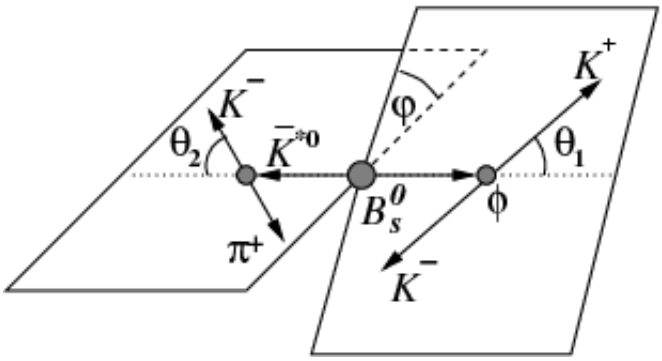


Figure 6: Definition of the angles of distribution of the $B_s^0 \to \phi\overline{K}^{*0}$ decay products.

3.4 Angular distribution

The $B_s^0 \to \phi\overline{K}^{*0} \to (K^+K^-)(K^-\pi^+)$ decay takes place through two spin-1 intermediate particles (vector mesons). The amplitude of this process, and therefore the angular distribution, can be described by three independent components related to the different linear polarization states of the vector mesons: A_0 (longitudinal polarization states), $A_\parallel$ (states with polarizations perpendicular to the momentum direction, but parallel between them) and $A_\perp$ (states with polarizations transverse to the direction of motion and perpendicular between them). These amplitudes can be obtained from the angles between the decay products[17]: θ_1, θ_2 and φ. The convention for the angles can be seen in Figure 6. This Figure also shows how θ_1 (θ_2) is the angle of emission of the K^+ (K^-) meson with respect to the opposite direction of the B_s^0 meson in the ϕ ($\overline{K}^{*0}$) rest frame. Similarly, φ is defined as the angle between the decay planes of $\overline{K}^{*0}$ and ϕ in the B_s^0 rest frame.

By analysing the angular distribution of the decay it is feasible to obtain the values of the polarization fractions (f_j, con $j = L, \parallel, \perp$) and the phase difference between A_0 and $A_\parallel$ ($\delta_\parallel$). One possible approach this study is the one performed in [1], where a flavour-averaged and time-integrated polarization analysis is applied. The analysis was executed under the assumptions that: the $\mathcal{CP}$-violation phase is zero and that the amount of B_s^0 and $\overline{B}_s^0$ mesons produced is the same. The analytic expression used to fit the data is (11), with some extra terms that account for the S-wave and interference contributions.

$$\begin{aligned}\frac{\mathrm{d}^3\Gamma}{\mathrm{d}\cos\theta_1 \mathrm{d}\cos\theta_2 \mathrm{d}\phi} &\propto |A_0|^2 \cos^2\theta_1 \cos^2\theta_2 + \frac{1}{2}\left|A_\parallel\right|^2 \sin^2\theta_1 \sin^2\theta_2 \cos^2\phi \\ &+ \frac{1}{2}|A_\perp|^2 \sin^2\theta_1 \sin^2\theta_2 \sin^2\phi + \frac{1}{2\sqrt{2}}|A_0|\left|A_\parallel\right| \cos\delta_\parallel \sin 2\theta_1 \sin 2\theta_2 \cos\phi \end{aligned} \quad (11)$$

From the fit to the angular distribution of the particles from the decay $B_s^0 \to \phi\overline{K}^{*0}$,

[17] These angles are defined with respect to the helicity frame.

the quantities $\delta_\parallel$, A_0, $A_\parallel$ and $A_\perp$ are obtained. Therefore, it is possible to determine the polarization fractions through:

$$f_j = \frac{|A_j|^2}{(|A_0|^2 + |A_\parallel|^2 + |A_\perp|^2)} \tag{12}$$

This paper does not include an angular distribution analysis, since its complexity surpasses that expected from a Bachelor of Science dissertation. The major physical value of conducting such an analysis in this decay channel leaves the door open to future in-depth studies of it.

4 Event Selection

This investigation is based in the analysis of two data samples collected with the LHCb detector. One from the 2011 data taking period, with an integrated luminosity of 1 fb^{-1} and a centre-of-mass energy of 7 TeV and the other from 2012, corresponding to an integrated luminosity of 2 fb^{-1} and $\sqrt{s} = 8$ TeV. Due to the different LHC working conditions during both runs, the data samples were analysed separately. Finally, the results from both samples are combined to obtain a more accurate measurement.

As mentioned earlier, the first step for selecting the events of interest is the trigger system. Afterwards, the LHCb software automatically applies a series of cuts, called *stripping*, to each channel (see Table (1)). The *stripping* drastically reduces the number of events stored for subsequent analysis, while maintaining a high efficiency for events of interest. Once saved, further cuts must be applied to improve the signal-to-noise ratio. Before delving into the details of the selections made, it is worth enumerating and describing the variables employed.

- **Mass:** Obviously, one of the variables used is the mass of the particles. Through the 4-momenta of the *granddaughter* particles[18] the mass of the *daughters* and the *mother* can be reconstructed.

