Search for new physics in charmless three-body B-meson decays

Émilie Bertholet

SORBONNE UNIVERSITÉ
École doctorale des Sciences de la Terre et de l'environnement et Physique de l'Univers, Paris (ED 560)

Laboratoire de Physique Nucléaire et de Hautes Énergies (LPNHE, UMR 7585)

Search for new physics in charmless three-body B-meson decays

by

Emilie Bertholet

Submitted in fulfillment of the requirements for the degree of

Doctor of Philosophy

Supervised by Prof. Eli Ben-Haïm and Dr. Matthew Charles

Defended on the 25th of September 2019 in front of the committee:

Prof.	Christophe	Balland	
Prof.	Eli	Ben-Haïm	Supervisor
Dr.	Matthew	Charles	Supervisor
Dr.	Sébastien	Descotes-Genon	Referee
Prof.	Tim	Gershon	
Dr.	Anton	Poluektov	
Dr.	Karim	Trabelsi	Referee
Dr.	Keri	Vos	
Dr.	Marco	Zito	

הכל צפוי והרשות נתונה
פרקי אבות ג טו

à mes parents
à mon frère
à mon oncle Gilou

Contents

Chapter I

Introduction

The Standard Model (SM) of particle physics, developed in the early 1970s, is the theory that currently best describes the fundamental interactions between elementary particles. It unifies three out of the four known fundamental interactions: the strong, weak and electromagnetic interactions. The strong force is responsible for the formation of bound states of quarks called hadrons, such as protons and neutrons, and allows for the existence of nuclei. The weak interaction is capable of changing fermion flavour, and thus enables radioactive decays in the nuclei. Finally, the electromagnetic interaction holds the electrons and nuclei together, forming atoms. The incredible success of the SM comes from the fact that many of its predictions have been confirmed experimentally, leading to numerous discoveries, such as, for instance, the top quark [1] in 1995, the tau neutrino [2] in 2000 and the Higgs boson [3, 4] in 2012. However, the SM is not the ultimate theory. For example, the gravitational force has not been included yet due to the complexities arising from the quantisation of fields of spin two. Moreover, cosmological models and astrophysical observations show that baryonic matter is only a small portion of the matter present in the universe and that we actually understand only a tiny part of the total energy content of the universe.

Assuming that the Universe started from an initial state with equal amounts of matter and antimatter, there must be a mechanism that generates an asymmetry so that during the evolution of the universe matter prevailed over antimatter, leading to today's observations of baryon asymmetry. The physics processes underlying baryogenesis must meet three criteria, called the Sakharov conditions [5], to explain the matter/antimatter asymmetry: baryon number violation, *CP* violation and a departure from thermal equilibrium. Different mechanisms in the SM can generate *CP* violation, all implying particles with

non-zero masses. The first source of *CP* violation in the SM comes from the weak sector. The quarks, providing that they have non-zero and non-degenerate masses, can change flavour under certain conditions when undergoing a decay through the weak interaction. This difference in the quark content between the initial and final states can lead to *CP* violation and thus, to different behaviours between matter and antimatter. A similar mechanism can occur in the neutrino sector, due to the small, but non-zero, mass of the neutrinos. A priori, in quantum chromodynamics, the theory of strong interaction, nothing prevents strong *CP* violation from occurring. However, no experimental observations of such violation have been found in strong decays: this is known as the strong *CP* problem [6]. A further problem is that, although *CP* violation occurs in the SM, the amount of *CP* violation is not large enough to explain the matter-antimatter imbalance in the universe.

An important part of today's research in high energy physics is to find ways to discover what exists beyond the SM. From the experimental point of view, it means either trying to detect new particles, aka "direct searches", or looking for discrepancies with the SM predictions by performing high-precision measurements, aka "indirect searches". The direct detection of new particles in collider experiments depends on the energy of the collision in the centre-of-mass. For pair production, the energy in the centre-of-mass must be at least twice the mass of the new particle. Currently, the LHC is providing proton-proton collisions at an energy of 13 TeV in the centre-of-mass. To probe for a higher energy scale without having to increase the energy of the collision, the indirect approach is adopted. The decay of a particle can proceed via intermediate, "virtual", particles. Theses virtual states, also called "off-shell" particles, cannot be detected directly, but the presence of a new particle may affect the physical observables. In that case, discrepancies with the theoretical prediction or with other measurements of the same quantity may be observed. The direct and indirect approaches are complementary; progress in particle physics over the years has been done thanks to both.

On the theoretical side, different approaches are also possible. One can, in a "top to bottom" approach, try finding the ultimate theory that describes the laws of physics at all energy scales and verify that the established results still hold. Conversely, in a "bottom to top" approach, the idea is to start from current experimental and theoretical knowledge and propose models that explain the tensions with the SM while being still consistent with past observations, and to design methods to extract experimental observables that are sensitive to new physics. This approach requires theorists and experimentalists to work closely together to find methods that minimise the theoretical and experimental uncertainty. The analysis described in chapter IV of this thesis is part of this

context.

The work presented in this thesis is centred on precision measurements in the flavour sector using three-body B meson decays to charmless final states. Chapters II and III give the theoretical and experimental grounds for understanding the analyses presented in the subsequent chapters. Chapter II introduces the main concepts of the SM, with an particular emphasis on the flavour sector and on CP violation. A deeper insight into the specifics of charmless B meson decays is given in chapter III. This chapter also discusses the framework of three-body decays and presents different methods of extraction of the weak phase γ, an observable related to CP violation. A method to extract γ with charmless three-body B meson decays, sensitive to new physics, is studied in chapter IV. The second analysis presented in this thesis concerns the study of the $B^0_{(s)} \to K^0_S h^\pm h'^\mp$ modes, which are decays of the B^0 or B^0_s mesons into charmless final states containing a K^0_S and two hadrons ($h = \pi, K$), and are among the inputs relevant to the γ extraction discussed in chapter IV. The decay mode $B^0_s \to K^0_S K^+ K^-$ is observed for the first time and its branching fraction is measured using data collected by the LHCb detector in chapter VI. The LHCb detector, described in chapter V, has been specifically designed for the study of b- and c-hadrons, and is thus well suited for this analysis.

Chapter II

General theoretical framework: Standard Model and *CP* violation

II.1 Symmetries

Symmetries have always played an important role in physics. A system is said to have a symmetry when the measurable quantities are left invariant under a specific transformation. At the beginning of the XX$^{\text{th}}$ century, the German mathematician Emmy Noether proved that each continuous, differentiable symmetry is associated with a conservation law. For instance, if the Lagrangian of a system is invariant under translations in space, the related conserved quantity is the momentum. Similarly, invariance under rotations in space is associated with angular momentum conservation and invariance under translations in time is related to energy conservation. These ideas are at the centre of modern-physics thinking and theory building, and in particular in quantum field theory, where the theories that describe particles and their interactions can be constructed from symmetry arguments.

This part describes the main concepts of symmetry in field theories. First, the notion of gauge symmetry is presented in a general manner and illustrated with examples taken from the Standard Model of particle physics (SM). Then, discrete symmetries, such as C, P and T are addressed. The aim is not to explain the construction of the SM from scratch or to give a mathematical description of group theory but rather to show the link between the two through the concept of symmetry. It is assumed that the reader has a basic knowledge in group theory, quantum field theory and the SM.

II.1.1 Gauge symmetries

A theory describing fundamental particles must agree with special relativity. In other words, the Lagrangian of the theory must be invariant under space-time transformations belonging to the Poincaré group[1], P(1,3). In field theory, for one time dimension and three space dimensions, this requirement leads to scalar, vector, and tensor fields for bosons and spinor fields for fermions. Symmetries that couple to the Poincaré group are called "external symmetries".

A Lagrangian can also have internal symmetries that are independent of the Poincaré group, which means that the generators of these symmetries commute with the generators of P(1,3). These symmetries are related to transformations of the fields; for example a global phase shift of a field ϕ such as $\phi \to e^{i\alpha}\phi$ belongs to the symmetry group U(1). Mathematically, continuous symmetries are described by Lie groups, which contain an infinite number of elements that can be arbitrarily close to the identity. For instance, U(N) is the group of N×N unitary matrices ($UU^\dagger = \mathbb{1}$) that satisfy $|\det(\mathrm{U})| = 1$. A representation of the group is $U = e^{i\theta T}$, where T is a generator of the group. The number of generators needed to describe a group corresponds to the dimension of the group.

Symmetries provide a fundamental principle that determines the form of the Lagrangian of a theory. They can be local (space-time dependent) or global (space-time independent). Gauge symmetries are internal symmetries, so they do not "see" P(1,3). In other words, at each point of the spacetime there is an internal space that is independent of P(1,3). If the field ϕ is invariant under this symmetry, then $\phi(x)$ is a vector of the internal space defined at point x. This is the reason why the gauge symmetry needs to be local, so that the symmetry is fulfilled in the internal space defined at each spacetime point.

In the simple case of an Abelian[2] gauge group, U(1), a local transformation of a field ϕ is of the form

$$\begin{aligned} \phi(x) &\to e^{i\alpha(x)}\phi(x), \\ \phi^*(x) &\to e^{-i\alpha(x)}\phi^*(x), \end{aligned} \tag{II.1}$$

which corresponds to a phase rotation. It is straightforward to see that the derivative of the field, $\partial_\mu\phi$, is not invariant under local gauge transformations. Therefore, ∂_μ is not the right quantity to use; instead, one must adopt the covariant derivative, D_μ, which has the property to transform the same way as the field:

$$D_\mu\phi(x) \to e^{i\alpha(x)} D_\mu\phi(x). \tag{II.2}$$

[1] The Poincaré group is the Lorentz group with translations added.

[2] Also called commutative group: the generators of the group commute with each other.

The covariant derivative can be seen as a generalisation of the derivative to non-Euclidian spaces. In the specific case of an Abelian transformation, it can be shown that

$$D_\mu \phi(x) = (\partial_\mu + ie\mathcal{A}_\mu(x))\phi(x), \tag{II.3}$$

where e is a constant and $\mathcal{A}_\mu$ is a vector field, satisfies Eq. (II.2). The field $\mathcal{A}_\mu$, called a connection, ensures that the covariant derivative transforms like the field ϕ. It connects the internal group at point x with its "copy" at point $x+dx$. This term is a direct consequence of the invariance under local phase rotations that was imposed on the field. From there we can deduce a systematic method to construct locally invariant Lagrangians: any combination of ϕ and $\partial_\mu\phi$ that is invariant under a global phase rotation can be made locally invariant by simply replacing the derivative by the covariant derivative.

We now need to find a kinematic term for $\mathcal{A}_\mu$. That is, a combination of $\mathcal{A}_\mu$ and its derivatives (exclusively) which is invariant under (II.1). This can be achieved by using, again, the properties of the covariant derivative. Inspecting (II.2), it is clear that the product of two covariant derivatives is invariant under the local transformation and so is their commutator,

$$[D_\mu, D_\nu]\phi(x) \to e^{i\alpha(x)}[D_\mu, D_\nu]\phi(x). \tag{II.4}$$

Using (II.3), the expression of the commutator as a function of $\mathcal{A}_\mu$ is given by

$$[D_\mu, D_\nu] = ie(\partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu) = ieF_{\mu\nu}. \tag{II.5}$$

The quantity $F_{\mu\nu}$ and its derivatives are then locally invariant. We now have the all the ingredients for the Lagrangian: ϕ, $D_\mu\phi$, $F^{\mu\nu}$, and $\partial_\mu F^{\mu\nu}$.

Up to now, nothing was specified about the generic field ϕ. Let us take the case of a Dirac fermion, $\psi(x)$, for example, an electron. Using the "building blocks" described previously it is possible to construct the most general Lagrangian including this field. This Lagrangian can, in principle, contain operators of any dimension. However, in the case of the electron in four space-time dimensions, if only operators up to dimension four are included it is possible to show that no divergences due to higher order corrections are present in the Lagrangian. In this case the theory is said to be "renormalisable"[3]. If we further require parity and time invariance symmetries[4] we obtain the QED Lagrangian

$$\mathcal{L}_{\text{QED}} = \overline{\psi}(i\gamma^\mu D_\mu - M)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{II.6}$$

[3]Renormalisation theory will not be discussed in this document. The interested reader can refer to [7].

[4]Discrete symmetries are discussed in the next section.

where γ^μ are the Dirac matrices and $\overline{\psi}$ is the Dirac adjoint, defined as $\overline{\psi} \equiv \psi^\dagger \gamma^0$. The connection $\mathcal{A}_\nu$ can be identified as the vector potential (i.e. the photon) and $F_{\mu\nu}$ is the familiar electromagnetic field tensor.

This example is an illustration of how fundamental is the notion of symmetry in field theory and, especially, gauge symmetry. By requiring a field to be locally invariant under a certain set of transformations a new vector field, $\mathcal{A}_\mu$, which is required by the covariant derivative, is introduced. The kinetic energy term related to this new field is then found by inspecting the commutator of the two covariant derivatives, which gives the strength field tensor. This method ensures that the quantities that are built obey to the symmetry laws that were imposed on the field.

This approach can be generalised to non-Abelian gauge symmetries. In this more general case, the generators, T^a of the group do not commute, and thus we have the relation

$$[T^a, T^b] = i f^{abc} T^c, \tag{II.7}$$

where f^{abc} are called structure constants. As an example, we can take the commutator of two vectors in a three-dimensional Euclidian space, which is nothing else than the cross product of these vectors. The corresponding structure constants are then the three-dimensional Levi-Civita symbols[5].

Since gauge symmetries groups are Lie groups, the commutation laws satisfy a Lie algebra. Hence, the commutators must obey the Jacobi identity

$$[T^a, [T^b, T^c]] + [T^b, [T^c, T^a]] + [T^c, [T^a, T^b]] = 0, \tag{II.8}$$

which leads to the following relation between the structure constants

$$f^{ade} f^{bcd} + f^{bde} f^{cad} + f^{cde} f^{abd} = 0. \tag{II.9}$$

A local, non-Abelian, gauge transformation of a field ϕ takes the from

$$\phi(x) \to e^{i\alpha^a(x) L_a} \phi(x), \tag{II.10}$$

where L_a are matrix representation of the generators. Equipped with the commutation relations and the transformation, the next step is to compute the covariant derivative, which is given by

$$D_\mu = \partial_\mu - i g \mathcal{A}^a_\mu L_a, \tag{II.11}$$

[5] ϵ^{ijk} is equal to 1 for even permutations of (i, j, k), -1 for odd permutations of (i, j, k), and 0 if any index is repeated.

where g is a constant and $\mathcal{A}^a_\mu$ are the connections. It is interesting to notice that the number of vector fields (connections) that are needed for the covariant derivative is the same as the number of independent generators of the gauge group. For instance, in the familiar case of SU(2), there are three independent generators, i.e. three Pauli matrices, therefore three vector fields $\mathcal{A}^a_\mu$ appear in the covariant derivative.

Finally, the field strength tensors $F^a_{\mu\nu}$ are obtained by studying the commutation relations of the covariant derivative,

$$[D_\mu, D_\nu] = -igF^a_{\mu\nu}. \tag{II.12}$$

After calculations, one obtains

$$F^a_{\mu\nu} = \partial_\mu \mathcal{A}^a_\nu - \partial_\nu \mathcal{A}^a_\mu + g f_{abc} \mathcal{A}^b_\mu \mathcal{A}^c_\nu, \tag{II.13}$$

where the first two terms are identical to the Abelian case and the last term reflects the non-commutative nature of the gauge group.

A more concrete example of a non-Abelian gauge theory is the theory that describes the strong interaction, namely quantum chromodynamics (QCD). To explain the existence of particles such as the Δ^{++}, which is composed of three up-quarks with parallel spins, without violating Pauli exclusion principle, an extra degree of freedom was postulated by Oscar W. Greenberg in 1964 [8]: the colour charge. In addition of coming in different flavours, quarks come also with different colours. Three colours (and three anti-colours) are necessary to explain the light states in the SM. The particles observed in nature are colourless: for example the three quarks of the Δ^{++} baryon must each have a different colour, which obviously solves the problem of having three identical quarks in the same state.

The gauge group of colour symmetry is SU(3) (which will be referred as colour SU(3) or $\mathrm{SU}(3)_C$ in order not to confuse it with flavour SU(3) that will be discussed later). A field in the colour space has then three elements, corresponding to the three colours,

$$\psi = \begin{pmatrix} \psi_r \\ \psi_g \\ \psi_b \end{pmatrix}. \tag{II.14}$$

The general expressions for the local gauge transformation (II.10), the covariant derivative (II.11), and the field strength tensors (II.13) can be directly used here. There are eight independent generators of SU(3), which can be represented by the "Gell-Mann matrices" (see, for example, Ref. [7]). This means that eight vectors fields $\mathcal{A}^a_\mu$ are needed in the covariant derivative. These fields correspond to the gluons that are the mediators of the strong force and the carriers of the colour charge.

II.1.2 Discrete symmetries

In addition to continuous symmetries, the Lagrangian can also be invariant under discrete symmetries. In this part we will focus on parity, P, charge, C and time reversal T symmetries. In order to avoid unnecessary complications, we will use the Hamiltonian formalism instead of the Lagrangian one. The two formalisms are equivalent and one can be deduced from the other by applying some transformations. The only reason for using the Hamiltonian formalism is that it gives a more direct grasp on the concepts that are discussed here.

Unlike the gauge transformations, which are described by groups with an infinite number of elements, the symmetry groups of C, P and T contain a finite number of elements: the identity I, and an element g that satisfies $g^2 = I$. Both P and T are symmetries of the spacetime while C is not. The statement that the physical observables remain unchanged under g implies that g can be represented by a unitary (or antiunitary) operator $U(g)$ such that $[U, H] = 0$, where H is the (Hermitian) Hamiltonian of the system. As a result, the eigenvalues of these operators are $\pm$ 1.

Parity

A parity transformation reverses the direction in space. The parity unitary operator, P, acts on a four-vector of the the Poincaré group, $(t, \vec{x})$, by changing the sign of the space coordinates: $(t, \vec{x}) \to (t, -\vec{x})$. Applied to a particle, this symmetry reverses the sign of the momentum while leaving the spin unchanged. Eigenstates of P, such as hadrons, can be classified according to the how they transform under parity,

$$\begin{array}{llll} \text{scalar:} & S \to S & \text{vector:} & V \to V \\ \text{pseudoscalar:} & P \to -P & \text{pseudovector:} & A \to -A \end{array} \tag{II.15}$$

The parity of a fermion is the opposite of the parity of its antiparticle while the parity of a boson is the same as the parity of its antiparticle. The parity of a quark is taken to be positive (this is an arbitrary choice) and so the parity of an antiquark is negative. Parity is a multiplicative quantum number, which means that the parity of a composite state is the product of the parities of the individual states. For systems that are not in their ground state the orbital momentum between the different particles has to be taken into account. In the case of a meson, which is a bound state of a quark and an antiquark, the parity is then $(+1)(-1)(-1)^l = (-1)^{l+1}$, where l is the orbital momentum between the two quarks.

Historically, parity symmetry was taken for granted and was assumed to be conserved in all interactions. However, an unusual behaviour was observed in

the weak decay of two particles, named θ and τ at the time. These two mesons had the same mass and the same quantum numbers (same spin, same charge etc.). Nevertheless, the preferred decay channel for the θ was into two pions ($P = +1$) whereas the τ decayed into three pions ($P = -1$). Assuming parity conservation, the two particles have opposite parities, and thus are different particles. However, it is quite surprising to find two particles that are identical in every aspect except parity. This was known as the $\theta - \tau$ puzzle. A solution to this puzzle was proposed by Lee and Yang in 1956 [9], by postulating that these particles were actually the same particle, which is now called the K^+, and that parity is actually not conserved in one of the decays. This triggered a whole era of theoretical and experimental searches for parity violation in weak decays.

Charge conjugation

The charge conjugation unitary operator, C, changes the sign of all charges — all internal quantum numbers — into their opposite. In other words, it transforms a particle into its antiparticle. Several particles can form together a C eigenstate, as for instance, a particle and its antiparticle. Like the parity group, the charge-conjugation group has two elements, I and C, and $C^2 = I$. Thus the eigenvalues of C are ± 1, and the transformation of a charge eigenstate under C reads

$$C|p\rangle = \pm|p\rangle = |\overline{p}\rangle \tag{II.16}$$

where $|p\rangle$ is an eigenvector of C. In terms of physics, two states that only differ by a global phase shift represent the same state. Applied to particles, it means the eigenstates of C are particles that are their own antiparticles, which leads to the fact that most particles are not C eigenstates.

Time reversal

Time reversal is another symmetry that applies to spacetime, reversing the sign of the time component: $(t, \vec{x}) \to (-t, \vec{x})$. Unlike in the case of P and C, the time reversal operator T is antiunitary. This ensures that this symmetry changes the sign of t while keeping the energy positive. Applied to a particle, this operator reverses its momentum and its spin projection.

CPT invariance

Experimentally, it was shown that C and P are violated in the weak interaction while strong, electromagnetic and gravitational interactions seem to conserve

C, P and T symmetries. The violation of the combined action of P and C, known as CP violation[6], is well established in K and B mesons, and was very recently observed in D mesons [10]. However, the combination of the three symmetries together is believed to be an exact symmetry. This is a consequence of the CPT theorem [11, 12] that states that it is not possible to construct a Lorentz-invariant quantum field theory with a Hermitian Hamiltonian that violates CPT. A violation of CPT symmetry would probably implies Lorentz-symmetry breaking. Another important consequence of this theorem is that CP violation must be compensated by a violation of T parity. Time-reversal violation was observed experimentally by the BaBar collaboration in 2012, using entangled states of B mesons [13].

II.1.3 Realisation of symmetries and spontaneous symmetry breaking

Formally, a symmetry of the Lagrangian can be realised in two modes: the Wigner-Weyl and the Nambu-Goldstone modes, depending on the transformation of the vacuum state.

Wigner-Weyl mode

The Wigner-Weyl mode corresponds to the case where both the Lagrangian and the vacuum state are invariant under a set of transformations. The relation $T^a|0\rangle = 0$, where T^a are the generators of the symmetry and $|0\rangle$ denotes the vacuum state, implies that the spectrum is degenerate. For instance, rotations in a three dimensional space are generated by three generators, L_i (where $i = x, y, z$), which commute with the Hamiltonian of the system. It can be shown that the spectrum is composed of multiplets of angular momentum l with $2l + 1$ degenerate values. These values are labelled by the projection of $\vec{L}$ along an arbitrary axis; the usual convention is $l_z = m$ and $m = -l, -l + 1, ..., l - 1, l$.

A more instructive example is given by chiral symmetry. Consider the free Dirac Lagrangian for quarks

$$\mathcal{L} = \overline{\psi}(i\not{\partial} - M)\psi, \tag{II.17}$$

where $\not{\partial} = \gamma^\mu \partial_\mu$. In this example, we neglect the colour symmetry so that ψ is

[6]The mechanism of CP violation will be discussed in details in the next sections.

a vector in the flavour space

$$\psi = \begin{pmatrix} \psi_u \\ \psi_d \\ \vdots \end{pmatrix}, \tag{II.18}$$

where the components ψ_f of ψ are Dirac spinors that represent the various quark flavours. The mass matrix M is diagonal in the flavour space. The Lagrangian of Eq. (II.17) is invariant under vector transformations of the form

$$\psi \to e^{i\vec{\alpha}\cdot\vec{T}}\psi. \tag{II.19}$$

If only two flavours are included, the corresponding symmetry group is then $\mathrm{SU}(2)_V$ (where the subscript V stands for vector), and the generators of the group, $\vec{T}$, are proportional to the Pauli matrices[7] $\vec{\sigma}$, $\vec{T} = \vec{\sigma}/2$. From Noether theorem, the conserved currents are given by

$$\vec{J}^{\mu}_V = \overline{\psi}\gamma^{\mu}\vec{T}\psi \tag{II.20}$$

Since these currents are conserved, their derivative must be zero. It can be shown, using Eq. (II.17), that

$$\partial_{\mu}\vec{J}^{\mu}_V = i\overline{\psi}[M, \vec{T}]\psi. \tag{II.21}$$

The requirement $\partial_{\mu}\vec{J}^{\mu}_V$ implies that M and $\vec{T}$ commute. By working out the commutator using the explicit forms of the matrices it appears that

$$[M, \vec{T}] = 0 \Rightarrow m_u = m_d. \tag{II.22}$$

It appears from this equations that the quark-mass states are degenerate, this symmetry is then realised in the Wigner-Weyl mode. It turns out that this $\mathrm{SU}(2)_V$ symmetry, also known as isospin symmetry, is a pretty good symmetry of QCD. Indeed, it is found that the masses of the proton and the neutron, which are an isospin doublet, are very close together: $m_p = 938.23\,\mathrm{MeV}$ and $m_n = 939.57\,\mathrm{MeV}$. The same goes for pions, which form a triplet under isospin: $m_{\pi^{\pm}} = 139.57\,\mathrm{MeV}$ and $m_{\pi^0} = 134.98\,\mathrm{MeV}$.

[7] The Pauli matrices are three 2×2 unitary Hermitian matrices given by $\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $\sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

The Dirac Lagrangian of Eq.(II.17) is also invariant under axial vector (pseudovector) transformations

$$\psi \to e^{i\vec{\alpha}\cdot\vec{T}\gamma_5}\psi, \tag{II.23}$$

where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. If we consider two flavours of quarks, the symmetry group is then $\mathrm{SU}(2)_A$ (where the subscript A stands for axial vector). The generators of the group depend on the Pauli matrices and on γ^5,

$$\vec{T}\gamma^5 = \frac{\vec{\sigma}\gamma^5}{2}. \tag{II.24}$$

Following the same procedure as in the isospin case, we find that the Noether currents are given by

$$\vec{J}^{\mu}_A = \overline{\psi}\gamma^{\mu}\vec{T}\gamma^5\psi. \tag{II.25}$$

This time, this is the anti-commutator of M and $\vec{T}$ that appears in the derivative of the current

$$\partial_{\mu}\vec{J}^{\mu}_A = i\overline{\psi}\{M, \vec{T}\}\gamma^5\psi. \tag{II.26}$$

The condition $\partial_{\mu}\vec{J}^{\mu}_A = 0$ is fulfilled only if the quarks have zero mass. The combined vector and axial symmetry is called chiral symmetry. This symmetry is a approximately true at energies large enough so that the quark masses can be neglected.

Nambu-Goldstone mode

When the Lagrangian is invariant but not the vacuum state, $T^a|0\rangle \neq 0$, the symmetry is said to be realised in the Nambu-Goldstone mode. The generators that satisfy $T^a|0\rangle \neq 0$ are called "broken" generators. Not all the generators have to break the invariance of the vacuum for the symmetry to be realised in this mode. The consequences of the presence of such broken generators is that the vacuum expectation value of the field is different from zero and that it exists operators that can create particles from the vacuum. Those particles are called Goldstone bosons. Using the fact that a symmetry is associated with the conservation of Noether currents, it can be shown that the Goldstone bosons must be massless [7].

To illustrate this, consider the Dirac Lagrangian used in the previous example Eq. (II.17), in the case of massless fermions ($M = 0$). It can be expressed in the Weyl basis, by projecting the state ψ on the left- and right-handed states[8]. The Lagrangian then reads

$$\mathcal{L} = \overline{\psi}_L i\not{\partial}\psi_L + \overline{\psi}_R i\not{\partial}\psi_R. \tag{II.27}$$

[8]This is done by using the projectors $P_{R/L} = \frac{1\pm\gamma^5}{2}$, $\psi = P_R\psi + P_L\psi$.

Expressed under this form, the Lagrangian has two disjoint parts, L and R, which indicates a $\mathrm{SU}(2)_R \otimes \mathrm{SU}(2)_L$ symmetry in the case of two flavours. As seen in the previous paragraph, this Lagrangian is invariant under both vector and axial symmetries. After projection on the left- and right-handed states the transformation takes the form

$$\psi_{R/L} \to e^{i\vec{\alpha}_{R/L}\cdot\vec{T}_{R/L}}\psi_{R/L}, \tag{II.28}$$

where the generators are given by $\vec{T}_{R/L} = \frac{1}{2}(1 \pm \gamma^5)\vec{T}$. Parity transforms left and right states into each other. If this symmetry is realised in the Wigner-Weyl mode it entails a hadron spectrum composed of degenerated multiplets with opposite parity, which is not what it is observed experimentally. Therefore, the symmetry is spontaneously broken. According to observation, the conserved symmetry is isospin, so that

$$\mathrm{SU}(2)_L \otimes \mathrm{SU}(2)_R = \mathrm{SU}(2)_V \otimes \mathrm{SU}(2)_A \to \mathrm{SU}(2)_V. \tag{II.29}$$

The ground state is not invariant under $\mathrm{SU}(2)_A$ (which rotates differently ψ_R and ψ_L fields), but it is invariant under isospin rotations (which are the same for ψ_R and ψ_L). As explained before, we expect as many (massless) Goldstone bosons as broken generators, three in the case of two flavours. Since the broken symmetry is axial, these particles should be pseudoscalar particles. They can be identified with the pions. Obviously, pions are not massless particles, so this symmetry is not exact. Nevertheless, it explains the rather small masses of these particles.

II.2 Weak interaction

In the previous section, we introduced the notion of symmetry and explored some of its consequences. Considering gauge symmetries, it is possible to construct invariant Lagrangians in a general and systematic way. Taking into account renormalisation conditions and extra symmetries one can build a theory that describes physical processes. We saw that gauge symmetries engender vector fields, the gauge bosons, which are massless. To give mass to gauge bosons another mechanism is needed: the so-called Brout-Englert-Higgs (BEH) mechanism [14], which is based on spontaneous symmetry breaking. By adding a scalar field to the Lagrangian, with a potential chosen so that the field acquires a non-zero vacuum expectation value, the gauge invariance can be spontaneously broken. This leads to massless Goldstone bosons, which couple with the gauge bosons giving them non-zero masses. These Goldstone

particles do not appear as physical states; it is possible to get rid of them by making an appropriate choice of gauge. Nevertheless, they are important for the conservation of the number of degrees of freedom. Indeed, a massless vector boson has two polarisations while a massive one has three. The Goldstone bosons are then said to be "eaten" by the gauge bosons. The BEH mechanism will not be discussed in more detail in this work.

From there, it is rather straight forward to understand the construction of the SM. We already introduced, in part II.1.1, $\mathrm{SU}(3)_C$ and $\mathrm{U}(1)$: the gauge groups of the strong interaction (mediated by massless gluons) and the electromagnetic interaction (mediated by massless photons), respectively. The remaining interaction to describe is the weak interaction, which is mediated by three massive gauge bosons $W^{\pm}$ and Z^0. A starting point is $\mathrm{SU}(2)$ gauge symmetry, with the addition of a scalar field that breaks spontaneously the symmetry. This theory leads to three massive gauge bosons with the same mass, which is not what is observed; the measured masses of the $W^{\pm}$ and Z^0 are $m_{W^{\pm}} = 80.385 \pm 0.015\,\mathrm{GeV}$ and $m_{Z^0} = 91.1876 \pm 0.0021\,\mathrm{GeV}$ [15]. A way to overcome this issue is to add a $\mathrm{U}(1)$ gauge symmetry and treat weak and electromagnetic interactions together. This model is known as the Glashow-Weinberg-Salam (GWS) theory [16–18].

The complete gauge group of the SM is then $\mathrm{SU}(3)_C \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$, where the strong interaction described by $\mathrm{SU}(3)_C$ and the electroweak interaction by $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$. We will focus on the electroweak interaction in the following paragraph. The SM Lagrangian corresponding to this symmetry group can be split into three parts

$$\mathcal{L}_{\mathrm{SM}} = \mathcal{L}_{\mathrm{kinetic}} + \mathcal{L}_{\mathrm{Higgs}} + \mathcal{L}_{\mathrm{Yukawa}}, \tag{II.30}$$

where $\mathcal{L}_{\mathrm{kinetic}}$ contains the kinetic terms of the form $\frac{1}{4}|F_{\mu\nu}|^2$ for gauge bosons and $\overline{\psi} i \not{D} \psi$ for fermions.The Higgs potential is contained in $\mathcal{L}_{\mathrm{Higgs}}$ and $\mathcal{L}_{\mathrm{Yukawa}}$ describes the interactions between the Higgs field and the fermions. This last term will be presented when discussing the mechanism of *CP* violation in part II.3.

Glashow-Weinberg-Salam theory

As argued before, a possible way to describe the electroweak interaction is by requiring a $\mathrm{SU}(2) \otimes \mathrm{U}(1)$ symmetry and adding a scalar field, ϕ, that spontaneously breaks the gauge symmetry. This field is a $\mathrm{SU}(2)$ doublet and has a non-zero expectation value $\langle \phi \rangle$. Using local gauge invariance under $\mathrm{SU}(2)$

rotations, the expectation value can be expressed in the form

$$\langle\phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix}, \tag{II.31}$$

where v is a real number. The transformation of the field ϕ is then given by

$$\phi \to e^{\frac{i}{2}\vec{\alpha}\cdot\vec{\sigma}} e^{\frac{i}{2}\beta}\phi, \tag{II.32}$$

where $\vec{\sigma}$ are the Pauli matrices and a charge $\frac{1}{2}$ is assigned to ϕ under U(1). By working out the matrices in the exponential, it is easy to see that the choice $\alpha^1 + \alpha^2 = 0$ and $\alpha^3 = \beta$ leave the ground state invariant. Therefore, there is one unbroken generator so that the theory contains one massless gauge boson and three massive bosons, corresponding to the three remaining broken generators.

The covariant derivative related to the field ϕ follows naturally from Eq. (II.32)

$$D_\mu\phi = \left(\partial_\mu - \frac{i}{2}gW_\mu^a\sigma^a - \frac{i}{2}g'B_\mu\right)\phi, \tag{II.33}$$

where W_μ^a are the gauge bosons of SU(2) and B_μ is the gauge boson of U(1), and g and g' are their respective coupling constants. The values for the masses of the gauge bosons are found by evaluating $|D_\mu\phi|^2$ at the expectation value given in Eq. (II.31). The physical gauge bosons, $W_\mu^\pm$, Z_μ^0 and A_μ, are obtained by making linear combinations of W_μ^a and B_μ:

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2), \\ Z_\mu^0 &= \frac{1}{\sqrt{g^2+g'^2}}(gW_\mu^3 - g'B_\mu), \\ A_\mu &= \frac{1}{\sqrt{g^2+g'^2}}(gW_\mu^3 + g'B_\mu), \end{aligned} \tag{II.34}$$

where it appears that the Z_μ^0 and the A_μ fields are orthogonal to each other. The corresponding masses are $m_W = gv/2$ for the $W^\pm$, $m_Z = (\sqrt{g^2+g'^2})v/2$ for the Z^0, and $m_A = 0$ as expected for the photon.

It was mentioned before that parity is not a symmetry of the weak interaction. Since this operation flips the handedness of a fermion, it is natural to expect left-handed (LH) and right-handed (RH) fields to transform differently under the gauge group. That is to say, LH and RH fields belong to different representations of the gauge group. More concretely, it means that LH and RH

fermions couple differently to the gauge bosons, so that the covariant derivative takes the general from

$$D_\mu = \partial_\mu - igW_\mu^a T^a - ig'YB_\mu, \tag{II.35}$$

with $T^a = \sigma^a/2$ for LH fermions and $T^a = 0$ for RH fermions, and the U(1) charge Y, called hypercharge, differs in terms of the handedness of the fermion fields. To understand the physics of these interactions, it is convenient to rewrite Eq. (II.35) in terms of the mass eigenstates of the gauge fields ($W_\mu^\pm$, Z_μ^0 and A_μ). This is achieved using "ladder operators", $T^\pm = (T^1 \pm iT^2)$. After rewriting the couplings in terms of physical observables, Eq. (II.35) becomes

$$D_\mu = \partial_\mu - ig\frac{1}{2}(W_\mu^+ T^+ + W_\mu^- T^-) - i\frac{g}{cos\theta_w}Z_\mu^0(T^3 - Q\sin^2\theta_w) - iQA_\mu, \tag{II.36}$$

where Q is the electric charge and θ_w is the weak mixing angle. This quantities are related to those appearing in Eq. (II.35) by

$$\begin{aligned} Q &= T^3 + Y, \\ e &= \frac{gg'}{\sqrt{g^2 + g'^2}}, \\ g &= \frac{e}{\sin\theta_w}, \\ \tan\theta_w &= \frac{g'}{g}, \end{aligned} \tag{II.37}$$

where e is the electromagnetic coupling. This also leads to the fact that the masses of the $W^\pm$ and the Z^0 are not independent and are related by $m_W = m_Z \cos\theta_w$.

It is known experimentally that the $W^\pm$ couples only to LH fermions. The LH fermions can be arranged into SU(2) doublets, $\psi_L = \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} e \\ \nu_e \end{pmatrix}_L$ etc., and the RH fermions into SU(2) singlets, $\psi_R = u_R, d_R, e_R$ etc. The corresponding quantum numbers are then $T = \frac{1}{2}$ with $T^3 = \pm\frac{1}{2}$ for the doublet and $T = 0$ with $T^3 = 0$ for the singlet. The transformations of these fields under SU(2) $\otimes$U(1) are then

$$\begin{aligned} \psi_L &\to e^{i\vec{T}\cdot\vec{\sigma}} e^{iY}\psi_L, \\ \psi_R &\to e^{iY}\psi_R. \end{aligned} \tag{II.38}$$

The values for Y can be obtained using the first equation in (II.37). For example, a RH electron, which has an electric charge $Q = -1$ has a hypercharge of

$Y = -1$. A summary of the different charges of up-type quarks, down-type quarks and leptons is given in Table I. As seen previously, when projecting onto LH and RH states the kinetic terms in the fermion Lagrangian take the form of Eq. (II.27). In the present case, the contents of the covariant derivative, Eq. (II.36), differ for the LH and the RH part, according to values of the quantum numbers T, T^3 and Q.

Table I: Quantum numbers of the electroweak interaction for quarks and leptons. Notice that since there is no right-handed neutrino in the SM, it does not appear in the table.

particle	T	T^3	Q	Y
u_L	1/2	1/2	2/3	1/3
d_L	1/2	-1/2	-1/3	1/3
u_R	0	0	2/3	4/3
d_R	0	0	-1/3	-2/3
ν_L	1/2	1/2	0	-1
e_L	1/2	-1/2	-1	-1
e_R	0	0	-1	-2

The LH and RH fields couple in the same way (parity is conserved) to the A_μ (the photon) and the coupling is proportional to the electric charge. The electromagnetic currents contain terms of the form $\overline{\psi}\gamma^\mu Q\psi$. Conversely, only the LH fields couple to the $W^\pm$ bosons, and because of the presence of the $T^\pm$ operators, the up- and down-type particles are mixed[9], so that the currents contain terms like $\overline{u}_L\gamma^\mu d_L$. In the case of the Z^0 boson, the current does not mix flavours[10], so it contains terms of the form $\overline{u_L}\gamma^\mu C_Z u_L$, where the coupling constant C_Z, depends on the values of T^3, Q and the weak mixing angle. The Lagrangian describing weak interactions between fermions of the i-th generation takes the form

$$\mathcal{L}_i = -\frac{g}{\sqrt{2}}\overline{\psi}_i\gamma^\mu\frac{(1-\gamma^5)}{2}(W^+_\mu T^+ + W^-_\mu T^-)\psi_i - \frac{g}{2\cos\theta_w}\overline{\psi}_i\gamma^\mu(C^i_V - C^i_A\gamma^5)\psi_i Z_\mu, \tag{II.39}$$

[9]For example $T^+\psi_L = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}u\\ d\end{pmatrix} = \begin{pmatrix}d\\ 0\end{pmatrix}$

[10]This means that flavour-changing neutral current is forbidden at tree-level in the SM. But it can still happen through loop processes.

where

$$\begin{aligned} C_V &= T_i^3 - Q_i \sin^2 \theta_w, \\ C_A &= T_i^3. \end{aligned} \tag{II.40}$$

As expected, the interaction terms related to the weak interaction violates "maximally" parity and charge conjugation. However they conserve *CP*.

II.3 *CP* violation

As seen in the previous section, the weak interaction terms contained in the Lagrangian, although they violate C and P separately, are *CP* conserving. The violation of *CP* symmetry in the SM is accounted for by the Kobayashi-Maskawa (KM) mechanism [19]. This mechanism is a generalisation to six quarks (three generations[11]) of the Cabibbo matrix [20], which was introduced by Nicola Cabibbo to preserve the universality of weak interaction in processes with s quarks, d quarks and leptons. This section will present the KM mechanism, the neutral-meson oscillations and the different types of *CP* violation. The focus will be on the quarks sector; a similar mechanism can be derived for neutrinos, leading to the PMNS matrix [21, 22], but this will not be addressed in this dissertation.

II.3.1 Kobayashi-Maskawa mechanism

In the theories described previously, the fermions remain massless. The mass terms for the fermions come from their interactions with the Higgs field. Before spontaneous symmetry breaking, the interaction terms of the quarks and the Higgs field can be expressed as

$$\mathcal{L}_{\text{Yukawa}} = -(Y^d)_{ij} \overline{\psi_L^i} \phi \, d_r^j - (Y^u)_{ij} \epsilon^{ab} \overline{\psi_{La}^i} \phi_b^\dagger u_r^j + \text{h.c.}, \tag{II.41}$$

where ψ_L^i is the SU(2) doublet corresponding to the i-th generation of quarks, u_r and d_r are RH up-type and down-type singlets, respectively, ϕ is the Higgs field and $Y^{d/u}$ are the Yukawa couplings. Given the fact that these interaction terms do not result from a gauge symmetry, no assumption can be made a priori on the the Yukawa couplings and they can be any $N \times N$ general complex matrices (assuming N generations of quarks). One can see immediately that if these matrices are real-valued the Lagrangian in Eq. (II.41) is *CP* conserving. Under spontaneous symmetry breaking, the Higgs field acquires a non-zero

[11] The mechanism is valid for at least three generations.

vacuum expectation value $\langle\Phi\rangle = (0, v/\sqrt{2})$, which yields to mass terms for the quark fields of the form

$$M^f_{ij} = \frac{v}{\sqrt{2}} Y^f_{ij}, \tag{II.42}$$

where $f = u, d$. Since the Yukawa matrices are general matrices, they are not diagonal and neither is M^f, which means that the flavour eigenstates do not coincide with the mass eigenstates. The fermion masses are obtained then by diagonalising the M^f matrices, introducing four unitary matrices $V^f_{L(R)}$

$$M^f = (V^f_L)^\dagger M^f_{\text{diag}} V_R, \tag{II.43}$$

where the mass matrices M^f_{diag} are diagonal and have real and positive eigenvalues. Similarly, the fermion fields themselves transform with the same $V^f_{L(R)}$ matrices

$$U_{L(R)} = V^u_{L,R} U'_L, \quad D_{L(R)} = V^d_{L,R} D'_L, \tag{II.44}$$

where, for N generations of quarks, $U^{(\prime)}$ is a vector that contains N up-type quarks, similarly $D^{(\prime)}$ contains N down-type quarks. The prime denotes quantities that are expressed in the flavour basis whereas the absence of prime relates to quantities in the mass basis. Finally, in the mass basis, the interaction between $W^\pm$ and the quarks takes the form

$$\mathcal{L}_W = -\frac{g}{\sqrt{2}} \overline{U}^i_L \gamma^\mu W^+_\mu (V_{\text{CKM}})_{ij} \overline{D}^j_L + \text{h.c.}, \tag{II.45}$$

where the matrix $V_{\text{CKM}} = (V^u_L)^\dagger V^d_L$, called the Cabibbo-Kobayashi-Maskawa (CKM) matrix, is unitary. In this basis, the interaction with a $W^\pm$ allows for flavour-changing transitions between up- and down-type quarks of different generations and also inside a generation.

II.3.2 CKM matrix

As discussed in the previous section, the CKM matrix is complex and unitary. The size of this matrix is directly related to the number of generations of quarks: for N generations the size of the matrix is $N \times N$ (similarly for leptons). In the most general case, a $N \times N$ complex matrix has $2N^2$ independent parameters. The number of degrees of freedom is reduced to N^2 after applying the unitarity condition. The relative phases between the quarks can be absorbed into each quark field, and thus, they are unphysical and can be rotated out. Finally, the number of parameters is reduced to $(N-1)^2$. Among these, $N(N-1)/2$ are

Euler angles and $(N-1)(N-2)/2$ are phases describing *CP* violation[12]. From there, it appears that at least three generations of quarks are needed to allow for *CP* violation.

With three generations of quarks, the CKM matrix is then a 3×3 matrix

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{II.46}$$

where the CKM matrix elements V_{ij} represent the couplings between up-type quarks (u, c, t) and down-type quarks (d, s, b). This matrix can be parametrised with 3 angles and one phase. Many parameterisations are possible; the most common one was proposed by Chau and Keung [23]. It is obtained by the product of three rotation matrices:

$$\begin{aligned} V_{\text{CKM}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \end{aligned} \tag{II.47}$$

where $s_{ij} \equiv \sin\theta_{ij}$ and $c_{ij} \equiv \cos\theta_{ij}$. The three angles θ_{12}, θ_{13} and θ_{23} are chosen so that the parameters c_{ij} and s_{ij} are positive. The parameter δ is the *CP* violating phase; it ensures that $V_{\text{CKM}} \neq V^*_{\text{CKM}}$, thus $V_{ij} \neq \overline{V_{ij}}$, allowing for *CP* violation. In the SM, considering massless neutrinos, this phase is the only source of *CP* violation[13].

It is observed experimentally [15] that $s_{13} \ll s_{23} \ll s_{12} \ll 1$. It is then convenient to express V_{CKM} in a form that exhibits this hierarchy. This can be achieved by expanding V_{CKM} in terms of $\lambda \approx |V_{us}| \approx 0.23$. This expansion is called the Wolfenstein parametrisation [24] and is widely used. At order λ^3 it reads

$$V_{CKM} = \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4), \tag{II.48}$$

[12]For leptons, in the case of Dirac neutrinos, the number of parameters is the same as in the CKM matrix. However, if the neutrinos are Majorana particles, i.e. if they are their own antiparticles, the number of phases is $N(N+1)/2$ because massive Majorana neutrino fields cannot absorb phases. Then, for three generations of leptons, there are three *CP* violating phases.

[13]In principle, a strong *CP* (and *P*) violating term is allowed in the SM. However, its value is constrained by the electric dipole moment of the neutron, which limits this strong *CP* violating term to the order 10^{-10}.

where the the Wolfenstein parameters A, λ, ρ, and η are related to the s_{ij} and c_{ij} by the following equations

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}, \quad s_{23} = A\lambda^2, \quad s_{13}e^{i\delta} = V_{ub}^* = A\lambda^3(\rho + i\eta). \tag{II.49}$$

The hierarchy between the transitions is emphasised by the parametrisation given in Eq. (II.48). For example, it is manifest that the transitions between quarks of the same generation are favoured ($V_{us} \approx V_{cs} \approx V_{tb} \propto \mathcal{O}(1)$) whereas the transitions between the first and the third generations are the most suppressed ($\mathcal{O}(\lambda^3)$). There is no explanation for this specific hierarchy in the SM.

The CKM parameters can be elegantly represented in terms of unitarity triangles using the unitarity relations

$$\sum_i V_{ij}V_{ik}^* = \delta_{jk}, \quad \sum_j V_{ij}V_{kj}^* = \delta_{ik}. \tag{II.50}$$

The six vanishing combinations can be represented as triangles in the complex plane. All the unitarity triangles have the same area, which is half the Jarlskog invariant [25], J, defined as

$$\Im(V_{ij}V_{kl}V_{il}^*V_{kj}^*) = J\sum_{m,n}\epsilon_{ikm}\epsilon_{jln}, \tag{II.51}$$

where ϵ_{ijk} is the Levi-Civita tensor. A phase-convention independent measure of *CP* violation is given by

$$\begin{aligned}\Im\det([M^u(M^u)^\dagger, M^d(M^d)^\dagger]]) =& 2J(m_t^2 - m_c^2)(m_t^2 - m_u^2)(m_c^2 - m_u^2)\\ & (m_b^2 - m_s^2)(m_b^2 - m_d^2)(m_s^2 - m_d^2),\end{aligned} \tag{II.52}$$

where M^f are the mass terms for the quark fields defined in Eq. (II.42). From this equation, it appears that J must be non-zero but also that the masses of the up-type and down-type quarks must be non-degenerated to have *CP* violation.

The most commonly used triangle for B^0-meson decays, often referred to as "*the* unitarity triangle", is obtained by taking the scalar product between the first and the third columns of the CKM matrix,

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{II.53}$$

All the sides of this triangle are of the same order of magnitude, $\mathcal{O}(\lambda^3)$. A sketch of this triangle is shown on Fig. II.1. The coordinates of the vertices

are obtained by dividing both sides of Eq. (II.53) by $V_{cd}V_{cb}^*$ and using the phase-convention-independent relation,

$$\overline{\rho} + i\overline{\eta} \equiv -\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}. \tag{II.54}$$

With this rescaling, one side of the triangle is thus equal to unity whereas the other two are given by

$$\begin{aligned} R_u &= \frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} = \sqrt{\overline{\rho}^2 + \overline{\eta}^2}, \\ R_t &= \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = \sqrt{(1-\overline{\rho})^2 + \overline{\eta}^2}. \end{aligned} \tag{II.55}$$

Phase-convention-independent quantities, such as the moduli of CKM matrix elements, $|V_{ij}|$, and their quadri-product, $V_{ij}V_{kl}V_{il}^*V_{kj}^*$, are particularly interesting because they are physically meaningful. For this reason the quantity $\overline{\rho} + i\overline{\eta}$ is preferred over $\rho + i\eta$. The Jarlskog invariant can thus be directly related to *CP* violation through

$$|\Im(V_{ij}V_{kl}V_{il}^*V_{kj}^*)| = \lambda^6 A^2 \overline{\eta}, \qquad (i \neq k, l \neq j). \tag{II.56}$$

From Eq. (II.54) the expressions for the three angles can be expressed as functions of the parameters $\overline{\rho}$ and $\overline{\eta}$:

$$\alpha = \arg\left(-\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right) = \arg\left(-\frac{1-\overline{\rho}-i\overline{\eta}}{\overline{\rho}+i\overline{\eta}}\right), \tag{II.57}$$

$$\beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right) = \arg\left(\frac{1}{1-\overline{\rho}-i\overline{\eta}}\right), \tag{II.58}$$

$$\gamma = \arg\left(-\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right) = \arg\left(\overline{\rho}+i\overline{\eta}\right). \tag{II.59}$$

In the case of the B_s^0 meson, another unitarity triangle is used. It involves CKM matrix elements related to transitions with the s quark

$$V_{us}V_{ub}^* + V_{cs}V_{cb}^* + V_{ts}V_{tb}^* = 0. \tag{II.60}$$

Unlike *the* unitarity triangle, the size of the sides of the B_s^0 triangle do not have the same order of magnitude. The first term in Eq. (II.60) is of order $\mathcal{O}(\lambda^4)$ whereas the other two terms are of order $\mathcal{O}(\lambda^2)$. This results in a "squashed" triangle. An interesting parameter to extract is the *CP*-violating phase ϕ_s, which is a mixing phase between B_s^0 and $\overline{B}_s^0$. It is defined as

$$\phi_s \equiv -2\beta_s = \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right). \tag{II.61}$$

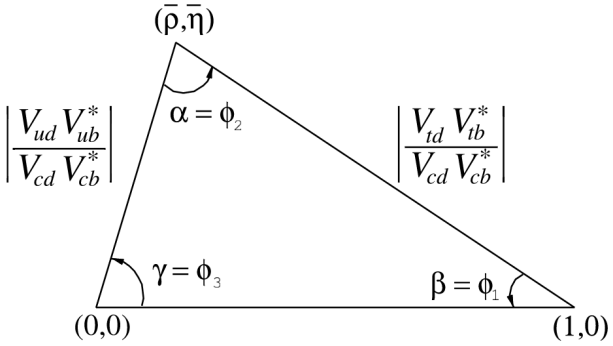


Figure II.1: Sketch of *the* unitarity triangle [15].

II.3.3 Neutral meson mixing

A consequence of the transitions between the different quark flavours is the possibility for a neutral meson to oscillate into its antiparticle[14]. This phenomenon has been observed experimentally, first in the K^0-$\overline{K}^0$ in 1956 [26], then in the B^0-$\overline{B}^0$ system in 1987 [27], in the B_s^0-$\overline{B}_s^0$ system in 2006 [28] and finally in the D^0-$\overline{D}^0$ system in 2007 [29, 30] . In the SM, mixing happens through box diagrams with the exchange of two W bosons. The diagrams describing B-meson oscillations is given in Fig. II.3.

At $t = 0$, the state describing a neutral meson, which is different from its own antiparticle, can be expressed as an admixture of flavour eigenstates M^0 and $\overline{M}^0$

$$|\psi(0)\rangle = a(0)|M^0\rangle + b(0)|\overline{M}^0\rangle, \tag{II.62}$$

At a time $t > 0$, the state evolves according to

$$|\psi(t)\rangle = a(t)|M^0\rangle + b(t)|\overline{M}^0\rangle + \sum_i c_i(t)|f_i\rangle, \tag{II.63}$$

where the sum runs over all the possible final states accessible by the system at time t. If t is very large compared to the strong interaction scale, Eq. (II.63) can be approximated by a two-state system. The effect of the other states, $|f_i\rangle$, is then accounted for by introducing a non-Hermitian part in the effective Hamiltonian describing the time evolution of the system. Since any matrix can be written as a sum of a Hermitian and an anti-Hermitian term, the Hamiltonian of the system, $\mathcal{H}$, can be written in the form

$$\mathcal{H} = \mathcal{M} - \frac{i}{2}\Gamma, \tag{II.64}$$

[14]Note that neutral baryon oscillation is forbidden by baryon number conservation.

where $\mathcal{M}$ and Γ are 2×2 Hermitian matrices related to the mass and the decay width of the state, respectively. The diagonal terms of these matrices are related to flavour-conserving transitions ($M^0 \to M^0$ and $\overline{M}^0 \to \overline{M}^0$) while the off-diagonal terms are related to flavour-changing transitions ($M^0 \to \overline{M}^0$ and $\overline{M}^0 \to M^0$). The matrix $\mathcal{M}$ is a dispersive term that describes transitions through off-shell intermediate states whereas Γ is an absorptive term associated with transitions through on-shell intermediate states. The Hamiltonian $\mathcal{H}$ is not Hermitian. Indeed, if $\mathcal{H}$ had been a Hermitian matrix, the absorptive part would have been zero and the state would have only oscillate with no possibility to decay.

In the most general case, without assuming *CPT* invariance, the eigenvectors of $\mathcal{H}$, $|M_L^0\rangle$ ("light") and $|M_H^0\rangle$ ("heavy") can be expressed as linear combinations of $|M^0\rangle$ and $|\overline{M}^0\rangle$

$$\begin{aligned} |M_L^0\rangle &\propto p\sqrt{1-z}|M^0\rangle + q\sqrt{1+z}|\overline{M}^0\rangle, \\ |M_H^0\rangle &\propto p\sqrt{1+z}|M^0\rangle - q\sqrt{1-z}|\overline{M}^0\rangle, \end{aligned} \tag{II.65}$$

where p, q and z are complex numbers, and the normalisation condition is such that $|p|^2 + |q|^2 = 1$ when $z = 0$. If *CPT* invariance is assumed then $z = 0$ and Eq. (II.65) simplifies. Other parametrisations of the mass eigenstates are possible; the one presented here is the most commonly used for describing B-meson decays. The diagonalisation of $\mathcal{H}$ yields

$$\left(\frac{q}{p}\right)^2 = \frac{\mathcal{M}_{12}^* - \frac{i}{2}\Gamma_{12}^*}{\mathcal{M}_{12} - \frac{i}{2}\Gamma_{12}}, \qquad z = \frac{\delta m - \frac{i}{2}\delta\Gamma}{\Delta m - \frac{i}{2}\Delta\Gamma}, \tag{II.66}$$

where

$$\begin{aligned} \Delta m &\equiv m_H - m_L, & \Delta\Gamma &\equiv \Gamma_H - \Gamma_L, \\ \delta m &\equiv \mathcal{M}_{11} - \mathcal{M}_{22}, & \delta\Gamma &\equiv \Gamma_{11} - \Gamma_{22}. \end{aligned} \tag{II.67}$$

The quantity Δm is positive by definition and is related to the frequency of the oscillations between the two states, whereas $\Delta\Gamma$ can be negative and, in fact, it has been established that $\Delta\Gamma < 0$ for K^0 and B_s^0 mesons [15]. If *CP* or T is conserved then

$$\arg\left(\frac{\Gamma_{12}}{\mathcal{M}_{12}}\right) = 0 \quad \Rightarrow \quad \left|\frac{q}{p}\right| = 1. \tag{II.68}$$

In addition, if *CP* is a symmetry then the mass eigenstates are orthogonal, which leads to

$$\langle M_L^0 | M_R^0 \rangle = 0 \quad \Rightarrow \quad |p|^2 - |q|^2 = 0. \tag{II.69}$$

Equations (II.68) and (II.69) are valid independently of *CPT* being conserved or not. Moreover, Eq. (II.69) is independent of T conservation as well.

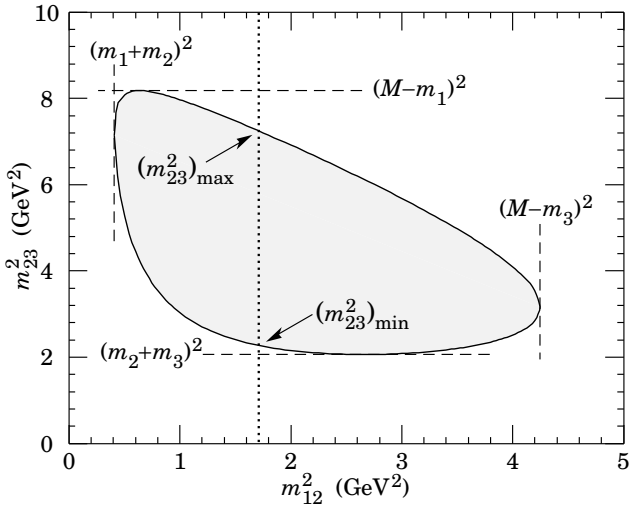


Figure II.2: Dominant diagrams corresponding to $B_{(s)}$ oscillation.

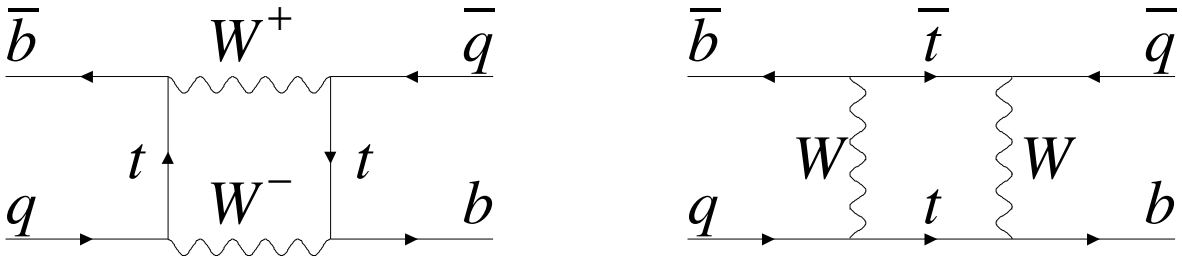


Figure II.3: Dominant diagrams corresponding to $B^0_{(s)}$ oscillation [15]. The contribution form the top quark is dominant in the loop, but contributions from u and c quarks are also possible.

Time evolution

According to quantum mechanics, an initial state evolves in time with the time-evolution operator

$$|M^0(t)\rangle = e^{-i\mathcal{H}t}|M^0(t=0)\rangle, \tag{II.70}$$

where $\mathcal{H}$ is the Hamiltonian of the system. The evolution in time of a state that is a pure $|M^0\rangle$ or $|\overline{M}^0\rangle$ at $t = 0$ is obtained by combining Eq. (II.70) and

Eq. (II.65). An initial state, which is pure at $t = 0$, evolves into a mixed state at $t > 0$ according to

$$\begin{aligned} |M^0(t)\rangle &= (g_+(t) + z g_-(t))|M^0\rangle - \sqrt{1-z^2}\frac{q}{p} g_-(t)|\overline{M}^0\rangle, \\ |\overline{M}^0(t)\rangle &= (g_+(t) - z g_-(t))|M^0\rangle - \sqrt{1-z^2}\frac{p}{q} g_-(t)|\overline{M}^0\rangle, \end{aligned} \tag{II.71}$$

where the functions $g_\pm(t)$ can be expressed in terms of the eigenvalues of the effective Hamiltonian defined in Eq. (II.64)

$$g_\pm(t) \equiv \frac{1}{2}\left(e^{-i(m_H - \frac{i}{2}\Gamma_H)t} \pm e^{-i(m_L - \frac{i}{2}\Gamma_L)t}\right). \tag{II.72}$$

The time-dependent decay rate is proportional to the number of decays of $M^0(t)$ (or its *CP* conjugate $\overline{M}^0(t)$) to a final state f (or $\overline{f}$) within t and $t + \mathrm{d}t$. It is given by

$$\begin{aligned} \Gamma(M^0(t) \to f) &= \mathcal{N}_f |\langle f|\mathcal{H}|M^0\rangle|^2, \\ \Gamma(\overline{M}^0(t) \to f) &= \mathcal{N}_f |\langle f|\mathcal{H}|\overline{M}^0\rangle|^2, \end{aligned} \tag{II.73}$$

where the time-independent normalisation factor $\mathcal{N}_f$, common to $M^0(t)$ and $\overline{M}^0(t)$, is obtained by integrating over the phase space. The time convention used hereafter is $0 < t < \infty$. A different convention is usually chosen for the study of decays of mesons produced in a coherent way (for example through $\Upsilon(4S) \to B^0\overline{B}^0$ in B factories) leading to slightly different (but equivalent) expressions for the decay rates. It is convenient to define phase-convention independent decay amplitudes of M^0 and $\overline{M}^0$

$$\begin{aligned} A_f &= \langle f|\mathcal{H}|M^0\rangle, \qquad \overline{A}_f = \langle f|\mathcal{H}|\overline{M}^0\rangle, \\ A_{\overline{f}} &= \langle \overline{f}|\mathcal{H}|M^0\rangle, \qquad \overline{A}_{\overline{f}} = \langle \overline{f}|\mathcal{H}|\overline{M}^0\rangle. \end{aligned} \tag{II.74}$$

Assuming *CPT*, $z = 0$, then Eq. (II.71) simplifies and the time-dependent decay rates, expressed in terms combinations of phase-convention independent quantities, are given by

$$\begin{aligned} \frac{\mathrm{d}\Gamma(M^0(t) \to f)}{\mathcal{N}_f e^{-\Gamma t}\mathrm{d}t} &= \left(|A_f|^2 + |\left(\frac{q}{p}\right)\overline{A}_f|^2\right)\cosh(\frac{\Delta\Gamma}{2}t) + \left(|A_f|^2 - |\left(\frac{q}{p}\right)\overline{A}_f|^2\right)\cos(\Delta m t) \\ &\quad + 2\Re\left(\left(\frac{q}{p}\right)A_f^*\overline{A}_f\right)\sinh(\frac{\Delta\Gamma}{2}t) - 2\Im\left(\left(\frac{q}{p}\right)A_f^*\overline{A}_f\right)\sin(\Delta m t), \\ \frac{\mathrm{d}\Gamma(\overline{M}^0(t) \to f)}{\mathcal{N}_f e^{-\Gamma t}\mathrm{d}t} &= \left(|\left(\frac{p}{q}\right)A_f|^2 + |\overline{A}_f|^2\right)\cosh(\frac{\Delta\Gamma}{2}t) - \left(|\left(\frac{p}{q}\right)A_f|^2 - |\overline{A}_f|^2\right)\cos(\Delta m t) \\ &\quad + 2\Re\left(\left(\frac{p}{q}\right)A_f\overline{A}_f^*\right)\sinh(\frac{\Delta\Gamma}{2}t) - 2\Im\left(\left(\frac{p}{q}\right)A_f\overline{A}_f^*\right)\sin(\Delta m t). \end{aligned} \tag{II.75}$$

In these formulae, the terms containing $|A_f|^2$ and $|\overline{A}_f|^2$ are related to the decay of the meson without a net oscillation whereas terms proportional to $|\left(\frac{q}{p}\right) A_f|^2$ and $|\left(\frac{q}{p}\right) \overline{A}_f|^2$ are related to decays that happen after net oscillation(s). The terms containing $\sinh(\frac{\Delta\Gamma}{2}t)$ and $\sin(\Delta m t)$ describe interference between the two. In the case of a multi-body final state, the decay amplitudes are phase space dependent, which means that the interference can vary between different phase-space regions, and they are strongly influenced by the resonant substructures of the decay. The expressions for the conjugate final states, $\mathrm{d}\Gamma(M^0(t) \to \overline{f})$ and $\mathrm{d}\Gamma(\overline{M}^0(t) \to \overline{f})$ can be obtained by substituting f with $\overline{f}$ in Eqs. (II.73) and (II.75).

The oscillation parameters for the neutral mesons differ significantly between K^0, B^0, B_s^0 and D^0, inducing different oscillation behaviours. The probabilities for a pure M^0 state, produced at time $t = 0$, to oscillate into M^0 or $\overline{M}^0$ are related to the square modulus of Eq. (II.72):

$$\begin{aligned}
\mathcal{P}(M^0 \to M^0) &= |g_+|^2 = \frac{1}{4}\left(e^{-\Gamma_L t} + e^{-\Gamma_H t} + 2e^{-\Gamma t}\cos\Delta m t\right), \\
\mathcal{P}(M^0 \to \overline{M}^0) &= \left|\frac{q}{p}\right|^2 |g_-|^2 = \frac{1}{4}\left|\frac{q}{p}\right|^2 \left(e^{-\Gamma_L t} + e^{-\Gamma_H t} - 2e^{-\Gamma t}\cos\Delta m t\right),
\end{aligned} \tag{II.76}$$

where the average width Γ is given by $\Gamma = (\Gamma_L + \Gamma_H)/2 = \tau^{-1}$. An illustration of time-dependent mixing for different neutral-meson systems is given in Fig. II.4. This figure illustrates several properties: in the kaon system the width difference, $\Delta\Gamma$, is very large compared to the averaged width and the mass difference, Δm, which gives the oscillation frequency; in the D^0-$\overline{D}^0$ system the mass and width differences are both very small; finally, in the B_s^0-$\overline{B}_s^0$ system, where $\Delta m/\Gamma = 26.79 \pm 0.08$ [15], the state initially produced as a B_s^0 oscillates many times before decaying.

II.3.4 Classification of *CP* violating effects

The *CP* violating effects can be classified according to the presence or absence of neutral-meson mixing. *CP* violation in the decay is characterised by a difference in the decay rate of a state and the decay rate of its *CP* conjugate. This is the only source of *CP* violation for non-mixing states such as baryons and charged mesons. In neutral-meson decays more sources of *CP* violation are available due to the mixing.