- $\boldsymbol{p_T}$ **:** This is the transverse momentum, i.e. the momentum of the particle projected onto the axis perpendicular to the beam. It is one of the most important variables since the *daughters* of the B mesons can reach high p_T values due to the large mass of the meson. This property permits the distinction of background events, usually coming from the combination of particles produced at the PV, from events coming from a B-meson decay.

[18] In the case of the B_s^0 meson decay, the *granddaughter* particles are K^+K^- (from the ϕ decay) and $K^-\pi^+$ (from the K^{*0} decay).

- **DOCA:** It stands for Distance of Closest Approach and it is applied to the tracks of the B-meson *daughters* (ϕK^{*0}). It also provides information about the quality of the secondary vertex.

- **Flight distance (ctau):** It is the distance travelled by a particle from the vertex where it was produced to the point where it decays.

- **PID:** This variable is related to particle identification. Specifically, it is a relation between the mass-hypothesis assigned to a track, i.e. the hypothesis of a certain track corresponding to a given particle. This variable is defined as:

$$\mathrm{PID_{hk}} = \ln\left(\frac{\mathcal{L}(h)}{\mathcal{L}(k)}\right) \tag{13}$$

  Where $\mathcal{L}(h)$ and $\mathcal{L}(k)$ are the probabilities for a given track of being particle "h" or "k", respectively. The LHCb detects electrons, photons, kaons, muons and protons, consequently, there is a PID value for each and every one of those particles except for photons.

- **IP and IPS:** The IP variable was already defined when describing the trigger system, where the high-IP property of the B mesons, due to their rather prolonged mean-life, was discussed. This magnitude is defined as the minimum distance between a specific vertex and the direction outlined by the velocity of the particle in question. The IPS or IP/χ^2 is a variant of the impact parameter weighted by the uncertainty of the reconstruction (χ^2).

- **Vertex $\boldsymbol{\chi^2}$ and $\boldsymbol{\chi^2/ndof}$:** The first one is the value of χ^2 for a vertex, while the second one ($\chi^2/ndof$) is the reduced χ^2, i.e. χ^2 over the number of degrees of freedom.

- **DIRA:** It is defined as the cosine of the angle between the momentum of the B meson and the flight direction from the primary to the secondary vertex.

- **Dm and phim:** These two variables are used for mass reconstruction, both help discard events that might be confused with $B_s^0 \to \phi\overline{K}^{*0}$ decays when, actually, they are $B_s^0 \to \phi\phi$ decays. In order to eliminate these events, the following cuts were applied: $|\mathrm{Dm} - 5366.77| > 50$; $|\mathrm{phim} - 1019.455| > 10$.

Table 1: *Stripping* cuts for $B_s^0 \to \phi\overline{K}^{*0}$

	Cuts
All tracks p_T	> 500 MeV
All tracks IPχ^2	> 9
$\text{PID}_{\text{K}\pi}(\text{K}^{\pm})$	> 0
$\text{PID}_{\text{K}\pi}(\pi^{\pm})$	< 10
K^{*0} mass window	$\pm$ 150 MeV
$p_T(\text{K}^{*0})$	> 900 MeV
K^{*0} vertex χ^2	< 9
ϕ mass window	$\pm$ 25 MeV
$p_T(\phi)$	> 900 MeV
ϕ vertex χ^2	< 9
B_s^0 mass window	$\pm$ 500 MeV
B_s^0 DOCA	< 0.3 mm
B_s^0 vertex χ^2/ndof	< 15

4.1 The TMVA package

The automatic cuts applied by the*stripping* do not discard enough background events, meaning an additional selection must be performed. Rectangular cuts can reduce the number of candidates coming from different channels than the one being analysed. However, they are not optimal in the sense that they do not maximise the signal significance[19]. Therefore, a new, more complex, kind of selection must be applied to the variables. In particular, the TMVA package [6] from ROOT is used. It provides tools for processing, evaluating and classifying events using multivariate techniques[20]. From all of the tools that conform the TMVA package, the BDT (*Boosted Decision Trees*) is the one used for this analysis.

The BDT is capable of analysing those variables that help differentiate signal from background and creates a classification based on their behaviour. In order to perform this tasks, the BDT requires two orthogonal input data samples[21], one rich in signal and the other rich in background events. Hence, when analysing a new sample, a new variable can be added to each event indicating the likelihood of that event being part of the signal or the background. This new variable is called BDT and takes values from -1 to 1. Therefore, an event with $BDT = -1$ is without doubt a background event and one with $BDT = 1$ is, indisputably, a signal event.

[19]The signal significance is a parameter that relates signal and background and it is defined as: $\frac{S}{\sqrt{S+B}}$. Where S is the signal and B the background.

[20]The multivariate analysis consists in statistical methods for determining the contributions from several factors to the event or result. This is performed by taking into account, simultaneously, the behaviour of several variables and their correlations.

[21]This means that they do not share any events and that none of them will be subjected to further analysis with the BDT created from this data.

4.1.1 Operating characteristics of a BDT

As explained above, the BDT requires both a signal-rich and a background-rich sample to find the variables useful for distinguishing between signal and background events. A BDT is a classification method based on a binary tree structure (see Figure 7). The program studies every single variable, making them go through several stages were the system analyses whether or not the variables verify a certain condition. The following stage depends on the variable meeting or not the condition. This process continues until the variable satisfies a certain criterion that stops the procedure. Eventually a series of regions containing all the data are formed. Each region is associated to the events verifying certain conditions. As a result, by analysing if the majority of the events in each region correspond to the signal or the background it is possible to create a method of classification. This entire process is called *training* of the BDT.

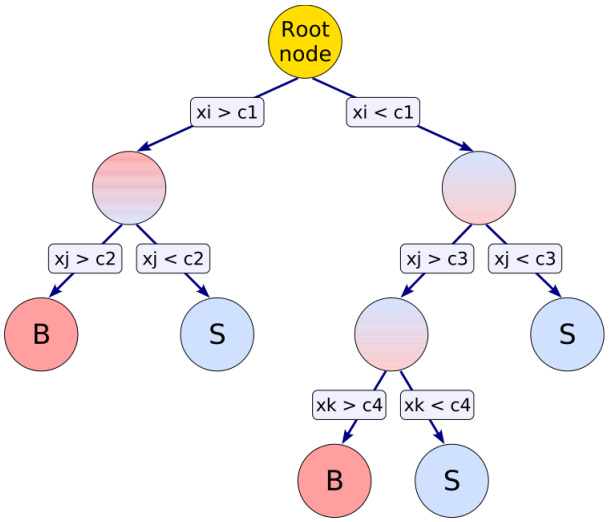


Figure 7: Schematic view of a decision tree.

4.1.2 Application of the BDT

Due to the different LHC working conditions during the 2011 and 2012 runs, separate BDTs were trained and applied for the analysis of data from each run. As mentioned earlier, for the training of the BDT a signal and a background sample are needed. In this paper, the samples used were constituted by real data rather than simulated data. The complete set of data from one year (contained in a type of file called N-tuple) was divided into two N-tuples[22] **A** and **B**. The basic idea is to train a BDT with N-tuple **A** to then apply it to N-tuple **B** and vice versa. This procedure was applied to the 2011 and 2012 data samples separately. The method is described in detail below.

First, a BDT is trained with the data in N-tuple **A**. In order to do so it is necessary to define which of the events in N-tuple **A** belong to the background and which to the signal. All events that fall within the mass window of the B_d ($|M(B_d^0) - 5280| < 30$ MeV) are considered signal, the rest are considered background. At first it may seem odd that the B_s^0 mass window is not included in the definition of signal events, there are two reasons for doing so: 1) using only the B_d^0 signal means profiting from a signal that has almost the exact topology of the B_s^0 and equivalent detection conditions, but avoiding the introduction

[22]The N-tuples were split in two by means of a python script specifically made for this purpose.

of a bias due to the use of the signal itself in the training; and 2) the B_s^0 mass window is background-dominated since the decay is strongly suppressed, thereby it would contaminate the signal sample with background. Once the BDT is trained with the data in **A**, it is applied to **B** obtaining a new N-tuple **B'**. This new N-tuple is the same as **B** but containing the BDT variable. Finally, the process is repeated using **B** as the training data sample and **A** as the *target* sample in order to obtain **A'**.

Before obtaining **A'** and **B'**, the training results must be checked using the graphic tools provided by the TMVA package. Of all the plots produced, 4 of them are particularly significant:

- **Histogram with the linear correlation coefficient of the input variables:**

  The plots resulting from the training of the BDT for the 2012 sample are shown in Figure 8 (the plots for the 2011 data are not shown due to space limitations). The ideal scenario would be for the correlation between different variables to be zero. This, however, does not occur, but they are mostly small correlations (except for a few variables). From this an other tests the program can create a variable ranking based on their correlation, giving more importance to those less correlated.

- **Normalised distribution of both the background and signal as a function of the variable BDT:**

  This plot allows the estimation of the optimal BDT cut value for the majority of the events selected to be signal and not background. Ideally, there would not be an overlap between the signal and background distributions, so it would be possible to find a BDT value that removed all the background events leaving only the signal. In reality this does not happen and one can only aspire to obtain distributions with a small overlapping region and, above all, with perfectly distinguishable and clearly separated maximums. This ensures that the cut applied to the BDT variable will not eliminate too many signal events. The plots obtained through the training of the BDT for the 2012 data are shown in Figure 9 (once again, the plots for the 2011 data are not included due to space limitations).

- **Plot of the significance as a function of the BDT value.**

  This is one the most important plots since it permits to find precisely the BDT cut value that optimises the discrimination between signal and background. Considering that this value is the one that maximises the significance, indicating the expected quantity for both signal and background events, the plot adapts to these values and delivers the maximum value of the significance and the BDT cut value associated to it. One of this plots, obtained with the 2011 data sample, is shown as an example in Figure 10.