CP violation in the decay

CP violation in the decay is defined by

$$\left|\frac{\overline{A}_{\overline{f}}}{A_f}\right| \neq 1. \tag{II.77}$$

In the absence of mixing, the *CP* asymmetry is given by

$$\mathcal{A}^{CP} = \frac{\Gamma(\overline{M} \to \overline{f}) - \Gamma(M \to f)}{\Gamma(\overline{M} \to \overline{f}) + \Gamma(M \to f)}, \tag{II.78}$$

and can be related to Eq (II.77) through

$$\mathcal{A}^{CP} = \frac{\left|\overline{A}_{\overline{f}}/A_f\right|^2 - 1}{\left|\overline{A}_{\overline{f}}/A_f\right|^2 + 1}. \tag{II.79}$$

Two sorts of phases enter in the decay amplitudes: "weak" phases, which are *CP* odd and "strong" phases, which are *CP* even. In the SM, the weak phases come from the couplings of the $W^{\pm}$ bosons. In contrast, the strong phases are related to contributions from intermediate on-shell states to the decay, such as $M \to f' \to f$. Most of the rescattering processes are long-distance QCD effects. Consequently, the strong phases cannot be computed using a pertubative approach and the interpretation of the result is model dependent (except in specific cases where the strong phases can be measured experimentally or ruled out by combining several decay modes). Nevertheless, it is worth to emphasise that the *CP* conserving strong phases are necessary to observe *CP* violation effects. In their absence $\overline{A}_{\overline{f}} = A_f^*$ and thus the ratio defined in Eq. (II.77) would always be equal to one. The decay amplitudes can be expressed as a sum of contributions

$$A_f = \sum_j A_j e^{i(\delta_j + \phi_j)}, \qquad \overline{A}_{\overline{f}} = \sum_j A_j e^{i(\delta_j - \phi_j)}, \tag{II.80}$$

where A_j, δ_j and ϕ_j are, respectively, the magnitude, and the strong and weak phases of the contribution j. Note that δ_j and ϕ_j are phase-convention dependent but quantities such as $\delta_j - \delta_k$ and $\phi_j - \phi_k$ with $j \neq k$ are phase-convention independent, and thus are physical observables. At least two different contributions with non-zero strong- and weak-phase differences are needed in order to be sensitive to *CP* violation. In that case, the *CP* asymmetry can be written in terms of the magnitudes and the phases as

$$\mathcal{A}^{CP} = \frac{2A_1A_2\sin(\phi_2 - \phi_1)\sin(\delta_2 - \delta_1)}{A_1^2 + A_2^2 + 2A_1A_2\cos(\phi_2 - \phi_1)\cos(\delta_2 - \delta_1)}. \tag{II.81}$$

CP violation in the mixing

CP violation can occur in the mixing when the probability for a transition $M^0 \to \overline{M}^0$ is different from the probability of the *CP* conjugate transition, $\overline{M}^0 \to M^0$. As mentioned in section II.3.3, this happens if the mass eigenstates are not equivalent to the *CP* eigenstates. In other words, if

$$\left|\frac{q}{p}\right| \neq 1. \tag{II.82}$$

Experimentally, one can measure the corresponding *CP* asymmetry, denoted $\mathcal{A}_{\rm SL}$, in semileptonic decays of neutral mesons to final states of the form $l^{\pm}X$, where this asymmetry is the only source of *CP* violation

$$\mathcal{A}_{\rm SL} = \frac{\frac{{\rm d}\Gamma}{{\rm d}t}(\overline{M}^0(t) \to l^+X) - \frac{{\rm d}\Gamma}{{\rm d}t}(M^0(t) \to l^-X)}{\frac{{\rm d}\Gamma}{{\rm d}t}(\overline{M}^0(t) \to l^+X) + \frac{{\rm d}\Gamma}{{\rm d}t}(M^0(t) \to l^-X)} = \frac{1 - \left|\frac{q}{p}\right|^4}{1 + \left|\frac{q}{p}\right|^4}. \tag{II.83}$$

The information on whether the decay occurred before or after meson oscillation is encoded into the lepton charge. For instance, a b-quark decays semileptonically to a negative lepton, $b \to l^-\overline{\nu}X$, whereas a $\overline{b}$-quark undergoes a decay to a positive lepton, $\overline{b} \to l^+\nu X$, so that the decay to the final state $l^-\overline{\nu}X$ can happen via $\overline{B}^0 \to l^-\overline{\nu}X$ or $B^0 \to \overline{B}^0 \to l^-\overline{\nu}X$. All the decays that appear in Eq. (II.83) are proceeding through oscillations. They are often referred to as "wrong sign" decays. The standard model prediction for $\mathcal{A}_{\rm SL}$ is below 10^{-3} for the B^0 system and below 10^{-4} for the B_s^0 system. Thus, $\left|\frac{q}{p}\right| = 1$ is a good approximation for most of the experimental measurements of *CP* violation in $B^0_{(s)}$ mesons.

Mixing-induced *CP* violation

This type of *CP* violation can occur when the final state is common to M^0 and $\overline{M}^0$. In that case it can be reached with and without net oscillations ($M^0 \to \overline{M}^0 \to f$ and $M^0 \to f$), so that effects of *CP* violation in the decay and in the mixing can interfere. The quantity of interest in that case is

$$\lambda_f \equiv \frac{q}{p}\frac{\overline{A}_f}{A_f}. \tag{II.84}$$

The final state f is not necessarily a *CP* eigenstate, for example $B_s^0 \to D_s^- K^+$ and $\overline{B}_s^0 \to D_s^- K^+$ have the same final state, which it is not a *CP* eigenstate.

In the most general case, if *CP* is not conserved, then

$$\arg(\lambda_f) + \arg(\lambda_{\overline{f}}) \neq 0. \tag{II.85}$$

In the case of a decay to a *CP* eigenstate, $f_{CP} = \overline{f}_{CP}$, then $|A_{f_{CP}}| = |\overline{A}_{f_{CP}}|$ and thus

$$\Im(\lambda_{f_{CP}}) \neq 0. \tag{II.86}$$

The related time-dependent *CP* asymmetry is given by

$$\mathcal{A}_f(t) = \frac{\frac{\mathrm{d}\Gamma}{\mathrm{d}t}(\overline{M}^0(t) \to f) - \frac{\mathrm{d}\Gamma}{\mathrm{d}t}(M^0(t) \to f)}{\frac{\mathrm{d}\Gamma}{\mathrm{d}t}(\overline{M}^0(t) \to f) + \frac{\mathrm{d}\Gamma}{\mathrm{d}t}(M^0(t) \to f)}. \tag{II.87}$$

In the $B^0_{(s)}$ mesons the deviation of $|q/p|$ from unity are small and can be neglected compared to the current experimental precision. Therefore, the time-dependent *CP* asymmetry for the B^0_s meson reads

$$\mathcal{A}_f^{B^0_s}(t) = \frac{S_f \sin(\Delta m t) - C_f \cos(\Delta m t)}{\cosh(\frac{\Delta\Gamma}{2}t) - A_f^{\Delta\Gamma}\sinh(\frac{\Delta\Gamma}{2}t)}, \tag{II.88}$$

with

$$S_f \equiv \frac{2\Im(\lambda_f)}{1+|\lambda_f|^2}, \quad C_f \equiv \frac{1-|\lambda_f|^2}{1+|\lambda_f|^2}, \quad A_f^{\Delta\Gamma} \equiv -\frac{2\Re(\lambda_f)}{1+|\lambda_f|^2}, \tag{II.89}$$

where S_f is related to *CP* violation with and without the mixing and C_f to *CP* violation in the decay. These observables are related to each other by $(C_f)^2 + (S_f)^2 + (A_f^{\Delta\Gamma})^2 = 1$. For certain final states, they can be related to angles of the unitarity triangle. In the case of the B^0 meson, this expression can be further simplified by taking $\Delta\Gamma = 0$:

$$\mathcal{A}_f^B(t) = S_f \sin(\Delta m t) - C_f \cos(\Delta m t). \tag{II.90}$$

If the B meson decays to a *CP* eigenstate then $C_f = 0$, due to the fact that $|A_{f_{CP}}| = |\overline{A}_{f_{CP}}|$. In that case, there is no *CP* violation in the decay, and *CP* violation in the mixing is negligible ($|q/p| \simeq 1$), which means that the only contribution to $\mathcal{A}_f^B(t)$ comes from interferences between decays with and without mixing, $\mathcal{A}_f^B(t) = \Im(\lambda_{f_{CP}})\sin(\Delta m t)$.

Measurements of mixing-induced *CP* violation in B^0 decays give access to the angle β of the unitarity triangle. Neutral-meson mixing proceed, at first order, via box diagrams as shown in Fig. II.3, with a dominant virtual top

quark in the loop for B^0 oscillations. Thus, neglecting contributions from other virtual particles and long-range interactions we have

$$\frac{q}{p} \approx \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}. \tag{II.91}$$

In the Wolfenstein phase convention V_{tb} is real and $V_{td} = |V_{td}|e^{-i\beta}$, which leads to

$$\frac{q}{p} \approx \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*} = e^{-i2\beta + \mathcal{O}(\lambda^4)}. \tag{II.92}$$

The time-dependent *CP* asymmetry for the B^0 meson decaying to a *CP* eigenstate is given in Eq. (II.90). If there is one dominant process with a single CKM phase, then $|A_f| = |\overline{A}_f|$, there is no *CP* violation in the decay and thus

$$C_f = 0, \qquad S_f = \sin(\arg \lambda_f) = \eta_f \sin 2\theta, \tag{II.93}$$

where η_f is the *CP* eigenvalue of the final state f and 2θ is the phase difference between the decays with and without net oscillations.

In processes of the form $b \to c\bar{c}s$, such as $B^0 \to J/\psi\, K_S^0$, the Cabbibo-suppressed penguin diagrams can be neglected with respect to experimental precision. Such processes are thus dominated by a single CKM phase. In that case, λ_f can be expressed as

$$\lambda_f = -\left(\frac{q}{p}\right)_{B^0} \frac{\overline{A}_f}{A_f} \left(\frac{p}{q}\right)_{K^0}, \tag{II.94}$$

where the K^0 mixing is taken into account by the term $\left(\frac{p}{q}\right)_{K^0}$ and $\left(\frac{q}{p}\right)_{B^0}$ is given in Eq. (II.92). Expressing Eq. (II.94) in terms of CKM matrix elements, it can be shown that

$$S_f = \sin 2\beta. \tag{II.95}$$

The modes $B^0 \to J/\psi\, K_S^0$ have been extensively studied by the B factories [32, 33] and, more recently by LHCb [34]. It is possible to resolve a part of the four-fold ambiguity on β for instance by adding information from the time-dependent angular analysis of $B^0 \to J/\psi\, K^*$ decays [35, 36], which gives access to $\cos 2\beta$.

II.3.5 Experimental constraints on the CKM matrix

The CKM parameters are fundamental parameters of the SM. It is thus very important to measure them with the highest precision possible. The best

precision is usually obtained by combining measurements coming from different decay modes, and imposing SM constraints such as the unitarity of $V_{\rm CKM}$.

Another angle of approach is to perform independent precision measurements of the parameters without imposing SM constraints. This way, it is possible to test the validity of the SM and set limits on new physics. For instance, the unitarity of the CKM matrix can be tested by measuring separately the matrix elements verifying if the unitarity relations hold. Recent measurements [15] give

$$\begin{aligned} |V_{ud}|^2 + |V_{us}|^2 + |V_{ub}|^2 &= 0.9994 \pm 0.0005, \\ |V_{cd}|^2 + |V_{cs}|^2 + |V_{cb}|^2 &= 1.043 \pm 0.034, \\ |V_{ud}|^2 + |V_{cd}|^2 + |V_{td}|^2 &= 0.9967 \pm 0.0018. \end{aligned} \tag{II.96}$$

The unitarity of the CKM matrix can also be tested by over-constraining the parameters of the unitarity triangle such as the angles (unitarity implies that $\alpha + \beta + \gamma = \pi$), the sides, or the position of the apex. Many physical observables can be used to set constraints on the unitarity triangle parameters. For instance, the study of neutral-meson mixing gives access to observables such as the kaon mixing parameter ϵ_K, and $\Delta m_{d,s}$ for $B^0_{(s)}$ mixing. The three angles are determined using processes involving a b-quark, such as $b \to u\overline{u}d$ for α, and $b \to c\bar{c}s$ for β. The best precision on γ is achieved combining many $B \to Dh$ decay modes. A deep understanding of B physics is therefore crucial to perform and interpret these measurements. The current experimental constraints on the unitarity triangle are shown in Fig. II.5 and the values of the parameters are given in Table II. Note that the small experimental value of the Jarlskog invariant is due to the strong hierarchy between the quark transitions, cf. Eq. (II.52), and is in accordance with the small *CP* violation observed in the SM.

The extraction of the CKM parameters is done by using a broad set of processes. Global fits to the data coming from many processes are performed to better constrain the observables. Some parameters, such as γ and $|V_{ub}|$, can be extracted from tree-dominated decays, whereas others are measured in loop processes, which are, in general, less clean theoretically but potentially sensitive to new-physics effects. This is very interesting because it means that consistency checks between different classes of observables (tree level, loop-induced, *CP*-conserving and *CP*-violating) can be used to better understand the *CP* violation mechanism and to test for new physics.

So far, the results are consistent with the SM expectations within the current experimental precision. However, there is still room for new physics given the fact that some parameters are not yet known very accurately, such

as, for example, the angles α and γ that are known at a precision of about five degrees. Over the years, the precision on the CKM parameters improved greatly due to an overall progress of the experimental techniques and an increase in size of the datasets, which not only increases the precision but also gives access to a broader variety of processes. The uncertainties on some parameters, such as the three angles of the unitarity triangle, are still dominated by experimental uncertainties and thus there is still room for improvement from the experimental point of view. As explained before, in most *CP* violation measurements it is not possible to completely disentangle the weak- and strong-interaction effects. The hadronic effects are, in most cases, non-perturbative and thus they have to be modelled theoretically. Therefore, improvements in the theory side are as important as those in the experimental side, as they contribute to lowering the uncertainties on the results. Progress in lattice QCD, for instance, lower the systematic uncertainties associated with the extraction of non-perturbative hadronic parameters, which are dominating in parameters associated with neutral-meson mixing and matrix elements such as V_{ub} and V_{cb}.

Table II: Current values for the parameters of the unitarity triangle, the Wolfenstein parameters and the Jarlskog invariant [37]. R_u and R_t are the sides of the triangle, defined in Eq. (II.55)

angle	value [deg]
α	$86.4^{+4.5}_{-4.3}$
β	$22.14^{+0.69}_{-0.67}$
γ	$72.1^{+5.4}_{-5.7}$
$\overline{\rho}$	$0.1577^{+0.0096}_{-0.0074}$
$\overline{\eta}$	$0.3493^{+0.0095}_{-0.0071}$
R_u	$0.3832^{+0.0087}_{-0.0065}$
R_t	$0.9128^{+0.0067}_{-0.0124}$
J	$3.172^{+0.094}_{-0.098} \times 10^{-5}$

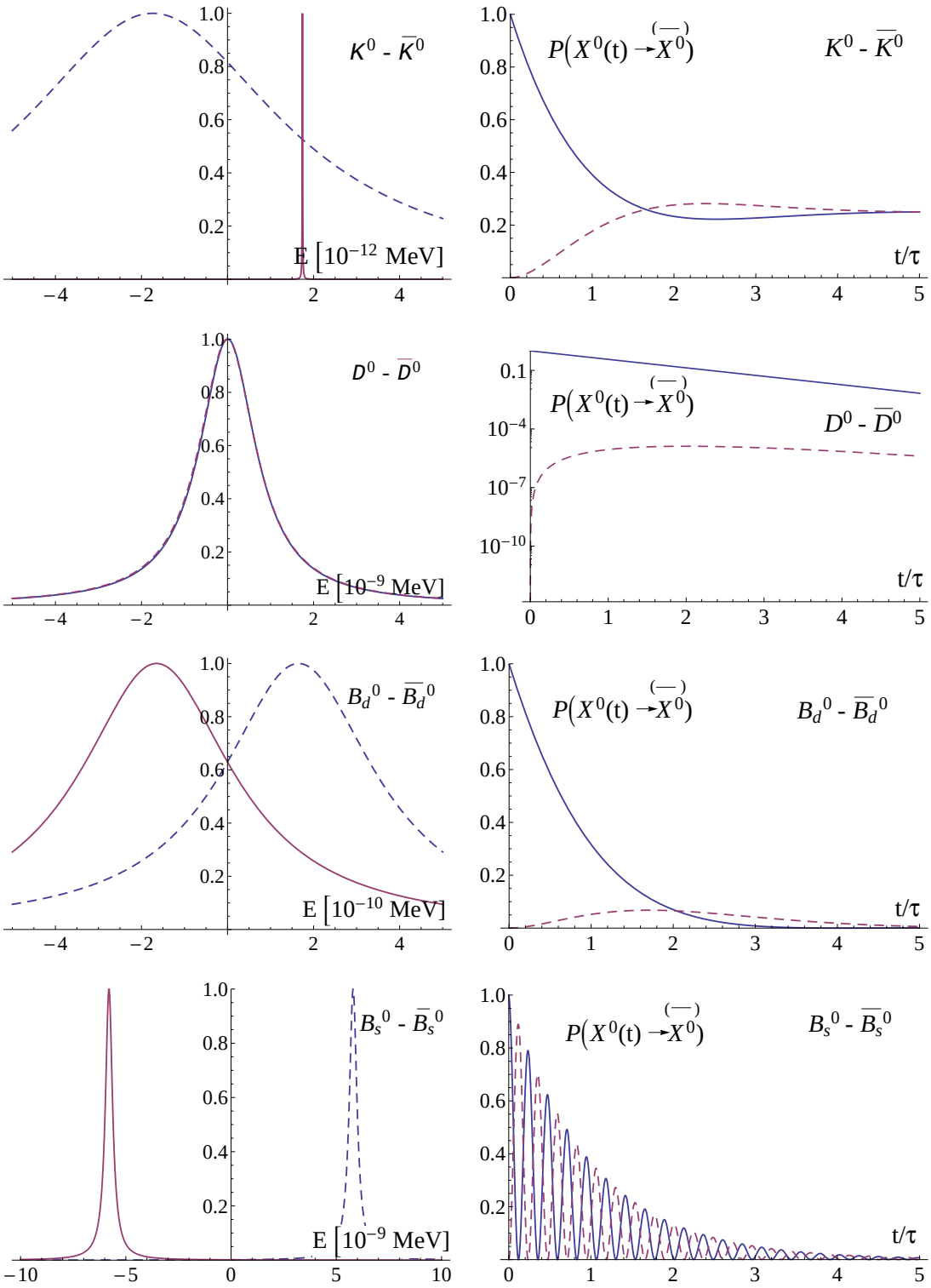


Figure II.4: Left: Mass and width differences of the mass-eigenstates of the different neutral-meson systems. One eigenstate is denoted by a full red line and the other by a dashed blue line. Right: Probabilities for an initially produced neutral meson to be found after the time t as itself (full blue line) or as its anti-particle (dashed red line) [31].

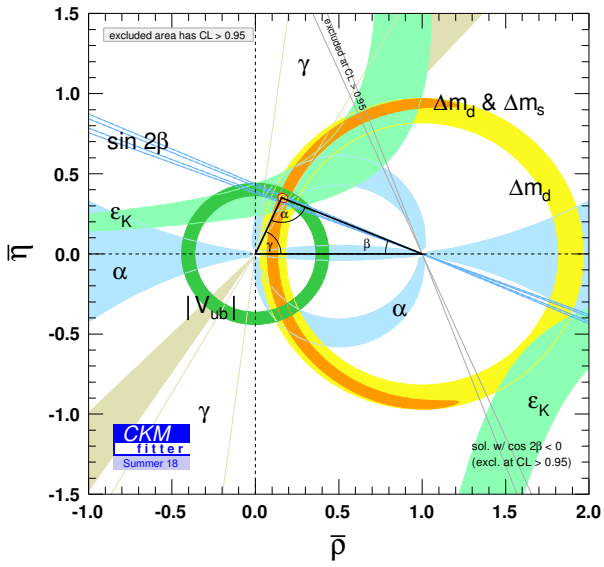


Figure II.5: Current experimental constraints on the unitarity triangle [37]. The coloured regions correspond to 95% CL.

Chapter III

Three-body charmless decays of B mesons

III.1 Charmless decays of B-mesons

The physics of charmless B-meson decays is very rich and gives access to a vast number of physical processes. The typical values of branching fractions of B-meson decays into charmless hadronic final states are below 10^{-5}, which means that large data samples (high integrated luminosity) and specific experimental techniques (regarding background rejection for instance) are needed in order to study them. Many observables can be measured in these decays, such as branching fractions, *CP* asymmetries, and CKM parameters, allowing to probe the dynamics of both weak and strong interactions. In particular, some B-meson decay modes are especially suited to measure precisely the angles of the unitarity triangle.

The B^0 and $B^\pm$ mesons were studied extensively at the B Factories, in which the energy of the e^+e^- collisions was chosen to be at the $\Upsilon(4S)$ resonance peak, leading to the production of a large amount of B^+B^- and $B^0\overline{B}^0$ pairs in a rather "clean" environment. Apart from a few runs at the $\Upsilon(5S)$ resonance in the Belle experiment, giving access to a small sample of B_s^0 mesons, the B_s^0 sector remained largely unexplored by the first-generation of B Factories. Colliders running at a higher energy, such as LEP, Tevatron, and the LHC, are (or were) able to produce, in addition to B^0 and $B^\pm$ mesons, more massive states such as B_s^0 mesons and b-baryons at the cost of a more complicated environment. In particular, the LHCb experiment is well suited to study charmless decays of both B^0 and B_s^0 mesons.

The study of *CP* violation in B_s^0-meson decays gives access to particular

CKM parameters, such as the mixing-angle ϕ_s defined in Eq. (II.61). This angle is directly related to the width difference $\Delta\Gamma_s$ through $\Delta\Gamma_s = \Gamma_{12}\cos\phi_s$. Parameters of B^0 and B_s^0 oscillations are used to constrain the unitarity triangle. For example, the oscillation frequencies (the width difference between the "light" and "heavy" states), which are measured to be $\Delta m_d = 506.4 \pm 1.9$ ns^{-1} and $\Delta m_s = 17.757 \pm 0.021$ ps^{-1} [38], provide strong constraints on the angle α.

Hadronic decays of B-mesons to charmless final states may receive rather large contributions from loop processes. For example, the creation of s quarks in the final state can only originate, in the SM, either from $b \to sq\bar{q}$ penguin diagrams (where q designates u, d or s) or $b \to u(\bar{u}s)$ tree diagrams. The former involve the CKM matrix elements $|V_{tb}|$ and $|V_{ts}|$ whereas the latter involve $|V_{ub}|$ and $|V_{us}|$, and are thus suppressed due to the small value of $|V_{ub}|$. As a consequence, B^0- (B_s^0-)meson decays in which the final state contains an odd (even) number of s quarks or kaons are favoured. Conversely to the charmless case, penguin diagrams in B-meson decays to charmed final states are suppressed as they imply the $|V_{cb}|$ and $|V_{ub}|$ CKM matrix elements.

In light of the above, the measurement of CKM parameters in processes implying B-meson decays to charmed final states is a powerful way to study SM-dominated effects, while similar measurements in loop-dominated B-meson decays open the possibility to probe new physics, as heavy virtual particles can contribute to the loop diagrams. Both approaches are important and complementary for a better understanding of the flavour sector. An interesting example of this complementarity is the measurement of the angle γ that can be obtained in both tree-dominated and loop-dominated decay modes. A short review of the main tree-level based methods is given in part III.4.1, and the method using decays containing an important contribution from loop-diagram studied in this thesis is described in section III.4.2.

Weak decays of B mesons are affected by short- and long- distance QCD effects. When studying weak-interaction processes, the contributions from strong interaction are difficult to disentangle, and the interpretation of the results in terms of weak observables only is arduous, and sometimes not possible. Furthermore, due to their mass, a large phase space is accessible in B meson decays, which encompasses different kinematic regimes. At low momentum transfer, a perturbative-QCD approach is no longer valid, which complicates the interpretation of the results as the strong parameters are less well known.

In three-body decays, the final state is no longer entirely determined by the knowledge of the initial state. The internal dynamics of the decay can be studied by means of an amplitude analysis. A common representation of three-body decays of scalar or pseudoscalar particles is the Dalitz plane, which gives a visual representation of the different sub-processes occurring in the

decay.

III.2 Effective field theory

The idea behind effective field theory (EFT) [39] is somewhat common to most, if not all, fields of physics, that is to identify the scale of interest and build a theory so that the different scales of the problem decouple. Within this approach, large and short scales are usually suppressed by the power ratio of the different scales entering the problem. For example, the study of the motion of a relatively "light" and "slow" body can usually be performed with classical mechanics, and if a more precise result is needed, relativistic corrections can be computed. Another example is the multipole expansion in electrodynamics where a precise knowledge of the charge distribution is not needed to describe the effects at large distance.

In quantum field theory the strategy is the same. After determining the domain of interest, i.e., the specific energy scale of the problem, one identifies the relevant symmetries and degrees of freedom, constructs the corresponding Lagrangian, quantises the field, and simplifies the calculations (or make them possible). This picture is complicated by the fact that in loop-diagrams the integration over the momenta runs over all scales, which makes the decoupling of the different energy scales a priori impossible. However, this apparent complexity can be overcome by using regularisation techniques [40, 41] (for example by placing a cutoff at a specific energy).

As mentioned previously, very distinct energy scales appear in the physics of B-meson decays, $m_{u,s,d} \ll \Lambda_{\rm QCD} \ll m_b \ll m_W$; $\Lambda_{\rm QCD}$ is the characteristic scale of the strong interaction below which perturbative calculations are no longer valid; its value is approximately 1 GeV. Weak interactions at a scale below the mass of the W boson can be described by an effective Hamiltonian[1] of the form

$$\mathcal{H}_{\rm eff} = \frac{G_{\rm F}}{\sqrt{2}} \sum_{p=u,c} \lambda_p^{(D)} \sum_i C_i(\mu) Q_i^p(\mu), \tag{III.1}$$

where $G_{\rm F}$ is the Fermi constant, and $\lambda_p^{(D)}$ represent CKM matrix elements defined as

$$\lambda_p^{(D)} \equiv V_{pb} V_{pD}^*, \quad \text{with} \quad p = u, c \ \text{ and } \ D = d, s. \tag{III.2}$$

The remaining terms in Eq. (III.1) are the coefficients C_i, called Wilson coefficients [42], which are effective couplings that contain *high-scale physics*, that is

[1] The Hamiltonian given here assumes only SM contributions and that $\lambda_u^{(D)} + \lambda_c^{(D)} + \lambda_t^{(D)} = 0$.

to say processes that take place at an energy scale higher than the typical scale, μ, of the problem. Generally, μ is chosen so that it separates the perturbative regime (high-scale) from the non-perturbative one. It is common to refer to the physics at a scale larger than μ as "short-distance" physics, whereas physics below μ is referred to as "long-distance" physics. The Wilson coefficients depend on μ and can be computed perturbatively to orders of the strong interaction coupling, $\alpha_s(\mu)$ (as well as perturbatively in the EW coupling, where the expansion parameter is $\ll 1$). The Q_i^p operators are effective operators (tree, gluonic and electroweak penguins, etc.) that also depend on μ and describe long-distance effects. In perturbation theory, the quantity $\sum_i C_i(\mu)Q_i^p(\mu)$ is μ independent, and so is the effective Hamiltonian in Eq. (III.1). The inclusion of NP in this picture is done by modifying the operators Q_i and the values of the Wilson coefficients.

To understand the different contributions to this effective Hamiltonian it is useful to consider its expression at first order. For a $\overline{B}$ meson that is weakly decaying into a charmless final state f it reads

$$\mathcal{A}(\overline{B} \to f) = \langle f|\mathcal{H}_{\text{eff}}|\overline{B}\rangle = \lambda_u^{(D)} A_f^u + \lambda_c^{(D)} A_f^c, \tag{III.3}$$

where the dependence on the Wilson coefficients and the effective operators is contained into the partial amplitudes $A_f^{u,c}$ and the CKM coefficients are factored out. The decay amplitude depends on three factors: the CKM factors, the Wilson coefficients and the hadronic matrix elements. Roughly speaking, the size of the Wilson coefficients for tree diagrams is of order 1 whereas for penguin diagrams it is around 0.1. For $b \to d$ transitions (i.e., $D = d$ in Eq. (III.3)), $\lambda_u^{(d)} \approx \lambda_c^{(d)} \approx \lambda^3$, and the amplitude A_f^c, which corresponds to the penguin transition, is suppressed with respect to the tree amplitude A_f^u by the Wilson coefficients. In contrast, for $b \to s$ transitions, despite the approximate factor ten between the Wilson coefficients corresponding to the tree and the penguin contributions, the CKM matrix elements related to the penguin diagram, $\lambda_c^{(s)}$, are of order λ^2 whereas $\lambda_u^{(s)} \approx \lambda^4$. Therefore, the penguin amplitude dominates for these transitions. Interferences between the favoured and the suppressed amplitudes can be a source of (direct) *CP* violation, and depend on the phases of the hadronic matrix elements $\langle f|Q_i^p|\overline{B}\rangle$. Here again, it appears that if there is no strong phase difference between the two contributing amplitudes the resulting *CP* asymmetry will be zero.

The main difficulty in obtaining theoretical predictions with this framework is the computation of the hadronic matrix elements that depend on the energy scale. The typical scale for the formation of hadronic bound states is around $\mu \approx 0.5\,\text{GeV}$, which is outside of the perturbative regime of QCD. Different

complementary approaches exist to overcome this problem. The factorisation approach is based on the identification of the different scales contributing to $\langle f|Q_i^p|\overline{B}\rangle$, such as $\Lambda_{\rm QCD}$, $\sqrt{m_b\Lambda_{\rm QCD}}$ and m_b. By doing so, it is then possible to simplify the computation of matrix elements by remarking that only the $\Lambda_{\rm QCD}$ scale requires a non-perturbative treatment while the other scales can be computed perturbatively. This reduces the complexity of obtaining $\langle f|Q_i^p|\overline{B}\rangle$ to the computation of form factors, decay constants and distribution amplitudes. The factorisation approach allows the calculation of matrix elements from first principles, but its accuracy is limited by $\Lambda_{\rm QCD}/m_b$ corrections. Several theoretical descriptions of hadronic B decays are based on this approach, the most common are QCD factorisation [43–46] (QCDF), perturbative QCD [47–51] (pQCD), and soft collinear effective theory [52–54] (SCET).

Another approach is to use flavour symmetries [55], such as flavour SU(3) (U-spin), where the masses of the u, d and s (d and s) quarks are considered to be equal. Under the flavour SU(3) assumption, for instance, the quark fields and the weak interaction Hamiltonian are decomposed into SU(3) representations and the matrix elements, $\langle f|Q_i^p|\overline{B}\rangle$, are expressed as reduced matrix elements and SU(3) Clebsh-Gordan coefficients. In practice, it is more convenient to use topological diagrams instead of the reduced matrix elements. Within this framework, the computation of matrix elements from first principles is no longer possible, but relations between different matrix elements, and thus different decay modes, can be obtained, opening possibilities for a large panel of searches. The values of the matrix elements are obtained from data and thus their precision depends on the size of the dataset. The accuracy of this approach depends on the size of the SU(3)-breaking corrections, which can be complicated to evaluate because they depend on the process and cannot be computed. These corrections are usually taken to be around 20-30%, based on the ratio of the kaon and pion decay constants.

III.3 3-body decays

III.3.1 Three-body kinematics

The interactions between initial- and final-state particles are can be described using the S-matrix, which is a $n \times n$ unitary matrix, where n is the number of channels. The probability for an initial state $|i\rangle$ to decay to a final state $|f\rangle$ is given by the matrix element

$$\mathcal{P}(i \to f) \approx |\langle f|S|i\rangle|^2. \tag{III.4}$$

In the absence of interactions, the S-matrix is simply the identity. It is thus convenient to express it as

$$S \equiv \mathbb{1} + iT, \tag{III.5}$$

where the T-matrix describes the non-trivial interactions between the initial and the final state. The four-momentum conservation is embedded into the S- and T-matrices and can be factored out

$$\langle f|iT|i\rangle = a \cdot i\mathcal{M}(i \to f), \tag{III.6}$$

where the factor a contains the four-momentum conservation; for an initial state composed of two particles of four-momenta k_1 and k_2 and a final state containing n particles of four-momenta p_j this term is $(2\pi)^4\delta^{(4)}\left(k_1 + k_2 - \sum_{j=1}^{n} p_j\right)$. The kinematics of the decay being factored out, the invariant matrix element $\mathcal{M}$ contains only the dynamics of the process.

Thus, the differential decay rate of a decay to a n-body final state, denoting the mass of the initial-state particle M and the energy and three-momentum of the j-th final state particle E_j and $\boldsymbol{p}_j$, respectively, can be expressed in terms of $\mathcal{M}$ as

$$\mathrm{d}\Gamma = \frac{(2\pi)^4}{2M}|\mathcal{M}(P \to \{p_1, ..., p_n\})|^2 \mathrm{d}\Phi_n(P; p_1, ..., p_n), \tag{III.7}$$

where Φ_n is the Lorentz invariant n-body phase space; it is given by

$$\mathrm{d}\Phi_n(P; p_1, ..., p_n) = \delta^{(4)}(P - \sum_{j=1}^{n} p_j) \prod_{j=1}^{n} \frac{\mathrm{d}^3\boldsymbol{p}_j}{(2\pi)^3 2E_j}. \tag{III.8}$$

Dalitz plot formalism

In two-body decays the kinematics of the final state is totally fixed by the knowledge of the initial state. This is not the case in three-body decays. For a final state that contains three particles, the number of degrees of freedom (dof) is twelve (each four-vector has four dof and there are three four-vectors). This number is reduced by the fact that the masses of the final state particles are known, which subtracts three dof. Applying energy-momentum conservation reduces the number of dof by four. In the case of a scalar (or pseudo scalar) particles decaying into three scalar (or pseudo scalar) particles there is no angular dependence (isotropy of the decay), and thus further two dof can be removed. Finally, the number of degrees of freedom is reduced to two. This can also be understood by the fact that in the rest frame of the decaying particle, $P = (M, 0)$ and the momenta of the final state particles form a plane.

A frequent choice of these free parameters is to take two of the three squared invariant masses of the final state particles, defined by

$$m_{ij}^2 = p_{ij}^2 = (p_i + p_j)^2 = m_i^2 + m_j^2 + 2(E_i E_j - \boldsymbol{p}_i \cdot \boldsymbol{p}_j). \tag{III.9}$$

The three invariant masses are not independent but are related by

$$m_{12}^2 + m_{23}^2 + m_{13}^2 = M^2 + m_1^2 + m_2^2 + m_3^2, \tag{III.10}$$

where m_1, m_2 and m_3 are the masses of the final-state particles and M is the mass of the initial particle. The plane of two squared invariant masses is called the Dalitz plane [56]. The knowledge of the of the final-state energies fixes the relative three-momenta and only their orientations with respect to the initial particle remain unknown. The partial decay rate can thus be expressed as

$$\mathrm{d}\Gamma = \frac{1}{(2\pi)^5}\frac{1}{16M}|\mathcal{M}|^2 \ \mathrm{d}E_1 \ \mathrm{d}E_2 \ \mathrm{d}\alpha \ \mathrm{d}(\cos\beta) \ \mathrm{d}\gamma, \tag{III.11}$$

where E_1, E_2 are the energies of the particles 1 and 2 in the rest frame of the initial particle and (α, β, γ) are the Euler angles, defining the orientation of the final-state momenta with respect to the initial particle. After integration over the Euler angles the partial decay rate becomes

$$\mathrm{d}\Gamma = \frac{1}{(2\pi)^3}\frac{1}{32M^3}\overline{|\mathcal{M}|^2}\mathrm{d}m_{23}^2\mathrm{d}m_{12}^2. \tag{III.12}$$

According to Eq. (III.9), the minimum value of m_{ij}^2 is reached when $\boldsymbol{p}_i$ and $\boldsymbol{p}_j$ are colinear. In this configuration, $\boldsymbol{p}_k$ points in the opposite direction to $\boldsymbol{p}_i$ and $\boldsymbol{p}_j$ and, according to momentum conservation, receives the largest fraction of the total momentum. Similarly, the maximum of m_{ij}^2 is reached when $\boldsymbol{p}_i$ and $\boldsymbol{p}_j$ go in opposite directions and $\boldsymbol{p}_k = 0$. Hence the m_{ij}^2 is bounded according to $(m_i + m_j)^2 \leq m_{ij}^2 \leq (M - m_k)^2$. For a fixed value of m_{12}^2, the minimum and maximum achievable values for m_{23}^2 are given by

$$\begin{aligned}
(m_{23}^2)_{min} &= (E_2^* + E_3^*) - \left(\sqrt{E_2^{*} - m_2^2} + \sqrt{E_3^{*} - m_3^2}\right), \\
(m_{23}^2)_{max} &= (E_2^* + E_3^*) - \left(\sqrt{E_2^{*} - m_2^2} - \sqrt{E_3^{*} - m_3^2}\right),
\end{aligned} \tag{III.13}$$

where E_2^* and E_3^* are the energies of the particles 2 and 3 in the rest frame of the system formed by 1 and 2. They are given by

$$E_2^* = \frac{m_{12}^2 - m_1^2 + m_2^2}{2m_{12}}, \qquad E_3^* = \frac{M^2 - m_{12}^2 - m_3^2}{2m_{12}}, \tag{III.14}$$

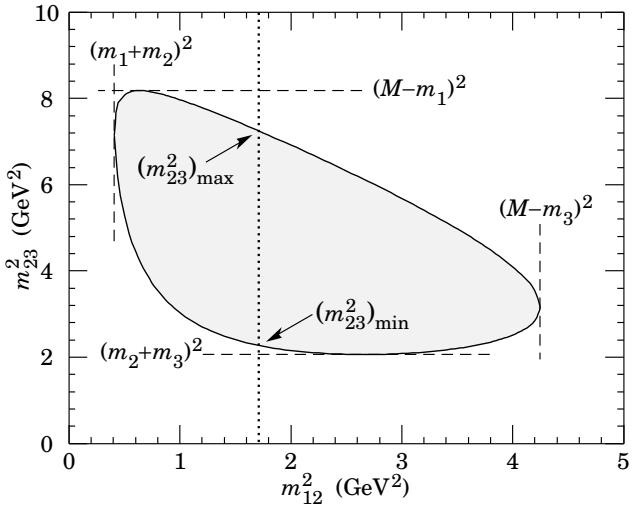


Figure III.1: Schematic view of a Dalitz plane with its kinematical boundaries [15].

The kinematic boundaries of the Dalitz plane can be visualised in Fig. III.1. The corners of the Dalitz plane correspond to configurations where one of the final state particles is at rest and the two others go in opposite direction. Along the boundaries the particles are colinear, with one particle going in one direction and the other two in the opposite direction. Inside the Dalitz plane, the three particles are no longer colinear.

The Dalitz plane is very convenient for visualising the dynamics of the decay. If there are no substructures present in the decay, then $|\mathcal{M}|^2$ is constant and the events are uniformly distributed over the Dalitz plane. On the contrary, if the decay $P_i \to P_1P_2P_3$ proceeds via intermediate states, for example $P_i \to RP_3$ where R is a resonance decaying to P_1P_2, then the events will be concentrated in a strip around the square mass of the intermediate state $m_R^2 = m_{12}^2$. The width of the strip corresponds to the decay width of the intermediate state. The distribution of the events within the band is generally not uniform and depends on angular momentum. It is thus possible to "see" the spin of the resonance by inspecting the repartition of the events on the Dalitz plane.

Square Dalitz plane

In general, the Dalitz-plane distributions of backgrounds are difficult to model with parametric functions; they are usually described by histograms taken from Monte Carlo simulation or from data sidebands. When the Dalitz-

plane boundary does not coincide with the histogram bin boundaries this may introduce biases. A common approach is to use a transformation that results in a square distribution: the square Dalitz plane. In charmless B-meson decays the events are clustered close to the boundaries of the large Dalitz plane. Another advantage of the commonly used square Dalitz plane transformation is that it gives a better resolution close to the boundaries.

The square Dalitz plane variables are defined, with a choice of two final state particles i and j, as

$$\begin{aligned} m' &\equiv \frac{1}{\pi}\arccos\left(2\frac{m_{ij}-m_{ij}^{\min}}{m_{ij}^{\max}-m_{ij}^{\min}}-1\right), \\ \theta' &\equiv \frac{1}{\pi}\theta_{ij}, \end{aligned} \tag{III.15}$$

where $m_{ij}^{\min} = m_i + m_j$ and $m_{ij}^{\max} = M - m_k$. The angle θ_{ij} is the helicity angle corresponding to the angle between the momenta of the particles k and i in the rest frame of the system formed by the particles i and j. According to these definitions, the two variables m' and θ' are defined between 0 and 1 and are dimensionless. The transformation of the $\mathrm{d}m_{ij}^2$ elements is given by

$$\mathrm{d}m_{ij}^2\mathrm{d}m_{jk}^2 \to |\det J|\mathrm{d}m'\mathrm{d}\theta', \tag{III.16}$$

where J is the Jacobian of the transformation, which satisfies

$$|\det J| = 4|\boldsymbol{p}_i^*||\boldsymbol{p}_k^*|\frac{\partial m_{ij}}{\partial m'}\frac{\partial\cos\theta_{ij}}{\partial\theta'}, \tag{III.17}$$

where the momenta $\boldsymbol{p}^*$ are defined in the rest frame of the corresponding particles i and j. Notice that, in this reference frame $|\boldsymbol{p}_i^*| = |\boldsymbol{p}_j^*|$. Thus

$$\begin{aligned} |\boldsymbol{p}_i^*| &= \sqrt{E_i^* - m_i^2} \\ |\boldsymbol{p}_k^*| &= \sqrt{E_k^* - m_k^2}, \\ \frac{\partial m_{ij}}{\partial m'} &= -\frac{\pi}{2}(m_{ij}^{\max}-m_{ij}^{\min})\sin(\pi m'), \\ \frac{\partial\cos\theta_{ij}}{\partial\theta'} &= -\pi\sin(\pi\theta'). \end{aligned} \tag{III.18}$$

An illustration of the transformation can be found in Fig. III.3 and the Jacobian is shown in Fig. III.2. As can be seen in Fig. III.3 (b), depending on its orientation in the Dalitz plane the same resonance can have different shapes in the square Dalitz plane, and if it does not decay to the particles i and j

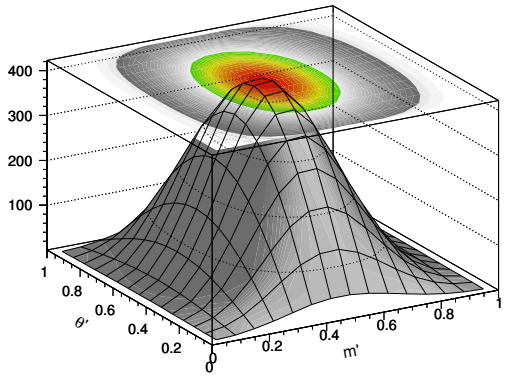
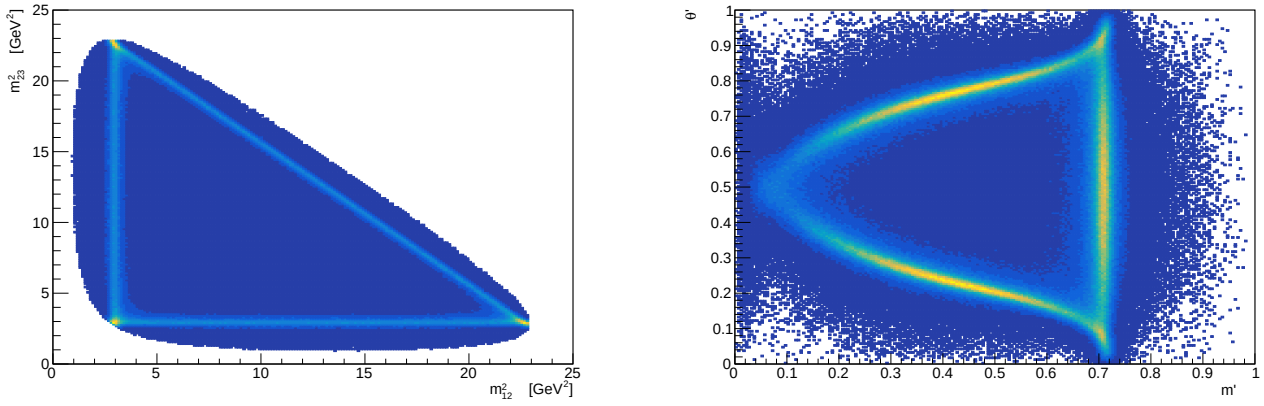


Figure III.2: Jacobian determinant of the transformation from the nominal Dalitz plane to the square Dalitz plane. The figure is from Ref. [57].

chosen to define m' it will not follow a straight line. Due to this feature the interpretation "by eye" of the decay dynamics is not straightforward anymore. The figure also illustrates the fact that, on the other hand, the events that are concentrated near the kinematical boundaries of the Dalitz plane are spread out, which is the reason why the square Dalitz plane representation is particularly suitable for analyses in bins of the phase space where rapid variations in the signal and background distributions are expected close to the Dalitz-plane boundaries.

III.3.2 Dalitz-plane amplitude analysis

A Dalitz-plane amplitude analysis is a very powerful tool to study three-body decays. It allows to establish the resonant structure of the decay and its dynamics, to characterise intermediate resonances by measuring their masses, widths, and spins. An amplitude analysis in the Dalitz plane also gives access to a variety of *CP* violating observables, such as *CP* asymmetries. These *CP* asymmetries can be measured in different regions of the *CP* or integrated over the phase space. A full time-dependent Dalitz-plot analysis with tagging information (i.e. information on the flavour of the decaying particle) can also be performed if the dataset is large enough. This type of analysis provides more information than a quasi-two-body analysis: the phases of the different contributions can be directly determined, without trigonometric ambiguities. Neutral-meson mixing can also be studied and the Dalitz-plot analysis of three-



(a) (b)

Figure III.3: Example of the transformation of a resonance in the Dalitz plane (a) to its equivalent in the square Dalitz plane (b). In this toy model, the kinematic boundaries of the Dalitz plane are fixed according to $B^0 \to K^0_S K^0_S K^0_S$ decay mode and one resonance, $f_0(1710)$, is replicated along the three directions. The variables m_{ij} and θ_{ij} in Eq. (III.15) are taken to be m_{12} and θ_{12} in this example.

body B^0 decays provides information on unitarity triangle parameters, such as $\sin 2\beta$.

The experimental study of three-body decays requires a careful modelling of the amplitude across the Dalitz plane. In the following section we will present the isobar model, which consists of a good approximation for describing decays that proceed mainly via quasi-two-body states.

III.3.3 Isobar model

The isobar model [58–60] is commonly used to describe the decay amplitude of multibody decays[2]. It is thus an important tool in amplitude analyses. In this context, the decay of a particle to a three-body final state is assumed to proceed through a resonance decaying into two particles, while the third particle (called the bachelor) remains unaffected; it is assumed to be energetic enough so it leaves the interaction region before interacting with its environment. This description is generally a good approximation for hadronic decays of B and

[2]This model is not limited to three particles in the final state and has been successfully used with more final-state particles, see for example Ref. [61].

D mesons that are expected to be dominated by two-body resonant processes. In the isobar model, the total amplitude, $\mathcal{A}$, of the decay is described as a coherent sum of n individual resonant and non-resonant decay channels $\mathcal{A}_j$. The total amplitude at each point[3] of the Dalitz plane is then given by

$$\mathcal{A}(m_{12}^2, m_{23}^2) = \sum_{j=1}^{n} \mathcal{A}_j(m_{12}^2, m_{23}^2). \tag{III.19}$$

Each individual amplitude is further decomposed as

$$\mathcal{A}_j(m_{12}^2, m_{23}^2) = c_j F_j(m_{12}^2, m_{23}^2), \tag{III.20}$$

where c_j are complex numbers called the *isobar coefficients*, giving the relative magnitudes and phases of the different components j. These coefficients are constant over the Dalitz plane. They describe the weak interaction and the momentum-independent part of the strong interaction. They are usually extracted from a fit to data. These parameters are potentially *CP* violating and *CP* asymmetries can be extracted directly form their values. The momentum-dependent part of the strong dynamics is encoded into the functions $F_j(m_{12}^2, m_{23}^2)$. Since these terms contain only the strong dynamics they are *CP* conserving. They can be expressed as products of invariant-mass and angular distributions

$$F_j(m_{12}^2, m_{23}^2) = R(m) B_l(|\boldsymbol{p}| r_{\mathrm{BW}}^P) B_l(|\boldsymbol{q}| r_{\mathrm{BW}}^R) Z_l(\boldsymbol{p}, \boldsymbol{q}), \tag{III.21}$$

where:

- $\boldsymbol{p}$ is the momentum of one of the resonance decay products and $\boldsymbol{q}$ is the momentum of the particle that does not originate from the resonance (the bachelor), both evaluated in the rest frame of the resonance[4].
- $R_i(m)$ is the resonance mass term (also called lineshape), which depends on the invariant mass, m, of the decay products of the resonance. The most common lineshapes are described below.
- $Z_l(\boldsymbol{p}, \boldsymbol{q})$ are the Zemach tensors [62, 63], which are angular distributions. Their expression depend on the spin of the resonance and will be given in the next section.

[3] In this part, we choose to express the Dalitz-plane coordinates as (m_{12}^2, m_{23}^2). This choice is of course arbitrary and any other couple of invariant masses can be used instead.

[4] In case of non-resonant components, the "resonance" in this description is simply replaced by the system of the two corresponding final-state particles.

- The B_l functions are the Blatt-Weiskopf factors [64]. Similarly to the Zemach tensors, these factors depend on the spin of the resonance. The parameters r_{BW}^P and r_{BW}^R are characteristic meson radii related to the parent and the resonance, respectively. The explicit functions will be given in the following.

Usually, the functions F_j are normalised in order to simplify the comparison between the isobar parameters. One possible normalisation consists in choosing the normalisation factor such that $\iint_{DP} |F_j(m_{12}^2, m_{23}^2)|^2 = 1$.

Quantum mechanical description of a resonance

A resonance, R, decaying into two particles a and b, $R \to ab$, can be approximated by an unstable (and non-relativistic) quantum-mechanical system. This description, inherited from nuclear physics, although not exact, is useful to understand the terms in Eq. (III.21). The wave function of the system can be written in spherical coordinates as a product between a radial term and an angular term. For spinless final-state particles, the potential entering the Hamiltonian of the system has a spherical symmetry and the wave function can be expressed in terms of spherical harmonics

$$\Psi(E, r, \theta, \phi) = \sum_{l,m} \rho_l(E, r) Y_{lm}(\theta, \phi), \tag{III.22}$$

where l is the orbital angular momentum and m its projection. Considering that the system has an axial symmetry, the solution can be expressed in term of Legendre polynomials, $P_l(\cos\theta)$,

$$\Psi(E, r, \theta, \phi) = \sum_{l=0}^{\infty} (2l+1) \rho_l(E, r) P_l(\cos\theta). \tag{III.23}$$

Note that in the case of a decay to particles with non-zero spins, the Legendre polynomials no longer provide a valid description and they are replaced by angular functions such as the Zemach tensors. The lineshapes and the Blatt-Weiskopf factors mentioned previously come from the study of the radial part of the wave function. More precisely, the lineshapes are obtained by studying the system at a fixed energy around that of the resonance, and the Blatt-Weiskopf factors arise when studying the system at a fixed $r = r_{\text{BW}}$, where r_{BW} is the effective radius of the resonance.

The term $\rho_l(E, r)$ must satisfy the radial Schrödinger equation. Under the convention $\hbar = c = 1$, it is written

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} + \frac{2m}{\hbar^2}(E - V(r)) - \frac{l(l-1)}{r^2} \right] (r \rho_l(E, r)) = 0, \tag{III.24}$$

where the reduced mass m is given by

$$m = \frac{m_a m_b}{m_a + m_b}. \tag{III.25}$$

In the form given in Eq. (III.24), it appears that the equation is equivalent to the one-dimension Schrödinger equation for a particle of mass m in an effective potential given by

$$V_{\text{eff}}(r) = V(r) + \frac{1}{2m}\frac{l(l+1)}{r^2}. \tag{III.26}$$

The orbital-dependent term of the effective potential is called the centrifugal (or kinematic) potential and describes the behaviour of the outgoing final state particles, whereas the first term, the dynamical potential, is linked to the properties of the resonance.

The potential entering the Hamiltonian of the system should reflect the properties of the resonance. As already motioned, in the case of spinless particles it should have a spherical symmetry. It should also describe the short-range strong interactions between the daughter particles of the resonance within its the effective radius. Furthermore, the barrier must be large enough to ensure the existence of a quasi-steady state solution of the corresponding Schrödinger equation. In the region $r > r_{\text{BW}}$ the interaction between the particles becomes negligible and the potential should tend towards zero fast enough to ensure asymptotic freedom of the daughter particles. These requirements can be fulfilled by considering a potential that is modelled as a combination of a well and an effective potential similar to Eq. (III.26), which can be written as

$$U(r) = \begin{cases} -U_0(r) & r \leq r_{\text{BW}} \\ V(r) + \frac{1}{2m}\frac{l(l+1)}{r^2} & r > r_{\text{BW}} \end{cases} \tag{III.27}$$

The crossing symmetry[5] allows to describe particle decay and scattering processes in a similar fashion. The asymptotical behaviour ($r \to \infty$) of a state corresponding to a particle scattered by a potential similar to that of Eq. (III.27) is given by

$$\psi(k, r, \theta) = e^{ikr\cos\theta} + f(k, \theta)\frac{e^{ikr}}{r} + \mathcal{O}(\frac{1}{r^2}), \tag{III.28}$$

where k is the wave number, $k \equiv \sqrt{2mE}$ and θ is the angle between the incoming and the scattered particles. The first exponential term in this

[5]The crossing symmetry states that the S-matrix for a process involving an incoming particle with momentum p is identical to the S-matrix for the same process with an outgoing particle of momentum $-p$.

equation corresponds to the propagation of a free particle in the direction $r\cos\theta$. The second term describes a spherical wave propagating far away from the potential. The function $f(k,\theta)$ is called the scattering amplitude and is, in general, independent of r. It can be decomposed in terms of partial waves

$$f(k,\theta) = \frac{1}{k}\sum_{l=0}^{\infty}(2l+1)f_l P_l(\cos\theta), \tag{III.29}$$

where f_l is the partial-wave amplitude with orbital momentum l. Its expression is given by

$$f_l = \frac{(\eta_l e^{2i\delta_l}-1)}{2i}, \tag{III.30}$$

where η_l is the elasticity parameter (also called inelasticity). It is defined between zero and one, and is equal to one for purely elastic scattering. The parameter δ_l corresponds to the phase shift of the l^{th} partial wave. The evolution of f_l with the energy as a trajectory in an Argand plot is shown in Fig. III.4. The differential cross section is then given by

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f(k,\theta)|^2, \tag{III.31}$$

where $\mathrm{d}\Omega$ represents the infinitesimal phase space element.

The scattering amplitude can be related to the Lorentz-invariant matrix element $\mathcal{M}$ through

$$\mathcal{M} = -8\pi\sqrt{s}f(k,\theta). \tag{III.32}$$

As stated before, the mass distributions are obtained by studying quasi-stationary states of the Schrödinger equation. For a resonance R with an average lifetime τ, the stationary condition, $\mathrm{d}\psi/\mathrm{d}t \approx 0$ for $t \ll \tau$, leads to the following solution

$$\psi(E,r,\theta,\phi,t) = \psi(E,r,\theta,\phi)e^{-iEt}, \tag{III.33}$$

where E is an eigenvalue of the Hamiltonian. Since the system can decay, E is complex (the energy of a decaying system is not a physical observable), and it can be written as

$$E = E_0 - i\frac{\Gamma}{2}, \tag{III.34}$$

where E_0 is the energy of the resonance and Γ is the decay rate. By expanding the partial wave amplitude around the resonance energy, E_0, it can be shown that

$$f_l = \frac{\Gamma/2}{E_0 - E - i\frac{\Gamma}{2}}, \tag{III.35}$$

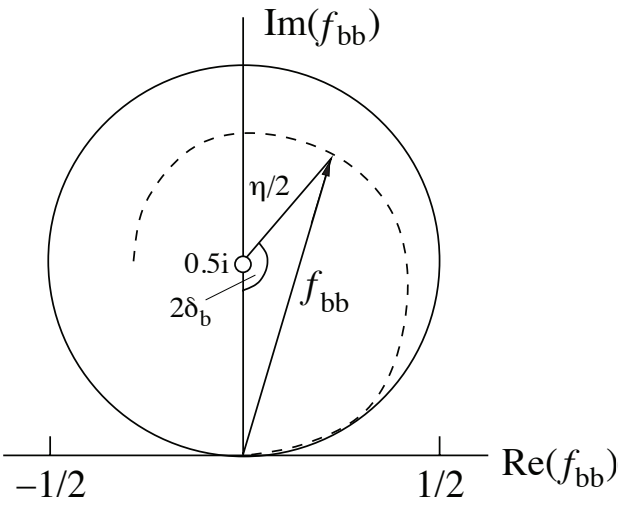


Figure III.4: Trajectory of a partial-wave amplitude, noted f_{bb} here, as a function of the energy. The unitary circle represents the case where the scattering is purely elastic ($\eta_l = 1$) [15]. The dashed line corresponds to inelastic scattering ($\eta_l < 1$). It clearly appears in this diagram that at $\delta_b = \pi/2$ the energy is real, which corresponds to the pole mass of the resonance.

which is the (non relativistic) Breit-Wigner distribution (BW). From this equation, it appears that the asymptotic value of the phase of f_l at $E \ll E_0$ ($E \gg E_0$) is 0 (π), that the change of phase occurs around the energy of the resonance and that the phase at $E = E_0$ is $\frac{\pi}{2}$. A sketch of the behaviour of the BW distribution can be seen in Fig. III.5. In actual Dalitz-plot analyses, more complex propagators that account for relativistic effects are used instead of the simple BW, they are discussed in a dedicated paragraph below.

Blatt-Weiskopf factors

These factors are obtained by studying the system around the effective radius $r_{\rm BW}$. They are related to the transmission coefficients $T(k, k_0; r_{\rm BW}, \infty)$ which describe the probability for the particle to be at $r_{\rm BW}$ and to leave the system with a wave vector k, relative to the probability to be at $r_{\rm BW}$ and to leave with a wave vector k_0. In the case of a decay of a particle into three final-state particles proceeding through a resonance, $P_i \to RP_3 \to P_1P_2P_3$, the amplitude will contain two Blatt-Weiskopf factors: the first describes the probability that the resonance R overcomes the centrifugal potential of RP_3, and the second

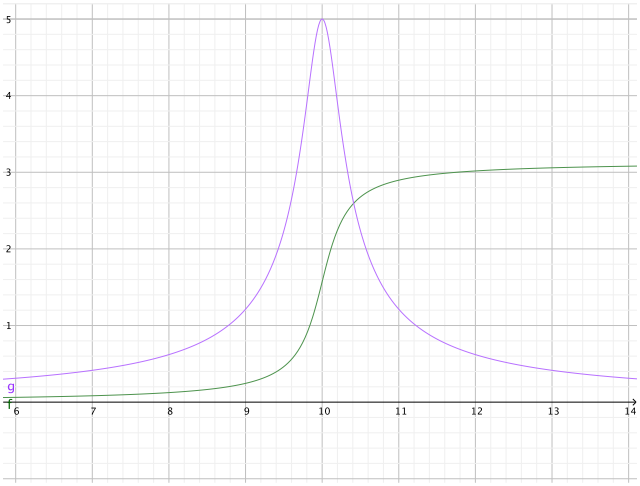


Figure III.5: Sketch of the magnitude (in purple) and phase (in green) of a BW distribution for $E_0 = 10$, $\Gamma = 0.5$ and an arbitrary normalisation. The change of phase around $E = E_0$ is clearly seen.

is related to the probability that the particle P_1 overcomes the centrifugal potential of P_1P_2.

The Blatt-Weiskopf [64] factors up to $l = 3$ are given by

$$\begin{aligned} B_{l=0}(z) &= 1, \\ B_{l=1}(z) &= \sqrt{\frac{z_0^2+1}{z^2+1}}, \\ B_{l=2}(z) &= \sqrt{\frac{z_0^4+3z_0^2+9}{z^4+3z^2+9}}, \\ B_{l=3}(z) &= \sqrt{\frac{z_0^6+6z_0^4+45z_0^2+225}{z^6+6z^4+45z^2+225}}, \end{aligned} \tag{III.36}$$

where $z = |\boldsymbol{p}|r_{\mathrm{BW}}^P$ or $z = |\boldsymbol{q}|r_{\mathrm{BW}}^R$, and z_0 is the value of z when m is equal to the nominal mass of the resonance m_0. The superscripts P and R refer, respectively, to the radius corresponding to the parent particle and the radius of the resonance[6].

[6]Different conventions exist the the definition of Blatt-Weiskopf factors. With the convention adopted here, the momentum of the bachelor particle is computed in the rest frame of the resonance. An alternative, but not equivalent, parametrisation uses the

Zemach tensors

The Zemach tensors [62, 63] give the angular probability distribution according to the orbital momentum, l, between the two daughters of the resonance. For any integer orbital momentum, they are proportional to

$$T_l(\boldsymbol{p},\boldsymbol{q}) \propto (-2|\boldsymbol{p}||\boldsymbol{q}|)^l P_l(\cos\theta), \tag{III.37}$$

where θ is the *helicity angle*, defined between p and q and $P_l(\cos\theta)$ are Legendre polynomials. In the case where the mother particle and its daughters have zero spins, l is simply the spin of the resonance. The explicit formulae up to $l = 3$ are given by

$$\begin{aligned} Z_{l=0}(\boldsymbol{p},\boldsymbol{q}) &= 1, \\ Z_{l=1}(\boldsymbol{p},\boldsymbol{q}) &= -2\boldsymbol{p}\cdot\boldsymbol{q}, \\ Z_{l=2}(\boldsymbol{p},\boldsymbol{q}) &= \frac{4}{3}\left(3(\boldsymbol{p}\cdot\boldsymbol{q})^2 - (|\boldsymbol{p}||\boldsymbol{q}|)^2\right), \\ Z_{l=3}(\boldsymbol{p},\boldsymbol{q}) &= -\frac{24}{15}\left(5(\boldsymbol{p}\cdot\boldsymbol{q})^3 - 3(\boldsymbol{p}\cdot\boldsymbol{q})(|\boldsymbol{p}||\boldsymbol{q}|)^2\right). \end{aligned} \tag{III.38}$$

For a resonance with a zero angular momentum, the probability distribution is equal to unity, meaning that there is no preferred direction in space, as expected in this case. It appears that these factors are not bound for large p, which means that the amplitude could be made infinitely large. Fortunately, this behaviour is counteracted by the Blatt-Weiskopf factors.

Alternatives to the Zemach tensor formalism exist, such as the helicity framework [65]. For spinless particles, the two formalisms are equivalent. A relativistic version of the Zemach formalism has also been derived [66].

Mass distributions

In the following, we will list the lineshapes used by in the BaBar analyses of the charmless three-body decays used to extract γ in chapter IV. The expressions that are given below are the actual lineshapes implemented in the Laura++ Dalitz-plot fitter [67], used in this work.

Relativistic Breit-Wigner

The Relativistic Breit-Wigner (RBW) is the most commonly used lineshape. It is particularly suited to narrow, isolated, resonances. This distribution

momentum of the bachelor particle computed in the rest frame of the parent particle. A covariant formulation is also available.

corresponds to the relativistic version of Eq. (III.35), that is expressed in terms of invariant masses instead of energies. It is given by

$$R(m) = \frac{1}{(m_0^2 - m^2) - im_0\Gamma(m)}, \tag{III.39}$$

where m_0 is the nominal mass of the resonance. The decay rate depends on the invariant mass and is given by

$$\Gamma(m) = \Gamma_0 \left(\frac{q}{q_0}\right)^{2l+1} \left(\frac{m_0}{m}\right) B_l^2(qr_{\text{BW}}^R), \tag{III.40}$$

where Γ_0 is the nominal width of the resonance and q_0 is $q(m = m_0)$. The dependency in mass of the width makes $R(m)$ non-analytic. Close to the peak of the resonance, the BW behaviour is recovered.

Flatté

The Flatté lineshape [68] is similar to the RBW but the width term is given by the sum of the widths of two coupled channels. These distributions are suited for resonances like the $f_0(980)$, $a_0(980)$, $K_0^*(1430)$. For example, in the case of the $f_0(980)$, which can decay in the KK or $\pi\pi$ channels, the mass-dependent width is given by

$$\Gamma(m) = \Gamma_{\pi\pi}(m) + \Gamma_{KK}(m), \tag{III.41}$$

with

$$\begin{aligned} \Gamma_{\pi\pi}(m) &= g_\pi \left(\frac{1}{3}\sqrt{1 - 4\frac{m_{\pi^0}^2}{m^2}} + \frac{2}{3}\sqrt{1 - 4\frac{m_{\pi^\pm}^2}{m^2}}\right), \\ \Gamma_{KK}(m) &= g_K \left(\frac{1}{2}\sqrt{1 - 4\frac{m_{K^0}^2}{m^2}} + \frac{1}{2}\sqrt{1 - 4\frac{m_{K^\pm}^2}{m^2}}\right), \end{aligned} \tag{III.42}$$

where g_π and g_K are couplings constants and the fractional coefficients arise from isospin conservation.

Gounaris-Sakurai

This Gounaris-Sakurai (GS) distribution [69] is an alternative to the RBW for broad vector mesons decaying into two pions such as the ρ resonances. In contrast to the RBW, the GS is an analytic function:

$$R(m) = \frac{1 + D\frac{\Gamma_0}{m_0}}{(m_0^2 - m^2) + f(m) - im_0\Gamma(m)}, \tag{III.43}$$

where D is a constant given by

$$D = \frac{3}{\pi}\frac{m_\pi^2}{q_0^2}\,\ln\left(\frac{m_0+2q_0}{2m_\pi}\right) + \frac{m_0}{2\pi\,q_0} - \frac{m_\pi^2 m_0}{\pi\,q_0^3}, \tag{III.44}$$

and the function $f(m)$ writes

$$f(m) = \Gamma_0\,\frac{m_0^2}{q_0^3}\,\left[\,q^2\,[h(m)-h(m_0)] + \left(\,m_0^2 - m^2\,\right)\,q_0^2\,\left.\frac{dh}{ds}\right|_{m_0}\right], \tag{III.45}$$

with

$$h(m) = \frac{2}{\pi}\,\frac{q}{m}\,\ln\left(\frac{m+2q}{2m_\pi}\right), \tag{III.46}$$

and

$$\left.\frac{dh}{ds}\right|_{m_0} = h(m_0)\left[(8q_0^2)^{-1} - (2m_0^2)^{-1}\right] + (2\pi m_0^2)^{-1}. \tag{III.47}$$

LASS

The LASS parametrisation [70] was developed to describe the $(K\pi)$ S-wave component. It is modelled as a $\overline{K}^{*0}(1430)$ interfering with a nonresonant component:

$$R(m) = \frac{m}{q\cot\delta_B - iq} + e^{2i\delta_B}\frac{m_0\Gamma_0\frac{m_0}{q_0}}{(m_0^2-m^2) - im_0\Gamma_0\frac{q}{m}\frac{m_0}{q_0}}\,, \tag{III.48}$$

$$\text{with}\ \cot\delta_B = \frac{1}{aq} + \frac{1}{2}rq\,, \tag{III.49}$$

where a is the scattering length and r is the effective range parameters.

Nonresonant terms

The nonresonant amplitude accounts for effects that cannot be described as quasi two-body decays. In most analyses, this component is found empirically by choosing the form that best describes the data. The most simple form used is a constant complex number (magnitude and phase) over the whole Dalitz plane. Another frequent choice consists of exponential form factors of the form $e^{-\alpha m^2}$, where the parameter α is determined from the data. However, these models are generally not sufficient to describe final-state interactions and other effects coming from non-perturbative QCD. Also, for some decay modes these simple form factors are not sufficient to describe the data and

some analyses use more complicated functions containing more free parameters, such as, for example $B \to KKK$ [71], where the nonresonant term is composed of two second-order invariant-mass polynomial lineshapes accounting for S- and P-wave terms.

Dalitz-plot observables

A Dalitz-plot analysis of a given decay mode usually proceeds via steps of increasing complexity. After the measurement of the branching fractions, the first step is generally a time-integrated and flavour-averaged analysis. Then, in the case of neutral mesons, if the data sample is large enough, a flavour-tagged time-dependent analysis can be performed, with a description for $|\mathcal{A}(t)|^2$ based on the formulae obtained in Eq (II.75).

A number of observables are easily accessible from the knowledge of the isobar parameters c_j and the hadronic functions $F_j(m^2_{12}, m^2_{23})$. The relative rate of an isobar component j is given by the (flavour-averaged) *fit fractions*, defined as

$$FF_j \equiv \frac{\iint_{\rm DP} \left(|\mathcal{A}_j(m^2_{12}, m^2_{23})|^2 + |\overline{\mathcal{A}}_j(m^2_{12}, m^2_{23})|^2\right) {\rm d}m^2_{12}{\rm d}m^2_{23}}{\iint_{\rm DP} \left(|\mathcal{A}(m^2_{12}, m^2_{23})|^2 + |\overline{\mathcal{A}}(m^2_{12}, m^2_{23})|^2\right) {\rm d}m^2_{12}{\rm d}m^2_{23}}, \qquad \text{(III.50)}$$

where the integrals are over the whole Dalitz plane, the total amplitude $\mathcal{A}$ is defined in Eq. (III.19) as a coherent sum of the components $\mathcal{A}_j$, and $\overline{\mathcal{A}}$ is similarly obtained with and their *CP* conjugates $\overline{\mathcal{A}}_j$:

$$\mathcal{A}_j(m^2_{12}, m^2_{23}) = c_j F_j(m^2_{12}, m^2_{23}), \quad \overline{\mathcal{A}}_j(m^2_{12}, m^2_{23}) = \overline{c}_j \overline{F}_j(m^2_{12}, m^2_{23}). \quad \text{(III.51)}$$

Interferences between two resonances j and k can be quantified by the corresponding *interference fit fractions*, given by

$$FF_{jk} \equiv 2\Re \frac{\iint_{\rm DP} \left(\mathcal{A}^*_j(m^2_{12}, m^2_{23})\mathcal{A}_k(m^2_{12}, m^2_{23}) + \overline{\mathcal{A}}^*_j(m^2_{12}, m^2_{23})\overline{\mathcal{A}}_k(m^2_{12}, m^2_{23})\right) {\rm d}m^2_{12}{\rm d}m^2_{23}}{\iint_{\rm DP} \left(|\mathcal{A}(m^2_{12}, m^2_{23})|^2 + |\overline{\mathcal{A}}(m^2_{12}, m^2_{23})|^2\right) {\rm d}m^2_{12}{\rm d}m^2_{23}}. \qquad \text{(III.52)}$$

If FF_{jk} is positive the interference between the two resonances is constructive, while in the case it is negative their interference is destructive. With the definitions of Eqs. (III.50) and (III.52) we have

$$\sum_j FF_j + \sum_{j<k} FF_{jk} = 1. \qquad \text{(III.53)}$$

Note that Eqs. (III.50) and (III.52) are flavour-averaged. If the flavour of the decaying meson is known, i.e. if the decay mode is flavour specific or if

flavour-tagging is used, it is possible to differentiate between $\mathcal{A}$ and $\overline{\mathcal{A}}$ and thus flavour-specific fit fractions can be computed as well.

When $\mathcal{A}$ and $\overline{\mathcal{A}}$ are distinguishable, for each intermediate state, a *CP* asymmetry can be assessed as

$$A_j^{CP} \equiv \frac{\int\int_{\rm DP}(|\overline{\mathcal{A}}_j(m_{12}^2, m_{23}^2)|^2 - |\mathcal{A}_j(m_{12}^2, m_{23}^2)|^2){\rm d}s_1{\rm d}s_2}{\int\int_{\rm DP}(|\overline{\mathcal{A}}(m_{12}^2, m_{23}^2)|^2 + |\mathcal{A}(m_{12}^2, m_{23}^2)|^2){\rm d}s_1{\rm d}s_2}. \tag{III.54}$$

This formula gives the *CP* asymmetry integrated over the whole Dalitz plane; it is also possible to obtain localised *CP* asymmetries by integrating over regions of the Dalitz plane. Since the functions $F_j(m_{12}^2, m_{23}^2)$ are *CP* even (they contain only the strong dynamics), they cancel out in Eq. (III.54) and thus the *CP* asymmetry can be written in terms of the isobar coefficients only as

$$A_j^{CP} = \frac{|\overline{c}_j|^2 - |c_j|^2}{|\overline{c}_j|^2 + |c_j|^2}. \tag{III.55}$$

Similarly, for a neutral particle decaying to a flavour eigenstate, mixing-induced *CP* asymmetry writes

$$S_j = \frac{2\Im(\overline{c}_j c_j^*)}{|\overline{c}_j|^2 + |c_j|^2}. \tag{III.56}$$

Note that, given the fact that the isobar coefficients contain a strong phase then $\overline{c}_j \neq c_i^*$. Since amplitude analyses give direct access to phases and to the interferences between resonances, it is possible to use this information to extract specific observables such as

$$\delta\phi(j,k) = \arg(\overline{c}_j, \overline{c}_k^*) - \arg(c_j, c_k^*), \tag{III.57}$$

which is a direct *CP* violation observable. Time-dependent mixing-induced *CP* asymmetry can be quantified by the angle $\beta_{\rm eff}$

$$\beta_{\rm eff,j} = \frac{1}{2}\arg(c_j, \overline{c}_j^*). \tag{III.58}$$

Usually the notation $\beta_{\rm eff}$ is used for a measurement based on loop processes, to distinguish it from β, which is measured with tree-level processes as, for instance, using the "golden mode" $B^0 \to J/\psi K_{\rm S}^0$. The difference between β and $\beta_{\rm eff}$ is due to SM QCD corrections and, potentially, the contributions from new physics in the loops.

III.3.4 Limitations and alternatives to the Isobar model

The isobar model is a convenient and powerful approximation. However, it is not completely satisfactory. First of all, the description of the Dalitz plane in terms of isobars is model dependent, and thus there is an irreducible, non-trivial, uncertainty due to the modelling. This model is also unsatisfactory from the theory point of view: indeed, the addition of more than one Breit-Wigner propagator in a channel violates the unitarity of the S-matrix, resulting in the non-conservation of the probability currents.

The accuracy of the model also depends on the knowledge of the hadronic parameters that enter the functions describing the strong dynamics (cf. the terms $F_j(m^2_{12}, m^2_{23})$ in Eq. (III.21)), which are not always well assessed by the theory or accessible from the data. Furthermore, the description of resonances is based on a quantum-mechanical approach, leading to barrier factors and angular functions that are non-relativistic. This may prove to be insufficient even if relativistic corrections of these terms exist.

As mentioned before, the Breit-Wigner-like propagators are well suited for narrow resonances, so the P- and D-waves, which usually proceed via narrow and isolated resonances, are in general rather well described by the isobar model. On the other hand, the S-wave is often more intricate and contains many overlapping broad states for which the Breit-Wigner approach does not give a satisfactory description. In addition to this, the BW parameters, such as the mass and the width, are usually reaction dependent and thus can vary between different decay modes.

In principle, three-body decays do not exclusively proceed via quasi-two-body decays, and long-distance effects, such as final-state interactions (e.g. rescattering), are not (or hardly) taken into account by the isobar model. Moreover, unlike the case of charm decays, where the contribution from nonresonant (NR) component is generally rather small, NR effects dominate in many decay modes of the B meson. This can be explained by the rather large phase space accessible to the B meson. For three-body charmless B decays the maximum value for the invariant mass is of order 5 GeV whereas most of the resonances are localised on the boundaries of the Dalitz plane, between 0.5 and 2 GeV, which explains the large size of the NR component. Such a large phasespace actually contains different kinematic regions that correspond to distinct QCD regimes. This indicates that a unique description for the whole Dalitz plane may not be sufficient.

To tackle some of these issues, several attempts have been made to improve the model, such as the addition of relativistic corrections to the lineshapes and to the angular distributions. In a recent LHCb analysis of the $B^\pm \to$

$\pi^{\pm}K^{+}K^{-}$ decay mode [72], the use of a lineshape describing the $\pi\pi \leftrightarrow KK$ rescattering [73], together with a NR form factor accounting for partonic interaction in the final state [74], result in a good description of the data and of the large CP asymmetry observed in the low invariant mass region.

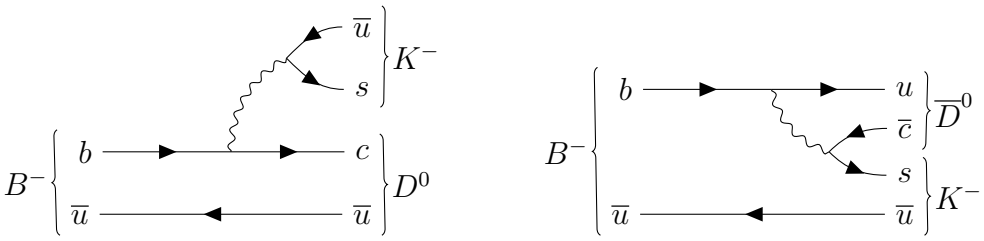
Other approaches than the isobar model have been developed, such as the K-matrix approximation [75, 76]. The K-matrix is hermitian, real and symmetric. It is defined in terms of the T-matrix introduced in Eq. (III.5) as $K \equiv (T^{-1} + i\mathbb{1})^{-1}$. This formalism is still model dependent and is based on approximations (the S-wave is supposed not to interact with the rest of the final state particles) but it has the advantage of preserving unitarity. Since the K-matrix approximation is more complex than the isobar model, it is generally used to parameterise the S-wave only, especially when it contains many broad, overlapping resonances; the other resonances being described using isobars. The improvement of Dalitz-plot analysis techniques can only be achieved via a joint effort between the theory and experimental communities. Previous LHCb results showing a large CP violation at low invariant mass of $B \to hh'h''$ decays [77], which cannot be explained only by resonant effects, generated a motivation for reaching a better understanding of the underlying physics. Another interesting recent result is the Dalitz plane analysis of $B^{+} \to \pi^{+}\pi^{+}\pi^{-}$ performed by LHCb [78, 79], where different but complementary approaches to describe the S-wave were compared: the isobar model, the K-matrix formalism, and a quasi-model-independent approach based on a binning of the phase space. The three methods proved to be consistent with each other showing the robustness of the different descriptions.

III.4 Extraction of the CKM angle γ

The expression of the angle γ in terms of CKM matrix elements does not involve the top quark, meaning that γ can be extracted from tree-level decays, which are unlikely to be affected by new physics [7]. Loop processes can also be used, which yield, in general, a less precise value of γ. On the other hand, these processes are more sensitive to new physics. It is then possible to compare the measurements of γ from "tree-level" and "loop-level" processes as a probe for new physics. It has to be noticed here that, in the case of new physics, the "loop-level" values of γ are likely to be process-dependent and thus they are expected to differ from one decay mode to another. In this part, we will

[7] This statement has to be mitigated: it was shown in Ref. [80] that new physics could affect tree-level contributions to the Wilson coefficients c_1 and c_2 up to 10% without being inconsistent with the current value of γ.

Figure III.6: Feynman diagrams of the favoured $B^- \to D^0 K^-$ decay (left) and the suppressed $B^- \to \overline{D}^0 K^-$ decay (right).



present the most common methods for the "tree-level" extraction and then discuss "loop-level" methods.

III.4.1 Extraction of the CKM angle γ from tree decays

The golden modes for the extraction of γ with "tree-level" decays are $B^\pm \to \overset{(-)}{D^0} K^\pm$, with $\overset{(-)}{D^0}$ decaying to a final state that is accessible by both D^0 and $\overline{D}^0$. The value of γ is then obtained by exploiting the interferences between $b \to c$ and $b \to u$ transitions. The decay of the $B^\pm$ to the D^0, which proceeds via a $b \to c$ transition, is singly Cabibbo suppressed, and thus it is favoured compared to the decay of the $B^\pm$ to the $\overline{D}^0$ that proceeds via a $b \to u$ transition and is suppressed both by colour and the CKM matrix element V_{ub} (see diagrams Fig. III.6). The ratio of the corresponding amplitudes is related to γ by

$$\frac{\mathcal{A}_{b\to u}}{\mathcal{A}_{b\to c}} = r_B e^{i(\delta_B \pm \gamma)}, \tag{III.59}$$

where $r_{\rm B}$ is the ratio of magnitudes, δ_B the strong phase difference between $A_{b\to u}$ and $A_{b\to c}$, and the $+(-)$ sign is associated with the decay of a meson containing a $\overline{b}$ (b) quark. The sensitivity to the relative phase of the two interfering amplitudes is governed by the magnitude of their ratio. For $B^\pm \to \overset{(-)}{D^0} K^\pm$ modes, $r_{\rm B}$ is approximately given by

$$r_{\rm B} = \left|\frac{\mathcal{A}(B^- \to \overline{D}^0 K^-)}{\mathcal{A}(B^- \to D^0 K^-)}\right| \approx \frac{1}{3}\left|\frac{V_{cs}V_{ub}^*}{V_{us}V_{cb}^*}\right| \approx 0.1, \tag{III.60}$$

where the factor 1/3 is due to colour suppression. Other similar modes, such as $B^\pm \to \overset{(-)}{D^0}\pi^\pm$, can be used to extract γ but with a reduced sensitivity to the phases due to a smaller value of $r_{\rm B}$.

Recalling that that the D^0 and $\overline{D}^0$ have access to the same final states f^D and $\overline{f}^D$, two different paths exist to reach the final state of the B meson. In the

case of the $B^- \to \overset{(-)}{D^0} K^-$ decay mode, these paths are $B^- \to \overline{D}^0(\to f^D)K^-$ and $B^- \to D^0(\to f^D)K^-$ (similarly for the *CP*-conjugate decays), so that the amplitudes can be written as

$$\begin{aligned}\mathcal{A}(B^- \to f^D K^-) &= \mathcal{A}(B^- \to D^0 K^-)\mathcal{A}(D^0 \to f^D) + \mathcal{A}(B^- \to \overline{D}^0 K^-)\mathcal{A}(\overline{D}^0 \to f^D),\\ \mathcal{A}(B^+ \to \overline{f}^D K^+) &= \mathcal{A}(B^+ \to \overline{D}^0 K^+)\mathcal{A}(\overline{D}^0 \to \overline{f}^D) + \mathcal{A}(B^+ \to D^0 K^+)\mathcal{A}(D^0 \to \overline{f}^D).\end{aligned} \tag{III.61}$$

Similarly to Eq. (III.59), a ratio of magnitudes and a strong phase can be defined for the D-meson decay: $\mathcal{A}(D^0 \to f^D)/\mathcal{A}(\overline{D}^0 \to f^D) = r_{\rm D} e^{\delta_D}$ (neglecting *CP* violation in the D sector). The decay rates can thus be obtained by expressing the amplitudes in Eq. (III.61) in terms of γ and the hadronic parameters. They write

$$\begin{aligned}\Gamma(B^- \to f^D K^-) &\propto r_{\rm B}^2 + r_{\rm D}^2 + 2 r_{\rm B} r_{\rm D} \cos(\delta_B - \gamma - \delta_D),\\ \Gamma(B^+ \to \overline{f}^D K^+) &\propto r_{\rm B}^2 + r_{\rm D}^2 + 2 r_{\rm B} r_{\rm D} \cos(\delta_B + \gamma - \delta_D),\end{aligned} \tag{III.62}$$

where the normalisation factors have been dropped given the fact that they disappear when computing the ratios of amplitudes. The following asymmetry can then be obtained:

$$A = \frac{\Gamma(B^- \to f^D K^-) - \Gamma(B^+ \to \overline{f}^D K^+)}{\Gamma(B^- \to f^D K^-) + \Gamma(B^+ \to \overline{f}^D K^+)}. \tag{III.63}$$

Depending on the final state of the D-meson decay, different methods are employed to extract the value of γ. The most commonly used are the Gronau-London-Wyler (GLW) method [81, 82] and the Atwood-Dunietz-Soni (ADS) method [83]. The GLW method uses D-meson decays to a *CP* eigenstate so that $f^D = \overline{f}^D = f^D_{\rm CP}$. In the particular case of a decay to a *CP* eigenstate, $r_{\rm D} = 1$. If $f^D_{\rm CP}$ is *CP* even (eg. K^+K^-) then $\delta_D = 0$, and if $f^D_{\rm CP}$ is *CP* odd (eg. $K^0_{\rm S}\pi^0$) then $\delta_D = \pi$. Thus, no external knowledge of the charm parameters is needed. Using the decay rates obtained in Eq. (III.62) the asymmetry writes

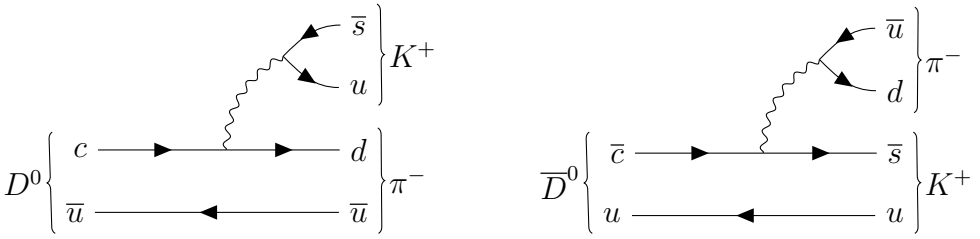
$$A_{\rm GLW} = \frac{\eta_{CP} 2 r_{\rm B} \sin\delta_B \sin\gamma}{1 + r_{\rm B}^2 \eta_{CP} 2 \cos\delta_B \cos\gamma}, \tag{III.64}$$

where η_{CP} is the *CP* eigenvalue. Along with $A_{\rm GLW}$ another observable can be obtained

$$R_{\rm GLW} = 1 + r_{\rm B}^2 \eta_{CP} 2 \cos\delta_B \cos\gamma. \tag{III.65}$$

Thus there are two observables, $\mathcal{A}_{\rm GLW}$ and $R_{\rm GLW}$, and three theoretical parameters, $r_{\rm B}$, δ_B and γ, which means that multiple solutions are expected.

Figure III.7: Example of final states that can be used in the ADS model. The doubly Cabbibo-suppressed decay of $D^0 \to K^+\pi^-$ (left) and the Cabbibo-allowed decay $\overline{D}^0 \to K^+\pi^-$.



The ADS method considers decays of the D-meson to flavour specific final states. Contrary to what happens in the GLW method, the hadronic parameters coming from the D-meson decay do not cancel out and thus input from the charm sector is needed. Like in the previous case, two observables can be derived:

$$\begin{aligned} \mathcal{A}_{\mathrm{ADS}} &= \frac{2r_{\mathrm{B}}r_{\mathrm{D}}\sin(\delta_B+\delta_D)\sin\gamma}{r_{\mathrm{B}}^2+r_{\mathrm{D}}^2+2r_{\mathrm{B}}r_{\mathrm{D}}\cos(\delta_B+\delta_D)\cos\gamma}, \\ R_{\mathrm{ADS}} &= r_{\mathrm{B}}^2+r_{\mathrm{D}}^2+2r_{\mathrm{B}}r_{\mathrm{D}}\cos(\delta_B+\delta_D)\cos\gamma. \end{aligned} \tag{III.66}$$

Comparing to GLW, the ADS method has an apparent complication due to the increase in the number of theoretical parameters. This difficulty is balanced by the fact that the ADS method allows to chose a final state of the D-meson decay that "compensates" the imbalance between the favoured and suppressed B-meson decay amplitudes. For example, using the $K\pi$ final state of the D-meson (see diagrams on Fig. III.7), the favoured $B^- \to D^0K^-$ is combined with the doubly Cabbibo-suppressed $D^0 \to K^+\pi^-$ and the suppressed $B^- \to \overline{D}^0K^-$ is combined with the favoured $\overline{D}^0 \to K^+\pi^-$, leading to a comparable size of the two amplitudes, and thus to an enhancement of CP violation effects.

Theoretically, the hadronic parameters can be obtained from the data and thus the value of γ can be completely determined from tree-level decays. On top of that, the theoretical uncertainty is of the order $\delta\gamma/\gamma \approx 10^{-7}$ [84], which is negligible compared to the current experimental precision. Given these considerations, from the theory point of view γ can be known to a very high precision.

However, on the experimental side, the measurement of γ is complicated due to the small decay rates of the modes that enter in the measurement, with typical branching fractions of the order 10^{-7}. Moreover, the size of the

interference effect is also small; the interference effects are expected to be of about 10%, as shown in Eq. (III.60). The experimental strategy adopted by LHCb is to measure γ in many decay modes[8], even in channels where the precision is not expected to be very good, and combine the results together. This strategy is challenging: many of the decay modes included in this combination contain neutral particles, such as K^0_S and π^0, which are very hard to reconstruct in LHCb. Another challenge comes from the statistical treatment, that is to say how to combine together, in a meaningful way, observables that come from different decay modes. This combination of results not only allows to obtain a very precise value for γ but also to extract the strong parameters like r_B and δ_B. The ADS and GLW methods described previously are not the only methods used to obtain the individual values of γ combined by LHCb. In particular, methods based on a binning of the Dalitz plane are also used, such as the GGSZ method [85, 86]. The value of γ is also extracted using time-dependent decay rates [87–89]. The latest LHCb γ combination [90] uses a likelihood function that is built from the product of probability density functions of 98 experimental observables, and 40 parameters are left free in the fit. The value of γ is found to be $\gamma = (74.0^{+5.0}_{-5.8})^\circ$, which dominated the current world average.

III.4.2 Extraction of the CKM angle γ from loop decays

In this part we essentially describe the theoretical framework of the method used to extract the CKM angle γ in chapter IV. At the end of the section other methods of extraction of γ from loop decays will be shortly mentioned.

In the previous paragraph we gave the bases for the precise extraction of γ with tree-level dominated decays. In the present approach, the aim is not to get the most precise value of γ but rather to exploit several decay modes that receive an important contribution from loop diagrams. The angle γ is extracted from charmless three-body decays of B mesons to three pseudo scalar particles, $B \to PPP$, where P is a kaon or a pion. There are 32 possible final states for these decays: 16 channels are available for $\overline{b} \to \overline{d}$ transitions and 16 for $\overline{b} \to \overline{s}$ transitions. Thus there are many possibilities to combine these channels into subsets to extract γ.

The amplitudes of the different decay modes can be expressed in terms of diagrams using an effective Hamiltonian. These diagrams are defined so that they absorb the CKM matrix elements. In the case of two-body decays there

[8] In principle, a very large number of decay modes are eligible: it is typically the number of possible $B \to D^0 h$ decay modes times the number of $\overset{(-)}{D^0} \to X$ decay modes, where X is a final state accessible by both D^0 and $\overline{D}^0$.

are nine diagrams: tree (T), colour-suppressed tree (C), gluonic penguins (P_{uc}, P_{tc}), electroweak penguin ($P_{\rm EW}$), colour-suppressed electroweak penguin ($P^C_{\rm EW}$), annihilation amplitude (A), exchange amplitude (E), penguin annihilation amplitude (PA). The three annihilation-type diagrams, A, E and PA involve the interaction with the spectator quark and are expected to be significantly smaller than the others, and are thus neglected. A similar treatment can be adopted for three-body decays. More diagrams have to be taken into account, given the fact that there are two possibilities to pop a pair of quarks out from the vacuum. The notation convention used here is the following: if the quark pair is between two final-state quarks that do not include the spectator quark, the diagram contains the subscript "1", conversely, if the spectator quark is involved, the diagram has the subscript "2". Moreover, for $\overline{b} \to \overline{s}$ transitions the diagrams are written with a prime. As an illustration, the different diagrams contributing to $B \to \pi\pi\pi$ are shown in Fig. III.8. Another noticeable difference with the two-body case is that in three-body, only the total momentum of the final state is constrained by momentum conservation, but not the momenta of the individual particles, therefore, the diagrams are momentum dependent.

A further complication arises when dealing with three-body decays because final states such as $K^0_{\rm S}\pi^+\pi^-$ are not *CP* eigenstates. Indeed, in the two-body case, $B^0 \to \pi^+\pi^-$, the relative angular momentum between the two pions is known: it is $l = 0$ so that the final state is *CP* even. In three-body decays such as $B^0 \to K^0_{\rm S}\pi^+\pi^-$, the angular momentum between the pions can be either *CP* even or *CP* odd. A way to experimentally disentangle these *CP* states is to use Dalitz-plot analyses and symmetrise the amplitude according to the symmetry of the $\pi\pi$ final state. For example, $B^0 \to K^0_{\rm S}\pi^+\pi^-$ is *CP* even if the $\pi^+\pi^-$ pair is symmetric, so that the corresponding amplitude is $\frac{1}{\sqrt{2}}\Big(\mathcal{A}(s_{K^0_{\rm S}\pi^+}, s_{K^0_{\rm S}\pi^-}) + \mathcal{A}(s_{K^0_{\rm S}\pi^-}, s_{K^0_{\rm S}\pi^+})\Big)$. Given all this, it is possible to write the amplitudes of a specific class of decays — $B \to \pi\pi\pi$, $B \to K\pi\pi$, $B \to KK\pi$, $B \to K\overline{K}K$ — in terms of common diagrams, by separating the cases where the isospin state is odd or even. By doing so, the different amplitudes contain a number of momentum-dependent strong parameters proportional to the number of diagrams and one momentum-independent weak phase, which is γ. The experimental observables are the branching fractions, and the direct and indirect *CP* asymmetries. If the number of observables is larger than the number of theoretical parameters, γ can be extracted by a fit without the need of external parameters input.

It was shown in Ref. [91] that this condition can be fulfilled for some classes of decays under the assumption of flavour SU(3) symmetry. Under flavour SU(3), all the particles are identical and thus the amplitudes and

the observables need to be symmetrised. With three final-state particles, the symmetry group is S_3 (permutation group), and then there are six possible symmetrisations: fully-symmetric (FS) — the amplitude is fully-symmetric under the exchange of two final state particles —, fully-antisymmetric — the amplitude is fully-antisymmetric (AS) under the exchange of two final state particles — and four mixed states. In the FS state, the *CP* odd and *CP* even relations hold and, in addition, there are relations between the electroweak penguin diagrams and the tree diagrams. This was first shown in two-body decays [92, 93] and then in three-body decays [94]. They write

$$\begin{aligned} P'_{EW1,2} &= \kappa T'_{1,2}, \\ P'^{C}_{EW1,2} &= \kappa C'_{1,2}, \end{aligned} \tag{III.67}$$

where κ is a real coefficient that depends on CKM matrix elements (as defined in Eq. (III.2)) and Wilson coefficients,

$$\kappa \equiv -\frac{3}{2}\frac{|\lambda_t^{(s)}|}{|\lambda_u^{(s)}|}\frac{c_9 + c_{10}}{c_1 + c_2}. \tag{III.68}$$

This relation depends on two assumptions: that flavour SU(3) symmetry holds and that $c_1/c_2 = c_9/c_{10}$. In $\bar{b} \to \bar{s}$ decays, the electroweak penguin and tree decays are suppressed with respect to P'_{tc}, which is the dominant contribution. Therefore the error relative to flavour SU(3) breaking is subdominant and the theoretical error on κ is mainly due to the approximate ratios of Wilson coefficients, which are supposed to hold to about 5%. Thanks to these relations, the expressions of the amplitudes simplify and the number of theoretical parameters is reduced, making the extraction of γ feasible. The computation of the expressions for the 32 $B \to PPP$ amplitudes in the fully-symmetric state, including rescattering effects to all orders in α_s, has been carried out in Ref. [95].

In principle, γ can be extracted using the $B \to K\pi\pi$ decay modes [94]. There are six such modes: $B^+ \to K^+\pi^+\pi^-$, $B^+ \to K^+\pi^0\pi^0$, $B^0 \to K^+\pi^0\pi^-$, $B^0 \to K^0\pi^+\pi^+$, and $B^0 \to K^0\pi^0\pi^0$. The drawback of this method is that, experimentally, it is complicated to reconstruct the π^0 mesons, while two of the six modes have two neutral pions in the final state. A way out is to combine several classes of decays, and more particularly $B \to K\pi\pi$ and $B \to K\overline{K}K$. In $B \to K\overline{K}K$, the pair of quarks that is popped from the vacuum is an $s\bar{s}$ pair, instead of $u\bar{u}$ or $d\bar{d}$ for $B \to K\pi\pi$, which means that the diagrams are, in principle, different for the two cases. However, under flavour SU(3) symmetry, they are basically the same and the relations between tree and electroweak

penguin diagrams hold. It is then possible to combine together $B \to K\pi\pi$ and $B \to K\overline{K}K$ decay modes and avoid the states involving two π^0 mesons. The amplitudes for the modes of interest, under flavour SU(3), are then given by

$$\begin{aligned}
2\mathcal{A}(B^0 \to K^+\pi^0\pi^-)_{sym} &= T'_1 e^{i\gamma} + C'_2 e^{i\gamma} - \kappa\,(T'_2 + C'_1)\,,\\
\sqrt{2}\mathcal{A}(B^0 \to K^0\pi^+\pi^-)_{sym} &= -T'_1 e^{i\gamma} - C'_1 e^{i\gamma} - \tilde{P}'_{uc} e^{i\gamma} + \tilde{P}'_{tc} + \ \kappa\left(\frac{1}{3}T'_1 + \frac{2}{3}C'_1 - \frac{1}{3}C'_2\right),\\
\sqrt{2}\mathcal{A}(B^+ \to K^+\pi^+\pi^-)_{sym} &= -T'_2 e^{i\gamma} - C'_1 e^{i\gamma} - \tilde{P}'_{uc} e^{i\gamma} + \tilde{P}'_{tc} + \ \kappa\left(\frac{1}{3}T'_1 - \frac{1}{3}C'_1 + \frac{2}{3}C'_2\right),\\
\sqrt{2}\mathcal{A}(B^0 \to K^+K^0K^-)_{sym} &= -T'_2 e^{i\gamma} - C'_1 e^{i\gamma} - \tilde{P}'_{uc} e^{i\gamma} + \tilde{P}'_{tc} + \ \kappa\left(\frac{1}{3}T'_1 - \frac{1}{3}C'_1 + \frac{2}{3}C'_2\right),\\
\mathcal{A}(B^0 \to K^0K^0\overline{K}^0)_{sym} &= \tilde{P}'_{uc} e^{i\gamma} - \tilde{P}'_{tc} + \ \kappa\left(\frac{2}{3}T'_1 + \frac{1}{3}C'_1 + \frac{1}{3}C'_2\right),
\end{aligned} \tag{III.69}$$

where the subscript "*sym*" denotes that the isospin pair is symmetric under the permutation of the two particles (*CP* even), $\tilde{P} = P_1 + P_2$, and κ is the coefficient that relates trees and electroweak penguin diagrams defined in Eq. (III.68). Note here that $\mathcal{A}(B^+ \to K^+\pi^+\pi^-)_{sym} = \mathcal{A}(B^0 \to K^+K^0K^-)_{sym}$ and thus they are not independent. These equations can be further simplified by defining effective diagrams

$$\begin{aligned}
A &\equiv -\tilde{P}'_{tc} + \kappa\left(\frac{2}{3}T'_1 + \frac{1}{3}C'_1 + \frac{1}{3}C'_2\right),\\
B &\equiv T'_1 + C'_2,\\
C &\equiv T'_2 + C'_1,\\
D &\equiv T'_1 + C'_1.
\end{aligned} \tag{III.70}$$

Hence, the amplitudes can be expressed in terms of five effective diagrams

$$\begin{aligned}
2\mathcal{A}_{\rm fs}(B^0 \to K^+\pi^0\pi^-) &= Be^{i\gamma} - \kappa C,\\
\sqrt{2}\mathcal{A}_{\rm fs}(B^0 \to K^0\pi^+\pi^-) &= -De^{i\gamma} - \tilde{P}'_{\rm uc} e^{i\gamma} - A + \kappa D,\\
\mathcal{A}_{\rm fs}(B^0 \to K^0K^0\overline{K}^0) &= \tilde{P}'_{\rm uc} e^{i\gamma} + A,\\
\sqrt{2}\mathcal{A}_{\rm fs}(B^0 \to K^+K^0K^-) &= -Ce^{i\gamma} - \tilde{P}'_{\rm uc} e^{i\gamma} - A + \kappa B,\\
\sqrt{2}\mathcal{A}_{\rm fs}(B^+ \to K^+\pi^+\pi^-) &= -Ce^{i\gamma} - \tilde{P}'_{\rm uc} e^{i\gamma} - A + \kappa B.
\end{aligned} \tag{III.71}$$

There are then ten theoretical parameters: the five magnitudes and the four relative phases of the effective diagrams, and γ. As mentioned before, there are up to three observables per decay mode: the decay rates and direct *CP*

asymmetries, that are accessible to all the modes ($B^+ \to K^+\pi^+\pi^-$ is not included in the computation of the observables since it is not independent from $B^0 \to K^0K^0\overline{K}^0$) and the indirect CP asymmetries that can be obtained only in modes where the final state is the same for B^0 and $\overline{B}^0$ ($B^0 \to K^0\pi^+\pi^-$, $B^0 \to K^0K^0\overline{K}^0$ and $B^0 \to K^+K^0K^-$). The number of experimental observables is thus eleven. With more observables that theoretical parameters γ can be extracted from a fit.

There are three sources of flavour SU(3) breaking in this approach. Firstly the SU(3) breaking related to the computation of the coefficient κ, defined in Eq.(III.68). This source, as explained previously, is subdominant. Secondly, there is an SU(3) breaking effect coming from the fact that the diagrams in $B \to K\pi\pi$ and $B \to K\overline{K}K$ are considered to be the same while in reality they differ: $B \to K\overline{K}K$ decay modes have a $s\overline{s}$ pair in the final state, while the $B \to K\pi\pi$ modes have a $u\overline{u}$ or $d\overline{d}$ pair. Even if the breaking effect may differ between a pair of quarks popped from the vacuum and a pair of quarks produced via the decay of a virtual state, the difference is assumed to be marginal. Therefore, to a good approximation, the SU(3) breaking is considered to be the same for each diagram. Since there is room for one more theoretical parameter, the SU(3) breaking related to the assumption of identical diagrams can be taken into account by introducing an additional factor, $\alpha_{\mathrm{SU(3)}}$, in the expressions for the amplitudes in Eq. (III.71). Finally, the third source of flavour SU(3) breaking , which can also be included in $\alpha_{\mathrm{SU(3)}}$, is the fact that kaons and pions are assumed to be the same particles while they are obviously not. The different methods to evaluate systematic uncertainties will be discussed in details in chapter IV. The expressions for the amplitudes including $\alpha_{\mathrm{SU(3)}}$ are then given by

$$\begin{aligned}
&2\mathcal{A}_{\mathrm{fs}}(B^0 \to K^+\pi^0\pi^-) = Be^{i\gamma} - \kappa C,\\
&\sqrt{2}\mathcal{A}_{\mathrm{fs}}(B^0 \to K^0\pi^+\pi^-) = -De^{i\gamma} - \tilde{P}'_{\mathrm{uc}}e^{i\gamma} - A + \kappa D,\\
&\mathcal{A}_{\mathrm{fs}}(B^0 \to K^0K^0\overline{K}^0) = \alpha_{\mathrm{SU(3)}}(\tilde{P}'_{\mathrm{uc}}e^{i\gamma} + A),\\
&\sqrt{2}\mathcal{A}_{\mathrm{fs}}(B^0 \to K^+K^0K^-) = \alpha_{\mathrm{SU(3)}}(-Ce^{i\gamma} - \tilde{P}'_{\mathrm{uc}}e^{i\gamma} - A + \kappa B),\\
&\sqrt{2}\mathcal{A}_{\mathrm{fs}}(B^+ \to K^+\pi^+\pi^-) = -Ce^{i\gamma} - \tilde{P}'_{\mathrm{uc}}e^{i\gamma} - A + \kappa B.
\end{aligned} \tag{III.72}$$

Note that now the $B^+ \to K^+\pi^+\pi^-$ and $B^0 \to K^+K^0K^-$ amplitudes differ. The number of experimental observables is then 13 whereas the number of theoretical parameters is 11 so that γ can still be extracted form a fit.

The method described in this section is particularly interesting given the fact that it is potentially sensitive to new physics because of the large contribution from loop-decays in charmless decay modes of the B-meson. The different

parameters are extracted from the data, including the hadronic parameters, which means that no external input (e.g. from QCDF) is needed, and thus the result is model independent[9]. Furthermore, this data-driven approach implies that the accuracy on the determination of the parameters depends on the size of the datasets, which grows with time and allows for more precise results. The experimental procedure to extract γ with this method is developed in chapter IV and carried out using fully-symmetrised amplitudes.

Other methods to extract γ from three-body charmless decays have been proposed, such as Ref. [96], which relies on isospin symmetry and for which a time-integrated Dalitz-plot analysis of $B_s^0 \to K^-\pi^+\pi^0$ decay mode is needed. More recently, a method based on U-spin symmetry [97], considering pairs of B^0 and B_s^0 decays was also designed. Even though these methods are promising, the lack of the necessary experimental results in the B_s^0-meson decay modes prevent from using them at present.

[9]Here we need to mitigate this point. Indeed, the experimental results that are used so far are model dependent (isobar model).

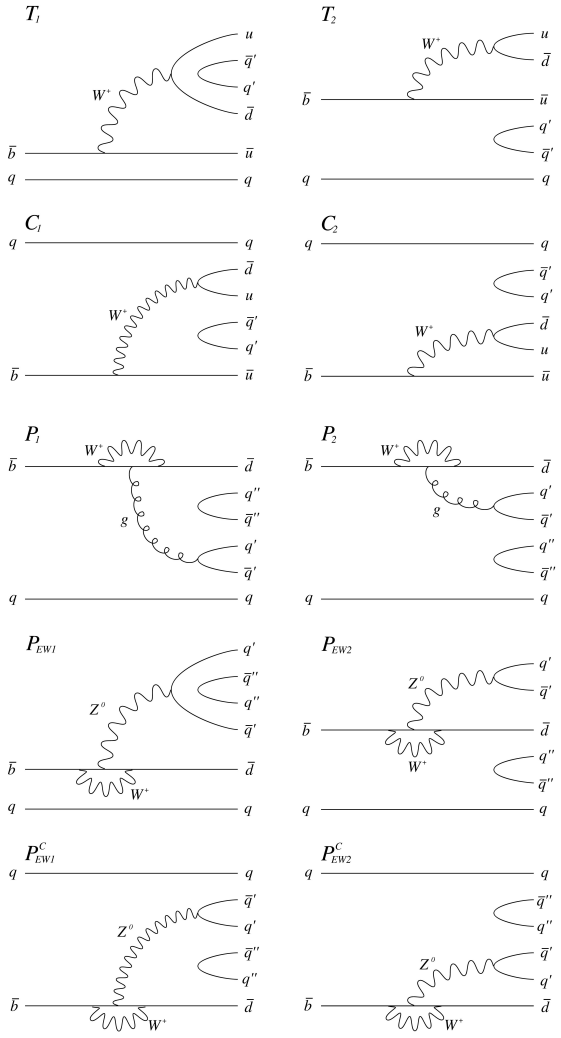


Figure III.8: List of diagrams contributing to $B \to \pi\pi\pi$ decay mode. Annihilation-type diagrams are neglected. [91].

Chapter IV

Extraction of the CKM phase γ using charmless 3-body decays of B mesons

IV.1 Introduction

A powerful test of the Standard Model (SM) consists in measuring the parameters of the unitarity triangle using many methods and decay modes. Any discrepancies from the SM expectations would suggest the presence of new physics (NP). As explained in part III.4, the angle γ can be extracted using processes dominated by tree-level transitions such as $B \to DP$ where B (D) represents a neutral or charged $B_{(s)}$ ($D_{(s)}$) meson and P is a pseudo scalar hadron (typically a kaon or a pion). The theoretical uncertainties associated with such methods are very small and thus it is possible to achieve high precision on the angle value. The world average value, $\gamma = (73.5^{+4.2}_{-5.1})^\circ$, is currently dominated by LHCb.

The angle γ can also be measured in decay modes containing a large contribution from loop diagrams. In this part we study a method to extract γ from three-body charmless $B \to PPP$ decay modes proposed by Bhattacharya, Imbeault and London [98]. The theoretical grounds for this method are described in part III.4. Several $B \to K\pi\pi$ and $B \to KK\bar{K}$ decay modes[1] are used to extract γ:

$$\begin{array}{lll} B^0 \to K^0 K^0 \overline{K}^0 & B^0 \to K^+\pi^0\pi^- & B^+ \to K^+\pi^+\pi^- \\ B^0 \to K^0 K^+ K^- & B^0 \to K^0_S \pi^+\pi^-. & \end{array} \tag{IV.1}$$

All these decay modes involve $\overline{b} \to \overline{s}$ transitions and receive contributions from

[1] Charge conjugate modes are implicitly included this part unless stated otherwise.

both tree and penguin diagrams. Since the dominant contribution to these modes comes from penguin diagrams, this method is particularly sensitive to new physics. The $B \to PPP$ modes given in (IV.1) are combined together under flavour SU(3) symmetry using diagrammatic analysis [91].

The aim of the present study is to provide a complete proof of principle of the method including fully-propagated experimental uncertainties. Full Dalitz-plot amplitude analysis of the different modes is required to obtain the observables. Since these results are not yet available in LHCb we used BaBar analysis results instead. The full study was published in Ref. [99].

We start with a brief review of the method in part IV.2. Part IV.3 describes the practical implementation along with the fitting procedure. Fit results for the baseline model are given in part IV.4. The systematic uncertainties are discussed in part IV.5. A key point of the method is to make sure that the effect of SU(3) breaking on the analysis is controlled. For that purpose, two different tests of SU(3)-breaking have been designed and are presented in part IV.6. Finally, the results of the analysis are summarised in part IV.7. Part IV.8 presents some theoretical and experimental perspectives that could be worth to explore in the future.

IV.2 Description of the method

As described in section III.4.2 the amplitudes of three-body charmless $B \to PPP$ can be written in terms of momentum-dependent diagrams. One can fix the symmetry of the final state in $B \to P_1P_2P_3$ by using its Dalitz plane [91]. To extract the value of γ with a fit, the number of theoretical parameters must be less than the number of observables. This condition is automatically fulfilled in the case of two-body decays. However, when using three-body decays, one needs to make further assumptions to reduce the number of parameters. In that order, flavour SU(3) symmetry is assumed.

We define the three Mandelstam variables $s_i \equiv (p_j + p_k)^2$, where p_i is the momentum of P_i, and $ijk = 123$, 231 or 312. These obey to $s_1 + s_2 + s_3 = m_B^2 + m_1^2 + m_2^2 + m_3^2$. The decay amplitude $\mathcal{A}(s_1, s_2)$ of a 3-body decay varies as a function of the position in the Dalitz plane and can be reconstructed experimentally via different methods, cf. part III.3.2. Under flavour SU(3) symmetry a kaon is the same as a pion and thus all the final state particles are identical. It is therefore necessary to symmetrise the amplitude. For three identical particles, six symmetrisation states are possible: one fully symmetric, one anti symmetric and four mixed states. In this study we chose to use amplitudes that are fully symmetric under permutations of the final-state

particles. The fully-symmetric amplitude is then given by

$$\begin{aligned}\mathcal{A}_{\rm fs}(s_1,s_2) = \frac{1}{\sqrt{6}}(&\mathcal{A}(s_1,s_2)+\mathcal{A}(s_2,s_1)+\mathcal{A}(s_1,s_3)\\ &+\mathcal{A}(s_3,s_1)+\mathcal{A}(s_3,s_2)+\mathcal{A}(s_2,s_3)).\end{aligned} \tag{IV.2}$$

The effect of the symmetrisation on the Dalitz plane and choice of points will be discussed later, in section IV.3.5. Under flavour SU(3) there are relations between the electroweak penguin and tree diagrams for $b \to s$ transitions [94]. For the fully-symmetric final state, these take the form

$$P'_{EWi} = \kappa T'_i \ , \quad P'^{C}_{EWi} = \kappa C'_i \ \ (i=1,2) \ \ ; \qquad \kappa \equiv -\frac{3}{2}\frac{|\lambda_t^{(s)}|}{|\lambda_u^{(s)}|}\frac{c_9+c_{10}}{c_1+c_2} \ , \tag{IV.3}$$

where the c_i are Wilson coefficients and $\lambda_p^{(s)} \equiv V_{pb}^* V_{ps}$ (the V_{ij} are elements of the CKM matrix). The amplitude T'_i are associated with colour-favoured tree diagrams while the C'_i represent the contributions of colour-suppressed ones. The proportionality coefficient κ is approximately equal to 0.5.

Obviously, flavour SU(3) is not an exact symmetry as, for example, the quarks u, d and s have rather different masses. When extracting γ one has to account for flavour SU(3)-breaking. This can be done by introducing SU(3) breaking parameters into the expressions of the amplitudes. Ideally, we need one SU(3) breaking parameter per diagram. However, this would result in a number of parameters that is too large to perform a fit. Consequently, it is assumed that the size of SU(3) breaking is the same for all diagrams and is small. Then, one single parameter, referred as $\alpha_{\rm SU(3)}$ in the following, is enough to account for flavour SU(3) breaking. This parameter relates $B \to K\pi\pi$ and $B \to KK\bar{K}$ decays and is equal to unity in the flavour SU(3) limit. As mentioned previously, in the case of three-body decays, the diagrams are momentum dependent and so is the parameter $\alpha_{\rm SU(3)}$. This means that the effect of SU(3)-breaking can vary in different regions of the Dalitz plane. The assumption that is made here is that the average of the variations on the Dalitz plane is close to one. More precisely, variations from unity are expected in localised regions of the Dalitz plane but they are supposed to compensate when averaging on many points in Dalitz plane. This assumption is shown to be valid in part IV.6.

This analysis makes use of the five $B \to PPP$ decays modes listed in Eq. (IV.1) to extract γ. As explained in part III.4.2, the fully-symmetric amplitudes of the five decay modes, $\mathcal{A}_{\rm fs}$, can be expressed as linear combinations of momentum-dependent effective diagrams, represented here by complex

parameters A, B, C, D and $\tilde{P}'_{uc}$:

$$\begin{aligned}
&2\mathcal{A}_{\rm fs}(B^0 \to K^+\pi^0\pi^-) = Be^{i\gamma} - \kappa C,\\
&\sqrt{2}\mathcal{A}_{\rm fs}(B^0 \to K^0\pi^+\pi^-) = -De^{i\gamma} - \tilde{P}'_{\rm uc}e^{i\gamma} - A + \kappa D,\\
&\mathcal{A}_{\rm fs}(B^0 \to K^0K^0\overline{K}^0) = \alpha_{\rm SU(3)}(\tilde{P}'_{\rm uc}e^{i\gamma} + A),\\
&\sqrt{2}\mathcal{A}_{\rm fs}(B^0 \to K^+K^0K^-) = \alpha_{\rm SU(3)}(-Ce^{i\gamma} - \tilde{P}'_{\rm uc}e^{i\gamma} - A + \kappa B),\\
&\sqrt{2}\mathcal{A}_{\rm fs}(B^+ \to K^+\pi^+\pi^-) = -Ce^{i\gamma} - \tilde{P}'_{\rm uc}e^{i\gamma} - A + \kappa B.
\end{aligned} \tag{IV.4}$$

where γ, the CKM phase, is the parameter of interest, $\alpha_{\rm SU(3)}$ is the flavour SU(3)-breaking parameter and κ is a constant defined in equation (IV.3). The effective diagrams relate to the diagrams defined in [100] through the equations given in (III.70).

As already mentioned in part III.4.2, if SU(3)-breaking is not taken into account, i.e. $\alpha_{\rm SU(3)} = 1$, then $\mathcal{A}_{\rm fs}(B^0 \to K^+K^0K^-) = \mathcal{A}_{\rm fs}(B^+ \to K^+\pi^+\pi^-)$. Hence $B^+ \to K^+\pi^+\pi^-$ decay mode can be dropped from the analysis. This way of extracting γ with four modes will be used as the baseline for the analysis. The reasons for this choice are related to stability and fit convergence.

From the amplitudes measured experimentally on the Dalitz plane, one can form a set of three linearly-independent observables per decay mode:

$$\begin{aligned}
X(s_1, s_2) &= |\mathcal{A}_{\rm fs}(s_1, s_2)|^2 + |\overline{\mathcal{A}}_{\rm fs}(s_1, s_2)|^2,\\
Y(s_1, s_2) &= |\mathcal{A}_{\rm fs}(s_1, s_2)|^2 - |\overline{\mathcal{A}}_{\rm fs}(s_1, s_2)|^2,\\
Z(s_1, s_2) &= {\rm Im}[\mathcal{A}^*_{\rm fs}(s_1, s_2)\overline{\mathcal{A}}_{\rm fs}(s_1, s_2)],
\end{aligned} \tag{IV.5}$$

where $\overline{\mathcal{A}}_{\rm fs}$ denotes the fully-symmetric amplitude of the conjugate mode and $\mathcal{A}^*_{\rm fs}$ is the complex conjugate. The observables X, Y, and Z are related to the effective *CP*-averaged branching fraction, the direct *CP* asymmetry, and the indirect *CP* asymmetry, respectively. For a given decay mode, their values depend on the position in the Dalitz plane. The observable Z has no physical meaning for flavour-specific final states and for charged B-mesons since meson mixing requires a neutral B-meson and the same final state for both B^0 and $\overline{B}^0$. Therefore, Z is not computed for $B^0 \to K^+\pi^0\pi^-$ and $B^0 \to K^0_{\rm S}\pi^+\pi^-$. Note that, to have access to indirect *CP* asymmetry experimentally, one needs to perform a time-dependent flavour tagged Dalitz plot.

In this study, we take as experimental inputs the amplitude models obtained by BaBar [2] in Refs. [71, 101–104]. The BaBar analysis of $B^0 \to K^0_{\rm S}K^0_{\rm S}K^0_{\rm S}$ [104] was time-integrated and *CP*-averaged. Since no distinction was made between B^0 and $\overline{B}^0$, then $\mathcal{A} = \overline{\mathcal{A}}$; the observables Y and

[2] Note that experimentally K^0 is observed as $K^0_{\rm S}$

Z vanish and only the observable X is accessible for this mode. As noted in Ref. [98], this implies a simplification in the expression of amplitudes compared with Eq. (IV.4). To be specific, the requirement that $Y = Z = 0$ implies that $\tilde{P}'_{\rm uc} = 0$, so that Eq. (IV.4) becomes

$$\begin{aligned}
&2\mathcal{A}_{\rm fs}(B^0 \to K^+\pi^0\pi^-) = Be^{i\gamma} - \kappa C,\\
&\sqrt{2}\mathcal{A}_{\rm fs}(B^0 \to K^0\pi^+\pi^-) = -De^{i\gamma} - A + \kappa D,\\
&\mathcal{A}_{\rm fs}(B^0 \to K^0K^0\overline{K}^0) = \alpha_{\rm SU(3)}A,\\
&\sqrt{2}\mathcal{A}_{\rm fs}(B^0 \to K^+K^0K^-) = \alpha_{\rm SU(3)}(-Ce^{i\gamma} - A + \kappa B),\\
&\sqrt{2}\mathcal{A}_{\rm fs}(B^+ \to K^+\pi^+\pi^-) = -Ce^{i\gamma} - A + \kappa B.
\end{aligned} \tag{IV.6}$$

Since for each mode the observables X, Y, Z defined in Eq. (IV.5) depend upon the fully-symmetric amplitude, and $\mathcal{A}_{\rm fs}$ is related to the theory parameters by Eqs. (IV.6), the observables may be written as functions of the theoretical parameters. Expressing them in terms of magnitudes and strong phases ($U = ue^{i\phi_u}$ for $U = A, B, C, D$), and setting $\phi_a = 0$ without loss of generality, the following relations are obtained,

$$\begin{aligned}
&X^{th}_{K^+\pi^+\pi^-}(s_1,s_2) = a^2 + (\kappa b)^2 + c^2 + 2ac\cos\phi_c\cos\gamma - 2\kappa ab\cos\phi_b - 2\kappa bc\cos(\phi_b-\phi_c)\cos\gamma\ ,\\
&Y^{th}_{K^+\pi^+\pi^-}(s_1,s_2) = -2\left(ac\sin\phi_c + \kappa bc\sin(\phi_b-\phi_c)\right)\sin\gamma\ ,\\
&X^{th}_{K^0_{\rm S}K^+K^-}(s_1,s_2) = (\alpha_{\rm SU(3)})^2 X^{th}_{K^+\pi^+\pi^-}\ ,\\
&Y^{th}_{K^0_{\rm S}K^+K^-}(s_1,s_2) = (\alpha_{\rm SU(3)})^2 Y^{th}_{K^+\pi^+\pi^-}\ ,\\
&Z^{th}_{K^0_{\rm S}K^+K^-}(s_1,s_2) = (\alpha_{\rm SU(3)})^2\left(-c^2\cos\gamma - ac\cos\phi_c + \kappa bc\cos(\phi_b-\phi_c)\right)\sin\gamma\ ,\\
&X^{th}_{K^0_{\rm S}\pi^+\pi^-}(s_1,s_2) = a^2 + (\kappa d)^2 + d^2 + 2ad\cos\phi_d\cos\gamma - 2\kappa ad\cos\phi_d - 2\kappa d^2\cos\gamma\ ,\\
&Y^{th}_{K^0_{\rm S}\pi^+\pi^-}(s_1,s_2) = -2ad\sin\phi_d\sin\gamma\ ,\\
&Z^{th}_{K^0_{\rm S}\pi^+\pi^-}(s_1,s_2) = \left(-d^2\cos\gamma - ad\cos\phi_d + \kappa d^2\right)\sin\gamma\ ,\\
&X^{th}_{K^+\pi^+\pi^0}(s_1,s_2) = \frac{1}{2}\left(b^2 + \kappa^2c^2 - 2\kappa bc\cos\gamma\cos(\phi_b-\phi_c)\right)\ ,\\
&Y^{th}_{K^+\pi^+\pi^0}(s_1,s_2) = \kappa bc\sin\gamma\sin(\phi_b-\phi_c)\ ,\\
&X^{th}_{K^0_{\rm S}K^0_{\rm S}K^0_{\rm S}}(s_1,s_2) = 2(\alpha_{\rm SU(3)})^2a^2\ .
\end{aligned} \tag{IV.7}$$

This set of equations is used to perform a fit to the theoretical parameters and extract γ. If γ is extracted at a single point (s_1, s_2) on the Dalitz plane, there are nine real, unknown parameters: four magnitudes (a, b, c, d), three strong phases (ϕ_b, ϕ_c, ϕ_d), γ, and $\alpha_{\rm SU(3)}$. From the experimental inputs, there

are eleven observables: three (X, Y, Z) for each of the modes $K^0_S K^+ K^-$ and $K^0_S \pi^+ \pi^-$, two (X, Y) for each of the modes $K^+ \pi^+ \pi^-$ and $K^+ \pi^+ \pi^0$, and one (X) for $K^0_S K^0_S K^0_S$. If $\alpha_{\mathrm{SU(3)}}$ is fixed to unity, there are instead eight unknown parameters and nine observables. In both cases, there are more observables than theory parameters, and γ may be extracted with a fit. Instead of using one point, it is possible, and preferable, to use several points. By doing so, the number of observables and theoretical parameters is increased. For N points on the Dalitz plane, one obtains $11N$ observables and $8N + 1$ unknown theoretical parameters when $\alpha_{\mathrm{SU(3)}}$ is allowed to vary. If $\alpha_{\mathrm{SU(3)}}$ is fixed to unity, there is $9N$ observables and $7N + 1$ unknowns. In both cases, for any (positive) number of points on the Dalitz plane, the number observables exceeds the number of unknowns, allowing for γ to be extracted.

IV.3 Practical implementation of the method

IV.3.1 Implementation of the decay modes

The BABAR analyses use the isobar formalism to parametrise the decay amplitude at each point on the Dalitz plane as a coherent sum over n resonant and non-resonant components

$$\mathcal{A}(s_1, s_2) = \sum_{j=1}^{n} c_j F_j(s_1, s_2) \tag{IV.8}$$

where the c_j's are the complex isobar coefficients, the F_j's are the functions that describe the strong interaction. A complete description of this formalism is given in part III.3.3.

The isobar coefficients and the definitions of the lineshapes given in BABAR's papers are used to reconstruct the amplitude of each mode as a function of the position in the Dalitz plane. This is achieved thanks to the LAURA^{++} software package [67], which is a Dalitz plot fitter. With this package, it is possible to obtain the variation of the amplitude over the Dalitz plane given a model, in our case, the isobar models described in the papers. Due different choices of conventions in the normalisation of the lineshapes between BABAR analyses and LAURA^{++} some adjustments have been made. The details of the implementation and the cross-checks can be found in appendix A.1.

The total amplitude on the Dalitz plane is defined up to a global constant. A global normalisation factor has to be introduced in order to be able to compare the amplitudes of the different decay modes together. For 3 different

particles the normalisation writes

$$\mathcal{N} = \sqrt{\frac{2}{\tau_B} \frac{\mathcal{B}}{\iint (|A(s_1,s_2)|^2 + |\overline{A}(s_1,s_2)|^2) \mathrm{ds}_1 \mathrm{ds}_2}}, \tag{IV.9}$$

where $\mathcal{B}$ is the branching fraction of the considered mode and τ_B is the lifetime of the B^0 or $B^\pm$ meson. In the case of 3 identical particles ($B^0 \to K^0_S K^0_S K^0_S$) we have

$$\mathcal{N}_{3K_s} = \sqrt{\frac{8}{9\tau_{B^0}} \frac{\mathcal{B}}{\iint |A(s_1,s_2)|^2 \mathrm{ds}_1 \mathrm{ds}_2}}. \tag{IV.10}$$

IV.3.2 Error propagation

The errors on the isobar parameters, taken from BaBar's papers are propagated to the observables. For each mode the derivatives of the observables with respect to the isobar parameters are computed analytically. It is then possible to compute the matrix of these derivatives at any desired point on the Dalitz plane:

$$G(s_1,s_2) = \begin{pmatrix} \frac{\partial X_{mode\,1}}{\partial a_1}(s_1,s_2) & \dots & \frac{\partial X_{mode\,1}}{\partial a_n}(s_1,s_2) \\ \frac{\partial Y_{mode\,1}}{\partial a_1}(s_1,s_2) & \dots & \frac{\partial Y_{mode\,1}}{\partial a_n}(s_1,s_2) \\ \frac{\partial Z_{mode\,1}}{\partial a_1}(s_1,s_2) & \dots & \frac{\partial Z_{mode\,1}}{\partial a_n}(s_1,s_2) \\ \vdots & \ddots & \end{pmatrix}. \tag{IV.11}$$

From the covariance matrix of the isobar parameters provided by BaBar , V_{a_j}, and the derivatives of the observables with respect to isobar parameters (IV.11), we can compute the covariance matrix of the observables, $V_{XYZ}(s_1,s_2)$, at any given point on the Dalitz plane:

$$V_{XYZ}(s_1,s_2) = G^T(s_1,s_2)\ V_{a_j}\ G(s_1,s_2), \tag{IV.12}$$

where G^T is the transpose of the matrix of derivatives defined in Eq. (IV.11). In the specific case of $B^0 \to K^0_S K^0_S K^0_S$ decay mode, no correlation matrix was quoted in the BaBar paper, and therefore the correlations are neglected. No correlations are expected between the different decay modes, so that V_{XYZ} is a block-diagonal matrix that contains the correlations between the observables

within each mode. This computation can be generalised to a combination of two points A and B as

$$V_{XYZ}\left((s_1,s_2)_A;(s_1,s_2)_B\right) = G^T\left((s_1,s_2)_A;(s_1,s_2)_B\right)\ V_{a_j}\ G\left((s_1,s_2)_A;(s_1,s_2)_B\right). \tag{IV.13}$$

From this equation we see that the correlations between observables belonging to the same decay mode and evaluated at different points on the Dalitz plane are also taken into account. It is worth noticing here that the size of the covariance matrix varies with the number of observables. The more Dalitz-plot points are added, the larger the number of observables. Such that, for N points, with $\alpha_{\mathrm{SU(3)}}$ fixed to unity, the covariance matrix sizes $9N \times 9N$ and if $\alpha_{\mathrm{SU(3)}}$ is released then the size is $11N \times 11N$.

IV.3.3 Fitting procedure

From the implementation described in the previous section, the observables and their corresponding covariance matrix can be obtained at any point in the Dalitz plane, and the corresponding observables can be computed using Eq. (IV.5). It is therefore possible to compare the observables obtained from experimental inputs to the ones obtained given a specific set of theoretical parameters $\{a,\ b,\ c,\ d,\ \phi_b,\ \phi_c,\ \phi_d,\ \alpha_{\mathrm{SU(3)}}\}$ by means of a χ^2 function.

A scan in γ is then performed: γ is fixed to consecutive values (with a step of 1°), and for each fixed value of γ the other free parameters are evaluated by minimising the χ^2 function. Doing so, we obtain a profile of the χ^2 as a function of γ. In principle, for a global minimisation, the final values of the fit parameters and the χ^2 should not depend on the initial values of the parameters. However, in the case of multidimensional fits, some dependency is observed due to the presence of secondary local minima. To obtain a robust estimate of the global minimum, the minimisation is repeated 500 times randomising the initial value of the parameters. The smallest value of χ^2 is retained for each fixed value of γ. Then, the preferred value(s) for γ is (are) the minimum (minima) of this profile. The asymmetric statistical uncertainty is then estimated as the change in γ required to produce a change of one unit in χ^2 from the minimum.

The minimisation is performed using `MINUIT2` minimisation package [105, 106]. All the non converging fits are rejected as well as the fits for those the internal correlation matrix of the minimiser has been forced to be positive definite.

Examples of γ-scans can be seen on figure IV.1. They are obtained following the procedure described here, using different random combinations of three

points in the Dalitz plane. In this example, SU(3) is fixed to one so that flavour SU(3)-breaking effects are not taken into account. Multiple solutions from γ appear on these plots. This feature is not entirely surprising given the trigonometric nature of Eq. (IV.6). It is also manifest that the scans are different from each other while exhibiting common characteristics: there are roughly two groups of three solutions separated by a "forbidden region" between 100° and 200°.

IV.3.4 Algorithm to extract the minima from a scan

The χ^2 functions obtained in the scans are not analytical so that a specific algorithm has been developed to extract numerically the minima in the different scans. This algorithm has been designed using an important number of scans with specific characteristics. The different parameters as well as the fitting function have been chosen in an empirical way by comparing the results of different options between many scans. The final algorithm follows these main steps:

1. Start at the first point.
2. Define the current window to be the range of γ spanned by the current point plus the next 19 consecutive points. Fit those 20 points with a 3^{rd}-order polynomial function.
3. Determine the minimum of the fitted polynomial (at $x = \gamma_{\text{min}}$, $y = \chi^2(\gamma_{\text{min}})$)
4. Reject the minimum $(\gamma_{\text{min}}, \chi^2(\gamma_{\text{min}}))$ if any of the following is true:
   - The value of γ_{min} is outside the window.
   - $\chi^2(\gamma_{\text{min}}) > 7$.
   - The polynomial fit is of poor quality (its fit χ^2 is greater than 5).
5. Move along one point, then go back to step 2 (unless the points have been exhausted).

Usually, when a minimum is identified, since the consecutive windows overlap, it will be present in several consecutive polynomial fits (steps 2–4). Due to statistical fluctuations, the value of γ_{min} will differ slightly between these; the average value is taken.

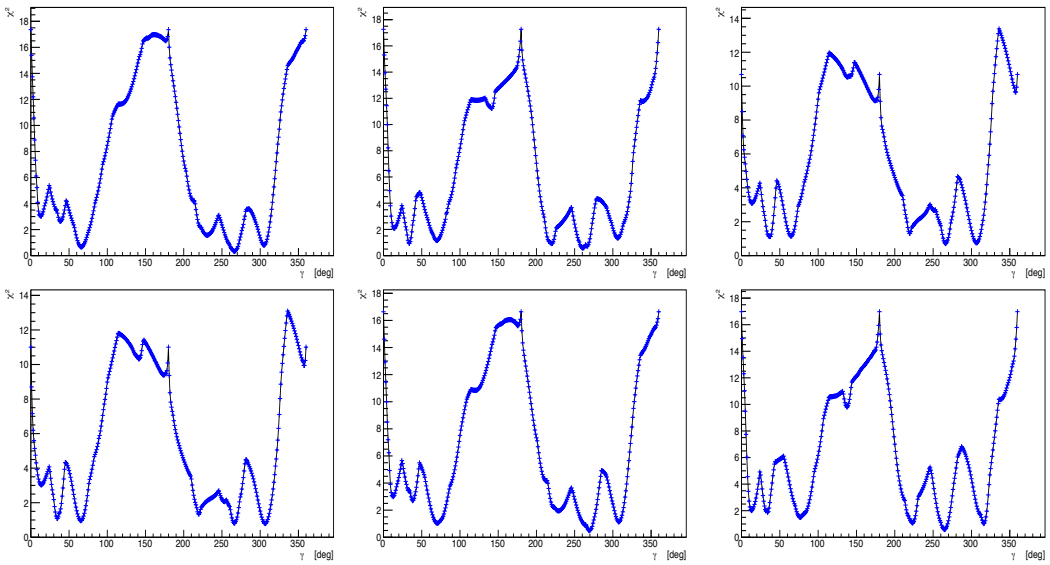


Figure IV.1: Examples of γ scans obtained taking random combinations of 3 points on the DP using 4 modes and fixing $\alpha_{\mathrm{SU(3)}}$ to unity.

IV.3.5 Choice of points on the Dalitz plane

As stated in pat IV.2, fully-symmetric versions of the amplitudes are used in this study. The fully-symmetric state is obtained by summing over the permutations of the 4-momenta of the final-state particles. In our case, this is achieved by permuting the invariant square masses of the final state particles. For reading convenience we recall Eq. (IV.2) here,

$$\begin{aligned}\mathcal{A}_{\mathrm{fs}}(s_1, s_2) = \frac{1}{\sqrt{6}}(&\mathcal{A}(s_1, s_2) + \mathcal{A}(s_2, s_1) + \mathcal{A}(s_1, s_3) \\ &+\mathcal{A}(s_3, s_1) + \mathcal{A}(s_3, s_2) + \mathcal{A}(s_2, s_3)).\end{aligned} \tag{IV.14}$$

It is appears from this equation that the symmetrised amplitude $\mathcal{A}_{\mathrm{fs}}$ has a sixfold symmetry in the Dalitz plane. In effect, the plane can be divided into six regions along the lines of zero cosine helicity as illustrated on figure IV.2.

The structure and information in each region is identical to the others[3]. It is therefore sufficient to consider points in one sixth of the symmetrised Dalitz plane.

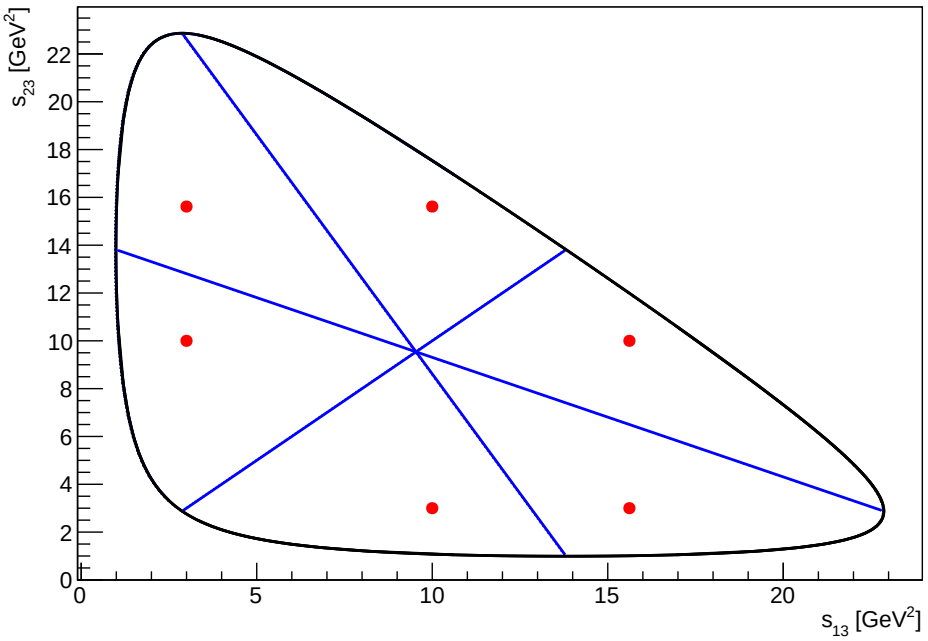


Figure IV.2: Symmetrisation example of $K^0_S K^0_S K^0_S$ Dalitz plane. The dots are the symmetric points corresponding to point (10,3), there are 6 points in total because of the 6 possible permutations of the masses. The blue lines correspond to $\cos\theta_H = 0$ where θ_H is the helicity angle, those lines divide the DP into 6 regions which are symmetric. Each of these regions contains the same information as the others.