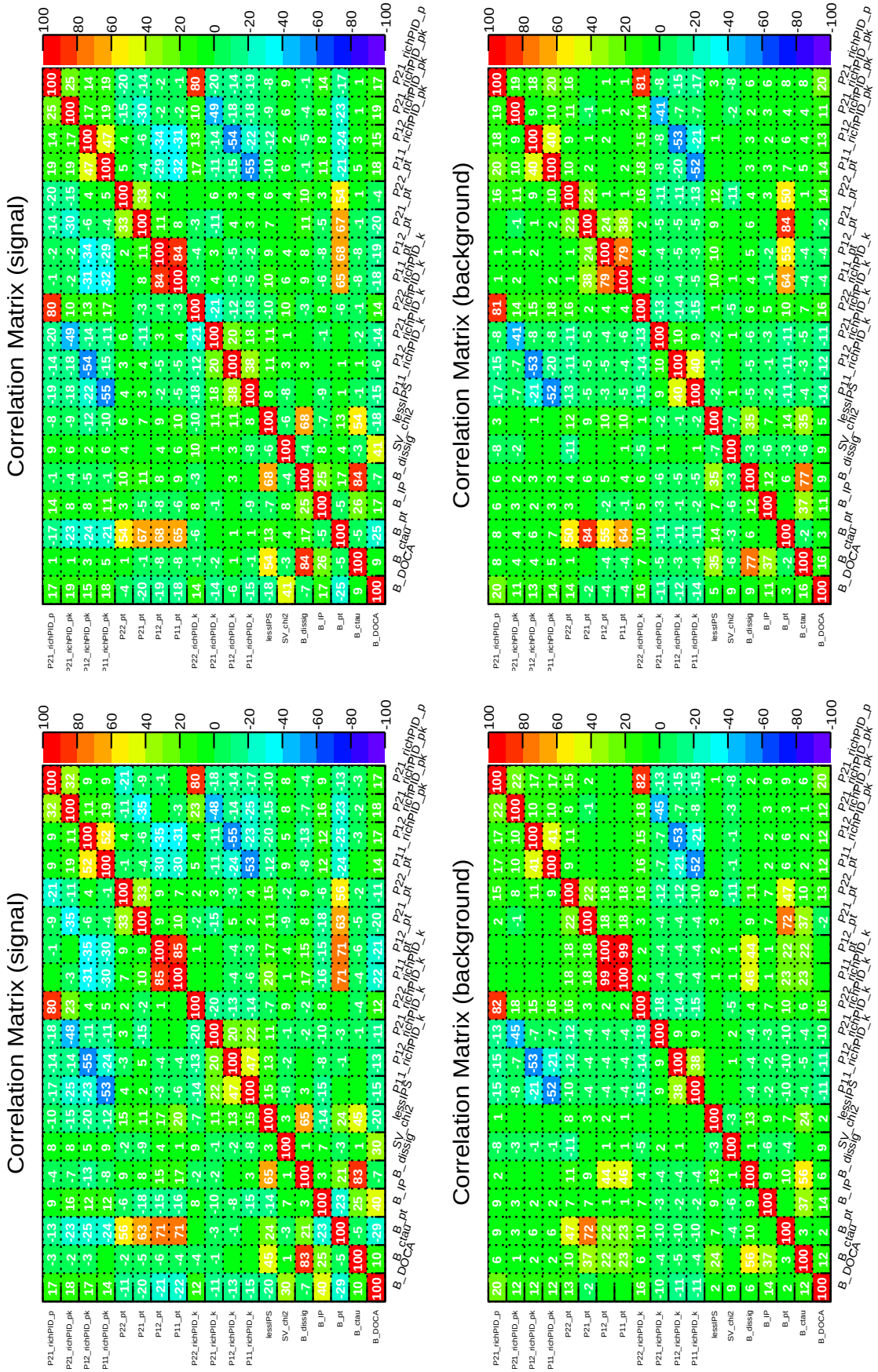
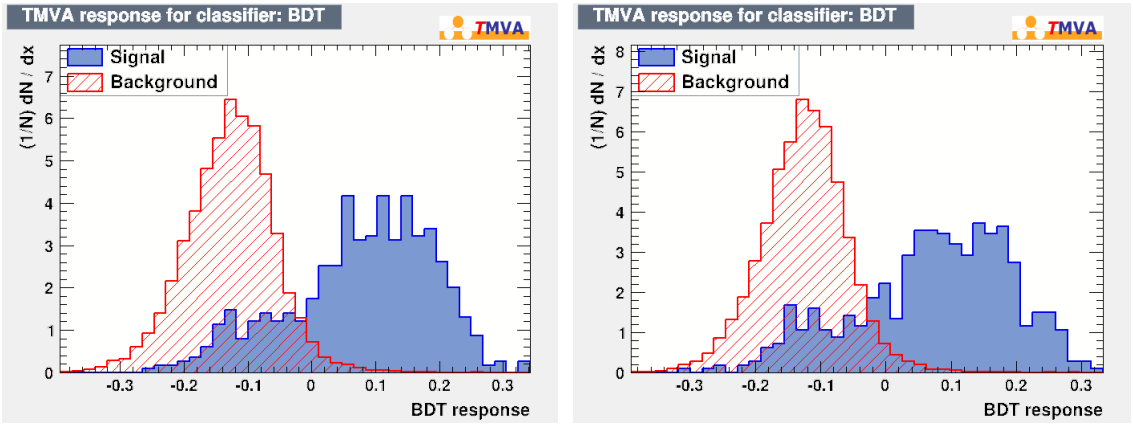


Figure 8: Linear correlation matrices for the signal and background 2012 training samples.