Considering the difference between the masses of kaons and pions it is evident that the phase space of the disintegration for $B \to K\pi\pi$ decay modes is larger than for $B \to KK\bar{K}$ modes. Thus the boundaries of the five Dalitz planes do not coincide, as illustrated on figure IV.3. The points used to extract γ must be present in all the decay modes. We are thus limited to chose points within the boundaries of the smallest Dalitz plane which is $B^0 \to K^0_S K^0_S K^0_S$. A consequence to this is that the information contained on the boundaries of the $B \to K\pi\pi$ modes cannot be used to extract γ.

In principle, it is possible to extract γ using only one point on the Dalitz

[3]Up to flavour SU(3)-breaking effects.

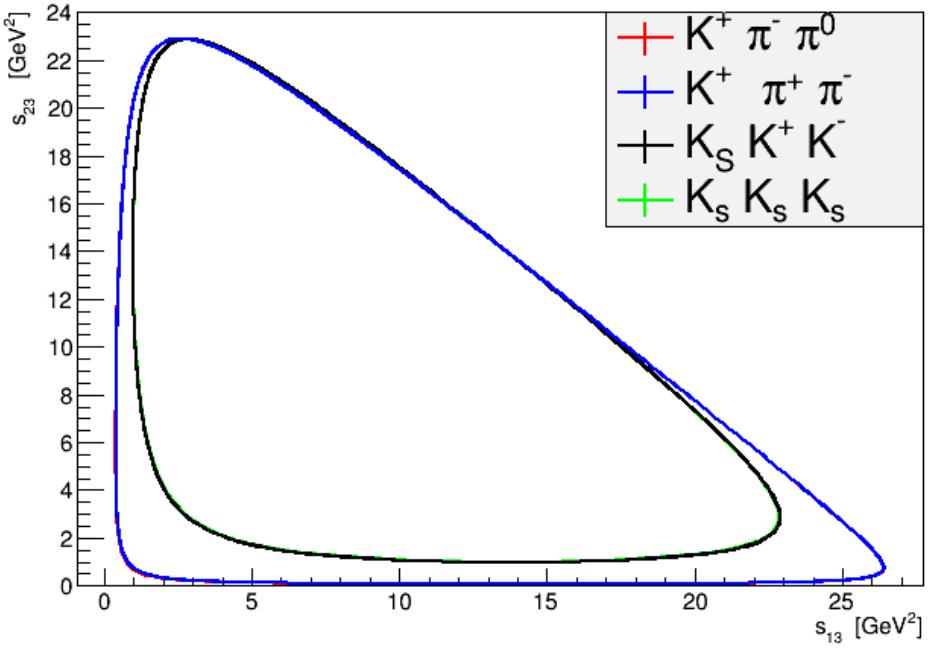


Figure IV.3: Kinematic limits of the different modes. The difference of mass between charged and neutral kaons is small so $K^0_S K^+ K^-$ and $K^0_S K^0_S K^0_S$ Dalitz plane are almost the same size. The same applies for $K^0_S \pi^+ \pi^-$ and $K^+ \pi^+ \pi^0$ Dalitz plane.

plane. Yet, the breaking of flavour SU(3) cannot be controlled when considering single points on the Dalitz plane. However, as shown in part IV.6, the effects of flavour SU(3)-breaking are small when averaging over a large number of points. For that reason, as well as to use maximum amount of information it is better to extract γ with the largest possible number of points. A priory, an arbitrary large number of points can be used with this method since the number of observables, which goes as $11N$, is always larger than the number of theoretical parameters, which goes as $8N + 1$. This condition is also fulfilled for $\alpha_{\mathrm{SU(3)}} = 1$, in that case we have $9N > 7N + 1$. In practice, some points can be very highly correlated, especially if they receive a large contribution coming from the same resonance. Those high correlations have an impact on the covariance matrix with becomes approximately singular and not invertible. This impose limitations to the choice of points: the number of points that can be used simultaneously in a fit is finite and small and all the combinations of points are not possible due to the correlations. Empirically, the maximum number of points that can be simultaneously used is found to be three. In order

to avoid experimenter's bias in the choice of points, hundreds of combinations of points are chosen randomly in the Dalitz plane with a requirement on the maximum correlations allowed in the correlation matrix. For the baseline fit, 501 three-points combinations have been made and the maximum of accepted correlations is set to 70%.

For each three-points set, a scan on γ is performed and the minima are extracted as explained previously in section IV.3.3. The final result is obtained by averaging on the central values and the statistical uncertainties extracted from the individual scans. Note that, fluctuations aside, the average uncertainty does not decrease as more scans are added.

IV.4 Baseline results

As mentioned before, the baseline result is obtained by fixing $\alpha_{\mathrm{SU(3)}}$ to one in the fit, i.e. the flavour SU(3)-breaking effects are ignored. The SU(3)-breaking effects are then taken into account as a systematic effect.

The results are obtained following the procedure described in the previous section: a total of 501 sets of three points are chosen randomly in the Dalitz plane with a requirement on the maximum allowed correlations of 70%. For each combination, a scan on γ is performed and the minima are extracted along with their asymmetric statistical uncertainties.

Since the points are randomly scattered across the Dalitz plane, each combination of points contains different information. As a result the χ^2 profiles obtained with the different combinations are similar but not identical: the central values and statistical uncertainties fluctuates, and, in some cases not all of the six minima are present. The distribution of the central values is shown in Figure IV.4 and the rates at which the minima are found are given in Table I. It appears that all the minima have a rate greater than 91%, which means that less than 10% of the scans exhibit missing minima.

In some instances, due to the shape of the χ^2 scan, a minimum can be identified but it is not is not deep enough to extract the statistical uncertainties, it is then referred as "poorly resolved". In practice, a minimum is considered poorly resolved when at least one of the statistical uncertainty crosses with the central value of the next minimum; i.e. for a given minimum i, $(\mu_i + \sigma_{R_i}) > \mu_{i+1}$ or $(\mu_i - \sigma_{L_i}) < \mu_{i-1}$ or both. Since no statistical uncertainties can be determined for these minima they are not included in the average from where the overall result is obtained. The potential bias on the result due to the exclusion of these minima is taken into account as a systematic effect and is described in section IV.5.

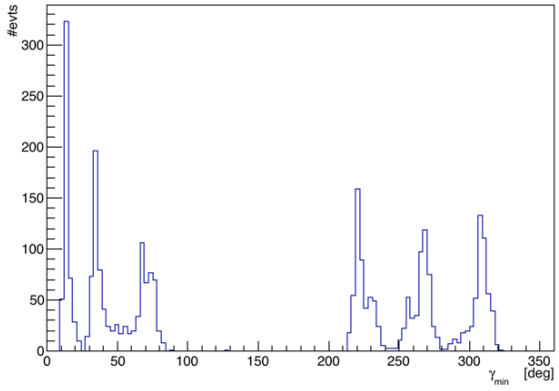


Figure IV.4: The minima found with four decay modes ($\alpha_{\mathrm{SU(3)}} = 1$). For each of the 501 sets of random combinations of three points in the Dalitz plot, a χ^2 scan for γ is performed and the minima γ_{min} are found. The histogram shows the accumulation of the minima across all 501 scans.

The averaged central value, μ, and asymmetric statistical uncertainties (σ_L, σ_R) along with the χ^2 are given in table II. Six distinct minima are found and the third one is compatible with the world average value $\gamma = (73.5^{+4.2}_{-5.1})^\circ$ [38]. The statistical uncertainties are between 4.3° and 10.9°. These results are compatible with the ones obtained in Ref. [98], where a preliminary implementation of the method was carried out without taking into account the correlations and systematic uncertainties. From the χ^2 values reported in Table II, the first minimum (12.9°), and, to a less extent, the fourth one (223.2°) are disfavoured, while the fifth one (266.4°) is favoured, which is as well in agreement with Ref. [98].

IV.5 Systematic uncertainties

The experimental statistical and systematic uncertainties on the amplitude models used as inputs are already included in the "statistical uncertainties" given in Table II, through the use of the correlation matrices and uncertainties provided in BaBar's papers. Two additional systematic uncertainties coming from the method are considered here. The first one is related to the poorly resolved minima described in the previous section and the second one concerns SU(3) breaking. The results are summarised in Table III.

As discussed previously, in some cases an extracted minimum is not well

Table I: The rates at which the different minima are obtained with four decay modes ($\alpha_{\mathrm{SU(3)}} = 1$). A total of 501 scans are used.

	Count	Fraction (%)
Minimum 1	484	96.6
Minimum 2	474	94.6
Minimum 3	461	92.0
Minimum 4	499	99.6
Minimum 5	487	97.2
Minimum 6	488	97.4

Table II: The minima found with four decay modes ($\alpha_{\mathrm{SU(3)}} = 1$). For each minimum, the central value for γ is given (μ), along with the asymmetric experimental uncertainty on the left- and right-hand sides (σ_L, σ_R) and the corresponding χ^2.

	μ	σ_L	σ_R	χ^2
Minimum 1	12.9°	4.3°	8.4°	3.61
Minimum 2	36.6°	6.1°	6.6°	1.99
Minimum 3	68.9°	8.6°	8.6°	2.07
Minimum 4	223.2°	7.5°	10.9°	2.15
Minimum 5	266.4°	10.8°	9.2°	1.40
Minimum 6	307.5°	8.1°	6.9°	1.74

separated from another nearby minimum. This means that the algorithm described in section IV.3.4 is not able to compute the statistical uncertainty on the corresponding central value. As explained in section IV.4, these poorly resolved minima are excluded from the final average. Disregarding these minima could bias the result. To estimate this effect, the central value is computed again, this time including the poorly resolved minima in the average. The systematic uncertainty is then obtained as

$$\sigma_{\text{poorly resolved}} = |\mu - \mu^{\text{all}}|, \tag{IV.15}$$

where μ is the central value obtained including only well-resolved minima in the average, and μ^{all} is the central value obtained when both well-resolved and not-well-resolved minima are included. The values obtained for each minimum

are given in Table III and are below 1.5°.

The baseline results are obtained using four out of the five modes by fixing $\alpha_{\mathrm{SU(3)}}$ to one in the fit. Physically, this means that flavour SU(3) is assumed to be exact and no breaking is taken into account. As explained in section IV.2 it is not possible to assess SU(3) breaking in a general way. Nevertheless, the $\alpha_{\mathrm{SU(3)}}$ parameter can be used to get an idea of the scale of the breaking. To that end, the analysis is repeated using the five modes and allowing for $\alpha_{\mathrm{SU(3)}}$ to vary in the fit. The exact same procedure is followed to obtain the results. The detailed results are given later in section IV.5.1. The systematic uncertainty related to flavour SU(3)-breaking is then assessed as

$$\sigma_{\mathrm{SU(3)}} = |\mu - \mu^{5\,\mathrm{modes}}|. \tag{IV.16}$$

where μ is the central value obtained with the baseline procedure and $\mu^{5\,\mathrm{modes}}$ is the central value obtained with the five-modes extraction. The results are summarised in Table III and are below 3°. More tests of the validity of flavour SU(3) symmetry hypothesis are presented in section IV.6.

The dominant systematic uncertainty comes from flavour SU(3) but its value is definitely smaller than the statistical uncertainties obtained in section IV.4, which means that this analysis is dominated by the statistical error.

Table III: Summary of the systematic uncertainties.

	Poorly resolved minima	Flavour SU(3)-breaking
Minimum 1	0.8°	1.0°
Minimum 2	0.3°	2.6°
Minimum 3	0.2°	2.4°
Minimum 4	0.7°	0.7°
Minimum 5	1.4°	1.3°
Minimum 6	0.7°	0.9°

IV.5.1 Results allowing for SU(3) breaking

To assess the systematic effect related to the fixed value of $\alpha_{\mathrm{SU(3)}}$ in the baseline fit, the analysis is repeated, this time using the five decay modes and allowing for $\alpha_{\mathrm{SU(3)}}$ to vary in the fit. The sensible increase in size of the covariance matrix — its size is now 33×33 whereas in the baseline case it is 27×27 —

complicates the procedure for finding points that can be used simultaneously. For this reason the rejection criterion on the correlations between sets of points was relaxed from 70% to 80%, and the number of random sets of three points was reduced from 501 to 401.

Compared to the baseline extraction, the fit behaviour is less stable with a convergence rate of about 80% (versus 100% in the baseline). The frequency with which the minima appear among the 401 scans is also reduced, especially for the two first minima, as shown in Table IV. The reduced stability is taken to be due to the increased number of free parameters, and the consequent increase in the size of the covariance matrix. The histogram of the distribution of the central values across the scans is shown in Figure IV.5.

The results averaged over the different scans and excluding the poorly resolved minima, are given in Table V. This table includes the the central values (μ), asymmetric experimental uncertainties (σ_L, σ_R), the χ^2, the recomputed systematic uncertainty due to poorly resolved minima ($|\mu - \mu^{\text{all}}|$), and the systematic uncertainty associated to SU(3) breaking. The results for the minima are compatible within the uncertainties with the ones obtained with the baseline procedure. The χ^2 values are globally larger than in the previous case, which can be explained by the increase in the size of the parameter space, but the ordering of the minima with respect to their χ^2 value is the same.

Table IV: The rates at which the different minima are obtained with five decay modes. A total of 401 scans are used.

	Count	Fraction (%)
minimum 1	306	76.3
minimum 2	329	82.0
minimum 3	372	92.3
minimum 4	383	95.5
minimum 5	378	94.3
minimum 6	391	97.5

IV.6 Studies of SU(3) breaking

In this approach, flavour SU(3) symmetry is assumed to hold when averaging on large number of points of the Dalitz plane. Fluctuations are foreseen in localised regions of the Dalitz plane — i.e. it is expected to observe flavour SU(3)-breaking when considering single points in the Dalitz plane — but

Table V: The minima found with five decay modes, allowing $\alpha_{\mathrm{SU(3)}}$ to vary in the fit. For each minimum, the central value for γ is given (μ), along with the asymmetric experimental uncertainty on the left- and right-hand sides (σ_L, σ_R) and the related χ^2. The quantities $|\mu - \mu^{\mathrm{all}}|$ and $|\mu^{4\,\mathrm{modes}} - \mu^{5\,\mathrm{modes}}|$ are taken as estimates of the systematic uncertainties due to poorly resolved minima and flavour SU(3) breaking, respectively.

	μ	σ_L	σ_R	χ^2	$\lvert\mu - \mu^{all}\rvert$	$\lvert\mu^{4\,\mathrm{modes}} - \mu^{5\,\mathrm{modes}}\rvert$
Minimum 1	11.9°	5.8°	9.1°	3.53	1.3	1.0
Minimum 2	39.2°	6.3°	6.7°	2.50	1.2	2.6
Minimum 3	71.3°	9.5°	9.3°	2.58	0.4	2.4
Minimum 4	223.9°	7.4°	9.5°	2.92	0.1	0.7
Minimum 5	265.0°	11.0°	10.0°	2.19	1.2	1.3
Minimum 6	308.4°	8.8°	7.0°	2.49	0.6	0.9

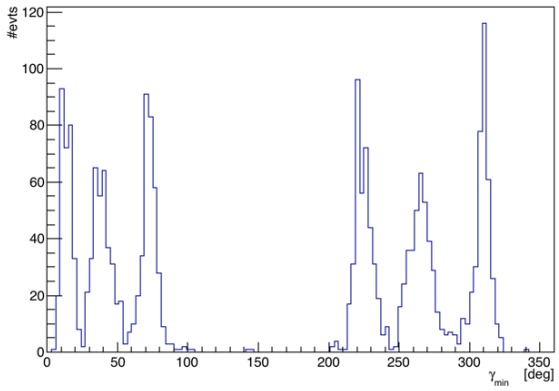


Figure IV.5: The minima found with five decay modes, with $\alpha_{\mathrm{SU(3)}}$ free to vary in the fit. For each of the 401 sets of random combinations of three points in the Dalitz plot, a χ^2 scan for γ is performed and the minima γ_{min} are found. The histogram shows the accumulation of the minima across all 401 scans.

these effects are assumed to partly cancel when averaging on many points. The results obtained in part IV.5.1 show that this hypothesis seems to hold. In this part, we present two further way to test this assumption and specifically to check that $\alpha_{\mathrm{SU(3)}}$ averages to one when considering many points. The first test is based on the theoretical expressions for the amplitudes given in Eq. (IV.4)

and involves comparing the amplitudes of two modes related by flavour SU(3) as a function of the position in the Dalitz plane. The second test consists in extracting the value of $\alpha_{\mathrm{SU(3)}}$ over the Dalitz plane from fits.

IV.6.1 Comparison of the amplitudes of $B^0 \to K_S K^+ K^-$ and $B^+ \to K^+\pi^+\pi^-$

From inspection of the last two lines of Eq. (IV.4), there is a linear relationship between the fully symmetric amplitudes for $B^0 \to K^0_S K^+ K^-$ and $B^+ \to K^+\pi^+\pi^-$:

$$\mathcal{A}_{\mathrm{fs}}(B^0 \to K^+ K^0 K^-) = \alpha_{\mathrm{SU(3)}}\, \mathcal{A}_{\mathrm{fs}}(B^+ \to K^+\pi^+\pi^-)\ . \tag{IV.17}$$

The value of the parameter $\alpha_{\mathrm{SU(3)}}$ can be inferred by comparing the values of the amplitudes of these two modes at different points on the Dalitz plane [95]. To obtain an estimate of $\alpha_{\mathrm{SU(3)}}$ at one point on the Dalitz plane, we define the following ratio:

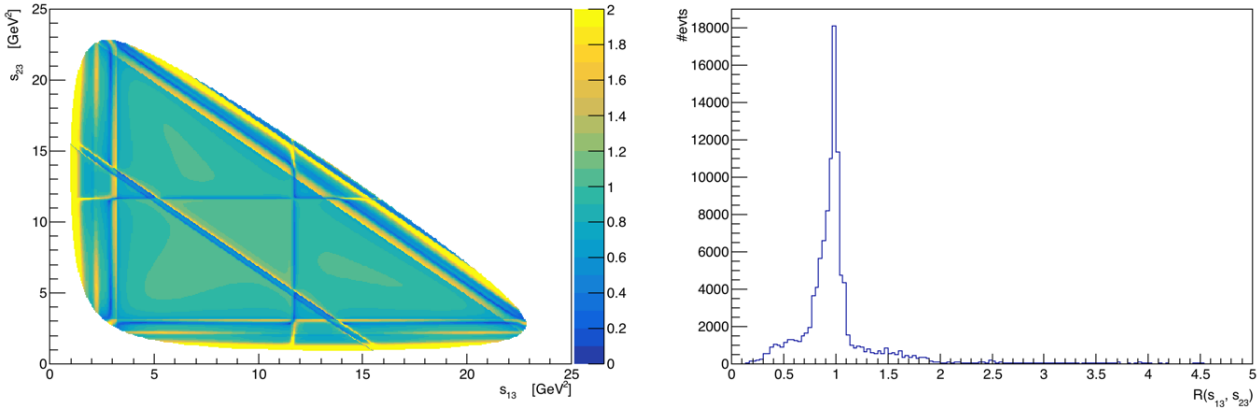
$$R(s_{13}, s_{23}) = \left| \frac{\mathcal{A}_{\mathrm{fs}}(B^+ \to K^+\pi^+\pi^-; s_{13}, s_{23}) + \mathcal{A}_{\mathrm{fs}}(B^- \to K^-\pi^-\pi^+; s_{13}, s_{23})}{\mathcal{A}_{\mathrm{fs}}(B^0 \to K^+ K^0_S K^-; s_{13}, s_{23}) + \mathcal{A}_{\mathrm{fs}}(\overline{B}^0 \to K^- K^0_S K^+; s_{13}, s_{23})} \right|, \tag{IV.18}$$

where $\mathcal{A}_{\mathrm{fs}}(X; s_{13}, s_{23})$ is the symmetrised amplitude for the decay mode X measured at point (s_{13}, s_{23}).

The ratio $R(s_{13}, s_{23})$ is computed at each point of a grid of more than 1M points uniformly distributed on the Dalitz plane. The variation of $R(s_{13}, s_{23})$ as a function of the position in the Dalitz plane is shown on Figure IV.6 (a). Significant deviations from one are observed, especially in regions containing resonant contributions, which is unsurprising, given that flavour SU(3) is broken by the mass difference between s and u, d quarks. A histogram of the values of R is shown in Fig. IV.6 (b). The distribution peaks near one, and the average value is 1.028, rather close to unity. These two plots show that despite being strongly violated locally, flavour SU(3) holds reasonably well when averaging across the phase space.

IV.6.2 Fitted value of $\alpha_{\mathrm{SU(3)}}$ over the Dalitz Plane

Another manner to study flavour SU(3) breaking is by determining $\alpha_{\mathrm{SU(3)}}$ from a fit. While it is difficult to interpret the values of $\alpha_{\mathrm{SU(3)}}$ extracted along with γ from the procedure described in part IV.5.1 — because of the use of combinations of three correlated points, it is not straight forward to understand



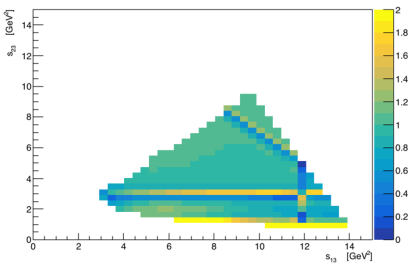
(a) (b)

Figure IV.6: (a) Ratio of amplitudes $R(s_{13}, s_{23})$ over the whole fully symmetrised DP. Note that the $R(s_{13}, s_{23})$ scale is truncated at 2.0. (b) Histogram of the different values of the ratio of amplitudes $R(s_{13}, s_{23})$.

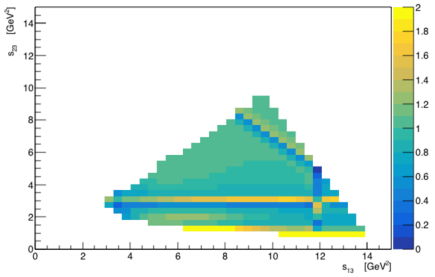
the correlations between the different values of $\alpha_{\mathrm{SU(3)}}$— we can use a similar strategy with single points on the Dalitz plane.

A uniform grid of 386 points chosen in one sixth of the Dalitz plane is exploited. For each single point, a χ^2 minimisation is carried out in a similar fashion as the procedure described in section IV.3.3, repeating the fit 500 times with random initial values of the parameters and allowing for $\alpha_{\mathrm{SU(3)}}$ to vary in the fit. The value of $\alpha_{\mathrm{SU(3)}}$ not only depends on the position on the Dalitz plane but also on γ.

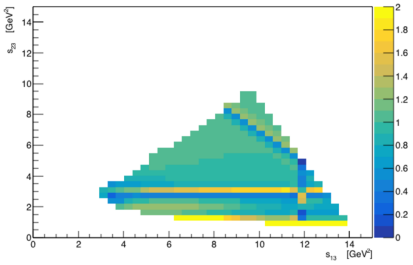
Figure IV.7 shows the fitted values of $\alpha_{\mathrm{SU(3)}}$ at each point of the grid for fixed values of γ that correspond approximately to the six minima given in Table II. The mean value of $\alpha_{\mathrm{SU(3)}}$, $\langle \alpha_{SU(3)} \rangle$ is also given in each case. The variation of $\alpha_{\mathrm{SU(3)}}$ between the six values of γ are negligible and the mean values are very close together and to unity. The observed variation of $\alpha_{\mathrm{SU(3)}}$ in Figure IV.7 is similar to the variation of $R(s_{13}, s_{23})$ in Figure IV.6 (a), the same pattern of large flavour SU(3) breaking is seen near resonances. The main conclusions form this test are the same as in the previous on and enforce the hypothesis of flavour SU(3) symmetry nearly conserved when averaging on many points on the Dalitz plane.



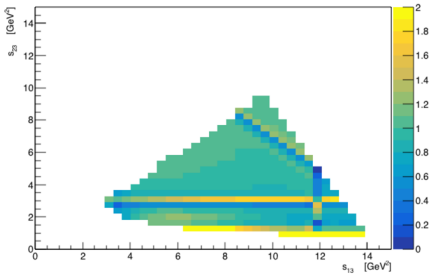
(a) $\gamma = 12°$, $\langle \alpha_{SU(3)} \rangle = 1.06$



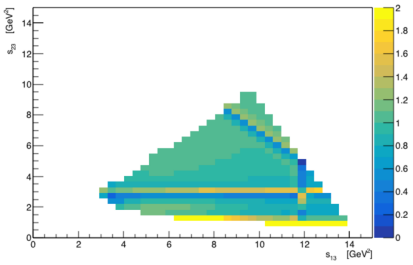
(b) $\gamma = 37°$, $\langle \alpha_{SU(3)} \rangle = 1.06$



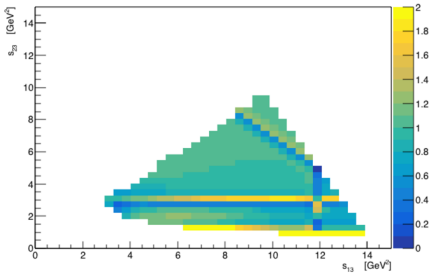
(c) $\gamma = 68°$, $\langle \alpha_{SU(3)} \rangle = 1.05$



(d) $\gamma = 223°$, $\langle \alpha_{SU(3)} \rangle = 1.06$



(e) $\gamma = 266°$, $\langle \alpha_{SU(3)} \rangle = 1.05$



(f) $\gamma = 307°$, $\langle \alpha_{SU(3)} \rangle = 1.05$

Figure IV.7: Plot of the fitted values of $\alpha_{SU(3)}$ across the DP for γ fixed to the values of the 6 solutions obtained in part IV.4. For each plot, the average value $\langle \alpha_{SU(3)} \rangle$ is also given.

IV.7 Summary and conclusion

The method of extracting the weak phase γ from three-body charmless decays of the B meson developed by Bhattacharya, Imbeault and London [98] is applied to amplitude models of five charmless three-body decays of B mesons obtained by the BABAR collaboration [71, 101–104]. Six solutions for γ are

found:

$$\begin{aligned}
\gamma_1 &= [\ 12.9\,^{+8.4}_{-4.3}\ \ (\text{stat}) \pm 1.3\ (\text{syst})]^\circ,\\
\gamma_2 &= [\ 36.6\,^{+6.6}_{-6.1}\ \ (\text{stat}) \pm 2.6\ (\text{syst})]^\circ,\\
\gamma_3 &= [\ 68.9\,^{+8.6}_{-8.6}\ \ (\text{stat}) \pm 2.4\ (\text{syst})]^\circ,\\
\gamma_4 &= [223.2\,^{+10.9}_{-7.5}\ (\text{stat}) \pm 1.0\ (\text{syst})]^\circ,\\
\gamma_5 &= [266.4\,^{+9.2}_{-10.8}\ (\text{stat}) \pm 1.9\ (\text{syst})]^\circ,\\
\gamma_6 &= [307.5\,^{+6.9}_{-8.1}\ \ (\text{stat}) \pm 1.1\ (\text{syst})]^\circ\ .
\end{aligned}$$

The six values obtained are well separated, and one is compatible with the Standard Model while the others are not. The presence of multiple solutions may reflect trigonometric ambiguities. The central values and statistical uncertainties are obtained assuming flavour SU(3) symmetry while the the systematic uncertainties include SU(3)-breaking effect and the impact of poorly resolved minima on the result. The dominant uncertainty is statistical and is below 11°, which is approximately twice the uncertainty on the world average. It is obtained by propagating the statistical and systematic uncertainties on the isobar parameters, determined by the BaBar analyses of the different modes, to the observables of the method; so that all the correlations are included.

The hypothesis of flavour SU(3) symmetry has been tested to a large extent, using different approaches. Besides extracting γ with an SU(3)-breaking parameter left free in the fit, two more tests were performed in order to asses the variation of the breaking across the Dalitz plane. As expected, strong local variations are seen, especially near resonances, but, when considering the average of the variations over the Dalitz plane they are found to be consistent within the hypothesis of flavour SU(3) symmetry within a few percents.

IV.8 Perspectives

The present study shows that information on the weak phase γ can be extracted from charmless three-body decays with a good precision. This was carried out using results from the BaBar collaboration that used the isobar model as a description of the amplitude variation across the Dalitz plane. Time-dependent Dalitz plot analyses of the same modes using the data collected at LHCb during the two runs of data taking are on their way and, given the increase of statistics, should provide a better description of the Dalitz plot. Experimental results from the Belle II experiment are also expected and will be complementary to the LHCb ones. Indeed, the charged modes, e.g. $B \to KK\bar{K}$, are a priori more suited to LHCb while the neutral ones, e.g. $B^0 \to K^0_S K^0_S K^0_S$, are better adapted

to Belle II. Given this, an interesting possibility would be a simultaneous fit of the physics parameters to datasets of both experiments using a framework such as JFIT [107]. It is worth stressing that, even if the isobar model was used for this study, it is not required by the method itself. Other descriptions of the amplitude can be used. As described in part III.3.4, other models are available to parametrise the Dalitz plane and some Dalitz-plot analyses of charmless B-decays in LHCb are investigating these alternative models [72, 78, 79].

As mentioned previously, since the kinematic boundaries of the Dalitz plot of the different modes do not coincide, we are limited to chose the points within the boundaries of $B \to K^0_S K^0_S K^0_S$ and a part of the information contained in the $B \to K^0_S \pi\pi$ is lost. Some tentatives of reparametrising the Dalitz plane of the different modes in a way such as the kinematical boundaries coincide have been made. An attempt using two cosines of the helicity angles, $\cos\theta_H$, as Dalitz-plot variables instead of the invariant masses of the daughters particles has been carried out. The key point is that by doing so, every Dalitz plane is contained within a square between -1 and 1, and so, in principle, all the information available in the different decay modes can be used. Changing the set of variables is not as trivial as it seems. The amplitude symmetrisation has to be expressed in terms of these new variables, and a definition of "coinciding points" between the different modes has to be chosen. Indeed, when reparametrising the Dalitz plane in terms of helicity angles, the resonances are not straight lines anymore and their shapes change quite significantly with the masses of the daughters particles; for example, $\chi_{c0} \to \pi^+\pi^-$ and $\chi_{c0} \to K^+K^-$ will have rather different shapes, so that a point lying on the χ_{c0} resonance in the $B^0 \to K^0_S \pi^+\pi^-$ mode may not be located on the χ_{c0} resonance of $B^0 \to K^0_S K^+K^-$ when using the same coordinates. This is actually a flavour SU(3)-breaking effect that is already present when using the standard Dalitz-plane coordinates but it becomes more striking when working with the helicity angles. Other complications arise due to the fact that, unlike the invariant square masses, the third value of the cosine of helicity angle cannot be deduced from the knowledge of the other two using a simple formula. Furthermore, the $\cos\theta_H$ can be expressed as functions of the invariant square masses but the inverse requires numerical computation. The procedure that was used here, was to take $B^0 \to K^0_S K^0_S K^0_S$, the only "real" flavour SU(3)-symmetric mode, as a benchmark to try to workout a meaningful symmetrisation. Unfortunately, the uncertainty on γ obtained with this method was too large to resolve all the minima. More investigations in this direction are needed and can result in alternative parametrisations that could be used to replace the baseline model or to assess systematic uncertainties.

Only fully-symmetric states have been used so far. A consequence of this symmetrisation is that, by construction, the vector resonances disappear.

The exploration of the other symmetrisations (fully-antisymmetric and mixed states) would add information and probably reduce the statistical uncertainties. Combining different symmetrisations together may resolve a part of the trigonometrical ambiguities (or maybe all of them), and therefore determine if the value of γ obtained from this method agrees or not with the tree-level value. The use of other symmetrisations implies considering a different set of decay modes. For example, the amplitude of $B^0 \to K^0_S K^0_S K^0_S$ vanishes when fully-antisymmetrised due to the fact that the resonances of spin 0 and 2 disappear when antisymmetrising. Moreover, since it cannot be assessed in a general way, the hypothesis of flavour SU(3) has to be tested for each set of modes and each symmetrisation. Unlike the fully-symmetric case, for antisymmetric and mixed amplitudes, flavour SU(3) breaking cannot be determined from a simple ratio of amplitudes. A more sophisticated treatment is then required.

Chapter V

The LHCb experiment

Elementary particles can be detected through their interaction with matter. A particle physics detector is made of several subdetectors that are dedicated to measure the properties, for instance momentum, position, energy etc., of different types of particles—for example, some subdetectors are designed to detect charged particles, while others can also detect neutral ones. The information provided by each subdetector is then combined to reconstruct the whole event.

The analysis presented in chapter VI uses data collected with the LHCb experiment at the Large Hadron Collider (LHC). The first section of this chapter describes the LHC accelerator complex and the second part gives an overview of the LHCb detector.

V.1 The LHC

The LHC [108] is a 27 km-circumference proton-proton[1] collider, part of the CERN accelerator complex. It was installed in the same tunnel as the former LEP machine, situated 45 m to 170 m underground, below the Franco-Swiss border. A part of the LEP injection chain is also re-used for the LHC. A schematic of the CERN accelerator complex is shown in Fig V.1. The LHC accelerates two proton beams in opposite direction using 8 radio-frequency (RF) cavities per beam. The beam is bent by superconducting dipole magnets that provide an 8.3 T magnetic field. In addition, quadrupole magnets are used to focus the beam and prevent it from deviating from its trajectory. There are

[1] Most of the time, the LHC accelerates and collides two proton beams, but special runs with proton-ion, ion-proton and ion-ion configurations are also performed.

a total of 9,593 magnets in the LHC.

Before reaching the LHC, the proton beams are accelerated through a series of smaller accelerators. The protons are obtained from a source of ionised hydrogen gas, and accelerated at first by a linear accelerator, LINAC2, to an energy of 50 MeV. They are then injected into the Proton Synchrotron Booster (PSB), a 157 m-circumference quadrupole ring, that splits the beam into bunches of about 10^{11} protons and accelerates them up to 1.4 GeV. Next, the proton bunches enter the Proton Synchrotron (PS) and are accelerated up to 25 GeV. The beam is also further separated to achieve the required structure to enter the LHC, consisting of a separation between the bunches, "bunch spacing", of about 25 ns. The last acceleration stage before the LHC is performed by the Super Proton Synchrotron (SPS) where the protons reach 450 GeV before being injected into the LHC storage ring where the beams are subsequently accelerated to the requested energy.

The energy of the beams in the LHC was increased over the years. During run I, collisions occurred at a centre-of-mass energy of 7 TeV in 2011 and 8 TeV in 2012. The beam energy was further increased during run II, providing collisions at a centre-of-mass energy of 13 TeV. Less than 5 minutes are necessary to fill each LHC beam, and it takes about 20 minutes to reach the run II energy (6.5 TeV per beam).

The collisions of the two beams happen at four interaction points where detectors are located. Two general purpose experiments, ATLAS [109] and CMS [110] are located on opposite sides of the LHC ring, at Point 1 and Point 5 respectively. The ALICE [111] detector, situated at Point 2, is dedicated mainly to the study of heavy-ion collisions and quark-gluon plasma. The LHCb [112] spectrometer is located at Point 8 and is specially designed for the study of flavour physics.

V.2 The LHCb detector

The LHCb detector is a single-arm forward spectrometer designed for precision measurements of beauty and charm hadrons and *CP* violation. Thanks to the large beauty and charm cross-sections at the LHC, the LHCb detector collected approximately 10^{12} heavy flavour decays in run I, which corresponds to the data taken during 2011 and 2012 at centre-of-mass energies of 7 TeV (in 2011) and 8 TeV (in 2012). At these energies, the beauty cross section is about two hundred times smaller than the total cross-section, and the charm cross-section about ten times smaller than the total, which implies that the decay modes of interest are buried under a large amount of background. The signals originating

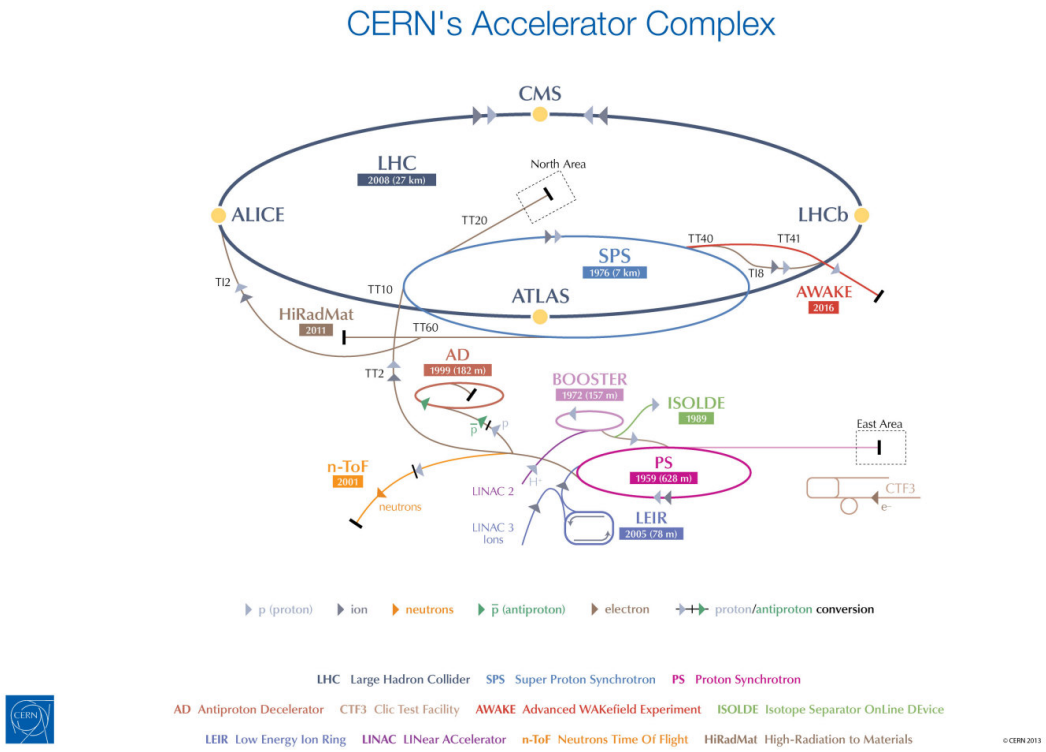


Figure V.1: Scheme of the CERN accelerator complex and the four collision points situated along the LHC ring [113].

from beauty and charm decays are generally characterised by a secondary decay vertex that is displaced from the proton-proton interaction point (the primary vertex), and a high transverse momentum. At high energies, the b and $\overline{b}$-hadrons are very correlated and are mainly produced in the same forward (or backward) cone around the beam. The LHCb detector is specially designed to recognise this typical signature and to provide the necessary information to make precision measurements of b- and c-hadron decays. The data collected by the LHCb detector thus far consist of two separate runs. The run I data, taken in 2011 and 2012 at centre-of-mass energies of 7 TeV and 8 TeV, respectively, correspond to an integrated luminosity of $3\,\text{fb}^{-1}$. The second run of data taking (2015–2018) consists of nearly $6\,\text{fb}^{-1}$ of pp collisions collected at a centre-of-mass energy of 13 TeV.

The layout of the LHCb detector is shown on Fig. V.2. The detector covers

a pseudorapidity range of approximately $2 < \eta < 5$ in the forward direction. A right-handed Cartesian coordinate system, with the z axis defined along the beam axis, the y axis in the vertical direction and the x in the horizontal plane is used. Alternatively, cylindrical polar coordinates (r, θ, z) are also used when needed. Positive values of z correspond to the forward (downstream) region, and negative values of z to the backward (upstream) region. The LHCb detector has excellent vertex resolution thanks to the vertex locator (see Sec. V.3.1), which is crucial for precise measurements of the displaced vertices from beauty and charm decays. The resulting decay time resolution is good enough to resolve the fast oscillations of the neutral B_s^0 meson. In addition to this, the good momentum and invariant mass resolution of the detector (Sec. V.3) as well as the charge particle identification (Sec. V.4) help to suppress the combinatorial background and to distinguish between heavy-flavour decays with similar topologies, such as the $B^0_{(s)} \to K^0_S h^+ h'^-$ decay modes that are studied in Chapter VI. The possibility to detect photons thanks to the calorimeters allows radiative decays to be reconstructed, as well as decays containing π^0 and η particles in the final state. The different subdetectors will be described in the following sections.

The operating conditions of the detector are also adapted to the study of heavy-flavour particles. As the lifetime of b- and c-hadrons is rather long, the typical distance travelled by a B^0 in the detector before decaying is around 7 mm, such that the decay vertex is significantly displaced with respect to primary vertex (PV). The algorithms to associate tracks and particles to their primary or secondary vertex of origin are more efficient with a low number of PVs per bunch-crossing. The pile-up, defined as the average number of visible interactions per bunch-crossing [114], is thus a critical parameter for detector performance. The instantaneous luminosity delivered by the LHC from 2012 on exceeds the original design luminosity of the LHCb detector by a factor of two, as illustrated on Fig V.3. Nevertheless it was proven that the detector is still efficient at this higher luminosity. Since 2011, a levelling procedure of the luminosity at the LHCb interaction point is in use. This levelling is performed by adjusting the transverse overlap of the beams so as to keep the luminosity approximately constant during a fill. This is particularly convenient because it allows to keep the running conditions and trigger configuration during a fill, and reduces the systematic uncertainties related to changes in the detector occupancy. Figure V.4 demonstrates the effects of the levelling by comparing the evolution of the luminosity for LHCb with ATLAS and CMS detectors during one fill in 2012. The levelling procedure has since been adopted by the general-purpose detectors as well.

Thanks to the versatility and performance of the detector, the LHCb physics

program has grown beyond the original idea of studying b- and c-hadron decays and now also includes, for instance, electroweak, heavy-ion, and soft-QCD physics analyses. The LHCb spectrometer also takes data during from special runs with proton-lead or lead-lead collisions. The capability to inject gas in the vertex locator, initially intended for measuring the beam size and luminosity, is now also used to study fixed-target collisions. The design of the detector is such that it is possible to collect fixed-target data and collision data simultaneously.

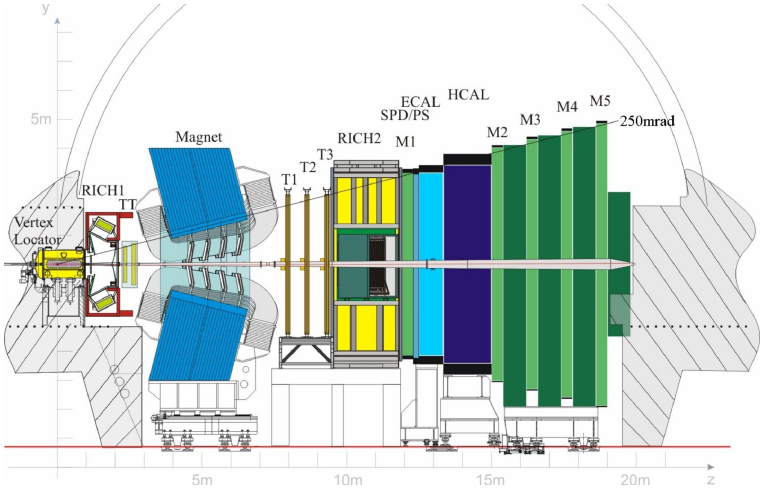


Figure V.2: Side view of the LHCb detector [115].

V.3 Track reconstruction

Reconstruction of the trajectories of charged particles is performed by the tracking system which consists of the vertex locator (VELO), which is situated around the interaction point, and the tracking stations that are placed on both sides of a dipole magnet with a bending power of 4 Tm in the x direction. The three tracking stations situated downstream of the magnet are referred as T-stations. Charged particles may interact with the sensors of the different tracking detectors. The positions of the hits are then used to reconstruct the trajectories of the charged particles. When passing through the magnet they are deflected by the magnetic field, which enables a measurement of the momentum of the track.

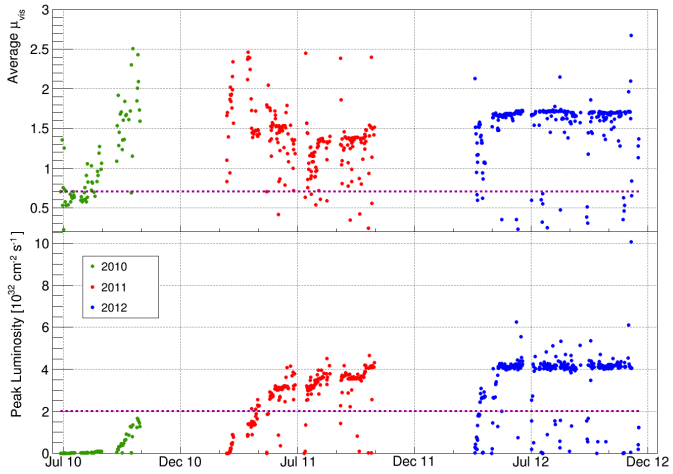


Figure V.3: The upper plot shows the pile-up, μ, at the LHCb interaction point over run I. The lower plot shows the instantaneous luminosity for the same period. The dotted lines show the design values [116].

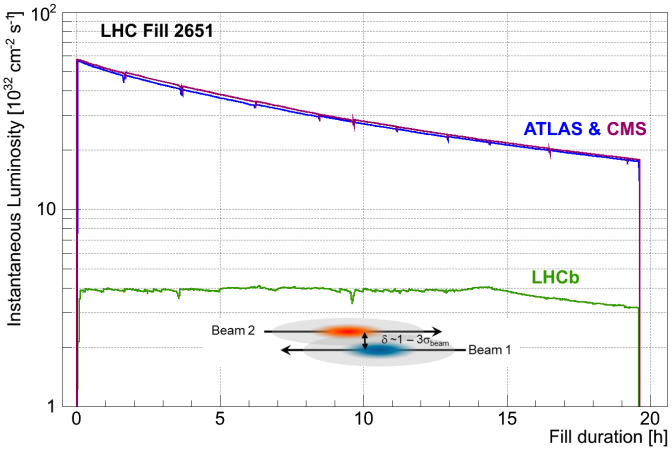


Figure V.4: Instantaneous luminosity at the LHCb, ATLAS and CMS detectors as a function of the time, during a specific fill from May 2012. It can be seen that LHCb was able to maintain a stable luminosity (within 5%) over 15 hours; after 15 hours the beams are colliding head-on at the maximum available luminosity (lower at LHCb than ATLAS and CMS due to the difference in the final focusing at the collision points [116]).

V.3.1 The vertex locator

The VELO [117] is a silicon detector composed of 21 stations placed along the beam, surrounding the interaction region, as shown in Fig. V.5. Each VELO station is composed of two semi-circular modules that contain R and Φ sensors. The R sensors that measure the r coordinate of the track (radial distance from the z axis), are composed of strips that are arranged in four segments of approximately 45°. The Φ sensors, that measure the ϕ coordinate of the track (azimutal angle), are divided into two regions, with inner and outer strips. A schematic view of the sensors is given in Fig. V.6. The design of the VELO is such that all the tracks inside the LHCb acceptance cross at least three VELO stations, which corresponds to the minimum number of hits required for a track to be labelled as "reconstructible". Tracks without momentum information in the forward and backward directions are also reconstructed to improve the quality of the PV reconstruction. The sensors are operated in a vacuum and cooled using a bi-phase CO_2 cooling system that maintains a temperature of $(-7 \pm 2)^{\circ}$C, and the readout is performed through routine lines that transport electrical signals to the edge of the sensor. Due to its proximity to the beam, the VELO detector is exposed to a high amount of radiation. To reduce the risk of serious damage from the beam impacting the sensors directly, the VELO modules are retractable: when not operating or during beam injection (i.e. outside STABLE_BEAMS), the two halves of the detector are opened. During data taking, the sensors are moved inward to a distance of 7 mm from the beam, with a small overlap between the two halves to help with the alignment.

The PV position resolution is 13 µm in the transverse plane and 71 µm along the beam axis for vertices with 25 tracks. For particles with a transverse momentum greater than 1 GeV, the impact parameter (IP) resolution is below 35 µm.

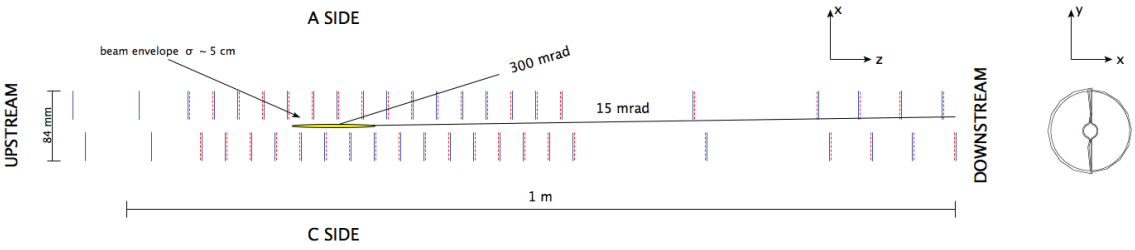


Figure V.5: View of the VELO senors in xz plane at $y = 0$ in the closed position [117]. The R sensors are represented with a with solid blue lines and the Φ sensors are shown by dashed red lines.

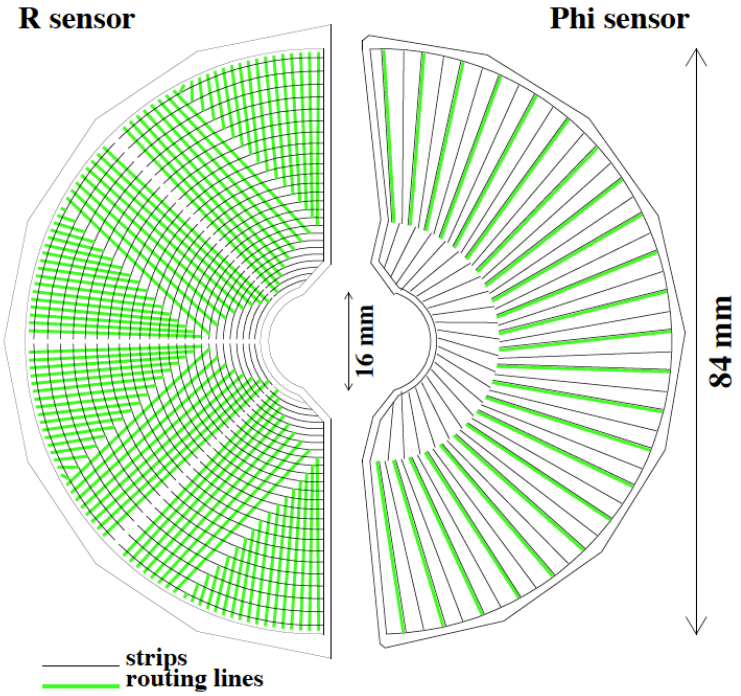


Figure V.6: Schematic representation of the R and Φ sensors of the VELO [117].

V.3.2 The Tracker Turicensis

The Tracker Turicensis (TT) is a silicon detector situated upstream of the magnet that covers the full acceptance of the LHCb detector. It is made of four layers of silicon microstrip sensors which are 9.64 cm wide, 9.44 cm long, and 500 µm thick. The layers are arranged to measure the so-called "stereo" coordinates in an x-u-v-x configuration: the two x layers have vertical strips while the u and v stations have their strips rotated by angles of $-5°$ and $+5°$ with respect to the x-layers. The TT detector is particularly important in the reconstruction of long-lived particles such as K^0_S and Λ hadrons that may fly a significant distance and decay after or near the end of the VELO (leaving fewer than three hits). The TT also improves the quality of the reconstruction of low-momentum tracks.

V.3.3 The Dipole Magnet

The magnet [118], located between the TT and the T-stations, is a warm dipole that provides an integrated magnetic field of about 4 Tm, mainly oriented along the y direction. A view of the magnet is shown on Fig. V.7. It consists of two trapezoidal coils of 25 tons each, inclined at 45° on the two transverse sides, arranged inside a 1450-ton iron yoke. The evolution of the y component of the magnetic field as a function of the z position is shown on Fig. V.8. It can be observed that the field is concentrated between TT and T1, with a field strength that is low in the tracking stations themselves and minimal inside the VELO.

The charged particles passing through the magnet are deflected in the x direction by the magnetic field, which allows for a measurement of their momentum with a relative precision of about 0.5% for tracks with momenta in the range 2 GeV to 100 GeV. Depending on their charge, the particles are deflected in opposite directions; if the detector efficiency is not exactly the same on both sides, charges asymmetries will be observed between conjugates modes. These left-right asymmetries, like any other asymmetries in the detection, are particularly problematic for *CP* violation measurements. To cope with this effect, the magnet has been designed in such a way that its polarity can be reversed, and data is taken with both up and down magnet polarities. If the amount of data taken with both polarities is similar and the detector conditions do not vary significantly over time, the asymmetry will approximately cancel when averaging results over the polarities.

For charged track reconstruction, a precise mapping of the magnetic field across the detector is essential. This was initially done in 2005 by measuring the field at about 500 000 points in the detector in steps of 4 cm in the x-y plane and 10 cm along the z axis. Corrections to this map were the provided in 2011, due to the influence of the magnetic field on the position of the subdetectors.

V.3.4 The Inner Tracker

The three T-stations situated downstream of the magnet are each composed of an Inner Tracker (IT) [119] and an Outer Tracker (see below). The IT is located around the beam pipe, in a region of high detector occupancy. The IT subdetector of each T-station includes a light-tight box, cooled to a temperature below 5°C, containing four layers of silicon-strip detectors organised in a stereo arrangement similar to the TT. The efficiency and the hit resolution of the IT are also similar to that of the TT. However, unlike the TT, the power and cooling systems, as well as the readout boards of the IT are inside the LHCb

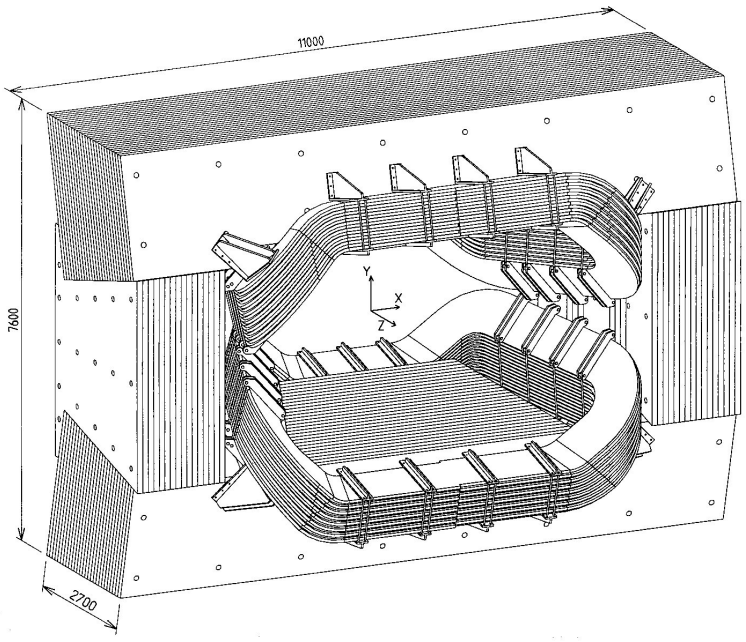


Figure V.7: Schematic of the LHCb dipole magnet [118].

acceptance, creating asymmetries in the reconstruction of the charged tracks.

V.3.5 The Outer Tracker

The Outer Tracker (OT) [120, 121] is a gaseous detector covering an area of approximately $5\times6\,\mathrm{m}^2$. Each of the three stations is made of four double layers of straw tubes, arranged in the same stereo setup as the IT and the TT. The drift tubes have a 4.9 mm inner diameter and are filled with a gas mixture of Ar (70%), CO_2 (28.5%) and O_2 (1.2%), which ensures a drift time below 50 ns and a drift coordinate resolution of 200 μm. Charged particles passing through the tubes ionise the gas mixture, creating ions that drift toward the anode wires that are maintained at a voltage of +1550V. The drift time of the ions provide a measurement of the position of the ionising particles.

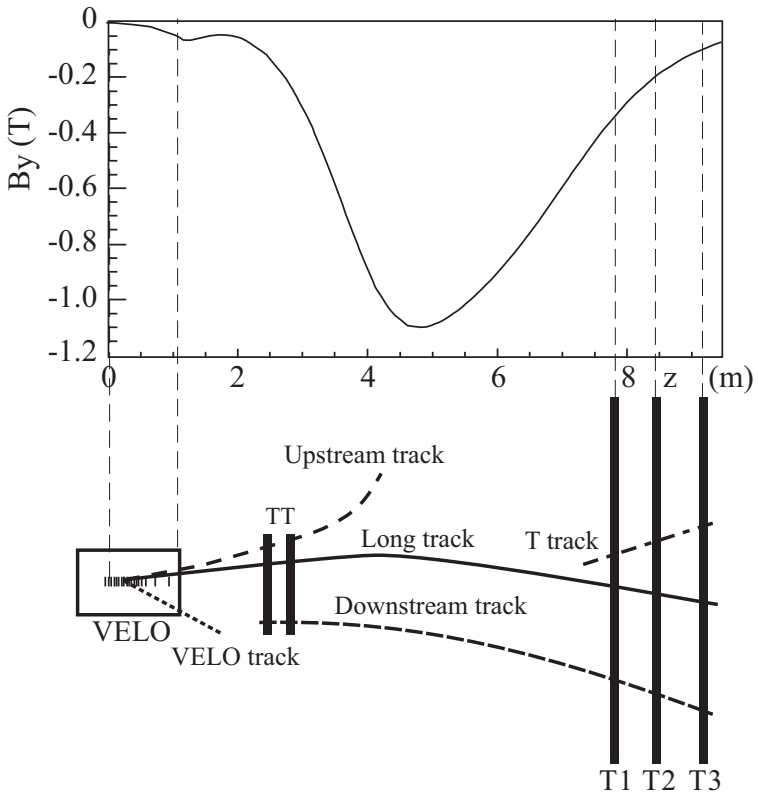


Figure V.8: The magnetic field component in the y direction (B_y) as a function of the z position. An illustration of the tracking system is shown for comparison [116].

V.3.6 Track and vertex reconstruction

Track reconstruction

The track reconstruction [122] is performed by combining the individual measurements provided by the tracking detectors into a particle trajectory. The design of the LHCb detector is such that the magnetic field in most of the tracking detectors is negligible, so that the track segments upstream and downstream to the magnet can be approximated by straight lines. Tracks are classified into different types according to which subdetectors contain hits. Figure V.9 gives an overview of the different track types. The analysis presented in chapter VI uses principally `Long` tracks, which go through the full tracking system and have hits in (at minimum) the VELO and the T-stations, and `Downstream` tracks that have hits in the TT and the T-stations but not in the

VELO. The `Downstream` tracks are useful to reconstruct long-lived particles that decay outside the VELO, such as K^0_S and Λ hadrons. Other types of tracks exists, such as VELO tracks that only have hits in the VELO and are principally useful for primary vertex reconstruction, and `upstream` tracks that have hits in the VELO and the TT, and T-tracks that only have hits in the T-stations downstream of the magnet. The latter are mainly used to improve the performance of the RICH detectors' particle identification algorithms.

The reconstruction of `Long` tracks starts by identifying VELO track segments that are straight lines with at least three hits in the VELO stations. Then, these segments are combined with information coming from the T-stations. Two algorithms are used to combine the VELO tracks with the hits in the T-stations. In the forward tracking algorithm [123], a VELO segment is combined with a single hit in a T-station to obtain a measurement of the trajectory and of the particle momentum. The search is then extended to further hits in the T-stations along the trajectory to improve the track quality. The track matching algorithm [124, 125] combines VELO track segments with T-tracks, which are required to have at least one hit in the x layers and one hit in the stereo layers of each T-station. Finally, the candidates produced by the two algorithms are combined to from `Long` tracks, and hits in the TT compatible with their trajectories are added to improve the momentum determination. The `Downstream` track reconstruction proceeds in the opposite direction: T-tracks from the T stations are extrapolated through the magnetic field and a search for compatible hits is performed in the TT.

The average reconstruction efficiency for `Long` tracks is above 96% in the momentum range $5\,\text{GeV} < p < 200\,\text{GeV}$ in the angular acceptance of LHCb [116]. The efficiency is slightly degraded for high-multiplicity events.

Primary vertex resolution

The PV resolution is measured by randomly splitting the set of tracks in an event in two and reconstructing the PVs with each subsample. The resolution is then obtained from the width of the distribution of the difference of the vertex positions. The resolution is strongly correlated with the track multiplicity. The x and y PV resolution as a function of the track multiplicity is shown in Fig. V.10.

Impact parameter resolution

The impact parameter is widely used in analyses to separate secondary tracks from those produced in the primary interaction. It is defined as the distance

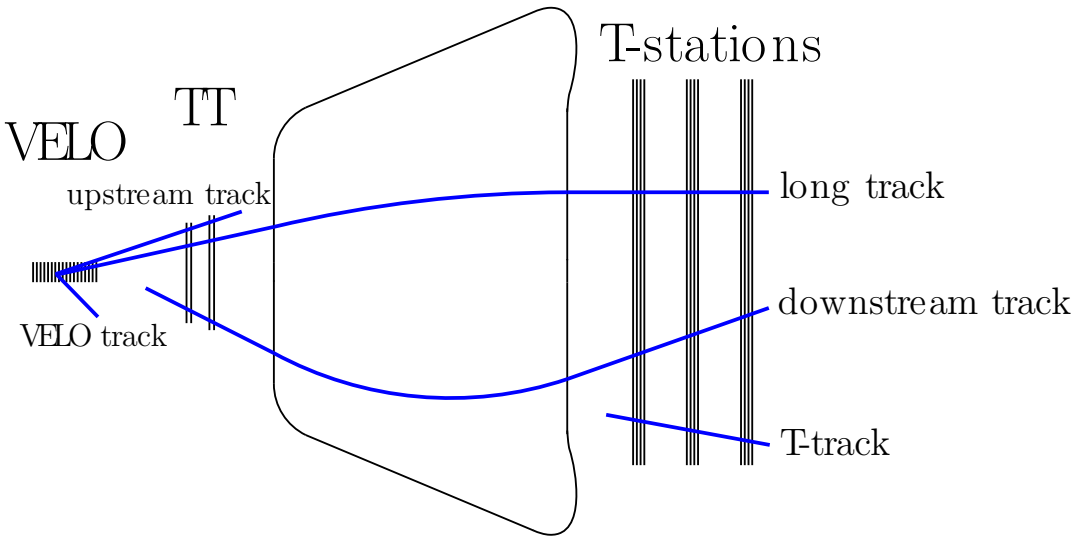


Figure V.9: Illustration of the track types in LHCb [126].

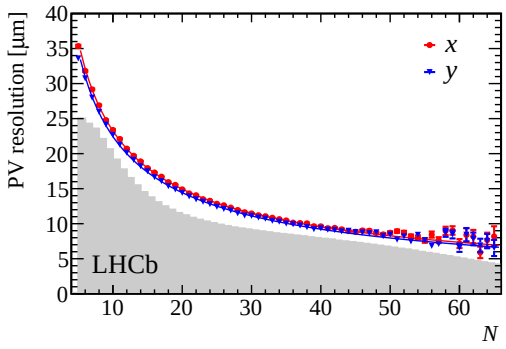


Figure V.10: Primary vertex resolution as a function of the track multiplicity, obtained with data collected in 2012. The resolution in the x coordinate is shown in red and the resolution in the y direction in blue [116].

between the point of closest approach of a particle to the PV and the PV itself. A good IP resolution is crucial for the identification of B and D mesons, taking advantage of the fact that they tend to fly a significant distance before decaying. The IP resolution mainly depends on the hit resolution in the tracking system, the distance a track must be extrapolated between its first measured hit and the vertex, and the amount of multiple scattering that the particle undergoes after interacting with the material in the detector. The VELO detector has been specifically designed to minimise theses factors. The IP resolution has a linear dependency on the inverse of the transverse momentum of the track, which is observed in both x and y directions. At asymptotically high p_{T}, the resolution is about 12 µm. The IP resolution as a function of the momentum and the transverse momentum is shown in Fig. V.11.

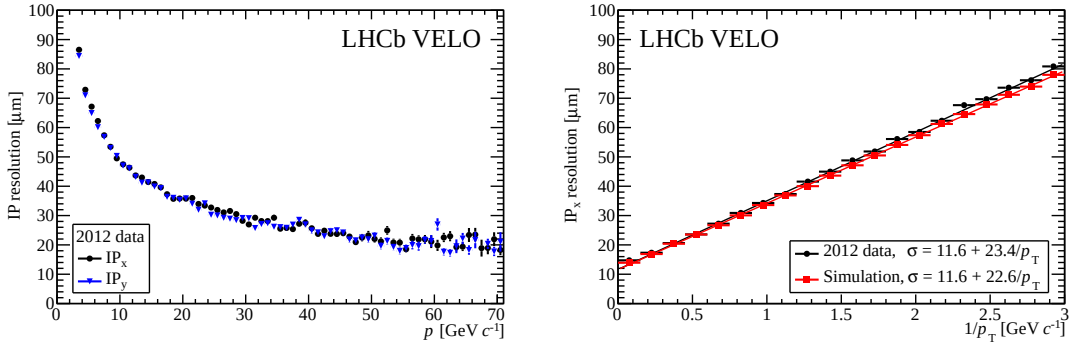


Figure V.11: (Left) IP resolution in the x and y directions as a function of the momentum, obtained from data collected in 2012. (Right) Comparison between data and simulation of the IP resolution in the x direction as a function of $1/p_{\mathrm{T}}$. [117].

K^0_{S} and Λ reconstruction

V^0 particles such as K^0_{S} and Λ are reconstructed from two tracks of opposite charge ($K^0_{\mathrm{S}} \rightarrow \pi^+ \pi^-$ and $\Lambda \rightarrow p \pi^-$). Due to the their lifetime, the decay of these particles can happen either inside the VELO, in which case the daughter particles are reconstructed as `Long` tracks, or outside (or at the very end of) the VELO, in which case the daughter tracks are then classified as `Downstream` tracks. Thanks to information from the VELO, the precision on the direction of `Long` tracks is better than that for `Downstream` tracks, which in turn leads to a significantly better mass resolution for V^0 candidates reconstructed with `Long` tracks. The typical mass resolution of a K^0_{S} reconstructed with `Long`

tracks is 3.5 MeV, and with `Downstream` tracks it is 7 MeV. Invariant mass spectra of K^0_S reconstructed with both types of tracks are shown in Fig. V.12.

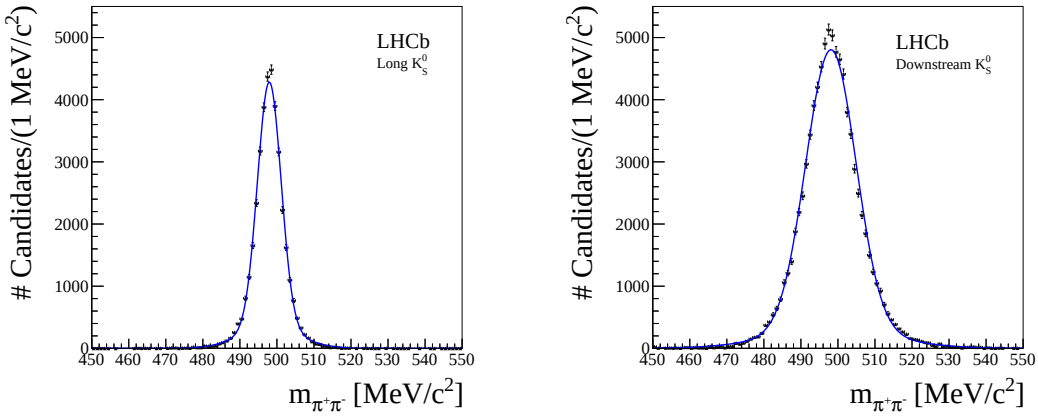


Figure V.12: Invariant mass of $K^0_S \to \pi^+ \pi^-$ candidates reconstructed from `Long` tracks (left) and `Downstream` tracks (right) [116].

V.4 Particle identification

The identification of particles is performed by the two Ring-Imaging Cherenkov (RICH) detectors, the hadronic (HCAL) and electromagnetic (ECAL) calorimeters, and the muon detector.

V.4.1 The Ring-Imaging Cherenkov detectors

The main purpose of the RICH system [127] is to provide identification capabilities for charged hadrons (pions, kaons and protons) that are used in the software trigger and at the analysis level. Information from the RICH detectors also contributes to charged lepton identification, together with the calorimeters and the muon detector.

Charged particles in a medium that propagate faster than the speed of light in that medium emit photons at a specific angle, θ, that depends on the velocity of the particle, βc, and the refractive index of the medium, n:

$$\cos\theta = \frac{1}{n\beta}. \tag{V.1}$$

The RICH detectors use several types of medium, called radiators, to measure the angle of emission of the Cherenkov photons and identify the particles.

The two RICH detectors, RICH1 and RICH2, are situated upstream and downstream of the magnet, respectively. This choice of location is justified by the strong correlation between the polar angle and the momentum of the particles; usually, to high-momentum particles correspond to low polar angles. Consequently, RICH1, covers the low and intermediate momentum region, from 2 GeV to 40 GeV, over the full angular acceptance of the LHCb detector (25–500 mrad), and RICH2 covers the high-momentum region, from 15 GeV to 100 GeV, over a reduced angular range of 15–120 mrad. The optical system, similar for both detectors, is composed of a spherical focusing primary mirror and a secondary flat mirror, which reflect the photons outside of the spectrometer acceptance. This allows the photodetectors and electronics to be installed outside the acceptance, reducing the amount of material that particles in the acceptance must cross. Fluorocarbon gases are used as radiators: C_4F_{10} with a refractive index of $n = 1.0014$ is used in RICH1, and CF_4 with $n = 1.0005$ is used in RICH2. The Cherenkov photons are reflected by the mirrors situated into the radiators and focused into ring images on plane of hybrid photon detectors (HPDs). Figure V.13 shows the reconstructed Cherenkov angle as a function of the momentum for isolated tracks. Depending on the mass of the particle, the events populate distinct bands. The hadrons form three clearly visible and well separated bands, and contributions from muons and electrons are also visible.

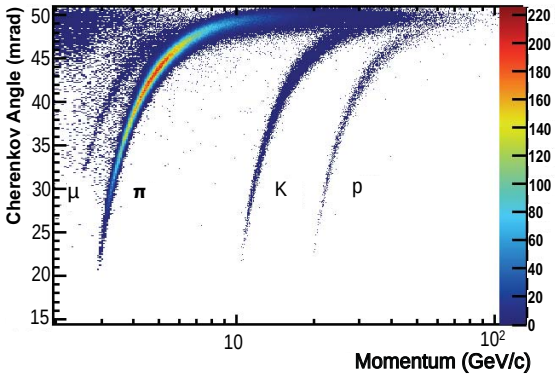


Figure V.13: Reconstructed Cherenkov angle as a function of the track momentum in the C_4F_10 radiator for isolated tracks [127].

V.4.2 The calorimeters

The calorimeters are principally used to identify electrons, photons and neutral pions and to provide a measurement of their energy and position. They are also used by the hardware trigger to select candidates with high transverse energy.

Particles interacting with the dense material contained in the calorimeters producing particle showers, and the energy deposited is subsequently measured by scintillator planes. The ECAL is designed to absorb and measure almost all of the energy of an electron or photon shower, with minimal leakage into the HCAL except at the highest energies.

The calorimeter system [128] is composed of a Scintillating Pad Detector (SPD), whose purpose is to improve the separation of electrons and photons, a preshower (PS) detector that is used to reduce the background from charged pions, an ECAL, and an HCAL. The ECAL, situated after the PS, is composed of 66 layers of 2 mm lead planes alternating with planes of scintillator pads of 4 mm, oriented perpendicular to the beam direction. Its energy resolution is $(10/\sqrt{E} \oplus 1.5)\%$ for E in units of GeV [128]. The energy deposits made by hadrons are measured in the HCAL, which consists of 26 modules of alternating layers of iron and scintillator planes perpendicular to the beam axis. The HCAL energy resolution is of $(80/\sqrt{E} \oplus 10)\%$, again for E in GeV [128].

V.4.3 The muon stations

The muon system [129, 130] is composed of five stations, one placed before the calorimeters (M1) and four placed after (M2–M5). They provide muon identification and contribute to the L0 trigger. Most of the particles that arrive in M2–M5 are muons; the other particles either decay before or are absorbed by the calorimeters. This is not true for M1; it is needed to improve the muon p_{T} measurement in the L0 trigger rather than for offline muon identification.

The muon detector is huge: it is composed of 1380 chambers and covers a total area of 435 m^2. The five rectangular stations use principally Multi-Wire proportional chambers (MWPCs), except in the highest-rate region of M1 that uses Gas Electron Multipliers (GEMs), which are more robust against radiation damage. The stations M2–M5 are interleaved with 80 cm thick iron absorbers to stop any remaining hadrons. They also absorb low-momentum muons; muon candidates must have momentum of at least 6 GeV to traverse the five muon stations.

V.4.4 Particle identification methods

The RICH, calorimeters and muon systems provide independent PID information. The separation between charged pions, kaons and protons is done by the RICH. Electrons and photons are identified using energy deposits (clusters) in the ECAL. Photons, being neutral particles, are identified from clusters not associated with any track nor signals in the SPD and PS. Neutral pions are reconstructed from two photons, which can be either two separate clusters in the ECAL (resolved π^0) or one large cluster, caused by two photons emitted with little angular separation (merged π^0).

For charged particles, the information provided by these particle identification systems is combined into sets of related variables. Two sets of variables are constructed. In the first approach, the individual log-likelihoods provided by each subdetector are added linearly to form a combined log-likelihood difference of the form

$$\Delta \log \mathcal{L}_{\text{comb}}(X - \pi), \tag{V.2}$$

where X corresponds to an electron, muon, kaon or proton hypothesis. These variables are called DLL variables. Another approach, developed later, is based on a multivariate technique (using a neural network), and takes into account the correlations between the subdetectors as well as additional information. The variables built in this way are denoted `ProbNN`, and are the ones used in the analysis presented in chapter VI.

V.5 Trigger

Because of the luminosity and the high collision rate of the LHC, an enormous quantity of data is produced and the resources to process and store all these data do not exist. The main purpose of the trigger system is to reduce the amount of data stored on disk by selecting events that are potentially interesting for physics analyses while rejecting the others.

The LHCb trigger consists of hardware and software levels. The hardware trigger, L0, is synchronised with the LHC clock. It reduces the event rate from 40 MHz to a little below 1 MHz, which is the maximum sustained rate at which the whole detector can be read out. The L0 trigger needs to take fast decisions, and for this purpose only partial information coming from the detector is used. There are four main L0 trigger lines: `L0Hadron`, `L0Electron`, `L0Photon`, and `L0Muon`. The hadron, electron and photon trigger decisions are based on an estimate of the transverse energy obtained from measurements of

energy deposits in the calorimeters, while the muon trigger decision is taken by estimating the transverse momentum using the muon detector.

The higher level trigger is implemented in software and divided into two sub-levels, HLT1 and HLT2. This division is performed for reasons of timing. Unlike L0 which uses information coming from individual subdetectors, the HLT performs an online event reconstruction. HLT1 performs a partial reconstruction using information from the VELO and the T-stations to reduce the rate of the data processed by HLT2. This trigger selects candidate-based quantities such as the p_{T}, p, the distance to the PV and the track quality. HLT2 performs a full event reconstruction on the events that passed HLT1 reducing the event rate to about 3–5 kHz in run I and 12.5 kHz in run II. Many trigger selections ("trigger lines") are available at this level depending on the purpose of the analysis. In the analysis presented in chapter VI, we use exclusively the so-called inclusive topological HLT lines [131], which makes decisions based on topological requirements on the final state. These lines are specifically designed to select generic n-body b-hadron decays. Their original design was based on a cut-based approach that was abandoned at the end of 2011 for a method using multivariate analysis with better performance. Another type of HLT2 line that is not used in the analysis presented in this thesis consists of exclusive lines that are designed to select specific final states. Diagrams representing the data processing by the different trigger levels are shown in Fig. V.14 for the 2012 and 2015 years of data taking, highlighting the differences between run I and run II.

The collection of tracks (or other detector activity) that causes a positive trigger decision is referred as a "trigger object". Trigger decisions are classified into three categories depending on which par of the event is responsible for the positive decision:

- **Trigger On Signal (TOS):** the trigger is fired by signal tracks independently of the rest of the event (i.e. the signal tracks alone are sufficient to fire the trigger). For the TOS criterion to be fulfilled, at least one trigger object must have all of its tracks overlapping with the signal. Two tracks are said to overlap if they share more than 70% of their hits (60% in case of muon tracks).

- **Trigger Independent of Signal (TIS):** the trigger is fired by "the rest of the event" independently of the presence of signal (i.e. the trigger would still have fired in the absence of the signal tracks). An event is TIS if at least one trigger object has no overlap with the signal. In the case of the TIS classification, two tracks are said to overlap if they share more

than 1% of their hits. Since tracks in LHCb have at most around 60 hits, in practice this criterion means that the tracks do not share any hit.

- **Trigger On Both (TOB):** events for which neither the signal alone nor the rest of the event is sufficient to fire the trigger, meaning that both are necessary. These events are neither TIS nor TOS.

Note that these categories are not mutually exclusive, so that an event can be TIS and TOS at the same time.

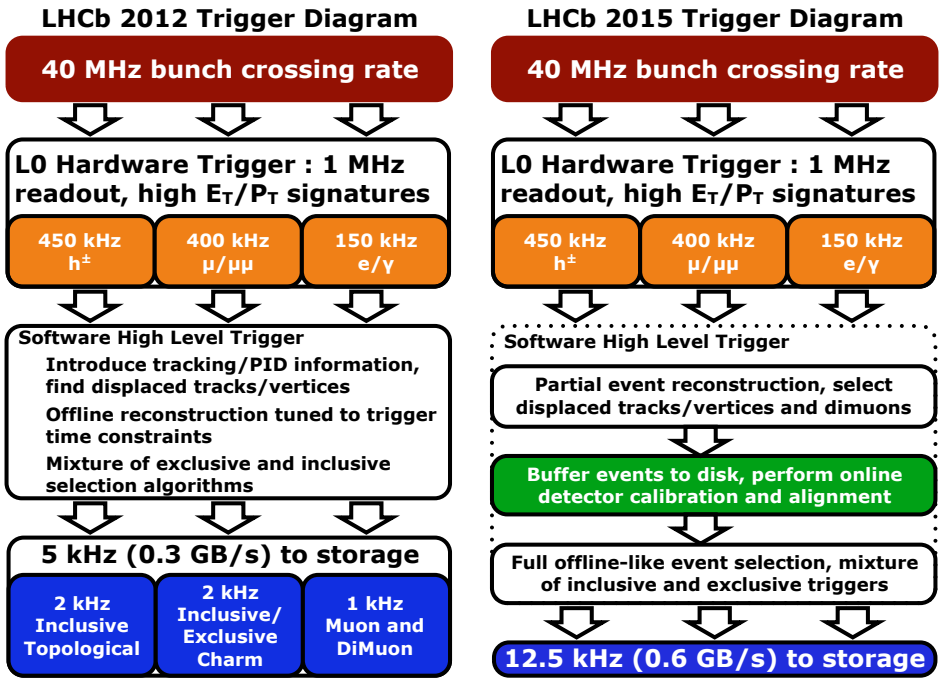


Figure V.14: Schematic diagram of the LHCb trigger during run I (left) and run II (right).

V.6 Simulation

Simulated data, also called Monte-Carlo (MC) data, are used extensively in most analyses at LHCb. In this work, MC samples are used to assess the efficiency of the selection, and in the training of the multivariate selection that reduces the background.

The simulation is performed in two main steps. The first step, the "generator phase", consists of simulating the pp collisions and the decays of the particles produced in those collisions. This is done using the PYTHIA generator. The version used here is PYTHIA 8 [132], which provides a standalone generator, with a specific LHCb configuration [133]. The hard process of the collision[2] is simulated by using parton distribution functions that describe the composition of the incoming protons as a function of transverse momentum, and generates the outgoing partons. The simulation of the showers generated by these partons, the hadronisation and the multiple interactions are also generated by PYTHIA. The decays of hadronic particles are described with EVTGEN [134], with final-state radiation handled by the PHOTOS generator [135]. The second step, called the "simulation phase", consists of simulating the interactions of the particles in the detector and the detector response. The GEANT4 toolkit [136, 137] implemented as described in Ref. [138] is used to simulate the LHCb detector.

[2]Note that PYTHIA is not specific to only pp collisions, e.g. $p\overline{p}$, e^+e^- and $\mu^+\mu^-$ collisions can also be handled by this generator.

Chapter VI

Update of the branching fraction measurements of $B^0_{(s)} \to K^0_S h^\pm h'^\mp$ modes

As explained in chapter III, charmless B-meson decays give access to a wide range of physics observables, such as CKM parameters, CP violation etc. The non-negligible contribution from loop diagrams in $B^0_{(s)} \to K^0_S h^\pm h'^\mp$ decay modes makes them an interesting playground for new physics searches. As a first step towards performing amplitudes analyses of each mode, and to quantify the potential sensitivity of such analyses with the upgraded detector, the branching fractions have to be determined as precisely as possible.

The decays $B^0_{(s)} \to K^0_S h^\pm h'^\mp$ represent six decay modes[1] $B^0_{(s)} \to K^0_S \pi^+ \pi^-$, $B^0_{(s)} \to K^0_S K^+ K^-$, $B^0_{(s)} \to K^0_S K^+ \pi^-$ and their conjugates. These six modes have similar topologies, such that misidentified decays from one mode will appear as structures in other final states; to take these cross-feeds into account, a simultaneous fit to all of the modes is used to measure their branching fractions. A previous LHCb analysis of these modes with $3\,\text{fb}^{-1}$ of run I data [139] was able to measure the branching fractions of five out of the six modes, while the sixth mode $B^0_s \to K^0_S K^+ K^-$ remains unobserved: its statistical significance in the published analysis was 2.4σ. The work described in this chapter focuses on the observation of the missing $B^0_s \to K^0_S K^+ K^-$ decay mode, and represents progress towards an update of the measurement of the branching fractions of the six decay modes with run I and run II data.

[1] In principle there are eight, since e.g. $B^0 \to K_S K^+ \pi^-$ and $B^0 \to K_S K^- \pi^+$ are distinct, but they cannot be distinguished at LHCb without flavour tagging.

The first two sections of this chapter give an overview of the analysis strategy (Sec. VI.1), and of the analysis tools (Sec. VI.2). The dataset and the reconstruction are described in section VI.3. The selection of the candidates is detailed in section VI.4 and the efficiencies corresponding to the different selection steps are given in section VI.5. The fit model and the fit results are discussed in sections VI.6 and VI.7. The systematic uncertainties are presented in section VI.8, the combination of the results is detailed in section VI.9 and the final result and the conclusion are given in section VI.10.

VI.1 Analysis strategy

The decay modes are reconstructed in four final states, namely $K^0_S\pi^+\pi^-$, $K^0_S K^+K^-$, $K^0_S K^+\pi^-$ and $K^0_S\pi^+K^-$ (with the latter two being essentially equivalent in the limit of zero production and detection asymmetries). In each final state, correctly reconstructed contributions from both B^0_d and B^0_s decays are expected, with one of the two favoured with respect to the other: in $K^0_S\pi^+\pi^-$ and $K^0_S K^+K^-$ the B^0_d decay dominates, whereas in $K^0_S K^+\pi^-$ and $K^0_S\pi^+K^-$ the dominant contribution is from the B^0_s. The branching fractions are measured with respect to a normalisation channel, $B^0 \to K^0_S\pi^+\pi^-$, which is the most abundant of the six modes. The master formula for computing the ratio of branching fractions is given by

$$\frac{\mathcal{B}(B^0_{(s)} \to K^0_S h^\pm h'^\mp)}{\mathcal{B}(B^0 \to K^0_S\pi^+\pi^-)} = \frac{\epsilon^{\rm Tot}_{B^0\to K^0_S\pi^+\pi^-}}{\epsilon^{\rm Tot}_{B^0_{(s)}\to K^0_S h^\pm h'^\mp}} \frac{N_{B^0_{(s)}\to K^0_S h^\pm h'^\mp}}{N_{B^0\to K^0_S\pi^+\pi^-}} \frac{f_d}{f_{d,s}}, \tag{VI.1}$$

where $f_{d,s}$ are the hadronisation fractions of the $B^0_{d,s}$ mesons. The ratio f_s/f_d has been measured by LHCb [140,141], and its value is found to be 0.259 ± 0.015 at $\sqrt{s} = 7\,\text{TeV}$; a similar value is found for $\sqrt{s} = 13\,\text{TeV}$ [142]. The signal yields, N, are extracted from a fit and the selection efficiency, $\epsilon^{\rm Tot}$, is obtained from MC simulation (after data-driven corrections). The efficiency term corresponds to that of the overall selection, within which various effects have to be taken into account: acceptance, trigger, reconstruction and selection efficiencies. The total efficiency can be expressed as a product of the different contributions

$$\epsilon^{\rm Tot} = \epsilon^{\rm Acc} \times \epsilon^{\rm Trigger|Acc} \times \epsilon^{\rm Reco|Trigger} \times \epsilon^{\rm Sel|Reco}, \tag{VI.2}$$

with each term (other than $\epsilon^{\rm Acc}$) computed with respect to the previous requirement. An interesting point of Eq. VI.1 is that, since the decay modes have similar topologies and kinematic properties, most of the systematic effects

related to the computation of the efficiencies (other than the hadron ID requirements) are expected to cancel to first order in the ratio.

Different types of background are present in the $B^0_{(s)} \to K^0_S h^+ h'^-$ decay modes. They can be classified into four categories depending on their origin:

- Combinatorial background that comes from random combinations of unrelated tracks of opposite charge is the largest source of background. Because of its nature, it does not appear as a peak in the invariant mass of the B meson, but it is rather present throughout and varies smoothly. The strategy to reduce this background is to use a multivariate classifier to differentiate between this background and the signal. The remaining combinatorial background present after the selection is included as a component in the final fit.

- Crossfeed backgrounds come from $B^0_{(s)} \to K^0_S h^+ h'^-$ decays with a misidentified particle. For example, events coming from $B^0_s \to K^0_S K^+ \pi^-$ with the kaon misidentified as a pion can be present in the $B^0 \to K^0_S \pi^+ \pi^-$ spectrum. This type of background is particularly dangerous because it creates peaking structures that are partly underneath to the signal peak. The crossfeeds are partially suppressed by the PID selection and the remaining contributions are modelled in the fit.

- Decays with an intermediate charm or charmonia particle (e.g. $B^0 \to J/\psi K_S$ with $J/\psi \to h^+ h^-$, $B^0 \to D^- \pi^+$ with $D^- \to K_S \pi^-$) form peaking structures that are visible in the invariant mass spectra of two-body particles. These contributions are vetoed. Such decays with a single misidentified particle are also taken into account, but those with double misidentification are neglected.

- Partially-reconstructed backgrounds arise from four-body (or more) final states with at least one particle that is not included in the reconstructed B candidate. This kind of background is predominant in the left-hand sideband of the mass spectrum and has to be modelled in the fit with a dedicated function based on the expected lineshapes of the main contributing decay modes (see Sec. VI.6).

The data sample used in this analysis consists of the data taken by the LHCb detector during run I and 2016, which corresponds to an integrated luminosity of $4.67\,\text{fb}^{-1}$. (The comparatively small 2015 sample is not used.) The data is divided into different categories according to the data-taking period and the K^0_S reconstruction class (`Long` or `Downstream`; see Sec. V.3.6). A

simultaneous unbinned extended maximum likelihood fit to all the samples is then performed to extract the signal yields in each category.

VI.2 Analysis tools

The tools used in this analysis are presented here with a focus on the way they are used in this particular case. The maximum likelihood estimator used in the fit is explained in Sec. VI.2.1, and the general principles of multivariate analysis are discussed in Sec. VI.2.2. The ${}_s\mathcal{P}lot$ method, used for background subtraction, is presented in Sec. VI.2.3 and the `TISTOS` method, used to evaluate the trigger efficiency, is discussed in Sec. VI.2.4. Apart from the `TISTOS` method, which is specific to LHCb, the other tools are commonly employed in high energy physics as well as in other fields of science.

VI.2.1 Maximum likelihood estimator

The goal of a physics analysis is to extract information on the parameters of interest from the distribution of events in the data. To do so, a model that describes the expected distribution of the observables given a set of parameters is built. Then, the parameters of the model are inferred using the distribution of the observables in the data sample. Different methods exist to estimate the best values for the parameters, such as the method based on a χ^2 function used to extract a value of the angle γ described in the chapter IV. In this analysis, we use another approach, the maximum likelihood estimator (MLE), which is widely used in particle physics. The main interest of such estimators is that they are unbiased and efficient on large data samples.

The distributions of observables are modelled with a probability density function (PDF), $f(\boldsymbol{x};\boldsymbol{\theta})$, where $\boldsymbol{x}$ is the observable (or a set of observables) and $\boldsymbol{\theta}$ are the parameters of interest. The PDFs are positive-definite and are normalised to unity over the range of $\boldsymbol{x}$. In practice, in order to properly describe the relation between the observables and the parameters of interest, it is often necessary to incorporate additional parameters in the PDF that are extracted along with the parameters of interest. These so-called nuisance parameters are implied here in the definition of $\boldsymbol{\theta}$. For several statistically independent data points x_i, the total likelihood is the product of the individual PDFs evaluated at the different points x_i (with an additional factor for the overall normalisation to make it an extended likelihood, as discussed below).

For N data points, the unbinned likelihood[2] reads

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i=1}^{N} f(x_i; \boldsymbol{\theta}). \tag{VI.3}$$

The maximum likelihood estimator $\hat{\boldsymbol{\theta}}$ is defined as the value of $\boldsymbol{\theta}$ for which the likelihood is maximal. It is often more convenient, for numerical evaluation, to work with the logarithm of $\mathcal{L}$. The negative log-likelihood (NLL) is then defined as

$$-\ln \mathcal{L}(\boldsymbol{\theta}) = -\sum_{i=1}^{N} \ln f(x_i; \boldsymbol{\theta}). \tag{VI.4}$$

The maximum likelihood estimator is then obtained by minimising the NLL, in other words by solving the likelihood equations simultaneously for the different parameters, $\boldsymbol{\theta} = \{\theta_1, \ ..., \ \theta_m\}$,

$$\frac{\partial \ln \mathcal{L}}{\partial \theta_i} = 0, \qquad i = 1, ..., m. \tag{VI.5}$$

For cases where the number of events is not fixed, but rather depends on $\boldsymbol{\theta}$, the normalisation to unity of the PDFs is no longer valid and one has to include the number of events into the definition of the likelihood. For repeated experiments, the observed number of events, n, is assumed to fluctuate according to a Poisson distribution around a mean value, μ, which corresponds to the expected number of events. Under these circumstances, Eq. VI.3 becomes

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{\mu^n}{n!} e^{-\mu} \prod_{i=1}^{n} f(x_i; \boldsymbol{\theta}). \tag{VI.6}$$

This is often called the extended maximum likelihood (EML).

This formalism can be generalised to models that contain more than one species. For example, for a model with two species, signal and background, the total PDF can be written as

$$f(\boldsymbol{x}; \boldsymbol{\theta}) = \frac{N_S}{N_S + N_B} f_S(\boldsymbol{x}; \boldsymbol{\theta}) + \frac{N_B}{N_S + N_B} f_B(\boldsymbol{x}; \boldsymbol{\theta}), \tag{VI.7}$$

where N_S and N_B are the signal and background yields. Substituting the PDF given in Eq. VI.7 into the EML defined in Eq. VI.6, and taking $\mu = N_S + N_B$,

[2]Depending on the specifics of the analysis, it can more efficient to use a binned likelihood in which case the product runs over bins rather than data points. This approach is not the one chosen for this analysis and thus is not discussed further here.

one is able to extract the yields N_S and N_B from a fit to the data. This procedure can be easily generalised to an arbitrary number of species. For N_p species and N events in total, the negative log-likelihood is given by

$$-\ln \mathcal{L}(\boldsymbol{\theta}) = \sum_{i=1}^{N_p} N_i - \sum_{e=1}^{N} \ln \left(\sum_{i=1}^{N_p} N_i f_i(x_i; \boldsymbol{\theta}) \right). \tag{VI.8}$$

The robustness of the fit can be improved by using external information on some parameters. One possibility to include such information into the fitting framework is to simply fix the relevant fit parameters to their estimated values. However, this procedure does not take into account the uncertainties on the fixed parameters. An alternative to fixing these parameters is to use constraints. They are implemented as "penalties" to the likelihood, $\mathcal{L}^c(\boldsymbol{\theta})$, of the form

$$\mathcal{L}^c(\boldsymbol{\theta}) = \mathcal{L}(\boldsymbol{\theta}) \times \mathcal{C}(\boldsymbol{\theta}), \tag{VI.9}$$

where $\mathcal{C}(\boldsymbol{\theta})$ is the PDF of the constraint. Multiple constraints can be included in the likelihood by taking the product of their PDFs. A common choice for $\mathcal{C}(\boldsymbol{\theta})$ is to use a Gaussian distribution with mean equal to the expected value of the parameter and the width corresponding to the uncertainty on this value.

VI.2.2 Multivariate Analysis

Multivariate analysis (MVA) techniques are commonly used in many fields of research involving a large amount of data. They have a broad range of applications in high energy physics, for instance for calibration, particle identification, event reconstruction, etc. In the present analysis a multivariate approach is used to separate signal-like events from combinatorial-background-like ones (denoted simply background hereafter). Combinatorial background arises from random combinations of unrelated tracks and is the primary source of background in the decay modes studied here. It is first suppressed by loose preselection cuts, and then further reduced by a multivariate selection. A set of judiciously chosen variables are given as inputs to the MVA algorithm that combines them into a single output variable. A cut on this output variable can be used to separate background and signal. In the next paragraphs we address the choice of variables, the methods to evaluate the performance of the classifier and the technique used to obtain the best cut value.

There are two main steps in a MVA. First, the algorithm has to "learn" how to recognise the characteristics of the signal and the background from

a training stage. The approach used here is a classification method with supervised statistical learning. In other words, the training (and, later, test) samples provided to the MVA are identified as signal and background, so that the true result is known by the algorithm, which uses this information during the learning process. The choice of the training samples is very important since they must have similar properties to the signal and background we want to separate. Here, we use simulated events to describe the signal. The combinatorial background is expected to vary smoothly and so the right-hand sideband of the data is used. The classifier algorithm combines a set of judiciously chosen variables into a single output by analysing their behaviour in the background and signal training samples. To evaluate the performance of the training, it is applied to the test samples that are statistically independent from, but similar in origin to, the training ones. Many different classification algorithms are available, from simple linear cuts (e.g Fisher) to more sophisticated methods, such as neural networks. The method used in this analysis is based on a Boosted Decision Tree algorithm (BDT), and will be described in the following section.

Once the training is finished and validated, it is applied to the sample of interest. In that step, the MVA uses the classifier that was built during the training and applies it to each event in the data sample, generating the corresponding output variable. Events with an output value above a pre-defined threshold are retained. The best cut point is determined by optimising a figure of merit (FoM).

Boosted Decision Tree classifier

The choice of the classification algorithm has to be adapted to the problem. In this analysis, we use a boosted decision tree (BDT) classifier with a gradient boosting (BDTG). This type of classifier uses binary decision trees, with two possible outputs (yes/no decisions). A sketch of an individual tree is shown on Fig. VI.1. The training starts from a node called "root node", then splits the events into two subsets according to a cut on a specific variable (one single variable at a time). The usual splitting criterion is found by maximising the signal/background separation, but alternatives are also available such as the Gini index which is based on the signal purity[3]. The procedure is repeated until the whole tree is built, i.e. until a stop criterion is fulfilled. Each cut depends on the outcome of the previous one, and the phase space is divided into regions that are classified as signal or background. This classification is

[3]The Gini index is defined as $p(1-p)$ where p is the signal purity. The criterion for splitting is then by Gini(initial node) − [Gini(left node) + Gini(right node)]

done according to the fraction of events of each type that end up in a particular final node (a "leaf node").

The strong point of decision tree methods is that monotonic transformations of the variables do not affect the decision trees, so that it is possible to use certain functions of the variables. For example it can be useful to work with the logarithm of a variable rather than with the variable itself. Another advantage is that they are not prone to the "curse of dimensionality", which arises in some methods when the dimension of the phase space is large compared to the available data points. The performance of decision tree classifiers is not degraded by the addition of more input variables. Indeed, the use of variables by the algorithm depends on their separation power, such that one variable may be used at several nodes while others might not be used at all, so that "weak variables" are simply ignored by the algorithm and do not contribute.

However, an important drawback of methods based on BDTs is that they are very sensitive to the training samples and thus to overtraining. Overtraining happens when the classifier learns features from the training sample that are specific to that sample and are not general characteristics of the signal and background, for instance being statistical fluctuations. This results in a deterioration of the generic performance of the algorithm. Various options to overcome this problem exist, such as boosting, which is used in BDT algorithms. In this approach, the algorithm is performed multiple times on reweighted versions of the training samples. Then, a weighted average of the individual trees is takem to combine them into a single classifier. The reweighting of the training events is done in a way such that the events that were misclassified in a tree are given a higher weight in the next tree[4]. This procedure stabilises the output and is less prone to overtraining. Other techniques to address the problem of overtraining include the use of bootstrap or random samples.

Discriminating variables

The input variables should be chosen according to their separation power between signal and background. Good variables have usually a different behaviour (a different distribution) in the signal and background samples. Depending on the algorithm used by the classifier, the correlations between the variables can also be used in the learning process. Depending on the analysis, there may be additional requirements that the variables chosen should fulfil. For instance, variables showing correlations with the mass of the $B^0_{(s)}$ meson should not be used, since this could bias the results by sculpting the mass

[4] An overall normalisation is applied so that the sum of weights remains constant.

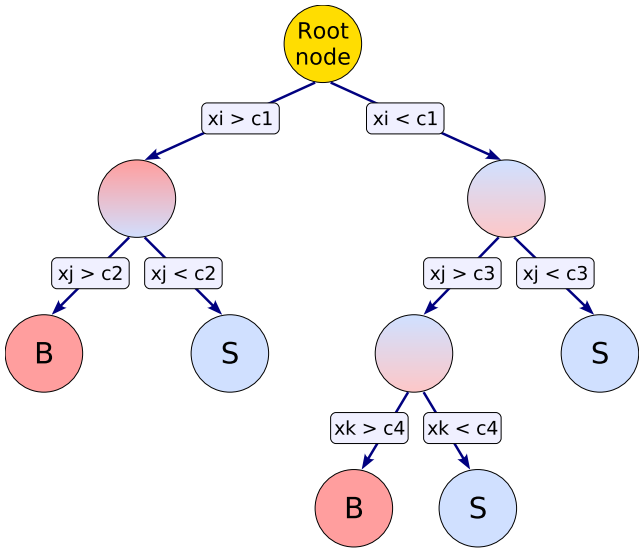


Figure VI.1: Schematic view of a decision tree [143]. In this example the procedure starts at the root node and the discriminating variable, x_i, that gives the best separation between signal and background is used to split the sample. This method is repeated, each time using the best variable, until a stop criterion is fulfilled. The final nodes ("leaves") are then classified as signal, S, or background, B, depending on the number of events from the signal and background training samples that are contained in the node.

spectrum[5]. In the same vein, when performing Dalitz plot analyses, variables that are correlated with the Dalitz plot and square Dalitz plane variables might also pose problems.

Practical aspects of MVA choice and training

There is no general, foolproof method to choose and implement an MVA. The choice of the algorithm and of the different variables has to be adapted to each situation. Usually, one proceeds by testing and evaluating different configurations. A way to compare different implementations is to construct their ROC curves (Receiver Operating Characteristic) when applied to the same test samples. The ROC curve describes the ability of the classifier to separate signal from background. The ROC curve can be constructed in several ways; usually, in high energy physics, one plots the background rejection as a function

[5]Not only linear but also non-linear correlations have to be properly checked.

of the signal efficiency. By construction, the signal efficiency decreases as the background rejection increases, and so a trade-off has to be made according to the ultimate goals of the analysis. (See Sec. VI.4.3 for the optimisation procedure used in this case.)

As mentioned previously, BDT algorithms are especially sensitive to overtraining. A way to check for potential overtraining is to superimpose the output variable distributions obtained with the training and test samples. Any significant disagreement between the curves (beyond statistical fluctuations) is a hint for overtraining. Overtraining can result in sub-optimal performance of the classifier when applied to data other than the training sample; the training signal sample will also give an overestimate of the efficiency.

VI.2.3 The ${}_s\mathcal{P}lot$ method

As described in Sec. VI.1, many of the events in the data samples are not signal but originate from one of the various sources of background. The ${}_s\mathcal{P}lot$ technique [144] is a tool to unfold the different contributions in a data sample. It allows distributions of so-called control variables to be reconstructed independently for each species without the need for prior knowledge of these variables. This unfolding is done via the use of discriminating variables, which are variables for which the distributions of all the sources are known (and different). Given a mixed sample of events that contains contributions from all species, these discriminating variables are used to infer the distributions of each individual species in the control variables on a statistical level. For this approach to be valid, the control and discriminating variables must not be correlated.

Let's consider a single discriminating variable y, whose behaviour is known for all the species N, and a control variable x, uncorrelated with y, and for which the individual distributions of the sources are unknown. For the present analysis, the discriminating variable is the invariant mass of the B meson candidate, obtained from the combination of its three daughters. A maximum likelihood fit is performed to extract the different components of the spectrum: to each species i is associated a PDF f_i and a yield N_i. From there it is possible, for each event e and species n, to compute a weight, referred to as an ${}_s\mathcal{W}eight$, that is defined as

$$ {}_s\mathcal{P}_n(y_e) = \frac{\sum_{j=1}^{N} V_{nj} f_j(y_e)}{\sum_{k=1}^{N} N_k f_k(y_e)}, \tag{VI.10} $$

where V_{nj} is the covariance matrix of the yields extracted from the likelihood

fit to the discriminating variable, defined as

$$V_{nj}^{-1} = \frac{\partial^2(-\ln \mathcal{L})}{\partial N_n \partial N_j}, \tag{VI.11}$$

where $\mathcal{L}$ is the likelihood function. defined in Eq. (VI.8).

The binned distribution of the control variable x, $\tilde{M}_n(\bar{x})$, is then given by

$$N_n \tilde{M}_n(\bar{x})\delta x \equiv \sum_{e \in \delta x} {}_s\mathcal{P}_n(y_e), \tag{VI.12}$$

where the x-bins are of width δx and are centred on $\bar{x}$, so that the sum runs over the events contained in each bin.

VI.2.4 The TISTOS method

The TISTOS method [145] is a method that allows the trigger efficiency to be determined from the data. In the introduction to this chapter, eq. (VI.2) gives the overall efficiency as a product of the efficiencies corresponding to the different steps of the selection. It is also possible to change the order of the terms and compute the trigger efficiency at the end, after all the selection steps are applied,

$$\epsilon^{\text{Tot}} = \epsilon^{\text{Acc}} \times \epsilon^{\text{Reco|Acc}} \times \epsilon^{\text{Sel|Reco}} \times \epsilon^{\text{Trig|Sel}}. \tag{VI.13}$$

Note that this equation gives exactly the same result as Eq. (VI.2) (since events must still pass every step to be selected); the different ordering of the terms is just a matter of convenience for the computation of the relative efficiencies.

The relative trigger efficiency is then computed as the ratio of the number of particles passing the trigger requirements, $N_{\text{Trig|Sel}}$, with respect to number of particles passing the whole selection other than the trigger, N_{Sel}. Thus,

$$\epsilon_{\text{Trig|Sel}} = \frac{N_{\text{Trig|Sel}}}{N_{\text{Sel}}}. \tag{VI.14}$$

The difficulty here is that only the data events that passed the hardware trigger are kept, so the number of events that entered the LHCb acceptance but were not retained by the trigger is unknown, and therefore we do not have access to N_{Sel}. The TISTOS method offers a way to estimate of N_{Sel} and $\epsilon_{\text{Trig|Sel}}$ under some assumptions.

As described in Sec. V.5, events selected by the trigger may be classified into several categories that are not mutually exclusive. In the TISTOS method, only the TOS (Trigger On Signal) and TIS (Trigger Independent of Signal)

categories are used. Events are classified as TOS if a positive trigger decision is generated by signal objects, and as TIS if the trigger is fired by objects from the rest of the event. An event may be both TIS and TOS[6]. Assuming that the TIS and TOS trigger decisions are not correlated—in other words, assuming that the signal candidates are uncorrelated with the rest of the event—the TIS efficiency of the subsample of the triggered events is the same as that of the whole sample. Under this hypothesis, the TIS efficiency on the whole sample of selected events is the same as the TIS efficiency measured on the TOS sample. That is to say,

$$\epsilon_{TIS|Sel} = \frac{N_{TISTOS|Sel}}{N_{TOS|Sel}}, \tag{VI.15}$$

where N_{TISTOS} is the number of events that have both TIS and TOS positive decisions. The TOS efficiency can be determined in the same way, replacing the denominator of Eq. VI.15 by $N_{TIS|Sel}$. The trigger efficiency can then be expressed in terms of quantities that can be obtained from data:

$$\epsilon_{\text{Trig|Sel}} = N_{\text{Trig|Sel}} \times \frac{N_{TISTOS|Sel}}{N_{TOS|Sel} \times N_{TIS|Sel}}. \tag{VI.16}$$

Uncertainty on the trigger efficiency obtained with the `TISTOS` method

The quantities involved in Eq. (VI.16) are not statistically independent from each other—for instance, $N_{TOS|Sel}$ and $N_{TIS|Sel}$ share a subset of the events. Comparing Eq. (VI.14) and (VI.16), we obtain

$$N_{\text{Sel}} = \frac{N_{TOS|Sel} \times N_{TIS|Sel}}{N_{TISTOS|Sel}}. \tag{VI.17}$$

This equation can be rewritten by separating the different terms into independent quantities,

$$N_{\text{Sel}} = \frac{(b+d)(c+d)}{d}, \tag{VI.18}$$

where $d = N_{TISTOS|Sel}$, $b = N_{TIS|Sel} - N_{TISTOS|Sel}$, the non-overlapping part of $N_{TIS|Sel}$ and similarly $c = N_{TOS|Sel} - N_{TISTOS|Sel}$. The uncertainty on N_{Sel}

[6]It is possible for events to be neither TIS nor TOS yet still fire the trigger. This is uncommon and usually associated with misreconstruction in the trigger. Such events will be discarded in the trigger selection (Sec. VI.3.1) and are not considered further here.

can then be computed from

$$\begin{aligned}\sigma^2_{N_{\text{Sel}}} &= \left(\frac{\partial N_{\text{Sel}}}{\partial b}\right)^2 \sigma_b^2 + \left(\frac{\partial N_{\text{Sel}}}{\partial c}\right)^2 \sigma_c^2 + \left(\frac{\partial N_{\text{Sel}}}{\partial d}\right)^2 \sigma_d^2 \\ &= \left(\frac{c+d}{d}\right)^2 b + \left(\frac{b+d}{d}\right)^2 c + \left(1 - \frac{bc}{d^2}\right)^2 d.\end{aligned} \tag{VI.19}$$

Similarly, we can express N_{Sel} into two independent contributions, as the sum of the events that passed the trigger, p with the events that did not, n, $N_{\text{Sel}} = p + n$, and $p = N_{\text{Trig|Sel}}$. The trigger efficiency is then

$$\epsilon_{\text{Trig|Sel}} = \frac{p}{p+n}. \tag{VI.20}$$

The error on $\epsilon_{\text{Trig|Sel}}$ is therefore given by

$$\sigma^2_{\epsilon_{\text{Trig|Sel}}} = \left(\frac{n}{(p+n)^2}\right)^2 p + \left(\frac{p}{(p+n)^2}\right)^2 \left(\sigma^2_{N_{\text{Sel}}} - p\right) \tag{VI.21}$$

This method relies heavily on the assumption that the TIS efficiency and the TOS efficiency are uncorrelated. In fact, this is not quite true since the b-quarks are usually produced in $b\bar{b}$ pairs, and the kinematics of the two b-hadrons are correlated (such that, for example, a high-p_{T} B with a better-than-average TOS efficiency is more likely to be accompanied by a high-p_{T} $\overline{B}$ so that the B also has a better-than-average TIS efficiency). This effect can be taken into account by measuring the trigger efficiency in kinematic bins. However, in this analysis the systematic effect related to this correlation will be neglected because it applies to both the signal and normalisation modes and therefore approximately cancels in the ratio of branching fractions.

VI.3 Dataset, Montecarlo samples, and reconstruction

The dataset used in this analysis consists of a data sample recorded by the LHCb experiment in 2011 and 2012, referred in the following as run I data, and a sample from 2016, referred to as run II data. Together, the data used correspond to an integrated luminosity of $4.67\,\text{fb}^{-1}$. As described in Sec. V, the data-taking conditions evolved over the years and the centre-of-mass energy of the collision increased from 7 TeV in 2011 to 8 TeV in 2012 and 13 TeV for the whole of run II (2015 to 2018).

Each of the modes considered here contains a neutral K^0_S, which is reconstructed from two tracks of opposite charge identified as pions. As explained in part V.3.6, different types of tracks are used in the reconstruction. The tracks used here for K^0_S reconstruction are `Downstream` tracks (not using any hits in the VELO) and the `Long` tracks (hits in the VELO and in the T-stations). The K^0_S is then reconstructed either from two `Downstream` pion tracks (DD category) or from two `Long` pion tracks (LL). The K^0_S candidates reconstructed in the LL category benefit from the information of the VELO and have better mass, momentum, and vertex resolutions than the ones reconstructed in the DD category. However, there are approximately twice as many candidates in the DD category, which motivates its inclusion.

The Monte Carlo (MC) samples are generated with conditions as similar as possible to the data-taking conditions. Different samples were generated for each combination of decay mode, year and magnet polarity. When generating three-body events, it is necessary to assume a model for their distribution in the phase space. For each sample used here, the events are generated with a uniform distribution in the square Dalitz plane. The samples consist of approximately 2 million events generated per year and per mode for run I and 4 million events per year and per mode for run II (summing the two magnet polarities).

The trigger configuration (TCK) often changes during data-taking. Generally, the MC samples are simulated using a single trigger configuration that is representative of conditions in the corresponding data-taking year—but an exception is made for 2012. Before mid-2012, the inclusive topological HLT2 lines used in this analysis (see next section) were not able to include K^0_S with `Downstream` daughters in the trigger decision. During the technical stop at the end of June 2012, major changes were made to these lines. Among other improvements, this allowed them to include the DD category for K^0_S and Λ candidates, resulting in an large increase in efficiency for the present analysis. Consequently, the 2012 data taking year is split here into two periods, 2012a and 2012b, corresponding to the data taken before and after this change. Accordingly, two different sets of MC samples were generated with trigger conditions corresponding to 2012a and 2012b data taking periods.

VI.3.1 Trigger selection

At the hardware trigger (L0) level, the event either is required to be fired by a hadron belonging to the signal (trigger on signal, TOS) or that any L0 physics trigger line has a positive decision triggered by any other track in the event (trigger independent of signal, TIS). The physics TIS triggers include decisions

on particles such as muons, electrons, photons and hadrons. The exact list of triggers and their configuration varies with the data taking conditions and the year. The complete list of trigger lines is given in appendix B.1.

At the software level, the HLT1 lines select all the events that passed the L0 trigger, with some additional requirements to retain only high transverse momentum tracks, well displaced from the primary vertices and with a good fit quality. At HLT2 stage, the full event reconstruction is performed in order to refine the selection. To this end, inclusive topological trigger lines are employed [146]. These lines are designed to trigger on partially reconstructed b-hadron decays with, at least, two charged daughters. Since the selection must remain generic (i.e. we don't want to select a specific decay but rather all possible B decays), cuts on variables such as the mass of the candidate or its impact parameter with respect to the primary vertex must remain very loose. The selection is performed in two steps, first a list of two-body "proto-candidates" is made from input particles that have the same "best" primary vertex (PV) — the PV for which the impact parameter (IP) is the smallest —, a distance of closest approach (DOCA) smaller than 2 mm, and whose vertex is downstream its best PV. In addition, very loose cuts on the invariant mass (less than 7 GeV) and the flight-distance χ^2 (larger than 100) are required. From there, n-body topological lines are constructed by further filtering the two-body proto-candidates. To select the candidates, the HLT2 topological lines use a multivariate classifier — called a bonsai boosted decision tree (BBDT) [147] — based on discretised kinematic and topological input variables. In this analysis, the HLT2 topological lines selecting two, three and four candidates are used. In 2011, similar lines using a simple cut-based selection were also available and were then abandoned in the following years. These lines are also included in the 2011 selection, along with the BBDT ones.

VI.3.2 Stripping selection

On top of the trigger, another filtering step, called stripping, is applied centrally in order to further reduce the amount of data written on disc. Similarly to the trigger step, the stripping is organised in different lines, that are specific to a decay, or a class of decays. The lines used in this analysis select candidates that are compatible with the $B^0_{(s)} \to K^0_S h^+ h'^-$ hypothesis. These lines have a similar structure for run I and run II although the details of the cuts differ between the two runs. The B candidates are formed by combining a K^0_S candidate along with two oppositely charged tracks, denoted h_1 and h_2 (or $h^\pm$) in the following. Prior to any processing, some sanity cuts are applied to reject events that are not of interest and save CPU time. The selected events are required to have

less than 250 `Long` tracks and, at least one primary vertex (PV). In a first step, all the charged tracks are assumed to be pions and they are subsequently refitted with different mass hypotheses.

The daughter particles are drawn from so-called "containers", which are lists of candidates that passed a certain set of cuts. The K^0_S candidate is formed by combining two oppositely charged tracks: the daughters of the DD K^0_S are taken from a `StdNoPIDsDownPions`, which contains a list of candidates compatible with `Downstream` pions, whereas for the LL case the candidates are taken from `StdLoosePions`, which lists `Long` tracks compatible with pions.

One of the most important difference in the stripping selection between the two runs is that, for run I the $h^\pm$ candidates are taken from `StdLoosePions` list whereas in run II `StdAllNoPIDsPions` is used. The container used in run I requires the tracks to have a transverse momentum greater than 250 MeV ($p_T(h^\pm) > 250$ MeV) and a minimum impact parameter χ^2 with respect to the PV bigger than four ($\min\chi^2_{IP}$ $(h^\pm) > 4$), whereas the run II container does not have these requirements. The motivation for this change is that the candidates with a low p_T are situated at the corners of the Dalitz plane: by increasing the number of such candidates we expect an improvement in the resolution on the edges of the Dalitz plane. While these events are particularly relevant for Dalitz plots analysis, they represent a tiny fraction of the total signal events and are not crucial for the branching fraction measurements. Another impact of the loosening of the stripping cuts in run II is an increase in the amount of background selected along with the signal. To reduce a part of this background, cuts on the HLT1 and HLT2 trigger decisions have been added.

In order to save CPU time the stripping cuts are applied in different steps. Before the mother candidates are formed, loose kinematical and topological "daughter cuts" are applied to the tracks. Then the mother candidates are formed simply by four-momentum addition. At this point, "combination cuts" are applied to the three daughters to reduce the number of candidates prior to the full vertex fit. Finally, "mother cuts" further refine the quality of the candidates after the fit. To avoid biasing the Dalitz plane distribution most of the cuts use topological variables such as the χ^2 of the vertex fit, χ^2_{vtx}, the flight distance χ^2 (χ^2 distance from the PV), χ^2_{FD}, the minimum impact parameter χ^2 of a track with respect to the related PV, $\min\chi^2_{IP}$, the distance of closest approach, DOCA, and the cosine of the direction angle (DIRA). The list of the stripping cuts applied in run II for both DD and LL categories are given in Table III, and the corresponding table for run I is given in appendix B.1.

After the stripping cuts are applied, the selected charged candidates initially considered as pions are re-fitted using the Decay Tree Fitter (DTF) package [148]. This algorithm, based on a Kalman filter, allows to change the mass hypothesis

of the candidate and re-fit the whole decay chain. During this procedure, the quantities related to the lifetimes of the particles, the vertex coordinates and vertex fit quality are recalculated. This package is also used to re-fit the candidates using constraints, such as fixing some of the particle masses to their nominal value or forcing the B candidate to originate from the best PV. Usually, for Dalitz plot analysis, the invariant mass variables that are used are obtained by fixing all their masses (the daughters as well as the mother particle masses), this ensures that all the events are inside the kinematic boundaries. In this analysis the B mass is obviously not constrained, but constraints on the PV and on the K^0_S mass are used in order to improve the B mass resolution. The DTF package is also used to re-fit the MC samples with different mass hypotheses in order to study the crossfeed background.

Table I: Stripping requirements for run II. As the stripping is applied in different steps, the same cut can appear multiple times. When this happens, only the last occurrence of the cut, the tightest version, is listed in this table.

Cut step	Candidate(s)	`StrippingB2KShh_DD_Run2_OS_Line`	`StrippingB2KShh_LL_Run2_OS_Line`

TriggerCuts	HLT1	`Hlt1TrackMVADecision`
		`Hlt1TwoTrackMVADecision`
		`Hlt1IncPhiDecision`(Hlt1PhiIncPhiDecision)
			idem
			idem
			idem

	HLT2	`Hlt2Topo{2,3,4}BodyDecision` `Hlt2IncPhiDecision`(Hlt2PhiIncPhiDecision)	idem idem
DaughterCuts	K^0_S	$p(K^0_S) > 6000.0\,\text{MeV}$ $\lvert m_{\pi^+\pi^-} - m^{\text{PDG}}_{K^0_S}\rvert < 30.0\,\text{MeV}$ $\chi^2_{\text{vtx}}(K^0_S) < 12.0$ $\chi^2_{\text{FD}}(K^0_S) > 50.0$	- $\lvert m_{\pi^+\pi^-} - m^{\text{PDG}}_{K^0_S}\rvert < 20.0\,\text{MeV}$ idem $\chi^2_{\text{FD}}(K^0_S) > 80.0$
	K^0_S daughters	$p(\pi^\pm) > 2.0\,\text{GeV}$ $\min\chi^2_{\text{IP}}\,(\pi^\pm) > 4$ - DOCA χ^2 of K^0_S daughters< 25 track $\chi^2/\text{ndf} < 4.0$ -	idem $\min\chi^2_{\text{IP}}\,(\pi^\pm) > 9$ $p_{\text{T}}(\pi^\pm) > 250\,\text{MeV}$ idem idem track ghost probability < 0.5
	$h_{1,2}$	track $\chi^2/\text{ndf} < 4.0$ track ghost probability < 0.5	idem idem
CombinationCut		$p_{\text{T}}(B) > 1000.0\,\text{MeV}$ $p_{\text{T}}(K^0_S) + p_{\text{T}}(h^+) + p_{\text{T}}(h^-) > 4200.0\,\text{MeV}$ at least 2 daughters with $p_{\text{T}} > 800\,\text{MeV}$ $(4000 < m_{K^0_S h^+ h'^-} < 6200)\,\text{MeV}$ DOCA χ^2 between pairs of daughters < 25	idem $p_{\text{T}}(K^0_S) + p_{\text{T}}(h^+) + p_{\text{T}}(h^-) > 3000.0\,\text{MeV}$ idem idem idem
MotherCut		$p_{\text{T}}(B) > 1500.0\,\text{MeV}$ $\chi^2_{\text{vtx}}(B) < 12.0$ DIRA$(B)> 0.999$ $\chi^2_{\text{FD}}(B) > 5$ $Z_{\text{vtx}}(K^0_S) - Z_{\text{vtx}}(B) > 15.0\,\text{mm}$ $\chi^2_{\text{IP}}\,(\pi^\pm) < 6.0$ sum of the χ^2_{IP} of the daughters w.r.t their PV > 50.0	idem idem idem idem idem - idem

VI.4 Selection

VI.4.1 Preselection

To reduce an important fraction of the the background prior to the multivariate analysis, high-signal efficiency cuts are applied on top of the stripping. To keep the variation of the efficiency over the Dalitz plane as flat as possible, most of the preselection is based on topological variables and loose cuts on the momenta of the B-meson daughters. A clearly isolated B vertex is required by comparing the χ^2 of the vertex formed with the signal candidate tracks and the χ^2 of the vertex formed by adding, in turn, each other Long track in the event (`B_STRIP_VTXISOCHI2ONETRACK`). To remove the K^0_S candidates reconstructed upstream to the B vertex, a separation between the two vertices larger than 30 mm is required in the direction of the beam. Tracks compatible with a muon are removed. Fiducial cuts are applied by constraining the momentum of the $h^\pm$ tracks to be in a range where the PID information obtained form the RICH detectors is reliable. As mentioned in before, in section VI.3.2, some of the cuts present in the run I stripping lines have loosen in the run II stripping lines. This resulted in a important increase in the volume of data after the stripping. The computing resources and the disc space needed to produce and save the nTuples being very large, it was decided to put back some of these cuts so that the nTuples have a manageable size. The cuts on the transverse momentum of the tracks of the charged daughters and on the $\min\chi^2_{\rm IP}$ have then been reintroduced. However, it is important to stress that, in the case one would like to use a similar selection to perform a Dalitz plot analysis, it may be worth to reconsider loosening these cuts. The preselection cuts are summarised in table II, and their acceptance rate, computed on MC, are given in table III.

Table II: Preselection cuts applied on top of the stripping.

Preselection cut	Description
$\chi^2_{\rm IsoVtx}(B) > 4$	B vertex isolation variable
$Z_{\rm vtx}(K^0_S) - Z_{\rm vtx}(B) > 30$	K^0_S vertex separation w.r.t. the B vertex
`h{1,2}_isMuon == 0`	Reject $h^\pm$ candidates compatible with the muon hypothesis
$3000 \leq p\ (h^\pm) \leq 100000$	Fiducial cut
$\min\chi^2_{\rm IP}\ (h^\pm) > 4$	Minimum IP χ^2 of the charged daughters with respect to the related PV
$p_{\rm T}(h^\pm) > 250$ MeV	Minimum transverse momentum of the charged daughters.

Table III: Percentage of number of events selected by each preselection cut, computed as the ratio between the number of events after the given cut and the number of events before the cut (i.e after trigger and stripping selections). This table is a comparison of the rates obtained in 2011 and 2016 $B^0 \to K^0_S\pi^+\pi^-$ MC samples. Since the cuts on $\min\chi^2_{\rm IP}$ and $p_{\rm T}$ ($h^\pm$) are included in the run I stripping, the rates for 2016 are computed with respect to the number of events after applying these cuts, so that the 2011 and 2016 selections shown here contain similar cuts.

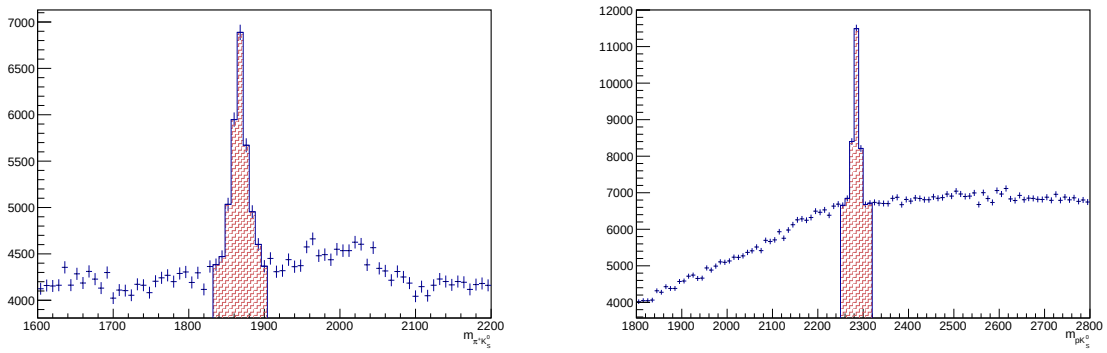
Preselection cut	2011		2016	
	LL [%]	DD [%]	LL [%]	DD [%]
$\chi^2_{\rm IsoVtx}(B) > 4$	90.2	93.0	86.5	87.99
$Z_{\rm vtx}(K^0_S) - Z_{\rm vtx}(B) > 30$	84.8	100	92.6	100
`h{1,2}_isMuon == 0`	92.3	95.6	90.0	91.55
$3000 \leq p\ (h^\pm) \leq 100000$	86.6	87.1	85.8	86.55
TOTAL	60.7	77.4	61.3	69.4

VI.4.2 Peaking backgrounds

The different $K^0_S h^+ h'^-$ final states can be reached via intermediate states containing charmed particles, such as $B^0 \to D^+\pi^-$ with $D^+ \to \pi^+ K^0_S$. These states have a long lifetime compared to the other structures in the decay, they thus do not interfere and are considered as background. These intermediate decays appear in the two-body invariant mass spectra of pairs of final state particles as peaking structures and are vetoed. Decays with one misidentified particle are also taken into account, as for instance $\Lambda_c^+ \to pK^0_S$ with the proton misidentified as a pion or a kaon. The different intermediate states are reconstructed with the corresponding mass hypothesis using the DTF package and vetoed using $|m^{\rm reco} - m^{\rm PDG}| < 30$ MeV/c for $D^\pm$, D^0, $D_s^\pm$ and Λ_c^+ and $|m^{\rm reco} - m^{\rm PDG}| < 48$ MeV/c for the charmonium states, χ_{c0} and J/ψ. The list of all the considered vetoed contributions is:

- $D^\pm \to \pi^\pm K^0_S,\ K^\pm K^0_S$
- $D_s^\pm \to \pi^\pm K^0_S,\ K^\pm K^0_S$
- $D^0 \to \pi^+\pi^-,\ K^+K^-,\ K^\pm\pi^\mp$
- $\Lambda_c^+ \to pK^0_S$, $\overline{\Lambda}_c^- \to \overline{p}K^0_S$
- $J/\psi \to K^+K^-,\ \pi^+\pi^-, \mu^+\mu^-$
- $\chi_{c0} \to K^+K^-,\ \pi^+\pi^-, \mu^+\mu^-$

Notice that, the probability to actually have muons at this stage of the selection is greatly reduced by the cut on the `isMuon` variable in the preselection step. Decays with two misidentified particles are not taken into account, so that, for example, the veto on $J/\psi \to K^+K^-$ is not considered in the $K^0_S\pi^+\pi^-$ spectrum.



(a) (b)

Figure VI.2: Example of vetoed peaking background in $B^0 \to K^0_S\pi^+\pi^-$ DD sample 2016, after trigger, stripping and preselection cuts. (a) The shaded region corresponds to the veto applied on $D^+ \to \pi^+K^0_S$ in the $\pi^+K^0_S$ invariant mass. (b) The shaded region corresponds to the veto applied on $\Lambda_c^+ \to pK^0_S$ in the $\pi^+K^0_S$ invariant mass with the proton misreconstructed as a pion.

VI.4.3 Multivariate Analysis

A multivariate analysis is performed after the preselection to further reduce the amount of combinatorial background. It is implemented using the ROOT TMVA package [143, 149], which provides a framework for the training and application phases and contains various classifiers that can be customised to improve the performances.

Training and test samples

The signal, training and test samples, consist in MC truth-matched events and the background is taken from the right-hand side band of the data ($m_{K^0_S\pi^+\pi^-} > 5425\,\text{MeV}$), which is assumed to be dominated by combinatorial background. The left-hand side band (LHSB) of the data is not used because of the important contribution from partially reconstructed backgrounds in this region.

The same selection is applied to both signal and background samples. This selection consists in trigger, stripping, preselection cuts, and the vetoes described in sections VI.3 and VI.4, and PID cuts taken from the previous analysis. To remove peaking backgrounds that can mislead the classifier, vetoes on $\Lambda_b^0 \to p\, K_S^0\, \pi$ are also applied.

To ensure that there is no bias in the training, the signal and data samples are randomly split into two sub-samples with approximately the same number of events. One sample is used for the training phase and the other one for the validation. Two separated MVAs are then performed, using each sample, in turn, for the training and the testing phase. For the analysis to be reproducible, the samples are split according to a pseudo-random number ($r \in [0, 1]$) computed for each event the following way,

$$r = ((134 \times \texttt{eventNumber} + \texttt{runNumber})\%531241)/531241, \qquad \text{(VI.22)}$$

where `runNumber` is an number proper to each run and `eventNumber` identifies each event within a given run. The events with $r > 0.5$ are used to train the first MVA (MVA1) while the other half of the sample (events with $r \leq 0.5$) is used for the validation. Reciprocally, the second MVA (MVA2) is trained with events with $r \leq 0.5$ and events with $r > 0.5$ are used for the testing.

Since the six $K_S^0 h^+ h'^-$ decay modes have rather similar topological properties, it has been decided to use exclusively $B^0 \to K_S^0 \pi^+ \pi^-$ decay mode for the training and apply the same training to all the decay modes. A different training is performed for each K_S^0 reconstruction category and year of data taking. The size of the samples in each category is given in Table IV.

Table IV: Number of events used in the signal and background samples to train (test) and validate MVA1. The yields for MVA2 are exactly the same after swapping training and test samples.

Year	DD		LL	
	Sig	Bkg	Sig	Bkg
2011	6169 (6101)	2230 (2174)	2905 (2819)	876 (852)
2012a	7480 (7361)	1692 (1643)	2530 (2449)	979 (950)
2012b	5314 (5365)	6722 (6570)	2549 (2533)	1501 (1513)
2016	29154 (29219)	114582 (114061)	11710 (11910)	50742 (50262)

Discriminating variables

The choice of variables in the current analysis is based on the previous one. For the reasons exposed in section VI.2, the input variables should not be correlated with the B mass or the Dalitz plot variables. In this context, when associated to a final state particle, only topological variables are considered, such as criteria on the separation and the quality of the vertices, the direction angle (DIRA) and the flight distance of the B and K^0_S mesons with respect to the primary vertex (PV). The impact parameters of the B daughters are known to be correlated with the Dalitz plot variables when considered separately, however, their sum does not show such correlations. Kinematic variables related to the transverse momentum an the pseudo rapidity of the B meson are also added. Because simulated events are used to describe the signal, only MC variables that are in good agreement with the data are considered. A first model containing the same variables as in the previous model was first tested, and is referred as "baseline model" here. Then, the addition of new variables was also considered, especially the so-called "cone" variables that are obtained by computing the asymmetry between the variable associated to the desired track and the value of this variable summing over all the tracks that are contained in a cone around the particle in question. For example, the p_T asymmetry of the B-candidate is defined as

$$p_T^{\text{asym}} = \frac{p_T(B) - p_T(\text{cone})}{p_T(B) + p_T(\text{cone})} \tag{VI.23}$$

where $p_T(\text{cone})$ refers to the transverse component of the sum of the momenta of all the particles inside a cone around the B-candidate in the $\eta - \phi$ plane. The radius of this cone is taken to be 1.5 rad. Cone variables of both B and K^0_S candidates are uses. A complete list of the variables included in the MVA is presented in Table V. The distributions for signal and background variables are shown in Fig VI.3 for 2016 LL sample, similar tendencies are seen in the DD sample and the other data taking years. From these plots it appears that variables such as the direction angle (`logBDIRA`), the sum over the χ^2_{IP} of the daughters ($\log \sum \chi^2_{\text{IP}}$) and the flight distance χ^2 of the B ($\log \chi^2_{\text{FD}}(B)$) have a good separating power. In contrast, the cone variables seem to have a rather poor separating power. However, as can be seen in the correlation matrices shown in Fig. VI.5, these variables are very correlated. It was observed that the addition of the cone variables one at a time in the MVA does not improve the performance, it is rather the correlations between them that are used by the BDT to learn the specificities of the signal and the background. These matrices represent the linear correlations only, other type of correlations that

are not represented here may play a role in the learning process. The model with all the variables included is referred as "final model". The list of variables differs slightly between the LL and DD samples: the LL sample uses one more variable, the flight distance of the K^0_S which is not included in the DD sample. Similarly the sum over the χ^2_{IP} of the daughters includes the K^0_S in the LL case whereas it is not included in DD case.

The ROC curves for the baseline model and the final one are given in Fig. VI.6. It appears on these curves that the addition of the cone variables improve significantly the classifier performance, and so the final model is chosen as the definitive model.

Table V: MVA input variables. The baseline model corresponds to the variables used in the previous analysis. In the current study, seven additional variables are also included.