(a) (b)

Figure 9: Normalised distributions of the background and signal events as a function of the BDT value using the 2012 data sample.

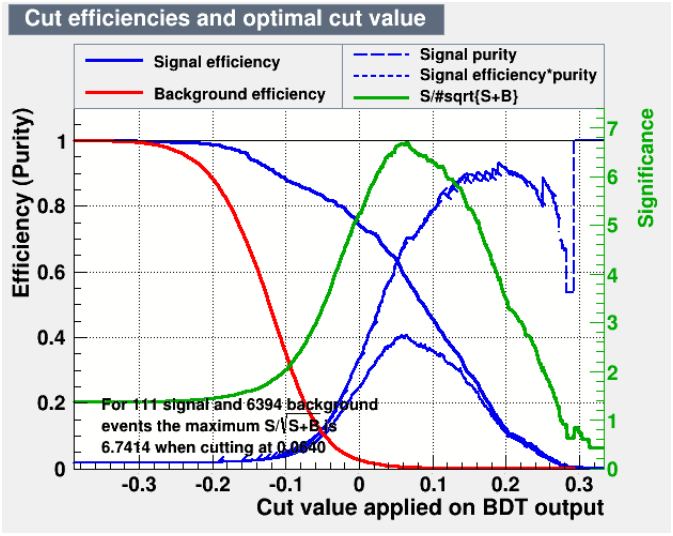


Figure 10: Plot of the significance versus BDT value.

5 Fit Model for the Four-Body Invariant Mass Spectrum

The identification of the signal events is performed through the analysis of the four-body invariant mass spectrum, namely $K^+K^-K^-\pi^+$. In order to fit the invariant mass spectrum, a number functions must be employed so as to fit both the signal and the background. These functions and the reasons for their use are described below.

- ***Crystal-Ball* + Gaussian:**

  A *Crystal-Ball* (eq. (14)) plus a concentric Gaussian distribution are used to fit the B_s^0 and B_d^0 peaks. The *Crystal-Ball* contains most of the statistics. The shape of this distribution is similar to that of a Gaussian, but with a *bremsstrahlung tail* that describes photon emission processes of the final state mesons when interacting with the medium. The Gaussian parametrizes events reconstructed with a lower resolution, i.e. events with particles that have most likely suffered a hard collision in one of the detectors.

$$f(x;\alpha,n,\vec{x},\sigma) = H \cdot \begin{cases} \exp\left(-\frac{(x-\vec{x})^2}{2\sigma^2}\right) & \text{para } \frac{x-\vec{x}}{\sigma} > -\alpha, \\ A \cdot \left(B - \frac{x-\vec{x}}{\sigma}\right)^{-n} & \text{para } \frac{x-\vec{x}}{\sigma} \leq -\alpha, \end{cases} \tag{14}$$

  where

$$A = \left(\frac{n}{|\alpha|}\right)^n \cdot \exp\left(-\frac{|\alpha|^2}{2}\right), \; B = \frac{n}{|\alpha|} - |\alpha|$$

  and the a and n parameters describe the bremsstrahlung tail. Large values of these parameters would indicate that the bremsstrahlung tail is negligible.

- **ARGUS convoluted with a Gaussian:**

  These functions represent partially reconstructed decays. The source of these events may be B_d^0 or $B_u^\pm$ decays to ϕ and K or K_i^* (excited states of K). The excited kaon decays into a pion that was either lost or excluded from the four-body calculation. The explicit representation of the function is:

$$f_P(m) \propto m' \left(1 - \frac{m'^2}{m_0^2}\right) \Theta(m_{PhysBkg} - m') e^{-k_{PhysBkg} \cdot m'} \otimes G(m - m'; \sigma_{PhysBkg}), \tag{15}$$

  where Θ is the Heaviside step function, $\otimes$ represents the convolution between both functions, m' is the variable over which the convolution integral is calculated, $G(m - m'; \sigma_{PhysBkg})$ is a Gaussian p.d.f. (probability density function) with standard deviation $\sigma_{PhysBkg}$, that represents the experimental resolution, and $m_{PhysBkg}$ that is a free parameter.

- **Decreasing exponential:**

$$f_C(m) \propto e^{-r_{Comb} \times m} \tag{16}$$

  This function accounts for the combinatorial background coming from the misidentification of particles.

- ***Toy Monte-Carlo* histogram:**

  To fit those events produced by a $\Lambda_b^0 \to pK\phi$ decay, where the proton has been misidentified as a pion or a kaon, a toy Monte-Carlo is used. The Monte-Carlo creates a histogram that represents the spectrum of the decay. Since it is not a function but a histogram, the only parameter to fit is the number of events.

5.1 Fit results

In order to fit the 2011 and 2012 data it was necessary to obtain the optimal BDT cut value that ensured the most effective discrimination between signal and background. Due to the method employed to construct the BDTs for both years (see section 4.1.2), there are two BDT cut values per year. As a result, the cut value used for each year is equal to the mean between the two values obtained for that specific year. As mentioned earlier, in order to find the BDT cut value, it is necessary to know the number of signal and background events expected before applying the BDT cut. Ideally, there would be a BDT cut value that perfectly separates signal from background. However, this does not happen in the real world and a few signal events are lost, nonetheless, this must be taken into account when optimising the BDT.