Variable	name	Description	Sample
Baseline model	p_T (B)	B transverse momentum	DD and LL
	$\eta(B)$	B pseudorapidity	DD and LL
	χ^2_{IP} (B)	IP χ^2 of the B w.r.t its PV	DD and LL
	χ^2_{FD} (B)	flight distance χ^2 of the B	DD and LL
	`B_DIRA`	cosine of the B direction angle	DD and LL
	χ^2_{vtx} (B)	χ^2_{vtx} of the B	DD and LL
	$\chi^2_{IsoVtx}(B)$	B vertex isolation variable	DD and LL
	$p_T^{asym}(B)$	p_T asymmetry for a cone of radius 1.5 rad around the B-candidate in the η-ϕ plane	DD and LL
	χ^2_{IP} (h^+)+χ^2_{IP} (h^-)	sum of the χ^2_{IP} of $h^\pm$ w.r.t their PV	DD
	χ^2_{FD} (K^0_S)	χ^2_{FD} of K^0_S (LL sample only)	LL
Additional variables	χ^2_{IP} (h^+)+χ^2_{IP} (h^-)+χ^2_{IP} (K^0_S)	sum of the χ^2_{IP} of the daughters w.r.t their PV	LL
	$p_T^{asym}(K^0_S)$	cone p_T asymmetry for the K^0_S candidate	DD and LL
	$p^{asym}(B)$	cone p asymmetry for the B candidate	DD and LL
	$p^{asym}(K^0_S)$	cone p asymmetry for the K^0_S candidate	DD and LL
	$\Delta\eta^{asym}(B)$	cone $\Delta\eta$ asymmetry for the B candidate	DD and LL
	$\Delta\eta^{asym}(K^0_S)$	cone $\Delta\eta$ asymmetry for the K^0_S candidate	DD and LL
	$\Delta\phi^{asym}(B)$	cone $\Delta\phi$ asymmetry for the B candidate	DD and LL
	$\Delta\phi^{asym}(K^0_S)$	cone $\Delta\phi$ asymmetry for the K^0_S candidate	DD and LL

Optimisation of the BDT cut

Once the classifier is trained it is applied to both the data and MC samples. The optimal cut point on the BTDG output is obtained by maximising a figure of merit (FoM). The choice of FoM is guided by the goals of the analysis and the observables of interest. For the branching fraction measurements of the observed modes, we chose to use a FoM that maximises the significance of the

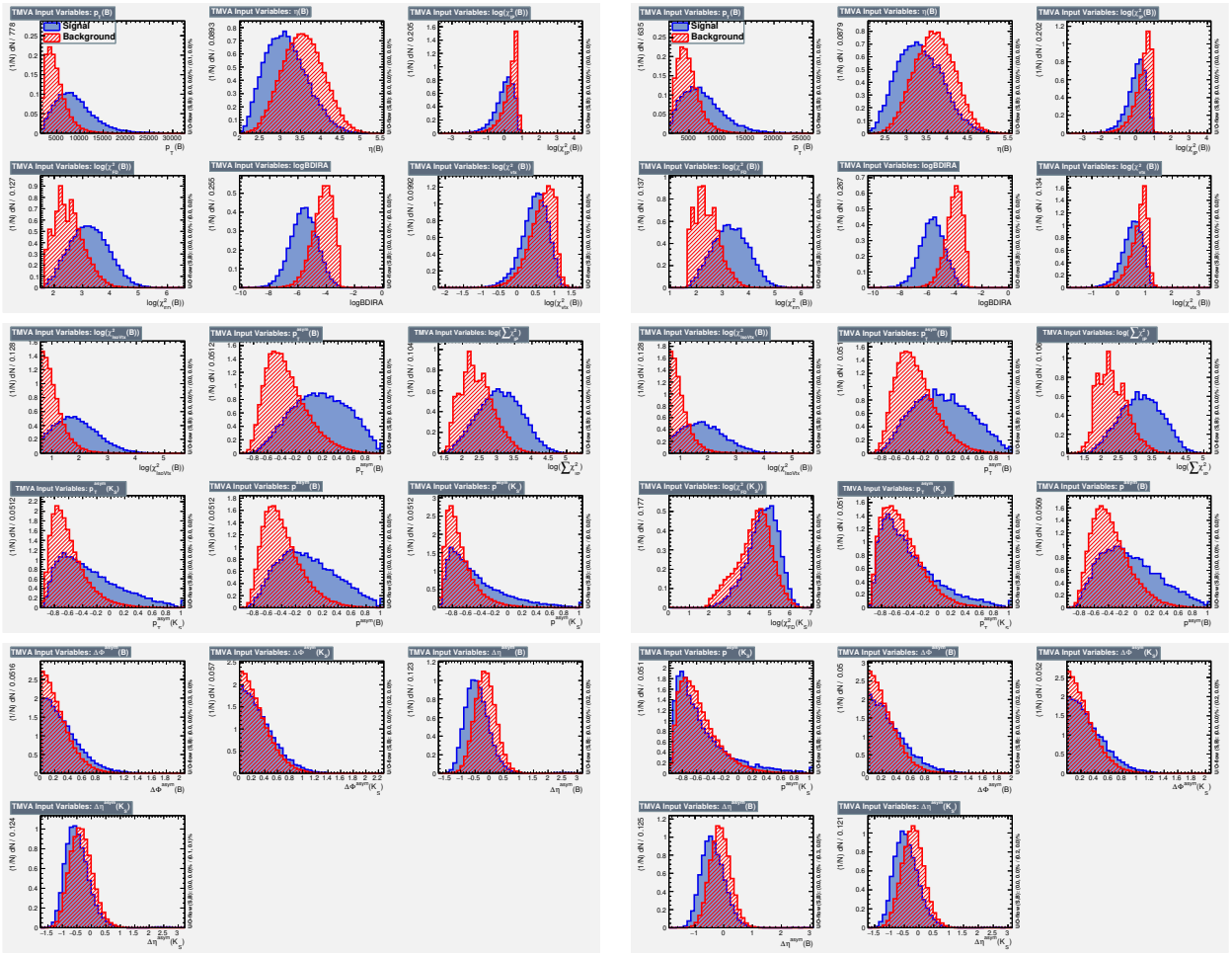


Figure VI.3: Distributions for the input signal (blue) and background (red) MVA variables for 2016 DD (left column) and LL (right column) samples.

signal with respect to the background

$$\frac{S}{\sqrt{S+B}}, \tag{VI.24}$$

where the number of signal events S is obtained by the product of the expected number of signal events S_0 and the efficiency of the BDT cut on MC, $S = S_0 \times \epsilon^{\text{MVA}}_{\text{MC}}$. The number of expected signal events S_0 is estimated from

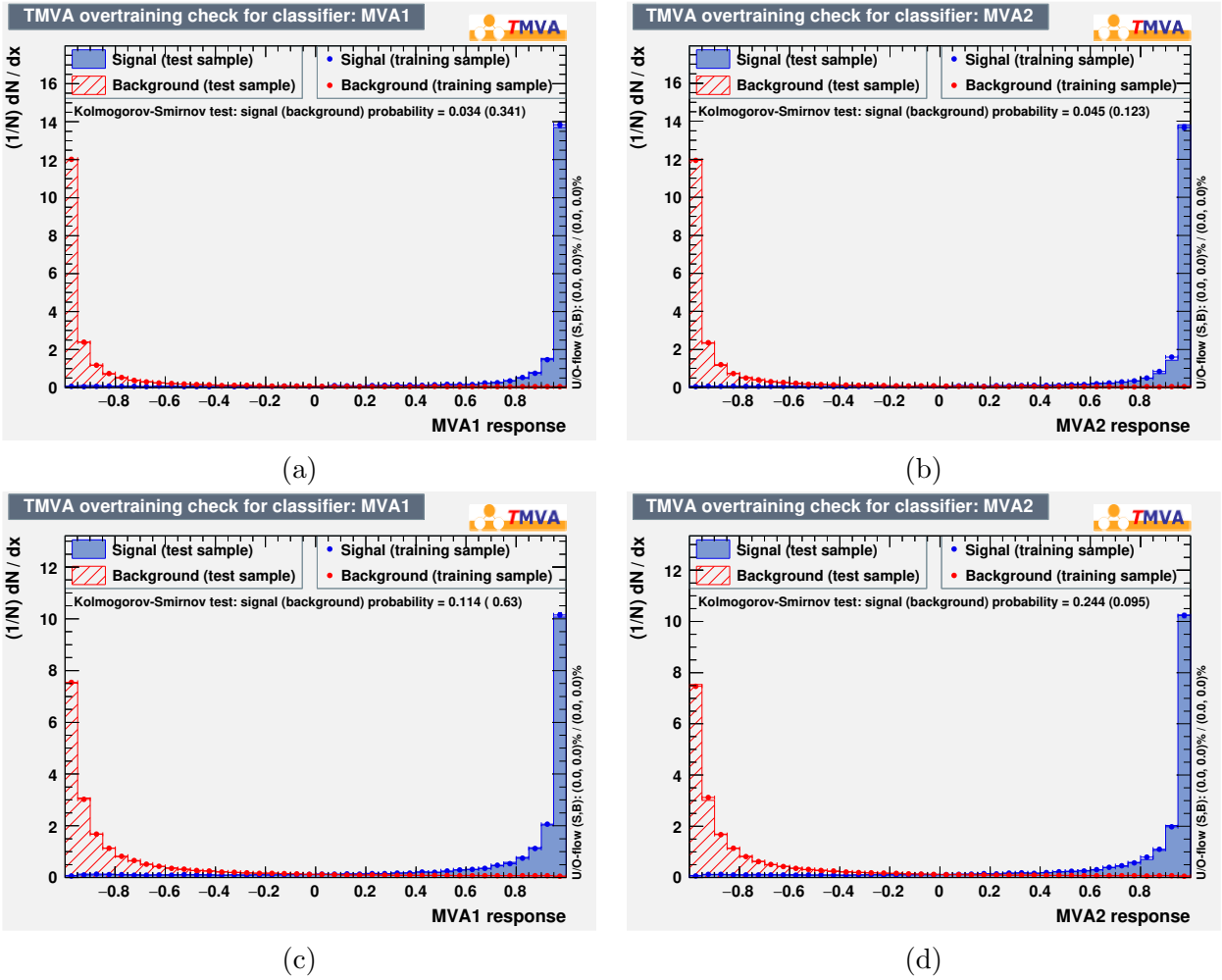


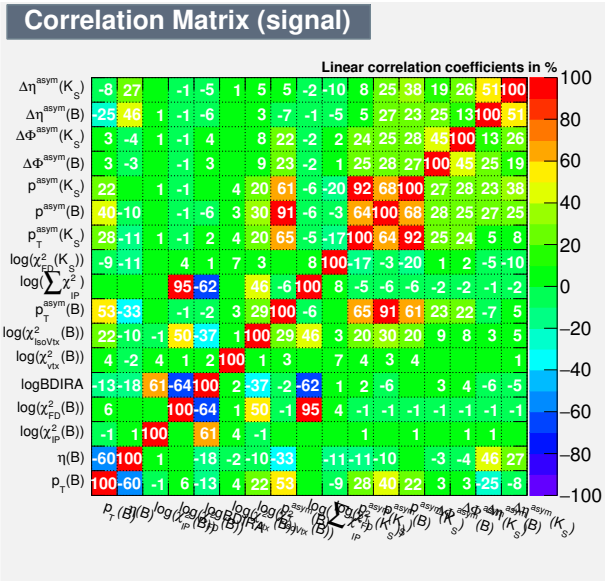
Figure VI.4: Overtraining tests for 2016 with the corresponding Kolmogorov-Smirnov test. The first row, (a) and (b), corresponds to the LL sample and the second row, (c) and (d), to the DD sample.

known branching fractions of the modes

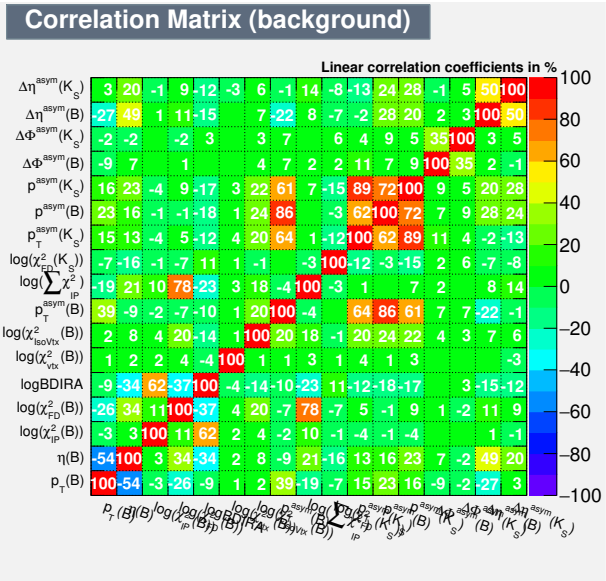
$$S_0(B^0_{(s)} \to K^0_S h^+ h'^-) = 2 \times \mathcal{L} \times \sigma(b\bar{b}) \times \mathcal{B}(B^0_{(s)} \to K^0_S h^+ h'^-) \times \mathcal{B}(K^0_S \to \pi^+\pi^-) \times f_{d/s} \times \epsilon^{\mathrm{MC}}, \tag{VI.25}$$

where $\mathcal{L}$ is the integrated luminosity of the corresponding data taking period, $\sigma(b\bar{b})$ is the $b\bar{b}$-cross section[7] measured by LHCb in [150], $f_{d/s}$ corresponds

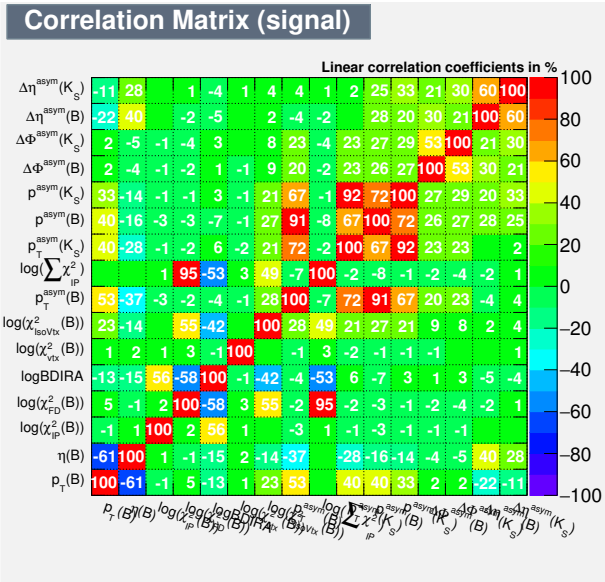
[7]The factor two in this formula comes from the fact that we are interested in the cross section of one b quark.



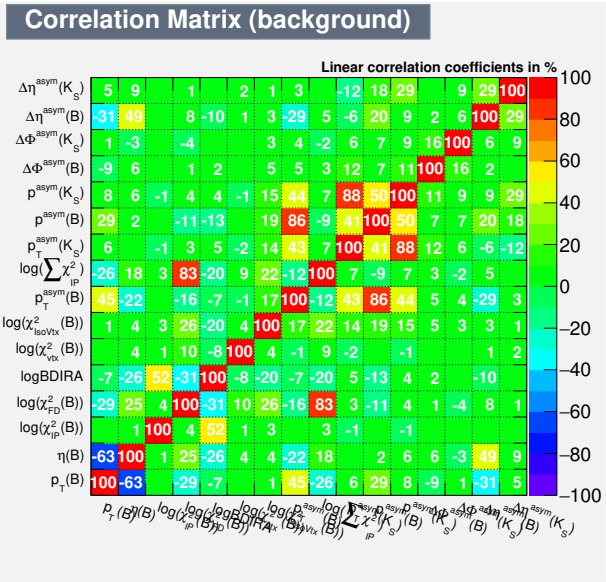
(a)



(b)



(c)



(d)

Figure VI.5: Linear correlations between the MVA input variables for signal and background training samples. (a) and (b) correspond to the 2016 LL samples and (c) and (d) correspond to the 2016 DD samples.

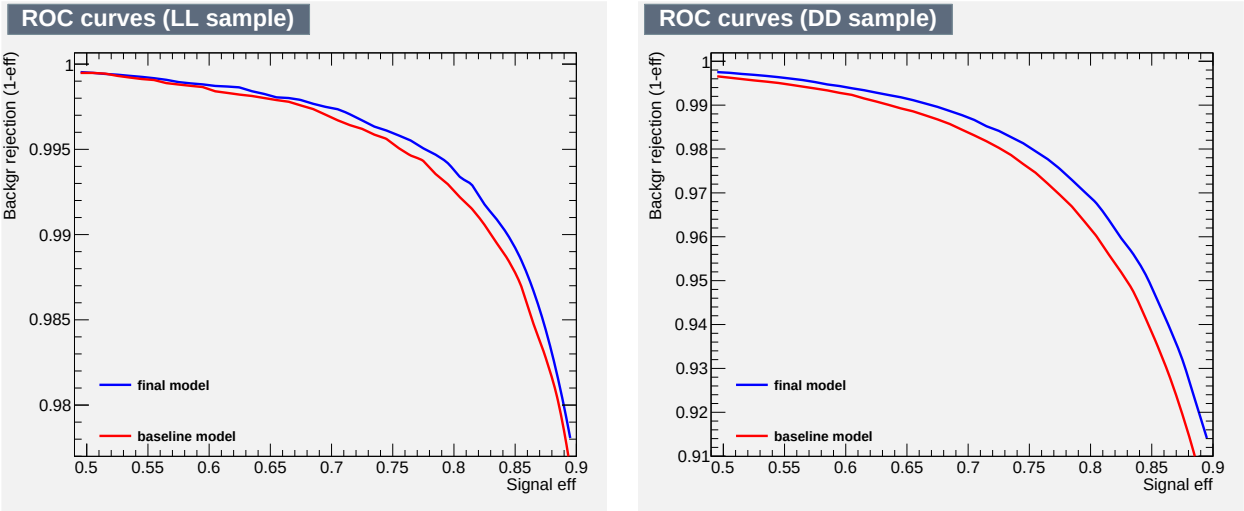


Figure VI.6: Comparison of the ROC curves obtained with the baseline and the final models for 2016. The final model that includes the cones variables has a better performance in both DD and LL samples.

to the hadronisation fraction of the $B^0_{(s)}$ meson, and ϵ^{MC} is the efficiency of the selection applied before the MVA stage evaluated on simulated data. The number of background events in the signal region, B, is taken from an exponential fit to the RHSB of the data, extrapolated to the signal region.

The figure of merit described previously is used for all the modes except for the yet unobserved $B^0_s \to K^0_S K^+ K^-$ for which we use a Punzi FoM

$$\frac{\epsilon_{\text{sig}}}{\frac{a}{2} + \sqrt{B}}, \tag{VI.26}$$

where assumptions on the expected number of signal are not required. The parameter a is generally adjusted according to the desired significance (in units of Gaussian standard deviations). It was noticed in this analysis that the value of a has a very little impact on the shape of the Punzi FoM and on the value of the cut point. It has then be decided to fix this parameter to five.

The figures of merit for $B^0 \to K^0_S \pi^+ \pi^-$, $B^0 \to K^0_S K^+ K^-$ and $B^0_s \to K^0_S K^+ K^-$ for 2016 are given in Fig. VI.7. Note that the maximum of the FoM merit gives an idea of the optimal cut point but remains an approximation, different maxima can be obtained using other FoMs or different values for S and B. This means that, in the final decision on the choice of the cut point, more parameters have to be considered than just the maximum value. In this analysis

we want to take advantage of the fact that the systematic errors associated to the MVA selection cancel at first order in the ratio of Eq. (VI.1) when using the same value for the cut in both the numerator and the denominator modes. The choice of the final value for the cut is then made by comparing the FoMs of all the modes and choosing a common cut point. For robustness reasons, the chosen cut is generally looser than the optimal one. Another point is that the number of training events in the 2011 and 2012a LL sample is rather low, cf. Table IV, and overtraining is thus present and especially for high values of the cut. In order to not be affected by overtraining, the cut for these years is chosen to be zero. The summary of the chosen cuts for all the years is given in Table VI and the their performance on signal and combinatorial background are given in Table VII for $B^0_s \to K^0_S K^+ K^-$ and $B^0 \to K^0_S \pi^+ \pi^-$.

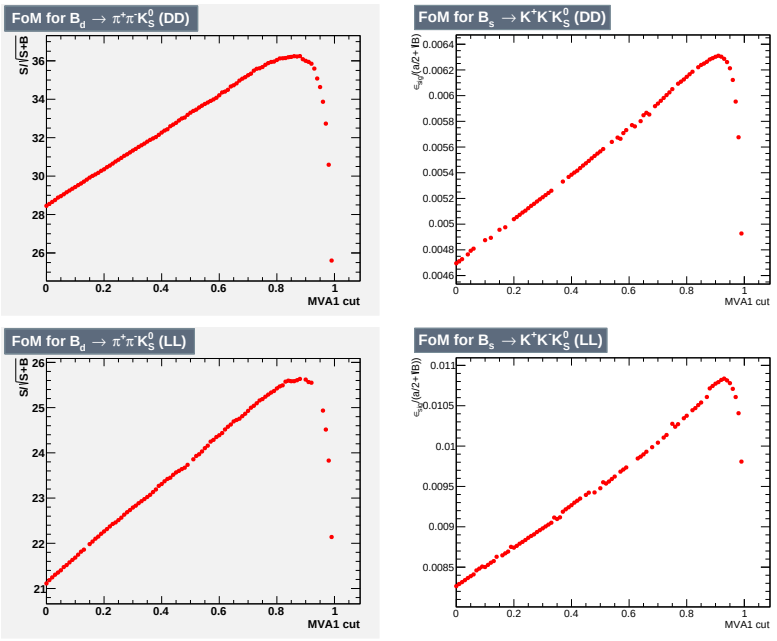


Figure VI.7: Figures of merit for the optimisation of 2016 MVA for the DD (first row) and LL samples (second row).

VI.4.4 Particle identification

The identification of the charged hadrons in LHCb is principally done by the two RICH detectors. Two types of variables are used for the particle identification

Table VI: Cut values for the MVA selection for each data taking period and each K^0_S reconstruction.

year	2011	2012a	2012b	2016
DD	0.0	0.4	0.5	0.8
LL	0.0	0.0	0.5	0.8

Table VII: Percentage of events passing the MVA requirements with respect to the previous steps of the selection. Only $B^0_s \to K^0_S K^+ K^-$ and $B^0 \to K^0_S \pi^+ \pi^-$ modes are shown here but the results are similar for all the final states and the B and B^0_s mesons within the same category (same year and same K^0_S reconstruction). The signal sample is taken from MC and the combinatorial background sample corresponds to the right-hand side band of the data.

Year	Decay mode	DD		LL	
		signal	comb.	signal	comb.
2011	$B^0_s \to K^0_S K^+ K^-$	89.7	18.7	91.9	21.0
	$B^0 \to K^0_S \pi^+ \pi^-$	88.9	10.5	91.2	9.0
2012a	$B^0_s \to K^0_S K^+ K^-$	81.9	12.9	93.9	14.2
	$B^0 \to K^0_S \pi^+ \pi^-$	80.6	6.0	94.1	7.1
2012b	$B^0_s \to K^0_S K^+ K^-$	81.5	9.1	85.9	13.7
	$B^0 \to K^0_S \pi^+ \pi^-$	80.1	3.9	86.2	4.1
2016	$B^0_s \to K^0_S K^+ K^-$	72.0	2.8	83.1	2.7
	$B^0 \to K^0_S \pi^+ \pi^-$	71.0	1.4	82.6	1.1

(PID): the DLL, which are based on differences of log likelihood between of a particular mass hypothesis and the pion hypothesis, and the `ProbNN` that are based on a Bayesian neural network. In this analysis the latter variables are used. The details of the particle identification in LHCb are given in part V.4.

Decays with one (or more) misidentified candidate contribute to the cross-feed background and create peaking structures in the mass spectrum. For $B^0_{(s)} \to K^0_S h^+ h'^-$ decay modes, pions misidentified as kaons (or vice versa) are particularly dangerous because their peak is situated very close to the mass peak. The main purpose of the PID selection is to reduce as much as possible the probability of misidenfication, reducing the amount of crossfeeds. The PID requirements also decrease the quantity of combinatorial background.

PID variables in simulated data

The PID variables in the MC are known to badly reproduce the corresponding variables in the data. In order to design PID cuts using MC information and to evaluate the efficiency of these cuts, it is necessary to correct the MC PID variables. Among the techniques usually used to apply these corrections, one is to completely replace the simulated PID variables by new ones that are randomly generated from calibration PDFs. The main drawback of this method is that the correlations among the PID variables are not preserved. In this analysis, we use a different method, called `PIDcorr` [151], which transforms the MC variables so that their distribution is similar to the data. This is done in an unbinned approach, using calibration PDFs that are described by a kernel density estimation procedure [152]. These PDFs are four dimensional and contain, in addition to the PID variables, kinematic variables such as the transverse momentum and the pseudo-rapidity of the tracks, and event multiplicity information. This method has the advantage to preserve the correlations between the PID variables, so that they can potentially be used into a MVA selection.

PID selection

Since the main concern comes from the $\pi \leftrightarrow K$ misidentification we use the `ProbNN` variables related to these particles, `ProbNN`K and `ProbNN`π, in addition with `ProbNN`p to suppress the misidentification of one of the charge candidates with a proton. Since the values of the `ProbNN` variables range between zero and one they can be interpreted as the probabilities of the different mass hypothesis. An illustration of the shapes of these variables for the $B^0 \to K^0_S K^+ K^-$ decay mode can be found in Fig VI.8. As can be see on the plot, the distributions are very flat with a sharp peak close to zero or one.

To design the cuts, we compare the behaviour of these variables in samples that represent the different species of signal and background considered:

- **signal:** simulated data.
- **crossfeed background:** simulated data of the mode with the misidentified particle reconstructed as the signal mode. For instance, in the case of $K^0_S K^+ K^-$ signal we use simulation of $K^0_S K^+ \pi^-$ events refitted with the $K^0_S K^+ K^-$ hypothesis using the DTF package.
- **combinatorial background:** data taken from the righ-hand side band of the mass spectrum.

The distributions of the PID variables in the different species described above for 2011 and 2016 data taking periods for the DD K^0_S reconstruction category, can be found respectively in Fig VI.9 and Fig VI.10 for $B^0 \to K^0_S\pi^+\pi^-$ and $B^0 \to K^0_S K^+K^-$ decay modes. A similar behaviour is found between DD and LL and between B and B^0_s modes. The distributions corresponding to 2012 samples are comparable to 2011.

Several sets of cuts have been tested and, in each case, the relative efficiencies have been computed for all the different species in order to assess the performance of the cuts on the backgrounds and on the signal. Another requirement of this analysis is that the samples corresponding to the different final states must be independent, leading to the fact that a loose cut on a variable in one sample may imply a tight cut in another sample. Considering all this, the optimisation focuses on the mode of interest, $B^0_s \to K^0_S K^+K^-$, and the control mode $B^0 \to K^0_S\pi^+\pi^-$ while the cuts for the other final states are deduced from these. The cuts that are retained for the baseline model are the ones that best suppress the crossfeeds in the $K^0_S K^+K^-$ spectrum while keeping as much signal as possible. Given the similarities of the `ProbNN`K and `ProbNN`π distributions between the different data taking periods and K^0_S reconstructions, the same cuts are used for all the categories. For the $K^0_S K^+K^-$ final state, a cut on `ProbNN`K larger than 0.5 for both charged hadrons is applied. For the $K^0_S\pi^+\pi^-$ final state we require `ProbNN`π greater than 0.5 and `ProbNN`K smaller than 0.5 for both charged daughters. Finally, for the $K^0_S K^+\pi^-$ final states, for the pion candidate, the requirements are `ProbNN`$\pi > 0.5$ and `ProbNN`$K < 0.5$ and for the kaon candidate `ProbNN`$\pi < 0.5$ and `ProbNN`$K > 0.5$. Very mild cuts are also applied on the `ProbNN`p variable. As can be seen on Figs. VI.9 and VI.10, the behaviour of this variable differs between run I and run II, different cuts are thus applied for both runs. For run I, the charged candidates are required to have a `ProbNN`p value that is less than 0.8, in order to remove potential misidentified protons. For run II, in the case of the kaon candidate, a "bump" can be seen in the distribution around 0.9. This feature is produced by tracks with a momentum smaller than 10 GeV, which corresponds to the Cherenkov threshold below which the RICH detectors cannot differentiate between a kaon and a proton. In this case, the `ProbNN`p cut is chosen to be below 0.9 for kaon candidates with a momentum smaller than 10 GeV and below 0.8 in the other cases. The relative efficiencies of the PID cuts on the signal, crossfeeds and combinatorial background for 2011 and 2016 are given in Table VIII.

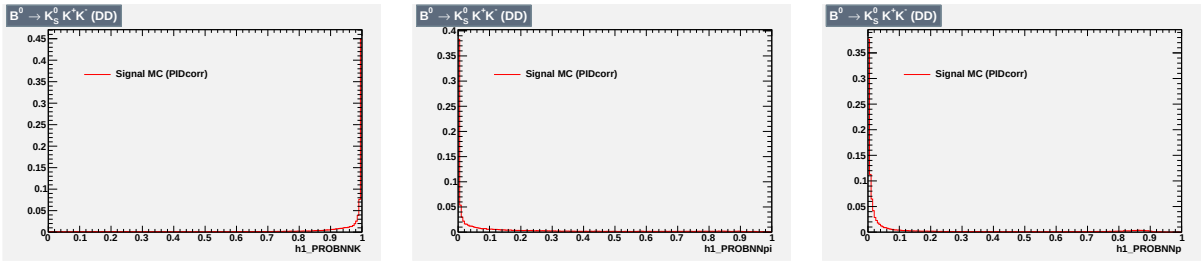


Figure VI.8: Distributions of the three `ProbNN` variables considered in the PID selection. The variables are taken from the simulation 2016 DD sample of $B^0 \to K^0_S K^+ K^-$ decay mode and have been corrected to reproduce the data with the `PIDcorr` method, they correspond to the h_1 candidate, here the K^+.

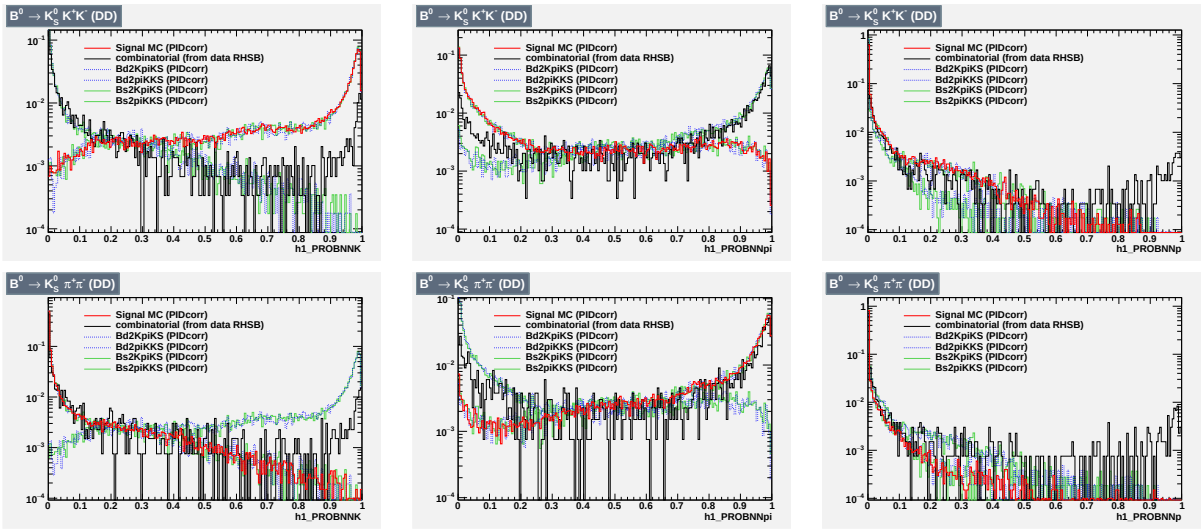


Figure VI.9: Distributions of `ProbNN`K, `ProbNN`π and `ProbNN`K variables for $B^0 \to K^0_S K^+ K^-$ (first row) and $B^0 \to K^0_S \pi^+ \pi^-$ (second row) 2011 DD samples, shown in log scale. The red curve corresponds to signal MC, the dashed blue curve to crossfeed coming from $B^0 \to K^0_S K^+ \pi^-$ and $B^0 \to K^0_S \pi^+ K^-$ with the π misidentified as a K, the green curve represents crossfeed form $B^0_s \to K^0_S K^+ \pi^-$ and $B^0_s \to K^0_S \pi^+ K^-$, and the black curve is the combinatorial background, taken from the data.

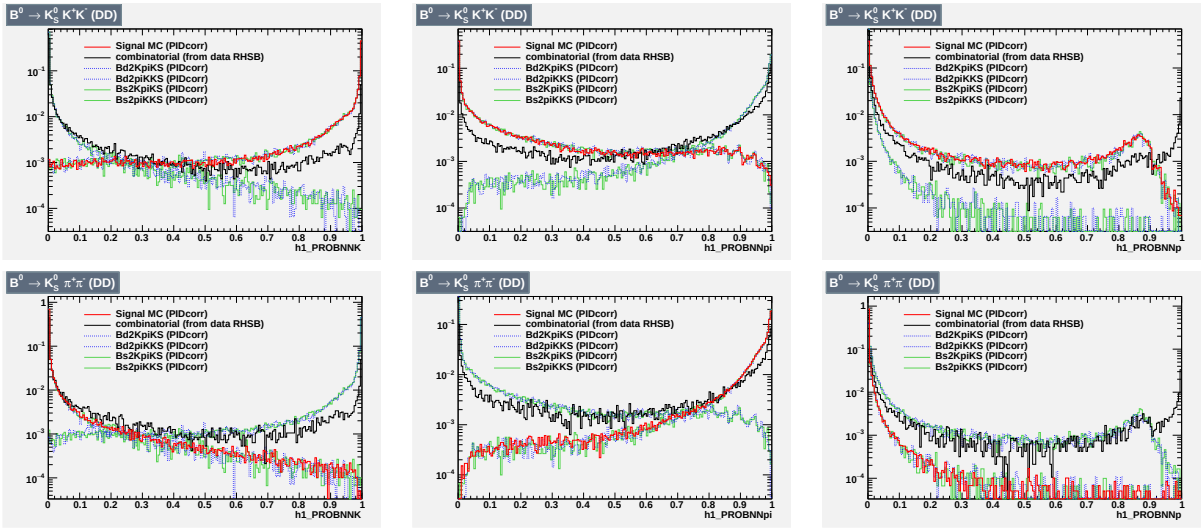


Figure VI.10: Distributions of `ProbNN`K, `ProbNN`π and `ProbNN`K variables for $B^0 \to K^0_S K^+ K^-$ (first row) and $B^0 \to K^0_S \pi^+ \pi^-$ (second row) 2016 DD samples, shown in log scale. The red curve corresponds to signal MC, the dashed blue curve to crossfeed coming from $B^0 \to K^0_S K^+ \pi^-$ and $B^0 \to K^0_S \pi^+ K^-$ with the π misidentified as a K, the green curve represents crossfeed form $B^0_s \to K^0_S K^+ \pi^-$ and $B^0_s \to K^0_S \pi^+ K^-$, and the black curve is the combinatorial background, taken from the data. In order to properly see the variations of the distributions are plotted in log scale.

VI.5 Efficiencies

In order to obtain the relative branching fractions of the $K^0_S h^+ h'^-$ decay modes, the yields extracted from the fit need to be corrected for efficiency. This section describes the different terms entering the computation of the efficiencies. These are summarised in Tables X, XI, XI, XII, and XIII (for $B^0 \to K^0_S \pi^+ \pi^-$ DD and LL, and for $B^0_s \to K^0_S K^+ K^-$ DD and LL).

The efficiencies are obtained from simulated signal samples and computed separately for each data taking period and K^0_S reconstruction class. As mentioned previously, the total efficiency can be expressed in a factorised form,

$$\epsilon^{\text{Tot}} = \epsilon^{\text{Gen}}\, \epsilon^{\text{Reco|Gen}}\, \epsilon^{\text{L0|Reco}}\, \epsilon^{\text{HLT|L0}}\, \epsilon^{\text{Preselection+Vetoes+MVA|HLT}}\, \epsilon^{\text{PID|MVA}}. \tag{VI.27}$$

The terms that enter into this equation are the following:

- ϵ^{Gen}: Generator-level efficiency. When generating MC samples, the

Table VIII: Percentage of events passing the PID requirements with respect to the MVA step measured on the samples corresponding to the different species described in part VI.4.4. Only the DD K^0_S reconstruction is shown here, the results for the LL K^0_S reconstruction are similar, with variations of a few percent. Likewise for the MVA step, the results for B and B^0_s differ by less than 1%. The results for 2012a and 2012b are similar to 2011. Two crossfeeds are considered per decay mode: for $B^0 \to K^0_S\pi^+\pi^-$ and $B^0_s \to K^0_S K^+K^-$ CF1 corresponds to $K^0_S K^+\pi^-$ and CF2 to $K^0_S\pi^+K^-$, and for $B^0 \to K^0_S K^+\pi^-$ CF1 corresponds to $K^0_S\pi^+\pi^-$ and CF2 to $K^0_S K^+K^-$.

Year	Decay mode	Sig	CF1	CF2	comb.
2011	$B^0_s \to K^0_S K^+K^-$	62.0	3.2	3.5	2.1
	$B^0 \to K^0_S\pi^+\pi^-$	65.0	10.3	10.7	40.1
	$B^0 \to K^0_S K^+\pi^-$	51.7	2.1	7.6	5.5
2016	$B^0_s \to K^0_S K^+K^-$	81.8	2.3	2.0	2.6
	$B^0 \to K^0_S\pi^+\pi^-$	89.5	6.4	5.5	36.0
	$B^0 \to K^0_S K^+\pi^-$	75.7	1.1	5.1	7.7

most CPU-consuming step is the detector simulation. In order to save computational resources and speed up the production of simulated data, some cuts are introduced at generator level to ensure that the detector simulation is run on meaningful events, i.e. on events that fall into the LHCb acceptance.

- $\epsilon^{\text{Reco|Gen}}$: Reconstruction efficiency. This contains the efficiency of the reconstruction algorithm that identifies and reconstructs the the signal candidates from information coming from the subdetectors. The stripping efficiency is also included in this step.

- $\epsilon^{\text{L0|Reco}}$: L0 trigger efficiency. Although the data/MC agreement for HLT1 and HLT2 triggers is good, this is not the case for the L0, especially for the hadronic L0 trigger line which is used in this analysis, a systematic uncertainty is assigned to take into account discrepancies between MC and data.

- $\epsilon^{\text{HLT|L0}}$: Efficiency of the HLT triggers (HLT1 and HLT2). The HLT response is well modelled in the MC and can be directly computed from the simulation samples.

- $\epsilon^{\text{Preselection+Vetoes+MVA|Reco}}$: This step concerns the efficiency of the MVA cut together with the preselection cuts and vetoes that were applied prior to the MVA selection.
- $\epsilon^{\text{PID|MVA}}$: PID efficiency. As explained in section VI.4.4, the MC PID variables do not reproduce the data well. A resampling technique was used so that the variables follow the same distribution as they would in data.

Because the MC simulation emulates the trigger decisions without discarding events that fail them, it offers more flexibility in the ordering of the terms in the factorisation. In this section, the efficiencies are presented following a similar sequence as in data, except for the stripping requirement, which is actually applied before the trigger in the MC. (Depending on the analysis this order may be varied; the trigger efficiency is sometimes computed last in order to simplify the evaluation of systematic uncertainties.)

MC truth matching

A small amount of background is present in the MC samples, mainly originating from misreconstructed events and combinatorial background. To determine the pure signal yield in the signal MC sample, various techniques are available. One can, for example, perform a maximum likelihood fit or use side band subtraction. Another approach, which is chosen here, is to use the information contained at "truth" level. At LHCb, each MC sample includes information on the true nature of reconstructed particles (i.e. whether they are matched to a generated particle) and the true values of their properties, that is to say the value of a variable at the generation of the particle, as well as the reconstructed values. The two usually differ because of resolution effects, hit inefficiencies, and small mistakes in the pattern recognition. An algorithm is used to match reconstructed candidate with the corresponding true particles. By imposing requirements on this truth matching, it is possible to select only reconstructed particles that correspond to the generated $B^0_{(s)} \to K^0_S h^+ h'^-$ signal.

VI.5.1 Generator level cuts

Some very loose cuts have been set at generator level to speed up MC generation. These cuts are looser that the stripping cuts and basically make sure that the daughter particles are inside the acceptance of the LHCb detector, for instance requiring their pseudorapidity to be between 2 and 5. The same requirements have been applied to all the decay modes. The complete list of the cuts can

be found in Table IX. The generator level cut efficiencies are similar for the B^0 and B^0_s mesons, and they are also rather consistent between the different data taking periods. Depending on the final state, they vary between 5.9% and 7.2% with a higher value for modes containing charged kaons.

Table IX: Cuts applied at MC generator level.

Cut	Description
	B **candidate**
$p_T(B) > 1500$ MeV	Transverse momentum of the B candidate
	$h^\pm$ **candidates**
$0.010 < \theta(h^\pm) < 0.400$	Charged daughters of the B within detector acceptance.
$1.8 < \eta(h^\pm) < 5.0$	Pseudo-rapidity of $h^\pm$ candidates.
3.0GeV < p(h$^\pm$) < 150GeV	Total momentum.
	K^0_S **candidate**
$2 < \eta(K^0_S) < 5$	Pseudo-rapidity of K^0_S candidate.
`ksTT = GVEV & ( GFAEVX ( GVZ , 1.e+10 ) < 240 * centimeter )`	K^0_S must decay less than 240 cm downstream of interaction point
	K^0_S **daughters**
$1.6 < \eta(\pi^\pm) < 5.2$	Pseudo-rapidity of $h^\pm$ candidates.
2.0GeV < p($\pi^\pm$) < 150GeV	Total momentum.
`bothPI = 2 == GNINTREE ( ( 'pi+' == GABSID ) )`	K^0_S must decay to two charged pions

VI.5.2 Reconstruction and stripping efficiencies

The reconstruction and stripping efficiencies are evaluated from signal MC samples using the same reconstruction algorithm and the same stripping version as in the data. The typical reconstruction efficiency with respect to the generator-level cut step for the DD samples is of about 7.5% in run I and for 4% in run II for all the modes. For the LL samples, the typical efficiency is of about 3% in run I and for 2% in run II for all the modes. The differences between run I and run II are mainly due to changes in the run II stripping lines and are compensated by changes in the relative trigger efficiency (see below).

VI.5.3 Trigger efficiency

Several possibilities exist to evaluate the L0 trigger efficiency. The baseline for the computation of the trigger efficiencies chosen here is the use of simulated samples. Since this method can be biased by the imperfect modelling of the trigger response in the MC, a systematic uncertainty related to this particular choice is assigned and is discussed in section VI.8.

The L0 efficiency relative to the previous step (stripping) is rather similar between the decay modes and between DD and LL samples. The values are about 50–55% for run I and about 97% for run II. The difference between

the two runs is explained by the presence of cuts on the trigger decisions in the run II stripping lines. Specifically, in run II the stripping requires a positive software trigger decision, and the software trigger in turn requires that the event fired at least one L0 trigger line. This explains why the stripping efficiency is lower and the relative trigger efficiency is higher in run II: the product remains approximately constant between run I and run II. Since our L0 selection accepts most events that fire a physics trigger (it accepts events that are TOS on L0Hadron or TIS on most L0 physics lines), it is not surprising to obtain a value for the efficiency in run II close to 100%. A similar pattern is observed for the HLT efficiency, and the same explanation applies.

VI.5.4 Preselection, vetoes and MVA efficiencies

The preselection cuts, the vetoes and the MVA optimisation are described in section VI.4. The efficiency of the preselection cuts is comparable between the decay modes and data taking periods. It is about 60% in the LL sample and 75% in the DD sample in run I and 60% and 70% in the LL and DD sample in run II, respectively. The signal efficiency of the vetoes is very similar between the different data taking periods and K^0_S reconstruction, ranging between 80% and 90% with the highest values seen for the $K^0_S K^\pm \pi^\mp$ final state and the smallest for $K^0_S \pi^+ \pi^-$. On the other hand, the MVA cut efficiency does vary between the different categories, which is explained by differences in the training and the chosen cut points; the efficiencies are similar between the modes within the same final state.

VI.5.5 PID efficiency

The PID efficiency is calculated using variables that are resampled using the `PIDcorr` package [151] to match the distributions in control samples of data, as explained in Sec. VI.4.4. A consistent set of cuts are applied for all run I samples, but the cut values are changed for run II (specifically, the `ProbNN`p cut differs). Partly as a result, the PID efficiencies are higher for 2016 than for run I. For the four final states $K^0_S \pi^+ \pi^-$ and $K^0_S K^+ K^-$, the relative efficiency is typically 60–70% for run I and 80–90% for runII; somewhat smaller values are found for $K^0_S K^\pm \pi^\mp$.

VI.5.6 Summary of the efficiencies

Summary tables of the efficiencies describes previously can be found in Tables X and XI for $B^0 \to K^0_S \pi^+ \pi^-$ and in Tables XII and XIII for $B^0_s \to K^0_S K^+ K^-$.

The variation of the total efficiency as a function of the position on the square Dalitz plane is shown in Fig. VI.11 for the 2016 data taking period. Clear structures (regions of low efficiency) can be observed; these are related to the vetoes that are applied to the samples. For illustration, the same plots but with the vetoes excluded are shown in Fig. VI.12.

Table X: Summary of the efficiencies for $B^0 \to K^0_S\pi^+\pi^-$ for the DD K^0_S reconstruction sample. All values are in percent (including the final row). The uncertainties are related to the limited size of the MC sample.

	2011	2012a	2012b	2016
ϵ^{Gen}	5.889 ± 0.019	6.011 ± 0.016	6.005 ± 0.017	6.299 ± 0.021
$\epsilon^{\text{Reco\|Gen}}$	8.034 ± 0.018	7.295 ± 0.017	7.350 ± 0.018	4.010 ± 0.010
$\epsilon^{\text{L0\|Reco}}$	55.950 ± 0.119	51.212 ± 0.120	54.511 ± 0.130	96.840 ± 0.044
$\epsilon^{\text{HLT\|L0}}$	37.227 ± 0.155	46.281 ± 0.167	37.158 ± 0.170	96.027 ± 0.050
$\epsilon^{\text{Presel\|HLT}}$	77.700 ± 0.218	75.654 ± 0.212	76.236 ± 0.246	68.779 ± 0.120
$\epsilon^{\text{Vetoes\|Presel}}$	80.028 ± 0.238	80.896 ± 0.223	79.820 ± 0.266	81.135 ± 0.122
$\epsilon^{\text{MVA\|Vetoes}}$	88.895 ± 0.209	80.645 ± 0.249	80.109 ± 0.296	70.996 ± 0.157
$\epsilon^{\text{PID\|MVA}}$	65.038 ± 0.337	72.663 ± 0.313	71.335 ± 0.375	89.458 ± 0.126
ϵ^{Tot}	$0.0354 \pm (3.296 \times 10^{-4})$	$0.0373 \pm (3.217 \times 10^{-4})$	$0.0311 \pm (3.166 \times 10^{-4})$	$0.0832 \pm (4.542 \times 10^{-4})$

Table XI: Summary of the efficiencies for $B^0 \to K^0_S\pi^+\pi^-$ for the LL K^0_S reconstruction sample. All values are in percent (including the final row). The uncertainties are related to the limited size of the MC sample.

	2011	2012a	2012b	2016
ϵ^{Gen}	5.889 ± 0.019	6.011 ± 0.016	6.005 ± 0.017	6.299 ± 0.021
$\epsilon^{\text{Reco\|Gen}}$	3.048 ± 0.012	2.542 ± 0.010	2.541 ± 0.011	1.807 ± 0.007
$\epsilon^{\text{L0\|Reco}}$	55.459 ± 0.193	50.624 ± 0.203	54.304 ± 0.220	96.735 ± 0.066
$\epsilon^{\text{HLT\|L0}}$	54.696 ± 0.260	54.231 ± 0.285	60.374 ± 0.294	97.360 ± 0.061
$\epsilon^{\text{Presel\|HLT}}$	61.391 ± 0.344	58.828 ± 0.382	59.658 ± 0.379	59.360 ± 0.188
$\epsilon^{\text{Vetoes\|Presel}}$	81.250 ± 0.352	80.854 ± 0.398	80.771 ± 0.394	81.242 ± 0.194
$\epsilon^{\text{MVA\|Vetoes}}$	91.269 ± 0.282	94.071 ± 0.266	86.175 ± 0.384	82.613 ± 0.209
$\epsilon^{\text{PID\|MVA}}$	72.723 ± 0.466	77.037 ± 0.488	77.885 ± 0.498	91.329 ± 0.171
ϵ^{Tot}	$0.0180 \pm (2.283 \times 10^{-4})$	$0.0145 \pm (1.948 \times 10^{-4})$	$0.0162 \pm (2.244 \times 10^{-4})$	$0.0390 \pm (2.791 \times 10^{-4})$

Table XII: Summary of the efficiencies for $B_s^0 \to K_S^0 K^+ K^-$ for the DD K_S^0 reconstruction sample. All values are in percent (including the final row). The uncertainties are related to the limited size of the MC sample.

	2011	2012a	2012b	2016
ϵ^{Gen}	6.692 ± 0.016	6.806 ± 0.018	6.808 ± 0.019	7.104 ± 0.025
$\epsilon^{\text{Reco\|Gen}}$	7.847 ± 0.018	7.198 ± 0.017	7.219 ± 0.018	3.860 ± 0.010
$\epsilon^{\text{L0\|Reco}}$	53.849 ± 0.117	48.930 ± 0.120	52.074 ± 0.130	96.658 ± 0.046
$\epsilon^{\text{HLT\|L0}}$	40.188 ± 0.157	49.103 ± 0.172	39.838 ± 0.177	95.934 ± 0.051
$\epsilon^{\text{Presel\|HLT}}$	76.995 ± 0.212	75.248 ± 0.212	76.159 ± 0.244	69.269 ± 0.122
$\epsilon^{\text{Vetoes\|Presel}}$	83.967 ± 0.211	85.354 ± 0.200	84.788 ± 0.236	85.396 ± 0.112
$\epsilon^{\text{MVA\|Vetoes}}$	89.656 ± 0.191	81.930 ± 0.236	81.544 ± 0.277	72.009 ± 0.154
$\epsilon^{\text{PID\|MVA}}$	61.986 ± 0.322	63.840 ± 0.326	62.950 ± 0.382	81.778 ± 0.156
ϵ^{Tot}	$0.0409 \pm (3.567 \times 10^{-4})$	$0.0395 \pm (3.504 \times 10^{-4})$	$0.0338 \pm (3.521 \times 10^{-4})$	$0.0892 \pm (4.557 \times 10^{-4})$

Table XIII: Summary of the efficiencies for $B_s^0 \to K_S^0 K^+ K^-$ for the LL K_S^0 reconstruction sample. All values are in percent (including the final row). The uncertainties are related to the limited size of the MC sample.

	2011	2012a	2012b	2016
ϵ^{Gen}	6.692 ± 0.016	6.806 ± 0.018	6.808 ± 0.019	7.104 ± 0.025
$\epsilon^{\text{Reco\|Gen}}$	2.915 ± 0.011	2.463 ± 0.010	2.457 ± 0.011	1.709 ± 0.006
$\epsilon^{\text{L0\|Reco}}$	52.790 ± 0.192	47.777 ± 0.206	51.325 ± 0.224	96.651 ± 0.069
$\epsilon^{\text{HLT\|L0}}$	56.790 ± 0.263	55.719 ± 0.296	61.780 ± 0.304	97.460 ± 0.061
$\epsilon^{\text{Presel\|HLT}}$	61.394 ± 0.343	57.493 ± 0.395	59.665 ± 0.390	59.377 ± 0.193
$\epsilon^{\text{Vetoes\|Presel}}$	85.223 ± 0.319	84.521 ± 0.381	85.651 ± 0.361	85.323 ± 0.181
$\epsilon^{\text{MVA\|Vetoes}}$	91.946 ± 0.265	93.869 ± 0.275	85.907 ± 0.387	83.083 ± 0.207
$\epsilon^{\text{PID\|MVA}}$	60.401 ± 0.496	60.909 ± 0.577	61.429 ± 0.584	80.500 ± 0.240
ϵ^{Tot}	$0.0170 \pm (2.255 \times 10^{-4})$	$0.0124 \pm (1.905 \times 10^{-4})$	$0.0143 \pm (2.235 \times 10^{-4})$	$0.0390 \pm (2.810 \times 10^{-4})$

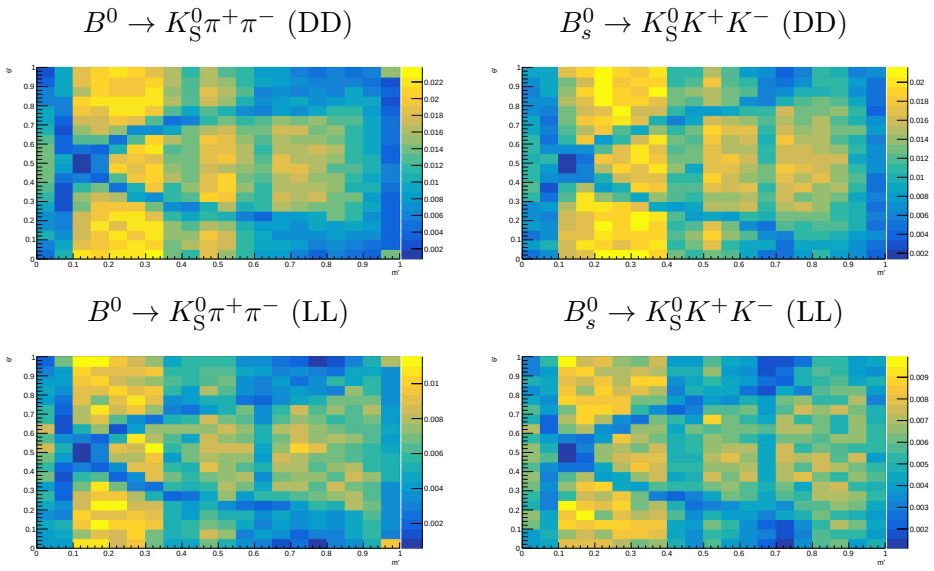


Figure VI.11: Variation of the selection efficiency across the square Dalitz plane for $B^0 \to K^0_S\pi^+\pi^-$ and $B^0_s \to K^0_S K^+K^-$ modes, obtained from 2016 MC samples. The generator level cuts efficiencies are not included here.

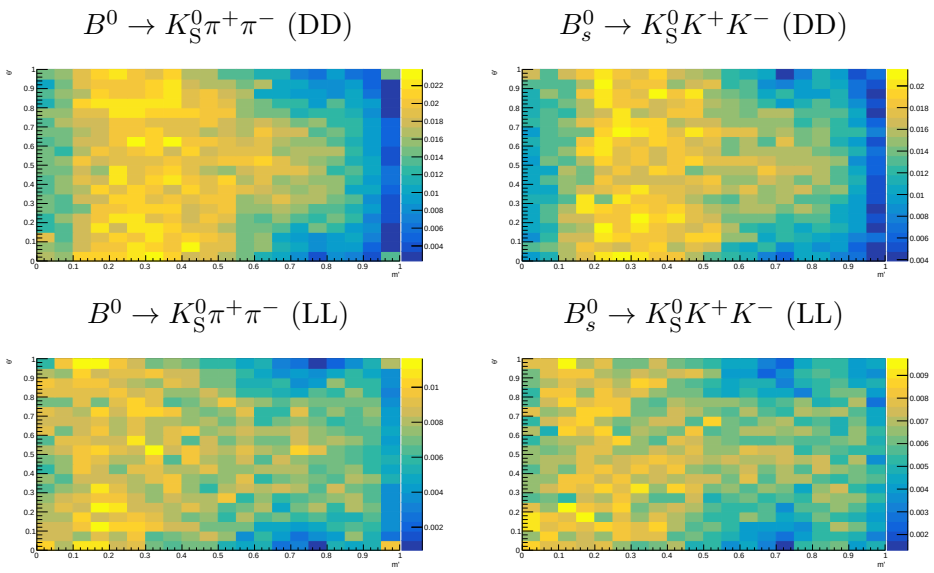


Figure VI.12: Variation of the selection efficiency across the square Dalitz plane for $B^0 \to K^0_S\pi^+\pi^-$ and $B^0_s \to K^0_S K^+K^-$ modes, obtained from 2016 MC samples. The generator level cuts efficiencies are not included here nor are the vetoes.

VI.6 Fit model

This section describes the different components included in the mass fit and how they are modelled. An unbinned extended maximum likelihood fit is performed on all of the data, separated by final state ($K^0_S\pi^+\pi^-$, $K^0_S K^+K^-$, $K^0_S K^+\pi^-$, $K^0_S\pi^+K^-$), data taking period (2011, 2012a, 2012b, 2016) and K^0_S reconstruction category (DD, LL), representing a total of 32 subsamples. All the subsamples are fitted simultaneously and the corresponding signal yields are extracted from the fit. The total PDF is a combination of signal and background PDFs. Three kinds of background are included in the fit: combinatorial, crossfeed, and partially reconstructed backgrounds; they will be described in the following sections. The stability and internal consistency of the fitting framework has been tested with pseudoexperiments, and no bias was found.

Given the potential number of parameters describing each of the spectra, a direct fit to the data with all parameters free is not feasible. Instead, many of the parameters governing the shapes of the different components are determined with fits to MC simulation and then fixed when fitting the data. Other constraints are also used (e.g. on the relative normalisation of components between modes; see below).

Signal model

The $B^0_{(s)} \to K^0_S h^+ h'^-$ signal modes are modelled using a double Crystal-Ball (CB) function, which consists of the sum of two single CBs [153] that share common mean and width parameters. A CB distribution has a Gaussian core with a power-law tail and is defined as

$$f_{CB}(t;n,\alpha,\sigma) = \mathcal{N}\cdot\begin{cases} e^{-t^2/2\sigma^2} & \text{if } t/\sigma > -\alpha \\ \left(\frac{n}{|\alpha|}\right)^n \left(\frac{n-\alpha^2}{|\alpha|} - \frac{t}{\sigma}\right)^{-n} e^{-\alpha^2/2} & \text{if } t/\sigma \leq -\alpha \end{cases}, \quad \text{(VI.28)}$$

where $\mathcal{N}$ is a normalisation factor, σ is the resolution of the Gaussian, and $t = m - \mu$ is the difference between the reconstructed mass m and the mean value μ of the Gaussian. The parameter n governs the power law tail and the sign of α determines the location of the tail, right-handed or left-handed, with respect to the mean value of the Gaussian. A double CB has two independent tails. The left tail is dominated by the radiative effects coming from the emission of photons by the final state particles that smears out the mass peak, along with stochastic effects related to the detector. The right tail accounts for non-Gaussian detector resolution effects.

The positions of the signal peaks, corresponding approximately to the B^0 and B_s^0 masses, are free to vary in the fit to data since the momentum scale calibration may be imperfect (and may differ between samples). The values of the parameters governing the tails of the CBs as well are their relative fraction are extracted from simulated samples and fixed in the fit to data. An example of the results such fits is shown in Fig. VI.13. The parameters related to the left tail, α_L and n_L, are free to vary in the MC fit between all the modes and categories, whereas the right tail parameters, α_R and n_R, are constrained to be identical between the modes.

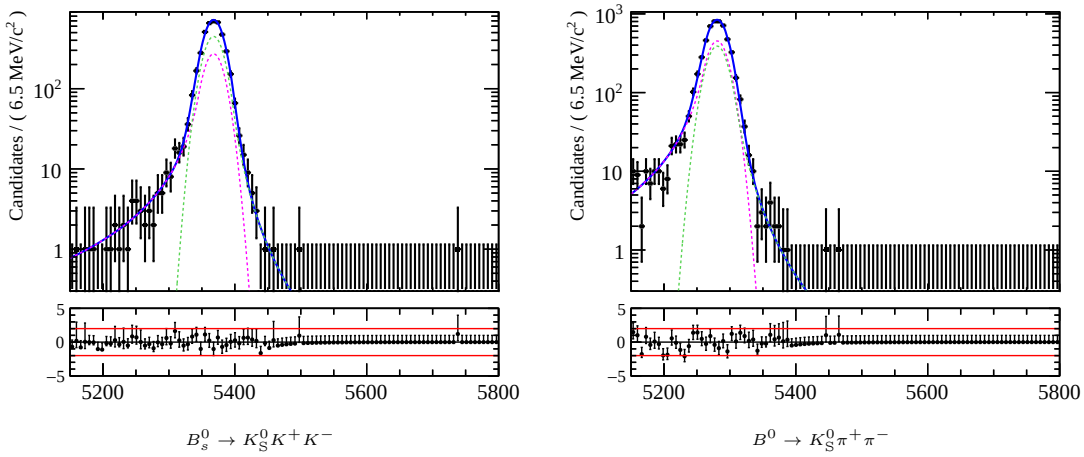


Figure VI.13: Examples of fits to the invariant mass of simulated data, corresponding to $B_s^0 \to K_S^0 K^+ K^-$ (left) and $B^0 \to K_S^0 \pi^+ \pi^-$ (right) signal modes. The samples were obtained with running conditions analogous to 2012b data taking period and the K_S^0 is reconstructed in the LL category. The x-axis shows the reconstructed candidate mass in MeV.

Combinatorial background model

The combinatorial background is described by a Chebychev polynomial of the first kind, of order one (i.e. a linear function). The single slope parameter is allowed to vary between the different data taking periods and K_S^0 reconstruction classes.

Crossfeed model

The crossfeed contributions are modelled by sum of two CBs. The CB parameters are determined by fitting simulated data that has been deliberately reconstructed under the wrong mass hypothesis (e.g. $K^0_S\pi^+\pi^-$ reconstructed as $K^0_S K^+\pi^-$) using the DecayTreeFitter package described in section VI.3.2. Crossfeed backgrounds are considered only if they have a single misidentified pion or kaon and produce a peak within the relevant mass range. Other potential misidentified decays are expected to be small and are thus neglected. Consequently, four crossfeed contributions are included in each spectra:

- In the $K^0_S K^+K^-$ and $K^0_S\pi^+\pi^-$ spectra, crossfeeds from the $B^0 \to K^0_S K^+\pi^-$, $B^0 \to K^0_S\pi^+K^-$, $B^0_s \to K^0_S K^+\pi^-$, and $B^0_s \to K^0_S\pi^+K^-$ decay modes reconstructed and selected as $K^0_S K^+K^-$ or $K^0_S\pi^+\pi^-$ are included in the model. The contribution from the B^0_s mesons are expected to be larger than the ones from the B^0 given that the B^0_s decay is favoured in the $K^0_S K^\pm\pi^\pm$ spectra.

- In the $K^0_S K^+\pi^-$ and $K^0_S\pi^+K^-$ spectra, crossfeeds from the $B^0 \to K^0_S\pi^+\pi^-$, $B^0 \to K^0_S K^+K^-$, $B^0_s \to K^0_S\pi^+\pi^-$, and $B^0_s \to K^0_S K^+K^-$ decay modes reconstructed and selected as $K^0_S K^+\pi^-$ or $K^0_S\pi^+K^-$ are included in the model. Unlike the former case, the favoured mode in the spectra with two pions or two kaons is the B^0 so that we expect more evens in the crossfeed coming from these modes.

All the parameters governing the shapes are taken from a fit to simulated data reconstructed with the corresponding mass hypothesis and are fixed in the fit to the data. Since the yields of these component is small, it cannot be totally free in the fit. Instead, Gaussian constraints are used. The ratio of the yields of the mis-identified decay (A reconstructed as B) and the true mode (A) can be expressed as a ratio of efficiencies

$$\frac{N(\text{A reconstructed as B})}{N(\text{A})} = \frac{\epsilon^{\text{sel}}_{\text{A-as-B}} \times \epsilon^{\text{MVA}}_{\text{A-as-B}} \times \epsilon^{\text{PID}}_{\text{A-as-B}}}{\epsilon^{\text{sel}}_{\text{A}} \times \epsilon^{\text{MVA}}_{\text{A}} \times \epsilon^{\text{PID}}_{\text{A}}}. \tag{VI.29}$$

Using this relation, it is possible to obtain the number of crossfeed events from the yield of the correctly reconstructed mode. Since by design the reconstruction and selection are very similar between different final states in this analysis, and in particular the same MVA selector and cut is used for all the modes (for a given data-taking period and K^0_S reconstruction class), Eq. (VI.29) reduces to the ratio of PID efficiencies:

$$\frac{N(\text{A reconstructed as B})}{N(\text{A})} = \frac{\epsilon^{\text{PID}}_{\text{A-as-B}}}{\epsilon^{\text{PID}}_{\text{A}}}. \tag{VI.30}$$

The efficiencies are obtained from MC simulation using PID variables resampled with `PIDcorr`. The mean value of the Gaussian constraint is thus the ratio of efficiencies while its width is obtained from the statistical uncertainties on the ratio.

The MC we use is generated with a flat distribution in the square Dalitz plane. As discussed at many points in this thesis, the data are not expected to be uniformly distributed. While a distribution uniform in the square Dalitz plane is not a bad first approximation—it produces more events close to the kinematic boundaries of the classic Dalitz plane, as is observed empirically in charmless B decays—it is clearly not correct. To take into account the lack of knowledge of the distribution of the data in the Dalitz plane, the widths of the Gaussian constraints have been multiplied by a factor of 2.

To obtain a better description of the efficiencies, one could reweight the simulated data using Dalitz plot analysis results. At present, Dalitz plot analyses have not been performed for all of the modes and some assumptions would be needed for the unknown ones. Because the crossfeed contributions are relatively small (due to the good PID performance of the LHCb detector), this level of precision is not needed yet.

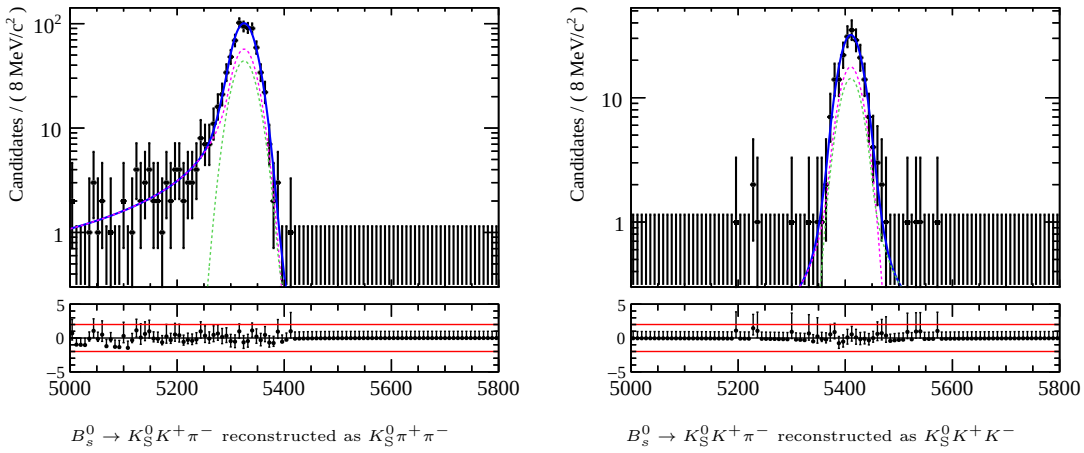


Figure VI.14: Examples of fits to simulated data for the crossfeeds coming from $B_s^0 \to K_S^0 K^+ \pi^-$ decays with a misidentified pion (left) or a misidentified kaon (right). The simulated sample corresponds to the 2012a conditions and the LL K_S^0 reconstruction category. The x-axis shows the reconstructed candidate mass in MeV.

Partially reconstructed backgrounds models

Backgrounds coming from partially reconstructed (PR) decays, with one or more missing particles, appear in the left-hand sideband of the mass spectrum. These backgrounds are modelled with an ARGUS distribution, parametrised as follows

$$f_{\text{ARGUS}}(m) = \mathcal{N} m \left(1 - \left(\frac{m}{m_t}\right)^2\right)^p e^{-\frac{1}{2}c^2\left(1-\left(\frac{m}{m_t}\right)^2\right)}, \tag{VI.31}$$

where $\mathcal{N}$ is a normalisation factor, m_t is a cutoff value, p controls the slope of the curve and c defines its curvature. This distribution is then convoluted with a Gaussian to take into account the detector resolution. The width of this Gaussian is constrained to the resolution of the signal decay mode. All the other parameters except m_t are extracted from fits to simulated data. The mass threshold is fixed to a physical value, which is computed as the mass difference between the $B^0_{(s)}$ meson and the missing particle.

A large number of decay modes can contribute to the PR backgrounds. In order to limit the number of components in the fit, we restrict them to two generic categories: decays with a missing meson — such as $B^0 \to K^{*0}\ (\to K^0_S\pi^0)\ \rho^0\ (\to \pi^+\pi^-)$ and $B^+ \to D^0\ (\to K^0_S\pi^+\pi^-)\ K^+$, where a pion or a kaon is not reconstructed — and radiative decays in which a photon is not reconstructed such as, for example, $B^0 \to \eta'(\to \rho^0\gamma)K^0_S$ and $B^0 \to K^0_S\pi^+\pi^-\gamma$.

Due to the lack of fully simulated samples corresponding to all the decay modes considered and all the data-taking periods, MC samples were obtained with a fast simulation method, called `RapidSim` [154], that can generate phase-space decays of B and D hadrons. Using this simplified framework, it is not possible to apply all the selection cuts to the samples (e.g. the vertex fit χ^2 cannot be estimated). Thus, the shape parameters were obtained from these MC samples without applying any cut. Examples of fits to simulated PRs are shown in Fig. VI.15.

In the fit to the data, the yields of different modes belonging to the same category are Gaussian-constrained using the known branching fractions of the different modes, under the assumption that their efficiencies are the same.

VI.7 Fit results

The results of the simultaneous fit to the data are shown on Fig. VI.16 and VI.17 for 2016 DD and LL samples. The fits corresponding to the run I data taking

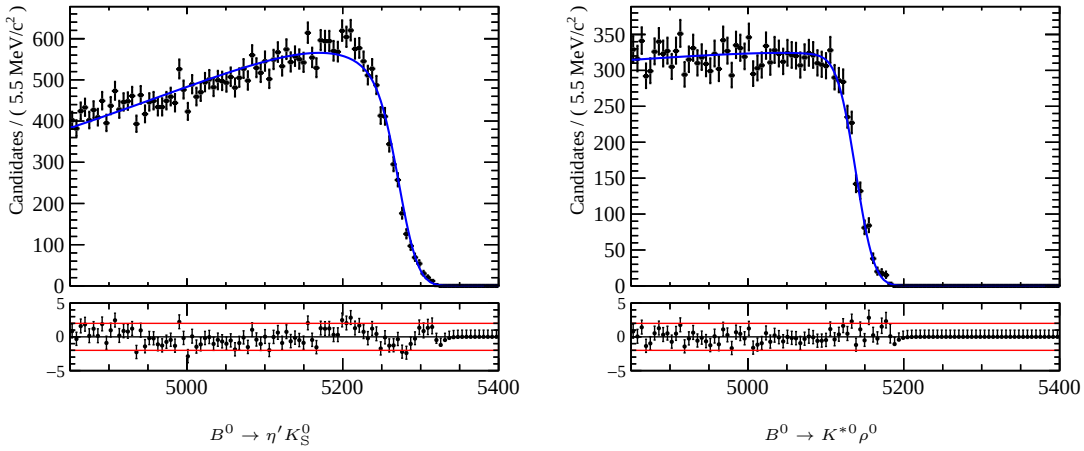


Figure VI.15: Examples of fits to simulated data for partially reconstructed backgrounds. The left plot corresponds to $B^0 \to \eta'(\to \rho^0\gamma)K^0_S$ decay mode where the γ is missing in the reconstruction. The right plot is $B^0 \to K^{*0}$ $(\to K^0_S\pi^0)$ ρ^0 $(\to \pi^+\pi^-)$ with the π^0 not reconstructed. Both samples were generated with the `RapidSim` framework. The x-axis shows $m(K^0_S\pi^+\pi^-)$ in MeV.

periods are shown in Appendix B.2. For each spectrum, the χ^2 probability indicates good fit quality. The extracted yields for each category are reported in Tables XIV and XV for the DD and LL samples, respectively. In total, 206 $B^0_s \to K^0_S K^+K^-$ and 12276 $B^0 \to K^0_S\pi^+\pi^-$ decays are observed. The total yields for the other final states are: 5162 $B^0_s \to K^0_S K^\pm\pi^\mp$ decays, 1361 $B^0 \to K^0_S K^\pm\pi^\mp$ decays, 1248 $B^0_s \to K^0_S\pi^+\pi^-$ decays and 5438 $B^0 \to K^0_S K^+K^-$ decays. Considering only the statistical uncertainties, and propagating the errors taking into account the correlations between the yields, the significance for $B^0_s \to K^0_S K^+K^-$ is above 7.5σ.

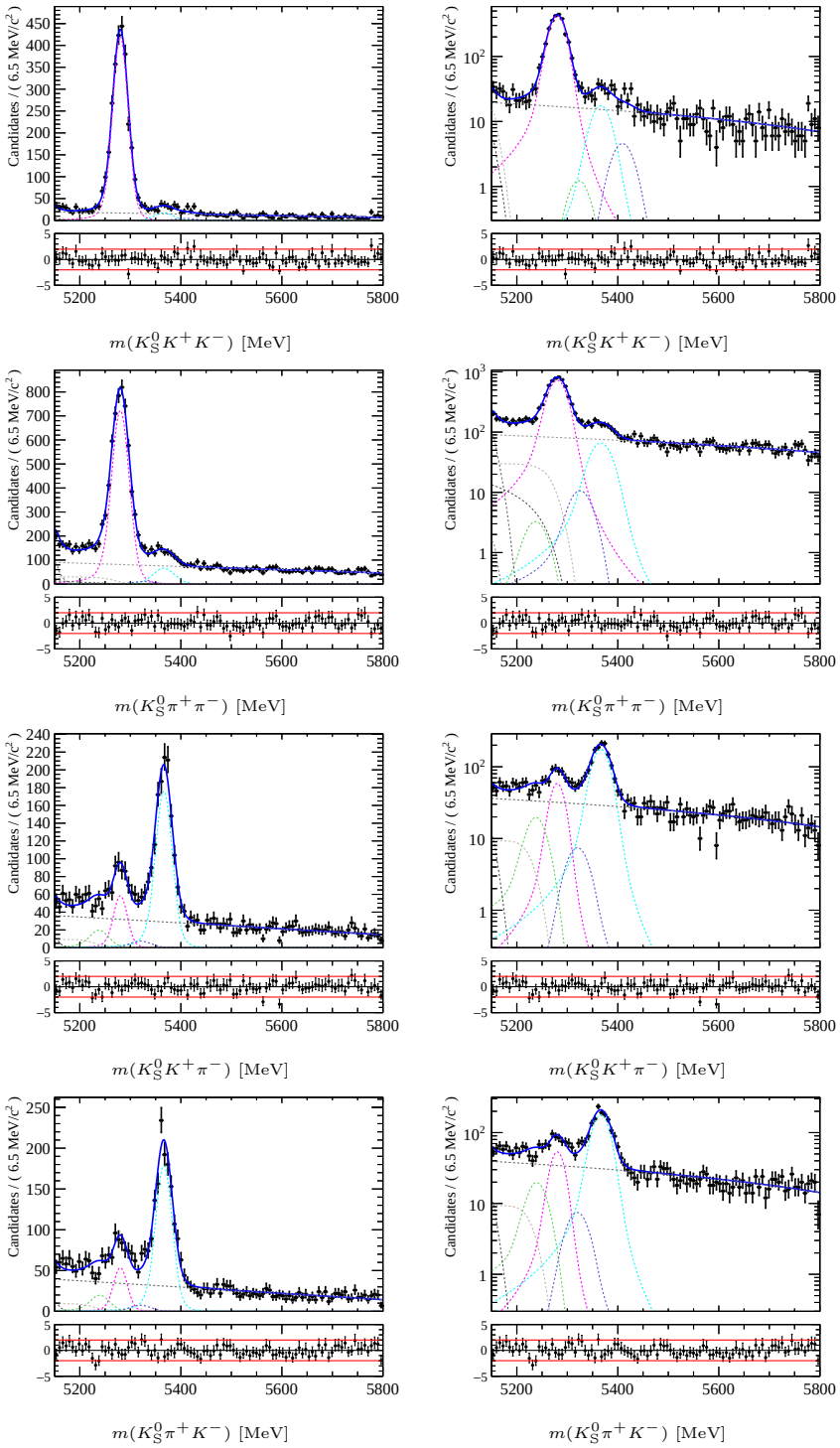


Figure VI.16: Simultaneous fit to the data in the 2016 DD samples, shown on linear (left) and logarithmic (right) scales. The four spectra, $K^0_S K^+ K^-$, $K^0_S \pi^+ \pi^-$, $K^0_S K^+ \pi^-$ and $K^0_S \pi^+ K^-$, are displayed from top to bottom. On each plot, the total PDF is shown in solid blue and the individual components are shown as dashed lines: the B^0 signal in pink, the B^0_s signal in light blue, the crossfeeds from B^0 in green and those from B^0_s in purple, the combinatorial background in grey, and the partially reconstructed backgrounds in grey to the left.

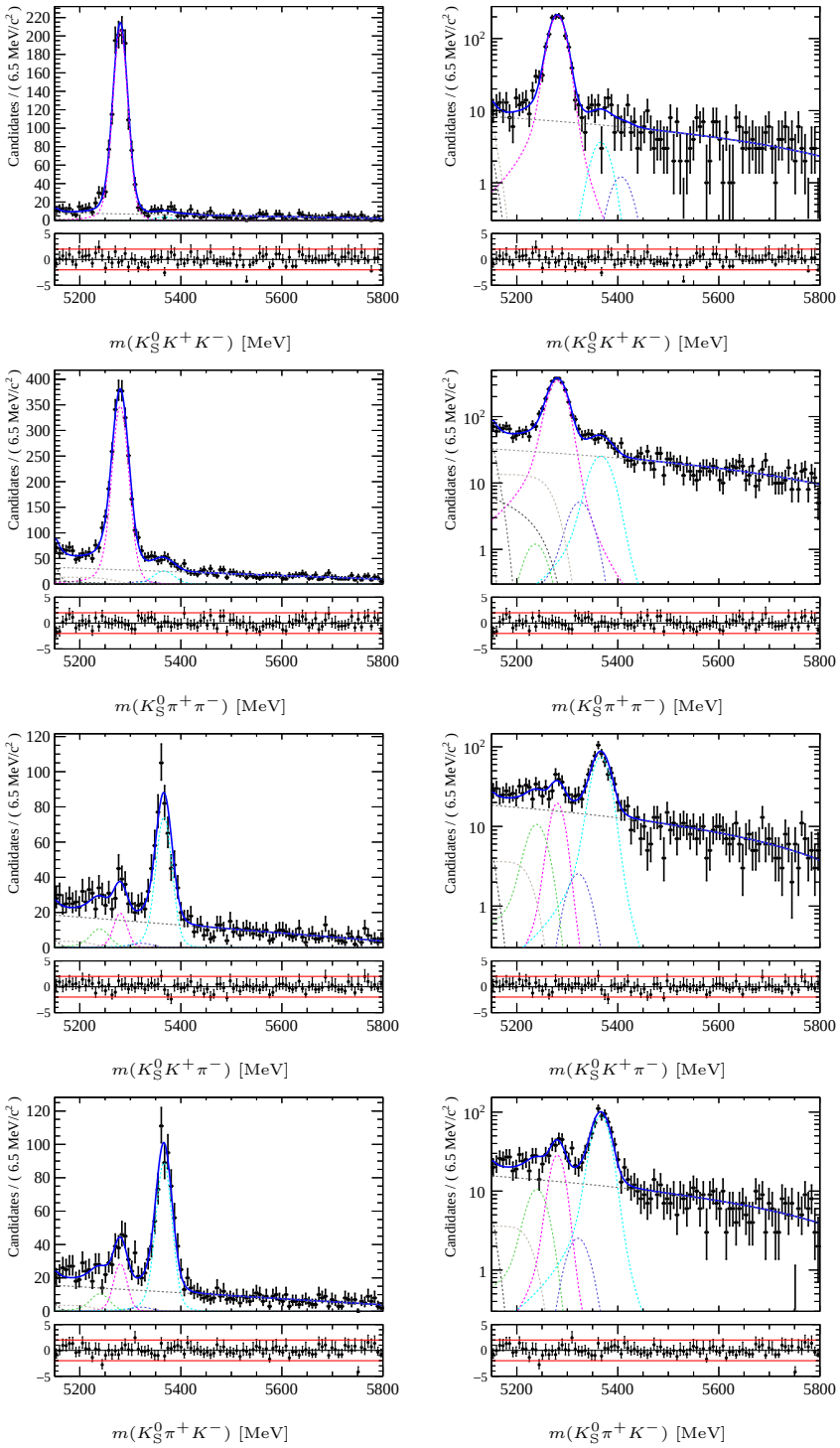


Figure VI.17: Simultaneous fit to the data in the 2016 LL samples, shown on linear (left) and logarithmic (right) scales. The four spectra, $K^0_S K^+ K^-$, $K^0_S \pi^+ \pi^-$, $K^0_S K^+ \pi^-$ and $K^0_S \pi^+ K^-$, are displayed from top to bottom. On each plot, the total PDF is shown in solid blue and the individual components are shown as dashed lines: the B^0 signal in pink, the B^0_s signal in light blue, the crossfeeds from B^0 in green and those from B^0_s in purple, the combinatorial background in grey, and the partially reconstructed backgrounds in grey to the left.

Table XIV: Signal yields extracted from the simultaneous fit to the data for the DD sample, with statistical uncertainties only.

	2011	2012a	2012b	2016
$B^0 \to K^0_S\pi^+\pi^-$	902 ± 37	567 ± 28	1559 ± 46	5292 ± 93
$B^0 \to K^0_S K^+K^-$	286 ± 18	171 ± 14	564 ± 24	2613 ± 56
$B^0_s \to K^0_S K^+\pi^-$	169 ± 15	89 ± 11	306 ± 20	1221 ± 44
$B^0_s \to K^0_S\pi^+K^-$	146 ± 15	110 ± 13	292 ± 20	1234 ± 45
$B^0 \to K^0_S K^+\pi^-$	48 ± 11	38 ± 9	79 ± 13	334 ± 31
$B^0 \to K^0_S\pi^+K^-$	66 ± 12	19 ± 7	69 ± 12	305 ± 31
$B^0_s \to K^0_S\pi^+\pi^-$	80 ± 22	35 ± 13	180 ± 24	586 ± 56
$B^0_s \to K^0_S K^+K^-$	18 ± 7	5 ± 5	5 ± 7	132 ± 21

Table XV: Signal yields extracted from the simultaneous fit to the data for the LL sample, with statistical uncertainties only.

	2011	2012a	2012b	2016
$B^0 \to K^0_S\pi^+\pi^-$	482 ± 26	344 ± 22	651 ± 29	2 479 ± 60
$B^0 \to K^0_S K^+K^-$	175 ± 13	96 ± 10	273 ± 16	1261 ± 38
$B^0_s \to K^0_S K^+\pi^-$	86 ± 11	52 ± 9	93 ± 11	498 ± 28
$B^0_s \to K^0_S\pi^+K^-$	93 ± 11	52 ± 9	121 ± 12	601 ± 30
$B^0 \to K^0_S K^+\pi^-$	26 ± 7	22 ± 7	17 ± 7	108 ± 18
$B^0 \to K^0_S\pi^+K^-$	22 ± 7	10 ± 6	37 ± 8	159 ± 10
$B^0_s \to K^0_S\pi^+\pi^-$	45 ± 13	35 ± 11	67 ± 13	220 ± 31
$B^0_s \to K^0_S K^+K^-$	8 ± 4	5 ± 3	7 ± 4	27 ± 11

VI.8 Systematic uncertainties

The value of the $B_s^0 \to K_S^0 K^+ K^-$ branching fraction is computed relative to the normalisation channel, $B^0 \to K_S^0 \pi^+ \pi^-$. As mentioned before, taking the ratio with respect to another decay mode that is kinematically and topologically similar implies that a great part of the systematic uncertainties cancel or are reduced (though some do not, notably the PID efficiency). The remaining systematic effects are taken into account and are mainly related to the fitting procedure and to differences between data and MC that can affect the efficiency estimates.

VI.8.1 Selection

Most of the efficiencies are directly computed from simulated data, as detailed in Sec. VI.5. Since the MC samples have a finite size, there is an uncertainty related to the size of the sample, which corresponds to the statistical uncertainty on the determination of the efficiencies.

VI.8.2 Tracking

It is known that the simulation of the tracking is not perfectly accurate and should be corrected. However, given the similarities in the topologies and the kinematics of the modes studied here, we expect the systematic error related to the tracking to largely cancel in the ratios of efficiencies. This is confirmed by the very small uncertainty, below 0.3%, associated to the tracking in the previous LHCb analysis [139].

VI.8.3 PID

A systematic uncertainty on the PID selection is evaluated by tightening and loosening the PID cuts. The tight PID selection is obtained by changing the value of the `ProbNN`π cut for pion candidates (`ProbNN`K for kaon candidates) from 0.5 to 0.4, and the loose one corresponds to a change in the same variables from 0.5 to 0.6. For each PID optimisation the whole fit procedure is performed again and the yields are extracted. The efficiency corrected yields — i.e. the yield divided by the corresponding efficiency — are then obtained for each PID optimisation (baseline, loose and tight). The small yields for $B_s^0 \to K_S^0 K^+ K^-$ do not allow to compute a systematic effect, indeed, the fluctuations of the yields between the different optimisations are dominated by the statistical uncertainty. The same argument applies to the yields of $B^0 \to K_S^0 \pi^+ \pi^-$ yields

in run I. To evaluate the systematic effect on the PID for kaons and pions it has been decided to use 2016 DD sample, assuming that the effect is similar in the other categories. The decay mode used to obtain the uncertainty on the pion identification is then $B^0 \to K^0_S \pi^+ \pi^-$ and for the uncertainty on the kaon PID we use $B^0 \to K^0_S K^+ K^-$, which are the most abundant channels. Table XVI give the yields and efficiencies for each of the PID optimisations. The tight optimisation gives an uncertainty of 2.8% for the pion and 7.8% for the kaon. For the loose optimisation we obtain 5% for the pion and 2% for the kaon. A naive average between the loose and tight uncertainties reads 3.9% for the pion and 4.9% for the kaon. Assuming that there is no correlations between the kaon and pion PIDs, the total uncertainty on the PID selection is then of about 6%.

Table XVI: Summary of the yields and efficiencies corresponding to the different PID optimisations considered to assess a systematic uncertainty on the PID. The yields are taken from a simultaneous fit to all the categories, and the efficiencies are obtained from MC. The numbers presented here correspond to the 2016 DD sample.

PID cut	$B^0 \to K^0_S \pi^+ \pi^-$		$B^0 \to K^0_S K^+ K^-$	
	Yield	Efficiency [$\times 10^{-4}$]	Yield	Efficiency [$\times 10^{-4}$]
baseline	5292	8.32398	2613	8.38667
loose	5297	8.57154	2619	9.11769
tight	4702	7.79315	2403	7.56077

VI.8.4 L0 trigger

To assess the bias caused by the discrepancies between the hardware trigger and its MC modelling, a specific procedure, the `TISTOS` method, is used to assess the trigger efficiency from data; it is detailed in Sec. VI.2.4. Three approaches have been used to compute the L0 trigger efficiency:

- Standard MC procedure: the first method consists of obtaining the efficiency by simply counting the events that passed or failed the L0 trigger cuts.
- `TISTOS` on MC: a value for the efficiency is obtained by applying the `TISTOS` method to the MC signal samples.

- `TISTOS` on data: a third value of the efficiency is obtained by applying the `TISTOS` method to the data samples.

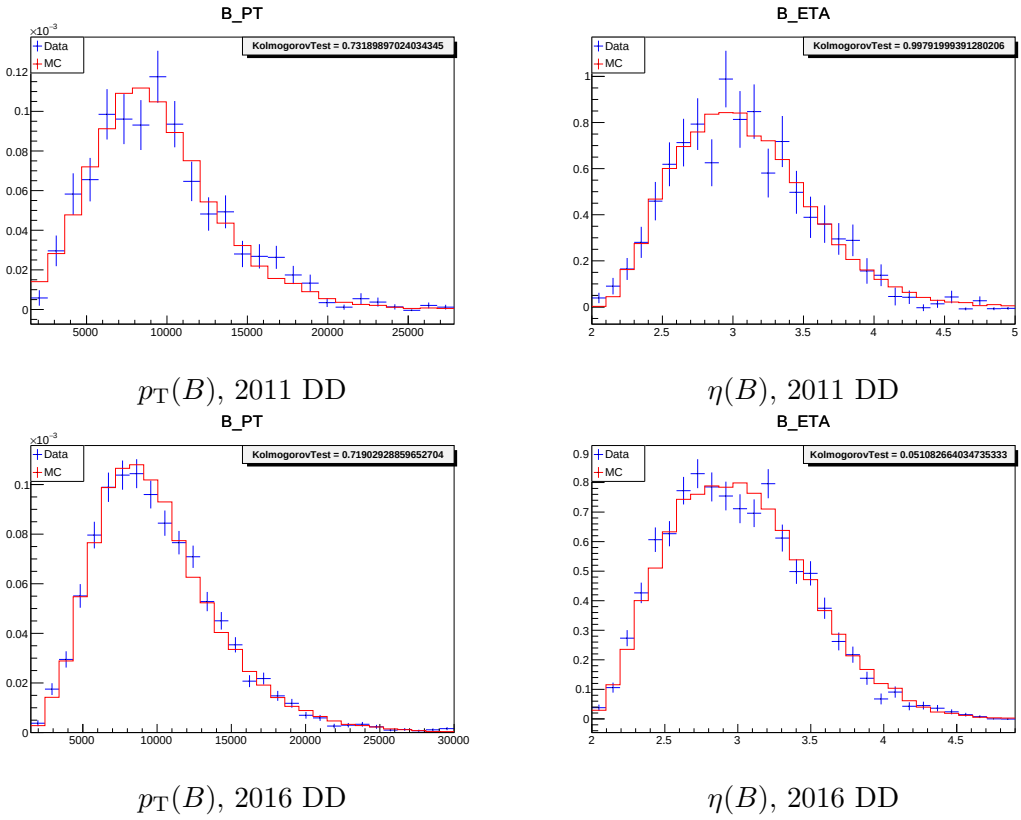
It was found that the results obtained by applying the `TISTOS` method to the simulated samples and to the data agree within their uncertainties. However, given the limited signal statistics in the data, the procedure could only be carried out for the favoured modes. Moreover, the size of the error obtained from the data is too large to be directly exploited, especially for the 2011 and 2012a data taking periods. The results from data are thus used as a cross-check with the values obtained from the simulated samples. Different values are obtained with the standard MC method compared to the `TISTOS`. Nonetheless, the ratios of efficiencies are in good agreement. A systematic uncertainty is therefore assigned by taking the difference between the ratio obtained with the standard MC procedure and the MC `TISTOS` method. The estimated systematic error on the ratio of $B_s^0 \to K_S^0 K^+ K^-$ and $B^0 \to K_S^0 \pi^+ \pi^-$ L0 trigger efficiencies varies between subsamples but is typically about 2%.

Another option to address the L0 efficiency issue would be to use correction maps that are obtained from calibration data samples of $D^{*+} \to D^0(\to \pi^+ K^-)\pi^+$ decays data, analogous to the PID efficiency correction. Unfortunately, these corrections are not available for the run II.

VI.8.5 B transverse momentum and pseudo-rapidity: agreement between data and MC

The kinematic distribution of the B mesons, as a function of p_T and η, may not be perfectly simulated in the MC. The efficiencies in general depend on the B kinematics, and in particular some kinematic variables (including these) are used in the MVA selection. The L0 efficiency, for example, depends on the transverse momentum of the B meson and is less efficient at low $p_T(B)$ than at high $p_T(B)$.

The ${}_s\mathcal{P}lot$ technique described in part VI.2.3, is used to extract the signal weights and compare the ${}_s\mathcal{W}eighted$ distributions of the variables in the data with the MC ones. Comparative plots of $B^0 \to K_S^0 \pi^+ \pi^-$ for 2016 and 2011 data taking periods are shown in Fig. VI.18. Good agreement is seen between the $\eta(B)$ and p_T (B) between the data and the simulation in all the categories. Thus, no corrections to these variables are needed.



$p_T(B)$, 2011 DD $\eta(B)$, 2011 DD

$p_T(B)$, 2016 DD $\eta(B)$, 2016 DD

Figure VI.18: Comparison between the distributions of the p_T (left) and η (right) of the B meson in signal MC and ${}_s\mathcal{W}eighted$ data for the $B^0 \to K^0_S\pi^+\pi^-$ decay mode in 2011 (upper) and 2016 (lower) DD samples. The blue markers correspond to ${}_s\mathcal{W}eighted$ data and the red curve to MC signal. Good agreement in seen between the two.

VI.8.6 Fit model

The values of the extracted yields can be influenced by two main types of systematic effects that are induced by the fit model. One is related to the parameters that are obtained from MC samples and are later fixed in the data fit, and the other is comes from the choice of the shapes of the PDFs used to model the different components. The uncertainty arising from the fixed parameters can be assessed by performing multiple fits to the data in which the values of the fixed parameters are varied according to the correlation matrix of the fit to MC. Since the simulated samples are large, this systematic effect is expected to be subdominant and the uncertainty due to the choice

of lineshapes is expected to dominate. (For the crossfeeds, the yields after selection in the MC are smaller due to the rejection of misidentified candidates, but the yields in data are similarly reduced so the relative effect remains small and is neglected here.)

To estimate the uncertainty due to the choice of lineshapes, the fit is performed again after changing, in turn, the signal, crossfeed and combinatorial background shapes. The alternative model for the signal and the crossfeeds is a single CB, and for the combinatorial background an exponential distribution is used.

Table XVII summarises the systematic effects associated with the fit shapes. The uncertainty arising from each lineshape is obtained by computing the difference between the mean value of the yields obtained with the baseline model and those obtained with the modified shape. (Part of the estimated uncertainty will be of statistical origin.) The dominant uncertainty comes from choice of the signal shape. In samples where the signal yield is small (especially the $B^0_s \to K^0_S K^+ K^-$ 2012b DD sample), the relative uncertainty can be large. However, these statistical fluctuations become less important when propagating the uncertainties to the final result and accounting for correlations between the yields.

Table XVII: Summary of the systematic effects due to the fit model. The difference in number of events with respect to the baseline model is shown for each alternative shape. The total uncertainty on a year obtained by summing in quadrature the different effects. This number is given for reference but is not used in the final propagation of the uncertainties (the errors are propagated individually). Note that only the last column is in % and that all the others show yields or changes in yield.

	Year	Yield	Sig shape	Comb shape	CF shape	Total fit model	relative error [%]
$B^0 \to K^0_S \pi^+ \pi^-$ DD	2011	902.39	9.92	14.58	3.45	18.0	2.0
	2012a	567.3	11.03	10.11	2.95	15.3	2.7
	2012b	1558.5	27.4	7.4	7.8	29.4	5.2
	2016	5292.1	128	53.1	12.2	139.1	2.6
$B^0_s \to K^0_S K^+ K^-$ DD	2011	17.555	0.819	0.053	1.037	1.3	7.5
	2012a	4.799	0.2775	0.2067	1.0245	1.1	22.5
	2012b	5.4045	1.1754	0.3603	4.3609	4.5	83.8
	2016	131.65	4.45	2.06	8.59	9.9	7.5
$B^0 \to K^0_S \pi^+ \pi^-$ LL	2011	481.56	7.42	9.48	1.67	12.2	2.5
	2012a	344.34	7.0	3.35	0.37	7.8	2.6
	2012b	650.67	13.41	14.28	1.75	19.7	3.0
	2016	2479.2	61.2	55.2	5.3	82.6	3.3
$B^0_s \to K^0_S K^+ K^-$ LL	2011	8.4136	0.1092	0.2162	0.8215	0.9	10.2
	2012a	4.6596	0.0549	0.2623	0.9005	0.9	20.2
	2012b	6.9345	0.5339	0.1255	0.577	0.8	11.5
	2016	26.814	1.329	0.0	0.64	1.5	5.5

VI.9 Combination of the results and extraction of the relative branching fraction of $B^0_s \to K^0_S K^+ K^-$

The branching fraction of $B^0_s \to K^0_S K^+ K^-$ relative to that of $B^0 \to K^0_S \pi^+ \pi^-$ is obtained using the master formula given at the beginning of this chapter, eq. (VI.1). The uncertainties on the yields are propagated taking into account their correlations (the correlation matrix is obtained in the fit procedure). The errors on the efficiencies—i.e. the systematic uncertainties due to the finite size of the MC samples, the L0 trigger and the PID—are considered uncorrelated and are propagated accordingly. Table XVIII shows the central value and uncertainties for the combination of all data-taking periods. This is shown separately for the two K^0_S reconstruction classes, and finally for their combination.

Table XVIII: Measured value of $\mathcal{B}(B^0_s \to K^0_S K^+ K^-)/\mathcal{B}(B^0 \to K^0_S \pi^+ \pi^-)$, along with the details of the different statistical and systematic effects. The results for the K^0_S reconstruction classes (DD and LL) are shown separately and then combined taking correlations into account. The relative error is computed with respect to the central value obtained by combining DD and LL samples together.

	DD [$\times 10^{-3}$]	LL [$\times 10^{-3}$]	Total [$\times 10^{-3}$]	Relative error [%]
Central value	54.4	51.3	52.8	
Stat	9.3	13.6	8.3	15.7
Sig shape	1.7	1.4	1.1	2.1
Comb shape	0.7	0.9	0.6	1.1
CF shape	4.1	2.5	2.4	4.5
PID	3.3	3.1	3.2	6.1
Trigger	1.1	1.0	1.1	2.1
MC size	0.3	0.4	0.3	0.6

VI.10 Summary and conclusion

The branching fraction of the decay mode $B^0_s \to K^0_S K^+ K^-$ is measured relative to that of $B^0 \to K^0_S \pi^+ \pi^-$ by analysing the LHCb data collected during run I and 2016, corresponding to an integrate luminosity of $4.67\,\text{fb}^{-1}$. The result

is obtained by performing a simultaneous unbinned extended likelihood fit to the four invariant mass spectra that contain the decay modes $B^0_{(s)} \to K^0_S h^\pm h'^\mp$. The fit returns the B^0 and B^0_s signal yields in each of the four final states, four data taking periods, and two K^0_S reconstruction categories, for a total of 32 samples fitted simultaneously. The result of the final combination is

$$\frac{\mathcal{B}(B^0_s \to K^0_S K^+ K^-)}{\mathcal{B}(B^0 \to K^0_S \pi^+ \pi^-)} = 0.053 \pm 0.008(\text{Stat}) \pm 0.004(\text{Syst}) \pm 0.003(\text{f}_\text{s}/\text{f}_\text{d}) \tag{VI.32}$$

The previous LHCb analysis [139], made with the run I data only, found a 90% confidence interval of 0.008–0.051; this new result is at the upper end of this range but compatible at the 2σ level and more precise. The yield of the $B^0_s \to K^0_S K^+ K^-$ decay mode has a statistical significance of 7.5 standard deviations with stat. errors only, and of almost 7σ when taking the fit shape systematic uncertainties into account as well. The ratio of branching fraction given in Eq. (VI.32) is more than 5σ from zero. This is thus the first observation of this channel.

The result above may be combined with the world-average value [15] $\mathcal{B}(B^0 \to K^0_S \pi^+ \pi^-) = (49.4 \pm 1.8) \times 10^{-6}$ to obtain:

$$\mathcal{B}(B^0_s \to K^0_S K^+ K^-) = (2.6 \pm 0.5) \times 10^{-6}. \tag{VI.33}$$

Chapter VII

Conclusion

During the three years of my PhD, I focused my research on the flavour sector, and more specifically on charmless B meson decays into three particles. My work is divided into two main parts. The first was a feasibility study of a method of extracting the weak phase γ based on flavour-SU(3) symmetry, potentially sensitive to new physics [98]. I presented a complete proof of principle, including fully-propagated experimental uncertainties. Using BABAR Dalitz-plot analysis results of $B^0 \to K^+\pi^0\pi^-$, $B^0 \to K^0_S K^0_S K^0_S$, $B^+ \to K^+\pi^+\pi^-$, $B^0 \to K^0_S K^+ K^-$ and $B^0 \to K^0_S \pi^+\pi^-$, I showed that this method is promising: six values of γ are obtained:

$$
\begin{aligned}
\gamma_1 &= [\ 12.9\,^{+8.4}_{-4.3}\ \ (\text{stat}) \pm 1.3\ (\text{syst})]^\circ, \\
\gamma_2 &= [\ 36.6\,^{+6.6}_{-6.1}\ \ (\text{stat}) \pm 2.6\ (\text{syst})]^\circ, \\
\gamma_3 &= [\ 68.9\,^{+8.6}_{-8.6}\ \ (\text{stat}) \pm 2.4\ (\text{syst})]^\circ, \\
\gamma_4 &= [223.2\,^{+10.9}_{-7.5}\ (\text{stat}) \pm 1.0\ (\text{syst})]^\circ, \\
\gamma_5 &= [266.4\,^{+9.2}_{-10.8}\ (\text{stat}) \pm 1.9\ (\text{syst})]^\circ, \\
\gamma_6 &= [307.5\,^{+6.9}_{-8.1}\ \ (\text{stat}) \pm 1.1\ (\text{syst})]^\circ\ ,
\end{aligned}
$$

one of which is compatible with the world average, $\gamma = (73.5^{+4.2}_{-5.1})^\circ$, while the others are not. The uncertainties on γ are below 11°, which is about twice the world-average uncertainty. I also studied the effect of SU(3) breaking on the result and showed that the SU(3)-breaking effects largely cancel when averaging on many points in the Dalitz plane. This result is very promising and shows that it is possible to extract γ from charmless three-body decays with an acceptable uncertainty. In the coming years, time-dependent Dalitz-plot analyses of these modes with LHCb and Belle II data will be available and an extraction of γ using these results, or directly from the data, will be possible.

This study, largely documented in chapter IV of this thesis, was also published in PRD [99].

The second part of my work is dedicated to the study of the decay modes $B^0_{(s)} \to K^0_S h^\pm h'^\mp$ with the LHCb detector using a dataset collected during run I (2011, 2012) and 2016, corresponding to an integrated luminosity of $4.67\,\text{fb}^{-1}$. I obtained a measurement of the branching fraction of $B^0_s \to K^0_S K^+ K^-$ decay mode relative to $B^0 \to K^0_S \pi^+ \pi^-$:

$$\frac{\mathcal{B}(B^0_s \to K^0_S K^+ K^-)}{\mathcal{B}(B^0 \to K^0_S \pi^+ \pi^-)} = 0.053 \pm 0.008(\text{Stat}) \pm 0.004(\text{Syst}) \pm 0.003(\text{f}_\text{s}/\text{f}_\text{d}).$$

This work follows on from a previous LHCb analysis [139] that was performed on run I data, in which all of the $B^0_{(s)} \to K^0_S h^\pm h'^\mp$ modes were observed except for $B^0_s \to K^0_S K^+ K^-$. After redesigning parts of the selection and adding significantly more data (2016), the $B^0_s \to K^0_S K^+ K^-$ decay mode was observed here for the first time, with a significance of almost 7 standard deviations (after including systematic uncertainties related to the fit model). The result was obtained by performing a simultaneous fit to the different $K^0_S h^\pm h'^\mp$ spectra ($K^0_S \pi^+ \pi^-$, $K^0_S K^+ K^-$, and $K^0_S K^\pm \pi^\mp$) so as to take cross-feeds into account correctly. A total of 32 samples were fitted simultaneously and the yields of the B^0 and the B^0_s mesons were extracted in each of the samples. As well as being of interest in its own right, this analysis prepares the ground for measurements of the branching ratios of all of the $B^0_{(s)} \to K^0_S h^\pm h'^\mp$ decay modes with the full run I and run II LHCb dataset, corresponding to an integrated luminosity of $9\,\text{fb}^{-1}$. The very encouraging results obtained with run I and 2016 ($4.67\,\text{fb}^{-1}$) data, notably the increase in yield of the normalisation channel $B^0 \to K^0_S \pi^+ \pi^-$ by a factor of 2.75 after adding 2016 compared to 2011–2012 alone, suggest that the precision of the measurement with the whole dataset will probably be excellent.

The next step after measuring the branching fractions is to perform Dalitz-plot analyses. Some time-integrated amplitude analyses have already been completed with run I data leading to interesting results, such as $B^0 \to K^0_S \pi^+ \pi^-$ [155], $B^0_s \to K^0_S K^+ \pi^-$ [156], $B^+ \to \pi^+ \pi^- \pi^+$ [78, 79] and $B^\pm \to K^+ K^- \pi^\pm$ [157]. The addition of the run II data will allow for more advanced analyses, for instance a time-dependent amplitude analysis of $B^0 \to K^0_S \pi^+ \pi^-$, or for time-integrated analyses of less abundant modes such as $B^0_s \to K^0_S \pi^+ \pi^-$. A time-integrated analysis of the decay $B^0 \to K^0_S K^+ K^-$ was attempted with run I data, see Ref. [158], but was inconclusive due to a large number of possible solutions for the isobar parameters (17 solutions were found). During my PhD, I dedicated a part of my time to trying to improve this result

by modifying the resonant model, by adding and/or removing resonant components with the aim of improving the fit quality and breaking the degeneracy. The addition of a $a_0(980)$ lineshape was tested, which appeared to decrease the number of solutions. However, due to a large destructive interference with the $f_0(980)$, the corresponding fit fractions became very high, and thus unphysical. The impact of adding information from flavour tagging was also assessed as well as the addition of more data. The main conclusion of this study was that the addition of the run II data will be critical, since the larger sample will both improve the discrimination and allow the model to be refined.

The second run of data-taking finished in late 2018 and analyses with the full run II dataset are ongoing. At the same time, the LHCb detector is being dismantled and will be replaced by an almost completely new detector [159] that will operate from 2021 at an expected instantaneous luminosity of $2 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The new design of the detector will allow it to be read at the full LHC bunch-crossing frequency of 40 MHz, a major improvement over the current data bandwidth limit of 1 MHz (which is mainly due to the readout and the hardware trigger). A total dataset corresponding to an integrated luminosity of $50\,\mathrm{fb}^{-1}$ (and with a substantially improved trigger efficiency after run II) is expected to be collected by the end of run IV, allowing for very precise measurements of many key observables, searching for small deviations from the SM predictions or looking for new phenomena.

Appendices

Appendix A

Supplementary material for the extraction of γ from three-body charmless decays of the B meson

A.1 Practical implementation of the different decay modes into Laura^{++} and cross checks

A.1.1 Introduction

In this appendix we give a detailed description of the implementation of the different amplitude models taken from BaBar analysis results into Laura^{++} [67]. Checks of the accuracy of the implementation are done comparing the values of the observables (fit fractions and CP asymmetries) extracted within our framework with the available results quoted in BaBar papers.

A.1.2 Normalisation factors

Most of the BaBar results quoted in the papers [71, 101–104], were not extracted using the Laura^{++} framework and use different conventions for the normalisation of the $F(s_1, s_2)$ distribution[1]. To account for this, we need to introduce normalisation factors which apply to each of the resonances of the different modes.

[1]Only the $B^+ \to K^+\pi^+\pi^-$ analysis was performed using Laura^{++} software.

$$\mathrm{N_i} = \sqrt{\frac{\frac{\mathrm{FF_i}}{\mathrm{FF_0}}}{\mathrm{a_i^2 + \overline{a}_i^2}}}, \tag{A.1}$$

where FF_i are the fit fractions as defined in part III.3.3, and a_i and $\overline{a}_i$ are the magnitudes of the isobar parameters quoted in the BABAR analyses cited earlier.

A.1.3 Computation of Q2B parameters within our framework

In order to check our implementation of the BABAR model we compare the quasi-two-body parameters (Q2B) we obtain in our framework with the ones quoted in the paper. The fit fraction for a intermediate state is defined as

$$FF_j \equiv \frac{\int\int_{\mathrm{DP}}(|A_j(s_1,s_2)|^2 + |\bar{A}_j(s_1,s_2)|^2)\mathrm{ds_1 ds_2}}{\int\int_{\mathrm{DP}}(|A(s_1,s_2)|^2 + |\bar{A}(s_1,s_2)|^2)\mathrm{ds_1 ds_2}}. \tag{A.2}$$

The interference fit fractions quantify the interferences between different states. They are given by

$$FF_{jk} \equiv 2\mathrm{Re}\frac{\int\int_{\mathrm{DP}}(\mathrm{A_j(s_1,s_2)A_k^*(s_1,s_2)} + \bar{\mathrm{A}}_\mathrm{j}(\mathrm{s_1,s_2})\bar{\mathrm{A}}_\mathrm{k}^*(\mathrm{s_1,s_2}))\mathrm{ds_1 ds_2}}{\int\int_{\mathrm{DP}}(|\mathrm{A(s_1,s_2)}|^2 + |\bar{\mathrm{A}}(\mathrm{s_1,s_2})|^2)\mathrm{ds_1 ds_2}}. \tag{A.3}$$

With this definition,

$$\sum_j FF_j + \sum_{j<k} FF_{jk} = 1. \tag{A.4}$$

The direct CP asymetry, $A_{CP}(j)$, for a given intermediate state writes

$$A_{CP}(j) \equiv \frac{\int\int_{\mathrm{DP}}(|\bar{A}_j(s_1,s_2)|^2 - |A_j(s_1,s_2)|^2)\mathrm{ds_1 ds_2}}{\int\int_{\mathrm{DP}}(|A(s_1,s_2)|^2 + |\bar{A}(s_1,s_2)|^2)\mathrm{ds_1 ds_2}}. \tag{A.5}$$

Since the lineshape terms $F(s_1,s_2)$ contain only the strong dynamics, we have

$$\iint_{\mathrm{DP}} |F(s_1,s_2)|^2 \mathrm{ds_1 ds_2} = \iint_{\mathrm{DP}} |\overline{\mathrm{F}}(\mathrm{s_1,s_2})|^2 \mathrm{ds_1 ds_2}, \tag{A.6}$$

we can thus express the direct CP asymmetry as a function of the isobar parameters

$$A_{CP}(j) = \frac{|\overline{c}_j|^2 - |c_j|^2}{|\overline{c}_j|^2 + |c_j|^2}. \tag{A.7}$$

As we will show in the following section, the results we obtained are consistent with the papers within the rounding errors.

$B^+ \to K^+\pi^+\pi^-$

From the five analyses, $K^+\pi^+\pi^-$ is the only one that was done using LAURA^{++} framework. This makes the implementation quite straightforward. The amplitude on the Dalitz plane is modelled using a Cartesian representation of the isobar parameters, given by

$$\begin{cases} A &= \sum_j [(x_j + \Delta x_j) + i(y_j + \Delta y_j)]F_j, \\ \bar{A} &= \sum_j [(x_j - \Delta x_j) + i(y_j - \Delta y_j)]\bar{F}_j. \end{cases} \tag{A.8}$$

The fitted values for the isobar parameters is given in table I

To check the validity of our implementation we compared the fit fractions and the $A_{CP}(j)$ we obtained with the ones in the paper. The results are consistent with the paper [101] and are summarised in tables II, III, IV and V. The Δx and Δy parameters for the non resonant and the $\omega(782)$ are fixed to zero in the fit, the value of their A_{CP} parameter is thus zero (cf. equation (A.7) and table V).

Table I: Results of fits to data for $B^+ \to K^+\pi^+\pi^-$ decay mode , with statistical, systematic, and model-dependent uncertainties. [101]

Resonance	x	y	Δx	Δy
$K^{*0}(892)\pi^+$	1.0 fixed	0.0 fixed	$-0.017\pm0.029\pm0.005\,^{+0.004}_{-0.006}$	$-0.238\pm0.228\pm0.062\,^{+0.144}_{-0.018}$
$(K\pi)^{*0}_0\pi^+$	$1.718\pm0.084\pm0.064\,^{+0.350}_{-0.055}$	$-0.727\pm0.108\pm0.080\,^{+0.331}_{-0.111}$	$-0.154\pm0.131\pm0.030\,^{+0.095}_{-0.010}$	$-0.285\pm0.337\pm0.091\,^{+0.221}_{-0.019}$
$\rho^0(770)K^+$	$0.683\pm0.075\pm0.045\,^{+0.015}_{-0.073}$	$-0.025\pm0.135\pm0.071\,^{+0.015}_{-0.073}$	$-0.160\pm0.049\pm0.024\,^{+0.094}_{-0.013}$	$0.169\pm0.096\pm0.057\,^{+0.133}_{-0.027}$
$f_0(980)K^+$	$-0.220\pm0.200\pm0.203\,^{+0.500}_{-0.095}$	$1.203\pm0.085\pm0.052\,^{+0.113}_{-0.045}$	$-0.109\pm0.143\pm0.087\,^{+0.037}_{-0.103}$	$0.047\pm0.045\pm0.012\,^{+0.046}_{-0.018}$
$\chi_{c0}K^+$	$-0.263\pm0.044\pm0.016\,^{+0.030}_{-0.014}$	$0.180\pm0.052\pm0.034\,^{+0.225}_{-0.022}$	$-0.033\pm0.049\pm0.012\,^{+0.017}_{-0.012}$	$-0.007\pm0.057\pm0.019\,^{+0.006}_{-0.111}$
Nonresonant	$-0.594\pm0.070\pm0.170\,^{+0.112}_{-0.035}$	$0.068\pm0.132\pm0.154\,^{+0.112}_{-0.099}$	0.0 fixed	0.0 fixed
$K^{*0}_2(1430)\pi^+$	$-0.301\pm0.060\pm0.030\,^{+0.012}_{-0.134}$	$0.424\pm0.060\pm0.045\,^{+0.012}_{-0.134}$	$0.032\pm0.078\pm0.024\,^{+0.057}_{-0.050}$	$0.007\pm0.086\pm0.017\,^{+0.025}_{-0.034}$
$\omega(782)K^+$	$-0.058\pm0.067\pm0.018\,^{+0.053}_{-0.011}$	$0.100\pm0.051\pm0.010\,^{+0.033}_{-0.032}$	0.0 fixed	0.0 fixed
$f_2(1270)K^+$	$-0.193\pm0.043\pm0.022\,^{+0.026}_{-0.033}$	$0.110\pm0.050\pm0.034\,^{+0.078}_{-0.073}$	$-0.089\pm0.046\pm0.019\,^{+0.034}_{-0.014}$	$0.125\pm0.058\pm0.021\,^{+0.034}_{-0.025}$
$f_X(1300)K^+$	$-0.290\pm0.047\pm0.064\,^{+0.047}_{-0.031}$	$-0.136\pm0.085\pm0.098\,^{+0.102}_{-0.031}$	$0.024\pm0.040\pm0.019\,^{+0.023}_{-0.018}$	$0.056\pm0.087\pm0.044\,^{+0.010}_{-0.036}$

Table II: $K^+\pi^+\pi^-$: Fit fractions as quoted in the paper [101] (%).

	$K^{*0}(892)$	$(K\pi)_0^{*0}$	$\rho^0(770)$	$f_0(980)$	χ_{c0}	NR	$K_2^{*0}(1430)$	$\omega(782)$	$f_2(1270)$	$f_X(1300)$
$K^{*0}(892)$	13.252	0.001	0.287	-0.116	-0.099	-0.000	0.000	-0.005	0.088	0.131
$(K\pi)_0^{*0}$	-	44.948	-0.068	-1.049	-0.115	1.468	-0.000	0.027	-0.501	-0.153
$\rho^0(770)$	-	-	6.543	0.000	-0.000	0.000	0.118	-0.371	0.000	0.000
$f_0(980)$	-	-	-	18.932	0.059	2.927	0.370	0.000	0.000	2.057
χ_{c0}	-	-	-	-	1.289	0.302	-0.054	-0.000	0.000	-0.038
NR	-	-	-	-	-	4.476	0.000	-0.000	0.000	-0.718
$K_2^{*0}(1430)$	-	-	-	-	-	-	3.399	-0.012	0.150	0.063
$\omega(782)$	-	-	-	-	-	-	-	0.168	-0.000	-0.000
$f_2(1270)$	-	-	-	-	-	-	-	-	0.912	0.000
$f_X(1300)$	-	-	-	-	-	-	-	-	-	1.333

Table III: $K^+\pi^+\pi^-$: Fit fractions obtained in our framework (%).

	$K^{*0}(892)$	$(K\pi)_0^{*0}$	$\rho^0(770)$	$f_0(980)$	χ_{c0}	NR	$K_2^{*0}(1430)$	$\omega(782)$	$f_2(1270)$	$f_X(1300)$
$K^{*0}(892)$	13.248	0.004	0.286	-0.114	-0.099	-0.001	-0.001	-0.004	0.088	0.131
$(K\pi)_0^{*0}$	-	44.936	-0.065	-1.049	-0.114	1.483	0.000	0.026	-0.501	-0.154
$\rho^0(770)$	-	-	6.534	0.000	-0.000	-0.000	0.117	-0.367	-0.000	-0.000
$f_0(980)$	-	-	-	18.923	0.059	2.925	0.372	0.000	-0.000	2.060
χ_{c0}	-	-	-	-	1.287	0.301	-0.054	0.000	0.000	-0.038
NR	-	-	-	-	-	4.481	-0.000	0.000	0.000	-0.718
$K_2^{*0}(1430)$	-	-	-	-	-	-	3.402	-0.012	0.151	0.063
$\omega(782)$	-	-	-	-	-	-	-	0.168	0.000	0.000
$f_2(1270)$	-	-	-	-	-	-	-	-	0.914	-0.000
$f_X(1300)$	-	-	-	-	-	-	-	-	-	1.333

Table IV: $K^+\pi^+\pi^-$: $(FF^{us} - FF^{paper}) \times 10^2$.

	$K^{*0}(892)$	$(K\pi)_0^{*0}$	$\rho^0(770)$	$f_0(980)$	χ_{c0}	NR	$K_2^{*0}(1430)$	$\omega(782)$	$f_2(1270)$	$f_X(1300)$
$K^{*0}(892)$	-0.004	0.003	-0.001	0.002	-0.000	-0.001	-0.001	0.001	-0.000	-0.000
$(K\pi)_0^{*0}$	-	-0.012	0.003	-0.000	0.001	0.015	0.000	-0.001	0.000	-0.001
$\rho^0(770)$	-	-	-0.009	0.000	-0.000	-0.000	-0.001	0.004	-0.000	-0.000
$f_0(980)$	-	-	-	-0.009	-0.000	-0.002	0.002	0.000	-0.000	0.003
χ_{c0}	-	-	-	-	-0.002	-0.001	0.000	0.000	0.000	-0.000
NR	-	-	-	-	-	0.005	-0.000	0.000	0.000	-0.000
$K_2^{*0}(1430)$	-	-	-	-	-	-	0.003	-0.000	0.001	0.000
$\omega(782)$	-	-	-	-	-	-	-	-0.000	0.000	0.000
$f_2(1270)$	-	-	-	-	-	-	-	-	0.002	-0.000
$f_X(1300)$	-	-	-	-	-	-	-	-	-	-0.000

Table V: $K^+\pi^+\pi^-$: A_{CP} (%).

	$K^{*0}(892)$	$(K\pi)_0^{*0}$	$\rho^0(770)$	$f_0(980)$	χ_{c0}	NR	$K_2^{*0}(1430)$	$\omega(782)$	$f_2(1270)$	$f_X(1300)$
Us	3.22	3.20	43.55	-10.67	-14.45	-	4.91	-	-84.85	27.42
Paper	3.2	3.2	44	-10.6	-14	-	5	-	-85	28
Residuals	0.02	0.00	0.45	-0.07	-0.45	-	0.09	-	0.15	0.58

$B^0 \to K^0_S \pi^+ \pi^-$

In this mode the amplitude is modelled the following way

$$\begin{cases} A &= \sum_j a_j e^{i\theta_j} F_j \\ \bar{A} &= \sum_j \bar{a}_j e^{i\bar{\theta}_j} \bar{F}_j. \end{cases} \tag{A.9}$$

The interference fit fractions are not quoted in BaBar paper. The results we obtained for the fit fractions and the A_{CP} are quoted table VII table VIII. They are consistent with the paper [102].

Table VI: Results of fit to data for the isobar amplitudes with statistical uncertainties. Both solutions are shown. [102]

	Solution I	
Isobar Amplitude	Magnitude	Phase (°)
$c_{f_0(980)K^0_S}$	4.0	0.0
$\bar{c}_{f_0(980)K^0_S}$	3.7 ± 0.4	-73.9 ± 19.6
$c_{\rho(770)K^0_S}$	0.10 ± 0.02	35.6 ± 14.9
$\bar{c}_{\rho(770)K^0_S}$	0.11 ± 0.02	15.3 ± 20.0
$c_{K^{*+}(892)\pi^-}$	0.154 ± 0.016	-138.7 ± 25.7
$\bar{c}_{K^{*-}(892)\pi^+}$	0.125 ± 0.015	163.1 ± 23.0
$c_{(K\pi)^{*+}_0\pi^-}$	6.9 ± 0.6	-151.7 ± 19.7
$\bar{c}_{(K\pi)^{*-}_0\pi^+}$	7.6 ± 0.6	136.2 ± 19.8
$c_{f_2(1270)K^0_S}$	0.014 ± 0.002	5.8 ± 19.2
$\bar{c}_{f_2(1270)K^0_S}$	0.011 ± 0.003	-24.0 ± 28.0
$c_{f_X(1300)K^0_S}$	1.41 ± 0.23	43.2 ± 22.0
$\bar{c}_{f_X(1300)K^0_S}$	1.24 ± 0.27	31.6 ± 23.0
c_{NR}	2.6 ± 0.5	35.3 ± 16.4
$\bar{c}_{NR}$	2.7 ± 0.6	36.1 ± 18.3
$c_{\chi_{c0}K^0_S}$	0.33 ± 0.15	61.4 ± 44.5
$\bar{c}_{\chi_{c0}K^0_S}$	0.44 ± 0.09	15.1 ± 30.0

Table VII: $K^0_S\pi^+\pi^-$: FF_j (%).

	$f_0(980)$	$\rho^0(770)$	$K^{*+}(892)$	$(K\pi)^{*+}_0$	$f_2(1270)$	$f_X(1300)$	NR	χ_{c0}
Us	13.71	8.54	10.93	44.91	2.29	3.58	11.43	1.03
Paper	13.8	8.6	11.0	45.2	2.3	3.6	11.5	1.04
Residuals	-0.09	-0.06	-0.07	-0.29	-0.01	-0.02	-0.07	-0.01

Table VIII: $K^0_S\pi^+\pi^-$: A_{CP}.

	$f_0(980)$	$\rho^0(770)$	$K^{*+}(892)$	$(K\pi)^{*+}_0$	$f_2(1270)$	$f_X(1300)$	NR	χ_{c0}
Us	-0.08	0.10	-0.21	0.10	-0.24	-0.13	0.04	0.28
Paper	-0.08	0.05	-0.21	0.09	-0.28	-0.13	-0.01	-0.29
Residuals	0.00	0.05	0.00	0.01	0.04	0.00	0.05	-0.01

$B^0 \to K^+\pi^-\pi^0$

In this mode the amplitude is given by

$$\begin{cases} A &= \sum_j a_j e^{i\theta_j} F_j, \\ \bar{A} &= \sum_j \bar{a}_j e^{i\bar{\theta}_j} \bar{F}_j. \end{cases} \tag{A.10}$$

The results we obtained on the fit fractions are summarised in tables IX, X and XI whereas the $A_{CP}(j)$ are given in table XII. They are all consistent with the paper [103].

Table IX: $K^+\pi^-\pi^0$: Fit fractions quoted in the paper (%)

	$\rho(770)^-$	$\rho(1450)^-$	$\rho(1700)^-$	$K^*(892)^+$	$(K\pi)^{*+}_0$	$K^*(892)^0$	$(K\pi)^{*0}_0$	NR
$\rho(770)^-$	17.61	7.22	0.88	0.47	-1.49	0.50	-0.78	0.00
$\rho(1450)^-$	-	6.34	-1.71	0.60	0.65	0.42	0.97	0.00
$\rho(1700)^-$	-	-	1.68	0.22	-0.72	0.23	-0.28	0.00
$K^*(892)^+$	-	-	-	7.05	0.00	-0.05	-0.10	0.00
$(K\pi)^{*+}_0$	-	-	-	-	30.30	-0.08	0.34	-0.08
$K^*(892)^0$	-	-	-	-	-	5.87	0.00	0.00
$(K\pi)^{*0}_0$	-	-	-	-	-	-	15.29	1.16
NR	-	-	-	-	-	-	-	7.49

Table X: $K^+\pi^-\pi^0$: Fit fractions obtained in our framework (%).

	$\rho(770)^-$	$\rho(1450)^-$	$\rho(1700)^-$	$K^*(892)^+$	$(Kpi)_0^{*+}$	$K^*(892)^0$	$(Kpi)_0^{*0}$	NR
$\rho(770)^-$	17.695	7.207	0.888	0.478	-1.498	0.506	-0.780	-0.000
$\rho(1450)^-$	-	6.299	-1.703	0.599	0.628	0.425	0.948	0.000
$\rho(1700)^-$	-	-	1.712	0.228	-0.733	0.238	-0.286	0.000
$K^*(892)^+$	-	-	-	7.149	-0.000	-0.046	-0.102	0.000
$(Kpi)_0^{*+}$	-	-	-	-	30.261	-0.077	0.331	-0.060
$K^*(892)^0$	-	-	-	-	-	5.978	-0.002	0.000
$(Kpi)_0^{*0}$	-	-	-	-	-	-	15.199	1.134
NR	-	-	-	-	-	-	-	7.385

Table XI: $K^+\pi^-\pi^0$: $(FF^{us} - FF^{paper}) \times 10^2$.

	$\rho(770)^-$	$\rho(1450)^-$	$\rho(1700)^-$	$K^*(892)^+$	$(Kpi)_0^{*+}$	$K^*(892)^0$	$(Kpi)_0^{*0}$	NR
$\rho(770)^-$	0.085	-0.013	0.008	0.008	-0.008	0.006	0.000	-0.000
$\rho(1450)^-$	-	-0.041	0.007	-0.001	-0.022	0.005	-0.022	0.000
$\rho(1700)^-$	-	-	0.032	0.008	-0.013	0.008	-0.006	0.000
$K^*(892)^+$	-	-	-	0.099	-0.000	0.004	-0.002	0.000
$(Kpi)_0^{*+}$	-	-	-	-	-0.039	0.003	-0.009	0.020
$K^*(892)^0$	-	-	-	-	-	0.108	-0.002	0.000
$(Kpi)_0^{*0}$	-	-	-	-	-	-	-0.091	-0.026
NR	-	-	-	-	-	-	-	-0.105

Table XII: $K^+\pi^-\pi^0$: $A_{CP}(j)$.

	$\rho(770)^-$	$\rho(1450)^-$	$\rho(1700)^-$	$K^*(892)^+$	$(K\pi)_0^{*+}$	$K^*(892)^0$	$(K\pi)_0^{*0}$	NR
Us	0.196	-0.092	-0.346	-0.289	0.075	-0.150	-0.153	0.101
Paper	0.2	-0.1	-0.36	-0.29	0.07	-0.15	-0.15	0.1
Residuals	-0.004	0.008	0.014	0.001	0.005	-0.000	-0.003	0.001

$B^0 \to K^0_S K^+ K^-$

In this mode the amplitude is modelled the following way

$$\begin{cases} A &= \sum_j c_j(1+b_j)e^{i(\phi_j+\delta_j)}F_j, \\ \bar{A} &= \sum_j c_j(1-b_j)e^{i(\phi_j-\delta_j)}\bar{F}_j. \end{cases} \tag{A.11}$$

The results we obtained on the fit fractions are summarised in tables XV, XVI and XVII whereas the $A_{CP}(j)$ are given in table XVIII. All the results are consistent with the paper [71].

Table XIII: *CP*-conserving isobar parameters for $B^0 \to K^0_S K^+ K^-$, for the global minimum. The NR coefficients are defined in . Phases are given in degrees. Only statistical uncertainties are given. [71]

Parameter		Value
$\phi(1020)K^0_S$	c	0.039 ± 0.005
	ϕ	20 ± 19
$f_0(980)K^0_S$	c	2.2 ± 0.5
	ϕ	40 ± 16
$f_0(1500)K^0_S$	c	0.22 ± 0.05
	ϕ	17 ± 16
$f'_2(1525)K^0_S$	c	0.00080 ± 0.00028
	ϕ	53 ± 23
$f_0(1710)K^0_S$	c	0.72 ± 0.11
	ϕ	110 ± 11
$\chi_{c0}K^0_S$	c	0.144 ± 0.023
	ϕ	-17 ± 29
NR		
a_{S0}	c	1.0 (fixed)
	ϕ	0 (fixed)
a_{S1}	c	1.25 ± 0.25
	ϕ	-149 ± 9
a_{S2}	c	0.58 ± 0.22
	ϕ	56 ± 15
a_{P0}	c	1.22 ± 0.22
	ϕ	65 ± 13
a_{P1}	c	0.28 ± 0.18
	ϕ	-68 ± 28
a_{P2}	c	0.42 ± 0.16
	ϕ	-131 ± 25

Table XIV: *CP*-violating parameters $\beta_{\rm eff}$, $\mathcal{A}^{CP}$, and S for $B^0 \to K^0_S K^+ K^-$. Central values correspond to the global minimum. Statistical uncertainties for $\beta_{\rm eff}$ and $\mathcal{A}^{CP}$ are determined from likelihood scans. [71]

Component	$\beta_{\rm eff}$ (deg)	$\mathcal{A}^{CP}(=-C)(\%)$	S
$\phi(1020)IK^0_S$	$21 \pm 6 \pm 2$	$-5 \pm 18 \pm 5$	$0.66 \pm 0.17 \pm 0.07$
$f_0(980)K^0_S$	$18 \pm 6 \pm 4$	$-28 \pm 24 \pm 9$	$0.55 \pm 0.18 \pm 0.12$
Other	$20.3 \pm 4.3 \pm 1.2$	$-2 \pm 9 \pm 3$	$0.65 \pm 0.12 \pm 0.03$

Table XV: $K^0_S K^+ K^-$: Fit fractions as quoted in the paper (%).

	$\phi(1020)$	$f_0(980)$	$f_0(1500)$	$f'_2(1525)$	$f_0(1710)$	χ_{c0}	NR (S)	NR (P)
$\phi(1020)$	13.1	0.0	0.0	0.0	0.0	0.0	0.0	0.2
$f_0(980)$	-	26.3	0.1	-0.0	14.4	-0.7	-81.2	0.0
$f_0(1500)$	-	-	2.1	-0.0	5.3	-0.1	-0.7	0.0
$f'_2(1525)$	-	-	-	0.5	-0.0	0.0	0.0	0.0
$f_0(1710)$	-	-	-	-	16.7	-0.2	-27.0	0.0
χ_{c0}	-	-	-	-	-	3.4	1.6	0.0
NR (S)	-	-	-	-	-	-	114.5	0.0
NR (P)	-	-	-	-	-	-	-	11.7

Table XVI: $K^0_S K^+ K^-$: Fit fractions obtained in our framework (%).

	$\phi(1020)$	$f_0(980)$	$f_0(1500)$	$f'_2(1525)$	$f_0(1710)$	χ_{c0}	NR (S)	NR (P)
$\phi(1020)$	13.23	0.00	-0.00	-0.00	-0.00	-0.00	-0.00	0.03
$f_0(980)$	-	26.56	-0.04	-0.00	14.31	-0.74	-81.74	0.00
$f_0(1500)$	-	-	2.12	0.00	5.26	-0.10	-0.95	-0.00
$f'_2(1525)$	-	-	-	0.50	0.00	0.00	0.00	-0.00
$f_0(1710)$	-	-	-	-	16.87	-0.16	-27.66	-0.00
χ_{c0}	-	-	-	-	-	3.43	1.60	0.00
NR (S)	-	-	-	-	-	-	115.65	0.00
NR (P)	-	-	-	-	-	-	-	11.82

Table XVII: $K^0_S K^+ K^-$: $(FF^{us} - FF^{paper}) \times 10^2$.

	$\phi(1020)$	$f_0(980)$	$f_0(1500)$	$f'_2(1525)$	$f_0(1710)$	χ_{c0}	NR (S)	NR (P)
$\phi(1020)$	0.130	0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.168
$f_0(980)$	-	0.260	-0.137	-0.000	-0.087	-0.045	-0.536	0.000
$f_0(1500)$	-	-	0.021	0.000	-0.037	0.002	-0.253	-0.000
$f'_2(1525)$	-	-	-	0.005	0.000	0.000	0.000	-0.000
$f_0(1710)$	-	-	-	-	0.165	0.041	-0.659	-0.000
χ_{c0}	-	-	-	-	-	0.034	-0.003	0.000
NR (S)	-	-	-	-	-	-	1.150	0.000
NR (P)	-	-	-	-	-	-	-	0.117

Table XVIII: $K^0_S K^+ K^-$: $A_{CP}(j)$ (%).

	$\phi(1020)$	$f_0(980)$	$f_0(1500)$	$f'_2(1525)$	$f_0(1710)$	χ_{c0}	NR (S)	NR (P)
Us	-6.0	-27.5	-2.0	-2.0	-2.0	-2.0	-2.0	-2.0
Paper	-5	-28	-2	-2	-2	-2	-2	-2
Residuals	-1.0	-0.5	0.0	0.0	0.0	0.0	0.0	0.0

$B^0 \to K^0_S K^0_S K^0_S$

In this mode the amplitude is modelled the following way

$$A = \sum_j a_j e^{i\theta_j} F_j. \tag{A.12}$$

The results we obtained on the fit fractions are summarised in tables XX, XXI and XXII. Due to a different choice of conventions, there is a factor of 2 between FF_{jk} formula in the paper and our formula, the fit fractions we quoted in this note are in accordance with formula (A.3). In this analysis no difference was made between B^0 and $\bar{B}^0$, there is thus no $A_{CP}(j)$.

The results are in agreement with the paper [104].

Table XIX: Summary of measurements of the quasi-two-body parameters. The quoted uncertainties are statistical only. The change in the log-likelihood ($-2\Delta \ln \mathcal{L}$) corresponds to the case where the magnitude of the amplitude of the resonance is set to 0. This number is used for the estimation of the statistical significance of each resonance. [104]

Mode	Parameter	Solution 1	Solution 2
$f_0(980)K^0_S$	FF	$0.44^{+0.20}_{-0.19}$	$1.03^{+0.22}_{-0.17}$
	Phase [rad]	0.09 ± 0.16	1.26 ± 0.17
	$-2\Delta \ln \mathcal{L}$	11.7	-
	Significance [σ]	3.0	-
$f_0(1710)K^0_S$	FF	$0.07^{+0.07}_{-0.03}$	$0.09^{+0.05}_{-0.02}$
	Phase [rad]	1.11 ± 0.23	0.36 ± 0.20
	$-2\Delta \ln \mathcal{L}$	14.2	-
	Significance [σ]	3.3	-
$f_2(2010)K^0_S$	FF	$0.09^{+0.03}_{-0.03}$	0.10 ± 0.02
	Phase [rad]	2.50 ± 0.20	1.58 ± 0.22
	$-2\Delta \ln \mathcal{L}$	14.0	-
	Significance [σ]	3.3	-
NR	FF	$2.16^{+0.36}_{-0.37}$	$1.37^{+0.26}_{-0.21}$
	Phase [rad]	0.0	0.0
	$-2\Delta \ln \mathcal{L}$	68.1	-
	Significance [σ]	8.0	-
$\chi_{c0}K^0_S$	FF	$0.07^{+0.04}_{-0.02}$	0.07 ± 0.02
	Phase [rad]	0.63 ± 0.47	-0.24 ± 0.52
	$-2\Delta \ln \mathcal{L}$	18.5	-
	Significance [σ]	3.9	-
	Total FF	$2.84^{+0.71}_{-0.66}$	$2.66^{+0.35}_{-0.27}$

Table XX: $K^0_S K^0_S K^0_S$: Fit fractions as quoted in the paper (using our conventions).

	$f_0(980)$	$f_0(1710)$	$f_2(2010)$	NR	χ_{c0}
$f_0(980)$	0.44	0.14	-0.04	-1.6	0.02
$f_0(1710)$	-	0.07	-0.02	-0.34	-0.0006
$f_2(2010)$	-	-	0.09	-0.04	0.0004
NR	-	-	-	2.16	-0.04
χ_{c0}	-	-	-	-	0.07

Table XXI: $K^0_S K^0_S K^0_S$: Fit fractions obtained in our framework.

	$f_0(980)$	$f_0(1710)$	$f_2(2010)$	NR	χ_{c0}
$f_0(980)$	0.4579	0.1432	-0.0185	-1.6744	0.0182
$f_0(1710)$	-	0.0729	-0.0046	-0.3617	-0.0010
$f_2(2010)$	-	-	0.0937	-0.0241	0.0008
NR	-	-	-	2.2480	-0.0230
χ_{c0}	-	-	-	-	0.0729

Table XXII: $K^0_S K^0_S K^0_S$: $FF^{us} - FF^{paper}$.

	$f_0(980)$	$f_0(1710)$	$f_2(2010)$	NR	χ_{c0}
$f_0(980)$	0.0179	0.0032	0.0215	-0.0744	-0.0018
$f_0(1710)$	-	0.0029	0.0154	-0.0217	-0.0004
$f_2(2010)$	-	-	0.0037	-0.0641	0.0004
NR	-	-	-	0.0880	0.0170
χ_{c0}	-	-	-	-	0.0029

Appendix B

Supplementary material for the branching fraction measurements of $B^0_{(s)} \to K^0_S h^\pm h'^\mp$ modes

B.1 Trigger and stripping lines included in the selection

The list of the various trigger and stripping lines included in the analysis presented in chapter VI are shown in this section for completeness.

Table I: Trigger lines included in the selection. For a given year, a logical OR is implied between the different lines of the same trigger level, that is to say L0, HLT1 and HLT2, and a logical AND is implied between the different levels.