The estimation of the expected number of events is very easy to attain. It is just a matter of counting, by using ROOT, how many events there are in a ± 30 MeV window around the B_s^0 mass (782 in 2011 and 6394 in 2012). This can be done because the number of signal events is negligible with respect to the number of background events. It is not that simple, though, to calculate the number of signal events expected. In order to obtain this value it is necessary to take into consideration the fact that the ratio between the amount of $B_d^0 \to \phi K^{*0}$ and $B_s^0 \to \phi\overline{K}^{*0}$ decays must be approximately constant. Knowing this quantity and the amount of events in the B_d^0 peak of the studied sample, it is possible to determine the number of B_s^0 events expected. The results obtained in the paper announcing the first observation[23] of the $B_s^0 \to \phi\overline{K}^{*0}$ decay [1] are shown below[24]:

$$N_s = 30 \pm 6 \; ; \; N_d = 1000 \pm 32 \; ; \quad \text{with a significance of: } 6.2\sigma \tag{17}$$

The general expression for the calculation of the amount of signal events expected is:

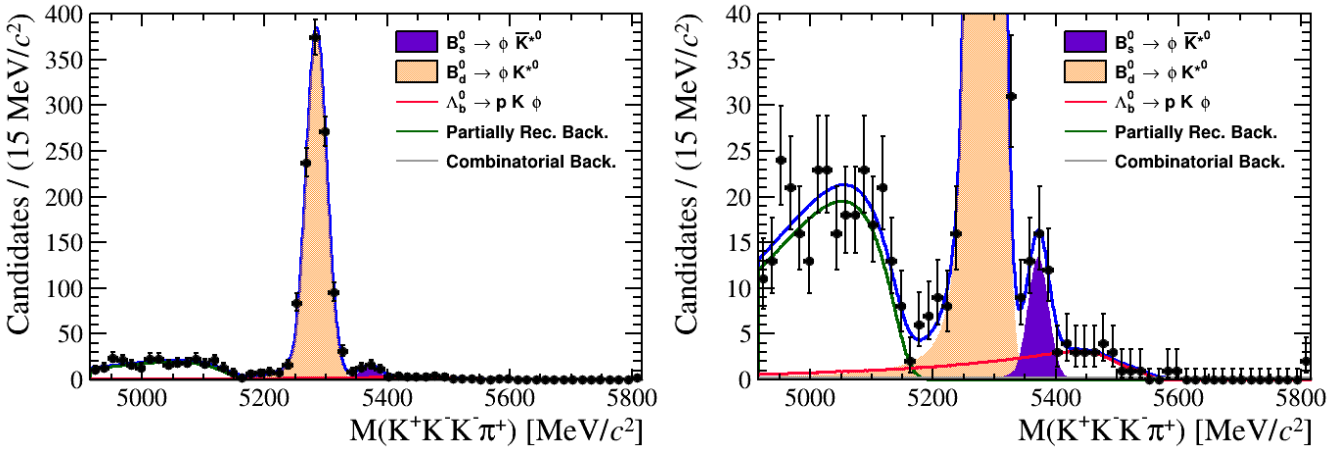
$$n_s = n_d \cdot 30/1000 \tag{18}$$

Where n_d is the number of events in the B_d^0 peak when no cut is applied to the BDT variable. In order to find out this quantity, a *blind fit*[25] is performed to the data with no cut on the BDT variable. This way, n_s can be estimated through (18), obtaining the optimal BDT cut value. With these results, the final fit to the data may be performed. Figures 11 and 12 show the plots with the fitting functions and the data from 2011 and 2012. Additionally, all results are summarised in Table 2.

[23] In this article [1], the analysis was performed with the data collected in 2011, the same data is used in this investigation.

[24] N_s and N_d are the measured number of events in the B_s^0 and B_d^0 peaks, respectively.

[25] The term *blind fit* comes from the fact that the fit is performed in such a way that it is impossible to see the results, graphic or numeric, within the B_s^0 mass window. The objective is to avoid any influences in the progress of the analysis due to the observation of the results, even if preliminary.



(a) (b)

Figure 11: Four-body ($K^+K^-K^-\pi^+$) invariant mass spectrum for the 2011 sample. The dots represent the data and the blue line the global fit. The shadowed purple area indicates the signal $B_s^0 \to \phi\overline{K}^{*0}$, while the shadowed orange region corresponds to the $B_d^0 \to \phi K^{*0}$ decay; the grey line symbolises the combinatorial background, the green line is the partially reconstructed background and the pink line the misidentified Λ_b^0 events. The full graph is shown on the left, while the right plot shows a zoom to the region of interest.

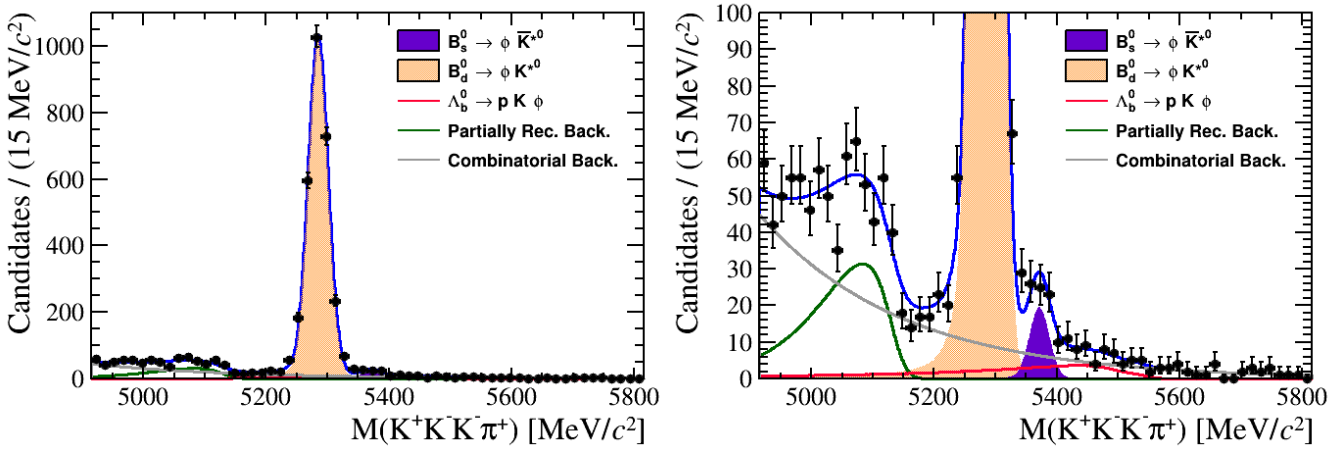
Table 2: Fit results.

Parameter	2011 Results	2012 Results
$N_{B_s^0 \to \phi\overline{K}^{*0}}$	39.5 ± 8.2	55 ± 11
$N_{B_d^0 \to \phi K^{*0}}$	1126 ± 35	2849 ± 58
Significance	5.70σ	5.52σ

When comparing the results obtained for the 2011 data with those published in 17 it is clear that they are consistent, taking into account the uncertainties in the measurements. As regards the 2012 data, it was expected that the number of events were twice that of 2011. However, the value observed is slightly smaller than expected, which may be due to statistical fluctuations or event misidentifications.

5.2 Branching fraction calculation

From the results in Table 2 and using equation (7) the branching fraction $\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})$ can be computed. Furthermore, the ratio between the branching fractions of the studied decays $\left(\frac{\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})}{\mathcal{B}(B_d^0 \to \phi K^{*0})}\right)$ can be calculated. Ergo, a comparison between the experimental result and the standard model prediction (see eq.(10)) may be performed. Table 3 gathers all the results obtained for the 2011 and 2012 data samples as well as the weighted mean of both.



(a) (b)

Figure 12: Four-body ($K^+K^-K^-\pi^+$) invariant mass spectrum for the 2012 sample. The dots represent the data and the blue line the global fit. The shadowed purple area indicates the signal $B_s^0 \to \phi\overline{K}^{*0}$, while the shadowed orange region corresponds to the $B_d^0 \to \phi K^{*0}$ decay; the grey line symbolises the combinatorial background, the green line is the partially reconstructed background and the pink line the misidentified Λ_b^0 events. The full graph is shown on the left, while the right plot shows a zoom to the region of interest.

Table 3: Branching fractions

Parameter	2011 Results	2012 Results	Weighted mean
$\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})$	$(1.34 \pm 0.31) \cdot 10^{-6}$	$(0.74 \pm 0.17) \cdot 10^{-6}$	$(0.88 \pm 0.15) \cdot 10^{-6}$
$\frac{\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})}{\mathcal{B}(B_d^0 \to \phi K^{*0})}$	0.137 ± 0.031	0.075 ± 0.016	0.088 ± 0.014

Comparing the results in Table 3 with the theoretical and experimental values presented in section 1, it is clear that they are compatible. Regarding the ratio between branching fractions, it can be observed that there is a factor of 2 between the measurements and the approximated value (see eq. (10)) calculated in section 3.3.2. This comes as no surprise since the value obtained in (10) is a mere approximation that can only give an idea the order of magnitude of the value.

The compatibility of the results with the values predicted by the SM indicates that the SM can correctly describe the physical processes occurring in these decays. This would imply, in principle, that the observables studied in this paper are not affected (at least significantly) by other possible particles or interactions beyond the standard model. For this to be confirmed, more accurate theoretical and experimental values are needed.

6 Conclusions

As a result of the study of the 2011 and 2012 data samples, the amount of candidates of $B_s^0 \to \phi\overline{K}^{*0}$ seen is 39.5 ± 8.2 in 2011 and 55 ± 11 in 2012, with a statistical significance of 5.70σ and 5.52σ, respectively. The number of events in 2012 was expected to be twice as large as that for 2011, since the LHCb luminosity increased by factor 2 from one year to the other. However, this was not observed, which can be ascribed to statistical fluctuations or to the misidentification of some candidate events.