Year	Trigger requirements	HLT1 trigger requirements	HLT2 trigger requirements
2011	B_L0HadronDecision_TOS	B_Hlt1TrackAllL0Decision_TOS	B_Hlt2Topo2BodyBBDTDecision_TOS
	B_L0DiMuonDecision_TIS		B_Hlt2Topo3BodyBBDTDecision_TOS
	B_L0MuonDecision_TIS		B_Hlt2Topo4BodyBBDTDecision_TOS
	B_L0ElectronDecision_TIS		B_Hlt2Topo2BodySimpleDecision_TOS
	B_L0PhotonDecision_TIS		B_Hlt2Topo3BodySimpleDecision_TOS
	B_L0HadronDecision_TIS		B_Hlt2Topo4BodySimpleDecision_TOS
2012	idem 2011	idem 2011	B_Hlt2Topo2BodyBBDTDecision_TOS
			B_Hlt2Topo3BodyBBDTDecision_TOS
			B_Hlt2Topo4BodyBBDTDecision_TOS
2016	B_L0HadronDecision_TOS	B_Hlt1TrackMVADecision_TOS	B_Hlt2Topo2BodyDecision_TOS
	B_L0DiMuonDecision_TIS	B_Hlt1TwoTrackMVADecision_TOS	B_Hlt2Topo3BodyDecision_TOS
	B_L0MuonDecision_TIS		B_Hlt2Topo4BodyDecision_TOS
	B_L0ElectronDecision_TIS		
	B_L0PhotonDecision_TIS		
	B_L0HadronDecision_TIS		
	B_L0MuonEWDecision_TIS		
	B_L0JetElDecision_TIS		
	B_L0JetPhDecision_TIS		

Table II: Stripping lines included in the run I selection. Different lines are used for the two K^0_S reconstruction categories. A hyphen in the table means that this cut is not included in the line and "idem" means that the cut is the same for both DD and LL lines. DOCA is the distance of closest approach and DIRA corresponds to the cosine of the direction angle.

Cut step	Candidate(s)	StrippingB2KShhDDLine	StrippingB2KShhLLLine
DaughterCuts	K^0_S	$p(K^0_S) > 6000.0\,\text{MeV}$	-
		$\lvert m_{\pi^+\pi^-} - m^{\text{PDG}}_{K^0_S}\rvert < 30.0\,\text{MeV}$	$\lvert m_{\pi^+\pi^-} - m^{\text{PDG}}_{K^0_S}\rvert < 20.0$
		$\chi^2_{\text{vtx}}(K^0_S) < 12.0$	idem
		$\chi^2_{\text{FD}}(K^0_S) > 50.0$	$\chi^2_{\text{FD}}(K^0_S) > 80.0$
	K^0_S daughters	$p(\pi^\pm) > 2.0\,\text{GeV}$	idem
		$\min\chi^2_{\text{IP}}\ (\pi^\pm) > 4$	$\min\chi^2_{\text{IP}}\ (\pi^\pm) > 9$
		-	$p_{\text{T}}(\pi^\pm) > 250\,\text{MeV}$
		DOCA χ^2 of K^0_S daughters< 25	idem
		-	track $\chi^2/\text{ndf} < 4.0$
		-	track ghost probability < 0.5
	$h^\pm$	track $\chi^2/\text{ndf} < 4.0$	idem
		track ghost probability < 0.5	idem
		$p_{\text{T}}(h^\pm) > 250\,\text{MeV}$	idem
		$\min\chi^2_{\text{IP}}\ (h^\pm) > 4$	idem
CombinationCut	B, K^0_S, $h^\pm$	$p_{\text{T}}(B) > 1000.0\,\text{MeV}$	idem
		$p_{\text{T}}(K^0_S) + p_{\text{T}}(h^+) + p_{\text{T}}(h^-) > 4200.0\,\text{MeV}$	$p_{\text{T}}(K^0_S) + p_{\text{T}}(h^+) + p_{\text{T}}(h^-) > 3000.0\,\text{MeV}$
		at least 2 daughters with $p_{\text{T}} > 800\,\text{MeV}$	idem
		$(4000 < m_{K^0_S h^+ h'^-} < 6200)\,\text{MeV}$	idem
		`AVAL_MAX(MIPDV(PRIMARY),PT)>0.05`	idem
		DOCA χ^2 of any pair of daughters< 5	idem
MotherCut	B, K^0_S, $h^\pm$	$p_{\text{T}}(B) > 1500.0\,\text{MeV}$	idem
		$\chi^2_{\text{vtx}}(B) < 12.0$	idem
		DIRA$(B) > 0.999$	DIRA> 0.9999
		$\min\chi^2_{\text{IP}}\ (B) < 6.0$	$\min\chi^2_{\text{IP}}\ (B) < 8.0$
		B flight distance w.r.t. any PV > 1.7 mm	B flight distance > 1.0 mm
		$\chi^2_{\text{FD}}(B) > 50$	idem
		$\chi^2_{\text{IP}}\ (h^+) + \chi^2_{\text{IP}}\ (h^-) > 50$	-

Table III: Stripping lines included in the run II selection. Different lines are used for the two K^0_S reconstruction categories. A hyphen in the table means that this cut is not included in the line and "idem" means that the cut is the same for both DD and LL lines. DOCA is the distance of closest approach and DIRA corresponds to the cosine of the direction angle.

Cut step	Candidate(s)	`StrippingB2KShh_DD_Run2_OS_Line`	`StrippingB2KShh_LL_Run2_OS_Line`
TriggerCuts	HLT1	`Hlt1TrackMVADecision`	idem
		`Hlt1TwoTrackMVADecision`	idem
		`Hlt1IncPhiDecision`(Hlt1PhiIncPhiDecision)	idem
	HLT2	`Hlt2Topo{2,3,4}BodyDecision`	idem
		`Hlt2IncPhiDecision`(Hlt2PhiIncPhiDecision)	idem
DaughterCuts	K^0_S	$p(K^0_S) > 6000.0\,\text{MeV}$	-
		$\lvert m_{\pi^+\pi^-} - m^{\text{PDG}}_{K^0_S}\rvert < 30.0\,\text{MeV}$	$\lvert m_{\pi^+\pi^-} - m^{\text{PDG}}_{K^0_S}\rvert < 20.0\,\text{MeV}$
		$\chi^2_{\text{vtx}}(K^0_S) < 12.0$	idem
		$\chi^2_{\text{FD}}(K^0_S) > 50.0$	$\chi^2_{\text{FD}}(K^0_S) > 80.0$
	K^0_S daughters	$p(\pi^\pm) > 2.0\,\text{GeV}$	idem
		$\min\chi^2_{\text{IP}}\ (\pi^\pm) > 4$	$\min\chi^2_{\text{IP}}\ (\pi^\pm) > 9$
		-	$p_{\text{T}}(\pi^\pm) > 250\,\text{MeV}$
		DOCA χ^2 of K^0_S daughters< 25	idem
		track $\chi^2/\text{ndf} < 4.0$	idem
		-	track ghost probability < 0.5
	$h_{1,2}$	track $\chi^2/\text{ndf} < 4.0$	idem
		track ghost probability < 0.5	idem
CombinationCut		$p_{\text{T}}(B) > 1000.0\,\text{MeV}$	idem
		$p_{\text{T}}(K^0_S) + p_{\text{T}}(h^+) + p_{\text{T}}(h^-) > 4200.0\,\text{MeV}$	$p_{\text{T}}(K^0_S) + p_{\text{T}}(h^+) + p_{\text{T}}(h^-) > 3000.0\,\text{MeV}$
		at least 2 daughters with $p_{\text{T}} > 800\,\text{MeV}$	idem
		$(4000 < m_{K^0_S h^+ h'^-} < 6200)\,\text{MeV}$	idem
		DOCA χ^2 between pairs of daughters < 25	idem
MotherCut		$p_{\text{T}}(B) > 1500.0\,\text{MeV}$	idem
		$\chi^2_{\text{vtx}}(B) < 12.0$	idem
		DIRA$(B) > 0.999$	idem
		$\chi^2_{\text{FD}}(B) > 5$	idem
		$Z_{\text{vtx}}(K^0_S) - Z_{\text{vtx}}(B) > 15.0\,\text{mm}$	idem
		$\chi^2_{\text{IP}}\ (\pi^\pm) < 6.0$	-
		sum of the χ^2_{IP} of the daughters w.r.t their PV > 50.0	idem

B.2 Results of the simultaneous fit to the run I data sample

The plots corresponding to the result of the fit to the run I data are shown here for all the categories, in linear and logarithmic scale.

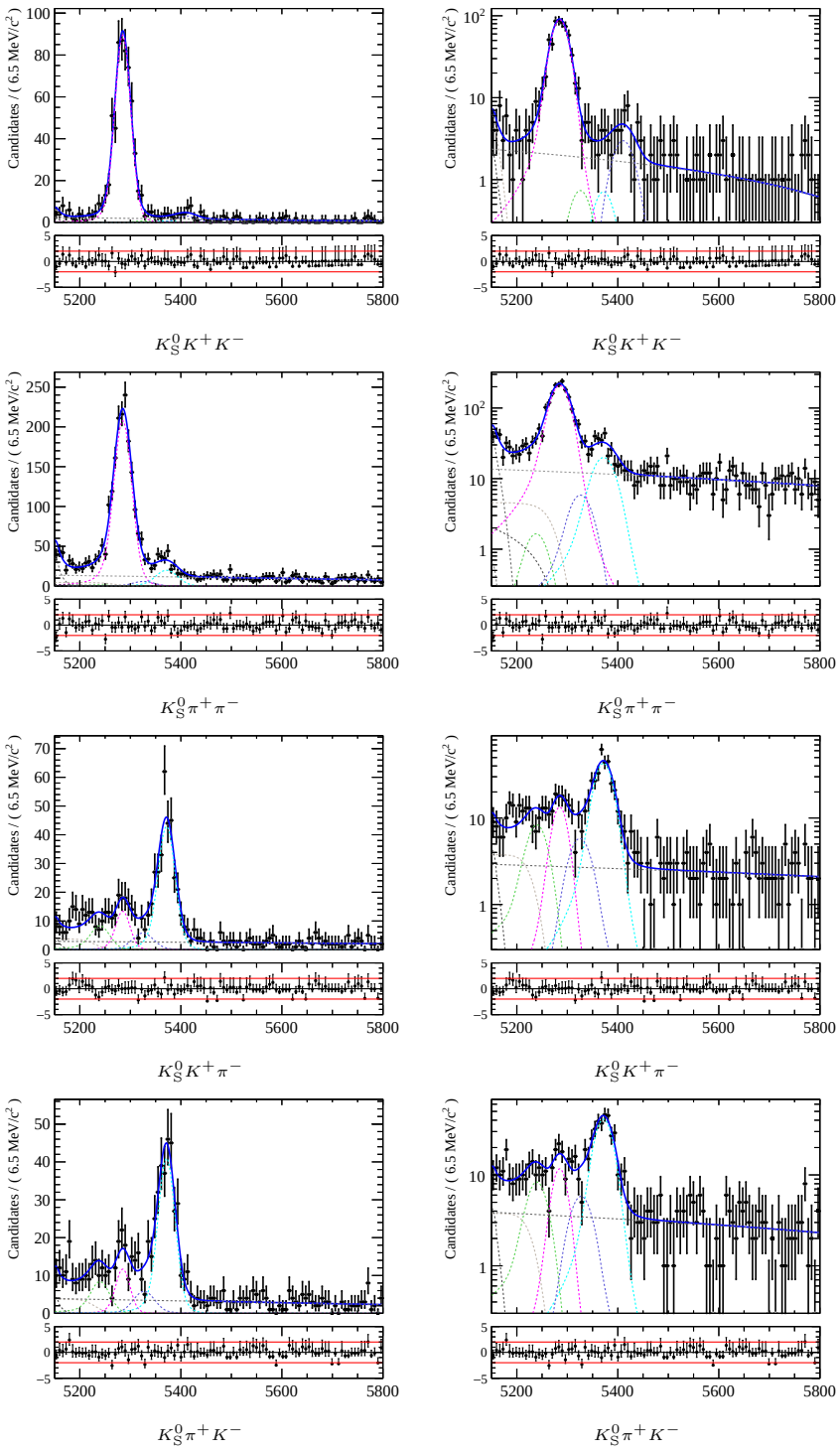


Figure B.1: Simultaneous fit to the data in the 2012b DD samples, shown on linear (left) and logarithmic (right) scales. The four spectra, $K^0_S K^+ K^-$, $K^0_S \pi^+ \pi^-$, $K^0_S K^+ \pi^-$ and $K^0_S \pi^+ K^-$, are displayed from top to bottom. On each plot, the total PDF is shown in solid blue and the individual components are shown as dashed lines: the B^0 signal in pink, the B^0_s signal in light blue, the crossfeeds from B^0 in green and those from B^0_s in purple, the combinatorial background in grey, and the partially reconstructed backgrounds in grey to the left.

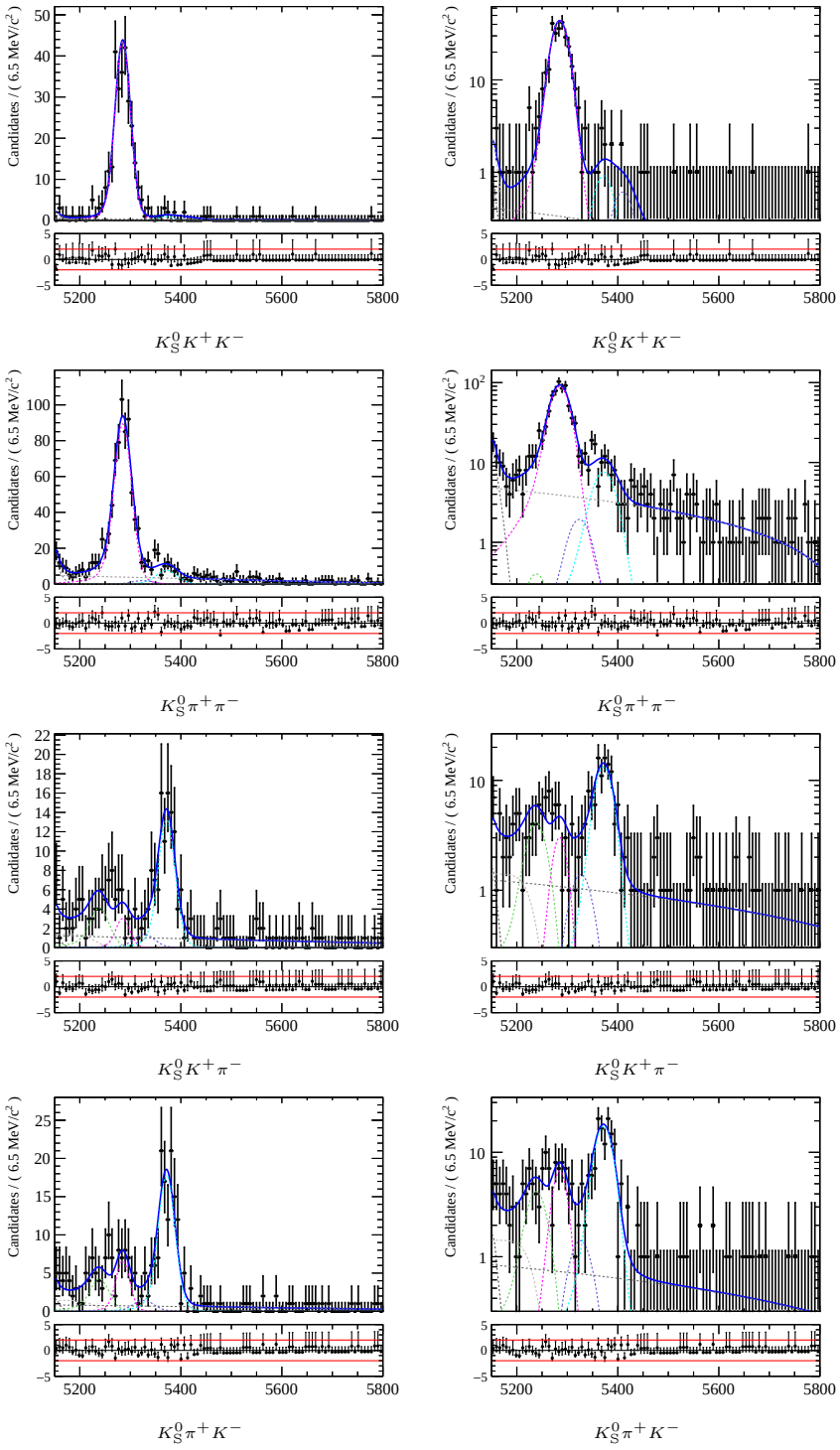


Figure B.2: Simultaneous fit to the data in the 2012b LL samples, shown on linear (left) and logarithmic (right) scales. The four spectra, $K^0_S K^+ K^-$, $K^0_S \pi^+ \pi^-$, $K^0_S K^+ \pi^-$ and $K^0_S \pi^+ K^-$, are displayed from top to bottom. On each plot, the total PDF is shown in solid blue and the individual components are shown as dashed lines: the B^0 signal in pink, the B^0_s signal in light blue, the crossfeeds from B^0 in green and those from B^0_s in purple, the combinatorial background in grey, and the partially reconstructed backgrounds in grey to the left.

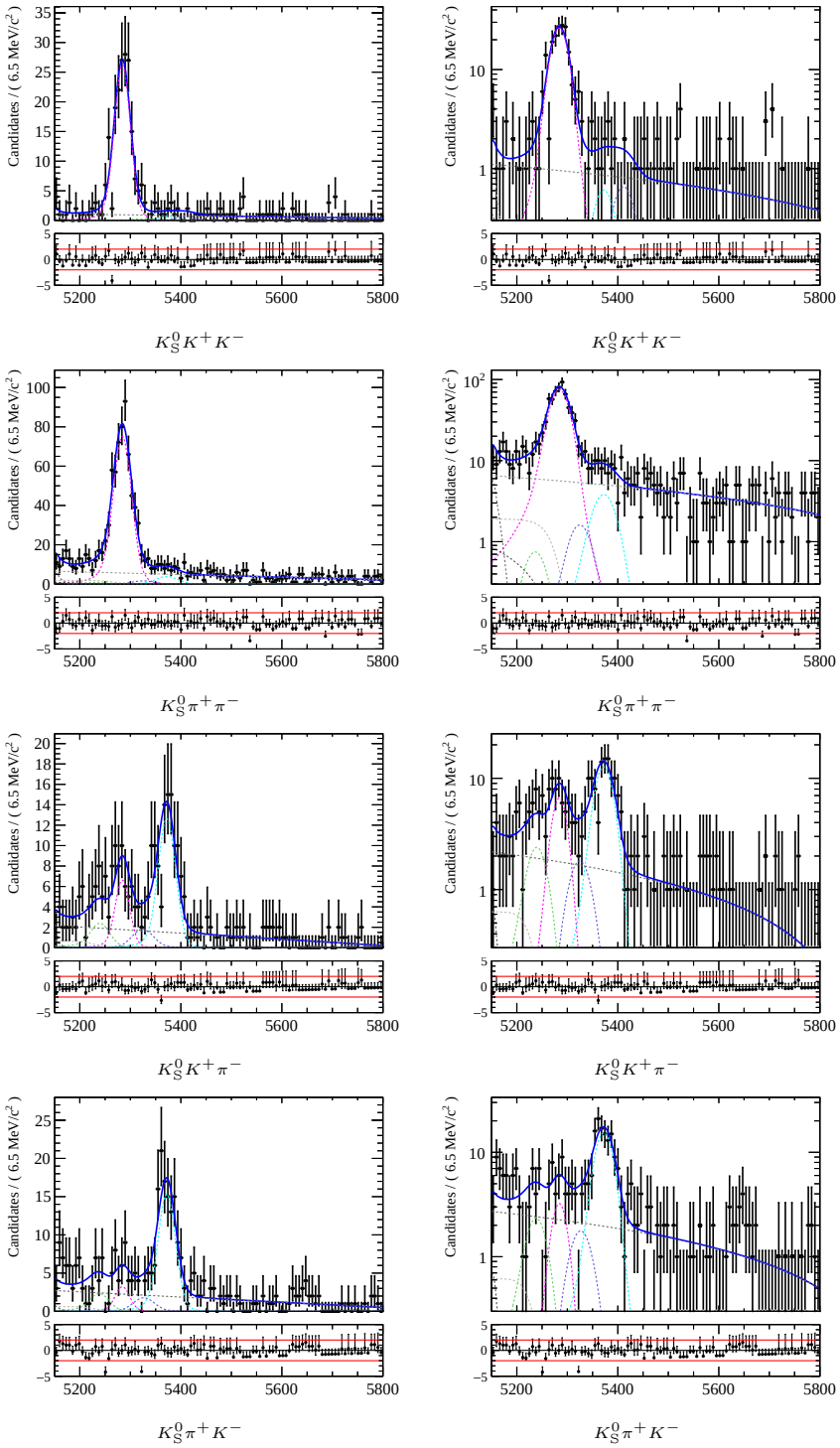


Figure B.3: Simultaneous fit to the data in the 2012a DD samples, shown on linear (left) and logarithmic (right) scales. The four spectra, $K^0_S K^+ K^-$, $K^0_S \pi^+ \pi^-$, $K^0_S K^+ \pi^-$ and $K^0_S \pi^+ K^-$, are displayed from top to bottom. On each plot, the total PDF is shown in solid blue and the individual components are shown as dashed lines: the B^0 signal in pink, the B^0_s signal in light blue, the crossfeeds from B^0 in green and those from B^0_s in purple, the combinatorial background in grey, and the partially reconstructed backgrounds in grey to the left.

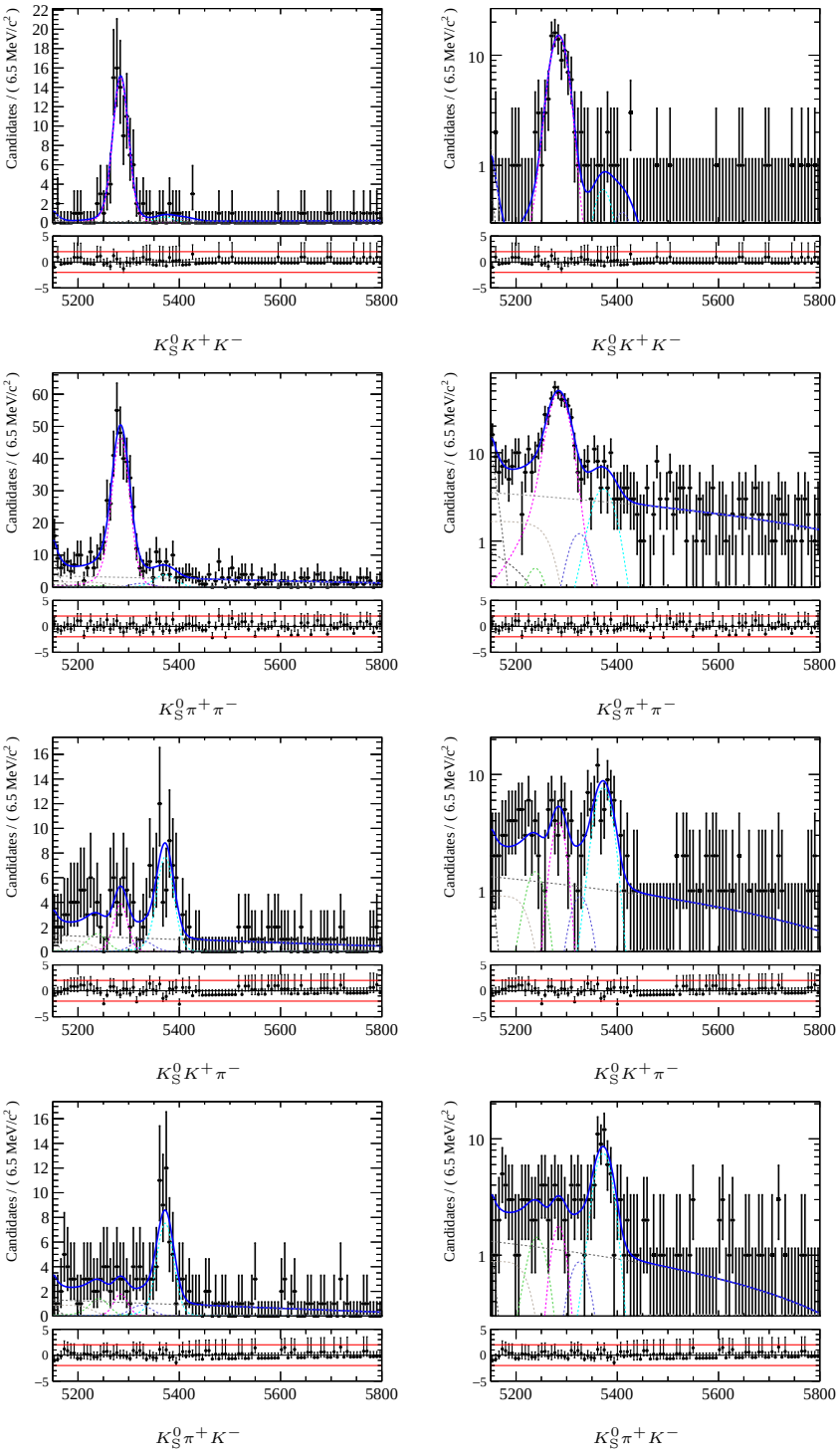


Figure B.4: Simultaneous fit to the data in the 2012a LL samples, shown on linear (left) and logarithmic (right) scales. The four spectra, $K^0_S K^+ K^-$, $K^0_S \pi^+ \pi^-$, $K^0_S K^+ \pi^-$ and $K^0_S \pi^+ K^-$, are displayed from top to bottom. On each plot, the total PDF is shown in solid blue and the individual components are shown as dashed lines: the B^0 signal in pink, the B^0_s signal in light blue, the crossfeeds from B^0 in green and those from B^0_s in purple, the combinatorial background in grey, and the partially reconstructed backgrounds in grey to the left.

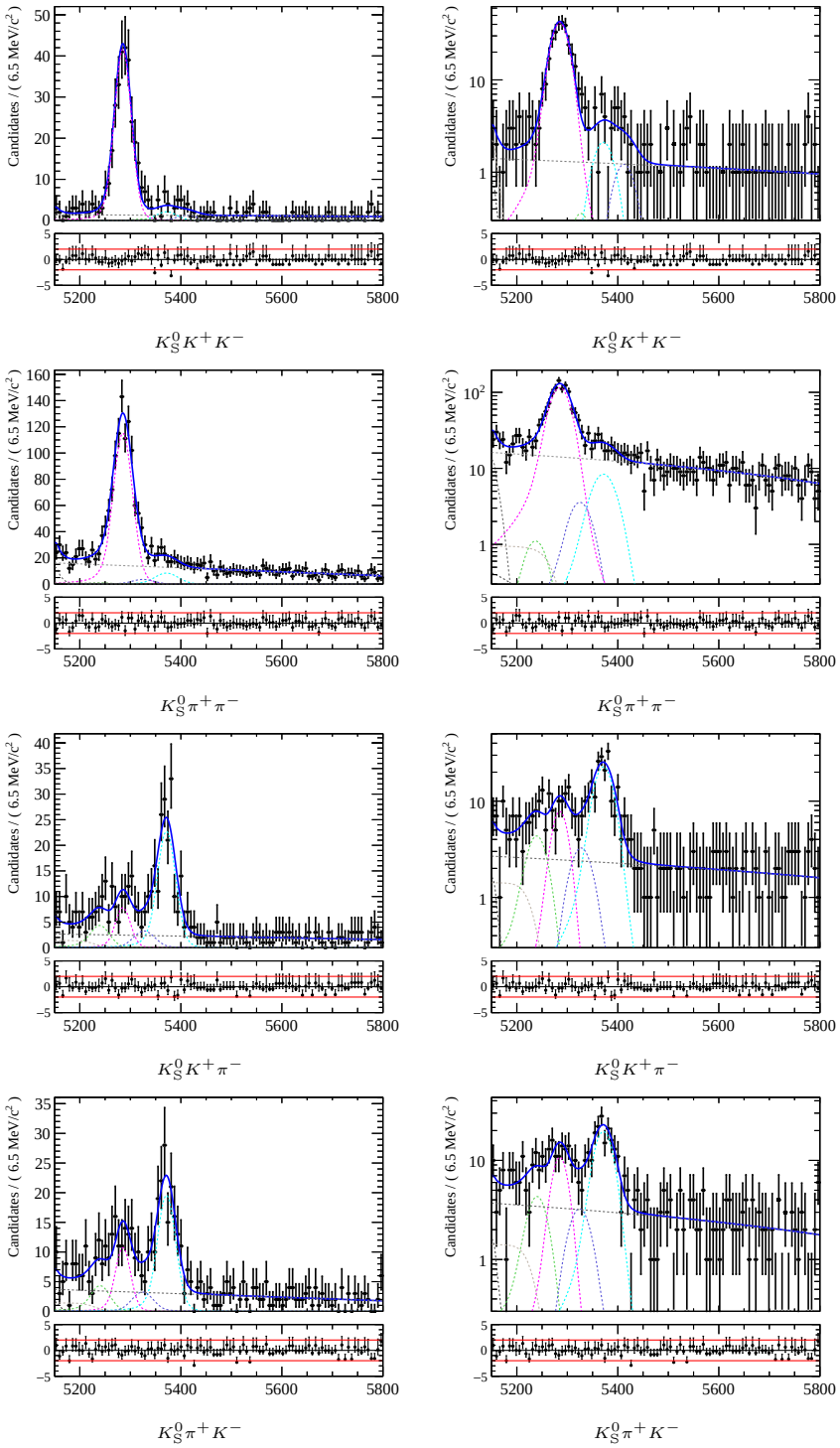


Figure B.5: Simultaneous fit to the data in the 2011 DD samples, shown on linear (left) and logarithmic (right) scales. The four spectra, $K^0_S K^+ K^-$, $K^0_S \pi^+ \pi^-$, $K^0_S K^+ \pi^-$ and $K^0_S \pi^+ K^-$, are displayed from top to bottom. On each plot, the total PDF is shown in solid blue and the individual components are shown as dashed lines: the B^0 signal in pink, the B^0_s signal in light blue, the crossfeeds from B^0 in green and those from B^0_s in purple, the combinatorial background in grey, and the partially reconstructed backgrounds in grey to the left.

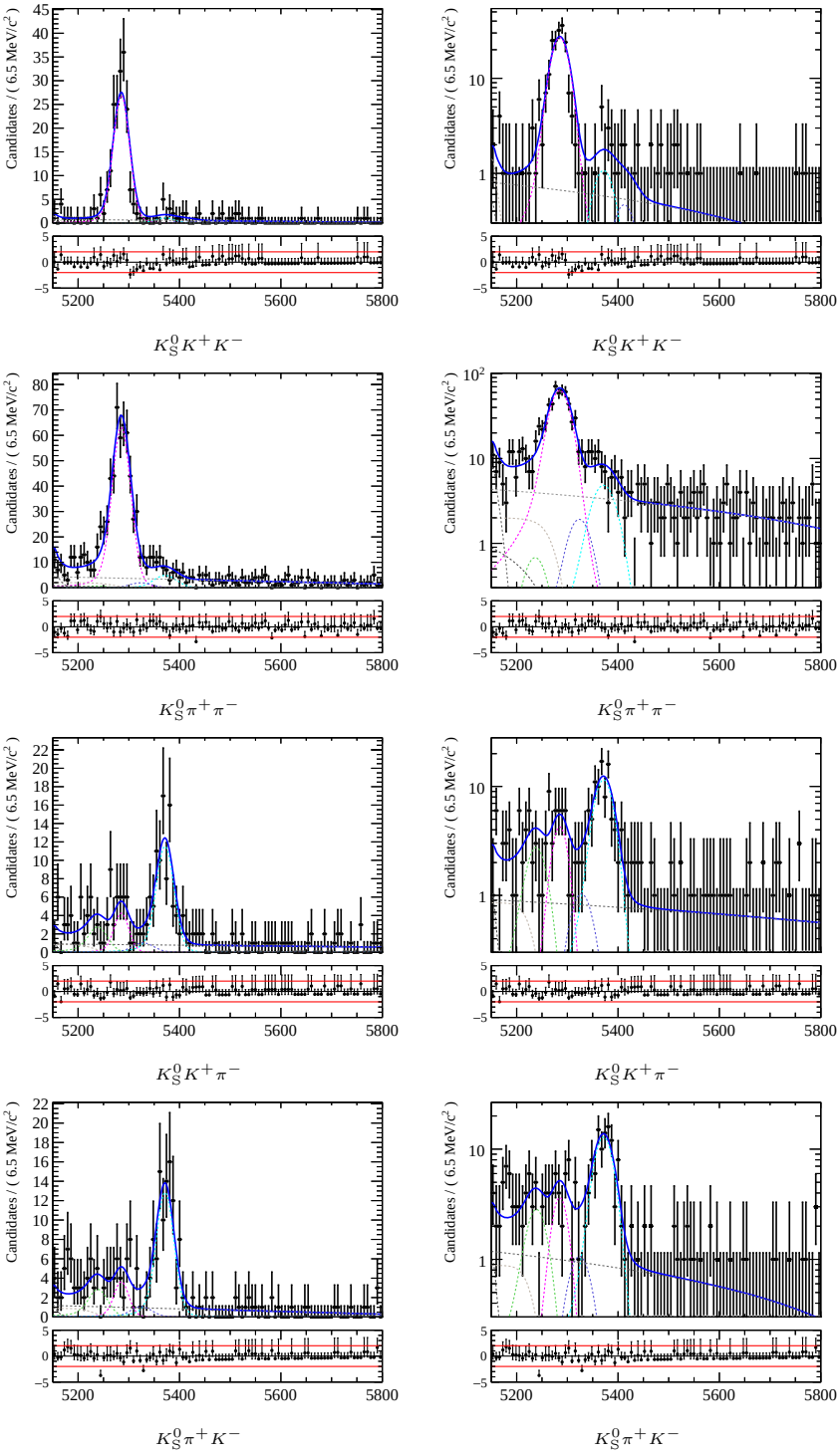


Figure B.6: Simultaneous fit to the data in the 2011 LL samples, shown on linear (left) and logarithmic (right) scales. The four spectra, $K^0_S K^+ K^-$, $K^0_S \pi^+ \pi^-$, $K^0_S K^+ \pi^-$ and $K^0_S \pi^+ K^-$, are displayed from top to bottom. On each plot, the total PDF is shown in solid blue and the individual components are shown as dashed lines: the B^0 signal in pink, the B^0_s signal in light blue, the crossfeeds from B^0 in green and those from B^0_s in purple, the combinatorial background in grey, and the partially reconstructed backgrounds in grey to the left.

Bibliography

[1] **CDF** Collaboration, F. Abe *et al.*, "Observation of top quark production in $\bar{p}p$ collisions," *Phys. Rev. Lett.* **74** (1995) 2626–2631, arXiv:hep-ex/9503002 [hep-ex].

[2] **DONUT** Collaboration, K. Kodama *et al.*, "Observation of tau neutrino interactions," *Phys. Lett.* **B504** (2001) 218–224, arXiv:hep-ex/0012035 [hep-ex].

[3] **ATLAS** Collaboration, G. Aad *et al.*, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Phys. Lett.* **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[4] **CMS** Collaboration, S. Chatrchyan *et al.*, "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC," *Phys. Lett.* **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[5] A. D. Sakharov, "Violation of *CP* in variance, *C* asymmetry, and baryon asymmetry of the universe," *Soviet Physics Uspekhi* **34** no. 5, (May, 1991) 392–393.

[6] T. Mannel, "Theory and phenomenology of *CP* violation," *Nuclear Physics B - Proceedings Supplements* **167** (2007) 170 – 174. Proceedings of the 7th International Conference on Hyperons, Charm and Beauty Hadrons.

[7] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.

[8] O. W. Greenberg, "Spin and unitary-spin independence in a paraquark model of baryons and mesons," *Phys. Rev. Lett.* **13** (Nov, 1964) 598–602.

[9] T. D. Lee and C. N. Yang, "Mass degeneracy of the heavy mesons," *Phys. Rev.* **102** (Apr, 1956) 290–291.

[10] **LHCb** Collaboration, R. Aaij *et al.*, "Observation of *CP* violation in charm decays," *Phys. Rev. Lett.* **122** no. 21, (2019) 211803, `arXiv:1903.08726 [hep-ex]`.

[11] G. Luders, "On the equivalence of invariance under time reversal and under particle-antiparticle conjugation for relativistic field theories," *Kong. Dan. Vid. Sel. Mat. Fys. Med.* **28N5** no. 5, (1954) 1–17.

[12] W. Pauli, "Exclusion principle, Lorentz group and reflection of space-time and charge,". In: Enz C.P., v. Meyenn K. (eds) Wolfgang Pauli. Vieweg+Teubner Verlag.

[13] **BaBar** Collaboration, J. P. Lees *et al.*, "Observation of Time Reversal Violation in the B^0 Meson System," *Phys. Rev. Lett.* **109** (2012) 211801, `arXiv:1207.5832 [hep-ex]`.

[14] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509.

[15] **Particle Data Group** Collaboration, M. Tanabashi *et al.*, "Review of Particle Physics," *Phys. Rev.* **D98** no. 3, (2018) 030001.

[16] S. L. Glashow, "The renormalizability of vector meson interactions," *Nucl. Phys.* **10** (1959) 107–117.

[17] A. Salam and J. C. Ward, "Weak and electromagnetic interactions," *Il Nuovo Cimento (1955-1965)* **11** no. 4, (Feb, 1959) 568–577.

[18] S. Weinberg, "A model of leptons," *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266.

[19] M. Kobayashi and T. Maskawa, "CP Violation in the Renormalizable Theory of Weak Interaction," *Prog. Theor. Phys.* **49** (1973) 652–657.

[20] N. Cabibbo, "Unitary symmetry and leptonic decays," *Phys. Rev. Lett.* **10** (Jun, 1963) 531–533.

[21] Z. Maki, M. Nakagawa, and S. Sakata, "Remarks on the Unified Model of Elementary Particles," *Progress of Theoretical Physics* **28** no. 5, (11, 1962) 870–880.

[22] B. Pontecorvo, "Inverse beta processes and nonconservation of lepton charge," *Sov. Phys. JETP* **7** (1958) 172–173. [Zh. Eksp. Teor. Fiz.34,247(1957)].

[23] L.-L. Chau and W.-Y. Keung, "Comments on the Parametrization of the Kobayashi-Maskawa Matrix," *Phys. Rev. Lett.* **53** (1984) 1802.

[24] L. Wolfenstein, "Parametrization of the kobayashi-maskawa matrix," *Phys. Rev. Lett.* **51** (Nov, 1983) 1945–1947.

[25] C. Jarlskog, "Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal CP nonconservation," *Phys. Rev. Lett.* **55** (Sep, 1985) 1039–1042.

[26] K. Lande, E. T. Booth, J. Impeduglia, L. M. Lederman, and W. Chinowsky, "Observation of long-lived neutral V particles," *Phys. Rev.* **103** (Sep, 1956) 1901–1904.

[27] **ARGUS** Collaboration, H. Albrecht *et al.*, "Observation of $B^0-\overline{B}^0$ mixing," *Physics Letters B* **192** no. 1, (1987) 245 – 252.

[28] **CDF Collaboration** Collaboration, A. Abulencia *et al.*, "Observation of $B_s^0 - \overline{B}_s^0$ oscillations," *Phys. Rev. Lett.* **97** (Dec, 2006) 242003.

[29] **BABAR Collaboration** Collaboration, B. Aubert *et al.*, "Evidence for $D^0-\overline{D}^0$ mixing," *Phys. Rev. Lett.* **98** (May, 2007) 211802.

[30] **Belle Collaboration** Collaboration, M. Starič *et al.*, "Evidence for $D^0-\overline{D}^0$ mixing," *Phys. Rev. Lett.* **98** (May, 2007) 211803.

[31] **BaBar, Belle** Collaboration, A. J. Bevan *et al.*, "The Physics of the B Factories," *Eur. Phys. J.* **C74** (2014) 3026, `arXiv:1406.6311 [hep-ex]`.

[32] **BaBar** Collaboration, B. Aubert *et al.*, "Measurement of time-dependent CP asymmetry in $B^0 \to c\bar{c}K^{(*)0}$ decays," *Phys. Rev.* **D79** (2009) 072009, arXiv:0902.1708 [hep-ex].

[33] I. Adachi *et al.*, "Precise measurement of the *cp* violation parameter sin2ϕ_1 in $B^0 \to (c\bar{c})K^0$ decays," *Phys. Rev. Lett.* **108** (2012) 171802, arXiv:1201.4643 [hep-ex].

[34] **LHCb** Collaboration, R. Aaij *et al.*, "Measurement of CP Violation in $B^0 \to J/\psi K^0_S$ Decays," *Phys. Rev. Lett.* **115** no. 3, (2015) 031601, arXiv:1503.07089 [hep-ex].

[35] **Belle** Collaboration, R. Itoh *et al.*, "Studies of CP violation in $B \to J/\Psi K^*$ decays," *Phys. Rev. Lett.* **95** (2005) 091601, arXiv:hep-ex/0504030 [hep-ex].

[36] **BaBar** Collaboration, B. Aubert *et al.*, "Ambiguity-free measurement of cos(2β): Time-integrated and time-dependent angular analyses of $B \to J/\psi K\pi$," *Phys. Rev.* **D71** (2005) 032005, arXiv:hep-ex/0411016 [hep-ex].

[37] **CKMfitter group** Collaboration, J. Charles *et al.*, "Current status of the standard model CKM fit and constraints on $\Delta F = 2$ new physics," *Phys. Rev.* **D91** (2015) 073007, arXiv:1501.05013 [hep-ph]. updated results and plots available at http://ckmfitter.in2p3.fr/.

[38] **HFLAV** Collaboration, Y. Amhis *et al.*, "Averages of b-hadron, c-hadron, and τ-lepton properties as of summer 2016," *Eur. Phys. J.* **C77** no. 12, (2017) 895, arXiv:1612.07233 [hep-ex].

[39] S. Weinberg, "Phenomenological lagrangians," *Physica A: Statistical Mechanics and its Applications* **96** no. 1, (1979) 327 – 340.

[40] K. G. Wilson and J. Kogut, "The renormalization group and the ϵ expansion," *Physics Reports* **12** no. 2, (1974) 75 – 199.

[41] T. Appelquist and J. Carazzone, "Infrared singularities and massive fields," *Phys. Rev. D* **11** (May, 1975) 2856–2861.

[42] K. G. Wilson, "Renormalization group and strong interactions," *Phys. Rev. D* **3** (Apr, 1971) 1818–1846.

[43] M. Beneke, G. Buchalla, M. Neubert, and C. Sachrajda, "QCD factorization for exclusive non-leptonic b-meson decays: general arguments and the case of heavy-light final states," *Nuclear Physics B* **591** no. 1, (2000) 313 – 418.

[44] M. Beneke, G. Buchalla, M. Neubert, and C. T. Sachrajda, "QCD factorization for $B \to \pi\pi$ decays: Strong phases and CP violation in the heavy quark limit," *Phys. Rev. Lett.* **83** (Sep, 1999) 1914–1917.

[45] M. Beneke, G. Buchalla, M. Neubert, and C. Sachrajda, "QCD factorization in $B \to \pi K$, $\pi\pi$ decays and extraction of wolfenstein parameters," *Nuclear Physics B* **606** no. 1, (2001) 245 – 321.

[46] M. Beneke and M. Neubert, "QCD factorization for $B \to PP$ and $B \to PV$ decays," *Nuclear Physics B* **675** no. 1, (2003) 333 – 415.

[47] Y.-Y. Keum, H. nan Li, and A. Sanda, "Fat penguins and imaginary penguins in perturbative QCD," *Physics Letters B* **504** no. 1, (2001) 6 – 14.

[48] Y.-Y. Keum, H.-n. Li, and A. I. Sanda, "Penguin enhancement and $B \to K\pi$ decays in perturbative qcd," *Phys. Rev. D* **63** (Feb, 2001) 054008.

[49] Y.-Y. Keum and H.-n. Li, "Nonleptonic charmless B decays: Factorization versus perturbative QCD," *Phys. Rev. D* **63** (Mar, 2001) 074006.

[50] C.-D. Lü, K. Ukai, and M.-Z. Yang, "Branching ratio and CP violation of $B \to \pi\pi$ decays in the perturbative QCD approach," *Phys. Rev. D* **63** (Mar, 2001) 074009.

[51] C.-D. Lü and M.-Z. Yang, "$B \to \pi\rho, \pi\omega$ decays in perturbative QCD approach," *The European Physical Journal C - Particles and Fields* **23** no. 2, (Mar, 2002) 275–287.

[52] C. W. Bauer, S. Fleming, and M. Luke, "Summing sudakov logarithms in $B \to X_s\gamma$ in effective field theory," *Phys. Rev. D* **63** (Dec, 2000) 014006.

[53] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, "An effective field theory for collinear and soft gluons: Heavy to light decays," *Phys. Rev. D* **63** (May, 2001) 114020.

[54] C. W. Bauer and I. W. Stewart, "Invariant operators in collinear effective theory," *Physics Letters B* **516** no. 1, (2001) 134 – 142.

[55] D. Zeppenfeld, "SU(3) relations for B-meson decays," *Zeitschrift für Physik C Particles and Fields* **8** no. 1, (Mar, 1981) 77–84.

[56] R. H. Dalitz and S. F. Tuan, "The phenomenological description of K-nucleon reaction processes," *Annals Phys.* **10** (1960) 307–351.

[57] **BaBar** Collaboration, J. P. Lees *et al.*, "Measurement of CP-violating asymmetries in $B^0 \to (\rho\pi)^0$ decays using a time-dependent Dalitz plot analysis," *Phys. Rev.* **D88** no. 1, (2013) 012003, arXiv:1304.3503 [hep-ex].

[58] G. N. Fleming, "Recoupling effects in the isobar model. i. general formalism for three-pion scattering," *Phys. Rev.* **135** (Jul, 1964) B551–B560.

[59] D. Morgan, "Phenomenological analysis of $i = \frac{1}{2}$ single-pion production processes in the energy range 500 to 700 mev," *Phys. Rev.* **166** (Feb, 1968) 1731–1759.

[60] D. J. Herndon, P. Söding, and R. J. Cashmore, "Generalized isobar model formalism," *Phys. Rev. D* **11** (Jun, 1975) 3165–3182.

[61] **LHCb** Collaboration, R. Aaij *et al.*, "Study of the $B^0 \to \rho(770)^0 K^*(892)^0$ decay with an amplitude analysis of $B^0 \to (\pi^+\pi^-)(K^+\pi^-)$ decays," *JHEP* **05** (2019) 026, arXiv:1812.07008 [hep-ex].

[62] C. Zemach, "Three-pion decays of unstable particles," *Phys. Rev.* **133** (Mar, 1964) B1201–B1220.

[63] C. Zemach, "Use of angular-momentum tensors," *Phys. Rev.* **140** (Oct, 1965) B97–B108.

[64] J. M. Blatt and V. F. Weisskopf, *Theoretical nuclear physics*. Springer, New York, 1952.

[65] M. Jacob and G. Wick, "On the general theory of collisions for particles with spin," *Annals of Physics* **7** no. 4, (1959) 404 – 428.

[66] A. Anisovich, E. Klempt, A. Sarantsev, and U. Thoma, "Partial wave decomposition of pion and photoproduction amplitudes," *Eur. Phys. J.* **A24** (2005) 111–128, arXiv:hep-ph/0407211 [hep-ph].

[67] J. Back *et al.*, "Laura++ : a Dalitz plot fitter," *Computer Physics Communications* **231** (2018) 198, arXiv:1711.09854 [hep-ex].

[68] S. Flatté, "Coupled-channel analysis of the $\pi\eta$ and $K\overline{K}$ systems near $K\overline{K}$ threshold," *Physics Letters B* **63** no. 2, (1976) 224 – 227.

[69] G. J. Gounaris and J. J. Sakurai, "Finite-width corrections to the vector-meson-dominance prediction for $\rho \to e^+e^-$," *Phys. Rev. Lett.* **21** (Jul, 1968) 244–247.

[70] D. Aston *et al.*, "A Study of $K^-\pi^+$ Scattering in the Reaction $K^-p \to K^-\pi^+n$ at 11-GeV/c," *Nucl. Phys.* **B296** (1988) 493–526.

[71] **BABAR** Collaboration, J. P. Lees *et al.*, "Study of CP violation in Dalitz-plot analyses of B $\to K^+K^-K^0_s$, $B \to K^+K^-K^+$, $B \to K^0_sK^0_sK^+$," *Phys. Rev.* **D85** (2012) 112010, arXiv:1201.5897 [hep-ex].

[72] **LHCb collaboration** Collaboration, R. Aaij *et al.*, "Amplitude analysis of $B^\pm \to \pi^\pm K^+K^-$ decays," LHCb-PAPER-2018-051 CERN-EP-2019-062, arXiv:1905.09244 [hep-ex]. submitted to Phys. Rev. Lett.

[73] J. R. Peláez and F. J. Ynduráin, "Pion-pion scattering amplitude," *Phys. Rev. D* **71** (Apr, 2005) 074016.

[74] J. H. A. Nogueira, I. Bediaga, A. B. R. Cavalcante, T. Frederico, and O. Lourenço, "*cp* violation: Dalitz interference, *cpt*, and final state interactions," *Phys. Rev. D* **92** (Sep, 2015) 054010.

[75] S. U. Chung, J. Brose, R. Hackmann, E. Klempt, S. Spanier, and C. Strassburger, "Partial wave analysis in k-matrix formalism," *Annalen der Physik* **507** no. 5, 404–430.

[76] I. Aitchison, "The k-matrix formalism for overlapping resonances," *Nuclear Physics A* **189** no. 2, (1972) 417 – 423.

[77] **LHCb collaboration** Collaboration, R. Aaij *et al.*, "Measurement of *CP* violation in the three-body phase space of charmless $B^\pm$

decays," *Phys. Rev.* **D90** (2014) 112004 CERN-PH-EP-2014-203, LHCb-PAPER-2014-044, arXiv:1408.5373 [hep-ex].

[78] **LHCb collaboration** Collaboration, R. Aaij *et al.*, "Amplitude Analysis of the $B^+ \to \pi^+\pi^+\pi^-$ decay," LHCb-PAPER-2019-017. in preparation.

[79] **LHCb collaboration** Collaboration, R. Aaij *et al.*, "Observation of several sources of *CP* violation in $B^+ \to \pi^+\pi^+\pi^-$ decays," LHCb-PAPER-2019-018. in preparation.

[80] J. Brod, A. Lenz, G. Tetlalmatzi-Xolocotzi, and M. Wiebusch, "New physics effects in tree-level decays and the precision in the determination of the quark mixing angle γ," *Phys. Rev.* **D92** no. 3, (2015) 033002, arXiv:1412.1446 [hep-ph].

[81] M. Gronau and D. London, "How to determine all the angles of the unitarity triangle from $B^0_d \to DK_S$ and $B^0_s \to D\phi$," *Phys. Lett.* **B253** (1991) 483–488.

[82] M. Gronau and D. Wyler, "On determining a weak phase from CP asymmetries in charged B decays," *Phys. Lett.* **B265** (1991) 172–176.

[83] D. Atwood, I. Dunietz, and A. Soni, "Enhanced CP violation with $B \to KD^0(\bar{D}^0)$ modes and extraction of the CKM angle γ," *Phys. Rev. Lett.* **78** (1997) 3257–3260, arXiv:hep-ph/9612433 [hep-ph].

[84] J. Brod and J. Zupan, "The ultimate theoretical error on γ from $B \to DK$ decays," *JHEP* **01** (2014) 051, arXiv:1308.5663 [hep-ph].

[85] A. Giri, Y. Grossman, A. Soffer, and J. Zupan, "Determining γ using $B^\pm \to DK^\pm$ with multibody D decays," *Phys. Rev. D* **68** (Sep, 2003) 054018.

[86] Y. Grossman, Z. Ligeti, and A. Soffer, "Measuring γ in $B^\pm \to K^\pm \left(KK^*\right)_D$ decays," *Phys. Rev. D* **67** (Apr, 2003) 071301.

[87] I. Dunietz and R. G. Sachs, "Asymmetry between inclusive charmed and anticharmed modes in B^0, $\overline{B}^0$ decay as a measure of *CP* violation," *Phys. Rev. D* **37** (Jun, 1988) 3186–3192.

[88] R. Aleksan, I. Dunietz, and B. Kayser, "Determining thecp-violating phase γ," *Zeitschrift für Physik C Particles and Fields* **54** no. 4, (Dec, 1992) 653–659.

[89] R. Fleischer, "New strategies to obtain insights into cp violation through $B_s^0 \to D_s^{\pm} K^{\mp}, D_s^{*\pm} K^{\mp}$, ..., $B^0 \to D^{\pm}\pi^{\mp}, D_s^{\pm}\pi^{\mp}$ decays," *Nuclear Physics B* **671** (2003) 459 – 482.

[90] **LHCb Collaboration** Collaboration, "Update of the LHCb combination of the CKM angle γ," Tech. Rep. LHCb-CONF-2018-002. CERN-LHCb-CONF-2018-002, CERN, Geneva, May, 2018.

[91] N. Rey-Le Lorier, M. Imbeault, and D. London, "Diagrammatic analysis of charmless three-body B decays," *Phys. Rev. D.* **84** (2011) 034040, `arXiv:1011.4972 [hep-ph]`.

[92] M. Neubert and J. L. Rosner, "New bound on γ from $B^{\pm} \to \pi K$ decays," *Phys. Lett.* **B441** (1998) 403–409, `arXiv:hep-ph/9808493 [hep-ph]`.

[93] M. Gronau, D. Pirjol, and T. M. Yan, "Model independent electroweak penguins in B decays to two pseudoscalars," *Phys. Rev.* **D60** (1999) 034021, `arXiv:hep-ph/9810482 [hep-ph]`. [Erratum: Phys. Rev.D69,119901(2004)].

[94] M. Imbeault, N. Rey-Le Lorier, and D. London, "Measuring γ in $B \to K\pi\pi$ Decays," *Phys. Rev.* **D84** (2011) 034041, `arXiv:1011.4973 [hep-ph]`.

[95] B. Bhattacharya, M. Gronau, M. Imbeault, D. London, and J. L. Rosner, "Charmless B→PPP decays: The fully-symmetric final state," *Phys. Rev.* **D89** no. 7, (2014) 074043, `arXiv:1402.2909 [hep-ph]`.

[96] M. Ciuchini, M. Pierini, and L. Silvestrini, "Hunting the CKM weak phase with time-integrated dalitz analyses of $B_s \to K\pi\pi$ decays," *Physics Letters B* **645** no. 2, (2007) 201 – 203.

[97] B. Bhattacharya and D. London, "Using U spin to extract γ from charmless $B \to PPP$ decays," *JHEP* **04** (2015) 154, `arXiv:1503.00737 [hep-ph]`.

[98] B. Bhattacharya, M. Imbeault, and D. London, "Extraction of the CP-violating phase γ using $B \to K\pi\pi$ and $B \to KK\bar{K}$ decays," *Phys. Lett.* **B728** (2014) 206–209, arXiv:1303.0846 [hep-ph].

[99] E. Bertholet, E. Ben-Haim, B. Bhattacharya, M. Charles, and D. London, "Extraction of the ckm phase γ using charmless three-body decays of B mesons," *Phys. Rev. D* **99** (Jun, 2019) 114011.

[100] N. Rey-Le Lorier and D. London, "Measuring γ with $B \to K\pi\pi$ and $B \to KK\bar{K}$ Decays," *Phys. Rev.* **D85** (2012) 016010, arXiv:1109.0881 [hep-ph].

[101] **BABAR** Collaboration, B. Aubert *et al.*, "Evidence for Direct CP Violation from Dalitz-plot analysis of $B^{\pm} \to K^{\pm}\pi^{\mp}\pi^{\pm}$," *Phys. Rev.* **D78** (2008) 012004, arXiv:0803.4451 [hep-ex].

[102] **BABAR** Collaboration, B. Aubert *et al.*, "Time-dependent amplitude analysis of $B^0 \to K^0_S\pi + \pi-$," *Phys. Rev.* **D80** (2009) 112001, arXiv:0905.3615 [hep-ex].

[103] **BABAR** Collaboration, J. P. Lees *et al.*, "Amplitude Analysis of $B^0 \to K^+\pi^-\pi^0$ and Evidence of Direct CP Violation in $B \to K^*\pi$ decays," *Phys. Rev.* **D83** (2011) 112010, arXiv:1105.0125 [hep-ex].

[104] **BABAR** Collaboration, J. P. Lees *et al.*, "Amplitude analysis and measurement of the time-dependent CP asymmetry of $B^0 \to K^0_S K^0_S K^0_S$ decays," *Phys. Rev.* **D85** (2012) 054023, arXiv:1111.3636 [hep-ex].

[105] F. James and M. Roos, "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations," *Comput. Phys. Commun.* **10** (1975) 343–367.

[106] F. James and M. Winkler, "MINUIT User's Guide,".

[107] E. Ben-Haim, R. Brun, B. Echenard, and T. E. Latham, "JFIT: a framework to obtain combined experimental results through joint fits," arXiv:1409.5080 [physics.data-an].

[108] L. Evans and P. Bryant, "LHC machine," *Journal of Instrumentation* **3** no. 08, (Aug, 2008) S08001–S08001.

[109] **The ATLAS Collaboration** Collaboration, and G Aad *et al.*, "The ATLAS experiment at the CERN large hadron collider," *Journal of Instrumentation* **3** no. 08, (Aug, 2008) S08003–S08003.

[110] **The CMS Collaboration** Collaboration, S. Chatrchyan *et al.*, "The CMS experiment at the CERN LHC," *Journal of Instrumentation* **3** no. 08, (Aug, 2008) S08004–S08004.

[111] **The ALICE Collaboration** Collaboration, K. Aamodt *et al.*, "The ALICE experiment at the CERN LHC," *Journal of Instrumentation* **3** no. 08, (Aug, 2008) S08002–S08002.

[112] **LHCb collaboration** Collaboration, A. A. Alves *et al.*, "The LHCb detector at the LHC," *Journal of Instrumentation* **3** no. 08, (Aug, 2008) S08005–S08005.

[113] F. Marcastel, "CERN's Accelerator Complex. La chaîne des accélérateurs du CERN,". General Photo.

[114] **LHCb collaboration** Collaboration, R. Aaij *et al.*, "Absolute luminosity measurements with the LHCb detector at the LHC," *JINST* **7** (2012) P01010 LHCb-PAPER-2011-015, CERN-PH-EP-2011-157, `arXiv:1110.2866 [hep-ex]`.

[115] **LHCb collaboration** Collaboration, "LHCb reoptimized detector design and performance: Technical Design Report," 2003.

[116] **LHCb collaboration** Collaboration, R. Aaij *et al.*, "LHCb detector performance," *Int. J. Mod. Phys.* **A30** (2015) 1530022 LHCB-DP-2014-002, CERN-PH-EP-2014-290, `arXiv:1412.6352 [hep-ex]`.

[117] R. Aaij *et al.*, "Performance of the LHCb Vertex Locator," *JINST* **9** (2014) P09007 LHCB-DP-2014-001, `arXiv:1405.7808 [physics.ins-det]`.

[118] **LHCb collaboration** Collaboration, "LHCb magnet: Technical Design Report," 2000.

[119] **LHCb collaboration** Collaboration, "LHCb inner tracker: Technical Design Report," 2002.

[120] **LHCb collaboration** Collaboration, "LHCb outer tracker: Technical Design Report," 2001.

[121] **LHCb Outer Tracker Group** Collaboration, R. Arink *et al.*, "Performance of the LHCb Outer Tracker," *JINST* **9** no. 01, (2014) P01002, arXiv:1311.3893 [physics.ins-det].

[122] **LHCb** Collaboration, R. Aaij *et al.*, "Measurement of the track reconstruction efficiency at LHCb," *JINST* **10** no. 02, (2015) P02007, arXiv:1408.1251 [hep-ex].

[123] O. Callot and S. Hansmann-Menzemer, "The Forward Tracking: Algorithm and Performance Studies," Tech. Rep. LHCb-2007-015. CERN-LHCb-2007-015, CERN, Geneva, May, 2007.

[124] M. Needham and J. Van Tilburg, "Performance of the track matching," Tech. Rep. LHCb-2007-020. CERN-LHCb-2007-020, CERN, Geneva, Mar, 2007.

[125] M. Needham, "Performance of the Track Matching," Tech. Rep. LHCb-2007-129. CERN-LHCb-2007-129, CERN, Geneva, Oct, 2007.

[126] R. Aaij *et al.*, "Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC," arXiv:1812.10790 [hep-ex].

[127] **LHCb RICH Group** Collaboration, M. Adinolfi *et al.*, "Performance of the LHCb RICH detector at the LHC," *Eur. Phys. J.* **C73** (2013) 2431, arXiv:1211.6759 [physics.ins-det].

[128] **LHCb collaboration** Collaboration, "LHCb calorimeters: Technical Design Report," 2000.

[129] **LHCb collaboration** Collaboration, "LHCb muon system: Technical Design Report," 2001.

[130] A. A. Alves, Jr. *et al.*, "Performance of the LHCb muon system," *JINST* **8** (2013) P02022, arXiv:1211.1346 [physics.ins-det].

[131] M. Williams, V. V. Gligorov, C. Thomas, H. Dijkstra, J. Nardulli, and P. Spradlin, "The HLT2 Topological Lines," Tech. Rep. LHCb-PUB-2011-002. CERN-LHCb-PUB-2011-002, CERN, Geneva, Jan, 2011.

[132] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput. Phys. Commun.* **178** (2008) 852–867, arXiv:0710.3820 [hep-ph].

[133] I. Belyaev *et al.*, "Handling of the generation of primary events in gauss, the LHCb simulation framework," *Journal of Physics: Conference Series* **331** no. 3, (Dec, 2011) 032047.

[134] D. J. Lange, "The evtgen particle decay simulation package," *Nucl. Instrum. Meth.* **A462** no. 1, (2001) 152 – 155.

[135] P. Golonka and Z. Was, "PHOTOS Monte Carlo: A Precision tool for QED corrections in Z and W decays," *Eur. Phys. J.* **C45** (2006) 97–107, arXiv:hep-ph/0506026 [hep-ph].

[136] **Geant4** Collaboration, S. Agostinelli *et al.*, "Geant4: a simulation toolkit," *Nucl. Instrum. Meth.* **A506** no. 3, (2003) 250 – 303.

[137] **Geant4** Collaboration, J. Allison *et al.*, "Geant4 developments and applications," *IEEE Transactions on Nuclear Science* **53** no. 1, (Feb, 2006) 270–278.

[138] M. Clemencic, G. Corti, S. Easo, C. R. Jones, S. Miglioranzi, M. Pappagallo, and P. R. and, "The LHCb simulation application, gauss: Design, evolution and experience," *Journal of Physics: Conference Series* **331** no. 3, (Dec, 2011) 032023.

[139] **LHCb collaboration** Collaboration, R. Aaij *et al.*, "Updated branching fraction measurements of $B^0_{(s)} \to K^0_S h^+ h'^-$ decays," *JHEP* **11** (2017) 027 LHCb-PAPER-2017-010, CERN-EP-2017-131, arXiv:1707.01665 [hep-ex].

[140] **LHCb collaboration** Collaboration, R. Aaij *et al.*, "Measurement of the fragmentation fraction ratio f_s/f_d and its dependence on B meson kinematics," *JHEP* **04** (2013) 001 CERN-PH-EP-2013-006, LHCb-PAPER-2012-037, arXiv:1301.5286 [hep-ex].

[141] **LHCb Collaboration** Collaboration, "Updated average f_s/f_d b-hadron production fraction ratio for 7 TeV pp collisions,".

[142] **LHCb collaboration** Collaboration, R. Aaij *et al.*, "Measurement of b-hadron fractions in 13 TeV pp collisions," arXiv:1902.06794 [hep-ex].

[143] A. Hoecker *et al.*, "TMVA 4 — Toolkit for Multivariate Data Analysis with ROOT. Users Guide.," `arXiv:physics/0703039 [physics]`.

[144] M. Pivk and F. R. Le Diberder, "sPlot: A statistical tool to unfold data distributions," *Nucl. Instrum. Meth.* **A555** (2005) 356–369, `arXiv:physics/0402083 [physics.data-an]`.

[145] S. Tolk, J. Albrecht, F. Dettori, and A. Pellegrino, "Data driven trigger efficiency determination at LHCb," Tech. Rep. LHCb-PUB-2014-039. CERN-LHCb-PUB-2014-039, CERN, Geneva, May, 2014.

[146] V. V. Gligorov, C. Thomas, and M. Williams, "The HLT inclusive B triggers," Tech. Rep. LHCb-PUB-2011-016. CERN-LHCb-PUB-2011-016. LHCb-INT-2011-030, CERN, Geneva, Sep, 2011. LHCb-INT-2011-030.

[147] V. V. Gligorov and M. Williams, "Efficient, reliable and fast high-level triggering using a bonsai boosted decision tree," *JINST* **8** (2013) P02013, `arXiv:1210.6861 [physics.ins-det]`.

[148] W. D. Hulsbergen, "Decay chain fitting with a Kalman filter," *Nucl. Instrum. Meth.* **A552** (2005) 566–575, `arXiv:physics/0503191 [physics]`.

[149] H. Voss, A. Hoecker, J. Stelzer, and F. Tegenfeldt, "TMVA - Toolkit for Multivariate Data Analysis with ROOT," *PoS* **ACAT** (2007) 040.

[150] **LHCb collaboration** Collaboration, R. Aaij *et al.*, "Measurement of the b-quark production cross-section in 7 and 13 TeV pp collisions," *Phys. Rev. Lett.* **118** (2017) 052002 LHCb-PAPER-2016-031, CERN-EP-2016-201, `arXiv:1612.05140 [hep-ex]`.

[151] R. Aaij *et al.*, "Selection and processing of calibration samples to measure the particle identification performance of the LHCb experiment in Run 2," *Eur. Phys. J. Tech. Instr.* **6** (2018) 1 LHCb-DP-2018-001, `arXiv:1803.00824 [hep-ex]`.

[152] A. Poluektov, "Kernel density estimation of a multidimensional efficiency profile," *JINST* **10** no. 02, (2015) P02011, `arXiv:1411.5528 [physics.data-an]`.

[153] T. Skwarnicki, *A study of the radiative CASCADE transitions between the Upsilon-Prime and Upsilon resonances*. PhD thesis, Cracow, INP, 1986.

[154] G. A. Cowan, D. C. Craik, and M. D. Needham, "RapidSim: an application for the fast simulation of heavy-quark hadron decays," *Comput. Phys. Commun.* **214** (2017) 239–246, arXiv:1612.07489 [hep-ex].

[155] **LHCb collaboration** Collaboration, R. Aaij *et al.*, "Amplitude analysis of the decay $\overline{B}^0 \to K^0_S \pi^+ \pi^-$ and first observation of *CP* asymmetry in $\overline{B}^0 \to K^*(892)^- \pi^+$," *Phys. Rev. Lett.* **120** (2018) 261801 LHCb-PAPER-2017-033, CERN-EP-2017-317, arXiv:1712.09320 [hep-ex].

[156] **LHCb collaboration** Collaboration, R. Aaij *et al.*, "Amplitude analysis of $B^0_s \to K^0_S K^\pm \pi^\mp$ decays," *JHEP* **06** (2019) 114 LHCb-PAPER-2018-045 CERN-EP-2019-017, arXiv:1902.07955 [hep-ex].

[157] **LHCb** Collaboration, R. Aaij *et al.*, "Amplitude analysis of $B^\pm \to \pi^\pm K^+ K^-$ decays," arXiv:1905.09244 [hep-ex].

[158] L. Henry, *Charmless hadronic three-body decays of neutral B mesons with a Kos in the final state in the LHCb experiment : branching fractions and an amplitude analysis*. PhD thesis, Paris U., VI-VII, 2016.

[159] **LHCb collaboration** Collaboration, "Framework TDR for the LHCb Upgrade: Technical Design Report," 2012.