The branching fraction attained by calculating the weighted mean of the branching fractions measured with the data samples from each year is:

$$\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0}) = (0.88 \pm 0.15) \cdot 10^{-6};$$

this value was obtained using the $B_d^0 \to \phi K^{*0}$ decay as a normalisation channel. This is the best measurement of this magnitude achieved so far, taking into account solely the statistical uncertainties[26]. On the one hand, the result is approximately twice that obtained based on QCD factorisation, which is $(0.4^{+0.5}_{-0.3}) \cdot 10^{-6}$ [2]. On the other hand, it is also slightly larger than the value obtained through perturbative QCD, $(0.65^{+0.33}_{-0.23}) \cdot 10^{-6}$ [3]. Notwithstanding, all results are compatible within 1σ. Similarly, the result is consistent with the experimental result obtained in [1]: $\left(1.10 \pm 0.24(stat) \pm 0.14(syst) \pm 0.08 \left(\frac{f_d}{f_s}\right)\right) \cdot 10^{-6}$.

Finally, the value for the ratio between branching fractions of the $B_s^0 \to \phi\overline{K}^{*0}$ and $B_d^0 \to \phi K^{*0}$ decays is:

$$\frac{\mathcal{B}(B_s^0 \to \phi\overline{K}^{*0})}{\mathcal{B}(B_d^0 \to \phi K^{*0})} = 0.088 \pm 0.014.$$

This result is twice as large as the value predicted by the SM $(0.043^{+0.075}_{-0.046})$ [2], although they are compatible within uncertainties. The result is also consistent with the one obtained in [1]: $0.113 \pm 0.024(stat) \pm 0.013(syst) \pm 0.009 \left(\frac{f_d}{f_s}\right)$.

All the results obtained along this investigation fit within the predictions of the standard model. An improvement in the precision of the theoretical and experimental estimations is necessary in order to confirm or discard the existence of particles or interactions beyond the SM that might affect this process.

[26] The analysis of systematic errors is beyond the objectives of this research project.

References

[1] LHCb Collaboration, *First observation of the decay* $B_s^0 \to \phi\overline{K}^{*0}$, JHEP 11 (2013) 092, arXiv:1306.2239v3.

[2] J. Beneke, J. Rohrer, and D. Yang, *Branching fractions, polarisation and asymmetries of* $B \to VV$ *decays*, Nucl. Phys. **B774** (2007) 64, arXiv:hep-ph/0612290.

[3] A. Ali et al., *Charmless non-leptonic B_s decays to PP, PV and VV final states in the pQCD approach*, Phys. Rev. **D76** (2007) 074018, arXiv:hep-ph/0703162.

[4] J. Beringer et al.(PDG), PR **D86**, 010001 (2012) and 2013 update for the 2014 edition (http://pdg.lbl.gov)

[5] LHCb Collaboration, *Updated average* f_s/f_d *b-hadron production fraction ratio for 7 TeV pp collisions*, CERN-LHCb-CONF-2013-011 (2013).

[6] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, PoS ACAT:040 (2007), arXiv:physics/0703039.

[7] http://home.web.cern.ch/

[8] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, P. Proudlock, *LHC Design Report - v.1: the LHC Main Ring*, CERN-2004-003-V-1 (2004).

[9] ATLAS Collaboration, *ATLAS: technical proposal for a general-purpose pp experiment at the Large Hadron Collider at CERN*, CERN-LHCC/94-43 (1994).

[10] CMS Collaboration, *Technical Proposal*, CERN-LHCC/94-38 (1994).

[11] ALICE Collaboration, *ALICE: Technical proposal for a Large Ion collider Experiment at the CERN LHC*, CERN-LHCC/95-71 (1995).

[12] LHCb Collaboration, *LHCb : Technical Proposal*, CERN-LHCC/98-004 (1998).

[13] http://lhcb-public.web.cern.ch/lhcb-public/

[14] V. Niess (LHCb Collaboration), *Commissioning of the LHCb preshower detector with cosmic rays and first LHC collisions*, J. Phys.: Conf. Ser. **293** 012060 (2011).

[15] I. Machikhiliyan (LHCb calorimeter group), *The LHCb electromagnetic calorimeter*, J. Phys.: Conf. Ser. **160** 012047 (2009).

[16] A. A. Alves Jr et al., *Performance of the LHCb muon system*, JINST **8** P02022 (2013).

[17] R. Aaij et al., *The LHCb trigger and its performance in 2011*, JINST **8** P04022 (2013).

[18] LHCb Collaboration, *LHCb trigger system : Technical Design Report*, CERN-LHCC-2003-031 (2003).

[19] LHCb Collaboration, *LHCb VELO (VErtex LOcator) : Technical Design Report*, CERN-LHCC-2001-0011 (2001).

[20] S. Perazzini, *Measurement of branching fractions and $\mathcal{CP}$ violation for charmless charged two-body B decays at LHCb*, Tese de doutoramento, Università di Bologna, Italia (2012).

[21] A. S. Dighe et al., *Angular distributions and lifetime differences in $B_s \to J/\psi\phi$ decays*, Phys.Lett. B369 (1996) 144-150, arXiv:hep-ph/9511363.

[22] P. Gandini, *Observation of CP Violation in $B^\pm \to DK^\pm$ Decays*, Springer (2013).